

· 特邀综述 ·

高等植物启动子元件鉴定与合成研究进展

宋想[†], 王璐瑶[†], 富博晓, 李双达, 魏媛媛, 洪艳*, 戴思兰*

北京林业大学园林学院, 林木花卉遗传育种教育部重点实验室, 城乡生态环境北京实验室, 花卉种质创新与分子育种
北京市重点实验室, 国家花卉工程技术研究中心, 北京 100083

摘要 启动子是高等植物中驱动基因表达不可或缺的重要调控序列。启动子元件的不同功能是导致基因表达效率和时空特异性差异的根本原因, 鉴定其结构和功能对理解植物生长发育进程、环境胁迫反应和进化发育均具有重要意义。随着高通量测序技术、人工智能和合成生物学的发展, 鉴定植物顺式作用元件组, 将符合设计需求的元件构建成人工调控元件的技术已逐渐兴起, 这些技术为分子育种中高效、精准和多样的基因调控奠定了基础。该文以启动子元件在分子设计中的重构应用为导向, 介绍了高等植物启动子的详细结构和功能, 及顺式作用元件的鉴定方法, 系统梳理了27类共计174个已鉴定到的诱导型和组织特异性元件及其应用。据此, 提出其设计方向和方法, 旨在为高等植物基因工程启动子元件的研究现状、发掘方法和重新设计提供指导, 提升分子设计育种中目的基因的表达效率和准确性。

关键词 启动子, 顺式作用元件, 诱导型, 组织特异性, 人工合成启动子

宋想, 王璐瑶, 富博晓, 李双达, 魏媛媛, 洪艳, 戴思兰 (2024). 高等植物启动子元件鉴定与合成研究进展. 植物学报 59, 691–708.

顺式作用元件(*cis-acting elements*)亦称顺式调控元件(*cis-regulatory elements*, CREs)是指存在于基因旁侧序列中, 能影响基因表达的DNA序列, 其中启动子(*promoter*)是基因能够准确表达最主要的DNA调控序列。研究者最初对植物休眠等现象的研究表明, 并非所有基因都是一成不变地表达, 这种具有时空特异性的生理过程由启动子控制(Hemberg, 1949; Amen, 1968)。启动子通过结合相应的转录因子(*transcription factors*, TFs)来控制基因转录水平的表达(Porto et al., 2014)。根据启动子上元件的不同, 其与转录因子的互作有不同的时空特异性, 可划分为组成型启动子(*constitutive promoter*)、组织特异性启动子(*tissue-specific promoter*)和诱导型启动子(*inducible promoter*)。目前, 研究者基于对启动子功能的理解, 实现了作物观赏性状(Noda et al., 2017)、经济性状(Wu et al., 2023)和抗逆性(Beliaev et al., 2023)等农艺性状的精准改良。相较于原核生物, 真核生物物种间启动子元件结构和数量差异极大, 这也是许多植物

在短期内形成遗传多样性的基础之一(Brázda et al., 2021)。相同植物中, 不同的天然启动子驱动基因表达的差异可能会高达数十倍(Annadana et al., 2002); 即使其结构不变, 也可通过甲基化等方式调节转录活性(Lucibelli et al., 2022), 影响花色(Han et al., 2020)、花期(Yang et al., 2022)、花型(Zhang et al., 2022)和花香(Yuan et al., 2021)等性状。因此, 对启动子结构和功能进行研究可丰富我们对物种进化关系、基因组保守性和植物环境适应机理的理解。

目前, 组成型启动子因活性稳定、研究充分及各组织驱动表达无差异等特点得到广泛应用(邓泽宜等, 2021), 但在分子育种技术发展过程中, 对目的基因表达产生了更多样的需求。例如, 需要基因在特定时期、部位或诱导条件下高强度表达, 因此亟须研发植物特异性启动子。另外, 天然启动子在转录效率、强度和诱导条件等方面存在局限性, 包括不必要的负反馈、代谢负担和能量损失等(Rushton, 2016)。为解决上述问题, 可对已知的具有不同功能的顺式作用元件

收稿日期: 2023-12-22; 接受日期: 2024-03-18

基金项目: 北京市自然科学基金(No.6222043)、国家自然科学基金(No.32271946)和北京林业大学“5·5工程”科研创新团队项目(No.BLRC2023A06)

[†] 共同第一作者

* 通讯作者。E-mail: hongy@bjfu.edu.cn; silandai@sina.com

进行突变或重组,使目的基因具有复合且高效的时空表达模式。但目前对启动子的研究大多集中于天然启动子的功能鉴定,缺乏顺式作用元件的研究数据。为此,本文综述了高等植物天然启动子的结构特征及其作用方式,目前鉴定获得的高等植物启动子顺式作用元件及其在基因工程中的应用,启动子元件的重组方法和最新应用,以及合成启动子在基因工程中的巨大应用潜力,以期对高等植物启动子的分析、设计和转基因应用提供思路。

1 高等植物启动子的结构及功能

植物启动子主要由核心启动子(core promoter)、近端

启动子元件(promoter proximal element)和远端启动子区(distal promoter region)组成,这些区域的定义取决于其与转录起始位点(transcription start site, TSS)的距离和自身功能(图1)。其中,核心启动子在转录起始位点前后约50 bp (± 50 bp),结构和反式作用因子均相对固定,是具有驱动基因表达功能的最短片段(Butler and Kadonaga, 2002);近端启动子元件跨越数百个碱基对,包括普遍存在的上游元件(CAAT-box和GC-box等)和较为特异的应答元件(promoter response elements, PREs) (响应光、干旱和低温等环境信号及生长发育信号)两种,决定基因转录的活性、特异性和效率,确保转录精确且有效地起始(张春

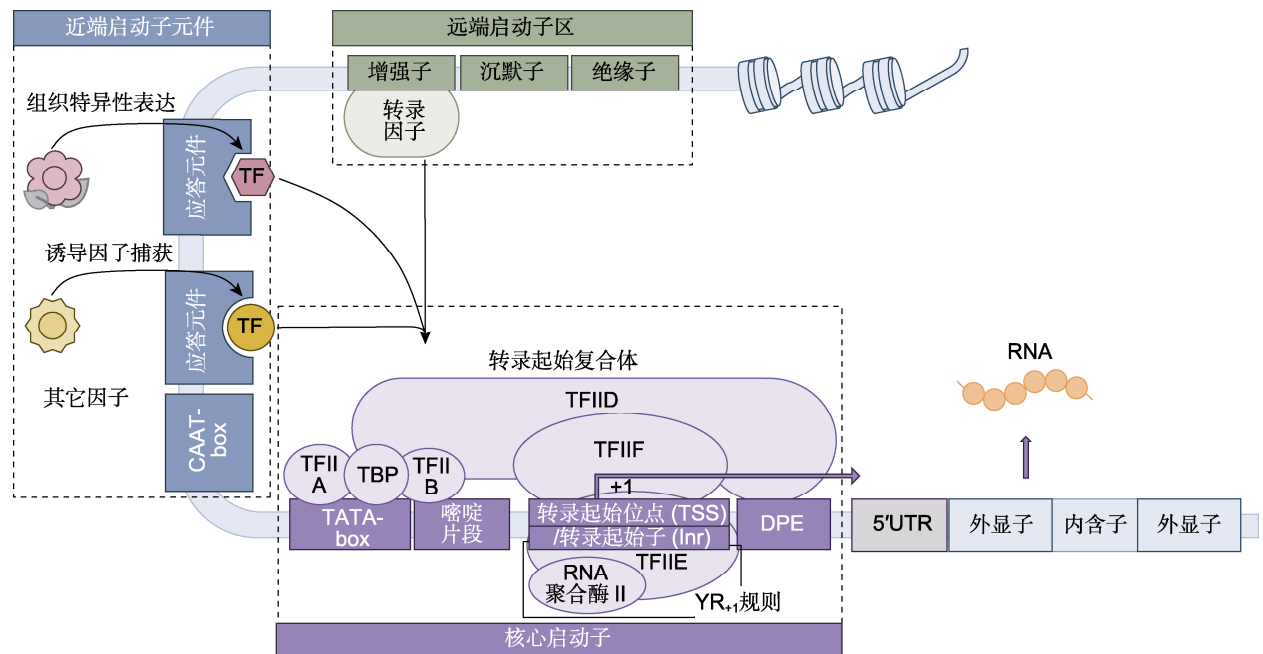


图1 植物启动子的结构及其作用方式(Chen and Xu, 2022; Yasmeen et al., 2023; Brooks et al., 2023)

植物的启动子区域与RNA聚合酶II是转录调控的中心。核心启动子区,一般转录因子TFIID的TBP区域识别TATA-box,与TFIIF、TFIIE和RNA聚合酶II等蛋白组装为转录起始复合体,识别并结合TATA-box或其下游核心启动子元件(DPE),协同转录起始位点(TSS)等核心元件,调控转录起始。嘧啶片段(Y patch)等元件有增强或稳定启动子活性的作用。CAAT-box等一般上游调控元件和应答元件与受调控的转录因子结合,参与基因表达调控;在基因的上游和下游还包含增强子等顺式作用元件,通过特定转录因子的协同作用调控基因的表达。

Figure 1 The structure and mode of actions of plant promoters (Chen and Xu, 2022; Yasmeen et al., 2023; Brooks et al., 2023) The promoter region of plants and RNA polymerase II are at the center of transcriptional regulation in plants. In the core promoter region, The TBP region of the general transcription factor TFIID recognizes the TATA-box and assembles with proteins such as TFIIF, TFIIE, and RNA polymerase II to form the transcription initiation complex that recognizes and binds TATA-box or downstream core promoter elements (DPE) to regulate the initiation of transcription together with transcription initiation sites (TSS) and other core elements. Some elements such as Y patch can enhance or stabilize promoter activity. Regulatory upstream elements and responsive elements (such as CAAT-box) bind to the regulated transcription factors and participate in the regulation of gene expression. *Cis*-acting elements such as enhancers exist on the upstream and downstream of genes which regulate gene expression through the synergistic effect of specific transcription factors.

