Development of a yeast cell factory for production of aromatic secondary metabolites

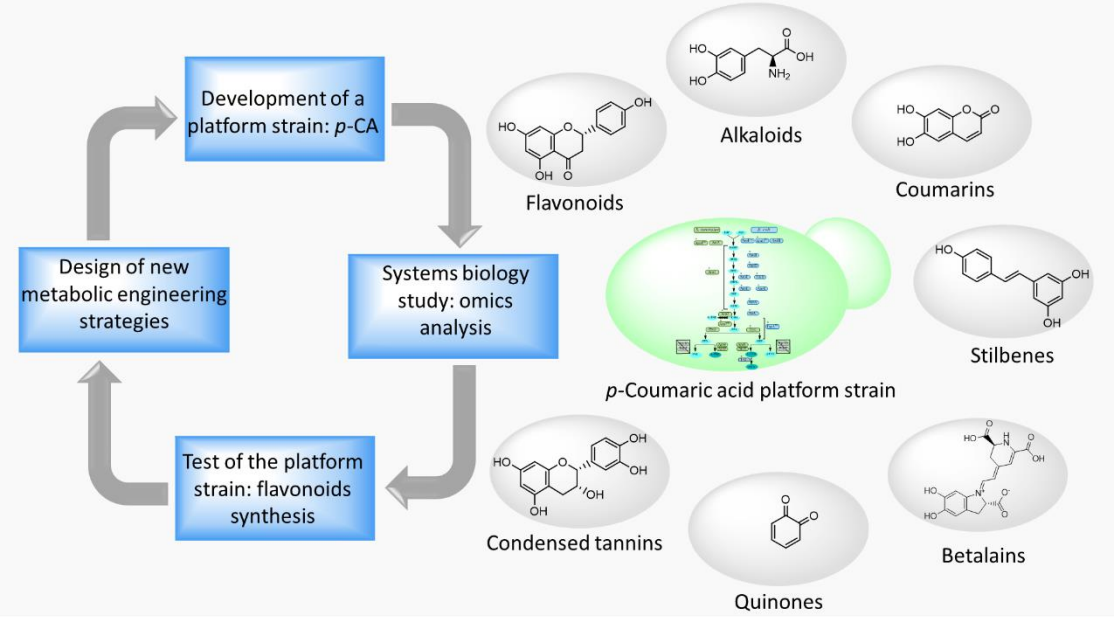
Rodriguez Prado, Edith Angelica

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Development of a yeast cell factory for production of aromatic secondary metabolites

PhD Thesis



Edith Angelica Rodriguez P.
July 2016

DTU Biosustain
The Novo Nordisk Foundation Center for Biosustainability

Development of a yeast cell factory for production of aromatic secondary metabolites

PhD Thesis by

Edith Angelica Rodriguez Prado

The Novo Nordisk Foundation Center for Biosustainability

Technical University of Denmark

Copenhagen, Denmark 2016

PhD Thesis

The Novo Nordisk Foundation Center for Biosustainability

Technical University of Denmark

Kemitorvet Bygning 220

2800 Kgs. Lyngby

Denmark

Cover illustration

Overview of the research approach adopted in this thesis and schematic illustration of secondary metabolites synthesized from aromatic amino acids.

Copenhagen, Denmark 2016

Preface

This thesis is written as a partial fulfillment of the requirements to obtain the PhD degree at the Technical University of Denmark, DTU. This thesis includes work carried out at The Novo Nordisk Foundation Center for Biosustainability and at the Department of Chemical and Biological Engineering, Chalmers University of Technology, from February 2013 until July of 2016 under the supervision of the Professor Jens Nielsen and Doctor Irina Borodina.

This PhD project was funded by The Novo Nordisk Foundation.

Angelica Rodriguez Prado

Copenhagen, July 2016.

Acknowledgements

I would like to thank my supervisor, Professor Jens Nielsen for giving me the opportunity to join his research group, your scientific career and expertise together with your kindness are an inspiration to me. Your positive attitude, advice and permanent support was essential for the accomplishment of each one of the research projects.

Also, I would like to thank my co-supervisor, Doctor Irina Borodina. It was a valuable experience to work with you. You are an excellent scientist and I have learned many things from you. Your professional feedback, support and advice contributed to my formation as a scientist.

I would like to thank Kanchana R. Kildegaard and Yun Chen for their guidance and support in the projects that I developed through the PhD program at CfB and at the Chalmers University of Technology. Both of you were great collaborators, your advice and support were very helpful.

I would like to thank my co-authors Tomas Strucko, Steen Stahlhut, Mette Kristensen, Mingji Li, Emre Ozdemir and Sakda Khoomrung for the scientific collaboration in the research projects; I learned many things on each project and all of you contributed to that.

I would like to thank Tim Snoek, Carina Holkenbrink, Tadas Jakociunas, Arun Rajkumar, Laura Dato and Jie Zhang for reading parts of this book. Your suggestions and comments were very helpful, thank you very much to all of you.

I would like to thank Mingji Li and Gheorghe Borja, both of you are great friends, thank you for your company and help. I would like to thank Carina Holkenbrig, Vratislav Stovicek, Marie I. Dam and all the members of the yeast community at CfB or your kindness and help in the laboratory.

Also, I would like to thank the persons from CfB and the Chalmers University of Technology who helped with the practical matters of the research projects: Anna Koza, Leif Väremo, Eduard Kerkhoven, Mark Bisschops, Konstantin Schneider, Marie Nordqvist, Ximena Rozo, Hanne

Christensen, Lars Petersen, Julia Karlsson, Svetlana Sherstyuk, Simo Jacobsen and Dushica Arsovska.

Finally, I would like to thank my grandmother and mother for your example and unconditional support through all these years. I would like to thank my boyfriend for his support especially in the last months of the PhD, also I would like to thank my sisters and brother for always being there for me.

Abstract

Aromatic secondary metabolites are compounds mainly synthesized by plants and fungi as a response to predators and environmental stresses. These compounds have a broad range of natural properties such as reduction of oxidative damage in cells, antibacterial effects and UV protection. Many of these properties can be useful for the treatment of different diseases and development of pharmaceutical products.

The low abundance of these compounds in natural sources together with technical challenges for the extraction of these compounds from plants, open up the possibility for synthesizing aromatic secondary metabolites in cell factories. In this research project, we developed a yeast platform strain for the production of *p*-coumaric acid an intermediate compound for the synthesis of aromatic secondary metabolites. Subsequently, we performed a systems biology analysis of the strain and finally we developed an array of yeast strains expressing flavonoid metabolic pathways containing up to ten heterologous genes.

The platform strain was capable of producing 1.93 ± 0.26 g L^{-1} of *p*-coumaric acid in fed-batch fermentation, which is the highest titer that has been reported for a yeast cell factory so far.

The systems biology analysis of the platform strain suggests that the strain has transcriptional downregulations in genes involved in the transport of amino acids and sugars, which could be a response to the stress triggered by the production of *p*-coumaric acid.

The platform strain was capable of synthesizing six different types of flavonoids, and some of the engineered strains produced significant titers of flavonoid compounds such as kaempferol and quercetin. Moreover, for the first time, we synthesized the flavonoids liquiritigenin, resokaempferol and fisetin in yeast.

Resume (Danish abstract)

Aromatiske sekundære metabolitter er forbindelser, der hovedsageligt syntetiseres af planter og svampe som en reaktion på prædation og miljømæssige stressorer. Disse forbindelser har en bred vifte af naturlige egenskaber såsom reduktion af oxidativt stress i celler, antibakterielle virkninger og UV-beskyttelse. Mange af disse egenskaber kan være nyttige til behandling af forskellige sygdomme og udvikling af farmaceutiske produkter.

Den begrænsede mængde af disse forbindelser i naturlige kilder kombineret med de tekniske udfordringer for udvinding af disse forbindelser fra planter, åbner for muligheden for at syntetisere aromatiske sekundære metabolitter i cellefabrikker. I dette forskningsprojekt udviklede vi en gær platform-stamme til fremstilling af *p*-coumarinsyre, et mellemprodukt til syntese af aromatiske sekundære metabolitter. Efterfølgende har vi udført en systembiologisk analyse af stammen og endelig har vi udviklet en vifte af gærstammer, der udtrykker flavonoid stofskifteveje, der indeholder op til ti heterologe gener.

Platform stammen var i stand til at producere 1.93 ± 0.26 g L^{-1} af *p*-coumarinsyre i fed-batch fermentering, hvilket er den højeste titer, der er blev rapporteret for en gærcelle fabrik hidtil.

Den systembiologiske analyse af platformstammen tyder på, at stammen har transkriptionelle nedreguleringer i gener involveret i transporten af aminosyrer og sukkerarter, som kunne være en reaktion på stress udløst af produktionen af *p*-coumarinsyre.

Platformstammen var i stand til at syntetisere seks forskellige typer flavonoider, og nogle af de genmodificerede stammer producerede betydelige koncentrationer af flavonoidforbindelser, som kaempferol og quercetin. Desuden syntetiserede vi for første gang flavonoiderne liquiritigenin, resokaempferol og fisetin i gær.

List of publications

Included in this thesis

1. Establishment of a yeast platform strain for production of *p*-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis.
   Rodriguez, A., Kildegaard, K.R., Li, M., Borodina, I., Nielsen, J., 2015. Metab. Eng. 31, 181-188. doi:10.1016/j.ymben.2015.08.003.

2. Metabolic response of *Saccharomyces cerevisiae* to the over-production of *p*–coumaric acid.
   Rodriguez, A., Chen, Y., Khoomrung, S., Özdemir, E., Borodina, I., Nielsen, J. 2016.

3. Metabolic engineering of yeast for fermentative production of flavonoids.
   Rodriguez, A., Strucko, T., Stahlhut, S.G., Kristensen, M., Svenssen, D.K., Forster, J., Nielsen. J., Borodina, I. 2016.

Additional work during the PhD study but not included in this thesis

4. Microbial production of the flavonoids garbanzol, resokaempferol and fisetin.
   Stahlhut, S.G., Siedler, S., Neves, A.R., Maury, J., Forster, J., Gaspar, P., Borodina, I., **Rodriguez, A**., Strucko, T. 2016. Patent WO 2016071505 (A1).

5. De novo production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*.
   Li, M., Kildegaard, K.R., Chen, Y., **Rodriguez, A**., Borodina, I., Nielsen, Metabolic Engineering. Metab. Eng. 32, 1–11. doi:10.1016/j. ymben. 2015.08.007.

6. CasEMBLR: Cas9-facilitated multi-loci genomic integration of in vivo assembled DNA parts in *Saccharomyces cerevisiae*.
   Jakociunas, T., Rajkumar, A.S., Zhang, J., Arsovska, D., **Rodriguez, A**., Jendresen, C.B., Skjoedt, M.L., Nielsen, A.T., Borodina, I., Jensen, M.K., Keasling, J.D., 2015. CasEMBLR: ACS Synth. Biol. 150317152158005. doi:10.1021/acssynbio.5b00007.

Contents

Preface ... **3**
Acknowledgements ... **4**
Abstract ... **6**
Resume (Danish abstract) ... **7**
List of publications ... **8**
Contents ... **9**
Overview ... **10**
1. Introduction ... **12**
1.1 Aromatic amino acids biosynthesis in yeast ... 12
1.2 Cell factory development ... 20
1.3 Aromatic secondary metabolites ... 25
1.4 Systems biology and development of cell factories ... 31
1.5 Strategies for strain engineering ... 35
References ... 40
2. Development of a platform strain for synthesis of *p*-coumaric acid ... **54**
Abstract ... 54
Introduction ... 54
Methods ... 55
Results ... 58
Discussion ... 60
3. Systems biology analysis of the *p*-coumaric acid platform strain ... **74**
Abstract ... 74
Introduction ... 75
Methods ... 77
Results ... 83
Discussion ... 90
4. Synthesis of flavonoids in yeast ... **101**
Abstract ... 101
Introduction ... 102
Methods ... 104
Results ... 111
Discussion ... 119
5. Conclusions and perspectives ... **131**

Overview

Microbial cell factories are applied in industrial biotechnology for production of biofuels, chemicals, pharmaceutical and cosmetic ingredients, etc. There is a great potential in developing novel cell factories, capable of making new classes of chemical compounds and utilizing a broad range of renewable carbon feedstocks, to replace some of the traditional petrochemical routes or extraction from natural sources.

In the last decade, many *S. cerevisiae* cell factories have been developed for the production of various chemicals. The development of genome editing and systems biology techniques have especially contributed to the advance in the development of cell factories. However, further work is required to decrease the time frame for the development of a cell factory and reach competitive prices for scaling the production of these compounds to industrial levels.

Systems biology studies are crucial for the development of cell factories. In other disciplines such as oncology and drug discovery, systems biology approaches have contributed to a better understanding and development of improved therapeutic products.

The aim of this thesis was the development of a yeast cell factory for the production of aromatic secondary metabolites. Beyond the development of a cell factory for producing a specific compound, this thesis is part of a general effort by the Novo Nordisk Center for Biosustainability for the development of knowledge and techniques to create cell factories for the production of chemicals and protein-based compounds.

In **Chapter 1**, I review the state-of-the-art in the topic of my thesis. **Chapter 2** describes my published study on rational metabolic engineering of *S. cerevisiae* platform strain with optimized production of *p*-coumaric acid (Rodriguez et al., Metab Eng. 31, 181-188). In **Chapter 3**, I studied the *p*-coumaric acid cell factory through systems biology to gain more knowledge about the metabolic and physiological changes that yeast cells undergo when they are producing aromatic secondary metabolites. Finally, I applied

this platform strain for production of various flavonoids (**Chapter 4**). **Chapter 5** contains the conclusions and perspectives based on the results obtained through the research project.

1. Introduction

1.1 Aromatic amino acids biosynthesis in yeast

Aromatic amino acids (AAAs) are fundamental for the synthesis of proteins in all cells. In plants, they are also the precursors of compounds involved in defense, growth and development. Since the production of AAAs requires high amounts of energy, this biosynthetic pathway is tightly regulated so the cells produce just the necessary quantities of these amino acids.

In *S. cerevisiae*, specific research about the AAAs pathway is limited, a complete characterization of the AAAs pathway was made by Braus (1991). Since this publication, there are some other studies related to specific enzymes of the pathway or applications for the synthesis of aromatic secondary metabolites (Duncan et al., 1988; Kunzler et al., 1992; Heimstaedt et al., 2005).

An appropriate understanding of the AAAs biosynthesis in yeast and other organisms may help to design strategies for the development of cell factories. In this section, I will give an overview of the AAAs biosynthetic pathway in *S. cerevisiae*, its regulations and I will highlight its difference compared to the AAAs pathway in *E. coli* and plants.

Aromatic amino acids pathway overview

AAAs are synthesized trough the shikimate pathway. In microorganisms, AAAs are used exclusively for protein synthesis, whereas in plants AAAs are the precursors of a broad range of aromatic secondary metabolites. Since plants need an almost permanent synthesis of some aromatic secondary metabolites, the differences in the shikimate pathway between yeast and plants are relevant to the development of cell factories.

Although bacteria, fungi and plants have similar AAAs biosynthetic pathways, there are some fundamental differences related to the pathway

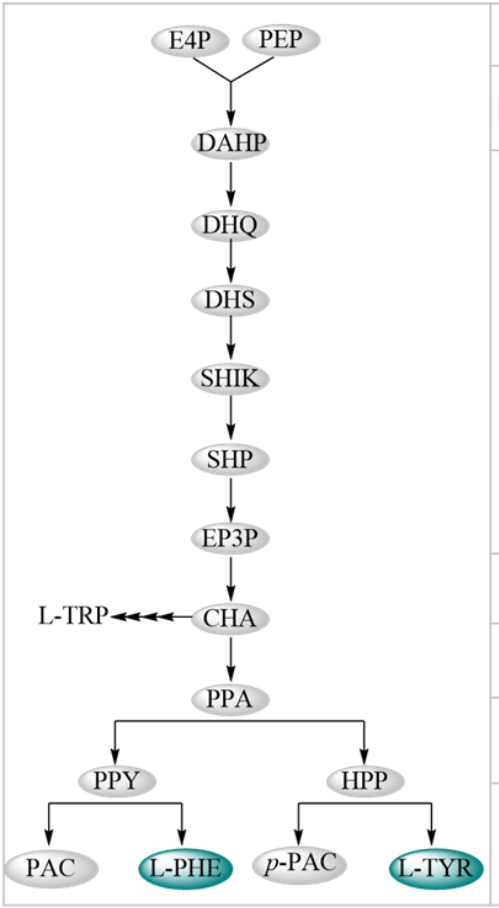
regulation and enzymes fusion arrangements (Richards et al., 2006; Hermann et al., 1999).

In *S. cerevisiae* the AAAs, L-tyrosine, L-phenylalanine and L-tryptophan are synthesized from erytrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP). The pathway has 7 common enzymatic steps for the three amino acids, after that the pathway is branched to L-tryptophan, L-tyrosine and L-phenylalanine; interestingly L-phenylalanine and L-tyrosine share the last enzymatic step catalyzed by the aminotransferases ARO8 and ARO9 (Figure 1).

The aromatic amino acids pathway is coupled to the central metabolism by the enzymes E4H and PEP. One of the more remarkable differences between the AAAs pathway in microorganisms and plants is the post-translational regulation: in plants it has not been reported allosteric regulation for the first enzymes of the pathway, whereas in microorganisms the allosteric feedback inhibition on DAHP synthase and chorismate mutase is well known.

DAHP synthases ARO4 and ARO3 are responsible for the aldol condensation of E4P and PEP for the production of 3-Deoxy-D-arabino-heptulosonate7-phosphate (DAHP). In yeast these enzymes are subject to allosteric regulation by L-tyrosine (ARO4) and L-phenylalanine (ARO3), in plants and other microorganisms DAHP synthases are not feedback regulated and are activated by L-tryptophan (Silakowski et al., 2000; Maeda and Dudareva, 2012).

In *S. cerevisiae*, the penta functional enzyme ARO1 covers five of the seven common enzymatic steps from the shikimate pathway, whereas in *E. coli* we have monofunctional enzymes and in plants, the second and third steps in the reaction are synthesized by a bifunctional enzyme (Maeda and Dudareva et al., 2012).



	S. cerevisiae	*E. coli*	*A. thaliana*
	Aro3 Aro4	AroF AroG AroH	DAHPS1 DAHPS2
	Aro1	AroB	DHQS
		AroD	DHD-SDH
		AroE YdiB	
		AroL AroK	SK1 SK2
		AroA	EPSP
	Aro2	AroC	CS1 CS2 CS3
	Aro7	PheA[1] TyrA[2]	CM1 CM2 CM3
	Pha2[1] Tyr1[2]		ADT-PDT1 2 3 4 5 6[1]
			ADH-PDH1 2[2] PPA-AT[1,2]
	Aro8 Aro9	TyrB AspC IlvB	PPY-AT HPP-AT

Figure 1. Schematic representation of the aromatic amino acids pathway in *S.cerevisiae*, *E. coli* and *A. thaliana*. E4P: erythrose 4-phosphate, PEP: phosphoenolpyruvate, DAHP: 3-deoxy-D-arabino-heptulosonic acid 7-phosphate, DHQ: 3-dehydroquinate, DHS: 3-dehydro-shikimate, SHIK: shikimate, SHP: shikimate-3-phosphate, EP3P: 5-enolpyruvylshikimate-3-phosphate, PPA: prephenate, PPY: phenylpyruvate, HPP: para-hydroxy-phenylpyruvate, PAC: phenylacetaldehyde, pPAC: para-hydroxy-acetaldehyde, L-PHE: L-phenylalanine, L-TYR: L-tyrosine, The superscript 1 indicates enzymes involved in the synthesis of L-phenylalanine and the superscript 2 indicates enzymes involved in the synthesis of L-tyrosine.

In *S. cerevisiae*, at this enzymatic step DAHP is converted in 3-dehydroquinate by five chemical reactions (alcohol oxidation, β-elimination of inorganic phosphate, carbonyl reduction, ring opening, and intramolecular aldol condensation). After that, 3-dehydroquinate is dehydrated to 3-dehydroshikimate, and reduced to shikimate. Subsequently, shikimate is phosphorylated to shikimate 3-phosphate and finally the enol-pyruvyl moiety from PEP is transferred to shikimate 3-phosphate to get the final compound synthesized by the pentafunctional enzyme: 5-enolpyruvylshikimate 3-phosphate (EPSP).

From the metabolic point of view, it is considered that a pentafunctional enzyme has advantages related to the channelling of intermediate compounds and equalized regulation of the enzymatic activity (Zhang, 2011; Proschell et al., 2015). Considering that AAAs synthesis is highly regulated in *S. cerevisiae*, it may be an advantage for the cell but not for the development of a cell factory as the synthesis of the intermediate compounds may be controlled to produce just the necessary amounts of AAAs.

The chorismate synthase ARO2, catalyze the last common enzymatic step between L-tryptophan, L-tyrosine and L-phenylalanine, at this point chorismate is synthesized by the elimination of the 3-phosphate and C6-*pro*-R hydrogen from EPSP. The cofactor for this reaction is flavin mononucleotide.

The protein sequence of chorismate synthase has been conserved in many organisms. There are two types of chorismate synthases according to their capacity to regenerate the reduced form of flavin mononucleotide. In plants and some bacteria, the reduction of flavonoid mononucleotides is performed by an external system whereas yeast has a bifunctional chorismate synthase with an additional oxidoreductase activity (Quevillon-Cheruel et al., 2004).

ARO7 chorismate mutase performs the last common enzymatic step between L-tyrosine and L-phenylalanine, at this enzymatic step chorismate is converted in prephenate through a pericyclic Claisen rearrangement. Chorismate mutases from *S. cerevisiae* and *E. coli* have significant differences: yeast has a monofunctional enzyme, whereas *E. coli* has a bifunctional enzyme combining the functions of chorismate mutase-prephenate dehydrogenase (Synthesis of L-tyrosine) or chorismate mutase-prephenate hydrolase (of L-phenylalanine).

Chorismate synthase is subject to allosteric regulation; its activity is stimulated by L-tryptophan and inhibited by L-tyrosine, this regulation allows the cell to control the flux of chorismic acid either to L-tryptophan biosynthesis or the synthesis of L-phenylalanine and L-tyrosine. The substitution of threonine by isoleucine at the position 226 in the C-terminal part of the protein increases the activity of the enzyme and removes the inhibitory response to L-tyrosine (Schmidheini et al., 1989).

PHA2 prephenate dehydratase catalyzes the conversion of prephenate to phenylpyruvate; yeast has a monofunctional enzyme whereas *E. coli* has a bifunctional enzyme as was mentioned in the previous paragraph. The protein sequence of prephenate dehydratase from yeast and arogenate dehydratase, its equivalent in plants have significant differences, despite the differences, arogenate dehydratases from *Arabidopsis thaliana* ADT1 and ADT2 can use prephenate for the synthesize phenylpyruvate (Bross et al., 2011).

TYR1 prephenate dehydrogenase catalyzes the oxidative carboxylation and dehydration of prephenate to *p*-hydroxyphenylpyruvate. This enzyme is transcriptionally regulated by amino acids, mostly L-phenylalanine. Mannhaupt et al. (1989) found that *TYR1* expression is affected by aromatic

amino acids: medium containing L-tyrosine or L-tryptophan decreased the expression of TYR1 in 20% and the combination of L-tyrosine or L.-tryptophan with L-phenylalanine dropped the expression levels arround 50%.

ARO8 and ARO9 aminotransferases I and II are involved in the transamination of phenylpyruvate and *p*-hydroxyphenylpyruvate into L-phenylalanine and L-tyrosine. ARO8 is mainly involved in L-tyrosine and L-phenylalanine biosynthesis whereas ARO9 takes part in the degradation of L-tryptophan. Interestingly in strains with *ARO8* deletion *ARO9* performs the biosynthetic functions of *ARO8*. The broad range of amino donors substrate of ARO8 and ARO9 suggest that the aminotransferases may also be involved in the metabolism of other amino acids such as leucine, lysine, and methionine (Urrestarazu et al., 1998).

Regulation of aromatic amino acids biosynthesis

The synthesis of aromatic amino acids is regulated at the transcriptional and post-translational level. Since the genes involved in the synthesis of aromatic amino acids are located on different chromosomes, the transcription is initiated by different promoters; also *S. cerevisiae* has a high level of transcription of genes involved in the synthesis of AAAs, even when the intracellular concentration of AAAs is high. Another important characteristic is that the synthesis of AAAs is part of an intricate regulatory network, interestingly when the cells face a deficit of a specific amino acid, the transcription of genes involved in the synthesis of other amino acids is triggered (Braus., 1991).

Amino acids biosynthetic promoters can be regulated by the GCN4 dependent system and the basal system. The GCN4 dependent system is activated when the cells are facing amino acids deficiencies, whereas the basal system is not activated by amino acids deficiencies and is in charge of the high basal level of transcription of genes involved in amino acids biosynthesis (Braus., 1991).

GCN4 is a transcriptional activator of the biosynthesis of more than 30 amino acids and its expression is regulated at the transcriptional and

translational level. The translation of GCN4 is induced when cells are deprived of amino acids triggering a transcriptional induction of genes involved in the synthesis of enzymes. The regulatory response of GCN4 enables cells to limit their consumption of amino acids in environments with a low concentration of nutrients. Also, the response of GCN4 is caused by the amino acids starvation occasioned by the feedback inhibition in medium with imbalanced concentrations of amino acids. (Braus., 1991; Hinnebusch., 2005).

The post-transcriptional regulation of AAAs biosynthesis is controlled at two enzymatic steps: at the beginning of the pathway where E4P and PEP are condensed to DAHP by DAHP synthase and at the eighth enzymatic step where chorismate is converted to prephenate. Subsequently prephenate can be directed either to the biosynthesis of L-tryptophan or the biosynthesis of L-tyrosine and L-phenylalanine. High concentrations of L-phenylalanine regulate the DAHP synthase ARO3. ARO4 is regulated by L-tyrosine and chorismate mutase ARO7 is inhibited by L-tyrosine.

The identification and subsequent elimination of the feedback regulation of DAHP synthase and chorismate mutase have contributed to the improvement of production of aromatic secondary metabolites (Luttik et al., 2008; Koopman et al., 2012). Besides to the allosteric regulation of some enzymes of the pathway, many other factors can affect the activity of the enzymes such as pH, substrate concentrations and inhibitors (Sauro et al., 2011).

In plants, the flux of aromatic amino acids to aromatic secondary metabolites is higher since different metabolites are necessary depending on the growth stage and stress conditions of the organism. The knowledge of the mechanisms and genes that are upregulated or downregulated under specific stress conditions is relevant for the development of microbial cell factories; this information can be used for the development of metabolic engineering strategies. Plants are present in different habitats and subsequently facing a different type of stress, their patterns of synthesis of aromatic secondary metabolites can be a valuable source of information for the development of yeast cell factories. Tzin et al. (2010) give some

examples of transcription factors whose high expression trigger elevated expression of genes from the shikimate pathway and aromatic amino acids metabolism. For example, DAHP synthase from *Solanum tuberose* is strongly expressed under pathogen stress conditions.

1.2 Cell factory development

The development of a cell factory is a process with a time frame of around 3-5 years and investment of hundreds of millions of dollars (Becker et al., 2012; Nielsen et al., 2014). It is expected that in the following years the cost and the time required for the development of a cell factory can decrease due to the development of new technologies. Also, the information that has been obtained in the last years allow us to have a better understanding of the biological process that regulates the synthesis of the production of a specific compound.

At the moment the best way for developing cell factories is through the use of the traditional platform strains such as *S. cerevisiae* and *E. coli*. However the steady development of technologies and knowledge may change this situation and in the future, we may have more flexibility and tools for the development of cell factories in other microorganisms. This section is a review of the cell factories that have been developed in the last years for the synthesis of aromatic secondary metabolites in *S. cerevisiae*, *E. coli* and other organisms.

S. cerevisiae and *E. coli* have been used as cell factories for diverse compounds. In comparison with other organisms extensive research in areas such as physiology and systems biology originated a broad knowledge of the organisms and their biological characteristics. The information gathered in the last years and the development of new techniques allowed the development and implementation of an extensive range of genome engineering tools in comparison to other organisms where many of their biological characteristics are still unknown making the development of tools for genome engineering more complex.

The production of aromatic secondary metabolites has been explored in different organisms. By the use of different techniques, such as codon optimization, protein engineering and enzyme tuning strategies it has been possible to produce diverse compounds such as flavonoids, stilbenes and alkaloids (Chemler et al., 2006; Hawkins et al., 2008; Li et al., 2015). In this

section, we review some of the cell factories developed for the synthesis of aromatic secondary metabolites in the last years.

Escherichia coli

E. coli is a widely used cell factory. The use of this organism has many advantages such as tolerance to organic acids, ability to metabolize 5 and 6 carbon sugars, fast growth and the availability of a broad range of established genome editing techniques. As *E. coli* has been a key model bacterium, a considerable amount of research on genetics, metabolism and physiology of *E. coli* has been performed. (Clomburg et al., 2010; Yu et al., 2011).

E. coli has been used for the synthesis of different aromatic secondary metabolites, however, many of these strains have to be feed with intermediate compounds such as L-phenylalanine or L-tyrosine. The feeding strategy is useful for testing heterologous pathways, however for the development of cell factories with potential industrial use, it is necessary to engineer organisms for the synthesis of secondary metabolites using glucose or another not expensive carbon source.

On this respect, Santos et al. (2011) developed a strain for overproduction of L-tyrosine, the pathway was optimized through the balance of gene expression, improvement of malonyl-CoA supply and identification of optimal variants of enzymes and promoters. Ssubsequently they assembled a heterologous pathway for naringenin synthesis, this strain was able to produce 29 mg L^{-1} of naringenin from glucose in minimal medium.

Leonard et al. (2006) developed a platform strain for production of flavonoids. Initially, they fused a P450 flavonoid 3′, 5′-hydroxylase with a P450 reductase; then they combined the fusion protein together with a biosynthetic pathway for production of kaempferol and quercetin in an *E. coli* strain. The engineered strain was able to produce 140 and 20 µg L^{-1} in cultures with minimal medium supplemented with *p*-coumaric acid.

Further work to improve the synthesis of flavonoids has been performed in *E. coli,* by improving the supply of malonyl-CoA and UDP-glucose through re-engineering some pathways from the central metabolism (Leonard et at., 2007; Leonard et al., 2008). Interestingly, they found that carbon channeling towards fatty acids is a competitive step for the synthesis of flavonoids and malonyl-CoA should be adjusted to the optimal level for the cells, since high amounts of malonyl-CoA are negative for the synthesis of flavonoids. The engineered strains obtained in the previous research produced 480 mg L^{-1} of pinocembrin, 155 mg L^{-1} of naringenin and 50 mg L^{-1} eridoctyol.

Recently different secondary metabolites have been synthesized by the use of co-cultures, this type of culture is particularly useful for the synthesis of molecules with complex biosynthetic pathways. By the use of co-cultures, it is possible to decrease the metabolic burden caused by the expression of long heterologous pathways; by the adjustment of the ratios of the strains in the culture it is possible to reach an optimal synthesis of the target compound and engineer the strain for production of specific substrates and co-factors (Zhou et al., 2014, Zhang et al., 2016).

Co-cultures have been used for the synthesis of flavonoids by Jones et al. (2016); they co-cultured a strain expressing the upstream part of the pathway (Malonyl-CoA dependent) and another strain expressing the downstream part (NADPH-dependent), they managed to get an improvement of 970-fold in the production of flavon-3-ols.

Yeast

Yeast is an ideal host strain for production of flavonoids. Due to its eukaryotic nature, the expression of plants heterologous pathways is less complicated in comparison to other organisms; for example prokaryotes cannot perform post-translational modifications and incorrect protein folding and membrane translation can be a drawback. (Chemler et al., 2008; Koopman et al., 2012). Spite of the advantages of yeast as a cell factory, around 85% of the flavonoids cell factories have been developed in *E. coli* (Pandey et al., 2016).

Jiang et al. (2015) conducted one of the first studies on production of flavonoids in *S. cerevisiae*. They assembled heterologous genes from different plant species for the synthesis of naringenin and pinocembrin, reaching 7 mg L^{-1} and 0.8 mg L^{-1} respectively. Also, they found that the pool of L-tyrosine was limiting the synthesis of *p*-coumaric acid, the precursor of naringenin.

Other flavonoids such as chrysin, apigenin and luteolin have been synthesized in *S. cerevisiae* using *p*-coumaric acid as a precursor. Leonard et al. (2005) tested the expression of a soluble flavonoid synthase (FSI) and a membrane-bound flavone synthase (FSII). They found that the expression of FSI had a positive effect on the production of flavonoids on this strain. Recently other approaches have been used for the synthesis of flavonoids. For example, flavonoids pathways for the synthesis of naringenin and kaempferol have been assembled in an artificial chromosome and expressed in *S. cerevisiae* (Naesby et al., 2009).

Koopman et al. (2012) engineered yeast strain for production of naringenin. They targeted the feedback regulation of the shikimate pathway enzymes DAHP synthase and chorismate mutase (ARO4 and ARO7). Also, they eliminated the synthesis of by-products by the elimination of phenylpyruvate decarboxylase. Finally, they optimized the heterologous pathway for naringenin synthesis by improving the gene expression in important heterologous genes; they produced 109 mg L^{-1} in aerobic batch fermentations.

Other cell factories

Streptomyces has been used as a cell factory for different secondary metabolites. The use of this organism as a cell factory has many advantages such as rapid growth, availability of many genome editing tools. Also, *Streptomyces* is a natural producer of many secondary metabolites, guarantying the supply of secondary metabolites precursors (Kim *et al*., 2015). *Streptomyces venezuelae* have been used as host strain for the production of flavonoids; the strain was optimized for the synthesis of

flavonoids by the assembly of a heterologous malonate assimilation pathway from *Streptomyces coelicolor.* This strain expressed heterologous genes involved in the synthesis of flavones and flavonones and it was possible to synthesize naringenin, pinocembrin apigenin and chrysin (Park et al., 2011).

The bacteria *Pseudomonas putida* is an attractive organism for the development of cell factories. Some of the advantages are its low nutritional needs and tolerance to extreme environmental conditions such as high temperature, drastic changes in the pH and high tolerance to toxins in the culture (Poblete-Castro et al., 2012). *P. putida* has also been engineered for production of the flavonoids precursor *p*-coumaric acid. Calero et al. (2016) developed a set of vectors for this species and increased the L-tyrosine availability in the strain, leading to the production of 197 mg L^{-1} of *p*-coumaric acid.

For many years it was thought that flavonoids synthesis was an evolutionary characteristic of terrestrial plants and that synthesis of these compounds allowed them to adapt to the UV radiation. Recent research demonstrated that many microalgae are natural synthesizers of flavonoids (Kovacik et al., 2010; Goiris et al., 2014). Since microalgae can produce these compounds, the opportunity of using these organisms for the synthesis of flavonoids can be explored.

Natural synthesis of flavonoids synthesis has been reported in the photosynthetic evolutionary lineages *Ochrophyta*, *Haptophyta*, *Rhodophyta*, *Chlorophyta* and *Cyanophyta.* From these lineages, *Cyanophyta* and *Chlorophyta* have been already explored as platforms for the production of biofuels (Larkum et al., 2011; Savakis et al., 2015). The disadvantage of using of microalgae or cyanobacteria for the synthesis of flavonoids is that only a few genome editing tools are developed and little information of their genotype is available. The development of some genome editing tools and systems biology studies in recent years, make it feasible to use microalgae as platforms for the production of flavonoids in the future (Daboussi et al., 2013; Gimpel et al., 2013).

1.3 Aromatic secondary metabolites

Plants produce a diverse range of secondary metabolites. Some of them are phenolic compounds that originate in the shikimate pathway or the acetate-mevalonate pathway. The phenolic secondary metabolites derived from the shikimate pathway cover a broad range of compounds such as flavonoids, alkaloids and coumarins (Figure 2).

The production of phenolic secondary metabolites has played a key role in the adaptation of higher plants to different environments, and their pharmaceutical properties have been the subject of recent research with promising results (Woelfle *et al.,* 2010; Khoo *et al.,* 2010; Bulzomi *et al.,* 2009).

Terrestrial and aquatic plants have followed different evolutionary tracks and adapted to different types of environments, and secondary metabolites have played a key role in this process (Buchanan et al., 2000). This section explores the chemistry of flavonoids, their natural properties and the potential use of these compounds in the industry.

Chemistry

Phenolic secondary metabolites derived from the shikimate pathway are characterized by the presence of at least one aromatic ring with one or more hydroxyl groups attached. Phenolic secondary metabolites cover compounds with simple structures with low molecular weight and compounds with multiple aromatic rings and more complex structures. These compounds can be classified according to the number and disposition of their carbon atoms, also some compounds conjugated to sugars and organic acids (Crozier et al., 2007).

The most studied secondary metabolites derived from aromatic amino acids are alkaloids and phenylpropanoids. Alkaloids can be synthesized from L-phenylalanine, L-tyrosine and other amino acids; these compounds are characterized by the presence of one or more nitrogen atoms in their structures. As an example, benzylisoquinoline alkaloids are synthesized

from L-tyrosine through series of decarboxylations, ortho-hydroxilations and deaminations giving as a result, the synthesis of the intermediates dopamine and 4-hydroxyphenylacetaldehyde. Subsequently, these compounds are condensed into (S)-norcoclaurine, the precursor of a broad range of benzylisoquinoline alkaloids (Crozier et al., 2007).

