Combinatorial assembly platform enabling engineering of genetically stable metabolic pathways in cyanobacteria

**George M. Taylor[1], Andrew Hitchcock [2] and John T. Heap [1,3,*]**

[1]Imperial College Centre for Synthetic Biology, Department of Life Sciences, Imperial College London, London SW7 2AZ, UK, [2]Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK and [3]School of Life Sciences, The University of Nottingham, Biodiscovery Institute, University Park, Nottingham NG7 2RD, UK

ABSTRACT

Cyanobacteria are simple, efficient, genetically-tractable photosynthetic microorganisms which in principle represent ideal biocatalysts for CO_2 capture and conversion. However, in practice, genetic instability and low productivity are key, linked problems in engineered cyanobacteria. We took a massively parallel approach, generating and characterising libraries of synthetic promoters and RBSs for the cyanobacterium *Synechocystis* sp. PCC 6803, and assembling a sparse combinatorial library of millions of metabolic pathway-encoding construct variants. Genetic instability was observed for some variants, which is expected when variants cause metabolic burden. Surprisingly however, in a single combinatorial round without iterative optimisation, 80% of variants chosen at random and cultured photoautotrophically over many generations accumulated the target terpenoid lycopene from atmospheric CO_2, apparently overcoming genetic instability. This large-scale parallel metabolic engineering of cyanobacteria provides a new platform for development of genetically stable cyanobacterial biocatalysts for sustainable light-driven production of valuable products directly from CO_2, avoiding fossil carbon or competition with food production.

GRAPHICAL ABSTRACT
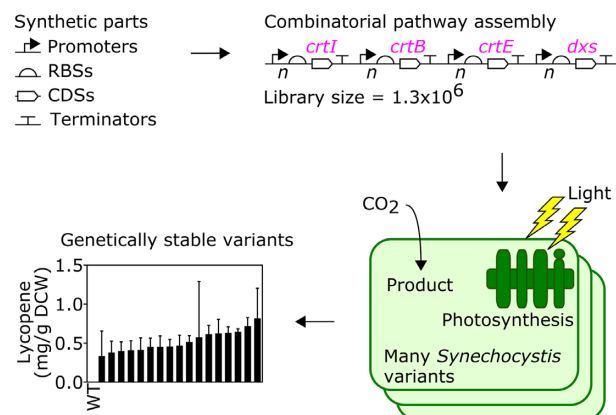
INTRODUCTION

Future sustainable economies consistent with net zero greenhouse gas emissions and limiting climate change will require radical changes in global carbon flows, including the development and large-scale deployment of technologies which capture and recycle carbon. Cyanobacteria are the simplest and most genetically-tractable organisms capable of oxygenic photosynthesis, using CO_2 and sunlight as their sole carbon and energy sources, respectively. Their photosynthetic yield ([1]) and growth rate ([2]) are similar to fast-growing microalgae, and greater than terrestrial plants ([3,4]). Therefore, cyanobacteria have great potential to serve as whole-cell biocatalysts in light-driven, carbon-negative bioprocesses converting atmospheric or waste CO_2 to products of interest, while avoiding competition with the food chain.

Cyanobacteria have been genetically modified to synthesise a wide range of non-native compounds ranging from commodity chemicals and biofuels to high-value products, including alcohols ([5–8]), organic acids ([8–11]), diols

*To whom correspondence should be addressed. Tel: +44 115 84 67956; Email: john.heap@nottingham.ac.uk

(12,13), terpenoids (14–21), sugars (22–26) and many others (27–32). However, production using cyanobacteria is not yet commercially competitive (33–35) with production by petrochemical or conventional heterotrophic biotechnology approaches, both of which benefit from the current low economic costs of using fossil carbon and emitting CO_2. There are also issues with cyanobacterial production itself, including lower growth rates and cell densities than many heterotrophs, self-shading (36), and the crucial genetic and metabolic design challenges of low productivity (37) and genetic instability (38). Importantly, these are linked: genetic instability and loss of metabolic pathways causes loss of productivity (8).

Heterologous DNA often confers a fitness cost on host cells, through the activity of encoded heterologous proteins, or the demands of their synthesis, or both (39). In the case of heterologous metabolic pathways, the diversion of central metabolites away from endogenous biosynthesis towards non-native products can limit growth. The relative expression levels of enzymes in heterologous pathways may also be poorly balanced, leading to accumulation of intermediates, which can be toxic (40,41). Heterologous metabolic pathways which impair growth are genetically unstable, because selection favours cells in which the pathways are inactivated by spontaneous point mutations or deletions, until such cells predominate. This effect is expected in all organisms, but appears to be quite pronounced in cyanobacteria (38), perhaps due to the lengthy incubations during strain construction (required to allow genetic segregation, so polyploid strains become homozygous) and cultivation, which provide more opportunity for mutations to occur and to be selected (8,32). Furthermore, heterologous genes have often been strongly overexpressed in cyanobacteria by inserting them downstream of a strong native promoter (5,9,42–45), which is unlikely to be optimal and probably contributes to instability (39,46).

In principle, optimal target concentrations of each enzyme in a pathway could be predicted by metabolic modelling, although this is a substantial undertaking. Furthermore, it is not currently possible to reliably design an individual DNA sequence to produce a set of several enzymes at specific target concentrations, due to the limited predictability of the impact of combinations of different promoters, ribosome-binding sites (RBSs) and codon usage on protein synthesis, particularly in different sequence contexts (47) and in lesser-studied organisms. Fortunately, modular DNA assembly provides an alternative solution, by enabling the efficient construction of large libraries of variants of pathway-encoding constructs. By systematically varying the expression of each enzyme combinatorially, it is typically possible to identify pathway variants which perform well by screening. The same assembly methods allow for optimisation by rational replacement of individual parts to fine-tune expression. This combinatorial approach to construction and optimisation of metabolic pathway-encoding constructs is a recent development, and while its effectiveness has been clearly demonstrated in model organisms like *Escherichia coli* and *Saccharomyces cerevisiae* (40,41,48,49), it is not yet standard practice, and in many other organisms, including cyanobacteria, it has not yet been established.

Here, we develop a platform for assembly and optimisation of metabolic pathway-encoding constructs in cyanobacteria and apply it to light-driven lycopene production from atmospheric CO_2 in the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter referred to as *Synechocystis*). The platform uses the organism-independent Start-Stop Assembly system, and incorporates synthetic promoters and synthetic RBSs we generated and characterised in *Synechocystis*, along with terminators we validated previously (50). Lycopene is both a valuable product and a useful model for the challenges of producing terpenoids in photoautotrophs, which is complicated by potentially deleterious competition for common precursors with the biosynthesis of chlorophyll, carotenoids, quinones and tocopherols. A large combinatorial library of lycopene production pathway variants was constructed and used to study stability and production characteristics. Genetically stable lycopene accumulating strains were readily obtained. This platform represents a generally-applicable solution to the genetic instability of heterologous metabolic pathways in cyanobacteria, and opens the way to sustainable light-driven production of valuable products directly from CO_2.

MATERIALS AND METHODS

Bacterial strains and growth conditions

E. coli strain DH10B was used for all DNA assembly, all other cloning, and all other *E. coli* experiments described in this study. *E. coli* DH10B cells were routinely grown in LB medium (tryptone 10 g l^{-1}, yeast extract 5 g l^{-1} and NaCl 5 g l^{-1}) at 37°C, with shaking at 225 rpm, or on LB agar plates (containing 15 g l^{-1} bacteriological agar) at 37°C. LB was supplemented with ampicillin (100 μg ml^{-1}), tetracycline (10 μg ml^{-1}) or kanamycin (50 μg ml^{-1}) as appropriate. *E. coli* DH10B cells were routinely transformed by electroporation (51).

Synechocystis sp. PCC 6803 (the glucose-tolerant derivative of the wild type, obtained from the Nixon lab at Imperial College London) was routinely grown in TES-buffered (pH 8.2) BG11 medium (52) (photoautotrophic growth). Cultures were supplemented with 5 mM glucose as appropriate (mixotrophic growth). Cultures were grown at 30°C with agitation at 150 rpm and in constant white light at 50 μmol m^{-2} s^{-1}. BG11 was supplemented with kanamycin (30 μg ml^{-1}) as appropriate.

Construction of synthetic promoter library and synthetic RBS library

Tables of oligonucleotides (Supplementary Table S3), plasmids (Supplementary Table S4) and synthetic DNA (Supplementary Table S5) are provided in the Supplementary Information. The synthetic promoter library (SPL) and RBS libraries were generated by inverse PCR with 5′-phosphorylated degenerate primers using pATM2 and pGT77 (SPL clone 3) as templates, respectively. After PCR, DpnI was added to the PCR product to digest the template. The PCR product was circularised by ligation with T4 DNA ligase. The ligation product was used to transform *E. coli*. The SPL was generated using the degenerate primers

oligoATM3 and oligoATM5. The nine RBS libraries were generated using the degenerate primers oligoATM51-60.

Construction of *Synechocystis* Start-Stop Assembly destination vector

To generate pGT270, the pSHUTTLE2 backbone was PCR-amplified (oligoGT373 and oligoGT374) and assembled with the Level 2 Start Stop Assembly cassette PCR-amplified from pStA212 (oligoGT375 and oligoGT376) by Splicing by Overlap Extension PCR (SOE-PCR). The PCR product was treated with DpnI to digest the template, then circularised by ligation with T4 DNA ligase. The ligation product was used to transform *E. coli*. The plasmids of transformant colonies were purified and sequence verified.