晓等, 2004); 远端启动子区包含增强子(enhancer)、沉默子(silencer)和绝缘子(insulator)等顺式作用元件, 位置分布广泛(杨鹏芳等, 2018; Schmitz et al., 2022)。在启动子中, 核心启动子必要且保守, 近端启动子元件通过与不同转录因子结合赋予启动子时空响应差异, 远端启动子区深刻影响启动子的驱动表达活性, 这些元件组成了严谨调控基因表达的顺式作用元件系统。

1.1 核心启动子

植物的核心启动子(图1)指导RNA聚合酶II (RNA Pol II)准确启动基因序列(包括一些较短的保守序列)转录。与哺乳动物和昆虫等的元件组成略有不同, 植物核心启动子包含TATA-box、起始子(initiator element, Inr)、下游核心启动子元件(downstream core promoter element, DPE)和嘧啶片段(Y patch)等, 尚未发现BRE等元件(Brooks et al., 2023)。核心启动子在TSS附近首先被TFIID识别, 该蛋白进一步招募RNA Pol II和TFIIF等其它一般转录因子(general transcription factor, GTF)组装成转录起始复合体(pre-initiation complex, PIC); PIC决定基因的基础转录水平、方向和位置(Chen and Xu, 2022)。TATA-box位于Inr转录起始位点上游-30—70 bp处, 为富含AT/TA的保守序列区域, 是动物和植物启动子中最保守的元件(Yamamoto et al., 2007), 由一般转录因子TFIID的TBP区域识别, 决定了RNA聚合酶II的位置(Jores et al., 2021)。酵母实验证实, 部分不包含TATA-box的启动子也可被TFIID蛋白识别并且结合(Chen and Xu, 2022)。植物中, TATA-box常与组织特异性和胁迫诱导型驱动表达模式有关(Molina and Grotewold, 2005)。Inr直接覆盖TSS, 保守序列为PTCA₊₁NTTP (A₊₁表示转录起始位置) (Nakamura et al., 2002)。嘧啶片段是位于TSS上游13 bp的TC基序, 能使基因稳定表达(Yamamoto and Obokata, 2008; Bernard et al., 2010)。当TATA-box存在时, Inr和嘧啶片段可与其协同启动转录, 且表现出较高的驱动强度(Huang et al., 1996; Yamamoto et al., 2007); 在部分不存在TATA-box的启动子中(如管家光合作用相关基因启动子TFIID (Nakamura et al., 2002))则可以通过识别下游核心启动子元件招募转录相关蛋白(Chen and Xu, 2022)。DPE通常存在于TSS下游约

+30 bp处, 保守元件为RGWCGTG, DPE与Inr元件也具有协同反应, 二者的间距对于DPE依赖性启动子的转录活性至关重要(王婧等, 2014)。拟南芥(*Arabidopsis thaliana*)和水稻(*Oryza sativa*)中还存在一种“YR₊₁规则”, 即TSS上游通常有1个1 bp的C或T核苷酸, 下游有1个1 bp的A或G核苷酸(Y₋₁, 即C₋₁或T₋₁; R₊₁, 即A₊₁或G₊₁) (Yamamoto et al., 2007)。此外, 植物启动子启动基因表达的过程中都有5'UTR和前导内含子参与, 这是由于转录后mRNA水平和基因功能同样依赖于上游元件长度(Sun et al., 2017; Oldfield et al., 2019)。综上表明, 核心启动子通过较为保守的元件和严格的保护性结构来保证其功能稳定。

核心启动子在启动基因表达功能上具有3个特点。(1) 核心启动子TSS、Inr和DPE等元件的位置保守性较高, 对招募基础转录物质和确定转录功能具决定性作用; (2) 同一基因的启动子在不同基因组中可能具有不同活性; (3) 核心启动子通常不能指导基因高水平表达(Jores et al., 2021)。在设计人工合成启动子时, 该区域需满足的条件是结构简单且能稳定驱动转录。35S启动子核心片段(-46—+1 bp处)被称为“最小35S启动子”, 是合成启动子中组装基本转录元件的首选, 该片段插入增强子后可驱动基因高水平表达(Amack and Antunes, 2020)。

1.2 近端启动子元件

近端启动子元件(图1)包括一般上游调控元件和应答元件(response elements)。启动子的结构和表达分析显示, -500 bp至TSS以及-250 bp至TATA-box区域是近端启动子元件与转录因子结合位点的富集区(Weirauch et al., 2014; Ksouri et al., 2021; Hiratsuka et al., 2022; Rozière et al., 2022)。一般上游调控元件主要包括CAAT-box和GC-box两种, 二者均具有增强转录活性的作用。CAAT-box的保守序列为GGCT-CAATCT, 一般位于Inr上游-80—70 bp, 与其互作的因子有CTF家族的成员CP1、CP2和核因子NF-1等(Hernández-Rodríguez et al., 2009); 该元件与RNA聚合酶的结合有关(夏江东等, 2006)。GC-box是富含GC/CG序列的元件, 常以多拷贝形式存在于-90 bp处; 识别该序列的转录激活因子为Sp1, 在缺少TATA-box的启动子中可能具有转录起点作用

(Blake et al., 1990)。应答元件通常位于基因的上游, 能被转录因子识别并结合, 调控基因的专一性表达(刘玉瑛和张江丽, 2007)。目前, PlantPAN 4.0植物启动子数据库中收录了3 428个用于转录因子结合位点预测的不冗余矩阵(Chow et al., 2019), 表明植物启动子的结构和功能丰富度极高。利用这一特征, 分子育种家可通过不同元件的排列和重复方式, 创造特定条件下驱动基因表达的合成启动子。

近端启动子元件在应用上具有以下特点: (1) 许多典型的近端启动子元件的序列和位置在不同物种(包括双子叶和单子叶植物)中高度保守(Yamamoto et al., 2007; Lieberman-Lazarovich et al., 2019; Uygun et al., 2019), 因此, 可根据特定顺式作用元件的位置和数量预测基因表达模式, 乃至解码整个植物生长阶段合成启动子的表达模式; (2) 基于应答元件的结构和功能保守性, 为提高启动子的设计效率, 可将一些经典的特异序列作为不同物种启动子的设计元件; (3) 近端启动子元件对基因表达的调控相对独立(Belcher et al., 2020), 表明可通过正交设计等方式筛选出功能最佳的合成启动子元件。

1.3 远端启动子区

远端启动子区(图1)序列距转录起始位点较远, 甚至分布在非编码内含子序列中, 包含增强子、沉默子和绝缘子, 对基因表达的调控通常不受与TSS距离远近、自身拷贝数和基因表达方向的限制(Schmitz et al., 2022)。远端区域的顺式作用元件在发育特异性基因表达调控中发挥重要作用(Peremarti et al., 2010)。增强子是其中报道较多的基因表达增强元件, 能够促进基因的转录(王婧等, 2014), 且位置分布广泛, 可通过与特定的转录因子协同作用(Mehrotra et al., 2011)增强基因的表达(Aysha et al., 2018)。沉默子的位置分布具有较高的随机性, 一般分为2类: 一类短且位置分散, 能与阻遏蛋白结合并阻止PIC组装; 另一类位于内含子或者TSS附近, 通过形成特定的空间构型阻止转录(胡朝阳等, 2019)。绝缘子通常位于非编码区, 是具有基因调控功能的边界DNA序列, 通过抑制邻近增强子的作用或者阻断异染色质的扩张来保证调控信息的准确传递(王海等, 2011)。远端启动子区(非启动子重构必要元件)的位置较为分散且可控性不强, 因此基因工程中仅应用了部分经典的增强子

元件。

2 高等植物启动子元件的获得与鉴定

天然启动子可严格调控基因的时空表达, 但也存在结构冗余、功能单一和受物种限制等问题。基于前文对启动子结构的介绍, 发掘和鉴定非编码基因组中的元件是了解基因如何实现表达特性、简化启动子设计和扩充元件库的必要前提。

2.1 启动子及其元件的获得

启动子及元件的获得方法包括PCR克隆、启动子捕获、转录组和基因组筛选等。随着分子育种对基因表达精度和效率要求的提高, 获得新功能且更稳定地驱动表达活性的启动子元件成为启动子鉴定的主流趋势。

2.1.1 PCR克隆方法

PCR克隆方法包括以下3种: (1) 常规PCR, 即根据基因组合成引物克隆启动子DNA序列; (2) 反向PCR, 使用限制性核酸内切酶对基因组PCR进行酶切, 将产物片段环化, 根据已知目的基因片段设计引物进行反向PCR即可获得启动子元件; (3) 染色体步移法(chromosome walking)是获得已知DNA片段侧翼序列的方法。其中, 热不对称交错PCR技术(thermal asymmetric interlaced PCR, TAIL-PCR)在启动子克隆时最为常用, 经3轮巢式PCR分步对扩增的特异性启动子片段PCR产物进行富集(康丹等, 2013)。PCR克隆方法需要根据已知启动子或基因片段进行引物设计, 优势在于能较高效地获得已知基因的完整且高保真序列。

2.1.2 启动子陷阱

启动子陷阱(promoter trap)由基因陷阱(gene trap)发展而来, 该方法通过将无启动子驱动的报告基因质粒插入受体细胞, 根据报告基因的表达情况探究上游DNA序列的驱动活性(李季等, 2015)。在此基础上, 将有特定启动子驱动报告基因设计的质粒插入受体细胞, 可进一步检测插入位点附近的增强子等元件, 优势是可获得具较强活性的启动子片段(Yang et al., 2013; Chen et al., 2013)。使用这种方法, 研究人员在拟南芥(Alvarado et al., 2004)、香蕉(*Musa spp.*)

(Santos et al., 2009)和甘蓝型油菜(*Brassica napus*) (Bade et al., 2003)中筛选出组成型、胁迫诱导型和组织特异性启动子。该方法不依赖已知基因获得启动子序列,但需要受体植物有完善的转化体系,且鉴定周期较长(李季等, 2015)。

2.1.3 转录组文库筛选

相对于基因组,转录组数据更易获得。Liu等(2014)为提高大豆(*Glycine max*)的抗病性,根据公开的转录组数据对易感病大豆在大豆胞囊线虫作用下被诱导表达的基因进行*de-novo*基序挖掘,合成了大豆胞囊线虫诱导型启动子;Yang等(2021)使用*de-novo*算法从杨树(*Populus spp.*) RNA-seq数据库中筛选了30个干旱以及34个盐胁迫候选DNA基序;Zhou等(2023)从公开的拟南芥RNA-seq数据集中挖掘并验证组成型启动子文库,开发了使启动子在远亲物种中也具有可翻译性的元件预测工具PCONS。日益丰富的数据库和工具为启动子元件开发提供了极大助力。

2.1.4 基因组文库筛选

获得一定丰度的基因组文库后,即可通过不同探针与基因组文库杂交,获得具有调控作用的非编码区序列,进一步结合染色质分析等辅助技术实现元件高通量筛选,构建植物顺式作用元件组,甚至获得未知元件。快速发现基因组水平上的DNA基序或顺式作用元件方法包括酶联免疫共沉淀测序(electrophoretic mobility shift assay sequencing, EMSA-seq)、高通量测序-荧光配体互作分析(high-throughput sequencing-fluorescence ligand immunoprecipitation, HiTS-FLIP)、ATAC-seq (assay for transposase-accessible chromatin with high throughput sequencing)与细胞核分选、DNA亲和纯化测序(DNA affinity purification sequencing, DAP-seq)、染色质免疫沉淀测序(chromatin immunoprecipitation sequencing, ChIP-seq)和交联免疫沉淀测序(cross-linked immunoprecipitation sequencing, CLIP-seq)等(Liu et al., 2002; Chen et al., 2008; Nutiu et al., 2011; Bartlett et al., 2017; Lu et al., 2017)。但上述方法每次检测的转录因子有限,并且活性转录因子周围会产生开放的染色质区域,该区域包含许多可对检测精度产生干扰的非活性元件。计算机构建相应模型扩大了顺式作用

