

# 在无天然游离质粒产甘油假丝酵母中非整合表达合成咖啡酸

董德金<sup>1,2</sup>, 王心仪<sup>1,2</sup>, 宗红<sup>1,2</sup>, 陆信曜<sup>1,2</sup>, 诸葛斌<sup>\*1,2</sup>

1 江南大学工业生物技术教育部重点实验室, 江苏 无锡 214122

2 江南大学生物工程学院工业微生物及多元醇研究设计中心, 江苏 无锡 214122

董德金, 王心仪, 宗红, 陆信曜, 诸葛斌. 在无天然游离质粒产甘油假丝酵母中非整合表达合成咖啡酸[J]. 微生物学通报, 2022, 49(10): 4104-4117

Dong Dejin, Wang Xinyi, Zong Hong, Lu Xinyao, Zhuge Bin. Synthesis of caffeic acid by episomal expression in *Candida glycerinogenes* without natural episomal plasmids[J]. Microbiology China, 2022, 49(10): 4104-4117

**摘要:**【背景】咖啡酸(3,4-二羟基肉桂酸)是一种有多种生物活性和药用价值的天然酚类化合物,产甘油假丝酵母(*Candida glycerinogenes*)具有咖啡酸前体代谢途径,高耐酸且生长与发酵速率快,是潜在高产咖啡酸的底盘细胞,但无游离载体将影响咖啡酸合成的深入研究。【目的】探索在无天然游离质粒的 *C. glycerinogenes* 中构建操作更简便、表达能力更强的游离载体合成咖啡酸的可行性。

【方法】筛选自主复制序列(autonomously replicating sequence, ARS), 构建适用于 *C. glycerinogenes* 合成咖啡酸的游离载体, 并通过改造其 ARS 位置、标记基因 *URA5* 启动子长度、基因表达元件和利用 Kozak 序列优化表达并合成咖啡酸。【结果】构建的 5 个分别含不同 ARS 的载体中, pTGAPU-CA-AOX1t- KLARS 在 *C. glycerinogenes* 中能自我复制并表达合成咖啡酸的基因, 而且当 ARS 位于目的基因表达元件上游、*URA5* 启动子截短 250 bp, 或分别采用 Kozak 序列与终止子 *URA5t* 后, 咖啡酸产量较改造前均有明显提升, 最高产量为初始产量的 3.73 倍, 达 29.1 mg/L, 高于前期整合表达产量。【结论】在 *C. glycerinogenes* 中非整合表达合成咖啡酸且优于整合表达, 为今后利用游离载体改造咖啡酸合成代谢途径提供了新工具, 同时为其他无游离质粒菌株构建非整合表达体系提供参考。

**关键词:** 产甘油假丝酵母; 咖啡酸; 游离载体; 载体改造; 自主复制序列

基金项目: 国家自然科学基金(31970033)

Supported by: National Natural Science Foundation of China (31970033)

\*Corresponding author: E-mail: bzhuge@126.com

Received: 2022-03-09; Accepted: 2022-04-09; Published online: 2022-05-05

## Synthesis of caffeic acid by episomal expression in *Candida glycerinogenes* without natural episomal plasmids

DONG Dejin<sup>1,2</sup>, WANG Xinyi<sup>1,2</sup>, ZONG Hong<sup>1,2</sup>, LU Xinyao<sup>1,2</sup>, ZHUGE Bin<sup>\*1,2</sup>

1 Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, Jiangsu, China

2 Lab of Industrial Microorganism & Research and Design Center for Polyols, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

**Abstract:** **[Background]** Caffeic acid (3,4-dihydroxycinnamic acid), a natural phenolic compound, has various biological activities and great medicinal value. *Candida glycerinogenes*, which harbors the caffeic acid precursor metabolic pathway, is a potential high caffeic acid-yielding chassis cell owing to the high acid tolerance, fast growth, and high fermenting rate. However, it has no episomal vectors, which limits the in-depth research on caffeic acid synthesis. **[Objective]** To explore whether it is feasible to construct episomal vectors with stronger expression of caffeic acid in *C. glycerinogenes* without natural episomal plasmids in a simpler way. **[Methods]** Autonomously replicating sequences (ARSs) were selected to establish the episomal plasmid for the synthesis of caffeic acid in *C. glycerinogenes*, and the plasmid was further optimized by modifying its ARS position, the marker gene promoter length, elements for gene expression, and using Kozak sequence. **[Results]** Of the five vectors constructed containing different ARSs, pTGAPU-CA-AOX1t-KLARS was able to self-replicate and express the genes for the synthesis of caffeic acid in *C. glycerinogenes*. After the vector was modified in different ways, such as ARS located upstream of the target gene expression element, *URA5* promoter truncated by 250 bp, using Kozak sequence or *URA5* terminator, the yield of caffeic acid was significantly higher. The maximum yield after the modification was 3.73 times higher than the initial yield, reaching 29.1 mg/L. **[Conclusion]** Caffeic acid was synthesized by episomal expression in *C. glycerinogenes*, with better results than that by integrated expression, which provides a new tool for future modification of the caffeic acid anabolic pathway with episomal vectors and a reference for the construction of episomal expression systems in other strains without episomal plasmids.

**Keywords:** *Candida glycerinogenes*; caffeic acid; episomal vector; vector modification; autonomously replicating sequence

咖啡酸(3,4-二羟基肉桂酸)是一种普遍存在于植物中的天然羟基肉桂酸,具有多种药理活性,如抗炎、抗氧化、抗癌、抗病毒、抗糖尿病和抗抑郁能力<sup>[1]</sup>,被广泛应用于医药和食品等健康领域。然而咖啡酸在植物中以低水平积累,提取困难且昂贵<sup>[2]</sup>。生物合成咖啡酸是以酪氨酸为前体,通过酪氨酸解氨酶(tyrosine ammonia lyases, TAL)与4-羟基苯乙酸-3-羟化酶

(4-hydroxyphenylacetate 3-hydroxylase, 4HPA3H)(由 *hpaB* 和 *hpaC* 编码)、对香豆酸-3-羟化酶(p-coumarate 3-hydroxylase, C3H)或细胞色素P450酶 CYP199A2 合作催化得到<sup>[3]</sup>(图1)。当前研究代谢改造合成咖啡酸最常用底盘细胞是大肠杆菌(*Escherichia coli*)和酿酒酵母(*Saccharomyces cerevisiae*)<sup>[4]</sup>,在 *E. coli* 中, Lin 等利用内源酶 4HPA3H 与荚膜红细菌(*Rhodobacter*

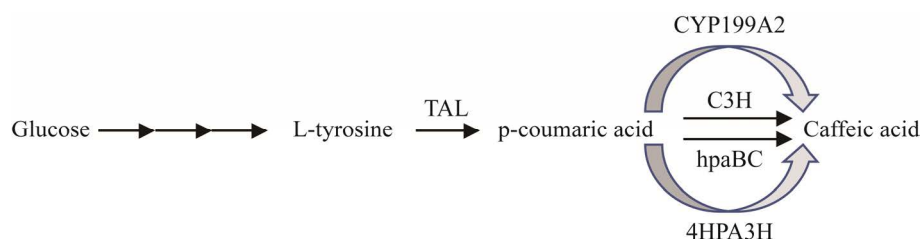


图 1 咖啡酸生物合成途径

Figure 1 Caffeic acid biosynthesis pathway.

