TU Delft

Delft University of Technology

Metabolic engineering of Saccharomyces cerevisiae for the production of aromatic compounds

Hassing, E.J.

DOI
10.4233/uuid:09108773-40cb-4db1-adf9-a6f586e03eca

Publication date
2022

Document Version
Final published version

Citation (APA)
Hassing, E. J. (2022). *Metabolic engineering of Saccharomyces cerevisiae for the production of aromatic compounds*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:09108773-40cb-4db1-adf9-a6f586e03eca

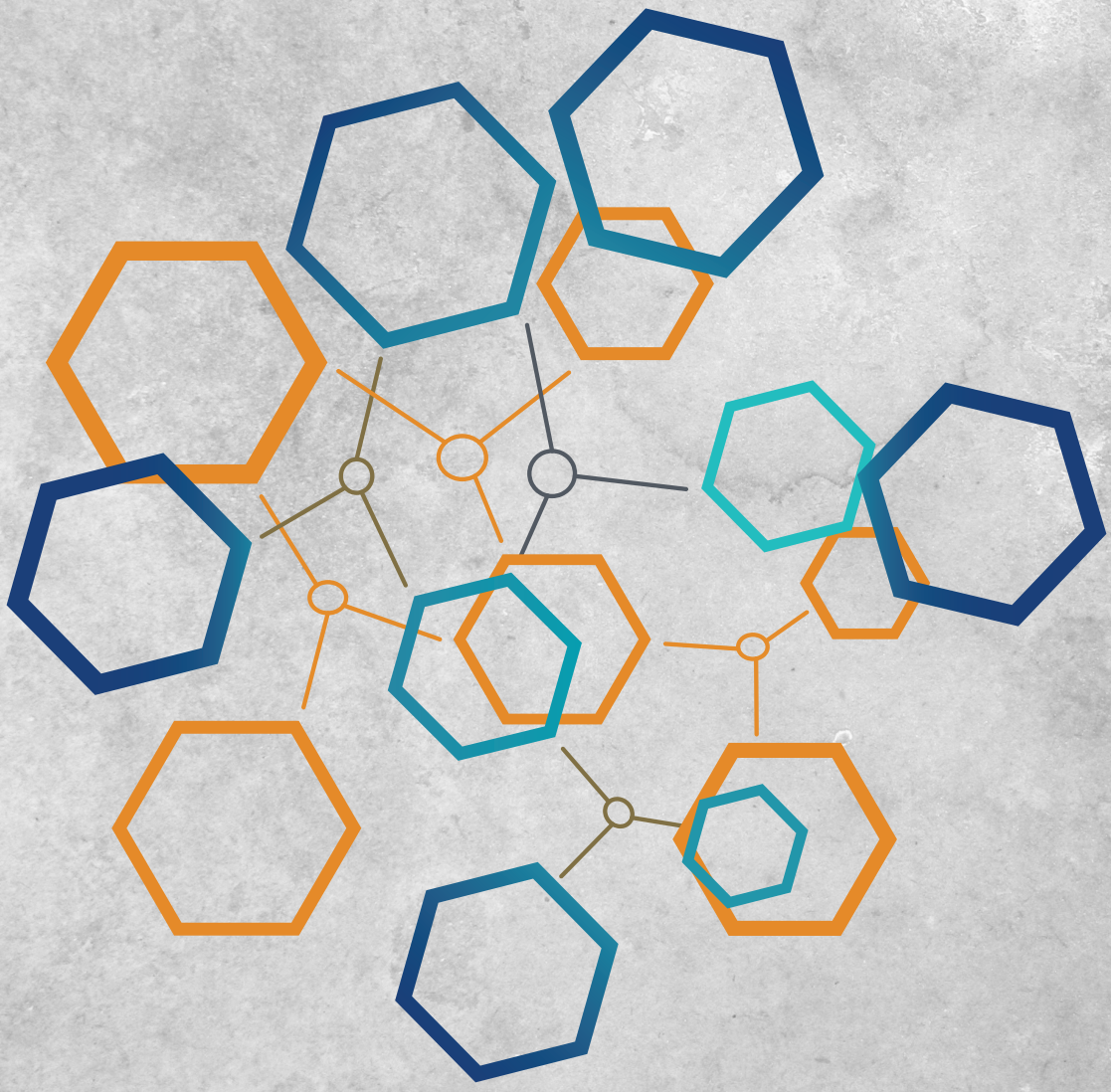
Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Metabolic engineering of *Saccharomyces cerevisiae* for the production of aromatic compounds



Else-Jasmijn Hassing

Metabolic engineering of *Saccharomyces cerevisiae* for the production of aromatic compounds

Dissertation

For the purpose of obtaining the degree of doctor

at Delft University of Technology

by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen

Chair of the Board for Doctorates

to be defended publicly on

Friday 28 January 2022 at 12:30 o'clock

By

Else-Jasmijn HASSING

Master of Science in Life, Science and Technology,

Delft University of Technology, The Netherlands

Born in Venlo, The Netherlands

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. dr. ir. J.-M.G Daran	Delft University of Technology, promotor
Prof. dr J.T. Pronk	Delft University of Technology, promotor

Independent Members:

Prof. dr. F. Hollmann	Delft University of Technology
Prof. dr. A.J.M. Driessen	University of Groningen
Dr. J.P. Morissey	University College Cork
Dr. B. Gasser	University of Natural Resources and Life Sciences BOKU
Dr. R. González Cabaleiro	Delft University of Technology

Reserve member:

Prof. dr. P.A.S. Daran-Lapujade	Delft University of Technology

The research presented in this thesis was performed at the Industrial Microbiology section, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, the Netherlands. The project was part of the CHASSY project and received funding from the European Union's Horizon 2020 Research and Innovation program under grant agreement No. 720824.

Cover	Kim Schotting
Layout	Else-Jasmijn Hassing
Printed by	ProefschriftMaken | www.proefschriftmaken.nl
ISBN	978-94-6423-618-7

© 2022 Else-Jasmijn Hassing
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronically, mechanically, by photo-copying, recording or otherwise, without prior written permission of the author.

Contents

Samenvatting 5

Summary 9

Chapter 1 13
General Introduction.

Chapter 2 37
Connecting central carbon and aromatic amino acid metabolism to improve *de novo* 2-phenylethanol production in *Saccharomyces cerevisiae.*

Chapter 3 71
Modular, synthetic chromosomes as new tools for large scale engineering of metabolism.

Chapter 4 103
Elimination of aromatic fusel alcohols as by-products of *Saccharomyces cerevisiae* strains engineered for phenylpropanoids production by 2-oxo acid decarboxylase replacement.

Outlook 135

Bibliography 139

Acknowledgements 157

Curriculum Vitae 162

List of Publications 163

Samenvatting

De wereldpopulatie is de afgelopen eeuw verdubbeld. Deze groei gaat vergezeld met een toenemende vraag naar voedsel, water, energie en consumentengoederen. De resulterende intensivering van landbouw, ontbossing, uitstoot van broeikasgassen en het hoge gebruik van natuurlijke grondstoffen zoals steenkool, mineralen en fossiele brandstoffen hebben geleid tot een opwarming van de aarde en uitputting van natuurlijke grondstoffen. Het toepassen van biotechnologie kan bijdragen aan de transitie naar een meer duurzame, circulaire, "bio-based" economie, door ontwikkeling van nieuwe processen voor het maken van diverse producten, zoals medicijnen, dranken, voedsel, chemicaliën en brandstoffen uit hernieuwbare grondstoffen. Bakkersgist is bekend om zijn natuurlijke vermogen om suikers om te zetten in CO_2 en ethanol. Hiervan wordt al eeuwenlang gebruik gemaakt voor het maken van onder andere alcoholische dranken en brood. Pure ethanol biedt bovendien, als transportbrandstof, een duurzamer alternatief voor benzine. Microbiële biotechnologie kan, dankzij de enorme groei in kennis, technieken voor DNA-bewerking en DNA-sequentiebepalingen, nu worden toegepast voor het maken van steeds meer producten. Dit betreft zowel producten die van nature in een micro-organisme voorkomen, maar ook producten die van nature alleen in andere organismen, zoals bijvoorbeeld in mensen of planten, voorkomen. Micro-organismen zijn bijvoorbeeld genetisch aangepast voor productie van complexe moleculen zoals menselijk insuline, de smaakstof vanilline en het antimalaria-medicijn artemisinine. Insuline, dat essentieel is voor behandeling van diabetici, werd voor deze ontwikkelingen geproduceerd met behulp van extractie uit varkenspancreas, terwijl artemisinine en vanilline uit planten werden gewonnen of, in het geval van vanilline, chemisch werd gesynthesiseerd. Microbiële productie van zulke industrieel waardevolle moleculen, vanuit een simpel substraat zoals glucose, levert een meer betrouwbare en duurzame productiemethode dan de eerder genoemde klassieke methoden. In dit proefschrift wordt speciaal de nadruk gelegd op de bakkersgist *Saccharomyces cerevisiae* en zijn toepassing in de microbiële productie van aromatische producten.

Er is een toenemende interesse in microbiële productie van aromatische moleculen, bijvoorbeeld in de smaak-en geurindustrie. De economische potentie van dit veld komt mede voort uit Europese wetgeving, die productie en verkoop van microbieel geproduceerde producten toestaat zolang het eindproduct vrij is van genetisch gemodificeerde organismes (GMOs). *S. cerevisiae* is van nature in staat om verschillende soorten aromatische producten te maken, maar deze productie is beperkt door een strakke regulatie van de betrokken reactieroutes. Veel andere industrieel interessante aromatische producten komen van nature alleen voor in planten. Om productie van deze aromatische verbindingen in *S. cerevisiae*

mogelijk te maken, is het niet alleen nodig om de betreffende genen vanuit planten in deze gist in te brengen, maar ook om het stofwisselingsnetwerk van de gist aan te passen voor snelle en efficiënte productvorming.

Hoofdstuk 1 geeft een inleiding in dit onderzoeksveld door de impact te bestuderen van het genetisch modificeren van de metabole routes van *S. cerevisiae* voor de biosynthese van aromatische producten. Hiertoe wordt eerst de transitie van *Saccharomyces* soorten beschreven van traditionele toepassingen naar de moderne biotechnologie. Vervolgens wordt behandeld hoe *S. cerevisiae* kan worden gebruikt voor de biosynthese van een diverse groep van aromatische verbindingen met toepassingen in onder andere de levensmiddelenindustrie, de chemische en de farmaceutische industrie. Hoewel voor beide klassen van producten vergelijkbare genetische modificatietechnieken worden ingezet, wordt hierbij onderscheid gemaakt tussen producten die gist van nature maakt, en heterologe producten die niet in wildtype gisten voorkomen. Naast zijn natuurlijke eigenschappen om CO_2 en ethanol te produceren, is *S. cerevisiae* ook in staat om smaak-actieve moleculen te maken zoals hogere alcoholen ("foezelalcoholen"), die afkomstig zijn van aminozuurmetabolisme via de Ehrlich-route. Extra nadruk wordt gelegd op de hogere foezelalcoholen 2-fenylethanol en *p*-hydroxyfenylethanol. Deze afbraakproducten van, respectievelijk, de aromatische aminozuren fenylalanine en tyrosine zijn op zichzelf waardevolle aromatische moleculen die onder andere worden gebruikt in de voedsel-en smaakstoffenindustrie. De aromatische aminozuren fenylalanine en tyrosine kunnen ehter ook dienen als bouwstoffen voor de biosynthese van fenylpropanoïden, een groep plant-natuurlijke moleculen waar o.a. stilbenoïden, tannines en flavonoïden onder vallen. Huidige productiemethodes voor deze moleculen zijn gebaseerd op directe extractie uit planten, wat doorgaans een inefficiënt en niet duurzaam proces oplevert. Microbiële productie, waarbij gebruik wordt gemaakt van hernieuwbare grondstoffen, biedt een veelbelovend alternatief voor de productie van aromatische moleculen. Echter, om industrieel relevante productopbrengsten, productconcentraties (titers) en productiesnelheden te behalen, moeten verschillende uitdagingen worden overwonnen. Ten eerste moet, in de gistcel, de beschikbaarheid van metabolieten die nodig zijn voor biosynthese van fenylpropanoïden worden verhoogd. Daarnaast moet de katalytische capaciteit van snelheidsbeperkende enzymen die betrokken zijn in fenylpropanoïdeproductie worden verhoogd. Een ander belangrijk doel in procesontwikkeling is het voorkomen van bijproducten, bijvoorbeeld door het aanpassen van de substraatspecificiteit van betrokken enzymen. Deze drie uitdagingen staan centraal in het onderzoek dat in dit proefschrift wordt beschreven.

In **Hoofdstuk 2** worden meerdere genetische strategieën bestudeerd om de flux door de shikimaatroute, de stofwisselingsroute die verantwoordelijk is voor synthese van aromatische aminozuren, in *S. cerevisiae* te verhogen. Een verhoogde flux door deze route resulteert in verhoogde productie van 2-fenylethanol. Dit natuurlijke product werd daarom als model

gebruikt om de impact van individuele en gecombineerde genetische aanpassingen te evalueren. Een combinatie van het opheffen van allostere regulatie van essentiële enzymen in shikimaatroute, deregulatie van cruciale reacties van de pentosefosfaatroute en de glycolyse, gecombineerd met het verlagen van de capaciteit van een concurrerende route die leidt tot productie van *p*-hydroxyfenylethanol, de van tyrosine afkomstige foezelalcohol, leidde tot een substantiële toename in 2-fenylethanolproductie in *S. cerevisiae*. Dit onderzoek leverde daarmee een basis voor de biosynthese van een diverse groep aromatische producten waarvan de synthese aan de shikimaatroute kan worden gekoppeld, zoals bijvoorbeeld fenylpropanoïden.

Hoofdstuk 3 bestudeert hoe constructie van neo-chromosomen (NeoChrs) kan worden ingezet om *S. cerevisiae*-stammen te verkrijgen waarin nieuwe stofwisselingsroutes en relevante reacties uit de centrale giststofwisseling snel kunnen worden geoptimaliseerd. *De novo* productie van de in voedsel toegepaste kleurstof anthocyaan pelargonidin 3-O-glucoside (P3G) in *S. cerevisiae* wordt hierbij gebruikt als modelproces. Hoewel P3G al eerder is geproduceerd in genetisch gemodificeerde microorganismen, waaronder ook bakkergist, blijft synthese in microbiële platforms tot nu toe inefficiënt. Voor omzetting van glucose in P3G zijn 27 reacties uit de centrale stofwisseling van *S. cerevisiae* nodig, in de glycolyse, pentosefosfaatroute, shikimaatroute en aromatische aminozuurbiosynthese, en tien reacties die van nature in planten voorkomen. Modulaire NeoChrs met daarop gistgenen, bacteriële genen en plantengenen vereist voor *de novo* P3G synthese werden zowel met een lineaire als een circulaire configuratie ontworpen en gebouwd. Een stam met een lineair NeoChr, met negen kopieen van de chalconsynthase, een van de benodigde genen, produceerde extracellulair P3G. Dit resultaat betekende een verbetering ten opzichte van eerdere studies waarbij dit molecuul vrijwel uitsluitend intracellulair bleef.

De heterologe route voor fenylpropanoïden, een grote groep aromatische verbindingen waar onder andere 3-fenylpropeenzuren, stilbenoïden, en flavonoïden toe behoren, begint met deaminering van de precursor L-fenylalanine en/of L-tyrosine via de shikimaatroute. Echter, zoals beschreven in **Hoofdstuk 2**, kunnen deze aromatische aminozuren ook worden omgezet in de Ehrlich route. Dit resulteert in de productie van aromatische foezelalcoholen en zuren, die in giststammen die zijn aangepast voor de productie van waardevolle aromatische producten worden beschouwd als ongewenste bijproducten. De synthese van foezelalcoholen en zuren vereist een van de vier 2-ketocarbonzuur-decarboxylases van *S. cerevisiae*. Eliminatie van deze vier enzymen door het uitschakelen van de verantwoordelijke genen is echter niet eenvoudig omdat deze decarboxylases tijdens groei op glucose een essentiële rol spelen in cytosolische acetyl-CoA biosynthese. **Hoofdstuk 4** beschrijft een zoektocht naar een decarboxylase dat wel de hiervoor benodigde pyrodruivenzuurdecarboxylase heeft, maar dat geen aromatische 2-ketocarbonzuren kan decarboxyleren. Uit elf geteste decarboxylases uit gisten en bacteriën,

lieten enzymen uit de bacteriën *Zymomonas mobilis* en G*luconacetobacter diazotrophicus* de gezochte substraatspecificiteit zien. Deze enzymen waren bovendien in staat om groei op glucose van een *S. cerevisiae*-stam waaruit de natieve pyruvaatdecarboxylases waren verwijderd, te herstellen. Deze decarboxylase-uitwisselingsstrategie leidde tot een volledige eliminatie van de productie van aromatische foezelalcoholen tijdens groei op glucose van een *p*-coumaarzuurproducerende *S. cerevisiae*-stam, alsmede een verbeterde productie van *p*-coumaarzuur.

Summary

Over the past century, the world population has increased fourfold. This increase in human population is accompanied by an increased demand in food, water, energy and consumer goods. The resulting intensification of agriculture, deforestation, emission of green-house gasses and high utilization of natural compounds such as coal, minerals, metals and fossil-fuels have resulted in global warming and a depletion of Earth's natural reserves. The application of biotechnology can aid in the transition to a more sustainable, circular and bio-based economy. For example, by offering novel production processes for a range of different compounds, such as therapeutics, beverages, food, chemicals and fuels from renewable sources. For instance, Baker's yeast is known for its natural ability to produce CO_2 and ethanol from sugars, characteristics that were historically exploited for the production of alcoholic beverages and bread. Today, bio-ethanol as transportation fuel made by yeasts also provides a more sustainable alternative to gasoline. Additionally, due to the enormous increase in knowledge and the establishment of genome editing tools and sequencing possibilities, biotechnology can now apply genetically engineered microbes to produce an ever-increasing range of products, both native or heterologous, to the microorganism. For example, micro-organisms have been engineered to produce complex molecules such as human insulin, the flavoring compound vanillin and the antimalarial drug artemisinin. Insulin, which is essential for treatment of diabetes, was conventionally produced by extraction from pig pancreas, while artemisinin and vanillin were extracted from plants, or in case of vanillin, also synthesized chemically. However, microbial production of such industrially valuable compounds, from simple substrates such as glucose or second-generation feedstocks, offers a more reliable and sustainable production method compared to these classical methods. In this thesis special emphasis is given to the Baker's yeast *Saccharomyces cerevisiae* and its application in the production of aromatic compounds.

There is an increasing interest in the microbial production of aromatic molecules, such as in the flavor and fragrance industry. The economic potential of this field is partly due to European legislation, that allows the production and sale of microbially produced molecules, as long as the final product is devoid of genetically modified organisms (GMOs). *S. cerevisiae* is able to natively synthesize several aromatic compounds, although their production is limited by tight regulation of the involved pathways. Many other industrially attractive aromatic compounds find their origin in plants. In order to establish yeast-based production of these aromatic molecules, it is necessary to both introduce plant genes, and modify, the metabolism of *S. cerevisiae* to obtain fast and efficient production.

Chapter 1 provides an introduction to this research field by reviewing the impact of metabolic engineering of *S. cerevisiae* for the biosynthesis of aromatic compounds. After describing the trajectory of *Saccharomyces* species from traditional applications to modern biotechnology, this chapter analyses how *S. cerevisiae* has been exploited for the synthesis of a broad group of aromatic molecules with applications in among others the food industry, chemical industry and pharmaceutical industry. Although the main metabolic engineering strategies are similar, a distinction is made between native and new-to-yeast compounds. Next to its natural ability to produce CO_2 and ethanol, the yeast *S. cerevisiae* is also naturally able to produce flavor-active compounds such as higher alcohols ("fusel alcohols") which are derived from amino acid metabolism via the Ehrlich pathway. A special interest is given to the fusel alcohols 2-phenylethanol and *p*-hydroxyphenylethanol. These degradation products of phenylalanine and tyrosine respectively, are also valuable aromatic compounds in their own regard as they are used in, among others, the food and flavor industry. Furthermore, the aromatic amino acids phenylalanine and tyrosine can also serve as precursors for the biosynthesis of phenylpropanoids, a class of plant natural compounds including among others stilbenoids, tannins and flavonoids. Current production of these molecules is often based on the inefficient and unsustainable extraction from plants. Microbial production using renewable feedstocks represents a promising alternative for aromatics production.

However, in order to obtain industrially relevant production yields, titers and rates, several challenges have to be addressed. First of all, the intracellular availability of the metabolites required for phenylpropanoid biosynthesis, has to be increased. Secondly, the catalytic activity of rate limiting enzymes involved in phenylpropanoid biosynthesis has to be improved. Finally, another important aspect during process development is elimination of by-products, for example by adjusting the substrate specificity of involved enzymes. These three challenges are central in the research presented in this dissertation.

In **chapter 2,** several metabolic engineering strategies aimed at improving the flux through the shikimate pathway, the pathway responsible for the synthesis of aromatic amino acids, were investigated. In *S. cerevisiae*, a flux increase through this pathway translates into an increase of 2-phenylethanol. This natural product was therefore used as a model to evaluate the impact of the individual and combined genetic modifications. A combination of alleviation of the allosteric regulation of the essential points of the shikimate pathway, deregulation of key reactions of the pentose phosphate pathway and glycolysis, and reducing the capacity of a competing branch, which leads to the production of *p*-hydroxyphenylethanol, a tyrosine-derived fusel alcohol, resulted in significantly improved 2-phenylethanol production in *S. cerevisiae*. This research also provided a base for the biosynthesis of a diverse group of aromatic products, whose synthesis can be coupled to the shikimate pathway, such as phenylpropanoids.

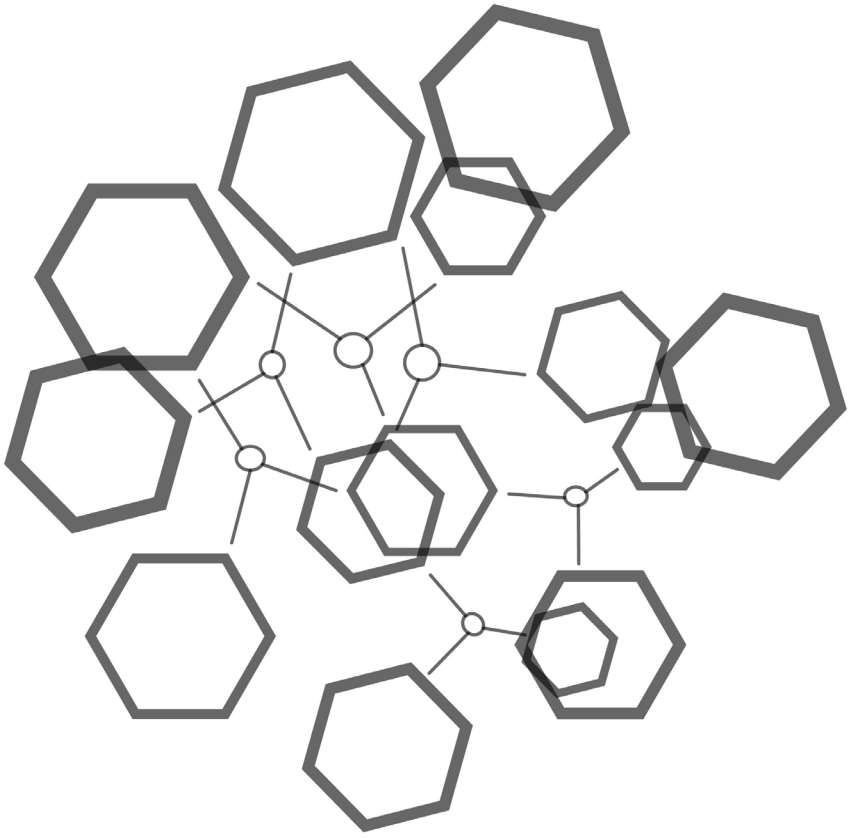
Chapter 3 explores how construction of neochromosomes (NeoChrs) can be deployed to construct *S. cerevisiae* strains where pathways and relevant reactions from central metabolism can fast and readily be optimized. *De novo* production of the industrial food dye anthocyanin pelargonidin 3-O-glucoside (P3G) in *S. cerevisiae* is showcased as proof of principle. Although already demonstrated in genetically engineered microorganisms, including Baker's yeast, the synthesis of P3G is highly inefficient in all microbial platforms tested hitherto. Furthermore, P3G synthesis in yeast is complex and requires the channeling of carbon through 27 reactions in the yeast central metabolism, including glycolysis, pentose phosphate pathway, shikimate pathway and aromatic amino acid biosynthesis, in addition to 10 reactions that originate from plants. To demonstrate the potential to express such complex pathways via synthetic chromosomes, modular NeoChrs in linear and circular configurations, harboring yeast genes, bacterial genes and plant genes required for P3G *de novo* synthesis, were designed and constructed. One strain harboring a linear Neochr, with nine copies of the gene encoding chalcone synthase, produced P3G extracellularly. This result signifies an improvement compared to previous studies where this compound remained almost exclusively intracellular.

The heterologous pathway for phenylpropanoids, a large family of aromatic molecules, which includes among others hydroxycinnamic acids, stilbenoids and flavonoids, starts with the deamination of the precursors L-phenylalanine and/or L-tyrosine, via the shikimate pathway. However, as discussed in **chapter 2**, the Ehrlich pathway can also convert these aromatic amino acids to aromatic fusel alcohols and acids, which are undesirable by-products in yeast strains engineered for production of aromatic compounds. Formation of these fusel alcohols and acids requires either of four *S. cerevisiae* 2-oxo-acid decarboxylases. However, elimination of this activity by deleting the responsible genes is not straightforward as these decarboxylases play a key role during growth on glucose in cytosolic acetyl-CoA biosynthesis. **Chapter 4**, describes the search for a decarboxylase that has the required pyruvate decarboxylase activity, but cannot decarboxylate aromatic 2-oxo acids. Out of eleven 2-oxo acid decarboxylases from yeast and bacteria, two bacterial enzymes from *Zymomonas mobilis* and *Gluconacetobacter diazotrophicus* showed the required substrate specificity. These enzymes were also able to restore growth on glucose of a *S. cerevisiae* strain in which all native decarboxylases were deleted. This decarboxylase swapping strategy completely eliminated aromatic fusel-alcohol production in glucose-grown batch cultures of an engineered coumaric acid producing *S. cerevisiae* strain, while simultaneously improving coumaric acid formation.

Chapter 1

General Introduction

One Ring To Make Them All



A small history of microbial biotechnology

Microbial biotechnology, defined as the application of microbial cultures for manufacturing products for human use, is among the oldest technologies in the world. Long before the science of microbiology appeared, mankind already used microbes for production of fermented foods and beverages. Archaeobotanical evidence showed that bread-like products were already consumed 14 millennia ago in north-eastern Jordan (1). Similarly, archaeological chemical evidence was recorded for production of fermented beverages in 7000 BC in an early Neolithic village in China (2), while evidence for Early Neolithic wine found in Georgia was dated to around 6000 BC (3). Moving from prehistory into recorded history, a Sumerian document originating from 3200-3000 BC, already described that beer was at that time a common product, part of centralized economy (4), while chemical evidence found by analysing ancient Sumerian jars confirmed that beer occurred in 4000 BC (5) in Mesopotamia.

It was not until the 17th century that, using his famous self-fabricated microscopes, Antonie van Leeuwenhoek became first to observe and report the existence of unicellular micro-organisms, for which he used the Dutch words 'kleine diertgens' ('little animals'). Up until this point, humans were unaware that tiny organisms, invisible to the naked eye, were involved in long-established processes for making products such as beer, wine, bread (6) and vinegar (7). Van Leeuwenhoek's discovery is often cited as the start of microbiology. However, it took another full century until, around 1840, it was discovered that microorganisms, in this case yeasts, were responsible for the fermentative production of ethanol (8). In 1857, Louis Pasteur confirmed the role of yeast cells in alcoholic fermentation (9). In the same period, Moritz Traube provided the first demonstration of cell free fermentation. He called the molecules responsible for fermentation "ferments", now better known as enzymes (10).

The discovery of microorganisms and their fermentative metabolism ushered in the era of chemical or industrial biotechnology, which covers the industrial use of microorganisms to produce a wide range of chemical compounds. Some of the early applications included production of α-amylases by *Bacillus amyloliquefaciens* (11), acetone and butanol by *Clostridium acetobutylicum* (12), glycerol by *Saccharomyces cerevisiae* (13), citric acid by *Aspergillus niger* (14) and L-ascorbic acid (vitamin C) by *Gluconobacter oxydans* (15).

These early industrial biotechnology applications all relied on natural characteristics of microorganisms isolated from nature. Additionally, their products had relatively simple structures and were closely associated to, or even part of, central carbon metabolism. The first complex compound that was efficiently produced at an industrial scale by a microorganism was the β-lactam antibiotic penicillin, discovered in 1928 by Alexander Fleming (16). The start of the Second World War led to a high need of penicillin, which urged the intensification of penicillin production. After WW2, process optimization became ever more strongly linked with optimization of the *Penicillium* strains responsible for penicillin production. For several decades, strain optimization relied exclusively on random mutagenesis using chemical muta-

gens, X-rays, or ultraviolet radiation (17).

The first half of the 20th century witnessed tremendous progress in the elucidation of enzymes and metabolic pathways, including central metabolic routes such as glycolysis (for an overview see work of Barnett (18)) and the tricarboxylic acid cycle (19). Another milestone was reached in 1953 with the elucidation of the structure of DNA (20). This breakthrough was followed by the first demonstration of recombinant-DNA technology (21) and the market introduction of recombinant human insulin as first product of a genetically engineered *Escherichia coli* strain (22, 23). In 1995, the first bacterial genome (of *Haemophilus influenza*) (24) was fully sequenced, followed one year later by the first eukaryotic genome (of the yeast *S. cerevisiae*) (25). Sequencing of these and other microbial genomes were stepping-stones towards the completion of the Human Genome Sequencing project in 2001 (26, 27). The combination of ever more detailed knowledge on metabolic pathways and genomes with the increasingly availability of powerful techniques for genetic modification led to a new field of research called metabolic engineering (28). In metabolic engineering, knowledge-based genetic modification of metabolic, transport and regulatory functions of cells is applied to extend and improve their industrial application. In addition, the complexity of genomes, metabolic networks and regulatory networks provided a powerful incentive for the development of genomics and of systems biology, which integrates experimental biology with mathematical modelling and synthetic biology. In the second decade of the 21st century, targeted strain engineering, now often referred to as 'genome editing' received a further boost by the identification of the Cas9 endonuclease and the demonstration of its use as a versatile tool for precise, fast and simultaneous multi-target genome editing (29, 30).

Microbial production of valuable products

Implementing synthetic biology techniques has significantly facilitated and accelerated engineering of microorganisms for the production of chemicals, both by optimizing their native metabolic networks and by introducing novel, heterologous or novel-to-nature enzymes and/or pathways. The types of products made can range from relatively simple molecules such as amino acids, organic acids and alcohols to structurally more complex compounds such as terpenoids, flavonoids, cannabinoids (31) and opioids (32, 33).

Compounds produced by wild-type or genetically modified micro-organisms can be differentiated into two major groups based on their connection to cellular energy metabolism. Assimilatory products require a net input of ATP for their formation from simple substrates. In contrast, formation of dissimilatory products leads to a net synthesis of ATP, as for example in the case of anaerobic ethanol and lactate production by fermentative organisms. Some of the challenges encountered in optimization of strain performance by metabolic engineering are agnostic to the type of product. For example, limitations in transport of substrates,

products and intermediates (34), formation of undesired by-products (35), suboptimal energy coupling of key reactions (36), toxicity of products or intermediates (37) and metabolic burdens (38) caused by high-level expression of key enzymes can occur for assimilatory as well as for dissimilatory products. Moreover, natural regulatory mechanisms, which have evolved over millions of years to maximize fitness in natural environments, may counteract high-level product formation by mechanisms such as transcriptional, translational and post-translational regulation. For example, biosynthesis of amino acids is generally tightly regulated to meet the demands of protein synthesis and accumulation of free amino acids is prevented by allosteric feed-back inhibition of key enzymes in their biosynthetic pathways. Elimination of such feed-back inhibition or repression mechanisms is then a prerequisite for achieving economically relevant product yields (39).

All heterologous product pathways that are implemented into microorganisms require supply of precursor molecules by their native metabolic networks. For example, production of isoprenoid-derived products such as farnesene and amorphadiene, as well as many other products, by engineered yeast strains depends on provision of acetyl-coenzyme A as a precursor (40). Since regulation networks in microbial cells have evolved to accurately 'tune' precursor supplies to the requirements for growth, modification of the host cell's metabolic network is often required to improve precursor supply. In addition to overcoming such regulation-related constraints, achieving high product yields may require changes to the native metabolic network to optimize energy coupling of precursor supply (41). For example, the native pathway for cytosolic acetyl-CoA synthesis is very ATP-intensive and has been replaced by alternative, ATP-independent pathways to improve yields of industrially relevant products on glucose (42, 43).

Finally, expression of heterologous genes can introduce requirements for synthesis of specific cofactors (33, 44), transport of metal ions (45, 46), expression of chaperones (47) or other cellular processes that are not, or not optimally, met by the host microorganism. In addition, these limitations need to be eliminated by further strain engineering to enable economically viable industrial production.

Natural production of flavour-active compounds by *S. cerevisiae*

The yeast *S. cerevisiae* and *Saccharomyces* hybrids have been used and domesticated for centuries because of their natural ability to ferment sugars to ethanol and CO_2 as well as their ability to produce flavour and aroma compounds. These two characteristics were essential for their initial applications: the production of alcoholic beverages and bread. Indeed, most industrial *Saccharomyces* strains appear to have been selected for high-level production of a cocktail of aroma compounds (48, 49)

The absolute and relative concentrations of individual aroma and flavour compounds are yeast-strain dependent and, moreover, depend on environmental conditions (50). For example, medium composition (nitrogen source, sugar used as substrate etc.) (51), cultivation temperature (51) and oxygen availability (52) all strongly affect flavour profiles (53). The most important aroma-related metabolites of yeasts include, but are not limited to, esters, aldehydes, phenols, carbonyl compounds, sulfur-containing compounds, organic acids and higher alcohols (49).

Despite their generally low concentrations in fermented products, volatile esters are very important contributors to the aroma of alcoholic beverages (54). Ethyl esters are formed through the condensation of ethanol and the medium-chain-length acyl-Coenzyme A molecules, with ethyl octanoate (fruity, floral odour) and ethyl hexanoate (pineapple odour) as important examples in alcoholic beverages (54). These volatile acetate esters are formed by condensation of a higher alcohol and acetyl-coenzyme A (55). Important aroma compounds belonging to this class include isoamyl acetate (IA) (53) which has a fruity, banana-like aroma, ethyl acetate (EA) which confers a solvent-like aroma and the aromatic ester phenylethyl acetate which has a rose/honey-like aroma (56). In addition to their acetate esters, higher alcohols are also important flavour active compounds in their own right and are the most abundant organoleptic compound in terms of concentration in beer (69, 70). The most abundant higher alcohols in beer are amyl alcohol, which affects beer drinkability and isobutyl alcohol, which has a sweet, musty odour (71), while 2-phenylethanol has a pronounced rose-like smell (72). In addition to the roles of these yeast metabolites as flavour and aroma compounds in the food and beverage industry, their characteristics can be exploited by the flavour and fragrance industry, for example by their application in perfumes, personal care products, cleaning products, nutraceuticals etc. (**Table 1**). Especially for these non-food applications, application of metabolic engineering to boost productivity and yield of individual aroma compounds by yeast strains offers interesting possibilities (73, 74).

Table 1. An overview of the main applications of yeast derived aromatic products and example compounds.

Application	Example Compound(s)	Reference
Beverages	Phenylethyl acetate	(56)
Fuels	Limonene	(57)
Preservatives/ Anti-Microbial	Eriodictyol	(58, 59)
Colorants	Pelargonidin 3-O-glucoside	(60-62)
Flavorings	Vanillin	(63)
Fragrances	2-phenylethanol	(64)
Nutraceuticals	Multitude of phenylpropanoids	(65, 66)
Therapeutics	Psilocybin, artemisinin, tyrosol	(40, 67, 68)

The Ehrlich Pathway

Yeasts produce several higher alcohols, also termed fusel alcohols, in a pathway first described by Felix Ehrlich in 1907 (75) and therefore referred to as the Ehrlich Pathway (76) (**Fig. 1**). This pathway describes the catabolism of the branched-chain amino acids valine, isoleucine and leucine, the sulphur-containing amino acid methionine and the aromatic amino acids phenylalanine, tyrosine and tryptophan. These amino acids can either be obtained from the medium or be derived from amino-acid biosynthesis.

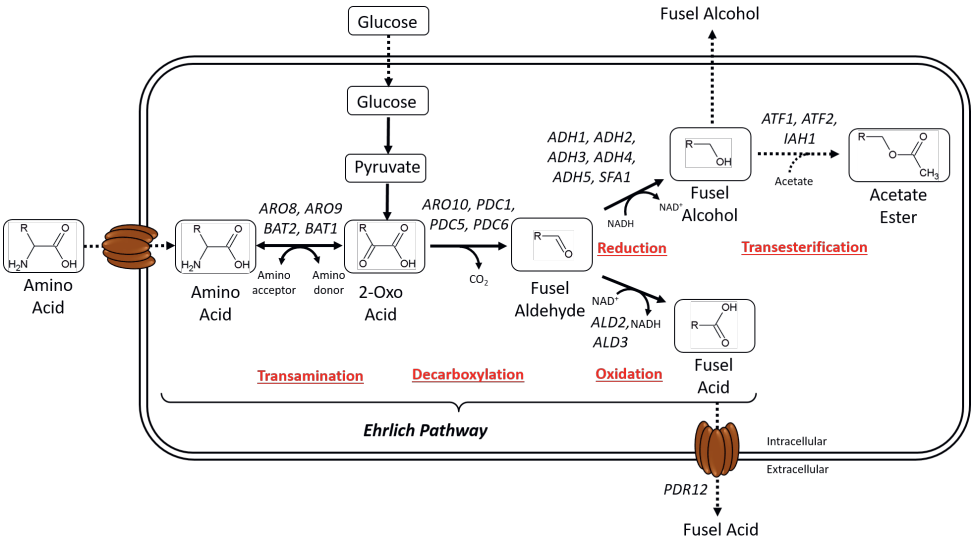


Figure 1. Schematic representation of the Ehrlich pathway in yeast for conversion of amino acids into fusel alcohols and fusel acids, and the subsequent formation of acetate esters. Amino acids can either be taken up from the medium or derived from amino acid biosynthesis. R indicates an amino acid specific moiety. Reaction mechanisms are underlined and in red. Gene names refer to structural genes in S. cerevisiae that encode enzymes able to catalyse the indicated reactions.

The Ehrlich pathway is initiated by transamination of an amino acid to the corresponding 2-oxo acid (2OA). In *S. cerevisiae,* valine, leucine and isoleucine can be transaminated by either the mitochondrial branched-chain amino-acid aminotransferase, Bat1, or by the cytosolic Bat2 isoenzyme (69). Two other *S. cerevisiae* transaminases, Aro8 and Aro9, were initially shown to catalyse the transamination of tyrosine, phenylalanine and tryptophan and are therefore annotated as aromatic amino acid transferases (77, 78). Further research showed activity of Aro8 with methionine, leucine and α-aminoadipate (79), while Aro9 was shown to have a broader substrate specificity than first assumed (79). Aro8 and Aro9 are therefore now considered to be broad-substrate-specificity amino acid transaminases (76). The 2-oxo acids formed by transamination are subsequently decarboxylated to yield the corresponding aldehyde (80, 81). This irreversible reaction is mainly catalysed by the thiamine-pyrophosphate

(TPP)-dependent phenylpyruvate decarboxylase (Aro10) (82, 83), but the *S. cerevisiae* pyruvate decarboxylases Pdc1, Pdc5 and Pdc6 can also catalyse this reaction. Depending on the redox status of the cells, the produced fusel aldehydes are oxidized to fusel acids or reduced to fusel alcohols (81) (**Table 2**).

Table 2. An overview of amino acids (AA) and their derived 2-oxo acids, fusel aldehydes, fusel acids and fusel alcohols generated via the Ehrlich pathway in yeasts. Adapted from Hazelwood *et al., 2008*.

AA	2-oxo acid	Fusel aldehyde	Fusel Acid	Fusel Alcohol
Ile	α-ketomethylvalerate	methylvaleraldehyde	methylvalerate	active amyl alcohol
Leu	α-ketoisocaproate	isomylaldehyde	isovalerate	isoamyl alcohol
Met	α-keto-Y-(methylthio) butyrate	methional	3-(methylthio) propanoate	methionol
Phe	phenylpyruvate	2-phenylacetaldehyde	2-phenylacetate	2-phenylethanol
Trp	3-indolepyruvate	3-indoleacetaldehyde	2-(indol-3-yl) ethanoate	tryptophol
Tyr	ρ-hydroxy	*p*-hydroxy	*p* -hydroxy	*p* -hydroxy
	phenylpyruvate	phenylacetaldehyde	phenylacetate	phenylethanol
Val	α-ketoisovalerate	isovaleraldehyde/ isobutanal	isobutyrate	isobutanol

As observed for other aroma and flavour compounds, production of fusel alcohols is strain dependent and, moreover, strongly depends on the availability of amino acids in growth media. When ammonium sulfate, urea, arginine, asparagine, proline or glutamic acid are used as sole, non-growth-limiting, nitrogen source, little to no fusel alcohols are produced (81, 84). In contrast, using the amino acids leucine, methionine and phenylalanine as nitrogen source leads to abundant fusel alcohol production by fermenting yeast cultures (81). Similarly, when supplying 2-oxo acids to yeast cultures, a direct proportional increase in higher alcohol concentration could be observed (55).

The shikimate pathway and aromatic amino acid biosynthesis in yeast

Of the fusel alcohols produced by yeast, the aromatic compounds *p*–hydroxyphenylethanol (*p*OH2PE) and 2-phenylethanol (2PE) are of special interest in the context of this thesis. 2-phenylethanol is extensively used by the food and fragrance industry due to its characteristic rose-like odor. Additionally its antifungal and antibacterial properties have raised interest for its application in disinfectants and cleaning products (64). The annual global production of 2-phenylethanol was estimated at around 10,000 tons in 2010 (85). *p*–Hydroxyphenylethanol, which has a bitter taste and occurs in high amounts in olive oil (86) and sake wine (67) has antioxidant and anti-inflammatory activities (67).

p–Hydroxyphenylethanol and 2-phenylethanol are derived from L-phenylalanine (PHE) and L-tyrosine (TYR), respectively. Synthesis of these aromatic amino acids via the so-called shikimate pathway starts by condensation of the glycolytic intermediate phosphoenolpyruvate (PEP) with erythrose-4-phoshate, which is derived from the pentose-phosphate pathway. This condensation yields the 7-carbon compound deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) which, in six subsequent reactions catalysed by Aro1-Aro4 and Aro7, is converted to chorismate (87-90). Chorismate is a precursor for tryptophan biosynthesis but can also be converted into prephenate (PPA) by chorismate mutase (Aro7). The prephenate dehydratase Pha2 and the prephenate dehydrogenase Tyr1, convert prephenate into the 2-oxo acids phenylpyruvate (PPY) or p-hydroxyphenylpyruvate (*p*OHPPY), respectively. These 2-oxo acids can then be transaminated into phenylalanine and tyrosine, respectively. Alternatively, they can be metabolized via the Ehrlich pathway, which leads to production of the fusel alcohols 2-phenylethanol and p-hydroxyphenylethanol respectively or, alternatively, the corresponding fusel acids phenylacetate (PAA) and *p*-hydroxyphenylacetate (*p*OHPAA) (**Fig. 2**). To produce high levels of fusel alcohols or other aromatic compounds, a first essential step is to eliminate allosteric feedback inhibition of several enzymes in the shikimate pathway. The DAHP synthases, Aro3 and Aro4, are feedback inhibited by phenylalanine and tyrosine respectively (39, 89) while the chorismate mutase Aro7 is inhibited by tyrosine but activat-

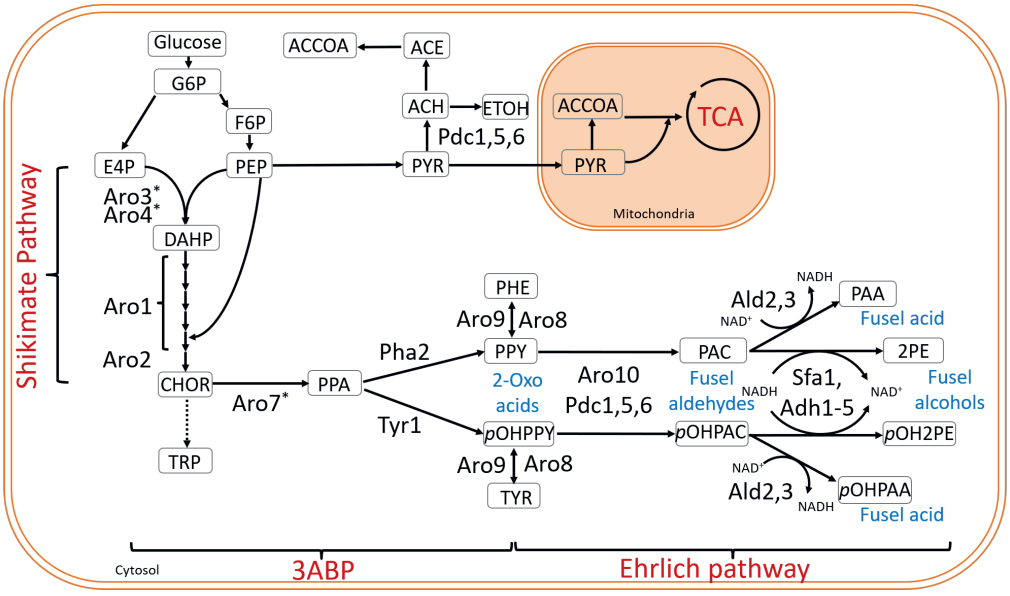


Figure 2. Metabolic pathways involved in biosynthesis and degradation of the aromatic amino acids phenylalanine and tyrosine in *S. cerevisiae*. G6P glucose-6-phosphate, F6P fructose-6-phosphate, ACE acetate, ACCOA acetyl-coA, TCA tricarboxylic acid cycle, PYR pyruvate, ETOH ethanol, PEP phosphoenolpyruvate, E4P erythose-4-phosphate, DAHP 3-deoxy-d-arabino-heptulosonate-7-phosphate, CHOR chorismate, PPA prephenate, PPY phenylpyruvate, PHE L-phenylalanine, PAC phenylacetaldehyde, PAA phenylacetate, 2PE 2-phenylethanol, *p*OHPPY *p*-hydroxyphenylpyruvate, TYR L-tyrosine, *p*OHPAC *p*-hydroxyphenylacetaldehyde, *p*OHPAA *p*-hydroxyphenylacetate, *p*OH2PE *p*-hydroxyphenylethanol, TRP tryptophan, 3ABP aromatic amino-acid biosynthesis. * Enzyme subject to allosteric feedback inhibition.

ed by tryptophan (91-93). Expression of feedback-resistant versions of these three proteins have been shown to increase the flux through the shikimate pathway such as Aro4^{K229L} (39), Aro3^{K222L} (94), Aro7^{T226I} (91) and Aro7^{G141S} (39, 92)

Phenylpropanoid pathway

In addition to the role of the shikimate and Ehrlich pathways in the production of fusel alcohols, these pathways are also highly relevant for expression in yeasts for heterologous pathways leading to aromatic products. Many bio-active aromatic compounds are derived from the phenylpropanoid pathway in plants, which starts from the transamination of the aromatic amino acids phenylalanine or tyrosine into cinnamic acid or coumaric acid, respectively. From coumaric acid different kinds of the secondary metabolite phenylpropanoids can be made. To cope with biotic and abiotic stresses and attract the right pollinators, plants have evolved to produce a wide range of secondary metabolites (95), which are not directly involved in growth but do confer an evolutionary advantage. For example, anthocyanins give color to flowering plants, which helps to attract pollinators, repel plant pest organisms and, for some anthocyanins, confers antimicrobial properties (96). The phenylpropanoids include hydroxycinnamic acids, stilbenoids, coumarins, flavonoids, lignin and hydrolysable tannins (**Fig. 3**). Many of these compounds are known to possess nutraceutical properties (for a review see Neelam et al. (66)).

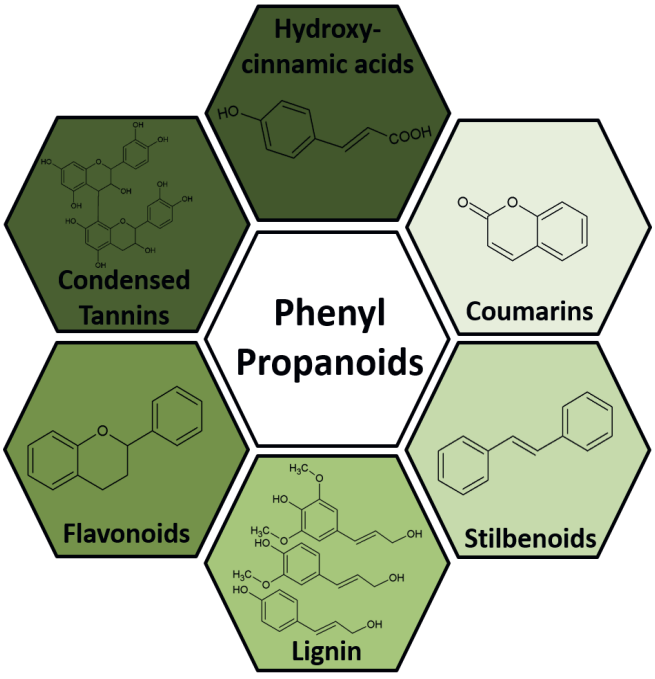


Figure 3. The main classes of phenylpropanoids and their core chemical structure. For lignin, the three main monolignols are presented: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. For tannins the basic structural unit, a catechin monomer, is shown.

Although the phenylpropanoid pathway is present in most plants (97), plant biomass typically contains very low concentrations of the derived bio-active products. Therefore, large amounts of plant biomass are needed for production of economically relevant amounts of an individual product (98, 99). Levels of relevant compounds in plants are subject to environmental, seasonal and geographical variation and on plant species or tissue (95). Moreover, plants often produce a range of structurally related phenylpropanoid products, which complicates purification of individual compounds. These factors, combined with the slow growth of plants are major impediments for economically feasible production of phenylpropanoids from plant biomass (100).

Chemical synthesis of phenylpropanoids only provides an alternative for extraction from plant biomass for compounds with simple structures. For more complex compounds such as glycosylated flavonoids, the presence of multiple active functional groups and modifications at particular positions (101) make chemical synthesis difficult and therefore requires biotransformation. Furthermore, chemical synthesis often requires the use of toxic chemical solvents and extreme reaction conditions, which affects product yield and scalability of production processes (102). Additionally, European and US legislation limits the use of chemically produced food additives (103) and consumers increasingly prefer chemicals from natural sources for applications such as food colouring (104).

All prokaryotes and eukaryotes that are amino acid prototrophs natively produce the phenylpropanoid precursors phenylalanine and tyrosine. Introducing the biosynthetic genes required for the synthesis of phenylpropanoids into industrial microorganisms could therefore open up interesting possibilities for production of these compounds. Indeed, multiple successful examples for production of valuable aromatic compounds by genetically modified bacteria and yeasts have been reported as reviewed by Yang et al. (105), Chen et al. (106) and Liu et al. (107). For production of more complicated aromatic structures such as decorated flavonoids, yeasts are currently the favourite production platform due to the higher similarity to plant cells in terms of compartmentalization, organelles, post-transcriptional and post-translational modification machinery (108, 109).

When phenylalanine is the precursor for phenylpropanoid synthesis (**Fig. 4**), it is first converted into cinnamic acid by the phenylalanine ammonia lyase (PAL). The cinnamate 4-hydroxylase (C4H) and a cytochrome P450 reductase (CPR) then convert cinnamic acid into coumaric acid. Alternatively, tyrosine can directly be converted into coumarate by tyrosine ammonia lyase (TAL). The latter reaction does not require the involvement of a p450 oxidoreductase.

From coumaric acid, several other hydroxycinnamic acids (C_6-C_3) can be produced such as caffeic acid, ferulic acid and sinapic acids. Hydroxycinnamic acids can be found in a wide variety of consumables such as beverages, cereals, vegetables and fruits (110). Similar to

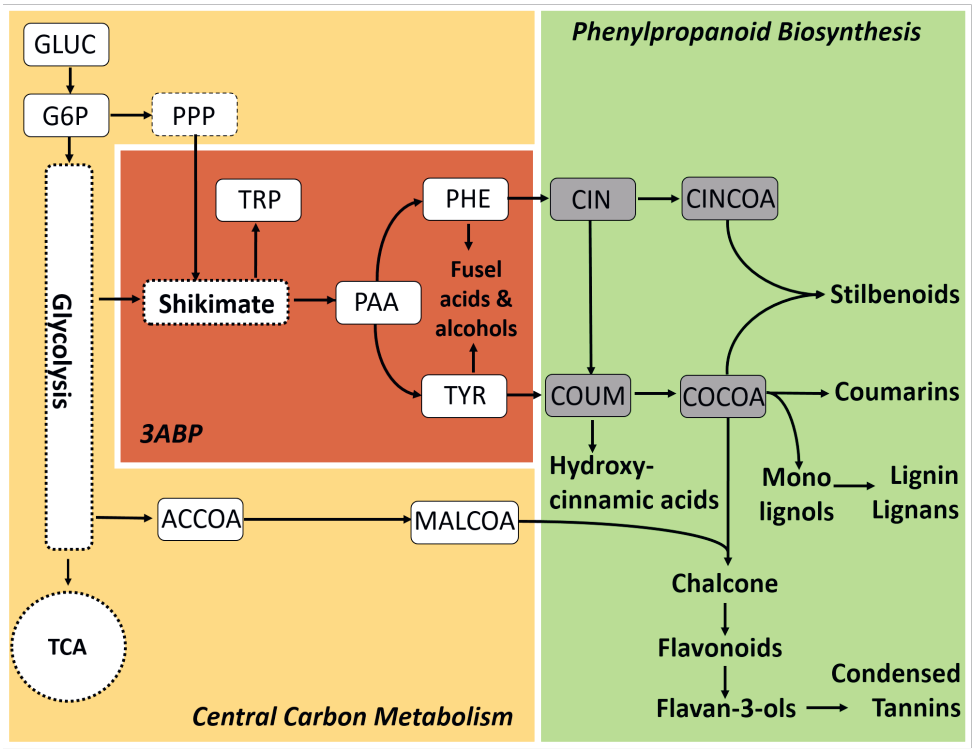


Figure 4. An overview of the main biosynthetic pathways involved in the biosynthesis of most phenylpropanoids. The precursors native to yeast are presented in white boxes, important intermediates of the phenylpropanoid pathway heterologous to yeast are depicted in grey boxes. Pathways are indicated with a dotted line. The central carbon metabolism is depicted in orange, the aromatic amino acid biosynthetic pathway (3ABP) in red and the phenylpropanoid biosynthetic pathway in green. GLUC glucose, G6P glucose-6-phosphate, PPP pentose-phosphate pathway, TCA tricarboxylic acid cycle, TRP tryptophan, TYR tyrosine, PHE phenylalanine, PAA, prephenate, COUM coumaric acid, CIN cinnamic acid, CINCOA cinnamoyl-CoA, COCOA coumaroyl-CoA, ACCOA acetyl-CoA, MACOA malonyl-CoA.

many phenylpropanoids, these molecules possess properties that are beneficial to health such as antioxidative, antibacterial, ant-proliferative, anti-inflammatory and anti-carcinogenic properties (111).

In a next reaction, cinnamic acid or coumaric acid are activated by the cinnamate-CoA ligase (CNL) or the 4-coumaric acid-CoA ligase (4CL) respectively, yielding cinnamoyl-CoA or coumaroyl-CoA. Stilbenoid synthesis from cinnamoyl-CoA or coumaroyl-CoA is initiated by stilbene synthases. Several yeast strains have successfully been engineered to produce resveratrol and other stilbenoids as reviewed previously (112, 113). Coumaroyl-CoA is also an important branchpoint in phenylpropanoid synthesis. For example, coumarins can be produced from p-coumaroyl-CoA but pathways starting from *p*-coumaric acid or cinnamic acid as precursors have also been described (114).

Surpassed only by cellulose, lignin is the second most abundant biopolymer on Earth. It gives mechanic support to plants and is therefore mainly found in cell walls and stems (115). Monolignols, which serve as the main building blocks for lignin, are synthesized from 4-coumaryl-CoA. The most common monolignols, p-coumaryl alcohol (unit H), coniferyl alcohol (unit G) and sinapyl alcohols (unit S), as well as other monolignols are polymerized to yield lignans (2 monolignol units) or lignins (>2 units).

Tannins are a very diverse group of high-molecular-weight polyphenols, whose main physiological role is to provide protection against abiotic and biotic stresses. Tannins, which occur in a wide variety of fruits, cereals and cacao (65) are divided into two groups, condensed tannins (proanthocyanidins) and hydrolysable tannins. The latter are produced from gallic acid (GAL), which is produced from 3-hydroshikimate, an intermediate of the shikimate pathway, while condensed tannins result from polymerization of flavan-3-ols or flavan-3-4-diols (Fig. 4)

Flavonoids: types and purposes

Flavonoids comprise a group of polyphenolic compounds that are naturally produced by a wide variety of plants. Over 9,000 different flavonoid molecules have currently been documented (116). Anthocyanins, a specific group of water-soluble flavonoids, are commonly used as cosmetics and as colouring agents (60). However, the interest in flavonoid production is primarily related to their use as nutraceuticals or even as therapeutics. Purported health benefits of flavonoids include anti-oxidant (117, 118), analgesics, anti-inflammatory (119), anti-microbial (58, 120), anti-thrombotic (121), anti-diabetic (122), anti-carcinogenic and anti-angiogenic properties (123).