元件的鉴定范围。例如,Shahmuradov等(2017)使用TSS Plant工具基于植物启动子序列的18个重要组成和信号特征,通过人工神经网络模型预测了8种单/双子叶植物基因组中的TSSs和TATA-box等7个核心启动子元件;Arnold等(2013)利用自转录活性调控区测序(self-transcribing active regulatory region sequencing, STARR-seq)技术评估了数百万个候选序列的增强子活性;Wei等(2018)使用活性转录因子鉴定(active TF identification, ATI)技术通过富集细胞核提取物,从提取的DNA文库中直接分离了蛋白质结合的DNA序列,并对其进行质谱鉴定,获得了STARR-seq技术基础上的负调控元件信息;Wen等(2023)对ATI技术中胞质成分和细胞核内容物提取方法进行了优化,开发了级联抽提辅助-活性转录因子鉴定(sequential extraction assisted-active transcription factor identification, sea-ATI)新方法,在7种植物中鉴定了41个高活性顺式作用元件基序模型(motif models),其中15个模型在现有数据库中未被收录。

2.2 启动子元件的功能鉴定

最初,阐明顺式作用元件的功能是通过天然启动子缺失或突变后进行功能分析。植物天然启动子元件功能的一般鉴定方法包括生物信息学预测、利用报告基因进行表达分析和蛋白质-DNA互作鉴定等,这些方法可辅助探究启动子或顺式作用元件的时空表达特性、上游分子的调控机制及确定关键调节序列(Yasmeen et al., 2023)。

生物信息学分析是根据已有启动子中较为保守元件的信息与未知功能的启动子进行比对,预测其功能。Plant CARE (Lescot et al., 2002)、PLACE (Che et al., 2005)和Plant Prom (Shahmuradov et al., 2003)等网站收录了大部分高等植物启动子元件的序列和功能等信息。Plant PAN (Chow et al., 2019)、TRANSFAC (Fornes et al., 2020)和JASPAR (Hammal et al., 2022)等网站提供转录因子的DNA结合位点预测,科研人员可根据转录因子功能对启动子的时空表达特性进行判断。此外,通过生物信息学分析工具也可进行特征序列预测,发掘未被收录的新元件。使用转录组测序(RNA sequencing, RNA-seq)、ChIP-seq以及ATAC-seq等方法可获得大量具有潜在蛋白结合能力或环境响应功能的DNA序列,基于期望最

大(expectation maximization, EM)等算法, 利用MEME等工具筛选其中具有相似特征的短序列, 可以获得具有潜在调控功能的顺式作用元件(Liu et al., 2002; Bailey et al., 2009)。Yang等(2021)通过这种方法获得了新的干旱诱导元件, 由这些元件合成的SD-13启动子具有较高的干旱响应性。

表达验证是最直观的启动子功能鉴定方法, 即使使用未知启动子或其点突变/截短片段驱动报告基因, 直接检测受体植物不同生长时期或不同部位中的报告基因表达量及表达产物。常见的植物外源报告基因有*GUS*、*GFPI*/*OFPI*/*RFP*荧光蛋白基因和*LUC*荧光素酶基因等, 其优势在于这些基因在植物中的合成量较低, 便于进行定性和定量检测。植物内源报告基因包括花青素苷生物合成的转录因子编码基因*MrMYB1*和*MrbHLH1* (张玺丽等, 2017), 以及甜菜色素合成的串联基因*RUBY*等(He et al., 2020)。由于上述报告基因可直接合成色素, 因此可用于受体植物启动子驱动活性的无创监测, 尤其是对诱导型启动子, 可获得不同条件下连续观测的理想报告基因产物。

启动子功能的发挥与上游转录因子密不可分, 因此检测其相互作用的关键序列和分子机制十分重要。染色质免疫共沉淀(chromatin immunoprecipitation, ChIP)法、DNase I足迹法、电泳迁移率检测(electrophoretic mobility shift assay, EMSA)法和酵母单杂交(yeast one-hybrid, Y1H)检测法等蛋白质-DNA互作方法(曾晓玲等, 2018)可对启动子上的关键作用元件进行鉴定, 以确定该启动子的核心功能区域, 为其人工改造奠定基础。例如, 在番茄(*Solanum lycopersicum*)中鉴定到果实特定表达的关键功能区*SIHDC-A*, 就是在验证其上游因子结合区的基础上确定的(Kim et al., 2022)。同时, 转录因子也是影响启动子活性的重要因素, 在跨物种或品种应用启动子时 also 需考虑是否存在上游转录因子。

3 启动子元件的应用

本文基于PLACE网站中收录的469个维管植物的启动子元件和文献中更新的启动子元件(刘晓娜等, 2007; 郑少缘和范燕萍, 2011; 智联腾等, 2011; Mizoi et al., 2012; 咎新丽等, 2013; 张雄飞等, 2013; Du et al., 2014; Dey et al., 2015; 方彦昊等, 2015; 曹译文等,

2017; 王睿等, 2017; Brackmann et al., 2018; 黄珂等, 2018; 崔文文等, 2020; 王美华等, 2021; 张雪等, 2023), 总结出27类共计174个常见的植物应答元件和组织特异性元件(表1), 其中许多已成功应用于顺式元件重组。人工合成启动子的优势在于: (1) 可有效减少冗余序列并提高驱动表达效率; (2) 由于近端启动子元件在不同物种中具有保守性, 合成启动子具有较高的可控性; (3) 集合多种因子的调控元件可实现更精准调控, 如仅在植物根部受到干旱胁迫时表达的启动子, 可在提高作物抗旱性的同时, 避免抗性基因过度表达影响植物的正常生长(Jameel et al., 2020)。

3.1 核心启动子元件和启动子增强元件

核心启动子元件提供基础的启动转录功能, 为重组所必须。使用最广泛的核心启动子是最小启动子CaMV 35S, 其提供低基础驱动活性, 具有稳定转录起始的功能(Liu and Stewart, 2016), 在许多研究中被广泛使用(Ranjan and Dey, 2012; Jameel et al., 2020; Kar et al., 2022)。核心启动子的不同装配方式可合成双向启动子(Bi-directional promoters, BDPs), BDPs最初是通过在启动子5'端反向连接单向最小35S启动子片段合成(Chaturvedi et al., 2006; Zhang et al., 2008), 可驱动启动子区相邻基因在相反方向上转录, 以实现单一载体的多基因表达(In et al., 2020)。

一些核心启动子元件也是增强启动子驱动活性的常见元件。例如, 含TATA-box的核心启动子比不含TATA-box的驱动活性高4倍; 在随机合成启动子中, 添加Inr、嘧啶片段或TATA-box元件, 均能提高启动子活性, 其中TATA-box的作用最强(Jores et al., 2021)。Noda等(2013)使用*NtADH-5'UTR*区作为增强子, 实现了*CmF3H*启动子驱动*GUS*基因在菊花(*Chrysanthemum × morifolium*)舌状花中高效表达。Bai等(2020)将水稻*P_{Osrbcs-550}*和*P_{D540-544}*反向装配, 实现了同一载体特异性驱动基因在绿色组织中双向表达, 且使用*OsAct1*内含子作为启动子的增强元件, 使驱动活性显著提高。Du等(2014)使用增强子成功构建了高驱动活性的合成启动子。

3.2 近端启动子元件

如前文所述, 许多应答元件在不同遗传背景的物种中具有保守性。例如, HY5、ABRE、DREB、CBF和ACE

表1 高等植物启动子的顺式作用元件

Table1 *Cis-acting elements of promoters in higher plants*

元件类型	时空响应因子	元件名称
诱导型	生物因子 病原体	W-box、S-box、D-box、EIRE、GCC-box、H-box、G-box、GST1、PRE2PRE4E-box、SARE、PR-1基序、ERE、NPR-1基序、JERE、E17元件、F元件、FELEMENT1 (大豆根瘤菌)、TDBA12NTCHN50 (TMV)和OSE1ROOTNODULE (根瘤菌)
	创伤	20NTNTNOS、W-box、WIN、WRE、WUN-box和as-1-like
	物理因子 光	10PEHVPSBD (蓝光、白光或UV-A光)、ACE (UV-B和UV-1/蓝光)和ELRE1 (UV)
	光强	3AF1-box、AT-1 box、BOX-C、BOX-I、CPRF、CYTO位点、DE1、DRE (高光强胁迫)、G-box、GT2、HY5、I-box、LRE-box、MNF1、PAL-box、PI、PII、PRE、SITE1、SORLIPs、T-box、Z-DNA-forming序列和GT1基序
	光周期	EVENING
	温度 低温	DRE、CARGATCONSENSUS、CBF、CRT、LTRE、TCA-like、CAT-box、HSE和GAATTC
	高温	
	干旱和盐胁迫	ABAD、ABRE、ACGT、DRE、EMBP1、MYB1、MYB2和SRENTTTO1、NACR/HDZFR、MYBR/MYR、MBS、Erd1、TC-rich重复、GT1基序、E-box和STRE
	化学因子 激素	生长素
		RE、AuxRE、TGA元件、偶联元件CE1和CE3
		脱落酸
		ABRE、ABRC、ABRERATCAL、ABRELATERD1、偶联元件CEs和STRE
		细胞分裂素
		CANBNNAPA、MYBGAHV和CARG1
		赤霉素
		GARE基序和as-1-like
		茉莉酸
		20NTNTNOS、JASE1、JERE、T/G-box、ACG基序和as-1-like
		乙烯
		EIN3、EREGCC、ERE、GCCCORE、YREGIONNTPRB1B和Y区
		水杨酸
		AS1、LS5、LS7、SARECAMV和TCA1基序
	糖	284MOTIF、314MOTIF、ACGTABOX、TATCCAY基序、BBOXSITE1STPATS-box、CMSRE1和SREATMSD
	氮	EM和GLMHVCHORD
	其它物质	PAS (镉)、IDE1 (缺铁)、IRO (缺铁)和SURECORE (硫)
组织特异性	根	RSE、ROOTMOTIFTAPOX1、RHERPATEXPA7、tef-box、TGA1a、MYCCONSENS-USAT、SORLIP1AT、RAV1AAT、LEAFYATAG、SURECOREATSULTR11、P1BS (根毛)、SP8BFIBSP8AIB (块根)、141NTG13 (根尖)、AUXREPSIAA4 (根尖)、OSE1-ROOTNODULE (根瘤感染细胞)、OSE2ROOTNODULE (根瘤感染细胞)、WUSATAg (根尖分生组织)和XYLAT (核心木质部)
	茎	ABF、as-2-box、BBOXSITE1STPAT (块茎)和TSSR (块茎)
	叶	GATFLK、LPSE1、LSE1、GRA、RAV1 (莲座叶)和TCCAAAA基序(抑制基因在叶片表达)
	绿色组织	GEAT、GSE1和GSE2
	其它组织	ACIIPVPAL2 (维管束)、BS1 (维管束)、DOF结合位点(保卫细胞)和LPRSE1 (抑制基因在茎和种子中表达)
	花	AGAMOUSAT、AGL1、CArG-box、TACPyAT和CHS启动子核心片段(PCHS、LCHS) GATA-box、CACT-box、CACG-box、MYBPLANT (花茎)和MYB26PS (花蕾)
	花粉	PS区、POLLEN1LELAT52、POLLEN1、5256-box、定量元件、GTGANTG10、VOZ结合序列、GTGA-box、telo-box、A9和TA29启动子片段(绒毡层)及anther box (花药)
	果实	TAAAG基序、E-box、SEF结合位点、AGTTAGG、TGTCACA和SIHDC-A核心启动子区
	种子	Skn-1、RY基序、O2位点、E-box、AACACORE、ABAD、AMYBOX1、CAREOSREP1、EM、ESP、GLMHVCHORD、Sph元件、TGACGT基序、A27zn、Glb1、GCN4基序、CANBNNAPA和CATGTAA (大麦糊粉层)

