plants

MDPI

Article

Isolation and Comprehensive in Silico Characterisation of a New *3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase 4* (*HMGR4*) Gene Promoter from *Salvia miltiorrhiza*: Comparative Analyses of Plant *HMGR* Promoters

Małgorzata Majewska *, Łukasz Kuźma and Piotr Szymczyk

Department of Biology and Pharmaceutical Botany, Medical University of Lodz, Muszyńskiego 1, 90-151 Lodz, Poland; lukasz.kuzma@umed.lodz.pl (Ł.K.); piotr.szymczyk@umed.lodz.pl (P.S.)
* Correspondence: malgorzata.majewska1@stud.umed.lodz.pl

Abstract: *Salvia miltiorrhiza* synthesises tanshinones with multidirectional therapeutic effects. These compounds have a complex biosynthetic pathway, whose first rate limiting enzyme is 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR). In the present study, a new 1646 bp fragment of the *S. miltiorrhiza HMGR4* gene consisting of a promoter, 5′ untranslated region and part of a coding sequence was isolated and characterised in silico using bioinformatics tools. The results indicate the presence of a TATA box, tandem repeat and pyrimidine-rich sequence, and the absence of CpG islands. The sequence was rich in motifs recognised by specific transcription factors sensitive mainly to light, salicylic acid, bacterial infection and auxins; it also demonstrated many binding sites for microRNAs. Moreover, our results suggest that *HMGR4* expression is possibly regulated during flowering, embryogenesis, organogenesis and the circadian rhythm. The obtained data were verified by comparison with microarray co-expression results obtained for *Arabidopsis thaliana*. Alignment of the isolated *HMGR4* sequence with other plant *HMGRs* indicated the presence of many common binding sites for transcription factors, including conserved ones. Our findings provide valuable information for understanding the mechanisms that direct transcription of the *S. miltiorrhiza HMGR4* gene.

Keywords: *HMGR4*; microRNA; promoter; *Salvia miltiorrhiza*; transcription factor; transcription factor binding site

1. Introduction

Salvia miltiorrhiza Bunge, also known as Red sage, or Chinese sage, is an important species used in traditional Chinese medicine. The dried root is used alone or in combination with other herbs to treat various ailments including cardiovascular diseases, menstrual disorders and insomnia [1,2]. In addition, it has been found to have potential in treating cancer [3], Parkinson's [4] and Alzheimer's disease [5], as well as renal deficiency [6], hepatocirrhosis [7], acute lung injury [8], fibrosis [9], neuropathic pain [10], diabetes mellitus [11], or alcohol dependence [12]. The main bioactive compounds responsible for such medical properties are quinone diterpenoids (e.g., tanshinones) and phenolic acids. The tanshinones are biosynthesised from intermediates generated in mevalonate (MVA) and methylerythritol phosphate (MEP) pathways [13]. The key rate-limiting enzyme in the MVA pathway converting 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) to MVA is HMG-CoA reductase (HMGR) [14]. The pivotal role of *HMGR* in plant metabolism is emphasised by the precise regulation of its function at the level of transcription, post-transcription, translation and post-translation [15,16]. To date, five *S. miltiorrhiza HMGR* gene sequences (*HMGR—HMGR4*) have been identified and deposited in the GenBank database [17–19]. A combination of cDNA sequence similarity searches with exon/intron structure indicates

that *HMGR* (EU680958.1) and *HMGR4* (JN831103.1), and *HMGR2* (FJ747636.1) and *HMGR3* (JN831102.1) are probably two pairs of duplicated genes, respectively [19]. Although the coding sequences of the *S. miltiorrhiza HMGR* genes have been identified and described, some of their promoter sequences remain unknown. These include the *HMGR4* promoter.

Promoter sequence analysis provides much valuable information for understanding the regulation of gene expression. Motifs such as the TATA box, CpG islands, tandem repeats or transcription factor binding sites (TFBSs) deserve special attention. Genome-wide analyses indicate that most in vivo functional TFBSs are located in the proximal promoter region [20,21]. These sites form clusters, thus improving interactions of corresponding transcription factors (TFs) to ensure a better execution of their regulatory functions [22]. An essential functional linkage exists between TFs and RNA polymerase II, acting as a large, conformationally flexible multiprotein complex known as a Mediator [23]. This complex regulates polymerase activity by transmitting signals from TFs. Groups of TFs form complex networks of dependencies and act in a coordinated manner in response to intracellular and environmental signals, thus directing many biological processes [24]. Gene expression is also regulated by the activity of microRNAs (miRNAs). They are mainly known as post-transcriptional and translational inhibitors of gene expression. The miRNAs cut the mRNA strands, destabilise the mRNA by shortening its poly(A) tail, and reduce the efficiency of the translation process [25,26]. However, studies on human and *Arabidopsis thaliana* indicate that miRNAs can also regulate gene expression during transcription by binding to promoter sequences [27–30]. In conclusion, to understand the mechanisms driving gene expression, it is necessary to also understand the nature of the promoter regions. The first step to achieving this goal requires use of bioinformatics tools.

A. thaliana is the most widely-studied plant in modern biology. Its wide appeal for scientists results from its fast growth rate, easy maintenance and small space requirements [31]. Moreover, the plant indicates a good tolerance to homozygosity and self-fertility [32]. Genetic studies are attracted by the small size (132 Mbp) of its completely sequenced genome [32]. Due to its similarity to other plants, *A. thaliana* has become the starting point for studying numerous aspects of plant cell-, molecular- and system-biology [33]. The abundance of research conducted on *A. thaliana* has led to the creation of numerous clones, cloning vectors, mutant lines, seeds, databases and online tools containing genomic, epigenomic, transcriptomic and proteomic data [34].

This work describes the isolation of a new *S. miltiorrhiza HMGR4* promoter and 5′ untranslated region (5′UTR) sequences, and their in silico characterisation via specialised databases such as PlantPan 2.0, TSSP and miRBase. Furthermore, the comprehensive in silico analysis presented herein presents valuable new information about the regulatory functions of *HMGR* promoters. The data can be used to create modified or synthetic promoters which could be active under certain controlled conditions [35]. In this way, numerous medically-important metabolites may be obtained.

2. Results

2.1. In Silico Analysis of the S. miltiorrhiza HMGR4 Promoter

The 1646 bp sequence obtained by the DNA walking method was deposited in GenBank under accession number KT921337.1 (Figure 1). The sequence contained a 51 bp coding region which perfectly coincided with *HMGR4* gene sequence JN831103.1 identified by Ma et al. [19]. The TATA box was located at bases −28 to −33 from TSS. One tandem repeat (at −1296 to −1353) and no CpG islands were found. A pyrimidine-rich sequence (PRS) was recognised in the 5′UTR.

Figure 1. Isolated *S. miltiorrhiza HMGR4* promoter sequence (1499 bp), 5′ untranslated region (5′UTR) (96 bp) and coding sequence fragment (51 bp). The potential TATA box, transcription start site (TSS), pyrimidine-rich sequence (PRS), tandem repeat and consensus sequences for hormone-, pathogen-, wounding-, light-, and anaerobic-responsive elements are signed and marked in pink on the strands. ABRE, abscisic acid-responsive element; ARE, anaerobic-responsive element; AuxRE, auxin-responsive element; BRRE, brassinosteroid-responsive element; ERE, ethylene-responsive element; LRE, light-responsive element; MJRE, methyl jasmonate-responsive element.

Moreover, 5369 TFBSs and 365 potentially interacting TFs described previously in *A. thaliana* were revealed using the PlantPan 2.0 tool (File S1). Each of the obtained TFs could interact with a number of binding sites. The TFBSs were identified at both strands of the examined sequence. The similarity score between the binding sites found in the *S. miltiorrhiza HMGR4* promoter and those detected in *A. thaliana* ranged from 0.7 to 1.0.

Additional analysis of the entire *HMGR4* promoter sequence revealed the following commonly-known consensus sequences: two auxin-responsive elements (AuxREs); seven salicylic acid (SA)-responsive elements, including W box and TCA-elements; two brassinosteroid-responsive elements (BRREs); two ethylene-responsive elements (EREs); two abscisic acid-responsive elements (ABREs); four methyl jasmonate-responsive elements (MJREs); one pathogen-responsive element Eli box 3; three wounding- and pathogen-responsive elements WRE3; three anaerobic-responsive elements (AREs); and twenty-four

light-responsive elements (LREs), including ATCT sequences, Box 4, GT1 motif, TCT motifs, GA motifs, AE box, Box I (Figure 1). The vast majority of these results were in agreement with data provided by PlantPan 2.0, which demonstrated the presence of numerous binding sites for TFs responding to these types of stimulation (Table S1). Only anaerobic-sensitive TFs were not found. The most commonly-observed TFs were those representing the following families: GATA (light); MYB-related and Dof (auxin); WRKY (SA, wounding and pathogen); NAC; NAM (brassinosteroid and abscisic acid (ABA)); MYB-related (ethylene); and CAMTA (MeJa) (Table S1). The consensus sequences listed above were distributed along the entire *HMGR4* promoter, with some being located only a few nucleotides from each other; this allows more precise regulation of gene expression by dimerization of binding TFs. Such sequences included two SA-responsive TCA elements and two LREs (Table 1).

Table 1. Transcription factor binding sites (TFBSs) with the potential to form dimers detected in the *S. miltiorrhiza HMGR4* promoter sequence.

Transcription Factor (TF) Family Name	Fragment of Promoter Sequence with Underlined TFBSs [a] and Potentially Interacting TF Pairs
-	AGAAAAATGGAATAAGAAGA two salicylic acid (SA)-responsive TCA elements (AGAAAA and AGAAGA) spaced by eight nucleotides
-	ATCTCCAATCT two LREs (ATCT) spaced by three nucleotides
HD-ZIP	atTTAATgtaTTCATAAATata ATML1/HDG1 and PDF2/ATML1 spaced by four nucleotides
WRKY	AGTCATAACAATGTCAA WRKY2/WRKY14/WRKY45/WRKY57/WRKY69 and WRKY2/WRKY14/WRKY45/WRKY57/WRKY69 spaced by six nucleotides
Dof	AAAGAAAAAAGA DOF5.4 and DOF5.4 spaced by two nucleotides

[a] Most conserved positions within a matrix were written in capital letters.

