

述评

Review

原核生物的启动子及其研究方法进展

张小梅 彭莹 龙雨欣 龙中儿*

江西师范大学生命科学学院, 南昌市鄱阳湖湿地微生物资源开发与利用重点实验室, 南昌, 330022

* 通信作者, longzhonger@163.com

摘要 启动子是调控转录起始的最基本元件, 对原核生物的生存和适应至关重要。深入研究原核生物的启动子有助于了解原核生物的基因表达调控机制, 有利于构建智能和高效的工程菌株, 进而有利于提高重要代谢物的产量, 在合成生物学和代谢工程等领域均具有重要意义。近年来, 随着原核生物启动子研究成果的不断积累, 大量启动子的序列被克隆、鉴定以及改造, 并被尝试用于生产实践。本文介绍了原核生物启动子的结构和类型, 详述了启动子预测、结构分析、功能鉴定及其开发应用的方法, 最后对该领域的研究进行展望, 以期为筛选和鉴定新型原核生物启动子、提高靶基因的高效表达及优化代谢途径提供思路。

关键词 原核生物; 启动子; 筛选; 功能鉴定; 启动子改造

Progress in Prokaryotic Promoters and Their Research Methods

ZHANG Xiaomei PENG Xuan LONG Yuxin LONG Zhonger*

Nanchang Key Laboratory of Microbial Resources Exploitation & Utilization from Poyang Lake Wetland, College of Life Sciences, Jiangxi Normal University, Nanchang, 330022

* Corresponding author, longzhonger@163.com

DOI: 10.13417/jgab043.000943

Abstract Promoters are the most fundamental elements responsible for regulating transcription initiation and are essential for the survival and adaptation of prokaryotes. An in-depth study of the prokaryotic promoters helps to understand the regulatory mechanism of gene expression, which is conducive to constructing intelligent and efficient engineering strains and then helps to improve the production of important metabolites, which is of great significance to the fields of synthetic biology and metabolic engineering. In recent years, with the continuous accumulation of research results on prokaryotic promoters, a large number of promoter sequences have been cloned, characterized as well as modified, and attempted to be used in production practice. This article introduces the structure and types of prokaryotic promoters and elaborates on the methods of promoter prediction, structural analysis, functional identification, and development and application. Finally, it prospects research in this field, intending to provide ideas for screening and identifying new prokaryotic promoters, improving efficient expression of target genes, and optimizing metabolic pathways.

Keywords Prokaryotes; Promoters; Screening; Functional identification; Promoter modification

原核生物基因的表达水平和转录活性受多个调控元件的影响, 包括启动子 (Guo et al., 2023)、核糖体结合位点 (ribosome binding site, RBS) (Abdulmalek and Yazgan-Karataş, 2023)、信号肽 (Liu et al., 2023) 和终止子 (Lammens et al., 2023) 等。其中, 启动子是负责转录起始的最基本元件, 可以被依赖

DNA 的 RNA 聚合酶 (RNA polymerase, RNAP) 特异性识别而启动目标基因的转录 (Steitz, 2009)。原核生物启动子的研究是生命科学领域中一个重要的领域, 它可以帮助我们更加深刻理解原核生物基因表达的调控机制 (Shen et al., 2022; Wu et al., 2023)、原核生物与环境之间的相互关系 (Chen et al.,

基金项目: 本研究由国家自然科学基金项目 (31960015) 资助。

引用格式: 张小梅, 彭莹, 龙雨欣, 等, 2024. 原核生物的启动子及其研究方法进展. 基因组学与应用生物学, 43(6): 943-957. [ZHANG X M, PENG X, LONG Y X, et al., 2024. Progress in prokaryotic promoters and their research methods. Genomics and Applied Biology, 43(6): 943-957.]

通信作者简介: 龙中儿, 教授, 从事微生物资源与生理生化研究。

收稿日期: 2023-11-26; 接受日期: 2024-02-03

2023), 在生命科学基础研究 (Zhang et al., 2023b)、工业微生物的改良 (Huo et al., 2023)、微生物重要代谢产物的生产 (Pan et al., 2022) 等领域均有重要意义。

目前, 有关原核生物启动子的研究主要集中在启动子的预测和分析、功能验证、结构分析以及开

发应用等四个方面 (图 1)。本文综述了该领域的研究进展, 具体内容包括原核生物启动子的结构及其类型、原核生物启动子的预测分析、原核生物启动子的结构分析和功能验证、原核生物启动子的改造及其开发应用。最后对原核生物启动子的研究进行展望, 以期为相关研究提供参考。

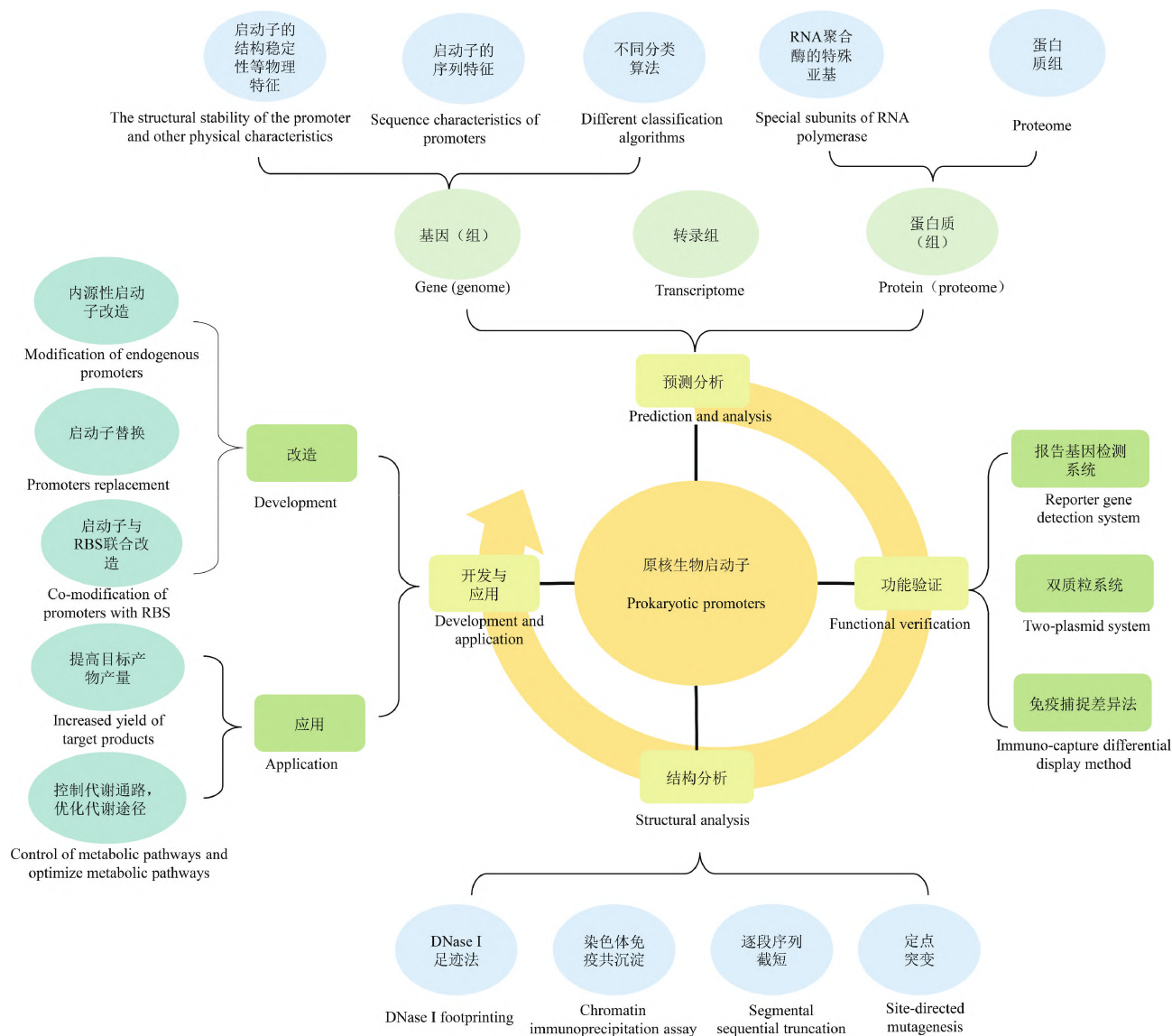


图 1 原核生物启动子研究的主要内容和方法

Figure 1 The main content and methods of research on prokaryotic promoter

1 原核生物启动子的结构

原核生物启动子通常位于转录起始位点 (transcription start site, TSS) 上游, 包含基因转录起始所需的复杂调控元件, 即相对于 TSS 的 -10 区和 -35 区、两区之间的间隔序列以及 TSS 上游保守序列 (图 2), 其中 -10 区和 -35 区构成启动子的核心序列。此外, 在大肠杆菌 (*Escherichia coli*) (Lu et al.,

2023)、枯草芽孢杆菌 (*Bacillus subtilis*) (Jacob et al., 2022)、戈登氏菌 (*Gordonia*) (Jaishankar et al., 2021) 和类球红细菌 (*Rhodobacter sphaeroides*) (Henry et al., 2021) 等的某些启动子上还存在扩展 -10 区。在转录过程中, 上游调控元件、-35 区和 -10 区之间的间隔序列长度以及扩展 -10 区都可能影响启动子的活性, 而核心启动子序列是转录所需的最基本结构。

