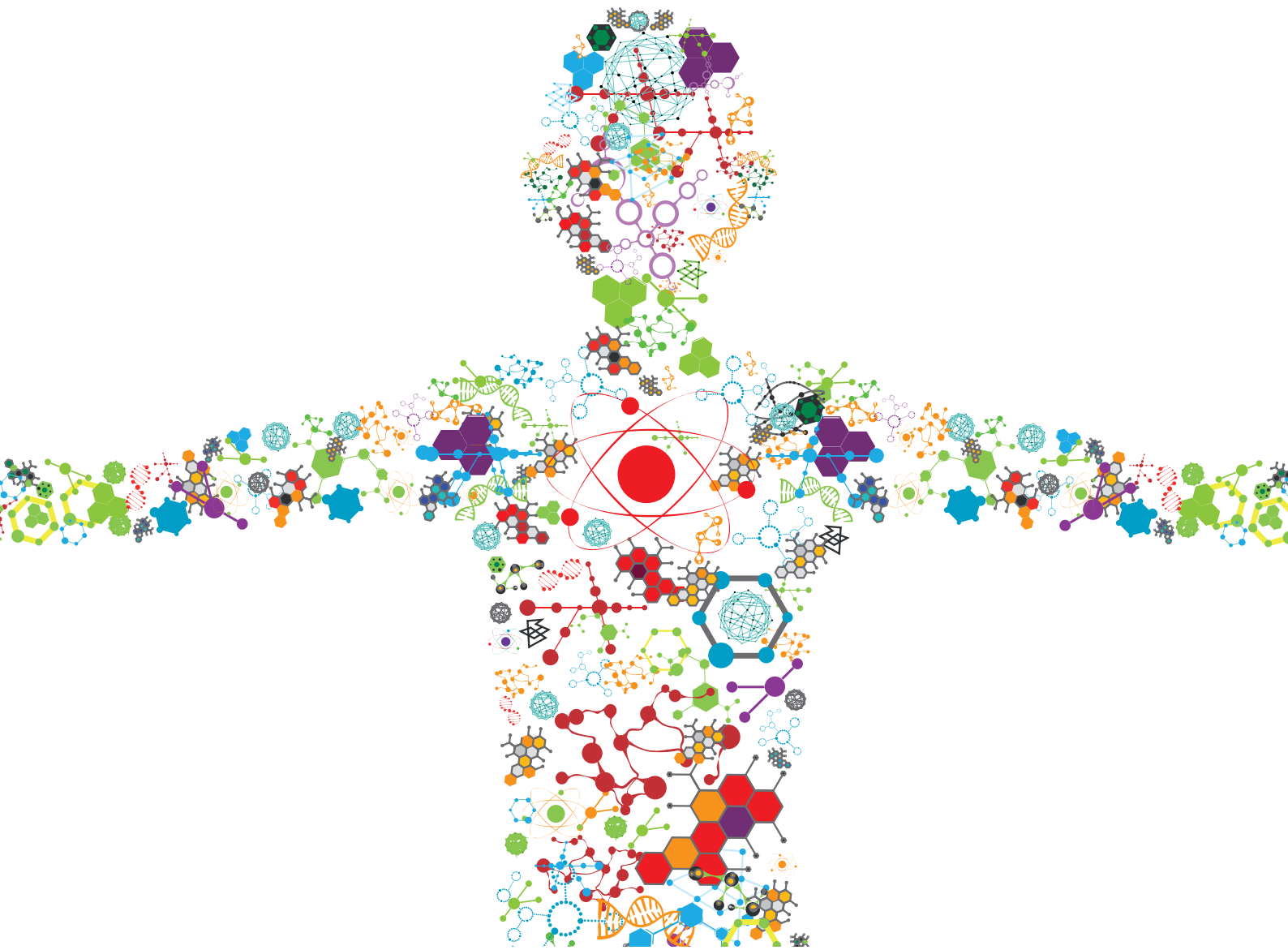
RECENT ADVANCES IN APPLICATION OF SYNTHETIC BIOLOGY FOR PRODUCTION OF BIOACTIVE COMPOUNDS

EDITED BY: Luan Luong Chu, Jae Kyung Sohng, Dipesh Dhakal and Jingwen Zhou

PUBLISHED IN: Frontiers in Bioengineering and Biotechnology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-309-4
DOI 10.3389/978-2-88974-309-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

RECENT ADVANCES IN APPLICATION OF SYNTHETIC BIOLOGY FOR PRODUCTION OF BIOACTIVE COMPOUNDS

Topic Editors:
Luan Luong Chu, Phenikaa University, Vietnam
Jae Kyung Sohng, Sun Moon University, South Korea
Dipesh Dhakal, University of Florida, United States
Jingwen Zhou, Jiangnan University, China

Citation: Chu, L. L., Sohng, J. K., Dhakal, D., Zhou, J., eds. (2022). Recent Advances in Application of Synthetic Biology for Production of Bioactive Compounds. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-309-4

Table of Contents

frontiers in Bioengineering and Biotechnology

EDITORIAL
published: 22 December 2021
doi: 10.3389/fbioe.2021.819475

Editorial: Recent Advances in Application of Synthetic Biology for Production of Bioactive Compounds

Luan Luong Chu[1,2]*, *Jingwen Zhou*[3], *Dipesh Dhakal*[4] and *Jae Kyung Sohng*[5,6]

[1]Faculty of Biotechnology, Chemistry and Environmental Engineering, Phenikaa University, Hanoi, Vietnam, [2]Bioresource Research Center, Phenikaa University, Hanoi, Vietnam, [3]Science Center for Future Foods, Jiangnan University, Wuxi, China, [4]Department of Medicinal Chemistry, University of Florida, Gainesville, FL, United States, [5]Department of Life Science and Biochemical Engineering, SunMoon University, Chungnam, South Korea, [6]Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, Chungnam, South Korea

Keywords: synthetic biology, metabolic engineering, bioactive compounds, *E. coli*, *S. cerevisiae*

Editorial on the Research Topic

Recent Advances in Application of Synthetic Biology for Production of Bioactive Compounds

Bioactive natural compounds broadly exhibit application in various fields, as chemical and food additives, agrochemical products, cosmetics, pharmaceuticals, and biofuels. Furthermore, unnatural bioactive compounds have been synthesized and shown a great effect on promoting better health. Both natural and unnatural bioactive compounds include polyphenols, flavonoids, terpenoids, alkaloids, carotenoids, stilbenes, and anthocyanins. Noticeably, bioactive compounds also contain glycosides, steroids, plant sterols, and peptides. The exploration, biosynthesis, and amplification of biological activities of natural and non-natural molecules have remained unchanged as one of the most exciting trends in biotechnological products.

Most of the bioactive compounds have been isolated from plants, microorganisms, marine organisms, and fungi. However, the yields of natural products are usually low and require time-consuming procedures for industrial production. Furthermore, chemical synthesis is a costly, harmful alternative and requires multi-step isolation and purification processes. Although plant molecular engineering has been significantly developed, using industrially preferred microorganisms is a promising approach for the biosynthesis of industrial products. In recent years, a wide variety of novel technologies for engineering plants and microbes have been developed to produce natural and non-natural compounds from renewable biomasses. Along with evolutionary engineering, metabolic engineering, and systems biology, synthetic biology is expected to further improve the productivity of the compounds. This topic focuses on providing an overview of the recent advances, emerging challenges, and future prospects of synthetic biology and metabolic engineering for the biomanufacturing of bioactive compounds.

High-value compounds have been synthesized and produced using various engineered platforms. For example, Liu et al. provided an overview of the microbial chassis for the production of fatty acid-derived chemicals (FACs). Interestingly, autotrophs (such as *Cupriavidus necator*, *Rhodococcus opacus*, *Synechococcus* sp. PCC 7002, and *Nostoc punctiforme*) can synthesize FACs from CO_2 using chemical sources, light, and electric energy; while heterotrophs (as *Escherichia coli*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, and *Aureobasidium*) were designed and engineered to produce FACs from organic carbon sources. Sajied et al. summarized the current *ex-planta* production of isoflavonoid *via* an artificial isoflavonoid biosynthesis pathway in engineered *E. coli* and *S. cerevisiae*. Furthermore, Wang et al. reviewed the fermentation processes in the production of water-soluble vitamins (vitamin B1, B2, B3, B5, B6, B7, B9, B12, and vitamin C) and fat-soluble

OPEN ACCESS

Edited and reviewed by:
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

***Correspondence:**
Luan Luong Chu
luan.chuluong@phenikaa-uni.edu.vn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *21 November 2021*
Accepted: *29 November 2021*
Published: *22 December 2021*

Citation:
Chu LL, Zhou J, Dhakal D and Sohng JK (2021) Editorial: Recent Advances in Application of Synthetic Biology for Production of Bioactive Compounds. Front. Bioeng. Biotechnol. 9:819475. doi: 10.3389/fbioe.2021.819475

vitamins (vitamin A/D/E and vitamin K) in microbial cell factories. Noticeably, Mutanda et al. summarized comprehensive information regarding the current biotechnological production of taxol in engineered plants (*A. thaliana*, *Nicotiana benthamiana*, *Nicotiana Sylvestris*), endophytic fungus (*Alternaria alternate*) as well as engineered microbes (*E. coli*, *S. cerevisiae*). They have provided a novel insight on the regulatory mechanism governing taxol biosynthesis by understanding the challenges with metabolic engineering and mining of transcriptomic data sets from *Taxus* species. The review demonstrated that plant genetic engineering has been significantly developed to improve the catalytic process and product specificity of the important compounds. Fu et al. demonstrated *Artemisia annua* L. with glandular trichomes can be used as a transgenic platform for high-level patchoulol production. Overexpressing farnesyl diphosphate synthase and patchoulol synthase genes along with alteration of the subcellular location resulted in increasing to 273 $\mu g\, g^{-1}$ DW of patchoulol.

Overexpression of genes in biosynthetic pathway and inhibition of genes in competitive pathway plays a key role in engineered microorganisms for bioactive compound biosynthesis. Wang et al. engineered *Pseudomonas Chlororaphis* P3 to produce gentisate from 3-hydroxybenzoate (3-HBA) and 4-hydroxybenzoate (4-HBA). The chromosome-integrated synthetic pathway and blocking the key conversion steps resulted in the production of 365 $mg\, L^{-1}$ of gentisate from 3-HBA. Similarly, Ma et al. reported the first report of the production of (+)-borneol using engineered *S. cerevisiae*. (+)-bornyl diphosphate synthase from *Cinnamomum burmanni* was modified by tailored truncation and adding Kozak sequences, then overexpressed in *S. cerevisiae* harboring reconstituted (+)-bornyl biosynthetic pathway. Moreover, Chen et al. reported the production of 2-keto-L-gulonic acid in an acetic acid bacterium *Gluconobacter oxydans*. They obtained 97 promoters and identified the activity of the strongest promoter (P_{2703}) from genome of *G. oxydan*. After that, gene succinate dehydrogenase (*SDH*) was overexpressed under the control of P_{2730} in *G. oxydan* WSH-003. An engineered strain reached 3.7 $g\, L^{-1}$ of 2-keto-L-gulonic acid. On another hand, activation of the silent biosynthetic gene clusters (BGCs) in the native host is an essential strategy to discover novel natural products. For example, Liu et al. reviewed the major approaches for activation of natural product BGCs in *Streptomyces*. They described the strategies to metabolic regulatory network for *in situ* activation of target BGCs, including promoter engineering, transcriptional regulation engineering, ribosome, and RNA polymerase engineering.

In general, the research topic collected excellent examples of advanced metabolic engineering assisted by synthetic biology. The engineered host in this research topic showed the ability to produce bioactive compounds on industrial fermentation. However, there is still need to develop advanced synthetic biology tools in both plant and microbial platforms. Despite information in databases, such as KEGG, BioCyc, MetaCyc, or BRENDA, and utilization of genomic editing tools as ZFNs, TALENs, or CRISPR/Cas system, having been rapidly increased in recent years, the efficient synthetic biology tools for non-conventional microbial platforms are still limited. These barriers have resulted in limitations of the titer, rate, and yield of bioactive compound products. In order to overcome the bottlenecks and further advances in engineered strains, Ramzi et al. provided an omics technology and machine learning (ML) platform as an efficient tool for optimizing biosynthetic pathways and enhancing the microbial production capacity. Moreover, ML-based synthetic biology-combined artificial intelligence (AI) has a promising approach to generate a super host with an enhanced metabolic pathway centered for industrial bioactive compound products. The intelligence-produced hosts are expected to not only generate novel biomolecules but also combine carbon-fixing autotrophs and heterotrophs with net-zero greenhouse gas emissions.

In summary, the current research topic summaries a valuable collection of articles focusing on recent development and application of synthetic biology and metabolic engineering for secondary metabolite production in plant and microbial platforms. Omics and ML-assisted synthetic biology and metabolic engineering are expected to play significant roles to overcome the emerging challenges faced in agricultural, medical, and environmental biotechnology.

AUTHOR CONTRIBUTIONS

LC wrote the manuscript. JZ, DD, and JS revised the manuscript. All authors approved the manuscript.

ACKNOWLEDGMENTS

The editors appreciate the contribution of all authors to this research topic, the constructive comment of all the reviewers, and the editorial support from Frontiers throughout the publication process.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 21 December 2020
doi: 10.3389/fbioe.2020.608918

Streamlining Natural Products Biomanufacturing With Omics and Machine Learning Driven Microbial Engineering

Ahmad Bazli Ramzi[1], Syarul Nataqain Baharum[1], Hamidun Bunawan[1] and Nigel S. Scrutton[2]*

[1] Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, Bangi, Malaysia, [2] EPSRC/BBSRC Future Biomanufacturing Research Hub, BBSRC/EPSRC Synthetic Biology Research Centre, Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, United Kingdom

OPEN ACCESS

Edited by:
Luan Luong Chu, Yeungnam University, South Korea

Reviewed by:
Dae-Hee Lee, Korea Research Institute of Bioscience and Biotechnology (KRIBB), South Korea
Hiromichi Minami, Ishikawa Prefectural University, Japan
Jian Zha, Shaanxi University of Science and Technology, China

***Correspondence:**
Ahmad Bazli Ramzi
bazliramzi@ukm.edu.my

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *22 September 2020*
Accepted: *18 November 2020*
Published: *21 December 2020*

Citation:
Ramzi AB, Baharum SN, Bunawan H and Scrutton NS (2020) Streamlining Natural Products Biomanufacturing With Omics and Machine Learning Driven Microbial Engineering. Front. Bioeng. Biotechnol. 8:608918. doi: 10.3389/fbioe.2020.608918

Increasing demands for the supply of biopharmaceuticals have propelled the advancement of metabolic engineering and synthetic biology strategies for biomanufacturing of bioactive natural products. Using metabolically engineered microbes as the bioproduction hosts, a variety of natural products including terpenes, flavonoids, alkaloids, and cannabinoids have been synthesized through the construction and expression of known and newly found biosynthetic genes primarily from model and non-model plants. The employment of omics technology and machine learning (ML) platforms as high throughput analytical tools has been increasingly leveraged in promoting data-guided optimization of targeted biosynthetic pathways and enhancement of the microbial production capacity, thereby representing a critical debottlenecking approach in improving and streamlining natural products biomanufacturing. To this end, this mini review summarizes recent efforts that utilize omics platforms and ML tools in strain optimization and prototyping and discusses the beneficial uses of omics-enabled discovery of plant biosynthetic genes in the production of complex plant-based natural products by bioengineered microbes.

Keywords: microbial engineering, synthetic biology, omics technology, machine learning, biomanufacturing, systems biology

INTRODUCTION

Omics-Enabled Discovery of Plant Biosynthetic Genes

Plant natural products represent an enormous resource for chemical and biotechnological production of biopharmaceuticals and natural products-based drugs where about 50–70% of all anti-infective agents in clinical use are being provided and inspired by natural products (Newman and Cragg, 2016). As of 2019, up to 41.3% of anti-infective agents including antiviral and anti-malarial drugs were derived from natural products, which underlines the importance of these compounds as therapeutic agents (Newman and Cragg, 2020). With the advent of systems biology and omics research that have focused on investigating biological mechanisms at systems levels (Kitano, 2002; Lister et al., 2009), a plethora of bioactive compounds and relevant biosynthetic

pathways has since been profiled and identified. This has culminated in the steady expansion of plant-based natural products datasets (Rai et al., 2017).

Driven by the increased availability of bioinformatics tools and high throughput instruments, including next generation sequencing (NGS) and mass spectrometry (MS), omics technologies have been prominently used as principal tools in systems biology research aimed at elucidating the underlying molecular mechanisms behind cellular functions and interplays among biomolecules in biological systems (Fridman and Pichersky, 2005; Sheth and Thaker, 2014; O'Brien et al., 2015). Omics technologies including DNA sequencing (genomics), RNA sequencing (RNA-seq; transcriptomics), and MS-based protein (proteomics) and metabolite (metabolomics) analyses have empowered the reconstruction of metabolic networks based on genome annotation and functional characterization of targeted biochemical reactions in a particular organism or system. The use of systems biology approaches in combination with computational methods has contributed to the generation of genome-scale metabolic models (GEMs) that are important in identifying all metabolic reactions and corresponding biosynthetic genes in various microbes and plants (Seaver et al., 2012; O'Brien et al., 2015).

Importantly, the adoption of single or multi-omics in natural products studies has seen the increment of omics-guided discovery of known and novel metabolites, biosynthetic genes, and regulatory elements from model and non-model plants. By employing transcriptome-guided gene mining and microbial engineering strategies, a number of natural products from previously incomplete and gapped pathways, such as opiate alkaloid noscapine and cannabinoids, have since been produced in microbial hosts, thereby opening up new and exciting opportunities in natural products biomanufacturing using bioengineered microbes as the preferred bioproduction platform (Li and Smolke, 2016; Luo et al., 2019; Courdavault et al., 2020). Biomanufacturing and commercialization of fermentation-based bioproducts, such as artemisinin, nootkatone, and β-farnesene, serve to demonstrate the feasibility and the bioeconomy potential of microbial engineering platforms in the production of fine chemicals and biopharmaceuticals (Benjamin et al., 2016; Ekas et al., 2019). In this mini review, recent applications of systems and synthetic biology approaches in the bioproduction of natural products are discussed where the advancement of natural products biomanufacturing using omics-driven microbial engineering and machine learning (ML)-assisted strain optimization strategies was further highlighted.

INTEGRATION OF SYSTEMS AND SYNTHETIC BIOLOGY FOR MICROBIAL PRODUCTION OF NATURAL PRODUCTS

Metabolic engineering and synthetic biology represent advanced bioproduction strategies that have allowed researchers to reprogram and modulate microbial metabolism using genetic and computational tools (Ramzi, 2018; Choi et al., 2019). Multi-omics approaches have been initially established for microbial systems leading to a growing number of reconstructed GEMs, especially in the universal chassis *Escherichia coli* and *Saccharomyces cerevisiae* where the computational sets of stoichiometric and mass-balanced metabolic reactions in the microbes were derived from genomics-guided experimental analysis including flux balance analysis (FBA) and elementary node analysis (Gu et al., 2019; Dahal et al., 2020). A host of systems biology, bioinformatics, and computer-aided design (CAD) tools has been developed and utilized to identify cellular metabolic bottleneck, pathway prediction, and gene design with the ultimate aim of enhancing bioproduction titers, rates, and yields (TRYs) by metabolically engineered microbes (Chae et al., 2017; Choi et al., 2019).

The advent of data-driven systems and synthetic biology has brought a renewed and ever-increasing interest in translating laboratory strains into commercial-level microbial prototypes using omics- and *in silico*-guided biomanufacturing platforms that are expected to accelerate the scale-up process and speed up industrial scale production of desired products (Lee and Kim, 2015; Carbonell et al., 2018; Dunstan et al., 2020). The incorporation of the iterative Design-Build-Test-Learn (DBTL) cycle in microbial engineering approaches has provided a biological engineering and *in silico*-assisted framework for strain design and prototyping invaluable for industrial biotechnology applications. As part of the efforts in converging predictive analytics in improving bioproduction capabilities, the employment of metabolome, proteome, transcriptome, and bioinformatics analyses of the plant resources and microbial chassis has provided a comprehensive data-driven means for modulating and streamlining the biomanufacturing process of high-value natural products guided by the DBTL bioengineering framework (Casini et al., 2018; Carqueijeiro et al., 2020). An overview of data-guided bioproduction of natural products using systems and synthetic biology approaches is illustrated in **Figure 1** where the implementation of omics technology and ML tools in improving top-down and bottom-up biomanufacturing strategies is further discussed in the following sections.

TOP-DOWN APPROACH: OMICS-GUIDED STRAIN DESIGN AND PATHWAY OPTIMIZATION

One of the key aspects of strain development using metabolic engineering and synthetic biology tools is the generation and characterization of biosynthetic genes as genetic parts in the pathway design of which the standardization in parts and plasmid assembly allows rapid strain prototyping *via* the DBTL iteration (Nielsen and Keasling, 2016; Robinson et al., 2020). In efforts to maximize TRYs of the natural products and precursor biosynthesis, omics-guided pathway analysis has been applied for a top-down microbial engineering approach by elucidating and identifying affected genes and proteins especially rate-limiting enzymes in engineered metabolic pathways (**Table 1**). In this top-down strain optimization approach, several omics platforms were employed in pathway debottlenecking and optimization in bioengineered microbial

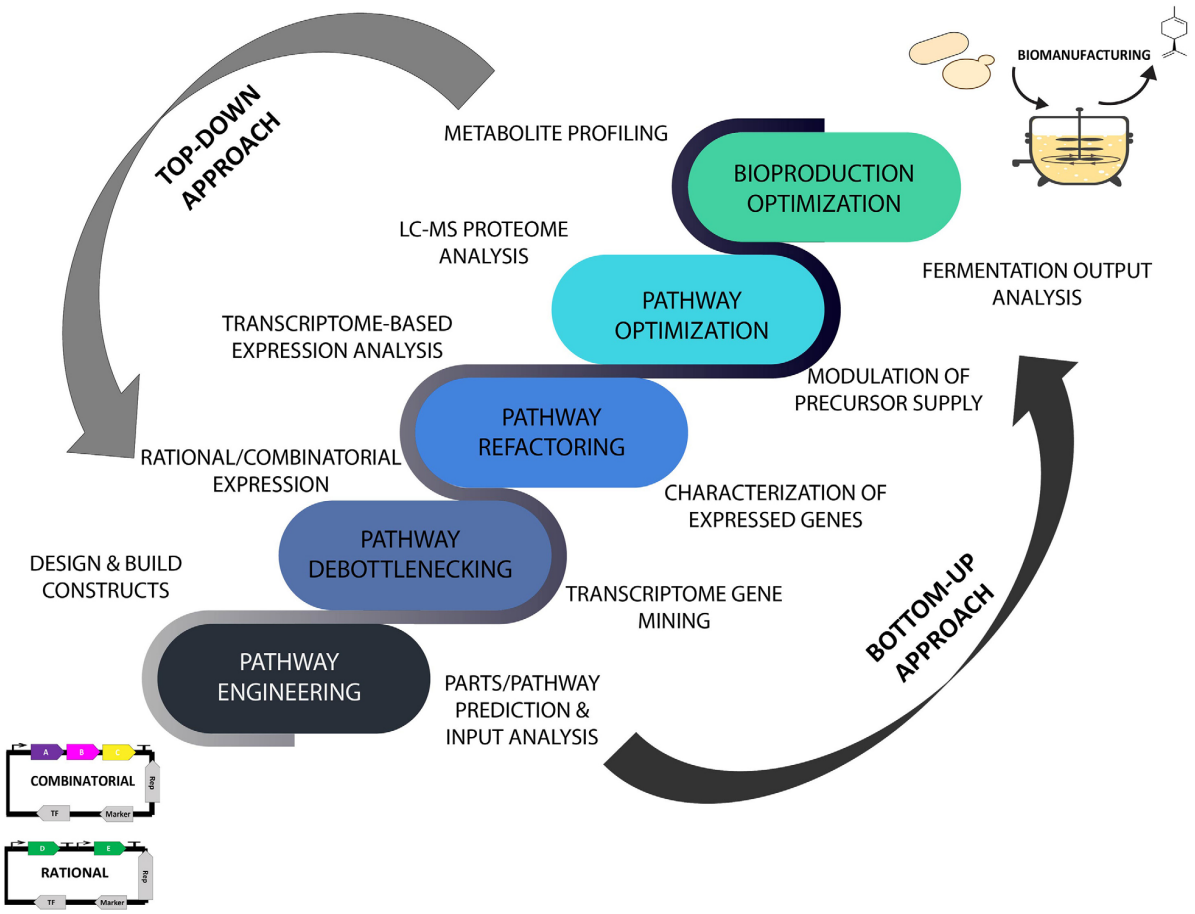


FIGURE 1 | Top-down and bottom-up biomanufacturing strategies driven by omics and machine learning tools using high throughput analytical and predictive engineering technologies. Top-down approach focused on strain optimization and bioproduction improvement. Bottom-up approach aimed at reconstruction and refactoring of biosynthetic pathways for synthesis of existing and new-to-nature bioproducts. LC-MS, liquid chromatography–mass spectrometry.

chassis that aimed at improving precursor supply and enhancing targeted natural product biosynthesis in a reverse engineering manner. With the focus on Test and Learn steps, proteome, metabolome, and bioinformatics analyses were conducted for the modulation of endogenous pathway intermediates, such as amino acids and isopentenyl pyrophosphate (IPP)-derived precursors, in bioengineered microbes. In particular, fine-tuning of IPP-related biosynthetic genes was found to be critical in optimizing terpenes bioproduction in engineered *E. coli* and *S. cerevisiae* owing to poor recombinant protein translation and precursor toxicity. Through proteome and transcriptome analyses of terpene-producing strains of *E. coli*, these pathway bottlenecks were debugged through codon optimization of the rate-limiting enzymes and the use of strong and regulated promoters, such as pTrc and pGadE (Redding-Johanson et al., 2011; Dahl et al., 2013). The application of principal component analysis of proteomics (PCAP) and multi-omics approaches in terpene-producing *E. coli* further demonstrated the importance of balanced and optimal protein expression, especially for HMG-CoA reductase, the key enzyme in the IPP-supplying mevalonate (MVA) pathway (Alonso-Gutierrez et al., 2015). In a seminal report by Brunk et al. (2016) on omics-guided microbial engineering, the combination of GEM, metabolomic, and proteomic analyses has allowed comprehensive pathway mapping and debottlenecking in MVA-derived terpene-overproducing *E. coli* by which several genes in the pentose phosphate pathway, tricarboxylic acid (TCA) cycle, and acetyl-CoA biosynthesis were found to be important in particular by downregulating pyruvate synthase (YDBK) gene that culminated in higher specific production of limonene. In genome engineered *S. cerevisiae*, the use of flux and metabolomic analysis has aided the functional expression of a heterologous 1-deoxy-D-xylulose 5-phosphate (DXP) pathway, the alternative IPP-producing pathway by which combinatorial expression of IspG (2-C-methyl-D-erythritol-2,4-cyclodiphosphate reductase) and IspH (4-hydroxyl-3-methylbut-2-enyl diphosphate reductase) enzymes was tested to overcome the poor conversion of 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (MEcPP) and the limited NADPH coenzyme availability (Kirby et al., 2016).

Omics tools have also been utilized in elucidating cellular changes in yeast chassis engineered to produce aromatic phenylpropanoids *via* the shikimate pathway using aromatic

amino acid L-phenylalanine or L-tyrosine as the main entry routes for phenylpropanoid biosynthesis. Metabolomic and transcriptomic analyses of p-coumaric acid (p-CA)-overproducing *S. cerevisiae* revealed distinct transcriptional changes of genes related to sugars and amino acids transport in S288c and CEN.PK background strains that aided in the efforts of systematically modulating the final production of p-CA with up to 20–50% improvement (Rodriguez et al., 2017). Further transcriptome-guided pathway optimization enabled enhanced p-CA bioproduction from xylose in which deletion of the tyrosine and tryptophan amino acid transporter TAT1 resulted in 50% increased of the p-CA titer (Borja et al., 2019). Similar transcriptome-assisted bioengineering strategies were employed to build and test multiple sets of yeast promoters including pINO1, pSED1, and pCCW12 that conferred increased naringenin production from p-CA in engineered *S. cerevisiae* (Gao et al., 2020). Evidently, the application of omic technologies in chassis optimization, especially in the Test and Learn synthetic biology cycle, is inordinately advantageous in pathway debottlenecking and increasing the TRYs of the desired natural products.

BOTTOM-UP APPROACH: OMICS-ENABLED PATHWAY ENGINEERING AND REFACTORING FOR NATURAL PRODUCTS BIOMANUFACTURING

The employment of single- and multi-omic tools has brought about a systematic biology-informed pipeline for discovering and biomanufacturing of new-to-nature plant-derived compounds using systems and synthetic biology platforms (Goh, 2018; Chen et al., 2020; Jamil et al., 2020). In the Design and Build steps, genes involved in plant and microbial natural products pathways are considered as important genetic parts by which reconstruction and combinatorial expression of the corresponding biosynthetic pathways have yielded a plethora of industrially important natural products and biochemicals in bioengineered microbes. Discovery of key and missing enzymes in plant biosynthetic pathways has been greatly expedited with transcriptome gene mining of non-model plants and expression of the candidate genes in microbial systems (Goh et al., 2018; Ku Bahaudin et al., 2018; Pyne et al., 2019). Two alkaloid-enriched plants specifically *Papaver somniferum* (opium poppy) and *Catharanthus roseus* (Madagascar periwinkle) have emerged as the model medicinal plants with regard to the employment of multi-omics approaches in the comprehensive analysis of the benzylisoquinoline alkaloids (BIAs) and monoterpenoid indole alkaloids (MIAs) biosynthetic pathways, respectively (Facchini and De Luca, 2008; Scossa et al., 2018). Using multi-omics strategies, the complete biosynthetic pathway of the anticancer drug vinblastine in *C. roseus* has been finally elucidated where a total of 31 steps are required for MIA compound synthesis from geranyl pyrophosphate (GPP) where the key redox and hydrolase enzymes for the conversion of stemmadenine to tabersonine or catharanthine were successfully identified *via* proteome analysis and transcriptome gene mining (Caputi et al., 2018). These omics-driven strategies were similarly employed for the identification and expression of terpene and phenylpropanoid biosynthetic genes from the aromatic plant *Polygonum minus* (*Persicaria minor*) essential for pathway reconstruction and natural product biosynthesis in engineered microbes (Ramzi et al., 2018; Rusdi et al., 2018; Tan et al., 2018).

Transcriptomic-Driven Design and Build of High-Value Natural Products in Microbial Chassis

One of the prominent examples of omics-enabled discovery and production of high-value natural products is the bioproduction of BIAs where candidate genes were obtained from the transcriptome datasets of BIA-accumulating plants, thereby representing a bottom-up approach in natural products biomanufacturing. The production of (*S*)- and (*R*)-reticuline was first demonstrated in engineered *S. cerevisiae* through BIA pathway reconstitution that includes the expression of the enzymes norcoclaurine synthase (NCS) and reticuline epimerase (CYP82Y2) from opium poppy *P. somniferum* (DeLoache et al., 2015; Farrow et al., 2015; **Table 1**). Through gene mining of *P. somniferum* transcriptome datasets, microbial expression of long and complex pathway of BIAs allowed the bioproduction of bioactive dihydrosanguinarine, thebaine, and hydrocodone compounds in engineered *S. cerevisiae* (Fossati et al., 2014; Galanie et al., 2015). Reconstruction and implantation of plant biosynthetic pathways can be modulated and programmed to exploit intrinsic amino acid pathways, such as L-phenylalanine, L-tryptophan, and L-ornithine, thereby removing the metabolic barriers for precursor and energy supply. Combinatorial and rational design strategies have enabled the biosynthesis of tropane alkaloids where *de novo* production of *N*-methylpyrrolinium, tropine, and cinnamoyl tropine has been attained through the incorporation and conversion of L-ornithine- and L-phenylalanine-derived intermediates, respectively, through the expression of corresponding *N*-methyl putrescine oxidase (MPO) from *Nicotiana tabacum* and tropane alkaloid biosynthesis genes from *Anisodus acutangulus* (Ping et al., 2019a,b; Srinivasan and Smolke, 2019).

Transcriptome analysis of antioxidant-rich medicinal plants, including *P. minus* and *Erigeron breviscapus*, revealed the candidate biosynthetic genes for phenylpropanoid-derived flavonoids and breviscapine that shared L-phenylalanine as the main intermediate compound in the plant biosynthetic pathway (Loke et al., 2017; Liu et al., 2018). The introduction of key biosynthetic genes, such as cinnamate-4-hydroxylase (C4H), flavone-6-hydroxylase (F6H), and flavonoid-7-*O*-glucuronosyltransferase (F7GAT), enabled pathway reconstruction and directed biosynthesis of the desired phenylpropanoid compounds in engineered *S. cerevisiae* using glucose as carbon source (Liu et al., 2018; Ramzi et al., 2018). Interestingly, the presence of endogenous MVA and squalene biosynthetic pathways in *S. cerevisiae* serves as a starting platform for transcriptome-enabled biosynthesis of

TABLE 1 | Omics-guided microbial engineering approaches for natural product and precursor biomanufacturing. Top-down approach mainly represented pathway debottlenecking and strain optimization for increasing bioproduction capacity. Bottom-up approach utilized transcriptome-enabled gene discovery for pathway engineering, refactoring, and bioproduction of industrially important natural products. MVA, mevalonate; PP, pentose-phosphate; TCA, tricarboxylic acid; DXP, 1-deoxy-D-xylulose 5-phosphate; BIA, benzylisoquinoline alkaloid.

Approach	Target metabolite (Chassis)	Key biosynthetic genes and parts	Omic-guided strategy	References
Top-down	**Terpene (*E. coli*)**	**MVA pathway** Mevalonate kinase (MK) and phosphomevalonate kinase (PMK) from *S. cerevisiae* under the control of *trc* promoter (*E. coli*)	Proteome-guided promoter characterization and pathway bottleneck debugging via codon optimization	Redding-Johanson et al., 2011
		MVA pathway Farnesyl pyrophosphate (FPP)-responsive promoters PybrL, PgadE, and PrstA controlling FPP biosynthetic genes	Promoter characterization and pathway intermediate toxicity measurement based on proteome and transcriptome dataset analysis of engineered *E. coli*	Dahl et al., 2013
		MVA pathway HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR) from *Staphylococcus aureus*; Terpene synthase from *Mentha spicata* and *Abies grandis*	Application of principal component analysis for enzyme characterization and improvement based on proteome dataset of engineered *E. coli*	Alonso-Gutierrez et al., 2015
		PP pathway Phosphogluconate dehydrogenase (GND), glucose-6-phosphate dehydrogenase (G6PDH2r) in *S. cerevisiae* **TCA cycle** Isocitrate dehydrogenase (ICDHyr), alpha-ketoglutarate dehydrogenase (AKGDH) in *S. cerevisiae* **Acetyl-CoA biosynthesis** Pyruvate synthase (YDBK) in *S. cerevisiae*	Pathway precursor supply mapping using multi-omics (Metabolomics and proteomics) and GEM analysis	Brunk et al., 2016
	Terpene precursor (*S. cerevisiae*)	**MVA-associated pathway** HMG-CoA synthase (ERG13) and membrane protein (PRM10) in *S. cerevisiae* **DXP pathway** DXP biosynthetic genes (*Dxs, IspC, IspD, IspE, IspF, IspH*) from *E. coli*; 2-methyl-butenyl-4-diphosphate (HMBPP) synthase (IspG) from *E. coli, Bacillus subtilis, B. thuringiensis* and *Thermus thermophilus* **DXP-related redox system** Flavodoxin/ferredoxin $NADP^+$-reductase (AtrFNR) from *Arabidopsis thaliana* and flavodoxin (Fld) from *E. coli* and *B. subtilis* **Iron-sulfur cluster (ISC) machinery** ISC operon (HscA, iscA, cyaY, iscS, iscU, hscB, fdx) and respiratory protein A (erpA) from *E. coli*	Genomics and metabolomics-assisted DXP pathway optimization	Kirby et al., 2016
	Phenylpropanoid precursor (*S. cerevisiae*)	**Coumaric acid biosynthesis** Tyrosine ammonia lyase (TAL) from *Flavobacterium johnsoniae*, shikimate kinase (aroL) from *E. coli*, tyrosine biosynthetic genes (m*ARO7*, m*ARO4, ARO10*) and pyruvate decarboxylase (PDC5) in *S. cerevisiae* **Amino acid and sugar transport** Tyrosine and tryptophan amino acid transporter (TAT1), polyamine transporter (TPO1), arginine transporter (ALP1), amino acids transporters (BAP2, AGP3), acetate transporter (ADY2) and galactose transporter (GAL2) in *S. cerevisiae*	Metabolic pathway characterization and optimization based on metabolomic and transcriptomic analysis	Rodriguez et al., 2017
		Coumaric acid biosynthesis Shikimate kinase (aroL) from *E. coli*, tyrosine ammonia-lyase (TAL) from *F. johnsoniae*, tyrosine biosynthetic genes (m*ARO7*, m*ARO4*, *ARO10*) and pyruvate decarboxylase (PDC5) from *S. cerevisiae* **PP pathway** Glucose-6-phosphate dehydrogenase (ZWF1), 6-phosphogluconolactonase (SOL3) and 6-phosphogluconate dehydrogenase (GND1) in *S. cerevisiae*	Transcriptome-guided metabolic pathway characterization and optimization	Borja et al., 2019

(Continued)

TABLE 1 | Continued

Approach	Target metabolite (Chassis)	Key biosynthetic genes and parts	Omic-guided strategy	References
	Flavonoid (*S. cerevisiae*)	**Characterized promoters** pTDH1, pPGK1, pINO1, pSED1 and pCCW12 in *S. cerevisiae* **Naringenin biosynthesis** 4-coumarate:CoA ligase (Ps4CL) from *Petroselinum crispum*, CHS chalcone synthase from *Petunia x hybrida* (PhCHS) and CHI chalcone isomerase from *Medicago sativa* (MsCHI)	Promoter characterization and yield improvement via transcriptomic analysis	Gao et al., 2020
Bottom-up	**BIA (*S. cerevisiae*)**	**(*S*)-reticuline biosynthesis** Norcoclaurine synthase (PsNCS) from *Papaver somniferum*	Identification and functional expression of norcoclaurine synthase in *S. cerevisiae* for (*S*)-reticuline production from L-tyrosine	Xiao et al., 2013; DeLoache et al., 2015
		(*R*)-reticuline biosynthesis Reticuline epimerase (PsCYP82Y2) from *P. somniferum*	Identification and functional expression of reticuline epimerase in *S. cerevisiae* for conversion of (*S*) to (*R*)-reticuline	Desgagné-Penix et al., 2012; Farrow et al., 2015
		Dihydrosanguinarine biosynthesis 6-*O*-methyltransferase (6*O*MT), coclaurine *N*-methyltransferase (CNMT), 4′-*O*-methyltransferase 2 (4′*O*MT2), truncated berberine bridge enzyme (BBEΔN), cheilanthifoline synthase (PsCFS) and stylopine synthase (PsSPS), cytochrome P450 reductase (PsCPR), tetrahydroprotoberberine cis-*N*-methyltransferase (TNMT), (S)-cis-*N*-methylstylopine 14-hydroxylase (MSH) from *P. somniferum*	Transcriptome gene mining and expression of 10-gene pathway from *P. somniferum* for biosynthesis of dihydrosanguinarine in *S. cerevisiae*	Xiao et al., 2013; Fossati et al., 2014
		Opioids biosynthesis 1,2-dehydroreticuline synthase-1,2-dehydroreticuline reductase (DRS-DRR), salutaridine synthase (SalSyn), salutaridine reductase (SalR), salutaridinol 7-O-acetyltransferase (SalAT), thebaine 6-*O*-demethylase (T6*O*DM) from *P. somniferum*	Transcriptome gene mining, characterization and complete biosynthesis of opioids thebaine and hydrocodone in bioengineered *S. cerevisiae*	Xiao et al., 2013; Matasci et al., 2014; Galanie et al., 2015
	Tropane alkaloids (*S. cerevisiae*)	**Tropane alkaloid biosynthesis** *N*-methyl putrescine oxidase (MPO) from *Nicotiana tabacum*	Transcriptome gene mining, characterization and functional expression of putrescine oxidase in *S. cerevisiae*	Matasci et al., 2014; Srinivasan and Smolke, 2019
		Tropane alkaloid biosynthesis Polyketide synthase (AaPYKS), cytochrome p450 (AaP450), tropinone reductase (AaTRI, AaTRII) from *Anisodus acutangulus*	Transcriptome gene mining, characterization and functional expression of tropane alkaloid biosynthetic genes in bioengineered *S. cerevisiae*	Cui et al., 2015; Ping et al., 2019b
	Tropane alkaloids precursor (*S. cerevisiae*)	**Tropane alkaloid biosynthesis** Diamine oxidase (DAO) from *A. acutangulus*	Transcriptome gene mining, characterization and functional expression of diamine oxidase in *S. cerevisiae*	Cui et al., 2015; Ping et al., 2019a
	Phenylpropanoid precursor (*S. cerevisiae*)	**Phenylpropanoid biosynthesis** Cinnamate-4-hydroxylase (C4H) from *P. minus*	Transcriptome gene mining and expression of cinnamate-4-hydroxylase in *S. cerevisiae*	Loke et al., 2017; Ramzi et al., 2018
	Breviscapine flavonoid (*S. cerevisiae*)	**Breviscapine flavonoid biosynthesis** Flavonoid-7-*O*-glucuronosyltransferase (F7GAT) and flavone-6-hydroxylase (F6H) from *E. breviscapus*	Transcriptome gene mining, characterization and reconstitution of complete breviscapine flavonoid pathway from *E. breviscapus* in bioengineered *S. cerevisiae*	Liu et al., 2018
	Triterpenoid saponin (*S. cerevisiae*)	**Triterpenoid saponin biosynthesis** Cucurbitadienol synthase (SgCDS), epoxide hydrolase (SgEPH3EPH), cytochrome p450 (SgCYP87D18) from *Siraitia grosvenorii*	Production of mogroside V compounds by bioengineered *S. cerevisiae* expressing *S. grosvenorii* enzymes	Tang et al., 2011; Itkin et al., 2016
	Cannabinoids (*S. cerevisiae*)	**Cannabinoids biosynthesis** Prenyltransferases (CsPT), tetraketide synthase (*C. sativa* TKS; CsTKS), olivetolic acid cyclase (CsOAC), acyl activating enzyme (AAE). Cannabinoid synthases THCAS and CBDAS from *C. sativa*	Transcriptome gene mining, characterization and reconstitution of cannabinoid biosynthetic pathway in bioengineered *S. cerevisiae*	van Bakel et al., 2011; Luo et al., 2019

cannabinoids and triterpenoid saponin that were naturally derived from *Cannabis sativa* L. and *Siraitia grosvenorii*, respectively. Complete biosynthesis of cannabinoids was demonstrated through the expression of *Cannabis* enzymes that include newly identified *Cannabis* candidate prenyltransferases that are responsible for the conversion of olivetolic acid and GPP supplied by native MVA and heterologous hexanoyl-CoA biosynthetic pathways, respectively (Luo et al., 2019). Using a MVA-dependent squalene pathway in *S. cerevisiae*, the biosynthesis of triterpenoid mogrol compounds was achieved *via* pathway reconstitution and heterologous expression of cucurbitadienol synthase, epoxide hydrolase, and cytochrome p450 identified from *S. grosvenorii* transcriptome (Itkin et al., 2016). Overall, the utilization and expression of transcriptome-derived plant biosynthetic genes represent an increasingly valuable and feasible strategy in pathway engineering and natural product biomanufacturing using bioengineered microbes as cell factories.

THE WAY FORWARD: STREAMLINING NATURAL PRODUCTS BIOMANUFACTURING WITH OMICS AND ML PLATFORMS

To date, model microbes, especially *E. coli* and *S. cerevisiae*, represent the most suitable natural product chassis for strain improvement and biological engineering using DBTL iteration and upscaling processes owing to increased availability of genetic parts and biological data, including GEMs and omics datasets. As discussed earlier, omics technologies have been valuable in enhancing synthetic biology applications, but progress remains in accelerating the Learn step needed to inform the next Design phase and consequent DBTL cycles important in improving the desirable specification and biomanufacturing capacities. Recent progress in advanced genomics and synthetic biology has seen the increased adoption of ML-based data training and non-biased predictive tools for analyzing biological datasets to complement the biology-informed systems biology approaches. The predictive ability of ML tools is empowered through training and learning of experimental data *via* statistical linkage and modeling of independent and dependent variables as input and output data, respectively (Radivojević et al., 2020). Critically, the employment of ML approaches in strain design and optimization is gaining much interest, which is expected to address limitations in biology-informed approaches and circumvent the needs for detailed mechanistic understanding and resource constraints (Carbonell et al., 2019; Presnell and Alper, 2019).

Improving DBTL Performance and Predictive Capacities With ML and Omics Tools

ML-based training of biological datasets has been successfully used in microbial hosts in the efforts to improve gene annotation, metabolic pathway optimization, and fermentation bioprocess parameters (Kim et al., 2020). The bioproduction of specialty and fine chemicals, such as dodecanol and limonene, has been demonstrated in engineered *E. coli* and *S. cerevisiae* using ML-generated predictive models, which enabled unbiased genetic designs and combination (Zhou et al., 2018; Jervis et al., 2019; Opgenorth et al., 2019). A key advantage of utilizing ML tools is the development of a pure *in silico* system applicable for the Design and Learn phases that enable the selection of high-performing biological system without the needs to perform extensive and costly *in vivo* screening experiments. In the Design step, several ML tools have been developed for optimizing gene expression and cellular protein synthesis through *de novo* and quantitative design of genetic parts including promoter, 5′-untranslated region (5′UTR), and ribosomal binding site (RBS) in addition to the use of ML-assisted directed evolution and semi-rational protein engineering strategies (Decoene et al., 2018; Jervis et al., 2019; Wu et al., 2019). By training of partial least square (PLS) regression model on fluorescence output of a yeast UTR (yUTR) library, a newly constructed yUTR calculator was used to accurately predict the outcome of translation initiation rates in *S. cerevisiae* (Decoene et al., 2018). Employment of the predictive yUTR calculator enabled a tailored *in vivo* p-CA production in tyrosine ammonia lyase (TAL1)-expressing *S. cerevisiae* in accordance to the strengths of *de novo* and native 5′UTR with weak and high predicted protein abundance (Decoene et al., 2018).

Modulation and improvement of terpenes production has also been demonstrated through ML-enabled fine-tuning of gene expression by synthetic promoters and RBS of the MVA and non-MVA DXP biosynthetic genes (Meng et al., 2013; Jervis et al., 2019). Using a mutated Trc promoter and RBS sequences for artificial neural network (NN)-based model training and test, the expression of 1-deoxy-D-xylulose-5-phosphate synthase (DXS) gene under the control low-strength synthetic s14 promoter enhanced the production of amorphadiene in engineered *E. coli* (Meng et al., 2013). A recent report (Jervis et al., 2019) has expanded the use of a feedforward NN-based ML model on *de novo* design and screening of synthetic RBS for MVA pathway engineering and bioproduction of limonene where model training was conducted on expression levels of HMG-CoA synthase (HMGS), HMG-CoA reductase (HMGR), MVA kinase (MK), and IPP delta-isomerase (IDI) using multiple combinations of RBS sequences. The constructed library of 32 RBS combinations was then built and tested in combination with terpene-synthesizing pGL403 plasmid construct that resulted in the identification and selection of high-performing *E. coli* strains with improved limonene titer over 1.5–3-fold (Jervis et al., 2019).

The combination of omics datasets and ML strategies is expected to drive the production of natural products and other biobased chemicals especially in terms of biosynthetic pathway inference, refactoring, and optimization. The self-organizing map (SOM) approach represents an unsupervised NN method useful in the identification of new enzymes using plant transcriptome datasets to complement common gene co-expression analysis, such as differentially expressed genes (DEG) method (Dugé de Bernonville et al., 2020). The SOM-assisted co-expression analysis of *Rauvolfia serpentina* transcriptome has led to the identification of sarpagan bridge enzyme (SBE) and vinorine

hydroxylase (VH) essential in sarpagan and ajmalan alkaloid biosynthesis that could be useful in the Build and Test of these high-value bioproducts in engineered microbial chassis (Dang et al., 2017, 2018). A supervised ML platform has been developed and tested using proteome and metabolome datasets of biofuel- and terpene-producing *E. coli* where the ML-driven model predictions yielded an accurate *in silico* pathway design and outperformed classical Michaelis–Menten kinetic modeling (Costello and Martin, 2018). In their report, a Tree-based Pipeline Optimization Tool (TPOT) was used for training data and succeeded in generating models for dynamically predicting medium level limonene-producing *E. coli* strains using experimental omics datasets, thus providing a pure ML and omic dataset-based virtual strain simulation and pathway construction (Costello and Martin, 2018). Interestingly, another recent report by Radivojević et al. (2020) leveraged on ensemble approach and probabilistic modeling methods to construct a ML-based Automated Recommendation Tool (ART) useful for improving microbial engineering and DBTL bioproduction performance by training of proteome datasets among a host of experimental data as input variables. By comparing limonene bioproduction improvement in engineered *E. coli* guided by experimentally tested PCAP, the ML models generated by ART were suggested to be able to match and further enhance the production of a given product through the DBTL cycle by recommending new inputs, such as transcriptome datasets and promoter strengths in the next Design phase. Following this, the integration of transcriptome, proteome, and/or metabolome datasets with ML methods is particularly useful in the development of mathematical models in the Test and Learn cycle that would guide and facilitate *in silico* optimization of the DBTL pipeline (Presnell and Alper, 2019; St. John and Bomble, 2019; Volk et al., 2020). Thanks to the growing list of genome, transcriptome, and GEM resources, further adoption and implementation of *in silico* and ML tools on these biological datasets are expected to bring about a markedly improved and accurate predictive engineering and retrosynthetic design of metabolic pathways to existing and new-to-nature chemicals (Lin et al., 2019; Zhang et al., 2020). In line with the emergence of data-driven 4th Industrial Revolution (4IR), the applications of omics and ML tools in strain and bioproduct development are set to be the cornerstone in industrial biomanufacturing of biobased chemicals and pharmaceuticals.

CONCLUSION AND PERSPECTIVES

Overall, it is envisioned that the employment of data-centered omics and ML platforms will lead to more streamlined and less resource-intensive biomanufacturing strategies and accelerate strain prototyping pipelines that have been a major stumbling block in the translation of bioproduct development from laboratory to market. Omics-guided microbial engineering and ML-assisted biomanufacturing will therefore bring about data-driven biomanufacturing pipelines that can be expanded to include metagenome datasets and accelerate the bioproduction of industrially relevant biomolecules and drugs tailored to the pressing needs of medical, agricultural, environmental, and industrial sectors.

AUTHOR CONTRIBUTIONS

ABR, SNB, HB, and NSS all contributed toward the writing and editing of this manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Yayasan Penyelidikan Antartika Sultan Mizan (YPASM) Fellowship Scheme (RB-2019-001) and Geran Universiti Penyelidikan (GUP-2019-065) research grants. The article processing fee for this work is funded by Dana Pecutan Penerbitan INBIOSIS UKM (PP-INBIOSIS-2020), a publication initiative grant awarded to the Institute of Systems Biology (INBIOSIS) under UKM Research University Grant.

REFERENCES

Alonso-Gutierrez, J., Kim, E.-M., Batth, T. S., Cho, N., Hu, Q., Chan, L. J. G., et al. (2015). Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering. *Metab. Eng.* 28, 123–133. doi: 10.1016/j.ymben.2014.11.011

Benjamin, K. R., Silva, I. R., Cherubim, J. P., McPhee, D., and Paddon, C. J. (2016). Developing commercial production of semi-synthetic artemisinin, and of β-farnesene, an isoprenoid produced by fermentation of Brazilian sugar. *J. Braz. Chem. Soc.* 27, 1339–1345. doi: 10.5935/0103-5053.20160119

Borja, G. M., Rodriguez, A., Campbell, K., Borodina, I., Chen, Y., and Nielsen, J. (2019). Metabolic engineering and transcriptomic analysis of *Saccharomyces cerevisiae* producing p-coumaric acid from xylose. *Microb. Cell Fact.* 18, 191. doi: 10.1186/s12934-019-1244-4

Brunk, E., George, K. W., Alonso-Gutierrez, J., Thompson, M., Baidoo, E., Wang, G., et al. (2016). Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow. *Cell Syst.* 2, 335–346. doi: 10.1016/j.cels.2016.04.004

Caputi, L., Franke, J., Farrow, S. C., Chung, K., Payne, R. M. E., Nguyen, T.-D., et al. (2018). Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Science* 360, 1235–1239. doi: 10.1126/science.aat4100

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., et al. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.* 1:66. doi: 10.1038/s42003-018-0076-9

Carbonell, P., Radivojevic, T., and García Martín, H. (2019). Opportunities at the intersection of synthetic biology, machine learning, and automation. *ACS Synth. Biol.* 8, 1474–1477. doi: 10.1021/acssynbio.8b00540

Carqueijeiro, I., Langley, C., Grzech, D., Koudounas, K., Papon, N., O'Connor, S. E., et al. (2020). Beyond the semi-synthetic artemisinin: metabolic engineering of plant-derived anti-cancer drugs. *Curr. Opin. Biotechnol.* 65, 17–24. doi: 10.1016/j.copbio.2019.11.017

Casini, A., Chang, F.-Y., Eluere, R., King, A. M., Young, E. M., Dudley, Q. M., et al. (2018). A pressure test to make 10 molecules in 90 days: external evaluation of methods to engineer biology. *J. Am. Chem. Soc.* 140, 4302–4316. doi: 10.1021/jacs.7b13292

Chae, T. U., Choi, S. Y., Kim, J. W., Ko, Y.-S., and Lee, S. Y. (2017). Recent advances in systems metabolic engineering tools and strategies. *Curr. Opin. Biotechnol.* 47, 67–82. doi: 10.1016/j.copbio.2017.06.007

Chen, Y., Banerjee, D., Mukhopadhyay, A., and Petzold, C. J. (2020). Systems and synthetic biology tools for advanced bioproduction hosts. *Curr. Opin. Biotechnol.* 64, 101–109. doi: 10.1016/j.copbio.2019.12.007

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., and Lee, S. Y. (2019). Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37, 817–837. doi: 10.1016/j.tibtech.2019.01.003

Costello, Z., and Martin, H. G. (2018). A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst. Biol. Appl.* 4:19. doi: 10.1038/s41540-018-0054-3

Courdavault, V., O'Connor, S. E., Oudin, A., Besseau, S., and Papon, N. (2020). Towards the microbial production of plant-derived anticancer drugs. *Trends Cancer* 6, 444–448. doi: 10.1016/j.trecan.2020.02.004

Cui, L., Huang, F., Zhang, D., Lin, Y., Liao, P., Zong, J., et al. (2015). Transcriptome exploration for further understanding of the tropane alkaloids biosynthesis in *Anisodus acutangulus*. *Mol. Genet. Genomics.* 290, 1367–1377. doi: 10.1007/s00438-015-1005-y

Dahal, S., Yurkovich, J. T., Xu, H., Palsson, B. O., and Yang, L. (2020). Synthesizing systems biology knowledge from omics using genome-scale models. *Proteomics* 20:e1900282. doi: 10.1002/pmic.201900282

Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., et al. (2013). Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* 31, 1039–1046. doi: 10.1038/nbt.2689

Dang, T. T. T., Franke, J., Carqueijeiro, I. S. T., Langley, C., Courdavault, V., and O'Connor, S. E. (2018). Sarpagan bridge enzyme has substrate-controlled cyclization and aromatization modes. *Nat. Chem. Biol.* 14, 760–763. doi: 10.1038/s41589-018-0078-4

Dang, T. T. T., Franke, J., Tatsis, E., and O'Connor, S. E. (2017). Dual catalytic activity of a cytochrome p450 controls bifurcation at a metabolic branch point of alkaloid biosynthesis in *Rauwolfia serpentina*. *Angew. Chem. Int. Ed. Engl.* 56, 9440–9444. doi: 10.1002/anie.201705010

Decoene, T., Peters, G., De Maeseneire, S. L., and De Mey, M. (2018). Toward predictable 5'UTRs in *Saccharomyces cerevisiae*: development of a yUTR calculator. *ACS Synth. Biol.* 7, 622–634. doi: 10.1021/acssynbio.7b00366

DeLoache, W. C., Russ, Z. N., Narcross, L., Gonzales, A. M., Martin, V. J. J., and Dueber, J. E. (2015). An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat. Chem. Biol.* 11, 465–471. doi: 10.1038/nchembio.1816

Desgagné-Penix, I., Farrow, S. C., Cram, D., Nowak, J., and Facchini, P. J. (2012). Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy. *Plant Mol. Biol.* 79, 295–313. doi: 10.1007/s11103-012-9913-2

Dugé de Bernonville, T., Papon, N., Clastre, M., O'Connor, S. E., and Courdavault, V. (2020). Identifying missing biosynthesis enzymes of plant natural products. *Trends Pharmacol. Sci.* 41, 142–146. doi: 10.1016/j.tips.2019.12.006

Dunstan, M. S., Robinson, C. J., Jervis, A. J., Yan, C., Carbonell, P., Hollywood, K. A., et al. (2020). Engineering *Escherichia coli* towards de novo production of gatekeeper (2S)-flavanones: naringenin, pinocembrin, eriodictyol and homoeriodictyol. *Synth. Biol.* 5:ysaa012. doi: 10.1093/synbio/ysaa012

Ekas, H., Deaner, M., and Alper, H. S. (2019). Recent advancements in fungal-derived fuel and chemical production and commercialization. *Curr. Opin. Biotechnol.* 57, 1–9. doi: 10.1016/j.copbio.2018.08.014

Facchini, P. J., and De Luca, V. (2008). Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. *Plant J.* 54, 763–784. doi: 10.1111/j.1365-313X.2008.03438.x

Farrow, S. C., Hagel, J. M., Beaudoin, G. A. W. W., Burns, D. C., and Facchini, P. J. (2015). Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. *Nat. Chem. Biol.* 11, 728–732. doi: 10.1038/nchembio.1879

Fossati, E., Ekins, A., Narcross, L., Zhu, Y., Falgueyret, J. P., Beaudoin, G. A. W., et al. (2014). Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in *Saccharomyces cerevisiae*. *Nat. Commun.* 5:3283. doi: 10.1038/ncomms4283

Fridman, E., and Pichersky, E. (2005). Metabolomics, genomics, proteomics, and the identification of enzymes and their substrates and products. *Curr. Opin. Plant Biol.* 8, 242–248. doi: 10.1016/j.pbi.2005.03.004

Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., Smolke, C. D., Interrante, M. F., et al. (2015). Complete biosynthesis of opioids in yeast. *Science* 349, 1095–1100. doi: 10.1126/science.aac9373

Gao, S., Zhou, H., Zhou, J., and Chen, J. (2020). Promoter-library-based pathway optimization for efficient (2S)-naringenin production from p-coumaric acid in *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* 68, 6884–6891. doi: 10.1021/acs.jafc.0c01130

Goh, H. H. (2018). Integrative multi-omics through bioinformatics. *Adv. Exp. Med. Biol.* 1102, 69–80. doi: 10.1007/978-3-319-98758-3_5

Goh, H. H., Ng, C. L., and Loke, K. K. (2018). Functional genomics. *Adv. Exp. Med. Biol.* 1102, 11–30. doi: 10.1007/978-3-319-98758-3_2

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., and Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biol.* 20:121. doi: 10.1186/s13059-019-1730-3

Itkin, M., Davidovich-Rikanati, R., Cohen, S., Portnoy, V., Doron-Faigenboim, A., Oren, E., et al. (2016). The biosynthetic pathway of the nonsugar, high-intensity sweetener mogroside V from *Siraitia grosvenorii*. *Proc. Natl. Acad. Sci.* 113, E7619–E7628. doi: 10.1073/pnas.1604828113.

Jamil, I. N., Remali, J., Azizan, K. A., Nor Muhammad, N. A., Arita, M., Goh, H.-H., et al. (2020). Systematic multi-omics integration (MOI) approach in plant systems biology. *Front. Plant Sci.* 11:944. doi: 10.3389/fpls.2020.00944

Jervis, A. J., Carbonell, P., Vinaixa, M., Dunstan, M. S., Hollywood, K. A., Robinson, C. J., et al. (2019). Machine learning of designed translational control allows predictive pathway optimization in *Escherichia coli*. *ACS Synth. Biol.* 8, 127–136. doi: 10.1021/acssynbio.8b00398

Kim, G. B., Kim, W. J., Kim, H. U., and Lee, S. Y. (2020). Machine learning applications in systems metabolic engineering. *Curr. Opin. Biotechnol.* 64, 1–9. doi: 10.1016/j.copbio.2019.08.010

Kirby, J., Dietzel, K. L., Wichmann, G., Chan, R., Antipov, E., Moss, N., et al. (2016). Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in *Saccharomyces cerevisiae*. *Metab. Eng.* 38, 494–503. doi: 10.1016/j.ymben.2016.10.017

Kitano, H. (2002). Systems biology: a brief overview. *Science* 295, 1662–1664. doi: 10.1126/science.1069492

Ku Bahaudin, K. N. A., Ramzi, A. B., Baharum, S. N., Sabri, S., Thean Chor, A. L., Sabri, S., et al. (2018). Current progress in production of flavonoids using systems and synthetic biology platforms. *Sains Malays.* 47, 3077–3084. doi: 10.17576/jsm-2018-4712-18

Lee, S. Y., and Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072. doi: 10.1038/nbt.3365

Li, Y., and Smolke, C. D. (2016). Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. *Nat. Commun.* 7:12137. doi: 10.1038/ncomms12137

Lin, G. M., Warden-Rothman, R., and Voigt, C. A. (2019). Retrosynthetic design of metabolic pathways to chemicals not found in nature. *Curr. Opin. Syst. Biol.* 14, 82–107. doi: 10.1016/j.coisb.2019.04.004

Lister, R., Gregory, B. D., and Ecker, J. R. (2009). Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr. Opin. Plant Biol.* 12, 107–118. doi: 10.1016/j.pbi.2008.11.004

Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., et al. (2018). Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. *Nat. Commun.* 9:448. doi: 10.1038/s41467-018-02883-z

Loke, K.-K., Rahnamaie-Tajadod, R., Yeoh, C.-C., Goh, H.-H., Mohamed-Hussein, Z. A., Zainal, Z., et al. (2017). Transcriptome analysis of *Polygonum minus* reveals candidate genes involved in important secondary metabolic pathways of phenylpropanoids and flavonoids. *PeerJ* 5:e2938. doi: 10.7717/peerj.2938

Luo, X., Reiter, M. A., d'Espaux, L., Wong, J., Denby, C. M., Lechner, A., et al. (2019). Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 567, 123–126. doi: 10.1038/s41586-019-0978-9

Matasci, N., Hung, L. H., Yan, Z., Carpenter, E. J., Wickett, N. J., Mirarab, S., et al. (2014). Data access for the 1,000 Plants (1KP) project. *Gigascience.* 3:2047-217X-3-17. doi: 10.1186/2047-217X-3-17

Meng, H., Wang, J., Xiong, Z., Xu, F., Zhao, G., and Wang, Y. (2013). Quantitative design of regulatory elements based on high-precision strength prediction using artificial neural network. *PLoS One* 8:e60288. doi: 10.1371/journal.pone.0060288

Newman, D. J., and Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* 79, 629–661. doi: 10.1021/acs.jnatprod.5b01055

Newman, D. J., and Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* 83, 770–803. doi: 10.1021/acs.jnatprod.9b01285

Nielsen, J., and Keasling, J. D. (2016). Engineering cellular metabolism. *Cell* 164, 1185–1197. doi: 10.1016/j.cell.2016.02.004

O'Brien, E. J., Monk, J. M., and Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell* 161, 971–987. doi: 10.1016/j.cell.2015.05.019

Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., et al. (2019). Lessons from two Design–Build–Test–Learn cycles of dodecanol production in *Escherichia coli* aided by machine learning. *ACS Synth. Biol.* 8, 1337–1351. doi: 10.1021/acssynbio.9b00020

Ping, Y., Li, X., Xu, B., Wei, W., Wei, W., Kai, G., et al. (2019a). Building microbial hosts for heterologous production of N-methylpyrrolinium. *ACS Synth. Biol.* 8, 257–263. doi: 10.1021/acssynbio.8b00483

Ping, Y., Li, X., You, W., Li, G., Yang, M., Wei, W., et al. (2019b). *De novo* production of the plant-derived tropine and pseudotropine in yeast. *ACS Synth. Biol.* 8, 1257–1262. doi: 10.1021/acssynbio.9b00152

Presnell, K. V., and Alper, H. S. (2019). Systems metabolic engineering meets machine learning: a new era for data-driven metabolic engineering. *Biotechnol. J.* 14:e1800416. doi: 10.1002/biot.201800416

Pyne, M. E., Narcross, L., and Martin, V. J. J. (2019). Engineering plant secondary metabolism in microbial systems. *Plant Physiol.* 179, 844–861. doi: 10.1104/pp.18.01291

Radivojević, T., Costello, Z., Workman, K., and Garcia Martin, H. (2020). A machine learning automated recommendation tool for synthetic biology. *Nat. Commun.* 11, 4879. doi: 10.1038/s41467-020-18008-4

Rai, A., Saito, K., and Yamazaki, M. (2017). Integrated omics analysis of specialized metabolism in medicinal plants. *Plant J.* 90, 764–787. doi: 10.1111/tpj.13485

Ramzi, A. B. (2018). Metabolic engineering and synthetic biology. *Adv. Exp. Med. Biol.* 1102, 81–95. doi: 10.1007/978-3-319-98758-3_6

Ramzi, A. B., Ku Bahaudin, K. N. A., Baharum, S. N., Che Me, M. L., Goh, H.-H., Hassan, M., et al. (2018). Rapid assembly of yeast expression cassettes for phenylpropanoid biosynthesis in *Saccharomyces cerevisiae*. *Sains Malays.* 47, 2969–2974. doi: 10.17576/jsm-2018-4712-05

Redding-Johanson, A. M., Batth, T. S., Chan, R., Krupa, R., Szmidt, H. L., Adams, P. D., et al. (2011). Targeted proteomics for metabolic pathway optimization: application to terpene production. *Metab. Eng.* 13, 194–203. doi: 10.1016/j.ymben.2010.12.005

Robinson, C. J., Carbonell, P., Jervis, A. J., Yan, C., Hollywood, K. A., Dunstan, M. S., et al. (2020). Rapid prototyping of microbial production strains for the biomanufacture of potential materials monomers. *Metab. Eng.* 60, 168–182. doi: 10.1016/j.ymben.2020.04.008

Rodriguez, A., Chen, Y., Khoomrung, S., Özdemir, E., Borodina, I., and Nielsen, J. (2017). Comparison of the metabolic response to over-production of p-coumaric acid in two yeast strains. *Metab. Eng.* 44, 265–272. doi: 10.1016/j.ymben.2017.10.013

Rusdi, N. A., Goh, H.-H., Sabri, S., Ramzi, A. B., Mohd Noor, N., and Baharum, S. N. (2018). Functional characterisation of new sesquiterpene synthase from the Malaysian herbal plant, *Polygonum minus*. *Molecules* 23, 1370. doi: 10.3390/molecules23061370

Scossa, F., Benina, M., Alseekh, S., Zhang, Y., and Fernie, A. R. (2018). The integration of metabolomics and next-generation sequencing data to elucidate the pathways of natural product metabolism in medicinal plants. *Planta Med.* 84, 855–873. doi: 10.1055/a-0630-1899

Seaver, S. M. D., Henry, C. S., and Hanson, A. D. (2012). Frontiers in metabolic reconstruction and modeling of plant genomes. *J. Exp. Bot.* 63, 2247–2258. doi: 10.1093/jxb/err371

Sheth, B. P., and Thaker, V. S. (2014). Plant systems biology: insights, advances and challenges. *Planta* 240, 33–54. doi: 10.1007/s00425-014-2059-5

Srinivasan, P., and Smolke, C. D. (2019). Engineering a microbial biosynthesis platform for de novo production of tropane alkaloids. *Nat. Commun.* 10:3634. doi: 10.1038/s41467-019-11588-w

St. John, P. C., and Bomble, Y. J. (2019). Approaches to computational strain design in the multiomics era. *Front. Microbiol.* 10:597. doi: 10.3389/fmicb.2019.00597

Tan, C. S., Hassan, M., Mohamed Hussein, Z. A., Ismail, I., Ho, K. L., Ng, C. L., et al. (2018). Structural and kinetic studies of a novel nerol dehydrogenase from *Persicaria minor*, a nerol-specific enzyme for citral biosynthesis. *Plant Physiol. Biochem.* 123, 359–368. doi: 10.1016/j.plaphy.2017.12.033

Tang, Q., Ma, X., Mo, C., Wilson, I. W., Song, C., Zhao, H., et al. (2011). An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis. *BMC Genomics* 12:343. doi: 10.1186/1471-2164-12-343

van Bakel, H., Stout, J. M., Cote, A. G., Tallon, C. M., Sharpe, A. G., Hughes, T. R., et al. (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12:R102. doi: 10.1186/gb-2011-12-10-r102

Volk, M. J., Lourentzou, I., Mishra, S., Vo, L. T., Zhai, C., and Zhao, H. (2020). Biosystems design by machine learning. *ACS Synth. Biol.* 9, 1514–1533. doi: 10.1021/acssynbio.0c00129

Wu, Z., Jennifer Kan, S. B., Lewis, R. D., Wittmann, B. J., and Arnold, F. H. (2019). Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8852–8858. doi: 10.1073/pnas.1901979116

Xiao, M., Zhang, Y., Chen, X., Lee, E. J., Barber, C. J. S., Chakrabarty, R., et al. (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.* 166, 122–134. doi: 10.1016/j.jbiotec.2013.04.004

Zhang, J., Petersen, S. D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A., Abeliuk, E., et al. (2020). Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. *Nat. Commun.* 11:4880. doi: 10.1038/s41467-020-17910-1

Zhou, Y., Li, G., Dong, J., Xing, X. H., Dai, J., et al. (2018). MiYA, an efficient machine-learning workflow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*. *Metab. Eng* 47, 294–302. doi: 10.1016/j.ymben.2018.03.020

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 22 January 2021
doi: 10.3389/fbioe.2020.622226

Engineering a Synthetic Pathway for Gentisate in *Pseudomonas Chlororaphis* P3

Songwei Wang[1], *Cong Fu*[1], *Kaiquan Liu*[2], *Jiajia Cui*[1], *Hongbo Hu*[1], *Wei Wang*[1] *and Xuehong Zhang*[1]*

[1] State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, [2] State Key Laboratory of Biobased Material and Green Papermaking (LBMP), Department of Bioengineering, Qilu University of Technology, Shandong Academy of Sciences, Jinan, China

Pseudomonas chlororaphis P3 has been well-engineered as a platform organism for biologicals production due to enhanced shikimate pathway and excellent physiological and genetic characteristics. Gentisate displays high antiradical and antioxidant activities and is an important intermediate that can be used as a precursor for drugs. Herein, a plasmid-free biosynthetic pathway of gentisate was constructed by connecting the endogenous degradation pathway from 3-hydroxybenzoate in *Pseudomonas* for the first time. As a result, the production of gentisate reached 365 mg/L from 3-HBA via blocking gentisate conversion and enhancing the gentisate precursors supply through the overexpression of the rate-limiting step. With a close-up at the future perspectives, a series of bioactive compounds could be achieved by constructing synthetic pathways in conventional *Pseudomonas* to establish a cell factory.

Keywords: ***Pseudomonas chlororaphis*** **P3, gentisate, biosynthesis, plasmid-free, bioactive compounds, cell factory**

OPEN ACCESS

Edited by:
Jingwen Zhou,
Jiangnan University, China

Reviewed by:
Mingfeng Cao,
University of Illinois at Urbana-Champaign, United States
Dae-Hee Lee,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), South Korea

***Correspondence:**
Xuehong Zhang
xuehzhang@sjtu.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *28 October 2020*
Accepted: *30 December 2020*
Published: *22 January 2021*

Citation:
Wang S, Fu C, Liu K, Cui J, Hu H, Wang W and Zhang X (2021) Engineering a Synthetic Pathway for Gentisate in Pseudomonas Chlororaphis P3. Front. Bioeng. Biotechnol. 8:622226. doi: 10.3389/fbioe.2020.622226

INTRODUCTION

Growing attention to environmental problems and energy crises has inspired the development of bio-based production of valuable bioproducts over the past few decades (Choi et al., 2015; Liao et al., 2016; Noda et al., 2017). Microbial-based synthetic biology and metabolic engineering are eco-friendly approaches for producing valuable biochemicals from sustainable carbon sources. Hydroxybenzoic acids and their derivatives are widely used as additives in foods, drugs, and cosmetics for antisepsis and flavor preservation, or as a monomer to synthesize bioactive compounds (Wang et al., 2018a; Shen et al., 2020). Besides, hydroxybenzoic acids play essential roles in microbial metabolism by serving as intermediates of the degradation of the aromatic compounds and contributing to the synthesis of various valuable secondary metabolites. Therefore, it is essential to explore the metabolism of hydroxybenzoic acids in microbial hosts to synthesize new natural products and improve the methods for overproduction of valuable compounds.

Recently, *Pseudomonas* has received significant attention in synthetic biology due to its robustness and metabolic versatility (Belda et al., 2016; Wang et al., 2020). Genome database and tools for gene editing make it possible that *Pseudomonas* become a cell factory for bio-industrial application (Poblete-Castro et al., 2012; Wang et al., 2020). *Pseudomonas chlororaphis* P3 (*P. chlororaphis* P3) is one phenazine-1-carboxamide (PCN) producing biocontrol strain obtained from *P. chlororaphis* HT66 with multiple rounds of mutation and selection with enhanced

shikimate pathway based on the isobaric tags for relative and absolute quantification (iTRAQ)-based quantitative proteomic analysis (Jin et al., 2016). Based on the efficient shikimate pathway and simple cultivation conditions, *P. chlororaphis* P3 has been genetically engineered for the synthesis of arbutin (Wang et al., 2018c).

Shikimate pathway is the leading pathway for the synthesis of numerous aromatic compounds. Besides the aromatic amino acids, folic acid, ubiquinone, and phenazine antibiotics are also synthesized through the shikimate pathway (Averesch and Krömer, 2018; Wang et al., 2018b; Cao et al., 2020). There have been many researches focused on the synthesis of hydroxybenzoic acid and its derivatives, such as the construction of cell factory based on 4-hydroxybenzoic acid (4-HBA) for the synthesis of arbutin, muconic acid (MA), vanillyl alcohol, and other value-added products (Bai et al., 2016; Chen et al., 2017c; Wang et al., 2018c). In addition, salicylic acid (SA) and MA could also be synthesized by introducing isochorismate synthase in *E. coli* (Lin et al., 2014). Gentisate (GA) is an important intermediate with high antiradical and antioxidant activities that can be used as a precursor for drugs. According to an earlier study, plasmid-based 3-HBA expression systems were established that use antibiotics and inducers to ensure the hereditary stability of engineered strains (Kallscheuer and Marienhagen, 2018; Zhou et al., 2019), thus leaving environmental footprints (Keen and Patrick, 2013). In this context, we engineered chromosome-integrated synthetic pathways for 3-HBA and GA in *P. chlororaphis* P3. Exogenous 3-HBA synthetic enzyme was introduced, and then GA was biosynthesized by connecting the endogenous degradation pathway from 3-HBA. Also, we tried novel GA synthesis from 4-HBA (**Figure 1**). Strategies used in this research revealed *Pseudomonas*' versatility as a bioengineering strain, and this green microbial synthetic approach demonstrated its great potential of relieving environmental problems.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Culture Conditions

All strains and plasmids constructed or used in this study are listed in the Supporting Information (**Supplementary Tables 1**, **2**). *E. coli* and *P. chlororaphis* were cultured in Lysogeny Broth (LB) medium (Tryptone 10 g/L, Yeast extract 5 g/L, NaCl 10 g/L) during the construction of strains. King's medium B (KB) (Glycerol 18 g/L, Tryptone 20 g/L, $MgSO_4{\cdot}7H_2O$ 1.498 g/L, K_2HPO_4 0.514 g/L) was used for secondary metabolites production in *P. chlororaphis*. Agar was supplemented at a final concentration of 1.5% before sterilization. The medium was supplemented with the following antibiotics: 50 mg/L kanamycin and 100 mg/L ampicillin for screening of positive clones. To induce a double exchange of homologous recombination, sucrose was added to a final concentration of 15% before sterilization. *P. chlororaphis* was cultured at 28°C, while *E. coli* was cultured at 37°C. The shake flasks were filled with 60 mL medium and maintained at 28°C, 220 rpm.

P. chlororaphis strains were activated on KB agar medium and cultured overnight at 28°C. Single colonies were isolated and then inoculated to ∼50 mL KB medium in a flask. The primary pre-cultures were incubated at 28°C overnight. At the beginning of fermentation, the bacterial suspension was inoculated into a 250 mL shake flask containing 60 mL KB broth to reach an initial OD_{600} of 0.02. Samples of 1 mL were collected every 12 h for the determination of cell growth and metabolic products. Each fermentation test was conducted in triplicate.

DNA Techniques

All primers were designed by Primer Premier 5.0 (PREMIER Biosoft, San Francisco, USA), and then synthesized by Personalbio (Shanghai, China) (**Supplementary Table 3**). A sequence of *hyg5* from *Streptomyces hygroscopicus* ATCC 29253 was codon-optimized and synthesized by Genewiz (Suzhou, China). To construct plasmid for expressing *hyg5* in *pykA* locus, the 500 bp upstream and downstream DNA fragments of *pykA*, P_{phz} promoter, and open reading frame (ORF) of *hyg5* were amplified by PCR using PrimerSTAR Max DNA Polymerase. Then, the PCR products were purified with HiPure Gel Pure DNA Mini Kit (Magen, Guangzhou, China) after agarose gel electrophoresis. Next, these DNA fragments were assembled into the restriction enzyme-digested pk18*mobsacB* using In-Fusion Cloning Kit (TaKaRa Bio, Beijing, China). After transformation, plasmids were collected by HiPure Plasmid Micro Kit (Magen, Guangzhou, China). Gene deletion or substitution plasmids were constructed using the same method as reported previously (Wang et al., 2018b). The corresponding nucleotide sequences are presented in **Supplementary Table 3**.

The principle of strain construction is homologous recombination-mediated by suicide plasmids. Here, *pykA* deletion was taken as an example. Similar to early study (Wang et al., 2018b), plasmid pk18-Δ*pykA* containing *pykA* upstream and downstream fragments were constructed using the method mentioned above, and then transferred into S17-1 (λ pir). S17 containing pk18-Δ*pykA* and P3 were inoculated into LB liquid medium and then cultured overnight. To maintain the stability of plasmid, 50 mg/L Kan was added into the medium. A few milliliters of cell suspension were then centrifugated, and the bacteria were mixed with LB liquid medium. After incubation at 28°C for 1–2 h, the mixture was incubated again on a LB solid medium plate at 28 °C for 24–36 h. The mixed bacterial cells were scraped from LB plate, resuspended in 200 μL LB liquid medium, coated on a new LB plate containing 50 mg/L Kan and 100 mg/L Amp, and incubated at 28°C. A single colony was selected, diluted with LB liquid medium to a specific proportion, and then coated on a LB plate containing 15% sucrose. After 36 h of culture, the colonies were selected and cultured on LB plates containing Kan or Amp. The colonies that grow on LB (Amp) plates but not on LB (Kan) plates are positive transformants. To screen the strains, PCR was conducted using pykA-1F and pykA-2R primers. After agarose gel electrophoresis, two or three suspected mutant strains were determined and cultured overnight. Genomic DNA was extracted by HiPure Bacterial DNA Kit (Magen, Guangzhou, China) and used as the template of PCR amplification to verify

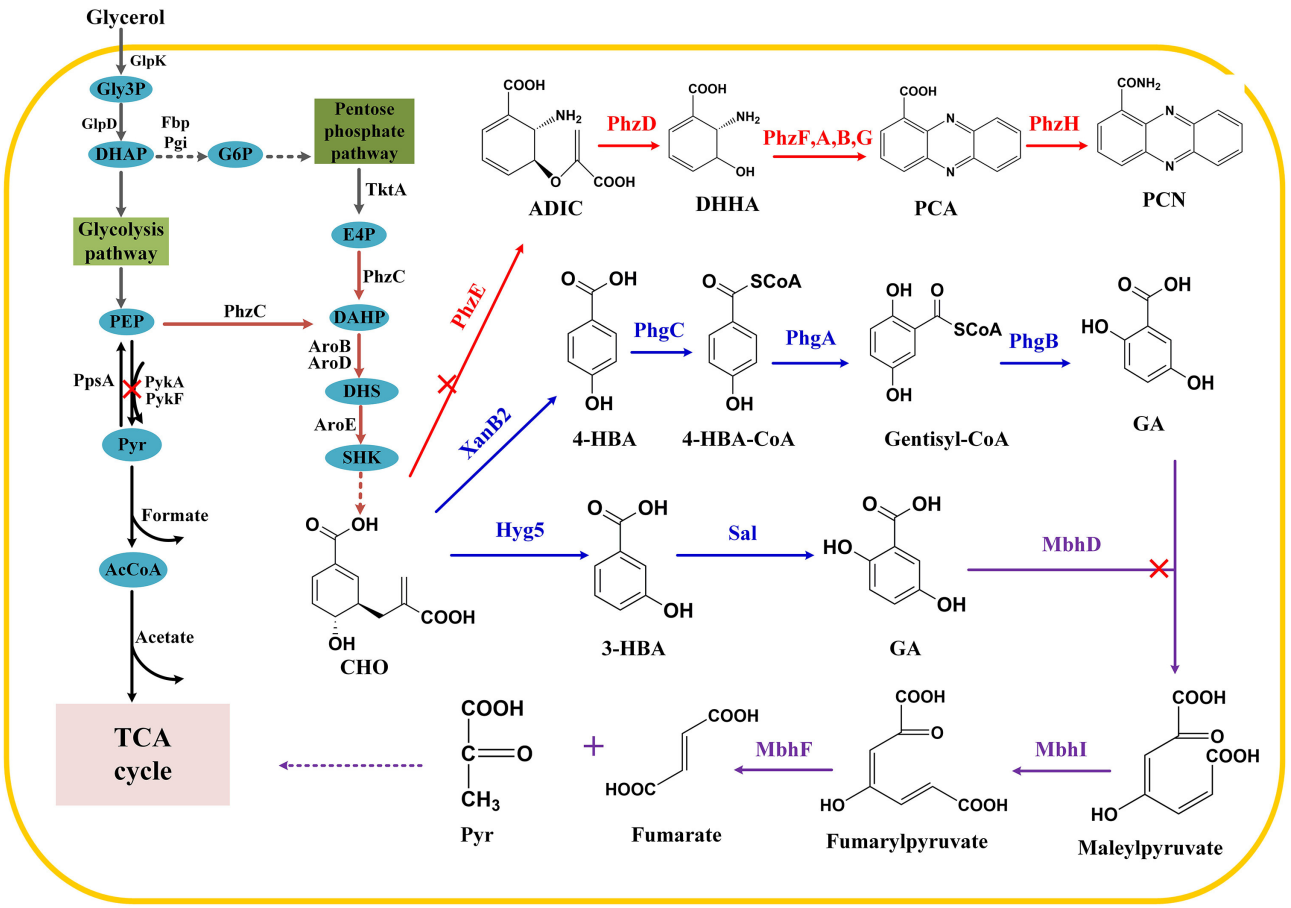


FIGURE 1 | Construction of synthetic pathway for gentisate in *P. chlororaphis* P3. Gly3P, glycerol-3-phosphate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; G6P, glucose-6-phosphate; Pyr, pyruvate; PEP, phosphoenolpyruvate; E4P, erythrose 4-phosphate; DAHP, 3-deoxy-D-arabino-heptulosonate 7-phosphate; DHS, 3-dehydroshikimate; SHK, shikimate; CHO, chorismate; 4-HBA, 4-hydroxybenzoate; 3-HBA, 3-hydroxybenzoate; ADIC, 2-amino-2-desoxyisochorismic acid; DHHA, trans-2,3-dihydro-3-hydroxyanthranilic acid; PCA, phenazine-1-carboxylic acid; PCN, phenazine-1-carboxamide. Main enzymes invloved, TktA: pyruvate synthase; PhzC DAHP synthase; Hyg5: 3-hydroxybenzoate synthase; Sal: salicylate hydroxylase; PykA PykF: pyruvate kinase; PpsA: PEP synthase; XanB2: chorismate lyase; PhzE: anthranilate synthase; PhzDFABD: phenazine synthetic protein; MhbD: gentisate 1,2-dioxygenase; MhbI: maleylacetoacetate isomerase.

pykA deletion, and no mutation occurred in homologous arms. In this way, a *pykA*-deleted strain was successfully constructed.

Whole-cell Transformation

BL21 (DE3) strains (i.e., BL21-Sal, BL21-PobA, and BL21-PobAM) were activated on LB plates, cultured overnight at 37°C, and then inoculated to 60 mL LB medium in shake flasks. Isopropyl-β-D–thiogalactopyranoside (IPTG) was used to induce 60 mL cultures at OD_{600} of 0.2–0.6. After overnight incubation at 16°C, the cells were gathered by centrifugal precipitation, washed with 50 mM phosphate buffer (pH = 8.0), and resuspended in 30 mL phosphate buffer at a final OD_{600} of 5.0. Following the addition of 30 μg NADH and 15 mg 3-HBA, the cell suspensions were incubated at 28°C with continuous shaking at 220 rpm. All samples were collected for analysis every 6 h.

Analytical Methods

The absorbance of cell suspension at 600 nm was determined with an ultraviolet spectrophotometer to measure the number of cells. High-performance liquid chromatography (HPLC) method was established for determining the contents of metabolites. Samples were collected during fermentation at specific time points and then centrifuged at 12,000 rpm for 5 min. Subsequently, the supernatant was filtered using nylon filters with an aperture span of 0.2 μm. The samples of 3-HBA and GA were detected using an Agilent Technologies 1,260 Infinity HPLC system with a C18 reversed-phase column at 30°C and 1 mL/min (constant flow rate). The concentrations of products were determined using an ultraviolet absorbance detector at 235 nm, and the injection volume was 20 μL. The mobile phase consisted of solvent A (methanol) and solvent B (water containing 0.1% formic acid). The separation of metabolites was

carried out via gradient elution under the following conditions: 0–2 min, 5% A; 2–10 min, a linear gradient of A from 5 to 15%; 10–20 min, a linear gradient of A from 15 to 25%; 20–25 min, a linear gradient of A from 25 to 30%; and 25–35 min, 5% A.

Statistical Analysis

All results of three independent experiments were averaged and presented as mean ± standard deviation (SD). Statistical differences among the means of two or more groups ($p < 0.05$) were determined using a one-way analysis of variance (ANOVA) followed by Duncan's multiple range test (SAS Institute Inc., Cary, NC, USA). The number of cells was monitored by measuring OD_{600} values, and the growth curve was fitted with a sigmoidal model. Quantification of the released compounds was performed according to the standard curve calibrated using each authentic compound.

Sequence Data Analysis

DNA sequences of the genes in *P. chlororaphis* were retrieved from the *Pseudomonas* Genome Database (http://www.pseudomonas.com/). Sequence homology searching was conducted using the NCBI nucleotide BLAST server. The amino acid sequences of 3-hydroxybenzoate 6-hydroxylases and salicylate hydroxylases from other strains were obtained from GenBank. The phylogenetic tree was constructed by MEGA 7.0 using the Neighbor-Joining method.

RESULTS

Tolerance of *P. chlororaphis* to 3-HBA and Gentisate

It is well-recognized that the accumulation of some metabolites, especially phenolic compounds, is highly toxic to cells (Adeboye et al., 2014). Therefore, to determine whether the excessive accumulation of 3-HBA and GA could affect cell growth, the tolerability of P3 to 3-HBA and GA were evaluated. The tolerance experiment was carried out in shake flasks to simulate the synthesis of secondary metabolites during fermentation. In consistent with the synthetic process, at the early logarithmic growth phase, different concentrations of 3-HBA and GA were supplemented into the medium at a final concentration of 0.5 to 4 g/L. After that, the cells were continuously cultured, and cell growth was monitored until the stationary phase. The results demonstrated that 3-HBA with a concentration of <2 g/L in medium showed no significant effect on cells' growth. When the concentration of 3-HBA reached 3 g/L, the growth of P3 was significantly inhibited, and cell growth was entirely blocked at a concentration of 4 g/L (**Figure 2A**). Although 3-HBA shows quite a toxicity to *Pseudomonas*, tolerance will not be a crucial factor in current research. When different concentration (0.5 to 4 g/L) of GA were supplemented to the medium, no significant affect was detected on the growth of *P. chlororaphis* P3 (**Figure 2B**). We can conclude that *P. chlororaphis* P3 is a good candidate for GA synthesis from 3-HBA.

Pathway Construction for the Synthesis of 3-HBA

To synthesize 3-HBA from glycerol in *P. chlororaphis*, efforts were first intensified on the upstream of shikimate pathway for accumulating chorismate. Shikimate pathway begins with the aldol condensation of metabolic intermediates (i.e., PEP and E4P) involved in the central carbon metabolism. PEP is a key central metabolite that mainly responsible for the synthesis of pyruvate catalyzed by pyruvate kinase. Therefore, weakening the conversion of PEP to pyruvate may increase the availability of PEP. As reported in *E. coli*, *pykA*, and *pykF* encoding pyruvate kinases, they are responsible for converting PEP to pyruvate in *Pseudomonas* (Meza et al., 2012). Once the two *pyk* genes were deleted, the flux of pyruvate to acetyl-CoA would be impeded, thus interfering with the normal growth of cells. Therefore, *pykA* deletion was carried out based on previous research (Wang et al., 2018c).

