Title

Synthetic biology strategies toward heterologous phytochemical production

Journal

Natural Product Reports, 35(9)

ISSN

0265-0568

Authors

Kotopka, Benjamin J

Li, Yanran

Smolke, Christina D

Publication Date

2018-09-19

DOI

10.1039/c8np00028j

Peer reviewed

Natural Product Reports

REVIEW

View Article Online
View Journal | View Issue

Cite this: *Nat. Prod. Rep.*, 2018, **35**, 902

Synthetic biology strategies toward heterologous phytochemical production

Benjamin J. Kotopka, †[a] Yanran Li †[b] and Christina D. Smolke *[ac]

Covering: 2006 to 2018

Phytochemicals are important sources for the discovery and development of agricultural and pharmaceutical compounds, such as pesticides and medicines. However, these compounds are typically present in low abundance in nature, and the biosynthetic pathways for most phytochemicals are not fully elucidated. Heterologous production of phytochemicals in plant, bacterial, and yeast hosts has been pursued as a potential approach to address sourcing issues associated with many valuable phytochemicals, and more recently has been utilized as a tool to aid in the elucidation of plant biosynthetic pathways. Due to the structural complexity of certain phytochemicals and the associated biosynthetic pathways, reconstitution of plant pathways in heterologous hosts can encounter numerous challenges. Synthetic biology approaches have been developed to address these challenges in areas such as precise control over heterologous gene expression, improving functional expression of heterologous enzymes, and modifying central metabolism to increase the supply of precursor compounds into the pathway. These strategies have been applied to advance plant pathway reconstitution and phytochemical production in a wide variety of heterologous hosts. Here, we review synthetic biology strategies that have been recently applied to advance complex phytochemical production in heterologous hosts.

Received 28th March 2018

DOI: 10.1039/c8np00028j

rsc.li/npr

[a]*Department of Bioengineering, Stanford University, 443 Via Ortega, MC 4245, Stanford, California 94305, USA. E-mail: csmolke@stanford.edu; Fax: +1 650 721 6602; Tel: +1 650 721 6371*

[b]*Department of Chemical and Environmental Engineering, 900 University Avenue, Bourns Hall, Riverside, California 92521, USA*

[c]*Chan Zuckerberg Biohub, San Francisco, CA 94158, USA*

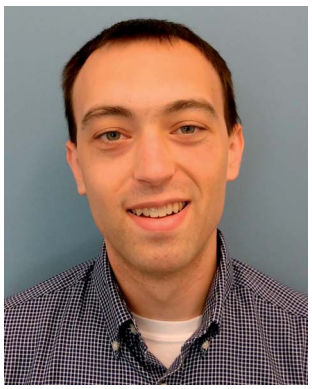
† B. Kotopka and Y. Li contributed equally to this work.

1 Introduction

Phytochemicals are non-nutritive, biologically active compounds synthesized by plants. This group of molecules exhibits a broad spectrum of bioactivities and is important for plant communication, sensing, and survival.[1–5] In addition, phytochemicals are pharmaceutically important, with many used as first-line drugs.[6] Most phytochemicals have complex structures; thus, their chemical syntheses are complicated,

inefficient, and costly. Therefore, phytochemicals are commonly commercially sourced *via* extraction from the natural host plants.[7] However, as the abundance of phytochemicals in the producing plants is often very low, and given the complex metabolite background of plants, efficient isolation of target phytochemicals from native producers can be challenging.[8,9] Plant cell culture of native producers is a potential alternative to the supply of these molecules,[10–12] but can encounter operational challenges such as providing adequate light and mixing while maintaining aseptic conditions and minimizing wall adhesion. In addition, many plant secondary metabolites are mainly produced in specialized differentiated tissues, which makes this approach more challenging.

With advances in recombinant DNA techniques and genome engineering, DNA synthesis, and synthetic biology, heterologous production – that is, transferring the synthetic machinery into a genetically tractable organism – is emerging as a compelling alternative to traditional production of high-value phytochemicals. Numerous examples of heterologous production of complex phytochemicals have been demonstrated in genetically tractable plants, such as *Nicotiana benthamiana*, and microbial hosts, such as *S. cerevisiae* and *Escherichia coli*. In this review, we discuss synthetic biology strategies recently applied to advance production of high-value phytochemicals in heterologous plant, bacteria, and yeast systems, and highlight the advantages and drawbacks of each system.



Benjamin J. Kotopka obtained a Bachelor's degree in molecular biology from Princeton University in 2013. He is currently a PhD student in Prof. Christina Smolke's group at Stanford University. His research centers on applying machine learning techniques to build predictive models of yeast promoter activity, in order to develop artificial promoters for use in metabolic engineering. He also works on characterizing plant biosynthetic gene clusters by reconstruction in yeast.

2 Heterologous phytochemical production in plant systems

The biosynthetic pathways of most phytochemicals are highly complex, including a number of membrane-bound proteins that require special intracellular infrastructure, which can make transferring such synthetic machinery to other organisms challenging. Compared with microbial hosts, heterologous plant hosts exhibit more similar microenvironments to the native producer plant for heterologous protein expression, and have similar core metabolic pathways to native producer plants, often providing a natural supply of chemical precursors for downstream biosynthetic pathways. However, heterologous plant systems, like the native producers, grow slower than microbial hosts, generally lack convenient stable genomic engineering approaches, and exhibit a complex metabolite background – potentially including compounds similar to the authentic target compound – that can lead to challenges in isolating the target phytochemical. Heterologous plant-based systems have also widely been used to investigate plant secondary metabolism and the biological and ecological functions of these phytochemicals. A number of recent reviews have focused on the engineering of complex metabolic pathways in plant hosts.[13,14] In this section, we focus on strategies and

Yanran Li obtained her Bachelor's degrees in Chemistry from Nankai University, and Chemical Engineering from Tianjin University, China, in 2007. From 2008 to 2012, she pursued her PhD degree in Professor Yi Tang's group at UCLA, working on elucidating and engineering the biosynthesis of bacterial and fungal aromatic polyketides. She then spent one year in Prof. Rustem Ismagilov's group at Caltech, using microfluidic devices to investigate microbial interactions. From 2013 to 2016, she was a postdoctoral fellow in Prof. Christina Smolke's group at Stanford, working on the engineered biosynthesis of plant alkaloids in yeast. In 2016, she was appointed as an Assistant Professor in the Department of Chemical and Environmental Engineering at UC Riverside, working on harnessing yeast to investigate plant secondary metabolism.

Christina Smolke is Professor and Associate Chair of Education in the Department of Bioengineering and, by courtesy, Chemical Engineering at Stanford University. Christina's academic research program develops foundational tools that drive transformative advances in our ability to engineer biology. Christina is also Co-Founder and CEO of Antheia, a biotechnology company that leverages advances in synthetic biology, genomics, informatics, and fermentation to transform how we make and discover medicines. Her research program has been recognized with a number of awards, including Chan Zuckerberg Biohub Investigator, Nature's 10, AIMBE College of Fellows, NIH Director's Pioneer Award, WTN Award in Biotechnology, and TR35 Award.

approaches to consider when engineering heterologous plant hosts for the production of phytochemicals.

2.1 Selecting a heterologous plant host for phytochemical production

Most plant hosts are genetically intractable or do not have well-developed molecular biology tools for expressing heterologous genes. Thus, a small number of model plants are often employed for heterologous phytochemical production. The primary model organism for studying plant biology, *Arabidopsis thaliana*, was the first plant to have its genome sequenced[15] and a set of genetic modification approaches have been developed for genetic modification, including nuclear transformation. These approaches were applied to heterologously express the maize terpene synthase *TPS10* in *A. thaliana*, yielding a transgenic plant which produced high levels of sesquiterpene products identical to those produced by maize.[16] The heterologous expression of *TPS10* in *A. thaliana* also facilitated the investigation of the ecological functions of *TPS10* products, by showing that compounds produced by this enzyme were sufficient to help females of the parasitoid *Cotesia marginiventris* locate their lepidopteran hosts.[16]

Tobacco plants of the genus *Nicotiana* are popular organisms for heterologous reconstitution of plant biosynthetic pathways. The tobacco BY-2 cell line, derived from the seedlings of *Nicotiana tabacum*, has been widely utilized in cell suspension culture[17] and as a model organism to study plant molecular biology and physiology.[18] *N. benthamiana* can be successfully infected by a large number of plant viruses, and thus has been employed as a model organism to investigate plant-pathogen interactions.[19] A number of biosynthetic pathways have been reconstituted in *N. tabacum* and *N. benthamiana* for the synthesis of various compounds, including sesquiterpenes,[20–28] diterpenes,[29–31] monoterpenes,[20,32] lignans,[33] glucosinolates,[34] and flavonoids,[35] with reconstituted pathways comprising up to ten heterologous enzymes.[33]

Although it does not share a similar metabolite background with the vascular plants, the moss *Physcomitrella patens*[36] has recently been employed for the heterologous production of sesquiterpenes[37] and diterpenes.[38] *P. patens* can be cultivated with standard plant tissue culturing techniques, and allows efficient and stable genomic integration *via* homologous recombination.[39,40] In addition, *Solanum lycopersicum* (tomato), as a well-studied model organism with comparatively well-established molecular biology techniques, has been utilized for the production of taxadiene,[41] a biosynthetic precursor to the anticancer agent paclitaxel (Taxol), because it contains an abundance of the upstream terpene building block geranylgeranyl pyrophosphate (GGPP).

Non-model plants have also been chosen as heterologous hosts for phytochemical production. One of the major reasons for using non-model plant hosts is the presence of specific upstream substrates in the selected organism. For example, *Artemisia annua* and ginseng (*Panax ginseng*) have been utilized to synthesize taxadiene,[42,43] due to the abundance of terpene precursors in these plants. *A. annua* grows relatively quickly and has an efficient genetic transformation system and ginseng root cultures can be transformed to functionally express heterologous plant enzymes. As another example, the lignan sesamin was synthesized from *Forsythia koreana*,[44,45] which is an abundant natural producer of various other lignans.

2.2 Genetic strategies for modifying multi-gene expression in plants

The introduction of heterologous genes and application of metabolic engineering strategies for enhanced phytochemical production require genetic modification of the plant host. Unlike microorganisms such as *E. coli* and *S. cerevisiae*, genetic engineering in plants is more complicated and time-consuming. *Agrobacterium*-mediated plant transformation of plasmids for expression of heterologous proteins is a common approach that has been established and streamlined in a large number of plants.[46] In one example, a pipeline based on *Agrobacterium*-mediated transient transformation in *N. benthamiana* was established that enabled high-throughput identification of hundreds of terpenoid biosynthetic enzymes within 7 days after transformation with *Agrobacterium*.[47]

While transient expression of single-gene constructs *via* agro-infiltration is well-established, recent advances in synthetic biology have enabled more powerful tools for controlling gene expression in plants. In particular, expressing multiple genes and achieving stable expression through chromosomal integration are facilitated by various multi-gene expression constructs (Fig. 1A) and integration strategies, which have been recently reviewed.[48] These techniques have been applied to synthesis of the sesquiterpene antimalarial drug artemisinin in *N. tabacum*: a "mega-vector" was constructed to stably express five biosynthetic genes and a selection marker, with the expression of each gene controlled by a distinct set of promoters and terminators.[24] Similarly, the final three steps of the benzylglucosinolate pathway were transferred from *Arabidopsis* to tobacco by expressing the three enzymes on a single peptide chain using a 2A polycistronic open reading frame (Fig. 1B), which enables the release of each protein C-terminally fused to the 2A self-cleaving peptide.[34]

In addition to introducing multiple heterologous genes, endogenous genes responsible for undesired side products may be inactivated to enhance the availability of precursors for the downstream biosynthetic enzymes. The inactivation of endogenous genes can be achieved through a number of strategies, including virus-induced gene silencing (Fig. 1C), RNA interference, and agrofiltration.[49] For example, the heterologous production of the sesquiterpene (+)-valencene in *N. benthamiana* was increased 2.8-fold by silencing the expression of the endogenous squalene synthase (SQS) and the 5-*epi*-aristolochene synthase (EAS) using an RNA interference approach.[27] In another study, taxadiene production in tobacco was increased 1.9-fold by enhancing the level of the key precursor GGPP by inactivating the phytoene synthase (*PSY*) enzyme, which uses GGPP as the substrate for carotenoid synthesis, by virus-induced gene silencing.[31]

Tools are also being developed to simplify existing techniques for genetic engineering of plant systems. One such tool

