

Article

Genomic Characterization and Functional Validation of Six *cis*-Regulatory Sequences in Medicinal Plant *Andrographis paniculata*

Xingbin Lv ^{1,2,†} , Hua Yang ^{2,†} , Yufang Hu ² , Qi Liang ^{2,3} , Shuyun Tian ², Lang Yang ² , Mingkun Huang ² , Ling Zhang ^{2,*}  and Yanqin Xu ^{1,*}

¹ School of Pharmacy, Jiangxi University of Traditional Chinese Medicine, 1688 Meiling Avenue, Xinjian District, Nanchang 330004, China; lv.xingbin@outlook.com

² Jiangxi Provincial Key Laboratory of Ex Situ Plant Conservation and Utilization, Lushan Botanical Garden, Chinese Academy of Sciences, 9 Zhiqing Road, Jiujiang 332900, China; yangh@lsbg.cn (H.Y.); yufanghu163@163.com (Y.H.); liangqi0611@163.com (Q.L.); shuytian@163.com (S.T.); langyang1124@outlook.com (L.Y.); huangmk@lsbg.cn (M.H.)

³ School of Life Sciences, Nanchang University, 999 Xuefu Avenue, Honggutan New District, Nanchang 330047, China

* Correspondence: linzh00@126.com (L.Z.); xuyanqin927@163.com (Y.X.)

† These authors contributed equally to this work.

Abstract: Plant *cis*-regulatory sequences (CRSs) are essential for gene expression and transcriptional regulation in plants. With the rapid development of plant biotechnologies, such as transgenesis and gene editing tools, plant-derived CRSs have gradually replaced traditional CRSs, like the 35S promoter, to avoid some negative effects caused by exogenous DNA sequences. In this study, we provided a comprehensive analysis of the published genomic data of *Andrographis paniculata* and identified six candidate *ApACR*s. Based on the transient assays, two candidate CRSs (*ApACR1* and *ApACR2*) demonstrated high transcriptional activities comparable to those of the 35S promoter in several plant species, in contrast to others with lower activities, suggesting their further potential applications in plant biotechnology and genetic engineering.

Keywords: *Andrographis paniculata*; medicinal plant; *cis*-regulatory sequences; epigenetic modifications; transcriptional regulation



Academic Editor: Juan Capel

Received: 6 December 2024

Revised: 6 January 2025

Accepted: 7 January 2025

Published: 9 January 2025

Citation: Lv, X.; Yang, H.; Hu, Y.; Liang, Q.; Tian, S.; Yang, L.; Huang, M.; Zhang, L.; Xu, Y. Genomic Characterization and Functional Validation of Six *cis*-Regulatory Sequences in Medicinal Plant *Andrographis paniculata*. *Horticulturae* **2025**, *11*, 63. <https://doi.org/10.3390/horticulturae11010063>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Andrographis paniculata (L.) Heynh (*A. paniculata*), belonging to the *Acanthaceae* family [1], is a traditional Chinese medicinal herb and has various medicinal effects such as anti-inflammatory [2], immunomodulatory [3], antioxidant [4], anti-cancer [5], antipyretic [6], hepatoprotection [7], and anti-HIV [8] effects. Andrographolide, the primary active compound in *A. paniculata* [9], has many significant medicinal properties, particularly anti-inflammatory and antibacterial properties, making *A. paniculata* known as a natural antibiotic [10]. Andrographolide accumulates mainly in leaf tissues, but at relatively low levels [11]. In addition, over-harvesting and environmental changes have limited the availability of wild *A. paniculata* resources, resulting in a continuous decline in germplasm reserves, which fail to meet the requirements of clinical applications [12].

Nowadays, research efforts have increasingly focused on developing effective cultivation techniques and molecular breeding strategies to increase the yield of *A. paniculata* [13]. However, molecular breeding relies heavily on a deep understanding of complex regulatory networks to improve active ingredients. Research in this area has mainly focused on

elucidating transcriptional regulatory mechanisms. Specifically, the interactions between *cis*-regulatory sequences (CRSs) and transcription factors (TFs) are at the core of the transcriptional regulatory network. CRSs, such as enhancers, promoters, and silencers, are critical for gene regulation, controlling intensities, stages, and regions by binding transcription factors [14–16]. In plants, the Cauliflower mosaic virus (CaMV) 35S promoter is one of the most extensively studied and widely experimentally used CRSs [17], consisting of a TATA box, inverted repeats, a CAAT box, and a core sequence [18]. Although the 35S promoter has been widely used due to its highly constitutive expression in many plants [19], its activity might be silenced by some protective mechanisms in plants, since it was derived from a virus [20,21]. Compared with the viral CRSs, plant-derived CRSs might provide greater safety and stability for plant transgenic breeding [22]. Therefore, developing plant-derived CRSs with comparable transcriptional activity has potential applications in further plant biotechnology like molecular breeding and transgenic experiments.

In this study, we comprehensively analyzed the published data on *A. paniculata*, including the transcriptome, transposase-accessible chromatin using sequencing (ATAC-Seq) [23], and histone modification chromatin immunoprecipitation sequencing (ChIP-seq) [24] in *A. paniculata*. Six CRSs (*ApACR1–ApACR6*) were selected using multi-omics analysis for functional validation. The transient assays showed that these *ApACRs*, especially *ApACR1* and *ApACR2*, have comparable activity with the 35S promoter in many plant species, indicating that these two CRSs might be used as alternatives to the 35S promoter for further transgenic experiments in *A. paniculata*.

2. Materials and Methods

2.1. Bioinformatic Analysis of *A. paniculata* Genomic Data

Transcriptome data of *A. paniculata* were obtained from NCBI under accession numbers SRX12305817 (<https://www.ncbi.nlm.nih.gov/sra/SRX12305817>, accessed on 28 September 2023), SRX12305811 (<https://www.ncbi.nlm.nih.gov/sra/SRX12305811>, accessed on 28 September 2023), and SRX12305810 (<https://www.ncbi.nlm.nih.gov/sra/SRX12305810>, accessed on 28 September 2023) [25]. Histone ChIP-seq and ATAC-seq data were downloaded from the National Genomics Data Center under project number PRJCA026190 (<https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA026190>, accessed on 20 December 2024) [26]. RNA-seq data were processed using TrimGalore (<https://github.com/FelixKrueger/TrimGalore>, version 0.6.10, accessed on 20 December 2024) to remove adapter sequences and low-quality reads. Filtered reads were then aligned to the *A. paniculata* reference genome (figshare, <https://doi.org/10.6084/m9.figshare.24986769>, accessed on 12 January 2024) using HISAT2 (version 2.2.1) [27]. Gene expression levels were calculated using Cuffnorm (version 2.2.1) (<http://cole-trapnell-lab.github.io/cufflinks/cuffnorm>, accessed on 20 December 2024) and represented by the fragments per kilobase of transcript per million mapped reads (FPKM) value.