Considering that phenazine is the main competitive secondary metabolite in the shikimate pathway, the synthesis of PCN should be blocked to ensure chorismate's maximum availability for other pathways. The formation of 2-amino-4-deoxychorismate (ADIC) is the first step of phenazine synthesis (Li et al., 2011), and thus *phzE* deletion was carried out in this study. 3-HBA is synthesized from chorismate via a reaction catalyzed by chorismatase/3-hydroxybenzoate synthase. It has been reported that *hyg5*, a 3-HBA synthase gene originated from *Streptomyces hygroscopicus* was inserted into plasmids, allowing *E. coli* and *C. glutamicum* to synthesize 3-HBA (13, 14) efficiently. Consequently, *hyg5* was integrated into the genome, *phzAB* locus under the control of native strong promoter P_{phz}, resulting in a derivative of P3-Hb0. As expected, P3-Hb0 lost the ability to synthesize PCN; however, 3-HBA was not accumulated in fermentation. When using 3-HBA as a sole carbon source to culture *P. chlororaphis* P3, a little colony growth was observed, indicating that 3-HBA can be degraded in *P. chlororaphis*. According to previous reports, there are two major pathways related to the aerobic degradation of 3-HBA: one is the conversion of 3-HBA to 3,4-dihydroxybenzoic acid (protocatechuic acid) catalyzed by 3-hydroxybenzoate 4-hydroxylase, and the product ultimately enters the protocatechuic acid pathway (Michalover et al., 1973); the other is the para-hydroxylation of 3-hydroxybenzoate to produce GA via 3-hydroxybenzoate 6-hydroxylase, and it enters the GA pathway (Groseclose et al., 1973). Since GA pathway is more common in *Pseudomonas*, sequence alignment was conducted using 3-hydroxybenzoate 6-hydroxylase as a template. Thus, Sal, which is annotated as salicylate hydroxylase in the database, has been identified and postulated to catalyze the degradation of 3-HBA, and then *sal* deletion was performed in P3-Hb0, resulting in a derivative of P3-Hb1.

When culturing P3-Hb0 and P3-Hb1 in shake flasks, the products were analyzed by HPLC. As shown in **Figure 3**, a new peak appeared in P3-Hb1 samples similar to the standard, while no chromatographic peak was found in P3-Hb0 samples over the corresponding time

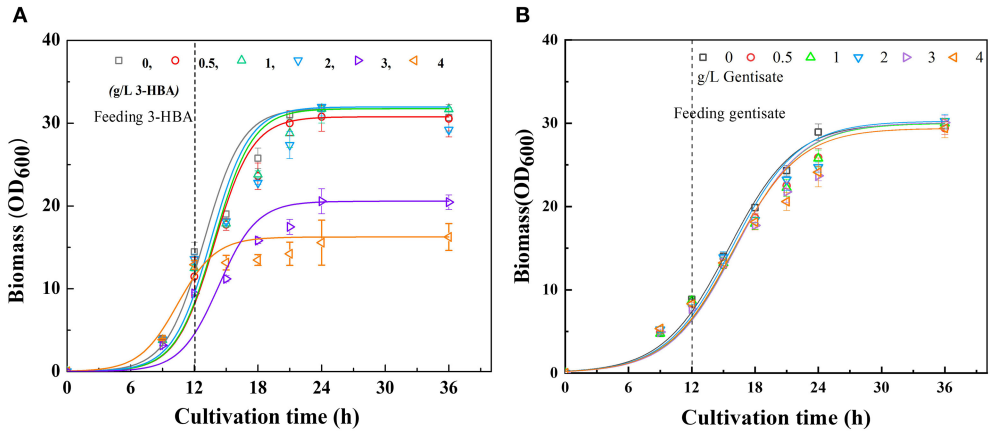


FIGURE 2 | Culture profiles of *P. chlororaphis* in KB medium supplemented with 0–4 g L^{-1} 3-HBA or gentisate at 12 h. **(A)** Time courses of a bacterial cell growth when supplemented with 0–4 g L^{-1} 3-HBA, **(B)** Time courses of a bacterial cell growth when supplemented with 0–4 g L^{-1} gentisate. Data are presented as the mean ± standard deviation of three independent experiments ($n = 3$).

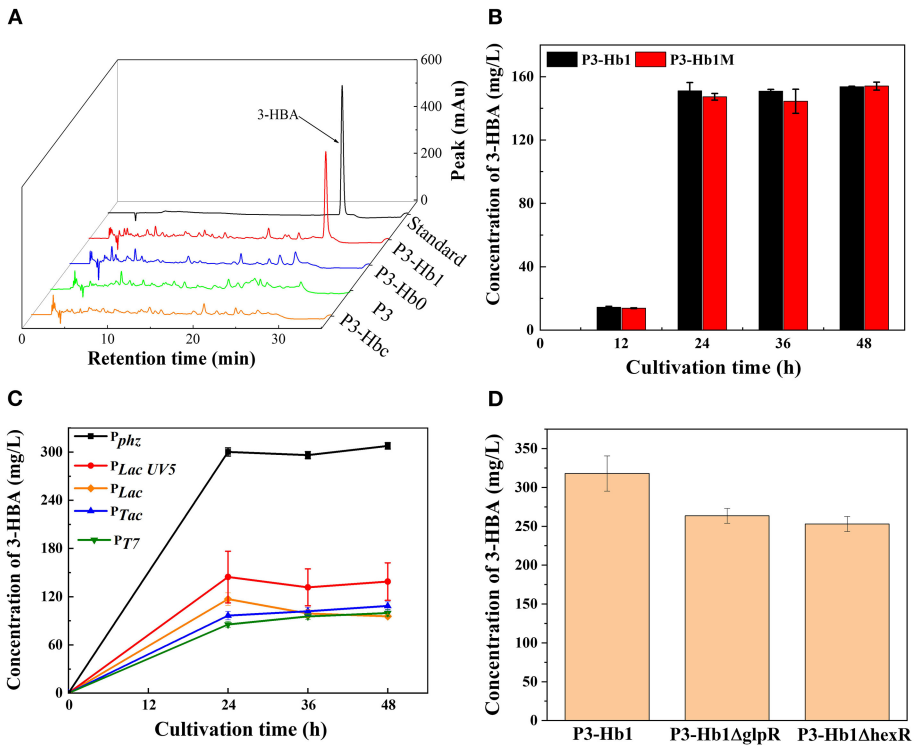


FIGURE 3 | Culture profiles of various 3-HBA-producing transformants. **(A)** HPLC profiles of various transformants; **(B)** The amount of produced 3-HBA in the cultures of P3-Hb1 and P3-Hb1M; **(C)** The amount of produced 3-HBA in the cultures when using different promoter to express *hyg5*; **(D)** The amount of produced 3-HBA in the cultures of P3-Hb1Δ*glpR* and P3-Hb1Δ*hexR*. Data are presented as the mean ± standard deviation of three independent experiments ($n = 3$).

points. To verify the accumulation of 3-HBA, P3-Hb1 sample was further analyzed with ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS). The mass spectrometer was operated in the negative ESI mode, and data acquisition was performed in selected-ion-monitoring (SIM) mode. The peak was observed at ~25.6 min, with the m/z of 137.02, which corresponds to the molecular ion of 3-HBA (**Supplementary Figure 1**). Collectively, the synthetic pathway of 3-HBA in P3 was successfully established, and the amount of 3-HBA produced from P3-Hb1 was 151 mg/L after 48 h of cultivation.

Improvement of a Rate-Limiting Step in 3-HBA Production on Multiple Levels

After the successful construction of 3-HBA synthetic pathway, the next step was to enhance the production of 3-HBA. P3-Hb1 produced 151 mg/L of 3-HBA, which was much less than the PCN quantity produced by P3. The supply of precursor chorismate increases the synthesis of 3-HBA, thus it is considered a rate-limiting step that catalyzes chorismate to 3-HBA. To improve the rate-limiting step, we attempted to optimize the expression of genes involved in 3-HBA production at multiple levels.

Firstly, *cuv10*, a candidate 3-hydroxybenzoate synthase gene from *S. hygroscopicus* was used as a substitute for *hyg5*. P3-Hbc was constructed by inserting *cuv10* into its genome under the control of promoter P_{phz} for a replacement. As shown in **Figure 3A**, the chromatographic peak of 3-HBA was not observed in P3-Hbc sample, indicating that *cuv10*-carrying P3-Hb0 cannot efficiently produce 3-HBA.

Secondly, an additional initiator codon sequence was inserted to the upstream region of ORF, *hyg5* possessed double initiator codon in P3-Hb1m consequently. According to a report that the substitution of the start codon to regulate the expression of some genes is one useful strategy in synthetic biology (Chen et al., 2017a). Based on this, we adopted the similar approach to promote the binding of ribosome and mRNA and ultimately increase translation levels. P3-Hb1m was fermented while P3-Hb1 as a control group and then sampled for HPLC analysis. **Figure 3B** shows the amounts of 3-HBA produced during the cultivation. There was no significant difference in the production levels of 3-HBA between these two strains, and both of them achieved a maximum yield of 151 mg/L after 48 h of cultivation.

Thirdly, *hyg5* was expressed under the control of different promoters. P_{phz} is a native strong promoter located at the upstream of the phenazine gene cluster, which can effectively regulate the transcription levels of the whole gene cluster. Four foreign constitutive promoters (i.e., P_{Lac}, P_{LacUV5}, P_{Tac} and P_{T7}) were cloned from *E. coli*, a native promoter (i.e., P_{phz}) was cloned from *P. chlororaphis*, then linked to *hyg5* and inserted into the genome of P3-Hb1. For construction of P_{T7}-based expression derivative, T7 RNA polymerase was inserted into the *phzE* locus under the control of native promoter. A total of five derivatives (i.e., P3-Hhb1, P3-Hhb2, P3-Hhb3, P3-Hhb4, and P3-Hhb5) were fermented, and the amount of 3-HBA produced in each derivative is shown in **Figure 3C**. It was found that P_{phz}-induced overexpression of *hyg5* could considerably enhance the production of 3-HBA, with a maximum level of 300 mg/L, which nearly doubled compared to P3-Hb1. However, other promoters did not exhibit a positive effect on the improvement of 3-HBA production.

Lastly, global metabolic regulation was concerned. It has been reported that the deletion of *glpR* in *P. putida* could eliminate its growth lag-phase and increase polyhydroxyalkanoates accumulation when cultured on glycerol (Escapa et al., 2013). Besides, the transcriptional factor HexR regulates the central carbohydrate metabolism globally. According to previous findings, pyruvate kinase is regulated by HexR (Leyn et al., 2011). Thus, *glpR* and *hexR* were deleted in P3-Hhb1, individually. Contrary to our expectation, the amount of 3-HBA decreased slightly (**Figure 3D**). We assumed that GlpR displays positive action on central carbohydrate metabolism in *P. chlororaphis*, once deleted, the precusor of PEP is decreased for shikimate pathway, more evidence should be revealed.

Biosynthesis of Gentisate From 3-HBA

As the efficient production of 3-HBA was achieved, we attempted to construct the pathway for GA from 3-HBA. During the 3-HBA synthetic pathway construction, Sal was found to catalyze the degradation of 3-HBA by adding a hydroxyl group to the benzene ring. Therefore, different hydroxybenzoic acid monooxygenases were screened for catalyzing 3-HBA to dihydroxybenzoic acid derivatives. According to a previous report, *p*-hydroxybenzoate hydroxylase encoded by *pobA* from *P. aeruginosa* was mutated into Y385F/T294A PobA (hereinafter referred to as PobAM). PobAM displayed a high catalytic activity toward 3,4-dihydroxybenzoic acid and catalyzed the formation of gallic acid (Chen et al., 2017b). Thus, the hydroxybenzoic acid monooxygenase genes (i.e., *sal*, *pobA* and *pobAM*) were linked to expression vector pET28a(+) and subsequently transferred into BL21(DE3). 3-HBA was added to the cell suspension culture with a final concentration of 500 mg/L, and the concentration changes of 3-HBA in the four groups are presented in **Figure 4A**. After 12 h of incubation, most of 3-HBA was catalyzed by Sal, while the concentration of 3-HBA did not differ significantly between PobA and PobAM groups. The chromatographic peaks of 3-HBA in the four groups at 12 h are shown in **Figure 4B**. Results demonstrated that a new substance appeared when catalyzing 3-HBA by Sal. In comparison with the HPLC profile of various hydroxybenzoic acids, it can be speculated that the new substance is GA. Furthermore, UPLC-MS/MS analysis also supported the speculation that Sal can catalyze the conversion of 3-HBA to GA (**Supplementary Figure 2**).

The degradation pathway of 3-HBA in *Pseudomonas* has also been clarified. As reported earlier, 3-HBA was converted to GA, and then maleylpyruvate was formed by GA 1,2-dioxygenase (MhbD) mediated ring-cleavage reaction. The final products of the above reactions are pyruvate and fumarate, which ultimately enter the TCA cycle (Lin et al., 2010; Wang et al., 2020; **Figure 1**). The genes encoding these enzymes are all located within a single gene cluster, involving a 3-HBA transporter (Xu et al., 2012; Wang et al., 2020). Thus, a pathway for GA accumulation was constructed by deleting *mhbD1* in P3-Hb0. Notably, the chromatographic peak of GA was observed, implying that GA was accumulated successfully in P3-GA1. Besides, the concentration of GA in fermentation broth reached a maximum of 105 mg/L after cultivation for 24 h and then decreased to 47 mg/L at 36 h.

Improvement of Gentisate Production

As mentioned above, GA has accumulated in *P. chlororaphis* P3 temporarily, the concentrations of GA were decreased during fermentation, suggesting another degradation pathway for GA in *P. chlororaphis*. After database searching, a gene was identified

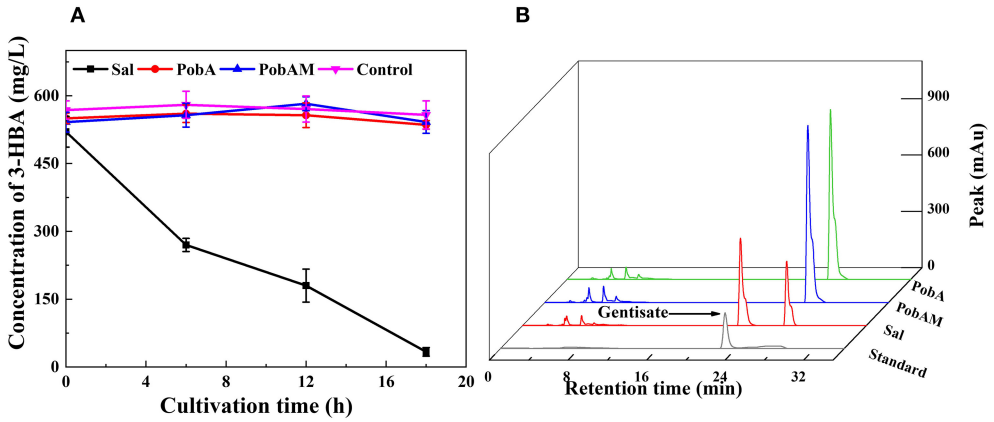


FIGURE 4 | 3-HBA consuming profiles by different monooxygenase. **(A)** Time courses of 3-HBA consuming, **(B)** HPLC profiles of various transformants for 3-HBA and gentisata.

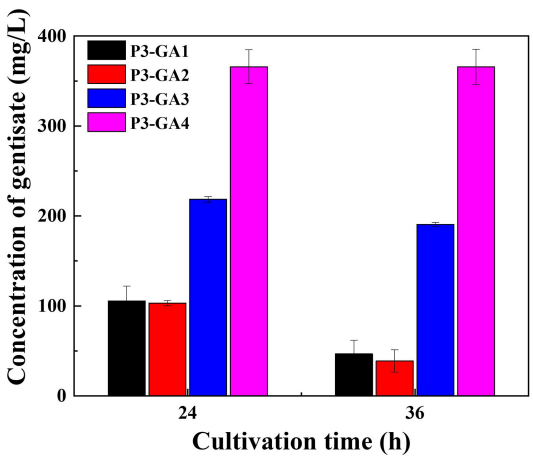


FIGURE 5 | Improvement of rate limiting steps for enhancement of gentisate production.

and named as *mhbD2*, due to its position on the antisense strand. It has been recorded to encode a GA 1,2-dioxygenase, according to the *Pseudomonas* Genome Database. When deleted *mhbD2*, the degradation of GA was similar to P3-GA1 (**Figure 5**).

Increasing the availability of precursors to enhance the production of GA may be an option to offset the degradation. Overexpression of *hyg5* was conducted in P3-Ga0 under the control of P_{phz} promoter, resulting in P3-GA3. To our expectation, the amount of GA was doubled to 218 mg/L at 24 h of fermentation. It was noteworthy that the yield of GA at 36 h was 190 mg/L, and the degradation of GA from 24 to 36 h was retarded with its improved production rates. Following *hyg5* overexpression, *hmgA* that encodes homogentisate 1,2-dioxygenase was inactived, since the structure of homogentisate is relatively similar to GA. After deleting *hmgA* in P3-GA3, the amount of GA reached 365 mg/L at 24 h of fermentation, which was 67% higher than P3-GA3. There was no significant difference in the amount of GA between 24 and 36 h, indicating that GA is no longer degraded during fermentation process (**Figure 5**). Therefore, a stable and effective GA biosynthetic pathway was successfully established in the present study, and 365 mg/L GA was produced in P3-GA4.

Biosynthesis of Gentisate From 4-HBA

4-HBA has recently emerged as a versatile intermediate for several value-added bioproducts, such as muconic acid, arbutin, gastrodin, xiamenmycin, and vanillyl alcohol using 4-HBA as the starting feedstock (Wang et al., 2018a,c). A novel reaction in the conversion of 4-HBA to GA was reported (Zhao et al., 2018), in which three genes (*phgABC*) catalyze the transformation of 4-HBA to GA via a route involving CoA thioester formation, hydroxylation concomitant with a 1, 2-shift of the acetyl CoA moiety and thioester hydrolysis (**Figure 1**). Using our earlier screened XanB2 for 4-HBA synthesis, we integrated P_{phz}-*xanB2* on *pykA* locus, with *phzE*, *pobA*, *mhbD1* and *hmgA* deleted. Then, *phgA-phgB-phgC* were integrated on *phzAB* locus under the control of native strong promoter P_{phz}, yielding one GA derivative GA-4HBA. When fermented in KB medium, unfortunately, no new peak appeared and no siginificantly 4-HBA reduced.

DISCUSSION

GA is an important chemical with high industrial values, together with other hydroxybenzoic acids, including salicylic acid, 4-HBA, 3-HBA, and so on. There have been many reports about hydroxybenzoic acid production and their derivatives in various microbial systems (Wang et al., 2018a). Although antibiotic compounds are concerned as 'emerging contaminants' (Keen and Patrick, 2013), the biosynthesis of GA independent of inducers and antibiotics has not been achieved previously. In this work, we engineered a chromosome-integrated synthetic pathway for GA production from 3-HBA in *Pseudomonas*.

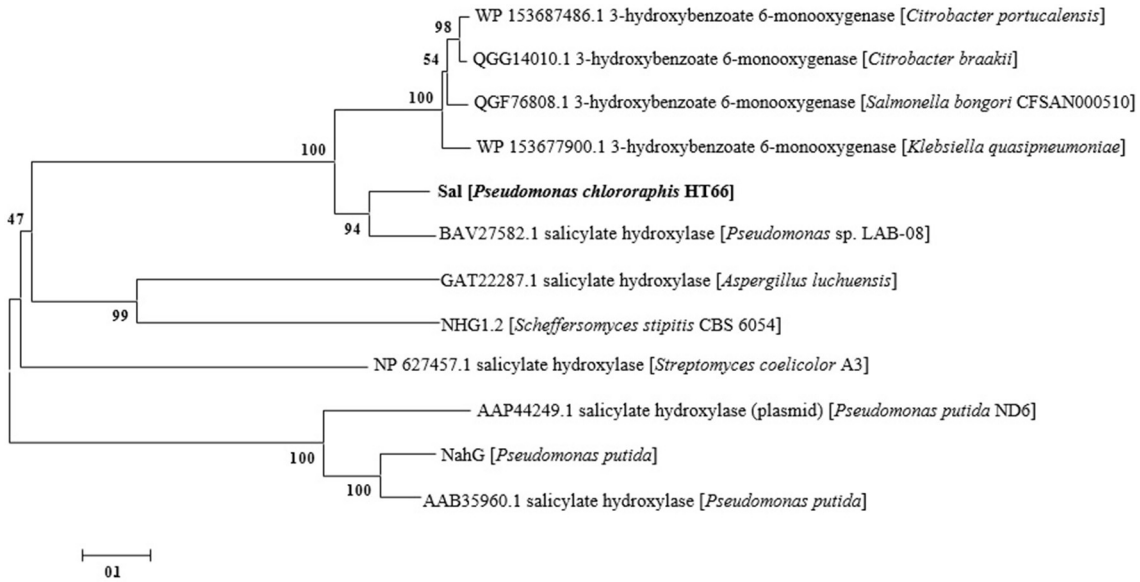


FIGURE 6 | Phylogenetic tree of of 3-hydroxybenzoate 6-hydroxylase, salicylate hydroxylase and Sal. The ruler at the bottom of figure indicates the horizontal distance equal to 10% sequence divergence.

3-HBA usually serves as an important platform chemical in microorganisms, in which its synthetic pathway can be found and reconstructed. We focused on the enhanced shikimate pathway in *P. chlororaphis* P3, the leading pathway for aromatic compound synthesis. Apart from introducing exogenous 3-hydroxybenzoate synthase to catalyzes 3-HBA from chorismate, another approach was to prevent the degradation of the products and their potential precursors in *Pseudomonas*. To enhance the synthesis of 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP), *pykA* deletion was conducted in P3. Afterwards, *phzE* that catalyzes chorismate to phenazines and *sal* that degrades 3-HBA were knocked out, individually, resulting in a chassis strain for 3-HBA synthesis from glycerol (**Figure 1**).

To enhance the production of 3-HBA and GA, multiple strategies were employed. Several reports have described that the feedback inhibition in the shikimate pathway may impede the production of target products (Kikuchi et al., 1997; Juminaga et al., 2012). However, on the basis of our previous findings, we speculated that such bottleneck mainly occurs during the conversion of chorismate to 3-HBA in *P. chlororaphis* (Wang et al., 2018b). Few 3-hydroxybenzoate synthase genes were reported, and the enzymes with higher activity than Hyg5 are not available for a replacement. According to reports, Cuv10 exhibits maximum activity at 26°C (pH 6.5), the culture conditions of *P. chlororaphis* may not meet the enzymatic properties of Cuv 10. Meanwhile, *cuv10* is a part of a native polyketide synthase gene, and it is probably non-functional as chorismatase in *Pseudomonas* (Jiang et al., 2013).

Moreover, the enhancement of the start codon did not upregulate the activity of *hyg5* The overexpression of *hyg5* under the control of native strong promoter P_{phz} significantly upregulated its expression levels, which in turn led to a double increase (300 mg/L) in 3-HBA production (**Figure 3C**). After that, *glpR* and *hexR* (encoding transcriptional regulator) were deleted to optimize metabolic regulation. The results demonstrated that the metabolic flux to shikimate pathway has been improved in *P. chlororaphis* P3, and the deletion of *glpR* and *hexR* might cause an imbalance in primary metabolism and energy flux, resulting in a negative effect on the production of 3-HBA (**Figure 3D**). Thus, the highest titer of 3-HBA in P3-Hhb1 was 300 mg/L. The reaction of Hyg5 undergoes an intramolecular arene oxide mechanism, starts with surmounting a high energy barrier (Dong and Liu, 2017), thereby resulting in its low level of activity.

Based on the whole-cell catalysis experiment of 3-HBA, we identified that endogenous Sal catalyzed the conversion of 3-HBA to GA with high efficiency, which displayed the characteristics of 3-hydroxybenzoate 6-hydroxylases. As shown in **Figure 6**, a phylogenetic tree was constructed with representative 3-hydroxybenzoate 6-hydroxylases and salicylate hydroxylases from other strains (Chen et al., 2018). As shown, Sal reveals as one of these two groups, suggesting that Sal may display unselective substrate adaptability and dual catalytic activity (Fang and Zhou, 2014). Both salicylate hydroxylase and 3-hydroxybenzoate 6-hydroxylases belong to the same family of flavin-dependent monooxygenases (Yang et al., 2011; Huijbers et al., 2014), with high sequence homology between them.

Upon assessing the production of GA, no 3-HBA was found in the samples, indicating that endogenous Sal catalyzed 3-HBA effectively. As mentioned above, the maximum amount of GA was achieved simultaneously (24 h of cultivation) as 3-HBA. However, the net conversion rate of 3-HBA to GA was only 62.2% (the mole ratio) at 24 h without residual 3-HBA, which was much lower than the theoretical conversion rate, indicating

that a third of GA was degraded. To eliminate the degradation caused by spontaneous oxidation, GA was added to KB medium and incubated at 28°C for two days. No significant change was detected in the culture, confirming that GA is relatively stable in the culture (**Supplementary Figure 3**).

GA 1,2-dioxygenase detected in other *Pseudomonas* shared a high homology level with MhbD1. There are few reports of other GA degradation genes. In this study, gene *mhbD2* was identified. Sequence alignment and analysis revealed that *mhbD2* was not associated with the synthesis or degradation of GA, and it probably encoded a member of fumarylacetoacetate hydrolase family protein. Thus, we attempted to enhance the carbon flux from chorismate to GA by overexpressing *hyg5*. Consequently, the maximum amount of GA produced was doubled, and the degradation was partially offset. Interestingly, when *hmgA* was deleted, the production of GA was improved entirely. It has not been reported that homogentisate 1,2-dioxygenase catalyzes GA previously, but it does indeed exist in *Pseudomonas*. Our results suggest that microbial catabolism is not only composed of one single pathway and contains an interrelated metabolic network. The unselective substrate adaptability of Sal and HmgA reflects the versatile metabolism of *Pseudomonas*, indicating that our platform strain P3 has a great potential to synthesize valuable chemicals via metabolic engineering. Unexpected, when *phgABC* were expressed for synthesis GA from 4-HBA based on NIH shift, no significant GA was accumulated. For the lower specific activity of PhgC against 4-HBA (Zhao et al., 2018), the higher concentration of 4-HBA may inhibit the expression of PhgC.

Various pathways could be linked to the shikimate pathway. At present, we have synthesized versatile platform compound GA and 3-HBA from simple carbon sources successfully, and many other GA-based value-added derivatives could be synthesized in the near future. Pathways can be designed to produce gallic acid via hydroxylation of 3-HBA and protocatechuic (Chen et al., 2017b). According to a report, industrially valuable maleate production was attained by extending chorismate and GA pathways (Noda et al., 2017). Besides, the hydrolysis of chorismate to 3-HBA involves the synthesis of macrocyclic polyketides, which has attracted great interests in treating metastatic and inflammatory diseases (Andexer et al., 2011). Therefore, connecting the shikimate pathway and other pathways with hydroxybenzoate acid as a node may become a powerful strategy for producing valuable bioproducts, including new to nature products.

In conclusion, chromosome-integrated synthetic pathway for GA from 3-HBA were constructed in *Pseudomonas* for the first time based on the enhanced shikimate pathway in P3 strain. The biosynthetic route of GA was constructed by connecting the endogenous degradation pathway and 3-HBA synthetic pathway. This study provides new insights into the possibility of using *Pseudomonas* to synthesize valuable compounds from renewable feedstocks, with a more environmentally responsible, eco-friendly strategy.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

SW and XZ conceived and designed the experiments. SW performed experiments, analyzed the experimental data, and drafted the manuscript. CF, KL, and JC assisted in experimental work and manuscript writing. HH and WW contributed reagents & materials. XZ revised the manuscript. All authors contributed to the final paper.

FUNDING

This work was financed by the National Key Science Research Projects (Grant No. 2019YFA09004302) and the National Natural Science Foundation of China (Grant No. 31670033).

ACKNOWLEDGMENTS

We would like to express our gratitude to the Instrumental Analysis Center of Shanghai Jiao Tong University for their skillful technical assistance in UPLC-Q/TOF MS analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.622226/full#supplementary-material

REFERENCES

Adeboye, P. T., Bettiga, M., and Olsson, L. (2014). The chemical nature of phenolic compounds determines their toxicity and induces distinct physiological responses in Saccharomyces cerevisiae in lignocellulose hydrolysates. *AMB Express* 4:46. doi: 10.1186/s13568-014-0046-7

Andexer, J. N., Kendrew, S. G., Nur-E-Alam, M., Lazos, O., Foster, T. A., Zimmermann, A. S., et al. (2011). Biosynthesis of the immunosuppressants FK506, FK520, and rapamycin involves a previously undescribed family of enzymes acting on chorismate. *Proc. Natl. Acad. Sci. U.S.A.*108, 4776–4781. doi: 10.1073/pnas.1015773108

Averesch, N. J. H., and Krömer, J. O. (2018). Metabolic engineering of the shikimate pathway for production of aromatics and derived compounds-present and future strain construction strategies. *Front. Bioeng. Biotechnol.* 6:32. doi: 10.3389/fbioe.2018.00032

Bai, Y. F., Yin, H., Bi, H. P., Zhuang, Y. B., Liu, T., and Ma, Y. H. (2016). *De novo* biosynthesis of Gastrodin in *Escherichia coli*. *Metab. Eng.* 35, 138–147. doi: 10.1016/j.ymben.2016.01.002

Belda, E., Van Heck, R. G. A., Lopez-Sanchez, M. J., Cruveiller, S., Barbe, V., Fraser, C., et al. (2016). The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. *Environ. Microbiol.* 18, 3403–3424. doi: 10.1111/1462-2920.13230

Cao, M., Gao, M., Suástegui, M., Mei, Y., and Shao, Z. (2020). 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. doi: 10.1016/j.ymben.2019.08.008

Chen, X., Tang, H. Z., Liu, Y. D., Xu, P., Xue, Y., Lin, K. F., et al. (2018). Purification and initial characterization of 3-hydroxybenzoate 6-hydroxylase from a halophilic martelella strain AD-3. *Front. Microbiol.* 9:1335. doi: 10.3389/fmicb.2018.01335

Chen, Z., Huang, J. H., Wu, Y., Wu, W. J., Zhang, Y., and Liu, D. H. (2017a). Metabolic engineering of *Corynebacterium glutamicum* for the production of 3-hydroxypropionic acid from glucose and xylose. *Metab. Eng.* 39, 151–158. doi: 10.1016/j.ymben.2016.11.009

Chen, Z. Y., Shen, X. L., Wang, J., Wang, J., Yuan, Q. P., and Yan, Y. J. (2017b). Rational engineering of p-hydroxybenzoate hydroxylase to enable efficient gallic acid synthesis via a novel artificial biosynthetic pathway. *Biotechnol. Bioeng.* 114, 2571–2580. doi: 10.1002/bit.26364

Chen, Z. Y., Shen, X. L., Wang, J., Wang, J., Zhang, R. H., Rey, J. F., et al. (2017c). Establishing an artificial pathway for *de novo* biosynthesis of vanillyl alcohol in *Escherichia coli*. *ACS Synth. Biol.* 6, 1784–1792. doi: 10.1021/acssynbio.7b00129

Choi, S., Song, C. W., Shin, J. H., and Lee, S. Y. (2015). Biorefineries for the production of top building block chemicals and their derivatives. *Metab. Eng.* 28, 223–239. doi: 10.1016/j.ymben.2014.12.007

Dong, L. H., and Liu, Y. J. (2017). Comparative studies of the catalytic mechanisms of two chorismatases: CH-fkbo and CH-Hyg5. *Proteins* 85, 1146–1158. doi: 10.1002/prot.25279

Escapa, I. F., Del Cerro, C., Garcia, J. L., and Prieto, M. A. (2013). The role of GlpR repressor in *Pseudomonas putida* KT2440 growth and PHA production from glycerol. *Environ. Microbiol.* 15, 93–110. doi: 10.1111/j.1462-2920.2012.02790.x

Fang, T., and Zhou, N.-Y. (2014). Purification and characterization of salicylate 5-hydroxylase, a three-component monooxygenase from *Ralstonia* sp. strain U2. *Appl. Microbiol. Biotechnol.* 98, 671–679. doi: 10.1007/s00253-013-4914-x

Groseclose, E. E., Ribbons, D. W., and Hughes, H. (1973). 3-hydroxybenzoate 6-hydroxylase from *Pseudomonas aeruginosa*. *Biochem. Biophys. Res. Commun.* 55, 897–903. doi: 10.1016/0006-291X(73)91228-X

Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D., and Van Berkel, W. J. H. (2014). Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* 544, 2–17. doi: 10.1016/j.abb.2013.12.005

Jiang, Y. H., Wang, H. X., Lu, C. H., Ding, Y. J., Li, Y. Y., and Shen, Y. M. (2013). Identification and characterization of the cuevaene A biosynthetic gene cluster in *streptomyces* sp. LZ35. *Chembiochem* 14, 1468–1475. doi: 10.1002/cbic.201300316

Jin, X. J., Peng, H. S., Hu, H. B., Huang, X. Q., Wang, W., and Zhang, X. H. (2016). iTRAQ-based quantitative proteomic analysis reveals potential factors associated with the enhancement of phenazine-1-carboxamide production in *Pseudomonas chlororaphis* P3. *Sci. Rep.* 6:27393. doi: 10.1038/srep27393

Juminaga, D., Baidoo, E. E. K., Redding-Johanson, A. M., Batth, T. S., Burd, H., Mukhopadhyay, A., et al. (2012). Modular engineering of L-tyrosine production in *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 89–98. doi: 10.1128/AEM.06017-11

Kallscheuer, N., and Marienhagen, J. (2018). *Corynebacterium glutamicum* as platform for the production of hydroxybenzoic acids. *Microb. Cell Fact.* 17:70. doi: 10.1186/s12934-018-0923-x

Keen, P. L., and Patrick, D. M. (2013). Tracking change: a look at the ecological footprint of antibiotics and antimicrobial resistance. *Antibiotics* 2, 191–205. doi: 10.3390/antibiotics2020191

Kikuchi, Y., Tsujimoto, K., and Kurahashi, O. (1997). Mutational analysis of the feedback sites of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase of *Escherichia coli*. *Appl. Environ. Microbiol.* 63, 761–762. doi: 10.1128/AEM.63.2.761-762.1997

Leyn, S. A., Li, X. Q., Zheng, Q. X., Novichkov, P. S., Reed, S., Romine, M. F., et al. (2011). Control of proteobacterial central carbon metabolism by the HexR transcriptional regulator: a case study in *Shewanella oneidensis*. *J. Biol. Chem.* 286, 35782–35794. doi: 10.1074/jbc.M111.267963

Li, Q. A., Mavrodi, D. V., Thomashow, L. S., Roessle, M., and Blankenfeldt, W. (2011). Ligand binding induces an ammonia channel in 2-amino-2-desoxyisochorismate (ADIC) synthase PhzE. *J. Biol. Chem.* 286, 18213–18221. doi: 10.1074/jbc.M110.183418

Liao, J. C., Mi, L., Pontrelli, S., and Luo, S. S. (2016). Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* 14, 288–304. doi: 10.1038/nrmicro.2016.32

Lin, L. X., Liu, H., and Zhou, N. Y. (2010). MhbR, a LysR-type regulator involved in 3-hydroxybenzoate catabolism via gentisate in *Klebsiella pneumoniae* M5a1. *Microbiol. Res.* 165, 66–74. doi: 10.1016/j.micres.2008.08.001

Lin, Y. H., Sun, X. X., Yuan, Q. P., and Yan, Y. J. (2014). Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in *Escherichia coli*. *Metab. Eng.* 23, 62–69. doi: 10.1016/j.ymben.2014.02.009

Meza, E., Becker, J., Bolivar, F., Gosset, G., and Wittmann, C. (2012). Consequences of phosphoenolpyruvate: sugar phosphotranferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in *Escherichia coli*. *Microb. Cell Fact.* 11:127. doi: 10.1186/1475-2859-11-127

Michalover, J. L., Ribbons, D. W., and Hughes, H. (1973). 3-Hydroxybenzoate 4-hydroxylase from *Pseudomonas testosteroni*. *Biochem. Biophys. Res. Commun.* 55, 888–896. doi: 10.1016/0006-291X(73)91227-8

Noda, S., Shirai, T., Mori, Y., Oyama, S., and Kondo, A. (2017). Engineering a synthetic pathway for maleate in *Escherichia coli*. *Nat. Commun.* 8:1153. doi: 10.1038/s41467-017-01233-9

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

Shen, Y., Sun, F., Zhang, L., Cheng, Y., Zhu, H., Wang, S. P., et al. (2020). Biosynthesis of depsipeptides with a 3-hydroxybenzoate moiety and selective anticancer activities involves a chorismatase. *J. Biol. Chem.* 295, 5509–5518. doi: 10.1074/jbc.RA119.010922

Wang, S., Cui, J., and Bilal, M. (2020). *Pseudomonas spp*. as cell factories (MCFs) for value-added products: from rational design to industrial applications. *Crit. Rev. Biotechnol.* 40, 1232–1249. doi: 10.1080/07388551.2020.1809990

Wang, S. W., Bilal, M., Hu, H. B., Wang, W., and Zhang, X. H. (2018a). 4-Hydroxybenzoic acid-a versatile platform intermediate for value-added compounds. *Appl. Microbiol. Biotechnol.* 102, 3561–3571. doi: 10.1007/s00253-018-8815-x

Wang, S. W., Bilal, M., Zong, Y. N., Hu, H. B., Wang, W., and Zhang, X. H. (2018b). Development of a plasmid-free biosynthetic pathway for enhanced muconic acid production in *Pseudomonas chlororaphis* HT66. *ACS Synth. Biol.* 7, 1131–1142. doi: 10.1021/acssynbio.8b00047

Wang, S. W., Fu, C., Bilal, M., Hu, H. B., Wang, W., and Zhang, X. H. (2018c). Enhanced biosynthesis of arbutin by engineering shikimate pathway in *Pseudomonas chlororaphis* P3. *Microb. Cell Fact.* 17:174. doi: 10.1186/s12934-018-1022-8

Xu, Y., Gao, X. L., Wang, S. H., Liu, H., Williams, P. A., and Zhou, N. Y. (2012). MhbT is a specific transporter for 3-hydroxybenzoate uptake by gram-negative bacteria. *Appl. Environ. Microbiol.* 78, 6113–6120. doi: 10.1128/AEM.01511-12

Yang, Y. F., Zhang, J. J., Wang, S. H., and Zhou, N. Y. (2011). Purification and characterization of the ncgl2923 -encoded 3-hydroxybenzoate 6-hydroxylase from *Corynebacterium glutamicum*. *J. Basic Microbiol.* 50, 599–604. doi: 10.1002/jobm.201000053

Zhao, H., Xu, Y., Lin, S., Spain, J. C., and Zhou, N. Y. (2018). The molecular basis for the intramolecular migration (NIH shift) of the carboxyl group during para-hydroxybenzoate catabolism. *Mol. Microbiol.* 110, 411–424. doi: 10.1111/mmi.14094

Zhou, Y. Y., Li, Z. H., Wang, X. N., and Zhang, H. R. (2019). Establishing microbial co-cultures for 3-hydroxybenzoic acid biosynthesis on glycerol. *Eng. Life Sci.* 19, 389–395. doi: 10.1002/elsc.201800195

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 26 January 2021
doi: 10.3389/fbioe.2021.613322

Opportunities and Challenges for Microbial Synthesis of Fatty Acid-Derived Chemicals (FACs)

*Yilan Liu[1], Mauricio Garcia Benitez[1], Jinjin Chen[1], Emma Harrison[1], Anna N. Khusnutdinova[1] and Radhakrishnan Mahadevan[1,2]**

[1] Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, [2] Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Global warming and uneven distribution of fossil fuels worldwide concerns have spurred the development of alternative, renewable, sustainable, and environmentally friendly resources. From an engineering perspective, biosynthesis of fatty acid-derived chemicals (FACs) is an attractive and promising solution to produce chemicals from abundant renewable feedstocks and carbon dioxide in microbial chassis. However, several factors limit the viability of this process. This review first summarizes the types of FACs and their widely applications. Next, we take a deep look into the microbial platform to produce FACs, give an outlook for the platform development. Then we discuss the bottlenecks in metabolic pathways and supply possible solutions correspondingly. Finally, we highlight the most recent advances in the fast-growing model-based strain design for FACs biosynthesis.

Keywords: fatty acid-derived chemicals, microbial chassis, systems engineering, model-assisted design, review

OPEN ACCESS

Edited by:
Dipesh Dhakal,
University of Florida, United States

Reviewed by:
Zongbao K. Zhao,
Chinese Academy of Sciences, China
Dae-Hee Lee,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), South Korea

***Correspondence:**
Radhakrishnan Mahadevan
krishna.mahadevan@utoronto.ca

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *02 October 2020*
Accepted: *04 January 2021*
Published: *26 January 2021*

Citation:
Liu Y, Benitez MG, Chen J, Harrison E, Khusnutdinova AN and Mahadevan R (2021) Opportunities and Challenges for Microbial Synthesis of Fatty Acid-Derived Chemicals (FACs). Front. Bioeng. Biotechnol. 9:613322. doi: 10.3389/fbioe.2021.613322

INTRODUCTION

Increasing consumption of petroleum-derived products leads to increasing atmospheric carbon dioxide (CO_2) levels and global warming (Sperry et al., 2019). Furthermore, the uneven distribution and unsustainability of fossil resources have motivated engineers to seek alternative sustainable solutions (Raslavičius et al., 2014; Chen et al., 2020). Compared with the traditional strategies to convert plant oils and animal fats into biodiesel, microbial synthesis of fuels, and chemicals presents several advantages. Firstly, feedstocks can be shifted from edible plant oils and animal fats to non-edible biomass feedstocks, especially CO_2. Secondly, due to the flexibility of pathways in microbial chassis, a large diversity of bioproducts can be produced in microbial cell factories. Among these bioproducts, fatty acid-derived chemicals (FACs) have attracted significant attention, because fatty acids (FAs) are essential metabolites in all organisms. FAs and their biosynthetic/catabolic intermediates can be used as precursors for a large diversity of FACs, which have an unprecedented wide application range: biofuels, pharmaceuticals, feed additives, and others. Thirdly, bioproducts are green alternatives to petroleum-based fuels, given the capacity of net-zero greenhouse gas emissions. Microbial chassis must be extensively designed and engineered to produce FACs at high titer, rate and yield from various substrates. Recent successes in model-based strain design have speed-up the Design-Build-Test-Learn (DBTL) cycle in metabolic engineering (Carbonell et al., 2018; Hamedirad et al., 2019; Opgenorth et al., 2019). Although FACs biosynthesis has been reviewed from different angles (Marella et al., 2018; Liu and Li, 2020), the purpose of this review

is to update the most recent advances in this fast-developing field, with an emphasis on possible synthetic microbial chassis and computational modeling for biosynthesis of FACs.

TYPES AND APPLICATIONS OF FATTY ACID-DERIVED CHEMICALS

With accelerating concerns over climate change and the environmental impact of conventional production methods, interest in the renewable microbial production of chemicals have grown (Liu and Nielsen, 2019; Cho et al., 2020; Li M. et al., 2020; Sgobba and Wendisch, 2020; Wu et al., 2020). Among these chemicals, FACs are of particular interest due to their various applications in biofuels, detergents, medicines, industrial lubricants, bioplastics, emulsifiers, food and feed additives, and others (**Supplementary Table 1**; Richardson and Mcallister, 1945; Geller and Goodrum, 2004; Bellou et al., 2016; Jiang W. et al., 2018; Li G. et al., 2020; Verma et al., 2020; Zerhusen et al., 2020). Different end groups and lengths of FACs lead to different physical and chemical properties, which in turn lead to different practical applications (**Figure 1**). In general, FACs can be mainly classified into free fatty acids (FFA), fatty alcohols, alka(e)nes, and fatty acid esters (FAEs) (Steen et al., 2010). Most naturally occurring FACs have an unbranched chain within the range of C3 to C28. Based on the chain length, they are generally classified into short-chain ($\leq$ 6), medium-chain (7–12), long-chain (13–20), and very-long-chain ($>$ 20) (Schönfeld and Wojtczak, 2016). However, the definitions can vary from one study to another (Rodriguez-Moya and Gonzalez, 2015).

Fatty acids, one of the most studied FACs, are n-alkyl carboxylic acids with an aliphatic chain. Short-chain fatty acids play an important role in human health. For example, butyric acid can be used in food and pharmaceutical industries (Jiang L. et al., 2018). Most free FFAs are converted into biofuels, or consumer products (Leber et al., 2015; Marella et al., 2018; Sathesh-Prabu et al., 2019). Fatty alcohols have hydroxyl groups attached to the aliphatic chain. Short-chain alcohols, such as 1-propanol and 1-butanol, can be used as gasoline or fuel additives (Guo et al., 2019; Zhou et al., 2019). Alcohols with a chain length of C8–C10 are important materials to produce detergents, lubricants, cosmetics, pharmaceuticals, and plastics. Those in C12–C14 range are used as lubricant additives, and those in C16–C18 range are used for drug delivery and defoamers (Dong and Mumper, 2006; Zheng et al., 2012). Among these, C11–C14 alcohols, the key intermediates for surfactants production, represent 55% of the market share (Fillet and Adrio, 2016). Alkanes are saturated hydrocarbons with the general chemical formula C_nH_{2n+2}, while alkenes are unsaturated hydrocarbons containing at least one C–C double bond. Alka(e)nes are an important class of FACs because of their high similarity to petroleum-derived fuels. Depending on the chain length, alka(e)nes have different applications, including drop-in fuels in gasoline (C3–C9), jet fuel (C8–C16), diesel (C10–C18), and lubricants (C16–C30) (Kang et al., 2017). In addition to the FACs mentioned above, there are other important fatty acid-based chemicals, such as fatty acid alkyl esters (FAAEs) and branched FACs that are used for certain applications due to their specific properties (Röttig et al., 2010; Ngo et al., 2013; Gupta et al., 2015; Teo et al., 2015; Bentley et al., 2016; Jiang et al., 2017; Shrestha and Yamamoto, 2018; Singh and Choudhury, 2018). For examples, branched fatty alcohol 4-methyl-pentanol is a common brake fluids (Shrestha and Yamamoto, 2018; **Supplementary Table 1**).

MICROBIAL CHASSIS FOR BIOSYNTHESIS OF FACS

Theoretically speaking, any microorganism can be used as a microbial chassis for biosynthesis of FACs, because fatty acid metabolic pathways exist in all living cells. Currently, most researches are devoted to model organisms such as *Escherichia coli* and *Saccharomyces cerevisiae* (Liu et al., 2016; Hu et al., 2019; Kim and Park, 2019; Yang et al., 2020). In our opinion, there are some other promising microorganisms, which have not been explored. In this review, microorganisms are classified into four groups including chemoautotroph, photoautotroph, heterotrophic prokaryotes, and heterotrophic eukaryotes. Their advantages, disadvantages, and the promising synthetic community strategy for microbial production of FACs will be discussed in detail (**Figure 2**).

Chemoautotrophs are organisms that can synthesize their own organic molecules through the fixation of carbon dioxide. Energy required for this process comes from the oxidation of inorganic molecules such as iron, sulfur, or magnesium (Thakur et al., 2018). Though research efforts on chemoautotrophic bacteria have started to gain attention, the application of chemoautotrophs at an industrial scale is still challenging, due to their slow growth pattern and the limited applicable genetic engineering tools. *Cupriavidus necator*, which has one of the highest growth rates among natural autotrophic bacteria, was successfully used to produce FACs from CO_2. However, the autotrophic production level of FACs was much lower compared to heterotrophic production on fructose (Crépin et al., 2016). Another chemolithotrophic oleaginous bacterium, *Rhodococcus opacus*, was engineered to produce fatty acids and fuels as high as 50.2 g/L, however this was carried out under heterotrophic condition with glucose as carbon source (Kim et al., 2019). Recently some chemoautotrophs were observed to utilize electricity as energy resources for biosynthesis, which make them promising microbial chassis (Geelhoed and Stams, 2011). Photoautotrophic microorganisms are cells that capture light energy to fix carbon. Among these microorganisms, cyanobacteria are the most studied because it is easy to genetically modify (Liu et al., 2011; Tan et al., 2011; Wang et al., 2013), and biosynthesis of FACs has already been proven feasible in them (Liu et al., 2011; Eungrasamee et al., 2019). For examples, an important Omega−3 fatty acid was produced by overexpression of desaturase *desA* and *desB* in *Synechococcus* sp. PCC 7002 (Santos-Merino et al., 2018), and fatty alcohols were also successfully produced in photosynthesis-driven cyanobacteria (Tan et al., 2011; Yunus and Jones, 2018). By overexpressing acyl-acyl carrier protein reductase (AAR) and aldehyde decarbonylase (AD), metabolically engineered

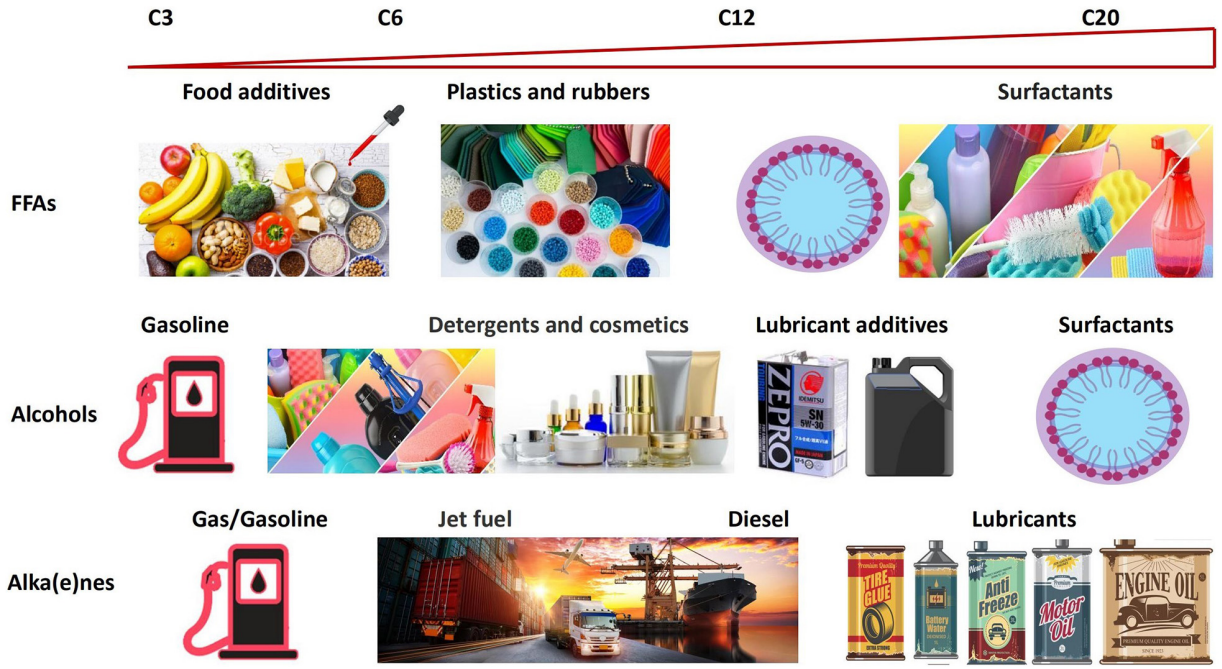


FIGURE 1 | Overview of types and applications of FACs. FACs have a wide application range based on their various chemical properties caused by different end groups and chain lengths.

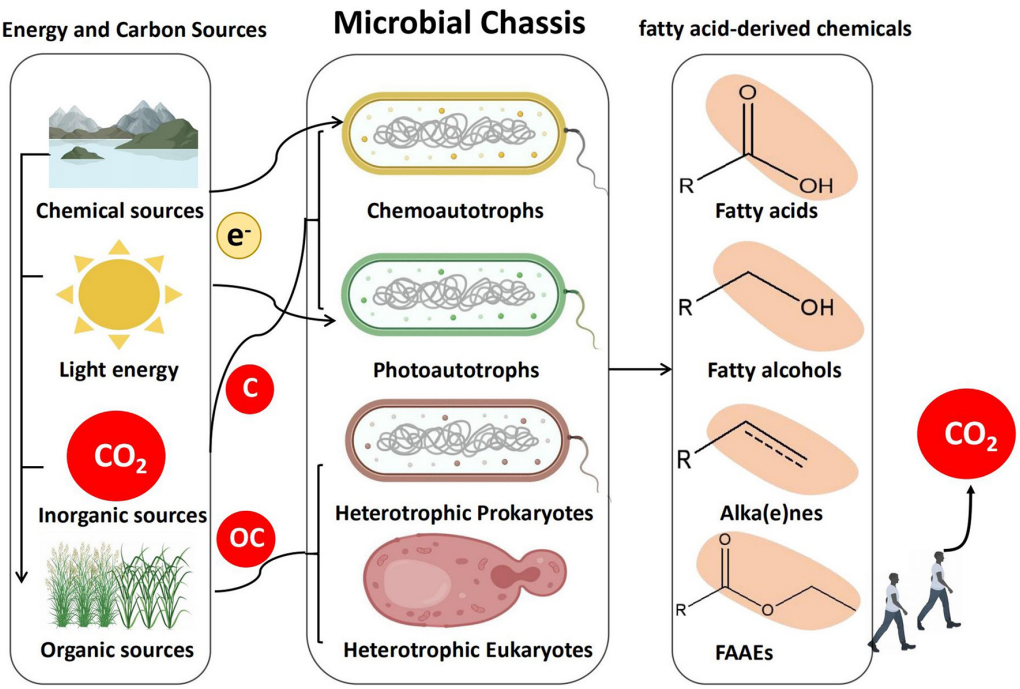


FIGURE 2 | Strategies for production of FACs in microbial chassis. Autotrophs and heterotrophs synthesize FACs from CO_2 and organic carbon (OC), respectively. Either platforms have their own advantages and limitations. Autotroph-heterotroph consortia represent a promising platform for microbial production of FACs.

cyanobacterium, *Nostoc punctiforme*, produced alkanes at levels up to 12% of their cell dry weight (Peramuna et al., 2015). In the case of the heterotrophs, both heterotrophic prokaryotes and eukaryotes have been widely used for FACs biosynthesis (Wang et al., 2010; Rutter and Rao, 2016; Wu et al., 2017; Xin et al., 2017; McNeil and Stuart, 2018; Zhou et al., 2018;

Wu et al., 2019). *E. coli,* the most commonly used heterotrophic prokaryote, was engineered to produce free fatty acids at a titer of up to 21.5 g/L (Xiao et al., 2016). The most commonly used heterotrophic eukaryote, *Saccharomyces cerevisiae* was designed and engineered to produce 33.4 g/L of extracellular free fatty acids (Yu et al., 2018). Some oleaginous heterotrophic eukaryotes, such as *Yarrowia lipolytica* and *Aureobasidium pullulans* show great potential for FACs biosynthesis, since just with simple adjustment they can reach much higher titers than the engineered *S. cerevisiae* (Xu et al., 2016; Xin et al., 2017). Although these microbial platforms have their own advantages, they also have their own limitations. For example, autotrophs can synthesize FACs from CO_2 via using solar, chemical and electric energy. Due to the abundance of CO_2 in the atmosphere and its role in driving global climate change, CO_2-assimilating microbes represent a unique and promising type of microbial chassis for FACs biosynthesis. However, autotrophs normally have limited growth rates and genetic engineering tools, resulting in difficulties to engineer metabolic pathways to produce specific FACs. Comparatively, many genetic engineering tools have been developed for fast-growing and metabolically versatile heterotrophs. But using organic carbon substrates makes them less environmental friendliness. Therefore, a platform combining different types of microorganism is required for more economical, environmental, and efficient microbial factories. Fortunately, recent advances in synthetic biology have made synthetic microbial communities possible (Johns et al., 2016; Tang, 2019; Liu et al., 2020). There are microalgae-microalgae, microalgae and bacteria, microalgae and molds communities constructed for FACs production (Magdouli et al., 2016). Ideally, a microbial community system could use autotrophs to fix CO_2 from the atmosphere and subsequently transfer the organic products to heterotrophs for FACs biosynthesis. One such system has already been reported, a *Synechococcus elongates-Pseudomonas putida* consortium was constructed to produce bioplastic (PHA, polyhydroxyakanotate) (García-Jiménez et al., 2018). Another microalgae-yeast co-culture was isolated from wastewater and identified to contain a number of microalgae and yeast species, which was also successfully used for fatty acid methyl esters production (Suastes-Rivas et al., 2019). From our perspective, these autotroph-heterotroph communities have the potential to produce specific FACs from CO_2. However, the slow rates of CO_2 fixation in autotrophs seriously affect their practical applications. We believe that direct evolution of these synthetic communities could be a promising solution to overcome this limitation (Chang et al., 2020).

SYSTEMS ENGINEERING STRATEGIES FOR BIOSYNTHETIC PATHWAY OPTIMIZATION

The metabolic pathway for FACs biosynthesis can be broadly divided into three steps: initiation, elongation and termination. It starts with the conversion of feed materials to the universal precursor acetyl-CoA through various conversion pathways. The most common pathway for the synthesis of acetyl-CoA is through glycolysis, which converts glucose into pyruvate, then can be decarboxylated to produce acetyl-CoA. However, the decarboxylation of pyruvate loses a carbon equivalent, thus limits the theoretical carbon yield, and constrains the commercialization potential. Fortunately, a non-oxidative pathway was built up to produce stoichiometric amounts of acetyl-CoA from hexose, pentose and triose phosphates without carbon loss (Bogorad et al., 2013). Another intriguing alternative to sugars is the potential to produce acetyl-CoA from one-carbon resources such as CO_2 and formate (Lu et al., 2019). After external carbon sources being converted into acetyl-CoA, it can be directly used as initiation blocks or transformed to propionyl-CoA, acetyl-acyl carrier protein (ACP), and propionyl-ACP for initiation. The initiation pattern determines the odd or even carbon chain of the produced FACs (**Supplementary Figure 1**; Dellomonaco et al., 2011; Park et al., 2020; Zhang et al., 2020).

In terms of elongation, fatty acid synthesis (FAS) and reverse beta-oxidation (RBO) pathways are the two identified routes for FACs biosynthesis. Although the four serial steps including condensation, reduction, dehydration, and reduction are similar in both FAS and RBO pathways (**Figure 3**), the iterative feeding strategies are different. In FAS pathway, acetyl-CoA was transferred into malonyl-ACP before being fed into the elongation cycle, while acetyl-CoA was directly fed into the elongation cycle in RBO pathway (**Figure 3**). FAS has been most widely studied and engineered to produce free fatty acids, alcohols, esters, and alkanes (Liu et al., 2016, 2018; Wenning et al., 2017; Yunus and Jones, 2018). However, RBO is widely accepted as the promising pathway for several reasons. Firstly, one ATP will be saved via the RBO pathway, as acetyl-CoA can be directly fed into the elongation cycle, while for elongation in FAS, acetyl-CoA must first be converted to malonyl-CoA via an ATP-consuming acetyl-CoA carboxylase. Secondly, most reductases from the FAS pathway have been shown to prefer NADPH as cofactors (Ratledge, 2004; Handke et al., 2011; Javidpour et al., 2014). In contrast, reductases from the RBO routes are mostly NADH-dependent (Lian and Zhao, 2014; Sheppard et al., 2016; Kim and Gonzalez, 2018). Since it has been demonstrated that cell has relatively high NADH/NAD^+ ratio under anaerobic condition (De Graef et al., 1999), RBO will be benefited in the anaerobic biosynthesis of FACs. Thirdly, RBO pathway is dependent on the universal CoA molecule, while FAS pathway is dependent on organism specific A, making RBO pathway more transferable in target microorganisms. Recently, it was reported that with the exception of condensation step, the remaining enzymes for other steps in FAS pathway: 3-ketoacyl-ACP reductase (*FabG*), 3-hydroxyacylACP dehydratase (*FabZ*), and enoyl-ACP reductase identified (*FabI*) can carry out similar conversions as in RBO in *E. coli* (Vick et al., 2015; Clomburg et al., 2018). It was reported that some of these enzymes show preference for acyl-ACP intermediates, such as *fabZ* from *E. coli* (Tsuge et al., 2003). These findings present both opportunities and challenges. On one hand, it offers the potential to employ FAS enzymes on various acyl-CoA intermediates, which can greatly expand the range of FACs produced by RBO. On the other hand, it will lead to promiscuous activity and make it difficult to produce specific FACs, as intracellular substrates

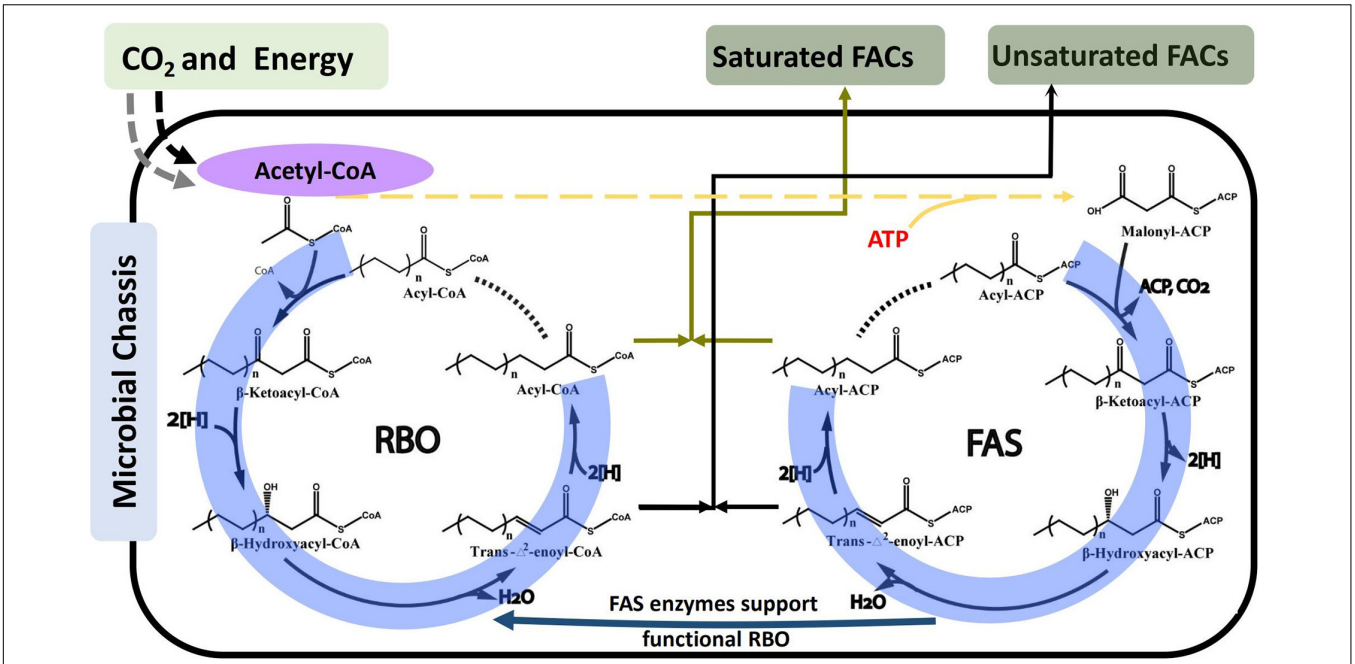


FIGURE 3 | Schematic of metabolic pathways for FACs biosynthesis via Reverse Beta Oxidation (RBO) and Fatty Acid Synthesis (FAS). The entry of CO_2 and energy to acetyl-CoA depends on the type of microbial chassis selected, as shown in **Figure 2**. The catalytic reaction cycles for both RBO and FAS include the same four serial steps: condensation, reduction, dehydration, and reduction. The major difference is that RBO employs a non-decarboxylative condensation mechanism to incorporate acetyl-CoA while FAS utilizes a decarboxylative condensation mechanism to incorporate malonyl-ACP.

will be automatically used by the endogenous FAS enzymes, which results in impure and unwanted products. Though both of FAS and RBO pathways can be used to produce FACs with different chain lengths, FAS pathway is preferred for long-chain FACs production, because it naturally has high efficiency; while for short-chain FACs, the RBO pathway is favored because it is easier to control product lengths than FAS pathway (Sheppard et al., 2016).

The termination step, which releases fatty acyl-CoA or fatty acyl-ACP from the elongation cycle, is the most important and widely investigated step, as it determines the types of FACs produced by microbial cell factories. For each type of FACs, there are multiple options for termination. For example, fatty alkenes and fatty alcohols can be generated from fatty acids, fatty acyl-ACPs and fatty acyl-CoAs (Liu et al., 2016; Liu and Li, 2020). Alkanes can be converted from fatty aldehydes by aldehyde decarbonylase or from fatty acids by photodecarboxylase (Eser et al., 2011; Sorigué et al., 2017). Even though numerous terminal pathway options have been found, it is still the major bottleneck for FACs biosynthesis for the following reasons: First, production of FACs other than FFAs is not efficient. According to our knowledge, the highest titer of mixed long chain FFAs (C14–C22) is 50.2 g/L using an oleaginous bacterium *Rhodococcus opacus* PD630 (Kim et al., 2019), while the highest titers of fatty alcohols and alka(e)nes is 12.5 and 2.54 g/L, respectively (Fatma et al., 2018). Considering the same upstream pathway, the low titers of fatty alcohols and alka(e)nes are perhaps caused by the low efficiency of enzymes in the termination step. Second, enzymes in terminal step naturally prefer longer chain substrates. Although great efforts have been made for short chain substrates, the problem is far from being resolved (Khara et al., 2013; Gajewski et al., 2017). Hence, screening and engineering of enzymes that prefer short chain substrates should be an important area of research.

MODEL-ASSISTED DESIGN FOR BIOSYNTHESIS OF FACS

Model-assisted design has shown to be successful in metabolic engineering (Teusink and Smid, 2006; Fatma et al., 2018; Ferreira et al., 2019; Das et al., 2020; **Figure 4**). Increasing information in databases, such as KEGG, BioCyc, BRENDA, MetRxn, and SEED (Shin et al., 2013; Long et al., 2015; Delépine et al., 2018; Choi et al., 2019), makes it possible to develop organism-specific reaction networks, *de novo* pathway predictions and even retrosynthetic design of metabolic pathways for non-natural chemicals (Medema et al., 2012; Tabei et al., 2016; Biz et al., 2019; Garcia and Trinh, 2019a). Model-assisted design facilitates efficient Design-Build-Test-Learn (DBTL) cycle, avoiding costly trial and error approaches (Long et al., 2015; Choi et al., 2019). There are two basic metabolic analysis algorithms for model-guided design in metabolic engineering: Flux balance analysis (FBA) and Elementary mode analysis (EMA) (Mahadevan et al., 2002; Klamt and Gilles, 2004; Machado and Herrgård, 2015). FBA uses linear optimization to find a set of reaction fluxes that satisfy both an objective function and a set of constraints limiting the solution space of the network representing a given growth

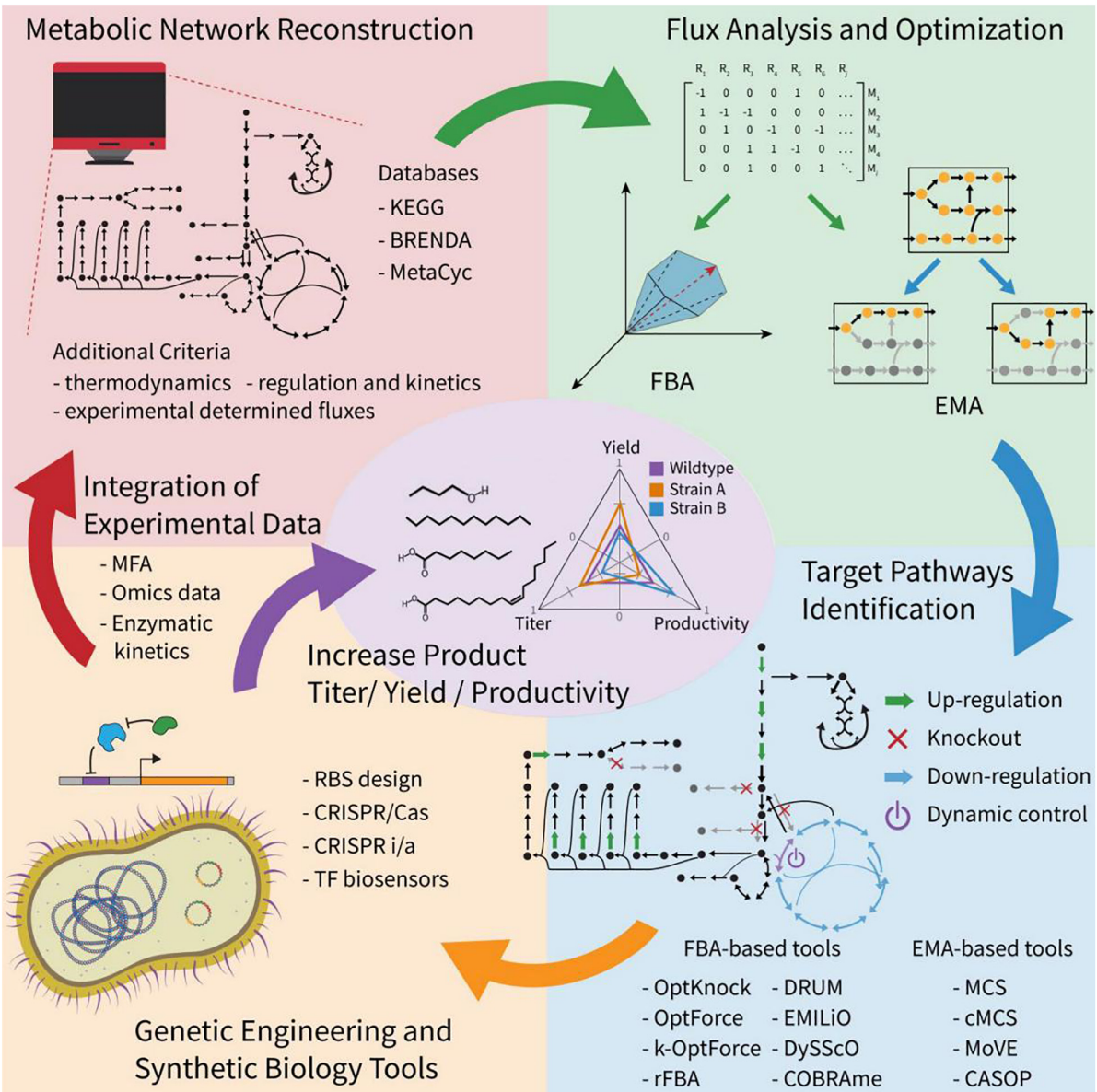


FIGURE 4 | Overview of computational model-assisted design algorithms. Computational methods have been developed to identify pathways to enable production of desired compounds using known or *de novo* metabolic pathways. Additional host metabolic strategies, including up/down-regulation, knockout, and dynamic control can be designed using a variety of constraint-based modeling algorithms.

condition (Orth et al., 2010). EMA calculates all the solutions with minimal support in the network that satisfy the steady state and other constraints (Trinh et al., 2009). Each solution in EMA is called an elementary mode (EM) and describes the topology of the metabolic network, which is useful in determining its properties and for rational design (Trinh et al., 2009). Most computational tools are derived from EMA or FBA for more specific purposes (Rodrigo et al., 2008; Trinh et al., 2009; Campodonico et al., 2014; Long et al., 2015; Garcia and Trinh, 2019b, 2020; Jiang et al., 2020).

Methods based on FBA, such as OptKnock, OptStrain, OptForce, dFBA, DySScO DynamicME, and COBRAme have been developed for strain engineering purposes, to identify a set of genetic interventions to increase the production of target compounds (Mahadevan et al., 2002; Burgard et al., 2003; Pharkya et al., 2004; Ranganathan et al., 2010; Zhuang et al., 2013; Lloyd et al., 2018; Yang et al., 2019). The two most used design programs based on FBA are OptKnock and OptForce. OptKnock is the first bi-level optimization framework for strain design, which can identify optimal reaction deletion strategies that couple cellular growth with the production of a target metabolite (Burgard et al., 2003). A successful case of OptKnock algorithm application is a growth-coupled strategy designed for biofuel production in *Synechocystis*, and it shows

that lowering the ATP/NADPH ratio in the cell is a promising strategy for biosynthesis of fatty alcohols and alkanes (Shabestary and Hudson, 2016). OptForce is a framework predicting genetic interventions such as overexpression and repression based on the comparison of an initial metabolic status and the desired overproduction goal (Ranganathan et al., 2010). It can prioritize the interventions according to their effects on the increment of the production, making it possible to start with the modifications that would have higher impact on the process. The OptForce algorithm has been successfully used for strategy design in *E. coli* for fatty acids production. Moreover, it can predict less intuitive interventions, such as the redirection of the flux through the Entner-Doudoroff pathway to produce NADPH and induce a growth arrest limiting the ATP production (Ranganathan et al., 2012; Tee et al., 2014). Recently, another study applied OptForce for the production of octanoic acid, achieving high selectivity (>70%) and an extracellular concentration up to 1 g/L of free octanoic acid in minimal medium via fed-batch culture (Liu et al., 2018). A metabolic model was constructed for long-chain alkane and alcohol production based on FBA analysis, and the engineered strain produced the maximum titers of hydrocarbons (Fatma et al., 2018). Besides, recently breakthroughs have been made in visualizing genome-scale metabolic flux networks, which improved understanding of the predicted solutions (Chazalviel et al., 2018; Hari and Lobo, 2020).

EMA-based tools have been used for strain engineering by applying the concept of minimal cut sets (MCS) (Klamt and Gilles, 2004). Constrained MCS (cMCS) was developed to remove limitations in MCS, where many solutions also eliminated growth (Hädicke and Klamt, 2011). Using cMCS, researchers identified sets of reactions to eliminate and enhanced the production of ethanol and isobutanol in *Clostridium thermocellum* and cyanobacteria, respectively (Erdrich et al., 2014; Thompson and Trinh, 2017). There has been interest in dynamic control strategies, which can dynamically regulate of flux through metabolite sensor, inducer, temperature, light and cell density (Lalwani et al., 2018; Liu et al). These provide the option to prioritize growth or production in a two-stage process, which can lead to higher yields, productivities and titers of FACs (Zhang et al., 2012; Lalwani et al., 2018; Raj et al., 2020). To accelerate the strain engineering process for enhanced chemical production, MODCELL and MODCELL 2 frameworks were developed for rapid generation of optimal production strains by systematically assembling a modular cell with an exchangeable production module (Trinh et al., 2015; Garcia and Trinh, 2019c). Moreover, MoVE, a newly developed tool based on MCS, can identify genetic interventions that allow the transition between growth and production states for dynamic control of the metabolism (Venayak et al., 2018).

Once engineering strategies are obtained from computational modeling, there are plenty of synthetic biology tools available to implement the suggested metabolic engineering interventions. For instance, CRISPR-based technologies make it possible to perform multiple knockouts, inhibitions, or activations of designed sets (Behler et al., 2018; Kaczmarzyk et al., 2018; Reis et al., 2019). Significant improvements have been achieved in FACs biosynthesis using model-based strain design strategies (Matsuda et al., 2011; Shabestary and Hudson, 2016; Fatma et al., 2018; Yu et al., 2018), however, there remain challenges to be addressed in future studies. For example, there is still a lack of methods to integrate large amounts of data into genome-scale models and provide user-friendly tools that allow users with no programming experience to exploit the potential of genome-scale metabolic models for rational design. In conclusion, model-based strain engineering is still in an early stage and its application has been limited to few chemical targets and tools. We expect that the development of novel user-friendly computational models can enable increased adoption of such tools for various types of FACs production.

CONCLUSION

The ongoing reliance on fossil fuels of human society is driving elevated atmospheric CO_2 and increasing global temperatures, thereby escalating the risk of widespread environmental disasters in the near future. We anticipate that microbial synthesis of products from CO_2, which can provide chemicals with near-zero net greenhouse gas emissions, will play as a game-changer in the future (Ediger, 2019). Great progress has been made in the areas of enzyme engineering, metabolic engineering, and model-assisted engineering to assist microbial production of FACs (Cao et al., 2016; Herman and Zhang, 2016; Kim et al., 2016; Zhou et al., 2016; Fatma et al., 2018; Marella et al., 2018; Kim and Park, 2019; Liu and Nielsen, 2019; Lynch et al., 2019). However, the present-day microbial cell factories still have major challenges to overcome, such as controlling the length and types of released FACs and improving the conversion efficiency via RBO. We expect that directed enzyme evolution and rational enzyme engineering will contribute to the production of target FACs through the RBO pathway. Recently, there are some machine learning-based algorithms developed for computational protein design, which can also be used in enzyme engineering (Masso and Vaisman, 2008; Fang, 2019; Zu Belzen et al., 2019). In addition, new methods for design and build of synthetic microorganism communities can contribute to the construction of novel microbial platforms, which combine carbon-fixing autotrophs with heterotrophs for efficient FACs biosynthesis with net-zero greenhouse gas emissions.

AUTHOR CONTRIBUTIONS

YL and RM conceived of the idea. YL, MB, JC, EH, and AK wrote the manuscript. YL, JC, EH, MB, and RM contributed to revising. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Ontario Ministry of Research and Innovation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.613322/full#supplementary-material

Supplementary Figure 1 | Initiation and feed in blocks in Reverse Beta Oxidation (RBO) **(A)** and Fatty Acid Synthesis (FAS) **(B)** pathways. Each turn of the FAS and RBO cycles leads to the addition of two carbons through incorporation of malonyl-ACP or acetyl-CoA, respectively. Consequently, the initiation blocks control over whether the produced FACs have odd or even carbon chains.

REFERENCES

Behler, J., Vijay, D., Hess, W. R., and Akhtar, M. K. (2018). CRISPR-based technologies for metabolic engineering in cyanobacteria. *Trends Biotechnol.* 36, 996–1010. doi: 10.1016/j.tibtech.2018.05.011

Bellou, S., Triantaphyllidou, I.-E., Aggeli, D., Elazzazy, A. M., Baeshen, M. N., and Aggelis, G. (2016). Microbial oils as food additives: recent approaches for improving microbial oil production and its polyunsaturated fatty acid content. *Curr. Opin. Biotechnol.* 37, 24–35. doi: 10.1016/j.copbio.2015.09.005

Bentley, G. J., Jiang, W., Guamán, L. P., Xiao, Y., and Zhang, F. (2016). Engineering *Escherichia coli* to produce branched-chain fatty acids in high percentages. *Metab. Eng.* 38, 148–158. doi: 10.1016/j.ymben.2016.07.003

Biz, A., Proulx, S., Xu, Z., Siddartha, K., Indrayanti, A. M., and Mahadevan, R. (2019). Systems biology based metabolic engineering for non-natural chemicals. *Biotechnol. Adv.* 37:107379. doi: 10.1016/j.biotechadv.2019.04.001

Bogorad, I. W., Lin, T.-S., and Liao, J. C. (2013). Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* 502, 693–697. doi: 10.1038/nature12575

Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* 84, 647–657. doi: 10.1002/bit.10803

Campodonico, M. A., Andrews, B. A., Asenjo, J. A., Palsson, B. O., and Feist, A. M. (2014). Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm. *GEM Path. Metab. Eng.* 25, 140–158. doi: 10.1016/j.ymben.2014.07.009

Cao, Y.-X., Xiao, W.-H., Zhang, J.-L., Xie, Z.-X., Ding, M.-Z., and Yuan, Y.-J. (2016). Heterologous biosynthesis and manipulation of alkanes in *Escherichia coli*. *Metab. Eng.* 38, 19–28. doi: 10.1016/j.ymben.2016.06.002

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., et al. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.* 1, 1–10. doi: 10.1007/978-3-319-31421-1_214-1

Chang, C.-Y., Vila, J. C., Bender, M., Li, R., Mankowski, M. C., Bassette, M., et al. (2020). Top-down engineering of complex communities by directed evolution. *bioRxiv* [Preprint]. doi: 10.1101/2020.07.24.214775v2

Chazalviel, M., Frainay, C., Poupin, N., Vinson, F., Merlet, B., Gloaguen, Y., et al. (2018). MetExploreViz: web component for interactive metabolic network visualization. *Bioinformatics* 34, 312–313. doi: 10.1093/bioinformatics/btx588

Chen, Y., Banerjee, D., Mukhopadhyay, A., and Petzold, C. J. (2020). Systems and synthetic biology tools for advanced bioproduction hosts. *Curr. Opin. Biotechnol.* 64, 101–109. doi: 10.1016/j.copbio.2019.12.007

Cho, I. J., Choi, K. R., and Lee, S. Y. (2020). Microbial production of fatty acids and derivative chemicals. *Curr. Opin. Biotechnol.* 65, 129–141. doi: 10.1016/j.copbio.2020.02.006

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., and Lee, S. Y. (2019). Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37, 817–837. doi: 10.1016/j.tibtech.2019.01.003

Clomburg, J. M., Contreras, S. C., Chou, A., Siegel, J. B., and Gonzalez, R. (2018). Combination of type II fatty acid biosynthesis enzymes and thiolases supports a functional β-oxidation reversal. *Metab. Eng.* 45, 11–19. doi: 10.1016/j.ymben.2017.11.003

Crépin, L., Lombard, E., and Guillouet, S. E. (2016). Metabolic engineering of Cupriavidus necator for heterotrophic and autotrophic alka (e) ne production. *Metab. Eng.* 37, 92–101. doi: 10.1016/j.ymben.2016.05.002

Das, M., Patra, P., and Ghosh, A. (2020). Metabolic engineering for enhancing microbial biosynthesis of advanced biofuels. *Renewab. Sustain. Energy Rev.* 119:109562. doi: 10.1016/j.rser.2019.109562

De Graef, M. R., Alexeeva, S., Snoep, J. L., and De Mattos, M. J. T. (1999). The steady-state internal redox state (NADH/NAD) reflects the external redox state and is correlated with catabolic adaptation in *Escherichia coli*. *J. Bacteriol.* 181, 2351–2357. doi: 10.1128/jb.181.8.2351-2357.1999

Delépine, B., Duigou, T., Carbonell, P., and Faulon, J.-L. (2018). RetroPath2.0: a retrosynthesis workflow for metabolic engineers. *Metab. Eng.* 45, 158–170. doi: 10.1016/j.ymben.2017.12.002

Dellomonaco, C., Clomburg, J. M., Miller, E. N., and Gonzalez, R. (2011). Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* 476, 355–359. doi: 10.1038/nature10333

Dong, X., and Mumper, R. (2006). The metabolism of fatty alcohols in lipid nanoparticles by alcohol dehydrogenase. *Drug Dev. Ind. Pharm.* 32, 973–980. doi: 10.1080/03639040600640006

Ediger, V. Ş (2019). An integrated review and analysis of multi-energy transition from fossil fuels to renewables. *Energy Proc.* 156, 2–6. doi: 10.1016/j.egypro.2018.11.073

Erdrich, P., Knoop, H., Steuer, R., and Klamt, S. (2014). Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. *Microb. Cell Fact.* 13:128.

Eser, B. E., Das, D., Han, J., Jones, P. R., and Marsh, E. N. G. (2011). Oxygen-independent alkane formation by non-heme iron-dependent cyanobacterial aldehyde decarbonylase: investigation of kinetics and requirement for an external electron donor. *Biochemistry* 50, 10743–10750. doi: 10.1021/bi2012417

Eungrasamee, K., Miao, R., Incharoensakdi, A., Lindblad, P., and Jantaro, S. (2019). Improved lipid production via fatty acid biosynthesis and free fatty acid recycling in engineered *Synechocystis* sp. PCC 6803. *Biotechnol. Biofuels* 12:8.

Fang, J. (2019). A critical review of five machine learning-based algorithms for predicting protein stability changes upon mutation. *Briefings Bioinform.* 21, 1285–1292. doi: 10.1093/bib/bbz071

Fatma, Z., Hartman, H., Poolman, M. G., Fell, D. A., Srivastava, S., Shakeel, T., et al. (2018). Model-assisted metabolic engineering of *Escherichia coli* for long chain alkane and alcohol production. *Metab. Eng.* 46, 1–12. doi: 10.1016/j.ymben.2018.01.002

Ferreira, R., Skrekas, C., Hedin, A., Sánchez, B. J., Siewers, V., Nielsen, J., et al. (2019). Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation. *ACS Synth. Biol.* 8, 2457–2463. doi: 10.1021/acssynbio.9b00258

Fillet, S., and Adrio, J. L. (2016). Microbial production of fatty alcohols. *World J. Microbiol. Biotechnol.* 32:152.

Gajewski, J., Pavlovic, R., Fischer, M., Boles, E., and Grininger, M. (2017). Engineering fungal de novo fatty acid synthesis for short chain fatty acid production. *Nat. Commun.* 8, 1–8.

Garcia, S., and Trinh, C. T. (2019a). Comparison of multi-objective evolutionary algorithms to solve the modular cell design problem for novel biocatalysis. *Processes* 7:361. doi: 10.3390/pr7060361

Garcia, S., and Trinh, C. T. (2019b). Modular design: implementing proven engineering principles in biotechnology. *Biotechnol. Adv.* 37:107403. doi: 10.1016/j.biotechadv.2019.06.002

Garcia, S., and Trinh, C. T. (2019c). Multiobjective strain design: a framework for modular cell engineering. *Metab. Eng.* 51, 110–120. doi: 10.1016/j.ymben.2018.09.003

Garcia, S., and Trinh, C. T. (2020). Harnessing natural modularity of metabolism with goal attainment optimization to design a modular chassis cell for production of diverse chemicals. *ACS Synth. Biol.* 9, 1665–1681. doi: 10.1021/acssynbio.9b00518

García-Jiménez, B., García, J. L., and Nogales, J. (2018). FLYCOP: metabolic modeling-based analysis and engineering microbial communities. *Bioinformatics* 34, i954–i963.

Geelhoed, J. S., and Stams, A. J. (2011). Electricity-assisted biological hydrogen production from acetate by Geobacter sulfurreducens. *Environ. Sci. Technol.* 45, 815–820. doi: 10.1021/es102842p

Geller, D. P., and Goodrum, J. W. (2004). Effects of specific fatty acid methyl esters on diesel fuel lubricity. *Fuel* 83, 2351–2356. doi: 10.1016/j.fuel.2004.06.004

Guo, X., Zhang, R. M., Gao, L. G., Zhang, X., and Xu, X. (2019). Computational kinetics of the hydrogen abstraction reactions of n-propanol and iso-propanol by OH radical. *Phys. Chem. Chem. Phys.* 21, 24458–24468. doi: 10.1039/c9cp04809j

Gupta, C., Prakash, D., and Gupta, S. (2015). A biotechnological approach to microbial based perfumes and flavours. *J. Microbiol. Exp.* 3, 11–18.

Hädicke, O., and Klamt, S. (2011). Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metab. Eng.* 13, 204–213. doi: 10.1016/j.ymben.2010.12.004

Hamedirad, M., Chao, R., Weisberg, S., Lian, J., Sinha, S., and Zhao, H. (2019). Towards a fully automated algorithm driven platform for biosystems design. *Nat. Commun.* 10, 1–10.

Handke, P., Lynch, S. A., and Gill, R. T. (2011). Application and engineering of fatty acid biosynthesis in *Escherichia coli* for advanced fuels and chemicals. *Metab. Eng.* 13, 28–37. doi: 10.1016/j.ymben.2010.10.007

Hari, A., and Lobo, D. (2020). *Fluxer: A Web Application to Compute, Analyze and Visualize Genome-Scale Metabolic Flux Networks*. Baltimore, MD: UMBC Faculty Collection.

Herman, N. A., and Zhang, W. (2016). Enzymes for fatty acid-based hydrocarbon biosynthesis. *Curr. Opin. Chem. Biol.* 35, 22–28. doi: 10.1016/j.cbpa.2016.08.009

Hu, Y., Zhu, Z., Nielsen, J., and Siewers, V. (2019). Engineering Saccharomyces cerevisiae cells for production of fatty acid-derived biofuels and chemicals. *Open Biol.* 9:190049. doi: 10.1098/rsob.190049

Javidpour, P., Pereira, J. H., Goh, E.-B., Mcandrew, R. P., Ma, S. M., Friedland, G. D., et al. (2014). Biochemical and structural studies of NADH-dependent FabG used to increase the bacterial production of fatty acids under anaerobic conditions. *Appl. Environ. Microbiol.* 80, 497–505. doi: 10.1128/aem.03194-13

Jiang, L., Fu, H., Yang, H. K., Xu, W., Wang, J., and Yang, S.-T. (2018). Butyric acid: applications and recent advances in its bioproduction. *Biotechnol. Adv.* 36, 2101–2117. doi: 10.1016/j.biotechadv.2018.09.005

Jiang, S., Wang, Y., Kaiser, M., and Krasnogor, N. (2020). NIHBA: a network interdiction approach for metabolic engineering design. *Bioinformatics* 36, 3482–3492. doi: 10.1093/bioinformatics/btaa163

Jiang, W., Gu, P., and Zhang, F. (2018). Steps towards 'drop-in'biofuels: focusing on metabolic pathways. *Curr. Opin. Biotechnol* 53, 26–32. doi: 10.1016/j.copbio.2017.10.010

Jiang, W., Qiao, J. B., Bentley, G. J., Liu, D., and Zhang, F. (2017). Modular pathway engineering for the microbial production of branched-chain fatty alcohols. *Biotechnol. Biofuels* 10:244.

Johns, N. I., Blazejewski, T., Gomes, A. L., and Wang, H. H. (2016). Principles for designing synthetic microbial communities. *Curr. Opin. Microbiol.* 31, 146–153. doi: 10.1016/j.mib.2016.03.010

Kaczmarzyk, D., Cengic, I., Yao, L., and Hudson, E. P. (2018). Diversion of the long-chain acyl-ACP pool in Synechocystis to fatty alcohols through CRISPRi repression of the essential phosphate acyltransferase PlsX. *Metab. Eng.* 45, 59–66. doi: 10.1016/j.ymben.2017.11.014

Kang, M.-K., Zhou, Y. J., Buijs, N. A., and Nielsen, J. (2017). Functional screening of aldehyde decarbonylases for long-chain alkane production by Saccharomyces cerevisiae. *Microb. Cell Fact.* 16:74.

