

ARTICLE

Received 14 Feb 2013 | Accepted 11 Sep 2013 | Published 17 Oct 2013

DOI: 10.1038/ncomms3595

Programming adaptive control to evolve increased metabolite production

Howard H. Chou^{1,2,3} & Jay D. Keasling^{1,2,3,4,5,6}

The complexity inherent in biological systems challenges efforts to rationally engineer novel phenotypes, especially those not amenable to high-throughput screens and selections. In nature, increased mutation rates generate diversity in a population that can lead to the evolution of new phenotypes. Here we construct an adaptive control system that increases the mutation rate in order to generate diversity in the population, and decreases the mutation rate as the concentration of a target metabolite increases. This system is called feedback-regulated evolution of phenotype (FREP), and is implemented with a sensor to gauge the concentration of a metabolite and an actuator to alter the mutation rate. To evolve certain novel traits that have no known natural sensors, we develop a framework to assemble synthetic transcription factors using metabolic enzymes and construct four different sensors that recognize isopentenyl diphosphate in bacteria and yeast. We verify FREP by evolving increased tyrosine and isoprenoid production.

¹UCSF-UCB Joint Graduate Group in Bioengineering, University of California, Berkeley, California 94720, USA. ²Joint BioEnergy Institute, Emeryville, California 94720, USA. ³Synthetic Biology Engineering Research Center, University of California, Berkeley, California 94720, USA. ⁴Department of Chemical & Engineering, University of California, Berkeley, California 94720, USA. ⁵California Institute for Quantitative Biosciences, University of California, Berkeley, California 94720, USA. ⁶Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. Correspondence and requests for materials should be addressed to J.D.K. (email: keasling@berkeley.edu).

Adaptation is a behaviour that allows cells to survive and thrive in constantly changing environmental conditions, and is characterized by rapid genetic change creating rare beneficial mutations¹. The appearance of microbial strains with accelerated mutation rates accompany periods of adaptation in both natural and laboratory environments^{2,3}, such as in the emergence of bacterial antibiotic resistance⁴. Models and experimental data of the adaptive process suggest that a ‘variable mutation rate’ strategy is one of the strategies used by nature to evolve traits, where a period of high mutation rate increases the genetic diversity of populations with low phenotypic diversity, and the mutation rate decreases with increased genetic diversity in the population^{5,6}.

Many mutagenesis strategies to generate diversity in the laboratory exist, but most industrially important phenotypes are not amenable to the high-throughput screens and selections required to isolate mutants exhibiting the desired traits⁷. Furthermore, directed evolution strategies that generate mutant libraries *in vitro* are limited by the ligation efficiency⁸, and those that use mutator strains with unregulated, high mutation rates to generate mutant libraries *in vivo*⁹ suffer from the accumulation of deleterious mutations that eventually lead to cell death. Although adaptation has proven useful for evolving certain phenotypes, its application has been limited to traits that are directly tied to growth^{10–12}. Therefore, a method capable of changing the mutation rate *in vivo* according to a particular phenotype, independent of whether it is linked to growth, could circumvent the constraints set by ligation inefficiencies, deleterious mutations, and assay availability.

Here we create such a method by implementing the ‘variable mutation rate’ strategy to evolve increased metabolite production using an adaptive control system we call feedback-regulated evolution of phenotype (FREP). We demonstrate the utility FREP by evolving increased tyrosine and isoprenoid production.

Results

Feedback-regulated evolution of phenotype. FREP consists of a two-module genetic circuit that dynamically controls the mutation rate of the genome (M) based on the concentration of a target metabolite (L) (Fig. 1a). The actuator module converts a transcriptional signal (T) into M , and the sensor module modifies T by converting L into a change in transcriptional signal (ΔT). M affects L over time as mutations that increase the metabolite production accumulate in the genome, creating a feedback loop that causes M to decrease as L increases (Fig. 1b). FREP can successfully modify L by changing M , because the changes to the system are manifested through transcription, which occurs at a much faster timescale than the propagation of those changes through growth and cell division. The sensor is assembled from two components: a transcription factor (TF) that binds the target metabolite and a promoter regulated by the TF. Depending on the metabolite, FREP could evolve a phenotype at either the population or single-cell level. If the metabolite is diffusible across the cell membrane and the rate of diffusion $\gg dL/dt$, then the effect of FREP is averaged across the entire population. However, if the metabolite is not diffusible across the cell membrane or its diffusion across the membrane is $\ll dL/dt$, then FREP acts on each

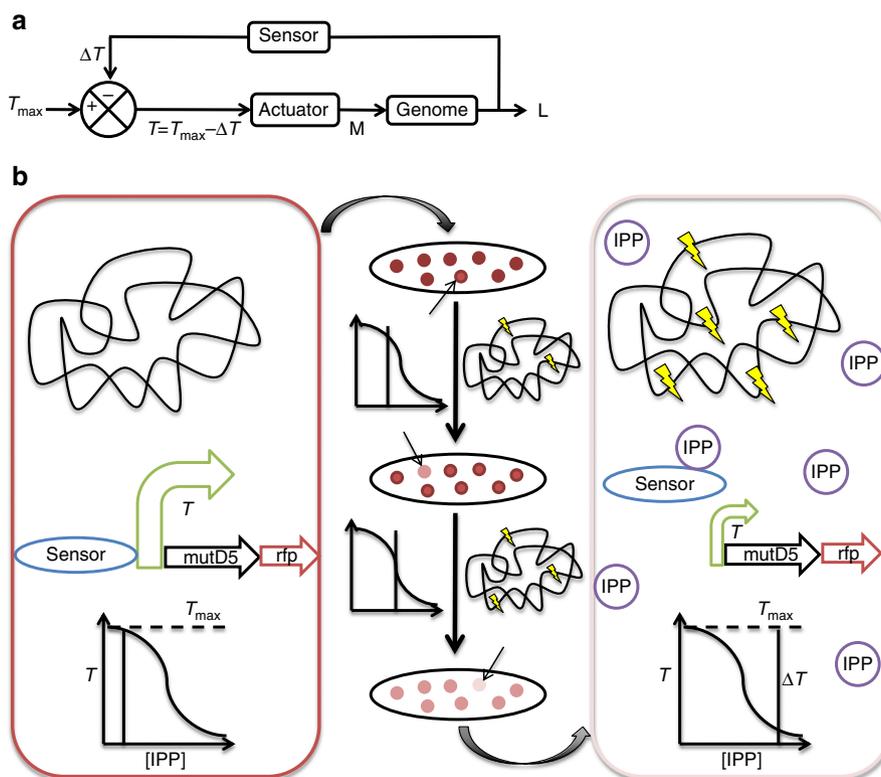


Figure 1 | FREP design. (a) FREP implementation of the variable mutation strategy using an adaptive control system. The sensor controls the change in transcriptional level (ΔT) in the system. The actuator converts the transcriptional level (T) into a mutation rate (M) that modifies the genome to produce the target phenotype gauged by L . As L increases, the sensor increases ΔT , which causes the actuator to decrease M . (b) Evolution of increased IPP production using FREP. The sensor activates the promoter that controls *mutD5* expression in the absence of the ligand IPP, and RFP acts as a reporter gene. MutD5 increases the mutation rate and mutations accumulate on the chromosome with each successive generation. Some of the mutations lead to increased IPP production, which decreases the ability of the sensor to activate expression of *mutD5* and *rfp* in those cells. The mutations that increase IPP production are fixed into successive generations by passing those cells with decreased *rfp* expression.

individual cell separately. Here we demonstrate the application of FREP to the evolution at either the population or single-cell level, and use tyrosine as a membrane permeable and isopentenyl diphosphate (IPP) as a membrane-impermeable metabolite.

