REVIEW PAPER

Synthetic biology approaches for the production of plant metabolites in unicellular organisms

Tessa Moses[1,2,#], Payam Mehrshahi[3,#], Alison Smith[3] and Alain Goossens[1,2,*]

[1] Ghent University, Department of Plant Biotechnology and Bioinformatics, 9052 Ghent, Belgium

[2] VIB Center for Plant Systems Biology, 9052 Ghent, Belgium

[3] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA

[#] These authors contributed equally to this manuscript

[*]Correspondence: algoo@psb.vib-ugent.be; Tel +32(0)93313851

E-mail addresses: "Tessa Moses" temos@psb.vib-ugent.be; "Payam Mehrshahi" pm579@cam.ac.uk; "Alison Smith" as25@cam.ac.uk; "Alain Goossens" algoo@psb.vib-ugent.be

Submission date: 23/01/2017

Word count: 7,971

Number of Tables: 1

Number of figures: 4 (all in colour in print and online)

Abstract

Synthetic biology is the repurposing of biological systems for novel objectives and applications. Through the coordinated and balanced expression of genes, both native and those introduced from other organisms, resources within an industrial chassis can be siphoned for the commercial production of high-value commodities. This developing interdisciplinary field has the potential to revolutionise drug discovery from higher plants, by providing a diverse array of tools, technologies and strategies for exploring the large chemically complex space of plant natural products using unicellular organisms. In this review we emphasise the key features that influence the generation of biorefineries and highlight technologies and strategic solutions that can be used to overcome engineering pitfalls with rational design. Also presented is a succinct guide to assist the selection of unicellular chassis most suited for the engineering and subsequent production of the desired natural product, in order to meet the global demand for plant natural products in a safe and sustainable manner.

Highlight:

Synthetic biology provides tools for exploring the chemical space of plant natural products, thereby revolutionising drug discovery. A suite of unicellular organisms can be used as production chassis.

Keywords: promoter engineering, protein engineering, biosensor, genome reduction, non-canonical bacteria, non-canonical yeast, cyanobacteria, microalgae

Running Title: Synbio for plant metabolites in unicellular organisms

Introduction

Synthetic biology is an evolving field that is emerging as a centrepiece of rational biological design. It is an interdisciplinary science that combines the expertise of biologists, engineers, physicists and computer scientists to gain in-depth understanding of the functioning of living systems, thereby allowing their (re)design for novel purposes, such as to be utilised for the sustainable production of biomolecules, biomaterials, and fine and bulk chemicals. A relatively older field in biology, that involves the optimisation of genetic and regulatory processes within cells to increase the production of certain compounds of interest, is metabolic engineering. As research publications in the field increase, there is an ever-growing misuse of the terminologies of synthetic biology and metabolic engineering (Stephanopoulos, 2012). Before we go further into the implications and applications of synthetic biology, we will define and distinguish the concepts of metabolic engineering and synthetic biology.

Metabolic engineering in its simplest definition is the directed modification of metabolic pathways in a living organism for the synthesis of a desired product. It encompasses the compilation of pathway genes ultimately to yield product in a high titre, high production rate and efficient manner. Optimisation of the production host with flux balancing involves coordinating both cell and heterologous pathway performance. While the construction and generation of engineered strains may take only a short time, improving these to the point of use for commercial production is likely a much more time-consuming process. In this context, the core enabling technology of synthetic biology can aid metabolic engineering (Stephanopoulos, 2012). The two fields overlap partially in their interest of pathway engineering, but synthetic biology is characterised by the use of defined 'parts' that are easily combined and exchanged, using standardised workflows and often combined with modelling and computational methods, with the aim to take advantage of high-throughput methodology such as robotics (Fig. 1). This may be extended to synthetic cells, genetic circuits and non-linear cell dynamics for the generation of novel gene networks and whole genome synthesis. However, developments in both metabolic engineering and synthetic biology strategies provide complementary methods to improve the *in vivo* performance of cells, in particular industrial microorganisms (Awan *et al.*, 2016; Keasling, 2012; Lechner *et al.*, 2016; Stephanopoulos, 2012).

Synthetic biology is reorienting and revolutionising drug discovery from plants, a major source of human medicine and high-value compounds (Frasch *et al.*, 2013). Higher plants produce and accumulate a diverse array of small molecules (referred to as plant natural products, secondary or specialised metabolites) that are most often not essential for their

fundamental physiological processes, but contribute greatly to their ecology. Humans have exploited the therapeutic potential of many natural products in the form of traditional herbal medicine, and more recently popularised them as flavours, fragrances, colorants and health promoting agents. Based on their chemical structure and biosynthetic origin, the plant natural products are classified into three major classes, the terpenoids, alkaloids and phenolics (Croteau *et al.*, 2000). The chemical complexity of these compounds and their scarce production in the plants themselves has stimulated efforts to introduce their biosynthetic pathways into heterologous organisms. Microbial systems offer many advantages because these have rapid doubling times, are robust under process conditions, are readily scalable, and have extensive engineering toolkits. In addition, they offer a simplified product purification pipeline owing to the lack of competing contaminants, and most of all make the production process cost-effective by converting inexpensive feedstock to high value compounds. The success with metabolic engineering for the production of the anti-malarial drug artemisinin in yeast and *E. coli* set the dawn for synthetic biology for the production of plant natural products in cellular factories (Paddon and Keasling, 2014). Today microbial cells are widely used as production systems, as well as screening platforms for target chemical production. Furthermore, synthetic cells have been engineered with genetic circuits to decipher disease mechanisms, drug mode of action mechanisms and to study cell-to-cell communication within bacterial consortia (Trosset and Carbonell, 2015).

Plant natural products are often synthesised by the action of multiple enzymes: artemisinin biosynthesis from the central precursor farnesyl pyrophosphate requires the catalytic action of four enzymes, a sesquiterpene synthase (amorpha-4,11-diene synthase, ADS), a cytochrome P450 monooxygenase (CYP71AV1), an alcohol dehydrogenase (ADH1), and an artemisinic aldehyde dehydrogenase (ALDH1) (Covello, 2008). To produce a plant natural product in a heterologous host, balancing the expression of all pathway genes within a host cell is critical and vital. It is also essential to acknowledge the burden caused by the expression of the introduced biosynthetic pathway enzymes, which not only siphons resources from the host's native pathways, but also may generate toxic pathway intermediates (Keasling, 2012). Synthetic biology tools that allow finely-tuned multi-gene expression, rational engineering of heterologous proteins in the context of the production chassis, quantitative uncomplicated visual screens for the target chemical, and reduction of the metabolic burden on the host through the elimination of non-essential genetic elements, are reviewed here as strategies that will become increasingly relevant for the engineering of smart cell factories (Fig. 2). In addition, computational modelling is a valuable approach to

synthetic biology as it allows evaluating possible cellular designs for a desired outcome by combining well characterised cellular modules (Kaznessis, 2009). While synthetic biology aspires to implement modelling for the production of natural products, thus far computational systems biology has only been effectively applied for the production of precursors and cofactors valuable for natural product biosynthesis in the widely used chassis *Escherichia coli* (Maia *et al.*, 2015) and *Saccharomyces cerevisiae* (Mustacchi *et al.*, 2006), and is starting to be applied in cyanobacteria (Westermark and Steuer, 2016) and algae (Perez-Garcia *et al.*, 2016). In this review we also bring to the spotlight several rising chassis with desirable characteristics for biorefinery, including non-canonical bacteria, yeasts and photosynthetic algae. This review provides a road map of the latest synthetic biology tools available for consideration and showcases the hosts available for rational engineering of high-value plant natural products in unicellular microbes.

Microbial Cell Factories

The biotechnological production of compounds ranging from small aromatics to more complex plant specialised metabolites has been achieved in microbial cell factories (Wang *et al.*, 2016). Microbial fermentation regularly forms the core of the biorefinery by converting the often inexpensive raw material to the desired chemicals, thus providing added value. Over the years, *E. coli* and *S. cerevisiae* have stood out as the workhorses of biotechnology (Fig. 3) and continue to be engineered as cell factories for new biorefineries (Wang *et al.*, 2016). However, in the evolving era of biotechnology many other unicellular microbes are being explored as chassis for the production of plant metabolites. Here we highlight some of their key features and present success stories of engineered microbial cell factories.

The canonical hosts: Escherichia coli and Saccharomyces cerevisiae

Escherichia coli is the most widely used prokaryotic host for the production of heterologous metabolites, owing to its rapid growth, high product yield, cost-effective nature and ease of scale-up. The availability of various expression vectors and strains, manipulation technologies, and the relative ease of product purification make it an attractive host for industrial applications. However, the lack of intracellular membranes and absence of posttranslational modifications limit its use as the chassis of choice for many plant natural products. Nonetheless, technological advancements to meet industrial requirements have resulted in the use of *E. coli* for the production of a range of complex plant specialised metabolites, including flavonoids (Katsuyama *et al.*, 2007) and its glycosides (Kim *et al.*,

2015), mono- (Jongedijk *et al.*, 2016; Tashiro *et al.*, 2016), sesqui- (Han *et al.*, 2016; Yang and Nie, 2016), di- (Biggs *et al.*, 2016) and tetraterpenoids (Li *et al.*, 2015; Ma *et al.*, 2016), as well as alkaloid precursors such as (*S*)-reticuline (Matsumura *et al.*, 2017) and geraniol (Chen *et al.*, 2016).

Similarly, the well-known baker's yeast *Saccharomyces cerevisiae* has been the eukaryotic host of choice, since it is fermentable, genetically tractable and generally recognised as safe (GRAS). Moreover, it is more effective than *E. coli* in expressing genes coding for membrane-bound cytochrome P450 monooxygenases (P450), which are key catalysts in several plant specialised metabolite pathways, although yeasts require the coexpression of plant P450 reductase genes for optimal electron delivery to the P450 enzyme (Werck-Reichhart and Feyereisen, 2000). Thus far many plant specialised metabolite pathways have been successfully reconstituted in yeast for the production of flavonoids and stilbenoids (Billingsley *et al.*, 2016), a variety of alkaloids (Kishimoto *et al.*, 2016), polyketides (Bond *et al.*, 2016), terpenoids (Brown *et al.*, 2015; Paddon and Keasling, 2014), and aromatic amino acid derived compounds such as muconic acid (Suastegui and Shao, 2016). Yeasts have been employed to refactor biosynthetic pathways, optimise production and industrially scale-up engineered organisms using genomic, synthetic biology (Fig. 1) and metabolic engineering (Fig. 3) strategies. Despite the many advantages (Fig. 4; Billingsley *et al.*, 2016), *S. cerevisiae* can be a challenging host to engineer, as some proteins prove to be non-functional or show low activity. In addition the native spliceosome activity of yeast is poor, which requires the removal of introns from heterologous protein coding sequences before introduction into this host. These features can result in poor precursor supply or sub-optimal flux through the heterologous pathway, leading to low titres of the product (Billingsley *et al.*, 2016; Kavscek *et al.*, 2015).

Non-canonical bacterial hosts: Bacillus subtilis, Corynebacterium glutamicum and Pseudomonas putida

The expansion of synthetic biology from traditional chassis to other organisms depends primarily on the biotechnological application that is desired. This has led to the identification of production hosts with fast growth, well-characterised regulation and the possibility to engineer using known metabolic networks.

Bacillus subtilis is a GRAS organism with a fast growth rate that has been extensively used for the industrial production of proteins. It has a wide substrate utilisation range, can survive under harsh conditions and produces innate cellulases that can degrade lignocellulosic

material to generate cheap carbon source for large-scale industrial fermentations. It has drawn tremendous attention because of its high isoprene production rate, making it theoretically an ideal host for terpenoid biosynthesis (Julsing *et al.*, 2007; Wagner *et al.*, 2000). It has been successfully used for the engineering of the sesquiterpenoid amorpha-4,11-diene, the precursor of artemisinin (Zhou *et al.*, 2013). However, *B. subtilis*, like all Gram-positive bacteria shows plasmid instability, and being incompletely characterised poses engineering challenges. Despite these shortcomings, a genome-reduced strain of *B. subtilis* was generated and used for the production of heterologous enzymes (Juhas *et al.*, 2014; Li *et al.*, 2016). Also the strengths of promoter candidates that can be used for transgene expression were evaluated recently in *B. subtilis* (Song *et al.*, 2016). These studies allow a deeper understanding of the host and facilitate future work to develop this organism as a production strain using synthetic biology approaches.

Corynebacterium glutamicum is a pigmented Gram-positive bacterium, which has been extensively used for the industrial production of amino acids. It also naturally synthesises the cyclic carotenoid decaprenoxanthin and its glucosides, which makes it a potential chassis for the engineering of plant terpenoids such as the carotenoids β-carotene and zeaxanthin (Heider *et al.*, 2014), and the sesquiterpenoid (+)-valencene (Frohwitter *et al.*, 2014). Through engineering of the isoprenoid pathway and co-expression of plant genes, lycopene and astaxanthin production in *C. glutamicum* was reported in a genome-reduced strain (Baumgart *et al.*, 2013; Heider and Wendisch, 2015).

Pseudomonas putida is a Gram-negative soil bacterium traditionally known for its xenobiotic degradation activity. Like *B. subtilis* and *C. glutamicum*, refactored *P. putida* strains with streamlined genomes were observed to outperform the wild-type strain in heterologous protein production (Lieder *et al.*, 2015). These bacteria are readily adaptable to selective pressures, can relatively easily develop resistance to drugs, and have multiple efflux pumps that contribute to their resistance to toxic compounds, making them good candidates for directed evolution experiments with the aim of producing specialised chemicals through synthetic pathway expression. *P. putida* strains are also well known for their ability to cope with high oxidative stress and in the process to generate high concentrations of reducing equivalents, which are advantageous for the synthesis of several plant specialised metabolites (Kim *et al.*, 2016). In addition to these natural advantages, molecular biology tools, such as inducible promoter systems and fully sequenced genomes of *P. putida* strains allow manipulation and thus boosting of the metabolic capabilities of this host as a microbial

factory. *P. putida* has been used as a production host for rhamnolipids, terpenoids, polyketides and non-ribosomal peptides (Loeschcke and Thies, 2015).

Non-canonical yeast hosts: Yarrowia lipolytica, Pichia pastoris, Hansenula polymorpha and Kluyveromyces lactis

Yarrowia lipolytica is a dimorphic, non-pathogenic and strictly aerobic GRAS yeast with promising properties to be utilised as a host for engineering the production of plant specialised metabolites, particularly because of its oleaginous nature (Markham and Alper, 2015). Several engineering tools have been developed in this host with the aim of utilising it for industrial production of valuable compounds; including an efficient transformation method (Yu *et al.*, 2016), gene editing and insertion into specific genomic loci with CRISPR-Cas9 (Gao *et al.*, 2016a; Schwartz *et al.*, 2016), expansion of the substrate range to utilise sugars that are abundant in lignocellulosic hydrolysates as a carbon source (Ledesma-Amaro *et al.*, 2016; Ledesma-Amaro and Nicaud, 2016), and the identification of genetic elements (promoters, terminators, etc.) for efficient gene expression (Blazeck *et al.*, 2013; Curran *et al.*, 2015; Shabbir Hussain *et al.*, 2016). With these tools *Y. lipolytica* has been engineered for the production of metabolites such as β-carotene and aromatic compounds (Gao *et al.*, 2014; Liu *et al.*, 2015).

Pichia pastoris is a methylotrophic yeast reputed for the industrial production of biopharmaceutical proteins, humanised glycoproteins in particular (Vogl *et al.*, 2013). It can typically achieve high-density fermentations, is tractable to genetic modifications and has highly evolved pathways for posttranslational protein modification. Besides being extensively used for protein production, *P. pastoris* has recently gained attention as a platform cell factory for pathway engineering, which has led to advancements in molecular and genome engineering toolboxes for this host (Kang *et al.*, 2017; Vogl *et al.*, 2014). The production of xanthophyllic compounds astaxanthin, lycopene, β-carotene and canthaxanthin was successfully engineered into *P. pastoris* (Araya-Garay *et al.*, 2012).