The search of the PlantPan 2.0 database revealed about 234 TFs that could interact with 915 TFBSs in the *HMGR4* proximal promoter (File S2). To complement the findings described above, the binding sites located in the proximal promoter were analysed in terms of their response to external factors. Such data were available for 666 TFBSs (File S2). The results indicated that *HMGR4* transcription may be most dependent on light, SA, bacterial infection and auxins, and less dependent on ABA, gibberellin, chitin, cold or salt stress (Figure 2). The response to these agents was mainly associated with the following TF families: GATA (light); WRKY (SA, bacterial infection, chitin); bZIP (auxins); MYB-related, WRKY and C2H2 (ABA); MYB-related and MADS box (gibberellins); C2H2 and WRKY (cold); and WRKY and MYB-related (salt stress) (Table S2). To determine the stage of *S. miltiorrhiza* development at which the *HMGR4* gene expression regulation is likely to occur, 450 TFBSs and interacting TFs located in its proximal promoter were examined (File S2). The findings indicated flowering, embryogenesis, organogenesis (root, shoot, leaf and flower development) and circadian rhythm (Figure 3).
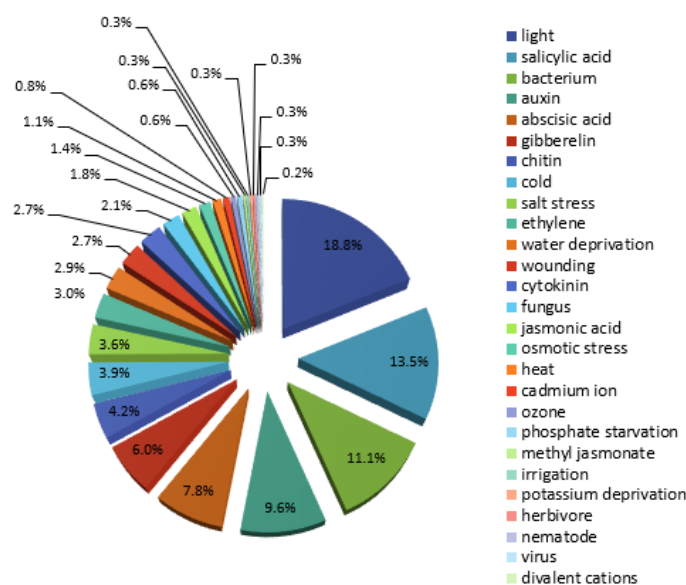
Figure 2. Classification of TFBSs found in the proximal *S. miltiorrhiza HMGR4* promoter with regard to their response to biotic and abiotic factors.
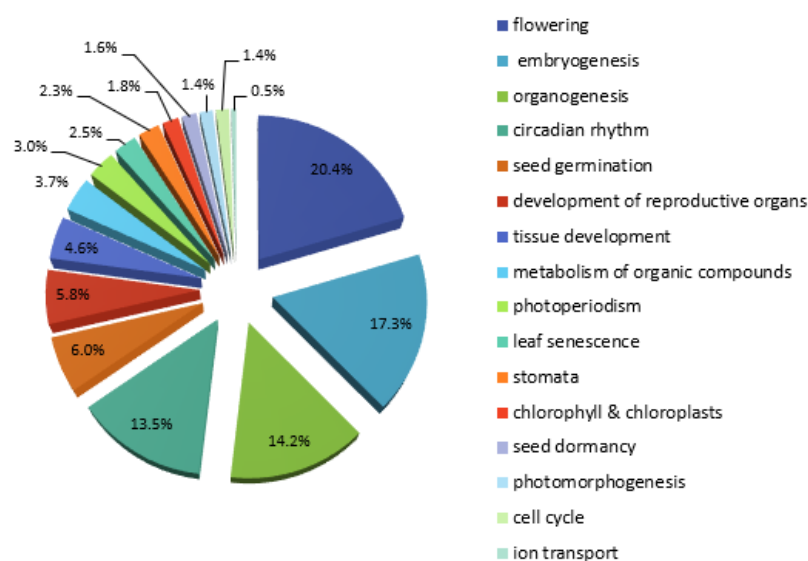

Figure 3. Classification of TFBSs found in the proximal *S. miltiorrhiza HMGR4* promoter with regard to their biological functions.

As *HMGR* genes are crucial for the production of intermediates in the biosynthesis pathway for tanshinones, a literature search was performed for TFs that positively regulate tanshinone production. Following this, based on the results obtained with the PlantPan 2.0 database, the *S. miltiorrhiza HMGR4* promoter sequence was searched for the presence of binding sites for the 20 identified TFs; of these, the results indicated the presence of the following five TFs: BHLH6 (MYC2) (14 binding sites), BHLH74 (2 sites), BZIP20 (32 sites), WRKY2 (8 sites) and WRKY61 (9 sites) (File S1) [36–40]. A number of TFBSs were also found to be located in the proximal promoter region. The obtained results suggest that *HMGR4* may play a role in the biosynthesis of tanshinones.

Furthermore, in silico analysis of the *HMGR4* promoter and 5′UTR revealed potentially interacting miRNAs (Table 2). In total, 12 mature miRNAs were found, 8 binding within the promoter and 4 within the 5′UTR.

Table 2. Identification of microRNAs (miRNAs) potentially interacting with the *S. miltiorrhiza HMGR4* promoter sequence and 5′UTR identified by the miRBase tool.

miRNA Name and Source	miRNA Sequence	Sequence Alignment Position Start/End	Strand	e-Value
HMGR4 **Promoter**				
miR1128 *Saccharum sp.*	UACUACUCCCUCCGUCCCAAA	350/368 405/423	+ −	0.75 4.2
miR6462c-5p *Populus trichocarpa*	AAGGGACAAAAAUGGCAUAAGA	259/279	−	3.5
miR1128 *Triticum aestivum*	UACUACUCCCUCCGUCCGAAA	350/368	+	4.2
miR1436 *Oryza sativa* and *Hordeum vulgare*	ACAUUAUGGGACGGAGGGAGU	354/368	−	6.2
miR5205a *Medicago truncatula*	CAUACAAUUUGGGACGGAGGGAG	355/374	−	9.1
miR8740 *Gossypium raimondii*	UAAUGAUGUGGCACAAUAUUA	634/653	−	9.1
miR11573a and miR11573b *Picea abies*	UUGGGGAGCGUAUUGUAGAUU	197/216	−	9.1
5′UTR of *HMGR4*				
miR477 *Gossypium raimondii*	CGAAGUCUUGGAAGAGAGUAA	59/75	−	3.2
miR6180 *Hordeum vulgare*	AGGGUGGAAGAAAGAGGGCG	55/69	−	3.9
miR4993 *Glycine max*	GAGCGGCGGCGGUGGAGGAUG	13/30	−	6.9
miR12107-5p *Citrus sinensis*	CUGAUGAGAGAGCGAAUGAUA	51/66	−	8.4

2.2. Microarray and Next-Generation Sequencing (NGS) Co-Expression Data Analysis

The Protein BLAST analysis revealed that the coding sequence of *S. miltiorrhiza HMGR4* (AEZ55673.1) was more similar to *A. thaliana HMGR1* (NP_177775.2), with an identity of 73.76%, than to *A. thaliana HMGR2* (NP_179329.1), with one of 69.48%. Additionally, a phylogenetic study of coding sequences indicated that *HMGR4* from *S. miltiorrhiza* (JN831103.1) and *HMGR1* from *A. thaliana* (NM_106299.4) were more closely related to each other (Figure S1). Therefore, the *A. thaliana HMGR1* gene (At1g76490) was used for further co-expression research. As a result of the conducted microarray analysis, 166 TF genes co-expressed with *A. thaliana HMGR1* in the *r* range of 0.5–1.0 were found: 41 in AtGenExpress Hormone and Chemical Compendium, 37 in AtGenExpress Abiotic Stress Compendium, 34 in AtGenExpress Pathogen Compendium, 25 in AtGenExpress Tissue Compendium, and 29 in AtGenExpress Plus—Extended Tissue Compendium (Table S3). The RNA-seq analysis did not identify any TF genes co-expressed with *A. thaliana HMGR1* in the WGCNA correlation range of 0.5–1.0.

2.3. Comparison of the in Silico HMGR4 Analysis Results with Microarray Co-Expression Data

The comparison identified 32 common TFs in the *S. miltiorrhiza HMGR4* promoter (Table 3), with the most well-represented being TFs from the homeodomain-leucine zipper (HD-ZIP) and WRKY families. The common TFs participated mainly in response to hormones (ABA, ethylene, jasmonic acid and cytokinins), other abiotic factors (light, salt stress, water deprivation, heat and iron ion) and biotic agents (bacteria), embryogenesis, organogenesis (root development), flowering or, finally, tissue development (epidermis) (Table 3).

Table 3. TFs common between in silico analysis of the *S. miltiorrhiza HMGR4* promoter and microarray co-expression studies with *A. thaliana HMGR1*.

TF Family Name	TF Gene Name and Locus	Processes in Which TF is Involved [a]	r-Value [b]	TFBS Motif and Localisation [c,d]	
Homeodomain; HD-ZIP	*ATHB-13*; At1g69780	cotyledon and leaf morphogenesis; primary root development; sucrose-signalling pathway	0.594	ATAAT 310; 309	AATAA 308; 307
	ATHB-16; At4g40060	regulation of timing of transition from vegetative to reproductive phase; repression of cell expansion during plant development; response to blue light	0.562	ATAAT 309	
	HDG1; At3g61150	maintenance of floral organ identity	0.507	ATTAA 161	**TTAAT 1218; 1331; 1332**
	ANL2; At4g00730	regulation of tissue-specific accumulation of anthocyanins; cellular organisation of primary root; cuticle hydrocarbon biosynthetic process; plant-type cell wall modification; root hair cell differentiation	0.524	**TTAAT 1218**	ATTAA 1161
	ATML1; At4g21750	cotyledon development; epidermal cell differentiation; seed dormancy and germination	0.583	**TTAAT 1332** **ATTTA** 1057; **1282**	**TAAAT** 987; **1272; 1346**
	PDF2; At4g04890	cotyledon development; epidermal cell differentiation; seed dormancy and germination; maintenance of floral organ identity	0.587	**ATTTA** 1057; **1282**	**TAAAT** 987; **1272; 1346**
Homeodomain; bZIP; HD-ZIP	*HAT5*; At3g01470	leaf morphogenesis; response to blue light and salt stress	0.521	ATAAT 307; 310	AATAA 308; 307
bZIP	*BZIP25*; At3g54620	positive regulation of seed maturation	0.666	CCACG 822 TACGT 46; 720 ACGTA 47; 721	AACGT 290;1188 ACGTT 291; 1189
WRKY	*WRKY2*; At5g56270	regulation of basal cell division patterns during early embryogenesis; establishment of cell polarity; longitudinal axis specification; pollen development	0.575/ 0.557	TGACT 5; 832 **AGTCA** 111; **1240**	TTGAC 831 **GTCAA** 913; 1147; **1251**
	WRKY14; At1g30650	-	0.576		