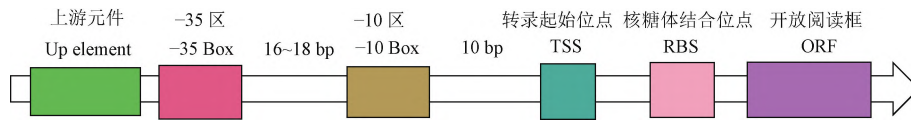
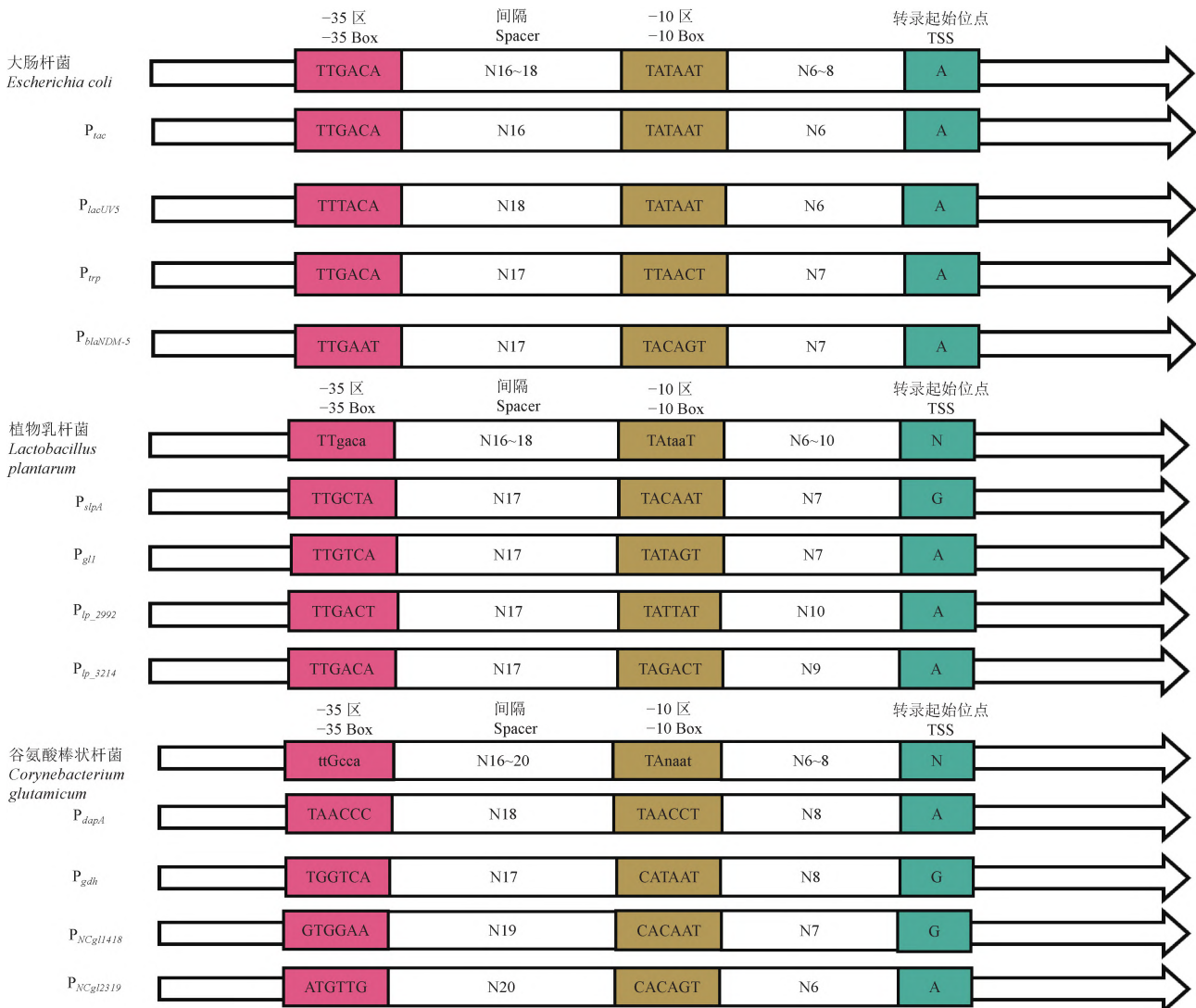


图2 原核生物中常见启动子核心序列的保守性

Figure 2 Conserved core sequences of common promoters from prokaryotes

在革兰氏阴性菌中，不同启动子的核心序列保守性较高，如大肠杆菌大多数启动子的-10区序列通常为TATAAT，-35区序列为TTGACA (Liu et al., 2004; Deng et al., 2018; Zhao et al., 2018; Li et al., 2022)。然而，在植物乳杆菌 (*Lactobacillus plantarum*) (McCracken et al., 2000; Spangler et al., 2022) 和谷氨酸棒状杆菌 (*Corynebacterium glutamicum*) (Pátek et al., 1996; Liu et al., 2020b; Huang et al., 2021b) 等革兰氏

阳性菌中，不同启动子的核心序列保守性较差 (图3)。存在这种差异的主要原因是不同原核生物中 σ 因子识别序列的特异性不同。 σ 因子是RNAP的辅助因子，与RNAP核心酶相互作用，形成RNAP全酶，可识别特定的调控序列，指导复合物至启动子区域 (Paget and Helmann, 2003; Feklistov et al., 2014)。这也导致了不同微生物之间有些启动子无法通用，或者呈现出来的转录效果存在显著差异。



启动子以其所启动转录的基因和/或操纵子 (排斜体字) 进行命名，前面冠以字母“P”表示。如启动色氨酸操纵子结构基因转录的启动子记为“P_{trp}”，启动谷氨酸脱氢酶基因转录的启动子记为“P_{gdh}”。启动子核心序列中大写字母表示超过80%的启动子在此位置出现的核苷酸，小写字母表示超过40%的启动子在该位置出现的核苷酸。

Promoters are named after the gene and/or operon (in italics) that they initiate transcribe, preceded by the letter "P". For example, the promoter that initiates the transcription of the structure gene of tryptophan operon is named as "P_{trp}", and the promoter that initiates the transcription of the glutamate dehydrogenase gene is named as "P_{gdh}". Capital letters in the core sequence of the promoter represent nucleotides present at the position in over 80% of promoters, while lowercase letters represent nucleotides present at the position in over 40% of promoters.

图3 3种重要原核生物的启动子保守序列

Figure 3 Conserved sequence of promoters from three important organisms

2 原核生物启动子的分类

按照不同的标准,原核生物启动子可以分为许多种。

2.1 基于转录水平的分类

根据对转录的调控强度不同,原核生物启动子可分为强启动子、中等强度启动子和弱启动子,其本质就是启动子与 σ 因子结合的能力不同。一般情况下,强启动子和中等强度启动子用于靶基因的高效表达,而弱启动子适用于弱化有毒物质编码基因的表达,或者用于降低某些特定代谢(如分支代谢)通路中某一支代谢途径中关键酶基因的表达,进而提高另一分支代谢途径中靶基因的表达量,以保持细胞内代谢通量的总量平衡。

2.2 基于作用方式的分类

根据作用方式的不同,启动子分为组成型启动子、诱导型启动子、生长依赖型启动子以及自诱导型启动子。

组成型启动子调控基因的表达水平比较稳定,通常用于提高重要化合物的产量,或调控合成生物学中关键基因的表达。如Li等(2023d)利用苏云金芽孢杆菌(*Bacillus thuringiensis*) BMB171中的组成型启动子P_{rsi}实现了Vip3Aa杀虫蛋白的高效生产。Kong等(2019)用2个组成型启动子来调节嗜热链球菌(*Streptococcus thermophilus*)胞外多糖(extracellular polysaccharides, EPS)生物合成基因簇中关键基因*epsA*和*epsE*的表达水平,提高了EPS的产量。此外,表征良好的组成型启动子也常常作为判断其他启动子强弱的对照启动子,如巴氏醋酸杆菌(*Acetobacter pasteurianus*)启动子P_{adh}(Gao et al., 2022)、李斯特菌(*Listeria monocytogenes*)高表达启动子(highly expressed *Listeria* promoter, P_{help})(Ma et al., 2021)以及链霉菌(*Streptomyces*)启动子P_{ermE*}、P_{hasO*}和P_{sp44}(Guo et al., 2023)等。组成型启动子

的研究成本低,但大多数为强启动子,难以控制靶基因的表达量,故不适于表达毒性蛋白相关基因。同时,在代谢通路中,并不是所有的基因都需要高表达,因此,在代谢工程中需要用一些弱启动子来平衡细胞代谢。

诱导型启动子启动下游基因转录的能力受诱导剂的影响。诱导剂通常包括物理诱导剂(如温度、光照和渗透压等)和化学诱导剂(如糖、营养物质、代谢物、金属离子、pH和各种溶剂等)。与组成型启动子相比,诱导型启动子因能按需控制基因的表达,常作为外源蛋白质的生产和代谢工程的关键工具(Toh et al., 2023)。当然,诱导型启动子也存在一些缺陷,如所依赖的诱导剂价格昂贵、对微生物的代谢产生压力、容易将外源污染物引入培养基等。因此,诱导型启动子在工业上有时也非最优选择。

生长依赖型启动子是指当菌体生长至某一特定阶段时可启动基因表达的启动子,不需添加任何诱导剂。如Yang等(2017)用枯草芽孢杆菌的3个生长依赖型启动子P_{abrB}、P_{spoVG}、P_{lyt}分别实现了酯酶、角蛋白酶和碱性多聚半乳糖醛酸裂解酶的高水平表达。Su等(2023)利用枯草芽孢杆菌的生长依赖型启动子P_{aprE},基于群体感应机制开发了自动调控表达系统,来调节细胞生长和角蛋白酶的生产,以降低角蛋白酶对细胞的毒性并促进生物量积累和角蛋白酶表达。这些研究成果加深了研究者对启动子调控策略的思考,但目前对这类启动子的研究主要集中在枯草芽孢杆菌(Kang et al., 2020; Zhu et al., 2022)和大肠杆菌(Wang et al., 2022)中,在其他原核生物中的相关报道较少。