FREP increases tyrosine production in *E. coli*. We performed FREP to increase production of the industrially important amino-acid tyrosine¹³ in *Escherichia coli* using the tyrosine-responsive TF TyrR¹⁴ to regulate expression of the mutator *mutD5* (ref. 15) (Fig. 2). The *mutD5* mutations are located in the *mutD* (*dnaQ*) gene that encodes the proofreading exonuclease of DNA polymerase III and leads to a reduction in mismatch repair. In this implementation, M should be high initially because the tyrosine concentration (L) is low, and M is reduced as beneficial mutations that increase tyrosine production appear. We modified TyrR and three TyrR-regulated promoters (P_{aroF} , P_{aroL} , P_{aroP}) to construct twenty different sensors, and screened their response to tyrosine in *E. coli* DJ106 and DJ166, two derivatives of BLR that produce different amounts of tyrosine. We monitored each sensor's output with the fluorescent protein mCherry (Supplementary Figs S1–S3). Sensor S_{aroF3} was the most sensitive to changes in tyrosine concentration, showing a 25% decrease in fluorescence from the lower to higher producing strain and a dynamic range of 0.44 RFU mM^{-1} per OD (Fig. 3a).

We tested FREP implemented with S_{aroF3} as the sensor and *mutD5* as the actuator in *E. coli* DJ238, expressing *mCherry* bicistronically with *mutD5* to monitor T and the relative mutator and metabolite levels in the cell. We reasoned that *mCherry* levels could decrease in response to either increased tyrosine production or mutations disrupting the sensor or *mCherry* expression. We isolated 10 colonies with the lowest fluorescence after 24 h and quantified tyrosine production to distinguish between the different scenarios. All 10 mutants demonstrated increased tyrosine production, and one exhibited greater than fivefold increase compared with the starting strain (Fig. 3b). Our observations indicate that raising M increased the diversity in the population to a level sufficient for mutants with increased L (tyrosine) to appear, and the sensor identified those mutants with increased metabolite production.

A synthetic TF for IPP in *E. coli*. To determine if FREP could evolve other traits, we implemented an adaptive control system to increase the production of isoprenoids, a class of compounds with a wide range of industrial applications, such as pharmaceuticals¹⁶ and biofuels¹⁷. Natural TFs for these compounds have not been discovered yet, so we developed a framework to rationally assemble synthetic TFs that could be used to regulate evolution

towards high isoprenoid-producing strains. Our strategy was to construct synthetic TFs reminiscent of natural TFs by taking advantage of their structural and functional modularity. The framework assembles synthetic TFs from three parts: Part1 is a metabolic enzyme that binds the target ligand, Part2 converts the binding signal into ΔT by regulating RNA polymerase binding to the target promoter, and Part3 joins Part1 and Part2 together (Fig. 4a).

For example, AraC regulates expression of arabinose utilization genes from the arabinose-inducible *araBAD* promoter (P_{BAD}) by preferentially binding different DNA sequences in the presence and absence of arabinose¹⁸. AraC has a distinct amino-terminal, ligand-binding domain (LBD) and carboxy-terminal, DNA-binding domain (DBD), and changes its ability to activate or repress P_{BAD} depending on whether the LBD has bound arabinose. We reasoned it should be possible to construct synthetic TFs for isoprenoids by replacing AraC's LBD with metabolic enzymes that naturally bind isoprenoids. We engineered a synthetic *E. coli* TF (chimeric protein IA, Fig. 4a) to respond to IPP, the central intermediate for all isoprenoid biosynthesis¹⁹, by fusing the AraC DBD (Part2) and linker

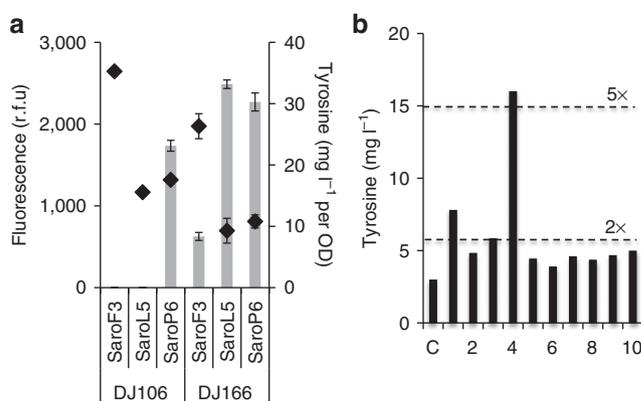


Figure 3 | FREP evolves increased tyrosine production. (a) Out of twenty sensors tested, the most sensitive sensors for each promoter (P_{aroF} , P_{aroL} , P_{aroP}) are compared for sensitivity to changes in tyrosine concentration *in vivo*. Bars represent tyrosine production and solid diamonds represent relative fluorescence units normalized to OD measured at 600 nm. DJ106 and DJ166 are variants of *E. coli* BLR, and DJ166 produces more tyrosine than DJ106. The error bars represent s.d. Each data point represents the average of three replicates. (b) Tyrosine production from 10 mutants evolved with FREP showing the lowest fluorescence after 24 h. C is the control not evolved with FREP.

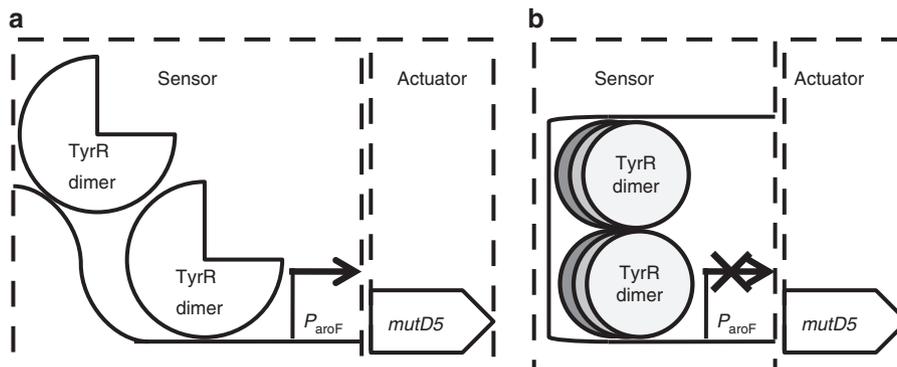


Figure 2 | FREP design to increase tyrosine production. (a) In one design, the sensor consists of TyrR and P_{aroF} , and the actuator consists of *mutD5*. TyrR dimers activate transcription from P_{aroF} in the absence of tyrosine. (b) Tyrosine-bound TyrR form hexamers that dimerize to repress transcription from P_{aroF} .

