

## 研究论文

DOI: 10.12211/2096-8280.2024-046

## 大肠杆菌中终止子对下游转录单元基因表达的影响

任家卫<sup>1,2</sup>, 张金鹏<sup>1,2</sup>, 徐国强<sup>1,2</sup>, 张晓梅<sup>3</sup>, 许正宏<sup>4</sup>, 张晓娟<sup>1,2</sup>

(<sup>1</sup> 江南大学生物工程学院工业生物技术教育部重点实验室, 江苏 无锡 214122; <sup>2</sup> 江南大学粮食发酵与食品生物制造国家工程研究中心, 江苏 无锡 214122; <sup>3</sup> 江南大学生命科学与健康工程学院, 江苏 无锡 214122; <sup>4</sup> 四川大学轻工科学与工程学院, 四川 成都 610065)

**摘要:** 在基因转录过程中, RNA聚合酶通过识别启动子序列启动转录过程, 而当其识别到位于3'-UTR的终止子序列后, 转录复合物解离, 转录过程终止。因此, 转录单元内的启动子和终止子分别发挥启动和终止转录的作用。然而, 对于下游的转录单元, 终止子除了终止转录本通读这一直接作用以外, 其与RNA聚合酶之间的解离可能会影响后续转录单元中启动子与RNA聚合酶的结合, 从而间接改变下游转录单元的表达。这种跨转录单元的终止子和启动子之间的互作关系目前尚缺少研究, 因此, 明确终止子对下游转录单元的转录反应强度的影响对于精准调控基因表达、开发高效终止子具有重要的意义。本研究通过 one-pot 技术将9种终止子、5种间隔序列和9种启动子进行组合构建了一个包含405种不同组合元件(终止子-间隔序列-启动子)的组装文库, 基于FlowSeq技术对文库中所有组合进行测序和荧光强度分析, 进而建立组合序列-下游基因表达之间的相关性。结果表明弱终止子、短间隔以及强启动子的组合更有利于提高下游基因的表达。通过对转录本的定量分析发现弱终止子不仅提高下游的渗漏转录(提高21~70倍), 同时也促进了下游启动子重新招募RNA聚合酶进行重启转录(提高2~3倍)。本研究解析了终止子对下游转录单元基因表达的调控效果和机制, 为利用终止子构建基因回路提供设计依据。

**关键词:** 终止子; 启动子; 基因表达调控; one-pot组装; FlowSeq技术

中图分类号: Q819 文献标志码: A

## Effect of terminators on the downstream transcript unit with gene expression in *Escherichia coli*

REN Jiawei<sup>1,2</sup>, ZHANG Jinpeng<sup>1,2</sup>, XU Guoqiang<sup>1,2</sup>, ZHANG Xiaomei<sup>3</sup>, XU Zhenghong<sup>4</sup>, ZHANG Xiaojuan<sup>1,2</sup>

(<sup>1</sup> Key Laboratory of Industrial Biotechnology of Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China; <sup>2</sup> National Engineering Research Center for Cereal Fermentation and Food Biomanufacturing, Jiangnan University, Wuxi 214122, Jiangsu, China; <sup>3</sup> School of Life Science and Health Engineering, Jiangnan University, Wuxi 214122, Jiangsu, China; <sup>4</sup> College of Light Industry Science and Engineering, Sichuan University, Chengdu 610065, Sichuan, China)

**Abstract:** During gene transcription, RNA polymerase initiates the process by recognizing the promoter sequence,

收稿日期: 2024-06-17 修回日期: 2024-08-22

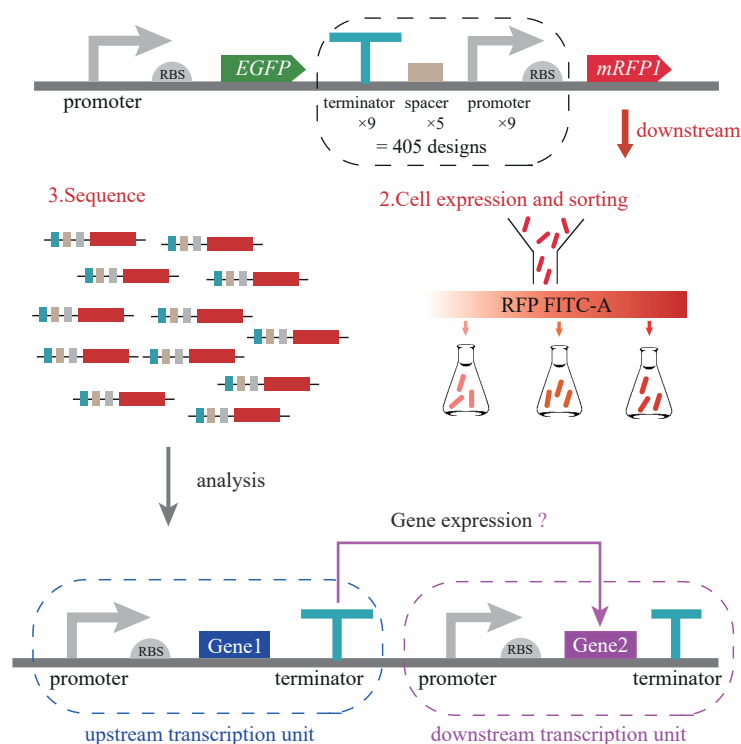
基金项目: 国家自然科学基金(32171421)

引用本文: 任家卫, 张金鹏, 徐国强, 张晓梅, 许正宏, 张晓娟. 大肠杆菌中终止子对下游转录单元基因表达的影响[J]. 合成生物学, 2025, 6(1): 213-227

Citation: REN Jiawei, ZHANG Jinpeng, XU Guoqiang, ZHANG Xiaomei, XU Zhenghong, ZHANG Xiaojuan. Effect of terminators on the downstream transcript unit with gene expression in *Escherichia coli*[J]. Synthetic Biology Journal, 2025, 6(1): 213-227

and terminates it upon recognizing the terminator sequence located at the 3'-UTR, leading to dissociation of the transcription complex. Therefore, promoters and terminators within the transcription unit play the role of initiating and terminating transcription, respectively. For downstream transcription units, in addition to the direct effect of terminating transcript read-through, the dissociation of the RNA polymerase from the terminator may affect the binding of the promoter to RNA polymerase in the subsequent transcription unit, thus indirectly altering the expression of the downstream transcription unit. This interplay between terminators and promoters across transcription units remains poorly understood, therefore, elucidating the impact of terminators on the transcriptional strength of downstream transcription units is of great significance for the precise regulation of gene expression and the development of efficient terminators. In this study, a library containing 405 different combinatorial elements (terminator-spacer-promoter) was constructed by combining nine terminators, five spacer sequences, and nine promoters using one-pot assembly technology. All combinations in the library were sequenced and analyzed in terms of fluorescence intensity based on the FlowSeq technology to establish the correlations between combinatorial sequences and downstream gene expression. The results showed that combinations of weak terminators, short spacers, and strong terminators were more favorable to enhance the expression of downstream genes, while combinations of strong terminators, long spacers, and weak terminators reduced the expression of downstream genes. Quantitative analysis of transcription revealed that weak terminators not only enhanced downstream leakage transcription (21~70-fold enhancement), but also facilitated downstream promoters to re-recruit RNA polymerase for re-promoted transcription (2~3-fold enhancement). This study has elucidated the effect and mechanism of terminators on the regulation of gene expression in the downstream transcription units, providing a design framework for the construction of gene circuits using terminators.

#### 1. Design of terminator library



**Keywords:** terminator; promoter; downstream gene expression regulation; one-pot assembly; FlowSeq technique

基因表达包含多个步骤,涉及多种基因表达调控元件,它们在不同水平发挥作用<sup>[1-6]</sup>。其中,终止子位于编码基因的3'-UTR区,当RNA聚合酶与终止子相遇时转录复合物发生解离,转录过程终止。原核体系中的终止子主要分为两类:ρ因子依赖性终止子和内源性终止子。其中内源性终止子终止转录主要依靠自身序列的折叠形成二级结构<sup>[7-9]</sup>,相较于ρ因子依赖型终止子来说更为简单且易于调控,围绕它展开的研究也更为丰富<sup>[10-12]</sup>。内源性终止子根据其终止效率的强弱在基因表达调控中的应用也各不相同,强终止子由于其较强的终止转录的能力,常插入目标基因下游,发挥及时终止转录、避免转录通读的作用,从而提高细胞经济性<sup>[13]</sup>;而弱终止子则常被用于插入特定基因上游,发挥弱化下游基因表达的作用<sup>[14]</sup>。

终止子序列组成和二级结构是影响终止子终止效率的重要因素<sup>[15]</sup>。Ahn等<sup>[16]</sup>的研究表明茎较长的终止子呈现更稳定的二级结构,终止效率也更高;Chen等<sup>[17]</sup>在大肠杆菌中设计了一个可以包含582条不同终止子的终止子文库,发现强终止子形成的二级结构中发卡基部位置处GC碱基含量较高,在环结构中则没有发现这样的规律;Cui等<sup>[18]</sup>对枯草芽孢杆菌中天然终止子的终止效率进行了评估,发现当终止子二级结构中U-tract位置处U碱基的含量增加时,终止子的终止效率也随之提高,同时,最小自由能ΔG的减小也会使终止子的终止效率提高;本课题组<sup>[19-20]</sup>的前期研究发现拥有稳定且较长的茎结构的终止子的终止效率更高,完美的U-tract结构可以进一步提高终止效率。

