Article

Mining of Root-Specific Expression Genes and Their Core Cis-Regulatory Elements in Plants

Shan Gao [†] [ID]**, Yi Chai** [†]**, Xinyue Zhou and Suhui Chen** *

Shanghai Key Laboratory of Plant Molecular Sciences, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China
* Correspondence: csh@shnu.edu.cn
† These authors contributed equally to this work.

Abstract: Mining tissue-specific genes is important for studying the processes of life activities within tissues, and it is a way of finding genes that regulate relevant traits. In recent years, the massive growth of expression data from various tissues has provided important opportunities for the large-scale analysis of tissue-specific genes. We found 489, 276, and 728 RTEGs (root tissue-specific expression genes) using 35 RNA-seq databases in 13 different tissues from three species of plants, e.g., Arabidopsis, rice, and maize, respectively, by bioinformatics methods. A total of 34 RTEGs in rice were found to be conserved in all three species, and 29 genes of them were unreported. Furthermore, 16 newly core cis-acting elements, named REM1-16 (root expression motif), were predicted by four well-known bioinformatics tools, which might determine the root tissue expression pattern. In particular, REM2 is conserved in not only Arabidopsis, but also rice. These cis-acting elements may be an important genetic resource that can be introduced into synthetic memory circuits to precisely regulate the spatiotemporal expression of genes in a user-defined manner.

Keywords: root-specific expression genes; cis-regulatory elements; plants

1. Introduction

Roots are vital to plant growth and development as they not only anchor the plant and sense its surroundings, but also anchor and absorb water and nutrients [1]. Two important symbioses in which roots are also involved are the mycorrhizal symbiosis and the root/bacterial mutualism, i.e., the fixation of atmospheric nitrogen. Many tissue-specific genes play key biological roles in organisms and are associated with a wide range of important agronomic traits in crops. However, the systematic identification of conserved RTEGs across species has been largely overlooked. Bridging this research gap holds the potential not only to elucidate core regulatory mechanisms governing root development, but also to identify evolutionarily conserved genetic targets that could be applied to the precise optimization of crop root architecture [2].

To improve the agronomic traits of crops, various types of promoters have been used to induce expression, including constitutive, spatiotemporal, and stress/chemical inducible promoters. Although many studies have successfully used constitutive promoters, they are accompanied by some side effects because the functional properties of the genes themselves are not taken into consideration. For example, the overexpression lines of *OsMADS26* and *OsbHLH142* driven by the *Ubi* promoter have triggered several abnormal phenotypes, including retarded root/shoot growth and sterility [3,4]. These unintended phenotypes

underscore a critical need for precision control tools, specifically tissue-specific promoters, which can restrict gene expression in target tissues while maintaining endogenous physiological balance [5].

Emerging opportunities lie in transcriptomic data. RNA-seq has now revolutionized the way the research community studies gene expression. Indeed, this technology offers the possibility of quantifying the expression levels of all genes at once. Significant amounts of RNA-seq data have been released in the past decade. The real challenge shifts from data generation to knowledge extraction. Nevertheless, few studies have systematically mined these resources to identify conserved RTEGs across species [6]. A DNA sequence motif, a subsequence of DNA sequence, is a short (about 6–12 bp), repeated, and conserved pattern of nucleotides, and it has many biological functions such as being a DNA binding site for a regulatory protein and controlling gene expression. There are hundreds of algorithms for motif extraction in the promoters of genes of interest, and these tools can be classified into three classes, such as enumeration, deterministic optimization, and probabilistic optimization [7,8].

We performed a comparative analysis of tissue-specific gene expression patterns through the systematic integration of 35 RNA-seq datasets across three model plant species (*Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays*). Our investigation identified 489, 276, and 728 root-specific expressed genes in Arabidopsis, rice, and maize, respectively, with 34 orthologous RTEGs conserved across all three species, suggesting the evolutionary conservation of these molecular pathways in plant root development. To decipher the transcriptional regulation of RTEGs, we employed four established de novo motif discovery algorithms for a comprehensive analysis of their promoter sequences. This integrative approach identified 16 core cis-regulatory elements, which were named REM1-16. Next, *cis*-regulatory element sequences could be used to fabricate synthetic promoters by either random ligation or rationally designed building block approaches. Synthetic promoters can precisely control the expression of target genes at a specific temporal and spatial distribution [9,10].