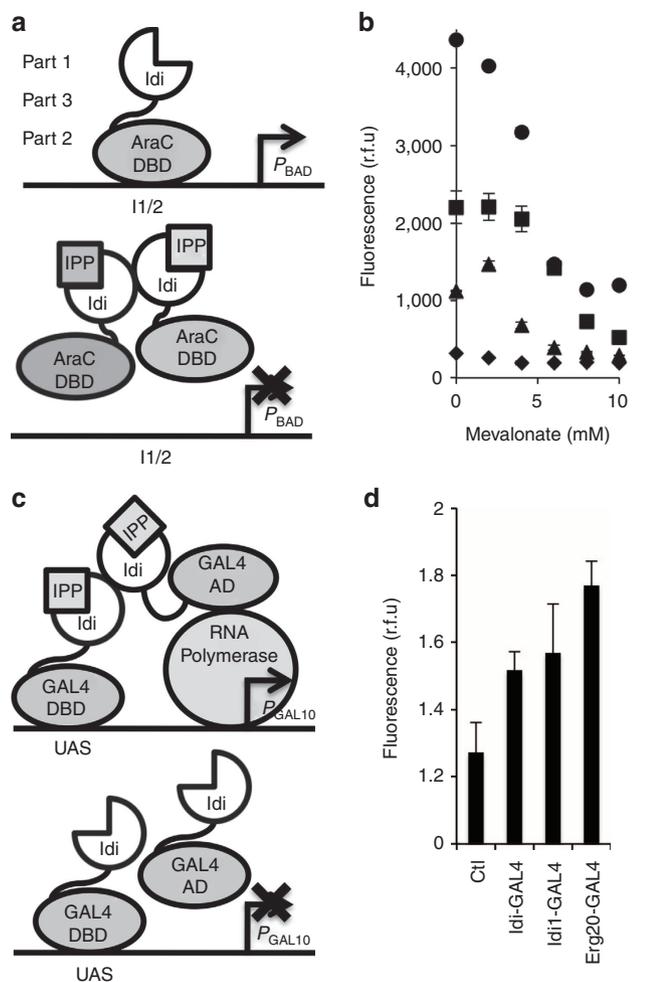


Figure 4 | Synthetic transcription factors respond to IPP. (a) A synthetic transcription factor (TF) consists of three parts: Part1 binds the target ligand, Part2 converts the binding signal into a change in RNA polymerase binding to the target promoter and Part3 is an amino-acid linker fusing Part1 and Part2 together. Here, a sensor with synthetic TF IA comprises Idi as Part1 and AraC's DBD and linker as Part2 and Part3, respectively. One model for how IA regulates P_{BAD} is IA binds the DNA sequence $I_{1/2}$, activating transcription from P_{BAD} in the absence of IPP (top), and IPP-bound IA dimerizes, preventing binding to $I_{1/2}$ and activation of P_{BAD} (bottom). (b) Output of four sensors, each with a different TF, to changing IPP concentrations in *E. coli* HC175 monitored with *mCherry*. Solid diamonds represent AC, solid triangles IA32, solid squares IA and solid circles IA44. The error bars represent s.d. Each data point represents the average of three replicates. (c) A sensor for detecting IPP in *S. cerevisiae*. The synthetic TF consists of Idi as Part1, GAL4's AD and DBD as Part2, and a 19-amino-acid linker as Part3. One model for P_{GAL10} regulation is that Idi dimerizes when bound to IPP, bringing the upstream activation sequence (UAS)-bound GAL4 DBD in close enough proximity with the GAL4 AD to activate transcription (top). In the absence of Part1 dimerization, there is no transcription from P_{GAL10} (bottom). (d) P_{GAL10} output from three sensors with synthetic TFs in *S. cerevisiae* MO219 induced with galactose. The synthetic TFs consist of Idi, Idi1 or Erg20 as Part1 fused to GAL4's AD and DBD. Ctl is the control without synthetic TFs. Output was monitored with the fluorescent protein yCitrine and normalized to fluorescence in the absence of galactose. The error bars represent s.d. Each data point represents the average of three replicates.

(Part3) with IPP isomerase (*idi²⁰*) (Part1). We chose Idi, because crystallographic data indicated that it dimerizes upon binding IPP²¹, suggesting that dimerization of Part1 could create at least

Table 1 | Fluorescence output from P_{BAD} and P_C . The promoters P_{BAD} and P_C were regulated by one of three TFs: AC, AraC, or IA. P_{BAD} was monitored using RFP and P_C with CFP.

	CFP		RFP	
	- Inducer	+ Inducer	- Inducer	+ Inducer
AC	2.8 ± 0.02	1.4 ± 0.2	0.19 ± 0.03	0.11 ± 0.02
AraC	1.5 ± 0.05	0.80 ± 0.03	0.22 ± 0.01	1.5 ± 0.002
IA	1.0 ± 0.06	0.64 ± 0.1	1.0 ± 0.07	0.28 ± 0.1

CFP, cyan fluorescent protein; RFP, red fluorescent protein. Fluorescence output was normalized to the output from the promoters regulated by IA in the absence of mevalonate (0 mM). Experiments were performed in HC175 induced with 0.1 mM IPTG. 10 mM arabinose was added in the case of '+ Inducer' for AraC, and 10 mM mevalonate was added in the cases of '+ Inducer' for AC and IA. Each data point represents the average of three experiments, and the s.d. values are provided.

two different conformational states for IA. We hypothesized that only one of the two states would activate transcription.

A sensor consisting of IA and P_{BAD} was tested by monitoring its output with *mCherry* in a modified strain of *E. coli* MG1655 able to convert mevalonate to IPP (HC175). A titration of mevalonate from 0–10 mM changed fluorescence by over threefold (Fig. 4b). There was no change in fluorescence when a synthetic TF consisting of only the AraC DBD and linker (AC) regulated P_{BAD} . We also evaluated expression from the divergent *araC* promoter (P_C) with *cfp* (Table 1). Combined with the P_{BAD} data, IA appears to regulate P_{BAD} and P_C nearly as tightly as AraC. Unlike AraC, IA represses P_{BAD} in the presence of ligand. Furthermore, both half-sites I_1 and I_2 upstream of P_{BAD} are necessary but interchangeable for IA regulation (Supplementary Table S1). These observations indicate IA can regulate T from P_{BAD} based on the concentration of L (IPP) with a fluorescent output of 518–2,202 r.f.u. per OD across a concentration range from 0–8 mM to provide a dynamic range of $210 \text{ r.f.u. mM}^{-1}$ per OD, assuming all of the mevalonate was converted to IPP.