Like *P. pastoris*, *Hansenula polymorpha* has been extensively used for humanised glycoprotein engineering. It is a methylotrophic, thermotolerant yeast that can utilise glucose and xylose for ethanol fermentation at temperatures up to 50°C (Suwannarangsee *et al.*, 2010). *H. polymorpha* also possesses tightly-regulated native promoters, which are a useful asset in synthetic pathway engineering (Saraya *et al.*, 2012). Krasovska and co-workers reported a glucose repression-deficient *H. polymorpha* strain in which genes can be induced with methanol in the presence of glucose. This lack of glucose repression is particularly

interesting as the strain can be cultured in an inexpensive glucose-based medium containing a minimal inducing concentration of methanol, instead of an induction medium with high-concentration of methanol as the primary carbon source (Krasovska *et al.*, 2007). Thus far plant natural product biosynthesis has not been engineered in *H. polymorpha*, but it has been successfully implemented for the production of the plant fatty acid γ-linolenic acid (Khongto *et al.*, 2010).

Kluyveromyces lactis is the most similar in native biosynthetic capacities to *S. cerevisiae* and often promoters and toolkits from the latter are used for heterologous gene expression in the former. *K. lactis* also has tightly-regulated native promoters. In addition, there is a high level of transferability in vectors for *K. lactis*, but they have low stability in the host and are lost over a few generations compared to *S. cerevisiae* (Rodicio and Heinisch, 2013).

Each of these non-conventional yeasts present different advantages, commonalities, and differences compared with *S. cerevisiae* and has been concisely summarised elsewhere (Wagner and Alper, 2016). In short, each has diverse catabolic repertoires which enables low cost bioprocesses, can produce high levels of heterologous proteins and are 'Crabtree-negative' as they do not undergo aerobic alcoholic fermentation, unlike *S. cerevisiae*. Nonconventional host systems like the organisms listed above require an expanded synthetic biology toolkit to compete with the well-established conventional production hosts. Highly controllable, programmable and tuneable expression cassettes, targeted recombination and metabolic engineering strategies for multi-gene pathway expression need to be developed and implemented in these hosts for their future application in synthetic biology.

The photosynthetic cell factories

Cyanobacteria and microalgae are a diverse group of aquatic, unicellular photosynthetic microorganisms that have been used as models to study fundamental biological processes (Dorrell and Smith, 2011). Their unicellular physiology, rapid cell division (doubling time of 5-8 hours in *Chlamydomonas* and 4-12 hours in cyanobacteria) and photosynthetic growth render them ideal for the production of a plethora of compounds, including specialty chemicals, pharmaceuticals, food supplements and biofuels. In part, algal biotechnology has been stimulated by the natural capacity of these organisms to synthesise and store compounds of commercial interest, including high-value products such as pigments, vitamins and antioxidants, and bulk compounds like glycerol and triacylglycerols with potential as biofuels (Gimpel *et al.*, 2015). Furthermore, they inhabit diverse aquatic and terrestrial environments,

adopting photosynthetic, heterotrophic and mixotrophic lifestyles. Therefore, unlike crop-based industries, large-scale cultivation of these organisms does not require precious arable-land (Pittman *et al.,* 2011), high water usage (Yang *et al.*, 2011), and input such as fertilisers (Yang *et al*., 2011; Scaife and Smith, 2016).

Cyanobacteria and microalgae are suitable industrial biotechnology platforms (Fig. 4), since like bacteria and yeast they can be grown at scale in controlled and contained environments. However, being photosynthetic and being able to fix carbon from inorganic sources, their cultivation is in principle more sustainable than that of heterotrophic chassis, nonetheless to be certain it is essential to carry out life cycle assessment for any particular process (Kazamia and Smith, 2014). Before cultivation of engineered strains is possible on the very large scales needed for economically profitable production however, many challenges have to be overcome (Klein-Marcuschamer *et al.*, 2013; Scaife and Smith, 2016; Scott *et al.*, 2010). Although these challenges can be addressed with the current genetic and molecular toolkits, the adoption of synthetic biology principles can accelerate the cumbersome process of strain development and optimisation (Fig. 1). To this end, current engineering technologies of photosynthetic cells need to be upgraded and new tools developed to ultimately produce economically competitive cyanobacterial- and algal-sourced biochemicals.

Cyanobacteria as hosts

There are over 100 cyanobacterial species with sequenced genomes, and technical and practical knowledge are enabling metabolic engineering of many of them. Four hosts in particular are widely exploited – *Anabaena* sp. PCC 7120, *Synechocystis* sp. PCC 6803, *Synechococcus* sp. PCC 7002 and *Synechococcus* sp. PCC 7942. Indeed, proof-of-principle metabolic engineering involving directed tailoring of native metabolism, or introduction of heterologous metabolic pathways, has been reported for a range of platform chemicals such as 2,3-butanediol, lactate, more specialist molecules like limonene and even potential biofuel molecules such as ethanol and isoprene (Table 1). These examples of metabolic engineering used classic genetic engineering approaches for the introduction of gene(s). The trial and error nature of selection and use of DNA parts (e.g. promoters, transgenes, terminators) and methodologies (e.g. optimisation for codon usage) consistently resulted in slow strain development and suboptimal outcomes, that can be particularly unpredictable when combined with the occurrence of genetic instability (Jones, 2014). For example, the first report of the heterologous expression of an isoprene synthase (*IspS*) gene in *Synechocystis* sp. PCC6803 (Lindberg *et al.*, 2010) resulted in trace levels of isoprene, a C-5 hydrocarbon that serves as a

feedstock for the rubber industry (Lindberg *et al.*, 2010). To boost isoprene production, codon optimisation increased *IspS* expression by 10-fold, while heterologous expression of the mevalonate pathway in *Synechocystis* improved pathway intermediate abundance, which in turn increased isoprene production by 2.5-fold (Bentley *et al.*, 2014). These improvements were still impeded by metabolic bottlenecks, which were elucidated and additional *IspS* enzymes from different organisms were tested for activity and product yield, allowing rational metabolic engineering to produce 0.54 g gDW^{-1} isoprene in *Synechococcus* sp. PCC7942 (Gao *et al.*, 2016b). The journey to engineer isoprene-producing cyanobacteria has highlighted the slow and time-consuming nature of current approaches towards engineering of a simple pathway that yields a simple molecule. It is therefore not surprising that a synthetic biology approach that combines rational design with systems-level analysis of impact upon introduction of heterologous genes would have the greatest impact on the speed and efficiency of the engineering of multi-step pathways that produce complex molecules like triterpenoids. Recently, synthetic biology approaches have been adopted for engineering of cyanobacteria with P450s, where requirement for reducing equivalents were met by plastidic non-cognate redox partners, photoreduced ferredoxin and flavodoxin (Gnanasekaran *et al.*, 2015; Goñi *et al.*, 2009; Lassen *et al.*, 2014; Wlodarczyk *et al.*, 2016). This novel strategy for re-routing electron partitioning allows engineers to overcome the bottleneck of insufficient reducing power, and design P450-dependent pathways to tap into the ample supply of reducing power of photosynthetic organisms (Nielsen *et al.*, 2016). Looking to the future of synthetic biology, inspiration can be drawn from work in *E. coli*, where systems-level analysis informed by genome-scale models for flux-balance analyses (Chubukov et al., 2016) are proving vital for production of metabolites at a commercial scale. Therefore the potential for similar genome-scale models, now available for *Cyanothece* sp. ATCC 51142, *Synechococcus* sp. PCC 7002, and *Synechocystis* sp. PCC 6803 (Knoop *et al.*, 2013; Nogales *et al.*, 2012; Vu *et al.*, 2012), cannot be overstated.

Microalgae as hosts

Compared with cyanobacteria, fewer microalgal genomes have been fully sequenced and annotated. Despite this, a wide range of microalgae has been genetically engineered, although it is routine only for a few species (Table 1). A particular issue for microalgae is that currently nuclear transformation is by random integration. Although homologous recombination has been reported for a few species, such as *Ostreococcus* and *Nannochloropsis*, this is not routine, and moreover is not possible in the most widely used algal hosts such as

Chlamydomonas reinhardtii and *Phaeodactylum tricornutum*. A further complication in *C. reinhardtii* is extensive transgene silencing. Strains of this green alga, UVM4 and UVM11, identified from a UV mutagenized population, due to their propensity for increased transgene expression (Neupert *et al.*, 2009), are helping to overcome this issue.

Another issue among microalgae is to improve the efficiency of transformation and overcome complications of DNA fragmentation caused by biolistics and electroporation. To this end, Karas and co-workers have developed extrachromosomal vectors or episomes for transgene expression in *P. tricornutum* (Karas *et al.*, 2015). Coupled with a conjugation method for transfer of the construct assembled in *E. coli* directly to diatom cells, transformation frequency has been markedly improved in this species. Of wider interest is the utility of episomes for effective transformation of another diatom species *Thalassiosira pseudonana* (Karas *et al.*, 2015).

In addition to nuclear transformations, in photosynthetic organisms the chloroplast genome is also amenable to genetic manipulation (Manuell *et al.*, 2007). The chloroplast genomes of >20 species are now accessible to genetic modification (Day and Goldschmidt-Clermont, 2011). Being the site of major anabolic pathways (e.g. carotenoid and fatty-acid biosynthesis), from an engineering perspective the ability to engineer this cellular compartment is of significant biotechnological importance, and the methodology is well established for several higher plants (Bock, 2015). Similarly, the *C. reinhardtii* chloroplast has been widely manipulated but to date chloroplast transformation is not routine for any other microalga (Purton *et al*., 2013). Compared with nuclear transformations, transformation of the chloroplast genome for transgene expression provides unique advantages, most importantly the ability to target transgene insertion via homologous recombination. In addition, high-level expression of transgene and compartmentalised over-accumulation of proteins containing disulphide bonds occur readily and undesired glycosylations are prevented (Tissot-Lecuelle *et al.*, 2014). Moreover, well-established bacterial resources have been demonstrated to function in the *C. reinhardtii* chloroplast, including the *aadA* antibiotic resistance marker (Goldschmidt-Clermont M, 1991), the β-glucuronidase reporter (Ishikura *et al*., 1999), and the lac repressor system (Kato *et al.*, 2007). Given that many native promoters are under strict regulation of essential photosynthetic processes, access to bacterial regulatory resources is particularly invaluable. Analogous to cyanobacteria, re-routing of electron partitioning has also been successfully performed in *Chlamydomonas,* where transplastomic expression of P450 reductase and a bifunctional *cis*-abienol synthase resulted in the

translocation of these enzymes into the chloroplast and production of *cis*-abienol, a precursor used in the perfume industry (Gangl *et al.*, 2015; Nielsen *et al.*, 2016; Nielsen *et al.*, 2013).

Synthetic biology approaches for engineering unicellular cell factories

Our growing understanding of the molecular basis of biological functions has permitted the engineering of biological systems for novel purposes. Modern genetics, molecular, structural, systems and microbiology tools have facilitated the rational engineering of natural systems, the unicellular microbes in particular. Understanding relationships between biomolecule structure and function allows the design of novel nucleic acid and protein sequences, incorporating non-canonical building blocks and thus assembling fully artificial systems with customised properties (Marner, 2009) based on the principles of synthetic biology. In this section we discuss the most promising technologies that are valuable for organism engineering (Fig. 2).

Standardising parts and assembly approaches

Well-established systems that enable targeted, temporal and spatial expression of endogenous and heterologous transgenes are at the forefront of synthetic biology driven engineering of microorganisms. The fundamental concept requires the use of standard parts that are readily assembled and easily interchanged, for engineering a particular organism, but ideally also between organisms for orthogonality (Scaife and Smith, 2016). For *E. coli* and yeast hosts knowledge and tools have been amassed over decades. Unicellular photosynthetic organisms such as cyanobacteria and microalgae are fast growing and can be cultured in a similar fashion to bacteria and yeast, and taking a synthetic biology approach now will facilitate their development as production platforms. In contrast, the slow generation time of higher plants can hamper advanced metabolic engineering, and so one approach would be to use the unicellular photosynthetic hosts for the generation and testing of plant DNA parts and technologies that can ultimately be adopted for use in higher plant biotechnology. Indeed advanced molecular biology techniques such as ZFN, TALENs and CRISPR/Cas9 have been developed, tested and optimised for many unicellular photosynthetic organisms. Examples of advances towards molecular tool-optimisation are many and reviewed elsewhere (Jinkerson and Jonikas, 2015; Scaife and Smith, 2016).

Nonetheless, in all cases the development of standardised DNA parts that adhere to a common syntax for rapid assembly is revolutionising the design and building of complex multi-gene heterologous pathways. The first of these was BioBricks, developed by the iGEM

competition (Knight, 2003), where the use of specific restriction enzymes at the end of each part enabled them to be assembled sequentially into longer and longer sequences. This was followed rapidly by Gibson assembly (Gibson *et al.*, 2009), which avoids the introduction of scars between the parts, and also enables several fragments to be assembled simultaneously. Further refinement came from methods such as Golden Gate (Engler *et al.*, 2008), which combined the ease of restriction-based cloning and simultaneous, seamless assembly of multiple parts.

Although the DNA parts that make up a genetic circuit may vary between organisms, through adoption of a syntax such as that proposed for Modular Cloning using Golden Gate assembly, libraries of synthetic parts are being developed and made available (Engler and Marillonnet, 2014; Lee *et al.*, 2015; Patron, 2014; Weber *et al.*, 2011). Access and ease by which expression cassettes can be designed and assembled, is enabling high throughput systematic testing of parts in the target organisms. In many instances, however the desired "fully predictable" parts have yet to be discovered. Indeed, successful metabolic engineering depends on expression of a transgene in a predictable fashion so as to modify metabolism reliably. Tightly regulated and inducible protein synthesis are important determinants of product yield, and help avoidance of unintended effects on endogenous metabolism that can culminate in pleiotropic defects.

Regulation of transgene expression by modular effectors

Metabolic flux in a cell is regulated by a series of interwoven regulatory controls, at the levels of transcription, translation, protein stability, and enzyme activity. In bacterial systems transcriptional regulation is probably one of the most important processes, and in a synthetic biology context, this may be the simplest access point to alter cellular flux. Accordingly, metabolic engineering efforts have long relied on the discovery and characterisation of effective promoters, in both bacterial and eukaryotic systems. More recently, other regulatory elements such as riboswitches are being employed as modular effectors that can be easily permuted with the coding sequence and other parts of the transcriptional units that are engineered into the host chassis (Berens et al., 2015; McKeague et al., 2016).

Selecting promoters is a key component of the design cycle, where a wide range of expression capacities of several orders of magnitude is required for the optimal expression of genes in a pathway. Over the years several naturally existing promoters have been characterised in several bacterial and yeast hosts, but they do not represent the dynamic range required for flux optimisation in a pathway. Promoter engineering aims at modulating the

transcriptional capacity of an existing promoter by altering its DNA sequence to generate synthetic promoter libraries. To achieve this, natural promoters are classically subjected to molecular approaches such as saturation mutagenesis of promoter regions (Hammer et al., 2006; Solem and Jensen, 2002), error-prone PCR (Hammer et al., 2006; Hauk et al., 2016; Reed et al., 2012; Tyo et al., 2011) or hybrid promoter engineering (Blazeck et al., 2012; Blazeck et al., 2011; Blazeck et al., 2013; Xu et al., 2014), and the resulting promoter variants are screened with quantifiable read-outs to define promoter strengths. However, this approach is only applicable to relatively short promoter sequences and for certain experimental purposes. Computational methods on the other hand allow prediction of promoter activity prior to functional characterisation, thereby reducing the number of promoter variants to be tested *in vivo*. *In silico* models like the position weight matrix (Sinha, 2006), partial least squares regression (Jonsson et al., 1993), and artificial neural networks (Meng et al., 2013) allow elucidating the effect of individual DNA bases and motifs on promoter activity.

