

Coordinated induction of multi-gene pathways in *Saccharomyces cerevisiae*

Jing Liang¹, Jonathan C. Ning¹ and Huimin Zhao^{1,2,*}

¹Department of Chemical and Biomolecular Engineering and ²Departments of Chemistry, Biochemistry, and Bioengineering, Center for Biophysics and Computational Biology, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received July 24, 2012; Accepted November 12, 2012

ABSTRACT

Bacterial operons are nature's tool for regulating and coordinating multi-gene expression in prokaryotes. They are also a gene architecture commonly used in the biosynthesis of many pharmaceutically important compounds and industrially useful chemicals. Despite being an important eukaryotic production host, *Saccharomyces cerevisiae* has never had such gene architecture. Here, we report the development of a system to assemble and regulate a multi-gene pathway in *S. cerevisiae*. Full pathways can be constructed using pre-made parts from a plasmid toolbox. Subsequently, through the use of a yeast strain containing a stably integrated gene switch, the assembled pathway can be regulated using a readily available and inexpensive compound—estradiol—with extremely high sensitivity (10 nM). To demonstrate the use of the system, we assembled the five-gene zeaxanthin biosynthetic pathway in a single step and showed the ligand-dependent coordinated expression of all five genes as well as the tightly regulated production of zeaxanthin. Compared with a previously reported constitutive zeaxanthin pathway, our inducible pathway was shown to have 50-fold higher production level.

INTRODUCTION

Coordinated gene expression is an important tool in the biosynthesis of natural products, fine chemicals and fuels (1–3). For example, the taxol biosynthetic pathway involves the synchronized action of at least 19 genes (4). In prokaryotes, genes involved in a biosynthetic pathway are often grouped together into an operon and are

transcribed as a polycistronic messenger RNA (mRNA) under the regulation of a single promoter (5,6). The operon structure makes it easy to regulate a large number of genes (7). In eukaryotes, especially filamentous fungi and even some plants, genes in the same pathway are also often grouped together into a gene cluster; but in contrast to that in prokaryotes, the genes are not under the same promoter and may be subjected to independent regulation (8,9). With rare exceptions, operons generally do not exist in eukaryotes (10).

Saccharomyces cerevisiae is an important industrial production host for heterologous pathways and is generally suited for expressing pathways from fungi and plants, especially those involving cytochrome P450s (1). Being a eukaryote, it does not recognize operons, and each gene in a pathway will need to have its own promoters. This requirement can make construction of multi-gene pathways cumbersome. Traditionally, genes in pathways are usually broken up into multiple plasmids—at times only one gene on a plasmid (11) or two genes under a divergent promoter (12). In either format, the construction of a long pathway will require many plasmids and, correspondingly, many selection markers. Thanks to the recently developed large-scale cloning methods, such as SLIC (13), Golden Gate (14), Gibson *et al.* (15) and DNA Assembler (16), large concatenations of promoters and genes can now be routinely made. However, large pathways constructed thus far are driven by constitutive promoters because, with the exception of Golden Gate, most of the assembly methods are homology dependent (17,18), and non-homologous inducible promoters that respond to the same inducer are limited. Furthermore, it is commonly believed that because of the active homologous recombination machinery in yeast, homologous promoters should be avoided even if they do not interfere with DNA assembly (1).

GAL-inducible promoters are commonly used for controlling gene expression in yeast (19). However,

*To whom correspondence should be addressed. Tel: +1 217 333 2631; Fax: +1 217 333 5052; Email: zhao5@illinois.edu

because of the high cost of the inducer, galactose, its use in industrial production is limited (20). To circumvent this problem, we have used an estrogen receptor-based gene switch to activate GAL promoters. Because of its nanomolar sensitivity, only 200 μmol of 17- β -estradiol is needed to fully induce a 20000 l reactor, at a cost of roughly \$1.20 (Sigma Aldrich catalogue 2012).

In this study, we present a system of constructing and activating multi-gene pathways in *S. cerevisiae*. The system consists of three parts: (i) a collection of inducible promoters; (ii) an estradiol-inducible yeast strain; and (iii) a method of DNA assembly that allows for easy construction of inducible pathways in one step from polymerase chain reaction (PCR) products. As a proof of concept, we assembled a five-gene zeaxanthin pathway, demonstrated the estradiol-dependent production of zeaxanthin and characterized the pathway's behaviour.

MATERIALS AND METHODS

Strains, plasmids, media, reagents and cell cultivation

TOP10 [F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*lacZ* Δ M15 Δ *lacX74 recA1 araD139 Δ (*ara-leu*) 7697 *galU galK rpsL* (StrR) *endA1 nupG*] and Stbl3 [F- *mcrB mrr hsdS20* (r_B^- , m_B^-) *recA13 supE44 ara-14 galK2 lacY1 proA2 rpsL20* (Str) *xyl-5* λ *leu mtl-1* r F- *mcrB mrr hsdS20* (r_B^- , m_B^-) *recA13 supE44 ara-14 galK2 lacY1 proA2 rpsL20* (StrR) *xyl-5* λ *leu mtl-1*] (Life Technologies, Carlsbad, CA, USA) were used for routine plasmid cloning and amplification in *Escherichia coli*. YM954 (*MATa; ura3-52 his3-200 ade2-100 lys2-801 trp1-901 leu2-3 112 gal4-542 gal80-538*) (kindly provided by Florence Vignols and Stanley Fields) was used for yeast experiments. Yeast strains were cultivated in either synthetic dropout medium (0.17% of Difco yeast nitrogen base without amino acids and ammonium sulphate, 0.5% of ammonium sulphate and 0.083% of amino acid dropout mix) or YPA medium (1% of yeast extract, 2% of peptone and 0.01% of adenine hemisulphate) supplemented with 2% of glucose as a carbon source. *Escherichia coli* strains were cultured in Luria broth (LB; Fisher Scientific, Pittsburgh, PA, USA).*

YZE-19 through 149 and YZE-PA were created in this study by integrating the P65-Gal-ER (6GE) gene switch into YM954 at the *LYS2* locus and are *lys2+*. pLys2-P_{ENO19,55,100,149}-6GE and pLys2-P_{ADH1}-6GE were linearized with DraIII and transformed into YM954 by chemical transformation as described elsewhere (21). Integrants were selected using SC-Lys plates and were confirmed by PCR as well as functional assay.

All restriction and DNA processing enzymes were obtained from New England Biolabs (Ipswich, MA, USA), and all chemicals were obtained from Sigma Aldrich (St Louis, MO, USA) unless otherwise specified. PCR reactions were carried out using Phusion DNA Polymerase (Dharmacon, Lafayette, CO, USA), and primers were from Integrated DNA Technologies (Coralville, IA, USA).