随着终止子领域研究的不断深入,越来越多的研究表明,除了终止基因的转录外,终止子还被广泛应用于下游基因表达的调控。Zhang等<sup>[14]</sup>在生产鸟氨酸的谷氨酸棒杆菌表达体系中插入不同强度的终止子去调控下游*argF*基因的表达,发现随着终止子强度的增加*argF*基因的表达强度也随之减弱;Du等<sup>[21]</sup>尝试在多个基因之间插入终止子,发现插入终止子后下游基因转录水平会下降,同时,随着插入的终止子数量的增加,下游基因转录水平会进一步下降。然而,目前关于终止子的研究大多是在转录单元内进行的,跨转录单元

的终止子调控机制尚不明确,系统解析终止子对下游转录单元中基因表达的调控效应,可以为构建基因回路提供设计依据。

为了解析终止子与下游启动子,以及它们中间间隔序列之间的联动或串扰效果,需要系统考察包含不同终止子、启动子和间隔序列各类组合序列,解析构效关系。one-pot组装技术是一种在T4连接体系上拓展优化的多片段组装技术,其基本原理是通过在多个片段间设计黏性末端并磷酸化,将磷酸化后的片段与线性化质粒置于同一试管中进行连接反应,完成片段与片段、片段与载体之间的连接。该方法可以一次性完成多个片段之间互相的随机组装,并将组装完成后的多片段组装产物与线性化质粒相连接,是构建高精度的组装文库的有效方法<sup>[22]</sup>。此外,传统关于终止子的研究主要通过通过在目的基因5'端或3'端插入终止子来检测目的基因转录和表达水平变化,其中最典型的就是通过荧光报告基因检测终止子的终止效率<sup>[17, 23-25]</sup>,尽管这一类方法可以准确分析终止子对目的基因的表达调控效果,但一次只能检测一种终止子的功能,具有通量小、耗时长等缺点。FlowSeq是一种将流式细胞分选与高通量测序相结合的技术,它的基本原理是利用流式细胞仪根据细胞的荧光特征等可检测因素对细胞进行分组,后续再使用高通量测序技术对分组后的细胞进行测序分析,快速获得不同表型亚群的基因信息<sup>[26-33]</sup>。one-pot组装技术结合FlowSeq技术,为收集大规模序列和调控效果信息、探究终止子在基因表达调控中的作用提供了解决方案。

本研究基于one-pot技术和FlowSeq技术,将9种不同强度的终止子、5种不同长度的间隔序列和9种强度不同的启动子进行组合,构建了一个包含405种不同组合元件(终止子-间隔序列-启动子)的组装文库<sup>[22]</sup>,按照下游红色荧光蛋白表达水平对文库进行分组,随后对文库进行DNA扩增并进行高通量测序<sup>[26-28]</sup>,从蛋白和转录本两个层面对终止子-间隔-启动子的组合调控效果进行分析,进一步明确了弱终止子对下游渗漏转录和重启转录两个方面的影响程度。本研究所得到的结论揭示了终止子在多基因和转录单元间的调控机制,为终

止子元件的理性开发、基因回路的理性构建提供了基础信息。

## 1 材料与方法

### 1.1 材料

#### 1.1.1 菌株与质粒

本论文所使用的菌株和质粒见表1。

#### 1.1.2 引物

本实验所使用引物见表2。

#### 1.1.3 培养基及试剂

LB培养基：10 g/L 蛋白胨，5 g/L 酵母提取物，10 g/L NaCl。

琼脂粉、蛋白胨、酵母粉、氯化钠和利福平购于国药集团化学试剂有限公司；Phanta 酶、RNA 提取试剂盒、反转录试剂盒和qPCR 试剂盒购于南京诺唯赞生物科技有限公司；限制性核酸内切酶、T4 多聚核苷酸激酶和 T4 连接酶购于 Takara 公司；BSA I 酶购于 NEB 公司。

### 1.2 方法

#### 1.2.1 菌株和生长条件

以大肠杆菌 BL21 (DE3) 为宿主构建质粒。菌株在 LB 培养基中培养。需要固体培养基时，在上述培养基中加入 2% 的琼脂粉。抗生素浓度为 50.0 mg/L 的卡那霉素。需要表达报告基因时，加

入 0.5 mmol/L 的 IPTG。

#### 1.2.2 反向 PCR 插入酶切位点

通过反向 PCR 的方式在质粒 PTK 上插入 BSA I 酶切位点，反向 PCR 的反应流程如下：①预变性 95 °C 3 min；②变性 95 °C 15 s；③退火 58 °C 15 s；④延伸 72 °C 5 min；⑤重复③、④两个步骤 35 次；⑥彻底延伸 72 °C 5 min。反应完成后使用胶回收试剂盒进行直接回收，然后使用 DPN I 酶 37 °C 酶切 1.5 h 消化模板 DNA，接着使用胶回收试剂盒进行直接回收，回收完成后加入 BL21 (DE3) 感受态中进行转化。

#### 1.2.3 One-pot 组装文库的构建

将终止子、间隔序列和启动子的正向链和反向链引物的干粉稀释至 10 μmol/L，各取 10 μL 混合，然后进行退火反应，反应程序如下：95 °C 5 min，94 °C 1 min，93 °C 1 min，92 °C 1 min，91 °C 1 min，90 °C 1 min，之后立即将样品放入沸水中冷却至室温。将退火后的终止子、间隔和启动子的双链 DNA 各取 1 μL 混合，接着取 20 μL 混合后的样品对其使用 T4 多聚核苷酸激酶进行磷酸化，结束后放 4 °C 备用。使用 BSA I 酶对插入了 BSA I 酶切位点的质粒进行酶切，酶切完成后使用胶回收试剂盒进行直接回收得到线性化质粒。接着，使用 T4 连接酶将磷酸化后的样品与线性化的质粒进行连接，连接的体系如下：线性化质粒 50 fmol，磷酸化混合组件 3 μL，T4 连接酶 1 μL，T4 连接 buffer 2 μL，无酶水加至 20 μL，随后 16 °C 过夜连接。最后，将连接产物分装到 10 只大肠杆菌 BL21 感受态细

表1 本研究所涉及的菌株和质粒信息

Table 1 Strains and plasmids used in this study

| Name                          | Description                                                                                                | Source     |
|-------------------------------|------------------------------------------------------------------------------------------------------------|------------|
| <i>Escherichia coli</i> JM109 |                                                                                                            | EUROSCARF  |
| <i>Escherichia coli</i> BL21  |                                                                                                            | EUROSCARF  |
| PTK                           | Dual Fluorescent Probe Plasmid, Kan <sup>r</sup> , T7 promoter, <i>EGFP</i> , <i>mRFP1</i> , T7 terminator | this study |
| PTK-pot                       | PTK derivatives carrying one-pot assembly libraries                                                        | this study |
| PTK-pheA-13-e                 | PTK derivatives carrying the terminator pheA, the spacer sequence sp-13, and the promoter dap-e            | this study |
| PTK-pheA-13-e12               | PTK derivatives carrying the terminator pheA, the spacer sequence sp-13, and the promoter dap-e12          | this study |
| PTK-pheA-40-e                 | PTK derivatives carrying the terminator pheA, the spacer sequence sp-40, and the promoter dap-e            | this study |
| PTK-pheA-40-e12               | PTK derivatives carrying the terminator pheA, the spacer sequence sp-40, and the promoter dap-e12          | this study |
| PTK-recA-13-e                 | PTK derivatives carrying the terminator recA, the spacer sequence sp-13 and the promoter dap-e             | this study |
| PTK-recA-13-e12               | PTK derivatives carrying the terminator recA, the spacer sequence sp-13 and the promoter dap-e12           | this study |
| PTK-recA-40-e                 | PTK derivatives carrying the terminator recA, the spacer sequence sp-40 and the promoter dap-e             | this study |
| PTK-recA-40-e12               | PTK derivatives carrying the terminator recA, the spacer sequence sp-40 and the promoter dap-e12           | this study |

