

大肠杆菌合成启动子的构建及在顺,顺-粘康酸生物合成中的应用

吴元庆^{1,2}, 张媛媛², 涂然², 刘浩¹, 王钦宏²

1 天津科技大学生物工程学院, 天津 300457

2 中国科学院天津工业生物技术研究所 中国科学院系统微生物工程重点实验室, 天津 300308

吴元庆, 张媛媛, 涂然, 等. 大肠杆菌合成启动子的构建及在顺,顺-粘康酸生物合成中的应用. 生物工程学报, 2013, 29(6): 760-771.

Wu YQ, Zhang YY, Tu R, et al. Construction of the synthetic promoters for *Escherichia coli* and application in the biosynthesis of *cis,cis*-muconic acid. Chin J Biotech, 2013, 29(6): 760-771.

摘 要: 启动子是基因表达调控的重要元件。在代谢工程和合成生物学研究中, 经常需要利用不同强度的启动子对代谢途径进行精细调控, 来实现代谢平衡, 降低中间产物积累, 提高目标产物合成。然而目前可获得的启动子难以满足以上要求, 而且不同来源的启动子通用性差, 缺乏标准化。针对这些问题, 设计了1条88个碱基对的启动子, 包含典型的-35区、-10区以及核糖体结合区。同时, 在转录起始位点上游6个碱基、-35与-10区间隔区14个碱基对中引入简并序列, 构建了合成启动子文库。利用合成启动子控制红色荧光蛋白mCherry的表达强度, 经过两轮筛选, 从5 000多个克隆中获得了720个不同强度的启动子。随机挑选35条不同强度的启动子进行测序分析, 结果表明不同强度的启动子具有碱基偏好性。对于强启动子, -13位点嘌呤碱基出现频率高, 转录起始区除-4位点外, 嘧啶碱基出现的频率高于嘌呤碱基, 而-10区与-35区间14个位点的嘌呤碱基与嘧啶碱基出现频率大致相当。最后选取5条不同强度启动子应用于顺,顺-粘康酸合成途径调控优化, 结果显示不同强度的启动子可以调节目标产物顺,顺-粘康酸的合成和中间产物儿茶酚的积累。

关键词: 合成启动子, 精细调控, 顺,顺-粘康酸, 代谢工程, 合成生物学

Received: December 27, 2012; **Accepted:** January 29, 2013

Supported by: National Basic Research Program of China (973 Program) (No. 2011CBA00800), Industrial Biotechnology Program of Tianjin Municipal Science and Technology Commission (No. 10ZCKFSY05300), Research Equipment Program of Chinese Academy of Sciences (No. YZ201153).

Corresponding author: Qinrong Wang. Tel/Fax: +86-22-84861950; E-mail: wang_qh@tib.cas.cn

国家重点基础研究发展计划 (973 计划) (No. 2011CBA00800), 天津市科学技术委员会工业生物技术专项 (No. 10ZCKFSY05300), 中国科学院科研装备项目 (No. YZ201153) 资助。

网络出版时间: 2013-03-13

网络出版地址: <http://www.cnki.net/kcms/detail/11.1998.Q.20130313.1517.001.html>

Construction of synthetic promoters for *Escherichia coli* and application in the biosynthesis of *cis,cis*-muconic acid

Yuanqing Wu^{1,2}, Yuanyuan Zhang², Ran Tu², Hao Liu¹, and Qinhong Wang²

¹ College of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, China

² Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Science, Tianjin 300308, China

Abstract: Promoter is one of important elements for gene expression and regulation. In the construction of recombinants for metabolic engineering and synthetic biology, it is necessary to have the promoters with varying strengths for fine-tuning metabolic pathway to reach the metabolic balance, decrease the accumulation of intermediate and increase the production of target metabolite. However, the natural promoters available are not completely suitable for fine-tuning metabolic pathway due to discrete strength, lack of versatility and standardization. To deal with this problem, in this study, a new 88 bp synthetic promoter, which contains the typical -35 box, -10 box as well as ribosome bind site, was designed. Then, the promoter library was constructed by introducing some degenerate base pairs in the sequence of 6 bp in the upstream of the initial transcription site and 14 bp in spacer region between -35 and -10 box. 720 promoters with varying strengths were screened out from a library of more than 5 000 clones via the expression of red fluorescent protein mCherry under the control of the synthetic promoter. The sequence analysis based on 35 promoters with varying strengths showed the promoters with varying strengths are base preference. The purine bases in -13 site and pyrimidine bases in the transcriptional initiation sequence are of high frequency; the purine and pyrimidine bases are of the similar frequency in the spacer sequence between -35 and -10 box in strong promoter. In the end, five characterized promoters with varying strengths were selected to tune the synthetic pathway of *cis,cis*-muconic acid in *Escherichia coli*. The results showed that the promoters with varying strengths can regulate the production of *cis, cis*-muconic acid and the accumulation of the intermediate catechol.