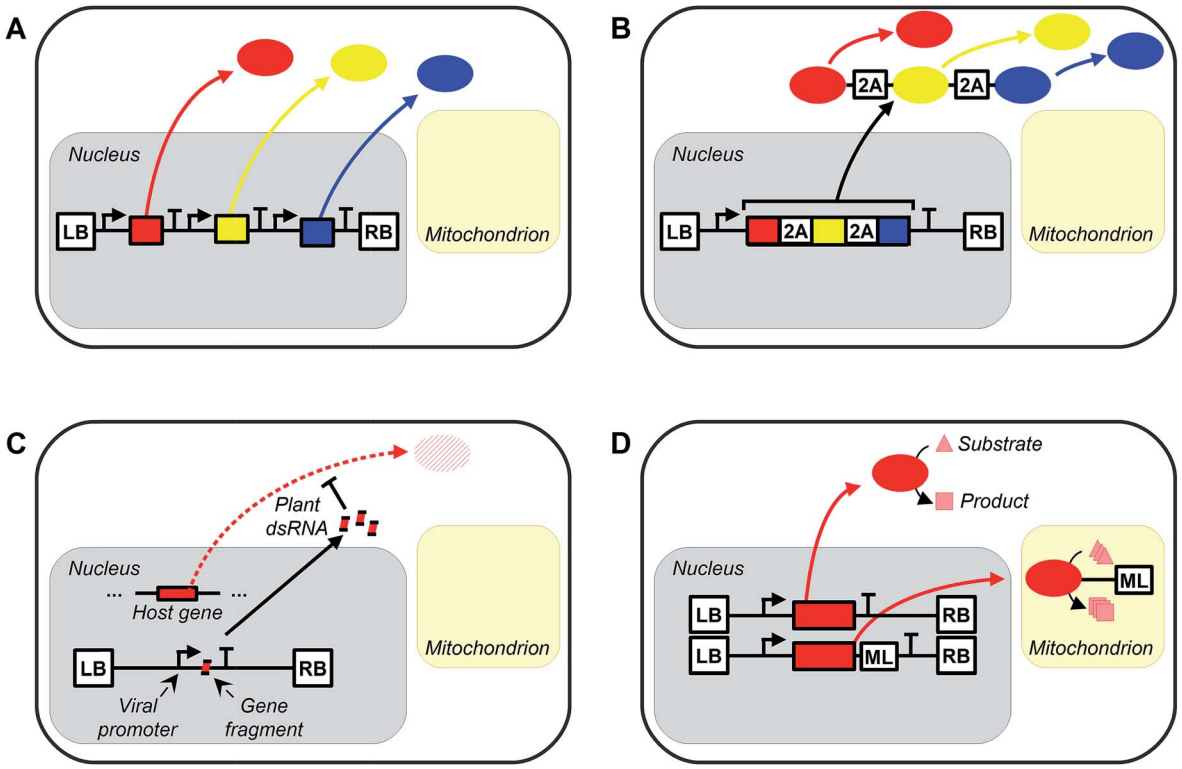


Fig. 1 Methods for gene expression control in heterologous plant systems. A variety of strategies have been developed to induce expression of heterologous genes or repress unwanted expression of host genes. Such methods include: (A) multigene expression constructs; (B) expression of multiple enzymes from a single polypeptide, separated by 2A self-cleaving peptide sequences; (C) RNAi-mediated silencing of host genes by virus-induced gene silencing (VIGS); (D) targeting of enzyme expression to plastid organelles containing a high concentration of enzyme substrates for enhanced efficiency. "LB", "RB": left and right boundary sequences for *Agrobacterium*-mediated expression (respectively); "ML": mitochondrial localization tag.

is a Gateway®-based high-throughput viral vector cloning system from *Narcissus Mosaic Virus* (NMV), which was developed in *N. benthamiana* and shown to enable production of flavonoid compounds, among other applications.[35] In addition, CRISPR/Cas9 genome editing systems have been recently demonstrated in *Arabidopsis* and *N. benthamiana*, and have been shown to enable mutagenesis of phytoene desaturase genes in both systems.[50,51] These tools will likely increase the ease of genome editing and thereby advance heterologous production of phytochemicals in these organisms.

2.3 Subcellular localization as a way to enhance phytochemical production

Plant metabolic pathways are highly compartmentalized, such that different sections of a biosynthetic pathway take place in different subcellular locations.[52] Plants even have two complete sets of terpene biosynthetic pathways: the cytosolic mevalonate (MVA) pathway, predominantly responsible for biosynthesis of C15-derived terpenes, and the plastidic nonmevalonate pathway (referred to either as the methyl-erythritol phosphate (MEP) or the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway), which mainly synthesizes monoterpenes (C10), diterpenes (C20), and carotenoids (C40). Thus, localizing the terpene biosynthetic enzymes in different cellular compartments can alter production levels of the corresponding terpenes (Fig. 1D). Redirecting amorpha-4,11-diene synthase (ADS) and farnesyl diphosphate synthase (FPS) from the cytosol to plastids in *N. tabacum* led to an improved production of the artemisinin precursor amorpha-4,11-diene, of more than 40 000-fold to over 25 μg g^{-1} fresh weight (FW).[22] Similarly, production of the monoterpene limonene was enhanced by up to tenfold through targeting the expression of limonene synthase (LMS) from the cytosol to plastids in *N. tabacum*.[20] This approach has also been applied to larger pathways; a 12-enzyme pathway was introduced into *N. tabacum* chloroplasts, allowing for the synthesis of artemisinic acid at levels of 0.1 mg per gram of fresh plant matter – comparable to the levels seen in the native producer, *Artemisia annua*. However, this engineering also substantially impacted the growth of the transplastomic plants, possibly due to the stress induced by overexpression of a large number of heterologous enzymes.[23]

Heterologous biosynthetic enzymes have been redirected to organelles other than the plastids to achieve phytochemical production. In particular, redirecting the expression of certain terpene biosynthetic enzymes to mitochondria may enhance the production of corresponding terpenes due to higher availability of the terpene substrate in this organelle.[53] The targeted expression of amorphadiene synthase (ADS) in the mitochondria of *N. tabacum* led to a six-fold enhancement in artemisinin

production compared to expression of ADS in the cytosol.[24] In addition, redirecting germacrene A synthase (GAS) from the cytosol to mitochondria resulted in a 15-fold increase in production of the sesquiterpene germacrene A in *N. benthamiana*.[26]

2.4 Pathway reconstitution in heterologous plants as an approach to elucidate biosynthesis schemes

Heterologous reconstitution of phytochemical biosynthesis in *A. thaliana* and tobacco has been widely utilized for the characterization of plant biosynthetic pathways. For example, putative biosynthetic enzymes for the sesquiterpene lactone parthenolide were proposed based on their expression profile in the native producer, feverfew (*Tanacetum parthenium*), characterized initially in yeast, and verified through reconstitution of the biosynthetic pathway in *N. benthamiana*.[25] Similarly, in reconstituting the early steps of the paclitaxel pathway in *Nicotiana sylvestris*, the activity of taxadiene 5α-hydroxylase (CYP725A4) in catalyzing the conversion of taxadiene to 5(12)-oxa-3(11)-cyclotaxane was identified.[30] These approaches were also used to study the biosyntheses of monoterpene indole alkaloids (MIAs) and the lignan podophyllotoxin. Candidate genes were proposed from transcriptome mining of *Catharanthus roseus* and *Podophyllum hexandrum*, the activities of these genes were reconstituted in *N. benthamiana*, and the biosyntheses of the MIA strictosidine and the podophyllotoxin derivative bietoposide aglycone, respectively, were elucidated and achieved.[32,33]

While the relative biological similarity of model plants to the native producers of phytochemicals makes them promising hosts for heterologous production, the historical difficulties of genetic modification in these systems has made progress in this area challenging. However, advances in synthetic biology are enabling more efficient genetic modification strategies and thus novel approaches to plant genetic engineering, which are facilitating projects ranging from individual enzyme characterization by expression in heterologous systems to reconstitution of entire pathways.

3 Bacterial production of phytochemicals

The rapid growth and high target molecule productivity of microbial cultures make microbial production of phytochemicals through metabolic engineering an attractive potential source of these compounds. Despite the evolutionary distance separating bacteria and plants, numerous plant biosynthetic pathways have been reconstituted in bacterial systems. Furthermore, heterologous pathways from non-plant organisms have been adopted to improve the supply of precursor compounds; in particular, while early efforts to synthesize isoprenoid compounds relied on the native nonmevalonate (*e.g.*, methylerythritol 4-phosphate, MEP) pathway to supply isoprenoid precursors,[54,55] recent work in this area generally imports enzymes from the eukaryotic mevalonate (MVA) pathway. The MVA pathway is speculated to outperform the native pathway because its enzymes are not inhibited by the bacterial host's regulatory machinery,[56] even though its theoretical maximum productivity is less than that of the native nonmevalonate pathway.[57] *E. coli*, the workhorse model organism for bacterial synthetic biology, is commonly used in phytochemical biosynthesis research. In addition, photosynthetic cyanobacteria, such as *Synechococcus*, have attracted interest due to their ability to grow in the absence of a carbon feedstock.[58–60]

Synthetic biologists have developed approaches that seek to address challenges associated with heterologous expression of biosynthetic pathways and high productivity. Progress in gene expression control strategies has facilitated optimization of enzyme expression levels. Synthetic biology techniques have also been applied to ensure heterologous plant enzymes are functional once expressed in the microbial environment. Additionally, pathway expression constructs can be optimized through selection and screening workflows enabled by the development of novel biosensors.

3.1 Tuning expression of plant biosynthetic genes in bacteria

Productivity of a bioprocess is often sensitive to enzyme expression level and timing. In some cases, strategies that express all pathway enzymes constitutively at high levels may be sufficient to achieve production goals. However, in other cases expression control techniques supporting graded, inducible, or even feedback-regulated expression may result in a more stable, energy-efficient, targeted design that ultimately leads to increased productivity toward the desired phytochemical product.

Numerous well-characterized promoters are available for achieving high-level constitutive or inducible protein expression in bacteria. The *lac* promoter and its hybrid derivatives, such as the *tac* promoter,[61] allow gene expression to be controlled by addition of the inducer molecule isopropyl β-D-1-thiogalactopyranoside (IPTG). Driving expression of all pathway enzymes with IPTG-inducible promoters is a common initial configuration in pathway engineering efforts. Inducer concentrations can also be titrated to identify an optimum enzyme expression level for phytochemical production. In one example, the arabinose-inducible P_{BAD} promoter was used to modulate expression of the enzymes synthesizing the terpenoid precursor isopentyl pyrophosphate (IPP) and thereby optimize production of the carotenoid lycopene.[55] The T7 bacteriophage promoter, which provides strong protein expression in the presence of the T7 RNA polymerase, is also a commonly used promoter system in metabolic engineering. Multiple expression systems can be combined to optimize enzyme expression levels; for instance, a variety of combinations of *trc* and T7 promoters, as well as high- and low-copy plasmids, were used to optimize expression of a five-enzyme pathway converting tyrosine to the stilbenoid compound resveratrol. This strategy achieved a titer of 35 mg L^{-1}, an almost 30-fold increase over an unoptimized strain.[62] While these common promoter systems are well-validated in *E. coli*, other promoters have been shown to

function better in other bacterial systems – when the pea plant chloroplast *psbA* promoter was used to overexpress spearmint limonene synthase in cyanobacteria, an over-100-fold increase in limonene productivity (to 885.1 μg per L per OD per d) was achieved relative to a system based on the IPTG-inducible *trc* promoter.[60]

Engineered promoters designed to cover a range of expression levels can be used to balance the expression of pathway enzymes to achieve an optimal flux through the pathway. Mutagenesis of a natural promoter has been used to generate families of promoter variants that span a range of expression strengths. In one application of this strategy to production of phytochemicals, mutagenized promoters generated through error-prone PCR of the constitutive bacteriophage P_L-λ promoter were used to optimize the relative production of enzymes in the lycopene pathway in *E. coli*, nearly doubling lycopene titers to over 2.0 mg L^{-1}.[63]

Inducible promoters that are controlled by a pathway-related compound can also be used to impart dynamic feedback regulation over parts of a pathway, which can dynamically balance expression levels and robustly optimize yield over changing conditions during the course of a fermentation. Flux can be dynamically redirected away from toxic side products by using promoters induced by the side product to drive expression of pathway genes (Fig. 2A). In one study illustrating this approach, a natural acetate-inducible promoter, *glnAp2*, was used to increase yields of lycopene by increasing expression of the isoprenoid biosynthetic enzymes isopentenyl-diphosphate isomerase and phosphoenolpyruvate synthase in the presence of acetate.[64] The regulatory strategy dynamically redirected metabolic flux away from the toxic side product acetate and into the desired pathway. Whereas constitutive overexpression of *pps* had previously been shown to reduce growth,[65] the dynamic regulatory strategy did not impart a growth defect to the system.[64] More recently, promoter engineering has been applied in the context of feedback control. Overaccumulation of a metabolic intermediate can be prevented by using a promoter repressed by the intermediate to downregulate the enzyme which produces it, or using promoters induced by the intermediate to upregulate the enzyme catalyzing the next reaction in the pathway (Fig. 2B). This approach has been explored using the *Bacillus subtilis*-derived transcription factor *FapR*. This transcription factor is inhibited by the compound malonyl-CoA, a precursor in biosynthesis of many compounds including the phytochemical triacetic acid lactone.[66] Engineered promoters that are either activated or repressed by malonyl-CoA *via FapR* regulation have been developed and used to regulate production of malonyl-CoA through feedback regulation of enzymes in its biosynthetic pathway.[67–69]

Assuming a promoter responsive to a metabolite of interest is available, implementing an effective feedback control strategy requires thorough characterization of the promoter's strength and sensitivity to the target metabolite. For example, characterization methods for a set of metabolite-sensitive bacterial

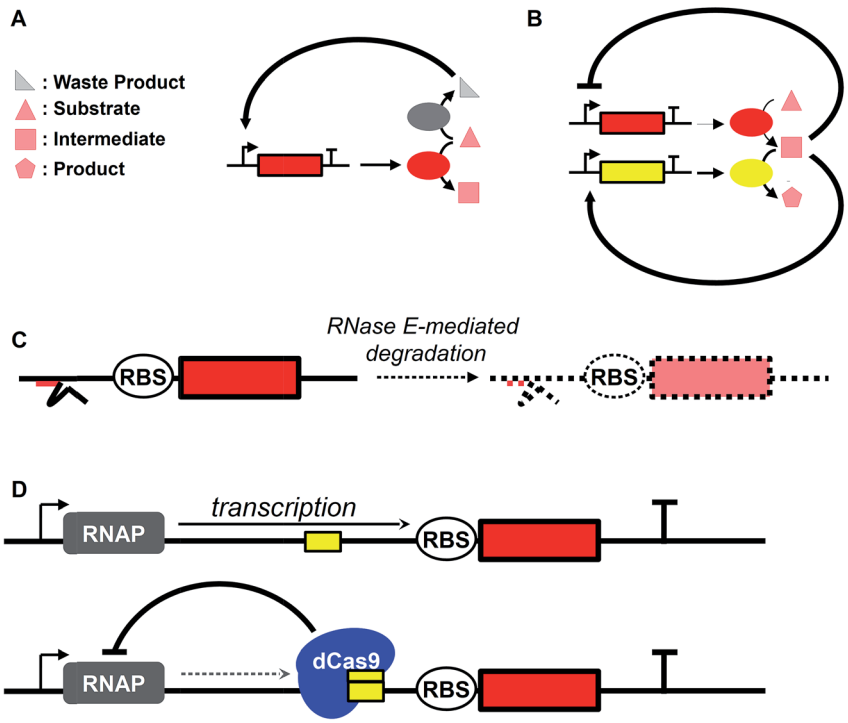


Fig. 2 Synthetic biology tools for expression control and pathway optimization in bacterial systems. Synthetic biology strategies incorporating transcriptional and post-transcriptional control have been applied for phytochemical pathway optimization in bacteria. (A) Metabolite-responsive promoters can be used to reduce levels of side products by increasing expression of desired pathway enzymes in response to side product accumulation. (B) Similarly, levels of pathway intermediates can be moderated by using feedback to control reactions producing and consuming these compounds. (C) sRNA is used to inhibit gene expression at the translational level. (D) dCas9 is used to inhibit transcription by binding a target site in a promoter region (yellow) and obstructing RNA polymerase (RNAP).

