Biotechnol Lett
DOI 10.1007/s10529-016-2133-0

ORIGINAL RESEARCH PAPER

Gene expression pattern analysis of a recombinant *Escherichia coli* strain possessing high growth and lycopene production capability when using fructose as carbon source

Wen Du · Yajian Song · Mengya Liu · Haixu Yang · Yao Zhang · Yanli Fan · Xuegang Luo · Zhongyuan Li · Nan Wang · Hongpeng He · Hao Zhou · Wenjian Ma · Tongcun Zhang

Received: 16 February 2016 / Accepted: 25 May 2016

Abstract
Objective *Escherichia coli* K12f-pACLYC has a high capability for growth and lycopene production when using fructose as carbon source and the transcription of genes involved was compared in glucose-grown and fructose-grown cells.
Results *Escherichia coli* K12f-pACLYC was grown on 10 g fructose l^{-1} and reached 4.6 g DCW l^{-1} with lycopene at 192 mg g DCW^{-1}, values that are 3-fold and 7-fold higher than when growing on glucose. Gene transcription profiles of fructose-grown and glucose-grown cells were compared. 384 differentially expressed genes (DEGs) with fold changes ≥4 were identified, and the transcription of genes involved in fructose uptake and metabolism, pyruvate catabolism, tricarboxylic acid cycle and oxidative phosphorylation varied significantly. These changes enhanced the metabolic flux into the Embden–Meyerhof–Parnas pathway and the tricarboxylic acid cylcle and coupled to oxidative phosphorylation. These enhanced activities provide more precursors, cofactors and energy needed for growth lycopene production.
Conclusion The high capability of *E. coli* K12f-pACLYC for growth and lycopene production when growing on fructose is due to transcriptional regulation, and the relevant genes were identified.

Keywords 2-C-Methyl-D-erythritol-4-phosphate pathway · Lycopene · Oxidative phosphorylation · Pyruvate · Transcriptome sequencing · Tricarboxylic acid cycle

Introduction

Lycopene is a tetraterpenoid (C40) carotenoid. Carotenoids are used in the feed, food and nutraceutical industries (Srivastava et al. 2013) but also as antioxidants and anti-carcinogens to prevent cardiovascular diseases, hepatic fibrogenesis or human papillomavirus persistence (Vaishampayan et al. 2007). Lycopene is mainly obtained from plants such as tomato (Lavecchia et al. 2008). Nevertheless, some studies have emerged on lycopene production by metabolic engineering based on *Escherichia coli* (Kim et al. 2009; Gallego-Jara et al. 2015; Kim et al. 2009).

Lycopene and other carotenoids are derived from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). In prokaryotes their

Electronic supplementary material The online version of this article (doi:10.1007/s10529-016-2133-0) contains supplementary material, which is available to authorized users.

W. Du · Y. Song · M. Liu · H. Yang · Y. Zhang · Y. Fan · X. Luo · Z. Li · N. Wang · H. He · H. Zhou · W. Ma · T. Zhang (✉)
College of Biotechnology, Tianjin University of Science and Technology, No. 29, 13th Street, Tianjin Economic and Technological Development Area, Tianjin 300457, China
e-mail: tony@tust.edu.cn

Published online: 05 July 2016

production is via the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway (Lee et al. 2002). Although *E. coli* possesses the MEP route for IPP synthesis, it still requires genes encoding geranylgeranyl pyrophosphate (GGPP) synthase (*crtE*), phytoene synthase (*crtB*) and phytoene desaturase (*crtI*) to synthesize lycopene. Therefore, construction of a recombinant strain that contains these enzymes is necessary to produce lycopene (Cunningham et al. 1994; Yoon et al. 2006).

Several strategies have been proposed to improve the production of lycopene or other carotenoids in *E. coli*, such as enhancing the metabolic flux towards the MEP pathway by overexpressing the genes encoding for 1-deoxy-D-xylulose-5-phosphate synthase and isopentenyl diphosphate isomerase (Lee et al. 2002; Yoon et al. 2006). Carotenoid production could also be improved by increasing the supplies of the precursors (pyruvate and glyceraldehyde-3-phosphate) by down-regulating pyruvate dehydrogenase (*aceE*), transaldolase B (*talB*) and phosphoglycerate mutase B (*gpmB*) gene, as well as amplifying fructose bisphosphate aldolase (*fbaA*) and triose phosphate isomerase (*tpi*) (Farmer et al. 2000). In addition, enhancing the pathways such as the TCA cycle and pentose phosphate (PPP) pathway can improve cofactor supply which lycopene production requires. Engineering multiple central metabolic modules together can improve β-carotene production significantly (Zhao et al. 2013).