Keywords: synthetic promoter, fine tuning, *cis,cis*-muconic acid, metabolic engineering, synthetic biology

近年来,随着合成生物学快速发展为代谢工程提供了强有力的工具,各种设计合成的功能元件逐渐应用于代谢工程领域,显示出巨大的应用潜力^[1-2]。例如,在代谢工程研究中对合成途径的基因进行表达与调控是重要研究内容,精确控制合成途径关键酶或异源基因的表达可以有效地提高产物合成能力和效率^[3-4]。启动子是基因表达调控的重要元件,但是细胞内源基因的启动子由于经常受到细胞自身代谢网络的复杂调控,难以提高其效率,而外源基因的启动子在宿主细胞中经常不能发挥作用,需要替换成其他适合宿

主菌的启动子实现基因的异源表达。另外合成途径的不同基因需要进行不同程度的表达,从而使整个代谢途径获得平衡,提高目标产物的合成,同时降低或阻止中间产物的积累^[5]。因此,在构建合成途径中需利用不同强度,同时不受复杂调控的启动子。

在大肠杆菌中,诱导型启动子如 P_{tac} 、 P_{bad} 及 T7 启动子等被广泛用于基因表达调控^[6-8]。这些启动子需要添加外源诱导物来起始基因的表达与调控。尽管通过添加不同浓度的诱导物在一定范围内可以实现基因不同强度的表达,但是更

多的诱导型启动子只能实现是或非的基因表达,难以进行精细表达调控。同时,诱导型启动子在使用成本高、对诱导物浓度超敏感等缺点^[9]。因此,近些年来设计和筛选新型合成启动子受到越来越多的关注^[10-11]。

设计和筛选合成启动子的方法包括定向进化、理性或半理性设计等^[12-14]。定向进化的方法是通过易错 PCR (Error-prone PCR, Ep-PCR) 在原始启动子整个序列上造成随机突变,形成启动子突变文库,然后结合报告基因的表达强度进行筛选,获得不同强度的启动子。由于易错 PCR 引入突变是分布在启动子的全序列上,许多突变是无效的。因此研究人员就根据启动子的特征保守序列(如-10, -35 区及间隔序列等),在保持某些特征保守序列不变的情况下,在其他序列通过引入简并碱基的理性设计或半理性设计方法,构建小规模、但保持突变多样性的启动子文库,从而快速获得不同强度的启动子。理性设计或半理性设计方法由于相对简单,目前是研究人员关注的热点。

本研究在结合多种启动子设计的基础上^[6,11],设计了一条相对简单、由 88 个碱基对组成的合成启动子,在保持-10 区、-35 区以及核糖体结合区等序列不变的情况下,在转录起始区、-10 区与-35 区间隔区引入简并碱基,直接合成两条包含启动子区及酶切位点的简并配对引物,对引物变性、退火处理后,形成启动子突变文库,然后以红色荧光蛋白 mCherry 为筛选标记,筛选出一系列不同强度的启动子。对筛出的启动子进行测序分析后,得到大肠杆菌合成启动子中相关区域碱基分布与启动子强度的关系。选择不同强度的启动子应用于生物尼龙单体己二酸前体物——

顺,顺-粘康酸^[15-18]合成途径的优化改造,实现了对合成途径的精细调控,得到目标产物顺,顺-粘康酸 (*cis,cis*-muconic acid, 简称 ccMA) 合成增加、中间产物积累减少的工程菌。我们设计、筛选的合成启动子也可以用于其他代谢途径的优化与调控,并为不同体系的启动子设计构建提供了参考。

1 材料与方法

1.1 材料

1.1.1 菌种

大肠杆菌 AB2834 购自美国 *Escherichia coli* Genetic Stock Center, 该菌是 3-脱氢莽草酸脱氢酶突变菌 (*aroE353*), 即 3-脱氢莽草酸到莽草酸合成途径是阻断的; 大肠杆菌感受态细胞 DH5 α 购自北京全式金生物技术有限公司。

1.1.2 质粒及引物

本研究用引物见表 1, 由生工生物工程(上海)股份有限公司合成。

本研究所用或构建的质粒见表 2。*aroZ*、*aroY* 和 *catA* 分别是来源于克雷伯氏肺炎球菌 *Klebsiella pneumoniae* 的 3-脱氢莽草酸脱水酶 (3-dehydroshikimate dehydratase)、原儿茶酸脱羧酶 (Protocatechuate decarboxylase) 和来源于乙酸钙不动杆菌 *Acinetobacter calcoaceticus* 的儿茶酚 1,2-双氧化酶 (Catechol 1,2-dioxygenase) 的基因^[15-16]。这 3 个基因组成的异源合成途径引入大肠杆菌后可以把芳香族氨基酸合成途径中的 3-脱氢莽草酸转化为顺,顺-粘康酸。Amp^R、Tc^R、Cm^R 和 Spe^R 分别表示氨苄青霉素、四环素、氯霉素和壮观霉素抗性。mCherry 是编码红色荧光蛋白的基因。

表 1 本研究所用引物

Table 1 Primers used in this study

Primer name	Primer sequence (5'-3')
Ptc-EcoRI-5	CACTGT <u><i>GAATTC</i></u> AAAACGATCTCAAGAAGATC
Ptc-3	GAATATTTTAACTTCCATGGCCGGTGCCTGACTGCGTTAG
CatA-tc-5	CTAACGCAGTCAGGCACCGGCCATGGAAGTTAAATATTC
CatA-EcoRI-3	CACTGAGA <u><i>GAATTC</i></u> GAACCATTTTGGTGTATTACAC
mCherry-5	CTAGAT <u><i>GGTACCA</i></u> AAGAAGGAGATATACATATGGTGAGCAAGGGCGAGGAGGAC
mCherry-3	CTAGAT <u><i>CTCGAGT</i></u> TAGCCGGCCTTGTACAGCTCGTC
KD8.292T- 5	AGAGAT <u><i>GGTACCG</i></u> TATTACACCGCTAGACGTGGACG
KD8.292T- 3	CTAGAT <u><i>CTCGAG</i></u> CCGACAGTAAGACGGGTAAGCCTGTTG
PL-F	<u><i>CTAGA</i></u> ATATGTTATCTCTGGCGGTGTTGACANNNNNNNNNNNNTGRTATAATNN NNNNTGAGCGGATAACAATTCAAGGAGGACAGCTC
PL-R	<u><i>CATGG</i></u> AGCTGTCCTCCTTGAAATTGTTATCCGCTCANNNNNNATTATAYCANNNNN NNNNNNNNNTGTCAACACCGCCAGAGATAACATATT

The underlined italic is the restriction sites.