Storing genetic parts in Start-Stop Assembly storage vector

Composite promoter-RBS parts were cloned in pStA0 to generate pGT287-290, 293–297, 300–303, 305–309, 311–315, 317–321 and 425–430 (Supplementary Table S6). This was achieved by inverse PCR with 5′- phosphorylated primers (oligoGT436-446 and oligoGT583) using pStA0 as a template. After PCR, DpnI was added to the PCR product to digest the template, the PCR product was circularised by ligation with T4 DNA ligase, and the ligation products were used to transform *E. coli*. The plasmids of transformant colonies were purified and sequence verified. To store in Start-Stop Assembly format one additional transcriptional terminator validated in *Synechocystis*, pGT424 (pStA0::ECK120015170) was generated by inverse PCR with 5′-phosphorylated degenerated primers (oligoGT579 and oligoGT578) using pStA0 as the template. DpnI was added to the PCR product to digest the template, and the PCR product was circularised by ligation with T4 DNA ligase, and the ligation product was used to transform *E. coli*. The plasmids of transformant colonies were purified and sequence verified.

Assembly of lycopene pathway constructs

The four individual lycopene pathway constructs (pGT432-435) and the combinatorial lycopene pathway construct library (pGT547) were assembled using Start-Stop Assembly (41). For each of the four individual constructs and the library, four Level 1 expression units encoding CrtI, CrtE, CrtB and DXS were assembled from parts or mixtures of parts. These four expression units were then assembled to generate the pathway-encoding constructs in the Level 2 destination vector pGT270. For pGT432, each of the four CDSs were assembled using promoter SPL clone 19 and RBS clone 21 (Supplementary Figure S10). For pGT433, each of the four CDSs were assembled using promoter SPL clone 3 and RBS clone 123 (Supplementary Figure S10). For pGT434, each of the four CDSs were assembled using promoter SPL clone 19 and RBS clone 48 (Supplementary Figure S10). For pGT435, each of the four CDSs were assembled using promoter SPL clone 45 and RBS clone 21 (Supplementary Figure S10). For the combinatorial library pGT547, an equimolar mixture of the 34 promoter-RBS composite parts (pGT287-290, 293–297, 300–303, 305–309,

311–315, 317–321 and 425–430) was used in Level 1 assemblies of the four CDSs. All white colonies from the transformation plates of the four Level 1 reactions were picked, resuspended as a mixture in P1 buffer and DNA was isolated using a miniprep procedure. The mixtures of Level 1 constructs were then used in the Level 2 assembly reaction to generate the lycopene pathway library (Supplementary Figure S12).

Synechocystis strain construction

Synechocystis cells were grown in BG11 medium and incubated under photoautotrophic conditions (30°C, 150 rpm, 50 μmol photons m^{-2} s^{-1} light) to an OD_{750} of 0.5. Following incubation, cells were harvested from 4 ml of culture by centrifugation at 3200 *g* for 15 min. Pellets were resuspended in 100 μl BG11, then 100 ng of plasmid DNA was added and the mixture was incubated at 100 μmol photons m^{-2} s^{-1} light for 60 min 30°C. Cells were transferred onto BG11 plates and incubated at 50 μmol photons m^{-2} s^{-1} light for 24 h at 30°C. Cells were collected and transferred onto BG11 plates supplemented with 30 μg ml^{-1} kanamycin. To achieve complete segregation, single colonies were subcultured four times on selective BG11 kanamycin plates, after which segregation was assessed by PCR using primers oligoAH48 and oligoAH49. Completeness of lycopene pathway constructs was assessed using three sets of primers: oligoAH34 and oligoGT603, oligoGT654 and oligoGT605, and oligoGT653 and oligoAH49.

Flow cytometry analysis of *Synechocystis*

Isolated transformant clones of *Synechocystis* were grown in 5 ml BG11 medium supplemented with kanamycin (30 μg ml^{-1}) and incubated in 25 ml vented culture flasks under photoautotrophic conditions (30°C, 150 rpm, 50 μmol photons m^{-2} s^{-1} light) for 48 h. Following incubation, cultures were diluted to an OD_{750} of 0.1 and incubated for a further 24 h. Cultures were diluted 1:50 in filtered phosphate-buffered saline (PBS) and were immediately subjected to flow cytometry analysis. A two-laser FACScan flow cytometer (Becton-Dickinson) equipped with an automated multi-well sampler (AMS) was used. The sample was analysed at a data rate of 1000–1500 events s^{-1} for 45 s. A 488 nm argon ion laser was used for excitation of the EYFP reporter. Detection used a 530 nm band-pass filter (FL1) with a gain of 700 V. For each sample of 1000–1500 events, the geometric mean of the distribution was used as the measure of central tendency. All part characterisation assays in *Synechocystis* were conducted using three biological replicates. The arithmetic mean was used as the measure of central tendency among the three replicates.

Flow cytometry analysis of *E. coli*

E. coli DH10B cells were transformed with purified plasmids. Individual transformant colonies were used to inoculate 200 μl LB broth supplemented with the appropriate antibiotic, in a 96 U-shaped 1.2 ml well plate covered with sterile breathable sealing film (Breathe Easy) and grown at

37°C at 700 rpm for 16 h in a Multitron shaker (Infors-HT). Cultures were subcultured by 1:1000 dilution into 200 μl fresh LB broth with the appropriate antibiotic and grown at 37°C at 700 rpm for 4 h. Cultures were diluted 1:50 in filtered PBS and were immediately subjected to flow cytometer analysis. A two-laser FACScan flow cytometer (Becton-Dickinson) equipped with an automated multiwell sampler (AMS) was used. The sample was analysed at a data rate of 1000–2000 events s^{-1} for 30 s. A 488 nm argon ion laser was used for excitation of the EYFP reporter. Detection used a 530 nm band-pass filter (FL1) with a gain of 900 V. For each sample of 1000–2000 events, the geometric mean of the distribution was used as the measure of central tendency. All part characterisation assays in *E. coli* were conducted using five biological replicates. The arithmetic mean was used as the measure of central tendency among the five replicates.

Spectrophotometric quantification of lycopene

E. coli DH10B cells were transformed with 5 μl of the combinatorial lycopene pathway library reaction product. Individual transformant colonies were used to inoculate 500 μl LB broth supplemented with the appropriate antibiotics, in a 96 U-shaped 1.2 ml well plate covered with sterile breathable sealing film (Breathe Easy) and grown at 30°C at 700 rpm for 48 h in a Multitron shaker (Infors-HT). Following incubation, the liquid cultures were centrifuged at 4000 rpm for 15 min before removing the supernatant and adding 100 μl of methanol and 200 μl of hexane to the bacterial cell pellets. After vortexing for 2 min, 100 μl of deionised water was added before vortexing the mixture again for 2 min. The solution underwent centrifugation at 4000 rpm for 20 min and the top hexane layer was transferred to a shallow 96-well plate and absorbance at 530 nm was determined (BMG LABTECH POLARstar Omega).

Growth of *Synechocystis* strains for lycopene measurement by LC-DAD

Isolated transformant clones of *Synechocystis* were grown in 5 ml BG11 medium supplemented with kanamycin (30 μg ml^{-1}) in 25 ml vented culture flasks under photoautotrophic conditions (30°C, 150 rpm, 50 μmol photons m^{-2} s^{-1} light) and incubated for 48 h. Following incubation, cultures were diluted to an OD$_{750}$ of 0.1 and incubated for a further two weeks.

Growth of *E. coli* strains for lycopene measurement by LC-DAD

E. coli DH10B cells were transformed with purified plasmids. Individual transformant colonies were used to inoculate 5 ml LB media supplemented with the appropriate antibiotic and grown at 30°C at 225 rpm for 24 h. Cultures were subcultured by 1:1000 dilution into 5 ml fresh LB and grown for 48 h.

Dry cell weight measurements

Dry cell weight (DCW) was determined by centrifugation of 2 ml of liquid culture at 16 900 *g* for 5 min to remove the supernatant. The bacterial cell pellet was washed in deionised water and then centrifuged again and dried at 70°C until a constant weight was obtained.

Lycopene extraction

To extract lycopene, 2 ml of liquid culture was centrifuged at 14 500 *g* for 5 min to remove the supernatant. The bacterial cell pellet was washed in deionised water and then centrifuged as before. The bacterial cell pellet was resuspended in 1 ml acetone and incubated at 55°C for 15 min to extract lycopene. The supernatant was obtained by filtration through a 0.22 μm pore-size nylon membrane for LC-DAD analysis.

LC-DAD quantification of lycopene

Lycopene was detected and measured using an Agilent LC system with UV/Vis diode array detector. Absorbance at 450 nm and 471 nm were monitored and the peak area corresponding to each component integrated to provide a measure of abundance. The LC column used was an Acquity UPLC Peptide BEH C18 column (2.1 × 100 mm, 1.7 μm, 300 Å, Waters). LC buffers were 50% (v/v) methanol in water (A) and 25% (v/v) ethyl acetate in acetonitrile (B). All solvents used were HPLC grade. The LC method was 6.5 min in total (0–1 min: 30% A, 70% B; 1–6 min: 0.1% A, 99.9% B; 6–6.5 min: 30% A, 70% B; at a flow rate of 0.3 ml/min) with 1.5 min of post-run time. The injection volume for the samples was 1.0 μl. Commercially available lycopene (Sigma-Aldrich) was dissolved in acetone as a standard and a standard curve was generated.

RESULTS

A platform for assembly and optimisation of metabolic pathways in cyanobacteria

To develop a general-purpose platform for assembly and optimisation of metabolic pathways in cyanobacteria we required both an efficient multi-part DNA assembly method and sets of expression control parts (promoters, RBSs and transcriptional terminators) with suitable properties. Several DNA assembly methods are suitable for combinatorial assembly (53), but most are designed for a specific organism. Start-Stop Assembly (41) offers an ideal general-purpose DNA assembly system, with a unique streamlined assembly hierarchy which typically requires construction of only one new vector to apply the system to a new organism. Therefore we constructed a new Level 2 Start-Stop Assembly vector pGT270 (Supplementary Figure S1) for the cyanobacterium *Synechocystis*.

