Enhancers in Plant Development, Adaptation and Evolution

Bliss M. Beernink, John P. Vogel and Li Lei*

U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

*Corresponding author: E-mail, LiLei@lbl.gov

Understanding plant responses to developmental and environmental cues is crucial for studying morphological divergence and local adaptation. Gene expression changes, governed by *cis*-regulatory modules (CRMs) including enhancers, are a major source of plant phenotypic variation. However, while genome-wide approaches have revealed thousands of putative enhancers in mammals, far fewer have been identified and functionally characterized in plants. This review provides an overview of how enhancers function to control gene regulation, methods to predict DNA sequences that may have enhancer activity, methods utilized to functionally validate enhancers and the current knowledge of enhancers in plants, including how they impact plant development, response to environment and evolutionary adaptation.

Introduction

At its most fundamental level, transcription requires the binding of RNA polymerase to a promoter, followed by the synthesis of a complementary RNA strand. However, this process typically requires additional *cis*-regulatory elements [CREs, e.g. enhancers (terminology and acronyms defined in Box 1)] and proteins [e.g. transcription factors (TFs)]. These additional elements allow cells to precisely control expression levels, timing and location and to modulate expression in response to external and internal cues. This intricate regulation is essential for the normal functioning of all organisms.

Enhancers are crucial in regulating gene expression and play a significant role in the complex regulatory networks of plants. They can modulate the transcription of target genes over long genomic distances, influencing critical processes such as development, stress responses and adaptation. Enhancers' ability to integrate multiple signals and facilitate precise spatial and temporal control of gene expression makes them indispensable for the dynamic and versatile regulatory needs of plants. Understanding enhancers' roles and mechanisms in plants can provide insights into plant biology and open new avenues for crop improvement and adaptation strategies.

CREs are specific DNA sequence motifs recognized by RNA polymerase, TFs and other regulatory proteins, often found clustered together to form *cis*-regulatory modules (CRMs). CRMs function as genomic regulatory codes that interact combinatorially with TFs to modulate gene expression (Davidson 2006). These interactions coordinate genome-wide transcriptional programs, enabling cells to respond dynamically to environmental conditions and developmental cues (reviewed in Schmitz et al. 2022, Marand et al. 2023).

CRMs can be categorized into distinct regulatory domains based on their function, including core promoters, enhancers, silencers and insulators. Promoters are identified as upstream CREs containing specific nucleotide sequences that RNA polymerases and TFs can bind. The minimal sequence required to bind an RNA polymerase and initiate transcription is referred to as the core promoter and typically spans 50–100 bp prior to the transcription starting site (Juven-Gershon and Kadonaga 2010). Promoters are classified based on the type of RNA polymerase they recruit and their transcription patterns (e.g. constitutive, inducible or tissue specific; reviewed in Villao-Uzho et al. 2023). Working in concert with promoters, enhancers are CRMs bound by additional TFs and cofactors capable of boosting transcription rates, thereby enhancing gene expression in a tissue-, developmental stage- and/or condition-specific manner. Although enhancers cannot initiate transcription independently like promoters, studies have shown that some promoters possess enhancer functions (Dao et al. 2017, Andersson and Sandelin 2020). In contrast to enhancers, silencers actively repress the expression of their target genes. Similarly, both enhancers and silencers are bound by TFs and associated cofactors and can be located up- or downstream, or within introns of their target genes functioning in an orientation-independent manner (Ogbourne and Antalis 1998, Doni Jayavelu et al. 2020, Ngan et al. 2020). Insulators, when bound by specific proteins, prevent expression of target genes when positioned between enhancers, silencers and promoters. Taken together, CRMs coordinate gene regulation by enhancing, suppressing or

A.
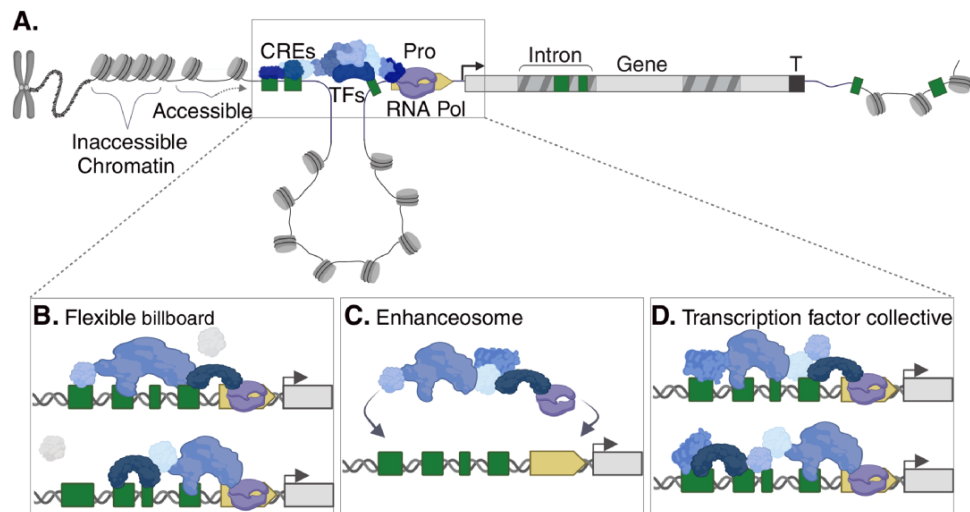
Fig. 1 *Cis*-regulatory module recruitment and binding of TFs. (A) *Cis*-regulatory modules, composed of clustered CREs, regulate transcription by binding TFs and other regulatory proteins. The arrangement of CREs is highly variable and CREs have been identified upstream, downstream and within introns of their target gene. It has also been determined that long distance gene regulation can occur from several megabases away with CREs being brought into proximity through chromatin looping. (B–D) There are three proposed models for how TFs are recruited and assembled at CREs. (B) The flexible billboard model asserts that TFs bind independently to TF-binding motifs found within CRMs resulting in a flexible architecture and grammar for TF binding. (C) The enhanceosome model purports that a TF complex forms prior to being recruited to the CRM. (D) The TF collective model suggests that each CRE can recruit multiple TFs and that while all of these recruited TFs are required, the order of binding and cooperation within the enhanceosome is more flexible. The gene being transcribed is indicated as a rectangle labeled as 'Gene', with an arrow representing the direction of transcription. CREs are represented as boxes labeled 'CREs', the promoter is represented as a pointed rectangle labeled 'Pro', introns are denoted by striped rectangles within the gene target, the terminator is indicated as a black square labeled 'T', the RNA polymerase (RNA Pol) is denoted as the shape bound to the promoter and TFs are represented as several shapes labeled 'TFs' showing the many different TFs binding together at the CRM. Created in BioRender. Beernink, B. (2024) BioRender.com/j29z706.

repressing transcription. Since some genes activate or suppress gene cascades, individual enhancers can affect the expression of many genes and may play key roles in development and responses to the environment.

Enhancers usually regulate their target genes via binding TFs and cofactors to form a transcription initiation complex at the promoter (reviewed in de Laat and Duboule 2013, Long et al. 2016, Li and X-d 2019). Three proposed models describe how TFs interact with enhancers to regulate gene expression (**Fig. 1**; Spitz and Furlong 2012; Jindal and Farley 2021). In all three models, the TF-binding sites contained in enhancers are key in determining gene regulation. First, the flexible billboard model (**Fig. 1B**), proposed by Kulkarni and Arnosti (2003), suggests that TFs bind TF-binding motifs independently, the enhancer order is fixed, but TFs have flexible binding syntax. Kulkarni and Arnosti (2003) found that a flexible arrangement of TFs predominated in developing embryonic *D. melanogaster* tissues, where gene expression diversity is crucial (Vockley et al. 2017).

The second model, the enhanceosome, was proposed by Panne et al. (2007) and posits that a specific TF complex forms prior to binding enhancers (**Fig. 1C**). Their study on the animal gene interferon-β (IFN-β) showed that an enhanceosome of eight TF proteins recognized and bound the entire enhancer. However, it is not clear if the enhanceosome binds to the DNA as a fully formed complex or if the assembly occurs as the TFs bind to the DNA (Panne et al. 2007, Panne 2008). It is

thought that in addition to IFN-β, other highly conserved genes may also follow the enhanceosome model in order to achieve regulatory precision. The inflexibility and evolutionary conservation of the enhanceosome model lies in stark contrast to the flexible billboard model, which allows expression diversity and evolutionary flexibility.

The third model, the TF collective model, developed by Junion et al. (2012), posits that the CRM uses multiple mechanisms, including protein–protein and protein–DNA binding, to recruit multiple TFs. This collective binding forms diverse enhanceosomes along the flexible motif composition of enhancers (**Fig. 1D**). This model was supported by evidence from studies on cardiac cell fate determination in *D. melanogaster* and further supported by Uhl et al. (2016) through reporter assays showing protein–DNA and protein–protein interactions in regulatory complex formation. These studies found that the combinatorial binding of TFs does not require a specific binding motif orientation, suggesting an alternative model of CRE cooperation that has extensive motif flexibility in comparison to the flexible billboard model. It has recently been suggested that enhancer activity and enhanceosome formation may be explained as a spectrum of the proposed models and that no single model is the correct model (Jindal and Farley 2021). In fact, Jores et al. (2024) showed that plant enhancers can exhibit both billboard-like and enhanceosome-like activity.

The first enhancer was identified from a DNA virus, simian virus 40, in 1981 (Banerji et al. 1981). Additional enhancers continued to be identified in mammalian systems, leading to the discovery of greater numbers of putative enhancers from more complex genomes [for a historical perspective on enhancer discovery, see Schaffner (2015)]. Today, over a million putative enhancers have been identified from mammalian genomes (reviewed in Göndör and Ohlsson 2018). In humans, the median enhancer size is 1.3 kb (Zhao et al. 2022, Thomas and Buecker 2023). Some enhancers cluster together in close genomic proximity, forming arrays, often known as super-enhancers, with a median size of 19.4 kb (Lovén et al. 2013). The concerted action of multiple enhancers has been found to alter gene expression of a single gene at different times from varied locations (Crocker et al. 2014, Long et al. 2016). The location of enhancers relative to their target genes are highly variable, being found upstream, downstream, within introns and distally (Lettice et al. 2002, Benko et al. 2009, Pennacchio et al. 2013; **Fig. 1A**). While not common, enhancers can also act from several megabases away when they are brought into proximity of the target gene through chromatin looping (Kadauke and Blobel 2009; Peng et al. 2019). Nearly one-third of identified enhancers are derived from transposable elements (TEs). Based on predicted CRMs in human and *Zea mays* (maize) genomes, as much as 25% and 30% of regulatory sequences may be TE-derived, respectively (Oka et al. 2017, Zhao et al. 2018, Pehrsson et al. 2019, Fagny et al. 2021). Transient reporter assays have confirmed that a significant number of TE-derived CRMs have enhancer activity in mammalian or maize cells (Lynch et al. 2011, Xie et al. 2013, Zhao et al. 2018).

Researchers have used characteristics of known enhancers including their correlation with evolutionarily conserved sequences, chromatin accessibility, certain histone marks and TF-binding sites to identify putative enhancers (Kellis et al. 2014). Genetic evidence from quantitative trait locus (QTL) mapping and genetic fine-mapping have been used to identify putative enhancers (Stam et al. 2002a, Clark et al. 2006, Salvi et al. 2007, Studer et al. 2011, Zheng et al. 2015). In these cases, because the DNA interval that co-segregated with the examined trait did not contain a coding gene, it was assumed that the sequence must be regulatory. In some cases, further experimentation validated the actual enhancer. The availability of whole genome sequences for many species has enabled the identification of many putative enhancers using comparative genomics (Weber et al. 2016; **Table 1**). These methods rely on the conservation of non-coding sequences over evolutionary time to identify functionally constrained sequences. Many of these conserved sequences are presumably enhancers. This approach is useful for taxonomic groups with highly conserved genomes, like mammals, but is of limited use for plants due to their dynamic genomes that make multi-species genome alignment difficult (Kellis et al. 2014). Another approach to identify putative enhancers is through the distribution of DNA marks and chromatin accessibility. Like genes, enhancers are found in areas of open chromatin.