Plasmid construction

The natural GAL promoters were PCR amplified directly from the genomic DNA of *S. cerevisiae*. The resulting PCR products were overlap extended with another PCR fragment encoding the green fluorescent protein (GFP) and the ADH1 terminator. The promoter-gene-terminator cassette was subsequently ligated into pRS416 centromeric vector (ATCC 87521) via BamHI and SacI to create pRS416-Gal1-GFP, pRS416-Gal2-GFP and pRS416-Gal10-GFP. The synthetic GAL upstream activation sequences (UAS) were synthesized from overlap extension of oligonucleotides. These were overlap extended with the TATA box region of Gal1 and Gal2 to create the SYN1 and SYN2 series of promoters, respectively, and subsequently cloned into pRS416 to create pRS416-SYN1-GFP and pRS416-SYN2-GFP. The DNA sequences of the SYN1 and SYN2 promoters can be found in the Supplementary Data.

The gene switch constructs were generated by ligation of either a P65 or GAL4 activation domain (AD) with an estrogen receptor ligand-binding domain (ER LBD) and a GAL4 DNA-binding domain (DBD). The P65 and ER LBD were PCR amplified from human genomic DNA, whereas the GAL4 AD and DBD were amplified from *S. cerevisiae* genomic DNA. The fragments were cloned directly into pRS414-PMT, which is a pRS414 vector (ATCC 87519) modified to contain an ADH1 promoter and ADH1 terminator interspaced by a multiple cloning site.

The gene switch integration plasmids, pLys2-P_{ENO19,55,100,149}-6GE and pLys2-P_{ADH1}-6GE, were created by ligating the promoter fragments (EcoRI, KpnI) and 6GE-T_{ADH1} (KpnI, MluI) into pLys2 (EcoRI, MluI), which is pNEB193 (New England Biolabs, Ipswich, MA, USA) with the *LYS2* gene inserted into the AatII site. P_{ENO19,55,100,149} promoters were obtained by random mutagenesis of the wild-type ENO promoter as described elsewhere (22).

The T-P plasmids were constructed by the ligation of the PGI1 terminator (AflII, EcoRI) and the GAL2 promoter (EcoRI, XbaI) into pFUS_A (AflII, XbaI) with appropriate BsaI sites introduced at the two ends of the T-P cassette. pFUS_A is an *E. coli* plasmid with a spectinomycin resistance marker as described elsewhere (23). The PGI1 terminator was PCR amplified from pRS416-Zeax, which is a constitutive zeaxanthin pathway as described elsewhere (16), whereas the GAL2 promoter was PCR amplified from the aforementioned pRS416-Gal2-GFP.

The Golden Gate receiver plasmid, containing the first promoter and the last terminator interspaced by a BsaI-flanked LacZ α cassette, was constructed by the overlap extension of the GAL2 promoter (from pRS416-Gal2-GFP), LacZ α cassette (from pFUS_A) and ADH1 terminator (from pRS416-Gal2-GFP). The overlap extended fragment was cloned into pRS416K2 via HindIII and NotI. pRS416K2 was modified from pRS416 by the removal of all BsaI sites, the LacZ α remnant, the fl origin and by the replacement of AmpR with KanR.

Pathway assembly

The Golden Gate reaction for pathway assembly was carried out in 20 μ l using 100 ng of receiver plasmid, 100 ng of each T-P plasmid, 10 ng/kb of the PCR gene fragments, 1 μ l of BsaI-HF and 1 μ l high-concentration T4 DNA ligase, in 1X T4 ligase buffer. The reaction was thermocycled as follows: 37°C for 10 min, (37°C for 5 min, 16°C for 10 min) repeated 10 times, 37°C for 10 min, 75°C for 5 min. The ligation product was transformed into TOP10 competent cells (Life Technologies, Carlsbad, CA, USA) by heat-shock, and white colonies on X-GAL/IPTG plates were screened for the correctly assembled pathway. Primer sequences used to obtain the PCR gene fragments are listed in Supplementary Table S1.

Measurement of GFP expression

GFP expression level was measured by flow cytometry using LSRII (BD Biosciences, Franklin Lakes, NJ, USA). Single colonies were picked into liquid synthetic dropout media and were grown overnight with or without induction. Small aliquots of the cultures were then washed and resuspended in phosphate buffered saline for flow cytometry analysis.

Measurement of zeaxanthin production

Cells from 6 ml of *S. cerevisiae* dropout media culture were collected by centrifugation, resuspended in 1 M of sorbitol, 10 mM of Tris buffer pH 7.4, 5 U/ml Zymolyase (Zymo Research, Irvine, CA, USA) and incubated at 37°C for 45 min. Cells were then collected again by centrifugation, and 600 μ l of methanol was used to extract the zeaxanthin directly. For quantification, 80 μ l of methanol extract was loaded onto an Agilent ZORBAX SB-C18 column and monitored at 450 nm on an Agilent 1100 series high-performance liquid chromatography (HPLC) (Agilent Technologies, Palo Alto, CA, USA). The pump program was 0.6 ml/min, 100% of methanol for 10 min. Elution was monitored at 450 nm with reference set at 360 nm. Authentic zeaxanthin from Sigma (St Louis, MO, USA) was used as a standard.

RNA preparation and quantitative PCR

Yeast colonies were picked from synthetic dropout plates into 3 ml synthetic dropout medium and grown overnight at 30°C with shaking at 250 r.p.m. to saturation. Hundred microlitres of overnight culture was then used to inoculate 3 ml of fresh media, and estradiol was added to the appropriate samples. After growing for 18 h, cells from 0.5 ml of each culture were harvested by centrifugation. RNA was then isolated using the RNeasy Mini Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's protocol. One microgram of the resulting total RNA was treated by 10 U of recombinant DNaseI (Takara Bio, Otsu, Shiga, Japan) in 20 μ l for 40 min according to the manufacturer's protocol. Complementary DNA (cDNA) was then synthesized using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol and was stored at -20°C until use. Quantitative PCR primers were designed using

the Primer3Plus software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) (Supplementary Table S2). Reactions were performed using Applied Biosystems 7900HT Fast Real-Time PCR System (Applied Biosystems, Austin, TX, USA) according to the manufacturer's protocol. Expression was quantified relative to the ALG9 gene using the standard curve method (24).

RESULTS

Inducible promoter collection

The availability of natural GAL responsive promoters is limited to the genes involved in galactose utilization. To address this limitation, we cloned the GAL1, GAL2 and GAL10 promoters into a GFP reporter and induced the GFP expression using Gal-P65, a constitutive gene switch made from the fusion between the GAL4 DNA-binding domain and the P65-activation domain. As shown in Figure 1A, the natural GAL promoters exhibited tight regulation—the uninduced fluorescence level was indistinguishable from the background fluorescence.