Escherichia coli K12f-pACLYC is a lycopene production strain preserved in our laboratory. It produces lycopene in high yield only when using fructose as substrate. To understand the underlying mechanisms, the gene transcription pattern was studied by transcriptome sequencing and quantitative PCR analysis when cells were grown on fructose and glucose respectively. We believe that these results will contribute to the metabolic engineering of lycopene production strain.

Materials and methods

Strains, plasmid and culture conditions

Escherichia coli K12f-pACLYC was constructed by introducing plasmid pACLYC into *E. coli* K12f, *an E. coli* K12 mutant strain with high capability of using fructose. The plasmid pACLYC contains three lycopene synthesis genes, *crtE*, *crtB* and *crtI* (Cunningham et al. 1994). Cells were grown at 37 °C in lysogeny broth (LB) or LB agar containing different carbon sources from 2 to 10 g l^{-1}. 50 μg chloramphenicol l^{-1} was added to maintain plasmid.

Growth and sugar concentration determination

Growth of *E. coli* K12f-pACLYC was determined from the OD_{600} value and correlated to dry cell weight (DCW) by the DCW/OD_{600} standard curve [$y = 0.522x - 0.365$ (x is OD_{600}, y is DCW, $R^2 = 0.991$]. Sugar concentration was measured using dinitrosalicylic acid (DNS) method.

Lycopene extraction and determination

One ml culture was centrifuged, cells were washed, extracted in 1 ml acetone at 55 °C for 15 min in the dark. The samples were re-centrifuged and lycopene content of the extracts was quantified by measuring the A_{475} value and comparing with a lycopene standard purchased from Sigma-Aldrich (Alper et al. 2005).

RNA isolation and library construction

Cells cultivated in the medium with 6 g fructose l^{-1} and 6 g glucose l^{-1} respectively were harvested in mid-growth phase (14 h for fructose and 8 h for glucose). Total RNA was isolated using TRIzol. RNA quality was characterized by agarose gel electrophoresis and spectrophotometry. Ribo-Zero rRNA Removal Kit (Gram-negative bacteria) (Epicentre, Chicago, USA) was used to remove bacterial rRNA from the total RNA samples according to the manufacturer's protocols. RNA-seq libraries were created and quantified using NEB Next Ultra II DNA Library Prep Kit for Illumina (NEB, CA, USA) according to the manufacturer's protocols.

Transcriptome sequencing and bioinformatics analysis

The sequencing of two samples (glucose-grown and fructose-grown) was performed on Illumina HiSeq 2500 system (Illumina, Santiago, USA). Differentially expressed genes (DEGs) were screened out using DEseq. To identify putative functions of the genes, the