*capsulatus*)来源的 TAL (RcTAL)成功以葡萄糖生产咖啡酸(12.1 mg/L)<sup>[2]</sup>, Li 等在 *S. cerevisiae* 中表达 RcTAL 和拟南芥的 C3H 与 P450 还原酶 CPR1, 并经代谢途径和培养基优化后咖啡酸产量达 11.4 mg/L<sup>[5]</sup>。除添加昂贵的前体物质外, 受限于通路基因表达强度或菌株自身特性等因素, 代谢通路未改造时 *E. coli* 与 *S. cerevisiae* 中咖啡酸初始产量均较低, 同时 *E. coli* 等模式菌株耐酸性差将是未来高产咖啡酸的瓶颈。

产甘油假丝酵母(*Candida glycerinogenes*) WL2002-5 是一株有自主知识产权与多重抗逆性能的工业安全菌株<sup>[6]</sup>, 在生物合成化合物方面有重要应用<sup>[7-9]</sup>。与 *E. coli* 和 *S. cerevisiae* 等模式菌株相比, *C. glycerinogenes* 不仅具有从葡萄糖到酪氨酸的完整代谢途径, 更重要的是具有高于模式菌种数倍甚至 10 倍的强耐酸性能, 同时还有生长与发酵速率快、耐高温和耐高渗透压等特性<sup>[10]</sup>, 是生物合成咖啡酸很有潜力的宿主。本课题组已成功在 *C. glycerinogenes* 中整合表达咖啡酸合成的相关基因, 初始产量为 14.71 mg/L<sup>[11]</sup>, 但二倍体 *C. glycerinogenes*<sup>[12]</sup>的整合操作烦琐复杂, 影响深层次的合成代谢改造。

本研究构建并优化 *C. glycerinogenes* 非整合表达体系, 以期解决 *C. glycerinogenes* 非整合表达合成咖啡酸这一限制性问题, 同时为后期生产咖啡酸改造代谢途径提供新工具。

## 1 材料与方法

### 1.1 菌株和质粒

研究中所使用的菌株和质粒如表 1 所示, 其中 *E. coli* JM109、*C. glycerinogenes* WL2002-5 (CCTCC M93018)、*C. glycerinogenes* WL2002-5/ $\Delta$ URA5、*Pichia pastoris* GS115 和 pRS414 为实验室保藏, 其余质粒为本研究构建。本研究使用的所有引物序列如表 2 所示。

### 1.2 主要试剂和仪器

蛋白胨和酵母粉, OXOID 公司; 酵母无氮基氮源(yeast nitrogen base, YNB)与酵母基因组 DNA 快速抽提试剂盒, 生工生物工程(上海)股份有限公司; *Ex Taq* DNA 聚合酶、限制性内切酶和 pMD™ 19-T Vector Cloning Kit, TaKaRa 公司; ABclonal MultiF Seamless Assembly Mix, 武汉爱博泰克生物科技有限公司; AceQ Universal SYBR qPCR Master Mix, 南京诺唯赞生物科技股份有限公司; 质粒提取试剂盒、柱回收试剂盒和凝胶回收试剂盒, 上海捷瑞生物工程有限公司。PCR 扩增仪和实时荧光定量基因扩增仪, Bio-Rad 公司; 高效液相色谱仪, 江苏汉邦科技有限公司; 紫外分光光度计, 北京普析通用公司。所用引物由无锡天霖技术有限公司合成。

### 1.3 培养基与培养方法

LB 培养基(g/L): 酵母提取物 5.0, 蛋白胨

表 1 本研究所用菌株和质粒

Table 1 Strains and plasmids used in this study

Strain and plasmid	Related characteristics
Strains	
<i>E. coli</i> JM109	Genetic cloning of hosts
<i>C. glycerinogenes</i> WL2002-5	Wild type; CCTCC M93018
<i>C. glycerinogenes</i> WL2002-5/ $\Delta$ URA5	Uracil nutritional deficient strains
<i>P. pastoris</i>	GS115
Plasmids	
pRS414	CEN/ARS, TEF1p, CYC1t, TRP1, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-414ARS	414ARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-PSARS	PSARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-KMARS	KMARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-SPARS	SPARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-upKLARS	KLARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-opKLARS	KLARS, GAPp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-Ko-CA-AOX1t-KLARS	KLARS, GAPp, Kozak, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-AOX1t-KLARS-T250	pTGAPU-CA-KLARS (Promoter truncation 250 bp of URA5)
pTGAPU-CA-AOX1t-KLARS-T500	pTGAPU-CA-KLARS (Promoter truncation 500 bp of URA5)
pTGAPU-CA-AOX1t-KLARS-T700	pTGAPU-CA-KLARS (Promoter truncation 700 bp of URA5)
pTGPDU-CA-AOX1t-KLARS	KLARS, GPDp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTSTL3U-CA-AOX1t-KLARS	KLARS, STL3p, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGMTU-CA-AOX1t-KLARS	KLARS, GMTp, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTCWPIU-CA-AOX1t-KLARS	KLARS, CWP1p, TAL, hpaB, hpaC, AOX1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-CYC1t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, CYC1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-URA5t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, URA5t, URA5, Amp <sup>r</sup>
pTGAPU-CA-TRP1t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, TRP1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-1t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, 1t, URA5, Amp <sup>r</sup>
pTGAPU-CA-2t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, 2t, URA5, Amp <sup>r</sup>
pTGAPU-CA-3t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, 3t, URA5, Amp <sup>r</sup>
pTGAPU-CA-4t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, 4t, URA5, Amp <sup>r</sup>
pTGAPU-CA-5t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, 5t, URA5, Amp <sup>r</sup>
pTGAPU-CA-G1t-KLARS	KLARS, GAPp, TAL, hpaB, hpaC, G1t, URA5, Amp <sup>r</sup>

10.0, 氯化钠 10.0; 抗性平板需添加 100  $\mu$ g/mL 的氨苄青霉素。

SD 培养基(g/L): YNB 6.7, 葡萄糖 20.0, pH 为 7.0。

固体培养基需要添加 20 g/L 的琼脂粉。

培养方法:*E. coli* JM109 在含有 100  $\mu$ g/mL 氨苄青霉素的 LB 培养基中于 37  $^{\circ}$ C 培养, 用以

本研究中所有载体的构建; *C. glycerinogenes* WL2002-5/ $\Delta$ URA5 用于所有酵母转化并于 30  $^{\circ}$ C 条件下利用 SD 培养基培养。

发酵培养: SD 培养基, 30  $^{\circ}$ C, 其中研究不同启动子时重组菌株特定培养条件。GAPp<sup>[13]</sup>: YNB 6.7 g/L, 葡萄糖 300 g/L, pH 为 7.0, 培养温度为 30  $^{\circ}$ C; GPDp<sup>[14]</sup>: YNB 6.7 g/L, 葡萄糖