Table 3. *Cont.*

TF Family Name	TF Gene Name and Locus	Processes in Which TF is Involved [a]	r-Value [b]	TFBS Motif and Localisation [c,d]	
	WRKY57; At1g69310	response to osmotic stress, salt stress and water deprivation	0.504		
	WRKY45; At3g01970	phosphate ion transport	0.515	TTGAC 830; 831	**AGTCA** 111; **1240**
	WRKY69; At3g58710	-	0.564	TGACT 5; 832	**GTCAA** 913; 1147; **1251**
Myb/SANT; ARR-B	*ARR2*; At4g16110	His-to-Asp phosphorelay signal transduction system; expression of nuclear genes for components of mitochondrial complex I; ethylene- and cytokinin-activated signalling pathways; promotion of cytokinin-mediated leaf longevity; root meristem growth; seed growth; stomatal movement	0.552/0.611	AATCT 15; 197; 919 **AGATT 1405**	AATCC 179; 204; 548
	ARR14; At2g01760	His-to-Asp phosphorelay signal transduction system; activation of some type-A response regulators in response to cytokinins	0.509/ 0.530		
Myb/SANT; MYB	*MYB6*; At4g09460	response to ethylene, abscisic acid (ABA), indole-3-acetic acid, and *Pseudomonas syringae* pv. *phaseolica*	0.599	ACCTA 886	
MYB-related	*RVE1*; At5g17300	morning-phased TF integrating circadian clock and auxin pathways; regulation of free indole-3-acetic acid level in time-of-day specific manner; negative regulation of freezing tolerance	0.501	ATATC 1166	
	RVE4; At5g02840	regulation of circadian rhythm	0.545/ 0.573	ATATC 1166	**GATAT 1215**
EIN3; EIL	*EIL1*; At2g27050	positive regulation of ethylene response pathway; cellular response to iron ion; defence response to bacterium	0.554	TGTAT 374; 759	ATACA 391
	EIL3; At1g73730	ethylene response pathway; sulphur metabolic process; cellular response to iron ion	0.525	ATGTA 757	

Table 3. *Cont.*

TF Family Name	TF Gene Name and Locus	Processes in Which TF is Involved [a]	r-Value [b]	TFBS Motif and Localisation [c,d]	
MADS box; MIKC	*AGL18*; At3g57390	negative regulation of flowering and short-day photoperiodism; pollen development	0.567	TTTCC 804; 801 TTTTG 805 **CAAAA 1396**	**AGAAA** 293; **1402** **GGAAA 1364; 1362** TTTTT 77; 78; 79; 80; 81; 82; 83; 84
	SVP; At2g22540	inhibition of floral transition in autonomous flowering pathway; identity of floral meristem; response to temperature stimulus	0.537	TTTCC 801	**GGAAA 1362**
NAC; NAM	*NAC055*; At3g15500	jasmonic acid-mediated signalling pathway; response to water deprivation	0.594	TACGT 44; 718 CGTA 720 ACAT 644	TTGAC 829 ACGTA 44; 718; 720
	NAC072; At4g27410	activator in ABA-mediated dehydration response	0.543	CGTA 720 TTGAC 829	ACAT 644
NF-YB	*NFYB5*; At2g47810		0.558	CTAAT 42	**TCAAT** 13; 1149; **1253**
NF-YC	*NFYC10*; At5g38140	protein heterodimerization activity	0.523	ATCGG 102; 131 CCCAT 139 CCAAG 149; 213 CCAAT 195; 202 GTTGG 389 ATGGG 959 ATTGC 1039 CGAAT 1115 **ATTAG 1389** CCTAT 890 TTTGG 811	AATGG 462; 847 **CCACT** 144; 380; **1364** CCATT 169; 266; 701 CCAAC 530 AGTGG 592 ATTGA 766; 1099 **ACAAT** 460; 800; **1247** ATTGT 599; 637; 1145; 1152 **CCAAA** 176; 534; **1261**; **1396** **CAAAT** 177; 469; 897; 1107; 1161; **1262**; **1313**
TBP	*TBP2*; At1g55520	required for basal transcription (facilitating the recruitment of TFIID to the promoter, forming a preinitiation complex with RNA polymerase)	0.645	ATATA 746; 739 TTTTA 1022; 1078 **TAAAA** 541; **1465**; 1024	**TATAT** 677; **1305**; 743; 736 **ATAAA** 1025; **1466**; **1344**; 298; 439; 1024; **1465** **TTTAT** 669; 1016; 1072; **1297**; **1223**; **1303**; 674; 1077; **1302**
TCP	*TCP21*; At5g08330	positive regulation of circadian clock	0.546	CCCAC 818	

Table 3. *Cont.*

TF Family Name	TF Gene Name and Locus	Processes in Which TF is Involved [a]	r-Value [b]	TFBS Motif and Localisation [c,d]	
AP2; ERF	*RAP2-3;* At3g16770	cell death; heat acclimation; ethylene-activated signalling pathway; response to cytokinin, jasmonic acid and other organism	0.502	TAAGA 494	
C2H2	*AZF2;* At3g19580	inhibition of plant growth under abiotic stress conditions; negative regulation of ABA signalling during seed germination; positive regulation of leaf senescence; jasmonate early signalling response; response to chitin and water deprivation; plants overexpressing AZF2 have increased sensitivity to salt stress and barely survive under high salt conditions	0.503	**ACACT** 29; **1462**	
Dof	*DOF5.4;* At5g60850	metal ion binding; binding of OBF TFs to OCS elements	0.727	CGTTA 685 AACGT 286; 1184 ACGTT 289; 1187 GCCTT 79 CCTTT 80; 807 AAAGT 109; 982 AAGGG 1031 **AAGGA 1372**	ACCTT 158; 184; 605; 612 GCTTT 369 AAAGC 400; 571 **AAAGA** 939; **1413**; **1420** **AAAGG** 1030; **1289**; **1371** TCTTT 252; 328; 586; 874 TCCTT 275; 338; 806; 1156 **ACTTT** 382; 509; 881; 1021; **1291**
SBP	*SPL3;* At2g33810	promotion of vegetative phase change and flowering; vegetative- to reproductive-phase transition of meristem	0.596	AGTAC 51; 972 GTACA 578; 973 TGTAC 577 GTACT 52 ATACG 43 CGTAG 46	CTTAC 275; 427; 625 CGTCC 360 GGACG 405 CGAAC 524 CTACG 717 CGTAA 720

[a] The roles of the TFs were assumed based on the UniProt database. [b] r-value between the TFBSs found in the *S. miltiorrhiza HMGR4* promoter and those detected in *A. thaliana* ranged from 0.5 to 1.0. [c] For TFBSs, only the most conserved positions within a matrix were listed. [d] TFBSs localised in proximal promoter region were put in bold.

These 32 common TFs were scanned with the Genomatix Pathway System for the presence of interactions between them. The identified relationships are presented in Figure 4. It was found that SVP, AGL18 and SPL3 proteins were involved together in the regulation of flowering, and PDF2 and ATML1 in epidermal specification in embryos, respectively. Furthermore, NAC3 and RD26 responded to high salinity, drought and ABA, while NAC3 and ZF2 supported resistance to the herbivore *Spodoptera littoralis*. EIL3 and EIL1 played roles in regulating the response to sulphur deficiency and in ethylene signalling, and EBP transcription was light modulated through the EIN2-EIN3/EIL1 pathway.
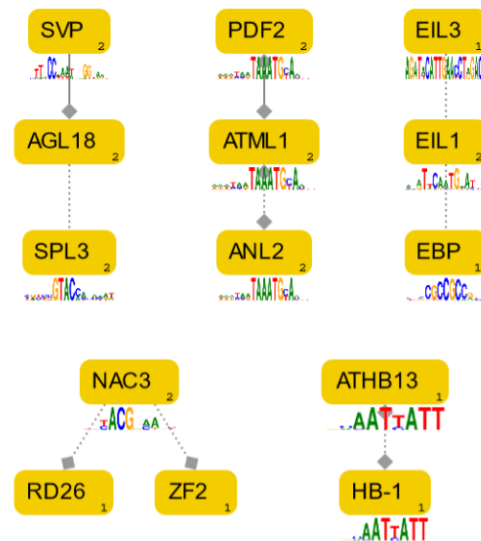
Figure 4. Interactions between TFs potentially binding to the *S. miltiorrhiza HMGR4* promoter found with the Pathway System tool. The presented dependencies are based on co-citation (dashed line) or expert-curation (solid line). Diamond-ended lines indicate that a given TF has a predicted binding site in dependent promoter sequence. The number in the lower right corner of TF indicates the number of interactions within the network (including those not displayed). EBP = RAP 2-3, NAC3 = NAC055, RD26 = NAC072, ZF2 = AZF2, HB-1 = HAT5.

Of the 32 common TFs that were recognised, 18 were found to interact with the TFBSs situated in the proximal promoter region (Table 3). The ability of these 18 TFs to bind to DNA as dimers or multimers was also tested. The TFBSs identified for HD-ZIP (ATML1, PDF2, HDG1), WRKY (WRKY2, WRKY14, WRKY45, WRKY57, WRKY69) and Dof (DOF5.4) are given in Table 1; all are closely located to each other, and only separated by a few nucleotides. The existence of experimentally-determined interactions between ATML1 and PDF2 proteins was confirmed by the BioGRID database.

2.4. Comparison of S. miltiorrhiza HMGR Promoters

The *S. miltiorrhiza HMGR1*, *HMGR2* and *HMGR4* promoter sequences were analysed using the Common TFs tool. Based on the findings, common binding sites for TFs were recognised, and these were classified into 22 matrix families (Figure 5, Table 4), with each single matrix family comprising identical or functionally-similar TFs identified by weight matrices. The following matrix families were found: *Arabidopsis* homeobox proteins (P$AHBP), L1 box (P$L1BX), MYB IIG-type binding sites (P$MIIG), DNA binding with one finger (P$DOFF), GT box elements (P$GTBX), MADS box proteins (P$MADS), MYB-like proteins (P$MYBL), MYB proteins with single DNA binding repeat (P$MYBS), NAC factors with transmembrane motif (P$NTMF), plant specific NAC proteins (P$NACF), transcription repressor KANADI (P$KAN1), W box family (P$WBXF), time-of-day-specific regulatory elements (P$TODS), nodulin consensus sequence 1 (P$NCS1), sweet potato DNA-binding factor with two WRKY domains (P$SPF1), zinc finger proteins (P$ZFAT), light-responsive elements (P$LREM), protein secretory pathway elements (P$PSPE), CGCG

box binding proteins (P\$CGCG), proteins involved in programmed cell death response (P\$PCDR), plant nitrate-responsive elements (P\$PNRE) and finally, stomatal carpenter (P\$SCAP). As can be seen in Figure 5, the distribution of the matrices within the *HMGR1* and *HMGR2* promoter sequences was strikingly similar. The above TF family analysis found that the *S. miltiorrhiza HMGR1*, *HMGR2* and *HMGR4* genes can be co-regulated in response to abiotic factors (auxins, gibberellins, ABA, SA, jasmonic acid, brassinosteroids, light, water deprivation, salt stress, cold or phosphate starvation), biotic factors (bacteria, fungi and viruses) and during root, stem, leaf and flower organogenesis (Table 4).
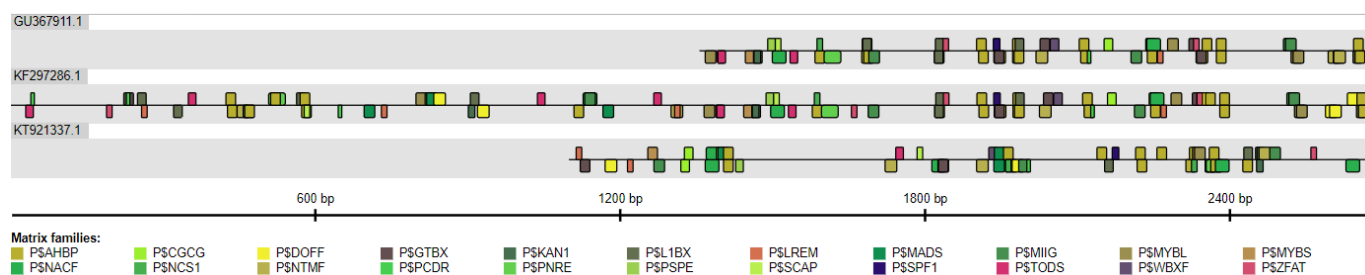