ATAC-seq data were pre-processed using TrimGalore. The processed reads were then aligned to the *A. paniculata* reference genome using Bowtie2 (version 2.5.4) [28], with the parameter-X set to 1000. Reads with a mapping quality (MAPQ) score greater than 30 were retained. Peak calling was performed using MACS2 (version 2.2.9.1) [29] with the following parameters: --nomodel--shift 100--extsize 200-q 0.01. Tn5 transposase-tagged genomic DNA served as a control for peak analysis. Peaks identified in both biological replicates were defined as accessible chromatin regions (ACRs). The genomic annotation of the ACRs was performed using HOMER (version 4.11) [30]. The visualization of ATAC-seq signals was performed using Deeptools (version 3.5.5) [31] and pyGenomTrack (version 3.6) [32]. ACR abundance was assessed and ranked using FeatureCounts (version 2.0.6) [33].

Histone ChIP-seq data were pre-processed using TrimGalore. Filtered reads were aligned to the reference genome using Bowtie2 with default parameters. Aligned reads with a MAPQ score greater than 30 were used for peak calling with MUSIC software (version 1.0.0-2) (<https://github.com/gersteinlab/MUSIC>, accessed on 20 December 2024).

2.2. Co-Expression Analysis and Network Construction

We performed a Pearson analysis using *ApACR* sequences and transcription factors, defining the Pearson correlation coefficient (PCC) ≥ 0.65 and p -value ≤ 0.05 as co-expressed gene pairs. We then scanned for potential MOTIF binding sites on *ApACR* sequences using FIMO (version 5.5.7 in the MEME suite with a p -value $< 1 \times 10^{-4}$) [34] according to the position weight matrix (PWM) of TF motifs downloaded from PlantPAN3.0 [35]. The interaction network between TFs and *ApACRs* was constructed using cytoscape (version 3.10.3) (<https://github.com/cytoscape/cytoscape>, accessed on 20 December 2024).

2.3. Cloning and Vector Construction of *ApACRs*

Six candidate *ApACRs* were obtained from *A. paniculata* multi-omics data using the following criteria: high associated gene expression, high *ApACR* abundance, high H3K27ac and H3K4me3 enrichment, and low H3K27me3 enrichment. Specific primers with *KpnI* and *PstI* restriction sites (Table S1) were designed based on the selected *ApACR* sequences (Table S2). Target fragments were amplified from *A. paniculata* genomic DNA using KOD high-fidelity DNA polymerase (TOYOBO, Osaka, Japan). The PCR reaction mixture (50 μ L) contained 5 μ L 10 \times PCR Buffer, 5 μ L 2 mM dNTPs, 3 μ L 25 mM MgSO₄, 1.5 μ L forward primer, 1.5 μ L reverse primer, 1 μ L DNA template, and 33 μ L ddH₂O. The PCR cycling conditions were as follows: 94 °C for 2 min; 35 cycles of 98 °C for 10 s, 58 °C for 30 s, and 68 °C for 30 s; and a final extension at 68 °C for 10 min. PCR products were verified by gel electrophoresis and purified by the Agarose gel DNA Recovery Kit (Coolaber, Science & Technology Co., Ltd., Beijing, China). The purified fragments were then ligated into the *mini35S-LUC* vector, which had been linearized by double digestion with *KpnI* and *PstI* (NEB, MA, USA). The digestion reaction contained 1 μ g of vector, 5 μ L 10 \times rCutSmart Buffer, 1 μ L *KpnI*, 1 μ L *PstI*, and ddH₂O to a final volume of 50 μ L. This reaction was incubated at 37 °C for 30 min and then purified. The ligation reaction (10 μ L) contained 2 μ L linearized plasmid, 6.5 μ L purified PCR product, 1 μ L 10 \times T4 Ligase Buffer, and 0.5 μ L T4 DNA Ligase. The reaction was incubated at 25 °C for 2 h and then transformed into *Escherichia coli* DH5 α competent cells. Positive clones were selected by kanamycin and colony PCR verification and then sequenced by Sanger sequencing (Shanghai Biological Engineering Co., Ltd, Shanghai, China). Constructs with correct sequences were designated as *ApACRs-mini35S::LUC* vectors.

2.4. Luciferase Reporter Assay of *ApACRs* Transcriptional Activity in Tobacco Leaves

Following the methods described in a previously published article [36], the *ApACR-mini35S::LUC*, *35S-mini35S::LUC* (positive control), and *mini35S::LUC* vectors (Figure S1) (negative control, a 46 bp truncated 35S promoter without transcriptional activity) [37] were transformed into *Agrobacterium tumefaciens* GV3101 (*pSoup*) and infiltrated into 1-month-old, fully expanded tobacco leaves. After the infiltration, plants were kept in the dark for 16 h, then returned to normal light conditions (16 h light/8 h dark cycle) for two days. The leaves were then harvested and a chemiluminescence detector (Tano 5200, Tanon Science Technology Co., Ltd, Shanghai, China) was used to visualize the LUC (luciferase) activity. Relative LUC/REN activities were measured using the Dual Luciferase Reporter Assay System (Promega, Madison, WI, USA) according to the user's manual.