The basic structure of a flavonoid consists of two phenyl rings (A and B) connected by a heterocyclic ring (C), together forming a C_6-C_3-C_6 carbon skeleton (Fig. 5). Based on the type of connection of the phenyl B ring to the C ring, compounds are classified as a flavonoid (C2), isoflavonoid (C3) or a neoflavonoid (C4). Flavonoids are further classified depending on the presence of a double bond, a ketone group or a hydroxyl group in the C ring.

Flavonoids are a highly diverse family of compounds due to these differences in basic structure as well as the variety of further modifications of the basic structure, including methylation, glycosylation, prenyltransferation, hydroxylation and acylation (101, 124). For example, a common modification strategy is the glycolysation of anthocyanidins, resulting in stable anthocyanin (125). The type and degree of modification affects molecular properties. For example, a higher number of hydroxyl groups (126) or a specific position of the hydroxyl group (127) leads to increased lipophilicity. Other modifications such as acrylation (128) or

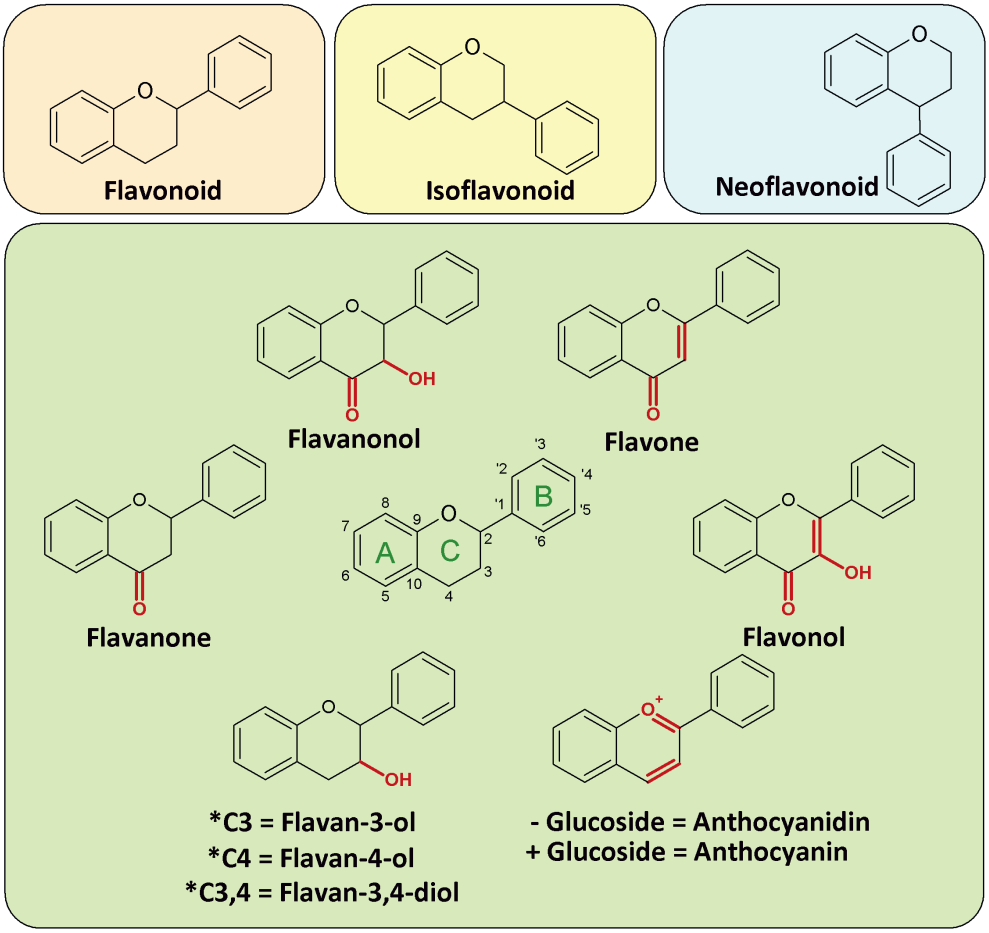


Figure 5. The different types of flavonoids. The central flavonoid core structure is depicted in the middle. The specific modification to the C-ring that determines the type of flavonoid is highlighted in red. *C3 and *C4 represent to which carbon atom the hydroxyl group is attached.

glycolysation (129) can increase the solubility and stability of flavonoids. Furthermore, the biological activity (129-131) of a flavonoid also differs depending on the type and level of modification. Synthesis of flavonoids starts with the conversion of one molecule of coumaroyl-CoA and three molecules of malonyl-CoA into naringenin chalcone in a reaction catalysed by chalcone synthase (Chs). Naringenin chalcone can spontaneously isomerize into naringenin or, alternative, this reaction can be catalysed by a chalcone isomerase (CHI). From naringenin onward, pathways branch off towards specific flavonoids and modification reactions (**Fig. 6**). So far, only *S. cerevisiae* and *Y. lipolytica* have been metabolically engineered for *de novo* production of several flavonoids such as the central flavonoid naringenin (132, 133) or higher-value derivatives (**Table 3**).

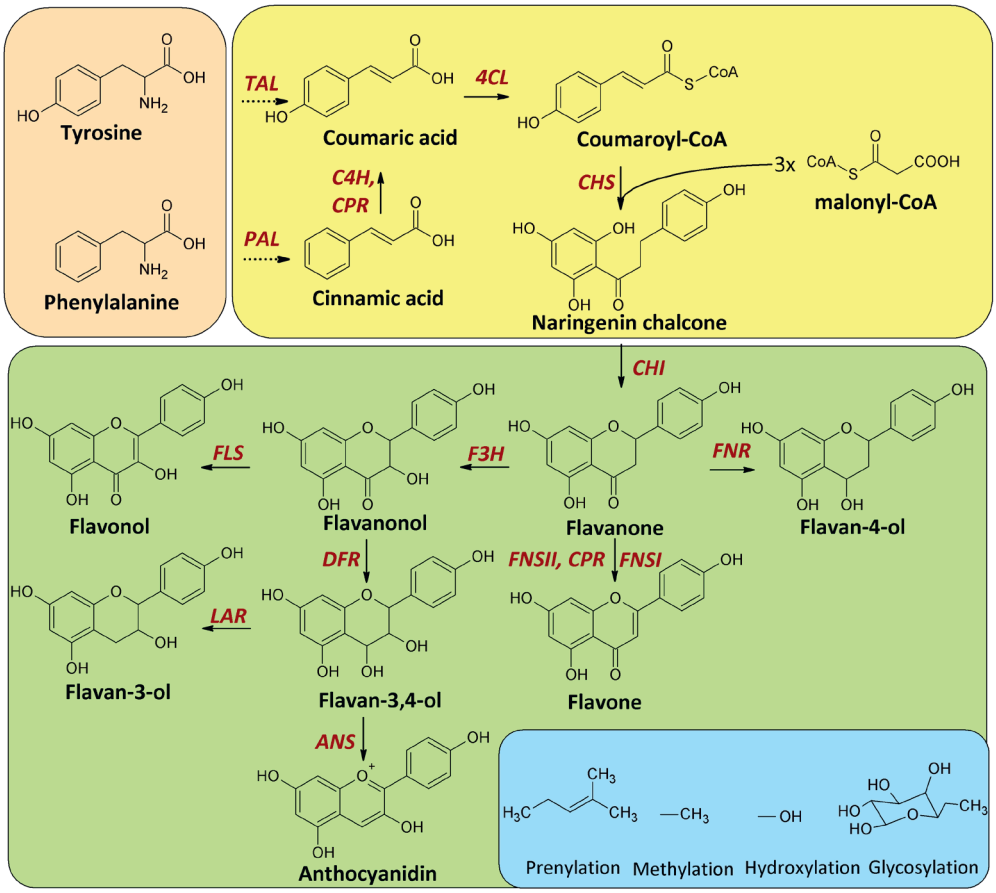


Figure 6. Flavonoid biosynthesis in yeast starting from the precursors native to *S. cerevisiae*: phenylalanine and/or tyrosine (indicated in a light red box). The yellow box depicts the pathway for synthesis of the central flavonoid intermediate naringenin (a flavanone). The green box shows reactions leading to the formation of different types of flavonoids. The blue box represents common modification strategies that flavonoid at different steps can be subjected to. Structural plant enzymes of key reactions are indicated in red. Dotted arrows indicate optional reactions as coumaric acid can be synthesized from tyrosine and/or from phenylalanine. Abbreviations: TAL tyrosine ammonia lyase, 4CL 4-coumarate-CoA ligase, PAL phenylalanine ammonia lyase, C4H cinnamate 4-hydroxylase, CPR cytochrome p450 reductase, CHS chalcone synthase, CHI chalcone isomerase, FLS flavonol synthase, F3H flavanone 3-hydroxylase, FNR Flavanone 4-reductase, LAR leucoanthocyanidin reductase, DFR dihydroflavonol 4-reductase, FNSI soluble flavone synthase, FNSII membrane-bound flavone synthase, ANS anthocyanidin synthase.

Table 3. Overview of the current status of *de novo* production of phenylpropanoids by yeast. Only one reference per compound is given as example. All strains were grown on synthetic medium with glucose as carbon source unless indicated otherwise. For most studies the yield (Y) per substrate was initially not mentioned and thus calculated. However, in most cases there was not sufficient information available (M.I. missing information). Org organism, Cult cultivation method, HCA hydroxycinnamaic acid, *Sc S. cerevisiae, Yl Y. lipolytica,* YPD yeast peptone medium, YNB yeast nitrogen base, Gal galactose, ceru cerulin, Fed Fed-Batch, Shake Shake-Flask, DWP Deep well plate, BioR Batch controlled bioreactor, FIT synthetic feed-in-time (FIT) fed-batch cultivation, ACD anthocyanidin, LIG lignin, NQ not quantified, N.A. not available, #µmol/gCDW

	Compound	**Org**	**Titer (mM)**	**Yield (mmol mol^{-1})**	**Cult.**	**Ref.**
HCA	Coumaric acid	*Sc*	76.2	170.1	Fed	(134)
Stilbenoid	Pinostilbene	*Sc*	0.023	M.I.	FIT	(135)
	Pterostilbene	*Sc*	0.14	M.I.	FIT	(135)
	Resveratrol	*Yl*	54.3	42.2	Fed	(136)
	Polydatin	*Sc*	1.40	10.1	BioR	(137)
(Dihydro)chalcone	Phloretin	*Sc*	0.156	M.I.	DWP	(138)
	Pinocembrin dihydrochalcone	*Sc*	0.010	M.I.	DWP	(138)
	Naringin dihydrochalcone	*Sc*	0.020	M.I.	DWP	(138)
	3-OH phloretin	*Sc*	0.099	M.I.	DWP	(138)
	Nothofagin	*Sc*	0.136	M.I.	DWP	(138)
	Phlorizin	*Sc*	0.149	M.I.	DWP	(138)
	Trilobatin	*Sc*	0.075	M.I.	DWP	(138)
ACD	Pelargonidin	*Sc*	N.A.	0.01#	BioR	(62)
Antho-cyanine	Cyanidin 3-O-glucoside	*Sc*	~0.0043	M.I.	DWP	(61)
	Delphinidin 3-O-glucoside	*Sc*	~0.0042	M.I.	DWP	(61)
	Pelargonidi 3-O-glucoside	*Sc*	N.A.	0.001#	Shake	(62)
	Pelargonidin 3-O-glucoside	*Sc*	~0.0040	M.I.	DWP	(61)
flavan	Afzelechin	*Sc*	~0.33	M.I.	DWP	(61)
	Catechin	*Sc*	~0.33	M.I.	DWP	(61)
	Gallocatechin	*Sc*	~0.082	M.I.	DWP	(61)
Flavanone	Liquiritin	*Sc*	0.0011	0.0032	Fed (Gal)	(139)
	Naringenin	*Yl*	3.3	M.I.	BioR + Pulse	(133)
	Eriodictyol	*Yl*	0.46	M.I.	Fed + Ceru	(140)
	Liquiritigenin	*Sc*	0.021	M.I.	FIT	(98)
	Pentahydroxy-flavanone	*Sc*	NQ	M.I.	DWP	(61)
	Pinocembrin	*Sc*	0.0031	M.I.	Shake (YPD)	(100)
	8-Prenyl naringenin	*Sc*	0.0035	M.I.	Shake	(141)

flava-nonol	dihydro kaempferol	*Sc*	~0.47	M.I.	DWP	(61)
	dihydro myricetin	*Sc*	~0.27	M.I.	DWP	(61)
	Taxifolin	*Sc*	1.11	1.72	Fed	(142)
Flavone	Breviscapine*	*Sc*	0.63	M.I.	Fed	(143)
	Chrysin	*Sc*	0.028	M.I.	Shake	(144)
	Baicalein	*Sc*	0.011	M.I.	Shake	(144)
	Baicalin	*Sc*	0.012	M.I.	Shake	(144)
	Norwogonin	*Sc*	0.026	M.I.	Shake	(144)
	Wogonin	*Sc*	0.016	M.I.	Shake	(144)
	Isowogonin	*Sc*	0.015	M.I.	Shake	(144)
	Moslosoo flavone	*Sc*	0.010	M.I.	Shake	(144)
Flavonol	Fisetin	*Sc*	0.0080	M.I.	FIT	(98)
	Kaempferol	*Sc*	0.23	M.I.	Fed	(145)
	Myricetin	*Sc*	~0.063	M.I.	DWP	(61)
	Quercetin	*Sc*	~0.12	M.I.	DWP	(61)
	resokaempferol	*Sc*	0.0019	M.I.	FIT	(98)
	Kaempferol 3-O-glucoside	*Sc*	~0.30	M.I.	DWP	(61)
	Myricetin 3-O-glucoside	*Sc*	~0.011	M.I.	DWP	(61)
	Quercetin 3-O-glucoside	*Sc*	~0.11	M.I.	DWP	(61)
Lig	Coniferyl alcohol	*Sc*	1.12	2.32	Fed	(142)

Improving phenylpropanoid biosynthesis

Although yeast has been engineered to produce many different phenylpropanoids, product titers in most cases remain well below 1.0 mM, especially when looking at compounds further down the phenylpropanoid pathway such as flavonoids (**Table 3**). Optimizing titers and yields of flavonoids and other phenylpropanoids by metabolic engineering is therefore an important prerequisite for further exploration of the biological activities of this diverse class of compounds and for development of economically viable microbial production platforms. Roughly speaking, there are two levels that need to be considered when addressing this challenge: improvements aimed at increasing the precursor supply for phenylpropanoids (phenylalanine, tyrosine and sometimes malonyl-CoA) and improvements of the phenylpropanoid pathway itself (**Table 4**).

Table 4. Metabolic engineering of yeast for the production of phenylpropanoids. The table lists the main issues encountered during *de novo* yeast-based production of phenylpropanoids, common metabolic engineering (Metab. Eng.) strategies that have been used to alleviate the problems and some important genes that are involved in each step. *Se: Salmonella enterica,* *heterologous gene.

Level	Problem	Metab. Eng. Strategies	Enzymes	Ref.
Limited Precursors Supply	Phenylalanine Tyrosine	Overexpressing (feedback resistant) alleles of the shikimate pathway. Expressing *E. coli* shikimate genes.	$Aro3^{K222L}$, $Aro4^{K229L}$ $Aro7^{T226I}$, $Aro7^{G141S}$ AroL*	(94) (39) (91) (92) (109)
	Erythrose-4-phosphate	Introducing heterologous pathway. Overexpression/deletion of key enzymes.	Xfpk* Pta* Tkl1 Zwf1	(146) (42) (147) (147)
	PEP	Lowering expression level.	Pyk1	(148)
	Malonyl-CoA	Relief of allosteric regulation. Overexpression of key enzymes.	$Acc1^{S659A}$ $Acc1^{S1157A}$ $Acc1^{S686A}$ $Se Acs^{L641P}$*	(149) (149) (150) (151)
Phenyl Propanoid Pathway	Substrate promiscuity	Screening of heterologous genes with limited substrate specificity.	PAL*/TAL* CHS*	(152) (153)
	Poor catalytic capacity	Overexpression of enzymes. Screening of heterologous genes with higher catalytic capacity.	CHS*/C4H* 4CL*	(132) (154)
	By-product formation	Deletion of competing pathways. Implementation of heterologous enzymes without moonlighting functions.	Pdc5/6, Aro10 Tsc13	(132) (35)

Improving the precursory supply for enhancing phenylpropanoid biosynthesis

Efficient production of phenylpropanoids depends on appropriate intracellular concentrations of the precursors phenylalanine, tyrosine and malonyl-CoA. In wild-type yeast cells, regulation mechanisms have evolved to tune provision of these intermediates to requirements for growth. Several approaches have been tested in order to increase the pool of these precursors to improve phenylpropanoid production.

Improving the pool of aromatic precursors.

The aromatic amino acids phenylalanine and tyrosine are important precursors for phenylpropanoid biosynthesis. Alleviating (part of) the feedback inhibition of shikimate pathway enzymes has successfully enhanced phenylpropanoid production in several studies. For example, implementation of the feedback resistant allele $ARO4^{G226S}$ resulted in increased titers of flavonoids (62, 132, 141). The feedback resistant allele $ARO4^{K229L}$ (39) has

been successfully implemented for coumaric acid (134, 155), stilbenoids (135, 137, 155), flavonoids, (98, 142, 156, 157) and taxifolin or coniferyl alcohol (106). Different feedback resistant alleles of the chorismate mutase have also been implemented in *S. cerevisiae,* either *ARO7*^{G141S} for flavonoid (98, 156), coumaric acid (134) and stilbenoid (135) production or *ARO7*^{G229S} for flavonoid (142), taxifolin and coniferyl alcohol (106) production. The resistant alleles *ARO4*^{Q166K} and *ARO7*^{T226I} from *S. cerevisiae* have also been successfully used to improve opioid production (33). Furthermore, feedback resistant alleles of *Y. lipolytica* have also been identified (*YlARO4*^{K221L} and *YlARO7*^{G139S}) and implemented to improve naringenin (133) or resveratrol production (136). Finally, implementation of (part of) the *E. coli* shikimate pathway in *S. cerevisiae* led to increased production of *p*-coumaric acid (109, 134, 155).

E4P and PEP are the two precursors feeding the shikimate pathway (**Fig. 2**). Suástegui *et al.* studied the E4P and PEP fluxes in *S. cerevisiae* using metabolic flux analysis and concluded that E4P was a limiting precursor for synthesis of aromatic compounds (158). Although PEP and E4P are both substrates for the shikimate pathway, no modifications to increase the pools of these two compounds have been introduced in flavonoid producing strains so far. However, the overexpression of *YlTKT* or *ZWF1* were shown to increase the E4P supply resulting in increased 2-phenylethanol (155) or coniferyl alcohol (142) production respectively. PEP pools were improved by adapting the promoter strength of *PYK1, PFK1 and PFK2* resulting in increased coumaric acid titers (134). Deleting *YlPYK1* also resulted in increased coumaric acid and resveratrol production (155).

Bergman *et al.* identified heterologous phosphoketolases that convert fructose-6-phosphate (F6P) into E4P and acetyl-phosphate, but which also converted xylose-5-phosphate (X5P) in glycerol-3-phosphate (G3P) and acetyl-phosphate. Preference for F6P or X5P differed for phosphoketolase orthologs (146). In *S. cerevisiae*, acetyl-phosphate was shown to be hydrolyzed by Gpp1 and Gpp2, yielding acetate as a product. To reduce acetate accumulation, *GGP1* was deleted (146). Acetyl-phosphate can be converted into the metabolic precursor cytosolic acetyl-CoA with expression of a phosphotransacetylase (42). Improved supply of E4P as well as cytosolic acetyl-CoA seems attractive for flavonoid production as E4P can be used as a substrate for the shikimate pathway while cytosolic acetyl-CoA is the precursor for malonyl-CoA production.

Improving the malonyl-CoA pool

To produce one flavonoid molecule, three molecules of malonyl-CoA are required, which are generated by carboxylation of acetyl-CoA in a reaction catalyzed by the cytosolic, biotin-dependent acetyl-CoA carboxylase Acc1. When combined with the expression of a feedback resistant allele, overexpression of Acc1 resulted in increased production of naringenin (159) and kaempferol (157) by engineered S. cerevisiae strains. Similarly, overexpression of *YlACC1*

led to increased eriodictyol, taxifolin and naringenin titers in *Y. lipolytica* (140, 160). Acc1 harbors three serine residues that can be phosphorylated by protein kinases, thereby repressing Acc1 activity. Substitution of these serine residues by alanine eliminates this posttranslational regulation and results in increased intracellular malonyl-CoA levels (149, 150). This strategy for increasing malonyl-CoA supply has been successfully applied to increase stilbenoid titers (134, 135, 161).

Expression of an acetyl-CoA synthase allele from *Salmonella enterica* (*Seacs*^{L641P}) insensitive to post-translational modification was shown to increase stilbenoid production (135) and breviscapine production (143) by engineered *S. cerevisiae* strains. In the latter study, malonyl-CoA pools were further increased by deleting *MLS1* and *CIT2* to eliminate competition for acetyl-CoA by cytosolic malate synthase and extra mitochondrial citrate synthase, respectively. Additional overexpression of *ALD6* and *ADH2* to enhance malonyl-CoA production even further contributed to a final titer of 0.63 mM breviscapine (143). The relevance of engineering the malonyl-CoA supply is further demonstrated by research on *Y. lipolytica.* Combined overexpression of *YlACS2, YlARO1* and *YlACC1* in *Y. lipolytica* improved production naringenin, taxifolin and eriodictyol by engineered strains (140).

Suboptimal substrate specificity of enzymes of the phenylpropanoid biosynthesis pathway.

Several key enzymes in the phenylpropanoid pathway show a broad substrate specificity. This problem already appears in the first reaction of the pathway, in which phenylalanine and tyrosine are converted into cinnamic acid or coumaric acid, respectively, by phenylalanine and tyrosine ammonia lyases (PAL and TAL). Some of these enzymes use both aromatic amino acids as substrates and are therefore called phenylalanine/tyrosine ammonia lyases (PAL/TAL; EC 4.3.1.25). Depending on the target phenylpropanoid product, a suitable ammonia lyase should be used (152).

During flavonoid biosynthesis, issues with substrate specificity are not limited to the ammonia lyase reaction (**Fig. 6**). In *S. cerevisiae* engineered for quercetin production, kaempferol was detected as a by-product (162). This problem was attributed to *Solanum tuberosum* (potato) flavonol synthase which also acted on dihydroquercetin, an intermediate in the biosynthetic pathway towards quercetin, thus yielding kaempferol. Similarly, anthocyanidin synthase (ACS), which converts flavan-3,4-diols into anthocyanidin, also converts these flavan-3,4-diols into flavonols and, moreover, converts flavanonols into flavonols (61). Yet another substrate-specificity challenge has been described for 4-coumarate-CoA ligase (4CL) which, in addition to coumaric acid, also activates cinnamic acid and caffeic acid (106, 163). Finally, the chalcone synthase which normally converts three molecules of malonyl-CoA and a CoA molecule (such as coumaroyl-CoA) can also condense two malonyl-CoA molecules resulting

in byproduct formation (164). Improving the specificity of key reactions, either by exploration of natural biodiversity or by protein engineering, is therefore an important objective in metabolic engineering of the production of aromatic plant molecules in yeasts.

Suboptimal kinetics of key enzymes of the phenylpropanoid biosynthesis pathway.

In addition to substrate promiscuity, the catalytic capacity of key enzymes can represent challenges in engineering yeast for production of flavonoids. The chalcone synthase is notorious for its low catalytic activity (165) with turnover rates as low as 0.017 s^{-1} (153). Suboptimal CHS activity was identified as an important bottleneck for naringenin production in engineered *S. cerevisiae* strains in which productivity was not constrained by availability of coumaric acid (156). High-level expression of three copies of the *CHS3* of *Arabidopsis thaliana* resulted in increased naringenin production by *S. cerevisiae* compared to a strain carrying a single *AtCHS* expression cassette (132). Accumulation of coumaric acid in a naringenin-producing strain further supported the notion that *in vivo* activity of 4CL and/or CHS limited production rates (35). Also, in the yeast *Y. lipolytica,* the chalcone synthase was implicated as a key factor in determining naringenin productivity. Expressing up to five chalcone synthase expression cassettes showed an incremental increase in flavonoid titers (140) while another recent study using *Y. lipolytica* used two copies (133). Overexpressing the entire phenylpropanoid pathway has also shown to be successful for the increased production of stilbenoids (135), taxifolin and coniferyl alcohol (142).

Elimination of by-product formation during phenylpropanoid biosynthesis.

Phenylalanine and tyrosine are precursors for the phenylpropanoid pathway but can also be degraded in the Ehrlich pathway resulting in fusel alcohols. In *S. cerevisiae*, four thiamine-pyrophosphate-dependent 2-oxo-acid decarboxylases can catalyse decarboxylation of phenylpyruvate and *p*-hydroxyphenylpyruvate: Pdc1, Pdc5, Pdc6 and Aro10 (76, 82). Of these four, Pdc1 and Pdc6 have the highest affinity for pyruvate (82). However, Pdc6 is normally only expressed under sulfur limitation (83). Aro10 has a broad substrate specificity, which however does not include pyruvate, and has the highest affinity for phenylpyruvate and *p*-hydroxyphenylpyruvate, followed by Pdc5 (82). Since deletion of all three pyruvate decarboxylase genes renders *S. cerevisiae* unable to grow on glucose in batch cultures (166-169) metabolic engineering strategies to decrease byproduct formation at the phenylpyruvate branch point mostly focus on deletion of *ARO10* and *PDC5*, sometimes along with *PDC6*. Deletion of these main phenylpyruvate decarboxylases in engineered *S. cerevisiae* strain has been shown to enable improved production of coumaric acid (134), stilbenoids (135),

naringenin (35, 132, 159) several other flavanones (98), pelargonidin (62) 8-prenyl naringenin (141) and kaempferol (159) by reducing, but not eliminating, the formation of aromatic fusel alcohols.

During production of naringenin by engineered *S. cerevisiae* strains, phloretic acid is formed as a by-product by reduction of *p*-coumaroyl-CoA. In a study by lehka *et al.*,(35) *S. cerevisiae* NAD(P)H-dependent reductases and oxidoreductases were screened to identify the enzyme responsible for this undesirable byproduct formation. The results showed that the enoyl reductase Tsc13 catalysed reduction of *p*-coumaroyl-CoA to phloretic acid. Because Tsc13 is essential for fatty-acid biosynthesis, it was replaced with a long-chain enoyl-CoA reductase from *Malus domestica (MdECR*). This modification virtually eliminated phloretic acid production and led to improved naringenin production (35). This metabolic engineering strategy subsequently also eliminated phloretic acid production by anthocyanin-producing and an 8-prenyl naringenin-producing producing *S. cerevisiae* strains (62, 141).

Scope of this thesis

Phenylpyruvate produced by the shikimate pathway is a key metabolite for synthesis of aromatic compounds by *S. cerevisiae*. It serves as precursor for L-phenylalanine but can also be decarboxylated to phenylacetaldehyde, whose subsequent reduction yields 2-phenylethanol. This fusel alcohol is a valuable compound applied in the flavor and fragrance industry. However, when yeast strains are designed and constructed for production of for example phenylpropanoids, fusel alcohols constitute undesirable by-products. Under standard cultivation conditions, wild-type *S. cerevisiae* strains produce little to no 2-phenylethanol, but bypassing the shikimate pathway by using phenylalanine as nitrogen source or deregulation of its key enzymes can lead to higher concentrations. The goal of the research described in the present thesis was to develop 'chassis' strains of *S. cerevisiae* for the production of aromatic compounds in which *in vivo* activity of the shikimate pathway was maximized, initially using 2-phenylethanol as a model product. Next, the production of by-products synthesized via the Ehrlich pathway was minimized.

In **Chapter 2** different metabolic engineering strategies were explored to increase supply of phenylpyruvate through the shikimate pathway by using 2-phenylethanol as a model product. Feedback insensitive alleles of crucial enzymes in the shikimate pathway were overexpressed in order to alleviate feedback inhibition. Subsequently, several strategies were tested for improving supply of the two precursors to the shikimate pathway: phospho-enol-pyruvate and erythrose-4-phosphate. The biosynthetic pathways for phenylpyruvate and *p*-hydroxyphenylpyruvate are identical up until the production of prephenate. Therefore, any improvements made upstream of prephenate will automatically also increase production of *p*-hydroxyphenylpyruvate and the derived fusel alcohol *p*-hydroxyphenylethanol.

Deleting the gene coding for the enzyme responsible for the conversion of prephenate into *p*-hydroxyphenylpyruvate would result in tyrosine auxotrophy, which is undesirable for industrial purposes. Therefore, strategies were explored to reduce *p*-hydroxyphenylpyruvate availability while retaining tyrosine prototrophy. Genetic modifications aimed at improving phenylpyruvate synthesis were combined with engineering of the Ehrlich pathway to redirect flux towards phenylethanol.

Chapter 3 explores production of flavonoids by yeast strains carrying a synthetic chromosome (NeoChr) with either a linear or a circular configuration. These NeoChr were constructed to contain expression cassettes for all genes required for conversion of glucose to the production of the anthocyanine flavonoid pelargonidin 3-O-glucoside, starting with hexokinase 2 (Hxk2) and ending with a plant anthocyanin 3-O-glucosyltransferase (*3GT*). Several metabolic engineering strategies tested in Chapter 2 were implemented in the NeoChr design to improve supply of aromatic precursors. In addition to analyzing the production of aromatic compounds, the impact of a linear or circular configuration of the NeoChr on yeast physiology was tested. Finally, the effect of integrating additional copies of the expressing cassette for chalcone synthase, a proposed rate-limiting enzyme in the flavonoid pathway, was evaluated.

The fusel alcohols 2-phenylethanol and *p*-hydroxyphenylethanol are well-known byproducts of phenylpropanoid-producing *S. cerevisiae* strains. The amount of fusel alcohols produced can be limited by deleting *ARO10, PDC5* and/or *PDC6* but additional deletion of *PDC1* would result in a strain incapable of growing on glucose in batch cultures. Therefore, significant amounts of fusel alcohols are invariably still found in *S. cerevisiae* strains engineered for production of flavonoids, stilbenoids, hydroxycinnamic acids and other aromatic products. In order to completely eliminate fusel alcohol production while retaining the ability of yeast to grow on glucose as sole carbon source, pyruvate decarboxylases from Prokaryotic and Eukaryotic origin were first screened for their substrate specificity in **Chapter 4.** Subsequently, enzymes that showed a high specificity for pyruvate were implemented in a coumaric-acid producing *S. cerevisiae* strain devoid of all native 2-oxo-acid decarboxylase in order to study whether fusel alcohol production could be completely eliminated.

Chapter 2

Connecting central carbon and aromatic amino acid metabolisms to improve *de novo* 2-phenylethanol production in *Saccharomyces cerevisiae*

Else-Jasmijn Hassing, Philip A. de Groot, Vita R. Marquenie, Jack T. Pronk, Jean-Marc G Daran

Essentially as published in
Metabolic Engineering
2019. 56, 165-180
Supplementary materials available online:
https://doi.org/10.1016/j.ymben.2019.09.011

The organic compound 2-phenylethanol (2PE) has a pleasant floral scent and is intensively used in the cosmetic and food industries. Microbial production of 2PE by phenylalanine bioconversion or *de novo* biosynthesis from sugar offer sustainable, reliable and natural production processes compared to chemical synthesis. Despite the ability of *S. cerevisiae* to naturally synthesize 2PE, *de novo* synthesis in high concentration and yield remains a metabolic engineering challenge. Here we demonstrate that improving phosphoenolpyruvate supply by expressing pyruvate kinase variants and eliminating the formation of *p*-hydroxyphenylethanol without creating tyrosine auxotrophy significantly contributed to improve 2PE production in *S. cerevisiae*. In combination with the engineering of the aromatic amino acid biosynthesis and Ehrlich pathway, these mutations enabled better connection between glycolysis and pentose phosphate pathway, thereby optimizing the carbon flux towards 2PE. However, an attempt to further connect these two parts of central carbon metabolism by redirecting fructose-6P towards erythrose-4P by expressing a phosphoketolase-phosphotransacetylase pathway did not result in better strains. The best performing strains were capable of producing 13mM of 2PE at a yield of 0.113 mol mol^{-1} which represents the highest yield for de novo produced 2PE in *S. cerevisiae* and other yeast species.

Abbreviations: FBR Feedback resistant allele, GLC glucose, G6P glucose-6-phosphate, F6P fructose-6-phosphate, ACP acetyl-P, ACE acetate, ACCOA acetyl-coA, TCA tricarboxylic acid cycle, PYR pyruvate, PEP phosphoenolpyruvate, E4P erythose-4-phosphate, DAHP 3-deoxy-D-arabino-heptulosonate-7-phosphate, CHR chorismate, SHK shikimate, PPA prephenate, PPY phenylpyruvate, PHE L-phenylalanine, PAC phenylacetaldehyde, PAA phenylacetate, 2PE 2-phenylethanol, *p*OHPPY *p*-hydroxyphenylpyruvate, TYR L-tyrosine, TRP tryptophan, *p*OHPAC p-hydroxyphenylacetaldehyde, *p*OHPAA *p*-hydroxyphenylacetate, *p*OH2PE *p*-hydroxyphenylethanol.

Introduction

Flavouring and fragrance compounds are used in a wide range of application sectors including food, beverage, cosmetic and perfume industries. The organic compound 2-phenylethanol (2PE) is characterized by a pleasant, fresh, rose and green phenolic muguet scent. Additionally, 2PE serves as a building block for other products such as phenylethyl acetate, a flavouring agent with rose, honey and raspberry notes (64).

Although the most cost effective 2PE production processes are based on chemical synthesis these methods are gradually replaced by biological production processes. The absence of substrate stereo-selectivity of the 2PE chemical synthesis processes leads to formation of by-products that require use of non-environmentally friendly purification steps. Additionally, the European and US regulations restricted food grade 2PE to natural sources which in this case includes botanical and microbiological sources, comprising fermentation products (103). The first 2PE natural source derives from plant and flower essential oils extraction, mostly rose petals (170). However, this supply is limited and impeded by costs of purification (171-173). The second 2PE natural source is the product of microbial metabolism. 2PE is the product of the Ehrlich pathway (75). Several amino acids (branched-chain and aromatic amino acids as well as methionine) can be assimilated by the Ehrlich pathway and converted in higher fusel alcohols. This pathway consists of three reactions, i) a transamination of the amino acid, ii) a decarboxylation of the 2-oxo acid formed in the preceding reaction and iii) a reduction of the resulting fusel aldehyde into the higher fusel alcohol (76). The yeast *Saccharomyces cerevisiae,* traditionally used in wine, beer and bread fermentation naturally produces, next to ethanol and CO_2, a range of higher alcohols and esters which have a strong impact on the sensory properties and quality of the products (143, 174). This characteristic has been exploited to produce 2PE from phenylalanine by whole-cell bioconversion (175-178). However, biotransformation often relies on a multi-stages approach, adding-up operation units and use of a substrate precursor that on the long term might limit the attractiveness of this method. *De novo* 2PE production using a low cost, renewable sugar such as glucose could represent a promising alternative to phenylalanine bioconversion.

In *S. cerevisiae,* the Ehrlich pathway precursor phenylpyruvate is directly derived from the shikimate pathway. This pathway starts with the condensation of the glycolytic intermediate phosphoenolpyruvate (PEP) and the pentose phosphate pathway intermediate erythrose 4-phospate (E4P) to form the first dedicated metabolite of the shikimate pathway 3-deoxy-*D*-arabino-heptulosonate-7-phosphate (DAHP). DAHP is then converted to chorismate in six consecutive steps catalysed by the enzymes Aro1-Aro4 and Aro7 (87-89, 179). Chorismate is the last common intermediate of all three aromatic amino acids. There, the pathway divides in two branches, one towards tryptophan (TRP) and the other towards phenylalanine (PHE) and

tyrosine (TYR). The latter branch is catalysed by the chorismate mutase Aro7. The product of this reaction, prephenate (PPA), is then either directed to the 2-oxo acid phenylpyruvate (PPY) or to *p*-hydroxyphenylpyruvate (*p*OHPPY) catalysed by the prephenate dehydratase Pha2 or the dehydrogenases Tyr1, respectively. The 2-oxo acids can then be transaminated to phenylalanine or tyrosine by the aromatic amino acid transferases I (Aro8) or II (Aro9) depending on the nitrogen source (77, 78). Alternatively, the resulting 2-oxo acid can get decarboxylated by a phenylpyruvate decarboxylase into a fusel aldehyde (80, 81). In *S. cerevisiae* this step is mainly catalysed by the thiamine pyrophosphate (TPP)-dependent 2-oxo acid decarboxylase, Aro10 (82, 83). Finally, fusel aldehydes can either be reduced to fusel alcohols (180) or oxidized to a fusel acids depending on the redox status of the cells (81) (**Fig. 1)**

Hitherto, pathway engineering strategies for *de novo* production of 2PE essentially targeted the shikimate and the Ehrlich pathways (181). In the case of aromatic amino acids, the DAHP synthases Aro3 and Aro4, are feedback inhibited by phenylalanine and tyrosine, respectively (39, 89) and the chorismate mutase Aro7 is allosterically controlled by tyrosine and tryptophan that act as an inhibitor or activator, respectively (91-93). Expression of a DAHP synthase allele encoding a feedback insensitive variant (*ARO4*^{K229L}) in *S. cerevisiae* was shown to be sufficient to yield *de novo* synthesis of aromatic amino acid derived Ehrlich metabolites (39). Co-expression of alleles encoding both Aro4^{K229L} and Aro3^{K222L} feedback insensitive variants further improved production of metabolites derived from the shikimate pathway (94). In contrast to the DAHP synthase, the sole alleviation of the feedback inhibition of Aro7 showed no impact on the formation of Ehrlich metabolites. However, the combined expression of these variant alleles promoted a significant 4-fold increase of the flux through the aromatic amino acid pathway yielding a 200-fold increase of extracellular concentration of the aromatic Ehrlich metabolites and improved fusel alcohol production (39). Although previously applied engineering strategies significantly increased 2PE biosynthesis (0.41 g L^{-1}, yield of 0.03 mol 2PE per mol glucose consumed), the product yields achieved are still one order of magnitude lower than the calculated glucose maximum theoretical yield (0.500 mol mol^{-1}).

The aim of this study was to define a metabolic engineering strategy to improve *de novo* production of 2PE by *S. cerevisiae*, using glucose as sole carbon source. To do so, central carbon metabolism was modified to optimize the supply of E4P and PEP to the shikimate pathway. This was combined with the overexpression, and mutations alleviating allosteric regulation, of genes of the shikimate and Ehrlich pathways. In addition, the flux towards by-products was decreased by down-tuning transcriptional regulation of competing metabolic branches (**Fig. 1**).

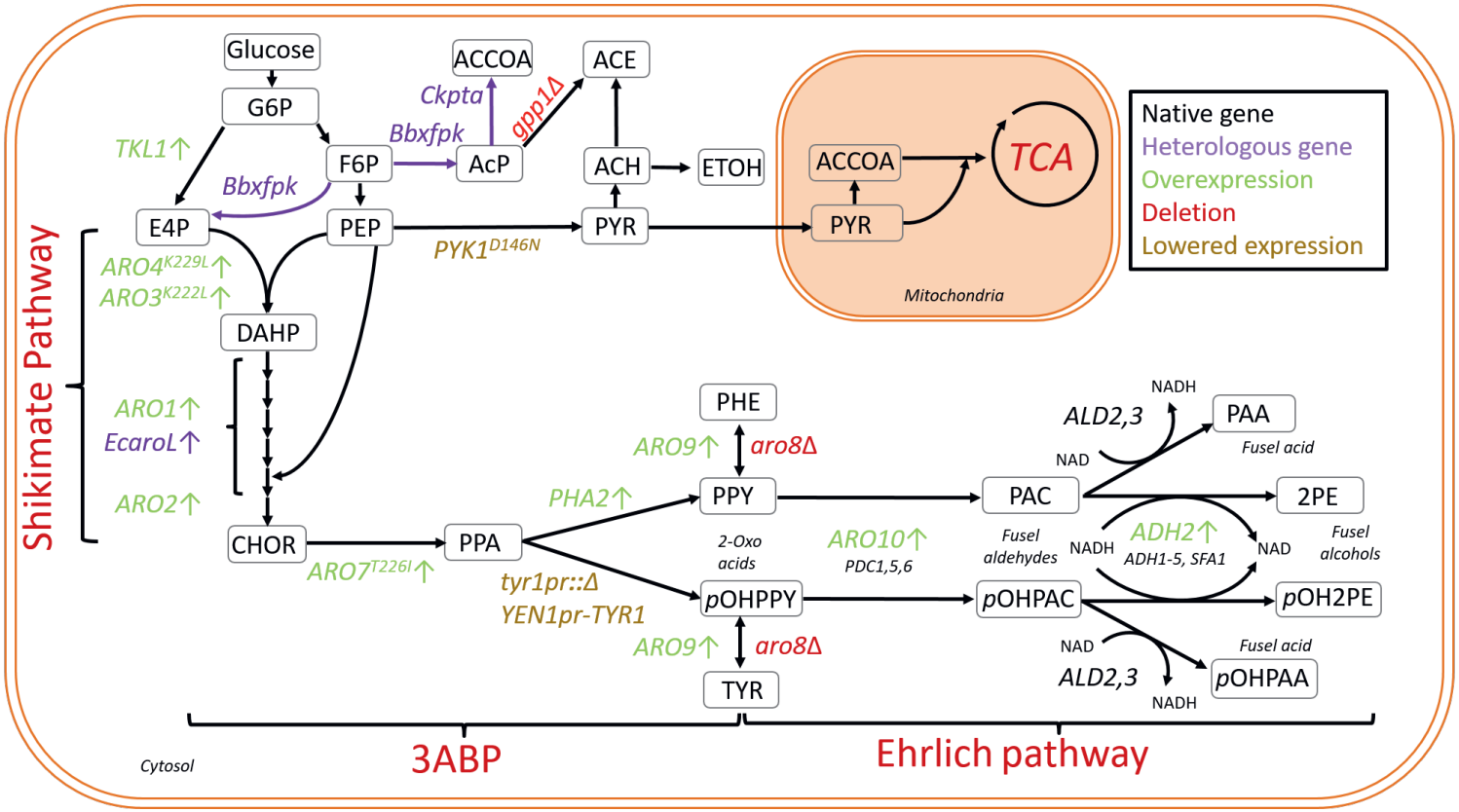


Figure 1. Adapted metabolic pathway for increased 2PE production in *S. cerevisiae*. Native enzymatic pathways are indicated with black arrows. Heterologous enzymatic pathways are indicated in purple. Overexpressed genes are highlighted in green, deleted genes are annotated in red, genes with lowered expression are shown in brown and heterologous enzymes are depicted in purple. FBR Feedback resistant allele, glucose, G6P glucose-6-phosphate, F6P fructose-6-phosphate, ACP acetyl-P, ACE acetate, ACCOA acetyl-coA, TCA tricarboxylic acid cycle, PYR pyruvate, PEP phosphoenolpyruvate, E4P erythose-4-phosphate, DAHP 3-deoxy-D-arabino-heptulosonate-7-phosphate, CHR chorismate, PPA prephenate, PPY phenylpyruvate, PHE L-phenylalanine, PAC phenylacetaldehyde, PAA phenylacetate, 2PE 2-phenylethanol, *p*OHPPY *p*-hydroxyphenylpyruvate, TYR L-tyrosine, *p*OHPAC *p*-hydroxyphenylacetaldehyde, *p*OHPAA *p*-hydroxyphenylacetate, *p*OH2PE *p*-hydroxyphenylethanol.

Materials and Methods

Strains and culture conditions

All yeast strains used in this study belong to the CEN.PK lineage (182)(**Table 1**). Strains were grown in 500 ml shake flasks containing 100 ml complex medium (YPD) with 10 g L^{-1} Bacto Peptone, 20 g L^{-1} glucose and 200 mg L^{-1} hygromycin or 100 mg L^{-1} nourseothricin when required. Alternatively, strains were grown in chemically defined (synthetic) medium (SM) containing 3 g L^{-1} KH_2PO4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 5 g L^{-1} $(NH_4)_2SO_4$, 1 mL L^{-1} trace element solution and 1 mL L^{-1} vitamin solution (183) containing 20 g L^{-1} glucose with 150 mg L^{-1} uracil, 76 mg L^{-1} phenylalanine and 225 mg L^{-1} tyrosine when required (184). Synthetic medium with urea as sole nitrogen source consisted of 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 5 g L^{-1} K_2SO_4, 2.3 g L^{-1} filter sterilized urea, 1 mL L^{-1} trace element solution and 1 mL L^{-1} vitamin solution containing 20 g L^{-1} glucose and 200 mg L^{-1} hygromycin when required (185). All strains were grown at 30˚C at 200 RPM. Solid medium was obtained by the addition of 2% (w/v) bacto peptone prior to heat sterilization.

Escherichia coli XL1 blue cells (Agilent Technologies, Santa Clara, CA) were used for plasmids storage and propagation. *E. coli* cells were grown in lysogeny broth (LB) with supplementation of 100 mg L^{-1} ampicillin, 25 mg L^{-1} chloramphenicol or 50 mg L^{-1} kanamycin. Solid medium was obtained by the addition of 2% (w/v) bacto peptone prior to heat sterilization.

Yeast and *E. coli* cultures were stored by adding 30% (v/v) glycerol to exponentially growing cells and storing the aliquots at -80 ˚C.

Table 1. Strains used in this study. *Bb Bifidobacterium breve, Ck Clostridium kluyveri, Ec Escherichia coli, Kl Kluyveromyces lactis* and *Sp Streptococcus pyogenes*

Strain	Relevant genotype	Reference
CEN.PK 113-7D	*MATa URA3 HIS3 LEU2 TRP1 MAL2-8c SUC2*	(182)
S. kudriavzevii CR85	*MATa/Matα*	(186)
S. eubayanus CBS12357	*MATa/Matα*	(187)
IMX581	*ura3-52 can1Δ::Spcas9-natNT2*	(188)
IME324	*ura3-52 can1Δ::Spcas9-natNT2* p426-TEF *(URA3)*	(189)
IMN002	*ura3-52 aro3Δ ARO4*$_p$*-ARO4Δ::TDH3*$_p$*-ARO4*^{K229L}	(39)
IMX1492	*ura3-52 s*$_p$*3Δ::Spcas9-natNT2 aro3Δ TDH3*$_p$*-ARO4*^{K229L}	This study
IMX1533	*ura3-52 s*$_p$*3Δ::Spcas9-natNT2 aro3Δ aro7Δ TDH3*$_p$*-ARO4*^{K229L}	This study
IMX1586	*ura3-52 spr3Δ::Spcas9-natNT2 aro3Δ aro7Δ TDH3*$_p$*-ARO4*^{K229L} *GPM1*$_p$*-ARO7*^{T226I}*-TEFt*$_t$ pUDR649 (*KlURA3,* gRNA-*ARO7*$_p$*)*	This study
IMX1593	*ura3-52 spr3Δ::Spcas9-natNT2 aro3Δ aro7Δ TDH3*$_p$*-ARO4*^{K229L} *ARO7*$_p$*::SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t*	This study
IMX1754	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ* pUDR409 (*KlURA3,* gRNA-*ARO3*)	This study
IMX1783	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ*	This study
IMX1952	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2*$_t$ shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1pt-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1*	This study
IMX1955	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t* pUDR372 (*klURA3,* gRNA-*SGA1*)	This study
IMX2029	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t*	This study
IMX2051	*ura3-52 can1Δ::Spcas9-natNT2 tyr1Δ*	This study
IMX2052	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ tyr1Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t*	This study
IMX2056	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t*	This study
IMC124	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3*$_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t* pUDC245 (*CEN6/ARS4, URA3, SkPDC1*$_p$*-EcaroG*^{P150L}*-ScPFK1t*)	This study

Table 1 (continued)

IMC125	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC246 (*CEN6/ARS4, URA3, SkADH1$_p$-EcaroB-ScFBA1t*)	This study
IMC126	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC247 (*CEN6/ARS4, URA3, SkFBA1$_p$-EcaroD-TDH3t*)	This study
IMC127	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC248 (*CEN6/ARS4, URA3, TDH2$_p$-EcaroE-PDC1t*)	This study
IMC128	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC249 (*CEN6/ARS4, URA3, SePDC1$_p$-EcaroL-ADH3t*)	This study
IMC129	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC250 (*CEN6/ARS4, URA3, SkFBA1$_p$-EcaroA-PGI1t*)	This study
IMC130	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC251 (*CEN6/ARS4, URA3, TDH3$_p$-EcaroC-PFK2t*)	This study
IMC131	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t^p* pUDC252 (*CEN6/ARS4, URA3, ENO2$_p$-EcpheA^{T326P}-GPM1t*)	This study
IMC132	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t* pGGKd019 (*CEN6/ARS4, URA3*)	This study
IMX2068	*ura3-52 can1Δ::Spcas9-natNT2 tyr1$_p$Δ::YEN1$_p$-TYR1*	This study
IMX2069	*ura3-52 can1Δ::Spcas9-natNT2 tyr1$_p$Δ::AGE1$_p$-TYR1*	This study
IMX2071	*ura3-52 can1Δ::Spcas9-natNT2 tyr1$_p$Δ::SEC18$_p$-TYR1*	This study
IMX2072	*ura3-52 can1Δ::Spcas9-natNT2 tyr1$_p$Δ::MRI$_p$-TYR1*	This study
IMX2073	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$::YEN1$_p$-TYR1*	This study

Table 1 (continued)

IMX2074	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::AGE1$_p$-TYR1*	This study
IMX2076	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::SEC18$_p$-TYR1*	This study
IMX2077	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ ARO1$_p$::TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::MRI$_p$-TYR1*	This study
IMX2102	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t PYK1^{A336S} pUDR574 (KlURA3,* gRNA*-PYK1.2)*	This study
IMX2106	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t PYK1^{DA336S}*	This study
IMX2107	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ ARO1$_p$::TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t PYK1^{D146N} pUDR577 (KlURA3,* gRNA*-PYK1.1)*	This study
IMX2108	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t PYK1^{D146N}*	This study
IMX2123	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::YEN1$_p$-TYR1 PYK1^{DA336S}*	This study
IMX2124	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::YEN1$_p$-TYR1 PYK1^{D146N}*	This study
IME471	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ TDH3$_p$-ARO4^{K229L}-ENO2t* shrDA *SkTDH3$_p$-PHA2-TEF2t* shrDB *SePDC1$_p$-ARO2-SSA1t* shrDC *SeFBA1$_p$-ARO9-ADH1t* shrDD *SeGPM1$_p$-ARO7^{T226I}-TEF1t* shrDE *ENO2$_p$-ARO1 sga1Δ::ENO2$_p$-TKL1-TDH3t tyr1$_p$Δ::YEN1$_p$-TYR1 PYK1^{D146N}* pUDE001 *(2μm URA3, TDH3$_p$-ARO10-CYC1t)*	This study

Table 1 (continued)		
IMX2179	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ* $TDH3_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t tyr1*$_p$*::YEN1*$_p$*-TYR1 X3::FBA1*$_p$*-ARO3*^{K222L}*-PGKIt* shrAF *PDC1*$_p$*-EcaroL-ADH3t PYK1*^{D146N} pUDE001 *(2μm URA3, TDH3*$_p$*-ARO10-CYC1t)*	This study
IMX2222	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ* $TDH3_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t tyr1*$_p$*Δ::YEN1*$_p$*-TYR1 X3::FBA1*$_p$*-ARO3*^{K222L}*-PGKIt* shrAF *PDC1*$_p$*-EcaroL-ADH3t gpp1Δ::PGK1p-Bbxfpk-ENO1t shrBA PGI1p -Ckpta-TDH1t PYK1*^{D146N} pUDE001 *(2μm URA3, TDH3*$_p$*-ARO10-CYC1t)*	This study
IMX2279	*ura3-52 can1Δ::Spcas9-natNT2 aro3Δ aro8Δ* $TDH3_p$*-ARO4*^{K229L}*-ENO2t* shrDA *SkTDH3*$_p$*-PHA2-TEF2t* shrDB *SePDC1*$_p$*-ARO2-SSA1t* shrDC *SeFBA1*$_p$*-ARO9-ADH1t* shrDD *SeGPM1*$_p$*-ARO7*^{T226I}*-TEF1t* shrDE *ENO2*$_p$*-ARO1 sga1Δ::ENO2*$_p$*-TKL1-TDH3t tyr1*$_p$*Δ::YEN1*$_p$*Δ-TYR1 X3::FBA1*$_p$*-ARO3*^{K222L}*-PGKIt* shrAF *PDC1*$_p$*-EcaroL-ADH3t X2::PYK1*$_p$*-ADH2-PDC1t PYK1*^{D146N} pUDE001 *(2μm URA3, TDH3*$_p$*-ARO10-CYC1t)*	This study

Molecular Biology Techniques

For cloning purposes, Phusion high fidelity polymerase (Thermo Scientific, Landsmeer, Netherlands) was used according to manufacturer's recommendations with the exception that the primer and polymerase concentrations were set at 200 nM and 0.03 μL^{-1} respectively. Genomic DNA used as template for PCR amplification was isolated using the YeaStar genomic DNA kit (Zymo Research, Irvine, CA) according to manufacturer's instructions. PCR products were purified by gel purification using Zymoclean kit (Zymo Research) according to manufacturer's recommendations using milliQ as eluent solvent. Alternatively, PCR products were digested for an hour using 1 μl DpnI FastDigest enzyme (Thermo Fisher Scientific). The digested products were then purified using the GenElute™ PCR clean-Up Kit (Sigma-Aldrich, St. Louis, MO). Diagnostic PCR was performed using DreamTaq PCR mastermix (Thermo Fisher Scientific). A list of diagnostic primers is provided in **Supplementary Table 1.** Plasmids were isolated from *E. coli* using the GenElute plasmid miniprep kit (Sigma-Aldrich) according to manufacturer's instructions using miliQ water as eluent.

Plasmid construction

All constructed plasmids were transformed to *E. coli* (XL1-Blue) cells according to the manufacturer's recommendations and grown under selective conditions. Plasmids used in this study are given in **Supplementary Table 2**.

Contruction of Cas9 reprogramming gRNA plasmids

Plasmids containing either single or double gRNA for Cas9 targeting of specific loci were constructed as described by Mans *et al.* (188)

Construction of the expression cassettes.

The construction of promoter, gene or terminator part plasmids compatible with Golden Gate assembly following the yeast tool kit principle (190) were constructed by first amplifying the region of interest with primers providing part type specific overhangs and assembling these fragments in either pUD565 (191) or pUD564 (**Supplementary Table 2**), two GFP dropout entry vectors containing chloramphenicol or kanamycin resistance genes, respectively.

Genomic DNA from *S. cerevisiae* CEN.PK113-7D, *Saccharomyces, kudriavzevii* CR85 or *Saccharomyces eubayanus* CBS12357 was used as template DNA for amplifying the promoter regions. Primers providing promoter part type specific overhangs (AACG and CATA) (Supplementary Table 1) were used to amplify the following fragments: $PDC1_p$ (9755 & 9756) $ENO2_p$ (9739 & 9740), $PGI1_p$ (9630 & 9631), $PYK1_p$ (10608 & 10609), $THD3_p$ (10753 & 10754), $PGK1_p$ (9421 & 9422), $SkADH1_p$ (9737 & 9738), $SkTDH3_p$ (9751 & 9752), $SkPDC1_p$ (9731 & 9732), $SkFBA1_p$ (9640 & 9641), $SePDC1_p$ (9729 & 9730), $SeFBA1_p$ (9409 & 9410) and $SeGPM1_p$ (9759 & 9760). Additionally, $ENO2_p$ was also ordered synthesized using the GeneArt gene Synthesis (ThermoFisher Scientific) with upstream the promoter specific YeastToolkit flank 'AAGCATCGTCTCATCGGTCTCAAACG' and downstream with 'TTATGCCGTCTCAGGTCTCACATA' (190). The promoter fragments were cloned into entry vector pUD565 using BsmBI mediated golden gate assembly resulting in part plasmids pGGKp025 ($PDC1_p$), pGGKp028 ($ENO2_p$), pGGKp033 ($PGI1_p$), pGGKp034 ($PYK1_p$), pGGKp035 ($THD3_p$), pGGKp036 ($PGKI_p$), pGGKp062 ($SkADH1_p$) pGGKp063 ($SkTDH3_p$), pGGKp064 ($SkPDC1_p$), pGGKp065 ($SkFBA1_p$), pGGKp074 ($SePDC1_p$), pGGKp075 ($SeFBA1_p$), pGGKp095 ($SeGPM1_p$) and pGGKp164 ($ENO2_p$).

Terminator constructs were all obtained from CEN.PK113-7D genomic DNA, this time using terminator part type specific overhang (ATCC and CAGC) primers (Supplementary Table 1) obtaining fragments $ADH1_t$ (10769 & 10770), $TEF2_t$ (10884 & 10885), $TEF1_t$ (10767 & 10768), $PYK1_t$ (10886 & 10887), $TDH3_t$ (10761 & 10762), $PDC1_t$ (10773 & 10774) and $GPM1_t$ (10759 & 10760). Additionally, the $PGI1_t$, $PFK1_t$, $PFK2_t$, $FBA1_t$, $TDH3_t$, $GPM1_t$, $PDC1_t$ and $ADH3_t$ regions were synthesized (GeneArt, ThermoFisher Scientific). In this case, the terminators were flanked upstream with the terminator specific YeastToolkit flank 'AAGCATCGTCTCATCGGTCTCAATCC' and downstream with 'TTATGCCGTCTCAGGTCTCACAGC' (190). All terminator fragments were assembled into entry vector pUD565 resulting in part plasmids pGGKp037 ($ADH1_t$), pGGKp038

(*TEF2*$_t$), pGGKp039 (*TEF1t*), pGGKp040 (*PYK1*$_t$), pGGKp041 (*TDH3*$_t$), pGGKp045 (*PDC1*$_t$), pGGKp048 (*GPM1*$_t$), pGGKp099 (*PGI1*$_t$), pGGKp101 (*PFK1*$_t$), pGGKp103 (*PFK2*$_t$), pGGKp105 (*FBA1*$_t$), pGGKp106 (*TDH3*$_t$) pGGKp107 (*GPM1*$_t$), pGGKp110 (*PDC1*$_t$) and pGGKp113 (*ADH3*$_t$).