自诱导型启动子是一类由 σ^S 识别并由起始膜蛋白基因、应激基因等抵抗环境胁迫相关基因表达的启动子。目前被报道的自诱导型启动子大多来源于大肠杆菌,如启动子P_{Jic}(Li et al., 2023a)。自诱导型启动子有利于降低目标蛋白工业化生产的成本,也是用于解决由过表达引起的代谢失衡的有效策

略,但天然自诱导型启动子因活性低而导致其应用受到限制。

3 原核生物启动子的预测分析

启动子的分析和预测对指导科学研究、确定控制基因转录起始的关键区域以及理解代谢调控网络具有重要意义。当前所采用的方法主要包括以下几种。

3.1 基于启动子的结构稳定性等物理特征进行预测分析

早期的原核生物启动子预测仅限于关键基序的分析(Huerta and Collado-Vides, 2003),但由于样本基因库和统计算法的缺乏,启动子预测的准确性不高。为快速区分原核生物某一 DNA 序列是否为启动子序列,研究人员根据原核生物启动子的结构稳定性和能量信息等物理特征开发了多种有效预测原核生物启动子的工具。

大部分原核生物启动子具有低稳定性、高曲率和低可弯曲性等特征,其中最为显著的特征就是启动子 DNA 双链与其周边序列稳定性的差异,Kanhere 和 Bansal(2005)据此提出了一种预测原核生物启动子的方法,并成功用于识别大肠杆菌(227 个)、枯草芽孢杆菌(89 个)和谷氨酸棒状杆菌(28 个)中的启动子。另外,根据 DNA 序列的其他物理特征预测启动子的文献也逐年增多,如 Chen 等(2020)基于启动子由不同结构的 DNA 序列决定,根据不同 DNA 局部结构的总能量不同,构建了一种用于启动子预测的统计物理模型,并用于预测大肠杆菌 K-12 中 σ^{38} 、 σ^{54} 和 σ^{70} 启动子,其准确性分别达到 87.0%、78.3% 和 86.4%。当然,单纯依据启动子的某一物理特征开发的预测工具其普适性和准确性有待提高,将启动子的多个特征相结合构建启动子预测模型或者将多个启动子预测模型联合使用,可以大幅度提高启动子预测的准确性和普适性。如 SEProm 是一种基于 DNA 结构和能量信息特征开发的启动子预测模型,几乎可用于预测所有的原核生物启动子,且具有较高的灵敏性及特异性(Mishra et al., 2020)。

3.2 基于启动子的序列特征进行预测分析

基因组测序技术的不断发展加速了基因组序列的累积过程,这为原核生物启动子预测提供了更为

有效的数据。研究人员发现,根据启动子在基因组上的有序分布、核苷酸偏好和 G+C 含量等序列特征,能快速地从基因组上识别启动子区域。

Jacques 等(2006)的研究表明,在许多细菌的基因组中,转录因子结合位点(transcription factor binding sites, TFBS)偏向于分布在编码区之间的区域,据此开发了一种既适用于启动子识别、又能预测出其他调控蛋白识别的 DNA 序列的工具。He 等(2018)依据 DNA 单链上特定位点的核苷酸序列偏好(或倾向性)建立了一种启动子预测模型,并用于预测大肠杆菌 K-12 中 σ^{70} 启动子,其准确性高达 94%。这类预测工具快速、高效,但由于这些模型的建立是基于已注释的基因组,其应用范围受到限制。为提高其应用潜能,有时会将启动子的序列特征与前述的结构特征相结合来开发启动子预测模型。如 Liu(2020a)等结合序列的 G+C 含量和弯曲性等特征构建启动子预测模型,并将其预测大肠杆菌、枯草芽孢杆菌和铜绿假单胞菌(*Pseudomonas aeruginosa*)的启动子,其准确度分别高达 81.5%、87.9% 和 90.4%。

3.3 基于 RNA 聚合酶的特殊亚基进行预测分析

启动子需要在 RNAP 和相应辅因子的共同作用下才能启动转录,这为原核生物启动子预测提供了一个新思路,即利用 RNAP 的特殊亚基识别启动子。

Spx 是一种在革兰氏阳性菌中具有高保守性的全局转录调控因子。Nakano 等(2010)的研究表明,Spx 与 RNAP 的 α 亚基 C 端(α CTD)相互作用形成的复合物可以识别 Spx 调控基因的启动子序列。Spx 被氧化为二硫化物形式的构象变化会导致 α 螺旋($\alpha 4$)结构发生改变,此时的 $\alpha 4$ 是激活转录以及 Spx- α CTD 复合物与启动子相互作用的特殊亚基,从而根据这一结合特征推测 Spx 调控基因的启动子序列。

该预测方式准确性高,但是操作比较复杂且应用范围窄,可辅助其他启动子的预测方式。

3.4 基于组学技术进行预测及分析

由于基因启动子的强度与基因的表达水平密切相关,随着测序技术和生物信息学的发展,转录组学测序和蛋白质组学分析已成为启动子预测、筛选的有效方法。转录组学测序又包括 RNA-seq 和微阵列两种技术。

RNA-seq 产生的大量 reads 可以映射到全基因

组序列,有利于筛选启动子、TSS和TFBS,并适用于大多数原核生物,如伯克霍尔德氏菌(*Burkholderiales*) (Ouyang et al., 2020)以及丙酮丁醇梭菌(*Clostridium acetobutylicum*) ATCC 824、贝氏梭菌(*Clostridium beijerinckii*) DSM 6423和贝氏梭菌 NCIMB 8052 (Hocq et al., 2022)等。此外,在RNA-seq过程中,一般将RPKM (reads per kilobase of transcript per million reads mapped)、FPKM (fragments per kilobase of transcript per million fragments mapped)或TPM (transcripts per million)值作为评判基因相对表达量的标准。因此, RNA-seq数据也可用于评估启动子的强度。如Shi等(2021)通过RNA-seq对球形红杆菌(*Rhodobacter sphaeroides*) JDW-710突变体和野生型菌株进行分析,以FPKM值作为基因相对表达量的评判标准,挑选出7个在JDW-710和野生型菌株中均表现出高转录活性的启动子,并用 P_{rsp_7571} 截短后得到的启动子 $P_{rsp_7571(T334*)}$ 过表达辅酶10合成途径中的两个关键基因 dxs 和 $ubie$,使得辅酶10的产量增加了28%。

基因微阵列技术将已知序列的DNA探针和多种寡聚核苷酸固定在芯片特定位置上,根据碱基互补配对原则,与荧光标记的待测基因进行杂交,之后根据杂交信号定量基因相对表达水平获得核酸序列信息,这有利于在特定条件下筛选启动子。如Jaisankar等(2021)根据NCBI中戈登氏菌IITR100的基因组序列信息,设计所有相应基因的探针,利用基因微阵列技术对生长至不同生长时期的戈登氏菌IITR100开展实验,获得差异表达基因信息,并据此筛选出2个生长依赖型启动子 P_{dsbA} 和 P_{glx} 。

蛋白质组学技术则是通过双向电泳凝胶可视化蛋白表达情况并联合质谱和微阵列测序数据确定高表达基因,筛选相应启动子。如Wang等(2019)通过蛋白质组学方法在查塔诺诺链霉菌(*Streptomyces chattanoogensis*) L10中筛选到1个可用于高效表达纳他霉素的强启动子 P_{groES} 。这种方式广泛用于链霉菌,有利于次级代谢产物特异性启动子的确定,但其灵敏性不高,可作为转录组筛选启动子的一种补充。

由于数据库等平台的信息共享,如今可以在网上获得许多原核生物的转录组测序数据,从而有利于启动子的识别。然而启动子强度并不是影响基因表达水平的唯一因素,还包括mRNA的稳

定性、细菌的培养条件和生长状况等其他因素。因此,为分离出有用的启动子,在分析组学数据的基础上还应该利用报告基因等方式系统地鉴定启动子的功能。

3.5 基于机器学习分类器进行的预测及分析

随着测序技术和计算机技术的不断发展,原核生物启动子预测也逐渐进入到依赖复杂的机器学习分类阶段,如线性判别分析、随机森林、支持向量机(support vector machine, SVM)和神经网络(neural networks, NN)等。目前,依赖SVM和NN算法的预测方式发展更为迅速。其中,SVM是一种对数据进行二元分类的算法,可表示复杂的线性关系。如Lin等(2023)建立了基于SVM的预测模型,用于分析产气克雷伯氏菌(*Klebsiella aerogenes*)的启动子序列和非启动子序列,其准确率高达95.8%。NN是一种模拟大脑神经网络的自适应机器学习技术。依赖人工神经网络的预测工具SAPPHIRE,可准确预测假单胞菌(*Pseudomonas*) σ^{70} 启动子(Coppens and Lavigne, 2020);依赖卷积神经网络技术的预测工具iPromoter-BnCNN,有效提高了对大肠杆菌中 σ 启动子识别和分类的准确性(Amin et al., 2020);依赖胶囊网络和卷积神经网络的预测工具iPromoter-CLA,实现了准确预测原核生物启动子的同时还能评估启动子的强度(Zhang et al., 2022)。

机器学习技术的不断完善使启动子预测方法的准确性得到了较大的提高。但现有的原核生物启动子预测工具大多依赖典型模式生物建立的数据库,如大肠杆菌启动子数据库RegulonDB (Santos-Zavaleta et al., 2019)、枯草芽孢杆菌启动子数据库DBTBS (Makita et al., 2004)和铜绿假单胞菌启动子数据库PRODORIC (Münch et al., 2003),导致有些预测工具的普适性不高,从而无法准确预测其他原核生物的启动子。因此,为提高启动子的预测性能,还需要进一步开发新的预测工具。