At the start of this study there was a lack of promoter and RBS parts suitable for controlling and tuning gene expression in cyanobacteria (although some examples have now been published (54,55)). Therefore we set out to generate synthetic libraries and screen samples of these in order to identify defined sets of these expression control parts providing an evenly-distributed, wide range of strengths suitable for controlling and tuning expression levels in cyanobacteria.

A synthetic promoter library (SPL) was designed by the method of Jensen and Hammer (56), which unlike classic

approaches (57–59) does not mutate the consensus sigma factor-binding −35 or −10 regions, but instead randomises the sequences surrounding these core elements. This is a proven method for generating libraries of synthetic promoters which span an evenly-distributed, wide range of strengths (56,60,61) allowing sets of promoters with desired properties to be obtained from small samples. An advantage of obtaining such sets of promoters by random isolation from very large degenerate libraries (instead of, for example, conventional individual 'down mutations' of natural promoters) is that all the isolated promoters are likely to have very different sequences, so will not cause homologous recombination problems if used together in the same constructs.

A total of 33 bp surrounding the −35 (TTGACA) and −10 (TATAAT) elements of a Type 1 (SigA σ^{70}-dependent) *Synechocystis* promoter (62) (Figure 1B) were randomised using degenerate primers (giving a library size of 7.4×10^{19}) and introduced in front of a yellow fluorescent protein coding sequence (CDS) reporter (*eyfp*) in the *E. coli-Synechocystis* vector pATM2 (Figure 1A). *E. coli* was transformed with the SPL then one hundred transformant colonies were picked (70 randomly and 30 spanning a range of fluorescence levels by visual inspection) and a screening cascade of flow cytometry and sequencing was used for validation (Figure 1A). Promoter clones with unintended mutations or which were non-functional or multimodal (showing sub-populations with differing fluorescence) in *E. coli* were discarded, resulting in a set of 37 promoter clones (Supplementary Table S1).

Synechocystis was transformed separately with each of the 37 promoter clone plasmids then transformants were subcultured until completely segregated (homozygous, Figure 1A). The fluorescence of these 37 segregated *Synechocystis* strains was determined by flow cytometry of mid-linear phase cultures grown under photoautotrophic conditions. We observed a wide distribution of promoter strengths with small increments of strength (Figure 1B, with clone numbers labelled in Supplementary Figure S8), which ranged from very weak (SPL clone 115) to 27-fold stronger (SPL clone 25). Interestingly, over 80% of the 37 synthetic promoters were stronger than the positive control strong native cyanobacterial promoter P$_{psbA2}$ (63) (pATM10), and the strongest synthetic promoter, SPL clone 25, was three times stronger. The 37 promoter clones were also characterised in *E. coli* where a wide distribution of promoter strengths was again observed (Supplementary Figure S2), although interestingly no correlation ($R^2 = 0.05$) was observed between promoter strengths in *Synechocystis* and *E. coli* (Supplementary Figure S3).

Next we set out to generate RBSs. In principle, computational tools such as the RBS Calculator (64) can be used to design specific RBS sequences with desired strengths tailored to a specific CDS. In practise, the published RBS design tools remain limited in their predictive ability, particularly for lesser-studied organisms such as cyanobacteria (55,65). Furthermore, for combinatorial assembly, prediction of the strengths of individual RBSs is less important than the diversity of strengths. Therefore, we chose a degenerate approach to design a synthetic RBS library, similar to the promoter library.

To determine what degenerate RBS library design would give a convenient broad distribution of RBS strengths, allowing a desired set of RBSs to be obtained from a small sample, a series of nine RBS libraries was generated (Supplementary Figure S4). Each differed only in the number of degenerate positions, from 20, giving a library size of 1.1×10^{12} (RBS library 1) to 29, giving a library size of 2.9×10^{17} (RBS library 9). The nine RBS libraries were introduced between a mid-strength promoter (SPL clone 3) from the promoter library and the *efyp* CDS in the *E. coli-Synechocystis* vector pGT77 (Figure 1A). *E. coli* was transformed with the nine RBS library assembly products and the fluorescence distributions among populations of cells containing each RBS library were compared in *E. coli* by flow cytometry (Supplementary Figure S4). A wide range of fluorescence intensities was observed in the RBS libraries with fewest degenerate positions, but as the number of degenerate positions increased, the range of fluorescence intensities narrowed, shifting the population towards that of the negative control (pATM2). A hundred transformant colonies were picked (70 randomly and 30 spanning a range of fluorescence levels by visual inspection) from RBS libraries 1, 4 and 9 and initially assessed in singlicate using the same screening cascade as before (Figure 1A) resulting in a set of 58 RBS clones (Supplementary Table S2), which were then characterised in triplicate, showing a wide distribution of strengths (Supplementary Figure S5).

Synechocystis was transformed separately with each RBS clone and the resulting strains were characterised by flow cytometry as before. We observed a wide distribution of fluorescence intensities that increased in small increments (Figure 1C) from very weak (RBS clone 249) to very strong (RBS clone 4) over a 152-fold range. A very weak correlation ($R^2 = 0.34$) was observed between RBS strengths in *Synechocystis* and *E. coli* (Supplementary Figure S6). Interestingly, the 58 RBS clones showed no correlation ($R^2 = 0.153$) between their measured fluorescence and their predicted ΔG_{tot} in either *Synechocystis* ($R^2 = 0.153$) or *E. coli* ($R^2 = 0.149$) (Supplementary Figure S7), as calculated by the reverse engineering mode of the RBS Calculator v2.0 (64), which is consistent with previous reports (55,65).

Using the entire set of 37 promoters and 58 RBSs would give up to 2146 expression levels per CDS, which would be excessive and redundant. Instead we chose a subset of six promoters and six RBSs evenly spanning the range of available strengths (Figure 1B and C, red bars) with the aim of generating a total of 36 diverse expression levels per CDS, allowing broad sampling of design space.

Rationally-designed individual pathway constructs were genetically unstable

The synthetic promoters, RBSs and Start-Stop Assembly represent a platform for assembly of metabolic pathways in cyanobacteria. To test this platform, and to tackle genetic instability, increased production of the terpenoid lycopene was adopted as a model system (Figure 2). Lycopene is a valuable product used in the pharmaceutical, food and cosmetic industries (66) with a strong red colour which can be used as a measure of pathway activity (40,41,67) including by visual inspection of colony colour, at least
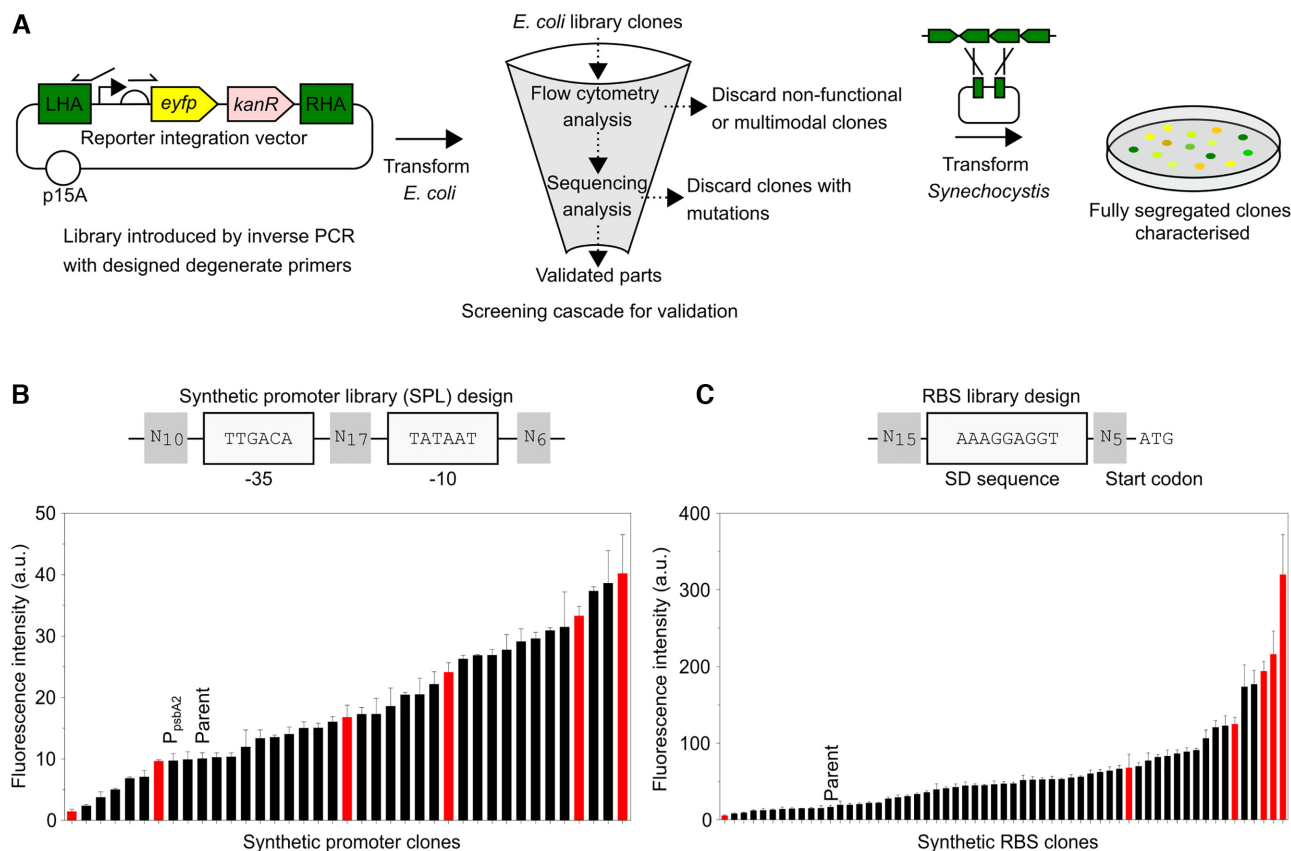
Figure 1. Parts enabling rational and combinatorial pathway assembly in cyanobacteria. (**A**) Workflow of library design, generation and characterisation. Libraries were generated by inverse PCR and *E. coli* was transformed with each library. Clones were picked (70 randomly and 30 spanning a range of fluorescence levels by visual inspection) and a screening cascade of flow cytometry and sequencing was used for validation. Clones were discarded if their fluorescence was multimodal or weaker than pATM2 (−ve control) or if they contained unintended mutations (at non-degenerate positions). *Synechocystis* was transformed and integration and full segregation at the previously-used *ndhB* locus (ssl0410 CDS) neutral site was confirmed before characterisation. (**B**) Synthetic promoter library (SPL) design conserved −35 and −10 consensus sequences, and randomised the surrounding sequences. *Synechocystis* SPL clones were characterised by flow cytometry (detail in Supplementary Figure S8). A subset of six promoters was chosen (red bars). 'Parent' plasmid pATM2 was used as the template to generate the SPL. (**C**) RBS library design. The Shine-Dalgarno sequence was conserved and the surrounding sequences were randomised. *Synechocystis* RBS library clones were characterised by flow cytometry (detail in Supplementary Figure S9). A subset of six RBSs was chosen (red bars). 'Parent' plasmid pGT77 was used as the template to generate the RBS library. Error bars represent the standard deviation of three independent biological replicates.

