CellPress

Review

Plant Enhancers: A Call for Discovery

Blaise Weber,[1,3] Johan Zicola,[2,3] Rurika Oka,[1,3] and Maike Stam[1,*]

Higher eukaryotes typically contain many different cell types, displaying different cellular functions that are influenced by biotic and abiotic cues. The different functions are characterized by specific gene expression patterns mediated by regulatory sequences such as transcriptional enhancers. Recent genome-wide approaches have identified thousands of enhancers in animals, reviving interest in enhancers in gene regulation. Although the regulatory roles of plant enhancers are as crucial as those in animals, genome-wide approaches have only very recently been applied to plants. Here we review characteristics of enhancers at the DNA and chromatin level in plants and other species, their similarities and differences, and techniques widely used for genome-wide discovery of enhancers in animal systems that can be implemented in plants.

Trends

Enhancers are one of the key elements in gene regulation in eukaryotes that allow correct temporal as well as tissue- and cell type-specific gene expression.

Thousands of enhancers have been discovered in animals, but only limited numbers are known in plants

Despite the limited number of features known for plant enhancers, they appear to share a number of common properties with the well-characterized animal enhancers. Plant-specific enhancer features are yet to be discovered.

The use of high-throughput sequencing-based methods enables the genome-wide discovery and characterization of plant enhancers.

Adaptation of powerful techniques such as STARR-seq, developed in the animal field, would greatly contribute to the identification and characterization of plant enhancers.

Enhancers in Gene Regulation

The vast majority of eukaryotes consist of numerous different cell types. In a given organism, the different cell types possess the same DNA, and it is fascinating that such diversity of cell types can arise from one and the same set of chromosomes. Cells of all organisms are in addition able to respond to abiotic and biotic environmental cues such as light, temperature, chemicals, and pathogens. The correct temporal and spatial regulation of gene expression is crucial for the successful production of highly specialized cell types and their response to external signals [1]. This is in large part accomplished through the activation and repression of the relevant ***cis*-regulatory elements** (see Glossary), such as **transcriptional enhancers** (hereafter referred to as enhancers) and **silencers**, at the correct moment in time and space [2,3]. Enhancers are non-coding DNA sequences that can be bound by multiple **transcription factors** (TFs) to activate the expression of genes located up to several Mb away (Figure 1A) [4,5]. Silencers are DNA elements that repress gene expression [3]. Both enhancers and silencers can be located up- or downstream of their target genes and function in an orientation-independent manner [6]. Enhancing and silencing functions can also be combined into one and the same DNA element, such as shown for the light-inducible and tissue-specific regulatory elements of *ab80* and *rbcS-3A* in pea (*Pisum sativum*) [7–10]. This review focuses on enhancers.

The general mechanisms by which enhancers are activated and trigger gene expression are well studied [11]. Enhancers are generally activated by the binding of pioneer TFs, followed by the recruitment of coactivators such as histone acetyltransferases and chromatin remodelers that together increase chromatin accessibility [12]. This increased accessibility promotes the binding of other TFs, leading to transcriptional activation of the target genes [12]. To do so, enhancers physically interact with the promoters of their cognate genes (Figure 1B). Ultimately, transcription is initiated by RNA polymerase II at the transcription start-site (TSS) of the gene [13].

In the past decades several examples of enhancers have been identified and characterized in different species, including yeast, fungi, animals, and plants (e.g., [14–21]). These examples have

[1]Swammerdam Institute for Life Sciences, Universiteit van Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[2]Max Planck Institute for Plant Breeding Research, Department Plant Developmental Biology, Carl-von-Linné-Weg 10, 50829 Köln, Germany
[3]These authors contributed equally to this work.

*Correspondence: m.e.stam@uva.nl (M. Stam).

 http://dx.doi.org/10.1016/j.tplants.2016.07.013

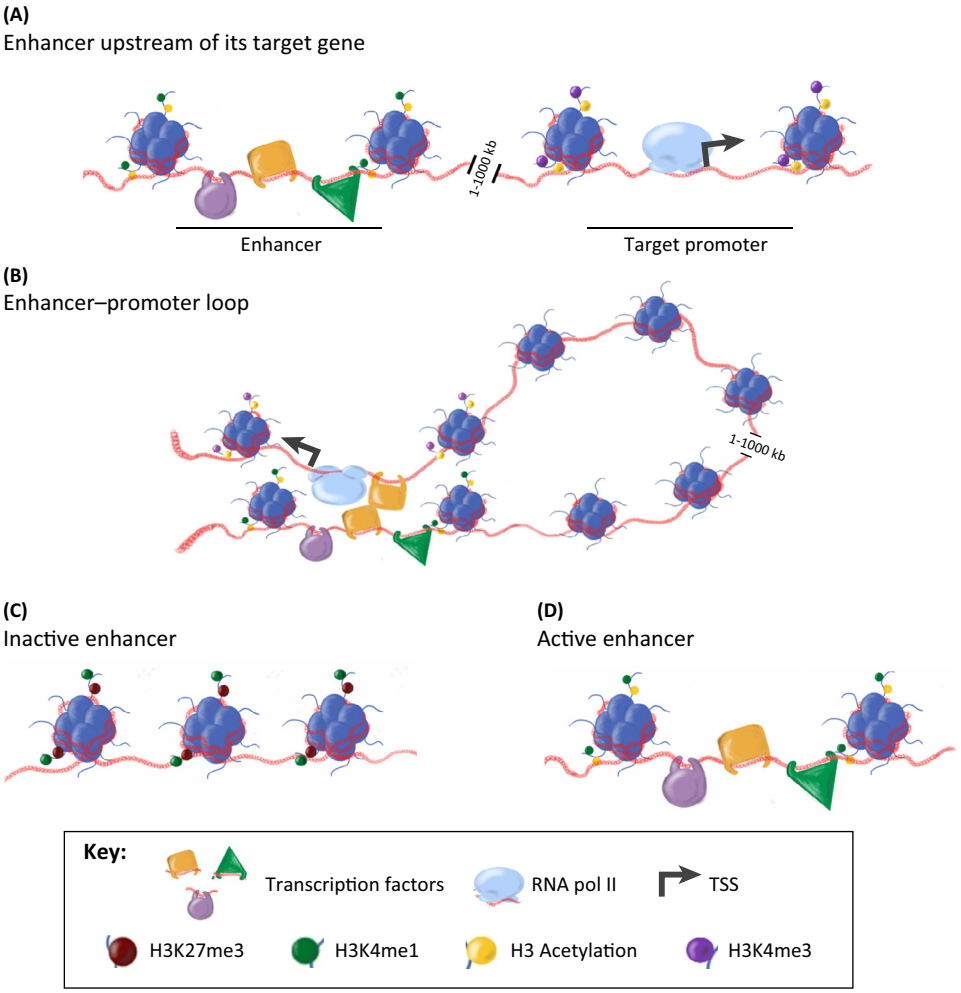

Figure 1. Schematic Illustrations of Chromatin Features and Associated Proteins Observed at Enhancer Regions and Target Genes in Animals.