In addition to validating novel promoter engineering strategies and defining promoter strengths, the synthetic promoters are also successfully implemented for natural product biosynthesis. In *E. coli*, a single promoter containing three unique *cis*-regulatory elements that can recognise three transcription factors simultaneously was developed using an *in silico* optimisation algorithm, to allow multiple environmental signals to modulate gene expression differently (Amores et al., 2015). Through a promoter engineering approach, the production of lycopene in *E. coli* was improved by modulating the expression of heterologously expressed genes (Shen *et al.*, 2015). In *S. cerevisiae*, synthetic promoters inducible by low pH were engineered from pH-unresponsive promoters (Rajkumar et al., 2016). Inducible and constitutive synthetic minimal promoters <120bp in length, have been designed by combining a short robust and generic core element with strong upstream activating sequences of ~10bp (Redden and Alper, 2015). In a distinct approach, nucleosome architecture was used to predict high activity promoters computationally. As nucleosome occupancy along arbitrary DNA sequences is predictive, promoters can be (re)designed for defined activity (Curran et al., 2014). Recently, a modular synthetic multi-gene promoter and transcription factor system consisting of adjustable and predictable expression levels ranging from negligible to very strong, independent from externally added inducers, was also developed for multiple gene pathway expression (Rantasalo et al., 2016).

Similarly, in microalgae such as *C. reinhardtii*, decades of research have elucidated several fundamental processes that may be built upon to facilitate transgene expression (Scaife and Smith, 2016). Characterised promoters used for controlled transgene expression

include those regulated by light (Fischer and Rochaix, 2001), nitrate (Ohresser *et al.*, 1997), copper (Quinn and Merchant, 1995), nitrogen (Matthijs *et al.*, 2016) and vitamin B12 (Helliwell *et al.*, 2014). In addition to high-expression native promoters (e.g. RuBisCo small subunit gene RBCS2 and PsaD), the development and use of synthetic promoters (e.g. heat shock protein 70A promoter fused to the RBCS2 promoter) for constitutive expression of transgenes has resulted in significant improvement in the frequency of stable transformants obtained (Schroda *et al.*, 2000).

In addition to promoters, functional non-coding RNA molecules known as riboswitches have been developed as synthetic regulatory elements in many bacteria and eukaryotes (Roth and Breaker, 2009). In general, the ligand-sensing domain (aptamer) of riboswitches is combined with a regulatory domain, or expression platform, which can control gene expression in a variety of ways. For example, in bacteria direct binding of a specific ligand to the aptamer domain can be used to attenuate transcription termination or translation initiation (Roth and Breaker, 2009), whereas in eukaryotes (plants, fungi as well as algae), ligand binding causes alternative splicing, which prevents protein production for example by introduction of an in-frame stop codon. In terms of practicality for large-scale application, the riboswitches can be regulated by nanomolar concentrations of ligand (Croft *et al.*, 2007; Moulin *et al.*, 2013). In *C. reinhardtii*, the riboswitch within the *THI4* gene of the thiamine biosynthesis pathway has already been developed into a transgene expression regulatory tool (Croft *et al.*, 2007; Moulin *et al.*, 2013). This system has also been adapted into a novel synthetic regulatory circuit to regulate the expression of plastid genes (Ramundo *et al.*, 2013). Riboswitches have also been developed as tools for the regulation of gene expression in cyanobacteria. A theophylline-dependent riboswitch has been shown to act as a stringent inducible protein production system in *S. elongatus* PCC 7942, able to achieve binary on/off expression states (Nakahira *et al.*, 2013). Similar to other riboswitch systems, the level of protein production can be fine-tuned with different theophylline concentrations, an attribute that highlights the flexibility of such expression platforms when integrated within genetic circuits.

Protein engineering

Natural proteins and enzymes are often engineered to meet the rigorous needs of industrial processes. It is also a powerful tool for synthetic biology where tailored protein properties suit the requirements of synthetic metabolic networks. Protein engineering is a form of synthetic biology that contributes significantly to metabolic engineering by improving enzymatic

activity, altering the substrate and product specificity of enzymes, managing metabolic flux using enzyme scaffold engineering, and by modulating the regulatory elements that tune protein levels (Foo *et al.*, 2012; Li, 2012; Li and Cirino, 2014). A key factor in protein engineering is the ability to precisely determine the three-dimensional structure of proteins, as they provide essential details about biological function and mechanism of action. Computational techniques aid understanding relationships between structure and function of proteins, which subsequently allow the modelling and design of primary amino acid sequences that will ultimately result in novel or modified enzyme functions. Computer-aided protein modelling and designing approaches, using either homology-based or *ab initio* methods, have been summarised elsewhere (Kelchtermans *et al.*, 2014; Khan *et al.*, 2016; Kingsley and Lill, 2015; Wei and Zou, 2016).

Directed evolution is a powerful tool for engineering proteins. Randomised variants of a gene encoding a protein of interest are expressed in a suitable host and combined with an appropriate screening method to identify mutant proteins with desired properties. With this approach, properties of proteins have been redesigned or improved through iterative cycles of mutagenesis, and amplification of the desired mutants to accumulate beneficial mutations in the final protein (Cobb *et al.*, 2013). With regards to plant specialised metabolism, several classes of enzymes have been engineered for modulating their catalytic efficiency, specificity and/or activity. One such class of enzymes is the plant uridine diphosphate-dependent glycosyltransferase (UGT) family, the members of which catalyse the addition of sugar residues to natural product structures such as flavonoids or alkaloids. Glycosylation plays a key role in enhancing solubility and stability of specialised metabolites in the aqueous environment of a cell, as well as detoxification, thus facilitating their accumulation and storage (Liang *et al.*, 2015). Advancements in plant UGT structural studies has bettered our understanding of catalytic mechanisms, which in turn allowed directed engineering for enhanced activity, altered regiospecificity and glycosylation pattern, to produce a plethora of bioactive glycosides (Wang, 2009). P450 enzymes are another superfamily that play critical roles in plant specialised metabolite biosynthesis. They display a wide substrate range and catalyse a diverse array of chemical transformations. Their ability to introduce oxygen atoms stereo- and regio-specifically into complex molecules, and to activate a backbone molecule through primarily oxygenation reactions, makes them particularly valuable for synthetic biology applications. For instance, soluble bacterial P450s have been engineered to oxidise diverse chemical scaffolds for the production of high value chemicals, highlighting the malleability as well as robustness of these proteins (Girvan and Munro, 2016).

Artificial protein scaffolding is another versatile approach (Fig. 2, see protein engineering box), that allows mimicking of natural machineries for multi-enzyme pathways, and has been extensively applied in synthetic biology. Generally in nature multi-cascade reactions occur in multi-enzyme complexes where micro-compartments facilitate individual reactions. This phenomenon known as metabolon formation or metabolic channelling allows avoidance of flux imbalances, diffusion and loss of intermediates, release of toxic pathway intermediates, and interference from competing metabolic networks (Bassard *et al.*, 2017; Jørgensen *et al.*, 2005). Therefore, the assembly of active centres of sequential pathway enzymes in close proximity is of great interest to increase overall catalytic efficiency of a heterologous pathway (Conrado *et al.*, 2008). In this phenomenon of metabolic channelling, enzymes are ordered sequentially to ensure that the local concentration of pathway intermediates is high around the enzyme complex. As most plant specialised metabolite biosynthetic pathways involve multiple enzymes, metabolic channelling in production hosts may well be an effective means of optimising production and increasing titres (Singleton *et al.*, 2014). A variety of target molecules have been synthesised through the scaffolding of enzymes with peptide motifs and their cognate adaptor domains (Horn and Sticht, 2015; Proschel *et al.*, 2015). In *E. coli* scaffolding has been effectively used for the increased production of glucaric acid (Moon *et al.*, 2010) and mevalonate, the precursor of sesqui- and triterpenoids (Dueber *et al.*, 2009). Similarly, in yeast resveratrol production was enhanced by the scaffolding of 4-coumarate:CoA ligase and stilbene synthase (Wang and Yu, 2012). Artificial scaffolding additionally offers the possibility of evaluating the effect of different enzyme ratios on product yield. For instance, the titre of glucaric acid produced in *E. coli* was more strongly affected by inclusion of scaffold interaction domains for the upstream enzymes, than for the later enzymes of the pathway (Moon *et al.*, 2010).

Like artificial scaffolding, the direct fusion of proteins has also been used to improve catalytic efficiency of enzymes. In nature for instance, the natural fusion of a P450 to an aldo-keto reductase in *Papaver* spp. was discovered to be essential for the synthesis of morphinan alkaloids (Farrow *et al.*, 2015; Winzer *et al.*, 2015). Artificial fusion of the *Panax ginseng* P450 protopanaxadiol synthase (PPDS) and the *Arabidopsis thaliana* P450 reductase ATR1 was shown to improve catalytic activity of the PPDS (Zhao *et al.*, 2016). Likewise, the in-frame fusion of a red clover P450 isoflavonone synthase with a rice P450 reductase after the removal of the membrane binding domains enabled the conversion of naringenin to genistein in *E. coli* (Kim *et al.*, 2009).

Biosensors

The quantification of plant natural products in a production host is often challenging, as they are seldom associated with a measurable phenotype. However, natural biological regulators can serve as molecular reporters in the presence of specific ligands, and they can be applied as biosensors to correlate the amount of ligand to the read out of a reporter protein. Typically biosensors are linked to a colorimetric output to report the concentration of metabolites. This enables the screening and selection of better performing strains in a high-throughput fashion. Fluorescent reporter proteins are broadly employed because of their compatibility with fluorescence-activated cell sorting (FACS), which allows for efficient screening in a short time (Binder *et al.*, 2012). The transcriptional regulators in bacteria that activate metabolite degradation and carbon assimilation pathways upon detection of plant flavonoids, are naturally existing regulatory mechanisms that can be applied as biosensors (Hirooka and Fujita, 2011; Marin *et al.*, 2013; Rao and Cooper, 1994). By combining such transcription factors with a fluorescent reporter signal, flavonoid biosensors for the quantification of naringenin, quercetin and kaempferol were successfully constructed in *E. coli* (Siedler *et al.*, 2014). Small molecule sensors can also be coupled to a selective advantage, so as to transform biosynthetic phenotypes into a fitness difference. Raman and co-workers used a sensor domain responsive to a target molecule, to control the expression of a reporter gene necessary for the survival of the strain under selective conditions. By converting the intracellular presence of the target molecule into a fitness advantage, a general strategy was developed for the enrichment of rare high producers within a population of variants. Using this approach the production of naringenin and glucaric acid in *E. coli* was improved significantly through evolution guided pathway optimisation (Raman *et al.*, 2014). In a more recent study, the use of ligand binding domains was proposed as a general strategy to construct small molecule biosensors for any target molecule. Feng and co-workers converted a single high-affinity ligand binding domain into multiple specific conditionally-stable scaffolds, which when fused to transcription factors generated biosensors that respond to their target ligands, the steroids digoxin and progesterone (Feng *et al.*, 2015).

Natural biosensors are ubiquitous detectors that perceive intra- or extracellular environmental signals like small molecules, ions, or physical parameters. They respond to detected signals by modifying cellular activity at a transcriptional, translational, or protein activity level. For instance, *E. coli* has 230 metabolite responsive transcription factors that sense and respond to sugars and their phosphates, amino acids and lipids (Binder *et al.*, 2012). The ability of biosensors to switch between an ‘on and off’ status in response to stimuli,

makes them applicable for various synthetic biology purposes. Biosensors that can sense and respond to small molecule stimuli within cells have a wide range of applications, including regulation and optimisation of metabolic pathways, measurement and imaging of metabolite concentration, and monitoring of environmental toxins (Hassani *et al.*, 2016; Mehrotra, 2016; Paige *et al.*, 2012; Raman *et al.*, 2014; Rogers *et al.*, 2016; Zhang *et al.*, 2012; Zhang and Keasling, 2011). There are several control elements within a cell that can modulate gene expression in response to a stimulus and therefore can be employed as biosensors, including transcriptional regulatory proteins, RNA based genetic riboswitch regulators, and functional enzymes (De Paepe *et al.*, 2016; Mehdizadeh Aghdam *et al.*, 2016; Palchetti, 2016). Transcription factors play a key role in biosensor design as they can be easily incorporated into synthetic circuits to control gene expression conditionally. Transcription factor based regulatory biosensors are particularly attractive due to the modularity of their constituent parts, the DNA binding domain/promoter pair, the ligand binding domain, the transcription activation domain, the reporter and the host strain, all of which can be modified according to the application (De Paepe *et al.*, 2016; Mahr and Frunzke, 2016).

In a distinct approach, rather than using fluorescent protein production as the output, an enzyme-coupled biosensor was developed in yeast to identify a tyrosine hydroxylase that could efficiently convert L-tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA), a pathway intermediate in the synthesis of benzylisoquinoline alkaloids (BIA). For this a plant biosynthetic enzyme that converts L-DOPA to a yellow fluorescent pigment was introduced into *S. cerevisiae* as a simple quantifiable biosensor read out. Using this biosensor an efficient tyrosine hydroxylase was identified from a mutant library of tyrosine hydroxylases, and used to synthesise the BIA intermediates (*S*)-reticuline and (*S*)-norcoclaurine from glucose in yeast (DeLoache and Russ, 2015).

Biosensors are an attractive phenotype screening platform, with promising design strategies that make their applicability in industrial processes a reality for the future. For instance, biosensors have been successfully used for the real-time monitoring of metabolite production (Rogers and Church, 2016), and multiple biosynthetic components have been screened in a chassis using multiple robust sensors that have low false positive rates and wide operational ranges (Rogers *et al.*, 2015). The lack of sensor domains for all industrially valuable compounds, makes biosensor directed metabolic engineering in production hosts challenging. However, further characterisation of biosensor systems and protein engineering is expected to help address these concerns and provide solutions for the field.

Minimal genomes

The natural variation in genome sizes across species has raised questions about the size of minimal genomes and the essential set of genes required for maintaining cellular life (Glass *et al.*, 2006). Through gene deletion approaches, researchers have generated smaller and increasingly stable bacterial genomes showing that a large proportion of the genomes can be readily deleted without affecting growth and survival of the cells. In *E. coli* over 15% genome reduction was achieved by deleting transposable elements and horizontally derived genes without significantly affecting growth (Csorgo *et al.*, 2012; Mizoguchi *et al.*, 2007; Umenhoffer *et al.*, 2010). Likewise genome reductions have also been achieved in *Schizosaccharomyces pombe* and *B. subtilis* (Ara *et al.*, 2007; Giga-Hama *et al.*, 2007).

Microbial production chassis during evolution have accumulated many genes other than those essential for growth and survival, to endure ever-changing environmental challenges. Amongst these are genes that are non-essential or even harmful for the industrial production of compounds with commercial value, like the genes that encode for enzymes that degrade the desired product or genes that code for biosynthetic pathways of undesired products which make downstream isolation of the desired compound demanding (Fujio, 2007). Production chassis with minimal genomes where the unnecessary and harmful genes are deleted, thus prove to be ideal for industrial production of commercially valuable compounds. The industrial actinobacteria *Streptomyces avermitilis*, with its endogenous specialized metabolism deleted, serves as an efficient production host for the expression of exogenous biosynthetic genes. Through the expression of a codon-optimised amorpha-4,11-diene synthase, the production of amorpha-4,11-diene was achieved in the genome-minimised strain of *S. avermitilis* optimised for terpenoid production (Komatsu *et al.*, 2010).

Similarly, a genome reduced strain of *S. cerevisiae* is being generated by an international consortium of researchers under the Sc2.0 project (http://syntheticyeast.org/sc2-0/). The designer chromosome SynIII, which has approximately one-sixth of the base pairs of the wild-type chromosome III, was the first to be successfully generated and yeast cells carrying either of the two chromosomes were phenotypically indistinguishable (Annaluru *et al.*, 2014). More recently, the consortium reported a highly modified strain with the genome reduced in size by nearly 8% (Richardson *et al.*, 2017). The genome reduced yeast strain would be an ideal host for the introduction of large biosynthetic pathways, like that of the alkaloids, for the production of multi-gene plant natural product pathways.

