Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Review

Sourcing DNA parts for synthetic biology applications in plants

Katie Vollen*, Chengsong Zhao*, Jose M Alonso and Anna N Stepanova

Transgenic approaches are now standard in plant biology research aiming to characterize gene function or improve crops. Recent advances in DNA synthesis and assembly make constructing transgenes a routine task. What remains nontrivial is the selection of the DNA parts and optimization of the transgene design. Early career researchers and seasoned molecular biologists alike often face difficult decisions on what promoter or terminator to use, what tag to include, and where to place it. This review aims to inform about the current approaches being employed to identify and characterize DNA parts with the desired functionalities and give general advice on basic construct design. Furthermore, we hope to share the excitement about new experimental and computational tools being developed in this field.

Address
Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC 27695, USA

Corresponding author: Stepanova, Anna N (atstepan@ncsu.edu)
* Equal contribution.

Current Opinion in Biotechnology 2024, **87**:103140

This review comes from a themed issue on **Plant Biotechnology**

Edited by **Andrew D Hanson** and **Catalin Voiniciuc**

Available online xxxx

https://doi.org/10.1016/j.copbio.2024.103140

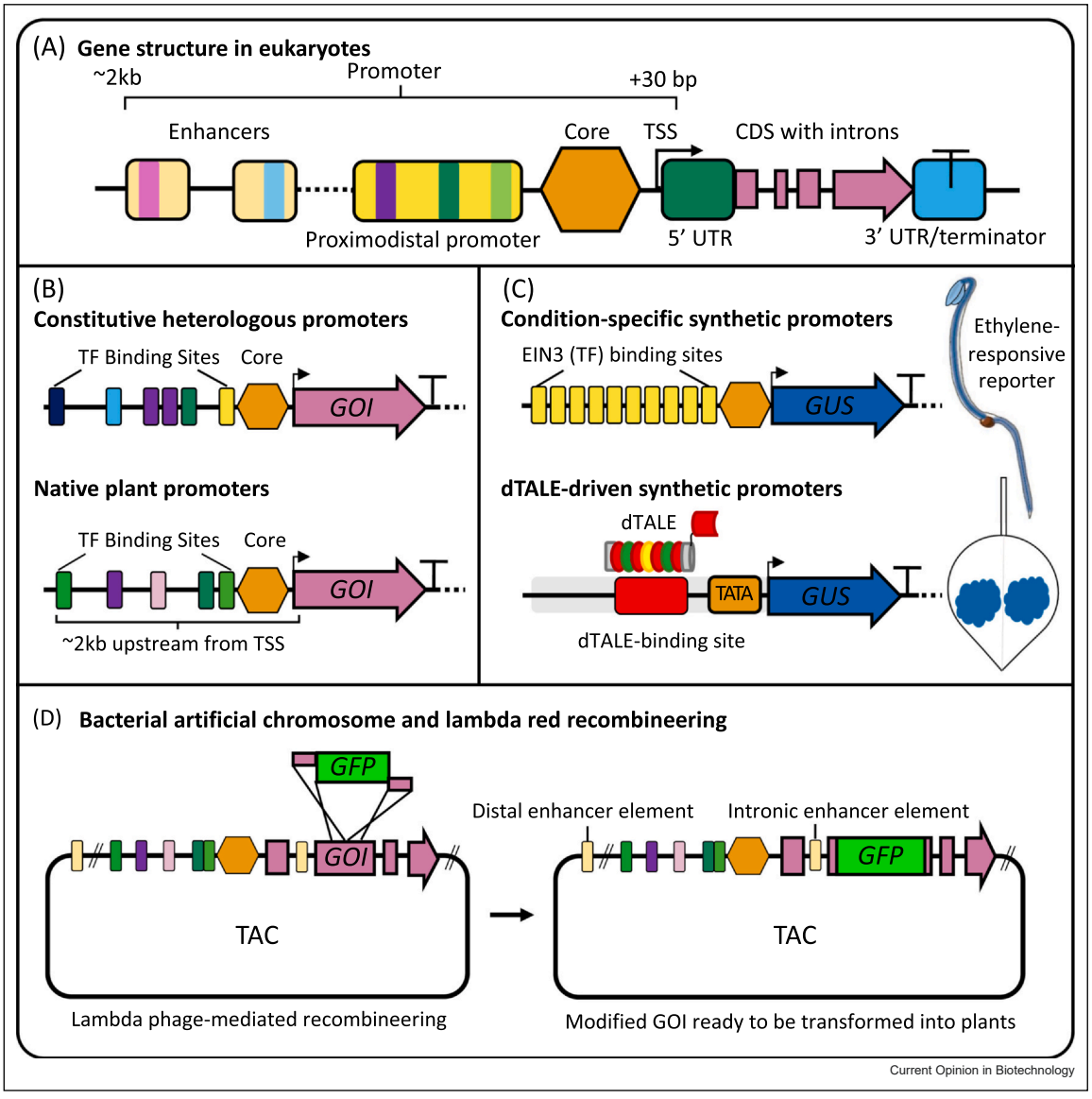
Introduction

Since RNA polymerase II (RNAPII) was first purified from sea urchin and rat cells more than 50 years ago [1], major progress has been made in understanding transcription of protein-coding genes in eukaryotes and identifying the molecular machinery orchestrating the processes of transcription initiation, promoter-proximal pausing, elongation, and termination [2], as well as controlling mRNA translation, mRNA and protein structure, stability, trafficking, and localization [3,4]. Although the level of our comprehension of these processes in plant systems has lagged behind that in select animal and yeast models, the cumulative body of gene expression regulation literature can be leveraged to develop a set of guidelines on how to best build synthetic plant genes. In very basic terms, to ensure robust expression of a transgene of interest, it is important to identify and include the appropriate DNA sequences for recruiting the transcription and translation machinery or for conferring the desired level of RNA or protein stability. In this review article, we discuss the key components of protein-coding synthetic genes — promoters and enhancers, 5' and 3' untranslated regions (UTRs), coding sequences (CDSs), and protein fusion tags — and how these are typically identified and employed to optimize transgene expression and stability *in planta* with an emphasis on the increasing role of computational tools.

Promoters and enhancers

When deciding on the optimal way to express a gene of interest (GOI), the choice between a native or synthetic promoter needs to be made (Figure 1a–c). If the goal of one's study is to capture the full expression pattern of a plant gene, for example, to visualize the sites of gene activity via a reporter fusion or demonstrate the ability of a tagged GOI to complement a loss-of-function mutant phenotype, the traditional choice of a 2 kb promoter region driving a cDNA–reporter fusion may not be adequate due to the lack of distal or intron-localized enhancers. For example, transgenic complementation lines containing *FLOWERING LOCUS T (FT)* cDNA driven by *FT* promoter fragments of 5.7 kb or more were able to rescue the mutant *ft-10*, whereas those with a 4.0 kb *FT* promoter could not [5]. This finding suggests that the more distal region of the promoter may be necessary for full gene activity. Interestingly, a reporter–cDNA fusion of the auxin biosynthesis gene *TAA1* driven by a long (10 kb) native promoter, *TAA1p:YPet-TAA1*$_{cDNA}$, was also not able to fully revert the respective *taa1 tar1* mutant root defects, whereas an equivalent construct with introns, *TAA1p:YPet-TAA1*$_{gDNA}$, could [6]. Consistently, the cDNA construct lacked a critical expression domain in the quiescent center of the root, implying that *TAA1* introns may harbor an intronic enhancer critical to this gene's expression. Thus, for recapitulating the full expression

Figure 1



Transgene design. **(a)** A typical eukaryotic gene. The gene consists of a promoter, 5′ UTR, coding region (often with introns), and the 3′/terminator region. Promoters are sets of regulatory sequences in the 5′ end of a gene that in plants typically extend from two or more kilobases (kbs) upstream of a gene's transcription start site (TSS, black arrow) to about 30 bp downstream of TSS. The sequences immediately upstream and downstream of the TSS recruit RNAPII and general transcription factors (TFs) and are referred to as the core or minimal promoter (orange hexagon). The sequences upstream of the core promoter that bind gene-specific TFs are often referred to as proximodistal promoters but may also be viewed as enhancers as some of these may be orientation and position independent. Some enhancers can also be found further upstream or downstream of the genes they control or be harbored in genes' introns. **(b)** Natural promoters. The proximodistal region of natural promoters contains TF-binding sites that impart tissue-specific and developmental regulation of gene expression. Constitutive heterologous promoters, such as the well-characterized *CaMV 35S* promoter, are often used to drive high levels of transgene expression in the whole plant. Native plant promoters more accurately reflect the tissue specificity and developmental context in which a native gene is expressed, though sometimes at insufficient expression levels to achieve the goals of a transgene. In an attempt to capture the native TF-binding sites that regulate native gene expression, approximately 2 kb of sequence upstream of the TSS is typically used as the promoter sequence. **(c)** Synthetic promoters. These can be created by placing TF-binding sites upstream of a core promoter. TF-specific binding sites, such as the EIN3 TF-binding site, can be used to confer condition specificity on gene expression, as is the case for the EIN3-binding site containing *EBSp:GUS* reporter for ethylene. Additionally, synthetic promoter systems rely on a synthetic TF made of a programmable DNA-binding domain (such as dTALE) fused with a transcriptional activation domain (red flag). Upstream of a natural minimal promoter or of a synthetic TATA box containing core promoter sequence are programmable DNA-binding domain-binding site(s) (dTALE-binding site) surrounded by neutral DNA sequence that does not contain any known TF-binding sites. **(d)** BAC and lambda red recombineering. Plant transformation–ready BACs (TACs) carry large pieces of plant gDNA containing all necessary regulatory sequences to achieve native plant gene expression patterns upon transformation into plants. TACs can be modified to include a reporter gene using recombineering tools. Exogenous DNA containing a reporter gene (*GFP*) and homology arms (pink overhangs) can be introduced into an exon of the GOI harbored by the TAC using lambda phage–mediated homologous recombination in recombineering strains of *Escherichia coli*. Following the incorporation of a reporter into the GOI, the TAC can be transformed into plants using standard *Agrobacterium*-mediated transformation methods.

pattern of a GOI, the detailed knowledge of its enhancer regions may be necessary. In the absence of an efficient knock-in approach in plants [7], this remains a major bottleneck in construct design.