promoters, including to the plant flavonoid naringenin,[70] were described by measuring the population distribution of fluorescent protein expression over a range of inducer concentrations in the context of both high- and low-copy plasmids. While the need for a metabolite-responsive promoter can limit the utility of this approach, systems biology methods have been applied to identify natural metabolite-responsive promoters. For instance, a microarray-based approach was used to find promoters responsive to the toxic isoprenoid precursor farnesyl pyrophosphate (FPP). A series of pathway constructs for production of amorpha-4,11-diene were constructed using these promoters to achieve dynamic FPP-responsive control over gene expression. The resulting optimized construct displayed a two-fold increase in amorphadiene titers (to 1.6 g L^{-1}) over similar constructs using inducible or constitutive promoters.[71]

Post-transcriptional regulation strategies for controlling enzyme expression in bacteria have also been demonstrated. The sequence of a ribosome binding site (RBS) can be altered to change the rate of translation initiation and thus enzyme expression levels. Computational tools have been developed to alter the strength of an RBS in a predictable manner, offering a rational strategy for tuning expression strength.[72,73] One study applying these tools to phytochemical production used a set of RBS variants generated using the "RBS Calculator"[72] to predictably alter expression levels of three mevalonate pathway genes, as well as amorphadiene synthase, in an *E. coli* strain producing amorphadiene. This enabled a five-fold increase in amorphadiene production (to 3.6 g L^{-1}) over the original strain.[74] Similarly, a library of RBS variants was used to modulate expression of enzymes in the mevalonate pathway for optimized production of the carotenoid β-carotene, yielding a 51% increase in cell density-normalized product titers.[75]

A number of other synthetic biology devices that modulate gene expression post-transcriptionally have been leveraged for phytochemical production in bacteria. An early example of engineered post-transcriptional control is a family of RNA sequences referred to as TIGRs (Tunable Intergenic Regions), which were designed to modulate RNA stability and translation rates when placed between genes in an operon.[76] In the context of phytochemical production, a set of TIGR variants was used to create a library of mevalonate pathway constructs expressing pathway enzymes at a range of levels. The best designs from the library were shown to reduce levels of the toxic intermediate 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) and boost mevalonate production by a factor of 7 to 6 mM after 24 hours.[76] Additionally, small regulatory RNAs (sRNAs), which inhibit translation and accelerate RNA degradation (Fig. 2C) in a sequence-dependent manner,[77] have been used to reduce the expression of large sets of targeted genes. Using this strategy, combinations of knockdowns of 130 *E. coli* genes were rapidly tested for impact on production of the quinolizidine alkaloid precursor cadaverine. The approach identified that knockdown of *murE*, an essential gene that would not have been identified by a more conventional gene-deletion approach such as screening a deletion library,[78] led to a 55% increase in cadaverine titer (to 2.15 g L^{-1}) relative to a previously reported *E. coli* strain.[79] sRNAs can also be used to finely modulate expression levels of heterologously expressed genes. In another example, a "pool" of sRNAs inducing varying degrees of repression was used to tune the expression of genes in a β-carotene pathway. This approach led to a strain producing 200 mg L^{-1} of β-carotene, a five-fold increase relative to the original, unoptimized strain.[80]

More recently, CRISPR-Cas9-based tools have been used to develop flexible strategies for achieving targeted gene repression. Catalytically dead Cas9 (dCas9) can be targeted to regions of DNA to block transcription initiation (Fig. 2D). The efficiency of knockdown is tunable by changing the position at which dCas9 binds or the degree of base pairing between the Cas9 gRNA and the target DNA.[81] In one application of this technique to phytochemical production, dCas9-based gene repression was used to achieve a 7.4-fold increase in titers of the flavonoid naringenin (to over 400 mg L^{-1}) by identifying genes whose repression increased levels of the precursor compound malonyl-CoA.[82] Additionally, this strategy was used to achieve a 2-fold increase in titers of the anthocyanin peonidin 3-*O*-glucoside (to 51 mg L^{-1}) by reducing production of the SAM-responsive transcriptional repressor *metJ*, which disregulated production of the cofactor *S*-adenosyl methionine.[83] The level of dCas9 repression can also be titrated by modulating the level of dCas9 *via* an inducible promoter. Such an approach was applied to optimizing expression of genes in the mevalonate pathway, resulting in an approximately 7-fold increase in lycopene titers (to approximately 70 mg L^{-1}).[84]

3.2 Strategies for functional expression of plant enzymes in bacteria

The vast evolutionary distance between plants and prokaryotes raises a number of challenges when attempting to reconstitute plant biosynthetic pathways in bacteria. Achieving functional expression of complex enzymes in such a different context can be difficult. In particular, the activities of members of the cytochrome P450 oxidoreductase enzyme family, which commonly catalyze oxidation and hydroxylation reactions, are dependent on proper folding and integration into the endoplasmic reticulum to interact with reductase partners also located within this endomembrane. Thus, plant P450s are often nonfunctional when heterologously expressed in prokaryotes, which lack an endoplasmic reticulum. An additional challenge is ensuring that each enzyme in the pathway expresses stably and at an appropriate level, so that potentially toxic intermediates do not accumulate. Progress in synthetic biology has made both individual enzymes and entire enzyme pathways easier to functionally express.

A number of strategies have been developed to address the challenges that arise in achieving high activities of plant cytochrome P450s in heterologous bacterial hosts, including redesigning pathways to remove the reliance on P450-catalyzed reactions and expressing the P450 enzyme as a soluble protein fusion. In one example of the former strategy, the bacterial enzymes *hpaB* and *hpaC*, which together act as a *p*-hydroxyphenylacetate 3-hydroxylase, were used in a redesigned pathway in *E. coli* in place of a plant cytochrome P450 that

catalyzes a similar reaction in the biosynthesis of the phenylpropanoid caffeic acid.[85] In a more complex instance of pathway redesign, a pathway for the production of the benzylisoquinoline alkaloid (BIA) reticuline was constructed in *E. coli* without the cytochrome P450 CYP80B1, which catalyzes a required hydroxylation on the 1-benzylisoquinoline alkaloid scaffold.[86] The plant biosynthetic pathway to reticuline begins with the condensation of dopamine and 4-hydroxyphenylacetaldehyde to form the 1-benzylisoquinoline scaffold norcoclaurine, which subsequently undergoes a series of three methylation and one hydroxylation reactions to form reticuline. The redesigned pathway leveraged the substrate promiscuity of the plant enzymes and engineered an upstream pathway that produces dopamine and 3,4-dihydroxyphenylacetaldehyde, which when condensed provide the 1-benzylisoquinoline scaffold with the hydroxyl group added in such a way that only the methylation reactions were required to form reticuline from this modified scaffold.

Strategies have been developed to functionally express plant cytochrome P450s in bacterial hosts. An important consideration for functional P450 expression is the approach used to engineer the protein's N-terminal domain. The native N-terminal domain, which anchors into the ER membrane in eukaryotes and is often insoluble, is usually truncated[87] or replaced with other sequences, such as an 8-residue peptide from bovine 17α hydroxylase, to improve solubility.[88,89] In concert with N-terminal engineering, P450s can be expressed as a soluble fusion with the appropriate reductase partner; an approach which mimics the structure of the well-studied bacterial P450 BM3 (Fig. 3A). This recoding approach has been shown to provide solubility and enable efficient coupling between the cytochrome P450 and reductase partner. In one example, a P450 flavonoid 3′-hydroxylase was fused to a P450 reductase with a glycine-serine-threonine tripeptide linker, enabling soluble, functional expression.[90] Similarly, the P450 IFS1 was functionally expressed in *E. coli* for the conversion of the flavonoid naringenin to genistein. Optimizing the P450 N-terminal truncation and choice of linker domain increased genistein productivity five-fold over the original fusion construct; ultimately, a titer of over 10 mg g^{-1} cell weight of genistein was achieved.[91] This approach has also been validated in the context of expressing the P450 taxadiene 5-α hydroxylase for generating oxygenated taxanes from taxadiene.[92] One drawback of the fusion approach is that it enforces a 1 : 1 stoichiometry of P450s and reductases, which may not be optimal. Further work on taxadiene 5-α hydroxylase expression found that expressing the P450 and the reductase separately allowed for independent optimization of their expression levels, which enabled efficient catalysis without stressing the cell by overexpression of insoluble protein, and helped lead to a more than ten-fold increase in oxygenated taxane titers, to over 500 mg L^{-1}.[89]

Despite this drawback, enzyme fusions have been shown to be useful in contexts beyond P450 expression; they can increase pathway efficiency by co-localizing enzymes that catalyze subsequent reactions in a pathway. For example, as part of an effort to increase yields of the sesquiterpene plant defense

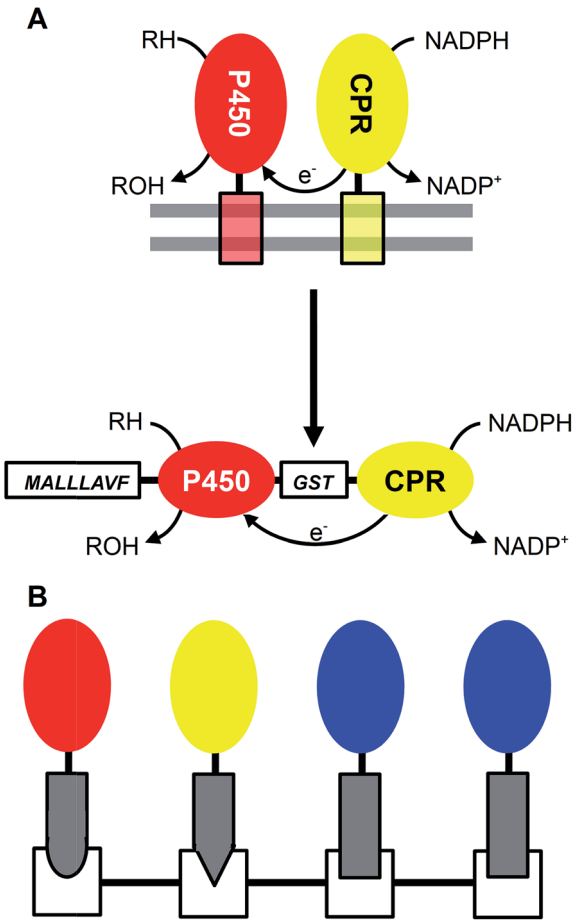


Fig. 3 Fusion and scaffolding approaches for heterologous enzyme expression. (A) P450–CPR complexes can be solubilized by protein fusion, circumventing the need for a eukaryotic endoplasmic reticulum. The solubilizing N-terminal peptide 'MALLLAVF' was used by Biggs *et al.*,[89] among others, and the linker peptide 'GST' was used by Zhu *et al.*[90] (B) Pathway enzyme colocalization at a defined stoichiometry *via* scaffolding, after the approach of Dueber *et al.*[94]

compound α-farnesene,[93] farnesyl pyrophosphate synthase and α-farnesene synthase, which acts on FPP, were fused with a $(GGGGS)_2$ linker peptide. The resulting strain achieved α-farnesene titers of 86.8 mg L^{-1} – approximately 50% more than a strain expressing the enzymes separately – and produced no detectable amount of the side product farnesol, supporting the conclusion that the protein fusion enabled efficient coupling between the two enzymes.

In order to co-localize enzymes while retaining control over ratios of enzyme expression, synthetic biologists have developed macromolecular scaffolds for enzyme co-localization (Fig. 3B). One such scaffold was designed using a set of protein–protein interaction pairs from metazoans. One member of each pair was fused to an enzyme of interest, and one or more copies of each interacting protein's partner were fused together to form a scaffold, where the relative numbers of each interacting domain included in the scaffold determined the stoichiometric ratio of co-localized enzymes. This system was used to co-localize mevalonate pathway enzymes in order to reduce the concentration of the toxic intermediate HMG-CoA

and boost mevalonate levels for a 77-fold improvement in product titer over a non-scaffolded implementation of the pathway.[94] Protein–nucleic acid interactions can also be used as the basis for a scaffold. Enzymes fused to zinc finger DNA-binding domains targeting specific DNA sequences were co-localized on DNA scaffolds designed with specified numbers of zinc finger binding sites to determine the ratio of co-localized enzymes. This strategy was applied to a pathway converting 4-coumaric acid to the stilbenoid compound resveratrol, leading to a five-fold increase in yield (to over 10 mg L^{-1}) *versus* a random-scaffold control.[95]

In addition to balancing enzyme activity levels, ensuring pathway stability is a common challenge. Heterologously expressed plant biosynthetic pathways impose stress on the host cell, creating a selective pressure for loss of function. Strategies for improving genetic stability can improve product yields. One such approach is chemically inducible chromosomal evolution (CIChE), in which the pathway's genes are integrated into a bacterial genome with an antibiotic resistance marker. *RecA*-mediated homologous recombination and growth in increasing concentrations of the antibiotic are used to increase the number of integrated pathway copies. Deleting *recA* then stabilizes the bacterial genome, yielding a strain that stably expresses integrated enzymes for long periods without antibiotic selection. This strategy was applied to increase titers of lycopene by 60% relative to a plasmid-based control (to over 12 000 ppm).[96] An extension of this system using copies of the *E. coli* essential gene *fabI* to confer resistance to increasing concentrations of the antibacterial compound triclosan was developed and used to overproduce lycopene without the need for antibiotic resistance markers at any stage of the strain construction process.[97]

3.3 Co-culture fermentation for enhanced efficiency towards synthesis of phytochemicals

Plant biosynthetic pathways can comprise a large number of biosynthetic enzymes, including membrane-bound proteins, which can lead to expression burden and growth deficiency when introduced into the heterologous microbial host. This expression burden can be shared among multiple microbial strains through compartmentalization, by breaking the complete pathway into separate modules and expressing each module in a dedicated microbial strain. Reconstructing a complex biosynthetic pathway across multiple production hosts can be used to address challenges associated with burden from expressing large numbers of heterologous enzymes – some of which may misfold or otherwise stress the cell – in a single host.