We purified IA to confirm it binds the I_1 and I_2 half-sites adjacent to P_{BAD} *in vitro*. Gel electrophoresis mobility shift assay²² experiments showed two bands when I_1 and I_1I_2 were substrates and three bands when the substrate was the DNA sequence from P_C to P_{BAD} (Supplementary Fig. S4). The additional band supports the observation that IA regulates both P_{BAD} and P_C , which have distinct binding sequences. The shifted DNA bands were less intense when IPP was added, indicating that IA's affinity for the binding sequences decreases in the presence of IPP. We confirmed that IPP modulates IA DNA binding using fluorescence resonance energy transfer by splitting I_1 and I_1I_2 into two DNA fragments each constituting half of the original sequence and tagged with either a fluorophore or quencher²³. Only the presence of IA and both half-sequences induced a change in fluorescence (Supplementary Fig. S5). Adding IPP decreased the change in fluorescence across all concentrations of IA tested (Supplementary Fig. S6). Thus, both *in vivo* and *in vitro* data are consistent with IA regulation of transcription from P_{BAD} according to changing IPP concentrations, and both I_1 and I_2 half-sites are necessary for this regulation.

A synthetic TF for IPP in yeast. To further evaluate our framework for assembling synthetic TFs, we constructed a synthetic TF for isoprenoids in *Saccharomyces cerevisiae* using the GAL4 protein, which regulates expression of GAL genes in response to galactose²⁴. Similar to AraC, the functional domains of GAL4 are structurally distinct, consisting of an activator domain (AD)

and DBD²⁵. We reused Idi as Part1 and fused it to the GAL4 AD and DBD (Part2), reasoning that Idi dimerization should bring the AD and DBD in close enough proximity to activate transcription from a GAL promoter (for example, P_{GAL10}). Part3 was a 19 amino-acid sequence having relatively high stability²⁶. This sensor (Fig. 4c) was tested by monitoring its output with the fluorescent protein γ Ecitrine in *S. cerevisiae* MO219, a genetically modified strain that increases isoprenoid production when induced with galactose²⁷. We observed a change in fluorescence greater than baseline after galactose induction (Fig. 4d). Two additional yeast TFs were constructed from yeast enzymes known to catalyse reactions with IPP as a substrate (Idi1 (ref. 28) and Erg20 (ref. 29)) as Part1 in place of Idi, and both showed even greater changes in fluorescence following induction. Induction led to an almost twofold increase in sensor output in response to increased isoprenoid levels using the synthetic TF constructed with Erg20. Combined with IA, these GAL4-based TFs highlight the modularity of our framework in assembling synthetic TFs for constructing sensors, alleviating the need to rely on pre-existing biological components.

FREP increases isoprenoid production in *E. coli*. Next, we modified the *E. coli* IPP TF IA using error-prone PCR to create IPP sensors with different dynamic ranges and maximum transcription levels (T_{max}). Out of the 60 variants screened (Supplementary Fig. S7), IA32 (L39M, S127C) showed half the T_{max} of IA and a dynamic range of 145 r.f.u.mM⁻¹ per OD, whereas IA44 (R267H) showed twice the T_{max} of IA and a dynamic range of 350 r.f.u./mM/OD (Fig. 3c). We implemented FREP using one of three synthetic TFs (AraC, IA32 or IA44) as part of the sensor and the *mutD5* actuator, and examined these constructs in *E. coli* MG1655 using Luria–Delbruck fluctuation analysis³⁰. The fluctuation analysis data was analysed using the fluctuation analysis calculator web tool in order to calculate the mutation frequency and rate from a mutation assay that is based on the frequency of rifampicin resistance in the population. Thirty colonies for each implementation were tested for rifampicin resistance, an orthogonal phenotype that could be quantified quickly. In general, we observed more rifampicin-resistant mutants with higher mutator expression, and a strong

correlation between relative mutator expression and mutation rate ($r=0.97$) (Supplementary Fig. S8). For example, IA32 and IA44 exhibited a fourfold difference in T_{max} and a 2.4-fold difference in M . A negative control consisting of a sensor with IA44 and no actuator generated no rifampicin-resistant mutants. These results show that increasing ΔT decreases M , consistent with our design, and suggest that dynamically controlling mutator expression changes the mutation rate in the cell. Furthermore, the ability to adjust the dynamic range of the TF allows the range of metabolite concentration FREP can sense to be controlled.

We performed FREP with IA44 to increase isoprenoid production in *E. coli* MG1655, and expressed *mCherry* bicistronically with the actuator to monitor changes in *in vivo* IPP concentration. Ten colonies with the lowest fluorescence after 24 h were made electrocompetent and transformed with a plasmid containing the lycopene synthase genes (pLyc). Lycopene measured from a random transformant for all 10 colonies was higher than the control without the actuator module. Six colonies had mutants producing on average 2,900 μg lycopene g⁻¹ dry cell weight (p.p.m.), a nearly threefold increase compared with the control that did not undergo FREP, which produced only 1,000 p.p.m. (Supplementary Fig. S9). Repeating the experiment with a sensor employing AraC induced with arabinose as a negative control (AraC does not respond to IPP) generated no mutants producing more lycopene than the initial strain, illustrating the importance of the feedback loop and sensor to select mutants exhibiting increased metabolite production.

Finally, we examined the ability of FREP to generate increased metabolite production in the context of a long-term experiment. We co-transformed pLyc with an IPP sensor and *mutD5* actuator into *E. coli* MG1655, and monitored the evolution of IPP production using lycopene as a reporter over 432 h. We quantified lycopene production every 72 h from 10 random colonies and only passaged the isolate demonstrating the highest production levels. After 432 h, lycopene production increased to 6,800 p.p.m. using IA44, 4,700 p.p.m. using IA32, and only 400 p.p.m. using AraC (Fig. 5a). A negative FREP control implemented with IA44 without an actuator produced 0 p.p.m. For the strains evolved using FREP implemented with IA44 and an actuator, we purified pLyc from cells isolated from each time point. Transforming those plasmids into *E. coli* MG1655 did not