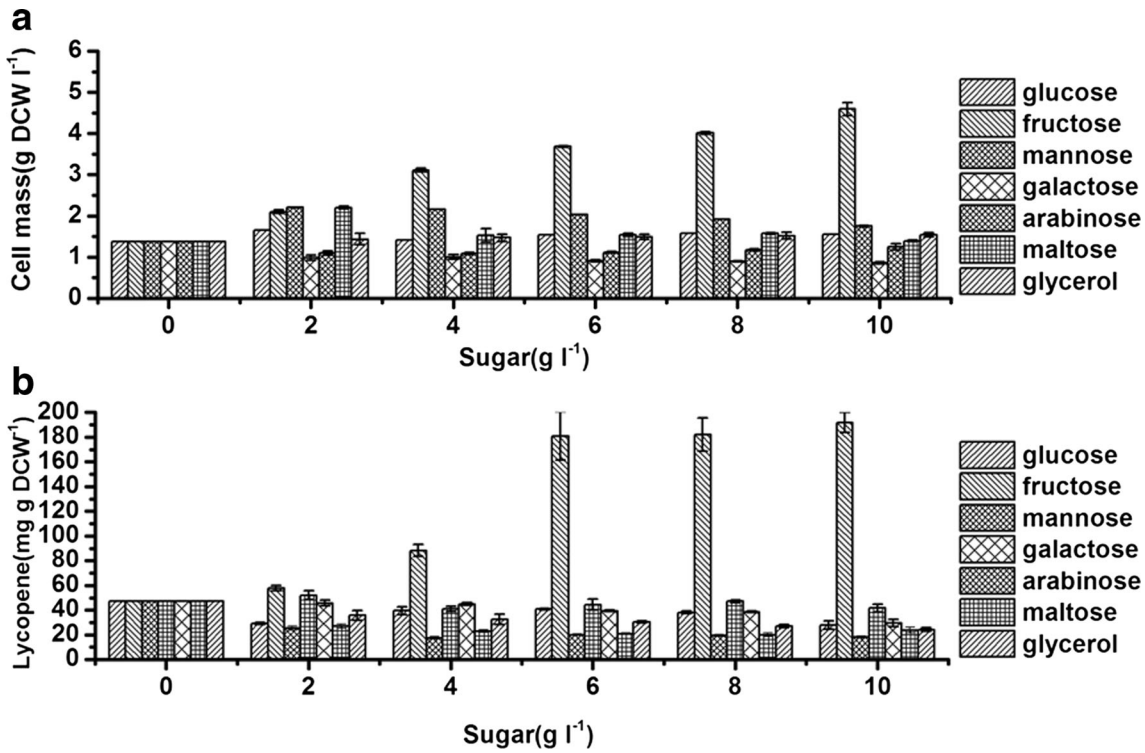

Fig. 1 Cell mass and lycopene yield when the strain was cultivated in media with different carbon sources. **a** Cell mass. **b** lycopene yield

sequences were aligned using BLASTX (E-value $\leq 0^{-5}$) against the following public protein databases: National Center for Biotechnology Information non-redundant (Nr) and nucleotide (Nt) database, Clusters of Orthologous Groups (COG) database, Gene Ontology (GO) database, and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The Blast2GO software package was used to compare and determine unigene GO annotations. Finally, WEGO software was used to obtain GO functional classifications for all annotated unigenes. Meanwhile, the DEGs annotated by KEGG were classified through the KEGG database. Transcriptome data including raw data and analysis data were deposited in the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE77674.

Quantitative real-time (qRT) PCR analysis

The qRT PCRs were performed in an Applied Biosystems StepOne real-time PCR system (Thermo, MA, USA). Three biological replicates were collected for each condition and amplified with three technical replicates, *recA* was used as the internal control. Primers used in qRT PCR analysis were shown in Supplementary Table 1.

Results

Growth and lycopene yield of *E. coli* K12f-pACLYC culturing with different carbon sources

We compared the biomass and lycopene accumulation of *E. coli* K12f-pACLYC growing on different sugars (Fig. 1) with best results using fructose. When 10 g fructose l^{-1} was added, the final cell mass and the lycopene yield reached 4.6 g DCW l^{-1} and 192 mg g DCW^{-1} (881 mg l^{-1}) respectively, which are 3-fold and 7-fold higher than that of glucose-grown cells.

Transcriptome sequencing of cells growing on fructose and glucose

Figure 2 shows the growth curve of *E. coli* K12f-pACLYC when cultivated in LB medium with either

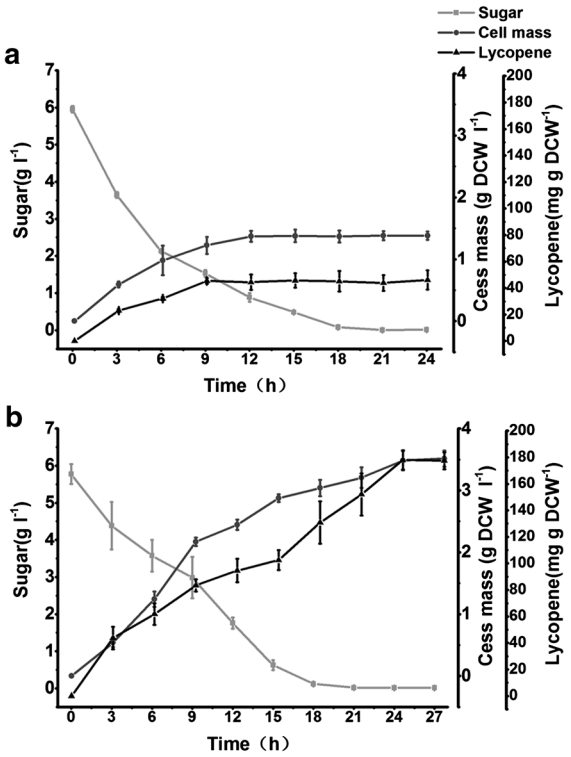

Fig. 2 Cell mass, sugar concentration and lycopene yield when the strain was cultivated on glucose or fructose, respectively. **a** 6 g glucose l^{-1} was added to the medium. **b** 6 g fructose l^{-1} was added to the medium