表 2 实验用质粒

Table 2 Plasmids used in this study

Plasmid name	Description	Source
pYIP-mCherry	mCherry, Amp ^R	Our lab
pACYC184	Tc ^R , Cm ^R	Our lab
pKD8.243	<i>aroZ</i> , <i>aroY</i> , Cm ^R	[13]
pKD8.292	<i>catA</i> , Spe ^R	[13]
pKD8.292T	Derived from pKD8.292 with the promoter of Tc ^R cassette (<i>Ptc</i>) to replace tac promoter (<i>Ptac</i>)	This study
pKD8.292T-mCherry	Derived from pKD8.292T with the insertion of mCherry under the control <i>Ptc</i>	This study
pKD8.292PL01-mCherry	Derived from pKD8.292T-mCherry with synthetic promoter PL01 to place promoter <i>Ptc</i>	This study
PKD8.292PL05-mCherry	Derived from pKD8.292T-mCherry with synthetic promoter PL05 to place promoter <i>Ptc</i>	This study
pKD8.292PL20-mCherry	Derived from pKD8.292T-mCherry with synthetic promoter PL20 to place promoter <i>Ptc</i>	This study
pKD8.292PL31-mCherry	Derived from pKD8.292T-mCherry with synthetic promoter PL31 to place promoter <i>Ptc</i>	This study
pKD8.292PL35-mCherry	Derived from pKD8.29T-mCherry with synthetic promoter PL35 to place promoter <i>Ptc</i>	This study

1.1.3 试剂

限制性内切酶均购自 Fermentas 公司, *TransStartTMFastPfu* DNA 聚合酶购自北京全式金生物技术有限公司; 莽草酸、顺,顺-粘康酸及抗生素 (氨苄青霉素、四环素、氯霉素和壮观霉素) 购自 Sigma 公司; 质粒提取试剂盒及 PCR 产物纯化回收试剂盒购自 Biomiga 公司; BCA 蛋白定量试剂盒 Protein Assay Kit 购自 Merck 公司, 蛋白提取试剂 Protein Extraction Reagent 购自 Novagen 公司; 其他配制培养基药品均购自天津国药集团有限公司, 分析纯。

1.1.4 培养基

液体 LB (Luria-Bertani) 培养基: 10 g 蛋白胨, 10 g NaCl 及 5 g 酵母膏。固体 LB 培养基: 10 g 蛋白胨, 10 g NaCl, 5 g 酵母膏及 1.5 g 琼脂粉。液体及固体 LB 培养基 pH 7.0, 1×10^5 Pa 灭菌 30 min。

发酵培养基: 含有 1% 葡萄糖的 M9 基本培养基。其中氨苄青霉素、四环素、氯霉素和壮观霉素在培养基中的工作浓度分别为 50、25、34 和 50 $\mu\text{g/mL}$, 莽草酸的终浓度为 0.04 g/L。氨苄青霉素、四环素、氯霉素、壮观霉素及莽草酸均用直径 0.22 μm 的滤膜过滤除菌后加入培养基中。

1.2 方法

1.2.1 质粒构建

质粒 pKD8.292T 的构建: 以质粒 pACYC184 为模板, 以引物 Ptc-EcoRI-5 和 Ptc-3 扩增出大约 220 bp 的四环素抗性基因组成型启动子 P_{tc} , 以质粒 pKD8.292 为模板, 以引物 CatA-tc-5 和 CatA-EcoRI-3 扩增出 1 080 bp 的 *catA* 基因, 然

后这两个 PCR 产物通过融合 PCR 获得 1.3 kb 融合 PCR 产物 P_{tc} -catA^[19]; P_{tc} -catA 和 pKD8.292 分别用 *EcoR* I 酶切、回收、连接后获得重组质粒 pKD8.292T, 这样组成型启动子 P_{tc} 替代原来质粒 (pKD8.292) 中的诱导型启动子 P_{tac} 。

质粒 pKD8.292T-mCherry 的构建: 以质粒 pYIP-mCherry 为模板, 以引物 mCherry-5 及 mCherry-3 扩增出携带 *Kpn* I 及 *Xho* I 酶切位点的红色荧光蛋白 mCherry 基因; 以质粒 pKD8.292T 为模板, 以引物 KD8.292-T-5 及 KD8.292-T-3 扩增出携带 *Kpn* I 及 *Xho* I 酶切位点的线性片段 KD8.292T, 分别用 *Kpn* I 及 *Xho* I 双酶切、回收、连接后获得重组质粒 pKD8.292T-mCherry。该质粒中引入与 *catA* 共用同一启动子的红色荧光蛋白 mCherry 基因, 作为筛选报告基因。

1.2.2 合成启动子的构建及筛选

将合成的启动子引物 PL-F 和 PL-R (表 1) 分别用灭菌去离子水配置成 10 $\mu\text{mol/L}$ 浓度, 等体积混合后加入适量 $10 \times$ T4 DNA 连接酶缓冲液, 置于 85 $^{\circ}\text{C}$ 水浴 10 min, 自然冷却至室温即可得到携带 *Xba* I 及 *Nco* I 突出末端的多样性启动子。

将质粒 pKD8.292T-mCherry 用限制性内切酶 *Xba* I 和 *Nco* I 处理后纯化回收, 用 T4 DNA 连接酶与合成启动子连接, 连接产物转化至感受态细胞 DH5 α , 涂布于有 50 $\mu\text{g/mL}$ 壮观霉素的固体 LB 平板, 37 $^{\circ}\text{C}$ 过夜培养, 得到合成启动子多样性单克隆。