(续表 1)

TRP1t-F	aaatggaagctccagtttaaGTAGATTA ACTTAAAAATATAAAAAAATTGAATGATTCTTATAG
TRP1t-R	ctgatattatccaagatgttgctagcGTT CAGAATTTACAAGATTCCAGATG
1t-F	tatatatatataactgtctAGAAATAAAGAGTATCATCTTTCAAAGCTAGCCAACATCTTTGGATAATATCAGAATGAG
1t-R	agacagttatatatatataTATATATATATATATATATATATATATATAA ACTGGAGCTTCCATTT
2t-F	tatatatttactgtctagattttaataaagagatcatctttcaaaGCTAGCCAACATCTTTGGATAATATCAGAATGAG
2t-R	ttgaaagatgatactctttatataaaatctagacagtaaaatataTTAAACTGGAGCTTCCATTT
3t-F	tataaactgtctagaaaaaagagatcatctttcaaaGCTAGCCAACATCTTTGGATAATATCAGAATGAG
3t-R	ttgaaagatgatactctttttctagacagttatataTTAAACTGGAGCTTCCATTT
4t-F	tataaactgtctagaaataaattttttcaaaGCTAGCCAACATCTTTGGATAATATCAGAATGAG
4t-R	ttgaaaaaaattttttctagacagttatataTTAAACTGGAGCTTCCATTT
5t-F	taactgtctagaaataaagagTATCATCTTTCAAAGCTAGCCAACATCTTTGGATAATATCAGAATGAG
5t-R	ctctttttctagacagttataTATAAAAAA AAAAAA TTAAACTGGAGCTTCCATTT
G1t-F	tataaactgtctagaaataaagagatcatctttcaaaGCTAGCCAACATCTTTGGATAATATCAGAATGAG
G1t-R	ttgaaagatgatactctttttctagacagttatataTTAAACTGGAGCTTCCATTT
RT-URA3-F	TGGTGTCTTCCGATTGCCGAATG
RT-URA3-R	TTGTTTCTTGGGCTGCCTCCTTC
RT-CA-F	ACATGGTGGTGTGGTGCTTCTG
RT-CA-R	ACTTCTGGGGTTGGTCTTCTTTCAC

注: F: 上游引物, R: 下游引物, 下划线部分为限制性酶切位点, 小写字母为同源重组连接时的同源序列

Note: F: Upstream primer, R: Downstream primer, underlined part is restriction enzyme cut site, lowercase letter is homologous sequence when homologous recombination is joined.

200 g/L, pH 为 7.0, 培养温度为 30 °C; STL3p<sup>[15]</sup>: YNB 6.7 g/L, 葡萄糖 300 g/L, pH 为 7.0, 培养温度为 30 °C; GMTp<sup>[16]</sup>: YNB 6.7 g/L, 葡萄糖 20 g/L, pH 为 2.0, 培养温度为 30 °C; CWP1p<sup>[17]</sup>: YNB 6.7 g/L, 葡萄糖 20 g/L, pH 为 7.0, 培养温度为 42 °C。250 mL 三角瓶 50 mL 装液量, 200 r/min 振荡培养, 初始接菌  $OD_{600}$  为 0.1, 培养 3 d。

## 1.4 载体构建

### 1.4.1 咖啡酸合成初始游离载体构建

以 *C. glycerinogenes* 基因组为模板, 利用引物对 GAPp-F/GAPp-R 与 URA5-F/URA5-R 通过 PCR 扩增获得 GAP 启动子和基因 URA5, 以巴斯德毕赤酵母(*Pichia pastoris*) GS115 基因组为模板, 利用 AOX1t-F/AOX1t-R 通过 PCR 扩增获得 AOX1 终止子, 自主复制序列(KLARS、414ARS、PSARS、KMARS 和 SPARS)与针对

*C. glycerinogenes* 密码子偏好性优化后通过融合蛋白技术获得的基因片段 CA: 5'-TAL-hpaB-hpaC-3' (其中 TAL 与 hpaB 之间有连接序列 GGTCTGGT)由无锡天霖技术有限公司合成; 利用 KLARS-F/KLARS-R 通过 PCR 扩增获得 KLARS, 同样方法获得 PSARS、KMARS、SPARS 和 414ARS, 利用 CA-F/CA-R 通过 PCR 扩增获得 CA, 经 Gibson Assembly 方法获得片段 5'-GAPp-CA-AOX1t-ARS-URA5-3', 利用 TA 克隆将其与 pMD<sup>TM</sup> 19-T 载体连接获得 pTGAPU-CA-AOX1t-ARS (图 2), 其中 ARS 分别为 KLARS、414ARS、PSARS、KMARS 和 SPARS。

### 1.4.2 优化 ARS 位点载体构建

利用引物对 upKLARS-F/upKLARS-R 通过 PCR 扩增获得 upKLARS, *Stu* I 酶切 pTGAPU-CA-AOX1t-KLARS, 经 Gibson Assembly 方法获得 pTGAPU-CA-AOX1t-upKLARS (up-KLARS),

同样方法获得 pTGAPU-CA-AOX1t-opKLARS (op-KLARS); 添加 Kozak 序列载体构建: 以 pTGAPU-CA-AOX1t-KLARS 为模板, 利用 Kozak-F/Kozak-R 经 PCR 扩增后直接转化 *E. coli* JM109 获得 pTGAPU-Ko-CA-AOX1t-KLARS (Ko); *URA5* 启动子截短载体构建: 以 pTGAPU-CA-AOX1t-KLARS 为模板, 分别利用引物对 T250-F/TP-R、T500-F/T500-R 和 T700-F/T700-R 经 PCR 扩增后直接转化 *E. coli* JM109 获得 pTGAPU-CA-AOX1t-KLARS-T250 (T250)、pTGAPU-CA-AOX1t-KLARS-T500 (T500) 和 pTGAPU-CA-AOX1t-KLARS-T700 (T700); 不同启动子载体构建: 以 *C. glycerinogenes* 基因组为模板, 利用 GPDp-F/GPDp-R 通过 PCR 扩增获得 *GPD* 启动子, *Stu* I 与 *Apa* I 双酶切 pTGAPU-CA-AOX1t-KLARS, 经 Gibson Assembly 方法获得 pTGAPU-CA-AOX1t-KLARS, 同样方法获得 pTSTL3U-CA-AOX1t-KLARS、pTGMTU-CA-AOX1t-KLARS 和 pTCWPIU-CA-AOX1t-KLARS; 不同终止子载体构建: 以 pRS414 为模板, 利用 CYC1t-F/CYC1t-R 通过 PCR 扩增获得 *CYC1* 终止子, 以 *C. glycerinogenes* 基因组为模板, 分别利用 *URA5t-F/URA5t-R* 和 *TRP1t-F/TRP1t-R* 通过 PCR 扩增获得 *URA5* 和 *TRP1* 的终止子, 利用 1t-F/1t-R 通过 PCR 扩增获得合成终止子 1t, 分别通过 Gibson Assembly 方法与 *Nhe* I 酶切后的 pTGAPU-CA-AOX1t-KLARS 连接, 获得 pTGAPU-CA-CYC1t-KLARS、pTGAPU-CA-URA5t-KLARS、pTGAPU-CA-TRP1t-KLARS 和 pTGAPU-CA-1t-KLARS, 同样方法获得 pTGAPU-CA-2t-KLARS、pTGAPU-CA-3t-KLARS、pTGAPU-CA-4t-KLARS、pTGAPU-CA-5t-KLARS 和 pTGAPU-CA-G1t-KLARS。