Thus, histone marks typically associated with genes can be used to identify non-coding sequences that may act as enhancers. The main methods used to identify open chromatin include DNase I-sequencing (DNase-seq), formaldehyde-assisted isolation of regulatory element-sequencing (FAIRE-seq), micrococcal nuclease-sequencing and assay for transposase accessible chromatin-sequencing (ATAC-seq; Kellis et al. 2014).The use of these methods to help identify putative enhancers will be discussed in the next section.

In this review, we focus on methods for identifying and functionally validating putative enhancers and highlight examples of the roles of enhancers play in plant development, biotic and abiotic stress responses and evolution.

Approaches to Identify and Validate Enhancers in Plants

Approaches to identify enhancers in plants

Enhancers are difficult to identify because they lack a singular defining characteristic like the open reading frames found in protein-coding genes. Plant enhancers are more difficult to identify than animal enhancers because the latter share some characteristic hallmarks (e.g. H3K4me1 in the flanking nucleosomes). Further complicating identification, enhancers can be located virtually anywhere with respect to the gene they regulate. This makes the lack of syntenic conservation between plant genomes particularly challenging to overcome. Enhancer identification requires, first, identification of putative enhancer sequences and, secondly, functional validation of the enhancer properties using molecular methods. Due to the labor-intensive nature of these experiments, only 200 plant enhancers have been functionally validated to date. Weber et al. (2016) review and summarize the studies identifying and functionally validating the first 20 plant enhancers. Since then, technological and sequencing advancements over the last decade have enabled the identification of nearly 130,000 putative plant enhancers, of which around 178 plant enhancers were functionally validated and summarized in **Table 1**. Below, we describe established and emerging methods to identify putative plant enhancers and approaches to increase the throughput of enhancer validation.

Experimental approaches

Individual enhancer identification Traditionally, enhancers have been identified by first locating genomic regions that harbor putative enhancers (e.g. enhancer trapping and QTL mapping) and then applying molecular methods to demonstrate the enhancer activity of specific DNA sequences (e.g. transformation with a reporter gene with the enhancer upstream of a minimal promoter; reviewed in Shlyueva et al. 2014). However, due to the labor required and the need to create transgenic plants, most studies stop after identifying the interval containing a putative enhancer. In plants, enhancer trapping introduces a reporter gene under the control of a minimal promoter randomly into the genome, and when the minimal promoter

Table 1 Identification of putative plant enhancers using sequencing approaches

Plant species	Putative enhancers identified	Method for identification	Enhancers validated	Method for functional validation	Impact	References
Arabidopsis thaliana	749	DNase-seq	6	CRISPR/Cas9 deletion	Primary root length	Zhao et al. (2022)
Arabidopsis thaliana	4,327	STARR-seq	30	Leaf infiltration assay		Tan et al. (2023)
Arabidopsis thaliana	12,216	ATAC-seq and DNase-seq	6	Protoplast reporter assay	Vegetative growth and PAMP-triggered immunity	Zhang et al. (2022)
Arabidopsis Thaliana	3,281	DNase-seq	15	Transgenic reporter assay		Meng et al. (2021)
Arabidopsis thaliana	4,844	DNase-seq and ChIP-seq	22	Transgenic reporter assay	Flower development	Yan et al. (2019)
Arabidopsis thaliana	10,044	DNase-seq	14	Transgenic report assay	Circadian regulation	Zhu et al. (2015)
Arabidopsis thaliana	12	DNase-seq	12	Leaf infiltration assay	-	Lin et al. (2019)
Cucumis sativus	82,756	Comp. genomics and ATAC-seq	2	CRISPR/Cas9 deletion	Cell size and fruit size	Xin et al. (2024)
Oryza sativa	9,642	STARR-seq	29	Protoplast reporter assay		Sun et al. (2019)
Solanum tuberosum	1	DNase-seq	1	CRISPR/Cas9 deletion	Cold stress response	Zhu et al. (2024)
Triticum aestivum	2,144–8,280	DNase-seq, H3K9ac, ChIP-seq and BS-seq	36	Protoplast reporter assay	Pathogenesis	Xie et al. (2022)
Zea mays	1,500	DNase-seq, H3K9ac, ChIP-seq and BS-seq	5	Previously validated enhancers		Oka et al. (2017)

This table describes recent publications that utilize high-throughput sequence approaches to identify putative plant enhancers and go on to functionally validate a portion of enhancers identified in their studies.

comes under the influence of endogenous enhancers, it is activated, leading to expression of the reporter gene (Chudalayandi 2011). This indicates that an enhancer is in the vicinity of the transgene, but further work is required to identify the actual enhancer sequence. For plants with small transformable genomes, unknown enhancers have been identified via enhancer-trap transgenic lines (Sundaresan et al. 1995, Wu et al. 2003, Gardner et al. 2009). McGarry and Ayre (2008) successfully identified enhancer *MATURE MINOR VEIN ELEMENT1* (*MMVE1*) in *Arabidopsis thaliana* (*Arabidopsis*) using enhancer trapping methods. However, for plants with large genomes or those recalcitrant to transformation, researchers apply traditional forward genetic analysis, such as QTL mapping and genetic fine-mapping approaches, to identify candidate enhancers. Forward genetics starts with a well-characterized phenotype, such as disease resistance, and works toward identifying the gene(s) responsible for the phenotype. QTL mapping uses statistical methods to link phenotypic data (trait measurements) and genotypic data (usually molecular markers) to explain the genetic basis of variation in complex traits (Falconer and Mackay 1996, Kearsey 1998, Lynch and Walsh 1998). For example, in maize, researchers used a forward genetics approach to identify several unannotated regions that presumably contain an enhancer, including the *Hepta-repeat b1* enhancer, the *teosinte branched 1* (*TB1*) enhancer, *Vegetative to generative transition 1* and *Distal Cis-Element* (*DICE*), which are

located 100, 70, 60 and 140 kb upstream of their (presumed) target genes, respectively (Stam et al. 2002a, Clark et al. 2006, Salvi et al. 2007, Studer et al. 2011, Zheng et al. 2015). Similarly, Du et al. (2020) identified *KERNEL ROW NUMBER4* (*KRN4*), an enhancer validated with reported gene assay, located 60 kb downstream of its target gene, *UNBRANCHED3* (*UB3*). Although enhancer-trap and forward genetic approaches have successfully identified several enhancers, more comprehensive high-throughput methods are desirable.

Genome-wide investigations Advances in sequencing technology have enabled several new approaches to identify genome-wide putative enhancers. These novel sequencing approaches detect hallmarks of enhancers including accessible chromatin regions (ACRs), TF-binding sites and certain chromatin interactions (Lieberman-Aiden et al. 2009, Bernstein et al. 2010, Spitz and Furlong 2012, Shlyueva et al. 2014, Weber et al. 2016, Oka et al. 2017, Lu et al. 2019, Ricci et al. 2019).

Several assays that detect ACRs are available including DNase-seq (Boyle et al. 2008), FAIRE-seq (Giresi et al. 2007) and ATAC-seq (Buenrostro et al. 2015). With DNase-seq, DNase I hypersensitive sites can be identified by partial digestion of chromatin with the endonuclease DNase I, followed by sequencing of the small fragments. The accessible fraction of the genome will be under-represented in the resulting sequence data (Hesselberth et al. 2009). DNase-seq robustly identifies

DNase I hypersensitive sites but is not very sensitive in predicting TF-binding motifs because of the intrinsic cleavage bias of DNase I (He et al. 2014). FAIRE-seq identifies protein-free DNA regions (i.e. free from nucleosomes) by cross-linking chromatin with formaldehyde, followed by sonication of chromatin and phenol/chloroform extraction of the nucleosome-free DNA fragments. FAIRE offers a lower resolution than DNase I–based assays because sonication produces higher background noise than DNase I digestion (Song et al. 2011). ATAC-seq relies on an active Tn5 transposase to insert adapters into ACR preferentially in vitro. Sequencing from these adapters enables quantitative measuring of chromatin accessibility throughout the genome (Zhu et al. 2015). Sequence enrichment relative to naked DNA enables defining of genome-wide ACRs, indicating TF-binding sites (Gross and Garrard 1988, Henikoff 2008). The earliest ACR studies in plants were completed in *Oryza sativa* (rice) and *Arabidopsis*, using a DNase-seq approach to identify ACRs that are potentially CRMs (Zhang et al. 2012a, 2012b, Jiang 2015). However, only a subset of ACRs are enhancers, so distance from the transcription start site is used to differentiate between promoters, genes and enhancers (Zhu et al. 2015, Yan et al. 2019). Zhu et al. (2015) applied the same approach to identify >10,000 putative distal enhancers and used reporter assays to detect and validate enhancer activity for 10 out of 14 candidates (**Table 1**). Zhang et al. (2022) used both ATAC-seq and DNase-seq to identify 12,216 candidate enhancers in *Arabidopsis* (**Table 1**). Lin et al. (2019) used DNase-seq to identify 12 *Arabidopsis* and six rice putative enhancers and validate them using a *Nicotiana benthamiana* leaf infiltration reporter system. Similarly, Zhu et al. (2024) used DNase-seq to identify an intronic enhancer impacting cold stress responses in *Solanum tuberosum* (potato; **Table 1**). Recently, single-cell ATAC-seq has proved highly effective for detecting cell-type-specific CRMs, at single-cell resolution in plants (Marand et al. 2021). This is particularly important because recent research on *Arabidopsis* and maize revealed that approximately one-third of detected ACRs by single-cell ATAC-seq are cell-type specific (Dorrity et al. 2021, Marand et al. 2021).

By definition, TF-binding sites are the most likely sequences to act as enhancers and two methods have been developed to identify TF-binding sites. Chromatin immunoprecipitation-sequencing (ChIP-seq; Haring et al. 2007, Johnson et al. 2007) is widely used to detect and characterize CRMs, including enhancers. Experimentally, the chromatin from cross-linked or native nuclei is fragmented and precipitated with an antibody to the protein of interest. DNA associated with the protein of interest is then purified and sequenced (Kaufmann et al. 2010). ChIP-seq can be used to detect TF-binding sites and histone modifications. Researchers can also use other approaches, like DNase-seq, to identify the ACRs and then identify the putative enhancers based on the epigenetic mark (i.e. H3K27ac) on the surrounding histones with CHIP-seq. For example, Yan et al. (2019) combined DNase-seq and ChIP-seq to identify 4,844 putative enhancers in *Arabidopsis*. Using similar approaches and combining with other RNA-sequencing technologies, researchers also identified thousands of putative enhancers in *Triticum aestivum* (wheat) and maize (Oka et al. 2017, Xie et al. 2022). However, ChIP-seq has several limitations (reviewed in Schmitz et al. 2022), including (I) the limited availability of antibodies that recognize TFs and other proteins of interest, (II) the lack of transgenic plants expressing functional epitope-tagged TFs, (III) the highly dynamic nature of TF–DNA interactions (Para et al. 2014) and (IV) difficulties in scaling throughput.