6 g glucose l^{-1} or 6 g fructose l^{-1}. Gene expression pattern of the strain in middle-growth phase on the two carbon sources was compared using transcriptome sequencing. 384 annotated differentially expressed genes (DEGs) with fold change ≥4 were identified. COG classification indicated that most DEGs are involved in carbohydrate transport, metabolism regulation or energy production and conversion (Supplementary Fig. 1), and pathway classification by KEGG database indicated that pathways such as ABC transporters, oxidative phosphorylation, TCA cycle, fructose and mannose metabolism, and phosphotransferase system contain most DEGs (Supplementary Fig. 2).

Transcription pattern of the genes related to lycopene synthesis

The transcription of *fruA* (b2167) and *fruB* (b2169) encoding PTS system fructose-specific enzyme II and *fruK* (b2168) encoding fructose-1-phosphate kinase was dramatically up-regulated by fructose, while genes in another pathway by which fructose enters EMP via fructose-6-phosphate was only slightly regulated (Fig. 3a). Gene transcription related to pyruvate metabolism changed significantly. Expression of *ldhA* (b1380), *poxB* (b0871), and *aceE* (b0114), which catalyzed pyruvate to generate lactate, acetate, and acetyl-CoA respectively, were all down-regulated by fructose.

Most genes in the TCA cycle were highly transcripted on fructose. The gene *pck* (b3403), encoding phosphoenolpyruvate carboxykinase which synthesizes phosphoenolpyruvate from oxaloacetate, was up-regulated by 42.8-fold (Fig. 3b). A large number of genes in oxidative phosphorylation pathway were also dramatically up-regulated by fructose, including NADH-quinone oxidoreductase subunit F-N (*nuoF-N*, b2276-2284), succinate dehydrogenase A-D (*sdhA-D*, b0721-0724) and F-type H^+-transporting ATPase subunit CD (*atpC, D*, b3731-3732) (Fig. 3c). No gene in isoprenoids biosynthesis pathway or PPP pathway was found to be regulated significantly. Gene ID, expression pattern and annotation of all the DEGs displayed in Fig. 3 were summarized in Supplementary Table 2.

Quantitative real-time PCR analysis of key genes

To evaluate the reliability of transcriptome sequencing results, the expression levels of 11 selected genes were determined by quantitative real-time (qRT) PCR. As shown in Table 1, the expression patterns of selected genes obtained using qRT PCR fit well with RNA-sequencing data, which demonstrated that the sequencing data is reliable.

Discussion

Introduction of genes, *crtEBI,* into *E. coli* enabled lycopene to accumulate (Lee et al. 2002). Further improvement of lycopene yield was obtained by gene engineering and fermentation optimization. A very high lycopene yield up to 45 mg g DCW^{-1} was reported by Kim et al. (2009). Zhu et al. (2015) used an *E. coli* mutant to produce lycopene by fed-batch fermentation and the yield reached 34.3 mg

