

研究报告

大肠杆菌染色体上严谨型启动子的构建

仇焕娜^{1,2,3} 赵东东^{2,3} 满淑丽¹ 毕昌昊^{2,3} 朱欣娜^{2,3*} 张学礼^{2,3*}

(1. 天津科技大学生物工程学院 天津 310018)

(2. 中国科学院天津工业生物技术研究所 天津 300308)

(3. 中国科学院系统微生物工程重点实验室 天津 300308)

摘要:【背景】启动子的渗漏表达是代谢工程和合成生物学较为关注的问题,探索严谨型启动子使之能像开关一样控制基因的表达有助于解决这一问题。【目的】为避免在质粒上研究启动子带来的弊端,本研究将在染色体上对严谨型启动子进行构建和评价。【方法】基于4种调控元件四环素 tetO、乳糖 lacO、阿拉伯糖 araC 和鼠李糖 rhaR 的序列,以及2种来源的启动子 PL 和 Plac 序列,设计和组合构建了6个启动子 PtetO2、PtetO3、PlacO2、PlacO3、PlacO+ara 和 PlacO+rha。应用 CRISPR/Cas9 系统将这6个启动子序列整合到大肠杆菌 ATCC 8739 染色体上,利用绿色荧光蛋白(Green fluorescent protein, GFP)的表达,分析这6个启动子的相对表达强度和严谨型控制情况。【结果】GFP 表达分析显示,启动子 PlacO+rha 为最佳严谨型启动子,在无诱导剂时表达为 0.02,有诱导剂时最大表达强度为 lacZ 基因启动子的 12 倍,相对控制范围为 600 倍。【结论】研究结果将为代谢工程和合成生物学中的精确调控基因表达奠定良好的应用基础。

关键词: 严谨型, 启动子, 绿色荧光蛋白, 大肠杆菌

Construction of promoters with tight regulation on chromosome of *Escherichia coli*

QIU Huan-Na^{1,2,3} ZHAO Dong-Dong^{2,3} MAN Shu-Li¹ BI Chang-Hao^{2,3}
ZHU Xin-Na^{2,3*} ZHANG Xue-Li^{2,3*}

(1. College of Biotechnology, Tianjin University of Science and Technology, Tianjin 310018, China)

(2. Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China)

(3. Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China)

Abstract: [Background] The leakage expression of promoters is a major concern in metabolic engineering and synthetic biology. To explore promoters with tight regulation like a switch would resolve this problem. [Objective] In order to avoid the shortcomings of using plasmid as research

Foundation items: National Natural Science Foundation of China (31770059, 31300089)

*Corresponding authors: ZHANG Xue-Li: Tel: 86-22-84861983; E-mail: zhang_xl@tib.cas.cn

ZHU Xin-Na: Tel: 86-22-84861946; E-mail: zhu_xn@tib.cas.cn

Received: October 19, 2017; **Accepted:** December 28, 2017; **Published online** (www.cnki.net): January 08, 2018

基金项目: 国家自然科学基金(31770059, 31300089)

*通信作者: 张学礼: Tel: 86-22-84861983; E-mail: zhang_xl@tib.cas.cn

朱欣娜: Tel: 86-22-84861946; E-mail: zhu_xn@tib.cas.cn

收稿日期: 2017-10-19; 接受日期: 2017-12-28; 网络首发日期(www.cnki.net): 2018-01-08

auxiliary, the aim of this study is to construct and assess promoter with tight regulation on chromosome. **[Methods]** Four elements including tetO, lacO, araC and rhaR regulated by TetR, LacI, AraC and RhaR respectively, were selected for designing six promoters including PtetO2, PtetO3, PlacO2, PlacO3, PlacO+ara and PlacO+rha. CRISPR/Cas9 system was applied to integrate these promoters into chromosome of *Escherichia coli* ATCC 8739. Promoter strength and relative expression with or without inducing agents were determined by the fluorescent quantities of GFP protein. **[Results]** Hybrid promoter PlacO+rha was the best one with tight regulation. The expression strength of PlacO+rha was 0.02 in absence of the inducer, and the maximum expression intensity was 12-fold higher than that of the induced *lacZ* promoter, of which the range of expression level was 600-fold. **[Conclusion]** This study would provide a good foundation for precisely regulating gene expression in metabolic engineering and synthetic biology.

Keywords: Tight regulation, Promoter, Green fluorescent protein, *Escherichia coli*

启动子是位于结构基因 5'-上游的 DNA 序列, 用于 RNA 聚合酶特异识别和结合。它控制着基因表达(转录)的起始时间和程度^[1-2]。细菌的基因表达调控主要发生在转录水平上, 而启动子在转录水平调控中居于关键的地位。随着合成生物学和代谢工程的发展, 细菌启动子获得了越来越多的应用。利用启动子可以调控基因的表达强度, 改造细菌的代谢通路, 或者方便地表达目的蛋白, 使其在细胞工厂和酶催化中应用广泛^[3]。

诱导型启动子 Plac、Ptac 和 Ptrc 其上含有 lacO 调控序列, 当无诱导剂存在时, 阻遏蛋白 LacI 和调控序列结合, 抑制启动子的表达^[4]。在常规的大肠杆菌中, LacI 的表达量不高, 导致非诱导条件下较高的表达, 系统严谨性不强。虽然突变的阻遏蛋白 LacIq 能部分解除非诱导条件下的高表达, 但仍然不够严谨^[5]。常用的 λ 噬菌体 PL 是一个组成型启动子, 其表达不需要诱导剂, 但这种启动子的表达不受控制^[6]。实际应用中生物合成途径中某些基因的表达, 需要对表达时间和表达量进行控制, 而不受控制的启动子很难实现这一点, 因此, 有必要构建出严谨型的启动子, 使基因的表达在可调控下进行。

人们常常在质粒上研究启动子的强度, 用于代谢途径中关键基因的表达, 实现提高目标产品的生产能力^[7]。然而质粒用于表达具有诸多弊端^[8]: 质粒的维持对宿主细胞会造成很大的代谢负荷, 尤其

是高拷贝的质粒; 很多质粒的遗传稳定性不好, 不利于工业化大规模的生产; 只有低拷贝数质粒的复制和细胞的繁殖是同步进行的, 因此也只有它们所有的细胞中保持一致的拷贝数; 很多化合物的合成需要在细胞中构建出一条复杂的代谢途径, 因此需要引入包含多个基因的长 DNA 片段, 而大部分质粒都比较难携带长 DNA 片段。

在质粒上研究启动子和基因的表达时, 由于含质粒的工程菌株不稳定性, 在确定了目标基因的表达水平后, 最终的目标还是将启动子和目标基因整合于染色体上, 但是整合于染色体上的基因表达强度会和质粒上有偏差, 在某个启动子控制下基因在染色体上的表达不是最优状态。因此, 通过在细胞染色体上直接研究启动子强度来调控基因表达的策略和质粒过表达相比有很多优势。