表2 本实验所用的引物

Table 2 Primers used in this study

| Primer        | Description                                                                |
|---------------|----------------------------------------------------------------------------|
| pheA-F        | TAGCAACAATAAGGCTCCCAAATCGGGGGGCTTTTTATTGAT                                 |
| pheA-R        | TTAGATCAATAAAAAAGGCCCCCGATTGGGAGGCCTTATTGTT                                |
| thrL-F        | TAGCTCAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTACT                              |
| thrL-R        | TTAGAGTAAAAAAAAGCCCGCACTGTCAGGTGCGGGCTTTTTTGA                              |
| rpsO-F        | TAGCCAGAAAAGGGGGCCTGAGTGGCCCCCTTTTTCAAGCT                                  |
| rpsO-R        | TTAGAGCTTGAAAAAGGGGCCACTCAGGCCCCCTTTTCTG                                   |
| arcA-F        | TAGCAATAAAAAACGCGCTAAAAAGCGCCGTTTTTTTGACG                                  |
| arcA-R        | TTAGCGTCAAAAAAACGCGCTTTTAGCGCCGTTTTTATT                                    |
| fhuE-F        | TAGCGTAAAAAAGGCAGCCATCTGGCTGCCTTAGTCTCCCA                                  |
| fhuE-R        | TTAGTGGGGAGACTAAGGCAGCCAGATGGCTGCCTTTTTAC                                  |
| hisI-F        | TAGCCCCTGCCCTTTTTCTTTAAACCGAAAAGATTACTTCGCGT                               |
| hisI-R        | TTAGACGCGAAGTAATCTTTTCGGTTTTAAAGAAAAAGGGCAGGG                              |
| recA-F        | TAGCAAGCAAAAGGGCCGAGATGCGACCCCTGTGTATCAAC                                  |
| recA-R        | TTAGGTTGATACACAAGGGTCGCATCTGCGGCCCTTTTGCTT                                 |
| T7-MOD-F      | TAGCAAACAGATAGGCCCTCtcgGAGGGCCtatctgtTTTTTTT                               |
| T7-MOD-R      | TTAGAAAAAaAcagataGGCCCTCcgaaGAGGGCCTATCTGTTT                               |
| SP-TE-F       | TAGCCCAATGTTTACTCATATCCAGTCACAGAACTGAACTATC                                |
| SP-TE-R       | TTAGGATAGTTCAGTTTCTGTGACTGGATATGAGTAAACATTGG                               |
| SPACE-5-F     | CTAACTATC                                                                  |
| SPACE-5-R     | TTATGATAG                                                                  |
| SPACE-13-F    | CTAAAACTGAACTATC                                                           |
| SPACE-13-R    | TTATGATAGTTCAGTTT                                                          |
| SPACE-20-F    | CTAAGTCACAGAACTGAACTATC                                                    |
| SPACE-20-R    | TTATGATAGTTCAGTTTCTGTGAC                                                   |
| SPACE-28-F    | CTAACATATCCAGTCACAGAACTGAACTATC                                            |
| SPACE-28-R    | TTATGATAGTTCAGTTTCTGTGACTGGATATG                                           |
| SPACE-35-F    | CTAAGTTTACTCATATCCAGTCACAGAACTGAACTATC                                     |
| SPACE-35-R    | TTATGATAGTTCAGTTTCTGTGACTGGATATGAGTAAAC                                    |
| SPACE-40-F    | CTAACCAATGTTTACTCATATCCAGTCACAGAACTGAACTATC                                |
| SPACE-40-R    | TTATGATAGTTCAGTTTCTGTGACTGGATATGAGTAAACATTGG                               |
| dap-A16-F     | ATAATTGTTTAACCCCAAATGAGGGAAGAAGGTATAATTGAACTCTCGCTCAAGGCGCAAGGAGCACACACA   |
| dap-A16-R     | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCAATTATACCTTCTCCCTCATTGGGGGTAAACAA     |
| dap-e-F       | ATAATTGTTTAGCCACCAAATGAGGGAAGAGGCACAATGGAACCTCTCGCTCAAGGCGCAAGGAGCACACACA  |
| dap-e-R       | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCCATTGTGCCTCTTCCCTCATTGGTGGCTAAACAA    |
| dapA-e10-F    | ATAATTGTTTGTACACCAAATGAGGGAATGTGGTATAATTGAACTCTCGCTCAAGGCGCAAGGAGCACACACA  |
| dapA-e10-R    | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCAATTATACCACATTCCCTCATTGGTGTCAAAACAA   |
| dapA-e11-F    | ATAATTGTTTGTACACCAAATGAGGGAATGTGCTATAATGGAACCTCTCGCTCAAGGCGCAAGGAGCACACACA |
| dapA-e11-R    | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCCATTATAGCACATTCCCTCATTGGTGTCAAAACAA   |
| dapA-e12-F    | ATAATTGTTTGTACACCAAATGAGGGAATGTGGTAGAGTGGAACCTCTCGCTCAAGGCGCAAGGAGCACACACA |
| dapA-e12-R    | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCCACTTACCACATTCCCTCATTGGTGTCAAAACAA    |
| dapA-e10-35-F | ATAATTGTTTAACCCCAAATGAGGGAATGTGGTATAATTGAACTCTCGCTCAAGGCGCAAGGAGCACACACA   |
| dapA-e10-35-R | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGAGTTCAATTATACCACATTCCCTCATTGGGGGTAAACAA    |
| J1-F          | ATAATTGACAATTTTCTTAAATTGTGTTACAATGGGTTTCGCTCAAGGCGCAAGGAGCACACACA          |
| J1-R          | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAAACCCATTGTAACACAATTTAAGAAAATTGTCAA          |
| J2-F          | ATAATTGACATTTTTTAGTTTTGAGTTACAATGGTTGTCGCTCAAGGCGCAAGGAGCACACACA           |
| J2-R          | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGACAACCATTGTAACCTAAAACTAAAAAATGTCAA           |
| SP-PRO-F      | ATAAGTTTACTCATATCCAGTCACAGAACTGAACTATCTCGCTCAAGGCGCAAGGAGCACACACA          |
| SP-TRO-R      | TCATTGTGTGTGCTCCTTGCGCCTTGAGCGAGATAGTTCAGTTTCTGTGACTGGATATGAGTAAAC         |
| T7-CG-F       | AATCCGGCCGCGGGGCCGCTTCGGCGGGCCCCGCGGCCGTTTTTTAACTGAACTATCG                 |
| T7-CG-R       | GATCCGATAGTTCAGTTTAAAAAACGCGCGGGGGCCCGGAAGCGGGCCCCGCGGCCG                  |
| WEAK-F        | AATTCAGCCCCCTAGTATAGGGGCGTATTTCTCAAACCTGAACTATCG                           |
| WEAK-R        | GATCCGATAGTTCAGTTTGAAGAAATACGCCCTTACTGAGGGGCTG                             |

胞中, 放冰上冷却 20 min, 随后在 42 °C 下加热 90 s, 再次放冰上冷却 5 min。之后向每个试管中加入 LB 培养基, 使最终体积达到 1 mL。在 37 °C、220 r/min 的条件下培养 1 h 后, 将所有菌液转移到 100 mL 添加了卡那霉素的 LB 液体培养基中过夜培养。

#### 1.2.4 流式细胞分选

取 2 mL 过夜培养后的菌液加入 100 mL 添加卡那霉素的 LB 液体培养基中, 并加入终浓度为 0.5 mmol/L 的 IPTG, 在 37 °C、220 r/min 条件下诱导培养 2 h。接着取 1 mL 菌液在 8000 r/min 转速下离心获得细胞, 用 PBS 溶液洗涤两次后, 再加入 1 mL PBS 溶液重悬细胞。随后使用流式细胞仪对细胞进行筛选, 流式细胞仪 (Becton Dickinson FACS Aria III) 的激发波长为 561 nm, 发射波长为 610/20 nm, 检测 *mRFP1* 荧光。使用 FlowJo v10 软件分析流式细胞仪的数据并计算平均荧光强度。最后根据下游红色荧光的强度将细胞进行分组。

#### 1.2.5 高通量测序

将收集到的细胞在 37 °C、220 r/min 的条件下复苏 2 h, 随后接种到 100 mL 添加卡那霉素的 LB 培养基中, 在 37 °C 下以 220 r/min 的速度振荡培养 12 h, 取 1.5 mL 菌液提取质粒, 使用目标引物进行 PCR 扩增获得 DNA 扩增产物, 随后将从每个组中各取 1000 ng DNA 扩增片段进行高通量测序, 测序平台为华大的 MGISEQ-2000 平台。

#### 1.2.6 数据分析

从 fastq 文件中提取终止子序列的过程如下: 使用 cutadapt (1.2.1) 软件切除序列, 并使用 prinseq-lite (0.19.5) 软件进行质量控制, 确保每个碱基的质量值大于 30; 此外, U-tract 位点前三个碱基中至少有一个 T 的序列也被视为终止序列并予以保留<sup>[34]</sup>; 利用 Vienna RNA package (2.5.0) 工具包中的 RNA 折叠和 RNA 评估软件预测二级结构, 并获得每个终止序列的最小自由能  $\Delta G$ 。

#### 1.2.7 相对荧光强度强度的计算

相对荧光加权强度 (relative fluorescence intensity of *mRFP1*) 的计算公式如下:

$$F_{Px} = \frac{Px \text{ 中组合元件的 reads}}{Px \text{ 中的总 reads}} \quad (1)$$

$$F_{re, Px} = \frac{F_{Px}}{F_{P4} + F_{P5} + F_{P6}} \quad (2)$$

$$\text{Relative fluorescence intensity of } mRFP1 = (-100 \times F_{re, P4}) + (650 \times F_{re, P5}) + (5000 \times F_{re, P6}) \quad (3)$$

式中,  $Px$  表示流式分选出的组 ( $P4$ 、 $P5$  和  $P6$  三组,  $x$  值为 4、5 和 6);  $-100$  为  $P4$  组中的平均荧光表达强度;  $650$  为  $P5$  组中的平均荧光表达强度;  $5000$  为  $P6$  组中的平均荧光表达强度。

#### 1.2.8 mRNA 水平检测

划线接种于固体 LB 平板 (Kan 抗性) 上并于 37 °C、220 r/min 条件下培养 12 h。接着挑取单菌接种到 10 mL LB 液体培养基 (Kan 抗性) 中, 37 °C、220 r/min 培养 12 h, 然后取 200  $\mu$ L 菌液加入到 10 mL LB 液体培养基 (Kan 抗性) 中, 37 °C、220 r/min 培养 1.5 h 使 OD 值在 0.4~0.6 之间。取 1 mL 菌液使用 RNA 提取试剂盒从中提取 RNA, 再通过反转录试剂盒将 RNA 逆转录为 cDNA, 最后通过 Quantitative Real-time PCR (qPCR) 的方式对 mRNA 水平进行检测, 使用的内参基因为 16sRNA。