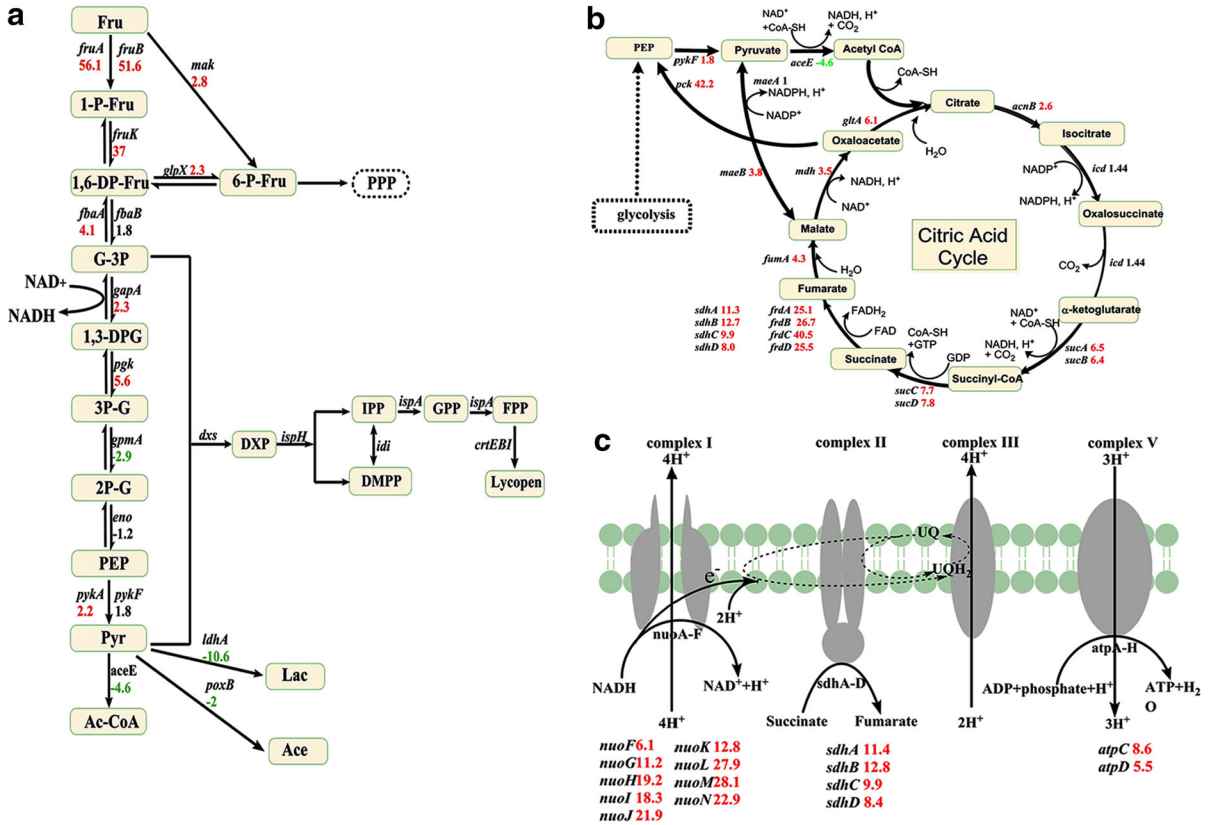

Fig. 3 Expression pattern of the genes in related pathways. **a** Expression pattern of the genes in EMP and pyruvate catabolism. **b** Expression pattern of the genes in TCA cycle. **c** Expression pattern of the genes in oxidative phosphorylation. Fold changes of the DEGs were determined by transcriptome sequencing on strain cultivated on fructose compared to glucose. *Numbers* in *red* represents up-regulated, *Numbers* in *green* with "-" represents down-regulated and Numbers in black represent up-regulated with the fold < 2

g DCW^{-1} (1.23 g l^{-1}). The overproduction and in situ extraction in organic/aqueous culture systems was carried out by Gallego-Jara et al. (2015) and the final specific lycopene yield reached 49.7 mg g DCW^{-1}. In the present study, when *E. coli* K12f-pACLYC was grown on 10 g fructose l^{-1}, lycopene reached 192 mg g DCW^{-1}, which is much higher than that on other carbon sources.

To understand the underlying reasons of the carbon source effect, gene transcription profiles were analyzed. In fructose-grown cells, the up-regulation of *fruA*,*fruB* and *fruK* can enhance the fructose-1,6-biphosphate production, which would contribute to the accumulation of G-3-P. On the other hand, transcription of the genes encoding pyruvate catabolic enzymes are all repressed, which would lead to the accumulation of pyruvate. The up-regulation of gene *pck* (b3403) can increase the formation of PEP from oxaloacetate. PEP is the precursor of pyruvate in the EMP pathway, so it is also likely to enhance the pyruvate production. As G-3-P and pyruvate are important precursors of the MEP pathway, these transcription changes would ensure the carbon supply of lycopene synthesis.

To produce one mol lycopene, 8 mol ATP and 16 mol NADPH are required (Alper et al. 2005). There are three major ways for the generation of NADPH in *E. coli*, including PPP pathway, TCA cycle and conversion of NADH to NADPH through membrane-bound transhydrogenase (PntA, B) (Sauer et al. 2004). In our research, PPP pathway exhibited no significant change between the two samples, but TCA cycle and oxidative phosphorylation modules were strongly induced by fructose. The high