基因编辑技术 CRISPR/Cas9 已经成功地应用于大肠杆菌, 实现了基因的插入、删除和替换^[9-10]。我们课题组已经开发出一种高效的基因组编辑技术, 只需一个质粒的构建和转化的 CRISPR/Cas9^[11]就能对基因进行编辑。CRISPR/Cas9 系统不仅消除了基因组编辑选择标记, 也提高了编辑工作效率^[12]。在本研究中我们利用 CRISPR/Cas9 将启动子整合于染色体上进行表达分析。

基于严谨型启动子对于合成生物学和代谢工程中的重要作用, 本研究基于 4 种调控元件四环素 tetO、乳糖 lacO、阿拉伯糖 araC 和鼠李糖 rhaR 的

结合序列, 设计和组合构建了 6 个启动子 PtetO2、PtetO3、PlacO2、PlacO3、PlacO+ara 和 PlacO+rha。应用 CRISPR/Cas9 系统将这 6 个启动子序列整合到大肠杆菌 ATCC 8739 染色体上, 利用绿色荧光蛋白 GFP 的表达分析了这 6 个启动子的相对表达

强度和严谨型控制情况。

1 材料与方法

1.1 菌种来源

本实验所用菌株和质粒见表 1。

表 1 实验所用的菌株和质粒

Table 1 The strains and plasmids used in this work

菌株和质粒 Strains and plasmids	遗传特征 Genetic characteristics	来源 Source
Strains		
<i>E. coli</i> ATCC 8739	Wild type	Lab collection
M1-46	ATCC 8739, FRT-km-FRT::M1-46::lacZ	[8]
S-PlacZ	ATCC 8739, Δ lacZ-PlacZ-gfp	This work
S-PlacO2	ATCC 8739, Δ lacZ-PlacO2-gfp	This work
S-PlacO3	ATCC 8739, Δ lacZ-PlacO3-gfp	This work
S- Δ araBAD	ATCC 8739, Δ araBAD	This work
S-PlacO+ara	ATCC 8739, Δ araBAD, Δ lacZ-PlacO+ara-gfp,	This work
S- Δ rhaBAD	ATCC 8739, Δ rhaBAD	This work
S-PlacO+rha	ATCC 8739, Δ rhaBAD, Δ lacZ-PlacO+rha-gfp	This work
S-tetR	ATCC 8739, Δ rhaBAD::M1-46-tetR	This work
S-PtetO2	ATCC 8739, Δ rhaBAD::M1-46-tetR, Δ lacZ-PtetO2-gfp	This work
S-PtetO3	ATCC 8739, Δ rhaBAD::M1-46-tetR, Δ lacZ-PtetO3-gfp	This work
Plasmids		
pUC57-Kan	plac, ColEI, kan	Lab collection
pACYC184-M	cat, replace tet with lacI and P _{trc} of pTrc99A-M	[13]
pRed_Cas9_ΔpoxB300	Derived from pKD46, <i>exo</i> , <i>bet</i> , <i>gam</i> , arabinose operon, Cas9, gRNA with N20 and with homologous arms	[11]
pCmGFP	A plasmid containing <i>gfp</i> gene	[14]
pSC101	<i>tet</i> , SC101, <i>repA</i>	Lab collection
pGFP	A plasmid modified from pCmGFP by removing the <i>Bsa</i> I restriction endonuclease site of <i>gfp</i> gene	This work
placZ	cat and P15A from pACYC184-M, N20-gRNA targeting <i>lacZ</i> gene expressed under constitute promoter, <i>lacZ</i> gene and its up and down homologous hand	This work
pΔlacZ-PlacZ-gfp	<i>lacZ</i> gene was replaced by <i>gfp</i> gene using placZ-N20 as template	This work
pΔlacZ-PlacO2-gfp	P _{native-lacZ} promoter was replaced by PlacO2 promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔlacZ-PlacO3-gfp	P _{native-lacZ} promoter was replaced by PlacO3 promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔlacZ-PlacO+ara-gfp	P _{native-lacZ} promoter was replaced by PlacO+ara promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔlacZ-PlacO+rha-gfp	P _{native-lacZ} promoter was replaced by PlacO+rha promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔlacZ-PlacO+tetO2-gfp	P _{native-lacZ} promoter was replaced by PlacO+tetO2 promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔlacZ-PlacO+tetO3-gfp	P _{native-lacZ} promoter was replaced by PlacO+tetO3 promoter using pΔlacZ-P _{native-lacZ} -gfp as template	This work
pΔaraBAD-N20	cat, P15A, N20-gRNA targeting <i>araBAD</i> gene expressed under constitute promoter, the up and down homologous sequence of <i>araBAD</i> gene, which used for deletion <i>araBAD</i> gene	This work
pΔrhaBAD-N20	cat, P15A, N20-gRNA targeting <i>rhaBAD</i> gene expressed under constitute promoter, the up and down homologous sequence of <i>rhaBAD</i> gene, which used for deletion <i>rhaBAD</i> gene	This work
pΔrhaBAD::M1-46-tetR	M1-46-tetR gene was inserted between the up and down homologous sequence of <i>rhaBAD</i> gene, which used for integrated <i>tetR</i> gene	This work
pRedCas9	kan, P _{BAD} -cas9, P _{BAD} -Red-recA, repA101(Ts)	[11]

1.2 培养基^[15]和培养条件

LB 培养基(g/L): 酵母膏 5.0, 蛋白胨 10.0, 氯化钠 10.0, 琼脂 15.0 (固体培养基)大肠杆菌于 37 °C、250 r/min 培养。

1.3 主要试剂和仪器

抗生素和诱导剂, Sigma-Aldrich 公司。抗生素的使用浓度分别为: 氨苄霉素 100 µg/mL, 卡那霉素 50 µg/mL, 氯霉素 34 µg/mL。诱导剂: IPTG、脱水四环素、L(+)阿拉伯糖、L(+)鼠李糖。PCR 高保真扩增 Phusion 酶、快速 T4 DNA 连接酶、限制性内切酶 *Bsa* I, NEB 生物公司; Trans T1 感受态细胞, 全式金生物技术有限公司。离心机, Sigma 实验室离心机公司; 荧光显微镜, 奥林巴斯公司; 荧光强度分析酶标仪, TECAN 酶标仪公司; 全自动凝胶成像仪, Bio-Rad 公司。

1.4 启动子合成

6 条启动子 PtetO2、PtetO3、PlacO2、PlacO3、PlacO+ara 和 PlacO+rha (含 RBS 结合位点, 表 2) 分别全基因合成于质粒 pUC57-Kan 上。为方便后续克隆, 每条启动子的上下游加了 20–30 bp 的特殊序列: 上游为 5'-TGGTTAATGACAGTGGGACGggtctca-3', 下游为 5'-tgagaccAACCTCGGCATTGCGTAA-3', 斜体小写为 *Bsa* I 酶切位点。