To experimentally define native proximal promoters and distal enhancer sequences important for transcription, several approaches can be employed, from classical promoter deletions for a specific GOI to modern high-throughput molecular and computational tools at the whole-genome level [8–11]. With the adoption of next-generation sequencing, genome-wide approaches have been implemented to pinpoint *cis*-regulatory elements that affect gene expression. In plants, DNase I hypersensitive site sequencing (DNase-seq), chromatin immunoprecipitation sequencing using anti-histone H3K9ac antibodies (H3K9ac ChIP-seq), bisulfite sequencing (BS-seq), the assay for transposase-accessible chromatin using sequencing (ATAC-seq), and self-transcribing active regulatory region sequencing (STARR-seq) have been employed to define accessible chromatin regions (ACRs) and putative enhancers [12–15]. A key takeaway from these studies is that there may be more putative enhancers in the genome than there are genes, and thus, many (if not all) genes may be regulated by distal elements, many of which are tissue specific or condition specific. Therefore, to truthfully recapitulate a gene's activity pattern, long-range, distally located enhancers may need to be included in a construct along with proximal and core promoter and intron sequences. However, on a gene-by-gene basis, it is still very challenging to translate the information from these whole-genome ACR studies into the design of individual constructs, and thus, multiple constructs often need to be made to explore and leverage native regulation of individual genes of interest [11].

One possible alternative to the still technically problematic knock-in strategy [7] is to build much larger, for example, 100 kb, constructs that should contain most or all of the regulatory sequences, for example, in the pseudogenomic context of a bacterial artificial chromosome (BAC) harboring a large piece of plant genomic DNA (gDNA) [16]. In that scenario, a reporter is integrated into the GOI carried by a transformable BAC clone via recombineering, that is, phage-protein-assisted (aka lambda red) homologous recombination in bacteria (Figure 1d) [17]. The BAC can be trimmed (also via recombineering) to preserve only the desired regions upstream and downstream of the gene and transformed into plants via standard transformation methods [18–21].

Another way around not knowing all of the regulatory sequences of a GOI is to use well-characterized constitutive heterologous promoters such as *35S* from the Cauliflower mosaic virus (CaMV), *FMV* from Figwort mosaic virus, *CmYLCV* from Cestrum yellow leaf curling virus, and *nopaline synthase* (*NOS*) promoter from *Agrobacterium tumefaciens* to maximize gene expression in plants [8–11]. For example, for the genetic constructs in most commercially available genetically modified plants, the *35S* promoter is used. Currently, this is the most common approach in plant sciences [22]. However, these strong ubiquitous promoters are not well suited when precise spatiotemporal expression patterns are required, as in many cases, overexpression of a gene may or may not be able to complement a respective mutant in full, be toxic to plants, or provide misleading functional information due to physiologically irrelevant levels or distribution of the resulting protein [23–25]. Likewise, the analysis of overexpressed gene-reporter fusions can give incorrect subcellular protein localization patterns upon overwhelming the protein trafficking machinery [26].

The third alternative for driving gene expression in desired patterns or at preferred levels is to utilize synthetic promoters where a known core promoter is preceded by native, heterologous, or synthetic proximodistal sequences (Figure 1c). These sequences serve the purpose of recruiting tissue-, stage-, or condition-specific transcription factors (TFs) to turn the GOI on in a controllable spatiotemporal manner. For example, to make a reporter gene responsive to the hormone ethylene in *Arabidopsis thaliana*, multiple copies of a binding site for a transcriptional master regulator of ethylene signaling, EIN3, were stacked upstream of a *35S* core promoter driving a histochemical marker *GUS* [27]. Likewise, photosynthetic tissue–specific and drought-inducible promoters were built in poplar [28]. The major advantage of well-designed synthetic promoters is that these can be developed to specifically recruit only the TF(s) of interest and thus have less background or leaky expression stemming from unrelated TF binding. For example, to temporally control transgene expression, inducible promoters can be generated by stacking the binding sites for synthetic TFs regulated by specific stimuli or chemical inputs. Common synthetic inducible promoter choices in plants include heat-, steroid-, ethanol-, copper-, and light-responsive systems [29].

To build a functional synthetic promoter regulated by native TFs, in theory, an enhancer-like sequence from any source can be placed upstream of a core promoter, as demonstrated by Jores et al. [30] via STARR-seq in agroinfiltrated *Nicotiana benthamiana* leaves for enhancer sequences sourced from the *35S* promoter, wheat and pea *CHLOROPHYLL A-B BINDING PROTEIN* genes *CAB-1* and *AB80*, and a pea *RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE SMALL SUBUNIT* gene *rbcS-E9* placed upstream of the *35S* core. Since promoter testing in this study was limited to leaf agroinfiltration assays in *N. benthamiana*, it remains to be seen if these synthetic sequences are universally functional in different tissues and plant species.

Brückner et al. [31] set out to rationally design the proximodistal sequences. The library of synthetic promoters these researchers constructed contains an 18-bp-long *cis*-element corresponding to the binding site of a single designer transcription activator-like effector (dTALE) and a *TATA box* flanked by degenerate sequences (19 bp upstream and 43 bp downstream; Figure 1c). Based on the *GUS* reporter activity in transient assays in *N. benthamiana*, the expression level of the synthetic promoters was inferred to range from around 5% to almost 100% of the *35S* promoter, suggesting that this simple promoter architecture is a useful platform for plant synthetic promoter design. Cai et al. [32] replaced the 18-bp-long dTALE-binding site from the Brückner et al. [31] design with various computationally developed *cis*-regulatory elements, as well as swapped the original 43-bp-long degenerate sequence for an unnamed core promoter that includes the transcription start site (TSS). A good correlation ($R^2 = 0.7076$) was found in a *Luciferase* (*Luc*) reporter assay between computationally predicted gene expression levels and the actual values determined experimentally for 24 *MinSyn* promoter sequences that were randomly selected from a library of 1000 constitutive *MinSyn*s.

Much like the many options of proximodistal sequences, the choice of core promoters is not limited by a handful of well-characterized promoters such as *35S*. Jores et al. [33] used STARR-seq to measure the strengths of 18 329 Arabidopsis, 34 415 maize, and 27 094 sorghum core promoters in transient assays in the context of synthetic genes containing histone *H3* 5'UTR sequences placed upstream of a barcoded green fluorescent protein gene (*GFP*). The presence of a *TATA box*, promoter GC content, and promoter-proximal TF-binding sites were all found to affect promoter strength, with the *TATA box* positioning ~30–40 bp upstream of the TSS being the most critical feature determining the level of gene expression. Furthermore, Jores et al. [33] designed novel synthetic promoters by generating 170-bp-long random sequences with nucleotide frequencies similar to an average Arabidopsis or maize promoter. These sequences were further modified by introducing a *TATA box* (TATAAATA) at position 133–140, a *Y patch* at position 147–154, and/or an *Initiator* element (yyyyTCAyyyy, where y indicates a change of A to T or G to C) at positions 147–154. The strongest synthetic core promoters Jores et al. [33] developed could reach activities comparable to the *35S* minimal promoter (–46 to +5 relative to the TSS), indicating that rationally designing synthetic core promoters of varying strength is possible. In addition, these researchers took a machine learning approach using a convolutional neural network to predict promoter strengths and used *in silico* evolution to design synthetic promoters with increased activity. After 3–10 rounds of sequence evolution, a prominent increase in promoter strength was observed. This work provides a great resource for expanding the synthetic core promoter options. It is, however, still necessary to validate these sequences in the context of transgenes to show that the enhanced activity of the new promoters results in greater protein expression, as higher levels of transgene activity at a transcriptional level do not always lead to increased protein levels in all tissues and conditions, presumably due to processes such as translational regulation via 5' UTRs that overlap with core promoters and post-transcriptional gene silencing [34].

Despite the growing arsenal of both natural and synthetic promoter elements and of our understanding of the grammar rules governing their activities, the design of promoters with prescribed spatiotemporal expression and strength characteristics remains extremely challenging. Although still in its early days, synthetic biology approaches based on Boolean logic are starting to be developed in plants that hold the promise of generating novel expression patterns using complex computational combinations of existing promoter elements [35–37].

To summarize, there are multiple choices of promoters for driving a transgene in plants. With the implementation of genome-wide studies in plants and the adoption of synthetic biology methods in species beyond Arabidopsis, we anticipate that the use of synthetic promoters can provide an unprecedented level of gene regulation and exceed the strength of standard constitutive promoters such as *35S*.

5' untranslated region/leader sequences

Traditionally, when making a construct for a GOI, a promoter is often fused directly to the coding region without including a 5'UTR sequence. However, a number of studies suggest that a 5'UTRs can have a profound effect on the expression of a GOI at both transcriptional and post-transcriptional level [38]. The inclusion of a 5'UTR can be used to enhance or reduce the activity of a GOI by controlling processes, such as transcription initiation (e.g. due to the inclusion or omission of downstream core promoter elements and TF-binding sites), transcription elongation (e.g. by affecting RNA structure and RNA Pol II promoter-proximal pausing), RNA stability (e.g. due to the presence of RNA destabilization cis-elements), and translation efficiency (e.g. by containing inhibitory upstream open reading frames [uORFs] or stable hairpins or by harboring internal ribosome entry sites) [39]. The two best-known translational enhancers are the 5' UTRs of the Tobacco mosaic virus (TMV) RNA and Alfalfa mosaic virus (AMV) RNA4 known in the field as omega and AMV leader sequences, respectively [40,41]. Besides these TMV and AMV sequences, other plant viral RNAs may also harbor efficient translational enhancers in their 5' ends [42]. In addition, a few plant-sourced 5' UTRs

have also been reported to enhance gene expression [43–45].

To expand the very limited toolbox of characterized 5' UTR/leader sequences, De Amicis et al. [46] leveraged the structure of the omega leader sequence [consisting of three octamer direct repeats of ACAAUUAC and a poly(CAA) region] to design the first synthetic 78 bp 5' UTR. These researchers incorporated the 5 bp 5' UTR of the *35S* transcribed sequence (+1 to +5 relative to TSS) and a cytosine and thymine (CT)-rich region into the omega backbone (containing a single octamer and nine CAA repeats) in the context of the pSTART vector. The resulting synthetic 5' UTR was found to be 8.6- to 12.5-fold stronger than the *gusA* leader in the pBI121 vector at supporting *gusA* reporter expression. Kanoria and Burma [47] developed a small synthetic 5' UTR (28 bp in length), *synJ*, which contained only the first 5 bp of the 5' UTR of *35S* transcript (+1 to +5 relative to TSS) and a near-perfect Kozak translation initiation context. These scientists found that *synJ* was equivalent to the omega leader sequence at enhancing *GUS* gene expression in transformed cotton callus and in the leaves of transgenic tobacco (*Nicotiana tabacum*) plants relative to the 5' UTRs in pBI121 and pRT100 vectors. Tanaka et al. [48] generated artificial synthetic 5' UTRs in rice (*Oryza sativa*) using an efficient machine learning model, named 'R-STEINER', that could predict the amount of protein of interest (POI) with a correlation coefficient of 0.89. Finally, Peyret et al. [49] used rational design to generate four synthetic 5' UTRs with desirable characteristics (such as low GC content, low secondary structure, repeats of an AAC motif, and a strong Kozak consensus sequence). All synthetic 5' UTRs these researchers created were superior in their performance relative to the control construct that harbored a modified 5' UTR from the bipartite Comovirus cowpea mosaic virus (CPMV) RNA-2 [50].