2. Results

2.1. Large-Scale Investigation of Root Tissue-Specific Expressed Genes in Plants

We selected three representative plant models, including one dicotyledonous plant (*Arabidopsis thaliana*) and two monocots (*Oryza sativa*, *Zea mays*), with the aim of identifying RTEGs or even conserved RTEGs to elucidate the presence of conserved pathways controlling root development in plants. To accurately map the expression profiles, we browsed all RNA-seq data in the NCBI database and finally selected 13 tissues covering almost all organs of plants from the vegetative stages to reproductive stages in three species, e.g., root, shoot, leaf, tillering bud, stigma, ovary, pollen, husk, pericarp, shoot apical meristem (SAM), inflorescence meristem (IM), embryo, and endosperm, and we were guided by the principle that the same tissue RNAseq database, in different species, should be chosen to have as similar a developmental period as possible (Supplementary Table S2).

A total of 35 high-quality RNA-seq data from 29 published papers [11–39] were downloaded, and we conducted an RNA-seq analysis by species (Supplementary Table S1). We performed a Principal Component Analysis (PCA) of RNAseq-based expression data from three plants (Supplementary Figure S1). The results show that all tissues were independent and there was good reproducibility of duplicate biological samples. Furthermore, we isolated RTEGs by improving the tissue specificity index algorithm. The general scheme is shown below (Figure 1).
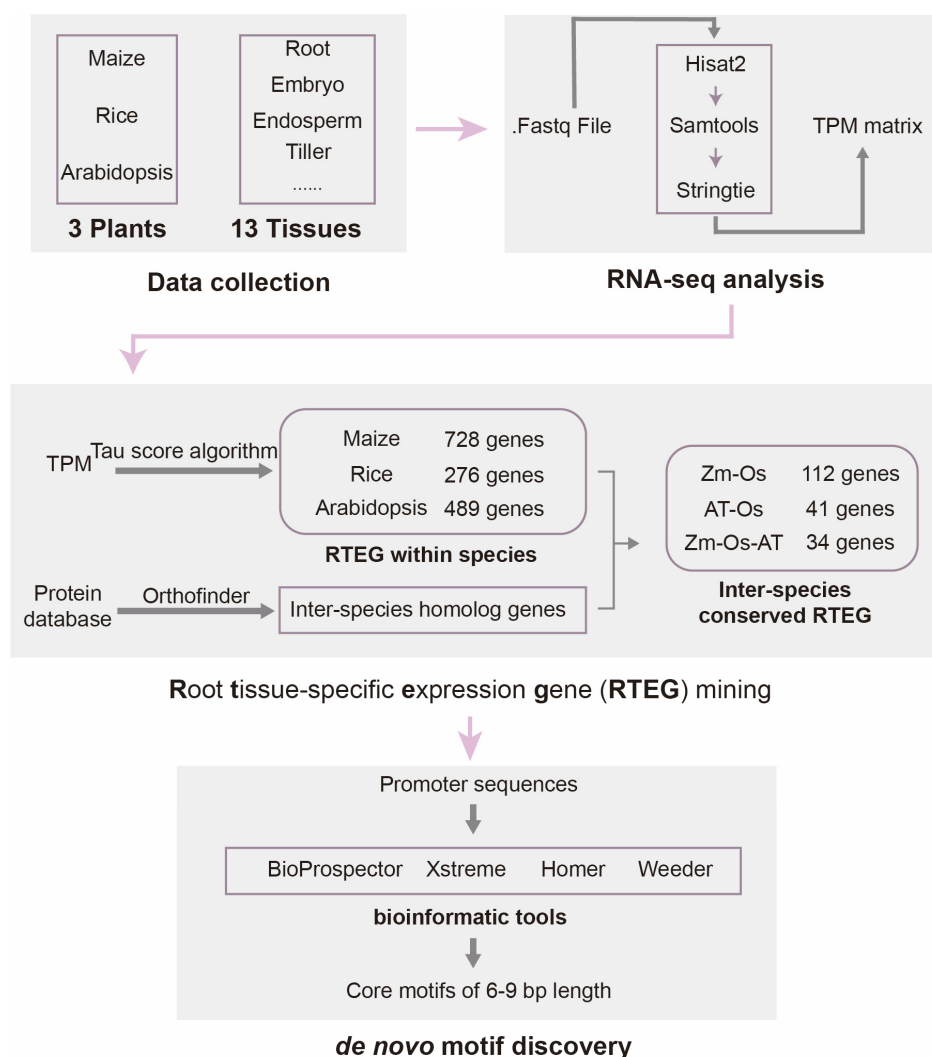
Figure 1. The general scheme for identifying root tissue-specific expression genes and de novo motif discovery.

Bioinformatics analyses included data collection, RNA-seq analysis, root tissue-specific expression gene mining, and de novo motif discovery. Technical details of the work are given in Section 4.