2.5. Luciferase Reporter Assay of ApACRs Transcriptional Activity in *Arabidopsis thaliana* and *Torenia fournieri* Protoplasts

Healthy and fully expanded leaves were harvested from 3 to 4-week-old *Arabidopsis* plants and digested in a 10 mL enzyme solution containing Cellulase (R10 or RS, Yakult, Tokyo, Japan) and Macerozyme (Yakult, Tokyo, Japan) for 30 min. The enzyme/protoplast solution was then washed with an equal volume of W5 solution (2 mM MES pH 5.7 containing 154 mM NaCl, 125 mM CaCl₂, and 5 mM KCl). Protoplasts were pelleted by centrifugation at 100× *g* for 1–2 min and resuspended in W5 solution and incubated on ice for 30 min. After incubation, the W5 solution was removed, and the protoplasts were resuspended in MMG solution (4 mM MES pH 5.7 containing 0.4 M mannitol and 15 mM MgCl₂) at room temperature. For DNA transfection, 10 µL plasmid DNA was mixed with 100 µL protoplasts, followed by the addition of 110 µL PEG solution (20–40% (wt/vol) PEG4000 (Merck, Darmstadt, Germany) containing 0.2 M mannitol and 100 mM CaCl₂ in ddH₂O.) The mixture was incubated at room temperature for 10 min, and the reaction was stopped by adding W5 solution. These protoplasts were pelleted by centrifugation at 100× *g* for 1–2 min, resuspended in WI solution (4 mM MES pH 5.7 containing 0.5 M mannitol and 20 mM KCl), and transferred to culture plates for incubation at room temperature. Following incubation, protoplasts were pelleted again by centrifugation at 100× *g* for 1–2 min and stored at –80 °C. The protoplast isolation and transformation of *T. fournieri* followed the previous study [38]. The constructs (obtained in Section 2.3) were transformed into the protoplasts of *A. thaliana* and *T. fournieri* and luciferase activities were measured as described in Section 2.4.

3. Results

3.1. Identification of Candidate Regulatory Sequences by Multi-Omics Analysis

It is known that CRSs are usually associated with gene expression and epigenetic modifications [39–41]. In this study, published data on the genome, RNA-seq, ChIP-seq, and ATAC-seq from *A. paniculata* were utilized to analyze gene expression levels (fragments per kilobase of exon model per million mapped fragments, FPKM), the location and abundance of ACRs (accessible chromatin regions), and histone modification features, aiming to identify candidate CRSs in *A. paniculata*.

The gene expression levels were positively correlated with ACR, H3K4me₃, and H3K27ac, while they were negatively correlated with H3K27me₃ (Figure 1A). The ACR, H3K4me₃, and H3K27ac peaks of high-expression genes were observed to be significantly high near the transcription start site (TSS), suggesting that highly expressed genes might be accompanied by H3K4me₃ and H3K27ac. In contrast, genes with low and no expression levels exhibited weaker signals in the TSS (Figure 1A). These findings are consistent with those of previous reports [26,42]. Figure 1B shows a positive correlation between gene expression levels and the enrichments of H3K27ac and H3K4me₃, while H3K27me₃ enrichment is associated with lower gene expression. These results demonstrate that highly expressed genes were usually marked by H3K27ac and H3K4me₃, particularly in promoter regions near the TSS. Therefore, ideal CRSs might be associated with high-expression genes and a high abundance of ACR, H3K4me₃, and H3K27ac signals, rather than H3K27me₃. Subsequently, six candidate CRSs (*ApACR1*–*ApACR6*) were selected from *A. paniculata* based on ACR abundance and the corresponding FPKM values for associated genes, H3K27ac, and H3K4me₃, while also considering H3K27me₃ enrichment during the selection process. (Figure 1C). Their characteristics are summarized in Table 1. The length of *ApACRs* ranged from 235 to 1058 bp, and they were located in the promoter region. The associated genes exhibited high expression levels, reaching up to 5973.6 FPKM, and were modified by active histone marks (H3K4me₃ and H3K27ac). For example, *ApACR1* exhibited high

accessibility (FPKM: 581) and was associated with the high expression of a nearby gene (FPKM:984) and strong enrichment of H3K4me3 (FPKM: 2368) and H3K27ac (FPKM: 2266). As mentioned above, the genomic features of *ApACRs* showed strong potential to promote its downstream gene expression, indicating that *ApACRs* might be able to recruit TFs for gene expression regulation.