The ORFs were also obtained from genomic DNA of CEN.PK113-7D using primers with gene part type specific overhangs (TATG and GGAT) (**Supplementary Table 1**). Primer pairs 16497/16498, 11580/11581 and 13106/13107 were used to obtain fragments *ADH2, TKL1* and *PHA2*. Because the ORF of *TYR1, ARO9, ARO8* and *ARO2* contained internal BsaI and/or BsmbI sites, these restriction sites were removed from the coding sequence by amplifying the gene in multiple fragments using overlap primers that introduced silent mutations, thereby removing the internal restriction site and adding an additional BsmBI site, which would be lost after BsmBI assembly. For the 5' end of the reverse primer the overlap 5'-CACGTCTCANNNN(primer sequence)-3' was used where NNNN is the 4-base stretch for replacing the internal site. For the forward primer the overlap 5'-TTCGTCTCTNNNN(primer sequence)-3' was used, again using the NNNN stretch for removing the same site. Since *ARO2* contained one BsmBI site close to the start of the gene, only the BsaI site was removed from this gene. The *TYR1* ORF was amplified using primer pairs 13111/13109 and 13108/13110, *ARO9* with 13104/13103 and 13102/13105, *ARO8* with 13100/13099 and 13098/13101 and *ARO2* with 13096/13093 and 13092/13097. The *ARO7*^{T226I,} *ARO3*^{K222L}, *ARO4*^{K229L}, *EcaroG*^{P150L}, *EcaroB, EcaroD, EcaroE, EcaroL, EcaroA, EcaroC* and *EcpheA*^{T326P} genes were synthesized using the Invitrogen GeneArt gene synthesis service (ThermoFisher Scientific). The sequences for the *E. coli* genes were derived from the annotated genome of strain K-12 substr. MG1655 (Bioproject accession number PRJNA225; assembly GCA_000005845.2) and ordered codon optimised. Any internal BsmBI/BsaI sites were removed from the sequence using silent mutations and the ORFs were flanked upstream with the gene specific YeastToolkit flank 'AAGCATCGTCTCATCGGTCTCAT' and downstream with 'TTATGCCGTCTCAGGTCTCAGGAT'. The amplified and synthesised fragments were cloned into entry vector pUD565, via BsmBI Golden Gate assembly, resulting in part plasmids pGGKp118 (*ARO7*^{T226I}), pGGKp119 (*EcaroG*^{P150L}), pGGKp120 (*EcaroB*), pGGKp121 (*EcaroD*), pGGKp122 (*EcaroE*), pGGKp123 (*EcaroL*), pGGKp124 (*EcaroA*), pGGKp125 (*EcaroC*), pGGKp126 *(EcpheA*^{T326PL}), pGGKp134 (*ARO4*^{K229L}), pGGKp165 (*ARO8*), pGGKp166 (*ARO9*), pGGKp167 (*PHA2*), pGGKp168 (*TYR1*), pGGKp170 (*ARO2*), pGGKp246 (*ARO3*^{K222L}) and pGGKp294 (*ADH2*). The amplified *TKL1* fragment was assembled into entry vector pUD564 resulting in part plasmid pGGKp067 (*TKL1*).

Correct construction of part plasmids would result in the replacement of the GFP gene by the target gene which allows visual screening of the colonies. Additional plasmid confirmation was done by colony PCR using primers 2012 and 2397. Correct removal of internal BsaI/

BsmBI sites of *ARO2, TYR1, ARO8* and *ARO9* was verified by Sanger sequencing (BaseClear, Leiden, NL).

To construct the expression cassettes to engineer 2PE production in *S. cerevisiae* first recipient GFP dropout plasmids were constructed by BsmBI Golden Gate assembly using the yeast tool kit principle outlined by Lee et al. (190). The first one, PGGKd005 was obtained by assembling the part plasmids pYTK002 and pYTK067 (left and right connectors), pYTK047 (GFP dropout), pYTK079 (*hghR*), pYTK081 (*CEN6/ARS4* yeast replication origin) with pYTK083 (bacterial origin of replication factor and ampicillin resistance (*bla*) gene). pGGKd015 was constructed by golden gate assembly of pYTK002 and pYTK067 (left and right connectors respectively), pYTK047 (GFP dropout) and pYTK095 (bacterial origin of replication factor and ampicillin resistance gene). pGGKd019 was constructed by combining pYTK002 and pYTK073 (left and right connectors), pYTK047 (GFP dropout), pYTK074 (URA3), pYTK081 (CEN6/ARS4 yeast replication origin) with pYTK084 (bacterial origin of replication factor and kanamycin resistance gene). Finally, pGGKd046 was assembled using part plasmids pYTK003 and pYTK072 (left and right connectors), pYTK047 (GFP dropout), pYTK078 (*natR*), pYTK081 (*CEN6/ARS4* yeast replication origin) with pYTK083 (bacterial origin of replication factor and ampicillin resistance gene).

Next, a specific promoter, gene and terminator part plasmid were assembled into one of the GFP dropout plasmids via BsaI mediated Golden Gate assembly resulting in expression cassettes. As example, pGGKp028 ($ENO2_p$), pGGKp041 ($TDH3_t$) and pGGKp067 (*TKL1*) were assembled into pGGKd015 resulting in pUD660 ($ENO2_p$-*TKL1*-$TDH3_t$). A full overview of the constructed expression cassettes and the part plasmids used for the construction is given in **Supplementary Table 2.** Colony PCR for initial screening was performed on randomly picked colonies using primers 10320 and 10327 for expression cassettes with the pGGKd005 or pGGKd015 backbone and primers 10320 and 10335 for the pGGKd019 or pGGkd046 backbone. Final plasmid confirmation was done by restriction analysis.

Strain construction

Integration of *Spcas9* into IMN002 was performed by assembling and integrating the *Spycas9* and natNT2 marker cassettes into the *SPR3* locus. The integration cassette containing the *cas9* gene was amplified using p414-$TEF1_p$-*cas9*-$CYC1_t$ (192) as DNA template with primers 12036 and 4653, adding homologous flanks for integration into the *SPR3* locus and a short homologous sequence (shr) (193) efficient strategy for assembly of multi-fragment expression vectors in *S. cerevisiae* facilitating homologous recombination. The natNT2 integration cassette was amplified using primer pair 3093 and 12037 with pUG-natNT2 (194) as template using an initial denaturation of 1 min instead of the advised 30 seconds and 20 seconds denaturation instead of 10 seconds. The constructed strain was stocked as IMX1492. Correspondingly,

IMX1492 was co-transformed with the linearized backbone of pMEL10 (188, 195) the double-stranded cassette (primers 12048 & 12049) containing the gRNA for *ARO7* together with the double stranded repair oligo (primers 12050 & 12051) for knocking out *ARO7*. The resulting *aro7Δ* strain was stocked as IMX1533 after plasmid recycling. Next, the $ARO7^{T226I}$ cassette was PCR amplified from pUD714 using primers 12248 and 12249, adding flanks for integration in the *ARO7* promoter region. IMX1533 was transformed with the purified PCR product, pUDR649 (gRNA-$ARO7_p$) leading to strain IMX1586 and IMX1593 ($3ABP^{FBR}$) before and after plasmid recycling respectively.

IMX581 was transformed with pGGKd019 (*CEN6/ARS4, URA3*) resulting in IME324. Additionally, IMX581 was transformed with pUDR409 (gRNA-*ARO3*) and the double-stranded repair oligo (primer 13594 & 13595) for deleting *ARO3,* resulting in strain IMX1754 and IMX1783 before and after plasmid recycling.

Next, the integration cassettes for all 3ABP genes except *ARO1* were obtained by PCR amplification using primers adding homologous flanks for integration into the *ARO1* promoter region or shr flanks facilitating *in vivo* assembly of the cassettes. Since *ARO1* is a large gene of 4767 bp, its overexpression was achieved by replacing the native promoter by the strong consecutive promoter $ENO2_p$. Primers 14173 & 14174 were used to amplify the $ARO4^{K229L}$ fragment using pUD759 as template DNA. The *PHA2* cassette was amplified using primers 12656 & 12666 and pUD754 as template DNA. Primers 14175 & 14176 were used to amplify the *ARO2* region using pUD760. pUD804 was used as template to amplify the *ARO9* fragment with primers 14177 & 14463. Primer pair 14302 & 14459 were used to obtain the $ARO7^{T226I}$ region with pUD714 as template. Finally, the *ENO2* promoter was amplified from pGGKp028 using primers 12660 & 14181. All obtained fragments were co-transformed in IMX1783 together with pUDR478 (gRNA-$ARO1_p$), targeting the *ARO1* promoter. Subsequently, pUDR478 was recycled and the strain containing overexpressed (feedback resistant) alleles of all 3ABP genes without plasmid was stocked as IMX1952 ($3ABP^{FBR\uparrow}$)

Integration of the overexpression cassette for *TKL1* was done by amplifying the region with primers 12047 and 14563 from pUD660 adding 60 bp homology flanks to *SGA1*. The obtained fragment was co-transformed together pUDR372 (gRNA-*SGA1*) resulting in strain IMX1955 and IMX2029 before and after plasmid removal respectively. Subsequently, *ARO8* was deleted in IMX2029 by transforming this strain with pUDR553 (gRNA-*ARO8)* and an *ARO8* repair oligo (primers 8856 & 8857). pUDR553 was then recycled and the strain was stocked as IMX2056. IMX2056 was then transformed with pUDC245-pUDC252 or pGGKd015 resulting in strains IMC124 ($EcaroG^{P150L}$), IMC125 (*EcaroB*), IMC126 (*EcaroD*), IMC127 (*EcaroE*), IMC128 (*EcaroL*), IMC129 (*EcaroA*), IMC130 (*EcaroC*), IMC131 ($EcpheA^{T326PL}$) and IMC132 (*URA3*).

The native *TYR1* promoter of IMX2029 was replaced by four different promoters (*AGE1*$_p$, *MRI1*$_p$, *SEC18*$_p$, and *YEN1*$_p$). Primer pairs 14948 & 14947, 14998 & 14952, 14999 & 14949 and 14945 & 14946 were used to amplify the *AGE1*$_p$, *MRI1*$_p$, *SEC18*$_p$ and *YEN1*$_p$ regions respectively. Obtained promoter regions contained 60 bp homology flanks to the *TYR1* promoter region and *TYR1*. The exact promoter sequence of *YEN1*$_p$ was already described and shown to lead to low gene expression (196). For *AGE1*$_p$, 800 nucleotides upstream of the gene were used as promoter region. For the *MRI1*$_p$ and *SEC18*$_p$ all the nucleotides until the upstream gene were used as promoter region. By Cas9 mediated editing using pUDR568 (gRNA-*TYR1*$_p$) and the individual promoter integration fragments, the native *TYR1* promoter was replaced by one of the four promoters in both the IMX581 and the IMX2029 background. After plasmid removal, for the IMX581 background this resulted in strains IMX2068 (*YEN1*$_p$-*TYR1*), IMX2069 (*AGE1*$_p$-*TYR1*), IMX2071 (*SEC18*$_p$-*TYR1*) and IMX2072 (*MRI1*$_p$-*TYR1*). For the IMX2029 background the strains IMX2073 (*YEN1*$_p$-*TYR1*), IMX2074 (*AGE1*$_p$-*TYR1*), IMX2076 (*SEC18*$_p$-*TYR1*) and IMX2077 (*MRI1*$_p$-*TYR1*) were constructed after discarding its plasmids. In order to study the effect of a complete *TYR1* knockout, IMX581 and IMX2029 was also transformed with pUDR550 (gRNA-*TYR1*) and a repair oligo consisting of an artificial gRNA (primers 14935 & 14936) resulting in strains IMX2051 and IMX2052, respectively, after plasmid removal.

IMX2056 was transformed with pUDR574 (gRNA-*PYK1.1*) using repair oligo 14853 & 14854 or with pUDR577 (gRNA-*PYK1.2*) using repair oligo 14851 & 14852, to obtain silent mutations in *PYK1*. The SNPs were confirmed using Sanger sequencing and stored as IMX2102 (*PYK1*^{A336S}) and IMX2107 (*PYK1*^{D147N}), respectively and as IMX2106 (*PYK1*^{A336S}) and IMX2108 (*PYK1*^{D147N}) after losing their plasmid.

In both IMX2106 and IMX2108, the native *TYR1*$_p$ promoter was replaced using the same cas9 plasmid (pUDR568) and the same *YEN1*$_p$ integration fragment as described before resulting in strains IMX2123 and IMX2124 respectively after discarding the plasmid. Next, IME471 was obtained by transforming IMX2124 with the episomal plasmid pUDE001 overexpressing *ARO10* (81).

Subsequently, IME471 was transformed with *EcaroL* and *ARO3*^{K222L} integration cassettes using pUDR599 (gRNA-*X3*) for the cas9 targeting of the *X3* region (197). pUDC249 served as template for amplifying *EcaroL*, whereas *ARO3*^{K222L} was obtained from pUD1037. *EcaroL* and *ARO3*^{K222L} integration cassettes were amplified using primers 15831 & 15697 and 15694 & 15830, respectively, obtaining strain IMX2179 (*EcaroL*$^{\uparrow}$*ARO3*$^{K222L\uparrow}$) after plasmid removal.

Next, the *Bbxfpk* expression cassette was amplified from pUDC289 using primer pair 15377 & 16009 and the *Ckpta* expression cassette was obtained using primers 15333 & 16008 with pUDC290 as template (146). Both fragments were cotransformed in IMX2179 together with

pUDR598 (gRNA-*GPP1*) for targeting *GPP1* resulting in strain IMX2222 after discarding the plasmid.

Finally, the *ADH2* ORF was amplified from pUD1091 using primers 12934 and 16499, adding flanks with homology to the *X2* intergenic region. IMX2179 was transformed with the obtained fragment and pUDR547 (gRNA-*X2)* targeting th*e X2* region. After recycling the plasmid, the strain was stocked as IMX2279 (*ADH2*$^{\uparrow}$). An overview of the strain genealogy is given in **Fig. 2**.

All transformed strains were re-streaked three times prior to stocking. Correct integration or deletion was checked using colony PCR using diagnostic primers given in **Supplementary Table 1.** When required, plasmid removal in all strains was performed by growing the strain without the selective pressure required for retaining the plasmid. After obtaining single colonies, a single colony was re-streaked on both selective and non-selective medium. The colony growing only on non-selective medium was assumed to have discarded its plasmid and was used to inoculate liquid medium and stocked (198, 199).

Characterization of tyrosine prototrophy

To study potential growth defects caused by tyrosine auxotrophy of the *TYR1* promoter replacement strains, spot plate assays were performed. The number of cells of exponentially growing strain cultures were counted using the Z2 Coulter counter (Beckman Coulter, Indianapolis, IN) according to manufacturer's protocol. The cultures were washed and dilutions of 10^6, 10^5, 10^4 and 10^3 cells per 0.5 mL were prepared using sterile demineralized H_2O. Of each of these dilutions, 5 µL was spot-plated on petri plates containing YPD, SMD (Ura) and SMD (Ura + Tyr). Plates were incubated at 30 ˚C for 48 hours (YPD) to 72 hours (SMD).

Strain characterization

Growth rates were obtained from aerobic batch cultures grown in independent biological duplicates. Exponentially growing cells from the same medium and temperature were used to inoculate 500 mL shake flasks containing 100 mL SMD with additional uracil when required at a starting OD_{660} of 0.5.

For 2PE production studies, all strains were inoculated at a starting OD_{660} of 0.2 in biological triplicates. The cells were grown for 24 or 48 hours based on the time required to deplete all glucose. Samples were centrifuged for 5 min at 16,050 *g* and the supernatant was collected for metabolite analysis.

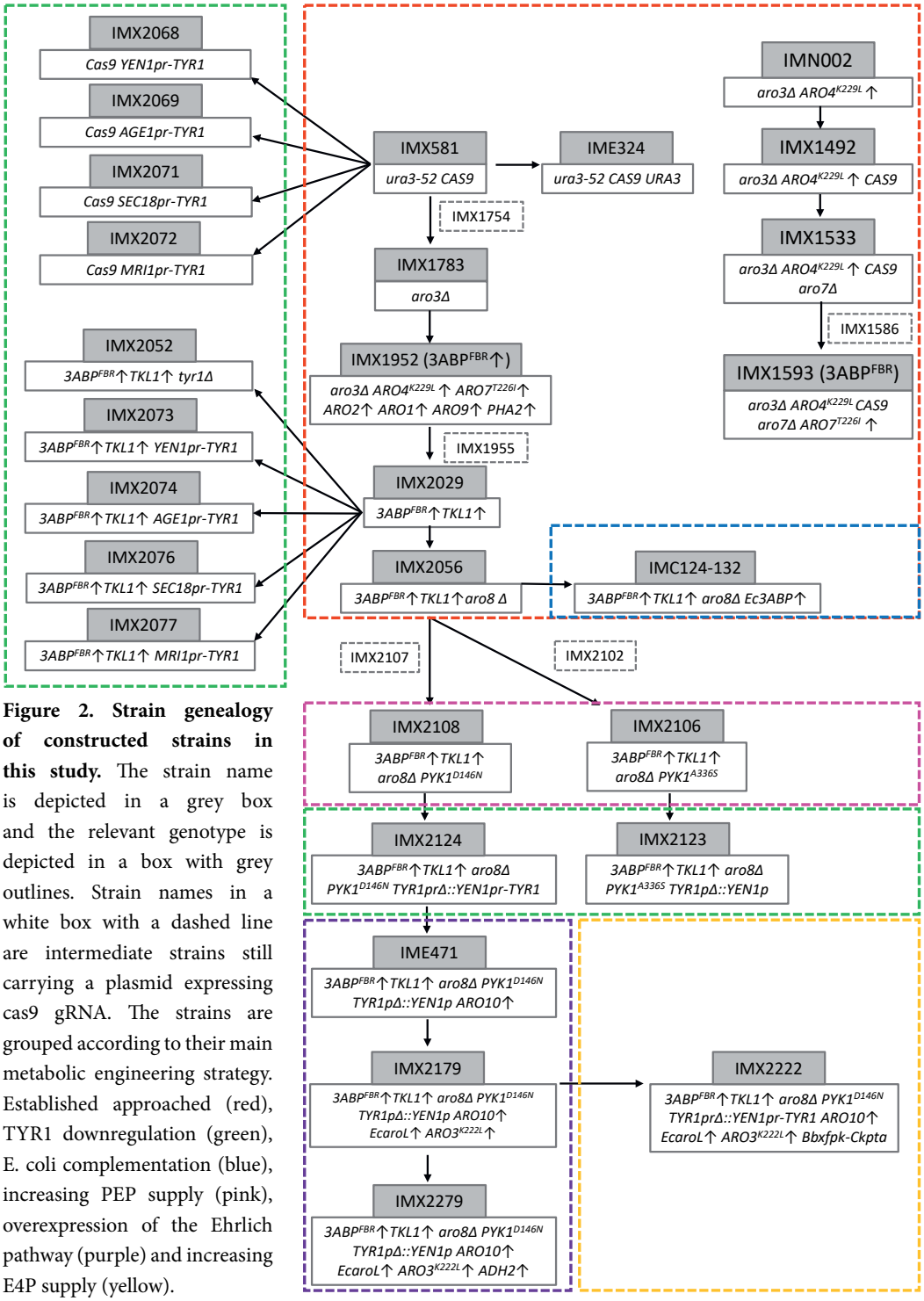


Figure 2. Strain genealogy of constructed strains in this study. The strain name is depicted in a grey box and the relevant genotype is depicted in a box with grey outlines. Strain names in a white box with a dashed line are intermediate strains still carrying a plasmid expressing cas9 gRNA. The strains are grouped according to their main metabolic engineering strategy. Established approached (red), TYR1 downregulation (green), E. coli complementation (blue), increasing PEP supply (pink), overexpression of the Ehrlich pathway (purple) and increasing E4P supply (yellow).

Extracellular organic acids, sugars and ethanol were determined by high performance liquid chromatography (HPLC) analysis using an Aminex HPX-87H ion-exchange column (Agilent, Santa Clara, CA) with 5 mM H_2SO_4 as mobile phase and a flow rate of 0.6 mL min^{-1} at 60°C. Glucose, glycerol, and ethanol were detected by a refractive-index detector (Agilent G1362A) and organic acids by a dual-wavelength absorbance detector (Agilent G1314F). Ehrlich metabolites (2PE, *p*OH2PE, PPA, PPY, CHR, SHK, *p*OHPPY, *p*OHPAA were measured using an Agilent Zorbax Eclipse plus C18 column (4.6 x 100 mm, 3.5 micron) (Agilent) with 0.020 M KH_2PO_4 (pH 2.0) and 1% acetonitrile as mobile phase with a flow rate of 0.8 mL min^{-1} at 40˚C. The amount of acetonitrile was increased to 10% in 6 minutes, followed by an increase to 40% until 23 minutes. From 23 to 27 minutes the amount of acetonitrile was reduced to 1% again and kept at this amount until 30 minutes. Detection of compounds occurred by means of a diode array and multiple wavelength detector (Agilent G1315C) at different wavelengths: 200 nm for *p*OH2PE, PPY, and PAA, 214 nm for shikimate and 2-phenylethanol, 280 nm for *p*OHPAA and 310 nm for *p*OHPPY.

Aerobic batch bioreactor cultures were performed in 2-L bioreactors (Applikon, Delft, The Netherlands) using a working volume of 1.2 L. The bioreactors were filled with synthetic medium with extra nitrogen source (SMN). SMN contained 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, and trace elements and 10 g L^{-1} $(NH_4)_2SO_4$ to avoid nitrogen limitation (132). After heat sterilization of the bioreactors, 20 g L^{-1} glucose, 0.2 g L^{-1} antifoam C (Sigma-Aldrich, Zwijndrecht, The Netherlands), and filter-sterilized vitamins solution were added (183).

The bioreactors were inoculated from exponentially growing cells in preculture shake flasks containing 100 mL of SMD to an initial biomass concentration of 0.08 g L^{-1}. The aerobic batch cultivations were performed at 30°C at a stirrer speed of 800 rpm. The culture pH was maintained at 5.0 by automated addition of 2M KOH and pressurized air was sparged through the bioreactors at 0.6 L min^{-1} to supply oxygen and strip produced carbon dioxide.

Optical density was measured by using a Jenway 7200 spectrophotometer (Jenway, Staffordshire, United Kingdom) at 660 nm. Cell dry weights were determined via filtration of 10 mL of well-mixed sample over dry nitrocellulose membrane filters with a pore size of 0.45 µm (Pall Corporation, Port Washington, NY). Filters were washed with demineralized water and dried in a microwave oven for 20 min at 360 W.

For analysis of carbon dioxide production and oxygen consumption in the bioreactor, the off-gas was first cooled in a condenser on the bioreactor (2°C) and dried with a Permapure MD-110-48P-4 dryer (Permapure, Lakewood, NJ), after which CO_2 and O_2 concentrations in the off-gas were measured with an NGA 2000 Rosemount gas analyser (Rosemount Analytical, Irvine, CA).

Results

Effects of the alleviation of allosteric regulation of the aromatic amino acid biosynthetic pathway, its overexpression and the supply of erythrose-4-P through transketolase are additive with respect to 2PE synthesis.

The aromatic amino acid biosynthesis pathway has been intensively exploited to produce a wide range of metabolites that derive from the aromatic amino acid biosynthetic pathway such as stilbenoids e.g. resveratrol (200, 201), hydroxycinnamic acids e.g. as *trans*-cinnamic acid (202, 203) or *p*-coumaric acid (98), flavonoids e.g. naringenin, kaempferol (98, 132), anthocyanin (61, 62) and even opioids (33). All these compounds share the yeast precursors phenylalanine and/or tyrosine. However, in contrast to the presented work, in these reference production strains the capacity to produce 2PE had to be reduced or eliminated since it represents an unwanted by-product. In a first step towards increasing 2PE production, we verified that some of the metabolic engineering strategies proposed for the biosynthesis of these phenylalanine and tyrosine derived compounds were also beneficial for de novo 2-phenylethanol production.

The alleviation of the allosteric feedback regulation of the DAHP synthases and to a lesser extend of the chorismate mutase activities have been shown to be instrumental for *de novo* production of 2PE. The strains IMN002 (*aro3*Δ *ARO4*$^{K229L\uparrow}$) (39) and IMX1593 (*aro3*Δ *ARO4*$^{K229L\uparrow}$ *ARO7*$^{T226I\uparrow}$) produced more than 1mM extracellular aromatic higher alcohols while the control strain IMX581 did not produce any. Both IMN002 and IMX1593 did produce detectable amounts of shikimate in the supernatant. Subsequently the native *ARO1* promoter was exchanged by the strong consecutive enolase 2 gene *ENO2* promoter (191) and the overexpression cassettes of ARO4^{K229L}, *ARO7*^{T226I} and the remaining 3ABP genes *ARO2, ARO9* and *PHA2* were introduced upstream of the *ARO1* locus assembling a 3ABP genes cluster on CHRIV. The resulting strain IMX1952 produced 56% more aromatic higher alcohols. The increase was mainly attributed to 2PE that increased by 86% while *p*OH2PE slightly decreased by 9% (**Fig. 3**).

Next, to maximize the flux towards erythrose-4-phosphate, the transketolase gene *TKL1* was overexpressed. In contrast to previous work (147), this overexpression was not coupled with the deletion of the glucose-6-phosphate dehydrogenase gene *ZWF1* since *ZWF1* deletion resulted in an excessive reduction in growth rate. The strain IMX2029 (*aro3*Δ *ARO4*$^{K229L\uparrow}$ *ARO7*$^{T226I\uparrow}$ 3ABP$^{\uparrow}$ *TKL1*$^{\uparrow}$) secreted an extra 40% and 118% of aromatic higher alcohols relative to IMX1952 and IMX1593.

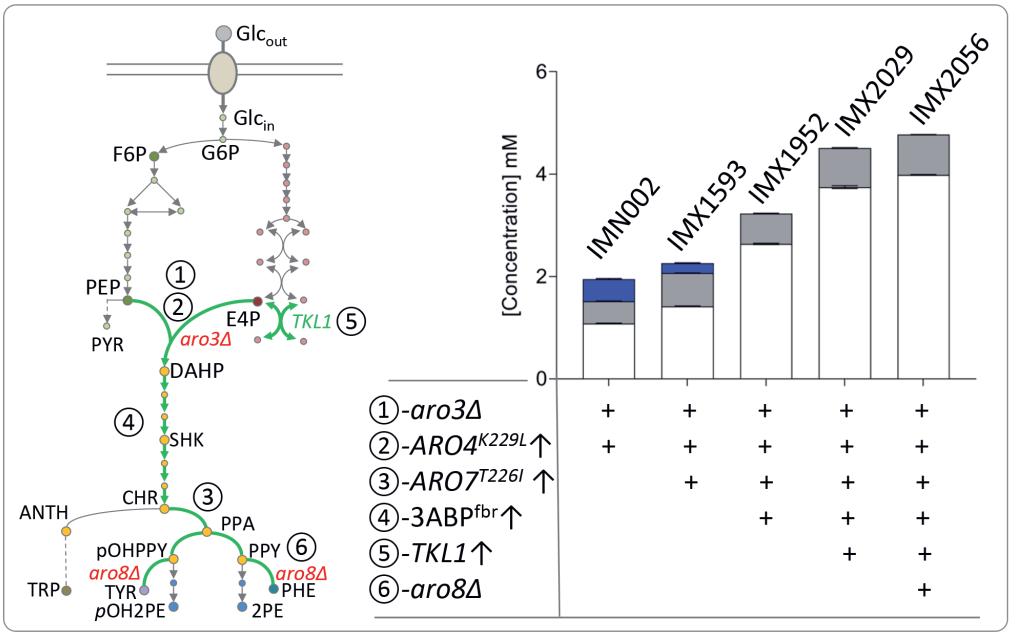


Figure 3. Recapitulation of validated modifications affecting the aromatic amino acid biosynthesis pathway. Left panel- schematic representation of the genetic modifications introduced in strains IMN002 (39), IMX1593, IMX1952, IMX2029 and IMX2056. The modifications include (1) the *ARO3* deletion, (2) alleviation of allosteric regulation of the DAHP synthase, (3) alleviation of allosteric regulation of the chorismate mutase, (4) overexpression of the aromatic amino acid biosynthetic pathway comprising feedback insensitive variants of Aro4 and Aro7, (5) overexpression of the transketolase *TKL1* and (6) the deletion of *ARO8*. The green arrows denote an overexpression. Right panel- Production of 2PE (white bar), *p*OH2PE (grey bar), shikimate (blue bar) in (mM). Strains were grown at 30 ˚C for 24 hours at 200 RPM in 500 mL shake flasks containing 100 mL minimal synthetic medium supplemented with 150 mg/L uracil when required and 20 g/L glucose as carbon source. The values represent averages ± mean deviations of data obtained from independent triplicate cultures. + denotes the presence of a set of genetic modifications in a given strain.

Finally, the deletion of *ARO8* that encodes an aromatic aminotransferase with broad substrate specificity was introduced in IMX2029. The 2PE and *p*OH2PE extracellular concentrations of IMX2056 (*aro3*Δ *aro8*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑) cultures were only 6% higher than that of the parental strain IMX2029 reaching 3.81 ± 0.03 mM and 0.75 ± 0.02 mM respectively after 24 hours of shake flask cultivations, still representing a small but significant increase (**Supplementary Table 3**).

Overexpression of *EcaroL* (shikimate kinase) eliminates the formation of shikimate and improves 2PE biosynthesis.

The first strains of the lineage IMN002 (*aro3*Δ *ARO4*^{K229L}↑) and IMX1593 (*aro3*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑) were excreting low but measurable shikimate, suggesting a catalytic limitation downstream of the shikimate point. In *S. cerevisiae* the main part of the shikimate

pathway is catalysed by a pentafunctional enzyme Aro1 which complicates tuning gene expression at individual activity levels. In *E. coli*, the shikimate pathway reactions are catalysed by enzymes encoded by individual genes. To systematically evaluate the impact of each single *E. coli* gene, the strain IMX2056 (*aro3*Δ *aro8*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑) was transformed with plasmids expressing a single (feedback resistant) *E. coli* gene (*EcaroG*^{P150L}, *EcaroB, EcaroD, EcaroE, EcaroL, EcaroA, EcaroC* and *EcpheA*^{T326P}). The strains IMC124 to IMC131 were compared to the reference strain IMC132 which corresponds to IMX2056 transformed with an empty *URA3* plasmid to restore strain prototrophy. Independent of whether the overexpressed genes were located upstream or downstream of shikimate in the pathway, the corresponding strains were still producing shikimate with strain IMC128 (*EcaroL)* producing the lowest amount. Most interestingly, this was accompanied by a 29% increase of the $\Sigma_{(2PE+2pOH2PE)}$ (**Fig. 4**). These results were in line with previously published results (109) that demonstrated that indeed the phosphorylation step of shikimate to shikimate 3-phosphate catalysed by shikimate kinase represented a target in the metabolic engineering of 2PE production in *S. cerevisiae*.

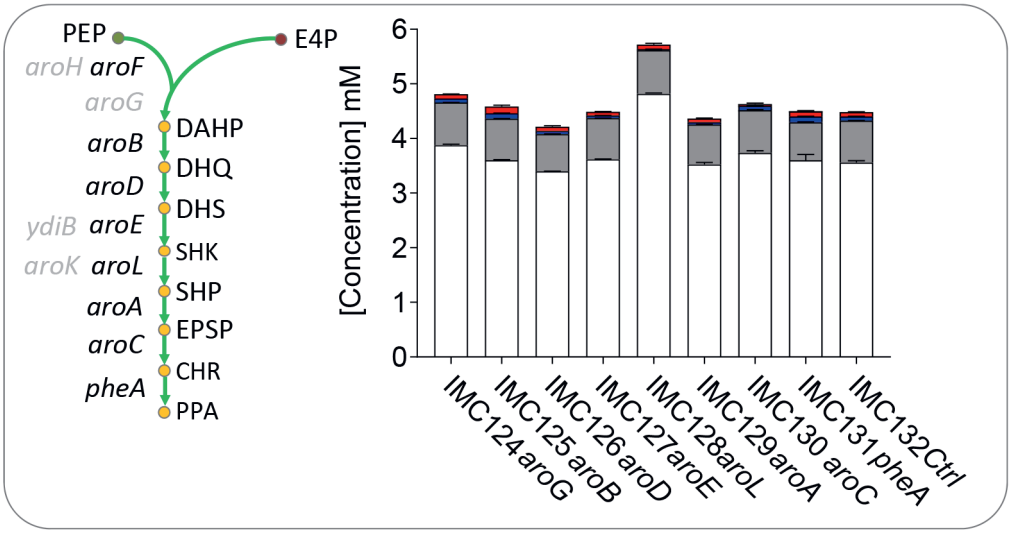


Figure 4. Effect of the overexpression of single gene from the *E. coli* shikimate pathway *aroA-E, G*^{P150} (that encodes a tyrosine insensitive DAHP synthase), *aroL* and *pheA*^{T326P} (that encodes a phenylalanine insensitive bi-functional chorismate mutase-prephenate dehydratase) on 2PE and *p*OH2PE biosynthesis in *S. cerevisiae*. Left panel-schematic representation of the *E. coli* shikimate pathway. The genes indicated in grey were not overexpressed. DAHP 3-deoxy-D-arabino-heptulosonate-7-phosphate, DHQ 3-dehydroquinate, DHS 3-dehydro-shikimate, SHK shikimate, SHP shikimate-3-phosphate, ESPS 5-enolpyruvyl-shikimate-3-phosphate, CHR chorismate and PPA prephenate. Right panel- Production of 2PE (white bar), *p*OH2PE (grey bar), SHK (blue bar) and phenylpyruvate (red bar) in (mM). Strains were grown at 30 ˚C for 24 hours at 200 RPM in 500 mL shake flasks containing 100 mL minimal synthetic medium supplemented with 150 mg/L uracil when required and 20 g/L glucose as carbon source. The values represent averages ± mean deviations of data obtained from independent triplicate cultures. + denotes the presence of a set of genetic modifications in a given strain.

Elimination of *p*-hydroxyphenylethanol (*p*OH2PE) biosynthesis leads to an increase of 2PE formation.

In the engineered strains IMX1593, IMX1952, IMX2029 and IMX2056 extracellular *p*-hydroxyphenylethanol (*p*OH2PE) concentrations were ranging from 15 to 30% of the total aromatic higher alcohol produced. The latest common intermediate of the phenylalanine and tyrosine branches, prephenate, can be either directed to phenylalanine in a reaction catalysed by Pha2 that performs the dehydration of prephenate to phenylpyruvate or is directed to tyrosine in a reaction catalysed by Tyr1 that performs the oxidative decarboxylation of prephenate (PPA) yielding *p*-hydroxyphenylpyruvate (*p*OHPPY). The deletion of *TYR1* would eliminate the production of *p*OH2PE but also introduce an auxotrophy for tyrosine which is a rather undesirable trait. Instead, fine-tuning the expression of *TYR1* could lower the production of the by-product while preventing tyrosine auxotrophy. For this purpose, the suitable replacement for the *TYR1* promoter should be weaker but also constitutively expressed to enable growth irrespective of the culture conditions. A microarray compendium of 170 steady-state chemostat cultures of the yeast *S. cerevisiae* that encompass 55 unique conditions was investigated (204, 205). In this compendium *TYR1* already exhibited a rather low expression (75.9 ± 23.4 AU) that is lower than average gene expression levels set at 110 in the compendium (206). In addition to a lower expression than *TYR1* the selected genes should exhibit a coefficient of variation of its expression level lower than 25% over the entire conditions set. Four genes with expression profiles fulfilling these requirements were identified in the compendium (**Table 2**).

The promoter sequences of *YEN1*, *AGE1*, *SEC18* and *MRI1* (832, 800, 362 and 257 bp respectively) were cloned and used to replace the chromosomal *TYR1* promoter in the strain control strain IMX581. A deletion of *TYR1* in the same strain (IMX2051) cancelled growth on medium lacking tyrosine, an auxotrophy that can be rescued by addition of

Table 2: Profiling overview of genes displaying lower expression than *TYR1* and stable expression (Coefficient of variation (CV)<15%) in a tranriptome compendium covering 55 different conditions. Expression values are the means (Avg) ± standard deviations (Std) of data from 170 GeneChip (Affymetrix YG-S98) analyses of samples issued from different steady-state chemostat cultivations (204, 205). Expression values are expressed in Affymetrix signal units (AU).

Gene name	Systematic name	Functional description	Avg (AU)	Std (AU)	CV (%)
TYR1	YBR166C	Prephenate dehydrogenase (NADP+)	75.9	23.4	31.0
YEN1	YER041W	Similarity to DNA repair protein Rad2p and Dsh1p	31.5	5.6	18.0
AGE1	YDR524C	Similarity to hypothetical human protein and *YIL044c*	40.1	9.8	24.4
SEC18	YBR080C	Cytoplasmic ATPase involved in protein transport between ER and Golgi	62.3	15.0	24.1
MRI1	YPR118W	Similarity to *M.jannahii* translation initiation factor, eIF-2B	75.0	16.9	23.0

tyrosine in the culture medium or growth on complex medium (YPD). Meanwhile, the four constructed strains IMX2068 (*YEN1*$_p$), IMX2069 (*AGE1*$_p$), IMX071 (*SEC18*$_p$) and IMX2072 (*MRI1*$_p$) exhibited growth on chemically defined medium lacking tyrosine demonstrating the functionality of the four promoters. The three strains IMX2069 (*AGE1*$_p$), IMX071 (*SEC18*$_p$) and IMX2072 (*MRI1*$_p$) could grow as fast as the control strain IMX581 in tyrosine free medium (**Supplementary Table 3**). In contrast IMX2068 (*YEN1*$_p$) was significantly slower than IMX581 and the other three strains with a specific maximum growth rate (μ) of 0.33± 0.01 h^{-1} instead of 0.38 ± 0.00 h-1.

Next, to evaluate the impact on the Ehrlich higher alcohol production, the *TYR1* promoter of strain IMX2029 (*aro3*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑) was replaced by one of the four selected regulatory sequences. In this genetic background, a *TYR1* deletion was also sufficient to prevent growth on medium lacking tyrosine, an auxotrophy that can be rescued by addition of tyrosine in the culture medium or growth on complex medium (YPD) (**Fig. 5**). In contrast to the *tyr1Δ* mutant (IMX2052) all strains carrying the exchanged promoter were able to grow on medium without tyrosine (**Fig. 5**). The strains IMX2073 (*YEN1*$_p$), IMX2074 (*AGE1*$_p$), IMX2076 (*SEC18*$_p$) and IMX2077 (*MRI1*$_p$) were also grown for 24h in chemically defined medium (SMD) and extracellular concentration of the Ehrlich metabolites were measured. Out of the four constructs only IMX2073 (*YEN1*$_p$) displayed a 90% reduced *p*OH2PE formation. Despite a slower growth rate, the strain could still consume all sugars in 24h and produce up to 6 mM of 2PE which represents a significant 60% increase relative to

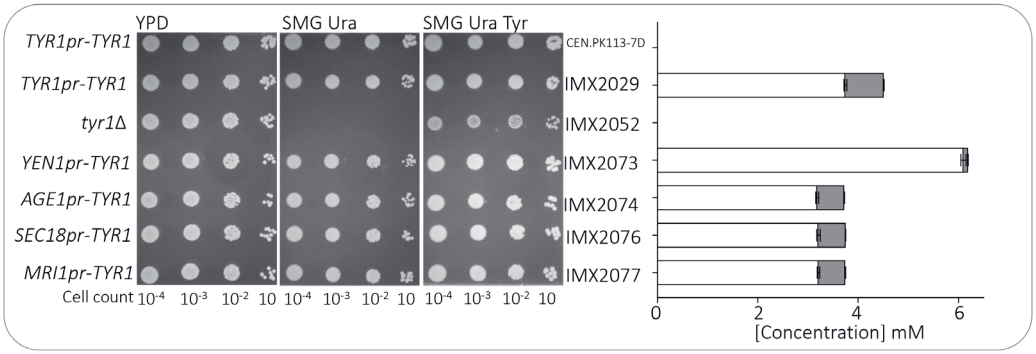


Figure 5: Modulation of *TYR1* expression. Left panel- Spot plate assay of 2PE overproducing IMX2029 (*aro3*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑), its derived progeny strains IMX2052 (*tyr1*Δ), IMX2073 (*YEN1*$_p$*-TYR1*), IMX2074 (*AGE1*$_p$*-TYR1*), IMX2076 (*SEC18*$_p$*-TYR1*) and IMX2076 (*MRI1*$_p$*-TYR1*) on YPD, SMD supplemented with 150 mg L^{-1} uracil and SMD supplemented with 150 mg L^{-1} uracil (SMD Ura) and 225 mg L^{-1} tyrosine (SMD Ura Tyr). Four different cell concentrations were spotted and were incubated for 48 hours (YPD) or 72 hours (SMD based media) at 30 ˚C. Right panel- Production of 2PE (white bar), *p*OH2PE (grey bar) in (mM). Strains were grown at 30 ˚C for 24 hours at 200 RPM in 500 mL shake flasks containing 100 mL minimal synthetic medium supplemented with 150 mg/L uracil when required and 20 g/L glucose as carbon source. The values represent averages ± mean deviations of data obtained from independent triplicate cultures.

the parental strain IMX2029 (**Fig. 5**). The *TYR1* promoter replacement by $YEN1_p$ has not only contributed to a near complete conversion of *p*OH2PE into 2PE but has also enabled a further increase of 2PE concentrations ($^{IMK2029}\Sigma_{(2PE+2pOH2PE)}$ = 4.5 ± 0.03 mM < $^{IMK2073}\Sigma_{(2PE+2pOH2PE)}$ = 6.2 ± 0.06 mM).

Novel strategies to improve precursor supply of the shikimate pathway.

The shikimate pathway (**Fig. 1**) is initiated by the condensation of the two metabolite precursors phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P). Further downstream in the shikimate pathway, a second molecule of PEP is consumed at the step of 3-phosphoshikimate 1-carboxyvinyltransferase, catalysed by Aro1. PEP is also a critical metabolite in the glycolysis since it serves as substrate of pyruvate kinase, an enzyme that in addition to forming pyruvate is a key metabolite in both respiratory and fermentative metabolism. Furthermore, PEP is also an intermediate in *S. cerevisiae* anaplerotic functions.

While investigating the reversibility of the PEP-carboxykinase (PEPCK) reaction, *PYK1* mutant alleles were identified and these mutations, although not thoroughly characterised, enabled growth on glucose of a pyruvate carboxylase *pyc1*Δ *pyc2*Δ double mutant expressing a PEPCK from *Actinobacillus succinogenes* (207). This result suggested that the flux through pyruvate kinase was decreased. The two variant alleles of *PYK1* ($PYK1^{D147N}$ and $PYK1^{A336S}$) previously identified were reconstructed in IMX2056 (*aro3*Δ *aro8Δ* $ARO4^{K229L}\uparrow$ $ARO7^{T226I}\uparrow$ 3ABP$\uparrow$ $TKL1\uparrow$) by CRISPR guided *in vivo* directed mutagenesis resulting in IMX2106 (*aro3*Δ *aro8Δ* $ARO4^{K229L}\uparrow$ $ARO7^{T226I}\uparrow$ 3ABP$\uparrow$ $TKL1\uparrow$ $PYK1^{A336S}$) and IMX2108 (*aro3*Δ *aro8Δ* $ARO4^{K229L}\uparrow$ $ARO7^{T226I}\uparrow$ 3ABP$\uparrow$ $TKL1\uparrow$ $PYK1^{D147N}$).

As anticipated, the disturbance of the flux through the pyruvate kinase had unwanted side effects. Both strains grew slower than the parental strain with a 17% (0.29 ± 0.00 h^{-1} vs 0.35 ± 0.00 h^{-}) and 34% (0.23 ± 0.00 h^{-1} vs 0.35 ± 0.00 h^{-1}) reduction of the IMX2106 and IMX2108 growth rate, relative to the parental strain IMX2056. As already observed for strain IMX2073 ($YEN1_p$-*TYR1*), the increase in Ehrlich pathway derived aromatic alcohols seemed to be inversely correlated with the strains growth rate. The slower strain IMX2108 ($PYK1^{D147N}$) showed the best performance, producing 7.4 ± 0.04 mM of 2PE against 6.1 ± 0.16 mM for IMX2106 ($PYK1^{A336S}$) and 4.0 ± 0.01 mM for the parent IMX2056 (**Fig. 6**).

The presence of the native *TYR1* regulation was corroborated with the production 1 mM of *p*OH2PE. To prevent the formation of the tyrosine derived higher fusel alcohol, the native *TYR1* promoters of IMX2106 and IMX2108 were replaced by the $YEN1_p$ as previously shown in strain IMX2073 ($YEN1_p$). The combination of expression of a *PYK1* mutant and the down-regulated *TYR1* expression resulted in even better performance.

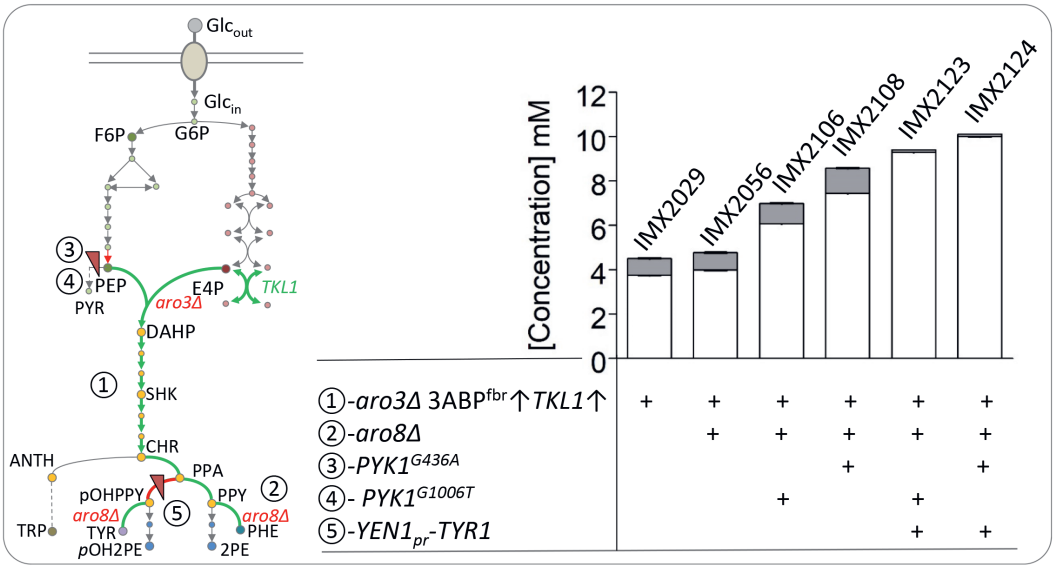


Figure 6. **Implementation of novel metabolic engineering strategies to increase 2PE production in *S. cerevisiae***. Left panel- schematic representation of the genetic modifications introduced in strains IMX2029, IMX2056, IMX2106, IMX2108, IMX2123 and IMX2124. The modifications include (1) the introduction of the overexpression of the aromatic amino acid biosynthetic pathway comprising mutant alleles of *ARO4* and *ARO7,* (2) the deletion of *ARO8,* (3)(4) the down tuning of the catalytic activity of the pyruvate kinase Pyk1 and (5) the alteration of *TYR1* expression. The green arrows denote an overexpression and the red arrows denotes either transcription down regulation or a reduction of the catalytic capacity of the enzyme. Right panel- Production of 2PE (white bar), *p*OH2PE (grey bar) in (mM). Strains were grown at 30 ˚C for 24 hours at 200 RPM in 500 mL shake flasks containing 100 mL minimal synthetic medium supplemented with 150 mg/L uracil when required and 20 g/L glucose as carbon source. The values represent averages ± mean deviations of data obtained from independent triplicate cultures. + denotes the presence of a set of genetic modifications in a given strain.

The best combination consisted of the association of the *YEN1*$_p$*-TYR1* expression cassette with the *PYK1*^{D147N} allele. The strain IMX2124 (*aro3*Δ *aro8*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑ *PYK1*^{D147N} *YEN1*$_p$*-TYR1*) reached the 10mM 2PE threshold (**Fig. 6**). The tuning of the flux of the pyruvate kinase by expressing mutant alleles showed to be a successful approach to increase the 2PE production in *S. cerevisiae.*

These two new impactful modifications were incrementally combined with the overexpression of the Ehrlich pathway broad substrate 2-oxo acid decarboxylase *ARO10* (IME471), the overexpression of the *E. coli* shikimate kinase *EcaroL (109)* and the *ARO3*^{K222L} allele encoding a tyrosine feedback insensitive DAHP synthase (94, 208) (IMX2179) and finally the overexpression of *ADH2* encoding an alcohol dehydrogenase (IMX2279) (209, 210). Hitherto, measurements of extracellular metabolites were performed after 24h, a time point at which all sugars were consumed, but newly constructed strains grew slower (μ_{IME471}=0.15 h^{-1}; $\mu_{IMX2179}$=0.16 h^{-1}; $\mu_{IMX2279}$=0.15 h^{-1}) and therefore were grown for 48h to allow completion of all sugars. All three strains showed improvement relative to the parental strain IMX2124 (**Fig. 7**).

The performances of IMX2179 and IMX2279 were undistinguishable from one another and produced the highest titer measured so far of 12.6 ± 0.3 mM and 12.7 ± 0.04 mM of 2PE respectively. The overexpression of *TKL1* has previously been shown to positively contribute to the production of 2PE by improving E4P supply. To further augment E4P availability to initiate the aromatic amino acid pathway, a new metabolic connection between the glycolysis and the pentose phosphate pathway was implemented. A heterologous pathway consisting of a phosphoketolase from *Bifidobacterium breve* (*Bbxfpk*) (146) and a phosphotransacetylase from *Clostridium kluyveri* (*Ckpta*) (42) was integrated in IMX2179 resulting in strain IMX2222. The phosphoketolase can convert fructose-6-phosphate into E4P and acetyl-phosphate redirecting a fraction of the glycolytic flux towards the pentose

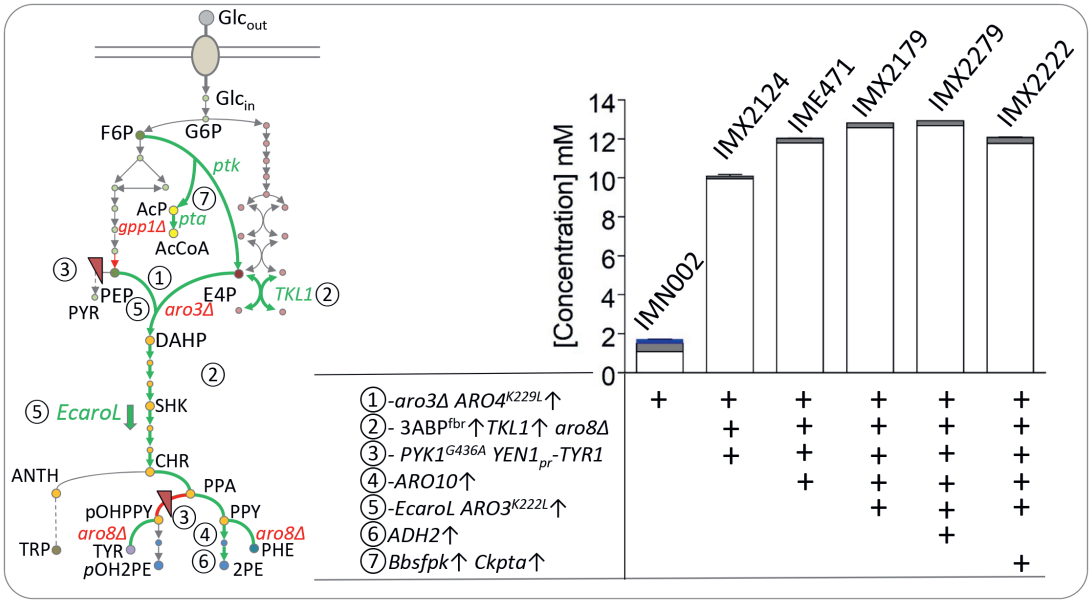


Figure 7: **Combination of strategies to increase precursors supply and overexpression of the Ehrlich pathway**. Left panel- schematic representation of the genetic modifications introduced in strains IMN002, IMX2124, IME471, IMX2179, IMX2279 and IMX2222. The modifications include (1) alleviation of allosteric regulation of the DAHP synthase (Aro4^{K229L}), (2) the combination of the *ARO8* deletion, overexpression of the aromatic amino acid biosynthetic pathway comprising mutant alleles of *ARO4* and *ARO7* and overexpression of the transketolase Tkl1, (3) the down tuning of the catalytic activity of the pyruvate kinase Pyk1 and the alteration of *TYR1* expression, (4) overexpression of the Ehrlich pathway 2-oxo decarboxylase Aro10, (5) overexpression of the *E. coli* shikimate kinase EcAroL and the feedback insensitive DAHP synthase variant Aro3^{K222L}, (6) Overexpression of the alcohol dehydrogenase Adh2 and (7) expression of the heterologous phosphoketolase by-pass including the *B. brevis* phosphoketolase (*Bb*Xfpk) and a phosphotransacetylase from *C. kluyveri* (*Ck*Pta) (146). The green arrows denote an overexpression and the red arrows denotes either transcription down regulation or a reduction of the catalytic capacity of the enzyme. Right panel- Production of 2PE (white bar), *p*OH2PE (grey bar), shikimate (blue bar) in (mM). Strains were grown at 30 ˚C for 48 hours at 200 RPM in 500 mL shake flasks containing 100 mL minimal synthetic medium supplemented with 150 mg/L uracil when required and 20 g/L glucose as carbon source. The values represent averages ± mean deviations of data obtained from independent triplicate cultures. + denotes the presence of a or a set of genetic modifications in a given strain.

phosphate pathway and E4P more specifically. The acetyl-phosphate hydrolysis was shown to be reduced upon deletion of *GPP1* encoding a glycerol-3-phosphate dehydrogenase (146). Therefore, to limit accumulation of acetate originating from acetyl-phosphate hydrolysis, the two bacterial genes were integrated at the *GPP1* locus resulting in its disruption. Additionally, the phosphotransacetylase can convert acetyl phosphate in acetyl-CoA which can be further used in yeast metabolism (42) . The strain IMX2222 grew faster than the parental strain ($\mu_{IMX2222}$=0.23 h^{-1} >$\mu_{IMX2179}$=0.16 h^{-1}), but unfortunately did not show an improvement relative to the best performing strain constructed so far: IMX2279, instead the extracellular concentration of 2PE and total aromatics was slightly decreased by 7.2 and 6.6 % respectively.

2PE production in controlled aerobic batch cultures.

To further evaluate the influence of the new metabolic engineering solutions describeded in this study, we compared the performance of strains at an early and late stage of the improvement progamme (**Fig. 8**): the strain IMX1955 (*aro3*Δ *ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑) that harboured a set of modifications available at the start of the study (39, 109, 147), the strain IMX2179 (*aro3*Δ *aro8Δ ARO4*^{K229L}↑ *ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑ *PYK1*^{D146N} *YEN1*$_p$*-TYR1 ARO10*↑ *EcaroL*↑ *ARO3*^{K222L}↑) the highest producing strain which also included the metabolic engineering targets identified in this study and IMX2222 (*aro3*Δ *aro8Δ gpp1Δ ARO4*^{K229L}↑*ARO7*^{T226I}↑ 3ABP↑ *TKL1*↑ *PYK1*^{D146N} *YEN1*$_p$*-TYR1 ARO10*↑ *EcaroL*↑ *ARO3*^{K222L}↑ *Bbxfpk*↑ *Ckpta*↑) the strain expressing the phosphoketolase-phostransactylase pathway were characterized under controlled batch conditions. The strain IMX1955 resulted from the transformation of IMX2029 with an *URA3* plasmid to restore prototrophy.

The three strains along with the reference IME324 (*Spycas9*) were cultured in a 2L batch bioreactor with 20 g·L^{-1} glucose at a fixed pH of 5.0. In these conditions (aerobic and glucose excess), *S. cerevisiae predominantly metabolizes glucose through fermentation metabolism (211, 212)* and exhibits a diauxic growth profile. As expected, IME324 did not produce any fusel alcohols or intermediates of the shikimate pathway when grown in these conditions. The strain had a growth rate of 0.35 h^{-1} (**Table 3**) and was completing ethanol consumption in a typical diauxic shift (**Fig. 8**). The intermediate 2PE producing strain IMX1955 produced 4.42 ± 0.25 mM of 2PE in 30 h. The best 2PE producing strain IMX2179 grew even slower with a growth rate 0.1 h^{-1} but produced 12.7 ± 0.25 mM of 2PE in 40 h. Interestingly IMX2179 showed an 88% reduction in *p*OH2PE ([*p*OH2PE]$_{IMX2179}$ =0.172 ± 0.00 mM vs [*p*OH2PE]$_{IMX1955}$ =1.38 ± 0.01 Mm) confirming the positive effect of downregulating the prephenate dehydrogenase *TYR1* gene.