For a Figure360 author presentation of Figure 1, see the figure online at http://dx.doi.org/10.1016/j.tplants.2016.07.013#mmc1.

(A) Enhancer located at a distance from its target gene. The presence of H3K4me1 and absence of H3K4me3 distinguishes enhancers from promoters. (B) Active enhancers physically interact with the promoter of their target gene through protein complexes. (C) Inactive enhancers are associated with H3K27me3 and H3K4me1. (D) Active enhancers are associated with nucleosome-depleted regions as well as H3K4me1, H3K9ac, and H3K27ac (annotated as H3 acetylation). Abbreviation: TSS, transcription start-site.

mainly been identified using low-throughput methods such as enhancer trapping, promoter deletion analysis, recombinant analysis, and quantitative trait locus mapping. The recent development of affordable next-generation sequencing technologies, in combination with the identification of general enhancer features, especially DNA and chromatin features, has allowed the genome-wide identification of enhancers in a high-throughput manner. This led to the discovery of over 43 000 enhancer candidates in the human genome [22] and up to 100 000 predicted enhancers in drosophila (*Drosophila melanogaster*) [23]. Remarkably, genes are often shown to be regulated by more than one enhancer [23–25]. The crucial roles of enhancers in gene regulation have been emphasized in studies linking enhancers not only with proper embryonic development and the specialization of cell types, but also with a large set of diseases

Glossary

Active enhancers: enhancers that are upregulating the expression of their target genes. They are located within accessible chromatin regions, are associated with activating histone marks and low levels of DNA methylation, and are bound by TFs.

Chromosome conformation capture (3C): 3C reveals the relative frequency of physical interactions between a given chromosomal fragment, called the bait or viewpoint, with other known fragments (one-to-one). Derivative techniques increase the number of detected interactions: (i) 4C (circular 3C) reveals all interactions of a given bait (one-to-all); (ii) 5C reveals all interactions for many baits (many-to-many); (iii) Hi-C reveals all interactions genome-wide (all-to-all).

***Cis*-regulatory elements:** non-coding DNA sequences that regulate gene expression by recruiting TFs. The elements can be located nearby or at a distance from their target genes.

DNA footprinting: method allowing the identification of protein binding sites using techniques such as DNAse I footprinting, DNase-seq, ChIP-seq or ATAC-seq. In combination with next-generation sequencing, footprinting allows to elucidate TF binding motifs. For example, when using DNase-seq for DNA footprinting, TF binding to DNA protects binding sites from DNase I cleavage. TF binding motifs can then be determined by sequence analysis of the protected fragments.

Inactive enhancers: enhancers that are silenced. They can be stably silenced or ready to be activated (also known as poised enhancers). Stably silenced enhancers are located in inaccessible chromatin regions and carry repressive histone marks (e.g., H3K9me2) and high DNA methylation. Poised enhancers are associated with both repressive (H3K27me3) and activating histone marks (e.g., H3ac), and display an increased level of accessibility compared to stably silenced enhancers.

Insulators: *cis*-regulatory elements that block the interaction between enhancers or silencers and non-target genes.

Position effect: the effect of the genomic location of an endogenous

including cancer [2,4,26–29]. Plant genomes are also very likely to contain numerous enhancers. So far, we do not know how many enhancers are present across plant species, and their chromatin features are poorly characterized, except for a few examples such as the enhancers of the *Pea plastocyanin* (*PetE*) gene in pea [30], the *booster1 (b1)* gene in maize (*Zea mays* ssp. *mays* L.) [14,31,32], and the enhancer of the *FLOWERING LOCUS T* (*FT*) gene in arabidopsis *(Arabidopsis thaliana)* [17,33] (Table 1). The first genome-wide study identifying enhancers in *Arabidopsis* based on chromatin features [34] reflects the renewed interest in plant enhancer elements.

What are the general properties of enhancers? Large-scale animal studies comparing several molecular features showed that, depending on their activity state, enhancers are characterized by different DNA and chromatin features [11]. **Inactive enhancers** are typified by low chromatin accessibility and the presence of specific histone modifications, for example trimethylation of lysine 27 of histone H3 (H3K27me3) (Figure 1C) [35,36]. **Active enhancers** generally display high chromatin accessibility and histone acetylation, enhancer RNA (eRNA) transcription, and low DNA methylation [22,35–40] (Figure 1D). Specific histone modifications, such as H3K4me1, mark enhancers independent of their activity level. Despite growing interest, so far no study has reported a comparison of different chromatin features of enhancers in plants.

In this review we (i) provide an overview on the current knowledge on enhancers in plants, including their molecular characteristics, (ii) discuss the potential commonalities and differences between plant and animal enhancers, and (iii) discuss and compare the different techniques available to identify and characterize enhancers in plants, focusing mainly on high-throughput methods based on next-generation sequencing approaches. Finally, we provide directions for future research.

Enhancers in Plants

One of the first enhancers described in plants dates back to 1985 when Simpson *et al*. [41] reported an enhancer of the chlorophyll *a/b*-binding protein gene *AB80* in pea. Since then, other enhancers have been identified and characterized in different plant species (examples are given in Table 1). At first, plant enhancers were primarily characterized using promoter deletion assays, electrophoretic mobility shift assays (EMSA), and DNAse I footprinting [15,42–44]; chromatin features were subsequently investigated as well. The enhancers of the *hydroxyproline-rich glycoprotein* (*HRGP*) gene in maize and the *PetE* gene in pea are among the first enhancers examined for accessible chromatin and histone acetylation, respectively [30,44]. Currently, a hepta-repeat region of the *b1* gene in maize is one of the enhancers for which the chromatin features are best-characterized. When active, the hepta-repeat displays several hallmarks of active enhancers in mammals, such as accessible chromatin, H3 acetylation, and low DNA methylation [14,31,32]. Most other plant enhancers are less well characterized, and the list of their associated characteristics is incomplete. Recent studies in *Arabidopsis* and rice (*Oryza sativa*) [34,45] are the first reporting the use of chromatin features to identify *cis*-regulatory elements in a genome-wide, high-throughput manner.

Characteristics of Enhancers

Enhancer regions display specific characteristics, including the presence of **TF binding motifs**, chromatin accessibility, particular histone modifications, eRNA expression, low DNA methylation, and physical interactions with their target genes [11]. Together, these signatures can help to identify enhancers in a genome. For better enhancer prediction and characterization, multiple features should be studied in parallel given that particular features can also be displayed by other *cis*-regulatory elements, TSSs or coding regions of genes. We discuss below enhancer characteristics identified in animals that can be used to identify and characterize plant enhancers.

or transgenic sequence on its transcriptional status.