### 1.5 载体拷贝数测定

提取阳性克隆子总 DNA, 以载体上的基因

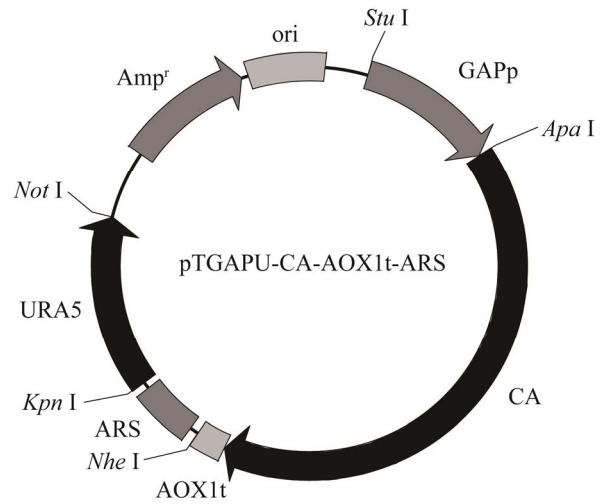


图2 咖啡酸合成初始游离载体示意图

Figure 2 Schematic diagram of the initial episomal plasmids for caffeic acid synthesis.

*TAL* 为目的基因, 以染色体上的基因 *URA3* 为内参基因, 使用 SYBR qPCR Master Mix 在实时荧光定量基因扩增仪上进行 qPCR, 引物序列见表 2。

### 1.6 测定方法

生物量以  $OD_{600}$  的值进行表达, 通过分光光度计测定。

利用高效液相色谱检测咖啡酸: 发酵液经 10 000 r/min 离心 15 min, 取 1 mL 上清液经 0.22  $\mu$ m 水系微孔滤膜过滤, 利用 HPLC 法检测发酵液中咖啡酸的浓度。色谱柱为 C18 反相色谱柱(4.6 mm $\times$ 250 mm, 5  $\mu$ m), 流动相为水:乙腈:乙酸体积比为 849:150:1, 柱温为 30  $^{\circ}$ C, 流速为 0.4 mL/min, 进样体积为 5  $\mu$ L。

### 1.7 统计分析

所有统计测试分析均使用软件 IBM SPSS Statistics 26。数据符合正态分布采用平均数 $\pm$ 标准差表示; 符合正态分布与方差齐性的两组间数据比较采用两独立样本 *t* 检验; 多组间数据比较符合条件采用单因素 ANOVA 检验, 组间两两比较采用 LSD/Duncan 法;  $P < 0.05$  表示数

据有显著性差异, 具有统计学意义。

## 2 结果与分析

### 2.1 产甘油假丝酵母系列游离载体的构建及咖啡酸合成

根据咖啡酸生物合成途径, 选择的通路基因 *TAL*、*hpaB* 和 *hpaC* 分别来源于约氏黄杆菌 (*Flavobacterium johnsoniae*)、铜绿假单胞菌 (*Pseudomonas aeruginosa*) 和肠道沙门氏菌 (*Salmonella enterica*), 合成基因片段 CA 用于 *C. glycerinogenes* 中构建咖啡酸合成途径<sup>[2]</sup>。

酵母游离载体属穿梭载体, 含表达外源基因的启动子和终止子、用于插入外源基因的单克隆或多克隆位点(multiple cloning site, MCS), 以及两套分别用于 *E. coli* 和酵母中的选择标记与自主复制序列(autonomously replicating sequence, ARS)<sup>[18]</sup>。在无天然游离质粒的 *C. glycerinogenes* 中非整合表达生产咖啡酸的首要任务是筛选出可用于游离载体构建的 ARS<sup>[19]</sup>。将选择的外源 ARS 分别命名为 KLARS<sup>[20]</sup>、414ARS、PSARS<sup>[21]</sup>、KMARS<sup>[22]</sup>和 SPARS<sup>[23]</sup> (表 3), 构建系列载体 pTGAPU-CA-AOX1t-ARS (图 2)通过 PEG/LiAc 法转化酵母进行研究。结果表明不同载体转化后产生的菌落数相差很大(表 4), 由 KLARS 构建的载体转化后所获得的转化子数最多, 达 225 个, 而由其余几种 ARS 构建的载体获得的转化子很

表 3 所获得的不同来源的 ARS

Table 3 ARS from different sources obtained

ARS	Source
KLARS	<i>Kluyveromyces lactis</i>
414ARS	<i>S. cerevisiae</i> pRS414
PSARS	<i>Pichia stipitis</i>
KMARS	<i>Kluyveromyces marxianus</i>
SPARS	<i>Schizosaccharomyces pombe</i>

少。最终选择由 KLARS 与 414ARS 构建的载体用以咖啡酸的合成, 检测两者重组菌株发酵液中咖啡酸的产量(表 4), 发现只有由 KLARS 构建的载体成功实现了咖啡酸在 *C. glycerinogenes* 中的合成, 产量达 7.8 mg/L。结果表明构建的载体性能差异较大, 利用 qPCR 检测质粒在 *C. glycerinogenes* 中的拷贝数(表 4), 发现由 KLARS 构建的载体拷贝数为 5, 其余载体拷贝数几乎检测不到。以上表明在 *C. glycerinogenes* 中通过游离载体 pTGAPU-CA-AOX1t-KLARS 实现了咖啡酸合成。

### 2.2 优化载体 ARS 位点增加咖啡酸产量

ARS 位于目的基因表达元件的不同位置会影响目的基因的表达<sup>[24]</sup>, 探究该游离载体 KLARS 的位置对咖啡酸合成的影响, 以 KLARS 位于目的基因表达元件下游的 pTGAPU-CA-AOX1t-KLARS (dn-KLARS)为模板, 改变 KLARS 位置获得位于目的基因表达元件上游与对侧的载体 up-KLARS 和 op-KLARS (图 3A), 用于咖啡酸