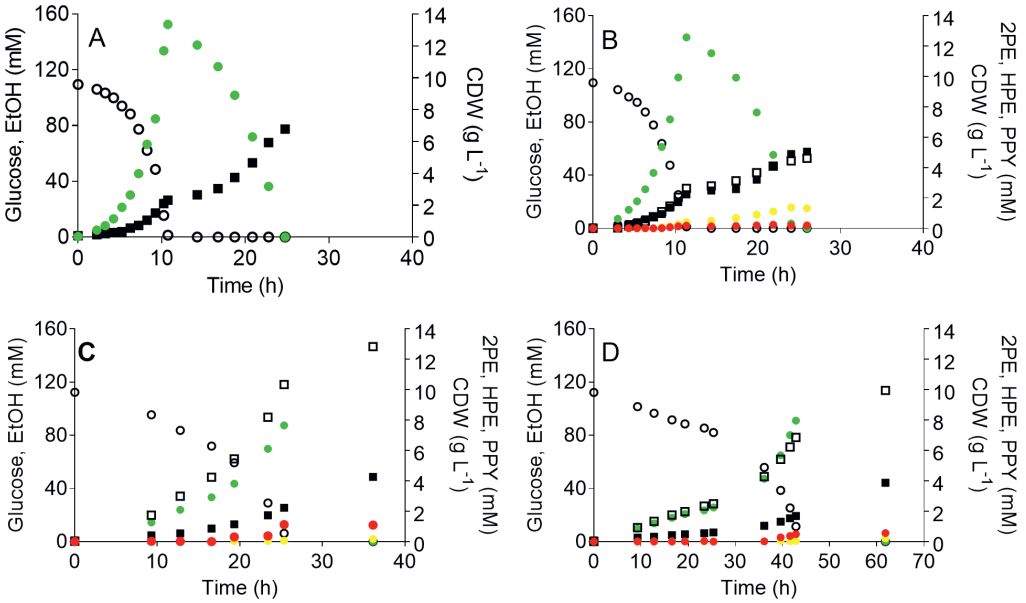


Figure 8. Product formation of the 2PE-producing strains S. cerevisiae IMX1955, IMX2179, IMX2222 and the control strain IME324 in bioreactors. Growth and extracellular metabolite formation were studied in aerobic and pH controlled (pH 5.0) batch cultures on glucose. The results shown are from a single representative experiment. Panel A: IME324, panel B: IMX1955, panel C: IMX2179 and panel D: IMX2222. ■ Cell dry weight (g L^{-1}), □ 2PE, ○ glucose, ● ethanol, ● *p*OH2PE, ● PPY (mM).

Conversely to what was observed in shake flask cultures the strain IMX2179 produced 1 mM PPY (phenylpyruvate) (**Fig. 8**). In these conditions the 2PE yield on glucose of the strain IMX2179 was calculated to be of 0.113 ± 0.002 mol/mol (**Table 3**) which represents nearly one quarter of the 2PE maximum theoretical yield on glucose (0.500 mol/mol or 0.339 g/g).

In line with the shake flask results the strain IMX2222 showed reduced performance compared to IMX2179. But contrasting with the shake flask cultures this strain grew much slower in a controlled bioreactor than in shake flask (0.06 ± 0.00 h^{-1} versus 0.23 + 0.00 h^{-1}). Additionally, IMX2179 and IMX2222 displayed an unusual growth profile, the off-gas CO_2 profile (**Supplementary Fig. 1**), revealed that the growth phase on glucose was not steady including three distinct phases. After an initial growth phase, the culture slowed down before resuming exponential growth. In both cases the glucose phase was followed by a diauxic ethanol phase. Although showing similar CO_2 profiles, IMX2222 exhibited an elongated growth phase twice longer than that of IMX2179.

Table 3. Fermentation performance indicators of 2PE producing strains. Specific growth rate (μ), yields (Y) of biomass (X) and ethanol (EtOH) on glucose (S) and yield (Y) of biomass, 2-phenylethanol (2PE), p-hydroxyphenylethanol (pOH2PE) and phenylpyruvate (PPY) on glucose in aerobic bioreactor batch cultures of S. cerevisiae strains IME324 (control strain), IMX1955, IMX2179 and IMX2222 carrying several genetic modifications for increased 2PE production.

Strain	**IME324**	**IMX1955**	**IMX2179**	**IMX2222**
Relevant genotype	*cas9*	*cas9 aro3Δ aro8Δ* $3ABP^{fbr}$↑ *TKL1*↑	*cas9 aro3Δ aro8Δ* $3ABP^{fbr}$↑ *TKL1*↑ $PYK1^{D146N}$ $YEN1_p$*-TYR1* *ARO10*↑ *EcaroL*↑ $ARO3^{K222L}$↑	*cas9 aro3Δ aro8Δ* *gpp1Δ* $3ABP^{fbr}$↑*TKL1*↑ $PYK1^{D146N}$ $YEN1_p$*-TYR1 ARO10*↑ *EcaroL*↑ $ARO3^{K222L}$↑ *Bbxfpk*↑ *Ckpta*↑
[a]μ_{max} (h^{-1})	0.35 ± 0.06	0.30 ± 0.01	0.10 ± 0.01	0.06 ± 0.00
[a]$Y_{X/S}$ (g g^{-1})	0.12 ± 0.00	0.11 ± 0.00	0.11 ± 0.00	0.09 ± 0.00
[a]$Y_{EtOH/S}$ (mol mol^{-1})	1.41 ± 0.00	1.33 ± 0.00	0.81 ± 0.02	0.88 ± 0.03
$Y_{X/S}$ (g g^{-1})	0.34 ± 0.00	0.26 ± 0.01	0.21 ± 0.00	0.18 ± 0.01
$Y_{2PE/S}$ (mol mol^{-1})	-	0.04 ± 0.00	0.11 ± 0.00	0.09 ± 0.00
$Y_{pOH2PE/S}$ (mol mol^{-1})	-	0.01 ± 0.00	0.002 ± 0.00	0.002 ± 0.000
$Y_{PPY/S}$ (mol mol^{-1})	-	0.002 ± 0.00	0.01 ± 0.00	0.005 ± 0.000

[a] Determined from the glucose phase only.

Discussion

Hitherto metabolic engineering strategies devised to improve the production of aromatic metabolites derived from the three aromatic amino acids were axed around the relief of regulatory constraints of the shikimate pathway (109, 132). Indeed, as most of the amino acid biosynthetic pathways, the phenylalanine, tyrosine and tryptophan biosynthetic routes are tightly regulated at transcriptional and (post)-translational levels (89, 91, 92). The alleviation of the metabolic control of the DAHP synthase and chorismate mutase was proven to be key to unlock the production of metabolites derived from the shikimate pathway (39). Furthermore, the relief of the native transcriptional control by overexpression of the genes of the shikimate pathway led to the construction of the best aromatic compounds (phenylpropanoids, flavonoids, alkaloids) producers so far (61, 62, 202, 213). Although well-established, these approaches were applied to produce compounds downstream of phenylalanine, but not 2-phenylethanol which represents a by-product for the formation of these compounds. In line with these previous works, the relief of allosteric and transcriptional regulation of the shikimate pathway positively impacted the formation of 2PE. While these modifications led to an increase in catalytic capacity of the pathway, our study more specifically focussed on the augmentation of the supply of the DAHP synthase substrates, PEP and E4P and elimination of metabolic by-products.

Firstly, improving PEP supply by tuning pyruvate kinase activity ended successfully. Earlier attempts to improve PEP supply for tyrosine biosynthesis by lowering *PYK1* expression or expressing specific mutant contrastingly did not lead to any improvement but instead showed to negatively impact the aromatic amino acid synthesis (148). Also, the phenotype of the strains carrying the *PYK1*^{D147N} allele used in this study suggests a weakening of the kinase activity (207), that might lead to an increase in intracellular PEP, the exact reasons at the origin of this improvement remains uncharacterized and further work such as in vitro enzymatic activity measurements or intracellular metabolites quantification will be necessary to provide a full explanation.

Conversely, the second approach consisting in the diversion of a fraction of the glycolytic flux into the non-oxidative branch of the pentose pathway by expressing a phosphoketolase (*Bb*Xfpk) -phosphotransacetylase (*CkPta*) pathway did not result in an increase of Ehrlich higher alcohol biosynthesis. Phosphoketolase does not only hydrolyze fructose-6-phosphate (C6) into erythrose-4-phosphate (E4P) and acetyl-phosphaste. Instead, most of these enzymes including the *Bifidobacterium breve* phosphoketolase exhibit higher activity with xylulose-5-phosphate (146). The significant accumulation of acetate in IMX2222 culture supernatant supported the functionality of the *Bifidobacterium* shunt (214, 215), despite the deletion of *GPP1* that prevents the excessive endogenous hydrolysis of acetyl-phosphate (146). The

absence of improvement might be associated with too low fructose-6-phosphate (F6P) intracellular concentration. Several genetic modifications have already been shown to affect intracellular F6P concentration such as deletion of one of the two phosphofructokinases, *PFK1* or *PFK2* (216) or deletion of one of the 6-phosphofructo-2-kinases, *PFK26/PFK27,* that catalyze the conversion of F6P to fructose-2,6-biphosphate (217). Implementation of these changes in IMX2222 could yield F6P concentrations compatible with the full expression of the metabolic potential of the *Bifidobacterium* shunt and eventually higher 2PE concentration.

While in shake flask the strain IMX2222 exhibited a better 2PE productivity than IMX2179, these results were reversed in bioreactors. Both strains displayed a reduced growth rate but this phenotype was exacerbated in IMX2222 (**Fig. 8** and **Fig. S1**). Controlled bioreactors, in contrast to shake flasks, are sparged with pressurized air stripping produced carbon dioxide. The *PYK1* mutations applied in this study arose during evolution in a pyruvate carboxylase negative strain (*pyc1Δ pyc2Δ*) expressing a phosphoenolpyruvate carboxykinase (*PEPCK*) of *Actinobacillus succinogenes* to enable a sufficiently high anaplerotic flux to sustain growth (207). Growth was only observed in strains carrying either of the *PYK1* mutations under high CO_2 concentrations. The thermodynamic potential of carboxylation reactions is enhanced under a CO_2 rich atmosphere. The $PYK1^{D146N}$ mutation has been engineered in both IMX2179 and IMX2222. The resulting decrease in pyruvate levels effects the maximal growth rate but could potentially also result in a lowered anaplerotic flux through Pyc1, Pyc2 and Mae1. These enzymes would favour from increased CO_2 conditions. A lower initial CO_2 and pyruvate levels might be detrimental as these strains might be unable to produce sufficient CO_2 and pyruvate derived metabolites required for biosynthetic reactions and fast growth. This effect could be even stronger for IMX2222 because the phosphoketolase will divert even more carbon away from pyruvate.

The points discussed so far have implications beyond 2PE as these strategies could be applied to the biosynthesis of a wide range of products that are based on the aromatic amino acid carbon skeleton. In contrast, metabolic engineering strategies aiming at the elimination of the by-product *p*OH2PE biosynthesis are limited to 2PE. The activation of the Ehrlich pathway indiscriminately stimulates the formation of all fusel alcohols and acids (39, 76, 82, 172, 200, 218). While fusel acids phenylacetate (PAA) and *p*-hydroxyphenylacetate (*p*OHPAA) were systematically absent, we could not exclude the presence of tryptophol (aka. Indole-3-ethanol), the Ehrlich alcohol derived from tryptophan. If so, a similar approach as applied for *TYR1* could be applied at the chorismate node by down-regulating the anthranilate synthase genes *TRP2* and *TRP3*.

The best engineered strains were able to produce up to 12.7 mM of 2PE but exhibited a strong reduction up to 75% in growth rate relative to the lineage parental strains CEN.PK113-7D or

IMX581 (**Table 3, Fig. 8**). This growth decrease was systematically observed in intermediate strains. Remarkably, these two strains characteristics were near perfectly inversely correlated (Coefficient of correlation R^2=0.9237) (**Fig. 9)**.

This linear model predicts a complete growth inhibition around 19.4 mM of 2PE ethanol, a value in line with 2PE inhibitory concentration found in literature that range between 10 and 25mM (219). This effect could be anticipated as 2PE has a partition coefficient octanol/water ($LogP_{O/W}$) of 1.36 denoting lipophilic characteristics. This is in agreement with some of the proposed 2PE toxicity mechanisms, such as increasing membrane fluidity (220) that consequently would cause leakage of ions (221), a reduced uptake of nutrients such as glucose and amino acids (222) and a reduction of the respiratory capacity, possibly due to changes in mitochondrial membrane fluidity (178).

Therefore, considering this toxicity, further improvement of the 2PE yield on glucose ($Y_{2PE/S}$) to approach the maximum theoretical yield of 0.500 mol mol^{-1} will certainly require to adopt fermentation modes including *in situ* product recovery techniques to continuously remove the product from the broth to prevent its toxicity. The strain performances reported in this study are the highest titers of 2PE ever achieved *de novo* in *S. cerevisiae* or any other *Saccharomycotina* yeasts (181). Further, the strain IMX2179 can be used as platform for the development of the next 2PE producing strain generation.

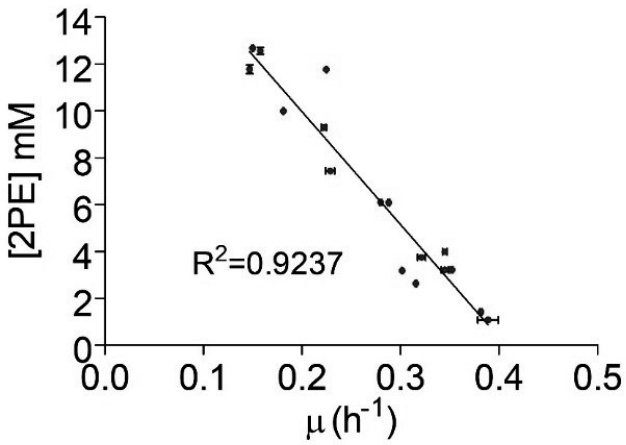


Figure 9: Correlation between growth rate and extracellularly produced 2PE in metabolically engineered *S. cerevisiae* strains. Data plotted in this graph correspond to the 2PE concentrations reported in Figures and the strain growth rates measured in the same experiments (**Table S.3** and **S.4**). The plotted values represent averages ± mean deviations of data obtained from independent triplicate cultures. The R^2 value represents the correlation coefficient of the linear regression through the data points. The line shows a Y-intercept of 19.6 mM which simulates full growth inhibition.

Financial support

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 720824

Declaration of interest

None.

Acknowledgements

We would like to thank Dr. John Morrissey and Dr. Arun Rajkumar for fruitful discussion. We would like to thank Mr. Erik de Hulster for the supervision of the fermentation and Francine Boonekamp for providing template plasmids for the construction of promoter and terminator parts.

Chapter 3

Modular, synthetic chromosomes as new tools for large scale engineering of metabolism.

Else-Jasmijn Hassing*, Eline D. Postma*, Venda Mangkusaputra, Jordi Geelhoed, Pilar de la Torre, Marcel van den Broek, Christiaan Mooiman, Martin Pabst, Jean-Marc Daran, Pascale Daran-Lapujade

*These authors contributed equally to this work.

Supplementary materials available online

https://doi.org/10.4121/14979876

Abstract

The construction of powerful cell factories requires intensive genetic engineering for the addition of new functionalities and the remodeling of native pathways and processes. The present study, using *Saccharomyces cerevisiae*, explores the potential of modular, specialized *de novo*-assembled neochromosomes (named NeoChrs) to facilitate extensive genome reprogramming. Linear and circular NeoChrs carrying 20 native and 21 heterologous genes were designed and constructed by *in vivo* assembly in yeast in two transformation steps leading to *de novo* production of anthocyanins, native to plants. Turned into exclusive expression platforms for heterologous and essential metabolic routes, the NeoChrs behaved like native chromosomes regarding mitotic and genetic stability, copy number, harmlessness for the host and editability by CRISPR/Cas9. This study paves the way for future microbial cell factories with modular genomes in which core metabolic networks, localized on satellite, specialized NeoChrs can be swapped for alternative configurations and serve as landing pads for the addition of functionalities.

Abbreviations: PHE L-phenylalanine, PPY phenylpyruvate, 2PE 2-phenylethanol, PAA phenylacetic acid, TYR tyrosine, *p*OHPPY *p*-hydroxyphenylpyruvate, *p*OH2PE *p*-hydroxyphenylethanol, *p*OHPAA *p*-hydroxyphenylacetic acid, CIN cinnamic acid, COUM coumaric acid, COCOA coumaroyl-CoA, NARCC naringenin chalcone, PHLOR phloretic acid, NAR naringenin, DHK dihydrokaempferol, LPE leucopelargonidin, KAE kaempferol, K3G kaempferol 3-O-glucoside, PEL pelargonidin, P3G pelargonidin 3-O-glucoside.

Enzymatic reactions: PAL phenylalanine ammonia lyase, TAL tyrosine ammonia lyase, CPR cytochrome P450 reductase, C4H cinnamate 4-hydroxylase, 4CL 4-coumarate CoA ligase, CHS chalcone synthase, CHI chalcone isomerase, F3H: flavanone 3-hydroxylase, DFR dihydroflavonol 4-reductase, ANS anthocyanidin synthase, 3GT anthocyanin 3-*O*-glucosyltransferase.

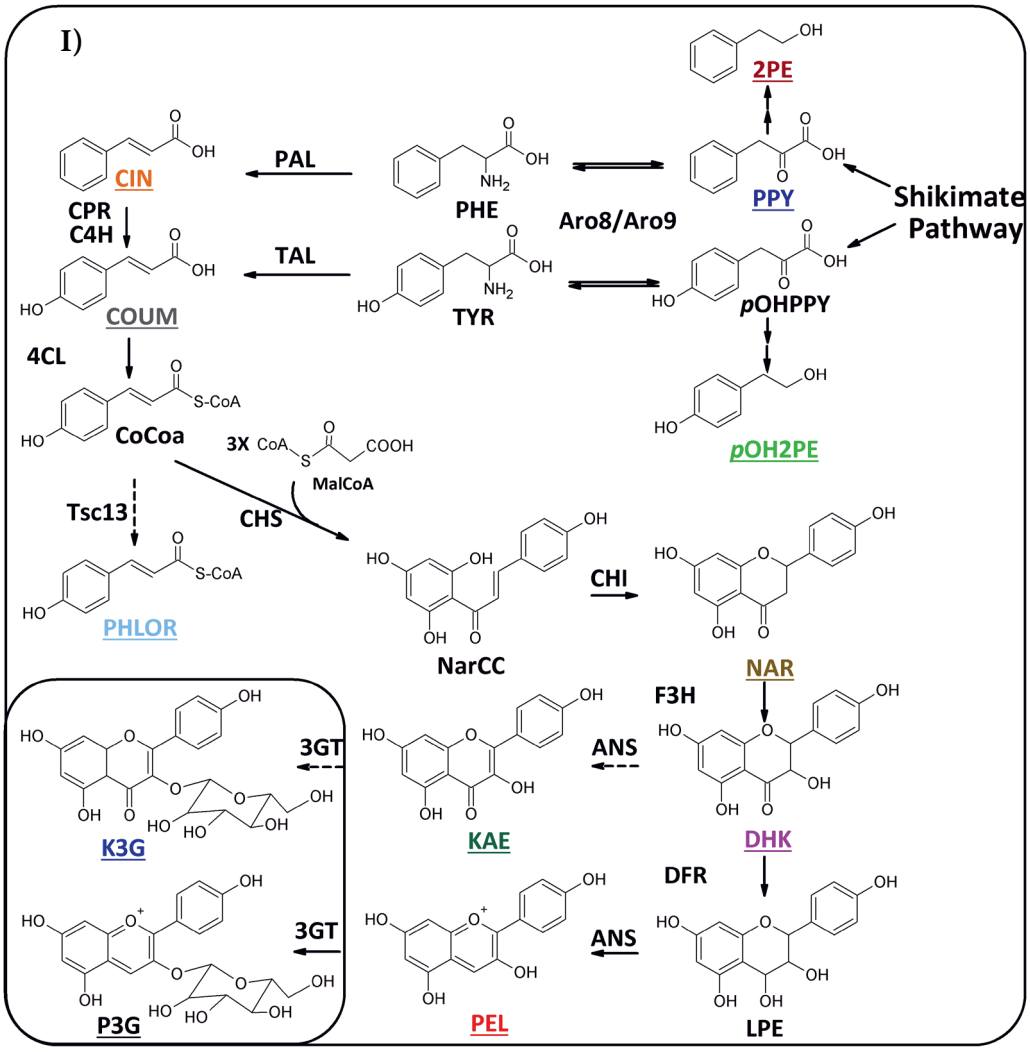
Introduction

While microbial cell factories have great potential for the sustainable production of fuels, chemicals and therapeutics, microbial processes are often less economically attractive than chemical, oil-based processes (32, 223). Increasing the cost efficiency of microbial cell factories requires high product titer, rate and yield, features only attainable by intensive genetic engineering of the microbial host via costly strain construction programs. These programs focus on the transplantation of new functionalities, as well as on the reprogramming of the microbial host metabolic networks in which the new functionalities are plugged. These networks of biochemical reactions are encoded by several hundreds of genes scattered over large mosaic microbial genomes. Even considering the CRISPR/Cas9 revolution, rewiring these biochemical networks remains a daunting challenge, in particular for eukaryotic cell factories characterized by a high degree of genetic redundancy (224). Microbial platforms in which core biochemical networks can be remodeled at will have a great role to play in the construction of powerful cell factories, but also to reach a deep fundamental understanding of these biochemical networks and their regulation.

The pathway swapping concept was developed to address this persistent challenge (225). This modular concept is based on the genetic reduction and clustering of genes belonging to a metabolic pathway, so the pathway can be easily swapped by any other design. This concept was demonstrated in the model and industrial yeast *Saccharomyces cerevisiae*, using glycolysis and alcoholic fermentation, a 12-step pathway catalyzed by a set of 26 enzymes, as an example. In the SwYG (Switchable Yeast Glycolysis) strain this set of 26 genes was reduced to 13 and relocalized to a single chromosomal locus, enabling the facile remodeling of the entire pathway (225). Scaling up pathway swapping to the set of core biochemical reactions required for most biotechnological applications (i.e., central carbon metabolism) would offer an unprecedented ability to deeply reprogram cell factories metabolism. However, it would require the integration of hundreds of genes on existing chromosomes, which is a potential source of genetic instability. This problem can be addressed by the implementation of supernumerary, *de novo* assembled synthetic chromosomes (named NeoChrs) as orthogonal expression platforms for genome remodeling (226).

The present study explores the potential of combining pathway swapping with NeoChrs to simultaneously equip *S. cerevisiae* with new heterologous routes and remodel native metabolic networks. *De novo* production from glucose of the anthocyanin pelargonidin 3-O-glucoside (P3G), a food and industrial dye, was attempted in *S. cerevisiae* as proof of principle (60). Although already demonstrated in *S. cerevisiae*, the synthesis of anthocyanins is highly inefficient in all microbial platforms tested hitherto (60-62). Requiring the channeling of carbon through 27 core metabolic catalytic steps in the glycolytic, pentose phosphate, shikimate

and aromatic amino acids biosynthesis pathway, as well as 10 reactions native to plants (**Fig. 1** and **Supplementary Fig. 1**), P3G is a perfect paradigm to demonstrate the potential of NeoChrs for metabolic engineering. Modular NeoChrs in linear and circular form harboring yeast native, bacterial and plant genes required for P3G *de novo* synthesis were designed and constructed (**Fig. 1** and **Supplementary Fig. 1**). The resulting strains were tested by in-depth genetic and physiological characterization, and compared to strains carrying test NeoChrs of equivalent size and mostly composed of non-coding DNA. Ultimately, the ability of NeoChrs to serve as landing pad for large pathways and to be edited by CRISPR/Cas9 for metabolic engineering purposes could be demonstrated.



II)

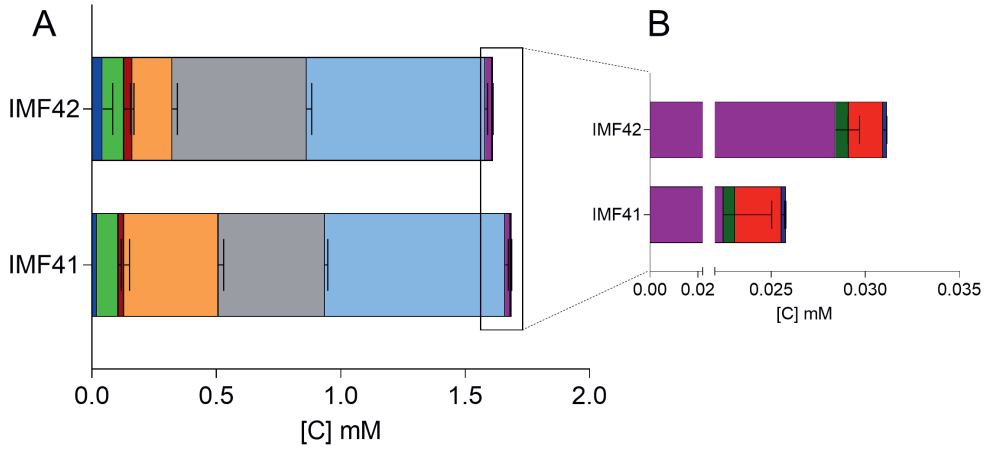


Figure 1. D*e novo* anthocyanin production using NeoChrs. A) Schematic overview of the anthocyanin production pathway. Compounds that can be measured by HPLC or GC are underlined, coumaroyl-CoA, narigenin-chalcone, and leucopelargonidin were not measured. The end products of the pathway are indicated in a frame. **B)** Extracellular production of aromatic intermediates of the anthocyanin pathway by engineered *S. cerevisiae* strains IMF41 (Circular NeoChr) and IMF42 (Linear NeoChr). The yeast strains were grown in shake-flask cultures with chemically defined minimal medium with 20 g L^{-1} glucose as carbon source and urea as nitrogen source for 72 hours. The colour of the bars corresponds to the colour of the metabolites in panel A. Phenylpyruvate (blue), *p*-hydroxyphenylethanol (green), 2-phenylethanol (dark red), cinnamic acid (orange), coumaric acid (grey), phloretic acid (sky blue), dihydrokaempferol (purple), kaempferol (dark green), pelargonidin (red) and kaempferol 3-O-glucoside (dark blue). Naringenin and pelargonidin 3-O-glucoside were not detected. The right panel shows a magnification of the data of the metabolites downstream of naringenin. The data represents the average ± standard deviation of independent biological triplicates.

PHE L-phenylalanine, PPY phenylpyruvate, 2PE 2-phenylethanol, TYR tyrosine, *p*OHPPY *p*-hydroxyphenylpyruvate, *p*OH2PE *p*-hydroxyphenylethanol, CIN cinnamic acid, COUM coumaric acid, CoCoa coumaroyl-CoA, NarCC naringenin chalcone, PHLOR phloretic acid, NAR naringenin, DHK dihydrokaempferol, LPE leucopelargonidin, KAE kaempferol, K3G kaempferol 3-O-glucoside, PEL pelargonidin, P3G pelargonidin 3-O-glucoside.

Materials and Methods.

Strains, growth medium and maintenance.

All *S. cerevisiae* strains used and constructed in this study are derived from the CEN.PK family (**Supplementary Table 1-2**) (182). The detailed protocols for *S. cerevisiae* and *E. coli* growth and maintenance are detailed in **Supplementary Methods 1**.

Molecular biology techniques

Genomic DNA extraction from *E.coli* and *S. cerevisiae*, different PCR techniques and gel-verification of the amplicons, methods for *E. coli* and yeast transformation and various kits used in this study are detailed in **Supplementary Methods 2**.

Plasmid construction

All plasmids used in this study are listed in **Supplementary Table 3.**

gRNA plasmids. GuideRNA (gRNA) plasmids (**Supplementary Table 3A**) for targeting Cas9 to specific loci were constructed as described by Mans *et al.* (188). Primers to construct and verify the gRNA plasmid are listed in **Supplementary Table 4-5**.

Golden Gate part plasmids. Part plasmids compatible with Golden Gate Assembly, harboring promoter, gene or terminator were constructed using the Yeast Toolkit principle (190). A range of part plasmids were present in-house and previously described (**Supplementary Table 3B**). Some of the promoters, genes and terminators flanked by *Bsa*I and *BsmB*I restriction sites were ordered from GeneArt (Thermo Fisher Scientific). The promoters and terminators listed in **Supplementary Table 3C** were subcloned by GeneArt (Thermo Fisher Scientific) in the entry vector pUD565 and could be directly used for construction of expression plasmids. The pentose phosphate pathway genes were ordered from GeneArt but subcloned in-house into entry vector pUD565 **(Supplementary Table 8D**). Finally, some part plasmids were made by amplifying the target region with primers containing part type specific overhangs (**Supplementary Table 6**) and assembling in entry vector pUD565 by *BsmB*I cloning. These plasmids are listed in **Supplementary Table 3E** with the respective primers and template for target region amplification. Internal *BsmB*I or *Bsa*I sites were removed as described by Hassing *et al.* (227)and verified by Sanger sequencing (Baseclear, Leiden, The Netherlands). For two parts (*coAtANS* #1 and *pSePYK1*) no correct *E.coli* part plasmid transformant was found. Therefore, the PCR fragments containing yeast toolkit flanks were directly assembled into expression cassettes (**Supplementary Table 3G**) Plasmids were verified by PCR and/or restriction analysis (**Supplementary Table 7**).

Golden Gate expression plasmids. All expression plasmids were made using *Bsa*I mediated

Golden Gate assembly (190, 227) in the pGGKd012 GFP dropout plasmid. All expression cassettes and the part plasmids (or PCR fragments) used for construction are outlined in **Supplementary Table 3F-G**. Plasmids were verified by PCR and/or restriction analysis (**Supplementary Table 8**). Correct assembly of some of the expression plasmids was verified by Sanger sequencing (Baseclear).

Gibson assembly expression plasmids. For two expression cassettes, namely: *pRPS3-coAtCPR1-tIDH2* and *pSeTPI1-At4CL3-tSDH2,* the individual parts contained too many *Bsa*I/*Bsm*BI sites, making Golden Gate assembly impossible. Therefore, the individual parts were amplified with PCR primers containing homology flanks and the expression cassettes were constructed via Gibson Assembly (**Supplementary Table 3H** and **8**).

Strain construction

Construction of strains harboring test NeoChrs

Test NeoChrs, NeoChr10 and NeoChr11 consisted of 43 fragments namely: 36 non-coding 2.5 kb *E.coli* DNA fragments, *CEN6/ARS4, ARS1, ARS417, mRuby2, mTurquoise2, HIS3* and two TeSS fragments (**Supplementary Table 9**). The TeSS fragments were amplified from the telomerator plasmid pLM092 (228). To prevent possible circularization of the linear NeoChr, in NeoChr11 the sequence of the right TeSS was changed by modifying a few base pairs (PCR amplification with a mutated primer). Fragments were amplified by PCR using primers with 60 bp SHR flanks (193) (**Supplementary Table 10**). NeoChrs were assembled by transforming 200 fmol of each non-coding 2.5 kb *E.coli* DNA fragment, 200 fmol of each fluorescent marker, 200 fmol of the TeSS fragments, and 100 fmol of *CEN6/ARS4*, *HIS3* and the two *ARS* fragments. Transformants were checked by FACS, CHEF and long-read nanopore sequencing. One transformant (NeoChr10.13) missing only 2113 bp of fragment 7A was stocked as strain IMF22.

Construction of strains harboring NeoChrs designed for anthocyanin production

Construction of the host strain IMX2770. The host strain for the anthocyanin NeoChrs, IMX2270, was constructed by engineering of the SwYG strain IMX589 in which the minor paralogs of glycolysis and fermentation have been deleted and the major paralogs relocalized as a cluster to the *sga1* locus on chromosome IX (225). Using CRISPR/Cas9 editing (188), the minor paralogs of the pentose phosphate pathway, *GND2, NQM1, SOL4* and *TKL2*, were deleted to facilitate the subsequent relocalization of the major homologues to the NeoChrs. Additionally, the *ARO10* gene was removed to minimize phenylethanol production (132), resulting in strain IMX2270. The construction steps leading from the SwYG strain to IMX2270 are detailed in **Supplementary Methods 3, Supplementary table 11-14**.

Assembly of NeoChr25 and NeoChr26. In strain IMX2270 the circular NeoChr26 and the linear NeoChr25 were assembled. The fragments for assembly were identical between the two NeoChrs namely: 7 pentose phosphate pathway genes (*ZWF1, TKL1, GND1, RKI1, TAL1, RPE1, SOL3*), 13 glycolysis genes (*HXK2, PGK1, FBA1, TPI1, TDH3, GPM1, ENO2, PGI1, PYK1, PDC1, ADH1, PFK1, PFK2*), 10 codon optimized *E. coli* shikimate pathway genes (*coEcAroG^{fbr}, coEcAroB, coEcAroE, coEcAroL, coEcAroA, coEcAroD, coEcAroC, coEcTyrA^{fbr}, coEcPheA^{fbr}, coEcTyrB*), *HIS3* selection marker, two fluorescent markers (*mRuby2, mTurquoise2*), *CEN6/ARS4* and two ARS sequences. For NeoChr26 (circular), the telomerator fragment was transformed, while for NeoChr25 (linear) two TeSS were used instead. The fragments were amplified by PCR from the corresponding expression cassettes for all transcriptional units or from genomic DNA and other plasmids for auxiliary parts using primers with 60 bp SHR flanks (**Supplementary Table 15-16**). For each fragment 200 fmol was used, with the exception of *HIS3* and *CEN6/ARS4* fragments, where 100 fmol were added instead. The pooled transformation mix were concentrated with Vivacon 500 (Sartorius AG, Gottingen, Germany) up to a final volume of 50 μl and transformed in IMX2270. Both the NeoChr25 (linear) and NeoChr26 (circular) transformants were screened by FACS and long-read nanopore sequencing, and NeoChr25 was additionally also verified by CHEF. After sequence confirmation using short-read whole genome sequencing, the strains were stocked as IMF27 (NeoChr25) and IMF29 (NeoChr26).

Removal of single locus glycolysis from native chromosome. In IMF29 (circular NeoChr) the glycolytic cassette was removed from the *sga1* locus by induction of two double strand breaks (DSB) in the flanks of the cassette with the gRNA plasmid pUDR413, and providing a 120 bp repair fragment homologous to the upstream and downstream region of *sga1* (**Supplementary Table 3** and **17**). For IMF27 (linear NeoChr) the glycolytic cassette was replaced by a *KlURA3* transcriptional-unit by providing a *KlURA3* expression unit as repair fragment, obtained by PCR amplification from pMEL10 with primers containing flanks homologous to *sga1* locus (**Supplementary Table 17**). After discarding the gRNA plasmids, the strains were stocked as IMF31 (linear NeoChr) and IMF32 (circular NeoChr). The circular NeoChr in strain IMF32 contained a mutation in the *RKI1* ORF. The mutation was repaired by inducing a DSB in the vicinity of the mutation with the gRNA plasmid pUDR756, targeting only the watermarked gene but not the native copy, and repairing it with a 120 bp repair fragment (**Supplementary Table 3** and **18**). The repair fragment contained the non-mutated sequence and a silent mutation of the PAM sequence. A correct transformant was confirmed by Sanger sequencing (Baseclear) and named IMF35.

Deletion of major PPP paralogs from their native chromosomal loci. In a first transformation round, *ZWF1, SOL3, GND1* and *RKI1* were removed from IMF31 and IMF35 using the gRNA plasmids pUDR703 and pUDR700 and 120 bp repair fragments (**Supplementary Table 1**

and **19**). Plasmid removal resulted in strains IMF33 and IMF36 with linear and circular NeoChr, respectively. In a second transformation *RPE1, TKL1* and *TAL1* were removed using the gRNA plasmids pUDR701 and pUDR702 and 120 bp repair fragments (**Supplementary Table 2** and **19**). After plasmid recycling, the strains were stocked as IMF34 and IMF40, for the linear and circular NeoChr strains respectively.

Integration of plant anthocyanin pathway. The next step in strain construction was the integration of single copies of the plant genes of the anthocyanin pathway on the NeoChrs, *AtPAL1, coRcTAL1, coAtC4H, coAtCPR1, At4CL3, AtCHI1, coAtCHS3, coAtF3H, coGhDFR, coAtANS, coAt3GT,* under expression of strong, constitutive yeast promoters. The fragments were amplified from their corresponding Golden Gate expression plasmids with primers containing 60 bp SHRs for *in vivo* assembly (**Supplementary Table 20-21**). They were integrated in a single locus of the linear (strain IMF34) and circular (strain IMF40) NeoChrs by induction of a DSB in the *E.coli* fragment 16AB using gRNA plasmid pUDR765 (**Supplementary Table 2-3**). The two outer fragments contained 60 bp homology to a stretch of 60 bp in the *E.coli* fragment16AB and in the flanking SHR DL. After verification of correct integration by PCR (**Supplementary Table 21**), the gRNA plasmid was removed by unselective growth on SMD and the strains were stocked as IMF41 (circular NeoChr) and IMF42 (linear NeoChr).

Integration of multiple copies of *coAtCHS3*. To integrate multiple copies of *coAtCHS3*, the *coAtCHS3* expression unit was amplified from pUDC352 using primers adding 60 bp homology flanks to the integration sites Chunk 7BC, chunk 15CD, SHR N and chunk 9CD in the linear NeoChr31 and *CAN1, X2, YPRCtau3* and *SPR3* in native chromosomes (**Supplementary Table 1,2** and **22**). The *coAtCHS3* expression units carrying homology with native chromosomes were transformed to IMF42 using the gRNA plasmids pUDR771 and pUDR772 for targeted editing. After recycling of the gRNA plasmids, strain IMF44 (Lin, 5x *CoAtCHS3* (4 copies in native genome)) was obtained. IMF44 was then transformed with suitable *coAtCHS3* cassettes and gRNA plasmids pUDR780 and pUDR781 targeting four loci in NeoChr31 resulting in strain IMF47, containing nine copies of the *coAtCHS3* (five on NeoChrs and four on native chromosomes).

***coAtANS* repair**. *coAtANS* was repaired by integration of a correct copy of *coAtANS* in the *mTurquoise* gene of NeoChr33 in IMF47. The sequence of pUDC398 (*pSeENO2-coAtANS-tFUM1)* was verified and the transcriptional unit was amplified using primers 18740/18741 (**Supplementary Table 23**) adding flanks with homology to *mTurquoise* on NeoChr33. IMF47 was transformed with the template DNA and pUDR400 (gRNA-*mTurquoise*) for targeting *mTurquoise*. After PCR confirmation (**Supplementary Table 23**), pUDR400 was recycled and the final strain was stocked as IMF48 (linear NeoChr34, 9x *coAtCHS3,* repaired *coAtANS*).

Strain verification by PCR and whole genome sequencing

All strains constructed by CRISPR/Cas9 were verified by diagnostic PCR. In addition, whole genome short read sequencing was performed in-house for IMF27, IMF29 on an Illumina MiSeq Sequencer (Illumina, San Diego, CA) as described previously (226, 229). Strains IMF41, IMF42, and IMF47 were sequenced on a NovaSeq 6000 at GenomeScan Leiden (GenomeScan, Leiden, NL). Long-read sequencing of IMF22, IMF27, IMF29, IMF41, IMF42, IMF47 and IMF48 was performed in-house on MinION flow cells using the SQK-LSK109 sequencing kit with the EXP-NBD104 expansion kit (Oxford Nanopore Technologies, Oxford, United Kingdom) (**Supplementary Methods 4**).

Characterization of NeoChrs fidelity, stability and toxicity.

Fluorescence-based sorting by flow cytometry. Flow cytometry employing an BDFACSAria™ II SORP Cell Sorter (BD Biosciences, Franklin Lakes, NJ), was used to sort single cells in the chromosome stability experiment, as well as for screening of fluorescent transformants bearing the constructed NeoChrs as both described earlier by Postma *et al.* (226)

Karyotyping with CHEF analysis. Chromosomes separation and size determination was performed using contour-clamped homogeneous electric field electrophoresis as described earlier by Postma *et al.* (226)

Quantification of strain viability and NeoChr stability. NeoChr stability was assessed based on plating cells from shake flask cultures on selective and non-selective growth medium with respect to the NeoChr as described by Postma *et al.* (226). On the second- and fifth-day strain viability and stability was measured, on the third and fourth day, only the cell density (OD_{660}) of overnight cultures was measured. For the test-chromosome bearing strains (IMF22 and IMF23) viability and stability was based on sorting of 96 cells, while for CEN.PK113-7D, IMF27, IMF29, IMF34, IMF40, IMF41, IMF42 and IMF47, 384 cells were sorted on a microtiter plate.

Determination of specific growth rates. For determination in shake flasks, strains were inoculated from a -80^{o}C freezer stock in a 500 mL shake flask with 100 mL medium and grown until late-exponential phase. Cells were transferred to fresh 100 mL medium and grown until mid-exponential phase. Finally, from these cultures, shake flasks containing 100 mL medium were inoculated at an OD_{660} of 0.3. Cell density was measured with a 7200 Visible spectrophotometer (Cole-Parmer, Staffordshire, United Kingdom). A maximum specific growth rate (μ_{max}) was calculated from at least biological duplicates, using five data points and at least two doublings in the exponential phase. For growth rate determination in multi titer plates, a Growth Profiler (Enzyscreen, Heemstede, NL) was used as previously described (226) using plate 06 and a: 0.084327, b:5.35 × 10^{-8} , c: 4.398348, d: -0.41959 as constants.

Growth rates were calculated from six biological replicates, using ODs between 1 and 8.

Characterization of aromatics production by the engineered strains

Shake flasks. The strains were inoculated in biological triplicates in 500 mL shake flasks containing 100 mL synthetic medium with 20 g L^{-1} glucose. Urea was used as sole nitrogen source to prevent acidification of the medium, thereby allowing respiration of the produced ethanol. After 72 hours of cultivation, the optical density was measured using a Jenway 7200 spectrometer (Jenway, Staffordshire, United Kingdom) at 660 nm and aromatics were quantified as described below.

Aerobic batch bioreactors. Strains IMF41, IMF42 and IMF48 were characterized in bioreactors. 2L bioreactors (Applikon, Delft, The Netherland) were filled with 1.0 L synthetic medium, containing 20 g L^{-1} glucose and ammonium sulfate as nitrogen source. Exponentially growing cells were used to inoculate the bioreactors at an initial biomass concentration of around 0.12 g L^{-1}. The cultivations were performed at 30 ˚C, 800 RPM using 0.5 L min^{-1} pressurized air to sparge the bioreactor with oxygen. Automated addition of 2 M KOH and 2 M H_2SO_4 ensured maintenance of the culture pH at 5.0. Samples were taken at regular time intervals for optical density, metabolite concentrations and cell dry weights determinations. Cell dry weight, organic acids, sugars and ethanol concentrations were measured as previously described (227, 230).

Analysis of aromatics. Extracellular naringenin and upstream aromatics were quantified by HPLC as previously described (227) and elaborated upon in **Supplementary Methods 5A**. Sample preparation and quantification of intracellular and extracellular flavonoids downstream of naringenin using LC-MS/MS is described in **Supplementary Methods 5B.** Detection of P3G is illustrated in **Supplementary Fig. 2.**

Data Availability

Genome sequence data are available at NCBI as Bioproject under the accession number PRJNA738851.

Results

Neochromosome genetic design: does circular or linear configuration matter?

In *S. cerevisiae* the highly efficient homology directed repair machinery can be exploited for the modular assembly of tailored synthetic chromosomes (226). The application of these *de novo* assembled NeoChrs as orthogonal expression platform for pathway engineering requires the fulfilment of several core, size-independent, properties: efficiency and fidelity of assembly, stability in terms of sequence fidelity as well as copy number during replication and segregation, and finally absence of toxicity towards the microbial host. These properties were well met in one particular NeoChr design in strain IMF23 (226), offering a promising proof of principle. The chromosomes in this study however were circular, and mimicking the linear configuration of yeast native chromosomes might further improve these core properties (231, 232). To test this hypothesis, a linear, 100 kb NeoChr was assembled *in vivo* from transcription-unit sized DNA parts and compared to the previously published, circular NeoChr.

The design of the linear test NeoChr in the present study was identical to the 100 kb circular NeoChr design in strain IMF23, consisting of 43 DNA fragments of 2.5 kb (226). To summarize, 36 of the DNA fragments used were non-coding in yeast and originated from *Escherichia coli*, the remaining 7 fragments encompassed selection markers, fluorescent reporters and the elements required for chromosome replication and segregation (**Fig. 2**). Direct assembly of linear chromosomes of this scale has never been attempted before, therefore particular attention was given to the DNA parts framing the chromosomes, containing the telomeric regions. The telomerator containing short telomere seed regions (TeSS) flanking an I-SceI recognition sequence (embedded as intron in a functional *URA3* selection marker), was previously shown to lead to stable linear chromosomes upon *in vivo* digestion (228, 233). The same TeSS were used for the direct assembly of linear chromosomes. This design, composed of 44 fragments and identical TeSS regions on the right and left arm of the NeoChr was named NeoChr10.

To evaluate the risk of unwanted recombination events between the two identical terminal TeSS fragments, which might cause circularization during *in vivo* assembly of all the DNA parts, a second design was tested. In this second design, one of the TeSS was mutagenized to prevent homologous recombination with the original TeSS, leading to NeoChr11. With the exception of the telomerator module (carrying the *URA3* selection marker), the design of the linear chromosomes (NeoChr10 and 11) and circular chromosome (NeoChr12) was identical, which enabled the direct comparison of the core properties between linear and circular chromosomes. The identification of correctly assembled chromosomes was performed

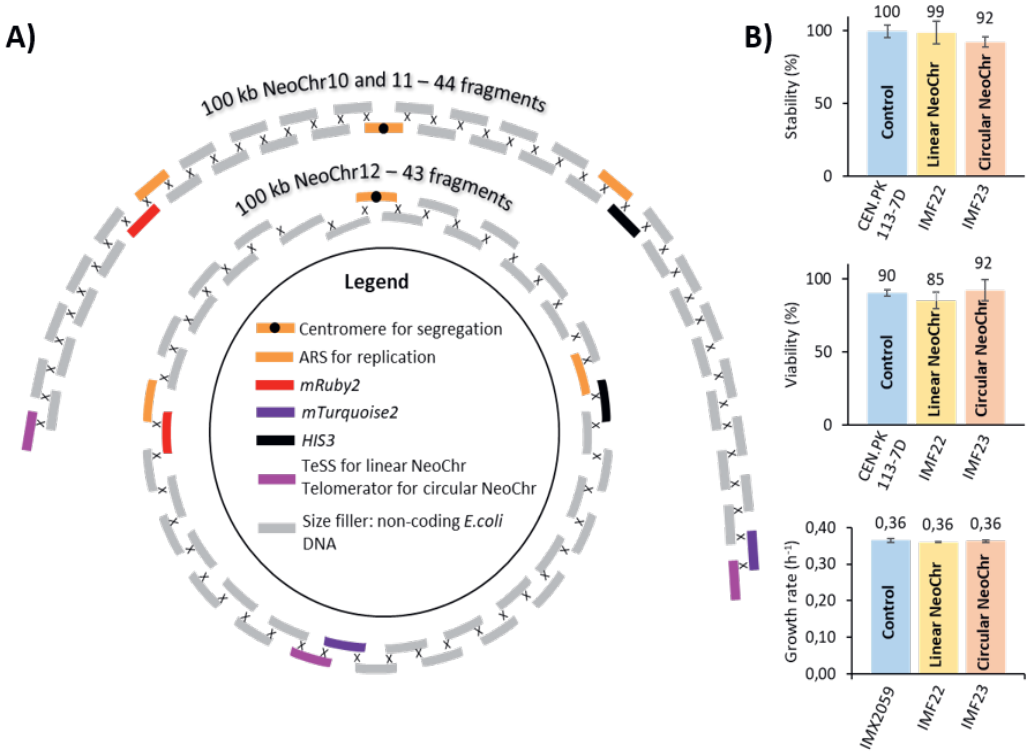


Figure 2. Construction and physiological characterization of test Neochromosomes. **A**) Schematic representation of the *in vivo* assembly design for linear and circular chromosomes. **B**) Physiological characterization of strains with a 100 kb *de novo* assembled linear (IMF22) and circular (IMF23) test NeoChr. **Top graph**, stability of NeoChrs calculated as ratio of the number of colonies on selective medium (with respect to the NeoChr) divided by the number of colonies on non-selective medium. Data represent the average and standard deviation of two days of measurements (day 1 and day 4) of culture duplicates. **Middle graph**, viability of strains, counted as the ratio between the number of colonies growing on non-selective YPD medium and the total number of individually plated cells. Bars represent the average and standard deviation of two days of measurements (day 1 and day 4) of culture duplicates. **Bottom graph**, specific growth rate of strains grown on selective SMD medium. Growth rates represent the average and standard deviation of six biological replicates for IMX2059 and IMF23 and two biological replicates for IMF22. None of the measurements show significant differences between strains (one-way ANOVA with Post-Hoc Tukey-Kramer, $p<0.05$).

by screening for expression of fluorescent markers by FACS, for chromosome size by CHEF and for DNA sequence by whole genome, long-read sequencing (**Table 1** and **Supplementary Fig. 3-4**).

As reported by Postma *et al.* (226), 36% of the transformants with circular chromosomes were true to the *in silico* design. Conversely, irrespective of the telomere sequence used, no linear NeoChrs faithful to the original design were found. Seven transformants of the linear NeoChr10 and one transformant of the NeoChr11, with correct size according to CHEF analysis (**Table 1**) were sequenced by Nanopore technology.

Table 1. NeoChrs assembly efficiency and fidelity. NeoChrs assembly fidelity tested by the screening pipeline. Correct sequence means that all fragments are present in the correct configuration with respect to the *in silico* design.

	Test NeoChrs			**NeoChrs**	
NeoChr	NeoChr10	NeoChr11	NeoChr12	NeoChr25	NeoChr26
Configuration	Linear	Linear	Circular	Linear	Circular
Chromosome size[a] (bp)	99228	99228	99228	99832	99832
Number of fragments	44	44	43	43	42
Auxotrophic markers	*HIS3*	*HIS3*	*HIS3* and *URA3*	*HIS3*	*HIS3* and *URA3*
Number of colonies	72	31	11	391[b]	94[b]
Correct fluorescence	9 (=13%)	7 (=20%)	8 (=72%)	12 (14%)	12 (80%)
Correct size	7 (=10%)	1 (=3%)	5 (=45%)	6 (=7%)[c]	ND[c]
Correct sequence	0 out of 7 screened	0 out of 1 screened	4 (=36%) out of 5 screened	1 out of 4 screened	3 out of 5 screened

[a] Excluding telomeres.
[b] Due to a bacterial infection on plates, the total number of colonies is not reliable. Yeast colonies were subsequently isolated by microscopy based on fluorescence (88 for Neochr25 and 15 for Neochr26).
[c] For NeoChr25 all 12 colonies were screened on CHEF, 6 colonies showed correct size while no band was observed for the other colonies. ND: not determined.

Presence of intact telomeres for all NeoChrs and absence of circularization for both NeoChr10 and 11, demonstrated that the short TeSS supplied were sufficient for the formation of functional telomeres and that homology between the two telomeric fragments was not a hurdle for the direct assembly of linear chromosomes. A wide variety of configurations was observed for these assembled linear NeoChrs (**Supplementary Fig. 4-5**), ranging from the absence of a single fragment (NeoChr10.13 and 10.47), a combination of missing and duplicated fragments (e.g., NeoChr10.62) and more complex configurations with missing, duplicated or inverted fragments and swapped regions (NeoChr10.54 and 11.19). These misassemblies revealed the difficulty encountered by yeast cells to assemble all supplied DNA parts, and demonstrated the intervention of non-homologous end joining, while homology directed repair is typically the preferred mode of DNA double strand break repair in *S. cerevisiae* (234-236). The large impact of linear versus circular configuration, despite otherwise identical design, on *in vivo* chromosome assembly possibly revealed differences in the accessibility of telomerator and TeSS fragments in the nucleus for repair and assembly with the other fragments. Nevertheless, two NeoChrs, NeoChr10.13 and 10.47, displayed a remarkably high degree of fidelity with the *in silico* design, with a single fragment (7A) missing while the remaining 43 fragments were correctly assembled. Overall, keeping in mind that circular chromosomes carried an additional auxotrophic marker as compared to linear chromosomes, these results demonstrated that circular chromosomes are superior in

terms of assembly efficiency and fidelity as compared to linear chromosomes.

The physiology of strains IMF22, carrying the linear NeoChr10.13, and IMF23, harboring a correct circular chromosome NeoChr12, was compared (**Fig. 2B**). Both strains grew as fast as the control strain on minimal, chemically defined medium and displayed the same viability on complex medium. During propagation of the strains over four days (ca. 25 generations), the fraction of the population containing both linear and circular chromosomes (measured as the ratio of colonies on selective over non-selective medium) remained unaltered and similar to the control strain, demonstrating the stability of the NeoChrs during cell division. Moreover, fluorescence analysis and sequencing showed that both NeoChr-configurations were present in a single copy per cell (**Supplementary Fig. 6-7**). To conclude, both linear and circular *de novo* assembled 100 kb NeoChrs were present in one copy per cell, stable and innocuous to their host. As the linear or circular nature of the test NeoChrs did not visibly affect the phenotype of the yeast strains, both configurations were further tested as platforms for metabolic engineering.

In vivo, *de novo* modular assembly of specialized NeoChrs for anthocyanin synthesis

Following the construction strategy as described above, linear and circular NeoChrs, which were expressing the genes native to yeast, bacteria and plants required for P3G production, were designed (**Fig. 3**). The precursors for P3G synthesis natively produced in yeast are L-tyrosine and L-phenylalanine (**Fig. 1**). For *de novo* P3G synthesis from glucose, carbon flows through glycolysis, the pentose phosphate pathway (PPP), cytosolic acetyl-CoA and malonyl-CoA synthesis, the shikimate pathway and aromatic amino acid biosynthesis (**Fig. 1**). The genes encoding enzymes in these pathways are scattered over the 16 yeast chromosomes and many have a high degree of genetic redundancy, making the remodeling of these metabolic routes extremely challenging. Expanding on the pathway swapping concept, the strain construction strategy was built on SwYG, a strain harboring a single locus, minimized glycolysis and fermentation pathway (13 genes involved in the conversion of glucose to ethanol (237)). This strain was further engineered by the genetic reduction of the pentose phosphate pathway (removal of the four minor paralogs *NQM1, GND2, TKL2, SOL4*) with the goal of relocating the genes encoding glycolysis, ethanol fermentation and the PPP from native chromosomes to the specialized NeoChrs (**Fig. 3-4** and **Supplementary Fig. 1**).

The biosynthesis of amino acids is tightly regulated in *S. cerevisiae*, particularly via feedback inhibition (39). Therefore, to improve the supply of tyrosine and phenylalanine for P3G synthesis, the entire pathways for aromatic amino acids synthesis from *E. coli* (10 genes) including key feedback resistant alleles (*coEcaroG*^{fbr} (238), *coEctyrA*^{fbr} (239) and *coEpheA*^{fbr} (240)) were integrated in the design of the specialized NeoChrs. Finally, as aromatic fusel

alcohols and acids are undesired by-products during flavonoid production (132), the 2-oxo acid decarboxylases responsible for their production (Pdc5, Pdc6 and Aro10) had to be removed (82, 83). In the SwYG strain, *PDC5* and *PDC6*, homologues of the major pyruvate decarboxylase encoded by *PDC1*, were already deleted. *ARO10,* that encodes a 2-oxo acid decarboxalyse with broad substrate specificity, was therefore deleted as well from the SwYG strain (**Fig. 4A**).

Figure 3. Genetic design of NeoChr for anthocyanin production. Genetic design of the *in vivo*-assembled linear NeoChr25, NeoChr26, NeoChr30 and NeoChr31. The arrows indicate the directionality of transcription. The names of the genes and auxiliary parts are indicated on the circular chromosome. All genes from *E. coli* were codon-optimized. The linear NeoChr differs from the circular NeoChr by the presence of two TeSS ends instead of the telomerator (carrying the *URA3* selection marker, pink parts). The eleven plant genes required for P3G biosynthesis were integrated into chunk 16AB from NeoChr25 and NoeChr26, resulting in linear NeoChr30 and circular NeoChr31 respectively.