1.5 启动子质粒的构建

1.5.1 质粒 pΔlacZ-PlacZ-gfp 的构建

质粒 pΔlacZ-PlacZ-gfp 含有在 PlacZ 启动子下的 *gfp* 基因, 及针对 *lacZ* 基因的 N20-gRNA。为构建质粒 pΔlacZ-PlacZ-gfp, 首先构建 placZ 质粒和 pGFP 质粒。

placZ 质粒含有氯霉素抗性基因 *cat*, 复制子 P15A, 识别 *lacZ* 基因的 lacZN20-gRNA 序列, *lacZ* 基因和该基因上下游同源臂序列。其中, 氯霉素抗性基因 *cat* 和复制子 P15A 序列, 以质粒 pACYC184-M^[13] 为模板, 用引物对 N20-B-F1/N20-B-R1 扩增获得; lacZN20-gRNA 的片段以质

粒 pRed_Cas9_ΔpoxB300^[11] 为模板, 用引物 lacZ-N20-B-F2/N20-B-R2 扩增; *lacZ* 基因及同源臂序列, 以 *E. coli* ATCC 8739 DNA 为模板, 用引物 lacZ-F1/lacZ-R1 扩增得到。这 3 个片段用 Golden Gate 技术策略^[16] 进行连接, 转化 TransT1 感受态细胞, 得到的克隆用引物 N20-B-YZ-F/lacZ-R1 进行验证, 阳性克隆提取质粒 DNA 送样测序分析, 得到正确的 placZ 质粒。

pGFP 质粒含有 *gfp* 基因, 为使 *gfp* 基因序列不含有 *Bsa* I 酶切位点, 从而方便运用 Golden Gate 策略, 对出发质粒 pCmGFP 上 *gfp* 的 *Bsa* I 酶切位点的序列用简并碱基进行了替换。以 pCmGFP 质粒为模板, 用引物 Gfp-GGA-F/Gfp-GGA-R (引物序列上含有改变的碱基, 且两条引物上含有互补序列) 进行 PCR 扩增, 得到的片段用 *Dpn* I 处理后, 直接转化 TransT1 感受态细胞, 提取质粒 DNA 测序分析, 得到正确的 pGFP 质粒。

在上述 2 个质粒构建的基础上, 构建 pΔlacZ-PlacZ-gfp 质粒。具体过程如下: 以 placZ 质粒为模板, 用引物 lacZ-B-F2/lacZ-B-R2 进行反向扩增, 得到含有 *cat*、P15A、lacZN20-gRNA、PlacZ 和 *lacZ* 基因外上下游同源臂序列片段; 以 pGFP 质粒为模板, 用引物 gfp-F/gfp-R 扩增得到 *gfp* 基因序列。用同样的 Golden Gate 策略将这 2 个片段连接起来, 转化 TransT1 感受态细胞, 得到的克隆用引物 lacZ-F1/gfp-YZ-down200 进行验证, 阳性克隆提取质粒 DNA 送样测序分析, 得到正确的 pΔlacZ-PlacZ-gfp 质粒。

所用引物见表 2。

1.5.2 启动子表达质粒系列 pΔlacZ-Pxx-gfp 的构建

分别以 6 个启动子 PtetO2、PtetO3、PlacO2、PlacO3、PlacO+ara 和 PlacO+rha 的全基因合成质粒为模板, 用引物 Promoter-F/Promoter-R 进行 PCR 扩增得到 6 个启动子序列。

表 2 本研究所用的引物及调控元件序列

Table 2 Primers and regulatory parts sequences used in this work

引物及调控元件 Primers and regulatory parts	序列 Sequences (5'→3')
Construction of placZ	
N20-B-F1	CCAGGTCTCAGTGCGCCATGAGAACGAACCATTG
N20-B-R1	CCAGGTCTCAGCTAAGATCTGACTCCATAACAGAGTACTCGC
lacZ-N20-B-F2	CCAGGTCTCATAGCTTTACAGGGCGGCTTCGTCT GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
N20-B-R2	CCAGGTCTCAACCGCTGGCTAAATACGGAAGGATCT
lacZ-F1	CCAGGTCTCACGGT GGATACGACGATACCGAAGA
lacZ-R1	CCAGGTCTCAGCAC TGACTTTCTCAATAAATGCCTCTAC
Construction of pGFP	
Gfp-GGA-F	GGACCACATGGTCCTTCTTGAGTT
gfp-GGA-R	CCATGTGGTCCCTCTTTTCGTTG
Construction of pΔlacZ-PlacZ-gfp	
lacZ-B-F2	CCAGGTCTCA CCAG TAATAACCGGGCAGGCCAT
lacZ-B-R2	CCAGGTCTCA AGCTGTTTCCTGTGTGAAATTGT
gfp-F	CCAGGTCTCAAGCT ATGGGTAAGGGAGAAGAAGCTTTTC
gfp-R	CCAGGTCTCACTGG TTATTTGTATAGTTCATCCATGCC
Constructing plasmid with designed promoter	
Promoter-F	TGGTTAATGACAGTGGGACGggtctca
Promoter-R	TTACGCAATGCCGAGGTTggtctca
GFP-B-P-F1	CCAGGTCTCATACCATGGGTAAGGGAGAAGAAGCTTTTC
GFP-B-Plac-R1	CCAGGTCTCATATATCACTGCCCCGCTTTCCAG
GFP-B-Para-R1	CCAGGTCTCATATGTCACCTGCCCCGCTTTCCAG
GFP-B-Prha-R1	CCAGGTCTCAAGATTCACTGCCCCGCTTTCCAG
GFP-B-Ptet-R1	CCAGGTCTCAATTTTCACTGCCCCGCTTTCCAG
Detecting promoter with <i>gfp</i> integration	
N20-B-YZ-F	TCGAGTAAGGATCTCCAGGC
lacZ-YZ-up	TGGTTGCCAACGATCAGA
gfp-YZ-down200	AAGTCGTGCCGCTTCATA
araBAD-YZ-up	GATTTCTGCGTCGGTCATT
rhaBAD-YZ-up	AAAAGTGTTACTGTGCTGAA
Construction of pAaraBAD-N20	
araBAD-N20-B-F2	CCAGGTCTCATAGCGCACATGGCGGTCGTGAGTTGTTTTAGAGCTAGAAATAG CAAGTTAAAATAAGGC
araBAD-F1	CCAGGTCTCACGGTCGGACGAAAGTAAACCCACTG
araBAD-R1	CCAGGTCTCACGGCAGTGTGACGCCGTGCAAATAAT
araBAD-F2	CCAGGTCTCAGCACGGTTTCGATGCTAACCACGAC
araBAD-R2	CCAGGTCTCAGCACGGTTTCGATGCTAACCACGAC