表 4 不同载体的转化子数、拷贝数和对应咖啡酸产量

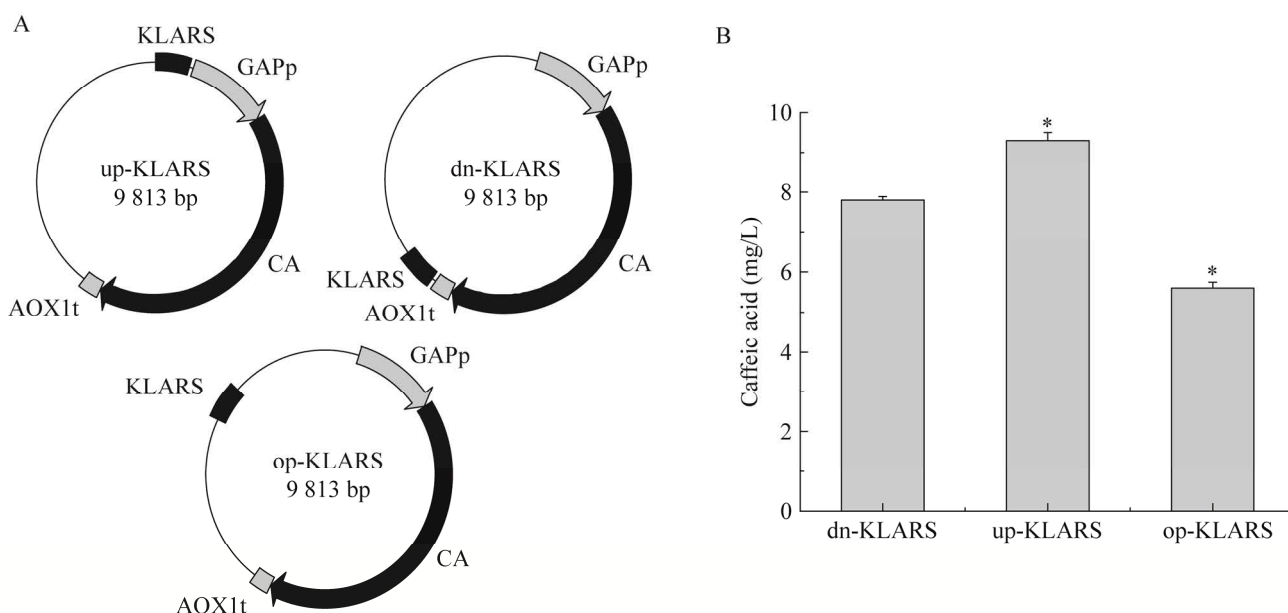
Table 4 Number of transformants, copy number and caffeic acid yield of different vectors

Plasmid	Number of transformers/5 $\mu$ g DNA	Caffeic acid yield (mg/L)	Copy number
pTGAPU-CA-AOX1t-KLARS	225 $\pm$ 13*	7.80 $\pm$ 0.09	5.0 $\pm$ 0.2
pTGAPU-CA-AOX1t-414ARS	19 $\pm$ 8	—	—
pTGAPU-CA-AOX1t-PSARS	6 $\pm$ 3	—	—
pTGAPU-CA-AOX1t-KMARS	7 $\pm$ 4	—	—
pTGAPU-CA-AOX1t-SPARS	5 $\pm$ 3	—	—

注: 表中数据为平均值 $\pm$ 标准差; —: 数值无法检测或未检测; \*:  $P < 0.05$

Note: Data in table are mean $\pm$ standard deviation. —: Not detected or not tested. Compared with other plasmids, \*:  $P < 0.05$ .





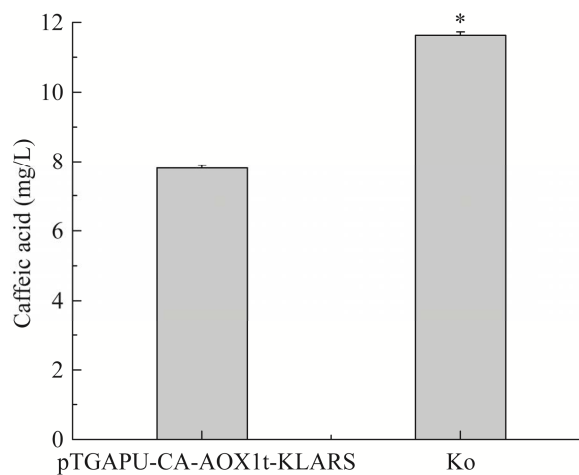
**图3 KLARS 位置对咖啡酸产量的影响** A: KLARS 位于载体不同位置的示意图; B: KLARS 位于载体不同位置时的咖啡酸产量。图中显示 3 次独立实验的平均值±标准差; \*:  $P < 0.05$

Figure 3 Effect of KLARS location on caffeic acid yield. A: Schematic diagram of KLARS at different positions of plasmid; B: Caffeic acid yield of KLARS at different positions of plasmid. Graphs show the mean±standard deviation of three independent experiments; Compared with pTGAPU-CA-AOX1t-KLARS, \*:  $P < 0.05$ .

合成。结果显示, up-KLARS 获得的产量最高, 达 9.3 mg/L, 较 dn-KLARS 提高约 19%; 相反地, op-KLARS 获得的产量比 dn-panARS 降低约 28%, 为 5.6 mg/L(图 3B), 表明 KLARS 位于目的基因表达元件上游时对咖啡酸合成最有益。

### 2.3 利用 Kozak 序列促进咖啡酸合成

Kozak 序列是真核生物 mRNA 起始密码子 AUG 两侧的序列, 其与翻译起始因子结合而影响 mRNA 的翻译<sup>[25]</sup>。为保证氨基酸序列不发生变化, 在载体 pTGAPU-CA-AOX1t-KLARS 目的基因(CA)起始密码子上游添加 Kozak 序列获得载体 Ko, 检测改造前后两者重组菌株发酵液的咖啡酸产量。结果显示, 改造后咖啡酸产量为 11.6 mg/L(图 4), 较改造前提升约 49%, 表明利用 Kozak 序列能明显促进 *C. glycerinogenes* 的咖啡酸合成。



**图4 Kozak 序列对咖啡酸产量的影响** 图中显示 3 次独立实验的平均值±标准差; \*:  $P < 0.05$

Figure 4 Effect of Kozak sequences on caffeic acid yield. Graphs show the mean±standard deviation of three independent experiments; Compared with pTGAPU-CA-AOX1t-KLARS, \*:  $P < 0.05$ .

## 2.4 调节 *URA5* 启动子长度提高咖啡酸产量

调节载体标记基因的启动子长度可以增加载体拷贝数, 从而提高产物产量<sup>[26]</sup>。由标记基因 *URA5* 启动子序列可知(图 5A), 以起始密码子 ATG 中 A 为+1, -65、-321、-415 这 3 个位点为 TATA 盒<sup>[27]</sup>, -117 是转录起始位点 TSS<sup>[28]</sup>。对载体 pTGAPU-CA-AOX1t-KLARS (T0) 的 *URA5* 启动子分别截短 250、500 和 700 bp, 得到载体 T250、T500 和 T700。结果表明(图 5B): *URA5* 启动子截短 250 bp 时, 3 个位点的 TATA

盒均未受影响, 咖啡酸产量为 14 mg/L, 较初始 T0 提升约 79%; 截短 500 bp 时, 失去-415 与 -321 两个位点的 TATA 盒, 产量为 4.3 mg/L, 降低约 45%; 截短 700 bp 致所有 TATA 盒和 TSS 位点丧失且无转化子。进一步提取含 T0、T250 和 T500 的重组菌株的总 DNA, 利用 qPCR 测定质粒拷贝数(图 5C), 可知 T250 拷贝数最高约为 9, T0 拷贝数约为 5, 而 T500 拷贝数约为 3, 这表明调节 *URA5* 启动子长度可以增加载体拷贝数, 从而提高咖啡酸产量。