Perspective

An ever increasing number of tools and strategies are being developed and applied to engineer and optimise the challenges faced for the production of high-value specialised metabolites in unicellular hosts. Advances in synthetic biology tools, development of high throughput computational engineering approaches and discovery of new molecular biology phenomena improve our ability to refactor biosynthetic pathways to meet the global demand for valuable natural products. The broadening of synthetic biology hosts from the canonical ones to others based on the biotechnological application sought, has allowed the identification of key features that make a good chassis. These primarily include a convenient lifestyle, rapid growth, well characterised regulatory elements and engineering tools, useful interactions with the environment and an accurate map of the hosts' metabolic network. In addition there is a need for reliable prediction of the physiological status of the chassis using metabolic models, whilst considering resource availability and investment versus product output. The dynamic description of cells and the available toolkits therefore form an integral part of efficiently engineering smart cell factories of the future. Although the metabolic capacities of unicellular hosts form the basis for many biotechnological advances, there are limits to what can be achieved with single strains. There is growing focus on the metabolic potential of microbial consortia to perform complex processes that are inefficient or impossible with a single species. This realisation of the potential of synthetic ecologies, necessitates better understanding of the key components that influence stable coexistence of multiple organisms (Escalante *et al.*, 2015). Incorporating evolutionary and ecological perspectives into the design of microbial consortia is further relevant to achieve stable and sustainable systems.

Acknowledgements

This work has been supported by the Special Research Funds from Ghent University (postdoctoral fellowship to T.M.) and the UK Biotechnology and Biological Sciences Research Council (BBSRC) grant reference BB/M018210/1.

Figure Legends

Figure 1: Schematic representation of the cyclical synthetic biology workflow for metabolic engineering. The workflow is standardised regardless of the DNA components or the host organism, and progresses through iterative cycles of the Design-Build-Test-Learn process to achieve predictable Outputs. In the Design phase, the engineer defines the goal, identifies the pathway(s) and gene(s) of interest, selects host-specific regulatory components (parts) that enable the desired expression of transgene(s), and opts for an assembly approach. The design is implemented in the Build phase and introduced into the host. Evaluating the engineered strains for performance of parts (transcripts and proteins) and production of target molecule(s) generates Test data, which is used during the Learn phase to analyse the components, providing an opportunity to deconstruct complex part combinations and generate knowledge, which can be used for production of tangible and predictable Output(s).

Figure 2: A schematic overview of the latest synthetic biology tools available for engineering high value plant natural products in unicellular industrial hosts. For each emerging tool, the basic design principles behind a selection of strategies are depicted.

Figure 3: A concise flow diagram depicting the key steps in engineering unicellular chassis for the production of high value plant natural products.

Figure 4: A visual guide to the key features to consider for chassis selection for the engineering and production of plant natural products in heterologous hosts.

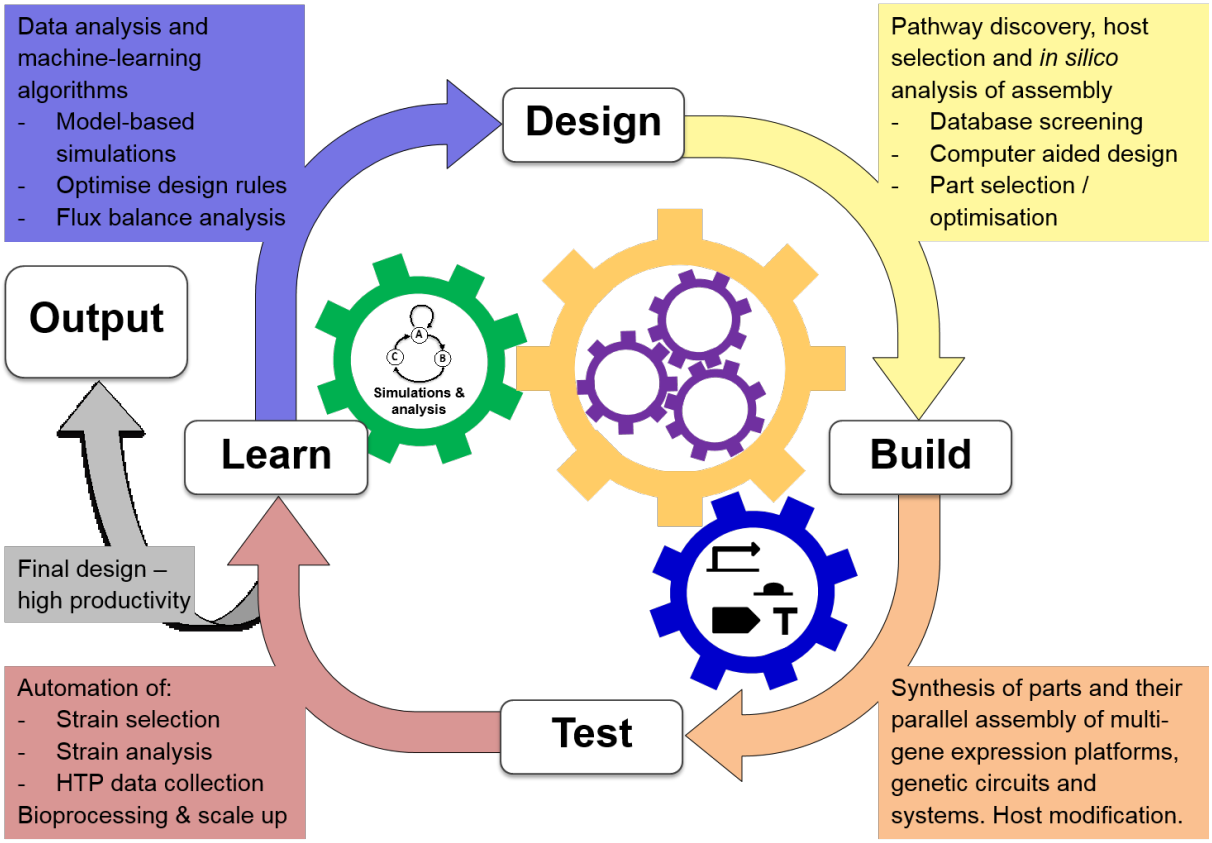


Figure 1: Schematic representation of the cyclical synthetic biology workflow for metabolic engineering. The workflow is standardised regardless of the DNA components or the host organism, and progresses through iterative cycles of the Design-Build-Test-Learn process to achieve predictable Outputs. In the Design phase, the engineer defines the goal, identifies the pathway(s) and gene(s) of interest, selects host-specific regulatory components (parts) that enable the desired expression of transgene(s), and opts for an assembly approach. The design is implemented in the Build phase and introduced into the host. Evaluating the engineered strains for performance of parts (transcripts and proteins) and production of target molecule(s) generates Test data, which is used during the Learn phase to analyse the components, providing an opportunity to deconstruct complex part combinations and generate knowledge, which can be used for production of tangible and predictable Output(s).

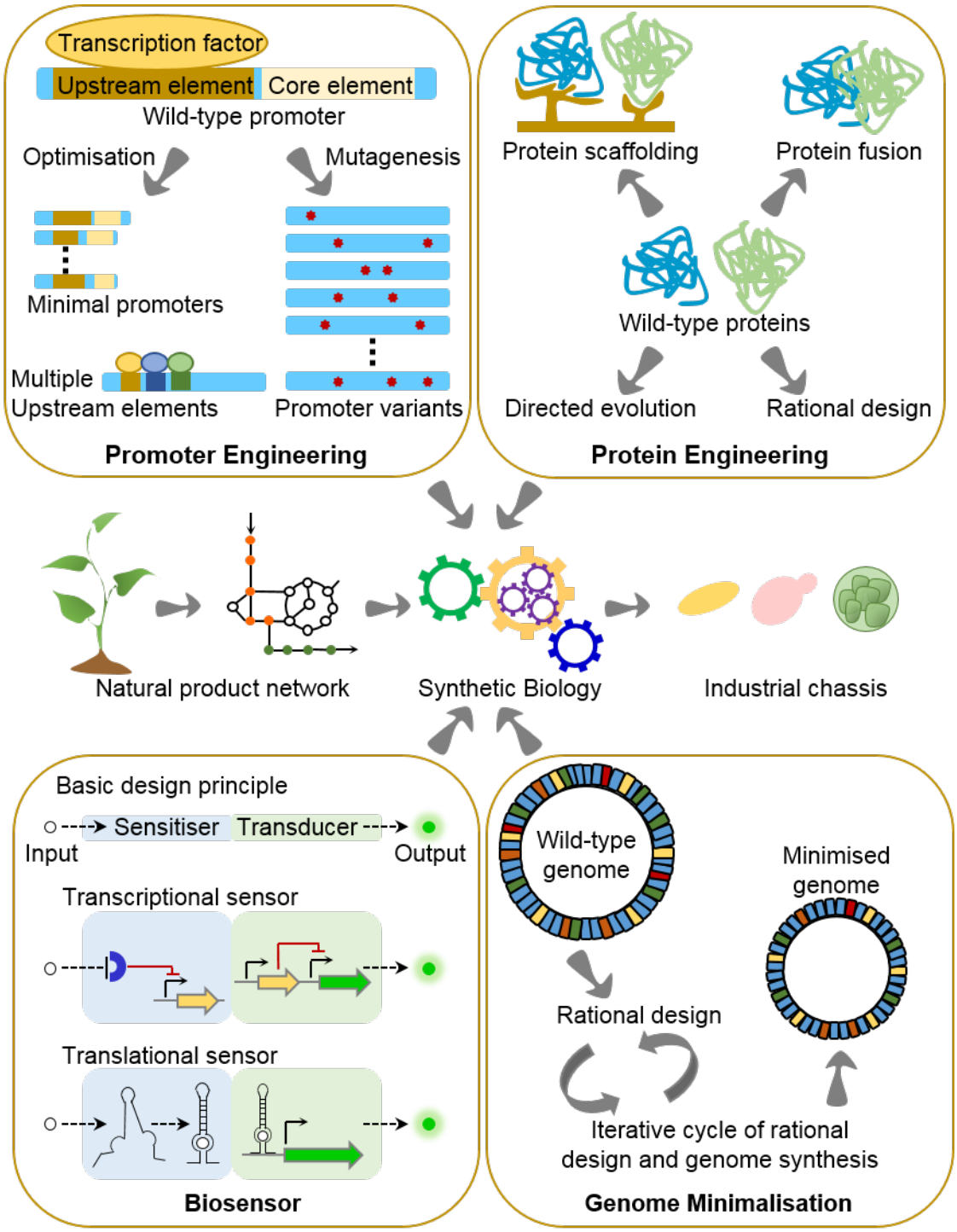


Figure 2: A schematic overview of the latest synthetic biology tools available for engineering high value plant natural products in unicellular industrial hosts. For each emerging tool, the basic design principles behind a selection of strategies are depicted.

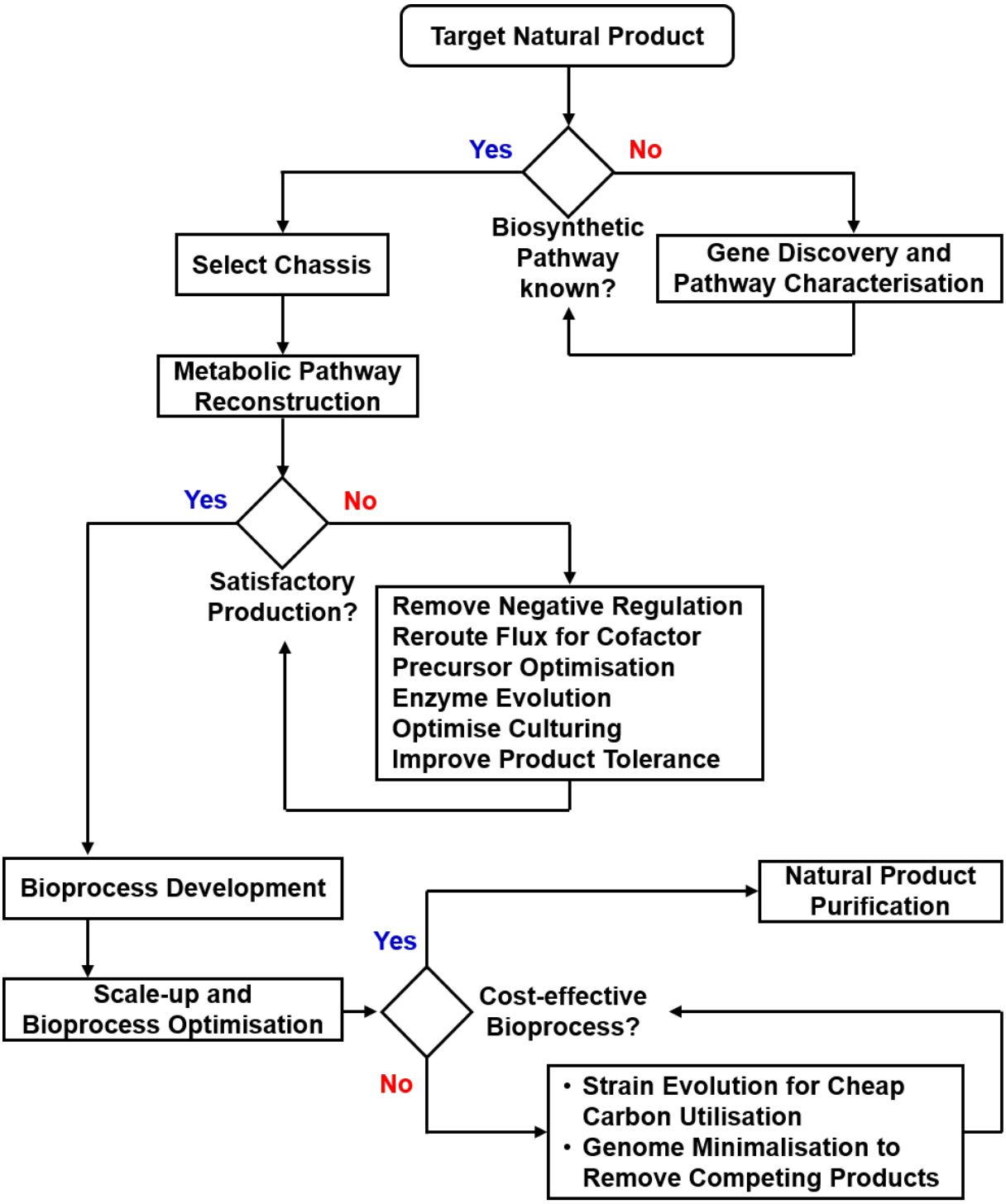


Figure 3: A concise flow diagram depicting the key steps in engineering unicellular chassis for the production of high value plant natural products.