#### 1.2.9 mRNA 水平的计算

mRNA 水平的计算方法如下:

$$\Delta C_{q,x} = C_{q,x} - C_{q,r} \quad (4)$$

$$\Delta C_{q,cb} = C_{q,cb} - C_{q,cr} \quad (5)$$

$$\Delta \Delta C_q = \Delta C_{q,x} - \Delta C_{q,cb} \quad (6)$$

$$\text{mRNA level} = 2^{-\Delta \Delta C_q} \quad (7)$$

式中,  $x$  表示样本,  $C_{q,x}$  是其目标基因的量化周期,  $C_{q,r}$  是其内参基因的量化周期;  $cb$  是对照样本,  $C_{q,cb}$  是其量化周期,  $C_{q,cr}$  是其内参基因的量化周期。

#### 1.2.10 渗漏 mRNA 和重启 mRNA 水平的计算

对下游 mRNA、渗漏 mRNA 和重启 mRNA 水平进行归一化处理, 方法如下:

渗漏 mRNA 水平计算见 1.2.9。

下游 mRNA 水平计算方法如下:

(1) 首先参考 1.2.9 的方法计算出  $\Delta \Delta C_{q,s}$ , 使用的数据为测量渗漏 mRNA 时得到的目标基因和内参基因的量化周期。

(2) 接着参考 1.2.9 的方法计算出  $\Delta \Delta C_{q,t}$ , 使用的数据为测量下游 mRNA 时得到的目标基因和内参基因的量化周期。

(3) 计算出渗漏 mRNA 占下游 mRNA 的比值, 公式如下:

$$\frac{\text{leakage transcription mRNA level}}{\text{downstream transcription unit mRNA level}} = \frac{1}{2^{-(\Delta \Delta C_{q,s} - \Delta \Delta C_{q,t})}} \quad (8)$$

(4) 归一化后下游 mRNA 计算公式如下:

$$\text{downstream transcription unit mRNA level} = \frac{\text{leakage transcription mRNA level} \times 1}{\text{leakage transcription mRNA level}} \quad (9)$$

重启 mRNA 水平计算公式如下:

$$\text{re-promoted transcription mRNA} = \text{downstream transcription unit mRNA level} - \text{leakage transcription mRNA level} \quad (10)$$

## 2 结果和分析

### 2.1 不同强度终止子对下游转录单元基因表达的影响强度分析

前期的研究结果表明, 终止子除了会影响转

录单元内的基因表达调控, 对于后续转录单元的基因表达同样会产生影响<sup>[19, 26]</sup>。为了探究终止子对后续转录单元基因表达的影响, 在双荧光探针质粒 PTK 的两个荧光报告基因之间插入了一个终止子和启动子, 形成了两个不同的转录单元 [图 1(a)], 其中上游转录单元的基因为 *EGFP*, 下游转录单元的基因为 *mRFP1*。随后分别选取课题组前期定量分析过的 3 种不同强度的终止子 T7-CG、T7 和 weak 以及 3 种不同强度的启动子 *dap-e12*、*dap-e* 和 *sp-pro*<sup>[19, 35]</sup>, 替换双荧光探针质粒上的两个报告基因之间的终止子和启动子, 得到了 9 种不同的突变质粒。分别对它们的下游转录单元中红色荧光报告基因的表达水平和 mRNA 水平进行定量分析, 结果如图 1(b)、(c) 所示, 不同强度的启动子调控下的红色荧光报告基因 mRNA 水平相差最高达到了 700 倍, 红色荧光表达水平相差最高也达到了

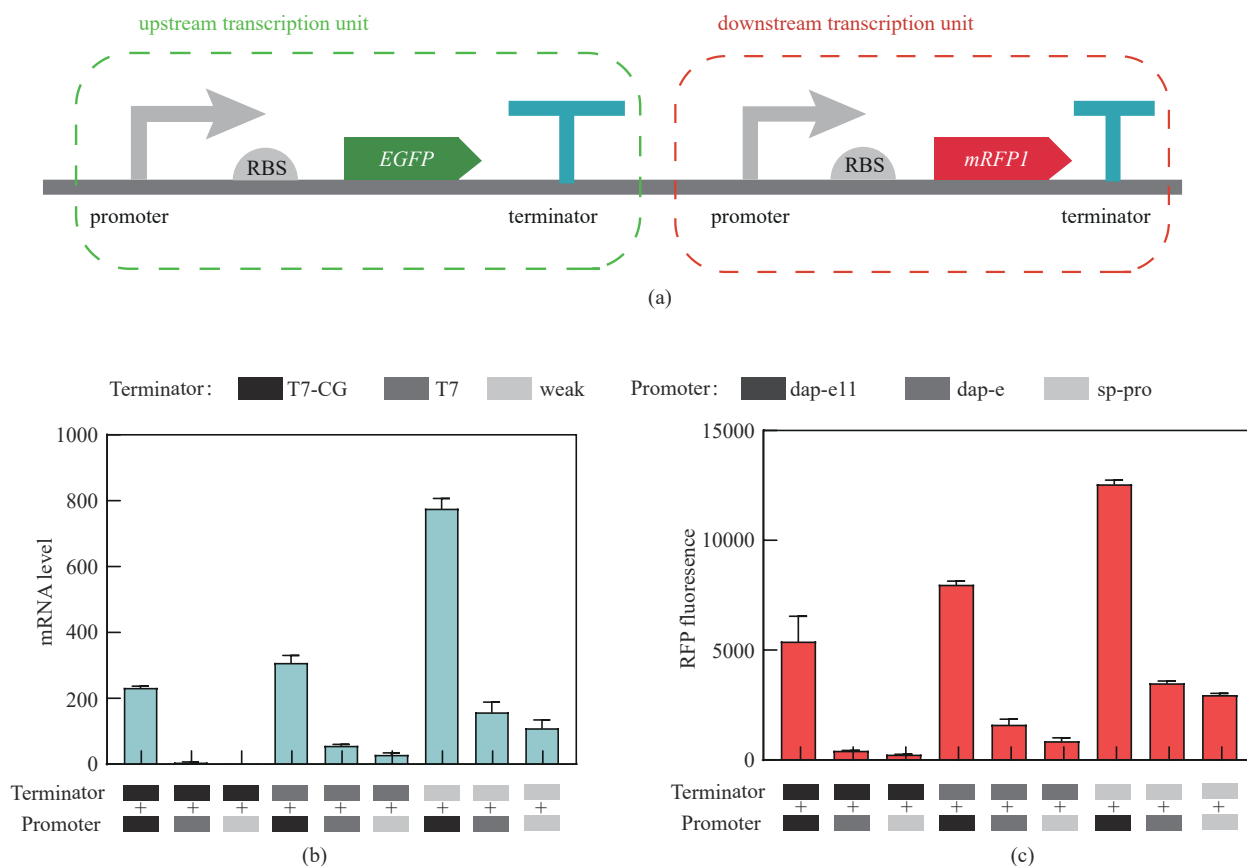


图1 不同强度终止子和启动子组合调控效果

质粒中上游和下游转录单元结构示意图 (a) 以及 9 种突变质粒中下游转录单元红色荧光转录水平 (b) 和表达水平 (c)

Fig. 1 Analysis of the regulatory effect of the combination of terminators and promoters with different strengths

Schematic structure of upstream and downstream transcriptional units in the plasmid (a), red fluorescent transcript levels (b), and expression levels (c) of downstream transcriptional units in the nine mutant plasmids

300倍。

启动子在一个转录单元中发挥着启动转录的功能,对基因表达调控的影响最大,这样的结果并不意外,然而当终止子强度发生变化时,下游转录单元的基因表达调控也会发生显著改变,弱终止子 weak 调控下的下游转录单元的红色荧光报告基因 mRNA 水平为强终止子 T7-CG 调控下的3~102倍,红色荧光表达水平相差了2~11倍,这一结果初步体现了弱终止子对后续转录单元的基因表达是更有利的,并且其对下游基因表达的影响是显著的。

## 2.2 通过 One-pot 结合 FlowSeq 技术解析终止子-间隔-启动子组合效应

为了扩大采集的样品量,系统探究终止子对后续转录单元基因表达的影响,设计了9种不同的终止子、5种不同的间隔序列和9种不同的启动子

元件,具体见表3(其中 SP-TE 是没有终止效率的一段序列, SP-PRO 是没有启动转录功能的一段序列)。采用 one-pot 技术将它们互相组合构建了一个包含405种不同组合元件(终止子-间隔-启动子)的组装文库<sup>[27]</sup>,使用 FlowSeq 技术根据下游报告基因(*mRFP1*)的荧光强度对细胞进行分选并测序,解析终止子对后续转录单元基因表达的影响。

one-pot 组装文库构建完成后,通过流式分选按照下游报告基因(*mRFP1*)的荧光强度对细胞进行分组,流式分选结果如图2(a)所示。其中 P4、P5 和 P6 组收集的细胞平均荧光表达强度为-100、650 和 5000。对 P4、P5 和 P6 组中收集到的细胞进行高通量测序,并按以下标准进行数据筛选,包括:①根据 Lesnik 等总结的终止子最基本的序列特征,即 U 通道前的开头3个碱基必须包含一个 U 碱基<sup>[36]</sup>;②剔除只检测到一个或两个读数的序列。