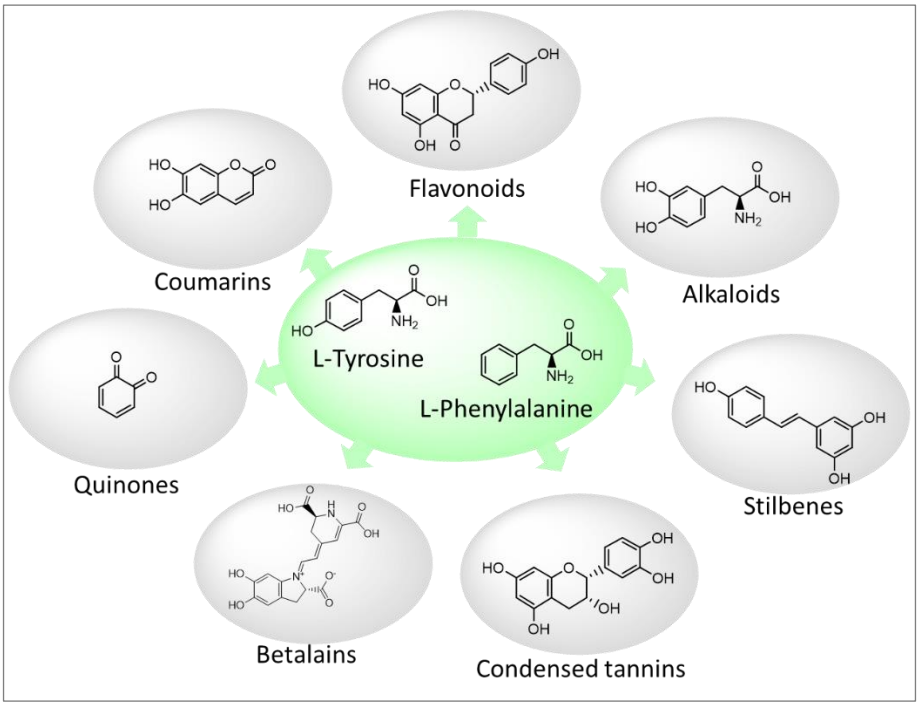


Figure 2. Examples of the aromatic secondary metabolites synthesized from L-phenylalanine and L-tyrosine: flavonoids, alkaloids, stilbenes, condensed tannins, betalains, quinones and coumarins.

Phenolic compounds cover a broad range of secondary metabolites such as flavonoids, lignins, coumarins and stilbenes. Aromatic secondary metabolites are mainly synthesized from L-phenylalanine. Flavonoids have in common the first biosynthetic steps catalyzed by chalcone synthase and chalcone isomerase, resulting in the formation of 2S-flavanones, naringenin and liquiritigenin. From this enzymatic step different reactions are catalyzed

giving origin to various subgroups of flavonoids such as isoflavonoids, flavones and flavonols.

Lignins are synthesized from coniferyl alcohol and *p*-coumaryl alcohol and the synthesis of coumarins requires hydroxylation, glycolysis and cyclization of cinnamic acid obtaining as a product *p*-coumaric acid that can be further hydroxylated and give origin to other coumarins.

Natural properties

Secondary metabolites have a fundamental role in plants evolution since the synthesis of these compounds allowed plants to conquer different environments and survived diverse types of stresses. Interestingly, some secondary metabolites had been found in unrelated plant families; this can be linked either to convergent evolution or differential expression of genes (Wink et al., 2014).

One of the best-known functions of aromatic secondary metabolites is their role in the multicomponent defense mechanism in plants; this mechanism includes signaling molecules, constitutive defenses and phytoalexins (Naoumkina et al., 2010). A wide range of studies have shown that secondary metabolites such as coumarins, flavonoids and lignings, can inhibit the growth of pathogens. Also, it was found that plants with higher concentrations of flavonoids were less susceptible to pathogens attacks (Shimizu et al., 2005; Bhuiyan et al., 2009; Ortuno et al., 2011).

Ultraviolet (UV) radiation may generate damage in DNA and accumulation of reactive oxygen species. Many studies have reported an increment in the accumulation of flavonoids such as quercetin and kaempferol in tissues under different grades of UV radiation (Winkel et al., 2002; Warren et al., 2003). A different study showed that plants with non-functional chalcone synthase (CHS) or chalcone isomerase (CHI) were less tolerant to UV radiation (Landry et al., 1995).

Secondary metabolites are also essential for the establishment of symbiosis with other organisms. Secretion of various compounds, including secondary

metabolites, by vascular plants contributes to the modification of the properties of the soil that surround the roots. Secondary metabolites act as chemotactic agents to establish mutualistic relations with specific microorganisms (Cheynier et al., 2013; Bertin et al., 2003). For example, *Medicago sativa* has been shown to excrete some flavonoids that induce transcription of genes related to fixation of nitrogen in its symbiotic species (Hartwig et al., 1990).

Pigments play a fundamental role in the attraction of pollinators and propagation of seeds. Aromatic secondary metabolites are precursors for the synthesis of some of these pigments. The most common pigments derived from aromatic amino acids are anthocyanins and betalains. Anthocyanins are flavonoids responsible for orange to blue pigmentation in different tissues of plants, and their color greatly depends on the number of hydroxyl groups in the second ring. Betalains are synthesized from L-tyrosine, and they are linked to colorations from yellow to red. Betalains are defined as condensation products of betalamic acid with amines and amino acids; the relation between the structure and properties of these compounds is still under investigation (Gandia-Herrero et al., 2010).

The production of phenolic secondary metabolites in plants has been documented as an adaptive response to oxidative stress. For example, the transport of flavonoids to zones of generation of reactive oxygen species has been reported as a measure to diminish the effect of the oxidative stress by reducing diverse forms of reactive oxigen (Agati et al., 2012).

Industrial applications

The health properties of aromatic secondary metabolites derived from L-phenylalanine and L-tyrosine has been widely studied. Many flavonoids have been attributed anti-bacterial, anti-ageing, anti-cancer and anti-Alzheimer properties (Choi *et al.,* 2012; Wedick *et al.,* 2012; Liu *et al.,* 2012; Raza *et al.,* 2015; Hamalainen *et al.,* 2015).

One of the most promising uses of aromatic secondary metabolites is their application as anticancer agents, which has been shown in many recent

publications using human cells or animals as models. For example, coumarin-chalcones has been tested in four lines of human cancer cells and some of them showed high cytotoxicity and selectivity for this type of cells. Other compounds that have shown positive results are myricetin and luteonin, these compounds inhibit proliferation and viability of human cancer cells. Other studies using animal models have shown that quercetin, fisetin and ampelopsin caused apoptosis and decreased the proliferation of cancer cells both *in vivo* and *in vitro* (Sashidhara et al; 2010; Angst et al., 2013; Sun et al., 2012; Ni et al., 2012).

Another interesting property of aromatic secondary metabolites is their anti-aging effects. One of the principal causes of skin aging is the presence of free radicals. Since many aromatic secondary metabolites are antioxidants, they are ideal additives for anti-aging products (Chuarienthong et al., 2010; Baxter et al., 2008; Jung et al., 2010). Cosmetics companies such as L'oreal, Estee Lauder and Procter and Gamble have integrated flavonoids in some of their anti-aging products. Patents from these companies related to anti-aging formulations where flavonoids are part of the components have been registered (N'guyen et al., 1994; Chen et al., 2006; Declerq et al., 2014).

Other studies have shown that flavonoids such as baicalein, baicalin scutellarin, hibifolin, and quercetin attenuated neuronal cell damage and induction of cell death and other negative effects of oxidative stress in neuronal cells cultured in medium containing reactive oxygen species (Gao et al., 2001; Zhu et al., 2007). Recent studies on cardiovascular diseases and chronical inflammations have shown that consumption of aromatic secondary metabolites have positive effects on people with high risk of cardiovascular disease or chronicle inflammation (Bisht et al., 2010; Macready et al., 2014).

The pigments betalains and anthocyanins have been used permanently in the food industry. Due to the possible side effects of synthetic dyes, it is expected that the market for natural pigments will continue to grow in the following years. Companies such as DDW color, Chr Hansen and Sensient are working on the development and production of natural dyes. Recent

research in pigments biotechnology has been focused on stability enhancement (Cavalcanti et al., 2011).

The pigments anthocyanins and betalains have been used as synthesizers in solar cells for conversion of visible light into electricity; experimental results have shown that it is possible to convert approximately 2% of a unit of sun light power into electricity (Calogero et al., 2012). The development of dye-sensitized solar cells needs further improvement to reach large-scale production of electricity; this is an innovative alternative for energy generation.

1.4 Systems biology and development of cell factories

Systems biology can be defined as the analysis of the cell interactions at different levels and the subsequent integration of this information for a better understanding of organisms. Systems biology analysis of cell factories is fundamental for the optimization of platform strains. Different omics techniques such as transcriptomics, proteomics and metabolomics allow collecting information at different levels and integrating it through computational methods.

For the third paper, we did an omics analysis of our *p*-coumaric acid producer strain and we got information about the metabolic changes that the yeast cells faced when they were producing aromatic compounds. In this section, we show the utility of the omics techniques that we used in our project, and we make an overview of other techniques that we did not use and that should be considered when engineering cell factories.

Omics techniques for the development of cell factories

For the development of efficient cell factories, it is necessary to overcome the limitations inherent to the chosen organism, as metabolic regulations, generation of by-products, toxicity, cofactor imbalance. (Kim et al., 2012). These limitations used to be approached through rational metabolic engineering and adaptive evolution.

Techniques such as genomics, transcriptomics, proteomics, metabolomics, lipidomics and localizomics, combined with appropriate tools for data integration may have a significant repercussion on the development of cell factories.

The quantification of intracellular and extracellular metabolites has contributed to the identification of regulated steps in the pathways of interest and enzymatic specificity in different platform strains. Metabolic profiling of intermediate compounds or metabolites with known physiological relevance also has been useful for to the improvement of cell factories (Harrison *et al*., 2013; Toya and Shimizu, 2013). For example, metabolomics analysis has

been used for the improvement of the resistance of a *S. cerevisiae* xylose-fermenting strains to weak acids: Hasunuma et al. (2011) found that yeast cells accumulated compounds involved in the non-oxidative pentose phosphate pathway when weak acids were added to the medium. Interestingly the overexpression of enzymes involved in the pentose phosphate pathway TAL and TKL improved the production of ethanol when weak acids were added to the medium.

Comparative genomics has been used for studying the genome of xylose-fermenting fungi. By comparing the genome of 5 fungi species with different xylose consumption phenotypes; they identified a considerable number of genes related to carbohydrate transport and metabolism involved in xylose assimilation, the subsequent expression of these genes in *Sacharomyces cerevisiae* had a positive effect on growth and assimilation of xylose (Wohlbach et al., 2011).

The use of transcriptomics analysis and *in silico* knockout simulations led to the improvement of the synthesis of L-valine in *E. coli.* The transcriptome analysis allowed them to identify downregulated genes such as *lrp,* encoding the global regulator Lrp and the L-valine exporter *ygaZH*. Subsequently, these genes were expressed in the strain and led the improvement of L-valine biosynthesis. The implementation of *in silico* knockout targets from a model improved the flux through the pentose phosphate pathway arising an increase in the synthesis of NADPH, a cofactor needed for L-valine biosynthesis. The final strain had an outstanding yield of 0.378 g of L-valine per gram of glucose (Park et al., 2007).

Lactococcus lactis is an organism widely used for the production of fermented dairy products. Comparative and functional genomics have been used to identify differences between industrial and non-industrial subspecies of *L. lactis.* Through the use of genomics, transcriptomics and proteomics it has been possible to determine biological traits relevant for the production and development of flavour, texture and stability of dairy products. This information has been used for subsequent strain development, for example, proteolysis plays a key role in flavour development of cheese, subsequently

many of the genes and enzymes involved in this process have been sequenced and widely studied (Kok et al., 2005, van Hylckama et al., 2006).

The development and improvement of omics techniques allow getting information at different biological levels such as DNA, proteins, metabolites and the combined used of this information allowed getting a better understanding of cells and organisms not only in metabolic engineering but also in human biology and diseases treatment.

Genome-scale models

One of the most important aims of systems biology is the development of accurate mathematical models representing metabolic interactions. In this regard, the development of omics techniques have contributed to gathering more information about metabolic processes at different levels.

Many genome-scale models have been developed in the last years for different platform strains such as *E. coli, S. cerevisiae* and *Clostridium thermocellum*. These models have been very useful for the improvement of the strains for biotechnological applications and in the long term all the information will contribute to a better understanding of the physiology of the organisms.

Mathematical models can contribute to the development of cell factories at different stages of the process. By the use of models, it is possible to get information about the optimal host strain, identification of targets that may contribute to the improvement of production yield and titer of the desired compound and optimization of the fermentation process (Almquist et al., 2014).

Genome-scale models have been used to improve the production of ethanol in yeast. Dikicioglu et al. (2008) predicted the production of ethanol by flux balance analysis for mutants with deletions of a subunit of the respiratory chain complex III; interestingly after the deletion of the subunits the strains reduced biomass yields and increased ethanol yields.

Vanillin production also has been improved through genome-scale stoichiometric modeling using MOMA as the biological objective function. Brochado et al. (2010) identified GDH1, an enzyme involved in ammonium metabolism and PDC1, an enzyme involved in the carboxyl reaction of pyruvate to acetaldehyde and other intermediate compounds involved in ethanol synthesis as targets for the improvement of vanillin production. Interestingly, the strains with the deletion of these genes improved the synthesis of *p*-CA around 2-folds in comparison to the control strain.

Another interesting example is the biosynthesis of succinic acid. By the use of a genome-scale model Agren et al. (2013) identified that the deletion of *dic1* improved the synthesis of succinic acid to 0.02 C-mol/C-mol, further analysis of the *dic1* deletion strain showed its relation with gene ontology processes involved in inter-compartmental transports and redox balance.

1.5 Strategies for strain engineering

Through the research projects, we developed and characterized a platform strain for the synthesis of aromatic secondary metabolites. This platform was used for the synthesis of flavonoids but since we observe accumulation of intermediate compounds in the engineered strains, further engineering strategies are necessary for improving the synthesis of aromatic secondary metabolites in yeast. In this section, we present an overview of several techniques and strategies that can be used for the improvement of the aromatic secondary metabolites cell factories.

Multivariate modular metabolic engineering

Multivariate modular metabolic engineering (MMME) is an approach developed by Ajkumar et al. (2010). Different promoters, homologous enzymes, enzyme localization and copy numbers are combined in different modules; these modules are defined according to the enzyme turnover and chemistry. Finally the optimal expression level for the synthesis of the target compound is identified using multivariate statistics. With this approach, it is possible to find limiting steps in the pathway, since multiple combinations of promoters, enzymes and copy numbers are tested. This method also allows identifying toxicity of intermediate compounds and unknown pathways competing for any of the intermediate compounds.

MMME and similar approaches have been used for development of cell factories both in *E. coli* and *S. cerevisiae*. The use of this method in *E. coli* cell factories made it possible to overcome the toxicity of cinnamoyl-CoA for the production of (2S)-pinocembrin and the increment of the central metabolites precursors for the production of β- carotene resveratrol (Wu et al., 2013a; 2013b). In *S. cerevisiae* this approach was used to improve the production of the terpenoids intermediate compound miltiradiene (Dai et al., 2012; Zhou et al., 2012; Wu et al., 2014).

To summarize this approach can also be very useful for the development of a flavonoids cell factory. On this pathway, we already have a limiting step for the conversion of *p*-coumaric acid to naringenin or liquiritigenin and after

overcoming this limiting step, the tuning of the enzymatic steps downstream can probably contribute to improvement of performance of the flavonoids cell factory.

Genome editing

The development of yeast cell factories involves engineering of the host strain metabolism together with the expression of heterologous pathways. The development of a platform strain can involve deletion, overexpression, down-regulation and up-regulation of genes. A critical part of the engineering process is the implementation of techniques that allow engineering multiple metabolic targets in a short period of time with a reasonable cost.

The efficient endogenous homologous recombination mechanism of *S. cerevisiae* had played a fundamental role in yeast strains development; this organism shows a high frequency of homologous recombination, which allowed development of different methods for genome editing (Srikrishnan., 2011; Jensen et al., 2013; Jakociunas et al., 2016).

The use of yeast-oligo mediated genome engineering YOGE has shown practical advantages for genome editing in *S. cerevisiae*, since single strand oligos have a reasonable price, and it is not necessary to perform a PCR amplification and subsequent purification. The efficiency of oligo-mediated recombination combined with a selection of desired phenotypes had shown positive results for the development of deletions and overexpressions in yeast (DiCarlo et al., 2013).

Partially synthetic chromosomes also had been developed for *S. cerevisiae*; this chromosome had the phenotype and fitness of a wild-type chromosome, lacking destabilizing elements and genetic flexibility. By the use of this technique, it is possible to perform combinatorial mutagenesis generating a broad range of phenotypes (Dymond et al., 2011; Shen et al., 2016).

In recent years, we have also seen a rapid advance in genome engineering techniques due to the emergence of new tools based on the CRISPR/Cas9 technology. These tools are very useful for selective targeting of DNA to

create double strand breaks. CRISPR/Cas9 successfully had been used for development of several different genome engineering strategies in *S. cerevisiae*. These techniques made it possible to generate simultaneous gene deletions, multigene chromosomal integration and single directed mutagenesis (Bao et al., 2014; Mans et al., 2015; Jakociunas et al., 2015a).

The CRÍSPR/Cas9 technology allows performing multiple deletion and integration of genes in a shorter period in comparison with other methods. For example, it is possible to perform single and multiple deletions of genes with efficiencies of 100% and up to 43% respectively (Ryan et al., 2014. Also, it is possible to do multi-loci integrations of up to six genes with efficiencies between 75 to 100%, and combined integration and deletion of genes with efficiencies of 58% (Mans et al., 2015; Jakociunas et al., 2015a).

The availability of selection markers limits the amount of genetic modifications that can be performed in a strain. Since the use of CRISPR/Cas9 for genome editing doesn't need selection markers, the use of this technique open new possibilities for the expression of long heterologous pathways in yeast (Jakociunas et al., 2015b; Solis-Escalante et al., 2013).

As the use of multivariate combinatorial strategies such as MMME for pathways with many enzymatic steps relies on the practical issues related to strains construction, the use of CRISPR/Cas9 technology for genome engineering is an excellent option for facilitating the development and improvement of cell factories.

Enzyme co-localization

Strategies such as the development of fusion proteins by the fusion of genes with a linker sequence can contribute to the improvement of yeast cell factories. The fusion has to be designed according to the sequential order of the enzymes in the pathway. With this approach, it is possible to increase the local concentration of intermediate compounds and improve their flux to subsequent enzymatic steps, and even decrease the flux of intermediate

compounds to unwanted metabolic pathways (Li et al., 2010; Wang et al., 2012).

Other options for enzyme co-localization are synthetic enzymes scaffold and compartmentalization. While synthetic enzyme scaffolds allows the binding of more than two enzymes and controlling the enzymes' stoichiometry. Further, compartmentalization allows localizing, even more, enzymes than the synthetic scaffold and also permits the control of some chemical characteristics of the compartment, such as pH and redox state. Compartmentalization has already been used as a strategy for the production of terpenoids and branched chain alcohols (Dueber et al., 2009; Farhi et al., 2013; Avalos et al., 2013; Woolston., 2013).

Protein Engineering

Many factors can affect the optimal performance of an enzyme such as allosteric regulation, enzyme specificity, cofactor imbalance or optimal temperature and pH (Sauro., 2011). The use of computational tools combined with specific knowledge about the protein sequence and structure contribute to the improvement of the enzyme performance through protein engineering (Chen et al., 2013).

Protein engineering allows identifying and targeting residues associated with inhibitory binding, designing strategies to modify the electrostatic conformation of the cofactor binding sites and designing of libraries with targets for improving enzyme specificity. Protein engineering is a permanent practice for the development of protein therapeutics and development of tailored enzymes in the biotechnology industry; however, it is not a permanent practice for metabolic engineering and strain development (Carter *et al.,* 2011; Bommarius *et al.,* 2013).

Protein engineering had been used in combination with metabolic engineering for production of diterpenoids as a strategy to overcome the insufficient downstream capacity. By the generation of a library of combinatorial mutations for the rate-limiting enzymes, a 2600 fold improvement in the production of levopimaradiene in *E. coli* was reached. In

S. cerevisiae, it was possible to improve the catalytic activity of the glucose regulator Hexokinase 2 in the presence of xylose (Leonard et al., 2010; Bergdahl et al., 2013).

References

Agati, G., Azzarello, E., Pollastri, S., Tattini, M., 2012. Flavonoids as antioxidants in plants: Location and functional significance. Plant Sci. 196, 67–76. doi:10.1016/j.plantsci.2012.07.014

Agren, R., Otero, J.M., Nielsen, J., 2013. Genome-scale modeling enables metabolic engineering of *Saccharomyces cerevisiae* for succinic acid production. J. Ind. Microbiol. Biotechnol. 40, 735–747. doi:10.1007/s10295-013-1269-3

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E.J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., Stephanopoulos, G., 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. Science 330, 70–4. doi:10.1126/science.1191652

Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., Jirstrand, M., 2014. Kinetic models in industrial biotechnology - Improving cell factory performance. Metab. Eng. 24, 38–60. doi:10.1016/j.ymben.2014.03.007

Angst, E., Park, J.L., Moro, A., Lu, Q., Li, G., King, J., Chen, M., Reber, H.A., Go, V.L.W., Eibl, G., Hines, O.J., 2014. The flavonoid quercetin inhibits pancreatic cancer growth in vitro and in vivo. Pancreas 42, 223–229. doi:10.1097/MPA.0b013e318264.

Avalos, J.L., Fink, G.R., Stephanopoulos, G., 2013. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. Nat. Biotechnol. 31, 335–41. doi:10.1038/nbt.2509

Baxter, R.A., 2008. Anti-aging properties of resveratrol: Review and report of a potent new antioxidant skin care formulation. J. Cosmet. Dermatol. 7, 2–7. doi:10.1111/j.1473-2165.2008.00354.

Becker, J., Wittmann, C., 2012. Bio-based production of chemicals, materials and fuels - Corynebacterium glutamicum as versatile cell factory. Curr. Opin. Biotechnol. 23, 631–640. doi:10.1016/j.copbio.2011.11.012

Bergdahl, B., Sandström, A.G., Borgström, C., Boonyawan, T., van Niel, E.W.J., Gorwa-Grauslund, M.F., 2013. Engineering Yeast Hexokinase 2 for Improved Tolerance Toward Xylose-Induced Inactivation. PLoS One 8, 1–10. doi:10.1371/journal.

Bertin, C., Yang, X., Weston, L.A., 2003. The role of root exudates and allelochemicals in the rhizosphere. Plant Soil 256, 67–83. doi:10.1023/A:1026290508166

Bhuiyan, N.H., Selvaraj, G., Wei, Y., King, J., 2009. Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defence in wheat against powdery mildew invasion. J. Exp. Bot. 60, 509–521. doi:10.1093/jxb/ern290

Bisht, K., Wagner, K.H., Bulmer, A.C., 2010. Curcumin, resveratrol and flavonoids as anti-inflammatory, cyto- and DNA-protective dietary compounds. Toxicology 278, 88–100. doi:10.1016/j.tox.2009.11.008

Bommarius, A.S., Blum, J.K., Abrahamson, M.J., 2011. Status of protein engineering for biocatalysts: How to design an industrially useful biocatalyst. Curr. Opin. Chem. Biol. 15, 194–200. doi:10.1016/j.cbpa. 2010.11.011

Braus, G.H., 1991. Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: a model system for the regulation of a eukaryotic biosynthetic pathway. Microbiol. Rev. 55, 349–70.

Brochado, A.R., Matos, C., Møller, B.L., Hansen, J., Mortensen, U.H., Patil, K.R., 2010. Improved vanillin production in baker's yeast through in silico design. Microb. Cell Fact. 9, 84. doi:10.1186/1475-2859-9-84

Bross, C.D., Corea, O.R.A., Kaldis, A., Menassa, R., Bernards, M.A., Kohalmi, S.E., 2011. Complementation of the pha2 yeast mutant suggests functional differences for arogenate dehydratases from Arabidopsis thaliana. Plant Physiol. Biochem. 49, 882–890. doi:10.1016/j.plaphy.2011.02.010

Buchanan, B. B., Gruissem, W., & Jones, R. L. (2000). Biochemistry and molecular biology of plants. New York: Wiley. p. 1286.

Bulzomi, P., Bolli, A., Galluzzo, P., Leone, S., Acconcia, F., Marino, M., 2010. Naringenin and 17beta-estradiol coadministration prevents hormone-induced human cancer cell growth. IUBMB Life 62, 51–60. doi:10.1002/iub.279

Calero, P., Jensen, S.I., Nielsen, A.T., 2016. Broad host range ProUSER vectors enable fast characterization of inducible promoters and optimization of *p*-coumaric acid production in Pseudomonas putida KT2440. ACS Synth. Biol. doi:10.1021/acssynbio.6b00081

Calogero, G., Yum, J.H., Sinopoli, A., Di Marco, G., Grätzel, M., Nazeeruddin, M.K., 2012. Anthocyanins and betalains as light-harvesting pigments for dye-sensitized solar cells. Sol. Energy 86, 1563–1575. doi:10.1016/j.solener.2012.02.018

Carter, P.J., 2011. Introduction to current and future protein therapeutics: A protein engineering perspective. Exp. Cell Res. 317, 1261–1269. doi:10.1016/j.yexcr.2011.02.013

Cavalcanti, R.N., Santos, D.T., Meireles, M.A.A., 2011. Non-thermal stabilization mechanisms of anthocyanins in model and food systems-An overview. Food Res. Int. 44, 499–509. doi:10.1016/j.foodres.2010.12.007

Cheynier, V., Comte, G., Davies, K.M., Lattanzio, V., Martens, S., 2013. Plant phenolics: Recent advances on their biosynthesis, genetics, and ecophysiology. Plant Physiol. Biochem. 72, 1–20. doi:10.1016/j.plaphy.2013.05.009

Chemler, J. a, Yan, Y., Koffas, M. a G., 2006. Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. Microb. Cell Fact. 5, 20. doi:10.1186/1475-2859-5-20

Chen, Z., Zeng, A.P., 2013. Protein design in systems metabolic engineering for industrial strain development. Biotechnol. J. 8, 523–533. doi:10.1002/biot.201200238

Choi, S.H., Ryu, M., Yoon, Y.J., Kim, D.M., Lee, E.Y., 2012. Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from Streptomyces antibioticus in batch and repeated batch modes. Biotechnol. Lett. 34, 499–505. doi:10.1007/s10529-011-0789-z

Clomburg, J.M., Gonzalez, R., 2010. Biofuel production in *Escherichia coli*: The role of metabolic engineering and synthetic biology. Appl. Microbiol. Biotechnol. 86, 419–434. doi:10.1007/s00253-010-2446-1

Chen, M., Sun, P., Date, A., Yuyama, E., Bissett, D.L. 2007. Skin care composition containing a flavonoid and vitamin b3. WO patent 2006020164 A1.

Chuarienthong, P., Lourith, N., Leelapornpisid, P., 2010. Clinical efficacy comparison of anti-wrinkle cosmetics containing herbal flavonoids. Int. J. Cosmet. Sci. 32, 99–106. doi:10.1111/j.1468-2494.2010.00522.

Crozier, A., Jaganath, I.B., Clifford, M.N., 2007. Phenols, Polyphenols and Tannins: An Overview. Plant Second. Metab. Occur. Struct. Role Hum. Diet 1–24. doi:10.1002/9780470988558.ch1

Daboussi, F., Leduc, S., Maréchal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D.F., Cavarec, L., Duchateau, P., 2014. Genome engineering empowers the diatom Phaeodactylum tricornutum for biotechnology. Nat. Commun. 5, 3831. doi:10.1038/ncomms4831

Dai, Z., Liu, Y., Huang, L., Zhang, X., 2012. Production of miltiradiene by metabolically engineered *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 109, 2845–2853. doi:10.1002/bit.24547

Declercq,L., Corstjens, H., Maes, D., Van, B.W., Schelkens, G. 2006. Topical compositions containing phosphorylated polyphenols. WO patent 2006029484 A1

Dicarlo, J.E., Conley, A.J., Penttila, M., Jantti, J., Wang, H.H., Church, G.M., 2013a. Yeast oligo-mediated genome engineering (YOGE). ACS Synth. Biol. 2, 741–749. doi:10.1021/sb400117c

Dikicioglu, D., Pir, P., Onsan, Z.I., Ulgen, K.O., Kirdar, B., Oliver, S.G., 2008. Integration of metabolic modeling and phenotypic data in evaluation and improvement of ethanol production using respiration-deficient mutants of *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 74, 5809–5816. doi:10.1128/AEM.00009-08

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A. V, Prather, K.L.J., Keasling, J.D., 2009. Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. 27, 5–8. doi:10.1038/nbt.1557

Duncan, K., Edwards, R.M., Coggins, J.R., 1988. The *Saccharomyces cerevisiae* ARO1 gene an example of the co-ordinate regulation of five enzymes on a single biosynthetic pathway. FEBS Lett. 241, 83–88. doi:10.1016/0014-5793(88)81036-6

Dymond, J.S., Richardson, S.M., Coombes, C.E., Babatz, T., Muller, H., Annaluru, N., Blake, W.J., Schwerzmann, J.W., Dai, J., Lindstrom, D.L., Boeke, A.C., Gottschling, D.E., Chandrasegaran, S., Bader, J.S., Boeke, J.D., 2011. Synthetic chromosome arms function in yeast and

generate phenotypic diversity by design. Nature 477, 471–6. doi:10.1038/nature10403

Farhi, M., Kozin, M., Duchin, S., Vainstein, A., 2013. Metabolic engineering of plants for artemisinin synthesis. Biotechnol. Genet. Eng. Rev. 29, 135–48. doi:10.1080/02648725.2013.821283

Gandía-Herrero, F., Escribano, J., García-Carmona, F., 2010. Structural implications on color, fluorescence, and antiradical activity in betalains. Planta 232, 449–460. doi:10.1007/s00425-010-1191-0

Gao, Z., Huang, K., Xu, H., 2001. Protective effects of flavonoids in the roots of scutellaria baicalensis georgi against hydrogen peroxide-induced oxidative stress in hs-sy5y cells. Pharmacol. Res. 43, 173–178. doi:10.1006/phrs.2000.0761

Gimpel, J.A., Specht, E.A., Georgianna, D.R., Mayfield, S.P., 2013. Advances in microalgae engineering and synthetic biology applications for biofuel production. Curr. Opin. Chem. Biol. 17, 489–495. doi:10.1016/j.cbpa.2013.03.038

Goiris, K., Muylaert, K., Voorspoels, S., Noten, B., De Paepe, D., E Baart, G.J., De Cooman, L., 2014. Detection of flavonoids in microalgae from different evolutionary lineages. J. Phycol. 50, 483–492. doi:10.1111/jpy.12180

Hamalainen, M., Nieminen, R., Vuorela, P., Heinonen, M., Moilanen, E., 2007. Anti-inflammatory effects of flavonoids: Genistein, kaempferol, quercetin, and daidzein inhibit STAT-1 and NF-κB activations, whereas flavone, isorhamnetin, naringenin, and pelargonidin inhibit only NF-κB activation along with their inhibitory effect on i. Mediators Inflamm. 2007. doi:10.1155/2007/45673

Harrison, S.J., Herrgård, M.J., 2013. The Uses and Future Prospects of Metabolomics and Targeted Metabolite Profiling in Cell Factory Development. Ind. Biotechnol. 9, 196–202. doi:10.1089/ind.2013.0008

Hartwig, U.A., Maxwell, C.A., Joseph, C.M., Phillips, D.A., 1990. Chrysoeriol and Luteolin Released from Alfalfa Seeds Induce nod Genes in Rhizobium meliloti. Plant Physiol. 92, 116–22. doi:10.1104/pp.92.1.116

Hasunuma, T., Sanda, T., Yamada, R., Yoshimura, K., Ishii, J., Kondo, A., 2011. Metabolic pathway engineering based on metabolomics confers

acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*. Microb. Cell Fact. 10, 2. doi:10.1186/1475-2859-10-2

Hawkins, K.M., Smolke, C.D., 2008. Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. Nat. Chem. Biol. 4, 564–573. doi:10.1038/nchembio.105

Helmstaedt, K., Strittmatter, A., Lipscomb, W.N., Braus, G.H., 2005. Evolution of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase-encoding genes in the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 102, 9784–9789. doi:10.1073/pnas.0504238102

Herrmann, K.M., Weaver, L.M., 1999. the Shikimate Pathway. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50, 473–503. doi:10.1146/annurev.arplant.50.1.473

Hinnebusch, A.G., 2005. Translational regulation of *gcn4* and the general amino acid control of yeast*. Annu. Rev. Microbiol. 59, 407–450. doi:10.1146/annurev.micro.59.031805.133833

Jakociunas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., Keasling, J.D., 2015a. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. Metab. Eng. 28, 213–222. doi:10.1016/j.ymben.2015.01.008

Jakociunas, T., Jensen, M.K., Keasling, J.D., 2015b. CRISPR/Cas9 advances engineering of microbial cell factories. Metab. Eng. 34, 44–59. doi:10.1016/j.ymben.2015.12.003

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., Borodina, I., 2014. EasyClone: Method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. FEMS Yeast Res. 14, 238–248. doi:10.1111/1567-1364.12118

Jiang, H., Wood, K. V, Morgan, J. a, 2005. Metabolic Engineering of the Phenylpropanoid Pathway in *Saccharomyces cerevisiae* Metabolic Engineering of the Phenylpropanoid Pathway in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 71, 2962–2969. doi:10.1128/AEM.71.6.2962

Jones, J.A., Vernacchio, V.R., Sinkoe, A.L., Collins, S.M., Ibrahim, M.H.A., Lachance, D.M., Hahn, J., Koffas, M.A.G., 2016. Experimental and computational optimization of an *Escherichia coli* co-culture for the efficient production of flavonoids. Metab. Eng. 35, 55–63. doi:10.1016/j.ymben.2016.01.006

Jung, S.K., Lee, K.W., Kim, H.Y., Oh, M.H., Byun, S., Lim, S.H., Heo, Y.S., Kang, N.J., Bode, A.M., Dong, Z., Lee, H.J., 2010. Myricetin suppresses UVB-induced wrinkle formation and MMP-9 expression by inhibiting Raf. Biochem. Pharmacol. 79, 1455–1461. doi:10.1016/ j.bcp.2010.01.004

Kim, I.K., Roldão, A., Siewers, V., Nielsen, J., 2012. A systems-level approach for metabolic engineering of yeast cell factories. FEMS Yeast Res. 12, 228–248. doi:10.1111/j.1567-1364.2011.00779.

Khoo, B.Y., Chua, S.L., Balaram, P., 2010. Apoptotic effects of chrysin in human cancer cell lines. Int. J. Mol. Sci. 11, 2188–2199. doi:10.3390/ijms11052188

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J. a, Pronk, J.T., Daran, J.-M., 2012. De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microb. Cell Fact. 11, 155. doi:10.1186/1475-2859-11-155

Kovacik, J., Klejdus, B., Backor, M., 2010. Physiological responses of scenedesmus quadricauda (Chlorophyceae) to UV-A and UV-C light. Photochem. Photobiol. 86, 612–616. doi:10.1111/j.1751-1097.2010.00708.

Kok, J., Buist, G., Zomer, A.L., Van Hijum, S.A.F.T., Kuipers, O.P., 2005. Comparative and functional genomics of lactococci. FEMS Microbiol. Rev. 29, 411–433. doi:10.1016/j.femsre.2005.04.004

Kunzler, M., Paravicini, G., Egli, C.M., Irniger, S., Braus, G.H., 1992a. Cloning, primary structure and regulation of the ARO4 gene, encoding the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae*. Gene 113, 67–74. doi:10.1016/0378-1119(92)90670-K.

Landry, L.C., Chapple, C.C.S., Last, R. 1, 1995. Arabidopsis Mutants Lacking Phenolic Sunscreens Exhibit Enhanced Ultraviolet-B lnjury and

Oxidative Damage'. Plant Physiol 109, 59–1. doi:10.1104/pp.109.4.1159

Larkum, A.W.D., Ross, I.L., Kruse, O., Hankamer, B., 2012. Selection, breeding and engineering of microalgae for bioenergy and biofuel production. Trends Biotechnol. 30, 198–204. doi:10.1016/j.tibtech.2011.11.003

Leonard, E., Yan, Y., Lim, K.H., Koffas, M.A.G., 2005. Investigation of two distinct flavone synthases for plant-specific flavone biosynthesis in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 71, 8241–8248. doi:10.1128/AEM.71.12.8241-8248.