in *E. coli*. Lycopene is produced endogenously by *Synechocystis* and other cyanobacteria as a precursor of all other carotenoids used for light harvesting and photoprotection, but does not naturally accumulate (Figure 2). A pathway-encoding construct was designed using homologs of three native CDSs (*dxs*, *crtE*, and *crtB*) to boost pathway flux (*dxs* is a known rate-limiting reaction in both *E. coli* (40,68,69) and cyanobacteria (14,15,70)) and a heterologous CDS (*ctrI*), which encodes a bacterial-type one-enzyme, four-step phytoene desaturase system to bypass the native CrtP and CrtQ desaturases (71) in the conversion of phytoene to lycopene (Figure 2).

We designed four individual lycopene pathway-encoding constructs for *Synechocystis* using different design strategies, in order to test their effects on genetic stability, aiming to identify a strategy allowing stable expression and enhanced lycopene production. The first design strategy (pGT432) used a traditional strong overexpression rationale with the strongest promoter and RBS driving expression of each CDS. The second design strategy (pGT433)

aimed to avoid excessive expression by using medium-strength promoters and RBSs for all four CDSs. The third design strategy (pGT434) followed recently-described principles for minimising metabolic burden, which have been demonstrated in *E. coli* (39). This design uses a strong promoter and weak RBS to reduce 'overinitiation' of translation, because high translation is more stressful for the cell than high transcription. As a high-burden control, the fourth pathway design strategy (pGT435) reversed this principle, using a weak promoter and strong RBS.
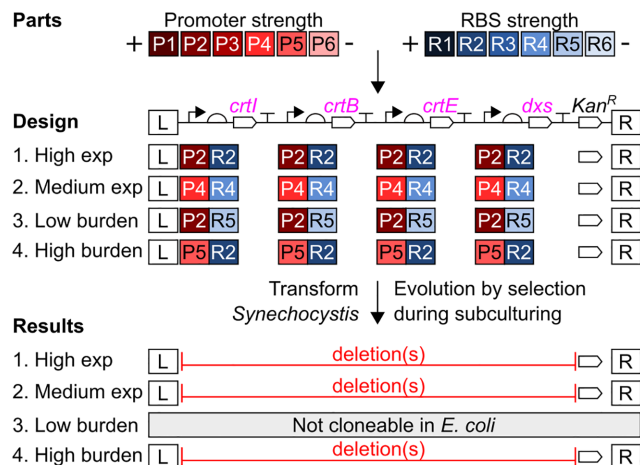
We opted to express each CDS in a monocistronic configuration (with its own promoter and terminator) to maximise their independence, and therefore included four terminators from a set we recently characterised in *Synechocystis* (50). The parts (promoters, RBSs, CDSs and terminators) were each stored and verified in Start-Stop Assembly format and the four pathways (pGT432-435) were individually assembled by hierarchical Start-Stop Assembly (Figure 3 and Supplementary Figure S10). After three independent attempts it was not possible to assemble the suppos-
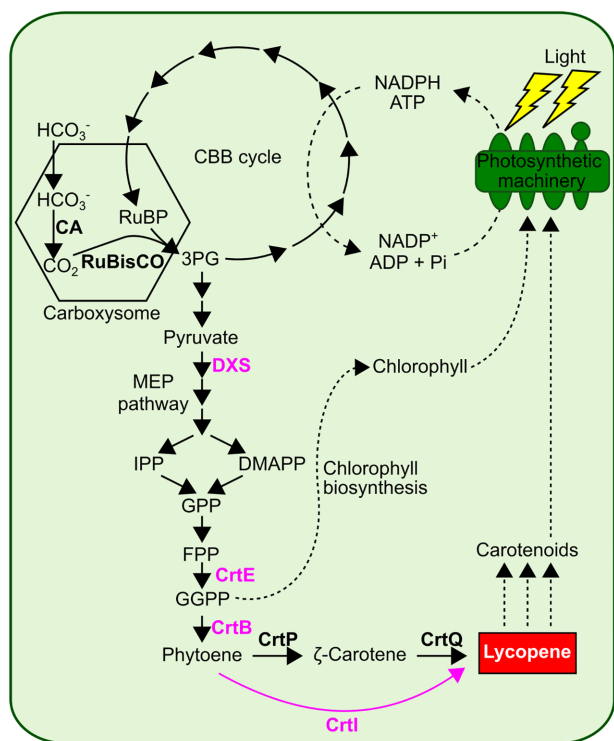
Figure 2. Metabolic design for enhanced lycopene production in *Synechocystis*. The lycopene overproduction construct encodes four enzymes shown in pink. Native metabolism of *Synechocystis* is shown in black. DXS (1-deoxy-D-xylulose 5-phosphate synthase) catalyses the synthesis of 1-deoxy-D-xylulose 5-phosphate from pyruvate and glyceraldehyde 3-phosphate. GGPP (Geranylgeranyl diphosphate) is synthesised by CrtE (Geranylgeranyl diphosphate synthase) which catalyses the condensation between FPP (Farnesyl diphosphate) and IPP (Isopentenyl diphosphate). CrtB (phytoene synthase) then catalyses the condensation of two GGPP molecules to generate phytoene. Finally, CrtI (phytoene desaturase) catalyses the synthesis of lycopene (shown in red box) from phytoene, bypassing the native two-enzyme desaturation pathway (CrtP and CrtQ) via ζ-carotene.

Figure 3. Rational design and genetic instability of individual lycopene overproduction constructs for *Synechocystis*. Four individual pathway constructs were designed: high expression (pGT432), medium expression (pGT433), low burden (pGT434) and high burden (pGT435). Each construct was individually assembled using Start-Stop Assembly in the destination vector pGT270 (Supplementary Figure S10). pGT434 was not cloneable in *E. coli* (in three independent attempts). *Synechocystis* was individually transformed with each pathway and PCR screening was used to assess genetic stability (detail in Supplementary Figure S11). In each case flanking sequences allowing integration at the *ndhB* neutral site in the *Synechocystis* genome were present and antibiotic selection pressure maintained the kanamycin resistance gene *kan^R*, but gross deletions in the lycopene pathway-encoding constructs were observed. Promoters in descending order of strength as characterised using EYFP: P1 = SPLc25, P2 = SPLc19, P3 = SPLc47, P4 = SPLc3, P5 = SPLc45 and P6 = SPLc17. RBSs in descending order of strength as characterised using EYFP: R1 = RBSc4, R2 = RBSc21, R3 = RBSc110, R4 = RBSc123, R5 = RBSc48 and R6 = RBSc249.

edly low-burden design pGT434, suggesting that unexpectedly it is toxic to *E. coli*. The other three designs (pGT432, pGT433 and pGT435) were successfully assembled, giving red-coloured *E. coli* colonies showing lycopene production.

To assess the high expression (pGT432), medium expression (pGT433) and high burden (pGT435) lycopene pathway constructs, *Synechocystis* was transformed separately with each plasmid and serially subcultured under antibiotic selection in an attempt to obtain fully segregated (homozygous) clones. Individual transformant colonies appeared more slowly than usual, after approximately two weeks instead of one, suggesting a burdensome effect of the pathways. After four subcultures we assessed genetic stability (Figure 3 and Supplementary Figure S11) and found that all three remaining lycopene pathway designs had acquired gross deletions which had become homozygous, presumably due to selection pressure acting against the burden imposed by the constructs. Therefore, due to genetic instability either in *E. coli* (pGT434) or *Synechocystis* (pGT432, pGT433 and pGT435), none of the four individual pathway designs was suitable for sustained overproduction of lycopene in *Synechocystis*, so an alternative approach was needed.