(待续)

(续表 2)

Construction of p Δ rhaBAD-N20	
rhaBAD-N20-B-F2	CCAGGTCTCATAGCTGTTTCCGGTTTTGAAAACGTGTTTATAGAGCTAGAAATAG CAAGTTAAAATAAGGC
rhaBAD-F1	CCAGGTCTCACGGTCTTTTACGCAGCAAGAGCAG
rhaBAD-R1	CCAGGTCTCAAATG GAACGTGATGATGTTTACAATTT
rhaBAD-F2	CCAGGTCTCA CATT ACTCAGTCCTGGTTTGTC
rhaBAD-R2	CCAGGTCTCAGCACCGGGAATACCACCAGACACT
Instruction of p Δ rhaBAD::M1-46-tetR	
rhaBAD-B-F1	CCAGGTCTCACGCACATTACTCAGTCCTGGTTTGTC
rhaBAD-B-R1	CCAGGTCTCAAGTGGAACGTGATGATGTTTACAATTT
M1-46-up-F	CCAGGTCTCACACTTTATCTCTGGCGGTGTTGAC
M1-46-down-R	CCAGGTCTCA AGCTGTTTCCTGGTTTAAACC
tetR-F	CCAGGTCTCAAGCTATGTCTAGATTAGATAAAAGTAAAGTGATTAAC
tetR-R	CCAGGTCTCATGCGTTAAGACCCACTTTCACATTTAAGT
p-up	TTATCTCTGGCGGTGTTGACA
Synthesis promoter	
PtetO2	AAATCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAG CACATCAGCAGGACGCACTGACCTTCACACAAGGAGATATACC (97 bp)
PtetO3	AAATCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAG CACATCCCTATCAGTGATAGAGATTACACAAGGAGATATACC (97 bp)
PlacO2	TATAAATTGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGATACTGA CCACATCAGGAGGACGCACTGACCTTCACACAAGGAGATATACC (97 bp)
PlacO3	TATAAATTGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGATACTGA CCACATTGTGAGCGGATAACAACCTTCACACAAGGAGATATACC (97 bp)
PlacO+ara	CATAGCATTTTATCCATAAAGATTAGCGGATCCTAAGCTTGACATTGTGAGCGG ATAACAATATAATGAGCACAATTGTGAGCGGATAACAACCTTCACACAAGGAG ATATACC (114 bp)
PlacO+ara	ATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTGAGTAACGAGAAGGTTTG ACATTGTGAGCGGATAACAATATAATGAGCACAATTGTGAGCGGATAACAACC TTCACACAAGGAGATATACC (126 bp)

注：下划线为靶向各基因的 N20 序列；小写 ggtctc 为 *Bsa* I 酶切位点。

Note: The underlines represent the N20 sequence of the targeting gene; The *Bsa* I is represented by lowercase letters (ggtctc).

以 p Δ lacZ-PlacZ-gfp 为模板，反向 PCR 扩增获得连接启动子的骨架片段。正向引物 GFP-B-P-F1 为扩增骨架通用引物。当反向引物为 GFP-B-Plac-R1 时，扩增用于连接 PlacO2 和 PlacO3 启动子的骨架片段；反向引物为 GFP-B-Ptet-R1 扩增用于连接 PtetO2 和 PtetO3 启动子的骨架片段；反向引物 GFP-B-Para-R1 和 GFP-B-Prha-R1 分别扩增用于连接 PlacO+ara 和 PlacO+rha 启动子的骨架片段。

用 Golden Gate 策略将骨架片段和相应的启动

子片段进行连接，转化 TransT1 感受态细胞，得到的克隆用引物 lacZ-F1/gfp-YZ-down200 进行验证，阳性克隆提取质粒 DNA 送样测序分析，得到正确的 p Δ lacZ-PtetO2-gfp、p Δ lacZ-PtetO3-gfp、p Δ lacZ-PlacO2-gfp、p Δ lacZ-PlacO3-gfp、p Δ lacZ-PlacO+ara-gfp 和 p Δ lacZ-PlacO+rha-gfp 质粒。

1.6 CRSIPR/Cas9 技术将启动子整合染色体上

利用 CRISPR/Cas9 基因编辑技术将启动子整合到染色体上，替换 *lacZ* 基因的启动子；同时将 *lacZ* 基因替换为绿色荧光蛋白 *gfp* 基因(图 1)。

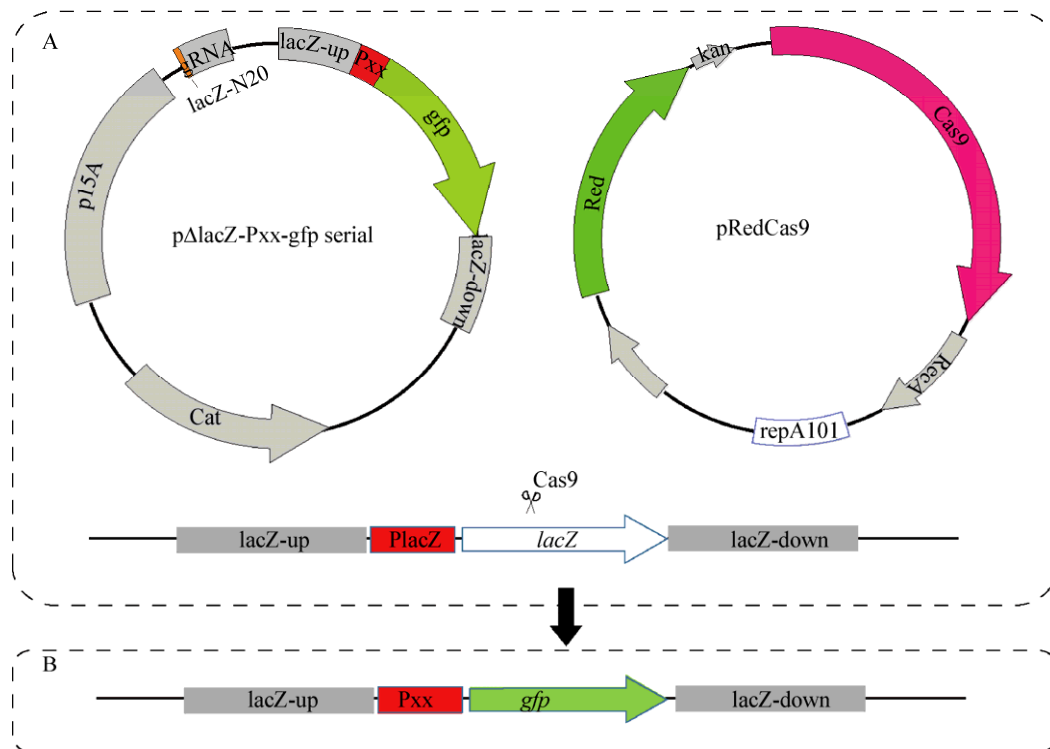


图1 基于 CRISPR/Cas9 技术将启动子整合到染色体上

Figure 1 Promoter integration into chromosomes by CRISPR/Cas9

Note: A: Co-transformation of pRedCas9 and pΔlacZ-Pxx-gfp into target strain; B: With the help of pRedCas9, promoter with *gfp* was inserted and replaced *lacZ* gene.