Figure 5. Distribution of matrix families common to the *S. miltiorrhiza* promoter sequences, i.e., *HMGR1* (GU367911.1), *HMGR2* (KF297286.1) and *HMGR4* (KT921337.1) identified by the Common TFs tool. Black lines correspond to the promoter sequences. Each matrix family is marked with a semicircular coloured symbol. The figure shows families found on the positive and negative strands. P\$AHBP, *Arabidopsis* homeobox proteins; P\$L1BX, L1 box; P\$MIIG, MYB IIG-type binding sites; P\$DOFF, DNA binding with one finger; P\$GTBX, GT box elements; P\$MADS, MADS box proteins; P\$MYBL, MYB-like proteins; P\$MYBS, MYB proteins with single DNA binding repeat; P\$NTMF, NAC factors with transmembrane motif; P\$NACF, plant specific NAC proteins; P\$KAN1, transcription repressor KANADI; P\$WBXF, W box family; P\$TODS, time-of-day-specific regulatory elements; P\$NCS1, nodulin consensus sequence 1; P\$SPF1, sweet potato DNA-binding factor with two WRKY domains; P\$ZFAT, zinc finger proteins; P\$LREM, light-responsive elements; P\$PSPE, protein secretory pathway elements; P\$CGCG, CGCG box binding proteins; P\$PCDR, proteins involved in programmed cell death response; P\$PNRE, plant nitrate-responsive elements; P\$SCAP, stomatal carpenter.

Table 4. Common TF matrix families found during in silico analysis of the *S. miltiorrhiza HMGR1*, *HMGR2*, *HMGR4* promoter sequences using the Common TFs tool.

TF Matrix Family	Processes in Which TF Is Involved [a]
Arabidopsis homeobox proteins (P\$AHBP)	root, leaf and anther development; seed maturation; meristem initiation and growth; xylem and phloem pattern formation; cell differentiation; determination of bilateral symmetry; transition from vegetative to reproductive phase; glucosinolate metabolic process; response to: auxin, gibberellin, ABA, water deprivation, blue light and salt stress
L1 box (P\$L1BX)	cotyledon development; seed germination and dormancy; epidermal cell differentiation; maintenance of floral organ identity
MYB IIG-type binding sites (P\$MIIG)	root, seed, stamen and xylem development; cellular cadmium ion homeostasis; gibberellin and flavonol biosynthesis; defence response to fungi; response to: ABA, chitin, salt stress, cold, water deprivation, phosphate starvation, potassium ion and light
DNA binding with one finger (P\$DOFF)	secondary shoot, cotyledon and seed development; cell wall modification; cell cycle; gibberellin biosynthesis; response to: SA, auxin, chitin, red light and cold
GT box elements (P\$GTBX)	shoot system and stomatal complex development; trichome morphogenesis; seed maturation and germination; cell size and growth; response to: auxin and water deprivation

Table 4. *Cont.*

TF Matrix Family	Processes in Which TF Is Involved [a]
MADS box proteins (P$MADS)	flower, ovule and seed coat development; seed maturation; meristem structural organisation; transition from vegetative to reproductive phase; short-day photoperiodism; circadian rhythm; response to auxin
MYB-like proteins (P$MYBL)	integument, anther and pollen development; leaf morphogenesis; seed growth and dormancy; endothelial cell proliferation; vacuole organisation; wax biosynthesis; long-day photoperiodism; defence response to bacteria and fungi; response to: SA, brassinosteroid, gibberellin, ABA, jasmonic acid, chitin, salt, water deprivation and cold
MYB proteins with single DNA binding repeat (P$MYBS)	leaf and lateral root development; leaf senescence; circadian rhythm; peroxidase activity; auxin and sulphate ion homeostasis; response to: ABA, phosphate starvation, absence of light and high light intensity
NAC factors with transmembrane motif (P$NTMF)	leaf and trichome morphogenesis; xylem development; seed germination; photoperiodism; membrane protein proteolysis; response to: gibberellin, salt stress
plant specific NAC proteins (P$NACF)	leaf and secondary shoot development; primary shoot apical meristem specification; formation of organ boundary; regulation of timing of organ formation; response to water deprivation
transcription repressor KANADI (P$KAN1)	phenylpropanoid metabolic process
W box family (P$WBXF)	induced systemic resistance; JA-mediated signalling pathway; phosphate ion transport; defence response to: bacteria, fungi and viruses; response to: SA, chitin and wounding
time-of-day-specific regulatory elements (P$TODS)	circadian rhythm; red or far-red light signalling pathway; response to temperature
nodulin consensus sequence 1 (P$NCS1)	nodule-specific expression
zinc finger proteins (P$ZFAT)	regulation of root development; phosphate ion homeostasis
light-responsive elements (P$LREM)	response to hypoxia
protein secretory pathway elements (P$PSPE)	SA induction of secretion-related genes via NPR1
CGCG box binding proteins (P$CGCG)	leaf senescence; defence response to: bacteria, fungi and insects; response to: cold, auxins and water deprivation
proteins involved in programmed cell death response (P$PCDR)	regulation of expression of vacuolar processing enzyme
plant nitrate-responsive elements (P$PNRE)	nitrate assimilation; stomatal movement; response to: nitrate and water deprivation
stomatal carpenter (P$SCAP)	stomatal movement
sweet potato DNA-binding factor with two WRKY domains (P$SPF1)	-

[a] The roles of the TFs were assumed based on the MatInspector (Genomatix) database.

The FrameWorker tool indicated the existence of 10,000 10-element-frameworks within the *S. miltiorrhiza HMGR1*, *HMGR2* and *HMGR4* promoters. Two selected models are provided in Figure 6. The frameworks were created based on 52 matrix families common to the tested sequences, some of which are mentioned above in the Common TF results section. The matrix families were located on the positive or negative strand of the promoters.
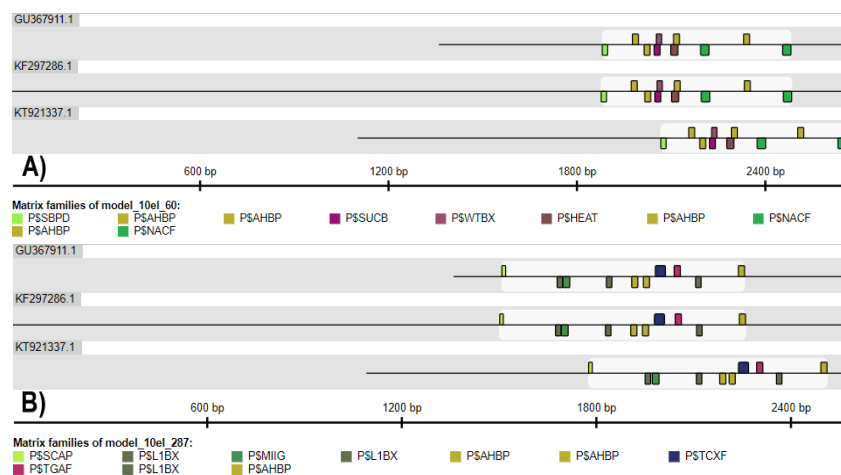
Figure 6. Selected 10-element-frameworks of TFBSs obtained for *S. miltiorrhiza* promoter sequences, i.e., *HMGR1* (GU367911.1), *HMGR2* (KF297286.1) and *HMGR4* (KT921337.1) using the FrameWorker tool. Black lines correspond to the promoter sequences. Each matrix family is marked with a semicircular coloured symbol. The figure shows families found on the positive and negative strands. (**A**) P\$SBPD, SBP-domain proteins; P\$AHBP, *Arabidopsis* homeobox proteins; P\$SUCB, sucrose box; P\$WTBX, WT box; P\$HEAT, heat shock factors; P\$NACF, plant specific NAC proteins; (**B**) P\$SCAP, stomatal carpenter; P\$L1BX, L1 box; P\$MIIG, MYB IIG-type binding sites; P\$AHBP, *Arabidopsis* homeobox proteins; P\$TCXF, CRC domain containing tesmin/TSO1-like CXC (TCX) factors; P\$TGAF, basic/leucine zipper-type TFs of the TGA-family.

The DiAlign TF tool analysis found the *HMGR1* (GU367911.1) and *HMGR2* (KF297286.1) promoters to demonstrate the greatest similarity (97%). In contrast, only 17% similarity was found between *HMGR4* (KT921337.1) and *HMGR2*, and 14% between *HMGR4* and *HMGR1*. The greatest number of identical areas was revealed in the proximal fragments of the analysed promoters, as well as in the beginning and the middle of the distal parts. Common overlapping TFBSs were identified in locations where all three tested sequences showed high local similarity; these included two binding sites for *Arabidopsis* homeobox proteins (P\$AHBP), one site for SBP domain proteins (P\$SBPD), one site for W box family proteins (P\$WBXF), one site for DNA binding with one finger factors (P\$DOFF), one GT box element (P\$GTBX) and, finally, one L1 box (P\$L1BX). The PlantPan 2.0 and MatInspector (Genomatix) databases indicated that P\$AHBP proteins are mainly involved in the response to hormones (auxins, ABA, cytokinins and gibberellins) and the initiation and development of shoot, root and flower meristems. P\$SBPD TFs are associated with inflorescence development, flowering and leaf epidermis differentiation. In turn, P\$WBXF matrix family responds to hormonal stimulation (SA, ABA, ethylene and jasmonic acid), other abiotic factors (salt stress, wounding, osmotic stress, heat, water deprivation and cold) and biotic agents (bacteria, fungi and viruses), and also participate in leaf senescence. P\$DOFF proteins are primarily involved in the regulation of flowering, circadian rhythm and in response to hormones (auxins and SA). P\$GTBX factors participate in the organogenesis of flowers and shoots. In addition, P\$L1BX proteins are needed for epidermis development and seed germination. These data are available in File S1 and Table 4.