In light of these studies convincingly demonstrating the ability of 5' UTR sequences to improve gene expression, it is advisable to include a well-studied viral, endogenous, or synthetic leader downstream of a core promoter immediately upstream of the CDS. It is, however, risky to incorporate an uncharacterized 5' UTR as it may harbor negative regulatory sequences such as uORFs [51].

Coding regions

In synthetic construct designs, the CDS is inherently the most variable part. While all constructs require a promoter, 5'UTR, and terminator sequence to ensure that the CDS is transcribed and translated in the correct cellular context, the user-defined CDS provides a readout or function to the construct (Figure 2). In CDS design, one must carefully consider what questions the experiment aims to answer to inform design decisions. For example, if one is interested in purifying their protein of interest (POI), they should consider expressing a tagged version. If one wants to capture the native pattern of expression and all possible splicing variants of a GOI, its intron-containing version should be used. Given that many CDS inputs are possible, this section will focus on common design considerations and applications for expressing a GOI, from monitoring gene activity to transgene sequence optimization.

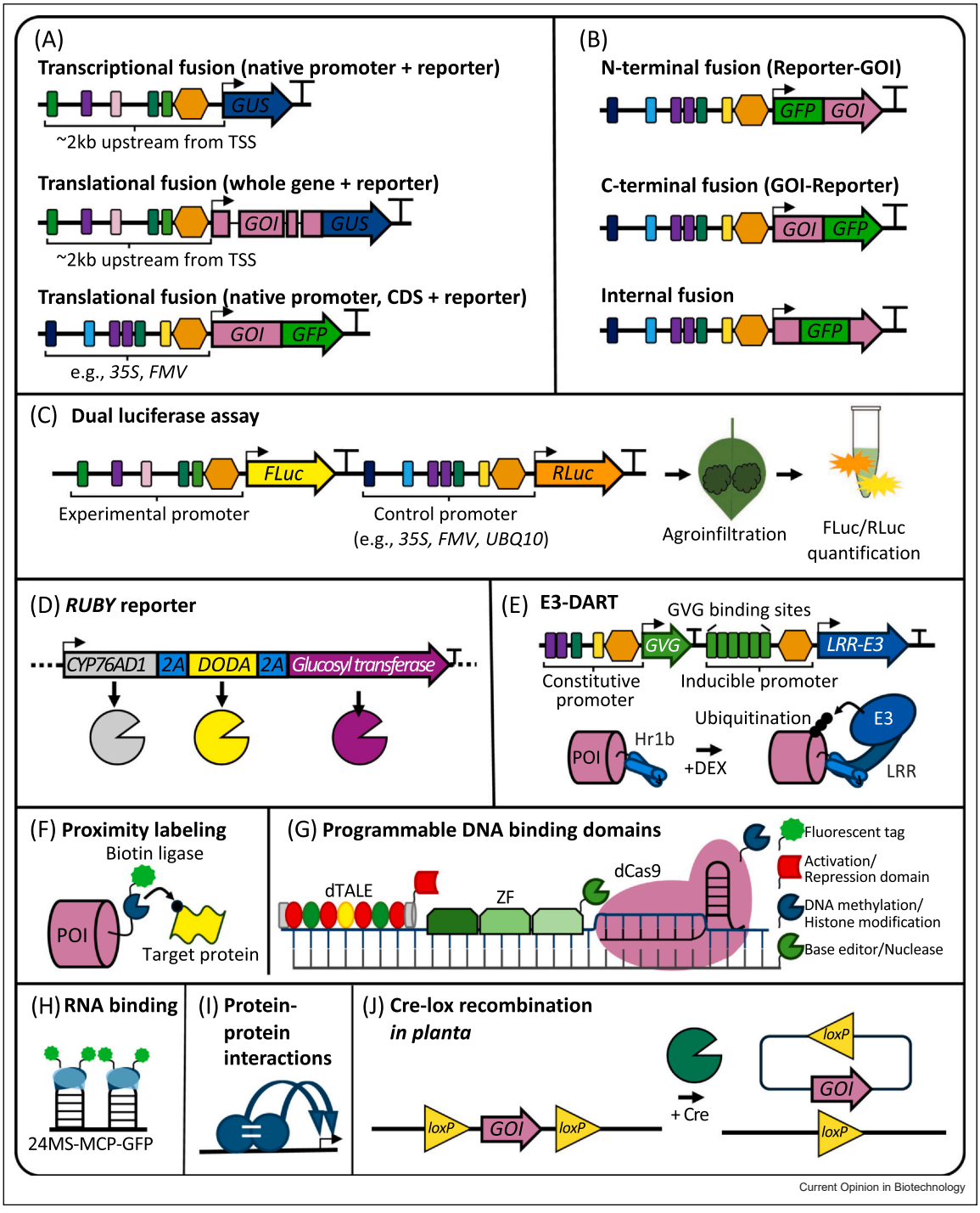
Reporter tags

Perhaps, the most common application of transgenes involves fusing a GOI to a reporter gene to visualize when and where the gene is active. Reporters can be fused to a gene's promoter to study GOI's transcription or in frame with the full or partial CDS to monitor POI levels and distribution at cellular and subcellular levels (Figure 2a). Histochemical markers such as beta-glucuronidase (GUS), fluorescent proteins (FPs) such as GFP, luminescent proteins such as Luc, and, more recently, a colorimetric reporter RUBY are widely adopted in plant research [52,53].

GUS, the oldest of the reporters used in plants, converts a colorless substrate, X-gluc, to an easy-to-see blue product. Despite some disadvantages of GUS reporters (such as relatively long protein half-life, lack of cellular resolution due to diffusion of cleaved product, and false-positive signal caused by native GUS activity) [52,54,55], the high sensitivity of GUS (with every molecule of the enzyme hydrolyzing multiple molecules of the substrate) makes this reporter a popular tool widely employed by the plant biology community. For example, Lauressergues et al. [56] utilized GUS transcriptional fusions to monitor the expression of micropeptides for nine different Arabidopsis pre-microRNAs . The GUS fusions used in this experiment demonstrated that the first adenine, thymine and guanine (ATG) start codon of each gene tested was sufficient to initiate translation, as indicated by GUS detection.

GFP and other FPs are, perhaps, the most versatile and ubiquitous reporters of gene expression. Applications of such reporters are vast and have been reviewed [57,58]. FPs can be used in transcriptional or translational reporters to provide tissue-level and subcellular localization information in a noninvasive or destructive manner (Figure 2a). The expression levels and pattern of a GOI can be inferred by using the intensity of FP expression as a proxy for transcriptional activity or protein abundance. For example, Wang et al. [59] used live-cell confocal microscopy imaging to monitor programmed cell death in Arabidopsis root cap cells. Researchers examined the breakdown of the nuclear envelope, ER membrane, and mitochondria by observing nuclear-, ER- and mitochondria-localized FP diffusion into the

Figure 2



(caption on next page)

Transgene applications. **(a)** Transcriptional and translational reporter fusions. A typical transcriptional fusion construct uses ~2 kb of sequence upstream of the TSS to capture most of the native elements controlling gene expression. The CDS is composed of a reporter (*GUS* in this figure), followed by a terminator, typically a well-characterized terminator such as *Tnos* from *Agrobacterium*, though the native terminator is sometimes used. The goal of a transcriptional fusion is to observe the patterns of gene transcription, that is, where and when RNA is produced. A translational fusion, on the other hand, serves the purpose of analyzing protein levels and distribution at the cellular and subcellular levels. Translational fusions may use a native promoter and terminator to capture native expression levels and patterns. In a translational fusion, a full-length or truncated CDS is fused in frame with a reporter gene to capture protein localization. Given that introns can harbor regulatory elements, the CDS of the GOI is depicted as broken up to represent the use of gDNA instead of cDNA. Sometimes, a translational fusion does not provide sufficient protein signal to detect subcellular localization. In those cases, one strategy to increase reporter signal is to use a strong constitutive promoter such as *35S* or *FMV* to drive the construct, though such reporter will not be suitable for monitoring tissue-specific protein expression patterns. **(b)** Varying tag location. When constructing a translational fusion, one must consider where the reporter tag will be fused with the GOI. N-terminal fusions involve the reporter CDS preceding the GOI CDS, whereas C-terminal fusions have the GOI CDS preceding the reporter CDS. N- and C- terminal fusions are the most common options when tagging a POI; however, one can also create an internal fusion by inserting the reporter CDS within the GOI CDS. Protein structure predictions are commonly used to infer the safe sites for internal reporter integration, but as with N- and C-terminal fusions, the functionality of the construct needs to be tested via respective loss-of-function mutant complementation. **(c)** Dual luciferase assays. These make use of two distinct enzymes, typically firefly Luciferase (*FLuc*) and Renilla Luciferase (*RLuc*), and can provide a high-throughput means to quantify the impact of regulatory elements on gene expression. The figure illustrates a tobacco leaf infiltration experiment in which promoter elements are the independent variable, and the dependent variable is *FLuc* expression. Given *FLuc* expression will vary between leaves and experiments, it is important to include *RLuc* under a constitutive promoter as an internal experimental control for normalization. 3–5 days following infiltration, leaf tissue is harvested and ground for Luc quantification with a fluorometer. **(d)** Synthetic *RUBY* gene structure. RUBY is a reporter composed of a single transcriptional unit that combines three enzymes for betalain (pigment) synthesis. The three enzyme genes are expressed under a single promoter, connected together by P2A peptides that induce ribosomal skipping, resulting in three separate proteins following translation. The resulting pigment can be visualized by eye. **(e)** E3-DART. This is an inducible protein degradation system, which involves fusion of a POI (pink) with the Hr1b domain (cyan) of the human target PKN1. The Hr1b domain can be bound by the LRR (blue) and ubiquitinated (black circles) by the E3 ligase domain of *Salmonella*-secreted protein H1. The construct depicted contains a constitutive promoter driving expression of GVG, a TF that binds dexamethasone (DEX). Upon DEX binding, GVG translocates to the nucleus and interacts with the promoter driving LRR-E3 expression. Thus, in the presence of DEX, LRR-E3 is expressed, while in the absence of DEX, GVG remains sequestered in the cytosol and LRR-E3 transcription is turned off. **(f)** Proximity labeling. This is a method used to study protein–protein interactions and organelle proteomes. In proximity labeling experiments, a POI (pink) is fused to a biotin ligase (blue, e.g. TurboID) and an FP reporter (green) to visualize protein localization and assess protein expression. The biotin ligase attaches a biotin (black) to proteins (yellow) within a given radius (~35 nm for TurboID), thus labeling all proteins within certain proximity to the POI. Biotin-labeled proteins are then affinity purified using streptavidin beads for mass spectrometry analysis. **(g)** Programmable DNA-binding domains. dTALEs, ZFs, and dCas9 can be fused to a variety of effector domains (FP tags, transcriptional activation or repression domains, DNA methylation or histone modification enzymes, nucleases, or base editors) to impart a function on the DNA-binding domain. **(h)** RNA aptamers. MS2 and other aptamers (striped structures) are leveraged along with their associated RNA-binding proteins (blue) tagged with an FP tag (green) to visualize target RNA in the cell. Binding of the FP fusion to the aptamer can be detected by monitoring protein fluorescence. Other effector proteins can be attached to aptamer-binding proteins, such as TurboID to capture RNA-protein interactions. **(i)** Protein–protein interactions. These can be leveraged to promote TF (blue circles) co-operativity by fusing protein interaction domains (white lines) to TFs to maximize their recruitment to target DNA (black line). **(j)** Cre-lox recombination. Cre recombinase target sites (LoxP, yellow) are integrated into the genome via a transgene. Site-specific recombinase Cre (green) recognizes them and induces precise genome modifications. Two common applications of this technology are controlled excision of DNA fragments, as shown in the figure, or targeted insertion of DNA from a donor DNA fragment into the genome.

cytoplasm. To observe dynamic cellular processes, an FP modified with decreased protein half-life can be used to increase the time resolution of a reporter [52]. A challenge associated with using FP fusions in plant tissues is the autofluorescent signal from plant cell walls and chloroplasts that overlaps with FP spectra [52]. Though modern microscopy techniques can filter autofluorescence to some degree, plant biologists are primarily limited to FPs that emit in red, green, and yellow spectra.