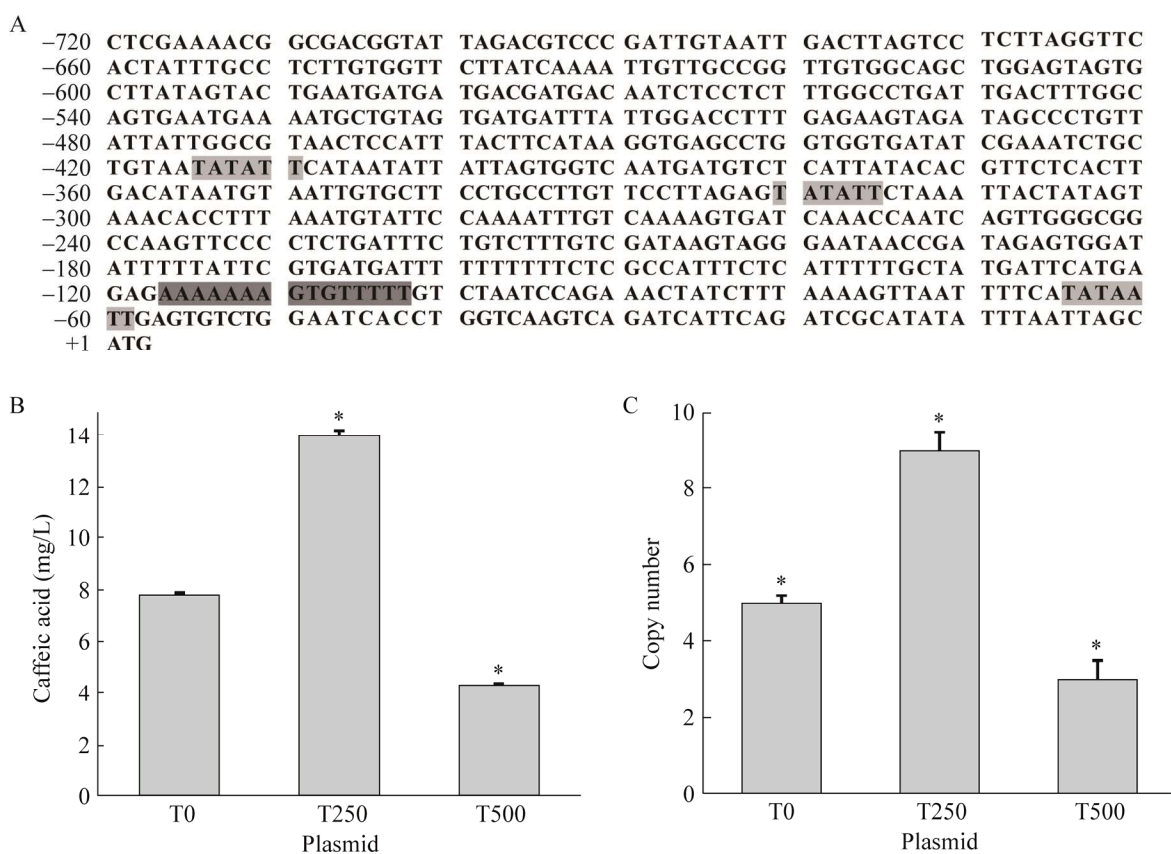


图 5 *URA5* 启动子不同程度截短对咖啡酸产量的影响 A: *URA5* 启动子序列; B: *URA5* 启动子截短前后的咖啡酸产量; C: *URA5* 启动子截短前后载体的拷贝数。图中显示 3 次独立实验的平均值±标准差; \*:  $P < 0.05$

Figure 5 Effect of different levels of truncation of the *URA5* promoter on the yield of caffeic acid. A: *URA5* promoter sequence; B: Caffeic acid yield before and after *URA5* promoter truncation; C: The copy number of plasmid before and after truncation of *URA5* promoter. Graphs show the mean±standard deviation of three independent experiments; Compared with pTGAPU-CA-AOX1t-KLARS, \*:  $P < 0.05$ .

## 2.5 不同启动子与终止子对咖啡酸合成的影响

启动子作为基本的顺式作用元件可以控制目标基因转录强度<sup>[29]</sup>, 终止子负责终止基因的转录, 其强弱会影响 mRNA 转录的终止效率<sup>[30]</sup>。不同诱导方式的启动子或不同终止子会影响目的基因的转录从而对产物的合成产生影响。

课题组前期研究中获得了多种典型诱导启动子。为探究多种诱导启动子在该游离载体上启动基因表达合成咖啡酸的可行性, 比较不同条件下对应诱导启动子对咖啡酸产量的影响, 利用本课题组库藏的 *C. glycerinogenes* 启动子: 渗透压诱导的 GAPp<sup>[13]</sup>、GPDp<sup>[14]</sup>和 STL3p<sup>[15]</sup>, 酸诱导的 GMTp<sup>[16]</sup>和高温诱导的 CWP1p<sup>[17]</sup>, 以 pTGAPU-CA-AOX1t-KLARS 为模板构建含不同启动子的载体, 将含不同载体的重组菌株在特定条件下培养 24 h 检测咖啡酸产量。结果表明利用 GAPp、GPDp、STL3p、GMTp 和 CWP1p

均检测到咖啡酸的合成, 产量分别为 7.8、5.1、4.0、3.1 和 3.5 mg/L (图 6A)。该结果说明不同诱导启动子在对应诱导条件下均能实现咖啡酸的合成, 而且 GAPp 启动强度最高。

为研究不同终止子对咖啡酸产量的影响, 以 pTGAPU-CA-AOX1t-KLARS (AOX1t) 为模板, 利用不同终止子替换 AOX1t, 终止子分别为: *S. cerevisiae* 终止子 CYC1t 和 *C. glycerinogenes* 内源终止子 URA5t、TRP1t, 分别针对终止子的效率元件、连接序列 1、定位元件、连接序列 2 和 poly A 位点等元素优化后合成的简短终止子 1t、2t、3t、4t 和 5t, 以及报道的最简终止子 G1t<sup>[31]</sup>。结果显示(图 6B)利用 URA5t 获得的咖啡酸产量最高, 达 29.1 mg/L, 是初始产量的 3.73 倍, TRP1t 获得的咖啡酸产量为 11.9 mg/L, 是初始产量的 1.53 倍, 而使用外源酵母终止子与合成终止子时咖啡酸产量则普遍偏低。总而言之, 各个终止子均能完成转录终止的作用而合成咖啡酸, 同时内源终止子表现出了显著优势。