Combinatorial design overcomes genetic instability

Next we explored a combinatorial approach to address the genetic instability observed with the individually-designed pathways. Using the same assembly system, monocistronic architecture and gene order as before, we designed a library of lycopene pathway encoding constructs in which the promoter and RBS of each CDS was varied combinatorially. The six promoters and six RBSs identified earlier were pre-assembled as composite promoter-RBS parts, of which 34 of 36 possible combinations were cloneable, giving a maximum combinatorial library size of 1.3×10^6. The pathway library (pGT547) was generated using hierarchical Start-Stop Assembly with suitable mixtures of parts used in place of individual parts in order to introduce the combinatorial diversity at Level 1, and to propagate the diversity to the complete pathways at Level 2 (Figure 4, Supplementary Figure S12). It was interesting to determine whether prescreening in *E. coli* would provide a useful prediction of performance in cyanobacteria, especially given that strain construction and analysis in *E. coli* is so much more rapid and convenient. Therefore, from the *E. coli* transformation, 20 clones (pGT437-456) were picked at random and 20 clones (pGT457-476) were chosen by prescreening 96 (randomly-picked) clones for their lycopene production in *E. coli*. Of the 20 chosen by prescreening in *E. coli*, ten were the strongest lycopene producers and the other ten were evenly-distributed throughout the observed range of
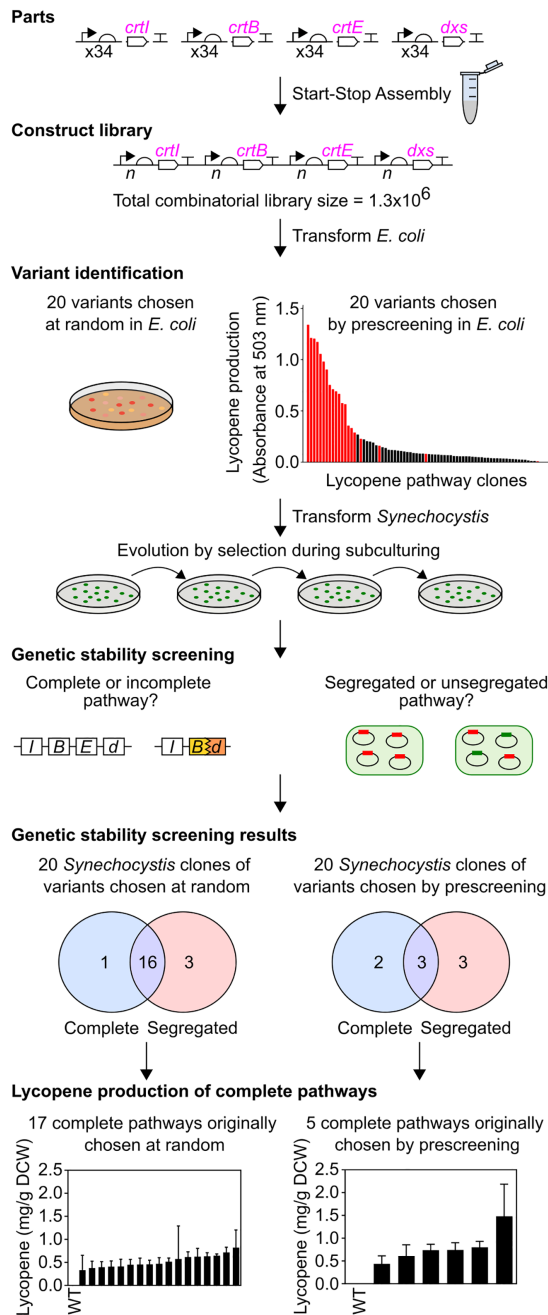
Figure 4. Genetic design, selection and evaluation of combinatorial lycopene pathway library. A library of lycopene pathway construct variants was assembled using Start-Stop Assembly in the destination vector pGT270. For combinatorial assembly equimolar mixtures of 34 composite promoter-RBS parts labelled 'x34' were used in Level 1 assemblies (Supplementary Figure S12). The uncertain representation of each promoter and RBS at each position following assembly is represented by 'n'. The maximum library size was $34^4 = 1.3 \times 10^6$. *E. coli* was transformed with the combinatorial library and pathway clones were either chosen at random or after prescreening in *E. coli* (see text). *Synechocystis* was transformed with the chosen pathway variants, after which PCR screening was used to assess the completeness and segregation of the resultant clones (PCR screening details shown in Supplementary Figure S11). *Synechocystis* clones containing complete pathway constructs were grown in photoautotrophic batch cultures with constant light for two weeks, then lycopene accumulation was measured (detail in Supplementary Figure S15). Lycopene concentration was determined as mg lycopene per g dry cell weight (DCW). Error bars represent the standard deviation of three independent biological replicates.

lycopene concentrations (Figure 4 and Supplementary Figure S13).

Synechocystis was transformed separately with each of the 40 library variant plasmids and subcultured under antibiotic selection. As previously observed, transformant colonies appeared slightly more slowly than usual, suggesting metabolic burden, and after four subcultures we assessed genetic stability in terms of completeness and segregation as before (Figure 4). Interestingly, 16 of the 20 randomly-chosen pathway variants were genetically stable (complete and segregated) in *Synechocystis* (Figure 4), whereas only three of the 20 pathway variants chosen by prescreening in *E. coli* were stable (Figure 4), as the other 17 contained deletions and/or failed to segregate.

The lycopene accumulation of the 22 *Synechocystis* clones containing complete pathway constructs was determined using small-scale batch cultures grown photoautotrophically for two weeks (Figure 4). Lycopene accumulation was detected in all 22 clones (unlike the wild type) but only small differences in lycopene levels were observed among the pathway variants, ranging between 0.3 and 1.5 mg/g DCW (Figure 4). Interestingly, the same 22 pathway variants gave considerably greater diversity of lycopene production in *E. coli*, ranging between 0 and 40 mg/g DCW (Supplementary Figure S14).

Promoter and RBS preferences are associated with coding sequences, not genetic stability

To assess the impact of the combinatorial design, the 40 pathway variants used to transform *Synechocystis* were sequenced. All expected parts were present in multiple variants. Some deletions or misassemblies could be identified (Supplementary Figure S17 and S18). As diversity of expression was introduced using the same mixture (the same tube) of promoters and RBSs for each CDS during the assembly (Supplementary Figure S12), any substantial differences in frequencies of parts among the complete pathway constructs may reflect biological selection rather than assembly bias (Figure 5 and Supplementary Figure S19). There were no obvious differences in part frequencies between the segregated and non-segregated variants, one of the key aspects of genetic stability. In contrast, part frequencies clearly differ among the four CDSs, for both segregated and non-segregated variants. For example, we observed all six RBSs for *crtB* whereas for *dxs* we only observed three of the six RBSs. This observation is consistent with biological selection, which has presumably led to a preference towards certain expression levels.

Increasing precursor supply enhances lycopene accumulation

Glucose supplementation has been used to increase the supply of precursors to product-forming pathways (72), so we tried this approach to increase lycopene accumulation by the engineered strains, and to test the hypothesis that the low diversity of lycopene accumulation in *Synechocystis* might be caused by limited precursor availability. We compared the lycopene production of low (pGT439), medium (pGT455) and high (pGT462) producing strains (Supplementary Figure S16) in photoautotrophic (light, no glucose) and mixotrophic (light and glucose) conditions which
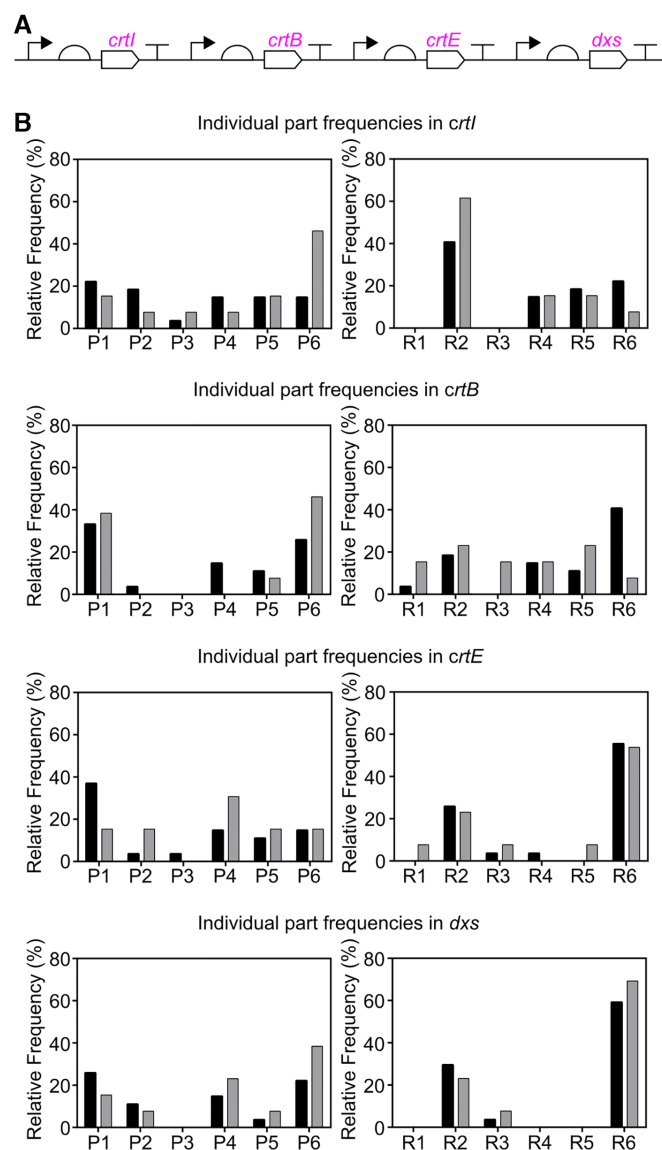
Figure 5. Frequencies of promoter and RBS parts in pathway variants. All 40 pathway variants used to transform *Synechocystis* were sequenced and the promoter and RBS upstream of each CDS were identified. (**A**) Schematic of construct library variants. (**B**) Relative frequencies of promoter and RBS parts: In segregated pathway variants, black bars; in nonsegregated pathway variants, grey bars. Promoters, in descending order of strength as characterised using EYFP: P1 = SPLc25, P2 = SPLc19, P3 = SPLc47, P4 = SPLc3, P5 = SPLc45 and P6 = SPLc17. RBSs, in descending order of strength as characterised using EYFP: R1 = RBSc4, R2 = RBSc21, R3 = RBSc110, R4 = RBSc123, R5 = RBSc48 and R6 = RBSc249.

boost precursor supply (72) (Figure 6). Supplementary glucose successfully increased accumulation by the medium and high producing strains (Figure 6). The modest level of increase suggests that the low diversity of lycopene accumulation in the library is not caused mainly by limited precursor availability, and may instead be due to increased flux from lycopene to other terpenoids or low tolerance to changes in production of these physiologically important products.