由于组合元件大多数在三组中都有被检测到,但在各组间分布频率存在差异,因此根据各组件

表3 文库中包含的组件

Table 3 Parts in the library

| Assembly   | Part Name   | Activity | Sequence(5'→3')                                           |
|------------|-------------|----------|-----------------------------------------------------------|
| Terminator | pheA        | Strong   | AATCATAAGGCCTGACAAATCGGGGGGATTTTTCTTGA                    |
|            | thrL        | ↓        | TCCAAAAAGCCCCGACCTGACAGTGACGGCTTTTTATTAC                  |
|            | T7-MOD      | weak     | AAACAGATAGGCCCTCTTCGGAGGGCCTATCTGTTTTTTTT                 |
|            | arcA        |          | AATATAAACGCGCTAAAACGCGCGGTTTTTTTTTGCC                     |
|            | rpsO        |          | CAGCAAAGGGTGCCTGAGTGCTCCCTTTTGCAAGC                       |
|            | fhuE        |          | GTAAAAAAGGCAGCCATCTGGCTGCCTTAGTCTCCCCA                    |
|            | hisl        |          | CCCTGCCCTTTTTCTTTAAAACCGAAAAGATTACTTCGCGT                 |
|            | recA        |          | AAGCAAAAGGGCCGCGAGATGCGACCCTTGTTGTATCAAC                  |
|            | SP-TE       |          | CCAATGTTTACTCATATCCAGTCACAGAACTGAACTATC                   |
| Spacer     | SP-13       | Short    | AAACTGAACTATC                                             |
|            | SP-20       | ↓        | GTCACAGAACTGAACTATC                                       |
|            | SP-28       | Long     | CATATCCAGTCACAGAACTGAACTATC                               |
|            | SP-35       |          | GTTTACTCATATCCAGTCACAGAACTGAACTATC                        |
|            | SP-40       |          | CCAATGTTTACTCATATCCAGTCACAGAACTGAACTATC                   |
| Promoter   | dapA-e12    | Strong   | TTGTTTTGACACCAAATGAGGGAATGTGGTAGAGTGGAATC                 |
|            | dapA-e11    | ↓        | TTGTTTTGACACCAAATGAGGGAATGTGCTATAATTGGAATC                |
|            | dapA-e10    | weak     | TTGTTTTGACACCAAATGAGGGAATGTGGTATAATTGGAATC                |
|            | dap-A16     |          | TAGGTTTTTTCGGGGTTGTTTAAACCCCAAATGAGGGAAGAAGGTATAATTGAACTC |
|            | J2          |          | TTGACATTTTTTAGTTTTGAGTTACAATGGTTG                         |
|            | dapA-e10-35 |          | TTGTTTAAACCCCAAATGAGGGAATGTGGTATAATTGAACTC                |
|            | J1          |          | TTGACAATTTCTTAAATTGTGTTACAATGGGTT                         |
|            | dap-e       |          | TTGTTTAGCCACCAAATGAGGGAAGAGGCACAATGGAATC                  |
|            | SP-PRO      |          | GTTTACTCATATCCAGTCACAGAACTGAAC                            |

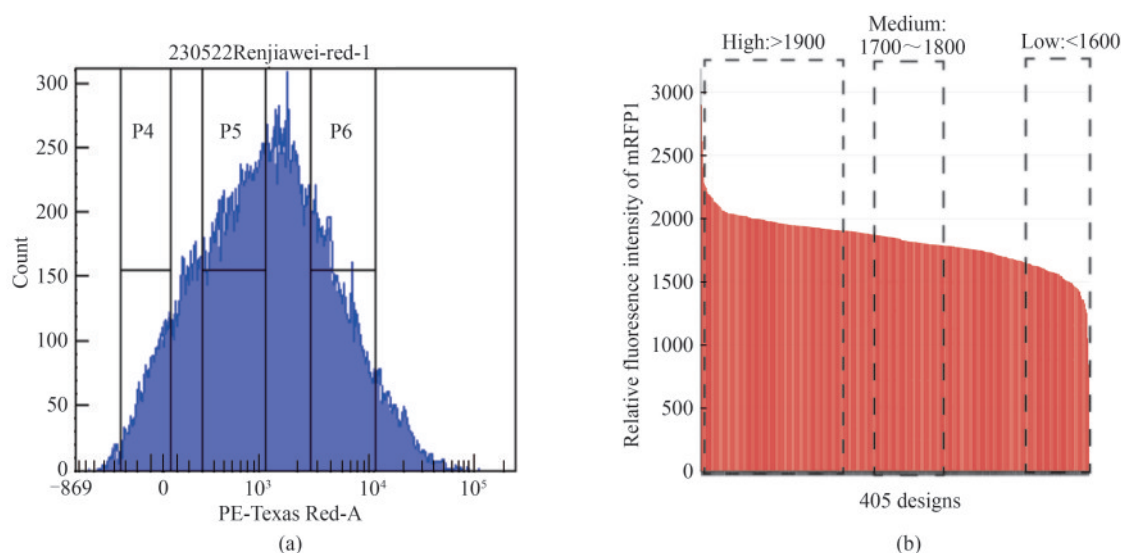


图2 组合调控元件（终止子-间隔-启动子）调控的下游基因表达强度

(a) 根据下游红色荧光表达水平对包含组合元件的细胞进行流式筛选；

(b) 根据下游相对荧光加权强度对组合元件进行高表达、中表达、低表达分组

Fig. 2 Expression of downstream genes regulated by the assemblies of terminator-spacer-promoter

(a) Screening of combinatorial elements by flow cytometry based on the red fluorescence for the expression of targeted genes; (b) Grouping of the assemblies of terminator-spacer-promoter into High, Medium, and Low expression groups based on the relative fluorescence intensity of *mRFP1* associated with the expression of targeted genes

在三组中的分布频率，对它们的荧光强度进行了加权计算，得出对应的相对加权荧光强度，并按照强度分为高表达组、中表达组和低表达组[图2(b)]。

各元件在高、中和低表达组中出现的频率如图3(a)、(b)所示。终止子元件中，无终止子活性的和弱终止子活性的SP-TE和recA在高表达组中出现频率较高，均超过了0.5；终止效率较高的T7-MOD、thrL和pheA在低表达组中出现频率较高，其中pheA出现的频率达到了0.59。间隔序列中，短的间隔序列SP-13和SP-20在高表达组中出现频率较高，均超过了0.4；长间隔序列SP-40在低表达组中出现频率较高（0.42）。启动子元件中，强启动子dap-e12和J2在高表达组中出现频率较高，分别为0.63和0.62；而弱启动子或无转录活性的dap-e和SP-PRO在低表达组中出现频率较高，达到了0.7。

上述结果体现弱终止子、短间隔以及强启动子都分别对提高下游转录单元的基因表达有促进作用。相反的强终止子、长间隔以及弱启动子的组合则会削弱下游转录单元的基因表达。而组合文库中，也可以发现上述的组合也体现了这一规

律，例如：弱终止子recA、短间隔sp-13和强启动子dap-e12组合后的元件的相对加权荧光强度为2408，为典型的“高表达”组元件；而强终止子thrL、长间隔sp-40和弱启动子SP-PRO组合后的元件的相对荧光加权强度为1254，为典型的“低表达”组元件[图2(b)]。

### 2.3 终止子对下游转录单元基因表达具有重要影响

上述研究通过one-pot组装文库结合FlowSeq技术，初步解析了不同强度的终止子、启动子与它们之间的间隔对下游转录单元基因表达的影响。然而启动子、终止子都在转录水平发挥调控作用，为了更加准确直接地表征调控组件对基因表达的调控效果，挑选了one-pot组装文库中典型的组合元件，采用qPCR定量分析下游转录单元基因转录水平差异。共计从文库中挑选出了8种不同的组合元件，它们由6种不同的元件组成（强终止子pheA、弱终止子recA、长间隔序列sp-40、短间隔序列sp-13、强启动子dap-e11和弱启动子dap-e）。结果如图4(a)所示，其中下游转录单元mRNA水平最高的组合元件为弱终止子recA-短间隔sp-13-