Silencers: *cis*-regulatory elements that downregulate the expression of their target gene.

TF binding motifs: short consensus DNA sequences to which TFs bind specifically. A TF binding motif provides the frequency at which each of the four nucleotides appears at different positions within the binding sequence. The frequencies indicate the flexibility and ambiguity in a given binding sequence.

TF binding sites: locations in a genome to which a given TF binds. Often TF binding sites are identified by ChIP using an antibody recognizing TFs.

Transcriptional enhancer: a *cis*-regulatory element which upregulates the expression of its target gene. It is often simply referred to as an enhancer. In yeast, an enhancer is called an upstream activating sequence (UAS).

Transcription factors (TFs): proteins that are recruited to *cis*-regulatory elements and regulate the expression of target genes. The majority of TFs can only bind to accessible chromatin; however, a special class of TFs, called pioneer transcription factors, can bind to *cis*-regulatory elements that are not accessible, which results in opening up the region for other protein factors.

Table 1. Examples of Currently Known Plant Enhancers and Their Associated Characteristics

Enhancer	Target gene	Organism	Location[a]	Chromatin accessibility[b]	Chromatin interaction[c]	Sequence conservation among or within species	TF binding motif	Histone mark[d]	Reporter assay[e]	DNA methylation[f]	Refs
Hepta-repeat	*booster1 (b1)*(GRMZM2G172795)	*Zea mays*	100 kb upstream	Yes	Yes	Yes, among different maize inbred lines[g]	ND[h]	H3ac (A)	Yes	Yes	[14,31,32,121]
Block C	*FLOWERING LOCUS T (FT)* (AT1g654800)	*Arabidopsis thaliana*	5 kb upstream	Yes	Yes[i]	Yes, among *Brasicaceae*	CCAAT-box, REalpha, I-box	H3K9K14ac (A)	Yes	Yes[i]	[17,33,53,75]
Region C	*LATERAL SUPPRESSOR (LAS)* (AT1G55580)	*Arabidopsis thaliana*	3.2 kb downstream	Yes	ND	Yes, among *Arabis alpina, Arabidopsis lyrata* and *Capsella rubella*	ND	H3K27me3 (I)	Yes	ND	[128]
P268 *PetE* enhancer	*pea plastocyanin gene (PetE)*	*Pisum sativum*	177 bp upstream	Yes	ND	ND	HMG-I/Y binding motif	H3ac and H4ac (A)	Yes	ND	[30]
P1-rr distal enhancer	*pericarp color1 (p1)* (GRMZM2G084799)	*Zea mays*	6 kb upstream	Yes	ND	ND	ND	ND	Yes	Yes	[124,129]
Vegetative to generative1 (Vgt1)	*Related to APETALA2 (ZmRap2.7)* (GRMZM2G700665)	*Zea mays*	70 kb upstream	ND	ND	Conserved among late flowering lines	ND	ND	ND	Yes	[16,130]
AB80 enhancer	*AB80 chlorophyll a/b binding protein (CAB) gene*	*Pisum sativum*	400 bp upstream	ND	ND	Yes, among the *CAB* genes in *N. plumbaginifolia* and wheat	G-boxes, GATA-box	ND	Yes	ND	[41,123,131]
Enhancer-like element	*Ribulose 1,5-biphosphate carboxylase small subunit (rbcS) genes SS3.6, E9, 3A, 3C*	*Pisum sativum*	400 bp upstream	ND	ND	Yes, among *Pisum sativum*, *Solanum lycopersicum*, wheat, *Nicotiana plumbaginifolia*, and *Antirrhinum majus*	Box II: homology with SV40 core enhancer GT motif Box III: homology with human β-interferon enhancer and adenovirus 5 E1A enhancer	ND	Yes	ND	[8,9,15]
TACPyAT repeats	*Chalcone synthase A (chsA)*	*Petunia hybrida*	−67 to −53 bp upstream	ND	ND	Yes, with *Antirrhinum majus*	Two TACPyAT motifs	ND	Yes	ND	[132,133]
L3 enhancer	Putative: *AT-hook motif nuclear-localized protein 22 (AHL22)* (AT2G45430)	*Arabidopsis thaliana*	4 kb upstream	Yes	ND	ND	ND	ND	Yes	ND	[34]

Table 1. (continued)

Enhancer	Target gene	Organism	Location[a]	Chromatin accessibility[b]	Chromatin interaction[c]	Sequence conservation among or within species	TF binding motif	Histone mark[d]	Reporter assay[e]	DNA methylation[f]	Refs
Distal *cis*-element (DICE)	*benzoxazinless1 (bx1)* (GRMZM2G085381)	*Zea mays*	140 kb upstream	ND	ND	ND	ND	ND	ND	Low methylation	[134]
tb1 enhancer	*teosinte branched1 (tb1) (AC233950.1_FG002)*	*Zea mays*	60 kb upstream	ND	ND	ND	ND	ND	Yes	ND	[135,136]
Egg apparatus-specific enhancer (*EASE*)	ND	*Arabidopsis thaliana*	ND	Yes	ND	Yes, among different *Arabidopsis* accessions	ND	ND	Yes	ND	[21]
MATURE MINOR VEIN ELEMENT1 (*MMVE1*)	ND	*Arabidopsis thaliana*	ND	ND	ND	Yes, among *Brassicaceae*	API and ARF binding motifs, CACGTG motif	ND	Yes	ND	[20]
HRGP enhancer	*Hydroxyproline-rich glycoprotein (HRGP)*	*Zea mays*	1380 to 220 bp upstream	Yes	ND	ND	ND	ND	ND	ND	[44]

[a]Location of enhancer relative to the TSS of its target gene.
[b]Correlation between enhancer activity and chromatin accessibility.
[c]Physical interaction observed with target gene upon gene activation.
[d]Enrichment in histone marks: (A) for active enhancer, (I) for inactive enhancer.
[e]Activity of enhancer measured in reporter assay (e.g., transient or transgenic minimal reporter assay or enhancer trap).
[f]DNA methylation at enhancer is associated with silencing of the enhancer.
[g]Sequence is conserved, but number of repeats varies.
[h]ND, not determined.
[i]Exact locations of interactions do not agree between publications.
[j]J. Zicola and F. Turck, unpublished data

CellPress

TF Binding Motifs

Enhancers are activated by the binding of TFs (Figure 1). This binding of TFs to DNA is specified by specific consensus sequences, known as TF binding motifs, and/or particular chromatin features, such as histone modifications [46]. Enhancers are enriched with multiple TF binding motifs. Nearly 600 different experimentally validated TF binding motifs have been reported in human (Transfac in 2003 [47]). In *Arabidopsis*, approximately 530 TF binding motifs have been experimentally determined, for example by mobility shift assays or DAP-seq [48,49]. In other plant species, such as maize, many TF binding motifs were predicted [50]; however, few studies have validated TF binding motifs experimentally [50,51].