4 原核生物启动子的功能验证

原核生物启动子的功能验证,包括鉴定启动子启动转录的能力及其强度,对理解原核生物的基因表达调控具有重要意义。具体方法包括以下三个方面。

4.1 报告基因检测系统

报告基因检测系统是当前研究原核生物启动子活性的重要工具之一。该方法首先需要在无启动子的载体上插入合适的报告基因；其次需要将预测的启动子片段克隆到含有报告基因的载体上，转入到相应的宿主中；最后通过报告基因的表达情况来表征启动子的强度，并鉴定所研究启动子的顺式作用元件。

目前已经报道的报告基因有编码氯霉素乙酰转移酶基因(Kozakai et al., 2021)、 β -半乳糖苷酶基因(Brandwein et al., 2023)、 β -葡萄糖醛酸酶基因(Chae et al., 2019)、儿茶酚 2,3-双加氧酶基因(Luo et al., 2015)、绿色荧光蛋白基因(Lee et al., 2023)、红色荧光蛋白基因(Gao et al., 2022)和荧光素酶基因(Yu et al., 2023)等 7 个。前 3 个报告基因检测系统具有良好的稳定性，但因需要通过裂解细胞、添加额外的底物或测定酶活性来确定启动子的活性而导致其应用受到限制。后 3 个报告基因由于操作简单、安全、定量范围广、可应用于活细胞而得到广泛应用，但是荧光蛋白基因报告系统在检测启动子活性时会受到细胞自身背景荧光的干扰。而儿茶酚 2,3-双加氧酶基因常作为链霉菌中启动子活性鉴定的报告基因(Guo et al., 2023)。因此，实验过程中需要视具体情况选择合适的报告基因。

4.2 免疫捕捉差异显示法

免疫捕捉差异显示法是筛选和鉴定环境诱导型启动子的免疫学方法。该方法依赖微型 Tn5 转座子衍生的启动子探针载体，以编码 β -半乳糖苷酶和 β -半乳糖苷透性酶的 *lacZY* 作为报告基因。其中，Tn5 转座子可直接在含有乳糖作为唯一碳源的培养基上选择操纵子融合突变体 *lacZY*，以及在 X-gal 指示剂培养基上选择卡那霉素和 β -半乳糖苷酶的表达，通过观察指示剂 X-gal 的显色情况可以判断质粒中插入的 DNA 片段是否具有启动子活性。此外，还设计了针对 β -半乳糖苷透性酶的单特异性抗体，以便于直接通过免疫磁珠捕获技术分离出携带不同启动子的细胞。如 Timms-Wilson 等(2000)用这种方法检测了根际细菌(*Rhizobacteria*)中环境诱导启动子的活性。

免疫捕捉差异显示法可以快速、准确地鉴定环境中诱导型启动子，但主要适用于根际细菌，且目前用这种方式鉴定启动子的研究较少。

4.3 双质粒系统

双质粒系统是依赖两个具有相容性质粒的检测系统。该系统先在其中一个质粒中表达 σ 因子，获得的 σ 因子与细菌中 RNAP 核心酶相互作用形成 RNAP 全酶，而 RNAP 全酶可识别克隆在另一个含有报告基因的探针质粒上的启动子。目前的双质粒系统有大肠杆菌双质粒系统和谷氨酸棒状杆菌双质粒系统。其中，大肠杆菌双质粒系统通常将 *lacZ α* 作为报告基因，运用更为广泛，如 Homerova 等(2004)通过这种方法从金黄色葡萄球菌(*Staphylococcus aureus*)中鉴定到 18 个 σ^B 依赖型启动子，并确定了转录因子起始位点；此外，该系统已成功用于鉴定来自天蓝色链霉菌(*Streptomyces coelicolor*) A3 (2) 中的 σ^B 识别启动子(Sevcikova et al., 2021)和结核分枝杆菌(*Mycobacterium tuberculosis*) 中的 σ^F 识别启动子(Homerova et al., 2007)。而谷氨酸棒状杆菌双质粒系统是以 *gfpuv* 为报告基因，Busche 等(2023)通过该系统鉴定了谷氨酸棒状杆菌中的 σ^E 识别启动子。

大肠杆菌双质粒系统无法识别那些需要体内进行转录激活的启动子，而谷氨酸棒状杆菌双质粒系统可以实现在体内鉴定启动子，二者都仅适用于鉴定 σ 因子依赖型启动子，因此应用受到限制。

5 原核生物启动子的结构分析

原核生物启动子的结构分析对理解基因的表达调控机制至关重要，常用方法包括以下几种。

5.1 DNase I 足迹法

DNase I 足迹法能精确鉴定 DNA 结合蛋白在 DNA 分子上的结合位点，因此有利于找出启动子上的 TFBS。先通过放射性或荧光对目标启动子序列进行末端标记，将纯化好的转录因子蛋白与启动子进行孵育，随后加入 DNase I，启动子会被消化成长度不同的片段，得到的 DNA 片段可通过凝胶电泳分离、自显影或荧光成像观察。此时，未与转录因子结合的启动子区域会被 DNase I 降解，而与转录因子结合的启动子区域不会被降解，因此会在凝胶上显示“足迹”，最后结合测序技术即可找出 TFBS。如 Kim 等(2019)通过 DNase I 足迹实验在铜绿假单胞菌 *mexEF-oprN* 基因的启动子上鉴定了两个与 MexE-F-OprN 外排系统的转录激活因子 MexT 结合的区域，分别为 TCCATAACCATCGAC 和 CCATGTAT-

CACTGTTTCGTGAT。这一研究结果增强了对 MexT 依赖的药物外排泵激活调控的理解,也为抗铜绿假单胞菌活性的药物开发提供了新线索。

在启动子研究中, DNase I 足迹法也常与凝胶阻滞实验(electrophoretic mobility shift assays, EMSA) 联用。其中, EMSA 用于判断转录因子与目标启动子是否结合, DNase I 足迹实验用于确定两者的结合位点。如 He 等(2023)通过 EMSA 和 DNase I 足迹实验表明耻垢分枝杆菌(*Mycobacterium smegmatis*) 中转录因子 GlnR 直接与启动子 P_{glnR} 上的 GTAAC 和 ATGTC 基序结合,从而揭示了 *glnR* 在氮胁迫下的转录调控机制。

值得注意的是,由于空间位阻,通过这种方式找出的 TFBS 序列一般比自身位点大 8~10 个碱基。除此之外,为保证 DNase I 足迹实验的准确性,研究人员需要严格控制 DNase I 的用量和反应时间。

5.2 染色体免疫共沉淀

原核生物的基因表达过程需要转录因子与启动子相结合,有研究者据此设计了染色质免疫共沉淀(chromatin immunoprecipitation assay, ChIP) 以分析蛋白质与 DNA 的结合位点(Gilmour and Lis, 1984; Solomon and Varshavsky, 1985)。其中,常用于启动子研究的方法包括 ChIP-chip 和 ChIP-seq。相比于 ChIP-chip, ChIP-seq 可以快速、有效地在全基因组范围内检测 DNA 与蛋白质的结合位点,有利于快速找出 TFBS。该技术在寻找启动子上的 TFBS 时依赖转录因子与启动子 DNA 的交联,通过细胞裂解和超声处理使启动子 DNA 片段化,然后与针对目的转录因子的抗体进行免疫沉淀,最后将 ChIP 富集的启动子 DNA 转化为用于下一代测序的文库,并将序列读数与参考基因组序列比对从而获得 TFBS(Wade, 2015)。如 Ramsey 等(2015)通过 ChIP-seq 确定转录因子 PigR 与 MglA-SspA 蛋白质复合物相互作用后可特异性与所调节启动子上的 TGCTGTA 结合,并控制 PigR 调节的毒力基因的表达。

虽然 ChIP-seq 的分辨率高,但仅使用 ChIP-seq 数据仍然无法以单核苷酸分辨率鉴定 DNA 结合蛋白的结合位点。因此,研究人员通过改良 ChIP-seq 开发了 ChIP-exo,可在免疫沉淀过程中对蛋白结合的 DNA 进行核酸外切酶处理,以实现单核苷酸分离(Yeh and Rhee, 2023),但目前 ChIP-exo 在原核生物中的应用较少。此外,Decker 等(2022)根据 ChIP-seq 和 ChIP-exo 数据开发了原核生物的染色质免疫