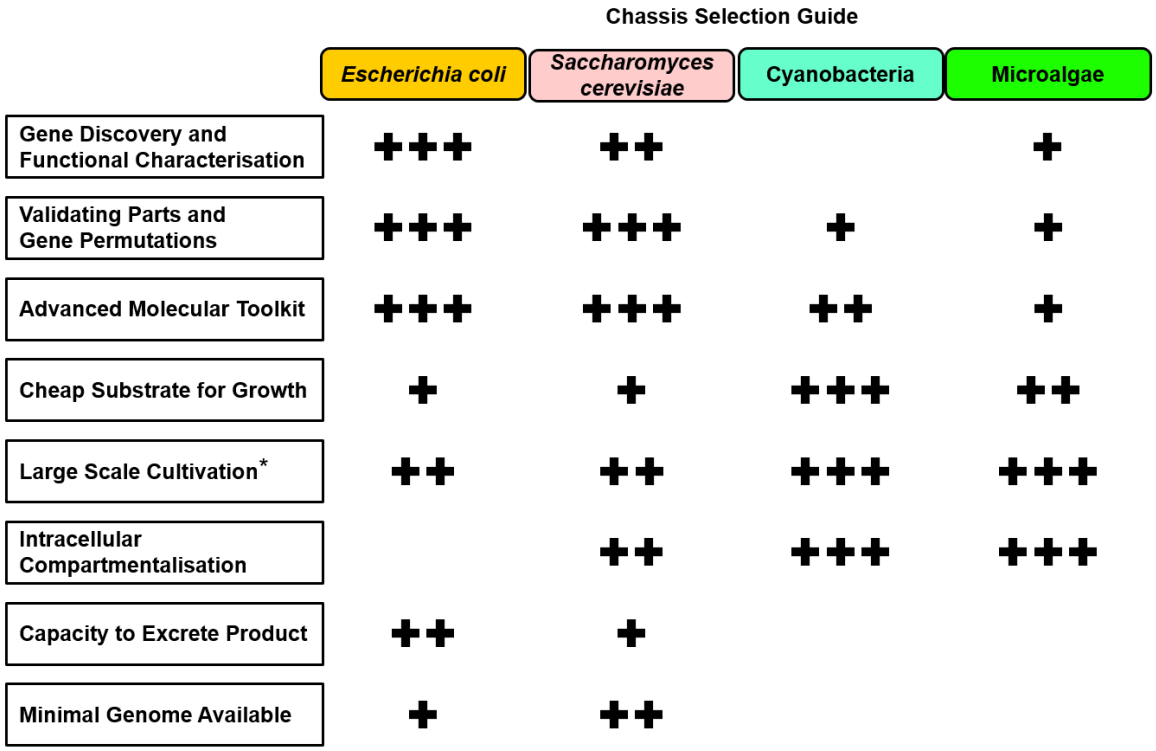


Figure 4: A visual guide to the key features to consider for chassis selection for the engineering and production of plant natural products in heterologous hosts.

References

Amores GR, Guazzaroni ME, Silva-Rocha R. 2015. Engineering Synthetic *cis*-Regulatory Elements for Simultaneous Recognition of Three Transcriptional Factors in Bacteria. *ACS Synthetic Biology* **4**, 1287-1294.

Angermayr SA, van der Woude AD, Correddu D, Vreugdenhil A, Verrone V, Hellingwerf KJ. 2014. Exploring metabolic engineering design principles for the photosynthetic production of lactic acid by *Synechocystis* sp. PCC6803. *Biotechnology for Biofuels* **7**, 99.

Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM, Cai Y, Zeller K, Agmon N, Han JS, Hadjithomas M, Tullman J, Caravelli K, Cirelli K, Guo Z, London V, Yeluru A, Murugan S, Kandavelou K, Agier N, Fischer G, Yang K, Martin JA, Bilgel M, Bohutski P, Boulier KM, Capaldo BJ, Chang J, Charoen K, Choi WJ, Deng P, DiCarlo JE, Doong J, Dunn J, Feinberg JI, Fernandez C, Floria CE, Gladowski D, Hadidi P, Ishizuka I, Jabbari J, Lau CY, Lee PA, Li S, Lin D, Linder ME, Ling J, Liu J, Liu J, London M, Ma H, Mao J, McDade JE, McMillan A, Moore AM, Oh WC, Ouyang Y, Patel R, Paul M, Paulsen LC, Qiu J, Rhee A, Rubashkin MG, Soh IY, Sotuyo NE, Srinivas V, Suarez A, Wong A, Wong R, Xie WR, Xu Y, Yu AT, Koszul R, Bader JS, Boeke JD, Chandrasegaran S. 2014. Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55-58.

Ara K, Ozaki K, Nakamura K, Yamane K, Sekiguchi J, Ogasawara N. 2007. *Bacillus* minimum genome factory: effective utilization of microbial genome information. *Biotechnology and Applied Biochemistry* **46**, 169-178.

Araya-Garay JM, Ageitos JM, Vallejo JA, Veiga-Crespo P, Sanchez-Perez A, Villa TG. 2012. Construction of a novel *Pichia pastoris* strain for production of xanthophylls. *AMB Express* **2**, 24.

Atsumi S, Higashide W, Liao JC. 2009. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature Biotechnology* **27**, 1177-1180.

Awan AR, Shaw WM, Ellis T. 2016. Biosynthesis of therapeutic natural products using synthetic biology. *Advanced Drug Delivery Reviews* **105**, 96-106.

Bassard JE, Møller BL, Laursen T. 2017. Assembly of Dynamic P450-Mediated Metabolons-Order Versus Chaos. *Current Molecular Biology Reports* **3**, 37-51.

Baumgart M, Unthan S, Ruckert C, Sivalingam J, Grunberger A, Kalinowski J, Bott M, Noack S, Frunzke J. 2013. Construction of a prophage-free variant of *Corynebacterium*

glutamicum ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Applied and Environmental Microbiology* **79**, 6006-6015.

Bentley FK, Zurbriggen A, Melis A. 2014. Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Molecular Plant* **7**, 71-86.

Berens C, Groher F, Suess B. 2015. RNA aptamers as genetic control devices: the potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnology Journal* **10**, 246-257.

Berla BM, Saha R, Immethun CM, Maranas CD, Moon TS, Pakrasi H. 2013. Synthetic biology of cyanobacteria: unique challenges and opportunities. *Frontiers in Microbiology* **4**, 1-14.

Biggs BW, Lim CG, Sagliani K, Shankar S, Stephanopoulos G, De Mey M. 2016. Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 3209-3214.

Billingsley JM, DeNicola AB, Tang Y. 2016. Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. *Current Opinion in Biotechnology* **42**, 74-83.

Binder S, Schendzielorz G, Stabler N, Krumbach K, Hoffmann K, Bott M, Eggeling L. 2012. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biology* **13**, R40.

Blazeck J, Garg R, Reed B, Alper HS. 2012. Controlling promoter strength and regulation in Saccharomyces cerevisiae using synthetic hybrid promoters. *Biotechnology and Bioengineering* **109**, 2884-2895.

Blazeck J, Liu L, Redden H, Alper H. 2011. Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Applied and Environmental Microbiology* **77**, 7905-7914.

Blazeck J, Reed B, Garg R, Gerstner R, Pan A, Agarwala V, Alper HS. 2013. Generalizing a hybrid synthetic promoter approach in *Yarrowia lipolytica*. *Applied Microbiology and Biotechnology* **97**, 3037-3052.

Bock R. 2015. Engineering Plastid Genomes: Methods, Tools, and Applications in Basic Research and Biotechnology. *Annual Reviews of Plant Biology* **66,** 211–241.

Bond C, Tang Y, Li L. 2016. *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases. *Fungal Genetics and Biology* **89**, 52-61.

Brown S, Clastre M, Courdavault V, O'Connor SE. 2015. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 3205-3210.

Chen F, Li W, Jiang L, Pu X, Yang Y, Zhang G, Luo Y. 2016. Functional characterization of a geraniol synthase-encoding gene from *Camptotheca acuminata* and its application in production of geraniol in *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology* **43**, 1281-1292.

Cobb RE, Sun N, Zhao H. 2013. Directed evolution as a powerful synthetic biology tool. *Methods* **60**, 81-90.

Conrado RJ, Varner JD, DeLisa MP. 2008. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current Opinion in Biotechnology* **19**, 492-499.

Covello PS. 2008. Making artemisinin. *Phytochemistry* **69**, 2881-2885.

Cordero BF, Couso I, León R, Rodríguez H, Vargas MÁ. 2011. Enhancement of carotenoids biosynthesis in *Chlamydomonas reinhardtii* by nuclear transformation using a phytoene synthase gene isolated from *Chlorella zofingiensis*. *Applied Microbiology and Biotechnology* **91**, 341-351.

Couso I, Vila M, Rodriguez H, Vargas MA, León R. 2011. Overexpression of an exogenous phytoene synthase gene in the unicellular alga *Chlamydomonas reinhardtii* leads to an increase in the content of carotenoids. *Biotechnology Progress* **27**, 54-60.

Croft MT, Moulin M, Webb ME, Smith AG. 2007. Thiamine biosynthesis in algae is regulated by riboswitches. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 20770-20775.

Croteau R, Kutchan TM, Lewis NG. 2000. Natural Products (Secondary Metabolites). In: Buchanan B, Gruissem W, Jones R, eds. *Biochemistry & Molecular Biology of Plants*: American Society of Plant Biologists, 1250-1318.

Csorgo B, Feher T, Timar E, Blattner FR, Posfai G. 2012. Low-mutation-rate, reduced-genome *Escherichia coli*: an improved host for faithful maintenance of engineered genetic constructs. *Microbial Cell Factories* **11**, 11.

Curran KA, Crook NC, Karim AS, Gupta A, Wagman AM, Alper HS. 2014. Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nature Communication* **5**, 4002.

Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. 2015. Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. *ACS Synthetic Biology* **4**, 824-832.

Davies FK, Work VH, Beliaev AS, Posewitz MC. 2014. Engineering Limonene and Bisabolene Production in Wild Type and a Glycogen-Deficient Mutant of *Synechococcus* sp. PCC 7002. *Frontiers in Bioengineering and Biotechnology* **2**, 21.

Day A, Goldschmidt-Clermont M. 2011. The chloroplast transformation toolbox: selectable markers and marker removal. *Plant Biotechnology J* **9**, 540-553.

De Paepe B, Peters G, Coussement P, Maertens J, De Mey M. 2016. Tailor-made transcriptional biosensors for optimizing microbial cell factories. *Journal of Industrial Microbiology & Biotechnology*, 1-23.

DeLoache WC, Russ ZN. 2015. An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nature Chemical Biology* **11**, 465-471.

Deng XD, Gu B, Li YJ, Hu XW, Guo JC, Fei XW. 2012. The roles of acyl-CoA: diacylglycerol acyltransferase 2 genes in the biosynthesis of triacylglycerols by the green algae *Chlamydomonas reinhardtii*. *Molecular Plant* **5**, 945-947.

Dorrell RG, Smith AG. 2011. Do Red and Green Make Brown?: Perspectives on Plastid Acquisitions within Chromalveolates. *Eukaryotic cell* **10**, 856-868.

Ducat DC, Avelar-Rivas JA, Way JC, Silver PA. 2012. Rerouting Carbon Flux To Enhance Photosynthetic Productivity. *Applied Environmental Microbiology* **78**, 2660-2668.

Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KL, Keasling JD. 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology* **27**, 753-759.

Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647.

Engler C, Marillonnet S. 2014. Golden Gate cloning. *Methods in Molecular Biology* **1116**, 119-131.

Escalante AE, Rebolleda-Gomez M, Benitez M, Travisano M. 2015. Ecological perspectives on synthetic biology: insights from microbial population biology. *Frontiers in Microbiology* **6**, 143.

Farrow SC, Hagel JM, Beaudoin GA, Burns DC, Facchini PJ. 2015. Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. *Nature Chemical Biology* **11**, 728-732.

Feng J, Jester BW, Tinberg CE, Mandell DJ, Antunes MS, Chari R, Morey KJ, Rios X, Medford JI, Church GM, Fields S, Baker D. 2015. A general strategy to construct small molecule biosensors in eukaryotes. *Elife* **4**, e10606.

Fischer N, Rochaix JD. 2001. The flanking regions of PsaD drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Molecular Genetics and Genomics* **265**, 888-894.

Foo JL, Ching CB, Chang MW, Leong SS. 2012. The imminent role of protein engineering in synthetic biology. *Biotechnology Advances* **30**, 541-549.

Frasch HJ, Medema MH, Takano E, Breitling R. 2013. Design-based re-engineering of biosynthetic gene clusters: plug-and-play in practice. *Current Opinion in Biotechnology* **24**, 1144-1150.

Frohwitter J, Heider SA, Peters-Wendisch P, Beekwilder J, Wendisch VF. 2014. Production of the sesquiterpene (+)-valencene by metabolically engineered *Corynebacterium glutamicum*. *Journal of Biotechnology* **191**, 205-213.

Fujio T. 2007. Minimum genome factory: innovation in bioprocesses through genome science. *Biotechnology and Applied Biochemistry* **46**, 145-146.

Gangl D, Zedler JAZ, Włodarczyk A, Jensen PE, Purton S, Robinson C. 2015. Expression and membrane-targeting of an active plant cytochrome P450 in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Phytochemistry* **110**, 22-28.

Gao S, Han L, Zhu L, Ge M, Yang S, Jiang Y, Chen D. 2014. One-step integration of multiple genes into the oleaginous yeast *Yarrowia lipolytica*. *Biotechnology Letters* **36**, 2523-2528.

Gao S, Tong Y, Wen Z, Zhu L, Ge M, Chen D, Jiang Y, Yang S. 2016a. Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. *Journal of Industrial Microbiology and Biotechnology* **43**, 1085-1093.

Gao X, Gao F, Liu D, Zhang H, Nie X, Yang C. 2016b. Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energy & Environmental Science* **9**, 1400-1411.

Gao ZX, Zhao H, Li ZM, Tan XM, Lu XF. 2012. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy & Environmental Science* **5**, 9857–9865.

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CAr, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343-345.

Giga-Hama Y, Tohda H, Takegawa K, Kumagai H. 2007. *Schizosaccharomyces pombe* minimum genome factory. *Biotechnology and Applied Biochemistry* **46**, 147-155.

Gimpel JA, Henriquez V, Mayfield SP. 2015. In Metabolic Engineering of Eukaryotic Microalgae: Potential and Challenges Come with Great Diversity. *Frontiers in Microbiology* **6**, 1376.

Girvan HM, Munro AW. 2016. Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Current Opinion in Chemical Biology* **31**, 136-145.

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, 3rd, Smith HO, Venter JC. 2006. Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 425-430.

Gnanasekaran T, Vavitsas K, Andersen-Ranberg J, Nielsen AZ, Olsen CE, Hamberger B, Jensen PE. 2015. Heterologous expression of the isopimaric acid pathway in *Nicotiana benthamiana* and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *Journal of Biological Engineering* **9**, 2831.

Goldschmidt-Clermont M. 1991. Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker for site-directed transformation of Chlamydomonas. *Nucleic Acids Research* **19**, 4083-2089.

Gong Y, Guo X, Wan X, Liang Z, Jiang M. 2011. Characterization of a novel thioesterase (PtTE) from *Phaeodactylum tricornutum*. *Journal of Basic Microbiology* **51**, 666-672.

Goñi G, Zöllner A, Lisurek M, Velázquez-Campoy A, Pinto S, Gómez-Moreno C, Hannemann F, Bernhardt R, Medina M. 2009. Cyanobacterial electron carrier proteins as electron donors to CYP106A2 from *Bacillus megaterium* ATCC 13368. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1794**, 1635-1642.

Hamilton ML, Haslam RP, Napier JA, Sayanova O. 2014. Metabolic engineering of *Phaeodactylum tricornutum* for the enhanced accumulation of omega-3 long chain polyunsaturated fatty acids. *Metabolic Engineering* **22**, 3-9.

Hammer K, Mijakovic I, Jensen PR. 2006. Synthetic promoter libraries--tuning of gene expression. *Trends in Biotechnology* **24**, 53-55.

Han GH, Kim SK, Yoon PK, Kang Y, Kim BS, Fu Y, Sung BH, Jung HC, Lee DH, Kim SW, Lee SG. 2016. Fermentative production and direct extraction of (-)-alpha-bisabolol in metabolically engineered *Escherichia coli*. *Microbial Cell Factories* **15**, 185.

Hassani S, Momtaz S, Vakhshiteh F, Maghsoudi AS, Ganjali MR, Norouzi P, Abdollahi M. 2016. Biosensors and their applications in detection of organophosphorus pesticides in the environment **91**, 109-130.

Hauk P, Stephens K, McKay R, Virgile CR, Ueda H, Ostermeier M, Ryu KS, Sintim HO, Bentley WE. 2016. Insightful directed evolution of *Escherichia coli* quorum sensing promoter region of the lsrACDBFG operon: a tool for synthetic biology systems and protein expression. **44**, 10515-10525.

Heider SA, Peters-Wendisch P, Netzer R, Stafnes M, Brautaset T, Wendisch VF. 2014. Production and glucosylation of C50 and C 40 carotenoids by metabolically engineered *Corynebacterium glutamicum*. *Applied Microbiology and Biotechnology* **98**, 1223-1235.

Heider SA, Wendisch VF. 2015. Engineering microbial cell factories: Metabolic engineering of *Corynebacterium glutamicum* with a focus on non-natural products. *Biotechnology Journal* **10**, 1170-1184.