In FP fusion design, it is important to consider how a protein fusion might impact protein function, folding, or localization. There are many computational programs that can predict cellular localization [60,61] and protein structure [62,63]. These tools can be useful in deciding where to attach a FP tag. For example, if a protein is expected to localize to the chloroplast, an N-terminal fusion would likely disrupt proper localization. Given the effects of a given protein fusion are difficult to foresee, it is prudent to generate both N- and C-terminal fusion constructs when possible (Figure 2b). Internal FP fusions may also be helpful when protein structure is known or can be modeled [64].

Luc and other luminescent reporters are optimal for quantifying gene expression in the context of transcriptional or translational fusions. Firefly Luc (*FLuc*) is an enzyme that catalyzes the oxidative decarboxylation of its substrate, D-luciferin to oxyluciferin, releasing a flash of light at 560 nm that can be easily quantified using a fluorimeter [65]. Due to the lack of native luminescent molecules in plant tissue, luciferase-based reporter systems have high signal-to-noise ratios. High-throughput studies, particularly in *N. benthamiana* leaf infiltration experiments, benefit from the use of a dual luciferase system to create an internal control for *Agrobacterium* infection efficiency and leaf developmental differences that affect overall protein expression. The dual luciferase system takes advantage of two distinct luciferase molecules: *FLuc*, described above, and *Renilla Luc (RLuc)*, which involves the conversion of its substrate, coelenteraxine, to coelenteramide in an oxygen-dependent, ATP-independent manner, releasing light at

480 nm [66]. Typically, *FLuc* expression is driven by experimental regulatory elements, while *RLuc* expression is driven by a constitutive promoter, enabling ratiometric analysis of protein abundance (Figure 2c) [67]. Some of the drawbacks associated with Luc-based assays are lack of tissue- or cell-level resolution, high variability in signal between replicates, uneven substrate penetration into samples, and the need for specialized equipment to detect the luminescence signal. To reduce assay costs and remedy nonuniform substrate penetration, Khakhar et al. [68] and Mitiouchkina et al. [69] implemented a fungal autoluminescent Luc-based reporter system in plants that does not require substrate input. By applying directed evolution as well as random and consensus mutagenesis to the fungal bioluminescence pathway and screening orthologous genes from different species of bioluminescent fungi, Shakova et al. [70] further optimized the system. The resulting synthetic pathway was transformed into six diverse plant species, including Arabidopsis, petunia, poplar, tobacco (*N. benthamiana* and *N. tabacum*), and chrysanthemum, resulting in visibly glowing plants. When compared with traditional *FLuc* in plant tissue culture, the engineered autoluminescent *Luc* was an order of magnitude brighter without substrate input [70]. Further development of autoluminnescent *Luc* pathways combined plant and fungal genes to create a more compact autoluminescent reporter system that functions well in yeast, mammals, and plants [71].

The newest reporter type implemented in plants is *RUBY*, an artificial gene containing the sequences of three enzymes required for betalain biosynthesis, producing a bright red-violet pigment that can be seen with the naked eye (Figure 2d) [53,72]. The fact that RUBY generates a robust visible signal that, unlike GUS, does not require substrate infiltration resulted in a rapid adoption of this reporter in many plant species. For example, Wang et al. [73] applied the *RUBY* marker in haploid inducer lines in maize and tomato plants, enabling rapid identification of haploid progeny by eye. Due to the relative stability of betalain in plant cells, the *RUBY* reporter is not suitable for tracking dynamic processes.

Degradation tags

In some cases, it may be desirable to make the POI short lived, for example, to enable its rapid turnover upon removal of a stimulus to reset the state of the cell, by adding a destabilization tag (aka degron) to the POI. For example, Khakhar et al. (2018) [74] combined a synthetic Cas9-based transcription repressor with a highly sensitive auxin-induced degron domain. The authors demonstrated that this system could be leveraged to reprogram development in Arabidopsis upon targeting the auxin transporter, *PIN-FORMED1* [74]. The newest tool for targeted protein degradation in plants, E3-targeted Degradation of Plant Proteins (E3-DART), takes advantage of the E3 catalytic activity of *Salmonella*-secreted protein H1 (SspH1) and its association with the human target protein kinase N1 (PKN1). Following induction with the glucocorticoid system, a POI fused with the HR1b domain of PKN1 is targeted for rapid protein degradation by the SspH1 Leucine-Rich Repeat and E3 ligase domain [75] (Figure 2e).

Other protein tags and functional domains

Besides the aforementioned CDS options, additional tags have been developed to further expand the capabilities of a transgene (Figure 2f–i). For example, a variety of affinity purification tags can be used depending on the requirements of protein yield, level of nonspecific binding, size of affinity tag, position of tag (N- or C-terminus, as described above), or live detection needs (whether the protein needs to be visualized) [76]. Similarly, an array of localization signals to target proteins to different subcellular compartments, such as the nucleus, chloroplast, mitochondria, endoplasmic reticulum, Golgi apparatus, and plasma membrane, have also been developed and are commonly used in plants [77]. Viral ribosomal skipping peptides such as P2A and T2A allow for the co-expression of multiple proteins from a single transcript (Figure 2d) [72]. Proximity labeling tags such as TurboID enable the characterization of molecular interactions that occur in the cell (Figure 2f) [78]. Various other functional domains enable researchers to study transcriptional regulation, promote protein–DNA, protein–RNA, and protein–protein interactions, or confer enzymatic activity (Figure 2g–i) [79–82].

Codon optimization and intron inclusion

All organisms exhibit codon-usage bias or nonrandom use of codons that encode identical amino acids (synonymous codons). It is important to consider optimal codon use when introducing a heterologous gene to achieve high levels of gene expression as codon usage can impact translational efficiency and cotranslational protein folding [83–85]. There are a number of computational codon optimization tools available to aid in synthetic gene design in plants (e.g. OPTIMIZER, CodonWizard, etc.) [86].

Traditional gene complementation experiments and translational fusions use cDNA sequences in the CDS position instead of gDNA. As mentioned above, introns can host functional elements that affect transcription and splicing. Thus, when expressing native genes, it is generally recommended to use the gDNA sequence. Likewise, the inclusion of introns in heterologous genes can enhance gene expression at post-transcriptional levels [87,88]. In crop engineering, it is common to use the intron-containing 5' UTR of maize *polyubiquitin-1* (*Ubi-1*) gene, particularly in monocot species due to low

expression resulting from the *35S* promoter [89]. One of the major advantages of including an intron in the CDS is it will abolish any leaky protein expression in bacteria (because prokaryotes lack the spliceosomal machinery) and thus ensure that the POI will not have an effect on bacterial health. Additionally, *Agrobacterium* can express a GOI from T-DNA, creating false-positive reporter gene signals in leaf infiltration experiments unless an intron is included in the CDS [35].

One of the major drawbacks to including introns in CDS design is that the exact mechanism by which introns increase gene expression is not well understood, and there is no clear and reliable guidance on the inclusion of intron sequences for a given GOI. A webtool to aid in intron insertion in transgenes called Intronserter is available [90], but it has not yet been widely adopted by the plant community. Nonetheless, in one representative study in energy cane (a *Saccharum* spp. hybrid), using this tool to augment the sequence of a garden nasturtium (*Tropaeolum majus*) *DIACYLGLYCEROL ACYLTRANSFERASE* gene (that was codon-optimized for *Sorghum bicolor* and equipped with a 110 bp intron from another *Sorghum* gene) resulted in a sevenfold enhancement of transgene expression in energy cane [91]. In general, it is advisable to use intron sequences that have been validated in previous studies and to create a parallel construct that does not include introns.

Perhaps, the most illustrative recent example of the beneficial effects of codon optimization and introns in synthetic constructs comes from the genome-editing study in Arabidopsis, where the efficiency of Cas9-mediated editing was increased from 0 to 70%–100% following maize codon optimization and the introduction of 13 Arabidopsis introns in the protein CDS [92]. This study demonstrates the largely understudied potential for increasing heterologous protein expression through codon optimization and intron inclusion.