沉淀数据库 proChIPdb,有利于研究人员快速寻找与目标启动子相互作用的转录因子和 TFBS 相关信息。

5.3 逐段序列截短

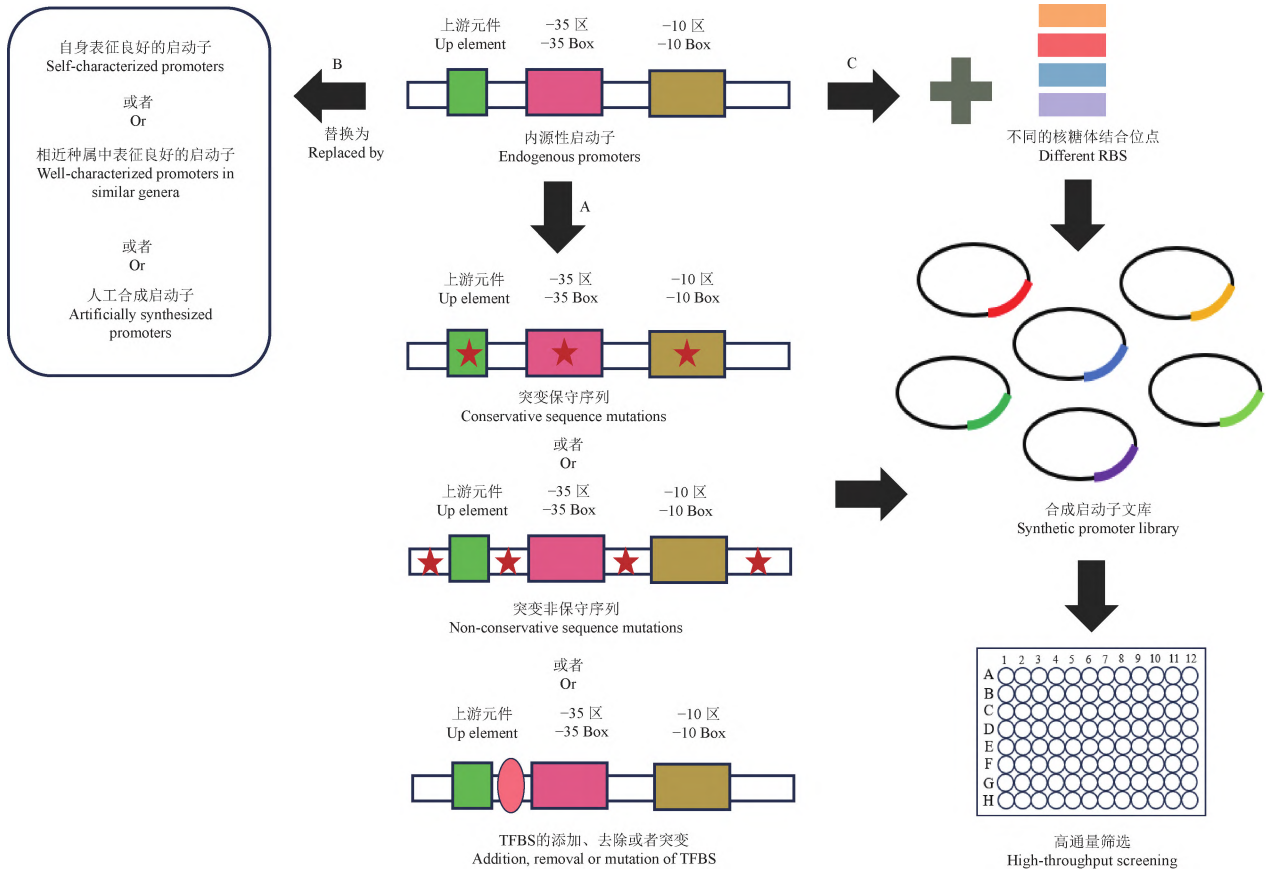
当某一启动子序列被确定后,可以进行启动子序列逐段截短,找到对启动子活性最关键的那段区域即启动子的核心序列。启动子序列逐段截短有 5'端逐段缺失、3'端逐段缺失和中间序列缺失等 3 种方式,将截短之后的片段连接到事先构建好的启动子探针载体上并导入宿主细胞,通过分析报告基因的表达情况以验证启动子的活性,从而确定目标启动子的核心区域。如 Sun 等(2021)构建了地衣芽孢杆菌(*Bacillus licheniformis*) 内源性启动子 P_{lth} 5'逐段截短且含报告基因 *egfp* 的载体,证明 P_{lth} 是由两个连续启动子区域组成的串联启动子,并用该启动子增强凝结芽孢杆菌(*Bacillus coagulans*) 中异源基因的表达量。

5.4 定点突变

定点突变是指通过 PCR、CRISPR/Cas9 等方法向目的片段中引入碱基替换、增添等变化,有利于确定启动子的 TATA 盒等参与转录调控的元件。这种方法需要通过生物信息学方法预测分析可能的转录调控区域,对这些区域进行定点突变,将获得的突变体克隆到无启动子的探针载体中,转化至宿主细胞后,分析细胞中报告基因的表达水平以确定突变前后启动子的活性,进而分析相应元件的位置。如杨巍等(2003)通过对乳酸乳球菌(*Lactococcus lactis*) 中启动子 P_{nisZ} 进行定点突变,确定了 P_{nisZ} 启动子的-35 区和-10 区。此外,也有研究者通过这种方式确定了嗜盐菌(*Haloferax mediterranei*) P_{lrp} 启动子的 TATA 盒及 TFBS,为揭示 Lrp 转录因子在胁迫条件下响应调控中的可能作用奠定了基础(Matarredona et al., 2023)。

6 原核生物启动子的改造

原核生物的天然启动子大多存在不同程度的缺陷。如表征良好的启动子数量有限;通用性差,在异源表达宿主中的活性低;灵活性差,不能满足基因表达的动态调控;大多数诱导型启动子需要添加诱导剂等。因此,对天然启动子进行改造具有重要意义,改造的方式有 3 种模式(图 4)。



A 表示内源性启动子的改造；B 表示启动子替换；C 表示启动子改造与核糖体工程结合；“★”表示突变位点。

A means modification of endogenous promoters; B means promoter replacement; C means combination of promoter modification and ribosome engineering; “★” means a mutation site.

图 4 启动子改造模式图

Figure 4 Diagram of the promoter modification model

6.1 对已知内源性启动子进行改造

对内源性启动子进行改造最常用的方法有：通过易错 PCR 或饱和突变的方法随机化启动子的保守区域或非保守区域构建合成启动子文库 (synthetic promoter library, SPL) (Li et al., 2023c; Yuan et al., 2023), 获得强度范围广的系列启动子, 可实现对不同基因进行不同程度的表达。然而, 从文库中筛选目的启动子费时费力, 常常需要结合高通量筛选技术或流式细胞仪技术。此外, 可以将启动子的核心序列改为保守的核苷酸序列, 如 σ 因子偏好的核苷酸序列或经典保守的核苷酸序列 (如 TATAAT 和 TT-GACA), 或者将启动子 UP 元件序列改为保守序列 (Zhou et al., 2019), 从而使启动子强度发生改变; 也可以将启动子的核心启动子序列替换为强启动子的核心序列以提高启动子的强度 (Li et al., 2023c); 还可以通过修饰启动子的 5'-UTR 构建文库, 如 Jia

等 (2023) 构建了乳酸链球菌 (*Pediococcus acidilactici*) DY15 启动子 5'-UTR 复合物文库, 并将获得的启动子-5'-UTR 复合物用于优化其转氨酶基因的表达水平和提高苯乳酸的产量。

对内源性启动子的改造还包括对 TFBS 的修饰。TFBS 是转录因子在调节基因表达过程中与启动子结合的特定 DNA 区域。通过在启动子序列上添加 TFBS 或对已包含的 TFBS 进行突变或去除, 以达到调节基因表达的目的。如在枯草芽孢杆菌中, Wu 等 (2020) 通过在启动子 P_{veg} 的 -10 区和 -35 区之间插入 GamR 结合位点 *gam02*, 得到 1 个能动态调控且强度和 P_{veg} 启动子一样的新启动子, 该启动子还可以用于调节细胞内 6-磷酸氨基葡萄糖的合成。而在链霉菌中, Wang 等 (2013) 通过去除启动子 P_{kasO} 核心区上游转录因子 ScbR 和 ScbR2 的共同结合位点获得启动子 P_{kasO3} , 随后随机突变去除 P_{kasO3} 核心启动

子区域内部的另一 *ScbR* 结合位点, 获得 1 个新的强启动子 P_{kasO^*} 。此外, 在大肠杆菌中, Zou 等(2021)通过突变 P_{s14} 启动子上转录因子 *MarR* 与 P_{marO} 的结合位点, 获得具有更高的响应强度和更宽的动态范围的启动子突变体 $P_{s14(12AII4T)}$ 和 $P_{s14(112AII4T)}$ 。

6.2 启动子替换

启动子替换是将自身或相近种属中表征良好的启动子或者人工合成的优选启动子替换原有的启动子。如 Pan 等(2022)通过对粘质沙雷氏菌(*Serratia marcescens*)进行 RNA-seq, 用获得的强组成型启动子 P_{rplJ} (P_{17})替换掉激活剂 *OmpR* 和 *PsrA* 相应基因原有的启动子, 可提高灵菌红素的产量。在枯草芽孢杆菌中, 通过 RNA-seq 得到 4 个启动子 P_{sodA} 、 P_{hag} 、 P_{tufA} 和 P_{fusA} , 这些启动子强度高于 P_{amyE} 启动子。随后, 用 $P_{sodA}+P_{mayE}+P_{fusA}$ 串联启动子替换 P_{amyE} 启动子, 得到的菌株其产出的普鲁兰酶的活性是原有菌株的 4.7 倍(Meng et al., 2018)。相对于其他启动子改造策略, 启动子替换策略针对性更强, 不必费时费力进行筛选, 但需要有大量可供选择的启动子。

6.3 启动子改造与 RBS 工程结合

启动子和 RBS 分别在转录和翻译的起始过程中发挥关键作用, 都是基因表达过程中的重要步骤。RBS 工程是一种有效调节基因表达水平的方式, 因此将原核生物启动子改造与 RBS 工程结合, 有利于实现同时在转录和翻译水平控制基因表达。如 Duan 等(2021)在谷氨酸棒状杆菌中偶联 6 个强度不同的启动子以及 11 个强度不同的 RBS 来调控报告基因 *egfp* 的表达, 结果表明 *egfp* 的表达水平至少在 10 倍的跨度内调节。此外, Sun 等(2023)联合改造钠弧菌(*Vibrio natriegens*)的启动子和 RBS, 构建突变文库, 为多种重组蛋白的表达提供了选择, 同时拓展了钠弧菌的应用范围。说明与 RBS 工程结合的启动子改造比单一启动子改造的调控范围更广, 为寻求平衡生物合成途径及提高目标代谢产量的方法提供了有意义的参考。