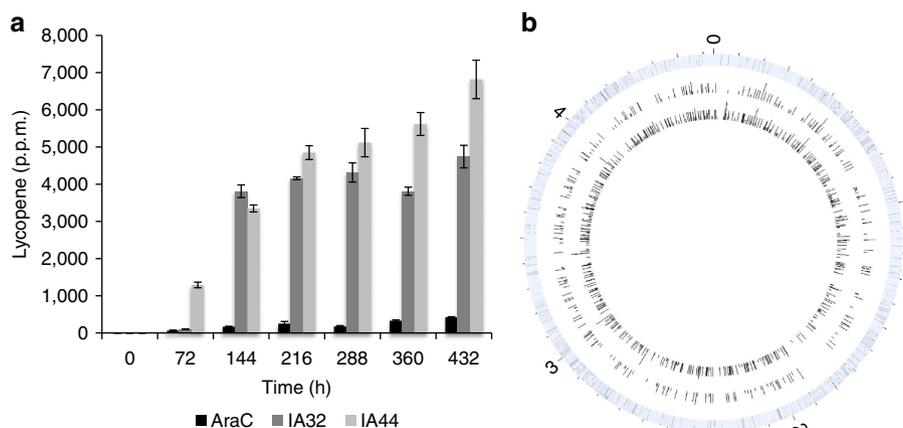


Figure 5 | FREP evolves increased IPP production. (a) Lycopene production sampled every 72 h over a 432-h period from *E. coli* MG1655 expressing pLyc, an IPP sensor and an actuator. The transcription factors used in the sensor were AraC (black bars), IA32 (dark grey bars) or IA44 (light grey bars). Lycopene production is presented as p.p.m. ($\mu\text{g g}^{-1}$ dry cell weight). The sensor with AraC was induced with 10 mM arabinose. The error bars represent s.d. Each data point represents the average of three replicates. (b) The strains isolated at each 72-h timepoint for lycopene analysis were resequenced. The outer blue ring represents the *E. coli* genome, the middle blue ring represents the location of the nonsynonymous SNPs and the inner ring shows all of the SNPs. The height of the bars in the inner two rings illustrates the relative occurrence of each SNP, where a taller bar indicates that the SNP was identified in more of the strains isolated.

lead to more lycopene production compared with the original plasmid (Supplementary Fig. S10). This observation indicates that FREP generated mutations on the chromosome that are necessary for increased IPP production. No difference in growth rate was observed between the control and mutant strains. The absence of any changes in growth rate might be a result of our selection method. Cells were grown on plates prior to selection, and selection for high producers was done based on the intensity of the red colour of the colonies. Therefore, mutants with slower growth rates or slower production rates were selected against, and the mutants we propagated and analysed were those that grew and produced lycopene the fastest. It is possible that some of the slower growers might be higher producers, and we missed these mutants based on our selection criteria.

We re-sequenced the strains isolated for high lycopene production as well as their parent strain. The parent strain matched the genome sequence of wild-type *E. coli* K12 MG1655 (GenBank: NC000913). The total number of SNPs and the number of non-synonymous SNPs increased with the amount of time FREP progressed until the final time point, where the number of SNPs remained the same and non-synonymous SNPs decreased. Between 0 and 360 h, the number of non-synonymous SNPs increased by 87, 250, 84, 4%, respectively. The decrease in the rate of accumulation of SNPs after 216 h confirms that the mutation rate decreased as the amount of IPP increased. The majority of the non-synonymous mutations is in the coding regions of the genome and is evenly distributed across the entire genome (Fig. 5b).

Mutations in *rrfD*, *rrfF*, *rrfG* and *ppiC* were maintained in all of the strains collected throughout the 432 h of evolution. The *rrf* genes encode the 50S subunit of the ribosome, and mutations in these genes are in accordance with the previous efforts to modify the *E. coli* and *S. cerevisiae* transcription machineries^{31,32}. These mutations suggest that modification to global mechanisms that affect the expression of multiple genes is an efficient mechanism to evolve new phenotypes. Related to the mechanism of translation, mutations were also found in *infB*, which is involved in translation initiation, *rhlB* and *rne*, which are involved in RNA catabolism. It is possible that some of the mutations accumulated might counter the deleterious effects of *mutD5*. Out of all of the genes that contain SNPs after undergoing FREP, only one gene, *aceE*, was previously identified to affect lycopene production³³. These results illustrate the difficulty in predicting all of the genes involved in producing a particular phenotype, and how FREP enables a more complete exploration of the genomic search space to evolve a particular phenotype. We successfully designed and implemented an adaptive control process capable of implementing a variable mutation strategy and selecting for mutants that demonstrate increased production of a target metabolite. FREP is different from simulation and omics-based strategies, because it does not rely on preexisting knowledge about the genes, RNA, proteins and their interactions that govern the trait being engineered to decide what changes to make^{33,34}. FREP is unique from other adaptive and metabolic evolution approaches^{10–12}, because it monitors the phenotype being engineered using a sensor that can be linked to a fluorescent readout, so the phenotype does not need to be linked to growth. Furthermore, FREP hijacks the cell's native mechanism for generating mutations and can be implemented in the absence of sophisticated recombination methods and automation equipment.

Discussion

Engineering the production of a particular metabolite is a complex, multi-dimensional problem with an extremely large

but discrete search space that must account for many variables. This makes the global optimum extremely difficult to identify. Therefore, most methods to engineer microbial metabolism depend on current knowledge of metabolism and focus on engineering a very limited set of genes and proteins known to be associated with the production of the metabolite of interest. FREP takes a novel approach that resembles the simulated annealing method of optimization based on the Metropolis Hastings algorithm for optimization of nonlinear systems. The large search space associated with optimizing metabolism makes direct sampling extremely difficult due to the practical limitations in generating that search space. Therefore, FREP generates a random library that attempts to represent that search space and tries to identify a good approximation of the global optimum in a fixed amount of time by changing the mutation rate. When the mutation rate is high, the probability that a new genome is fixed into the population is low. However, the mutation rate decreases as the concentration of the target metabolite increases, and the probability that a new genome is fixed into the population increases as the mutation rate decreases. The allowance for the mutation rate to increase rather than always decrease prevents FREP from getting trapped in a local optimum.

We demonstrated the application of FREP by evolving *E. coli* to increase tyrosine and IPP production, and isolating the evolved strains by monitoring the actuator level with a fluorescent protein. We confirmed that FREP was able to evolve increased production of metabolites that are either permeable or impermeable to the cell membrane. Metabolites that are membrane impermeable appear to be easier to evolve using FREP than those that are permeable, as exemplified by the fact that two out of ten tyrosine mutants and six out of ten IPP mutants showed a more than twofold increase in production. Once the level of a target trait saturates using a particular TF, its dynamic range could be altered to enable further increases in the level of the trait using FREP. Although the final population evolved using FREP could be heterogeneous with slight variations in production, the same sensor could be used to identify those individuals with the highest production. For example, by transforming the population of mutants with a construct that contains the sensor and not the mutator, the concentration of the metabolite can be translated into a fluorescence signal that can be sorted using FACS in order to identify the highest producing individual. Nevertheless, selecting the mutants using the sensor would still be superior to sorting them using FACS. Additionally, we presented a framework to rationally construct synthetic TFs using metabolic enzymes that enable the development of orthogonal sensors less likely to interact with existing cellular networks without being limited to the molecular recognition properties and control functions of naturally occurring TFs.

