Enhanced squalene biosynthesis in *Yarrowia lipolytica* based on metabolically engineered acetyl-CoA metabolism

Yu-Ying Huang[a], Xing-Xing Jian[a], Yu-Bei Lv[a], Ke-Qing Nian[a], Qi Gao[a], Jun Chen[a], Liu-Jing Wei[a,*], Qiang Hua[a,b]

[a] *State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai, 200237, PR China*
[b] *Shanghai Collaborative Innovation Center for Biomanufacturing Technology, 130 Meilong Road, Shanghai, 200237, PR China*

A B S T R A C T

As a bioactive triterpenoid, squalene is widely used in the food industry, cosmetics, and pharmacology. Squalene's major commercial sources are the liver oil of deep-sea sharks and plant oils. In this study, we focused on the enhancement of squalene biosynthesis in *Yarrowia lipolytica*, with particular attention to the engineering of acetyl-CoA metabolism based on genome-scale metabolic reaction network analysis. Although the over-expression of the rate-limiting endogenous *ylHMG1* (3-hydroxy-3-methylglutaryl-CoA reductase gene) could improve squalene synthesis by 3.2-fold over that by the control strain, the availability of the key intracellular precursor, acetyl-CoA, was found to play a more significant role in elevating squalene production. Analysis of metabolic networks with the newly constructed genome-scale metabolic model of *Y. lipolytica* iYL_2.0 showed that the acetyl-CoA pool size could be increased by redirecting carbon flux of pyruvate dehydrogenation towards the ligation of acetate and CoA or the cleavage of citrate to form oxaloacetate and acetyl-CoA. The over-expression of either acetyl-CoA synthetase gene from *Salmonella enterica* (*acs**) or the endogenous ATP citrate lyase gene (*ylACL1*) resulted in a more than 50% increase in the cytosolic acetyl-CoA level. Moreover, iterative chromosomal integration of the *ylHMG1*, *asc**, and *ylACL1* genes resulted in a significant improvement in squalene production (16.4-fold increase in squalene content over that in the control strain). We also found that supplementation with 10 mM citrate in a flask culture further enhanced squalene production to 10 mg/g DCW. The information obtained in this study demonstrates that rationally engineering acetyl-CoA metabolism to ensure the supply of this key metabolic precursor is an efficient strategy for the enhancement of squalene biosynthesis.

1. Introduction

Squalene ($C_{30}H_{50}$) is a high-value triterpenoid. Currently, its commercial production mainly comes from the liver oil of deep-sea sharks and plant seeds, which cannot meet the increasing demand for this compound. Increased consumer demand has prompted the development of microbial bioprocesses for squalene production. The most commonly used microbial hosts for squalene biosynthesis are microalga, followed by *S. cerevisiae* and *E. coli*. Genetic modifications of squalene synthase, squalene epoxidase, and HMGR were usually performed together with medium optimization to enhance squalene production (Ghimire et al., 2009; Kaya et al., 2014; Paramasivan and Mutturi, 2017). Microalga has natural advantages over bacteria and yeasts for squalene production; the microalga *Thraustochytrid aurantiochytrium* sp. 18W-13a accumulated more than 1 g/L and 198 mg/g

dry cell weight (DCW) of squalene without any genetic modification after 4 d of culture (Kaya et al., 2014). However, *S. cerevisiae* and *E. coli* produced less than 100 mg/L with genetic modification of the squalene synthase gene and the MVA or MEP pathway genes. For example, metabolic engineering of *E. coli* produced 4.1 mg/L of squalene by heterologous expression of the squalene synthesis pathway genes *hopA*, *hopB,* and *hopD* from *Streptomyces peucetius* ATCC 27952, and the squalene level was elevated to 11.8 mg/L when coupled with modulation of the MEP pathway genes *dxs* and *idi* (Ghimire et al., 2009). Integration of two copies of *tHMG1* into the chromosome coupled with regeneration of NADPH by overexpression of *pox5* yielded 58.6 mg/g DCW of squalene (28.4 mg/L), a 27.5-fold increase over the control strain of *S. cerevisiae* (Paramasivan and Mutturi, 2017). Achieving high squalene yield or productivity still remains a challenge in bacteria and yeasts.