Co-culture of multiple *E. coli* strains, each harboring part of a complete biosynthetic pathway, has been used to increase production of phytochemicals including resveratrol[98] and other flavonoids.[99] In contrast, co-culture of *E. coli* and *S. cerevisiae* strains has been used to address challenges associated with plant P450 expression in bacteria. One such bacterial–yeast co-culture was used to produce the benzylisoquinoline alkaloids magnoflorine, corytuberine, and scoulerine; the scaffold molecule reticuline was synthesized in *E. coli* and subsequently converted to other BIAs in engineered *S. cerevisiae* strains expressing pathway cytochrome P450s.[100] In another study, a mutualistic yeast–bacteria co-culture system (Fig. 4) was developed by engineering an *E. coli* strain that metabolized xylose to acetate, which in turn inhibited bacterial growth and served as the carbon source for the yeast strain.[98] By expressing a taxadiene biosynthetic pathway in *E. coli* and taxadiene-oxygenating P450s in *S. cerevisiae*, a co-culture system produced oxygenated taxanes and other oxygenated plant isoprenoids. Bacteria–yeast co-culture leverages the fast growth rate and efficient metabolic processes of *E. coli* and the ability of *S. cerevisiae* to functionally express plant P450s without extensive enzyme engineering.

3.4 Screening and selection for pathway optimization

Given the complexity of biological systems, a complementary strategy to rationally designing a relatively small set of genetic

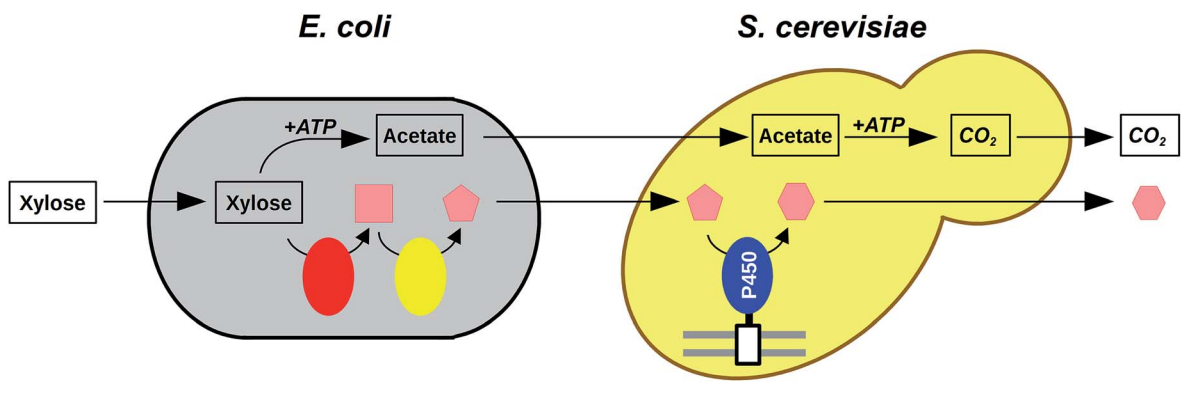


Fig. 4 Subdividing metabolic pathways among cocultured strains. Heterologous pathways can be expressed in consortia of bacteria and yeast, with the yeast expressing enzymes that may be challenging to functionally express in bacteria, such as plant cytochrome P450s. A mutualistic relationship can be engineered between the two organisms by engineering yeast to subsist on a bacterial fermentation product, ensuring that both microbes are available to carry out their designated roles. Xylose and acetate, shown here as carbon sources for *E. coli* and *S. cerevisiae*, respectively, were used by Zhou *et al.*[98] in one implementation of this strategy.

constructs is to generate a large library of candidate constructs and screen them for desired activities. In a screen, a phenotype of interest (such as production of a phytochemical) is measured for each member of the library – either directly by an analytical technique such as LC-MS, or indirectly, by coupling the phenotype of interest to another trait that can be measured in higher throughput, such as pigment formation or fluorescence. In a selection, the phenotype of interest is coupled to the strain's ability to survive, such that strains with the desired activities are enriched in the library over time. High-throughput screens and/or efficient selections enable researchers to sample a larger design space.

While production of almost any natural product can be measured by mass spectrometry, this is typically a relatively low-throughput approach that allows for characterizing only thousands of design variants. However, some compounds produce a visible phenotype such as a change in colony color, enabling higher-throughput screens for increased production levels. Lycopene, which produces red colonies that can be visually screened on agar plates, is commonly used in plate-based screening strategies. The related pigment astaxanthin has also been used for screening carotenoid production.[101] However, lycopene has been shown to degrade in the presence of hydrogen peroxide and other reactive oxygen species produced by oxidative stress, meaning it may not always be an accurate reporter of isoprenoid pathway flux.[102]

Synthetic biology strategies can be used to correlate the production of phytochemicals to a readily-screenable phenotype. Developing and characterizing biosensors – genetically encoded systems that respond to a stimulus, such as presence or absence of a compound – is a major area of interest in synthetic biology, which has been discussed in detail in recent reviews.[103,104] Many biosensors can be coupled to a screenable phenotype or to cell growth to enable screens or selections for a natural product (or a pathway intermediate). In one example, production of the isoprenoid precursor mevalonate was optimized through a biosensor-enabled screen. A mevalonate biosensor was first engineered through saturation mutagenesis of the transcriptional repressor AraC[105] and a GFP-based screen for variants which enabled significant increases in GFP production in the presence of mevalonate when GFP expression was controlled by the *araC*-regulated PBAD promoter. A library of mevalonate pathway variants was then constructed by modifying an existing plasmid expressing several genes from the pathway[56] and randomizing the ribosome binding site controlling production of the key enzyme HMG-CoA reductase. The resulting mevalonate-responsive *araC* variant was used to screen a library of 10^5 mevalonate pathway constructs by using the *araC* biosensor to drive expression of the reporter gene *lacZ*. The best constructs identified in this screen produced approximately 17 mM mevalonate – a three-fold increase over the original construct.[105] In another study leveraging a biosensor-enabled screen, a biosensor for the phytochemical triacetic acid lactone (TAL) was developed by directed evolution of *araC*, and used to screen a library of mutants of the TAL-generating enzyme 2-pyrone synthase, naturally found in the plant *Gerbera hybrida*, for increased activity in *E. coli*, leading to a 19-fold improvement in the enzyme's catalytic efficiency.[66] This biosensor was further applied to screen for increased production of TAL, as well as its precursor, malonyl-CoA, in an *E. coli* library generated by random transposon insertion, leading to the isolation of a strain producing over 30 mg per L per OD_{600} of TAL, a 4.2-fold improvement over the unmutagenized control.[106]

Beyond the design of the selection strategy, another important consideration in optimizing a heterologous pathway *via* selection or screening is the approach used for library generation. These approaches can induce changes throughout the bacterial genome, or target a small set of loci. Simple genome-wide approaches have been shown to be useful in the context of lycopene overproduction. "Shotgun" approaches, in which libraries of restriction enzyme-digested genomic fragments are transformed into *E. coli* and screened for product overproduction by colony phenotype, have been used to identify genes whose overexpression increases production of lycopene.[107,108] For instance, combinatorial optimization of 15 candidate genes for overexpression was performed leading to a strain that produced over 16 000 ppm lycopene, double the levels produced from a control strain.[107] Random genome-wide overexpression to increase lycopene titers has also been achieved by generating mutagenized libraries of the transcription-inducing sigma factor σ,[76] changing its binding specificity and causing it to upregulate randomly chosen sets of genes.[109] This approach led to a 50% increase in lycopene production (to 6.3 mg L^{-1}) over the original strain.

Strategies have also been developed that allow for a more targeted mutagenesis strategy. One such strategy is Multiplex Automated Genome Engineering (MAGE), in which cells are repeatedly transformed with oligonucleotide libraries targeting the RBSes of genes of interest to efficiently generate pathway libraries that sample a range of expression levels of the target genes.[110] In one study applying this approach to phytochemical production, a selection system was developed by using *ttgR*, a transcriptional repressor that is inactivated in the presence of naringenin, to control the expression of *tolC*, a gene which confers resistance to sodium dodecyl sulfate (SDS) but sensitivity to colicin E1.[111] The system was used to couple growth in SDS-containing media to naringenin production while exposure to colicin E1 was used to eliminate escape mutants that had evolved constitutive *tolC* expression. The selection system was used to carry out four rounds of MAGE and achieved a 36-fold increase in naringenin titers up to 39 mg L^{-1}, where further optimization in bioreactor culture resulted in a titer of over 60 mg L^{-1}. The MAGE strategy was also used to evolve higher-titer lycopene strains,[110] modulating expression levels of genes identified as relevant for lycopene production in earlier shotgun screens[107,108] and in computational modeling of the lycopene pathway.[112]

While the quick growth and potential for high product titers associated with bacterial cultures makes bacteria attractive heterologous hosts for phytochemical pathways, the evolutionary distance between plants and bacteria can pose challenges for efficient engineering. The recent development of strategies that facilitate the functional expression of plant enzymes in bacteria, as well as the process of selecting optimal variants of designed

pathways, is helping to facilitate the development of bacterial systems for producing phytochemicals.

4 Yeast production of phytochemicals

Yeast systems are widely utilized in the heterologous production of plant secondary metabolites.[113] Like bacterial systems, yeast systems such as *Saccharomyces cerevisiae* are metabolically and genetically well-studied,[114] especially as compared to plant systems. In addition, compared to bacterial systems, yeast exhibits a more similar microenvironment and metabolite profile to plants, which can lead to higher activity of certain plant enzymes and less expression burden than in bacteria. In particular, although multiple studies have shown *E. coli* to be a feasible organism for the heterologous expression of plant P450s, it is generally believed that functional expression of this group of endoplasmic reticulum (ER)-bound enzymes in yeast is relatively easier.[115] However, as unicellular eukaryotic organisms, yeast are quite distinct from plants. Therefore, difficulties in heterologous expression of certain types of plant enzymes and inefficient conversion of pathway intermediates remain as challenges in yeast systems. Recently, a number of studies have demonstrated the potential of non-*Saccharomyces* yeasts as heterologous hosts for the biosynthesis of plant molecules.[116,117] However, *S. cerevisiae* remains the most common choice when utilizing yeast to synthesize plant molecules,[118] largely because *S. cerevisiae* is one of the best-studied and genetically accessible microorganisms.[114] In this section, we discuss synthetic biology strategies utilized in yeast, to reconstruct plant pathways and to achieve efficient yeast-based bio-production of phytochemicals.

4.1 Genetic modification strategies for multi-gene pathway assembly

Plant biosynthetic pathways can require up to 20–30 enzymes for the synthesis of phytochemicals in heterologous hosts. *S. cerevisiae* features efficient homologous recombination, and a number of cloning techniques have been developed based on this feature,[119] which makes the reconstruction of pathways of such complexity feasible. Recently, yeast recombination has been adopted to assemble complex pathways in yeast on a plasmid or on the chromosome, an approach referred to as "DNA assembler".[120] This approach has been used for reconstructing multi-gene pathways (Fig. 5A). To synthesize the MIA strictosidine, a precursor of the anticancer drug vinblastine, 21 different enzymes, 15 of them from the MIA-producing plant *Catharanthus roseus*, were required to be introduced into the *S. cerevisiae* genome. Each gene was inserted between a promoter and terminator, and the expression cassettes for the enzymes were grouped into 7 modules to be introduced into 7 loci in the *S. cerevisiae* genome.[121] Similarly, the expression of 21 and 23 enzymes from plants, mammals, bacteria and yeast was required for the synthesis of the opioid alkaloids thebaine and hydrocodone, respectively. These biosynthetic pathways were each split into 6 or 7 genetic modules that were then integrated into the *S. cerevisiae* genome.[122] More recently, the complete biosynthesis of noscapine was established in *S. cerevisiae*, with a titers up to 2.2 mg L^{-1} achieved through protein engineering and metabolic engineering.[123] This *de novo* noscapine biosynthesis requires the expression of 25 heterologous enzymes from plants and mammals, which not only demonstrates the potential of yeast as a producer of highly complex phytochemicals, but also an alternative route for the synthesis of novel derivative compounds.

To assemble a multi-gene pathway, each gene is typically placed between a promoter and terminator to form an expression cassette, and a set of expression cassettes are assembled using homologous recombination. Therefore, multiple sets of promoters and terminators are desired for more efficient pathway assembly.[124,125] Recently, a synthetic biology platform was developed to screen for a set of synthetic minimal inducible and constitutive yeast promoters assembled from libraries of core elements, upstream activating sequences, and a neutral AT-rich spacer.[126] While these synthetic promoters retain similar transcription levels to the parent promoters, they are only 1/6 of the original length.[126] Similarly, libraries of synthetic short terminators have also been designed and verified through fusing the consensus sequences of yeast terminators with short random sequences.[127,128] The development of such synthetic short promoters and terminators that retain gene expression activities has the potential to enable more efficient assembly of multi-gene pathways in yeast, thereby facilitating phytochemical production.