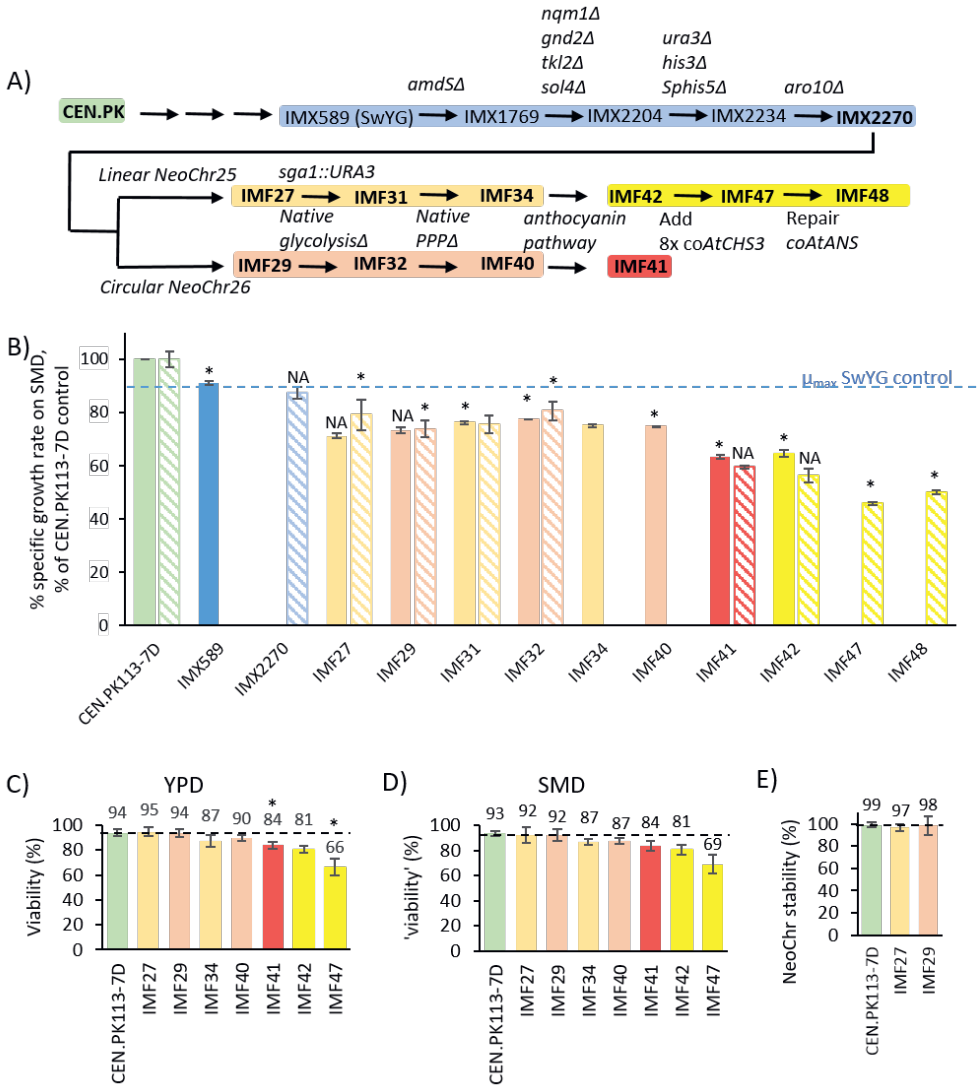


Figure 4. Construction and Physiological characterization of strains with NeoChrs engineered for anthocyanin production. A, overview of the strains constructed in this study for pathway engineering. **B,** Specific growth rate on SMD medium of engineered strains expressed as % of the control CEN.PK113-7D. The same colour-coding has been used as for panel A. Solid-filled bars: average and standard deviations of biological duplicates grown in shake flasks. Pattern-filled bars: average and standard deviations of three to six biological replicates grown in plate reader. For strain IMF27, uracil was supplemented to the medium. For IMX2270 uracil and histidine were supplemented to the medium. **C** and **D**, viability calculated as the number of colonies on YPD (C) or SMD (D) divided by the total number of plated cells (384 FACS-sorted cells). **E,** stability of the NeoChrs measured as number of transformants on selective medium (SMD (supplemented with Uracil)) divided by the number of colonies on non-selective medium (YPD). For panels C, D and E, bars represent the average and standard deviation of biological duplicates measured on two days (day 1 and day 4). Significant differences with respect to the first ancestor in the same graph are indicated with an asterisk (two-tailed paired homoscedastic t-test $P < 0.05$). In Panel B NA indicates data for which statistics could not be calculated because the ancestor was grown in a different set-up than the tested strain (shake-flask or plate reader).

To summarize the NeoChr design for P3G synthesis, the specialized synthetic chromosomes were intended as exclusive expression platforms for the glycolytic, fermentative, pentose phosphate pathways (20 genes), the bacterial shikimate and amino acid biosynthesis pathway (10 genes), and the 11 plant genes involved in the synthesis of P3G from tyrosine and phenylalanine.

The genetic design of the specialized NeoChrs was particularly challenging. This study presents the very first attempt of pathway construction of this magnitude, and information to rationally design an optimal configuration for NeoChr assembly with high efficiency, stability and expression. The first design consideration was high fidelity assembly and maintenance of the Neochrs. While the present study embraces the remarkable ability of *S. cerevisiae* to recombine homologous sequences, homology directed repair (HDR) might cause unwanted recombinations within the NeoChr during assembly and propagation. To prevent unwanted recombinations, homology within the NeoChr was kept to a minimum. However, maintaining a low homology for promoter and terminators regions for 41 transcription units was challenging. The 20 yeast genes were framed by their native promoter and terminator, however the 21 bacterial and plant genes also required terminators and strong constitutive promoters with unique sequences. Such promoters are not abundant in the *S. cerevisiae* molecular toolbox, and most often of glycolytic origin. The heterologous genes were therefore expressed from a selection of previously characterized *S. cerevisiae* promoters (190, 241) which also showed high and condition-independent expression in *S. cerevisiae* from a large transcriptome compendium (204), and promoters from *S. eubayanus* and *S. kudriavzevii* that have little sequence similarity but share functionality with their *S. cerevisiae* relatives (191) (**Supplementary Table 24**). The promoter regions of the transcription units on the NeoChr design did not display homologous sequences longer than 43 bp (*pScADH1*/*pSkADH1* and *pScTDH3*/*pSkTDH3*). Terminators, that generally have a milder impact on gene expression than promoters, especially when paired with strong promoters (242-244), were selected from a set of in-house central carbon metabolism terminators (**Supplementary Table 24**).

The second consideration for the genetic design of the NeoChrs (**Fig. 3**) was the spatial organization of transcriptional units and other elements along the chromosomes, which might impact both maintenance and gene expression of NeoChrs. To facilitate future modular pathway remodeling, genes encoding glycolysis, fermentation, PPP, shikimate and aromatic amino acids pathways were clustered per pathway. To prevent gene silencing resulting from chromatin structure near centromere and telomeres, transcriptional units were separated from these elements by 5 kb non-coding DNA fragments originating from *E. coli*. To facilitate the screening of correctly assembled chromosomes, fluorescent reporters (*mRuby2*, *mTurquoise2*) and auxotrophic markers (*HIS3* and *URA3* for circular and *HIS3* only for linear NeoChrs) were evenly spaced along the NeoChrs. Finally, while highly expressed genes are scattered

across native chromosomes (245), the designed NeoChrs are transcriptional hotspots with high density of heavily transcribed genes. To prevent potential clashes between the replication and transcription machineries, the directionality of transcription units was chosen to follow the same orientation as replication (233).

To test the suitability of NeoChrs for both *de novo* assembly of pathways and as landing pads for pathways a two-steps construction approach was followed. In a first step, 100 kb circular and linear chromosomes carrying all yeast and bacterial genes were assembled *de novo* in yeast (NeoChr26 and 25, respectively) (**Fig. 4A).** In a second step, the genes for anthocyanin synthesis from plant and from *Rhodobacter capsulatus* (*coRcTAL1* only) were simultaneously *de novo* assembled and integrated as a module in the NeoChrs (**Fig. 3**).

De novo assembly of fully coding NeoChrs verified by fluorescence, CHEF and sequencing, confirmed the higher assembly efficiency of circular over linear NeoChrs observed with the test NeoChrs (**Table 1, Supplementary Fig. 8-9** and **Supplementary Table 25-26**). Remarkably, one linear NeoChr was faithful to the *in silico* design. Genome re-sequencing of a transformant with this NeoChr (NeoChr25, strain IMF27) and sequencing of a transformant with a correctly assembled circular NeoChr (NeoChr26, strain IMF29), identified few mutations, with a single non-synonymous mutation in a coding region of NeoChr26 (*RKI1*, **Supplementary Table 25**). Summarizing, the assembly of the linear and circular NeoChrs was successful and few mutations in the native and synthetic chromosomes were caused by the addition of the NeoChrs (**Supplementary Table 25-26**). Next, exclusive expression of glycolysis, fermentation and PPP from the specialized NeoChrs was achieved by Cas9-mediated excision of the single locus glycolysis from ChrIX and multiplex deletion of the seven genes of the minimized PPP (**Fig. 4A**). The targeted deletion of the seven PPP genes scattered over the native chromosomes was facilitated by the insertion of watermarks on their near-identical copy carried by the NeoChrs (229)

As previously observed, genetic relocalization of the glycolytic and fermentative pathways in SwYG caused a mild (IMX589, 10%) reduction in growth rate as compared to the control strain CEN.PK113-7D (**Fig. 4B**, (225)). Additional genetic reduction of the PPP and deletion of *ARO10* (strain IMX2270) had no visible additional phenotypic effect (**Fig. 4B**). Remarkably, NeoChrs with 30 yeast and bacterial genes were as stable during cell division as native chromosomes in control strain CEN.PK113-7D, and strains IMF27 and IMF29 were as viable as the control strain. Their growth rate was slightly decreased (10-16%) with respect to their parental strain IMX2270 **(Fig. 4B**). The linear or circular topology affected neither NeoChr stability and viability nor specific growth rate. Finally, exclusive expression of the yeast glycolytic, fermentative and pentose phosphate pathway from the NeoChrs (strains IMF34 (linear) and IMF40 (circular) did not further decrease viability nor the specific growth

rate, which remained high at ca. 0.30 h^{-1} (**Fig. 4B** and **C**)

Expression of the anthocyanin biosynthetic pathway from a NeoChr

Inspired by earlier work (62), synthesis of P3G was attempted by integration of 11 anthocyanin genes from *Arabidopsis thaliana* (*AtPAL1, coAtCPR1, coAtC4H, At4CL3, coAtCHS3, AtCHI1, coAtF3H, coAt3GT, coAtANS), Gerbera hybrida* (*coGhDFR*) and *Rhodobacter capsulatus (coRcTAL1)*. As basal design, single copies of these 11 genes, interspaced by an ARS sequence, were integrated in the 16AB *E. coli* DNA chunk on the circular and linear NeoChrs (**Fig. 3** and **4A),** leading to NeoChr30 (IMF41) and NeoChr31 (IMF42), respectively. The two strains grew with a similar specific growth rate, but as compared to their parental strain devoid of anthocyanin pathway, they displayed a 15-17% slower specific growth rate (IMF34 and IMF40, *p*-values of 0.01 and 0.00, respectively, **Fig. 4B**), in agreement with previous observations (62). The viability of IMF41 was similar to that of IMF42, but IMF41 was slightly less viable than its ancestor IMF40 (7% lower, *p*-values of 0.03, **Fig. 4C** and **D**). Flow cytometric DNA quantification, short-and long-read sequencing showed that both linear and circular chromosomes were present in a single copy in IMF41 and IMF42 (**Supplementary Fig. 6** and **7**).

In both IMF41 and IMF42, insertion of the plant genes on the NeoChr caused the duplication and inversion of the last four genes in the anthocyanin pathway, *coAtF3H*, *coGhDFR*, *coAtANS* and *coAt3GT*. Sequence analysis revealed that this unexpected event resulted from homologous recombination between the *ScFBA1* promoter (upstream of *ScFBA1*) and the *SeFBA1* promoter (upstream of *coAtC4H*), although slightly different regions of the promoters recombined in IMF41 and IMF42 (**Supplementary Fig. 10**). The two promoters share 58% homology with identical stretches of 24 nucleotides. The absence of homology in the region where the second recombination event occurred suggested the non-homologous recombination of the two *pSePYK1*-SHR CJ ends (**Supplementary Fig. 10**). The presence of the same chromosomal rearrangement in both the linear and circular chromosome, involving rare micro-homology (7-24 bp of 100% homology) and non-homologous recombination events, and the absence of *ARS106*, suggested that the *ARS106* fragment was supplied in suboptimal concentration during transformation. As the same duplication occurred in the strains with linear and circular NeoChr, the impact of NeoChr shape can still be investigated. Plant proteins are typically difficult to express in a yeast environment, a positive rather than negative effect of this duplication on anthocyanin production was therefore anticipated. If desired the duplicated genes can be removed using CRISPR/Cas9 editing.

Short-and long-read sequencing revealed the occurrence of a 46 bp insertion at the beginning of the *coAtANS* gene. The presence of this mutation in the *coAtANS* expression cassette showed that it occurred prior to assembly in the NeoChr, and was not caused by *in vivo* assembly. This

insertion led to an early stop codon, however the presence of a start codon further along the protein still enabled the synthesis of a 277 amino acids long peptide (**Supplementary Fig. 11**). This 80 amino acids shortening from the N-terminus most likely strongly affected the anthocyanidin synthase functionality in IMF41 and IMF42. With the exception of a single non-synonymous mutation in the gene encoding the 4-coumarate:CoA ligase (*At4CL3*, Thr-15-Ala) in IMF42, plant genes were exempt of mutations.

Grown in aerobic shake flask cultures in chemically defined medium, IMF41 and IMF42 produced detectable levels of anthocyanins. The total amount of aromatics produced by the two strains was similar (1.69±0.08 and 1.61±0.06 mM) and anthocyanins down to dihydrokaempferol (DHK) could be reliably detected and quantified (**Fig. 1, Table 2**).

The production of DHK revealed the functional expression of at least eight out of the eleven NeoChr-borne plant genes in *S. cerevisiae*. Considering that metabolites downstream of *At*4Cl3 such as phloretic acid and dihydrokaempferol were detected in both IMF41 and IMF42, the mutation of *At4CL3* had no detectable effect (**Fig. 1, Table 2).**

Over 98% of the produced aromatic products were intermediates (cinnamic and coumaric acid) or byproducts upstream of chalcone synthase (phloretic acid) (**Fig. 1, Table 2**), a result in line with previous reports of chalcone synthase being a bottleneck for anthocyanin production, and a typical target for gene-dosage engineering (62, 132). While the synthesis of the end products P3G and K3G is challenging, detectable levels of other metabolites downstream of dihydrokaempferol were expected (kaempferol, pelargonidin and leucopelargonidin) (62). The absent or barely detectable levels of these metabolites in culture supernatants (**Fig. 1** and **Supplementary Table 2**) confirmed the poor functionality of the truncated anthocyanidin synthase *At*Ans. Overall, these results showed that both the linear and circular structure of the NeoChrs did not affect the physiology of the engineered strains or anthocyanin production, and that the two-step assembly of a NeoChr expressing the *de novo* anthocyanin synthesis pathway enabled anthocyanin production.

Table 2. Determination of the intermediates of the anthocyanin pathway in *S. cerevisiae* strains IMF41 (Cir NeoChr, 1X *coAtCHS3*), IMF42 (Lin NeoChr, 1X *coAtCHS3*), IMF47 (Lin NeoChr, 9X *coAtCHS3*) and IMF48 (Lin NeoChr, 9X *coAtCHS3* repaired *coAtANS*), grown in aerobic shake flask batch cultures on glucose (20 g L^{-1}) and urea. The data represents the average ± standard deviation of independent biological triplicates. Intermediates of the anthocyanin pathway coumaroyl-CoA, naringenin-chalcone, and leucopelargonidin were not measured. * Indicates statistical significance when comparing IMF47 or IMF48 to IMF42, and # when comparing IMF48 to IMF47 (Student *t*-test, two-tailed, homoscedastic, *p*-value threshold 0.05).

(mM)	IMF41	IMF42	IMF47	IMF48
Phenylpyrurate	2.00E-02 ± 0.00E+00	4.33E-02 ± 4.04E-02	BD[a]	BD[a]
2-Phenylethanol	8.67E-02 ± 1.15E-02	8.67E-02 ± 2.52E-02	3.23E-01 ± 3.51E-02*	1.97E-01 ± 2.62E-03*#
***p*-Hydroxyphenylethanol**	2.33E-02 ±2.08E-02	3.33E-02 ±5.77E-03	1.10E-01 ±1.73E-02*	BD[a]
Cinnamic acid	3.80E-01 ±2.00E-02	1.60E-01 ± 2.00E-02	0.00E+00 ± 0.00E+00*	1.46E-01 ± 3.06E-03#
Coumaric acid	4.27E-01 ± 1.15E-02	5.40E-01 ± 2.00E-02	7.13E-01 ± 1.15E-02*	7.64E-01 ± 6.24E-03*#
Phloretic acid	7.25E-01 ± 1.21E-02	7.18E-01 ± 8.72E-03	5.09E-01 ± 1.15E-03*	1.07E+00 ± 7.66E-03*#
Naringenin	BD[a]	BD[a]	BD[a]	BD[a]
Dihydrokaempferol	2.25E-02 ± 2.57E-03	2.84E-02 ± 1.27E-03	9.49E-02 ± 4.29E-03*	3.83E-01 ± 8.34E-02*#
Kaempferol	6.33E-04 ± 3.44E-05	6.96E-04 ± 3.46E-05	2.14E-03 ± 3.06E-04*	1.39E-02 ± 5.72E-03*#
Pelargonidin	2.45E-03 ± 1.60E-04	1.81E-03 ± 1.98E-04	5.76E-03 ± 4.11E-04*	3.65E-02 ± 7.64E-03*#
Kaempferol 3-O-glucoside	2.24E-04 ± 1.83E-05	2.24E-04 ± 5.98E-06	4.39E-04 ± 2.41E-05*	5.53E-03 ± 1.13E-03*#
Pelargonidin 3-O-glucoside	BD[a]	BD[a]	BD[a]	2.14E-05 ± 3.93E-06*#
Total aromatics before *At*Chs	1.66 ± 0.06	1.58 ± 0.06	1.66 ± 0.05	1.98 ± 0.01*#
Total anthocyanins after *At*Chs	0.03 ± 0.00	0.03 ± 0.00	0.1 ± 0.0*	0.45 ± 0.10*#
Total aromatics	1.69 ± 0.06	1.61 ± 0.08	1.76 ± 0.05	2.43 ± 0.11*#

Neochromosome engineering for improved de novo anthocyanin biosynthesis

The two engineering targets identified to improve anthocyanin synthesis were implemented by CRISPR/Cas9 editing of the NeoChrs. Targeting the linear NeoChr31 (strain IMF42), four additional *coAtCHS3* copies were integrated in four different loci on the NeoChr spaced by essential genes (in Chunk 7BC, chunk 15CD, SHR N and chunk 9CD, **Fig 3**). In an attempt to further boost anthocyanin synthesis, four additional *coAtCHS3* copies were integrated in native chromosomes (at *CAN1, X2, YPRCtau3* and *SPR3* loci (192, 246-248), resulting in a total of nine *coAtCHS3* copies (IMF47). Finally, the coAtANS gene was repaired leading to strain IMF48. The structure of the NeoChr was not affected by these additional genetic interventions.

The performance of IMF48, IMF42 and IMF41 was compared in pH-controlled, aerated bioreactors, a culture tool previously shown to enable higher anthocyanin titers than shake flasks (62). The strains were grown in minimal, chemically defined medium without additional growth supplements (*e.g.*, yeast extract or peptone), to purely evaluate *de novo* production of anthocyanins from glucose. The physiology of the three strains was similar, despite the eight additional *coAtCHS3* copies carried by IMF48 (**Table 3**). Irrespective of the strain, flavonoids were produced both during exponential growth on glucose and after the diauxic shift (**Fig. 5**).

Table 3. Physiological characterization of anthocyanin-producing strains grown in bioreactors. A) The specific growth rate (μ) and the yield (Y) of biomass (X) and ethanol (ETOH) on glucose (S). **B)** The overall yield (Y) of glycerol (GLYC), pyruvate (PYR), coumaric acid (COUM), phloretic acid (PHLOR) and dihydrokaempferol (DHK) on glucose and ethanol (S) during aerobic bioreactor batch cultivation of IMF41 (Cir NeoChr, 1x *coAtCHS3*), IMF42 (Lin NeoChr, 1x *coAtCHS3*),) and IMF48 (Lin NeoChr, 9x *coAtCHS3*, repaired *coAtANS*).

A)	[a]μ_{MAX} (h^{-1})	[a]$Y_{X/S}$ ($g\ g^{-1}$)	[a]$Y_{ETOH/S}$ ($mol\ mol^{-1}$)
IMF41 (Cir)	0.23 ± 0.00	0.12 ± 0.00	1.44 ± 0.05
IMF42 (Lin)	0.22 ± 0.01	0.12 ± 0.00	1.41 ± 0.02
IMF48 (Lin, 9x *coAtCHS3,* repaired *coAtANS*)	0.20 ± 0.01	0.13 ± 0.00	1.29 ± 0.05

B)	$Y_{GLYC/S}$ ($mol\ mol^{-1}$)	$Y_{PYR/S}$ ($mol\ mol^{-1}$)	$Y_{X/S}$ ($mol\ mol^{-1}$)	$Y_{coum/S}$ ($\mu mol\ mol^{-1}$)	$Y_{PHLOR/S}$ ($\mu mol\ mol^{-1}$)	$Y_{DHK/S}$ ($\mu mol\ mol^{-1}$)
IMF41 (Cir)	0.056 ± 0.005	0.029 ± 0.001	0.29 ± 0.01	7.30 ± 0.08	4.71 ± 0.20	0.28 ± 0.01
IMF42 (Lin)	0.048 ± 0.002	0.029 ± 0.001	0.30 ± 0.01	6.70 ± 0.12	5.02 ± 0.15	0.28 ± 0.00
IMF48 (Lin, 9x *coAtCHS3* repaired *coAtANS*)	0.052 ± 0.003	0.019 ± 0.003	0.27 ± 0.01	20.2 ± 0.59	5.89 ± 0.35	2.85 ± 0.02

[a] Determined for the glucose phase only

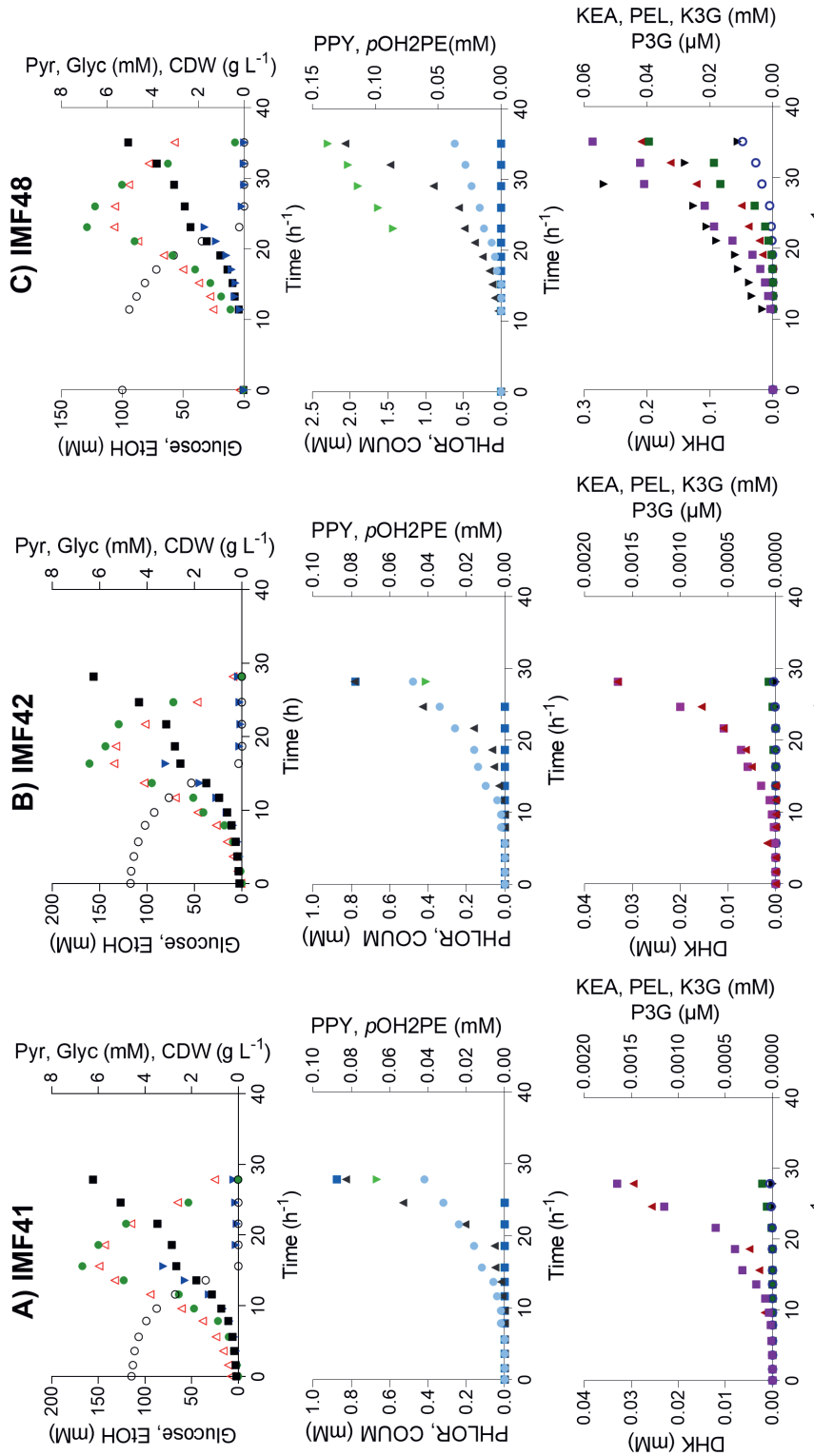


Figure 5. Substrates and products profiles during aerobic batch cultivation in bioreactors of IMF41, IMF42 and IMF48. In columns **A)** IMF41 (Cir, 1x *coAtCHS3*), **B)** IMF42 (Lin, 1x *coAtCHS3*), and **C)** IMF48 (Lin, 9x *coAtCHS3*, *coAtANS*), were grown in biological duplicates at 30°C in aerobic batch cultures in bioreactors, in chemically defined medium with 20 g L^{-1} glucose as sole carbon source (SMD). One representative cultivation is shown. The duplicate is depicted in **Supplementary Fig. 12. Row 1)** ■ CDW (g L^{-1}), ○ Glucose (mM), ● EtOH (mM), ▼ Pyr (mM), △ Glyc (mM) **Row 2)** ■ PPY (mM), ▲ COUM (mM), ● PHLOR (mM), ▼ *p*OH2PE (mM) **Row 3)** ■ DHK (mM), ▲ PEL (mM), ■ KAE (mM) ○ K3G (mM) ▼ P3G (µM).

Table 4. Characterization of anthocyanin production in bioreactors. Determination of the intermediates of the anthocyanin pathway in *S. cerevisiae* strains IMF41, IMF42 and IMF48. Extracellular and intracellular metabolite concentration at the end of aerobic and pH controlled (pH 5.0) bioreactor batch cultures on glucose. The concentrations (mM) of metabolites upstream from, and including, naringenin were measured by HPLC while metabolites downstream were quantified by LC-MS/MS. The data represents the average ± standard deviation of independent biological duplicates. Intermediates of the anthocyanin pathway coumaroyl-CoA, naringenin chalcone, and leucopelargonidin were not measured. An asterisk (*) indicates statistical significance when comparing extracellular concentrations of IMF48 to concentrations of IMF41 and IMF42 (Student *t*-test, *p*-value threshold 0.05, two-tailed, homoscedastic).

Compound	**IMF41**		**IMF42**		**IMF48**	
	Extracellular (mM)	Intracellular (µmol/g dry weights)	Extracellular (mM)	Intracellular (µmol/g dry weights)	Extracellular (mM)	Intracellular (µmol/g dry weights)
Phenylpyruvate	6.68E-02 ± 0.00E+00	NM	4.47E-02 ± 4.95E-03	NM	BD	NM
2-Phenylethanol	BD	NM	BD	NM	BD	NM
p-Hydroxy-phenylethanol	8.94E-02 ± 2.55E-03	NM	7.87E-02 ± 9.90E-04	NM	1.40E-01 ± 2.83E-03*	NM
Cinnamic acid	BD	NM	BD	NM	BD	NM
Coumaric acid	8.46E-01 ± 2.38E-02	NM	7.89E-01 ± 6.22E-03	NM	2.05E+00 ± 1.98E-02*	NM
Phloretic acid	5.46E-01 ± 3.24E-02	NM	5.92E-01 ± 1.15E-02	NM	5.98E-01 ± 2.42E-02	NM
Naringenin	BD	NM	BD	NM	BD	NM
Dihydrokaempferol	3.29E-02 ± 1.64E-04	7.26E-02 ± 4.44E-02	3.32E-02 ± 2.01E-04	1.41E-01 ±1.93E-02	2.90E-01 ± 3.98E-03*	1.15E+00 ±2.00E-01
Kaempferol	1.04E-04 ± 8.51E-06	1.20E-03 ± 8.84E-04	8.61E-05 ± 1.64E-05	1.58E-03 ± 6.99E-04	3.55E-02 ± 5.63E-03*	3.64E+00 ± 1.18E+00
Pelargonidin	1.46E-03 ±1.99E-05	6.25E-03 ± 3.96E-03	1.64E-03 ±2.73E-05	1.00E-02 ± 3.26E-03	4.14E-02 ±5.63E-04*	4.96E-01 ± 7.76E-02
Kaempferol -O-glucoside	2.74E-05 ± 5.38E-07	3.42E-05 ± 2.88E-05	2.17E-05 ± 6.21E-07	1.18E-04 ±8.03E-05	9.77E-03 ± 1.09E-04*	1.44E-02 ± 1.70E-03
Pelargonidin 3-O-glucoside	BD	BD	BD	BD	1.45E-05 ± 4.78E-06*	2.51E-04 ± 1.95E-05
Total aromatics (a)	1.47 ± 0.04	-	1.14 ± 0.03	-	3.3 ± 0.23*	-
Total flavonoids (b)	0.03 ± 0.002	-	0.03 ± 0.002	-	0.51 ± 0.19*	-
Fraction of flavonoids (=b/a)	2%	-	2%	-	15%*	-

Confirming the results from shake flask cultures, the structure of the NeoChr in IMF41 and IMF42 affected neither yeast physiology (**Table 3**) nor flavonoid formation (**Table 4, Fig. 5**). Overall, strain IMF48 outperformed IMF41 and IMF42 carrying a single *coAtCHS3* copy and a truncated *coAtANS* (**Table 3-4, Fig. 5**). IMF48 produced 14-fold more flavonoids downstream of chalcone synthase than IMF41 and IMF42, and the production of pelargonidin, kaempferol and K3G, metabolites downstream of *At*Ans, was increased by ca. 30-, 350- and 500-fold, respectively in IMF48 as compared to IMF41 and IMF42 (**Table 4**). P3G, not detected intracellularly nor extracellularly in IMF41 and IMF42 cultures, was produced with a titer of 0.049 ± 0.007 μM by IMF48 at the end of the glucose consumption phase (**Fig. 5C**). Notably, this is the first report of exclusive extracellular P3G production by yeast cultures.

The concentration of P3G decreased by ca. three-fold at the end of the diauxic phase, presumably due to degradation by periplasmic β-glucosidases (e.g., Exg1), upon glucose exhaustion (**Fig. 5**). These enzymes are interesting targets for further improvement of P3G production. IMF48 accumulated more coumaric acid and p-hydroxyphenylethanol, aromatic intermediates upstream of *At*Chs, an unexpected response to the increased chalcone synthase expression that suggested some yet unidentified regulation of the plant enzymes in the yeast environment (**Table 4**). No extracellular naringenin was detected for any of the tested strains. The absence of naringenin but increased intra- and extracellular dihydroxykaempferol concentration with increasing *coAtCHS3* copy number and flux through the anthocyanin pathway, showed that flavonoid hydroxylase (F3H) was most probably not limiting P3G production and was therefore not a target for genetic engineering at that stage. This apparent overcapacity of flavonoid hydroxylase might be related to its presence in two copies on the NeoChrs. Conversely, the strong increase in flavonoid concentration upstream of the last enzyme, 3GT (kaempfarol and pelargonidin, **Table 4**), suggested that despite the duplication of its gene, this enzyme might be limiting for K3G and P3G production and a target for further improvement. The titers of produced flavonoids were modest. However, the fraction of flavonoids over all aromatic compounds produced by IMF48 was 15%, representing a substantial improvement compared to the 2% produced by IMF41 and IMF42. Additionally, the total production of aromatic compounds was increased by 2.3-fold in IMF48.

The combined co*AtCHS3* copy number increase and *coAtANS* repair markedly affected the flux distribution in the aromatic pathway. To estimate the respective contribution of chalcone and anthocyanidin synthase to these changes, the parental strain of IMF48, a strain carrying nine *coAtCHS3* copies, but with impaired *coAtANS* was tested (IMF47, **Fig. 4A, Table 3.**). In line with increased chalcone synthase activity, IMF47 produced three times more flavonoids than IMF42, while the production of metabolites upstream of chalcone synthase was unchanged. However, P3G was still not detected in IMF47 cultures. The fraction of flavonoids over

total aromatics was 6% in IMF47 (**Table 3**). Repair of *coAtANS* in IMF48 further increased flavonoids production by ca. four-fold and slightly, but significantly, increased total aromatics production as compared to IMF47 (1.4-fold, **Table 3**). Altogether these data revealed that both increased *coAtCHS3* copy number and *coAtANS* repair contributed to the improvement of anthocyanin production in IMF48.

Discussion

The present study illustrates the amazing potential of *S. cerevisiae* for fast and extensive genome remodeling via synthetic chromosome engineering. Supernumerary, specialized chromosomes could be easily and rapidly assembled, in a single transformation round, from 30 transcription units and 13 accessory DNA parts. Beyond this technical 'tour de force', the NeoChrs resembled native chromosomes in terms of replication and segregation. The stable maintenance of the NeoChrs at one copy number and their harmlessness to the host are important features for their implementation as metabolic engineering platforms. Plant-derived chemicals have a broad range of biotechnological applications, but their *de novo* microbial production requires remodeling of the host metabolism, as well as functional transplantation of the plant pathway. Equipped with pathway swapping, NeoChrs enabled the relative facile implementation of three complex interventions: i) the remodeling of native metabolic networks (glycolysis and PPP), ii) the provision of an optimized metabolic route from prokaryotic origin (shikimate pathway) as surrogate for a native route, and iii) the implementation of a new, heterologous pathway from plant origin (anthocyanin synthesis). These modifications readily enabled the synthesis of anthocyanins from glucose, and the first report of exclusive extracellular P3G production in yeast or any other microbial host. The implementation of NeoChrs not only accelerated strain construction, and thereby genome remodeling, but also prevented interferences with native chromosomes. While chromosome construction was carefully designed in two steps to probe the limits of *in vivo* NeoChr assembly, the present data suggest that single step chromosome assembly would have readily enabled anthocyanin production in a shorter time frame. Once assembled, the NeoChrs could be edited using CRISPR/Cas9. CRISPR/Cas9 editing efficiency is locus-dependent, and a small set of robust genomic integration sites have been validated to date (188, 246, 248). Carefully designed, mostly non-coding NeoChrs, harboring strategically located, optimized CRISPR/Cas9 programmed sites, could become ideal landing pads for large sets of (heterologous) genes. Finally, the successful assembly of chromosomes from 43 parts suggests that the limit of *in vivo* assembly has not been reached and even larger chromosomes can be assembled if required.

Synthetic Genomics is a young research field (249, 250), and the design principles for optimal, tailor-made NeoChrs are ill-defined. The synthetic chromosomes rebuilt in the Sc2.0 initiative uses native chromosomes as scaffolds and yet reproduces the native organization of the chromosomes, albeit omitting non-essential elements and adding some short sequences (251, 252). With the possibility to construct and test any genetic design, *de novo*-assembled NeoChrs are fantastic testbeds to explore the genetic and physiological impact of chromosome sequence and structure. In this study, some design guidelines were formulated and tested, leading to stable chromosomes, easy to screen and with functionally expressed transcription units. The

100-130 kb linear and circular NeoChrs in this study showed equal growth rate and mitotic stability. This is in good agreement with earlier studies showing approximate equal stability of linear and circular chromosomes of 100 kb, while beneath this size circular chromosomes appear more stable and above this size linear chromosomes are more stable (233, 253-256). A remarkable observation was the low assembly efficiency of linear chromosomes as compared to their circular counterparts. As opposed to transcription units, telomeres and centromeres have specific, cell-cycle dependent localizations within the nucleus (257-260). It is conceivable that, for linear chromosomes equipped with one centromere and two telomeres, accessibility and spatial organization of these DNA parts conflict with homologous recombination, a difficulty potentially alleviated for circular chromosomes equipped with a single centromere and no telomeres. Additionally, the frequent occurrence of non-homologous repair in the linear NeoChrs, an otherwise rare mechanism active throughout the entire cell cycle (236), might indicate temporal incompatibility between telomere and centromere availability and homologous recombination, mostly active during the S/G_2 phase (236). For biotechnological applications, circular chromosomes, easier to isolate from yeast, are therefore recommended, and can be equipped with a telomerator to enable ulterior linearization if required (228, 233).

In their current design, the NeoChrs are extremely information-dense, with short intergenic regions and a concatenation of highly transcribed genes. While this genetic design leads to functional expression of the NeoChr-born genes and is not harmful for the host cells, many fundamental questions regarding optimal genetic design remain to be systematically explored, such as the cell's tolerance to 'transcriptional hotspots' (225, 226, 261), the impact of transcription units localization, orientation and distancing on gene expression, or the requirement for multiple selection markers for chromosome stability. Furthermore, while NeoChrs harboring spaced homologous sequences were genetically stable, the presence of homology between DNA parts might lead to unwanted recombination events during *in vivo* assembly of the NeoChrs. In this work the DNA parts were designed as 'homology-free' as possible, a design principle difficult to apply considering the limited availability of strong, constitutive promoters. Homology between DNA parts could be kept to a minimum thanks to the implementation of promoters from *S. cerevisiae* relatives (191). Nevertheless, the highest homology between promoters was still 75% (with stretches of identical sequences up to 42 nucleotides (*pSkTDH3* and *pTDH3*)). Along this line the *S. cerevisiae* toolbox could be further enriched by mining non-*Saccharomyces* yeasts genetic diversity for functional but sequence divergent promoters (191, 262). In a more distant future, progress in the design of synthetic promoters should enable the construction of libraries of *in silico*-designed, artificial promoters with minimal homology (263, 264)

From the present study, we can envisage future microbial cell factories with modular genomes in which core metabolic network and processes, localized on satellite, specialized NeoChrs

can be swapped for alternative configurations. Following the work on 32% reduction of the 111 genes of yeast central carbon metabolism (265), the present strategy for *in vivo* assembly of NeoChrs can be applied to construct yeast strains carrying specialized NeoChrs as exclusive expression platforms for central carbon metabolism. Combined with pathway swapping (225), such strains would enable fast and easy remodeling of large sets of core cellular functions. As yeast is tolerant to chromosome ploidy variation (266), other strategically designed NeoChr could carry other industrially-relevant pathways (e.g. nitrogen or fatty acids metabolism) and processes (e.g. protein secretion). Additional NeoChrs, similar to the test NeoChrs from the present study, but tailored for CRISPR/Cas9 targeting, could also serve as landing pads dedicated to the addition of functionalities.

Financial support

This project was funded by the European Union's Horizon 2020 research and innovation program under grant agreements No. 720824 (CHASSY Project) and No. 648141 (AdLibYeast European Research Council (ERC)).

Declaration of interest

None

Acknowledgments

We thank Carol de Ram for technical support for the analysis of metabolites by LC-MS.

Chapter 4

Elimination of aromatic fusel alcohols as by-products of *Saccharomyces cerevisiae* strains engineered for phenylpropanoid production by 2-oxo-acid decarboxylase replacement

Else-Jasmijn Hassing, Joran Buijs, Nikki Blankerts, Marijke A. Luttik, Erik A. de Hulster, Jack T. Pronk, Jean-Marc Daran

Essentially as published in
Metabolic Engineering Communications
2021. 13, e00183
Supplementary materials available online:
https://doi.org/10.1016/j.mec.2021.e00183

Abstract

Engineered strains of the yeast *Saccharomyces cerevisiae* are intensively studied as production platforms for aromatic compounds such as hydroxycinnamic acids, stilbenoid and flavonoid. Heterologous pathways for production of these compounds use L-phenylalanine and/or L-tyrosine, generated by the yeast shikimate pathway, as aromatic precursors. The Ehrlich pathway can convert these precursors to aromatic fusel alcohols and acids, which are undesirable by-products of yeast strains engineered for production of high-value aromatic compounds. Activity of the Ehrlich pathway requires either of four *S. cerevisiae* 2-oxo-acid decarboxylases (2-OADCs): Aro10 and the pyruvate-decarboxylase isoenzymes Pdc1, Pdc5, and Pdc6. Elimination of pyruvate-decarboxylase activity from *S. cerevisiae* is not straightforward as it plays a key role in cytosolic acetyl-CoA biosynthesis during growth on glucose. In a search for pyruvate decarboxylases that do not decarboxylate aromatic 2-oxo acids, eleven yeast and bacterial 2-OADC-encoding genes were investigated. Orthologs from *Kluyveromyces lactis* (*KlPDC1*), *Kluyveromyces marxianus* (*KmPDC1*), *Yarrowia lipolytica* (*YlPDC1*), *Zymomonas mobilis* (*Zmpdc1*) and *Gluconoacetobacter diazotrophicus* (*Gdpdc1.2* and *Gdpdc1.3*) complemented a Pdc$^-$ strain of *S. cerevisiae* for growth on glucose. Enzyme-activity assays in cell extracts showed that these genes encoded active pyruvate decarboxylases with different substrate specificities. In these *in vitro* assays, *Zm*Pdc1, *Gd*Pdc1.2 or *Gd*Pdc1.3 did not show activity with phenylpyruvate. Replacing Aro10 and Pdc1,5,6 by these bacterial decarboxylases completely eliminated aromatic fusel-alcohol production in glucose-grown batch cultures of an engineered coumaric acid producing *S. cerevisiae* strain. These results outline a strategy to prevent formation of an important class of by-products in 'chassis' yeast strains for production of non-native aromatic compounds.

Abbreviations: 2-OADC 2-oxo acid decarboxylase, *p*OHPPY *p*-hydroxyphenylpyruvate, *p*OH2PE *p*-hydroxyphenylethanol, *p*OHPAA *p*-hydroxyphenylacetic acid, CIN cinnamic acid, COUM coumaric acid, 2-OADC 2-oxo acid decarboxylase. E4P erythose-4-phosphate, PEP phosphoenolpyruvate, DAHP 3-deoxy-D-arabino-heptulosonate-7-phosphate, CHOR chorismate, PPA prephenate, PPY phenylpyruvate, PHE L-phenylalanine, PAC phenylacetaldehyde, 2PE 2-phenylethanol, *p*OHPPY *p*-hydroxyphenylpyruvate, TYR L-tyrosine, *p*OHPAC *p*-hydroxyphenylacetaldehyde, *p*OH2PE *p*-hydroxyphenylethanol, PYR pyruvate, ACH acetaldehyde, ACE acetate, ETOH ethanol, ACCOA acetyl-CoA.

Introduction

The aromatic amino acids L-phenylalanine and L-tyrosine are precursors of many industrially relevant compounds belonging to the phenylpropanoid family of aromatic compounds (66), including hydroxycinnamic acids (152), stilbenoids (162) and flavonoids (267). These compounds have diverse applications in the food, chemical, pharmaceutical and cosmetic industries (66). Current production processes mostly depend on petroleum-based chemical processes (268) or direct extraction from plants (269). However, the chemical processes involved are often inefficient and unsustainable (101, 102, 270) while plant extraction processes are limited by biomass availability, low extraction yields and low purity of the final products (98, 99). To overcome these pitfalls and meet the increasing demand for biologically and renewably sourced aroma and flavour compounds, microbial production from renewable feedstocks offers a promising alternative (269).

Development of microbial platforms for *de novo* production of aromatic compounds has been intensively studied in the yeasts *Saccharomyces cerevisiae* and *Yarrowia lipolytica* (107). The yeast shikimate pathway for aromatic amino-acid biosynthesis, a focal point in these metabolic engineering studies, is initiated by condensation of phospho-enol-pyruvate (PEP) and erythrose-4-phosphate (E4P) to form 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP). This seven-carbon intermediate is converted into chorismate via a series of biochemical reactions (**Fig. 1**). At chorismate, the pathway for L-tryptophan synthesis branches off. For biosynthesis of L-phenylalanine and L-tyrosine, chorismate is converted to prephenate, from which either phenylpyruvate or *p*-hydroxyphenylpyruvate are formed. Transamination of these two aromatic 2-oxo acids then yields L-phenylalanine and L-tyrosine, respectively.

In addition to high-level functional expression of heterologous pathway enzymes (134, 142, 271), elimination of allosteric feed-back inhibition of the shikimate-pathway enzymes DHAP synthase (Aro3 and Aro4) and chorismate mutase (Aro7) (39, 91, 92, 94, 272, 273), increasing the capacity of the shikimate pathway (40, 62, 98, 132-134, 136, 142, 155, 271) and improving supply of its precursors PEP and E4P (134, 142, 155) have enabled increased titers and yields of phenylpropanoid in *S. cerevisiae*. However, these metabolic engineering strategies also lead to increased formation of aromatic fusel alcohols (2-phenylethanol, *p*-hydroxyphenylethanol) (62, 132, 134) and fusel acids (phenylacetic acid, *p*-hydroxyphenylacetic acid) (155). During production of high-value phenylpropanoid such as hydroxycinnamic acid, stilbenoid and flavonoid, formation of these undesired by-products represents a drain of precursors and goes at the expense of product titers and yields.

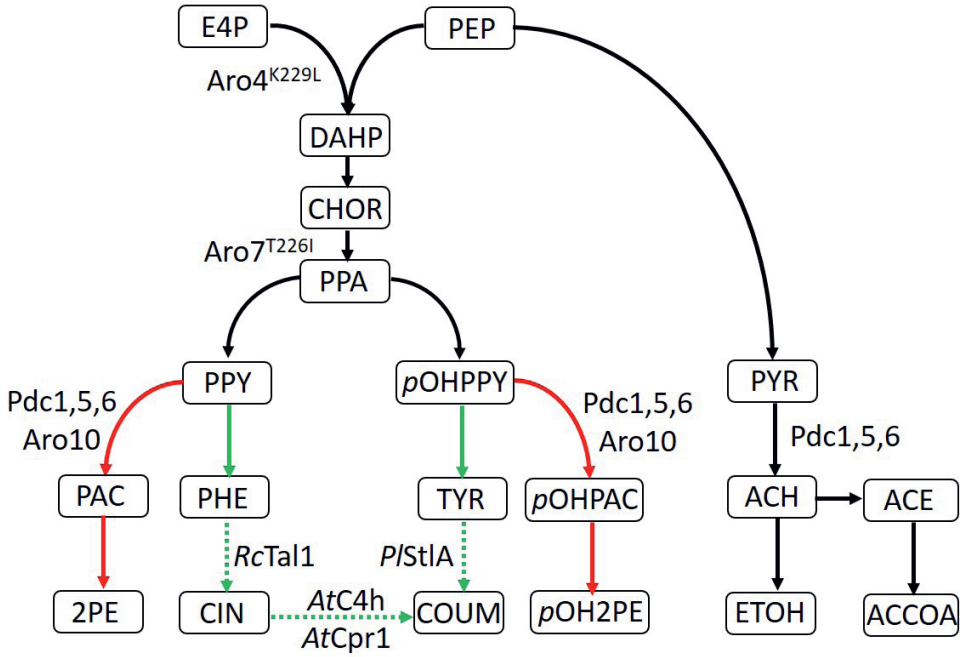


Figure 1. Metabolic pathways involved in production of the fusel alcohols 2-phenylethanol and *p*-hydroxyphenylethanol and their relation to coumaric acid production in *S. cerevisiae*. The 2-oxo acids phenylpyruvate and *p*-hydroxyphenylpyruvate can either be decarboxylated into the fusel aldehydes phenylacetaldehyde or *p*-hydroxyphenylacetaldehyde, respectively, or be transaminated into L-phenylalanine or L-tyrosine, respectively. These two aromatic amino acids can both serve as substrate in the phenylpropanoid pathway (indicated with dotted arrows), synthesizing compounds such as cinnamic acid and coumaric acid. In order to make more phenylpropanoids, it would be beneficial to eliminate the decarboxylation step which results in the formation of by-products in the form of the fusel alcohols 2-phenylethanol and *p*-hydroxyphenylethanol. However, not all genes encoding for the enzymes (Pdc1, Pdc5, Pdc6, Aro10) catalyzing this decarboxylation step can be deleted as these enzymes are also responsible for the decarboxylation of pyruvate resulting in acetaldehyde, which in turn is converted into acetyl-CoA, an important metabolite. The aim of this study is to eliminate the formation of fusel alcohols (indicated with red arrows), thereby increasing the flux towards phenylpropanoid biosynthesis (indicated with green arrows) while retaining the ability of *S. cerevisiae* to decarboxylate pyruvate into acetaldehyde. E4P erythose-4-phosphate, PEP phosphoenolpyruvate, DAHP 3-deoxy-D-arabino-heptulosonate-7-phosphate, CHOR chorismate, PPA prephenate, PPY phenylpyruvate, PHE L-phenylalanine, PAC phenylacetaldehyde, 2PE 2-phenylethanol, *p*OHPPY *p*-hydroxyphenylpyruvate, TYR L-tyrosine, *p*OHPAC *p*-hydroxyphenylacetaldehyde, *p*OH2PE *p*-hydroxyphenylethanol, PYR pyruvate, ACH acetaldehyde, ACE acetate, ETOH ethanol, ACCOA acetyl-CoA.

Fusel alcohols and acids are formed via the Ehrlich pathway for degradation of branched-chain, aromatic, and sulfur-containing amino acids (75, 76). In the Ehrlich pathway, transamination of amino acids yields the corresponding 2-oxo acids, which are subsequently decarboxylated. The resulting aldehyde is then either oxidised or reduced by yeast aldehyde dehydrogenases and alcohol dehydrogenases to yield fusel acids and fusel alcohols, respectively (76). The irreversible decarboxylation reaction in the Ehrlich pathway is catalysed by thiamine-pyrophosphate-dependent 2-oxo acid decarboxylases (2-OADC), which in *S. cerevisiae* are

encoded by *PDC1*, *PDC5*, *PDC6* and *ARO10*. Pdc1, Pdc5 and Pdc6 show a preference for the linear-chain 2-oxo acids pyruvate, 2-oxobutyrate and 2-oxopentanoate (82), while Aro10 shows no activity with linear chain 2-oxo acids, but does decarboxylate branched-chain and aromatic 2-oxo acids at high rates (81-83). Aro10 is a main contributor to 2-phenylethanol production by *S. cerevisiae* (82, 83, 227), but Pdc5 also shows a distinct activity with phenylpyruvate (82). Strains expressing heterologous pathways for flavonoid production from which *ARO10, PDC5* and *PDC6* were deleted, still exhibited residual 2-phenylethanol formation, indicating that Pdc1 still decarboxylated 2-phenylpyruvate at low rates (132).

Deletion of the three pyruvate-decarboxylase genes *PDC1*, *PDC5*, *PDC6* renders *S. cerevisiae* unable to grow on glucose in batch cultures. During aerobic glucose-limited growth, Pdc$^-$ strains are auxotrophic for C_2-compounds (ethanol or acetate) due to involvement of pyruvate decarboxylases in cytosolic acetyl-CoA synthesis (166, 167). Although metabolic engineering strategies have been published to bypass this acetyl-CoA requirement (reviewed by Van Rossum *et al.* (43)) and to mitigate the glucose sensitivity of Pdc- strains (274), deletion of all four 2-OADC genes in S. cerevisiae is not a straightforward strategy to eliminate formation of aromatic fusel alcohols and acids. Still, complete elimination of these by-products, without negative impacts on growth on glucose or a need to rewire central carbon metabolism, would be an attractive attribute of *S. cerevisiae* 'chassis' strains for production of phenylpropanoid.

The goal of the present study was to identify heterologous pyruvate decarboxylases that show activity with pyruvate but not with aromatic 2-oxo acids and to investigate whether they can functionally replace the native yeast enzymes and thereby prevent formation of aromatic Ehrlich-pathway products. To this end, a set of 11 2-oxo acid decarboxylases from yeasts (*Kluyveromyces lactis, K. marxianus* and *Y. lipolytica*) and bacteria (*Zymomonas mobilis* and G*luconacetobacter diazotrophicus*) were expressed in a 2-OADC-deficient *S. cerevisiae* strain. Enzyme assays with cell extracts of the resulting strains were used to assess substrate specificities and affinities of each of the decarboxylases for pyruvate and phenylpyruvate. Subsequenty, they were used to replace the native 2-OADC in a *S. cerevisiae* strain engineered for the production of the phenylpropanoid compound, coumaric acid. The impact of replacing the native yeast 2-OADC with the two best-performing heterologous pyruvate-decarboxylases was evaluated in aerobic, pH-controlled bioreactor cultures.

Materials and Methods

Strains and growth media.

S. cerevisiae strains used in this study were derived from the CEN.PK lineage (182, 275) (**Table 1)**. *Escherichia coli* XL1 blue (Agilent Technologies, Santa Clara, CA) was used for plasmid propagation and storage. *S. cerevisiae* and *E. coli* strains were stored at -80 °C as described previously (188). Complex YP (yeast extract/peptone) medium was prepared and sterilized as described previously (188) and, when required, was supplemented with 200 mg L^{-1} hygromycin (InvivoGen, San Diego, CA). As a carbon source, YP medium was supplemented with either 20 g L^{-1} glucose (YPD) or a mixture of 2% (v/v) glycerol and 2% (v/v) ethanol (YPEG) (198). Chemically defined synthetic medium (SM) containing mineral salts, trace elements and vitamins was prepared and autoclaved as described by Verduyn *et al.* (183). SM was supplemented with either 20 g L^{-1} glucose (SMD), 2% (v/v) ethanol (SME) or a mixture of 2% (v/v) glycerol and 2% (v/v) ethanol (SMEG) (198). When required, 150 mg L^{-1} uracil (Sigma-Aldrich, St Louis, MO) (184) was supplied to synthetic media. Lysogeny broth (LB) for growth of *E. coli* strains was prepared as described by Bertani *et al.* (276) and supplied with 25 mg L^{-1} chloramphenicol (Sigma-Aldrich), 100 mg L^{-1} ampicillin (Sigma-Aldrich) or 50 mg L^{-1} kanamycin (Sigma-Aldrich) as required. Solid media were prepared by adding 20 g L^{-1} Bacto Peptone (Brunschwig Chemie BV, Amsterdam, Netherlands) prior to autoclaving.

Table 1. Strains used in this study. Abbreviations: *Sc Saccharomyces cerevisiae, Km Kluyveromyces marxianus, Kl Kluyveromyces lactis, Yl Yarrowia lipolytica, Gd Gluconacetobacter diazotrophicus, Zm Zymomonas mobilis, Rc Rhodobacter capsulatus, Pl Photorhabdus luminescens*, Co codon optimized, pr promoter, t terminator, pdc↓ *pdc5Δ pdc6Δ aro10Δ,* Pdc- *pdc1Δ pdc5Δ pdc6Δ aro10Δ*, 3ABP aromatic amino acid biosynthetic pathway, COUM coumaric acid, FBR feedback resistant, 2μm multicopy.

Strain	Description	Genotype	Ref
CEN. PK113-7D	Reference	MATa *MAL2-8c SUC2*	(182)
CEN. PK711-7C	Pdc- *thi3Δ*	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10:: loxP-kan-loxP thi3::loxP-kan-loxP*	(81)
IMZ001	CEN.PK711-7C p426GPD	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10:: loxP-kan-loxP thi3::loxP-kan-loxP* p426GPD (*URA3*)	(81)
IMZ002	CEN.PK711-7C pUDE001	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10:: loxP-kan-loxP thi3::loxP-kan-loxP* pUDE001 (*URA3 TDH3pr-ScARO10-CYC1t*)	(81)
IMZ024	CEN.PK711-7C pUDE002	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10:: loxP-kan-loxP thi3::loxP-kan-loxP* pUDE002 (*URA3 TDH3pr-ScPDC5-CYC1t*)	(82)
IMZ031	CEN.PK711-7C pEXP214-PDC6.2	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10:: loxP-kan-loxP thi3::loxP-kan-loxP* pEXP214-PDC6.2 (*URA3 PGK1pr-ScPDC6-CYC1t*)	(82)
IMX1593	3ABP^{FBR}	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t*	(227)
IMX1758	3ABP^{FBR} pdc↓ pUDR406	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t pdc5Δ pcd6Δ aro10Δ pUR406* (*URA3,* gRNA-*PDC5/PDC6, ARO10*)	This study
IMX1789	3ABP^{FBR} pdc↓	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t pdc5Δ pcd6Δ aro10Δ*	This study
IMX2656	3ABP^{FBR} pdc↓ COUM	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-CoPlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc5Δ pcd6Δ aro10Δ*	This study
IMX2668	3ABP^{FBR} Pdc- COUM	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-CoPlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ*	This study
IME418	CEN.PK711-7C pUDE833	MATa *MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE833 (*URA3, TDH3pr-KlPDC5-CYC1t*)	This study
IME419	CEN.PK711-7C pUDE837	MATa *MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE837 (*URA3 TDH3pr-YlPDC1-CYC1t*)	This study
IME420	CEN.PK711-7C pUDE838	MATa MAL2-8c SUC2 *ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE838 (*URA3 TDH3pr-KmPDC1-CYC1t*)	This study
IME421	CEN.PK711-7C pUDE827	MATa *MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE827 (*URA3 TDH3pr-CoZmpdc1-CYC1t*)	This study
IME422	CEN.PK711-7C pUDE829	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE829 (*URA3 TDH3pr-KmPDC5-CYC1t*)	This study
IME423	CEN.PK711-7C pUDE828	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE828 (*URA3 TDH3pr-KmARO10-CYC1t*)	This study
IME424	CEN.PK711-7C pUDE834	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE834 (*URA3 TDH3pr-KlARO10-CYC1t*)	This study

Table 1 (continued)

IME425	CEN.PK711-7C pUDE832	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE832 (*URA3 TDH3pr-CoGdpdc1.1-CYC1t*)	This study
IME474	CEN.PK711-7C pUDE881	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE881 (*URA3 TDH3pr-CoGdpdc1.2-CYC1t*)	This study
IME495	CEN.PK711-7C pUDE882	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE882 (*URA3 TDH3pr-CoGdpdc1.3-CYC1t*)	This study
IME615	CENPK711-7C pUDE1037	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE1037 (*URA3 TDH3pr-KlPDC1-CYC1t*)	This study
IME656	IMX2668 pGGKd017	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ pdc1Δ* pGGKd017 (*URA3*)	This study
IME658	IMX2668 pUDE827	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE827 (*URA3 TDH3pr-CoZmpdc1-CYC1t*)	This study
IME659	IMX2668 pUDE838	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE838 (*URA3 TDH3pr-KmPDC1-CYC1t*)	This study
IME660	IMX2668 pUDE837	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE837 (*URA3 TDH3pr-YlPDC1-CYC1t*)	This study
IME661	IMX2668 pUDE1037	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE1037 (*URA3 TDH3pr-Klpdc1-CYC1t*)	This study
IME662	IMX2668 pUDE881	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE881 (*URA3 TDH3pr-CoGdpdc1.2-CYC1t*)	This study
IME663	IMX2668 pUDE882	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-*Co*PlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE882 (*URA3 TDH3pr-CoGdpdc1.3-CYC1t*)	This study
IME667	CENPK711-7C pUDE1099	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE1099 (*URA3 TDH3pr-ScPDC1-CYC1t*)	This study
IME668	CENPK711-7C pUDE1101	*MATa MAL2-8c SUC2 ura3 pdc1Δ pdc5Δ pcd6Δ aro10Δ thi3Δ* pUDE1101 (*URA3 TDH3pr-ScPDC1-CYC1t ENO2pr-ScTHI3-GPM1t*)	This study

Table 1 (continued)			
IME677	IMX2668 pUDE1099	*MATa ura3-52 spr3Δ::Spcas9*-natNT2 *aro3Δ aro7Δ TDH3p-ARO4^{K229L}-ARO4t ARO7pr::SeGPM1pr-ARO7^{T226I}-TEF1t x3::SePDC1pr-CoPlstlA-ENO2t shrB TPI1pr-CoAtC4H-ADH1t PGI1or-CoAtCPR1-PGI1t shrC SkTDH3pr-CoRctal1-ADH1t pdc1Δ pdc5Δ pcd6Δ aro10Δ* pUDE1099 (*URA3 TDH3pr-ScPDC1-CYC1t*)	This study

Table 2. Plasmids used in this study. Abbreviations: *Sc Saccharomyces cerevisiae, Km Kluyveromyces marxianus, Kl Kluyveromyces lactis, Yl Yarrowia lipolytica, Gd Gluconacetobacter diazotrophicus, Zm Zymomonas mobilis, Rc Rhodobacter capsulatus; Pl Photorhabdus luminescens*, Co codon optimized, pr promoter, t terminator, DO dropout.