Chromatin Accessibility

The degree of chromatin accessibility impacts the binding of TFs to regulatory sequences [25,52] (Figure 1). Chromatin accessibility depends on the local nucleosome occupancy and binding of chromatin-associated proteins (Figure 1A,D). Active *cis*-regulatory elements such as promoters and enhancers are localized in accessible genomic regions, also known as nucleosome-depleted regions (NDRs) [25]. NDRs have been mapped genome-wide in *Arabidopsis*, maize, and rice [45,53,54]. Consistent with having a regulatory role, NDRs are enriched at **TF binding sites** and conserved non-coding sequences in *Arabidopsis* and maize [53,54]. Moreover, several intergenic NDRs identified in *Arabidopsis* were validated as enhancers in transgenic experiments [34].

Histone Modifications

Histone modifications or marks are post-translational modifications of histones that have different roles in gene regulation, including the modulation of chromatin accessibility [55] (Figure 1). Nucleosomes at enhancer regions have been shown to carry specific histone marks. In animals, H3K4me1 is found at both active and inactive enhancers [35]. Acetylation at lysine 9, 12, 14, and 27 of H3 (H3K9ac, H4K12ac, H3K14ac, and H3K27ac) characterizes active enhancers [37,38,56], while H3K27me3 marks inactive enhancers [35]. All of these marks are, however, present at TSSs and/or coding regions as well, hampering the unequivocal identification of enhancer sequences from a single histone mark. Data from an additional mark, for example H3K4me3, which is preferentially enriched at TSSs, can be used to distinguish TSSs from enhancers [57–59].

Which histone marks best indicate plant enhancers and their activity states is not yet entirely clear. Knowledge on such marks is slowly emerging. For instance, the active pea *PetE* and maize *b1* enhancers were reported to be enriched in H3/H4ac and H3K9/K14ac, respectively [30,31]. Furthermore, intergenic NDRs in rice were strongly associated with H4K12ac, but also H3K27me3 [45]. Moreover, a recent study in *Arabidopsis* revealed a positive correlation between inactive enhancers and H3K27me3, and between active enhancers and H3K27ac, with the former correlation being clearer than the latter [34]. Together, the current results indicate that active plant enhancers are generally associated with H3 and H4 acetylation, while inactive enhancers appear to be associated with H3K27me3. At the same time, we must emphasize that more research will be necessary to identify the histone modifications that best detect active and inactive enhancers in plants.

eRNAs

In animals, the presence of enhancer transcripts (eRNAs) have been shown to provide a good indication of active enhancers [22,40]. eRNAs are non-coding, relatively short (<2 kb), capped, mostly non-polyadenylated, and unspliced RNA, and are rapidly degraded by exosomes. Animal enhancers are often transcribed bidirectionally [22,40] and, although the absolute eRNA transcript levels are much lower than those of protein-coding genes, they correlate with those of their target genes [22]. Insight into potential roles for eRNAs is emerging. Some eRNAs are, for

example, necessary to recruit TFs to enhancers [60] or to mediate enhancer–promoter interactions [61]. It can, however, not be excluded that part of the eRNAs have no role in gene regulation and may be products of leaky RNA pol II expression [62–64]. Recent findings suggested a significant association of non-coding RNAs with NDRs in *Arabidopsis* [34]. The presence, characteristics, and roles of eRNAs in gene regulation in plants remains to be further investigated.

DNA Methylation

DNA methylation is associated with transcriptional silencing in both animals and plants [65], and, when present at enhancers, downregulates the expression of target genes [39,66]. In plants this is, for example, observed for DNA methylation at regulatory sequences of *FLOWERING WAGENINGEN* (*FWA*), *TOO MANY MOUTHS* (*TMM*), and *FT* in *Arabidopsis* [67–69], and *pericarp color1* (*p1*) and *b1* in maize [31,70]. In human and mouse, the DNA methylation level at numerous enhancers is dynamically regulated, negatively correlating with the activity of enhancers, allowing the identification of tissue-specific enhancers [39]. Except for a study in tomato [66], there is little evidence that DNA methylation at *cis*-regulatory elements is regulated in a dynamic manner in plants.

Chromatin Interactions

Enhancers and target genes must be in close proximity to allow enhancers to activate transcription (Figure 1B). Chromosomal conformation studies indeed provide ample evidence that enhancers and their target genes physically interact with each other [32,71]. In mammals, CTCF, cohesin, the mediator complex, and sometimes also eRNAs have been shown to mediate enhancer–promoter interactions [61,72,73]. In addition to general protein factors and protein complexes, sequence-specific TFs are also required for enhancer–promoter interactions. The erythroid Krüppel-like transcription factor (EKLF) is, for instance, involved in establishing chromatin interactions at the active β-globin locus [74]. Chromatin interactions between distant enhancers and their target genes are also reported in plants. In maize, the 100 kb upstream hepta-repeat enhancer interacts with the TSS region of the *b1* gene when *b1* is expressed [32]. Similarly in *Arabidopsis*, two independent studies revealed interactions between the TSS and upstream regulatory regions of *FT* [33,75]. The results, however, do not agree on the exact genomic identity of the interacting regions.

Techniques for Enhancer Identification

In the following section we review techniques that are currently used for identifying enhancers in a genome-wide, high-throughput manner, mainly using next-generation sequencing technologies. Techniques for validating enhancer candidates are also discussed. Each technique has inherent methodological biases and limitations; therefore, to predict enhancers with higher accuracy, combining different approaches is preferred over the use of a single method. For a summary, see Table 2.

Based on DNA Sequence: TF Binding Motif Scan

Enhancers are bound by TFs, therefore scanning genomes for TF binding motifs can contribute to the identification of enhancers [76]. However, the presence of a TF binding motif does not guarantee the functional binding of TFs *in vivo*, because TF binding motifs are typically less than 10 nt in length and therefore can appear in a genome by chance [77,78]. In addition, TF binding is not always highly sequence-specific [46] and may therefore be hard to predict by motif scanning. Because TFs often function in complexes, detecting clusters of TF binding motifs reduces the number of false positives [79]. In plants other than *Arabidopsis*, motif scanning is limited by the relatively low number of known TF binding motifs. In such case, putative TF binding motifs can be determined by **DNA footprinting** using DNase-seq, ChIP-seq, or by analyzing promoter sequences of coexpressed genes using multiple tissue or multiple time point data [50,51,80,81].