DISCUSSION

Genetic stability is essential for industrial microorganisms, so the reported genetic instability of genetically modified cyanobacteria (38) is a key challenge for cyanobacterial synthetic biology (8,32,73–76). Despite its importance, genetic instability of heterologous DNA in cyanobacteria has not been systematically studied. Several aspects of the present study provide new insight into this problem. Firstly, we used a large-scale combinatorial approach to metabolic pathway construction. Such approaches have proven useful for simultaneously optimising expression of multiple genes in other organisms (40,41,48,49), but have not been established in cyanobacteria. Only three small-scale examples of combinatorial designs have been reported, the first varying a small number of alternative enzyme-encoding sequences (65) and the other two varying only the RBSs in a two-enzyme pathway (77,78). In these cases, the total number of variants constructed was small (5–25 variants) compared to the libraries of millions of variants described in *E. coli* and *S. cerevisiae* (40,41,49). Here, construction of a large library (over a million variants) and evaluation of numerous variants (rather than the typical one or two) provided more scope to overcome the low tolerance to heterologous genes and/or poor control of expression presumably involved in reported examples of genetic instability. Secondly, increased production of the target product lycopene is likely to be sensitive to genetic instability, and hence able to reveal design issues. Lycopene is a terpenoid, one of the largest and most diverse classes of natural products, with existing or emerging applications in the pharmaceutical, flavour, fragrance and fuel industries (66). However, modifying production of terpenoids in photoautotrophs is challenging, because GGPP is a precursor of carotenoids, chlorophyll, ubiquinones, plastoquinone and phylloquinone (79). These pigments/cofactors are important components of the photosynthetic machinery, involved in their function and/or protection (79), and are therefore essential in photoautotrophs. As a result, the production of any terpenoid competes with the biosynthesis of these essential products. Thirdly, our constructs allowed direct comparisons between *E. coli* and *Synechocystis* to be made at each stage, from the performance of individual promoter and RBS parts, to the relative lycopene accumulation of assembled pathway variants, to their associated genetic stability.

Genetic instability was prevalent in the combinatorial lycopene pathway library, as anticipated. However, many variants were genetically stable, over many generations during subculturing in the segregation process, and yet crucially many of these stable variants were also able to accumulate lycopene. Due to the slower than usual growth observed, this subculturing took approximately eight weeks, which effectively represents a laboratory evolution experiment, and a substantial window of opportunity for cells to escape the burden imposed by variants through spontaneous mutation. Genetic stability despite this long evolution window indicates real robustness and bodes well for the prospect of overproducing other terpenoids of interest. The high-producing strain (pGT462) gave a lycopene titre of 6 mg/L (equating to 1.48 mg/g DCW in Figure 4), which
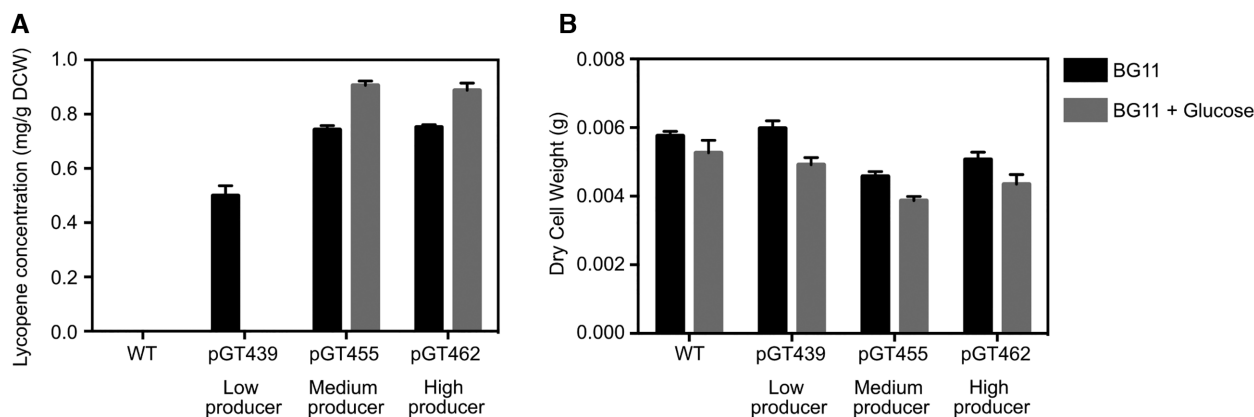
Figure 6. Effect of glucose supplementation on lycopene accumulation. The low, medium and high producing *Synechocystis* clones were grown in photoautotrophic conditions (light, no glucose; black bars) or mixotrophic conditions (light and glucose; grey bars) in batch cultures for two weeks with constant light, then (**A**) lycopene production and (**B**) dry cell weight (DCW) were determined. Error bars represent the standard deviation of three independent biological replicates.

is similar to reported titres for other terpenoids from engineered cyanobacteria (80), with the exception of one report of high production of isoprene (81).

The synthetic promoters characterised using the fluorescent reporter behaved very differently in *E. coli* and *Synechocystis* (Supplementary Figure S3). The RNA polymerases of these organisms differ somewhat (82) although the complete lack of correlation was surprising. The synthetic RBSs did show a modest correlation between *E. coli* and *Synechocystis* (Supplementary Figure S6), although perhaps weaker than might have been expected given the near-identical anti-Shine Dalgarno sequences of the two organisms (83). Besides RBS strength (translation initiation rate), other factors such as mRNA structure and ribosome availability may have contributed to the variance observed between these sets of constructs in the two organisms. The strengths of these synthetic RBSs were not computationally predictable for either organism (Supplementary Figure S7). It was therefore to be expected that the pathway variants combinatorially assembled using these parts produced different relative amounts of lycopene in *E. coli* and *Synechocystis* (Supplementary Figure S20), particularly given the profoundly different physiology, metabolism and growth rates of these organisms (84). Interestingly, most of the pathway variants which produced large amounts of lycopene during prescreening in *E. coli* were not genetically stable in *Synechocystis* (Figure 4). These apparently conflicting observations might be explained if there is an underlying correlation between the performance of pathway variants in the two organisms, but *Synechocystis* is more sensitive to metabolic burden caused by high-expressing or high-producing variants, leading to mutations which in turn prevent higher production being observed. This would be consistent with both the reported genetic instability of recombinant cyanobacteria, and with the expected sensitivity of cyanobacteria to perturbations in the production of lycopene (Figure 3).

This work directly shows that the previously anecdotally-reported genetic instability of heterologous DNA in cyanobacteria (38) is real, but the large fraction of stable overproducers in our sample (80% of variants chosen at random, 15% of those chosen by prescreening in *E. coli*) reveals that the problem can be overcome through suitable genetic design. The lack of association between the frequencies of promoter and RBS parts and the genetic stability of variants containing them (Figure 5) does not provide principles for the rational design of individual genetically stable constructs. Faced with limited predictive ability in a design-build-test-learn (DBTL) cycle framework, improvements could be obtained either by accelerating the cycles or taking highly parallel approaches (49). CRISPR-Cas9 has been used to accelerate segregation of recombinant cyanobacteria (85), and chromosomal integration can be replaced by replicative plasmids, although these are inherently unstable (8). In this first cyanobacterial example of a highly parallel approach to metabolic pathway engineering, pathway construct variants with the desired characteristics (genetic stability combined with target accumulation) were obtained in a single round. This is particularly noteworthy and useful in relatively slow-growing organisms like cyanobacteria, in which multiple iterative DBTL cycles could be prohibitively time-consuming.

Although previous studies have not directly dissected genetic instability in recombinant cyanobacteria (38), the problem is indirectly addressed by an emerging body of research on coupling production of target compounds with growth, in order to generate a selection pressure for sustained production (86,87). Such selection can only act on the genetic variants it encounters, so if used with suboptimal conventional individual pathway designs, the growth coupling strategy could force cells to maintain burdensome pathway-encoding constructs, leading to poor performance. However, if growth coupling were used in combination with the combinatorial approach reported here, efficient systems might be developed, giving optimal performance in terms of productivities and yields, along with long-term stability.

This study reveals simple principles for rational design of combinatorial libraries of pathway-encoding constructs, from which individual variants with desired properties can be obtained by screening a modest number of variants:

Synthetic promoter and RBS parts are generated, characterised and used to implement a broad, sparse sampling of design space with few prior assumptions. The parts and vectors needed for this platform are available to the community, and should allow cyanobacterial synthetic biology to much more effectively tackle the genetic instability 'elephant in the room' (38) and develop stable cyanobacterial biocatalysts for sustainable light-driven production of valuable products directly from CO_2.