Leonard, E., Yan, Y., Koffas, M.A.G., 2006. Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. Metab. Eng. 8, 172–181. doi:10.1016/j.ymben.2005.11.001

Leonard, E., Lim, K.-H., Saw, P.-N., Koffas, M. a G., 2007. Engineering central metabolic pathways for high-level flavonoid production in *Escherichia coli*. Appl. Environ. Microbiol. 73, 3877–86. doi:10.1128/AEM.00200-07

Leonard, E., Yan, Y., Fowler, Z.L., Li, Z., Lim, C.G., Lim, K.H., Koffas, M.A.G., 2008. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. Mol. Pharm. 5, 257–265. doi:10.1021/mp7001472

Leonard, E., Ajikumar, P.K., Thayer, K., Xiao, W.-H., Mo, J.D., Tidor, B., Stephanopoulos, G., Prather, K.L.J., 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. Proc. Natl. Acad. Sci. U. S. A. 107, 13654–13659. doi:10.1073/pnas.1006138107

Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I., Nielsen, J., 2015. Author ' s Accepted Manuscript De novo production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae* To appear in : Metabolic Engineering. Metab. Eng. 32, 1–11. doi:10.1016/j.ymben.2015.08.007

Liu, Z., Li, X., 2014. The kinetics of ethanol fermentation based on adsorption processes. Kem. u Ind. 63, 259–264. doi:10.15255/KUI.2013.023

Luttik, M. a H., Vuralhan, Z., Suir, E., Braus, G.H., Pronk, J.T., Daran, J.M., 2008. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: Quantification of metabolic impact. Metab. Eng. 10, 141–153. doi:10.1016/j.ymben.2008.02.002

Macready, A.L., George, T.W., Chong, M.F., Alimbetov, D.S., Jin, Y., Vidal, A., Spencer, J.P.E., Kennedy, O.B., Tuohy, K.M., Minihane, A., Gordon, M.H., Lovegrove, J.A., Study, F., 2014. Flavonoid-rich fruit and vegetables improve microvascular reactivity and inflammatory status in men at risk of cardiovascular disease — FLAVURS : a randomized controlled trial 1 – 5. Am. J. Clin. Nutr. 99, 479–489. doi:10.3945/ajcn.113.074237.Epidemiologic

Maeda, H., Dudareva, N., 2012. The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. Annu. Rev. Plant Biol. 63, 73–105. doi:10.1146/annurev-arplant-042811-105439

Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.A., Daran, J.M.G., 2015. CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 15, 1–15. doi:10.1093/femsyr/fov004

Mannhaupt G, Stucka R, Pilz U, Scharzlose C, Feldmann H., 1989. Characterization of the prephenate dehydrogenase-encoding gene, TYR1, from *Saccharomyces cerevisiae*. 85(2):303–11. doi:10. 1016/0378-1119(89)90422-8.

Naesby, M., Nielsen, S.V., Nielsen, C.A., Green, T., Tange, T.O., Simón, E., Knechtle, P., Hansson, A., Schwab, M.S., Titiz, O., Folly, C., Archila, R.E., Maver, M., van Sint Fiet, S., Boussemghoune, T., Janes, M., Kumar, a S.S., Sonkar, S.P., Mitra, P.P., Benjamin, V.A.K., Korrapati, N., Suman, I., Hansen, E.H., Thybo, T., Goldsmith, N., Sorensen, A.S., 2009. Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. Microb. Cell Fact. 8, 45. doi:10.1186/1475-2859-8-45

Naoumkina, M.A., Zhao, Q., Gallego-Giraldo, L., Dai, X., Zhao, P.X., Dixon, R.A., 2010. Genome-wide analysis of phenylpropanoid defence pathways. Mol. Plant Pathol. 11, 829–846. doi:10.1111/j.1364-3703.2010.00648.

N'guyen. 1994. Cosmetic or dermopharmaceutical composition containing in association a lauroylmethionate of a basic amino acid and at least a polyphenol. EP patent 0595694 A1.

Ni, F., Gong, Y., Li, L., Abdolmaleky, H.M., Zhou, J.R., 2012. Flavonoid ampelopsin inhibits the growth and metastasis of prostate cancer in vitro and in mice. PLoS One 7. doi:10.1371/journal.pone.0038802

Nielsen, J., Fussenegger, M., Keasling, J., Lee, S.Y., Liao, J.C., Prather, K., Palsson, B., 2014. Engineering synergy in biotechnology. Nat. Chem. Biol. 10, 319–22. doi:10.1038/nchembio.1519

Ortuno, A., Diaz, L., Alvarez, N., Porras, I., Garcia-Lidon, A., Del Rio, J.A., 2011. Comparative study of flavonoid and scoparone accumulation in different Citrus species and their susceptibility to Penicillium digitatum. Food Chem. 125, 232–239. doi:10.1016/j.foodchem.2010.09.012

Park, S.R., Ahn, M.S., Han, A.R., Park, J.W., Yoon, Y.J., 2011. Enhanced flavonoid production in Streptomyces venezuelae via metabolic engineering. J. Microbiol. Biotechnol. 21, 1143–1146. doi:10.4014/jmb.1108.08012

Pandey, R.P., Parajuli, P., Koffas, M.A.G., Sohng, J.K., 2016. Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. Biotechnol. Adv. doi:10.1016/j.biotechadv.2016.02.012

Poblete-Castro, I., Becker, J., Dohnt, K., Santos, V.M. Dos, Wittmann, C., 2012. Industrial biotechnology of Pseudomonas putida and related species. Appl. Microbiol. Biotechnol. 93, 2279–2290. doi:10.1007/s00253-012-3928-0

Pröschel, M., Detsch, R., Boccaccini, A.R., Sonnewald, U., 2015. Engineering of Metabolic Pathways by Artificial Enzyme Channels. Front. Bioeng. Biotechnol. 3, 168. doi:10.3389/fbioe.2015.00168

Quevillon-Cheruel, S., Leulliot, N., Meyer, P., Graille, M., Bremang, M., Blondeau, K., Sorel, I., Poupon, A., Janin, J., Van Tilbeurgh, H., 2004. Crystal Structure of the Bifunctional Chorismate Synthase from *Saccharomyces cerevisiae*. J. Biol. Chem. 279, 619–625. doi:10.1074/jbc.M310380200

Raza, S.S., Khan, M.M., Ahmad, A., Ashafaq, M., Islam, F., Wagner, A.P., Safhi, M.M., Islam, F., 2013. Neuroprotective effect of naringenin is mediated through suppression of NF-KB signaling pathway in experimental stroke. Neuroscience 230, 157–171. doi:10.1016/ j.neuroscience. 2012.10.041

Richards, T.A., Dacks, J.B., Campbell, S.A., Blanchard, J.L., Foster, P.G., McLeod, R., Roberts, C.W., 2006. Evolutionary origins of the eukaryotic shikimate pathway: Gene fusions, horizontal gene transfer, and endosymbiotic replacements. Eukaryot. Cell 5, 1517–1531. doi:10.1128/EC.00106-06

Ryan, O.W., Skerker, J.M., Maurer, M.J., Li, X., Tsai, J.C., Poddar, S., Lee, M.E., DeLoache, W., Dueber, J.E., Arkin, A.P., Cate, J.H.D., 2014. Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife 3, 1–15. doi:10.7554/eLife.03703

Santos, C.N.S., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab. Eng. 13, 392–400. doi:10.1016/j.ymben.2011.02.002

Sashidhara, K. V., Kumar, A., Kumar, M., Sarkar, J., Sinha, S., 2010. Synthesis and in vitro evaluation of novel coumarin–chalcone hybrids as potential anticancer agents. Bioorg. Med. Chem. Lett. 20, 7205–7211. doi:10.1016/j.bmcl.2010.10.116

Sauro, H.M.,2011. Enzyme Kinetics for Systems Biology. Ambrosius Publishing, Seattle, ISBN:0982477317.

Savakis, P., Hellingwerf, K.J., 2015. Engineering cyanobacteria for direct biofuel production from CO2. Curr. Opin. Biotechnol. 33, 8–14. doi:10.1016/j.copbio.2014.09.007

Schmidheini, T., Mösch, H.U., Evans, J.N., Braus, G., 1990. Yeast allosteric chorismate mutase is locked in the activated state by a single amino acid substitution. Biochemistry 29, 3660–8. doi:066/0

Shen, Y., Stracquadanio, G., Wang, Y., Yang, K., Mitchell, L.A., Xue, Y., Cai, Y., Chen, T., Dymond, J.S., Kang, K., Gong, J., Zeng, X., Zhang, Y., Li, Y., Feng, Q., Xu, X., Wang, J., Wang, J., Yang, H., Boeke, J.D., Bader, J.S., 2016. SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. Genome Res. 26, 36–49. doi:10.1101/gr.193433.115

Shimizu, B.I., Miyagawa, H., Ueno, T., Sakata, K., Watanabe, K., Ogawa, K., 2005. Morning glory systemically accumulates scopoletin and scopolin after interaction with Fusarium oxysporum. Zeitschrift fur Naturforsch. - Sect. C J. Biosci. 60, 83–90.

Silakowski, B., Kunze, B., Müller, R., 2000. Stigmatella aurantiaca Sg a15 carries genes encoding type I and type II 3-deoxy- d-arabino-heptulosonate-7-phosphate synthases: Involvement of a type II synthase in aurachin biosynthesis. Arch. Microbiol. 173, 403–411. doi:10.1007/s002030000162

Solis-Escalante, D., Kuijpers, N.G.A., van der Linden, F.H., Pronk, J.T., Daran, J.M., Daran-Lapujade, P., 2014. Efficient simultaneous excision of multiple selectable marker cassettes using I-SceI-induced double-strand DNA breaks in *Saccharomyces cerevisiae*. FEMS Yeast Res. 14, 741–754. doi:10.1111/1567-1364.12162

Sun, F., Zheng, X.Y., Ye, J., Wu, T.T., Wang, J.L., Chen, W., 2012. Potential anticancer activity of myricetin in human T24 bladder cancer cells both in vitro and in vivo. Nutr. Cancer 64, 599–606. doi:10.1080/01635581.2012.665564

Toya, Y., Shimizu, H., 2013. Flux analysis and metabolomics for systematic metabolic engineering of microorganisms. Biotechnol. Adv. 31, 818–826. doi:10.1016/j.biotechadv.2013.05.002

Tzin, V., Galili, G., 2010. New Insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. Mol. Plant 3, 956–972. doi:10.1093/mp/ssq048

Urrestarazu, A., Vissers, S., Iraqui, I., Grenson, M., 1998. Phenylalanine- and tyrosine-auxotrophic mutants of *Saccharomyces cerevisiae* impaired in transamination. Mol. Gen. Genet. 257, 230–237. doi:10.1007/s004380050643

van Hylckama Vlieg, J.E., Rademaker, J.L., Bachmann, H., Molenaar, D., Kelly, W.J., Siezen, R.J., 2006. Natural diversity and adaptive responses of Lactococcus lactis. Curr. Opin. Biotechnol. 17, 183–190. doi:10.1016/j.copbio.2006.02.007

Wang, C., Yoon, H.S., Jang, J.H., Chung, R.Y., Kim, Y.J., Choi, S.E., Kim, W.S., 2011.Metabolic engineering of *Escherichia coli* for a-farnesene production. Metab.Eng. doi:10.1016/j.ymben.2011.08001.

Warren, J.M., Bassman, J.H., Fellman, J.K., Mattinson, D.S., Eigenbrode, S., 2003. Ultraviolet-B radiation alters phenolic salicylate and flavonoid composition of Populus trichocarpa leaves. Tree Physiol. 23, 527–535. doi:10.1093/treephys/23.8.527

Wedick, N.M., Pan, A., Cassidy, A., Rimm, E.B., Sampson, L., Rosner, B., Willett, W., Hu, F.B., Sun, Q., Van Dam, R.M., 2012. Dietary flavonoid intakes and risk of type 2 diabetes in US men and women. Am. J. Clin. Nutr. 95, 925–933. doi:10.3945/ajcn.111.028894

Wink, M. (2014) Evolution of Secondary Metabolism in Plants, in Ecological Biochemistry: Environmental and Interspecies Interactions (eds G.-J. Krauss and D. H. Nies), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany. doi: 10.1002/9783527686063.ch3

Winkel-Shirley, B., 2002. Biosynthesis of flavonoids and effects of stress biosynthesis of flavonoids and effects of stress. Curr. Opin. Plant Biol. 5, 218–223. doi:10.1016/S1369-5266(02)00256-X

Woelfle, U.; Simon-Haarhaus, B.; Merfort, I.; Schempp, C. Reseda luteola L. extract displays antiproliferative and pro-apoptotic activities that are related to its major flavonoids. Phytother. Res. 24:1033–1035; 2009.

Wohlbach, D.J., Kuo, A., Sato, T.K., Potts, K.M., Salamov, A.A., Labutti, K.M., Sun, H., Clum, A., Pangilinan, J.L., Lindquist, E.A., Lucas, S., Lapidus, A., Jin, M., Gunawan, C., Balan, V., Dale, B.E., Jeffries, T.W., Zinkel, R., Barry, K.W., Grigoriev, I. V, Gasch, A.P., 2011. Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. Proc. Natl. Acad. Sci. U. S. A. 108, 13212–7. doi:10.1073/pnas.1103039108

Woolston, B.M., Edgar, S., Stephanopoulos, G., 2013. Metabolic Engineering: Past and Future. Annu. Rev. Chem. Biomol. Eng. 4, 259–288. doi:10.1146/annurev-chembioeng-061312-103312

Wu, J., Du, G., Zhou, J., Chen, J., 2013. Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. Metab. Eng. 16, 48–55. doi:10.1016/j.ymben.2012.11.009

Wu, J., Liu, P., Fan, Y., Bao, H., Du, G., Zhou, J., Chen, J., 2013. Multivariate modular metabolic engineering of *Escherichia coli* to

produce resveratrol from l-tyrosine. J. Biotechnol. 167, 404–411. doi:10.1016/j.jbiotec.2013.07.030

Wu, J., Zhou, T., Du, G., Zhou, J., Chen, J., 2014. Modular optimization of heterologous pathways for de novo synthesis of (2S)-naringenin in *Escherichia coli*. PLoS One 9, e101492. doi:10.1371/journal.pone.0101492

Yu, C., Cao, Y., Zou, H., Xian, M., 2011. Metabolic engineering of *Escherichia coli* for biotechnological production of high-value organic acids and alcohols. Appl. Microbiol. Biotechnol. 89, 573–583. doi:10.1007/s00253-010-2970-z

Zhang, Y.H.P., 2011. Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol. Adv. 29, 715–725. doi:10.1016/j.biotechadv.2011.05.020

Zhang, H., Wang, X., 2016. Modular co-culture engineering, a new approach for metabolic engineering. Metab. Eng. 37, 114–121. doi:10.1016/ j.ymben.2016.05.007

Zhou, Y.J., Gao, W., Rong, Q., Jin, G., Chu, H., Liu, W., Yang, W., Zhu, Z., Li, G., Zhu, G., Huang, L., Zhao, Z.K., 2012. Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. J. Am. Chem. Soc. 134, 3234–3241. doi:10.1021/ja2114486

Zhou, K., Qiao, K., Edgar, S., Stephanopoulos, G., 2015. Distributing a metabolic pathway among a microbial consortium enhances production of natural products. Nat. Biotechnol. 33, 377–383. doi:10.1038/nbt.3095

Zhu, J.T.T., Choi, R.C.Y., Chu, G.K.Y., Cheung, A.W.H., Gao, Q.T., Li, J., Jiang, Z.Y., Dong, T.T.X., Tsim, K.W.K., 2007. Flavonoids possess neuroprotective effects on cultured pheochromocytoma PC12 cells: A comparison of different flavonoids in activating estrogenic effect and in preventing β-amyloid-induced cell death. J. Agric. Food Chem. 55, 2438–2445. doi:10.1021/jf063299zDevelopment of a platform strain for synthesis of *p*-coumaric acid

2. Development of a platform strain for synthesis of *p*-coumaric acid

Metabolic Engineering 31 (2015) 181–188

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Establishment of a yeast platform strain for production of *p*-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis

Angelica Rodriguez [a], Kanchana R. Kildegaard [a], Mingji Li [a], Irina Borodina [a], Jens Nielsen [a,b,*]

[a] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle allé 6, 2970 Hørsholm, Denmark
[b] Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE412 96 Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 18 February 2015
Received in revised form
10 May 2015
Accepted 7 August 2015
Available online 18 August 2015

Keywords:
p-Coumaric acid
Saccharomyces cerevisiae
Shikimate kinase

ABSTRACT

Aromatic amino acids are precursors of numerous plant secondary metabolites with diverse biological functions. Many of these secondary metabolites are already being used as active pharmaceutical or nutraceutical ingredients, and there are numerous exploratory studies of other compounds with promising applications. *p*-Coumaric acid is derived from aromatic amino acids and, besides being a valuable chemical building block, it serves as precursor for biosynthesis of many secondary metabolites, such as polyphenols, flavonoids, and some polyketides.

Here we developed a *p*-coumaric acid-overproducing *Saccharomyces cerevisiae* platform strain. First, we reduced by-product formation by knocking out phenylpyruvate decarboxylase *ARO10* and pyruvate decarboxylase *PDC5*. Second, different versions of feedback-resistant DAHP synthase and chorismate mutase were overexpressed. Finally, we identified shikimate kinase as another important flux-controlling step in the aromatic amino acid pathway by overexpressing enzymes from *Escherichia coli*, homologous to the pentafunctional enzyme Aro1p and to the bifunctional chorismate synthase-flavin reductase Aro2p. The highest titer of *p*-coumaric acid of 1.93 ± 0.26 g L^{-1} was obtained, when overexpressing tyrosine ammonia-lyase *TAL* from *Flavobacterium johnsoniaeu*, DAHP synthase $ARO4^{K229L}$, chorismate mutase $ARO7^{G141S}$ and *E. coli* shikimate kinase II (*aroL*) in *Δpdc5Δaro10* strain background. To our knowledge this is the highest reported titer of an aromatic compound produced by yeast.

The developed *S. cerevisiae* strain represents an attractive platform host for production of *p*-coumaric-acid derived secondary metabolites, such as flavonoids, polyphenols, and polyketides.

1. Introduction

Aromatic amino acids are precursors of many secondary metabolites produced in plants, where they have a key role in the plant development, adaptation and defense mechanisms (Maeda and Dudareva, 2012; Perez-Gregorio et al., 2014). Among the plant secondary metabolites many alkaloids, flavonoids, tannins and lignins find applications as nutraceutical and pharmaceutical ingredients (Scotti, 2012). Exploratory research of this type of compounds shows promising results (Winkel-Shirley, 2001; Hawkins and Smolke, 2008; Bhan et al., 2013; Leonard et al., 2009), but the limiting factor for a wider use of these secondary metabolites is the lack of efficient extraction systems from plants, or a competent microbial biosynthetic alternative that can produce these compounds in high yields (Santos et al., 2011). There is therefore much interest in developing a microbial cell factory platform that can be used for production of secondary metabolites derived from aromatic amino acids. As the biosynthesis of many plant secondary metabolites involves P450 enzymes, which are often difficult to express in bacteria, the yeast *Saccharomyces cerevisiae* is well suited as a cell factory platform for production of these products. Since many of the flavonoids are intended for nutraceutical applications, using *S. cerevisiae* as the host may be a further advantage due to the long history of its application in food and beverage production (Krivoruchko et al.,

* Corresponding author at: Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE412 96 Gothenburg, Sweden.
E-mail address: nielsenj@chalmers.se (J. Nielsen).

http://dx.doi.org/10.1016/j.ymben.2015.08.003
1096-7176/

2011; Siddiqui et al., 2012) Furthermore, *S. cerevisiae* has proven well amenable for genetic engineering and industrial-scale fermentation and is currently used for production of active pharmaceutical ingredients, dietary supplements, chemicals and fuels (Hong and Nielsen, 2012; Nielsen et al., 2013; Borodina and Nielsen, 2014; Li and Borodina, 2015; Borodina et al., 2015).

In order to enable efficient production of secondary metabolites derived from aromatic amino acids it is necessary to optimize aromatic amino acid biosynthesis. *S. cerevisiae* has been metabolically engineered for improved production of aromatic amino acids through the elimination of feedback inhibition of key enzymes and elimination of by-product formation. Luttik et al. (2008) explored the use of mutated DAHP synthase *ARO4* and chorismate mutase *ARO7* to avoid feedback inhibition and got an increment of 200-fold of aromatic compounds in comparison to the reference strain. Furthermore, through expression of heterologous pathways, it has been possible to produce secondary metabolites derived from aromatic amino acids. Thus, Koopman et al. (2012) produced the flavonoid naringenin from glucose using a background strain with a triple knockout of the most active phenylpyruvate decarboxylases, which prevented formation of the by-product phenylethanol, and overexpressing feedback-resistant DAHP synthase and chorismate mutase.

Escherichia coli and *S. cerevisiae* have been extensively engineered in order to obtain flavonoid-producing cell factories. Among the flavonoids produced in *E. coli* from glucose are naringenin, pinocembrin, and kaempferol 3-O-rhamnoside, with titers of 84 mg L^{-1}, 40 mg L^{-1}, and 57 mg L^{-1}, respectively (Santos et al., 2011; Wu et al., 2014; Yang et al., 2014). Several other flavonoids were produced by supplementing the broth with intermediate compounds, e.g., pinocembrin (429 mg L^{-1}), naringenin (119 mg L^{-1}), eriodictyol (52 mg L^{-1}), quercetin (23.78 mg L^{-1}), and resveratrol (2.3 g L^{-1}) (Leonard et al., 2007; Pandey and Sohng, 2013; Lim et al., 2011). Flavonoids, such as resveratrol (0.31 mg L^{-1}), genistein (7.7 mg L^{-1}), kaempferol (4.6 mg L^{-1}), and quercetin (0.38 mg L^{-1}), have been produced by engineered *S. cerevisiae*, when supplemented with naringenin (Trantas et al., 2009). Using glucose as carbon source, Koopman et al. (2012) produced 102 mg L^{-1} of naringenin. Jendresen et al. (2015) has reported several novel highly active tyrosine ammonia-lyases and shown their activity in *E. coli*, *Lactococcus lactis* and *S. cerevisiae*. These examples prove the potential of *E. coli* and *S. cerevisiae* for flavonoids production, but also evidence the need for development of a platform strain capable of high-level production of aromatic metabolites, e.g. *p*-Coumaric acid from which many secondary metabolites are derived (Santos et al., 2011). We therefore here developed a *S. cerevisiae* strain that overproduces *p*-coumaric acid, and besides representing a starting point for further development of a process for commercial *p*-coumaric acid, we believe this strain can be used as a platform strain for production of flavonoids and other coumaric-acid derived secondary metabolites.

2. Materials and methods

2.1. Plasmids and strains construction

The background strain for this research was *S. cerevisiae* CEN.PK102-5B. Cloning was carried out using *E. coli* strain DH5α. All the fragments used for overexpression of genes were amplified by PCR using primers and templates as described in Supplementary Tables 1 and 2. The fragment encoding chorismate mutase from *C. guilliermondii* was identified through BLAST search, by comparing the full amino acids sequence of Aro7p from *S. cerevisiae* against *C. guilliermondii* protein sequences in GenBank. Homology of 60% was found between Aro7p from *S. cerevisiae* and the hypothetical protein PGUG_00476 from *C. guilliermondii* (Genbank accession number: XM_001487049.1).

Tyrosine ammonia-lyase *TAL* from *F. johnsoniaeu* was as described before (Jendresen et al., 2015). The amplified products were cloned along with strong constitutive promoters into Easy-Clone integrative plasmids by USER cloning (Jensen et al., 2014). The clones with correct inserts were confirmed by sequencing. The list of the constructed vectors can be found in Table 1 and the details on the cloning are given in Supplementary Table 3.

Transformation of yeast cells was carried out by the lithium acetate method (Gietz et al., 2002). The strains were selected on synthetic drop-out medium (Sigma-Aldrich), selecting for *URA*, *HIS* and *LEU* markers. The yeast strains constructed in this study are listed in Table 1.

The mutant genes, sc*ARO4*^{fbr} (K229L), sc*ARO7*^{fbr} (G141S), ec*aroG*^{fbr1} (L175D) and ec*aroG*^{fbr2} (S180F), were constructed by site-directed mutagenesis method (Zheng et al., 2004), using primers and templates described in Supplementary Tables 1 and 3. The *ARO4*, *ARO7* and *aroG* wild-type genes were amplified from the genomic DNA of *S. cerevisiae* and *E. coli* NST 74. The DNA fragments were gel-purified and cloned into vector pESC-*URA*-ccdB-USER or pESC-*HIS*-ccdB-USER, the derived plasmids pCfB761, pCfB775, pCfB1075 and pCfB1076 were confirmed by DNA sequencing. These plasmids were used as the templates for site-directed mutagenesis reactions. The complementary primers with nucleotide substitutions for mutagenesis were designed for each mutation according to the guidelines stated by Zheng et al., 2004. The reactions were incubated with DpnI for 1 h before transformation into competent DH5α *E. coli* cells, the strains were grown overnight on LB agar (Amp) plates at 37 °C. Colonies were then selected, and plasmid DNA was extracted and sequenced over the region of the mutation. Successful mutants were verified by DNA sequencing and then re-cloned into the integrative expression vectors (Supplementary Tables 2 and 3).

2.2. Deletions of ARO10 and PDC5

The double knockout strain Δ*aro10*Δ*pdc5* was constructed by an iterative replacement of the targeted genes with the *URA3* cassette in the strain CEN.PK102-5B by bi-partite method (Erdeniz et al., 1997). The knockout fragments were transformed into *S. cerevisiae* and transformants were selected on SC-Ura yeast synthetic drop-out media. The knockouts were confirmed by PCR on genomic DNA preparations. The *URA3* marker was looped-out via direct repeats by growing the yeast on 5-fluoroorotic acid (5-FOA) plates, and the second gene was knocked-out in the same way and the URA3 marker was removed again.

For the single knockout strains (Δ*aro10* and Δ*pdc5*), the target genes were replaced by a *LEU2* cassette in the strain CEN.PK102-5B. The knockout fragments were transformed into *S. cerevisiae* and transformants were selected on SC-Leu yeast synthetic drop-out media. The knockouts were confirmed by PCR on genomic DNA preparations.

The gene fragments, carrying the upstream and downstream fragments of the marker cassettes *URA3* and *LEU2* and of the targeted genes *ARO10* and *PDC5*, were generated by PCR amplification using the method developed by Reid et al. (2002). Primers and templates used for targeting the genes are indicated in Supplementary Tables 1 and 2. The upstream fragments of the targeted genes (*PDC5*_UP or *ARO10*_UP), the downstream fragments of the targeted genes (*PDC5*_DOWN or *ARO10*_DOWN), the upstream and downstream fragment of the markers (2/3_*URA3*_UP, 2/3_*URA3*_DOWN, for the double knockout strain and 2/3_*LEU2*_UP and 2/3_*LEU2*_DOWN for the single knockout strain construction) from *Kluyveromyces lactis* were amplified using primers described in Supplementary Table 1. To generate a

Table 1
Plasmids and strains used in this study.

Plasmid ID	Genotype		Source
Parental plasmids			
pCfB0054	Episomal replication vector, pESC, *HIS*, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}		Jensen et al. (2014)
pCfB0055	Episomal replication vector, pESC, *HIS*, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}		Jensen et al. (2014)
pCfB255	Integrative plasmid, pX-2-loxP, *klURA3*, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}		Jensen et al. (2014)
pCfB257	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-T_{ADH1}, P_{PGK1}-P_{PGK1}		Jensen et al. (2014)
pCfB258	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}		Jensen et al. (2014)
Plasmids used for directed mutagenesis			
pCfB744	Episomal replication vector, pESC, *URA*, P_{TEF1}-*scARO7*-T_{ADH1}		This study
pCfB745	Episomal replication vector, pESC, *HIS*, P_{TEF1}-*scARO4*-T_{ADH1}		This study
pCfB761	Episomal replication vector, pESC, *URA*, P_{TEF1}-*scARO7*^{fbr}-T_{ADH1}		This study
pCfB775	Episomal replication vector, pESC, *HIS*, P_{TEF1}-*scARO4*^{fbr}-T_{ADH1}		This study
PCfB1074	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*ecAroG*-T_{ADH1}		This study
PCfB1075	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}- *ecaroG*^{fbr1}-T_{ADH1}		This study
PCfB1076	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}- *ecaroG*^{fbr2}-T_{ADH1}		This study
Integrative plasmids			
pCfB826	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*scARO7*^{fbr}-T_{ADH1}, P_{PGK1}-*scARO4*^{fbr}-T_{CYC1}		This study
pCfB827	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*scARO7*^{fbr}-T_{ADH1}, P_{PGK1}-*ecaroF*^{fbr}-T_{CYC1}		This study
pCfB830	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*cgARO7*-T_{ADH1}, P_{PGK1}-*scARO4*^{fbr}-T_{CYC1}		This study
pCfB831	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*cgARO7*-T_{ADH1}, P_{PGK1}-*ecaroF*^{fbr}-T_{CYC1}		This study
pCfB1077	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*scARO7*^{fbr}-T_{ADH1}, P_{PGK1}-*ecaroG*^{fbr1}-T_{CYC1}		This study
pCfB1078	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*cgARO7*-T_{ADH1}, P_{PGK1}-*ecaroG*^{fbr1}-T_{CYC1}		This study
pCfB1080	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*scARO7*^{fbr}-T_{ADH1}, P_{PGK1}-*ecaroG*^{fbr2}-T_{CYC1}		This study
pCfB1081	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*cgARO7*-T_{ADH1}, P_{PGK1}-*ecaroG*^{fbr2}-T_{CYC1}		This study
pCfB1221	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*scTYR1*-T_{ADH1}		This study
pCfB1226	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*ectyrA*^{fbr}-T_{ADH1}, P_{PGK1}-*scARO4*^{fbr}-T_{CYC1}		This study
pCfB1227	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*ectyrA*^{fbr}-T_{ADH1}, P_{PGK1}- *ecaroF*^{fbr}-T_{CYC1}		This study
pCfB1228	Integrative plasmid, pX-4-loxP, *spHIS5*, P_{TEF1}-*ectyrA*^{fbr}-T_{ADH1}, P_{PGK1}-*ecaroG*^{fbr1}-T_{CYC1}		This study
pCfB1964	Integrative plasmid, pX-2-loxP, Kl*URA3*, P_{TEF1}-fjTAL-T_{ADH1}		This study
pCfB2733	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}- *scARO1*-T_{ADH1}, P_{PGK1}- *scARO2*-T_{CYC1}		This study
pCfB2739	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*ecaroB*-T_{ADH1}		This study
pCfB2741	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*ecaroE*-T_{ADH1}		This study
pCfB2742	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*ecydiB*-T_{ADH1}		This study
pCfB2743	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{PGK1}-*ecaroK*-T_{CYC1}		This study
pCfB2745	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*ecaroA*-T_{ADH1}		This study
pCfB2746	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{PGK1}-*ecaroD*-T_{CYC1}		This study
pCfB2747	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{PGK1}-*ecaroL*-T_{CYC1}		This study
pCfB2749	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{PGK1}-*ecaroC*-T_{CYC1}		This study
pCfB2740	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{TEF1}-*scARO1*-T_{ADH1}		This study
pCfB2748	Integrative plasmid, pX-3-loxP, *klLEU2*, P_{PGK1}-*scARO2*-T_{CYC1}		This study
Parent and template strains			
Strain ID	**Strain**		**Source**
ST10	*S. cerevisiae* CEN.PK102-5B (MATa *ura3-52 his3 Δ 1 leu2-3/112* MAL2-8c SUC2)		Peter Kötter
ST700	*E. coli* NST 74 (ATCC 31884)		ATCC
ST679	*C. guilliermondii* (ATCC 6260)		ATCC
Knockout strains			
Strain ID	**Parent strain**	**Characteristics**	**Source**
ST4034	ST10	MATa *aro10Δ: LEU2*	This study
ST3532	ST10	MATa *pdc5Δ: LEU2*	This study
ST691	ST10	MATa *aro10Δ pdc5Δ*	This study
Strains transformed with integrative plasmids			
Strain ID	**Parent strain**	**Integrated plasmids**	**Source**
ST4068	ST10	pCfB1964, pCfB257, pCfB258	This study
ST4069	ST10	pCfB255, pCfB257, pCfB258	This study

Table 1 (*continued*)

Plasmid ID	Genotype		Source
ST4072	ST4034	pCfB1964, pCfB258	This study
ST4073	ST4034	pCfB255, pCfB258	This study
ST4070	ST3532	pCfB1964	This study
ST4071	ST3532	pCfB255	This study
ST4048	ST691	pCfB1964, pCfB257, pCfB258	This study
ST4050	ST691	pCfB255, pCfB257, pCfB258	This study
ST2645	ST691	pCfB1964, pCfB257, pCfB826	This study
ST4040	ST691	pCfB1964, pCfB1221, pCfB826	This study
ST4049	ST691	pCfB1964, pCfB257, pCfB830	This study
ST4044	ST691	pCfB1964, pCfB1221, pCfB830	This study
ST4038	ST691	pCfB1964, pCfB257, pCfB1226	This study
ST4041	ST691	pCfB1964, pCfB1221, pCfB827	This study
ST4051	ST691	pCfB1964, pCfB257, pCfB827	This study
ST4045	ST691	pCfB1964, pCfB1221, pCfB831	This study
ST4052	ST691	pCfB1964, pCfB257, pCfB831	This study
ST4037	ST691	pCfB1964, pCfB257, pCfB1227	This study
ST4053	ST691	pCfB1964, pCfB257, pCfB1077	This study
ST4042	ST691	pCfB1964, pCfB1221, pCfB1077	This study
ST4054	ST691	pCfB1964, pCfB257, pCfB1078	This study
ST4046	ST691	pCfB1964, pCfB1221, pCfB1078	This study
ST4039	ST691	pCfB1964, pCfB257, pCfB1228	This study
ST4055	ST691	pCfB1964, pCfB257, pCfB1080	This study
ST4043	ST691	pCfB1964, pCfB1221, pCfB1080	This study
ST4056	ST691	pCfB1964, pCfB257, pCfB1081	This study
ST4047	ST691	pCfB1964, pCfB1221, pCfB1081	This study
ST3213	ST691	pCfB1964, pCfB826	This study
ST4057	ST3213	pCfB2739	This study
ST4065	ST3213	pCfB2746	This study
ST4062	ST3213	pCfB2741	This study
ST4063	ST3213	pCfB2742	This study
ST4066	ST3213	pCfB2743	This study
ST4058	ST3213	pCfB2747	This study
ST4064	ST3213	pCfB2745	This study
ST4060	ST3213	pCfB2749	This study
ST4067	ST3213	pCfB2733	This study
ST4061	ST3213	pCfB2740	This study
ST4059	ST3213	pCfB2748	This study

complete gene targeting substrate, the upstream fragments (*PDC5*_UP or *ARO10*_UP), were fused to the 2/3 upstream fragment of the markers (2/3_*URA3*_UP for the double knockout and 2/3_*LEU2*_UP for the single knockout), in the same way, the downstream fragments were fused to the downstream fragment of the markers (2/3_*URA3*_DOWN for the double knockout and 2/3_*LEU2*_DOWN for the single knockout). The two fusion PCR fragments per targeted gene were transformed simultaneously into the *S. cerevisiae* strain and selected in SC-Ura or Sc-Leu medium according to the selection marker. The correct transformants were confirmed by PCR, using primers described in Supplementary Table 1.

2.3. Media and cultivations

Synthetic complete (SC) medium as well as drop-out media (SC-Ura, SC-Leu, SC-His) and agar plates were prepared using premixed drop-out powders from Sigma-Aldrich. Synthetic fed-batch medium for *S. cerevisiae* M-Sc.syn-1000 (FIT) was purchased from M2P labs GmbH (Germany). The medium was supplemented with the supplied vitamins solution (final 1% v/v) and the enzyme mix (final concentration 0.5% v/v) immediately prior to use.

At least six single colonies originating from independent transformants were inoculated in 0.5 ml drop-out SC liquid medium without uracil, histidine, and/or leucine in 96-deep well

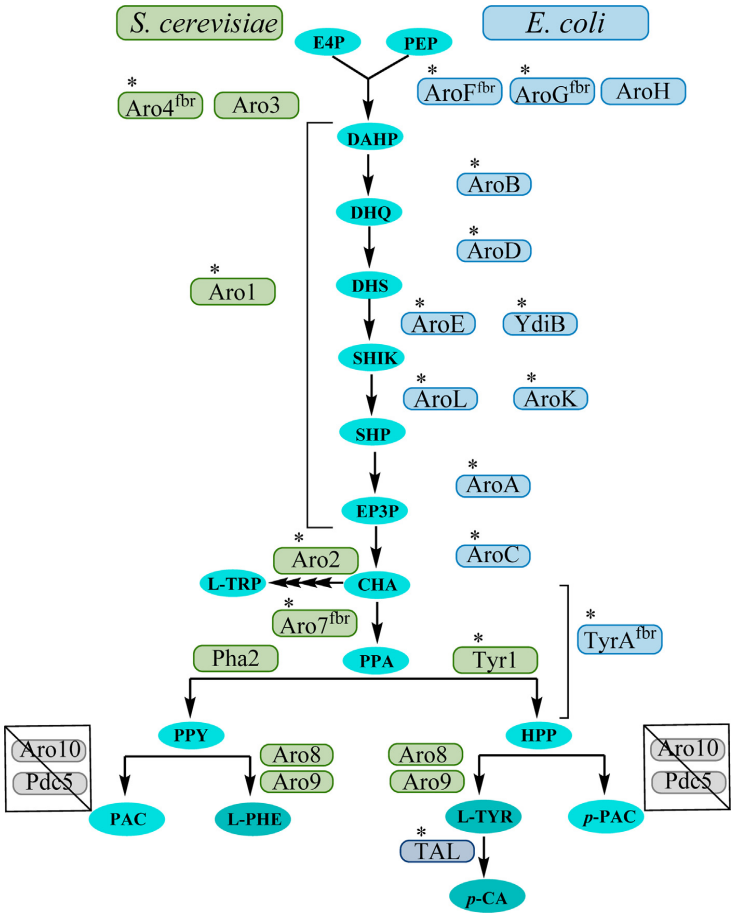


Fig. 1. Schematic representation of the engineered *p*-coumaric acid production pathway in *S. cerevisiae*. E4P: erythrose 4-phosphate, PEP: phosphoenolpyruvate, DAHP: 3-deoxy-D-arabino-heptulosonic acid 7-phosphate, DHQ: 3-dehydroquinate, DHS: 3-dehydro-shikimate, SHIK: shikimate, SHP: shikimate-3-phosphate, EP3P: 5-enolpyruvylshikimate-3-phosphate, PPA: prephenate, PPY: phenylpyruvate, HPP: para-hydroxy-phenylpyruvate, PAC: phenylacetaldehyde, pPAC: para-hydroxy-acetaldehyde, L-PHE: L-phenylalanine, L-TYR: L-tyrosine, *p*-CA: p-coumaric acid, TAL: tyrosine ammonia-lyase. The star key indicates overexpressed enzymes; enzymes in gray boxes represent knockouts, "fbr" indicates feedback-resistant.

microtiter plates with air-penetrable lid (EnzyScreen, NL). The plates were incubated at 30 °C with 250 rpm agitation at 5 cm orbit cast overnight. 50 μl of the overnight cultures were used to inoculate 0.5 ml synthetic fed-batch medium in a 96-deep well plate. Fermentation was carried out for 72 h at the same conditions as above.