Table 2. Summary of Currently Used or Promising Techniques To Identify Enhancers

Technique	Studied aspect	Advantages	Disadvantages	Refs
TF binding motif scan	TF binding motifs	Identifies TF binding sites	High false-discovery rate; prior knowledge of TF binding motif required	[81]
DNase-seq	Open chromatin	TF binding motifs can be detected	DNase I can introduce cleavage bias, affecting TF binding motif detection	[45,53,80]
ChIP-seq	Histone modifications, TFs, chromatin-associated proteins	A wide range of targets can be studied	Relies on the availability of high-quality antibodies or tagged proteins	[89–91]
RNA-seq	Transcript levels	eRNA levels implicate enhancer activity, detects directionality of transcription	eRNA expression is low, high sequencing depth is required	[40]
CAGE	Transcript levels	eRNA levels implicate enhancer activity; detects directionality of transcription	Only detects capped eRNAs	[22]
GRO-seq	Nascent transcript levels	eRNA transcription implicates enhancer activity	Challenging technique	[93,94]
STARR-seq	Enhancer mapping (and activity)	High-throughput identification and validation of enhancers in parallel	Minimal promoter used influences the set of identified enhancers	[96,97]
BS-seq	DNA methylation	Single bp resolution	High sequence depth needed; incomplete BS conversion affects data interpretation	[66,98,103]
Enhancer trapping	Enhancer activity	Visualizes tissue-specific pattern mediated by endogenous *cis*-regulatory sequences	Difficult to locate trapped enhancers	[20,21,106,109]
3C technology	Chromatin interactions	Identifies promoter–enhancer interactions	Challenging technique; trade-off between number of observed interactions and resolution	[71,111,112,117]
Reporter assay	Transcriptional activity	Confirms activity and tissue specificity of enhancer candidates	Potential expression bias arising from test conditions and the minimal promoter used	[15,119,120,124,125]

Assaying Chromatin Accessibility

DNase-seq

Active enhancer sequences are usually located in NDRs and are therefore sensitive to nuclease activity [82]. Therefore, DNase-seq is a very valuable tool to identify *cis*-regulatory sequences. With this method, DNase I hypersensitive sites (DHSs) can be identified by partial digestion of chromatin with the endonuclease DNase I, followed by sequencing of the small fragments representing the accessible fraction of the genome (DNase-seq) [80]. Alternatively, the DNAse I-digested ends can be sequenced, followed by identification of DHSs [83]. DNase-seq robustly identifies DHSs, but is not very sensitive in predicting TF binding motifs because of the intrinsic cleavage bias of DNAse I [84].

ATAC-seq: Assay for Transposase-Accessible Chromatin

ATAC-seq, a technique in which the engineered transposase Tn5 ligates accessible DNA to sequencing adapters, was shown to provide a good alternative to DNase-seq in human [85]. Importantly, with ATAC-seq, highly comparable results to DNase-seq could be obtained using 200-fold fewer cells. Like DNase-seq, ATAC-seq can also be used for DNA footprinting [85]. ATAC-seq datasets have not yet been reported for plant tissue, but the method seems attractive, especially for analyzing tissues that are difficult to collect in large quantities.

FAIRE: Formaldehyde-Assisted Isolation of Regulatory Elements

Another method to identify accessible chromatin is FAIRE-seq [86]. FAIRE identifies protein-free DNA regions (i.e., free from nucleosomes) by crosslinking tissues or cells 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 provides higher background noise than DNase I digestion [87]. FAIRE-qPCR applied at the maize *b1* locus revealed FAIRE enrichment at the active hepta-repeat enhancer, demonstrating the potential of FAIRE to identify plant enhancers [32]. A FAIRE-seq protocol was developed for *Arabidopsis* but no genome-wide FAIRE study has yet been reported in plants [88].

ChIP-seq: Chromatin Immunoprecipitation Sequencing

ChIP-seq can identify DNA regions based on their associated modifications or proteins (e.g., histone marks, TFs, and polymerases) [11,89,90]. For instance, in animals, antibodies recognizing H3K27ac and the histone acetyltransferase p300 can be used to detect active enhancers [38,91]. The modifications or proteins targeted by ChIP influence the number and types of enhancers identified. To identify TF binding sites with almost base-pair resolution, ChIP can be coupled to an exonuclease treatment (ChIP-exo; exonucleases remove DNA not bound by TFs) [92]. ChIP has been adapted for use in plants [89,90]. However, the most relevant combinations of histone marks or TFs for enhancer identification remain to be determined [31,34,45].

Assaying Transcriptional Activity

RNA-seq BasedMethods

In animals, the production of eRNAs provides a good indication of enhancer activity [22]. eRNAs can be identified by different techniques. When sequencing RNAs (RNA-seq), a high sequencing depth is required to detect the low-abundant eRNAs [40]. By using cap analysis of gene expression (CAGE), where only the 5' ends of RNAs are sequenced, eRNAs can be detected at lower sequencing depth [22]. In addition to being low abundant, eRNAs are sensitive to degradation; therefore, genome-wide nuclear run-on assays (GRO-seq), which measure nascent transcript production, may provide higher sensitivity to detect eRNAs than CAGE [93,94]. The recently developed GRO-cap technique, which allows the detection of nascent capped transcripts, maps TSSs with higher accuracy than GRO-seq and could be an interesting technique to apply in plants for eRNA detection [95].

STARR-seq: Self-Transcribing Active Regulatory Region Sequencing

STARR-seq is a technique developed in *Drosophila* to capture sequences with enhancer activity [96]. With this technique, random fragments from sheared genomic DNA are cloned between a minimal promoter incapable of driving high expression, and a polyadenylation sequence. The resulting plasmids are transfected into cells, after which fragments with enhancer activity can enhance their own transcription. Hence, sequencing of polyadenylated transcripts isolated from the transfected cells reveals the sequence and transcriptional strength of cloned DNA fragments. Not all minimal promoter–enhancer combinations lead to transcriptional activation [97]. Therefore, the minimal promoter used determines the enhancers identified. STARR-seq has not yet been implemented in plants.

Bisulfite (BS)-seq

As mentioned before, low DNA methylation levels can indicate enhancers [39]. Genome-wide DNA methylation levels can be measured using BS conversion, which converts unmethylated cytosines to thymines, followed by sequencing (BS-seq) [98]. BS-seq was first implemented in *Arabidopsis*, and then used in several other plant species [99–101]. BS-seq offers single-base resolution of genome-wide DNA methylation profiles [102], allowing the precise delimitation of poorly methylated regions using computational tools [103]. Note that incomplete BS conversion and sequence polymorphisms can affect data interpretation.