DATA AVAILABILITY

Nucleic acid sequences are available in the Supplementary Data. Plasmids are available from Addgene. Flow cytometry data are available from FlowRepository with ID numbers FR-FCM-Z3NQ, FR-FCM-Z3NR, FR-FCM-Z3XZ, FR-FCM-Z3XY and FR-FCM-Z3X2.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Schuurmans,R.M., van Alphen,P., Schuurmans,J.M., Matthijs,H.C.P. and Hellingwerf,K.J. (2015) Comparison of the photosynthetic yield of cyanobacteria and green algae: Different methods give different answers. *PLoS One*, **10**, e0139061.
2. Lürling,M., Eshetu,F., Faassen,E.J., Kosten,S. and Huszar,V.L.M. (2013) Comparison of cyanobacterial and green algal growth rates at different temperatures. *Freshw. Biol.*, **58**, 552–559.
3. Melis,A. (2009) Solar energy conversion efficiencies in photosynthesis: minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.*, **177**, 272–280.
4. Knoot,C.J., Ungerer,J., Wangikar,P.P. and Pakrasi,H.B. (2018) Cyanobacteria: promising biocatalysts for sustainable chemical production. *J. Biol. Chem.*, **293**, 5044–5052.
5. Gao,Z., Zhao,H., Li,Z., Tan,X. and Lu,X. (2012) Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy Environ. Sci.*, **5**, 9857–9865.
6. Namakoshi,K., Nakajima,T., Yoshikawa,K., Toya,Y. and Shimizu,H. (2016) Combinatorial deletions of *glgC* and *phaCE* enhance ethanol production in *Synechocystis* sp. PCC 6803. *J. Biotechnol.*, **239**, 13–19.
7. Lan,E.I., Ro,S.Y. and Liao,J.C. (2013) Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy Environ. Sci.*, **6**, 2672–2681.
8. Yunus,I.S. and Jones,P.R. (2018) Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols. *Metab. Eng.*, **49**, 59–68.
9. Angermayr,S.A. and Hellingwerf,K.J. (2013) On the use of metabolic control analysis in the optimization of cyanobacterial biosolar cell factories. *J. Phys. Chem. B*, **117**, 11169–11175.
10. Xue,Y., Zhang,Y., Grace,S. and He,Q. (2014) Functional expression of an Arabidopsis p450 enzyme, p-coumarate-3-hydroxylase, in the cyanobacterium *Synechocystis* PCC 6803 for the biosynthesis of caffeic acid. *J. Appl. Phycol.*, **26**, 219–226.
11. Kawahara,A., Sato,Y., Saito,Y., Kaneko,Y., Takimura,Y., Hagihara,H. and Hihara,Y. (2016) Free fatty acid production in the cyanobacterium *Synechocystis* sp. PCC 6803 is enhanced by deletion of the cyAbrB2 transcriptional regulator. *J. Biotechnol.*, **220**, 1–11.
12. Oliver,J.W.K., Machado,I.M.P., Yoneda,H. and Atsumi,S. (2013) Cyanobacterial conversion of carbon dioxide to 2, 3-butanediol. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 1249–1254.
13. Li,H. and Liao,J.C. (2013) Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO_2 to 1,2-propanediol. *Microb. Cell Fact.*, **12**, 4.
14. Englund,E., Andersen-Ranberg,J., Miao,R., Hamberger,B. and Lindberg,P. (2015) Metabolic engineering of *Synechocystis* sp. PCC 6803 for production of the plant diterpenoid manoyl oxide. *ACS Synth. Biol.*, **4**, 1270–1278.
15. Englund,E., Shabestary,K., Hudson,E.P. and Lindberg,P. (2018) Systematic overexpression study to find target enzymes enhancing production of terpenes in *Synechocystis* PCC 6803, using isoprene as a model compound. *Metab. Eng.*, **49**, 164–177.
16. Formighieri,C. and Melis,A. (2017) Heterologous synthesis of geranyllinalool, a diterpenol plant product, in the cyanobacterium *Synechocystis*. *Appl. Microbiol. Biotechnol.*, **101**, 2791–2800.
17. Davies,F.K., Work,V.H., Beliaev,A.S. and Posewitz,M.C. (2014) Engineering limonene and bisabolene production in wild type and a glycogen-deficient mutant of *Synechococcus* sp. PCC 7002. *Front. Bioeng. Biotechnol.*, **2**, 21.
18. Bentley,F.K., García-Cerdán,J.G., Chen,H.-C. and Melis,A. (2013) Paradigm of monoterpene (β-phellandrene) hydrocarbons production via photosynthesis in cyanobacteria. *Bioenergy Res.*, **6**, 917–929.
19. Reinsvold,R.E., Jinkerson,R.E., Radakovits,R., Posewitz,M.C. and Basu,C. (2011) The production of the sesquiterpene β-caryophyllene in a transgenic strain of the cyanobacterium *Synechocystis*. *J. Plant Physiol.*, **168**, 848–852.
20. Englund,E., Pattanaik,B., Ubhayasekera,S.J.K., Stensjö,K., Bergquist,J. and Lindberg,P. (2014) Production of squalene in *Synechocystis* sp. PCC 6803. *PLoS One*, **9**, e90270.
21. Betterle,N. and Melis,A. (2019) Photosynthetic generation of heterologous terpenoids in cyanobacteria. *Biotechnol. Bioeng.*, **116**, 2041–2051.
22. Du,W., Liang,F., Duan,Y., Tan,X. and Lu,X. (2013) Exploring the photosynthetic production capacity of sucrose by cyanobacteria. *Metab. Eng.*, **19**, 17–25.
23. Ducat,D.C., Avelar-Rivas,J.A., Way,J.C. and Silver,P.A. (2012) Rerouting carbon flux to enhance photosynthetic productivity. *Appl. Environ. Microbiol.*, **78**, 2660–2668.
24. Niederholtmeyer,H., Wolfstädter,B.T., Savage,D.F., Silver,P.A. and Way,J.C. (2010) Engineering cyanobacteria to synthesize and export hydrophilic products. *Appl. Environ. Microbiol.*, **76**, 3462–3466.
25. Qiao,Y., Wang,W. and Lu,X. (2020) Engineering cyanobacteria as cell factories for direct trehalose production from CO2. *Metab. Eng.*, **62**, 161–171.
26. Lin,P.-C., Zhang,F. and Pakrasi,H.B. (2020) Enhanced production of sucrose in the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Sci. Rep.*, **10**, 390.
27. Yunus,I.S., Wichmann,J., Wördenweber,R., Lauersen,K.J., Kruse,O. and Jones,P.R. (2018) Synthetic metabolic pathways for photobiological conversion of CO_2 into hydrocarbon fuel. *Metab. Eng.*, **41**, 201–211.

28. Atsumi,S., Higashide,W. and Liao,J.C. (2009) Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.*, **27**, 1177–1180.

29. Savakis,P., Tan,X., Du,W., Branco dos Santos,F., Lu,X. and Hellingwerf,K.J. (2015) Photosynthetic production of glycerol by a recombinant cyanobacterium. *J. Biotechnol.*, **195**, 46–51.

30. Zhou,J., Zhang,H., Zhang,Y., Li,Y. and Ma,Y. (2012) Designing and creating a modularized synthetic pathway in cyanobacterium *Synechocystis* enables production of acetone from carbon dioxide. *Metab. Eng.*, **14**, 394–400.

31. Lee,H.J., Choi,J., Lee,S.-M., Um,Y., Sim,S.J., Kim,Y. and Woo,H.M. (2017) Photosynthetic CO_2 conversion to fatty acid ethyl esters (FAEEs) using engineered cyanobacteria. *J. Agric. Food Chem.*, **65**, 1087–1092.

32. Takahama,K., Matsuoka,M., Nagahama,K. and Ogawa,T. (2003) Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the *psbAI* locus. *J. Biosci. Bioeng.*, **95**, 302–305.

33. Johnson,T.J., Jahandideh,A., Johnson,M.D., Fields,K.H., Richardson,J.W., Muthukumarappan,K., Cao,Y., Gu,Z., Halfmann,C., Zhou,R. *et al.* (2016) Producing next-generation biofuels from filamentous cyanobacteria: an economic feasibility analysis. *Algal Research*, **20**, 218–228.

34. Barsanti,L. and Gualtieri,P. (2018) Is exploitation of microalgae economically and energetically sustainable? *Algal Research*, **31**, 107–115.

35. Tredici,M.R., Rodolfi,L., Biondi,N., Bassi,N. and Sampietro,G. (2016) Techno-economic analysis of microalgal biomass production in a 1-ha Green Wall Panel (GWP®) plant. *Algal Research*, **19**, 253–263.

36. Chisti,Y. (2013) Constraints to commercialization of algal fuels. *J. Biotechnol.*, **167**, 201–214.

37. Angermayr,S.A., Gorchs Rovira,A. and Hellingwerf,K.J. (2015) Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol.*, **33**, 352–361.

38. Jones,P.R. (2014) Genetic instability in cyanobacteria - an elephant in the room? *Front Bioeng Biotechnol*, **2**, 12.

39. Ceroni,F., Algar,R., Stan,G.-B. and Ellis,T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods*, **12**, 415–418.

40. Zelcbuch,L., Antonovsky,N., Bar-Even,A., Levin-Karp,A., Barenholz,U., Dayagi,M., Liebermeister,W., Flamholz,A., Noor,E., Amram,S. *et al.* (2013) Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.*, **41**, e98.

41. Taylor,G.M., Mordaka,P.M. and Heap,J.T. (2018) Start-Stop Assembly: a functionally scarless DNA assembly system optimized for metabolic engineering. *Nucleic Acids Res.*, **47**, e17.

42. Bentley,F.K., Zurbriggen,A. and Melis,A. (2014) Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Mol. Plant*, **7**, 71–86.