PlacO2 和 PlacO3 启动子的整合: 将 pRedCas9 和启动子表达质粒如 pΔlacZ-PlacO2-gfp 或 pΔlacZ-PlacO3-gfp 同时电转化于 *E. coli* ATCC 8739 感受态细胞中, 于卡那霉素和氯霉素双抗平板上 30 °C 生长。挑取单克隆转接到 2 mL LB [含卡那霉素和氯霉素; 2% L(+)-阿拉伯糖], 30 °C、250 r/min 过夜诱导重组和切割未发生重组的 DNA。稀释涂 LB 卡那霉素和氯霉素双抗平板[2% L(+)-阿拉伯糖] 30 °C 生长。挑取 10 个单克隆, 用引物 lacZ-YZ-up/gfp-YZ-down200 进行菌落 PCR 验证。阳性克隆的 PCR 反应体系送样测序分析。获得 PlacO2 和 PlacO3 启动子整合菌株 S-PlacO2 和 S-PlacO3。同时, 利用相同的策略, 将 pΔlacZ-Plac-gfp 中的 PlacZ-gfp 整合, 得到对照菌株 S-Plac。

对于 PlacO+ara 和 PlacO+rha 启动子整合, 分别从 S-ΔaraBAD 和 S-ΔrhaBAD 菌株(材料与方法 1.7)出发, 利用相同的启动子整合策略, 将 PlacO+ara-gfp 和 PlacO+rha-gfp 整合, 得到启动子整合菌株 S-PlacO+ara 和 S-PlacO+rha。

对于 PtetO2 和 PtetO3 启动子的整合, 从 S-tetR 菌株(1.8 节)出发, 利用相同的启动子整合策略, 将 PtetO2-gfp 和 PtetO3-gfp 整合, 得到启动子整合菌株 S-tetO2 和 S-tetO3。

1.7 *araBAD* 和 *rhaBAD* 基因的敲除

araBAD 基因敲除质粒的构建: 以质粒 placZ 为模板, 用引物 N20-B-F1/N20-B-R1 进行反向扩增获得通用骨架片段, 含有 *cat* 和 P15A 序列; 以相同质粒为模板, 用引物 araBAD-N20-B-F2/N20-B-R2 扩增, 获得 araBADN20-gRNA 的片段; 以

E. coli ATCC 8739 DNA 为模板, 分别用引物对 araBAD-F1/araBAD-R1 和 araBAD-F2/araBAD-R2 获得 *araBAD* 基因上下游同源臂序列。4 个片段用 Golden Gate 技术策略^[16]进行组装, 转化 TransT1 感受态细胞, 得到的克隆用引物 N20-B-YZ-F/araBAD-BsaI-R2 进行验证, 阳性克隆提取质粒 DNA 送样测序分析, 正确质粒命名为 pΔaraBAD-N20。从 *E. coli* ATCC 8739 出发, 制备电转化感受态细胞, 利用 CRISPR/Cas9 基因编辑系统, 共转化 pRedCas9 和 pΔaraBAD-N20, 实现 *araBAD* 基因的敲除, 得到菌株 S-ΔaraBAD, 具体过程同 1.6 所述。

rhaBAD 基因敲除质粒 pΔrhaBAD-N20 的构建同上, 相应引物名称用 rhaBAD 替换 araBAD, 其敲除技术方法同 *araBAD* 基因的敲除, 得到菌株 S-ΔrhaBAD。

1.8 *tetR* 基因的整合

以质粒 pΔrhaBAD-N20 为模板, 用引物 rhaBAD-B-BsaI-F1/rhaBAD-B-BsaI-R1 进行反向扩增, 得到骨架片段; 以 M1-46 基因组为模板^[8], 用引物 M1-46-up-F/M1-46-R 扩增得到 M1-46 启动子; 以质粒 pSC101 为模板, 用引物 tetR-F/tetR-R 扩增 *tetR* 基因序列。用 Golden Gate 策略将这 3 个片段组装, 转化 TransT1 感受态细胞, 得到的克隆用引物 N20-B-YZ-F/tetR-R 进行验证, 阳性克隆提取质粒 DNA 送样测序分析, 得到正确的 pΔrhaBAD::M1-46-tetR 质粒。从 *E. coli* ATCC 8739 出发, 制备电转化感受态细胞, 利用 CRISPR/Cas9 基因编辑系统, 共转化 pRedCas9 和 pΔrhaBAD::M1-46-tetR, 实现 *tetR* 基因的整合, 得到菌株 S-tetR, 具体过程同 1.6 所述。

1.9 启动子相对表达强度分析及荧光显微镜观察

挑取单克隆到 1 mL LB 试管 37 °C、250 r/min 培养过夜; 按 1% 转接到 1 mL LB 试管, 37 °C、250 r/min 培养 1 h 后加诱导剂 IPTG (终浓度为

0.1 mmol/L), 继续培养 3 h 后测荧光和荧光显微镜观察。取 200 μL 到 96 孔黑色透明板, 用酶标仪测荧光及 OD_{550} ; 荧光强度值=(样品荧光强度-培养基背景荧光强度)/ OD 。同时取 5 μL 培养液于载玻片, 在 100 倍油镜下用奥林巴斯荧光显微镜进行观察; GFP 荧光强度测试条件: 激发光波长 488 nm; 发射光 520 nm; Gain 60。Exposure 值为 400 s, Gain 10, Saturation 10。

2 结果与分析

2.1 启动子的设计

启动子的表达, 由 RNA 聚合酶和启动子的结合速率决定^[17]。RNA 聚合酶和启动子结合速率低, 抑制蛋白就会竞争性结合, 启动子的表达受到抑制。众所周知, 这种类型的启动子即使不诱导, 仍然会有泄露的表达, 如 Plac 启动子。另外一种类型的启动子, 在不存在诱导剂的情况下, RNA 聚合酶和启动子的结合速率高, 启动子强烈表达, 缺点是这种启动子的表达不受控制。因此, 有必要设计出一类严谨型启动子, 在抑制蛋白存在时, RNA 聚合酶不能有效地与启动子结合, 启动子暂不表达; 在有激活剂(诱导剂)存在时, RNA 聚合酶能有效地与启动子结合, 启动子表达。