Enhancer Trapping

With enhancer trapping, enhancers are detected by random genomic insertion of a reporter gene driven by a minimal promoter that is not sufficient to drive expression by itself [104,105] Expression of the reporter gene can be observed when inserted adjacent to an endogenous enhancer activating the gene [106]. For plants, typically a reporter gene (e.g., *GPF*) driven by the minimal CaMV-35S promoter is used. Numerous enhancer-trapping lines were isolated in *Arabidopsis* [107,108] and rice [109,110]. Most lines showed tissue-specific expression, hence this method allows the identification of regulatory sequences mediating expression patterns of interest. In reality, however, only few enhancers have been identified using this method, for example the minor vein phloem-specific enhancer *MATURE MINOR VEIN ELEMENT1* (*MMVE1*) [20] and *Egg Apparatus-Specific Enhancer* (*EASE*) [21]. This suggests that it is difficult to identify the trapped enhancers; indeed, enhancers can be located distally to enhancer traps. When studying a large, complex genome, enhancer identification will be even more cumbersome.

Characterization of Enhancers and their Target Genes

3C-Based Techniques

Chromosome conformation capture (3C) and its derivatives (e.g., 4C, 5C and Hi-C) measure relative interaction frequencies between different genomic regions [71,111]. In short, interacting chromosomal regions are crosslinked, followed by restriction enzyme digestion and intramolecular ligation of the interacting fragments. Finally, interaction frequencies are quantified using qPCR (3C) or sequencing. The main strength of these methods is their ability to identify target genes of enhancers (and vice versa). 3C and 4C are the method of choice when focusing on specific enhancers or genes [112,113], while a Hi-C protocol with a resolution of 1 kb allows genome-wide studies of enhancer–promoter interactions [114]. To avoid loss of sequencing capacity through interactions other than between enhancers and promoters, methods such as Capture C and Hi-Cap have been developed [115,116]. In plants, 3C was first implemented at the *b1* locus in maize, identifying interactions between the hepta-repeat enhancer and the *b1* gene [32,117]. The outcome of 3C and 4C studies on enhancer–promoter or other functional interactions in *Arabidopsis* is more cumbersome [33,75,118]. This can be explained by the compact genome size of *Arabidopsis*, hampering both the identification of the exact interacting sequences and the detection of relevant interactions above the background level of random ligation events.

Reporter Assay

The gold standard for testing enhancer sequences is a reporter assay [11]. Usually, with this method, a candidate enhancer and a control fragment are cloned upstream of a minimal promoter driving a reporter gene, followed by introduction into the tissue or cell type of interest and measurement of reporter gene activity. In plants, one can either use transient reporter assays or generate stable transgenic lines [15,41,119–121]. For transient assays, often *Agrobacterium tumefaciens*-mediated transient assays (ATTA) are performed, typically in tobacco leaves [119]. However, reporter assays are preferably performed in the plant species and tissues the enhancer candidates were derived from, because TFs and TF binding motifs may not be conserved between species [121–123]. Methods such as particle bombardment and protoplast

transformation allow candidate enhancers to be tested in their own genetic background [124,125].

Concluding Remarks and Future Perspectives

Similar to the situation in other organisms, enhancers play a crucial role in gene regulation in plants. Unlike plant enhancers, animal enhancers are very well characterized for their general properties (see Outstanding Questions). Based on features of the few well-characterized enhancers in plants, it appears that plant and animal enhancers share several characteristics, such as high chromatin accessibility, enrichment in histone acetylation, and low DNA methylation levels. These shared features can be used for the discovery of new enhancers and subsequent in-depth characterization of the properties of plant enhancers, in both active and inactive states. It will be interesting to find out whether eRNAs also have a role in enhancer function in plants. We stress that more genome-wide characterization of enhancers will be necessary to determine which combination of marks is best to identify enhancers in different activity states and to discriminate enhancers from other *cis*-regulatory elements, such as silencers and **insulators**. For the time being, validation of candidate enhancer sequences is necessary.

In this review we have discussed the advantages and drawbacks of the different techniques used to identify and characterize enhancers. Importantly, a combination of techniques offers a higher accuracy for genome-wide enhancer detection than a single method. As an example of an innovative genome-wide technique, we highlight STARR-seq, which uses the power of high-throughput sequencing to quantitatively assess enhancer activity [96]. Other revolutionary techniques such as CRISPR/Cas9 [126] allow the functionality of putative enhancer sequences to be tested *in vivo*, circumventing issues associated with transgenic reporter assays (e.g., transgene silencing and **position effects** [127]). Application of these novel techniques can greatly contribute to the identification and characterization of plant enhancers. Meanwhile, differences between plant species and their genomes need to be considered, as well as how these may impact on the performance of a specific technique. In conclusion, we expect the knowledge concerning transcriptional regulation by enhancers, and thereby knowledge of their regulatory potential in plants, to increase significantly in the very near future, and believe that many plant scientists will strongly benefit of such insights.

Acknowledgments

We gratefully acknowledge the financial support of the European Commission 7th Framework-People-2012-ITN Project EpiTRAITS (Epigenetic Regulation of Economically Important Plant Traits; 316965) to B.W., J.Z., R.O., and M.S.

References

1. Waddington, C.H. (1940) The genetic control of wing development in *Drosophila*. *J. Genet.* 41, 75–113
2. Sakabe, N.J. *et al.* (2012) Transcriptional enhancers in development and disease. *Genome Biol.* 13, 238
3. Kolovos, P. *et al.* (2012) Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin* 5, 1
4. Lettice, L.A. *et al.* (2002) Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci.* 99, 7548–7553
5. Benko, S. *et al.* (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nat. Genet.* 41, 359–364
6. Ogbourne, S. and Antalis, T.M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* 331, 1–14
7. Simpson, J. *et al.* (1986) Light-inducible and tissue-specific pea lhcp gene expression involves an upstream element combining enhancer- and silencer-like properties. *Nature* 323, 551–554
8. Kuhlemeier, C. *et al.* (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
9. Fluhr, R. *et al.* (1986) Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts. *EMBO J.* 5, 2063–2071
10. Coruzzi, G. *et al.* (1984) Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. *EMBO J.* 3, 1671–1679
11. Shlyueva, D. *et al.* (2014) Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15, 272–286
12. Iwafuchi-Doi, M. and Zaret, K.S. (2014) Pioneer transcription factors in cell reprogramming. *Genes Dev.* 28, 2679–2692
13. Jonkers, I. and Lis, J.T. (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 16, 167–177
14. Stam, M. (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev.* 16, 1906–1918
15. Timko, M.P. *et al.* (1985) Light regulation of plant gene expression by an upstream enhancer-like element. *Nature* 318, 579–582
16. Salvi, S. *et al.* (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci.* 104, 11376–11381

Outstanding Questions

It is important to establish which combination of characteristics is best to identify plant enhancers. Do they have plant-specific chromatin features or do they show the same features as in other species, and is the enhancer profile conserved across the entire plant kingdom?

One of the characteristics of active enhancers in animals is their transcriptional activity. It still needs to be investigated if plant enhancers are also transcribed upon activation and, if so, if eRNAs in plants are involved in similar functions as in animals.