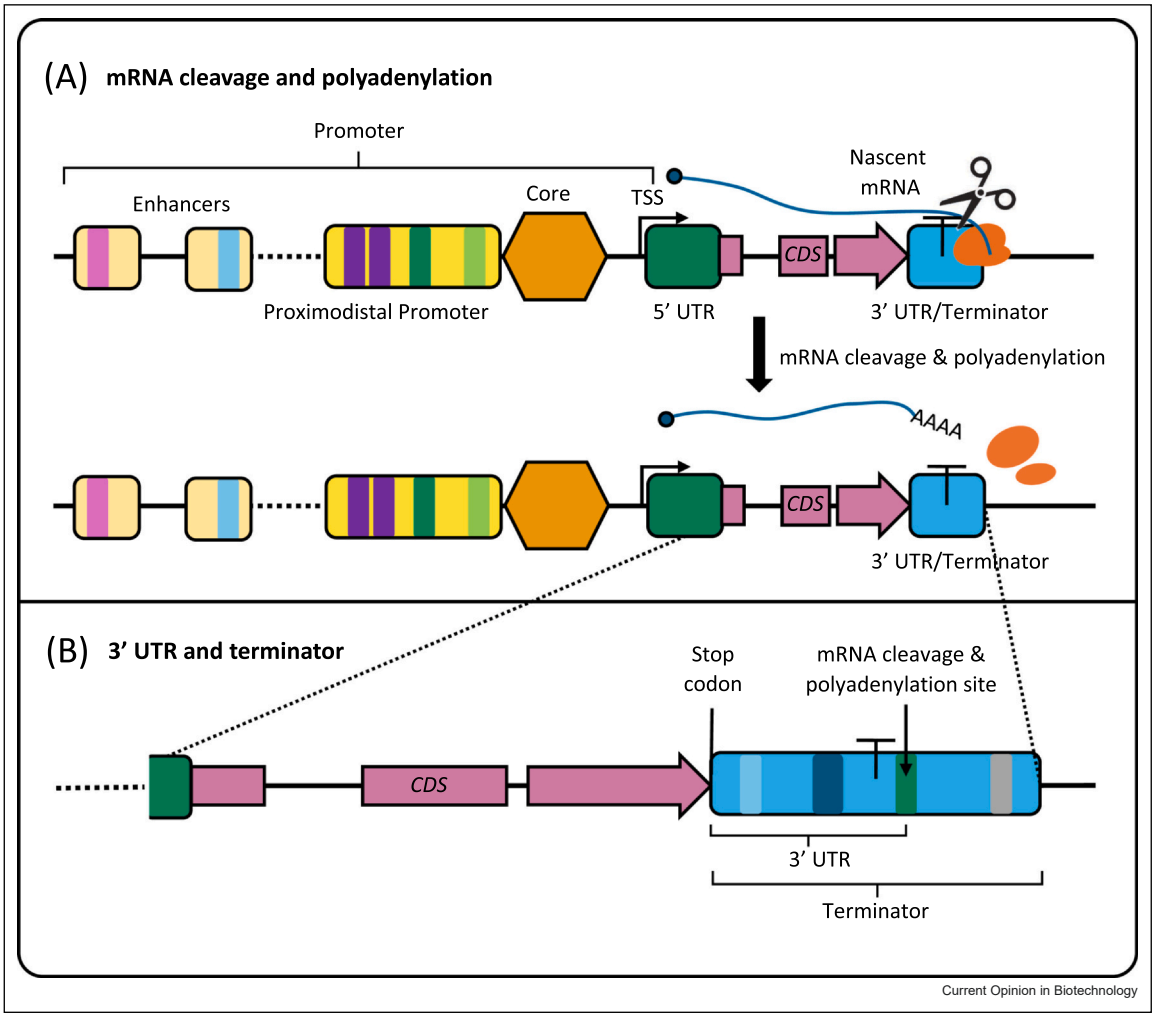
3' untranslated region/terminators

When a transcribing RNAPII finishes reading the coding part of a gene, it continues copying the 3' UTR and the rest of the terminator region (Figure 3). The terminator is thus a transcribed sequence of a gene that spans the mRNA cleavage and polyadenylation site and serves to recruit a set of the 3' end processing and polyadenylation factors [2]. The poly(A) tail is critical to mRNA stability and plays an important role in mRNA export to the cytoplasm from the nucleus [93]. In mammals and yeast, the machinery responsible for these 3' end processing events has been well defined, and the plant homologs of a majority of these factors have been identified [94]. RNA-seq approaches have been instrumental to defining the 3' UTRs of most protein-coding genes in the genome, but the more downstream terminator sequences not contained in the 3' UTR are not well mapped.

When building a plant transgene of interest, to ensure its optimal expression, traditionally, a terminator from constitutive genes with high levels of expression, such as that of CaMV *35S* (*T35S*), *Agrobacterium NOS* (*Tnos*) and *octopine synthase* (*OCS*), or plant housekeeping genes, such as *ACTIN* (*ACT*) or *UBIQUITIN* (*UBQ*), is commonly used [94]. However, the 3' end processing and polyadenylation factors do not function in isolation, and numerous lines of evidence show that these proteins physically interact or functionally cross-talk with the general TF TFIIB, at least in humans [95,96]. Thus, for a chosen promoter, different terminators may significantly affect the reporter gene expression [94]. Mitsuhara et al. [97] compared the *T35S* and *Tnos* terminators combined with a series of chimeric promoters in transient and stable expression systems in tobacco (*N. tabacum*) and rice (*O. sativa*). The authors found that the *T35S* terminator was more effective than the *Tnos*. To expand the choice of available terminators, Diamos and Mason [98] systematically compared 20 different plant and viral terminators in combination with the *35S* promoter and TMV 5' UTR and found that in *N. benthamiana* transient assays, eight terminators significantly enhance the expression of the reporter genes (*GFP* or *DsRed*) relative to the *T35S* or *Tnos* terminators. Similarly, Tian et al. [99] compared 13 plant and viral terminators combined with the cassava vein mosaic virus (*CsVMV*) promoter in *N. benthamiana* leaves and *N. tabacum* BY2 cells. These authors found that the terminator of the Arabidopsis *HEAT SHOCK PROTEIN18.2* (*AtHsp18.2*) gene produced 1.4- and 2.4-fold higher expression of the reporter gene compared with the *T35S* and *Tnos* terminators, respectively. Recently, Gorjifard et al. [100] measured the activity of over 50,000 terminators from *Arabidopsis* and maize in combinations with the *35S* promoter using STARR-seq in tobacco leaves and maize protoplasts. These authors found that thousands of Arabidopsis and maize terminators were better than the *Tnos* and *Agrobacterium mannopine synthase* (*MAS*) terminators at enhancing *GFP* reporter gene expression, with a handful of these outperforming the *T35S* terminator. The authors concluded that the optimal terminators for the *35S* promoter in dicots are that of *At3G46230* (*HEAT SHOCK PROTEIN17.4*), *At2G05530* (a Glycine-rich protein gene), and *At4G39730* (*PLAT DOMAIN POTEIN1*), whereas in monocots, the terminators of maize genes *Zm00001d016542* (*anthranilate 1, 2-dioxygenase)*, *Zm00001d047961* (unknown), and *Zm00001d017119* (*glucose-6-phosphate dehydrogenase 5*) are the best.

Interestingly, in multiple plant systems (tobacco leaves, sugarcane leaf segments, and sorghum), a double terminator combining the *T35S* and *Tnos* sequences was reported to

Figure 3



Transcription termination. **(a)** mRNA cleavage and polyadenylation. RNAPII (orange ovals) transcribes the 5' UTR, the CDS, and the 3' UTR/terminator. The resulting nascent RNA is cleaved roughly in the middle of the terminator region, and a poly(A) tail is added. **(b)** 3' UTR and terminator. The 3' UTR is the 5'-most part of the terminator (cyan) that extends from the stop codon of the CDS to the mRNA cleavage polyadenylation site. This is the part of the terminator that is straightforward to infer from RNA-seq data. It contains several cis-elements (colored boxes) important for 3' end processing, including the polyadenylation site (AAUAAA or a related sequence). Addition of the 3' UTR in the 3' end of a construct is usually insufficient to obtain efficient transcription termination, as downstream terminator elements are lacking. The full terminator extends past the mRNA cleavage and polyadenylation site and includes additional cis-elements (gray box) necessary for the mRNA cleavage and polyadenylation machinery to process the 3' end of the transcript.

significantly increase the *eYFP* reporter gene expression relative to either terminator alone in constructs driven by the maize *Ubi-1* promoter, as shown by Beyene et al. [101]. Diamos and Mason [98] found that combining terminators in tandem produced synergistic effects and that seven double terminators significantly exceeded the strength of the *T35S-Tnos* double terminator. In addition, instead of combining two terminators, Meshcheriakova et al. [102] fused the 3' UTR of CPMV RNA-2 to the *Tnos* terminator and found that these sequences increased *GFP* expression in *N. benthamiana* transient assays threefold relative to the *NOS* terminator alone.

To avoid reusing natural regulatory elements, including classical terminators/3' UTRs derived from plant viruses, Peyret et al. [49] designed eight synthetic 3' UTRs based on the properties of highly expressed genes of plant viruses, such as low 3' UTR GC content and the presence of the polyadenylation signal AAUAAA, as well as CA and UUUU motifs. Some synthetic 3' UTRs also contained the Y-loop structure from the 3' UTR of CPMV RNA-2. However, none of the synthetic 3' UTRs were better at supporting gene expression than the 3' UTR of CPMV RNA-2. Gorjifard et al. [100] used the DenseNet model for *in silico* evolution of 222

terminators (111 terminators from Arabidopsis and maize each). After 10 rounds of evolution, several terminators generated by an *in silico* evolution approach had greater strength than the *T35S* terminator in *N. benthamiana* leaves and maize protoplasts, indicating that combining iterative in *silico* and STARR-seq is a promising strategy for optimizing DNA parts.

With this body of literature in mind, when deciding on the best terminator for one's construct of interest, especially, if a novel synthetic promoter is used, the well-studied viral and bacterial terminators of *35S*, *NOS*, and *OCS* genes and those from highly expressed plant genes such as *RUBISCO* or *HEAT SHOCK PROTEIN* are typically chosen [94,103] and continue to be the safest option. However, if a readthrough transcription is of concern, especially in multigene constructs with the potential for silencing, a double terminator may be preferred, and the seven aforementioned double terminators identified by Diamos and Mason [98] and the *T35S-Tnos* double terminator evaluated by Beyene et al. [101] should be considered.

Other DNA parts

Not all transgenes are intended for protein expression. If the goal of a transgene is, for example, to monitor RNA levels and distribution, then a DNA part encoding a functional RNA may need to be included in the construct. RNA aptamers are short RNA sequences that fold into tertiary structures that can be applied to studying RNA localization. One common approach adopted in plants involves the use of hairpin-shaped RNA aptamers, such as that from the bacteriophage MS2, alongside a sequence-specific RNA-binding protein; in this case, MS2 coat protein (MCP) tagged with an FP (Figure 2g) [104]. RNA transcripts of a GOI are tagged with one or more MS2 aptamers (typically placed downstream of the stop codon), and the fluorescence of MCP-FP recruited to the MS2 aptamers is tracked to infer the tagged RNA trafficking and localization. Aptamer/FP-based detection is the current standard technique for live-cell imaging of RNA in plants. For example, Alamos et al. [105] tagged RNA transcripts with either MS2 or another aptamer, PP7. Aptamer-binding bacteriophage MCP and PP7 coat protein were fused to a GFP tag, which enabled the identification of active transcription sites using laser-scanning confocal microscopy to visualize RNAPII activity in Arabidopsis and *N. benthamiana* under different treatment conditions. Live-cell RNA imaging techniques have enabled researchers to understand dynamic processes in plant cells with an unprecedented detail and will likely continue to transform our current understanding of biological processes.