7 原核生物启动子的开发应用

随着人工合成、随机突变和启动子替换等启动子改造手段的运用, 研究者获得了许多符合目标需求的启动子, 并广泛应用于各种微生物中, 以实现

高效产酶、开发抑菌剂以及优化代谢途径等目标。

7.1 有效提高目标产物的产量

原核生物大肠杆菌、谷氨酸棒状杆菌和枯草芽孢杆菌等常作为异源基因的表达宿主, 启动子在基因表达中起着关键作用, 开发具有不同强度和功能的启动子, 可有效提高目标产物的产量, 实现目标产物的高效表达。

如 Rao 等(2023)通过组合不同 σ 依赖型启动子核心序列获得新型启动子, 并对这些启动子进行表征和分析, 使用表征效果好的启动子提高了蛋白质和重要代谢产物的产量, 为地衣芽孢杆菌的基因调控和多样化产品生产提供了有力支持。Cao 等(2023)对大肠杆菌中的启动子 P_{luxI} 进行改造, 以 P_{luxI} 突变体 $P_{luxI(T-38C)}$ 为启动子的自诱导表达系统在高密度发酵条件下表达的藻酸盐裂解酶活性是 P_{T7} 启动子的 4.34 倍。此外, 工程化的 P_{tac} 启动子也可以减少抗体 Fab' 片段在大肠杆菌中的表达泄露(Schofield et al., 2016)。

7.2 控制代谢通路, 优化代谢网络

启动子可以调节特定代谢途径中的基因表达水平, 从而控制代谢通量。通过启动子改造获得的不同强度启动子可优化代谢网络, 进而显著提高工业应用所需微生物细胞工厂的能力。

在工程菌中重构生物合成途径时, 不是每一个基因都需要在强启动子的控制下表达, 而是需要精确地控制途径中每一种酶基因的表达, 尤其是一些关键酶, 可避免不必要的中间代谢物积累。Hwang 等(2018)通过工程化溶氧依赖型启动子 P_{nar} 获得强度范围广的合成 P_{nar} 启动子文库, 随后, 运用这些启动子来控制 D-乳酸和 2,3-丁二醇对应基因的表达水平。选择活性强的合成 P_{nar} 启动子控制编码 D-乳酸脱氢酶的 *ldhD* 基因的表达, 较大幅度提高了 D-乳酸的产量。在 2,3-丁二醇合成途径中, *aldB* 基因的表达强度与 2,3-丁二醇的产量无关, 而 *ilvBN* 和 *bdhI* 基因的表达水平增强有利于提高 2,3-丁二醇的产量, 因此, 用弱的合成 P_{nar} 启动子控制 *aldB* 基因的表达水平, 同时用活性强的合成 P_{nar} 启动子分别控制 *ilvBN* 和 *bdhI* 基因的表达水平, 最终使 2,3-丁二醇的产量增加 72%。另外, Zhang 等(2023a)通过 CRISPR/Cas9 介导启动子改造, 并用其中 3 种强度不同的启动子来微调红霉素生物合成途径中关键基

因的表达量，突破了红霉素生物合成的瓶颈，也为提高其他重要放线菌天然产物的产量提供了有价值的参考。

8 结论与展望

原核生物启动子的研究有利于开发和表征更多新启动子，揭示细菌的生命活动机制，同时为合成生物学、工业微生物的改良以及环境治理等前沿领域提供了线索 (Huang et al., 2021a; Shemyakina et al., 2021; Zhang et al., 2021)。在合成生物学新技术、高通量生物检测技术、单细胞测序技术乃至生物信息学等交叉研究领域协同发展的背景下，原核生物启动子研究已经取得了令人瞩目的进展和突破。本文介绍了原核生物启动子的结构和种类，重点介绍了原核生物启动子的筛选、预测和功能验证方法以及其开发和应用情况，旨在为相关研究提供参考。当然，原核生物启动子的研究目前仍面临许多挑战，后续的研究需要优先解决的问题及其方案如下：

(1) 现有原核生物启动子预测工具大多基于典型的模式生物，而不同生物之间的启动子结构和序列具有多样性，导致非模式生物的启动子有时难以准确预测。在现有启动子预测模型及已知启动子结构和序列的基础上，利用大数据技术进一步优化算法，构建既能预测模式生物启动子又能预测非模式生物启动子的预测模型，预期可提高启动子筛选的系统性和高通量性。

(2) 当前，启动子功能验证主要依赖于报告基因检测系统，验证的准确性和灵敏度依赖于报告基因表达产物的检测分析，如前文中所述的氯霉素乙酰转移酶 (Kozakai et al., 2021)、 β -半乳糖苷酶 (Brandwein et al., 2023)、 β -葡萄糖醛酸酶 (Chae et al., 2019)、儿茶酚 2,3-双加氧酶 (Luo et al., 2015)、绿色荧光蛋白 (Lee et al., 2023)、红色荧光蛋白 (Gao et al., 2022) 和荧光素酶 (Yu et al., 2023) 等，在弱启动子的功能验证过程中容易形成假阴性。挖掘更为有效的报告基因，或者构建对现有报告基因产物更为灵敏、准确的分析检测方法，可以有效提高启动子特别是弱启动子功能验证的准确性。

(3) 现有表征良好的原核生物启动子数量和多样性都非常有限。在上述启动子筛选和功能验证方法取得突破后，可在更广的生物范围内挖掘各具特

色的启动子，特别是生长依赖型启动子和自诱导型启动子，以适应不断增长的生物学基础研究和工业生产实践的需求。

总之，未来的原核生物启动子研究将会更加注重启动子筛选的系统性、高通量性和精准性，以及多学科交叉技术的开发应用，高通量测序技术、计算算法和人工智能等领域的发展有望为启动子的预测和开发带来新突破 (Li et al., 2023b; Zhao et al., 2022)。另外，筛选启动子以外的其他转录调控元件以及其对基因的转录调控规律的研究也必将成为今后原核生物启动子研究领域的重要内容。

作者贡献

张小梅负责论文初稿的撰写和修改；彭萱和龙雨欣负责绘图、文稿修改和参考文献的整理；龙中儿是论文的构思者及负责人，指导论文写作和修改，并负责论文定稿。全体作者都已阅读并同意最终的文本。

参考文献

- 杨巍, 陈秀珠, 袁静, 等, 2003. *nisZ* 启动子结构与功能的研究. 微生物学报, 43(2): 194-199. [YANG W, CHEN X Z, YUAN J, et al., 2003. Studies on structure and function of *nisZ* promoter. Acta Microbiologica Sinica, 43(2): 194-199.]
- ABDULMALEK H W, YAZGAN-KARATAŞ A, 2023. Improvement of bacilysin production in *Bacillus subtilis* by CRISPR/Cas9-mediated editing of the 5'-untranslated region of the *bac* operon. J. Microbiol. Biotechnol., 33(3): 410-418.
- AMIN R, RAHMAN C R, AHMED S, et al., 2020. iPromoter-BnCNN: a novel branched CNN-based predictor for identifying and classifying sigma promoters. Bioinformatics, 36(19): 4869-4875.
- BRANDWEIN J N, SCULTHORPE T S, RIDDER M J, et al., 2023. Factors impacting the regulation of *nos* gene expression in *Staphylococcus aureus*. Microbiol. Spectr., 11(5): e0168823.
- BUSCHE T, DOSTÁLOVÁ H, RUCKÁ L, et al., 2023. Overlapping SigH and SigE sigma factor regulons in *Corynebacterium glutamicum*. Front. Microbiol., 13: 1059649.
- CAO Z N, LIU Z, ZHANG G L, et al., 2023. *P_{luxI}* mutants with different promoting period and their application for quorum sensing regulated protein expression. Food Sci. Hum. Wellness, 12(5): 1841-1849.