Methods

Oligonucleotides and DNA sequencing. All oligonucleotides were obtained from Integrated DNA Technologies and are presented in Supplementary Table S2.

The function of each pair of oligonucleotides used to construct the plasmids in this study is described in Supplementary Table S3. DNA sequencing was performed by Quintara Biosciences.

Strains and plasmids availability. Strains, plasmids and plasmid sequences (in Genbank format) are deposited in the private instance of the JBEI registry and will be moved to the public instance (<http://www.public-registry.jbei.org>) after publication. Strains and plasmids are available from Addgene (<http://www.addgene.org>), and listed in Supplementary Table S4.

Strains. We cloned the kanamycin cassette from pKD4 into pMevB³⁵ to construct pMevB-Kan using the primers Kan-F and Kan-R. We engineered EHC175 by amplifying *mk*, *pmk* and *pmd* of the mevalonate operon with the kanamycin cassette from pMevB-Kan with the primers IdIKO-F and IdIKO-R, and knocking out *idi* in *E. coli* MG1655 with the PCR product according to Datsenko &

Wanner³⁶. *E. coli* DJ106, DJ166 and DJ238 were gifts from Dr. Darmawi Juminaga. *S. cerevisiae* MO219 was a gift from Dr. Mario Ouellet. Genes and promoter sequences amplified from the *E. coli* chromosome were from MG1655. Genes amplified from the *S. cerevisiae* chromosome were from BY4742.

Construction of pLyc. We cut *crte*, *crtI* and *crtB* from pT-LYCM4 (gift from Dr. Adrienne McKee) using the restriction enzymes SpeI and HindIII. The cut fragment was gel purified and ligated into pBAD18-Cm using T4 DNA ligase (Fermentas) in order to construct pLyc. The ligation product was transformed into *E. coli* DH10B using electroporation.

Construction of plasmids containing *E. coli* IPP sensors. pCtI-RFP-*S*_{AraC} (*S*_{AraC}: sensor containing AraC) was constructed by removing HindIII from pBAD24 using QuickChange PCR, cloning in the DNA sequence from *araC* from pBAD24M-gfp³⁷ using ClaI and EcoRI, cloning in *mutD* amplified from *E. coli*, and cloning in *mCherry* 3' of the *araBAD* promoter, *P*_{BAD}.

The chimeric protein IA was constructed by fusing *idi* to the C-terminus of *araC* using SOEing PCR. *idi* was amplified from *E. coli*, and the linker and C-terminus of *araC* were amplified from pBAD24. These two PCR products were templates for SOEing PCR to amplify the chimeric protein. We cloned IA into pCtI-RFP-*S*_{AraC} by replacing AraC to make pCtI-RFP-*S*_{IA}.

Mutants of IA were generated using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions. We cloned the IA mutants into pCtI-RFP-*S*_{IA}, transformed the constructs into EchC175, and screened for changes in RFP expression relative to IA in the presence (10 mM) and absence (0 mM) of mevalonate with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich). RFP was measured using a Spectramax M2 (Molecular Devices) exciting at 587 nm and measuring emission at 610 nm. We isolated two mutants: IA32 and IA44. pCtI-RFP-*S*_{AC} was constructed by amplifying the C-terminal domain of AraC, and cloning into pCtI-RFP-*S*_{AraC}. pCtI-RFP-*S*_{IA44} was digested with ClaI and KpnI, and the fragment containing IA44 to mutD was cloned into pBAD24 to construct pCtI-*S*_{IA44}.

Construction of plasmids containing *mutD5* mutator module. The mutator *mutD5* was a gift from Dr. Adrienne McKee and cloned into pCtI-*S*_{IA44}, pCtI-RFP-*S*_{AraC}, pCtI-RFP-*S*_{IA44}, and pCtI-RFP-*S*_{aroF3} using the primers MutD-F and MutD-R to make pMut-*S*_{IA44}, pMut-RFP-*S*_{AraC}, pMut-RFP-*S*_{IA44}, and pMut-RFP-*S*_{aroF3}. pMut-*S*_{AC}, pMut-*S*_{AraC}, pMut-*S*_{IA32} were constructed by cutting the TF from pCtI-RFP-*S*_{AC}, pCtI-RFP-*S*_{AraC}, pCtI-RFP-*S*_{IA32} using ClaI and HindIII, and cloning the fragments into pMut-*S*_{IA44}.

Characterization of *E. coli* IPP sensor modules. We measured expression of RFP from *P*_{BAD} controlled by one of the TFs (AraC, AC, IA, IA32 or IA44) by transforming pCtI-RFP-*S* (*S* designates a sensor with one of the TFs) into EchC175 and plating on LB agar plates with ampicillin and kanamycin. We picked three clones from each plate, grew each clone in LB medium with antibiotics overnight and inoculated each culture into fresh EZ Rich Defined Medium (Teknova) with antibiotics to an initial Abs₆₀₀ of 0.05 the following day. Each fresh culture was grown for 3 h at 37 °C, induced with isopropyl β-D-1-thiogalactopyranoside (0.1 mM) and mevalonate (0–10 mM) (or 0–10 mM arabinose for AraC), and grown for an additional 17 h at 37 °C. We measured RFP fluorescence exciting at 495 nm and measuring emission at 520 nm, and Abs₆₀₀ using a Spectramax M2.

To determine IA's binding sequence upstream of *P*_{BAD}, we amplified I₁I₁ using the primer IIII-F, I₂I₁ using the primer I2I1-F or I₂I₂ using the primer I2I2-F, all paired with AraReg-R and using pCtI-RFP-*S*_{IA} as template. The PCR products were cloned into pCtI-RFP-*S*_{IA} to replace the I₁I₂ sequence to make pCtI-RFP-*S*_{IA}-I₁I₁, pCtI-RFP-*S*_{IA}-I₂I₁ and pCtI-RFP-*S*_{IA}-I₂I₂. RFP expression from the modified binding sequences was determined as described above.

We amplified CFP using the primers CFP-F and CFP-R, and inserted it 3' of the TF expressed 3' of *P*_C into pCtI-RFP-*S*_{AC}, pCtI-RFP-*S*_{AraC}, and pCtI-RFP-*S*_{IA} to make pCtI-CFP-RFP-*S*_{AC}, pCtI-CFP-RFP-*S*_{AraC}, and pCtI-CFP-RFP-*S*_{IA}. RFP and CFP expression from these constructs were determined as indicated above, and CFP fluorescence was measured using a Spectramax M2 exciting at 433 nm and measuring emission at 475 nm.