To address the shortcomings of ChIP-seq for identifying TF-binding sites, DNA affinity purification sequencing (DAP-seq) was developed (O'Malley et al. 2016, Bartlett et al. 2017). DAP-seq combines affinity purification of genomic DNA fragments and next-generation sequencing to provide a high-throughput platform for mapping the genome-wide DNA-binding sites of single TF or TF families. Specifically, target TFs are fused with affinity tags, such as HaloTag, and expressed in vitro. Genomic DNA extracted from plant tissues is fragmented by sonication, and Illumina compatible sequencing adapters are ligated to the DNA fragments to create a genomic DNA library. After incubating the Halo-TF with the DNA library, the TF/DNA complex is purified by HaloTag ligand–conjugated magnetic beads, unbound DNA is washed away and TF-bound DNA is sequenced (Li and Huang 2022). O'Malley et al. (2016) first used this approach to profile genome-wide, context-independent binding sites for 529 *Arabidopsis* TFs. However, DAP-seq cannot identify TF-binding sites that require heteromultimeric interactions for binding, causing many TFs to fail under the current framework (Marand et al. 2023). For example, only 30% of the 1,812 *Arabidopsis* TFs tested in a recent large-scale cistrome mapping screen yielded TF-binding sites (O'Malley et al. 2016, Bartlett et al. 2017). A recently developed method, Systematic Evolution of Ligands by Exponential Enrichment coupled to Consecutive Affinity Purification, has been developed to identify sites where cooperative binding of TF occurs (Jolma et al. 2015, Ibarra et al. 2020, X. Li et al., submitted for publication). For example, using this approach, Jolma et al. (2015) analyzed 9,400 TF–TF–DNA interaction and identified 618 new heterodimeric motifs from 316 TF–TF interactions with human cell lines, but this has not been applied in plants. Some enhancers are activated by combinations of TFs (reviewed in Spitz and Furlong 2012) such that the binding of only one or a few TFs may be insufficient to activate transcription. Thus, it is hard to know which combinations of TF-binding sites and enhancers actually activate transcription.

One of the hallmarks of enhancers is that they come into close physical proximity to the promoter of the gene they regulate. Thus, methods that detect the spatial relationship between regions of DNA can be used to identify regions containing putative enhancers. Chromosome conformation capture (3C) and its derivatives (e.g. 4C, 5C and Hi-C) measure relative interaction frequencies between different genomic regions and can be used to identify putative enhancer regions (de Wit and de Laat 2012, Bodega and Lanzuolo 2021). Briefly, these methods cross-link DNA, while it is still in the cell so that adjacent DNA

strands become covalently bound. The DNA is then digested with restriction enzymes to produce small fragments covalently bound by the cross-link. The DNA is then ligated under conditions that favor ligation of the covalently linked fragments. The DNA is then sequenced or amplified by qPCR to detect sequences that are composed of chimeric DNA, sequences that are not adjacent in the genome. These methods can identify both putative enhancers and their target genes. Hi-C allows genome-wide detection of putative enhancer–promoter interactions (Rao et al. 2014). In plants, 3C was the first utilized to find the *Hepta-repeat b1* enhancer locus in maize by identifying interactions between the *hepta-repeat* enhancer and the B1 gene Louwers et al. 2009a; Louwers et al. 2009b).

Bioinformatic detection of putative enhancers. Comparative genomics is an elegant computational approach to predict CREs including enhancers. This approach assumes that DNA sequences involved in gene regulation are more conserved than non-functional DNA across a broad phylogenetic range (Hardison and Taylor 2012). Non-coding sequences occupy a large portion of plant genomes. Interestingly, in several studies most genome-wide associations were located in non-coding regions, suggesting that polymorphisms in control elements like enhancers are responsible for considerable phenotypic variation (Wallace et al. 2014, Zhang and Lupski 2015, Nishizaki and Boyle 2017, Giral et al. 2018). Comparison of conserved non-coding sequences (CNSs) across species can identify regions under purifying selection suggesting functional constraints like regulatory functions. In a study of 502 human genes, CREs were found to be more prevalent in CNSs than in non-conserved non-coding areas (Levy et al. 2001). Thus, identification of CNSs by comparative genomics can be used to identify putative CREs. However, the relatively short length of CREs within CRMs, along with the increased sequence turnover of sequences around CREs, makes the identification of CNSs challenging (Van de Velde et al. 2016), and since plant CNSs are considerably smaller and less numerous than those in mammals, they are particularly challenging to identify. However, this method has resulted in the discovery of multiple CRMs in plants, such as the *AGAMOUS* (*AG*) and *FLOWERING LOCUS T FT* loci in *Arabidopsis* and the miR164 locus in *Brassicaceae* (Hong et al. 2003, Adrian et al. 2010, Jain et al. 2018). Recently, Xin et al. (2024) applied comparative genomics to identify 82,756 CNSs found in cucurbit species, and ~20% of them overlapped with ACRs identified by ATAC-seq. Thus, the identification of CNSs is useful for identifying CRMs. However, since many CRMs are less conserved, they will not be identified as CNSs and additional information, such as TF-DNA-binding data or chromatin structure, data is needed to pinpoint their location (Lieberman-Aiden et al. 2009, Bernstein et al. 2010, Spitz and Furlong 2012, Shlyueva et al. 2014, Weber et al. 2016, Oka et al. 2017, Lu et al. 2019, Ricci et al. 2019).

Machine learning is a powerful computational approach that is being widely applied to make inferences from large datasets in many fields. In particular, deep learning, a type of machine learning, has been used to predict CREs in plants (Shen et al. 2021, Sielemann et al. 2021, Liu et al. 2021a, Wang et al. 2021a, Yan et al. 2022, Hu et al. 2023, Peleke et al. 2024). For example, SeqConv (Shen et al. 2021) and TSPTFBS (Liu et al. 2021a) employed a deep learning model, called deep convolutional neural network, to build predictive models for maize and *Arabidopsis* using TF-binding sites identified with ChIP-seq and DAP-seq, respectively. Although these two methods significantly outperform conventional methods in plants, they still suffered some drawbacks, such as requiring training of a predictive model for each TF and ignoring that DNA is a complex, three-dimensional macromolecule. Sielemann et al. (2021) and Wang et al. (2021a) showed improved performance by incorporating the 3D shape of DNA into the model. In 2022, Plant-Bind successfully applied deep learning to predict the potential TF-binding sites of 315 TFs in *Arabidopsis* and showed high suitability of transfer learning in maize (Yan et al. 2022). In 2024, DeepCRE was developed by using data from *Arabidopsis*, *Solanum lycopersicum* (tomato), *Sorghum bicolor* (sorghum) and maize to predict the CREs. A couple of tools based on deep learning were also developed to predict the CRMs in plants. For example, Klie et al. (2023) developed a toolkit called Elucidating the Utility of Genomic Elements with Neural nets, which was trained using the plant CRMs identified by Jores et al. (2021). Using the same training dataset, Deng et al. (2023) developed a web-based tool, called iCREPCP, to predict plant core promoters. However, the use of machine learning to predict enhancers in plants is still in its infancy. The main limitation is the paucity of ground truth data, previously characterized plant enhancers, which makes it hard to train models.

Approaches to validate enhancers in plants

Reporter assays. Reporter assays are the benchmark method for validating and characterizing enhancer sequences (Shlyueva et al. 2014). The general approach starts by cloning a candidate enhancer adjacent to a minimal promoter that is upstream of a reporter gene (e.g. GFP or GUS). The resulting construct is then introduced into a cell/organism through transient or stable transformation. The cell/organism is then examined for expression of the reporter gene. Since the minimal promoter can only promote low-level transcription by itself, any expression of the reporter gene is generally below detectable levels, especially for GFP- and GUS-based assays. Therefore, high levels of GFP or GUS protein production are linked to the candidate enhancer activity directed by the DNA sequence in question (Simpson et al. 1985, Timko et al. 1985, Pwee and Gray 1993, Sparkes et al. 2006, Belele et al. 2013). Transient assays such as *N. benthamiana* leaf infiltration assays or plant protoplast transfection systems are commonly used for reporter assays as an efficient screening tool. However, these systems can only detect enhancers that utilize machinery present in leaves or protoplasts. Thus, the validation and characterization of many enhancers require the creation of stable transgenic lines and the examination of many organs/cell types/environmental conditions for reporter gene expression.
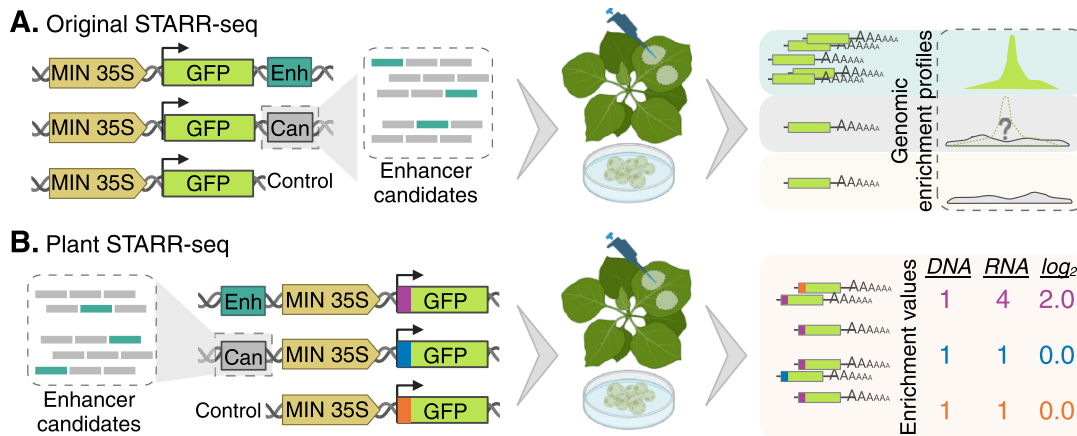
Fig. 2 STARR-sequencing (STARR-seq) for plant enhancer identification. This figure illustrates two versions of STARR-seq used for identification of plant enhancers from genomic DNA fragments. Enhancer candidates (Can) were screened in transient reporter assays to determine their enhancer activity. (A) Original STARR-seq (Arnold et al. 2013). Enhancer candidates were cloned into the 3′ UTR of constructs containing a minimal cauliflower mosaic virus 35S promoter (MIN 35S), which has weak transcription activity without an enhancer, and GFP, a reporter gene. A no-enhancer negative control (Control) is also included. Constructs were tested in transient reporter assays, such as leaf infiltration or protoplast assays. The messenger RNA is then sequenced, and transcript levels are mapped back to the genome to generate genomic enrichment profiles, identifying candidate enhancer activity. (B) Plant STARR-seq (Jores et al. 2020) is a modified version of the original method, where the candidate enhancers are cloned into the 5′ UTR of constructs containing a MIN 35S, GFP and a barcode for sequencing (indicated by the small boxes inside the GFP box). A no-enhancer negative control (Control) is included. Constructs are tested in transient reporter assays. The transcribed barcodes are sequenced and their enrichment relative to the input DNA is calculated giving a distinct enrichment value for each tested enhancer candidate. Enhancers are indicated by boxes labeled 'Enh', candidate enhancers are indicated by gray rectangles labeled 'Can', the MIN 35S promoter is represented by pointed rectangles with arrows indicating transcription direction and GFP is represented by rectangles labeled 'GFP'. Created in BioRender. Beernink, B. (2024) BioRender.com/u85z459.