To reduce the number of promoter–terminator pairs required for pathway assembly, multiple-gene pathways have been assembled through fusing them in a single reading frame using ribosome-skipping sequences called 2A sequences[34] (discussed in the context of plant genetic engineering in Section 2.2). In one study, 2A sequences from various viruses were tested in *S. cerevisiae*, and the one from the virus *Thosea asigna* (T2A) was verified as a functional ribosome-skipping sequence in *S. cerevisiae*. Three genes encoding enzymes involved in the synthesis of β-carotene from the carotenoid-producing yeast *Xanthophyllomyces dendrorhous* were then expressed on a single expression cassette using the verified T2A, which enabled production of β-ionone (at a level of 0.22 mg g^{-1} dry weight) by introducing the expression of a carotenoid-cleavage dioxygenase from raspberry (*Rubus ideaus* RiCCD1).[129]

In addition to promoters and terminators, when assembling complex heterologous pathways in the yeast genome, multiple well-characterized integration loci are required. In addition to the loci such as *URA3*, *HIS3*, *LEU2*, and *TRP1* that are commonly inactivated to enable the use of auxotrophic markers, multiple integration sites have been examined and verified for the genomic integration of heterologous genes without affecting yeast growth.[130–132] To mimic the natural clustering of genes encoding enzymes responsible for the synthesis of certain molecules on microbial chromosomes, 14 integration sites separated by genes essential for survival or growth were verified on chromosomes X, XI, and XII.[131,132] A multi-gene pathway can thus be constructed through grouping multiple expression cassettes in a single locus with DNA assembler, or integrating each gene into the verified clustered locus. No direct

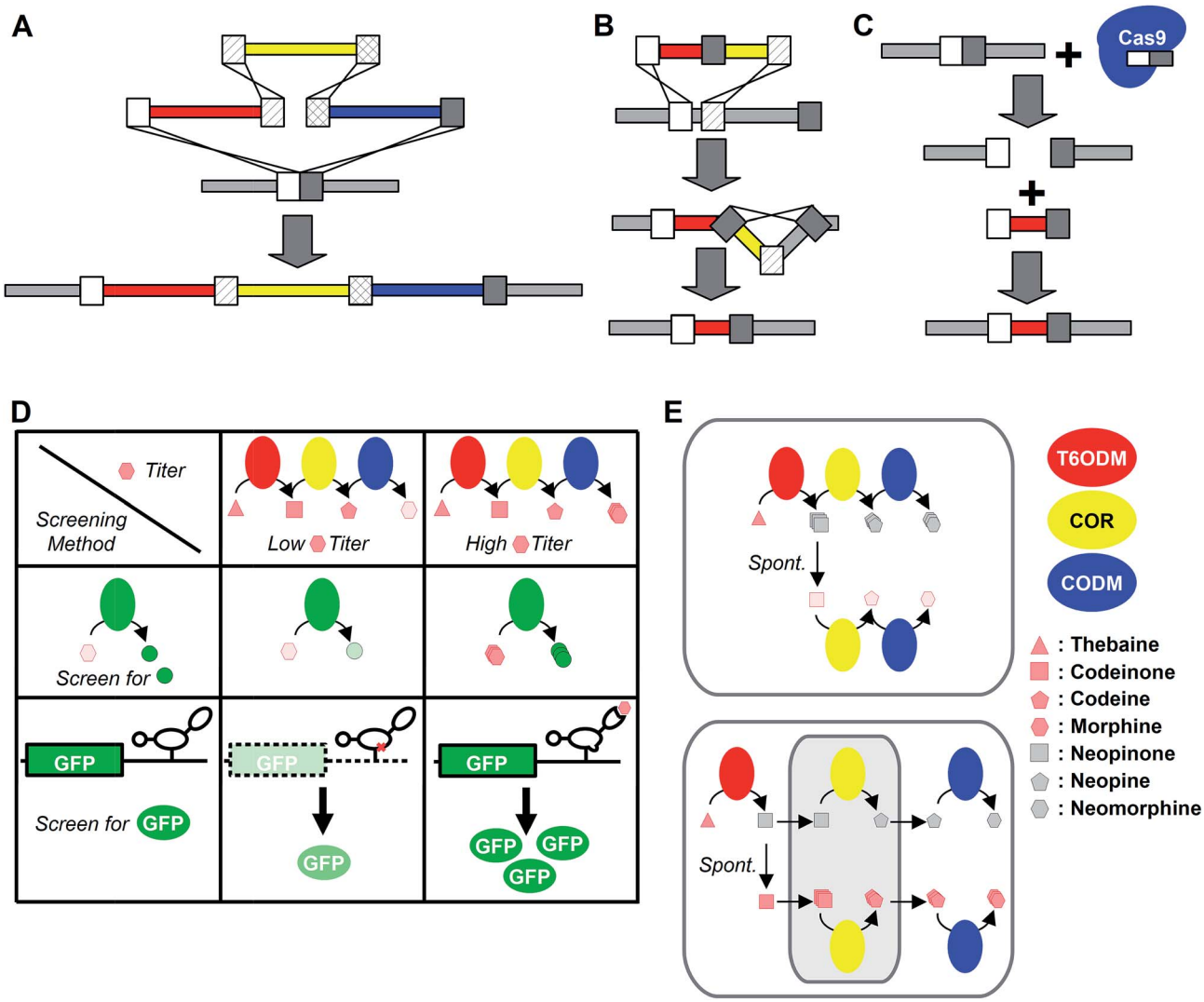


Fig. 5 Synthetic biology tools for genetic engineering and pathway optimization in yeast systems. (A) The "DNA Assembler" method for multigene integration relies on regions of shared homology between integration cassettes and the genome to assemble large pathways in a single step. (B) Homologous recombination can be leveraged to excise selection markers (yellow) *via* a counterselection strategy for scarless integration of desired heterologous genes (red).[134] (C) Cas9-mediated cleavage increases the efficiency of homology-guided integration. (D) Biosensor-aided screens in yeast have been implemented by expression of an enzyme that converts a product of interest to a colored, screenable compound, and by using a product-sensitive riboswitch to couple expression of green fluorescent protein (GFP) to the presence of the product. (E) Native organelles in yeast, such as the endoplasmic reticulum (gray), have been utilized for compartmentalization strategies. In one example, compartmentalization was used to separate enzymes in a pathway converting the opioid thebaine into morphine, providing more time for a spontaneous reaction ("Spont.") converting thebaine to codeinone to take place and reducing flux into an undesired side pathway leading to neomorphine.[148] T6ODM: thebaine 6-*O*-demethylase; COR: codeinone reductase; CODM: codeine *O*-demethylase.

comparison has been conducted yet to determine which approach is more optimal in multi-gene pathway reconstruction. In most cases, the Cre-LoxP-mediated marker recycling system is employed to rescue the selection marker,[133] but this method normally leaves scar sequences in the yeast chromosome. A set of yeast integrating plasmids was developed to facilitate a scarless integration after the rescue of the selection marker (Fig. 5B), by employing two rounds of homologous recombination.[134]

Integration efficiency is a key consideration in choosing a strain assembly method. Recently developed CRISPR/Cas9-based assembly systems (Fig. 5C) enable efficient, marker-free pathway assembly in yeast.[135,136] Complementing the CRISPR/Cas9 system with the efficient homologous recombination machinery in *S. cerevisiae*, a one-step method for marker-free yeast genome editing at multiple loci was developed, called CasEMBLR.[137] The application of CasEMBLR was validated by successfully constructing a carotenoid pathway and tyrosine pathway, through assembling 15 DNA fragments into three targeted loci and 10 DNA fragments into two loci, respectively.[137] Similarly, a vector toolkit called EasyClone-MarkerFree was established for marker-less integration of genes into *S. cerevisiae*, and showed efficient genetic engineering in a diploid industrial strain for production of 3-hydroxypropionic acid.[138] More recently, the expression levels and

integration efficiencies at 23 different chromosomal loci targeted by established Cas9-sgRNA plasmids were examined, both of which exhibited variety along the genome.[139]

4.2 Biosensor-driven protein engineering for higher pathway efficiency

The biosyntheses of phytochemicals, when compared to their microbial counterparts, are poorly understood. The biosynthetic pathways of most phytochemicals are not fully elucidated; enzymes catalyzing certain steps in the proposed pathway are unknown. In addition, due to differences in the microenvironment between plant and microbial hosts, plant biosynthetic enzymes often exhibit low efficiency in the microbial host, including yeast. Thus, protein evolution or engineering has become an especially useful approach towards obtaining the enzyme catalyzing the missing step, or enhancing the enzyme activity in a heterologous host.[140] For example, the biosynthetic route to make L-DOPA, one precursor of reticuline, has not been fully elucidated in plants: the enzyme catalyzing the 3-hydroxylation of tyrosine has not been identified. However, beet CYP76AD1 was recently found to catalyze the conversion of L-DOPA to *cyclo*-DOPA.[141] Protein evolution was then applied on CYP76AD1 for altered and enhanced activity as a tyrosine hydroxylase, through establishing a fluorescence-based screen (Fig. 5D) harnessing the further conversion of the product (L-DOPA) to the fluorescent compound betaxanthin, which is partially retained in the yeast cell's vacuole, by a plant DOPA dioxygenase. Screening of libraries constructed by error-prone PCR and DNA shuffling led to a variant which produced 2.8-fold more L-DOPA *in vivo* than the wild-type enzyme, with compounds further downstream in the reticuline pathway exhibiting further increases in titer.[142] In another example, a fluorescence-based screen for the methylxanthine alkaloid theophylline was established by using a theophylline-sensitive ribozyme switch to couple expression of green fluorescent protein (GFP) to theophylline levels (Fig. 5D). Seven rounds of screening libraries of variants of the bacterial P450 enzyme BM3 for caffeine demethylase activity increased *in vivo* enzyme activity by 33-fold.[143]

4.3 Strategies for gene expression tuning for optimizing phytochemical production

Higher protein expression levels can be desired to increase intracellular enzyme concentrations, which can lead to increased enzyme reaction rates. One common strategy to increase the expression level of a target protein is through introducing extra copies of the expression cassette of the target enzyme on a plasmid or into the host's chromosome. This approach was applied to synthesize the apocarotenoid β-ionone, by introducing the carotenoid cleavage dioxygenase from *Petunia hybrida* (*PhCCD1*), lycopene cyclase (*CrtYB*), and phytoene desaturase (*CrtI*) into the yeast genome.[144] The titer of β-ionone was increased 8.5-fold (to 0.62 mg g^{-1} dry weight) by increasing the expression levels of *CrtYB*, *CrtYB* with *CrtI*, or *CrtYB* with *CrtI* and *PhCCD1*, which was achieved through introducing high-copy plasmids encoding the corresponding expression cassettes.[144]

The expression of heterologous enzymes on high-copy plasmids does not always lead to higher production of the target molecules. In particular, because plant cytochrome P450s and reductases (CPRs) are ER membrane-bound proteins, higher expression of this class of enzymes does not always lead to higher enzyme activity. In one study demonstrating this concept, cheilanthifoline synthase from *Eschscholzia californica* (*EcCFS*) and one CPR from *Arabidopsis thaliana* (*ATR1*) were introduced into yeast on high-copy plasmids, low-copy plasmids, or through genome integration. The lowest cheilanthifoline titer was achieved with expression from high-copy plasmids, whereas expressing cheilanthifoline synthase from a more stable low-copy plasmid resulted in a 2.6-fold yield increase.[145] Through fusing green fluorescent protein (GFP) with *EcCFS* and imaging fluorescence in the cell population, this study also indicated that the high-copy plasmids encoding these enzymes were not as well maintained as the low-copy plasmids in yeast (90% for low-copy plasmid *versus* 10.7% for high-copy plasmid).[145] Similarly, when expressing the three yeast β-carotene synthetic enzymes within a single open reading frame using the 2A peptide, western blot analysis indicated that expressing with a low-copy plasmid leads to a higher level of the heterologous enzymes in *S. cerevisiae*.[129] In another study, focusing on the synthesis of methylxanthines in *S. cerevisiae*, higher methylxanthine production was observed when chromosomally introducing expression of biosynthetic enzymes compared with expression from a plasmid.[146] Recently, through testing truncated, weakened promoters to regulate the expression of antibiotic resistance markers, synthetic biologists were able to engineer high-copy plasmids into tunable-copy plasmids regulated by varying antibiotic concentration.[147] This strategy, named Pathway Optimization by Tuning Antibiotic Concentrations (POTAC), was utilized to enhance the titer of lycopene and *n*-butanol by 10- and 100-fold, respectively (to approximately 11 mg g^{-1} dry weight lycopene and approximately 110 mg L^{-1} *n*-butanol).[147]

Under some circumstances, a certain ratio between the expression levels of different synthetic enzymes in the pathway may lead to a relatively higher titer of target molecule. This strategy was recently demonstrated on the downstream enzymes in the morphinan biosynthetic pathway, in which thebaine 6-*O*-demethylase (T6ODM), codeinone reductase (COR), and codeine *O*-demethylase (CODM) are required to convert thebaine to a mixture of morphine and neomorphine in *S. cerevisiae*.[148] The titers of morphine and neomorphine vary from strains harboring different copy numbers of these three enzymes.[148] The morphine titer nearly doubled when increasing the copy number of CODM (T6ODM : COR : CODM = 1 : 1 : 3) compared to increasing the copy number of COR (1 : 3 : 1).[148]

To combine the advantages of high-level expression from multi-copy plasmids and stable chromosomal expression, a number of multi-copy chromosomal integration strategies have recently been developed. More than 100 δ sequences of the Ty element are distributed along the genome of *S. cerevisiae*, and this element has been developed as a recombination site

for one-to-multiple-copy integration of heterologous genes.[149] In addition to multiple-copy integration of a single gene, δ integration was recently employed for the combinatorial integration of several genes simultaneously onto the yeast genome, introducing extra copies of enzymes involved in the synthesis of FPP so as to enhance the titer of amorpha-4,11-diene.[150] Additionally, by using the CRISPR/CAS system together with δ integration, a Di-CRISPR (δ-integration CRISPR-Cas) platform was established recently in *S. cerevisiae* that enables efficient multicopy genomic integration (up to 18 copies) of a multi-gene pathway of 24 kb.[151]