2.5. The Conservation of Plant HMGR Promoters

MEGA X software alignment of 36 sequences spanning the proximal promoters and 5′UTRs of the plant *HMGR* genes revealed the presence of conserved regions; these are marked in blue in Figure 7. These regions were detected in both the 5′UTRs and proximal promoters. In most of the tested sequences, PRS was identified within the preserved areas. Additional analysis with the DiAlign TF tool revealed the presence of conserved TFBSs. However, no TFBS was found to be conserved in any of the analysed sequences. The most conserved site was the TATA box, detected in 41.7% of the sequences. Two

preserved binding sites for TFs belonging to the P$AHBP (*Arabidopsis* homeobox protein) and P$GCCF (GCC box family) families were detected in 27.8% of cases. Sites for P$TDTF (transposase-derived proteins), P$MYBL (MYB-like proteins) and P$L1BX (L1 box) proteins were identified in 25% of the sequences. Binding sites for P$DREB (dehydration responsive element binding factors) and P$ROOT (root hair-specific *cis*-elements in angiosperms) families were found in 22.2%. These TFBSs are highlighted in red in Figure 7.
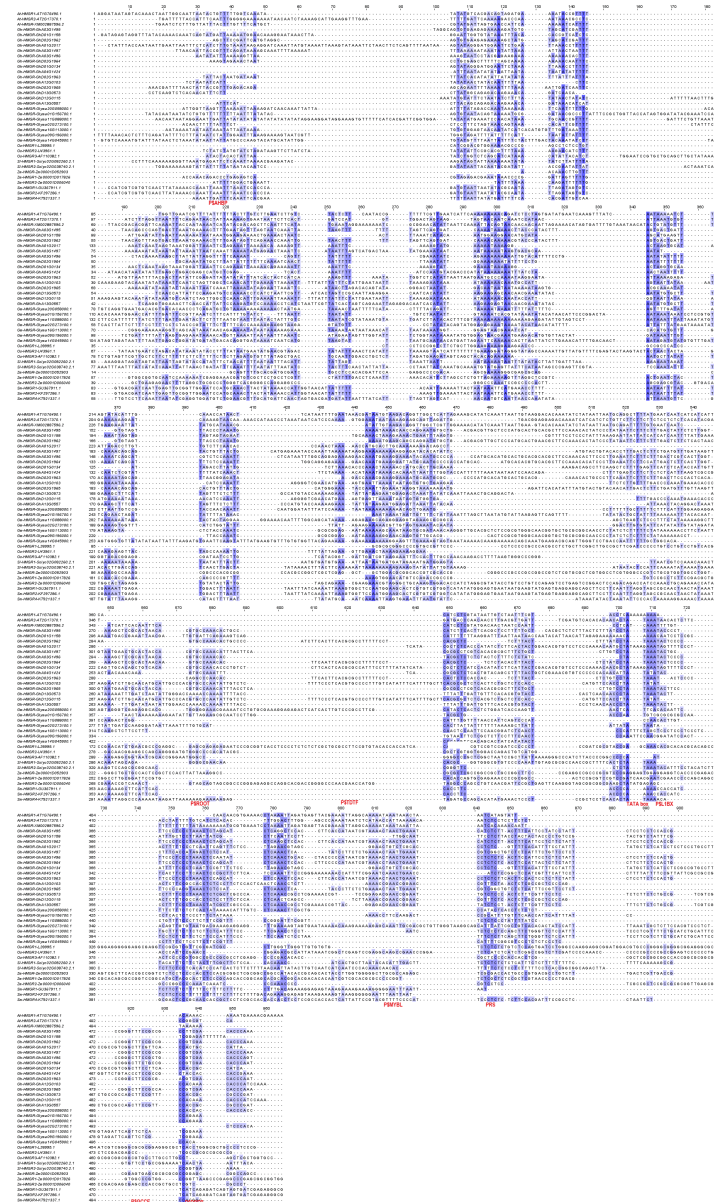

Figure 7. Alignment of proximal promoter regions and 5′UTRs of *Arabidopsis thaliana* (At), *Arabidopsis lyrata* (Al), *Gossypium hirsutum* (Gh), *Glycine max* (Gm), *Oryza sativa* (Os), *Solanum lycopersicum* (Sl), *Zea mays* (Zm) and *Salvia miltiorrhiza* (Sm) *HMGR* genes. Conserved nucleotides are marked in blue. The darker the colour, the greater the degree of conservation within the analysed sequences. The proximal promoters are present at the beginning and the 5′UTRs at the end of the displayed sequences, respectively. Conserved TFBSs identified by DiAlign TF tool are highlighted in red. P$AHBP, *Arabidopsis* homeobox protein; P$GCCF, GCC box family; P$TDTF, transposase-derived proteins; P$MYBL, MYB-like proteins; P$L1BX, L1 box; P$DREB, dehydration responsive element binding factors; P$ROOT, root hair-specific cis-elements in angiosperms.

Apart from the conserved TATA box and PRS motifs, the investigated *S. miltiorrhiza HMGR4* promoter was found to contain several other common binding sites for TFs, belonging to P$AHBP (*Arabidopsis* homeobox protein), P$GTBX (GT box elements), P$DOFF (DNA binding with one finger) and P$L1BX (L1 box); these were shared by 8–13% of the tested sequences. Nucleotide pairwise alignment of *S. miltiorrhiza HMGR4* with the remaining tested sequences found 13 to 31% identity (mean 24.7%).

3. Discussion

Our study presents new data regarding the isolated *S. miltiorrhiza HMGR4* promoter and 5′UTR and compares these sequences with other plant *HMGR*s.

Initially, the sequences were examined for the presence of certain distinctive motifs (Figure 1). One such motif found in the investigated sequences is the TATA box, which is estimated to be present in 30–50% of all known promoters [41] and 29% of *A. thaliana* promoters [42]. It was also detected in the *S. miltiorrhiza HMGR2* promoter [43]. Previous studies in human and yeast models indicate that the TATA box is more common in promoters of highly-regulated genes and in those stimulated by stress factors and extracellular signals [44–48]; in contrast, TATA-less genes are more constitutively expressed and associated with key processes such as cell growth [44–48]. In addition, promoters containing the TATA box appear to have a more conserved sequence than those that do not [49].

The *HMGR4* promoter is also characterised by the presence of a single tandem repeat. This motif is estimated to be present in only 25% of promoters [50], and is absent from the *S. miltiorrhiza HMGR2* promoter [43]. As tandem repeats are more prone to mutation, which affects the length of the repeat and thus local nucleosome positioning and gene expression rate, genes whose promoters have tandem repeats show higher rates of transcription divergence [50].

Both the *HMGR2* promoter and the studied *HMGR4* promoter lack CpG islands [43]. The cytosines in the CG dinucleotides of the islands can be methylated, thus inhibiting gene expression [51,52]. However, the CpG cluster is not required for methylation since, in plants, it can also occur within the CHG and CHH sequences (H = A, T or C) [53].

A PRS is also detected in the 5′UTR of the described sequence. This is a rather rare observation, but not an unprecedented one, as a PRS has also been found in the 5′UTR of the *S. miltiorrhiza HMGR2* gene [43]. It is believed to take part in the organisation of the spliceosomal complex [54].

Furthermore, the examined *S. miltiorrhiza HMGR4* promoter sequence turned out to be rich in TFBSs recognised by specific TFs (File S1). The conducted research indicates that the number of promoter regulatory elements and interacting proteins positively correlates with divergence of gene expression [55]. The *HMGR4* proximal promoter was found to contain consensus sequences mainly related to the response to light, SA, bacterial infection, auxins, ABA, gibberellin, chitin, cold or, finally, salt stress (Figure 2), suggesting that these factors may participate in gene regulation. One previous paper investigating the influence of external agents on *S. miltiorrhiza HMGR4* found that treatment with 200 μM MeJa had no significant effect on *HMGR4* expression in either leaves or roots [19]. It is important to note that the effect of these factors has been examined on other *S. miltiorrhiza HMGR* genes. Chen et al. found that only 100% red light slightly increased the expression of *HMGR* in hairy root culture, while other light types (e.g., 100% far-red, 100% blue, red:far-red, blue:far-red, red:blue, red:blue:UV) had an inhibitory effect [56]. In contrast, Wang et al. noted that UV-B enhanced the expression of *HMGR1* in roots almost 5-fold compared to an untreated control [57]. Incubation of hairy root culture with 100 μM SA raised *HMGR* transcript level, peaking at three-times higher than baseline after 36 h [58]. In turn, 200 μM SA has been found to have a differential effect on *HMGR2* promoter in leaf material [43]. A decrease in its activity was observed after 12, 24 and 48 h of treatment, while a 2.5- to 3-fold rise compared to the calibrator values was observed after 72 and 96 h. Bacteria appeared to be good activators of *HMGR* expression. The addition of *Pseudomonas brassicacearum* subsp. *neoaurantiaca* and *Pseudomonas thivervalensis* to *S. miltiorrhiza* root culture resulted in

2.1- and 1.5-fold enhancements in HMGR enzyme activity, respectively [59]. In addition, *Streptomyces pactum* Act12 increased *HMGR1* expression by more than a factor of 35 on day 14 relative to the calibrator [60]. Exposure to 2.85 µM IAA and 2.88 µM gibberellic acid improved the activity of the *HMGR2* promoter, resulting in manifold higher expression compared to the calibrator in 96 h [43]. In turn, 200 µM and 10 µM ABA upregulated *HMGR1* and *HMGR2*, respectively [43,61]. Salt stress (50 mM, 100 mM, 200 mM, 300 mM NaCl) enhanced the expression and enzymatic activity of *HMGR* in leaves and roots over 48 h of exposure [62]. It was also found that 200 mM NaCl inhibited the level of *HMGR1* transcript in leaves and roots as compared to the calibrator [63].

As non-coding regions are generally not highly conserved, from an evolutionary perspective, finding such motifs in the promoter or 5′UTR sequences suggest they have functional importance [64].

Within the studied 36 *HMGR* sequences, the most frequently-identified conserved motif was the TATA box (Figure 7). This is a known sequence that has been conserved from *Archaebacteria* to humans [65]. The other TFBSs discussed in the Results section were shared by a much smaller number of tested sequences (27.8% or fewer).

The study also examined the possibility that more complex structures could be created by TFs interacting with the *HMGR4* promoter. TFs participate in the regulation of gene expression as monomers, dimers (homo- and heterodimers) or multimers. Dimers and multimers are often preferred by nature because they allow specific interactions with the promoter sequence and bind with high affinity [66]. One TF monomer can create dimers or multimers with different functions, thus mediating the regulation of various genes, by forming bonds with multiple, but not random, protein partners [67]. Our analyses revealed the presence of closely-related TFBSs for the following TFs in the *S. miltiorrhiza HMGR4* proximal promoter: HD-ZIP (ATML1, PDF2 and HDG1), WRKY (WRKY2, WRKY14, WRKY45, WRKY57 and WRKY69) and Dof (DOF5.4) (Table 1). The HD-ZIP proteins are unique to the plant kingdom. TFs from the family are unable to bind to DNA as monomers [68]. They form homo- and heterodimers via the leucine zipper motif [67]. Meanwhile, ATML1 was able to create a heterodimer with its paralogue PDF2 in studies on *Nicotiana benthamiana* and *A. thaliana* [69,70], and to form homodimers in vitro [69,71]. It has been shown that WRKY TFs can interact with DNA as monomers or create homo- and heterodimers, especially those with a leucine zipper motif [72–74], WRKY2 protein was found to form homodimers in *Hordeum vulgare* [75], while WRKY45 created homodimers in vitro by exchanging β4-β5 strands in *Oryza sativa* [72]. The Dof TFs have a multifunctional domain that allows them to bind to DNA and interact with other proteins [76] and to establish homo- and heterodimers.