括号中的文字表示该元件响应的具体环境、位置或生长发育信号。

Text in brackets indicates the specific environment, location, or growth/development signal that the element responds to.

等多种环境胁迫与激素诱导元件中均包含有ACGT基序,是植物响应非生物胁迫的核心序列(Khan et al., 2022)。ABRE是保守程度和胁迫响应性较高的启动子元件,在拟南芥、水稻和玉米(*Zea mays*)等植物中广泛存在,重构后的合成启动子可提高受体的抗逆性(Jameel et al., 2020)。但需要注意的是,由于启动子发挥作用需要招募特定转录因子,类似功能的不同元件可能存在细微差别。例如,水稻具有根部表达特性的启动子*GRP7*、*MTL*以及*RCg2*,它们在不同级别的根以及表皮/分生区具有不同的表达模式(Xue et al., 2016)。同时,由于植物不同组织和器官参与的生理过程具有时空联系,部分元件也具有诱导型和组织特异性响应这两种功能(Khan et al., 2022)。因此,有效利用丰富的近端启动子元件,是通过分子育种手段实现性状精准改良的关键,图2总结了其时空内在联系和元件应用方向。

3.2.1 诱导型启动子元件

诱导型启动子元件可与接收特定外源或内源信号的反式作用因子结合,驱动下游基因大量表达,使基因表达具有时间特异性。目前对诱导型元件及其互作因子已有较为详尽的研究(Misra and Ganesan, 2021)。诱导型启动子的串联、重复及与其它元件组合可以产生驱动活性更强的启动子。Hou等(2012)基于多个胁迫

反应元件构建了3个人工合成启动子(EKCM、EK-CRM以及ECCRM)驱动*GUS*报告基因表达,合成启动子的活性在干旱、低温和高盐条件下显著高于天然胁迫诱导型*rd29A*启动子;脱水条件下, EKCM启动子驱动的*GUS*酶活性是天然*rd29A*启动子的1.29倍。Yang等(2021)利用MEME suite等工具,从杨树RNA-seq数据库中获得16个响应干旱和34个响应盐胁迫的候选DNA基序,并通过qRT-PCR方法检测胁迫处理后的15个胁迫诱导型合成启动子的驱动效果,发现其中5个合成启动子比天然启动子具有更高的干旱和盐胁迫响应性。

3.2.2 组织特异性启动子元件

组织特异性启动子元件在植物生长发育过程中调控基因在特定部位表达,天然组织特异性启动子常存在物种特异性和驱动不稳定性(Du et al., 2014)。Du等(2014)使用百合(*Lilium brownii* var. *viridulum*)和矮牵牛(*Petunia hybrida*)的*CHS*启动子关键片段连接35S/OCS启动子增强区域,合成了比本源启动子驱动活性更强并且仅在花部表达的特异性启动子。Wang等(2015)为实现在水稻绿色组织中特异性表达,根据组织特异性元件信息,使用植物顺式作用元件数据库(PLACE)和水稻cDNA微阵列数据库(CREP)预测、*GUS*检测以及定点诱变法鉴定了潜在的绿色组织特

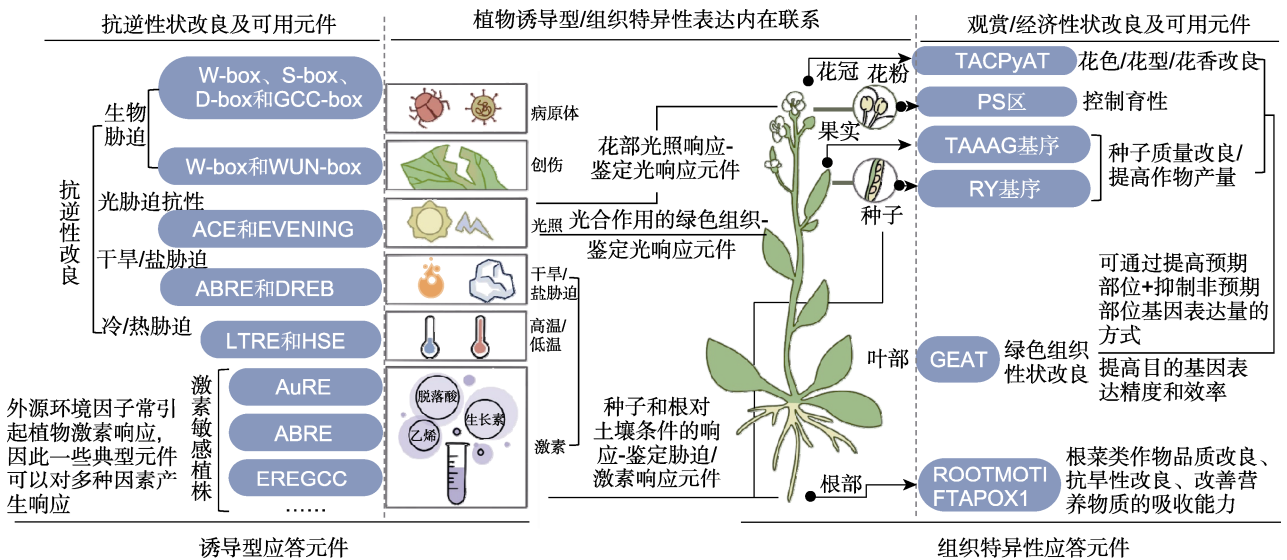


图2 近端诱导型和组织特异性启动子元件的内在联系及其在分子育种中的应用方向
Figure 2 The internal connection between proximal inducible and tissue-specific promoter elements and their application in molecular breeding

异性元件GEAT及其侧翼序列,并且筛选得到5个绿色组织特异性启动子GSSP1、GSSP3、GSSP5、GSSP6和GSSP7(Wang et al., 2015)。Kim等(2022)在番茄中鉴定出包含GATA、HSF以及AP1转录因子结合位点的SIHDC-A核心功能区域(果实特异性表达区域),并将其串联重复后与35S最小启动子重组,相比35S启动子,合成启动子仅在果实部位的驱动活性显著提高。发掘组织特异性启动子元件,为减轻目的基因对受体植物的代谢负荷、提高基因表达效率和抑制转基因产品扩散并提升其安全性奠定了基础。

4 高等植物合成启动子的设计方法

启动子元件数据充分,可人工构建启动子,方法是将需要的关键顺式作用元件按不同位置、间距和数量单独或重复地组合到现有天然或核心启动子中,可构建序列更短且表达准确的载体。与天然启动子在控制基因时空特异性表达和效率上的局限性相比,设计人工启动子旨在合成序列更短、活性更强且受目的条件控制的顺式作用元件,以协调多种基因在不同植物代谢和信号通路中的表达(Dey et al., 2015)。引入外源DNA元件也可增加调控元件的多样性,且相对于引

入同源序列,可以更好地减少基因沉默(Xue et al., 2016)。合成启动子的流程包括鉴定天然启动子功能,以及依据已知启动子元件进行重组设计并进行评估。启动子的重组设计方式主要包括点突变、随机重组、合成定向进化、定向重组和计算机生物学辅助设计等(图3)。

4.1 点突变

点突变是基因工程中改造顺式作用元件最常用的方法之一,通过PCR和基因编辑等方法在天然启动子片段中插入突变以创建或者删除目标顺式作用元件(Kumar et al., 2012),原理是通过酶切-PCR替换特定片段,或使用基因编辑产生多位点突变(Lieberman-Lazarovich et al., 2019),使目标启动子增减1个或多个靶向顺式作用元件(Ranjan and Dey, 2012)。与原始的GH3启动子相比,使用点突变顺式作用元件产生的DR5启动子对生长素的响应更敏感(Chen et al., 2013)。Efremova等(2020)通过突变草莓(*Fragaria × ananassa*) *pro-SmAMP2*启动子的ACTCAT以及CAANNNNATC负调控元件,产生了解除脯氨酸诱导且组成型驱动活性提高的启动子。利用CRISPR-