At the end of the cultivation OD_{600} was measured as following: 10 μl of the sample was mixed with 190 μl water and absorbance was measured at 600 nm wavelength in microplate reader BioTek Synergy MX (BioTek). The culture broth was spun down and the supernatant was analyzed for *p*-coumaric acid concentration using HPLC.

2.4. Quantification of p-coumaric acid

Quantification of *p*-coumaric acid was performed on HPLC (Thermo), equipped with a Discovery HS F5 150 mm × 2.1 mm column (particle size 3 μm). Samples were analyzed using a gradient method with two solvents: 10 mM ammonium formate pH 3.0 (A) and acetonitrile (B) at 1.5 ml min^{-1}. The program started with 5% of solvent B (0–0.5 min), after which its fraction was increased linearly from 5% to 60% (0.5–7.0 min) and maintained at 60% for 2.5 min (7.0–9.5 min). Then the fraction of solvent B was decreased back to 5% (9.5–9.6 min) and remained at 5% until the end (9.6–12 min). *p*-Coumaric acid was detected by absorbance at 277 nm and the peak (retention time 4.7 min) area was integrated with Chromeleon 7 and used for quantification by fitting with a standard curve. For all the strains at least three biological replicates were analyzed.

3. Results

3.1. Deletion of phenylpyruvate and pyruvate decarboxylases

In order to avoid production of aromatic alcohols and direct the pathway flux to aromatic amino acids, we performed single knockouts of *ARO10* (phenylpyruvate decarboxylase), *PDC5* (pyruvate decarboxylase), and a double knock out of *ARO10* and *PDC5*.

Furthermore, a tyrosine ammonia-lyase *TAL* from *F. johnsoniae* was overexpressed in these strains in order to produce *p*-coumaric acid (Jensen et al., 2014) (Fig. 1).

The reference strain without deletions was able to produce 0.24 ± 0.03 g L^{-1} of *p*-coumaric acid (Fig. 2). The strain with the single deletion of *PDC5* produced 0.30 ± 0.09 g L^{-1} of *p*-coumaric acid, while the strain carrying the knockout of *ARO10* produced 0.29 ± 0.04 g L^{-1}. The highest production (0.55 ± 0.13 g L^{-1}) was obtained in the strain with the double knockout of *ARO10* and *PDC5*.

3.2. Effect of the elimination of the feedback inhibition of DAHP synthase and chorismate mutase on p-coumaric acid production

The enzymes DAHP synthase and chorismate mutase from the aromatic amino acids pathway are feedback-inhibited by L-tyrosine and L-phenylalanine (Hartmann et al., 2003; Luttik et al., 2008). In order to enhance the activity of these enzymes, we overexpressed feedback-resistant variants of DAHP synthase and chorismate mutase in the *Δaro10Δpdc5* strain. For this we selected 4 variants of DAHP synthase, a mutated feedback-resistant $ARO4^{K229L}$ from *S. cerevisiae*, an *aroF* from *E. coli* NST 74 (ATCC 31884) and two mutant variants of *aroG*, also from *E. coli*, which were constructed through replacement of the residues L175D and S180F of the hydrophobic domain. For chorismate mutase, there were selected a mutated feedback-resistant $ARO7^{G141S}$ from *S. cerevisiae*, a naturally feedback-resistant chorismate mutase from *C. guilliermonii* and *tyrA* from the *E. coli* strain NST 74 (ATCC 31884) (Tribe, 1987). The chorismate mutase from *C. guilliermondii* had been reported as non-feedback inhibited (Bode and Birnbaum, 1991). The variants of chorismate mutase from *S. cerevisiae* and *C. guilliermondii* were complemented with prephenate dehydrogenase *TYR1* from *S. cerevisiae* in order to get equivalent overexpressions to the bifunctional chorismate mutase-prephenate dehydrogenase TyrA from *E. coli* NST 74 (ATCC 31884). The mutations were selected from previous studies, where they had been reported as feedback-insensitive mutations: the mutations in *ARO4* and *ARO7* were reported by Luttik et al. (2008), the mutation of *aroG* (L175D) was reported by Hu et al. (2003) and the mutation of *aroG* (S180F) was reported by Ger et al. (1994).

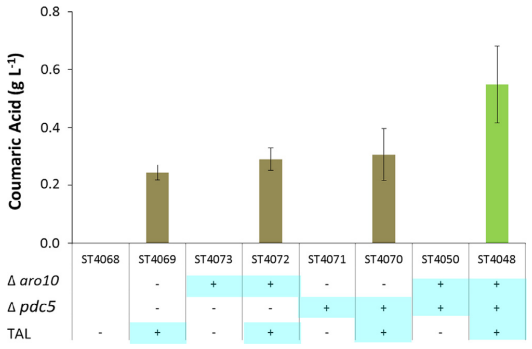


Fig. 2. Effect of knockouts of *ARO10* and *PDC5* on *p*-coumaric acid production in strains overexpressing tyrosine-ammonia lyase *FjTAL* from *F. johnsoniaeu*.

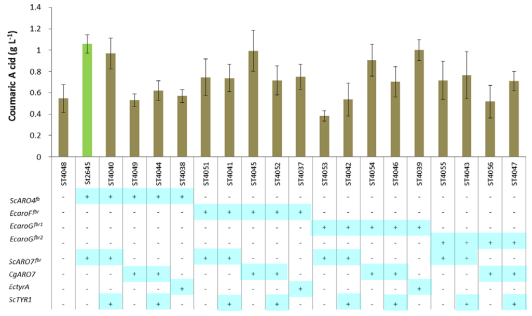


Fig. 3. Production of *p*-coumaric acid upon overexpression of feedback inhibition-resistant DAHP synthase and chorismate mutase in strains with overexpression of *FjTAL* and deletion of *PDC5* and *ARO10*.

All the strains overexpressing DAHP synthase and chorismate mutase were evaluated for their ability to produce *p*-coumaric acid and a two-way ANOVA was conducted to analyze the effect of DAHP synthases and chorismate mutases. The *p*-coumaric acid production was normally distributed for all the combinations of DAHP synthases and chorismate mutases as assessed by a Shapiro–Wilk's test (p-Value > 0.05). The results showed that the overexpression of chorismate mutase alone did not have a significant effect on the production of *p*-coumaric acid (p-Value 0.399), while the overexpression of DAHP synthase and the combined overexpression of DAHP synthase and chorismate mutase had a significant effect on the production of the compound (p-Values 0.010 and 0.0005 correspondingly) (Supplementary Tables 4 and 5).

From the strains overexpressing $scARO4^{fbr}$, the best combination was obtained, when overexpressing at the same time $scARO7^{fbr}$, for the strains overexpressing *ecaroF* the best producer was obtained in combination with $cgARO7^{fbr}$. Although the two residues replaced in the *aroG* strains are located in the same region, the *p*-coumaric acid production after the overexpression of the mutated AroG enzymes was different. The replacement L175D seems to generate a more active AroG, since two of the strains carrying this mutation ($ecaroG^{fbr1}$-*ectyrA* and $ecaroG^{fbr1}$-$cgARO7$) had production of over 0.9 g L^{-1} in contrast to the strains with the replacement S180F, where the titer did not exceed 0.8 g L^{-1}. It was not possible to see a general trend of the effect of Tyr1p in the production of *p*-coumaric acid, but in 4 of the 8 strains overexpressing this enzyme a negative effect was observed (Fig. 3).

The overexpression of DAHP synthase and chorismate mutase mostly had a positive effect on production of *p*-coumaric acid, however there were some exceptions: *scARO4*-*cgARO7*, *scARO4*-*ectyrA*, *ecaroG*-*ARO7* and *ecaroG*-*scARO7*-*scTYR1* all resulted in the same or lower titer than in the reference strain.

3.3. p-Coumaric acid production after overexpression of ARO1 and ARO2 from S. cerevisiae and their analogous from E. coli

In *S. cerevisiae*, the five steps to synthetize the aromatic intermediate compound 5-enolpyruvylshikimate-3-phosphate (EPSP) from DAHP are catalyzed by the pentafunctional enzyme Aro1p, while in other organisms such as plants and bacteria each step is performed by monofunctional enzymes (Fig. 1). In order to find flux-controlling steps in this common branch of the aromatic amino acid pathway, analogous enzymes to Aro1p and Aro2p from *E. coli* were overexpressed in *S. cerevisiae*. Strains overexpressing *ARO1* and *ARO2* from *S. cerevisiae* were also constructed with the purpose of making a comparative analysis between the strains overexpressing enzymes from *E. coli* and *S. cerevisiae*. The background strain for this experiment was the strain ST3213 (*aro10Δpdc5Δ* $ARO4^{K229L}$ $ARO7^{G141S}$). The control strain for this experiment was the strain ST3213 transformed with the empty integrative plasmid pCfB257 instead of plasmids carrying overexpression cassettes for *ARO1*, *ARO2* or their analogous.

The overexpression of the monofunctional enzymes from *E. coli* had a positive effect on the *p*-coumaric acid production; the only exception was the overexpression of shikimate kinase AroK (Fig. 4A).

The strains overexpressing dehydroquinate synthase (AroB), shikimate dehydrogenase (YdiB), EPSP synthase (AroA) or shikimate kinase (AroL) produced more than 1.6 g L^{-1} of *p*-coumaric acid. The strain with the highest improvement was the strain overexpressing the isoenzyme of shikimate kinase AroL, producing 1.93 ± 0.26 g L^{-1} of *p*-coumaric acid.

Overexpression of the pentafunctional enzyme Aro1p and the bifunctional chorismate synthase-flavin reductase Aro2p from *S. cerevisiae* had a positive effect in the *p*-coumaric acid production. The strain overexpressing Aro1p produced 1.68 ± 0.19 g L^{-1} of *p*-coumaric acid, the strain overexpressing Aro2p produced 1.40 ± 0.12 g L^{-1}, and the simultaneous overexpression of Aro1p and Aro2p increased the production of *p*-coumaric acid to 1.71 ± 0.12 g L^{-1}. The production of *p*-coumaric acid was very similar between the strain overexpressing Aro1p and the strain overexpressing Aro1p and Aro2p simultaneously. None of the strains expressing native versions of Aro1p and Aro2p produced more *p*-coumaric acid than the strain overexpressing AroL from *E. coli* (Fig. 4B).

4. Discussion

This study describes engineering of *S. cerevisiae* for production of *p*-coumaric acid from glucose, leading to a final production titer of 1.93 ± 0.26 g L^{-1} on feed-in-time medium in deep-well plates, which represents a 7.9-fold improvement in comparison to the non-optimized strain. The optimized strain also produced 1.89 g L^{-1} *p*-coumaric acid in controlled fed-batch fermentation on mineral medium (Supplementary Fig. 1). To the best of our knowledge, this is the highest titer of *de novo* production of an aromatic compound reported for *S. cerevisiae*. There are studies of flavonoids production from glucose in *E. coli* and *S. cerevisiae*;

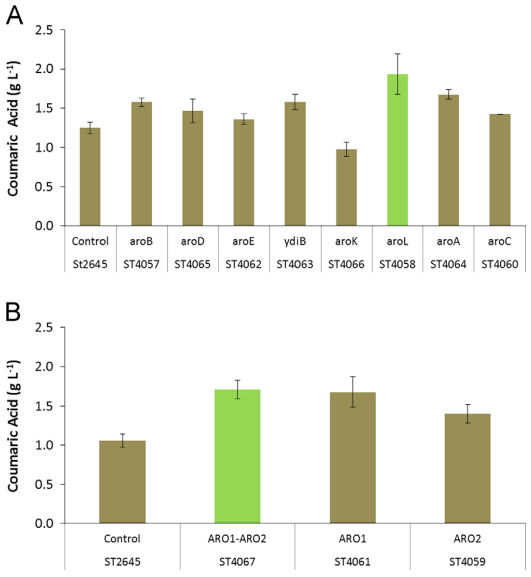


Fig. 4. Production of *p*-coumaric acid upon overexpression of *ARO1* and *ARO2* from *S. cerevisiae* and their analogous from *E. coli* along with overexpression of $ARO4^{K229L}$, $ARO7^{G141S}$, *FjTAL* and knockouts of *ARO10* and *PDC5*. (A) Genes from *E. coli*; *aroB*: 3-dehydroquinate synthase, *aroD*: 3-dehydroquinate dehydratase, *aroE*: shikimate dehydrogenase, *ydiB*: shikimate dehydrogenase – quinate dehydrogenase, *arok*: shikimate kinase I, *aroL*: shikimate kinase II, *aroA*: EPSP synthase, *aroC*: chorismate synthase. (B) Genes from *S. cerevisiae; ARO1*: pentafunctional enzyme, *ARO2*: bifunctional chorismate synthase and flavin reductase. Control strain ST2645 (*FjTAL, aro10Δpdc5Δ*, $ARO4^{K229L}$, $ARO7^{G141S}$).

however none of them has *p*-coumaric acid as the final product. Some studies reported the accumulation of *p*-coumaric acid in parallel to the production of other aromatic compounds. In *E. coli*, Santos et al. (2011) and Wu et al. (2014) reported accumulation of *p*-coumaric acid to 79 and 70 mg L^{-1}, respectively, when they were producing 84 and 101 mg L^{-1} of naringenin. In *S. cerevisiae*, Koopman et al. (2012) reported around 61 mg L^{-1} of accumulated *p*-coumaric acid, when they were producing 65 mg L^{-1} of naringenin in bioreactors.

The single knockouts of *PDC5* and *ARO10* had a positive effect on the production *p*-coumaric acid. The PDC5 strain was auxotrophic for histidine, and the effect of the PDC5 knockout in comparison to ARO10 may be different in a prototrophic strain. The simultaneous deletions lead a 2-folds improvement in comparison to the reference strain (Fig. 2).This was expected as deletion of the two genes had been previously reported to improve production of some other tyrosine- and phenylalanine-derived compounds, i.e., naringenin (Koopman et al., 2012)

The results obtained from the 2-factor Anova showed that overexpression of chorismate mutase alone did not have a statistically significant effect on production of *p*-coumaric acid. This is consistent with results obtained by Luttik et al. (2008), where overexpression of a native or a feedback resistant Aro7 did not have significant impact on the production of the aromatic fusel alcohols, unless Aro4 was overexpressed as well.

Besides Aro4p and Aro7p that have been widely reported as feedback controlled enzymes, there must be other enzymes limiting the production of aromatic amino acids. To this end, we showed that the activity of the pentafunctional enzyme Aro1p in *S. cerevisiae* is also limiting for the biosynthesis of chorismate. With exception of the strain overexpressing AroK, all the strains overexpressing the Aro1p and Aro2p analogous enzymes from *E. coli* had a higher production of *p*-coumaric acid (Fig. 4A). These results show that these enzymes also exhibit flux control and further improvements can be obtained through the overexpression of heterologous enzymes with higher activity than Aro1p and Aro2p from *S. cerevisiae*.

The strain overexpressing AroL from *E. coli* had the highest production of *p*-coumaric acid, indicating that conversion of shikimate to shikimate-3-phosphate has the highest flux control of these five steps. These results are supported by the observation from Luttik et al. (2008), where they reported accumulation of shikimate in the culture supernatants after overexpression of $scAro4^{fbr}$ in *S. cerevisiae*. Previous studies in *E. coli* also reported increased flux of intermediate compounds to L-tyrosine, when shikimate kinase II was overexpressed (Takai et al., 2005; Juminaga et al., 2012).

The overexpression of AroK did not increase the production of *p*-coumaric acid in the constructed strain. This may be due to the low affinity of AroK to shikimate, i.e. the K_m for shikimate of AroK is more than 20 mM whereas the K_m of AroL is only 0.2 mM (DeFeyter and Pittard, 1986). The contribution of AroK to the shikimate kinase activity in *E. coli* is therefore minimal (DeFeyter and Pittard, 1986) and its overexpression did not have a significant effect in *p*-coumaric acid production in *E. coli*. Previously there was not detected any feedback inhibition of AroL or AroK by aromatic amino acids, chorismic acid or prephenic acid (DeFeyter and Pittard, 1986). Overexpression of the pentafunctional protein (Aro1p) and the bifunctional chorismate synthase-flavin reductase (Aro2p) from *S. cerevisiae* had a positive effect on the production of *p*-coumaric acid, though the titer was still lower than of the strain overexpressing AroL alone.

The *p*-coumaric acid overproducing strain that we describe can be further engineered to include overproduction of malonyl-CoA, a common precursor for biosynthesis of polyphenols and flavonoids. It has been previously shown that increasing malonyl-CoA supply improves production of naringenin (Koopman et al., 2012),

flavonone (Leonard et al., 2007), and 7-O-methyl aromadendrin (Malla et al., 2012). Increased flux towards malonyl-CoA can be achieved by overexpression of deregulated acetyl-CoA carboxylase (Shi et al., 2014) and by further increase of acetyl-CoA biosynthesis as described previously (Krivoruchko et al., 2015).

In conclusion, we describe metabolic engineering strategies that lead towards a platform yeast strain, producing high levels of *p*-coumaric acid, which besides being a product of commercial interest by itself, also serves as an intermediate compound for aromatic secondary metabolites. We also demonstrate that heterologous expression of tyrosine-ammonia lyase *TAL* is well suitable for evaluation of metabolic engineering targets for improving the flux through the aromatic amino acid biosynthetic pathway. Through combination of several different strategies we improved the production of *p*-coumaric acid 7.9-fold, and we are therefore confident that our strain represents a good platform stain for production of *p*-coumaric acid derived secondary metabolites by *S. cerevisiae.*

Acknowledgments

This work was financed by the Novo Nordisk Foundation. We thank Christian Bille Jendresen and Mette Kristensen for assistance with HPLC analytical method. We would also like to acknowledge Tobias Klein, Konstantin Schneider, Inger Rosenstand and Gheorghe Manuel Borja Zamfir for their help with fermentations in bioreactors.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2015.08.003.

References

Bhan, N., Xu, P., Koffas, M. a G., 2013. Pathway and protein engineering approaches to produce novel and commodity small molecules. Curr. Opin. Biotechnol. 24, 1137–1143. http://dx.doi.org/10.1016/j.copbio.2013.02.019.

Bode, R., Birnbaum, D., 1991. Regulation of chorismate mutase activity of various yeast species by aromatic amino acids. Antonie van Leeuwenhoek 59, 9–13. http://dx.doi.org/10.1007/BF00582113.

Borodina, I., Nielsen, J., 2014. Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. Biotechnol. J. 9, 609–620. http://dx.doi.org/10.1002/biot.201300445.

Borodina, I., Kildegaard, K.R., Jensen, N.B., Blicher, T.H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M.J., Rosenstand, I., Öberg, F., Forster, J., Nielsen, J., 2015. Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via β-alanine. Metab. Eng. 27, 57–64. http://dx.doi.org/10.1016/j.ymben.2014.10.003.

DeFeyter, R.C., Pittard, J., 1986. Purification and properties of shikimate kinase II from *Escherichia coli* K-12. J. Bactiol 165, 331–333.

Erdeniz, N., Mortensen, U.H., Rothstein, R., 1997. Cloning-free PCR-based allele replacement methods. Genome Res. 7, 1174–1183. http://dx.doi.org/10.1101/gr.7.12.1174.

Ger, Y.M., Chen, S.L., Chiang, H.J., Shiuan, D., 1994. A single Ser-180 mutation desensitizes feedback inhibition of the phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthetase in *Escherichia coli*. J. Biochem. 116, 986–990.

Gietz, B.R.D., Woods, R. A, 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA / polyethylene glycol method. Methods Enzymol. 350, 87–96.

Hartmann, M., Schneider, T.R., Pfeil, A., Heinrich, G., Lipscomb, W.N., Braus, G.H., 2003. Evolution of feedback-inhibited beta /alpha barrel isoenzymes by gene duplication and a single mutation. Proc. Natl. Acad. Sci. U.S.A 100, 862–867. http://dx.doi.org/10.1073/pnas.0337566100.

Hawkins, K.M., Smolke, C.D., 2008. Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. Nat. Chem. Biol. 4, 564–573. http://dx.doi.org/10.1038/nchembio.105.

Hong, K.-K., Nielsen, J., 2012. Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. Cell. Mol. Life Sci. 69, 2671–-2690. http://dx.doi.org/10.1007/s00018-012-0945-1.

Hu, C., Jiang, P., Xu, J., Wu, Y., Huang, W., 2003. Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of *Escherichia coli*. J. Basic Microbiol. 43, 399–406. http://dx.doi.org/10.1002/jobm.200310244.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., Borodina, I., 2014. EasyClone: method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. FEMS Yeast Res. 14, 238–248. http://dx.doi.org/10.1111/1567-1364.12118.

Jendresen, C.B., Stahlhut, S.G., Li, M., Gaspar, P., Siedler, S., Forster, J., Maury, J., Borodina, I., Nielsen, A.T., 2015. Highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and Saccharomyces cerevisiae. Appl. Environ. Microbiol. 81, 4458–4476. http://dx.doi.org/10.1128/AEM.00405-15.

Juminaga, D., Baidoo, E.E.K., Redding-Johanson, A.M., Batth, T.S., Burd, H., Mukhopadhyay, A., Petzold, C.J., Keasling, J.D., 2012. Modular engineering of L-tyrosine production in *Escherichia coli*. Appl. Environ. Microbiol. 78, 89–98. http://dx.doi.org/10.1128/AEM.06017-11.

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J. a, Pronk, J.T., Daran, J.-M., 2012. De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microb. Cell Fact. 11, 155. http://dx.doi.org/10.1186/1475-2859-11-155.

Krivoruchko, A., Siewers, V., Nielsen, J., 2011. Opportunities for yeast metabolic engineering: lessons from synthetic biology. Biotechnol. J. 6, 262–276. http://dx.doi.org/10.1002/biot.201000308.

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., Nielsen, J., 2015. Microbial acetyl-CoA metabolism and metabolic engineering. Metab. Eng. 28, 28–42. http://dx.doi.org/10.1016/j.ymben.2014.11.009.

Leonard, E., Lim, K.-H., Saw, P.-N., Koffas, M. a G., 2007. Engineering central metabolic pathways for high-level flavonoid production in *Escherichia coli*. Appl. Environ. Microbiol. 73, 3877–3886. http://dx.doi.org/10.1128/AEM.00200-07.

Leonard, E., Runguphan, W., O'Connor, S., Prather, K.J., 2009. Opportunities in metabolic engineering to facilitate scalable alkaloid production. Nat. Chem. Biol. 5, 292–300. http://dx.doi.org/10.1038/nchembio.160.

Li, M., Borodina, I., 2015. Application of synthetic biology for production of chemicals in yeast *Saccharomyces cerevisiae*. FEMS Yeast Res. 15, 1–12. http://dx.doi.org/10.1111/1567-1364.12213.

Lim, C.G., Fowler, Z.L., Hueller, T., Schaffer, S., Koffas, M.A.G., 2011. High-yield resveratrol production in engineered *Escherichia coli*. Appl. Environ. Microbiol. 77, 3451–3460. http://dx.doi.org/10.1128/AEM.02186-10.

Luttik, M.A.H., Vuralhan, Z., Suir, E., Braus, G.H., Pronk, J.T., Daran, J.M., 2008. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. Metab. Eng. 10, 141–153. http://dx.doi.org/10.1016/j.ymben.2008.02.002.

Maeda, H., Dudareva, N., 2012. The shikimate pathway and aromatic amino acid biosynthesis in plants. Annu. Rev. Plant Biol. 63, 73–105. http://dx.doi.org/10.1146/annurev-arplant-042811-105439.

Malla, S., Koffas, M.A G., Kazlauskas, R.J., Kim, B.G., 2012. Production of 7-O-methyl aromadendrin, a medicinally valuable flavonoid, in *Escherichia coli*. Appl. Environ. Microbiol. 78, 684–694. http://dx.doi.org/10.1128/AEM.06274-11.

Nielsen, J., Larsson, C., van Maris, A., Pronk, J., 2013. Metabolic engineering of yeast for production of fuels and chemicals. Curr. Opin. Biotechnol. 24, 398–404. http://dx.doi.org/10.1016/j.copbio.2013.03.023.

Pandey, R.P., Sohng, J.K., 2013. Genetics of Flavonoids. In: K.G., Ramawat, J.M., Mérillon (Eds.), Natural Products. Springer, Berlin Heidelberg, pp. 1617–1645. http://dx.doi.org/10.1007/978-3-642-22144-6.

Perez-Gregorio, M.R., Regueiro, J., Simal-Gándara, J., Rodrigues, a.S., Almeida, D.P.F., 2014. Increasing the added-value of onions as a source of antioxidant flavonoids: a critical review. Crit. Rev. Food Sci. Nutr. 54, 1050–1062. http://dx.doi.org/10.1080/10408398.2011.624283.

Reid, R.J.D., Lisby, M., Rothstein, R., 2002. Cloning-free genome alterations in *Saccharomyces cerevisiae* using adaptamer-mediated PCR. Methods Enzymol. 350, 258–277. http://dx.doi.org/10.1016/S0076-6879(02)50968-X.

Santos, C.N.S., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab. Eng. 13, 392–400. http://dx.doi.org/10.1016/j.ymben.2011.02.002.

Scotti, L., 2012. SAR, QSAR and docking of anticancer flavonoids and variants: a review. Curr. Top. Med. Chem. 12, 2785–2809.

Siddiqui, M.S., Thodey, K., Trenchard, I., Smolke, C.D., 2012. Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. FEMS Yeast Res. 12, 144–170. http://dx.doi.org/10.1111/j.1567-1364.2011.00774.

Shi, S., Chen, Y., Siewers, V., 2014. Improving Production of Malonyl Coenzyme A-Derived Metabolites 5, 1–8, http://dx.doi.org/10.1128/mBio.01130-14. (Editor).

Takai A, Nishi R, Joe Y, Ito H., 2005. L-Tyrosine producing bacterium and a method for producing L-tyrosine. US Patent application no. 2005/0277179 A1.

Tribe DE., 1987. Novel microorganism and method. US Patent 4,681,852; Jul 21 1987.

Trantas, E., Panopoulos, N., Ververidis, F., 2009. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids *in Saccharomyces cerevisiae*. Metab. Eng. 11, 355–366. http://dx.doi.org/10.1016/j.ymben.2009.07.004.

Winkel-Shirley, B., 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 126, 485–493. http://dx.doi.org/10.1104/pp.126.2.485.

Wu, J., Zhou, T., Du, G., Zhou, J., Chen, J., 2014. Modular optimization of heterologous pathways for de novo synthesis of (2S)-naringenin in *Escherichia coli*. PLoS One 9, e101492. http://dx.doi.org/10.1371/journal.pone.0101492.

Yang, S.-M., Han, S.H., Kim, B.-G., Ahn, J.-H., 2014. Production of kaempferol 3-O-rhamnoside from glucose using engineered *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 41, 1311–1318. http://dx.doi.org/10.1007/s10295-014-1465-9.

Zheng, L., Baumann, U., Reymond, J.-L., 2004. An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res. 32, e115. http://dx.doi.org/10.1093/nar/gnh110.

Supplementary material

Supplementary Table 1. Primers used in this study.

USER cloning		
ID	**Name**	**Sequence (5'-3')**
1396	Sc_ARO4_2_fw	ATCTGTCAUAAAACAATGAGTGAATCTCCAATGTTCG
1397	Sc_ARO4_2_rv	CACGCGAUTCATTTCTTGTTAACTTCTCTTCTTTG
1398	Sc_ARO7_1_fw	AGTGCAGGUAAAACAATGGATTTCACAAAACCAGAAAC
1399	Sc_ARO7_1_rv	CGTGCGAUTCACTCTTCCAACCTTCTTAGCAAG
1468	Cg_ARO7_1_fw	AGTGCAGGUAAAACAATGGACTTCACTAAACCAGAAACTG
1469	Cg_ARO7_1_rv	CGTGCGAUTCACTTGTATTTTGCAACCACCG
1642	Ec_aroG_fw	AGTGCAGGUAAAACAATGAATTATCAGAACGACGATTTACGC ATCA
1644	Ec_aroF_2_fw	ATCTGTCAUAAAACAATGCAAAAAGACGCGCTGAATAACG
1645	Ec_aroG_2_fw	ATCTGTCAUAAAACAATGAATTATCAGAACGACGATTTACGCA TCA
1650	Ec_aroG_rv	CGTGCGAUTTACCCGCGACGCGCTTTTAC
1652	Ec_aroF_2_rv	CACGCGAUTTAAGCCACGCGAGCCGT
1653	Ec_aroG_2_rv	CACGCGAUTTACCCGCGACGCGCTTTTAC
1691	Fj_TAL_1_fw	AGTGCAGGUAAAACAATGAACACCATCAACGAATATCTGAGC
1692	Fj_TAL_1_rv	CGTGCGAUTTAATTGTTAATCAGGTG
2173	Sc_ARO1_1_fw	AGTGCAGGUAAAACAATGGTGCAGTTAGCCAAAG
2174	Sc_ARO1_1_rv	CGTGCGAUCTACTCTTTCGTAACGGCATC
2179	Sc_ARO2_2_fw	ATCTGTCAUAAAACAATGTCAACGTTTGGGAAACTG
2180	Sc_ARO2_2_rv	CACGCGAUTTAATGAACCACGGATCTGGA
2241	Sc_TYR1_1_fw	AGTGCAGGUAAAACAATGGTATCAGAGGATAAGATTGAG
2242	Sc_TYR1_1_rv	CGTGCGAUTTATGTATTTCTTTTTTCAGCGGC
3087	Ec_tyrA_1_fw	AGTGCAGGUAAAACAATGGTTGCTGAATTGACCG
3088	Ec_tyrA_1_rv	CGTGCGAUTCATTGTCTGTTATCGTTGGCT
6777	Ec_aroB_1_fw	AGTGCAGGUAAAACAATGGAGAGGATTGTCGTTACT
6778	Ec_aroB_1_rv	CGTGCGAUTTACGCTGATTGACAATCGG
6779	Ec_aroD_2_fw	ATCTGTCAUAAAACAATGAAAACCGTAACTGTAAAAGATC
6780	Ec_aroD_2_rv	CACGCGAUTTATGCCTGGTGTAAAATAGTTAAT
6781	Ec_aroE_1_fw	AGTGCAGGUAAAACAATGGAAACCTATGCTGTTTTTGG
6782	Ec_aroE_1_rv	CGTGCGAUTCACGCGGACAATTCCTC
6783	Ec_ydiB_1_fw	AGTGCAGGUAAAACAATGGATGTTACCGCAAAATACG

ID	Name	Sequence (5'-3')
6784	Ec_ydiB_1_rv	CGTGCGAUTCAGGCACCGAACCC
6785	Ec_aroL_2_fw	ATCTGTCAUAAAACAATGACACAACCTCTTTTTCTGA
6786	Ec_aroL_2_rv	CACGCGAUTCAACAATTGATCGTCTGTGC
6787	Ec_aroK_2_fw	ATCTGTCAUAAAACAATGGCAGAGAAACGCAATAT
6788	Ec_aroK_2_rv	CACGCGAUTTAGTTGCTTTCCAGCATGT
6789	Ec_aroA_1_fw	AGTGCAGGUAAAACAATGGAATCCCTGACGTTACAAC
6790	Ec_aroA_1_rv	CGTGCGAUTCAGGCTGCCTGGCTAAT
6791	Ec_aroC_2_fw	ATCTGTCAUAAAACAATGGCTGGAAACACAATTGG
6792	Ec_aroC_2_rv	CACGCGAUTTACCAGCGTGGAATATCAGT
Knockouts		
ID	**Name**	**Sequence (5'-3')**
91	Kl_URA3_Start_fw	TGGCAATTCCCGGGGATC
92	Kl_URA3_Start_rv	CGCTTCCCATCCAGCATTTC
93	Kl_URA3_End_fw	CTGTCGTTCCATTGAAAGC
94	Kl_URA3_End_rv	TAGGGCGAATTGGGTACC
150	Kl_LEU2_Start_rv	CAGAAGCATAACTACCCATTCC
151	Kl_LEU2_End_fw	TGGAAGAGGCAAGCACGTTAGC
476	Kl_LEU2_Start_fw	TGGCAATTCCCGGGGATCACGCTGCAGGTCGACAAC
479	Kl_LEU2_End_rv	TAGGGCGAATTGGGTACCGCCACTAGTGGATCTGATATCAC
1368	Sc_PDC5_Start_fw	CGTAAACCTGCATTAAG
1369	Sc_PDC5_Start_rv	GATCCCCGGGAATTGCCATTGTGTTGTTCTCTTTG
1370	Sc_PDC5_End_fw	GGTACCCAATTCGCCCTAGATTCAACGTTTGTGTA
1371	Sc_PDC5_End_rv	CTAAGATCATAGCTAAAGG
1372	Sc_ARO10_Start_fw	GGATAGCCGTCATTTAC
1373	Sc_ARO10_Start_rv	GATCCCCGGGAATTGCCAGAGGGTTGATCAGTTAAA
1374	Sc_ARO10_End_fw	GGTACCCAATTCGCCCTACTACCAATTGTTCGTTT
1375	Sc_ARO10_End_rv	CGATAGGAATGACAGAA
Point-directed mutagenesis		
ID	**Name**	**Sequence (5'-3')**
1404	Sc_ARO4_K229L_fw	CATTTCATGGGTGTTACTTTGCATGGTGTTGCTGCTATC
1405	Sc_ARO4_K229L_rv	GATAGCAGCAACACCATGCAAAGTAACACCCATGAAATG
1406	Sc_ARO7_G141S_fw	GATAAGAATAACTTCAGTTCTGTTGCCACTAG
1407	Sc_ARO7_G141S_rv	CTAGTGGCAACAGAACTGAAGTTATTCTTATC

ID	Name	Sequence (5'-3')
2070	Ec_aroG_L175D_fw	GGTGCACCGCGAAGATGCATCAGGGCTTTC
2071	Ec_aroG_L175D_rv	GAAAGCCCTGATGCATCTTCGCGGTGCACC
2102	Ec_aroG_S180F_fw	CATCAGGGCTTTTTTGTCCGGTCGG
2103	Ec_aroG_s180F_rv	CCGACCGGACAAAAAAGCCCTGATG
Verification		
ID	**Name**	**Sequence (5'-3')**
902	Sc_X-2-out-seq_rv	GAGAACGAGAGGACCCAACAT
904	Sc_X-3-out-seq_rv	CCGTGCAATACCAAAATCG
906	Sc_X-4-out-seq_rv	GACGGTACGTTGACCAGAG
1384	Sc_PDC5_Start_fw	AAAGCCTCCATATCCAAAG
1385	Sc_PDC5_End_rv	AGGTATGGTTAAAGATCACAC
1386	Sc_ARO10_Start_fw	ACCGAAATTTAAAAAAGCAG
1387	Sc_ARO10_End_rv	GTTTTCGGATAAAACTTCTTC
2220	Sc_ColoPCR_fw	CCTGCAGGACTAGTGCTGAG

Supplementary Table 2. List of Biobricks generated by PCR amplification.