Recently, the non-conventional oleaginous yeast *Y. lipolytica* has been considered an attractive microorganism for terpenoid production owing to its oleaginous nature and ability of exploiting low-value carbons to produce value-added chemicals in an environmentally friendly manner (Gao et al., 2016; Liu et al., 2015; Papanikolaou et al., 2003; Plácido and Capareda, 2016). With certain well-suited mechanisms for oleaginous metabolism, some metabolically engineered *Y. lipolytica* strains could even accumulate lipids up to 90% DCW under optimized conditions (Blazeck et al., 2014). The DuPont company has made several lipid-related commercial products including New Harvest™ EPA oil for human dietary supplement and EPA-rich *Y. lipolytica* biomass for sustainably farmed salmon Verlasso® (https://www.verlasso.com/) based on metabolic engineering of *Y. lipolytica* (Xue et al., 2013). Recent studies also demonstrated that *Y. lipolytica* could be used for terpenoid production by metabolic engineering of the MVA pathway and heterologous expression of the corresponding terpenoid synthases (Cao et al., 2016, 2017; Gao et al., 2017; Matthaus et al., 2014; Yang et al., 2016). All of these results suggest an enormous potential for *Y. lipolytica* to synthesize liposoluble chemicals in an economically feasible manner. The most commonly used strategies to achieve large-scale production of terpenoids in *Y. lipolytica* include screening corresponding synthases, optimizing the MVA pathway, downregulating of competing pathways, and optimizing of fermentation processes.

Engineering central carbon metabolism to supply adequate precursors has been commonly employed to improve the biosynthesis of acetyl-CoA derived chemicals (Meadows et al., 2016). Acetyl-CoA is a key node in central metabolism and numerous researchers are focusing on metabolically engineering acetyl-CoA metabolism for the biosynthesis of a wide range of industrially interesting chemicals. The commonly employed strategies for enhancing the acetyl-CoA supply include the optimal expression of the PDH bypass or acetyl-CoA synthetase (Chen et al., 2013b), expression of the phosphoketolase pathway (Kocharin et al., 2013), and functional expression of the pyruvate dehydrogenase complex in the cytosol (Kozak et al., 2014). An extended review on the engineering of microbial acetyl-CoA metabolism can be found elsewhere (Krivoruchko et al., 2015).

Recently, genome-scale metabolic models (GEMs) have been widely employed to understand genome-scale genotype-phenotype relationships of microorganisms. Based on the model-predicted information of metabolic flux distributions, rational strategies for metabolically engineering the targeted intracellular metabolisms could be accomplished (Chen et al., 2013a; Feist et al., 2009; Gruchattka and Kayser, 2015). The metabolic network model of *Y. lipolytica* has recently been investigated intensively and five versions of the genome-scale metabolic network models have been constructed since 2012, namely iYL619_PCP, iNL859, iMK735, iYali4, and iYL_2.0 (Kavscek et al., 2015; Kerkhoven et al., 2016; Loira et al., 2012; Pan and Hua, 2012; Wei et al., 2017). The latest version of the *Y. lipolytica* metabolic network model iYL_2.0 (645 genes, 1083 metabolites, and 1471 reactions) was developed in-house and employed to study the acetyl-CoA related metabolism in *Y. lipolytica* (Wei et al., 2017).

In this work, the enhancement of squalene biosynthesis in *Y. lipolytica* was investigated, with a particular focus on the understanding and engineering acetyl-CoA metabolism (Fig. 1). Based on the *in silico* metabolic reaction network analysis and using the newly constructed model system iYL_2.0, an integrated push-block-pull strategy including the push of carbon fluxes from intracellular precursors to the formation of acetyl-CoA, block of further conversion of acetyl-CoA by competing pathways, and pull of acetyl-CoA consumption flux towards the target metabolite was performed to improve the acetyl-CoA supply and squalene production. Optimal expression of *acs** coupled with overexpression of endogenous *ylACL1* and *ylHMG1* significantly increased squalene production. Moreover, 10 mg/g DCW squalene was successfully synthesized by culturing the genetically engineered strain in YPD medium with supplementation of citrate, more than 50-fold improvement over the starting strain. This study demonstrated that rational engineering of acetyl-CoA metabolism is an efficient strategy to develop *Y. lipolytica* as a source of squalene biosynthesis.