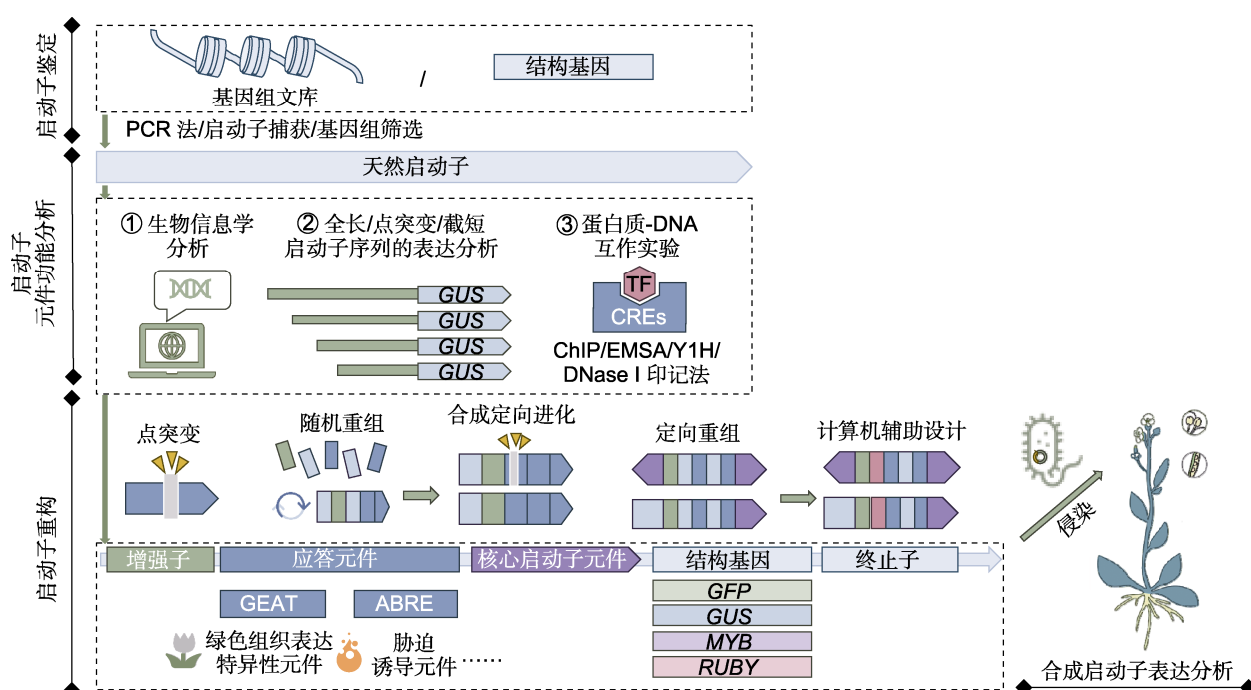


图3 高等植物启动子的鉴定、元件分析、重组和功能分析流程

Figure 3 Process of identification, element analysis, recombination and functional analysis of promoters in higher plants

Cas9等基因编辑工具,可在点突变的基础上提高编辑效率,扩大编辑范围,实现指定区域转录因子结合基序的引入、广泛基因组的靶标编辑和表观遗传标记的靶向编辑(Shi et al., 2023)。

4.2 随机重组

随机重组是通过DNase I酶将启动子序列酶切为2个及以上片段,并使用DNA聚合酶随机重组,使重组启动子片段包含不同启动子片段的多个拷贝(Ranjan and Dey, 2012)。这种重组方法产生的合成启动子数量多,但功能可控性较差。Ranjan和Dey (2012)使用该方法合成了6组共300个重组启动子,各组中,4%–34%的启动子获得了相比CaMV 35S启动子更高的驱动活性。

4.3 合成定向进化

进化选择压是进行定向筛选的动力之一,引入选择压可以对点突变和随机重组等方式产生的大量未知合成启动子进行有效选择。合成定向进化(synthetic directed evolution, SDE)是利用局部序列多样化和选择压力进化筛选出具有更强适应性的启动子(Yasmeen et al., 2023)。该方法通过对启动子进行多轮突变和检测,使其在时空选择压下产生驱动活性增强等有益突变(Oliva et al., 2019; Rao et al., 2021),不仅可用于筛选检测大量随机重组启动子中的最优者,还可产生定向突变。合成定向进化的优势在于可将受体植株的所有遗传物质作为整体来诱导定向进化,并可在多基因控制的数量性状改良上产生可控性更强的启动子元件。

Oliva等(2019)通过CRISPR-Cas9介导的基因编辑技术向3个蔗糖转运基因(*SWEET11*、*SWEET13*以及*SWEET14*)启动子中引入随机突变,易感病品种接种63株Xoo (白叶枯病菌)菌株后,未表现出感病性状,受体植株的抗病性增强。Liu等(2021)在玉米*ZmCLE7*和*ZmFCP1*启动子区域添加靶向突变,使*ZmCLE7*基因在发育的果穗原基中上调表达,从而提高产量。这种方法也可与有性杂交方法相结合,产生更大数量的潜在突变,且不需要二次遗传转化即可将突变启动子快速整合到植物中(Rodriguez-Leal et al., 2019)。

4.4 启动子定向重组

人工筛选特定元件进行重组也可降低启动子元件随机重组的不可控性。基于对已有顺式作用元件作用的理解,可使用天然启动子及其片段进行串联重复,进而按需进行设计,创造新的启动子。启动子定向重组的优势在于可有目的地提高启动子的驱动效率,将天然启动子的不同特性进行整合,也可合成双向启动子(3.1节)。启动子定向重组的方法是将2个或多个重要的启动子片段或结构域通过PCR等方法连接起来。Gupta等(2021)将已知功能的RTBV启动子片段嵌合插入完整的MMV启动子中,产生了3个较CaMV 35S启动子驱动活性更强的启动子。Jameel等(2020)将大豆的根特异性启动子以及胁迫诱导型启动子元件(ABRE、MYB和E2F-VARIANT等11种元件)与CaMV 35S最小启动子进行串联合成,构建了具有胁迫诱导和根部组织表达特性的人工启动子。

合成高敏感性启动子元件可用于生成植物传感器的高精度“开关”。其可以报告基因的表达作为信号并辅以快速数字化检测,用于环境监测等场景。例如,利用组氨酸激酶信号转导途径构建的人工Plant-Pho启动子可以用于监测环境污染物、爆炸物和化学试剂(Antunes et al., 2011)。缺氧响应启动子元件ERF1和ERF2在多个物种的缺氧反应基因启动子中富集,可作为监测植物低氧条件和种子缺氧萌发的有效元件(Gasch et al., 2016)。以M-EZ2增强子为基础的磷酸盐缺乏响应BDP元件对磷浓度的变化反应敏感(Araceli et al., 2017)。Persad-Russell等(2022)在马铃薯(*S. tuberosum*)中合成了一种植物病原体传感器,使用合成4xS-box启动子-Q-system转录因子的报告系统能在细菌感染早期检测细胞外部信号,防止病原在作物中扩散。

4.5 计算机辅助设计

基于计算机提供的合成启动子驱动表达模型可预测重组启动子元件的功能,提高启动子的设计效率和准确率(Zrimec et al., 2020)。数字化模型的建立通常需要满足以下3个条件。(1) 具有数据量充足的顺式作用元件/DNA-蛋白质相互作用信息库;(2) 设计恰当的数据模型,如位置权重矩阵(position weight matrix, PWM)等,用于预测启动子特征以及表达谱的普遍规

律; (3) 具有大规模、快速和精准的基因表达谱检测平台。除已知功能的顺式作用元件外, 其余的侧翼序列(flanking sequence)也可以通过AI模型进行优化(Yasmeen et al., 2023)。例如, DeepSEED是一种用于设计合成启动子的AI辅助侧翼序列优化方法, 可提高真核细胞中77.8%被测启动子的诱导活性(Zhang et al., 2023)。

5 结论与展望

启动子与转录因子互作是高等植物转录调控的核心, 启动子的特异性元件是其发挥作用的关键。合成启动子在转化研究和植物生物技术方面具有广阔的应用前景。基因在不同物种和生理过程的表达具有差异性, 植物的反式作用因子通过接收特定的环境和发育信号, 与保守DNA基序相结合, 这一基础调控过程为基因表达赋予了时空特异性。随着生物技术的发展, 基因工程需要构建更精准和高效的表达调控载体。本文综述了目前已知的高等植物启动子各部分结构及其功能, 总结了具有不同特异性和强度的27类、174个典型顺式作用元件, 以及启动子设计和重构方法, 为植物基因工程中启动子的准确与高效利用提供了参考。

5.1 合成启动子的应用前景

合成启动子在基因工程研究中具有很大的优势, 尤其在基因编辑技术和合成生物学等新技术方面展现出强大的应用潜力。(1) 提高基因表达精度和效率, 增加受体基因组的多样性, 减少基因沉默, 为探究基因表达模式和提高作物产量奠定基础; (2) 可作为环境监测与保护的敏感“开关”, 通过加入识别物理或化学信号的合成启动子并与报告基因融合, 可作为植物传感器的一个模块(Liu et al., 2011); (3) 合成精细调控且具有复合功能的启动子, 用于1至多个靶蛋白的精准表达。在基因编辑技术中, 由于靶基因的近端启动子区域包含gRNA结合序列的单个或多个拷贝, 可逐步调节植物中内源基因的表达。合成启动子可以精细调节代谢通路中限速酶编码基因的表达, 在基因编辑和植物代谢工程中具有广阔的应用前景(Selma et al., 2019; Persad-Russell et al., 2022); (4) 提高转基因产品的安全性: 转基因植物在生长过程中会受到杂交、与其它种群和土壤的物质交换及外源基因漂移等因素的影响, 产生环境安全风险, 这是限制转基因产

品商业化的重要因素之一(梁晋刚等, 2021)。根据特异性启动子可有效降低目的基因在非预期时期或部位的表达特性, 可控制外源基因在受体植物中的表达部位、调节植物育性和降低植物的保护地栽培成本, 进而限制转基因产品与其它物种和土壤中的物质交换、基因传递和栽培范围, 有助于转基因产品的推广应用。

5.2 设计合成启动子的挑战

当前启动子应用和改造面临的挑战之一是特异性元件不足。对于植物启动子及其关键元件的功能鉴定研究工作仍然十分有限(Zhang et al., 2023)。使用Plant CARE和PLACE启动子元件预测数据库只能提供已有的元件数据。今后可通过以下方法扩充启动子元件库。(1) 将微阵列和RNA-seq表达公开数据集与基因组序列信息共同分析, 利用MEME、MDscan、MotifSampler、Weeder、BioProspector、W-AlignACE和SCOPE等生物信息学软件预测新元件并验证设计的靶向序列; (2) 基因组分析结合计算机进行高通量筛选, 有助于发掘新元件。例如, Wen等(2023)通过改良ATI技术, 高通量筛选了15个新发现的顺式作用元件模型和负调控模型; (3) 利用基因编辑方法对靶位点进行批量突变, 可提高顺式作用元件的功能分析效率并用于植物数量性状改良。

其次, 为提高合成启动子的有效性, 还需要注意DNA基序对转录因子亲和力的影响。启动子能够产生时空特异性的关键是应答元件, 这些CREs的保守功能源于顺式调节模型(cis-regulatory modules, CRMs)的保守性。CRMs在不同环境因子、组织部位和物种间均具有稳定性(Diehl and Boyle, 2018)。Inukai等(2017)研究发现, DNA构象和侧翼序列对其亲和性也会产生影响。今后可使用正交设计对目的元件进行多种组合, 通过插入间隔序列和保护序列等提高植物的遗传稳定性(Jopcik et al., 2014; He et al., 2023), 并辅以计算机模型预测和表达分析筛选最优启动子。

综上所述, 引入人工合成启动子的基因工程载体设计毋庸置疑会向更简洁、高效和精准指导转录方向发展。如果合成启动子变得更加“智能”, 必将迅速且广泛应用于提高作物产量和质量, 及进化生物学和大数据分析等多个研究领域。