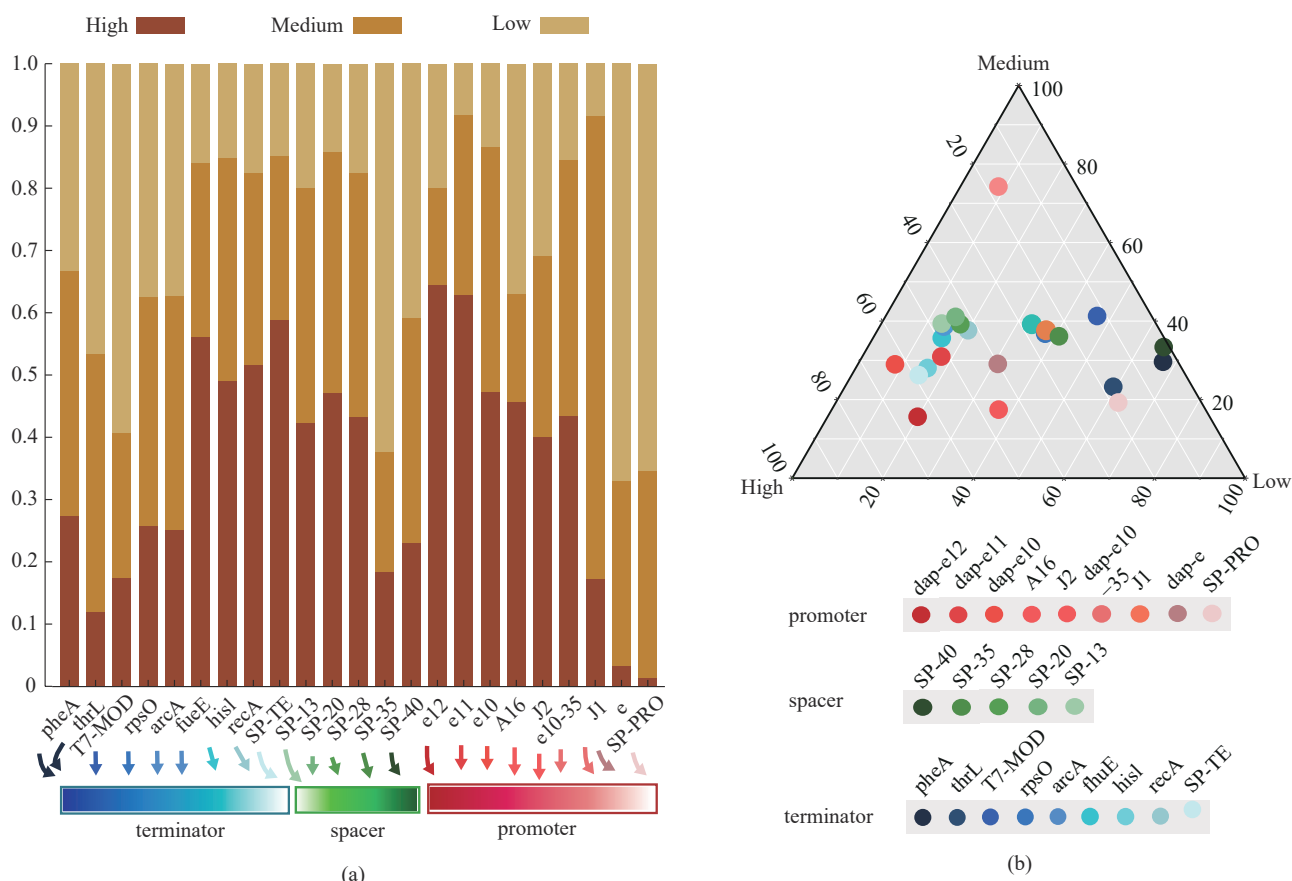


图3 组件中的各元件在高、中和低表达组的出现频率

(a) 高、中和低表达组中不同元件出现频率的叠加；(b) 高、中和低表达组中不同元件出现频率的三元图

Fig. 3 Frequency of each part used in the assemblies at the “High”, “Medium”, and “Low” groups

(a) Stacked histograms of the occurrence frequency of different components in the “High”, “Medium”, and “Low” groups; (b) Ternary results of the occurrence frequency of different components in the “High”, “Medium”, and “Low” groups

强启动子 *dap-e11*；相反，下游转录单元 mRNA 水平最低的组合元件为强终止子 *pheA*-长间隔 *sp-40*-弱启动子 *dap-e*，两个组合元件调控下的下游转录单元 mRNA 水平相差了 10 倍。对这 8 种不同组合元件调控下的下游转录单元红色荧光蛋白表达进行了检测 [图 4(b)]，发现了相似的规律，即弱终止子、较短的间隔序列以及强启动子调控下的下游转录单元荧光蛋白表达水平是所有组合中对基因表达积极作用最明显的。2.2 节的结论也发现它们各自调控下的基因表达最强，而它们组合后，也体现出正向的叠加作用。在构建多转录单元的基因回路时，遵循这一设计依据，理论上可以提高下游目标基因的表达。

进一步统计各元件对下游转录单元 mRNA 水平和荧光蛋白表达量的影响幅度，结果如图 5 所示。当组合元件中终止子和启动子都相同时，短

间隔 *sp-13* 调控下的下游转录单元 mRNA 水平是长间隔序列 *sp-40* 调控下的 1.2~1.6 倍 [图 5(a)]；当组合元件中终止子和间隔序列都相同时，较强启动子 *dap-e11* 调控下的下游转录单元 mRNA 水平是较弱启动子 *dap-e* 调控下的 1.5~2.1 倍 [图 5(b)]；当组合元件中间隔序列和启动子都相同时，弱终止子 *recA* 调控下的下游转录单元 mRNA 水平是强终止子 *pheA* 调控下的 3.5~4.8 倍 [图 5(c)]。下游红色荧光蛋白表达水平的定量分析结果也体现类似的结果。

总结认为，除了公认的启动子调控下游基因表达外，终止子和间隔序列也在一定程度上影响着下游转录单元的基因表达，其中间隔序列的影响较小，但终止子的影响是不可忽略的，弱终止子 *recA* 调控下的下游转录单元 mRNA 水平最高达到了强终止子 *pheA* 调控下的 4.8 倍，红色荧光蛋白

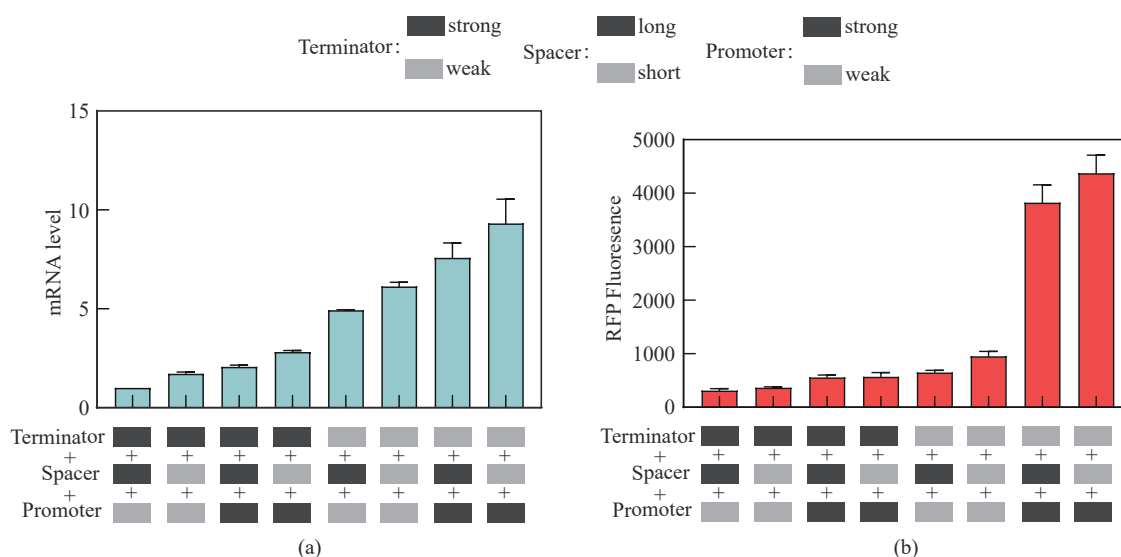


图4 典型组合元件调控的下游基因mRNA水平 (a) 和红色荧光蛋白表达水平 (b)

(强弱启动子分别为dap-e11和dap-e, 强弱终止子分别为pheA和recA, 长短间隔序列分别为sp-40和sp-13)

Fig. 4 Expression levels of mRNA and red fluorescent protein with the downstream transcription unit regulated by typical assemblies selected from the one-pot assembly library

(The strong and weak promoters were dap-e11 and dap-e, the strong and weak terminators were pheA and recA, and the long and short interval sequences were sp-40 and sp-13, respectively.)

表达水平最高相差了6.9倍, 因此选用强度合适的终止子可以有效调控下游转录单元的基因表达。

## 2.4 弱终止子促进下游转录单元的重启转录和渗漏转录

上述结果表明, 弱终止子有利于下游转录单元的基因表达, 然而, 这一促进作用是单纯由于弱终止子没有有效关闭转录造成的下游转录通读 (leakage transcription mRNA) 从而导致渗漏转录现象的发生, 还是由于弱终止没有造成RNA聚合酶及时解离, 从而提高RNA聚合酶的循环效率, 促进了其与下游转录单元中启动子的结合, 从而提高了重启转录效率 (re-promoter mRNA)。为了分辨弱终止子对下游基因表达的促进机制, 设计了两对引物 (上游引物不同, 下游引物一致), 分别针对不同片段进行扩增, 具体引物设计位置如图6(a)所示。其中, 引物F1在上游EGFP基因的末端位置, 引物R在下游mRFP1的中部, 扩增的片段为渗漏转录mRNA; 引物F2在下游转录单元RBS位置, 引物R在下游mRFP1的中部, 扩增的片段为渗漏转录和重启转录两种情况的加合。利用这两对引物对转录本逆转录形成的cDNA进行

定量分析, 通过对两对引物qPCR得到的结果进行比较, 可以得到下游转录单元中启动子招募RNA聚合酶启动转录生成的重启转录mRNA。

图6(b)和图6(c)分别表示下游mRNA水平和渗漏转录mRNA水平, 从图6(c)中不难看出, 拥有弱终止子recA的组合元件调控下的渗漏转录mRNA水平比任意一个拥有强终止子pheA的组合元件调控下的渗漏转录mRNA水平都要高。需要强调的是, 如图6(d)所示, 弱终止子调控下的渗漏转录最高达到了下游转录单元mRNA水平的30%, 相比之下, 强启动子调控下的渗漏转录的情况占比较低, 占了下游渗漏转录mRNA水平的1%~5%。重启转录mRNA水平结果如图6(e)所示, 弱终止子调控下的重启转录mRNA水平相较于强终止子提高了2~3倍。因此上述结果表明, 弱终止子不仅更有利于下游渗漏转录水平的提高, 同时也更有利于提高下游转录单元中启动子重新招募RNA聚合酶, 重新启动转录。这两方面的共同作用, 导致了下游转录单元基因的高表达。