In addition to tuning the copy number of expression cassettes encoding target synthetic enzymes, another major approach to alter the expression level of heterologous protein is through using different types of promoters. Strong promoters, such as the inducible *GAL* promoters and constitutive *TEF1* promoter,[139,152,153] are normally adopted to regulate the expression of heterologous enzymes for relatively high expression level. However, expression of multiple heterologous enzymes under strong, constitutive promoters is not always desirable. For example, three cytochrome P450s are required to synthesize the alkaloid noscapine from the precursor compound canadine in yeast. A recent study demonstrated that using either weaker or late-stage promoters to drive expression of the P450 that functions in the downstream part of the noscapine pathway leads to an enhanced titer of noscapine.[154] In addition to using native promoters to control the timing and expression level of target genes, synthetic promoters have also been developed to accomplish fine-tuning of gene expression.[155] A set of synthetic core promoter sequences, which are designed from *S. cerevisiae* natural strong core promoters, were fused to the *cis*-regulatory modules of the *Pichia pastoris AOX1* promoter.[155] This resulting library of promoters exhibits activities ranging from 0.3% to 70.6% of the wild-type *AOX1* promoter,[155] which will enable precise expression level regulation. In addition, orthogonal regulation of gene expression has been attempted in *S. cerevisiae* through introducing NAC transcription factors from *A. thaliana*, activation domains from yeast or plant, and cognate binding sites into the yeast transcription system, resulting in even higher transcription activation output than the commonly utilized constitutive promoter *TDH3*.[156]

4.4 Central metabolism engineering to increase pathway substrate and cofactor levels

In addition to developing strategies to alter or tune the expression level of heterologous genes, strategies to engineer central metabolism in yeast have been developed to increase the productivity of heterologous pathways by overproducing the substrates and cofactors needed to synthesize the desired phytochemicals. Many of these strategies have been applied to production of terpenoids *via* the mevalonate pathway. Plant terpenoids span a large group of high-value phytochemicals,[157] which are synthesized from the universal five-carbon building blocks IPP and dimethylallyl diphosphate (DMAPP). *S. cerevisiae* utilizes the mevalonate (MVA) pathway to synthesize IPP and DMAPP, which are then conjugated to FPP by ERG20p. A number of synthetic biology tools and metabolic engineering strategies have been applied to enhance the supply of IPP and DMAPP, while at the same time inhibiting the native pathway for the cell membrane component ergosterol, which consumes these terpene substrates.[158] While this pathway is also often imported into bacteria for isoprenoid production, the reverse process of importing the bacterial nonmevalonate MEP pathway into yeast was achieved relatively recently.[159] Efforts to produce plant terpenes in yeast to date have relied on the native MVA pathway, although the nonmevalonate MEP pathway has a higher theoretical yield of terpene precursors.[57]

Gene expression tuning has been applied to increase flux through the MVA pathway and FPP production. One common strategy to enhance the level of metabolites along the MVA pathway is to overexpress a truncated version of HMG-CoA reductase isoenzyme 1, which lacks an anchoring transmembrane domain and is more soluble and active.[160–163] In addition, overexpression of key endogenous yeast enzymes, such as Erg20p and Idi1p, has been applied to enhance production of FPP.[121,160,162,164] Another common strategy is to upregulate the expression level of a global transcription factor regulating sterol biosynthesis, Upc2p, or its mutant Upc2-1p,[160–163,165] leading to increased levels of enzymes involved in the synthesis of IPP and ergosterol from squalene.[160,165,166]

Beyond increasing flux through the MVA, another strategy to increase target production is to eliminate side products. However, certain native metabolites generated from MVA intermediates are required by the cell, such that the enzymes responsible for their production are essential for cell viability. Downregulating of enzymes responsible for the conversion of FPP to downstream primary metabolites, such as Erg9p, has been achieved through replacing the native *ERG9* promoter with the glucose-sensing *HXT1* promoter,[163] a methionine-repressible *MET3* promoter,[160] or a copper-regulated *CTR3* promoter.[167] In addition, tRNA synthesis requires DMAPP, overexpression of negative regulators in tRNA biosynthesis such as Mfa1p can increase the production of target terpenoids.[121,164] Lastly, geranylgeranyl diphosphate synthase (GGPPS) or geranyl pyrophosphate synthase (GPPS) from heterologous hosts have been employed to draw metabolite flux from the terpene pathway towards GGPP and GPP, building blocks for diterpenes[161] and monoterpenes,[121] respectively. Similar approaches to those used to optimize the MVA have been applied to the primary metabolic pathways towards the synthesis of tyrosine and malonyl-CoA, which are involved in the synthesis of benzylisoquinoline alkaloids[122,123,168,169] and flavonoids.[170–172]

In addition to the supply of pathway substrates, strategies to enhance the supply of pathway cofactors have been developed and employed to enhance the efficiencies of heterologous pathways in yeast. In one study, enzymes synthesizing NADPH, such as NAD-dependent glutamate dehydrogenase *GDH2* (ref. 163) and $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase *gapN*,[173] were introduced into an engineered yeast strain to increase NADPH supply, and therefore the efficiency of pathways involving enzymes such as cytochrome P450s. In another example, an enzyme involved in synthesizing SAM, *S*-adenosylmethionine synthetase *SAM2*, was upregulated to enhance the

production of the monoterpene strictosidine by increasing cofactor supply for a plant methyltransferase in the pathway.[121]

Another strategy to enhance the efficiency of heterologous pathways is to adjust metabolite concentrations in the cell, so as to push the reaction balance towards the desired product. To enhance the synthesis of resveratrol in *S. cerevisiae*, the low-affinity, high-capacity bacterial *araE* transporter – which natively transports arabinose, a molecule that, like resveratrol, belongs to the family of polyhydric alcohols – was introduced into a resveratrol-producing yeast strain, and led to a 2.44-fold increase in resveratrol production.[174] Metabolite export can also be enhanced by changes to the culture medium, as opposed to the engineered strain itself; in one study, when utilizing yeast to synthesize the triterpene saponins, methylated β-cyclodextrin was applied to the culture. This compound sequesters triterpenes and facilitates their export to the medium, leading to an enhanced production of saponins.[175]

Heterologous pathways can be better adapted to yeast central metabolism using inducible promoters. In one study, through using ergosterol-repressed promoters to regulate the expression of Erg9p, the metabolite flux towards the synthesis of ergosterol from FPP was reduced *via* feedback regulation, and thus the production of amorpha-4,11-diene from FPP was enhanced.[176] Promoters which respond to stress conditions can also be used in this approach: when a set of low pH-responsive promoters were developed and utilized to drive expression of lactate dehydrogenase in an industrial yeast strain in later stages of fermentation, a tenfold enhancement of the production of the target molecule lactic acid was observed *versus* a control strain driving lactate dehydrogenase expression with the *TEF1* promoter.[177] However, this set of low pH-responsive or stress-responsive promoters has not yet been applied to plant heterologous pathways in yeast.

4.5 Heterologous enzyme compartmentalization for enhanced protein–protein interaction

Biosynthetic enzymes associated with plant secondary metabolism are commonly distributed across different organelles and cell types. However, when transferring the synthetic machinery from plants to yeast, all pathway enzymes are typically expressed within single yeast cells. Yeast subcellular compartmentalization of heterologous enzymes has been employed to examine the effects of altered localization on the activity of heterologous enzymes and pathways. For example, enhanced production of sesquiterpenes was achieved when targeting FPP to mitochondria, which indicates that there are significant amounts of DMAPP and IPP in *S. cerevisiae* mitochondria.[178] Similarly, L-ornithine, an intermediate of L-arginine biosynthesis, can be synthesized in yeast in the mitochondria and cytosol. Relocalizing the mitochondrial L-ornithine biosynthetic pathway to the cytosol leads to an increase of L-ornithine production, while relocalization from the cytosol to the mitochondria leads to decreased L-ornithine production.[179] These examples highlight that the effect of relocalization strategies can vary on different pathways, and thus such strategies need to be carefully considered and designed based on the specifics of the pathway and the host cell's metabolism. In addition, heterologous enzymes have also been de-localized to direct specificity towards certain reactions. One study applying this concept focused on the conversion of morphine from thebaine, which requires at least three biosynthetic enzymes – codeinone reductase (COR) and the dioxygenases thebaine 6-*O*-demethylase (T6ODM) and codeine *O*-demethylase (CODM) – and relies on the spontaneous conversion of neopinone (product of T6ODM) to codeinone (substrate of COR) (Fig. 5E). However, data from the heterologous expression of these enzymes in the yeast cytosol indicated that neopinone can also be reduced by COR. The pathway was redesigned such that COR and T6ODM were localized to distinct subcellular compartments to allow additional time for neopinone to be converted to codeinone before coming in contact with COR, which enhanced selectivity toward the morphine branch of the pathway from 44% to 86%.[148]

The development of synthetic biology techniques for engineering more efficient enzymes and optimizing yields from complex heterologous pathways in yeast has enabled the rapid reconstruction of plant pathways and higher titers of phytochemicals. Additionally, yeast's intracellular compartments have been exploited to enable engineered localization strategies, further increasing product selectivity. Yeast production of phytochemicals also enables the potential of synthesizing novel "unnatural natural products".[26,123,180,181] Heterologous phytochemical pathway reconstruction will benefit from continued advances in synthetic biology; precise expression tools, biosensors, and other techniques will facilitate the process of plant pathway reconstitution.

5 Conclusion

Phytochemicals have been vital to medicine and society for millennia, but the supply of many of these compounds is limited by the need to cultivate the organism that produces them. In recent years, progress in synthetic biology has enabled reconstituting metabolic pathways in tractable heterologous hosts. Pathway reconstruction in plant, bacterial, and yeast hosts has enabled phytochemical production for research and industrial applications. Ongoing progress in identifying the underlying biosynthetic pathways in native producers, such as identification of biosynthetic gene clusters in plant genomes,[182] continues to fuel progress in this area.

Several major challenges are currently impeding efforts to produce phytochemicals in heterologous hosts rapidly, cheaply, and at commercially relevant titers and scales. First, the biosyntheses of the majority of phytochemicals, especially those of complex structures, are under-investigated. Thus, the knowledge of biosynthetic machinery to transfer from native producers to heterologous hosts is limited. However, heterologous reconstruction of plant pathways in a simpler plant model organism or yeast has proven to be a useful strategy for enzyme discovery, thereby filling in the missing links in phytochemical biosyntheses. Second, although pathways involving large numbers of enzymes, including membrane-bound proteins, have been successfully reconstituted in heterologous hosts, the resulting product titers are much lower than would be needed

for a commercially viable bioprocess. This challenge could potentially be addressed through engineering the hosts themselves to enable higher expression levels of heterologously expressed enzymes. Additionally, the process of strain engineering and optimization could be facilitated by the continued development of biosensors – particularly by biosensor design strategies generalizable to a large number of target phytochemicals – to enable high-throughput screening of strain libraries.

The abundance of most phytochemicals in nature is low, and heterologous production of these valuable molecules is considered to be a promising alternative for their supply. Recent advances in synthetic biology strategies for phytochemical production in heterologous plant, bacterial, and yeast hosts demonstrate the potential of heterologous bioproduction in addressing challenges in the sourcing of these molecules, and also as a platform for foundational investigations into plant secondary metabolism. In the future, many phytochemicals may be primarily obtained not from the original plant source, but from bioengineered organisms developed using synthetic biology tools.

6 Conflicts of interest

The authors declare no conflict of interest regarding the publication of this article.

7 References

1 R. A. Dixon, *Nature*, 2001, **411**, 843–847.
2 L. J. Cseke, *Natural products from plants*, CRC/Taylor & Francis, Boca Raton, FL, 2nd edn, 2006.
3 N. Dudareva, F. Negre, D. A. Nagegowda and I. Orlova, *Crit. Rev. Plant Sci.*, 2006, **25**, 417–440.
4 A. E. Osbourn and V. Lanzotti, *Plant-derived natural products*, Springer, 2009.
5 A. Staniek, H. Bouwmeester, P. D. Fraser, O. Kayser, S. Martens, A. Tissier, S. van der Krol, L. Wessjohann and H. Warzecha, *Biotechnol. J.*, 2014, **9**, 326–336.
6 S. M. K. Rates, *Toxicon*, 2001, **39**, 603–613.
7 M. S. Hussain, S. Fareed, S. Ansari, M. A. Rahman, I. Z. Ahmad and M. Saeed, *J. Pharm. BioAllied Sci.*, 2012, **4**, 10–20.
8 H.-J. Bart, in *Industrial Scale Natural Products Extraction*, Wiley-VCH Verlag GmbH & Co. KGaA, 2011, pp. 1–25, DOI: 10.1002/9783527635122.ch1.
9 J. Azmir, I. S. M. Zaidul, M. M. Rahman, K. M. Sharif, A. Mohamed, F. Sahena, M. H. A. Jahurul, K. Ghafoor, N. A. N. Norulaini and A. K. M. Omar, *J. Food Eng.*, 2013, **117**, 426–436.
10 I. Smetanska, *Adv. Biochem. Eng./Biotechnol.*, 2008, **111**, 187–228.
11 S. A. Wilson and S. C. Roberts, *Plant Biotechnol. J.*, 2012, **10**, 249–268.
12 S. Roberts and M. Kolewe, *Nat. Biotechnol.*, 2010, **28**, 1175–1176.
13 G. Farre, R. M. Twyman, P. Christou, T. Capell and C. Zhu, *Curr. Opin. Biotechnol.*, 2015, **32**, 54–60.
14 G. Farre, D. Blancquaert, T. Capell, D. Van Der Straeten, P. Christou and C. Zhu, *Annu. Rev. Plant Biol.*, 2014, **65**, 187–223.
15 I. Arabidopsis Genome, *Nature*, 2000, **408**, 796–815.
16 C. Schnee, T. G. Kollner, M. Held, T. C. J. Turlings, J. Gershenzon and J. Degenhardt, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1129–1134.
17 R. H. Smith, *Plant tissue culture: techniques and experiments*, Academic Press, 2013.
18 T. Nagata, Y. Nemoto and S. Hasezawa, *Int. Rev. Cytol.*, 1992, **132**, 1–30.
19 M. M. Goodin, D. Zaitlin, R. A. Naidu and S. A. Lommel, *Mol. Plant-Microbe Interact.*, 2008, **21**, 1015–1026.
20 S. Wu, M. Schalk, A. Clark, R. B. Miles, R. Coates and J. Chappell, *Nat. Biotechnol.*, 2006, **24**, 1441–1447.
21 T. W. J. M. van Herpen, K. Cankar, M. Nogueira, D. Bosch, H. J. Bouwmeester and J. Beekwilder, *PLoS One*, 2010, **5**, e14222.
22 Y. Zhang, G. Nowak, D. W. Reed and P. S. Covello, *Plant Biotechnol. J.*, 2011, **9**, 445–454.
23 B. Saxena, M. Subramaniyan, K. Malhotra, N. S. Bhavesh, S. D. Potlakayala and S. Kumar, *J. Biosci.*, 2014, **39**, 33–41.
24 M. Farhi, E. Marhevka, J. Ben-Ari, A. Algamas-Dimantov, Z. Liang, V. Zeevi, O. Edelbaum, B. Spitzer-Rimon, H. Abeliovich, B. Schwartz, T. Tzfira and A. Vainstein, *Nat. Biotechnol.*, 2011, **29**, 1072–1074.
25 Q. Liu, D. Manzano, N. Tanic, M. Pesic, J. Bankovic, I. Pateraki, L. Ricard, A. Ferrer, R. de Vos, S. van de Krol and H. Bouwmeester, *Metab. Eng.*, 2014, **23**, 145–153.
26 Q. Liu, M. Majdi, K. Cankar, M. Goedbloed, T. Charnikhova, F. W. Verstappen, R. C. de Vos, J. Beekwilder, S. van der Krol and H. J. Bouwmeester, *PLoS One*, 2011, **6**, e23255.
27 K. Cankar, E. Jongedijk, M. Klompmaker, T. Majdic, R. Mumm, H. Bouwmeester, D. Bosch and J. Beekwilder, *Biotechnol. J.*, 2015, **10**, 180–189.
28 H. M. Ting, B. Wang, A. M. Rydén, L. Woittiez, T. Herpen, F. W. Verstappen, C. Ruyter-Spira, J. Beekwilder, H. J. Bouwmeester and A. Krol, *New Phytol.*, 2013, **199**, 352–366.
29 K. Bruckner and A. Tissier, *Plant methods*, 2013, **9**, 46.
30 D. Rontein, S. Onillon, G. Herbette, A. Lesot, D. Werck-Reichhart, C. Sallaud and A. Tissier, *J. Biol. Chem.*, 2008, **283**, 6067–6075.
31 M. M. Hasan, H. S. Kim, J. H. Jeon, S. H. Kim, B. Moon, J. Y. Song, S. H. Shim and K. H. Baek, *Plant Cell Rep.*, 2014, **33**, 895–904.
32 K. Miettinen, L. Dong, N. Navrot, T. Schneider, V. Burlat, J. Pollier, L. Woittiez, S. van der Krol, R. Lugan, T. Ilc, R. Verpoorte, K. M. Oksman-Caldentey, E. Martinoia, H. Bouwmeester, A. Goossens, J. Memelink and D. Werck-Reichhart, *Nat. Commun.*, 2014, **5**, 3606.
33 W. Lau and E. S. Sattely, *Science*, 2015, **349**, 1224–1228.
34 F. Geu-Flores, C. E. Olsen and B. A. Halkier, *Planta*, 2009, **229**, 261–270.
35 H. Zhang, L. Wang, D. Hunter, C. Voogd, N. Joyce and K. Davies, *Plant methods*, 2013, **9**, 28.