To expand the dynamic range of the promoter collection, two sets of synthetic GAL responsive promoters were created. DNA fragments containing a varying number of consensus UAS were joined to the TATA region of GAL1 and GAL2, creating the SYN1 and SYN2 collections of promoters, respectively. As shown in Figure 1B and C, the induction strength of the promoter can be reduced by decreasing the number of UAS. In contrast to the tightly regulated GAL1 promoter, the SYN1 collection of promoters had noticeably higher basal levels. The SYN2 collection of promoters did not suffer from increased basal level, and it was generally weaker than the natural promoters. With the natural promoters and the SYN2 collection of promoters, induction strength across a wide dynamic range can be selected while the pathway of interest is maintained in tight control. Further work on the engineering of GAL inducible promoters have recently been described by Blazeck *et al.* (25), which provided an even wider range of GAL inducible promoters to choose from.

Gene switch engineering and strain development

Domain selection and permutation

To identify the optimum configuration for the gene switch, we tested different orders of functional domains and different activation domains. The gene switches were co-transformed with P_{GAL10}-GFP and tested for their inducibility using 100 nM of estradiol. As shown in Figure 2A, P65-Gal-ER (6GE) had the best compromise between induction level and basal expression level and was selected for the development of a series of estradiol-inducible yeast strains.

Estradiol-inducible YZE strains

After identifying the best gene switch configuration, we moved on to integrate the gene switch into YM954 so as to create a more convenient induction system as well as to reduce the number of plasmids required. The gene switch

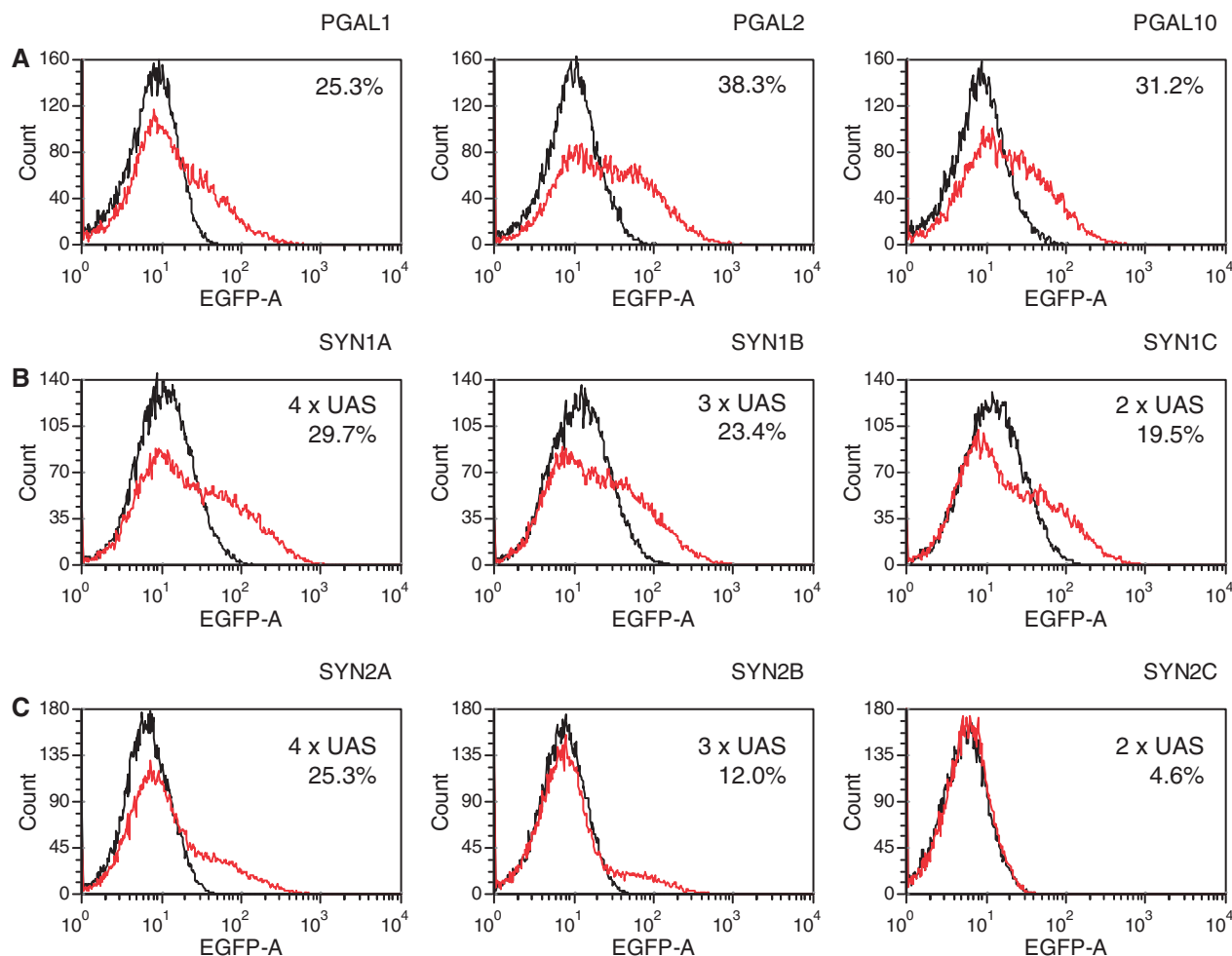


Figure 1. The behaviour of GAL inducible promoters, monitored via GFP expression and induced using Gal-P65. The GFP channel histograms obtained from flow cytometry are shown. Black indicates uninduced sample and red indicates induced sample. The sample identity and the percentage of positives are shown within each histogram. (A) The behaviour of natural GAL inducible promoters cloned from *S. cerevisiae*. (B) The behaviour of synthetic GAL inducible promoters, which were generated by fusing the synthetic UAS sequences with the TATA box of the GAL1 promoter. (C) Similar to (B), but fused with the TATA box of the GAL2 promoter.

was integrated into the *Lys2* locus without the use of antibiotic selection, thus leaving the most common auxotrophic markers intact and allowing the use of antibiotics for further genome engineering. As the gene switch protein is not catalytically involved in the production of any desired compound, its overexpression will likely be detrimental to production. By placing the gene switch behind promoters of different strengths, the gene switch expression level will be varied in the different strains. Gene switch expression in YZE-19, YZE-55, YZE-100 and YZE-149 are driven by the ENO promoter and its derivatives—the numbering indicates their relative percentage strength to the wild-type ENO promoter. The expression in YZE-PA is driven by the ADH1 promoter. The functionality of these strains was confirmed by observing fluorescence from the induction of P_{GAL10} -GFP on estradiol addition. The flow cytometry GFP channel histograms from the strains were shown in Figure 2B. The mean fluorescence value shifted towards the right at higher promoter strengths. YZE-PA was not shown, but its histogram was indistinguishable from that of YZE-55.