## 3 结论

在原核生物的基因表达调控中, 除了终止基

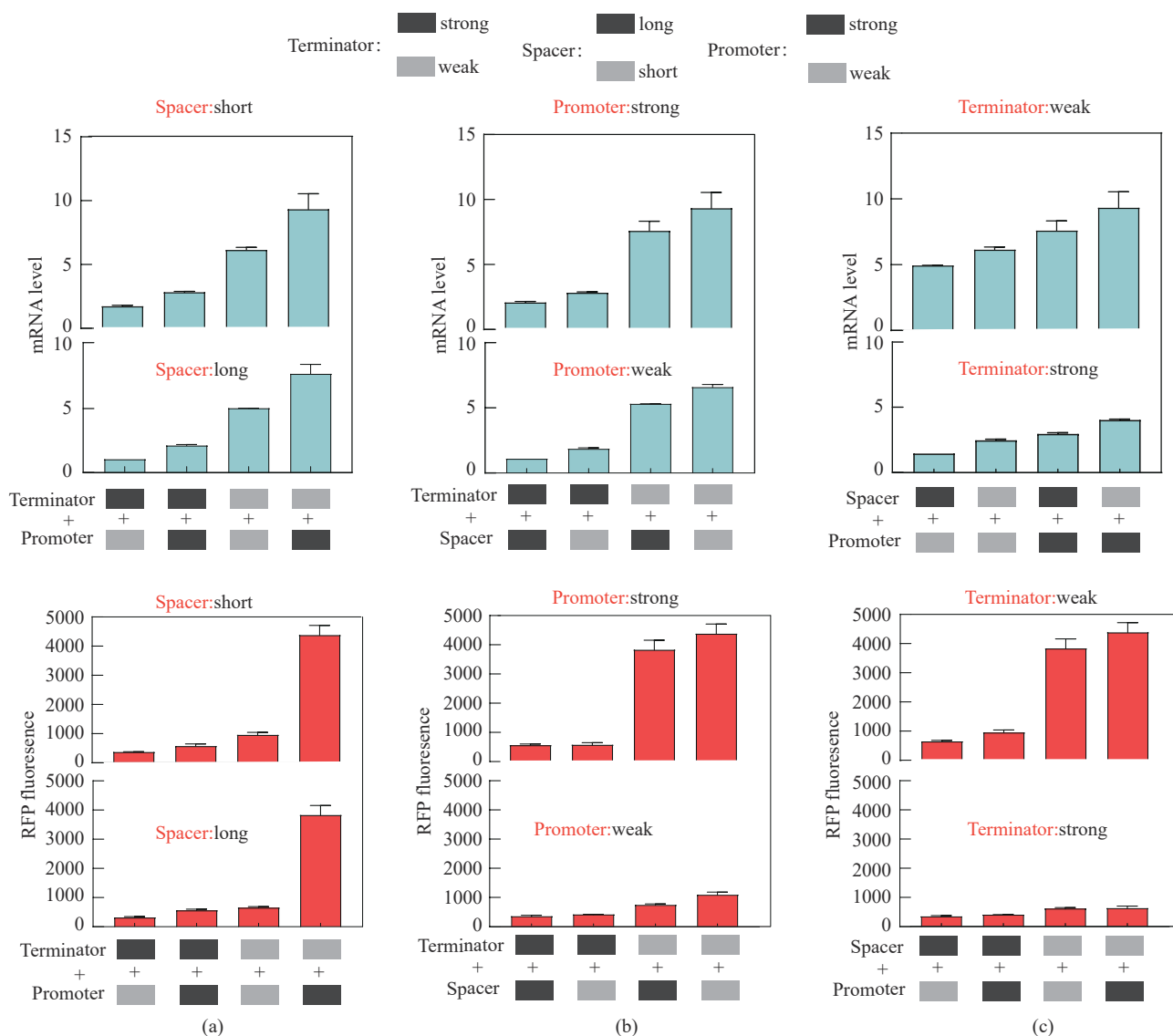


图5 不同强度启动子和终止子组合对下游转录单元 mRNA 水平和荧光蛋白的影响

(a) 固定间隔序列的情况下不同终止子和启动子组合对基因表达的影响；(b) 固定启动子的情况下不同间隔序列和终止子组合对基因表达的影响；(c) 固定终止子的情况下不同启动子和启动子组合对基因表达的影响；强弱启动子分别为 *dap-e11* 和 *dap-e*，强弱终止子分别为 *pheA* 和 *recA*，长短间隔序列分别为 *sp-40* 和 *sp-13*

**Fig. 5** Influence of the combination of promoters and terminators on expression of the mRNA and fluorescence protein

(a) Effect of different terminator and promoter combinations on gene expression with fixed interval sequences; (b) Effect of different spacer sequences and terminator combinations on gene expression with fixed promoters; (c) Effects of different promoters and combinations of promoters on gene expression with fixed terminators. The strong and weak promoters were *dap-e11* and *dap-e*, the strong and weak terminators were *pheA* and *recA*, and the long and short interval sequences were *sp-40* and *sp-13*, respectively.

因转录，避免转录通读外，终止子还影响着上游基因 mRNA 稳定性和下游基因表达的调控，然而，目前尚缺乏跨转录单元的终止子调控机制的相关研究，明确这一机制对于终止子元件的开发以及基因回路的构建有着重要的指导意义。

本研究基于 *one-pot* 技术构建了一个组装文库，同时基于 *FlowSeq* 技术对文库进行高效的基因型和

表型信息的采集。明确了弱终止子、短间隔和强终止子更有利于提高下游基因的表达。同时基于典型组合的转录本和蛋白表达量的分析也表明弱终止子调控下不仅通过渗漏转录提高了下游基因表达，同时也促进了下游转录单元中启动子重新招募 RNA 聚合酶从而进行的重启转录过程。本研究明确的弱终止对下游转录单元的影响是未来合

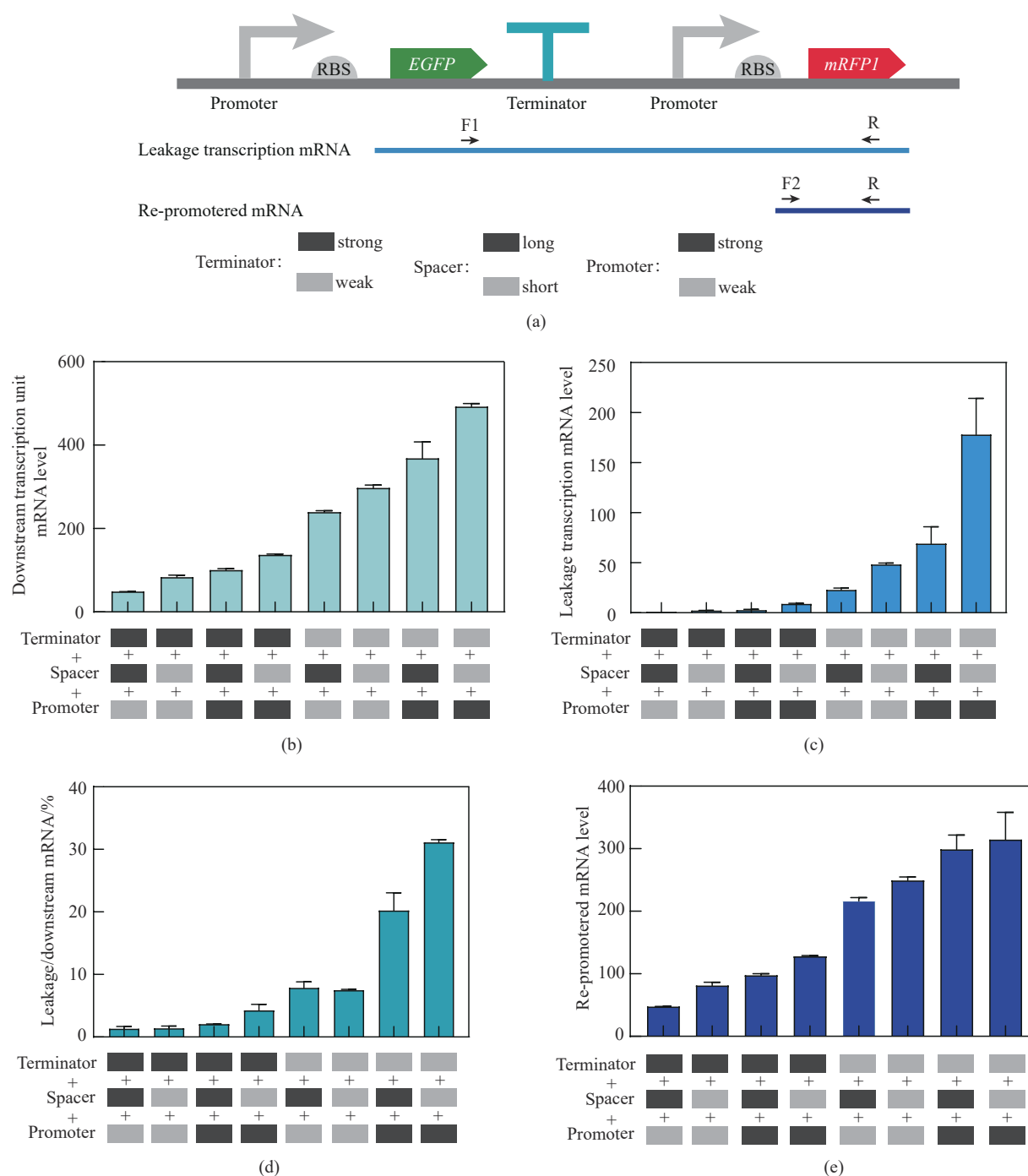


图6 不同组合元件调控下的下游整体mRNA、渗漏转录mRNA以及重启mRNA水平分析

(a) 通过不同引物的qPCR获得渗漏转录mRNA和重启转录mRNA示意图；(b) 不同组合元件调控下的下游mRNA水平；(c) 不同组合元件调控下的渗漏转录mRNA水平；(d) 渗漏转录mRNA在下游mRNA中的占比；(e) 不同组合元件调控下的重启转录的mRNA水平；强弱启动子分别为dap-e11和dap-e，强弱终止子分别为pheA和recA，长短间隔序列分别为sp-40和sp-13

Fig. 6 Schematic representation of downstream mRNA, leakage transcription mRNA, and re-promoted mRNA levels under the regulation of different combinatorial elements.