- CHAE J P, PAJARILLO E A, HWANG I C, et al., 2019. Construction of a bile-responsive expression system in *Lactobacillus plantarum*. *Food Sci. Anim. Resour.*, 39(1): 13-22.
- CHEN M M, REN G J, ZHANG X, et al., 2023. DegU-mediated suppression of carbohydrate uptake in *Listeria monocytogenes* increases adaptation to oxidative stress. *Appl. Environ. Microbiol.*, 89(10): e0101723.
- CHEN Y L, GUO D H, LI Q Z, 2020. An energy model for recognizing the prokaryotic promoters based on molecular structure. *Genomics*, 112(2): 2072-2079.
- COPPENS L, LAVIGNE R, 2020. SAPPHERE: a neural network based classifier for $\sigma 70$ promoter prediction in *Pseudomonas*. *BMC Bioinformatics*, 21(1): 415.
- DECKER K T, GAO Y, RYCHEL K, et al., 2022. proChIPdb: a chromatin immunoprecipitation database for prokaryotic organisms. *Nucleic Acids Res.*, 50(D1): D1077-D1084.
- DENG C, LI J H, SHIN H D, et al., 2018. Efficient expression of cyclodextrin glycosyltransferase from *Geobacillus stearothermophilus* in *Escherichia coli* by promoter engineering and downstream box evolution. *J. Biotechnol.*, 266: 77-83.
- DUAN Y T, ZHAI W J, LIU W J, et al., 2021. Fine-tuning multi-gene clusters via well-characterized gene expression regulatory elements: case study of the arginine synthesis pathway in *C. glutamicum*. *ACS Synth. Biol.*, 10(1): 38-48.
- FEKLÍSTOV A, SHARON B D, DARST S A, et al., 2014. Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.*, 68: 357-376.
- GAO L, WU X D, LI C Y, et al., 2022. Exploitation of strong constitutive and stress-driven promoters from *Acetobacter pasteurianus* for improving acetic acid tolerance. *J. Biotechnol.*, 350: 24-30.
- GILMOUR D S, LIS J T, 1984. Detecting protein-DNA interactions *in vivo*: distribution of RNA polymerase on specific bacterial genes. *Proc. Natl. Acad. Sci. USA*, 81(14): 4275-4279.
- GUO W L, XIAO Z H, HUANG T T, et al., 2023. Identification and characterization of a strong constitutive promoter *strYp* for activating biosynthetic genes and producing natural products in *Streptomyces*. *Microb. Cell Fact.*, 22(1): 127.
- HE J M, KANG X M, WU J C, et al., 2023. Transcriptional self-regulation of the master nitrogen regulator GlnR in *Mycobacteria*. *J. Bacteriol.*, 205(4): e0047922.
- HE W Y, JIA C Z, DUAN Y C, et al., 2018. 70ProPred: a predictor for discovering sigma70 promoters based on combining multiple features. *BMC Syst. Biol.*, 12(Suppl 4): 44.
- HENRY K K, ROSS W, GOURSE R L, 2021. *Rhodobacter sphaeroides* CarD negatively regulates its own promoter. *J. Bacteriol.*, 203(17): e0021021.
- HOCQ R, JAGTAP S, BOUTARD M, et al., 2022. Genome-wide TSS distribution in three related Clostridia with normalized capp-switch sequencing. *Microbiol. Spectr.*, 10(2): e0228821.
- HOMEROVA D, BISCHOFF M, DUMOLIN A, et al., 2004. Optimization of a two-plasmid system for the identification of promoters recognized by RNA polymerase containing *Staphylococcus aureus* alternative sigma factor sigmaB. *FEMS Microbiol. Lett.*, 232(2): 173-179.
- HOMEROVA D, SURDOVA K, MIKUSOVA K, et al., 2007. Identification of promoters recognized by RNA polymerase containing *Mycobacterium tuberculosis* stress-response sigma factor sigma(F). *Arch. Microbiol.*, 187(3): 185-197.
- HUANG C C, SHAN L H, CHEN Z Y, et al., 2021a. Differential effects of homologous transcriptional regulators NicR2A, NicR2B1, and NicR2B2 and endogenous ectopic strong promoters on nicotine metabolism in *Pseudomonas* sp. strain JY-Q. *Appl. Environ. Microbiol.*, 87(3): e02457-20.
- HUANG J W, CHEN J Z, WANG Y, et al., 2021b. Development of a hyperosmotic stress inducible gene expression system by engineering the MtrA/MtrB-dependent *NCgl1418* promoter in *Corynebacterium glutamicum*. *Front. Microbiol.*, 12: 718511.
- HUERTA A M, COLLADO-VIDES J, 2003. Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J. Mol. Biol.*, 333(2): 261-278.
- HUO K Y, WANG S Q, ZHAO W W, et al., 2023. Creating an efficient 1, 2-dichloroethane-mineralizing bacterium by a combination of pathway engineering and promoter engineering. *Sci. Total Environ.*, 878: 163140.
- HWANG H J, LEE S Y, LEE P C, 2018. Engineering and application of synthetic *nar* promoter for fine-tuning the expression of metabolic pathway genes in *Escherichia coli*. *Biotechnol. Biofuels*, 11: 103.
- JACOB H, GENG H, SHETTY D, et al., 2022. Distinct interaction mechanism of RNA polymerase and ResD at proximal and distal subsites for transcription activation of nitrite reductase in *Bacillus subtilis*. *J. Bacteriol.*, 204(2): e0043221.
- JACQUES P E, RODRIGUE S, GAUDREAU L, et al., 2006. Detection of prokaryotic promoters from the genomic distribution of hexanucleotide pairs. *BMC Bioinformatics*, 7: 423.
- JAISHANKAR J, BHATO A L, PATIL N, et al., 2021. Microarray profiling and identification of core promoter sequence in *Gordonia*. *Genomics*, 113(6): 4327-4336.
- JIA Y Z, HUANG C, MAO Y, et al., 2023. Screening and constructing a library of promoter-5'-UTR complexes with gra-

- dient strength in *Pediococcus acidilactici*. ACS Synth. Biol., 12(6): 1794-1803.
- KANG X M, CAI X, HUANG Z H, et al., 2020. Construction of a highly active secretory expression system in *Bacillus subtilis* of a recombinant amidase by promoter and signal peptide engineering. Int. J. Biol. Macromol., 143: 833-841.
- KANHERE A, BANSAL M, 2005. A novel method for prokaryotic promoter prediction based on DNA stability. BMC Bioinformatics, 6: 1.
- KIM S, KIM S H, AHN J, et al., 2019. Crystal structure of the regulatory domain of MexT, a transcriptional activator of the MexEFOPrN efflux pump in *Pseudomonas aeruginosa*. Mol. Cells, 42(12): 850-857.
- KONG L H, XIONG Z Q, SONG X, et al., 2019. Characterization of a panel of strong constitutive promoters from *Streptococcus thermophilus* for fine-tuning gene expression. ACS Synth. Biol., 8(6): 1469-1472.
- KOZAKAI T, SHIMOFUSA Y, NOMURA I, et al., 2021. Construction of a reporter system for *Bifidobacteria* using chloramphenicol acetyltransferase and its application for evaluation of promoters and terminators. Biosci. Microbiota Food Health, 40(2): 115-122.
- LAMMENS E-M, PUTZEYS L, BOON M, et al., 2023. Sourcing phage-encoded terminators using ONT-cappable-seq for SynBio applications in *Pseudomonas*. ACS Synth. Biol., 12(5): 1415-1423.
- LEE M J, KIM B, KIM K, et al., 2023. Heterologous protein production using *Psychrobacter* sp. PAMC 21119 analyzed with a green fluorescent protein-based reporter system. Protein Expr. Purif., 212: 106352.
- LI J Y, TONG S, AMIN F R, et al., 2023a. Enhancing the activity of a self-inducible promoter in *Escherichia coli* through saturation mutation and high-throughput screening. Fermentation, 9(5): 468.
- LI H F, ZHANG J Y, ZHAO Y M, et al., 2023b. Predicting *Corynebacterium glutamicum* promoters based on novel feature descriptor and feature selection technique. Front. Microbiol., 14: 1141227.
- LI H R, JIA M J, QI Q S, et al., 2023c. Engineered probiotic *Lactobacillus plantarum* WCSF I for monitoring and treatment of *Staphylococcus aureus* infection. Microbiol. Spectr., 11(6): e0182923.
- LI X L, ZHANG Y L, ZHAN Y D, et al., 2023d. Utilization of a strong promoter combined with the knockout of protease genes to improve the yield of Vip3Aa in *Bacillus thuringiensis* BMB171. Pest Manag. Sci., 79(5): 1713-1720.
- LI Y M, ZHANG R, WANG S L, 2022. A natural novel mutation in the *bla_{NDM-5}* promoter reducing carbapenems resistance in a clinical *Escherichia coli* strain. Microbiol. Spectr., 10(1): e0118321.
- LIN Y, SUN M L, ZHANG J J, et al., 2023. Computational identification of promoters in *Klebsiella aerogenes* by using support vector machine. Front. Microbiol., 14: 1200678.
- LIU M F, TOLSTORUKOV M, ZHURKIN V, et al., 2004. A mutant spacer sequence between -35 and -10 elements makes the *P_{lac}* promoter hyperactive and cAMP receptor protein-independent. Proc. Natl. Acad. Sci. USA, 101(18): 6911-6916.
- LIU X, GUO Z R, HE T, et al., 2020a. Prediction and analysis of prokaryotic promoters based on sequence features. Biosystems, 197: 104218.
- LIU X L, YU J, CHE Q, et al., 2023. Enhanced phycocyanobilin production in *Escherichia coli* by fusion-expression of apo-proteins with signal peptides. Fermentation, 9(9): 851.
- LIU X X, ZHAO Z H, DONG G B, et al., 2020b. Identification, repair and characterization of a benzyl alcohol-inducible promoter for recombinant proteins overexpression in *Corynebacterium glutamicum*. Enzyme Microb. Technol., 141: 109651.
- LU Q, CHEN T Y, WANG J N, et al., 2023. Structural insight into the mechanism of σ^{32} -mediated transcription initiation of bacterial RNA polymerase. Biomolecules, 13(5): 738.
- LUO Y Z, ZHANG L, BARTON K W, et al., 2015. Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. ACS Synth. Biol., 4(9): 1001-1010.
- MA J F, JI Q Y, WANG S Y, et al., 2021. Identification and evaluation of a panel of strong constitutive promoters in *Listeria monocytogenes* for improving the expression of foreign antigens. Appl. Microbiol. Biotechnol., 105(12): 5135-5145.
- MAKITA Y, NAKAO M, OGASAWARA N, et al., 2004. DBT-BS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics. Nucleic Acids Res., 32(Database issue): D75-D77.
- MATARREDONA L, CAMACHO M, BAUTISTA V, et al., 2023. Lrp as a potential transcriptional regulator involved in stress response in *Haloferax mediterranei*. Biochimie, 209: 61-72.
- MCCRACKEN A, TURNER M S, GIFFARD P, et al., 2000. Analysis of promoter sequences from *Lactobacillus* and *Lactococcus* and their activity in several *Lactobacillus* species. Arch. Microbiol., 173(5): 383-389.
- MENG F Q, ZHU X Y, NIE T, et al., 2018. Enhanced expression of pullulanase in *Bacillus subtilis* by new strong promoters mined from transcriptome data, both alone and in combination. Front. Microbiol., 9: 2635.