NAME	TEMPLATE_FOR_PCR	FW_PRIMER_FOR_PCR	RV_PRIMER_FOR_PCR
BB0101 (URA3_2/3_START)	p0047 (pX-3-ccdB)	URA3_2/3_START_fw (ID91)	URA3_2/3_START_rv (ID92)
BB0102 (URA3_2/3_END)	p0047 (pX-3-ccdB)	URA3_2/3_END_fw (ID93)	URA3_2/3_END_rv (ID94)
BB0261 (Sc_Aro4->)	CEN.PK113-5D gDNA	Sc_aro4_U2_fw (ID1396)	Sc_aro4_U2_rv (ID1397)
BB0262 (Sc_Aro7<-)	CEN.PK113-5D gDNA	Sc_aro7_U1_fw (ID1398)	Sc_aro7_U1_rv (ID1399)
BB0299 (Cg_Aro7<-)	*C. guilliermondii* ATCC 6260 gDNA	Cg_Aro7_fw (ID1468)	Cg_Aro7_rv (ID1469)
BB0355 (Ec31884_AroG<-)	EcoMG1655 ATCC 31884 gDNA	AroG_1 Ec31884 Fw (1642)	AroG_1 Ec31884 Rv (1650),
BB0361 (Sc_Aro7_G141S<-)	p0761(pESC-URA-ARO7pm)	Sc_aro7_U1_fw (ID1398)	Sc_aro7_U1_rv (ID1399)
BB0364 (Sc_Aro4_K229L->)	p0775 (pESC-HIS-ARO4pm)	Sc_aro4_U2_fw (ID1396)	Sc_aro4_U2_rv (ID1397)
BB0380 (Fj_tal<-)	*F. johnsoniaeu* codon-optimized synthetic gene	Fj_Tal_U1_fw (ID1691)	Fj_TalU1_rv (ID1692)
BB0420 (Ec_AroG_L175D->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroG_24 fw (ID1645)	Ec_AroG_2 rv (ID1653)
BB0421 (Ec_AroG_S180F->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroG_2 fw (ID1645)	Ec_AroG_2 rv (ID1653)
BB0443 (ScAro1<-)	CEN.PK113-5D gDNA	Sc_Aro1 1-fw (ID2173)	Sc_Aro1 1-rv (ID2174)
BB0444 (ScAro2->)	CEN.PK113-5D gDNA	Sc_Aro2 2-fw (ID2179)	Sc_Aro2 2-rv (ID2180)
BB0447(Ec_TyrA_F1_1<- Ec)	EcoMG1655 ATCC 31884 gDNA	Ec_TyrA_1-fw (ID3087)	Ec_TyrA_1-rv (ID3088)
BB0453 (ScTyr1<-)	CEN.PK113-5D gDNA	Sc_Tyr1_1-fw (ID2241)	Sc_Tyr1_1-rv (ID2242)
BB0456 (Ec_AroF_2mt->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroF_2 fw (ID1644)	Ec_AroF_2 rv (ID1652)
BB0497 (Ec_AroB<-)	EcoMG1655 ATCC 31884 gDNA	Ec_AroB_U1_fw (ID6777)	Ec_AroB_U1_rv (ID6778)
BB0498 (Ec_AroD->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroD_U2_fw (ID6779)	Ec_AroD_U2_rv (ID6780)
BB0499 (Ec_AroE<-)	EcoMG1655 ATCC 31884 gDNA	Ec_AroE_U1_fw (ID6781)	Ec_AroE_U1_rv (ID6782)
BB0500 (Ec_Ydib<-)	EcoMG1655 ATCC 31884 gDNA	Ec_Ydib_U1_fw (ID6783)	Ec_Ydib_U1_rv (ID6784)
BB0501 (Ec_AroL->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroL_U2_fw (ID6785)	Ec_AroL_U2_rv (ID6786)

NAME	TEMPLATE_FOR_PCR	FW_PRIMER_FOR_PCR	RV_PRIMER_FOR_PCR
BB0502 (Ec_aroK->)	EcoMG1655 ATCC 31884 gDNA	Ec_aroK_U2_fw (ID6787)	Ec_aroK_U2_rv (ID6788)
BB0503 (Ec_AroA<-)	EcoMG1655 ATCC 31884 gDNA	Ec_AroA_U1_fw (ID6789)	Ec_AroA_U1_rv (ID6790)
BB0504 (Ec_AroC->)	EcoMG1655 ATCC 31884 gDNA	Ec_AroC_U2_fw (ID6791)	Ec_AroC_U2_rv (ID6792)
BB245 (KlLEU2_2/3_START)	p0019 (pUG73)	LEU2_2/3_START_fw (ID476)	LEU2_2/3_START_rv (ID150)
BB249 (Sc_PDC5_UP)	CEN.PK113-7D gDNA	Sc_PDC5_UP_fw (ID1368)	Sc_PDC5_UP_rv (ID1369)
BB250 (Sc_PDC5_DOWN)	CEN.PK113-7D gDNA	Sc_PDC5_END_fw (ID1370)	Sc_PDC5_END_rv (ID1371)
BB251 (Sc_ARO10_UP)	CEN.PK113-7D gDNA	Sc_ARO10_UP_fw (ID1372)	Sc_ARO10_UP_rv (ID1373)
BB252 (Sc_ARO10_DOWN)	CEN.PK113-7D gDNA	Sc_ARO10_END_fw (ID1374)	Sc_ARO10_END_rv (ID1375)
BB253 (Sc_PDC5_UP_URA3_2/3_START)	BB249, BB0101	Sc_PDC5_UP_fw (ID1368)	URA3_2/3_START_rv (ID92)
BB254(URA3_2/3_END_Sc_PDC5_DOWN)	BB250, BB0102	URA3_2/3_END_fw (ID94)	Sc_PDC5_END_rv (ID1371)
BB255(Sc_ARO10_UP_URA3_2/3_START	BB251, BB0101	Sc_ARO10_UP_fw (ID1372)	URA3_2/3_END_rv (ID94)
BB256(URA3_2/3_ENDSc_ARO10_DOWN	BB0102, BB252	URA3_2/3_END_fw (ID94)	Sc_ARO10_END_rv (ID1375)
BB681 (LEU2_2/3_END)	p0019 (pUG73)	LEU2_2/3_END_fw (ID151)	KlLEU2_2/3START_rv (ID479)
BB826 (Sc_PDC5_UP_LEU2_2/3_START)	BB249, BB245	Sc_PDC5_UP_fw (ID1368)	LEU2_2/3_START_rv (ID150)
BB827(LEU2_2/3_END_Sc_PDC5_DOWN)	BB681, BB250	LEU2_2/3_END_fw (ID151)	Sc_PDC5_END_rv (ID1371)
BB828(Sc_ARO10_UP_LEU2_2/3_START	BB251, BB245	Sc_ARO10_UP_fw (ID1372)	LEU2_2/3_START_rv (ID150)
BB829(LEU2_2/3_ENDSc_ARO10_DOWN)	BB681, BB252	LEU2_2/3_END_fw (ID151)	Sc_ARO10_END_rv (ID1375)

Supplementary Table 3. Plasmids construction.

Plasmid construction for site-directed mutagenesis			
Plasmid	**Parent plasmid (template for PCR)**	**FW_PRIMER_FOR_PCR**	**RV_PRIMER_FOR_PCR**
pCfB761	p0744 (pESC-URA-ARO7_G141S)	ID1406:aro7_15T_fw	ID1407:aro7_15T_rv
pCfB775	p0745 (pESC-HIS-ARO4_K229L)	ID1404:aro4_TT18AA_fw	ID1405:aro4_TT18AA_rv
pCfB1075	p01074 (pX-4-LoxP-SpHiS5 EcAroG)	AroG L175D _EC Fw (2070)	AroG L175D _EC RV (2071)
pCfB1076	p01074 (pX-4-LoxP-SpHiS5 EcAroG)	AroG S180F_EC Fw (2102)	AroG S180F_EC Rv (2103)

Plasmid construction by USER cloning				
Plasmid	**Parent plasmid**	**Biobrick 1**	**Biobrick 2**	**Promoter**
pCfB744	p0054 (pESC-URA-ccdB-USER)	BB0262 (ScAro7<-)		BB0008 (PTEF1<-)
pCfB745	p0055(pESC-HIS-ccdB-USER)		BB0261 (ScAro4->)	BB0009 (PPGK1->)
pCfB826	pCfB258 (pX-4-loxP-SpHiS5)	BB0361 (ScAro7_G141S<-)	BB0364 (ScAro4_K229L->)	BB0010 (<-PTEF1-PPGK1->)
pCfB827	pCfB258 (pX-4-loxP-SpHiS5)	BB0361 (ScAro7_G141S<-)	BB0456 (Ec31884_AroF_2mt->)	BB0010 (<-PTEF1-PPGK1->)
pCfB830	pCfB258 (pX-4-loxP-SpHiS5)	BB0299 (CgAro7<-)	BB0364 (ScAro4_K229L->)	BB0010 (<-PTEF1-PPGK1->)
pCfB831	pCfB258 (pX-4-loxP-SpHiS5)	BB0299 (CgAro7<-)	BB0456 (Ec31884_AroF_2mt->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1074	pCfB258 (pX-4-loxP-SpHiS5)	BB0355 (Ec31884_AroG<-)		BB0008 (PTEF1<-)
pCfB1077	pCfB255 (pX-2-loxP-KlURA3)	BB0361 (ScAro7_G141S<-)	BB0420 (Ec31884_AroG_L175D->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1078	pCfB255 (pX-2-loxP-KlURA3)	BB0420 (Ec31884_AroG_L175D->)	BB0420 (Ec31884_AroG_L175D->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1080	pCfB255 (pX-2-loxP-KlURA3)	BB0361 (ScAro7_G141S<-)	BB0421 (Ec31884_AroG_S180F->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1081	pCfB255 (pX-2-loxP-KlURA3)	BB0299 (CgAro7<-)	BB0421 (Ec31884_AroG_S180F->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1221	pCfB257 (pX-3-loxP-KlLEU2)	BB0453 (ScTyr1<-)		BB0008 (PTEF1<-)
pCfB1226	pCfB255 (pX-2-loxP-KlURA3)	BB0447(TyrA_F1_1<- Ec)	BB0364 (ScAro4_K229L->)	BB0010 (<-PTEF1-PPGK1->)

Plasmid	Parent plasmid	Biobrick 1	Biobrick 2	Promoter
pCfB1227	pCfB255 (pX-2-loxP-KlURA3)	BB0447(TyrA_F1_1<- Ec)	BB0456 (Ec31884_AroF_2mt->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1228	pCfB255 (pX-2-loxP-KlURA3)	BB0447(TyrA_F1_1<- Ec)	BB0420 (Ec31884_AroG_L175D->)	BB0010 (<-PTEF1-PPGK1->)
pCfB1964	pCfB255 (pX-2-loxP-KlURA3)	BB0380 (Flavobacterium<-)		BB0008 (PTEF1<-)
pCfB2733	pCfB257 (pX-3-loxP-KlLEU2)	BB0443 (ScAro1<-)	BB0444 (ScAro2->)	BB0010 (<-PTEF1-PPGK1->)
pCfB2739	pCfB257 (pX-3-loxP-KlLEU2)	BB0497 (Ec_AroB<-)		BB0008 (PTEF1<-)
pCfB2741	pCfB257 (pX-3-loxP-KlLEU2)	BB0499 (Ec_AroE<-)		BB0008 (PTEF1<-)
pCfB2742	pCfB257 (pX-3-loxP-KlLEU2)	BB0500 (Ec_Ydib<-)		BB0008 (PTEF1<-)
pCfB2743	pCfB257 (pX-3-loxP-KlLEU2)	BB0502 (Ec_aroK->)		BB0009 (PPGK1->)
pCfB2745	pCfB257 (pX-3-loxP-KlLEU2)	BB0503 (Ec_AroA<-)		BB0008 (PTEF1<-)
pCfB2746	pCfB257 (pX-3-loxP-KlLEU2)	BB0498 (Ec_AroD->)		BB0009 (PPGK1->)
pCfB2747	pCfB257 (pX-3-loxP-KlLEU2)	BB0501 (Ec_AroL->)		BB0009 (PPGK1->)
pCfB2749	pCfB257 (pX-3-loxP-KlLEU2)	BB0504 (Ec_AroC->)		BB0009 (PPGK1->)
pCfB2373	pCfB257 (pX-3-loxP-KlLEU2)	BB0443 (ScAro1<-)		BB0008 (PTEF1<-)
pCfB2748	pCfB257 (pX-3-loxP-KlLEU2)	BB0444 (ScAro2->)		BB0009 (PPGK1->)

Supplementary Table 4. Two-factor ANOVA analysis of influence of DAHP synthase and chorismate mutase overexpression on production of *p*-coumaric acid.

Dependent Variable: gL_CA					
Source	Type III Sum of Squares	Freedom degree	Mean Square	F	p-value
Corrected Model	2.949[a]	18	.164	9.885	.000
Intercept	45.229	1	45.229	2729.472	.000
DAHPsyn	.203	3	.068	4.078	.010
CHORmut	.068	4	.017	1.028	.399
DAHPsyn * CHORmut	2.529	11	.230	13.874	.000
Error	1.210	73	.017		
Total	57.530	92			
Corrected Total	4.158	91			

Supplementary Table 5. Production of *p*-coumaric upon overexpression of different variants of DAHP synthase and chorismate mutase.

DAHP synthase	Chorismate mutase	*p*-Coumaric acid (g L^{-1})
ScARO4fbr	*ScARO7fbr*	
ScARO4fbr	*ScARO7fbr*	1.16
ScARO4fbr	*ScARO7fbr*	1.14
ScARO4fbr	*ScARO7fbr*	0.97
ScARO4fbr	*ScARO7fbr*	1.03
ScARO4fbr	*ScARO7fbr*	0.96
ScARO4fbr	*ScARO7fbr*	1.11
ScARO4fbr	*ScARO7fbr*	0.97
ScARO4fbr	*ScARO7fbrTyr1*	0.99
ScARO4fbr	*ScARO7fbrTyr1*	1.08
ScARO4fbr	*ScARO7fbrTyr1*	0.76
ScARO4fbr	*ScARO7fbrTyr1*	1.05
ScARO4fbr	*CgARO7*	0.59
ScARO4fbr	*CgARO7*	0.57
ScARO4fbr	*CgARO7*	0.45
ScARO4fbr	*CgARO7*	0.55
ScARO4fbr	*CgARO7*	0.48
ScARO4fbr	*CgARO7Tyr1*	0.48
ScARO4fbr	*CgARO7Tyr1*	0.64
ScARO4fbr	*CgARO7Tyr1*	0.69
ScARO4fbr	*CgARO7Tyr1*	0.67
ScARO4fbr	*EcTyrA*	0.52
ScARO4fbr	*EcTyrA*	0.64
ScARO4fbr	*EcTyrA*	0.57
ScARO4fbr	*EcTyrA*	0.62
ScARO4fbr	*EcTyrA*	0.49
EcAroFfbr	*ScARO7fbr*	0.75
EcAroFfbr	*ScARO7fbr*	0.88
EcAroFfbr	*ScARO7fbr*	0.73

DAHP synthase	Chorismate mutase	*p*-Coumaric acid (g L^{-1})
EcAroFfbr	*ScARO7fbr*	0.74
EcAroFfbr	*ScARO7fbr*	0.93
EcAroFfbr	*ScARO7fbrTyr1*	0.82
EcAroFfbr	*ScARO7fbrTyr1*	0.83
EcAroFfbr	*ScARO7fbrTyr1*	0.72
EcAroFfbr	*ScARO7fbrTyr1*	0.79
EcAroFfbr	*CgARO7*	0.67
EcAroFfbr	*CgARO7*	1.17
EcAroFfbr	*CgARO7*	1.15
EcAroFfbr	*CgARO7*	0.9
EcAroFfbr	*CgARO7*	1.1
EcAroFfbr	*CgARO7*	0.97
EcAroFfbr	*CgARO7Tyr1*	0.7
EcAroFfbr	*CgARO7Tyr1*	0.86
EcAroFfbr	*CgARO7Tyr1*	0.71
EcAroFfbr	*CgARO7Tyr1*	0.67
EcAroFfbr	*CgARO7Tyr1*	0.96
EcAroFfbr	*CgARO7Tyr1*	0.53
EcAroFfbr	*CgARO7Tyr1*	0.61
EcAroFfbr	*CgARO7Tyr1*	0.69
EcAroFfbr	*EcTyrA*	0.88
EcAroFfbr	*EcTyrA*	0.6
EcAroFfbr	*EcTyrA*	0.69
EcAroFfbr	*EcTyrA*	0.84
EcAroFfbr	*EcTyrA*	0.84
EcAroFfbr	*EcTyrA*	0.67
EcAroGfbr1	*ScARO7fbr*	0.37
EcAroGfbr1	*ScARO7fbr*	0.34
EcAroGfbr1	*ScARO7fbr*	0.44

DAHP synthase	Chorismate mutase	*p*-Coumaric acid ($g\ L^{-1}$)
EcAroGfbr1	*ScARO7fbrTyr1*	0.55
EcAroGfbr1	*ScARO7fbrTyr1*	0.38
EcAroGfbr1	*ScARO7fbrTyr1*	0.69
EcAroGfbr1	*CgARO7*	0.82
EcAroGfbr1	*CgARO7*	0.99
EcAroGfbr1	*CgARO7*	1.03
EcAroGfbr1	*CgARO7*	0.78
EcAroGfbr1	*CgARO7*	1.09
EcAroGfbr1	*CgARO7*	0.73
EcAroGfbr1	*CgARO7Tyr1*	0.8
EcAroGfbr1	*CgARO7Tyr1*	0.58
EcAroGfbr1	*CgARO7Tyr1*	0.74
EcAroGfbr1	*CgARO7Tyr1*	0.87
EcAroGfbr1	*CgARO7Tyr1*	0.53
EcAroGfbr1	*EcTyrA*	1.01
EcAroGfbr1	*EcTyrA*	1
EcAroGfbr1	*EcTyrA*	1.04
EcAroGfbr2	*ScARO7fbr*	0.67
EcAroGfbr2	*ScARO7fbr*	0.58
EcAroGfbr2	*ScARO7fbr*	0.57
EcAroGfbr2	*ScARO7fbr*	0.78
EcAroGfbr2	*ScARO7fbr*	1
EcAroGfbr2	*ScARO7fbrTyr1*	0.66
EcAroGfbr2	*ScARO7fbrTyr1*	0.51
EcAroGfbr2	*ScARO7fbrTyr1*	0.91
EcAroGfbr2	*ScARO7fbrTyr1*	0.99
EcAroGfbr2	*CgARO7*	0.63
EcAroGfbr2	*CgARO7*	0.39
EcAroGfbr2	*CgARO7*	0.38
EcAroGfbr2	*CgARO7*	0.67
EcAroGfbr2	*CgARO7Tyr1*	0.69

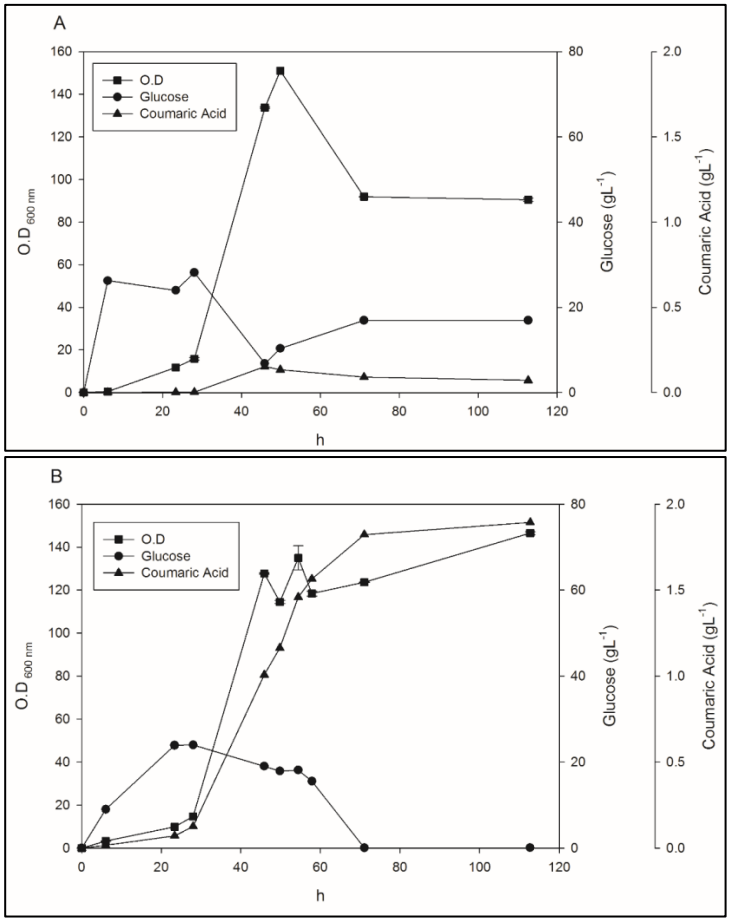
DAHP synthase	Chorismate mutase	*p*-Coumaric acid ($g\ L^{-1}$)
EcAroGfbr2	*CgARO7Tyr1*	0.72
EcAroGfbr2	*CgARO7Tyr1*	0.61
EcAroGfbr2	*CgARO7Tyr1*	0.83

Supplementary Materials and Methods

Controlled fed-batch cultivations

Glycerol stocks of strains ST4069 and ST4058 were inoculated into 50 ml synthetic drop-out medium in a 250-ml shake flasks (SC ura-his-leu-) and incubated at 30 °C in an orbital shaker (200 rpm) for 24 hours. The cell suspension was up concentrated to 5 ml by centrifugation and used to inoculate 0.5 L fermentation medium in Sartorius bioreactors with a maximal working volume of 1 L. The fermentation medium contained per litter: 15.0 g $(NH_4)_2SO_4$, 6.0 g KH_2PO_4, 1.0 g MgSO4·7H_2O, 4 ml trace metal solution, 2 ml vitamins solution, 0.4 ml antifoam A (Sigma-Aldrich), and 40g dextrose. Dextrose was autoclaved separately. Vitamins solution was added to the autoclaved medium by sterile filtration. The fermentation was carried out at 30 °C with agitation rate at 800 rpm. The pH was maintained at 5.0 by automatic addition of 2 N NaOH. Carbon dioxide concentration in the off-gas was monitored by an acoustic gas analyser (model number 1311, Bruël&Kjær). Once the glucose was exhausted, which was observed from the decline in CO_2 production and was also confirmed by residual glucose detection using glucose strips Glucose MQuant™ (MerckMillipore), the feed was started at 5 g h^{-1}. Constraint feed rate was maintained throughout the fed-batch phase. The total volume of feed solution used per reactor was 0.5 L. The feed solution contained per litter: 45 g $(NH_4)_2SO_4$, 18 g KH_2PO_4, 3 g $MgSO_4·7H_2O$, 12 ml trace metals solution, 6 ml vitamins solution, and 0.6 ml antifoam A. Dextrose and vitamins were added to feed solution in the same way as to batch fermentation medium.

The reactors were sampled twice a day to measure OD_{600} and metabolites. For metabolites analysis the sample was centrifuged and the supernatant was stored at −20°C until HPLC analysis.



Supplementary Figure 1. Fed-batch cultivation of non-optimized control strain ST4069 (A) and of the optimized *p*-coumaric acid-overproducing strain ST4058 (B).

3. Systems biology analysis of the *p*-coumaric acid platform strain

Metabolic response of *Saccharomyces cerevisiae* to the over-production of *p*–coumaric acid

Angelica Rodriguez[1], Yun Chen[2], Sakda Khoomrung[2], Emre Özdemir[1], Irina Borodina[1], Jens Nielsen[1,2]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle allé 6, 2970 Hørsholm, Denmark

[2]Department of Biology and Biological Engineering, Chalmers University of Technology, SE412 96 Gothenburg, Sweden

Abstract

The development of robust and efficient cell factories requires understanding of the metabolic changes triggered by the production of the targeted compound. Here we aimed to understand how production of *p*-coumaric acid, a precursor of multiple secondary aromatic metabolites, influences *S. cerevisiae*. We evaluated the growth and *p*-coumaric acid production in batch and chemostat cultivations and analyzed the transcriptome and intracellular metabolome during steady state in low- and high-producers of *p*-coumaric acid in two strain backgrounds, S288c or CEN.PK.

We found that the same genetic modifications resulted in higher production of *p*-coumaric acid in the CEN.PK background than in the S288c background. Moreover, the CEN.PK strain was less affected by the genetic engineering as was evident from fewer changes in the transcription profile and intracellular metabolites concentrations. Surprisingly, for both strains we found the largest transcriptional changes in genes involved in transport of

amino acids and sugars, which were downregulated. Additionally, in S288c amino acid and protein biosynthesis processes were also affected.

Based on these results, we propose several potential metabolic engineering strategies for further improvement of *p*-coumaric acid production by yeast.

Introduction

Plants produce a wide range of secondary metabolites as a protective mechanism to stresses caused by bacterial or viral infections, ultraviolet radiation, wounds, and other biotic and abiotic factors. Nearly 15% of these metabolites are phenolic compounds derived from the aromatic amino acids L-tyrosine, L-phenylalanine or L-tryptophan (Wink *et al*., 2010). Numerous aromatic secondary metabolites are available on the market as therapeutic agents, dyes, fragrances, and flavors. The majority of these compounds are currently synthesized chemically or isolated from plants (Bourgaud *et al*., 2001), however recently there have been significant advances in engineering industrial microbes, e.g., *Escherichia coli* and *S. cerevisiae*, for production of aromatic secondary metabolites by fermentation. A few biotech-derived aromatics are already on the market, such as phenylalanine, resveratrol, vanillin, steviol glucoside and others. Additionally, many aromatic metabolites have been produced in microbial cell factories at proof-of-concept levels, i.e., the strains, fermentation and downstream processes need further development before the production becomes economically feasible. These compounds include naringenin, genistein, kaempferol, fisetin, melatonin, resveratrol, and many others (Koopmann *et al*., 2012; Trantas *et al.,* 2009; Santos *et al.,* 2011; Leonard *et al.,* 2006; Stahlhut *et al.,* 2015; Krivoruchko and Nielsen, 2015; Li et al., 2016; Germann *et al*., 2016).

An important step towards improved microbial cell factories is a better understanding of how the engineered cells respond to production of target compounds (Nielsen and Keasling, 2016). For this purpose, ‘omic-level

characterization of the strains is useful since the organism can be studied at different levels and the information can be assessed in the context of cellular metabolism (Kim *et al*., 2012 and Curran *et al*., 2012). There are a few successful examples of applying systems biology for guiding metabolic engineering strategies. Otero *et al*. (2013) obtained a 30-fold improvement in succinic acid production in *S. cerevisiae* based on the integrative analysis of physiology and transcriptome data. Park *et al.* (2012) engineered an efficient L-valine-producing *E. coli* by using transcriptomic analysis together with *in silico* models (Park *et al*., 2007). A multi-omic analysis of two different *E. coli* strains allowed Yoon *et al*. (2012) to identify an optimal strain for production of recombinant proteins. This study is one of the few that considered the differences between strains of the same species, when selecting the suitable host organism. In *S. cerevisiae*, a considerable number of differences have been found in the genomes of two widely used strains, CEN.PK and S288c. These differences are mainly related to the presence of 13,787 single nucleotide polymorphisms, 939 of them related to 158 genes involved in metabolic functions with enrichment in the galactose uptake and ergosterol biosynthetic pathways. Moreover, 83 genes, mainly located in sub-telomeric regions of S288c, are absent in the CEN.PK strain (Otero *et al*., 2010 and Nijkamp *et al*., 2012).

The strain CEN.PK is widely used for industrial biotechnology research and applications, whereas the strain S288c is widely used in genetic studies. Recently the strain S288c has also been used for the production of some metabolites, such as vanillin-β-glucoside, 2-phenylethanol and methionol (Strucko *et al*., 2012; Yin *et al*., 2015a and Yin *et al*., 2015b). In the particular case of vanillin-β-glucoside, the engineered S288c strain produced 10-fold more product than the CEN.PK strain engineered in the same way and this effect was associated with several single nucleotide polymorphisms in the shikimate pathway genes.

In this study we aimed to investigate firstly how the production of *p*-coumaric acid (*p*-CA) influences the host and secondly whether these effects depend on the strain background. To answer these two questions, we performed transcriptome and intracellular metabolome analysis on S288c and CEN.PK strains, which either only expressed an enzyme for making the product (low-producers) or were additionally optimized towards production of aromatic products (high-producers).

Methods

Plasmids and strains

E. coli DH5*α* was used for cloning procedures. The fragments used for overexpression of genes were amplified by PCR using primers and templates as described in the Supplementary Table 1. The fragments were amplified from the genomic DNA of *S. cerevisiae* CEN.PK102-5B *(*MATa *ura*3-52 *his*3Δ1 *leu*2-3/112 MAL2-8c SUC2*)* and *E. coli* NST 74*.* The gene encoding tyrosine ammonia-lyase from *Flavobacterium johnsoniae* (FjTAL) was as in (Rodriguez et al., 2015). The amplified gene-encoding fragments were cloned together with strong constitutive promoters into EasyClone integrative plasmids by USER cloning (Jensen *et al*., 2014). The clones were tested for correct insertion of gene/promoter fragments by colony PCR using the primers summarized in Supplementary Table 1 and the resulting plasmids were verified by sequencing. The list of the constructed vectors can be found in Table 1.

S. cerevisiae CEN.PK113-7D was obtained from Peter Kötter (Johann Wolfgang Goethe-University Frankfurt, Germany). The strain BY4741, a derivative of strain S288c, was obtained from EUROSCARF. Transformation of yeast cells was performed using the lithium acetate method (Gietz and Woods, 2002). The strains were selected on synthetic drop-out medium

(Sigma-Aldrich) and the genetic modifications were confirmed by colony PCR. The yeast strains used in this study are listed in Table 1.

Media and cultivations

We prepared a mineral medium for the batch fermentation according to Verduyn *et al*. (1992). Glucose concentration in batch medium was 20 g l^{-1} glucose. The feed medium for chemostats was prepared in the same way, but the amount of glucose was reduced to 10 g l^{-1} and the medium was supplemented with 0.2 mL L^{-1} of 2M KOH and one drop of antifoam 204 (Sigma A-8311) per 20 L of medium. The pre-culture was done by inoculating a yeast colony into 50 ml of mineral medium in a 250-ml baffled shake flask and incubating the culture with shaking at 200 rpm at 30°C for around 12 hours. When the pre-culture reached OD_{600} of ca. 2, it was used to inoculate a bioreactor to a starting optical density of 0.05.

The fermentations were performed in DasGip 1-L stirrer-pro vessels (Eppendorf, Jülich, Germany), using the working volume of 500 ml. The temperature was 30°C, agitation was at 600 rpm and aeration at 1 vvm. pH was monitored with a pH sensor (Mettler Toledo, Switzerland) and pH was maintained at 5.0 ± 0.05 by automatic addition of 2M KOH. Dissolved oxygen was above 30% throughout the fermentation as measured by the polarographic oxygen sensor (Mettler Toledo, Greifensee, Switzerland). The completion of the batch phase was determined by monitoring CO2 in the exhaust gas, when the 2nd CO2 peak, corresponding to ethanol consumption phase, declined. We then initiated constant feed to obtain glucose-limited steady-state with dilution rate of 0.100±0.005 h^{-1}. The volume was kept constant using an overflow pump. The samples for transcriptome and metabolome analysis were taken after 3 residence times of steady-state growth. Four technical replicates were taken from each reactor for transcriptome and metabolome analyses. Each strain was fermented twice to obtain 2 biological replicates.

Table 1. Plasmids and strains used in this study.

Strains				
Strain ID	**Parent strain**	**Genotype**	**Plasmids**	**Source**
ST10	-	CEN.PK102.5B MAT*a ura*3-52*his*3 Δ 1*leu*2-3/112 *MAL*2-8^{c} *SUC*2		Entian and Kötter, 2007
ST144	-	S288c MAT*a his*3Δ0 l*eu*2Δ0 *met*15Δ0 *ura*3Δ0		Brachnan *et al*., 1998
ST691	ST10	CEN.PK MAT*a aro10*Δ *pdc5*Δ *ura*3-52*his*3 Δ 1*leu*2-3/112 *MAL*2-8^{c} *SUC*2		Rodriguez *et al*., 2015
ST4360	ST144	S288c MAT*a his*3Δ0 *leu*2Δ0 *ura*3Δ0		This study
ST4195	ST144	S288c MAT*a aro*10Δ *pdc*5Δ *ura*3Δ0 *his*3Δ0 *leu*2Δ0		This study
ST4408	ST10	P_{TEF1}->Fj_*TAL*	pCfB257, pCfB258, p03523	This study
ST4288	ST691	P_{TEF1}->Fj_*TAL*, P_{PGK1}->Ec_*aroL*, P_{TEF1}->Sc_*ARO7*^{G141S}, P_{PGK1}->Sc_*ARO4*^{K229L}	pCfB257, pCfB826, p03524	This study
ST4353	ST4360	P_{TEF1}->Fj_*TAL*, P_{PGK1}->Ec_*aroL*, P_{TEF1}->Sc_*ARO7*^{G141S}, P_{PGK1}->Sc_*ARO4*^{K229L}	pCfB257, pCfB258, p03523	This study
ST4397	ST4195	P_{TEF1}->Fj_*TAL*	pCfB257, pCfB826, p03524	This study
Integrative plasmids				
Name	**Parent plasmid**	**Properties**		**Reference**
pCfB257		X-3, loxP, *KlLEU2*		Jensen *et al*., 2014
pCfB258		X-4, loxP, *SpHIS*5		Jensen *et al*., 2014
pCfB390		XI-3-loxP-*KlURA*3		Jensen *et al*., 2014
pCfB0826	pCfB258	X-4, loxP, *SpHIS*5, BB0361(Sc*ARO7*^{G141S}<-), BB0010(<-P_{TEF1}-P_{PGK1}->), BB0364(Sc*ARO4*^{K229L}->)		Rodriguez *et al*., 2015
pCfB03523	pCfB390	XI-3-loxP, *KlURA*3, BB0380(Fj_*TAL*<-), BB0008(P_{TEF1}<-),		This study
pCfB03524	pCfB390	XI-3-loxP, *KlURA*3, BB0380 (Fj_*TAL*<-), BB0010(<-P_{TEF1}-P_{PGK1}->), BB0501(Ec*aroL*->)		This study

Analytical methods

For analysis of extracellular metabolites and the biomass, we withdrew ca. 3-ml samples from the reactor. 1 ml of the sample was centrifuged at 11,000 x g for 5 min and stored at -20ºC until HPLC analysis for glucose and organic acids. For *p*-CA analysis in the optimized strains (ST4288 and ST4353) we mixed 1 volume of sample with 9 volumes of 50% ethanol, whereas for the non-optimized strains (ST4408 and ST4397) we mixed 1 volume of sample with 1 volume of 50% ethanol. This was done to dissolve the *p*-CA that may have precipitated from the broth due to poor solubility in water. These samples were also centrifuged at 11,000 g for 5 min and stored at -20ºC until further analysis.

The analysis of glucose, glycerol, ethanol, and organic acids was performed on Dionex Ultimate 3000 high-performance liquid chromatography (HPLC) system (Dionex Softron GmbH, Germany), with an Aminex HPX-87H column (Bio-Rad) at 65°C, using 5 mM H_2SO_4 as the mobile phase with a flow rate of 0.6 ml/min.

Quantification of *p*-CA was performed as described in Rodriguez et al. (2015) using a HPLC (Thermo Fisher Scientific), with a Discovery HS F5 150mm X 2.1mm column (particle size 3µm). The samples were analyzed using a gradient method with two solvents: (A) 10 mM ammonium formate pH 3.0 and (B) acetonitrile at 1.5 ml min^{1}. The *p*-CA was detected by absorbance at 277 nm and the retention time was 4.7 min. The area under the curve was integrated with Chromeleon software 7. The quantification of *p*-CA was performed based on 5 points calibration curve in the range of 0.1 mM to 1 mM. For the dry cell weight measurement 5 mL of culture broth was filtered through a 0.45 µm filter membrane, after that the membrane was dried at 95°C for 24 hours and cooled down in a desiccator. The dry cell weight was calculated by measuring the weight increment of the dried filter.

Transcriptome analysis

Samples for RNA extraction were taken after four retention times of steady-state fermentation by rapidly withdrawing 5 ml of culture and injecting it into a 50 ml falcon tube with ca. 30 ml of crushed ice, the samples were immediately centrifuged at 4,000 rpm for 5 minutes at -20°C. The supernatant was discarded, the pellet was frozen in liquid nitrogen and stored at -80°C until further analysis. The RNA extraction was performed using the RNeasy Mini Kit (QIAGEN). The DNA was removed from the sample using Turbo DNA-free Kit (Ambion). The purified RNA samples were analyzed with a 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA) and stored at -80°C until further analysis.

The sequencing libraries were prepared in four replicates using a TruSeq® Stranded mRNA sample preparation kit LT (Illumina Inc.). The final concentration of each cDNA library was measured by Qubit® 2.0 Fluorimeter and Qubit dsDNA Broad Range assay (Life Technologies). Average dsDNA library size was determined by using the Agilent DNA 1000 kit on an Agilent 2100 Bioanalyzer (Agilent Technologies). Libraries were normalized and pooled in 10 mM Tris-Cl, pH 8.0, plus 0.05% Tween 20 to the final concentration of 10 nM. Denaturated in 0.2N NaOH, 1.2 pm pool of 16 libraries in 1300 µl ice-cold HT1 buffer was loaded into the flow cell provided in the NextSeq 500/550 Mid Output v2 Reagent kit (150 cycles, Illumina Inc.). Libraries were sequenced on the NextSeq (Illumina Inc.) platform with a paired end protocol and read lengths of 75 nt.

Metabolomics analysis

Sampling, quenching and washing of the intracellular metabolites was performed as described by Canelas *et al*. (2009). The analysis of amino acids and other organic acids was performed according to Khoomrung *et al*. (2015). The analysis of the derivatized metabolites was performed using a Focus GC ISQ-LT single quadrupole GC-MS (Thermo Fisher Scientific, USA). The column was a Zebron ZB-1701 GC column (30 m 0.25 mm I.D.,

0.25-mm film thickness, Phenomenex, Macclesfield, UK). The metabolites were identified by comparing their retention times and mass spectrum profiles with the authentic standards or the mass spectra from the National Institute of Standards and Technology (NIST), USA library. The data was processed using the Quan browser function in the Xcalibur software version 2.2 (Thermo Fisher Scientific).