Helliwell KE, Scaife MA, Sasso S, Araujo APU, Purton S, Smith AG. 2014. Unraveling vitamin B12-responsive gene regulation in algae. *Plant Physiology* **165**, 388-397.

Hirooka K, Fujita Y. 2011. Identification of aromatic residues critical to the DNA binding and ligand response of the Bacillus subtilis QdoR (YxaF) repressor antagonized by flavonoids. *Bioscience, Biotechnology, and Biochemistry* **75**, 1325-1334.

Horn AH, Sticht H. 2015. Synthetic Protein Scaffolds Based on Peptide Motifs and Cognate Adaptor Domains for Improving Metabolic Productivity. *Frontiers in Bioengineering and Biotechnology* **3**, 191.

Hsieh H-J, Su C-H, Chien L-J. 2012. Accumulation of lipid production in *Chlorella minutissima* by triacylglycerol biosynthesis-related genes cloned from *Saccharomyces cerevisiae* and *Yarrowia lipolytica*. *Journal of Microbiology* **50**, 526-534.

Ishikura K, Takaoka Y, Kato K, Sekine M, Yoshida K, Shinmyo A. 1999. Expression of a foreign gene in *Chlamydomonas reinhardtii* chloroplast. *Journal of Bioscience and Bioengineering* **87**, 307-314.

Jacobsen JH, Frigaard NU. 2014. Engineering of photosynthetic mannitol biosynthesis from CO_2 in a cyanobacterium. *Metabolic Engineering* **21**, 60-70.

Jinkerson RE, Jonikas MC. 2015. Molecular techniques to interrogate and edit the *Chlamydomonas* nuclear genome. *The Plant Journal* **82**, 393-412.

Jones PR. 2014. Genetic Instability in Cyanobacteria - An Elephant in the Room? *Frontiers in Bioengineering and Biotechnology* 2, 7098.

Jongedijk E, Cankar K, Buchhaupt M, Schrader J, Bouwmeester H, Beekwilder J. 2016. Biotechnological production of limonene in microorganisms. *Applied Microbiology and Biotechnology* **100**, 2927-2938.

Jonsson J, Norberg T, Carlsson L, Gustafsson C, Wold S. 1993. Quantitative sequence-activity models (QSAM)--tools for sequence design. *Nucleic Acids Research* **21**, 733-739.

Jørgensen K, Rasmussen AV, Morant M, Nielsen AH, Bjarnholt N, Zagrobelny M, Bak S, Møller BL. 2005. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology* **8**, 280-291.

Juhas M, Reuss DR, Zhu B, Commichau FM. 2014. *Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering. *Microbiology* **160**, 2341-2351.

Julsing MK, Rijpkema M, Woerdenbag HJ, Quax WJ, Kayser O. 2007. Functional analysis of genes involved in the biosynthesis of isoprene in *Bacillus subtilis*. *Applied Microbiology and Biotechnology* **75**, 1377-1384.

Kang Z, Huang H, Zhang Y, Du G, Chen J. 2017. Recent advances of molecular toolbox construction expand *Pichia pastoris* in synthetic biology applications. *World Journal of Microbiology and Biotechnology* **33**, 19.

Karas BJ, Diner RE, Lefebvre SC, McQuaid J, Phillips APR, Noddings CM, Brunson JK, Valas RE, Deerinck TJ, Jablanovic J, Gillard JTF, Beeri K, Ellisman MH, Glass JI, Hutchison III CA, Smith HO, Venter JC, Allen AE, Dupont CL, Weyman PD. 2015. Designer diatom episomes delivered by bacterial conjugation. *Nature Communication* **6**, 6925.

Kato K, Marui T, Kasai S, Shinmyo A. 2007. Artificial Control of Transgene Expression in *Chlamydomonas reinhardtii* Chloroplast Using the lac Regulation System from *Escherichia coli*. *Journal of Bioscience and Bioengineering* **104**, 207-213.

Katsuyama Y, Funa N, Miyahisa I, Horinouchi S. 2007. Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in *Escherichia coli*. *Chemistry and Biology* **14**, 613-621.

Kavscek M, Strazar M, Curk T, Natter K, Petrovic U. 2015. Yeast as a cell factory: current state and perspectives. *Microbial Cell Factories* **14**, 94.

Kazamia E, Smith AG. 2014. Assessing the environmental sustainability of biofuels. *Trends in Plant Science* **19**, 615–618.

Kaznessis YN. 2009. Computational methods in synthetic biology. *Biotechnology Journal* **4**, 1392-1405.

Keasling JD. 2012. Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering* **14**, 189-195.

Kelchtermans P, Bittremieux W, De Grave K, Degroeve S, Ramon J, Laukens K, Valkenborg D, Barsnes H, Martens L. 2014. Machine learning applications in proteomics research: how the past can boost the future. *Proteomics* **14**, 353-366.

Khan FI, Wei DQ, Gu KR, Hassan MI, Tabrez S. Current updates on computer aided protein modeling and designing. *International Journal of Biological Macromolecules*. **85**, 48-62.

Khongto B, Laoteng K, Tongta A. 2010. Fermentation process development of recombinant *Hansenula polymorpha* for gamma-linolenic acid production. *Journal of Microbiology and Biotechnology* **20**, 1555-1562.

Kim BG, Yang SM, Kim SY, Cha MN, Ahn JH. 2015. Biosynthesis and production of glycosylated flavonoids in *Escherichia coli*: current state and perspectives. *Applied Microbiology and Biotechnology* **99**, 2979-2988.

Kim DH, Kim BG, Jung NR, Ahn JH. 2009. Production of genistein from naringenin using *Escherichia coli* containing isoflavone synthase-cytochrome P450 reductase fusion protein. *Journal of Microbiology and Biotechnology* **19**, 1612-1616.

Kim J, Salvador M, Saunders E, Gonzalez J, Avignone-Rossa C, Jimenez JI. 2016. Properties of alternative microbial hosts used in synthetic biology: towards the design of a modular chassis. *Essays in Biochemistry* **60**, 303-313.

Kingsley LJ, Lill MA. 2015. Substrate tunnels in enzymes: structure-function relationships and computational methodology. *Proteins* **83**, 599-611.

Kishimoto S, Sato M, Tsunematsu Y, Watanabe K. 2016. Evaluation of Biosynthetic Pathway and Engineered Biosynthesis of Alkaloids. *Molecules* **21**.

Klein-Marcuschamer D, Chisti Y, Benemann JR, Lewis D. 2013. A matter of detail: assessing the true potential of microalgal biofuels. *Biotechnology and Bioengineering* **110**, 2317-2322.

Knight T. 2003. Idempotent Vector Design for Standard Assembly of Biobricks. URL: http://hdl.handle.net/1721.1/21168

Knoop H, Gründel M, Zilliges Y, Lehmann R, Hoffmann S, Lockau W, Steuer R. 2013. Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of *Synechocystis* sp. PCC 6803. *PLoS Computational Biology* **9**, e1003081.

Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. 2010. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2646-2651.

Krasovska OS, Stasyk OG, Nahorny VO, Stasyk OV, Granovski N, Kordium VA, Vozianov OF, Sibirny AA. 2007. Glucose-induced production of recombinant proteins in *Hansenula polymorpha* mutants deficient in catabolite repression. *Biotechnology and Bioengineering* **97**, 858-870.

Lan EI, Chuang DS, Shen CR, Lee AM, Ro SY, Liao JC. 2015. Metabolic engineering of cyanobacteria for photosynthetic 3-hydroxypropionic acid production from CO_2 using *Synechococcus elongatus* PCC 7942. *Metabolic Engineering* **31**, 163-170.

Lassen LM, Nielsen AZ, Olsen CE, Bialek W, Jensen K, Møller BL, Jensen PE. 2014. Anchoring a Plant Cytochrome P450 via PsaM to the Thylakoids in *Synechococcus* sp. PCC 7002: Evidence for Light-Driven Biosynthesis. *PLoS One* **9**, e102184.

Lechner A, Brunk E, Keasling JD. 2016. The Need for Integrated Approaches in Metabolic Engineering. *Cold Spring Harbor Perspectives in Biology* **8**.

Ledesma-Amaro R, Lazar Z, Rakicka M, Guo Z, Fouchard F, Coq AC, Nicaud JM. 2016. Metabolic engineering of *Yarrowia lipolytica* to produce chemicals and fuels from xylose. *Metabolic Engineering* **38**, 115-124.

Ledesma-Amaro R, Nicaud JM. 2016. Metabolic Engineering for Expanding the Substrate Range of *Yarrowia lipolytica*. *Trends in Biotechnology* **34**, 798-809.

Lee ME, DeLoache WC, Cervantes B, Dueber JE. 2015. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synthetic Biology* **4**, 975-986.

Li C, Ying LQ, Zhang SS, Chen N, Liu WF, Tao Y. 2015. Modification of targets related to the Entner-Doudoroff/pentose phosphate pathway route for methyl-D-erythritol 4-phosphate-dependent carotenoid biosynthesis in *Escherichia coli*. *Microbial Cell Factories* **14**, 117.

Li H, Liao JC. 2013. Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO_2 to 1,2-propanediol. *Microbial Cell Factories* **12**, 4.

Li Y. 2012. Beyond Protein Engineering: its Applications in Synthetic Biology. *Enzyme Engineering* **1**.

Li Y, Cirino PC. 2014. Recent advances in engineering proteins for biocatalysis. *Biotechnology and Bioengineering* **111**, 1273-1287.

Li Y, Zhu X, Zhang X, Fu J, Wang Z, Chen T, Zhao X. 2016. Characterization of genome-reduced *Bacillus subtilis* strains and their application for the production of guanosine and thymidine. *Microbial Cell Factories* **15**, 94.

Liang DM, Liu JH, Wu H, Wang BB, Zhu HJ, Qiao JJ. 2015. Glycosyltransferases: mechanisms and applications in natural product development. *Chemical Society Reviews* **44**, 8350-8374.

Lieder S, Nikel PI, de Lorenzo V, Takors R. 2015. Genome reduction boosts heterologous gene expression in *Pseudomonas putida*. *Microbial Cell Factories* **14**, 23.

Lindberg P, Park S, Melis A. 2010. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic Engineering* **12**, 70-79.

Liu HH, Ji XJ, Huang H. 2015. Biotechnological applications of *Yarrowia lipolytica*: Past, present and future. *Biotechnology Advances* **33**, 1522-1546.

Liu J, Zhong Y, Sun Z, Huang J, Sandmann G, Chen F. 2010. One amino acid substitution in phytoene desaturase makes *Chlorella zofingiensis* resistant to norflurazon and enhances the biosynthesis of astaxanthin. *Planta* **232**, 61-67.

Loeschcke A, Thies S. 2015. *Pseudomonas putida*-a versatile host for the production of natural products. *Applied Microbiology and Biotechnology* **99**, 6197-6214.

Ma T, Deng Z, Liu T. 2016. Microbial production strategies and applications of lycopene and other terpenoids. *World Journal of Microbiology and Biotechnology* **32**, 15.

Mahr R, Frunzke J. 2016. Transcription factor-based biosensors in biotechnology: current state and future prospects. *Applied Microbiology and Biotechnology* **100**, 79-90.

Maia P, Rocha M, Rocha I. 2015. *In Silico* Constraint-Based Strain Optimization Methods: the Quest for Optimal Cell Factories. *Microbiology and Molecular Biology Reviews* **80**, 45-67.

Manuell AL, Beligni MV, Elder JH, Siefker DT, Tran M, Weber A, McDonald TL, Mayfield SP. 2007. Robust expression of a bioactive mammalian protein in *Chlamydomonas* chloroplast. *Plant Biotechnology Journal* **5**, 402-412.

Marin AM, Souza EM, Pedrosa FO, Souza LM, Sassaki GL, Baura VA, Yates MG, Wassem R, Monteiro RA. 2013. Naringenin degradation by the endophytic diazotroph *Herbaspirillum seropedicae* SmR1. *Microbiology* **159**, 167-175.

Markham KA, Alper HS. 2015. Synthetic Biology for Specialty Chemicals. *Annual Review of Chemical and Biomolecular Engineering* **6**, 35-52.

Marner WD, 2nd. 2009. Practical application of synthetic biology principles. *Biotechnology Journal* **4**, 1406-1419.

Marshall KE, Robinson EW, Hengel SM, Paša-Tolić L, Roesijadi G. 2012. FRET Imaging of Diatoms Expressing a Biosilica-Localized Ribose Sensor. *PLoS One* **7**, e33771.

Matsumura E, Nakagawa A, Tomabechi Y, Koyanagi T, Kumagai H, Yamamoto K, Katayama T, Sato F, Minami H. 2017. Laboratory-scale production of (S)-reticuline, an

important intermediate of benzylisoquinoline alkaloids, using a bacterial-based method. *Bioscience, Biotechnology, and Biochemistry* **81**, 396-402.

Matthijs M, Fabris M, Broos S, Vyverman W, Goossens A. 2016. Profiling of the Early Nitrogen Stress Response in the Diatom *Phaeodactylum tricornutum* Reveals a Novel Family of RING-Domain Transcription Factors. *Plant Physiology* **170**, 489-498.

McKeague M, Wong RS, Smolke CD. 2016. Opportunities in the design and application of RNA for gene expression control. *Nucleic Acids Research* **44**, 2987-2999.

Mehdizadeh Aghdam E, Hejazi MS, Barzegar A. 2016. Riboswitches: From living biosensors to novel targets of antibiotics. *Gene* **592**, 244-259.

Mehrotra P. 2016. Biosensors and their applications - A review. *Journal of Oral Biology and Craniofacial Research* **6**, 153-159.

Meng H, Wang J, Xiong Z, Xu F, Zhao G, Wang Y. 2013. Quantitative design of regulatory elements based on high-precision strength prediction using artificial neural network. *PLoS One* **8**, e60288.

Mizoguchi H, Mori H, Fujio T. 2007. *Escherichia coli* minimum genome factory. *Biotechnology and Applied Biochemistry* **46**, 157-167.

Moon TS, Dueber JE, Shiue E, Prather KL. 2010. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metabolic Engineering* **12**, 298-305.

Moulin M, Nguyen GTDT, Scaife MA, Smith AG, Fitzpatrick TB. 2013. Analysis of *Chlamydomonas* thiamin metabolism in vivo reveals riboswitch plasticity. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 14622-14627.

Mustacchi R, Hohmann S, Nielsen J. 2006. Yeast systems biology to unravel the network of life. *Yeast* **23**, 227-238.

Nakahira Y, Ogawa A, Asano H, Oyama T, Tozawa Y. 2013. Theophylline-dependent riboswitch as a novel genetic tool for strict regulation of protein expression in Cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant and Cell Physiology* **54**, 1724-1735.

Neupert J, Karcher D, Bock R. 2009. Generation of Chlamydomonas strains that efficiently express nuclear transgenes. *The Plant Journal* **57**, 1140-1150.

Nielsen AZ, Mellor SB, Vavitsas K, Wlodarczyk AJ, Gnanasekaran T, Perestrello Ramos H de Jesus M, King BC, Bakowski K, Jensen PE. 2016. Extending the biosynthetic repertoires of cyanobacteria and chloroplasts. *The Plant Journal* **87**, 87-102.

Nielsen AZ, Ziersen B, Jensen K, Lassen LM, Olsen CE, Møller BL, Jensen PE. 2013. Redirecting Photosynthetic Reducing Power toward Bioactive Natural Product Synthesis. *ACS Synthetic Biology* **2**, 308-315.

Nogales J, Gudmundsson S, Knight EM, Palsson BO, Thiele I. 2012. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2678-2683.

Ohresser M, Matagne RF, Loppes R. 1997. Expression of the arylsulphatase reporter gene under the control of the nit1 promoter in *Chlamydomonas reinhardtii*. *Current Genetics* **31**, 264-271.

Oliver JWK, Machado IMP, Yoneda H, Atsumi S. 2013. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1249-1254.

Paddon CJ, Keasling JD. 2014. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology* **12**, 355-367.