Part Plasmids			
Name	**Description**	**Part Type**	**Source**
JA_NM_1 Sc_CoPlstlA	*camR CoPlstlA*	3	This study
pYTK001	*camR GFP* entry vector	Insert	(190)
pYTK002	*camR* ConLS connector	1	(190)
pYTK003	*camR* ConLS' connector	1	(190)
pYTK008	*camR* ConL1 connector	1	(190)
pYTK047	*camR GFP* DO	234r	(190)
pYTK055	*camR ENO2t*	5	(190)
pYTK067	*camR* ConR1 connector	5	(190)
pYTK072	*camR* ConRE connector	5	(190)
pYTK073	*camR* ConRE' connector	5	(190)
pYTK074	*camR URA3*	6	(190)
pYTK082	camR 2*μ*m	7	(190)
pYTK084	*camR kanR-ColE1*	8	(190)
pUD565	*camR GFP* entry vector	Insert	(191)
pGGKp025	*camR PDC1pr*	2	(227)
pGGKp027	*camR FBA1pr*	2	This study
pGGKp028	*camR ENO2pr*	2	(227)
pGGKp035	*camR TDH3pr*	2	(227)
pGGKp037	*camR ADH1t*	4	(227)
pGGKp039	*camR TEF1t*	4	(227)
pGGKp045	*camR PDC1t*	4	(227)
pGGKp048	*camR GMP1t*	4	(227)
pGGKp063	*camR SkTDH3p*	2	(227)
pGGKp074	*camR SePDC1p*	2	(227)
pGGKp182	*camR CYC1t*	4	This study
pGGKp183	*camR KmARO10*	3	This study

Table 2 (continued)

pGGKp184	*camR KmPDC1*	3	This study
pGGKp185	*camR* putative *KmPDC5*	3	This study
pGGKp211	*camR CoZmpdc1*	3	This study
pGGKp212	*camR CoGdpdc1.1*	3	This study
pGGKp213	*camR KlPDC5*	3	This study
pGGKp214	*camR KlARO10*	3	This study
pGGKp254	*camR CoGdpdc1.2*	3	This study
pGGKp255	*camR CoGdpdc1.3*	3	This study
pGGKp314	*camR KlPDC1*	3	This study
pGGKp315	*camR ScTHI3*	3	This study
pGGKp327	*CoRctal1*	3	This study
pGGKp337	*ScPDC1*	3	This study

Expression cassettes

Plasmid	Genotype	Parts used	Source
pGGKd017	2*μ*m *ampR-ColE1 URA3 GFP* DO	pYTK002, pYTK047, pYTK72, pYTK074, pYTK082, pYTK083	(277)
pGGKd071	2*μ*m *ampR-ColE1 conLS conR1 URA3 GFP* DO	pYTK002, pYTK047, pYTK67, pYTK074, pYTK082, pYTK083	This study
pGGKd072	2*μ*m *ampR-ColE1 conL1 conRS URA3 GFP* DO	pYTK003, pYTK047, pYTK72, pYTK074, pYTK082, pYTK083	This study
pGGKd073	2*μ*m *kanR-ColE1 conLS' conRS' URA3 GFP* DO	pYTK008, pYTK047, YTK073, pYTK074, pYTK082, pYTK084	This study
pUDE827	2*μ*m *ampR URA3 TDH3pr-CoZmpdc1-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp211	This study
pUDE828	2*μ*m *ampR URA3 TDH3pr-KmARO10-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp183	This study
pUDE829	2*μ*m *ampR URA3 TDH3pr-KmPDC5-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp185	This study
pUDE832	2*μ*m *ampR URA3 TDH3pr-CoGdpdc1.1-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp212	This study
pUDE833	2*μ*m *ampR URA3 TDH3pr-KlPDC5-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp213	This study
pUDE834	2*μ*m *ampR URA3 TDH3pr-KlARO10-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp214	This study
pUDE837	2*μ*m *ampR URA3 TDH3pr-YlPDC1-CYC1t*	Gibson assembly, pGGKd017, pGGKd035, pGGKp182	This study
pUDE838	2*μ*m *ampR URA3 TDH3pr-KmPDC1-CYC1t*	pGGKd017, pGGKp035, pGGKp182 pGGKp184	This study
pUDE881	2*μ*m *ampR URA3 TDH3pr -CoGdpdc1.2-CYC1t*	pGGKd017, pGGKp035, pGGKp254, pGGKp182	This study
pUDE882	2*μ*m *ampR URA3 TDH3pr-CoGdpdc1.3-CYC1t*	pGGKd017, pGGKp035, pGGKp255, pGGKp182	This study

Table 2 (continued)

pUDE1019	2μm *ampR URA3 SePDC1pr-CoPlstlA-ENO2t*	pGGKd017, pGGKp074, pYTK055, JA_NM 1_ Sc_ *coPlstIA*	This study
pUDE1037	2μm *ampR URA3 TDH3p-KlPDC1-CYC1t*	pGGKd017, pGGKp035, pGGKp314, pGGKp182	This study
pUDE1049	2μm *ampR URA3 ENO2pr-ScTHI3-GPM1t*	pGGkd072, pGGKp028, pGGKp315, pGGKp048	This study
pUDE1088	2μm *ampR URA3 SkTDH3pr-CoRctal1-CYC1t*	pGGkd017, pGGKp037, pGGKp063, pGGKp327	This study
pUDE1099	2μm *ampR URA3 TDH3pr-ScPDC1-CYC1t*	pGGKd017, pGGKp035, pGGKp182, pGGKp337	This study
pUDE1100	2μm *ampR URA3 TDH3pr-ScPDC1-CYC1t*	pGGKd071, pGGKp035, pGGKp182, pGGKp337	This study
pUDE1101	2μm *kanR URA3 TDH3pr-ScPDC1-CYC1t, ENO2pr-ScTHI3-GPM1t*	pGGd073, pUDE1049, pUDE1100	This study

Cas9 Plasmids

Name	Relevant Genotype	Primer(s) used for gRNA	Source
pROS10	2μm *bla URA3* gRNA-*CAN1*.Y gRNA-*ADE2.Y*	N.A.	(188)
pROS12	2μm *bla* hphNT1 gRNA-*CAN1*.Y gRNA-*ADE2.Y*	N.A.	(188)
pUDR406	2μm *bla URA3* gRNA-*PDC5/PDC6* gRNA-*ARO10*	7246 & 13614	This Work
pUDR470	2μm *bla* hphNT1 gRNA-*PDC1* (2x)	6178	This Work
pUDR599	2μm *bla* hphNT1 gRNA-*X3* (2x)	15832	(227)

Miscellaneous Plasmids

Name	Relevant Genotype	Source
pUDE172	Centromeric plasmid, *URA3 TDH3pr-AtPAL1-CYC1t TPIp-CoAtC4H-ADHt PGIpr-CoAtCPR1-PGIt*	(132)
pUDI069	Integration plasmid, *TRP1 TDH3pr-CoRctal1-CYC1t*	(132)
pE_MGV14	2μm *TDH3pr* natNT2 *bla HXT7p ,-CoPlstlA-CYC1t*	(132)

Molecular Biology Techniques

DNA templates for cloning were amplified with Phusion high-fidelity polymerase (Thermo Fisher Scientific, Landsmeer, Netherlands) according to manufacturer's protocol, with the exception that a primer concentration of 200 nM and 0.04 U μL^{-1} of polymerase were used. The YeaStar genomic DNA kit (Zymo Research, Irvine, CA) was used to isolate genomic DNA as template for PCR amplification. The Zymoclean kit (Zymo Research) was used to purify PCR products by gel purification according to manufacturer's recommendations using milliQ water as eluent. Alternatively, PCR products were first incubated for 1 h with DpnI FastDigest enzyme (Thermo Fisher Scientific) to digest template DNA and subsequently purified using the GenElute™ PCR Clean-Up Kit (Sigma-Aldrich). Diagnostic PCR was performed with

DreamTaq PCR mastermix (Thermo Fisher Scientific) and with oligonucleotide primers shown in **Supplementary Table 1.** The GenElute plasmid miniprep kit (Sigma-Aldrich) was used to isolate plasmids from *E. coli*.

Construction of plasmids and expression cassettes

Plasmids used and constructed in this study are shown in **Table 2**. Constructed plasmids were transformed to *E. coli* (XL1-Blue) cells according to the supplier's recommendations and grown under selective conditions.

Plasmids containing gRNAs for Cas9-mediated genome editing were constructed as described by Mans et al. (188). The resulting gRNA plasmids pUDR406 (gRNA-P*DC5/PDC6* and gRNA-*ARO10*), pUDR470 (gRNA-*PDC1* 2x) and pUDR599 (gRNA-*X3* 2x) (227) were used to target *PDC5, PDC6, PDC1, ARO10* and the *X3*-locus (248), respectively.

The expression cassettes used in this study were constructed using the Yeast Toolkit (190). In brief, promoter, gene and terminator fragments (parts) are amplified with part type specific overhangs containing restriction sites (BsmBI and BsaI). Using Golden Gate assembly with the corresponding restriction enzyme, BsmBI, the individual parts are initially assembled in a universal entry vector, resulting in a part plasmid. Next, a promoter, gene and terminator part plasmid are assembled into an expression cassette using BsaI-mediated golden gate assembly, resulting in an expression cassette containing a transcriptional unit.

Initially, DNA fragments carrying the *CYC1* terminator (*CYC1t*) and *FBA1* promoter *(FBA1pr*) fragments were amplified from genomic DNA from *S. cerevisiae* CEN.PK113-7D with oligonucleotide primers 14039/14040 and 9419/9420 adding terminator (ATCC and CAGC) or promoter (AACG and ATAC) part type specific overhangs (**Supplementary Table 1**) (190). Open reading frames of genes from *S. cerevisiae* CEN.PK113-7D (182), *K. marxianus* NRBC 1777 (NITE Biological Resource Center, Japan) (278) or *K. lactis* CBS 2359 (279) were amplified from genomic DNA using primers with gene-part type specific overhangs (TATG and GGAT) (**Supplementary Table 1**)**.** Primers ASR_A023F/ASR_A023Rcorr, 13940/13941 and 16851/16852 and 17630/17631 were used to amplify *KmPDC1, KlARO10, ScPDC1* and *ScTHI3* respectively. The ORFs of *KlPDC5*, *KlPDC1, KmPDC5 and KmARO10* were amplified in several fragments to remove internal BsaI and BsmBI sites from the coding sequence (227) using primer pairs 13932/13933 and 13934/13935 for *KlPDC5,* 13939/14138, 14137/13938 and 13937/13936 for *KlPDC1,* ASR_A024F/ASR_A024MR and ASR_A024MF/AR_A024Rcorr for *KmPDC5* and ASR_A022F/ASR_A022MR and ASR_A022MF/ASR_A022Rcorr for *KmARO10*. *CoRctal1* was amplified using pUDI069 (132) as template with primers 17825/17826. Correct removal of the internal BsaI/BsmBI sites of *KmPDC5, KmARO10, KlPDC5* and *KlPDC1* was confirmed by Sanger sequencing (BaseClear, Leiden,

Netherlands). A codon-optimized, based on yeast glycolytic codon usage (280), open reading frame of the phenylalanine ammonia lyase gene *CoPlstLA* from *Photorhabdus luminescens* was amplified from plasmid MGV14 (*CoPlstLA*) (202) with primers ASR_N009F/ASR_N009R. Codon optimisation of the *Gdpdc1* and *Zmpdc1* coding regions was performed using the Jcat Codon Adaptation Tool (281). The codon regions were custom synthesized by Invitrogen GeneArt (Thermo Fisher Scientific) service. The sequence of *Zmpdc1* was derived from the annotated genome of strain *Z. mobilis subsp. mobilis* ATCC 10988 (Bioproject accession number PRJNA30987) (282). Since three different sequences of *Gdpdc1* have been reported for *G. diazotrophicus* strain ATCC 49037, codon-optimized coding sequences for *Gdpdc1.1* (283), *Gdpdc1.2* (284) and *Gdpdc1.3* (285). Coding sequences were flanked upstream and downstream with the gene specific Yeast Toolkit flanks 'AAGCATCGTCTCATCGGTCTCAT' and 'TTATGCCGTCTCAGGTCTCAGGAT' respectively (190).

The amplified and synthesised fragments of *CYC1t, KmARO10, KmPDC1, KmPDC5* and *CoPlstLA* were cloned into entry vector pYTK001 (190), via BsmBI Golden Gate assembly, obtaining part plasmids pGGKp182 (*CYC1t),* pGGKp183 (*KmARO10*), pGGKp184 (*KmPDC1*), pGGKp185 (*KmPDC5*) and JA_NM 1_Sc_coPlstlA (*CoPlstLA). FBA1pr, Zmpdc1, Gdpdc1.1-3, KlPDC1*, *KlARO10, ScPDC1, KlPDC5, CoRctal1* and *ScTHI3* were also assembled via BsmBI Golden Gate assembly but into entry vector pUD565 (191), resulting in part plasmids pGGKp027 (*FBA1pr*), pGGKp211 (*Zmpdc1),* pGGKp212 (*Gdpdc1.1*), pGGKp213 (*KlPDC5*), pGGKp214 (*KlARO10*), pGGKp254 (*Gdpdc1.2*), pGGKp255 (*Gdpdc1.3*), pGGKp314 (*KlPDC1*), pGGKp315 (*ScTHI3*), pGGKp327 (*CoRctal1*) and pGGKp337 (*ScPDC1*). Part plasmid were confirmed by colony PCR using primers 2012 and 2397 for the pUD565 entry vector and with primers 14036 and 14977 for YTK001 entry vector.

The GFP dropout plasmid pGGKd017 (*URA3*$^{\uparrow}$) (277) was used as backbone to construct expression cassettes expressing a single 2-oxo acid decarboxylase. As example, the Golden Gate assembly of pGGkp035 (*TDH3pr)*, pGGkp182 (*CYC1t*) and pGGKp211 (*Zmpdc1*) using pGGKd017 as a backbone resulted in the construction of pUDE827 (*URA3, TDH3pr-Zmpdc1-CYC1t*). A full overview of all part plasmids that were used to construct the expression cassettes is presented in **Table 2**. Correct construction was verified by diagnostic PCR and restriction analysis.

Additionally, a multi-expression cassette plasmid was constructed expressing *ScPDC1* and *ScTHI3.* For this purpose, three additional GFP dropout plasmids were first constructed. The part plasmids pYTK002 and pYTK067 (ConLS and ConR1 connectors), pYTK047 (GFP dropout), pYTK074 (*URA3*), pYTK082 (2μm,) with pYTK083 (ColE1 *bla*) were assembled via BsmBI Golden Gate assembly (190) resulting in pGGKd071 (multigene cassette #1). Additionally, pYTK003 and pYTK072 (ConL1 and ConRE connectors), pYTK047,

pYTK074, pYTK082 were assembled via BsmBI Golden Gate assembly (190) resulting in pGGKd072 (multigene cassette #2). Finally, pGGKd073, a GFP multigene dropout plasmid, was constructed by assembling pYTK008 and pYTK073 (ConLS' and ConRE' connectors), pYTK047, pYTK074, pYTK082 with pYTK084 (ColE1 nptII).

After this, again using BsaI mediated golden gate assembly, pGGkd072 (multigene cassette #2), pGGKp028 (*ENO2pr)*, pGGKp315 (*ScTHI3*) and pGGKp048 (*GPM1t*) were assembled resulting in pUDE1049 (*ScTHI3*, multigene cassette #2). Next, pGGkd071 (multigene cassette #1), pGGKp035 (*TDH3pr*), pGGKp337 (*ScPDC1*) and pGGKp182 (*CYC1t*) were assembled resulting in pUDE1100 (*ScPDC1* multigene cassette #1). Finally, pGGKd073 (multigene dropout), pUDE1049 and pUDE1100 were assembled using a BsmBI golden gate assembly into pUDE1101 (*ScPDC1, ScTHI3*) (Table 2). Final plasmid confirmation was done by restriction analysis.

The expression cassette bearing *YlPDC1* was assembled using Gibson assembly. The gene *YlPDC1* (YALI0D10131g, Genome Resources for Yeast Chromosomes database (https://gryc.inra.fr)) was PCR amplified from genomic DNA of *Y. lipolytica* W29 (277, 286) using primers 14187/14188. The *TDH3pr* and *CYC1t* were amplified from pGGKp035 and pGGKp182 using primers 14185/14186 and 14189/14190, respectively. The linear pGGKd017 backbone was amplified using primers 14183/14184. The plasmid pUDE837 (*YlPDC1*) was constructed using Gibson assembly of the promoter, gene and terminator fragments. Correct construction of pUDE837 was confirmed by restriction analysis.

Strain construction

The thiamine-pyrophosphate-dependent-decarboxylase-negative strain *S. cerevisiae* CENPK711-7C (*ura3Δ pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ*) (81) was transformed with 2-oxo acid decarboxylase-expressing episomal plasmids resulting in strains (pUDE833 (*KlPDC5*$^{\uparrow}$)), IME419 (pUDE837 (*YlPDC1*$^{\uparrow}$)), IME420 (pUDE838 (*KmPDC1*$^{\uparrow}$)), IME421 (pUDE827 (*Zmpdc1*$^{\uparrow}$)), IME422 (pUDE829 (*KmPDC5*$^{\uparrow}$)), IME423 (pUDE828 (*KmARO10*$^{\uparrow}$)), IME424 (pUDE834(*KlARO10*$^{\uparrow}$)),IME425(pUDE832(*Gdpdc1.1*$^{\uparrow}$)),IME474(pUDE881(*Gdpdc1.2*$^{\uparrow}$)), IME495 (pUDE882 (*Gdpdc1.3*$^{\uparrow}$)), IME615 (pUDE1037 (*KlPDC1*$^{\uparrow}$)), IME667 (pUDE1099 (*ScPDC1*$^{\uparrow}$)) and IME668 (pUDE1101 (*ScPDC1*$^{\uparrow}$ *ScTHI3*$^{\uparrow}$).

S. cerevisiae IMX1593 (*ura3Δ Spcas9 aro3Δ aro7Δ ARO4*$^{K229L\uparrow}$ *ARO7*$^{T226I\uparrow}$) (227) was used as starting point for construction of a coumaric acid producing strain. Transformation with pUDR406 (gRNA-*PDC5/PDC6, ARO10*) and repair fragments consisting of oligonucleotides 7247/7248 for *ARO10*, 7717/7718 for *PDC5* and 7935/7936 for *PDC6,* yielded strain IMX1758. After curing of pUDR406 (188) strain IMX1789 was obtained, into which expression cassettes for coumaric acid biosynthesis were integrated. Cassettes for expression of *CoRctal1,*

CoPlstlA and *CoAtC4H/CoAtCPR1* were amplified using primer pairs 12044/18181, 12040/18183 and 4640/18180 and plasmids pUDE1088, pUDE1019 and pUDE172 (132), respectively, as templates. Strain IMX1789 (*pdc5Δ pdc6Δ aro10Δ*) was co-transformed with pUDR599 (gRNA-*X3*) and the three expression cassettes *CoRctal1, CoPlstlA* and C*oAtC4H/ CoAtCPR1* containing homologous flanks to the *X3* locus or a short homologous sequence (shr) (193) to allow homologous recombination of the flanks and integration into the edited *X3* locus (248) resulting in strain IMX2656 after curing pUDR599. In the final step, *PDC1* was deleted by co-transforming strain IMX2656 (coumaric acid producing *pdc5Δ pdc6Δ aro10Δ*) with pUDR470 (gRNA-*PDC1*) and a repair fragment consisting of annealed oligonucleotides 7719 and 7720, resulting in strain IMX2668 after curing the gRNA plasmid. The resulting strain IMX2668 (coumaric acid producing, *pdc1Δ pdc5Δ pdc6Δ aro10Δ*) was transformed with episomal plasmids expressing an individual 2-oxo-acid decarboxylase or with pGGKd017, an empty backbone plasmid, as negative control. This yielded strains IME656 (pGGKd017 (*URA3*$^{\uparrow}$)), IME658 (pUDE827 (*Zmpdc1*$^{\uparrow}$)), IME659 (pUDE838 (*KmPDC1*$^{\uparrow}$)), IME660 (pUDE837 (*YlPDC1*$^{\uparrow}$)), IME661 (pUDE1037 (*KlPDC1*$^{\uparrow}$)), IME662 (pUDE881 (*Gdpdc1.2*$^{\uparrow}$)), IME663 (pUDE882 (*Gdpdc1.3*$^{\uparrow}$)) and IME677 (pUDE1099 (*ScPDC1*$^{\uparrow}$)).

Growth studies

Shake-flask cultures were grown in 500 mL shake flasks containing 100 mL medium and incubated at 30 ˚C in an Innova incubator shaker (New Brunswick Scientific, Edison, NJ). Precultures on SMEG were inoculated from frozen stock cultures. These precultures were used to inoculate shake flasks containing SMEG and SMD, at an initial OD_{660} of 0.2. Independent duplicate cultures were grown for each combination of yeast strain and medium composition. Specific growth rates were calculated from a minimum number of six data points collected during exponential growth and covering 3-4 doublings of OD_{660}. Ehrlich pathway products were quantified in supernatant samples of triplicate stationary phase (72 h) shake-flask cultures.

Aerobic bioreactor batch cultures on SMD supplemented with 0.2 g L^{-1} antifoam C (Sigma-Aldrich) were grown in 2L bioreactors (Applikon, Delft, Netherlands) with a working volume of 1.0 L. Oxygen was supplied by continuously sparging the culture with pressurized air at 0.5 L min^{-1}. Exponentially growing shake-flask cultures on SMD were used to inoculate the bioreactors at an initial biomass concentration of around 0.1 g L^{-1}. Cultures were grown at 30 °C and stirred at 800 rpm with a Rushton impeller. The culture pH was maintained at 5.0 by automated addition of 2 M KOH or 2 M H_2SO_4. Optical density at 660 nm was measured with a Jenway 7200 spectrophotometer (Jenway, Staffordshire, United Kingdom). Biomass dry weight was measured as described previously (230). Off-gas from the bioreactors was

cooled using a condenser and dried using a Permapure MD-110-48P-4 dryer (Permapure, Lakewood, NJ). CO_2 and O_2 concentrations in the off-gas were measured with a NGA 2000 Rosemount gas analyser (Rosemount, Analytical, Irvine, CA).

Concentrations of glucose, ethanol and extracellular organic acids in culture supernatants were measured by high performance liquid chromatography (HPLC) as described before (227). The Ehrlich pathway metabolites 2-phenylethanol, *p*-hydroxyphenylethanol, phenylacetate, phenylpyruvate, coumaric acid and cinnamic acid were also measured by HPLC as described before (227). Aromatic compounds were detected by a diode-array multiple-wavelength detector (Agilent G1315C), at wavelengths of 200 nm for phenylacetate, 210 nm for phenylpyruvate, 214 nm for 2-phenylethanol and *p*-hydroxyphenylethanol, 270 nm for cinnamic acid and 280 nm for coumaric acid.

Enzyme-activity assays in cell extracts

Cell extracts of *S. cerevisiae* strains were prepared from late exponential phase (OD_{660} of approximately 8) shake-flask cultures grown on SMEG or SME medium. After 10 min centrifugation at 4696 x *g*, cell pellets were washed twice with 20 mL 10 mM potassium phosphate buffer (pH 7.5) containing 2 mM EDTA, resuspended in 4 mL buffer and stored at -20 °C. Prior to the enzyme assays, biomass samples were thawed, resuspended and washed with 100 mM potassium phosphate buffer (pH 7.5) containing 2 mM $MgCl_2$ and 2 mM dithiothreitol. When cell extracts were prepared for experiments to estimate kinetic parameters, which took several hours, complete (TM), Mini Protease Inhibitor Co. (Sigma-Aldrich) was added as protease inhibitor according to manufacturer's recommendations. Cell extracts were prepared by sonication with 0.7 mm diameter glass beads using a MSE sonicator (150-W output, 7-nm peak-to-peak amplitude) at 0 **°C**. After four bursts of 30 s with 30 s cooling intervals, debris was removed by centrifugation using a Sorvall SS34-rotor (Thermo Fisher Scientific) for 20 min at 47.000 x*g* operated at 4 **°C**. The clear supernatants were used as cell extracts and kept on ice during experiments. Pyruvate-decarboxylase activity in cell extracts was measured as described previously (287). Phenylpyruvate-decarboxylase activity was assayed essentially as described before (83) but with 5 mM instead of 2 mM phenylpyruvate. K_m values for pyruvate were obtained by measuring pyruvate-decarboxylase activities at concentrations ranging from 0.1 mM to 50 mM, followed by nonlinear regression of the obtained results with GraphPad Prism (version 9.02, GraphPad Software, San Diego, CA). Datasets were fitted with Michaelis-Menten as well as allosteric sigmoidal kinetics.

Protein homology and phylogenetic Tree

The amino acid sequences (**Supplementary Part 1**) of the 2-oxo acid decarboxylases compared in this study were aligned using Clustal Ω (288). A heat map displaying sequence similarity

was generated using GraphPad Prism. A phylogenetic tree of the aligned protein sequences was constructed with SeaView5 (289) applying the LG model (290) with default parameter settings using 100 Bootstrap replicates as support level for internal branches.

Results

Selection of heterologous pyruvate-decarboxylases with a potentially narrow substrate specificity

Heterologous pyruvate-decarboxylases with a potentially better substrate selectivity for pyruvate were selected based on three criteria: i) homology with *S. cerevisiae* pyruvate decarboxylases, ii) a demonstrated or proposed role in pyruvate decarboxylation, iii) absence of evidence for activity with aromatic 2-oxo acids. A resulting set of 11 decarboxylases was selected comprised of *Kl*Pdc1 (291), *Kl*Pdc5 (292) and *Kl*Aro10 (291) from K. *lactis*, *Km*Pdc1, *Km*Pdc5 and *Km*Aro10 from *K. marxianus* (293), *Yl*Pdc1 from *Y. lipolytica* (291) and four bacterial pyruvate decarboxylases: *Zm*Pdc1 from *Z. mobilis* (294) and *Gd*Pdc1.1 (283), *Gd*Pdc1.2 (284) and *Gd*Pdc1.3 (285) from *G. diazotrophicus* (**Table 3)**. In subsequent experiments, these heterologous enzymes were compared with the native *S. cerevisiae* 2-oxo acid decarboxylases *Sc*Pdc1, *Sc*Pdc5, *Sc*Pdc6 and *Sc*Aro10.

Table 3. Heterologous 2-oxo-acid decarboxylase (2-OADC) genes investigated in this study. Published information on activity with pyruvate and with the aromatic 2-oxo acids phenylpyruvate (PPY) and *p*-hydroxyphenylpyruvate (*p*OHPPY) is presented. The right-hand column indicates for which of these genes the coding sequences were codon-optimized (Co) for expression in *S. cerevisiae* in the present study.

Organism	Gene	Activity with pyruvate	Activity with PPY or *p*OHPPY	Co
K. marxianus	*KmPDC1*	Yes (292)	Unknown	No
	KmPDC5	No (292)	Unknown	No
	KmARO10	Unknown	Unknown	No
K. lactis	*KlPDC1*	Yes (295)	Unknown	No
	KlPDC5	Unknown	Unknown	No
	KlARO10	Unknown	Suggested (296)	No
Y. lipolytica	*YlPDC1*	Suggested (297)	Unknown	No
Z. mobilis	*Zmpdc1*	Yes (298)	not PPY (299) low *p*OHPPY (299)	Yes
G. diazotrophicus	*Gdpdc1.1*	Yes (283)	Not *p*OHPPY (283)	Yes
	Gdpdc1.2	Unknown	Unknown	Yes
	Gdpdc1.3	Unknown	Unknown	Yes

A phylogenetic tree of the amino-acid sequences of the selected 2-OADCs generated by multiple-sequence alignment using Clustal Ω (288) showed a clear segregation of the eukaryotic (yeast) and bacterial sequences (**Fig. 2**). As anticipated, sequences of the *S. cerevisiae* pyruvate decarboxylases Pdc1 and Pdc5 clustered with those of the *K. marxianus* and *K. lactis* Pdc1 orthologs. Conversely, *Kl*Pdc5 and *Km*Pdc5 had 76 % similarity to one another but only 34% similarity to *Sc*Pdc5. Despite the phylogenetic distance of the yeasts

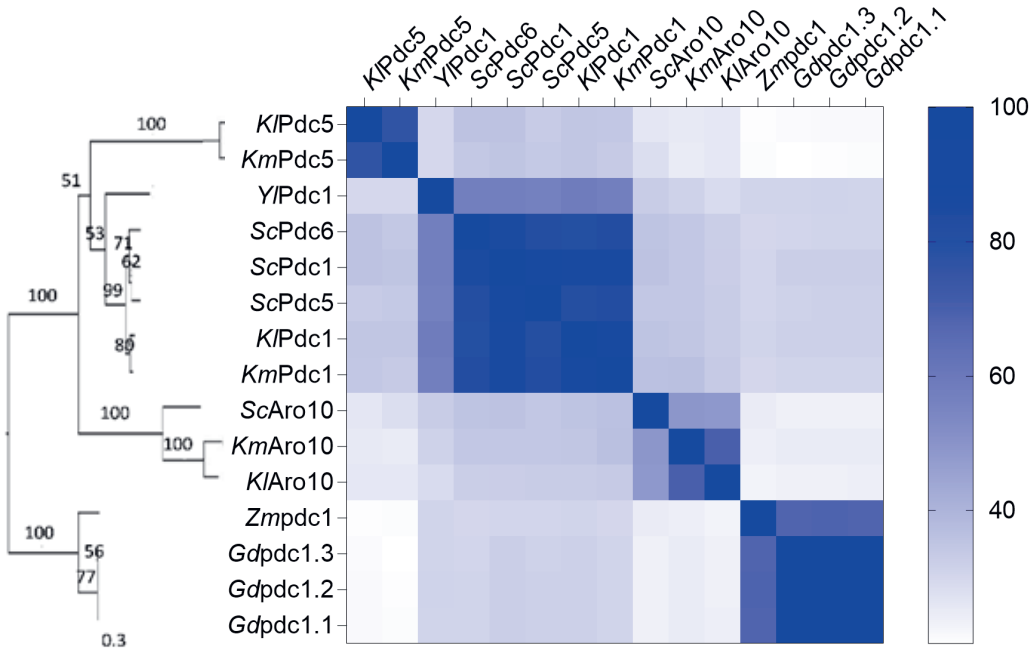


Figure 2. Phylogenetic tree and protein homology of the different 2-oxo acid decarboxylase candidates screened in this study. The phylogenetic tree was constructed as described in section 2.7 and the number of amino acid substitutions per site are represented by the scale bar. Protein homology is represented in a heatmap. The greater the intensity of blue, the higher the amino acid homology between two candidates.

S. cerevisiae and *Y. lipolytica*, *Yl*Pdc1 was more closely related to *Sc*Pdc1 than to the other selected proteins.

In vivo pyruvate-decarboxylase activity of heterologous enzymes expressed in *S. cerevisiae*

To assess *in vivo* pyruvate decarboxylase activity of the 11 selected enzymes upon introduction in *S. cerevisiae*, they were expressed from an episomal plasmid and under control of the strong consecutive *TDH3* promoter in the pyruvate-decarboxylase-negative strain CENPK711-7C (*ura3Δ pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ*). When precultures on SMEG were transferred to SMD, the empty-vector control strain IMZ001 (CENPK711-7C no 2-OADC, *URA3*$^{\uparrow}$) and the *Sc*Aro10-expressing strain IMZ002 (CENPK711-7C *ScARO10*$^{\uparrow}$) failed to grow. This observation was consistent with the inability of pyruvate-decarboxylase-negative *S. cerevisiae* strains to grow on glucose as sole carbon source (166, 167). Also strains IME418 (CENPK711-7C *KlPDC5*$^{\uparrow}$), IME423 (CENPK711-7C *KmARO10*$^{\uparrow}$), IME424 (CENPK711-7C *KlARO10*$^{\uparrow}$) and IME425 (CENPK711-7C *Gdpdc1.1*$^{\uparrow}$), did not show growth on SMD after seven days of incubation, while the same strains were fully grown on SMEG. These results indicated that the heterologous genes introduced into these strains were either not

functionally expressed or did not encode a functional pyruvate decarboxylase. In contrast, strains IME419 (CENPK711-7C *YlPDC1*↑), IME615 (CENPK711-7C *KlPDC1*↑), IME420 (CENPK711-7C *KmPDC1*↑), IME421 (CENPK711 *Zmpdc1*↑), IME474 (CENPK711-7C *Gdpdc1.2*↑), IME495 (CENPK711-7C *Gdpdc1.3*↑), and the positive control strain IME667 (CENPK711-7C *ScPDC1*↑) all showed instantaneous growth on SMD (**Fig. 3**).

Specific growth rates on SMD of these strains, including the positive-control strain IME667, were between 0.12 and 0.15 h^{-1}. These growth rates were ca. 3-fold lower than that of the reference strain CEN.PK113-7D (0.42 ± 0.00 h^{-1}), which retains all 2-oxo acid decarboxylase genes in their native genetic context (**Fig. 3**). The platform strain CENPK711-7C used to individually express the 2-oxo acid decarboxylases carried a deletion of *THI3*, a gene that was originally assumed to encode a fifth *S. cerevisiae* 2-oxo acid decarboxylase (83, 300, 301) but was later shown to instead encode a protein involved in thiamine homeostasis (302, 303). To investigate if inactivation of *THI3* was responsible for the unexpectedly

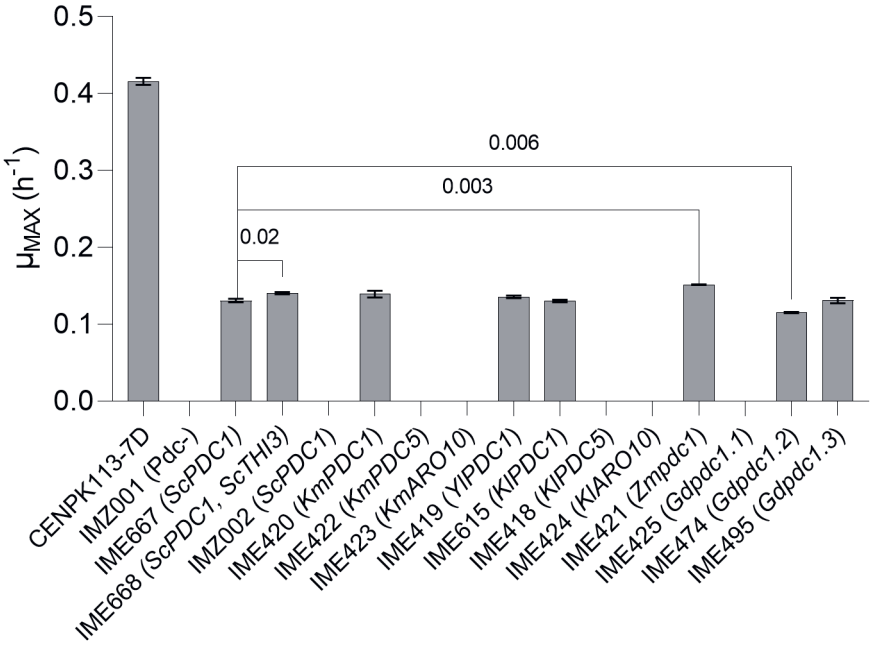


Figure 3. Specific growth rates of CEN.PK711-7C (*pdc1Δ pdc5Δ pdc6Δ aro10Δ* thi3Δ) expressing individual 2-oxo acid decarboxylases. Strains IMZ001 (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ* 2μm-*URA3)* and CEN.PK113-7D (*PDC1 PDC5 PDC6 ARO10 THI3*) were used as references. Strains IMZ001, CEN.PK113-7D, IME667 (*ScPDC1*↑), IME668 (*ScPDC1 ScTHI3*↑), IMZ002 (*ScARO10*↑), IME615 (*KlPDC1*↑), IME422 (*KmPDC5*↑), IME424 (*KlARO10*↑), IME420 (*KmPDC1*↑), IME423 (*KmARO10*↑), IME418 (*KlPDC5*↑), IME419 (*YlPDC1*↑), IME421 (*Zmpdc1*↑), IME425 (*Gdpdc1.1*↑), IME474 (*Gdpdc1.2*↑) and IME495 (*Gdpdc1.3*↑) expressed different decarboxylases genes from episomal multicopy plasmids and under control of *ScTDH3*pr. For each strain, duplicate cultures were grown on synthetic medium containing 2% glucose (SMD) at 30 °C. Strains with a significantly different ($p < 0.05$, *t*-test) specific growth rate than strain IME667 (*ScPDC1*↑) are indicated with the corresponding *p*-value.

low specific growth rate of the tested strains, strain IME668 (CEN.PK711-7C *ScPDC1*↑ *ScTHI3*↑) was constructed. Its specific growth rate on SMD was only 10% higher than that of strain IME667 (CENPK711-7C *ScPDC1*↑) and therefore still much lower than that of strain CEN.PK113-7D (**Fig. 3**).

In vitro comparison of substrate specificity of 2-oxo acid decarboxylase from various origin.

To assess the substrate specificities of the selected heterologous pyruvate decarboxylases, enzyme activity assays were performed in cell extracts. In view of the goal of this study to eliminate production of aromatic fusel alcohols and acids, these assays focused on their activities with pyruvate and phenylpyruvate as substrates (**Table 4**).

Table 4. Pyruvate and phenylpyruvate decarboxylase activities in cell extracts of *S. cerevisiae* strain CEN.PK711-7C (*pdc1*Δ *pdc5*Δ *pdc6*Δ *aro10*Δ *thi3*Δ) expressing individual 2-OADC genes from a multicopy plasmid. Cell extracts were prepared from late-exponential-phase shake-flask cultures on SMEG. Pyruvate decarboxylase and phenylpyruvate activities were assayed on duplicate cultures.

Strain	Genotype	Mean V_{MAX} (μmol · min^{-1} · (mg of protein)$^{-1}$)	
		Pyruvate	Phenylpyruvate
CEN.PK113-7D	wt	0.17 ± 0.0	BD
IMZ001	Pdc$^-$	BD	BD
IME667	*ScPDC1*	8.86 ± 0.5	0.098 ± 0.00
IMZ024	*ScPDC5*	2.05 ± 0.1	0.101 ± 0.01
IMZ031	*ScPDC6*	1.31 ± 0.0	0.030 ± 0.00
IMZ002	*ScARO10*	BD	0.105 ± 0.00
IME420	*KmPDC1*	6.20 ± 0.3	0.122 ± 0.00
IME423	*KmARO10*	BD	0.174 ± 0.00
IME418	*KlPDC5*	BD	BD
IME615	*KlPDC1*	5.58 ± 0.3	0.103 ± 0.01
IME422	*KmPDC5*	BD	BD
IME424	*KlARO10*	BD	0.152 ± 0.01
IME419	*YlPDC1*	3.01 ± 0.1	0.145 ± 0.00
IME421	*Zmpdc1*	8.79 ± 0.5	BD
IME425	*Gdpdc1.1*	BD	BD
IME474	*Gdpdc1.2*	0.25 ± 0.0	BD
IME495	*Gdpdc1.3*	0.22 ± 0.0	BD

BD, below detection limit. For pyruvate decarboxylase activity the detection limit was <0.04 μmol mg of protein^{-1} min^{-1}, for phenylpyruvate the limit was at <20 nmol mg of protein^{-1} min^{-1}

Absence of pyruvate decarboxylase activity in cell extracts of strains IMZ002 (*ScARO10*↑), IME423 (*KmARO10*↑), IME424 (*KlARO10*↑), IME418 (*KlPDC5*↑) and IME422 (*KmPDC5*↑) correlated with their inability to grow on SMD. In contrast, cell extracts of strains expressing *Yl*Pdc1 (IME419), *Km*Pdc1 (IME420), *Kl*Pdc1 (IME615), *Zm*Pdc1 (IME421), *Gd*Pdc1.2 (IME474) or *Gd*Pdc1.3 (IME495), as well as the strains expressing *Sc*Pdc1 (IME667), *Sc*Pdc5 (IMZ024) and *Sc*Pdc6 (IMZ031) all showed pyruvate-decarboxylase activities (**Table 4**). The highest activities, above 3 µmol (mg protein)$^{-1}$ min^{-1}, were observed in cell extracts of strains expressing Pdc1 orthologs from yeasts (**Table 4**). Cell extracts of strains expressing either of the two *G. diazotrophicus* decarboxylases (*Gd*Pdc1.2 and *Gd*Pdc1.3) exhibited a 35-fold lower pyruvate-decarboxylase activity than those of a strain expressing *Sc*Pdc1. However, pyruvate-decarboxylase activities of cell extracts of the strain expressing *Zm*Pdc1, the other bacterial pyruvate decarboxylase, were close to those observed with the *Sc*Pdc1-expressing strain IME667 (**Table 4**).

As anticipated, cell extracts of strains expressing yeast Aro10 orthologs showed phenylpyruvate-decarboxylase activity, although activities were two orders of magnitude lower than pyruvate-

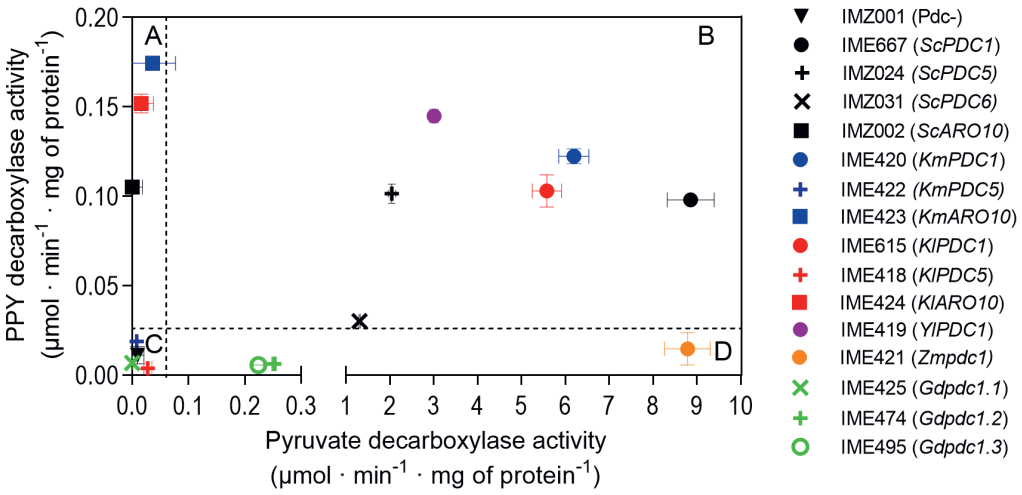


Figure 4. Specific decarboxylase activities for pyruvate and phenylpyruvate (PPY) in cell extracts of CENPK711-7C (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ)* expressing individual 2-oxo acid decarboxylase genes from a multicopy plasmid. All strains were grown duplicates at 30 °C on synthetic medium containing 2% (v/v) glycerol and 2% (v/v) ethanol as carbon source (SMEG). Black: *S. cerevisiae*, blue: *K. marxianus*, purple; *Y. lipolytica*, red: *K. lactis*, orange; *Z. mobilis,* green: *G. diazotrophicus.* Strains tested were IMZ001 (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ URA3*↑), IMZ002 (*ScARO10*↑), IMZ024 (*ScPDC5*↑), IMZ031 (*ScPDC6*↑), IME418 (*KlPDC5*↑), IME419 (*YlPDC1*↑), IME420 (*KmPDC1*↑), IME421 (*Zmpdc1*↑), IME422 (*KmPDC5*↑), IME423 (*KmARO10*↑), IME424 (*KlARO10*↑), IME425 (*Gdpdc1.1*↑), IME474 (*Gdpdc1.2*↑), IME495 (*Gdpdc1.3*↑), IME615 (*KlPDC1*↑) and IME667 (*ScPDC1*↑). The dotted line indicates the detection limit for decarboxylase activity, which was <0.04 µmol·mg protein^{-1}·min^{-1} for pyruvate as substrate and <20 nmol mg of protein^{-1} min^{-1} for phenylpyruvate as substrate. This results in the visualization of 4 classes: enzymes with decarboxylase activity for A) PPY but not pyruvate, B) both PPY and pyruvate C) no activity for either substrate and D) activity for pyruvate, but not PPY.

decarboxylase activities observed in cell extracts of strains expressing yeast or *Z. mobilis* Pdc1 homologs (**Table 4**). Three of the heterologous 2-oxo decarboxylases with demonstrated *in vivo* and *in vitro* pyruvate-decarboxylase activity upon expression in *S. cerevisiae* (*Gd*Pdc1.2, *Gd*Pdc1.3 and *Zm*Pdc1) showed no activity with 5 mM phenylpyruvate as substrate (**Table 4**). These enzymes were therefore identified as promising candidates for replacing the native 2-OADCs in *S. cerevisiae* strains engineered for production of phenylpropanoid (**Fig. 4**).

To estimate the Michaelis constant (K_m) of the heterologous pyruvate decarboxylases for pyruvate, enzyme activity assays with cell extracts of strains expressing the prokaryotic enzymes and *Sc*Pdc1 yeast orthologs (*Kl*Pdc1, *Km*Pdc1 and *Yl*Pdc1) were performed at pyruvate concentrations ranging from 0.1 to 50 mM (Fig. 5).

To investigate whether, similar to *S. cerevisiae* pyruvate decarboxylase (304), the heterologous pyruvate decarboxylases exhibit cooperativity, the data was fitted by non-linear regression to substrate-saturation Michaelis-Menten kinetics as well as to sigmoidal allosteric Hill kinetics (**Table 5**). Consistent with literature (82), cell extracts containing *Sc*Pdc1 showed a

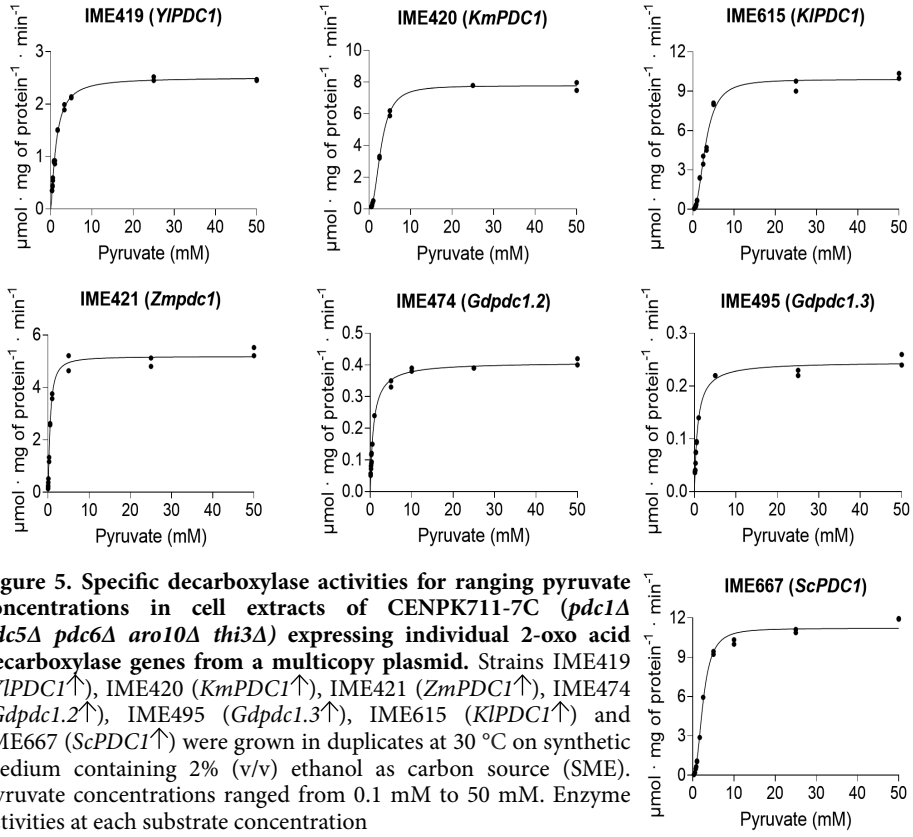


Figure 5. Specific decarboxylase activities for ranging pyruvate concentrations in cell extracts of CENPK711-7C (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ)* expressing individual 2-oxo acid decarboxylase genes from a multicopy plasmid. Strains IME419 (*YlPDC1*↑), IME420 (*KmPDC1*↑), IME421 (*ZmPDC1*↑), IME474 (*Gdpdc1.2*↑), IME495 (*Gdpdc1.3*↑), IME615 (*KlPDC1*↑) and IME667 (*ScPDC1*↑) were grown in duplicates at 30 °C on synthetic medium containing 2% (v/v) ethanol as carbon source (SME). Pyruvate concentrations ranged from 0.1 mM to 50 mM. Enzyme activities at each substrate concentration

Hill coefficient of 2.4, while a similar cooperativity was observed for cell extracts containing the *Kluyveromyces* enzymes *Kl*Pdc1 and *Km*Pdc1. In contrast, assays with cell extracts containing either *Y. lipolytica* Pdc1 or one of the three bacterial enzymes (*Zm*pdc1, *Gd*pdc1.2 or *Gd*pdc1.3), yielded a Hill coefficient close to one and absence of a sigmoidal relation between substrate concentration and enzyme activity (Table 5, Fig. 5), thus indicating absence of cooperativity (Table 5). In these assays, the *Z. mobilis* pyruvate decarboxylase *Zm*Pdc1 showed a 4-fold lower K_m than *Sc*Pdc1 and a higher Vmax/Km ratio than the *G. diazotrophicus* pyruvate decarboxylases.

Table 5. Specific pyruvate decarboxylase activity, K_m and the Hill coefficient for cell free extracts of *S. cerevisiae* strain CEN.PK711-7C (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ*) expressing individual 2-OADC genes. All strains were grown aerobically at 30 °C, 200 RPM in shake flasks containing 100 mL synthetic medium with 2% w/v ethanol as carbon source (SME). The cell extracts were prepared from late-exponential-phase shake-flask cultures. Different pyruvate concentrations were used as substrate for measuring pyruvate decarboxylase activity ranging from 0.1 mM to 50 mM. Enzyme activities were assayed from duplicate cultures.

Strain	Genotype	K_m (mM) ± SD	V_{MAX} (μmol min^{-1} (mg of protein)$^{-1}$ ± SD	Hill coefficient ± SD	V_{MAX}/K_m ratio
IME667	*ScPDC1*	2.5 ± 0.0	11.2 ± 0.0	2.4 ± 0.1	4.5
IME420	*KmPDC1*	2.9 ± 0.1	7.8 ± 0.1	2.4 ± 0.1	2.7
IME419	*Ylpdc1*	1.3 ± 0.0	2.5 ± 0.0	1.3 ± 0.0	1.9
IME615	*Klpdc1*	3.1 ± 0.1	9.9 ± 0.1	2.3 ± 0.0	3.2
IME421	*Zmpdc1*	0.6 ± 0.0	5.4 ± 0.1	1.3 ± 0.1	8.8
IME474	*Gdpdc1.2*	0.8 ± 0.1	0.4 ± 0.0	1.0 ± 0.0	0.5
IME495	*Gdpdc1.3*	0.8 ± 0.0	0.2 ± 0.0	1.0 ± 0.0	0.3

*V*max and *Km* values were obtained by performing a nonlinear regression of specific decarboxylase activity over the substrate concentration using either a Michaelis-Menten model or a sigmoidal allosteric model. The Hill coefficients were calculated using the Hill equation; A Hill coefficient of 2.0 indicates positive cooperativity.

Decarboxylase swapping in a coumaric acid-producing *S. cerevisiae* strain

To investigate whether replacement of the native yeast 2-OADCs (Pdc1, Pdc5, Pdc6 and Aro10) by heterologous pyruvate decarboxylases ('decarboxylase swapping') can eliminate formation of by-products in *S. cerevisiae* strains engineered for phenylpropanoid production, a tester strain producing coumaric acid was constructed. To this end, *ARO10, PDC5* and *PDC6* were first deleted from the previously constructed strain IMX1593, which overexpresses feedback-insensitive alleles of the DAHP synthase and chorismate mutase (*aro3Δ* $ARO4^{K229L\uparrow}$ $ARO7^{T226I\uparrow}$) (227). Subsequently, expression cassettes for *PlstlA, Rctal1t, AtC4H* and *AtCPR1,* which encode for respectively, a phenylalanine ammonia lyase, tyrosine ammonia lyase, cinnamic acid hydroxylase and its cytochrome p450 reductase, required for the activation of the cytochrome P450, were integrated at the *X3* locus on CHRX (248). Deletion of the pyruvate

decarboxylase gene *PDC1* yielded the 2-OADC-negative, coumaric acid producing platform strain IMX2668. This strain was transformed with multi-copy plasmids carrying expression cassettes for the different 2-oxo acid decarboxylases with specificity for pyruvate. All these strains grew on SMD in shake-flask cultures, albeit slower than the *ScPDC1*-expressing reference strain (IME677) (**Table 6**). Consistent with their low pyruvate-decarboxylase activities in cell extracts, the lowest specific growth rates were observed for the two strains expressing the *G. diazotrophicus* pyruvate decarboxylases (**Table 6**).

Table 6. Specific growth rates of IMX2668, a coumaric acid producing background strain, fully devoid of all native 2-OADCS (*pdc1Δ, pdc5Δ, pdc6Δ, aro10Δ*), expressing individual 2-OADC genes from a multicopy plasmid. All strains were grown aerobically in biological duplicates at 30 °C, 200 RPM in shake flasks containing 100 mL synthetic medium with 2% w/v glucose as carbon source (SMD). CEN.PK113-7D (*PDC1 PDC5 PDC6 ARO10*), IME656 (pGGKd017 (*URA3*↑), IME677 (*ScPDC1*↑), IME658 (*Zmpdc1*↑), IME659 (*KmPDC1*↑), IME660 (*YlPDC1*↑), IME661 (*KlPDC1*↑), IME662 (*Gdpdc1.2*↑) and IME663 (*Gdpdc1.3*↑).

Strain	Genotype	μ_{MAX} (h^{-1})
CEN.PK113-7D	Ref.	0.39 ± 0.00
IME656	IMX2668 *URA3*↑	0.00 ± 0.00
IME677	IMX2668 *ScPDC1*↑	0.27 ± 0.00
IME658	IMX2668 *Zmpdc1*↑	0.20 ± 0.00
IME659	IMX2668 *KmPDC1*↑	0.28 ± 0.00
IME660	IMX2668 *YlPDC1*↑	0.18 ± 0.00
IME661	IMX2668 *KlPDC1*↑	0.23 ± 0.00
IME662	IMX2668 *Gdpdc1.2*↑	0.11 ± 0.00
IME663	IMX2668 *Gdpdc1.3*↑	0.17 ± 0.00

In line with the results of the enzyme activity assays, the coumaric acid-producing strains IME659 (*KmPDC1*↑), IME660 (*YlPDC1*↑) and IME661 (*KlPDC1*↑) produced 2-phenylethanol and *p*-hydroxyphenylethanol, at concentrations ranging from 0.26 mM to 0.69 mM (**Fig. 6**). In contrast, strains IME658 (*Zmpdc1*↑), IME662 (*Gdpdc1.2*↑) and IME663 (*Gdpdc1.3*↑) did not show detectable concentrations of these aromatic fusel alcohols. Cultures of IME658 (*Zmpdc1*↑) reached 20% higher final coumaric acid concentration than the reference strain IME677 (*ScPDC1*↑). In the shake-flask cultures, strains IME662 (*Gdpdc1.2*↑) and IME663 (*Gdpdc1.3*↑), did not consume all glucose and did not produce detectable amounts of ethanol (**Supplementary Fig. 2**).