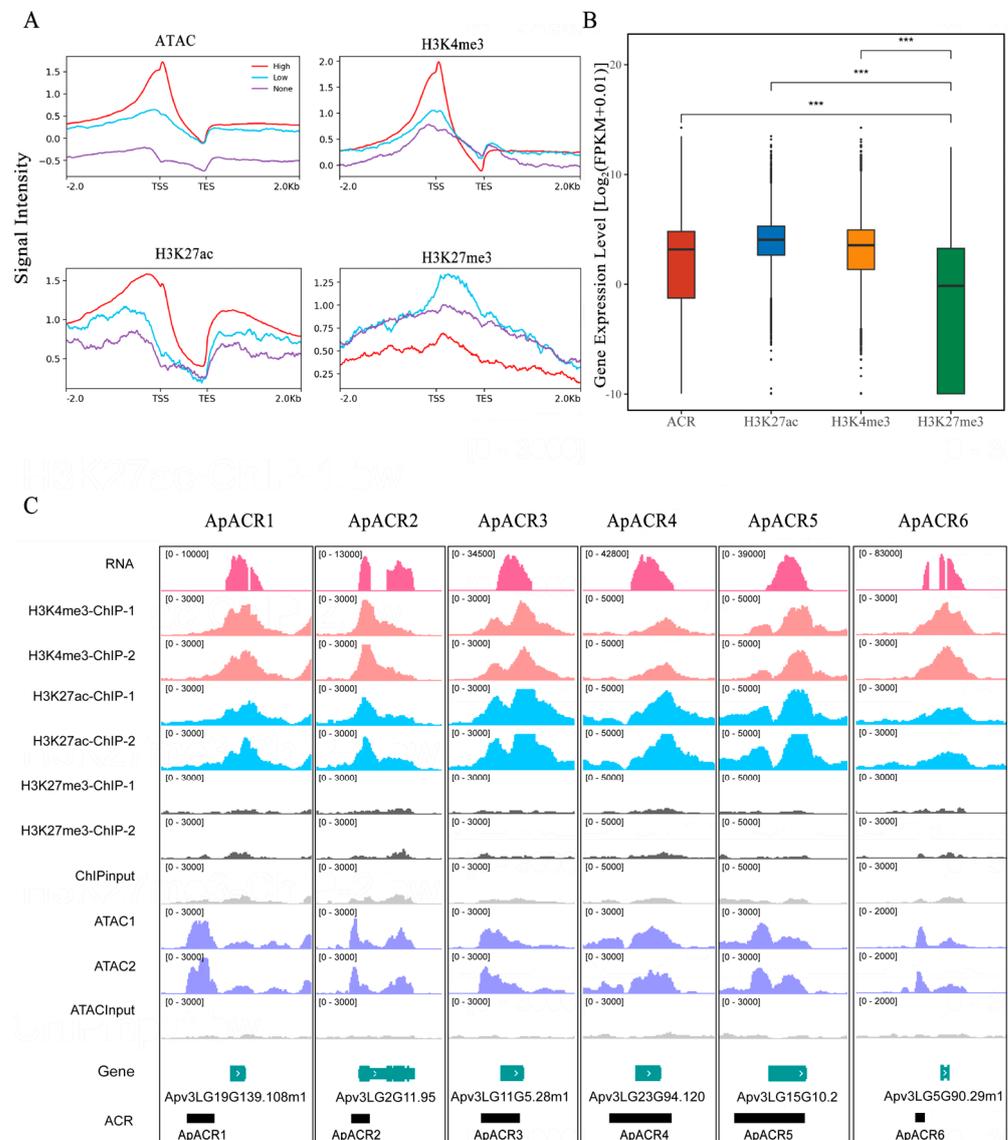


Figure 1. The genomic features of six selected *ApACRs*. **(A)** Correlation between gene expression level and four genomic features (ACR, H3K4me3, H3K27ac, and H3K27me3). Genes with high, low, and no expression were represented by $\text{FPKM} \geq 1$, $0 < \text{FPKM} < 1$, and $\text{FPKM} = 0$, respectively. FPKM: fragments per kilobase of transcript per million mapped reads; TSS: transcription start site; TES: transcription termination site. **(B)** Comparison of expression levels of genes modified by ACRs, H3K27ac, H3K4me3, and H3K27me3, respectively. Significant differences in gene expression levels were observed in genes associated with three active epigenetic signals (ACR, H3K27ac, and H3K4me3), compared with the repressive mark H3K27me3. **(C)** IGV screenshot showing genomic features of candidate *ApACRs*, which were associated with relatively high levels of gene expression and active histone modifications signals (H3K27ac and H3K4me3), except for H3K27me3. (***) indicated significant difference by Student's *t*-test ($p < 0.01$).

Table 1. Genomic features of six candidate *ApACRs*.

	<i>ApACR1</i>	<i>ApACR2</i>	<i>ApACR3</i>	<i>ApACR4</i>	<i>ApACR5</i>	<i>ApACR6</i>
Len ¹	798	498	628	950	1058	235
Chr ²	LG19	LG2	LG11	LG23	LG15	LG5
GC% ³	42.04%	41.56%	57.07%	43.43%	50.71%	51.69%
ACR FPKM ⁴	581	733	483	670	775	154
AssocGene ⁵	19G139.108	2G11.95	11G5.28	23G94.120	15G10.2	5G90.29
Gene Expr ⁶	984.8	880.6	2839.9	2910.8	3460.0	5973.6
Orientation ⁷	+	+	+	+	+	+
H3K27me3 Stat ⁸	None	None	None	None	None	None
H3K27me3 Loc ⁹	/	/	/	/	/	/
H3K27me3 Abund ¹⁰	/	/	/	/	/	/
H3K4me3 Stat ¹¹	Yes	Yes	Yes	Yes	Yes	Yes
H3K4me3 Loc ¹²	LG19: 13,948,418–13,953,227	LG2: 1,117,947–1,119,893	LG11: 553,591–554,806	LG23: 9,402,964–9,405,183	LG15: 986,833–990,657	LG5: 9,046,331–9,047,561
H3K4me3 Abund ¹³	2368	8896	4177	4070	9607	4931
H3K27ac Stat ¹⁴	Yes	Yes	Yes	Yes	Yes	Yes
H3K27ac Loc ¹⁵	LG19: 13,948,262–13,951,616	LG2: 1,118,020–1,118,963	LG11: 553,544–555,226	LG23: 9,404,241–9,407,453	LG15: 988,185–989,968	LG5: 9,046,317–9,047,562
H3K27ac Abund ¹⁶	2266	6003	9962	13,725	14,178	4315

Note: ¹ Length (bp); ² chromosome location; ³ GC content; ⁴ ACR abundance (FPKM); ⁵ associated gene; ⁶ associated gene expression level (FPKM); ⁷ associated gene orientation; ⁸ H3K27me3 modification status; ⁹ H3K27me3 modification location; ¹⁰ H3K27me3 abundance (FPKM); ¹¹ H3K4me3 modification status; ¹² H3K4me3 modification location; ¹³ H3K4me3 abundance (FPKM); ¹⁴ H3K27ac modification status; ¹⁵ H3K27ac modification location; ¹⁶ H3K27ac abundance (FPKM). FPKM: fragments per kilobase of transcript per million mapped reads.

3.2. *ApACRs* Were Potentially Targeted by Multiple Transcription Factors

The published RNA-seq data of *A. paniculata* from NCBI were utilized to identify the TFs co-expressed with *ApACRs*, which were associated with genes. In this study, co-expressed TFs, with a Pearson correlation coefficient (PCC) value over 0.65 and a *p*-value less than 0.05, were considered to be associated with the *ApACRs*. A total of 965 TFs, co-expressed with *ApACRs*, were identified as potentially involved in their regulation (Table S3). Moreover, *ApACR* sequences were analyzed by FIMO using the public TF motifs data to identify the potential TF binding sites in these *ApACRs*. Based on the analysis of co-expression and motif scanning (Table S4), the binding positions of different types of TFs in each *ApACR*, illustrated in Figure 2A,B show the number of binding sites for primary TF types. The interaction network between TFs and *ApACRs* was constructed using cytoscape (Figure 2C). *ApACR1* harbored the highest number of predicted TF binding sites (up to 32), including WRKY, bZIP, and MYB families. In contrast, *ApACR2*, *ApACR3*, *ApACR4*, *ApACR5*, and *ApACR6* had a lower number of predicted binding sites, with 5, 8, 12, 12, and 1 binding sites, respectively, indicating their binding TFs were distinct from each other. *ApACR2* might interact with MYB TFs; *ApACR3* showed interaction with AP2/ERF and MYB families; *ApACR4* and *ApACR5* were bound both with MYB and AP2/ERF families; a single bHLH TF motif was only found in *ApACR6* (Figure 2B). In addition, some TF binding sites shared by *ApACRs* were also observed (Figure 2C). For instance, two MYB-related TFs (Apv3LG1G103.1 and Apv3LG10G137.112) were predicted to bind to both *ApACR1* and *ApACR2*, and Apv3LG23G90.151 (bHLH) was found in both *ApACR1* and *ApACR6*. As previously reported, these transcription factor families play critical roles in processes such as growth and development [43–45] and the biosynthesis of andrographolide [46]. The analysis of the TFs-*ApACRs* interaction network indicated that the mechanism of *ApACRs* with high activation capacity in *A. paniculata* is complicated and needs further investigation. Nevertheless, these observations support their candidate roles for the further development of plant-derived CRSs.

that these *ApACRs*, especially *ApACR1* and *ApACR2*, could serve as effective alternatives to the 35S promoter for gene expression studies in various plant species.