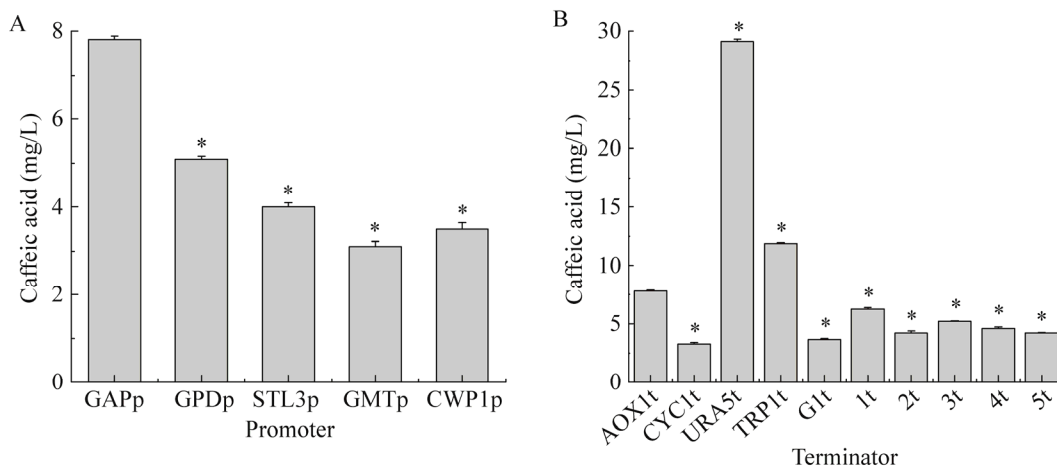


图 6 不同启动子与终止子对咖啡酸产量的影响 A: 不同启动子对咖啡酸产量的影响; B: 不同终止子对咖啡酸产量的影响。图中显示 3 次独立实验的平均值±标准差; \*:  $P < 0.05$

Figure 6 Effect of different promoters and terminators on caffeic acid yield. A: Effect of different promoters on caffeic acid yield; B: Effect of different terminators on caffeic acid yield. Graphs show the mean±standard deviation of three independent experiments; Compared with pTGAPU-CA-AOX1t-KLARS, \*:  $P < 0.05$ .

### 3 讨论与结论

利用非整合方式表达合成咖啡酸的基因, 关键是获得在 *C. glycerinogenes* 中能自我复制的游离载体。ARS 赋予载体独立于染色体外的自主复制能力<sup>[19]</sup>, 筛选 ARS 是构建游离载体必不可少的一步。本研究从不同来源的 5 个 ARS 中选择的 KLARS 所构建的载体转化子数与拷贝数均最高, 而这在酵母中是高性能游离载体的直观表现<sup>[32]</sup>, 由于 KLARS 来源于乳酸克鲁维酵母, 相较于本研究中其他几株酵母, *C. glycerinogenes* 与乳酸克鲁维酵母有更高的相似性, 因此推测 KLARS 在 *C. glycerinogenes* 中更容易被识别并起作用。研究发现 KLARS 在载体中的位置对表达合成咖啡酸有明显影响, 而且 KLARS 位于目的基因表达元件上游时对咖啡酸合成能力促进最大, 这与有关报道表明 ARS 位于下游时更有助于目的基因表达<sup>[24]</sup>存在不同, 这可能是 KLARS 在 *C. glycerinogenes* 中更有益于增强内源启动子 GAPp 的性能, 或可能 KLARS 位于目的基因表达元件上游时更适合目的基因在 *C. glycerinogenes* 中表达, 因此 ARS 位置应根据不同基因和宿主进行调整以适应表达要求。研究表明真核生物中 Kozak 序列对翻译起始有重要作用<sup>[25]</sup>, 游离载体融入 Kozak 序列后可强化咖啡酸合成的基因表达而使产物明显增加。载体标记基因启动子截短策略已被广泛应用于增加基因表达量而提高产物产量<sup>[26]</sup>, 所构建的 pTGAPU-CA-AOX1t-KLARS 中 *URA5* 启动子截短 250 bp 时咖啡酸产量明显提高, 研究表明载体标记基因启动子部分缺陷会导致标记基因表达强度降低, 为了应对筛选条件, 菌体需更高拷贝数的载体弥补单个标记基因降低的表达量, 从而获得含更高拷贝数载体的重组菌株<sup>[26]</sup>, 但这种截短强度应考虑不可过多涉及 TATA 盒,

特别不能删除 TSS 序列, 本研究利用 qPCR 也证实咖啡酸产量的提高是由于载体拷贝数的增加。研究中所选启动子均来自 *C. glycerinogenes*, 其在游离载体上均能实现目的基因的表达, 考察这些诱导启动子的性能有助于了解不同启动子所构建游离载体的实际应用。此外, 内源终止子性能明显强于外源终止子, 利用 *URA5t* 时咖啡酸产量最高, 这表明 *C. glycerinogenes* 偏好于自身来源的终止子序列, 值得注意的是过于简短的合成终止子序列可能会影响转录终止的修饰而降低终止效率<sup>[31]</sup>, 这在本研究中通过表达合成咖啡酸基因得到验证。

总而言之, 本研究构建并优化用于 *C. glycerinogenes* 咖啡酸合成的游离载体, 使产量达 29.1 mg/L, 明显高于课题组前期整合表达获得的初始咖啡酸产量(14.71 mg/L)<sup>[11]</sup>, 另一方面也高于 *E. coli* 或 *S. cerevisiae* 利用商业载体从葡萄糖直接获得的初始产量(12.1 mg/L<sup>[2]</sup>或 0.183 mg/L<sup>[5]</sup>), 这表明了 *C. glycerinogenes* 中利用非整合表达体系改造代谢途径高产咖啡酸的潜力。游离载体的获得为今后在 *C. glycerinogenes* 中调控生产提供了分子操作工具, 同时为其他无游离质粒菌株构建游离载体和调控生产提供了参考。

### REFERENCES

- [1] Zhang HR, Stephanopoulos G. Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars[J]. *Applied Microbiology and Biotechnology*, 2013, 97(8): 3333-3341
- [2] Lin YH, Yan YJ. Biosynthesis of caffeic acid in *Escherichia coli* using its endogenous hydroxylase complex[J]. *Microbial Cell Factories*, 2012, 11: 42
- [3] Rodrigues JL, Araújo RG, Prather KLJ, Kluskens LD, Rodrigues LR. Heterologous production of caffeic acid from tyrosine in *Escherichia coli*[J]. *Enzyme and Microbial Technology*, 2015, 71: 36-44
- [4] 王震, 李霞, 元英进. 微生物异源合成咖啡酸及其酯