Paige JS, Nguyen-Duc T, Song W, Jaffrey SR. 2012. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194.

Palchetti I. 2016. New Trends in the Design of Enzyme-based Biosensors for Medical Applications. *Mini Reviews in Medicinal Chemistry* **16**, 1125-1133.

Patron NJ. 2014. DNA assembly for plant biology: techniques and tools. *Current Opinion in Plant Biology* **19**, 14-19.

Peng K-T, Zheng C-N, Xue j, Chen X-Y, Yang W-D, Liu J-S, Bai W, Li H-Y. 2014. Delta 5 Fatty Acid Desaturase Upregulates the Synthesis of Polyunsaturated Fatty Acids in the Marine Diatom *Phaeodactylum tricornutum*. *Journal of Agricultural and Food Chemistry* **62**, 8773-8776.

Perez-Garcia O, Lear G, Singhal N. 2016. Metabolic Network Modeling of Microbial Interactions in Natural and Engineered Environmental Systems. *Frontier in Microbiology* **7**, 673.

Pittman JK, Dean AP, Osundeko O. 2011. The potential of sustainable algal biofuel production using wastewater resources. *Bioresource Technology* **102,** 17–25.

Proschel M, Detsch R, Boccaccini AR, Sonnewald U. 2015. Engineering of Metabolic Pathways by Artificial Enzyme Channels. *Frontiers in Bioengineering and Biotechnology* **3**, 168.

Purton S, Szaub JB, Wannathong T, Young R, Economou CK. 2013. Genetic engineering of algal chloroplasts: Progress and prospects. *Russian Journal of Plant Physiology* **60,** 491–499.

Quinn JM, Merchant S. 1995. Two copper-responsive elements associated with the *Chlamydomonas* Cyc6 gene function as targets for transcriptional activators. *The Plant Cell Online* **7**, 623-628.

Rajkumar AS, Liu G, Bergenholm D, Arsovska D, Kristensen M, Nielsen J, Jensen MK, Keasling JD. 2016. Engineering of synthetic, stress-responsive yeast promoters. *Nucleic Acids Research* **44**, e136.

Raman S, Rogers JK, Taylor ND, Church GM. 2014. Evolution-guided optimization of biosynthetic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 17803-17808.

Ramundo S, Rahire M, Schaad O, Rochaix J-D. 2013. Repression of essential chloroplast genes reveals new signaling pathways and regulatory feedback loops in chlamydomonas. *The Plant Cell* **25**, 167-186.

Rantasalo A, Czeizler E, Virtanen R, Rousu J, Lahdesmaki H, Penttila M, Jantti J, Mojzita D. 2016. Synthetic Transcription Amplifier System for Orthogonal Control of Gene Expression in *Saccharomyces cerevisiae*. *PLoS One* **11**, e0148320.

Rao JR, Cooper JE. 1994. Rhizobia catabolize nod gene-inducing flavonoids via C-ring fission mechanisms. *Journal of Bacteriology* **176**, 5409-5413.

Redden H, Alper HS. 2015. The development and characterization of synthetic minimal yeast promoters. *Nature Communication* **6**, 7810.

Reed B, Blazeck J, Alper H. 2012. Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes. *Journal of Biotechnology* **158**, 75-79.

Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CL, Chandrasegaran S, Cai Y, Boeke JD, Bader JS. 2017. Design of a synthetic yeast genome. *Science* **355**, 1040-1044.

Rodicio R, Heinisch JJ. 2013. Yeast on the milky way: genetics, physiology and biotechnology of *Kluyveromyces lactis*. *Yeast* **30**, 165-177.

Rogers JK, Church GM. 2016. Genetically encoded sensors enable real-time observation of metabolite production. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 2388-2393.

Rogers JK, Guzman CD, Taylor ND, Raman S, Anderson K, Church GM. 2015. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Research* **43**, 7648-7660.

Rogers JK, Taylor ND, Church GM. 2016. Biosensor-based engineering of biosynthetic pathways. *Current Opinion in Biotechnology* **42**, 84-91.

Roth A, Breaker RR. 2009. The Structural and Functional Diversity of Metabolite-Binding Riboswitches. *Annual Review of Biochemistry* **78**, 305-334.

Saraya R, Krikken AM, Kiel JA, Baerends RJ, Veenhuis M, van der Klei IJ. 2012. Novel genetic tools for *Hansenula polymorpha*. *FEMS Yeast Research* **12**, 271-278.

Scaife MA, Smith AG. 2016. Towards developing algal synthetic biology. *Biochemical Society Transactions* **44**, 716-722.

Schroda M, Blöcker D, Beck CF. 2000. The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. *The Plant Journal* **21**, 121-131.

Schwartz C, Shabbir-Hussain M, Frogue K, Blenner M, Wheeldon I. 2016. Standardized Markerless Gene Integration for Pathway Engineering in *Yarrowia lipolytica*. *ACS Synthetic Biology* doi: 10.1021/acssynbio.6b00285.

Scott SA, Davey MP, Dennis JS, Horst I. 2010. Biodiesel from algae: challenges and prospects. *Current Opinion in Biotechnology* **21**, 277-286.

Shabbir Hussain M, Gambill L, Smith S, Blenner MA. 2016. Engineering Promoter Architecture in Oleaginous Yeast *Yarrowia lipolytica*. *ACS Synthetic Biology* **5**, 213-223.

Shen CR, Liao JC. 2012. Photosynthetic production of 2-methyl-1-butanol from CO_2 in cyanobacterium *Synechococcus elongatus* PCC7942 and characterization of the native acetohydroxyacid synthase. *Energy & Environmental Science* **5**, 9574-9583.

Shen HJ, Hu JJ, Li XR, Liu JZ. 2015. Engineering of *Escherichia coli* for Lycopene Production Through Promoter Engineering. *Current Pharmaceutical Biotechnology* **16**, 1094-1103.

Siedler S, Stahlhut SG, Malla S, Maury J, Neves AR. 2014. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. *Metabolic Engineering* **21**, 2-8.

Singleton C, Howard TP, Smirnoff N. 2014. Synthetic metabolons for metabolic engineering. *Journal of Experimental Botany* **65**, 1947-1954.

Sinha S. 2006. On counting position weight matrix matches in a sequence, with application to discriminative motif finding. *Bioinformatics* **22**, e454-463.

Solem C, Jensen PR. 2002. Modulation of gene expression made easy. *Applied Environmental Microbiology* **68**, 2397-2403.

Song Y, Nikoloff JM, Fu G, Chen J, Li Q, Xie N, Zheng P, Sun J, Zhang D. 2016. Promoter Screening from *Bacillus subtilis* in Various Conditions Hunting for Synthetic Biology and Industrial Applications. *PLoS One* **11**, e0158447.

Steinbrenner J, Sandmann G. 2006. Transformation of the Green Alga *Haematococcus pluvialis* with a Phytoene Desaturase for Accelerated Astaxanthin Biosynthesis. *Applied Environmental Microbiology* **72**, 7477-7484.

Stephanopoulos G. 2012. Synthetic biology and metabolic engineering. *ACS Synthetic Biology* **1**, 514-525.

Suastegui M, Shao Z. 2016. Yeast factories for the production of aromatic compounds: from building blocks to plant secondary metabolites. *Journal of Industrial Microbiology and Biotechnology* **43**, 1611-1624.

Suwannarangsee S, Oh DB, Seo JW, Kim CH, Rhee SK, Kang HA, Chulalaksananukul W, Kwon O. 2010. Characterization of alcohol dehydrogenase 1 of the thermotolerant methylotrophic yeast *Hansenula polymorpha*. *Applied Microbiology and Biotechnology* **88**, 497-507.

Tashiro M, Kiyota H, Kawai-Noma S, Saito K, Ikeuchi M, Iijima Y, Umeno D. 2016. Bacterial Production of Pinene by a Laboratory-Evolved Pinene-Synthase. *ACS Synthetic Biology* **5**, 1011-1020.

Trentacoste EM, Shrestha RP, Smith SR, Gle C, Hartmann AC, Hildebrand M, Gerwick WH. 2013. Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19748-19753.

Trosset JY, Carbonell P. 2015. Synthetic biology for pharmaceutical drug discovery. *Drug Design, Development and Therapy* **9**, 6285-6302.

Tyo KE, Nevoigt E, Stephanopoulos G. 2011. Directed evolution of promoters and tandem gene arrays for customizing RNA synthesis rates and regulation. *Methods in Enzymology* **497**, 135-155.

Umenhoffer K, Feher T, Baliko G, Ayaydin F, Posfai J, Blattner FR, Posfai G. 2010. Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microbial Cell Factories* **9**, 38.

Varman AM, Xiao Y, Pakrasi HB, Tang YJ. 2013. Metabolic Engineering of *Synechocystis* sp. Strain PCC 6803 for Isobutanol Production. *Applied Environmental Microbiology* **79**, 908-914.

Vogl T, Hartner FS, Glieder A. 2013. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Current Opinion in Biotechnology* **24**, 1094-1101.

Vogl T, Ruth C, Pitzer J, Kickenweiz T, Glieder A. 2014. Synthetic core promoters for *Pichia pastoris*. *ACS Synthetic Biology* **3**, 188-191.

Vu TT, Stolyar SM, Pinchuk GE, Hill EA, Kucek LA, Brown RN, Lipton MS, Osterman A, Fredrickson JK, Konopka AE, Beliaev AS, Reed JL. 2012. Genome-Scale Modeling of Light-Driven Reductant Partitioning and Carbon Fluxes in Diazotrophic Unicellular Cyanobacterium *Cyanothece* sp. ATCC 51142. *PLoS Computational Biology* **8**, e1002460.

Wagner JM, Alper HS. 2016. Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. *Fungal Genetics and Biology* **89**, 126-136.

Wagner WP, Helmig D, Fall R. 2000. Isoprene biosynthesis in *Bacillus subtilis* via the methylerythritol phosphate pathway. *Journal of Natural Products* **63**, 37-40.

Wang B, Pugh S, Nielsen DR, Zhang W, Meldrum DR. 2013a. Engineering cyanobacteria for photosynthetic production of 3-hydroxybutyrate directly from CO_2. *Metabolic Engineering* **16**, 68-77.

Wang W, Liu X, Lu X. 2013b. Engineering cyanobacteria to improve photosynthetic production of alka(e)nes. *Biotechnology for Biofuels* **6**, 69.

Wang J, Guleria S, Koffas MA, Yan Y. 2016. Microbial production of value-added nutraceuticals. *Current Opinion in Biotechnology* **37**, 97-104.

Wang X. 2009. Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS Letters* **583**, 3303-3309.

Wang Y, Yu O. 2012. Synthetic scaffolds increased resveratrol biosynthesis in engineered yeast cells. *Journal of Biotechnology* **157**, 258-260.

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. 2011. Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* **6**, e19722.

Wei L, Zou Q. 2016. Recent Progress in Machine Learning-Based Methods for Protein Fold Recognition. *International Journal of Molecular Sciences* **17**, E2118.

Werck-Reichhart D, Feyereisen R. 2000. Cytochromes P450: a success story. *Genome Biology* **1**, Reviews3003.

Westermark S, Steuer R. 2016. Toward Multiscale Models of Cyanobacterial Growth: A Modular Approach. *Frontiers in Bioengineering and Biotechnology* **4**, 95.

Winzer T, Kern M, King AJ, Larson TR, Teodor RI, Donninger SL, Li Y, Dowle AA, Cartwright J, Bates R, Ashford D, Thomas J, Walker C, Bowser TA, Graham IA. 2015. Plant science. Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. *Science* **349**, 309-312.

Wlodarczyk A, Gnanasekaran T, Nielsen AZ. 2016. Metabolic engineering of light-driven cytochrome P450 dependent pathways into *Synechocystis* sp. PCC 6803. *Metabolic Engineering* **33**, 1-11.

Xiong W, Morgan JA, Ungerer J, Wang B, Maness PC. 2015. The plasticity of cyanobacterial metabolism supports direct CO_2 conversion to ethylene. *Nature Plants* **1**, 15053.

Xu P, Wang W, Li L, Bhan N, Zhang F, Koffas MA. 2014. Design and kinetic analysis of a hybrid promoter-regulator system for malonyl-CoA sensing in *Escherichia coli*. *ACS Chemical Biology* **9**, 451-458.

Xue J, Niu Y-F, Huang T, Yang W-D, Liu J-S, Li H-Y. 2015. Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. *Metabolic Engineering* **27**, 1-9.

Yang J, Nie Q. 2016. Engineering *Escherichia coli* to convert acetic acid to beta-caryophyllene. *Microbial Cell Factories* **15**, 74.

Yang J, Xu M, Zhang X, Hu Q, Sommerfeld M, Chen Y. 2011. Life-cycle analysis on biodiesel production from microalgae: Water footprint and nutrients balance. *Bioresource Technology* **102**, 159–165.

Yu AQ, Pratomo N, Ng TK, Ling H, Cho HS, Leong SS, Chang MW. 2016. Genetic Engineering of an Unconventional Yeast for Renewable Biofuel and Biochemical Production. *Journal of Visualized Experiments*.

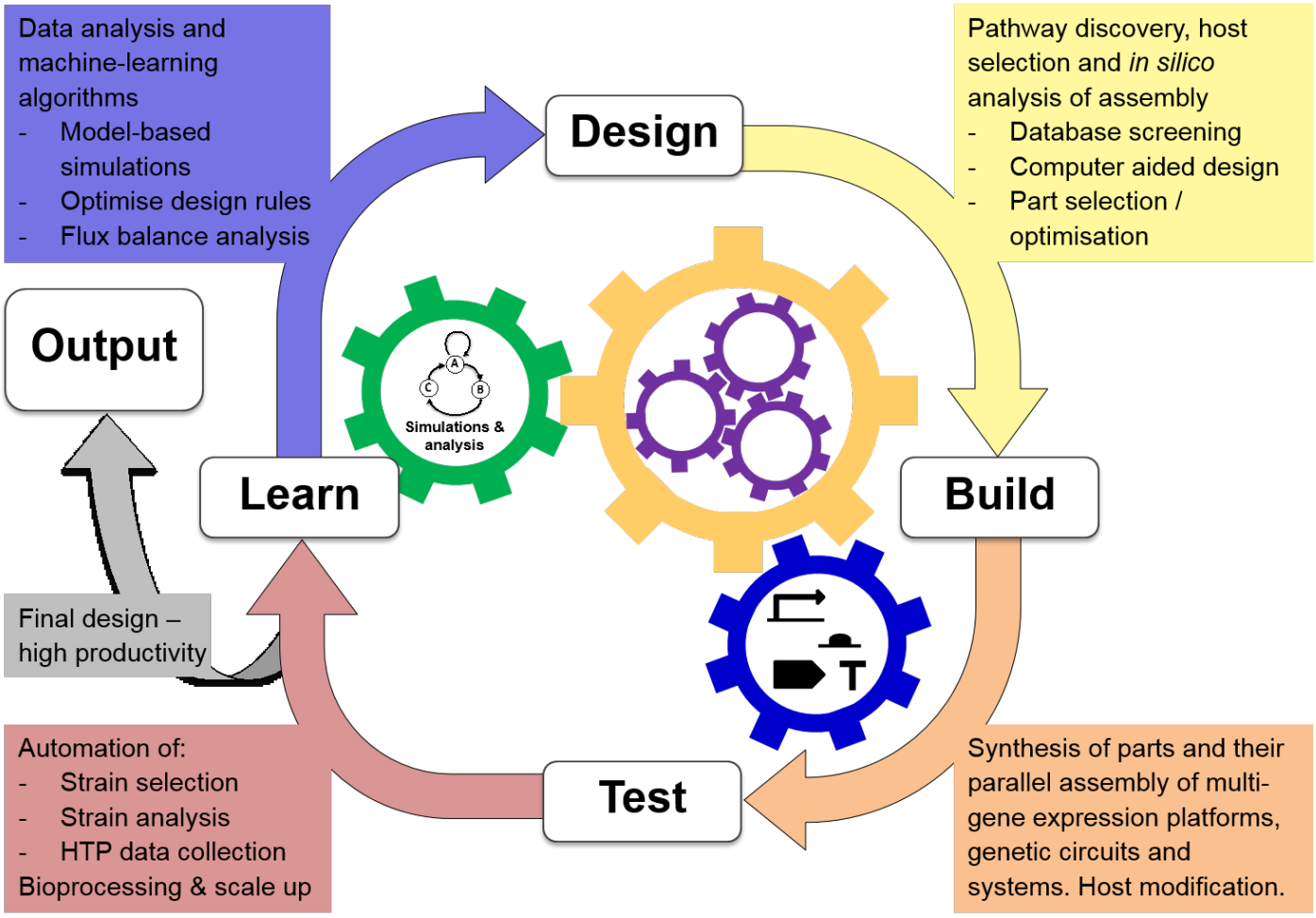
Zhang F, Carothers JM, Keasling JD. 2012. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology* **30**, 354-359.