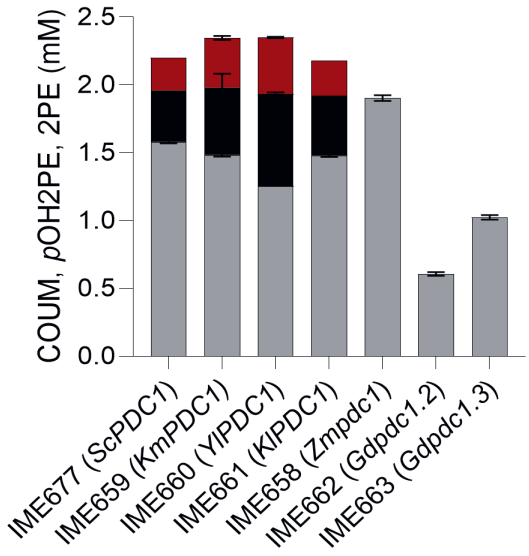


Figure 6. Metabolite profile of the coumaric acid producing reference strain IMX2668 (*Scpdc1Δ, Scpdc5Δ, Scpdc6Δ, Scaro101Δ*) expressing individual 2-OADC genes from a multicopy vector. IME677 (*ScPDC1*↑), IME658 (*Zmpdc1*↑) IME659 (*KmPDC1*↑), IME660 (*YlPDC1*↑), IME661 (*KlPDC1*↑), IME662 (*Gdpdc1.2*↑) and IME663 (*Gdpdc1.3*↑) were grown at 30 °C in biological triplicates on synthetic medium containing glucose as sole carbon source (SMD). All strains were inoculated at OD660 = 0.2 and grown for 72 h until they reached stationary phase. Red: 2-phenylethanol (2PE), black: p-hydroxyphenylethanol (*p*OH2PE), grey: coumaric acid (COUM). The concentrations of ethanol, glucose, pyruvate and glycerol are depicted in **Supplementary Fig. 2**.

To more accurately quantify the impact of decarboxylase swapping on coumaric acid production, the coumaric acid-producing reference strain IME677 (*ScPDC1*↑), as well as strains IME658 (*Zmpdc1*↑) and IME663 (*Gdpdc1.3*↑) were grown aerobically on SMD in pH-controlled bioreactors (**Fig. 7**).

Under these conditions, the reference strain IME677 produced 2.2 mM of coumaric acid and displayed the typical diauxic growth pattern of aerobic glucose-grown batch cultures of *S. cerevisiae,* with an initial respiro-fermentative growth phase followed by a respiratory ethanol consumption phase (211) (**Fig. 7A.I**). As observed in shake-flask cultures, strain IME677 (*ScPDC1*↑) produced 2-phenylethanol (0.10 mM) and *p*-hydroxyphenylethanol (0.15 mM) (**Fig. 7A.II**), which together corresponded to 12% of the total extracellular aromatic metabolites. In addition, this strain excreted detectable amounts of phenylpyruvate (0.10 mM) (**Fig. 7A.II**).

Although growing 30% slower than strain IME677, strain IME658 (*Zmpdc1*↑) also displayed a respiro-fermentative growth (**Fig. 7B.I**). However, in contrast to the *Sc*Pdc1-expressing reference strain, this *Zm*pdc1-expressing strain did not produce detectable levels of aromatic fusel alcohols (**Fig. 7B.II**). As in strain IME677, phenylpyruvate was detected (0.12 mM). In comparison with strain IME677, strain IME658 showed a higher coumaric acid titer (2.5 mM vs 2.2 mM) and, in contrast to the reference strain, excreted the coumaric acid precursor cinnamic acid (0.1 mM). In the absence of aromatic fusel alcohol production, strain IME658 (*Zmpdc1*↑) therefore excreted 14% more coumaric acid and 24% more coumaric acid precursors than the reference strain IME677 (*ScPDC1*↑). These improvements were also

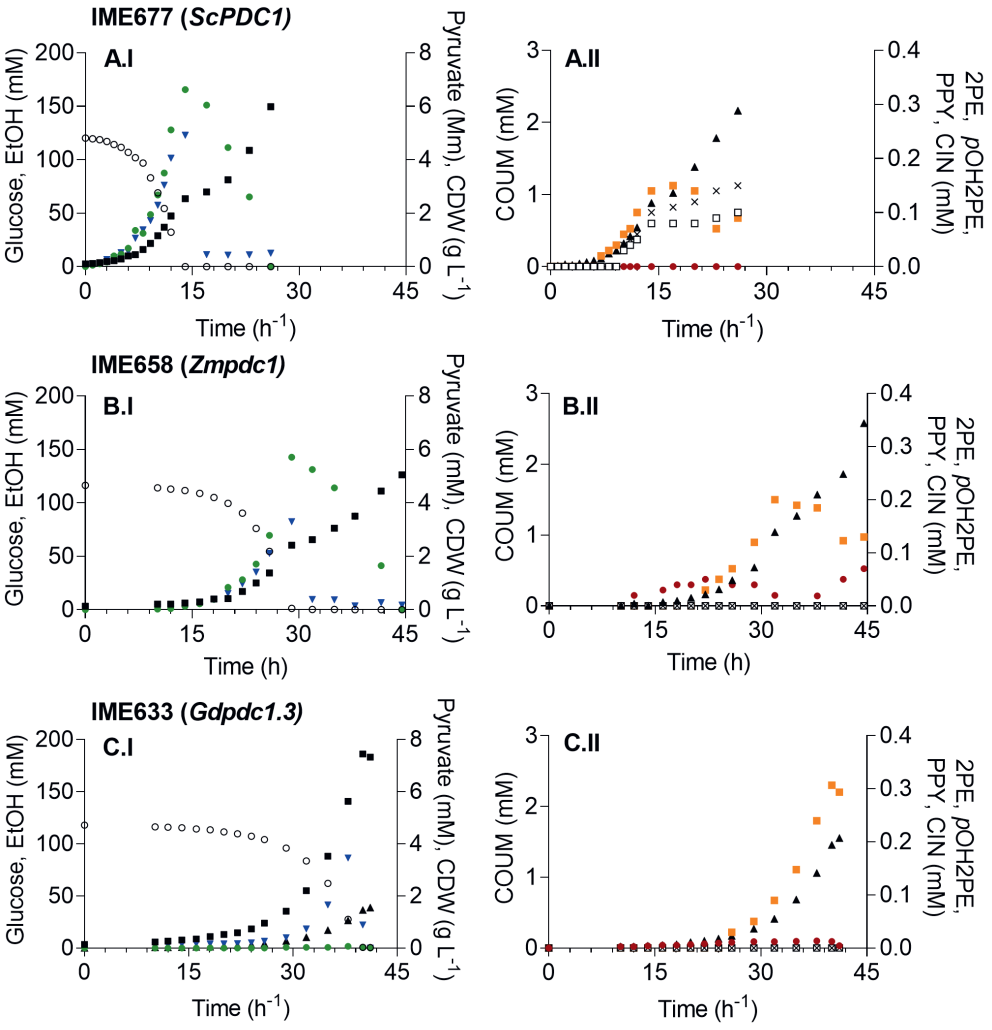


Figure 7. Aerobic batch cultivation in bioreactors of coumaric acid producing strains, expressing *ScPDC1, Kmpdc1* or *Gdpdc1.3* as sole pyruvate decarboxylase from a multicopy vector. All strains were grown aerobically at 30 °C, pH 5.0, in biological duplicates on synthetic medium containing glucose as sole carbon source (SMD). The results shown are from a single representative cultivation. Panels A: control strain IME677 (*ScPDC1*↑), panels B: IME658 (*Zmpdc1*↑) and panels C: IME663 (*Gdpdc1.3*↑). The left column (I) depicts the concentration of the cell dry weight (CDW) [■ (g L^{-1})], glucose [○ (mM)], ethanol [● (mM); EtOH] and pyruvate [▼ (mM); PYR] over time whereas the right column (II) shows the concentration of the aromatic metabolites coumaric acid [▲ (mM); COUM], cinnamic acid [● (mM); CIN], 2-phenylethanol [× (mM); 2PE], *p*-hydroxyphenylethanol [□ (mM); *p*OH2PE] and phenylpyruvate [■ (mM); PPY].

observed in the molar yields of coumaric acid on glucose ($Y_{COUM/S}$) in these strains, which were 22.0 ± 0.2 mmol mol^{-1} for strain IME658 and 18.4 ± 0.7 mmol mol^{-1} for strain IME677 (**Table 7**).

Table 7. Performance of aerobic batch cultures of 2-OADC expressing, coumaric acid producing strains. Specific growth rate (μ) and yields (Y) of biomass (X) and ethanol (EtOH) on glucose (S) during the glucose phase ([a]), the yields of biomass (X), 2-phenylethanol (2PE), *p*-hydroxyphenylethanol (*p*OH2PE) and coumaric acid (COUM) on glucose during the entire cultivation and the accumulated titer of all measured aromatics (2PE, *p*OH2PE, COUM, phenylpyruvate (PPY) and cinnamic acid (CIN)) of the *S. cerevisiae* strains IME677 (control strain, *ScPDC1*↑), IME658 (*Zmpdc1*↑) and IME633 (*Gdpdc1.3*↑).

Strain	**IME677**	**IME658**	**IME663**
Relevant genotype	*ScPDC1*	*Zmpdc1*	*Gdpdc1.3*
[a]μ (h^{-1})	0.29 ± 0.01	0.20 ± 0.02	0.14 ± 0.00
[a]$Y_{X/S}$ (g g^{-1})	0.12 ± 0.00	0.11 ± 0.00	0.35 ± 0.00
[a]$Y_{EtOH/S}$ (mol mol^{-1})	1.37 ± 0.01	1.20 ± 0.05	0.00 ± 0.00
$Y_{X/S}$ (g g^{-1})	0.27 ± 0.01	0.24 ± 0.01	0.35 ± 0.00
$Y_{2PE/S}$ (mmol mol^{-1})	1.08 ± 0.16	0.00 ± 0.00	0.00 ± 0.00
$Y_{pOH2PE/S}$ (mmol mol^{-1})	1.29 ± 0.04	0.00 ± 0.00	0.00 ± 0.00
$Y_{COUM/S}$ (mmol mol^{-1})	18.39 ± 0.50	21.97 ± 0.26	12.90 ± 0.48
∑Aromatics (mM)	2.58 ± 0.11	2.66 ± 0.08	1.78 ± 0.10

[a]Determined from the glucose phase only.

Strain IME663 (*Gdpdc1.3*↑) showed a 50% lower specific growth rate in the bioreactor cultures than strain IME677 (*ScPDC1*↑) (Table 7). In comparison to the other two coumaric acid-producing strains, it did not produce detectable amounts of ethanol and reached 30% higher final biomass concentrations. In contrast to the shake-flask cultures of strain IME663, the bioreactor cultures consumed all glucose. Although aromatic fusel alcohols were not detected in culture supernatants, strain IME663 reached a lower coumaric acid titer than strain IME677 (1.5 mM vs 2.2 mM) and, additionally, produced nearly three-fold higher extracellular phenylpyruvate concentrations (0.3 mM).

Discussion

Microbial thiamine-pyrophosphate-dependent pyruvate decarboxylases (EC 4.1.1.1) exhibit different kinetic properties and substrate specificities (81, 82, 305, 306). By exploring this diversity, we identified bacterial pyruvate decarboxylases that did not, or very slowly, decarboxylate phenylpyruvate in vitro and could functionally replace the native *S. cerevisiae* pyruvate decarboxylases in glucose-grown cultures in vivo. Replacing all native *S. cerevisiae* 2-OADCs in a coumaric acid producing strain by bacterial decarboxylases from *G. diazotrophicus* or *Z. mobilis* eliminated formation of aromatic by-products via the Ehrlich pathway. Moreover, the coumaric acid-producing strain *S. cerevisiae* IME658, in which the native yeast 2-OADCs were replaced by *Z. mobilis pdc1,* did not produce aromatic fusel alcohols and showed a higher coumaric acid yield than the congenic strain IME677 that instead expressed *ScPDC1*.

While our study provided a clear proof of principle for the 'decarboxylase swapping' approach, the *Zmpdc1*-expressing strain grew 30% slower than the *ScPDC1* expressing strain. These different growth rates occurred despite high and similar pyruvate-decarboxylase activities in cell extracts of SMEG-grown cultures of *pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ* strains carrying the same *Zmpdc1* and *ScPDC1* expression vectors (**Table 4** and **Fig. 4**). Slower growth of the *Zmpdc1*-expressing strain may be related to a reported 20-fold higher sensitivity of *Zm*Pdc1 to inhibition by its product acetaldehyde (307). It would therefore be interesting to express the acetaldehyde-tolerant variant *Zm*Pdc1^{W392M} (308, 309). Alternatively, as proposed earlier for a *ScPDC1*-overexpressing *S. cerevisiae* strain (310), reduced growth rates on SMD of strains expressing pyruvate-decarboxylase genes from episomal-multicopy plasmids may reflect protein-burden effects. Further metabolic engineering and/or adaptive laboratory evolution (311) can be applied to identify optimal expression levels of these pyruvate-decarboxylases. When impacts on specific growth rate can be prevented, the 2-oxo acid decarboxylase swapping strategy outlined in this study should be applicable for reduction of by-product formation by yeast strains engineered for production of a wide range of phenylpropanoids including stilbenoids, flavonoids and hydroxycinnamic acids.

Applicability of *Zm*Pdc1 in yeast biotechnology may extend beyond prevention of aromatic by-product formation. Previous research showed that, in contrast to *Sc*Pdc1 (82), *Z. mobilis* pyruvate decarboxylase does not decarboxylate the 2-oxo acids 3-methyl-2-oxopentanoate, 4-methyl-2-oxopentanoate and 3-methyl-2-oxobutanoate (299, 312), which are derived from isoleucine, leucine and valine, respectively. Elimination of these volatile fusel alcohol by-products during yeast-based ethanol production may enable reduced processing costs (313).

Expression of 2-OADCs from an episomal-multicopy plasmid in the CEN.PK711-7C background (*pdc1Δ pdc5Δ pdc6Δ aro10Δ thi3Δ*) resulted in specific growth rates for all

strains between 0.12 and 0.15 h-1 (Fig. 3). When CEN.PK113-7D is grown on synthetic medium with glucose as carbon source and ammonium sulphate as nitrogen source, *ScPDC1* transcript levels are 10-fold higher compared to the other decarboxylases (81) and serves as the main decarboxylase under these conditions. However, IME667 (*ScPDC1*$^{\uparrow}$) only had a specific growth rate of 0.13 h^{-1} whereas a specific growth rate close to CEN.PK113-7D (0.42 h^{-1}) was expected. Simultaneous expression of *ScPDC1* and *ScTHI3* (IME668), did not lead to a higher growth rate. Surprisingly, expression of the same 2-OADC multicopy plasmids in a newly constructed, *pdc1Δ pdc5Δ pdc6Δ aro10Δ* coumaric acid producing strain (IMX2668) resulted in strains with much higher specific growth rates (0.11 to 0.28 h^{-1} (Table 6)). Therefore a genetic defect, besides the *THI3* deletion, in the CEN.PK711-7C background and its transformants is causing the low specific growth which may find its origin in the use loxP-Cre recombinase during the construction of the strain (82); a method that can cause chromosomal recombination (314). Whole genome sequencing might elucidate the exact cause.

In addition to outlining a metabolic engineering strategy for minimizing by-product formation, our results provided new insights in the diversity of microbial 2-OADCs. Except for *Sc*Aro10, *Km*Aro10 and *Kl*Aro10, the genes evaluated in this study were annotated as structural genes encoding pyruvate decarboxylases. Based on the inability of *Km*Pdc5 and *Km*Pdc5 to complement the growth defect of a pyruvate-decarboxylase-negative *S. cerevisiae* strain and the absence of in vitro decarboxylase activity with pyruvate or phenylpyruvate, further research is required to investigate their catalytic activity. A predicted pyruvate decarboxylase (*Yl*Pdc1; YALI0D10131g) in *Y. lipolytica*, which showed low sequence similarity with other yeast pyruvate decarboxylases, is active and in contrast to other yeast pyruvate-decarboxylases (315) did not exhibit cooperativity for its substrate (**Table 5**). This result is intriguing in view of the inability of this yeast to produce ethanol (316) and because it is generally assumed that, in *Y. lipolytica*, cytosolic acetyl-CoA, which is a key precursor for lipid synthesis by this oleaginous yeast, originates from the activity of the ATP-citrate lyase (317). Combined with previously reported aldehyde dehydrogenases (291) and an acetyl-CoA synthetase (316, 318), *Yl*Pdc1 could provide an alternative, energetically less efficient (43), bypass. Further research should establish the physiological relevance of *Yl*Pdc1 in its native host.

The present study, which explored only a fraction of the natural diversity of 2-OADCs, illustrates for further screening, mutagenesis and targeted protein engineering to tailor catalytic and regulatory properties of these key enzymes to specific applications in biotechnology.

Financial support

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 720824.

Declaration of interest

None

Acknowledgements

We acknowledge Arun Rakjumar for supplying part plasmids expressing *K. marxianus* enzymes, Mislav Oreb for providing us with the plasmid pE_MGV14. We thank Agnes Hol and Haje Huffstadt for their contributions to plasmid construction and Christiaan Mooiman for technical support.

Supplementary Information

Supplementary Table 1. All primers used in this study

Supplementary Figure 2. Metabolite profile of the coumaric acid producing reference strain IMX2668 (*Scpdc1Δ, Scpdc5Δ, Scpdc6Δ, Scaro10Δ*) expressing individual 2-OADC genes from a multicopy vector.

Supplementary Data Set 1. The amino acid sequences of all 2-oxo acid decarboxylases used in this study or references.

Outlook

Microbial production of valuable aromatic compounds has gained more and more attention as an attractive, more sustainable and reliable alternative to chemical or plant-based production. Among these compounds, plant secondary metabolites such as phenylpropanoids receive special interest because of their potential health benefits and their use as supplements to enhance the flavor, fragrance, color or shelf-life of foods. However, in order to reach industrial titers, production rates and yields, several hurdles need to be overcome. In this dissertation some of the challenges involved in microbial production were addressed. First of all, several successful metabolic engineering strategies were established to increase the intracellular availability of phenylalanine and tyrosine, metabolites required for phenylpropanoid biosynthesis. Secondly the plant-natural anthocyanin pelargonidin 3-O-glucoside was successfully produced in an engineered yeast strain by among others improving the copy number of the chalcone synthase, one of the rate limiting enzymes in phenylpropanoid biosynthesis. Finally, a 2-oxo acid decarboxylase was identified with limited substrate specificity for pyruvate, resulting in the elimination of the common by-products 2-phenylethanol and *p*-hydroxyphenylethanol during *de novo* phenylpropanoid production in yeast. However, as also highlighted in chapter 3, anthocyanin titers and yields remain low, intermediates of the pathway tend to accumulate and instead of a single compound, several closely related compounds are produced. These, and more, issues need to be improved in order to make yeast-based production of aromatic compounds economically viable.

During phenylpropanoid biosynthesis the overall metabolic capacity through a dedicated reaction network needs to be increased by optimizing the flux through the considered pathways to achieve efficient conversion of substrate without accumulation of intermediates or the formation of by-products. However, many enzymes in this pathway exhibit poor catalytic efficiency or substrate promiscuity. This was also illustrated in **chapter 3** by the accumulation of several intermediates of the pelargonidin 3-O-glucoside pathway and the formation of by-products such as kaempferol and kaempferol 3-O-glucoside, resulting from the broad substrate specificity of the anthocyanin synthase (ANS) and the anthocyanin 3-O-glucosyltransferase (3GT). Most studies trying to tackle these two issues focus on screening heterologous enzymes with higher catalytic activity or a limited substrate specificity. Although this approach has been successful in some cases (135, 152, 154, 162, 203), most enzymes still perform suboptimal and have to be further improved through non-targeted or targeted protein engineering methods. Typically, classical non-targeted methods include random mutagenesis or error-prone PCR in which gene libraries are generated and used to transform a microbial host to test either in vitro or in vivo the performance of individual variants. Alternatively, rational, targeted

protein modifications rely on quantitative models of protein properties and requires a holistic and predictive understanding of structural stability and quantitative molecular function of the protein studied. Combined with overexpression, using highly active promoters or through an increase in gene copy number by either expression from a multi-copy plasmid or multi-copy chromosomal integration in repetitive regions can be used to relieve metabolic bottlenecks and thus improve strain performance.

Speculating on substrate channeling and generation of a metabolon, strategies based on enzymes fusion can also be contemplated to further improve flux through a metabolic pathway. This approach might contribute to increased local substrate concentrations resulting in enhanced reaction efficiencies while simultaneously minimizing possible toxicity effects (101). For example, the fusion of the 4-coumarate-CoA ligase and the stilbene synthase resulted in a 15-fold increase in resveratrol production relative to regular expression of the enzyme separately (319). However, the limited information on chimeric proteins does not allow a good prediction of the impact of a fusion between two proteins. Therefore one has to anticipate potential misfolding and protein aggregation phenomena (108). To mitigate the unpredictable consequence of protein fusion and retain advantage of metabolic channeling, synthetic scaffolds represent a workable alternative (101, 108). This approach co-localizes all the enzymes of a pathway into synthetic enzymatic complexes in a programmable manner, by using designed interactions between specific protein-protein interaction domains and ligands. This optimizes the local concentrations of metabolic intermediates, while avoiding accumulation of intermediates or potential toxic effects. It would also allow for high local concentrations of the metabolites, which would increase the probability of a substrate-enzyme encounter (320). Consequently this could improve the kinetics of the enzymes by, for instance, lowering the in vivo apparent K_m) (321). Finally, adjusting the number of interaction domain repeats that connect enzymes to the enzymatic complex enable the modulation of relative fluxes of each enzyme to enhance production levels (320, 321).

For all of these aforementioned metabolic engineering solutions a measure to quantify the effectiveness of the strategies is required. Although accurate analytical methods to measure most intermediates of the phenylpropanoid pathway are available, they are not compatible with high throughput measurements. Genetically-encoded biosensors can potentially transform information about a specific metabolite into an optical output e.g., fluorescence readout, and therefore facilitate molecule detection. Biosensors enable detection and quantification of specific metabolites in single cells and have the potential to enhance strain development. For example, a naringenin biosensor (322) could be used for screening modifications aimed at steps up until naringenin, such as an improved chalcone synthase. Alternatively, a malonyl-CoA biosensor could be used to assess modifications aimed at improving the malonyl-CoA supply (323). Of course, there are several pitfalls with biosensors as well, such as poor product

sensitivity, limited dynamic range and applicability in *Saccharomyces* yeasts which may render the construction of a tailor-made, phenylpropanoid specific, biosensor challenging.

Hitherto, a pure (mono)culture is by far the most common approach for microbial fermentation, however modular co(mixed)-cultures could provide several benefits for phenylpropanoid biosynthesis (324). Such an approach based on the division of labor would contribute to a reduced metabolic burden (38), enable cross-feeding strategies that can increase the efficiency of the pathway and provide robustness to the system (325). Additionally, genes can be expressed in an environment that is most suitable for optimal expression. Furthermore, an optimal level between the metabolic fluxes of the pathway modules can be obtained by varying the subpopulations ratio of the co-culture (324, 326, 327). Finally, pathways can be shaped in a modular fashion, thereby facilitating additional metabolic engineering steps or interchanging microbial strains. The modular, co-culture engineering concept has emerged as a new approach to efficiently conduct heterologous biosynthesis and could greatly contribute to successful production of natural products (324). To illustrate this, a co-culture with two different *S. cerevisiae* strains was recently performed. One strain with precursor modifications (*ACC1*$^{\uparrow}$ *ARO4*$^{K229L\uparrow}$ *aro10Δ pdc5*Δ) containing the biosynthetic pathway up until naringenin was combined with a strain also containing the remaining two genes of the kaempferol pathway. This led to an increase in naringenin, but not kaempferol, production, together with up to 50% improvement of cell growth, demonstrating that a co-culture could indeed aid in reducing the metabolic burden (159) but also that this concept needs to be further optimized.

Cross-species co-cultures have also been attempted such as a co-culture between *S. cerevisiae* and *E. coli* for the production of naringenin (328). Here, the natural (and improved) ability of *E.coli* to produce aromatic amino acids was combined with the potential of yeast to functionally express the naringenin biosynthetic pathway genes. In another study an *E. coli* strain with improved tyrosine production and the coumaric acid biosynthetic genes was co-cultured with a *S. cerevisiae* strain with improved malonyl-CoA production carrying the required genes for resveratrol biosynthesis (329). Of course, using cross-species co-cultures is likely more challenging since different micro-organisms may require different nutrients/growth factors, thrive under different cultivation conditions and respond differently to environmental stresses.

To conclude, despite several proof of concept studies (**Chapter 1, Table 3**) that have reported *de novo* production of phenylpropanoids by yeasts, key performance indicators such as titer, productivity and yield are orders of magnitude lower than maximal theoretical values that are critical targets for industrial application. Many challenges therefore still need to be addressed before we can harness the full potential of yeast-based production of plant secondary metabolites.

Bibliography

1. Arranz-Otaegui, A., Gonzalez Carretero, L., Ramsey, M.N., Fuller, D.Q. & Richter, T. Archaeobotanical evidence reveals the origins of bread 14,400 years ago in northeastern Jordan. *PNAS* **115**, 7925-7930 (2018).
2. McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., Nuñez, A., Butrym, E.D., Richards, M.P., Wang, C., Cheng, G., Zhao, Z. & Wang, C. Fermented beverages of pre- and proto-historic China. *PNAS* **101**, 17593–17598 (2004).
3. McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M.P., Smith, K.E., Hall, G.R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., Failla, O., Cola, G., Mariani, L., Boaretto, E., Bacilieri, R., This, P., Wales, N. & Lordkipanidze, D. Early Neolithic wine of Georgia in the South Caucasus. *PNAS* **114**, E10309-E10318 (2017).
4. Damerow, P. Sumerian Beer: The Origins of Brewing Technology in Ancient Mesopotamia. *Cuneiform Digital Library Journal* **2**, 1-20 (2012).
5. Michel, R.H., McGovern, P.E. & Badler, V.R. Chemical evidence for ancient beer. *Nature* **360**, 24 (1993).
6. Buchholz, K. & Collins, J. The roots - a short history of industrial microbiology and biotechnology. *Appl. Microbiol. Biotechnol.* **97**, 3747-3762 (2013).
7. He, G.Q., Liu, T.J., Sadiq, F.A., Gu, J.S. & Zhang, G.H. Insights into the microbial diversity and community dynamics of Chinese traditional fermented foods from using high-throughput sequencing approaches. *J. Zhejiang Univ. Sci. B.* **18**, 289-302 (2017).
8. Schlenk, F. Early Research on Fermentation–A Story of Missed Opportunities. *Trends Biochem. Sci.* **10**, 252-254 (1985).
9. Pasteur, L. Mémoire sur la fermentation alcoolique. *CR Acad. Sci.* **45**, 1032-1036 (1857).
10. Traube, M. Zur Theorie der Gährungs- und Verwesungs erscheinungen, wie der Fermentwirkungen überhaupt. *Annalen der Physik* **179**, 331-344 (1958).
11. Payen, A. & Persoz, J.F. in Annales de chimie et de physique, Vol. 53 73-92 (1833).
12. Weizmann, C. & Rosenfeld, B. The activation of the butanol-acetone fermentation of carbohydrates by *Clostridium acetobutylicum*. *Biochem. J.* **31**, 619–639. (1937).
13. Neuberg, C. & Reinfurth, E. Natürliche und erzwungene Glycerinbildung bei der alkoholischen Gärung. *BiochemZ* **92**, 32 (1918).
14. Max, B., Salgado, J.M., Rodríguez1, N., Cortés, S., Converti, A. & Domínguez, J.M. Biotechnological production of citric acid. *Braz. J. Microbiol.* **41**, 862-875 (2010).
15. Grzybowski, A. & Pietrzak, K. Tadeusz Reichstein (1897-1996): a cofounder of modern steroid treatment in dermatology. *Clin. Dermatol.* **30**, 243-247 (2012).
16. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J Exp. Pathol.* **10**, 226–236 (1929).
17. Campos Muñiz, C., Cuadra Zelaya, T.E., Rodríguez Esquivel, G. & Fernández, F.J. Penicillin and cephalosporin production: a historical perspective. *Rev. Latinoam. Microbiol.* **49**, 88-98 (2007).
18. Barnett, J.A. A history of research on yeasts 5: the fermentation pathway. *Yeast* **20**, 509-543 (2003).
19. Krebs, H.A. & Kornberg, H.L. in Energy Transformations in Living Matter 212-298 (Springer, Berlin, Heidelberg; 1957).
20. Watson, J.D.C., F. H. C. The structure of DNA. *Cold Spring Harb Symp Quant Biol* **18**, 123-131 (1953).
21. Cohen, S.N., Chang, A.C.Y., Boyer, H.W. & Helling, R.B. Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *PNAS* **70**, 3240-3244 (1973).
22. Johnson, I.S. Human insulin from recombinant DNA technology. *Science* **219**, 632-637 (1983).
23. Ladisch, M.R. & Kohlmann, K.L. Recombinant human insulin. *Biotechnol. Prog.* **1092**, 469-478 (1992).

24. Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F. & Kerlavage, A.R. Whole-genome random sequencing and assembly of *Haemophilus influenzae Rd*. *Science* **269**, 496-512 (1995).
25. Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S.G. Life with 6000 Genes. *Science* **274**, 563-567 (1996).
26. Venter, J.G., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M. & Huson, D.H. The Sequence of the Human Genome. *Science* **291**, 1304-1351 (2001).
27. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Felsenfeld, A., Wetterstrand, K., Patrinos, A. & Morgan, M. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
28. Bailey, J.E. Toward a science of metabolic engineering. *Science* **252**, 1668-1675 (1991).
29. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. & Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
30. Wright, A.V., Nunez, J.K. & Doudna, J.A. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **164**, 29-44 (2016).
31. Luo, X., Reiter, M.A., d'Espaux, L., Wong, J., Denby, C.M., Lechner, A., Zhang, Y., Grzybowski, A.T., Harth, S., Lin, W., Lee, H., Yu, C., Shin, J., Deng, K., Benites, V.T., Wang, G., Baidoo, E.E.K., Chen, Y., Dev, I., Petzold, C.J. & Keasling, J.D. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123-126 (2019).
32. Kiesling, J.D. Manufacturing Molecules Through Metabolic Engineering. *Science* **330**, 1355-1358 (2010).
33. Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F. & Smolke, C.D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095-1100 (2015).
34. van Maris, A.J., Konings, W.N., van Dijken, J.P. & Pronk, J.T. Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. *Metab. Eng.* **6**, 245-255 (2004).
35. Lehka, B.J., Eichenberger, M., Bjorn-Yoshimoto, W.E., Vanegas, K.G., Buijs, N., Jensen, N.B., Dyekjaer, J.D., Jenssen, H., Simon, E. & Naesby, M. Improving heterologous production of phenylpropanoids in *Saccharomyces cerevisiae* by tackling an unwanted side reaction of Tsc13, an endogenous double-bond reductase. *FEMS Yeast Res.* **17**, fox004 (2017).
36. Bellissimi, E., van Dijken, J.P., Pronk, J.T. & van Maris, A.J. Effects of acetic acid on the kinetics of xylose fermentation by an engineered, xylose-isomerase-based *Saccharomyces cerevisiae* strain. *FEMS Yeast Res.* **9**, 358-364 (2009).
37. Alper, H., Moxley, J., Nevoigt, E., Fink, G.R. & Stephanopoulos, G. Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production. *Science* **314**, 1565-1568 (2006).
38. Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S. & Koffas, M.A.G. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **34**, 652-664 (2016).
39. Luttik, M.A., Vuralhan, Z., Suir, E., Braus, G.H., Pronk, J.T. & Daran, J.M. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. *Metab. Eng.* **10**, 141-153 (2008).
40. Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K.,

Renninger, N.S. & Newman, J.D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-532 (2013).

41. de Kok, S., Kozak, B.U., Pronk, J.T. & van Maris, A.J. Energy coupling in *Saccharomyces cerevisiae*: selected opportunities for metabolic engineering. *FEMS Yeast Res.* **12**, 387-397 (2012).
42. Meadows, A.L., Hawkins, K.M., Tsegaye, Y., Antipov, E., Kim, Y., Raetz, L., Dahl, R.H., Tai, A., Mahatdejkul-Meadows, T., Xu, L., Zhao, L., Dasika, M.S., Murarka, A., Lenihan, J., Eng, D., Leng, J.S., Liu, C.L., Wenger, J.W., Jiang, H., Chao, L., Westfall, P., Lai, J., Ganesan, S., Jackson, P., Mans, R., Platt, D., Reeves, C.D., Saija, P.R., Wichmann, G., Holmes, V.F., Benjamin, K., Hill, P.W., Gardner, T.S. & Tsong, A.E. Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* **537**, 694-697 (2016).
43. van Rossum, H.M., Kozak, B.U., Pronk, J.T. & van Maris, A.J.A. Engineering cytosolic acetyl--coenzyme A supply in *Saccharomyces cerevisiae*: Pathway stoichiometry, free-energy conservation and redox-cofactor balancing. *Metab. Eng.* **36**, 99-115 (2016).
44. Perli, T., van der Vorm, D.N.A., Wassink, M., van den Broek, M., Pronk, J.T. & Daran, J.M. Engineering heterologous molybdenum-cofactor-biosynthesis and nitrate-assimilation pathways enables nitrate utilization by *Saccharomyces cerevisiae*. *Metab. Eng.* **65**, 11-29 (2021).
45. Milne, N., Luttik, M.A.H., Cueto Rojas, H.F., Wahl, A., van Maris, A.J.A., Pronk, J.T. & Daran, J.M. Functional expression of a heterologous nickel-dependent, ATP-independent urease in *Saccharomyces cerevisiae*. *Metab. Eng.* **30**, 130-140 (2015).
46. Verhoeven, M.D., Lee, M., Kamoen, L., van den Broek, M., Janssen, D.B., Daran, J.G., van Maris, A.J. & Pronk, J.T. Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci. Rep.* **7**, 46155 (2017).
47. Guadalupe-Medina, V., Wisselink, W., Luttik, M.A.H., de Hulster, E., Daran, J., Pronk, J.T. & van Maris, A.J.A. Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. *Biotechnol. Biofuels* **6**, 125 (2013).
48. Hyma, K.E., Saerens, S.M., Verstrepen, K.J. & Fay, J.C. Divergence in wine characteristics produced by wild and domesticated strains of *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **11**, 540-551 (2011).
49. Steensels, J., Meersman, E., Snoek, T., Saels, V. & Verstrepen, K.J. Large-scale selection and breeding to generate industrial yeasts with superior aroma production. *Appl Environ Microbiol* **80**, 6965-6975 (2014).
50. Hirst, M.B. & Richter, C.L. Review of Aroma Formation through Metabolic Pathways of *Saccharomyces cerevisiae* in Beverage Fermentations. *American Journal of Enology and Viticulture* **67**, 361-370 (2016).
51. Hiralal, L., Olaniran, A.O. & Pillay, B. Aroma-active ester profile of ale beer produced under different fermentation and nutritional conditions. *J. Biosci. Bioeng.* **117**, 57-64 (2014).
52. Rojas, V., Gil, J.V., Pinaga, F. & Manzanares, P. Studies on acetate ester production by non-*Saccharomyces* wine yeasts. *Int. J. Food Microbiol.* **70**, 283-289 (2001).
53. Verstrepen, K.J., Derdelinckx, G., Winderickx, J., Thevelein, J.M., Pretorius, I.S. & Delvaux, F.R. Flavor-active esters: Adding fruitiness to beer. *J. Biosci. Bioeng.* **96**, 110-118 (2003).
54. Saerens, S.M., Delvaux, F.R., Verstrepen, K.J. & Thevelein, J.M. Production and biological function of volatile esters in *Saccharomyces cerevisiae*. *Microb. Biotechnol.* **3**, 165-177 (2010).
55. Nordström, K. Formation of esters from alcohols by brewer's yeast. *J. Inst. Brew.* **70**, 328-336 (1964).
56. Scognamiglio, J., Jones, L., Letizia, C.S. & Api, A.M. Fragrance material review on phenylethyl alcohol. *Food Chem. Toxicol.* **50**, S224-S239 (2012).
57. Brennan, T.C., Williams, T.C., Schulz, B.L., Palfreyman, R.W., Kromer, J.O. & Nielsen, L.K. Evolutionary Engineering Improves Tolerance for Replacement Jet Fuels in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **81**, 3316-3325 (2015).
58. Aboody, M.S.A. & Mickymaray, S. Anti-Fungal Efficacy and Mechanisms of Flavonoids. *Antibiotics* **9**, 45 (2020).
59. Mandalari, G., Bennett, R.N., Bisignano, G., Trombetta, D., Saija, A., Faulds, C.B., Gasson, M.J. & Narbad, A. Antimicrobial activity of flavonoids extracted from bergamot (*Citrus ber-*

gamia Risso) peel, a byproduct of the essential oil industry. *J. Appl. Microbiol.* **103**, 2056-2064 (2007).

60. Zha, J., Wu, X. & Koffas, M.A. Making brilliant colors by microorganisms. *Curr. Opin. Biotechnol.* **61**, 135-141 (2020).
61. Eichenberger, M., Hansson, A., Fischer, D., Durr, L. & Naesby, M. *De novo* biosynthesis of anthocyanins in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **18**, foy046 (2018).
62. Levisson, M., Patinios, C., Hein, S., de Groot, P.A., Daran, J.M., Hall, R.D., Martens, S. & Beekwilder, J. Engineering *de novo* anthocyanin production in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **17**, 103 (2018).
63. Brochado, A.R., Matos, C., Møller, B., Hansen, J., Mortensen, U.H. & Patil, K.R. Overexpression of O-methyltransferase leads to improved vanillin production in baker's yeast only when complemented with model-guided network engineering. *Microb. Cell Factories* **9**, 84 (2010).
64. Etschmann, M.M., Bluemke, W., Sell, D. & Schrader, J. Biotechnological production of 2-phenylethanol. *Appl. Microbiol. Biotechnol.* **59**, 1-8 (2002).
65. Smeriglio, A., Barreca, D., Bellocco, E. & Trombetta, D. Proanthocyanidins and hydrolysable tannins: occurrence, dietary intake and pharmacological effects. *Br. J Pharmacol.* **174**, 1244-1262 (2017).
66. Neelam, Khatkar, A. & Sharma, K.K. Phenylpropanoids and its derivatives: biological activities and its role in food, pharmaceutical and cosmetic industries. *Crit. Rev. Food Sci. Nutr.* **60**, 2655-2675 (2020).
67. Yang, H., Xue, Y., Yang, C., Shen, W., Fan, Y. & Chen, X. Modular Engineering of Tyrosol Production in *Escherichia coli*. *J. Agric. Food Chem.* **67**, 3900-3908 (2019).
68. Milne, N., Thomsen, P., Molgaard Knudsen, N., Rubaszka, P., Kristensen, M. & Borodina, I. Metabolic engineering of *Saccharomyces cerevisiae* for the de novo production of psilocybin and related tryptamine derivatives. *Metab. Eng.* **60**, 25-36 (2020).
69. Eden, A., Simchen, G. & Benvenisty, N. Two yeast homologs of ECA39, a target for c-Myc regulation, code for cytosolic and mitochondrial branched-chain amino acid aminotransferases. *J. Biol. Chem.* **271**, 20242-20245 (1996).
70. Pires, E.J., Teixeira, J.A., Branyik, T. & Vicente, A.A. Yeast: the soul of beer's aroma - a review of flavour-active esters and higher alcohols produced by the brewing yeast. *Appl. Microbiol. Biotechnol.* **98**, 1937-1949 (2014).
71. Humia, B.V., Santos, K.S., Barbosa, A.M., Sawata, M., Mendonca, M.D.C. & Padilha, F.F. Beer Molecules and Its Sensory and Biological Properties: A Review. *Molecules* **24**, 1568 (2019).
72. Etschmann, M.M.W., Sell, D. & Schrader, J. Medium optimization for the production of the aroma compound 2-phenylethanol using a genetic algorithm. *J. Mol. Catal. B Enzym.* **29**, 187-193 (2004).
73. Gupta, C., Prakash, D. & Gupta, S. A Biotechnological Approach to Microbial Based Perfumes and Flavours. *J. Microbiol. Exp.* **2**, 00034 (2015).
74. Gustavsson, M. & Lee, S.Y. Prospects of microbial cell factories developed through systems metabolic engineering. *Microb. Biotechnol.* **9**, 610-617 (2016).
75. Ehrlich, F. Über die Bedingungen der Fuselölbildung und über ihren Zusammenhang mit dem Eiweissaufbau der Hefe. *Ber. Dtsch. Chem. Ges.* **40**, 1027–1047 (1907).
76. Hazelwood, L.A., Daran, J.M., van Maris, A.J., Pronk, J.T. & Dickinson, J.R. The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. *Appl. Environ. Microbiol.* **74**, 2259-2266 (2008).
77. Iraqui, I., Vissers, S., Andre, B. & Urrestarazu, A. Transcriptional Induction by Aromatic Amino Acids in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 3360–3371 (1999).
78. Iraqui, I., Vissers, S., Cartiaux, M. & Urrestarazu, A. Characterisation of *Saccharomyces cerevisiae ARO8* and *ARO9* genes encoding aromatic aminotransferases I and II reveals a new aminotransferase subfamily. *Mol. Gen. Genet.* **257**, 238-248 (1998).
79. Urrestarazu, A., Vissers, S., Iraqui, I. & Grenson, M. Phenylalanine- and tyrosine-auxotrophic mutants of *Saccharomyces cerevisiae* impaired in transamination. *Mol. Gen. Genet.* **257**, 230-237 (1998).
80. Boer, V.M., Tai, S.L., Vuralhan, Z., Arifin, Y., Walsh, M.C., Piper, M.D., de Winde, J.H., Pronk,

J.T. & Daran, J.M. Transcriptional responses of *Saccharomyces cerevisiae* to preferred and non-preferred nitrogen sources in glucose-limited chemostat cultures. *FEMS Yeast Res.* **7**, 604-620 (2007).

81. Vuralhan, Z., Luttik, M.A., Tai, S.L., Boer, V.M., Morais, M.A., Schipper, D., Almering, M.J., Kotter, P., Dickinson, J.R., Daran, J.M. & Pronk, J.T. Physiological characterization of the *ARO10*-dependent, broad-substrate-specificity 2-oxo acid decarboxylase activity of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **71**, 3276-3284 (2005).
82. Romagnoli, G., Luttik, M.A., Kotter, P., Pronk, J.T. & Daran, J.M. Substrate specificity of thiamine pyrophosphate-dependent 2-oxo-acid decarboxylases in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **78**, 7538-7548 (2012).
83. Vuralhan, Z., Morais, M.A., Tai, S.L., Piper, M.D. & Pronk, J.T. Identification and characterization of phenylpyruvate decarboxylase genes in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **69**, 4534-4541 (2003).
84. Äyräpää, T. Biosynthetic formation of higher alcohols by yeast. Dependence on the nitrogenous nutrient level of the medium. *J. Inst. Brew.* **77**, 266-276 (1971).
85. Hua, D. & Xu, P. Recent advances in biotechnological production of 2-phenylethanol. *Biotechnol. Adv.* **29**, 654-660 (2011).
86. Di Benedetto, R., Vari, R., Scazzocchio, B., Filesi, C., Santangelo, C., Giovannini, C., Matarrese, P., D'Archivio, M. & Masella, R. Tyrosol, the major extra virgin olive oil compound, restored intracellular antioxidant defences in spite of its weak antioxidative effectiveness. *Nutr. Metab. Cardiovasc. Dis.* **17**, 535-545 (2007).
87. Ball, S.G., Wickner, R.B., Cottarel, G., Schaus, M. & Tirtiaux, C. Molecular cloning and characterization of *ARO7-OSM2*, a single yeast gene necessary for chorismate mutase activity and growth in hypertonic medium. *Mol. Gen. Genet.* **205**, 326-330 (1986).
88. Duncan, K., Edwards, R.M. & Coggins, J.R. 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 (1988).
89. Helmstaedt, K., Strittmatter, A., Lipscomb, W.N. & Braus, G.H. Evolution of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase-encoding genes in the yeast *Saccharomyces cerevisiae*. *PNAS* **102**, 9784–9789 (2005).
90. Jones, D.G.L., Reusser, U. & Braus, G.H. Molecular cloning, characterization and analysis of the regulation of the *AR02* gene, encoding chorismate synthase, of *Saccharomyces cerevisiae*. *mol. Microbiol.* **5**, 2143-2152 (1991).
91. Schnappauf, G., Lipscomb, W.N. & Braus, G.H. Separation of inhibition and activation of the allosteric yeast chorismate mutase. *PNAS* **95**, 2868–2873 (1998).
92. Schnappauf, G., Krappmann, S. & Braus, G.H. Tyrosine and tryptophan act through the same binding site at the dimer interface of yeast chorismate mutase. *J. Biol. Chem.* **273**, 17012–17017 (1998).
93. Helmstaedt, K., Heinrich, G., Lipscomb, W.N. & Braus, G.H. Refined molecular hinge between allosteric and catalytic domain determines allosteric regulation and stability of fungal chorismate mutase. *PNAS* **10**, 6631–6636 (2002).
94. Reifenrath, M. & Boles, E. Engineering of hydroxymandelate synthases and the aromatic amino acid pathway enables *de novo* biosynthesis of mandelic and 4-hydroxymandelic acid with *Saccharomyces cerevisiae*. *Metab. Eng.* **45**, 246-254 (2018).
95. Dong, N.Q. & Lin, H.X. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. *J. Integr. Plant Biol.* **63**, 180-209 (2021).
96. Alappat, B. & Alappat, J. Anthocyanin Pigments: Beyond Aesthetics. *Molecules* **25**, 5500 (2020).
97. Cao, M., Gao, M., Suastegui, M., Mei, Y. & Shao, Z. Building microbial factories for the production of aromatic amino acid pathway derivatives: From commodity chemicals to plant-sourced natural products. *Metab. Eng.* **58**, 94-132 (2020).
98. Rodriguez, A., Strucko, T., Stahlhut, S.G., Kristensen, M., Svenssen, D.K., Forster, J., Nielsen, J. & Borodina, I. Metabolic engineering of yeast for fermentative production of flavonoids. *Bioresour. Technol.* **245**, 1645-1654 (2017).

99. Zhang, J. Flavonoids in grapefruit and commercial grapefruit juices concentration, distribution, and potential health benefits. *Proc. Fla. State Hort. Soc.* **120**, 288-294 (2007).
100. Jiang, H., Wood, K.V. & Morgan, J.A. Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **71**, 2962-2969 (2005).
101. Zha, J., Wu, X., Gong, G. & Koffas, M.A.G. Pathway enzyme engineering for flavonoid production in recombinant microbes. *Metab. Eng. Commun.* **9**, e00104 (2019).
102. Chemler, J.A. & Koffas, M.A.G. Metabolic engineering for plant natural product biosynthesis in microbes. *Curr. Opin. Biotechnol.* **19**, 597-605 (2008).
103. The-European-Parliament-and-the-council-of-the-European-union, Vol. L 354/34 34-50 (Official Journal of the European Union; 2008).
104. Cortez, R., Luna-Vital, D.A., Margulis, D. & Gonzalez de Mejia, E. Natural Pigments: Stabilization Methods of Anthocyanins for Food Applications. *Compr. Rev. Food Sci. Food Saf.* **16**, 180-198 (2017).
105. Yang, D., Park, S.Y., Park, Y.S., Eun, H. & Lee, S.Y. Metabolic Engineering of *Escherichia coli* for Natural Product Biosynthesis. *Trends Biotechnol.* **38**, 745-765 (2020).
106. Chen, R., Yang, S., Zhang, L. & Zhou, Y.J. Advanced Strategies for Production of Natural Products in Yeast. *iScience* **23**, 27 (2020).
107. Liu, Q., Liu, Y., Chen, Y. & Nielsen, J. Current state of aromatics production using yeast: achievements and challenges. *Curr. Opin. Biotechnol.* **65**, 65-74 (2020).
108. Delmulle, T., De Maeseneire, S.L. & De Mey, M. Challenges in the microbial production of flavonoids. *Phytochem. Rev.* **17**, 229-247 (2017).
109. Rodriguez, A., Kildegaard, K.R., Li, M., Borodina, I. & Nielsen, J. Establishment of a yeast platform strain for production of *p*-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. *Metab. Eng.* **31**, 181-188 (2015).
110. Shahidi, F. & Chandrasekara, A. Hydroxycinnamates and their *in vitro* and *in vivo* antioxidant activities. *Phytochemistry Reviews* **9**, 147-170 (2009).
111. Cheng, J.-C., Dai, F., Zhou, B., Yang, L. & Liu, Z.-L. Antioxidant activity of hydroxycinnamic acid derivatives in human low density lipoprotein: Mechanism and structure–activity relationship. *Food Chem.* **104**, 132-139 (2007).
112. Shrestha, A., Pandey, R.P. & Sohng, J.K. Biosynthesis of resveratrol and piceatannol in engineered microbial strains: achievements and perspectives. *Appl. Microbiol. Biotechnol.* **103**, 2959-2972 (2019).
113. Thapa, S.B., Pandey, R.P., Park, Y.I. & Kyung Sohng, J. Biotechnological Advances in Resveratrol Production and its Chemical Diversity. *Molecules* **24**, 2571 (2019).
114. Bourgaud, F., Hehn, A., Larbat, R., Doerper, S., Gontier, E., Kellner, S. & Matern, U. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochem. Rev.* **5**, 293-308 (2006).
115. Boerjan, W., Ralph, J. & Baucher, M. Lignin biosynthesis. *Annu. Rev. Plant Biol.* **54**, 519-546 (2003).
116. Ferrer, J.L., Austin, M.B., Stewart, C., Jr. & Noel, J.P. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.* **46**, 356-370 (2008).
117. Van Acker, S.A.B.E., Van den Berg, D.J., Tromp, M.N.J.L., Griffioen, D.H., Van Bennekom, W.P., Van der Vijgh, W.J.F. & Bast, A. Structural aspects of antioxidant activity of flavonoids. *Free Radic. Biol. Med.* **20**, 331-342 (1996).
118. Santos, C.M.M. & Silva, A.M.S. The Antioxidant Activity of Prenylflavonoids. *Molecules* **25**, 696 (2020).
119. Ferraz, C.R., Carvalho, T.T., Manchope, M.F., Artero, N.A., Rasquel-Oliveira, F.S., Fattori, V., Casagrande, R. & Verri, W.A., Jr. Therapeutic Potential of Flavonoids in Pain and Inflammation: Mechanisms of Action, Pre-Clinical and Clinical Data, and Pharmaceutical Development. *Molecules* **25**, 762 (2020).
120. Dan, W. & Dai, J. Recent developments of chalcones as potential antibacterial agents in medicinal chemistry. *Eur. J. Med. Chem.* **187**, 111980 (2020).
121. Bojic, M., Males, Z., Antolic, A., Babic, I. & Tomicic, M. Antithrombotic activity of flavonoids

and polyphenols rich plant species. *Acta Pharm* **69**, 483-495 (2019).
122. Caro-Ordieres, T., Marin-Royo, G., Opazo-Rios, L., Jimenez-Castilla, L., Moreno, J.A., Gomez-Guerrero, C. & Egido, J. The Coming Age of Flavonoids in the Treatment of Diabetic Complications. *J. Clin. Med.* **9**, 346 (2020).
123. Abbaszadeh, H., Keikhaei, B. & Mottaghi, S. A review of molecular mechanisms involved in anticancer and antiangiogenic effects of natural polyphenolic compounds. *Phytother. Res.* **33**, 2002-2014 (2019).
124. Marquez Campos, E., Stehle, P. & Simon, M.C. Microbial Metabolites of Flavan-3-Ols and Their Biological Activity. *Nutrients* **11**, 2260 (2019).
125. Zhao, C.L., Yu, Y.Q., Chen, Z.J., Wen, G.S., Wei, F.G., Zheng, Q., Wang, C.D. & Xiao, X.L. Stability-increasing effects of anthocyanin glycosyl acylation. *Food Chem.* **214**, 119-128 (2017).
126. Oteiza, P.I., Erlejman, A.G., Verstraeten, S.V., Keen, C.L. & Fraga, C.G. Flavonoid-membrane interactions: a protective role of flavonoids at the membrane surface? *Clin. Dev. Immunol.* **12**, 19-25 (2005).
127. Sinha, R., Joshi, A., Joshi, U.J., Srivastava, S. & Govil, G. Localization and interaction of hydroxyflavones with lipid bilayer model membranes: a study using DSC and multinuclear NMR. *Eur. J. Med. Chem.* **80**, 285-294 (2014).
128. Tanaka, Y., Sasaki, N. & Ohmiya, A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* **54**, 733-749 (2008).
129. Zhao, C.L., Chen, Z.J., Bai, X.S., Ding, C., Long, T.J., Wei, F.G. & Miao, K.R. Structure-activity relationships of anthocyanidin glycosylation. *Mol. Divers.* **18**, 687-700 (2014).
130. Isogai, S., Okahashi, N., Asama, R., Nakamura, T., Hasunuma, T., Matsuda, F., Ishii, J. & Kondo, A. Synthetic production of prenylated naringenins in yeast using promiscuous microbial prenyltransferases. *Metab. Eng. Commun.* **12**, e00169 (2021).
131. Kretzschmar, G., Zierau, O., Wober, J., Tischer, S., Metz, P. & Vollmer, G. Prenylation has a compound specific effect on the estrogenicity of naringenin and genistein. *J. Steroid Biochem. Mol. Biol.* **118**, 1-6 (2010).
132. Koopman, F., Beekwilder, J. & Crimi, B. *De novo* production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **11**, 155 (2012).
133. Palmer, C.M., Miller, K.K., Nguyen, A. & Alper, H.S. Engineering 4-coumaroyl-CoA derived polyketide production in *Yarrowia lipolytica* through a beta-oxidation mediated strategy. *Metab. Eng.* **57**, 174-181 (2020).
134. Liu, Q., Yu, T., Li, X., Chen, Y., Campbell, K., Nielsen, J. & Chen, Y. Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. *Nat. Commun.* **10**, 4976 (2019).
135. Li, M., Schneider, K., Kristensen, M., Borodina, I. & Nielsen, J. Engineering yeast for high-level production of stilbenoid antioxidants. *Sci. Rep.* **6**, 36827 (2016).
136. Sáez-Sáez, J., Wang, G., Marella, E.R., Sudarsan, S., Cernuda Pastor, M. & Borodina, I. Engineering the oleaginous yeast *Yarrowia lipolytica* for high-level resveratrol production. *Metab. Eng.* **62**, 51-61 (2020).
137. Liu, T., Liu, Y., Li, L., Liu, X., Guo, Z., Cheng, J., Zhu, X., Lu, L., Zhang, J., Fan, G., Xie, N., Lu, J. & Jiang, H. *De Novo* Biosynthesis of Polydatin in *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* **69**, 5917-5925 (2021).
138. Eichenberger, M., Lehka, B.J., Folly, C., Fischer, D., Martens, S., Simon, E. & Naesby, M. Metabolic engineering of *Saccharomyces cerevisiae* for *de novo* production of dihydrochalcones with known antioxidant, antidiabetic, and sweet tasting properties. *Metab. Eng.* **39**, 80-89 (2017).
139. Yin, Y., Li, Y., Jiang, D., Zhang, X., Gao, W. & Liu, C. *De novo* biosynthesis of liquiritin in *Saccharomyces cerevisiae*. *Acta Pharm. Sin. B.* **10**, 711-721 (2020).
140. Lv, Y., Marsafari, M., Koffas, M., Zhou, J. & Xu, P. Optimizing Oleaginous Yeast Cell Factories for Flavonoids and Hydroxylated Flavonoids Biosynthesis. *ACS Synth. Biol.* **8**, 2514-2523 (2019).
141. Levisson, M., Araya-Cloutier, C., de Bruijn, W.J.C., van der Heide, M., Salvador Lopez, J.M., Daran, J.M., Vincken, J.P. & Beekwilder, J. Toward Developing a Yeast Cell Factory for the Production of Prenylated Flavonoids. *J. Agric. Food Chem.* **67**, 13478-13486 (2019).

142. Yang, J., Liang, J., Shao, L., Liu, L., Gao, K., Zhang, J.L., Sun, Z., Xu, W., Lin, P., Yu, R. & Zi, J. Green production of silybin and isosilybin by merging metabolic engineering approaches and enzymatic catalysis. *Metab. Eng.* **59**, 44-52 (2020).
143. Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., Kui, L., Cheng, X., Ruan, J., Fan, W., Chen, J., Long, G., Zhao, Y., Cai, J., Wang, W., Ma, Y., Dong, Y., Yang, S. & Jiang, H. Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. *Nat. Commun.* **9**, 448 (2018).
144. Liu, X., Cheng, J., Zhu, X., Zhang, G., Yang, S., Guo, X., Jiang, H. & Ma, Y. De Novo Biosynthesis of Multiple Pinocembrin Derivatives in *Saccharomyces cerevisiae*. *ACS Synth Biol* **9**, 3042-3051 (2020).
145. Duan, L., Ding, W., Liu, X., Cheng, X., Cai, J., Hua, E. & Jiang, H. Biosynthesis and engineering of kaempferol in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **16**, 165 (2017).
146. Bergman, A., Siewers, V., Nielsen, J. & Chen, Y. Functional expression and evaluation of heterologous phosphoketolases in *Saccharomyces cerevisiae*. *AMB Express* **6**, 115 (2016).
147. Curran, K.A., Leavitt, J.M., Karim, A.S. & Alper, H.S. Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metab. Eng.* **15**, 55-66 (2013).
148. Gold, N.D., Gowen, C.M., Lussier, F.X., Cautha, S.C., Mahadevan, R. & Martin, V.J. Metabolic engineering of a tyrosine-overproducing yeast platform using targeted metabolomics. *Microb. Cell Fact.* **14**, 73 (2015).
149. Shi, S., Chen, Y., Siewers, V. & Nielsen, J. Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. *mBio* **5**, e01130-01114 (2014).
150. Chen, X., Yang, X., Shen, Y., Hou, J. & Bao, X. Screening Phosphorylation Site Mutations in Yeast Acetyl-CoA Carboxylase Using Malonyl-CoA Sensor to Improve Malonyl-CoA-Derived Product. *Front. Microbiol.* **9**, 47 (2018).
151. Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.K. & Keasling, J.D. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab. Eng.* **9**, 160-168 (2007).
152. Vannelli, T., Wei Qi, W., Sweigard, J., Gatenby, A.A. & Sariaslani, F.S. Production of *p*-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metab. Eng.* **9**, 142-151 (2007).
153. Abe, I., Watanabe, T. & Noguchi, H. Enzymatic formation of long-chain polyketide pyrones by plant type III polyketide synthases. *Phytochemistry* **65**, 2447-2453 (2004).
154. Xiong, D., Lu, S., Wu, J., Liang, C., Wang, W., Wang, W., Jin, J.M. & Tang, S.Y. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metab. Eng.* **40**, 115-123 (2017).
155. Gu, Y., Ma, J., Zhu, Y., Ding, X. & Xu, P. Engineering *Yarrowia lipolytica* as a Chassis for *De Novo* Synthesis of Five Aromatic-Derived Natural Products and Chemicals. *ACS Synth. Biol.* **9**, 2096-2106 (2020).
156. Gao, S., Lyu, Y., Zeng, W., Du, G., Zhou, J. & Chen, J. Efficient Biosynthesis of (2*S*)-Naringenin from *p*-Coumaric Acid in *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* **68**, 1015-1021 (2020).
157. Lyu, X., Ng, K.R., Lee, J.L., Mark, R. & Chen, W.N. Enhancement of Naringenin Biosynthesis from Tyrosine by Metabolic Engineering of *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* **65**, 6638-6646 (2017).
158. Suastegui, M., Guo, W., Feng, X. & Shao, Z. Investigating strain dependency in the production of aromatic compounds in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **113**, 2676-2685 (2016).
159. Lyu, X., Zhao, G., Ng, K.R., Mark, R. & Chen, W.N. Metabolic Engineering of *Saccharomyces cerevisiae* for *De Novo* Production of Kaempferol. *J. Agric. Food Chem.* **67**, 5596-5606 (2019).
160. Lv, Y., Edwards, H., Zhou, J. & Xu, P. Combining 26s rDNA and the Cre-loxP System for Iterative Gene Integration and Efficient Marker Curation in *Yarrowia lipolytica*. *ACS Synth. Biol.* **8**, 568-576 (2019).
161. Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I. & Nielsen, J. *De novo* producti-

on of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* **32**, 1-11 (2015).