High-throughput reporter assays. Although reporter assays are the benchmark method to validate enhancers, they are labor-intensive and cannot simultaneously evaluate many candidate enhancers. Recently, Jores et al. (2020) adapted a method called self-transcribing active regulatory region sequencing (Plant STARR-seq), which is more sensitive in plants (**Fig. 2**). STARR-seq was initially developed in *D. melanogaster* (Arnold et al. 2013) for genome-wide identification of functional enhancers. Plant STARR-seq is set up similarly to a reporter assay, where a candidate enhancer sequence is cloned upstream of a minimal 35S promoter and linked to a barcode at the start of the coding sequence of the reporter gene. However, when a candidate enhancer upregulates transcription, its activity is detected by next-generation sequencing instead of assaying reporter gene expression (**Fig. 2**). This strategy was used to identify enhancers in *Arabidopsis*, rice and maize protoplasts (Singer et al. 2012, Sun et al. 2019, Tan et al. 2023). In addition to detecting enhancers, Plant STARR-seq can be used to understand the biology underlying enhancer activity. For example, Jores et al. (2020) found that plant enhancers are less active when they are inserted in the 3′ untranslated region (UTR) of the reporter genes, but they can show strong and condition-specific activity when inserted just upstream of the promoter.

Gene editing with CRISPR/Cas9 deletion. While high-throughput parallel reporter assays like Plant STARR-seq can rapidly validate the activity of the enhancers, they rely on the transient expression of a reporter construct and, therefore, lack the genomic context surrounding the candidate enhancers. Thus, they can only validate enhancers that are not context dependent and are able to function under the conditions used for the assay. CRISPR/Cas9 genome editing is being used to overcome these limitations and characterize CRMs, including enhancers, in their native genomic context by selectively deleting enhancers and looking at the resultant gene expression changes. For example, researchers have used CRISPR/Cas9-mediated deletions to determine the effect of CRMs, mainly promoters, on gene regulation at specific loci (Rodríguez-Leal et al. 2017, Hendelman et al. 2021, Liu et al. 2021b, Wang et al. 2021b, Song et al. 2022). Similarly, Zhao et al. (2022) used CRISPR/Cas9 gene editing to validate six enhancers related to the length of the primary root in the early *Arabidopsis* development. Recently, Xin et al. (2024) used CRISPR/Cas9 to delete candidate enhancers in *Cucumis sativus* (cucumber) and validate the function of two enhancers by showing their impact on fruit and cell size (**Table 1**). While this approach enables researchers to validate and characterize enhancers in their native genomic context, it is laborious and time-consuming (Jores et al. 2023). Thus, a logical approach is to use large-scale reporter assays like Plant STARR-seq as a preliminary filter to select candidate sequences for CRISPR/Cas9 gene editing or reporter assays to provide a more comprehensive and contextual understanding of the function of the most promising enhancers.
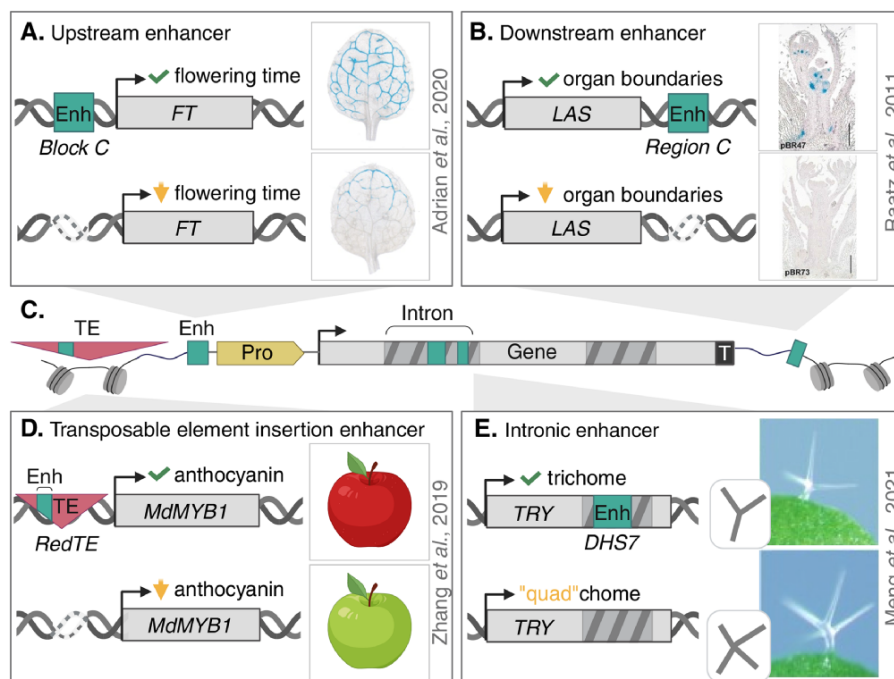
Fig. 3 Enhancer impacts on plant phenotypes despite location variation within plant genomes. (A) The enhancer, *Block C*, is located upstream of the promoter and the gene being targeted for expression. *Block C* is demonstrated to upregulate *FT* gene expression in a reporter assay where an FT–GUS fusion results in significantly more visualized GUS protein after staining. The increased pigment accumulation indicates that higher amounts of the GUS–florigen fusion protein are circulating through the *Arabidopsis* leaf vasculature in plants with an intact enhancer region, compared to plants where the enhancer was deleted (Adrian et al. 2010). (B) The enhancer, *Region C*, is located downstream from the gene target, *LAS*. *LAS* is a gene involved in *Arabidopsis* organ boundary determination for leaf and reproductive axials. These findings demonstrate that *Region C* confers GUS–LAS fusion protein, as visualized by in situ hybridization, significantly higher accumulation of GUS–LAS protein in axial regions when compared to the *Region C* enhancer deletion mutant (Raatz et al. 2011). (C) Diagram showing the variable locations of Enhancers (Enh) with respect to the gene target. The TE is represented by a triangle, Enhancers are indicated by boxes labeled 'Enh', the promoter is represented by a pointed rectangle with the arrow indicating the transcription directionality. The gene target is a box labeled 'Gene' with the introns represented as striped boxes, and the transcriptional terminator indicated as a box labeled 'T'. (D) In some cases, TE insertions provide enhancer elements that will impact plant phenotypes. In the case of apples, it was found that a retrotransposon insertion provided an enhancer element, *RedTE*, that positively impacted gene expression of the *M. domestica* (apple; *Md*) *MYB1* gene. In apple trees, where the retrotransposon-inserted enhancer is present, the skin of apples is red due to high levels of pigment accumulation from higher levels of anthocyanin production controlled through *MbMYB1* expression. However, in trees where the *RedTE* enhancer is not present, the skin of the apples remains green (Zhang et al. 2019). (E) Some enhancers are located within introns of target genes. The enhancer, *DHS7*, upregulated the *TRY* gene of *Arabidopsis* impacting trichome formation. When the enhancer is present, trichomes form normally producing three branches, but when the enhancer is deleted, the trichomes develop abnormally, forming four to five branches (Meng et al. 2021). Created in BioRender. Beernink, B. (2024) BioRender.com/c53u001.

The Vital Role of Enhancers in Plant Development and Response to Environmental Cues

Plant development

Plant enhancers play a critical role during vegetative and reproductive development by tightly regulating the expression of developmental genes. For example, Adrian et al. (2010) determined that the transition of *Arabidopsis* from vegetative to reproductive development is controlled by an enhancer, *Block C*, located 5 kb upstream of the gene, *FT*, that produces the flowering signal florigen. When the *Block C* enhancer is deleted, the translational fusion of GUS and *FT* results in visually reduced protein accumulation in the leaf vasculature of the plant, indicating that florigen is suppressed. Significantly reduced expression of *FT* due to the *Block C* enhancer deletion also resulted in delaying flowering time (**Fig. 3A**).

Similarly, in situ hybridization revealed that deletion of the enhancer, *Region C*, in *Arabidopsis* silences the gene, *LATERAL SUPPRESSOR* (*LAS*). *LAS* is the main regulator of axillary meristems that determine the shape and size of leaves and petals by defining the boundary of these organs. Deletion analysis revealed that the downstream enhancer, *Region C*, contains the enhancer element impacting *LAS* gene transcription, demonstrating that this conserved enhancer modulates gene expression at organ boundaries to orchestrate organ development (Raatz et al. 2011; **Fig. 3B**).

Interestingly, TEs have been found to contain CRMs, including motifs for TF binding, that act as enhancers for nearby genes (Karttunen et al. 2023). For example, a retrotransposon insertion containing an enhancer approximately 4 kb upstream of the *Malus domestica* (apple) *MYB1* gene contains TF-binding domains that impact the fruit color by upregulating the genes that produce anthocyanins. Apple varieties lacking this enhancer, *RedTE*, are green rather than red because anthocyanins are not produced in the fruit skin (Zhang et al. 2019; **Fig. 3D**).

Gene editing was used to disrupt an intronic enhancer revealing its role in tissue-specific gene expression in *Arabidopsis*. When the enhancer, *DHS7*, was deleted and its two TF-binding sites were lost, plants flowered later than their wild-type counterparts due to suppression of *FT* expression. Additionally, *DHS7* mutants had more trichome branches, indicating that this enhancer also regulates trichome development (Meng et al. 2021; **Fig. 3E**).

Several other examples of enhancers impacting plant development have been reported. A maize *Hydroxyproline-rich glycoprotein* (*HRGP*) enhancer regulates the *HRGP* gene, a cell wall protein, highly expressed in plumules, the embryonic shoots that emerge from the seed during germination (Valles et al. 1991). Maize inflorescence branching and length are regulated by the *KRN4* enhancer of the *UB3* gene (Du et al. 2020). The distal enhancer, *DICE*, impacts later developmental stages in the maize by enhancing the *BENZOXAZINLESS* gene (Zheng et al. 2015). In *Arabidopsis*, the *Egg apparatus-specific enhancer* impacts female gametophyte development (Yang et al. 2005), and the *MMVE1* enhancer promotes phloem mobile signals that lead to early flowering (McGarry and Ayre 2008). The *Vtg1* enhancer, a miniature TE insertion, upregulates the *Arabidopsis* gene, *RELATED TO APETALA2*, associated with early flowering (Salvi et al. 2007, Castelletti et al. 2014). Enhancers were identified and characterized as regulating the *Arabidopsis ETTIN* gene controlling gynoecium development, and the *BLADE ON PETIOLE2* gene controls floral organ development (Yan et al. 2019). A more comprehensive list of characterized plant enhancers is covered in the Weber et al. (2016) review and in **Table 1**.

Plant response to environmental cues

Gene expression has been found to vary in response to environmental stresses, such as cold, drought, salt, heat, light, nutrient and biotic stressors. Previous studies have identified many TFs and TF-binding sites involved in abiotic stress responses. For example, *Drought-responsive element* (*DRE*)/*C-repeat* (*CRT*) was the first low-temperature-responsive *CRE* identified in *Arabidopsis* by Yamaguchi-Shinozaki and Shinozaki (1994) (Baker et al. 1994). Other motifs associated with cold-induced genes include the myelocytomatosis viral oncogene homolog and myeloblastosis viral oncogene homolog (*MYB*) binding motif, identified in *Arabidopsis* in response to dehydration and also chilling and freezing (Abe et al. 2003, Agarwal et al. 2006). The CREs, *DRE* and *CRT*, with a similar 9-bp conserved sequence, are involved in drought stress response (Baker et al. 1994,

Yamaguchi-Shinozaki and Shinozaki 1994, 2006, Jiang et al. 1996, Stockinger et al. 1997, Thomashow 1999). Although many CREs regulating abiotic stress response have been identified, most are located in promoter regions. Recent studies on the genome-wide identification of ACRs in *Arabidopsis* during photomorphogenesis suggested that ACRs are enriched for specific TF-binding sites (Tantale et al. 2016). Profiling ACRs in *Arabidopsis* under heat, cold, salt or drought stresses identified examples of both constitutive and conditionally specific ACRs. In sorghum, drought-responsive accessible chromatin signatures were enriched for *ABSCISIC ACID RESPONSE ELEMENT*–binding sites (R. K. Parvathaneni *et al.*, submitted for publication). However, it is unknown how many of those ACRs are enhancers. Up to now, only five enhancers—*AB80* enhancer, *Enhancer-like element*, *rbcS-E9 enhancer-like sequence*, *Cab-1 enhancer-like sequence* and *TACPyAT repeats*—are characterized and suggested that they involved in light response (Simpson et al. 1985, Fluhr et al. 1986, Kuhlemeier et al. 1987, Nagy et al. 1987, van der Meer et al. 1990). While many putative enhancers and DNA regions that may play a role in plant responses to the environment have been identified, we are at the early stages of determining their actual function.