36 H. T. Simonsen, D. P. Drew and C. Lunde, *Perspect. Med. Chem.*, 2009, **3**, 1–6.
37 X. Zhan, Y. H. Zhang, D. F. Chen and H. T. Simonsen, *Front Plant Sci.*, 2014, **5**, 636.
38 A. Anterola, E. Shanle, P. F. Perroud and R. Quatrano, *Transgenic Res.*, 2009, **18**, 655–660.
39 D. G. Schaefer, *Annu. Rev. Plant Biol.*, 2002, **53**, 477–501.
40 S. S. Bach, B. C. King, X. Zhan, H. T. Simonsen and B. Hamberger, *Methods Mol. Biol.*, 2014, **1153**, 257–271.
41 K. Kovacs, L. Zhang, R. S. Linforth, B. Whittaker, C. J. Hayes and R. G. Fray, *Transgenic Res.*, 2007, **16**, 121–126.
42 M. Li, F. Jiang, X. Yu and Z. Miao, *BioMed Res. Int.*, 2015, **2015**, 504932.
43 M. Cha, S. H. Shim, S. H. Kim, O. T. Kim, S. W. Lee, S. Y. Kwon and K. H. Baek, *BMB Rep.*, 2012, **45**, 589–594.
44 H. J. Kim, E. Ono, K. Morimoto, T. Yamagaki, A. Okazawa, A. Kobayashi and H. Satake, *Plant Cell Physiol.*, 2009, **50**, 2200–2209.
45 J. Murata, E. Matsumoto, K. Morimoto, T. Koyama and H. Satake, *PLoS One*, 2015, **10**, e0144519.
46 A. Ziemienowicz, *Biocatal. Agric. Biotechnol.*, 2014, **3**, 95–102.
47 S. S. Bach, J. E. Bassard, J. Andersen-Ranberg, M. E. Moldrup, H. T. Simonsen and B. Hamberger, *Methods Mol. Biol.*, 2014, **1153**, 245–255.
48 R. Bock, *Plant Mol. Biol.*, 2013, **83**, 21–31.
49 R. Lu, A. M. Martin-Hernandez, J. R. Peart, I. Malcuit and D. C. Baulcombe, *Methods*, 2003, **30**, 296–303.
50 J. F. Li, J. E. Norville, J. Aach, M. McCormack, D. D. Zhang, J. Bush, G. M. Church and J. Sheen, *Nat. Biotechnol.*, 2013, **31**, 688–691.
51 V. Nekrasov, B. Staskawicz, D. Weigel, J. D. G. Jones and S. Kamoun, *Nat. Biotechnol.*, 2013, **31**, 691–693.
52 U. Heinig, M. Gutensohn, N. Dudareva and A. Aharoni, *Curr. Opin. Biotechnol.*, 2013, **24**, 239–246.
53 I. F. Kappers, A. Aharoni, T. W. van Herpen, L. L. Luckerhoff, M. Dicke and H. J. Bouwmeester, *Science*, 2005, **309**, 2070–2072.
54 V. J. Martin, Y. Yoshikuni and J. D. Keasling, *Biotechnol. Bioeng.*, 2001, **75**, 497–503.
55 S. W. Kim and J. D. Keasling, *Biotechnol. Bioeng.*, 2001, **72**, 408–415.
56 V. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman and J. D. Keasling, *Nat. Biotechnol.*, 2003, **21**, 796–802.
57 E. Gruchattka, O. Hadicke, S. Klamt, V. Schutz and O. Kayser, *Microb. Cell Fact.*, 2013, **12**, 84.
58 F. K. Davies, V. H. Work, A. S. Beliaev and M. C. Posewitz, *Front Bioeng. Biotechnol.*, 2014, **2**, 21.
59 H. Kiyota, Y. Okuda, M. Ito, M. Y. Hirai and M. Ikeuchi, *J. Biotechnol.*, 2014, **185**, 1–7.
60 X. Wang, W. Liu, C. Xin, Y. Zheng, Y. Cheng, S. Sun, R. Li, X. G. Zhu, S. Y. Dai, P. M. Rentzepis and J. S. Yuan, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 14225–14230.
61 H. A. de Boer, L. J. Comstock and M. Vasser, *Proc. Natl. Acad. Sci. U. S. A.*, 1983, **80**, 21–25.
62 J. Wu, P. Liu, Y. Fan, H. Bao, G. Du, J. Zhou and J. Chen, *J. Biotechnol.*, 2013, **167**, 404–411.
63 H. Alper, C. Fischer, E. Nevoigt and G. Stephanopoulos, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12678–12683.
64 W. R. Farmer and J. C. Liao, *Nat. Biotechnol.*, 2000, **18**, 533–537.
65 R. Patnaik, W. D. Roof, R. F. Young and J. C. Liao, *J. Bacteriol.*, 1992, **174**, 7527–7532.
66 S. Y. Tang, S. Qian, O. Akinterinwa, C. S. Frei, J. A. Gredell and P. C. Cirino, *J. Am. Chem. Soc.*, 2013, **135**, 10099–10103.
67 D. Liu, Y. Xiao, B. S. Evans and F. Zhang, *ACS Synth. Biol.*, 2015, **4**, 132–140.
68 P. Xu, L. Li, F. Zhang, G. Stephanopoulos and M. Koffas, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 11299–11304.
69 F. Zhang, J. M. Carothers and J. D. Keasling, *Nat. Biotechnol.*, 2012, **30**, 354–359.
70 J. K. Rogers, C. D. Guzman, N. D. Taylor, S. Raman, K. Anderson and G. M. Church, *Nucleic Acids Res.*, 2015, **43**, 7648–7660.
71 R. H. Dahl, F. Zhang, J. Alonso-Gutierrez, E. Baidoo, T. S. Batth, A. M. Redding-Johanson, C. J. Petzold, A. Mukhopadhyay, T. S. Lee, P. D. Adams and J. D. Keasling, *Nat. Biotechnol.*, 2013, **31**, 1039–1046.
72 H. M. Salis, *Methods Enzymol.*, 2011, **498**, 19–42.
73 S. Kosuri, D. B. Goodman, G. Cambray, V. K. Mutalik, Y. Gao, A. P. Arkin, D. Endy and G. M. Church, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 14024–14029.
74 F. F. Nowroozi, E. E. K. Baidoo, S. Ermakov, A. M. Redding-Johanson, T. S. Batth, C. J. Petzold and J. D. Keasling, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 1567–1581.
75 L. Ye, C. Zhang, C. Bi, Q. Li and X. Zhang, *Microb. Cell Fact.*, 2016, **15**, 202.
76 B. F. Pfleger, D. J. Pitera, C. D. Smolke and J. D. Keasling, *Nat. Biotechnol.*, 2006, **24**, 1027–1032.
77 D. Lalaouna, M. Simoneau-Roy, D. Lafontaine and E. Masse, *Biochim. Biophys. Acta*, 2013, **1829**, 742–747.
78 T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner and H. Mori, *Mol. Syst. Biol.*, 2006, **2**, 20060008.
79 D. Na, S. M. Yoo, H. Chung, H. Park, J. H. Park and S. Y. Lee, *Nat. Biotechnol.*, 2013, **31**, 170–174.
80 A. Ghodasara and C. A. Voigt, *Nucleic Acids Res.*, 2017, **45**, 8116–8127.
81 L. S. Qi, M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin and W. A. Lim, *Cell*, 2013, **152**, 1173–1183.
82 J. Wu, G. Du, J. Chen and J. Zhou, *Sci. Rep.*, 2015, **5**, 13477.
83 B. F. Cress, Q. D. Leitz, D. C. Kim, T. D. Amore, J. Y. Suzuki, R. J. Linhardt and M. A. Koffas, *Microb. Cell Fact.*, 2017, **16**, 10.
84 S. K. Kim, G. H. Han, W. Seong, H. Kim, S. W. Kim, D. H. Lee and S. G. Lee, *Metab. Eng.*, 2016, **38**, 228–240.
85 Y. H. Lin and Y. J. Yan, *Microb. Cell Fact.*, 2012, **11**, 42.
86 A. Nakagawa, H. Minami, J. S. Kim, T. Koyanagi, T. Katayama, F. Sato and H. Kumagai, *Nat. Commun.*, 2011, **2**, 326.
87 A. Nakagawa, E. Matsumura, T. Koyanagi, T. Katayama, N. Kawano, K. Yoshimatsu, K. Yamamoto, H. Kumagai, F. Sato and H. Minami, *Nat. Commun.*, 2016, **7**, 10390.
88 M. C. Y. Chang, R. A. Eachus, W. Trieu, D. K. Ro and J. D. Keasling, *Nat. Chem. Biol.*, 2007, **3**, 274–277.
89 B. W. Biggs, C. G. Lim, K. Sagliani, S. Shankar, G. Stephanopoulos, M. De Mey and P. K. Ajikumar, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 3209–3214.