162. Trantas, E., Panopoulos, N. & Ververidis, F. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metab. Eng.* **11**, 355-366 (2009).
163. Hwang, E.I., Kaneko, M., Ohnishi, Y. & Horinouchi, S. Production of plant-specific flavanones by *Escherichia coli* containing an artificial gene cluster. *Appl. Environ. Microbiol.* **69**, 2699-2706 (2003).
164. Jiang, C., Liu, X., Chen, X., Cai, Y., Zhuang, Y., Liu, T., Zhu, X., Wang, H., Liu, Y., Jiang, H. & Wang, W. Raising the production of phloretin by alleviation of by-product of chalcone synthase in the engineered yeast. *Sci. China Life Sci* **63**, 1734–1743 (2020).
165. Zuurbier, K.W.M., Fung, S., Scheffer, J.J.C. & Verpoorte, R. Assay of chalcone synthase activity by high-performance liquid chromatography. *Phytochemistry* **34**, 1225-1229 (1993).
166. Flikweert, M.T., de Swaaf, M., van Dijken, J.P. & Pronk, J.T. Growth requirements of pyruvate-decarboxylase-negative *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **174**, 73-79 (1999).
167. Flikweert, M.T., Van Der Zanden, L., Janssen, W.M., Steensma, H.Y., Van Dijken, J.P. & Pronk, J.T. Pyruvate decarboxylase: An indispensable enzyme for growth of *Saccharomyces cerevisiae* on glucose. *Yeast* **12**, 247-257 (1996).
168. Flikweert, M.T., van Dijken, J.P. & Pronk, J.T. Metabolic responses of pyruvate decarboxylase-negative *Saccharomyces cerevisiae* to glucose excess. *Appl Environ Microbiol* **63**, 399-404 (1997).
169. Ter Schure, E.G., Flikweert, M.T., Van Dijken, J.P., Pronk, J.T. & Verrips, C.T. Pyruvate Decarboxylase Catalyzes Decarboxylation of Branched-Chain 2-Oxo Acids but Is Not Essential for Fusel Alcohol Production by *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol* **64**, 1303–1307 (1998).
170. Xu, P., Hua, D. & Ma, C. Microbial transformation of propenylbenzenes for natural flavour production. *Trends Biotechnol.* **25**, 571-576 (2007).
171. Eikani, M.H., Golmohammad, F., Rowshanzamir, S. & Mirza, M. Recovery of water-soluble constituents of rose oil using simultaneous distillation-extraction. *Flavour Frag. J.* **20**, 555-558 (2005).
172. Kim, B., Cho, B.R. & Hahn, J.S. Metabolic engineering of *Saccharomyces cerevisiae* for the production of 2-phenylethanol via Ehrlich pathway. *Biotechnol. Bioeng.* **111**, 115-124 (2014).
173. Longo, M.A. & Sanromán, M.A. Production of Food Aroma Compounds Microbial and enzymatic methodologies. *Food Technol. Biotechnol.* **44**, 335–353 (2006).
174. Cordente, A.G., Schmidt, S., Beltran, G., Torija, M.J. & Curtin, C.D. Harnessing yeast metabolism of aromatic amino acids for fermented beverage bioflavouring and bioproduction. *Appl. Microbiol. Biotechnol.* **103**, 4325-4336 (2019).
175. Eshkol, N., Sendovski, M., Bahalul, M., Katz-Ezov, T., Kashi, Y. & Fishman, A. Production of 2-phenylethanol from L-phenylalanine by a stress tolerant *Saccharomyces cerevisiae* strain. *J. Appl. Microbiol.* **106**, 534-542 (2009).
176. Etschmann, M.M. & Schrader, J. An aqueous-organic two-phase bioprocess for efficient production of the natural aroma chemicals 2-phenylethanol and 2-phenylethylacetate with yeast. *Appl. Microbiol. Biotechnol.* **71**, 440-443 (2006).
177. Stark, D., Münch, T., Sonnleitner, B., Marison, I.W. & von Stockar, U. Extractive bioconversion of 2-phenylethanol from L-phenylalanine by *Saccharomyces cerevisiae*. *Biotechnol. Prog.* **18**, 514–523 (2002).
178. Stark, D., Zala, D., Münch, T., Sonnleitner, B., Marison, I.W. & von Stockar, U. Inhibition aspects of the bioconversion of L-phenylalanine to 2-phenylethanol by *Saccharomyces cerevisiae*. *Enzyme Microb. Technol.* **32**, 212–223 (2003).
179. Jones, D.G.L., Reusser, U. & Braus, G.H. Molecular cloning, characterization and analysis of the regulation of the *AR02* gene, encoding chorismate synthase, of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **5**, 2143-2153 (1991).
180. Dickinson, J.R., Salgado, L.E. & Hewlins, M.J. The catabolism of amino acids to long chain and complex alcohols in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 8028-8034 (2003).

181. Wang, Y., Zhang, H., Lu, X., Zong, H. & Zhuge, B. Advances in 2-phenylethanol production from engineered microorganisms. *Biotechnol. Adv.* **37**, 403-409 (2019).
182. Entian, K. & Kötter, P. 25 Yeast Genetic Strain and Plasmid Collections. *Methods Microbiol.* **36**, 629-666 (2007).
183. Verduyn, C., Postma, E., Scheffers, W.A. & van Dijken, J.P. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501-517 (1992).
184. Pronk, J.T. Auxotrophic yeast strains in fundamental and applied research. *Appl. Environ. Microbiol.* **68**, 2095-2100 (2002).
185. Luttik, M.A., Kötter, P., Salomons, F.A., van der Klei, I.J., van Dijken, J.P. & Pronk, J.T. The *Saccharomyces cerevisiae ICL2* gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-CoA metabolism. *J. Bacteriol.* **182**, 7007-7013 (2000).
186. Lopes, C.A., Barrio, E. & Querol, A. Natural hybrids of *S. cerevisiae x S. kudriavzevii* share alleles with European wild populations of *Saccharomyces kudriavzevii*. *FEMS Yeast Res* **10**, 412-421 (2010).
187. Libkind, D., Hittinger, C.T., Valerio, E., Goncalves, C., Dover, J., Johnston, M., Goncalves, P. & Sampaio, J.P. Microbe domestication and the identification of the wild genetic stock of lager--brewing yeast. *Proc Natl Acad Sci U S A* **108**, 14539-14544 (2011).
188. Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J. & Daran, J.M. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **15**, fov004 (2015).
189. Papapetridis, I., van Dijk, M., van Maris, A.J.A. & Pronk, J.T. Metabolic engineering strategies for optimizing acetate reduction, ethanol yield and osmotolerance in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **10**, 107 (2017).
190. Lee, M.E., DeLoache, W.C., Cervantes, B. & Dueber, J.E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth Biol* **4**, 975-986 (2015).
191. Boonekamp, F.J., Dashko, S., van den Broek, M., Gehrmann, T., Daran, J.M. & Daran-Lapujade, P. The Genetic Makeup and Expression of the Glycolytic and Fermentative Pathways Are Highly Conserved Within the *Saccharomyces* Genus. *Front. Genet.* **9**, 504 (2018).
192. DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J. & Church, G.M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336-4343 (2013).
193. Kuijpers, N.G.A., Solis-Escalante, D., Bosman, L., van den Broek, M., Pronk, J.T., Daran, J. & Daran-Lapujade, P. A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb. Cell Fact.* **12**, 47 (2013).
194. de Kok, S., Yilmaz, D., Suir, E., Pronk, J.T., Daran, J.M. & van Maris, A.J. Increasing free--energy (ATP) conservation in maltose-grown *Saccharomyces cerevisiae* by expression of a heterologous maltose phosphorylase. *Metab. Eng.* **13**, 518-526 (2011).
195. Mans, R., Wijsman, M., Daran-Lapujade, P. & Daran, J.M. A protocol for introduction of multiple genetic modifications in *Saccharomyces cerevisiae* using CRISPR/Cas9. *FEMS Yeast Res* **18**, foy063 (2018).
196. Papapetridis, I., Goudriaan, M., Vazquez Vitali, M., de Keijzer, N.A., van den Broek, M., van Maris, A.J.A. & Pronk, J.T. Optimizing anaerobic growth rate and fermentation kinetics in *Saccharomyces cerevisiae* strains expressing Calvin-cycle enzymes for improved ethanol yield. *Biotechnol. Biofuels* **11**, 17 (2018).
197. Strucko, T., Buron, L.D., Jarczynska, Z.D., Nodvig, C.S., Molgaard, L., Halkier, B.A. & Mortensen, U.H. CASCADE, a platform for controlled gene amplification for high, tunable and selection-free gene expression in yeast. *Sci. Rep.* **7**, 41431 (2017).
198. Mans, R., Hassing, E.J., Wijsman, M., Giezekamp, A., Pronk, J.T., Daran, J.M. & van Maris, A.J.A. A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **17**, fox085 (2017).
199. Wijsman, M., Swiat, M.A., Marques, W.L., Hettinga, J.K., van den Broek, M., Torre Cortes, P., Mans, R., Pronk, J.T., Daran, J.M. & Daran-Lapujade, P. A toolkit for rapid CRISPR-*Sp*Cas9

assisted construction of hexose-transport-deficient *Saccharomyces cerevisiae* strains. *FEMS Yeast Res.* **19**, foy107 (2019).

200. Shen, L., Nishimura, Y., Matsuda, F., Ishii, J. & Kondo, A. Overexpressing enzymes of the Ehrlich pathway and deleting genes of the competing pathway in *Saccharomyces cerevisiae* for increasing 2-phenylethanol production from glucose. *J. Biosci. Bioeng.* **122**, 34-39 (2016).
201. Shin, S.Y., Han, N.S., Park, Y.C., Kim, M.D. & Seo, J.H. Production of resveratrol from *p*-coumaric acid in recombinant *Saccharomyces cerevisiae* expressing 4-coumarate:coenzyme A ligase and stilbene synthase genes. *Enzyme Microb. Technol.* **48**, 48-53 (2011).
202. Gottardi, M., Grun, P., Bode, H.B., Hoffmann, T., Schwab, W., Oreb, M. & Boles, E. Optimisation of trans-cinnamic acid and hydrocinnamyl alcohol production with recombinant *Saccharomyces cerevisiae* and identification of cinnamyl methyl ketone as a by-product. *FEMS Yeast Res.* **17**, fox091 (2017).
203. Gottardi, M., Knudsen, J.D., Prado, L., Oreb, M., Branduardi, P. & Boles, E. *De novo* biosynthesis of *trans*-cinnamic acid derivatives in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **101**, 4883-4893 (2017).
204. Knijnenburg, T.A., Daran, J.M., van den Broek, M.A., Daran-Lapujade, P.A., de Winde, J.H., Pronk, J.T., Reinders, M.J. & Wessels, L.F. Combinatorial effects of environmental parameters on transcriptional regulation in *Saccharomyces cerevisiae*: a quantitative analysis of a compendium of chemostat-based transcriptome data. *BMC Genomics* **10**, 53 (2009).
205. Knijnenburg, T.A., de Winde, J.H., Daran, J.M., Daran-Lapujade, P., Pronk, J.T., Reinders, M.J. & Wessels, L.F. Exploiting combinatorial cultivation conditions to infer transcriptional regulation. *BMC Genomics* **8**, 25 (2007).
206. Piper, M.D., Daran-Lapujade, P., Bro, C., Regenberg, B., Knudsen, S., Nielsen, J. & Pronk, J.T. Reproducibility of oligonucleotide microarray transcriptome analyses. An interlaboratory comparison using chemostat cultures of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**, 37001-37008 (2002).
207. Zelle, R.M., Trueheart, J., Harrison, J.C., Pronk, J.T. & van Maris, A.J. Phosphoenolpyruvate carboxykinase as the sole anaplerotic enzyme in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **76**, 5383-5389 (2010).
208. Bruckner, C., Oreb, M., Kunze, G., Boles, E. & Tripp, J. An expanded enzyme toolbox for production of cis, cis-muconic acid and other shikimate pathway derivatives in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **18**, foy017 (2018).
209. Liu, C., Zhang, K., Cao, W., Zhang, G., Chen, G., Yang, H., Wang, Q., Liu, H., Xian, M. & Zhang, H. Genome mining of 2-phenylethanol biosynthetic genes from *Enterobacter sp. C*GMCC 5087 and heterologous overproduction in *Escherichia coli*. *Biotechnol. Biofuels* **11**, 305 (2018).
210. Wang, Z., Jiang, M., Guo, X., Liu, Z. & He, X. Reconstruction of metabolic module with improved promoter strength increases the productivity of 2-phenylethanol in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **17**, 60 (2018).
211. de Deken, H.R. The crabtree effect: a regulatory system in yeast. *J. gen. Microbiol.* **44**, 149-156 (1966).
212. Van Hoek, P., van Dijken, J.P. & Pronk, J.T. Effect of Specific Growth Rate on Fermentative Capacity of Baker's Yeast. *Appl. Environ. Microbiol.* **64**, 4226–4233 (1998).
213. Hawkins, K.M. & Smolke, C.D. Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. *Nat. Chem. Biol.* **4**, 564-573 (2008).
214. Fleige, C., Kroll, J. & Steinbuchel, A. Establishment of an alternative phosphoketolase-dependent pathway for fructose catabolism in *Ralstonia eutropha* H16. *Appl. Microbiol. Biotechnol.* **91**, 769-776 (2011).
215. Suzuki, R., Katayama, T., Kim, B.J., Wakagi, T., Shoun, H., Ashida, H., Yamamoto, K. & Fushinobu, S. Crystal structures of phosphoketolase: thiamine diphosphate-dependent dehydration mechanism. *J. Biol. Chem.* **285**, 34279-34287 (2010).
216. Boles, E., Heinisch, J. & Zimmermann, F.K. Different signals control the activation of glycolysis in the yeast *Saccharomyces cerevisiae*. *Yeast* **9**, 761-770 (1993).
217. Boles, E., Göhlmann, H.W.H. & Zimmermann, F.K. Cloning of a second gene encoding 6-

phosphofructo-2-kinase in yeast, and characterization of mutant strains without fructose-2,6-bisphosphate. *Mol. Microbiol.* **20**, 65-76 (1996).

218. Romagnoli, G., Knijnenburg, T.A., Liti, G., Louis, E.J., Pronk, J.T. & Daran, J.M. Deletion of the *Saccharomyces cerevisiae ARO8* gene, encoding an aromatic amino acid transaminase, enhances phenylethanol production from glucose. *Yeast* **32**, 29-45 (2015).
219. Hazelwood, L.A., Tai, S.L., Boer, V.M., de Winde, J.H., Pronk, J.T. & Daran, J.M. A new physiological role for Pdr12p in *Saccharomyces cerevisiae*: export of aromatic and branched-chain organic acids produced in amino acid catabolism. *FEMS Yeast Res.* **6**, 937-945 (2006).
220. Ingram, L.O. & Buttke, T.M. Effects of alcohols on micro-organisms. *Adv. Microb. Physiol.* **25**, 253–300 (1984).
221. Seward, R., Willetts, J.C., Dinsdale, M.W. & Lloyd, D. The effects of ethanol, hexan-1-ol, and 2-phenylethanol on cider yeast growth, viability, and energy status; synergistic inhibition. *J. Inst. Brew.* **102**, 439-443 (1996).
222. Lester, G. Inhibition of growth, synthesis, and permeability in *Neurospora crassa* by phenethyl alcohol. *J. Bacteriol.* **90**, 29-37 (1965).
223. Chubukov, V., Mukhopadhyay, A., Petzold, C.J., Keasling, J.D. & Martin, H.G. Synthetic and systems biology for microbial production of commodity chemicals. *NPJ Syst. Bio.l Appl.* **2**, 16009 (2016).
224. Escalera-Fanjul, X., Quezada, H., Riego-Ruiz, L. & Gonzalez, A. Whole-Genome Duplication and Yeast's Fruitful Way of Life. *Trends Genet.* **35**, 42-54 (2019).
225. Kuijpers, N.G., Solis-Escalante, D., Luttik, M.A., Bisschops, M.M., Boonekamp, F.J., van den Broek, M., Pronk, J.T., Daran, J.M. & Daran-Lapujade, P. Pathway swapping: Toward modular engineering of essential cellular processes. *PNAS* **113**, 15060-15065 (2016).
226. Postma, E.D., Dashko, S., van Breemen, L., Taylor Parkins, S.K., van den Broek, M., Daran, J.M. & Daran-Lapujade, P. A supernumerary designer chromosome for modular in vivo pathway assembly in S*accharomyces cerevisiae*. *Nucleic Acids Res.* **49**, 1769-1783 (2021).
227. Hassing, E.J., de Groot, P.A., Marquenie, V.R., Pronk, J.T. & Daran, J.G. Connecting central carbon and aromatic amino acid metabolisms to improve *de novo* 2-phenylethanol production in *Saccharomyces cerevisiae*. *Metab. Eng.* **56**, 165-180 (2019).
228. Mitchell, L.A. & Boeke, J.D. Circular permutation of a synthetic eukaryotic chromosome with the telomerator. *PNAS* **111**, 17003-17010 (2014).
229. Boonekamp, F.J., Dashko, S., Duiker, D., Gehrmann, T., van den Broek, M., den Ridder, M., Pabst, M., Robert, V., Abeel, T., Postma, E.D., Daran, J.M. & Daran-Lapujade, P. Design and Experimental Evaluation of a Minimal, Innocuous Watermarking Strategy to Distinguish Near-Identical DNA and RNA Sequences. *ACS Synth. Biol.* **9**, 1361-1375 (2020).
230. Postma, E., Scheffers, W.A. & van Dijken, J.P. Kinetics of growth and glucose transport in glucose-limited chemostat cultures of *Saccharomyces cerevisiae* CBS 8066. *Yeast* **5**, 159-165 (1989).
231. McClintock, B. The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes. *Genetics* **23**, 315 (1938).
232. Xie, Z.X., Li, B.Z., Mitchell, L.A., Wu, Y., Qi, X., Jin, Z., Jia, B., Wang, X., Zeng, B.X., Liu, H.M., Wu, X.L., Feng, Q., Zhang, W.Z., Liu, W., Ding, M.Z., Li, X., Zhao, G.R., Qiao, J.J., Cheng, J.S., Zhao, M., Kuang, Z., Wang, X., Martin, J.A., Stracquadanio, G., Yang, K., Bai, X., Zhao, J., Hu, M.L., Lin, Q.H., Zhang, W.Q., Shen, M.H., Chen, S., Su, W., Wang, E.X., Guo, R., Zhai, F., Guo, X.J., Du, H.X., Zhu, J.Q., Song, T.Q., Dai, J.J., Li, F.F., Jiang, G.Z., Han, S.L., Liu, S.Y., Yu, Z.C., Yang, X.N., Chen, K., Hu, C., Li, D.S., Jia, N., Liu, Y., Wang, L.T., Wang, S., Wei, X.T., Fu, M.Q., Qu, L.M., Xin, S.Y., Liu, T., Tian, K.R., Li, X.N., Zhang, J.H., Song, L.X., Liu, J.G., Lv, J.F., Xu, H., Tao, R., Wang, Y., Zhang, T.T., Deng, Y.X., Wang, Y.R., Li, T., Ye, G.X., Xu, X.R., Xia, Z.B., Zhang, W., Yang, S.L., Liu, Y.L., Ding, W.Q., Liu, Z.N., Zhu, J.Q., Liu, N.Z., Walker, R., Luo, Y., Wang, Y., Shen, Y., Yang, H., Cai, Y., Ma, P.S., Zhang, C.T., Bader, J.S., Boeke, J.D. & Yuan, Y.J. "Perfect" designer chromosome V and behavior of a ring derivative. *Science* **355**, 1046 (2017).
233. Walker, R.S.K. in Institute for Bioengineering (University of Edinburgh, Edinburgh; 2017).
234. Aylon, Y. & Kupiec, M. New insights into the mechanism of homologous recombination in

yeast. *Mutat. Res.* **566**, 231-248 (2004).
235. Pâques, F. & Haber, J.E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol.* **63**, 349–404 (1999).
236. Zhao, X., Wei, C., Li, J., Xing, P., Li, J., Zheng, S. & Chen, X. Cell cycle-dependent control of homologous recombination. *Acta Biochim. Biophys. Sin.* **49**, 655-668 (2017).
237. Kuijpers, N.G., Solis-Escalante, D., Luttik, M.A., Bisschops, M.M., Boonekamp, F.J., van den Broek, M., Pronk, J.T., Daran, J.-M. & Daran-Lapujade, P. Pathway swapping: Toward modular engineering of essential cellular processes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15060-15065 (2016).
238. Hu, C., Jiang, P., Xu, J., Wu, Y. & Huang, W. 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 (2003).
239. Lütke-Eversloh, T. & Stephanopoulos, G. Feedback inhibition of chorismate mutase/prephenate dehydrogenase (TyrA) of *Escherichia coli:* generation and characterization of tyrosine-insensitive mutants. *Appl. Environ. Microbiol.* **71**, 7224-7228 (2005).
240. Zhou, H., Liao, X., Wang, T., Du, G. & Chen, J. Enhanced L-phenylalanine biosynthesis by co-expression of pheA^{fbr} and aroF^{wt}. *Bioresour. Technol.* **101**, 4151-4156 (2010).
241. Keren, L., Zackay, O., Lotan-Pompan, M., Barenholz, U., Dekel, E., Sasson, V., Aidelberg, G., Bren, A., Zeevi, D., Weinberger, A., Alon, U., Milo, R. & Segal, E. Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* **9**, 701 (2013).
242. Curran, K.A., Karim, A.S., Gupta, A. & Alper, H.S. Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng.* **19**, 88-97 (2013).
243. Wei, L., Wang, Z., Zhang, G. & Ye, B. Characterization of Terminators in *Saccharomyces cerevisiae* and an Exploration of Factors Affecting Their Strength. *Chembiochem* **18**, 2422-2427 (2017).
244. Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., Moriya, H. & Matsuyama, T. A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a "terminatome" toolbox. *ACS Synth. Biol.* **2**, 337-347 (2013).
245. Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B. & Kinzler, K.W. Characterization of the yeast transcriptome. *Cell* **88**, 243–251 (1997).
246. Flagfeldt, D.B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* **26**, 545-551 (2009).
247. Kuijpers, N.G., Chroumpi, S., Vos, T., Solis-Escalante, D., Bosman, L., Pronk, J.T., Daran, J.M. & Daran-Lapujade, P. One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **13**, 769-781 (2013).
248. Mikkelsen, M.D., Buron, L.D., Salomonsen, B., Olsen, C.E., Hansen, B.G., Mortensen, U.H. & Halkier, B.A. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* **14**, 104-111 (2012).
249. Coradini, A.L.V., Hull, C.B. & Ehrenreich, I.M. Building genomes to understand biology. *Nat. Commun.* **11**, 6177 (2020).
250. Zhang, W., Mitchell, L.A., Bader, J.S. & Boeke, J.D. Synthetic Genomes. *Annu. Rev. Biochem.* **89**, 77-101 (2020).
251. Pretorius, I.S. & Boeke, J.D. Yeast 2.0-connecting the dots in the construction of the world's first functional synthetic eukaryotic genome. *FEMS Yeast Res.* **18**, foy032 (2018).
252. Richardson, S.M., Mitchell, L.A., Stracquadanio, G., Yang, K., Dymond, J.S., DiCarlo, J.E., Lee, D., Huang, C.L.V., Chandrasegaran, S., Cai, Y., Boeke, J.D. & Bader, J.S. Design of a synthetic yeast genome. *Science* **355**, 1040–1044 (2017).
253. Clarke, L. & Carbon, J. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**, 504-509 (1980).
254. Hieter, P., Mann, C., Snyder, M. & Davis, R.W. Mitotic stability of yeast chromosomes: A colo-

ny color assay that measures nondisjunction and chromosome loss. *Cell* **40**, 381-392 (1985).
255. Murray, A.W. & Szostak, J.W. Construction of artificial chromosomes in yeast. *Nature* **305**, 189-193 (1983).
256. Sleister, H.M., Mills, K.A., BIackwell, S.E., Killary, A.M., Murray, J.C. & Malone, R.E. Construction of a human chromosome 4 YAC pool and analysis of artificial chromosome stability. *Nucleic Acids Res.* **20**, 3419-3425 (1992).
257. Ebrahimi, H. & Donaldson, A.D. Release of yeast telomeres from the nuclear periphery is triggered by replication and maintained by suppression of Ku-mediated anchoring. *Genes Dev.* **22**, 3363-3374 (2008).
258. Straatman, K.R. & Louis, E.J. Localization of telomeres and telomere-associated proteins in telomerase-negative *Saccharomyces cerevisiae*. *Chromosome Res.* **15**, 1033-1050 (2007).
259. Taddei, A. & Gasser, S.M. Structure and function in the budding yeast nucleus. *Genetics* **192**, 107-129 (2012).
260. Wellinger, R.J. & Zakian, V.A. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: beginning to end. *Genetics* **191**, 1073-1105 (2012).
261. Hamperl, S. & Cimprich, K.A. Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell* **167**, 1455-1467 (2016).
262. Langenberg, A., Bink, F.J., Wolff, L., Walter, S., von Wallbrunn, C., Grossmann, M., Heinisch, J.J. & Schmitz, H. Glycolytic functions are conserved in the genome of the wine yeast *Hanseniaspora uvarum*, and pyruvate kinase limits its capacity for alcoholic fermentation. *Appl. Environ. Microbiol.* **83**, e01580-01517 (2017).
263. Chen, B., Lee, H.L., Heng, Y.C., Chua, N., Teo, W.S., Choi, W.J., Leong, S.S.J., Foo, J.L. & Chang, M.W. Synthetic biology toolkits and applications in *Saccharomyces cerevisiae*. *Biotechnol. Adv.* **36**, 1870-1881 (2018).
264. Machens, F., Balazadeh, S., Mueller-Roeber, B. & Messerschmidt, K. Synthetic Promoters and Transcription Factors for Heterologous Protein Expression in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* **5**, 63 (2017).
265. Solis-Escalante, D., Kuijpers, N.G., Barrajon-Simancas, N., van den Broek, M., Pronk, J.T., Daran, J.M. & Daran-Lapujade, P. A Minimal Set of Glycolytic Genes Reveals Strong Redundancies in *Saccharomyces cerevisiae* Central Metabolism. *Eukaryot. Cell* **14**, 804-816 (2015).
266. Gorter de Vries, A.R., Knibbe, E., van Roosmalen, R., van den Broek, M., de la Torre Cortes, P., O'Herne, S.F., Vijverberg, P.A., El Masoudi, A., Brouwers, N., Pronk, J.T. & Daran, J.G. Improving Industrially Relevant Phenotypic Traits by Engineering Chromosome Copy Number in *Saccharomyces pastorianus*. *Front. Genet.* **11**, 518 (2020).
267. Falcone Ferreyra, M.L., Rius, S.P. & Casati, P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* **3**, 222 (2012).
268. Das, S.K., Reddy, K.A. & Mukkanti, K. Total synthesis of phenylpropanoid glycosides, grayanoside A and syringalide B, through a common intermediate. *Carbohydr. Res.* **342**, 2309-2315 (2007).
269. Trantas, E.A., Koffas, M.A., Xu, P. & Ververidis, F. When plants produce not enough or at all: metabolic engineering of flavonoids in microbial hosts. *Front. Plant Sci.* **6**, 7 (2015).
270. Bhan, N., Xu, P. & Koffas, M.A.G. Pathway and protein engineering approaches to produce novel and commodity small molecules. *Curr. Opin. Biotechnol.* **24**, 1137-1143 (2013).
271. Liu, X., Lin, J., Hu, H., Zhou, B. & Zhu, B. *De novo* biosynthesis of resveratrol by site-specific integration of heterologous genes in *Escherichia coli*. *FEMS Microbiol. Lett.* **363**, fnw061 (2016).
272. Hartmann, M., Schneider, T.R., Pfeil, A., Heinrich, G., Lipscomb, W.N. & Braus, G.H. Evolution of feedback-inhibited beta/alpha barrel isoenzymes by gene duplication and a single mutation. *PNAS* **100**, 862–867 (2003).
273. Krappmann, S., Lipscomb, W.N. & Braus, G.H. Coevolution of transcriptional and allosteric regulation at the chorismate metabolic branch point of *Saccharomyces cerevisiae*. *PNAS* **97**, 13585-13590 (2000).
274. Oud, B., Flores, C.L. & Gancedo, C. An internal deletion in *MTH1* enables growth on glucose of pyruvate-decarboxylase negative, non-fermentative *Saccharomyces cerevisiae*. *Microb. Cell*

Fact. **11**, 131 (2012).

275. Salazar, A.N., Gorter de Vries, A.R., van den Broek, M., Wijsman, M., de la Torre Cortes, P., Brickwedde, A., Brouwers, N., Daran, J.G. & Abeel, T. Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113-7D. *FEMS Yeast Res.* **17**, fox074 (2017).
276. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).
277. Wronska, A.K., Haak, M.P., Geraats, E., Bruins Slot, E., van den Broek, M., Pronk, J.T. & Daran, J.M. Exploiting the Diversity of *Saccharomycotina* Yeasts To Engineer Biotin-Independent Growth of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **86**, e00270-00220 (2020).
278. Inokuma, K., Ishii, J., Hara, K.Y., Mochizuki, M., Hasunuma, T. & Kondo, A. Complete Genome Sequence of *Kluyveromyces marxianus* NBRC1777, a Nonconventional Thermotolerant Yeast. *Genome Announc.* **3**, e00389-00315 (2015).
279. Juergens, H., Varela, J.A., Gorter de Vries, A.R., Perli, T., Gast, V.J.M., Gyurchev, N.Y., Rajkumar, A.S., Mans, R., Pronk, J.T., Morrissey, J.P. & Daran, J.G. Genome editing in *Kluyveromyces* and *Ogataea* yeasts using a broad-host-range Cas9/gRNA co-expression plasmid. *FEMS Yeast Res.* **18**, foy012 (2018).
280. Wiedemann, B. & Boles, E. Codon-optimized bacterial genes improve L-Arabinose fermentation in recombinant *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **74**, 2043-2050 (2008).
281. Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D.C. & Jahn, D. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* **33**, W526-531 (2005).
282. Pappas, K.M., Kouvelis, V.N., Saunders, E., Brettin, T.S., Bruce, D., Detter, C., Balakireva, M., Han, C.S., Savvakis, G., Kyrpides, N.C. & Typas, M.A. Genome sequence of the ethanol-producing *Zymomonas mobilis subsp. mobilis* lectotype strain ATCC 10988. *J. Bacteriol.* **193**, 5051-5052 (2011).
283. van Zyl, L.J., Schubert, W.D. & Tuffin, M.I. Structure and functional characterization of pyruvate decarboxylase from *Gluconacetobacter diazotrophicus*. *BMC Struct Biol* **14**, 21 (2014).
284. Bertalan, M., Albano, R., de Padua, V., Rouws, L., Rojas, C., Hemerly, A., Teixeira, K., Schwab, S., Araujo, J., Oliveira, A., Franca, L., Magalhaes, V., Alqueres, S., Cardoso, A., Almeida, W., Loureiro, M.M., Nogueira, E., Cidade, D., Oliveira, D., Simao, T., Macedo, J., Valadao, A., Dreschsel, M., Freitas, F., Vidal, M., Guedes, H., Rodrigues, E., Meneses, C., Brioso, P., Pozzer, L., Figueiredo, D., Montano, H., Junior, J., de Souza Filho, G., Martin Quintana Flores, V., Ferreira, B., Branco, A., Gonzalez, P., Guillobel, H., Lemos, M., Seibel, L., Macedo, J., Alves-Ferreira, M., Sachetto-Martins, G., Coelho, A., Santos, E., Amaral, G., Neves, A., Pacheco, A.B., Carvalho, D., Lery, L., Bisch, P., Rossle, S.C., Urmenyi, T., Rael Pereira, A., Silva, R., Rondinelli, E., von Kruger, W., Martins, O., Baldani, J.I. & Ferreira, P.C. Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. *BMC Genomics* **10**, 450 (2009).
285. Giongo, A., Tyler, H.L., Zipperer, U.N. & Triplett, E.W. Two genome sequences of the same bacterial strain, *Gluconacetobacter diazotrophicus* PAl 5, suggest a new standard in genome sequence submission. *Stand Genomic Sci.* **2**, 309-317 (2010).
286. Magnan, C., Yu, J., Chang, I., Jahn, E., Kanomata, Y., Wu, J., Zeller, M., Oakes, M., Baldi, P. & Sandmeyer, S. Sequence Assembly of *Yarrowia lipolytica* Strain W29/CLIB89 Shows Transposable Element Diversity. *PLoS One* **11**, e0162363 (2016).
287. Postma, E., Verduyn, C., Scheffers, W.A. & P., v.D.J. Enzymic analysis of the Crabtree effect in glucose-limited. *Appl. Environ. Microbiol* **55**, 468-477 (1989).
288. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D. & Higgins, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
289. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221-224 (2010).

290. Le, S.Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307-1320 (2008).

291. Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., de Montigny, J., Marck, C., Neuve´glise, C., Talla, E., Goffard, N., Frangeul, L., Aigle, M., Anthouard, V., Babour, A., Barbe, V., Barnay, S., Blanchin, S., Beckerich, J., Beyne, E., Bleykasten, C., Boisramé, A., Boyer, J., Cattolico, L., Confanioleri, F., de Daruvar, A., Despons, L., Fabre, E., Fairhead, C., Ferry-Dumazet, H., Groppi, A., Hantraye, F., Hennequin, C., Jauniaux, N., Joyet, P., Kachouri, R., Kerrest, A., Koszul, R., Lemaire, M., Lesur, I., Ma, L., Muller, H., Nicaud, J., Nikolski, M., Oztas, S., Ozier-Kalogeropoulos, O., Pellenz, S., Potier, S., Richard, G., Straub, M., Suleau, A., Swennen, D., Tekaia, F., Wésolowski-Louvel, M., Westhof, E., Wirth, B., Zeniou-Meyer, M., Zivanovic, I., Bolotin-Fukuhara, M., Thierry, A., Bouchier, C., Caudron, B., Scarpelli, C., Gaillardin, C., Weissenbach, J., Wincker, P. & Souciet, J. Genome evolution in yeasts. *Nature* **430**, 35–44 (2004).

292. Choo, J.H., Han, C., Lee, D.W., Sim, G.H., Moon, H.Y., Kim, J.Y., Song, J.Y. & Kang, H.A. Molecular and functional characterization of two pyruvate decarboxylase genes, *PDC1* and *PDC5*, in the thermotolerant yeast *Kluyveromyces marxianus*. *Appl. Microbiol. Biotechnol.* **102**, 3723-3737 (2018).

293. Lertwattanasakul, N., Kosaka, T., Hosoyama, A., Suzuki, Y., Rodrussamee, N., Matsutani, M., Murata, M., Fujimoto, N., Suprayogi, Tsuchikane, K., Limtong, S., Fujita, N. & Yamada, M. Genetic basis of the highly efficient yeast *Kluyveromyces marxianus*: complete genome sequence and transcriptome analyses. *Biotechnol. Biofuels* **8**, 47 (2015).

294. Neale, A.D., Scopes, R.K. & Kelly, J.M. Alcohol production from glucose and xylose using *Escherichia coli* containing *Zymomonas mobilis* genes. *Appl. Microbiol. Biotechnol.* **29**, 162-167 (1988).

295. Bianchi, M.M., Tizzani, L., Destruelle, M., Frontali, L. & Wésolowski-Louvel, M. The 'petite-negative' yeast *Kluyveromyces lactis* has a single gene expressing pyruvate decarboxylase activity. *Mol. Microbiol.* **19**, 27-36 (1996).

296. Uzunov, Z.G., Petrova, V.Y., Ivanov, S.L. & Kujumdzieva, A.V. *In Silico* Study of *Aro* Genes Involved in the Ehrlich Pathway: Comparison between *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. *Biotechnol. Biotechnol. Equip.* **25**, 133-137 (2014).

297. Beopoulos, A., Chardot, T. & Nicaud, J.M. *Yarrowia lipolytica:* A model and a tool to understand the mechanisms implicated in lipid accumulation. *Biochimie* **91**, 692-696 (2009).

298. Bringer-Meyer, S., Schimz, K.L. & Sahm, H. Pyruvate decarboxylase from *Zymomonas mobilis*. Isolation and partial characterization. *Arch. Microbiol.* **146**, 105-110 (1986).

299. Siegert, P., McLeish, M.J., Baumann, M., Iding, H., Kneen, M.M., Kenyon, G.L. & Pohl, M. Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. *Protein Eng. Des. Sel.* **18**, 345-357 (2005).

300. Dickinson, J.R., Harrison, S.J., Dickinson, J.A. & Hewlins, M.J. An investigation of the metabolism of isoleucine to active Amyl alcohol in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**, 10937-10942 (2000).

301. Dickinson, J.R., Harrison, S.J. & Hewlins, M.J. An investigation of the metabolism of valine to isobutyl alcohol in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 25751-25756 (1998).

302. Mojzita, D. & Hohmann, S. Pdc2 coordinates expression of the *THI* regulon in the yeast *Saccharomyces cerevisiae*. *Mol. Genet. Genom.* **276**, 147-161 (2006).

303. Nosaka, K., Onozuka, M., Konno, H. & Akaji, K. Thiamin-dependent transactivation activity of *PDC2* in *Saccharomyces cerevisiae*. *FEBS Lett.* **582**, 3991-3996 (2008).

304. Hübner, G., Weidhase, R. & Schellenberge, A. The mechanism of substrate activation of pyruvate decarboxylase: a first approach. *Eur. J. Biochem.* **92**, 175-181 (1978).

305. Milne, N., van Maris, A.J., Pronk, J.T. & Daran, J.M. Comparative assessment of native and heterologous 2-oxo acid decarboxylases for application in isobutanol production by *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **8**, 204 (2015).

306. Stribny, J., Romagnoli, G., Perez-Torrado, R., Daran, J.M. & Querol, A. Characterisation of the broad substrate specificity 2-keto acid decarboxylase Aro10p of *Saccharomyces kudriavzevii*

and its implication in aroma development. *Microb. Cell Fact.* **15**, 51 (2016).
307. Goetz, G., Iwan, P., Hauer, B., Breuer, M. & Pohl, M. Continuous production of (R)-phenylacetylcarbinol in an enzyme-membrane reactor using a potent mutant of pyruvate decarboxylase from *Zymomonas mobilis*. *Biotechnol. Bioeng.* **74**, 317–325 (2001).
308. Bruhn, H., Pohl, M., Grötzinger, J. & Kula, M. The replacement of Trp392 by alanine influences the decarboxylase/carboligase activity and stability of pyruvate decarboxylase from *Zymomonas mobilis*. *Eur. J. Biochern.* **234**, 650-655 (1995).
309. Yun, H. & Kim, B.-G. Enzymatic production of (R)-phenylacetylcarbinol by pyruvate decarboxylase from *Zymomonas mobilis*. *Biotechnol. Biotechnol. Equip.* **13**, 372-376 (2008).
310. van Hoek, P., Flikweert, M.T., Van Dijken, J.P., van der Aart, Q.J.M., Steensma, H.Y. & Pronk, J.T. Effects of Pyruvate Decarboxylase Overproduction on Flux Distribution at the Pyruvate Branch Point in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **64**, 2133–2140 (1998).
311. Mans, R., Daran, J.G. & Pronk, J.T. Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production. *Curr. Opin. Biotechnol.* **50**, 47-56 (2018).
312. Chang, A.K., Nixon, P.F. & Duggleby, R.G. Effects of Deletions at the Carboxyl Terminus of *Zymomonas mobilis* Pyruvate Decarboxylase on the Kinetic Properties and Substrate Specificity. *Biochemistry* **39**, 9430-9437 (2000).
313. Mayer, F.D., Feris, L.A., Marcilio, N.R., Staudt, P.B., Hoffmann, R. & Baldo, V. Influence of Fusel Oil Components on the Distillation of Hydrous Ethanol Fuel (Hef) in a Bench Column. *Braz. J. Chem. Eng.* **32**, 585-593 (2015).
314. Solis-Escalante, D., van den Broek, M., Kuijpers, N.G., Pronk, J.T., Boles, E., Daran, J.M. & Daran-Lapujade, P. The genome sequence of the popular hexose-transport-deficient *Saccharomyces cerevisiae* strain EBY.VW4000 reveals LoxP/Cre-induced translocations and gene loss. *FEMS Yeast Res.* **15**, fou004 (2015).
315. König, S., Spinka, M. & Kutter, S. Allosteric activation of pyruvate decarboxylases. A never-ending story? *J. Mol. Catal. B Enzym.* **61**, 100-110 (2009).
316. Gatter, M., Ottlik, S., Kovesi, Z., Bauer, B., Matthaus, F. & Barth, G. Three alcohol dehydrogenase genes and one acetyl-CoA synthetase gene are responsible for ethanol utilization in *Yarrowia lipolytica*. *Fungal Genet. Biol.* **95**, 30-38 (2016).
317. Zhang, H., Wu, C., Wu, Q., Dai, J. & Song, Y. Metabolic Flux Analysis of Lipid Biosynthesis in the *Yeast Yarrowia lipolytica* Using 13C-Labled Glucose and Gas Chromatography-Mass Spectrometry. *PLoS One* **11**, e0159187 (2016).
318. Kujau, M., Weber, H. & Barth, G. Characterization of mutants of the yeast *Yarrowia lipolytica* defective in acetyl-coenzyme A synthetase. *Yeast* **8**, 193-203 (1992).
319. Zhang, Y., Li, S., Li, J., Pan, X., Cahoon, R.E., Jaworski, J.G., Wang, X., Jez, J.M., Chen, F. & Yu, O. Using Unnatural Protein Fusions to Engineer Resveratrol Biosynthesis in Yeast and Mammalian Cells. *J. Am. Chem. Soc.* **128**, 13030-13031 (2006).
320. Thomik, T., Wittig, I., Choe, J.Y., Boles, E. & Oreb, M. An artificial transport metabolon facilitates improved substrate utilization in yeast. *Nat. Chem. Biol.* **13**, 1158-1163 (2017).
321. Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L. & Keasling, J.D. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753-759 (2009).
322. Xiu, Y., Jang, S., Jones, J.A., Zill, N.A., Linhardt, R.J., Yuan, Q., Jung, G.Y. & Koffas, M.A.G. Naringenin-responsive riboswitch-based fluorescent biosensor module for *Escherichia coli* co-cultures. *Biotechnol. Bioeng.* **114**, 2235-2244 (2017).
323. Li, S., Si, T., Wang, M. & Zhao, H. Development of a Synthetic Malonyl-CoA Sensor in *Saccharomyces cerevisiae* for Intracellular Metabolite Monitoring and Genetic Screening. *ACS Synth. Biol.* **4**, 1308-1315 (2015).
324. Wang, R., Zhao, S., Wang, Z. & Koffas, M.A. Recent advances in modular co-culture engineering for synthesis of natural products. *Curr. Opin. Biotechnol.* **62**, 65-71 (2020).
325. Xu, P., Marsafari, M., Zha, J. & Koffas, M. Microbial Coculture for Flavonoid Synthesis. *Trends Biotechnol.* **38**, 686-688 (2020).
326. Akdemir, H., Silva, A., Zha, J., Zagorevski, D.V. & Koffas, M.A.G. Production of pyranoantho-

cyanins using *Escherichia coli* co-cultures. *Metab. Eng.* **55**, 290-298 (2019).
327. Wang, X., Li, Z., Policarpio, L., Koffas, M.A.G. & Zhang, H. *De novo* biosynthesis of complex natural product sakuranetin using modular co-culture engineering. *Appl. Microbiol. Biotechnol.* **104**, 4849-4861 (2020).
328. Zhang, W., Liu, H., Li, X., Liu, D., Dong, X.T., Li, F.F., Wang, E.X., Li, B.Z. & Yuan, Y.J. Production of naringenin from *D*-xylose with co-culture of *E. coli* and *S. cerevisiae*. *Eng. Life Sci.* **17**, 1021-1029 (2017).
329. Yuan, S.F., Yi, X., Johnston, T.G. & Alper, H.S. *De novo* resveratrol production through modular engineering of an *Escherichia coli-Saccharomyces cerevisiae* co-culture. *Microb. Cell Fact.* **19**, 143 (2020).

Acknowledgments

What a ride… Unfortunately Corona dominated in the end but when looking back I realize again what a great time I have had and that was mostly due to the people at IMB and thanks to my friends and family.

First of all, I would like to thank my promotors. **Jean-Marc,** thank you for your daily supervision and guidance over the last years. Especially during the final (busy) weeks you invested a great deal of time in reading and correcting this dissertation and I really appreciate it! Also, I am really grateful for your seemingly endless supply of new ideas/angles/topics and experimental insights! Finally, congratulations on becoming a full professor! **Jack**, thank you for your never-ending enthusiasm. Your excited replies on milestones always put a huge smile on my face! Even though we would not meet that regularly, you were informed about my projects and whenever I needed to talk to you, you always managed to find some time. Afterwards, I would walk out with more confidence, energy and enthusiasm. I have such respect for how you seem to be able to teach, supervise, do amazing science, lead a department while staying approachable!

Dear **Erik,** for a self-proclaimed cynical man I think you have a heart of gold and you are one of the pillars of IMB. No matter the time or need, you are always there to help people out, from taking samples, performing emergency operations on bioreactors or having nice scientific discussions! Thank you for taking the time to teach a molecular biologist like me how to run bioreactors. Furthermore, it is so nice that you are always up for joining social activities. I lost count how many beers we shared but I hope the count will keep rising in the future!

Dear **Marijke,** you are like the mom of IMB. You are such a warm, kind and selfless person! I deeply enjoyed all our talks about characters, physiology, personal colors etc. Also, thank you so much for teaching, but more importantly, motivating me with all the enzymes measurements. Finally, thank you for all your food support over the years. Whenever I needed it, you were there with a cookie, some cheese or even a second dinner!

Dear **Robert**, before I started my Master Thesis, I thought one had to be crazy to do a PhD. Five years later, I know for sure you have to be crazy… Nevertheless, your enthusiasm and passion for science during my MEP also showed me how exciting and fun it can be. Thank you for sharing the wonders of biotechnology, it is indeed pretty cool what we all do! **Pascale**, our joint adventure, with at times a seemingly endless streak of bad luck, resulted in an amazing research line. It was very interesting to be part of such a cutting-edge project and I hope that one day in the future, an artificial cell will be possible! I wish you all the luck!

Of course, I would also like to thank our MSD ladies or our BT's Angels as I like to call them. **Janine, Astrid and Apilena**, you help all of us so much! Always thinking along, extremely helpful and with a smile on your face! You are truly indispensable for the department of Biotechnology! Astrid, good luck with your new role, I hope you will enjoy it and Apilena, thank you for all the nice food you shared with me! Also, thanks to all the other staff members who helped running the labs from IMB and provided assistance whenever necessary: **Pilar, Christiaan, Maria, Susan, Raul** and **Marcel**

My PhD was part of the European Consortium **CHASSY.** I am so fortunately to have been part of such a fun project. I got to meet many people, visit places and learned new science. **Flip,** it was fun having you aboard as Technician within CHASSY. You were always willing to help me out, of course with fermentations but also with molecular work if needed. Even more so, I really enjoyed all our legendary CHASSY drinks. Finally, it was great having an eating buddy during our trips. Someone that would join me at the buffet (three times is the charm) or order a second course (those burgers were way too small). Dear **Mario,** you were an essential part of all our CHASSY fun. Your ability (read: ordering shots) to escalate even the most innocent drinks always resulted in good stories, better pictures and fond memories. I really admire your open, enthusiastic and unbiased attitude towards life, how you stay true to yourself and your commitment to living sustainable. Furthermore, you taught me a great deal about genomic models! I wish you all the luck!

During my PhD. I have had the privilege to superviso several students. **Haje, Casper, Agnes, Vita, Nikki, Panos, Joran and Tobias**, it was great fun with all of you! Even though I was the supervisor, I learned a substantial amount from you too and I am very grateful for your contributions to this thesis. Not everything was published, but the knowledge definitely helped in realizing this thesis. I hope that whenever you smell roses, you will think back at your time at IMB with a smile!

Also, a big shout-out to all the external partners within CHASSY. I really enjoyed our regular meetings, both the scientific discussions but perhaps even more so all the activities after 5 pm! There are too many people to mention by name, but I would like to especially thank **John Morrissey, Verena Siewers, Mislav Oreb, Eckard Boles** and **Tristan Rossignol** for all your contributions and teaching me about topics outside my own field. Special thanks go to **Arun Rajkumar.** We shared information and materials over the years and it was very nice to work together with someone on the same subject to share both our failures and successes!

During my PhD I also had several collaborations outside CHASSY. Not all collaborations resulted in articles (yet) but I still enjoyed them, especially because it allowed me to learn something outside my field of expertise. So thank you, **Frank Hollman, Adrie Straathof, Michael Neasby** and **Martin Pabst.**

Next, there are some colleagues that I would like to mention. **Eline and Jordi,** as I said before our anthocyanin work was at times hectic and stressful but I am really proud of the result. The sum is definitely better than the two parts and it was really nice working together! **Wijb,** wijbie, wijbPD, I think it all started with an apple... I just could not resist pranking my very serious neighbour and that was the start of years of fun! I really enjoyed our "Good cop, bad cop" routine during PBV and thank you for putting up with Lemon Tree every morning! Also thank you for all your insights and help with my research lines, you really are a brilliant scientist! Dear **Jonna**, things might have turned out slightly different than how we expected during our MEP (how young and innocent we were), but it was really nice to (again) work together towards a deadline. It was easier to have someone else around who shared the same stress, so that we could have a coffee/vent/complain together and then get right back to it! **Nicolò** you are such a caring, helpful and funny person (at times)! We had so many great moments together and I am really glad you came to NL! Also thank you for always being willing to be my sparring partner about TCA cycles/acetyl-CoA/Pdc- strains. Finally, I would like to thank everybody else from **IMB**, past and present, for all the interest you ever showed in my work and all the nice activities we had together.

Of course, also thanks to **Anna,** my paranymph. Perhaps the best thing of having done my PhD at IMB is having met you as a friend for life. From the very first day we shared so much joy, laughter and tears. You light up the room with your energy, enthusiasm but even more so your open and genuine caring attitude towards others. Our talks, wine & dine evenings and amazing hiking trips were so much fun! I hope that when we are old and grey, we can retire and drink our home-made wine together. Dear Anna, you are a successful man, please never forget that you deserve nothing but the best.

Also thanks to many of my friends, especially **Rolien, Jorik, Martijn (Marley), Eveline, Veronica, Kim, Marlike, Iris, Floor, Jet** and **Liset** who, although they did not understand half of it, were always interested in the progress of my PhD, provided welcome distractions and were understanding whenever I was absent or did not reply because of work (sorry!). Also **Kim,** thanks a lot for all your help with the design of the cover, it turned out absolutely perfect!

Being the youngest "mini" of six children has many advantages. My siblings make sure I keep my feet firmly on the ground while simultaneously always supporting (and teasing) me. **Robert-Jan/Robbie,** after so many years I finally "caught-up" with you and we can have scientific discussions. However, no worries, I will always stay the kid sister that wants to go to Toverland together with **Laura**, **Ruben** and **Caro**! **Anne-Roos/Roosie** having seen how much you grew once you moved to Delft and hearing you sing "Op een klein stationnetje" has always been my main motivation for going to Delft, a choice I have never regretted, so thank

you! Enjoy your recently completed family with **Daan, Vesper, Axel** and of course the newest addition, **Stella**! **Carlijne/Carrie** over the past few years we grew much closer and I am very happy that we did. I really enjoy hanging out with you, **Maurits** and happy little Buddha **Flippie**! **Rianne-Fleur/Rietje** thank you for, on the one hand always confronting me with the (sometimes) ugly truth, yet on the other hand always being there for me when I needed you! You are the most considerate and reliable person I know! Also thank you and **Thommy boy** for always welcoming me in your warm house and beautiful family with **Anna Phileine** and **Sebastiaan**! Finally, **Jorick/Jojo,** my paranymph. We came a long way from "pletmachines", "throwing darts" (I still remember…) and breaking barbies. Now, you are one of my best friends and I am proud of the man you became! I am also so glad to see how happy you and **Birgit** are and I wish you all the best together with **Viggo Bear.**

Of course, six kids don't come out of thin air. Dear mom and dad, the two of you have been indispensable during my PhD. You always listened to me whenever I was struggling and supported me to keep going. Dear **Mom**, as a true lioness you have always defended your children whenever necessary. Your warm concern about my well-being and your constant reminders to be proud of myself and to pursue happiness have kept me on track all these years. I couldn't wish for a better mom or a better emotional compass! Dear **dad**, my unofficial recruiter/personal agents/biggest fan. You don't know how much I appreciate your never-ending interest in my PhD, personal life and future career. I stopped correcting you with CrispA (it should be CrispR), but I don't think many other dads are so up to date with the content of their child's PhD. Thank you for always pushing me to reach just a bit further, because of that I grew enormously as a person. I love you both and no matter how old I get; I will always enjoy coming home!

Finally, dear **Wouter,** you turned 2020 from the worst to the best year (so far) and made me realize that no matter how hard I worked for this PhD, there are things in this world that are much more important. I love you!

Curriculum Vitae

Else-Jasmijn Hassing was born on the 31th of January, 1992, in Venlo, The Netherlands. She followed her pre-university education at College Den Hulster in Venlo and graduated in 2010. Jasmijn subsequently enrolled in the Bachelor program Life, Science & Technology at Leiden University and Delft, University of Technology. Throughout her entire studies Jasmijn was active in teaching activities such as giving exam trainings in Biology and Chemistry, even having been awarded the Best Teacher Award (2x). She finalized her bachelor with an internship at the immunology group of Prof. Dr. Frits Koning at Leiden University Medical Center, studying the effect of celiac disease on the composition of cell receptors. After obtaining her bachelor degree, Jasmijn took a year off from her studies for personal development. In this year, she functioned as a quality and safety inspector during her internship at DSM in Buenos Aires, Argentina, travelled through South-America and wrote a novel called "Het Vijfde Element". In 2014, Jasmijn enrolled in the Master Program Cell Factory at Delft, University of technology. In the second year of her masters, she did her Master End Project at the Industrial Microbiology group in the Biotechnology department at Delft, University of Technology, under supervision of Dr. Robert Mans. Here she studied the transport mechanism of lactic acid in the yeast *Saccharomyces cerevisiae,* resulting in a publication. The second year was completed with an industrial internship at Heineken, in the Innovation & Research – Fermentation & Yeast department under supervision of Dr. Niels Kuijpers. Here she studied the implementation of different hybridization techniques in order to obtain new lager yeasts. In 2016, Jasmijn obtained her Master Degree with honours. After working for two months as teamleader at PostNL, she travelled to New Zealand for another two months. After returning to the Netherlands, Jasmijn started a PhD at TU Delft in the Industrial Microbiology group under supervision of Prof. Dr. Jean-Marc Daran and Prof. Dr. Jack T. Pronk. This PhD project, of which the results are described in this thesis, focuses on engineering *S. cerevisiae* for the production of aromatic compounds. Jasmijn is currently working as Industrial Strain Engineer at the Protein Brewery.

List of Publications

In this thesis:

Hassing, E., de Groot, P.A., Marquenie, V.R., Pronk, J.T., and Daran, J. Connecting central carbon and aromatic amino acid metabolisms to improve *de novo* 2-phenylethanol production in *Saccharomyces cerevisiae*. *Metab. Eng.* **56**, 165-180 (2019).

Hassing, E., Buijs, J., Blankerts, N., Luttik., M. A., de Hulster, E., Pronk, J. T., Daran, J. Elimination of aromatic fusel alcohols as by-products of *Saccharomyces cerevisiae* strains engineered for phenylpropanoids production by 2-oxo acid decarboxylase replacement. *Metab. Eng. Com.* **13,** e00183 (2021).

Postma, E. D[#]., **Hassing, E[#].,** Mangkusaputra, V., Geelhoed, J., de la Torre, P., van den Broek, M., Mooiman, C., Pabst, M., Daran, J., Daran-Lapujade, P. Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. (2021). (Submitted)

[#]These authors contributed equally.

Other publications

Mans, R., **Hassing, E.,** Wijsman, M., Giezekamp, A., Pronk, J.T., Daran, J., van Maris, A. J. A. A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in *Saccharomyces cerevisiae. FEMS Yeast Research* **17**, fox085 (2017).