Khara, B., Menon, N., Levy, C., Mansell, D., Das, D., Marsh, E. N. G., et al. (2013). Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenase. *ChemBioChem* 14, 1204–1208. doi: 10.1002/cbic.201300307

Kim, H. M., Chae, T. U., Choi, S. Y., Kim, W. J., and Lee, S. Y. (2019). Engineering of an oleaginous bacterium for the production of fatty acids and fuels. *Nat. Chem. Biol.* 15, 721–729. doi: 10.1038/s41589-019-0295-5

Kim, S., Cheong, S., Chou, A., and Gonzalez, R. (2016). Engineered fatty acid catabolism for fuel and chemical production. *Curr. Opin. Biotechnol.* 42, 206–215. doi: 10.1016/j.copbio.2016.07.004

Kim, S., and Gonzalez, R. (2018). Selective production of decanoic acid from iterative reversal of β-oxidation pathway. *Biotechnol. Bioeng.* 115, 1311–1320. doi: 10.1002/bit.26540

Kim, S.-K., and Park, Y.-C. (2019). Biosynthesis of ω-hydroxy fatty acids and related chemicals from natural fatty acids by recombinant *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 103, 191–199. doi: 10.1007/s00253-018-9503-6

Klamt, S., and Gilles, E. D. (2004). Minimal cut sets in biochemical reaction networks. *Bioinformatics* 20, 226–234. doi: 10.1093/bioinformatics/btg395

Lalwani, M. A., Zhao, E. M., and Avalos, J. L. (2018). Current and future modalities of dynamic control in metabolic engineering. *Curr. Opin. Biotechnol.* 52, 56–65. doi: 10.1016/j.copbio.2018.02.007

Leber, C., Polson, B., Fernandez-Moya, R., and Da Silva, N. A. (2015). Overproduction and secretion of free fatty acids through disrupted neutral lipid recycle in Saccharomyces cerevisiae. *Metab. Eng.* 28, 54–62. doi: 10.1016/j.ymben.2014.11.006

Li, G., Huang, D., Sui, X., Li, S., Huang, B., Zhang, X., et al. (2020). Advances in microbial production of medium-chain dicarboxylic acids for nylon materials. *React. Chem. Eng.* 5, 221–238. doi: 10.1039/c9re00338j

Li, M., Hou, F., Wu, T., Jiang, X., Li, F., Liu, H., et al. (2020). Recent advances of metabolic engineering strategies in natural isoprenoid production using cell factories. *Nat. Product Rep.* 37, 80–99. doi: 10.1039/c9np00016j

Lian, J., and Zhao, H. (2014). Reversal of the β-oxidation cycle in Saccharomyces cerevisiae for production of fuels and chemicals. *ACS Synth. Biol.* 4, 332–341. doi: 10.1021/sb500243c

Liu, D., Mao, Z., Guo, J., Wei, L., Ma, H., Tang, Y., et al. (2018). Construction, model-based analysis, and characterization of a promoter library for fine-tuned gene expression in Bacillus subtilis. *ACS Synth. Biol.* 7, 1785–1797. doi: 10.1021/acssynbio.8b00115

Liu, K., and Li, S. (2020). Biosynthesis of fatty acid-derived hydrocarbons: perspectives on enzymology and enzyme engineering. *Curr. Opin. Biotechnol.* 62, 7–14. doi: 10.1016/j.copbio.2019.07.005

Liu, X., Sheng, J., and Curtiss Iii, R. (2011). Fatty acid production in genetically modified cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6899–6904. doi: 10.1073/pnas.1103014108

Liu, Y., Chen, J., Crisante, D., Lopez, J. M. J., and Mahadevan, R. (2020). Dynamic cell programming with quorum sensing-controlled CRISPRi circuit. *ACS Synth. Biol.* 9, 1284–1291. doi: 10.1021/acssynbio.0c00148

Liu, Y., Chen, S., Chen, J., Zhou, J., Wang, Y., Yang, M., et al. (2016). High production of fatty alcohols in *Escherichia coli* with fatty acid starvation. *Microb. Cell Fact.* 15:129.

Liu, Y., and Nielsen, J. (2019). Recent trends in metabolic engineering of microbial chemical factories. *Curr. Opin. Biotechnol.* 60, 188–197. doi: 10.1016/j.copbio.2019.05.010

Lloyd, C. J., Ebrahim, A., Yang, L., King, Z. A., Catoiu, E., O'brien, E. J., et al. (2018). COBRAme: a computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput. Biol.* 14:e1006302. doi: 10.1371/journal.pcbi.1006302

Long, M. R., Ong, W. K., and Reed, J. L. (2015). Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* 34, 135–141. doi: 10.1016/j.copbio.2014.12.019

Lu, X., Liu, Y., Yang, Y., Wang, S., Wang, Q., Wang, X., et al. (2019). Constructing a synthetic pathway for acetyl-coenzyme A from one-carbon through enzyme design. *Nat. Commun.* 10, 1–10.

Lynch, M., Louie, M., Copley, S., Spindler, E., Prather, B., Lipscomb, M., et al. (2019). *Microorganisms and Methods for The Production of Fatty Acids and Fatty Acid Derived Products. Google Patents*. Wayzata, MN: Cargill, Inc.

Machado, D., and Herrgård, M. J. (2015). Co-evolution of strain design methods based on flux balance and elementary mode analysis. *Metab. Eng. Commun.* 2, 85–92. doi: 10.1016/j.meteno.2015.04.001

Magdouli, S., Brar, S. K., and Blais, J.-F. (2016). Co-culture for lipid production: advances and challenges. *Biomass Bioenergy* 92, 20–30. doi: 10.1016/j.biombioe.2016.06.003

Mahadevan, R., Edwards, J. S., and Doyle Iii, F. J. (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi: 10.1016/s0006-3495(02)73903-9

Marella, E. R., Holkenbrink, C., Siewers, V., and Borodina, I. (2018). Engineering microbial fatty acid metabolism for biofuels and biochemicals. *Curr. Opin. Biotechnol.* 50, 39–46. doi: 10.1016/j.copbio.2017.10.002

Masso, M., and Vaisman, I. I. (2008). Accurate prediction of stability changes in protein mutants by combining machine learning with structure based computational mutagenesis. *Bioinformatics* 24, 2002–2009. doi: 10.1093/bioinformatics/btn353

Matsuda, F., Furusawa, C., Kondo, T., Ishii, J., Shimizu, H., and Kondo, A. (2011). Engineering strategy of yeast metabolism for higher alcohol production. *Microb. Cell Fact.* 10:70. doi: 10.1186/1475-2859-10-70

McNeil, B. A., and Stuart, D. T. (2018). Optimization of C16 and C18 fatty alcohol production by an engineered strain of Lipomyces starkeyi. *J. Ind. Microbiol. Biotechnol.* 45, 1–14. doi: 10.1007/s10295-017-1985-1

Medema, M. H., Van Raaphorst, R., Takano, E., and Breitling, R. (2012). Computational tools for the synthetic design of biochemical pathways. *Nat. Rev. Microbiol.* 10, 191–202. doi: 10.1038/nrmicro2717

Ngo, H. L., Dunn, R. O., and Hoh, E. (2013). C18-unsaturated branched-chain fatty acid isomers: characterization and physical properties. *Eur. J. Lipid Sci. Technol.* 115, 676–683. doi: 10.1002/ejlt.201200323

Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., et al. (2019). Lessons from two design–build–test–learn cycles of dodecanol production in *Escherichia coli* aided by machine learning. *ACS Synth. Biol.* 8, 1337–1351. doi: 10.1021/acssynbio.9b00020

Orth, J. D., Thiele, I., and Palsson, B. Ø (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi: 10.1038/nbt.1614

Park, Y.-K., Ledesma-Amaro, R., and Nicaud, J.-M. (2020). De novo biosynthesis of odd-chain fatty acids in Yarrowia lipolytica enabled by modular pathway engineering. *Front. Bioeng. Biotechnol.* 7:484. doi: 10.3389/fbioe.2019.00484

Peramuna, A., Morton, R., and Summers, M. (2015). Enhancing alkane production in cyanobacterial lipid droplets: a model platform for industrially relevant compound production. *Life* 5, 1111–1126. doi: 10.3390/life5021111

Pharkya, P., Burgard, A. P., and Maranas, C. D. (2004). OptStrain: a computational framework for redesign of microbial production systems. *Genome Res.* 14, 2367–2376. doi: 10.1101/gr.2872004

Raj, K., Venayak, N., and Mahadevan, R. (2020). Phenotypic design choices for enhanced two-stage microbial chemical production processes. *bioRxiv* [Preprint]. doi: 10.1101/803023v2

Ranganathan, S., Suthers, P. F., and Maranas, C. D. (2010). OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput. Biol.* 6:e1000744. doi: 10.1371/journal.pcbi.1000744

Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorrodi, A. R., Yoon, J. M., Fu, Y., et al. (2012). An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metab. Eng.* 14, 687–704. doi: 10.1016/j.ymben.2012.08.008

Raslavičius, L., Keršys, A., Mockus, S., Keršienė, N., and Starevičius, M. (2014). Liquefied petroleum gas (LPG) as a medium-term option in the transition to sustainable fuels and transport. *Renew. Sustain. Energy Rev.* 32, 513–525. doi: 10.1016/j.rser.2014.01.052

Ratledge, C. (2004). Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie* 86, 807–815. doi: 10.1016/j.biochi.2004.09.017

Reis, A. C., Halper, S. M., Vezeau, G. E., Cetnar, D. P., Hossain, A., Clauer, P. R., et al. (2019). Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nat. Biotechnol.* 37, 1294–1301. doi: 10.1038/s41587-019-0286-9

Richardson, A. S., and Mcallister, W. H. (1945). *Detergent Composition*. Google Patents EP19980947649. Wayzata, MN: Cargill, Inc.

Rodrigo, G., Carrera, J., Prather, K. J., and Jaramillo, A. (2008). DESHARKY: automatic design of metabolic pathways for optimal cell growth. *Bioinformatics* 24, 2554–2556. doi: 10.1093/bioinformatics/btn471

Rodriguez-Moya, M., and Gonzalez, R. (2015). Proteomic analysis of the response of *Escherichia coli* to short-chain fatty acids. *J. Proteom.* 122, 86–99. doi: 10.1016/j.jprot.2015.03.033

Röttig, A., Wenning, L., Bröker, D., and Steinbüchel, A. (2010). Fatty acid alkyl esters: perspectives for production of alternative biofuels. *Appl. Microbiol. Biotechnol.* 85, 1713–1733. doi: 10.1007/s00253-009-2383-z

Rutter, C. D., and Rao, C. V. (2016). Production of 1-decanol by metabolically engineered *Yarrowia lipolytica*. *Metab. Eng.* 38, 139–147. doi: 10.1016/j.ymben.2016.07.011

Santos-Merino, M., Garcillán-Barcia, M. P., and De La Cruz, F. (2018). Engineering the fatty acid synthesis pathway in Synechococcus elongatus PCC 7942 improves omega-3 fatty acid production. *Biotechnol. Biofuels* 11:239.

Sathesh-Prabu, C., Shin, K. S., Kwak, G. H., Jung, S.-K., and Lee, S. K. (2019). Microbial production of fatty acid via metabolic engineering and synthetic biology. *Biotechnol. Bioprocess Eng.* 24, 1–18.

Schönfeld, P., and Wojtczak, L. (2016). Short-and medium-chain fatty acids in energy metabolism: the cellular perspective. *J. Lipid Res.* 57, 943–954. doi: 10.1194/jlr.r067629

Sgobba, E., and Wendisch, V. F. (2020). Synthetic microbial consortia for small molecule production. *Curr. Opin. Biotechnol.* 62, 72–79. doi: 10.1016/j.copbio.2019.09.011

Shabestary, K., and Hudson, E. P. (2016). Computational metabolic engineering strategies for growth-coupled biofuel production by synechocystis. *Metab. Eng. Commun.* 3, 216–226. doi: 10.1016/j.meteno.2016.07.003

Sheppard, M. J., Kunjapur, A. M., and Prather, K. L. (2016). Modular and selective biosynthesis of gasoline-range alkanes. *Metab. Eng.* 33, 28–40. doi: 10.1016/j.ymben.2015.10.010

Shin, J. H., Kim, H. U., Kim, D. I., and Lee, S. Y. (2013). Production of bulk chemicals via novel metabolic pathways in microorganisms. *Biotechnol. Adv.* 31, 925–935. doi: 10.1016/j.biotechadv.2012.12.008

Shrestha, K., and Yamamoto, S. (2018). *Tractor Hydraulic Fluid Compositions. Google Patents*. Wayzata, MN: Cargill, Inc.

Singh, N., and Choudhury, B. (2018). Potential of *Lentibacillus* sp. NS12IITR for production of lipids with enriched branched-chain fatty acids for improving biodiesel properties along with hydrocarbon co-production. *Extremophiles* 22, 865–875. doi: 10.1007/s00792-018-1043-6

Sorigué, D., Legeret, B., Cuiné, S., Blangy, S., Moulin, S., Billon, E., et al. (2017). An algal photoenzyme converts fatty acids to hydrocarbons. *Science* 357, 903–907. doi: 10.1126/science.aan6349

Sperry, J. S., Venturas, M. D., Todd, H. N., Trugman, A. T., Anderegg, W. R., Wang, Y., et al. (2019). The impact of rising CO2 and acclimation on the response of US forests to global warming. *Proc. Natl. Acad. Sci. U.S.A.* 116, 25734–25744. doi: 10.1073/pnas.1913072116

Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., Mcclure, A., et al. (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463:559. doi: 10.1038/nature08721

Suastes-Rivas, J. K., Hernández-Altamirano, R., Mena-Cervantes, V. Y., Valdez-Ojeda, R., Toledano-Thompson, T., Tovar-Gálvez, L. R., et al. (2019). Efficient production of fatty acid methyl esters by a wastewater-isolated microalgae-yeast co-culture. *Environ. Sci. Pollut. Res.* 27, 28490–28499. doi: 10.1007/s11356-019-07286-1

Tabei, Y., Yamanishi, Y., and Kotera, M. (2016). Simultaneous prediction of enzyme orthologs from chemical transformation patterns for de novo metabolic pathway reconstruction. *Bioinformatics* 32, i278–i287.

Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., and Lu, X. (2011). Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab. Eng.* 13, 169–176. doi: 10.1016/j.ymben.2011.01.001

Tang, L. (2019). Microbial interactions. *Nat. Methods* 16:19.

Tee, T. W., Chowdhury, A., Maranas, C. D., and Shanks, J. V. (2014). Systems metabolic engineering design: fatty acid production as an emerging case study. *Biotechnol. Bioeng.* 111, 849–857. doi: 10.1002/bit.25205

Teo, W. S., Ling, H., Yu, A.-Q., and Chang, M. W. (2015). Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid short-and branched-chain alkyl esters biodiesel. *Biotechnol. Biofuels* 8:177.

Teusink, B., and Smid, E. J. (2006). Modelling strategies for the industrial exploitation of lactic acid bacteria. *Nat. Rev. Microbiol.* 4, 46–56. doi: 10.1038/nrmicro1319

Thakur, I. S., Kumar, M., Varjani, S. J., Wu, Y., Gnansounou, E., and Ravindran, S. (2018). Sequestration and utilization of carbon dioxide by chemical and biological methods for biofuels and biomaterials by chemoautotrophs: opportunities and challenges. *Bioresour. Technol.* 256, 478–490. doi: 10.1016/j.biortech.2018.02.039

Thompson, R. A., and Trinh, C. T. (2017). Overflow metabolism and growth cessation in Clostridium thermocellum DSM1313 during high cellulose loading fermentations. *Biotechnol. Bioeng.* 114, 2592–2604. doi: 10.1002/bit.26374

Trinh, C. T., Liu, Y., and Conner, D. J. (2015). Rational design of efficient modular cells. *Metab. Eng.* 32, 220–231. doi: 10.1016/j.ymben.2015.10.005

Trinh, C. T., Wlaschin, A., and Srienc, F. (2009). Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol.* 81:813. doi: 10.1007/s00253-008-1770-1

Tsuge, T., Taguchi, K., and Doi, Y. (2003). Molecular characterization and properties of (R)-specific enoyl-CoA hydratases from *Pseudomonas aeruginosa*:

metabolic tools for synthesis of polyhydroxyalkanoates via fatty acid ß-oxidation. *Int. J. Biol. Macromol.* 31, 195–205. doi: 10.1016/s0141-8130(02)00082-x

Venayak, N., Von Kamp, A., Klamt, S., and Mahadevan, R. (2018). MoVE identifies metabolic valves to switch between phenotypic states. *Nat. Commun.* 9, 1–9.

Verma, M. L., Kishor, K., Sharma, D., Kumar, S. and Sharma, K. D. (2020). "Microbial production of omega-3 polyunsaturated fatty acids," in *Biotechnological Production of Bioactive Compounds*, eds M. L. Verma, and A. Chandel, (Amsterdam: Elsevier), 293–326. doi: 10.1016/b978-0-444-64323-0.00010-2

Vick, J. E., Clomburg, J. M., Blankschien, M. D., Chou, A., Kim, S., and Gonzalez, R. (2015). *Escherichia coli* enoyl-acyl carrier protein reductase (FabI) supports efficient operation of a functional reversal of the β-oxidation cycle. *Appl. Environ. Microbiol.* 81, 1406–1416. doi: 10.1128/aem.03521-14

Wang, J.-P., Wu, L.-X., Xu, F., Lv, J., Jin, H.-J., and Chen, S.-F. (2010). Metabolic engineering for ethylene production by inserting the ethylene-forming enzyme gene (efe) at the 16S rDNA sites of *Pseudomonas* putida KT2440. *Bioresour. Technol.* 101, 6404–6409. doi: 10.1016/j.biortech.2010.03.030

Wang, W., Liu, X., and Lu, X. (2013). Engineering cyanobacteria to improve photosynthetic production of alka (e) nes. *Biotechnol. Biofuels* 6:69. doi: 10.1186/1754-6834-6-69

Wenning, L., Yu, T., David, F., Nielsen, J., and Siewers, V. (2017). Establishing very long-chain fatty alcohol and wax ester biosynthesis in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 114, 1025–1035. doi: 10.1002/bit.26220

Wu, J., Wang, Z., Duan, X., Zhou, P., Liu, P., Pang, Z., et al. (2019). Construction of artificial micro-aerobic metabolism for energy-and carbon-efficient synthesis of medium chain fatty acids in *Escherichia coli*. *Metab. Eng.* 53, 1–13. doi: 10.1016/j.ymben.2019.01.006

Wu, J., Zhang, X., Xia, X., and Dong, M. (2017). A systematic optimization of medium chain fatty acid biosynthesis via the reverse beta-oxidation cycle in *Escherichia coli*. *Metab. Eng.* 41, 115–124. doi: 10.1016/j.ymben.2017.03.012

Wu, X., Zha, J., and Koffas, M. A. (2020). Microbial production of bioactive chemicals for human health. *Curr. Opin. Food Sci.* 32, 9–16. doi: 10.1016/j.cofs.2019.12.007

Xiao, Y., Bowen, C. H., Liu, D., and Zhang, F. (2016). Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. *Nat. Chem. Biol.* 12:339. doi: 10.1038/nchembio.2046

Xin, F.-H., Zhang, Y., Xue, S.-J., Chi, Z., Liu, G.-L., Hu, Z., et al. (2017). Heavy oils (mainly alkanes) over-production from inulin by *Aureobasidium melanogenum* 9-1 and its transformant 88 carrying an inulinase gene. *Renew. Energy* 105, 561–568. doi: 10.1016/j.renene.2017.01.004

Xu, P., Qiao, K., Ahn, W. S., and Stephanopoulos, G. (2016). Engineering Yarrowia lipolytica as a platform for synthesis of drop-in transportation fuels and oleochemicals. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10848–10853. doi: 10.1073/pnas.1607295113

Yang, D., Park, S. Y., Park, Y. S., Eun, H., and Lee, S. Y. (2020). Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol.* 38, 745–765. doi: 10.1016/j.tibtech.2019.11.007

Yang, L., Ebrahim, A., Lloyd, C. J., Saunders, M. A., and Palsson, B. O. (2019). DynamicME: dynamic simulation and refinement of integrated models of metabolism and protein expression. *BMC Syst. Biol.* 13:2. doi: 10.1186/s12918-018-0675-6

Yu, T., Zhou, Y. J., Huang, M., Liu, Q., Pereira, R., David, F., et al. (2018). Reprogramming yeast metabolism from alcoholic fermentation to lipogenesis. *Cell* 174, 1549.e14–1558.e14.

Yunus, I. S., and Jones, P. R. (2018). Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols. *Metab. Eng.* 49, 59–68. doi: 10.1016/j.ymben.2018.07.015

Zerhusen, C., Bollmann, T., Gödderz, A., Fleischer, P., Glüsen, B., and Schörken, U. (2020). Microbial synthesis of nonionic long-chain sophorolipid emulsifiers obtained from fatty alcohol and mixed lipid feeding. *Eur. J. Lipid Sci. Technol.* 122:1900110. doi: 10.1002/ejlt.201900110

Zhang, F., Carothers, J. M., and Keasling, J. D. (2012). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30:354. doi: 10.1038/nbt.2149

Zhang, L.-S., Liang, S., Zong, M.-H., Yang, J.-G., and Lou, W.-Y. (2020). Microbial synthesis of functional odd-chain fatty acids: a review. *World J. Microbiol. Biotechnol.* 36, 1–9.

Zheng, Y.-N., Li, L.-L., Liu, Q., Yang, J.-M., Wang, X.-W., Liu, W., et al. (2012). Optimization of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered *Escherichia coli*. *Microb. Cell Fact.* 11:65. doi: 10.1186/1475-2859-11-65

Zhou, A., Zhang, C., Li, Y., Li, S., and Yin, P. (2019). Effect of hydrogen peroxide additive on the combustion and emission characteristics of an n-butanol homogeneous charge compression ignition engine. *Energy* 169, 572–579. doi: 10.1016/j.energy.2018.12.076

Zhou, Y. J., Buijs, N. A., Zhu, Z., Qin, J., Siewers, V., and Nielsen, J. (2016). Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nat. Commun.* 7, 1–9.

Zhou, Y. J., Hu, Y., Zhu, Z., Siewers, V., and Nielsen, J. (2018). Engineering 1-alkene biosynthesis and secretion by dynamic regulation in yeast. *ACS Synth. Biol.* 7, 584–590. doi: 10.1021/acssynbio.7b00338

Zhuang, K., Yang, L., Cluett, W. R., and Mahadevan, R. (2013). Dynamic strain scanning optimization: an efficient strain design strategy for balanced yield, titer, and productivity. DySScO strategy for strain design. *BMC Biotechnol.* 13:8. doi: 10.1186%2F1472-6750-13-8

Zu Belzen, J. U., Bürgel, T., Holderbach, S., Bubeck, F., Adam, L., Gandor, C., et al. (2019). Leveraging implicit knowledge in neural networks for functional dissection and engineering of proteins. *Nat. Mach. Intellig.* 1, 225–235. doi: 10.1038/s42256-019-0049-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 04 February 2021
doi: 10.3389/fbioe.2020.621127

High-Level Patchoulol Biosynthesis in *Artemisia annua* L.

*Xueqing Fu[1†], Fangyuan Zhang[1,2†], Yanan Ma[1], Danial Hassani[1], Bowen Peng[1], Qifang Pan[1], Yuhua Zhang[3], Zhongxiang Deng[3], Wenbo Liu[3], Jixiu Zhang[3], Lei Han[3], Dongfang Chen[3], Jingya Zhao[1], Ling Li[1], Xiaofen Sun[1] and Kexuan Tang[1]**

[1] Joint International Research Laboratory of Metabolic & Developmental Sciences, Key Laboratory of Urban Agriculture (South) Ministry of Agriculture, Plant Biotechnology Research Center, Fudan-Shanghai Jiaotong University (SJTU)-Nottingham Plant Biotechnology R&D Center, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China, [2] Southwest University-Tibet Agriculture and Animal Husbandry College (SWU-TAAHC) Medicinal Plant Joint R&D Centre, School of Life Sciences, Southwest University, Chongqing, China, [3] Corporate R&D Division, Firmenich Aromatics (China) Co. Ltd., Shanghai, China

Terpenes constitute the largest class of secondary metabolites in plants. Some terpenes are essential for plant growth and development, membrane components, and photosynthesis. Terpenes are also economically useful for industry, agriculture, and pharmaceuticals. However, there is very low content of most terpenes in microbes and plants. Chemical or microbial synthesis of terpenes are often costly. Plants have the elaborate and economic biosynthetic way of producing high-value terpenes through photosynthesis. Here we engineered the heterogenous sesquiterpenoid patchoulol production in *A. annua*. When using a strong promoter such as 35S to over express the avian farnesyl diphosphate synthase gene and patchoulol synthase gene, the highest content of patchoulol was 52.58 μg/g DW in transgenic plants. When altering the subcellular location of the introduced sesquiterpene synthetase via a signal peptide, the accumulation of patchoulol was observably increased to 273 μg/g DW. This case demonstrates that *A. annua* plant with glandular trichomes is a useful platform for synthetic biology studies.

Keywords: patchoulol, *Artemisia annua* L., synthetic biology, sesquiterpenoids, terpenes

OPEN ACCESS

Edited by:
Jingwen Zhou,
Jiangnan University, China

Reviewed by:
Xiaoya Chen,
Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences (CAS), China
Sheng-Hong Li,
Kunming Institute of Botany, China

***Correspondence:**
Kexuan Tang
kxtang@sjtu.edu.cn;
kxtang1@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *25 October 2020*
Accepted: *30 December 2020*
Published: *04 February 2021*

Citation:
Fu X, Zhang F, Ma Y, Hassani D, Peng B, Pan Q, Zhang Y, Deng Z, Liu W, Zhang J, Han L, Chen D, Zhao J, Li L, Sun X and Tang K (2021) High-Level Patchoulol Biosynthesis in Artemisia annua L. Front. Bioeng. Biotechnol. 8:621127. doi: 10.3389/fbioe.2020.621127

INTRODUCTION

Plants synthesize and secrete a good deal of secondary metabolites, some of which are considerable, economically in industry, agriculture, and pharmaceuticals (Balandrin et al., 1985; Pichersky and Gershenzon, 2002). Terpenes comprise the largest class of secondary metabolites in plants (Kappers et al., 2008). Many of them, such as phytohormones (abscisic acid, brassinosteroid, and gibberellin), sterols and carotenoid pigments, play critical roles in plant growth, development, membrane components, and photosynthesis (Bohlmann and Keeling, 2008). In addition, the majority of plant terpenes are involved in the interaction of plant with the environment and other organisms (Gershenzon and Dudareva, 2007). For instance, some terpenes bear antibacterial and antifungal activity (Rastogi et al., 1998; Lunde and Kubo, 2000). They can also hold the protective role in plants defense system against insects, mollusks, fish and nematodes (Lorimer et al., 1996; Ito et al., 1997; Laurent et al., 2003; Quintana et al., 2003). Terpenes can sometimes act a tool of communication among organisms. For instance, when the predators attack aphids, they normally release a kind of terpenoid, (*E*)-β-farnesene, as an alarm pheromone, to disperse and leave the host. Besides, (*E*)-β-farnesene is also released to attract natural enemies of aphids at the same time in plants (Hardie and Minks, 1999; Kunert et al., 2005).

On the basis of the number of five-carbon (isoprene) units, terpenoids are classified into hemiterpenes (half-terpenes), monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), triterpenes (C_{30}), tetraterpenes (C_{40}), polyterpenes, and meroterpenes (Croteau, 2000). In plants, terpenoids are synthesized from two precursors, dimethylallyl diphosphate (DMADP) and isopentenyl diphosphate provided by MVA (mevalonate) and MEP (2-C-methyl-D-erythritol-4-phosphate) pathway (Rohmer et al., 1996; Rohmer, 1999; Lange et al., 2000). In plants, terpenoids are synthesized from two precursors, dimethylallyl diphosphate (DMADP) and isopentenyl diphosphate provided by MVA (mevalonate) and MEP (2-C-methyl-D-erythritol-4-phosphate) pathway (Daviet and Schalk, 2010). Afterwards terpene synthases catalyze the cyclization of GPP (geranyl diphosphate), FPP (farnesyl diphosphate) and GGPP (geranylgeranyl diphosphate) to generate the carbon skeletons of terpenoids. Finally, the enzymes further modify the terpene backbone to synthesize plenty of natural terpene derivatives, for example, cytochrome P450 monooxygenases (P450) (Cheng et al., 2007) (**Figure 1**).

Overexploitation and wasteful consumption of natural resources for high-value terpenes compounds, may drive the species to extinction and alter the environment. For instance, *Taxus chinensis* is famous for Taxol, an effective anti-cancer drug (Ru et al., 2006). However, *T. chinensis* is facing extinction because of deforestation (Zhang and Ru, 2010). To date, many efforts have been made to manipulate terpene metabolism in microorganism (Carter et al., 2003; Martin et al., 2003), fungus (Jackson et al., 2003; Ro et al., 2006; Westfall et al., 2012; Paddon et al., 2013), and plants (Wu et al., 2006; Farhi et al., 2011; Zhan et al., 2014; Wang et al., 2016) to synthetically produce more high-value chemicals.

Engineering terpene metabolism in plants is an innovative and attractive strategy to provide high-value terpenes. Plants have the elaborate biosynthetic ability and a cheaper way of using photosynthesis to produce high-value terpenes (Wu et al., 2006). For example, the metabolic engineered tobaccos stably transformed with the deregulated 3-hydroxy-3-methylglutaryl-coenzyme A reductase gene (*tHMG*), *ADS*, *CYP71AV1*, *CPR*, and artemisinic aldehyde reductase gene (*DBR2*), produced artemisinin, although the content in the transgenic tobaccos was much lower than that in *A. annua* plants (Farhi et al., 2011). In another study, the transgenic tobacco accumulated high-levels of terpenes, containing patchoulol, amorpha-4,11-diene, and limonene, via overexpressing farnesyl diphosphate synthase gene from avian and the terpene synthase gene to control the carbon flux (Wu et al., 2006). The results from these researches indicate the necessity of deeper exploring and demonstration of terpenes biosynthesis pathway genes with respect to over-expression strategies (Wu et al., 2006). In recent years, various strategies for engineering of triterpene squalene metabolism in tobacco have been developed. For instance, co-expression of farnesyl diphosphate synthase gene *FPS* from avian and yeast squalene synthase gene *SQS*, driven by trichome-specific gene promoter in the chloroplast, resulted in the accumulation of squalene to a high level in transgenic tobacco (Wu et al., 2012). These findings suggested that the accumulation of high-value terpenes could be observably elevated by directing the biosynthesis to other subcellular compartments.

Patchoulol, a volatile sesquiterpenoid isolated from leaves of *Pogostemon cablin* plants, is an important ingredient in fragrance products like perfumes, soaps and cosmetics. Plant patchouli is the only commercial source of this compound (Srikrishna and Satyanarayana, 2005). Limited natural resources lead to the fluctuation in the price of patchoulol between 30 and 200 US dollar/kg (Zhan et al., 2014). A sesquiterpene cyclase enzyme, patchoulol synthase was identified to catalyze FPP to form patchoulol in patchouli plants (Deguerry et al., 2006). With the development of molecular and synthetic biology, the engineered tobacco could produce 0.030 mg/g fresh weight (FW) patchoulol (Wu et al., 2006). Biotechnological production of patchoulol has been carried out in *Physcomitrella patens*, and the highest yield of patchoulol was 1.34 mg/g dry weight (Zhan et al., 2014). Besides, Albertsen et al. reported that expression of FPPS of yeast fused with PTS from *P. cablin* in *Saccharomyces cerevisiae* increased the production of patchoulol compared with the accumulation produced by PTS (Albertsen et al., 2011). Subsequently, several strategies were adopted to increase patchoulol content in *S. cerevisiae*. The shaken flask contained 59.2 $\pm$ 0.7 mg/L patchoulol, and a final production was 466.8 $\pm$ 12.3 mg/L (20.5 $\pm$ 0.5 mg/g dry cell weight) after fermentation optimization (Ma et al., 2019).

Here we demonstrated, an engineered heterogenous sesquiterpenoid patchoulol production in *A. annua*. The highest content of patchoulol was 52.58 μg/g dry weight in the transgenic plants by overexpressing farnesyl diphosphate synthase gene and patchoulol synthase gene. Furthermore, the accumulation of patchoulol was increased to 273 μg/g dry weight by altering the subcellular location of the introduced sesquiterpene synthetase expression.

RESULT

Patchoulol Was Produced in *FPS*+*PTS*-overexpressing Transgenic *A. annua* Plants

To engineer the heterogenous sesquiterpenoid patchoulol production in *A. annua*, patchoulol synthase gene (*PTS*) from *P. cablin* and farnesyl diphosphate synthase gene (*FPS*) from an avian were chosen (Tarshis et al., 1994, Deguerry et al., 2006). Despite the cloned and identified *FPS* from *A. annua*, the avian *FPS* was observed not to be operated by the transcriptional or post-translational regulatory mechanisms in plants (Wu et al., 2006). To generate *FPS*+*PTS*-overexpressing transgenic *A. annua* plants, both *PTS* and *FPS* were transformed into *A. annua* plants via *A. tumefaciens* EHA105. Analysis of PCR showed that 25 independent lines were obtained. We then performed qRT-PCR to test the expression of both *FPS* and *PTS* genes in the transgenic lines. Among these, six transgenic lines showed a combination of high levels of *FPS* and *PTS* overexpression (**Figure 2A**). The patchoulol content in leaves was measured by GC-MS and the sesquiterpene alcohol patchoulol was

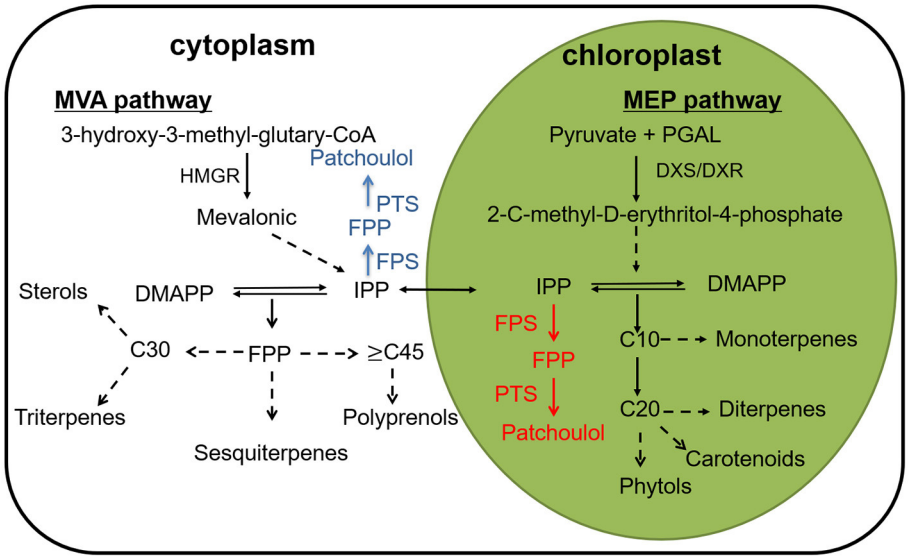


FIGURE 1 | A depiction of the terpene biosynthesis, along with a conceptualization for patchoulol biosynthesis to the cytoplasm (blue) and to the chloroplast (red) compartments. HMGR, 3-hydroxy-3- methylglutaryl coenzyme A reductase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose5-phosphate reductase; FPS, farnesyl diphosphate synthase; FPP, farnesyl pyrophosphate; PTS, patchoulol synthase.

successfully identified in *FPS*+*PTS*-overexpressing transgenic lines (**Supplementary Figure 1**). Quantification of patchoulol in transgenic lines showed that the six lines produced 23.51–52.58 μg patchoulol/g dry weight (**Figure 2B**), which was twice the transgenic tobaccos (Wu et al., 2006). The expressions of both *FPS* and *PTS* were higher than other transgenic lines in FPS+PTS-6 and 10, so they accumulated the highest patchoulol content.

Repressing the Artemisinin Biosynthetic Pathway Is Useful for Getting a High Yield of Patchoulol

Blocking the competitive biosynthetic pathways is an effective approach for increasing the sesquiterpene content. Former studies have reported, diverse array of sesquiterpene synthases in *A. annua*. For sesquiterpene biosynthesis in *A. annua*, FPSs convert the common precursor FPP into an array of cyclized products, such as amorpha-4,11-diene via the action of ADS (Bouwmeester et al., 1999), β-caryophyllene by β-caryophyllene synthase (CPS) (Cai et al., 2002), β-farnesene by β-farnesene synthase (BFS) (Picaud et al., 2005), germacrene A by germacrene A synthase (GAS) (Bertea et al., 2006) and epi-cedrol by epi-cedrol synthase (ECS) (Mercke et al., 1999) respectively. Artemisinin, an important sesquiterpene isolated from *A. annua*, is ∼0.1–1% of the dry weight in this plant, and artemisinin biosynthesis occurs in the cytosol in *A. annua* (Wallaart et al., 2001; Abdin et al., 2003). We speculated that silencing *ADS* gene competing for FPS with TPS by RNAi technology would enhance the patchoulol content in transgenic lines. To assess whether blocking the artemisinin biosynthetic pathway could increase the yield of patchoulol, the *FPS*+ *PTS* and *ADS*-RNAi constructs were co-transformed into *A. annua* plants resulting in 18 independent transgenic lines for further analysis. In transgenic lines, ADSi +FPS+ PTS −1, −5, −12, −18, and −23 had a combination of high levels of *FPS* and *PTS* transcripts (**Figure 3C**), as well as low *ADS* transcript level (**Figure 3A**). Compared with the wild-type, the artemisinin content was reduced to 42–55% in transgenic lines (**Figure 3B**). The results from patchoulol content GC-MS analysis showed that FPS+PTS+ADSi lines produced 41.13–83.23 μg patchoulol/g DW (**Figure 3D**). Consistent with the hypothesis, blocking the competing pathway could be an applicable approach for getting higher yield of patchoulol.

The Localization of Heterologous Proteins

Many efforts had been made to introduce the terpene synthases into the cellular compartments, in which the terpene is naturally synthesized, to compete for substrates or overcome prospective rate-limiting steps. For instance, tobacco was used to produce heterologous patchoulol, in which, both *FPS* and *PTS* were expressed in tobacco, and the final yield was about 0.3 μg/g FW (Wu et al., 2006). Furthermore, higher level of patchoulol (30 μg/g FW) was observed in transgenic tobacco when chloroplast-targeting signal sequence from the signal peptide of *Arabidopsis* RUBISCO small unit (tpFPS and tpPTS) was fused with the amino terminus of both FPS and PTS (Lee et al., 2006; Wu et al., 2006). To confirm the localization of FPS and PTS, the full-length of *FPS* and *PTS* were fused with GFP (Green Fluorescent Protein), respectively. The recombinant plasmids were transiently expressed in tobacco leaves. The results showed that both FPS-GFP and PTS-GFP fusion proteins were separated from the fluorescence of chloroplasts (**Figures 4A,B**). When the chloroplast-targeting signal sequence (TP) was targeted to the

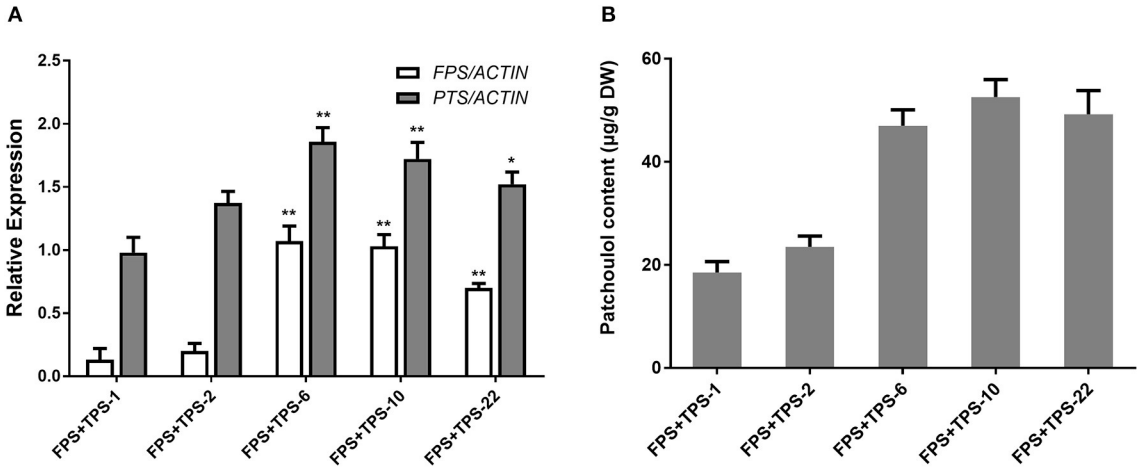


FIGURE 2 | Engineering patchoulol biosynthesis in the cytoplasm of *A. annua* leaves. **(A)** Relative expression of *FPS* and *PTS* in *FPS+PTS* transgenic *A. annua* lines. **(B)** The patchoulol content in *FPS+PTS* transgenic *A. annua* lines. *ACTIN* was used as internal control. T0 transgenic lines were used for analysis. The error bars represent the means ± SD from three biological replicates. All data represent the means ± SD of three replicates. $^{**}P < 0.05$, $^{*}P < 0.01$, student's *t*-test.

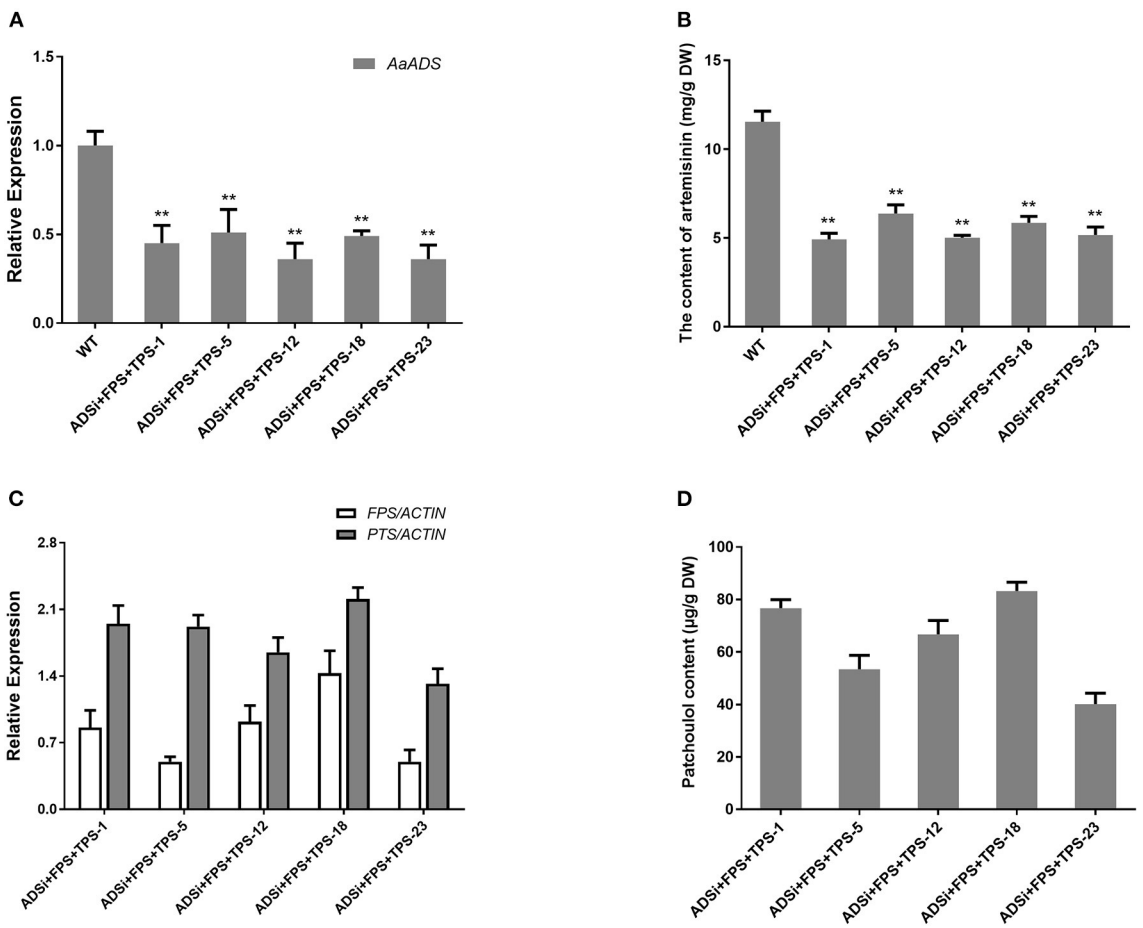


FIGURE 3 | Blocking the artemisinin biosynthesis increased patchoulol content in *ADSi+ FPS+PTS* transgenic *A. annua* plants. **(A)** Relative expression of *ADS* in *ADSi+ FPS+PTS* transgenic *A. annua* lines. **(B)** The artemisinin content in *ADSi+ FPS+PTS* transgenic *A. annua* lines. **(C)** Relative expression of *FPS* and *PTS* in *ADSi+ FPS+PTS* transgenic *A. annua* lines. **(D)** The patchoulol content in *ADSi+ FPS+PTS* transgenic *A. annua* lines. T0 transgenic lines were used for analysis. *ACTIN* was used as internal control. The error bars represent the means ± SD from three biological replicates. All data represent the means ± SD of three replicates. $^{**}P < 0.05$, $^{*}P < 0.01$, student's *t*-test.

amino terminus of both FPS and PTS, tpFPS-GFP and tpPTS-GFP fusion proteins were completely matched the fluorescence of chloroplasts (**Figures 4A,B**).

Engineering the Patchoulol Biosynthesis in the Chloroplast Compartment Enhanced the Accumulation of Patchoulol

When a particular terpene was produced in a certain compartment that this biosynthesis could not normally occur by diverting carbon flux at earlier intermediates, the high level of target product was obtained. For example, engineering of six genes encoding cytoplasmic MVA pathway, to chloroplast increased the levels of mevalonate, carotenoids, sterols, and squalene, suggesting the possible enhancement of overall terpene biosynthesis, despite of its organelles where it takes place (Kumar et al., 2011). Therefore, both FPS and PTS were further targeted into chloroplast via TP (tpFPS and tpPTS) (Lee et al., 2006) resulting in the more than 25 independent transgenic lines, which were further found to have higher expression levels of *tpFPS* and *tpPTS* (**Figure 5A**). The results from patchoulol content measurement by GC-MS revealed that, co-expression of *tpFPS* and *tpTPS* targeting the chloroplast compartment, could significantly enhance the patchoulol accumulation up to 273 μg/g DW (91 μg/g FW) (**Figure 5B**), which was 5–11-folds higher than those levels synthesized in the cytosol. The transgenic lines exhibited normal growth characteristics (**Supplementary Figure 2**).

DISCUSSION

Recently great efforts have been made to engineer terpene metabolism in plant (Lewinsohn et al., 2001; Wu et al., 2006, 2012; Zhan et al., 2014). *A. annua* is a traditional Chinese medicinal plant and is famous for artemisinin. A stable and efficient *Agrobacterium* mediated transformation system of *A. annua* has been established. Besides, there are a large amount of trichomes (glandular trichomes and T-shaped trichomes) on the leaves in *A. annua*, where large quantities of terpenes are synthetized and stored to protect plants against insects, pathogens, and herbivores (Wagner, 1991; Duke and Paul, 1993; Pichersky and Gershenzon, 2002). Recently the biochemistry of trichomes has been studied in *A. annua*. Numerous information from former studies on trichome development, provides a great opportunity for engineering of terpene biosynthesis in the specific target cellular compartment.

The Advantage of Using *A. annua* as Platform for Synthetic Botany

In this work, we could apply and develop a novel approach and technique to engineer a cultivar of *A. annua*, with 273 μg/g DW patchoulol production with no any alteration of artemisinin biosynthesis. This strategy enables us to improve the economic value of medicinal plants. With the development of synthetic botany, many approaches have been made to produce valuable secondary metabolites. For instance, biosynthesis of β-carotene and anthocyanin in rice and the production of artemisinin in tobacco. However, due to the lack of specific storage cells in target plants, the heterogeneous biosynthesis of volatility chemicals, including mono- or sesquiterpene, still remain a big challenge in synthetic biology (Houshyani et al., 2013).

For instance, Wu et al. have constructed the patchoulol biosynthesis pathway in the tobacco plastid with the same strategy used in this study. However, the patchoulol content in the best-performing transgenic tobacco line was reported to be 30 μg/g FW, while the highest patchoulol content in transgenic *A. annua* in our study, reached to the level of 273 μg/g DW (91 μg/g FW) which could be possibly resulted from the possession of numerous glandular trichomes on the epidermal cells of *A. annua* leaves, where the accumulation and storage of artemisinin and lots of other mono- or sesquiterpene takes place. Furthermore, there is no significant difference in the artemisinin contents between patchoulol produced in transgenic and wild type *A. annua* (**Supplementary Figure 3**), suggesting the transgenic *A. annua* to be a reliable sources of anti-malaria agent, artemisinin and flavor component patchoulol production. Beside the above evidences, the establishment of an efficient *Agrobacterium* mediated transformation system in *A. annu,* gives a noticeable credit to it for being a worthy candidate in biosynthetic biology.

The Enzyme Targeted Cellular Compartment Is Crucial for Patchoulol Biosynthesis in *A. annua*

It has been well-studied that the terpenoids are biosynthesized from two independent compartmentally pathways: the MVA and MEP pathways. The MEP pathway, located in plastid, is dominantly responsible for the biosynthesis of mono- and diterpenes. The cytoplasm located MVA pathway is mainly responsible for the biosynthesis of sesquiterpenes. In *A. annua*, the artemisinin biosynthesis depends on cytoplasm directed MVA pathway (Newman and Chappell, 1999; Weathers et al., 2006). However, only about 0.001% dry weight of dihydroartemisinic alcohol (the artemisinin precursor) is reported to be produced in tobacco when expressing the artemisinin biosynthetic genes in the cytoplasm. On the contrast, the artemisinic acid accumulation reached to a maximum of about 0.004% when transferring the entire artemisinin biosynthetic genes into the chloroplast (Saxena et al., 2014). Wu et al. also reported that targeting the *PTS* to the plastid increased the patchoulol accumulation as well. In this study, targeting *PTS* into *A. annua* plastid produced higher patchoulol compared to its expression in cytoplasm. These results suggested that targeting or expressing the enzymes in plastid might be a powerful tool for synthetic botany.

Pathway Block Is a Useful Method for Improving Patchoulol Accumulation

Besides higher production and accumulation of patchoulol through targeting the *PTS* into *A. annua* plastid, we also investigated the influence of pathway blockage toward the patchoulol biosynthesis in cytoplasm. Obviously, the risk will be risen when the metabolic pathway is composed of more than

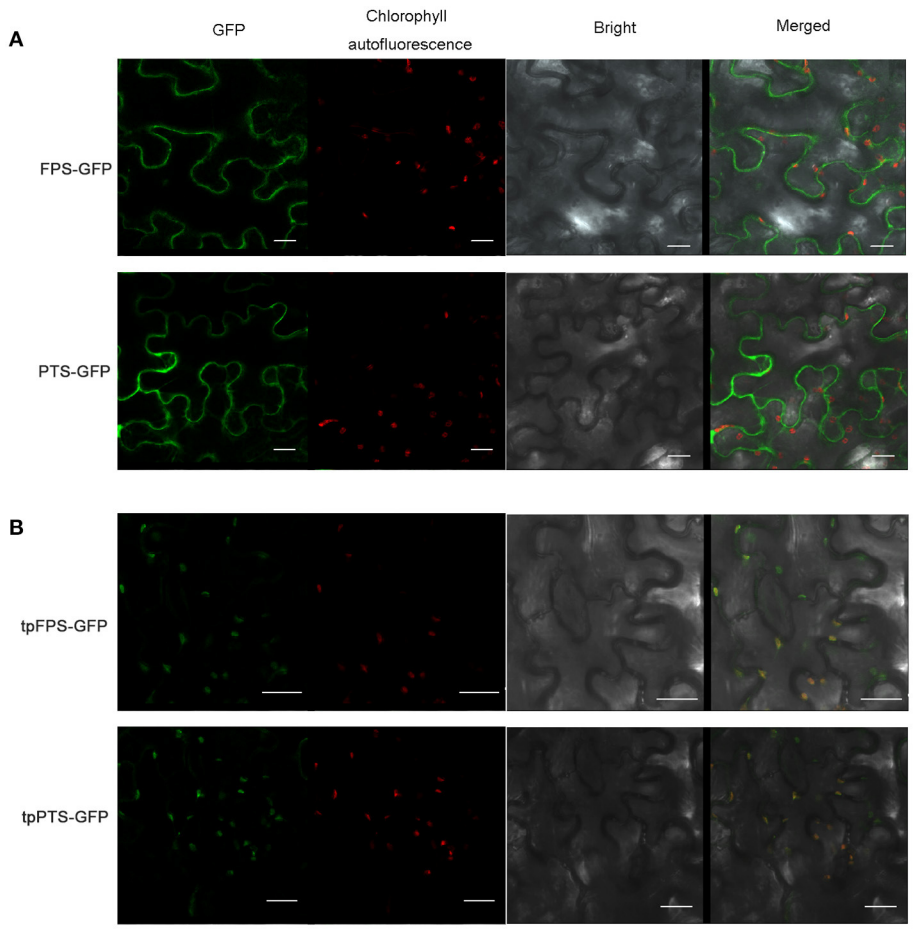


FIGURE 4 | The subcellular localization of heterologous proteins. **(A)** Subcellular localization of FPS-GFP and PTS-GFP in tobacco leaf epidermal cells. **(B)** Subcellular localization of tpFPS-GFP and tpPTS-GFP in tobacco leaf epidermal cells. GFP: green fluorescent protein, Bars = 20 μm.

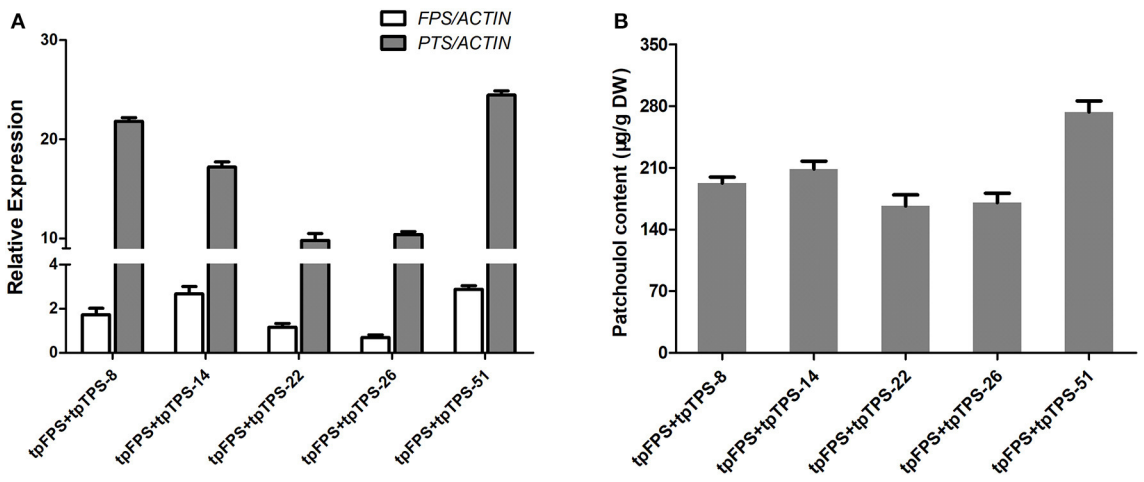


FIGURE 5 | Engineering the patchoulol biosynthesis in the chloroplast compartment increased the patchoulol content in *tpFPS+tpPTS* transgenic *A. annua* plants. **(A)** Relative expression of *FPS* and *PTS* in *tpFPS+tpPTS* transgenic *A. annua* lines. **(B)** The patchoulol content in *tpFPS+tpPTS* transgenic *A. annua* lines. *ACTIN* was used as internal control. The error bars represent the means ± SD from three biological replicates.

two genes. For improving the patchoulol yield in cytoplasm, we blocked the artemisinin biosynthesis through RNA interference for *ADS* gene. The results showed that the patchoulol content had significantly increased in the ADSi lines affirming the positive role of pathway blockage in elevation of patchoulol biosynthesis. However, the patchoulol content in ADSi+PTS+FPS lines was still lower than that in transgenic plastid targeted *PTS* in *A. annua*, which could be the result of incomplete blockage of artemisinin biosynthesis by RNA interference. In addition, the presence of other competitive pathway along with artemisinin biosynthesis could lead to the deficiency of RNAi for higher production of patchoulol.

MATERIALS AND METHODS

Plant Material and Growth Conditions

A. annua seeds originated from Chongqing province, were developed in Shanghai. *A. annua* plants were cultured in the greenhouse (16/8 h day/night, 25°C). Tobacco (*Nicotiana benthamiana*) was grown under the same conditions as *A. annua*.

Vectors Construction and the Transformation of *A. annua*

The patchoulol synthase gene (*PTS*) from *P. cablin* and farnesyl diphosphate synthase (*FPS*) from an avian were, respectively, inserted into the AscI/XhoI and SpeI/KpnI sites of the helper pTDUA vector. For the plastid-targeted expression, the transit peptide signal sequence of the RUBISCO protein in *Arabidopsis* (GenBank accession NM23202) was added to the 5' end of *FPS* and *PTS*, respectively. *PTS* driven by the cassava mosaic promoter and *FPS* driven by the 35S promoter were subsequently transferred to the pDONR vector through Gateway recombination reaction. The expression vectors were provided by Firmenich.

The 300 bp fragment of *AaADS* (GenBank accession AF138959) was cloned into the intermediate cloning vector pDONR, and transferred to the pHELLSGATE12 vector by LR recombination reaction (Invitrogen, Carlsbad, CA, USA). The information about the primers used are listed in **Supplementary Table 1**. The recombinant plasmids were transferred into *Agrobacterium tumefaciens* strain EHA105, then used to introduced into *A. annua* plants (Zhang et al., 2009). After 3–4 months, the regenerated plants were obtained.

Transcript Analysis of Terpene Synthase Genes

The first leaves were collected from 3 month-old *A. annua* plants for RNA extraction. Total RNA was extracted using the RNAprep Pure Plant Kit (Tiangen, Beijing, China). The first-strand cDNA for qRT-PCR was synthesized using the PrimeScript II first Strand cDNA Synthesis Kit (Takara, Shiga, Japan). RT-qPCR was performed using SuperReal PreMix Plus (Tiangen, Beijing, China). *β-ACTIN* was used as the reference gene. RT-qPCR was performed in a Roche LightCycler96 (Roche). According to the manufacturer's instructions, amplification was carried out using SYBR Green qPCR MasterMix (Takara, Shiga, Japan). The profile for SYBR Green qPCR was 95°C for 10 min, followed by 40 cycles of 95°C for 20 s, 55°C for 20 s, and 72°C for 20 s. The primers are represented in **Supplementary Table 1**. Three biological repeats were measured for each sample. The transcripts level was calculated using the $2^{-\Delta Ct}$ method. *ADS* expression was analyzed using $2^{-\Delta\Delta Ct}$ (Kilambi et al., 2013).

Subcellular Localization

To determine the subcellular localization of FPS and PTS proteins, the green fluorescent protein (GFP) was fused to the C-terminal domain of FPS, PTS, tpFPS, and tpPTS under the control of the CaMV35S promoter. The recombinant plasmids (pHB-FPS-GFP, pHB-PTS-GFP, pHB-tpFPS-GFP, and pHB-tpPTS-GFP) were, respectively, transferred into GV3101. The strains GV3101 harboring the recombinant plasmids were, respectively, co-transformed into *N. benthamiana* leaves with the strain containing p19 plasmid (Voinnet et al., 2003). GFP signals were observed by confocal laser microscopy (Leica TCS SP5-II) after 48 h incubation.

Quantification of Artemisinin by HPLC-ELSD

Quantification of artemisinin was performed as described previously (Zhang et al., 2009). Leaves from 3 month-old *A. annua* were collected, dried for 48 h at 50°C and pulverized into powder. 0.1 g dried-leaf powder was used for the ultrasonic extraction with 1 mL methanol for 30 min. Then the mixture was centrifuged at 12, 000 rpm for 5 min. The supernatants were filtered using filters (0.22 μm). The samples were analyzed by the HITACHI *2695* HPLC system coupled with a SANCO ELSD180 detector. The conditions were as follows: mobile phase, water/methanol (20:80, v/v); column, YMC-Pack ODS-A C18; flowrate, 1 ml/min. Artemisinin was set at 5.577 for artemisinin. The artemisinin standard was purchased from Sigma. Three biological repeats were measured for each sample.

Quantification of Patchoulol by GC-MS

Quantification of patchoulol was identified and quantified using gas chromatography and mass spectrometry (GC-MS). Five hundred milligram leaves were collected from 3 month-old *A. annua* and rapidly ground into powder in liquid nitrogen. The powder was used for the ultrasonic extraction with 3 ml ethyl acetate for 20 min. Then the samples were centrifuged at 5,000 g for 10 min. And the supernatants were filtered using filters (0.22 μm). Meanwhile, 500 mg leaves of *A. annua* were collected and dried at 60°C overnight. The weights were accurately measured for the calculation of the dry weight. Quantification was achieved based on the standard patchoulol (Aladdin, China). GC-MS analysis was performed on a GC-MS 7890/5975C (Agilent) according to the methods described previously (Zhan et al., 2014). Dodecane was used as the internal standards.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

XF and KT designed the research. XF and FZ performed the experiments and wrote the first draft of the manuscript. YM, BP, QP, ZD, WL, JZhang, JZhao, and XS analyzed the data. XF, KT, and DH drafted the manuscript. YZ, DC, and LL revised the manuscript. All authors approved the manuscript.

FUNDING

This work was funded by National Key R&D Program of China (2018YFA0900600), the National Science Foundation of China (18Z103150043), and SJTU Trans-med Awards Research (20190104).

ACKNOWLEDGMENTS

We thank Dr. Zhang Y. H. from Firmenich for providing the patchoulol and farnesyl diphosphate synthase genes and the pBDON vector. We also thank the Instrumental Analysis Center of the Shanghai Jiao Tong University for assistance with GC-MS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.621127/full#supplementary-material

REFERENCES

Abdin, M. Z., Israr, M., Rehman, R. U., and Jain, S. K. (2003). Artemisinin, a novel antimalarial drug: biochemical and molecular approaches for enhanced production. *Planta Med.* 69, 289–299. doi: 10.1055/s-2003-38871

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

Balandrin, M. F., Klocke, J. A., Wurtele, E. S., and Bollinger, W. H. (1985). Natural plant chemicals: sources of industrial and medicinal materials. *Science* 228,1154–1160. doi: 10.1126/science.3890182

Bertea, C. M., Voster, A., Verstappen, F. W., Maffei, M., Beekwilder, J., and Bouwmeester, H. J. (2006). Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Arch. Biochem. Biophys.* 448, 3–12. doi: 10.1016/j.abb.2006.02.026

Bohlmann, J., and Keeling, C. I. (2008). Terpenoid biomaterials. *Plant J.* 54, 656–669. doi: 10.1111/j.1365-313X.2008.03449.x

Bouwmeester, H. J., Wallaart, T. E., Janssen, M. H., van Loo, B., Jansen, B. J., Posthumus, M. A., et al. (1999). Amorpha-4, 11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* 52, 843–854. doi: 10.1016/S0031-9422(99)00206-X

Cai, Y., Jia, J. W., Crock, J., Lin, Z. X., Chen, X. Y., and Croteau, R. (2002). A cDNA clone for β-caryophyllene synthase from *Artemisia annua*. *Phytochemistry* 61, 523–529. doi: 10.1016/S0031-9422(02)00265-0

Carter, O. A., Peters, R. J., and Croteau, R. (2003). Monoterpene biosynthesis pathway construction in *Escherichia coli*. *Phytochemistry* 64, 425–433. doi: 10.1016/S0031-9422(03)00204-8

Cheng, A. X., Lou, Y. G., Mao, Y. B., Lu, S., Wang, L. J., and Chen, X. Y. (2007). Plant terpenoids: biosynthesis and ecological functions. *J. Integr. Plant Biol.* 49,179–186. doi: 10.1111/j.1744-7909.2007.00395.x

Croteau, R. (2000). "Natural products (secondary metabolites)," in *Biochemistry and Molecular Biology of Plants,* eds B. B. Bob, G. Wilhelm, and L. J. Russell (New York, NY; Hoboken, NJ: Wiley), 1250–1381.

Daviet, L., and Schalk, M. (2010). Biotechnology in plant essential oil production: progress and perspective in metabolic engineering of the terpene pathway. *Flavour Frag. J.* 25,123–127. doi: 10.1002/ffj.1981

Deguerry, F., Pastore, L., Wu, S. Q., Clark, A., Chappell, J., and Schalk, M. (2006). The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Arch. Biochem. Biophys.* 454,123–136. doi: 10.1016/j.abb.2006.08.006

Duke, S. O., and Paul, R. N. (1993). Development and fine structure of the glandular trichomes of *Artemisia annua* L. *Int. J. Plant Sci.* 154, 107–118.

Farhi, M., Marhevka, E., Ben-Ari, J., Algamas-Dimantov, A., Liang, Z., Zeevi, V., et al. (2011). Generation of the potent anti-malarial drug artemisinin in tobacco. *Nat. Biotechnol.* 29, 1072–1074. doi: 10.1038/nbt.2054

Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nat. Chem. Biol.* 3, 408–414. doi: 10.1038/nchembio.2007.5

Hardie, J., and Minks, A. K. (1999). Pheromones of non-lepidopteran insects associated with agricultural plants. *Austral. Entomol.* 39, 97–100. doi: 10.2307/3496420

Houshyani, B., Assareh, M., Busquets, A., Ferrer, A., Bouwmeester, H. J., and Kappers, I. F. (2013). Three-step pathway engineering results in more incidence rate and higher emission of nerolidol and improved attraction of *Diadegma semiclausum*. *Metab. Eng.* 15, 88–97. doi: 10.1016/j.ymben.2012.10.002

Ito, H., Muranaka, T., Mori, K., Jin, Z. X., and Yoshida, T. (1997). Dryofragin and aspidin PB, piscicidal components from *Dryopteris fragrans*. *Chem. Pharm. Bull.* 45, 1720–1722. doi: 10.1248/cpb.45.1720

Jackson, B. E., Hartwells, E. A., and Matsuda, S. P. (2003). Metabolic engineering to produce sesquiterpenes in yeast. *Org. Lett.* 5,1629–1632. doi: 10.1021/ol034231x

Kappers, I. F., Dicke, M., and Bouwmeester, H. J. (2008). "Terpenoids in plant signaling, chemical ecology," in *Wiley Encyclopedia of Chemical Biology,* eds T. P. Begley (New York, NY; Hoboken, NJ: Wiley, Inc.), 1–8.

Kilambi, H. V., Kumar, R., Sharma, R., and Sreelakshmi, Y. (2013). Chromoplast-specific carotenoid-associated protein appears to be important for enhanced accumulation of carotenoids in hp1 tomato fruits. *Plant Physiol.* 161:2085. doi: 10.1104/pp.112.212191

Kumar, S., Hahn, F. M., Baidoo, E., Kahlon, T. S., Wood, D. F., Mcmahan, C. M., et al. (2011). Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. *Metab. Eng.* 14,19–28. doi: 10.1016/j.ymben.2011.11.005

Kunert, G., Otto, S., Röse, U. S. R., Gershenzon, J., and Weisser, W. W. (2005). Alarm pheromone mediates production of winged dispersal morphs in aphids. *Ecol. Lett.* 8, 596–603. doi: 10.1111/j.1461-0248.2005.00754.x

Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13172–13177. doi: 10.1073/pnas.240454797

Laurent, P., Braekman, J. C., Daloze, D., and Pasteels, J. (2003). Biosynthesis of defensive compounds from beetles and ants. *Eur. J. Org. Chem.* 2003, 2733–2743. doi: 10.1002/ejoc.200300008

Lee, D. W., Lee, S., Lee, G. J., Lee, K. H., Kim, S., Cheong, G. W., et al. (2006). Functional characterization of sequence motifs in the transit peptide of arabidopsis small subunit of Rubisco. *Plant Physiol.* 140, 466–483. doi: 10.1104/pp.105.074575

Lewinsohn, E., Schalechet, F., Wilkinson, J., Matsui, K., Tadmor, Y., Nam, K., et al. (2001). Enhanced levels of the aroma and flavor compound *S*-linalool by metabolic engineering of the terpenoid pathway in tomato fruits. *Plant Physiol.* 127,1256–1265. doi: 10.1104/pp.010293

Lorimer, S. D., Perry, N. B., Foster, L. M., and Burgess, E. J. (1996). A nematode larval motility inhibition assay for screening plant extracts and natural products. *J. Agr. Food Chem.* 44, 2842–2845. doi: 10.1021/jf9602176

Lunde, C. S., and Kubo, I. (2000). Effect of polygodial on the mitochondrial ATPase of *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother.* 44, 1943–1953. doi: 10.1128/aac.44.7.1943-1953.2000

Ma, B., Liu, M., Li, Z. H., Tao, X., and Wang, F. Q. (2019). Significantly enhanced production of patchoulol in metabolically engineered *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* 67, 8590–8598. doi: 10.1021/acs.jafc.9b03456

Martin, V. J., Pitera, D., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* 21, 796–802. doi: 10.1038/nbt833

Mercke, P., Crock, J., Croteau, R., and Brodelius, P. E. (1999). Cloning, expression, and characterization of *epi*-cedrol synthase, a sesquiterpene cyclase from *Artemisia annua* L. *Arch. Biochem. Biophys.* 369, 213–222. doi: 10.1006/abbi.1999.1358

Newman, J. D., and Chappell, J. (1999). Isoprenoid biosynthesis in plants: carbon partitioning within the cytoplasmic pathway. *Crit. Rev. Biochem. Mol.* 34, 95–106. doi: 10.1080/10409239991209228

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., Mcphee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi: 10.1038/nature12051

Picaud, S., Brodelius, M., and Brodelius, P. E. (2005). Expression, purification and characterization of recombinant (E)-β-farnesene synthase from *Artemisia annua*. *Phytochemistry* 66, 961–967. doi: 10.1016/j.phytochem.2005.03.027

Pichersky, E., and Gershenzon, J. (2002). The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr. Opin. Plant Biol.* 5, 237–243. doi: 10.1016/S1369-5266(02)00251-0

Quintana, A., Reinhard, J., Faure, R., Uva, P., Bagneres, A., Massiot, G., et al. (2003). Interspecific variation in terpenoid composition of defensive secretions of European reticulitermes termites. *J. Chem. Ecol.* 29, 639–652. doi: 10.1023/A:1022868603108

Rastogi, N., Abaul, J., Goh, K. S., Devallois, A., Philogene, E., and Bourgeois, P. (1998). Antimycobacterial activity of chemically defined natural substances from the caribbean flora in guadeloupe. *Fems. Immunol. Med. Mic.* 20, 267–273. doi: 10.1111/j.1574-695X.1998.tb01136.x

Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943. doi: 10.1038/nature04640

Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S., and Sahm, H. (1996). Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative non-mevalonate pathway for terpenoid biosynthesis. *J. Am. Chem. Soc.* 118, 2564–2566. doi: 10.1021/ja9538344

Rohmer, M. (1999). "Isoprenoids including carotenoids and steroids," in *Comprehensive Natural Products Chemistry*, ed D. E. Cane (Oxford), 2.

Ru, W. M., Zhang, J. T., Zhang, F., and Zhang, G. P. (2006). Eco-environmental characteristics and endangered causes of Taxus mairei, an endemic to China. *Bull. Bot. Res.* 26, 624–628.

Saxena, B., Subramaniyan, M., Malhotra, K., Bhavesh, N. S., Potlakayala, S. D., and Kumar, S. (2014). Metabolic engineering of chloroplasts for artemisinic acid biosynthesis and impact on plant growth. *J. Biosci.* 39, 33–41. doi: 10.1007/s12038-013-9402-z

Srikrishna, A., and Satyanarayana, G. (2005). An enantiospecific total synthesis of (–)-patchouli alcohol. *Tetrahedron. Asymmetry* 16, 3992–3997. doi: 10.1016/j.tetasy.2005.11.012

Tarshis, L. C., Yan, M. J., Poulter, C. D., and Sacchettini, J. C. (1994). Crystal-structure of recombinant farnesyl diphosphate synthase at 2.6-agnstrom resolution. *Biochemistry* 33, 10871–10877. doi: 10.1021/bi00202a004

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* 33, 949–956. doi: 10.1046/j.1365-313X.2003.01676.xs

Wagner, G. J. (1991). Secreting glandular trichomes: more than just hairs. *Plant Physiol.* 96, 675–679. doi: 10.1104/pp.96.3.675

Wallaart, T. E., Bouwmeester, H. J., Hille, J., Poppinga, L., and Maijers, N. C. (2001). Amorpha-4, 11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. *Planta* 212, 460–465. doi: 10.1007/s004250000428

Wang, B., Kashkooli, A. B., Sallets, A., Ting, H. M., Ruijter, N. C. A. D., Olofsson, L., et al. (2016). Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metab. Eng.* 38,159–169. doi: 10.1016/j.ymben.2016.07.004

Weathers, P. J., Elkholy, S., and Wobbe, K. K. (2006). Artemisinin: the biosynthetic pathway and its regulation in *Artemisia annua*, a terpenoid-rich species. *In Vitro Cell Dev. Biol. Plant* 42, 309–317. doi: 10.1079/IVP2006782

Westfall, P. J., Pitera, D. J., Lenihan, J. R., Eng, D., Woolard, F. X., Regentin, R., et al. (2012). Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U.S.A.* 109, E111–E118. doi: 10.1073/pnas.1110740109

Wu, S. Q., Jiang, Z., Chase, K., Nybo, S. E., Husodo, S., Williams, R., et al. (2012). Engineering triterpene metabolism in tobacco. *Planta* 236, 867–877. doi: 10.1007/s00425-012-1680-4

Wu, S. Q., Schalk, M., Clark, A., Miles, R. B., Coates, R., and Chappell, J. (2006). Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat. Biotechnol.* 24, 1441–1447. doi: 10.1038/nbt1251

Zhan, X., Zhang, Y. H., Chen, D. F., and Simonsen, H. T. (2014). Metabolic engineering of the moss *Physcomitrella patens* to produce the sesquiterpenoids patchoulol and α/β-santalene. *Front. Plant Sci.* 5:636. doi: 10.3389/fpls.2014.00636

Zhang, J. T., and Ru, W. M. (2010). Population characteristics of endangered species *Taxus chinensis* var. mairei and its conservation strategy in Shanxi, China. *Popul. Ecol.* 52, 407–416. doi: 10.1007/s10144-009-0192-y

Zhang, L., Jing, F. Y., Li, F., Li, M. Y., Wang, Y. Y., Wang, G. F., et al. (2009). Development of transgenic *Artemisia annua* (Chinese wormwood) plants with an enhanced content of artemisinin, an effective anti-malarial drug, by hairpin-RNA-mediated gene silencing. *Biotechnol. Appl. Biochem.* 52, 199–207. doi: 10.1042/BA20080068

Conflict of Interest: YZ, ZD, WL, JZhang, LH, and DC were employed by company Firmenich Aromatics (China) Co. Ltd.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 05 February 2021
doi: 10.3389/fbioe.2021.632269

Recent Advances in Metabolic Engineering, Protein Engineering, and Transcriptome-Guided Insights Toward Synthetic Production of Taxol

*Ishmael Mutanda[1], Jianhua Li[1], Fanglin Xu[1,2,3] and Yong Wang[1]**

[1] Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China, [2] University of Chinese Academy of Sciences, Beijing, China, [3] Key Laboratory of Plant Stress Biology, State Key Laboratory of Cotton Biology, School of Life Sciences, He'nan University, Kaifeng, China

OPEN ACCESS

Edited by:
Jingwen Zhou, Jiangnan University, China

Reviewed by:
Kohsuke Honda, Osaka University, Japan
Leonardo Rios Solis, University of Edinburgh, United Kingdom

***Correspondence:**
Yong Wang
yongwang@cemps.ac.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *22 November 2020*
Accepted: *11 January 2021*
Published: *05 February 2021*

Citation:
Mutanda I, Li J, Xu F and Wang Y (2021) Recent Advances in Metabolic Engineering, Protein Engineering, and Transcriptome-Guided Insights Toward Synthetic Production of Taxol. Front. Bioeng. Biotechnol. 9:632269. doi: 10.3389/fbioe.2021.632269

The diterpenoid paclitaxel (Taxol®) is a blockbuster anticancer agent that was originally isolated from the Pacific yew (*Taxus brevifolia*) five decades ago. Despite the wealth of information gained over the years on Taxol research, there still remains supply issues to meet increasing clinical demand. Although alternative Taxol production methods have been developed, they still face several drawbacks that cause supply shortages and high production costs. It is highly desired to develop biotechnological production platforms for Taxol, however, there are still gaps in our understanding of the biosynthetic pathway, catalytic enzymes, regulatory and control mechanisms that hamper production of this critical drug by synthetic biology approaches. Over the past 5 years, significant advances were made in metabolic engineering and optimization of the Taxol pathway in different hosts, leading to accumulation of taxane intermediates. Computational and experimental approaches were leveraged to gain mechanistic insights into the catalytic cycle of pathway enzymes and guide rational protein engineering efforts to improve catalytic fitness and substrate/product specificity, especially of the cytochrome P450s (CYP450s). Notable breakthroughs were also realized in engineering the pathway in plant hosts that are more promising in addressing the challenging CYP450 chemistry. Here, we review these recent advances and in addition, we summarize recent transcriptomic data sets of *Taxus* species and elicited culture cells, and give a bird's-eye view of the information that can be gleaned from these publicly available resources. Recent mining of transcriptome data sets led to discovery of two putative pathway enzymes, provided many lead candidates for the missing steps and provided new insights on the regulatory mechanisms governing Taxol biosynthesis. All these inferences are relevant to future biotechnological production of Taxol.

Keywords: Taxol, protein engineering, transcriptome, paclitaxel, taxadien-5α-ol, taxane-5α-hydroxylase, metabolic engineering

INTRODUCTION

The blockbuster antitumor drug paclitaxel (Taxol®) is a highly functionalized plant diterpenoid discovered in the late 1960s in pacific yew (*Taxus brevifolia*) plants. Though it went through a long developmental phase, it later became the most successful plant natural product in use as an effective chemotherapy drug since its initial Food and Drug Administration (FDA) approval in 1992 (Wani et al., 1971; Wani and Horwitz, 2014). Taxol was initially approved for treatment of refractory ovarian cancer and breast carcinomas, but over the years it has found expanded clinical uses in treatment of non-small cell lung carcinoma, Kaposi's sarcoma and cancers of the lung, breast, bladder, prostate, esophagus and pancreas when used alone or in combination therapies with other antineoplastic agents.

Original production process involved direct isolation from the bark of yew plants, but the process is very destructive and inefficient, requiring three trees (∼12 kg bark material) to produce 1 g pure Taxol or ca. 3,000 yew trees to produce only 1 kg of the drug (Malik et al., 2011; McElroy and Jennewein, 2017). To add to this, Taxol is produced as a complex mixture with hundreds of other taxanes in *Taxus* plant tissue, making the purification of this highly hydrophobic compound very laborious and environmentally damaging due to use of organic solvents (Wani and Horwitz, 2014). Total chemical synthesis routes to Taxol were developed (Holton et al., 1994; Nicolaou et al., 1994), but are not commercially viable owing to the many steps required and cost considerations. Inspired by the success of Taxol, and the need to address supply and ecological challenges, alternative production platforms have been developed: (i) extraction of 10-deacetylbaccatin III (10-DAB) or baccatin III (BIII) from renewable twigs of *Taxus* species, that can be semi-synthetically modified to Taxol, and (ii) plant cell fermentation (PCF) involving use of *Taxus* cell suspension cultures (Fett-Netto et al., 1992; Mountford, 2010; Malik et al., 2011).

Despite several optimization efforts to improve these methods, they still suffer from several drawbacks that limit their capacity and also drive the price of Taxol high; (i) production from twigs is still dependent on yew trees, thus is susceptible to weather and environmental factors, (ii) long maturity time of the trees and (iii) extraction process still involves use of organic solvents. The PCF route presents obvious advantages over direct extraction from field-grown *Taxus* twigs, but it suffers from its own shortcomings; (i) instability of cell lines in the long fermentation periods (ii) use of endogenous biosynthetic pathways, thus no genetic engineering targets to improve flux toward Taxol and (iii) poor Taxol yields in the fermenters, even with elicitors (Choi et al., 2000; Ketchum and Croteau, 2006).

For these reasons, it is highly desirable to establish biotechnological production systems for Taxol production that will address most of these issues. Production in more amenable, fast-growing heterologous hosts offers more advantages in terms of control and manipulation of metabolic flux by improving enzyme expression, pathway regulation, availability of cofactors and engineering competing pathways. However, there are challenges that prevent development of desired sustainable Taxol biotechnological platforms: the biosynthetic pathway is long and complicated, involving 19 expected steps from geranylgeranyl diphosphate (GGPP) the universal precursor of diterpenoids (**Figure 1**) and there are still many gaps in our understanding of Taxol biosynthesis and its regulatory mechanisms. Several cytochrome P450 (CYP450) hydroxylases and an epoxidase remain missing (**Figure 1**). Even for the steps with known enzymes, expression of functional pathway enzymes in heterologous hosts, especially the CYP450s is proving to be non-trivial.

In this review, we discuss various strategies conducted in the last 5 years to develop heterologous systems for Taxol biosynthesis. Building from the breakthrough strategy 10 years ago (Ajikumar et al., 2010), recent advances in synthetic biology have been applied in several studies to optimize pathway expression, develop enzyme screening platforms and engineer pathway enzymes to get variants with improved catalytic performance and higher specificity. We also highlight the use of computational tools that were leveraged in gaining insights in the catalytic mechanisms of several enzymes and lastly, we summarize recent Taxol-related transcriptomes and how this publicly available resource was recently used in inferring regulatory mechanisms, mining the missing genes and providing many lead candidates for missing steps. In the coming years, we expect this knowledge to be pivotal in development of biotechnological systems for production of this critical drug.

CHALLENGES WITH METABOLIC ENGINEERING OF TAXOL PATHWAY

Metabolic engineering of the Taxol pathway in heterologous hosts is highly desirable as it will establish a versatile, direct route to the critical anticancer drug. Biosynthetic production systems are expected to improve supply of Taxol, as well as lower the production costs through improved efficiency and use of inexpensive sugars as starting material in microbial factories. Another potential avenue under exploration is production in cheaper, high biomass and fast-growing plants like tobacco that can be engineered for high-level production of non-native metabolites using carbon resources from photosynthesis. Before realizing this dream, there are several challenges with Taxol biosynthesis that should be addressed.

The Missing Pathway Enzymes

Taxol is derived from the C_5 terpenoid universal precursors, dimethylallyl diphosphate (DMAPP) and isopentenyl pyrophosphate (IPP) through geranylgeranyl diphosphate (GGDP), the C_{20} universal precursor of all diterpenoids. The structure of the Taxol molecule is complex and likewise, the biosynthetic pathway is equally complex; five decades after its discovery, the pathway is not yet fully understood, with several steps still undefined, and several enzymes still missing. These gaps in our understanding of the biosynthetic pathway deprive us of the critical framework to manipulate Taxol biosynthesis and develop heterologous production systems. The pathway is hypothesized to involve 19 steps from GGPP

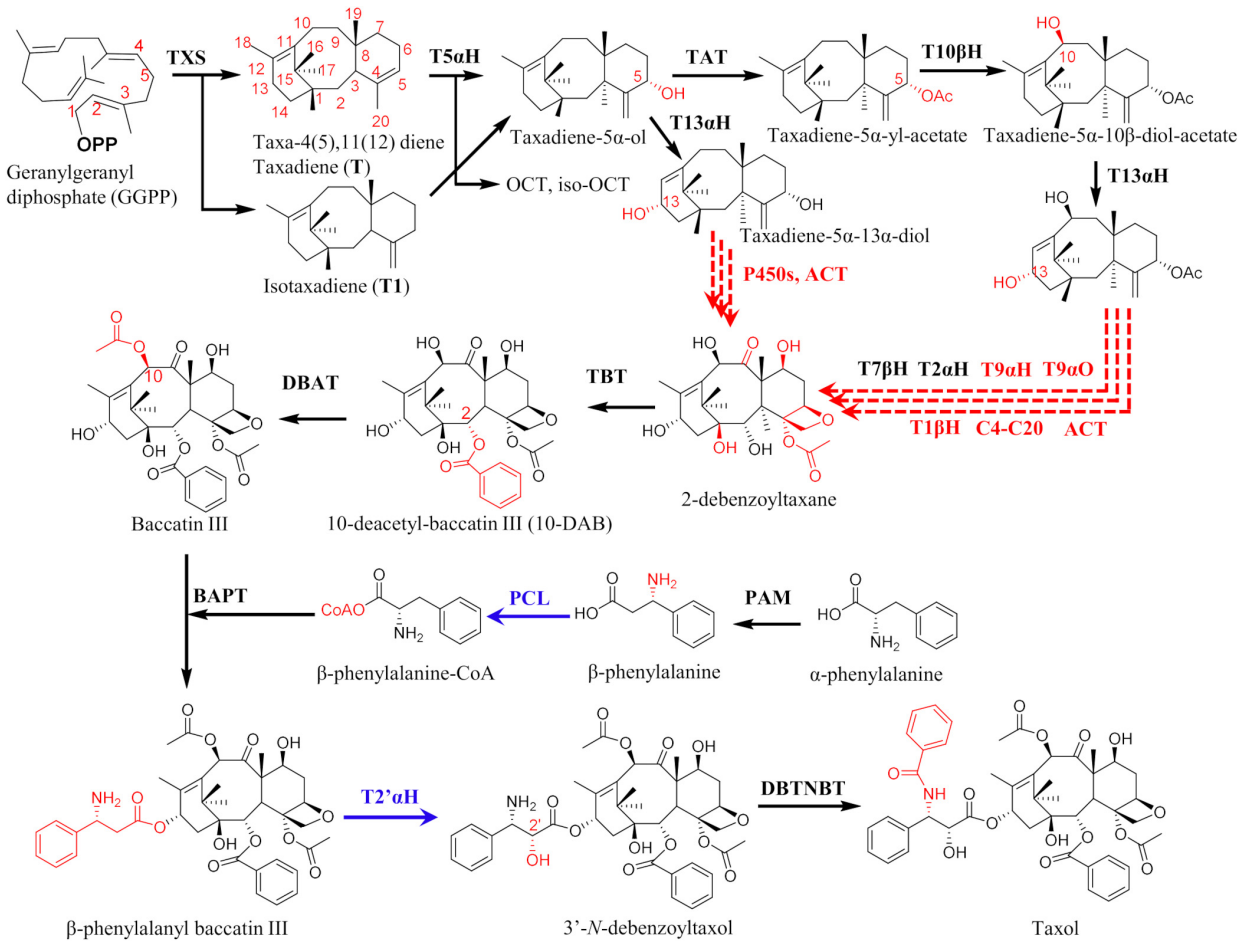


FIGURE 1 | Taxol biosynthetic pathway from geranylgeranyl diphosphate, showing carbon positions. Enzymes in red are not yet characterized and steps in red dotted arrows are not yet fully elucidated. Blue arrows—genes encoding putative PCL and T2′αH were recently isolated, though not yet applied in metabolic engineering designs. Enzyme abbreviations: TXS, taxadiene synthase; T5αH, taxane-5α-hydroxylase; TAT, taxane-5α-ol-O-acetyltransferase; T10βH, taxane-10β-hydroxylase; T13αH, taxane-13α-hydroxylase; T1βH, taxane 1β-hydroxylase; T9αH, taxane 9α-hydroxylase; T9αO, taxane 9α-dioxygenase; T2αH, taxane 2α-hydroxylase; T7βH, taxane 7β-hydroxylase; C4-C20, C4-C20 epoxidase; TBT, taxane-2α-O-benzoyl transferase; DBAT, 10-deacetylbaccatin III-10-O-acetyltransferase; PAM, phenylalanine aminomutase; PCL, phenylalanine-CoA ligase; BAPT, C-13 phenylpropanoyl-CoA transferase; T2′αH, taxane 2′α-hydroxylase; DBTNBT, Debenzoyl taxol *N*-benzoyl transferase. Products of T5aH—OCT, 5(12)-oxa-3 (11)-cyclotaxane and iso-OCT, 5(13)-oxa-3(11)-cyclotaxane, ACT, acyl-CoA transferase.

(**Figure 1**), involving terpene cyclization, 9 cytochrome P450-catalyzed hydroxylations, 3 acylations, acetylations, oxetane ring formation, benzoylations, and phenylisoserine side chain attachment. Exceptional work in the last 25 years has contributed genes and enzymes of up to 14 steps; from taxadiene synthase (TXS) that catalyzes the first and committed step to cyclize GGPPS to taxadiene (Hezari et al., 1995; Wildung and Croteau, 1996), five CYP450s [taxane 5α-hydroxylase (T5αH), taxane 2α-hydroxylase (T2αH), taxane 7β-hydroxylase (T7βH), taxane 10β-hydroxylase (T10βH) and taxane 13α-hydroxylase (T13αH)] (Hefner et al., 1996; Jennewein et al., 2001; Kaspera and Croteau, 2006; Schoendorf et al., 2001; Walker et al., 2000; Walker and Croteau, 2001), five acyl CoA transferases (ACTs) [taxane-5α-ol-O-acetyltransferase (TAT), taxane-2α-O-benzoyl transferase (TBT), 10-deacetylbaccatin III-10-O-acetyltransferase (DBAT), C-13 phenylpropanoyl-CoA transferase (BAPT) and Debenzoyl taxol *N*-benzoyl transferase (DBTNBT)] and phenylalanine aminomutase (PAM).

A putative ACT that activates β-phenylalanine to β-phenylalanine-CoA was isolated from *T. baccata* cell cultures (Ramírez-Estrada et al., 2016), and was identified as β-phenylalanine-CoA ligase (PCL) after functional characterization. However, a recently isolated homolog of this PCL displaying 99% sequence identity showed no detectable activity against both β-phenylalanine and 3-phenylisoserine *in vitro* (Srividya et al., 2020), thus, the suitability of this clone for heterologous expression in other hosts might remain an open question. Other still missing enzymes are the CYP450s taxane 1β-hydroxylase (T1βH), taxane 9α-hydroxylase (T9αH), taxane 9α-oxidase (T9αO), and C4,5 epoxidase (**Figure 1**). Taxane 2′α-hydroxylase (T2′αH) was recently isolated from mining *T. baccata* transcriptome (Sanchez-Muñoz et al., 2020), and the

details are reviewed in the section on transcriptomes below. The missing enzymes, together with other challenges outlined below need to be resolved before successful production of Taxol by synthetic biology systems. Other enzymes have been identified that channel flux toward off-pathway intermediates that do not lead to Taxol, such as the taxane 14β-hydroxylase (T14βH), the recently discovered ACTs that attach different groups to 3′-*N*-debenzoyltaxol (Srividya et al., 2020) and many more that are expected considering the chemical diversity of taxanes in plant tissues.

Poor Expression of Functional Enzymes and Challenges in CYP450 Chemistry

Oxygenation chemistry of Taxol presents a formidable barrier to both pathway elucidation and metabolic engineering efforts, starting with the first hydroxylation of the C5 position of taxadiene by T5αH (Hefner et al., 1996; Jennewein et al., 2004). All *Taxus* CYP450s identified and functionally characterized for activity in Taxol biosynthesis are from the CYP725A subfamily, including T5αH (CYP725A4). T5αH accepts both taxa-4(5),11(12) diene (taxadiene, **T**) and its close alternative isomer taxa-4(20),11(12) diene (isotaxadiene, **T1**), that are both products of TXS to form taxadiene-5α-ol (T-5α-ol) (**Figure 1**). This CYP450 presents several challenges due to the fact that it accepts both **T** and **T1** as substrates, its inherent product promiscuity, poor expression in heterologous hosts and low catalytic activity (it has been shown to convert less than 10% of taxadiene to T-5α-ol). Most research efforts in the past 5 years were focused on T5αH, and more synthetic biology tools have been dedicated to T5αH than any other pathway enzyme to overcome this bottleneck on the second step of the pathway, thus, this review also has more comprehensive discussions on this enzyme.