Besides expressing RNA aptamer fusions, transgenic expression of noncoding RNAs, such as miRNA, circular RNA, long noncoding RNA, or CRISPR guideRNAs, etc., is routine in plants [106]. Additionally, a transgene or its part may be intended to function at the DNA level and serve to deliver recombination sites into the plant genome (Figure 2j). Upon co-expression of a heterologous recombinase, *in planta* removal, addition, or inversion of sequences of interest can be triggered. Recombination sites are often used in conjunction with selectable marker and reporter genes so that the recombination events can be visualized. For example, Chamness et al. [107] generated transgenic *N. benthamiana* lines harboring an inactive *RUBY* reporter separated from a dual *35S* promoter by a Kanamycin resistance marker, *NptII*, flanked by different recombination sites. Upon expressing a recombinase to excise the *NptII* gene, betalain accumulation was observed, indicative of the efficient recombination. Furthermore, insulators and insulator-like elements can be incorporated into transgenic constructs to overcome challenges associated with the positional effects of T-DNA insertions and unwanted interactions of transgenes with endogenous genetic elements. Insulators are DNA elements that can block enhancer–promoter interactions and create chromosomal boundaries that shield transgenes from heterochromatin [108]. To date, there are few well-characterized true insulators in plants, with matrix attachment regions (MARs) being the best-studied class of insulator-like elements. Although most MARs lack the enhancer-blocking activity of true insulators, they can shield a transgene from the surrounding chromosomal environment by facilitating the formation of chromatin loops that separate the genome into independently regulated domains [109].

Finally, additional design decisions on what vectors to employ, which molecular cloning technologies to utilize, whether to include linkers, scars, or stuffers between DNA parts, and how to arrange the genes in a construct are all important for the success of one's project. Even minor considerations such as the order, spacing, and relative orientations (head-to-head, head-to-tail, or tail-to-tail) of genes in a plasmid may have profound effects on transgene functionality due to the promoter of one gene potentially serving as an enhancer for a neighboring gene or the possible leakiness of gene terminators resulting in a transcriptional readthrough and construct silencing [110]. In the end, testing multiple construct designs remains the safest option for most applications.

Concluding remarks

In the past few years, transgenic approaches in plants shifted toward an early-stage adoption of synthetic biology as an enabling tool to overcome some of the limitations of traditional constructs made of well-characterized natural parts. The fact that most of the synthetic DNA elements described to date were originally characterized in a limited set of conditions (often in just one species, developmental stage, tissue, environment, and assay) restricts the ability of a researcher to extrapolate the behavior of man-made parts from one biological system or experimental setting to

another. Therefore, we anticipate that a more universal, systematic, standardized testing of select DNA parts would need to be implemented before their broad adoption by the plant biology community. The functional validation work would need to take into account not only how specific DNA parts behave in the context of a given construct but also consider its interactions with all other DNA parts in a library to identify optimal combinations and minimize its interference with other components of the cell. We foresee that high-throughput strategies to measure the effects of synthetic DNA elements and their architecture on different aspects of gene expression, likely in combination with mathematical modeling and machine learning–based approaches, will become more mainstream in plant sciences and will augment and empower the design of growingly more functionally complex DNA constructs. The ultimate goal would be to move from modeling the effect of individual DNA components to predicting the behaviors of whole genes, pathways, cells, and biological systems. In the meantime, we hope this brief overview of what is currently feasible in plant sciences can serve as a starting point for a beginner looking to design an optimal construct and for a professional aiming to build new tools to advance the horizons of plant biology.

CRediT authorship contribution statement

Katie Vollen: Conceptualization; writing – original draft; visualization; **Chengsong Zhao**: Conceptualization; writing – original draft; **Jose M. Alonso**: Conceptualization; writing – review & editing; **Anna Stepanova**: Conceptualization; writing – original draft, review & editing; visualization.

Data Availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

The work in the Alonso-Stepanova lab is supported by the National Science Foundation grants 1444561 and 2327912 to JMA and ANS and 1750006 to ANS. We would like to thank Anna Yaschenko for her help with the graphical abstract and all members of the Alonso-Stepanova lab and anonymous reviewers for their constructive feedback.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Roeder RG, Rutter WJ: **Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms**. *Nature* 1969, **224**:234-237.

2. • Muniz L, Nicolas E, Trouche D: **RNA polymerase II speed: a key player in controlling and adapting transcriptome composition**. *EMBO J* 2021, **40**.
This review describes the approaches to measure RNAPII speed, the factors involved in this process, and the importance of transcription speed regulation.

3. Hirose Y, Manley JL: **RNA polymerase II and the integration of nuclear events**. *Genes Dev* 2000, **14**:1415-1429.

4. Wu Q, Bazzini AA: **Translation and mRNA stability control**. *Annu Rev Biochem* 2023, **92**:227-245.

5. Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F: **cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in Arabidopsis**. *Plant Cell* 2010, **22**:1425-1440.

6. Brumos J, Robles LM, Yun J, Vu TC, Jackson S, Alonso JM, Stepanova AN: **Local auxin biosynthesis is a key regulator of plant development**. *Dev Cell* 2018, **47**:306-318.e5.

7. D'Orso F, Forte V, Baima S, Possenti M, Palma D, Morelli G: **Methods and techniques to select efficient guides for CRISPR-mediated genome editing in plants**. In *A Roadmap for Plant Genome Editing*. Edited by Ricroch A, Eriksson D, Miladinović D, Sweet J, Van Laere K, Woźniak-Gientka E. Springer Nature Switzerland; 2024:89-117.

8. • Brooks EG, Elorriaga E, Liu Y, Duduit JR, Yuan G, Tsai C-J, Tuskan GA, Ranney TG, Yang X, Liu W: **Plant promoters and terminators for high-precision bioengineering**. *Biodes Res* 2023, **5**:0013.
The structure and functions of plant promoters and terminators are reviewed, with the emphasis on their application in plant engineering.

9. Khan A, Nasim N, Pudhuvai B, Koul B, Upadhyay SK, Sethi L, Dey N: **Plant synthetic promoters: advancement and prospective**. *Agriculture* 2023, **13**:298.

10. Villao-Uzho L, Chávez-Navarrete T, Pacheco-Coello R, Sánchez-Timm E, Santos-Ordóñez E: **Plant promoters: their identification, characterization, and role in gene regulation**. *Genes* 2023, **14**:1226.

11. Yasmeen E, Wang J, Riaz M, Zhang L, Zuo K: **Designing artificial synthetic promoters for accurate, smart, and versatile gene expression in plants**. *Plant Commun* 2023, **4**:100558.

12. Lu Z, Hofmeister BT, Vollmers C, DuBois RM, Schmitz RJ: **Combining ATAC-seq with nuclei sorting for discovery of *cis*-regulatory regions in plant genomes**. *Nucleic Acids Res* 2017, **45**:e41.

13. Oka R, Zicola J, Weber B, Anderson SN, Hodgman C, Gent JI, Wesselink J-J, Springer NM, Hoefsloot HCJ, Turck F, *et al.*: **Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize**. *Genome Biol* 2017, **18** (n/a).

14. • Ricci WA, Lu Z, Ji L, Marand AP, Ethridge CL, Murphy NG, Noshay JM, Galli M, Mejía-Guerra MK, Colomé-Tatché M, *et al.*: **Widespread long-range *cis*-regulatory elements in the maize genome**. *Nat Plants* 2019, **5**:1237-1249.
This report provides functional molecular evidence for the existence of numerous long-range cis-regulatory elements in the maize genome and implicates them in controlling gene expression via the formation of chromatin loops.

15. Sun J, He N, Niu L, Huang Y, Shen W, Zhang Y, Li L, Hou C: **Global quantitative mapping of enhancers in rice by STARR-seq**. *Genom Prote Bioinform* 2019, **17**:140-153.

16. Mozo T, Fischer S, Meier-Ewert S, Lehrach H, Altmann T: **Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome**. *Plant J* 1998, **16**:377-384.

17. Sawitzke JA, Thomason LC, Costantino N, Bubunenko M, Datta S, Court DL: **Recombineering: in vivo genetic engineering in *E. coli, S. enterica*, and beyond**. *Methods Enzym* 2007, **421**:171-199.

18. Zhou R, Benavente LM, Stepanova AN, Alonso JM: **A recombineering-based gene tagging system for Arabidopsis**. *Plant J* 2011, **66**:712-723.

19. Alonso JM, Stepanova AN: **Arabidopsis transformation with large bacterial artificial chromosomes**. *Methods Mol Biol* 2014, **1062**:271-283.

20. Alonso JM, Stepanova AN: **A recombineering-based gene tagging system for Arabidopsis**. *Methods Mol Biol* 2015, **1227**:233-243.

21. •• Brumos J, Zhao C, Gong Y, Soriano D, Patel AP, Perez-Amador MA, Stepanova AN, Alonso JM: **An improved recombineering toolset for plants**. *Plant Cell* 2020, **32**:100-122.
A simple, scalable, customizable recombineering toolset and protocols were developed to allow a variety of genetic modifications to be generated in plant genes harbored by BACs in *E. coli*.

22. Amack SC, Antunes MS: ***CaMV35S* promoter — a plant biology and biotechnology workhorse in the era of synthetic biology**. *Curr Plant Biol* 2020, **24**:100179.

23. Lee S, Moon JS, Ko T-S, Petros D, *et al*.: **Overexpression of Arabidopsis phytochelatin synthase paradoxically leads to hypersensitivity to cadmium stress**. *Plant Physiol* 2003, **131**:656-663.

24. Jung S, Lee H, Lee Y, Kang K, Kim YS, Grimm B, Back K: **Toxic tetrapyrrole accumulation in protoporphyrinogen IX oxidase-overexpressing transgenic rice plants**. *Plant Mol Biol* 2008, **67**:535-546.

25. Li S, Wang C, Zhou L, Shou H: **Oxygen deficit alleviates phosphate overaccumulation toxicity in *OsPHR2* overexpression plants**. *J Plant Res* 2014, **127**:433-440.

26. Zhang Y, Chen M, Liu T, Qin K, Fernie AR: **Investigating the dynamics of protein-protein interactions in plants**. *Plant J* 2023, **114**:965-983.

27. • Fernandez-Moreno J-P, Yaschenko AE, Neubauer M, Marchi AJ, Zhao C, Ascencio-Ibanez JT, Alonso JM, Stepanova AN: **A rapid and scalable approach to build synthetic repetitive hormone-responsive promoters**. *Plant Biotechnol J* 2024, https://doi.org/10.1111/pbi.14313.
A simple, single-tube reaction method was developed to generate repetitive promoter sequences that are not amenable to direct commercial DNA synthesis. The strategy was employed to generate a novel synthetic ethylene-responsive reporter and validate it in stable Arabidopsis lines.

28. • Yang Y, Chaffin TA, Shao Y, Balasubramanian VK, Markillie M, Mitchell H, Rubio-Wilhelmi MM, Ahkami AH, Blumwald E, Neal Stewart CJ: **Novel synthetic inducible promoters controlling gene expression during water-deficit stress with green tissue specificity in transgenic poplar**. *Plant Biotechnol J* 2024, https://doi.org/10.1111/pbi.14289.
Generation and testing of two novel poplar synthetic promoters that are active in photosynthetic tissues is described, a drought-inducible promoter and a constitutive promoter.