Table 1 Quantitative real-time (qRT) PCR analysis of relative genes

Gene ID	Name	Annotation	Function description or Pathway located	Fold change (sequencing)	Fold change (qRT PCR)
b1603	*pntA*	pyridine nucleotide transhydrogenase	NADP/NADPH transformation	2.17	2.76 ± 0.39
b3731	*atpC*	F1 sector of membrane-bound ATP synthase	ATP synthesis	8.57	3.45 ± 0.68
b0726	*sucA*	2-oxoglutarate decarboxylase	TCA cycle	6.54	6.22 ± 0.68
b0727	*sucB*	dihydrolipoyltranssuccinase	TCA cycle	6.41	6.39 ± 1.99
b0721	*sdhC*	succinate dehydrogenase	TCA cycle	9.92	13.91 ± 3.58
b2278	*nuoL*	NADH:ubiquinone oxidoreductase	Electron transfer	27.86	25.77 ± 4.42
b3403	*pck*	phosphoenolpyruvate carboxykinase	Pyruvate metabolism	42.81	26.44 ± 2.35
b2167	*fruA*	fused fructose-specific PTS enzymes IIB component/IIC components	Fructose transport	56.1	38.04 ± 0.96
b2168	*fruK*	fructose-1-phosphate kinase	Fructose metabolism	37.53	40.93 ± 16.1
b0114	*aceE*	pyruvate dehydrogenase	Pyruvate metabolism	0.22	0.21 ± 0.05
b1380	*ldhA*	fermentative D-lactate dehydrogenase	Pyruvate metabolism	0.09	0.22 ± 0.10

transcription of the genes in TCA cycle leads to the increase of the reducing power including NADH and NADPH. The slight up-regulation of *pntA* is also likely to accelerate the conversation of NADH to NADPH. To produce energy, TCA cycle should be coupled to oxidative phosphorylation to generate ATP through electron transfer chain. These transcription changes will supply abundant cofactors and energy for growth and lycopene synthesis.

There are numerous genes participating in lycopene synthesis, so revealing their effects in this physiological process is highly required. This study will provide the candidate targets and theoretical support to the metabolic engineering of lycopene production strains.

Acknowledgments This work was supported by the foundation of the Key Laboratory of Industrial Fermentation Microbiology of the Ministry of Education and the Tianjin Key Lab of Industrial Microbiology (Tianjin University of Science and Technology) (No. 2013IM102).

Supplementary information Supplementary Table 1—Primers used in quantitative real-time (qRT) PCR analysis.

Supplementary Table 2—Gene ID, name, annotation, and expression pattern and of all the DEGs mentioned involved in the paper.

Supplemetary Fig. 1—COG function classifications of DEGs.

Supplemetary Fig. 2—KEGG function classifications of DEGs.

References

Alper H, Jin YS, Moxley JF, Stephanopoulos G (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. Metab Eng 7:155–164

Cunningham FX Jr, Sun Z, 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

Farmer WR, Liao JC (2000) Improving lycopene production in *Escherichia coli* by engineering metabolic control. Nat Biotechnol 18:533–537

Gallego-Jara J, De Diego T, Del Real A, Ecija-Conesa A et al (2015) Lycopene overproduction and in situ extraction in organic-aqueous culture systems using a metabolically engineered *Escherichia coli*. AMB Express 5:65

Kim SW, Keasling JD (2001) Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. Biotechnol Bioeng 72:408–415

Kim SW, Kim JB, Ryu JM, Jung JK et al (2009) High-level production of lycopene in metabolically engineered *E. coli*. Process Biochem 44:899–905

Lavecchia R, Zuorro A (2008) Improved lycopene extraction from tomato peels using cell-wall degrading enzymes. Eur Food Res and Technol 228:153–158

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

Sauer U, Canonaco F, Heri S, Perrenoud A et al (2004) The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. J Biol Chem 279:6613–6619

Srivastava S, Srivastava AK (2013) Lycopene; chemistry, biosynthesis, metabolism and degradation under various abiotic parameters. Int J Food Sci Tech 52:41–53

Vaishampayan U, Hussain M, Banerjee M, Seren S (2007) Lycopene and soy isoflavones in the treatment of prostate cancer. Nutr Cancer 59:1–7

Yoon SH, Kim JE, Lee SH, Park HM et al (2006) Engineering the lycopene synthetic pathway in *E. coli* by comparison of the carotenoid genes of *Pantoea agglomerans* and *Pantoea ananatis*. Appl Microbiol Biotechnol 74:131–139

Zhao J, Li Q, Sun T, Zhu X et al (2013) Engineering central metabolic modules of *Escherichia coli* for improving β-carotene production. Metab Eng 17:42–50

Zhu F, Lu L, Fu S, Zhong XF et al (2015) Targeted engineering and scale up of lycopene overproduction in *Escherichia coli*. Process Biochem 50:341–346