To understand the origin of T5αH bottleneck, we present its history in heterologous expression in different hosts. Following successful cloning, functional expression and characterization of eight enzymes of the upper segment of the Taxol pathway, DeJong et al. (2006) sought to construct the first five sequential steps of the pathway leading to taxadien-5α-acetoxy-10β-ol in yeast (*Saccharomyces cerevisiae*). Expression and enzyme kinetics of T5αH in this construct was the lowest compared to other enzymes, leading to the first detection of a bottleneck at this first oxygenation step that restricted flux toward downstream steps. Two years later, an attempt to introduce TXS and T5αH in wild tobacco (*Nicotiana sylvestris*), targeting trichomes after knocking down production of cembratrien-diols similarly failed to produce the desired T-5α-ol, but instead led to production of a cyclic ether, 5(12)-oxa-3(11)-cyclotaxane (OCT) (Rontein et al., 2008) (summary of these metabolic engineering constructs are in **Table 1**). Intrigued by this observation, T5αH was expressed in yeast in the same study, and again OCT was observed. The bottleneck was again encountered in a carefully optimized *Escherichia coli* strain that was engineered via a multivariate modular metabolic engineering (MMME) approach that achieved a 15,000-fold increase in taxadiene production (titers of ∼1 g/L) but lost optimality and titers on introduction of T5αH (Ajikumar et al., 2010). Recent studies corroborated lack of selectivity and product promiscuity of T5αH *in vitro* and in several heterologous hosts like *E. coli*, *S. cerevisiae*, *Yarrowia lipolytica,* and *Nicotiana benthamiana*, showing a product profile dominated by OCT and its close isomer, iso-OCT, with T-5α-ol and several other monooxygenated diterpenes as minor compounds (Yadav, 2014; Biggs et al., 2016a,b; Edgar et al., 2016; Sagwan-Barkdoll and Anterola, 2017; Li et al., 2019). A number of optimization and engineering strategies that leveraged advances in synthetic biology have been applied to overcome the T5αH bottleneck as discussed in the sections below.

Poor Pathway Flux and the Branched Nature of the Pathway

Despite that much of the Taxol pathway enzymes and genes have been identified, success in engineering at least the known segments of the pathway in heterologous systems have been mainly hampered by inherent poor enzyme catalysis, protein interdependency issues and product promiscuity of key enzymes that lead to a highly branched pathway. Carbon flux is channeled toward off-target products by the promiscuous enzymes, presenting a formidable challenge to synthetic biologists that cannot be solved by simple redirection of IPP and DMAPP precursors. The first enzyme, TXS is now known to have a broad substrate profile including verticillenes and reports of cembrene A, in addition to **T** and **T1**, as discussed below. The T5αH- catalyzed second step splits **T** into many other products, dominated by OCT and its isomer, iso-OCT, branching the pathway into many directions. Several other downstream enzymes also accept different substrates and have broad product profiles, which present an emerging picture of a highly branched pathway. This lack of linearity is at the core of the problems with the pathway that has not only slowed metabolic engineering advances, but has even prevented pathway elucidation and identification of key downstream intermediates. Not surprisingly, recent work has focused on improving the catalytic efficiency and selectivity of pathway enzymes through experimental methods supported by computational tools to improve our mechanistic understanding of catalysis, as covered in the next sections.

Lack of Knowledge on the Regulatory Mechanisms and *in planta* Transport Mechanisms

Not much is known on the regulation of the Taxol pathway at transcription, translation and post-translational levels. Elicitors like methyl jasmonate (MeJA) have been used for a relatively long time in improving secondary metabolism in *Taxus* cell suspension cultures, but the mechanisms through which MeJA activates Taxol biosynthesis have not been elucidated in detail. Knowledge of the regulatory mechanisms, identification of key transcription factors and any feedback loops in the pathway is critical in informing metabolic engineering efforts. Likewise, there are also many gaps in our understanding of the transport mechanisms of taxane intermediates in plant cells. This wealth of information will be very vital in designing a biosynthetic route to Taxol especially in plant cells.

TABLE 1 | Heterologous production of early taxane metabolites in different platforms.

Host	Details	Achievements	References
E. coli	Multivariate-modular metabolic engineering (MMME) approach to optimize the MEP pathway and GGPP-TXS as two operons under inducible promoters	Taxadiene accumulation to 1 g/L. Introduction of T5αH-CPR disrupted taxadiene balance and achieved ∼58 mg/L taxadiene-5α-ol and equal amounts of 5(12)-oxa-3(11)-cyclotaxane (OCT)	Ajikumar et al., 2010
E. coli	TXS and elaborately optimized T5αH and CYP450 reductase partner optimizations through N-terminal modifications, fusion linked chimera protein expression, and controlled promoter strength.	Oxygenated taxanes (570 mg/L) was achieved in a bioreactor	Biggs et al., 2016a
E. coli	Co-expression of dxs, idi, GGPPS, and TXS	Taxadiene (1.3 mg/L) in shake flask	Huang et al., 2001
E. coli and S. cerevisiae	A synthetic consortium was designed and genes for taxadiene-5α-10β-diol-acetate production were designed in 2 modules, taxadiene module in *E. coli* and acetylation and CYP450-oxygenation chemistry in *S. cerevisiae* for a stable co-culture fermentation using xylose as carbon source	33 mg/L oxygenated taxanes were achieved with TXS and T5αH, and adding T10βH and TAT achieved 1.0 mg/L of the target monoacylated, dioxygenated taxane	Zhou et al., 2015
Bacillus subtilis	Overexpression of all MEP pathway genes (*dxs*, *ispD*, *ispF*, *ispH*, *ispC*, *ispE*, *ispG*) together with *ispA*, GGPPS and TXS	Taxadiene accumulated to 1.98 mg/L/OD_{600} (17.8 mg/L) in shake flask	Abdallah et al., 2019
S. cerevisiae	Multi-step pathway construction of 5 taxoid biosynthetic genes (GGPPS, TXS, T5αH, TAT, and T10βH) to attempt taxadiene-5α-acetoxy-10β-ol production.	All 5 recombinant proteins were successfully expressed and had measurable activity. Only Taxadiene (1 mg/L) and trace amounts of taxadiene-5α-ol (∼25 μg/L) was detected. No advanced metabolites were detected	DeJong et al., 2006
S. cerevisiae	Heterologous expression of a truncated 3-hydroxyl-3-methylglutaryl-CoA reductase (tHMGR), a mutant regulatory protein, UPC2-1, GGPPS from *Sulfolobus acidocaldarius* and TXS	Taxadiene (8.7 mg/L) and geranylgeraniol (33 mg/L) accumulated in shake flasks after 48 h fermentations	Engels et al., 2008
S. cerevisiae	Heterologous TXS and GGPPS (from *Taxus cuspidate* × *Taxus baccata*) and overexpression of erg20 and tHMGR	Taxadiene accumulated to 72.8 mg/L	Ding et al., 2014
S. cerevisiae	A CRISPR/Cas 9 toolkit was tested on TXS expression optimization in yeast. 10 protein tags and 5 promoters of different strengths were tested. Fusion of TXS to MBP under the strong GAL1 promoter achieved highest titer	Taxadiene titer of 20 mg/L was achieved	Apel et al., 2017
S. cerevisiae	TXS-ERG20 fusion protein was constructed with MBP tag for improved solubility, together with promoter strength and growth temperature optimization	High taxadiene titer in yeast of 129 mg/L was achieved in a bioreactor	Nowrouzi et al., 2020
A. thaliana	Chimeric TXS cDNA constitutively expressed in *A. thaliana*	Taxadiene (∼20 ng/g DW) in seedlings and leaves, but however observed stunted growth and reduced photosynthetic pigments. Induction with the synthetic glucocorticoid (dexamethasone) improved yields to 600 ng/g DW	Besumbes et al., 2004
Tobacco (*Nicotiana sylvestris*)	TSX and T5αH were stably expressed in tobacco trichome cells	Taxadiene (no reported yield) was detected while expected taxadiene-5α-ol was not detected in leaf extracts. Instead, only OCT was detected. Yeast microsomes also produced OCT only	Rontein et al., 2008
Tomato fruits	TXS was stably transformed into a yellow-fruited tomato line which lacks a functional phytoene synthase	160 mg/kg from freeze-dried tomatoes	Kovacs et al., 2007
Tobacco (*Nicotiana benthamiana*)	TXS, truncated T5αH and cytochrome P450 reductase were inserted into the chloroplast compartment and precursor pathway was overexpressed	Taxadiene—56.6 μg/g FW and Taxadiene-5α-ol was detected for the first time in a heterologous plant platform at 1.3 μg/g fresh weight	Li et al., 2019
Ginseng (*Panax ginseng*) roots	Stable transformation of TXS from *Taxus brevifolia* into ginseng roots	TXS-transgenic ginseng accumulated 9.1 μg/g DW. Methyl jasmonate treatment improved yields to 14.6–15.9 μg/g DW	Cha et al., 2012
Physcomitrella patens (moss)	Stable constitutive expression of TXS using a ubiquitin promoter	Taxadiene accumulated to 0.05% FW of plant tissue. No adverse effects on growth were noted	Anterola et al., 2009
Alternaria alternata (endophytic fungus)	Co-overexpression of isopentenyl diphosphate (*idi*), truncated 3-hydroxy-3-methylglutaryl-CoA reductase (tHMG1) and TXS under different promoter strengths	Detection of 61.9 μg/L taxadiene after 14 days of fermentation	Bian et al., 2017

ADVANCES IN METABOLIC ENGINEERING, PROTEIN ENGINEERING AND MECHANISTIC INSIGHTS ON TAXOL RELATED EZYMES

Synthetic biology tools have been widely applied in advancing Taxol biosynthesis research from gene discovery to pathway designs and construction in heterologous hosts, resulting in successes in detection and accumulation of taxane intermediates. Recent metabolic engineering strategies and achievements in different hosts are summarized in **Table 1**, together with approaches used in previous years, for comparison. Computational tools are an enabling technology that has also been at the forefront in unraveling mechanistic insights of key enzymes to guide protein engineering strategies for use in metabolic engineering constructs, thus we discuss these advances together in this section.

Quantum Mechanics/Molecular Mechanics (QM/MM) and Computational Modeling of a Catalytically Active TXS to Enable Enzyme Engineering

The class 1 terpene cyclase taxadiene synthase (TXS) is the first and rate-limiting enzyme of the Taxol biosynthesis pathway. It accepts the acyclic C20 diterpenoid precursor, (*E,E,E*)-geranylgeranyl diphosphate (GGPP) and cyclizes it to the endocyclic diterpene olefin taxa-4(5),11(12)-diene (taxadiene, **T**) and four other minor products: taxa-4(20), 11(12)-diene (isotaxadiene, **T1**), verticillia-3(4),7(8),12(13)-triene (**V**), verticillia- 4(20),7(8),11(12)-triene (**V1**) and verticillia-3(4),7(8),11(12)-triene (**V2**) (Koepp et al., 1995; Lin et al., 1996; Schrepfer et al., 2016; Li et al., 2019) and also the isomer taxa-3(4),11(12)-diene (Williams et al., 2000b; Li et al., 2019). Though the X-ray crystal structure of TXS was solved almost a decade ago (Köksal et al., 2011), it lacks N-terminal residues and is in an open, catalytically inactive form, which does not provide much clues relevant to the mechanism and intricate architecture of the active form. In addition, the reported structure is bound to the fluorinated substrate, 2-fluoro-geranylgeranyl diphosphate (2-F-GGPP) in a non-productive orientation (Hong and Tantillo, 2011; Schrepfer et al., 2016).

Consequently, labeling and computational tools have been used to get insights on the energetics of proton transfer and carbocation formation in the TXS active site. The generally agreed mechanism follows TXS-catalyzed cleavage of the pyrophosphate moiety (PPi) from GGPP to form charged carbocations followed by subsequent cyclizations and proton transfer, leading to the mixed product profile stated above (Williams et al., 2000a; Köksal et al., 2011). Using QM calculations, an indirect, two-step protein transfer sequence was proposed (Gutta and Tantillo, 2007; Hong and Tantillo, 2011). However, these gas phase calculations assumed a passive role of the TXS protein and placed much emphasis on substrate reactivity. To improve on this prior computational work and identify the role of the deprotonating bases in the TXS active site, more recent work focused on building closed, active models of the protein. Two groups have so far successfully used homology modeling and loop modeling based on the closed structure of bornyl diphosphate synthases (BPPS) as a template to model the missing N-terminal residues and build catalytically active models that can be used in docking the productive substrate, GGPP (Schrepfer et al., 2016; Freud et al., 2017). Such structural model based molecular mechanics calculations have the potential to inform enzyme engineering strategies to generate pathway enzyme variants with improved catalytic efficiency, less product promiscuity or tailored product profile to enable selective attenuation of the carbocation products. For example, closed TXS model construction and molecular mechanics by Schrepfer et al. (2016) led to identification of a conserved amino acid network responsible for an extended hydrogen (H) bonding (water- and amino-acid mediated) involving the PPi and R754, R768, Y835, R580, and N-terminal Y89 in the closed TXS-GGPP complex. Targeted mutagenesis of these five residues lead to loss of activity (**Table 2**), and the W753 residue was identified as a key deprotonating base for the Cation B (cembren-15-yl cation) that resulted in cembrene A formation when mutated to a histidine (W753H) or in a double mutant (W753H/C830A) (**Table 2**; Ansbacher et al., 2018; Schrepfer et al., 2016). In fact, the biotechnological application of these TXS-derived mutants was demonstrated, with W753H yielding 8 mg/L cembrene A and V584M yielding 11 mg/L verticilla-3,7,12(13)-triene in 30 L batch fermentations (Schrepfer et al., 2016).

Using the TXS model constructed by Freud et al. (2017), the crucial role of W753 and the analogous Y841 in controlling highly charged carbocations in the hydrophobic TXS active site were highlighted in simulations by Ansbacher et al. (2018). However, recent molecular dynamics (MD) and QM/MM calculations place more emphasis on the role of the reactive carbocations, the retained PPi and active site water molecules in controlling carbocation cascades and product distribution, rather than the TXS residues (Escorcia et al., 2018; van Rijn et al., 2019). The calculations corroborate the critical role of R580 in deprotonation of carbocation C, but suggest this is through water-assisted deprotonation via one or two water bridges, rather than direct interaction with PPi and closure of the active site (as a part of the RXR motif of terpene synthases) as suggested by Freud et al. (2017). Another interesting take from these computational and experimental data sets was the manipulation of the taxadiene and iso-taxadiene product percentage, as this has been shown to be important in determining the selectivity of the subsequent CYP450-catalyzed step as described below. C830S yielded the highest iso-taxadiene without much reduction in overall activity (**Table 2**), and such a mutant can be leveraged for improving the selectivity of T5αH as demonstrated by Edgar et al. (2017) (described below).

Strategies Toward Overcoming the T5αH Bottleneck

Optimization of Oxygenation Chemistry

Realizing the magnitude of the bottleneck posed by T5αH, Biggs et al. (2016a) carried out an extensive study to optimize

TABLE 2 | Mutagenesis of enzymes of the Taxol pathway to manipulate catalytic fitness, activity and product distribution profile.

Target	Mutation	% Activity	Taxadiene (T)	Iso-taxadiene (T1)	Verticillenes	Cembrene A	References
TXS							
TXS	Wild type	100	93.2	4.7	2.1	N.D	Schrepfer et al., 2016
Y89	Y89A/E/F	Lost activity*					
R580	R580A/E/H	Lost activity					
R754	R754A/E/H	Lost activity					
R768	R768A/E/H	Lost activity					
Y835	Y835A/F/W	Lost activity					
V584	V584K/L/M	89–92	14–30	0.6	70–85.6		
	V584N/S/P/R	Lost activity					
S587	S587D/Y/K/L/G	Lost activity					
	S587A	21.8	8.9	N.D	32.2	58.9	
Y609	Y609G	N.A	N.D	N.D	100	N.A	Edgar et al., 2017
V610	V610H/S/F/A	Lost Activity					Schrepfer et al., 2016
S713	S713T	97.4	92.7	5.1	2	N.D	Schrepfer et al., 2016
	S713A/L	Lost activity					Schrepfer et al., 2016
V714	V714A/I	8.7–10.4	94	5.7–5.9	N.D	N.D	
	V714T/G/P	Lost activity					
G715	G715A/S	Lost Activity					
S713	S713T	97.4	N.D	N.D	N.D	100	
	S713A/L	Lost activity					
W753	W753H	51.3	N.D	N.D	N.D	100	Schrepfer et al., 2016; Ansbacher et al., 2018
	W753H/C830A	48.5	N.D	N.D	N.D	100	Schrepfer et al., 2016
	W753A/C/E/L/V	Lost activity					
C830	C830A	88.5	93.1	4.8	2.1	N.D	
	C830S	92.4	79.3	7.3	13.4	N.D	
F834	F834A/G	25.6–28.3	38–87	6.1–6.5			
Y835	Y835F	3.5	100	N.D	N.D	N.D	
	Y835A/W	Lost activity					
Y841	Y841F	41.3	N.D	N.D	43.6	56.4	
	Y841A/T	Lost activity					
Y688¶	Y688L	2.4-fold increase in **T1**, and corresponding increase in T-5α-ol					Edgar et al., 2017
T5αH							
T380	T380S	Produced a dihydroxylated product—5(12)-oxa-3(11)-cyclo-taxan-10-ol					Edgar et al., 2017
K131	K131R	Improved turnover but lost selectivity compared to wild type					Yadav, 2014
V374	V374L	Improved selectivity at the expense of turnover					
S302	S302A	Lost activity and no change in fold-change					
DBAT							
G38	G38R	2.15-fold increase in baccatin III					Li et al., 2017
H162	H162A	Lost activity					Li et al., 2017
	H162A/R63H	3-fold increase in catalytic activity compared to wild type					You et al., 2018
R363	R363A	Lost activity					Li et al., 2017
	R363H	26-fold increase in catalytic activity compared to wild type					You et al., 2018
G361	G361A	Lost activity					Li et al., 2017
I164	I164A	Lost activity					Li et al., 2017
D166	D166H	15-fold increase in catalytic activity compared to wild type					You et al., 2018
	D166H/R363H	60-fold increase in catalytic activity compared to wild type					You et al., 2018
I43/D390	I43S/D380R	3.3-fold increase in catalytic efficiency using vinyl acetate and 3-fold using acetyl CoA					Lin et al., 2018

**Enzyme activity in comparison to the wild type TXS. Activity of 0–2% was qualified as "lost activity". N.D., Not detectable; N.A., Not assayed. ¶ These mutants led to corresponding increases in taxadiene-5α-ol after introduction of T5αH-CPR.*

P450 chemistry in *E. coli* as a proof of concept and achieved a fivefold increase in oxygenated taxanes, reaching the highest oxygenated diterpene titer to date (∼570 mg/L). The work built from the MMME breakthrough (Ajikumar et al., 2010) and aimed to optimize downstream P450 chemistry through chromosomal integration of the upstream pathways and use

of different strategies to optimize T5αH and cytochrome P450 reductase (CPR) partner interactions. A protein interdependency of the oxidative module and the upstream modules was uncovered through targeted proteomics, and was determined as a key obstacle of T5αH expression and to be responsible for reduction of upstream metabolites on introduction of T5αH. Strategies used to uncouple this interdependency and optimize the pathway proteins were varying promoter strength, N-terminal modifications, gene copy number optimization and optimization of CPR interactions (**Figure 2A**). Construction of the T5αH-CPR module as an operon resulted in higher titers compared to a linked, chimera construct; and a weaker Trc promoter demonstrated overall higher yields compared to a stronger T7 promoter. These results suggested the need for a lowered CPR expression as evidenced by the benefits of an operon construction and Trc promoter. To address the solubility issue that is common with P450 enzymes, truncation of the hydrophobic, membrane-targeting N-termini of both T5αH and CPR was undertaken, and three leader peptides of different solubilities (8RP, MA, and 2B1) were attached for N-terminal modifications. Despite the significant improvement in solubility of the expressed respective proteins that was noted, there were no benefits of these modifications on T5αH performance, in fact, increased hydrophilicity (2B1-T5αH/2B1-CPR) resulted in accumulation of taxadiene that was not converted to oxidized taxanes. Ultimately, this work developed a strain with a chromosomally integrated MMME module and T5aH-CPR in an operon construct in a low copy plasmid under a weak Trc promoter as the most optimal that achieved ∼570 mg/L oxygenated taxanes in a benchtop bioreactor.

As demonstrated by Biggs et al. (2016a), optimizing promoter strength to strike a balanced expression of pathway enzymes or achieving dynamic expression is a very powerful tool in metabolic engineering and synthetic biology. In the context of the Taxol pathway, this tool was recently exploited to finely tune expression of TXS and GGPPS using bidirectional promoters (BDPs) in the yeast *Pichia pastoris* (Vogl et al., 2018). BDPs allow not only differential expression of genes utilizing differences in promoter strength, but can be used to explore constitutive and inducible promoters, and for timing where one gene needs to be expressed after another (**Figure 2B**). When GGPPS was expressed under a constitutive promoter, no taxadiene was detected, but when GGPPS was under a depressed promoter and TXS under a different promoter, taxadiene titers increased by nearly 60-fold (Vogl et al., 2018). Though this strategy was not extended to T5αH, the same study already provided a promising example with a human P450 and its CPR partner that improved by fivefold under an optimal BDP. Another example of the versatility of manipulating promoters for dynamic pathway expression to overcome the T5αH bottleneck is a recent report using riboregulated switchable feedback promoters (rSFPs) (Glasscock et al., 2019). These novel rSFPs are created through using a natural, stress-response promoter in conjunction with a target sequence that is inserted between the promoter and the gene. A separate *trans*-acting regulator, called small transcription activating RNA (STAR) is introduced in a second plasmid to bind and activate the rSFP, creating a gated ON/OFF switch for the downstream gene (**Figure 2B**). Working with the best optimized strain from Biggs et al. (2016a), this rSFP tool was used to screen membrane envelope stress-response promoter library with the goal to create a promoter that is responsive to the stress caused by introduction of T5αH/CPR to the upper pathways. Optimization of the timing and expression magnitude of T5αH/CPR with rSFP in this way resulted in a notable improvement of 2.4-fold (25.4 mg/L) in oxygenated taxanes and 3.6-fold (39.0 mg/L) increase in total taxanes (Glasscock et al., 2019).

CRSIPR/Cas9 technology is a fast and precise enabling tool that is increasingly being used in synthetic biology applications. This technology was applied to build a cloning—free screening toolkit for promoter strength and solubility tag optimization in *S. cerevisiae* (Apel et al., 2017) to enable quick exploration of different constructs. As a proof of concept, the toolkit was applied to build an expression context library for TXS to explore localization tags, solubility tags and promoter strength. The fast toolkit identified a solubility problem with TXS in yeast as the major cause of poor catalysis, and identified the best optimized construction of TXS with MBP solubility tag and a strong GAL1 promoter that increased titers 25-fold compared to an optimized strain.

T5αH Protein Engineering

While all the above tools are promising for optimizing T5αH expression, it is important to note that the improvements are in titers of "oxygenated taxanes", a term describing a mixture of different mono- and doubly oxygenated taxanes monitored at m/z 288 and m/z 304 in the GC-MS chromatograms. The only product of T5αH that has been shown to lead to Taxol is T-5α-ol, thus, though increases in total oxygenated taxanes are promising, it is highly desired to address product promiscuity of the enzyme, and devise ways to improve only the desired product (i.e., T-5α-ol). Protein engineering is an effective tool that has already been harnessed to address the T5αH bottleneck, and to improve other enzymes of the Taxol pathway (**Figure 3**). Notwithstanding major advancements in directed evolution as a subfield that recently won a Nobel price, the major challenges with leveraging this tool in engineering T5αH and other P450s of the Taxol pathway is the dearth of mechanistic knowledge of catalysis and the lack of high throughput screening assays to quickly screen the large number of generated mutants. Assuming that the experimentally observed product promiscuity of T5αH was due to competing regiospecific proton abstraction by the oxyferryl species of the P450 enzyme on taxadiene as proposed previously (Hefner et al., 1996; Jennewein et al., 2004), a computational method was developed to guide mutagenesis and improve catalytic efficiency and selectivity (Yadav, 2014). A total of 53 mutants were designed targeting amino acids residues around the active site of the enzyme, basing on an energy-minimized homology model of T5αH that was developed using six P450 structures as template. Assessment of the mutant library identified six variants (5 single and 1 triple mutant) that improved both in turnover and T-5α-ol selectivity, though the identities of the mutants were not clearly stated and the mechanistic basis of the improvements was not discussed. Three mutants from the study were however used to infer mechanistic basis on the

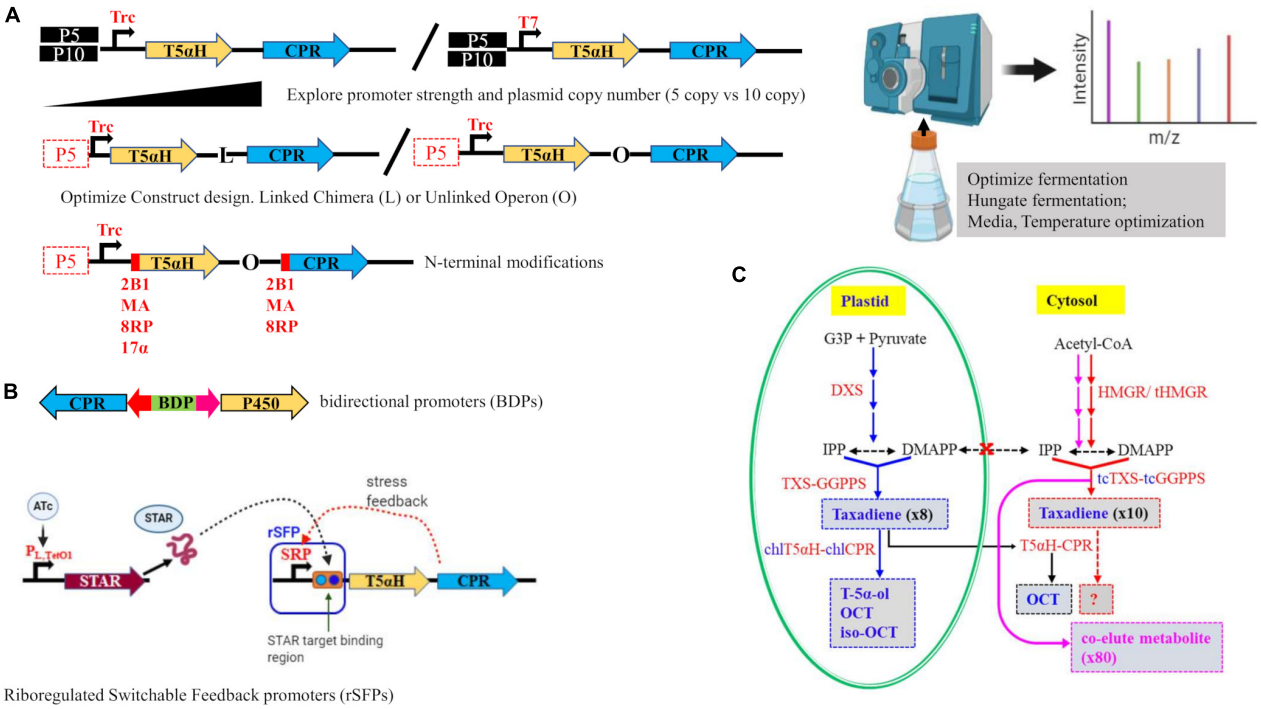


FIGURE 2 | Tools and strategies used to overcome the T5αH bottleneck. **(A)** Optimization of the P450 module in a previously optimized *E. coli* strain with chromosomally integrated MEP and TG modules was achieved using several approaches (see Biggs et al., 2016a). Top panel—A high copy number plasmid (P10) was compared to a low copy number (P5), and expression strength was explored using a strong T7 promoter and a weak (Trc) promoter; Middle panel—Using a low copy P5 plasmid and weak Trc promoter, interactions of the P450 and its reductase partner CPR were explored using a linked chimera construct (L) or an unlinked operon design (O); Lower panel—N-terminal modifications were tested on truncated versions of T5αH and CPR to enhance solubility using three leader peptides—2B1, MA an 8RP; 17α was also used in another study (Rouck et al., 2017). Right panel—Further optimizations including targeted proteomics and optimization of media and fermentation conditions led to highest titers of oxygenated taxanes. **(B)** Targeting promoters for dynamic regulation of T5αH catalysis. Top—Bidirectional promoters (BDPs) that differ in inducibility and strength can be used for fine-tuning and timing of expression of proteins in the Taxol pathway. Vogl et al. (2018) tested this concept and achieved a 50x-improvement in taxadiene titers. Lower—Glasscock et al. (2019) introduced riboregulated Switchable feedback promoter (rSFP) that uses a natural stress-response promoter (SRP) in combination with a second plasmid containing a small transcription activating RNA (STAR) that can be induced by an inducer like anhydrotetracycline (ATc) to create an external ON/OFF gate that achieved a 2.4-fold increase in oxygenated taxanes compared to Biggs et al. (2016a). **(C)** Compartmentalized engineering in plant tissue (*N. benthamiana*) targeting chloroplasts (Li et al., 2019) and cytosol (De La Peña and Sattely, 2020) was conducted. Plastid—chloroplastic targeting (blue arrows) resulted in production of T-5α-ol for the first time, while cytosolic targeting (red arrows) resulted in 10-fold increase in taxadiene. Purple arrows—upregulation of HMGR without cytosol-targeted tcTXS-tcGGPPS led to 80x increase in an unidentified compound that co-eluted with taxadiene. Black arrows—introduction of DXS and TXS-GGPPS without chloroplastic targeted chlT5αH-chlCPR results in detection of OCT only. Red dotted arrow, cytosolic engineering of taxadiene with further introduction of T5αH-CPR has not yet been attempted.

observed changes; S302A, K131R, and V374L (**Table 2**). Product promiscuity was slightly improved in the identified mutants, but T-5α-ol remained a minor product while OCT and iso-OCT dominated the product profile.

A different T5αH mechanism involving an epoxide intermediate, as opposed to an exclusive radical intermediate, was suggested by the Stephanopoulos group (Edgar et al., 2016). Under this proposal, the two main products of TXS (**T** and **T1**) undergo catalysis through different transition states en-route to T-5α-ol; **T1** follows the radical-rebound mechanism (abstraction at the C20- position), while **T** follows epoxidation route to an unstable epoxide intermediate that decomposes non-selectively to several products like OCT, iso-OCT, and T-5α-ol (**Figure 3A**). The proposal was supported by several lines of evidence, and it was experimentally determined that incubating taxadiene with TXS leads to several products, but incubating with iso-taxadiene produced a single peak of T-5α-H. Another independent chemical synthesis study provided direct evidence that indeed **T** can be epoxidized in a regio- and diastereoselective manner to yield the intermediate taxadiene-4(5)-epoxide that can be further rearranged into T-5α-ol, OCT, and iso-OCT (Barton et al., 2016). These observations are significant, and support that T5αH catalysis is substrate-driven, and that the broad product profile is a result of non-selective epoxide degradation rather than T5αH non-selectivity as previously assumed. Based on this alternative mechanistic proposal, three targets can be manipulated for enhancing selectivity of T5αH hydroxylation; (i) manipulating TXS selectivity to enhance **T1** in the product profile, (ii) engineering T5αH to improve turnover and (iii) searching for an epoxidase enzyme to improve selectivity of the hypothesized epoxidase intermediate (**Figure 3B**). All three approaches were conducted and TXS engineering to enhance **T1** titers resulted in variants that displayed 2.4-fold improvement in **T1** leading to a 2.4-fold improvement in T-5α-ol titers in *E. coli*

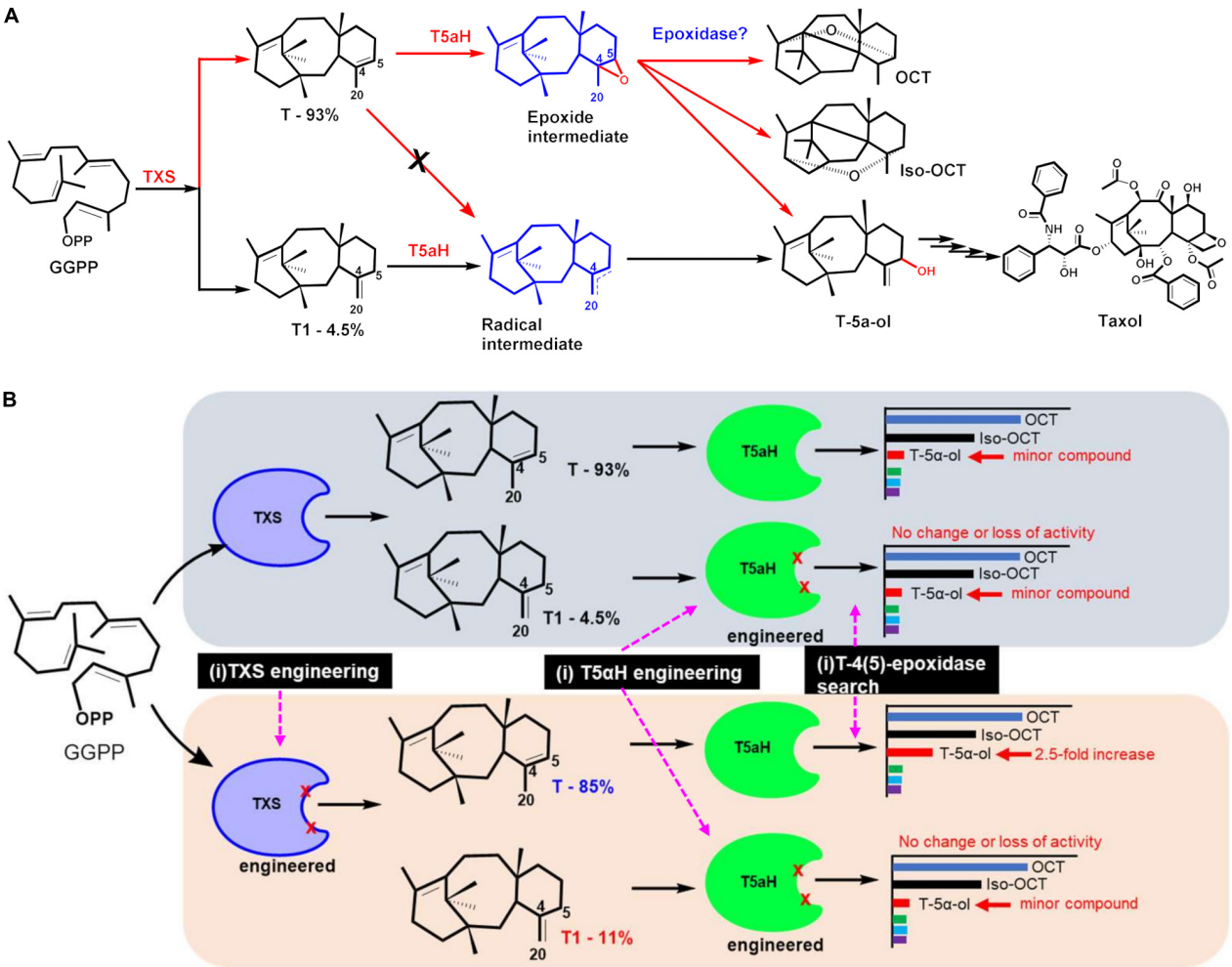


FIGURE 3 | Improving T5αH catalysis through mechanism—informed protein engineering strategies (based on Edgar et al., 2016, 2017). **(A)** A novel mechanism of T5αH catalysis was proposed, via an unstable epoxidase intermediate. Under this proposal, the mechanism of T5αH catalysis is substrate controlled. Taxadiene **(T)** undergoes catalysis through the proposed epoxide (red arrows), leading to several products like OCT, T-5α-ol and iso-OCT, while iso-taxadiene **(T1)** follows the radical rebound mechanism (black arrows) that exclusively leads to the desired T-5α-ol product and ultimately to Taxol. **(B)** Exploring engineering targets to improve T5αH catalytic specificity toward T-5α-ol. Three strategies: (i) TXS engineering to improve iso-taxadiene yield, (ii) T5αH engineering to improve catalytic efficiency and specificity and (iii) mining for an epoxide that could convert the epoxidase to T-5α-ol. Engineered TXS with improved T1 percentage improved T-5α-ol yield 2.5-fold.

after reconstituting T5αH-CPR (Edgar et al., 2017). A total of 14 residues consisting of polar residues near the TXS active site, residues closer to PPi, and two residues in the N-terminus that caps the active site were investigated by saturation mutagenesis, and the most successful mutant was Y688L. In fact, in addition to the critical PPi assisted abstraction, docking simulations also identified Y688 as a critical residue for abstraction of the hydrogen on C-5 position. Mutation of N-terminal residues led to a loss of activity, a result that was in agreement with Schrepfer et al. (2016). Mutagenesis of the P450 enzyme and mining of a novel T-4(5)-epoxidase were not successful, but with the increasing amount of *Taxus* transcriptome datasets and the substrate synthesis methodology reported by Barton and colleagues, further querying, *in vitro* screening and activity-guided fractionation are promising methods in searching for this enzyme. In addition to T-5α-ol-enhancing mutations, other TXS mutants were identified that yield verticillines (Q609G), or other unidentified compounds, most likely cembrene A (**Table 2**).

Semi-Rational Designing of DBAT

Besides the first two enzymes of the Taxol pathway, other downstream enzymes have also been targeted for engineering enhanced catalysis. For example, DBAT was engineered with the aid of a computational semi-rational strategy, leading to a double mutant (I43S/D390R) that not only improved catalytic efficiency but could more efficiently utilize vinyl acetate, a cheaper alternative acyl donor compared to acetyl CoA (Lin et al., 2018). Compared to the wild type DBAT, the double mutant improved catalytic efficiency by 3.3-fold when using vinyl acetate, and 2.99-fold when using acetyl CoA (**Table 2**). Employing a similar computationally guided semi-rational design, Li et al. (2017) employed structure

modeling, molecular docking, alanine scanning and saturation mutagenesis to engineer DBAT for acetylation of the C10 position of 10-deacetyltaxol (DT) with the aim to utilize the C7-gylcosylated Taxol analog (7-β-xylosyl-10-deacetyltaxol) for Taxol biosynthesis. The double mutant engineered in the study (G38R/F301V) demonstrated 6-fold improvement in catalytic efficiency (acetylation of DT to Taxol) compared to wild type DBAT, and in addition, a mutant (G38R) was identified with 2.15-fold improved catalytic efficiency in converting the native substate (10-deacetyl baccatin) to baccatin III. Molecular docking also revealed the critical catalytic role of His162 in DBAT catalysis, and alanine scanning identified 4 residues that led to a complete loss of activity when mutated to alanine (H162A, R363A, G361A, and I164A). Two of these residues (H162 and R363) and an additional active site residue (D166) were further investigated by another group by a similar computational strategy supported by site directed mutagenesis (You et al., 2018). Postulating the benefits of histidine residues in the DBAT catalytic pocket, these residues were mutated to histidine, leading to D166H, R363H, H162A/R363H, and D166H/R363H that demonstrated superior catalytic activities (15-, 26-, 3-, and 60-times improvements compared to wild type DBAT, respectively) (**Table 2**).

Computational tools are increasingly becoming useful in gaining mechanistic insights on the catalysis cycles of Taxol pathway enzymes. The above examples highlight how versatile and powerful the tool was leveraged for several proteins. Cytochrome P450s are very sensitive to mutations, thus computational and phylogenetically guided mutagenesis studies of T5αH have so far not yet identified significantly improved mutants for solving this bottleneck. Nevertheless, given the lack of a X-ray crystal structure of T5αH or any CYP450 of the Taxol pathway, homology modeling and molecular docking have thus far provided three models (Edgar et al., 2017; Rouck et al., 2017; Yadav, 2014) that can guide semi-rational design strategies. Several residues were identified in the resulting docking conformations that are potentially involved in stabilizing the hydrophobic taxadiene, reaction intermediates and the heme that could shed light on the mechanism of catalysis.

Metabolic Engineering in Plant Hosts

Plant hosts are expected to offer a more favorable environment for the challenging functional expression of plant CYP540 that currently frustrates engineering of the pathway in microbial hosts. Taxol pathway genes are located in different organelles in the plant cell (cytosol, endoplasmic reticulum membranes, and chloroplasts) but the nature of the exchange of intermediates among these different locations is not known. We considered inter-organellar transport of taxane intermediates to be the major barrier blocking access of endoplasmic reticulum-localized CYP450s to their diterpenoid substrate produced in the plastid in plant cells and used a compartmentalized engineering strategy that led to production of T-5α-ol in a heterologous plant host for the first time (Li et al., 2019; **Figure 2C**). A very recent report similarly targeted TXS and GGPPS in the cytosol and overexpressed the mevalonate pathway rate limiting enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) and observed a 10-fold improvement in taxadiene yields in *N. benthamiana* leaves (De La Peña and Sattely, 2020). Given the robustness of the mevalonate pathway compared to the MEP pathway, cytosolic targeting could potentially open new avenues for engineering taxanes in plant cells. It remains to be tested if this cytosolic compartmentalization strategy could be leveraged for engineering production of oxygenated taxanes.

NOVEL CANDIDATE GENES AND INSIGHTS FROM RECENT TRANSCRIPTOME DATA

Advances in sequencing technology in the last decade opened avenues for gene discovery and pathway elucidation of Taxol biosynthesis. Almost all novel genes identified in the early days were discovered through leveraging homology-based cloning, random sequencing of cDNA libraries from MeJA-elicited *Taxus* cell cultures, differential display of mRNA-reverse transcription-PCR, screening of EST libraries and use of available substrates and surrogate substrates (see reviews by Walker and Croteau, 2001; Kaspera et al., 2006). The advent of high throughput RNA sequencing technologies presented a very powerful tool that enabled transcriptome and genome sequencing, generating a huge amount of data that often is challenging to analyze and infer meaningful biological relevance. To date, powerful sequencing technologies using next-generation sequencing (NGS) and third-generation sequencing (TGS) platforms have been applied in transcriptomic studies of *Taxus* tissues and cell suspension cultures. Recent years have witnessed an increased interest in sequencing Taxol-related transcriptomes, revealing unprecedented insights into regulatory mechanisms of the pathway, unraveling mechanistic links to plant hormone signal transduction pathways and providing several lead candidates for the missing pathway genes—with some that have already been confirmed functional.

A transcriptome study was conducted with MeJA-elicited *T. baccata* suspension cells using high throughput complementary DNA-amplified fragment length polymorphism (cDNA-AFLP) that provided a total of 15 candidate transcripts identified as potential lead candidate genes encoding the six remaining enzymes (PCL, T1βH, T9O, C4-C20 epoxidase, T2′αH, and oxomutase) of the Taxol pathway. Functional characterization of these candidates led to the isolation of PCL (Ramírez-Estrada et al., 2016). A combined transcriptomic assembly of *Taxus chinensis* cultured cells and *in silico* mining of publicly available transcriptome data sets covered a comprehensive list of CYP450 genes, creating a valuable resource for searching the missing enzymes and for finding alternative P450s for bottleneck enzymes like T5αH (Liao et al., 2017). A total of 118 full length and 175 partial length *T. chinensis* P450s were identified, including the five known P450s of the pathway (CYP725A1—T10βH; CYP725A2—T13αH; CYP725A4—T5αH, CYP725A5—T7βH, CYP725A6—T2αH) and six novel CYP725A subfamily genes (CYP725A9, CYP725A11,

CYP725A16, CYP725A20, CYP725A22, CYP725A23). The same three sets of publicly available transcriptome datasets from *T. chinensis* cultured cells were mined for WRKY transcription factors leading to identification of 61 transcripts of TcWRKY of which six selected genes were all upregulated by MeJA (Zhang et al., 2018a).

Other recent Taxol-related transcriptomes covered a taxol-producing endophytic fungi *Cladosporium cladosporioides* MD2 (Miao et al., 2018), profiling of time-series reprogramming of *Taxus x media* genes following MeJA treatment (Mao et al., 2018), comparison of wild type *T. yunnanensis* with a high Taxol and 10-DAB—yielding new cultivar (He et al., 2018), comparative transcriptomes of *T. media*, *T. marei,* and *T. cuspidata* that differ in Taxol content (Zhou et al., 2019) and Iso-Seq of *T. cuspidata* tissues (Kuang et al., 2019). The Iso-Seq transcriptome identified nine CYP450s and seven acyl transferases (ACTs) as possible lead candidates for Taxol biosynthesis. The utility of transcriptome data sets in novel gene discovery, unraveling of biosynthetic and regulatory mechanisms was demonstrated in numerous examples discussed above. An in-depth computational and experimental analysis of the cDNA-AFLP dataset previously analyzed by Ramírez-Estrada et al. (2016) was recently conducted by the same group, leading to identification of transcript TB506 as a putative Taxane 2′a hydroxylase (T2′αH) (Sanchez-Muñoz et al., 2020). Molecular docking was conducted to confirm binding of such a huge substrate as 3′N-dehydroxydebenzoyltaxol and possible conformations were achieved. Functional expression and characterization of T2′αH activity was confirmed in *Pisum sativum* protoplasts, opening a biotechnological route to Taxol from its available intermediates 10-DAB and BIII.

In another recent study relevant to biotechnological production of Taxol, a library of 17 acyl CoA transferases (ACTs) was mined from three publicly available RNA-Seq data sets (from MeJA-elicited *T. media* suspension cell cultures) and screened for activation of different organic acids for N-substitution of 3-phenylisoserine side chain of taxoids (Srividya et al., 2020). In addition to identification of a candidate ACT with high specificity for generating CoA ester of benzoic acid (leading to Taxol formation) the study identified and functionally characterized several ACTs responsible for inserting different groups on this position, leading to several observed taxoids like Taxol B (insertion of a hexanoic acid), Taxol C (tiglic acid) and Taxol D (butyric acid). An ACT clone with 99% identity to the one described above (Ramírez-Estrada et al., 2016) was isolated, but activity screening of the clone, and all other candidates by Srividya et al. (2020) did not yield a positive hit.

Biochemical assays have been used to confirm several regulatory mechanisms inferred from deep sequencing studies. Using a GUS reporter assay with promoters of seven pathway genes from *Taxus cuspidata* cells, three basic helix-loop-helix (bHLH) transcription factors (TcMYC1, TcMYC2, and TcMYC4) were identified as negative regulators of MeJA-induced Taxol biosynthesis through their interaction with E-boxes in the promoters of Taxol pathway genes (Lenka et al., 2015). A more recent study, however, used GUS reporter assays in combination with yeast-one-hybrid, yeast-two-hybrid and *in vitro* assays and revealed TcMYC2a as a positive regulator of TXS in JA signaling (Zhang et al., 2018b). It relays its positive signal through binding JAZ proteins, and interacting with promoters of ERF15, ERF12, and TXS through the T/G-box, G-box, and E-box in their promoters.

PERSPECTIVES AND CONCLUDING REMARKS

Most genes of the Taxol pathway were isolated and their encoded proteins were functionally characterized is different systems. However, successes in functional expression of individual genes were not replicated when sequential genes were constructed into a pathway to synthesize intermediates. The most successful heterologous production of a taxane intermediate was 1 g/L achieved 10 years ago through an MMME approach, but introduction of T5αH that catalyzes the second step to the strain led to a dramatic loss of both optimality and titers. The past 5 years witnessed a significant increase in application of synthetic biology tools together with the emerging enabling technologies for gene assemblies in enzyme discovery and metabolic engineering of taxane intermediates. Several synthetic biology tools have been leveraged to optimize T5αH expression and activity, including truncations, promoter optimization, CPR optimization and compartmentalized engineering in plant organelles and use of riboregulated switchable feedback promoters (rSFPs). Computational and experimental approaches were used to improve our understanding of the catalytic mechanism of TXS and T5αH, shedding more light on the transition states governing the observed product profile. Models of the closed TXS and those of T5αH and DBAT were also built that provided intricate details of the active site architecture to guide semi-rational protein engineering strategies to improve catalytic activity and alter product profile.

As highlighted throughout this review, most strategies in the past 5 years were focused on T5αH. This is because this enzyme catalyzes the most important bottleneck of the pathway. With a taxadiene conversion rate of less than 10%, and a product profile comprising several monooxygenated and few dioxygenated taxanes, it's not surprising that strategies that aimed to increase supply of precursors, or those aiming to improve catalytic activity did not achieve much improvements since this is a major branching point of the pathway that splits flux into several off-target taxanes. The successes recorded through optimized expression (operon constructs, use of low strength promoters and low copy plasmids), a clever approach to couple expression of the protein to cell envelope stress through rSFPs, compartmentalized engineering in chloroplasts of plant cells and engineering of TXS to favor iso-taxadiene product that proved to exclusively lead to T-5α-ol are approaches that we expect to be further developed in the near future, utilizing such enabling tools as the CRISPR/Cas9 toolkit developed for *S. cerevisiae*. *Nicotiana benthamiana*, a high biomass plant with available technologies for DNA manipulation and agrobacterium-mediated transient expression systems is very promising as a chassis for heterologous

expression of the Taxol pathway and is increasingly being favored for production of many other terpenoids. The Taxol pathway that comprises nine CYP450s that trigger membrane envelope stress on their expression in host cells as demonstrated with T5αH, is a very good candidate for expression in *N. benthamiana*.

Advances is high throughput sequencing technologies have enabled generation of several Taxol-related transcriptomes, and recent mining of these publicly available resources have led to isolation of two missing genes, PCL and T2′αH, though wide testing of PCL in different heterologous hosts is yet to be conclusive. Dozens of lead candidate genes for the missing steps were identified through mining transcriptomes, and with more screening platforms being established, we anticipate full elucidation of the pathway in the near future. Furthermore, we anticipate gas phase QM/MM calculations and experimental methods to reveal the mechanism of catalysis of T5αH that will guide protein engineering to overcome the bottleneck, as was done for TXS and DBAT. As sequencing technology advances in the coming years, we also expect a high-resolution genome of *Taxus* species to be assembled that will complement transcriptome data sets and accelerate gene discovery of the remaining CYP450s, PCL, and an epoxidase. Synthetic biology tools are expected to play an increasingly important role in enzyme discovery, construction and optimization of the pathway in different chassis and silencing of competing pathways.

AUTHOR CONTRIBUTIONS

IM and JL drafted the manuscript with support from FX. IM, JL, FX, and YW revised the manuscript. YW supervised this work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Key R&D Program of China (2018YFA0900600), the National Natural Science Foundation of China (Grant nos. 22077129, 31670099, 31700261, and 41876084) and Research Program of State Key Laboratory of Bioreactor Engineering, the Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-002-15), the Program of Shanghai Academic Research Leader (20XD1404400), the Strategic Priority Research Program "Molecular mechanism of Plant Growth and Development" of CAS (XDB27020202 and XDB27020103), the Construction of the Registry and Database of Bioparts for Synthetic Biology of the Chinese Academy of Sciences (No. ZSYS-016), the International Partnership Program of Chinese Academy of Sciences (No. 153D31KYSB20170121), and the National Key Laboratory of Plant Molecular Genetics, SIPPE, CAS.

REFERENCES

Abdallah, I. I., Pramastya, H., Van Merkerk, R., and Sukrasno Quax, W. J. (2019). Metabolic engineering of *Bacillus subtilis* toward taxadiene biosynthesis as the first committed step for taxol production. *Front. Microbiol.* 10:218. doi: 10.3389/fmicb.2019.00218

Ajikumar, P. K., Xiao, W. H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., et al. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 330, 70–74. doi: 10.1126/science.1191652

Ansbacher, T., Freud, Y., and Major, D. T. (2018). Slow-starter enzymes: role of active-site architecture in the catalytic control of the biosynthesis of taxadiene by Taxadiene synthase. *Biochemistry* 57, 3773–3779. doi: 10.1021/acs.biochem.8b00452

Anterola, A., Shanle, E., Perroud, P. F., and Quatrano, R. (2009). Production of taxa-4(5),11(12)-diene by transgenic *Physcomitrella patens*. *Transgen. Res.* 18, 655–660. doi: 10.1007/s11248-009-9252-5

Apel, A. R., Espaux, L., Wehrs, M., Sachs, D., Li, A., Tong, G. J., et al. (2017). A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 45, 496–508. doi: 10.1093/nar/gkw1023

Barton, N. A., Marsh, B. J., Lewis, W., Narraidoo, N., Seymour, G. B., Fray, R., et al. (2016). Accessing low-oxidation state taxanes: is taxadiene-4(5)-epoxide on the Taxol biosynthetic pathway? *Chem. Sci.* 7, 3102–3107. doi: 10.1039/c5sc03463a

Besumbes, Ó, Sauret-Güeto, S., Phillips, M. A., Imperial, S., Rodríguez-Concepción, M., and Boronat, A. (2004). Metabolic engineering of isoprenoid biosynthesis in *Arabidopsis* for the production of taxadiene, the first committed precursor of taxol. *Biotechnol. Bioeng.* 88, 168–175. doi: 10.1002/bit.20237

Bian, G., Yuan, Y., Tao, H., Shi, X., Zhong, X., Han, Y., et al. (2017). Production of taxadiene by engineering of mevalonate pathway in *Escherichia coli* and endophytic fungus Alternaria alternata TPF6. *Biotechnol. J.* 12, 1–11. doi: 10.1002/biot.201600697

Biggs, B. W., Lim, C. G., Sagliani, K., Shankar, S., Stephanopoulos, G., De Mey, M., et al. (2016a). Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3209–3214. doi: 10.1073/pnas.1515826113

Biggs, B. W., Rouck, J. E., Kambalyal, A., Arnold, W., Lim, C. G., De Mey, M., et al. (2016b). Orthogonal assays clarify the oxidative biochemistry of Taxol P450 CYP725A4. *ACS Chem. Biol.* 11, 1445–1451. doi: 10.1021/acschembio.5b00968

Cha, M., Shim, S. H., Kim, S. H., Kim, O. T., Lee, S. W., Kwon, S. Y., et al. (2012). Production of taxadiene from cultured ginseng roots transformed with Taxadiene synthase gene. *BMB Rep.* 45, 589–594. doi: 10.5483/BMBRep.2012.45.10.085

Choi, H. K., Kim, S. I., Son, J. S., Hong, S. S., Lee, H. S., and Lee, H. J. (2000). Enhancement of paclitaxel production by temperature shift in suspension culture of *Taxus chinensis*. *Enzyme Microb. Technol.* 27, 593–598. doi: 10.1016/S0141-0229(00)00255-6

De La Peña, R., and Sattely, E. S. (2020). Rerouting plant terpene biosynthesis enables momilactone pathway elucidation. *Nat. Chem. Biol.* 17, 205–212. doi: 10.1038/s41589-020-00669-3

DeJong, J. H. M., Liu, Y., Bollon, A. P., Long, R. M., Jennewein, S., Williams, D., et al. (2006). Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 93, 212–224. doi: 10.1002/bit.20694

Ding, M. Z., Yan, H. F., Li, L. F., Zhai, F., Shang, L. Q., Yin, Z., et al. (2014). Biosynthesis of taxadiene in *Saccharomyces cerevisiae*: selection of geranylgeranyl diphosphate synthase directed by a computer-aided docking strategy. *PLoS One* 9:e0109348. doi: 10.1371/journal.pone.0109348

Edgar, S., Li, F. S., Qiao, K., Weng, J. K., and Stephanopoulos, G. (2017). Engineering of Taxadiene synthase for improved selectivity and yield of a key taxol biosynthetic intermediate. *ACS Synth. Biol.* 6, 201–205. doi: 10.1021/acssynbio.6b00206

Edgar, S., Zhou, K., Qiao, K., King, J. R., Simpson, J. H., and Stephanopoulos, G. (2016). Mechanistic insights into Taxadiene epoxidation by taxadiene-5α-hydroxylase. *ACS Chem. Biol.* 11, 460–469. doi: 10.1021/acschembio.5b00767

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

Escorcia, A. M., van Rijn, J. P. M., Cheng, G. J., Schrepfer, P., Brück, T. B., and Thiel, W. (2018). Molecular dynamics study of Taxadiene synthase catalysis. *J. Comput. Chem.* 39, 1215–1225. doi: 10.1002/jcc.25184

Fett-Netto, A. G., DiCosmo, F., Reynolds, W. F., and Sakata, K. (1992). Cell culture of *Taxus* as a source of the antineoplastic drug taxol and related taxanes. *Nat. B* 10, 1572–1575. doi: 10.1038/nbt1292-1572

Freud, Y., Ansbacher, T., and Major, D. T. (2017). Catalytic control in the facile proton transfer in Taxadiene synthase. *ACS Catal.* 7, 7653–7657. doi: 10.1021/acscatal.7b02824

Glasscock, C. J., Lazar, J. T., Biggs, B. W., Arnold, J. H., Kang, M. K., Tullman-ercek, D., et al. (2019). Dynamic control of pathway expression with riboregulated switchable feedback promoters. *bioRxiv* [Preprint]. doi: 10.1101/529180

Gutta, P., and Tantillo, D. J. (2007). A promiscuous proton in taxadiene biosynthesis? *Org. Lett.* 9, 1069–1071. doi: 10.1021/ol070007m

He, C. T., Li, Z. L., Zhou, Q., Shen, C., Huang, Y. Y., Mubeen, S., et al. (2018). Transcriptome profiling reveals specific patterns of paclitaxel synthesis in a new *Taxus yunnanensis* cultivar. *Plant Physiol. Biochem.* 122, 10–18. doi: 10.1016/j.plaphy.2017.10.028

Hefner, J., Rubenstein, S. M., Ketchum, R. E. B., Gibson, D. M., Williams, R. M., and Croteau, R. (1996). Cytochrome P450-catalyzed hydroxylation of taxa-4(5),11(12)-diene to taxa-4(20),11(12)-dien-5a-o1: the first oxygenation step in taxol biosynthesis. *Chem. Biol.* 3, 479–489. doi: 10.1016/S1074-5521(96)90096-4

Hezari, M., Lewis, N. G., and Croteau, R. (1995). Purification and Characterization of Taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of Taxol biosynthesis. *Arch. Biochem. Biophys.* 322, 437–444.

Holton, R. A., Somoza, C., Kim, H. B., Liang, F., Biediger, R. J., Boatman, P. D., et al. (1994). First total synthesis of taxol. 1. Functionalization of the B ring. *J. Am. Chem. Soc.* 116, 1597–1598. doi: 10.1021/ja00083a066

Hong, Y. J., and Tantillo, D. J. (2011). The taxadiene-forming carbocation cascade. *J. Am. Chem. Soc.* 133, 18249–18256. doi: 10.1021/ja2055929

Huang, Q., Roessner, C. A., Croteaul, R., Scott, A. I., Huang, Q. L., and Roessner, C. (2001). Engineering *Escherichia coli* for the synthesis of taxadiene. *Bioorgan. Med. Chem.* 9, 2237–2242.

Jennewein, S., Long, R. M., Williams, R. M., and Croteau, R. (2004). Cytochrome P450 taxadiene 5a-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. *Chem. Biol.* 11, 379–387. doi: 10.1016/j

Jennewein, S., Rithner, C. D., Williams, R. M., and Croteau, R. B. (2001). Taxol biosynthesis: taxane 13 alpha-hydroxylase is a cytochrome P450-dependent monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13595–13600. doi: 10.1073/pnas.251539398

Kaspera, R., and Croteau, R. (2006). Cytochrome P450 oxygenases of taxol biosynthesis. *Phytochem. Rev.* 5, 433–444. doi: 10.1007/s11101-006-9006-4

Kaspera, R., Croteau, R., Rodney, C., and Kaspera, R. (2006). Cytochrome P450 oxygenases of taxol biosynthesis. *Phytochem. Rev.* 5, 433–444.

Ketchum, R. E. B., and Croteau, R. B. (2006). The *Taxus* metabolome and the elucidation of the taxol® biosynthetic pathway in cell suspension cultures. *Biotechnol. Agric. Forest* 57, 291–309. doi: 10.1007/3-540-29782-0_21

Koepp, A. E., Hezari, M., Zajicek, J., Vogel, B. S., LaFever, R. E., Lewis, N. G., et al. (1995). Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of taxol biosynthesis in Pacific yew. *J. Biol. Chem.* 270, 8686–8690. doi: 10.1074/JBC.270.15.8686

Köksal, M., Jin, Y., Coates, R. M., Croteau, R., and Christianson, D. W. (2011). Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* 469, 116–122. doi: 10.1038/nature09628

Kovacs, K., Zhang, L., Linforth, R. S. T., Whittaker, B., Hayes, C. J., and Fray, R. G. (2007). Redirection of carotenoid metabolism for the efficient production of taxadiene [taxa-4(5),11(12)-diene] in transgenic tomato fruit. *Transgen. Res.* 16, 121–126. doi: 10.1007/s11248-006-9039-x

Kuang, X., Sun, S., Wei, J., Li, Y., and Sun, C. (2019). Iso-Seq analysis of the *Taxus cuspidata* transcriptome reveals the complexity of Taxol biosynthesis. *BMC Plant Biol.* 19:210. doi: 10.1186/s12870-019-1809-8

Lenka, S. K., Nims, N. E., Vongpaseuth, K., Boshar, R. A., Roberts, S. C., and Walker, E. L. (2015). Jasmonate-responsive expression of paclitaxel biosynthesis genes in *Taxus cuspidata* cultured cells is negatively regulated by the bHLH transcription factors TcJAMYC1, TcJAMYC2, and TcJAMYC4. *Front. Plant Sci.* 6:115. doi: 10.3389/fpls.2015.00115

Li, B. J., Wang, H., Gong, T., Chen, J. J., Chen, T. J., Yang, J. L., et al. (2017). Improving 10-deacetylbaccatin III-10-β-O-acetyltransferase catalytic fitness for Taxol production. *Nat. Commun.* 8, 1–13. doi: 10.1038/ncomms15544

Li, J., Mutanda, I., Wang, K., Yang, L., Wang, J., and Wang, Y. (2019). Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in *Nicotiana benthamiana*. *Nat. Commun.* 10, 1–12. doi: 10.1038/s41467-019-12879-y

Liao, W., Zhao, S., Zhang, M., Dong, K., Chen, Y., Fu, C., et al. (2017). Transcriptome assembly and systematic identification of novel cytochrome P450s in *Taxus chinensis*. *Front. Plant Sci.* 8:1468. doi: 10.3389/fpls.2017.01468

Lin, S. L., Wei, T., Lin, J. F., Guo, L. Q., Wu, G. P., Wei, J., et al. (2018). Bio-production of Baccatin III, an important precursor of paclitaxel by a cost-effective approach. *Mol. Biotechnol.* 60, 492–505. doi: 10.1007/s12033-018-0090-7

Lin, X., Hezari, M., Koepp, A. E., Floss, H. G., and Croteau, R. (1996). Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of taxol biosynthesis in Pacific yew. *Biochemistry* 35, 2968–2977. doi: 10.1021/bi9526239

Malik, S., Cusidó, R. M., Mirjalili, M. H., Moyano, E., Palazón, J., and Bonfill, M. (2011). Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: a review. *Process Biochem.* 46, 23–34. doi: 10.1016/j.procbio.2010.09.004

Mao, R., Chen, J., Chen, Y., and Guo, Z. (2018). Identification of early jasmonate-responsive genes in *Taxus* × media cells by analyzing time series digital gene expression data. *Physiol. Mol. Biol. Plants* 24, 715–727. doi: 10.1007/s12298-018-0527-2

McElroy, C., and Jennewein, S. (2017). "Taxol® biosynthesis and production: from forests to fermenters," in *Biotechnology of Natural Products*, eds W. Schwab, B. Lange, and M. Wüst (Cham: Springer), 145–185. doi: 10.1007/978-3-319-67903-7_7

Miao, L. Y., Mo, X. C., Xi, X. Y., Zhou, L., De, G., Ke, Y. S., et al. (2018). Transcriptome analysis of a taxol-producing endophytic fungus *Cladosporium cladosporioides* MD2. *AMB Express* 8:41. doi: 10.1186/s13568-018-0567-6

Mountford, P. G. (2010). "The Taxol ® story - development of a green synthesis via plant," in *Green Chemistry in the Pharmaceutical Industry*, eds P. J. Dunn, A. S. Wells, and M. T. Williams (Weinheim: Wiley-VCH Verlag GmbH & Co), 145–160.

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

Nowrouzi, B., Li, R., Walls, L. E., D'Espaux, L., Malci, K., Lungang, L., et al. (2020). Enhanced production of taxadiene in *Saccharomyces cerevisiae*. *Microb. Cell Fact* 19:200.

Ramírez-Estrada, K., Altabella, T., Onrubia, M., Moyano, E., Notredame, C., Osuna, L., et al. (2016). Transcript profiling of jasmonate-elicited *Taxus* cells reveals a β-phenylalanine-CoA ligase. *Plant Biotechnol. J.* 14, 85–96. doi: 10.1111/pbi.12359

Rontein, D., Onillon, S., Herbette, G., Lesot, A., Werck-Reichhart, D., Sallaud, C., et al. (2008). CYP725A4 from yew catalyzes complex structural rearrangement of taxa-4(5),11(12)-diene into the cyclic ether 5(12)-oxa-3(11)-cyclotaxane. *J. Biol. Chem.* 283, 6067–6075. doi: 10.1074/jbc.M708950200

Rouck, J. E., Biggs, B. W., Kambalyal, A., Arnold, W. R., De Mey, M., Ajikumar, P. K., et al. (2017). Heterologous expression and characterization of plant Taxadiene-5α-Hydroxylase (CYP725A4) in *Escherichia coli*. *Protein Expr. Purif.* 132, 60–67. doi: 10.1016/j.pep.2017.01.008

Sagwan-Barkdoll, L., and Anterola, A. M. (2017). Taxadiene-5α-ol is a minor product of CYP725A4 when expressed in *Escherichia coli*. *Biotechnol. Appl. Biochem.* 65, 294–305. doi: 10.1002/bab.1606

Sanchez-Muñoz, R., Perez-Mata, E., Almagro, L., Cusido, R. M., Bonfill, M., Palazon, J., et al. (2020). A novel hydroxylation step in the Taxane biosynthetic pathway: a new approach to Paclitaxel production by synthetic biology. *Front. Bioeng. Biotechnol.* 8:410. doi: 10.3389/fbioe.2020.00410

Schoendorf, A., Rithner, C. D., Williams, R. M., and Croteau, R. B. (2001). Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1501–1506. doi: 10.1073/pnas.98.4.1501

Schrepfer, P., Buettner, A., Goerner, C., Hertel, M., Rijn, J., Van Wallrapp, F., et al. (2016). Identification of amino acid networks governing catalysis in the

closed complex of class I Terpene synthases. *Proc. Natl. Acad. Sci. U.S.A.* 113, E958–E967. doi: 10.1073/pnas.1519680113

Srividya, N., Lange, I., Hartmann, M., Li, Q., Mirzaei, M., and Lange, B. M. (2020). Biochemical characterization of acyl activating enzymes for side chain moieties of Taxol and its analogs. *J. Biol. Chem.* 295, 4963–4973. doi: 10.1074/jbc.RA120.012663

van Rijn, J. P. M., Escorcia, A. M., and Thiel, W. (2019). QM/MM study of the Taxadiene synthase mechanism. *J. Comput. Chem.* 40, 1902–1910. doi: 10.1002/jcc.25846

Vogl, T., Kickenweiz, T., Pitzer, J., Sturmberger, L., Weninger, A., Biggs, B. W., et al. (2018). Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nat. Commun.* 9:3589. doi: 10.1038/s41467-018-05915-w

Walker, K., and Croteau, R. (2001). Taxol biosynthetic genes. *Phytochemistry* 58, 1–7. doi: 10.1016/S0031-9422(01)00160-1

Walker, K., Schoendorf, A., and Croteau, R. (2000). Molecular cloning of a taxa-4(20),11(12)-dien-5α-ol-O-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Arch. Biochem. Biophys.* 374, 371–380. doi: 10.1006/abbi.1999.1609

Wani, M. C., and Horwitz, S. B. (2014). Nature as a remarkable chemist: a personal story of the discovery and development of Taxol. *Anticancer. Drugs* 25, 482–487. doi: 10.1097/CAD.0000000000000063

Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., and Mcphail, A. T. (1971). Plant antitumor agents.VI. The isolation and structure of Taxol, a novel antileukemic and antitumor agent from *Taxus* brevifolia2. *J. Am. Chem. Soc.* 93, 2325–2327. doi: 10.1021/ja00738a045

Wildung, M. R., and Croteau, R. (1996). A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. *J. Biol. Chem.* 271, 9201–9204. doi: 10.1074/jbc.271.16.9201

Williams, D. C., Carroll, B. J., Jin, Q., Rithner, C. D., Lenger, S. R., Floss, H. G., et al. (2000a). Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of taxol catalyzed by recombinant Taxadiene synthase. *Chem. Biol.* 7, 969–977. doi: 10.1016/S1074-5521(00)00046-6

Williams, D. C., Wildung, M. R., Jin, A. Q., Dalal, D., Oliver, J. S., Coates, R. M., et al. (2000b). Heterologous expression and characterization of a "Pseudomature" form of Taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the Multistep Diterpene Cyclization reaction. *Arch. Biochem. Biophys.* 379, 137–146. doi: 10.1006/abbi.2000.1865

Yadav, V. G. (2014). Unraveling the multispecificity and catalytic promiscuity of taxadiene monooxygenase. *J. Mol. Catal. B Enzym.* 110, 154–164. doi: 10.1016/j.molcatb.2014.10.004

You, L. F., Huang, J. J., Wei, T., Lin, S. L., Jiang, B. H., Guo, L. Q., et al. (2018). Enhanced catalytic activities and modified substrate preferences for taxoid 10β-O-acetyl transferase mutants by engineering catalytic histidine residues. *Biotechnol. Lett.* 40, 1245–1251. doi: 10.1007/s10529-018-2573-9

Zhang, M., Chen, Y., Nie, L., Jin, X., Liao, W., Zhao, S., et al. (2018a). Transcriptome-wide identification and screening of WRKY factors involved in the regulation of taxol biosynthesis in *Taxus chinensis*. *Sci. Rep.* 8, 1–12. doi: 10.1038/s41598-018-23558-1

Zhang, M., Jin, X., Chen, Y., Wei, M., Liao, W., Zhao, S., et al. (2018b). TcMYC2a, a basic helix-loop-helix transcription factor, transduces JA-signals and regulates taxol biosynthesis in *Taxus chinensis*. *Front. Plant Sci.* 9:863. doi: 10.3389/fpls.2018.00863

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

Zhou, T., Luo, X., Yu, C., Zhang, C., Zhang, L., Song, Y. B., et al. (2019). Transcriptome analyses provide insights into the expression pattern and sequence similarity of several taxol biosynthesis-related genes in three *Taxus* species. *BMC Plant Biol.* 19:33. doi: 10.1186/s12870-019-1645-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 11 February 2021
doi: 10.3389/fbioe.2021.631863

Bornyl Diphosphate Synthase From *Cinnamomum burmanni* and Its Application for (+)-Borneol Biosynthesis in Yeast

Rui Ma[1,2†], Ping Su[1,3†], Juan Guo[1], Baolong Jin[1], Qing Ma[1], Haiyan Zhang[1], Lingli Chen[1], Liuying Mao[1], Mei Tian[1], Changjiangsheng Lai[1], Jinfu Tang[1], Guanghong Cui[1] and Luqi Huang[1,2]**

[1] State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China, [2] School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou, China, [3] Department of Chemistry, The Scripps Research Institute, Jupiter, FL, United States

OPEN ACCESS

Edited by:
Jingwen Zhou,
Jiangnan University, China

Reviewed by:
Yu Deng,
Jiangnan University, China
Jingyu Wang,
Westlake Institute for Advanced Study (WIAS), China

***Correspondence:**
Guanghong Cui
guanghongcui@163.com
Luqi Huang
huangluqi01@126.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 21 November 2020
Accepted: 25 January 2021
Published: 11 February 2021

Citation:
Ma R, Su P, Guo J, Jin B, Ma Q, Zhang H, Chen L, Mao L, Tian M, Lai C, Tang J, Cui G and Huang L (2021) Bornyl Diphosphate Synthase From Cinnamomum burmanni and Its Application for (+)-Borneol Biosynthesis in Yeast. Front. Bioeng. Biotechnol. 9:631863. doi: 10.3389/fbioe.2021.631863

(+)-Borneol is a desirable monoterpenoid with effective anti-inflammatory and analgesic effects that is known as soft gold. (+)-bornyl diphosphate synthase is the key enzyme in the (+)-borneol biosynthesis pathway. Despite several reported (+)-bornyl diphosphate synthase genes, relatively low (+)-borneol production hinders the attempts to synthesize it using microbial fermentation. Here, we identified the highly specific (+)-bornyl diphosphate synthase CbTPS1 from *Cinnamomum burmanni*. An *in vitro* assay showed that (+)-borneol was the main product of CbTPS1 (88.70% of the total products), and the K_m value was 5.11 ± 1.70 μM with a k_{cat} value of 0.01 s^{-1}. Further, we reconstituted the (+)-borneol biosynthetic pathway in *Saccharomyces cerevisiae*. After tailored truncation and adding Kozak sequences, the (+)-borneol yield was improved by 96.33-fold to 2.89 mg·L^{-1} compared with the initial strain in shake flasks. This work is the first reported attempt to produce (+)-borneol by microbial fermentation. It lays a foundation for further pathway reconstruction and metabolic engineering production of this valuable natural monoterpenoid.

Keywords: (+)-borneol, (+)-bornyl diphosphate synthase, *Cinnamomum burmanni*, metabolic engineering, *Saccharomyces cerevisiae*

INTRODUCTION

The monoterpene borneol is a highly desirable natural product widely used in medicine, spice, and chemical fields since ancient times (Wojtunik-Kulesza et al., 2019). It has a broad spectrum of bidirectional regulation on the central nervous system (Zhang et al., 2017; Zheng et al., 2018); anti-inflammatory (Zou et al., 2017; Ji et al., 2020) and antimicrobial activities (Xin et al., 2020); and increases biofilm barrier permeability (Song et al., 2018; Chen et al., 2019). Borneol is divided into (+)-borneol and (−)-borneol according to optical rotations. Natural (+)-borneol has primarily

Abbreviations: BPPS, bornyl diphosphate synthase; CIAP, calf intestinal alkaline phosphatase; DMAPP, dimethylallyl diphosphate; GC-MS, gas chromatography coupled with mass spectrometry; GPP, geranyl diphosphate; IPP, Isopentenyl diphosphate; MVA, mevalonate pathway; ORF, open reading frame; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

been extracted from *Cinnamomum camphora* (L.) Presl and *C. burmanni* (Nees et T.Nees) Blume (borneol-type) since the 1980s in China (Chen et al., 2010). However, the slow growth rate, low (+)-borneol levels, and restricted cultivation area mean that the yield of natural (+)-borneol is far from meeting the market demand. Borneol synthesized by chemical methods thus occupies most of the market share; however, a certain number of toxic compounds, such as isoborneol may exist in synthetic borneol. Thus, it is necessary to explore other methods to produce natural (+)-borneol.

Due to the clear genetic background and lack of susceptibility to phage infections, *Saccharomyces cerevisiae* is the preferred host for metabolic engineering (Kirby and Keasling, 2009; Liu et al., 2019; Nielsen, 2019). Many monoterpenoids, such as geraniol, limonene, linalool, and α-terpineol (**Figure 1**) have been produced in *S. cerevisiae* (Amiri et al., 2016; Cao et al., 2016, 2017; Zhang et al., 2019, 2020). Isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) derived from the mevalonate pathway (MVA) are the precursors of all terpenoids in *S. cerevisiae*, and geranyl diphosphate (GPP) is the direct precursor of monoterpenes catalyzed by farnesyl diphosphate synthase (ERG20) (Jiang et al., 2017). Thus, in order to produce monoterpenes in yeast, ERG20 is usually mutated or rationally designed into GPP synthase (Ignea et al., 2014; Zhao et al., 2016; Jiang et al., 2017). The upstream MVA pathway genes, *tHMG1* and *IDI1*, are frequently overexpressed in yeast (Zhao et al., 2016; Zhang et al., 2019). Guo et al. (2018) even overexpressed all MVA pathway genes (*ERG10*, *ERG13*, *tHMG1*, *ERG12*, *ERG8*, *ERG19*, *IDI1*, *ERG20*) to increase the GPP pool. Modification of the monoterpene synthase, including translational fusion and truncation of transit peptides at the N-terminus of the enzymes, is also an effective strategy for increasing the production of terpenes (Jongedijk et al., 2015; Ignea et al., 2019; Hu et al., 2020).

Like other monoterpenes, bornyl diphosphate synthase (BPPS) is the key enzyme involved in (+)-borneol biosynthesis. It catalyzes the universal precursor GPP to form (+)-bornyl diphosphate, and is then dephosphorylated to produce the target product (+)-borneol (**Figure 1**). BPPSs have been identified from several plants, including *Salvia officinalis* (SBS), *Lavandula angustifolia* (LaBPPS), *Lippia dulcis* (LdBPPS), and *Amomum villosum* (AvBPPS) (Wise et al., 1998; Despinasse et al., 2017; Hurd et al., 2017; Wang et al., 2018). However, all these enzymes produced multiple products, such as α-pinene, β-pinene, camphene, and limonene, with the largest amount of (+)-borneol produced by SBS, accounting for 57.8% of the total products. However, there have been no attempts to produce this valuable product by microbial cell factories.

Here we report a high-specificity (+)-borneol BBPS gene (*CbTPS1*) from *C. burmanni*. Among the products with GPP as substrate in an *in vitro* assay, (+)-borneol accounted for 88.70% of the total. We thus aimed to construct a (+)-borneol biosynthesis pathway in *S. cerevisiae*. To reach the target, eight genes involved in the MVA pathway were overexpressed. Truncated transit peptides and adding the Kozak sequence of CbTPS1 further improved the (+)-borneol production. Our work provides a good example for (+)-borneol production in microbial fermentation.

MATERIALS AND METHODS

Plant Materials and Chemicals

Leaves of *Cinnamomum burmanni* (Nees et T.Nees) Blume were obtained from Guangdong Huaqingyuan Technology Co., Ltd. Plant leaf material grown in natural conditions was picked in May 2019. *C. burmannii* was identified by Prof. Cui Guanghong of China Academy of Chinese Medical Sciences and stored at −80°C for further usage (Storage Number: YXS201905). GPP, geraniol, α-pinene, β-pinene, α-phellandrene, limonene, α-terpineol, (+)-borneol and (−)-borneol standards were purchased from Sigma-Aldrich Chemical Co., United States.

RNA Extraction, cDNA Synthesis

The total RNA from *C. burmannii* leaves was extracted using a quick RNA isolation kit (HuaYueYang Biotechnology, China) based on the manufacturer's protocol, and then digested and purified by RNase-free DNase I (TaKaRa, Japan). An aliquot containing 1 μg total RNA was used to synthesize the first-strand cDNA with TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotechnology, China) according to the manufacturer's guidelines.

BPPS Candidate Selection and Analysis

Transcriptomic libraries of the *C. burmannii* leaves were shipped to the Novogene Company[1] for library construction and RNA-seq. The Illumina-derived nucleotide sequences reported in this paper have been submitted to China National Center for Bioinformation[2] under accession number CRA003558. To mine the BPPS candidate genes, TBLASTN analysis of BPPSs in the *C. burmannii* transcriptome was carried out using BioEdit software (Su et al., 2018). *SBS* (GenBank Accession Number: AAC26017), *LaBPPS* (GenBank Accession Number: AJW68082), *LdBPPS* (GenBank Accession Number: ATY48638), and *AvBPPS* (GenBank Accession Number: AWW87313) were used as the query sequences. The *CbTPS1* (GenBank Accession Number: MW196671) sequence was analyzed using NCBI[3]. The open reading frames (ORFs) were identified using the ORF Finder[4], and deduced amino acid sequences were identified using ExPASy[5]. Multiple sequence alignments were conducted using CLC Bio Sequence Viewer 6[6]. The chloroplast transit peptide of CbTPS1 was predicted by ChloroP[7].

All statistical analyses were conducted using SPSS version 23.0 (SPSS Inc., Chicago, IL, United States) for windows. One-way analysis of variance was used to compare the mean difference in (+)-borneol of strains. The *P*-value of less than 0.05 considered statistically significant.

[1]https://www.novogene.com/
[2]https://bigd.big.ac.cn/
[3]http://www.ncbi.nlm.nih.gov/
[4]http://www.ncbi.nlm.nih.gov/gorf/gorf.html
[5]http://web.expasy.org/translate/
[6]http://www.clcbio.com
[7]http://www.cbs.dtu.dk/services/ChloroP/

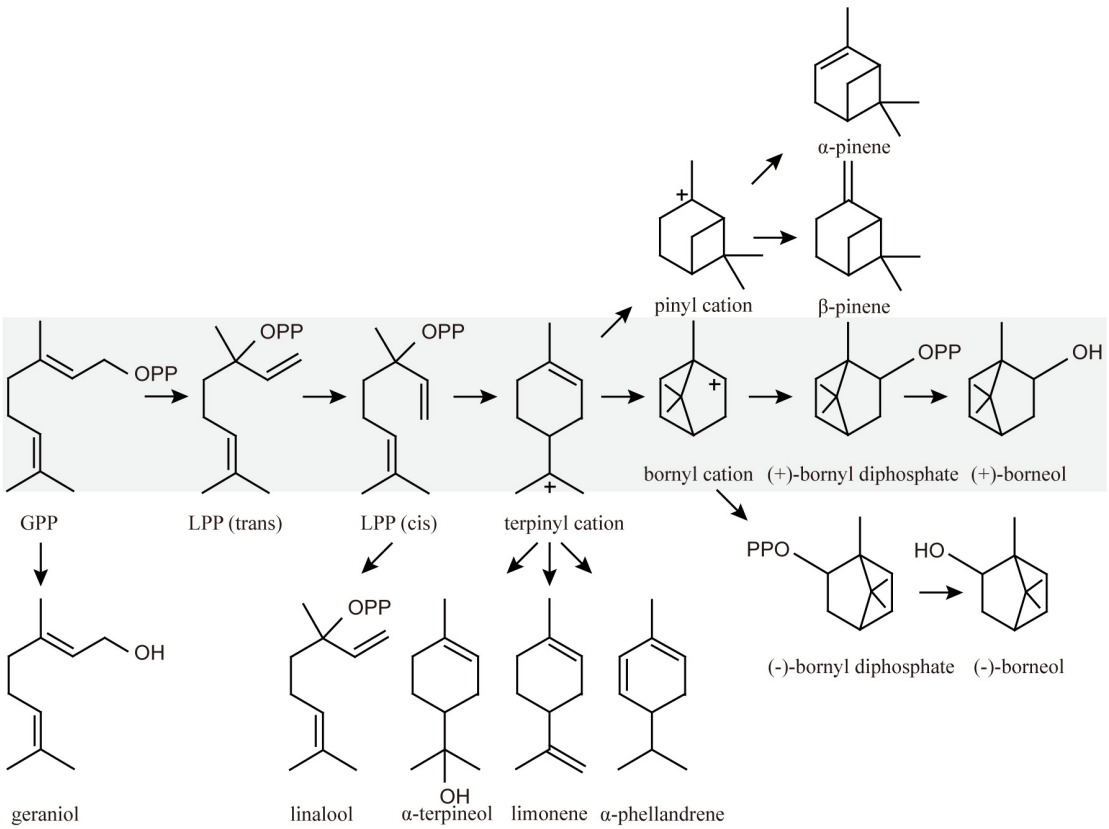


FIGURE 1 | Proposed mechanism for BPPS. The primary pathway leads to the formation of borneol (gray) and other monoterpenoid products.

Gene Cloning, Protein Expression and Purification

The ORF was cloned using specially designed primers (**Supplementary Table 1**). Phusion High-Fidelity PCR Master Mix (New England BioLabs, United States) was used for amplification reaction according to the included protocol. PCR products were purified, and then ligated into the pEASY®-Blunt Simple Cloning Vector (TransGen Biotech, China) and transformed into *E. coli* DH5α cells. Positive colonies were verified by sequencing (Beijing RuiBo Biotechnology Co., Ltd., China) and then subcloned into the pET-32a (+) expression vector (Novagen, United States) according to the protocol of the pEASY®-Uni Seamless Cloning and Assembly Kit (TransGen Biotech, China) (**Supplementary Table 1**).

Recombinant proteins were expressed and purified following the methods described previously (Ma et al., 2020), with some modifications as follows: the 200 mL bacterial solution was centrifuged (5,000 × g, 5 min, 4°C) to collect the cell pellets, and resuspended in 5 ml assay buffer (50 mM HEPES, pH 7.2, 10 mM $MgCl_2$, 5 mM dithiothreitol), and then a sonicator was used to lyse cells. The lysates were centrifuged (12,000 × g, 30 min, 4°C) to produce crude protein. And then the His-tagged purified proteins were eluted using a buffer equivalent to the binding buffer but supplemented with different concentrations of imidazole (50, 100, 250, 350, and 500 mM). Fractions containing the target protein were pooled together and concentrated to a volume of 1 mL using an Amicon Ultra-15 centrifugal filter unit with an Ultracel-30 membrane (Merck Millipore, Germany). Protein concentrations were determined using the Bradford Assay (Cowin Biotech, China). The protein samples were assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

In vitro Enzyme Assays and Kinetic Assays

In vitro enzyme assays followed the method described below: enzyme assays were performed in 300 μL, containing 50 mM HEPES (Ph 7.2), 10 mM $MgCl_2$, 5 mM DTT, 1 mM PMSF, 380 nM of the enzyme and 50 μM GPP, incubated for 1 h at 30°C. Then 1.5 μL calf intestinal alkaline phosphatase (CIAP) (TaKaRa, Japan) was added, followed by incubation for 2 h at 37°C to allow enzymatic dephosphorylation. Time-course experiments were carried out to obtain the initial speed of the enzymatic reaction from 1 to 180 min (**Supplementary Figure 1**). Then, 3 min was used in the kinetic assays. The enzyme assays were performed in a 300 μL reaction volume at 30°C. A concentration that ranged from 0.125 to 150 μM GPP substrate was used. After 3 min incubation, the reaction was terminated at 80°C for 3 min, followed by quenching in ice, and then added 1.5 μL CIAP, followed by incubation for 30 min at 37°C. Assay products

were extracted twice with 300 μL of hexane and samples were concentrated under a gentle nitrogen flow. The samples were then redissolved with 100 μL of hexane before analysis with gas chromatography coupled with mass spectrometry (GC-MS) (described below).

The GraphPad Prism version 5 for Windows (GraphPad Software, La Jolla California United States)[8] was used to obtain kinetic parameters by fitting the obtained data to the Michaelis-Menten equation. All assays were performed in triplicate.

Construction of (+)-Borneol Producing Strains

The initial strain used in this study was CEN.PK2-1D derived from *S. cerevisiae* (**Table 1**). All the endogenous genes (*ERG10*, *ERG13*, *tHMG1*, *ERG12*, *ERG8*, *ERG19*, *IDI1*, *ERG20*) involved in the MVA pathway were amplified from CEN.PK2-1D genomic DNA. The mutant of *ERG20*, $ERG20^{F96W-N127W}$, used in this work was reported to possess higher efficiency for monoterpene production (Jiang et al., 2017). The M2S integration method was applied to integrate gene expression cassettes into the yeast chromosome (Li et al., 2016). Briefly, *ERG10* and *ERG13* were amplified with the addition of a *Bsa*I digestion site and ligated with head-to-head promoters (*pGAL1-pGAL10*) into the terminator vector T1-(TPI1-PGI1), resulting in the plasmid T1-(*ERG10-ERG13*). Two terminators were inserted into the scaffold plasmid, with dedicated homologous arms L1 and L2 lying on both sides. Similarly, plasmids T2-(*tHMG1-tHMG1*), T3-(*tHMG1-ERG12*), T4-(*ERG8-ERG19*), and T5-(*IDI1-*$ERG20^{F96W-N127W}$) were generated with dedicated homologous arms L2 and L3, L3 and L4, L4 and L5, L5 and L6, respectively. Each expression cassette with designed homologous arms was amplified individually. The integration site *YPRCΔ15* was chosen as the target locus, and *URA3* was chosen as the selection marker. The upstream homologous arm *YPRCΔ15*-UP was amplified from CEN.PK2-1D genomic DNA; *URA3* cassette including the promoter was amplified from pESC-URA vector; and L1 arm was amplified from terminator vector T1. These three parts were assembled to form the selection marker module *YPRCΔ15*UP-URA3-L1 through overlap extension PCR. The downstream homologous arm *YPRCΔ15*DOWN was amplified from CEN.PK2-1D genomic DNA and the L6 arm was amplified from terminator vector T5, and they were then combined to generate the downstream homologous arm module L6-*YPRCΔ15*DOWN. All the amplified fragments were used to co-transform CEN.PK2-1D for assembly and integration, and transformants were selected on synthetic drop in medium-Ura (SD-Ura) containing 20 $g{\cdot}L^{-1}$ glucose and 18 $g{\cdot}L^{-1}$ agar. Positive transformants were verified by sequencing, yielding the strain MD.

For (+)-borneol production, the yeast codon-optimized CbTPS1 as well as three truncated variants of CbTPS1 (at positions S10, S32 and C37) were cloned into the *Bam*HI site of the pESC-Leu vector (Agilent Technologies, United States) according to the pEASY-Uni Seamless Cloning and Assembly Kit (TransGen Biotech, Beijing, China), yielding the plasmids pESC-LEU:*CbTPS1*, pESC-LEU:*t10-CbTPS1*, pESC-LEU:*t32-CbTPS1*, and pESC-LEU:*t37-CbTPS1*. Further, yeast-specific Kozak sequence was added in front of the START codon ATG of CbTPS1 and the three truncated variants, generating pESC-LEU:*CbTPS1K*, pESC-LEU:*t10-CbTPS1K*, pESC-LEU:*t32-CbTPS1K*, and pESC-LEU:*t37-CbTPS1K*. Plasmids with the correct sequence were transferred to the host strain MD using Frozen-EZ Yeast Transformation II™ (Zymo Research, United States) to obtain the (+)-borneol producing strains (**Table 1**). All the primers used are listed in **Supplementary Table 1**.

TABLE 1 | Information of strains and vectors used in this study.