Golden gate pathway assembly

Pathway assembly

Because of the limited availability of optimal inducible promoters, we designed the pathway assembly system based on the Golden Gate cloning method, so that a single good inducible promoter can be re-used to control multiple genes if necessary. Details of Golden Gate cloning have been described by Carola *et al.* (14). Briefly, each fragment in the assembly is flanked by type II restriction sites, for example, BsaI, which cleaves the DNA away from its recognition sequence, leaving a user-defined 4 bp overhang. By judiciously choosing unique 4 bp overhangs, these fragments can be ligated together in a specific order to form the pathway. Because the restriction sites are lost when the correct ligation happens, the restriction and ligation reaction can be carried out simultaneously in one tube.

To allow for modular assembly of pathways, we constructed a toolbox that contains plasmids carrying the intergenic terminator-promoters and a receiver plasmid that has the first promoter and the last terminator.

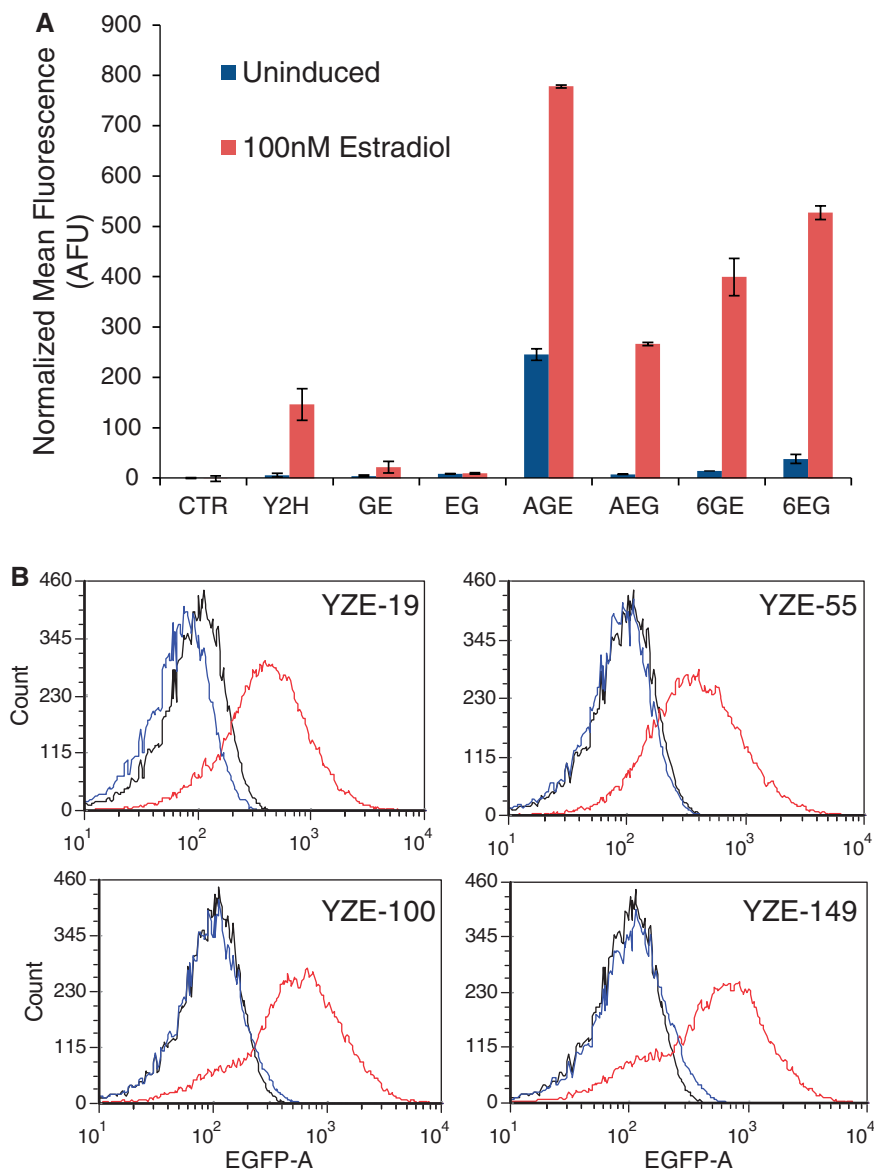


Figure 2. (A) Gene switch activity monitored via GFP expression and measured using flow cytometry. GFP expression levels are represented by the arithmetic means of GFP fluorescence (arbitrary unit). The mean of negative control has been subtracted from that of the samples. Blue columns represent uninduced samples, and red columns represent induced samples. Values are the average of two independent samples, and error bar indicates the standard error. Labels: CTR is negative control with empty plasmids, and Y2H is a yeast two-hybrid system positive control from (26). GE, EG, AGE, AEG, 6GE and 6EG are different combinations and permutations of Gal4 DBD(G), ER LBD(E), P65(6) and Gal4 AD(A) ordered from the N-terminus of the resulting gene switch protein. (B) Inducibility of the YZE strains monitored via GFP expression. The GFP channel histograms are shown. Black indicates negative control without gene switch, blue indicates uninduced sample and red indicates induced sample.

The assembly scheme is shown in Figure 3A. For proof of concept, these plasmids were put together with the PCR products of the zeaxanthin pathway genes (Crt E, B, I, Y and Z), and an inducible zeaxanthin pathway was constructed in a one-step Golden Gate reaction. For this pathway, $T_{PGI1-P_{GAL2}}$ was selected for all intergenic terminator-promoters, P_{GAL2} was selected as the first promoter, and T_{ADH1} was selected as the last terminator (Figure 3B and Supplementary Figure S2). As all five genes are under the same promoter, it allows for easier characterization of the inducible pathway. The assembly efficiency for the five-gene pathway was $\sim 16\%$ (2/12), and an alternative pathway that has $T_{TPI1-P_{GAL10}}$ in the T_1-P_2

position was also constructed with similar efficiency. When the assembled pathway was co-transformed with a plasmid carrying the P65-Gal-ER gene switch, the inducible expression of zeaxanthin occurred only in the presence of estradiol (Figure 4A).

Plasmid stability

The intergenic T-P regions are ~ 800 -bp long, and as $T_{PGI1-P_{GAL2}}$ was used for all T-P positions, the resulting plasmid contained four 800 bp direct repeats. It is commonly believed that plasmids with repetitive sequences are unstable in *S. cerevisiae* because of the presence of a highly active homologous recombination