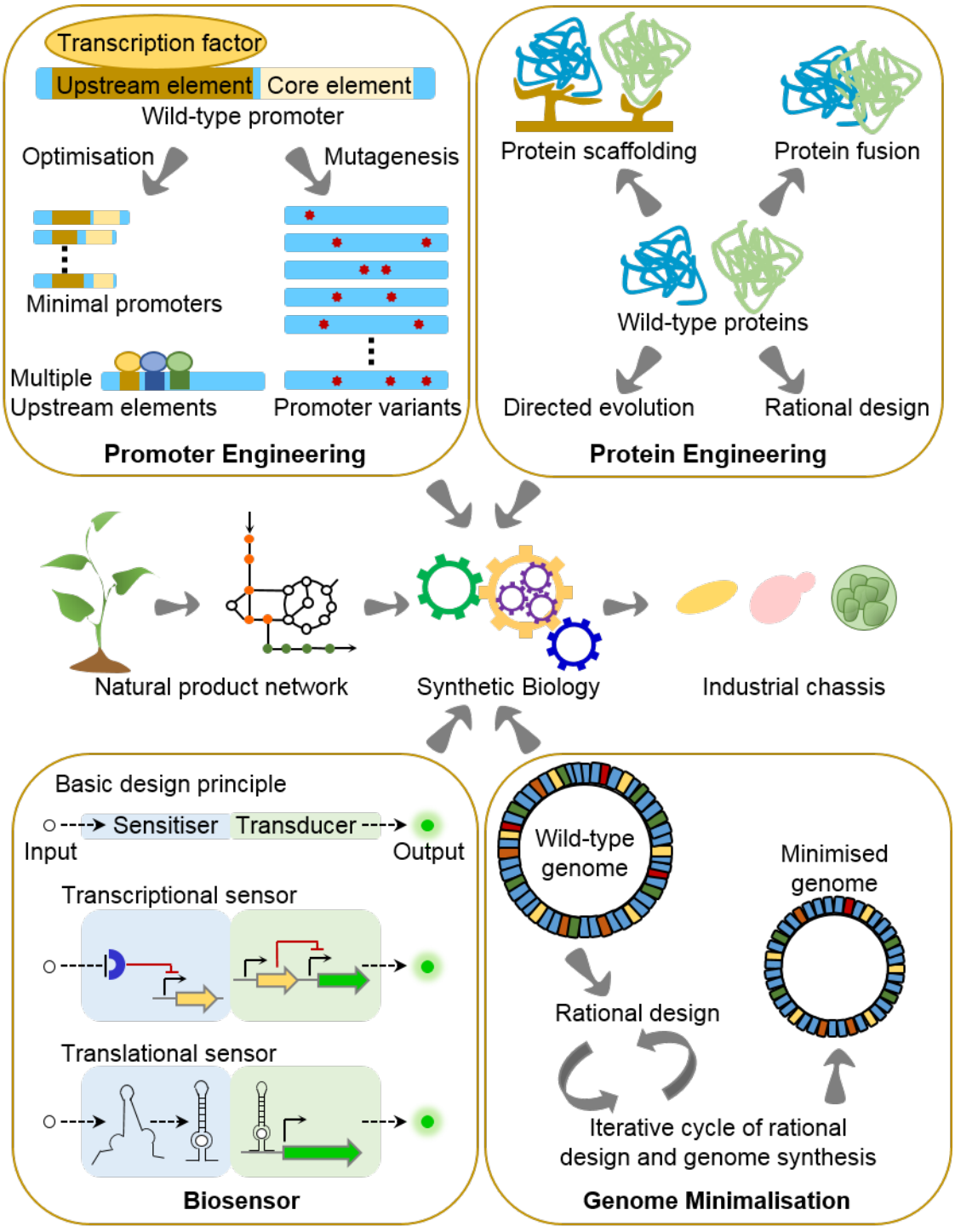
Zhang F, Keasling J. 2011. Biosensors and their applications in microbial metabolic engineering. *Trends in Microbiology* **19**, 323-329.

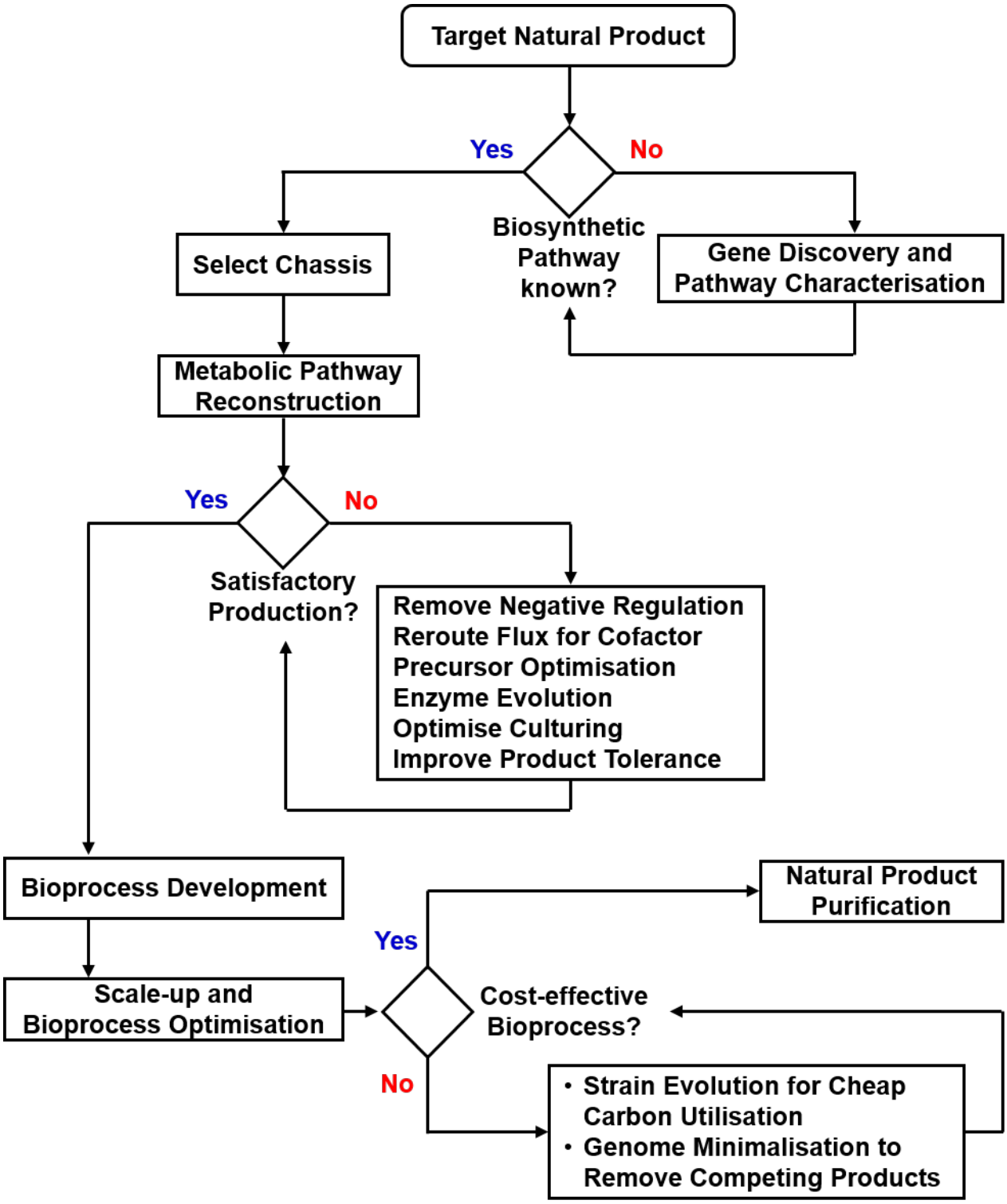
Zhao F, Bai P, Liu T, Li D, Zhang X, Lu W, Yuan Y. 2016. Optimization of a cytochrome P450 oxidation system for enhancing protopanaxadiol production in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering* **113**, 1787-1795.

Zhang J, Hao Q, Bai L, Xu J, Yin W, Song L, Xu L, Guo X, Fan C, Chen Y, Ruan J, Hao S, Li Y, Wang RR-C, Hu Z. 2014. Overexpression of the soybean transcription factor GmDof4 significantly enhances the lipid content of *Chlorella ellipsoidea*. *Biotechnology for Biofuels* **7**, 128.

Zhou K, Zou R, Zhang C, Stephanopoulos G, Too HP. 2013. Optimization of amorphadiene synthesis in *Bacillus subtilis* via transcriptional, translational, and media modulation. *Biotechnology and Bioengineering* **110**, 2556-2561.


Data analysis and machine-learning algorithms
- Model-based simulations
- Optimise design rules
- Flux balance analysis
Design
Pathway discovery, host selection and *in silico* analysis of assembly
- Database screening
- Computer aided design
- Part selection / optimisation
Output
Simulations & analysis
Learn
Build
Final design – high productivity
Automation of:
- Strain selection
- Strain analysis
- HTP data collection
Bioprocessing & scale up
Test
Synthesis of parts and their parallel assembly of multi-gene expression platforms, genetic circuits and systems. Host modification.


Transcription factor
Upstream element
Core element
Wild-type promoter
Optimisation
Mutagenesis
Minimal promoters
Multiple
Upstream elements
Promoter variants
Promoter Engineering
Protein scaffolding
Protein fusion
Wild-type proteins
Directed evolution
Rational design
Protein Engineering
Natural product network
Synthetic Biology
Industrial chassis
Basic design principle
Input
Sensitiser
Transducer
Output
Transcriptional sensor
Translational sensor
Biosensor
Wild-type
genome
Minimised
genome
Rational design
Iterative cycle of rational
design and genome synthesis
Genome Minimalisation


Target Natural Product
Yes
No
Biosynthetic Pathway known?
Select Chassis
Gene Discovery and Pathway Characterisation
Metabolic Pathway Reconstruction
Yes
No
Satisfactory Production?
Remove Negative Regulation
Reroute Flux for Cofactor
Precursor Optimisation
Enzyme Evolution
Optimise Culturing
Improve Product Tolerance
Bioprocess Development
Natural Product Purification
Yes
Scale-up and Bioprocess Optimisation
Cost-effective Bioprocess?
No
• Strain Evolution for Cheap Carbon Utilisation
• Genome Minimalisation to Remove Competing Products

Chassis Selection Guide

	Escherichia coli	*Saccharomyces cerevisiae*	Cyanobacteria	Microalgae
Gene Discovery and Functional Characterisation	+++	++		+
Validating Parts and Gene Permutations	+++	+++	+	+
Advanced Molecular Toolkit	+++	+++	++	+
Cheap Substrate for Growth	+	+	+++	++
Large Scale Cultivation*	++	++	+++	+++
Intracellular Compartmentalisation		++	+++	+++
Capacity to Excrete Product	++	+		
Minimal Genome Available	+	++		

* ++ indoor in enclosed bioreactors of several thousand litres, +++ outdoor cultivation at the hectare scale

Table 1: Examples of metabolite production systems in cyanobacteria and microalgae

Host	Compound	Reference	Heterologous (H) / Endogenous (E) pathway	Considerations for engineering of chassis organism
Cyanobacteria				
Synechococcus sp. PCC 7942	2,3-Butanediol	Oliver *et al.* 2013	H	Doubling time 12–24 hours. Autotrophic metabolism. Genetic manipulation via homologous recombination (HR) allows targeted engineering. Genome-scale metabolic model available (Berla *et al.*, 2013).
	Isobutaraldehyde	Atsumi *et al.* 2009	H	
	Sucrose	Ducat *et al.* 2012	H	
	Isoprene	Gao *et al.* 2016	H	
	Isobutanol	Atsumi *et al.* 2009	H	
	2-Methyl-1-butanol	Shen and Liao 2012	H	
	Lactate	Angermayr *et al.* 2014	H	
	1,2-Propanediol	Li and Liao 2013	H	
Synechococcus sp. PCC 7002	Limonene	Davies *et al.* 2014	H	Doubling time 3.5 hours. Mixotrophic and autotrophic metabolism. Genetic manipulation via HR possible. Genome-scale metabolic model available (Berla *et al.*, 2013).
	Bisabolene	Davies *et al.* 2014	H	
Synechocystis sp. PCC 6803	Ethanol	Gao *et al.* 2011	H	Doubling time 6–12 hours. Mixotrophic and autotrophic metabolism. Genetic manipulation via HR possible. Genome-scale metabolic model available (Berla *et al.*, 2013).
	3-Hydroxypropionate	Lan *et al.* 2015	H	
	Ethylene	Xiong *et al.* 2015	H	
	Isobutanol	Varman *et al.* 2013	H	
	3-Hydroxybutyrate	Wang *et al.* 2013a	H	
	Heptadecane	Wang *et al.* 2013b	H	
	Mannitol	Jacobsen and Frigaard 2014	H	

Microalgae				
Chlamydomonas reinhardtii	Hydrogen*	Li *et al.* 2010	E	Doubling time 5-8 hours. Both autotrophic and mixotrophic growth possible. Genomes of nucleus, chloroplast and mitochondrion sequenced and highly annotated. Transformation of all genomes routine. Comprehensive genetic, molecular and genomic resources available. Genome editting only possible through direct delivery of Cas9 ribonucleoprotein (RNP). Some limited knowledge of scale-up in bioreactors. Genome-scale metabolic model available.
	Lipids	Deng *et al.* 2012	E	
	Carotenoids	Couso *et al*. 2011	H	
		Cordero *et al*. 2011	H	
Phaeodactylum tricornutum	Lipids	Xue *et al*. 2015	E	Doubling time 18-29 hours. Genome sequenced and annotated. Nuclear genome can be transformed, also introduction of episomes by conjugation. Molecular tools expanding rapidly, and amiRNA and CRISPR/Cas9 genome editing established.
		Peng *et al*. 2014	E	
	Enhanced FA profile	Gong *et al*. 2011	E	
		Hamilton *et al*. 2014	H	
Thalassiosira pseudonana	Lipids	Trentacoste *et al*. 2013	E	Doubling time ~24 hours. Genome sequenced. Nuclear genome can be transformed routinely. Molecular tools expanding rapidly, including amiRNA. CRISPR/Cas9 genome editing reported.
	Ribose sensor	Marshall *et al.* 2012	H	
Chlorella zofigiensis	Carotenoids	Liu *et al*. 2010	E	Can grow rapidly in phototrophic, heterotrophic and mixotrophic conditions and naturally accumulates carotenoids, including astaxanthin, thus ideal for large-scale cultivation. Transformation reported.
Chlorella minutissima	Lipids	Hsieh *et al.* 2012	E	Transformation by HR reported.
Chlorella ellipsoidea	Lipids	Zhang *et al*. 2014	H	Fast growth under autotrophic or heterotrophic conditions. Transformation reported. Use of plant promoter/terminators demonstrated.
Haematococcus pluvialis	Carotenoids	Steinbrenner and Sandmann, 2006	E	Doubling time 13-72 hours. Naturally accumulates carotenoids, including astaxanthin, and being employed in large-scale cultivation. Molecular, genetic and genomic resources are very limited.

Asterisk indicates trait engineering via the chloroplast genome.