为更好地描述组合设计的启动子, 我们对启动子上的序列相对于-10 区和-35 区的位置进行命名。将-35 区上游的序列称为位置 IV, -10 区和-35 区之间的序列称为位置 III, -10 区下游的序列称为位置 II, 核糖体 RBS 结合序列称为位置 I。我们对 2 种来源的启动子(PL 和 Plac)及 4 种调控元件 tetO、lacO、ara 和 rha 进行组合设计(图 2), 组合的策略为将调控元件按一定的规律插入且替换位置 IV、III 或者 II 上的序列, 以期从中得到具有严谨控制的启动子。 λ 噬菌体来源的 PL 启动子是一个强启动子, 该启动子受 cI 蛋白的严格调控抑制^[18]。用四环素操纵子的调控区序列 tetO 在位置 IV 和 III 替换 cI 结合序列, 得到 97 bp 的 PtetO2; 调控区序列 tetO 除替换启动子 PL 位置 IV 和 III 外,

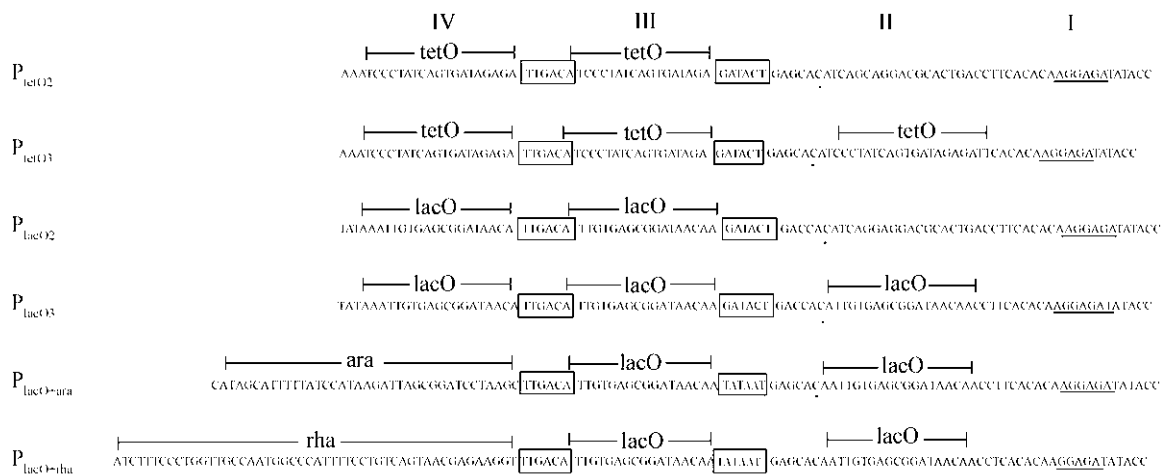


图2 启动子序列设计

Figure 2 Design promoters

Note: The -10 and -35 hexamers are boxed and the transcriptional start site is dotted. Bars indicated the binding site of LacO (lacO), TetO (tetO), AraC (ara) and RhaR (rha) in promoters.

还替换启动子下游序列位置 II, 得到 97 bp 的 P_{tetO3}。类似的, 用 lacO 调控序列替换 PL 启动子的位置 IV, III 或者 IV, III 和 II, 分别得到 97 bp 的启动子 PlacO2 和 PlacO3。Plac 来源的启动子, 将 lacO 调控序列替换到启动子 II 和 III 的位置, 将 ara 调控序列替换到 IV 的位置得到 114 bp 的 Plac+ara 启动子; 将 rha 调控序列替换到 IV 的位置得到 126 bp 的 Plac+rha 启动子。

2.2 启动子的染色体整合

利用 CRISPR/Cas9 基因编辑系统, 本研究将 6 个合成的组合启动子序列 P_{tetO2}、P_{tetO3}、PlacO2、PlacO3、PlacO+ara、PlacO+rha 及对照启动子 Plac, 连同 *gfp* 基因同时整合于染色体上 *lacZ* 基因位点, 替换野生型的 PlacZ-*lacZ* 基因元件(图 1)。具体的整合过程见 1.6。

PlacO2、PlacO3 和 Plac 启动子序列及 *gfp* 整合于 *E. coli* ATCC 8739 中; P_{tetO2} 和 P_{tetO3} 启动子序列及 *gfp* 整合于 S-tetR 菌株中, 该菌株在 *E. coli* ATCC 8739 菌株中整合了 *tetR* 基因, 其编码蛋白 TetR 能调控含 tetO 序列类型启动子的表达; PlacO+ara 启动子序列及 *gfp* 整合于 *E. coli*

S-araBAD 菌株中, 该菌株从 *E. coli* ATCC 8739 敲除了阿拉伯糖代谢途径基因 *araBAD*, 使阿拉伯糖作为诱导剂时不被细胞利用而消耗; PlacO+rha 启动子序列及 *gfp* 整合到 S-rhaBAD 菌株中, 该菌株从 *E. coli* ATCC 8739 敲除了鼠李糖代谢基因 *rhaBAD*, 使鼠李糖作为诱导剂时不被细胞利用而消耗。得到 6 个组合启动子整合菌株 S-tetO2、S-tetO3、S-lacO2、S-lacO3、S-araBAD、S-rhaBAD 和启动子对照菌株 S-Plac。

用引物 lacZ-YZ-up/*gfp*-YZ-down200 对整合后的组合启动子菌株进行 PCR 分析, 得到大小约 700 bp 的目标条带(图 3)。并对 PCR 产物进行测序分析, 测序分析结果显示得到正确的启动子的染色体整合菌株。

2.3 启动子表达强度的分析及荧光观察

对染色体上的 6 个组合型启动子的表达强度进行了评价: 用启动子控制下的 *gfp* 荧光强度作为指示, 用酶标仪对荧光强度进行检测, 将对照启动子 PlacZ, 在诱导剂下的表达强度值设为 1, 其他启动子的表达强度和 PlacZ 启动子的表达强度相比较得出。分析结果列于表 3。

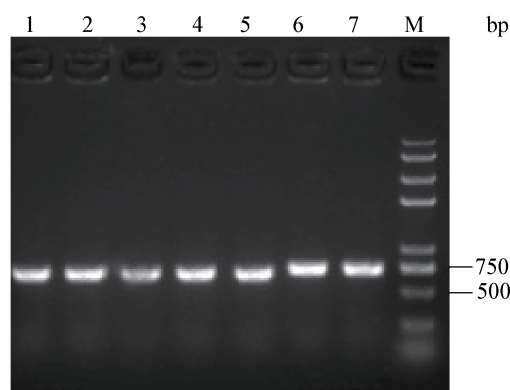


图3 整合型启动子菌株的PCR鉴定

Figure 3 Determination of strains with promoter integration by PCR

Note: M: DNA marker trans 2 000 plus; 1-7: PCR products from strains with promoter Plac, PtetO2, PtetO3, PlacO2, PlacO3, PlacO+ara, PlacO+rha.