90 S. Zhu, J. Wu, G. Du, J. Zhou and J. Chen, *Appl. Environ. Microbiol.*, 2014, **80**, 3072–3080.
91 E. Leonard and M. A. Koffas, *Appl. Environ. Microbiol.*, 2007, **73**, 7246–7251.
92 P. K. Ajikumar, W. H. Xiao, K. E. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer and G. Stephanopoulos, *Science*, 2010, **330**, 70–74.
93 C. Wang, S. H. Yoon, H. J. Jang, Y. R. Chung, J. Y. Kim, E. S. Choi and S. W. Kim, *Metab. Eng.*, 2011, **13**, 648–655.
94 J. E. Dueber, G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. Prather and J. D. Keasling, *Nat. Biotechnol.*, 2009, **27**, 753–759.
95 R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnsek, N. Tomsic, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Bencina, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala and M. P. DeLisa, *Nucleic Acids Res.*, 2012, **40**, 1879–1889.
96 K. E. Tyo, P. K. Ajikumar and G. Stephanopoulos, *Nat. Biotechnol.*, 2009, **27**, 760–765.
97 S. Goh and L. Good, *BMC Biotechnol.*, 2008, **8**, 61.
98 K. Zhou, K. J. Qiao, S. Edgar and G. Stephanopoulos, *Nat. Biotechnol.*, 2015, **33**, 377–383.
99 J. A. Jones, V. R. Vernacchio, A. L. Sinkoe, S. M. Collins, M. H. A. Ibrahim, D. M. Lachance, J. Hahn and M. A. G. Koffas, *Metab. Eng.*, 2016, **35**, 55–63.
100 H. Minami, J. S. Kim, N. Ikezawa, T. Takemura, T. Katayama, H. Kumagai and F. Sato, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7393–7398.
101 C. Wang, M. K. Oh and J. C. Liao, *Biotechnol. Prog.*, 2000, **16**, 922–926.
102 M. Bongers, P. K. Chrysanthopoulos, J. B. Y. H. Behrendorff, M. P. Hodson, C. E. Vickers and L. K. Nielsen, *Microb. Cell Fact.*, 2015, **14**, 193.
103 L. Dekker and K. M. Polizzi, *Curr. Opin. Chem. Biol.*, 2017, **40**, 31–36.
104 D. Liu, T. Evans and F. Zhang, *Metab. Eng.*, 2015, **31**, 35–43.
105 S. Y. Tang and P. C. Cirino, *Angew. Chem., Int. Ed.*, 2011, **50**, 1084–1086.
106 H. Li, W. Chen, R. Jin, J. M. Jin and S. Y. Tang, *Microb. Cell Fact.*, 2017, **16**, 187.
107 Y. S. Jin and G. Stephanopoulos, *Metab. Eng.*, 2007, **9**, 337–347.
108 M. J. Kang, Y. M. Lee, S. H. Yoon, J. H. Kim, S. W. Ock, K. H. Jung, Y. C. Shin, J. D. Keasling and S. W. Kim, *Biotechnol. Bioeng.*, 2005, **91**, 636–642.
109 H. Alper and G. Stephanopoulos, *Metab. Eng.*, 2007, **9**, 258–267.
110 H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest and G. M. Church, *Nature*, 2009, **460**, 894–898.
111 S. Raman, J. K. Rogers, N. D. Taylor and G. M. Church, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 17803–17808.
112 H. Alper, Y. S. Jin, J. F. Moxley and G. Stephanopoulos, *Metab. Eng.*, 2005, **7**, 155–164.
113 Y. Luo, B. Z. Li, D. Liu, L. Zhang, Y. Chen, B. Jia, B. X. Zeng, H. Zhao and Y. J. Yuan, *Chem. Soc. Rev.*, 2015, **44**, 5265–5290.
114 I. Borodina and J. Nielsen, *Biotechnol. J.*, 2014, **9**, 609–620.
115 H. Renault, J. E. Bassard, B. Hamberger and D. Werck-Reichhart, *Curr. Opin. Plant Biol.*, 2014, **19**, 27–34.
116 C. Madzak, *Appl. Microbiol. Biotechnol.*, 2015, **99**, 4559–4577.
117 J. M. Wagner and H. S. Alper, *Fungal Genet. Biol.*, 2016, **89**, 126–136.
118 C. C. Hilde, D. P. Pia and H. M. Uffe, *Curr. Biotechnol.*, 2017, **6**, 194–204.
119 K. L. Tee and T. S. Wong, in *Directed Enzyme Evolution: Advances and Applications*, ed. M. Alcalde, Springer International Publishing, Cham, 2017, pp. 201–227, DOI: 10.1007/978-3-319-50413-1_8.
120 Z. Shao, H. Zhao and H. Zhao, *Nucleic Acids Res.*, 2009, **37**, e16.
121 S. Brown, M. Clastre, V. Courdavault and S. E. O'Connor, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 3205–3210.
122 S. Galanie, K. Thodey, I. J. Trenchard, M. Filsinger Interrante and C. D. Smolke, *Science*, 2015, **349**, 1095–1100.
123 Y. Li, S. Li, K. Thodey, I. Trenchard, A. Cravens and C. D. Smolke, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E3922–E3931.
124 M. Deaner and H. S. Alper, *Adv. Biochem. Eng./Biotechnol.*, 2018, **162**, 21–44.
125 J. Gilman and J. Love, *Biochem. Soc. Trans.*, 2016, **44**, 731–737.
126 H. Redden and H. S. Alper, *Nat. Commun.*, 2015, **6**, 7810.
127 M. MacPherson and Y. Saka, *ACS Synth. Biol.*, 2017, **6**, 130–138.
128 K. A. Curran, N. J. Morse, K. A. Markham, A. M. Wagman, A. Gupta and H. S. Alper, *ACS Synth. Biol.*, 2015, **4**, 824–832.
129 J. Beekwilder, H. M. van Rossum, F. Koopman, F. Sonntag, M. Buchhaupt, J. Schrader, R. D. Hall, D. Bosch, J. T. Pronk, A. J. van Maris and J. M. Daran, *J. Biotechnol.*, 2014, **192**, 383–392.
130 E. C. Tatsis, I. Carqueijeiro, T. Duge de Bernonville, J. Franke, T. T. Dang, A. Oudin, A. Lanoue, F. Lafontaine, A. K. Stavrinides, M. Clastre, V. Courdavault and S. E. O'Connor, *Nat. Commun.*, 2017, **8**, 316.
131 M. D. Mikkelsen, L. D. Buron, B. Salomonsen, C. E. Olsen, B. G. Hansen, U. H. Mortensen and B. A. Halkier, *Metab. Eng.*, 2012, **14**, 104–111.
132 N. B. Jensen, T. Strucko, K. R. Kildegaard, F. David, J. Maury, U. H. Mortensen, J. Forster, J. Nielsen and I. Borodina, *FEMS Yeast Res.*, 2014, **14**, 238–248.
133 A. Nagy, *Genesis*, 2000, **26**, 99–109.
134 M. S. Siddiqui, A. Choksi and C. D. Smolke, *FEMS Yeast Res.*, 2014, **14**, 1171–1185.
135 V. Stovicek, C. Holkenbrink and I. Borodina, *FEMS Yeast Res.*, 2017, **17**, fox030.
136 T. Jakociunas, M. K. Jensen and J. D. Keasling, *Metab. Eng.*, 2016, **34**, 44–59.
137 T. Jakociunas, A. S. Rajkumar, J. Zhang, D. Arsovska, A. Rodriguez, C. B. Jendresen, M. L. Skjodt, A. T. Nielsen, I. Borodina, M. K. Jensen and J. D. Keasling, *ACS Synth. Biol.*, 2015, **4**, 1226–1234.
138 M. M. Jessop-Fabre, T. Jakociunas, V. Stovicek, Z. Dai, M. K. Jensen, J. D. Keasling and I. Borodina, *Biotechnol. J.*, 2016, **11**, 1110–1117.

139 A. Reider Apel, L. d'Espaux, M. Wehrs, D. Sachs, R. A. Li, G. J. Tong, M. Garber, O. Nnadi, W. Zhuang, N. J. Hillson, J. D. Keasling and A. Mukhopadhyay, *Nucleic Acids Res.*, 2017, **45**, 496–508.
140 M. J. Smanski, H. Zhou, J. Claesen, B. Shen, M. A. Fischbach and C. A. Voigt, *Nat. Rev. Microbiol.*, 2016, **14**, 135–149.
141 G. J. Hatlestad, R. M. Sunnadeniya, N. A. Akhavan, A. Gonzalez, I. L. Goldman, J. M. McGrath and A. M. Lloyd, *Nat. Genet.*, 2012, **44**, 816–820.
142 W. C. DeLoache, Z. N. Russ, L. Narcross, A. M. Gonzales, V. J. Martin and J. E. Dueber, *Nat. Chem. Biol.*, 2015, **11**, 465–471.
143 J. K. Michener and C. D. Smolke, *Metab. Eng.*, 2012, **14**, 306–316.
144 J. Lopez, K. Essus, I. K. Kim, R. Pereira, J. Herzog, V. Siewers, J. Nielsen and E. Agosin, *Microb. Cell Fact.*, 2015, **14**, 84.
145 I. J. Trenchard and C. D. Smolke, *Metab. Eng.*, 2015, **30**, 96–104.
146 M. McKeague, Y. H. Wang, A. Cravens, M. N. Win and C. D. Smolke, *Metab. Eng.*, 2016, **38**, 191–203.
147 J. Z. Lian, R. Jin and H. M. Zhao, *Biotechnol. Bioeng.*, 2016, **113**, 2462–2473.
148 K. Thodey, S. Galanie and C. D. Smolke, *Nat. Chem. Biol.*, 2014, **10**, 837–844.
149 A. Sakai, Y. Shimizu and F. Hishinuma, *Appl. Microbiol. Biotechnol.*, 1990, **33**, 302–306.
150 J. Yuan and C. B. Ching, *Biotechnol. Bioeng.*, 2014, **111**, 608–617.
151 S. Shi, Y. Liang, M. M. Zhang, E. L. Ang and H. Zhao, *Metab. Eng.*, 2016, **33**, 19–27.
152 S. Partow, V. Siewers, S. Bjorn, J. Nielsen and J. Maury, *Yeast*, 2010, **27**, 955–964.
153 J. Sun, Z. Shao, H. Zhao, N. Nair, F. Wen, J. H. Xu and H. Zhao, *Biotechnol. Bioeng.*, 2012, **109**, 2082–2092.
154 Y. Li and C. D. Smolke, *Nat. Commun.*, 2016, **7**, 12137.
155 R. M. C. Portela, T. Vogl, C. Kniely, J. E. Fischer, R. Oliveira and A. Glieder, *ACS Synth. Biol.*, 2017, **6**, 471–484.
156 G. Naseri, S. Balazadeh, F. Machens, I. Kamranfar, K. Messerschmidt and B. Mueller-Roeber, *ACS Synth. Biol.*, 2017, **6**, 1742–1756.
157 D. Tholl, *Adv. Biochem. Eng./Biotechnol.*, 2015, **148**, 63–106.
158 C. E. Vickers, T. C. Williams, B. Peng and J. Cherry, *Curr. Opin. Chem. Biol.*, 2017, **40**, 47–56.
159 J. Kirby, K. L. Dietzel, G. Wichmann, R. Chan, E. Antipov, N. Moss, E. E. K. Baidoo, P. Jackson, S. P. Gaucher, S. Gottlieb, J. LaBarge, T. Mahatdejkul, K. M. Hawkins, S. Muley, J. D. Newman, P. Liu, J. D. Keasling and L. Zhao, *Metab. Eng.*, 2016, **38**, 494–503.
160 D. K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong and J. D. Keasling, *Nature*, 2006, **440**, 940–943.
161 B. Engels, P. Dahm and S. Jennewein, *Metab. Eng.*, 2008, **10**, 201–206.
162 G. Scalcinati, C. Knuf, S. Partow, Y. Chen, J. Maury, M. Schalk, L. Daviet, J. Nielsen and V. Siewers, *Metab. Eng.*, 2012, **14**, 91–103.
163 G. Scalcinati, S. Partow, V. Siewers, M. Schalk, L. Daviet and J. Nielsen, *Microb. Cell Fact.*, 2012, **11**, 117.
164 J. Liu, W. Zhang, G. Du, J. Chen and J. Zhou, *J. Biotechnol.*, 2013, **168**, 446–451.
165 X. Yan, Y. Fan, W. Wei, P. Wang, Q. Liu, Y. Wei, L. Zhang, G. Zhao, J. Yue and Z. Zhou, *Cell Res.*, 2014, **24**, 770–773.
166 B. S. Davies, H. S. Wang and J. Rine, *Mol. Cell. Biol.*, 2005, **25**, 7375–7385.
167 C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, D. R. Polichuk, K. H. Teoh, D. W. Reed, T. Treynor, J. Lenihan, M. Fleck, S. Bajad, G. Dang, D. Dengrove, D. Diola, G. Dorin, K. W. Ellens, S. Fickes, J. Galazzo, S. P. Gaucher, T. Geistlinger, R. Henry, M. Hepp, T. Horning, T. Iqbal, H. Jiang, L. Kizer, B. Lieu, D. Melis, N. Moss, R. Regentin, S. Secrest, H. Tsuruta, R. Vazquez, L. F. Westblade, L. Xu, M. Yu, Y. Zhang, L. Zhao, J. Lievense, P. S. Covello, J. D. Keasling, K. K. Reiling, N. S. Renninger and J. D. Newman, *Nature*, 2013, **496**, 528–532.
168 I. J. Trenchard, M. S. Siddiqui, K. Thodey and C. D. Smolke, *Metab. Eng.*, 2015, **31**, 74–83.
169 L. Narcross, E. Fossati, L. Bourgeois, J. E. Dueber and V. J. J. Martin, *Trends Biotechnol.*, 2016, **34**, 228–241.
170 Y. Wang, S. Chen and O. Yu, *Appl. Microbiol. Biotechnol.*, 2011, **91**, 949–956.
171 M. Li, K. R. Kildegaard, Y. Chen, A. Rodriguez, I. Borodina and J. Nielsen, *Metab. Eng.*, 2015, **32**, 1–11.
172 S. Y. Shin, S. M. Jung, M. D. Kim, N. S. Han and J. H. Seo, *Enzyme Microb. Technol.*, 2012, **51**, 211–216.
173 K. Kocharin, V. Siewers and J. Nielsen, *Biotechnol. Bioeng.*, 2013, **110**, 2216–2224.
174 Y. Wang, C. Halls, J. Zhang, M. Matsuno, Y. Zhang and O. Yu, *Metab. Eng.*, 2011, **13**, 455–463.
175 T. Moses, J. Pollier, L. Almagro, D. Buyst, M. Van Montagu, M. A. Pedreno, J. C. Martins, J. M. Thevelein and A. Goossens, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 1634–1639.
176 J. Yuan and C. B. Ching, *Microb. Cell Fact.*, 2015, **14**, 38.
177 A. S. Rajkumar, G. Liu, D. Bergenholm, D. Arsovska, M. Kristensen, J. Nielsen, M. K. Jensen and J. D. Keasling, *Nucleic Acids Res.*, 2016, **44**, e136.
178 M. Farhi, E. Marhevka, T. Masci, E. Marcos, Y. Eyal, M. Ovadis, H. Abeliovich and A. Vainstein, *Metab. Eng.*, 2011, **13**, 474–481.
179 J. Qin, Y. J. Zhou, A. Krivoruchko, M. Huang, L. Liu, S. Khoomrung, V. Siewers, B. Jiang and J. Nielsen, *Nat. Commun.*, 2015, **6**, 8224.
180 A. Moglia, C. Comino, S. Lanteri, R. de Vos, P. de Waard, T. A. van Beek, L. Goitre, S. F. Retta and J. Beekwilder, *Metab. Eng.*, 2010, **12**, 223–232.
181 A. Eudes, E. E. Baidoo, F. Yang, H. Burd, M. Z. Hadi, F. W. Collins, J. D. Keasling and D. Loque, *Appl. Microbiol. Biotechnol.*, 2011, **89**, 989–1000.
182 H. W. Nutzmann, A. Huang and A. Osbourn, *New Phytol.*, 2016, **211**, 771–789.