挑取单克隆至每孔有 200 mL LB 液体培养基 (含 50 $\mu\text{g/mL}$ 壮观霉素) 的 96 孔板中, 于 37 $^{\circ}\text{C}$ 振荡培养 4 h 后, 在激发波长 587 nm, 发射

波长 610 nm 的条件下^[20]检测红色荧光强度,同时测定细胞生长 (OD_{600}),以红色荧光强度除以 OD_{600} ,即得单位 OD_{600} 相对红色荧光强度 (Relative Intensity of Red Fluorescent per OD_{600} ,简称 RIRF/ OD_{600}),依此进行合成启动子强度的筛选。

1.2.3 合成启动子序列分析

从筛选出的不同强度启动子按照一点强度间隔,从低到高随机选出 35 条启动子进行测序分析。对获得的序列进行不同区域,包括转录起始区和间隔区的碱基分布统计,解析碱基分布与启动子强度的关系。

1.2.4 合成启动子在大肠杆菌顺,顺-粘康酸生产中的应用

将筛选得到的不同合成启动子克隆与质粒 pKD8.243 共转至大肠杆菌 AB2834 感受态细胞,得到顺,顺-粘康酸合成的不同工程菌株。将不同工程菌株分别接种到有氯霉素及壮观霉素的液体 LB 培养基中 37 °C、220 r/min 振荡培养过夜,将该培养液以 1% 接种量接至同样液体 LB 培养基中,37 °C、220 r/min 振荡培养 8 h,然后以 20% 的接种量接至发酵培养基,37 °C、220 r/min 振荡培养 36 h,收集发酵液离心后去上清液利用 HPLC 分析发酵产物。HPLC 分析条件如下:分析柱 Eclipse XDB-C18 column (Agilent

Technologies),上样量 10 μ L,流动相是混合比为 9:1 的 2% 乙醇与纯甲醇,流速为 1 mL/min,柱温 30 °C,UV-Vis 吸收波长为 260 nm。每个样品重复 3 次取平均值和标准偏差。

1.2.5 儿茶酚 1,2-双氧化酶酶活测定

将不同工程菌株分别接种到有氯霉素及壮观霉素的液体 LB 培养基中 37 °C、220 r/min 振荡培养过夜,将该培养液以 1% 接种量接至同样液体 LB 培养基中,37 °C、220 r/min 振荡培养 12 h,离心收集菌体。菌体用蛋白抽提试剂处理后离心获得上清液用于蛋白浓度和儿茶酚 1,2-双氧化酶活测定,酶活测定方法见参考文献[16]和[18]。

2 结果与分析

2.1 合成启动子序列设计

Harleyd 等^[21]对 263 条大肠杆菌启动子序列比对后发现存在着保守序列,即-35 区与-10 区,在两保守序列间间隔序列最优长度为 17 bp,该序列的碱基分布对启动子强度有很大影响。对原核生物的研究表明^[13,22], -10 区上游隔 1 bp 处也有个保守框,即 TG, -13 位点碱基为 A 或 G (二者出现的理论频率均为 50%),位于转录起始+1 位点上游 6 个位点 (-1~-6) 的转录起始区的碱基也具兼并性。据此设计了如图 1 所示合成启动子

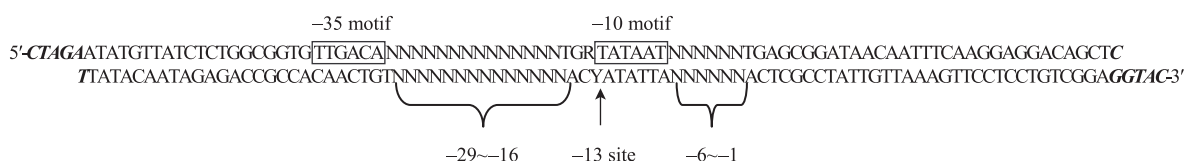


图 1 设计的合成启动子序列

Fig. 1 Sequence design of synthetic promoters.

序列。去掉两端的酶切位点序列 (这里是 *Xba* I 和 *Nco* I) 后, 启动子全长为 88 个碱基对, 其中, N 代表 A、G、T、C 4 个碱基且 4 个碱基在各位点出现的理论频率为 25%, R 与 Y 代表碱基 A 与 G 且 2 个碱基在该位点出现的理论频率为 50%, 框内即为 2 个保守区域序列, -35 区与 -10 区, 大括号内为随机突变区。

为了简化实验操作, 首先合成两条寡聚核苷酸链 PL-F 和 PL-R (表 1), 然后等摩尔混合这两条寡聚核苷酸链, 通过变性和退火处理就获得了合成启动子 DNA。同时在设计合成启动子序列时, 引入酶切位点的粘性末端, 其中 5' 端为限制性内切酶 *Xba* I 的粘性末端, 3' 端为限制性内切酶 *Nco* I 的粘性末端 (图 1 加粗斜体部分)。获得的合成启动子 DNA 可直接用于 DNA 连接反应, 无需进一步酶切处理。

2.2 启动子文库的构建与筛选

将上述设计合成的合成启动子 DNA 与进行过相应酶切处理的质粒 pKD8.292T-mCherry 连接, 转化大肠杆菌细胞 DH5 α 后形成了启动子文库。挑选启动子文库中克隆, 以启动子控制红色荧光蛋白 mCherry 的表达来筛选和检测启动子的强度 ($RIRF/OD_{600}$)。同时以含有质粒 pKD8.292T-mCherry 作为组成型启动子强度的参考。