PlacO2 和 PlacO3 组合型启动子, 其表达在不加诱导剂的表达强度和加诱导剂 IPTG 的相对表达强度相差无几, 为 0.8–1.0 左右, 说明这对组合型启动子渗漏表达严重, 即使在有诱导剂的情况下, 也没有提高启动子的表达强度。因此, PlacO2 和 PlacO3 这组由 PL 启动子和 lacO 调控元件组合而来的启动子, 组合后并没有达到预期的目标, 不可用于基因的严谨表达。

PtetO2 和 PtetO3 启动子, 其表达在不加诱导剂时的相对表达强度为 0.02, 加诱导剂脱水四环素后, PtetO2 和 PtetO3 的相对表达强度分别为 4.5

和 3.7, 其表达控制范围分别为 225 倍和 185 倍。这组启动子在无诱导剂存在下表达强度低, 在有诱导剂存在下表达较高。因此, PtetO2 和 PtetO3 这组由 PL 和 tetO 调控元件组合而来的启动子, 可用于基因的严谨型表达控制, 为严谨型启动子。

PlacO+ara 组合启动子在不加诱导剂时, 相对表达强度为 0.6, 在加诱导剂时, 相对表达强度为 6。这个组合启动子虽然在加诱导剂时相对表达强度较高, 但是不加诱导剂时也有渗漏表达较多。因此, 由 Plac 和 araC 调控元件组合的 PlacO+ara 启动子, 不能成为严谨型启动子。而 PlacO+rha 组合启动子, 其在不加诱导剂时的相对表达强度为 0.02, 在有诱导剂存在时其相对表达强度为 12, 表达控制范围达到 600 倍。由 Plac 和 rhaR 调控元件组合的 PlacO+rha 组合启动子, 在无诱导剂时基本上无渗漏表达, 在有诱导剂时表达强度很高, 达到了预期目标。因此, PlacO+rha 是一个很好的严谨型启动子, 可用于严谨型基因表达。

同时, 用荧光显微镜对启动子的表达进行了观察, 荧光显微镜观察的结果和用酶标仪测得的相对表达强度的结果具有相关性, 启动子表达强度高的, 绿色荧光亮度高。在此, 只显示了 3 个严谨型启动子 PtetO2、PtetO3 和 PlacO+rha 的荧光观察情况(图 4)。从图 4 中可以看出, 相对于对照菌株 S-PlacZ, 菌株 S-PtetO2 和 S-PtetO3 的绿色荧光强度居中, S-PlacO+rha 菌株的绿色荧光强度最高。

表3 启动子表达强度

Table 3 Promoter strength with or without inducer

启动子整合型菌株 Strains	相对表达强度 Relative strength			OD ₅₅₀	诱导剂 Inducer
	加诱导剂 With inducer	不加诱导剂 Without inducer	表达控制范围 Control range		
S-PlacZ	1.00	0.80		1.4	0.1 mmol/L IPTG
S-PtetO2	4.50	0.02	225.00	1.2	1 µg/mL Anhydrotetracycline
S-PtetO3	3.70	0.02	185.00	1.4	1 µg/mL Anhydrotetracycline
S-PlacO2	0.97	0.84	1.15	1.4	0.1 mmol/L IPTG
S-PlacO3	0.82	0.77	1.06	1.4	0.1 mmol/L IPTG
S-PlacO+ara	6.00	0.60	10.00	1.4	0.1 mmol/L IPTG+0.2% Arabinose
S-PlacO+rha	12.00	0.02	600.00	1.4	0.1 mmol/L IPTG+0.2% Rhamnose

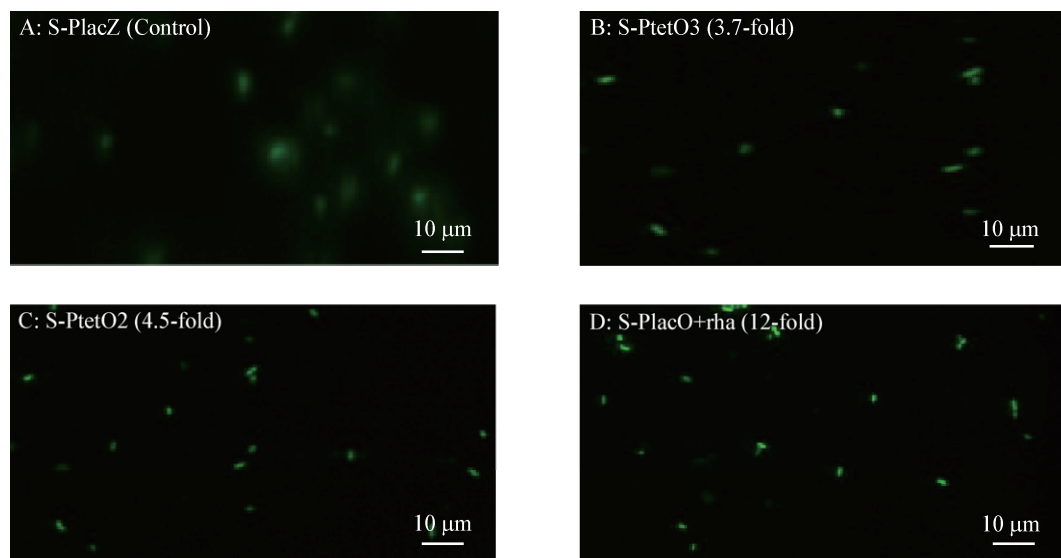


图4 严谨型启动子荧光显微镜观察

Figure 4 Observation of promoter strength by fluorescence microscope

3 讨论和结论

本研究扩大了对基因功能的研究。首先,能严格控制基因的表达,如在生理条件下,某些如热激调控子 σ_{32} 、伴侣蛋白 DnaJ、在细胞分裂中的重要成分 FtsZ 等的基因低表达时对细胞影响的研究。第二,获得的严谨型启动子 PtetO2 和 PlacO+rha 元件,使单个或多个基因进行独立或组合调控,研究某个蛋白浓度的变化或协同表达后对表型的贡献。

启动子-调控序列的组合有以下原则:17–19 bp 的调控序列能充分与阻遏蛋白结合,将这段序列置于间隔区–10 和–35 之间,能有效地阻止 RNA 聚合酶和启动子的结合,从而严格控制启动子的表达^[18];另一个比较好的调控序列位置是–10 区下游 lacO 调控区,这个位置虽不能很好抑制 RNA 聚合酶和启动子的结合,但组合起来利用也有一定的效果^[19]。

PL 启动子和 tetO 调控序列组合,将 tetO 序列置于 PL 启动子的位置 III IV 或 II III IV (图 2),得到组合型启动子 PtetO2 和 PtetO3,这些启动子的活性受到 TetR 和脱水四环素的控制,启动子强度