作者贡献声明

洪艳和戴思兰: 指导修改文章; 宋想: 文章起草和撰写; 王璐瑶: 文章撰写和校对; 富博晓、李双达和魏媛媛: 文稿校对。

参考文献

- 曹译文, 宋阳, 渠可心, 王丕武 (2017). 大豆组织特异启动子的克隆与功能分析. 中国油料作物学报 **39**, 771–777.
- 崔文文, 迟婧, 冯艳芳, 耿丽丽, 刘荣梅 (2020). 人工合成根特异启动子SRSP的功能分析. 生物工程学报 **36**, 700–706.
- 邓泽宜, 宋想, 洪艳, 戴思兰 (2021). 启动子在观赏植物基因工程中的应用综述. 园艺学报 **48**, 1250–1264.
- 方彦昊, 南文斌, 梁永书, 张汉马 (2015). 植物组织特异性基因表达技术及其应用. 植物生理学报 **51**, 797–805.
- 胡朝阳, 唐培培, 邓炎春, 孙伟娟, 姚勤 (2019). 转录水平调控中的负调控元件——沉默子. 生命科学 **31**, 686–692.
- 黄珂, 黄格格, 薛满德, 龙艳, 袁潜华, 裴新梧 (2018). 普通野生稻根特异启动子的克隆与鉴定. 生物技术通报 **34**(8), 87–92.
- 康丹, 方小艳, 游腾飞, 王婷婷, 睦安平, 杨星勇 (2013). 染色体步移技术克隆已知序列侧翼启动子的研究进展. 农业生物技术学报 **21**, 355–366.
- 李季, 黄天带, 华玉伟, 黄华孙 (2015). 启动子陷阱技术在植物启动子克隆研究中的应用. 热带农业科学 **35**(9), 46–50.
- 梁晋刚, 张开心, 张旭冬, 王颢潜, 陈子言, 刘鹏程, 张秀杰 (2021). 中国农业转基因生物环境安全检测标准体系现状与展望. 中国油料作物学报 **43**, 1–14.
- 刘晓娜, 付畅, 黄永芬 (2007). 种子特异性启动子研究进展. 植物学通报 **24**, 218–225.
- 刘玉瑛, 张江丽 (2007). 真核生物启动子预测相关数据库资源概述. 安徽农业科学 **35**, 7418–7419.
- 王海, 张倩, 方向东 (2011). 绝缘子调控基因的表达. 中国生物化学与分子生物学报 **27**, 493–498.
- 王婧, 李冰, 刘翠翠, 朱阵, 张继瑜 (2014). 启动子结构和功能研究进展. 生物技术通报 (8), 40–45.
- 王美华, 高洁, 李玉莲, 张淑娟, 宋国琦, 张荣志, 李玮, 李吉虎, 李根英 (2021). 籽粒高效特异性表达启动子的克隆与表达分析. 山东农业科学 **53**(5), 45–50.
- 王睿, 朱梦琳, 高方远, 任野胜, 陆贤军, 任光俊, 林拥军 (2017). 水稻组织特异型人工合成启动子的设计、构建及功能鉴定. 作物学报 **43**, 789–794.

- 夏江东, 程在全, 黄兴奇, 季鹏章, 熊华斌 (2006). 矮牵牛花色 *CHS-A* 基因启动子(Pchsa)的克隆及序列分析. 西南农业学报 **19**, 676–678.
- 杨鹏芳, 段国琴, 胡晓炜, 缪秀梅, 南淑珍, 张丽静 (2018). 高等植物启动子研究概述. 分子植物育种 **16**, 1482–1493.
- 咎新丽, 高英, 陈玉玲, 赵开军 (2013). 病原菌诱导型启动子顺式作用元件及其互作的转录因子. 植物学报 **48**, 219–229.
- 曾晓玲, 赵昶灵, 文国松, 丁灿, 张洪玲, 徐率, 古朝山 (2018). 启动子结构、功能预测和验证方法的研究进展. 分子植物育种 **16**, 3915–3925.
- 张春晓, 王文棋, 蒋湘宁, 陈雪梅 (2004). 植物基因启动子研究进展. 遗传学报 **31**, 1455–1464.
- 张玺丽, 殷学仁, 李方, 陈昆松, 刘晓芬 (2017). *MrMYB1-MrbHLH1*: 一个有潜力的园艺植物转基因可视化报告基因. 园艺学报 **44**, 2296–2304.
- 张雄飞, 刘雅莉, 姜倩, 祁银燕, 杜灵娟 (2013). 通过重叠PCR构建2个增强型植物花特异双向启动子. 浙江大学学报(农业与生命科学版) **39**, 34–41.
- 张雪, 程荔书, 张军, 魏爽, 赵艳 (2023). 高等植物不同类型启动子及其相关顺式元件研究进展. 高师理科学刊 **43**(4), 60–67.
- 郑少缘, 范燕萍 (2011). 花器官特异启动子的研究进展. 安徽农业科学 **39**, 19067–19069.
- 智联腾, 赵倩, 敖光明, 于静娟 (2011). 天然彩色棉纤维特异表达启动子 *LTP3* 的克隆及其在烟草中的表达特异性. 热带生物学报 **2**, 133–137.
- Alvarado MC, Zsigmond LM, Kovács I, Cséplő Á, Koncz C, Szabados LM (2004). Gene trapping with firefly luciferase in *Arabidopsis*. Tagging of stress-responsive genes. *Plant Physiol* **134**, 18–27.
- Amack SC, Antunes M (2020). CaMV35S promoter—a plant biology and biotechnology workhorse in the era of synthetic biology. *Curr Plant Biol* **24**, 100179.
- Amen RD (1968). A model of seed dormancy. *Bot Rev* **34**, 1–31.
- Annadana S, Beekwilder MJ, Kuipers G, Visser PB, Outchkourov N, Pereira A, Udayakumar M, De Jong J, Jongsma MA (2002). Cloning of the chrysanthemum *UEP1* promoter and comparative expression in florets and leaves of *Dendranthema grandiflora*. *Transgenic Res* **11**, 437–445.
- Antunes MS, Morey KJ, Smith JJ, Albrecht KD, Bowen TA, Zdunek JK, Troupe JF, Cuneo MJ, Webb CT, Helling HW, Medford JI (2011). Programmable ligand de-

- tection system in plants through a synthetic signal transduction pathway. *PLoS One* **6**, e16292.
- Araceli OA, Alfredo CR, Javier MM, Luis HE** (2017). A phosphate starvation-driven bidirectional promoter as a potential tool for crop improvement and *in vitro* plant biotechnology. *Plant Biotechnol J* **15**, 558–567.
- Arnold CD, Gerlach D, Stelzer C, Boryń ŁM, Rath M, Stark A** (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074–1077.
- Aysha J, Noman M, Wang FW, Liu WC, Zhou YG, Li HY, Li XW** (2018). Synthetic promoters: designing the *cis* regulatory modules for controlled gene expression. *Mol Biotechnol* **60**, 608–620.
- Bade J, van Grinsven E, Custers J, Hoekstra S, Ponstein A** (2003). T-DNA tagging in *Brassica napus* as an efficient tool for the isolation of new promoters for selectable marker genes. *Plant Mol Biol* **52**, 53–68.
- Bai JY, Wang X, Wu H, Ling F, Zhao Y, Lin YJ, Wang R** (2020). Comprehensive construction strategy of bidirectional green tissue-specific synthetic promoters. *Plant Biotechnol J* **18**, 668–678.
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS** (2009). MEME-SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**, W202–W208.
- Bartlett A, O'Malley RC, Huang SSC, Galli M, Nery JR, Gallavotti A, Ecker JR** (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat Protoc* **12**, 1659–1672.
- Belcher MS, Vuu KM, Zhou A, Mansoori N, Ramos AA, Thompson MG, Scheller HV, Loqué D, Shih PM** (2020). Design of orthogonal regulatory systems for modulating gene expression in plants. *Nat Chem Biol* **16**, 857–865.
- Beliaev DV, Yourieva NO, Tereshonok DV, Derevyagina MK, Meleshin AA** (2023). Early blight resistance of transgenic potato plants expressing the *ProSmAMP1* gene for antimicrobial peptides under the control of a light-inducible *cab* promoter. *Russ J Plant Physiol* **70**, 57.
- Bernard V, Brunaud V, Lecharny A** (2010). TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics* **11**, 166.
- Blake MC, Jambou RC, Swick AG, Kahn JW, Azizkhan JC** (1990). Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol Cell Biol* **10**, 6632–6641.
- Brackmann K, Qi JY, Gebert M, Jouannet V, Schlamp T, Grünwald K, Wallner ES, Novikova DD, Levitsky VG, Agustí J, Sanchez P, Lohmann JU, Greb T** (2018). Spatial specificity of auxin responses coordinates wood formation. *Nat Commun* **9**, 875.
- Brázda V, Bartas M, Bowater RP** (2021). Evolution of diverse strategies for promoter regulation. *Trends Genet* **37**, 730–744.
- Brooks EG, Elorriaga E, Liu Y, Duduit JR, Yuan GL, Tsai CJ, Tuskan GA, Ranney TG, Yang XH, Liu WS** (2023). Plant promoters and terminators for high-precision bioengineering. *Biodes Res* **5**, 0013.
- Butler JEF, Kadonaga JT** (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* **16**, 2583–2592.
- Chaturvedi CP, Sawant SV, Kiran K, Mehrotra R, Lodhi N, Ansari SA, Tuli R** (2006). Analysis of polarity in the expression from a multifactorial bidirectional promoter designed for high-level expression of transgenes in plants. *J Biotechnol* **123**, 1–12.
- Che DS, Jensen S, Cai LM, Liu JS** (2005). BEST: binding-site estimation suite of tools. *Bioinformatics* **21**, 2909–2911.
- Chen X, Guo LQ, Fan ZC, Jiang T** (2008). W-AlignACE: an improved Gibbs sampling algorithm based on more accurate position weight matrices learned from sequence and gene expression/ChIP-chip data. *Bioinformatics* **24**, 1121–1128.
- Chen XZ, Xu YH** (2022). Structural insights into assembly of transcription preinitiation complex. *Curr Opin Struct Biol* **75**, 102404.
- Chen YR, Yordanov YS, Ma C, Strauss S, Busov VB** (2013). *DR5* as a reporter system to study auxin response in *Populus*. *Plant Cell Rep* **32**, 453–463.
- Chow CN, Lee TY, Hung YC, Li GZ, Tseng KC, Liu YH, Kuo PL, Zheng HQ, Chang WC** (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res* **47**, D1155–D1163.
- Dey N, Sarkar S, Acharya S, Maitil B** (2015). Synthetic promoters *in planta*. *Planta* **242**, 1077–1094.
- Diehl AG, Boyle AP** (2018). Conserved and species-specific transcription factor co-binding patterns drive divergent gene regulation in human and mouse. *Nucleic Acids Res* **46**, 1878–1894.
- Du LJ, Lou Q, Zhang XF, Jiao SZ, Liu YL, Wang YJ** (2014). Construction of flower-specific chimeric promoters and analysis of their activities in transgenic *Torenia*. *Plant*