第 1 轮筛选检测了启动子文库中的 5 320 个克隆, 在第 1 轮的基础上, 第 2 轮筛选以 $RIRF/OD_{600}$ 值相差 10 为单位梯度, 由小至大筛选出 $RIRF/OD_{600}$ 相对稳定的 720 个克隆, 结果见图 2。由图可以看出, 所检测的 720 个启动子中约有 600 个 (从大约第 50 号到大约 650 号) 强

度是线性增加的。另外这 720 个启动子中, 最大强度和最小强度之间比值约为 15 倍。这些启动子基本满足基因线性表达调控的需求。

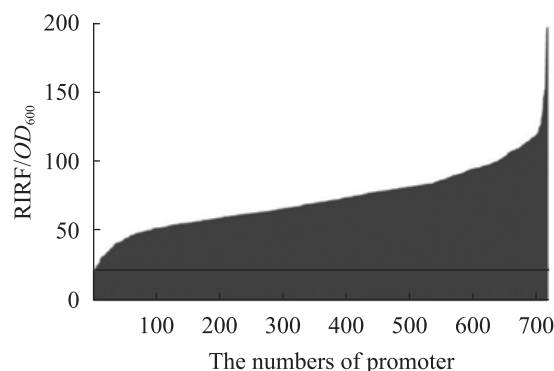


图 2 合成启动子单位 OD_{600} 相对红色荧光强度 (图中水平黑线是组成型启动子 P_{tc} 的相对强度)

Fig. 2 Relative intensity of red fluorescent per OD_{600} ($RIRF/OD_{600}$) for synthetic promoters. Solid black line: the relative strength of P_{tc} constitutive promoter.

2.3 启动子序列分析及表征

为找到启动子强度与突变位点序列之间的对应统计关系, 从含 720 条启动子的文库中随机筛选出 35 条 (PL01-PL35) 不同强度启动子进行测序分析, 结果如图 3 所示。将 35 条启动子根据各自相对强度 ($RIRF/OD_{600}$) 大小与参考启动子 (组成型启动子 P_{tc}) 的相对强度相比, 分为 3 组: 即弱启动子 (Weak promoter) 为参考启动子的 2 倍以下 (包括 2 倍); 中等强度启动子 (Moderate promoter) 为 2 倍以上、3 倍以下 (包括 3 倍); 强启动子 (Strong promoter) 为对照值 3 倍以上, 分别分析随机突变区域碱基分布。

对于不同强度启动子的 -10 区的 -13 位点, 启动子强度越大, 碱基 G 出现的频率越比碱基 A 出现的频率大。进一步分析发现, 在强弱启动子中二者出现的频率比恰好相反。在强启动子中碱

基 G 和碱基 A 出现的频率是 2 : 1, 在弱启动子中是 1 : 2, 而在中等强度启动子出现频率接近 1。

原核生物中, -10 区与-35 区间隔序列大约是 16~19 bp。小于 15 bp 大于 20 bp 都会降低启动子活性^[23], 而研究表明, 启动子活性最好时距离为 17 bp^[21], 为揭示间隔区 17 bp 中除-15 至-13 以外各位点碱基出现频率与启动子强弱关系, 分析上述 3 组启动子, 结果见图 4A。图 4A 结果表明在强启动子中, 此区域碱基 A、G、T、C 出现频率大致相当; 弱启动子中, 嘧啶碱基出现频率远大于嘌呤碱基出现频率; 中等强度启动子中, A 与 C 的频率较大, 与弱启动子相比, 嘌呤碱基的频率增大。

由于在原核生物细胞中某一基因表达经常取决于转录及翻译的起始效率^[23]。为了说明该区域各位点碱基的分布与启动子强度的关系, 不仅分析了各位点各碱基出现的实际频率, 还分析了各位点嘌呤碱基与嘧啶碱基出现的频率, 结果见图 4B。在强启动子的 6 个位点中除-4 位点外, 嘌呤碱基均比嘧啶碱基出现的频率低; 在中等强度启动子嘌呤碱基与嘧啶碱基出现频率差异不大; 弱启动子的-2 位点为碱基 C, 仅在-3 与-6 位点出现嘧啶碱基, 嘌呤碱基出现的频率明显低于嘧啶碱基出现的频率。合成启动子特定区域碱基分布规律研究将为启动子进一步理性或半理性设计提供基础。