(a) Schematic representation of leakage transcription mRNA and re-promoted mRNA; (b) Downstream mRNA levels under the regulation of different combinatorial elements; (c) Leakage transcription mRNA levels under the regulation of different combinatorial elements; (d) Percentage of Leakage transcription mRNA in downstream mRNA; (e) Re-promoted mRNA levels under the regulation of different combinatorial elements; The strong and weak promoters were dap-e11 and dap-e, the strong and weak terminators were pheA and recA, and the long and short interval sequences were sp-40 and sp-13, respectively.

成生物学领域中多顺反子和多转录单元的基因回路构建中需要遵循的依据。本研究得到的元件组合也是高效的基因表达调控元件,具有实际应用价值。

## 参 考 文 献

- [1] LAURSEN B S, SØRENSEN H P, MORTENSEN K K, et al. Initiation of protein synthesis in bacteria[J]. *Microbiology and Molecular Biology Reviews*, 2005, 69(1): 101-123.
- [2] MORENO J M, SØRENSEN H P, MORTENSEN K K, et al. Macromolecular mimicry in translation initiation: a model for the initiation factor IF2 on the ribosome[J]. *IUBMB Life*, 2000, 50(6): 347-354.
- [3] RAMAKRISHNAN V. Ribosome structure and the mechanism of translation[J]. *Cell*, 2002, 108(4): 557-572.
- [4] ZHOU S H, DU G C, KANG Z, et al. The application of powerful promoters to enhance gene expression in industrial microorganisms[J]. *World Journal of Microbiology & Biotechnology*, 2017, 33(2): 23.
- [5] XU N, WEI L, LIU J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis[J]. *World Journal of Microbiology & Biotechnology*, 2019, 35(2): 33.
- [6] ZALATAN F, PLATT T. Effects of decreased cytosine content on rho interaction with the rho-dependent Terminator *trp t'* in *Escherichia coli*[J]. *Journal of Biological Chemistry*, 1992, 267(27): 19082-19088.
- [7] HENKIN T M, YANOFSKY C. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions[J]. *BioEssays*, 2002, 24(8): 700-707.
- [8] ABE H, ABO T, AIBA H. Regulation of intrinsic Terminator by translation in *Escherichia coli*: transcription termination at a distance downstream[J]. *Genes to Cells*, 1999, 4(2): 87-97.
- [9] RAY-SONI A, BELLECOURT M J, LANDICK R. Mechanisms of bacterial transcription termination: all good things must end[J]. *Annual Review of Biochemistry*, 2016, 85: 319-347.
- [10] LEE D N, PHUNG L, STEWART J, et al. Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences[J]. *Journal of Biological Chemistry*, 1990, 265(25): 15145-15153.
- [11] KIREEVA M L, KASHLEV M. Mechanism of sequence-specific pausing of bacterial RNA polymerase[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(22): 8900-8905.
- [12] CHAN C L, WANG D G, LANDICK R. Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways [J]. *Journal of Molecular Biology*, 1997, 268(1): 54-68.
- [13] MAIRHOFER J, WITTEW A, CSERJAN-PUSCHMANN M, et al. Preventing T7 RNA polymerase read-through transcription-a synthetic termination signal capable of improving bioprocess stability[J]. *ACS Synthetic Biology*, 2015, 4(3): 265-273.
- [14] ZHANG B, YU M, ZHOU Y, et al. Improvement of L-ornithine production by attenuation of *argF* in engineered *Corynebacterium glutamicum* S9114[J]. *AMB Express*, 2018, 8(1): 26.
- [15] JU X W, LI D Y, LIU S X. Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria[J]. *Nature Microbiology*, 2019, 4(11): 1907-1918.
- [16] AHN J H, KANG T J, KIM D M. Tuning the expression level of recombinant proteins by modulating mRNA stability in a cell-free protein synthesis system[J]. *Biotechnology and Bioengineering*, 2008, 101(2): 422-427.
- [17] CHEN Y J, LIU P, NIELSEN A A K, et al. Characterization of 582 natural and synthetic terminators and quantification of their design constraints[J]. *Nature Methods*, 2013, 10(7): 659-664.
- [18] CUI W J, LIN Q, HU R C, et al. Data-driven and *in silico*-assisted design of broad host-range minimal intrinsic terminators adapted for bacteria[J]. *ACS Synthetic Biology*, 2021, 10(6): 1438-1450.
- [19] ZHAI W J, DUAN Y T, ZHANG X M, et al. Sequence and thermodynamic characteristics of terminators revealed by FlowSeq and the discrimination of terminators strength[J]. *Synthetic and Systems Biotechnology*, 2022, 7(4): 1046-1055.
- [20] 翟伟绩. 大肠杆菌终止子文库构建及构效关系分析[D]. 无锡: 江南大学, 2022.
- ZHAI W J. The construction of the terminator library and analysis of the structure-activity relationship[D]. Wuxi: Jiangnan University, 2022.
- [21] DU L P, GAO R, FORSTER A C. Engineering multigene expression *in vitro* and *in vivo* with small terminators for T7 RNA polymerase[J]. *Biotechnology and Bioengineering*, 2009, 104(6): 1189-1196.
- [22] TARNOWSKI M J, GOROCHOWSKI T E. Massively parallel characterization of engineered transcript isoforms using direct RNA sequencing[J]. *Nature Communications*, 2022, 13(1): 434.
- [23] LI R, ZHANG Q, LI J B, et al. Effects of cooperation between translating ribosome and RNA polymerase on termination efficiency of the rho-independent terminator[J]. *Nucleic Acids Research*, 2016, 44(6): 2554-2563.
- [24] LALANNE J B, TAGGART J C, GUO M S, et al.

- Evolutionary convergence of pathway-specific enzyme expression stoichiometry[J]. *Cell*, 2018, 173(3): 749-761.e38.
- [25] GOROCHOWSKI T E, ESPAH BORUJENI A, PARK Y, et al. Genetic circuit characterization and debugging using RNA-seq[J]. *Molecular Systems Biology*, 2017, 13(11): 952.
- [26] KOSURI S, GOODMAN D B, CAMBRAY G, et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(34): 14024-14029.
- [27] CAMBRAY G, GUIMARAES J C, ARKIN A P. Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in *Escherichia coli*[J]. *Nature Biotechnology*, 2018, 36: 1005-1015.
- [28] EVFRATOV S A, OSTERMAN I A, KOMAROVA E S, et al. Application of sorting and next generation sequencing to study 5'-UTR influence on translation efficiency in *Escherichia coli* [J]. *Nucleic Acids Research*, 2017, 45(6): 3487-3502.
- [29] YOON B J. Hidden Markov models and their applications in biological sequence analysis[J]. *Current Genomics*, 2009, 10(6): 402-415.
- [30] KUO S T, JAHN R L, CHENG Y J, et al. Global fitness landscapes of the Shine-Dalgarno sequence[J]. *Genome Research*, 2020, 30(5): 711-723.
- [31] OSTERMAN I A, CHERVONTSEVA Z S, EVFRATOV S A, et al. Translation at first sight: the influence of leading codons [J]. *Nucleic Acids Research*, 2020, 48(12): 6931-6942.
- [32] VERMA M, CHOI J, COTTRELL K A, et al. A short translational ramp determines the efficiency of protein synthesis[J]. *Nature Communications*, 2019, 10(1): 5774.
- [33] GAO R, YU K, NIE J K, et al. Deep sequencing reveals global patterns of mRNA recruitment during translation initiation[J]. *Scientific Reports*, 2016, 6: 30170.
- [34] LESNIK E A, SAMPATH R, LEVENE H B, et al. Prediction of rho-independent transcriptional terminators in *Escherichia coli*[J]. *Nucleic Acids Research*, 2001, 29(17): 3583-3594.
- [35] DUAN Y T, ZHAI W J, LIU W J, et al. Fine-tuning multi-gene clusters via well-characterized gene expression regulatory elements: case study of the arginine synthesis pathway in *C. glutamicum*[J]. *ACS Synthetic Biology*, 2021, 10(1): 38-48.
- [36] HUDSON A J, WIEDEN H J. Rapid generation of sequence-diverse terminator libraries and their parameterization using quantitative Term-Seq[J]. *Synthetic Biology*, 2019, 4(1): ysz026.



**通讯作者:** 张晓娟(1982—),女,教授。研究方向为酿造食品微生物开发与应用、新型工业菌株的基因表达调控元件开发。

E-mail: zhangxj@jiangnan.edu.cn



**第一作者:** 任家卫(1999—),男,硕士研究生。研究方向为终止子调控元件开发。

E-mail: 297887816@qq.com

广告索引:天津大学合成生物技术全国重点实验室(后彩一)/安及义实业(上海)有限公司(后彩二)/北京擎科生物科技股份有限公司(后彩三)/上海润度生物科技有限公司(后彩四)/安徽华恒生物科技股份有限公司(后彩五)/诚志生命科技有限公司(封三)