Strains or vectors	Description	Source
CEN.PK2-1D	*MATα, URA3-52, TRP1-289, LEU2-3112, HIS3Δ1, MAL2-8C, SUC2*	EUROSCARF
MD	CEN.PK2-1D, *YPRCΔ15 URA3-*P_{GAL1}*-ERG10-*T_{TPI1}*-*P_{GAL10}*-ERG13-*T_{PGI}*-*P_{GAL1}*-tHMG1-*T_{ADH1}*-*P_{GAL10}*-tHMG1-*T_{CYC1}*-*P_{GAL1}*-tHMG1-*T_{FBA1}*-*P_{GAL10}*-ERG12-*T_{PDC1}*-*P_{GAL1}*-ERG8-*T_{RPS2}*-*P_{GAL10}*-ERG19-*T_{TDH1}*-*P_{GAL1}*-IDI1-*T_{CCW12}*-*P_{GAL10}*-ERG20*$^{F96W-N127W}$*-*T_{RPL9A}	This study
MD-1	MD, pESC-LEU::*CbTPS1*	This study
MD-2	MD, pESC-LEU::*CbTPS1K*	This study
MD-3	MD, pESC-LEU::*t10-CbTPS1*	This study
MD-4	MD, pESC-LEU::*t10-CbTPS1K*	This study
MD-5	MD, pESC-LEU::*t32-CbTPS1*	This study
MD-6	MD, pESC-LEU::*t32-CbTPS1K*	This study
MD-7	MD, pESC-LEU::*t37-CbTPS1*	This study
MD-8	MD, pESC-LEU::*t37-CbTPS1K*	This study
T1-(TPI1-PGI)	Terminator vector with terminators TPI1 and PGI	This study
T2-(ADH1-CYC1)	Terminator vector with terminators ADH1 and CYC1	This study
T3-(FBA1-PDC1)	Terminator vector with terminators FBA1 and PDC1	This study
T4-(RPS2-TDH1)	Terminator vector with terminators RPS2 and TDH1	This study
T5-(CCW12-RPL9A)	Terminator vector with terminators CCW12 and RPL9A	This study

Shake Flask Fermentation

For shake flask fermentation, the positive strains were cultured in flasks (50 ml) containing 10 ml of synthetic drop-out medium without leucine and uracil (SD-Leu-Ura) (FunGenome, China) at 30°C and 200 rpm for 48 h. Next, the cells were collected and induced by GAL promoters in 10 ml of YPL (1% yeast extract, 2% peptone, and 2% galactose) medium at 30°C and 200 rpm for 48 h. The fermentation products were extracted with an equal volume of ethyl acetate for 1 h, and centrifuged at 13,000 × g for 10 min to separate the upper organic phase for analyzing by GC-MS (described below). The calibration curves for content determination are shown in **Supplementary Figure 2**. All assays were performed in triplicate.

[8]http://www.graphpad.com

Analysis Using GC-MS

The assay was carried out using a Trace 1310 series GC with a TSQ8000 MS detector (Thermo Fisher Scientific, United States). A TR-5 ms capillary column (30 m × 0.25 mm i.d., 0.25 μm film thickness; Thermo Fisher Scientific, United States) was used. The carrier gas for GC was helium at a flow rate of 1.0 mL·min^{-1}. The oven program was as follows: 50℃ for 2 min, linear ramp up at a rate of 5℃·min^{-1} to 230℃, held at 230℃ for 5 min, followed with a linear ramp up at a rate of 10℃·min^{-1} to 300℃, held at 300℃ for 2 min. The injector temperature and transfer line temperature were 280℃.

A chiral column, Agilent CycloSil-B (30 m × 0.25 mm i.d., 0.25 μm film thickness), was used to identify the chirality of the assay product and the content of borneol and camphor in *C. camphora* leaves. The carrier gas for GC was helium at a flow rate of 1.0 mL·min^{-1}. The oven program was as follows: 50℃ for 2 min, followed by a gradient from 50℃ to 180℃ at 5℃·min^{-1}, then 10℃·min^{-1} to 230℃, held at 230℃ for 2 min. The injector temperature was 200℃, and the transfer line temperature was 230℃.

RESULTS

Transcriptome-Based Discovery of (+)-Bornyl Diphosphate Synthase in 2

Based on the high abundance of (+)-borneol in the leaves of *C. burmanni* (Shi et al., 2013), we used RNA isolated

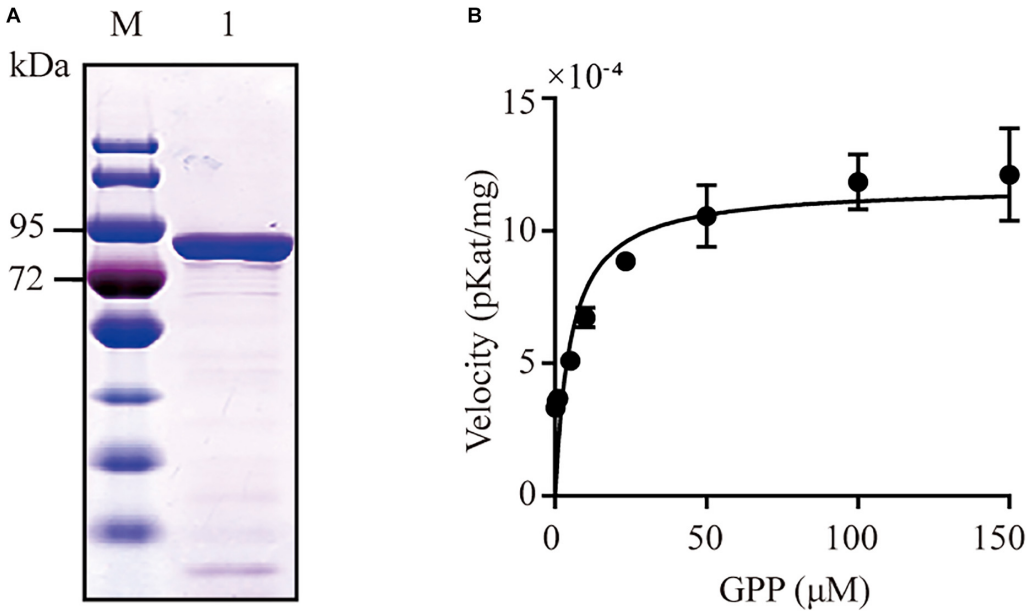


FIGURE 2 | SDS-PAGE and kinetic assays analysis of CbTPS1. **(A)** SDS-PAGE of CbTPS1 *in vitro* assays. **(B)** Velocity of CbTPS1 at increasing GPP concentrations.

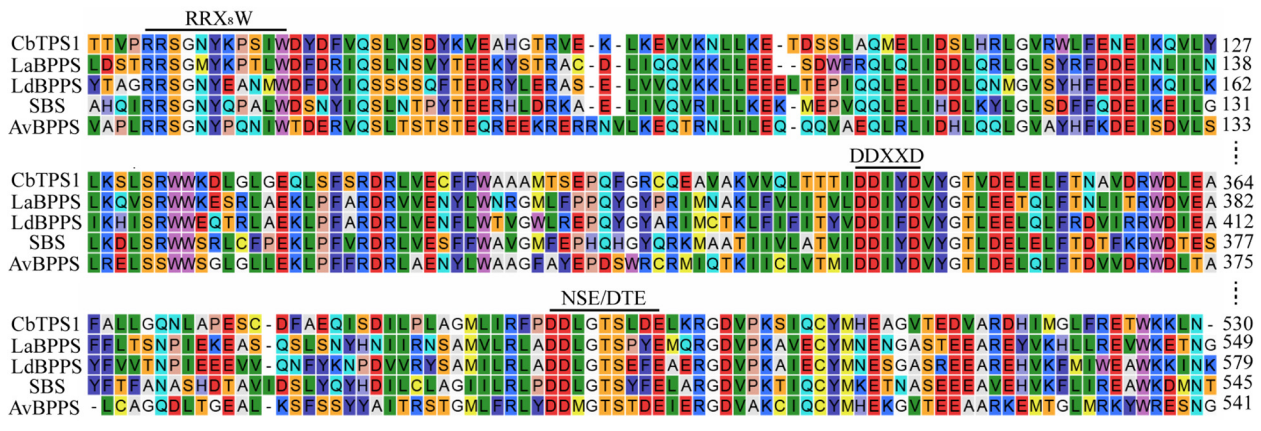


FIGURE 3 | Partial alignment of CbTPS1 protein sequence with those of other known BPPSs. LaBPPS (accession No. AJW68082); LdBPPS (accession No. ATY48638); SBS (accession No. AAC26017); AvBPPS (accession No. AWW87313). The conserved RRX$_8$W, DDXXD and NSE/DTE motifs are underlined.

from young leaves to produce the transcriptome sequences. The reported BPPS genes were further queried against the *de novo* assembly of these sequences, showing that trinity_1267_c0_g1_i1 had the highest identity with all the reported genes. Trinity_1267_c0_g1_i1 was present as full-length sequence. It was further cloned using specific primers and annotated as CbTPS1.

CbTPS1 has an open reading frame of 1,812 bp that encodes a 603-residue enzyme with a calculated molecular mass of 69.1 kDa (**Figure 2A**). It was classified into the TPS-b subfamily, which contains three motifs of typical terpene synthases, namely the RRX_8W motif responsible for monoterpenoid cyclization (Chen et al., 2011); and the DDXXD and NSE/DTE motifs in the C-terminal domain, which are responsible for metal-dependent ionization and substrate binding (Chen et al., 2011). Homologous alignment analysis showed that CbTPS1 shared highest sequence identities with SBS (41.75%) from *Salvia officinalis* (Wise et al., 1998; **Figure 3**), followed by AvBPPS (40.10%) from *Amomum villosum* (Wang et al., 2018), LaBPPS (38.68%) from *Lavandula angustifolia* (Despinasse et al., 2017), and LdBPPS (36.20%) from *Lippia dulcis* (Hurd et al., 2017).

Functional Analysis of CbTPS1

The recombinant protein of CbTPS1 was expressed in *E. coli* Transetta (DE3) cells using the pET-32a (+) expression vector, and then its function was identified with GPP as a substrate. CIAP was then added to remove the diphosphate group from the intermediate product. The purified CbTPS1 produced several monoterpenes (**Figure 4A**). Borneol was predominant (88.70%) with small amounts of α-pinene (2.70%), β-pinene (0.76%), α-phellandrene (1.20%), limonene (2.37%), and other minor monoterpenoids (4.27%). CbTPS1 was further examined for its catalytic properties and the K_m value was 5.11 ± 1.70 μM with a k_{cat} value of 0.01 s^{-1} (**Figure 2B**). In parallel, no product formation was found when the empty vector was transformed into *E. coli* Transetta (DE3) cells, and no product was produced in the absence of CIAP.

A chiral column was used to identify the chirality of borneol. Based on the results of GC-MS analysis (**Figure 4B**), a single product (peak 1) corresponding to the authentic standard (+)-borneol was detected. When the authentic standard (+)-borneol was added to the reaction product, only peak 1 was detected. However, a new product (peak 7) was detected when the authentic standard (−)-borneol was added. This result further proved that (+)-borneol was produced with GPP as a substrate.

Reconstituting the MVA Pathway in Yeast for (+)-Borneol Production

When the codon-optimized CbTPS1 was overexpressed in yeast CEN.PK2-1D, (+)-borneol could not be detected (**Figure 5A**). In addition, geraniol (the dephosphorylated GPP, precursor of (+)-borneol) was not detected in CEN.PK2-1D (**Figure 5B**). Hence, we reconstituted the MVA pathway in CEN.PK2-1D by overexpressing all the MVA pathway genes (*ERG10*, *ERG13*, *tHMG1*, *ERG12*, *ERG8*, *ERG19*, *IDI1*, $ERG20^{F96W-N127W}$) to increase the precursor pool (**Figure 5C**). The obtained chassis strain MD can accumulate 12.52 $mg{\cdot}L^{-1}$ geraniol (**Figure 5B**). Then CbTPS1 was overexpressed in strain MD, and (+)-borneol was generated with a yield of 0.03 $mg{\cdot}L^{-1}$ (strain MD-1) (**Figure 5A**).

Improving the (+)-Borneol Yield by Tailored Truncations

To obtain a higher (+) borneol titer, we engineered the CbTPS1 by further structure optimization. Most terpene synthases in plants have N-terminal plastidic transit peptidases, and will be hydrolyzed after the protein is targeted to the plastid (Bohlmann et al., 1998; Zybailov et al., 2008; Rowland et al., 2015). However, this affects the catalytic activity because yeast cannot digest the transit peptide. Thus, we truncated the chloroplast transit peptide according to the prediction of ChloroP[7]; CbTPS1 was truncated at the C37 position in the N-terminus, and named t37-CbTPS1. (+)-borneol was detected by GC-MS (**Figure 6A**). The truncated t37-CbTPS1 showed a significant increase of (+)-borneol production to 1.53 $mg{\cdot}L^{-1}$ (strain MD-7) (**Figure 6B**).

Hamilton compared 96 *Saccharomyces cerevisiae* sequences, and analyzed the window of 100 bases around the START codon (Hamilton et al., 1987). They found that 50% of highly expressed genes use the UCU serine codon as the second triplet, which indicated that UCU following the START codon ATG could increase gene expression. Therefore, we designed two truncated proteins with ATG followed by the UCU codon. Both amino acids at positions 10 (TCC) and 32 (TCA) of CbTPS1 are serine, which is the same as the amino acid encoded by UCU, so the codon corresponding to the truncated site was mutated to TCT to increase the (+) borneol titer, resulting in t10-CbTPS1 and t32-CbTPS1, respectively. The (+) borneol titer of the two truncated proteins increased significantly. The titer of truncated t10-CbTPS1 was 1.48 $mg{\cdot}L^{-1}$ (strain MD-3), which was 49.33-fold higher than untruncated CbTPS1, and the titer of truncated t32-CbTPS1 was 72-fold higher than CbTPS1, up to 2.16 $mg{\cdot}L^{-1}$ (strain MD-5) (**Figure 6B**).

Improving the (+)-Borneol Yield by Adding Kozak Sequence

The Kozak sequence is roughly the first six important nucleotides upstream of the START codon in *S. cerevisiae*, which are used for gene translation and expression. In yeast, the Kozak sequence is mostly "AAAAAA" (Hamilton et al., 1987; Li et al., 2017; Hernández et al., 2019). On the basis of truncation, yeast-specific Kozak sequence was added in front of START codon ATG of the codon-optimized CbTPS1 and three truncated proteins to increase (+)-borneol yield. The modified proteins were named CbTPS1K, t10-CbTPS1K, t32-CbTPS1K, and t37-CbTPS1K. The results showed that the yield increased at different levels after adding the Kozak sequence. The highest (+)-borneol titer was achieved in strain MD-6 containing t32-CbTPS1K, which is 96.33-fold higher than that in the strain harboring wild-type CbTPS1, producing 2.89 $mg{\cdot}L^{-1}$ (+)-borneol (**Figure 6B**).

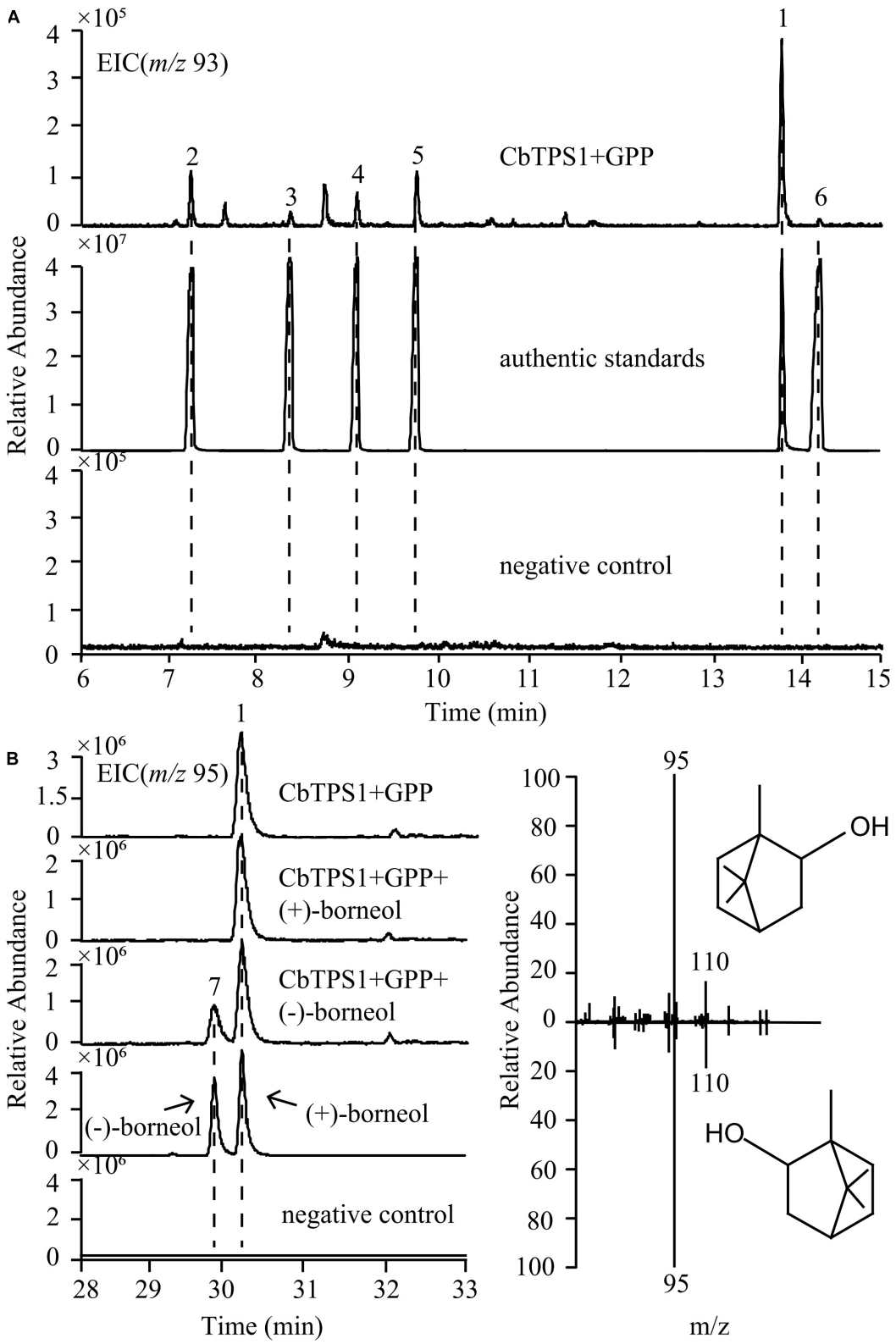


FIGURE 4 | GC-MS analysis of *in vitro* assays with CbTPS1. **(A)** Extracted ion chromatograms of *m/z* 93 *in vitro* assays with purified CbTPS1 and GPP as a substrate. Peak 1, (+)-borneol, Peak 2, α-pinene, Peak 3, β-pinene, Peak 4, α-phellandrene, Peak 5, limonene, Peak 6, α-terpineol. **(B)** Chromatogram of borneol product compared with authentic standards (+)- and (−)-borneol. Peak 1, (+)-borneol, Peak 7, (−)-borneol. Corresponding mass spectrum of (+)-borneol (upper halves) and (−)-borneol (lower halves). EIC, Extracted ion chromatograms.

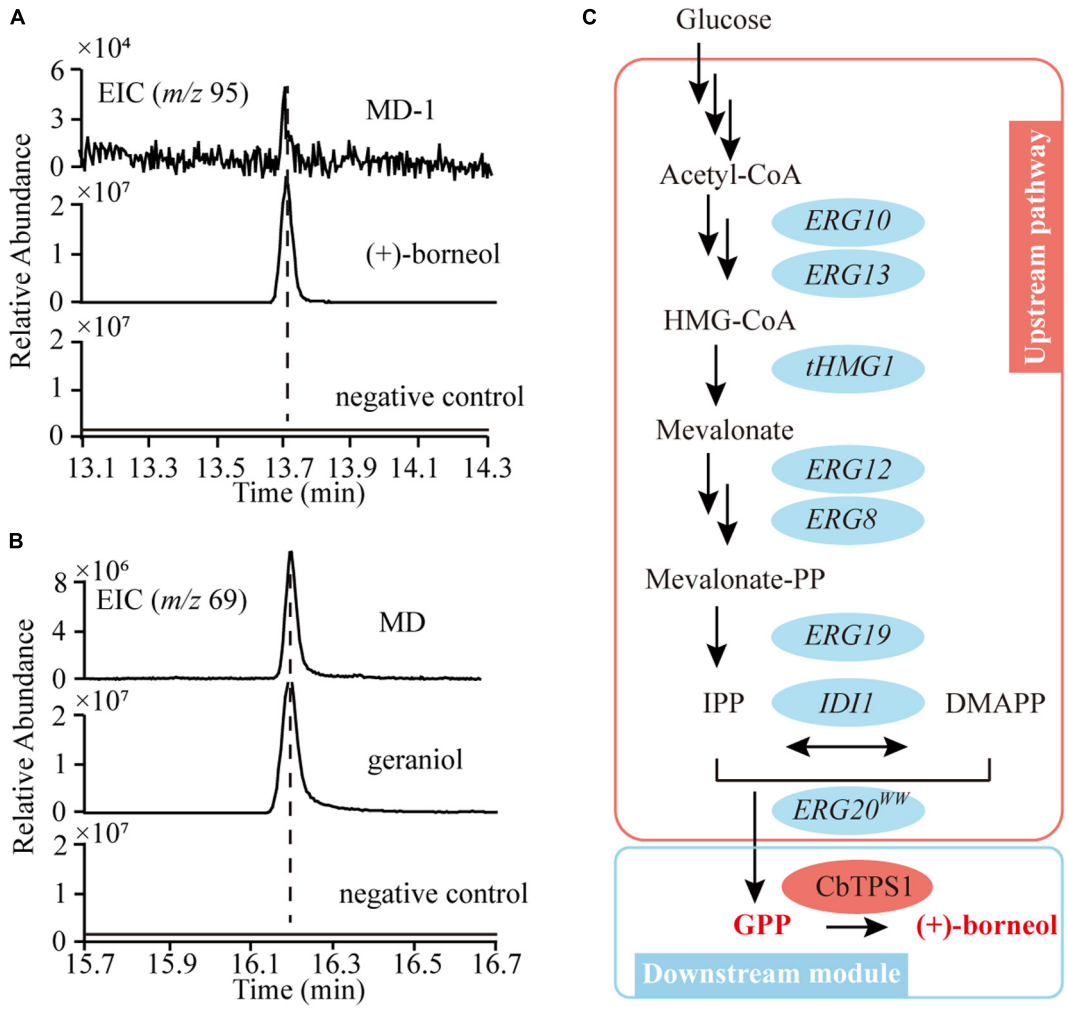


FIGURE 5 | The biosynthetic pathway of GPP and (+)-borneol in *S. cerevisiae*. **(A)** Extracted ion chromatograms of *m/z* 95 of (+)-borneol production in CEN.PK2-1D (negative control) and the MD-1 strain. **(B)** Extracted ion chromatograms of *m/z* 69 of geraniol production in CEN.PK2-1D (negative control) and the MD strain. **(C)** Reconstitution of the MVA pathway in yeast for GPP production (pink ellipses), and the biosynthetic pathway of (+)-borneol (blue ellipses).

DISCUSSION

Due to the insufficient supply of natural products, the role of microbial production of valuable compounds has emerged as an attractive alternative source. Microbial production is a promising choice to substitute for chemical synthesis or phytoextraction (Kirby and Keasling, 2009; Nielsen, 2019). High-efficiency gene elements are vital for metabolic engineering. In this study, we identified a (+)-bornyl diphosphate synthase (CbTPS1) from *C. burmannii* that catalyzed GPP to form (+)-borneol under the hydrolysis of CIAP. This is the first time an enzyme related to (+)-borneol synthesis was mined from *C. burmannii*, and it has the highest specificity for (+)-borneol production (Wise et al., 1998; Despinasse et al., 2017; Hurd et al., 2017; Wang et al., 2018). The K_m value of CbTPS1 (5.11 μM) for GPP is consistent with SBS (3.0 μM) (Wise et al., 1998) and slightly lower than other reported monoterpene synthases (13.10–26.12 μM), which indicated CbTPS1 had a higher affinity for GPP. Its k_{cat}/K_m (1.99×10^{-3} $s^{-1}/\mu M$) is similar to that of other efficient and highly specific monoterpene synthases (3.55×10^{-3}–1.23×10^{-2} $s^{-1}/\mu M$) (Morehouse et al., 2017; Ignea et al., 2019; Dusséaux et al., 2020). Thus, it gives us an opportunity to reconstruct the (+)-borneol biosynthetic pathway in *S. cerevisiae*.

In *S. cerevisiae,* GPP is mainly produced by FPP synthase (ERG20) to serve as the intermediate product of FPP synthesis, thus, it should be consumed rapidly. As a result, when there is no engineering of ERG20, no (+)-borneol or geraniol was detected in CEN.PK2-1D. In order to increase the GPP pool, we further overexpressed all MVA pathway genes and mutated the 96F and 127N of ERG20 to obtain the strain MD, which generated the target product (+)-borneol. However, compared with the accumulation of geraniol (12.52 $mg \cdot L^{-1}$), the yield of (+)-borneol product was relatively low (0.03 $mg \cdot L^{-1}$). Thus, modified proteins were used to improve the expression and activity of CbTPS1. After steady modification, strain MD6 was obtained with the highest yield of (+)-borneol (2.89 $mg \cdot L^{-1}$).

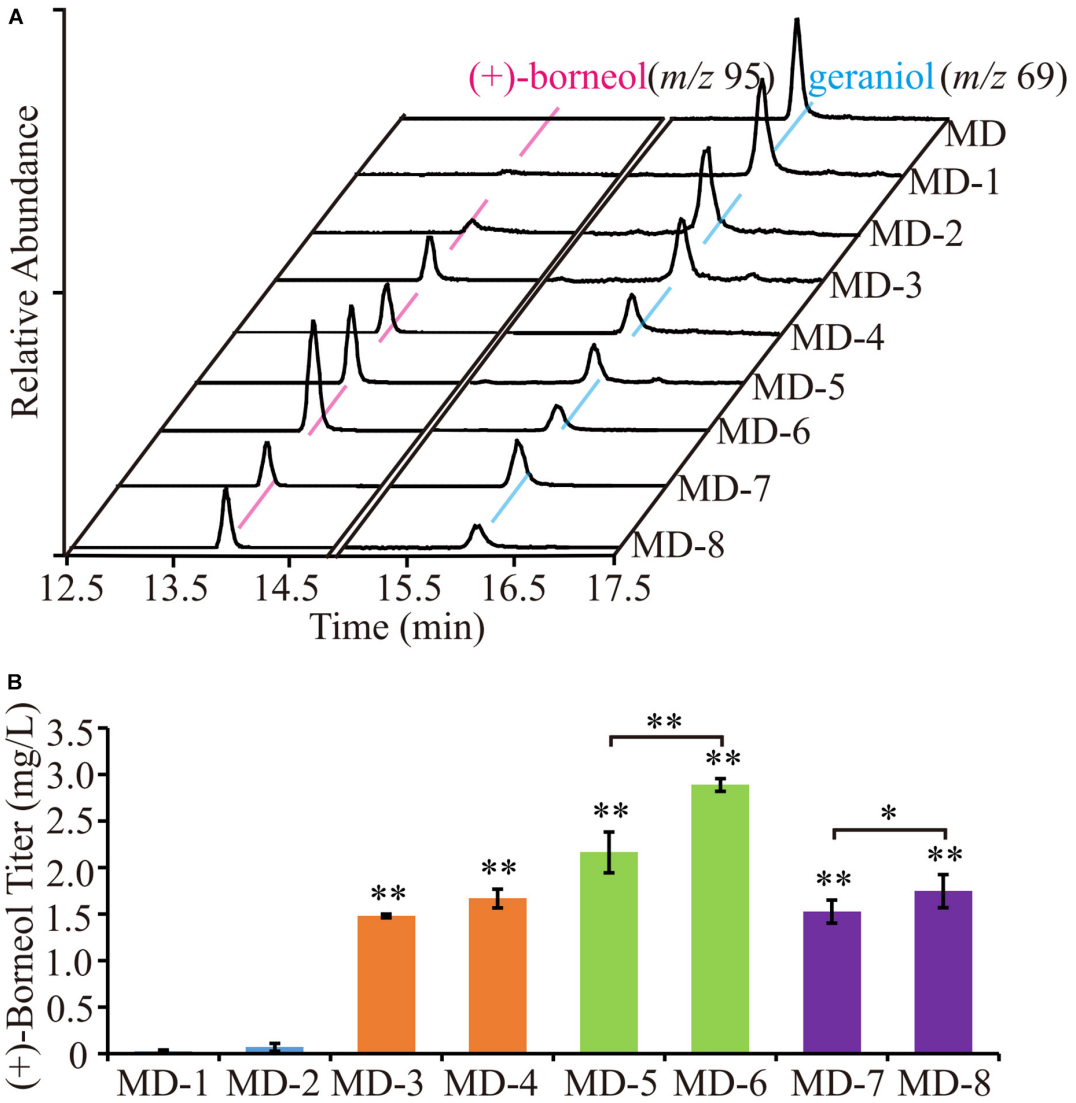


FIGURE 6 | The (+)-borneol production of strains expressing truncations of the CbTPS1. **(A)** GC-MS analysis of the fermentation products of strains expressing the truncated proteins. (**B**) The titer of (+)-borneol product in strains expressing the truncated proteins. (*) represent means which are significantly different at $p < 0.05$; (**) represent means which are significantly different at $p < 0.01$.

Thus, the combination of truncation and using Kozak sequence is an effective strategy for improving (+)-borneol productivity.

Though more than 20 mg·L^{-1} of linalool, α-terpineol, and limonene were produced in yeast (Cao et al., 2016; Zhang et al., 2019, 2020), the yields of most monoterpenes are still lower than the sesquiterpenes and diterpenes (Zebec et al., 2016; Zhao et al., 2016; Jiang et al., 2017), such as artemisinic acid (25 g·L^{-1}) (Paddon et al., 2013) and miltiradiene (3.5 g·L^{-1}) (Hu et al., 2020). The efficiency of forming the final product is influenced by many factors. Reduced efficiency is partially due to the high toxicity of many monoterpenes, such as pinene and limonene, to *S. cerevisiae* because they alter membrane properties or damage the cell wall (Brennan et al., 2013; Demissie et al., 2019). Two-phase extractive fermentation is usually used to alleviate the toxicity of monoterpenes (Brennan et al., 2012). We next will attempt more protein modification of CbTPS1, such as translational fusion (Ignea et al., 2019) and directed evolution of enzymes (Qu et al., 2019). In addition, optimizing the fermentation strategy, by selecting suitable solvent, and optimizing the carbon sources and fermentation parameters will further enhance production (Zhou et al., 2019). Thus, we have good reason to believe that *S. cerevisiae* could be a promising platform for a feasible, scalable, and economic route to the overproduction of (+)-borneol derivatives in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: Genome Sequence Archive

(Genomics, Proteomics & Bioinformatics 2017) in BIG Data Center (Nucleic Acids Res 2018), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers CRA003558 (http://bigd.big.ac.cn/gsa).

AUTHOR CONTRIBUTIONS

LH and GC conceived and designed the experiment. JG and BJ operated GC-MS. QM, HZ, and LC analyzed the data. LM provided the materials. MT purified protein. RM and PS were involved in all experimental steps and wrote this manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key R&D Program of China (2020YFA0908000 and 2018YFA0900600), Major Program of National Natural Science Foundation of China (81891010 and 81891013), the National Natural Science Foundation of China (81822046), and Key project at central government level: The ability to establish sustainable use of valuable Chinese medicine resources (2060302).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.631863/full#supplementary-material

REFERENCES

Amiri, P., Shahpiri, A., Asadollahi, M. A., Momenbeik, F., and Partow, S. (2016). Metabolic engineering of *Saccharomyces cerevisiae* for linalool production. *Biotechnol. Lett.* 38, 503–508. doi: 10.1007/s10529-015-2000-4

Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4126–4133. doi: 10.1073/pnas.95.8.4126

Brennan, T. C. R., Krömer, J. O., and Nielsen, L. K. (2013). Physiological and transcriptional responses of *Saccharomyces cerevisiae* to d-limonene show changes to the cell wall but not to the plasma membrane. *Appl. Environ. Microbiol.* 79, 3590–3600. doi: 10.1128/AEM.00463-13

Brennan, T. C. R., Turner, C. D., Krömer, J. O., and Nielsen, L. K. (2012). Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 109, 2513–2522. doi: 10.1002/bit.24536

Cao, X., Lv, Y. B., Chen, J., Imanaka, T., Wei, L. J., and Hua, Q. (2016). Metabolic engineering of oleaginous yeast *Yarrowia lipolytica* for limonene overproduction. *Biotechnol. Biofuels* 9, 1–11. doi: 10.1186/s13068-016-0626-7

Cao, X., Wei, L. J., Lin, J. Y., and Hua, Q. (2017). Enhancing linalool production by engineering oleaginous yeast *Yarrowia lipolytica*. *Bioresour. Technol.* 245, 1641–1644. doi: 10.1016/j.biortech.2017.06.105

Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E. (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* 66, 212–229. doi: 10.1111/j.1365-313X.2011.04520.x

Chen, M., Hua, Y., Huang, L., Lin, S., and Yang, G. (2010). Analysis on chemical components of leaf oil from reproductive progenies of *Cinnamomum camphora chvar.* Borneo. *Chin. J. Tradit. Chin. Med. Inf.* 17, 37–40. doi: 10.3969/j.issn.1005-5304.2010.08.017

Chen, Z.-X., Xu, Q.-Q., Shan, C.-S., Shi, Y.-H., Wang, Y., Chang, R. C.-C., et al. (2019). Borneol for regulating the permeability of the blood-brain barrier in experimental ischemic stroke: preclinical evidence and possible mechanism. *Oxid. Med. Cell. Longev.* 2019:2936737. doi: 10.1155/2019/2936737

Demissie, Z. A., Tarnowycz, M., Adal, A. M., Sarker, L. S., and Mahmoud, S. S. (2019). A lavender ABC transporter confers resistance to monoterpene toxicity in yeast. *Planta* 249, 139–144. doi: 10.1007/s00425-018-3064-x

Despinasse, Y., Fiorucci, S., Antonczak, S., Moja, S., Bony, A., Nicolè, F., et al. (2017). Bornyl-diphosphate synthase from *Lavandula angustifolia*: a major monoterpene synthase involved in essential oil quality. *Phytochemistry* 137, 24–33. doi: 10.1016/j.phytochem.2017.01.015

Dusséaux, S., Wajn, W. T., Liu, Y., Ignea, C., and Kampranis, S. C. (2020). Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl. Acad. Sci. U.S.A.* 117, 31789–31799. doi: 10.1073/pnas.2013968117

Guo, X. J., Xiao, W. H., Wang, Y., Yao, M. D., Zeng, B. X., and Liu, H. (2018). Metabolic engineering of *Saccharomyces cerevisiae* for 7-dehydrocholesterol overproduction. *Biotechnol. Biofuels* 11:192. doi: 10.1186/s13068-018-1194-9

Hamilton, R., Watanabe, C. K., and De Boer, H. A. (1987). Compilation and comparison of the s ce context around the AUG startcodons in *Saccharomyces cemisiae* mRNAs. *Nucleic Acids Res.* 15, 3581–3593. doi: 10.1093/nar/15.8.3581

Hernández, G., Osnaya, V. G., and Pérez-Martínez, X. (2019). Conservation and variability of the AUG initiation codon context in Eukaryotes. *Trends Biochem. Sci.* 44, 1009–1021. doi: 10.1016/j.tibs.2019.07.001

Hu, T., Zhou, J., Tong, Y., Su, P., Li, X., Liu, Y., et al. (2020). Engineering chimeric diterpene synthases and isoprenoid biosynthetic pathways enables high-level production of miltiradiene in yeast. *Metab. Eng.* 60, 87–96. doi: 10.1016/j.ymben.2020.03.011

Hurd, M. C., Kwon, M., and Ro, D. K. (2017). Functional identification of a *Lippia dulcis* bornyl diphosphate synthase that contains a duplicated, inhibitory arginine-rich motif. *Biochem. Biophys. Res. Commun.* 490, 963–968. doi: 10.1016/j.bbrc.2017.06.147

Ignea, C., Pontini, M., Maffei, M. E., Makris, A. M., and Kampranis, S. C. (2014). Engineering monoterpene production in yeast using a synthetic dominant negative geranyl diphosphate synthase. *ACS Synth. Biol.* 3, 298–306. doi: 10.1021/sb400115e

Ignea, C., Raadam, M. H., Motawia, M. S., Makris, A. M., Vickers, C. E., and Kampranis, S. C. (2019). Orthogonal monoterpenoid biosynthesis in yeast constructed on an isomeric substrate. *Nat. Commun.* 10, 1–15. doi: 10.1038/s41467-019-11290-x

Ji, J., Zhang, R., Li, H., Zhu, J., Pan, Y., and Guo, Q. (2020). Analgesic and anti-inflammatory effects and mechanism of action of borneol on photodynamic therapy of acne. *Environ. Toxicol. Pharmacol.* 75:103329. doi: 10.1016/j.etap.2020.103329

Jiang, G. Z., Yao, M. D., Wang, Y., Zhou, L., Song, T. Q., Liu, H., et al. (2017). Manipulation of GES and ERG20 for geraniol overproduction in *Saccharomyces cerevisiae*. *Metab. Eng.* 41, 57–66. doi: 10.1016/j.ymben.2017.03.005

Jongedijk, E., Cankar, K., Ranzijn, J., van der Krol, S., Bouwmeester, H., and Beekwilder, J. (2015). Capturing of the monoterpene olefin limonene produced in *Saccharomyces cerevisiae*. *Yeast* 32, 159–171. doi: 10.1002/yea.3038

Kirby, J., and Keasling, J. D. (2009). Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol.* 60, 335–355. doi: 10.1146/annurev.arplant.043008.091955

Li, J., Liang, Q., Song, W., and Marchisio, M. A. (2017). Nucleotides upstream of the Kozak sequence strongly influence gene expression in the yeast *S. cerevisiae*. *J. Biol. Eng.* 11, 1–14. doi: 10.1186/s13036-017-0068-1

Li, S., Ding, W., Zhang, X., Jiang, H., and Bi, C. (2016). Development of a modularized two-step (M2S) chromosome integration technique for integration of multiple transcription units in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 9, 1–11. doi: 10.1186/s13068-016-0645-4

Liu, Z., Zhang, Y., and Nielsen, J. (2019). Synthetic biology of yeast. *Biochemistry* 58, 1511–1520. doi: 10.1021/acs.biochem.8b01236

Ma, R., Su, P., Jin, B., Guo, J., Tian, M., Mao, L., et al. (2020). Molecular cloning and functional identification of a high-efficiency (+)-borneol dehydrogenase from *Cinnamomum camphora* (L.) Presl. *Plant Physiol. Biochem.* 158, 363–371. doi: 10.1016/j.plaphy.2020.11.023

Morehouse, B. R., Kumar, R. P., Matos, J. O., Olsen, S. N., and Oprian, D. D. (2017). Functional and structural characterization of a (+)-Limonene Synthase from *Citrus sinensis*. *Biochemistry* 56, 1706–1715. doi: 10.1021/acs.biochem.7b00143

Nielsen, J. (2019). Yeast systems biology: model organism and cell factory. *Biotechnol. J.* 14:e1800421. doi: 10.1002/biot.201800421

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi: 10.1038/nature12051

Qu, G., Li, A., Acevedo-Rocha, C. G., Sun, Z., and Reetz, M. T. (2019). The crucial role of methodology development in directed evolution of selective enzymes. *Angew. Chem.* 59, 13204–13231. doi: 10.1002/anie.201901491

Rowland, E., Kim, J., Bhuiyan, N. H., and Van Wijk, K. J. (2015). The *Arabidopsis* chloroplast stromal n-terminome: complexities of amino-terminal protein maturation and stability. *Plant Physiol.* 169, 1881–1896. doi: 10.1104/pp.15.01214

Shi, S., Wu, Q., Su, J., Li, C., Zhao, X., Xie, J., et al. (2013). Composition analysis of volatile oils from flowers, leaves and branches of *Cinnamomum camphora chvar*, Borneol in china. *J. Essent. Oil Res.* 25, 395–401. doi: 10.1080/10412905.2013.809323

Song, H., Wei, M., Zhang, N., Li, H., Tan, X., Zhang, Y., et al. (2018). Enhanced permeability of blood-brain barrier and targeting function of brain via borneol-modified chemically solid lipid nanoparticle. *Int. J. Nanomed.* 13, 1869–1879. doi: 10.2147/IJN.S161237

Su, P., Guan, H., Zhao, Y., Tong, Y., Xu, M., Zhang, Y., et al. (2018). Identification and functional characterization of diterpene synthases for triptolide biosynthesis from *Tripterygium wilfordii*. *Plant J.* 93, 50–65. doi: 10.1111/tpj.13756

Wang, H., Ma, D., Yang, J., Deng, K., Li, M., Ji, X., et al. (2018). An integrative volatile terpenoid profiling and transcriptomics analysis for gene mining and functional characterization of avbpps and avps involved in the monoterpenoid biosynthesis in *Amomum villosum*. *Front. Plant Sci.* 9:846. doi: 10.3389/fpls.2018.00846

Wise, M. L., Savage, T. J., Katahira, E., and Croteau, R. (1998). Monoterpene synthases from common sage (*Salvia officinalis*). *J. Biol. Chem.* 273, 14891–14899. doi: 10.1074/jbc.273.24.14891

Wojtunik-Kulesza, K. A., Kasprzak, K., Oniszczuk, T., and Oniszczuk, A. (2019). Natural monoterpenes: much more than only a scent. *Chem. Biodivers.* 16:e1900434. doi: 10.1002/cbdv.201900434

Xin, Y., Zhao, H., Xu, J., Xie, Z., Li, G., Gan, Z., et al. (2020). Borneol-modified chitosan: antimicrobial adhesion properties and application in skin flora protection. *Carbohydr. Polym.* 228:115378. doi: 10.1016/j.carbpol.2019.115378

Zebec, Z., Wilkes, J., Jervis, A. J., Scrutton, N. S., Takano, E., and Breitling, R. (2016). Towards synthesis of monoterpenes and derivatives using synthetic biology. *Curr. Opin. Chem. Biol.* 34, 37–43. doi: 10.1016/j.cbpa.2016.06.002

Zhang, C., Li, M., Zhao, G. R., and Lu, W. (2019). Alpha-terpineol production from an engineered *Saccharomyces cerevisiae* cell factory. *Microb. Cell Fact.* 18, 1–9. doi: 10.1186/s12934-019-1211-0

Zhang, Q.-L., Fu, B. M., and Zhang, Z.-J. (2017). Borneol, a novel agent that improves central nervous system drug delivery by enhancing blood-brain barrier permeability. *Drug Deliv.* 24, 1037–1044. doi: 10.1080/10717544.2017.1346002

Zhang, Y., Wang, J., Cao, X., Liu, W., Yu, H., and Ye, L. (2020). High-level production of linalool by engineered *Saccharomyces cerevisiae* harboring dual mevalonate pathways in mitochondria and cytoplasm. *Enzyme Microb. Technol.* 134:109462. doi: 10.1016/j.enzmictec.2019.109462

Zhao, J., Bao, X., Li, C., Shen, Y., and Hou, J. (2016). Improving monoterpene geraniol production through geranyl diphosphate synthesis regulation in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 100, 4561–4571. doi: 10.1007/s00253-016-7375-1

Zheng, Q., Chen, Z.-X., Xu, M.-B., Zhou, X.-L., Huang, Y.-Y., Zheng, G.-Q., et al. (2018). Borneol, a messenger agent, improves central nervous system drug delivery through enhancing blood-brain barrier permeability: a preclinical systematic review and meta-analysis. *Drug Deliv.* 25, 1617–1633. doi: 10.1080/10717544.2018.1486471

Zhou, J., Hu, T., Gao, L., Su, P., Zhang, Y., Zhao, Y., et al. (2019). Friedelane-type triterpene cyclase in celastrol biosynthesis from *Tripterygium wilfordii* and its application for triterpenes biosynthesis in yeast. *New Phytol.* 223, 722–735. doi: 10.1111/nph.15809

Zou, L., Zhang, Y., Li, W., Zhang, J., Wang, D., Fu, J., et al. (2017). Comparison of chemical profiles, anti-inflammatory activity, and UPLC-Q-TOF/MS-based metabolomics in endotoxic fever rats between synthetic Borneol and natural Borneol. *Molecules* 22:1446. doi: 10.3390/molecules22091446

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One* 3:1994. doi: 10.1371/journal.pone.0001994

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 18 February 2021
doi: 10.3389/fbioe.2021.632230

Recent Advances in Silent Gene Cluster Activation in *Streptomyces*

Zhenyu Liu[1†], Yatong Zhao[1†], Chaoqun Huang[1] and Yunzi Luo[1,2]*

[1] Key Laboratory of Systems Bioengineering (Ministry of Education), Frontier Science Center for Synthetic Biology, School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2] Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin University, Tianjin, China

Natural products (NPs) are critical sources of drug molecules for decades. About two-thirds of natural antibiotics are produced by *Streptomyces. Streptomyces* have a large number of secondary metabolite biosynthetic gene clusters (SM-BGCs) that may encode NPs. However, most of these BGCs are silent under standard laboratory conditions. Hence, activation of these silent BGCs is essential to current natural products discovery research. In this review, we described the commonly used strategies for silent BGC activation in *Streptomyces* from two aspects. One focused on the strategies applied in heterologous host, including methods to clone and reconstruct BGCs along with advances in chassis engineering; the other focused on methods applied in native host which includes engineering of promoters, regulatory factors, and ribosomes. With the metabolic network being elucidated more comprehensively and methods optimized more high-thoroughly, the discovery of NPs will be greatly accelerated.

Keywords: *Streptomyces*, natural products, biosynthetic gene cluster, heterologous expression, *in situ* activation, synthetic biology

OPEN ACCESS

Edited by:
Jingwen Zhou,
Jiangnan University, China

Reviewed by:
Guohui Pan,
Chinese Academy of Sciences, China
Zihe Liu,
Beijing University of Chemical Technology, China

***Correspondence:**
Yunzi Luo
yunzi.luo@tju.edu.cn;
luoyunzi827@aliyun.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *22 November 2020*
Accepted: *25 January 2021*
Published: *18 February 2021*

Citation:
Liu Z, Zhao Y, Huang C and Luo Y (2021) Recent Advances in Silent Gene Cluster Activation in Streptomyces. Front. Bioeng. Biotechnol. 9:632230. doi: 10.3389/fbioe.2021.632230

INTRODUCTION

Natural products (NPs) are major sources of drug molecules, including antibiotic, anticancer, antifungal, antiparasitic, and immunosuppressive compounds. *Streptomyces* plays a central role in the discovery of NPs, and the genes responsible for NPs biosynthesis are generally clustered in a continuous region of the genome termed as biosynthetic gene clusters (BGCs). With the rapid development of sequencing technologies, especially the third generation sequencing technology (Loman and Pallen, 2015), more and more genomic information of *Streptomyces* was clarified. Analysis of sequenced *Streptomyces* genome data revealed that a single *Streptomyces*' genome generally encodes 25–50 BGCs, ∼90% of which are silent or cryptic under standard laboratory growth conditions (Walsh and Fischbach, 2010; Rutledge and Challis, 2015; Mao et al., 2018). Therefore, to increase the production of the encoded natural product, methods to unlock or up-regulate these so called "silent" gene clusters have become the interest of research in recent years.
Numbers of methods have been developed to activate silent BGCs in recent years (Rutledge and Challis, 2015; Onaka, 2017; Mao et al., 2018; Lewis, 2020). Powerful bioinformatics approaches for genome mining and identification of NPs BGCs are well summarized in some recent reviews (Lee et al., 2020; Ren et al., 2020; Van Santen et al., 2020; Kenshole et al., 2021). Herein, we provide a concise overview as an introductory guide to the recent advances in silent BGCs activation in *Streptomyces* from two aspects, involving heterologous reconstruction and *in situ* activation

(**Figure 1**). For heterologous reconstruction, we discussed different cloning strategies, biosynthetic pathways reconstruction methods, and chassis strain engineering approaches. For *in situ* activation, we summarized the methods including promoter engineering, transcription factors operating, and ribosome engineering (**Table 1**).

HETEROLOGOUS EXPRESSION OF TARGET BGCs

Heterologous expression is an efficient and established approach to unlock silent or cryptic gene clusters that have been identified by genome mining. Compared to expressions in native hosts, heterologous expression owns several advantages. (1) It can express BGCs whose native host is uncultivable or grows slowly under laboratory growth conditions; (2) heterologous host usually holds mature genetic manipulation tools; (3) the background information of the heterologous host is clearer (Xu and Wright, 2019). Generally, heterologous expression includes three steps-cloning of the target BGCs; engineering of the target BGCs; and transformation to the selected heterologous host. In this section, we will briefly update on the progress of heterologous expression from these three aspects.

Cloning of Large BGCs

Traditional methods for cloning large BGCs generally employ genomic library constructed by cosmid, fosmid,

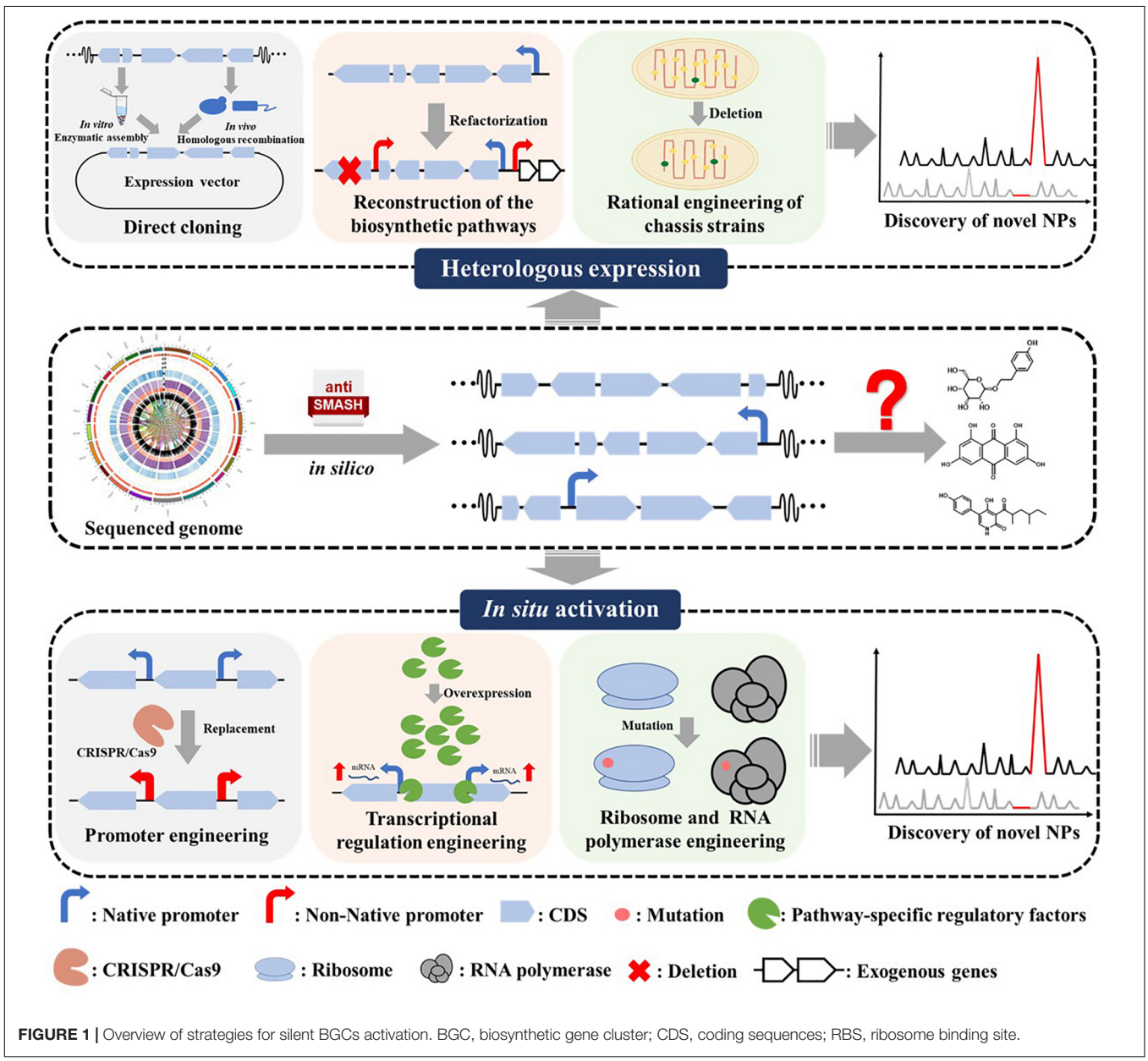


FIGURE 1 | Overview of strategies for silent BGCs activation. BGC, biosynthetic gene cluster; CDS, coding sequences; RBS, ribosome binding site.

TABLE 1 | Examples of BGCs activation in *Streptomyces*.

Hosts	Targets	Approaches	Expression strains	Effect	References
Heterologous host	Cloning BGC	Environmental DNA (eDNA) cosmid libraries	*S. albus* J1074	To clone the complete malacidin BGC from environmental metagenome samples	Hover et al., 2018
		Streptomyces bacterial artificial chromosome system (pSBAC)	*S. lividans and S. coelicolor*	To clone a 60 kb pikromycin BGC	Pyeon et al., 2017
		Exonuclease Combined with RecET recombination (ExoCET)		To clone a 106 kb salinomycin BGC	Wang et al., 2018
		Cas9-Assisted Targeting of chromosome segments (CATCH)		To clone a 36 kb jadomycin BGC	Jiang et al., 2015
		Combining CRISPR/Cas9 system with *in vitro* λ packaging system	*S. avermitilis* MA-4680	To clone 27.4 kb Tü3010 and 40.7 kb sisomicin BGC	Tao et al., 2019
	Reconstruction	Regulatory sequences cassettes	*S. albus* J1074	To activate the actinorhodin BGC.	Ji et al., 2018
		RedEx	*S. albus* J1074	To activate the spinosyn BGC and butenyl-spinosyn A production is 2.36 mg/l, spinosyn Jproduction is 7.34 mg/l	Song et al., 2020
	Chassis strain	Deletion phosphofructokinases (encoded by *pfk*)	*S. albus* J1074	To increase the precursor level of NADPH and heterologous expression of actinorhodin	Kallifidas et al., 2018
		Deletion genomic regions	*S. Chattanoogensis* L320	To delete 0.7 Mb non-essential genomic regions	Bu et al., 2019
Native host	Promoter	Identification strong promoters	*S. griseus*	To activate a PTM BGC.	Luo et al., 2015
		Characterization native or synthetic promoters and Ribosomal binding sites (RBSs)	*S. avermitilis*	To activate and overproduce the lycopene BGC.	Bai et al., 2015
		Identification a strong promoter *groESp*	*S. chattanoogensis* L10	To activate the natamycin BGC and to increase yield by 20%.	Wang K. et al., 2019
		Promoter engineering of the PAS-LuxR (*pimM*)	*Streptomyces*	To activate the polyene BGC.	Barreales et al., 2018
		CRISPR-Cas9 knock-in strategy	*Streptomyces*	To activate multiple BGCs of *Streptomyces* and to trigger the production of a pentangular polyketide.	Zhang et al., 2017
	Regulator	Expression of *bldA*	*S. coelicolor*	To activate the actinorhodin, undecylprodigiosin and methylenomycin BGCs.	Hackl and Bechthold, 2015)
		Overexpression of *slnR*	*S. albus*	To activate the salinomycin BGC.	Zhu et al., 2017)
		Heterologous expression of *papR2*	*S. lividans*	To activate the undecylprodigiosin BGC.	Krause et al., 2020)
		Expression of *gdmRIII*	*S. autolyticus*	To positively control the biosynthesis of geldanamycin.	Jiang et al., 2017
		Overexpression of *toyA*	*S. diastatochromogenes*	To activate the toyocamycin BGC and toyocamycin highest titer is 456.3 mg/l.	Xu et al., 2019
		Expression of *avel*	*S. avermitilis*	To activate the melanin BGC.	Liu et al., 2019
	Ribosome and RNA polymerase	Overexpression of exogenous *rpsL* and *rpoB* genes containing beneficial mutations	Marine *Streptomycete*	To activate the piloquinone and homopiloquinone BGCs	Zhang Q. et al., 2020
		Mutation RNA polymerase: *rpoB* (H437Y)	*S. chattanoogensis* L10 (CGMCC 2644)	To activate the anthrachamycin BGC.	Li Z.Y. et al., 2019
		Mutation RNA polymerase:guanosine-tetraphosphate (ppGpp)	*S.sp.* SoC090715LN-16 S55-50-5	To identify/overproduce a novel isoindolinomycin.	Thong et al., 2018

BAC (Bacterial Artificial Chromosomal), and PAC (P1-derived Artificial Chromosome) vectors (Blodgett et al., 2005; Jones et al., 2013; Xu et al., 2016). In a recent research, 90 *Actinomycetes* NP BGCs have been successfully heterologous expressed, and about 83% of them were constructed via the cosmid/fosmid library method (Nah et al., 2017).

However, these techniques are often time-consuming as well as laborious.

TAR (Transformation-Associated Recombination) cloning is a powerful and reliable system to directly clone large size BGCs (Orr-Weaver et al., 1981). The ends of the linearized TAR cloning vector contain specific homologous sequences of target BGCs as hooks to stimulate homologous recombination (Kouprina and Larionov, 2016, 2019). Bonet et al. (2015) reported the first case of heterologous expression of a natural product BGC from the marine *Streptomyces Salinispora* via the TAR-mediate pCAP01 vector. Later, Kang et al. (2016) built a mCRISTAR platform that combines CRISPR/Cas9 with TAR to simultaneously replace multiple promoters in the tetarimycin BGC. The system was further improved as mpCRISTAR (Multiple Plasmids-based CRISPR/Cas9 and TAR) by employing multiple plasmids, each harboring one or two unique guide RNAs. Based on mpCRISTAR, six or eight promoters can be simultaneously replaced with an efficiency of 68 and 32%, respectively (Kim et al., 2020).

Meanwhile, approaches based on site-specific recombinase systems have also been developed to directly clone BGCs. The integrase-mediated recombination (IR) system employs phage ΦBT1 *attP-attB-int* system to induce site-specific recombination (Du et al., 2015). Liu et al. (2009) described a versatile *E. coli-Streptomyces* shuttle vector system, pSBAC, employing the ΦBT1 IR system. Pyeon et al. further optimized the above system with additional restriction recognition sites on pSBAC to simplify the cloning procedure. They successfully cloned the 80 kb tautomycetin BGC and 60 kb pikromycin BGC for heterologous expression with the modified system (Pyeon et al., 2017). Zhang et al. (1998, 2000) developed a powerful Red/ET recombineering tool to assemble large DNA fragments using homologous recombination in *E. coli*. Later, Wang et al. (2018) upgraded this system by employing T4 polymerase to facilitate annealing between the linear target DNA and vector *in vitro*, and they termed the system as ExoCET. They used ExoCET to successfully cloned the intact 106 kb salinomycin BGC from *S. albus*.

Apart from *in vivo* cloning technologies, there are various *in vitro* cloning strategies. Gibson assembly has been well applied in multi-segment assembly *in vitro* (Gibson et al., 2009). For example, the 41 kb conglobatin BGC from *S. conglobatus* was cloned through Gibson assembly (Zhou et al., 2015). However, Gibson assembly is inefficient for large DNA fragments with high G + C content (Casini et al., 2014; Li et al., 2015). Therefore, Jiang et al. (2015) combined Gibson assembly with CRISPR-Cas9, termed as CATCH (Cas9-Assisted Targeting of CHromosome segments). They successfully cloned the 36 kb *jad* gene cluster from *S.venezuelae* and the 32 kb *ctc* gene cluster from *S. aureofaciens* into the p15A vector via CATCH. Similarly, Tao et al. (2019) illustrated an *in vitro* one-step targeted cloning approach combining CRISPR/Cas9 system with *in vitro* λ packaging system, and the pathways of Tü3010 (27.4 kb) and sisomicin (40.7 kb) were successfully cloned, respectively.

In summary, each strategy for cloning large BGCs has pros and cons. The methods of genomic library construction are random, but they are beneficial for metagenome-driven natural product discovery (Katz et al., 2016; Hover et al., 2018). The pSBAC is suitable for cloning large DNA fragments with specific restriction digestion sites, which are not generally available at both ends of target BGCs. As for tools based on homologous recombination, like TAR and Red/ET, both are commonly used in cloning large DNA fragments but may introduce some undesired recombination. Although CRISPR tools solve the limitation of insufficient restriction sites, it still faces the bottleneck of isolating targeted BGCs from the genomic DNA. With the increasing number of sequenced genomes, developing high-throughput cloning tools becomes imminent, such as combining current tools with automated platforms (Burger et al., 2020).

Reconstruction of the Biosynthetic Pathways

It has been reported that the complexity of the regulatory network in host cells was a major challenge for metabolic engineering (Shao et al., 2013). Therefore, reconstruction and heterologous expression of the biosynthetic pathways can release them from the complex metabolic network. Nevertheless, BGCs controlled by promoters of different strengths increase the complexity of the reconstruction (Horbal et al., 2018). At present, the reconstruction process mainly includes: (1) gene substitution, (2) enzyme evolution, (3) promoter replacement, (4) transcriptional repressor knockout (Li L. et al., 2019). For example, Alberti et al. (2019) successfully activated the *scl* BGC by inactivating the transcriptional repressors via CRISPR/Cas9. AGOS (Artificial Gene Operon assembly System) is a plug and play method designed for the construction of artificial gene operons through Red/ET mediated recombination. Four gene operons of novobiocin BGCs were heterologously integrated into the genome of *S. coelicolor* M1146 via AGOS, leading to the production of novobiocin and novobiocin precursors (Basitta et al., 2017). Marín et al. cloned the synthetic genes encoding tyrosine ammonia lyase, 4-coumaroyl CoA ligase, chalcone synthase, chalcone isomerase and flavone synthase into a high copy number shuttle vector including a strong promoter *ermE*p*. The final plasmid pAPI was transformed into the heterologous host *S. albus* and successfully produced apigenin at 0.08 mg/L (Marín et al., 2017). In another case, Song et al. (2020) refactored the spinosyn BGCs via RedEx to test whether the ethyl group at C-21 of spinosyn A can be replaced by butene group.

Among the above mentioned methods, promoter replacement is the most effective and well established method to activate silent BGCs, especially in *Streptomyces* (Luo et al., 2013). For example, Luo et al. identified strong promoters from *S. albus* J1074, whose strength is 200–1,300% the strength of the well-known strong promoter *ermE**p. They used a plug-and-play scaffold to successfully activate the silent PTM BGC of *S. griseus* in three widely used *Streptomyces* chassis strains (Luo et al., 2015). Ji et al. (2018) used synthetic regulatory sequences cassettes to successfully activate actinorhodin BGCs. In another case, the silent streptophenazine BGC in marine *Streptomyces S.sp* is non-transcriptional active in heterologous environment. After introducing four constitutive promoters (*ermE*p/actIp/sp44/p21*) at different positions in

the BGC, the production of streptophenazine was detected (Bauman et al., 2019).

All in all, thorough reconstruction of BGCs often leads to the activation of silent BGCs and the discovery of new NPs (Luo et al., 2016; Zhao et al., 2019). However, compared with *E. coli* and other model strains, the genetic manipulation tools in *Streptomyces* are still limited. Therefore, new methods are urgently needed and some new developments are well summarized in other reviews (Tan and Liu, 2017; Tao et al., 2018; Zhao Y. et al., 2020).

Rational Engineering of Chassis Strains

Streptomyces are rich in inherently valuable secondary metabolites. Therefore, series of *Streptomyces* species have been developed as chassis to express heterologous BGCs, such as *S. coelicolor, S. lividans,* and *S. albus* (Ziermann and Betlach, 1999; Zhou et al., 2012; Myronovskyi et al., 2018). A suitable surrogate expression host should contain several essential features: (1) a variety of natural product precursors which are conducive to construct abundant complex molecules; (2) a simplified secondary metabolite background; (3) an efficient transportation system to transfer various bioactive compounds; (4) a known regulatory network; (5) a mature fermentation and upscaling process; (6) powerful genetic manipulation tools (Baltz, 2010, 2016; Myronovskyi and Luzhetskyy, 2019; Xu and Wright, 2019). Researchers did a comprehensive and detailed introduction of *Streptomyces* species used as heterologous hosts from 2010 to 2018 in another excellent review (Myronovskyi and Luzhetskyy, 2019).

The commonly used chassis engineering strategy is to reduce the background of secondary metabolism (Lee et al., 2019). Non-essential genomic regions and secondary metabolic genes mainly appearing in the end region of the chromosome. They are not stable and prone to chromosomal rearrangements, hence knocking out of them may generate clean background chassis strains. Ahmed et al. (2020) developed a set of *S. lividans* chassis strains. The *S. lividans* ΔYA11 was obtained by deleting 11 gene clusters (228.5 kb) and inserting two *attB* sites. Bu et al. (2019) rational constructed two genome-reduced *Streptomyces* chassis strains, the *S. chattanoogensis* L320 and L321, through multiple computational approaches and site-specific recombination systems, with non-essential genomic regions deletion of 1.3 and 0.7 Mb, respectively. Sometimes, the low yield of heterologous produced NPs may be due to insufficient precursors in the expression hosts. Therefore, increasing the supply of the precursors is a promising strategy. Borodina et al. (2008) rationally engineered the *S. coelicolor* A3(2) strain by deleting the phosphofructokinases (encode by *pfkA2*) gene, thus the precursor level of NADPH was increased and the production of actinorhodin and undecylprodigiosin were upregulated correspondingly. Dang et al. (2017) knocked out *pfk* in *S. hygroscopicus* ATCC 29253. The titer of rapamycin increased by 30.8% in the engineered strain. Kallifidas et al. (2018) successfully heterologous expressed actinorhodin in *S. albus*, and then further increased its yield by knocking out pfk_{SA}.

Currently, a set of powerful bioinformatics approaches are developed to design chassis strains rationally (Ren et al., 2020). Meanwhile, the powerful genetic editing tool CRISPR has been applied in *Streptomyces* for genome engineering (Tong et al., 2019a,b; Zhang J. et al., 2020; Zhao Y.W. et al., 2020). These techniques are expected to accelerate the development of *Streptomyces* chassis strains.

Heterologous expression has numerous advantages, but some limitations still exist. (1) The size of SM-BGCs is highly variable (1–100 kb), and most are more than 10 kb. Currently, there is no certain method that is universal, large-size endurable, efficient, and high-throughput; (2) Because of the complicated metabolic networks of *Streptomyces*, clarify the interaction between the host strain and the heterologous BGCs is hard; (3) At present, *Streptomyces* chassis compatible with all NPs' production has not been reported; (4) Current genetic manipulation tools of *Streptomyces* are not applicable in all species, thus more powerful and universal genetic tools are needed.

IN SITU ACTIVATION OF TARGET BGCs

The expression of NPs BGCs in *Streptomyces* is governed by a complex metabolic regulatory network. The production of antibiotics can be greatly enhanced by rewiring the regulatory network (Xia et al., 2020). Therefore, a better understanding and manipulation of the regulatory network in these silent BGCs could help to activate BGCs. In this section, we described different strategies to manipulate the regulatory modules in the native hosts for silent BGCs activation.

Promoter Engineering

With regard to cluster activation, promoter elements are of indisputable importance as they are responsible for efficient transcription, which is the first stage of gene expression (Myronovskyi and Luzhetskyy, 2016). Promoter engineering employs a set of regulatory sequences with known functions, to release the following gene expression from the native complex regulations. Constitutive promoters commonly used to activate gene expression include: the promoter of the erythromycin resistance gene *ermE* of *S. erythraea*, *ermEp*1 and its derivatives (Bibb et al., 1985); the phage I19 originated promoter *SF14p* (Labes et al., 1997); and the engineered *kasOp** promoter (Takano et al., 2005; Wang et al., 2013). Inducible promoters commonly used to activate gene expression include: the thiostrepton-inducible promoter *PtipA* (Holmes et al., 1993), the synthetic resorcinol-inducible and cumate-inducible promoters (Horbal et al., 2014), and the synthetic tetracycline-inducible promoter *tcp*830 (Rodríguez-García et al., 2005).

Since the strategy of knocking in promoters with multiple operon structure by homologous double-crossover recombination is often time-consuming and laborious, (Zhang et al., 2017) reported an effective CRISPR-Cas9 knock-in strategy in *Streptomyces*, and this one-step strategy was applied to activate multiple silent BGCs in five *Streptomyces* species. Similarly, Tong et al. (2015) also adopted the CRISPR-Cas9 system (deemed CRISPRi) to control the expression of target genes in *Actinomycetes*. The combination of the CRISPR system

and promoter engineering approaches makes the experimental operation and procedure relatively simple and efficient.

At present, progress in activating silent BGCs in *Streptomyces* through comprehensive multi-promoter insertion is limited. Constructing promoters with a wide range of transcription initiation activities, transcription strength and robustness would promote effective activation of silent BGCs, and gene expression balance needs to be taken into considerations as well. In short, promoter-based gene expression activation methods still need improvement.

Transcriptional Regulation Engineering

The biosynthesis of NPs in *Streptomyces* is regulated by precise regulatory systems, in which transcription factors (TFs) regulate the initiation level of transcription by binding to DNAs. In the era of synthetic biology, coordination of TFs regulations sometimes can activate silent BGCs, such as overexpression of positive regulatory genes or inactivation of negative regulatory genes in *Streptomyces*. For example, *bldA* of *S. coelicolor* can activate the expression of the antibiotics actinorhodin, undecylprodigiosin, and methylenomycin BGCs (Cuthbertson and Nodwell, 2013; Bhukya and Anand, 2017). Guo et al. (2018) used gene deletion, complementation, and overexpression to determine the MarR family transcriptional regulator (MFR) SAV4189 as an activator of avermectin biosynthesis in *S. avermitilis*. In addition to pathway-specific regulatory factors, global regulatory factors can also activate silent BGCs. For example, through genome sequencing analysis, gene knockout, and transcriptional analysis, the global regulator AdpA was found to be able to activate nikkomycin biosynthesis, and repress the biosynthesis of oviedomycin at the same time (Xu et al., 2017). Recently, Wang B. et al. (2019) reported a transcription factor decoy strategy for targeted activation of large silent polyketide synthase and non-ribosomal peptide synthetase, and discovered a novel oxazole family compound. Li et al. (2020) developed a base editing system that combines CRISPR-Cas9 with site-specific recombination to achieve successful genome editing in *Streptomyces* by programmed mutation of target genes, thereby achieving product biosynthesis (such as hygromycin B). Owing to their simplicity and ease of use, these strategies can be scaled up readily for the discovery of natural products in *Streptomyces*.

Ribosome Engineering

Ribosome engineering is an approach to discover microbes with certain spontaneous mutations in their ribosome or RNA polymerase, through screening antibiotic-resistant mutants on Petri dishes (Zhu et al., 2019). It is suitable for gene activation and strain improvement, resulting in the identification of novel secondary metabolites, as well as the enhancement of enzyme production and tolerance to toxic chemicals (Ochi, 2017).

The *rpoB* gene (encoding the RNA polymerase β-subunit) can activate silent BGCs in various *Streptomyces* by rifampicin resistance mutations (up to 70 times at the transcription level). Analysis of the metabolite profile showed that *rpoB* mutants produced many metabolites undetectable in wild-type strains (Tanaka et al., 2013). Li Z.Y. et al. (2019) used site-directed mutagenesis to generate ten mutants with point mutations in the highly conserved region of *rpsL* (encoding the ribosomal protein S12) or *rpoB*. Among them, L10/*RpoB* (H437Y) activated anthrachamycin biosynthesis in *S. chattanoogensis* L10 (CGMCC 2644). Zhang et al. designed a TTO (Transcription–Translation in One) method using a plug-and-play plasmid system to directly overexpress exogenous *rpsL* (encoding ribosomal protein S12) and *rpoB* (encoding RNA polymerase β subunit) genes containing beneficial mutations. This method overcomes the false positive problem in the traditional ribosome engineering method and was successfully applied to activate the silent BGCs in three *Streptomyces* strains, thus discovering two aromatic polyketide antibiotics (Zhang Q. et al., 2020). Moreover, the ppGpp can interact with RNA polymerase and affect the production of antibiotics (Artsimovitch et al., 2004). It is suggested that RNA polymerases carrying specific *rif* mutations in the β-subunit can functionally mimic modification induced by binding of ppGpp (Xu et al., 2002). So, some studies showed that *rif* mutations could alter the gene expression patterns of ppGpp. Thong et al. (2016) screened mutants resistant to rifampicin and found an unknown metabolite.

At present, in addition to the conventional modification of ribosomes through mutagenesis, other ribosomal regulatory elements have also been engineered. Siu and Chen (2019) proposed a new class of riboregulators called toehold-gated gRNA (thgRNA) by integrating toehold riboswitches into sgRNA scaffolds and demonstrated their programmability for multiplexed regulation in *E. coli* with minimal cross-talks. In the future, this approach could also be tested in *Streptomyces* for gene expression regulation.

Promoter engineering can activate a single gene expression in BGCs, and it can also activate the full-length BGCs to produce the corresponding NPs. This method can be further developed for high-throughput activation of silent BGCs. Knockout of negative regulatory genes is one method to explore new NPs. However, in *Streptomyces*, the traditional gene knockout strategy is often completed by plasmid-mediated homologous recombination, which is usually time-consuming and laborious. Due to the differences in the source, structure and functions of BGCs, more attempts and innovations are needed to unlock the transcriptional regulation of BGCs.

DISCUSSION

At present, in addition to the methods mentioned in this review, the silent BGCs can also be activated by changing the culture conditions. Bode et al. (2002) defined it as one strain many compounds (OSMAC), that is, by adjusting the culture parameters of *Streptomyces*, such as medium composition, culture temperature, pH, aeration, and container type, to induce the expression of silent BGCs. Later, on the basis of OSMAC, other strategies were derived, such as the addition of low-concentration antibiotics, signal molecules and histone deacetylase inhibitors and other inducers (Seyedsayamdost, 2014), as well as co-cultivation strategies. In 2019, a review discussed the use of microbial culture techniques to expand

the range of NPs available in the laboratory in recent years, mainly including methods such as adding physical scaffolds, adding small molecule elicitors, and co-cultivating with another microorganism (Tomm et al., 2019). Although these methods are relatively economical and simple, they are particularly suitable for *Streptomyces* species with incomplete genome information or genetic isolation defects.

Due to the high investment and low return rate of silent BGCs activation, the discovery of new NPs has entered a bottleneck. Through combining bioinformatics analysis with multi-omics data to explore the genomic data, insights to regulate and activate BGCs could be elicited. These methods can not only act alone to produce NPs, but can also be combined with each other. There is still an urgent requirement to develop better methods to activate silent BGCs. For example, structured data can be used to further elucidate the detailed mechanism, automation can help improving high-throughput capabilities, and AI can be employed to assist experiment design. Perhaps combining *in situ* activation with simulation analysis, heterologous expression and other strategies, more precise transcription activation could be achieved for silent BGCs exploration.

AUTHOR CONTRIBUTIONS

ZL, YZ, and CH: writing—original draft. YL: writing—review and editing and project administration. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Key-Area Research and Development Program of Guangdong Province (2020B0303070002), the National Natural Science Foundation of China (Grant No. 32071426), the National Key R&D Program of China (2018YFA0903300), and the Natural Science Foundation of Tianjin Province (19JCYBJC24200).

REFERENCES

Ahmed, Y., Rebets, Y., Estevez, M. R., Zapp, J., Myronovskyi, M., and Luzhetskyy, A. (2020). Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb. Cell Fact.* 19:5. doi: 10.1186/s12934-020-1277-8

Alberti, F., Leng, D. J., Wilkening, I., Song, L., Tosin, M., and Corre, C. (2019). Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *Chem. Sci.* 10, 453–463. doi: 10.1039/c8sc03814g

Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M. N., Hosaka, T., Ochi, K., et al. (2004). Structural basis for transcription regulation by alarmone ppGpp. *Cell* 117, 299–310. doi: 10.1016/s0092-8674(04)00401-5

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., Miao, J., et al. (2015). Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc. Natl. Acad. Sci. U.S.A.* 112, 12181–12186. doi: 10.1073/pnas.1511027112

Baltz, R. H. (2010). *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J. Ind. Microbiol. Biotechnol.* 37, 759–772. doi: 10.1007/s10295-010-0730-9

Baltz, R. H. (2016). Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other *Actinomycetes*. *J. Ind. Microbiol. Biotechnol.* 43, 343–370. doi: 10.1007/s10295-015-1682-x

Barreales, E. G., Vicente, C. M., De Pedro, A., Santos-Aberturas, J., and Aparicio, J. F. (2018). Promoter engineering reveals the importance of heptameric direct repeats for DNA binding by *Streptomyces* antibiotic regulatory protein-large ATP-binding regulator of the LuxR family (SARP-LAL) regulators in *Streptomyces natalensis*. *Appl. Environ. Microbiol.* 84:e00246-18. doi: 10.1128/AEM.00246-18

Basitta, P., Westrich, L., Rösch, M., Kulik, A., Gust, B., and Apel, A. K. (2017). AGOS: a Plug-and-play method for the assembly of artificial gene operons into functional biosynthetic gene clusters. *ACS Synth. Biol.* 6, 817–825. doi: 10.1021/acssynbio.6b00319

Bauman, K. D., Li, J., Murata, K., Mantovani, S. M., Dahesh, S., Nizet, V., et al. (2019). Refactoring the cryptic streptophenazine biosynthetic gene cluster unites Phenazine, Polyketide, and Nonribosomal peptide biochemistry. *Cell Chem. Biol.* 26, 724–736.e727. doi: 10.1016/j.chembiol.2019.02.004

Bhukya, H., and Anand, R. (2017). TetR regulators: a structural and functional perspective. *J. Indian Inst. Sci.* 97, 245–259. doi: 10.1007/s41745-017-0025-5

Bibb, M. J., Janssen, G. R., and Ward, J. M. (1985). Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene* 38, 215–226. doi: 10.1016/0378-1119(85)90220-3

Blodgett, J. A. V., Zhang, J. K., and Metcalf, W. W. (2005). Molecular cloning, sequence analysis, and heterologous expression of the Phosphinothricin Tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736. *Antimicrob. Agents Chemother.* 49, 230–240. doi: 10.1128/aac.49.1.230-240.2005

Bode, H. B., Bethe, B., Höfs, R., and Zeeck, A. (2002). Big effects from small changes: possible ways to explore nature's chemical diversity. *Chembiochem* 3, 619–627.

Bonet, B., Teufel, R., Crusemann, M., Ziemert, N., and Moore, B. S. (2015). Direct capture and heterologous expression of *Salinispora* natural product genes for the biosynthesis of enterocin. *J. Nat. Prod.* 78, 539–542. doi: 10.1021/np500664q

Borodina, I., Siebring, J., Zhang, J., Smith, C. P., Van Keulen, G., Dijkhuizen, L., et al. (2008). Antibiotic overproduction inStreptomyces coelicolorA3(2) mediated by phosphofructokinase deletion. *J. Biol. Chem.* 283, 25186–25199. doi: 10.1074/jbc.m803105200

Bu, Q. T., Yu, P., Wang, J., Li, Z. Y., Chen, X. A., Mao, X. M., et al. (2019). Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb. Cell Fact* 18:16. doi: 10.1186/s12934-019-1055-7

Burger, B., Maffettone, P. M., Gusev, V. V., Aitchison, C. M., Bai, Y., Wang, X., et al. (2020). A mobile robotic chemist. *Nature* 583, 237–241. doi: 10.1038/s41586-020-2442-2

Casini, A., Macdonald, J. T., Jonghe, J. D., Christodoulou, G., Freemont, P. S., Baldwin, G. S., et al. (2014). One-pot DNA construction for synthetic biology: the Modular overlap-directed assembly with linkers (MODAL) strategy. *Nucleic Acids Res.* 42:e7. doi: 10.1093/nar/gkt915

Cuthbertson, L., and Nodwell, J. R. (2013). The TetR family of regulators. *Microbiol. Mol. Biol. Rev.* 77, 440–475. doi: 10.1128/MMBR.00018-13

Dang, L., Liu, J., Wang, C., Liu, H., and Wen, J. (2017). Enhancement of rapamycin production by metabolic engineering in *Streptomyces hygroscopicus* based on genome-scale metabolic model. *J. Ind. Microbiol. Biotechnol.* 44, 259–270. doi: 10.1007/s10295-016-1880-1

Du, D., Wang, L., Tian, Y., Liu, H., Tan, H., and Niu, G. (2015). Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in *Streptomyces*. *Sci. Rep.* 5:8740. doi: 10.1038/srep08740

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. III, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Guo, J., Zhang, X., Lu, X., Liu, W., Chen, Z., Li, J., et al. (2018). SAV4189, a MarR-family regulator in *Streptomyces avermitilis*, activates Avermectin biosynthesis. *Front. Microbiol.* 9:1358. doi: 10.3389/fmicb.2018.01358

Hackl, S., and Bechthold, A. (2015). The Gene *bldA*, a regulator of morphological differentiation and antibiotic production in *Streptomyces*. *Arch. Pharm.* 348, 455–462. doi: 10.1002/ardp.201500073

Holmes, D. J., Caso, J. L., and Thompson, C. J. (1993). Autogenous transcriptional activation of a thiostrepton-induced gene in *Streptomyces lividans*. *EMBO J.* 12, 3183–3191.

Horbal, L., Fedorenko, V., and Luzhetskyy, A. (2014). Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other *Actinobacteria*. *Appl. Microbiol. Biotechnol.* 98, 8641–8655. doi: 10.1007/s00253-014-5918-x

Horbal, L., Marques, F., Nadmid, S., Mendes, M. V., and Luzhetskyy, A. (2018). Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab. Eng.* 49, 299–315. doi: 10.1016/j.ymben.2018.09.010

Hover, B. M., Kim, S.-H., Katz, M., Charlop-Powers, Z., Owen, J. G., Ternei, M. A., et al. (2018). Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* 3, 415–422. doi: 10.1038/s41564-018-0110-1

Ji, C. H., Kim, J. P., and Kang, H. S. (2018). Library of synthetic *Streptomyces* regulatory sequences for use in promoter engineering of natural product biosynthetic gene clusters. *ACS Synth. Biol.* 7, 1946–1955. doi: 10.1021/acssynbio.8b00175

Jiang, M., Yin, M., Wu, S., Han, X., Ji, K., Wen, M., et al. (2017). GdmRIII, a TetR family transcriptional regulator, controls Geldanamycin and Elaiophylin biosynthesis in *Streptomyces autolyticus* CGMCC0516. *Sci. Rep.* 7:4803. doi: 10.1038/s41598-017-05073-x

Jiang, W., Zhao, X., Gabrieli, T., Lou, C., Ebenstein, Y., and Zhu, T. F. (2015). Cas9-Assisted Targeting of chromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* 6:8101. doi: 10.1038/ncomms9101

Jones, A. C., Gust, B., Kulik, A., Heide, L., Buttner, M. J., and Bibb, M. J. (2013). Phage P1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. *PLoS One* 8:e69319. doi: 10.1371/journal.pone.0069319

Kallifidas, D., Jiang, G., Ding, Y., and Luesch, H. (2018). Rational engineering of *Streptomyces albus* J1074 for the overexpression of secondary metabolite gene clusters. *Microb. Cell Fact* 17:25. doi: 10.1186/s12934-018-0874-2

Kang, H. S., Charlop-Powers, Z., and Brady, S. F. (2016). Multiplexed CRISPR/Cas9- and TAR-mediated promoter engineering of natural product biosynthetic gene clusters in yeast. *ACS Synth. Biol.* 5, 1002–1010. doi: 10.1021/acssynbio.6b00080

Katz, M., Hover, B. M., and Brady, S. F. (2016). Culture-independent discovery of natural products from soil metagenomes. *J. Ind. Microbiol. Biotechnol.* 43, 129–141.

Kenshole, E., Herisse, M., Michael, M., and Pidot, S. J. (2021). Natural product discovery through microbial genome mining. *Curr. Opin. Chem. Biol.* 60, 47–54. doi: 10.1016/j.cbpa.2020.07.010

Kim, H., Ji, C. H., Je, H. W., Kim, J. P., and Kang, H. S. (2020). mpCRISTAR: multiple plasmid approach for CRISPR/Cas9 and TAR-mediated multiplexed refactoring of natural product biosynthetic gene clusters. *ACS Synth. Biol.* 9, 175–180. doi: 10.1021/acssynbio.9b00382

Kouprina, N., and Larionov, V. (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* 125, 621–632. doi: 10.1007/s00412-016-0588-3

Kouprina, N., and Larionov, V. (2019). TAR cloning: perspectives for functional genomics, biomedicine, and biotechnology. *Mol. Ther. Methods Clin. Dev.* 14, 16–26. doi: 10.1016/j.omtm.2019.05.006

Krause, J., Handayani, I., Blin, K., Kulik, A., and Mast, Y. (2020). Disclosing the potential of the SARP-Type regulator PapR2 for the activation of antibiotic gene clusters in *Streptomycetes*. *Front. Microbiol.* 11:225. doi: 10.3389/fmicb.2020.00225

Labes, G., Bibb, M., and Wohlleben, W. (1997). Isolation and characterization of a strong promoter element from the *Streptomyces ghanaensis* phage I19 using the gentamicin resistance gene (*aacC1*) of Tn 1696 as reporter. *Microbiology* 143(Pt 5), 1503–1512. doi: 10.1099/00221287-143-5-1503

Lee, N., Hwang, S., Kim, J., Cho, S., Palsson, B., and Cho, B.-K. (2020). Mini review: genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in *Streptomyces*. *Comput. Struct. Biotechnol. J.* 18, 1548–1556. doi: 10.1016/j.csbj.2020.06.024

Lee, N., Hwang, S., Lee, Y., Cho, S., Palsson, B., and Cho, B. K. (2019). Synthetic biology tools for novel secondary metabolite discovery in *Streptomyces*. *J. Microbiol. Biotechnol.* 29, 667–686. doi: 10.4014/jmb.1904.04015

Lewis, K. (2020). The science of antibiotic discovery. *Cell* 181, 29–45. doi: 10.1016/j.cell.2020.02.056

Li, L., Liu, X., Jiang, W., and Lu, Y. (2019). Recent advances in synthetic biology approaches to optimize production of bioactive natural products in *Actinobacteria*. *Front. Microbiol.* 10:2467. doi: 10.3389/fmicb.2019.02467

Li, Z. Y., Bu, Q. T., Wang, J., Liu, Y., Chen, X. A., Mao, X. M., et al. (2019). Activation of anthrachamycin biosynthesis in *Streptomyces chattanoogensis* L10 by site-directed mutagenesis of *rpoB*. *J. Zhejiang. Univ. Sci. B* 20, 983–994. doi: 10.1631/jzus.B1900344

Li, L., Zhao, Y., Ruan, L., Yang, S., Ge, M., Jiang, W., et al. (2015). A stepwise increase in pristinamycin II biosynthesis by *Streptomyces pristinaespiralis* through combinatorial metabolic engineering. *Metab. Eng.* 29, 12–25. doi: 10.1016/j.ymben.2015.02.001

Li, S., Liu, Q., Zhong, Z., Deng, Z., and Sun, Y. (2020). Exploration of Hygromycin B biosynthesis utilizing CRISPR-Cas9-associated base editing. *ACS Chem. Biol.* 15, 1417–1423. doi: 10.1021/acschembio.0c00071

Liu, H., Jiang, H., Haltli, B., Kulowski, K., Muszynska, E., Feng, X., et al. (2009). Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile *Escherichia coli-Streptomyces* artificial chromosome vector, pSBAC. *J. Nat. Prod.* 72, 389–395. doi: 10.1021/np8006149

Liu, L., Cheng, Y., Lyu, M., Zhao, X., Wen, Y., Li, J., et al. (2019). AveI, an AtrA homolog of *Streptomyces avermitilis*, controls avermectin and oligomycin production, melanogenesis, and morphological differentiation. *Appl. Microbiol. Biotechnol.* 103, 8459–8472. doi: 10.1007/s00253-019-10062-3

Loman, N. J., and Pallen, M. J. (2015). Science China life sciences twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* 13, 787–794. doi: 10.1038/nrmicro3565

Luo, Y., Enghiad, B., and Zhao, H. (2016). New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Nat. Prod. Rep.* 33, 174–182. doi: 10.1039/c5np00085h

Luo, Y., Huang, H., Liang, J., Wang, M., Lu, L., Shao, Z., et al. (2013). Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* 4:2894. doi: 10.1038/ncomms3894

Luo, Y., Zhang, L., Barton, K. W., and Zhao, H. (2015). Systematic Identification of a Panel of Strong Constitutive Promoters from *Streptomyces albus*. *ACS Synth. Biol.* 4, 1001–1010. doi: 10.1021/acssynbio.5b00016

Mao, D., Okada, B. K., Wu, Y., Xu, F., and Seyedsayamdost, M. R. (2018). Recent advances in activating silent biosynthetic gene clusters in bacteria. *Curr. Opin. Microbiol.* 45, 156–163. doi: 10.1016/j.mib.2018.05.001

Marín, L., Gutiérrez-Del-Río, I., Yagüe, P., Manteca, Á, Villar, C. J., and Lombó, F. (2017). De novo biosynthesis of Apigenin, Luteolin, and Eriodictyol in the *Actinomycete Streptomyces albus* and production improvement by feeding and spore conditioning. *Front. Microbiol.* 8:921. doi: 10.3389/fmicb.2017.00921

Myronovskyi, M., and Luzhetskyy, A. (2016). Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* 33, 1006–1019. doi: 10.1039/c6np00002a

Myronovskyi, M., and Luzhetskyy, A. (2019). Heterologous production of small molecules in the optimized *Streptomyces* hosts. *Nat. Prod. Rep.* 36, 1281–1294. doi: 10.1039/c9np00023b

Myronovskyi, M., Rosenkranzer, B., Nadmid, S., Pujic, P., Normand, P., and Luzhetskyy, A. (2018). Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* 49, 316–324. doi: 10.1016/j.ymben.2018.09.004

Nah, H. J., Pyeon, H. R., Kang, S. H., Choi, S. S., and Kim, E. S. (2017). Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. *Front. Microbiol.* 8:394. doi: 10.3389/fmicb.2017.00394

Ochi, K. (2017). "Chapter 9 - Cryptic pathways and implications for novel drug discovery," in *Microbial Resources*, ed. I. Kurtböke (Cambridge, MA: Academic Press), 189–203.