	CTAGAAATATGTTATCTCTGCGCGTGTGACA	NNNNNNNNNNNNNTGR	TATAAT	NNNNNN	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	RIRF/ <i>OD</i> ₆₀₀
PL01	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TTTCCGTTTTTTCATGA	TATAAT	CGCTCC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	31.46
PL02	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GTITTTAAITTTGATGG	TATAAT	TCCTCG	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	38.06
PL03	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	AAACAAAGGAAAAATGA	TATAAT	CGATCG	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	39.86
PL04	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GCGCCAAAGAACTCTGG	TATAAT	TCCAAC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	45.90
PL05	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GAACCAAAAAACATATGG	TATAAT	CCAAGA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	46.55
PL06	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ATTGCGTTCGCTTTGTA	TATAAT	GTAGTT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	47.52
PL07	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	AACACAAAAAATGTA	TATAAT	CAACGG	TGAGCGGATTACAATTTCAAGGAGGACAGCTC	47.67
PL08	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	AAAATCCTAAAAGATGG	TATAAT	AGCAGA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	48.36
PL09	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GAACACACACGAAATGA	TATAAT	AACCAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	49.71
PL10	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CAACCCACAGAGAACTGA	TATAAT	CGAGCC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	51.52
PL11	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TGAGCACAAACAAATGG	TATAAT	GAAGGC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	51.90
PL12	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ACCACGAGACCAGTGG	TATAAT	ACTCTA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	52.17
PL13	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ACAACAATAAGAAATGA	TATAAT	CATCGA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	52.22
PL14	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TCGGCACGAAAGACCTGA	TATAAT	AGAAGG	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	53.98
PL15	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ACAGAACAGATAAAATGG	TATAAT	ACCCAC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	55.13
PL16	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TGTGACTCTGTTTGTGG	TATAAT	GCTGTT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	56.73
PL17	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ATAAAAAACACAAATGA	TATAAT	ACCAAC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	56.91
PL18	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CGCCCAAGACCGCGTGG	TATAAT	GCACAC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	57.53
PL19	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GCGTTGTTTTTGTAGTGG	TATAAT	TTTTGT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	58.36
PL20	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	AACGCCAAAAACAGTGA	TATAAT	GAAAAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	58.54
PL21	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	AAAAACCAACACAGTGG	TATAAT	CAAGAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	61.98
PL22	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GACAACCAAGCAATGA	TATAAT	AACCAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	63.29
PL23	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ACGAGCTGGACAACTGG	TATAAT	GCCACA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	67.15
PL24	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TTCTTGCGTGTATATGG	TATAAT	AGTGTT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	67.91
PL25	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ACGAGCTGGACAACTGG	TATAAT	GCCACA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	68.29
PL26	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	ATGTGGCTTGTGTTGTA	TATAAT	TTCITT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	76.71
PL27	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TTGGACTTTCTATTTGG	TATAAT	GAATTG	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	77.69
PL28	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CGCTGTTGCTTTTGTGG	TATAAT	TTGTTG	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	84.76
PL29	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TGTTGTGGCAGTGATGA	TATAAT	TTTGGT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	87.37
PL30	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GCCGTTCCITGCTCTGA	TATAAT	TTGTTT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	88.06
PL31	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	GAGAAAAACCAAAATGG	TATAAT	ACGACC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	88.45
PL32	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	TTTTCGTTTGTAGTGA	TATAAT	TTGTTT	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	91.17
PL33	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CCGAAACCGCGAAGATGG	TATAAT	AAGCAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	92.12
PL34	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CACAAAAACCAACCTGG	TATAAT	ACACCC	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	98.79
PL35	TCTAGAAATATGTTATCTCTGCGCGTGTGACA	CAAAAAAGAGGCCATGG	TATAAT	CAACAA	TGAGCGGATAACAATTTCAAGGAGGACAGCTC	104.58

图 3 35 条不同强度启动子序列及相对强度

Fig. 3 Sequence of 35 synthetic promoters with different strength. RIRF/*OD*₆₀₀: the relative intensity of red fluorescent per *OD*₆₀₀.

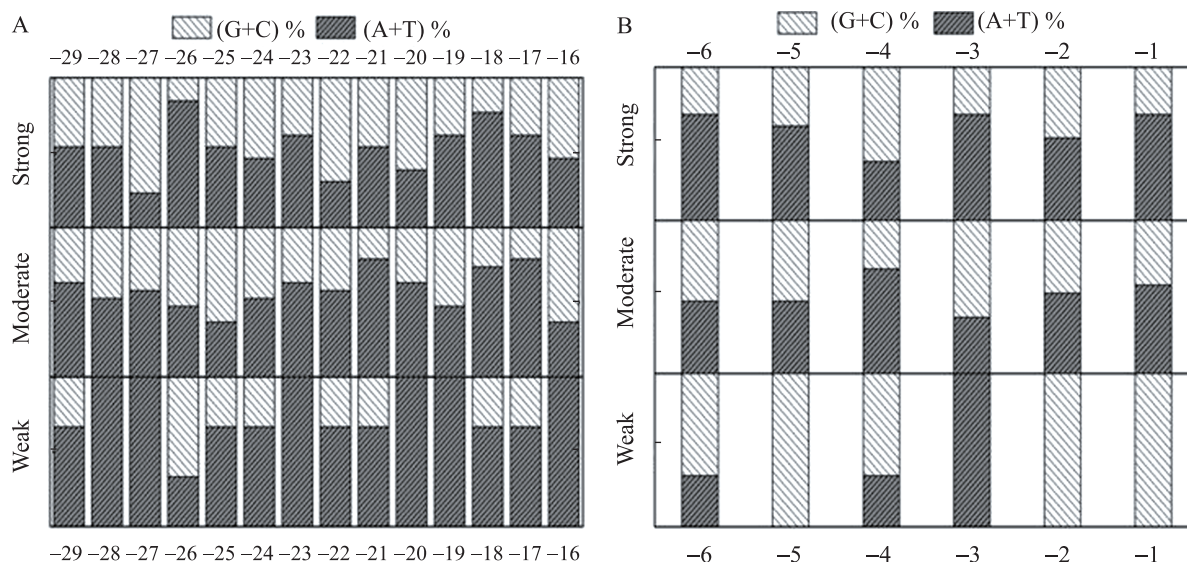


图4 不同合成启动子不同区域碱基分布规律

Fig. 4 Base preference of different motif in synthetic promoters. (A) The motif between -35 and -10. (B) The sequence of transcriptional initiation.