Construction of plasmids containing tyrosine sensor modules. We replaced *P*_C with CP20 (ref. 38) in pCtI-RFP-*S*_{AraC} and cloned in *tyrR* amplified from *E. coli* to construct pCtI-RFP-TyrR.

pCtI-RFP-*S*_{aroF0} was constructed by amplifying the promoter region of *aroF* from *E. coli* and cloning the PCR product into pCtI-RFP-TyrR to replace *P*_{BAD}. pCtI-RFP-*S*_{aroF1} was constructed by mutating TyrR to make TyrR E274Q³⁹ and cloning the PCR product into pCtI-RFP-*S*_{aroF0}. pCtI-RFP-*S*_{aroF2} was constructed by mutating TyrR to make TyrR N316K⁴⁰, and cloning the PCR product into pCtI-RFP-*S*_{aroF0}. pCtI-RFP-*S*_{aroF3} was constructed by incorporating the N316K mutation into TyrR E274Q and cloning the PCR product into pCtI-RFP-*S*_{aroF0}. The N-terminus of TyrR was also truncated to different lengths to generate TyrR Δ43, TyrR Δ93, and TyrR Δ187 (ref. 41) to make pCtI-RFP-*S*_{aroF4}, pCtI-RFP-*S*_{aroF5} and pCtI-RFP-*S*_{aroF6}.

pCtI-RFP-*S*_{aroL0} was constructed by replacing *P*_{BAD} with the promoter region of *aroL* from *E. coli* and cloning the PCR product into pCtI-RFP-TyrR. TyrR from pCtI-RFP-*S*_{aroF1}, pCtI-RFP-*S*_{aroF2} and pCtI-RFP-*S*_{aroF3} were amplified and cloned into pCtI-RFP-*S*_{aroL0} to construct pCtI-RFP-*S*_{aroL1}, pCtI-RFP-*S*_{aroL2} and pCtI-RFP-*S*_{aroL3}, respectively. The TyrR boxes 1, 2, 3 of the promoter *P*_{aroL} were also modified to tune TyrR regulation of the promoter⁴². pCtI-RFP-*S*_{aroL4} was constructed by modifying the sequences of box 1 and 2 of *P*_{aroL} in pCtI-RFP-*S*_{aroL0}. pCtI-RFP-*S*_{aroL5} was constructed by modifying the box 3 sequence of *P*_{aroL} in pCtI-RFP-*S*_{aroL0}.

pCtI-RFP-*S*_{aroP0} was constructed by replacing *P*_{BAD} with the promoter region of *aroP* from *E. coli* and cloning the PCR product into pCtI-RFP-TyrR. TyrR from pCtI-RFP-*S*_{aroF1}, pCtI-RFP-*S*_{aroF2} and pCtI-RFP-*S*_{aroF3} were amplified and cloned into pCtI-RFP-*S*_{aroP0} to construct pCtI-RFP-*S*_{aroP1}, pCtI-RFP-*S*_{aroP2} and pCtI-RFP-*S*_{aroP3}, respectively. pCtI-RFP-*S*_{aroP4} was constructed by modifying the *P*₂ sequence of *P*_{aroP} to make *P*_{2up}⁴³ in pCtI-RFP-*S*_{aroP0}. pCtI-RFP-*S*_{aroP5} and pCtI-RFP-*S*_{aroP6} were constructed by amplifying TyrRΔ43 and TyrRΔ93 from pCtI-RFP-*S*_{aroF4} and pCtI-RFP-*S*_{aroF5}, respectively and cloning the PCR products into pCtI-RFP-*S*_{aroP4}.

Characterization of tyrosine sensor modules. Plasmids containing each of the 20 tyrosine sensors described above were transformed into *E. coli* DJ106 and DJ166, and plated on LB agar with ampicillin. Clones were grown overnight in LB medium with ampicillin, inoculated into EZ Rich Defined Medium the next day to an initial Abs₆₀₀ of 0.05, and tyrosine production was quantified after 20 h. RFP fluorescence and Abs₆₀₀ were measured as described earlier. The experiment was repeated in triplicate for *S*_{aroF3}, *S*_{aroL5} and *S*_{aroP6}.

Yeast synthetic TFs and IPP sensor modules. The TEF promoter was amplified and cloned into pESC-Ura to make pESC-*P*_{TEF}. yEcitrine was amplified and cloned into pESC-*P*_{TEF} behind *P*_{gal10} to make pESC-YFP-*P*_{TEF}. The *cycl* terminator and TEF promoter were fused using SOEing PCR to make the PCR product *P*_{TEF2}. The activator (AD) and DNA-binding domains (DBD) of *gal4* were amplified from *S. cerevisiae* and the PCR products were fused 5' of *E. coli* *idi*, *S. cerevisiae* *idi1* or *S. cerevisiae* *erg20* using SOEing PCR. Each AD/DBD pair of fusion constructs was cloned into pESC-YFP-*P*_{TEF} behind *P*_{TEF} and separated by *P*_{TEF2} to make pESC-YFP-*S*_{Idi-GAL4}, pESC-YFP-*S*_{Idi1-GAL4}, and pESC-YFP-*S*_{Erg20-GAL4}.

Characterization of yeast IPP sensor modules. pESC-YFP-*P*_{TEF}, pESC-YFP-*S*_{Idi-GAL4}, pESC-YFP-*S*_{Idi1-GAL4} and pESC-YFP-*S*_{Erg20-GAL4} were transformed into SMO219 using electroporation, plated on synthetic defined (s.d.) agar without uracil and with 2% glucose, and grown at 30 °C for 3 days. s.d. medium was composed of 1 × CSM without the appropriate amino acids (Sunrise Science Products) and 1 × Difco Yeast Nitrogen Base without amino acids (BD), prepared according to the manufacturers' instructions. Three clones from each plate were grown overnight in s.d. medium without uracil and with 2% glucose, inoculated into fresh medium without uracil and with 1.8% galactose and 0.2% glucose the following day to an initial Abs₆₀₀ of 0.05, and grown for 3 days at 30 °C. YFP fluorescence was measured using a Spectramax M2 (Molecular Devices) exciting at 516 nm and measuring emission at 529 nm and normalized to OD measured at 600 nm.

References

- Elena, S. F., Cooper, V. S. & Lenski, R. E. Punctuated evolution caused by selection of rare beneficial mutations. *Science* **272**, 1802–1804 (1996).
- Desai, M. M. & Fisher, D. S. The balance between mutator and nonmutators in asexual populations. *Genetics* **188**, 997–1014 (2011).
- Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**, 703–705 (1997).
- Zhang, Q. *et al.* Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science* **333**, 1764–1767 (2011).
- Stich, M., Manrubia, S. C. & Lázaro, E. Variable mutation rates as an adaptive strategy in replicator populations. *PLoS One* **5**, e11186 (2010).
- Giraud, A. *et al.* Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* **291**, 2606–2608 (2001).
- Dietrich, J. A., McKee, A. E. & Keasling, J. D. High-throughput metabolic engineering: advances in small-molecule screening and selection. *Ann. Rev. Biochem.* **79**, 563–590 (2010).
- Hibert, E. G. *et al.* Directed evolution of biocatalytic processes. *Biomol. Eng.* **22**, 11–19 (2005).
- Greener, A., Callahan, M. & Jerseth, B. An efficient random mutagenesis technique using an *E. coli* mutator strain. *Mol. Biotechnol.* **7**, 188–195 (1997).
- Herring, C. D. *et al.* Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat. Genet.* **38**, 1406–1412 (2006).
- Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
- Portnoy, V. A., Bezdán, D. & Zengler, K. Adaptive laboratory evolution – harnessing the power of biology for metabolic engineering. *Curr. Opin. Biotechnol.* **22**, 590–594 (2011).