An increase in genome size could provide more room for additional *cis*-regulatory elements and a larger average distance between these elements and their target genes. Is there a correlation between the genome size and (i) the number of enhancers and/or (ii) the distance between enhancers and their target genes?

Enhancers can be located upstream, downstream, or within the coding region of a gene. In addition, it has been shown that animal enhancers can skip adjacent genes and control more distantly located ones. Do plant enhancers only regulate immediately adjacent genes or can they also regulate more-distant genes?

Enhancers are shown to act on genes from a distance. Does the distance between enhancers and their target genes determine the effect of an enhancer on the transcription level of the target genes?

In animals, enhancers have been reported to regulate more than one gene, and genes are on average regulated by more than one enhancer. In plants, are there enhancers that activate more than one gene, and how many enhancers, on average, regulate a single gene?

17. Adrian, J. *et al.* (2010) Cis-regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in *Arabidopsis*. *Plant Cell Online* 22, 1425–1440
18. Hamamoto, H. *et al.* (2000) Tandem repeat of a transcriptional enhancer upstream of the sterol 14α-demethylase gene (CYP51) in *Penicillium digitatum*. *Appl. Environ. Microbiol.* 66, 3421–3426
19. Petrascheck, M. *et al.* (2005) DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res.* 33, 3743–3750
20. 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
21. Yang, W. *et al.* (2005) An egg apparatus-specific enhancer of *Arabidopsis*, identified by enhancer detection. *Plant Physiol.* 139, 1421–1432
22. Andersson, R. *et al.* (2014) An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–461
23. Kvon, E.Z. *et al.* (2014) Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* 512, 91–95
24. Sanyal, A. *et al.* (2012) The long-range interaction landscape of gene promoters. *Nature* 489, 109–113
25. Thurman, R.E. *et al.* (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75–82
26. Melton, C. *et al.* (2015) Recurrent somatic mutations in regulatory regions of human cancer genomes. *Nat. Genet.* 47, 710–716
27. Aran, D. *et al.* (2013) DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol.* 14, R21
28. Sagai, T. *et al.* (2005) Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132
29. Maurano, M.T. *et al.* (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190–1195
30. Chua, Y.L. *et al.* (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
31. Haring, M. *et al.* (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation: paramutation and chromatin structure. *Plant J.* 63, 366–378
32. Louwers, M. *et al.* (2009) Tissue- and expression level-specific chromatin looping at maize b1 epialleles. *Plant Cell* 21, 832–842
33. Liu, L. *et al.* (2014) Induced and natural variation of promoter length modulates the photoperiodic response of FLOWERING LOCUS T. *Nat. Commun.* 5, 4558
34. Zhu, B. *et al.* (2015) Genome-wide prediction and validation of intergenic enhancers in *Arabidopsis* using open chromatin signatures. *Plant Cell* 27, 2415–2426
35. Rada-Iglesias, A. *et al.* (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283
36. Zentner, G.E. *et al.* (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* 21, 1273–1283
37. Karmodiya, K. *et al.* (2012) H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13, 424
38. Creyghton, M.P. *et al.* (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* 107, 21931–21936
39. Stadler, M.B. *et al.* (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480, 490–495
40. Kim, T.-K. *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187
41. Simpson, J. *et al.* (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
42. Green, P.J. *et al.* (1987) Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the rbcS-3A gene. *EMBO J.* 6, 2543–2549
43. Nagano, Y. *et al.* (2001) A novel class of plant-specific zinc-dependent DNA-binding protein that binds to A/T-rich DNA sequences. *Nucleic Acids Res.* 29, 4097–4105
44. Valles, M.P. *et al.* (1991) Nuclease sensitivity of a maize HRGP gene in chromatin and in naked DNA. *Plant Sci.* 78, 225–230
45. Zhang, W. *et al.* (2012) High-resolution mapping of open chromatin in the rice genome. *Genome Res.* 22, 151–162
46. Kumar, S. and Bucher, P. (2016) Predicting transcription factor site occupancy using DNA sequence intrinsic and cell-type specific chromatin features. *BMC Bioinformatics* 17, 41–50
47. Matys, V. *et al.* (2003) TRANSFAC®: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31, 374–378
48. Mathelier, A. *et al.* (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 42, D142–D147
49. O'Malley, R.C. *et al.* (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292
50. Yu, C.-P. *et al.* (2015) Transcriptome dynamics of developing maize leaves and genomewide prediction of cis elements and their cognate transcription factors. *Proc. Natl. Acad. Sci.* 112, E2477–E2486
51. Bolduc, N. *et al.* (2012) Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev.* 26, 1685–1690
52. Zaret, K.S. and Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241
53. Zhang, W. *et al.* (2012) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* 24, 2719–2731
54. Vera, D.L. *et al.* (2014) Differential nuclease sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional DNA elements in maize. *Plant Cell* 26, 3883–3893
55. Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395
56. Nagarajan, S. *et al.* (2015) H4K12ac is regulated by estrogen receptor-alpha and is associated with BRD4 function and inducible transcription. *Oncotarget* 6, 7305–7317
57. Heintzman, N.D. *et al.* (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39, 311–318
58. Hoffman, B.G. *et al.* (2010) Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. *Genome Res.* 20, 1037–1051
59. Bonn, S. *et al.* (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* 44, 148–156
60. Sigova, A.A. *et al.* (2015) Transcription factor trapping by RNA in gene regulatory elements. *Science* 350, 978–981
61. Li, W. *et al.* (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520
62. Li, G. *et al.* (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98
63. De Santa, F. *et al.* (2010) A large fraction of extragenic RNA Pol II transcription sites overlap enhancers. *PLoS Biol* 8, e1000384
64. Li, W. *et al.* (2016) Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* 17, 207–223
65. 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
66. Zhong, S. *et al.* (2013) Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* 31, 154–159

67. Kinoshita, Y. *et al.* (2006) Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats: FWA gene silencing in *A. thaliana*. *Plant J.* 49, 38–45
68. Deng, S. *et al.* (2014) Transcriptional silencing of *Arabidopsis* endogenes by single-stranded RNAs targeting the promoter region. *Plant Cell Physiol.* 55, 823–833
69. 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
70. 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
71. de Wit, E. and de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* 26, 11–24
72. Splinter, E. *et al.* (2006) CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes Dev.* 20, 2349–2354
73. Kagey, M.H. *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435
74. Drissen, R. *et al.* (2004) The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev.* 18, 2485–2490
75. Cao, S. *et al.* (2014) A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in *Arabidopsis*. *Plant Cell* 26, 1009–1017
76. Spitz, F. and Furlong, E.E.M. (2012) Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613–626
77. Stewart, A.J. *et al.* (2012) Why transcription factor binding sites are ten nucleotides long. *Genetics* 192, 973–985
78. Wunderlich, Z. and Mirny, L.A. (2009) Different gene regulation strategies revealed by analysis of binding motifs. *Trends Genet.* 25, 434–440
79. Gonzalez, S. *et al.* (2012) ReLA, a local alignment search tool for the identification of distal and proximal gene regulatory regions and their conserved transcription factor binding sites. *Bioinformatics* 28, 763–770
80. Hesselberth, J.R. *et al.* (2009) Global mapping of protein–DNA interactions in vivo by digital genomic footprinting. *Nat. Methods* 6, 283–289
81. Reiss, D.J. *et al.* (2015) cMonkey 2: automated, systematic, integrated detection of co-regulated gene modules for any organism. *Nucleic Acids Res.* 43, e87
82. Gross, D.S. and Garrard, W.T. (1988) Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57, 159–197
83. Boyle, A.P. *et al.* (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 311–322
84. He, H.H. *et al.* (2014) Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. *Nat. Methods* 11, 73–78
85. Buenrostro, J.D. *et al.* (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218
86. Giresi, P.G. *et al.* (2007) FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Res.* 17, 877–885
87. Song, L. *et al.* (2011) Open chromatin defined by DNase I and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* 21, 1757–1767
88. Omidbakhshfard, M.A. *et al.* (2014) A step-by-step protocol for formaldehyde-assisted isolation of regulatory elements from *Arabidopsis thaliana*: FAIRE protocol for plants. *J. Integr. Plant Biol.* 56, 527–538
89. Haring, M. *et al.* (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods* 3, 11
90. Kaufmann, K. *et al.* (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
91. Visel, A. *et al.* (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858
92. Rhee, H.S. and Pugh, B.F. (2011) Comprehensive genome-wide protein–DNA interactions detected at single-nucleotide resolution. *Cell* 147, 1408–1419
93. Core, L.J. *et al.* (2008) Nascent RNA sequencing Reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848
94. Melgar, M.F. *et al.* (2011) Discovery of active enhancers through bidirectional expression of short transcripts. *Genome Biol.* 12, R113
95. Core, L.J. *et al.* (2014) Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* 46, 1311–1320
96. Arnold, C.D. *et al.* (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077
97. Zabidi, M.A. *et al.* (2015) Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 518, 556–559
98. Krueger, F. *et al.* (2012) DNA methylome analysis using short bisulfite sequencing data. *Nat. Methods* 9, 145–151
99. Cokus, S.J. *et al.* (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219
100. Regulski, M. *et al.* (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res.* 23, 1651–1662
101. Takuno, S. and Gaut, B.S. (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl. Acad. Sci.* 110, 1797–1802
102. Stevens, M. *et al.* (2013) Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Res.* 23, 1541–1553
103. Burger, L. *et al.* (2013) Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Res.* 41, e155
104. Wilson, C. *et al.* (1989) P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes Dev.* 3, 1301–1313
105. Aichinger, C. *et al.* (2003) Identification of plant-regulated genes in *Ustilago maydis*. *Mol. Genet. Genomics* 270, 303–314
106. Chudalayandi, S. (2011) Enhancer trapping in plants. *Methods Mol. Biol. Clifton NJ* 701, 285–300
107. Sundaresan, V. *et al.* (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* 9, 1797–1810
108. Gardner, M.J. *et al.* (2009) GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. *J. Exp. Bot.* 60, 213–226
109. Wu, C. *et al.* (2003) Development of enhancer trap lines for functional analysis of the rice genome. *Plant J. Cell Mol. Biol.* 35, 418–427
110. Johnson, A.A.T. *et al.* (2005) Spatial control of transgene expression in rice (*Oryza sativa* L.) using the GAL4 enhancer trapping system. *Plant J.* 41, 779–789
111. Dekker, J. *et al.* (2002) Capturing chromosome conformation. *Science* 295, 1306–1311
112. van de Werken, H.J.G. *et al.* (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* 9, 969–972
113. Gheldof, N. *et al.* (2010) Cell-type-specific long-range looping interactions identify distant regulatory elements of the CFTR gene. *Nucleic Acids Res.* 38, 4325–4336
114. Rao, S.S.P. *et al.* (2014) A 3D map of the kuman genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680
115. Hughes, J.R. *et al.* (2014) Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* 46, 205–212
116. Sahlén, P. *et al.* (2015) Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol.* 16, 156

117. Louwers, M. *et al.* (2009) Studying physical chromatin interactions in plants using chromosome conformation capture (3C). *Nat. Protoc.* 4, 1216–1229

118. Grob, S. *et al.* (2013) Characterization of chromosomal architecture in *Arabidopsis* by chromosome conformation capture. *Genome Biol.* 14, R129

119. Sparkes, I.A. *et al.* (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1, 2019–2025

120. 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

121. Belele, C.L. *et al.* (2013) Specific tandem repeats are sufficient for paramutation-induced trans-generational silencing. *PLoS Genet.* 9, e1003773

122. Copley, R.R. *et al.* (2007) Functional conservation of Rel binding sites in drosophilid genomes. *Genome Res.* 17, 1327–1335

123. Lamppa, G. *et al.* (1985) Light-regulated and organ-specific expression of a wheat Cab gene in transgenic tobacco. *Nature* 316, 750–752

124. Sidorenko, L. *et al.* (1999) Characterization of the regulatory elements of the maize P-rr gene by transient expression assays. *Plant Mol. Biol.* 39, 11–19

125. Faraco, M. *et al.* (2011) *One protoplast is not the other! Plant Physiol.* 156, 474–478

126. Bortesi, L. and Fischer, R. (2015) The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* 33, 41–52

127. Matzke, A.J.M. and Matzke, M.A. (1998) Position effects and epigenetic silencing of plant transgenes. *Curr. Opin. Plant Biol.* 1, 142–148

128. Raatz, B. *et al.* (2011) Specific expression of LATERAL SUPPRESSOR is controlled by an evolutionarily conserved 3' enhancer. *Plant J.* 68, 400–412

129. Lund, G. *et al.* (1995) Tissue-specific DNase I-sensitive sites of the maize P gene and their changes upon epimutation. *Plant J.* 7, 797–807

130. Castelletti, S. *et al.* (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL Vgt1. *G3 GenesGenomesGenetics* 4, 805–812

131. Nagy, F. *et al.* (1987) The 5'-proximal region of the wheat Cab-1 gene contains a 268-bp enhancer-like sequence for phytochrome response. *EMBO J.* 6, 2537–2542

132. van der Meer, I.M. *et al.* (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

133. van der Meer, I.M. *et al.* (1992) The TACPyAT repeats in the chalcone synthase promoter of *Petunia hybrida* act as a dominant negative cis-acting module in the control of organ-specific expression. *Plant J.* 2, 525–535

134. Zheng, L. *et al.* (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

135. Clark, R.M. *et al.* (2006) A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* 38, 594–597

136. Studer, A. *et al.* (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat. Genet.* 43, 1160–1163