2.4 启动子文库的应用

顺,顺-粘康酸是合成尼龙单体己二酸的前体物。野生型大肠杆菌不能合成顺,顺-粘康酸,需要引入由3-脱氢莽草酸脱水酶、原儿茶酸脱羧酶和儿茶酚1,2-双氧化酶组成的异源合成途径,将大肠菌中间产物3-脱氢莽草酸,通过原儿茶酸和儿茶酸转化为儿茶酚转化为顺,顺-粘康酸^[15-16]。

为了验证启动子强度对异源顺,顺-粘康酸合成途径的精细调控,挑选了5条不同强度启动子,分别为PL01、PL05、PL20、PL31及PL35,其中PL01为弱启动子,PL05及PL20为中等强度启动子,PL31及PL35为强启动子,其强度见图5A。首先测定了5株菌儿茶酚1,2-双氧化酶的比活力,结果见表3。由表可知,合成启动子均可以调节儿茶酚1,2-双氧化酶的比活力,且强度越高,儿茶酚1,2-双氧化酶的比活力越高。与对照的组成型启动子 P_{tc} 相比,可以大幅度提高

酶的活性,而且不同合成启动子之间的活性也可以形成较大的差异。根据表中数据显示,合成启动子之间的比酶活可以相差13倍。

相应地,顺,顺-粘康酸发酵分析结果(图5)也表明合成启动子可以有效地调节工程菌生产顺,顺-粘康酸的能力。与对照的组成型启动子 P_{tc} 相比,其产量有大幅提高,最高产量提高了约21倍,合成启动子之间的产量差别也多达2倍。

另外,由表3和图5可知,对某一特定途径,并不是启动子越强产物的产量就越高,启动子强度应有一个阈值,超过这一阈值,产量增加很少或是下降,而在这一阈值范围内,启动子强度稍有加强产物产量即大幅提高,如合成启动子PL01、PL05及PL20对顺,顺-粘康酸合成的影响。还有随着启动子强度的增加,中间产物儿茶酚的积累就逐渐减少,甚至在发酵液中不能检测到(图5B)。

表 3 不同强度启动子对儿茶酚 1,2-双氧化酶活的影响

Table 3 Effect of synthetic promoters with varying strength on catechol 1,2-dioxygenase specific activity

Strains	Promoter	Activity of catechol 1,2-dioxygenase (U/mg)
AB2834/pKD8.243/pKD8.292T-mCherry	<i>P_{tc}</i>	0.0017±0.0001
AB2834/pKD8.243/pKD8.292PL01-mCherry	PL01	0.0732±0.0037
AB2834/pKD8.243/pKD8.292PL05-mCherry	PL05	0.1559±0.0078
AB2834/pKD8.243/pKD8.292-PL20-mCherry	PL20	0.1867±0.0093
AB2834/pKD8.243/pKD8.292-PL31-mCherry	PL31	0.6465±0.0323
AB2834/pKD8.243/pKD8.292-PL35-mCherry	PL35	0.9881±0.0494

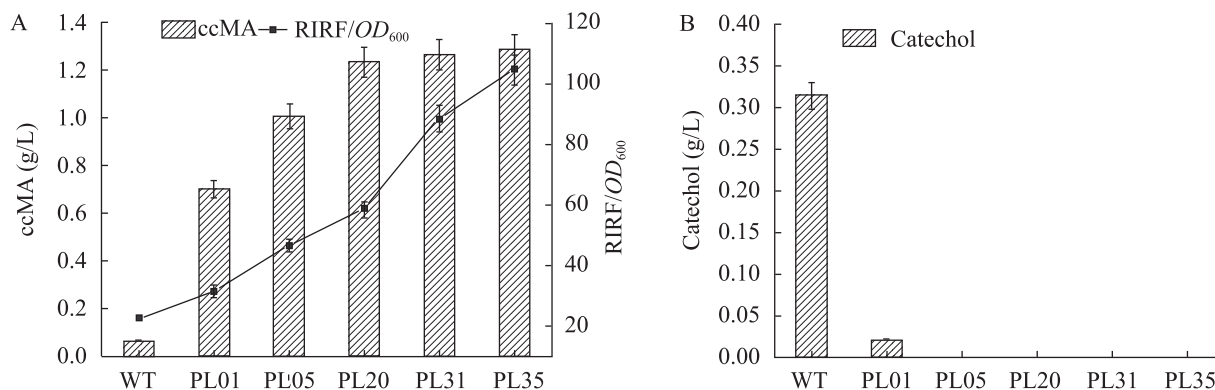


图 5 不同启动子影响顺,顺-粘康酸发酵生产

Fig. 5 Effect of synthetic promoters with varying strength on *cis,cis*-muconic acid production. (A) The relationship of production of ccMA and RIRF/OD₆₀₀. (B) The accumulation of catechol.

3 讨论

启动子文库的构建可以通过 Ep-PCR 的方法得到,也可以通过将突变区域序列设计到 PCR 引物上,在扩增目标基因时得到目的基因的启动子文库,比较两种方法可以发现 Ep-PCR 的要求较高,而且难以得到足够的突变体尤其是有效突变体^[6]。本研究采用 DNA 合成方法得到多样性突变启动子文库,操作简单,在短时间内即可以获得大量突变体。通过两轮启动子文库的筛选,

得到表达强度差异超过 15 倍的启动子文库,应用于酶活测定及发酵的启动子均比野生型启动子强,达到了预期目的,而且研究结果进一步说明了启动子能够对途径实现精细准确的调控。

在启动子文库的设计中,保守序列间间隔长度为 17 bp,虽然早有研究表明在原核生物中保持-35 区与-10 区间距离至关重要^[23],但对于代谢工程和合成生物学研究来说,为使元件实现多样性及普遍性,还可以改变该区域序列长度,如 19、18、16、14 bp 等。此外,原核生物中 RNA

聚合酶的 σ 因子对启动子的识别及转录起始有较大影响^[24],而RNA聚合酶的 α 亚基也能够识别启动子-35区上游罕见的富含A+T碱基的保守序列UP元件DNA,使启动子强度变化1.5~90倍^[25]。另有研究表明,启动子中-10区的保守性没有-35区保守性强,其-10位点的碱基T也可以替换为其他碱基从而降低启动子的表达强度^[26]。而改变作为细菌中调控蛋白质翻译的DNA序列核糖体结合位点(Ribosome binding sites, RBSs)序列,将其同启动子序列一起亦可以得到表达强度差异很大的文库^[27]。这为将来进一步设计提供了方向。