13. Juminaga, D. *et al.* Modular engineering of *L*-tyrosine production in *Escherichia coli*. *Appl. Env. Microbiol.* **78**, 89–98 (2012).
14. Pittard, J., Camakaris, H. & Yang, J. The TyrR regulon. *Mol. Microbiol.* **55**, 16–26 (2005).
15. Schaaper, R. M. Mechanisms of mutagenesis in the *Escherichia coli* mutator *mutD5*: role of DNA mismatch repair. *Proc. Natl Acad. Sci. USA* **85**, 8126–8130 (1988).
16. Chang, M. C. Y. & Keasling, J. D. Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* **2**, 674–681 (2006).
17. Keasling, J. D. & Chou, H. Metabolic engineering delivers next-generation biofuels. *Nat. Biotechnol.* **26**, 298–299 (2008).
18. Soisson, S. M., MacDougall-Shackleton, B., Schleif, R. & Wolberger, C. Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421–425 (1997).
19. Lange, B. M., Rujan, T., Martin, W. & Croteau, R. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl Acad. Sci. USA* **97**, 13172–13177 (2000).
20. Hahn, F. M., Hurlburt, A. P. & Poulter, C. D. *Escherichia coli* open reading frame 696 is *idi*, a nonessential gene encoding isopentenyl diphosphate isomerase. *J. Bacteriol.* **181**, 4499–4504 (1999).
21. De Ruyck, J., Oudjama, Y. & Wouters, J. Monoclinic form of isopentenyl diphosphate isomerase: a case of polymorphism in biomolecular crystals. *Acta. Cryst.* **F64**, 239–242 (2008).
22. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Prot.* **2**, 1849–1861 (2007).
23. Heyduk, T. & Heyduk, E. Molecular beacons for detecting DNA binding proteins. *Nat. Biotechnol.* **20**, 171–176 (2002).
24. Traven, A., Jelicic, B. & Sopta, M. Yeast Gal4: a transcriptional paradigm revisited. *EMBO* **7**, 496–499 (2006).
25. Fields, S. & Song, O.-k. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246 (1989).
26. Robinson, C. R. & Sauer, R. T. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl Acad. Sci. USA* **95**, 5929–5934 (1998).
27. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
28. Mayer, M. P., Hahn, F. M., Stillman, D. J. & Poulter, C. D. Disruption and mapping of *IDI*, the gene for isopentenyl diphosphate isomerase in *Saccharomyces cerevisiae*. *Yeast* **8**, 743–748 (1992).
29. Fischer, M. J., Meyer, S., Claudel, P., Bergdoll, M. & Karst, F. Metabolic engineering of monoterpene synthesis in yeast. *Biotechnol. Bioeng.* **108**, 1883–1892 (2011).
30. Rosche, W. A. & Foster, P. L. Determining mutation rates in bacterial populations. *Methods* **20**, 4–17 (2000).
31. Alper, H., Moxley, J., Nevoigt, E., Fink, G. R. & Stephanopoulos, G. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**, 1565–1568 (2006).
32. Alper, H. & Stephanopoulos, G. Global transcription machinery engineering: A new approach for improving cellular phenotype. *Met. Eng.* **9**, 258–267 (2007).
33. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat. Biotechnol.* **23**, 612–616 (2005).
34. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–899 (2009).
35. Martin, V. J. *et al.* Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **7**, 796–802 (2003).
36. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
37. Lee, S. K. *et al.* Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl. Env. Microbiol.* **73**, 5711–5715 (2007).
38. Jensen, P. D. & Hammer, K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Enl. Microbiol.* **64**, 82–87 (1998).
39. Kwok, T., Yang, J., Pittard, A. J., Wilson, T. J. & Davidson, B. E. Analysis of an *Escherichia coli* mutant TyrR protein with impaired capacity for tyrosine mediated repression, but still able to activate at sigma 70 promoters. *Mol. Micro.* **17**, 471–481 (1995).
40. Koyanagi, T., Katayama, T., Suzuki, H. & Kumagai, H. Altered oligomerization properties of N316 mutants of *Escherichia coli* TyrR. *J. Bacteriol.* **190**, 8238–8243 (2008).
41. Cui, J. & Somerville, R. L. Mutational uncoupling of the transcriptional activation function of the TyrR protein of *Escherichia coli* K-12 from the repression function. *J. Bacteriol.* **175**, 303–306 (1993).
42. Lawley, B. & Pittard, A. J. Regulation of *aroL* expression by TyrR protein and Trp repressor in *Escherichia coli* K-12. *J. Bacteriol.* **176**, 6921–6930 (1994).
43. Yang, J., Wang, P. & Pittard, A. J. Mechanism of repression of the *aroP* P2 promoter by the TyrR protein of *Escherichia coli*. *J. Bacteriol.* **181**, 6411–6418 (1999).

Acknowledgements

This work was supported by the Joint BioEnergy Institute, contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy and by the Synthetic Biology Engineering Research Center (SynBERC) through National Science Foundation grant 0540879. We thank Dr. Darmawi Juminaga (Joint Bioenergy Institute) for the gifts of *E. coli* DJ106, DJ166, and DJ238, Dr. Mario Ouellet (Joint Bioenergy Institute) for the gift of *S. cerevisiae* MO219, Dr. Adrienne McKee (Joint Bioenergy Institute) for the gifts of pT-LYCM4 and *mutD5*, and Dr. James Carothers for critical review and editing of the manuscript.

Author contributions

H.H.C. designed and performed the experiments. H.H.C. and J.D.K. conceived the project, analysed the data, and wrote the paper.

Additional information

Genbank accession numbers: Sequences of *E. coli* evolved with FREP for increased IPP production have been deposited in the NCBI Sequence Read Archive and are available through GenBank. The accession numbers for the sequences of the samples at the time points 0, 72, 144, 216, 288, 360 and 432 h are SRR956947, SRR957648, SRR958168, SRR958194, SRR 958276, SRR958416 and SRR958812, respectively.

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Conflict of Interest: J.D.K. has a financial interest in Amyris, LS9 and Lygos. The remaining authors declare no competing financial interest.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Chou, H. H. & Keasling, J. D. Programming adaptive control to evolve increased metabolite production. *Nat. Commun.* **4**:2595 doi: 10.1038/ncomms3595 (2013).