As miRNA is believed to regulate plant promoter activity at the transcription level, the investigated *HMGR4* sequence was searched for miRNA binding sites and interacting miRNAs [27]. Of the 12 miRNAs potentially binding to the *HMGR4* promoter and 5′UTR sequences (Table 2), non-conserved miR1128 and miR1436 were detected during deep sequencing in *S. miltiorrhiza* [77]. However, their significance in the regulation of gene expression has not yet been investigated at the experimental level.

The results of our present in silico analysis of the *HMGR4* promoter and 5′UTR sequences constitute a strong basis for planning future necessary experiments on *S. miltiorrhiza*.

4. Materials and Methods

4.1. Plant Material

S. miltiorrhiza plants were cultivated from seeds provided by the Garden of Medicinal Plants of the Medical University of Lodz. The plants were grown in pots containing composite soil at 26 ± 2 °C under natural light. Six-month-old plants were used for the experiment.

4.2. Isolation of the S. miltiorrhiza HMGR4 Promoter Sequence

Genomic DNA used for isolation of the *HMGR4* promoter was obtained from young, fresh *S. miltiorrhiza* leaves and stems according to the method proposed by Khan et al. [78]. The DNA was analysed using a NanoPhotometer P300 (Implen, Munich, Germany) to determine its quantity and quality based on A_{260}/A_{280} and A_{260}/A_{230} ratios. The *HMGR4* promoter region was isolated using GenomeWalker Universal Kit (Takara Bio, Kusatsu, Japan) according to the manufacturer's instructions. A 5′-terminal fragment of the *HMGR4* gene, deposited in GenBank under accession number JN831103.1, was used as a target for designing GSP1 and GSP2 specific primers (Table S4). The PCR reactions were performed using the Advantage 2 PCR Kit (Takara Bio, Kusatsu, Japan). The amplified DNA fragments were TOPO-TA cloned into a pCRII-TOPO vector (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The inserts were Sanger sequenced (CoreLab, Medical University of Lodz, Lodz, Poland) with specific primers listed in Table S4. The *HMGR4* promoter sequence was assembled using CodonCode Aligner software version 8.0.2 (CodonCode Corporation, Centerville, MA, USA).

The isolation and sequencing of the *S. miltiorrhiza HMGR4* promoter took approximately one month.

4.3. In Silico Analysis of the S. miltiorrhiza HMGR4 Promoter Sequence

The obtained *HMGR4* promoter sequence was characterised in silico using available tools and databases [79]. The promoter, TATA box, TSS and 5′UTR positions were identified with TSSP software (Softberry Inc., Mount Kisco, NY, USA) [80]. Tandem repeats, CpG islands, TFBSs and TFs were detected using PlantPan 2.0 [81]. The promoter sequence was screened for the presence of commonly-known consensus motifs reported in the published literature. Assuming that the functional TFBSs are concentrated mainly in the proximal promoters, special attention was paid to the promoter region lying within 300 bp from the TSS. The miRBase tool was used to search for miRNA binding sites and interacting miRNAs in the obtained promoter and 5′UTR sequences [82].

4.4. Microarray and NGS Co-Expression Data Analysis

Protein BLAST (NCBI, Bethesda, MD, USA) and MEGA X version 10.2.6 (Pennsylvania State University, State College, PA, USA) [83] were employed to determine which of the *A. thaliana HMGR* genes is a homologue of the *S. miltiorrhiza HMGR4* gene. Analyses were performed on coding sequences. Expression Angler (BAR, Toronto, ON, Canada) [84] and *Arabidopsis* RNA-seq Database [85] were utilised to find TF genes co-expressed with the selected *A. thaliana HMGR* gene. The Expression Angler tool has access to the expression results for approximately 22,000 *Arabidopsis* genes, while the *Arabidopsis* RNA-seq Database integrates 28,164 publicly available *Arabidopsis* RNA-seq libraries. The following microarray dataset compendiums were used during the study: AtGenExpress Hormone and Chemical, AtGenExpress Abiotic Stress, AtGenExpress Pathogen, AtGenExpress Tissue, and AtGenExpress Plus—Extended Tissue. The Pearson's correlation coefficient (r) ranging from 0.50 to 1.00 (moderate to strong positive correlation) was applied to identify co-regulated genes. Information on the detected TFs was obtained from the UniProt database [86]. The collected results were compared with the in silico data found by PlantPan 2.0. The occurrence of interactions between the received common TFs was determined using Pathway System (Genomatix, Munich, Germany) and the BioGRID database version 4.4.201 [87].

4.5. Comparison of S. miltiorrhiza HMGR Promoters

The entire available *S. miltiorrhiza HMGR* promoter sequences, i.e., *HMGR1* (GU367911.1), *HMGR2* (KF297286.1), and *HMGR4* (KT921337.1) were analysed with Common TFs, Frame-Worker and DiAlign TF tools from Genomatix, Munich, Germany. Common TFs was used for preliminary analysis of the common TFBSs and interacting TFs located anywhere in the investigated promoters. The search only included sites that were common to all three sequences. The similarity of the matrix to the tested sequences was set to the highest

possible value, i.e., 0.05. The FrameWorker tool permitted the common, most complex framework of TFBSs to be extracted from the input promoters. Frameworks are defined as TFBSs that occur in the same order and in a specified space range in all of the sequences. The DiAlign TF allowed for multiple alignment of the studied *HMGR* promoters, and revealed conserved regions and TFBSs located therein. The analyses using the Genomatix tools were based on matrix library version 11.3 and default search criteria.

4.6. Assessment of Conservation of Plant HMGR Promoters

The conservation of the 36 available *HMGR* promoters derived from plants such as *A. thaliana*, *Arabidopsis lyrata*, *Glycine max*, *Gossypium hirsutum*, *Oryza sativa*, *Solanum lycopersicum*, *Zea mays* and *S. miltiorrhiza* was assessed by aligning their sequences. Proximal promoters and 5'UTR sequences (each sequence 500 bases long) were obtained from Plant-Pan 3.0 [88] and NCBI Nucleotide databases with the participation of the UniProt [86]. Alignments were performed using the MUSCLE algorithm from the MEGA X software, version 10.2.6 [83]. TFBSs located in the conserved regions of the compared sequences were recognised with the DiAlign TF tool (Genomatix, Munich, Germany).

5. Conclusions

Regulation of *S. miltiorrhiza HMGR4* gene expression can occur during flowering, embryogenesis, organogenesis and circadian rhythm, and are influenced mainly by factors such as light, SA, bacterial infection and auxins.

The presence of binding sites for TFs that promote the biosynthesis of tanshinones may indicate that the *S. miltiorrhiza HMGR4* gene plays an important role in the production of these metabolites.

A comparison of TFBSs and TFs in the *S. miltiorrhiza HMGR1*, *HMGR2*, and *HMGR4* promoter sequences indicates that these genes can be co-regulated in response to abiotic and biotic factors, and during organogenesis.

The *S. miltiorrhiza HMGR4* promoter is not highly conserved.

Future research on the *S. miltiorrhiza HMGR4* promoter could be developed towards preparing promoter deletion mutants, and studying their transcriptional activity [89]. Moreover, mutagenesis of particular TFBSs could be suitable for experimental verification of their importance in response to biotic or abiotic factors [89]. The TFs or other regulatory proteins could be isolated using a yeast-one hybrid (Y1H) system and the promoter segments as bait [90]. Isolated TFs could be functionally characterised by studying their DNA binding properties, and their potential to increase expression of specific genes [91,92]. These studies could be verified by chromatin immunoprecipitation-sequencing (ChIP-seq) of DNA fragments that are associated with particular proteins [93]. Finally, regulatory networks of TFs and other proteins playing a pivotal role in the response to certain external factors could be built using transcriptomic RNA sequencing, and weighted gene co-expression network analysis (WGCNA) [94,95].

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

References

1. Xu, Y.; Wan, R.; Lin, Y.; Yang, L.; Chen, Y.; Liu, C. Recent advance on research and application of Salvia miltiorrhiza. *Asian J. Pharmacodyn. Pharmacokinet.* **2007**, *7*, 99–130.
2. Cheng, T.O. Cardiovascular effects of Danshen. *Int. J. Cardiol.* **2007**, *121*, 9–22. [CrossRef] [PubMed]
3. Zhang, H.S.; Wang, S.Q. Salvianolic acid B from Salvia miltiorrhiza inhibits tumor necrosis factor-alpha (TNF-alpha)-induced MMP-2 upregulation in human aortic smooth muscle cells via suppression of NAD(P)H oxidase-derived reactive oxygen species. *J. Mol. Cell. Cardiol.* **2006**, *41*, 138–148. [CrossRef] [PubMed]
4. Zhang, H.N.; An, C.N.; Zhang, H.N.; Pu, X.P. Protocatechuic acid inhibits neurotoxicity induced by MPTP in vivo. *Neurosci. Lett.* **2010**, *474*, 99–103. [CrossRef] [PubMed]
5. Wong, K.K.; Ho, M.T.; Lin, H.Q.; Lau, K.F.; Rudd, J.A.; Chung, R.C.; Fung, K.P.; Shaw, P.C.; Wan, D.C. Cryptotanshinone, an acetylcholinesterase inhibitor from Salvia miltiorrhiza, ameliorates scopolamine-induced amnesia in Morris water maze task. *Planta Med.* **2010**, *76*, 228–234. [CrossRef]
6. Kang, D.G.; Oh, H.; Sohn, E.J.; Hur, T.Y.; Lee, K.C.; Kim, K.J.; Kim, T.Y.; Lee, H.S. Lithospermic acid B isolated from Salvia miltiorrhiza ameliorates ischemia/reperfusion-induced renal injury in rats. *Life Sci.* **2004**, *75*, 1801–1816. [CrossRef]
7. Wu, Z.M.; Wen, T.; Tan, Y.F.; Liu, Y.; Ren, F.; Wu, H. Effects of salvianolic acid a on oxidative stress and liver injury induced by carbon tetrachloride in rats. *Basic Clin. Pharmacol. Toxicol.* **2007**, *100*, 115–120. [CrossRef]
8. Li, J.; Xu, M.; Fan, Q.; Xie, X.; Zhang, Y.; Mu, D.; Zhao, P.; Zhang, B.; Cao, F.; Wang, Y.; et al. Tanshinone IIA ameliorates seawater exposure-induced lung injury by inhibiting aquaporins (AQP) 1 and AQP5 expression in lung. *Respir. Physiol. Neurobiol.* **2011**, *176*, 39–49. [CrossRef]
9. Che, X.H.; Park, E.J.; Zhao, Y.Z.; Kim, W.H.; Sohn, D.H. Tanshinone II A induces apoptosis and S phase cell cycle arrest in activated rat hepatic stellate cells. *Basic Clin. Pharmacol. Toxicol.* **2010**, *106*, 30–37. [CrossRef]
10. Isacchi, B.; Fabbri, V.; Galeotti, N.; Bergonzi, M.C.; Karioti, A.; Ghelardini, C.; Vannucchi, M.G.; Bilia, A.R. Salvianolic acid B and its liposomal formulations: Anti-hyperalgesic activity in the treatment of neuropathic pain. *Eur. J. Biochem.* **2011**, *44*, 552–558. [CrossRef]
11. Yang, X.Y.; Qiang, G.F.; Zhang, L.; Zhu, X.M.; Wang, S.B.; Sun, L.; Yang, H.G.; Du, G.H. Salvianolic acid A protects against vascular endothelial dysfunction in high-fat diet fed and streptozotocin-induced diabetic rats. *J. Asian Nat. Prod. Res.* **2011**, *13*, 884–894. [CrossRef] [PubMed]
12. Serra, S.; Vacca, G.; Tumatis, S.; Carrucciu, A.; Morazzoni, P.; Bombardelli, E.; Colombo, G.; Gessa, G.L.; Carai, M.A. Anti-relapse properties of IDN 5082, a standardized extract of Salvia miltiorrhiza, in alcohol-preferring rats. *J. Ethnopharmacol.* **2003**, *88*, 249–252. [CrossRef]
13. Lu, S. *The Salvia Miltiorrhiza Genome*; Springer: Cham, Switzerland, 2019; pp. 129–139.
14. Liao, P.; Hemmerlin, A.; Bach, T.J.; Chye, M.L. The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* **2016**, *34*, 697–713. [CrossRef] [PubMed]
15. Hemmerlin, A. Post-translational events and modifications regulating plant enzymes involved in isoprenoid precursor biosynthesis. *Plant Sci.* **2013**, *203–204*, 41–54. [CrossRef] [PubMed]
16. Leivar, P.; Antolín-Llovera, M.; Ferrero, S.; Closa, M.; Arró, M.; Ferrer, A.; Boronat, A.; Campos, N. Multilevel control of Arabidopsis 3-hydroxy-3-methylglutaryl coenzyme A reductase by protein phosphatase 2A. *Plant Cell* **2011**, *23*, 1494–1511. [CrossRef]
17. Liao, P.; Zhou, W.; Zhang, L.; Wang, J.; Yan, X.; Zhang, Y.; Zhang, R.; Li, L.; Zhou, G.; Kai, G. Molecular cloning, characterization and expression analysis of a new gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from Salvia miltiorrhiza. *Acta Physiol. Plant.* **2009**, *31*, 565–572. [CrossRef]
18. Dai, Z.; Cui, G.; Zhou, S.F.; Zhang, X.; Huang, L. Cloning and characterization of a novel 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from Salvia miltiorrhiza involved in diterpenoid tanshinone accumulation. *J. Plant Physiol.* **2011**, *168*, 148–157. [CrossRef]
19. Ma, Y.; Yuan, L.; Wu, B.; Li, X.; Chen, S.; Lu, S. Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in Salvia miltiorrhiza. *J. Exp. Bot.* **2012**, *63*, 2809–2823. [CrossRef]