29. Omelina ES, Yushkova AA, Motorina DM, Volegov GA, Kozhevnikova EN, Pindyurin AV: **Optogenetic and chemical induction systems for regulation of transgene expression in plants: use in basic and applied research**. *Int J Mol Sci* 2022, **23**.

30. • Jores T, Tonnies J, Dorrity MW, Cuperus JT, Fields S, Queitsch C: **Identification of plant enhancers and their constituent elements by STARR-seq in tobacco leaves**. *Plant Cell* 2020, **32**:2120-2131.
This study employed STARR-seq in transiently transformed *N. benthamiana* leaves to compare the effect of transcriptional enhancer elements sourced from the 35S and several natural plant promoters.

31. Bruckner K, Schafer P, Weber E, Grutzner R, Marillonnet S, Tissier A: **A library of synthetic transcription activator-like effector-activated promoters for coordinated orthogonal gene expression in plants**. *Plant J* 2015, **82**:707-716.

32. • Cai Y-M, Kallam K, Tidd H, Gendarini G, Salzman A, Patron NJ: **Rational design of minimal synthetic promoters for plants**. *Nucleic Acids Res* 2020, **48**:11845-11856.
Natural plant and plant pathogen promoters were analyzed to identify unique and common cis-elements. Minimal synthetic plant promoters of different strengths were engineered and tested in stable Arabidopsis transformants and in isolated protoplasts of three additional plant species.

33. •• Jores T, Wrightsman JT, Buckler T, Cuperus ES, Fields JT, Queitsch S: **C: synthetic promoter designs enabled by a comprehensive analysis of plant core promoters**. *Nat Plants* 2021, **7**:842-855.
STARR-seq was implemented to quantify the strength of a nearly complete set of natural promoters from Arabidopsis, maize, and sorghum. This data set was leveraged to develop strong synthetic promoters.

34. Rajasundaram D, Selbig J, Persson S, Klie S: **Co-ordination and divergence of cell-specific transcription and translation of genes in arabidopsis root cells**. *Ann Bot* 2014, **114**:1109-1123.

35. •• Brophy JAN, Magallon KJ, Duan L, Zhong V, Ramachandran P, Kniazev K, Dinneny JR: **Synthetic genetic circuits as a means of reprogramming plant roots**. *Science* 2022, **377**:747-751.
Synthetic orthogonal TFs were tested in agroinfiltrated *N. benthamina* plants and leveraged to generate and evaluate a series of genetic logic gates in *N. benthamiana* transient assays and stable Arabidopsis transformants. The utility of the approach was validated by tweaking the Arabidopsis root architecture using a basic BUFFER gate.

36. Cordero T, Rosado A, Majer E, Jaramillo A, Rodrigo G, Daròs J-A: **Boolean computation in plants using post-translational genetic control and a visual output signal**. *ACS Synth Biol* 2018, **7**:2322-2330.

37. • Lloyd JPB, Ly F, Gong P, Pflueger J, Swain T, Pflueger C, Fourie E, Khan MA, Kidd BN, Lister R: **Synthetic memory circuits for stable cell reprogramming in plants**. *Nat Biotechnol* 2022, **40**:1862-1872.
Recombinase-based genetic switches were implemented in Arabidopsis and leveraged to develop a series of genetic logic gates.

38. Srivastava AK, Lu Y, Zinta G, Lang Z, Zhu J-K: **UTR-dependent control of gene expression in plants**. *Trends Plant Sci* 2018, **23**:248-259.

39. • Ryczek N, Łyś A, Makałowska I: **The functional meaning of 5'UTR in protein-coding genes**. *Int J Mol Sci* 2023, **24**:2976, https://doi.org/10.3390/ijms24032976.
The structure and functions of 5' UTRs are reviewed.

40. Gallie DR, Kado CI: **A translational enhancer derived from tobacco mosaic virus is functionally equivalent to a Shine-Dalgarno sequence**. *Proc Natl Acad Sci USA* 1989, **86**:129-132.

41. Jobling SA, Gehrke L: **Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence**. *Nature* 1987, **325**:622-625.

42. Fan Q, Treder K, Miller WA: **Untranslated regions of diverse plant viral RNAs vary greatly in translation enhancement efficiency**. *BMC Biotechnol* 2012, **12** 22–22.

43. Anthonisen IL, Salvador ML, Klein U: **Specific sequence elements in the 5' untranslated regions of *rbcL* and *atpB* gene mRNAs stabilize transcripts in the chloroplast of *Chlamydomonas reinhardtii***. *RNA* 2001, **7**:1024-1033.

44. Martinez-Trujillo M, Limones-Briones V, Chavez-Barcenas T, Herrera-Estrella L: **Functional analysis of the 5' untranslated region of the *sucrose phosphate synthase* rice gene (*sps1*)**. *Plant Sci* 2003, **165**:9-20.

45. Satoh J, Kato K, Shinmyo A: **The 5'-untranslated region of the tobacco *alcohol dehydrogenase* gene functions as an effective translational enhancer in plant**. *J Biosci Bioeng* 2004, **98**:1-8.

46. De Amicis F, Patti T, Marchetti S: **Improvement of the pBI121 plant expression vector by leader replacement with a sequence combining a poly(CAA) and a CT motif**. *Transgenic Res* 2007, **16**:731-738.

47. Kanoria S, Burma PK: **A 28 nt long synthetic 5'UTR (synJ) as an enhancer of transgene expression in dicotyledonous plants**. *BMC Biotechnol* 2012, **12**.

48. Tanaka H, Suzuki Y, Yamasaki S, Yoshino K, Kato K, Nakamura S: **R-STEINER: generation method of 5'UTR for increasing the amount of translated proteins**. *IPSJ Trans Bioinform* 2018, **11**:14-23.

49. Peyret H, Brown JKM, Lomonossoff GP: **Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions**. *Plant Methods* 2019, **15**.

50. Sainsbury F, Thuenemann EC, Lomonossoff GP: **pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants**. *Plant Biotechnol J* 2009, **7**:682-693.

51. Zhang H, Si X, Ji X, Fan R, Liu J, Chen K, Wang D, Gao C: **Genome editing of upstream open reading frames enables translational control in plants**. *Nat Biotechnol* 2018, **36**:894-898.

52. de Ruijter NCA, Verhees J, van Leeuwen W, van der Krol AR: **Evaluation and comparison of the GUS, LUC and GFP reporter system for gene expression studies in plants**. *Plant Biol* 2003, **5**:103-115.

53. •• He Y, Zhang T, Sun H, Zhan H, Zhao Y: **A reporter for noninvasively monitoring gene expression and plant transformation**. *Hortic Res* 2020, **7**:152.

A RUBY reporter was developed in this study by fusing the coding regions of three enzymes involved in betalain biosynthesis and separating the three via a P2A peptide.

54. Caissard J-C, Guivarc'h A, Rembur J, Azmi A, Chriqui D: **Spurious localizations of diX-indigo microcrystals generated by the histochemical GUS assay**. *Transgenic Res* 1994, **3**:176-181.

55. Hu CY, Chee PP, Chesney RH, Zhou JH, Miller PD, O'Brien WT: **Intrinsic GUS-like activities in seed plants**. *Plant Cell Rep* (1) 1990, **9**:5.

56. Lauressergues D, Ormancey M, Guillotin B, San Clemente H, Camborde L, Duboé C, Tourneur S, Charpentier P, Barozet A, Jauneau A, *et al.*: **Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity**. *Cell Rep* 2022, **38**:110339.

57. Ckurshumova W, Caragea AE, Goldstein RS, Berleth T: **Glow in the dark: fluorescent proteins as cell and tissue-specific markers in plants**. *Mol Plant* 2011, **4**:794-804.

58. Tanz SK, Castleden I, Small ID, Millar AH: **Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants**. *Front Plant Sci* 2013, **4**.

59. Wang J, Bollier N, Buono RA, Vahldick H, Lin Z, Feng Q, Hudecek R, Jiang Q, Mylle E, Van Damme D, *et al.*: **A developmentally controlled cellular decompartmentalization process executes programmed cell death in the Arabidopsis root cap**. *Plant Cell* 2023,941-962, https://doi.org/10.1093/plcell/koad308

60. Sperschneider J, Catanzariti A-M, DeBoer K, Petre B, Gardiner DM, Singh KB, Dodds PN, Taylor JM: **LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell**. *Sci Rep* 2017, **7**:44598.

61. Sahu SS, Loaiza CD, Kaundal R: **Plant-mSubP: a computational framework for the prediction of single- and multi-target protein subcellular localization using integrated machine-learning approaches**. *AoB Plants* 2020, **12**:plz068.

62. • Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, *et al.*: **Highly accurate protein structure prediction with AlphaFold**. *Nature* 2021, **596**:583-589.

A powerful machine learning – based method for three-dimensional protein structure prediction was developed that outperformed all previously available prediction methods.

63. • Bertoline LMF, Lima AN, Krieger JE, Teixeira SK: **Before and after AlphaFold2: an overview of protein structure prediction**. *Front Bioinform* 2023, **3**:1120370.

This review describes the state of the art in the protein structure modeling field.

64. Cheng X, Mwaura BW, Chang Stauffer SR, Bezanilla M: **A fully functional ROP fluorescent fusion protein reveals roles for this GTPase in subcellular and tissue-level patterning**. *Plant Cell* 2020, **32**:3436-3451.

65. de Wet JR, Wood KV, DeLuca M, Helinski DR, Subramani S: **Firefly *luciferase* gene: structure and expression in mammalian cells**. *Mol Cell Biol* 1987, **7**:725-737.

66. Matthews JC, Hori K, Cormier MJ: **Purification and properties of *Renilla reniformis* luciferase**. *Biochemistry* 1977, **16**:85-91.

67. Vazquez-Vilar M, Quijano-Rubio A, Fernandez-del-Carmen A, Sarrion-Perdigones A, Ochoa-Fernandez R, Ziarsolo P, Blanca J, Granell A, Orzaez D: **GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data**. *Nucleic Acids Res* 2017, **45**:2196-2209.

68. Khakhar A, Starker CG, Chamness JC, Lee N, Stokke S, Wang C, Swanson R, Rizvi F, Imaizumi T, Voytas DF: **Building customizable auto-luminescent luciferase-based reporters in plants**. *Elife* 2020, **9**, https://doi.org/10.7554/eLife.52786

69. Mitiouchkina T, Mishin AS, Somermeyer LG, Markina NM, Chepurnyh TV, Guglya EB, Karataeva TA, Palkina KA, Shakhova ES, Fakhranurova LI, *et al.*: **Plants with genetically encoded autoluminescence**. *Nat Biotechnol* 2020, **38**:944-946.

70. •• Shakhova ES, Karataeva TA, Markina NM, Mitiouchkina T, Palkina KA, Perfilov MM, Wood MG, Hoang TT, Hall MP, Fakhranurova LI, *et al.*: **An improved pathway for autonomous bioluminescence imaging in eukaryotes**. *Nat Methods* 2024, **21**:406-410.