- Mol Biol Rep* **32**, 234–245.
- Efremova LN, Strelnikova SR, Gazizova GR, Minkina EA, Komakhin RA** (2020). A synthetic strong and constitutive promoter derived from the *Stellaria media* pro-SmAMP1 and pro-SmAMP2 promoters for effective transgene expression in plants. *Genes* **11**, 1407.
- Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranasic D, Santana-Garcia W, Tan G, Chèneby J, Ballester B, Parcy F, Sandelin A, Lenhard B, Wasserman WW, Mathelier A** (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **48**, D87–D92.
- Gasch P, Funding M, Müller JT, Lee T, Bailey-Serres J, Mustroph A** (2016). Redundant ERF-VII transcription factors bind to an evolutionarily conserved *cis*-motif to regulate hypoxia-responsive gene expression in *Arabidopsis*. *Plant Cell* **28**, 160–180.
- Gupta D, Dey N, Leelavathi S, Ranjan R** (2021). Development of efficient synthetic promoters derived from pararetrovirus suitable for translational research. *Planta* **253**, 42.
- Hamma IF, deLangen P, Bergon A, Lopez F, Ballester B** (2022). ReMap 2022: a database of human, mouse, drosophila and *Arabidopsis* regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res* **50**, D316–D325.
- Han ML, Yin J, Zhao YH, Sun XW, Meng JX, Zhou J, Shen T, Li HH, Zhang F** (2020). How the color fades from *Malus halliana* flowers: transcriptome sequencing and DNA methylation analysis. *Front Plant Sci* **11**, 576054.
- He SF, Zhang ZW, Lu WY** (2023). Natural promoters and promoter engineering strategies for metabolic regulation in *Saccharomyces cerevisiae*. *J Ind Microbiol Biotechnol* **50**, kuac029.
- He YB, Zhang T, Sun H, Zhan HD, Zhao YD** (2020). A reporter for noninvasively monitoring gene expression and plant transformation. *Hortic Res* **7**, 152.
- Hemberg T** (1949). Growth-inhibiting substances in terminal buds of *Fraxinus*. *Physiol Plant* **2**, 37–44.
- Hernández-Rodríguez CS, Ferré J, Herrero S** (2009). Genomic structure and promoter analysis of pathogen-induced repeat genes from *Spodoptera exigua*. *Insect Mol Biol* **18**, 77–85.
- Hiratsuka T, Makita Y, Yamamoto YY** (2022). Sequence-based evaluation of promoter context for prediction of transcription start sites in *Arabidopsis* and rice. *Sci Rep* **12**, 6976.
- Hou L, Chen LJ, Wang JY, Xu DF, Dai LX, Zhang H, Zhao YX** (2012). Construction of stress responsive synthetic promoters and analysis of their activity in transgenic *Arabidopsis thaliana*. *Plant Mol Biol Rep* **30**, 1496–1506.
- Huang WB, Wong JM, Bateman E** (1996). TATA elements direct bi-directional transcription by RNA polymerases II and III. *Nucleic Acids Res* **24**, 1158–1163.
- In S, Lee HA, Woo J, Park E, Choi D** (2020). Molecular characterization of apathogen-inducible bidirectional promoter from hot pepper (*Capsicum annuum*). *Mol Plant Microbe Interact* **33**, 1330–1339.
- Inukai S, Kock KH, Bulyk ML** (2017). Transcription factor-DNA binding: beyond binding site motifs. *Curr Opin Genet Dev* **43**, 110–119.
- Jameel A, Noman M, Liu WC, Ahmad N, Wang FW, Li XW, Li HY** (2020). Tinkering *cis* motifs jigsaw puzzle led to root-specific drought-inducible novel synthetic promoters. *Int J Mol Sci* **21**, 1357.
- Jopcik M, Moravcikova J, Matusikova I, Libantova J** (2014). Spacer length-dependent protection of specific activity of pollen and/or embryo promoters from influence of *CaMV35S* promoter/enhancer in transgenic plants. *Plant Cell Tissue Organ Cult* **118**, 507–518.
- Jores T, Tonnies J, Wrightsman T, Buckler ES, Cuperus JT, Fields S, Queitsch C** (2021). Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat Plants* **7**, 842–855.
- Kar S, Bordiya Y, Rodriguez N, Kim J, Gardner EC, Gollihar JD, Sung S, Ellington AD** (2022). Orthogonal control of gene expression in plants using synthetic promoters and CRISPR-based transcription factors. *Plant Methods* **18**, 42.
- Khan ZH, Dang S, Memaya MB, Bhadouriya SL, Agarwal S, Mehrotra S, Gupta D, Mehrotra R** (2022). Genome-wide analysis of AAAG and ACGT *cis*-elements in *Arabidopsis thaliana* reveals their involvement with genes downregulated under jasmonic acid response in an orientation independent manner. *G3-Genes Genom Genet* **12**, jkac057.
- Kim HM, Park SH, Park SY, Ma SH, Do JH, Kim AY, Jeon MJ, Shim JS, Joung YH** (2022). Identification of essential element determining fruit-specific transcriptional activity in the tomato *HISTIDINE DECARBOXYLASE A* gene promoter. *Plant Cell Rep* **41**, 1721–1731.
- Ksouri N, Castro-Mondragón JA, Montardit-Tarda F, vanHelden J, Contreras-Moreira B, Gogorcena Y**

- (2021). Tuning promoter boundaries improves regulatory motif discovery in nonmodel plants: the peach example. *Plant Physiol* **185**, 1242–1258.
- Kumar D, Patro S, Ghosh J, Das A, Maiti IB, Dey N** (2012). Development of asalicylic acid inducible minimal sub-genomic transcript promoter from Figwort mosaic virus with enhanced root- and leaf-activity using TGACG motif rearrangement. *Gene* **503**, 36–47.
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S** (2002). PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* **30**, 325–327.
- Lieberman-Lazarovich M, Yahav C, Israeli A, Efroni I** (2019). Deep conservation of *cis*-element variants regulating plant hormonal responses. *Plant Cell* **31**, 2559–2572.
- Liu L, Gallagher J, Arevalo ED, Chen R, Skopelitis T, Wu QY, Bartlett M, Jackson D** (2021). Enhancing grain-yield-related traits by CRISPR-Cas9 promoter editing of maize *CLE* genes. *Nat Plants* **7**, 287–294.
- Liu WS, Mazarei M, Peng YH, Fethe MH, Rudis MR, Lin JY, Millwood RJ, Arelli PR, Stewart CN Jr** (2014). Computational discovery of soybean promoter *cis*-regulatory elements for the construction of soybean cyst nematode-inducible synthetic promoters. *Plant Biotechnol J* **12**, 1015–1026.
- Liu WS, Mazarei M, Rudis MR, Fethe MH, Stewart CN Jr** (2011). Rapid *in vivo* analysis of synthetic promoters for plant pathogen phytosensing. *BMC Biotechnol* **11**, 108.
- Liu WS, Stewart CN Jr** (2016). Plant synthetic promoters and transcription factors. *Curr Opin Biotechnol* **37**, 36–44.
- Liu XS, Brutlag DL, Liu JS** (2002). An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nat Biotechnol* **20**, 835–839.
- Lu ZF, Hofmeister BT, Vollmers C, DuBois RM, Schmitz RJ** (2017). Combining ATAC-seq with nuclei sorting for discovery of *cis*-regulatory regions in plant genomes. *Nucleic Acids Res* **45**, e41.
- Lucibelli F, Valoroso MC, Aceto S** (2022). Plant DNA methylation: an epigenetic mark in development, environmental interactions, and evolution. *Int J Mol Sci* **23**, 8299.
- Mehrotra R, Gupta G, Sethi R, Bhalothia P, Kumar N, Mehrotra S** (2011). Designer promoter: an artwork of *cis* engineering. *Plant Mol Biol* **75**, 527–536.
- Misra S, Ganesan M** (2021). The impact of inducible promoters in transgenic plant production and crop improvement. *Plant Gene* **27**, 100300.
- Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K** (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim Biophys Acta Gene Regul Mech* **1819**, 86–96.
- Molina C, Grotewold E** (2005). Genome wide analysis of *Arabidopsis* core promoters. *BMC Genomics* **6**, 25.
- Nakamura M, Tsunoda T, Obokata J** (2002). Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J* **29**, 1–10.
- Noda N, Aida R, Kishimoto S, Ishiguro K, Fukuchi-Mizutani M, Tanaka Y, Ohmiya A** (2013). Genetic engineering of novel bluer-colored chrysanthemums produced by accumulation of delphinidin-based anthocyanins. *Plant Cell Physiol* **54**, 1684–1695.
- Noda N, Yoshioka S, Kishimoto S, Nakayama M, Douzono M, Tanaka Y, Aida R** (2017). Generation of blue chrysanthemums by anthocyanin B-ring hydroxylation and glucosylation and its coloration mechanism. *Sci Adv* **3**, e1602785.
- Nutiu R, Friedman RC, Luo SJ, Khrebtukova I, Silva D, Li RB, Zhang L, Schroth GP, Burge CB** (2011). Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. *Nat Biotechnol* **29**, 659–664.
- Oldfield AJ, Henriques T, Kumar D, Burkholder AB, Cinghu S, Paulet D, Bennett BD, Yang PY, Scruggs BS, Lavender CA, Rivals E, Adelman K, Jothi R** (2019). NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. *Nat Commun* **10**, 3072.
- Oliva R, Ji CH, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom JS, Li CH, Nguyen H, Liu B, Auguy F, Sciallano C, Luu VT, Dossa GS, Cunnac S, Schmidt SM, Slamet-Loedin IH, Cruz CV, Szurek B, Frommer WB, White FF, Yang B** (2019). Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat Biotechnol* **37**, 1344–1350.
- Peremarti A, Twyman RM, Gómez-Galera S, Naqvi S, Farré G, Sabalza M, Miralpeix B, Dashevskaya S, Yuan DW, Ramessar K, Christou P, Zhu CF, Bassie L, Capell T** (2010). Promoter diversity in multigene transformation. *Plant Mol Biol* **73**, 363–378.
- Persad-Russell R, Mazarei M, Schimel TM, Howe L, Schmid MJ, Kakeshpour T, Barnes CN, Brabazon H, Seaberry EM, Reuter DN, Lenaghan SC, Stewart CN Jr** (2022). Specific bacterial pathogen phytosensing is ena-