Genetic and Epigenetic Variations in Enhancers and Their Significance in Evolution and Adaptation

Sequence variation of enhancers and their evolution

Some variations in enhancers are closely related to plant domestication and adaptation. A classic example is the presence/absence variation in the enhancer, *Hopscotch1*, located 60 kb upstream of the *TB1* gene. The presence of *Hopscotch1* enhances the expression levels of its target gene *TB1* in domesticated maize and results in unbranched maize plants, compared to its wild relative teosinte (Studer et al. 2011). The maize *hepta-repeat b1* enhancer, located 100 kb upstream of the transcription start site of the *B1* gene (Stam et al. 2002b, Belele et al. 2013), is responsible for the dark purple color kernel and leaves in some maize varieties. Purple maize is adapted to growth at high elevations because its high anthocyanin levels result in high antioxidant, antiradical and antimutagenic activities (Mansilla et al. 2020). In *Arabidopsis lyrata*, enhancer variants result in increased basal expression of cold-responsive genes (Akama et al. 2014, He et al. 2016).

Understanding the evolutionary forces shaping *cis*-regulatory variation is key to determining their role in adaptation and identifying molecular targets of natural selection. de Meaux (2018) explained that both positive selection and purifying selection influence the evolutionary rates of *cis*-regulatory elements and amino acids. Specifically, positive selection plays a significant role in shaping the *cis*-regulatory variants' adaptation by modifying gene expression in a tissue- or environment-specific manner, as seen in the domestication of maize (Studer

et al. 2011) and traits like flower morphology and size in the model plant *Capsella rubella* (Steige et al. 2015). These trait variations have been linked to enriched genes related to essential adaptive functions such as metal homeostasis, cold response and climate adaptation (Krämer 2005, Akama et al. 2014, Lasky et al. 2014, He et al. 2016). Conversely, purifying selection plays a vital role in shaping *cis*-regulatory divergence by removing harmful *trans*-acting variants while allowing beneficial *cis*-regulatory variants to accumulate, particularly in genes under lower purifying selection pressure, contributing to complex evolutionary patterns across species like *Arabidopsis* and *Capsella* (He et al. 2012, 2016, Fyon et al. 2015, Steige et al. 2015, 2017). Since *cis*-regulatory mutations are directly linked to the coding sequence they regulate, analyzing both variations together can reveal the evolutionary forces involved (de Meaux 2018). For example, Joly-Lopez et al. (2020) found that around 2% of the rice genome showed weak negative selection, frequently at candidate regulatory sites, including a novel set of 1,000 potentially active enhancers.

Epigenetic variations in enhancers and their adaptations

DNA methylation. Epigenetic modifications including DNA methylation and histone modifications can affect how enhancers regulate their target genes. DNA methylation is usually associated with silencing gene expression in animals and plants (Law and Jacobsen 2010). Consistently, methylation of enhancers downregulates the expression of target genes (Zentner et al. 2011, Zhong et al. 2013). In plants, this is observed for DNA methylation at regulatory sequences of *FLOWERING WAGENINGEN* (*FWA*), *TOO MANY MOUTHS* (*TMM*) and *FT* in *Arabidopsis* (Kinoshita et al. 2007, Deng et al. 2014, Deng and Chua 2015), and the *PERICARP COLOR1* (*P1*) gene in maize (Sidorenko and Peterson 2001, Haring et al. 2010). *FWA* and *FT* are responsible for flowering time (Koornneef et al. 1991), and methylated enhancers upstream of these genes cause a late-flower phenotype (Kinoshita et al. 2007; Deng et al. 2014). Flowering time is critical for adaptation to their natural environment because flowering at the appropriate time is crucial for seed production and survival (Izawa 2007). Loss-of-function mutations in the *Arabidopsis* gene *TMM* result in a clustered stomata phenotype on cotyledons as does methylation of enhancers upstream of *TMM* (Kinoshita et al. 2007, Deng et al. 2014, Deng and Chua 2015). The formation of stomata clusters was posited to reflect an adaptation to dry environments aimed at reducing water loss from plant leaves (Franks and Casson 2014). The maize *P1* gene encodes a *myb*-homologous transcriptional regulator of biosynthesis of a red phlobaphene pigment that accumulates in husks and floral organs, including silks, kernel pericarp, cob and tassel glumes (Lechelt et al. 1989). Increased methylation of the enhancers upstream of the *P1* gene causes reduced color in the maize kernels (Sidorenko and Peterson 2001). Although evidence suggests that red kernels are associated with corn adaptation, variation in this trait offers breeders the opportunity to select

varieties with different degrees of redness to satisfy human preferences.

Box 1. The definition of acronyms and terminologies in this review

---

CRE: *cis*-regulatory elements. CREs are specific DNA sequence motifs recognized by RNA polymerase, TFs and other regulatory proteins, often found clustered together to form *cis*-regulatory modules (CRMs).

CRMs: *cis*-regulatory modules. C*is*-regulatory elements are often clustered together to form CRMs. CRMs can be categorized into distinct regulatory domains based on their function, including core promoters, enhancers, silencers and insulators.

TF: transcription factors. TFs are proteins possessing domains that bind to the DNA of promoter or enhancer regions and interact with RNA polymerase or other TFs, consequently regulating gene expression.

ACRs: Accessible chromatin regions. ACRs are areas of chromatin that is loosened or open and accessible or available for binding by TFs, cofactors, polymerases and other proteins.

Promoter: Promoters are identified as an upstream *cis*-regulatory element containing specific nucleotide sequences that RNA polymerases and TFs can bind.

Core promoter: The core promoter is the minimal sequence required to bind an RNA polymerase and initiate transcription typically spanning 50–100 bp upstream of the transcription starting site.

Enhancer: An enhancer refers to a CRM bound by additional TFs and cofactors capable of boosting transcription rates.

Silencer: A silencer is a CRM bound by TFs and associated cofactors to actively repress the expression of their target genes.

Insulator: An insulator is a CRM bound by TFs and associated cofactors, which prevents other CRM activation or silencing of target genes when positioned between other CRMs.

CNSs: Conserved non-coding sequences. CNSs refers to a DNA sequence of non-coding DNA that is evolutionarily conserved. These sequences are of interest for their potential to regulate gene expression.

Quantitative trait locus (QTL) analysis: QTL analysis is a statistical method that links two types of information—phenotypic data (trait measurements) and genotypic data (usually molecular markers)—in an attempt to explain the genetic basis of variation in complex traits.

Histone modifications. DNA in eukaryotic nuclei is packaged into nucleosomes, composed of two turns of DNA wound around a histone octamer complex containing two molecules each of histones H2A, H2B, H3 and H4 (Wolffe and Hayes 1999). How tightly DNA is packed around histone octamers determines its accessibility to transcription machinery. Histone modifications (e.g. methylation and acetylation) influence the density of histone packing and are thus essential for gene regulation via modulation of chromatin accessibility (Bannister and Kouzarides 2011).

Nucleosomes at enhancers have been shown to carry specific histone marks. For example, H3K4me and H3K27ac are characteristic active enhancers in animals (Banerji et al. 1981, Creyghton et al. 2010, Yan et al. 2018, Barral and Déjardin 2023).

However, which histone marks correlate with enhancer activity in plants is not yet clear. Current research suggests that active plant enhancers are generally associated with H3 and H4 K acetylation, while inactive enhancers appear to be associated with H3K27me3 (Weber et al. 2016, Tan et al. 2023). For example, the active *Pisum sativum* (pea) *PetE* and maize *b1* enhancers were reported to be enriched in H3/H4ac and H3K9/K14ac, respectively (Chua et al. 2003, Haring et al. 2010). The *PET E* gene encodes the plastocyanin protein, which transfers electrons from the CYTOCHROME F protein in the CYTOCHROME BF protein complex to the P700 reaction center of photosystem I. In pea, the *PET E* gene is expressed only in photosynthetic tissues, and its transcription is activated by light (Last and Gray 1989, Chua et al. 2001). *PET E* is essential for plants to adapt to light, and activation of its enhancer is associated with H3/H4ac and H3K9/K14ac. So, levels of H3/H4ac and H3K9/K14ac in the enhancer of *PET E* could affect plants' adaptation to light.

The effect of histone marks varies among plant species. For instance, H3K27ac serves as an active enhancer mark in rice (Du et al. 2013) and maize (Zhang et al. 2015) but not in *Arabidopsis* (Yan et al. 2019). Intergenic nucleosome-depleted regions in rice, usually active promoter and enhancer regions, are strongly associated with H4K12ac and H3K27me3 (Zhang et al. 2012a). In contrast, in *Arabidopsis*, inactive enhancers are positively correlated with H3K27me3 (Zhu et al. 2015), and H3K27ac is not a hallmark for active enhancers (Yan et al. 2019) as data from rice (Du et al. 2013) and maize (Zhang et al. 2015) indicate. However, this variation could also be due to varying data quality across different studies and differences in the authors' interpretation of results. Thus, additional studies are required to fully understand the effect of histone modifications on enhancer activity in plants.

Future Challenges and Perspectives

Despite the critical role enhancers play in regulating plant gene expression, relatively few plant enhancers have been validated and functionally characterized. Many challenges remain for the discovery, characterization, functional validation of enhancers and identification of enhancers' target genes. The initial challenge is to improve the identification of putative enhancers within plant genomes. While several methods to predict sequences that may act as enhancers have been developed, they are all rather crude because they are based on general characteristics like chromatin accessibility or sequence conservation. Methods that integrate multiple properties of DNA across the genome may provide more accurate predictions and machine learning/artificial intelligence may be useful in this regard as larger datasets are developed.

A subsequent challenge is to functionally characterize putative enhancers within the context of the plant and/or genome at scale. Although massively parallel reporter assays, like Plant STARR-seq, can rapidly validate the activity of putative enhancers, these assays rely on the transient expression of a reporter gene under artificial conditions. Thus, they lack the genomic, environmental and developmental context necessary to demonstrate activity and understand the full regulatory dimensions of all enhancers. Enhancer deletions mediated by CRISPR/Cas modifications does not suffer from the contextual limitations of Plant STARR-seq, but generating edited plants is slow and labor-intensive. While many of the sequencing methods currently being used for putative enhancer identification are a vast improvement over previous methods, contextual limitations are still a challenge that needs to be addressed.