Onaka, H. (2017). Novel antibiotic screening methods to awaken silent or cryptic secondary metabolic pathways in *Actinomycetes*. *J. Antibiot.* 70, 865–870. doi: 10.1038/ja.2017.51

Orr-Weaver, T. L., Szostak, J. W., and Rothstein, R. J. (1981). Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci. U.S.A.* 78, 6354–6358. doi: 10.1073/pnas.78.10.6354

Pyeon, H. R., Nah, H. J., Kang, S. H., Choi, S. S., and Kim, E. S. (2017). Heterologous expression of pikromycin biosynthetic gene cluster using *Streptomyces* artificial chromosome system. *Microb. Cell Fact* 16:96. doi: 10.1186/s12934-017-0708-7

Ren, H., Shi, C., and Zhao, H. (2020). Computational tools for discovering and engineering natural product biosynthetic pathways. *iScience* 23, 100795. doi: 10.1016/j.isci.2019.100795

Rodríguez-García, A., Combes, P., Pérez-Redondo, R., Smith, M. C., and Smith, M. C. (2005). Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Res.* 33:e87. doi: 10.1093/nar/gni086

Rutledge, P. J., and Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* 13, 509–523. doi: 10.1038/nrmicro3496

Seyedsayamdost, M. R. (2014). High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7266–7271. doi: 10.1073/pnas.1400019111

Shao, Z., Rao, G., Li, C., Abil, Z., Luo, Y., and Zhao, H. (2013). Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth. Biol.* 2, 662–669. doi: 10.1021/sb400058n

Siu, K. H., and Chen, W. (2019). Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function. *Nat. Chem. Biol.* 15, 217–220. doi: 10.1038/s41589-018-0186-1

Song, C., Luan, J., Li, R., Jiang, C., Hou, Y., Cui, Q., et al. (2020). RedEx: a method for seamless DNA insertion and deletion in large multimodular polyketide synthase gene clusters. *Nucleic Acids Res.* 48:gkaa956.

Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G., Nihira, T., et al. (2005). A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol. Microbiol.* 56, 465–479. doi: 10.1111/j.1365-2958.2005.04543.x

Tan, G. Y., and Liu, T. (2017). Rational synthetic pathway refactoring of natural products biosynthesis in *Actinobacteria*. *Metab. Eng.* 39, 228–236. doi: 10.1016/j.ymben.2016.12.006

Tanaka, Y., Kasahara, K., Hirose, Y., Murakami, K., Kugimiya, R., and Ochi, K. (2013). Activation and products of the cryptic secondary metabolite biosynthetic gene clusters by rifampin resistance (*rpoB*) mutations in *Actinomycetes*. *J. Bacteriol.* 195, 2959–2970. doi: 10.1128/JB.00147-13

Tao, W., Chen, L., Zhao, C., Wu, J., Yan, D., Deng, Z., et al. (2019). *In Vitro* packaging mediated one-step targeted cloning of natural product pathway. *ACS Synth. Biol.* 8, 1991–1997. doi: 10.1021/acssynbio.9b00248

Tao, W., Yang, A., Deng, Z., and Sun, Y. (2018). CRISPR/Cas9-based editing of *Streptomyces* for discovery, characterization, and production of natural products. *Front. Microbiol.* 9:1660. doi: 10.3389/fmicb.2018.01660

Thong, W. L., Shin-Ya, K., Nishiyama, M., and Kuzuyama, T. (2016). Methylbenzene-containing Polyketides from a *Streptomyces* that spontaneously acquired rifampicin resistance: structural elucidation and biosynthesis. *J. Nat. Prod.* 79, 857–864. doi: 10.1021/acs.jnatprod.5b00922

Thong, W. L., Shin-Ya, K., Nishiyama, M., and Kuzuyama, T. (2018). Discovery of an antibacterial isoindolinone-containing tetracyclic polyketide by cryptic gene activation and characterization of its biosynthetic gene cluster. *ACS Chem. Biol.* 13, 2615–2622. doi: 10.1021/acschembio.8b00553

Tomm, H. A., Ucciferri, L., and Ross, A. C. (2019). Advances in microbial culturing conditions to activate silent biosynthetic gene clusters for novel metabolite production. *J. Ind. Microbiol. Biotechnol.* 46, 1381–1400. doi: 10.1007/s10295-019-02198-y

Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S. Y. (2015). CRISPR-Cas9 Based engineering of actinomycetal genomes. *ACS Synth. Biol.* 4, 1020–1029. doi: 10.1021/acssynbio.5b00038

Tong, Y., Weber, T., and Lee, S. Y. (2019a). CRISPR/Cas-based genome engineering in natural product discovery. *Nat. Prod. Rep.* 36, 1262–1280. doi: 10.1039/C8NP00089A

Tong, Y., Whitford, C. M., Robertsen, H. L., Blin, K., Jørgensen, T. S., Klitgaard, A. K., et al. (2019b). Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20366–20375. doi: 10.1073/pnas.1913493116

Van Santen, J. A., Kautsar, S. A., Medema, M. H., and Linington, R. G. (2020). Microbial natural product databases: moving forward in the multi-omics era. *Nat. Prod. Rep.* doi: 10.1039/D0NP00053A

Walsh, C. T., and Fischbach, M. A. (2010). Natural products version 2.0: connecting genes to molecules. *J. Am. Chem. Soc.* 132, 2469–2493. doi: 10.1021/ja909118a

Wang, B., Guo, F., Dong, S. H., and Zhao, H. (2019). Activation of silent biosynthetic gene clusters using transcription factor decoys. *Nat. Chem. Biol.* 15, 111–114. doi: 10.1038/s41589-018-0187-0

Wang, K., Chen, X. A., Li, Y. Q., and Mao, X. M. (2019). Identification of a secondary metabolism-responsive promoter by proteomics for over-production of natamycin in *Streptomyces*. *Arch. Microbiol.* 201, 1459–1464. doi: 10.1007/s00203-019-01710-3

Wang, H., Li, Z., Jia, R., Yin, J., Li, A., Xia, L., et al. (2018). ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Res.* 46:e28. doi: 10.1093/nar/gkx1249

Wang, W., Li, X., Wang, J., Xiang, S., Feng, X., and Yang, K. (2013). An engineered strong promoter for *Streptomycetes*. *Appl. Environ. Microbiol.* 79, 4484–4492. doi: 10.1128/AEM.00985-13

Xia, H. Y., Li, X. F., Li, Z. Q., Zhan, X. Q., Mao, X. M., and Li, Y. Q. (2020). The application of regulatory cascades in *Streptomyces*: yield enhancement and metabolite mining. *Front. Microbiol.* 11:406. doi: 10.3389/fmicb.2020.00406

Xu, J., Song, Z., Xu, X., Ma, Z., Bechthold, A., and Yu, X. (2019). ToyA, a positive pathway-specific regulator for toyocamycin biosynthesis in *Streptomyces diastatochromogenes* 1628. *Appl. Microbiol. Biotechnol.* 103, 7071–7084. doi: 10.1007/s00253-019-09959-w

Xu, J., Tozawa, Y., Lai, C., Hayashi, H., and Ochi, K. (2002). A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Genet. Genom.* 268, 179–189. doi: 10.1007/s00438-002-0730-1

Xu, J., Zhang, J., Zhuo, J., Li, Y., Tian, Y., and Tan, H. (2017). Activation and mechanism of a cryptic oviedomycin gene cluster via the disruption of a global regulatory gene, *adpA*, in *Streptomyces ansochromogenes*. *J. Biol. Chem.* 292, 19708–19720. doi: 10.1074/jbc.M117.809145

Xu, M., Wang, Y. M., Zhao, Z. L., Gao, G. X., Huang, S. X., Kang, Q. J., et al. (2016). Functional Genome mining for metabolites encoded by large gene clusters through Heterologous expression of a whole-genome bacterial artificial chromosome library in *Streptomyces spp*. *Appl. Environ. Microbiol.* 82, 5795–5805. doi: 10.1128/aem.01383-16

Xu, M., and Wright, G. D. (2019). Heterologous expression-facilitated natural products' discovery in *Actinomycetes*. *J. Ind. Microbiol. Biotechnol.* 46, 415–431. doi: 10.1007/s10295-018-2097-2

Zhang, J., Zhang, D., Zhu, J., Liu, H., Liang, S., and Luo, Y. (2020). Efficient multiplex genome editing in *Streptomyces* via engineered CRISPR-Cas12a systems. *Front. Bioeng. Biotechnol.* 8:726. doi: 10.3389/fbioe.2020.00726

Zhang, Q., Ren, J.-W., Wang, W., Zhai, J. A., Yang, J., Liu, N., et al. (2020). A versatile transcription-translation in one approach for activation of cryptic biosynthetic gene clusters. *ACS Chem. Biol.* 15, 2551–2557. doi: 10.1021/acschembio.0c00581

Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., et al. (2017). CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* 13, 607–609. doi: 10.1038/nchembio.2341

Zhang, Y., Buchholz, F., Muyrers, J. P., and Stewart, A. F. (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* 20, 123–128. doi: 10.1038/2417

Zhang, Y., Muyrers, J. P., Testa, G., and Stewart, A. F. (2000). DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* 18, 1314–1317. doi: 10.1038/82449

Zhao, Q., Wang, L., and Luo, Y. (2019). Recent advances in natural products exploitation in *Streptomyces* via synthetic biology. *Eng. Life Sci.* 19, 452–462. doi: 10.1002/elsc.201800137

Zhao, Y., Li, G., Chen, Y., and Lu, Y. (2020). Challenges and advances in genome editing technologies in *Streptomyces*. *Biomolecules* 10:734. doi: 10.3390/biom10050734

Zhao, Y. W., Tian, J. Z., Zheng, G. S., Chen, J., Sun, C. W., Yang, Z. Y., et al. (2020). Multiplex genome editing using a dCas9-cytidine deaminase fusion in *Streptomyces*. *Sci. China Life Sci.* 63, 1053–1062. doi: 10.1007/s11427-019-1559-y

Zhou, M., Jing, X., Xie, P., Chen, W., Wang, T., Xia, H., et al. (2012). Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of *Streptomyces coelicolor*. *FEMS Microbiol. Lett.* 333, 169–179. doi: 10.1111/j.1574-6968.2012.02609.x

Zhou, Y., Murphy, A. C., Samborskyy, M., Prediger, P., Dias, L. C., and Leadlay, P. F. (2015). Iterative mechanism of macrodiolide formation in the anticancer compound conglobatin. *Chem. Biol.* 22, 745–754.

Zhu, S., Duan, Y., and Huang, Y. (2019). The Application of ribosome engineering to natural product discovery and yield improvement in *Streptomyces*. *Antibiotics* 8:133. doi: 10.3390/antibiotics8030133

Zhu, Z., Li, H., Yu, P., Guo, Y., Luo, S., Chen, Z., et al. (2017). SlnR is a positive pathway-specific regulator for salinomycin biosynthesis in *Streptomyces albus*. *Appl. Microbiol. Biotechnol.* 101, 1547–1557. doi: 10.1007/s00253-016-7918-5

Ziermann, R., and Betlach, M. C. (1999). Recombinant polyketide synthesis in *Streptomyces*: engineering of improved host strains. *Biotechniques* 26, 106–110. doi: 10.2144/99261st05

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 09 April 2021
doi: 10.3389/fbioe.2021.673844

Identification of Gradient Promoters of *Gluconobacter oxydans* and Their Applications in the Biosynthesis of 2-Keto-L-Gulonic Acid

Yue Chen[1,2,3,4], *Li Liu*[1,2,4], *Shiqin Yu*[1,2,3,4], *Jianghua Li*[1,3], *Jingwen Zhou*[1,2,4]* and *Jian Chen*[1,2]*

[1] Key Laboratory of Industrial Biotechnology, Ministry of Education and School of Biotechnology, Jiangnan University, Wuxi, China, [2] National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi, China, [3] Science Center for Future Foods, Jiangnan University, Wuxi, China, [4] Jiangsu Provisional Research Center for Bioactive Product Processing Technology, Jiangnan University, Wuxi, China

OPEN ACCESS

Edited by:
Dawei Zhang,
Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, China

Reviewed by:
Sanjay Kumar Singh Patel,
Konkuk University, South Korea
Hui Xu,
Institute of Applied Ecology (CAS), China

***Correspondence:**
Jingwen Zhou
zhoujw1982@jiangnan.edu.cn
Jian Chen
jchen@jiangnan.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 28 February 2021
Accepted: 22 March 2021
Published: 09 April 2021

Citation:
Chen Y, Liu L, Yu S, Li J, Zhou J and Chen J (2021) Identification of Gradient Promoters of Gluconobacter oxydans and Their Applications in the Biosynthesis of 2-Keto-L-Gulonic Acid. Front. Bioeng. Biotechnol. 9:673844. doi: 10.3389/fbioe.2021.673844

The acetic acid bacterium *Gluconobacter oxydans* is known for its unique incomplete oxidation and therefore widely applied in the industrial production of many compounds, e.g., 2-keto-L-gulonic acid (2-KLG), the direct precursor of vitamin C. However, few molecular tools are available for metabolically engineering *G. oxydans*, which greatly limit the strain development. Promoters are one of vital components to control and regulate gene expression at the transcriptional level for boosting production. In this study, the low activity of SDH was found to hamper the high yield of 2-KLG, and enhancing the expression of SDH was achieved by screening the suitable promoters based on RNA sequencing data. We obtained 97 promoters from *G. oxydans*'s genome, including two strong shuttle promoters and six strongest promoters. Among these promoters, P_{3022} and P_{0943} revealed strong activities in both *Escherichia coli* and *G. oxydans*, and the activity of the strongest promoter (P_{2703}) was about threefold that of the other reported strong promoters of *G. oxydans*. These promoters were used to overexpress SDH in *G. oxydans* WSH-003. The titer of 2-KLG reached 3.7 g/L when SDH was under the control of strong promoters P_{2057} and P_{2703}. This study obtained a series of gradient promoters, including two strong shuttle promoters, and expanded the toolbox of available promoters for the application in metabolic engineering of *G. oxydans* for high-value products.

Keywords: 2-keto-L-gulonic acid, *Gluconobacter oxydans*, promoters, L-sorbose, sorbose dehydrogenase

INTRODUCTION

Gluconobacter oxydans has been widely applied in the industrial production of L-sorbose from D-sorbitol (De Wulf et al., 2000), dihydroxyacetone from glycerol (De La Morena et al., 2020), 1-amino-L-sorbose from 1-amino-D-sorbitol (Schedel, 2000), and levan-type fructans from sucrose (Hövels et al., 2020). Furthermore, *G. oxydans* is also an excellent workhorse for the biosynthesis of 2-keto-D-gluconate (Li et al., 2016; Zhou et al., 2020), 5-keto-D-gluconate (Merfort et al., 2006), xylonic acid (Hahn et al., 2020; Shen et al., 2020), 5-keto-D-fructose (Battling et al., 2020;

Hoffmann et al., 2020), and many other products (Deppenmeier et al., 2002; De Muynck et al., 2007). The wide applications of *G. oxydans* are mainly due to its unique dehydrogenases in the periplasm (Deppenmeier et al., 2002). Many protein engineering approaches have been used to improve the catalytic efficiency of these dehydrogenases, including enzyme immobilization (Kim et al., 2016), cofactor regeneration (Gao et al., 2019), ligand docking, and molecular dynamics simulations (Selvaraj et al., 2016). On the other hand, many metabolic engineering strategies of *G. oxydans* were based on the overexpression of related dehydrogenases or enzymes associated with the respiratory chains (Li et al., 2016; Yuan et al., 2016). However, the expression of key enzymes were often impeded by the accessible promoters.

Though the first genome of *G. oxydans* was reported in 2005 (Prust et al., 2005), only a few studies on the promoters of *G. oxydans* have been carried out. Generally, most studies directly selected some promoters from high expression level of genes. Nishikura-Imamura et al. (2014) cloned the putative promoter region of *G. oxydans* PQQ-dependent alcohol dehydrogenase (P_{adhAB}) to overexpress 3-dehydroquinate dehydratase. Mientus et al. (2017) characterized promoters of six membrane-bound dehydrogenases of *G. oxydans* 621H, and used the constitutive promoter of the alcohol dehydrogenase and the glucose-repressed promoter of inositol dehydrogenase to construct a shuttle vector system. Though some progress were achieved in promoter discovery, it was hard to apply the promoters to other metabolic pathways because the promoters were relatively weak without systematic comparison. Saito et al., found some strong promoters of *G. oxydans*, such as P_{tufB}, P_{0169}, and P_{264} (Saito et al., 1997; Yuan et al., 2016; Blank and Schweiger, 2018). Moreover, Kallnik and Hu reported some promoters of different strengths in *G. oxydans* (Kallnik et al., 2010; Hu et al., 2015). Nevertheless, the available promoters are still insufficient, especially the strong promoters are highly needed to support engineering *G. oxydans* for their industrial application.

RNA sequencing (RNA-Seq) in the study of prokaryotic and eukaryotic organisms has become more accessible in the last decade (Poulsen and Vinther, 2018; Stark et al., 2019). It has become an excellent strategy to mine strong promoters in many microorganisms. Lee et al. (2015) screened a novel strong promoter P_{TN0510} from *Thermococcus onnurineus* by RNA-Seq and applied it to the production of H_2. Liao et al. (2015) identified a strong promoter, P_{r2}, from the RNA-Seq data of *Bacillus amyloliquefaciens* and verified it by measuring beta-galactosidase activity. Several studies about the transcriptome analysis of *G. oxydans* has been reported to reveal the secretion pathways of PQQ (Wan et al., 2017) and the response to osmotic and oxidative stress of 2-keto-L-gulonic acid (2-KLG) (Fang et al., 2020). Kranz et al. (2018) provided deep insights into the transcriptional landscapes of *G. oxydans* including promoters and other regulatory elements. However, no further experimental studies were performed to characterize these promoters and regulatory elements. Thus, RNA-Seq of *G. oxydans* WSH-003 was first conducted in this study, followed with the characterization of promoters by using mCherry as a report to compare the strength of the screened promoters.

2-KLG is an important precursor of vitamin C in industry (Wang et al., 2018). However, there are only a few *G. oxydans* strains that can produce 2-KLG naturally (Hoshino et al., 1990; Saito et al., 1998; Chen et al., 2019), although many sequenced *G. oxydans* possess the entire set of 2-KLG biosynthesis genes (Wang et al., 2018). In our previous study, we identified the key SDH from a *G. oxydans* that was able to naturally produce 2-KLG, and successfully constructed a high-throughput screening platform for an FAD-dependent SDH (Shan et al., 2020). Different from SSDHs from *Ketogulonigenium vulgare*, SDH from *G. oxydans* showed higher substrate specificity to L-sorbose and did not require PQQ as a cofactor (Saito et al., 1997; Wang et al., 2018). In the present study, a group of gradient promoters was identified and applied in the biosynthesis of 2-KLG in the strain *G. oxydans* WSH-003. The titer of 2-KLG reached 3.7 g/L when used the strongest promoter (P_{2703}) to overexpress SDH. The results implied the low expression level of SDH may be the main problem for 2-KLG production in many *G. oxydans* strains. In conclusion, this study obtained a series of gradient promoters, and these promoters revealed promising prospects in metabolic engineering of *G. oxydans* for high-value products.

MATERIALS AND METHODS

Genes, Plasmids, and Strains

Escherichia coli JM109 was used for plasmid construction. *G. oxydans* WSH-003 was used for PCR amplification of promoters and protein expression. *G. oxydans* ATCC 621H was used for PCR amplification of promoters. *G. oxydans* WSH-004 was screened in our previous research (Chen et al., 2019). *G. oxydans* WSH-003-Δ*gdh* was used for 2-KLG production. The plasmids p2-5 and pBBR1MCS-5 were used to overexpress mCherry (Li et al., 2020) and sorbose dehydrogenase in *G. oxydans*, respectively. All strains and plasmids are listed in **Table 1**. The nucleotide sequences of p2-5 and pBBR1MCS-5 were listed in **Supplementary Table 5**.

TABLE 1 | Plasmids and strains used in this study.

Plasmids or strains	Characteristics	Sources
Plasmids		
p2-5	Km^R, shuttle vector of *G. oxydans*, used for overexpressing of mCherry	Stored in Lab
pBBR1MCS-5	Gm^R, shuttle vector of *G. oxydans*, used for overexpressing of sorbose dehydrogenase	Yuan et al., 2016
Strains		
E. coli JM109	Used for plasmid construction	Stored in Lab
G. oxydans WSH-003	Cef^R, used for PCR amplification of promoters and protein expression	Hu et al., 2015
G. oxydans WSH-004	Cef^R, used for PCR amplification of sorbose dehydrogenase	Chen et al., 2019
G. oxydans ATCC 621H	Cef^R, used for PCR amplification of promoters	Prust et al., 2005
G. oxydans WSH-003-Δ*gdh*	Cef^R and Km^R, deletion of glucose dehydrogenase (GenBank: AHKI01000025, from 1155 to 3575), used for 2-KLG producing	Stored in Lab

RNA Sequencing and Data Analysis

The strain *G. oxydans* WSH-003 was cultured to mid-log phase in sorbitol medium (50 g/L sorbitol and 10 g/L yeast extract) at 30°C with shaking at 220 rpm. Then the cells were harvested and washed twice with PBS. The total RNA was extracted by RNeasy Mini Kit (Qiagen, Hilden, Germany), and ribosomal RNAs were removed by Ribo-Zero^{TM} rRNA Removal Kits (Epicentre, Wisconsin, United States). The RNA sequencing libraries were constructed by TruSeq RNA Sample Preparation Kit v2 (Illumina, California, United States) and sequenced on MiSeq (Illumina, California, United States) using MiSeq Reagent Kit v3 (Illumina, California, United States). RNA sequencing was performed by Shanghai Biotechnology Corporation (Shanghai Biotechnology Co., Shanghai, China). The abundance of transcripts was determined using bowtie2 (Langmead and Salzberg, 2012; Bolger et al., 2014) and cufflinks (Trapnell et al., 2010) by mapping the appropriate reads to the genome of *G. oxydans* WSH-003 (Gao et al., 2012).

Genetic Operations

All the promoters were worked with a native ribosomal binding site (RBS), because it was hard to find a proper RBS with guaranteed strength. Promoters P_{tufB} and P_{dnak} and all the screened potential promoters were obtained by PCR amplification from the genomic DNA of *G. oxydans* WSH-003. Promoters P_{264} and P_{hp0169} were obtained by PCR amplification from the genomic DNA of *G. oxydans* ATCC 621H. The gene *sdh* was PCR-amplified from the genomic DNA of *G. oxydans* WSH-004. The gene *mCherry* was kept in our laboratory and obtained by PCR amplification. The *mCherry* gene was first ligated into the vector p2-5 to form the skeleton plasmid p2-5-mCherry by a one-step cloning kit (Takara, Dalian, China). Then different promoters were individually inserted into the plasmid p2-5-mCherry by the one-step cloning kit (Takara, Dalian, China). The *sdh* gene and gradient promoters were ligated into the vector pBBR1MCS-5 in the same way. All promoters and genes were verified by Sanger sequencing (Sangon Biotech, Shanghai, China). All vectors were constructed and amplified in the strain *E. coli* JM109. The vectors p2-5 and pBBR1MCS-5 were transferred by electroporation into *G. oxydans* WSH-003 (Zhang et al., 2010), which were selected using kanamycin and gentamycin, respectively. All primers are listed in **Supplementary Table 1**.

Fluorescence Intensity Assay

Single colonies of *G. oxydans* WSH-003 were picked into 14 mL tubes containing 2 mL of sorbitol medium and cultured for 24 hours at 30°C with shaking at 220 rpm. Then 2% of these cultures were inoculated into a 250 mL flask containing 25 mL of sorbitol medium and cultured at 30°C with shaking at 220 rpm. The cell fluorescence and cell density (OD_{600}) were measured every 4 hours on a Synergy H1 Hybrid Multi-Mode Microplate Reader (BioTek Instruments, Winooski, VT, United States) with excitation and emission wavelengths of 580 and 610 nm, respectively. The relative activity of mCherry was defined as the ratio of relative fluorescence unit (RFUs)

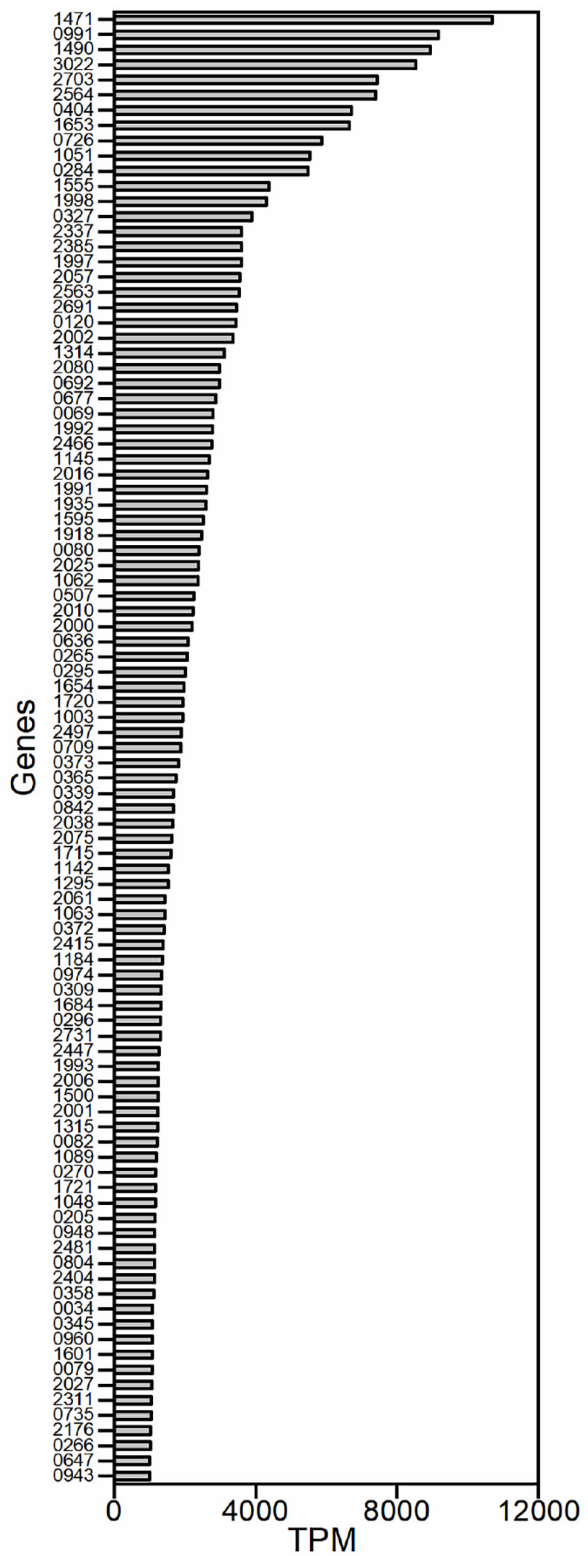


FIGURE 1 | The strongly transcribed genes of *G. oxydans* WSH-003 obtained by RNA sequencing. The promoters were selected based on transcript per million (TPM) values of the transcriptome. The abscissa axis shows the TPM values. The ordinate axis shows the different genes in *G. oxydans* WSH-003.

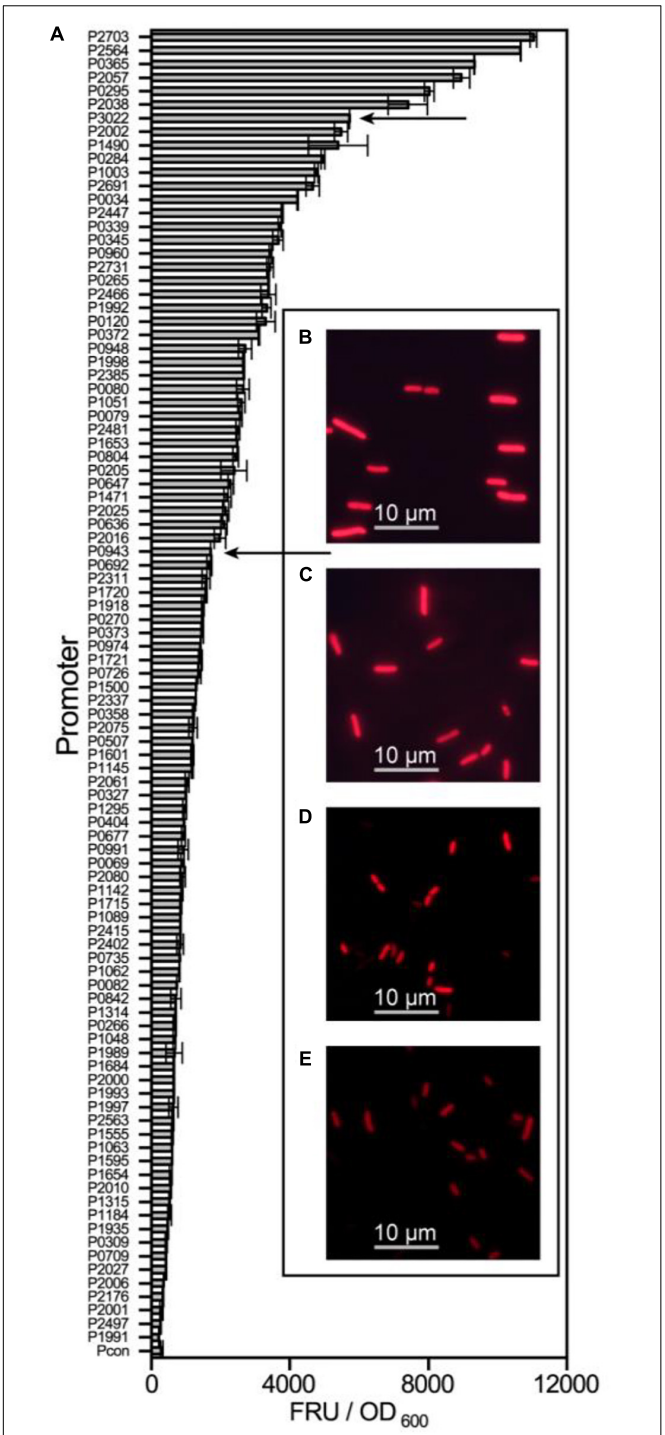


FIGURE 2 | The strength determination of screened promoters of *G. oxydans* WSH-003. The strength of screened promoters and microscope pictures of two shuttle promoters. **(A)** The strength of screened promoters as measured by evaluating mCherry activity. The abscissa axis shows the relative activity of mCherry. The ordinate axis shows different promoters of *G. oxydans* WSH-003. Arrows indicate the two strong shuttle promoters P_{3022} and P_{0943}. **(B)** *E. coli* JM109 expressing mCherry under the control of P_{3022}. **(C)** *E. coli* JM109 expressing mCherry under the control of P_{0943}. **(D)** *G. oxydans* WSH-003 expressing mCherry under the control of P_{3022}. **(E)** *G. oxydans* WSH-003 expressing mCherry under the control of P_{0943}.

divided by the optical density (OD_{600}). The strain *G. oxydans* WSH-003 harboring p2-5-mCherry without promoters was used as the control.

Culture Conditions for 2-KLG Production

The fermentation medium was formed with sorbitol medium (50 g/L sorbitol and 10 g/L yeast extract) containing 20 g/L $CaCO_3$. Single colonies of *G. oxydans* WSH-003 were picked into 250 mL flasks containing 25 mL of sorbitol medium and cultured for 24 hours at 30°C with shaking at 220 rpm. Then, 10% of these cultures were inoculated into 250 mL flasks containing 25 mL of fermentation medium and cultured at 30°C with shaking at 220 rpm.

Analysis of 2-KLG Production

The cell concentration was measured using a Microplate Reader (BioTek Instruments, Winooski, VT, United States). The concentrations of D-sorbitol, L-sorbose, and 2-KLG were detected by HPLC using an Aminex HPX-87H column (BioRad, Hercules, CA, United States) at 35°C with 5 mmol/L H_2SO_4 as the eluent at a flow rate of 0.5 mL/min (Chen et al., 2019).

Statistical Analysis

The results were interpreted with mean values and its standard. The paired two-tailed Student *t* tests were performed to demonstrate statistically significant differences between data points. A p-value of $\leq$ 0.05 was thought to be statistically significant. For data illustration, bar charts with error bars were used.

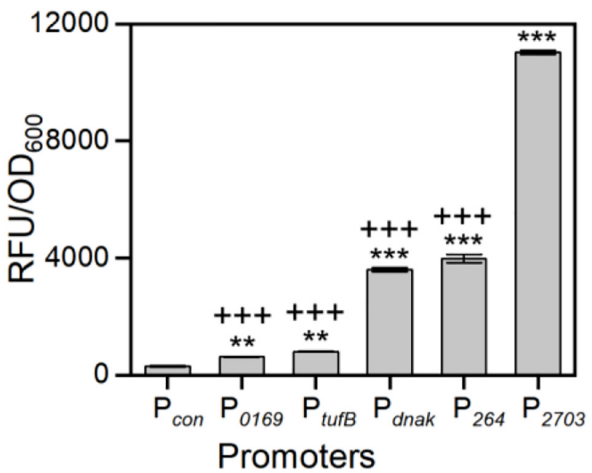


FIGURE 3 | The comparison of screened strongest promoter with reported strong promoters. The abscissa axis shows different promoters of *G. oxydans*. P_{con} is the negative control. P_{0169} and P_{264} are reported strong promoters of *G. oxydans* ATCC 621H. P_{tufB} and P_{dnak} are reported strong promoters of *G. oxydans* WSH-003. P_{2703} is the strongest promoter screened in this study. The ordinate axis shows the relative activity of mCherry. $^{**}P < 0.01$ compared with the P_{con} by two-tailed *t* test; $^{***}P < 0.001$ compared with the P_{con} by two-tailed *t* test. $^{+++}P < 0.001$ compared with the P_{2703} by two-tailed *t* test.

RESULTS

The RNA Sequencing Results

The isolated mRNAs of *G. oxydans* WSH-003 were subjected to high-throughput Illumina paired-end sequencing to obtain a global view of the transcriptome after removing ribosomal RNAs. The RNA sequencing data was submitted to sequence read archive (SRA) with the accession number of PRJNA706889. A total of 19.84 million mapped reads with an average length of 100 bp were obtained; 19.31 million reads were uniquely mapped to the genome of *G. oxydans* WSH-003, which represented a 500-fold coverage of the genome. The transcriptome data were analyzed with the software bowtie2 (Langmead and Salzberg, 2012; Bolger et al., 2014) and cufflinks (Trapnell et al., 2010). Only 188 genes were transcribed with transcript per million (TPM) values higher than 1000, and nearly 95% of the predicted 3545 genes were transcribed with TPM values below 1000 (**Supplementary Table 2**).

Because we aimed to identify strong promoters in *G. oxydans*, the genes that exhibited strong transcriptional activity were studied. Genes were excluded from our analysis if they encoded tRNAs or lacked a RBS. Genes that belonged to a gene cluster were also excluded because the same promoter probably controlled their transcription. Gene clusters were defined as contiguous genes with similar functions and with spacers smaller than 50 nucleotides. A total of 97 potential promoters were obtained based on the TPM values in the transcriptome (**Figure 1**). The sequences around 500 bp upstream of the open reading frame (ORF) were chosen as the potential promoters because little information on the promoter elements of *G. oxydans* has been reported.

Evaluation of Promoter Strength by Measuring mCherry Expression

All 97 potential promoters were transferred into *G. oxydans* WSH-003 to determine their strength by measuring mCherry expression. The fluorescence intensity was assayed every 4 hours. The highest value of relative mCherry activity was defined as the relative strength of the promoter. Most of the screened promoters showed remarkable intensity compared to the control (**Figure 2**). Among the promoters, the six strongest promoters were P_{2703}, P_{2564}, P_{0365}, P_{2057}, P_{0295}, and P_{2038}. Besides, it was found that most of the screened promoters had the highest strength at about 36 hours, when the strain was grown in stationary phase (**Supplementary Figure 1**).

As mentioned previously, a few strong *G. oxydans* promoters have been reported (Hu et al., 2015; Li et al., 2016; Blank and Schweiger, 2018). To verify the strength of our strongest screened promoters, we also obtained four reported strong promoters (P_{tufB}, P_{dnak}, P_{hp0169}, and P_{264}) from the genomes of WSH-003 and ATCC 621H. Compared with these four reported strong promoters, the promoter P_{2703} has the highest strength, which was about 2.8-fold higher than that of P_{264} and about 3.1-fold higher than that of P_{dnak} (**Figure 3**). The results showed that P_{2703} was the strongest promoter discovered in *G. oxydans* at present. Interestingly, two strong shuttle promoters (P_{3022} and

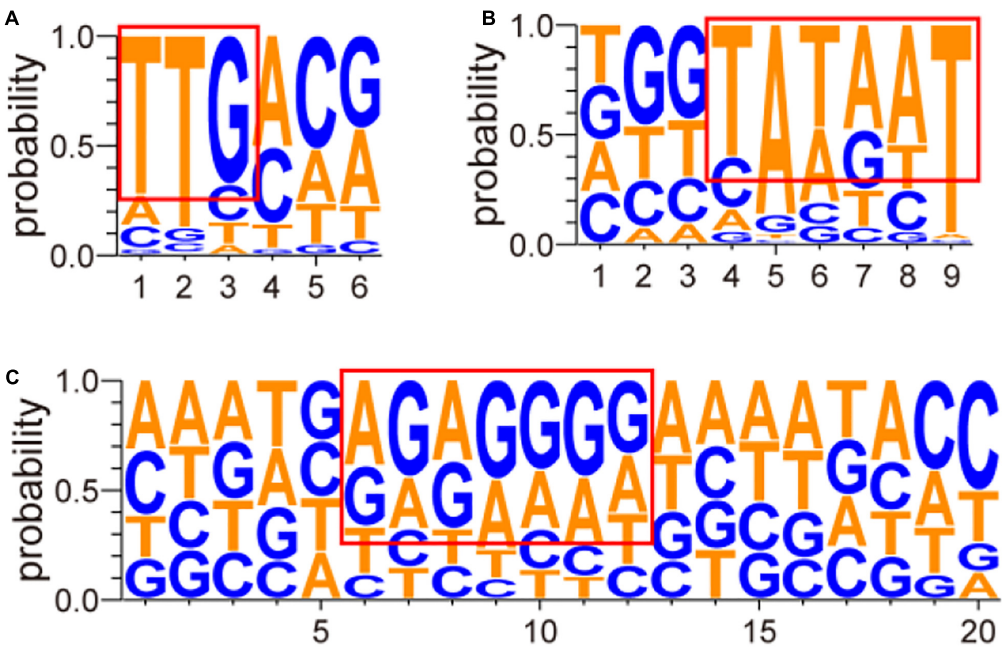


FIGURE 4 | Sequence analysis of conserved nucleotides of screened promoters. The sequence analysis of conserved nucleotides of screened promoters was conducted using Softberry (Salamov and Solovyevand, 2011) and Neural Network Promoter Prediction (Reese, 2001). The abscissa axis shows the positions of nucleotides in different promoters. The 20 nucleotides before the initiation codon "ATG" or "GTG" were selected for the prediction of RBS. The ordinate axis shows the probability of each nucleotide. In the figure, **(A)** stands for the prediction result of the conserved region near position −35, **(B)** stands for the prediction result of the conserved region near position −10, and **(C)** stands for the prediction result of the conserved RBS.

P_{0943}) in *E. coli* and *G. oxydans* were discovered in this study. The *E. coli* JM109 containing the above plasmids showed visible red fluorescence (**Supplementary Figure 2**). When detecting the strength of these promoters in *G. oxydans* WSH-003, strong red fluorescence could also be observed (**Figure 2**). The shuttle promoters were often applied to the construction of shuttle vectors to express resistance genes or used to build a broad host expression system. To the best of our knowledge, such strong shuttle promoters have rarely been reported in *E. coli* and *G. oxydans*, although many shuttle promoters have been reported in other strains (Yang et al., 2018; Khan et al., 2020).

Analysis of Screened Promoters

The structures of these screened promoters were also analyzed in this study. The promoter sequences are listed in **Supplementary Table 3**. Analysis of the above promoter region and transcription start site was performed by Softberry[1] (Salamov and Solovyevand, 2011) and Neural Network Promoter Prediction[2] (Reese, 2001). The analyzed results are listed in **Supplementary Table 4** and shown in **Figure 4** as mapped by the website[3] (Schneider and Stephens, 1990; Crooks et al., 2004). As shown in **Figure 4**, a "TTGnnn" region, with a highly conserved "TTG," near position −35 and a "TATAAT" region near position −10 were found in the screened promoters. High frequencies of "A" or "G" were also observed at transcription initiation sites. At last, a region enriched in "A" and "G" was discovered at about eight nucleotides before the initiation codon "ATG" or "GTG" (**Figure 4C**), and many "AGGAg" regions were observed when strong promoters were analyzed (**Supplementary Table 4**).

Application of Promoters for Improving the Production of 2-KLG

Strain *G. oxydans* has potential in the biosynthesis of 2-KLG from D-sorbitol (Wang et al., 2018). It has been reported that SDH was an essential dehydrogenase in the conversion of L-sorbose to form 2-KLG in *G. oxydans* (Hoshino et al., 1990; Saito et al., 1998). In our previous study, it was also found that the rate-limiting step of the fermentation is the enzyme activity of SDH (Chen et al., 2019). In this study, a group of gradient promoters was selected to overexpress SDH in *G. oxydans* WSH-003. As shown in **Figure 5**, when weak promoters were used, the strains produced almost no 2-KLG; when medium-strength promoters were used, the strains could only synthesize 2-KLG with yields lower than 2.0 g/L; when strong promoters were applied, the strains could synthesize 2-KLG with yields of up to 3.7 g/L. The highest conversion yields achieved about 7% in mole number using strong promoters, almost twofold higher than that of those using medium-strength promoters. In addition, all the strains achieved similar biomass (**Figure 5**). Taken together, it can be concluded that the titer of 2-KLG increased with the enhancement of promoters or, in other words, 2-KLG production was positively related to the expression of SDH. These results demonstrated that the activity of SDH was indeed a rate-limiting step in the fermentation of 2-KLG.

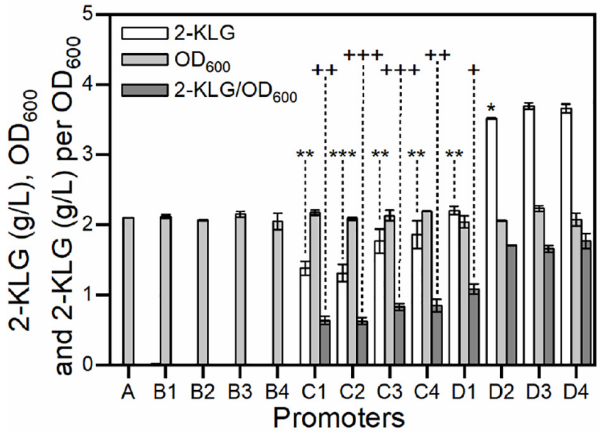


FIGURE 5 | Effect of over-expressing SDH by gradient promoters on the growth of *G. oxydans* and the yield of 2-KLG. The abscissa axis shows different promoters controlling the expression of SDH in *G. oxydans*. **(A)** negative control; **(B1–B4)** weak promoters P_{1142}, P_{0991}, P_{1295}, and P_{0327}; **(C1–C4)** medium-strength promoters P_{0647}, P_{0205}, P_{0804}, and P_{1653}; **(D1–D4)** strong promoters P_{3022}, P_{2038}, P_{2057}, and P_{2703}. The ordinate axis shows the titer of 2-KLG (g/L), OD_{600} values and the titer of 2-KLG per OD_{600}. $^{*}P < 0.05$ 2-KLG titer compared with D3 by two-tailed *t* test; $^{**}P < 0.01$ 2-KLG titer compared with D3 by two-tailed *t* test; $^{***}P < 0.001$ 2-KLG titer compared with D3 by two-tailed *t* test; D3 and D4 are not statistically significant in 2-KLG titer with $p = 0.0667$. $^{+}P < 0.05$ 2-KLG titer per OD_{600} compared with D3 by two-tailed *t* test; $^{++}P < 0.01$ 2-KLG titer per OD_{600} compared with D3 by two-tailed *t* test; $^{+++}P < 0.001$ 2-KLG titer per OD_{600} compared with D3 by two-tailed *t* test; D2 and D3 are not statistically significant in 2-KLG titer per OD_{600} with $p = 0.227$; D4 and D3 are not statistically significant in 2-KLG titer per OD_{600} with $p = 0.083$.

DISCUSSION

Gluconobacter oxydans is an excellent host to produce 2-KLG, which is an essential precursor of vitamin C (Saito et al., 1998; Gao et al., 2014; Wang et al., 2018). With the help of gradient promoters screened, a series of 2-KLG-producing strains have been obtained. These 2-KLG-producing strains showed the highest titer when the strongest promoters were used, while almost no production when weak promoters were employed. It was consistent with the study of Saito et al. (1997) that the productivity of 2-KLG could be improved by optimizing promoters. The degradation of 2-KLG by a class of aldo-keto reductases was reported in some *G. oxydans* strains (Sugisawa et al., 1990; Saito et al., 1997) and *Aspergillus niger* (Kuivanen et al., 2017). That may be the reason that *G. oxydans* WSH-003 could not accumulate 2-KLG when the expression level of sorbose dehydrogenase was low. These results may explain why many *G. oxydans* strains possess the entire set of 2-KLG biosynthesis genes, but only a few strains produce 2-KLG naturally (Wang et al., 2018).

In this study, a group of promoters with different strengths was obtained based on RNA-Seq data of whole transcripts. The

[1]http://www.softberry.com/

[2]https://www.fruitfly.org/seq_tools/promoter.html

[3]http://weblogo.threeplusone.com/

relative strength of these promoters covered a range of about 28 times, from 400 to 11,000, while reported promoters covered about 10 times, from 400 to 4000. Among them, the activity of the newly discovered strongest promoter P_{2703} was approximately threefold that of the reported strong promoters P_{264} and P_{dnak}. Besides, two promoters P_{0943} and P_{3022} showed high activity in both *E. coli* and *G. oxydans*, revealing great potential in the construction of a shuttle expression system. The promoter region and the transcription start site of the screened promoters were also analyzed. The two strong shuttle promoters P_{3022} and P_{0943} had an excellent linear discriminant function (LDF) value, which may be the reason why these two promoters had high activities in both *E. coli* and *G. oxydans*.

In recent years, researchers have conducted many studies on the promoters of prokaryotes, especially model microorganisms such as *E. coli* (Schuller et al., 2020), *Bacillus subtilis* (Castillo-Hair et al., 2019), and *Corynebacterium glutamicum* (Dostalova et al., 2019). Based on these studies, researchers could have a detailed knowledge of these promoters' structures and transcription factors. In this study, we found a "TTGnnn" region, with a highly conserved "TTG," nearing position −35 and a "TATAAT" region nearing position −10, which was in accordance with the results in many other bacteria like *E. coli*. However, the "TATnnT" region nearing position −10 was not observed in the strong promoters of *G. oxydans* (**Supplementary Table 4**). This result is in agreement with the study of Kranz, and the main reason may be that the promoter motif is recognized by alternative sigma factors except σ^{70} (Kranz et al., 2018). The higher frequency of "A" or "G" at transcription initiation sites supported the theory that purine nucleotides are related to the increased transcription initiation rates (Mendoza-Vargas et al., 2009). Consistent with a prediction by Kranz et al. (2018) the conserved RBS motif "AGGAg" was also found in the strong promoters of *G. oxydans*.

With the development of synthetic biology, many methods can be applied to improve the strength of promoters, for example, randomization of the non-conserved region of the promoters (Siegl et al., 2013), error-prone PCR (Swagatika et al., 2019), hybrid or cascade promoters (Zhou et al., 2017), the design of RBS by RBS Calculator (Salis, 2011), and the use of a promoter library based on machine learning (Zhao et al., 2020). On the other hand, many other promoters like shuttle promoters and inducible promoters are also crucial in protein engineering and metabolic engineering. A strong shuttle promoter, P_{bs}, for *B. subtilis*, *E. coli*, and *Saccharomyces cerevisiae* was constructed by Yang et al. (2018). Three broad-spectrum promoters (P_{bs1}, P_{bs2}, and P_{bs3}) with different strengths, were generated by random mutation and characterized. In a recent study, a newly tunable L-arabinose-inducible P_{BAD} promoter was discovered to be useful in *G. oxydans* 621H, and the activity of this promoter was affected by the pH of the medium (Fricke et al., 2020). In summary, the identification of gradient promoters in this study expanded the toolbox of available promoters, and these promoters revealed promising prospects in metabolic engineering of *G. oxydans* for high-value products. With further research, more serviceable promoters of *G. oxydans* are expected to be discovered and constructed.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in online repositories: https://www.ncbi.nlm.nih.gov/biosample/?term=SAMN18147321.

AUTHOR CONTRIBUTIONS

YC, LL, and SY performed the experiments and data analysis. YC and JZ wrote the manuscript and conceived the study. JL, JZ, and JC coordinated the project. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2019YFA0904900), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (32021005), and the National Science Fund for Excellent Young Scholars (21822806).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.673844/full#supplementary-material

REFERENCES

Battling, S., Wohlers, K., Igwe, C., Kranz, A., Pesch, M., Wirtz, A., et al. (2020). Novel plasmid-free *Gluconobacter oxydans* strains for production of the natural sweetener 5-ketofructose. *Microb. Cell Fact.* 19, 1–15.

Blank, M., and Schweiger, P. (2018). Surface display for metabolic engineering of industrially important *acetic acid bacteria*. *Peerj* 6, e4626. doi: 10.7717/peerj.4626

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Castillo-Hair, S. M., Baerman, E. A., Fujita, M., Igoshin, O. A., and Tabor, J. J. (2019). Optogenetic control of *Bacillus subtilis* gene expression. *Nat. Commun.* 10, 1–11. doi: 10.1128/mmbr.57.1.1-33.1993

Chen, Y., Liu, L., Shan, X., Du, G., Zhou, J., and Chen, J. (2019). High-throughput screening of a 2-keto-L-gulonic acid-producing *Gluconobacter oxydans* strain based on related dehydrogenases. *Front. Bioeng. Biotechnol.* 7:385.

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

De La Morena, S., Wojtusik, M., Santos, V. E., and Garcia-Ochoa, F. (2020). Kinetic modeling of dihydroxyacetone production from glycerol by *Gluconobacter oxydans* ATCC 621 resting cells: Effect of fluid dynamics conditions. *Catalysts* 10, 101. doi: 10.3390/catal10010101

De Muynck, C., Pereira, C., Naessens, M., Parmentier, S., Soetaert, W., and Vandamme, E. (2007). The genus *Gluconobacter oxydans*: comprehensive overview of biochemistry and biotechnological applications. *Crit. Rev. Biotechnol.* 27, 147. doi: 10.1080/07388550701503584

De Wulf, P., Soetaert, W., and Vandamme, E. J. (2000). Optimized synthesis of L-sorbose by C5-dehydrogenation of D-sorbitol with *Gluconobacter oxydans*. *Biotechnol. Bioeng.* 69, 339–343. doi: 10.1002/1097-0290(20000805)69:3<339::aid-bit12>3.0.co;2-e

Deppenmeier, U., Hoffmeister, M., and Prust, C. (2002). Biochemistry and biotechnological applications of *Gluconobacter* strains. *Appl. Microbiol. Biotechnol.* 60, 233. doi: 10.1007/s00253-002-1114-5

Dostalova, H., Busche, T., Holátko, J., Rucka, L., Štěpánek, V., Barvik, I., et al. (2019). Overlap of promoter recognition specificity of stress response sigma factors SigD and SigH in *Corynebacterium glutamicum* ATCC 13032. *Front. Microbiol.* 9:3287.

Fang, J., Wan, H., Zeng, W., Li, J., Chen, J., and Zhou, J. (2020). Transcriptome analysis of *Gluconobacter oxydans* WSH-003 exposed to elevated 2-keto-L-gulonic acid reveals the responses to osmotic and oxidative stress. *Appl. Biochem. Biotechnol.* 193, 128–141. doi: 10.1007/s12010-020-03405-8

Fricke, P. M., Link, T., Gätgens, J., Sonntag, C., Otto, M., Bott, M., et al. (2020). A tunable L-arabinose-inducible expression plasmid for the acetic acid bacterium *Gluconobacter oxydans*. *Appl. Microbiol. Biotechnol.* 104, 9267–9282. doi: 10.1007/s00253-020-10905-4

Gao, H., Li, J., Sivakumar, D., Kim, T.-S., Patel, S. K., Kalia, V. C., et al. (2019). NADH oxidase from *Lactobacillus reuteri*: a versatile enzyme for oxidized cofactor regeneration. *Int. J. Biol. Macromol.* 123, 629–636. doi: 10.1016/j.ijbiomac.2018.11.096

Gao, L., Hu, Y., Liu, J., Du, G., Zhou, J., and Chen, J. (2014). Stepwise metabolic engineering of *Gluconobacter oxydans* WSH-003 for the direct production of 2-keto-L-gulonic acid from D-sorbitol. *Metab. Eng.* 24, 30–37. doi: 10.1016/j.ymben.2014.04.003

Gao, L., Zhou, J., Liu, J., Du, G., and Chen, J. (2012). Draft genome sequence of *Gluconobacter oxydans* WSH-003, a strain that is extremely tolerant of saccharides and alditols. *J. Bacteriol.* 194, 4455–4456. doi: 10.1128/jb.00837-12

Hahn, T., Torkler, S., Van Der Bolt, R., Gammel, N., Hesse, M., Möller, A., et al. (2020). Determining different impact factors on the xylonic acid production using *Gluconobacter oxydans* DSM 2343. *Process Biochem.* 94, 172–179. doi: 10.1016/j.procbio.2020.04.011

Hoffmann, J. J., Hövels, M., Kosciow, K., and Deppenmeier, U. (2020). Synthesis of the alternative sweetener 5-ketofructose from sucrose by fructose dehydrogenase and invertase producing *Gluconobacter* strains. *J. Biotechnol.* 307, 164–174. doi: 10.1016/j.jbiotec.2019.11.001

Hoshino, T., Sugisawa, T., Tazoe, M., Shinjoh, M., and Fujiwara, A. (1990). Metabolic pathway for 2-keto-L-gulonic acid formation in *Gluconobacter melanogenus* IFO 3293. *Agric Biol Chem.* 54, 1211–1218. doi: 10.1271/bbb1961.54.1211

Hövels, M., Kosciow, K., Kniewel, J., Jakob, F., and Deppenmeier, U. (2020). High yield production of levan-type fructans by *Gluconobacter japonicus* LMG 1417. *Int. J. Biol. Macromol.* 164, 295–303. doi: 10.1016/j.ijbiomac.2020.07.105

Hu, Y., Wan, H., Li, J., and Zhou, J. (2015). Enhanced production of L-sorbose in an industrial *Gluconobacter oxydans* strain by identification of a strong promoter based on proteomics analysis. *J. Ind. Microbiol. Biotechnol.* 42, 1039–1047. doi: 10.1007/s10295-015-1624-7

Kallnik, V., Meyer, M., Deppenmeier, U., and Schweiger, P. (2010). Construction of expression vectors for protein production in *Gluconobacter oxydans*. *J. Biotechnol.* 150, 460–465. doi: 10.1016/j.jbiotec.2010.10.069

Khan, N., Yeung, E., Farris, Y., Fansler, S. J., and Bernstein, H. C. (2020). A broad-host-range event detector: expanding and quantifying performance between *Escherichia coli* and *Pseudomonas species*. *Synth. Biol.* 5, ysaa002.

Kim, T.-S., Patel, S. K., Selvaraj, C., Jung, W.-S., Pan, C.-H., Kang, Y. C., et al. (2016). A highly efficient sorbitol dehydrogenase from *Gluconobacter oxydans* G624 and improvement of its stability through immobilization. *Sci. Rep.* 6, 1–11.

Kranz, A., Busche, T., Vogel, A., Usadel, B., Kalinowski, J., Bott, M., et al. (2018). RNAseq analysis of α-proteobacterium *Gluconobacter oxydans* 621H. *BMC genomics* 19:24.

Kuivanen, J., Arvas, M., and Richard, P. (2017). Clustered genes encoding 2-keto-L-gulonate reductase and L-idonate 5-dehydrogenase in the novel fungal D-glucuronic acid pathway. *Front. Microbiol.* 8:225.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Meth.* 9, 357. doi: 10.1038/nmeth.1923

Lee, S. H., Kim, M.-S., Jung, H. C., Lee, J., Lee, J.-H., Lee, H. S., et al. (2015). Screening of a novel strong promoter by RNA sequencing and its application to H_2 production in a hyperthermophilic archaeon. *Appl. Microbiol. Biotechnol.* 99, 4085–4092. doi: 10.1007/s00253-015-6444-1

Li, K., Mao, X., Liu, L., Lin, J., Sun, M., Wei, D., et al. (2016). Overexpression of membrane-bound gluconate-2-dehydrogenase to enhance the production of 2-keto-D-gluconic acid by *Gluconobacter oxydans*. *Microb. Cell Fact.* 15, 121.

Li, N., Zeng, W., Xu, S., and Zhou, J. (2020). Obtaining a series of native gradient promoter-5'-UTR sequences in *Corynebacterium glutamicum* ATCC 13032. *Microb. Cell Fact.* 19, 1–11.

Liao, Y., Huang, L., Wang, B., Zhou, F., and Pan, L. (2015). The global transcriptional landscape of *Bacillus amyloliquefaciens* XH7 and high-throughput screening of strong promoters based on RNA-seq data. *Gene* 571, 252–262. doi: 10.1016/j.gene.2015.06.066

Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., et al. (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* 4:e7526. doi: 10.1371/journal.pone.0007526

Merfort, M., Herrmann, U., Bringer-Meyer, S., and Sahm, H. (2006). High-yield 5-keto-D-gluconic acid formation is mediated by soluble and membrane-bound gluconate-5-dehydrogenases of *Gluconobacter oxydans*. *Appl. Microbiol. Biotechnol.* 73, 443–451. doi: 10.1007/s00253-006-0467-6

Mientus, M., Kostner, D., Peters, B., Liebl, W., and Ehrenreich, A. (2017). Characterization of membrane-bound dehydrogenases of *Gluconobacter oxydans* 621H using a new system for their functional expression. *Appl. Microbiol. Biotechnol.* 101, 3189–3200. doi: 10.1007/s00253-016-8069-4

Nishikura-Imamura, S., Matsutani, M., Insomphun, C., Vangnai, A. S., Toyama, H., Yakushi, T., et al. (2014). Overexpression of a type II 3-dehydroquinate dehydratase enhances the biotransformation of quinate to 3-dehydroshikimate in *Gluconobacter oxydans*. *Appl. Microbiol. Biotechnol.* 98, 2955–2963. doi: 10.1007/s00253-013-5439-z

Poulsen, L. D., and Vinther, J. (2018). RNA-Seq for bacterial gene expression. *Curr. Protoc. Nucleic Acid Chem.* 73, e55.

Prust, C., Hoffmeister, M., Liesegang, H., Wiezer, A., Fricke, W. F., Ehrenreich, A., et al. (2005). Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. *Nat. Biotechnol.* 23, 195–200. doi: 10.1038/nbt1062

Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput. Chem.* 26, 51–56. doi: 10.1016/s0097-8485(01)00099-7

Saito, Y., Ishii, Y., Hayashi, H., Imao, Y., Akashi, T., Yoshikawa, K., et al. (1997). Cloning of genes coding for L-sorbose and L-sorbosone dehydrogenases from *Gluconobacter oxydans* and microbial production of 2-keto-L-gulonate, a precursor of L-ascorbic acid, in a recombinant *G. oxydans* strain. *Appl. Environ. Microbiol.* 63, 454–460. doi: 10.1128/aem.63.2.454-460.1997

Saito, Y., Ishii, Y., Hayashi, H., Yoshikawa, K., Noguchi, Y., Yoshida, S., et al. (1998). Direct fermentation of 2-keto-L-gulonic acid in recombinant *Gluconobacter oxydans*. *Biotechnol. Bioeng.* 58, 309–315. doi: 10.1002/(sici)1097-0290(19980420)58:2/3<309::aid-bit30>3.0.co;2-4

Salamov, V. S. A., and Solovyevand, A. (2011). "Automatic annotation of microbial genomes and metagenomic sequences," in *Metagenomics and its applications in agriculture, biomedicine and environmental studies*, ed. R. W. Li (New York: Nova Science Publishers), 61–78.

Salis, H. M. (2011). The ribosome binding site calculator. *Methods Enzymol.* 498, 19–42. doi: 10.1016/b978-0-12-385120-8.00002-4

Schedel, M. (2000). Regioselective oxidation of aminosorbitol with *Gluconobacter oxydans*, key reaction in the industrial 1-deoxynojirimycin synthesis. *Biotechnology* 8, 295–308. doi: 10.1002/9783527620913.ch7

Schneider, T. D., and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18, 6097–6100. doi: 10.1093/nar/18.20.6097

Schuller, A., Cserjan-Puschmann, M., Tauer, C., Jarmer, J., Wagenknecht, M., Reinisch, D., et al. (2020). *Escherichia coli* σ^{70} promoters allow expression rate control at the cellular level in genome-integrated expression systems. *Microb. Cell. Fact.* 19, 1–11.

Selvaraj, C., Krishnasamy, G., Jagtap, S. S., Patel, S. K., Dhiman, S. S., Kim, T.-S., et al. (2016). Structural insights into the binding mode of D-sorbitol with

sorbitol dehydrogenase using QM-polarized ligand docking and molecular dynamics simulations. *Biochem. Eng. J.* 114, 244–256. doi: 10.1016/j.bej.2016.07.008

Shan, X., Liu, L., Zeng, W., Chen, J., and Zhou, J. (2020). High throughput screening platform for a FAD-dependent L-sorbose dehydrogenase. *Front. Bioeng. Biotechnol.* 8:194.

Shen, Y., Zhou, X., and Xu, Y. (2020). Enhancement of *Gluconobacter oxydans* resistance to lignocellulosic-derived inhibitors in xylonic acid production by overexpressing thioredoxin. *Appl. Biochem. Biotechnol.* 191, 1072–1083. doi: 10.1007/s12010-020-03253-6

Siegl, T., Tokovenko, B., Myronovskyi, M., and Luzhetskyy, A. (2013). Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab. Eng.* 19, 98–106. doi: 10.1016/j.ymben.2013.07.006

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat. Rev. Genet.* 20, 631–656. doi: 10.1038/s41576-019-0150-2

Sugisawa, T., Hoshino, T., Masuda, S., Nomura, S., Setoguchi, Y., Tazoe, M., et al. (1990). Microbial production of 2-keto-L-gulonic acid from L-sorbose and D-sorbitol by *Gluconobacter melanogenus*. *Agric. Biol. Chem.* 54, 1201–1209. doi: 10.1271/bbb1961.54.1201

Swagatika, P., Nikhil, K., Barkha, R., Pashupathi, M., Parthasarathi, B., and Ajay, K. (2019). Effect of mutation resulted from error prone PCR on the strength of promoter activity. *J. Exp. Zool. India* 22, 981–986.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Wan, H., Xia, Y., Li, J., Kang, Z., and Zhou, J. (2017). Identification of transporter proteins for PQQ-secretion pathways by transcriptomics and proteomics analysis in *Gluconobacter oxydans* WSH-003. *Front. Chem. Sci. Eng.* 11:72–88. doi: 10.1007/s11705-016-1580-4

Wang, P., Zeng, W., Xu, S., Du, G., Zhou, J., and Chen, J. (2018). Current challenges facing one-step production of L-ascorbic acid. *Biotechnol. Adv.* 36, 1882–1899. doi: 10.1016/j.biotechadv.2018.07.006

Yang, S., Liu, Q., Zhang, Y., Du, G., Chen, J., and Kang, Z. (2018). Construction and characterization of broad-spectrum promoters for synthetic biology. *ACS Synth. Biol.* 7, 287–291. doi: 10.1021/acssynbio.7b00258

Yuan, J. F., Wu, M. B., Lin, J. P., and Yang, L. R. (2016). Combinatorial metabolic engineering of industrial *Gluconobacter oxydans* DSM2343 for boosting 5-keto-D-gluconic acid accumulation. *BMC Biotechnol.* 16:42.

Zhang, L., Lin, J., Ma, Y., Wei, D., and Sun, M. (2010). Construction of a novel shuttle vector for use in *Gluconobacter oxydans*. *Mol. Biotechnol.* 46, 227–233. doi: 10.1007/s12033-010-9293-2

Zhao, M., Zhou, S., Wu, L., and Deng, Y. (2020). Machine learning-based promoter strength prediction derived from a fine-tuned synthetic promoter library in *Escherichia coli*. *bioRxiv*

Zhou, S., Ding, R., Chen, J., Du, G., Li, H., and Zhou, J. (2017). Obtaining a panel of cascade promoter-5′-UTR complexes in *Escherichia coli*. *ACS Synth. Biol.* 6, 1065–1075. doi: 10.1021/acssynbio.7b00006

Zhou, X., Shen, Y., Xu, Y., and Balan, V. (2020). Directing cell catalysis of glucose to 2-keto-D-gluconic acid using *Gluconobacter oxydans* NL71. *Process Biochem.* 94, 365–369. doi: 10.1016/j.procbio.2020.04.038

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 17 June 2021
doi: 10.3389/fbioe.2021.661562

Microbial Cell Factories for Green Production of Vitamins

Yanyan Wang[1,2,3†], Linxia Liu[2,3,4†], Zhaoxia Jin[1] and Dawei Zhang[2,3,4,5*]*

[1] School of Biological Engineering, Dalian Polytechnic University, Dalian, China, [2] Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, China, [3] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, [4] National Technology Innovation Center of Synthetic Biology, Tianjin, China, [5] University of Chinese Academy of Sciences, Beijing, China

Vitamins are a group of essential nutrients that are necessary to maintain normal metabolic activities and optimal health. There are wide applications of different vitamins in food, cosmetics, feed, medicine, and other areas. The increase in the global demand for vitamins has inspired great interest in novel production strategies. Chemical synthesis methods often require high temperatures or pressurized reactors and use non-renewable chemicals or toxic solvents that cause product safety concerns, pollution, and hazardous waste. Microbial cell factories for the production of vitamins are green and sustainable from both environmental and economic standpoints. In this review, we summarized the vitamins which can potentially be produced using microbial cell factories or are already being produced in commercial fermentation processes. They include water-soluble vitamins (vitamin B complex and vitamin C) as well as fat-soluble vitamins (vitamin A/D/E and vitamin K). Furthermore, metabolic engineering is discussed to provide a reference for the construction of microbial cell factories. We also highlight the current state and problems encountered in the fermentative production of vitamins.

Keywords: vitamins, metabolic engineering, microbial cell factory, chemical synthesis, biosynthesis

OPEN ACCESS

Edited by:
Dipesh Dhakal, University of Florida, United States

Reviewed by:
Ana Margarida Goncalves Carvalho Dias, New University of Lisbon, Portugal
Angel León-Buitimea, Universidad Autonoma de Nuevo Leon, Mexico

***Correspondence:**
Zhaoxia Jin
jinzx2018@163.com
Dawei Zhang
zhang_dw@tib.cas.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *31 January 2021*
Accepted: *12 May 2021*
Published: *17 June 2021*

Citation:
Wang Y, Liu L, Jin Z and Zhang D (2021) Microbial Cell Factories for Green Production of Vitamins. Front. Bioeng. Biotechnol. 9:661562. doi: 10.3389/fbioe.2021.661562

INTRODUCTION

Vitamins are essential for proper growth and health of animals, that cannot produce vitamins by themselves or that synthesize insufficient amount to cover all their needs (Capone and Sentongo, 2019; Suter, 2020). The methods of producing vitamins are based either on chemical synthesis or fermentative production (Yuan et al., 2020).

There are at least 30 kinds of different compounds considered "vitamins," more than 20 vitamins of which are known to be necessary for biological health. Vitamins are either water-soluble or fat-soluble. As the name suggests, a water-soluble vitamin dissolves in water easily and insoluble in organic solvents. After absorption, the body stores very little of such proteins, and most are excreted with urine (Berdanier and Adkins, 2019). Fat-soluble vitamins are dissolved in fats but not in water, and which are stored in the liver or fatty tissues for future use. While vitamins are essential nutrients for all living things, many plants and microorganisms can synthesize them naturally by themselves. By contrast, humans and other animals need to acquire sufficient vitamins with their diet or through supplements to maintain optimal health (Blake and Konings, 2019).

Traditionally, vitamin production strains have been improved through mutagenesis and metabolic engineering, which can be conducted either through chemical or biological means

(Vandamme and Revuelta, 2016b). The main chemical strategies include chemical mutagenesis, application of N^+ ion beam, ultraviolet radiation or laser mutagenesis. The biological methods mainly include the construction and mutagenesis of the starting strain, genetic modification, synthetic biotechnology, optimization of media and culture conditions, construction of biofilm reactors, etc. (Nie et al., 2013; Song et al., 2014). A series of biotechnological methods are used to transform the metabolic network of cells to construct a programmable "chassis" and "programmable" whole, which can be used to develop an effective assembly strategy, test the adaptability of external components and modules after loading, forming a fine-tuned and customized biological application system. To drive the iterative evolution of other industrial strains, and effectively promote the transformation and renewal of high vitamin producing strains. Chemical methods are usually expensive, environment-unfriendly, waste-prone, and the costly waste disposal. However, the microbial fermentation method has attracted much attention due to low cost, low energy consumption and easy waste recycling. At present, the fermentation method has been recognized by researchers, and it is more environment-friendly and safe than chemical methods. As the fermentation technology matures, this approach is increasingly being used in industry to increase the production of different vitamins. For example, fermentation processes for the production of vitamin B_2 (VB_2), vitamin B_{12} (VB_{12}), vitamin C, and vitamin K2 have all been industrialized successfully.

Acevedo-Rocha et al. (2019) reviewed the fermentation of B vitamins from the aspect of sustainability. In this review, we mainly discuss vitamins that can be produced by green fermentation processes. It covers water-soluble vitamins, including vitamin C and vitamin B complex (thiamine, riboflavin, niacin, pantothenic acid, pyridoxine, biotin, folate, and cobalamin) as well as the fat-soluble vitamin E and vitamin K. Here, we discussed the producing microorganisms, advanced biological methods and metabolic bottlenecks of different vitamins.

WATER-SOLUBLE VITAMINS

B Vitamins

The global demand for B vitamins is growing due to wide applications in food, pharmaceuticals, feed, and other fields. Although most vitamins are manufactured by chemical synthesis, successful industrial bioprocesses have been established for the production of VB_2 and VB_{12}. The underlying extraordinary achievement in metabolic engineering is discussed in this article.

Vitamin B_1

Vitamin B_1, which is also known as thiamine, was the first B vitamin to be identified. Thiamine pyrophosphate (TPP), the active form of thiamine, can inhibit the activity of cholinesterase, reduce skin inflammation, prevent seborrheic dermatitis, or eczema, and improve skin health. Thiamine biosynthesis results from the coupling of the pyrimidine and the thiazole moieties to form thiamine phosphate (Dorrestein et al., 2004; Jurgenson et al., 2009; Cea et al., 2020). *Escherichia coli, Salmonella typhimurium,* and *Bacillus subtilis* are the most thoroughly studied thiamine production organisms (Begley et al., 1999).

In chassis cell *S. typhimurium*, the thiamine pyrimidine moiety can be produced through *de novo* purine biosynthesis or independently of the *purF* gene through the alternative pyrimidine biosynthesis (APB) pathway (Downs and Roth, 1991; Downs, 1992). According to the phenotypic characteristics of the *abpA* mutant, follow-up studies concluded that the functional APB pathway is essential for thiamine synthesis when *S. typhimurium* grows in the presence of exogenous purines (Downs and Petersen, 1994). Research has shown that overexpression of *thiA*, *nmtA,* and *thiP* in *Aspergillus oryzae* can increase the vitamin B_1 yield fourfold compared to the wild-type (Tokui et al., 2011). Based on the riboswitch mechanism, mutations in the genes of thiamine pyrophosphate kinase activity (*thiN*) and thiamine-related transport proteins (YkoD and YuaJ) were introduced in *B. subtilis* TH95. It was recently reported that thiamine biosynthesis is strictly regulated by TPP riboswitches in bacteria/eukaryotes and transcriptional repressors in archaea (Hwang et al., 2017). *E. coli* has emerged as the preferred cell factory for TPP production after a riboswitch-based biosensors enabled the discovery of thiamine transporters, combined with overexpression of the native *thiFSGHCE* and *thiD* genes, which are closely related to Fe-S metabolism (**Figure 1A** and **Table 1**; Cardinale et al., 2017).

However, *thiC/thiH* in the thiamine biosynthetic pathway is involved in Fe–S metabolism and is inhibited by *S*-adenosylmethionine (SAM) metabolites, and the catalytic activity of ThiC enzyme (**Figure 2**) is very low (k_{cat} = 0.002 s^{-1}) which is one of the main metabolic bottlenecks (Palmer and Downs, 2013). In addition, the cost of chemical production of thiamine is very low, and the production of engineered strains needs to be increased to be expected to be industrialized.

Vitamin B_2

Riboflavin is an important precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) (Balasubramaniam et al., 2019; Andreieva et al., 2020). Riboflavin insufficiency manifests as persistent anemia (Shi et al., 2014). The biosynthesis of riboflavin begins with guanosine triphosphate and ribose-5-phosphate, followed by six enzymatic steps (Fischer and Bacher, 2005). Burgess et al. (2004) found that overexpression of the *ribABCGH* genes can increase riboflavin production. Later, it was found that there were both nucleotide substitutions and deletions in the regulatory region of the *rib* operon. By deregulating the *rib* operon and purine pathway of *B. subtilis*, riboflavin production was greatly improved. The specific genetic engineering steps included overexpression of the *ribA* gene and deletion of the *purR* gene, after which maximum output of riboflavin reached more than 826.52 mg/L (**Figure 1B**; Shi et al., 2014). In *Candida famata* overexpression of *sef1* and *imh3* was combined with classic mutagenesis methods to construct the high riboflavin-producing strain AF-4. As a result, 1026 ± 50 mg/L of riboflavin can be produced during a fed-batch cultivation in a lab-scale fermenter. This research has

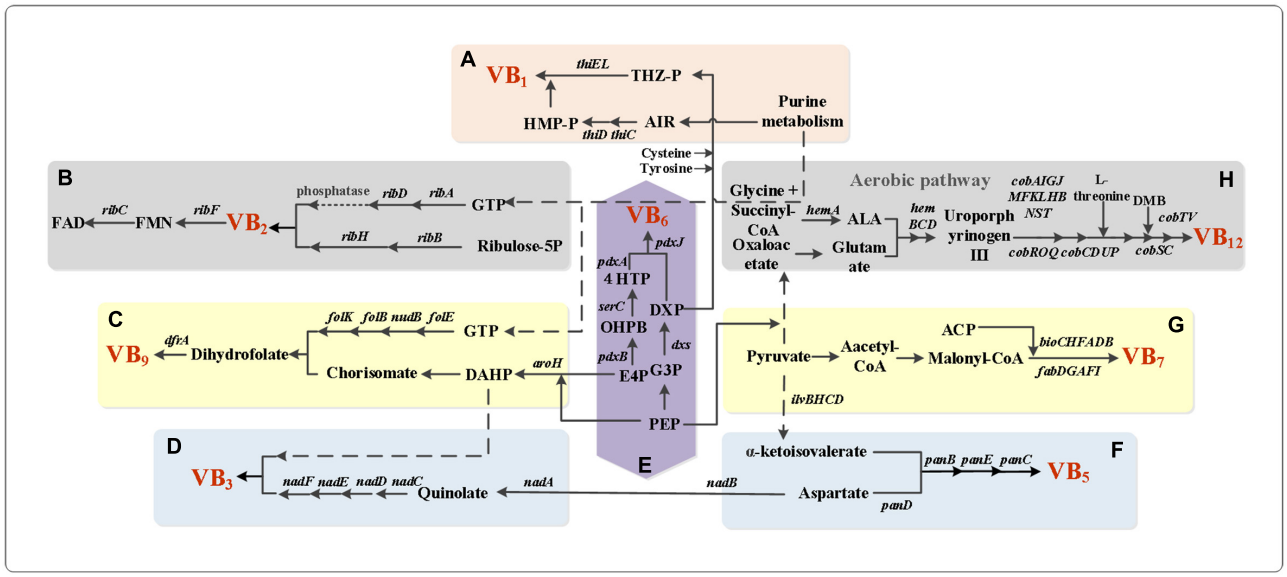


FIGURE 1 | Metabolic network pathway of B vitamins. **(A)** Biosynthesis pathway of thiamine in *E. coli*. ThiC/ThiD, phosphomethylpyrimidine synthase; ThiE, thiamine-phosphate pyrophosphorylase; ThiL, thiamine-monophosphate kinase. **(B)** Biosynthesis pathway of riboflavin in *B. subtilis*. RibA, GTP cyclohydrolase II; RibB, 3,4-dihydroxy 2-butanone 4-phosphate synthase; RibD, diaminohydroxyphosphoribosylaminopyrimidine deaminase; RibH, 6,7-dimethyl-8-ribityllumazine synthase. RibF, FMN adenylyltransferase; RibC, riboflavin synthase. **(C)** Vitamin B_9 biosynthesis pathway in *B. subtilis*. AroH, chorismate mutase; FolE, GTP cyclohydrolase IA; NudB, dihydroneopterin triphosphate diphosphatase; FolB, 7,8-dihydroneopterin aldolase/epimerase/oxygenase; FolK, 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase; DfrB, dihydrofolate reductase. **(D)** Vitamin B_3 biosynthesis pathway in *E. coli*. NadB, L-aspartate oxidase; NadA, quinolinate synthase; NadC, nicotinate-nucleotide pyrophosphorylase; NadD, nicotinate-nucleotide adenylyltransferase; NadE/NadF, NAD + synthase. **(E)** *De novo* biosynthesis pathway of vitamin B_6. PdxB, erythronate-4-phosphate dehydrogenase; SerC, phosphoserine aminotransferase; PdxA, 4-hydroxythreonine-4-phosphate dehydrogenase; PdxJ, pyridoxine 5-phosphate synthase; Dxs, 1-deoxy-D-xylulose-5-phosphate synthase. **(F)** Pathway for *de novo* synthesis of vitamin B_5. *ilvBHCD*, increased the transcription levels of the *ilv* genes; *panDBEC*, pantothenate biosynthetic genes. **(G)** Biosynthesis pathway of Vitamin B_7 in *E. coli*. BioC, malonyl-CoA *O*-methyltransferase; BioH, pimeloyl-[acyl-carrier protein] methyl ester esterase; BioF, 8-amino-7-oxononanoate synthase; BioA, 8-amino-7-oxononanoate aminotransferase; BioD, dethiobiotin synthetase; BioB, biotin synthase; FabD, *S*-malonyltransferase; FabG, 3-oxoacyl-(acyl-carrier protein) reductase; FabA, 3-hydroxyacyl-(acyl-carrier protein) dehydratase; FabF, 3-oxoacyl-(acyl-carrier-protein) synthase II; FabI, enoyl-(acyl-carrier protein) reductase I. **(H)** The aerobic pathway in the synthesis pathway of cobalamin. HemA, glutamyl-tRNA reductase; ALA, δ-aminolevulinate; HemB, porphobilinogen synthase; HemC, hydroxymethylbilane synthase; HemD, uroporphyrinogen-III synthase; CobA, uroporphyrin-III C-methyltransferase; CobI, precorrin-2 C(20)-methyltransferase; CobG, precorrin-3B synthase; CobJ, precorrin-3B C17-methyltransferase; CobF, precorrin-6A synthase; CobK, precorrin-6A/cobalt-precorrin-6A reductase; CobL, precorrin-6B methyltransferase; CobH, precorrin-8X/cobalt-precorrin-8 methylmutase; CobB, cobyrinic acid a,c-diamide synthase; cobNST, hydrogenobyrinic-acid-a,c-diamide:cobalt cobalt-ligase; CobR, cob(II)yrinic acid a,c-diamide reductase; CobO/CobP, corrinoid adenosyltransferase; CobQ, adenosylcobyric acid synthase; CobS/CobV, adenosylcobinamide-GDP ribazoletransferase; CobC, cobalamin biosynthesis protein; CobD, adenosylcobinamide-phosphate synthase; CobU/CobT, nicotinate-nucleotide–dimethylbenzimidazole phosphoribosyltransferase.

made a great contribution to industrial production of riboflavin (Dmytruk et al., 2011, 2014). The two most important industrial producers are *Ashbya gossypii* and *B. subtilis*.

In *A. gossypii*, malate synthase in the glyoxylate cycle is essential for riboflavin production. Deletion of the malate synthase gene (*ACR268C*) decreased riboflavin production 10-fold compared to the wild-type strain. Conversely, overexpression of the *ACR286C* gene significantly increased the yield of riboflavin by 70%. These results demonstrated that malate synthase is a new target for improving the production of riboflavin (Sugimoto et al., 2009). Abbas and Sibirny (2011) introduced the *icl* gene, overexpressed the *gly1*, *prs2,4,* and *prs3* genes, as well as knocking out the *vma4*, *shm2,* and *bas1* genes, resulting in riboflavin production of more than 20 g/L.

In *B. subtilis*, Schwechheimer et al. (2016) overexpressed riboflavin biosynthesis genes, decreased the activity of the flavin kinase RibCF, and improved the *de novo* purine synthesis and pentose supply, after which the riboflavin yield reached more than 26 g/L. At present, the bottleneck of riboflavin production is mainly due to the poor genetic stability of the engineered strain, and more by-products produced by fermentation, which restrict the high yield of riboflavin.

Vitamin B_3

Niacin is the precursor in the synthesis of the pyridine coenzymes NAD (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) (Chand and Savitri, 2016; Chauhan and Poddar, 2019; Tannous et al., 2020). It is found at relatively high concentrations in internal organs of animal, muscle tissues, and fruits. Currently, niacin is mainly used as a feed additive to increase the utilization of feed protein, or as a pharmaceutical intermediate in the synthesis of various drugs. So far, there is no systematic description of a commercial fermentation process of nicotinic acid (NA) or nicotinamide (NAM). Industrial production methods are mainly ammonia oxidation and electrolytic oxidation, but the former has

TABLE 1 | Water-soluble vitamins produced by biotechnological methods.

Vitamins	Strains	Biotechnological method	Medium and precursor	Yield	References
Vitamin B_1	*B. subtilis* TH95	Mutation of gene encoding thiamine pyrophosphate kinase activity (*thiN*) and thiamine-related transport protein (*ykoD* and *yuaJ*).	MM	1.27 mg/L	Schyns et al., 2005
	E. coli	TPP biosensor (plasmid pTPP_Bios); Overexpression of native *thiFSGH*; *thiC*; *thiE*; and *thiD*; Genetic-metabolic coupling.	MM	0.80 mg/L	Cardinale et al., 2017
	A. oryzae	Overexpression of *thiP*, *thiA*, and *nmtA*.	CD-Dex medium (5% dextrin)	4-fold > WT	Tokui et al., 2011
Vitamin B_2	*B. subtilis*	Decrease the activity of flavinase RibCF activity; Overexpression of riboflavin biosynthetic genes; improved the *de novo* purine synthesis and pentose supply.	MM	>26 g/L	Schwechheimer et al., 2016
	A. gossypii	Introduced the *icl* gene; Overexpression of *gly1*, *prs2,4,* and *prs3* genes; Knocked out *vma4*, *shm2,* and *bas1* genes.	YPD; Plant oil	>20 g/L	Abbas and Sibirny, 2011
	Candida famata	Conventional mutagenesis by overexpression of *sef1* and *imh3*.	YPD; Fluorophenilalanine	1026 ± 50 mg/L	Dmytruk et al., 2011
Vitamin B_3	*Yeast*	Knock out NR importer Nrt1in the NR-non-salvaging genotype *nrkI*, *urhI*, *pmpI* (strain PAB038).	2x YPD; Nicotinic acid	8 mg/L	Belenky et al., 2011
	E. coli	Expressing *R. hodochrous* nitrile hydratase.	LB medium; 2YT medium	508 g/L	Wang et al., 2017
Vitamin B_5	*C. glutamicum*	Deletion of *ilvA* gene and overexpression of *ilvBNCD* and *panBC* genes	MM	1000 mg/L	Leonardi and Jackowski, 2007
	B. subtilis	Overexpression of *ilvBHCD* and *panBCDE*; Overexpression of SerA and GlyA of the enzymes of the glycine cleavage cycle.	MM	82–86 g/L	Hohmann et al., 2016
Vitamin B_6	*E. coli*	Overexpression of native Epd, PdxJ, and Dxs enzymes	MM	78 mg/L	Hoshino et al., 2004
	S. meliloti IFO14782	Overexpression of *E. coli* Epd and native PdxJ enzyme.	MM	1.30 g/L	Hoshino et al., 2007
	B. subtilis	Overexpression of *E. coli* PdxA and *S. meliloti* PdxJ enzymes.	MM	65 mg/L	Commichau et al., 2015
Vitamin B_7	*Agrobacterium*/ *Rhizobium* HK4	Overexpression of a strong biotin operon from *E. coli*; Use of a powerful artificial tac promoter and introduct of a modified RBS in front of BioB.	MM; Betaine; Diaminononanoic acid	110 mg/L	Streit and Entcheva, 2003
	E. coli	Overexpression of native biotin operon from a high-copy number plasmid	MM; H-medium	11 mg/L	Ifuku et al., 1995
	B. subtilis	Overexpression of native biotin operon and selection on *S*-2-aminoethyl-L-cysteine.	MM	21 mg/L	Van Arsdell et al., 2005
Vitamin B_9	*A. gossypii* (ATCC 10895)	Overexpression of *FOL* genes and deletion of *AgMEY7*; Deletion of *AgADE12* and *AgRIB1* at the same time.	MA2 rich medium	7 mg/L	Serrano-Amatriain et al., 2016
Vitamin B_{12}	*S. meliloti* (MC5-2)	High throughput screening of mutants using riboswitch ARTP-irradiation was used to induce random mutations; Deletion of *cobI*; Overexpression of *hemE*.	MM; Cobalt chloride; DMBI	156 ± 4.20 mg/L	Cai et al., 2018
	P. denitrificans	Random mutagenesis and genetic engineering; Overexpression of *cobF-cobM* gene cluster and *cogA* and *cobE* genes; Optimize the best PH range; Optimize promoters.	Betaine; Beet molasses; Choline chloride	214.30 mg/L	Li et al., 2008
	E. coli	Heterologously expressed the *hemO*, *hemB*, *hemC*, and *hemD* genes etc.; Optimizing of fermentation conditions.	CM medium	0.67 mg/L	Fang et al., 2018
	Propionibacterium shermanii	Overexpression of biosynthetic genes.	MM; DMBI	206 mg/L	Sych et al., 2016
Vitamin C	*S. cerevisiae and Zygosaccharomyces bailii*	Overexpressing the endogenous D-arabinono-1,4-lactone oxidase and L-galactose dehydrogenase (overexpression of *lgdh* and *alo1*).	MM	100 mg/L	Sauer et al., 2004
	K. vulgare DSM 4025	Oxidation and lactonization.	L/D-sorbose; Glycerol; Baker's yeast	1.37 g/L	Sugisawa et al., 2005

(Continued)

TABLE 1 | Continued

Vitamins	Strains	Biotechnological method	Medium and precursor	Yield	References
	X. campestris 2286	Lactonation under oxidative stress; Direct synthesis of glucose (carbohydrate source) induced by free radicals (HOCL treatment).	MM; K_2HPO_4; Urea	20.40 g/L	Rao and Sureshkumar, 2000
	G. oxydans and *K. vulgare* and *B. endophyticus*	Cell–cell interaction; One step 2-KGA fermentation.	D-sorbitol	73.70 g/L (2-KGA)	Ma et al., 2019

MM, minimal media. CM, complete medium. YPD, yeast extract peptone dextrose. LB medium, Luria-Bertani medium. 2YT medium, Yeast extract and tryptone. CM medium, Cramer-Myers medium. DMBI, 5,6-dimethylbenzimidazole. YD, yeast extract and glucose. MYGP, malt extract, yeast extract, peptone, agar and dextrose.

high production costs and needs to be above 300°C during the reaction, and the latter has low costs of production, however, the efficiency of electrolysis is not high, which limits the industrial production of niacin (Chand and Savitri, 2016).

Recent reports describe the use of recombinant *E. coli* expressing *Rhodococcus rhodochrous* nitrile hydratase for vitamin B_3 production. At low cell density, nicotinamide was produced in fed-batch mode, and the product concentration reached 390 g/L. After high-density culture in 5 L bioreactor, the concentration of nicotinamide reached 508 g/L in 60 min (**Figure 1D**; Wang et al., 2017). Belenky et al. (2011) showed that the disruption of *nrt1* results in increased export of nicotinamide riboside (NR). Moreover, disruption of the niacin transporter Tna1 can also increase the output of niacin, revealing that cells regulate the intracellular NAD^+ metabolic process by balancing the transport of niacin, the precursor of NAD^+. On the basis of adding 5 mM niacin, yeast cells can produce 8 mg/L nicotinamide mononucleotide (Belenky et al., 2011).

Vitamin B_5

Vitamin B_5, also known as pantothenic acid, is composed of pantoic acid and β-alanine (β-Ala), which is a precursor of coenzyme A (Leonardi and Jackowski, 2007). It plays an important role in maintaining the health of skin and blood. Its general function is to participate in the production of energy in the body, but it can also control the fat metabolism, and is also an essential nutrient for the brain and nerves. There are chemical and microbial synthesis methods for the synthesis of pantothenic acid, whereby microbial methods can be used to directly synthesize optically pure D-pantothenic acid.

Sahm and Eggeling (1999) adopted a series of methods to increase the production of pantothenic acid, including the deletion of the *ilvA* gene and the overexpression of the *ilvBNCD* and *panBC* genes. The pantothenic acid production of the best strain reached 1000 mg/L (**Figure 1F**; Sahm and Eggeling, 1999). Huser et al. (2005) also used *Corynebacterium glutamicum* to produce pantothenic acid. They deleted the *ilvA* gene, inhibited the expression of the *ilvE* gene and overexpressed the *ilvBNCD* gene. The final titer of pantothenate reached 1.75 g/L (Huser et al., 2005). Studies have shown that the specific activity of pantothenic acid synthase PanC of *C. glutamicum* is 205.10 U/mg. Adding substrates (D-pantothenic acid and β-Ala) to *E. coli* containing the enzyme can be produced 97.10 U/mg within 32 h, the conversion rate of pantothenic acid was 99.10%. However, the reported work had production defects, which required the addition of exogenous substrate pantothenic acid, and the high market price of pantothenic acid seriously restricted the industrialization of this method. Another chassis organism that is commonly used to produce pantothenic acid is *B. subtilis*. Hohmann et al. (2016) clarified the highest production of pantothenic acid by overexpressing *ilvBHCD*, *panBCDE*, *serA*, and *glyA*, as well as the enzymes of the glycine cleavage cycle the purpose is to increase the number of precursors for pantothenic acid synthesis (**Figure 1F**). The maximal output of the best strain reached 82–86 g/L during a 48 h fed-batch fermentation, opening up a new chapter of vitamin production in the biological world (Hohmann et al., 2016).