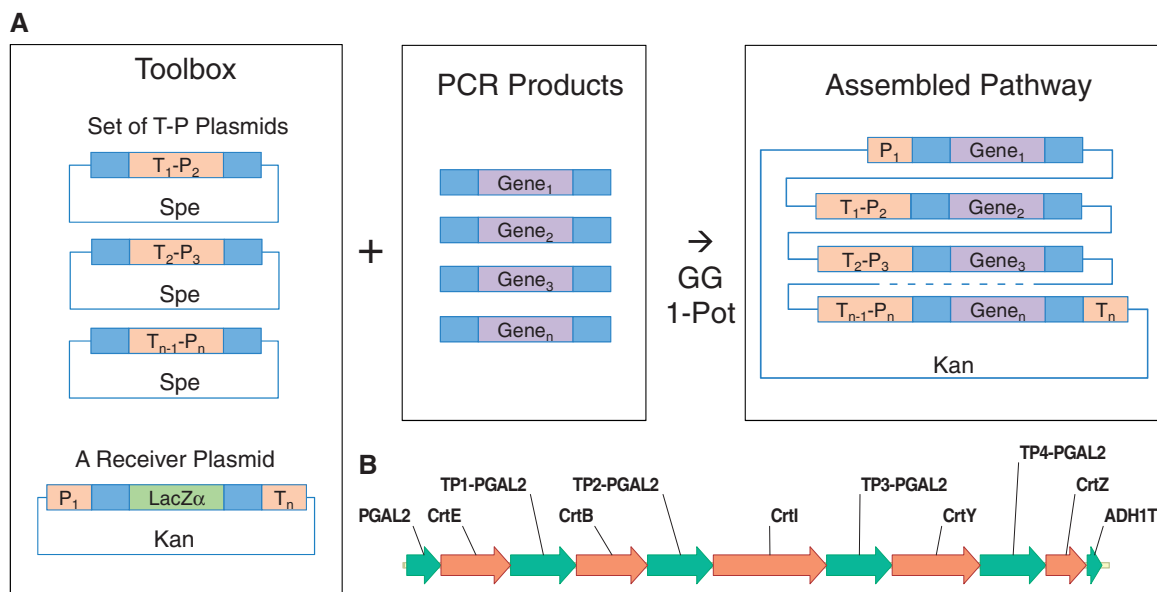


Figure 3. (A) The inducible pathway assembly scheme. The T-P plasmids have BsaI-excisible terminator and promoter fragments, and the receiver plasmid has BsaI-excisible LacZ α marker for blue-white screening. The toolbox plasmids together with the PCR products of the pathway genes can be used to assemble an inducible pathway in a Golden Gate one-pot (GG 1-Pot) reaction. BsaI sites are in light blue, Kan is the kanamycin resistance gene, Spe is the spectinomycin resistance gene. (B) A schematic of the assembled inducible zeaxanthin pathway. All 4 intergenic T-P fragments are identical except for the 4 bp overhang region.

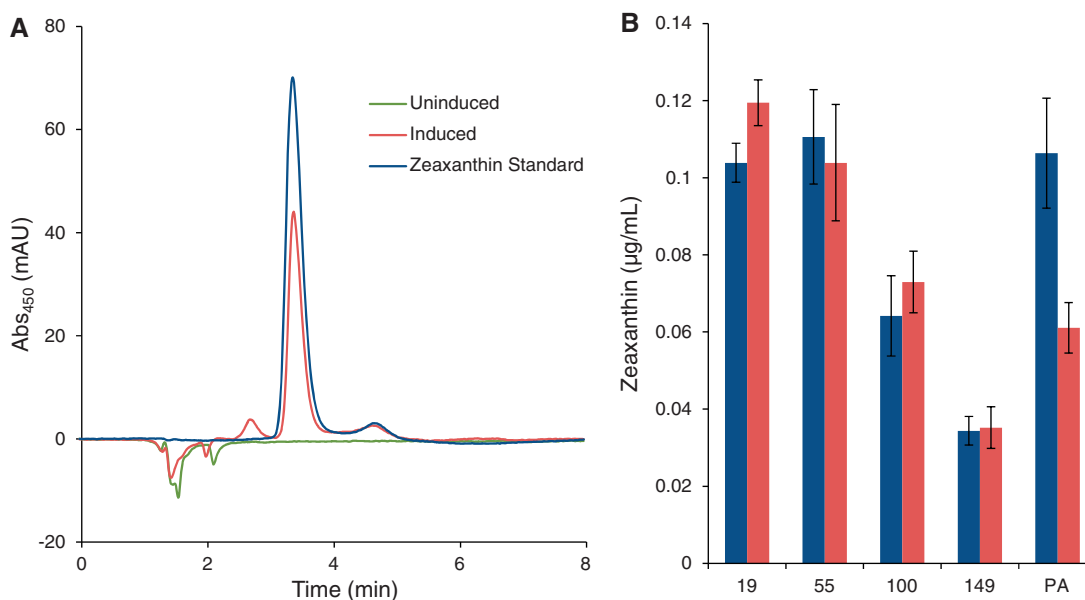


Figure 4. (A) HPLC traces showing the inducible production of zeaxanthin. (B) Zeaxanthin production in different YZE hosts. Blue columns represent data from the first 24-h samples. At the end of the 24 h, cells were re-inoculated into fresh media for another 24-h. Red columns represent data from the second 24-h samples. The values represent the average of four independent samples, and the error bar represents the standard error.

machinery (1,27). We, therefore, performed a test of plasmid stability to see whether the pathway will remain intact over multiple generations of propagations.

As the complete zeaxanthin pathway produces a yellow compound (and thus yellow colonies) readily detectable by visual inspection, we assessed the plasmid stability by monitoring the number of white versus yellow colonies over continuous culture. The aforementioned zeaxanthin pathway was separately transformed into YZE-55 and

YZE-PA and was then plated onto selective plates. The initial colonies were collected, diluted and re-plated with estradiol. The number of yellow and white colonies in this first re-plate was labelled as day 0 in Table 1. A yellow colony was then picked for uninduced continuous culture, diluting daily at 1/100, and an aliquot of the culture was plated with estradiol every 2 days (Supplementary Figure S1). As shown in Table 1, there was no significant increase in the number of white colonies after 8 days of continuous

Table 1. Assessment of plasmid stability. White colony is an indication that the yeast cell has lost at least part of the zeaxanthin pathway

Integrand	Time	Total colony	White colony	Percentage of white colonies
YZE-55	Day 0	60	12	20.00
	Day 2	243.5	0.5	0.21
	Day 4	195	1	0.51
	Day 6	360	1	0.28
	Day 8	267.5	1.5	0.56
YZE-PA	Day 0	110	39	35.45
	Day 2	295	1.5	0.51
	Day 4	195.5	0.5	0.26
	Day 6	404	2	0.50
	Day 8	267	1	0.37

The colony counts are the average of two independent experiments. For detail experiment scheme, see Supplementary Figure S1 in the Supplementary Information.

culture. With the exception of the day 0 re-plate, there were hardly any white colonies. A cell growing from the day 0 re-plate to day 8 saturated culture had gone through about 63 generations of divisions, more than enough to saturate a 20 000 l batch reactor, and yet, only ~0.5% of the cells had lost the pathway. This showed that the zeaxanthin pathway, despite multiple direct repeat sequences, could be stably propagated in *S. cerevisiae*.