Identification of enhancer–target gene pairs poses additional difficulties for functional characterization of plant enhancers. In mammalian systems, enhancers regulate multiple genes, and genes are often controlled by multiple enhancers (Peng and Zhang 2018). While a few plant studies have determined enhancer–gene interactions, more research is needed (Zhao et al. 2022). Enhancers can act on their target genes from a distal location, although how this distance affects transcription in plants remains unclear. Enhancers can be upstream, downstream or within a gene's intronic region (**Figs. 1**, **3**); therefore, understanding how enhancers regulate adjacent versus distant genesis is crucial. It is also unknown if enhancers affect alternative splicing to favor specific isoforms. Recent findings suggest that operon-like enhancers exist in plants that coordinate the expression of gene clusters, but their full impact is still uncertain (Zhao et al. 2022).

While only a limited number of plant enhancers have been functionally characterized, it is clear that enhancers play important roles in plant development and responses to environmental stressors such as drought, salinity, light, nutrient and cold. However, much more detailed information is required to fully understand the biological roles of enhancers. Currently, researchers are employing deep learning models to efficiently design synthetic, cell-type-specific enhancers in flies, starting from random sequences, and this optimization process allows detailed tracing of enhancer features at single-nucleotide resolution (Taskiran et al. 2023). With more data on plant enhancers, this approach could likely be adapted for plants to fine-tune the expression in specific cells. Additionally, advances in single-cell genomics, such as single-cell ATAC-seq, will improve our understanding of the gene regulatory networks present in specific cells and how they impact plant responses to environmental and developmental cues.

Genetic variation and epigenetic variation in enhancers play a pivotal role in plant evolution and adaptation by influencing the expression of genes affecting traits adaptive to local environments or domestication. While epigenetic modifications have been shown to regulate enhancer activity and limited evidence suggests that enhancers are shaped by both positive selection and purifying selection, key questions remain: Which enhancers experience positive selection? Which undergoes purifying selection? Are all epigenetic modifications adaptive? Are species range limits and species-specific adaptation linked to enhancer epigenetic modifications? Understanding these dynamics will offer valuable insights into plant gene regulation, helping improve crop resilience and productivity.

While we are still in the early stages of understanding the biology and impact of enhancers, it is not too soon to think about how to apply our knowledge to crop improvement. Future applications might include (I) using genome editing to manipulate enhancers to improve plant performance (II) mining natural variation in enhancers in breeding programs including genomic selection. Additionally, enhancers could be used to fine-tune expression of transgenes in a cell-type-specific manner.

Data Availability

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

Funding

Author Contributions

B.M.B. and L.L. drafted the manuscript, and J.P.V. revised the manuscript. All of the authors contributed to writing and revising.

Disclosures

The authors have no conflicts of interest to declare.

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K., et al. (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* 15: 63–78.

Adrian, J., Farrona, S., Reimer, J.J., Albani, M.C., Coupland, G., Turck, F., et al. (2010) *cis*-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in *Arabidopsis*. *Plant Cell* 22: 1425–1440.

Agarwal, M., Hao, Y., Kapoor, A., Dong, C.-H., Fujii, H., Zheng, X., et al. (2006) A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *J. Biol. Chem*. 281: 37636–37645.

Akama, S., Shimizu-Inatsugi, R., Shimizu, K.K., Sese, J., et al. (2014) Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*. *Nucleic Acids Res*. 42: e46.

Andersson, R. and Sandelin, A. (2020) Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet*. 21: 71–87.

Arnold, C.D., Gerlach, D., Stelzer, C., Boryń, M.Ł., Rath, M. and Stark, A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339: 1074–1077.

Baker, S.S., Wilhelm, K.S. and Thomashow, M.F. (1994) The 5′-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol. Biol*. 24: 701–713.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27: 299–308.

Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res*. 21: 381–395.

Barral, A. and Déjardin, J. (2023) The chromatin signatures of enhancers and their dynamic regulation. *Nucleus* 14: 2160551.

Bartlett, A., O'Malley, R.C., Huang, -S.-S.C., Galli, M., Nery, J.R., Gallavotti, A., et al. (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc*. 12: 1659–1672.

Belele, C.L., Sidorenko, L., Stam, M., Bader, R., Arteaga-Vazquez, M.A. and Chandler, V.L. (2013) Specific tandem repeats are sufficient for paramutation-induced trans-generational silencing. *PLoS Genet*. 9: e1003773.

Benko, S., Fantes, J.A., Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., et al. (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nat. Genet*. 41: 359–364.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., et al. (2010) The NIH roadmap epigenomics mapping consortium. *Nat. Biotechnol*. 28: 1045–1048.

Bodega B. and Lanzuolo C. (eds.) (2021) *Capturing Chromosome Conformation. Methods and Protocols* Vol. 2157. New York, NY: Humana Press.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., et al. (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132: 311–322.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., et al. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523: 486–490.

Castelletti, S., Tuberosa, R., Pindo, M. and Salvi, S. (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL Vgt1. *G3 (Bethesda)* 4: 805–812.

Chua, Y.L., Brown, A.P.C. and Gray, J.C. (2001) Targeted histone acetylation and altered nuclease accessibility over short regions of the pea plastocyanin gene. *Plant Cell* 13: 599–612.

Chua, Y.L., Watson, L.A. and Gray, J.C. (2003) The transcriptional enhancer of the pea plastocyanin gene associates with the nuclear matrix and regulates gene expression through histone acetylation. *Plant Cell* 15: 1468–1479.

Chudalayandi, S. (2011) Enhancer trapping in plants. *Methods Mol. Biol*. 701: 285–300.

Clark, R.M., Wagler, T.N., Quijada, P. and Doebley, J. (2006) A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet*. 38: 594–597.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., et al. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A*. 107: 21931–21936.

Crocker, J., Abe, N., Rinaldi, L., McGregor, A.P., Frankel, N., Wang, S., et al. (2014) Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* 160: 191–203.

Dao, L.T.M., Galindo-Albarrán, A.O., Castro-Mondragon, J.A., Andrieu-Soler, C., Medina-Rivera, A., Souaid, C., et al. (2017) Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet*. 49: 1073–1081.

Davidson, E.H. (2006) The Regulatory Genome: Gene Regulatory Networks in Development and Evolution, Vol. 1. Academic Press, Cambridge, MA.

de Laat, W. and Duboule, D. (2013) Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502: 499–506.

de Meaux, J. (2018) *Cis*-regulatory variation in plant genomes and the impact of natural selection. *Am. J. Bot.* 105: 1788–1791.

Deng, K., Zhang, Q., Hong, Y., Yan, J. and Hu, X. (2023) iCREPCP: a deep learning-based web server for identifying base-resolution cis-regulatory elements within plant core promoters. *Plant Commun.* 4: 100455.

Deng, S. and Chua, N.-H. (2015) Inverted-repeat RNAs targeting FT intronic regions promote FT expression in *Arabidopsis*. *Plant Cell Physiol.* 56: 1667–1678.

Deng, S., Dai, H., Arenas, C., Wang, H., Niu, Q.-W. and Chua, N.-H. (2014) Transcriptional silencing of *Arabidopsis* endogenes by single-stranded RNAs targeting the promoter region. *Plant Cell Physiol.* 55: 823–833.

de Wit, E. and de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization. *Genes. Dev.* 26: 11–24.

Doni Jayavelu, N., Jajodia, A., Mishra, A. and Hawkins, R.D. (2020) Candidate silencer elements for the human and mouse genomes. *Nat. Commun.* 11: 1061.

Dorrity, M.W., Alexandre, C.M., Hamm, M.O., Vigil, A.-L., Fields, S., Queitsch, C., et al. (2021) The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution. *Nat. Commun.* 12: 3334.

Du, Y., Liu, L., Peng, Y., Li, M., Li, Y., Liu, D., et al. (2020) UNBRANCHED3 expression and inflorescence development is mediated by UNBRANCHED2 and the distal enhancer, KRN4, in Maize. *PLoS Genet.* 16: e1008764.

Du, Z., Li, H., Wei, Q., Zhao, X., Wang, C., Zhu, Q., et al. (2013) Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. Japonica. *Mol. Plant* 6: 1463–1472.

Fagny, M., Kujjer, M.L., Stam, M., Joets, J., Turc, O., Rozière, J., et al. (2021) Identification of key tissue-specific, biological processes by integrating enhancer information in maize gene regulatory networks. *Front. Genet.* 11: 606285.

Falconer, D.S. and Mackay T.F.C. (1996) *Introduction to Quantitative Genetics*. 4th edn. Longman Publishing Group, Essex.

Fluhr, R., Kuhlemeier, C., Nagy, F. and Chua, N.H. (1986) Organ-specific and light-induced expression of plant genes. *Science* 232: 1106–1112.

Franks, P.J. and Casson, S. (2014) Connecting stomatal development and physiology. *New Phytol.* 201: 1079–1082.

Fyon, F., Cailleau, A., Lenormand, T. and Zhang, J. (2015) Enhancer runaway and the evolution of diploid gene expression. *PLoS Genet.* 11: e1005665.

Gardner, M.J., Baker, A.J., Assie, J.-M., Poethig, R.S., Haseloff, J.P. and Webb, A.A.R. (2009) GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. *J. Exp. Bot.* 60: 213–226.

Giral, H., Landmesser, U. and Kratzer, A. (2018) Into the wild: GWAS exploration of non-coding RNAs. *Front. Cardiovasc. Med.* 5: 181.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R. and Lieb, J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.* 17: 877–885.

Göndör, A. and Ohlsson, R. (2018) Enhancer functions in three dimensions: beyond the flat world perspective. *F1000Res.* 7: F1000 Faculty Rev-618.

Gross, D.S. and Garrard, W.T. (1988) Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57: 159–197.

Hardison, R.C. and Taylor, J. (2012) Genomic approaches towards finding cis-regulatory modules in animals. *Nat. Rev. Genet.* 13: 469–483.

Haring, M., Bader, R., Louwers, M., Schwabe, A., van Driel, R. and Stam, M. (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J.* 63: 366–378.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C., Stam, M., et al. (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods* 3: 11.

He, F., Arce, A.L., Schmitz, G., Koornneef, M., Novikova, P., Beyer, A., et al. (2016) The footprint of polygenic adaptation on stress-responsive cis-regulatory divergence in the *Arabidopsis* genus. *Mol. Biol. Evol.* 33: 2088–2101.

He, F., Zhang, X., Hu, J., Turck, F., Dong, X., Goebel, U., et al. (2012) Genome-wide analysis of *cis*-regulatory divergence between species in the *Arabidopsis* genus. *Mol. Biol. Evol.* 29: 3385–3395.

He, H.H., Meyer, C.A., Hu, S.S., Chen, M.-W., Zang, C., Liu, Y., et al. (2014) Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. *Nat. Methods* 11: 73–78.

Hendelman, A., Zebell, S., Rodriguez-Leal, D., Dukler, N., Robitaille, G., Wu, X., et al. (2021) Conserved pleiotropy of an ancient plant homeobox gene uncovered by cis-regulatory dissection. *Cell* 184: 1724–1739.

Henikoff, S. (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* 9: 15–26.

Hesselberth, J.R., Chen, X., Zhang, Z., Sabo, P.J., Sandstrom, R., Reynolds, A.P., et al. (2009) Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nat. Methods* 6: 283–289.

Hong, R.L., Hamaguchi, L., Busch, M.A. and Weigel, D. (2003) Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *Plant Cell* 15: 1296–1309.

Hu, X., Fernie, A.R. and Yan, J. (2023) Deep learning in regulatory genomics: from identification to design. *Curr. Opin. Biotechnol.* 79: 102887.

Ibarra, I.L., Hollmann, N.M., Klaus, B., Augsten, S., Velten, B., Hennig, J., et al. (2020) Mechanistic insights into transcription factors cooperativity and its impact on protein-phenotype interactions. *Nat. Commun.* 11: 124.