This study utilized directed evolution to optimize the bioluminescence pathway using enzymes sourced from diverse bioluminescent fungi. Prominent autonomous luminescence (i.e. not requiring addition of the substrate) was achieved in six different plant species, as well as in yeast and mammalian cell lines.

71. •• Palkina KA, Karataeva TA, Perfilov MM, Fakhranurova LI, Markina NM, Somermeyer LG, Garcia-Perez E, Vazquez-Vilar M, Rodriguez-Rodriguez M, Vazquez-Vilriales V, *et al.*: **A hybrid pathway for self-sustained luminescence**. *Sci Adv* 2024, **10**:eadk1992.

A hybrid synthetic bioluminescent pathway that leverages a combination of fungal and plant enzymes was implemented in *N. benthamiana*, yeast, and mammalian cell cultures. A compact-sized hispidin synthase gene sourced from *Plumbago zeylanica* replaces fungal hispidin synthase and phosphopantetheinyl transferase genes utilized in earlier versions of the pathway.

72. Sharma P, Yan F, Doronina VA, Escuin-Ordinas H, Ryan MD, Brown JD: **2A peptides provide distinct solutions to driving stop-carry on translational recoding**. *Nucleic Acids Res* 2012, **40**:3143-3151.

73. Wang D, Zhong Y, Feng B, Qi X, Yan T, Liu J, Guo S, Wang Y, Liu Z, Cheng D, *et al.*: **The *RUBY* reporter enables efficient haploid identification in maize and tomato**. *Plant Biotechnol J* 2023, **21**:1707-1715.

74. Khakhar A, Leydon AR, Lemmex AC, Klavins E, Nemhauser JL: **Synthetic hormone-responsive transcription factors can monitor and re-program plant development**. *eLife* 2018, **7**:e34702.

75. Huang L, Rojas-Pierce M: **Rapid depletion of target proteins in plants by an inducible protein degradation system**. *Plant Cell* 2024, https://doi.org/10.1093/plcell/koae072

76. Rozov SM, Deineko EV: **Strategies for optimizing recombinant protein synthesis in plant cells: classical approaches and new directions**. *Mol Biol* 2019, **53**:157-175.

77. Kim DH, Hwang I: **Direct targeting of proteins from the cytosol to organelles: the ER versus endosymbiotic organelles**. *Traffic* 2013, **14**:613-621.

78. • Xu S-L, Shrestha R, Karunadasa SS, Xie P-Q: **Proximity labeling in plants**. *Ann Rev Plant Biol* 2023, **74**:285-312.

This manuscript provides a thorough overview of the proximity labeling approaches and is accompanied by excellent graphical schematics.

79. Liu W, Yuan JS, Stewart Jr CN: **Advanced genetic tools for plant biotechnology**. *Nat Rev Genet* 2013, **14**:781-793.

80. Pan C, Sretenovic S, Qi Y: **CRISPR/dCas-mediated transcriptional and epigenetic regulation in plants**. *Curr Opin Plant Biol* 2021, **60**:101980.

81. Viola IL, Gonzalez DH: **Chapter 2 — Methods to study transcription factor structure and function**. In *Plant*

Transcription Factors. Edited by Gonzalez DH. Academic Press; 2016:13-33.

82. Nandwani N, Surana P, Negi H, Mascarenhas NM, Udgaonkar JB, Das R, Gosavi S: **A five-residue motif for the design of domain swapping in proteins**. *Nat Commun* 2019, **10**:452(x).

83. Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C: **Design parameters to control synthetic gene expression in *Escherichia coli***. *PLoS One* 2009, **4**:e7002.

84. Quax TEF, Claassens NJ, Söll D, van der Oost J: **Codon bias as a means to fine-tune gene expression**. *Mol Cell* 2015, **59**:149-161.

85. Liu Y: **A code within the genetic code: codon usage regulates co-translational protein folding**. *Cell Commun Signal* 2020, **18**:145.

86. Cheema N, Papamichail G, Papamichail D: **Chapter 12 — Computational tools for synthetic gene optimization**. In *New Frontiers and Applications of Synthetic Biology*. Edited by Singh V. Academic Press; 2022:171-189.

87. Callis J, Fromm M, Walbot V: **Introns increase gene expression in cultured maize cells**. *Genes Dev* 1987, **1**:1183-1200.

88. Samadder P, Sivamani E, Lu J, Li X, Qu R: **Transcriptional and post-transcriptional enhancement of gene expression by the 5′ UTR intron of rice *rubi3* gene in transgenic rice cells**. *Mol Genet Genom* 2008, **279**:429-439.

89. Laxa M: **Intron-mediated enhancement: a tool for heterologous gene expression in plants?** *Front Plant Sci* 2017, **7**.

90. Jaeger D, Baier T, Lauersen KJ: **Intronserter, an advanced online tool for design of intron containing transgenes**. *Algal Res* 2019, **42**:101588.

91. • Cao VD, Luo G, Korynta S, Liu H, Liang Y, Shanklin J, Altpeter F: **Intron-mediated enhancement of *DIACYLGLYCEROL ACYLTRANSFERASE1* expression in energycane promotes a step change for lipid accumulation in vegetative tissues**. *Biotechnol Biofuels Bioprod* 2023, **16**:153.
This study provides an illustrative recent example of how including an intron in a transgene can boost the expression of GOI and, in this case, enhance lipid production.

92. Grützner R, Martin P, Horn C, Mortensen S, Cram EJ, Lee-Parsons CWT, Stuttmann J, Marillonnet S: **High-efficiency genome editing in plants mediated by a Cas9 gene containing multiple introns**. *Plant Commun* 2020, **2**:100135.

93. Kuehner JN, Pearson EL, Moore C: **Unravelling the means to an end: RNA polymerase II transcription termination**. *Nat Rev Mol Cell Biol* 2011, **12**:283-294.

94. • de Felippes FF, Waterhouse PM: **Plant terminators: the unsung heroes of gene expression**. *J Exp Bot* 2023, **74**:2239-2250.
This timely review sheds light on the previously underappreciated role of terminators in regulating gene expression.

95. Bratkowski M, Unarta IC, Zhu L, Shubbar M, Huang X, Liu X: **Structural dissection of an interaction between transcription initiation and termination factors implicated in promoter-terminator cross-talk**. *J Biol Chem* 2018, **293**:1651-1665.

96. Al-Husini N, Medler S, Ansari A: **Crosstalk of promoter and terminator during RNA polymerase II transcription cycle**. *Biochim Biophys Acta Gene Regul Mech* 2020, **1863**:194657.

97. Mitsuhara I, Ugaki M, Hirochika H, Ohshima M, Murakami T, Gotoh Y, Katayose Y, Nakamura S, Honkura R, Nishimiya S, *et al.*: **Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants**. *Plant Cell Physiol* 1996, **37**:49-59.

98. Diamos AG, Mason HS: **Chimeric 3' flanking regions strongly enhance gene expression in plants**. *Plant Biotechnol J* 2018, **16**:1971-1982.

99. • Tian C, Zhang Y, Li J, Wang Y: **Benchmarking intrinsic promoters and terminators for plant synthetic biology research**. *Biodes Res* 2022,9834989.
A set of 19 natural promoters and 13 natural terminators was examined by dual luciferase assays in agroinfiltrated *N. benthamiana* leaves and tobacco BY2 cells. 105 promoter – terminator pairs were also compared leading to an over 300-fold difference in reporter gene expression.

100. •• Gorjifard S, Jores T, Tonnies J, Mueth NA, Bubb K, Wrightsman T, Buckler ES, Fields S, Cuperus JT, Queitsch C: Arabidopsis and Maize Terminator Strength Is Determined by GC Content, Polyadenylation Motifs and Cleavage Probability; bioRxiv ⟨doi:10.1101/2023.06.16.545379⟩..
The activity of over 50,000 natural terminators derived from *Arabidopsis thaliana* and *Zea mays* was measured in *N. benthamiana* leaves and maize protoplasts via STARR-seq. Many of the terminators tested exceeded the strength of Agrobacterium terminators traditionally used in transgenic applications. Furthermore, efficient synthetic terminators were developed via *in silico* evolution of plant terminators.

101. Beyene G, Buenrostro-nava MT, Damaj MB, Gao S, Molina J, Mirkov TE: **Unprecedented enhancement of transient gene expression from minimal cassettes using a double terminator**. *Plant Cell Rep* 2011, **30**:13-25.

102. Meshcheriakova YA, Saxena P, Lomonossoff GP: **Fine-tuning levels of heterologous gene expression in plants by orthogonal variation of the untranslated regions of a nonreplicating transient expression system**. *Plant Biotechnol J* 2014, **12**:718-727.

103. Bernardes WS, Menossi M: **Plant 3' regulatory regions from mRNA-encoding genes and their uses to modulate expression**. *Front Plant Sci* 2020, **11**:1252.

104. Zhang F, Simon AE: **A novel procedure for the localization of viral RNAs in protoplasts and whole plants**. *Plant J* 2003, **35**:665-673.

105. •• Alamos S, Reimer A, Niyogi KK, Garcia HG: **Quantitative imaging of RNA polymerase II activity in plants reveals the single-cell basis of tissue-wide transcriptional dynamics**. *Nat Plants* 2021, **7**:1037-1049.
This study employed laser-scanning confocal microscopy and RNA aptamer FP labeling techniques to visualize RNAPII activity in *Arabidopsis thaliana* and *N. benthamiana* in response to stress.

106. Bhar A, Roy A: **Emphasizing the role of long non-coding RNAs (lncRNA), circular RNA (circRNA), and micropeptides (miPs) in plant biotic stress tolerance**. *Plants* 2023, **12**:3951.

107. •• James CC, Jon PC, Anna JC, Daniel FV: Viral Delivery of Recombinases to Activate Heritable Genetic Switches in Plants; bioRxiv. 2024. ⟨doi:10.1101/2024.03.03.583219⟩..
Tobacco rattle virus vector was employed to deliver one of four recombinases (Cre, FLP, CinH, and Integrase13) into transgenic tobacco plants harboring cognate recombinase sites flanking a selectable marker gene, *NptII*. Successful recombination was monitored by visualizing the excision of *NptII* upon which a functional *RUBY* reporter was reconstituted.

108. Kurbidaeva A, Purugganan M: **Insulators in plants: progress and open questions**. *Genes* 2021, **12**.

109. Dolgova AS, Dolgov SV: **Matrix attachment regions as a tool to influence plant transgene expression**. *3 Biotech* 2019, **9**:176.

110. Jiang W, Sun L, Yang X, Wang M, Esmaeili N, Pehlivan N, Zhao R, Zhang H, Zhao Y: **The effects of transcription directions of transgenes and the *gypsy* insulators on the transcript levels of transgenes in transgenic Arabidopsis**. *Sci Rep* 2017, **7**:14757.