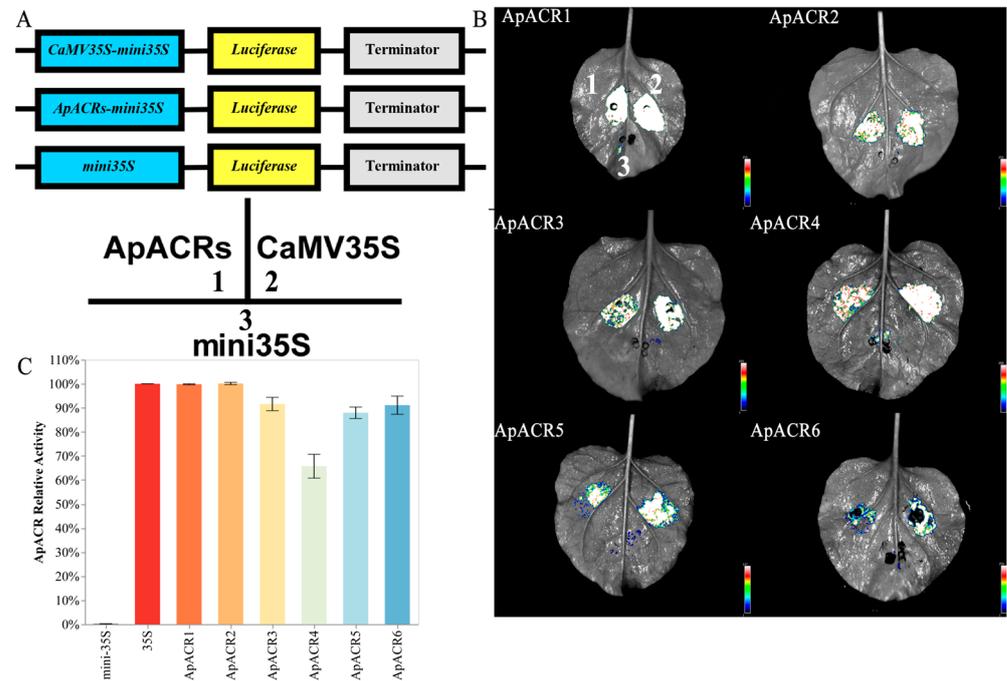


Figure 3. The transient expression assay of the activity of six *ApACRs* in tobacco leaves. (A) The structures of *ApACRs-mini35S* (marked in 1), *35S-mini35S* (marked in 2), and *mini35S* (marked in 3) vectors. (B) The transcription activities of *ApACRs* assessed in tobacco leaves by the transient expression assay. Three constructs were used: *ApACRs-mini35S* (experimental group); *35S-mini35S* as a positive control; and *mini35S*, a 46 bp truncated 35S promoter (negative control). The numbers 1,2,3 are shown in the lower panel of (A). (C) The relative transcriptional activities (LUC/REN) of the *ApACRs*. LUC represents firefly luciferase activity; REN represents renilla luciferase activity. Error bars represent the standard error of five biological replicates.

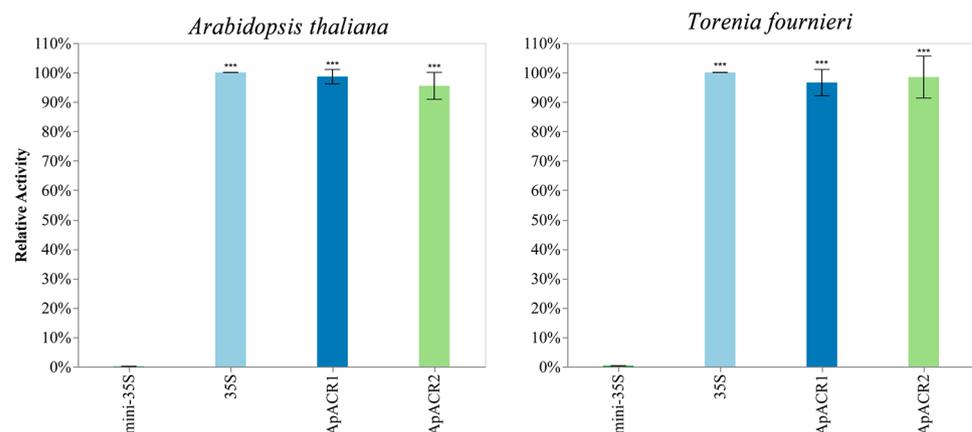


Figure 4. Relative transcriptional activities (LUC/REN) of *ApACR1* and *ApACR2* in *A. thaliana* and *T. fournieri* protoplasts, measured as relative luciferase activity. *Mini35S* served as negative control, and 35S promoter as a positive control. Error bars represent standard error of five biological replicates. LUC represents firefly luciferase activity; REN represents renilla luciferase activity. Asterisks (***) indicated significant difference compared to entry control (*mini-35S*) by Student's *t*-test ($p < 0.01$). Error bars indicate standard deviation of five biological replicates.

4. Discussion

CRSs are essential for driving the expression of associated genes in plants. Due to some negative effects caused by the virus-derived 35S promoter (e.g., gene silencing) in plants [30], plant-derived CRSs are now being considered as an alternative to the 35S promoter [47]. According to previous studies, plant-derived CRSs might be more effective than the traditional 35S promoter in regulating the expression of plant secondary metabolism-related genes [48]. For example, optimizing CRSs that controlled the expression levels of genes involved in betaine biosynthesis could maximize betaine production in tobacco [49]. To develop potential CRSs in a medicinal plant *A. paniculata*, we combined multi-omics analysis and transient expression assay to identify two *ApACRs* (*ApACR1* and *ApACR2*) with comparable transcriptional activity to the 35S promoter in several plant species, indicating a potential application role in further plant biotechnology. A promising application of the newly discovered *ApACR1* and *ApACR2* is enhancing the biosynthesis of valuable plant secondary metabolites, including pharmacologically active compounds such as andrographolide from *A. paniculata*. Increasing the expression of key genes (e.g., *ApCPS2*) involved in andrographolide biosynthesis may be an effective strategy to ultimately increase andrographolide production [25]. *ApACR1* and *ApACR2* will also expand the available toolkit for molecular breeding and genetic engineering in other plants. For example, the rapid development of gene editing technologies, such as CRISPR-Cas9, provide exciting possibilities for utilizing different types of CRSs in economically important crops or medicinal plants to improve editing efficiency *in vivo* [50].