- MISHRA A, DHANDA S, SIWACH P, et al., 2020. A novel method SEProm for prokaryotic promoter prediction based on DNA structure and energetics. *Bioinformatics*, 36 (8): 2375-2384.
- MÜNCH R, HILLER K, BARG H, et al., 2003. PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res.*, 31(1): 266-269.
- NAKANO M M, LIN A, ZUBER C S, et al., 2010. Promoter recognition by a complex of Spx and the C-terminal domain of the RNA polymerase alpha subunit. *PLoS ONE*, 5 (1): e8664.
- OUYANG Q, WANG X, ZHANG N, et al., 2020. Promoter screening facilitates heterologous production of complex secondary metabolites in *Burkholderiales* strains. *ACS Synth. Biol.*, 9(2): 457-460.
- PAGET M S B, HELMANN J D, 2003. The sigma70 family of sigma factors. *Genome Biol.*, 4(1): 203.
- PAN X W, YOU J J, TANG M, et al., 2022. Improving prodigiosin production by transcription factor engineering and promoter engineering in *Serratia marcescens*. *Front. Microbiol.*, 13: 977337.
- PÁTEK M, EIKMANN B J, PÁTEK J, et al., 1996. Promoters from *Corynebacterium glutamicum*: cloning, molecular analysis and search for a consensus motif. *Microbiology*, 142 (5): 1297-1309.
- RAMSEY K M, OSBORNE M L, VVEDENSKAYA I O, et al., 2015. Ubiquitous promoter-localization of essential virulence regulators in *Francisella tularensis*. *PLoS Pathog.*, 11 (4): e1004793.
- RAO Y, WANG J Q, YANG X Y, et al., 2023. A novel toolbox for precise regulation of gene expression and metabolic engineering in *Bacillus licheniformis*. *Metab. Eng.*, 78: 159-170.
- SANTOS-ZAVALETA A, SALGADO H, GAMA-CASTRO S, et al., 2019. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.*, 47(D1): D212-D220.
- SCHOFIELD D M, TEMPLAR A, NEWTON J, et al., 2016. Promoter engineering to optimize recombinant periplasmic Fab' fragment production in *Escherichia coli*. *Biotechnol. Prog.*, 32 (4): 840-847.
- SEVCIKOVA B, REZUCHOVA B, MAZURAKOVA V, et al., 2021. Cross-recognition of promoters by the nine SigB homologues present in *Streptomyces coelicolor* A3(2). *Int. J. Mol. Sci.*, 22(15): 7849.
- SHEMYAKINA A O, GRECHISHNIKOVA E G, NOVIKOV A D, et al., 2021. A set of active promoters with different activity profiles for superexpressing *Rhodococcus* strain. *ACS Synth. Biol.*, 10(3): 515-530.
- SHEN L F, ZHANG J X, XUE J, et al., 2022. Regulation of ECP fimbriae-related genes by the transcriptional regulator RcsAB in *Klebsiella pneumoniae* NTUH-K2044. *J. Basic Microbiol.*, 62(5): 593-603.
- SHI T, ZHANG L, LIANG M D, et al., 2021. Screening and engineering of high-activity promoter elements through transcriptomics and red fluorescent protein visualization in *Rhodobacter sphaeroides*. *Synth. Syst. Biotechnol.*, 6 (4): 335-342.
- SOLOMON M J, VARSHAVSKY A, 1985. Formaldehyde-mediated DNA-protein crosslinking: a probe for *in vivo* chromatin structures. *Proc. Natl. Acad. Sci. USA*, 82 (19): 6470-6474.
- SPANGLER J R, COOPER D N, MALANOSKI A P, et al., 2022. Promoter identification and optimization for the response of *Lactobacillus plantarum* WCFS1 to the gram-negative pathogen-associated molecule N-3-oxododecanoyl homoserine lactone. *ACS Biomater. Sci. Eng.*, 9(9): 5111-5122.
- STEITZ T A, 2009. The structural changes of T7 RNA polymerase from transcription initiation to elongation. *Curr. Opin. Struct. Biol.*, 19(6): 683-690.
- SU C, GONG J S, WU Z X, et al., 2023. Development of a growth-dependent system to regulate cell growth and keratinase production in *B. subtilis*. *J. Agric. Food Chem.*, 71 (5): 2421-2429.
- SUN W H, JIANG B, ZHAO D Y, et al., 2021. Integration of metabolic pathway manipulation and promoter engineering for the fine-tuned biosynthesis of malic acid in *Bacillus coagulans*. *Biotechnol. Bioeng.*, 118(7): 2597-2608.
- SUN Y J, XU J Q, ZHOU H S, et al., 2023. Recombinant protein expression chassis library of *Vibrio natriegens* by fine-tuning the expression of T7 RNA polymerase. *ACS Synth. Biol.*, 12(2): 555-564.
- TIMMS-WILSON T M, ELLIS R J, BAILEY M J, 2000. Immunocapture differential display method (IDDM) for the detection of environmentally induced promoters in rhizobacteria. *J. Microbiol. Methods*, 41(1): 77-84.
- TOH W K, TEO Y L, TOR X Y, et al., 2023. Development of constitutive and IPTG-inducible integron promoter-based expression systems for *Escherichia coli* and *Agrobacterium tumefaciens*. *3 Biotech*, 13(3): 91.
- WADE J T, 2015. Mapping transcription regulatory networks with ChIP-seq and RNA-seq. *Adv. Exp. Med. Biol.*, 883: 119-134.
- WANG K, CHEN X A, LI Y Q, et al., 2019. Identification of a secondary metabolism-responsive promoter by proteomics for over-production of natamycin in *Streptomyces*. *Arch. Microbiol.*

- ol., 201(10): 1459-1464.
- WANG W S, LI X, WANG J, et al., 2013. An engineered strong promoter for *Streptomyces*. *Appl. Environ. Microbiol.*, 79(14): 4484-4492.
- WANG Y Y, LU J F, XU M C, et al., 2022. Efficient biosynthesis of isopropanol from ethanol by metabolically engineered *Escherichia coli*. *ACS Sustain. Chem. Eng.*, 10(41): 13857-13864.
- WU L J, TANG L Q, HE Y C, et al., 2023. BysR, a LysR-type pleiotropic regulator, controls production of occidiofungin by activating the LuxR-type transcriptional regulator AmbR1 in *Burkholderia* sp. strain JP2-270. *Microbiol. Spectr.*, 11(2): e0268422.
- WU Y K, CHEN T C, LIU Y F, et al., 2020. Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in *Bacillus subtilis*. *Nucleic Acids Res.*, 48(2): 996-1009.
- YANG S, DU G C, CHEN J, et al., 2017. Characterization and application of endogenous phase-dependent promoters in *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.*, 101(10): 4151-4161.
- YEH S Y, RHEE H S, 2023. The ChIP-exo method to identify genomic locations of DNA-binding proteins at near single base-pair resolution. *Methods Mol. Biol.*, 2599: 33-48.
- YU G L, LI X C, DUAN Q Y, et al., 2023. Systematic identification of endogenous strong constitutive promoters from the diazotrophic rhizosphere bacterium *Pseudomonas stutzeri* DSM4166 to improve its nitrogenase activity. *Microb. Cell Fact.*, 22(1): 91.
- YUAN Y, KONG D M, WU J, et al., 2023. Expression element optimization and molecular modification to enhance the secretory expression of chitinase from *Bacillus licheniformis* in *Bacillus subtilis*. *Process Biochem.*, 131: 32-40.
- ZHANG X M, WANG Y, ZHANG Y, et al., 2023a. CRISPR/Cas9-mediated multi-locus promoter engineering in ery cluster to improve erythromycin production in *Saccharopolyspora erythraea*. *Microorganisms*, 11(3): 623.
- ZHANG Y P, ZHANG J H, XIAO J M, et al., 2023b. comCDE (competence) operon is regulated by CcpA in *Streptococcus pneumoniae* D39. *Microbiol. Spectr.*, 11(3): e0001223.
- ZHANG Y T, LIU H L, LIU Y J, et al., 2021. A promoter engineering-based strategy enhances polyhydroxyalkanoate production in *Pseudomonas putida* KT2440. *Int. J. Biol. Macromol.*, 191: 608-617.
- ZHANG Z M, ZHAO J P, WEI P J, et al., 2022. iPromoter-CLA: identifying promoters and their strength by deep capsule networks with bidirectional long short-term memory. *Comput. Methods Programs Biomed.*, 226: 107087.
- ZHAO M, TAO X Y, WANG F Q, et al., 2018. Establishment of a low-dosage-IPTG inducible expression system construction method in *Escherichia coli*. *J. Basic Microbiol.*, 58(9): 806-810.
- ZHAO M, YUAN Z Q, WU L T, et al., 2022. Precise prediction of promoter strength based on a *de novo* synthetic promoter library coupled with machine learning. *ACS Synth. Biol.*, 11(1): 92-102.
- ZHOU C Y, YE B, CHENG S, et al., 2019. Promoter engineering enables overproduction of foreign proteins from a single copy expression cassette in *Bacillus subtilis*. *Microb. Cell Fact.*, 18(1): 111.
- ZHU Y F, DU S S, YAN Y F, et al., 2022. Systematic engineering of *Bacillus amyloliquefaciens* for efficient production of poly- γ -glutamic acid from crude glycerol. *Bioresour. Technol.*, 359: 127382.
- ZOU Y S, LI C Y, ZHANG R H, et al., 2021. Exploring the tunability and dynamic properties of MarR-PmarO sensor system in *Escherichia coli*. *ACS Synth. Biol.*, 10(8): 2076-2086.

(责任副主编 李伟辉)
(责任执行主编 陈玲玲)