43. Deng,M.D. and Coleman,J.R. (1999) Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ. Microbiol.*, **65**, 523–528.

44. Ranade,S., Zhang,Y., Kaplan,M., Majeed,W. and He,Q. (2015) Metabolic engineering and comparative performance studies of *Synechocystis* sp. PCC 6803 strains for effective utilization of xylose. *Front. Microbiol.*, **6**, 1484.

45. Chaves,J.E., Rueda-Romero,P., Kirst,H. and Melis,A. (2017) Engineering isoprene synthase expression and activity in cyanobacteria. *ACS Synth. Biol.*, **6**, 2281–2292.

46. Glick,B.R. (1995) Metabolic load and heterologous gene expression. *Biotechnol. Adv.*, **13**, 247–261.

47. Mutalik,V.K., Guimaraes,J.C., Cambray,G., Mai,Q.-A., Christoffersen,M.J., Martin,L., Yu,A., Lam,C., Rodriguez,C., Bennett,G. *et al.* (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods*, **10**, 347–353.

48. Awan,A.R., Blount,B.A., Bell,D.J., Shaw,W.M., Ho,J.C.H., McKiernan,R.M. and Ellis,T. (2017) Biosynthesis of the antibiotic nonribosomal peptide penicillin in baker's yeast. *Nat. Commun.*, **8**, 15202.

49. Smanski,M.J., Bhatia,S., Zhao,D., Park,Y., Woodruff,B.A.L., Giannoukos,G., Ciulla,D., Busby,M., Calderon,J., Nicol,R. *et al.* (2014) Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.*, **32**, 1241–1249.

50. Kelly,C.L., Taylor,G.M., Šatkutė,A., Dekker,L. and Heap,J.T. (2019) Transcriptional terminators allow Leak-Free chromosomal integration of genetic constructs in cyanobacteria. *Microorganisms*, **7**, 263.

51. Lessard,J.C. (2013) Chapter twenty seven - transformation of E. coli via electroporation. In: Lorsch,J. (ed). *Methods in Enzymology*. Academic Press, Vol. **529**, pp. 321–327.

52. Rippka,R. (1988) [1]Isolation and purification of cyanobacteria. In: *Methods in Enzymology*. Academic Press, Vol. **167**, pp. 3–27.

53. Casini,A., Storch,M., Baldwin,G.S. and Ellis,T. (2015) Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.*, **16**, 568–576.

54. Wang,B., Eckert,C., Maness,P.-C. and Yu,J. (2018) A genetic toolbox for modulating the expression of heterologous genes in the cyanobacterium *Synechocystis* sp. PCC 6803. *ACS Synth. Biol.*, **7**, 276–286.

55. Markley,A.L., Begemann,M.B., Clarke,R.E., Gordon,G.C. and Pfleger,B.F. (2015) Synthetic biology toolbox for controlling gene expression in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *ACS Synth. Biol.*, **4**, 595–603.

56. Jensen,P.R. and Hammer,K. (1998) The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.*, **64**, 82–87.

57. Pribnow,D. (1975) Bacteriophage T7 early promoters: nucleotide sequences of two RNA polymerase binding sites. *J. Mol. Biol.*, **99**, 419–443.

58. Rothmel,R.K. and LeClerc,J.E. (1989) Mutational analysis of the lac regulatory region: second-site changes that activate mutant promoters. *Nucleic Acids Res.*, **17**, 3909–3925.

59. Bakke,I., Berg,L., Aune,T.E.V. and Brautaset,T. (2009) Random mutagenesis of the Pm promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.*, **75**, 2002–2011.

60. Tornøe,J., Kusk,P., Johansen,T.E. and Jensen,P.R. (2002) Generation of a synthetic mammalian promoter library by modification of sequences spacing transcription factor binding sites. *Gene*, **297**, 21–32.

61. Mordaka,P.M. and Heap,J.T. (2018) Stringency of synthetic promoter sequences in clostridium revealed and circumvented by tuning promoter library mutation rates. *ACS Synth. Biol.*, **7**, 672–681.

62. Imamura,S. and Asayama,M. (2009) Sigma factors for cyanobacterial transcription. *Gene Regul. Syst. Bio.*, **3**, 65–87.

63. Mohamed,A., Eriksson,J., Osiewacz,H.D. and Jansson,C. (1993) Differential expression of the *psbA* genes in the cyanobacterium *Synechocystis* 6803. *Mol. Gen. Genet.*, **238**, 161–168.

64. Salis,H.M., Mirsky,E.A. and Voigt,C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, **27**, 946–950.

65. Oliver,J.W.K., Machado,I.M.P., Yoneda,H. and Atsumi,S. (2014) Combinatorial optimization of cyanobacterial 2, 3-butanediol production. *Metab. Eng.*, **22**, 76–82.

66. Niu,F.-X., Lu,Q., Bu,Y.-F. and Liu,J.-Z. (2017) Metabolic engineering for the microbial production of isoprenoids: carotenoids and isoprenoid-based biofuels. *Synthetic and Systems Biotechnology*, **2**, 167–175.

67. Coussement,P., Bauwens,D., Maertens,J. and De Mey,M. (2017) Direct combinatorial pathway optimization. *ACS Synth. Biol.*, **6**, 224–232.

68. Kim,S.W. and Keasling,J.D. (2001) Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnol. Bioeng.*, **72**, 408–415.

69. Matthews,P.D. and Wurtzel,E.T. (2000) Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. *Appl. Microbiol. Biotechnol.*, **53**, 396–400.

70. Gnanasekaran,T., Vavitsas,K., Andersen-Ranberg,J., Nielsen,A.Z., Olsen,C.E., Hamberger,B. and Jensen,P.E. (2015) Heterologous expression of the isopimaric acid pathway in *Nicotiana benthamiana* and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *J. Biol. Eng.*, **9**, 24.

71. Sandmann,G. (2002) Molecular evolution of carotenoid biosynthesis from bacteria to plants. *Physiol. Plant.*, **116**, 431–440.

72. Varman,A.M., Xiao,Y., Pakrasi,H.B. and Tang,Y.J. (2013) Metabolic engineering of *Synechocystis* sp. strain PCC 6803 for isobutanol production. *Appl. Environ. Microbiol.*, **79**, 908–914.

73. Angermayr,S.A., Paszota,M. and Hellingwerf,K.J. (2012) Engineering a cyanobacterial cell factory for production of lactic acid. *Appl. Environ. Microbiol.*, **78**, 7098–7106.

74. Jacobsen,J.H. and Frigaard,N.-U. (2014) Engineering of photosynthetic mannitol biosynthesis from CO_2 in a cyanobacterium. *Metab. Eng.*, **21**, 60–70.

75. Du,W., Angermayr,S.A., Jongbloets,J.A., Molenaar,D., Bachmann,H., Hellingwerf,K.J. and Branco Dos Santos,F. (2017) Nonhierarchical flux regulation exposes the fitness burden associated with lactate production in *Synechocystis* sp. PCC6803. *ACS Synth. Biol.*, **6**, 395–401.

76. Kusakabe,T., Tatsuke,T., Tsuruno,K., Hirokawa,Y., Atsumi,S., Liao,J.C. and Hanai,T. (2013) Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metab. Eng.*, **20**, 101–108.

77. Bartasun,P., Prandi,N., Storch,M., Aknin,Y., Bennett,M., Palma,A., Baldwin,G., Sakuragi,Y., Jones,P.R. and Rowland,J. (2019) The effect of modulating the quantity of enzymes in a model ethanol pathway on metabolic flux in *Synechocystis* sp. PCC 6803. *PeerJ*, **7**, e7529.

78. Lin,P.-C., Zhang,F. and Pakrasi,H.B. (2021) Enhanced limonene production in a fast-growing cyanobacterium through combinatorial metabolic engineering. *Metab Eng Commun*, **12**, e00164.

79. Pattanaik,B. and Lindberg,P. (2015) Terpenoids and their biosynthesis in cyanobacteria. *Life*, **5**, 269–293.

80. Lin,P.-C. and Pakrasi,H.B. (2019) Engineering cyanobacteria for production of terpenoids. *Planta*, **249**, 145–154.

81. Gao,X., Gao,F., Liu,D., Zhang,H., Nie,X. and Yang,C. (2016) Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energy Environ. Sci.*, **9**, 1400–1411.

82. Kelly,C.L., Taylor,G.M., Hitchcock,A., Torres-Méndez,A. and Heap,J.T. (2018) A Rhamnose-Inducible system for precise and temporal control of gene expression in cyanobacteria. *ACS Synth. Biol.*, **7**, 1056–1066.

83. Ma,J., Campbell,A. and Karlin,S. (2002) Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J. Bacteriol.*, **184**, 5733–5745.

84. Asplund-Samuelsson,J., Janasch,M. and Hudson,E.P. (2018) Thermodynamic analysis of computed pathways integrated into the metabolic networks of *E. coli* and *Synechocystis* reveals contrasting expansion potential. *Metab. Eng.*, **45**, 223–236.

85. Li,H., Shen,C.R., Huang,C.-H., Sung,L.-Y., Wu,M.-Y. and Hu,Y.-C. (2016) CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. *Metab. Eng.*, **38**, 293–302.

86. Shabestary,K. and Hudson,E.P. (2016) Computational metabolic engineering strategies for growth-coupled biofuel production by *Synechocystis*. *Metab. Eng. Commun.*, **3**, 216–226.

87. Du,W., Jongbloets,J.A., van Boxtel,C., Pineda Hernández,H., Lips,D., Oliver,B.G., Hellingwerf,K.J. and Branco dos Santos,F. (2018) Alignment of microbial fitness with engineered product formation: obligatory coupling between acetate production and photoautotrophic growth. *Biotechnol. Biofuels*, **11**, 38.