20. Hernandez-Garcia, C.M.; Finer, J.J. Identification and validation of promoters and cis-acting regulatory elements. *Plant Sci.* **2014**, *217–218*, 109–119. [CrossRef]

21. Yu, C.P.; Lin, J.J.; Li, W.H. Positional distribution of transcription factor binding sites in Arabidopsis thaliana. *Sci. Rep.* **2016**, *6*, 25164. [CrossRef]

22. Hussain, T.; Rehman, N.; Inam, S.; Ajmal, W.; Afroz, A.; Muhammad, A.; Zafar, Y.; Ali, G.M.; Khan, M.R. Homotypic clusters of transcription factor binding sites in the first large intron of AGL24 MADS-box transcription factor are recruited in the enhancement of floral expression. *Plant Mol. Biol. Rep.* **2019**, *37*, 24–40. [CrossRef]

23. Allen, B.L.; Taatjes, D.J. The Mediator complex: A central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 155–166. [CrossRef] [PubMed]

24. Khan, D.; Ziegler, D.J.; Kalichuk, J.L.; Hoi, V.; Huynh, N.; Hajihassani, A.; Parkin, I.A.P.; Robinson, S.J.; Belmonte, M.F. Gene expression profiling reveals transcription factor networks and subgenome bias during Brassica napus seed development. *Plant J.* **2022**, *109*, 477–489. [CrossRef]

25. Engels, B.M.; Hutvagner, G. Principles and effects of microRNA-mediated post-transcriptional gene regulation. *Oncogene* **2006**, *25*, 6163–6169. [CrossRef] [PubMed]

26. Pillai, R.S.; Bhattacharyya, S.N.; Filipowicz, W. Repression of protein synthesis by miRNAs: How many mechanisms? *Trends Cell Biol.* **2007**, *17*, 118–126. [CrossRef]

27. Yang, G.; Li, Y.; Wu, B.; Zhang, K.; Gao, L.; Zheng, C. MicroRNAs transcriptionally regulate promoter activity in Arabidopsis thaliana. *J. Integr. Plant Biol.* **2019**, *61*, 1128–1133. [CrossRef]

28. Ye, R.; Wang, W.; Iki, T.; Liu, C.; Wu, Y.; Ishikawa, M.; Zhou, X.; Qi, Y. Cytoplasmic assembly and selective nuclear import of Arabidopsis Argonaute4/siRNA complexes. *Mol. Cell* **2012**, *46*, 859–870. [CrossRef]

29. Place, R.F.; Li, L.C.; Pookot, D.; Noonan, E.J.; Dahiya, R. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1608–1613. [CrossRef]

30. Kim, D.H.; Saetrom, P.; Snøve, O., Jr.; Rossi, J.J. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16230–16235. [CrossRef]

31. Holland, C.K.; Jez, J.M. Arabidopsis: The original plant chassis organism. *Plant Cell Rep.* **2018**, *37*, 1359–1366. [CrossRef]

32. Woodward, A.W.; Bartel, B. Biology in Bloom: A Primer on the Arabidopsis thaliana Model System. *Genetics* **2018**, *208*, 1337–1349. [CrossRef] [PubMed]

33. Provart, N.J.; Alonso, J.; Assmann, S.M.; Bergmann, D.; Brady, S.M.; Brkljacic, J.; Browse, J.; Chapple, C.; Colot, V.; Cutler, S.; et al. 50 years of Arabidopsis research: Highlights and future directions. *New Phytol.* **2016**, *209*, 921–944. [CrossRef]

34. Cantó-Pastor, A.; Mason, G.A.; Brady, S.M.; Provart, N.J. Arabidopsis bioinformatics: Tools and strategies. *Plant J.* **2021**, *108*, 1585–1596. [CrossRef] [PubMed]

35. Dey, N.; Sarkar, S.; Acharya, S.; Maiti, I.B. Synthetic promoters in planta. *Planta* **2015**, *242*, 1077–1094. [CrossRef]

36. Shi, M.; Wang, Y.; Wang, X.; Deng, C.; Cao, W.; Hua, Q.; Kai, G. Simultaneous promotion of tanshinone and phenolic acid biosynthesis in Salvia miltiorrhiza hairy roots by overexpressing Arabidopsis MYC2. *Ind. Crops Prod.* **2020**, *155*, 112826. [CrossRef]

37. Deng, C.; Hao, X.; Shi, M.; Fu, R.; Wang, Y.; Zhang, Y.; Zhou, W.; Feng, Y.; Makunga, N.P.; Kai, G. Tanshinone production could be increased by the expression of SmWRKY2 in Salvia miltiorrhiza hairy roots. *Plant Sci.* **2019**, *284*, 1–8. [CrossRef]

38. Zhang, Y.; Xu, Z.; Ji, A.; Luo, H.; Song, J. Genomic survey of bZIP transcription factor genes related to tanshinone biosynthesis in Salvia miltiorrhiza. *Acta Pharm. Sin. B* **2018**, *8*, 295–305. [CrossRef] [PubMed]

39. Zhang, X.; Luo, H.; Xu, Z.; Zhu, Y.; Ji, A.; Song, J.; Chen, S. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in Salvia miltiorrhiza. *Sci. Rep.* **2015**, *5*, 11244. [CrossRef]

40. Chen, Y.; Wang, Y.; Guo, J.; Yang, J.; Zhang, X.; Wang, Z.; Cheng, Y.; Du, Z.; Qi, Z.; Huang, Y.; et al. Integrated Transcriptomics and Proteomics to Reveal Regulation Mechanism and Evolution of SmWRKY61 on Tanshinone Biosynthesis in Salvia miltiorrhiza and Salvia castanea. *Front. Plant Sci.* **2021**, *12*, 820582. [CrossRef]

41. Shahmuradov, I.A.; Umarov, R.K.; Solovyev, V.V. TSSPlant: A new tool for prediction of plant Pol II promoters. *Nucleic Acids Res.* **2017**, *45*, e65. [CrossRef]

42. Molina, C.; Grotewold, E. Genome wide analysis of Arabidopsis core promoters. *BMC Genom.* **2005**, *6*, 25. [CrossRef]

43. Szymczyk, P.; Grąbkowska, R.; Skała, E.; Żebrowska, M.; Balcerczak, E.; Jeleń, A. Isolation and characterization of a 3-hydroxy-3-methylglutaryl coenzyme A reductase 2 promoter from Salvia miltiorrhiza. *J. Plant Biochem. Biotechnol.* **2018**, *27*, 223–236. [CrossRef]

44. Tirosh, I.; Weinberger, A.; Carmi, M.; Barkai, N. A genetic signature of interspecies variations in gene expression. *Nat. Genet.* **2006**, *38*, 830–834. [CrossRef]

45. Yang, C.; Bolotin, E.; Jiang, T.; Sladek, F.M.; Martinez, E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* **2007**, *389*, 52–65. [CrossRef] [PubMed]

46. Basehoar, A.D.; Zanton, S.J.; Pugh, B.F. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **2004**, *116*, 699–709. [CrossRef]

47. Bae, S.H.; Han, H.W.; Moon, J. Functional analysis of the molecular interactions of TATA box-containing genes and essential genes. *PLoS ONE* **2015**, *10*, e0120848. [CrossRef] [PubMed]

48. de Jonge, W.J.; O'Duibhir, E.; Lijnzaad, P.; van Leenen, D.; Groot Koerkamp, M.J.; Kemmeren, P.; Holstege, F.C. Molecular mechanisms that distinguish TFIID housekeeping from regulatable SAGA promoters. *EMBO J.* **2017**, *36*, 274–290. [CrossRef] [PubMed]

49. Cliften, P.; Sudarsanam, P.; Desikan, A.; Fulton, L.; Fulton, B.; Majors, J.; Waterston, R.; Cohen, B.A.; Johnston, M. Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* **2003**, *301*, 71–76. [CrossRef] [PubMed]

50. Vinces, M.D.; Legendre, M.; Caldara, M.; Hagihara, M.; Verstrepen, K.J. Unstable tandem repeats in promoters confer transcriptional evolvability. *Science* **2009**, *324*, 1213–1216. [CrossRef]

51. Davey, C.; Pennings, S.; Allan, J. CpG methylation remodels chromatin structure in vitro. *J. Mol. Biol.* **1997**, *267*, 276–288. [CrossRef] [PubMed]

52. Bell, A.C.; Felsenfeld, G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **2000**, *405*, 482–485. [CrossRef] [PubMed]