Data analysis

The alignment of sequencing reads to the reference genome was performed using TopHat, the assembly and quantification of the expression levels was developed with Cufflinks and a preliminary analysis of the data was performed with CummeRbund, the tree methods were used as described by Trapnell *et al*. (2012). Paired comparisons were performed between optimized and non-optimized strains on each background: the strain ST4288 was compared with the strain ST4408 and the strain ST4397 was compared with the strain ST4353.

The gene set analysis was performed using the R package Piano (Väremo *et al*., 2013), a platform for integrative analysis of omics data. The p-values and the fold changes were used as input data and two types of analysis were performed with this program: first a gene-set analysis with the reporter algorithm for gene ontology (GO) and second a gene set analysis using the reporter metabolites. The gene-metabolite network was obtained from the *S. cerevisiae* metabolic model iTO977 (Osterlund *et al*., 2013). The gene sets and reporter metabolites with a distinct directional p-value < 0.05 were chosen for the analysis.

The network topology analysis was performed using Kiwi a tool for visualization and interpretation of gene sets analysis (Väremo *et al*., 2014). This tool allows integrating the results of the gene set analysis with a gene set interaction network. The input for this analysis was a gene set

interaction network obtained from the *S. cerevisiae* metabolic model iTO977 and the results from the gene set analysis using reporter metabolites**.**

For the metabolomics data, we did a PCA analysis in MATLAB to identify the differences between the four strains object of this research and to establish the differences between the engineered and non-engineered strain we did volcano plots based on the results of a t-test comparison between optimized and non-optimized strains on each background.

Results

Physiological characterization of low and high producers of *p*-coumaric acid

To understand the fundamental metabolic changes triggered by the overproduction of *p*-CA, and the response of different background strains to these changes, we constructed two strains in each of the genetic backgrounds (CEN.PK and S288c). The “low-producers” were generated by overexpressing tyrosine ammonia lyase from *Flavobacterium johnsoniae* under control of the P_{TEF1} promoter. The “high-producers” were created by additional overexpression of *aroL* from *Escherichia coli* under control of the P_{TEF1} promoter, *ARO7*^{G141S} and *ARO4*^{K229L} from *S. cerevisiae* under control of the promoters P_{TEF1} and P_{PGK1} respectively and deletion of *ARO10* and *PDC5*. The resulting 4 strains were analyzed in batch and glucose-limited chemostat cultivations.

The concentrations of *p*-CA in batch and continuous cultivations were higher in the CEN.PK strains in comparison to the S288c strains with the same genetic modifications (Table 2). Glycerol yield was higher in the S288c strains in comparison to the CEN.PK strains. In batch fermentations, the optimized strains had lower biomass yield and accumulated more acetate than the non-optimized ones.

Table 2. Physiological data of the strains grown in batch and chemostat cultivations.

Background strain	**CEN.PK**		**S288c**	
Strain ID	ST4408	ST4288	ST4397	ST4353
Optimized for *p*-CA production	No	Yes	No	Yes
Maximum specific growth rate μ_{max} (h^{-1})	0.334 ± 0.006	0.294 ± 0.006	0.292 ± 0.005	0.271 ± 0.009
Final titer of *p*-CA (g L^{-1})	0.202 ± 0.005	2.405 ± 0.054	0.081 ± 0.005	2.018 ± 0.000
Biomass yield on glucose (g $glucose^{-1}$)	0.477 ± 0.044	0.268 ± 0.093	0.422 ± 0.005	0.352 ± 0.053
p-CA yield on glucose (g $glucose^{-1}$)	0.001 + 0.000	0.013 + 0.000	0.000 ± 0.000	0.012 ± 0.000
Glycerol yield on glucose (g $glucose^{-1}$)	0.019 ± 0.001	0.018 ± 0.002	0.048 ± 0.000	0.053 ± 0.001
Acetate yield on glucose (g $glucose^{-1}$)	0.010 ± 0.003	0.014 ± 0.002	0.006 ± 0.002	0.011 ± 0.001
Ethanol yield on glucose (g $glucose^{-1}$)	0.309 ± 0.001	0.293 ± 0.004	0.289 ± 0.013	0.297 ± 0.013
Final biomass dry weight (g $glucose^{-1}$)	13.198 ± 0.279	12.566 ± 0.056	11.481 ± 0.131	10.401 ± 0.785
Chemostat (steady-state data)				
Titer of *p*-CA (g L^{-1})	0.117 ± 0.000	0.507 ± 0.013	0.081 ± 0.005	0.410 ± 0.024
Biomass dry weight (g CDW L^{-1})	6.197 ± 0.207	6.226 ± 0.242	6.167 ± 0.220	6.159 ± 0.125
Glucose (g L^{-1})	ND	ND	ND	ND
Glycerol (g L^{-1})	ND	0.021 ± 0.004	ND	0.013 ± 0.001
Acetate (g L^{-1})	ND	ND	ND	ND
Ethanol (g L^{-1})	ND	ND	ND	ND

ND – not detected.

Transcriptional response of the strains to the synthesis of *p*-coumaric acid

For analysis of the differential gene expression, we did pairwise comparisons between the optimized and non-optimized strains for *p*-CA production: CEN.PK strains (ST4288 and ST4408) and the S288c strains (ST4353 and ST4397).

Significantly up and down-regulated gene sets were identified through a gene set analysis using GO terms (p adjusted value < 0.05). The engineered high-producing CEN.PK strain did not have any gene sets that

were significantly upregulated in comparison to the low-producing CEN.PK strain. For S288c strain, however, gene sets related to DNA helicase activity, telomere maintenance and ribonuclease activity were upregulated (Figure 1). Among the down-regulated gene sets were transport functions and iron metabolism, which was observed for both strain backgrounds. Additionally, S288c strain had remarkable downregulations in gene sets related to the synthesis of amino acids and proteins (Figure 1).

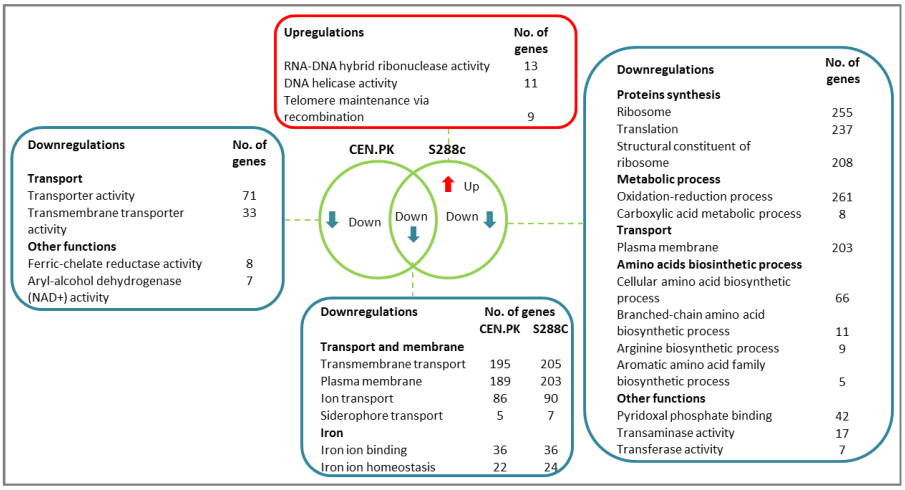


Figure 1. Gene sets with significant differences in the optimized strains for *p*-CA production in comparison to the non-optimized strains.

To elucidate the biological connections between the gene sets, we did a network analysis of the gene sets using metabolite reporters and used the network visualization tool Kiwi for visualizing the results (Väremo *et al*., 2014). The network analysis allows integrating the information from the gene set analysis with the metabolites interactions from a metabolic model.

For the CEN.PK strain we can see significant downregulations correlated to eight amino acids, hydron H(+) and galactose. All the amino acids mapped in the network have in common upregulations in *AGP1* and *GAP1* together

with downregulations in *BAP3* and *BAP2* and *TAT1*, all of them are involved in the transport of amino acids (Figure 2A, Supplementary Figure 1).

One of the metabolites correlated to downregulations is L-tyrosine, the precursor of *p*-CA, the network analysis shows that on top of the correlation to transport downregulation, this metabolite is related to a strong upregulation in the aromatic aminotransferase II *ARO9* involved in the conversion of *p*-hydroxyphenylpyruvate into L-tyrosine. Finally, the metabolite Hydron is mainly correlated to downregulations in genes involved in transport functions (*BAP2*, *TAT1*, *ALP1*, *TPO1*, *BIO5*, *VHT1*) and D-galactose is mainly correlated to downregulation of hexose transport (*HXT10*, *HXT14*).

The strain S288c had downregulations correlated to 5 amino acids (L-Methionine, L-tyrosine, L-tryptophan, L-glutamate and L-ornithine), three sugars (D-fructose, alpha-D-glucose and alpha-D-mannose), acetaldehyde and 2-oxoglutarate (Figure 2B, Supplementary Figure 2).

The amino acids reported in the network analysis had in common the downregulation of *BAP2*, *TAT1* and the upregulation of *AGP1*. These genes are correlated to amino acid transmembrane transporter activity. We found two aromatic amino acids in the network L-tyrosine and L-tryptophan, they share downregulations in genes related to amino acids transport (*BAP2* and *TAT1*), and they differ in the upregulations when L-tyrosine has a strong upregulation in *ARO9* and *ALD3*. L-tryptophan has a strong upregulation in *MSY1* (Figure 2B, Supplementary Figure 2).

Finally, the metabolites 2-oxoglutarate, L-glutamate and L-ornithine are correlated to downregulations in genes involved in the synthesis and transport of amino acids. L-glutamate has a central role in the metabolic network and it is correlated to downregulation of *ASN1*, *ADE4*, *CAR2* and *TRP2*, genes involved in the synthesis of amino acids. The three sugars

reported in the network are correlated to downregulations in genes involved in the transport of sugars (*HXT2* and *HXT16*).

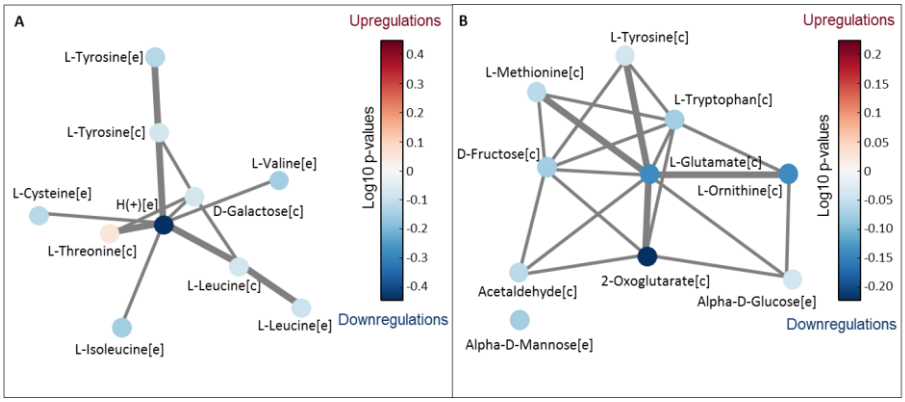
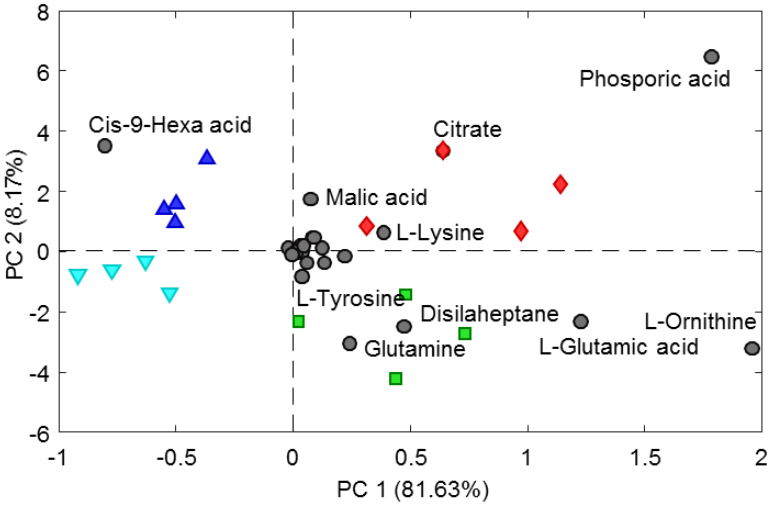


Figure 2. Network topology analysis identified metabolites with significant differences in the optimized strains in comparison to the non-optimized strains. The nodes are resized according to the gene-set significance, the colors reflect the direction of change of the gene-set, the edges between two metabolites symbolized how close they are in the metabolite-metabolite network and the thickest edges link the metabolites that are in close proximity to each other. A. comparison between CEN.PK strains; B. Comparison between S288c strains.

Changes of intracellular metabolome in response to the synthesis of *p*-coumaric acid

To identify the differences in the metabolome caused by overproduction of *p*-CA and by different genetic backgrounds, we did a PCA analysis. The first component of the PCA accounted for 82% of the variability and showed significant differences between the background strains CEN.PK and S288c, the metabolites with higher contributions to this component are phosphoric acid, disilaheptane, L-ornithine, glutamic acid, lysine and citrate, all of them related to higher concentration of the metabolites in the CEN.PK background whereas the strain S288c is related to higher values of cis-9-hexadecanoic acid (Figure 3).

The second component explains 8% of the variability and establishes the differences between optimized and non-optimized strains for the two backgrounds tested, the metabolites with higher contributions to this component are glutamine and L-tyrosine with higher concentrations for the engineered strains whereas the non-optimized strains are related to higher concentrations of malic acid.



Background strain	Low *p*-CA production	High *p*-CA production
CEN.PK	ST4408	ST4288
S288c	ST4397	ST4353
Metabolites		

Figure 3. Score and loading plot from the principal component analysis based on the metabolome data of optimized and non-optimized strains in the *S. cerevisiae* backgrounds CEN.PK and S288c.

A t-test was performed comparing the non-optimized and optimized strain of each background, aiming to identify which metabolites has significant differences when the cells are producing *p*-CA,. The significant differences for both of the background strains tested are mainly related to the low

concentration of metabolites in the optimized strains. For the CEN.PK strains, we found significant differences in four metabolites: two amino acids (L-valine and L-threonine), malic and citric an acid. For the S288c strain, we found significant differences in 5 amino acids, phosphoric acid, malic acid, citric acid and cis-9 hexadecanoic acid (Figure 4). There is bigger variance between the low and high-producing S288c strains than CEN.PK strains.

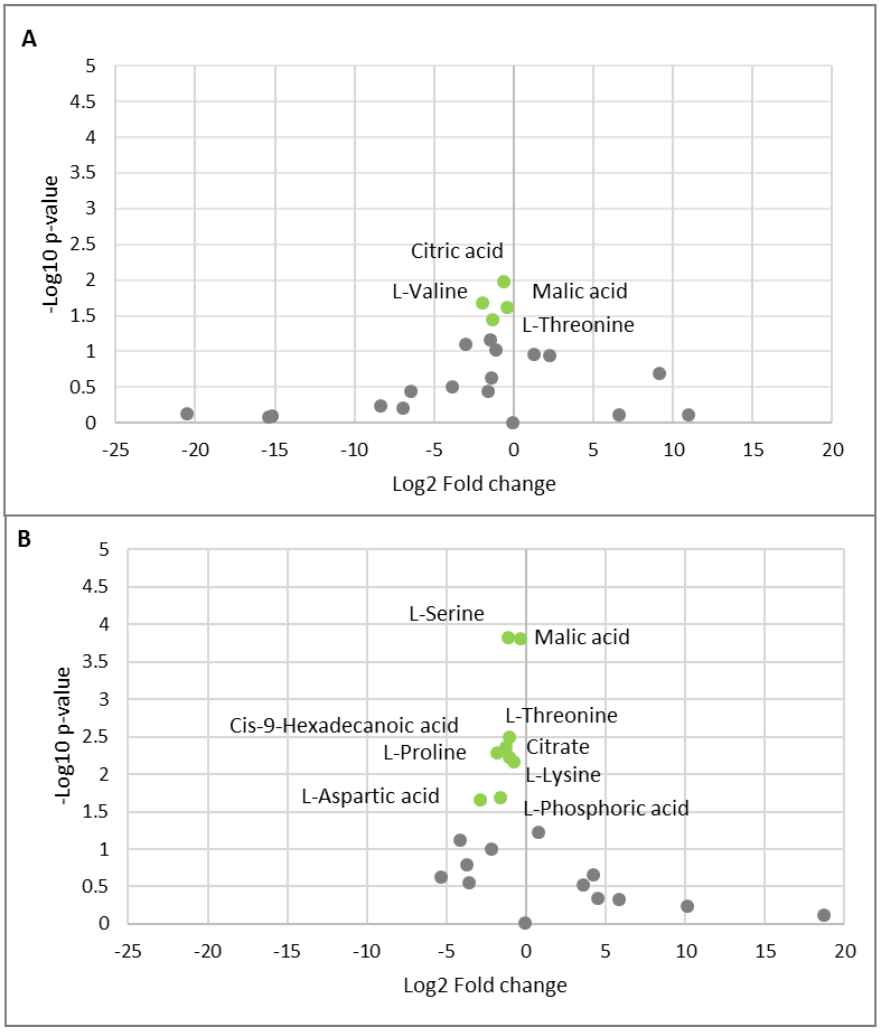


Figure 4. Volcano plot based on the statistical significance of the t-test and fold change from the comparison of the reporter metabolites between optimized and non-optimized strain on each background. A. CEN.PK strains, B. S288c strains.

Discussion

In our study, engineered high-producer strain with CEN.PK background gave 20-25% higher *p*-CA titers in batch and continuous cultivations in comparison with S288c strain engineered identically. Moreover, the CEN.PK-producer had a 25% higher μ_{max} than S288c-producer. Transcriptome analysis showed that the CEN.PK strain was less affected by engineering towards higher *p*-CA production than the S288c strain, as the number of significantly up/down-regulated genes was correspondingly 652 and 1927 amongst others, strain S288c had downregulations in gene sets involved in amino acid and protein biosynthesis. This suggests that CEN.PK may be a better platform strain for production of aromatic compounds than S288c strain.

The transcriptome analysis also revealed downregulations in transport functions in the engineered strains of both backgrounds, which could be a response to the stress triggered by production of *p*-CA. Previous studies on plasma membrane integrity and ethanol stress in *S. cerevisiae* have reported that yeast cells react to chemical stress by downregulating the transport of some metabolites and by decreasing gene expression in energy-demanding processes (Stanley *et al*., 2010, Madeira *et al*., 2009 and Leao & Van uden, 1984).

For the engineered S288c strain, metabolomics analysis revealed lower concentrations of phosphoric acid, L-ornithine and glutamic acid, which correlates with downregulation of the gene sets involved in the synthesis of proteins and amino acids. Interestingly previous research on *S. cerevisiae* stress responses to oxidative stress had shown diminution in the synthesis

of proteins as a prevention mechanism under potentially error-prone conditions (Shenton *et al*., 2003).

We observed higher accumulation of glycerol by the engineered strains than in non-optimized strains. Synthesis of glycerol is known as an important factor in the control of osmoregulation and of redox balance (Hohmann *et al*., 2007; Muzzei *et al*., 2009). Another interesting metabolite is cis-9-hexadecanoic acid; this metabolite had higher concentrations in the S288c strain in comparison to CEN.PK strain; we propose that the stress originating from the production of *p*-CA may trigger the synthesis of this fatty acid in the S288c strains. It had been previously reported that genetically engineered strains with a higher concentration of cis-9-hexadecanoic acid were more tolerant to temperature and oxidative stress (Steels *et al*., 1994; Jamieson *et al*., 1998).

In the particular case of the reporter metabolite L-tyrosine, the downregulation of the transport activity is accompanied by the upregulation of *ARO9*. The upregulation of *ARO9* can be a response to the increment of the flux of intermediate aromatic compounds toward L-tyrosine production and can be positive for the production of *p*-CA. The downregulation in transport functions can be a response to the high concentration of aromatic compounds, for example, Hueso *et al*. (2012) reported downregulation in the transport of leucine under intracellular acidification and a subsequent improvement in acid growth after overexpression of the transporter gene *BAP2*.

In conclusion, we found that the transport of amino acids and sugars were highly downregulated probably as a response to the stress produced by the secreted aromatic compounds. Both the transcriptome and metabolomics analysis showed that strain CEN.PK was less affected than S288c by increased *p*-CA production and hence suits better as a cell factory for production of this compound.

As a strategy for further improvement of *p*-CA cell factory, we suggest making a combinatorial evaluation of the effect of the upregulation and downregulation of the genes that we found were significantly affected in their expression levels, such as aminotransferases *ARO9* and *ARO8* and transport genes *AGP1*, *GAP1, BAP3*, *BAP2* and *TAT1*. This approach will allow identifying if the downregulations and upregulations of these genes are a response to stress with a negative effect on the production of aromatic secondary metabolites or if it was a physiological adjustment to the flux of intermediate compounds toward L-tyrosine with a positive effect on the production of aromatic secondary metabolites.

Acknowledgements

This work was financed by the Novo Nordisk Foundation. We thank Anna Koza for performing the RNA sequencing, Leif Väremo for his advice for the use of the packages Piano and Kiwi, Eduard Kerkoven for his suggestions on the transcriptome data analysis. Also we thank Klara Bojanovic, Kanchana R Kildegaard and Arun Rajkumar for their advice on RNA isolation.

Contributions

AR, IB and JN conceived and designed the study and analyzed the results. AR, YC and SK did the experimental work. AR, YC, EO and SK processed and analyzed the data. AR and IB drafted the manuscript and all authors read, edited and approved the final manuscript.

Supplementary material

Supplementary Table 1. Primers and Biobricks used in this study.

BIOBRICKS

Name	Template for PCR	FW primer	RV Primer
BB0380 (Fj_tal<-)	*Flavobacterium johnsoniae* codon-optimized synthetic gene	ID1691	ID1692
BB0501 (Ec_AroL->)	Genomic DNA of EcoMG1655 ATCC 31884	ID6785	ID6786
BB0361 (Sc_Aro7_G141S<-)	p0761(pESC-URA-ARO7pm)	ID1398	ID1399
BB0364 (Sc_Aro4_K229L->)	p0775 (pESC-HIS-ARO4pm)	ID1396	ID1397
BB1163(Sc_MET15_2/3_Up)	Genomic DNA of CEN.PK102.5B	ID11784	ID11785
BB1164(Sc_MET15_2/3_Down)	Genomic DNA of CEN.PK102.5B	ID11786	ID11787

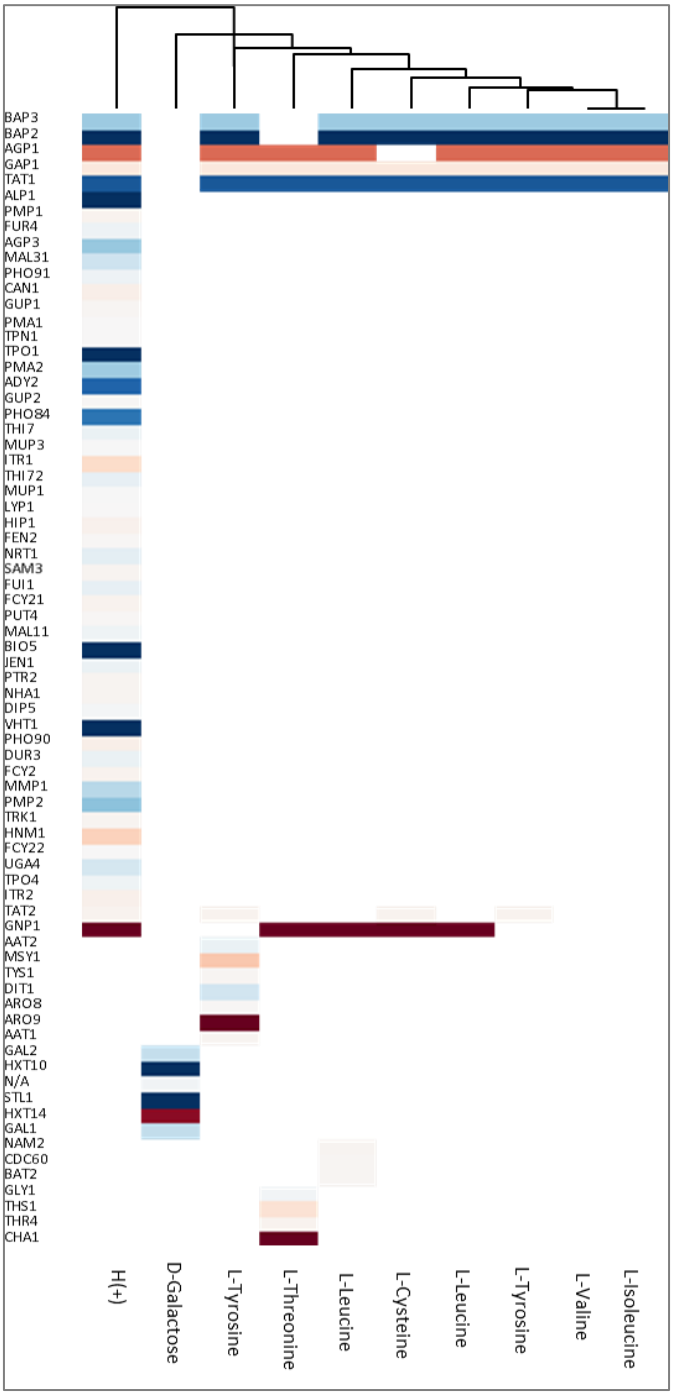
PRIMERS

For BioBricks construction

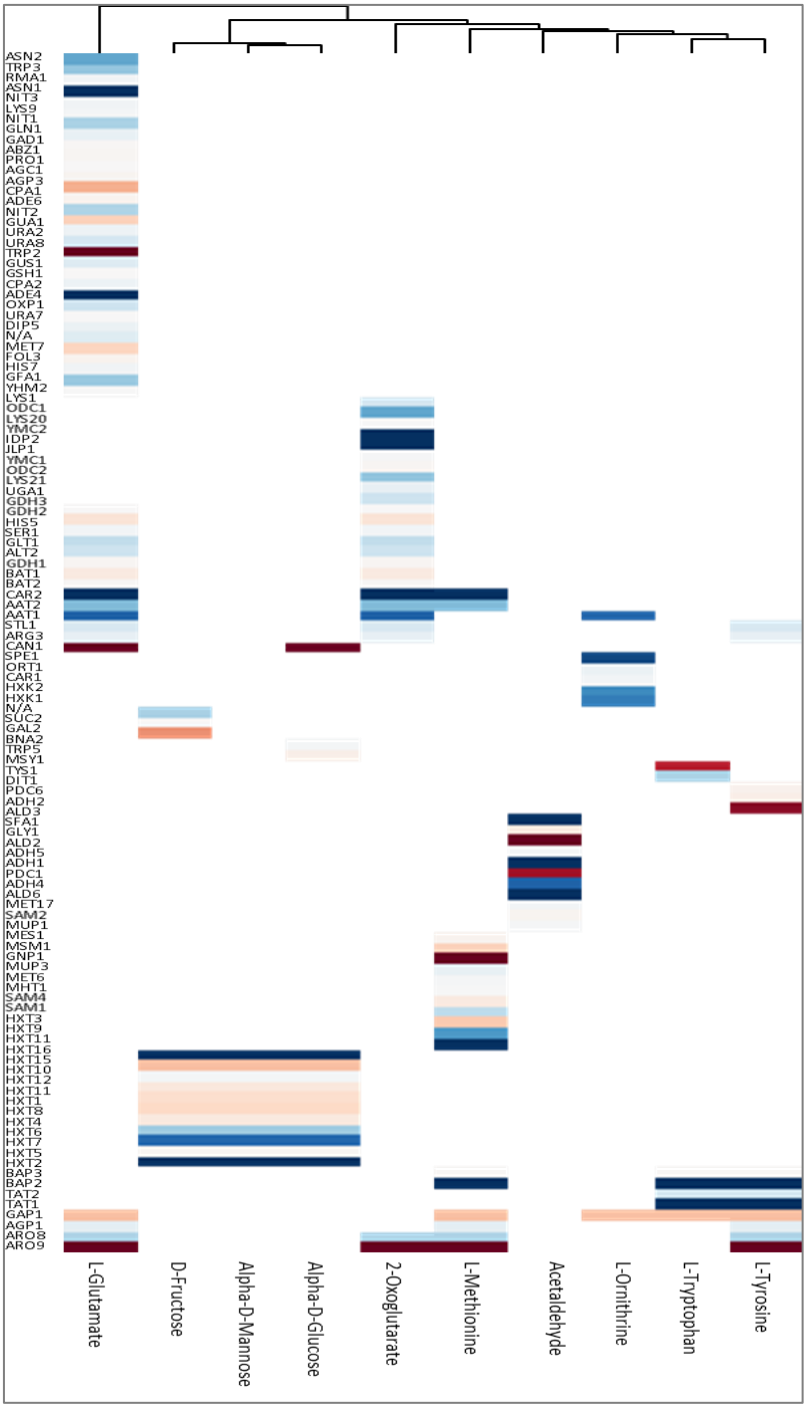
ID	Name	Sequence (5'-3')
1691	Fj_TAL_1_fw	AGTGCAGGUAAAACAATGAACACCATCAACGAATATCTGAGC
1692	Fj_TAL_1_rv	CGTGCGAUTTAATTGTTAATCAGGTG
6785	Ec_aroL_2_fw	ATCTGTCAUAAAACAATGACACAACCTCTTTTTCTGA
6786	Ec_aroL_2_rv	CACGCGAUTCAACAATTGATCGTCTGTGC
1398	Sc_ARO7_1_fw	AGTGCAGGUAAAACAATGGATTTCACAAAACCAGAAAC
1399	Sc_ARO7_1_rv	CGTGCGAUTCACTCTTCCAACCTTCTTAGCAAG
1396	Sc_ARO4_2_fw	ATCTGTCAUAAAACAATGAGTGAATCTCCAATGTTCG
1397	Sc_ARO4_2_rv	CACGCGAUTCATTTCTTGTTAACTTCTCTTCTTTG
11784	Sc_Met15_U1_Fw	GTTCTCGTCGAATGCTAGGTC
11785	Sc_Met15_U1_Rv	TTCACCGTGTCTTTCAGCTC
11786	Sc_Met15_U2_Fw	TCGAGGCTAGATTTGTTGAAGG
11787	Sc_Met15_U2_Rv	TTAATGACGTTCGGCTGGAG

For verification of correct insertion in the chromosomes

ID	Name	Sequence (5'-3')
904	Sc_X-3-out-seq_rv	CCGTGCAATACCAAAATCG
906	Sc_X-4-out-seq_rv	GACGGTACGTTGACCAGAG
912	XI-3- down-out-sqID912	CACATTGAGCGAATGAAACG
2220	Sc_ColoPCR_fw	CCTGCAGGACTAGTGCTGAG



Supplementary Figure 1. Heat map from the network analysis for the CEN.PK strains.



Supplementary Figure 2. Heat map from the network analysis for the S288c strains.

References

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., Boeke, J.D., 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14, 115–132. doi: 10.1002/(SICI) 1097-0061 (19980130) 14:2 <115:AID-YEA204 >3.0. CO; 2-2

Bourgaud, F., Gravot, A., Milesi, S., Gontier, E., 2001. Production of plant secondary metabolites: A historical perspective. Plant Sci. 161, 839–851. doi:10.1016/S0168-9452(01)00490-3

Canelas, A.B., Ten Pierick, A., Ras, C., Seifar, R.M., Van Dam, J.C., Van Gulik, W.M., Heijnen, J.J., 2009. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. Anal. Chem. 81, 7379–7389. doi:10.1021/ac900999t

Curran, K.A., Alper, H.S., 2012. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. Metab. Eng. 14, 289–297. doi: 10.1016/j.ymben.2012.04.006

Entian, K.-D., Kötter, P.,2007.25Yeast genetic strain and plasmid collections. Method Microbiol. 36, 629–666.

Germann, S.M., Baallal Jacobsen, S.A., Schneider, K., Harrison, S.J., Jensen, N.B., Chen, X., Stahlhut, S.G., Borodina, I., Luo, H., Zhu, J., Maury, J., Forster, J., 2016. Glucose-based microbial production of the hormone melatonin in yeast *Saccharomyces cerevisiae*. Biotechnol. J. doi:10.1002/biot.201500143

Gietz, B.R.D., Woods, R. a, Peg, D.N. a, 2002. [4] Transformation of Yeast by Lithium Acetate / Single-Stranded Carrier DNA / Polyethylene Glycol Method 350, 87–96.

Hohmann, S., Krantz, M., Nordlander, B., 2007. Yeast Osmoregulation, Methods in Enzymology. Elsevier Masson SAS. doi:10.1016/S0076-6879(07)28002-4

Hueso, G., Aparicio Sanchis, R., Montesinos, C., Lorenz, S., Murguía, J.R., Serrano, R., 2011. A novel role for protein kinase Gcn2 in yeast tolerance to intracellular acid stress. Biochem. J. 441, 255–264. doi:10.1042/BJ20111264

Jamieson, D.J., 1998. Oxidative stress responses of the yeast *Saccharomyces cerevisiae*. Yeast 14, 1511–1527. doi:10.1002/(SICI)1097-0061(199812)14:16<1511::AIDYEA356> 3.0. CO; 2-S

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., Borodina, I., 2014. EasyClone: method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. FEMS Yeast Res. 14, 238–248. doi: 10.1111/1567-1364.12118

Kim, I.K., Roldão, A., Siewers, V., Nielsen, J., 2012. A systems-level approach for metabolic engineering of yeast cell factories. FEMS Yeast Res. 12, 228–248. doi:10.1111/j.1567-1364. 2011.00779.x

Khoomrung, S., Martinez, J.L., Tippmann, S., Jansa-Ard, S., Buffing, M.F., Nicastro, R., Nielsen, J., 2015. Expanded metabolite coverage of *Saccharomyces cerevisiae* extract through improved chloroform/methanol extraction and tert-butyldimethylsilyl derivatization. Anal. Chem. Res. 6, 9–16. doi:10.1016/j.ancr. 2015.10.001

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J. a, Pronk, J.T., Daran, J.-M., 2012. De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae* Microb. Cell Fact. 11, 155. doi:10.1186/1475-2859-11-155

Krivoruchko, A., Nielsen, J., 2015. Production of natural products through metabolic engineering of *Saccharomyces cerevisiae*. Curr. Opin. Biotechnol. 35, 7–15. doi:10.1016/j.copbio.2014.12.004

Leao, C.n van Uden, N., 1984. Effects of ethanol and other alkanols on the general amino acid permease of *Saccharomyces cerevisiae*. Biotechnol Bioeng 26, 403–405.