Vitamin B_6

There are six forms of vitamin B_6, including pyridoxine (PN), pyridoxal (PL), and pyridoxamine (PM), as well as their respective phosphate derivatives. It is a water-soluble vitamin, which exists in the form of phosphate in the body. The most versatile form of vitamin B_6 is pyridoxal 5′-phosphate (PLP), which is a cofactor of many proteins and enzymes in all organisms. As the most widely available commercial form, PN hydrochloride is extensively used in the pharmaceutical and food industries (Eliot and Kirsch, 2004).

Two *de novo* synthesis routes have been reported the 1-deoxyxylulose 5-phosphate (DXP)-dependent pathway and the DXP-independent pathway (Tanaka et al., 2005). From large-scale screening studies of different strains, found that the Gram-negative bacterium *Sinorhizobium meliloti* is the best producer of vitamin B_6, reaching a titer of 103 mg/L of B_6 isoforms within 168 h. Vitamin B_6 production was further increased to 1.30 g/L by expressing the *E. coli epd* gene and the native *dxs* gene in this *S. meliloti* strain (**Figure 1E**; Hoshino et al., 2007). *E. coli* and *B. subtilis* were also engineered to produce vitamin B_6. The vitamin B_6 production was enhanced to 78 mg/L within 31 h in *E. coli*, and *B. subtilis* produced 65 mg/L of PN when supplied with the precursor 4-hydroxy-L-threonine (4HT) (Hoshino et al., 2004). At present, the industry mainly adopts the oxazole method to produce vitamin B_6, and the current research also focuses on the improvement of the oxazole method synthesis process. In the process of biosynthesis, the PdxJ enzyme activity is very low ($k_{cat} = 0.07\ s^{-1}$), and the reaction step catalyzed by this enzyme is the rate-limiting step in the VB_6 biosynthetic pathway. The intermediate metabolite 4-phosphate hydroxy-threonine (4HTP)

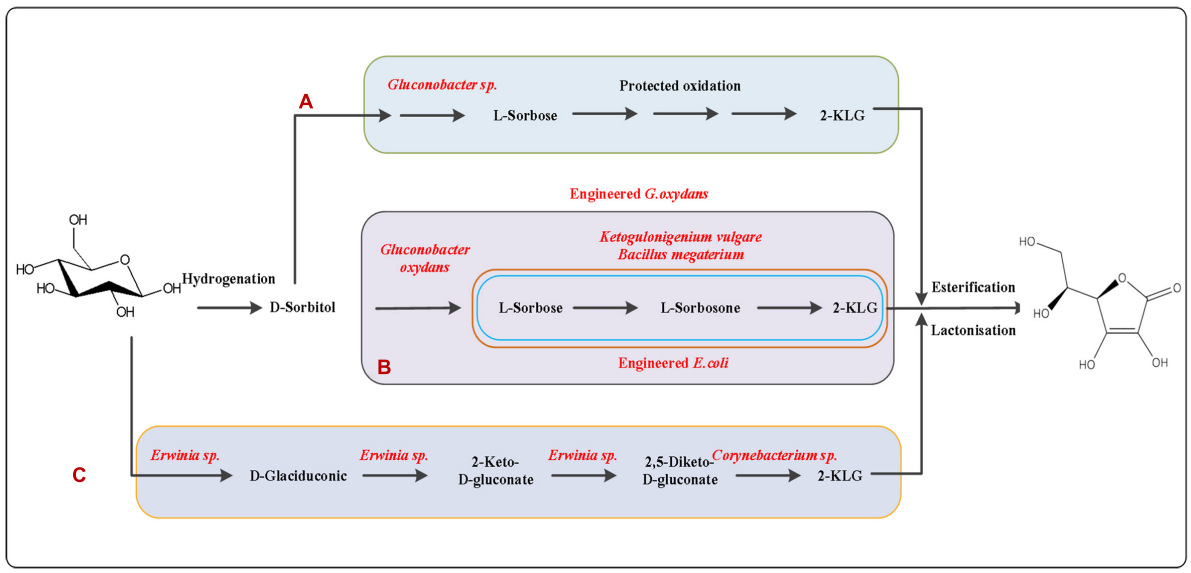


FIGURE 2 | L-ascorbic acid production pathway. **(A)** Classical Reichstein process: the seven-step Reichstein process involving six chemical transformations and one fermentation by *Gluconobacter oxydans*. **(B)** Two steps fermentation by mixed bacteria: *G. oxydans*, *Bacillus megaterium* and *Ketogulonicigenium vulgare*; **(C)** 2,5-Diketo-D-gluconic acid pathway: mixed or two stage fermentations with *Erwinia* sp. and *Corynebacterium* sp. or other species for the direct synthesis of 2-KLG.

is cytotoxic and is also the main bottleneck of biosynthesis. Therefore, the fermentative production of vitamin B_6 requires more effort to meet the commercial demand.

Vitamin B_7

Biotin is indispensable for the normal metabolism of fats and proteins (Lin and Cronan, 2011; Selvam et al., 2019). It is a nutrient necessary for human growth, development and normal function. Biotin combines with enzymes to participate in the process of carbon dioxide fixation and carboxylation in the body. The current large-scale production of D-biotin is mainly based on the Sternbach synthetic route, and the current industrial production method were improved on this basis. Unless biosynthetic methods can obtain high output at low cost, it is difficult to shake the position of chemical synthesis technology in industrial production. Nevertheless, it was recently reported that some microorganisms can overproduce biotin, which has been elaborated in *C. glutamicum, Mesorhizobium loti,* and *S. meliloti*.

In *Agrobacterium* and *Rhizobium* HK40, overexpression of the biotin operon from *E. coli* driven by the powerful *tac* promoter and introducing a modified RBS in front of *bioB* resulted in a biotin yield of 110 mg/L (**Figure 1G**; Streit and Entcheva, 2003). If the native biotin operon is overexpressed in *B. subtilis*, most enzymes will be strongly inhibited by the by-product of SAM. However, the high demand for SAM by biotin synthase and 7,8-diaminononanoate synthase is still a bottleneck that must be addressed in future research. If lysine is supplied to *B. subtilis*, BioK will use lysine as the amino donor of the biotin precursor to promote the production of biotin precursor (dephosphorization biotin), and the fermentation process used carbon-limited fed-batch growth conditions with computer control of dissolved oxygen concentrations, but the maximal titer can only reach 21 mg/L biotin. Therefore, improving the catalytic mechanism of biotin synthase is also a challenge for future research (Van Arsdell et al., 2005; Lin and Cronan, 2011).

Vitamin B_9

Naturally occurring folic acid is mostly found in the form of polyglutamic acid, and the biologically active form of folic acid is tetrahydrofolate (Myszczyszyn et al., 2019). Deficiency can lead to reduced hemoglobin content in red blood cells, impaired cell maturation and megaloblastic anemia (Lucock, 2000). *B. subtilis* or *A. gossypii* were successfully engineered to produce folic acid.

Jagerstad and Jastrebova (2013) achieved a 5-methyltetrahydrofolate (THF) titer of 0.95 mg/L by increasing the supply of precursor substances and blocking the catabolic pathway of THF in *B. subtilis*. With the continuous research progress, *A. gossypii* has attracted increasing interest as the chassis strain for folic acid production. *A. gossypii* can synthesize 0.04 mg/L of folic acid naturally, which can reach 6.59 mg/L after metabolic engineering treatment. This is also the highest production value reported to date (**Figure 1C**; Serrano-Amatriain et al., 2016). Since the commercial chemical synthesis of folic acid is cheap, unless the environmentally unfriendly part of the chemical synthesis process is restricted, there is still a long way to go for the fermentation of this product.

Vitamin B_{12}

Cobalamin is the only vitamin containing metal elements. Cobalamin is the general term for a class of corrin compounds containing cobalt (Osman et al., 2021). It is the largest and most complex vitamin molecule discovered so far. Vitamin B_{12} deficiency leads to increased formation of ring sideroblasts in pre-myelodysplastic syndromes (Kitago et al., 2020).

Vitamin B_{12} is synthesized by microorganisms through *de novo* synthesis or salvage synthesis in nature, but higher-animals and plants cannot produce it (**Figure 1H**; Fang et al., 2017). Although in the 19th century, researchers have completed the full chemical synthesis of vitamin B_{12}, the chemical synthesis method is too complicated and expensive, so the world's major suppliers rely on microbial fermentation to produce vitamins. *Pseudomonas denitrificans* and *Propionibacterium freudenreichii* being widely used in industrial fermentation to produce vitamin B_{12}. In order to improve the productivity of vitamin B_{12}, researchers have adopted a random mutagenesis method to construct a vitamin B_{12} overproducing strains by ultraviolet rays, nitrosoguanidine (NTG), nitrosomethylurethane and ethyleneimine (Blanche et al., 1992, 1995a,b). *Propionibacterium shermanii* was reported to produce vitamin B_{12} with a maximum titer of 200 mg/L (Sych et al., 2016). However, the aerobic *P. denitrificans* remains the most used industrial host, and the effect is obvious. Moreover, *P. denitrificans* has stronger production capacity than the anaerobic strain that produces vitamin B_{12} and is widely used in industrial production. Xia et al. (2015) increased the production of vitamin B_{12} to 198 $\pm$ 4.60 mg/L by optimizing the fermentation medium using response surface method. Our research group used *E. coli* strain MG1655 (DE3) as the starting strain to achieve *de novo* synthesis of vitamin B_{12} (Fang et al., 2018). Cai et al. (2018) later used riboswitch elements in *S. meliloti* for the first time, and successfully developed a flow cytometry high-throughput screening system for high-yield VB_{12} strains. The vitamin B_{12} titer of the best strain, *S. meliloti* MC5–2, reached 156 $\pm$ 4.20 mg/L, but the yield was still relatively low. At the same time, they also emphasized that the titer of vitamin B_{12} is greatly dependent on the medium composition (Cai et al., 2018).

Vitamin C

Vitamin C, also known as L-ascorbic acid (LAA), is an important cofactor for multiple enzyme reaction in the body (Paciolla et al., 2019; Kawahori et al., 2020). It can act as an antioxidant to scavenge free radicals and reduce oxidative stress, so a rapidly expanding market is the application of LAA as an additive to cosmetic products (Timoshnikov et al., 2020). Vitamin C deficiency can result in scurvy. Recently, researchers used biochemical methods combined with DNA recombination technology to produce vitamin C.

At present, L-AA is commercially manufactured via the classic seven-step Reichstein process using D-glucose as the initial substrate. The process involves six chemical steps and one fermentation steps for the oxidation of D-sorbitol to 2-keto-L-gulonic acid (2-KGA) by *Gluconobacter oxydans* and *Bacillus megaterium* (**Figure 2A**). Sugisawa et al. (2005) reported for the first time that *Ketogulonigenium vulgare* DSM 4025 can produce 1.37 g/L of L-AA under static culture conditions. Kim et al. (1996, 1998) reported that the respective enzymes from *Candida albicans* and *S. cerevisiae* convert not only D-arabinose to D-arabinono-1,4-lactone but also L-galactose to L-galactono-1,4-lactone *in vitro*. Experiments have shown that budding yeast cells overexpressing the endogenous D-arabinono-1,4-lactone oxidase and L-galactose dehydroge-nase can produce about 100 mg/L of L-ascorbic acid (Sauer et al., 2004). A microbiological consortium composed of *G. oxydans*, *K. vulgare,* and *B. endophyticus* was constructed to produce 2-KGA, and a final yield of 73.70 g/L was obtained within 30 h (**Figure 2B**; Ma et al., 2019). This result holds promise for the construction of a microbial cell factory for the production of vitamin C. However, it has been reported that mixed-bacteria fermentation can be unstable due to competition between the individual strains for nutrients and other factors. Therefore, mixed-bacteria fermentation technology has poor stability and low efficiency, which also hinders the pace of industrial production of vitamin C. Nevertheless, fermentation is expected to become the mainstream way of vitamin C production in the future if stable single strains can be used instead of mixed bacteria fermentation, while also shortening the production cycle.

FAT-SOLUBLE VITAMINS

Vitamin A

Vitamin A mainly includes β-carotene, α-carotene, and β-cryptoxanthin (Wise et al., 2021). β-carotene, a provitamin A carotenoid, is divided into all-*trans* and *cis* isomers (Yang et al., 2021). All-*trans*-β-carotene is the major isomer found in unprocessed carotene-rich plant foods, followed by its 9- and 13-*cis* isomers. β-carotene is an antioxidant, which not only inhibits singlet oxygen but also inhibits lipid peroxidation, thereby playing an important role in the prevention of disease (Kawata et al., 2018).

Carotene is mainly produced by fungi, some bacteria, and algae. For example, Yoon et al. increased the supply of IPP (isopentenyl diphosphate) and DMAPP (dimethylallyl diphosphate) through the introduction of foreign MVA (mevalonate) pathway (**Figure 3D**), thereby enhancing the production of carotenoids. The final engineered *E. coli* with a whole MVA pathway and β-carotene synthesis gene can produce β-carotene of 465 mg/L (**Figure 3C**; Yoon et al., 2009). Adenosine-triphosphate (ATP) and nicotinamide adenine dinucleotide phospha (NADPH) are two important cofactors in β-carotene biosynthesis pathway. Zhao et al. (2013) used *E. coli* as host cells, constructed and optimized a central metabolic module to increase the supply of ATP and NADPH in β-carotene synthesis pathway, thereby improving the yield of the β-carotene. Finally, the best strain CAR005 increased the β-carotene production to 2.1 g/L with a yield of 60 mg/g DCW in fed-batch fermentation (Zhao et al., 2013). Larroude et al. (2018) overexpressed heterologous carotene synthase (Crt) in *Yarrowia lipolytica* to make it produce high β-carotene. The fermentation yield of the engineered strain obtained by screening the best promoter was 1.5 g/L. By optimizing the fermentation conditions and using fed-batch fermentation, the yield of β-carotene was further increased production titer of 6.5 g/L and 90 mg/g DCW (Larroude et al., 2018). However, the insufficient number of precursors seriously hindered the industrialization process of β-carotene in the process of β-carotene synthesis in the future.

Vitamin D

Vitamin D refers to a group of fat-soluble secosteroids responsible for increasing intestinal absorption of magnesium, calcium, and phosphate, and many other biological effects. The most important compounds are vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol) in vitamin D. Vitamin D can increase intestinal absorption of calcium, magnesium, and phosphate, and can prevent many diseases (Yuan et al., 2020).

It is well known that the precursor of vitamin D_2 is ergosterol (Papoutsis et al., 2020). Vitamin D_2 is widely used in medical, food and other industries. The current commercial production of ergosterol is mainly produced by yeast fermentation. Tan et al. improved the production of ergosterol by optimizing the fermentation medium and screening high ergosterol producing strains (**Figure 3F**). The results show that dissolved oxygen (DO) can be used as the effective control parameter for yeast fed-batch fermentation. The total yield of ergosterol can be increased to 1.16 g/L when DO was controlled at 12% (±1%) and pulse fed-batch was used (Tan et al., 2003).

Vitamin D_3 cannot play a direct role in human and animals, but it can produce the physiologically active form 25-hydroxyvitamin D_3 (25-OH-VD_3) through the metabolism in liver. At present, the production process of 25-OH-VD_3 mainly includes chemical synthesis and light irradiation. The chemical reaction steps are cumbersome, some of them need halogen reagents, and the racemates are generated during the reaction, which makes the separation difficult. Therefore, more and more researchers pay attention to the fermentation of 25-OH-VD_3 by microorganisms. The strains used in microbial biosynthesis mainly include *Rhodococcus*, *Streptomyces*, *Pseudonocardia* sp., and *Mycobacterium*. Vitamin D_3 hydroxylase (Vdh) is a kind of cytochrome P450 monooxygenase, which can catalyze the two-step hydroxylation of vitamin D_3 (VD_3) to produce 25-OH-VD_3 and 1α,25-dihydroxyvitamin D_3. Yasutake et al. (2013) used nisin, a natural bioactive antimicrobial peptide, to treat *Rhodococcus* cells containing hydroxylase, and they found that 573 mg/L 25-OH-VD_3 can be synthesized. Although the current industrial production of vitamin D_3 is mainly dominated by chemical synthesis, microbial synthesis methods are more sustainable and do not produce impurities during the biosynthesis process, thus it will be taken priority in the future industrial production.

Vitamin E

Vitamin E is a group of lipid-soluble antioxidants, including tocopherols and tocotrienols (Muñoz and Munné-Bosch, 2019; Zeng Z. et al., 2020). These compounds are composed of an oxygen-containing double ring system with a hydrophobic prenyl side chain (Blake and Konings, 2019). Lack of vitamin E affects the function of T and B immune cells (Moriguchi and Muraga, 2000). Additionally, patients with severe impairment due to Alzheimer's disease improved significantly after receiving α-tocopherol (Sano et al., 1997). Considering various physiological effects of tocopherols, they are widely used in the manufacture of human dietary supplements, food preservatives and cosmetics. There are four different tocopherol compounds, named α, β, γ, and δ tocopherol. Among the four forms of vitamin E, α-tocopherol is the most biologically active (Kaiser et al., 1990). In nature, α-tocopherol is produced by photosynthetic organisms, e.g., eukaryotic algae and green plants, some prokaryotic cyanobacteria, such as *Synechocystis*, which can accumulate vitamin E in large amounts (**Figure 3B**; Taketomi et al., 1983).

Recently, *Euglena gracilis* was found to be suitable for the production of high-value products, such as amino acids and ascorbic acid (Schwarzhans et al., 2015). *E. gracilis* is the most promising host for the commercial production of α-tocopherol, with a high growth rate and α-tocopherol content, which accounts for more than 97% of the total tocopherol accumulated by *E. gracilis*. Tani and Tsumura added precursors such as homogentisate and L-tyrosine to *E. gracilis* growth medium, which increased the accumulation of α- tocopherol to 143.60 mg/L corresponding to 5.1 mg/g dry cell weight (DCW) (Tani and Tsumura, 1989). Durmaz (2007). explored the effect of nitrogen source and concentration on the accumulation of α-tocopherol in *Nannochloropsis oculata*. When sodium nitrate and ammonium chloride were used as inorganic nitrogen source, the highest content of α-tocopherol reached 2.32 ± 0.04 mg/g dry weight (DW) (**Table 2**). The research showed that higher concentrations of nitrogen in the form of NO_3^+ and NH_4^+ can promote production the of α-tocopherol (Durmaz, 2007).

To balance cell growth and product synthesis, Shen et al. (2020) recently combined heterologous genes from photosynthetic organisms with the endogenous shikimate and mevalonate pathways (MEP) to construct a strain of *S. cerevisiae* that produces tocotrienols (**Figure 3A**). By incorporating a newly designed cold-shock-triggered temperature control system, the phased control of cell biomass and tocotrienol accumulation by the engineered strains was successfully realized. The final total tocotrienol titer reached 320 mg/L in a 5 L fermenter, which laid the foundation for the production of natural vitamin E in a fully fermentative process (**Figure 3B**; Shen et al., 2020).

In general, compared with chemical total synthesis, the method of obtaining vitamin E directly through biotechnology has low yield and high cost, and is not suitable for large-scale production. Although chemical total synthesis is currently the main production method of vitamin E, there are still many problems with this technology, such as complex synthesis routes, high technical barriers, etc. Therefore, the development of safer and more efficient synthesis technology has become the main problem to improve the current situation of vitamin E.

Vitamin K

Vitamin K is a fat-soluble vitamin, which also called blood coagulation vitamin in virtue of the function of promoting blood coagulation and preventing osteoporosis (Henrik, 1973; Schwalfenberg, 2017; Zhou et al., 2019). There are two naturally occurring types of vitamin K, called vitamin K1 (phylloquinone/phytomenadione) and vitamin K2 (menaquinone, MK) (Holvik et al., 2019). Vitamin K1 is synthesized by plants, while vitamin K2 is synthesized by microorganisms and can be divided into 14 isoforms depending on the number of isoprenoid units connected to the menaquinone ring (**Figure 3E**; Schwalfenberg, 2017). Among

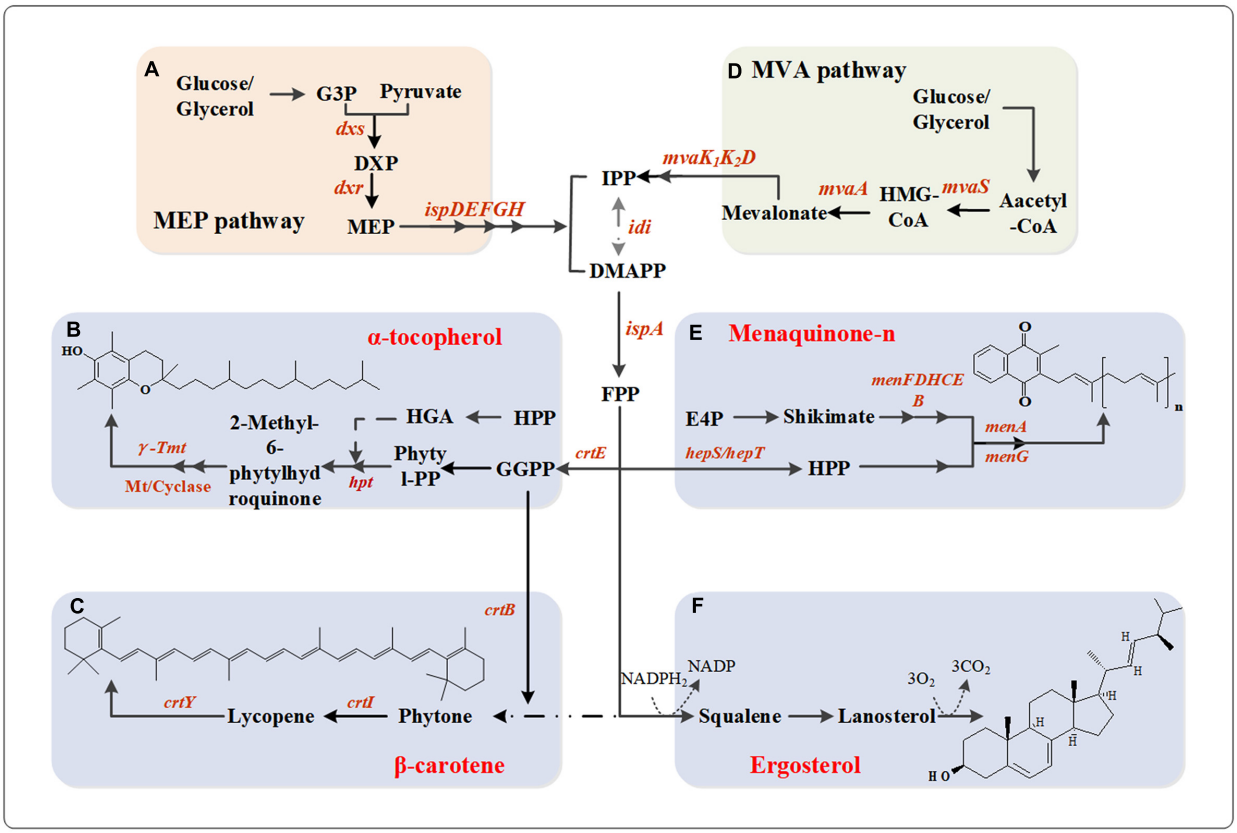


FIGURE 3 | Metabolic network pathway of vitamin A/D/E and vitamin K. **(A)** MEP pathway. Dxs, 1-deoxy-D-xylulose-5-phosphate synthase; Dxr, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; IspD, 2-*C*-methyl-D-erythritol 4-phosphate cytidylyltransferase; IspE, 4-diphosphocytidyl-2-*C*-methyl-D-erythritol kinase; IspF, 2-*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase; IspG, 1-hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate synthase; IspH, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; **(B)** α-tocopherol biosynthesis pathway. Idi, isopentenyl diphosphate isomerase; IspA, geranyltranstransferase; CrtE, GGPP synthase; hpt, hypoxanthine phosphoribosyltransferase; γ-Tmt, γ-tocopherol methyl-transferase; Mt, methyl-transferase. **(C)** β-carotene biosynthesis pathway. CrtB, phytoene synthase; CrtI, phytoene desaturase; CrtY, lycopene cyclase. **(D)** MVA pathway. MvaS, HMG-CoA synthase; MvaA, HMG-CoA reductase; MvaK1, mevalonate kinase; MvaK2, phosphomevalonate kinase; MvaD, diphosphomevalonate decarboxylase; **(E)** Menaquinone-*n* biosynthesis pathway. HepS/HepT, heptaprenyl diphosphate synthase component I/II; MenF, isochorismate synthase; MenD, 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase; MenH, demethylmenaquinone methyltransferase; MenC, *o*-succinylbenzoate synthase; MenE, *o*-succinylbenzoate-CoA ligase; MenB, 1,4-dihydroxy-2-naphthoyl-CoA synthase; MenA, 1,4-dihydroxy-2-naphthoate heptaprenyltransferase; MenG, demethylmenaquinone methyltransferase. **(F)** Ergosterol biosynthesis pathway.

them, menaquinone-7 (MK-7) is the most effective subtype of vitamin K with a very long half-life in circulation. Notably, MK-7 can be synthesized in the *cis*, *trans*, and *cis/trans* forms, but only the all-*trans* form is biologically active (Szterk et al., 2018).

The biosynthesis of MK-7 from the embden-meyerhof-parnas (EMP) pathway, pentose phosphate pathway (PPP), MVA pathway and menadione synthesis (MK) pathway (**Figure 3C**). A number of microorganisms have been used to produce MK-7, including *B. subtilis, E. coli*, lactic acid bacteria, *Flavobacterium* sp., and *B. amyloliquefaciens* (Sharma et al., 1993; Morishita et al., 1999; Sato et al., 2001; Wu, 2011; Taguchi et al., 2014).

Bacillus subtilis isolated from natto (a traditional Japanese food), the strain found in the eponymous fermented Japanese beans, has been certified by the FDA as a food-safe and has a strong ability to produce MK-7. Accordingly, *B. subtilis natto* was used as a parent strain to develop some of the industrial strains currently on the market. In industrial production, *B. subtilis natto* fermentation broth was sprayed and dried, and the dry powder from the fermentation broth was subjected to solvent extraction. The obtained extract was condensed into a paste and then purified by chromatography. Using this process, the final yield of MK-7 can reach 200–300 mg/L in the fermentation cycle of 16–24 h (Chen et al., 2016).

Cui et al. (2019) developed a bifunctional quorum-sensing system in *B. subtilis* 168 to engineer the synthesis modules of MK-7. The resulting strain was capable of producing 360 mg/L MK-7 in shak flasks and 200 mg/L MK-7 in 15-L bioreactor (Cui et al., 2019). Recently, comparative transcriptomics revealed that cell membrane and electron transfer engineering in *B. subtilis* can improve the synthesis of MK-7. The resulting strain reached a product titer of 410 mg/L after 6 days in shake-flask culture, which is the highest value reported to date (Cui et al., 2020). In the current market environment, the production of natural all-trans MK-7 is via liquid fermentation of *B. subtilis*

TABLE 2 | Fat-soluble vitamins produced by biotechnological methods.

Strain	Biotechnological method	Main culture substances	Yield	References
Vitamin A				
E. coli	Glycerol as the carbon source and harboring the whole MVA pathway.	2YT medium; Glycerol	465 mg/L	Yoon et al., 2009
E. coli	Overexpression of *crt* genes, *dxs*, *idi*, *sucAB*, *sdhABCD,* and *talB*.	LB medium	2.1 g/L	Zhao et al., 2013
Y. lipolytica	Expressing the heterologous pathway and screen the best combination of promoters for each of the studied genes.	YPD medium; MM medium; YNB medium	6.5 g/L	Larroude et al., 2018
Vitamin D				
S. cerevisiae	DO was kept at 12% (±1%) and pulse fed-batch was used.	MM medium	1.16 g/L VD_2	Tan et al., 2003
R. erythropolis	Insert the gene-expression cassette encoding *Bacillus megaterium* glucose dehydrogenase-IV into the chromosome of *R. erythropolis*.	MM medium	573 mg/L VD_3	Yasutake et al., 2013
Vitamin E				
E. gracilis	Add effective additives (homogentisate and L-tyrosine); Optimize the concentration of ethanol and protein.	KH medium; Homogentisate; L-tyrosine	5.10 mg/L	Tani and Tsumura, 1989
E. gracilis	Determination of the amount of α-tocopherol produced under photoautotrophically, heterotrophically or photoheterotrophically.	MM; Methane	8.60 ± 0.22 mg/L	Grimm et al., 2015
Stichococcus bacillaris	Ballon bioreactor culture with MeJa as inducer.	Methyl jasmonate (MeJa); Algal culture	0.60 mg/g (DW)	Sivakumar et al., 2014
Nannochloropsis oculata	Optimize the carbon source of the medium ($NO3^+$-N and $NH4^+$-N) and harvest time.	F/2 medium; Ammonium chloride	2.32 ± 0.04 mg/g (DW)	Durmaz, 2007
S. cerevisiae	Gene cloning from various photosynthetic organisms; Codon optimization and protein truncation.	SD medium	320 mg/L	Shen et al., 2020
Vitamin K (MK-4/MK-7)				
B. subtilis natto	Optimum media conditions and screening producing strain (Different nutrients of the culture medium will affect the yield of MK-7).	Glycerol	62.32 ± 0.34 mg/L	Berenjian et al., 2011
B. subtilis natto	Fermentation using soybean extract and screening highest MK7 yielding strain from commercially available natto.	Soy granules; Amylase	67.01 ± 0.18 mg/kg	Mahanama et al., 2011
B. subtilis	Deletion of *PAS-A, kinB, spoIIA, spo0IIE, dhbB, and ptsG;* Overexpresion of *menF*, *menB*, *menE*, *entC*, *ppsA*, *aroK*, *ispA*, *hepS/T*, *kdpG*, *dxr*, *dxs*, *fni*, *menA*.	LB Medium	200 mg/L	Cui et al., 2019
B. subtilis	Overexpresion of *BS*20- *qcrA-C and tatAD-CD.*	LB Medium	410 mg/L	Cui et al., 2020

YNB medium, yeast nitrogen base; DW, Dry weight; SD medium, synthetic dextrose medium; MYP agar, mannitol, egg yolk and polymyxin agar; TBAB, tetrabutylammonium bromide; CDW, cell dry weight.

natto, which is safe, natural and controllable, and occupies the mainstream position in the market. Compared with the chemically synthesized of *trans*-MK-7, it has a higher yield and fewer impurities.

CONCLUSION

The fermentative production of vitamins using bacteria, yeasts or microalgae has many advantages over traditional chemical synthesis methods. From the aspects of safety, biological activity, absorption rate, etc., vitamins manufactured by biological methods can be more suitable for both internal and external applications (Yuan et al., 2020). Although the fermentation of VB_2 and VB_{12} has technologically matured and is being applied in industrial production, fermentation methods for the remaining B-group vitamins have yet to be developed or require significant yield improvement.

Vitamin C has a large market, and its production method is mainly based on single-bacteria fermentation, which eliminates the dependency of associated bacteria by replacing accompanying bacteria with associated active agents (Vandamme and Revuelta, 2016a). However, the current market situation indicates that vitamin C production has overcapacity, the downstream processing is complicated, and the market demand is concentrated in the field of medicine and food. For these reasons, the momentum of price increase will remain slow in the future. Vitamin C fermentation technology can explore the mechanism of a variety of accompanying bacteria, establish their anabolism database, or use isotope technology to label and trace the individual metabolites. It is also possible to design heterologous assembly modules for 2-KGA synthesis, and study adaptation mechanisms in microbial chassis cells, so as to achieve higher productivity (Liu et al., 2011).

The current biosynthesis product of vitamin A is mainly focused on β-carotene. The biosynthesis of β-carotene

has successfully established a large-scale production process through classical and reasonable microbial metabolic engineering. However, due to the high barriers of intermediate industry and the complex process of synthesis and metabolism, the future research will face more difficult challenges. At present, the industrial production of vitamin D is mainly through the chemical synthesis of active 25-OH-VD_3 and 1α,25-dihydroxyvitamin D_3, but the biggest obstacle is the assured quality and security of supply for raw materials, which must be cholesterol with purity greater than 95% (NF grade). Therefore, the key to solving the problem of raw materials is to develop more production bacteria, optimize their metabolic pathways and make them highly productive. With the continuous optimization and technological progress of vitamin E fermentation, the overall cost of the industry has fallen, which promoted the growth of the industry. Unfortunately, although studies have shown that photosynthetic microorganisms have considerable potential for the production of tocopherols, light-driven fermentation is costly, which makes commercialization difficult. However, due to the considerable potential *E. gracilis* and the conditions of the cultivation environment, the construction of a specifically designed photo-bioreactor may be a feasible research direction for the production of tocopherol. Moreover, the controllable temperature-sensitive control system may also be a key control technology for vitamin E production. Among vitamin K producing bacteria, *B. subtilis natto* seems to be the most promising candidate for MK production. Many researchers have optimized the design of fermentation modes, medium components, and culture conditions. They have also applied genetic engineering and other means to increase MK production (Szterk et al., 2018). However, to achieve higher industrial output, the technology needs to be further improved. Some studies have used biofilm reactors, which may become a promising new area for future research.

Recently, our research group used *E. coli* MG1655 (DE3) as chassis strains and achieved the *de novo* synthesis of vitamin B_{12} via metabolic engineering and optimization of fermentation conditions. In addition, we have not only proved that *E. coli* is a microbial biosynthesis platform for the production of vitamin B_{12}, also provides an encouraging example of how the dozens of proteins in a complex biosynthetic pathway can be transferred between organisms to promote industrial production (Fang et al., 2018). In addition, our research group is also doing metabolism research on vitamin B_2, B_6, B_7, and vitamin K.

In general, the development of synthetic biotechnology provides new opportunities for the construction of vitamin cell factories. First, high-throughput screening of high-yield strains, the CRISPR/Cas9 genome editing technology, and automatic gene assembly technology provide important technical means for the mining and genetic modification of chassis cells (Chang et al., 2019; Zeng W. et al., 2020; Zhang and Showalter, 2020). Further, the output of vitamin products in different dimensions will be increased by transforming the complex and multi-enzyme pathways required for the production of vitamins, establishing microbial flora with controllable functions and stability, and application of some advanced engineering technology, such as the cold-shock-triggered temperature control system, dynamic control of gene expression systems, different types of biosensors, cell-free systems and computer-aided design, etc. (Koo et al., 2020; Marucci et al., 2020; Sachsenhauser et al., 2020; Shen et al., 2020; Glasscock et al., 2021). Additionally, the modular and orthogonal strategies are increasingly supporting the construction of vitamin cell factories (Liu et al., 2015). The mining and design of biological components, the assembly and integration of elements and modules, and the optimization and adaptation of the fermentation system are also important for efficient production of vitamins (Santos-Merino et al., 2019). However, there will be many challenges in the field of synthetic biotechnology in the future, including the compatibility between flexible biological systems and rigid engineering systems, or the universality of biological system reconstruction. It will be necessary and important to advance the existing technology, combine it with new strategies, and conduct interdisciplinary research to establish novel microbial cell factories for the industrial fermentation of most vitamins. All in all, we firmly believe that the industrialization of the fermentation production of vitamins is expected to become a broader, safer and more sustainable manufacturing with the continuous advancement of synthetic biotechnology and metabolic engineering.

AUTHOR CONTRIBUTIONS

YW and LL: manuscript planning, writing, and revision. ZJ and DZ: manuscript revision and writing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the National Key R&D Program of China (no. 2019YFA0905300), the National Natural Science Foundation of China under grants nos. 31670604 and 31970324, the Natural Science Foundation of Liaoning Province of China under grant no. 2019020758, the Science and Technology Project of Liaoning Education Department under grant no. 819001110761, and the Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (no. TSBICIP-CXRC-004).

REFERENCES

Abbas, C. A., and Sibirny, A. A. (2011). Genetic control of biosynthesis and transport of riboflavin and flavin nucleotides and construction of robust biotechnological producers. *Microbiol. Mol. Biol. Rev.* 75, 321–360. doi: 10.1128/MMBR.00030-10

Acevedo-Rocha, C. G., Gronenberg, L. S., Mack, M., Commichau, F. M., and Genee, H. J. (2019). Microbial cell factories for the sustainable manufacturing

of B vitamins. *Curr. Opin. Biotechnol* 56, 18–29. doi: 10.1016/j.copbio.2018.07.006

Andreieva, Y., Petrovska, Y., Lyzak, O., Liu, W., and Sibirny, A. (2020). Role of the regulatory genes SEF1, VMA1 and SFU1 in riboflavin synthesis in the flavinogenic yeast *Candida famata* (Candida flareri). *Yeast* 37, 497–504. doi: 10.1002/yea.3503

Balasubramaniam, S., Christodoulou, J., and Rahman, S. (2019). Disorders of riboflavin metabolism. *J. Inherited Metab. Dis.* 42, 608–619.

Begley, T. P., Downs, D. M., Ealick, S. E., Mclafferty, F. W., Loon, A. P. G. M. V., Taylor, S., et al. (1999). Thiamin biosynthesis in prokaryotes. *Arch. Microbiol.* 171, 293–300. doi: 10.1007/s002030050713

Belenky, P., Stebbins, R., Bogan, K. L., Evans, C. R., and Brenner, C. (2011). Nrt1 and Tna1-independent export of NAD+ precursor vitamins promotes NAD+ homeostasis and allows engineering of vitamin production. *PLoS One* 6:e19710. doi: 10.1371/journal.pone.0019710

Berdanier, C. D., and Adkins, T. K. (2019). Water-soluble vitamins. *Adv. Nutr. Micronutr.*

Berenjian, A., Mahanama, R., Talbot, A., Biffin, R., Regtop, H., Valtchev, P., et al. (2011). Efficient media for high menaquinone-7 production: response surface methodology approach. *N. Biotechnol.* 28, 665–672. doi: 10.1016/j.nbt.2011.07.007

Blake, C. J., and Konings, E. J. M. (2019). Committee on food nutrition: fat-soluble vitamins: water-soluble Vitamins. *J. AOAC Int.* 88, 325–330. doi: 10.1093/jaoac/88.1.325

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., and Battersby, A. R. (1995a). Vitamin B12: how the problem of its biosynthesis was solved. *Angewandte Chemie Int. Edition* 34, 383–411. doi: 10.1002/chin.199527327

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., and Battersby, A. R. (1995b). Vitamin B12: wie das problem seiner biosynthese gelöst wurde. *Angewandte Chemie* 107, 421–452. doi: 10.1002/ange.19951070404

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Levy-Schil, S., and Thibaut, D. (1992). Polypeptides involved in the biosynthesis of cobalamines and/or cobamides, dna sequences coding for these polypeptides, and their preparation and use. *WO* 1991011518:A1.

Burgess, C., O'Connell-Motherway, M., Sybesma, W., Hugenholtz, J., and van Sinderen, D. (2004). Riboflavin production in Lactococcus lactis: potential for in situ production of vitamin-enriched foods. *Appl. Environ. Microbiol.* 70, 5769–5777. doi: 10.1128/AEM.70.10.5769-5777.2004

Cai, Y., Xia, M., Dong, H., Qian, Y., Zhang, T., Zhu, B., et al. (2018). Engineering a vitamin B12 high-throughput screening system by riboswitch sensor in *Sinorhizobium meliloti*. *BMC Biotechnol.* 18:27. doi: 10.1186/s12896-018-0441-2

Capone, K., and Sentongo, T. (2019). The ABCs of nutrient deficiencies and toxicities. *Pediatr. Ann.* 48, e434–e440. doi: 10.3928/19382359-20191015-01

Cardinale, S., Tueros, F. G., and Sommer, M. O. A. (2017). Genetic-metabolic coupling for targeted metabolic engineering. *Cell Rep.* 20, 1029–1037. doi: 10.1016/j.celrep.2017.07.015

Cea, P. A., Araya, G., Vallejos, G., Recabarren, R., and Castro-Fernandez, V. (2020). Characterization of hydroxymethylpyrimidine phosphate kinase from mesophilic and thermophilic bacteria and structural insights into their differential thermal stability. *Arch. Biochem. Biophys.* 688:108389. doi: 10.1016/j.abb.2020.108389

Chand, T., and Savitri, B. (2016). *Vitamin B3, Niacin*. New York, NY: John Wiley & Sons, Ltd.

Chang, H., Wang, C., Wang, P., Zhou, J., and Li, B. (2019). DNA assembly technologies: a review. *Sheng Wu Gong Cheng Xue Bao* 35, 2215–2226. doi: 10.13345/j.cjb.190273

Chauhan, N., and Poddar, R. (2019). In silico pharmacophore modeling and simulation studies for searching potent antileishmanials targeted against Leishmania donovani nicotinamidase. *Comp. Biol. Chem.* 83:107150. doi: 10.1016/j.compbiolchem.2019.107150

Chen, J., Duan, P. L., Chen, L. H., Qiu, R. X., Cai, L. G., Hong, Z. L., et al. (2016). *Bacillus Subtilis Natto and Method for Producting Protein MK-7*. CN 104262129B. Beijing: China National Intellectual Property Administration.

Commichau, F. M., Alzinger, A., Sande, R., Bretzel, W., Reuss, D. R., Dormeyer, M., et al. (2015). Engineering *Bacillus subtilis* for the conversion of the antimetabolite 4-hydroxy-l-threonine to pyridoxine. *Metab. Eng.* 29, 196–207. doi: 10.1016/j.ymben.2015.03.007

Cui, S., Lv, X., Wu, Y., Li, J., Du, G., Ledesma-Amaro, R., et al. (2019). Engineering a bifunctional Phr60-Rap60-Spo0A quorum-sensing molecular switch for dynamic fine-tuning of menaquinone-7 synthesis in *Bacillus subtilis*. *ACS Synth. Biol.* 8, 1826–1837. doi: 10.1021/acssynbio.9b00140

Cui, S., Xia, H., Chen, T., Gu, Y., Lv, X., Liu, Y., et al. (2020). Cell membrane and electron transfer engineering for improved synthesis of menaquinone-7 in *Bacillus subtilis*. *iScience* 23:100918. doi: 10.1016/j.isci.2020.100918

Dmytruk, K., Lyzak, O., Yatsyshyn, V., Kluz, M., Sibirny, V., Puchalski, C., et al. (2014). Construction and fed-batch cultivation of *Candida famata* with enhanced riboflavin production. *J. Biotechnol.* 172, 11–17. doi: 10.1016/j.jbiotec.2013.12.005

Dmytruk, K. V., Yatsyshyn, V. Y., Sybirna, N. O., Fedorovych, D. V., and Sibirny, A. A. (2011). Metabolic engineering and classic selection of the yeast *Candida famata* (Candida flareri) for construction of strains with enhanced riboflavin production. *Metab. Eng.* 13, 82–88. doi: 10.1016/j.ymben.2010.10.005

Dorrestein, P. C., Zhai, H., Mclafferty, F. W., and Begley, T. P. (2004). The biosynthesis of the thiazole phosphate moiety of thiamin. *Chem. Biol.* 11, 1373–1381. doi: 10.1016/j.chembiol.2004.08.009

Downs, D. M. (1992). Evidence for a new, oxygen-regulated biosynthetic pathway for the pyrimidine moiety of thiamine in *Salmonella typhimurium*. *J. Bacteriol.* 174, 1515–1521. doi: 10.1128/jb.174.5.1515-1521.1992

Downs, D. M., and Petersen, L. (1994). apbA, a new genetic locus involved in thiamine biosynthesis in *Salmonella typhimurium*. *J. Bacteriol.* 176, 4858–4864. doi: 10.1007/BF02182164

Downs, D. M., and Roth, J. R. (1991). Synthesis of thiamine in *Salmonella typhimurium* independent of the purF function. *J. Bacteriol.* 173, 6597–6604. doi: 10.1128/jb.173.20.6597-6604.1991

Durmaz, Y. (2007). Vitamin E (α-tocopherol) production by the marine microalgae *Nannochloropsis oculata* (Eustigmatophyceae) in nitrogen limitation. *Aquaculture* 272, 717–722. doi: 10.1016/j.aquaculture.2007.07.213

Eliot, A. C., and Kirsch, J. F. (2004). Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* 73, 383–415. doi: 10.1146/annurev.biochem.73.011303.074021

Fang, H., Kang, J., and Zhang, D. (2017). Microbial production of vitamin B12: a review and future perspectives. *Microbial Cell Factories* 16:15. doi: 10.1186/s12934-017-0631-y

Fang, H., Li, D., Kang, J., Jiang, P., Sun, J., and Zhang, D. (2018). Metabolic engineering of *Escherichia coli* for de novo biosynthesis of vitamin B12. *Nat. Commun.* 9:4917. doi: 10.1038/s41467-018-07412-6

Fischer, M., and Bacher, A. (2005). Biosynthesis of flavocoenzymes. *Nat. Prod. Rep.* 22, 324–350. doi: 10.1039/b210142b

Glasscock, C. J., Biggs, B. W., Lazar, J. T., Arnold, J. H., Burdette, L. A., Valdes, A., et al. (2021). Dynamic control of gene expression with riboregulated switchable feedback promoters. *ACS Synth Biol.* doi: 10.1021/acssynbio.1c00015 Online ahead of print.

Grimm, P., Risse, J. M., Cholewa, D., Muller, J. M., Beshay, U., Friehs, K., et al. (2015). Applicability of *Euglena gracilis* for biorefineries demonstrated by the production of α-tocopherol and paramylon followed by anaerobic digestion. *J. Biotechnol.* 215, 72–79. doi: 10.1016/j.jbiotec.2015.04.004

Henrik, D. (1973). CLIV. The antihaemorrhagic vitamin of the chick. *Nutrition Rev.* 31:121. doi: 10.1111/j.1753-4887.1973.tb05050.x

Hohmann, H. P., Dijl, J. M. V., Krishnappa, L., and Prágai, Z. (2016). *Host Organisms: Bacillus subtilis*. Hoboken, NJ: John Wiley & Sons, Ltd.

Holvik, K., Fryland, L., Haugen, M., Henjum, S., and Parr, C. L. (2019). Assessment of dietary intake of vitamin K and maximum limits for vitamin K in food supplements. *Eur. J. Nutrition Food Safety* 9, 96–98. doi: 10.9734/ejnfs/2019/v9i230044

Hoshino, T., Ichikawa, K., and Nagahashi, Y. (2007). *Microorganism and Process for Preparing Vitamin B6*. Washington, DC: U.S. Patent and Trademark Office. U.S. Patent No 20060127992 A1.

Hoshino, T., Ichikawa, K., and Tazoe, M. (2004). *Recombinant Microorganism for the Production*. Washington, DC: U.S. Patent and Trademark Office. U.S. Patent No EP 1 543 139 B1.

Huser, A. T., Chassagnole, C., Lindley, N. D., Merkamm, M., Guyonvarch, A., Elisakova, V., et al. (2005). Rational design of a *Corynebacterium glutamicum* pantothenate production strain and its characterization by metabolic flux analysis and genome-wide transcriptional profiling. *Appl. Environ. Microbiol.* 71, 3255–3268. doi: 10.1128/AEM.71.6.3255-3268.2005

Hwang, S., Cordova, B., Abdo, M., Pfeiffer, F., and Maupin-Furlow, J. A. (2017). ThiN as a versatile domain of transcriptional repressors and catalytic enzymes of thiamine biosynthesis. *J. Bacteriol.* 199:JB.00810-16. doi: 10.1128/JB.00810-16

Ifuku, O., Koga, N., Haze, S., Kishimoto, J., Arai, T., and Wachi, Y. (1995). Molecular analysis of growth inhibition caused by overexpression of the biotin operon in *Escherichia coli*. *Biosci. Biotechnol. Biochem.* 59, 184–189. doi: 10.1271/bbb.59.184

Jagerstad, M., and Jastrebova, J. (2013). Occurrence, stability, and determination of formyl folates in foods. *J. Agric. Food Chem.* 61, 9758–9768. doi: 10.1021/jf4028427

Jurgenson, C. T., Begley, T. P., and Ealick, S. E. (2009). The structural and biochemical foundations of thiamin biosynthesis. *Annu. Rev. Biochem.* 78, 569–603. doi: 10.1146/annurev.biochem.78.072407.102340

Kaiser, S., Di, M. P., Murphy, M. E., and Sies, H. (1990). Physical and chemical scavenging of singlet molecular oxygen by tocopherols. *Arch. Biochem. Biophys.* 277, 101–108. doi: 10.1016/0003-9861(90)90556-E

Kawahori, K., Kondo, Y., Yuan, X., Kawasaki, Y., and Hashimoto, K. (2020). Ascorbic acid during the suckling period is required for proper DNA demethylation in the liver. *Sci. Rep.* 10:21228. doi: 10.1038/s41598-020-77962-7

Kawata, A., Murakami, Y., Suzuki, S., and Fujisawa, S. (2018). Anti-inflammatory activity of β-carotene, lycopene and tri-n-butylborane, a scavenger of reactive oxygen species. *In Vivo* 32, 255–264. doi: 10.21873/invivo.11232

Kim, S. T., Huh, W. K., Kim, J. Y., Hwang, S. W., and Kang, S. O. (1996). D-Arabinose dehydrogenase and biosynthesis of erythroascorbic acid in Candida albicans. *Biochim. Biophys. Acta* 1297, 1–8. doi: 10.1016/0167-4838(96)00077-5

Kim, S. T., Huh, W. K., Lee, B. H., and Kang, S. O. (1998). D-arabinose dehydrogenase and its gene from Saccharomyces cerevisiae. *Biochim. Biophys. Acta* 1429, 29–39. doi: 10.1016/S0167-4838(98)00217-9

Kitago, M., Kase, Y., Iwata, Y., Suwa, Y., Tsuchiya, H., Hanata, N., et al. (2020). Deteriorating anemia in an 86-year-old man was improved by prednisolone. *Geriatr. Gerontol. Int.* 20, 1091–1092. doi: 10.1111/ggi.14024

Koo, J., Yang, J., and Park, H. (2020). Cell-free systems: recent advances and future outlook. *Biotechnol. Bioprocess Eng.* 25, 955–961. doi: 10.1007/s12257-020-0013-x

Larroude, M., Celinska, E., Back, A., Thomas, S., Nicaud, J. M., and Ledesma-Amaro, R. (2018). A synthetic biology approach to transform *Yarrowia lipolytica* into a competitive biotechnological producer of β-carotene. *Biotechnol. Bioeng.* 115, 464–472. doi: 10.1002/bit.26473

Leonardi, R., and Jackowski, S. (2007). Biosynthesis of pantothenic acid and coenzyme a. *EcoSal Plus* 2:10.1128/ecosallus.3.6.3.4. doi: 10.1128/ecosalplus.3.6.3.4

Li, K. T., Liu, D. H., Chu, J., Wang, Y. H., Zhuang, Y. P., and Zhang, S. L. (2008). An effective and simplified pH-stat control strategy for the industrial fermentation of vitamin B(12) by *Pseudomonas* denitrificans. *Bioprocess Biosyst. Eng.* 31, 605–610. doi: 10.1007/s00449-008-0209-5

Lin, S., and Cronan, J. E. (2011). Closing in on complete pathways of biotin biosynthesis. *Mol. Biosyst.* 7, 1811–1821. doi: 10.1039/c1mb05022b

Liu, L., Xu, W., Jia, H., and Wu, J. (2011). A novel and effective multi-constrained QoS routing scheme in WMNs. *Front. Electrical Electronic Eng. China* 6:507–514. doi: 10.1007/s11460-011-0118-2

Liu, S., Dicker, K. T., and Jia, X. (2015). Modular and orthogonal synthesis of hybrid polymers and networks. *Chem. Commun. (Camb)* 51, 5218–5237. doi: 10.1039/c4cc09568e

Lucock, M. (2000). Folic acid: nutritional biochemistry, molecular biology, and role in disease processes. *Mol. Genet. Metab.* 71, 121–138. doi: 10.1006/mgme.2000.3027

Ma, Q., Bi, Y. H., Wang, E. X., Zhai, B. B., Dong, X. T., Qiao, B., et al. (2019). Integrated proteomic and metabolomic analysis of a reconstructed three-species microbial consortium for one-step fermentation of 2-keto-L-gulonic acid, the precursor of vitamin C. *J. Ind. Microbiol. Biotechnol.* 46, 21–31. doi: 10.1007/s10295-018-2096-3

Mahanama, R., Berenjian, A., Valtchev, P., Talbot, A., Biffin, R., Regtop, H., et al. (2011). Enhanced production of menaquinone 7 via solid substrate fermentation from *Bacillus subtilis*. *Int. J. Food Eng.* 7, 1–2. doi: 10.2202/1556-3758.2314

Marucci, L., Barberis, M., Karr, J., Ray, O., Race, P. R., de Souza Andrade, M., et al. (2020). Computer-aided whole-cell design: taking a holistic approach by integrating synthetic with systems biology. *Front. Bioeng. Biotechnol.* 8:942. doi: 10.3389/fbioe.2020.00942

Moriguchi, S., and Muraga, M. (2000). Vitamin E and immunity. *Vitamins Hormones-advances Res. Appl.* 59, 305–336.

Morishita, T., Tamura, N., Makino, T., and Kudo, S. (1999). Production of menaquinones by lactic acid bacteria. *J. Dairy Sci.* 82, 0–1903. doi: 10.3168/jds.S0022-0302(99)75424-X

Muñoz, P., and Munné-Bosch, S. (2019). Vitamin E in plants: biosynthesis, transport, and function. *Trends Plant Sci.* 24, 1040–1051. doi: 10.1016/j.tplants.2019.08.006

Myszczyszyn, A., Krajewski, R., Ostapów, M., and Hirnle, L. (2019). Folic acid – role in the body, recommendations and clinical significance. *Pielegniarstwo XXI wieku / Nursing in the 21st Century* 18, 50–59. doi: 10.2478/pielxxiw-2019-0007

Nie, G., Yang, X., Liu, H., Wang, L., and Gong, G. (2013). N+ ion beam implantation of tannase-producing *Aspergillus niger* and optimization of its process parameters under submerged fermentation. *Annals Microbiol.* 63, 279–287. doi: 10.1007/s13213-012-0471-2

Osman, D., Cooke, A., Young, T. R., Deery, E., and Warren, M. J. (2021). The requirement for cobalt in vitamin B12: a paradigm for protein metalation. *Biochimica et Biophysica Acta (BBA) - Mol. Cell Res.* 1868:118896. doi: 10.1016/j.bbamcr.2020.118896

Paciolla, C., Fortunato, S., Dipierro, N., Paradiso, A., and Pinto, M. C. D. (2019). Vitamin C in plants: from functions to biofortification. *Antioxidants* 8:519. doi: 10.3390/antiox8110519

Palmer, L. D., and Downs, D. M. (2013). The thiamine biosynthetic enzyme ThiC catalyzes multiple turnovers and is inhibited by S-adenosylmethionine (AdoMet) metabolites. *J. Biol. Chem.* 288, 30693–30699. doi: 10.1074/jbc.M113.500280

Papoutsis, K., Grasso, S., Menon, A., Brunton, N. P., Lyng, J. G., Jacquier, J.-C., et al. (2020). Recovery of ergosterol and vitamin D2 from mushroom waste-potential valorization by food and pharmaceutical industries. *Trends Food Sci. Technol.* 99, 351–366. doi: 10.1016/j.tifs.2020.03.005

Rao, Y. M., and Sureshkumar, G. K. (2000). Direct biosynthesis of ascorbic acid from glucose by *Xanthomonas campestris* through induced free-radicals. *Biotechnol. Lett.* 22, 407–411. doi: 10.1023/B:DOBS.0000033277.54135.71

Sachsenhauser, V., Deng, X., Kim, H. H., Jankovic, M., and Bardwell, J. C. A. (2020). Yeast tripartite biosensors sensitive to protein stability and aggregation propensity. *ACS Chem. Biol.* 15, 1078–1088. doi: 10.1021/acschembio.0c00083

Sahm, H., and Eggeling, L. (1999). D-Pantothenate synthesis in *Corynebacterium glutamicum* and use of panBC and genes encoding L-valine synthesis for D-Pantothenate overproduction. *Appl. Environ. Microbiol.* 65, 1973–1979. doi: 10.1016/S0027-5107(99)00041-X

Sano, M., Ernesto, C., Thomas, R. G., Klauber, M. R., Schafer, K., Grundman, M., et al. (1997). A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. *New England J. Med.* 336, 1216–1222. doi: 10.1056/nejm199704243361704

Santos-Merino, M., Singh, A., and Ducat, D. (2019). New applications of synthetic biology tools for cyanobacterial metabolic engineering. *Front. Bioeng. Biotechnol.* 7:33. doi: 10.3389/fbioe.2019.00033

Sato, T., Yamada, Y., Ohtani, Y., Mitsui, N., Murasawa, H., and Araki, S. (2001). Efficient production of menaquinone (vitamin K2) by a menadione-resistant mutant of *Bacillus subtilis*. *J. Ind. Microbiol. Biotechnol.* 26, 115–120. doi: 10.1038/sj/jim/7000089

Sauer, M., Branduardi, P., Valli, M., and Porro, D. (2004). Production of L-ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. *Appl. Environ. Microbiol.* 70, 6086–6091. doi: 10.1128/AEM.70.10.6086-6091.2004

Schwalfenberg, G. K. (2017). Vitamins K1 and K2: the emerging group of vitamins required for human health. *J. Nutr. Metab.* 2017:6254836. doi: 10.1155/2017/6254836

Schwechheimer, S. K., Park, E. Y., Revuelta, J. L., Becker, J., and Wittmann, C. (2016). Biotechnology of riboflavin. *Appl. Microbiol. Biotechnol.* 100, 2107–2119. doi: 10.1007/s00253-015-7256-z

Schwarzhans, J.-P., Cholewa, D., Grimm, P., Beshay, U., Risse, J.-M., Friehs, K., et al. (2015). Dependency of the fatty acid composition of Euglena gracilis

on growth phase and culture conditions. *J. Appl. Phycol.* 27, 1389–1399. doi: 10.1007/s10811-014-0458-4

Schyns, G., Potot, S., Geng, Y., Barbosa, T. M., Henriques, A., and Perkins, J. B. (2005). Isolation and characterization of new thiamine-deregulated mutants of *Bacillus subtilis*. *J. Bacteriol.* 187, 8127–8136. doi: 10.1128/JB.187.23.8127-8136.2005

Selvam, S., Ramaian Santhaseela, A., Ganesan, D., Rajasekaran, S., and Jayavelu, T. (2019). Autophagy inhibition by biotin elicits endoplasmic reticulum stress to differentially regulate adipocyte lipid and protein synthesis. *Cell Stress Chaperones* 24, 343–350. doi: 10.1007/s12192-018-00967-9

Serrano-Amatriain, C., Ledesma-Amaro, R., Lopez-Nicolas, R., Ros, G., Jimenez, A., and Revuelta, J. L. (2016). Folic acid production by engineered *Ashbya gossypii*. *Metab. Eng.* 38, 473–482. doi: 10.1016/j.ymben.2016.10.011

Sharma, V., Meganathan, R., and Hudspeth, M. E. (1993). Menaquinone (vitamin K2) biosynthesis: cloning, nucleotide sequence, and expression of the menC gene from *Escherichia coli*. *J. Bacteriol.* 175, 4917–4921. doi: 10.1128/jb.175.15.4917-4921.1993

Shen, B., Zhou, P., Jiao, X., Yao, Z., Ye, L., and Yu, H. (2020). Fermentative production of Vitamin E tocotrienols in *Saccharomyces cerevisiae* under cold-shock-triggered temperature control. *Nat. Commun.* 11:5155. doi: 10.1038/s41467-020-18958-9

Shi, Z., Zhen, S., Wittert, G. A., Yuan, B., Zuo, H., and Taylor, A. W. (2014). Inadequate riboflavin intake and anemia risk in a Chinese population: fve-year follow up of the Jiangsu Nutrition Study. *PLoS One* 9:e88862. doi: 10.1371/journal.pone.0088862

Sivakumar, G., Jeong, K., and Lay, J. O. Jr. (2014). Biomass and RRR-α-tocopherol production in *Stichococcus bacillaris* strain siva2011 in a balloon bioreactor. *Microb. Cell Fact.* 13:79. doi: 10.1186/1475-2859-13-79

Song, J., Liu, H., Wang, L., Dai, J., and Zheng, Z. (2014). Enhanced production of vitamin K2 from *Bacillus subtilis* (natto) by mutation and optimization of the fermentation medium. *Braz. Arch. Biol. Technol.* 57, 606–612. doi: 10.1590/S1516-8913201402126

Streit, W. R., and Entcheva, P. (2003). Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Appl. Microbiol. Biotechnol.* 61, 21–31. doi: 10.1007/s00253-002-1186-2

Sugimoto, T., Kanamasa, S., Kato, T., and Park, E. Y. (2009). Importance of malate synthase in the glyoxylate cycle of *Ashbya gossypii* for the efficient production of riboflavin. *Appl. Microbiol. Biotechnol.* 83, 529–539. doi: 10.1007/s00253-009-1972-1

Sugisawa, T., Miyazaki, T., and Hoshino, T. (2005). Microbial production of L-Ascorbic acid from D-Sorbitol, L-Sorbose, L-Gulose, and L-Sorbosone by *Ketogulonicigenium vulgare* DSM 4025. *Biosci. Biotechnol. Biochem.* 69, 659–662. doi: 10.1271/bbb.69.659

Suter, P. M. (2020). The B-vitamins. *Essential Toxic Trace Elements Vitamins Hum. Health* 2020, 217–239 doi: 10.1016/B978-0-12-805378-2.00017-6

Sych, J. M., Lacroix, C., and Stevens, M. J. A. (2016). *Vitamin B12 – Physiology, Production and Application*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Szterk, A., Zmyslowski, A., and Bus, K. (2018). Identification of cis/trans isomers of menaquinone-7 in food as exemplified by dietary supplements. *Food Chem.* 243, 403–409. doi: 10.1016/j.foodchem.2017.10.001

Taguchi, H., Shibata, T., Duangmanee, C., and Tani, Y. (2014). Menaquinone-4 production by a sulfonamide-resistant mutant of *Flavobacterium sp*. *Agricultural Biol. Chem.* 53, 3017–3023. doi: 10.1080/00021369.1989.10869800

Taketomi, H., Soda, K., and Katsui, G. (1983). Results of screening test in tocopherols in microbial realm. *Vitamins* 57, 133–138.

Tan, T., Zhang, M., and Gao, H. (2003). Ergosterol production by fed-batch fermentation of *Saccharomyces cerevisiae*. *Enzyme Microbial Technol.* 33, 366–370. doi: 10.1016/s0141-0229(03)00132-7

Tanaka, T., Tateno, Y., and Gojobori, T. (2005). Evolution of vitamin B6 (pyridoxine) metabolism by gain and loss of genes. *Mol. Biol. Evol.* 22, 243–250. doi: 10.1093/molbev/msi011

Tani, Y., and Tsumura, H. (1989). Screening for tocopherol-producing microorganisms and α-tocopherol production by *Euglena gracilis* Z. *Agricultural Biol. Chem.* 53, 305–312. doi: 10.1271/bbb1961.53.305

Tannous, C., Booz, G. W., Altara, R., Muhieddine, D. H., and Zouein, F. A. (2020). Nicotinamide adenine dinucleotide: biosynthesis, consumption, and therapeutic role in cardiac diseases. *Acta Physiol.* 231:e13551. doi: 10.1111/apha.13551

Timoshnikov, V. A., Kobzeva, T. V., Polyakov, N. E., and Kontoghiorghes, G. J. (2020). Redox interactions of vitamin C and iron: inhibition of the pro-oxidant activity by deferiprone. *Int. J. Mol. Sci.* 21:3967. doi: 10.3390/ijms21113967

Tokui, M., Kubodera, T., Gomi, K., Yamashita, N., and Nishimura, A. (2011). Construction of a thiamine pyrophosphate high-producing strain of *Aspergillus oryzae* by overexpression of three genes involved in thiamine biosynthesis. *J. Biosci. Bioeng.* 111, 388–390. doi: 10.1016/j.jbiosc.2010.12.011

Van Arsdell, S. W., Perkins, J. B., Yocum, R. R., Luan, L., Howitt, C. L., Chatterjee, N. P., et al. (2005). Removing a bottleneck in the *Bacillus subtilis* biotin pathway: BioA utilizes lysine rather than S-adenosylmethionine as the amino donor in the KAPA-to-DAPA reaction. *Biotechnol. Bioeng.* 91, 75–83. doi: 10.1002/bit.20488

Vandamme, E. J., and Revuelta, J. L. (2016a). "Industrial fermentation of vitamin C," in *Industrial Biotechnology of Vitamins, Biopigments, and Antioxidants*, eds E. J. Vandamme and J. L. Revuelta (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA), 161–192. doi: 10.1002/9783527681754.ch7

Vandamme, E. J., and Revuelta, J. L. (2016b). "Vitamins, biopigments, antioxidants and related compounds: a historical, physiological and (bio)technological perspective," in *Industrial Biotechnology of Vitamins, Biopigments, and Antioxidants*, eds E. J. Vandamme and J. L. Revuelta (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA).

Wang, Z., Liu, Z., Cui, W., and Zhou, Z. (2017). Establishment of bioprocess for synthesis of nicotinamide by recombinant *Escherichia coli* expressing high-molecular-mass nitrile hydratase. *Appl. Biochem. Biotechnol.* 182, 1458–1466. doi: 10.1007/s12010-017-2410-y

Wise, L. A., Wesselink, A. K., Bethea, T. N., Brasky, T. M., Wegienka, G., Harmon, Q., et al. (2021). Intake of lycopene and other carotenoids and incidence of uterine leiomyomata: a prospective ultrasound study. *J. Acad. Nutr. Diet.* 121, 92–104. doi: 10.1016/j.jand.2020.08.013

Wu, W.-J. (2011). Isolation and identification of *Bacillus amyloliquefaciens* BY01 with high productivity of menaquinone for Cheonggukjang production. *J. Korean Soc. Appl. Biol. Chem.* 54, 783–789. doi: 10.3839/jksabc.2011.118

Xia, W., Chen, W., Peng, W. F., and Li, K. T. (2015). Industrial vitamin B12 production by *Pseudomonas denitrificans* using maltose syrup and corn steep liquor as the cost-effective fermentation substrates. *Bioprocess Biosyst. Eng.* 38, 1065–1073. doi: 10.1007/s00449-014-1348-5

Yang, Y., Li, R., Hui, J., Li, L., and Zheng, X. (2021). β-Carotene attenuates LPS-induced rat intestinal inflammation via modulating autophagy and regulating the JAK2/STAT3 and JNK/p38 MAPK signaling pathways. *J. Food Biochem.* 45:e13544. doi: 10.1111/jfbc.13544

Yasutake, Y., Nishioka, T., Imoto, N., and Tamura, T. (2013). A single mutation at the ferredoxin binding site of P450 Vdh enables efficient biocatalytic production of 25-hydroxyvitamin D(3). *Chembiochem* 14, 2284–2291. doi: 10.1002/cbic.201300386

Yoon, S. H., Lee, S. H., Das, A., Ryu, H. K., Jang, H. J., Kim, J. Y., et al. (2009). Combinatorial expression of bacterial whole mevalonate pathway for the production of β-carotene in *E. coli*. *J. Biotechnol.* 140, 218–226. doi: 10.1016/j.jbiotec.2009.01.008

Yuan, P., Cui, S., Liu, Y., Li, J., and Liu, L. (2020). Metabolic engineering for the production of fat-soluble vitamins: advances and perspectives. *Appl. Microbiol. Biotechnol.* 104, 935–951. doi: 10.1007/s00253-019-10157-x

Zeng, W., Guo, L., Xu, S., Chen, J., and Zhou, J. (2020). High-throughput screening technology in industrial biotechnology. *Trends Biotechnol.* 38, 888–906. doi: 10.1016/j.tibtech.2020.01.001

Zeng, Z., Han, N., Liu, C., Buerte, B., Zhou, C., Chen, J., et al. (2020). Functional dissection of HGGT and HPT in barley vitamin E biosynthesis via CRISPR/Cas9-enabled genome editing. *Annals Bot.* 126, 929–942. doi: 10.1093/aob/mcaa115

Zhang, Y., and Showalter, A. M. (2020). CRISPR/Cas9 genome editing technology: a valuable tool for understanding plant cell wall biosynthesis and function. *Front. Plant Sci.* 11:589517. doi: 10.3389/fpls.2020.589517

Zhao, J., Li, Q., Sun, T., Zhu, X., Xu, H., Tang, J., et al. (2013). Engineering central metabolic modules of *Escherichia coli* for improving β-carotene production. *Metab. Eng.* 17, 42–50. doi: 10.1016/j.ymben.2013.02.002

Zhou, J., Chen, J., Duan, L., Zhang, M., Guochun, Q. U., Feng, X. U., et al. (2019). Update on the mechanism of vitamin K2 (MK-7) in preventing and treating

osteoporosis. *Chinese J. Osteoporosis*. 25, 539–545 doi: 10.3969/j.issn.1006-7108.2019.04.023

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 01 July 2021
doi: 10.3389/fbioe.2021.673270

Recent Advances in Heterologous Synthesis Paving Way for Future Green-Modular Bioindustries: A Review With Special Reference to Isoflavonoids

*Moon Sajid, Shane Ramsay Stone and Parwinder Kaur**

UWA School of Agriculture and Environment, University of Western Australia, Perth, WA, Australia

Isoflavonoids are well-known plant secondary metabolites that have gained importance in recent time due to their multiple nutraceutical and pharmaceutical applications. In plants, isoflavonoids play a role in plant defense and can confer the host plant a competitive advantage to survive and flourish under environmental challenges. In animals, isoflavonoids have been found to interact with multiple signaling pathways and have demonstrated estrogenic, antioxidant and anti-oncologic activities *in vivo*. The activity of isoflavonoids in the estrogen pathways is such that the class has also been collectively called phytoestrogens. Over 2,400 isoflavonoids, predominantly from legumes, have been identified so far. The biosynthetic pathways of several key isoflavonoids have been established, and the genes and regulatory components involved in the biosynthesis have been characterized. The biosynthesis and accumulation of isoflavonoids in plants are regulated by multiple complex environmental and genetic factors and interactions. Due to this complexity of secondary metabolism regulation, the export and engineering of isoflavonoid biosynthetic pathways into non-endogenous plants are difficult, and instead, the microorganisms *Saccharomyces cerevisiae* and *Escherichia coli* have been adapted and engineered for heterologous isoflavonoid synthesis. However, the current *ex-planta* production approaches have been limited due to slow enzyme kinetics and traditionally laborious genetic engineering methods and require further optimization and development to address the required titers, reaction rates and yield for commercial application. With recent progress in metabolic engineering and the availability of advanced synthetic biology tools, it is envisaged that highly efficient heterologous hosts will soon be engineered to fulfill the growing market demand.

Keywords: synthetic biology, heterologous synthesis, green bioindustries, isoflavonoids, plant secondary metabolites

OPEN ACCESS

Edited by:
Luan Luong Chu,
Phenikaa University, Vietnam

Reviewed by:
Jingyu Wang,
Westlake Institute for Advanced Study (WIAS), China
Lorenzo Pasotti,
University of Pavia, Italy

***Correspondence:**
Parwinder Kaur
parwinder.kaur@uwa.edu.au

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *27 February 2021*
Accepted: *27 May 2021*
Published: *01 July 2021*

Citation:
Sajid M, Stone SR and Kaur P (2021) Recent Advances in Heterologous Synthesis Paving Way for Future Green-Modular Bioindustries: A Review With Special Reference to Isoflavonoids. Front. Bioeng. Biotechnol. 9:673270. doi: 10.3389/fbioe.2021.673270

INTRODUCTION

Living organisms like plants, fungi and unicellular prokaryotes and eukaryotes produce a myriad of chemicals, broadly classified as primary metabolites and secondary metabolites. Primary metabolites are fundamental compounds of life as they are involved in vital cellular processes such as respiration and photosynthesis. Secondary metabolites are a specialized class of chemicals,

and they are usually produced under specific conditions and give producing organisms an additional advantage to survive, compete or attack other organisms (Nabavi et al., 2020). Secondary metabolites are usually divided into three classes (alkaloids, terpenoids, and phenylpropanoids) based on their chemical structure and precursor primary metabolite. The phenylpropanoid is the largest class of plant secondary metabolites, which play an important role in plant growth and development (Vogt, 2010). The phenylpropanoid pathway has different branches that lead to different groups of compounds including chalcones, flavanones, isoflavonoids, and anthocyanins (Novelli et al., 2019). The structure and function of most of these groups have been well researched and documented in the literature.

The isoflavonoids are a large group of plant secondary metabolites and possess a 3-phenylchroman skeleton, which is biogenetically derived from the 2-phenylchroman skeleton of the parent flavonoid (**Figure 1**). Isoflavonoids are predominantly present in Papilionoideae, a subfamily of Leguminosae (Dixon and Sumner, 2003). More than 2,400 isoflavonoids from over 300 plants have been identified so far (Veitch, 2007, 2009, 2013; Al-Maharik, 2019). Isoflavonoids play multiple roles in host plant, and their role in plant defense and plant–rhizobia relationships is the most significant (Larose et al., 2002). Due to their significance for the host plant, biosynthetic pathways involved in the synthesis and accumulation of many isoflavonoids have been explored. Following that, several attempts have been made to increase the content of isoflavonoids in endogenous as well as in related plants. However, due to the complexity of plant secondary metabolism, no significant improvement has been achieved.

Isoflavonoids have played a distinctive role in the history of disease prevalence across continents. It is believed that the difference in the prevalence of cancer across continents is linked with the preference for soy foods (Ko, 2014). Soybean products are rich in organic chemicals that are structurally similar to 17β-estradiol, a human sex hormone (**Figure 1**). Due to this structural similarity, isoflavonoids play an important role in cellular signaling pathways and control multiple functions in humans and are commonly known as phytoestrogens (Prasad et al., 2010). Together with this, isoflavonoids are commonly used in cosmetics, nutraceuticals, pharmaceuticals, food, and beverage industry; however, the pharmaceutical sector holds the largest share in the market due to potential applications of isoflavonoids in chronic and cardiovascular diseases. The market size of isoflavonoids was over US$ 13.5 billion in 2018 and is estimated to reach US$ 30 billion by 2025 (Ahuja and Mamtani, 2019). Isoflavonoids are presently extracted from plants; however, alternative production platforms are also being explored for their sustainable production to maintain a constant supply in the growing market.

Due to a range of potential applications, the demand for isoflavonoids is growing in recent times. However, the issue of traditional extraction and low yield in plants along with recent climate change and competition to use cultivatable land have questioned their availability for the general public. Therefore, the present paper is aimed to discuss the biosynthetic pathways and potential applications of isoflavonoids (as a subclass of flavonoids) generally but specifically about seven key isoflavonoids: daidzein, formononetin, pisatin, medicarpin, coumestrol, genistein, and biochanin-A with a special focus on their biosynthesis in heterologous hosts.

INDUSTRIAL APPLICATIONS OF ISOFLAVONOIDS

Isoflavonoids are commonly present in low amounts in seeds and roots of the Leguminosae/Fabaceae family including several commonly consumed plants like barley, broccoli, cauliflower, fava beans, lupine, kudzu, and soy (Prasad et al., 2010; **Table 1**). Traces of isoflavonoids are also present in red wine and in other plants like alfalfa, red clover and linseed (Pilsakova et al., 2010). Quite interestingly, isoflavonoids have also been identified from at least 59 non-leguminous plant families (i.e., Iridaceae, Rosaceae, and Liliaceae), as it is commonly believed that isoflavonoids' biosynthetic machinery is not widely distributed in plant families except legumes (Lapčík, 2007).

Role in Plants

With increasing climate and environmental pressures, the potential utilization of isoflavonoids *in planta* to enhance plant resistance against herbivore insects and to improve the interactions of the plant with the rhizobiome has resulted in increased interest and research (Dillon et al., 2017).

Isoflavonoids produce a spectrum of benefits for the host plant (**Figure 2**). Isoflavonoids play an important role in plant defense, as they possess a range of antimicrobial activities (commonly analyzed *in vitro*) (Dixon, 1999). They are famous as plant defensive chemicals and are active against vertebrates, molluscs, herbivorous insects, and microorganisms (Dakora and Phillips, 1996; Nwachukwu et al., 2013). For example, the well-known isoflavonoid pterocarpans, maackiain, and pisatin play an important role as phytoalexins in the interaction between *Nectria haematococca* and the host plant *Pisum sativum* (garden pea) (Wasmann and VanEtten, 1996; Enkerli et al., 1998). Both of these pterocarpans are targets of fungal virulence factors and detoxification enzymes, which indicates their importance for the host plant. Recently, Dillon and colleagues have shown that UV-B-induced accumulation of genistein enhances resistance of field-grown soybean plants against *Anticarsia gemmatalis* neonates (Dillon et al., 2017). A 30% reduction in survival and 45% reduction in mass gain of larvae was documented, and the authors have concluded that UV-B-induced accumulation of isoflavonoids increases the resistance of plants against *A. gemmatalis* (Dillon et al., 2017). An overview of UV-B-based induction of isoflavonoids is described in section "Regulation of Isoflavonoid Biosynthesis in Plants."

Isoflavonoids are not only active inside the cell but also play a beneficial role in the rhizosphere. The role of isoflavonoids in the induction of nodulation genes and as allelopathic agents has also been documented (Dixon, 1999). Daidzein, secreted by soybean roots, acts as a signaling molecule for nodulation and alters the structure and functioning of rhizosphere communities (Okutani et al., 2020). In addition to this, isoflavonoids play a role in the

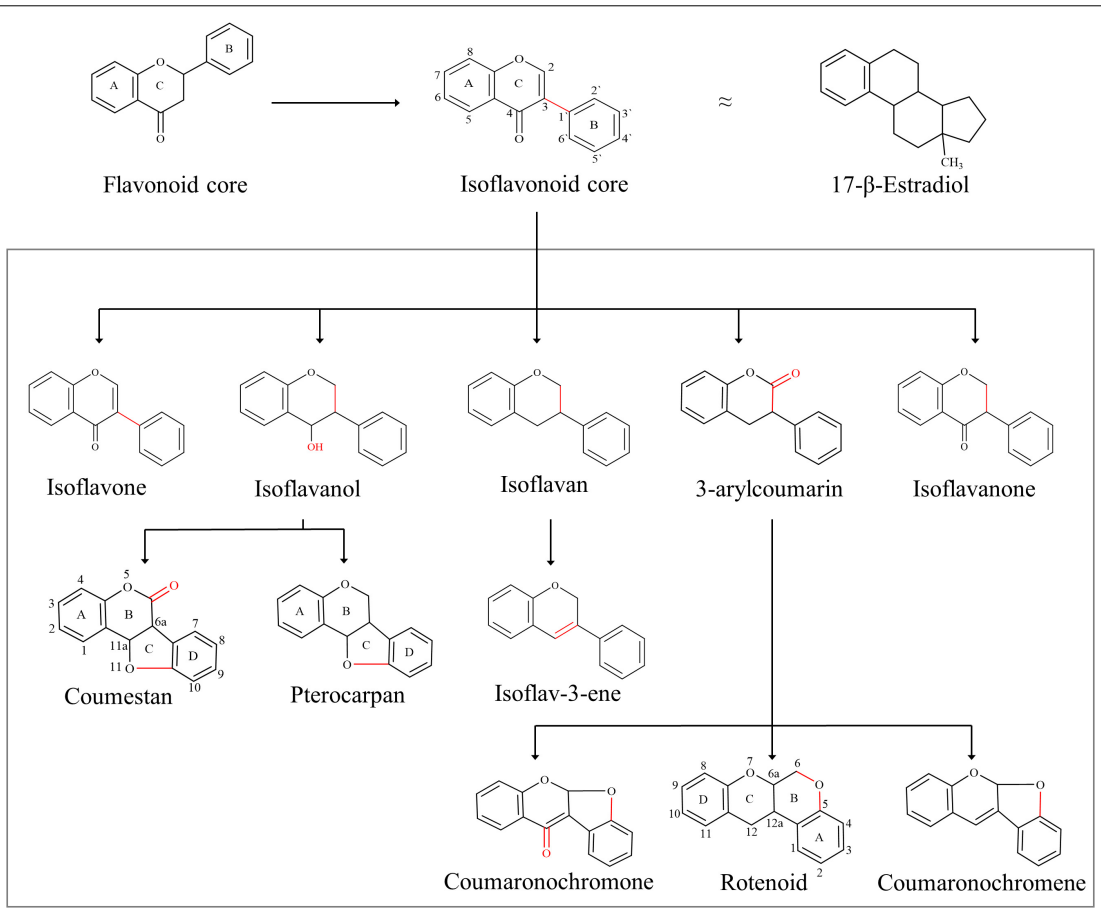


FIGURE 1 | Basic skeleton of isoflavonoids: isoflavonoids are structurally different to flavonoids, with the B-ring migration from position 2 to 3, which in turn leads to the structural similarities to estrogen, e.g., 17β-estradiol. Isoflavonoid diversity is regulated by simple functional additions such as hydroxyl, which in turn can generate additional rings into the backbone, e.g., pterocarpan and coumestan. The addition of ketones can also generate additional rings, for example, rotenoid and coumaronochromone.

induction of transcription of genes involved in the production of the Nod factor. These Nod factors are rhizobial signaling molecules that make plants receptive to symbiotic root infection. Some isoflavonoids are very specific and only induce the production of Nod factors in compatible hosts, thus playing an important role in host selection (Aoki et al., 2000). Isoflavonoids are also involved in developing mutualistic interaction with compatible fungal species. The role of isoflavonoid in initiating the spore germination, hyphal growth and root colonization as well as the formation of arbuscule inside the root cell has been documented (Larose et al., 2002).

Role in Human Health

Several epidemiological studies have shown that an isoflavonoid-rich diet is associated with a low risk of chronic diseases like menopausal, diabetes, cancer, and cardiovascular diseases (Kozłowska and Szostak-Węgierek, 2017). Due to chemical similarity with 17β-estradiol, isoflavonoids can bind with estrogen receptors (ERs) such as ER-α and ER-β (Chen et al., 2018). Due to this affinity, isoflavonoids interfere with cellular signaling mechanisms and play an important role in cellular growth and protection (**Figure 3**). Isoflavonoid aglycones (without glucose) are sometimes biologically more active and available than glycones, as glucose moiety has a strong effect on their function and absorbance in the human gastrointestinal tract (Lee et al., 2018). For example, in its un-glycosylated form, the affinity of genistein is comparable with that of 17β-estradiol, but in its glycosylated forms, its affinity is up to 100–500 times less (Kuiper et al., 1997; Breinholt and Larsen, 1998).

The application of isoflavonoids in human health is a diverse topic, which is not in the scope of the present review. Therefore, in the following paragraphs, a short and precise overview of their health applications has been covered. Interested readers are request to consult recent review papers for further details (Zaheer and Humayoun Akhtar, 2017; Das et al., 2020; Hu et al., 2020; Liu et al., 2020).

Estrogenic Properties

The applications of isoflavonoids as phytoestrogens are one of the most exciting areas of interest in clinical research and nutrition. The declining level of estrogen hormone in aging women is a main cause of osteoporosis, and soy isoflavonoids can substitute

TABLE 1 | Concentration of key isoflavonoids in common food and forage legumes.

Botanical name(common name)	Daidzein (μg/g)	Genistein (μg/g)	FMN (μg/g)	Biochanin-A (μg/g)	CMS (μg/g)	References
Arachis hypogaea(peanut)	0.49	0.82	0.06	0.06	0	Mazur et al., 1998
Cajanus cajan(pigeon pea)	0.14	7.37	318 (leaf)	405 (leaf)	tr	Mazur et al., 1998; Wei et al., 2013
Cicer arietinum(chick pea)	0.11	0.01 (flour)	0.02 (flour)	0.78 (flour)	0.05	Mazur, 1998; Megías et al., 2016
Glycine max(soybean)	327[a]	363[a]	0.40[b]	0.15[b]	1.85[b]	Mazur et al., 1998; Cho et al., 2020
Phaseolus vulgaris(kidney beans)	62[c]	77[c]	0.04	0.26	0.02	Mazur et al., 1998; de Lima et al., 2014
Pisum sativum(split peas)	0.07	0.22	0.04	0.05	tr	Mazur et al., 1998
Vicia faba(fava bean)	51 (stem)	0.48 (stem)	0.06	tr	0	Mazur et al., 1998; Fuentes-Herrera et al., 2020
Vigna mungo(Urd bean)	0.30	0.60	0	0.81	0.09	Mazur et al., 1998
Vigna radiata(mung bean)	26.30 (sprouts)	15.70 (sprouts)	0.07	0.14	tr	Mazur et al., 1998; Silva et al., 2013
Vigna unguiculata(cowpea)	0.30	0.55	0	0	tr	Mazur et al., 1998

Isoflavonoid content was tested in seeds unless specified.
FMN, formononetin; CMS, coumestrol; tr, present in trace amounts.
[a]*Daepung no. 2 variety.*
[b]*Santa Rosa variety.*
[c]*Soja variety.*

the natural estrogen and control bone loss (Chen et al., 2018; Zakłos-Szyda et al., 2020). A meta-analysis has shown that soy isoflavonoid intake for 6 months has a beneficial effect on bone mineral density (BMD), especially on the lumbar spine (Wei et al., 2012). That is why the prevalence of osteoporosis is low in the Chinese and Japanese populations as compared with that in European and Americans. The beneficial role of isoflavonoids is explained due to their molecular similarity with natural estrogen and binding affinity with ERs, especially with ER-β. In addition to this, isoflavonoids intake has positive effects on learning and memory expression, and isoflavonoids are also involved in controlling hot flashes in menopause women (Henderson et al., 2000; Li et al., 2015). Several isoflavonoids including genistein and daidzein as well their derivative have been identified and characterized for their beneficial estrogenic properties.

Antioxidant Properties

The role of isoflavonoids as antioxidants is well established; and it is sometimes believed that the antioxidant potential of isoflavonoids is comparable with that of the well-known antioxidant vitamin E (Djuric et al., 2001). Antioxidant properties are achieved either by the regulation of gene expression of antioxidant enzymes, for example, catalase, or by inhibiting the secondary oxidant production like hydrogen peroxide (Mortensen et al., 2009). Genistein and daidzein are well-known metal chelator and radical scavenger; and due to the presence of three hydroxyl groups, the former is better than the latter (Han et al., 2009). Genistein increases the production of superoxide dismutase (SOD), which scavenges free radicals (Kuriyama et al., 2013). Conversion of daidzein to equol is an important physiological phenomenon, as equol has 100-fold higher ER binding affinity and a greater antioxidant ability than daidzein (Zaheer and Humayoun Akhtar, 2017). Isoflavonoids support disease prevention like type 2 diabetes and health maintenance by repressing the oxidative stress (Umeno et al., 2016). However, it is largely unclear to which extent isoflavonoids mediate antioxidant activities *in vivo*, which is difficult to access, and their antioxidant activities are mostly accessed *in vitro*.

Anticancer Properties

Isoflavones are also well-known for their anticancer potential. Their role in the increase of prostacyclin level, activation of the endothelial level of nitric oxide synthase, inhibition of cell proliferation and DNA synthesis, relaxation of vessels and reduction of plaque, and maintenance of circulatory systems has also been documented (Cano et al., 2010; Tay et al., 2019). Isoflavonoids also interact with epigenetic modifications and are involved in the hypermethylation of tumor suppressor genes, and the underlying mechanism of methylation and acetylation of histones in breast cancer cell lines has also been revealed (Dagdemir et al., 2013).

Genistein, daidzein, formononetin, and, to some extent, coumestrol are well-studied isoflavonoids due to their anticancer potential. Genistein is a strong anticancer candidate isoflavonoid compound with a 50% inhibitory concentration (IC_{50}) value of 37.5 μM against human topoisomerase II (Mizushina et al., 2013). Genistein is known to inhibit protein tyrosine kinase (PTK) and DNA topoisomerases I and II and, therefore, affects a range of cellular activities specifically in carcinogenesis and other neurodegenerative diseases (Kuriyama et al., 2013). Genistein and daidzein can induce cell cycle arrest at S, G_2/M and G_1 stages in various cancer cells (Bossard et al., 2012; Adjakly et al., 2013;

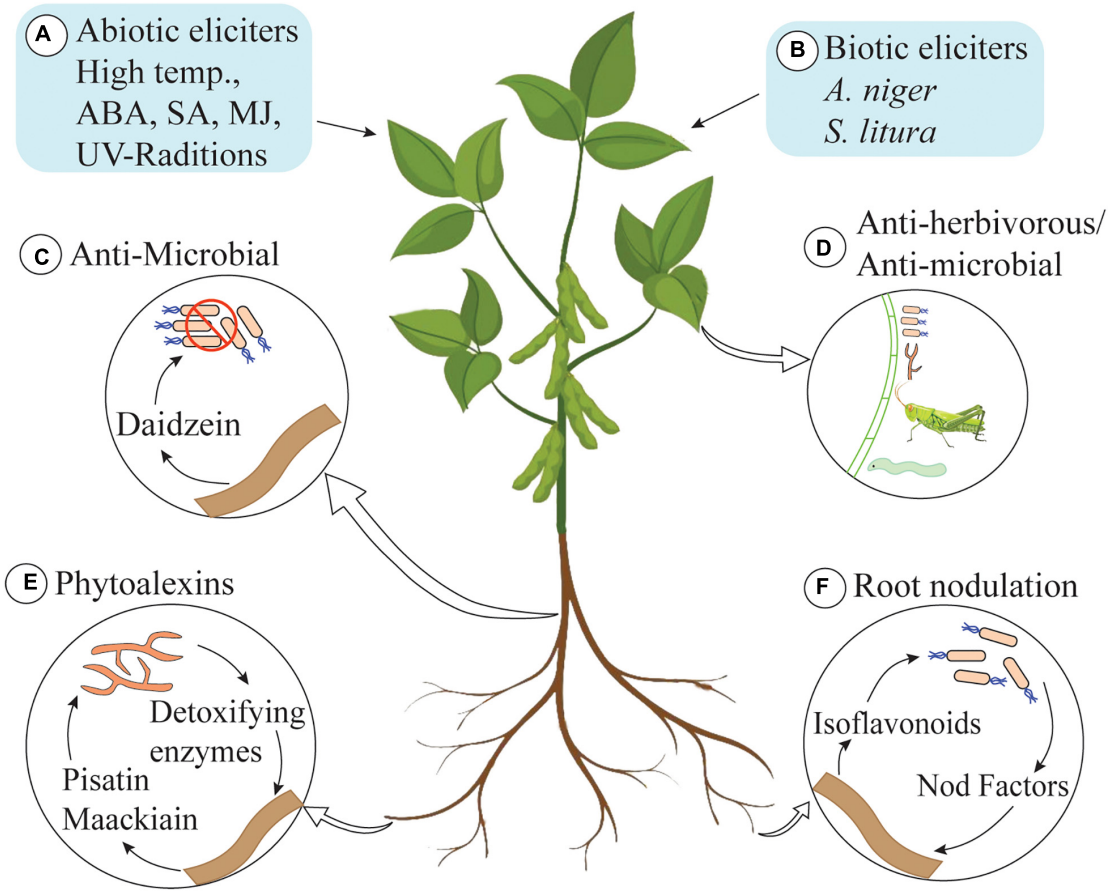


FIGURE 2 | Role of isoflavonoids in plants. **(A)** Abiotic elicitors: high temperature, abscisic acid (ABA), salicylic acid (SA), methyl jasmonate (MJ), and UV-radiations. **(B)** Biotic elicitors: *Aspergillus niger* and *Spodoptera litura* trigger isoflavonoid biosynthesis in plants. **(C)** Daidzein secreted by roots kills bacteria in rhizosphere and shapes the rhizosphere communities. **(D)** Isoflavonoids also possess anti-herbivorous and antimicrobial activities. **(E)** Pisatin and maachiain kill pathogenic fungi in rhizosphere (some fungi produce detoxifying enzymes to degrade pisatin and maachiain to cause infection in roots of host plant). **(F)** Some isoflavonoids induce nodulation genes in receptive bacterial species and play an important role in root-nodule formation.