- bled by a synthetic promoter-transcription factor system in potato. *Front Plant Sci* **13**, 873480.
- Porto MS, Pinheiro MPN, Batista VGL, dos Santos RC, de Albuquerque Melo Filho P, de Lima LM** (2014). Plant promoters: an approach of structure and function. *Mol Biotechnol* **56**, 38–49.
- Ranjan R, Dey N** (2012). Development of vascular tissue and stress inducible hybrid-synthetic promoters through DOF-1 motifs rearrangement. *Cell Biochem Biophys* **63**, 235–245.
- Rao GS, Jiang WJ, Mahfouz M** (2021). Synthetic directed evolution in plants: unlocking trait engineering and improvement. *Synth Biol* **6**, ysab025.
- Rodriguez-Leal D, Xu C, Kwon CT, Soyars C, Demesa-Arevalo E, Man J, Liu L, Lemmon ZH, Jones DS, Van Eck J, Jackson DP, Bartlett ME, Nimchuk ZL, Lippman ZB** (2019). Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. *Nat Genet* **51**, 786–792.
- Rozière J, Guichard C, Brunaud V, Martin ML, Coursol S** (2022). A comprehensive map of preferentially located motifs reveals distinct proximal *cis*-regulatory sequences in plants. *Front Plant Sci* **13**, 976371.
- Rushton PJ** (2016). What have we learned about synthetic promoter construction? *Methods Mol Biol* **1482**, 1–13.
- Santos E, Remy S, Thiry E, Windelinckx S, Swennen R, Sági L** (2009). Characterization and isolation of a T-DNA tagged banana promoter active during *in vitro* culture and low temperature stress. *BMC Plant Biol* **9**, 77.
- Schmitz RJ, Grotewold E, Stam M** (2022). *Cis*-regulatory sequences in plants: their importance, discovery, and future challenges. *Plant Cell* **34**, 718–741.
- Selma S, Bernabé-Orts JM, Vazquez-Vilar M, Diego-Martin B, Ajenjo M, Garcia-Carpintero V, Granell A, Orzaez D** (2019). Strong gene activation in plants with genome-wide specificity using a new orthogonal CRISPR/Cas9-based programmable transcriptional activator. *Plant Biotechnol J* **17**, 1703–1705.
- Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV** (2003). PlantProm: a database of plant promoter sequences. *Nucleic Acids Res* **31**, 114–117.
- Shahmuradov IA, Umarov RK, Solovyev VV** (2017). TSSPlant: a new tool for prediction of plant Pol II promoters. *Nucleic Acids Res* **45**, e65.
- Shi L, Su J, Cho MJ, Song H, Dong XO, Liang Y, Zhang ZY** (2023). Promoter editing for the genetic improvement of crops. *J Exp Bot* **74**, 4349–4366.
- Sun M, Ding JY, Li DL, Yang GP, Cheng ZN, Zhu QB** (2017). NUDT21 regulates 3'-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. *Cancer Lett* **410**, 158–168.
- Uygun S, Azodi CB, Shiu SH** (2019). *Cis*-regulatory code for predicting plant cell-type transcriptional response to high salinity. *Plant Physiol* **181**, 1739–1751.
- Wang R, Zhu ML, Ye RJ, Liu ZX, Zhou F, Chen H, Lin YJ** (2015). Novel green tissue-specific synthetic promoters and *cis*-regulatory elements in rice. *Sci Rep* **5**, 18256.
- Wei B, Jolma A, Sahu B, Orre LM, Zhong F, Zhu FJ, Kivioja T, Sur I, Lehtiö J, Taipale M, Taipale J** (2018). A protein activity assay to measure global transcription factor activity reveals determinants of chromatin accessibility. *Nat Biotechnol* **36**, 521–529.
- Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, Zheng H, Goity A, van Bakel H, Lozano JC, Galli M, Lewsey MG, Huang EY, Mukherjee T, Chen XT, Reece-Hoyes JS, Govindarajan S, Shaulsky G, Walhout AJM, Bouget FY, Ratsch G, Larrondo LF, Ecker JR, Hughes TR** (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443.
- Wen CJ, Yuan Z, Zhang XT, Chen H, Luo L, Li WY, Li T, Ma NN, Mao F, Lin DM, Lin ZX, Lin CT, Xu TD, Lü PT, Lin JC, Zhu FJ** (2023). Sea-ATI unravels novel vocabularies of plant active cistrome. *Nucleic Acids Res* **51**, 11568–11583.
- Wu B, Meng JH, Liu HB, Mao DH, Yin HR, Zhang ZY, Zhou XC, Zhang B, Sherif A, Liu HY, Li XH, Xiao JH, Yan WH, Wang L, Li XW, Chen W, Xie WB, Yin P, Zhang QF, Xing YZ** (2023). Suppressing a phosphohydrolase of cytokinin nucleotide enhances grain yield in rice. *Nat Genet* **55**, 1381–1389.
- Xue GP, Rae AL, White RG, Drenth J, Richardson T, McIntyre CL** (2016). A strong root-specific expression system for stable transgene expression in bread wheat. *Plant Cell Rep* **35**, 469–481.
- Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T** (2007). Identification of plant promoter constituents by analysis of local distribution of shorts equences. *BMC Genom* **8**, 67.
- Yamamoto YY, Obokata J** (2008). PPDB: a plant promoter database. *Nucleic Acids Res* **36**, D977–D981.
- Yang MM, Zhang WW, Ji SY, Cao PH, Chen YL, Zhao X** (2013). Generation of an artificial double promoter for pro-

- tein expression in *Bacillus subtilis* through a promoter trap system. *PLoS One* **8**, e56321.
- Yang YL, Lee JH, Poindexter MR, Shao YH, Liu WS, Lenaghan SC, Ahkami AH, Blumwald E, Stewart CN Jr** (2021). Rational design and testing of abiotic stress-inducible synthetic promoters from poplar *cis*-regulatory elements. *Plant Biotechnol J* **19**, 1354–1369.
- Yang ZF, Yan HD, Wang JP, Nie G, Feng GY, Xu XH, Li DD, Huang LK, Zhang XQ** (2022). DNA hypermethylation promotes the flowering of orchard grass during vernalization. *Plant Physiol* **190**, 1490–1505.
- Yasmeen E, Wang J, Riaz M, Zhang LD, Zuo KJ** (2023). Designing artificial synthetic promoters for accurate, smart, and versatile gene expression in plants. *Plant Commun* **4**, 100558.
- Yuan X, Ma KF, Zhang M, Wang J, Zhang QX** (2021). Integration of transcriptome and methylome analyses provides insight into the pathway of floral scent biosynthesis in *Prunus mume*. *Front Genet* **12**, 779557.
- Zhang CX, Gai Y, Wang WQ, Zhu YY, Chen XM, Jiang XN** (2008). Construction and analysis of a plant transformation binary vector pBDGG harboring a bi-directional promoter fusing dual visible reporter genes. *J Genet Genomics* **35**, 245–249.
- Zhang F, Lu CF, Qi S, Dai SL** (2022). Difference analysis of *CYC2*-like genes expression and DNA methylation between the two types of florets in *Chrysanthemum lavandulifolium*. *J Plant Growth Regul* **41**, 1316–1330.
- Zhang PC, Wang HC, Xu HW, Wei L, Liu LY, Hu ZR, Wang XW** (2023). Deep flanking sequence engineering for efficient promoter design using DeepSEED. *Nat Commun* **14**, 6309.
- Zhou JP, Liu GQ, Zhao YX, Zhang R, Tang X, Li L, Jia XY, Guo YC, Wu YC, Han YS, Bao Y, He Y, Han QQ, Yang H, Zheng XL, Qi YP, Zhang T, Zhang Y** (2023). An efficient CRISPR-Cas12a promoter editing system for crop improvement. *Nat Plants* **9**, 588–604.
- Zrimec J, Börlin CS, Buric F, Muhammad AS, Chen R, Siewers V, Verende IV, Nielsen J, Töpe IM, Zeleznik A** (2020). Deep learning suggests that gene expression is encoded in all parts of a co-evolving interacting gene regulatory structure. *Nat Commun* **11**, 6141.

Advances in Identification and Synthesis of Promoter Elements in Higher Plants

Xiang Song[†], Luyao Wang[†], Boxiao Fu, Shuangda Li, Yuanyuan Wei, Yan Hong^{*}, Silan Dai^{*}

Beijing Key Laboratory of Ornamental Plants Germplasm Innovation and Molecular Breeding, National Engineering Research Center for Floriculture, Beijing Laboratory of Urban and Rural Ecological Environment, Key Laboratory of Genetics and Breeding in Forest Tree and Ornamental Plants of Ministry of Education, School of Landscape and Architecture, Beijing Forestry University, Beijing 100083, China

Abstract Promoter is an indispensable regulatory sequence for driving gene expression in higher plants. Different promoter elements cause diverse driving efficiency and space-time specificity. Identifying the structures and functions of promoter elements contributes to a better understanding of the growth and development, multi-stress tolerance, and evolution of plants. With the development of high-throughput sequencing technologies, artificial intelligence and synthetic biology, the techniques for identifying *cis*-acting elements and constructing artificial biological components that meet the design requirements has gradually emerged, providing a foundation for efficient, precise, and diverse gene regulation in molecular breeding. This article targets on the application of promoter reconstruction in molecular design, introducing the detailed structure and function of higher plant promoters and the methods of *cis*-acting element identification. We summarized a total of 174 inducible, tissue-specific promoter elements in 27 categories and their applications on artificial modification and synthesis. At the end, we proposed the future directions and methods of the promoter designs. This review will be helpful for the further functional analyses of promoters in higher plants and their applications on molecular design breeding.

Key words promoter, *cis*-acting elements, inducible, tissue-specificity, artificial synthetic promote

Song X, Wang LY, Fu BX, Li SD, Wei YY, Hong Y, Dai SL (2024). Advances in identification and synthesis of promoter elements in higher plants. *Chin Bull Bot* **59**, 691–708.

[†] These authors contributed equally to this paper

^{*} Authors for correspondence. E-mail: hongy@bjfu.edu.cn; silandai@sina.com

(责任编辑: 孙冬花)

通讯作者/团队简介

戴思兰, 教授, 从事观赏植物遗传育种研究30余年, 尤其是对菊花历史文化, 品种资源收集、整理和评价到花色、花型、开花期和抗逆性等观赏品质形成的遗传调控机理, 菊花优异新种质创制及产业化栽培技术等进行了全面研究。主持国家自然科学基金项目等30余项国家和省部级科研项目。获国家林业局认定的科技成果4项。主持编写国家级高等教育规划教材, 发表学术论文350余篇, 出版专著《中国菊花全书》和《百菊图话》。

洪艳, 副教授, 硕士生导师, 以观赏植物菊花为研究对象, 对其观赏性状形成的分子机理等科学问题及育种和生产中的实际应用问题展开研究。主持国家自然科学基金面上项目等国家级和省部级科研项目4项, 以第一作者/通讯作者身份累计发表学术论文10余篇、授权国家专利1项。曾获首都绿化美化先进个人, 北京高校教师教学创新大赛二等奖, 北京市优秀本科生毕业论文指导教师等荣誉。