Leonard, E., Runguphan, W., O'Connor, S., Prather, K.J., 2009. Opportunities in metabolic engineering to facilitate scalable alkaloid production. Nat. Chem. Biol. 5, 292–300. doi:10.1038/nchembio.160

Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I., Nielsen, J., 2015. Author' s Accepted Manuscript De novo production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae* To

appear in: Metabolic Engineering. Metab. Eng. 32, 1–11. doi:10.1016/j.ymben.2015.08.007

Madeira, A., Leitão, L., Soveral, G., Dias, P., Prista, C., Moura, T., Loureiro-Dias, M.C., 2010. Effect of ethanol on fluxes of water and protons across the plasma membrane of *Saccharomyces cerevisiae*. FEMS Yeast Res. 10, 252–258. doi:10.1111/j.1567-1364. 2010.00607.x

Muzzey, D., Gomez-Uribe, C.A., Mettetal, J.T., van Oudenaarden, A., 2009. A Systems-Level Analysis of Perfect Adaptation in Yeast Osmoregulation. Cell 138, 160–171. doi:10.1016/j.cell.2009.04.047

Nielsen, J., Keasling, J.D., 2016. Engineering Cellular Metabolism. Cell 164, 1185–1197. doi:10.1016/j.cell.2016.02.004

Nijkamp, J.F., van den Broek, M., Datema, E., de Kok, S., Bosman, L., Luttik, M.A., Daran-Lapujade, P., Vongsangnak, W., Nielsen, J., Heijne, W.H.M., Klaassen, P., Paddon, C.J., Platt, D., Kötter, P., van Ham, R.C., Reinders, M.J.T., Pronk, J.T., de Ridder, D., Daran, J.-M., 2012. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. Microb. Cell Fact. 11, 36. doi:10.1186/1475-2859-11-36

Österlund, T., Nookaew, I., Bordel, S., Nielsen, J., 2013. Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. BMC Syst. Biol. 7, 36. doi:10.1186/1752-0509-7-36

Otero, J.M., Cimini, D., Patil, K.R., Poulsen, S.G., Olsson, L., Nielsen, J., 2013. Industrial Systems Biology of *Saccharomyces cerevisiae* Enables Novel Succinic Acid Cell Factory. PLoS One 8, 1–10. doi: 10.1371/journal.pone.0054144

Otero, J.M., Vongsangnak, W., Asadollahi, M.A., Olivares-Hernandes, R., Maury, J., Farinelli, L., Barlocher, L., Osterås, M., Schalk, M., Clark, A., Nielsen, J., 2010. Whole genome sequencing of *Saccharomyces cerevisiae*: from genotype to phenotype for improved metabolic engineering applications. BMC Genomics 11, 723. doi:10.1186/1471-2164-11-723

Park, J.H., Lee, K.H., Kim, T.Y., Lee, S.Y., 2007. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome

analysis and in silico gene knockout simulation. Proc. Natl. Acad. Sci. U. S. A. 104, 7797–802. doi:10.1073/pnas.0702609104

Rodriguez, A., Kildegaard, K.R., Li, M., Borodina, I., Nielsen, J., 2015. Establishment of a yeast platform strain for production of *p*-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. Metab. Eng. 31, 181-188. doi:10.1016/j.ymben. 2015.08.003

Santos, C.N.S., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab. Eng. 13, 392–400. doi:10.1016/j.ymben. 2011.02.002

Stahlhut, S.G., Siedler, S., Malla, S., Harrison, S.J., Maury, J., Neves, A.R., Forster, J., 2015. Assembly of a novel biosynthetic pathway for production of the plant flavonoid fisetin in *Escherichia coli*. Metab. Eng. 31, 84–93. doi:10.1016/j.ymben.2015.07.002

Stanley, D., Bandara, A., Fraser, S., Chambers, P.J., Stanley, G.A., 2010. The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. J. Appl. Microbiol. 109, 13–24. doi:10.1111/j.1365-2672. 2009.04657.x

Steels, E.L., Learmonth, R.P., Watson, K., 1994. Stress tolerance and membrane lipid unsaturation in *Saccharomyces cerevisiae* grown aerobically or anaerobically. Microbiology 140 (Pt 3, 569–76. doi:10.1099/00221287-140-3-569

Shenton, D., Grant, C.M., 2003a. Protein S-thiolation targets glycolysis and protein synthesis in response to oxidative stress in the yeast *Saccharomyces cerevisiae*. Biochem. J. 374, 513–519. doi:10.1042/BJ20030414

Strucko, T., Magdenoska, O., Mortensen, U.H., 2015. Benchmarking two commonly used *Saccharomyces cerevisiae* strains for heterologous vanillin-β-glucoside production. Metab. Eng. Commun. 2, 99–108. doi:10.1016/j.meteno. 2015.09.001

Trantas, E., Panopoulos, N., Ververidis, F., 2009. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. Metab. Eng. 11, 355–66. doi:10.1016/j.ymben. 2009.07.004

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–78. doi:10.1038/nprot.2012.016

Värremo, L., Nielsen, J., Nookaew, I., 2013. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic Acids Res. 41, 4378–4391. doi:10.1093/nar/gkt111

Väremo, L., Gatto, F., Nielsen, J., Varemo, L., Gatto, F., Nielsen, J., 2014. Kiwi: a tool for integration and visualization of network topology and gene-set analysis. BMC Bioinformatics 15, 408. doi:10.1186/s12859-014-0408-9

Verduyn, C., Postma, E., Scheffers, W.A., van Dijken, J.P., 1992. Effect of Benzoic Acid on Metabolic Fluxes in Yeasts: Yeast 8, 501–517.

Wink, M., 2010. Annual plant reviews volume 40, Biochemistry of plant Secondary Metabolites 40. doi:10.1002/9781444320503

Yin, S., Lang, T., Xiao, X., Liu, L., Sun, B., Wang, C., 2015a. Significant enhancement of methionol production by co-expression of the aminotransferase gene ARO8 and the decarboxylase gene ARO10 in *Saccharomyces cerevisiae*. FEMS Microbiol. Lett. 362, 1–7. doi:10.1093/femsle/fnu043

Yin, S., Zhou, H., Xiao, X., Lang, T., Liang, J., Wang, C., 2015b. Improving 2-Phenylethanol production via Ehrlich pathway using genetic engineered *Saccharomyces cerevisiae* strains. Curr. Microbiol. 70, 762–767. doi:10.1007/s00284-015-0785-y

Yoon, S.H., Han, M.-J., Jeong, H., Lee, C.H., Xia, X.-X., Lee, D.-H., Shim, J.H., Lee, S.Y., Oh, T.K., Kim, J.F., 2012. Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. Genome Biol. 13, R37. doi:10.1186/gb-2012-13-5-r37

4. Synthesis of flavonoids in yeast

Metabolic engineering of yeast for fermentative production of flavonoids

Angelica Rodriguez[1*], Tomas Strucko[1*], Steen G. Stahlhut[1], Mette Kristensen[1], Daniel Killerup Svenssen[1], Jochen Forster[1], Jens Nielsen[1,2], Irina Borodina[1†]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle allé 6, 2970 Hørsholm, Denmark

[2]Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE412 96 Gothenburg, Sweden

* These authors contributed equally to this work.

† Corresponding author e-mail address: irbo@biosustain.dtu.dk

Abstract

Flavonoids, secondary metabolites produced in plants and fungi, offer numerous health benefits owing to their antioxidant, anti-inflammatory, anti-carcinogenic and other biological activities. Because of the low abundance of these compounds in natural sources and challenges with extraction from plant material, there is a lot of interest in producing flavonoids by fermentation using cell factories. Here we engineered *Saccharomyces cerevisiae* to produce a range of flavonoids: naringenin, liquiritigenin, kaempferol, resokaempferol, quercetin and fisetin in defined medium with glucose as the sole carbon source. Up to eight heterologous genes from plants were overexpressed in a yeast platform strain, previously optimized for production of the flavonoid precursor *p*-coumaric acid. For the first time resokaempferol and fisetin were produced by yeast. The concentration of most of the compounds reached 1-5 mg L^{-1} in the medium, but kaempferol

and quercetin were secreted at higher titers of about 20 mg L^{-1}. The engineered strains simultaneously accumulated 10-15 mg L^{-1} of *p*-coumaric acid, suggesting that the *p*-coumaroyl-CoA ligase *4CL* and possibly chalcone synthase *CHS* were limiting the carbon flux into flavonoids. Overexpression of an additional copy of *CHS* and *CHR* genes indeed increased the titer of fisetin from 1.65 ± 0.10 mg L^{-1} to 2.28 ± 0.07 mg L^{-1}.

This work demonstrates the potential of flavonoid-producing yeast cell factories.

Key Words

Flavonoids, naringenin, liquiritigenin, kaempferol, resokaempferol, quercetin, fisetin, *Saccharomyces cerevisiae.*

Introduction

Flavonoids are aromatic secondary metabolites naturally synthesized by plants and fungi from aromatic amino acids L-phenylalanine and L-tyrosine. The structure of flavonoids is characterized by two phenolic rings and one heterocyclic ring. The main differences between flavonoids are related to the hydroxylation patterns, the position of the second aromatic ring and the saturation of the heterocyclic ring (Grotewold *et al.,* 2008). Flavonoids play a fundamental role in the physiology of plants; their natural functions include UV protection, reduction of oxidative damage in cells, and antibacterial effects (Cushnie and Lamb 2011; Roberts and Paul 2006; Agati *et al.,* 2011).

Research on human cells showed positive properties of flavonoids as reducing agents and protectors of oxidative reactions in age-related diseases, cancer and cardiovascular diseases (Woelfle *et al.,* 2010; Khoo *et al.,* 2010; Bulzomi *et al.,* 2012). Moreover, some flavonoids, e.g., naringenin, have neuroprotective and antioxidant properties. Liquiritigenin has been reported as a protective agent against oxidative stress in osteoblasts, and

some liquiritigenin derivatives show antitumor and antidiabetic activity (Choi *et al.,* 2012; Wedick *et al.,* 2012; Liu *et al.,* 2012; Raza *et al.,* 2015; Hamalainen *et al.,* 2015). Kaempferol, quercetin and fisetin showed anti-cancer, cardio-protective and anti-inflammatory effects (Hamalainen *et al.,* 2015; Chen and Chen, 2013; Nabavi *et al.,* 2012). Experiments with rats showed that doses of 50 mg kg^{-1} of quercetin inhibit the migration of melanoma cells. Other experiments with humans showed that doses of 150 mg day^{-1} of quercetin had positive effects on the health of people with high risk of cardiovascular disease (Caltagirone *et al.,* 2000; Khan *et al.,* 2008; Mukhtar *et al.,* 2015). Fisetin has neuroprotective, neurotrophic and anti-amyloid properties, which makes it a promising therapeutic agent for neurodegenerative disorders such as Huntington and Alzheimer diseases (Maher *et al.,* 2011; Currais *et al.,* 2014).

The main obstacles for large-scale production of flavonoids in plants are the long culture periods, requirements for specific cultivation conditions and low abundance. Additionally, extraction and purification processes add cost and result in product loss and degradation (Wang *et al.,* 2011; Routray and Orsat, 2012). The concentration of flavonoids in different varieties and sources of fruits oscillate between 30-4000 mg Kg^{-1} of dry weight (Hertog *et al.,* 1992; Paganga *et al.,* 1999; Miean and Suhaila, 2001; Crozier *et al.,* 1997). This means that for the production of 1 kg of flavonoids, it is required to process 250-1000 Kg of dry weight of fruits or vegetables.

For a sustainable supply of flavonoids, it can be an advantage to engage genetically engineered microbial cell factories, e.g., *E. coli* or *S. cerevisiae*. Koopmann *et al. (2012)* reported the production of 108.90 mg L^{-1} of naringenin from glucose using an engineered yeast strain. Furthermore, feeding naringenin to engineered cells allowed production of other flavonoids, such as genistein, kaempferol and quercetin (Trantas *et al.,* 2009). Production of naringenin in *E. coli* reached 29 mg L^{-1} using glucose

as carbon source, while kaempferol and quercetin have been produced using *p*-coumaric acid as precursor and fisetin using L-tyrosine as a precursor (Santos *et al.,* 2011; Leonard *et al.,* 2006; Stahlhut *et al.,* 2015).

Here we established a yeast cell factory for *de novo* production of flavonoids from glucose. The target flavonoids included: naringenin, liquiritigenin, kaempferol, resokaempferol, quercetin, and fisetin. Intermediate compounds and by-products were quantified to characterize regulated steps in the pathways and to identify targets for future improvements of the flavonoid-producing yeast cell factories.

Methods

Plasmid construction

The plasmids with the biosynthetic pathways for flavonoids production were constructed using EasyClone 2.0 integrative plasmids with auxotrophic and dominant selection markers (Stovicek *et al.,* 2015) that are targeting well-defined integration sites previously described by (Mikkelsen *et al*. 2012).

The genes used for flavonoids production were: 4-coumaroyl-CoA ligase from *Petroselinum crispum* (*4CL*), chalcone synthase from *Petunia hybrida* (*CHS*), chalcone reductase from *Astragalus mongholicus* (*CHR*), chalcone isomerase from *Medicago sativa* (*CHI*), flavanone 3-hydroxylase from *Astragalus mongholicus* (*F3H*), flavonol synthase from *Arabidopsis thaliana* (*FLS*), cytochrome P450 reductase from *Catharanthus roseus* (*CPR*), a cytochrome P450 flavonoid monooxygenases from *Fragaria ananassa* and from *Petunia hybrida* (*FMO*) (Figure 2). All the genes were synthesized by GeneArt (LifeTechnologies) in codon-optimized versions for *E. coli* as stated in Stahlhut *et al.,* (2015*)*.

DNA fragments (BioBricks), encoding genes or promoters, were amplified by PCR using primers and templates as described in Supplementary Tables 1

and 2. Individual BioBricks were assembled into the integrative plasmids by USER cloning as described previously (Stovicek *et al.*, 2015). Finally, the resulting vectors (Table 1) were verified by sequencing.

Table 1. Plasmids used in this study

Parental Plasmids		
Plasmid ID	**Genotype**	**Source**
pCfB2399	pXI-5-lox P-amdSYMsyn, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
pCfB2197	pXII-1-lox P-NatMXsyn3, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
pCfB2225	pXII-2-lox P-KanMXsyn, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
pCfB2337	pXII-5-lox P-HPHMXsyn, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
pCfB2224	pXI-2-lox P-KanMXsyn, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
pCfB2855	pXII-2-lox P-amdSYM, P_{TEF1}-T_{ADH1}, P_{PGK1}-T_{CYC1}	Stovicek *et al.,* 2015
Integrative plasmids for pathway expression		
Plasmid ID	**Genotype**	**Source**
pCfB0848	pXI-2-loxP-URA3, P_{TEF1}-CYB5 T_{ADH1}, P_{PGK1}-ATR2- T_{CYC1}	Li *et al.,* 2016
pCfB1018	pXI-5-loxP-HIS5, P_{TEF1}-PAL2-$T_{ADH1,}$ P_{PGK1}-C4H T_{CYC1}	Li *et al.,* 2016
pCfB3437	pXI-5-lox P-amdSYMsyn, P_{TEF1}-Ph_CHS-T_{ADH1}, P_{PGK1}-Pc_4CL-T_{CYC1}	This study
pCfB2879	pXII-1-lox P-NatMXsyn3, P_{TEF1}-Ms_CHI-T_{ADH1}, P_{PGK1}-Am_CHR-T_{CYC1}	This study
pCfB2893	pXII-1-lox P-NatMXsyn3, P_{TEF1}-Ms_CHI-T_{ADH1}	This study
pCfB3643	pXII-2-lox P-KanMXsyn, P_{TEF1}-At_FLS-T_{ADH1}, P_{PGK1}-At_F3H-T_{CYC1}	This study
pCfB3654	pXII-5-lox P-HPHMXsyn, P_{TEF1}-Ph_FMO- Cr_CPR-T_{ADH1}	This study
pCfB3655	pXII-5-lox P-HPHMXsyn, P_{TEF1}-Fa_FMO-Cr_CPR-T_{ADH1}	This study
pCfB4753	pXII-2-lox P-amdSYM, P_{TEF1}-Ph_CHS-T_{ADH1}, P_{PGK1}-Pc_4CL-T_{CYC1}	This study
pCfB4754	pXI-2-lox P-KanMXsyn, P_{TEF1}-At_FLS-T_{ADH1}, P_{PGK1}-At_F3H-T_{CYC1}	This study
pCfB4856	pXII-4-lox P-BleSyn, P_{TEF1}-Ph_CHS-Am_CHR- T_{ADH1}	This study
pCfB4857	pXII-4-lox P-BleSyn, P_{TEF1}-Ph_CHS-T_{ADH1}, P_{PGK1}- Am_CHR -T_{CYC1}	This study

Strains construction

The background strains for this research were three engineered *S. cerevisiae* strains (Table 2). The ST4757 strain was engineered to produce *p*-coumaric acid using the phenylalanine ammonia-lyase (PAL) pathway and

the strains ST4069 and ST2645 produced *p*-coumaric acid via the tyrosine ammonia-lyase (TAL) pathway. The background strains were transformed with *Not*I-linearized integrative vectors using lithium acetate protocol (Gietz and Woods, 2002). The transformants were selected on synthetic drop-out medium (Sigma-Aldrich), selecting for *URA3*, *HIS5* and *LEU2* markers. For selection on acetamide, the media contained 0.17 % yeast nitrogen base without amino acids and ammonium sulfate and 6.6 g L^{-1} of potassium sulfate and 0.6 g L^{-1} acetamide. For selection of dominant markers *NatMX*, *KanMX, BleMX* or *HphMX*, the medium was supplemented with 100 mg L^{-1} nourseothricin, 200 mg L^{-1} G418 disulfate salt, 10 mg L^{-1} phleomycin or 200 mg L^{-1} hygromycin B respectively. Correct integration of the vectors was verified by yeast colony PCR. The yeast strains used in this study are listed in Table 2.

Media and cultivations

For selection of yeast transformants and for routine cultivations we prepared synthetic complete (SC) medium as well as drop-out media (SC-Ura, SC-Leu, SC-His) and agar plates using pre-mixed drop-out powders from Sigma-Aldrich. Synthetic fed-batch medium for *S. cerevisiae* M-Sc.syn-1000 (FIT) was purchased from M2P Labs GmbH (Germany). The medium was supplemented with vitamins solution (final 1% v/v) and the enzyme mix (final concentration 0.5% v/v) immediately before use.

For the quantification of production of flavonoids for each strain, we tested five single colonies originating from independent transformants. The colonies were inoculated in 0.5 ml drop-out SC liquid medium without uracil, histidine, and/or leucine in 96-deep well microtiter plates with air-penetrable lid (EnzyScreen, The Netherlands). The plates were incubated at 30°C with 250 rpm agitation at 5 cm orbit cast overnight. 50 µl of the overnight cultures were used to inoculate 0.5 ml synthetic fed-batch medium in a 96-deep well plate. Fermentation was carried out for 72 hours at the same conditions as

before. At the end of cultivation, samples for metabolite analysis and optical density (OD) were taken. OD_{600} was estimated as follows: 10 µl of fermentation broth was mixed with 190 µl water in a 96 well microtiter plate and absorbance was measured at 600 nm wavelength in microplate reader BioTek Synergy MX (BioTek). 200 µl of the culture were mixed with 200 µl of absolute ethanol, the mixture was centrifuged at 2,272 g for 15 minutes and 250 µl of supernatant were analyzed for flavonoids concentration by HPLC.

Table 2. Strains used in this study

Parent strains				
Strain ID	**Genotype**			**Source**
ST4069	*Mata* P_{TEF1}->Fj_*TAL*			Rodriguez *et al.,* 2015
ST2645	*Mata* P_{TEF1}->Fj_*TAL*, P_{TEF1}->Sc_*ARO7*^{G141S}, P_{PGK1}->Sc_*ARO4*^{K229L}, *Δaro10 Δpdc5*			Rodriguez *et al.,* 2015
ST4757	*Matα* P_{TEF1}->*PAL2,* P_{PGK1}->*C4H,* P_{PGK1}->*CYB5,* P_{TEF1}->*ATR2*			This study
Strains transformed with integrative plasmids				
Strain ID	**Parent strain**	**Integrated Plasmids**	**Integrated flavonoid pathway genes**	**Source**
ST5066	ST4069	pCfB3437, pCfB2893	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI*	This study
ST5067	ST4069	PCFB3437, pCfB2879	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*CHR*	This study
ST5068	ST2645	PCfB3437, pCfB2893	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI*	This study
ST5069	ST2645	PCfB3437, pCfB2879	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*CHR*	This study
ST5070	ST5068	pCfB3643	P_{PGK1}->*4CL,* P_{PGK1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS*	This study
ST5071	ST5069	pCfB3643	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS*	This study
ST5072	ST5070	pCfB3654	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS, P_{TEF1}->Fa_FMO CPR*	This study
ST5073	ST5070	pCfB3655	P_{TEF1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS,* P_{TEF1}->*Ph_FMO-CPR*	This study
ST5074	ST5071	pCfB3654	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*CHR,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS,* P_{TEF1}->*Fa_FMO-CPR*	This study
ST5075	ST5071	pCfB3655	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*CHR,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS,* P_{TEF1}->*Ph_FMO-CPR*	This study
ST4972	ST4757	PCfB4753, PCfB4754, PCfB2879, PCfB3654	P_{PGK1}->*4CL,* P_{TEF1}->*CHS,* P_{TEF1}->*CHI,* P_{PGK1}->*CHR,* P_{PGK1}->*F3H,* P_{TEF1}->*FLS,* P_{TEF1}->*Ph_FMO-CPR*	This study
ST5401	ST5074	pCfB4856	P_{TEF1}->*CHS:CHR (fusion)*	This study
ST5402	ST5074	pCfB4587	P_{TEF1}->*CHS,* P_{PGK1}->*CHR*	This study

Quantification and identification of flavonoids

Quantification of flavonoids was performed on a Dionex Ultimate 3000 HPLC equipped with a Discovery HS F5 150mm X 4.6 mm column (particle size 5µm) connected to a UV detector (277, 290, 333 and 370nm). Samples were analyzed using a gradient method with two solvents: 10 mM ammonium formate pH 3.0 (A) and acetonitrile (B). For *p*-coumaric acid, naringenin and liquiritigenin detection, a flow rate of 1.5 ml min^{-1} was used. The program started with 5% of solvent B (0-0.5 min), after which its fraction was increased linearly from 5% to 60% (0.5-7 min), then the fraction was maintained at 60% (7-9.5 min), after that the fraction was decreased from 60% to 5% (9.5-9.6 min), finally, the fraction was maintained at 5% (9.6 to 12 min). *p*-Coumaric acid was detected at 5.6 min (333), liquiritigenin at 6.8 (277) and naringenin at 7.5 min (290 nm). For kaempferol, resokaempferol, quercetin and fisetin detection, a flow rate of 1.5 ml min^{-1} was used. The program started with 20% of solvent B (0-2 min), after which its fraction was increased linearly from 20% to 45% (2-20 min), then the fraction was decreased from 45% to 20% (20-22 min) and maintained at 20% for 2 minutes (22-24 min). Fisetin was detected at 10.3 min (333 nm), liquiritigenin at 12.8 min (277 nm), resokaempferol at 13.3 min (370 nm) and quercetin at 13.9 min (370 nm).

The pure compounds *p*-coumaric acid, naringenin, kaempferol, quercetin and fisetin were purchased from Sigma Aldrich Co. (Denmark); liquiritigenin was purchased from Tocris Bioscience (United Kingdom) and and resokaempferol was purchased from Extrasynthese (France). The compounds were used to generate calibration curves; the areas were integrated with Chromeleon 7 and used for quantification. The flavonoids were identified by comparing the retention times and UV absorbance spectra of the samples with authentic compounds. For all the strains five biological replicates were analyzed.

Identification of fisetin

For fisetin identification, the samples were dried at room temperature under reduced pressure using a Centrifugal Vacuum Concentrator (Savant Speed Vacs Concentrator, Thermofisher Scientific, Waltham Ma), followed by reconstitution using a 0.1% solution of formic acid in LC-MS grade acetonitrile. 20 µl of each sample were then injected and analyzed on a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, San Jose, CA) connected to an Orbitrap Fusion Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA). The UHPLC used a Supelco Discovery HS F5-3, 15cm x 2.1mm, 3 µm column. Temperature was 30°C and flow rate was 0.6 mL/min with a mobile phase of 95% 0.1 % formic acid (mobile phase A) and 5% acetonitrile with 0.1% formic acid (mobile phase B) for 2 min followed by a linear gradient to 95% mobile phase B over 12 minutes. This gradient was held for 2 min after which it was changed immediately to 95% mobile phase A and 5% mobile phase B and held for 6 min. The sample was passed on to the MS equipped with a heated electrospray ionization source (HESI) in a positive-ion mode with nitrogen as nebulizer gas (52 a.u.). The cone and probe temperatures were 356°C and 420°C, respectively. Probe gas flow was 16 a.u. and spray voltage was 3500 V. Scan range was 150 to 1000 Da and time between scans was 100 ms.

Fisetin was detected when expressing both the *FMO* from *F. ananassa* and *P. hybrida.* The observed ion 287 in positive mode is congruent with the ions recognized in the fisetin standard (Figure 1). The spectra for the infusion of the fisetin standard gave an ion with m/z 287.0549, which corresponds to 0.35 ppm of ion with the ionic formula of $C_{15}H_{11}O_6$ in positive ion mode; this is indicative of a compound with a molecular formula of $C_{15}H_{10}O_6$.

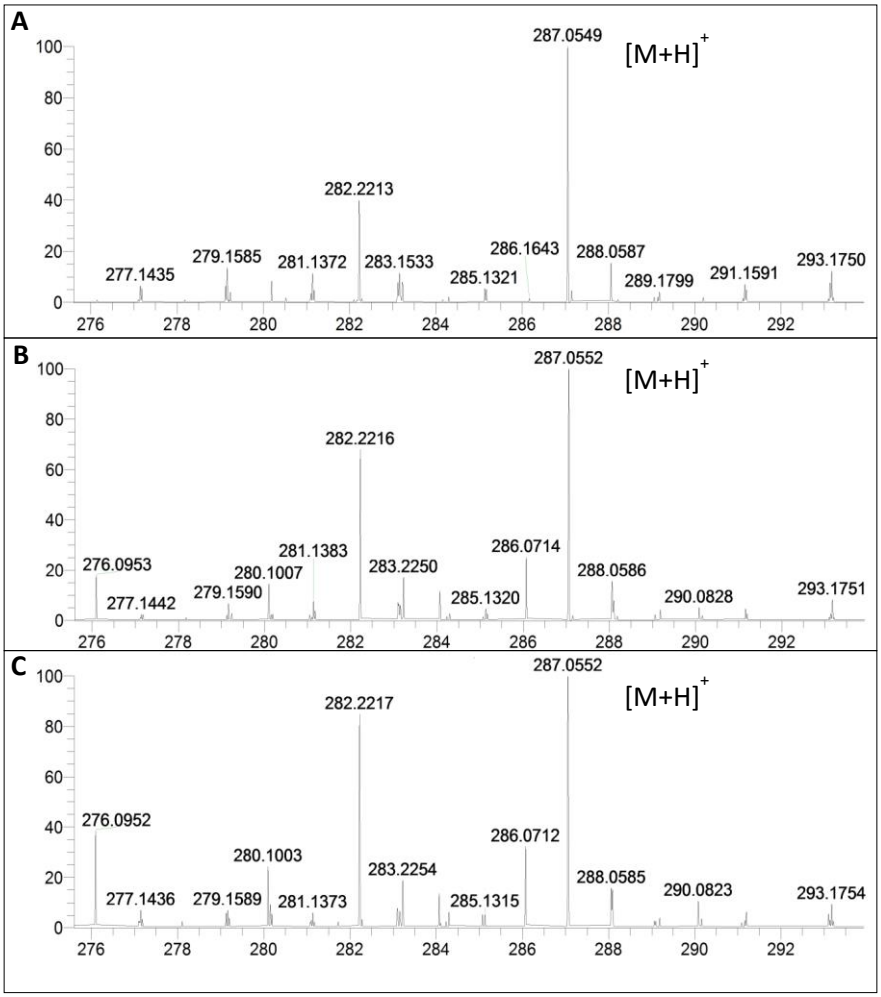


Figure 1. Spectra of infused isolated peaks of fisetin, positive spectrum. **A.** Fisetin standard, **B.** Fisetin synthesized using a *FMO* from *P. hybrida*, **C.** Fisetin synthesized using a *FMO* from *F. ananassa*.

Results

Flavonoids pathway design

Previously, we constructed a novel biosynthetic pathway for the production of fisetin starting from L-tyrosine in *E. coli* (Fig. 2) (Stahlhut et al. 2015). In this work, we set out to develop a fermentation process for *de novo* production of a portfolio of related flavonoids from a cheap carbon source and using yeast *S. cerevisiae* as the host. The biosynthetic pathway towards flavonoids is depicted in Fig. 2. First, *p*-coumaric acid must be activated by a 4-coumaroyl-CoA ligase *4CL*, and for this we used a variant from *P. crispum*. Next, we used chalcone synthase *CHS* from *P. hybrida* and chalcone isomerase *CHI* from *M. sativa* to convert the resulting *p*-coumaroyl-CoA and three molecules of malonyl-CoA into naringenin. To obtain isoliquiritigenin, we additionally overexpressed chalcone reductase *CHR* from *A. mongholicus*, the resulting naringenin and isoliquiritigenin producing strains were further engineered towards production of respectively kaempferol and resokaempferol by overexpression of flavanone 3-hydroxylase *F3H* from *A. mongholicus* and flavonol synthase *FLS* from *A. thaliana.* Lastly, to enable production of quercetin and fisetin, we tested two different cytochrome P450 flavonoid monooxygenases *FMO*s, one from *F. ananassa* and another from *P. hybrida*, which were overexpressed in the kaempferol producing strain to obtain quercetin and in the resokaempferol producing strain to obtain fisetin, respectively. Each of the tested *FMO* was fused in-frame with cytochrome P450 reductase *CPR* from *C. roseus.* Moreover, fisetin was synthesized either from L-tyrosine or L-phenylalanine by the use of tyrosine ammonia-lyase (TAL) or phenylalanine ammonia-lyase (PAL).

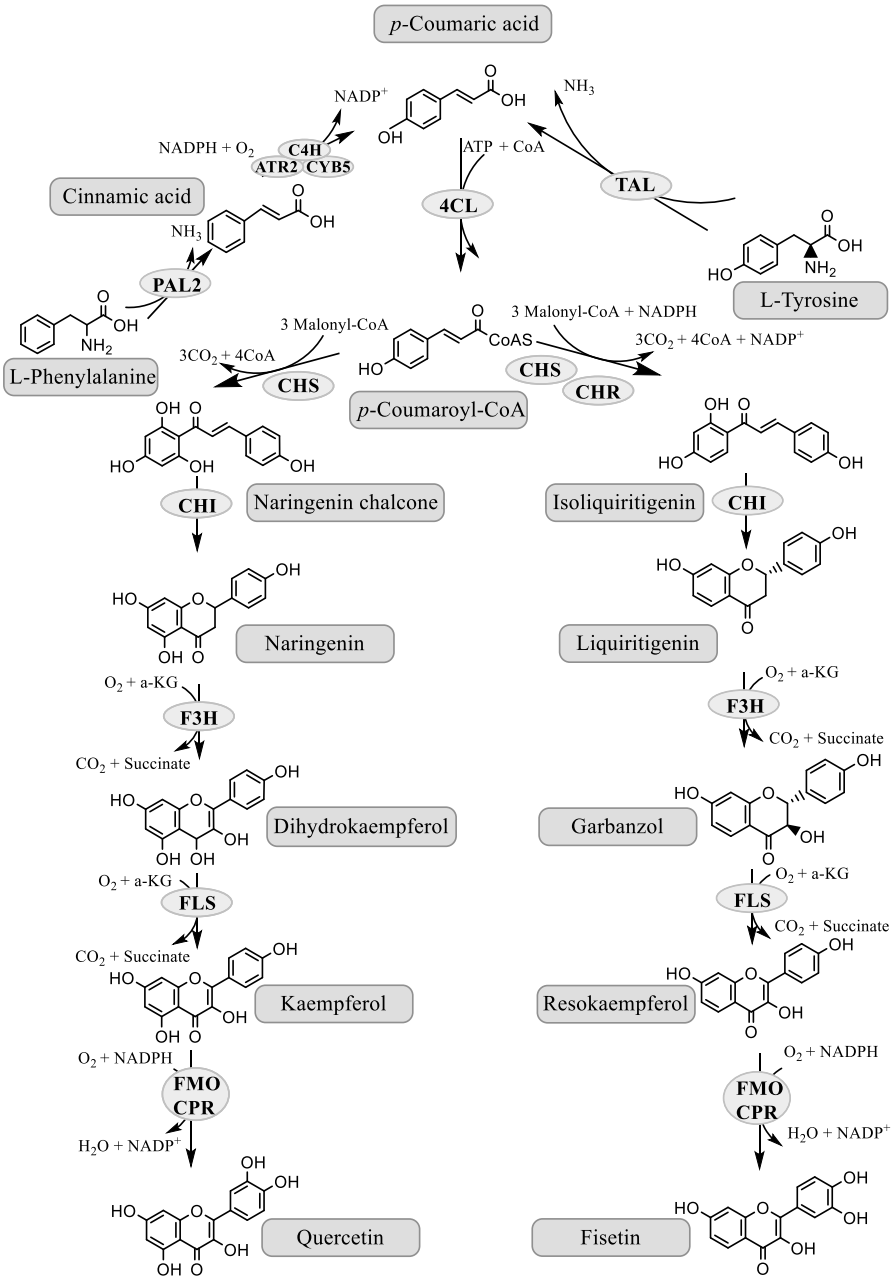

p-Coumaric acid
NADP+
NH3
NADPH + O2
C4H
ATR2 CYB5
ATP + CoA
TAL
Cinnamic acid
4CL
NH3
L-Tyrosine
PAL2
L-Phenylalanine
3 Malonyl-CoA
3 Malonyl-CoA + NADPH
3CO2 + 4CoA
3CO2 + 4CoA + NADP+
CHS
CHS
CHR
p-Coumaroyl-CoA
CHI
Naringenin chalcone
Isoliquiritigenin
CHI
Naringenin
Liquiritigenin
O2 + a-KG
F3H
CO2 + Succinate
Dihydrokaempferol
Garbanzol
O2 + a-KG
FLS
CO2 + Succinate
Kaempferol
Resokaempferol
O2 + NADPH
FMO
CPR
H2O + NADP+
Quercetin
Fisetin

Figure 2. Flavonoid biosynthetic pathways engineered into yeast *S. cerevisiae*. *PAL2*: phenylalanine ammonia-lyase; *C4H:* cinnamate 4-hydroxylyase; *ATR2* cytochrome p450 reductase; *CYB5*: cytochrome b5 electron carrier; *TAL*: tyrosine ammonia-lyase; *4CL*: 4-coumaroyl-CoA ligase; *CHS*: chalcone synthase; *CHR*: chalcone reductase; *CHI*: chalcone isomerase; *F3H*: flavanone 3-hydroxylase; *FLS*: flavonol synthase; *FMO*: flavonoid 3'-monooxygenase; *CPR*: cytochrome P450 reductase.

Biosynthesis of naringenin and liquiritigenin

For production of naringenin and liquiritigenin we explored a low-producer (ST4069) and a high-producer (ST2645) strain of *p*-coumaric acid as the starting platforms. As the low-producer we employed strain ST4069, which produced 0.24 ± 0.03 g L^{-1} of *p*-coumaric acid in our previous study; as the high-producer we used strain ST2645, which reached a titer of 1.93 ± 0 .26 g L^{-1} (Rodriguez *et al.,* 2015). Strain ST4069 has overexpression of a *TAL* gene, while strain ST2645 has additional modifications increasing the flux towards aromatic amino acids: knock-outs of *ARO10* and *PDC5* genes and over-expressions of *ARO4*^{fbr} and *ARO7*^{fbr}. To obtain naringenin production, we overexpressed 4-coumaroyl-CoA ligase *4CL*, chalcone synthase CHS and chalcone isomerase CHI. For liquiritigenin production we additionally overexpressed chalcone reductase CHR. The titers of naringenin and liquiritigenin were 3-fold higher in the strain background with optimized *p*-coumaric acid production: 1.55 ± 0 .13 mg L^{-1} of naringenin and 5.31 ± 0.48 mg L^{-1} of liquiritigenin. We also observed accumulation of *p*-coumaric acid in the medium, particularly in the optimized strains up to 8 mg L^{-1} of *p*-coumaric acid were measured. We also noted accumulation of naringenin (~1 mg L^{-1}) alongside liquiritigenin, indicating incomplete reduction by CHR.

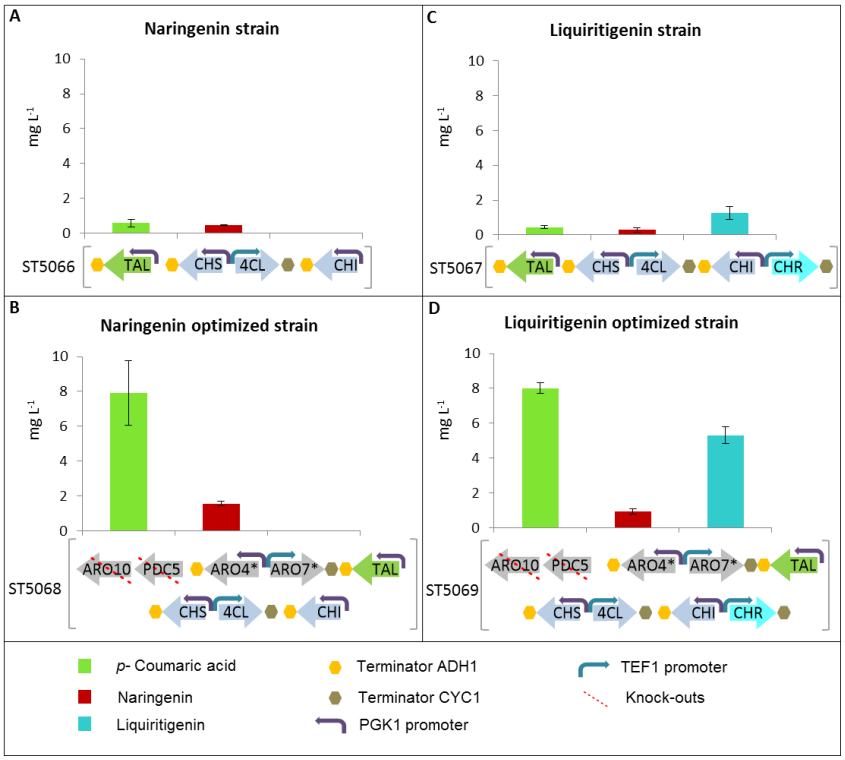


Figure 3. Strains engineered for biosynthesis of naringenin and liquiritigenin. A. Naringenin pathway was integrated into a low-level *p*-CA strain ST4069; B. Naringenin pathway was integrated into a high-level *p*-CA strain ST2645; C. Liquiritigenin pathway was integrated into a low-level *p*-CA strain ST4069; D. Liquiritigenin pathway was integrated into a high-level *p*-CA strain ST2645. The strains were cultivated in FIT medium for 72 hours, the mean value of extracellular concentration of compounds was calculated from five biological replicates.