总之,我们的启动子设计、筛选和应用的研究不仅为启动子进一步的理性或半理性设计奠定了基础,而且获得的启动子元件库可以用于应用于生物合成途径调节与控制。

REFERENCES

- [1] Wang JS, Qi QS. Synthetic biology for metabolic engineering—a review. *Chin J Biotech*, 2009, 25(9): 1296–1302 (in Chinese).
王俊姝, 祁庆生. 合成生物学与代谢工程. *生物工程学报*, 2009, 25(9): 1296–1302.
- [2] Zhao XQ, Bai FW, Li Y. Application of systems biology and synthetic biology in strain improvement for biofuel production. *Chin J Biotech*, 2010, 26(7): 880–887 (in Chinese).
赵心清, 白凤武, 李寅. 系统生物学和合成生物学研究在生物燃料生产菌株改造中的应用. *生物工程学报*, 2010, 26(7): 880–887.
- [3] Tyo KE, Alper HS, Stephanopoulos GN. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotech*, 2007, 25(3): 132–137.
- [4] Wang T, Ma X, Zhu H, et al. Available methods for assembling expression cassettes for synthetic biology. *Appl Microbial Biotechnol*, 2012, 93(5): 1853–1863.
- [5] Blazeck J, Alper HS. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Biotechnol J*, 2013, 8(1): 46–58.
- [6] Jensen PR, Westerhoff HV, Michelsen O. The use of lac-type promoters in control analysis. *Eur J Biochem*, 1993, 211(1-2): 181–191.
- [7] Guzman LM, Belin D, Carson MJ et al. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacter*, 1995, 177(14): 4121–4130.
- [8] Zhou Q, Chen JC, Chen GQ. Application of high efficiency promoters in microbial production of 4-hydroxybutyric acid. *Chin J Biotech*, 2012, 28(1): 48–55 (in Chinese).
周琴, 陈金春, 陈国强. 高效启动子在微生物生产4-羟基丁酸中的应用. *生物工程学报*, 2012, 28(1): 48–55.
- [9] Brautaset T, Lale R, Valla S. Positively regulated bacterial expression systems. *MicrobBiotechnol*, 2009, 2(1): 15–30.
- [10] Zhao X, Jiang R, Bai F. Directed evolution of promoter and cellular transcription machinery and its application in microbial metabolic engineering—a review. *Chin J Biotech*, 2009, 25(9): 1312–1315 (in Chinese).
赵心清, 姜如娇, 白凤武. 启动子和细胞全局转录机制的定向进化在微生物代谢工程中的应用. *生物工程学报*, 2009, 25(9): 1312–1315.
- [11] Ruth C, Glieder A. Perspectives on synthetic promoters for biocatalysis and biotransformation. *Chembiochem*, 2010, 11(6): 761–765.
- [12] Alper H, Fischer C, Nevoigt E, et al. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci USA*, 2005, 102(36): 12678–12683.
- [13] Schlabach MR, Hu JK, Li M, et al. Synthetic design of strong promoters. *Proc Natl Acad Sci USA*, 2009, 107(6): 2538–2543.
- [14] Hammer K, Mijakovic I, Jensen PR. Synthetic promoter libraries—tuning of gene expression. *Trends Biotechnol*, 2006, 24(2): 53–55.

- [15] Draths KM, Frost JW. Environmentally compatible synthesis of adipic acid from D-glucose. *J Am Chem Soc*, 1994, 116(1): 399–400.
- [16] Frost JW, Draths KM. Synthesis of adipic acid from biomass-derived carbon sources: US, 5487987. 1996-01-20.
- [17] Niu W, Draths KM, Frost JW. Benzene-free synthesis of adipic acid. *Biotechnol Prog*, 2002, 18(2): 201–211.
- [18] Zhang YY. Strain selection and fermentation optimization for adipic acid precursorsbiorefinery [D]. Tianjin: Tianjin Univ Sci Technol, 2011 (in Chinese).
张媛媛. 己二酸前体物的菌种选育与发酵优化 [D]. 天津: 天津科技大学, 2011.
- [19] Yon J, Fried M. Precise gene fusion by PCR. *Nucleic Acids Res*, 1989, 17(12): 4895.
- [20] Shaner NC, Campbell RE, Steinbach PA, et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol*, 2004, 22(12): 1567–1572.
- [21] Calvin BH, Robert PR. Analysis *E. coli* promoter sequences. *Nucleic Acids Res*, 1987, 15(5): 2343–2361.
- [22] Jensen PR, Hammer K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol*, 1998, 64(1): 82–87.
- [23] Zhu YX, Li Y, Zheng XF. *Modern Molecular Biology*. 3rd ed. Beijing: Higher Education Press, 2007: 84 (in Chinese).
朱玉贤, 李毅, 郑晓峰. 现代分子生物学. 北京: 高等教育出版社, 2007: 78, 84.
- [24] Gruber TM, Gross CA. Assay of *Escherichia coli* RNA polymerase: sigma-core interactions. *Methods Enzymol*, 2003, 370: 206–212.
- [25] Ross W, Aiyar SE, Salomon J, et al. *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J Bacteriol*, 1998, 180(20): 5375–5383.
- [26] Guo T, Kong J, Zhang L, et al. Fine tuning of the lactate and diacetyl production through promoter engineering in *Lactococcus lactis*. *PLoS ONE*, 2012, 7(4): e36296.
- [27] Lu J, Tang J, Liu Y, et al. Combinatorial modulation of galP and glk gene expression for improved alternative glucose utilization. *Appl Microbiol Biotechnol*, 2012, 93(6): 2455–2462.

(本文责编 陈宏宇)