Inducible pathway characterization

Protein level induction

Although the promoters were characterized individually before the pathway assembly, putting multiple promoters together in the same plasmid may result in position dependency. Therefore, we sought to investigate whether the protein induction levels depend on where the gene is situated within the long concatenation of gene cassettes. Using the same toolbox plasmids as aforementioned, five more pathways were assembled, each with a different zeaxanthin pathway gene replaced with the GFP gene. The GFP genes in the five plasmids are thus under P_{GAL2} promoters. Assuming that neighbouring structural genes do not interfere with promoter activity, any difference in induction level can be attributed to positional effect. As shown in Figure 5A, GFP at all five positions was estradiol inducible. The basal level expressions were indistinguishable from background fluorescence. The protein induction level seemed to dip at the third and fourth position, but the expression level at all five positions varied by no more than two-fold. This showed that the genes at all five positions can be effectively regulated, even though the exact induction strength has some position dependency.

mRNA level induction

Next, we investigated whether all five pathway genes can be induced simultaneously. Because the proteins were not tagged, we instead measured the mRNA-level induction of the zeaxanthin pathway genes. As shown in Figure 5B, all five zeaxanthin pathway genes can be induced, and the trend in relative induction strength is

the same as what we have observed from the GFP fluorescence measurement. This gave further evidence of the positional dependency of promoter strength. Because mRNA and protein abundance are not linearly correlated, a big change in mRNA level does not correspond to a big change in protein expression level, and only the trend is consistent. Interestingly, when mRNA induction ratio and protein abundance as reported by GFP fluorescence are plotted against each other, a power-law trend emerged (Supplementary Figure S3), which coincides with the trend observed by Beyer *et al.* (28) in their genome scale study of transcription–translation relationships in *S. cerevisiae*.

Inducible strain comparison

Production of zeaxanthin was compared in the five YZE strains that differed in their promoters driving the gene switch expression. The inducible zeaxanthin pathway was transformed into the five strains and tested for their zeaxanthin production at 24 and 48 h after induction. After sampling at 24 h, the culture was re-inoculated 1/100 into fresh media. As shown in Figure 4B, high gene switch promoter strengths, for example, YZE-100 and YZE-149, are bad for production. YZE-19, 55 and YZE-PA showed similar production level at 24 h, but YZE-PA's production decreased after re-inoculation. YZE-19 and YZE-PA were picked for further comparison.

Production time curve

YZE-19 and YZE-PA, each harbouring an inducible pathway plasmid, were compared with YM954 harbouring a constitutive pathway. The constitutive pathway was obtained from a previous publication and was used as a benchmark in our comparison (16). To make the result more comparable, the constitutive pathway was re-cloned into pRS416K2, which is the backbone of the inducible pathway, and YZE's parent strain, YM954, was used to host the pathway. As shown in Figure 5C, YZE-19 is the best producer. At 72 h post induction, YZE-19 accumulated roughly twice the zeaxanthin compared with YZE-PA and ~50-fold more compared with the constitutive pathway. The measured concentration by HPLC was 1.5 $\mu\text{g}/\text{ml}$, which corresponds to 15 ng/OD₆₀₀ cells or ~75 $\mu\text{g}/\text{g}$ dry cell weight.

Ligand titration

To identify an optimal inducer concentration, we performed a ligand titration in YZE-19 and YZE-PA. The strains harbouring the inducible pathway were subjected to different estradiol concentration, and the zeaxanthin production was measured after 48 h. In agreement with the time course experiment, YZE-19 is a better producer. As shown in Figure 5D, it is in fact more sensitive to estradiol than YZE-PA. Peak induction in YZE-19 can be obtained at 10^{-8} M of estradiol.

DISCUSSION

In this study, we have presented a system for the construction and regulation of multi-gene pathways in *S. cerevisiae*. Genes of an entire pathway can be

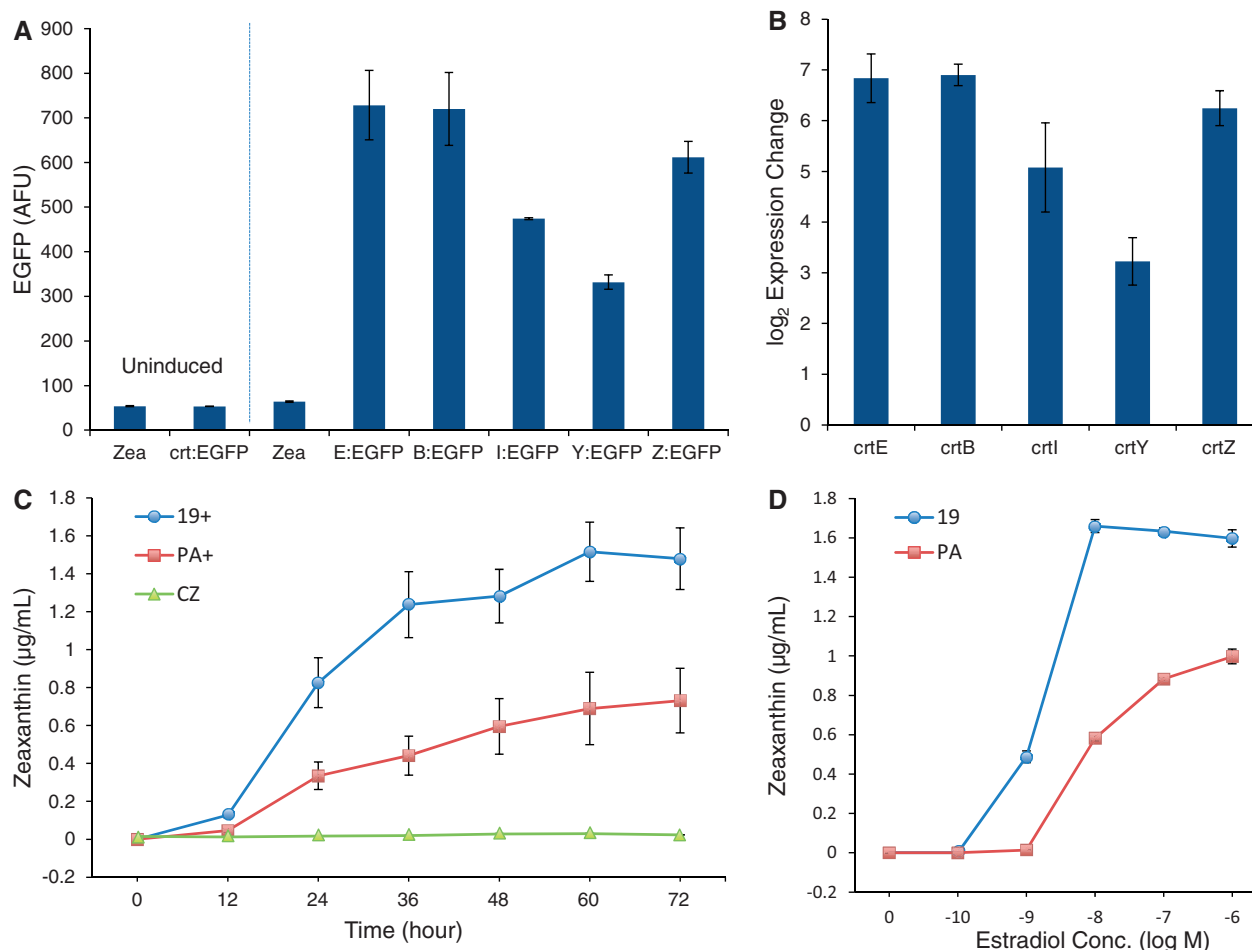


Figure 5. (A) Protein-level induction at the different positions using EGFP as the reporter. The genes along the pathway are arranged from 5' to 3' in the order Crt E-B-I-Y-Z. Zea is the negative control with the five-gene inducible zeaxanthin pathway without any EGFP gene in it. E, B, I, Y, Z:EGFP have their respective zeaxanthin pathway gene replaced by EGFP. The uninduced EGFP-replaced pathways have identical basal expression levels, and they have been collapsed under crt:EGFP. The values represent the average of four independent samples, and the error bar represents the standard error. (B) mRNA level induction at the different positions as measured by quantitative reverse transcriptase-PCR. The value represents the average of two independent samples, and the error bar represents the standard error. (C) Production time course that benchmarks the inducible pathway in two different YZE strains to a constitutive pathway in YM954. The values represent the average of three independent samples, and the error bar represents standard error. (D) Ligand-titration curve of the zeaxanthin production against the estradiol concentration. The values represent the average of two independent samples, and the error bar represents standard error.

concatenated together, each driven by an inducible promoter, forming a transcriptional unit that can be regulated as a whole—similar to what can be achieved by a bacterial operon. By using inducible promoters of different strengths for different genes, it is possible to balance the flux through the pathway, which is the main challenge in metabolic engineering and synthetic biology (29).

We currently have a set of six promoters that are tightly regulated, with induction strength that spans an 8-fold difference from the weakest to the strongest. More inducible promoters have recently been developed by Blazeck *et al.* (25), and they can easily be incorporated into our pathway construction system. As indicated in Figure 5D, induction strength can depend on the gene switch's expression level and can, therefore, differ in the different YZE strains. When using promoter strength to tune pathway gene expressions, it is also possible to pick induction strength through varying the ligand concentration.

However, ligand titration curves need to be obtained in the host strain to find out what the induction level is at any given ligand concentration. Furthermore, in this non-saturation range, small differences in ligand concentration can cause large differences in induction strength, making it difficult to control the induction level. For the aforementioned reasons, we find that it is best to vary induction strength by varying maximum induction levels.

Based on the results of gene switch configuration assessment, we integrated the best gene switch into the chromosome to create five inducible yeast strains differing in gene switch expression levels. When used in conjunction with the inducible pathways, no additional plasmid is necessary. As shown in Figure 4B, high gene switch expression can be detrimental to the production of the desired product. Because of high clonal variation in adaptation, when a pathway is transformed into any given strain and induced on a plate, colonies with different degrees of yellow colouration (indicating zeaxanthin production)

will occur. We picked both light- and dark-yellow colonies, assayed their product and confirmed that they were all producing zeaxanthin, just at different levels (data not shown). High zeaxanthin producing colonies occur at a frequency of roughly 1 in 10, and the adaptation can be propagated. When comparing the zeaxanthin production level in the different strains, we selected the darkest yellow colonies from each plate.

Using the Golden Gate assembly method and our plasmid toolbox, the five-gene zeaxanthin pathway can be assembled from their PCR products in a single step. This corresponds to a simultaneous ligation of 10 fragments forming a 13.3 kb plasmid with an efficiency of 16%. A similar efficiency was observed for the assembly of two zeaxanthin pathways and five GFP-replaced pathways. As the correct pathways could be picked from a reasonable number of colonies, no optimization was performed. For longer pathways, optimization of the Golden Gate reaction condition will probably be necessary, and should optimization fail, a two-stage hierarchical assembly can instead be used. In that case, the toolbox will be expanded to include two to three ampicillin-resistant intermediate plasmids. PCR products of the genes and the T-P fragments from T-P plasmids will first be assembled into the intermediate plasmids, up to five genes at a time. The intermediate plasmids will then be used as the substrate for a second Golden Gate reaction, which puts the intermediate assemblies together into the final receiver plasmid. If a target gene contains the BsaI restriction site, it is possible to PCR-amplify the gene in multiple segments, breaking and replacing each natural BsaI site through primers. These fragments can then be used in the 1-step Golden Gate assembly system as per normal. Another way to work around the restriction site is to use another type II restriction enzyme that gives a 4-bp 5' overhang for the problematic target gene, and break the 1-step assembly to 2-step assembly involving first digesting and then ligating.

The initial transformations of the assembled zeaxanthin pathway have a significant fraction of white colonies (Table 1), but when a yellow colony is picked and propagated, only 0.5% of the white colonies can be observed after 8 days of continuous cultivation, with no increase in the percentage of white colonies over the 8 days. This suggests that the plasmid may be less stable in *E. coli*, such that the initial plasmid preparation contained a small portion of truncated recombinant plasmids, and the proportion was magnified through the preferential transformation of smaller plasmids. Interestingly, when the plasmid for yeast transformation was prepared in the Stb13 strain instead of the TOP10 or DH5 α strains, the fraction of white colonies decreased drastically. As Stb13 is a strain better suited for amplifying unstable plasmids, it gives further evidence for plasmid instability in *E. coli*.

Ideally, promoter strength should be independent of its position in the pathway, but this was not the case, at least for the GAL2 promoter. We observed that the induction strength tended to dip towards the middle of the pathway, and this has been supported at both protein (Figure 5A) and mRNA (Figure 5B) levels. One possible cause of this observation could be the one-dimensional search

mechanism of DNA-binding proteins. DNA-binding proteins look for their cognate sequences in many ways, and one of the major search mechanisms is by first binding non-specifically to DNA, then performing a one-dimensional search along the DNA (30–32). The promoters at the flanks have access to a large ring of non-cognate DNA, that is, the plasmid backbone, and the gene switch will, therefore, have a higher chance of finding these promoters. In contrast, the promoters in the middle will have a much smaller binding area, and they will have to rely more on a three-dimensional diffusion mechanism. More experiments are needed to verify this explanation and come up with a way to predict the modification of induction strength because of positional effect.

Despite some positional dependency of induction strength, the pathway is tightly regulated, with no measurable production in the absence of estradiol. To benchmark the induced production, we used a constitutive zeaxanthin pathway previously reported in literature (16). Under the condition of our production time curve experiment, the inducible pathway resulted in ~50-fold higher production level than the constitutive pathway. In making the comparison, we note that the genes on the constitutive pathway have five different constitutive promoters. As constitutive promoters can sometimes be repressed under certain culture conditions, we may have just picked a poor production condition for the constitutive pathway. Although we have not demonstrated the superior production capacity of an inducible pathway because of the aforementioned reason, we have at least demonstrated the greater reliability and predictability of an inducible system. Indeed, in another previously reported zeaxanthin production using *S. cerevisiae*, the authors achieved a similar production level as our inducible pathway using a mixture of inducible and constitutive promoters (2,33).

We chose GAL4 DBD and its corresponding UAS in our study because the GAL4 DBD-based induction system has been well characterized, and it is commonly used for inducible expression in yeast. There is, however, a drawback in the system—being a native yeast protein, there are multiple regulatory and interaction targets for GAL4 DBD, and some of these interactions can lead to unexpected phenotypes. For example, the YZE strains will not flocculate when induced, but if un-induced, will flocculate when the carbon source is exhausted. This could be because of interactions of GAL4 with other cellular signalling pathways, with the overall effect of tricking the cell into acting as if there is still galactose around.

With the tools introduced in this study, biosynthetic pathways can be easily re-factored and studied in *S. cerevisiae*. More importantly, entire pathways can now be predictably and reliably induced using estradiol, a commonly available chemical. If desired, the ER LBD can also be engineered to generate new orthogonal ligand-receptor pairs, giving us the ability to regulate multiple pathways simultaneously and orthogonally (26,34,35). The ability to coordinate the expression of multiple genes in *S. cerevisiae* will be a useful addition to the toolbox of microbiologists, metabolic engineers and synthetic biologists.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2 and Supplementary Figures 1–3.

ACKNOWLEDGEMENTS

The authors thank Dr Florence Vignols and Dr Stanley Fields for sharing their yeast strains. J. Liang acknowledges the fellowship support from the Agency for Science, Technology, and Research (A*STAR), Singapore.

FUNDING

Funding for open access charge: National Institutes of Health [GM077596].

Conflict of interest statement. None declared.

REFERENCES

- Siddiqui, M.S., Thodey, K., Trenchard, I. and Smolke, C.D. (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res.*, **12**, 144–170.
- Pscheidt, B. and Glieder, A. (2008) Yeast cell factories for fine chemical and API production. *Microb. Cell Fact.*, **7**, 25.
- Du, J., Shao, Z. and Zhao, H. (2011) Engineering microbial factories for synthesis of value-added products. *J. Ind. Microbiol. Biotechnol.*, **38**, 873–890.
- Jennewein, S., Wildung, M.R., Chau, M., Walker, K. and Croteau, R. (2004) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc. Natl Acad. Sci. USA*, **101**, 9149–9154.
- Lawrence, J.G. and Roth, J.R. (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics*, **143**, 1843–1860.
- Lawrence, J.G. (1999) Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. *Curr. Opin. Genet. Dev.*, **9**, 642–648.
- Keasling, J.D. (2008) Synthetic biology for synthetic chemistry. *ACS Chem. Biol.*, **3**, 64–76.
- Osbourn, A.E. and Field, B. (2009) Operons. *Cell. Mol. Life Sci.*, **66**, 3755–3775.
- Field, B., Fiston-Lavier, A.S., Kemen, A., Geisler, K., Quesneville, H. and Osbourn, A.E. (2011) Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl Acad. Sci. USA*, **108**, 16116–16121.
- Blumenthal, T. (2004) Operons in eukaryotes. *Brief. Funct. Genomic. Proteomic.*, **3**, 199–211.
- Engels, B., Dahm, P. and Jennewein, S. (2008) Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metab. Eng.*, **10**, 201–206.
- Dejong, J.M., Liu, Y., Bollon, A.P., Long, R.M., Jennewein, S., Williams, D. and Croteau, R.B. (2006) Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.*, **93**, 212–224.
- Li, M.Z. and Elledge, S.J. (2007) Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat. Methods*, **4**, 251–256.
- Engler, C., Kandzia, R. and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A. and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
- Shao, Z., Zhao, H. and Zhao, H. (2009) DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.*, **37**, e16.
- Zotchev, S.B., Sekurova, O.N. and Katz, L. (2012) Genome-based bioprospecting of microbes for new therapeutics. *Curr. Opin. Biotechnol.*, **23**, 941–947.
- Merryman, C. and Gibson, D.G. (2012) Methods and applications for assembling large DNA constructs. *Metab. Eng.*, **14**, 196–204.
- Romanos, M.A., Scorer, C.A. and Clare, J.J. (1992) Foreign gene expression in yeast: a review. *Yeast*, **8**, 423–488.
- Nevoigt, E. (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.*, **72**, 379–412.
- Gietz, R.D. and Schiestl, R.H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.*, **2**, 31–34.
- Alper, H., Fischer, C., Nevoigt, E. and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl Acad. Sci. USA*, **102**, 12678–12683.
- Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J. and Voytas, D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.*, **39**, e82.
- Morrison, T.B., Weis, J.J. and Wittwer, C.T. (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques*, **24**, 954–958, 960–962.
- Blazek, J., Garg, R., Reed, B. and Alper, H.S. (2012) Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol. Bioeng.*, **109**, 2884–2895.
- Chockalingam, K., Chen, Z., Katzenellenbogen, J.A. and Zhao, H. (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc. Natl Acad. Sci. USA*, **102**, 5691–5696.
- Raymond, C.K., Powder, T.A. and Sexson, S.L. (1999) General method for plasmid construction using homologous recombination. *Biotechniques*, **26**, 134–138, 140–141.
- Beyer, A., Hollunder, J., Nasheuer, H.-P. and Wilhelm, T. (2004) Post-transcriptional expression regulation in the yeast *Saccharomyces cerevisiae* on a genomic scale. *Mol. Cell Proteomics*, **3**, 1083–1092.
- Du, J., Yuan, Y., Si, T., Lian, J. and Zhao, H. (2012) Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Res.*, **40**, e142.
- Slutsky, M. and Mirny, L.A. (2004) Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.*, **87**, 4021–4035.
- Wunderlich, Z. and Mirny, L.A. (2008) Spatial effects on the speed and reliability of protein-DNA search. *Nucleic Acids Res.*, **36**, 3570–3578.
- Tafvizi, A., Mirny, L.A. and van Oijen, A.M. (2011) Dancing on DNA: kinetic aspects of search processes on DNA. *Chemphyschem*, **12**, 1481–1489.
- Ausich, R.L., Brinkhaus, F.L., Mukharji, I., Proffitt, J.H., Yarger, J.G. and Yen, H.C. (1997) Biosynthesis of zeaxanthin and glycosylated zeaxanthin in genetically engineered hosts, US Patent 5684238.
- Chockalingam, K. and Zhao, H. (2005) Creating new specific ligand-receptor pairs for transgene regulation. *Trends Biotechnol.*, **23**, 333–335.
- McLachlan, M.J., Chockalingam, K., Lai, K.C. and Zhao, H. (2009) Directed evolution of orthogonal ligand specificity in a single scaffold. *Angew. Chem. Int. Ed.*, **48**, 7783–7786.