分别是野生型 PlacZ 的 4.5 倍和 3.7 倍。另外,脱水四环素是比四环素更为有效的诱导剂,它结合 TetR 的能力 35 倍于四环素,因此能在相当低的浓度下诱导启动子的表达^[20],本研究中所用浓度在 1 $\mu\text{g/mL}$ 就能充分诱导 PtetO2 和 PtetO3 的表达,且对细胞生长没有影响。PtetO3 的强度略低于 PtetO2,说明当 tetO 同时在位置 II III IV 时,对启动子的活性略有影响。

PlacO2 和 PlacO3 启动子在有诱导剂和无诱导剂的情况下,表达强度和对照菌株相当。这可能是由于染色体上 *lacI* 基因为单拷贝, LacI 的表达量不足,产生的阻遏蛋白不能充分地与 PlacO2 和 PlacO3 操纵子上的 *lacO* 序列结合,因此启动子泄露表达,从而这 2 个启动子都不能起到严谨控制的作用。若增加染色体上 *lacI* 基因的表达,提供足够的 LacI 阻遏蛋白,可能会实现 PlacO2 和 PlacO3 的严谨控制,相关研究正在探索。

为更好地应用利用阿拉伯糖和鼠李糖作为诱导剂,我们分别敲除了阿拉伯糖和鼠李糖代谢途径基因 *araBAD* 和 *rhaBAD*,用于评价 PlacO+ara 和 PlacO+rha 启动子。这 2 个启动子来源于 Plac 启动

子,除了在 II 和 III 位子上含 lacO 操纵子序列外,在-35 区上游 CRP/cAMP 结合区用 ara 序列替换或 rha 结合序列替换。从诱导情况来看,这 2 个启动子都能较强地表达;从严谨型控制来看,Plac+rha 的严谨型更好,在不添加诱导剂时几乎不表达,相对表达量为 0.02,添加诱导剂后表达强度为 12,表达差异范围为 600。Lutz 等^[21]在不同拷贝数的质粒上对组合型启动子进行了研究,得到了具有严谨性控制的启动子 P_{tetO-1}。遗憾的是,该启动子在不同拷贝数的质粒上表现出不一样的严谨型控制能力。另外,启动子用于质粒上进行基因的表达具有诸多弊端,深化应用于工业生产也有很大的局限性。不同于 Lutz 等的研究,本研究在染色体上对组合型启动子的严谨性进行了探索,得到了最佳严谨型启动子 Plac+rha,本研究结果将为代谢工程和合成生物学中的精确调控基因表达奠定良好的应用基础。

REFERENCES

- [1] Jin XM, Ma YB. Advances in *Pichia pastoris* promoters used for heterogeneous gene expression[J]. Journal of Microbiology, 2015, 35(3): 71-74 (in Chinese)
金晓媚, 马雁冰. 用于异源基因表达的毕赤酵母启动子研究进展[J]. 微生物学杂志, 2015, 35(3): 71-74
- [2] Xu YQ, Ma CQ, Tao F, et al. Bacterial promoter recognition and application[J]. Chinese Journal of Biotechnology, 2010, 26(10): 1393-1403 (in Chinese)
徐友强, 马翠卿, 陶飞, 等. 细菌启动子识别及应用研究进展[J]. 生物工程学报, 2010, 26(10): 1393-1403
- [3] Alper H, Fischer C, Nevoigt E, et al. Tuning genetic control through promoter engineering[J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(36): 12678-12683
- [4] Dykxhoorn DM, St Pierre R, Linn T. A set of compatible *tac* promoter expression vectors [J]. Gene, 1996, 177(1/2): 133-136
- [5] Arfman N, Worrell V, Ingram LO. Use of the *tac* Promoter and *lacI^q* for the controlled expression of *Zymomonas mobilis* fermentative genes in *Escherichia coli* and *Zymomonas mobilis* [J]. Journal of Bacteriology, 1992, 174(22): 7370-7378
- [6] Kincade JM, deHaseth PL. Bacteriophage lambda promoters P_L and P_R sequence determinants of *in vivo* activity and of sensitivity to the DNA gyrase inhibitor, coumermycin[J]. Gene, 1991, 97(1): 7-12
- [7] Jensen PR, Hammer K. Artificial promoters for metabolic optimization[J]. Biotechnology Bioengineering, 1998, 58(2/3): 191-195
- [8] Lu J, Tang JL, Liu Y, et al. Combinatorial modulation of *galP* and *glk* gene expression for improved alternative glucose utilization [J]. Applied Microbiology and Biotechnology, 2012, 93(6): 2455-2462
- [9] Jiang WY, Bikard D, Cox D, et al. RNA-guided editing of bacterial genomes using CRISPR-Cas9 systems[J]. Nature Biotechnology, 2013, 31(3): 233-239
- [10] Jiang Y, Chen B, Duan CL, et al. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system[J]. Applied and Environmental Microbiology, 2015, 81(7): 2506-2514
- [11] Zhao DD, Yuan SL, Xiong B, et al. Development of a fast and easy method for *Escherichia coli* genome editing with CRISPR/Cas9[J]. Microbial Cell Factories, 2016, 15(1): 205
- [12] Li YF, Lin ZQ, Huang C, et al. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing [J]. Metabolic Engineering, 2015, 31: 13-21
- [13] Zhao J, Li QY, Sun T, et al. Engineering central metabolic modules of *Escherichia coli* for improving β -carotene production[J]. Metabolic Engineering, 2013, 17: 42-50
- [14] Srikhanta YN, Dowd SJ, Edwards JL, et al. Phasevarions mediate random switching of gene expression in pathogenic *Neisseria*[J]. PLoS Pathogens, 2009, 5(4): e1000400
- [15] Sambrook J, Fritsch EF, and Maniatis T, Molecular Cloning: A Laboratory Manual[M]. C.S.H. Laboratory, Cold Spring Harbor, New York, USA, 3rd edition, 2001: 1595-1595
- [16] Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability[J]. PLoS One, 2008, 3(11): e3647
- [17] Lanzer M, Bujard H. Promoters largely determine the efficiency of repressor action[J]. Proceedings of the National Academy of Sciences of the United States of America, 1988, 85(23): 8973-8977
- [18] Knaus R, Bujard H. PL of coliphage lambda: an alternative solution for an efficient promoter[J]. The EMBO Journal, 1988, 7(9): 2919-2923
- [19] Kammerer W, Deuschle U, Gentz R, et al. Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process[J]. The EMBO Journal, 1986, 5(11): 2995-3000
- [20] Takahashi M, Degenkolb J, Hillen W. Determination of the equilibrium association constant between tet repressor and tetracycline at limiting Mg^{2+} concentrations: A generally applicable method for effector-dependent high-affinity complexes[J]. Analytical Biochemistry, 1991, 199(2): 197-202
- [21] Lutz R, Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements[J]. Nucleic Acids Research, 1997, 25(6): 1203-1210