2. Materials and methods

2.1. Strains, media, and culture conditions

The auxotrophic *Y. lipolytica* strain Po1f (Leu⁻, Ura⁻), which was kindly provided by Prof. Catherine Madzak (Institute National de la Recherche Agronomique/AgroParisTech, France), was used as the host strain. The yeast strains were cultured and fermented at 30 °C and 220 rpm in YPD medium (10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose; for solid medium 15 g/L agar was added). YNB medium (6.7 g/L yeast nitrogen base without amino acids, 10 g/L glucose; for solid medium 16 g/L agar was added) containing the appropriate nutrients (0.1 g/L leucine or 0.1 g/L uracil) was used for screening transformants of *Y. lipolytica*. For strain culture and activation, the *Y. lipolytica* strains were first inoculated in 15 mL glass tubes containing 5 mL YPD medium, and then grown at 30 °C and 220 rpm. For the fermentation, the pre-cultured *Y. lipolytica* strains were transferred to 250-mL shake flasks containing 50 mL YPD with an initial OD_{600} of 0.01, and cultivated at 30 °C and 220 rpm.

The *E. coli* strain JM109 was used for cloning and plasmid propagation. The cells were grown at 37 °C and 220 rpm in Luria-Bertani (LB) medium (5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl; for solid medium 15 g/L agar was added) containing 100 mg/L kanamycin or 100 mg/L ampicillin when necessary. All strains used in this study are listed in Table 1.

2.2. Construction of plasmids and transformation of Y. lipolytica

All vectors used in this study are listed in Table 2 and primers used for the construction of vectors are listed in Table S1. For gene expression, the integrative plasmids pINA1269 and pINA1312 were used as platform vectors (Madzak et al., 2004), and all of the structural genes were cloned into these two backbone vectors. Selected acetyl-CoA-related gene targets, the acetyl-CoA synthetase gene (*ylACS*, NCBI Gene ID: 2908441), pyruvate decarboxylase gene (*ylPDC*, NCBI Gene ID: 2910997), aldehyde dehydrogenase gene (*ylALD*, NCBI Gene ID: 2909058), ATP citrate lyase isozyme gene (*ylACL1*, NCBI Gene ID: 2910381), and ATP citrate lyase isozyme gene (*ylACL2*, NCBI Gene ID: 2912101), were amplified from *Y. lipolytica* Po1f genomic DNA with the Phanta Super-Fidelity DNA Polymerase (Vazyme Biotech Co., Ltd, Nanjing, China). Similarly, the pyruvate carboxylase gene (*ylPYC*, NCBI Gene ID: 2909579) was amplified from the cDNA of *Y. lipolytica*, and the *S. enterica* acetyl-CoA synthetase gene with an L641 P mutation (*acs**, NCBI Gene ID: 1255801) was codon-optimized and synthesized by Generay Biotech Co., Ltd. (Shanghai, China). Plasmid pINA1269-acs* was synthesized by Generay Biotech. In addition, the MVA pathway gene target, 3-hydroxy-3-methylglutaryl-CoA reductase gene (*ylHMG1*, NCBI Gene ID: 2912214), was amplified from the *Y. lipolytica* Po1f genomic DNA while the truncated 3-hydroxy-3-methylglutaryl-CoA reductase gene (*tHMG1*, NCBI Gene ID: 854900) was amplified from *S. cerevisiae* CEN.PK113-5D genomic DNA. All of the PCR products and synthetic gene fragments were designed with both ends containing a ∼20-bp region that was homologous to the plasmid backbone to facilitate Gibson assembly. Purified acetyl-CoA metabolism related genes were assembled with a *Pml*I- and *Bam*HI-digested pINA1269 plasmid, while purified MVA pathway genes were assembled with *Pml*I- and *Bam*HI-digested pINA1312 plasmid, using the ClonExpress® II One Step Cloning Kit (Vazyme Biotech). Thus, the plasmids pINA1269-ylACS, pINA1269-ylPDC, pINA1269-ylALD, pINA1269-ylACL1, pINA1269-ylACL2, pINA1269-ylPYC, pINA1269-acs*, pINA1312-ylHMG1, and pINA1312-tHMG1 were obtained. To assemble multiple genes in pINA1269, the genes in the promoter-gene cassette-terminator form were amplified from the constructed plasmids and ligated to target