Although we observed high transcriptional activities in these *ApACRs* in *A. paniculata*, the regulatory mechanism is unclear. As shown in Figure 2, several TF-binding motifs were predicted in these *ApACRs*, suggesting that these motifs play a key role in maintaining the high activity of these *ApACRs*. The interactions between TFs and *ApACRs* might be confirmed by other methods, such as high throughput sequencing methods, transient ChIP-seq [51], or *in vitro* DAP-seq [52], etc. Since the design of artificial CRSs by machine learning using big genomic data has been reported [53,54], our observation of these motif sequences might serve as an important genomic resource for further designing ideal CRSs and applying them to transgenic experiments in medicinal plants.

5. Conclusions

Taken together, our study could broaden the knowledge of some valuable CRSs, and the high transcriptional activities of *ApACR1* and *ApACR2* might serve as important resources for further plant biotechnological approaches, especially in *A. paniculata*.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/horticulturae11010063/s1>: Table S1: Primer Sequence Information for *ApACRs*; Table S2: Sequence Information for *ApACRs*; Table S3: Co-expression of *ApACRs* with Transcription Factors; Table S4: Motif Information for *ApACRs* Binding; Figure S1: The structure and the sequence annotation of *mini35S-LUC* reporter.

Author Contributions: Writing—original draft preparation, X.L. and H.Y.; methodology, M.H. and Y.X.; formal analysis, L.Y. and Q.L.; funding acquisition, L.Z. and M.H.; validation, Y.H. and S.T.; data curation and software, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (82260745) and Jiangxi Provincial Natural Science Foundation (20232BAB216120, China).

Data Availability Statement: Data is contained within the article.

Acknowledgments: We thank Shiqing Fan for the bioinformatic analysis of the genomic data of *A. paniculata*.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Okhuarobo, A.; Falodun, J.E.; Erharuyi, O.; Imieje, V.; Falodun, A.; Langer, P. Harnessing the Medicinal Properties of *Andrographis paniculata* for Diseases and beyond: A Review of Its Phytochemistry and Pharmacology. *Asian Pac. J. Trop. Dis.* **2014**, *4*, 213–222. [[CrossRef](#)]
2. Li, X.; Yuan, W.; Wu, J.; Zhen, J.; Sun, Q.; Yu, M. Andrographolide, a Natural Anti-Inflammatory Agent: An Update. *Front. Pharmacol.* **2022**, *13*, 920435. [[CrossRef](#)]
3. Chalichem, N.S.S.; Nabhan, P.; Bethapudi, B.; Agarwal, N.; Murugan, S.K.; Deepak, M. Assessment of Immunomodulatory Activity of AP-Bio[®] (KalmCold[®]), a Standardized Extract of *Andrographis paniculata* Using In Vivo and Ex Vivo Models. *Pharmacogn. Mag.* **2024**, *20*, 727–732. [[CrossRef](#)]
4. Sheeja, K.; Shihab, P.K.; Kuttan, G. Antioxidant and Anti-Inflammatory Activities of the Plant *Andrographis paniculata* Nees. *Immunopharmacol. Immunotoxicol.* **2006**, *28*, 129–140. [[CrossRef](#)]
5. Paul, S.; Roy, D.; Pati, S.; Sa, G. The Adroitness of Andrographolide as a Natural Weapon Against Colorectal Cancer. *Front. Pharmacol.* **2021**, *12*, 731492. [[CrossRef](#)]
6. Dai, Y.; Chen, S.-R.; Chai, L.; Zhao, J.; Wang, Y.; Wang, Y. Overview of Pharmacological Activities of *Andrographis paniculata* and Its Major Compound Andrographolide. *Crit. Rev. Food Sci. Nutr.* **2019**, *59* (Suppl. S1), S17–S29. [[CrossRef](#)] [[PubMed](#)]
7. Chao, W.-W.; Lin, B.-F. Hepatoprotective Diterpenoids Isolated from *Andrographis paniculata*. *Chin. Med.* **2012**, *3*, 136–143. [[CrossRef](#)]
8. Uttekar, M.M.; Das, T.; Pawar, R.S.; Bhandari, B.; Menon, V.; Nutan; Gupta, S.K.; Bhat, S.V. Anti-HIV Activity of Semisynthetic Derivatives of Andrographolide and Computational Study of HIV-1 Gp120 Protein Binding. *Eur. J. Med. Chem.* **2012**, *56*, 368–374. [[CrossRef](#)] [[PubMed](#)]
9. Kaewdech, A.; Nawalerspanya, S.; Assawasuwannakit, S.; Chamroonkul, N.; Jandee, S.; Sripongpun, P. The Use of *Andrographis paniculata* and Its Effects on Liver Biochemistry of Patients with Gastrointestinal Problems in Thailand during the COVID-19 Pandemic: A Cross Sectional Study. *Sci. Rep.* **2022**, *12*, 18213. [[CrossRef](#)] [[PubMed](#)]
10. Yan, J.; Wei, Y.; Hu, H.; Long, F.; Gou, S. Resource Investigation of Medicinal Species of *Andrographis* in Nationwid. *Lishizhen Med. Mater. Medica Res.* **2013**, *24*, 1997–1999.
11. Tundis, R.; Patra, J.K.; Bonesi, M.; Das, S.; Nath, R.; Das Talukdar, A.; Das, G.; Loizzo, M.R. Anti-Cancer Agent: The Labdane Diterpenoid-Andrographolide. *Plants* **2023**, *12*, 1969. [[CrossRef](#)]
12. Chen, X.; Ren, J.; Yang, J.; Zhu, Z.; Chen, R.; Zhang, L.; Chen, X.; Ren, J.; Yang, J.; Zhu, Z.; et al. A Critical Review of *Andrographis paniculata*. *Med. Plant Biol.* **2023**, *2*, 15. [[CrossRef](#)]
13. Hossain, S.; Urbi, Z.; Karuniawati, H.; Mohiuddin, R.B.; Moh Qrimida, A.; Allzrag, A.M.M.; Ming, L.C.; Pagano, E.; Capasso, R. *Andrographis paniculata* (Burm. f.) Wall. Ex Nees: An Updated Review of Phytochemistry, Antimicrobial Pharmacology, and Clinical Safety and Efficacy. *Life* **2021**, *11*, 348. [[CrossRef](#)] [[PubMed](#)]
14. Worsley-Hunt, R.; Bernard, V.; Wasserman, W.W. Identification of Cis-Regulatory Sequence Variations in Individual Genome Sequences. *Genome Med.* **2011**, *3*, 65. [[CrossRef](#)] [[PubMed](#)]
15. Lu, Z.; Ricci, W.A.; Schmitz, R.J.; Zhang, X. Identification of Cis-Regulatory Elements by Chromatin Structure. *Curr. Opin. Plant Biol.* **2018**, *42*, 90–94. [[CrossRef](#)] [[PubMed](#)]
16. Schmitz, R.J.; Grotewold, E.; Stam, M. Cis-Regulatory Sequences in Plants: Their Importance, Discovery, and Future Challenges. *Plant Cell* **2021**, *34*, 718–741. [[CrossRef](#)]
17. Schnurr, J.; Guerra, D. The CaMV-35S promoter is sensitive to shortened photoperiod in transgenic tobacco. *Plant Cell Rep.* **2020**, *19*, 279–282. [[CrossRef](#)] [[PubMed](#)]
18. Tang, T.; Xie, S.; Zhu, X.; Xu, J.; Tang, J.; Wang, X. CaMV35S Promoter and Its Application and Detection in Transgenic Crops. *Acta Agric. Zhejiangensis* **2019**, *31*, 161–170. [[CrossRef](#)]
19. Benfey, P.N.; Chua, N.-H. The Cauliflower Mosaic Virus 35 S Promoter: Combinatorial Regulation of Transcription in Plants. *Science* **1990**, *250*, 959–966. [[CrossRef](#)]
20. Okumura, A.; Shimada, A.; Yamasaki, S.; Horino, T.; Iwata, Y.; Koizumi, N.; Nishihara, M.; Mishiba, K. CaMV-35S Promoter Sequence-Specific DNA Methylation in Lettuce. *Plant Cell Rep.* **2016**, *35*, 43–51. [[CrossRef](#)]
21. Custers, J.B.M.; Snepvangers, S.C.H.J.; Jansen, H.J.; Zhang, L.; van Lookeren Campagne, M.M. The 35S-CaMV Promoter Is Silent during Early Embryogenesis but Activated during Nonembryogenic Sporophytic Development in Microspore Culture. *Protoplasma* **1999**, *208*, 257–264. [[CrossRef](#)]