Biosynthesis of kaempferol and resokaempferol

The naringenin- (strain ST5068) and liquiritigenin- (strain ST5069) producing strains were further engineered for production of respectively kaempferol and resokaempferol by overexpressing the genes encoding flavanone 3-hydroxylase F3H and flavonol synthase FLS. As a result, we obtained two new strains, the strain ST5070 for the production of kaempferol and the strain ST5071 for the production of resokaempferol. Strain ST5070 produced

26.57 ± 2.66 mg L^{-1} of kaempferol as well as two pathway intermediates: 10.75 ± 0.87 mg L^{-1} of *p*-coumaric acid and 0.83 ± 0.05 of naringenin (Figure 4A). The high titer of kaempferol was surprising, considering that the parent strain produced less than 2 mg L^{-1} of kaempferol precursor, naringenin. The depletion of naringenin and higher capacity to produce kaempferol imply that the conversion steps of naringenin into kaempferol are very efficient. The resokaempferol producing strain (ST5071) accumulated 0.51 ± 0.03 mg L^{-1} of resokaempferol, 14.54 ± 1.96 mg L^{-1} of *p*-coumaric acid, 11.39 ± 0.33 mg L^{-1} of kaempferol, and less than 1 mg L^{-1} of naringenin and liquiritigenin (Figure 4B). In the resokaempferol strain, we observed by-product kaempferol, which results from incomplete reduction of *p*-coumaroyl-CoA by CHR.

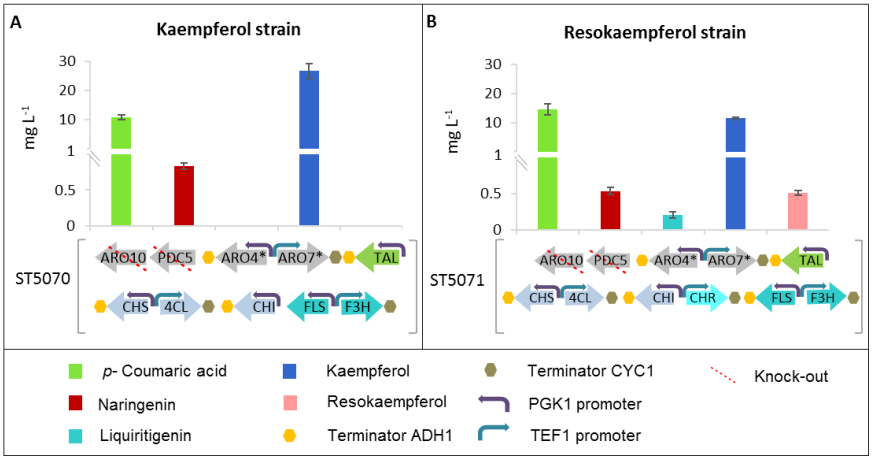


Figure 4. Strains engineered for biosynthesis of kaempferol and resokaempferol. A. Kaempferol producing strain. B. Resokaempferol producing strain. The strains were cultivated in FIT medium for 72 hours, the mean value of extracellular concentration of compounds was calculated from five biological replicates.

Biosynthesis of fisetin and quercetin

For the production of quercetin and fisetin, it is necessary to overexpress cytochrome P450 flavonoid monooxygenase (*FMO*) and cytochrome P450 reductase (*CPR*). In this study, we evaluated two *FMO* variants, one from *F. ananassa* and one from *P. hybrida,* that were fused in-frame to the *CPR* from *C. roseus* by a flexible glycine-serine linker (5'-GGGTCGAC-3').

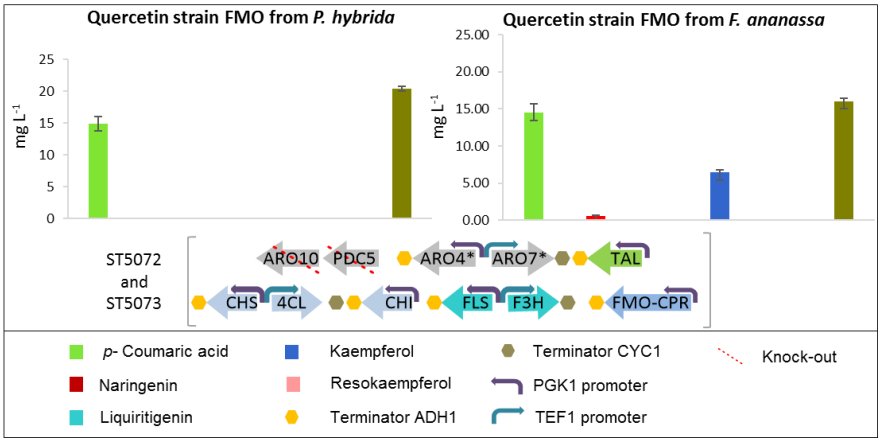


Figure 5. Strains engineered for biosynthesis of quercetin. The strains were cultivated in FIT medium for 72 hours, the mean value of extracellular concentration of compounds was calculated from five biological replicates.

The genes were overexpressed in the kaempferol producing strain (ST5068) to obtain 20.38 ± 2.57 mg L^{-1} and 16.04 ± 0.37 of quercetin in the strains with FMO from *P. hybrida* and *F. ananassa,* respectively (Figure 5). For fisetin biosynthesis, the host strain was the strain engineered for resokaempferol production (ST5069). Upon overexpression of *FMO-CPR* fusions, we obtained 1.65 ± 0.10 and 1.20 ± 0.19 mg L^{-1} of fisetin, respectively (Figure 6A). The strains overexpressing the *FMO* of *F. ananassa* additionally accumulated kaempferol and naringenin, while the strains with *FMO* from *P.*

hybrida did not secrete these intermediates, we therefore continued working with this strain (ST5074).

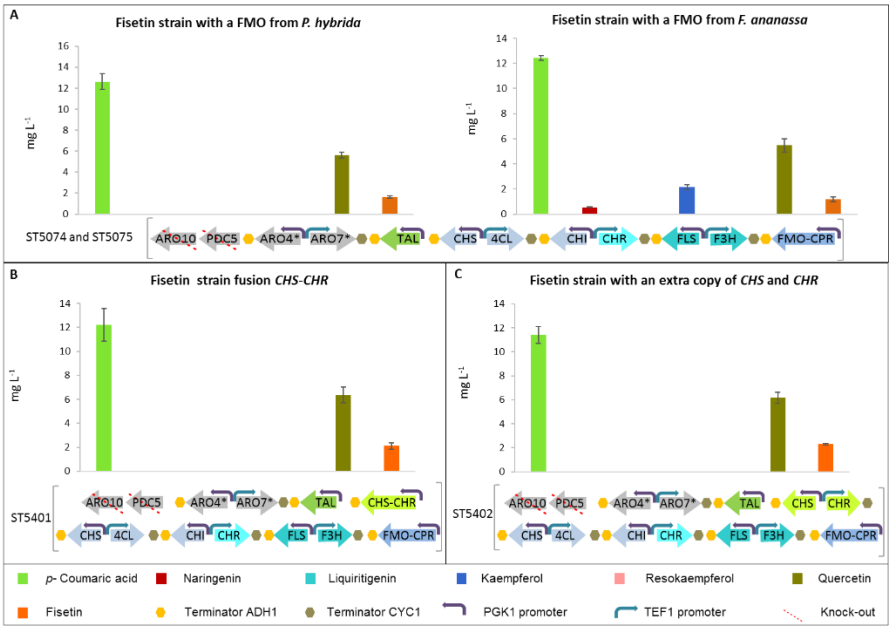


Figure 6. Strains engineered for biosynthesis of fisetin. A. Fisetin producer strains expressing FMOs from *P. hybrida* and *F. ananassa*. The background strain for these strains was the strain ST5071. B. Fisetin producer strain with an additional copy of fussed *CHS* and *CHR. C.* Fisetin producer strain with an additional copy of CHS and CHR. The background strain from the strains ST5401 and ST5402 was the strain ST5074. The strains were cultivated in FIT medium for 72 hours, the mean value of extracellular concentration of compounds was calculated from five biological replicates.

Improving fisetin production via heterologous pathway optimization

The strains designed to produce fisetin, accumulated significant amounts of quercetin. We attempted to direct more flux into fisetin by exploiting the metabolic channeling effect, which had proven effective in other studies (Albertsen *et al.,* 2011; Stahlhut *et al.,* 2015). We overexpressed a fusion of CHS and CHR proteins in the fisetin-producing strain ST5074, resulting in ST5401. A control strain, ST5402, expressing an additional copy of non-

fused *CHS* and *CHR* was constructed as well. We hereby obtained a significant improvement in production of fisetin, and the strains ST5401 and ST5402 produced 2.11 ± 0.26 mg L^{-1} and 2.29 ± 0.07 mg L^{-1} respectively in comparison to 1.65 ± 0.01 mg L^{-1} for the control strain ST5074 (Figure 6B and 6C). Interestingly, the final product (quercetin) of the competing metabolic pathway accumulated to similar concentrations. This result strongly indicates that protein fusion did not improve carbon flux toward fisetin production using this particular setup. Lastly, the precursor (*p*-coumaric acid) was the most abundant metabolite (approx. 12 mg L^{-1}) detected in both the CHS-CHR fusion and control strains.

Biosynthesis of flavonoid via phenylalanine

Several studies have previously demonstrated that *p*-coumaric acid, the first intermediate of the *de novo* flavonoid pathway, can be synthesized in *S. cerevisiae* via two routes: through tyrosine ammonia-lyase TAL or phenylalanine ammonia-lyase PAL (Koopmann *et al*., 2012; Li *et al*., 2015). We, therefore, set out to test whether supplying the metabolic precursor via the PAL route would be beneficial for flavonoid production as compared to the TAL route. Analysis of the basic strain (ST4757) revealed significant accumulation of *p*-coumaric acid, setting out the stage for construction of a flavonoid-producing cell factory. The remaining parts of the *de novo* pathway were reengineered in a similar manner as in the TAL strains, however, for the last biosynthetic step only the *FMO* from *P. hybrida* was tested as it showed better performance as described above. Interestingly, HPLC analysis of the final PAL-based strain (ST4972) revealed a different metabolic profile (Figure 7) as compared to the corresponding TAL-based strain. More specifically, we observed a 5.7-fold higher accumulation of *p*-coumaric acid (70 mg L^{-1}), while fisetin was only found in trace amounts and therefore could not be quantified.

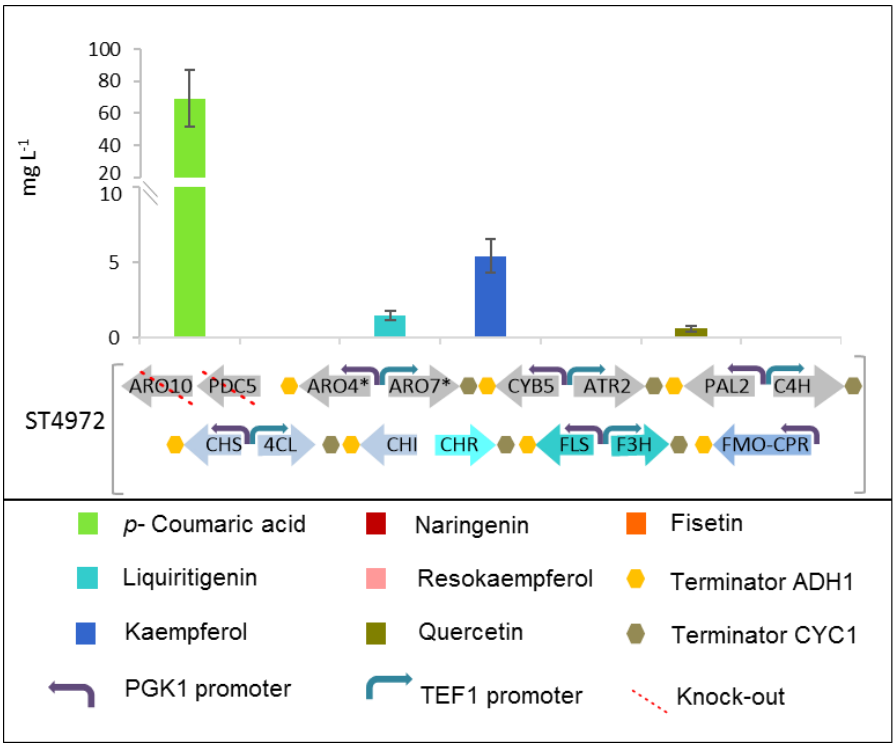


Figure 7. Strain engineered for biosynthesis of fisetin using the PAL pathway. The background strain for the strains ST4972 was the strain ST4757. The strains were cultivated in FIT medium for 72 hours, the mean value of extracellular concentration of compounds was calculated from five biological replicates.

Discussion

This work represents a proof-of-concept study for biosynthesis of flavonoids in engineered yeast cell factories. Here, we have successfully developed an array of yeast strains expressing heterologous flavonoid metabolic pathways containing up to ten genes, where some of the strains produced significant flavonoid titers in double-digit mg L^{-1} range. Moreover, we have successfully demonstrated that flavonoids can be synthesized via two different precursor-supplying routes. Metabolic engineering strategies applied in this study resulted in significantly improved titers for several metabolites of potential

commercial interest. The compounds quercetin and kaempferol were the ones with the highest extracellular concentrations detected in this study.

When assembling novel heterologous pathways for production of the complex metabolites it is important to engineer (or isolate a microbial host) with a sufficient pools of precursor metabolites for a given product; and to ensure the balanced enzymatic activities of the multistep metabolic pathways in order to avoid accumulation of intermediates and by-products. Therefore, we tested several engineered platform strains with different production capacities of the first intermediate *p*-coumaric acid. In all cases, when the strains with improved flux towards *p*-coumaric acid via TAL route were implemented an improvement in production of the downstream flavonoids was observed. However, this also resulted in high amounts of non-metabolized *p*-coumaric acid indicating existing metabolic regulated steps in the downstream flavonoid pathway. The existence of the limiting steps is clearly seen in the first examples, i.e. naringenin and liquiritigenin production. The low concentration of the compounds can be related to the low activity of CHS or CHI. Another explanation for low titers of naringenin and liquiritigenin could be a spontaneous chemical reaction into other not measured metabolites as well as instability of naringenin in an aqueous medium.

Although the precursors naringenin and liquiritigenin were secreted at amounts below 9 mg L^{-1}, it was possible to obtain the flavonoids kaempferol and resokaempferol using these strains as hosts. The overexpression of *F3H* and *FLS* did not decrease the levels of *p*-coumaric acid as its titers were comparable to the ones of the naringenin and liquiritigenin strains ~10 mg L^{-1}. The strain engineered for kaempferol production (strain ST5070) produced 26.57 ± 2.66 mg L^{-1} of kaempferol despite the fact that the host strain produced 1.55 ± 0.13 mg L^{-1} of the precursor naringenin. The strains engineered for liquiritigenin, resokaempferol and fisetin production (ST5069,

ST5071, ST5074 and ST5075) accumulated significant amounts of by-products; to optimize the conversion of *p*-coumaroyl-CoA into liquiritigenin, it is necessary to overcome the flux of the precursors to the production of naringenin chalcone and its derivate compounds.

For the final metabolic steps of the flavonoid pathway two different FMOs were tested for quercetin and fisetin production, where the *FMO* from *P. hybrida* displayed better enzymatic activities towards production of the final metabolites. The quercetin and fisetin strain expressing the *FMO* from *P. hybrida* did not accumulate the by-products naringenin and kaempferol, whereas the strains expressing the *FMO* from *F. ananassa* accumulated more than 2 mg L^{-1} of kaempferol and around 0.5 mg L^{-1} of naringenin.

Surprisingly, the production of the downstream flavonoids using the PAL pathway was much worse as compared to the similar strains using TAL route. In addition, the distribution of intermediates of the two competing pathways (quercetin and fisetin) was different too. Despite the fact that PAL strains accumulated 5.7-fold more coumaric acid as compared to the TAL strains, significantly lower amounts of quercetin and only trace levels of fisetin were observed. We speculate, that such a change in metabolic profile could be due to possible inhibitory effects of high levels of p-coumaric acid or negative interactions between the different P450s of the PAL and flavonoid pathways.

The accumulation of *p*-coumaric acid was common for all the engineered strains. The accumulation of this precursor indicates a limiting step for the conversion of *p*-coumaric acid into *p*-coumaroyl-CoA. Previous research reported the same limiting step in *E. coli* and yeast; they associated the accumulation of *p*-coumaric acid to the lack of malonyl-coA or low activity of the *4CL* for conversion of *p*-coumaric acid into *p*-coumaroyl-CoA (Stahlhut *et al.,* 2015, Koopman *et al.,* 2012). A first step to overcoming the limitation on

these enzymatic steps can be the combinatorial screening of other versions of the enzymes. Also, the over-expression on the limiting steps of the pathways can be explored (Yamada *et al.,* 2010; Zhou *et al.,* 2012; Zhang *et al.,* 2012). Indeed, an improvement by 28% was observed after an additional copy of CHS and CHR was introduced. Heterologous co-expression of genes originating from the same species has shown positive effects in the specificity and kinetic in the pathway for naringenin production (Koopman *et al.,* 2012). We think that this can be useful to test if the expression of genes from the same species has a positive effect on fisetin production in yeast. Another possibility for improvements could be testing several different *S. cerevisiae* strain backgrounds, as miniature genetic and phenotypic differences might result in significant differences of the final yields of a given metabolite (Strucko *et al*., 2015).

In higher plants, flavonoid biosynthetic enzymes are co-localized and assembled in protein complexes, the successive channeling of the intermediate compounds decreased the formation of by-products. In our study, however, the engineered fusion CHS-CHR protein did not result in the increased carbon flux towards fisetin biosynthesis branch. On the other hand, the increase of both quercetin and fisetin was observed proving the fact that these two enzymes are limiting steps for synthesis of flavonoids.

In summary, to our knowledge this is the first example where flavoinoids, resokaempferol and fisetin have been synthetized in yeast. Additionally, we found low enzymatic activity in two key steps: first, the conversion of *p*-coumaric acid into the first scaffolds in the flavonoids pathway (*4CL*, *CHS*, and *CHI*) and second the reduction of *p*-coumaroyl CoA into isoliquiritigenin (*CHR*).

Acknowledgements

This work was financed by the Novo Nordisk Foundation. We thank Vratislav Stovicek for providing the integrative vectors for stable gene expression and Mingji Li for providing the parts for the construction of the PAL pathway. Tomas Strucko, Steen Stahlhut and Jochen Förster have been funded by BacHBerry co-funded by the European Commission in the 7th Framework Programme, Project No. FP7-613793) and the ERA-IB project DeYeastLibrary (https://www.patil.embl.de/deyeastlibrary).

Supplementary material

Supplementary Table 1. Primers used in this study.

USER cloning		
ID	**Name**	**Sequence (5'-3')**
7628	(Pc_4Cl_U2_fw)	ATCTGTCAUAAAACAATGGGAGACTGTGTAGCAC
7629	(Pc_4Cl_U2_rv),	CACGCGAUTCATTATTTGGGAAGATCACCGGATG
7632	(Ph_CHS_U1_fw)	AGTGCAGGUAAAACAATGACCATGGTTACCGTTGAAG
7633	(Ph_CHS_U1_rv)	CGTGCGAUTCATTAGGTTGCAACGCTATGCAG
7636	(Ms_CHI_U1_fw)	AGTGCAGGUAAAACAATGACCATGGCAGCAAGC
7637	(Ms_CHI_U1_rv)	CGTGCGAUTCATTAGTTGCCGATTTTAAAGGCACC
7640	(Ms_CHR_U2_fw)	ATCTGTCAUAAAACAATGACCATGGGTAGCGTTG
7641	(Ms_CHR_U2_rv)	CACGCGAUTCATTAGTCATCATACAGATCATTCAGACC
7644	(At_F3H_U2_fw)	ATCTGTCAUAAAACAATGGCTCCAGGAACTTTGAC
7645	(At_F3H_U2_rv)	CACGCGAUTCACTAAGCGAAGATTTGGTCGACAG
7648	(At_FLS_U1_fw)	AGTGCAGGUAAAACAATGGAGGTCGAAAGAGTCC
7649	(At_FLS_U1_rv)	CGTGCGAUTCATCAATCCAGAGGAAGTTTATTGAGC
7654	(Fa_FMO-Cr_CPR_U1_fw)	AGTGCAGGUAAAACAATGGCGATTACCCTGCTG
7655	(Fa_FMO-Cr_CPR_U1_rv)	CGTGCGAUTCATTACCAAACGTCACGCAGATAAC
7658	(Ph_FMO-Cr_CPR_U1_fw)	AGTGCAGGUAAAACAATGGCGATTCTGTATACCGTG
7659	(Ph_FMO-Cr_CPR_U1_rv)	CGTGCGAUTCATTACCAAACGTCACGCAGATAAC
Verification		
ID	**Name**	**Sequence (5'-3')**
892	Sc_XII-1-down-out-sq	GGACGACAACTACGGAGGAT
894	Sc_XII-2-down-out-sq	GGCCCTGATAAGGTTGTTG
900	Sc_XII-5-down-out-sq	GTGGGAGTAAGGGATCCTGT
2220	Sc_Colo_pcr_fw	CCTGCAGGACTAGTGCTGAG
8419	Sc_XI-5_down-out-sq	GCATGGTCACCGCTATCAGC

Supplementary Table 2. List of Biobricks generated by PCR amplification, description of the templates can be found in Stahlhut et al., (2015).

NAME	TEMPLATE_FOR_PCR	FW_PRIMER_FOR_PCR	RV_PRIMER_FOR_PCR
BB0653 (Pc_4Cl->)	pCDF-4cl-2Pc	ID7628 (Pc_4Cl_U2_fw)	ID7629 (Pc_4Cl_U2_rv)
BB0655(Ph_CHS<-)	pET-chsPh-chiMs	ID7632 (Ph_CHS_U1_fw)	ID7633 (Ph_CHS_U1_rv)
BB0656 (Ms_CHI<-)	pET-chsPh-chiMs	ID7636 (Ms_CHI_U1_fw)	ID7637 (Ms_CHI_U1_rv)
BB0658 (Am_CHR->)	pRSF-chrAm	ID7642 (Am_CHR_U2_fw)	ID7643 (Am_CHR_U2_rv)
BB0659 (At_F3H->)	pCDFf3hAt-fls-1At	ID7644 (At_F3H_U2_fw)	ID7645 (At_F3H_U2_rv)
BB0660 (At_FLS <-)	pCDFf3hAt-fls-1At	ID7648 (At_FLS_U1_fw)	ID7649 (At_FLS_U1_rv)
BB0661 (Fa_FMO-Cr_CPR<-)	pACYCf30hFxa2-cprCr	ID7654 (Fa_F3H-Cr_CPR_U1_fw)	ID7655 (Fa_F3H-Cr_CPR_U1_rv)
BB0680 (Ph_FMO-Cr_CPR<-)	pACYCf30hPh-cprCr	ID7658 (Ph_F3H-Cr_CPR_U1_fw)	ID7659 (Ph_F3H-Cr_CPR_U1_rv)

References

Agati, G., Biricolti, S., Guidi, L., Ferrini, F., Fini, A., Tattini, M., 2011. The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in L. vulgare leaves. J. Plant Physiol. 168, 204–212. doi: 10.1016/j.jplph.2010.07.016

Albertsen, L., Chen, Y., Bach, L.S., Rattleff, S., Maury, J., Brix, S., Nielsen, J., Mortensen, U.H., 2011. Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. Appl. Environ. Microbiol. 77, 1033–1040. doi:10.1128/AEM.01361-10

Bulzomi, P., Galluzzo, P., Bolli, A., Leone, S., Acconcia, F., Marino, M., 2012. The pro-apoptotic effect of quercetin in cancer cell lines requires ERβ-dependent signals. J. Cell. Physiol. 227, 1891–1898. doi:10.1002/jcp.22917

Caltagirone, S., Rossi, C., Poggi, A., Ranelletti, F.O., Natali, P.G., Brunetti, M., Aiello, F.B., Piantelli, M., 2000. Flavonoids apigenin and quercetin inhibit melanoma growth and metastatic potential. Int. J. Cancer 87, 595–600. doi:10.1002/1097-0215(20000815)87:4<595: AID-IJC21>3.0.CO;2-5

Chen, A.Y., Chen, Y.C., 2013. A review of the dietary flavonoid, kaempferol on human health and cancer chemoprevention. Food Chem. 138, 2099–2107. doi: 10.1016/j.foodchem.2012.11.139

Choi, E.M., 2012. Liquiritigenin isolated from Glycyrrhiza uralensis stimulates osteoblast function in osteoblastic MC3T3-E1 cells. Int. Immunopharmacol. 12, 139–143. doi: 10.1016/j.intimp.2011.11.003

Crozier, A., Lean, M.E.J., McDonald, M.S., Black, C., 1997. Quantitative analysis of the flavonoid content of commercial tomatoes, onions, lettuce, and celery. J. Agric. Food Chem. 45, 590–595. doi:10.1021/jf960339y

Currais, A., Prior, M., Dargusch, R., Armando, A., Ehren, J., Schubert, D., Quehenberger, O., Maher, P., 2014. Modulation of p25 and inflammatory pathways by fisetin maintains cognitive function in Alzheimer's disease transgenic mice. Aging Cell 13, 379–390. doi:10.1111/acel.12185

Cushnie, T.P.T., Lamb, A.J., 2011. Recent advances in understanding the antibacterial properties of flavonoids. Int. J. Antimicrob. Agents 38, 99–107. doi: 10.1016/j.ijantimicag.2011.02.014

Grotewold E., 2006. The Science of flavonoids, Springer Science Business Media, Inc, New York.

Hämäläinen, M., Nieminen, R., Vuorela, P., Heinonen, M., Moilanen, E., 2007. Anti-inflammatory effects of flavonoids: Genistein, kaempferol, quercetin, and daidzein inhibit STAT-1 and NF-κB activations, whereas flavone, isorhamnetin, naringenin, and pelargonidin inhibit only NF-κB activation along with their inhibitory effect on i. Mediators Inflamm. 2007. doi:10.1155/2007/45673

Hertog, M.G., Hollman, P.C., Katan, M.B., 1992. Content of potentially anticarcinogenic flavonoids of 28 vegetables and 9 fruits commonly consumed in The Netherlands. J. Agric. Food Chem. 40, 2379–2383. doi:10.1021/jf00024a011

Khan, N., Afaq, F., Syed, D.N., Mukhtar, H., 2008. Fisetin, a novel dietary flavonoid, causes apoptosis and cell cycle arrest in human prostate cancer LNCaP cells. Carcinogenesis 29, 1049–1056. doi:10.1093/carcin/bgn078

Khoo, B.Y., Chua, S.L., Balaram, P., 2010. Apoptotic effects of chrysin in human cancer cell lines. Int. J. Mol. Sci. 11, 2188–2199. doi:10.3390/ijms11052188

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J. a, Pronk, J.T., Daran, J.-M., 2012. *De novo* production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microb. Cell Fact. 11, 155. doi:10.1186/1475-2859-11-155

Leonard, E., Ajikumar, P.K., Thayer, K., Xiao, W.-H., Mo, J.D., Tidor, B., Stephanopoulos, G., Prather, K.L.J., 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. Proc. Natl. Acad. Sci. U. S. A. 107, 13654–13659. doi:10.1073/pnas.1006138107

Leonard, E., Yan, Y., Koffas, M.A.G., 2006. Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. Metab. Eng. 8, 172–181. doi:10.1016/j.ymben.2005.11.001

Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I., Nielsen, J., 2015. *De novo* production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. Metabolic Engineering. Metab. Eng. 32, 1–11. doi: 10.1016/j.ymben.2015.08.007

Li, M., Schneider, K., Kristensen, M., Borodina, I., Nielsen, J., 2016. Engineering yeast for high-level production of stilbenoid antioxidants.

Liu, Y., Xie, S., Wang, Y., Luo, K., Wang, Y., Cai, Y., 2012. Liquiritigenin inhibits tumor growth and vascularization in a mouse model of HeLa cells. Molecules 17, 7206–7216. doi:10.3390/molecules17067206

Maher, P., Dargusch, R., Bodai, L., Gerard, P.E., Purcell, J.M., Lawrence Marsh, J., 2011. Erk activation by the polyphenols fisetin and resveratrol provides neuroprotection in multiple models of Huntington's disease. Hum. Mol. Genet. 20, 261–270. doi:10.1093/hmg/ddq460

Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., & Halkier, B. A. (2012). Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metabolic Engineering, 14(2), 104–11. http://doi.org/10.1016/j.ymben.2012.01.006

Mukhtar, E., Adhami, V.M., Sechi, M., Mukhtar, H., 2015a. Dietary flavonoid fisetin binds to β-tubulin and disrupts microtubule dynamics in prostate cancer cells. Cancer Lett. 367, 173–183. doi: 10.1016/j.canlet.2015.07.030

Nabavi, S.M., Nabavi, S.F., Eslami, S., Moghaddam, A.H., 2012. In vivo protective effects of quercetin against sodium fluoride-induced oxidative stress in the hepatic tissue. Food Chem. 132, 931–935. doi: 10.1016/j.foodchem.2011.11.070

Paganga, G., Miller, N., Rice-Evans, C. a, 1999. The polyphenolic content of fruit and vegetables and their antioxidant activities. What does a serving constitute? Free Radic. Res. 30, 153–162. doi:10.1080/10715769900300161

Raza, S.S., Khan, M.M., Ahmad, A., Ashafaq, M., Islam, F., Wagner, A.P., Safhi, M.M., Islam, F., 2013. Neuroprotective effect of naringenin is mediated through suppression of NF-kB signaling pathway in experimental stroke. Neuroscience 230, 157–171. doi: 10.1016/j.neuroscience.2012.10.041

Roberts, M.R., Paul, N.D., 2006. Seduced by the dark side: Integrating molecular and ecological perspectives on plant defence against pests and pathogens. New Phytol. 170, 677–699. doi:10.1111/j.1469-8137.2006. 01707.x

Rodriguez, A., Kildegaard, K.R., Li, M., Borodina, I., Nielsen, J., 2015. Establishment of a yeast platform strain for production of p-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. Metab. Eng. 31, 181-188. doi: 10.1016/j.ymben.2015.08.003

Routray, W., Orsat, V., 2012. Microwave-Assisted Extraction of Flavonoids: A Review. Food Bioprocess Technol. 5, 409–424. doi:10.1007/s11947-011-0573-z

Santos, C.N.S., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab. Eng. 13, 392–400. doi: 10.1016/j.ymben.2011.02.002

Stahlhut, S.G., Siedler, S., Malla, S., Harrison, S.J., Maury, J., Neves, A.R., Forster, J., 2015. Assembly of a novel biosynthetic pathway for production of the plant flavonoid fisetin in *Escherichia coli*. Metab. Eng. 31, 84–93. doi:10.1016/j.ymben.2015.07.002

Stovicek, V., Borja, G.M., Forster, J., Borodina, I., 2015. EasyClone 2.0: expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. J. Ind. Microbiol. Biotechnol. 42, 1519–1531. doi:10.1007/s10295-015-1684-8

Strucko, T., Magdenoska, O., & Mortensen, U. H. U. H. (2015). Benchmarking two commonly used *Saccharomyces cerevisiae* strains for heterologous vanillin-β-glucoside production. Metabolic Engineering Communications, 2, 99–108. http://doi.org/10.1016/j.meteno.2015.09.001

Trantas, E., Panopoulos, N., Ververidis, F., 2009. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. Metab. Eng. 11, 355–66. doi: 10.1016/j.ymben.2009.07.004

Wedick, N.M., Pan, A., Cassidy, A., Rimm, E.B., Sampson, L., Rosner, B., Willett, W., Hu, F.B., Sun, Q., Van Dam, R.M., 2012. Dietary flavonoid

intakes and risk of type 2 diabetes in US men and women. Am. J. Clin. Nutr. 95, 925–933. doi:10.3945/ajcn.111.028894

Wang, Y., Chen, S., Yu, O., 2011. Metabolic engineering of flavonoids in plants and microorganisms. Appl. Microbiol. Biotechnol. 91, 949–56. doi:10.1007/s00253-011-3449-2

Yamada, R., Tanaka, T., Ogino, C., Fukuda, H., Kondo, A., 2010. Novel strategy for yeast construction using delta-integration and cell fusion to efficiently produce ethanol from raw starch. Appl. Microbiol. Biotechnol. 85, 1491–1498. doi:10.1007/s00253-009-2198-y

Zhang, B., Rong, C., Chen, H., Song, Y., Zhang, H., Chen, W., 2012. *De novo* synthesis of trans-10, cis-12 conjugated linoleic acid in oleaginous yeast Yarrowia Lipolytica. Microb. Cell Fact. 11, 51. doi:10.1186/1475-2859-11-51

Zhou, H., Cheng, J. sheng, Wang, B.L., Fink, G.R., Stephanopoulos, G., 2012. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. Metab. Eng. 14, 611–622. doi: 10.1016/j.ymben.2012.07.011

5. Conclusions and perspectives

There is a permanent interest in the development of a yeast cell factory for the synthesis of aromatic secondary metabolites since many of these metabolites have a broad range of pharmaceutical and nutraceutical properties. This thesis contributed to a better understanding of the aromatic amino acids pathway in yeast and the physiological challenges that the cell faces after being engineered for the synthesis of *p*-coumaric acid (*p*-CA), a building block for the synthesis of many secondary metabolites. The *p*-CA platform strain was successfully employed for the synthesis of aromatic secondary metabolites using as a test case the synthesis of flavonoids.

Metabolic engineering for the synthesis of aromatic secondary metabolites in yeast.

Through the combination of several different strategies we improved the production of *p*-CA in yeast: reduction of by-product formation, followed by the elimination of the allosteric regulation of enzymes of the shikimate pathway and finally identification of flux-limiting steps in the shikimate pathway.

E. coli and *S. cerevisiae* had been engineered in previous research for the production of different aromatic secondary metabolites, however still is necessary further work for the development of a platform strain capable of high-level production of aromatic metabolites. Our platform strain reached a titer of 1.93 ± 0.26 g L^{-1} of *p*-CA; besides representing a starting point for further development of a process for commercial *p*-CA production, we believe this strain can be used as a platform strain for production of flavonoids and other *p*-CA derived secondary metabolites.

Systems biology and development of cell factories

Through the use of transcriptomics, metabolomics and physiological analysis we investigated how the production of *p*-CA influences the host and whether

these effects depend on the strain background. We found that the transport of amino acids and sugars were highly downregulated in the engineered strains of both backgrounds(CEN.PK and S288c); probably, as a response to the stress produced by the secreted *p*-CA.

Since the CEN.PK strain was less affected than S288c strain by increased *p*-CA production as was evident from fewer changes in the transcription profile, intracellular metabolites concentrations, and high *p*-CA titers; we consider that CEN.PK strain suits better as a cell factory for the production of *p*-CA.

Further improvement of the *p*-CA cell factory can be reached by the use of a combinatorial evaluation of the upregulation and downregulation of the genes that were significantly affected in their expression levels, (Aminotransferases, amino acids permeases and transporters). This approach will allow identifying if the downregulations and upregulations of these genes are a response to stress with a negative effect on the production of *p*-CA, or physiological adjustment to the flux of intermediate compounds toward L-tyrosine with a positive effect on the platform strain.

Cell factories for the synthesis of flavonoids

We successfully expressed heterologous flavonoid metabolic pathways in the previously engineered *p*-CA platform strain: we developed a set of strains for the synthesis of the flavonoids naringenin, liquiritigenin, kaempferol, resokaempferol, quercetin and fisetin and for the first time liquiritigenin, resokaempferol and fisetin were synthesized in yeast.

The compounds quercetin and kaempferol were the ones with the highest extracellular concentrations detected in this study; these strains secreted 20.38 ± 2.57 mg L^{-1} and 26.57 ± 2.66 mg L^{-1}respectivelly. We demonstrated that flavonoids can be synthesized via two different precursor supplying routes: PAL and TAL pathways, interestingly we found that the production of flavonoids was better in the strains engineered with the TAL pathway. The differences between the PAL and TAL pathway could be due to possible

inhibitory effects of high levels of *p*-CA or negative interactions between the different P450s of the PAL and flavonoid pathways.

By the use of strains with improved flux towards *p*-CA via TAL route, the synthesis of the downstream flavonoids was improved; however, the engineered strains accumulated significant amounts of *p*-CA, which suggested that the *p*-coumaroyl-CoA ligase *4CL* and possibly chalcone synthase *CHS* were limiting the carbon flux into flavonoids. Overexpression of an additional copy of *CHS* and *CHR* genes indeed increased the titer of fisetin from 1.65 ± 0.101 mg L^{-1} to 2.28 ± 0.073 mg L^{-1}.

Since we observed accumulation of intermediate compounds in the strains engineered for flavonoids synthesis, indicating that some enzymes are limiting the flux of intermediate compounds. Further improvement of the strains can be reached by the expression of extra copies of the enzymes involved in the limiting steps. At this respect approaches such as multivariate metabolic engineering can be useful since it is possible to test different promoters, heterologous genes and copy numbers aiming to reach optimal expression levels of the enzymes involved in the synthesis of flavonoids.

Traditional cell factories such as *E. coli* and *S. Cerevisiae* have been used for the development of a broad range of platform strains for production of a diverse range of biological and pharmaceutical products. In the particular case of aromatic secondary metabolites, it can be useful to explore alternative cell factories: aromatic secondary metabolites are produced in higher plants and they share an evolutionary trait with microalgae. Since it has been proved that that microalgae are also natural producers of aromatic secondary metabolites; it will be feasible to adapt this organism for the production of aromatic compounds.

The Novo Nordisk Foundataion
Center for Biosustainability
Technical University of Denmark

Kemitorvet Bygning 220
2800 Kgs. Lyngby
Denmark
Tlf. +45 45 25 80 00

www.biosustain.dtk.dk