Izawa, T. (2007) Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. *J. Exp. Bot.* 58: 3091–3097.

Jain, A., Anand, S., Singh, N.K. and Das, S. (2018) Sequence and functional characterization of MIRNA164 promoters from *Brassica* shows copy number dependent regulatory diversification among homeologs. *Funct. Integr. Genomics* 18: 369–383.

Jiang, C., Iu, B. and Singh, J. (1996) Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter *Brassica* napus. *Plant Mol. Biol.* 30: 679–684.

Jiang, J. (2015) The 'dark matter' in the plant genomes: non-coding and unannotated DNA sequences associated with open chromatin. *Curr. Opin. Plant Biol.* 24: 17–23.

Jindal, G.A. and Farley, E.K. (2021) Enhancer grammar in development, evolution, and disease: dependencies and interplay. *Dev. Cell* 56: 575–587.

Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316: 1497–1502.

Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., et al. (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* 527: 384–388.

Joly-Lopez, Z., Platts, A.E., Gulko, B., Choi, J.Y., Groen, S.C., Zhong, X., et al. (2020) An inferred fitness consequence map of the rice genome. *Nat Plants* 6: 119–130.

Jores, T., Hamm, M., Cuperus, J.T. and Queitsch, C. (2023) Frontiers and techniques in plant gene regulation. *Curr. Opin. Plant Biol.* 75: 102403.

Jores, T., Tonnies, J., Dorrity, M.W., Cuperus, J.T., Fields, S. and Queitsch, C. (2020) Identification of plant enhancers and their constituent elements by STARR-seq in tobacco leaves. *Plant Cell* 32: 2120–2131.

Jores, T., Tonnies, J., Mueth, N.A., Romanowski, A., Fields, S., Cuperus, J.T., et al. (2024) Plant enhancers exhibit both cooperative and additive interactions among their functional elements. *Plant Cell* 36: 2570–2586.

Jores, T., Tonnies, J., Wrightsman, T., Buckler, E.S., Cuperus, J.T., Fields, S., et al. (2021) Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat Plants* 7: 842–855.

Junion, G., Spivakov, M., Girardot, C., Braun, M., Gustafson, E.H., Birney, E., et al. (2012) A transcription factor collective defines cardiac cell fate and reflects lineage history. *Cell* 148: 473–486.

Juven-Gershon, T. and Kadonaga, J.T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* 339: 225–229.

Kadauke, S. and Blobel, G.A. (2009) Chromatin loops in gene regulation. *Biochim. Biophys. Acta* 1789: 17–25.

Karttunen, K., Patel, D., Xia, J., Fei, L., Palin, K., Aaltonen, L., et al. (2023) Transposable elements as tissue-specific enhancers in cancers of endodermal lineage. *Nat. Commun.* 14: 5313.

Kaufmann, K., Muiño, J.M., Østerås, M., Farinelli, L., Krajewski, P. and Angenent, G.C. (2010) Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nat. Protoc.* 5: 457–472.

Kearsey, M.J. (1998) The principles of QTL analysis (a minimal mathematics approach). *J. Exp. Bot.* 49: 1619–1623.

Kellis, M., Wold, B., Snyder, M.P., Bernstein, B.E., Kundaje, A., Marinov, G.K., et al. (2014) Defining functional DNA elements in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* 111: 6131–6138.

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W.J.J., Koornneef, M., et al. (2007) Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *Plant J.* 49: 38–45.

Klie, A., Laub, D., Talwar, J.V., Stites, H., Jores, T., Solvason, J.J., et al. (2023) Predictive analyses of regulatory sequences with EUGENe. *Nat. Comput. Sci.* 3: 946–956.

Koornneef, M., Hanhart, C.J. and van der Veen, J.H. (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 229: 57–66.

Krämer, U. (2005) MTP1 mops up excess zinc in *Arabidopsis* cells. *Trends Plant Sci.* 10: 313–315.

Kuhlemeier, C., Fluhr, R., Green, P.J. and Chua, N.H. (1987) Sequences in the pea rbcS-3A gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. *Genes. Dev.* 1: 247–255.

Kulkarni, M.M. and Arnosti, D.N. (2003) Information display by transcriptional enhancers. *Development* 130: 6569–6575.

Lasky, J.R., Des Marais, D.L., Lowry, D.B., Povolotskaya, I., McKay, J.K., Richards, J.H., et al. (2014) Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. *Mol. Biol. Evol.* 31: 2283–2296.

Last, D.I. and Gray, J.C. (1989) Plastocyanin is encoded by a single-copy gene in the pea haploid genome. *Plant Mol. Biol.* 12: 655–666.

Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11: 204–220.

Lechelt, C., Peterson, T., Laird, A., Chen, J., Dellaporta, S.L., Dennis, E., et al. (1989) Isolation and molecular analysis of the maize P locus. *Mol. Gen. Genet.* 219: 225–234.

Lettice, L.A., Horikoshi, T., Hearney, S.J.H., van Baren, M.J., van der Linde, H.C., Breedveld, G.J., et al. (2002) Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. U. S. A.* 99: 7548–7553.

Levy, S., Hannenhalli, S. and Workman, C. (2001) Enrichment of regulatory signals in conserved non-coding genomic sequence. *Bioinformatics* 17: 871–877.

Li, M. and Huang, -S.-S.C. (2022) DNA affinity purification sequencing (DAP-seq) for mapping genome-wide transcription factor binding sites in plants. *In Springer Protocols Handbooks.* Edited by Bilichak A. and Laurie J.D.. pp. 293–303. Springer US, New York, NY.

Li, X. and X-d, F. (2019) Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.* 20: 503–519.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326: 289–293.

Lin, Y., Meng, F., Fang, C., Zhu, B., and Jiang, J. (2019) Rapid validation of transcriptional enhancers using agrobacterium-mediated transient assay. *Plant Methods* 15: 21.

Liu, L., Gallager, J., Arevalo, E.D., Chen, R., Skopelitis, T., Wu, Q., et al. (2021b) Enhancing grain-yield-related traits by CRISPR-Cas9 promoter editing of maize CLE genes. *Nat. Plants* 7: 287–294.

Liu, L., Zhang, G., He, S., Hu, X. and Birol, I. (2021a) TSPTFBS: a Docker image for trans-species prediction of transcription factor binding sites in plants. *Bioinformatics* 37: 260–262.

Long, H.K., Prescott, S.L. and Wysocka, J. (2016) Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* 167: 1170–1187.

Louwers, M., Bader, R., Haring, M., van Driel, R., de Laat, W. and Stam, M. (2009b) Tissue- and expression level-specific chromatin looping at maize b1 epialleles. *Plant Cell* 21: 832–842.

Louwers, M., Splinter, E., van Driel, R., de Laat, W. and Stam, M. (2009a) Studying physical chromatin interactions in plants using Chromosome Conformation Capture (3C). *Nat. Protoc.* 4: 1216–1229.

Lovén, J., Hoke, H.A., Yin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., et al. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153: 320–334.

Lu, Z., Marand, A.P., Ricci, W.A., Ethridge, C.L., Zhang, X. and Schmitz, R.J. (2019) The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat Plants* 5: 1250–1259.

Lynch, M. and Walsh, B. (1998) *Genetics and Ananlysis of Quantitative Traits.* Sinauer Associates, Inc., Sunderland, MA.

Lynch, V.J., Leclerc, R.D., May, G. and Wagner, G.P. (2011) Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* 43: 1154–1159.

Mansilla, P.S., Nazar, M.C. and Pérez, G.T. (2020) Flour functional properties of purple maize (*Zea mays* L.) from Argentina. Influence of environmental growing conditions. *Int. J. Biol. Macromol.* 146: 311–319.

Marand, A.P., Chen, Z., Gallavotti, A. and Schmitz, R.J. (2021) A cis-regulatory atlas in maize at single-cell resolution. *Cell* 184: 3041–3055.e21.

Marand, A.P., Eveland, A., Kaufmann, K. and Springer, N.M. (2023) cis-Regulatory elements in plant development, adaptation, and evolution. *Annu. Rev. Plant Biol.* 74: 111–137.

McGarry, R.C. and Ayre, B.G. (2008) A DNA element between At4g28630 and At4g28640 confers companion-cell specific expression following the sink-to-source transition in mature minor vein phloem. *Planta* 228: 839–849.

Meng, F., Zhao, H., Zhu, B., Zhang, T., Yang, M., Li, Y., et al. (2021) Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in *Arabidopsis thaliana*. *Plant Cell* 33: 1997–2014.

Nagy, F., Boutry, M., Hsu, M.Y., Wong, M. and Chua, N.H. (1987) The 5′-proximal region of the wheat Cab-1 gene contains a 268-bp enhancer-like sequence for phytochrome response. *EMBO J.* 6: 2237–2542.

Ngan, C.Y., Wong, C.H., Tjong, H., Wang, W., Goldfeder, R.L., Choi, C., et al. (2020) Chromatin interaction analyses elucidate the roles of PRC2-bound silencers in mouse development. *Nat. Genet.* 52: 264–272.

Nishizaki, S.S. and Boyle, A.P. (2017) Mining the unknown: assigning function to noncoding single nucleotide polymorphisms. *Trends Genet.* 33: 34–45.

Ogbourne, S. and Antalis, T.M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* 331: 1–14.

Oka, R., Zicola, J., Weber, B., Anderson, S.N., Hodgman, C., Gent, J.I., et al. (2017) Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. *Genome Biol.* 18: 137.

O'Malley, R.C., Huang, -S.-S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., et al. (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165: 1280–1292.

Panne, D. (2008) The enhanceosome. *Curr. Opin. Struct. Biol.* 18: 236–242.

Panne, D., Maniatis, T. and Harrison, S.C. (2007) An atomic model of the interferon-β enhanceosome. *Cell* 129: 1111–1123.

Para, A., Li, Y., Marshall-Colón, A., Varala, K., Francoeur, N.J., Moran, T.M., et al. (2014) Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 111: 10371–10376.

Pehrsson, E.C., Choudhary, M.N.K., Sundaram, V. and Wang, T. (2019) The epigenomic landscape of transposable elements across normal human development and anatomy. *Nat. Commun.* 10: 5640.

Peleke, F.F., Zumkeller, S.M., Gültas, M., Schmitt, A. and Szymański, J. (2024) Deep learning the *cis*-regulatory code for gene expression in selected model plants. *Nat. Commun.* 15: 3488.

Peng, Y., Xiong, D., Zhao, L., Ouyang, W., Wang, S., Sun, J., et al. (2019) Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. *Nat. Commun.* 10: 2632.

Peng, Y. and Zhang, Y. (2018) Enhancer and super-enhancer: positive regulators in gene transcription. *Animal. Model. Exp. Med.* 1: 169–179.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A. and Bejerano, G. (2013) Enhancers: five essential questions. *Nat. Rev. Genet.* 14: 288–295.

Pwee, K.H. and Gray, J.C. (1993) The pea plastocyanin promoter directs cell-specific but not full light-regulated expression in transgenic tobacco plants. *Plant J.* 3: 437–449.

Raatz, B., Eicker, A., Schmitz, G., Fuss, E., Muller, D., Rossmann, S., et al. (2011) Specific expression of LATERAL SUPPRESSOR is controlled by an evolutionarily conserved 3′ enhancer. *Plant J.* 68: 400–412.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkova, I.D., Robinson, J.T., et al. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159: 1665–1680.

Ricci, W.A., Lu, Z., Ji, L., Marand, A.P., Ethridge, C.L., Murphy, N.G., et al. (2019) Widespread long-range cis-regulatory elements in the maize genome. *Nat. Plants* 5: 1237–1249.

Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171: 470–480.e8.

Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., et al. (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. U. S. A.* 104: 11376–11381.

Schaffner, W. (2015) Enhancers, enhancers—from their discovery to today's universe of transcription enhancers. *Biol. Chem.* 396: 311–327.

Schmitz, R.J., Grotewold, E. and Stam, M. (2022) Cis-regulatory sequences in plants: their importance, discovery, and future challenges. *Plant Cell* 34: 718–741.

Shen, W., Pan, J., Wang, G. and Li, X. (2021) Deep learning-based prediction of TFBSs in plants. *Trends Plant Sci.* 26: 1301–1302.

Shlyueva, D., Stampfel, G. and Stark, A. (2014) Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15: 272–286.

Sidorenko, L.V. and Peterson, T. (2001) Transgene-induced silencing identifies sequences involved in the establishment of paramutation of the maize p1 gene. *Plant Cell* 13: 319–335.

Sielemann, J., Wulf, D., Schmidt, R. and Bräutigam, A. (2021) Local DNA shape is a general principle of transcription factor binding specificity in *Arabidopsis thaliana*. *Nat. Commun.* 12: 6549.

Simpson, J., Timko, M.P., Cashmore, A.R., Schell, J., Montagu, M.V. and Herrera-Estrella, L. (1985) Light-inducible and tissue-specific expression of a chimaeric gene under control of the 5′-flanking sequence of a pea chlorophyll a/b-binding protein gene. *EMBO J.* 4: 2723–2729.

Singer, S.D., Liu, Z. and Cox, K.D. (2012) Minimizing the unpredictability of transgene expression in plants: the role of genetic insulators. *Plant Cell Rep.* 31: 13–25.

Song, L., Zhang, Z., Grasfeder, L.L., Boyle, A.P., Giresi, P.G., Lee, B.-K., et al. (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* 21: 1757–1767.

Song, X., Meng, X., Guo, H., Cheng, Q., Jing, Y., Chen, M., et al. (2022) Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* 40: 1403–1411.

Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C. (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1: 2019–2025.

Spitz, F. and Furlong, E.E.M. (2012) Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13: 613–626.

Stam, M., Belele, C., Dorweiler, J.E. and Chandler, V.L. (2002a) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes. Dev.* 16: 1906–1918.

Stam, M., Belele, C., Ramakrishna, W., Dorweiler, J.E., Bennetzen, J.L. and Chandler, V.L. (2002b) The regulatory regions required for B′ paramutation and expression are located far upstream of the maize b1 transcribed sequences. *Genetics* 162: 917–930.

Steige, K.A., Laenen, B., Reimegård, J., Scofield, D.G. and Slotte, T. (2017) Genomic analysis reveals major determinants of *cis*-regulatory variation in *Capsella grandiflora*. *Proc. Natl. Acad. Sci. U. S. A.* 114: 1087–1092.

Steige, K.A., Reimegård, J., Koenig, D., Scofield, D.G. and Slotte, T. (2015) *Cis*-regulatory changes associated with a recent mating system shift and floral adaptation in *Capsella*. *Mol. Biol. Evol.* 32: 2501–2514.

Stockinger, E.J., Gilmour, S.J. and Thomashow, M.F. (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc. Natl. Acad. Sci. U. S. A.* 94: 1035–1040.

Studer, A., Zhao, Q., Ross-Ibarra, J. and Doebley, J. (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat. Genet.* 43: 1160–1163.

Sun, J., He, N., Niu, L., Huang, Y., Shen, W., Zhan, Y., et al. (2019) Global quantitative mapping of enhancers in rice by STARR-seq. *GPB* 17: 140–153.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D., Dean, C., et al. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes. Dev.* 9: 1797–1810.

Tan, Y., Yan, X., Sun, J., Wan, J., Li, X., Huang, Y., et al. (2023) Genome-wide enhancer identification by massively parallel reporter assay in *Arabidopsis*. *Plant J.* 116: 234–250.

Tantale, K., Mueller, F., Kozulic-Pirher, A., Lesne, A., Victor, J.-M., Robert, M.-C., et al. (2016) A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* 7: 12248.

Taskiran, I.I., Spanier, K.I., Dickmänken, H., Kempynck, N., Pančíková, A., Ekşi, E.C., et al. (2023) Cell-type-directed design of synthetic enhancers. *Nature* 626: 212–220.

Thomas, H.F. and Buecker, C. (2023) What is an enhancer? *BioEssays* 45: e2300044.

Thomashow, M.F. (1999) PLANT COLD ACCLIMATION: freezing tolerance genes and regulatory mechanisms. *Annu. Rev. Plant Mol. Biol.* 50: 571–599.

Timko, M.P., Kauch, A.P., Castresana, C., Fassler, J., Herrera-Estrella, L., Van der Broeck, G., et al. (1985) Light regulation of plant gene expression by an upstream enhancer-like element. *Nature* 318: 579–582.

Uhl, J.D., Zandvakili, A., Gebelein, B. and Stathopoulos, A. (2016) A hox transcription factor collective binds a highly conserved distal-less cis-regulatory module to generate robust transcriptional outcomes. *PLoS Genet.* 12: e1005981.

Valles, M.P., Bernues, J., Azorin, F. and Puigdomènech, P. (1991) Nuclease sensitivity of a maize HRGP gene in chromatin and in naked DNA. *Plant Sci.* 78: 225–230.

van der Meer, I.M., Spelt, C.E., Mol, J.N. and Stuitje, A.R. (1990) Promoter analysis of the chalcone synthase (chsA) gene of *Petunia hybrida*: a 67 bp promoter region directs flower-specific expression. *Plant Mol. Biol.* 15: 95–109.

Van de Velde, J., Van Bel, M., Vaneechoutte, D. and Vandepoele, K. (2016) A collection of conserved noncoding sequences to study gene regulation in flowering plants. *Plant Physiol.* 171: 2586–2598.

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

Vockley, C.M., McDowell, I.C., D'Ippolito, A.M. and Reddy, T.E. (2017) A long-range flexible billboard model of gene activation. *Transcription* 8: 261–267.

Wallace, J.G., Bradbury, P.J., Zhang, N., Gibon, Y., Stitt, M. and Buckler, E.S. (2014) Association mapping across numerous traits reveals patterns of functional variation in maize. *PLoS Genet.* 10: e1004845.

Wang, S., Zhang, Q., Shen, Z., He, Y., Chen, Z.-H., Li, J., et al. (2021a) Predicting transcription factor binding sites using DNA shape features based on shared hybrid deep learning architecture. *Mol. Ther. Nucleic Acids* 24: 154–163.

Wang, X., Aguirre, L., Rodríguez-Leal, D., Hendelman, A., Benoit, M. and Lippman, Z.B. (2021b) Dissecting cis-regulatory control of quantitative trait variation in a plant stem cell circuit. *Nat Plants* 7: 419–427.

Weber, B., Zicola, J., Oka, R. and Stam, M. (2016) Plant enhancers: a call for discovery. *Trends Plant Sci.* 21: 974–987.

Wolffe, A.P. and Hayes, J.J. (1999) Chromatin disruption and modification. *Nucleic Acids Res.* 27: 711–720.

Wu, C., Li, X., Yuan, W., Chen, G., Kilian, A., Li, J., et al. (2003) Development of enhancer trap lines for functional analysis of the rice genome. *Plant J.* 35: 418–427.

Xie, M., Hong, C., Zhang, B., Lowdon, R.F., Xing, X., Li, D., et al. (2013) DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nat. Genet.* 45: 836–841.

Xie, Y., Chen, Y., Li, Z., Zhu, J., Liu, M., Zhang, Y., et al. (2022) Enhancer transcription detected in the nascent transcriptomic landscape of bread wheat. *Genome Biol.* 23: 109.

Xin, H., Liu, X., Chai, S., Yang, X., Li, H., Wang, B., et al. (2024) Identification and functional characterization of conserved cis-regulatory elements responsible for early fruit development in cucurbit crops. *Plant Cell* 36: 2272–2288.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994) A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* 6: 251–264.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu. Rev. Plant Biol.* 57: 781–803.

Yan, J., Chen, S.-A.-A., Local, A., Liu, T., Qiu, Y., Dorighi, K.M., et al. (2018) Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. *Cell Res.* 28: 204–220.

Yan, W., Chen, D., Schumacher, J., Durantini, D., Engelhorn, J., Chen, M., et al. (2019) Dynamic control of enhancer activity drives stage-specific gene expression during flower morphogenesis. *Nat. Commun.* 10: 1705.

Yan, W., Li, Z., Pian, C. and Wu, Y. (2022) PlantBind: an attention-based multi-label neural network for predicting plant transcription factor binding sites. *Brief. Bioinform.* 23: bbac425.

Yang, W., Jefferson, R.A., Huttner, E., Moore, J.M., Gagliano, W.B. and Grossniklaus, U. (2005) An egg apparatus-specific enhancer of *Arabidopsis*, identified by enhancer detection. *Plant Physiol.* 139: 1421–1432.

Zentner, G.E., Tesar, P.J. and Scacheri, P.C. (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* 21: 1273–1283.

Zhang, F. and Lupski, J.R. (2015) Non-coding genetic variants in human disease. *Hum. Mol. Genet.* 24: R102–10.

Zhang, L., Hu, J., Han, X., Li, J., Gao, Y., Richards, C.M., et al. (2019) A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nat. Commun.* 10: 1494.

Zhang, W., Garcia, N., Feng, Y., Zhao, H. and Messing, J. (2015) Genomewide histone acetylation correlates with active transcription in maize. *Genomics* 106: 214–220.

Zhang, W., Wu, Y., Schnable, J.C., Zeng, Z., Freeling, M., Crawford, G.E., et al. (2012a) High-resolution mapping of open chromatin in the rice genome. *Genome Res.* 22: 151–162.

Zhang, W., Zhang, T., Wu, Y. and Jiang, J. (2012b) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* 24: 2719–2731.

Zhang, Y., Tang, M., Huang, M., Xie, J., Cheng, J., Fu, Y., et al. (2022) Dynamic enhancer transcription associates with reprogramming of immune genes during pattern triggered immunity in *Arabidopsis*. *BMC Biol.* 20: 165.

Zhao, H., Yang, M., Bishop, J., Teng, Y., Cao, Y., Beall, B.D., et al. (2022) Identification and functional validation of super-enhancers in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 119: e2215328119.

Zhao, H., Zhang, W., Chen, L., Wang, L., Marand, A.P., Wu, Y., et al. (2018) Proliferation of regulatory DNA elements derived from transposable elements in the maize genome. *Plant Physiol.* 176: 2789–2803.

Zheng, L., McMullen, M.D., Bauer, E., Schön, C.-C., Gierl, A. and Frey, M. (2015) Prolonged expression of the BX1 signature enzyme is associated with a recombination hotspot in the benzoxazinoid gene cluster in *Zea mays*. *J. Exp. Bot.* 66: 3917–3930.

Zhong, S., Fei, Z., Chen, Y.-R., Zheng, Y., Huang, M., Vrebalov, J., et al. (2013) Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* 31: 154–159.

Zhu, B., Zhang, W., Zhang, T., Liu, B. and Jiang, J. (2015) Genome-wide prediction and validation of intergenic enhancers in *Arabidopsis* using open chromatin signatures. *Plant Cell* 27: 2415–2426.

Zhu, X., Chen, A., Butler, N.M., Zeng, Z., Xin, H., Wang, L., et al. (2024) Molecular dissection of an intronic enhancer governing cold-induced expression of the vacuolar invertase gene in potato. *Plant Cell* 36: 1985–1999.