53. Henderson, I.R.; Jacobsen, S.E. Epigenetic inheritance in plants. *Nature* **2007**, *447*, 418–424. [CrossRef] [PubMed]

54. Saulière, J.; Sureau, A.; Expert-Bezançon, A.; Marie, J. The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Mol. Cell. Biol.* **2006**, *26*, 8755–8769. [CrossRef] [PubMed]

55. Tirosh, I.; Barkai, N. Evolution of gene sequence and gene expression are not correlated in yeast. *Trends Genet.* **2008**, *24*, 109–113. [CrossRef]

56. Chen, I.J.; Lee, M.S.; Lin, M.K.; Ko, C.Y.; Chang, W.T. Blue light decreases tanshinone IIA content in Salvia miltiorrhiza hairy roots via genes regulation. *J. Photochem. Photobiol. B* **2018**, *183*, 164–171. [CrossRef]

57. Wang, C.H.; Zheng, L.P.; Tian, H.; Wang, J.W. Synergistic effects of ultraviolet-B and methyl jasmonate on tanshinone biosynthesis in Salvia miltiorrhiza hairy roots. *J. Photochem. Photobiol. B* **2016**, *159*, 93–100. [CrossRef]

58. Hao, X.; Shi, M.; Cui, L.; Xu, C.; Zhang, Y.; Kai, G. Effects of methyl jasmonate and salicylic acid on tanshinone production and biosynthetic gene expression in transgenic Salvia miltiorrhiza hairy roots. *Biotechnol. Appl. Biochem.* **2015**, *62*, 24–31. [CrossRef]

59. Yan, Y.; Zhang, S.; Zhang, J.; Ma, P.; Duan, J.; Liang, Z. Effect and mechanism of endophytic bacteria on growth and secondary metabolite synthesis in Salvia miltiorrhiza hairy roots. *Acta Physiol. Plant.* **2014**, *36*, 1095–1105. [CrossRef]

60. Yan, Y.; Zhang, S.; Yang, D.; Zhang, J.; Liang, Z. Effects of Streptomyces pactum Act12 on Salvia miltiorrhiza hairy root growth and tanshinone synthesis and its mechanisms. *Appl. Biochem. Biotechnol.* **2014**, *173*, 883–893. [CrossRef]

61. Yang, D.; Ma, P.; Liang, X.; Wei, Z.; Liang, Z.; Liu, Y.; Liu, F. PEG and ABA trigger methyl jasmonate accumulation to induce the MEP pathway and increase tanshinone production in Salvia miltiorrhiza hairy roots. *Physiol. Plant.* **2012**, *146*, 173–183. [CrossRef]

62. Yu, W.; Yu, Y.; Wang, C.; Zhang, Z.; Xue, Z. Mechanism by which salt stress induces physiological responses and regulates tanshinone synthesis. *Plant Physiol. Biochem.* **2021**, *164*, 10–20. [CrossRef] [PubMed]

63. Li, X.; Li, J.; Islam, F.; Najeeb, U.; Pan, J.; Hou, Z.; Shou, J.; Qin, J.; Xu, L. 5-Aminolevulinic acid could enhance the salinity tolerance by alleviating oxidative damages in Salvia miltiorrhiza. *Food Sci. Technol.* **2021**, *42*, 1–9. [CrossRef]

64. Setlow, J.K. *Genetic Engineering. Principles and Methods*, 1st ed.; Springer Science Business Media LLC: New York, NY, USA, 1997; pp. 15–47.

65. Reeve, J.N. Archaeal chromatin and transcription. *Mol. Microbiol.* **2003**, *48*, 587–598. [CrossRef]

66. Matthews, J.M. *Protein Dimerization and Oligomerization in Biology*; Springer Science+Business Media: New York, NY, USA, 2012; pp. 1–170.

67. Amoutzias, G.D.; Robertson, D.L.; Van de Peer, Y.; Oliver, S.G. Choose your partners: Dimerization in eukaryotic transcription factors. *Trends Biochem. Sci.* **2008**, *33*, 220–229. [CrossRef] [PubMed]

68. Ariel, F.D.; Manavella, P.A.; Dezar, C.A.; Chan, R.L. The true story of the HD-Zip family. *Trends Plant Sci.* **2007**, *12*, 419–426. [CrossRef]

69. San-Bento, R.; Farcot, E.; Galletti, R.; Creff, A.; Ingram, G. Epidermal identity is maintained by cell-cell communication via a universally active feedback loop in Arabidopsis thaliana. *Plant J.* **2014**, *77*, 46–58. [CrossRef]

70. Rombolá-Caldentey, B.; Rueda-Romero, P.; Iglesias-Fernández, R.; Carbonero, P.; Oñate-Sánchez, L. Arabidopsis DELLA and two HD-ZIP transcription factors regulate GA signaling in the epidermis through the L1 box cis-element. *Plant Cell* **2014**, *26*, 2905–2919. [CrossRef]

71. Abe, M.; Takahashi, T.; Komeda, Y. Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J.* **2001**, *26*, 487–494. [CrossRef]

72. Cheng, X.; Zhao, Y.; Jiang, Q.; Yang, J.; Zhao, W.; Taylor, I.A.; Peng, Y.L.; Wang, D.; Liu, J. Structural basis of dimerization and dual W-box DNA recognition by rice WRKY domain. *Nucleic Acids Res.* **2019**, *47*, 4308–4318. [CrossRef]

73. Eulgem, T.; Rushton, P.J.; Robatzek, S.; Somssich, I.E. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* **2000**, *5*, 199–206. [CrossRef]

74. Ciolkowski, I.; Wanke, D.; Birkenbihl, R.P.; Somssich, I.E. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol. Biol.* **2008**, *68*, 81–92. [CrossRef] [PubMed]

75. Shen, Q.H.; Saijo, Y.; Mauch, S.; Biskup, C.; Bieri, S.; Keller, B.; Seki, H.; Ulker, B.; Somssich, I.E.; Schulze-Lefert, P. Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* **2007**, *315*, 1098–1103. [CrossRef] [PubMed]

76. Yanagisawa, S. Dof DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. *Eur. J. Biochem.* **1997**, *250*, 403–410. [CrossRef] [PubMed]

77. Xu, X.; Jiang, Q.; Ma, X.; Ying, Q.; Shen, B.; Qian, Y.; Song, H.; Wang, H. Deep sequencing identifies tissue-specific microRNAs and their target genes involving in the biosynthesis of tanshinones in Salvia miltiorrhiza. *PLoS ONE* **2014**, *9*, e111679. [CrossRef]

78. Khan, S.; Qureshi, M.I.; Kamaluddin; Alam, T.; Abdin, M.Z. Protocol for isolation of genomic DNA from dry and fresh roots of medicinal plants suitable for RAPD and restriction digestion. *Afr. J. Biotechnol.* **2007**, *6*, 175–178. [CrossRef]

79. Majewska, M.; Wysokińska, H.; Kuźma, Ł.; Szymczyk, P. Eukaryotic and prokaryotic promoter databases as valuable tools in exploring the regulation of gene transcription: A comprehensive overview. *Gene* **2018**, *644*, 38–48. [CrossRef]

80. Shahmuradov, I.A.; Gammerman, A.J.; Hancock, J.M.; Bramley, P.M.; Solovyev, V.V. PlantProm: A database of plant promoter sequences. *Nucleic Acids Res.* **2003**, *31*, 114–117. [CrossRef]

81. Chow, C.N.; Zheng, H.Q.; Wu, N.Y.; Chien, C.H.; Huang, H.D.; Lee, T.Y.; Chiang-Hsieh, Y.F.; Hou, P.F.; Yang, T.Y.; Chang, W.C. PlantPAN 2.0: An update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res.* **2016**, *44*, D1154–D1160. [CrossRef]

82. Griffiths-Jones, S.; Grocock, R.J.; van Dongen, S.; Bateman, A.; Enright, A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **2006**, *34*, D140–D144. [CrossRef]

83. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]

84. Toufighi, K.; Brady, S.M.; Austin, R.; Ly, E.; Provart, N.J. The Botany Array Resource: E-Northerns, Expression Angling, and promoter analyses. *Plant J.* **2005**, *43*, 153–163. [CrossRef] [PubMed]

85. Zhang, H.; Zhang, F.; Yu, Y.; Feng, L.; Jia, J.; Liu, B.; Li, B.; Guo, H.; Zhai, J. A Comprehensive Online Database for Exploring ~20,000 Public Arabidopsis RNA-Seq Libraries. *Mol. Plant* **2020**, *13*, 1231–1233. [CrossRef] [PubMed]

86. UniProt, C. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]

87. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **2021**, *30*, 187–200. [CrossRef]

88. Chow, C.N.; Lee, T.Y.; Hung, Y.C.; Li, G.Z.; Tseng, K.C.; Liu, Y.H.; Kuo, P.L.; Zheng, H.Q.; Chang, W.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]

89. Luján, M.A.; Soria-García, Á.; Claver, A.; Lorente, P.; Rubio, M.C.; Picorel, R.; Alfonso, M. Different Cis-Regulatory Elements Control the Tissue-Specific Contribution of Plastid ω-3 Desaturases to Wounding and Hormone Responses. *Front. Plant Sci.* **2021**, *12*, 727292. [CrossRef]

90. Yuan, L.; Perry, S.E. *Plant Transcription Factors*, 1st ed.; Humana Press: Totowa, NJ, USA, 2011; pp. 45–66.

91. Zhang, X.; Ge, F.; Deng, B.; Shah, T.; Huang, Z.; Liu, D.; Chen, C. Molecular Cloning and Characterization of PnbHLH1 Transcription Factor in Panax notoginseng. *Molecules* **2017**, *22*, 1268. [CrossRef]

92. Chen, R.; Cao, Y.; Wang, W.; Li, Y.; Wang, D.; Wang, S.; Cao, X. Transcription factor SmSPL7 promotes anthocyanin accumulation and negatively regulates phenolic acid biosynthesis in Salvia miltiorrhiza. *Plant Sci.* **2021**, *310*, 110993. [CrossRef]

93. Chen, X.; Bhadauria, V.; Ma, B. ChIP-Seq: A Powerful Tool for Studying Protein-DNA Interactions in Plants. *Curr. Issues Mol. Biol.* **2018**, *27*, 171–180. [CrossRef]

94. Chen, H.; Yang, Q.; Fu, H.; Chen, K.; Zhao, S.; Zhang, C.; Cai, T.; Wang, L.; Lu, W.; Dang, H.; et al. Identification of Key Gene Networks and Deciphering Transcriptional Regulators Associated With Peanut Embryo Abortion Mediated by Calcium Deficiency. *Front. Plant Sci.* **2022**, *13*, 814015. [CrossRef]

95. Lei, X.; Liu, Z.; Xie, Q.; Fang, J.; Wang, C.; Li, J.; Wang, C.; Gao, C. Construction of two regulatory networks related to salt stress and lignocellulosic synthesis under salt stress based on a Populus davidiana × P. bolleana transcriptome analysis. *Plant Mol. Biol.* **2022**. [CrossRef] [PubMed]