22. Jiang, P.; Zhang, K.; Ding, Z.; He, Q.; Li, W.; Zhu, S.; Cheng, W.; Zhang, K.; Li, K. Characterization of a Strong and Constitutive Promoter from the *Arabidopsis* Serine Carboxypeptidase-like Gene AtSCPL30 as a Potential Tool for Crop Transgenic Breeding. *BMC Biotechnol.* **2018**, *18*, 59. [[CrossRef](#)] [[PubMed](#)]
23. Buenrostro, J.D.; Giresi, P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position. *Nat. Methods* **2013**, *10*, 1213–1218. [[CrossRef](#)] [[PubMed](#)]
24. Braunstein, M.; Rose, A.B.; Holmes, S.G.; Allis, C.D.; Broach, J.R. Transcriptional Silencing in Yeast Is Associated with Reduced Nucleosome Acetylation. *Genes Dev.* **1993**, *7*, 592–604. [[CrossRef](#)]
25. Sun, W.; Leng, L.; Yin, Q.; Xu, M.; Huang, M.; Xu, Z.; Zhang, Y.; Yao, H.; Wang, C.; Xiong, C.; et al. The Genome of the Medicinal Plant *Andrographis paniculata* Provides Insight into the Biosynthesis of the Bioactive Diterpenoid Neoandrographolide. *Plant J.* **2019**, *97*, 841–857. [[CrossRef](#)] [[PubMed](#)]
26. Huang, M.; Hu, Y.; Zhang, L.; Yang, H.; Feng, C.; Jiang, C.; Xie, N.; Liu, D.; Chen, S.; Wang, J.; et al. Decoding the Chromatin Accessibility in *Andrographis paniculata* Genome, a Case Study of Genome-Wide Investigation of the *Cis*-Regulatory Elements in Medicinal Plants. *Acta Pharm. Sin. B* **2024**, *14*, 4179–4182. [[CrossRef](#)]
27. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [[CrossRef](#)]
28. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
29. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoutte, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-Based Analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [[CrossRef](#)]
30. Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y.C.; Laslo, P.; Cheng, J.X.; Murre, C.; Singh, H.; Glass, C.K. Simple Combinations of Lineage-Determining Transcription Factors Prime *Cis*-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **2010**, *38*, 576–589. [[CrossRef](#)]
31. Ramírez, F.; Dünder, F.; Diehl, S.; Grüning, B.A.; Manke, T. deepTools: A Flexible Platform for Exploring Deep-Sequencing Data. *Nucleic Acids Res.* **2014**, *42*, W187–W191. [[CrossRef](#)] [[PubMed](#)]
32. Lopez-Delisle, L.; Rabbani, L.; Wolff, J.; Bhardwaj, V.; Backofen, R.; Grüning, B.; Ramírez, F.; Manke, T. pyGenomeTracks: Reproducible Plots for Multivariate Genomic Datasets. *Bioinformatics* **2021**, *37*, 422–423. [[CrossRef](#)]
33. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**, *30*, 923–930. [[CrossRef](#)]
34. Grant, C.E.; Bailey, T.L.; Noble, W.S. FIMO: Scanning for Occurrences of a given Motif. *Bioinformatics* **2011**, *27*, 1017–1018. [[CrossRef](#)] [[PubMed](#)]
35. Cn, C.; Ty, L.; Yc, H.; Gz, L.; Kc, T.; Yh, L.; Pl, K.; Hq, Z.; Wc, C. PlantPAN3.0: A New and Updated Resource for Reconstructing Transcriptional Regulatory Networks from ChIP-Seq Experiments in Plants. *Nucleic Acids Res.* **2019**, *47*, D1155–D1163. [[CrossRef](#)]
36. Huang, M.; Zhang, L.; Yung, W.-S.; Hu, Y.; Wang, Z.; Li, M.-W.; Lam, H.-M. Molecular Evidence for Enhancer–Promoter Interactions in Light Responses of Soybean Seedlings. *Plant Physiol.* **2023**, *193*, 2287–2291. [[CrossRef](#)]
37. Amack, S.C.; Antunes, M.S. CaMV35S Promoter—A Plant Biology and Biotechnology Workhorse in the Era of Synthetic Biology. *Curr. Plant Biol.* **2020**, *24*, 100179. [[CrossRef](#)]
38. Zhang, L.; Yung, W.-S.; Wang, Z.; Li, M.-W.; Huang, M. Optimization of an Efficient Protoplast Transformation System for Transient Expression Analysis Using Leaves of *Torenia Fournieri*. *Plants* **2022**, *11*, 2106. [[CrossRef](#)] [[PubMed](#)]
39. Beernink, B.M.; Vogel, J.P.; Lei, L. Enhancers in Plant Development, Adaptation and Evolution. *Plant Cell Physiol.* **2024**; *online ahead of print*, pcae121. [[CrossRef](#)]
40. Kellis, M.; Wold, B.; Snyder, M.P.; Bernstein, B.E.; Kundaje, A.; Marinov, G.K.; Ward, L.D.; Birney, E.; Crawford, G.E.; Dekker, J.; et al. Defining Functional DNA Elements in the Human Genome. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6131–6138. [[CrossRef](#)] [[PubMed](#)]
41. Preissl, S.; Gaulton, K.J.; Ren, B. Characterizing *Cis*-Regulatory Elements Using Single-Cell Epigenomics. *Nat. Rev. Genet.* **2023**, *24*, 21–43. [[CrossRef](#)] [[PubMed](#)]
42. Zhang, L.; Yung, W.-S.; Hu, Y.; Wang, L.; Sun, W.; Huang, M.; Zhang, L.; Yung, W.-S.; Hu, Y.; Wang, L.; et al. Establishment of a Convenient ChIP-Seq Protocol for Identification of the Histone Modification Regions in the Medicinal Plant *Andrographis paniculata*. *Med. Plant Biol.* **2023**, *2*, 6. [[CrossRef](#)]
43. Hao, Y.; Zong, X.; Ren, P.; Qian, Y.; Fu, A. Basic Helix-Loop-Helix (bHLH) Transcription Factors Regulate a Wide Range of Functions in *Arabidopsis*. *Int. J. Mol. Sci.* **2021**, *22*, 7152. [[CrossRef](#)] [[PubMed](#)]
44. Wu, Y.; Wen, J.; Xia, Y.; Zhang, L.; Du, H. Evolution and Functional Diversification of R2R3-MYB Transcription Factors in Plants. *Hortic. Res.* **2022**, *9*, uhac058. [[CrossRef](#)]
45. Wang, H.; Chen, W.; Xu, Z.; Chen, M.; Yu, D. Functions of WRKYs in Plant Growth and Development. *Trends Plant Sci.* **2023**, *28*, 630–645. [[CrossRef](#)] [[PubMed](#)]

46. Guan, R.; Xu, S.; Lu, Z.; Su, L.; Zhang, L.; Sun, W.; Zhang, Y.; Jiang, C.; Liu, Z.; Duan, L.; et al. Genomic Characterization of bZIP Transcription Factors Related to Andrographolide Biosynthesis in *Andrographis paniculata*. *Int. J. Biol. Macromol.* **2022**, *223*, 1619–1631. [[CrossRef](#)] [[PubMed](#)]
47. Kummari, D.; Palakolanu, S.R.; Kishor, P.B.K.; Bhatnagar-Mathur, P.; Singam, P.; Vadez, V.; Sharma, K.K. An Update and Perspectives on the Use of Promoters in Plant Genetic Engineering. *J. Biosci.* **2020**, *45*, 119. [[CrossRef](#)]
48. Ali, S.; Kim, W.-C. A Fruitful Decade Using Synthetic Promoters in the Improvement of Transgenic Plants. *Front. Plant Sci.* **2019**, *10*, 1433. [[CrossRef](#)]
49. Tian, C.; Zhang, Y.; Li, J.; Wang, Y. Benchmarking Intrinsic Promoters and Terminators for Plant Synthetic Biology Research. *BioDesign Res.* **2022**, *2022*, 9834989. [[CrossRef](#)]
50. Zhang, Y.; Liang, Z.; Zong, Y.; Wang, Y.; Liu, J.; Chen, K.; Qiu, J.-L.; Gao, C. Efficient and Transgene-Free Genome Editing in Wheat through Transient Expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* **2016**, *7*, 12617. [[CrossRef](#)]
51. Wang, G.; Li, X.; An, Y.; Zhang, J.; Li, H. Transient ChIP-Seq for Genome-Wide In Vivo DNA Binding Landscape. *Trends Plant Sci.* **2021**, *26*, 524–525. [[CrossRef](#)] [[PubMed](#)]
52. Bartlett, A.; O'Malley, R.C.; Huang, S.C.; Galli, M.; Nery, J.R.; Gallavotti, A.; Ecker, J.R. Mapping Genome-Wide Transcription-Factor Binding Sites Using DAP-Seq. *Nat. Protoc.* **2017**, *12*, 1659–1672. [[CrossRef](#)] [[PubMed](#)]
53. Li, T.; Xu, H.; Teng, S.; Suo, M.; Bahitwa, R.; Xu, M.; Qian, Y.; Ramstein, G.P.; Song, B.; Buckler, E.S.; et al. Modeling 0.6 Million Genes for the Rational Design of Functional *Cis*-Regulatory Variants and de Novo Design of *Cis*-Regulatory Sequences. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e2319811121. [[CrossRef](#)]
54. Mansoor, S.; Karunathilake, E.M.B.M.; Tuan, T.T.; Chung, Y.S. Genomics, Phenomics, and Machine Learning in Transforming Plant Research: Advancements and Challenges. *Hortic. Plant J.* **2024**; *in press*. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.