- 类衍生物研究进展[J]. 中国生物工程杂志, 2020, 40(7): 91-99
- Wang Z, Li X, Yuan YJ. Advances in production of caffeic acid and its ester derivatives in heterologous microbes[J]. China Biotechnology, 2020, 40(7): 91-99 (in Chinese)
- [5] Li YZ, Mao JW, Liu QL, Song XF, Wu YZ, Cai M, Xu HJ, Qiao MQ. De novo biosynthesis of caffeic acid from glucose by engineered *Saccharomyces cerevisiae*[J]. ACS Synthetic Biology, 2020, 9(4): 756-765
- [6] Zhuge J, Fang HY, Wang ZX, Chen DZ, Jin HR, Gu HL. Glycerol production by a novel osmotolerant yeast *Candida glycerinogenes*[J]. Applied Microbiology and Biotechnology, 2001, 55(6): 686-692
- [7] Zhao ML, Shi DC, Lu XY, Zong H, Zhuge B, Ji H. Ethanol fermentation from non-detoxified lignocellulose hydrolysate by a multi-stress tolerant yeast *Candida glycerinogenes* mutant[J]. Bioresource Technology, 2019, 273: 634-640
- [8] Zhao ML, Shi DC, Lu XY, Zong H, Zhuge B. Co-production of 1,2,4-butanetriol and ethanol from lignocellulose hydrolysates[J]. Bioresource Technology, 2019, 282: 433-438
- [9] Wang YQ, Zhang ZY, Lu XY, Zong H, Zhuge B. Genetic engineering of an industrial yeast *Candida glycerinogenes* for efficient production of 2-phenylethanol[J]. Applied Microbiology and Biotechnology, 2020, 104(24): 10481-10491
- [10] 赵美琳, 诸葛斌, 陆信曜, 宗红, 施丁昌. 工业酵母抗逆机理研究进展[J]. 微生物学通报, 2019, 46(5): 1155-1164
- Zhao ML, Zhuge B, Lu XY, Zong H, Shi DC. Research progress in stress tolerance of industrial yeasts[J]. Microbiology China, 2019, 46(5): 1155-1164 (in Chinese)
- [11] Wang XH, Zhao C, Lu XY, Zong H, Zhuge B. Production of caffeic acid with co-fermentation of xylose and glucose by multi-modular engineering in *Candida glycerinogenes*[J]. ACS Synthetic Biology, 2022, 11(2): 900-908
- [12] 宋保平, 诸葛斌, 方慧英, 诸葛健. 产甘油假丝酵母 (*Candida glycerinogenes*)染色体倍性分析[J]. 微生物学报, 2011, 51(3): 326-331
- Song BP, Zhuge B, Fang HY, Zhuge J. Analysis of the chromosome ploidy of *Candida glycerinogenes*[J]. Acta Microbiologica Sinica, 2011, 51(3): 326-331 (in Chinese)
- [13] Zhang C, Zhuge B, Zhan XB, Fang HY, Zong H, Zhuge J. Cloning and characterization of a novel NAD(+)-dependent glyceraldehyde-3-phosphate dehydrogenase gene from *Candida glycerinogenes* and use of its promoter[J]. Yeast: Chichester, England, 2013, 30(4): 157-163
- [14] 丁春生, 饶志明, 诸葛斌, 沈微, 陈献忠, 方慧英, 诸葛健. 利用荧光蛋白研究产甘油假丝酵母胞浆 3-磷酸甘油脱氢酶基因 *CgGPD* 启动子[J]. 微生物学报, 2008, 48(8): 1013-1018
- Ding CS, Rao ZM, Zhuge B, Shen W, Chen XZ, Fang HY, Zhuge J. Analysis of *CgGPD* gene promoter from *Candida glycerinogenes* by fluorescent protein[J]. Acta Microbiologica Sinica, 2008, 48(8): 1013-1018 (in Chinese)
- [15] 朱佳莉, 诸葛斌, 方慧英, 宗红, 陆信曜, 张成, 张炜极. 新型渗透压调控的工业酵母启动子[J]. 微生物学报, 2015, 55(11): 1385-1391
- Zhu JL, Zhuge B, Fang HY, Zong H, Lu XY, Zhang C, Zhang WJ. New osmo-regulatory promoters in the industrial yeast[J]. Acta Microbiologica Sinica, 2015, 55(11): 1385-1391 (in Chinese)
- [16] 诸葛斌, 周家豪, 何强, 陆信曜, 宗红. 一种在酸条件下诱导表达的启动子: 中国, CN107828790B[P]. 2020-06-09
- Zhuge B, Zhou JH, He Q, Lu XY, Zong H. Promoter for inducing expression under acidic conditions: China, CN107828790B[P]. 2020-06-09 (in Chinese)
- [17] Wang YD, Lin YP, Lu XY, Zhuge B, Zong H. Selection and application of novel high temperature inducible promoters in the tolerant yeast *Candida glycerinogenes*[J]. Journal of Bioscience and Bioengineering, 2020, 130(1): 1-5
- [18] Gnügge R, Rudolf F. *Saccharomyces cerevisiae* shuttle vectors[J]. Yeast, 2017, 34(5): 205-221
- [19] Newlon CS, Theis JF. The structure and function of yeast ARS elements[J]. Current Opinion in Genetics & Development, 1993, 3(5): 752-758
- [20] Liachko I, Dunham MJ. An autonomously replicating sequence for use in a wide range of budding yeasts[J]. FEMS Yeast Research, 2014, 14(2): 364-367
- [21] Yang VW, Marks JA, Davis BP, Jeffries TW. High-efficiency transformation of *Pichia stipitis* based on its *URA3* gene and a homologous autonomous replication sequence, *ARS2*[J]. Applied and Environmental Microbiology, 1994, 60(12): 4245-4254
- [22] Iborra F, Ball MM. *Kluyveromyces marxianus* small DNA fragments contain both autonomous replicative and centromeric elements that also function in

- Kluyveromyces lactis*[J]. Yeast, 1994, 10(12): 1621-1629
- [23] Brun C, Dubey DD, Huberman JA. pDblet, a stable autonomously replicating shuttle vector for *Schizosaccharomyces pombe*[J]. Gene, 1995, 164(1): 173-177
- [24] Camattari A, Goh A, Yip LY, Tan AHM, Ng SW, Tran A, Liu GW, Liachko I, Dunham MJ, Rancati G. Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications[J]. Microbial Cell Factories, 2016, 15(1): 139
- [25] Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes[J]. Cell, 1986, 44(2): 283-292
- [26] Kazemi Seresht A, Nørgaard P, Palmqvist EA, Andersen AS, Olsson L. Modulating heterologous protein production in yeast: the applicability of truncated auxotrophic markers[J]. Applied Microbiology and Biotechnology, 2013, 97(9): 3939-3948
- [27] Basehoar AD, Zanton SJ, Pugh BF. Identification and distinct regulation of yeast TATA box-containing genes[J]. Cell, 2004, 116(5): 699-709
- [28] Zhang ZH, Dietrich FS. Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE[J]. Nucleic Acids Research, 2005, 33(9): 2838-2851
- [29] Tang HT, Wu YL, Deng JL, Chen NZ, Zheng ZH, Wei YJ, Luo XZ, Keasling JD. Promoter architecture and promoter engineering in *Saccharomyces cerevisiae*[J]. Metabolites, 2020, 10(8): 320
- [30] Wei LN, Wang ZX, Zhang GL, Ye BC. Characterization of terminators in *Saccharomyces cerevisiae* and an exploration of factors affecting their strength[J]. ChemBioChem, 2017, 18(24): 2422-2427
- [31] Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. Short synthetic terminators for improved heterologous gene expression in yeast[J]. ACS Synthetic Biology, 2015, 4(7): 824-832
- [32] Abdel-Banat BMA, Hoshida H, Akada R. Various short autonomously replicating sequences from the yeast *Kluyveromyces marxianus* seemingly without canonical consensus[J]. Current Research in Microbial Sciences, 2021, 2: 100053