Shafiee et al., 2016). In a recent study, the authors have concluded that genistein and daidzein can play an important role in cancer cell metastasis, tumorigenesis and stem-like properties, and they have potential as alternative therapies for ovarian cancer patients (Chan et al., 2018).

The IC_{50} value of formononetin ranges from 10 to 300 μM when tested against various cancer cell lines. Formononetin is also able to efficiently inhibit tumor growth *in vivo*, and it is effective against many types of tumors including breast, bone, colon, nasopharyngeal and multiple myeloma cells (Qi et al., 2016; Kim C. et al., 2018; Park et al., 2018). In most of the studies, 1–200 μM (0.3–53.7 μg/ml) of concentration of formononetin was tested, and a variable response has been observed on different cell lines (Tay et al., 2019). Coumestrol is also an important anticancer isoflavonoid candidate molecule with an IC_{50} value of 228 nM tested against casein kinase 2 (CK2) (Liu S. et al., 2013). Selective reduction in CK2 activity has been seen for coumestrol in a dose-dependent manner in various cancer cell lines (Liu S. et al., 2013; Park et al., 2015; Kim et al., 2017). *In silico* modeling has suggested that coumestrol binds to ATP binding pocket of haspin kinase to suppress its activity and results in inhibition of cancer cell proliferation (Kim et al., 2017).

ISOFLAVONOID CHEMISTRY, BIOSYNTHESIS AND REGULATION

Isoflavonoid Chemistry

Isoflavonoids are a diverse and distinctive subclass of flavonoids, and despite their limited distribution in the plant kingdom, isoflavonoids are structurally very diverse (Stobiecki and Kachlicki, 2006; **Figure 1**). The number and complexity of substitution on the basic 3-phenylchroman skeleton along with the different levels of oxidation and the presence of additional heterocyclic rings are responsible for such an outstanding diversity. Isoflavonoids are further divided into several groups, which are shown in **Figure 1** (Dixon and Steele, 1999).

The ring system of isoflavones is derived from two different pathways: the A-ring is derived from the acetate pathway, whereas the B- and C-rings are formed from the shikimate

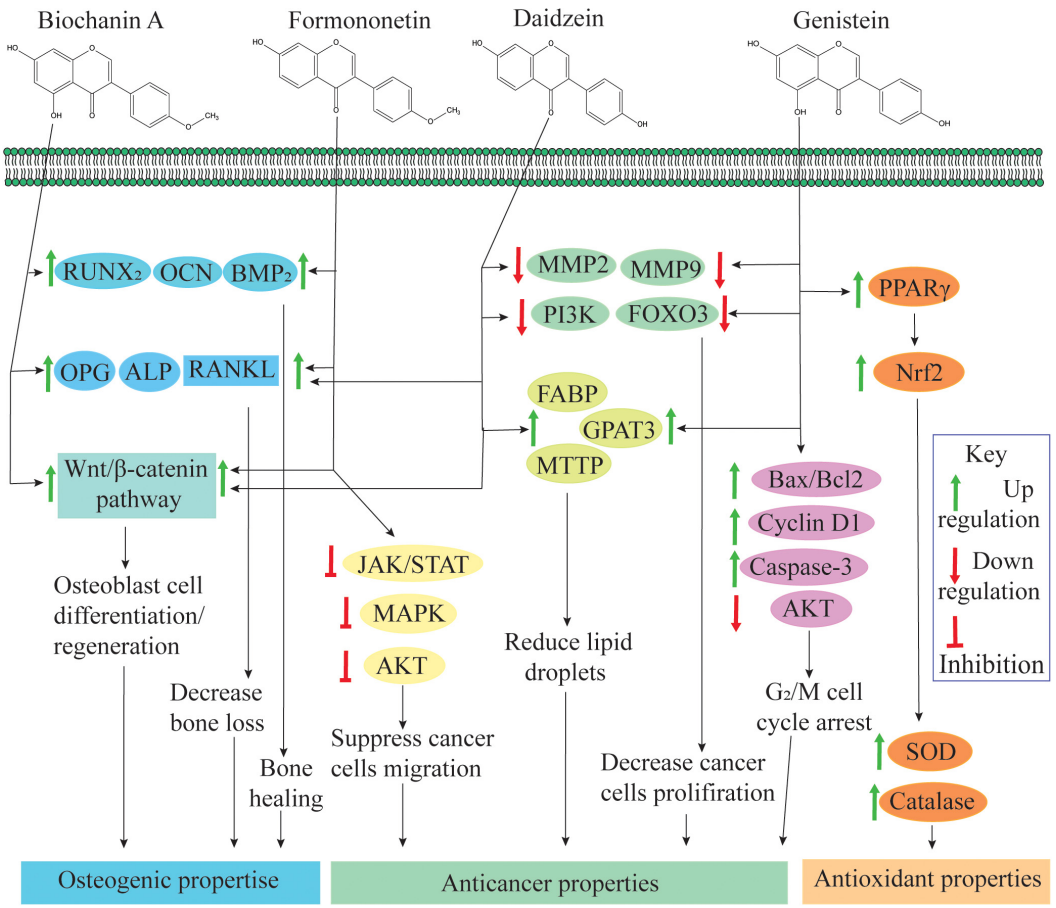


FIGURE 3 | Role of isoflavonoids in human health: biochanin-A and formononetin are involved in bone healing and regeneration effects by upregulating runt-related transcription factor 2 (RUNX2), osteocalcin (OCN) and bone morphogenetic protein 2 (BMP2) expression at the injury site (Singh et al., 2017). Biochanin-A, formononetin, and daidzein also regulate osteoprotegerin (OPG), alkaline phosphatase (ALP), and receptor activator of nuclear factor κβ ligand (RANKL) expression; and these compounds are actively involved in osteogenic activities (Zakłos-Szyda et al., 2020). Additionally, an isoflavone mixture (biochanin-A, formononetin, and daidzein) promotes osteoblast cell differentiation and proliferation through the activation of the Wnt/β-catenin pathway (Chen et al., 2018). Additionally, formononetin inactivates signaling pathways, namely, Janus kinase/signal transducers, and activators of transcription (JAK/STAT) pathway, protein kinase B (PKB or AKT) pathway and mitogen-activated protein kinase (extracellular signal regulated kinase 1/2) [MAPK (ERK1/2)] pathways and suppresses cell migration, invasion and angiogenesis (Qi et al., 2016; Park et al., 2018). Daidzein and genistein can inhibit expression of matrix metalloproteinase-2/9 (MMP2/9), phosphatidylinositol-3-kinase (PI3K) and forkhead box O-3 (FOXO3) and reduce the cancer cell proliferation. Through activation of fatty acid-binding protein (FABP), glycerol-3-phosphate acyltransferase 3 (GPAT3) and microsomal triglyceride transfer protein (MTTP), genistein and daidzein reduce lipid droplet accumulation. Both of these situations induce apoptosis in cancer cells (as reviewed by Hsiao et al., 2020). Genistein activates the expression of Bcl-2-associated X-protein (Bax/Bcl-2), cyclin D1 and caspase-3 pathways and suppressed PI3K/AKT phosphorylation, which results in genistein-induced G2/M cell cycle arrest (Shafiee et al., 2016). Genistein activates peroxisome proliferator-activated receptor gamma (PPARγ) and enhances expression of superoxide dismutase (SOD) and catalase via Nrf2 activation (Chan et al., 2018).

pathway, resulting in a basic C6-C3-C6 skeleton (Ververidis et al., 2007). The basic skeleton is then further decorated with various rounds of glycosylation, methylation and hydroxylation reactions performed by several enzymes and enzyme complexes. This versatile decoration of the basic isoflavonoid skeleton is responsible for the enormous diversity of isoflavonoids. In general, the most common hydroxylation sites of isoflavonoids are 5,7,2′,3′-, and 4′-C, and common C- and O-glycosylation and methylation sites are 6,7, 8-, and 4′-C (Stobiecki and Kachlicki, 2006). Naturally, isoflavonoids are present in a glycosylated form in a plant cell, and the dominant glycosidic form is β-D-glycoside. Other glycosylated forms like 6″-*O*-acetyl-glycoside and 6″-*O*-malonyl-glycoside are also possible; however, the aglycone form is sometimes biologically more active (Ko, 2014). Pterocarpans are an interesting group of isoflavonoids due to the presence of the fourth ring, which is formed due to fusion of 4-C keto group with 6′-C; and due to this fusion, the ring system of pterocarpans is renamed (Whitten et al., 1997).

Isoflavonoid Biosynthesis in Plants

The isoflavonoids are synthesized via the phenylpropanoid pathway, utilizing flavonoid intermediates. First, the aromatic amino acids (phenylalanine) are transformed into the *p*-coumaroyl CoA by a set of three enzymes: phenylalanine ammonia lyase (PAL), *trans*-cinnamate-4-hydroxylase (C4H) and 4-coumaroyl CoA lyase (4CL). Some plant species have a

promiscuous PAL, which is also able to incorporate tyrosine in the pathway (Rosler et al., 1997). The next set of three important enzymes—chalcone synthase (CHS), chalcone isomerase (CHI), and chalcone reductase (CHR)—is responsible for producing naringenin and liquiritigenin from *p*-coumaroyl CoA. A detailed overview of flavonoid biosynthesis has been recently published, and readers are requested to consult Nabavi et al. (2020) and references therein for further details. Both naringenin and liquiritigenin are important flavanones, which are intermediates for various other flavonoid subgroups such as anthocyanins, proanthocyanidins, flavonols, and flavones as well as precursors for isoflavonoids.

Leguminous plants produce isoflavonoids via two different routes, which, however, share many of their chemical reactions and biogenetic machinery. Migration of B-ring from the C-2 position to C-3 position is the first committed and unique step in isoflavonoid biosynthesis, which is catalyzed by isoflavone synthase (IFS), a cytochrome P450 class CYP93C enzyme (Steele et al., 1999; Jung et al., 2000; **Supplementary Table 1**). The immediate product of this reaction is 2,7,4′-trihydroxyisoflavanone, which is an unstable compound and dehydrated to corresponding isoflavanone, i.e., genistein or daidzein, either spontaneously or with the action of another enzyme, 2-hydroxyisoflavanone dehydratase (HIDH) (Steele et al., 1999; Akashi et al., 2005). Daidzein, formononetin, genistein, biochanin-A (isoflavones), pisatin, medicarpin (pterocarpans), and coumestrol (coumestans) are key isoflavonoids that are well-known for their potential pharmaceutical applications (Du et al., 2010). Biosynthesis of key isoflavonoids is discussed in detail in the following paragraphs and shown in **Figure 4** and **Supplementary Figures 1–4**.

Formononetin is an important isoflavone that is synthesized from daidzein is a single-step reaction catalyzed by isoflavanone 4′-*O*-methyltransferase (I4′OMT). I4′OMT, first identified from *Medicago truncatula*, transfer a methyl group from *S*-adenosyl-L-methionine (SAM) to the 4′-C position of daidzein (Liu et al., 2006). Other carbons of the basic skeleton can also be methylated; for example, in alfalfa, 7-C is methylated to produce isoformononetin by isoflavone-7-*O*-methyltransferase (I7OMT) (Zubieta et al., 2001). Sometimes, both 4′-C and 7-C sites are methylated, and the resulting product is known as dimethyldaidzein (Preedy, 2012).

Pisatin is an important phytoalexin that belongs to the pterocarpan group of isoflavonoids; and like coumestans, these compounds have two asymmetric carbons, C-6a and C-11a (Slade et al., 2005). Only dextrorotatory pterocarpans [(+)-pterocarpans] possess the antimicrobial activity and are produced in a few plant species such as peanut (*Arachis hypogaea*) (Strange et al., 1985). Pisatin is the first chemically identified (+)-pterocarpan that is exclusively synthesized by pea (*P. sativum*) (Cruickshank and Perrin, 1960). Starting from formononetin, the first chemical reaction in (+)-pisatin biosynthesis pathway is catalyzed by isoflavone 3′-hydroxylase (I3′H), a P450 class CYP81E9 enzyme that adds OH group at 3′-C of formononetin to form calycosin (**Supplementary Figure 1**). Calycosin is then converted to pseudobaptigenin by the action of pseudobaptigenin synthase (PBS) (Clemens and Barz, 1996; Liu et al., 2003). The next step is the formation of 2′,7-dihydroxy-4′,5′-methylenedioxyisoflavone (DMD) by another P450 class CYP81E1/E7 enzyme known as isoflavone 2′-hydroxylase (I2′H) (Akashi et al., 1998; Liu et al., 2003). The next chemical reaction is catalyzed by isoflavone reductase (IFR), a unique enzyme of the pathway that introduces chirality in pterocarpan biosynthesis that converts DMD to (3*R*)-sophorol (Tiemann et al., 1987; Uchida et al., 2017). The next step to IFR is the formation of (3*R*,4*R*)-2′-hydroxyisoflavanol from (3*R*)-sophorol by the help of 2′-hydroxyisoflavanone 4-reductase (I4′R). Pea I4′R is also known as sophorol reductase (SOR), as it specifically converts (3*R*)-sophorol to (3*R*,4*R*)-7,2′-dihydroxy-4′5′-methylenedioxyisoflavanol [*cis*-(−)-DMDI] (DiCenzo and VanEtten, 2006). RNA-mediated downregulation of IFS and SOR genes in pea resulted in decreased accumulation of (+)-pisatin, which indicates that (+)-pisatin synthesis proceeds through (3*R*)-sophorol and *cis*-(−)-DMDI intermediates (Kaimoyo and VanEtten, 2008). The next step to SOR is the catalyzation of isoflav-3-enes synthase (I3S), identified very recently, which converts *cis*-(−)-DMDI to 7,2′-dihydroxy-4′,5′-methylenedioxyisoflav-3-ene (DMDIF) (Uchida et al., 2020). The enzyme involved in the next step, conversion of achiral DMDIF to (+)-6*a*-hydroxymaackiain (+)-6*a*-(HMK), is not yet identified; however, it is believed that (+)-6*a*-(HMK) is the direct precursor of (+)-pisatin. The final methylation reaction is catalyzed by (+)-6*a*-hydroxymaackiain 3-*O*-methyltransferase (HMM), which converts (+)-6*a*-(HMK) to (+)-pisatin (Wu et al., 1997).

Medicarpin is another important pterocarpan that is formed from formononetin (**Supplementary Figure 2**). I2′H and IFR catalyze the initial chemical reaction in medicarpin biosynthesis (Paiva et al., 1991). I2′H performs oxidation reaction at 2′-C and produces 2′-hydroxy formononetin, which is then reduced to (−)-vestitone by IFR (Tiemann et al., 1987). Two reactions are catalyzed by vestitone reductase (VR) and 7,2′-dihydroxy-4′-methoxyisoflavanol dehydratase (DMID); and (−)-vestitone is reduced and dehydrated to (−)-medicarpin, a major phytoalexin of alfalfa (*Medicago sativa*) (Guo et al., 1994).

Coumestrol belongs to the coumestans group of isoflavonoid; and like pterocarpans, these compounds also have two asymmetric carbons, C-6a and C-11a; however, only *cis*-configurations are sterically possible and present in nature (Whitten et al., 1997). The biosynthetic pathway of coumestrol synthesis is not completely understood; however, few steps have been genetically tested, and the remaining are predicted based on differential gene expression and clustering analysis (Ha et al., 2019; **Supplementary Figure 3**). The pathway starts from daidzein, and the first two chemical reactions involve the conversion of daidzein to 2′-hydroxydaidzein to (3*R*)-2′-hydroxydihydrodaidzein catalyzed by I2′H and IFR, respectively (Dewick et al., 1970; Berlin et al., 1972). Dehydration of (3*R*)-2′-hydroxydihydrodaidzein to 3,9-dihydroxyterocarp-6*a*-en and following chemical reactions up to coumestrol biosynthesis are NAD(P)-dependent redox reactions catalyzed by unidentified NAD(P)-linked oxidoreductases. Overall, seven genes are predicted to be

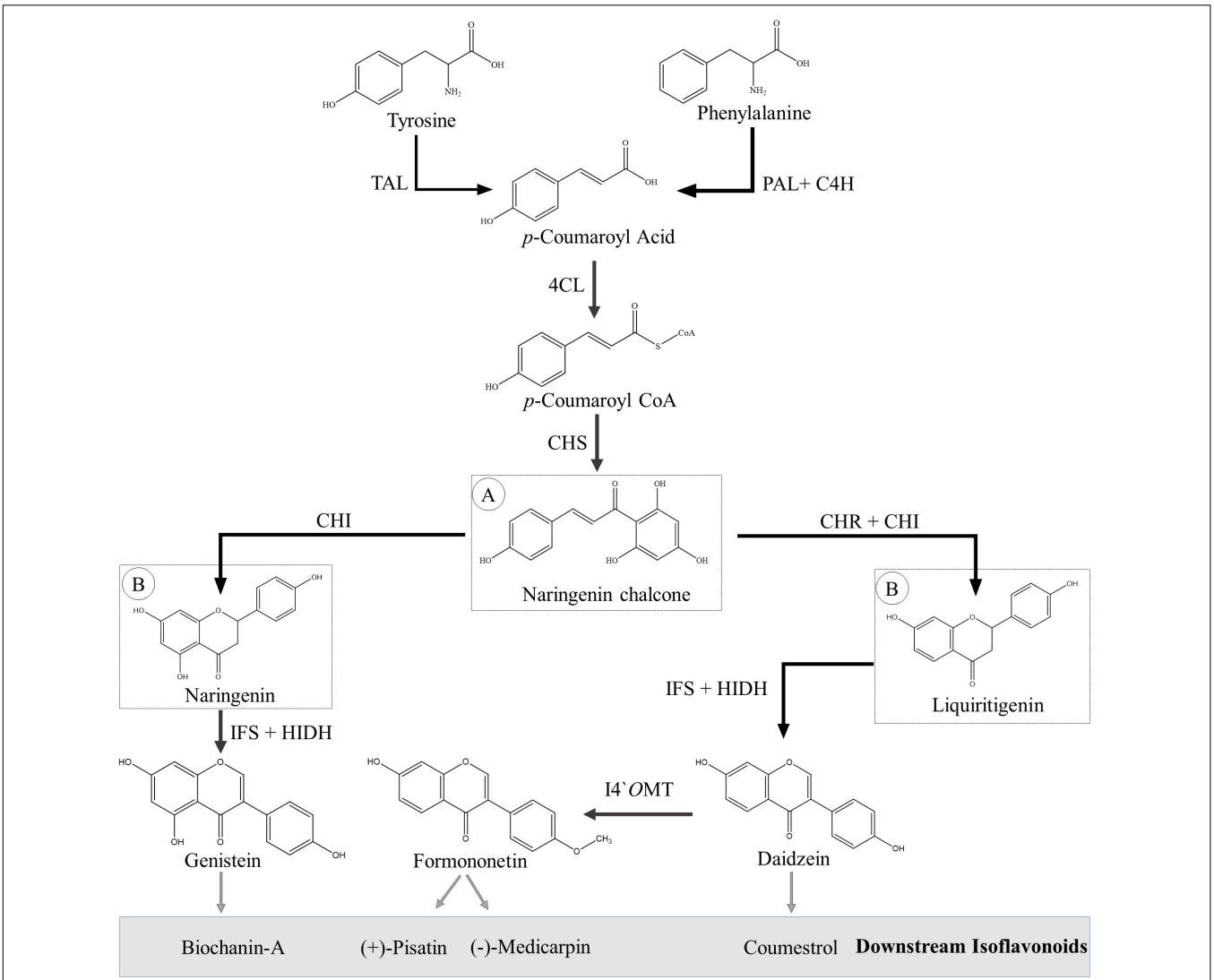


FIGURE 4 | Biosynthesis of basic isoflavonoids from amino acids (phenylalanine and tyrosine). **(A)** Flavonoid precursor molecule. **(B)** Isoflavonoid precursor molecules. Abbreviations, 4CL, 4-coumarate CoA ligase; C4H, *trans*-cinnamate 4-hydroxylase; CHI, chalcone isomerase; CHR, chalcone reductase; CHS, chalcone synthase; HIDH, 2-hydroxyisoflavanone dehydratase; I4′OMT, isoflavone 4′-*O*-methyltransferase; IFS, isoflavone synthase; PAL, phenylalanine ammonia lyase; TAL, tyrosine ammonia lyase.

involved in coumestrol biosynthesis starting from daidzein (Ha et al., 2019). These genes are supposed to encode NAD(P)-linked oxidoreductases, which are responsible for catalyzing NAD(P)-dependent oxidation reactions for coumestrol biosynthesis (Ha et al., 2019).

Genistein and biochanin-A (isoflavones) are also well-known isoflavonoids and are synthesized through the same pathway that starts from naringenin (second route of isoflavonoid biosynthesis) (**Figure 4** and **Supplementary Figure 4**). IFS is responsible for the migration of B-ring from the C-2 position to C-3 position, as mentioned earlier; and the resulting product is 2-hydroxy-2,3-dihydrogenistein (Jung et al., 2000). Like 2,7,4′-trihydroxyisoflavanone, 2-hydroxygenistein is an unstable compound and dehydrated to genistein either spontaneously or with the action of another enzyme, HIDH (Akashi et al., 2005). Genistein is then methylated at the 4′-C position by I4′OMT, and biochanin-A is formed (Liu et al., 2006).

Regulation of Isoflavonoid Biosynthesis in Plants

The isoflavonoid biosynthesis pathways are exceptionally complicated, and the overall accumulation of isoflavonoids inside a plant cell depends not only on pathway-specific enzymes but also on their interaction with other enzymes (Burbulis and Winkel-Shirley, 1999). Genes involved in isoflavone synthesis have shown a functional differentiation, which is because of two recent genome-duplication events: a soybean lineage-specific duplication (13 million years ago) and an early legume duplication (59 million years ago) (Schmutz et al., 2010;

Chu et al., 2014). In the following paragraph, several factors that influence isoflavonoid biosynthesis in the host plant and the underlying molecular mechanisms are discussed.

Abiotic and biotic factors are key regulators of isoflavonoid pathway genes because anything that upregulated the expression of CHS and IFS will have a strong effect on the overall synthesis and accumulation (Dhaubhadel et al., 2003, 2007; Cheng et al., 2008). This is so because CHS and IFS are important enzymes in isoflavonoid biosynthesis, as the farmer directs the C-flow from the phenylpropanoid pathway to flavonoids, and the latter diverts C-flow from the flavonoid pathway to isoflavonoids. In one study, the authors have reported that the accumulation of isoflavonoids and the expression of four key genes (CHS7, CHS8, IFS1, and IFS2) were increased in soybean plants under high temperature stress (Chennupati et al., 2012). However, scientific evidence to explain a direct positive correlation between temperature, gene expression and accumulation of isoflavonoid content is very few. The correlation between the expression of CHS and IFS genes, isoflavonoid accumulation and effect of biotic and abiotic stresses in soybean has also been investigated (Chen et al., 2009; Devi et al., 2020). In biotic factors, *Aspergillus niger* has been tested as an elicitor of isoflavonoid biosynthesis at 0.1% concentration and a nearly fourfold increase in *IFS1* expression, and 5.9-fold increase in isoflavonoid accumulation has been documented as compared with control plants (Devi et al., 2020). Recently, Murakami and colleagues have demonstrated that the content of daidzein and formononetin is increased following the herbivory of *Spodoptera litura* and foliar applications of *S. litura* oral sections (Murakami et al., 2014).

The effect of phytohormones as elicitors for isoflavonoid biosynthesis has also been tested, as plant hormones are extensively studied as signaling molecules involved in defense response to environmental signals in plants (Creelman and Mullet, 1997; Draper, 1997). In one study, effect of salicylic acid (SA) and methyl jasmonate (MJ) was analyzed, and it was noted that SA was more active than MJ. A fivefold increase in the expression of *IFS2*, onefold increase in the expression of *CHS8* and a 4.5-fold increase in isoflavonoid contents were observed as compared with control when SA was applied at the concentration of 10 μM. Similarly, when MJ was applied at the same concentration, a maximum of onefold increase in the expression of *IFS1* and 3.75-fold increase in isoflavonoid accumulation were seen over control (Devi et al., 2020). In another recent study, an overall correlation between abscisic acid (ABA) and UV-B-induced isoflavonoid accumulation and its molecular basis have been investigated. ABA along with guanosine-3′,5′-cyclic monophosphate (cGMP) can upregulate the expression of *CHS* and *IFS* genes and ultimately result in higher accumulation of isoflavonoids under UV-B treatment in soybean. ABA causes inhibition of type 2C protein phosphate (PP2C) and activation of SNF1-related protein kinase (SnRK) and upregulate the expression of *CHS* and *IFS* and finally results in higher accumulation of isoflavonoids in the plant cell (Jiao and Gu, 2019). Overall, the role of the IFS enzyme is more critical than that of CHS, as plants always try to accumulate transcript of *IFS* in higher concentrations than *CHS*. Therefore, it is speculated that either the IFS enzyme has translational regulation or it is enzymatically very slow and has a short lifespan.

Some transcription factors (TFs) of the MYB class involved in isoflavonoid biosynthesis have been identified. The first interesting candidate is R1-type MYB TF GmMYB176, which regulates the expression of *CHS8* and thus controls the overall flavonoid/isoflavonoid contents in plants (Yi et al., 2010). Similarly, the R2R3-type MYB TFs GmMYB39 and GmMYB100 are reported to downregulate the expression of structural genes of the isoflavonoids pathway, thus negatively controlling isoflavonoid biosynthesis in plants. Recently, a soybean TF, GmMYB29, has been characterized for its positive role in isoflavonoid biosynthesis. Comparative genomic analyses have shown that GmMYB29 has maintained the highly conserved R2R3 domain and small amino acid motif in the C-terminal region, which is related to stress resistance in plants (Chu et al., 2017). The expression pattern of GmMYB29 is similar to that of *IFS2*, which supports the hypothesis that the GmMYB29 is a regulator of the *IFS2* gene (Höll et al., 2013). A positive correlation between the expression of GmMYB29 and the accumulation of isoflavonoids in different tissues of plants has also been documented (Liu X. et al., 2013; Yan et al., 2015). It is said that the soybean genome has 4,343 predicted TF encoding genes, which are roughly equal to 6.5% of the total number of genes in the plant (Doerge, 2002). Therefore, it seems that more TF involved in isoflavonoid biosynthesis and regulation will be identified and characterized in the future.

Gene expression is also regulated at the post-transcriptional level, which is largely mediated by two small RNA classes: microRNA (miRNA) and short interfering RNA (siRNA) (Khraiwesh et al., 2010). SiRNA-based regulation of the flavonoid pathway was first reported by Tuteja et al. (2009). The authors have reported tissue-specific (seed coat) silencing of the *CHS* by siRNA in soybean. In *Arabidopsis*, miRNAs like miR156, miR163, miR393 and miR828 are reported to be involved in the regulation of synthesis of secondary metabolites (Bulgakov and Avramenko, 2015; Gupta et al., 2017a). Additionally, miR156-SPL (Squamosa Promoter Binding Protein like) target pair destabilizes WD40-bHLH-MYB TF and negatively regulates anthocyanin biosynthesis in plants (Gou et al., 2011). The role of miRNA in the secondary metabolism of several medicinal plants like *Catharanthus roseus*, *Papaver somniferum,* and *Picrorhiza kurroa* have also been reported (Hao et al., 2012; Pani and Mahapatra, 2013; Boke et al., 2015; Vashisht et al., 2015). As most of the isoflavone biosynthesis and accumulation occur in the developing seed, five new miRNA and their target genes that were predicted to be involved in isoflavonoid biosynthesis have been identified (Gupta et al., 2017b). Interestingly, expression correlation analysis of Gma-miRNA26/28 and their corresponding targets 4CL and I7′*O*MT genes have shown a perfect negative correlation across all stages and genotypes studied so far. It means that decreasing the expression of Gma-miRNA26 has resulted in increase of the expression of 4CL, which could potentially divert the flux toward the synthesis of phenylpropanoid pathways and also resulted in increased accumulation of total isoflavone in the respective plant (Gupta et al., 2017b).

HETEROLOGOUS BIOSYNTHESIS OF ISOFLAVONOIDS

Isoflavonoids are naturally mostly produced in legumes or pea family; however, bulk production of isoflavonoid faces some challenges due to their low content in parent plants. Therefore, alternative cost-effective production platforms are required to meet the growing demand and to ensure availability throughout the year. Model plants like *Nicotiana benthamiana* are considered useful transient expression hosts, as necessary cofactors and substrate pool are likely to be maintained *in planta* (Cravens et al., 2019). However, genetic manipulations, even for model plants, are difficult and slow as compared with microorganisms, and thus, microbial hosts are usually preferred.

Microorganisms are excellent production hosts for plant natural products (PNPs) due to low genetic complexity, ease in genetic manipulation, availability of genetic tool kit and genetic tractability. During the last couple of decades, researchers have engineered artificial isoflavonoid biosynthesis pathways in *Saccharomyces cerevisiae* and *Escherichia coli* (Cress et al., 2013; **Table 2**). At least seven enzymes (PAL/TAL, 4CL, CHS, CHI, CHR, IFS, and IFD) are required for the *de novo* synthesis of parent isoflavonoids: daidzein and genistein. The expression and functionality of plant P450 class enzymes are not optimal in heterologous hosts; therefore, the heterologous synthesis of isoflavonoids is challenging.

Engineering *Saccharomyces cerevisiae* for Isoflavonoid Production

Among the many production chassis available, *S. cerevisiae* is most commonly used for heterologous PNP synthesis, as being eukaryote, it has most of the cellular compartments found in the plant cell. Availability of genetic tool kit, high rate of genetic recombination, ease in genomic manipulation and integration along with generally recognised as safe (GRAS) status are some of the additional benefits for using yeast as a heterologous host (Cravens et al., 2019).

Initial attempts to synthesize isoflavonoids in yeast were focused to convert a flavonoid precursor into the corresponding isoflavonoid. Akashi et al. were the first to express the *IFS* gene from licorice in yeast and have successfully synthesized genistein from naringenin (Akashi et al., 1999). Kim et al. have reported that an engineered yeast strain can produce up to 20 mg/L of genistein when the necessary IFS/CPR is expressed and naringenin is added in the medium (Kim et al., 2005). *De novo* synthesis (construction of the complete pathway in the engineered microbial host) of isoflavonoids has also been achieved (Trantas et al., 2009; Rodriguez et al., 2017). An engineered yeast strain, overexpressing seven heterologous enzymes, was able to produce genistein from different precursors added in the growth medium. The final yield of genistein was 7.7 mg/L (28.5 μM), 0.14 mg/L (0.52 μM), and 0.1 mM (0.4 μM) when 0.5 mM of naringenin, 1 mM of *p*-coumaric acid and 10 mM of phenylalanine were added in the media, respectively (Trantas et al., 2009). An interesting case was the synthesis of quercetin, which was not detected in the medium when phenylalanine was used as a precursor (eight heterologous enzyme reactions), but the engineered strain was able to synthesize 0.26 mg/L (0.9 μM) of quercetin when *p*-coumaric acid was added in the medium (six heterologous enzyme reactions) (Trantas et al., 2009). Recently, with further advancement in knowledge, the synthesis of quercetin has been achieved (eight heterologous enzyme reactions); however, it is speculated that *de novo* synthesis of other isoflavonoids is difficult and requires multiple rounds of genetic engineering (Rodriguez et al., 2017). Employing yeast as a production host for the heterologous synthesis of isoflavonoids has distinctive advantages in the functional expression of plant P450 class enzymes; however, further knowledge and optimization are still needed for the synthesis of key isoflavonoids.

Engineering *Escherichia coli* for Isoflavonoid Production

E. coli has also been extensively used for heterologous biosynthesis of natural products due to the availability of genetic tools, ease in the engineering of native biochemical pathways, simple cultivation techniques and rapid growth (Yang et al., 2020). However, *E. coli* was not the first choice of genetic engineers for the heterologous biosynthesis of isoflavonoids due to issues in the expression of IFS and other plant P450 class enzymes.

The functional expression of IFS was the first bottleneck for the synthesis of isoflavonoids in prokaryotic hosts; therefore, efforts were put forward to express plant P450 enzymes in *E. coli*. Kim et al. (2009) were the first to engineer IFS of red clover by in-frame fusion with a CPR from rice, and the resulting chimeric protein was able to synthesize genistein (up to 15.1 mg/L) from naringenin. In the same year, Leonard et al. have also expressed an engineered chimeric IFS in *E. coli* and have successfully converted naringenin and liquiritigenin into genistein and daidzein, respectively. The *IFS* gene from *Glycine max* and CPR from *C. roseus* were used, and the membrane-spanning regions were replaced with a mammalian leader sequence (Leonard et al., 2009). Co-cultivation of *S. cerevisiae* and *E. coli* strategy was also exploited to boost the final yield of the desired isoflavonoid. In one study, when a naringenin-producing *E. coli* strain was co-cultured with *IFS* expressing *S. cerevisiae* strain, up to 6 mg/L of genistein was detected in the medium (Katsuyama et al., 2007). In recent years, significant progress has been made in heterologous product synthesis, and many PNPs are now being produced and commercialized. Some of these developments are discussed in section "Advance Genetic Engineering Approaches for *ex-planta* Isoflavonoids Biosynthesis."

ADVANCED GENETIC ENGINEERING APPROACHES FOR *EX-PLANTA* ISOFLAVONOID BIOSYNTHESIS

Significant advancement in biotechnology and synthetic biology has been made during the last couple of decades, and heterologous biosynthesis of many PNPs has been successfully

TABLE 2 | Isoflavonoid biosynthesis using heterologous hosts (microorganisms).

Compound	Host strains	Precursors	Titer or productivity	Heterologous enzymes	Approaches	References
Daidzein	*Escherichia coli*	Liquiritigenin	18 mg/g CDW	1	Enzyme engineering for functional expression of IFS in prokaryotic system	Leonard and Koffas, 2007
	Saccharomyces cerevisiae	Liquiritigenin	n.e.	1	Functional analysis of IFS in yeast microsome-based system	Akashi et al., 1999
Daidzin	*S. cerevisiae*	Daidzein	n.e.	1	*In vivo* functional analysis of UDP-glycosyltransferases (UGT)	Li et al., 2014
Ononin	*S. cerevisiae*	Formononetin	n.e.	1	*In vivo* functional analysis of UDP-glycosyltransferases (UGT)	Li et al., 2014
4′–*O*–Methyl daidzein	*E. coli*	Daidzein	102.88 mg/L	2	Methylation of parent compound to improve absorption and bioavailability	Koirala et al., 2019
Daidzein-7-*O*-phosphate	*Bacillus amyloliquefaciens*	Daidzein	n.e.	–	Screening of hosts for efficient biotransformation of daidzein	Kim K.-M. et al., 2018
3′-Hydroxydaidzein	*E. coli*	Daidzein	75 mg/L	1	Screening for candidate enzyme for regioselective hydroxylation	Lee et al., 2014
Genistein	*S. cerevisiae*	Naringenin	n.e.	1	Functional analysis of IFS in yeast microsome-based system	Akashi et al., 1999
	S. cerevisiae	Naringenin	0.87 mg/L	2	Functional expression of IFS and CPR in yeast	Kim et al., 2005
	E. coli	Naringenin	10 mg/g CDW	1	Enzyme engineering for functional expression of IFS in prokaryotic system	Leonard and Koffas, 2007
	E. coli/*S. cerevisiae*	Tyrosine	6 mg/L	5	Co-culturing to achieve higher titer and to manage IFS expression (in yeast)	Katsuyama et al., 2007
	S. cerevisiae	*p*-Coumaroyl *N*-acetylcysteamine	340 μg/L	3	Synthesis of isoflavonoids from modified precursors	Katsuyama et al., 2007
	S. cerevisiae	Phenylalanine	0.1 mg/L	7	Construction of complete pathway for *de novo* synthesis of (iso)flavonoids	Trantas et al., 2009
	E. coli	Naringenin	0.67 mg/L	1	Enzyme engineering for functional expression of IFS in prokaryotic system	Kim et al., 2009
	E. coli	Naringenin	35 mg/L	1	Enzyme engineering for functional expression of IFS in prokaryotic system	Kim, 2020
	E. coli	*p*-Coumaric acid	18.6 mg/L	4	Enzyme engineering for functional expression of IFS in prokaryotic system	Kim, 2020
4′–*O*–Methyl genistein	*E. coli*	Genistein	46.81 mg/L	2	Methylation of parent compound to improve absorption and bioavailability	Koirala et al., 2019

CDW, cell dry weight; n.e., not estimated.

achieved. As discussed in section "Heterologous Biosynthesis of Isoflavonoids," many attempts have also been made to synthesize isoflavonoids in the microbial host, as heterologous synthesis has several advantages over natural and organic synthesis. Some of these are cost-effectiveness, environmentally friendly, on-demand production and being easy to operate, and it is also a desired technology in the green economy initiative. Therefore, in the following section, some advancements in synthetic biology and metabolic engineering are discussed, which will be helpful to produce industrially acceptable titre, rate and yield (TRY) of heterologous synthesis of most of the key isoflavonoids.

Enzyme Engineering Approaches

Most of the enzymes involved in the synthesis of isoflavonoids are yet to be characterized; thus, identification and characterization of specific and highly efficient enzymes are a key requirement

for the successful heterologous biosynthesis of isoflavonoids. Increasing availability of genetic information is helping enzyme identification and characterization efforts, and genome sequencing projects will speed up the process (One Thousand Plant Transcriptomes Initiative, 2019). If genome sequences are available, it becomes very easy to identify or discover new enzymes by comparing plants that produce or lack a specific compound, which will then help us to explain which enzyme might be the candidate for the biosynthesis of that compound. Following the gene identification, approaches like codon optimization, promoter and other regulatory sequence selections and, finally, protein engineering can be used to improve the enzyme activity and expression level (Lee et al., 2016). Therefore, engineering of the heterologous host that can produce an isoflavonoid of interest at industrially acceptable yield is not easy and straightforward.

As mentioned in Subsection "Isoflavonoid biosynthesis in plants," enzymes involved in the synthesis of key isoflavonoids are identified but not well characterized, and suboptimal expression of an enzyme results in bottlenecks in a diversion of C-flux from central metabolism to a final product (**Supplementary Table 1**). Enzyme engineering can help to address some of the issues of low expression and/or low activity of plant-origin enzymes in a heterologous host such as issues of incorrect folding, feedback inhibition and suboptimal pH (Cravens et al., 2019). Enzyme localization by either synthetic or RNA-based scaffolds is a powerful strategy under enzyme engineering approaches. An interesting example in this regard is the use of a synthetic scaffold to cluster a group of three enzymes of the mevalonate pathway, and 77-fold increase in the final product was observed (Dueber et al., 2009). Together with this, protein fusion can effectively reduce pathway competition by bringing the active sites of enzymes near each other and facilitate product channeling. A truncated flavonoid 3′-hydroxylase (tF3′H), a plant P450 class enzymes and a truncated P450 reductase (tCPR) were expressed together as a fusion protein, and successful synthesis of eriodictyol was achieved in *E. coli* (Zhu et al., 2014). Directed evolution is another approach that can help to increase enzyme activity, stability and substrate specificity. Directed evolution becomes even easier if the relationship between enzyme structure, sequence and function are relatively well established.

The chemical space accessed by the heterologous pathway can be expanded to a range of new products by adding additional enzymes in the pathway. In this way, it is easy to transform natural products into their halogenated, hydroxylated, methylated and glycosylated forms by using an enzyme that can accept that natural product as a substrate. Additionally, protein engineering techniques can help to engineer existing enzymes to accept the molecule of interest, as is demonstrated for halogenase enzyme (Payne et al., 2015). The possibility to synthesize novel derivatives is another interesting opportunity. Once a pathway has been established, it is easy to add, replace or remove any pathway enzyme to produce novel derivatives.

Pathway Engineering Approaches

Techniques used to construct metabolic pathway in the heterologous host are extensively reviewed in the literature and includes methods for multigene genomic integration, gene editing and combinatorial enzyme expression (Li et al., 2019). Recent advances in genome engineering make it possible to establish heterologous pathways in previously un-cultivatable organisms. The introduction of the CRISPR/Cas9 system for genome engineering has revolutionized metabolic engineering, as the engineering efficiency of CRISPR/Cas9 sometimes reaches 100% in *E. coli* (Cho et al., 2018). CRISPR/Cas9 variants such as catalytically dead CRISPR-associated protein 9 (dCas9) are even more useful, as they are helpful to control gene expression and to divert metabolic C-flow toward the product of interest. However, approaches to engineer a metabolic pathway in a heterologous host are not yet clear and straightforward.

Modular pathway engineering approaches are helping the researchers to get a higher titer of the product of interest. This approach involves the engineering of a parent strain to increase substrate supply, improvement in the overall flux of the selected pathway and elimination or downregulation of side products. The first interesting example in this regard is the development of a yeast strain for high *p*-coumaric acid production, as it is an important intermediate in the isoflavonoid pathway. Rodriguez and colleagues have developed a yeast strain capable of producing 1.9 g/L of *p*-coumaric acid through a combination of six genetic changes in the central metabolism (Rodriguez et al., 2015). As discussed in section "Heterologous Biosynthesis of Isoflavonoids," the expression of CHS and IFS is not optimal, and to optimize the overall flux of the pathway, researchers have concluded that the optimal copy number for CHS and IFS is 5 and 2, respectively (Lv et al., 2019). Following further optimization of pH and carbon/nitrogen ratio (C/N), the engineered strain was capable of producing 252.4 mg/L of naringenin from glucose in a shake flask culture (Lv et al., 2019). Side product formation, i.e., phloretic acid formation via reduction of *p*-coumaroyl CoA, is also a hurdle in getting a commercially acceptable yield, as it results in metabolic C loss. Researchers have finally identified the enzyme, enoyl reductase Tsc13, responsible for phloretic acid formation. However, Tsc13 is an essential enzyme and cannot be deleted. Therefore, yeast *TSC13* is replaced with a plant *TSC13*, and the unwanted side reaction is eliminated while retaining the natural function of Tsc13 as such (Lehka et al., 2017). Thus, significant progress has been undertaken to engineer a strain able to produce a commercially acceptable yield of an isoflavonoid molecule of interest.

The development of bio-foundries is considered revolutionary progress in pathway construction and validation (Chao et al., 2017). A bio-foundry is a collection of wet lab robotics and software developed to systematize the construction, assembly and testing of pathway parts in the host strain. Until now, bio-foundry-based approaches are used to construct only a short pathway (< 5 enzymes) or already validated pathways (Casini et al., 2018). It is not clear when enzyme discovery and characterization will be automated for long pathways (> 5 enzymatic pathway), as most of the approaches are custom-made up to now; however, the potential is huge, and once optimized, such automation will revolutionize metabolic engineering projects.

Co-culture Approaches

It is now possible to transform and express over 25 genes in a single strain as demonstrated by Li et al. (2018). However, it is demonstrated that when the number of genes required for the biosynthesis of a compound is increased, the performance of a heterologous system is decreased, and strain optimization becomes difficult and laborious (Trantas et al., 2009). To address such issues, co-culture approach, in which two or more organisms expressing complementary genes are grown together, is an interesting and useful strategy.

The co-culture approach (growing two or more organisms having different modules of the same pathway together) can help metabolic engineering in multiple ways. The effectiveness of co-culturing approach is the most obvious in the strain re-optimization process, as it becomes easy to identify and manipulate the issue of heterologous expression in a strain expression in only a part/module of the complete pathway as compared with traditional monoculture approach (Jones and Koffas, 2016). In a traditional monoculture approach, an extension of an established heterologous pathway requires additional genes to be transformed and expressed in the previously engineered strain. This might destabilize the parent strain in a genetic or fermentation perspective and demands further rounds of optimization. Genetic re-optimization is laborious work, and sometimes, it becomes very difficult to regain the previously achieved fluxes (Wu et al., 2016). However, in poly-culture approaches, previous genetic optimizations are preserved on the one hand, and establishment/optimization of a new module into a new strain are easy (as compared with already engineered strain) on the other hand; and finally, few fermentation optimization steps are required to adjust the new strain in co-culture (Li et al., 2018). Together with this, co-culturing can also help to address the issue of the expression of plant P450 class enzymes in prokaryotic hosts, and in this way, the potential of prokaryotic systems can be used for the synthesis of isoflavonoids of interest. For example, a tyrosine-producing *E. coli* and naringenin producing yeast strain (using D-xylose as a C source) were co-cultured for naringenin biosynthesis. The optimized co-culture was able to produce up to 21.16 mg/L of naringenin from simple sugars, and the final titer was eightfolds higher than the monoculture of the engineered yeast strain (Zhang et al., 2017).

Detection and quantification of heterologous products in the microbial host by using genetic biosensors is an exciting area of research and advantages offered by co-culture opportunities have increased their importance (de Frias et al., 2018). For example, a naringenin-sensitive TF FdeR (from *Herbaspirillum seropedicae*) when combined with a green fluorescent protein (GFP) can serve as genetically encoded biosensor and helps in the detection and reporting of intracellular naringenin level (Siedler et al., 2014). Another interesting example is the development of the Systematic Evolution of Ligands by Exponential Enrichment (SELEX) method, which can activate the expression of downstream genes following activation by naringenin. A nearly threefold increase in the expression of reporter genes in *E. coli* strain expressing the riboswitch was noted when 200 mg/L of naringenin was added in the media (Jang et al., 2017). In another study, an RNA riboswitch-based biosensor module was used to control the growth of naringenin-responsive *E. coli* strains grown together in a co-culture system. A positive correlation was seen between the naringenin production (by one strain) and the expression of a reporter gene (in the second strain) (Xiu et al., 2017). The authors have concluded that a naringenin-responsive biosensor has helped the second strain to control metabolic burden, and it also allowed the authors to do module-specific strain optimization comparatively easily and efficiently.

BOX 1 Emerging developments in synthetic biology approaches for heterologous flavonoid biosynthesis.

Current challenges	Emerging developments
Characterization of flavonoid biosynthetic pathways and gene identification	**Genome sequencing projects and machine learning approaches will help in** ✓ Identification of microbial homologous of plant enzymes ✓ Annotation and characterization of genes from plant hosts
Slow enzyme kinetics	**Protein engineering and directed evolution will help in** ✓ Enhancing and optimizing enzyme activity and stability in prokaryotic systems ✓ Engineering binding epitopes for enhanced pathway partner protein interactions
Limitations in metabolic engineering	**CRISPR/Cas9-based genomic integration approaches will help in** ✓ Integration of natural product biosynthetic pathways into non-endogenous microorganisms ✓ Manipulation of microbial chassis ✓ Use of inexpensive microbial feed stock to reduce synthesis cost **Co-culturing approaches will help to** ✓ Manage gene expression, reduce metabolic burden and allow rapid testing of new biosynthetic pathways

CONCLUSION AND FUTURE DIRECTIONS

The isoflavonoids are an important group of plant secondary metabolites that play multiple significant roles in plants as well as in humans. The diversity of isoflavonoids, as well as other PNPs, is a result of an ongoing process of evolution that has generated rich and diverse enzyme sets and will continue to do so in the future. Genome sequencing projects for the identification of candidate genes and TFs will continue to produce a wealth of knowledge for isoflavonoid biosynthesis and engineering in plants.

In plants, the isoflavonoid-mediated natural plant defense mechanisms are a potential tool to address pre-harvest crop losses due to pests and diseases, along with reducing the use of toxic and expensive pesticides. Therefore, an attractive avenue for further research is to engineer beneficial isoflavonoid biosynthesis

pathways into non-endogenous commercial crop plants. However, this approach has technical and social limitations, in terms of the complex interactions and interdependencies of functional genes with added challenges from the public perception and regulatory requirements for the use of genetically modified organisms.

The widespread availability of isoflavonoids for use in agricultural, nutraceutical and pharmaceutical applications is limited due to low yield in plants; therefore, alternative approaches such as *ex-planta* biosynthesis synthetic methods are being investigated. Most of the reports published so far have used *E. coli* and *S. cerevisiae* as heterologous hosts; however, interest in using other available microbial chassis is growing. Therefore, the identification and engineering of enzymes and other genetic components, as well as the exploration of new hosts, will set the direction of future research in the heterologous synthesis of isoflavonoids. Developments in few key areas mentioned in **Box 1** will pave the way of successful biosynthesis of isoflavonoids in the near future.

The successful development of commercially scalable *ex-planta* production platforms for isoflavonoids will have the further benefit of opening exciting avenues toward the biosynthesis and exploitation of alternative PNPs. The realization of these efforts will pave the way toward future economically sustainable, socially beneficial, green-modular bioindustries.

AUTHOR CONTRIBUTIONS

MS, SRS, and PK conceived, designed, and wrote the initial draft of this review article. SS and PK reviewed and edited the manuscript. All authors have read and approved the contents of this manuscript.

ACKNOWLEDGMENTS

MS acknowledges the University of Western Australia for providing international fee scholarship and University Postgraduate Award. PK is supported by the University of Western Australia with additional research funding from Ex Planta Pty Ltd. toward Bio-Synthesis of Isoflavones—Formononetin (FMN) Proof of Concept project. We gratefully acknowledge the computational resources and support from the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.673270/full#supplementary-material

REFERENCES

Adjakly, M., Ngollo, M., Boiteux, J.-P., Bignon, Y.-J., Guy, L., and Bernard-Gallon, D. (2013). Genistein and daidzein: different molecular effects on prostate cancer. *Anticancer Res.* 33, 39–44.

Ahuja, K., and Mamtani, K. (2019). *Industry Analysis Report, Regional Outlook, Application Development Potential, Price Trends, Competitive Market Share & Forecast, 2019–2025*. Selbyville, DE: Global Market Insights.

Akashi, T., Aoki, T., and Ayabe, S. (1998). CYP81E1, a cytochrome P450 cDNA of licorice (*Glycyrrhiza echinata* L.), encodes Isoflavone 2′-hydroxylase. *Biochem. Biophys. Res. Commun.* 251, 67–70. doi: 10.1006/bbrc.1998.9414

Akashi, T., Aoki, T., and Ayabe, S. (1999). Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. *Plant Physiol.* 121, 821–828. doi: 10.1104/pp.121.3.821

Akashi, T., Aoki, T., and Ayabe, S. (2005). Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. *Plant Physiol.* 137, 882–891. doi: 10.1104/pp.104.056747

Al-Maharik, N. (2019). Isolation of naturally occurring novel isoflavonoids: an update. *Nat. Prod. Rep.* 36, 1156–1195. doi: 10.1039/c8np00069g

Aoki, T., Akashi, T., and Ayabe, S. (2000). Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. *J. Plant Res.* 113:475. doi: 10.1007/pl00013958

Berlin, J., Dewick, P. M., Barz, W., and Grisebach, H. (1972). Biosynthesis of coumestrol in *Phaseolus aureus*. *Phytochemistry* 11, 1689–1693. doi: 10.1016/0031-9422(72)85020-9

Boke, H., Ozhuner, E., Turktas, M., Parmaksiz, I., Ozcan, S., and Unver, T. (2015). Regulation of the alkaloid biosynthesis by mi RNA in opium poppy. *Plant Biotechnol. J.* 13, 409–420. doi: 10.1111/pbi.12346

Bossard, C., Busson, M., Vindrieux, D., Gaudin, F., Machelon, V., Brigitte, M., et al. (2012). Potential role of estrogen receptor beta as a tumor suppressor of epithelial ovarian cancer. *PLoS One* 7:e44787. doi: 10.1371/journal.pone.0044787

Breinholt, V., and Larsen, J. C. (1998). Detection of weak estrogenic flavonoids using a recombinant yeast strain and a modified MCF7 cell proliferation assay. *Chem. Res. Toxicol.* 11, 622–629. doi: 10.1021/tx970170y

Bulgakov, V. P., and Avramenko, T. V. (2015). New opportunities for the regulation of secondary metabolism in plants: focus on microRNAs. *Biotechnol. Lett.* 37, 1719–1727. doi: 10.1007/s10529-015-1863-8

Burbulis, I. E., and Winkel-Shirley, B. (1999). Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proc. Natl. Acad. Sci.* 96, 12929–12934. doi: 10.1073/pnas.96.22.12929

Cano, A., García-Pérez, M. Á, and Tarín, J. J. (2010). Isoflavones and cardiovascular disease. *Maturitas* 67, 219–226. doi: 10.1016/j.maturitas.2010.07.015

Casini, A., Chang, F.-Y., Eluere, R., King, A. M., Young, E. M., Dudley, Q. M., et al. (2018). A pressure test to make 10 molecules in 90 days: external evaluation of methods to engineer biology. *J. Am. Chem. Soc.* 140, 4302–4316. doi: 10.1021/jacs.7b13292

Chan, K. K., Siu, M. K., Jiang, Y., Wang, J., Leung, T. H., and Ngan, H. Y. (2018). Estrogen receptor modulators genistein, daidzein and ERB-041 inhibit cell migration, invasion, proliferation and sphere formation via modulation of FAK and PI3K/AKT signaling in ovarian cancer. *Cancer Cell Int.* 18, 1–14. doi: 10.3892/etm.2020.9238

Chao, R., Mishra, S., Si, T., and Zhao, H. (2017). Engineering biological systems using automated biofoundries. *Metab. Eng.* 42, 98–108. doi: 10.1016/j.ymben.2017.06.003

Chen, H., Seguin, P., Archambault, A., Constan, L., and Jabaji, S. (2009). Gene expression and isoflavone concentrations in soybean sprouts treated with chitosan. *Crop Sci.* 49, 224–236. doi: 10.2135/cropsci2007.09.0536

Chen, L., Teng, H., Jia, Z., Battino, M., Miron, A., Yu, Z., et al. (2018). Intracellular signaling pathways of inflammation modulated by dietary flavonoids: the most recent evidence. *Crit. Rev. Food Sci. Nutr.* 58, 2908–2924. doi: 10.1080/10408398.2017.1345853

Cheng, H., Yu, O., and Yu, D. (2008). Polymorphisms of IFS1 and IFS2 gene are associated with isoflavone concentrations in soybean seeds. *Plant Sci.* 175, 505–512. doi: 10.1016/j.plantsci.2008.05.020

Chennupati, P., Seguin, P., Chamoun, R., and Jabaji, S. (2012). Effects of high-temperature stress on soybean isoflavone concentration and expression of key genes involved in isoflavone synthesis. *J. Agric. Food Chem.* 60, 12421–12427. doi: 10.1021/jf3036319

Cho, C.-H., Jung, Y. S., Nam, T. G., Rha, C.-S., Ko, M.-J., Jang, D., et al. (2020). pH-adjusted solvent extraction and reversed-phase HPLC quantification of isoflavones from soybean (*Glycine max* (L.) Merr.). *J. Food Sci.* 85, 673–681. doi: 10.1111/1750-3841.15051

Cho, S., Shin, J., and Cho, B.-K. (2018). Applications of CRISPR/Cas system to bacterial metabolic engineering. *Int. J. Mol. Sci.* 19:1089. doi: 10.3390/ijms19041089

Chu, S., Wang, J., Cheng, H., Yang, Q., and Yu, D. (2014). Evolutionary study of the isoflavonoid pathway based on multiple copies analysis in soybean. *BMC Genet.* 15:76. doi: 10.1186/1471-2156-15-76

Chu, S., Wang, J., Zhu, Y., Liu, S., Zhou, X., Zhang, H., et al. (2017). An R2R3-type MYB transcription factor, GmMYB29, regulates isoflavone biosynthesis in soybean. *PLoS Genet.* 13:e1006770. doi: 10.1371/journal.pgen.1006770

Clemens, S., and Barz, W. (1996). Cytochrome P450-dependent methylenedioxy bridge formation in *Cicer arietinum*. *Phytochemistry* 41, 457–460. doi: 10.1016/0031-9422(95)00618-4

Cravens, A., Payne, J., and Smolke, C. D. (2019). Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nat. Commun.* 10:2142.

Creelman, R. A., and Mullet, J. E. (1997). Biosynthesis and action of jasmonates in plants. *Annu. Rev. Plant Biol.* 48, 355–381. doi: 10.1146/annurev.arplant.48.1.355

Cress, B. F., Linhardt, R. J., and Koffas, M. A. (2013). Isoflavonoid production by genetically engineered microorganisms. *Nat. Prod. Phytochem. Bot. Metab. Alkaloids Phenolics Terpenes* 2, 1647–1681. doi: 10.1007/978-3-642-22144-6_53

Cruickshank, I. A. M., and Perrin, D. R. (1960). Isolation of a phytoalexin from *Pisum sativum* L. *Nature* 187, 799–800. doi: 10.1038/187799b0

Dagdemir, A., Durif, J., Ngollo, M., Bignon, Y.-J., and Bernard-Gallon, D. (2013). Histone lysine trimethylation or acetylation can be modulated by phytoestrogen, estrogen or anti-HDAC in breast cancer cell lines. *Epigenomics* 5, 51–63. doi: 10.2217/epi.12.74

Dakora, F. D., and Phillips, D. A. (1996). Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. *Physiol. Mol. Plant Pathol.* 49, 1–20. doi: 10.1006/pmpp.1996.0035

Das, S., Sharangi, A. B., Egbuna, C., Jeevanandam, J., Ezzat, S. M., Adetunji, C. O., et al. (2020). "Health benefits of isoflavones found exclusively of plants of the fabaceae family," in *Functional Foods and Nutraceuticals*, eds C. Egbuna and G. Dable Tupas (Cham: Springer), 473–508. doi: 10.1007/978-3-030-42319-3_22

de Frias, U. A., Pereira, G. K. B., Guazzaroni, M.-E., and Silva-Rocha, R. (2018). Boosting secondary metabolite production and discovery through the engineering of novel microbial biosensors. *BioMed Res. Int.* 2018:7021826.

de Lima, P. F., Colombo, C. A., Chiorato, A. F., Yamaguchi, L. F., Kato, M. J., and Carbonell, S. A. M. (2014). Occurrence of isoflavonoids in Brazilian common bean germplasm (*Phaseolus vulgaris L.*). *J. Agri. Food Chem.* 62, 9699–9704. doi: 10.1021/jf5033312

Devi, M. A., Kumar, G., and Giridhar, P. (2020). Effect of biotic and abiotic elicitors on isoflavone biosynthesis during seed development and in suspension cultures of soybean (*Glycine max* L.). *3 Biotech* 10:98.

Dewick, P. M., Barz, W., and Grisebach, H. (1970). Biosynthesis of coumestrol in *Phaseolus aureus*. *Phytochemistry* 9, 775–783. doi: 10.1016/s0031-9422(00)85180-8

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of CHS7 and CHS8 genes for isoflavonoid synthesis in soybean seeds. *Plant Physiol.* 143, 326–338. doi: 10.1104/pp.106.086306

Dhaubhadel, S., McGarvey, B. D., Williams, R., and Gijzen, M. (2003). Isoflavonoid biosynthesis and accumulation in developing soybean seeds. *Plant Mol. Biol.* 53, 733–743. doi: 10.1023/b:plan.0000023666.30358.ae

DiCenzo, G. L., and VanEtten, H. D. (2006). Studies on the late steps of (+) pisatin biosynthesis: evidence for (−) enantiomeric intermediates. *Phytochemistry* 67, 675–683. doi: 10.1016/j.phytochem.2005.12.027

Dillon, F. M., Chludil, H. D., and Zavala, J. A. (2017). Solar UV-B radiation modulates chemical defenses against *Anticarsia gemmatalis* larvae in leaves of field-grown soybean. *Phytochemistry* 141, 27–36. doi: 10.1016/j.phytochem.2017.05.006

Dixon, R. A. (1999). "Isoflavonoids: biochemistry, molecular bilolgy, and biological functions," in *Comprehensive Natural Products Chemistry. Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivative*, Vol. 1, ed. U. Sankawa (Amsterdam: Elsevier), 773–823. doi: 10.1016/b978-0-08-091283-7.00030-8

Dixon, R. A., and Steele, C. L. (1999). Flavonoids and isoflavonoids–a gold mine for metabolic engineering. *Trends Plant Sci.* 4, 394–400. doi: 10.1016/s1360-1385(99)01471-5

Dixon, R. A., and Sumner, L. W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. *Plant Physiol.* 131, 878–885. doi: 10.1104/pp.102.017319

Djuric, Z., Chen, G., Doerge, D. R., Heilbrun, L. K., and Kucuk, O. (2001). Effect of soy isoflavone supplementation on markers of oxidative stress in men and women. *Cancer Lett.* 172, 1–6. doi: 10.1016/s0304-3835(01)00627-9

Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* 3, 43–52. doi: 10.1038/nrg703

Draper, J. (1997). Salicylate, superoxide synthesis and cell suicide in plant defence. *Trends Plant Sci.* 2, 162–165. doi: 10.1016/s1360-1385(97)01030-3

Du, H., Huang, Y., and Tang, Y. (2010). Genetic and metabolic engineering of isoflavonoid biosynthesis. *Appl. Microbiol. Biotechnol.* 86, 1293–1312. doi: 10.1007/s00253-010-2512-8

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27:753. doi: 10.1038/nbt.1557

Enkerli, J., Bhatt, G., and Covert, S. F. (1998). Maackiain detoxification contributes to the virulence of *Nectria haematococca* MP VI on chickpea. *Mol. Plant Microbe Interact.* 11, 317–326. doi: 10.1094/mpmi.1998.11.4.317

Fuentes-Herrera, P. B., Delgado-Alvarado, A., Herrera-Cabrera, B. E., Luna-Guevara, M. L., and Olvera-Hernández, J. I. (2020). Quantification of isoflavones in stems of faba bean (*Vicia faba L.*). *Rev. Fac. Cienc. Agrar.* 52, 43–51.

Gou, J.-Y., Felippes, F. F., Liu, C.-J., Weigel, D., and Wang, J.-W. (2011). Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* 23, 1512–1522. doi: 10.1105/tpc.111.084525

Guo, L., Dixon, R. A., and Paiva, N. L. (1994). Conversion of vestitone to medicarpin in alfalfa (*Medicago sativa* L.) is catalyzed by two independent enzymes. Identification, purification, and characterization of vestitone reductase and 7, 2′-dihydroxy-4′-methoxyisoflavanol dehydratase. *J. Biol. Chem.* 269, 22372–22378. doi: 10.1016/s0021-9258(17)31799-4

Gupta, O. P., Karkute, S. G., Banerjee, S., Meena, N. L., and Dahuja, A. (2017a). Contemporary understanding of miRNA-based regulation of secondary metabolites biosynthesis in plants. *Front. Plant Sci.* 8:374. doi: 10.3389/fpls.2017.00374

Gupta, O. P., Nigam, D., Dahuja, A., Kumar, S., Vinutha, T., Sachdev, A., et al. (2017b). Regulation of isoflavone biosynthesis by miRNAs in two contrasting soybean genotypes at different seed developmental stages. *Front. Plant Sci.* 8:567. doi: 10.3389/fpls.2017.00567

Ha, J., Kang, Y.-G., Lee, T., Kim, M., Yoon, M. Y., Lee, E., et al. (2019). Comprehensive RNA sequencing and co-expression network analysis to complete the biosynthetic pathway of coumestrol, a phytoestrogen. *Sci. Rep.* 9:1934.

Han, R.-M., Tian, Y.-X., Liu, Y., Chen, C.-H., Ai, X.-C., Zhang, J.-P., et al. (2009). Comparison of flavonoids and isoflavonoids as antioxidants. *J. Agric. Food Chem.* 57, 3780–3785. doi: 10.1021/jf803850p

Hao, D.-C., Yang, L., Xiao, P.-G., and Liu, M. (2012). Identification of Taxus microRNAs and their targets with high-throughput sequencing and degradome analysis. *Physiol. Plant.* 146, 388–403. doi: 10.1111/j.1399-3054.2012.01668.x

Henderson, V. W., Paganini-Hill, A., Miller, B. L., Elble, R. J., Reyes, P. F., Shoupe, D., et al. (2000). Estrogen for Alzheimer's disease in women: randomized, double-blind, placebo-controlled trial. *Neurology* 54, 295–295.

Höll, J., Vannozzi, A., Czemmel, S., D'Onofrio, C., Walker, A. R., Rausch, T., et al. (2013). The R2R3-MYB transcription factors MYB14 and MYB15 regulate

stilbene biosynthesis in *Vitis vinifera*. *Plant Cell* 25, 4135–4149. doi: 10.1105/tpc.113.117127

Hsiao, Y.-H., Ho, C.-T., and Pan, M.-H. (2020). Bioavailability and health benefits of major isoflavone aglycones and their metabolites. *J. Funct. Foods* 74:104164. doi: 10.1016/j.jff.2020.104164

Hu, C., Wong, W.-T., Wu, R., and Lai, W.-F. (2020). Biochemistry and use of soybean isoflavones in functional food development. *Crit. Rev. Food Sci. Nutr.* 60, 2098–2112. doi: 10.1080/10408398.2019.1630598

One Thousand Plant Transcriptomes Initiative (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574:679. doi: 10.1038/s41586-019-1693-2

Jang, S., Jang, S., Xiu, Y., Kang, T. J., Lee, S.-H., Koffas, M. A., et al. (2017). Development of artificial riboswitches for monitoring of naringenin in vivo. *ACS Synth. Biol.* 6, 2077–2085. doi: 10.1021/acssynbio.7b00128

Jiao, C., and Gu, Z. (2019). Cyclic GMP mediates abscisic acid-stimulated isoflavone synthesis in soybean sprouts. *Food Chem.* 275, 439–445. doi: 10.1016/j.foodchem.2018.09.071

Jones, J. A., and Koffas, M. A. G. (2016). Optimizing metabolic pathways for the improved production of natural products. *Methods Enzymol.* 575, 179–193. doi: 10.1016/bs.mie.2016.02.010

Jung, W., Yu, O., Lau, S.-M. C., O'Keefe, D. P., Odell, J., Fader, G., et al. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nat. Biotechnol.* 18, 208–212. doi: 10.1038/72671

Kaimoyo, E., and VanEtten, H. D. (2008). Inactivation of pea genes by RNAi supports the involvement of two similar O-methyltransferases in the biosynthesis of (+)-pisatin and of chiral intermediates with a configuration opposite that found in (+)-pisatin. *Phytochemistry* 69, 76–87. doi: 10.1016/j.phytochem.2007.06.013

Katsuyama, Y., Miyahisa, I., Funa, N., and Horinouchi, S. (2007). One-pot synthesis of genistein from tyrosine by coincubation of genetically engineered *Escherichia coli* and *Saccharomyces cerevisiae* cells. *Appl. Microbiol. Biotechnol.* 73, 1143–1149. doi: 10.1007/s00253-006-0568-2

Khraiwesh, B., Arif, M. A., Seumel, G. I., Ossowski, S., Weigel, D., Reski, R., et al. (2010). Transcriptional control of gene expression by microRNAs. *Cell* 140, 111–122. doi: 10.1016/j.cell.2009.12.023

Kim, B.-G. (2020). Biological synthesis of genistein in *Escherichia coli*. *J. Microbiol. Biotechnol.* 30, 770–776. doi: 10.4014/jmb.1911.11009

Kim, C., Lee, S.-G., Yang, W. M., Arfuso, F., Um, J.-Y., Kumar, A. P., et al. (2018). Formononetin-induced oxidative stress abrogates the activation of STAT3/5 signaling axis and suppresses the tumor growth in multiple myeloma preclinical model. *Cancer Lett.* 431, 123–141. doi: 10.1016/j.canlet.2018.05.038

Kim, D. H., Kim, B. G., Lee, H. J., Lim, Y., Hur, H. G., and Ahn, J.-H. (2005). Enhancement of isoflavone synthase activity by co-expression of P450 reductase from rice. *Biotechnol. Lett.* 27, 1291–1294. doi: 10.1007/s10529-005-0221-7

Kim, D.-H., Kim, B.-G., Jung, N.-R., and Ahn, J.-H. (2009). Production of genistein from naringenin using *Escherichia coli* containing isoflavone synthase-cytochrome P450 reductase fusion protein. *J. Microbiol. Biotechnol.* 19, 1612–1616. doi: 10.4014/jmb.0905.05043

Kim, J.-E., Lee, S.-Y., Jang, M., Choi, H.-K., Kim, J. H., Chen, H., et al. (2017). Coumestrol epigenetically suppresses cancer cell proliferation: coumestrol is a natural haspin kinase inhibitor. *Int. J. Mol. Sci.* 18:2228. doi: 10.3390/ijms18102228

Kim, K.-M., Park, J.-S., Choi, H., Kim, M.-S., Seo, J.-H., Pandey, R. P., et al. (2018). Biosynthesis of novel daidzein derivatives using *Bacillus* amyloliquefaciens whole cells. *Biocatal. Biotransformation* 36, 469–475. doi: 10.1080/10242422.2018.1461212

Ko, K.-P. (2014). Isoflavones: chemistry, analysis, functions and effects on health and cancer. *Asian Pac. J. Cancer Prev.* 15, 7001–7010. doi: 10.7314/apjcp.2014.15.17.7001

Koirala, N., Pandey, R. P., Thuan, N. H., Ghimire, G. P., Jung, H. J., Oh, T.-J., et al. (2019). Metabolic engineering of *Escherichia coli* for the production of isoflavonoid-4′-O-methoxides and their biological activities. *Biotechnol. Appl. Biochem.* 66, 484–493. doi: 10.1002/bab.1452

Kozłowska, A., and Szostak-Węgierek, D. (2017). "Flavonoids–food sources, health benefits, and mechanisms involved," in *Bioactive Molecules in Food*, eds J.-M. Mérillon and K. G. Ramawat (Cham: Springer), 1–27. doi: 10.1007/978-3-319-54528-8_54-1

Kuiper, G. G., Carlsson, B. O., Grandien, K. A. J., Enmark, E., Häggblad, J., Nilsson, S., et al. (1997). Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors α and β. *Endocrinology* 138, 863–870. doi: 10.1210/endo.138.3.4979

Kuriyama, I., Takahashi, Y., Yoshida, H., and Mizushina, Y. (2013). Inhibitory effect of isoflavones from processed soybeans on human dna topoisomerase II activity. *J. Plant Biochem. Physiol.* 1, 106–112.

Lapčík, O. (2007). Isoflavonoids in non-leguminous taxa: a rarity or a rule? *Phytochemistry* 68, 2909–2916. doi: 10.1016/j.phytochem.2007.08.006

Larose, G., Chênevert, R., Moutoglis, P., Gagné, S., Piché, Y., and Vierheilig, H. (2002). Flavonoid levels in roots of *Medicago sativa* are modulated by the developmental stage of the symbiosis and the root colonizing arbuscular mycorrhizal fungus. *J. Plant Physiol.* 159, 1329–1339. doi: 10.1078/0176-1617-00896

Lee, D.-H., Kim, M. J., Park, S.-H., Song, E.-J., Nam, Y.-D., Ahn, J., et al. (2018). Bioavailability of isoflavone metabolites after korean fermented soybean paste (doenjang) ingestion in estrogen-deficient Rats. *J. Food Sci.* 83, 2212–2221. doi: 10.1111/1750-3841.14214

Lee, H., Kim, B.-G., and Ahn, J.-H. (2014). Production of bioactive hydroxyflavones by using monooxygenase from *Saccharothrix espanaensis*. *J. Biotechnol.* 176, 11–17. doi: 10.1016/j.jbiotec.2014.02.002

Lee, P.-G., Kim, J., Kim, E.-J., Jung, E., Pandey, B. P., and Kim, B.-G. (2016). P212A mutant of dihydrodaidzein reductase enhances (S)-equol production and enantioselectivity in a recombinant *Escherichia coli* whole-cell reaction system. *Appl. Environ. Microbiol.* 82, 1992–2002. doi: 10.1128/aem.03584-15

Lehka, B. J., Eichenberger, M., Bjørn-Yoshimoto, W. E., Vanegas, K. G., Buijs, N., Jensen, N. B., et al. (2017). 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. doi: 10.1093/femsyr/fox004

Leonard, E., and Koffas, M. A. (2007). Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. *Appl. Environ. Microbiol.* 73, 7246–7251. doi: 10.1128/aem.01411-07

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

Li, J., Li, Z., Li, C., Gou, J., and Zhang, Y. (2014). Molecular cloning and characterization of an isoflavone 7-O-glucosyltransferase from *Pueraria lobata*. *Plant Cell Rep.* 33, 1173–1185. doi: 10.1007/s00299-014-1606-7

Li, L., Liu, X., Wei, K., Lu, Y., and Jiang, W. (2019). Synthetic biology approaches for chromosomal integration of genes and pathways in industrial microbial systems. *Biotechnol. Adv.* 37, 730–745. doi: 10.1016/j.biotechadv.2019.04.002

Li, L., Lv, Y., Xu, L., and Zheng, Q. (2015). Quantitative efficacy of soy isoflavones on menopausal hot flashes. *Br. J. Clin. Pharmacol.* 79, 593–604. doi: 10.1111/bcp.12533

Li, Y., Li, S., Thodey, K., Trenchard, I., Cravens, A., and Smolke, C. D. (2018). Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 115, E3922–E3931.

Liu, C.-J., Deavours, B. E., Richard, S. B., Ferrer, J.-L., Blount, J. W., Huhman, D., et al. (2006). Structural basis for dual functionality of isoflavonoid O-methyltransferases in the evolution of plant defense responses. *Plant Cell* 18, 3656–3669. doi: 10.1105/tpc.106.041376

Liu, C.-J., Huhman, D., Sumner, L. W., and Dixon, R. A. (2003). Regiospecific hydroxylation of isoflavones by cytochrome p450 81E enzymes from *Medicago truncatula*. *Plant J.* 36, 471–484. doi: 10.1046/j.1365-313x.2003.01893.x

Liu, L., Chen, X., Hao, L., Zhang, G., Jin, Z., Li, C., et al. (2020). Traditional fermented soybean products: processing, flavor formation, nutritional and biological activities. *Crit. Rev. Food Sci. Nutr.* 13, 1–19. doi: 10.1080/10408398.2020.1848792

Liu, S., Hsieh, D., Yang, Y.-L., Xu, Z., Peto, C., Jablons, D. M., et al. (2013). Coumestrol from the national cancer Institute's natural product library is a novel inhibitor of protein kinase CK2. *BMC Pharmacol. Toxicol.* 14:36. doi: 10.1186/2050-6511-14-36

Liu, X., Yuan, L., Xu, L., Xu, Z., Huang, Y., He, X., et al. (2013). Over-expression of GmMYB39 leads to an inhibition of the isoflavonoid biosynthesis in soybean (*Glycine max*. L). *Plant Biotechnol. Rep.* 7, 445–455. doi: 10.1007/s11816-013-0283-2

Lv, Y., Marsafari, M., Koffas, M., Zhou, J., and Xu, P. (2019). Optimizing oleaginous yeast cell factories for flavonoids and hydroxylated flavonoids biosynthesis. *ACS Synth. Biol.* 8, 2514–2523. doi: 10.1021/acssynbio.9b00193

Mazur, W. (1998). 11 Phytoestrogen content in foods. *Baillieres Clin. Endocrinol. Metab.* 12, 729–742. doi: 10.1016/S0950-351X(98)80013-X

Mazur, W. M., Duke, J. A., Wähälä, K., Rasku, S., and Adlercreutz, H. (1998). Isoflavonoids and lignans in legumes: nutritional and health aspects in humans. *J. Nutr. Biochem.* 9, 193–200. doi: 10.1016/S0955-2863(97)00184-8

Megías, C., Cortés-Giraldo, I., Alaiz, M., Vioque, J., and Girón-Calle, J. (2016). Isoflavones in chickpea (*Cicer arietinum*) protein concentrates. *J. Funct. Foods* 21, 186–192. doi: 10.1016/j.jff.2015.12.012

Mizushina, Y., Shiomi, K., Kuriyama, I., Takahashi, Y., and Yoshida, H. (2013). Inhibitory effects of a major soy isoflavone, genistein, on human DNA topoisomerase II activity and cancer cell proliferation. *Int. J. Oncol.* 43, 1117–1124. doi: 10.3892/ijo.2013.2032

Mortensen, A., Kulling, S. E., Schwartz, H., Rowland, I., Ruefer, C. E., Rimbach, G., et al. (2009). Analytical and compositional aspects of isoflavones in food and their biological effects. *Mol. Nutr. Food Res.* 53, S266–S309.

Murakami, S., Nakata, R., Aboshi, T., Yoshinaga, N., Teraishi, M., Okumoto, Y., et al. (2014). Insect-induced daidzein, formononetin and their conjugates in soybean leaves. *Metabolites* 4, 532–546. doi: 10.3390/metabo4030532

Nabavi, S. M., Šamec, D., Tomczyk, M., Milella, L., Russo, D., Habtemariam, S., et al. (2020). Flavonoid biosynthetic pathways in plants: versatile targets for metabolic engineering. *Biotechnol. Adv.* 38:107316. doi: 10.1016/j.biotechadv.2018.11.005

Novelli, S., Gismondi, A., Di Marco, G., Canuti, L., Nanni, V., and Canini, A. (2019). Plant defense factors involved in *Olea europaea* resistance against *Xylella fastidiosa* infection. *J. Plant Res.* 132, 439–455. doi: 10.1007/s10265-019-01108-8

Nwachukwu, I. D., Luciano, F. B., and Udenigwe, C. C. (2013). The inducible soybean glyceollin phytoalexins with multifunctional health-promoting properties. *Food Res. Int.* 54, 1208–1216. doi: 10.1016/j.foodres.2013.01.024

Okutani, F., Hamamoto, S., Aoki, Y., Nakayasu, M., Nihei, N., Nishimura, T., et al. (2020). Rhizosphere modelling reveals spatiotemporal distribution of daidzein shaping soybean rhizosphere bacterial community. *Plant Cell Environ.* 43, 1036–1046. doi: 10.1111/pce.13708

Paiva, N. L., Edwards, R., Sun, Y., Hrazdina, G., and Dixon, R. A. (1991). Stress responses in alfalfa (*Medicago sativa* L.) 11. Molecular cloning and expression of alfalfa isoflavone reductase, a key enzyme of isoflavonoid phytoalexin biosynthesis. *Plant Mol. Biol.* 17, 653–667. doi: 10.1007/bf00037051

Pani, A., and Mahapatra, R. K. (2013). Computational identification of microRNAs and their targets in *Catharanthus roseus* expressed sequence tags. *Genom. Data* 1, 2–6. doi: 10.1016/j.gdata.2013.06.001

Park, G., Baek, S., Kim, J.-E., Lim, T., Lee, C. C., Yang, H., et al. (2015). Flt3 is a target of coumestrol in protecting against UVB-induced skin photoaging. *Biochem. Pharmacol.* 98, 473–483. doi: 10.1016/j.bcp.2015.08.104

Park, S., Bazer, F. W., Lim, W., and Song, G. (2018). The O-methylated isoflavone, formononetin, inhibits human ovarian cancer cell proliferation by sub G0/G1 cell phase arrest through PI3K/AKT and ERK1/2 inactivation. *J. Cell. Biochem.* 119, 7377–7387. doi: 10.1002/jcb.27041

Payne, J. T., Poor, C. B., and Lewis, J. C. (2015). Directed evolution of RebH for site-selective halogenation of large biologically active molecules. *Angew. Chem.* 127, 4300–4304. doi: 10.1002/ange.201411901

Pilsakova, L., Riecanský, I., and Jagla, F. (2010). The physiological actions of isoflavone phytoestrogens. *Physiol. Res.* 59:651. doi: 10.33549/physiolres.931902

Prasad, S., Phromnoi, K., Yadav, V. R., Chaturvedi, M. M., and Aggarwal, B. B. (2010). Targeting inflammatory pathways by flavonoids for prevention and treatment of cancer. *Planta Med.* 76, 1044–1063. doi: 10.1055/s-0030-1250111

Preedy, V. R. (2012). *Isoflavones: Chemistry, Analysis, Function and Effects*. London: Royal Society of Chemistry.

Qi, C., Xie, M., Liang, J., Li, H., Li, Z., Shi, S., et al. (2016). Formononetin targets the MAPK and PI3K/Akt pathways to induce apoptosis in human nasopharyngeal carcinoma cells in vitro and in vivo. *Int. J. Clin. Exp. Med.* 9, 1180–1189.

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

Rodriguez, A., Strucko, T., Stahlhut, S. G., Kristensen, M., Svenssen, D. K., Forster, J., et al. (2017). Metabolic engineering of yeast for fermentative production of flavonoids. *Bioresour. Technol.* 245, 1645–1654. doi: 10.1016/j.biortech.2017.06.043

Rosler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997). Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiol.* 113, 175–179. doi: 10.1104/pp.113.1.175

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183.

Shafiee, G., Saidijam, M., Tavilani, H., Ghasemkhani, N., and Khodadadi, I. (2016). Genistein induces apoptosis and inhibits proliferation of HT29 colon cancer cells. *Int. J. Mol. Cell. Med.* 5:178.

Siedler, S., Stahlhut, S. G., Malla, S., Maury, J., and Neves, A. R. (2014). Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. *Metab. Eng.* 21, 2–8. doi: 10.1016/j.ymben.2013.10.011

Silva, L. R., Pereira, M. J., Azevedo, J., Gonçalves, R. F., Valentão, P., de Pinho, P. G., et al. (2013). *Glycine max* (L.) Merr., *Vigna radiata* L. and *Medicago sativa* L. sprouts: A natural source of bioactive compounds. *Food Res. Inter.* 50, 167–175. doi: 10.1016/j.foodres.2012.10.025

Singh, K. B., Dixit, M., Dev, K., Maurya, R., and Singh, D. (2017). Formononetin, a methoxy isoflavone, enhances bone regeneration in a mouse model of cortical bone defect. *Br. J. Nutr.* 117, 1511–1522. doi: 10.1017/S0007114517001556

Slade, D., Ferreira, D., and Marais, J. P. (2005). Circular dichroism, a powerful tool for the assessment of absolute configuration of flavonoids. *Phytochemistry* 66, 2177–2215. doi: 10.1016/j.phytochem.2005.02.002

Steele, C. L., Gijzen, M., Qutob, D., and Dixon, R. A. (1999). Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. *Arch. Biochem. Biophys.* 367, 146–150. doi: 10.1006/abbi.1999.1238

Stobiecki, M., and Kachlicki, P. (2006)."Isolation and identification of flavonoids," in *The Science of Flavonoids*, ed. E. Grotewold (Cham: Springer), 47–69. doi: 10.1007/0-387-28822-8_2

Strange, R. N., Ingham, J. L., Cole, D. L., Cavill, M. E., Edwards, C., Cooksey, C. J., et al. (1985). Isolation of the phytoalexin medicarpin from leaflets of *Arachis hypogaea* and related species of the tribe Aeschynomeneae. *Z. Naturforsch. C* 40, 313–316. doi: 10.1515/znc-1985-5-605

Tay, K.-C., Tan, L. T.-H., Chan, C. K., Hong, S. L., Chan, K.-G., Yap, W. H., et al. (2019). Formononetin: a review of its anticancer potentials and mechanisms. *Front. Pharmacol.* 10:820. doi: 10.3389/fphar.2019.00820

Tiemann, K., Hinderer, W., and Barz, W. (1987). Isolation of NADPH: isoflavone oxidoreductase, a new enzyme of pterocarpan phytoalexin biosynthesis in cell suspension cultures of *Cicer arietinum*. *FEBS Lett.* 213, 324–328. doi: 10.1016/0014-5793(87)81515-6

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

Tuteja, J. H., Zabala, G., Varala, K., Hudson, M., and Vodkin, L. O. (2009). Endogenous, tissue-specific short interfering RNAs silence the chalcone synthase gene family in Glycine max seed coats. *Plant Cell* 21, 3063–3077. doi: 10.1105/tpc.109.069856

Uchida, K., Akashi, T., and Aoki, T. (2017). The missing link in leguminous pterocarpan biosynthesis is a dirigent domain-containing protein with isoflavanol dehydratase activity. *Plant Cell Physiol.* 58, 398–408. doi: 10.1093/pcp/pcw213

Uchida, K., Aoki, T., Suzuki, H., and Akashi, T. (2020). Molecular cloning and biochemical characterization of isoflav-3-ene synthase, a key enzyme of the biosyntheses of (+)-pisatin and coumestrol. *Plant Biotechnol.* 37, 301–310. doi: 10.5511/plantbiotechnology.20.0421a

Umeno, A., Horie, M., Murotomi, K., Nakajima, Y., and Yoshida, Y. (2016). Antioxidative and antidiabetic effects of natural polyphenols and isoflavones. *Molecules* 21:708. doi: 10.3390/molecules21060708

Vashisht, I., Mishra, P., Pal, T., Chanumolu, S., Singh, T. R., and Chauhan, R. S. (2015). Mining NGS transcriptomes for miRNAs and dissecting their role in regulating growth, development, and secondary metabolites production in

different organs of a medicinal herb, *Picrorhiza kurroa*. *Planta* 241, 1255–1268. doi: 10.1007/s00425-015-2255-y

Veitch, N. C. (2007). Isoflavonoids of the leguminosae. *Nat. Prod. Rep.* 24, 417–464. doi: 10.1039/b511238a

Veitch, N. C. (2009). Isoflavonoids of the leguminosae. *Nat. Prod. Rep.* 26, 776–802. doi: 10.1039/b616809b

Veitch, N. C. (2013). Isoflavonoids of the leguminosae. *Nat. Prod. Rep.* 30, 988–1027. doi: 10.1039/c3np70024k

Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., and Panopoulos, N. (2007). Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: chemical diversity, impacts on plant biology and human health. *Biotechnol. J. Healthcare Nutr. Technol.* 2, 1214–1234. doi: 10.1002/biot.200700084

Vogt, T. (2010). Phenylpropanoid biosynthesis. *Mol. Plant* 3, 2–20. doi: 10.1093/mp/ssp106

Wasmann, C. C., and VanEtten, H. D. (1996). Transformation-mediated chromosome loss and disruption of a gene for pisatin demethylase decrease the virulence of *Nectria haematococca* on pea. *Mol. Plant Microbe Interact.* 9, 793–803. doi: 10.1094/mpmi-9-0793

Wei, P., Liu, M., Chen, Y., and Chen, D.-C. (2012). Systematic review of soy isoflavone supplements on osteoporosis in women. *Asian Pac. J. Trop. Med.* 5, 243–248. doi: 10.1016/s1995-7645(12)60033-9

Wei, Z., Zu, Y., Fu, Y., Wang, W., Luo, M., Zhao, C., et al. (2013). Ionic liquids-based microwave-assisted extraction of active components from pigeon pea leaves for quantitative analysis. *Sep. Purif. Technol.* 102, 75–81. doi: 10.1016/j.seppur.2012.09.031

Whitten, P. L., Kudo, S., and Okubo, K. K. (1997). "Isoflavonoids," in *Handbook of Plant and Fungal Toxicants*, ed. J. P. Felix D'Mello (Boca Raton, FL: CRC Press), 117–137.

Wu, G., Yan, Q., Jones, J. A., Tang, Y. J., Fong, S. S., and Koffas, M. A. (2016). Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 34, 652–664. doi: 10.1016/j.tibtech.2016.02.010

Wu, Q., Preisig, C. L., and VanEtten, H. D. (1997). Isolation of the cDNAs encoding (+) 6a-hydroxymaackiain 3-O-methyltransferase, the terminal step for the synthesis of the phytoalexin pisatin in *Pisum satvium*. *Plant Mol. Biol.* 35, 551–560.

Xiu, Y., Jang, S., Jones, J. A., Zill, N. A., Linhardt, R. J., Yuan, Q., et al. (2017). Naringenin-responsive riboswitch-based fluorescent biosensor module for *Escherichia coli* co-cultures. *Biotechnol. Bioeng.* 114, 2235–2244. doi: 10.1002/bit.26340

Yan, J., Wang, B., Zhong, Y., Yao, L., Cheng, L., and Wu, T. (2015). The soybean R2R3 MYB transcription factor GmMYB100 negatively regulates plant flavonoid biosynthesis. *Plant Mol. Biol.* 89, 35–48. doi: 10.1007/s11103-015-0349-3

Yang, D., Park, S. Y., Park, Y. S., Eun, H., and Lee, S. Y. (2020). Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol.* 38, 745–765. doi: 10.1016/j.tibtech.2019.11.007

Yi, J., Derynck, M. R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010). A single-repeat MYB transcription factor, GmMYB176, regulates CHS8 gene expression and affects isoflavonoid biosynthesis in soybean. *Plant J.* 62, 1019–1034.

Zaheer, K., and Humayoun Akhtar, M. (2017). An updated review of dietary isoflavones: nutrition, processing, bioavailability and impacts on human health. *Crit. Rev. Food Sci. Nutr.* 57, 1280–1293. doi: 10.1080/10408398.2014.989958

Zakłos-Szyda, M., Budryn, G., Grzelczyk, J., Pérez-Sánchez, H., and Żyżelewicz, D. (2020). Evaluation of isoflavones as bone resorption inhibitors upon interactions with receptor activator of nuclear factor-κB ligand (RANKL). *Molecules* 25:206. doi: 10.3390/molecules25010206

Zhang, W., Liu, H., Li, X., Liu, D., Dong, X.-T., Li, F.-F., et al. (2017). Production of naringenin from D-xylose with co-culture of E. coli and S. cerevisiae. *Eng. Life Sci.* 17, 1021–1029. doi: 10.1002/elsc.201700039

Zhu, S., Wu, J., Du, G., Zhou, J., and Chen, J. (2014). Efficient synthesis of eriodictyol from L-tyrosine in *Escherichia coli*. *Appl. Environ. Microbiol.* 80, 3072–3080. doi: 10.1128/aem.03986-13

Zubieta, C., Dixon, R. A., and Noel, J. P. (2001). Crystal structures of chalcone O-methyltransferase and isoflavone O-methyltransferase reveal the structural basis for substrate specificity in plant O-methyltransferases. *Nat. Struct. Biol.* 8, 271–279.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership