Available online at www.sciencedirect.com
ScienceDirect
Metabolic Engineering 9 (2007) 337–347
METABOLIC ENGINEERING
www.elsevier.com/locate/ymben

Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*

Yong-Su Jin[a,b], Gregory Stephanopoulos[a,*]

[a]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b]*Department of Food Science and Biotechnology, Sungkyunkwan University, Suwon, Republic of Korea*

Received 17 October 2006; received in revised form 2 March 2007; accepted 11 March 2007
Available online 12 April 2007

Abstract

Identification of multiple gene targets that exhibit different modes of action toward a desired phenotype is a crucial step in strain improvement. Target identification methods based on traceable genetic perturbations and stoichiometric modeling have been employed before for the mining of putative overexpression and knock-out targets. Most search methods are sequential and, as such, quite limited in the space they can explore. In this study, we investigate a multi-dimensional search approach whereby unknown interactions of gene targets identified by different search methods are assessed by employing orthogonal search strategies. To this end, we combined knock-out and overexpression gene targets, identified through systematic and combinatorial approaches, respectively, in order to improve lycopene production in *Escherichia coli*. Specifically, we first identified multiple overexpression targets by screening genomic libraries of *E. coli* in a sequential-iterative manner. Targets so identified confirmed previously amplified genes in the non-mevalonate pathway (*dxs* and *idi*) and some regulatory genes (*rpoS* and *appY*). Additionally, this method revealed novel gene targets (*yjiD*, *ycgW*, *yhbL*, *purDH*, and *yggT*). A two-dimensional search was subsequently undertaken, whereby the selected overexpression targets were combined with the knock-out targets predicted by stoichiometric modeling. All combinations of single (*rpoS, appY, yjiD, ycgW,* and *yhbL*), double (*yjiD-ycgW*) and triple (*yjiD-ycgW-yhbL*) overexpressions with four gene deletion backgrounds, including single (*ΔgdhA*, or *ΔaceE*), double (*ΔgdhAΔaceE*), and triple (*ΔgdhAΔaceEΔfdhF*) knockouts, were constructed and evaluated for lycopene production. Investigation of the metabolic landscape spanned by these 40 strains identified the best-engineered strain ($T5_P$*-dxs,* $T5_P$*-idi,* $rrnB_P$*-yjiD-ycgW, ΔgdhΔaceEΔfdhF*, pACLYC), which accumulated 16,000 ppm (16 mg/g cell) of lycopene within 24 h in a batch shake flask with 5 g/L of glucose in M9 minimal medium.

Keywords: Gene targeting; Lycopene; Combinational search

1. Introduction

Improvement of metabolic phenotype through directed genetic modifications is the main goal of metabolic engineering (Bailey, 1991; Stephanopoulos and Sinskey, 1993; Stephanopoulos, 1999; Stephanopoulos and Kelleher, 2001). Recent progress in molecular genetics methods makes it possible to knock-out or overexpress targeted genes in most microorganisms (Baudin et al., 1993; Mumberg et al., 1994; Mumberg et al., 1995; Lutz and Bujard, 1997; Datsenko and Wanner, 2000; Voss and Steinbuchel, 2006). Therefore, a key question in metabolic engineering is how to identify gene targets that have direct or indirect impact on a particular phenotype of interest.

Rational or hypothesis-driven methods typically focus on the improvement of a particular biosynthetic capacity through engineering at the pathway level (Stephanopoulos and Simpson, 1997; Koffas et al., 2003). The resulting phenotypes are often suboptimal due to distant effects of the genetic modifications or unknown regulatory interactions. In order to address these problems, two different approaches have been proposed: *systematic* and *combinatorial* gene target searches. In the systematic approach, global

*Corresponding author. Department of Chemical Engineering, Massachusetts Institute of Technology, Room 56-469, Cambridge, MA 02139. Fax: +617 253 3122.
E-mail address: gregstep@mit.edu (G. Stephanopoulos).

1096-7176/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2007.03.003

stoichiometric modeling is employed to analyze exclusively stoichiometric interactions and identify knock-out targets that can enhance product synthesis through increased precursor supply (Edwards et al., 2001; Segre et al., 2002; Burgard et al., 2003; Stephanopoulos et al., 2004; Shlomi et al., 2005; Alper et al., 2005b). As stoichiometric models cannot capture complex kinetic and regulatory interactions, combinatorial approaches, implemented by traceable genetic perturbation methods and high-throughput screening, are more amenable for the identification of both regulatory and metabolic genes. Reliable high-throughput screening process is a prerequisite of the combinatorial approach (Stephanopoulos, 2002; Tyo et al., 2006). There is also a key drawback which applies to both systematic and combinatorial approaches: these approaches are less effective for exploring mutant phenotype spaces when compared to random chemical mutagenesis.

While successful in improving product synthesis, the above methods are limited in the sense that they probe the sequence space along a pre-specified, sequential, gene-by-gene trajectory (Hwang et al., 2002; Misra et al., 2002; Gill et al., 2002a, b). As such, there is no assurance that such searches can reach a global phenotypic maximum. In order to expand the sequence space over which such gene target searches are conducted, we undertook a hybrid approach which combines systematic and combinatorial approaches for the identification of gene targets in the context of lycopene production in *Escherichia coli*.

Lycopene is a potential nutraceutical with anti-cancer and anti-oxidative properties (Dimascio et al., 1989; Giovannucci et al., 1995) that is abundant in plants and vegetables, especially in tomato (Giovannucci, 1999). Wild-type *E. coli* does not have a lycopene biosynthesis pathway. Therefore introduction of the lycopene biosynthesis operon (*crtEBI*) from *Erwinia euridova* (Misawa and Shimada, 1998; Sandmann et al., 1999; Lee and Schmidt-Dannert, 2002) was required to produce lycopene in *E. coli.* The resulting strains were able to accumulate lycopene in the cell but the levels of lycopene production were not significant. Previous efforts employing overexpression of target genes for enhancing lycopene production can be summarized by three approaches: (1) the improvement of metabolic capacity of the isoprenoids biosynthesis pathway by the overexpression of *dxs* and *idi* (Matthews and Wurtzel, 2000; Kim and Keasling, 2001; Sandmann et al., 2006), (2) the appropriate precursors supply through the reconfiguration of the central metabolic pathway (Farmer and Liao, 2000; Farmer and Liao, 2001), (3) the overexpression of global regulatory proteins (*rpoS*, *appY*, *ycgW*(*elbA*), and *yhbL*(*elbB*) that elicit lycopene production in unknown manners (Hemmi et al., 1998; Kang et al., 2005). Additionally, combinations of the knock-out targets (*gdhA*, *aceE*, and *fdhF*) identified through the stoichiometric modeling approach increased lycopene production significantly in the background of *dxs* and *idi* overexpression (Alper et al., 2005b). We also note that bioreactor operation can impact in a significant way

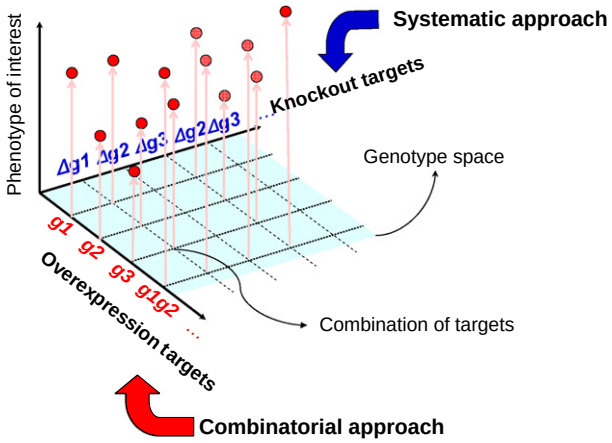

Fig. 1. Systematic diagram of a two-dimensional target identification method consisting of overexpression and knock-out targets search.

product accumulation (Kiss and Stephanopoulos, 1991; Kiss and Stephanopoulos, 1992; Alper et al., 2006).

In order to identify additional gene targets and investigate the interactions among the identified targets, we performed a multi-dimensional search consisting of combinations of overexpression targets and knock-out targets (Fig. 1). By combining targets identified from various searches, we were able to construct an engineered *E. coli* strain which exhibited four-fold increase in lycopene accumulation as compared to the parental strain (WS140).

2. Materials and methods

2.1. Microbial strains, primers, plasmids, and culture conditions

E. coli DH5α, K12, and WS140 (Yuan et al., 2006) strains (Table 1) were used for the overexpression target search with pACLYC plasmid (Cunningham et al., 1994) containing the *crtEBI* operon. Molecular cloning and manipulation of plasmids were done with the *E. coli* DH5α strain. Primers used in PCR are listed in Table 2. All *E. coli* strains were grown at 37 °C in a shaking incubator (225 rpm). Initial evaluation of lycopene accumulation by engineered strains was performed in 50 ml of M9 minimal media with appropriate antibiotics in a 250-ml flask. For the further characterization of lycopene production by selected strains, 100 ml of 2XM9 medium containing double amount of salts in 1X media except for $CaCl_2$ and $MgSO_4$ in a 250-ml flask was used. All fermentation experiments were initiated with a 1% (vol/vol) inoculation from an overnight 5-ml culture.

2.2. Genomic library construction and screening of libraries

The isolated genomic DNA from the *E. coli* K12 strain was partially digested with *Tsp*509 I. DNA fragments with different sizes were rescued from agarose gels and ligated into the *Eco*RI site of the pZE plasmid which was already

dephorphorylated in both ends. Constructed libraries were transformed into *E. coli* DH5α, K12, WS140, and related strains, which were already harboring the pACLYC plasmid in order to identify the putative overexpression targets for improved lycopene production. Transformants were plated into LB or M9 medium plus agar plates. High lycopene accumulating transformants were visually screened. Typically the agar plates were incubated at 37 °C overnight and kept at room temperature for 48 h for red-color development.

Table 1
Strains and plasmids used in this study

Name	Description	Reference
Strains		
WS140	*dxs, idi*, and *ispFD* genes are overexpressed through replacement of native promoter with a strong one ($_P$T5)	(Yuan et al., 2006)
WS140 ΔgdhA	Deletion of *gdhA* gene	(Alper et al., 2005b)
WS140 ΔaceE	Deletion of *aceE* gene	(Alper et al., 2005b)
WS140 ΔgpmB	Deletion of *gpmB* gene	(Alper et al., 2005b)
WS140 ΔgdhFΔaceE	Deletion of *gdhA* and *aceE* genes	(Alper et al., 2005b)
WS140 ΔgdhFΔgpmB	Deletion of *gdhA* and *gpmB* genes	(Alper et al., 2005b)
WS140 ΔgdhFΔaceEΔfdhF	Deletion of *gdhA, aceE*, and *fdhF* genes	(Alper et al., 2005b)
Plasmids		
pACLYC	pAC184 derivative containing crtEBI operon	(Cunningham et al., 1994)
pKD46	Lamda recombinase	(Datsenko and Wanner, 2000)
pZE	A medium copy plasmid	(Lutz and Bujard, 1997)
pUC19	A high copy plasmid	Invitrogen

2.3. Insert characterization and confirmation

We excluded the chance that the transformants accumulated high lycopene because of unknown chromosomal mutations by checking the consistency of lycopene production through the retransformation of the isolated plasmids into new competent cells with the pACLYC plasmid. The insert sequences were sequenced by two primers adjacent to the insert. By BLASTing the sequence against *E. coli* genome sequence, putative open reading frames inside the inserts were identified. If multiple open frames existed in the insert, we did a selective gene disruption analysis to identify the acting open reading frame impacting lycopene production. The selective gene disruption was performed by digesting the open reading frame with a unique restriction enzyme and subsequent re-ligation of the plasmid after removing the overhanging sequence. We also confirmed the effect of overexpression of the identified target genes through a "clean" overexpression experiment. For this purpose, we did a PCR amplification of putative ORF with primers (Table 2) containing 5′-*Kpn*I and 3′-*Mlu*I. The amplified insert ligated into the *Kpn*I-*Mlu*I site of pZE vectors (Lutz and Bujard, 1997) in order to place the insert under the control of strong promoters (P_L-TetO variants).

Table 2
Primers and their sequences used in this study

Name	Description	Sequence (5′→3′)
Jin112	*Kpn*I-*appY*	CGGGTACCGCTGTCAATCATTTGATTTAAT
Jin113	*Mlu*I-*appY*	AGCACGCGTGGGCTACGACATTGATAGAATA
Jin140	*Kpn*I-*yjiD*	GCCGGTACCGTGCGCAAAATGATGCGACA
Jin141	*Mlu*I-*yjiD*	GCCACGCGTTTAGCTGACATTCTCCAGCG
Jin143	*Kpn*I-*rpoS*	GTGGTACCGCATGTTCCGTCAAGGGATCAC
Jin144	*Mlu*I-*rpoS*	GCCACGCGTTTACTCGCGGAACAGCGCTT
Jin162	*Kpn*I-*yhbL*	GTGGTACCGCATGATCACAATGAAGAAAAT
Jin163	*Mlu*I-*yhbL*	GCACGCGTTATCATTCAGCCAGAACCAGCA
Jin164	*Kpn*I-*ycgW*	GTGGTACCGCATGAAGTGGATAGTAATTGA
Jin165	*Mlu*I-*ycgW*	GCACGCGTTAGCATATCGAGCATATTTATA
Jin168	*Eco*RI-*yjiD*	GGCGAATTCTTAGCTGACATTCTCCAGCG
Jin169	*Eco*RI-RBS-*ycgW*	GGGAATTCAGGAGGACAGCTATGAAGTGGATAGTAATTGA
Jin170	*Hind*III-*ycgW*	GCAAGCTTTAGCATATCGAGCATATTTATA
Jin171	*Hind*III-RBS-*yhbL*	GGAAGCTTAGGAGGACAGCTATGATCACAATGAAGAAAAT
Jin178	*lacZ*-5′UTR-Kan	TATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTGGCAACCGAGCGTTCTGAAC
Jin179	*lacZ*-3′UTR-mluI	ATGGATTTCCTTACGCGAAATACGGGCAGACATGGCCTGCCCGGTTATTAGAGCGTTCACCGACAAACAA
Jin185	RT-PCR-5-*yjiD*	CATTACCGTCATTACTTGTTTCAGG
Jin186	RT-PCR-3-*yjiD*	GCTCATGAAATAACAGTGCAGAAC
Jin189	RT-PCR-5-*yhbL*	TAACTGGCGAAGCGATGACGGAAAC
Jin190	RT-PCR-3-*yhbL*	GCACAATCAACGCATCCAGTTCAGC
Jin191	RT-PCR-5-*ycgW*	TCAATCTACTATATTTCTCCCTCCTGGCAG
Jin192	RT-PCR-3-*ycgW*	CCTTGTTGAATGGTGCGATGTGATAAA

2.4. Integration of an overexpression cassette into E. coli chromosome

In order to bypass the detrimental effect from the presence of multi-copy plasmids on lycopene production, the targets genes under the control of a strong promoter ($rrnB_P$) were delivered into the locus of *LacZ* in *E. coli* chromosome by a λ recombinase-mediated gene replacement method (Datsenko and Wanner, 2000). First, single gene overexpression cassettes and synthetic operons were constructed. By designing primers containing a unique restriction enzyme site and ribosomal binding sites, we created two synthetic operons: *yjiD-ycgW* and *yjiD-ycgW-yhbL*. The following four DNA fragments, *KpnI-yjiD-EcoR*I, *EcoR*I-RBS-*ycgW-Mlu*I, *EcoR*I-RBS-*ycgW-Hind*III, and *Hind*III-RBS-*yhbL-Mlu*I, were prepared and ligated into the pZE plasmids to create the abovementioned two synthetic operons under the control of a *rrnB* promoter. Two primers (Jin178 and Jin179) were used to generate linear fragments containing overexpression target genes with *rrnB* promoters which are surrounded by *LacZ* flanking regions. Blue/white screening after transformation of the linear fragments served as an initial screening method of the correct replacement between the overexpression cassette and *LacZ* gene. We also confirmed the correct replacement by the colony PCR method. Once the correct gene replacement was established in one strain background, we transferred it into the other strain background using P1 phage transduction.

2.5. Lycopene assay

Accumulated lycopene in *E. coli* was extracted from 1 ml of culture at 24 h. The cell pellet was washed with water and then extracted in 1 ml of acetone at 55 °C for 15 min with intermittent vortexing every 5 min. The amount of lycopene in the supernatant was quantified by the absorbance at 475 nm using a standard curve of pure lycopene (Sigma). Cell mass was calculated by correlating dry cell weight with OD600 for use in ppm (mg lycopene/g dry cell weight) calculations.

3. Results

3.1. Combinatorial search for overexpression targets through genomic library screening

We screened *E. coli* genomic DNA fragments enhancing, upon amplification, lycopene production using *E. coli* genomic libraries on multi-copy plasmids. Genomic DNA fragments in different sizes (0.5–1, 1–3, and 3–8 kbp) were isolated after a partial digestion of *E. coli* genomic DNA and subsequently ligated into a medium-copy (pZE21) and high-copy (pUC19) number plasmid. The constructed shotgun libraries were transformed into the *E. coli* DH5α strain containing pACLYC plasmid to screen the DNA fragments which improve lycopene production. Red color formation becomes more intense in *E. coli* colonies accumulating large amounts of lycopene intracellularly, hence we screened for colonies exhibiting intense red color compared to the background colonies in the plate. More than 30,000 transformants with each library were screened. The transformed plasmids were then isolated and retransformed into *E. coli* DH5α (pACLYC) strain to test whether the inserts in the isolated plasmids were responsible for the formation of the intense red color. Following confirmation, we identified insert sequences by comparing the sequencing results of the 5′ and 3′ regions of the inserts against the *E. coli* K12 genome sequence. Depending on the size of inserts and the copy number of the plasmid used for the library construction, we identified different inserts with some consensus from the screening (Table 3). By using *E. coli* DH5α (pACLYC) strain as a parental strain of the transformation of the libraries, we identified a total of 37 inserts which represent 16 independent regions of the *E. coli* genome from the screening with the libraries

Table 3
Summary of the identified inserts from the screening based on color development in the plate

Strains	Copy number of plasmids used for the library construction	Average size of inserts (kbp)	Medium used in the plate screening	Identified inserts (frequency/number of sequenced plasmids)
DH5α	Medium	5	LB	*dxs* (13/19), *idi* (4/19), *rpoS* (1/19), and *torC* (1/19)
DH5α	Medium	2	LB	*rpoS* (2/8), *appY* (1/8), *ydgK* (2/8), *dxs* (1/8), *yeiA* (1/8), and *yedR* (1/8)
DH5α	High	5	LB	*idi* (4/10), *yjiD* (2/10), *dxs* (1/10), *rpoS* (1/10), *torT* (1/10), and *yhbL* (1/10)
WS140	Medium	5	LB	*rpoS* (3/3)
WS140	Medium	5	M9	*rpoS* (4/7), *yggT* (1/7), *purDH* (1/7), and *yfjN* (1/7)
WS140 *ΔgdhA*, *ΔaceE*, and *ΔfdhF*	Medium	5	M9	*rpoS* (6/8), *yjiD* (1/8), and *ycgW* (1/8)

containing 1–3 and 3–8 kbp inserts. Although we performed a screening with the library containing 0.5–1.0 kbp inserts, we could not identify any colony exhibiting an intense red-color as compared to the control strain. Open reading frame searches in these identified inserts revealed that multiple inserts contained *dxs*, *idi*, and *rpoS* genes. Of the 37 isolated plasmids, thirteen contained *dxs*, eight contained *idi*, and four contained *rpoS*. The *dxs* was a major ORF (13/19) identified from screening with the library based on a medium-copy plasmid, whereas the *idi* was a major ORF (4/10) identified screening with the library based on a high-copy plasmid (Table 3). The *rpoS* gene was consistently found (1/19, 2/8, and 1/10) from the searches regardless of the insert sizes and plasmids' copy numbers. We were also able to identify a global regulator (*appY*), oxido-reductases (*torC*, *ydgK*, and *yeiA*), and hypothetical proteins (*yedR*, and *yhbL*) from the inserts (Table 3).

3.2. Sequential search for overexpression targets for lycopene production

To find multiple gene targets, we performed a sequential-iterative search where additional gene targets are identified after implementing genetic modifications suggested from the previous search. Since *dxs* and *idi* emerged as the most common targets from the previous screening with DH5α, we used an *E. coli* strain (WS140) which exhibit higher expression levels of *dxs* and *idi* gene through the replacement of native promoters with a strong promoter ($T5_P$) (Fig. 2). By using WS140 (pACLYC) strain as a host strain, we identified four inserts which improve lycopene production. Interestingly, all four inserts contained *rpoS* gene which was indeed pointed out as a third target in addition to *dxs* and *idi* from the previous screening. We also tried this sequential-iterative search in different environments by changing medium. Unlike the previous screenings, we used agar plates containing M9 minimal medium with 5 g/L of glucose instead of LB plates. The *rpoS* gene was still the dominant target (4/7) as in the LB medium, but new gene targets (*yggT*, *purDH*, and *yfjN*) now emerged. We further extended this sequential search using the strain with knockouts of three genes (*gdhA*, *aceE*, and *fhdhF*) as a host strain for an overexpression target search. The *rpoS* (6/8) was again the consensus target. *yjiD* (1/8) and *ycgW* (1/8) were also identified as overexpression targets in the triple knock-outs background (Fig. 2).

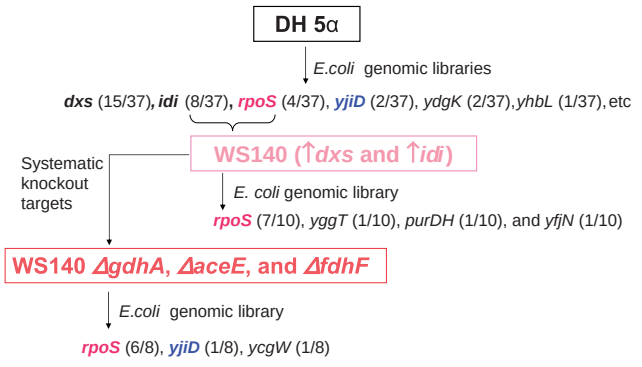

Fig. 2. Identification of overexpression targets impacting lycopene production under different strain backgrounds.

3.3. Validation of new targets by transformation of the isolated plasmids

Of the isolated plasmids from the screenings, plasmids containing *rpoS*, *torC*, *ydgK*, *appY*, *yggT*, *purDH*, *yfjN* in medium-copy plasmids, and plasmids containing *yjiD* and *ycgW* in high-copy plasmids were transformed into the WS140 strain with the pACLYC plasmid for a quantitative measurement of the improvement in lycopene production by each insert. Lycopene accumulation in 50 ml of M9 medium with 5 g/L of glucose in a 250 ml flask was measured after 24 h (Fig. 3). All tested plasmids containing inserts accumulated higher amounts of lycopene as compared to the control strain containing empty plasmids. The medium-copy plasmids containing *rpoS*, *appY*, *yggt*, *purDH*, and *yfjN* improved lycopene production about three fold (6000–9000 vs. 2400 ppm), but the effect of plasmids containing *torC* and *ydgK* on lycopene production was more moderate (around 3500 ppm). The high-copy plasmids containing *yjiD* and *ycgW* also improved lycopene production drastically compared to the control strain (3500–4500 vs 150 ppm). This result is consistent with a previous report that low copy plasmids are better than high-copy plasmids (Jones et al., 2000). One notable observation was that lycopene production was severely hindered by the transformation of multi-copy plasmids, especially with the high-copy plasmid. WS140 (pACLYC) stain without additional plasmid accumulated about 4300 ppm of lycopene in M9 medium with 5 g/L of glucose, but the control strain containing a medium-copy plasmid without inserts produced only 2100 ppm of lycopene under the same conditions. The control strain containing a high-copy plasmid accumulated only 150 ppm of lycopene.

3.4. Combination of overexpression targets with knock-out targets

In order to probe a new dimension in the lycopene metabolic landscape, the newly identified overexpression targets (*rpoS*, *appY*, *torC*, *ydgK*, *yggT*, *purDH*, *yfjN*, and *yjiD*) (Table 4) were combined with the single, double, and triple knockouts, which were previously identified by a systematic approach (Alper et al., 2005b). Of particular interest was the *yjiD* locus as its promoter had also been identified as a knock-out target previously (Alper et al., 2005c). Since *yjiD* was originally identified in a high-copy number plasmid, the open reading frame of *yjiD* was cloned into a medium-copy vector under the control of a variant of P_L lambda promoter (Alper et al., 2005a). Including seven strains with the control plasmid, a total of 63 transformants (7 knock-out strains × 9 plasmids) were

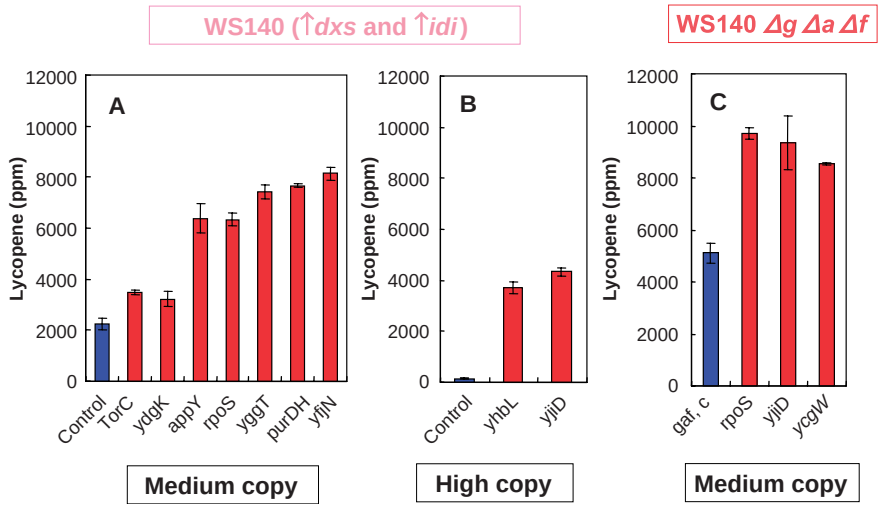

Fig. 3. Improvement of lycopene accumulation after overexpression of identified targets in M9 medium with 5 g/L of glucose: (A) Targets identified using a medium-copy plasmid. (B) Targets identified using a high-copy plasmid. (C) Targets identified using a medium-copy plasmid under a triple knock-out background.

Table 4
Known and putative function of the identified genes in this study

Gene name	b-number	Known or putative function[a] of the gene product
dxs	b0420	1-deoxy-D-xylulose 5-phosphate (DXP) synthase
idi	b2889	Isopentenyl diphosphate isomerase
rpoS	b2741	RNA polymerase sigma 38
torC	b0996	Trimethylamine N-oxide reductase, cytochrome *c*-type subunit
appY	b0564	Regulatory protein affecting *appA* and other genes
ydgk	b1626	Putative oxidoreductase
yeiA	b2147	Putative dihydrothymine dehydrogenase
yedR	b1963	Hypothetical protein
yjiD	b4326	Hypothetical protein
torT	b0994	Inducer-binding protein in *TorSR* two-component signal transduction system
yhbL	b3210	Sigma cross-reacting protein 27A, may be involved in the early steps of isoprenoid biosynthesis (*elbB*)
yggT	b2952	Putative resistance protein, hypothetical integral membrane protein
purDH	b4005/6	Phosphoribosylamine-glycine ligase and AICAR transformylase/IMP cyclohydrolase
yfjN	b2630	Putative cell division protein
ycgW	b1160	Orf, hypothetical protein, may be involved in the early steps of isoprenoid biosynthesis (*elbA*)

[a]Putative function predictions are based on EcoCyc (Encyclopedia of *Escherichia coli* K12 Genes and Metabolism, http://ecocyc.org).

constructed. By measuring lycopene production of the 63 strains in M9 minimal medium with 5 g/L of glucose, a two-dimensional search combining overexpression and knock-out targets was performed. The metabolic landscape is shown in Fig. 4. The features of the metabolic landscape can be summarized as follows: (1) most of the overexpression targets, except for *ydgK* and *torC*, improved lycopene production regardless of knock-out backgrounds; (2) the responsiveness of each knock-out strain to the overexpression targets was fairly similar; (3) a global maximum of lycopene production (9000 ppm) was found in the combination of overexpression of *yjiD* and *ΔgdhAΔaceE ΔfdhF*. This result suggests that performance of most of the knock-out strains must have been limited by regulatory or metabolic barriers, although they are designed optimally for lycopene production at the stoichiometric level.

3.5. Improvement of lycopene production in enriched media

We next investigated interactions between environmental conditions and gene overexpression, focusing on *rpoS* and *yjiD* (Table 5). We selected 2X M9 medium which has been shown to be more effective for lycopene production as compared to 1X M9 medium (Alper et al., 2005c). Overexpression of *yjiD* increased lycopene production about 70–100% in all seven of our tested strains, each with different genetic backgrounds, as compared to the control. Overexpression of *yjiD* drastically improved lycopene production in the double (*ΔgdhA ΔgpmB* and *ΔgdhA ΔaceE*) and triple mutants (*ΔgdhA*, *ΔaceE*, and *ΔfdhF*) as compared to the single mutants (*ΔaceE* and *ΔgpmB*). Overexpression of the *rpoS* gene also resulted in significant increases in lycopene accumulation in all seven strains. However, the improvement of lycopene production

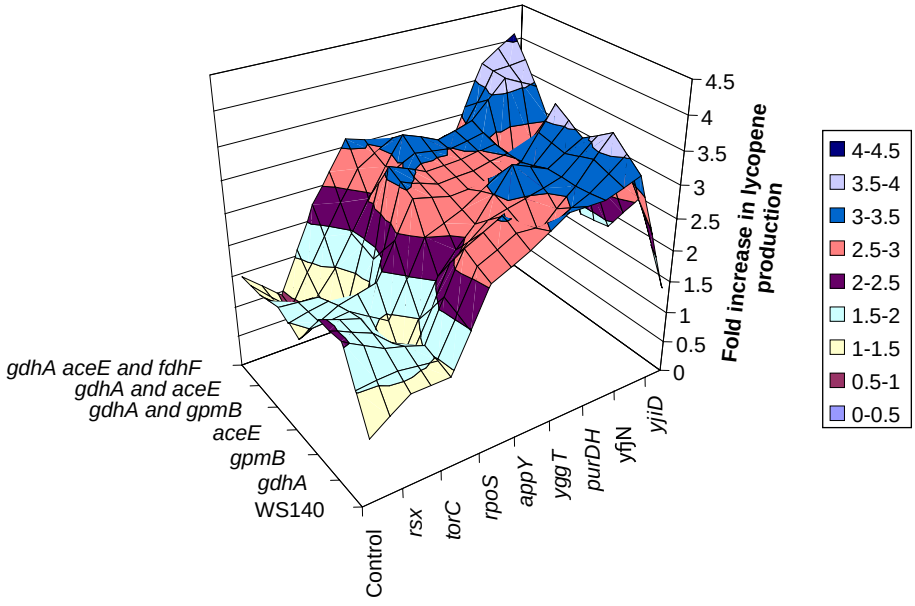

Fig. 4. The metabolic landscape of lycopene production by combination of the isolated overexpression targets with the knock-out strains which are constructed based on stoichiometric modeling. The open reading frame of *yjiD* in a high-copy plasmid was transferred into a medium-copy number plasmid under the control of a variant of P_L lambda promoter in the M9 medium with 5 g/L of glucose.

Table 5
Improvement of lycopene accumulation after overexpression of gene targets in the 2X M9 medium

Strains	Control	*rpoS*	*yjiD*
WS140	4290±391	10812±156	11135±840
ΔgdhA	7650±177	11810±64	12794±646
ΔaceE	6972±591	10297±17	10301±129
ΔgpmB	8669±318	10608±445	11253±649
ΔgdhAΔaceE	8861±686	12173±113	13513±94
ΔgdhAΔgpmB	8789±79	10123±207	13621±678
ΔgdhAΔaceEΔfdhF	8578±203	11755±499	14442±269

was rather uniform regardless of strain backgrounds. In general, lycopene accumulation in the 2X medium was significantly improved as compared to M9 medium.

3.6. Overexpression of the selected gene targets through stable integration

To bypass the inhibitory effect of multi-copy plasmids, we integrated overexpression cassettes into the *E. coli* genome. A strong promoter (*rrnB*) was chosen to ensure overexpressions of the target genes even with a single copy integration of the expression cassette. A subset of genes (*rpoS*, *appY*, *ycgW*, *yhbL*, and *yjiD*) was selected based on previous results. These genes were integrated into different strains containing single, double, and triple knockouts. We also overexpressed multiple genes by constructing synthetic operons contains *yjiD-ycgW* and *yjiD-ycgW-yhbL*. RT-PCR experiments with specific primers to *yjiD*, *ycgW*, and *yhbL* were performed to monitor the expression of the genes in the synthetic operon in two different strain backgrounds (Fig. 5). The expression patterns of *yjiD*, *ycgW*, and *yhbL* were well in accordance with the intended design of the synthetic operons.

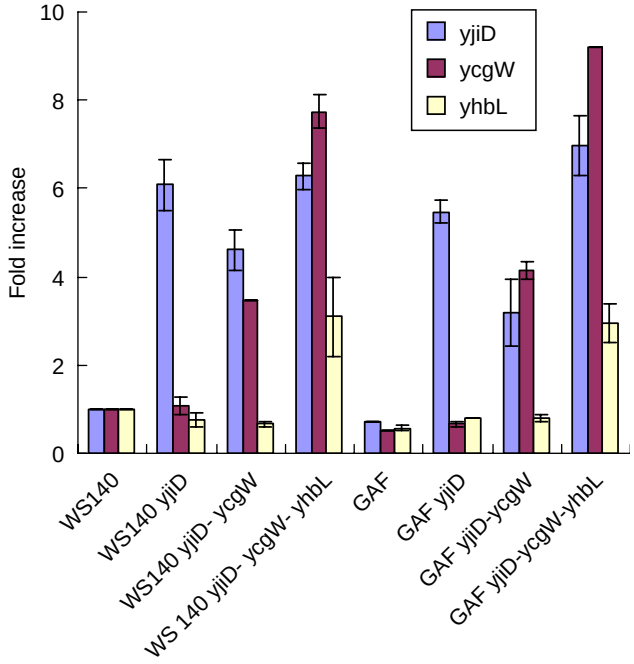

Fig. 5. Differential expression levels of *yjiD*, *ycgW*, and *yhbL* in the engineered strains through the integration of the overexpression cassettes in the 2XM9 medium with 5 g/L of glucose.

3.7. Mapping of the 2-D genotype and phenotype landscape

We constructed a set of 40 strains containing exhaustive combinations of the four knockouts (*ΔgdhA*, *ΔaceE*,

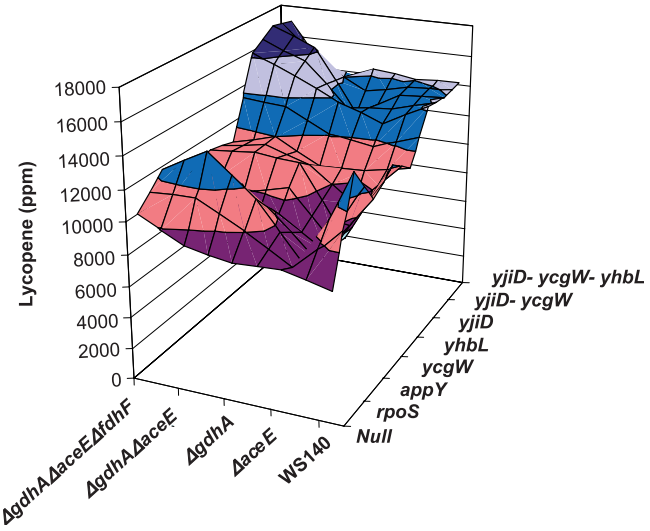

Fig. 6. Lycopene production by forty strains containing exhaustive combinations of the four knockouts (*ΔgdhA, ΔaceE*, *ΔgdhAΔaceE*, and *ΔgdhAΔaceEΔfdhF*) and seven overexpressions (*rpoS, appY, yjiD, ycgW, yhbL*, *yjiD-ycgW*, and *yjiD-ycgW-yhbL*) through the integration of the overexpression cassette in the 2X M9 medium with 5 g/L of glucose.

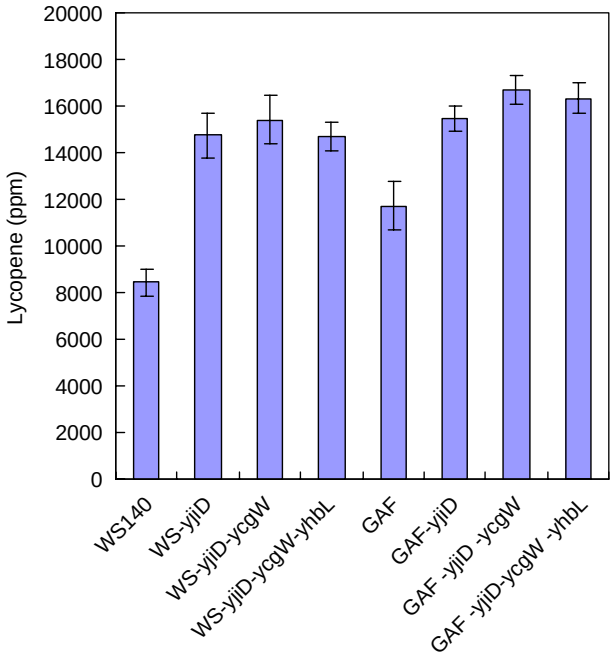

Fig. 7. Lycopene production by the selected strains in the 2X M9 medium. Error bar represent standard deviation of triple replicates.

ΔgdhAΔaceE, and *ΔgdhAΔaceEΔfdhF*) and seven overexpressions (*rpoS, appY, yjiD, ycgW, yhbL*, *yjiD-ycgW*, and *yjiD-ycgW-yhbL*) through integration to exclude the inhibitory effect. Also, we used 2X M9 medium with 5 g/L of glucose for evaluating the performances of the 40 strains. Fig. 6 shows the amounts of lycopene accumulation by the strains within 24 h. In general, these integrated strains accumulated more lycopene as compared to the strains containing the same inserts in plasmids for overexpression. Again, *yjiD* was noticeable as a potential overexpression target for lycopene accumulation. A global maximum (16,000 ppm) in lycopene production was observed from the strain containing a triple knockout (*ΔgdhAΔaceEΔfdhF*) and double overexpression (*yjiD-ycgW*). We also performed shaker flask experiments in triplicates in order to evaluate reproducibility of the lycopene production by the eight selected strains (Fig. 7). Simultaneous overexpression of *yjiD* and *ycgW* was better for lycopene production than *yjiD* overexpression in two strain backgrounds.

4. Discussion

Previous target identification methods for metabolic engineering mostly relied on one-dimensional searches whereby only one type of perturbations was probed for the improvement of a metabolic phenotype. For instance, Kang et al. (Kang et al., 2005) used a shotgun library screening method for the identification of overexpression targets which enhance lycopene production in *E. coli*. Genes of *dxs*, *appY*, *crl*, and *rpoS* were confirmed to be enhancing lycopene production upon overexpression. Alper et al. (2005c) employed systematic (stoichiometric model based) and combinatorial (random transposon insertion based) approaches for the identification of knock-out targets which improve lycopene production in *E. coli*. A recent study by Alper et al. (2005c) demonstrated a way of partially overcoming the problem due to the capability of exhaustive search through genetic perturbations by combining knock-out targets identified by different approaches. The resulting strain containing the combined knockouts showed two-fold improvement in lycopene accumulation. In this study, we attempted to expand the search space further using a two-dimensional search consisting of overexpression and knock-out target searches based on orthogonal strategies. Therefore, we were able to explore metabolic phenotypes with another level of complexity. The two-dimensional search yielded an engineered *E. coli* strain which is capable of accumulating more than 16,000 ppm of lycopene, higher than lycopene accumulation levels (10,000–13,000 ppm) by the engineered strains harboring only one type of perturbation.

One of the open questions in the quest of genetic targets for metabolic engineering is how *multiple* targets are identified. In principle, there should be a set of genes conferring the optimal phenotype. However, such optimal set cannot be experimentally discovered as it is impossible to construct and evaluate all possible combinations of these genes. Therefore, a proper search strategy is required to identify multiple targets. We have proposed a sequential search method where additional gene targets are explored after implementation of the identified targets from previous rounds of the search (Jin et al., 2005; Alper et al., 2005b). Using the sequential search method, we have identified triple knock-out targets for the improvement of lycopene production in *E. coli* (Alper et al., 2005b) and double

overexpression targets for the improvement of growth and ethanol production from xylose in recombinant *S. cerevisiae* (Jin et al., 2005). In this study, we used the same strategy for the identification of multiple overexpression targets for the improvement of lycopene accumulation in *E. coli*. From the initial screening of overexpression targets using *E. coli* DH5α, *dxs* and *idi* genes were identified as potential limiting enzymes in the biosynthesis of lycopene. Indeed overexpression of the *dxs* and *idi* gene were previously reported to enhance the production of lycopene in *E. coli* (Matthews and Wurtzel, 2000; Kim and Keasling, 2001). Therefore, we performed another round of search using the WS140 strain, which is engineered to overexpress *dxs* and *idi* genes. As expected, new overexpression targets, such as *rpoS*, *yggT* and *purDH* genes, were identified, while *dxs* and *idi* genes were not screened any more. In yet another round of the sequential search, we screened overexpression targets under the background of triple knockouts (Fig. 2). Interestingly, the combination of two identified targets (*yjiD* and *ycgW*) from the search resulted in an engineered strain showing the best performance for lycopene accumulation.

Of the 15 identified overexpression targets (Table 4) *dxs*, *idi*, *torC*, and *purDH* are four metabolic enzymes for which gene product functions are clearly known. As putative metabolic enzymes, *ydgK* and *yeiA* were identified. Known and putative regulatory genes, such as *rpoS*, *appY*, *torT*, *yjiD*, *ycgW*, and *yhbL*, were also identified from our screenings. Some of the targets, such as *dxs*, *idi, rpoS, appY, ycgW(elbA)*, and *yhbL(elbB)*, were previously mentioned as overexpression targets (Hemmi et al., 1998; Kim and Keasling, 2001; Lee and Schmidt-Dannert, 2002; Kang et al., 2005), but others, such as *torC*, *purDH, ydgK, yeiA, yjiD*, and *yfjN* have not been mentioned before. We also found there is an interesting overlap in the results between this overexpression target search and the knockout targets identified by transposon mutagenesis. We identified *yjiD* as an overexpression target, whereas Alper et al. (2005c) found that the insertion of a transposon element between the promoter region and the *yjiD* open reading frame (denoted as Δ_p*yjiD*) improved lycopene accumulation in *E. coli*. We performed an RT-PCR experiment to check the effect of Δ_p*yjiD* on the expression level of *yjiD* gene. Actually the transcript level of *yjiD* from the strain harboring Δ_p*yjiD* was significantly higher than that of the control strain. We also compared the mRNA levels of *yjiD* in *yjiD* overexpressing strain and Δ_p*yjiD* and found the level of *yjiD* transcript in Δ_p*yjiD* was lower than that of *yjiD* overexpressing strain (data not shown), which suggests the effect of *yjiD* overexpression on lycopene accumulation may not be gene dosage dependant. Still, these results are consistent with the finding of *yjiD* as an overexpression target. The putative function of the gene product of *yjiD* gene is not known yet. However, it is likely to be a non-metabolic enzyme because the length of the translated amino acid sequence is rather short (only 120 amino acids).

The best lycopene accumulating strain, WS140 (*$T5_P$-dxs, $T5_P$-idi, $rrnB_P$-yjiD-ycgW, ΔgdhΔaceEΔfdhF*) includes four overexpressions and three knockouts, which might exhibit three different putative mechanisms for improving lycopene production: push, pull, and global regulation favoring lycopene production. First, the strain contained three knockouts of metabolic enzymes in the central carbon metabolism, which might push precursors into the non-mevalonate pathway. Second, the strain had overexpressions of *dxs* and *idi*, which pull precursor metabolites from a central metabolic pathway into the non-mevalonate pathway. Finally, the strain also experiences an overexpression of the *yjiD* and *ycgW*, which presumably affects lycopene production through a regulatory manner.

We also found a significant overlap between our identified targets and the mutator genes identified through genomic library screening. Yang et al. (2004) searched genes which create mutator phenotypes when the putative genes are overexpressed using a genomic library in a multi-copy plasmid. Interestingly, the genes identified from their screening (*rpoS, appY, yjiD*, and *ycgW*) coincided with the identified targets in our study. These genes are classified as genes eliciting low levels of mutagenesis with Lac^+ revertant frequency close to the control in their report (Yang et al., 2004). This suggests that the main functions of these gene products might not be connected with mutational pathways directly but somehow the overexpresions of these genes affect the mutator phenotypes. How did two independent studies exploring totally different phenotypes identify the same gene targets? Although it is difficult to answer the question because we still do not know the function of the genes (*yjiD* and *ycgW*), we can speculate on possible answers from the experimental procedures employed in both studies. We incubated plates overnight for a colony formation at 37 °C and left plates at room temperature for two to three days until colonies became pigmented. They also incubated plates four days at 37 °C for the detection of mutator phenotypes. Both studies incubated plates for an extended period of time followed by a standard overnight incubation of plates for colony formation. This suggests that two assays (lycopene and the mutator phenotype) could be subject to the stationary metabolism of *E. coli*. Indeed, *rpoS* and *appY* are known to play important roles in the stationary or energy limited metabolism in bacteria (Atlung et al., 1997; Jiang et al., 2001). Therefore, we can speculate that the four gene targets identified from both studies could have direct or indirect connections with the stationary metabolism in *E. coli*, which is consistent with the results presented in this study since most of the lycopene accumulation in *E. coli* happens in the stationary phase.

A shotgun library in a multi-copy plasmid provides a convenient platform for the identification of overexpression targets. Although it still offers fast screenings of putative gene targets against intended phenotypes, the presence of multi-copy plasmids in cells could limit the capacity of targeted metabolite production because of

the metabolic burden caused by a high demand of carbon and energy sources for the synthesis of the multi-copy plasmids. We observed such detrimental effects of multi-copy plasmids on lycopene accumulation in *E. coli*. When a medium-copy plasmid was used, the amount of lycopene accumulation decreased about 40% (2500 ppm vs. 4300 ppm). The effect of a high-copy plasmid on lycopene accumulation was more drastic. Only 150 ppm of lycopene accumulation was observed when a high-copy plasmid was used. Although there were significant increases in lycopene production after introduction of target genes in multi-copy plasmids, integration of expression cassettes into the genome for target gene overexpression proved to be a much better way of archiving overexpression. The integration of the overexpression cassette resulted in a significant increase of lycopene accumulation as compared to the overexpression through multi-copy plasmids. One possible way of bypassing the problems with multi-copy plasmids is to use a promoter delivery method for overexpression target identification. For instance, we can insert a strong or inducible promoter inside a transposon element for the overexpression of adjacent genes after transposon insertion (Judson and Mekalanos, 2000). However, the capacity of the method will be very limited because of the presence of many operons in the bacterial genome.

In this study, we have demonstrated the improvement of lycopene production in *E. coli* using a strategy based on a two-dimensional search. However, the search could be extended into three or four dimensions by adding additional axis. For instance, knock down of knock-out gene targets, or more precise modulation of the over-expression targets could be included as a new axis. By adding another layer of complexity through the multi-dimensional search, we will be able to explore bigger mutant phenotype space and even understand the complex interactions which might have been undetectable in one-dimensional searches.

Acknowledgments

We would like to thank Dr. Wonchul Suh for providing the *E. coli* WS140 strain. This study was financially supported by the Dupont-MIT Alliance and the Singapore-MIT Alliance (SMA2).

References

Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., 2005a. Tuning genetic control through promoter engineering. Proc. Natl. Acad. Sci. USA 102, 12678–12683.

Alper, H., Jin, Y.S., Moxley, J.F., Stephanopoulos, G., 2005b. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. Metabol. Eng. 7, 155–164.

Alper, H., Miyaoku, K., Stephanopoulos, G., 2005c. Construction of lycopene-overproducing *Escherichia coli* strains by combining systematic and combinatorial gene knockout targets. Nat. Biotechnol. 23, 612–616.

Alper, H., Miyaoku, K., Stephanopoulos, G., 2006. Characterization of lycopene-overproducing *Escherichia coli* strains in high cell density fermentations. Appl. Microbiol. Biotechnol. 72, 968–974.

Atlung, T., Knudsen, K., Heerfordt, L., Brondsted, L., 1997. Effects of sigmaS and the transcriptional activator *AppY* on induction of the *Escherichia coli hya* and *cbdAB-appA* operons in response to carbon and phosphate starvation. J. Bacteriol. 179, 2141–2146.

Bailey, J.E., 1991. Toward a science of metabolic engineering. Science 252, 1668–1675.

Baudin, A., Ozierkalogeropoulos, O., Denouel, A., Lacroute, F., Cullin, C., 1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res. 21, 3329–3330.

Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. 84, 647–657.

Cunningham, F.X., Sun, Z.R., Chamovitz, D., Hirschberg, J., Gantt, E., 1994. Molecular-structure and enzymatic function of lycopene cyclase from the *Cyanobacterium synechococcus* sp. Strain Pcc7942. Plant Cell 6, 1107–1121.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA 97, 6640–6645.

Dimascio, P., Kaiser, S., Sies, H., 1989. Lycopene as the most efficient biological carotenoid singlet oxygen quencher. Arch. Biochem. Biophys. 274, 532–538.

Edwards, J.S., Ibarra, R.U., Palsson, B.O., 2001. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. Nat. Biotechnol. 19, 125–130.

Farmer, W.R., Liao, J.C., 2000. Improving lycopene production in *Escherichia coli* by engineering metabolic control. Nat. Biotechnol. 18, 533–537.

Farmer, W.R., Liao, J.C., 2001. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. Biotechnol. Prog. 17, 57–61.

Gill, R.T., Katsoulakis, E., Schmitt, W., Taroncher-Oldenburg, G., Misra, J., Stephanopoulos, G., 2002a. Genome-wide dynamic transcriptional profiling of the light-to-dark transition in *Synechocystis* sp. strain PCC 6803. J. Bacteriol. 184, 3671–3681.

Gill, R.T., Wildt, S., Yang, Y.T., Ziesman, S., Stephanopoulos, G., 2002b. Genome-wide screening for trait conferring genes using DNA microarrays. Proc. Natl. Acad. Sci. USA 99, 7033–7038.

Giovannucci, E., 1999. Tomatoes, tomato-based products, lycopene, and cancer: Review of the epidemiologic literature. J. Natl. Cancer Inst. 91, 317–331.

Giovannucci, E., Ascherio, A., Rimm, E.B., Stampfer, M.J., Colditz, G.A., Willett, W.C., 1995. Intake of carotenoids and retinol in relation to risk of prostate cancer. J. Natl. Cancer Inst. 87, 1767–1776.

Hemmi, H., Ohnuma, S., Nagaoka, K., Nishino, T., 1998. Identification of genes affecting lycopene formation in *Escherichia coli* transformed with carotenoid biosynthetic genes: Candidates for early genes in isoprenoid biosynthesis. J. Biochem. 123, 1088–1096.

Hwang, D., Schmitt, W.A., Stephanopoulos, G., Stephanopoulos, G., 2002. Determination of minimum sample size and discriminatory expression patterns in microarray data. Bioinformatics 18, 1184–1193.

Jiang, G.R., Nikolova, S., Clark, D.P., 2001. Regulation of the ldhA gene, encoding the fermentative lactate dehydrogenase of *Escherichia coli*. Microbiology 147, 2437–2446.

Jin, Y.S., Alper, H., Yang, Y.T., Stephanopoulos, G., 2005. Improvement of xylose uptake and ethanol production in recombinant *Saccharomyces cerevisiae* through an inverse metabolic engineering approach. Appl. Environ. Microbiol. 71, 8249–8256.

Jones, K.L., Kim, S.W., Keasling, J.D., 2000. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab. Eng. 2, 328–338.

Judson, N., Mekalanos, J.J., 2000. TnAraOut, A transposon-based approach to identify and characterize essential bacterial genes. Nat. Biotechnol. 18, 740–745.

Kang, M.J., Lee, Y.M., Yoon, S.H., Kim, J.H., Ock, S.W., Jung, K.H., Shin, Y.C., Keasling, J.D., Kim, S.W., 2005. Identification of genes affecting lycopene accumulation in *Escherichia coli* using a shot-gun method. Biotechnol. Bioeng. 91, 636–642.

Kim, S.W., Keasling, J.D., 2001. Metabolic engineering of the non-mevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. Biotechnol. Bioeng. 72, 408–415.

Kiss, R.D., Stephanopoulos, G., 1991. Metabolic activity control of L-lysine fermentation by restrained growth fed-batch strategies. Biotechnol. Prog. 7, 501–509.

Kiss, R.D., Stephanopoulos, G., 1992. Metabolic characterization of a L-lysine producing strain by continuous culture. Biotechnol. Bioeng. 39, 565–574.

Koffas, M., Jung, G.Y., Stephanopoulos, G., 2003. Engineering metabolism and product formation in *Corynebacterium glutamicum* by coordinated gene overexpression. Metabol. Eng. 5, 32–41.

Lee, P.C., Schmidt-Dannert, C., 2002. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. Appl. Microbiol. Biotechnol. 60, 1–11.

Lutz, R., Bujard, H., 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. Nucleic Acids Res. 25, 1203–1210.

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

Misawa, N., Shimada, H., 1998. Metabolic engineering for the production of carotenoids in non-carotenogenic bacteria and yeasts. J. Biotechnol. 59, 169–181.

Misra, J., Schmitt, W., Hwang, D., Hsiao, L.L., Gullans, S., Stephanopoulos, G., Stephanopoulos, G., 2002. Interactive exploration of microarray gene expression patterns in a reduced dimensional space. Genome Res. 12, 1112–1120.

Mumberg, D., Muller, R., Funk, M., 1994. Regulatable promoters of *Saccharomyces cerevisiae*—comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res. 22, 5767–5768.

Mumberg, D., Muller, R., Funk, M., 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene 156, 119–122.

Sandmann, G., Albrecht, M., Schnurr, G., Knorzer, O., Boger, P., 1999. The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli*. Trend Biotechnol. 17, 233–237.

Sandmann, G., Romer, S., Fraser, P.D., 2006. Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. Metabol. Eng. 8, 291.

Segre, D., Vitkup, D., Church, G.M., 2002. Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. USA 99, 15112–15117.

Shlomi, T., Berkman, O., Ruppin, E., 2005. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proc. Natl. Acad. Sci. USA 102, 7695–7700.

Stephanopoulos, G., 1999. Metabolic fluxes and metabolic engineering. Metabol. Eng. 1, 1–11.

Stephanopoulos, G., 2002. Metabolic engineering by genome shuffling. Nat. Biotechnol. 20, 666–668.

Stephanopoulos, G., Alper, H., Moxley, J., 2004. Exploiting biological complexity for strain improvement through systems biology. Nat. Biotechnol. 22, 1261–1267.

Stephanopoulos, G., Kelleher, J., 2001. How to make a superior cell. Science 292, 2024–2026.

Stephanopoulos, G., Simpson, T.W., 1997. Flux amplification in complex metabolic networks. Chem. Eng. Sci. 52, 2607–2627.

Stephanopoulos, G., Sinskey, A.J., 1993. Metabolic engineering: issues and methodologies. Trends Biotechnol. 11, 392–396.

Tyo, K.E., Zhou, H., Stephanopoulos, G., 2006. A high throughput screen for poly-3-hydroxybutyrate in *Escherichia coli* and *Synechocystis* PCC6803. Appl. Environ. Microbiol. 72, 3412–3417.

Voss, I., Steinbuchel, A., 2006. Application of a KDPG-aldolase gene-dependent addiction system for enhanced production of cyanophycin in *Ralstonia eutropha* strain H16. Metabol. Eng. 8, 66.

Yang, H.J., Wolff, E., Kim, M., Diep, A., Miller, J.H., 2004. Identification of mutator genes and mutational pathways in *Escherichia coli* using a multicopy cloning approach. Molec. Microbiol. 53, 283–295.

Yuan, L.Z., Rouviere, P.E., LaRossa, R.A., Suh, W., 2006. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *Escherichia coli*. Metabol. Eng. 8, 79.