2.2. Functional Classification and GO Analysis of RTEGs

In total, 489, 276, and 728 RTEGs were identified from Arabidopsis, rice, and maize, respectively, and were used to perform a gene ontology (GO) analysis (Supplementary Figures S2–S4, Supplementary Table S3). The results show that RTEGs in Arabidopsis are enriched in organic cyclic compound binding, heterocyclic compound binding, catalytic activity, transition metal ion binding, transferase activity, oxidoreductase activity, metal ion binding, hydrolase activity, and cation binding. However, RTEGs in rice are not enriched in organic cyclic compound binding, heterocyclic compound binding, nucleic acid binding, and calcium ion binding. RTEGs in rice are functionally similar to those in maize and are mainly participate in transition metal ion binding, tetrapyrrole binding, substrate-specific transmembrane transporter, oxidoreductase activity, ion binding, heme binding. In short, although RTEG functions are generally similar in different species, there are differences that may arise from the soil in which they live or the specific needs for certain nutrients.

2.3. Identification of Conserved RTEGs

Across species, roots share common functions, such as anchoring the plant, sensing its surroundings, and fixing and absorbing water and nutrients. We inferred all orthogroups in Arabidopsis, rice, and maize using OrthoFinder [40] and isolated conserved RTEGs across species to identify unreported genes that may be involved in the above processes. In the comparison of rice and maize RTEGs, there were 112 conserved RTEGs in rice and 119 conserved RTEGs in maize. In the comparison of rice and Arabidopsis RTEGs, there were 22 conserved RTEGs in Arabidopsis and 41 conserved RTEGs in rice. Surprisingly, we identified 34 conserved RTEGs in rice, whose homologs in Arabidopsis and maize also exhibit root-specific expression patterns (Figure 2A,B). This evolutionary conservation strongly suggests that these genes may function in conserved regulatory pathways in roots. Notably, 29 of the 34 RTEGs have not been previously reported, highlighting their potential as novel candidates for future functional studies. Moreover, we collected 26 different tissues at various development stages to perform RT-PCR to validate the accuracy of the bioinformatic analyses (Figure 2C). The results show that *OsTIP2* was expressed in roots only at an early stage, whereas *OsPT2* was expressed in roots in different stages of plant life.

2.4. Motif Discovery

Based on the RTEGs obtained from rice, maize, and Arabidopsis, we extracted a total of 1493 gene promoters from these three species and attempted to uncover the key motifs that regulate the spatial expression of the genes, particularly those in the root tissues, using bioinformatics tools. We selected four representative software packages for de novo cis-acting element analysis, namely BioProspect [41], Weeder [42], Homer [43], and Xstreme [44], and inferred a total of 523 root-specific cis-acting elements. Of these, 40 cis-acting elements each for RTEGs in Arabidopsis, maize, and rice were obtained using a BioProspector analysis. The numbers of cis-acting elements of RTEGs in Arabidopsis, maize, and rice obtained by Xstreme analysis were 36, 28, and 27, respectively. The numbers of cis-acting elements of RTEGs obtained by Homer analysis were 79, 89, and 81, respectively. The numbers of cis-acting elements of RTEGs obtained by Weeder analysis were 24, 21, and 18, respectively. Among them, the BioProspector, Xstreme, and Homer analyses were carried out based on a comparison between RTEGs and embryo-specific expression genes or endosperm-specific expression genes as the background, whereas the Weeder analysis directly mined cis-acting elements for RTEGs.

Since the lengths of sequences obtained from different analysis software vary, we performed sequence continuity truncation of cis-acting elements obtained from software analysis, with 6–9 bp as the main length, in an attempt to find sequences that appeared in more than one software. Through Weeder, XSTREME, Homer, and BioProspector, we identified a total of 91 cis-acting elements, named REM1-91. Sixteen of these key sequences are shown in Table 1, of which Arabidopsis has seven key motifs, named REM1-7; maize has six key motifs, REM8-13; and rice has four root-specific expression cis-acting elements, REM2 and REM14-16. One key motif, REM2, exists in both Arabidopsis and rice. The results show that most of the key sequences exhibited species specificity, and only very few sequences were conserved among species, implying that gene expression patterns in plants may be determined by different sequences and mechanisms. In addition, we analyzed 16 key motifs, namely REM1-REM16, in the three species to compare their frequency differences in embryo/endosperm and root-specific expressed gene promoters. This was calculated as follows: Frequency of occurrence = number of tissue-specific expressed genes containing the key motif in their promoter sequence/total number of tissue-specific expressed genes × 100%. Through the analysis, we found that the key motifs were distributed in the promoters of embryo/endosperm-specific expressed genes in addition to the RTEGs,

but the frequency of occurrence in the RTEGs was significantly higher than that in the other tissues (Figure 3A–C). It is implied that REM1-16 may be the key motifs determining the root expression of the genes.
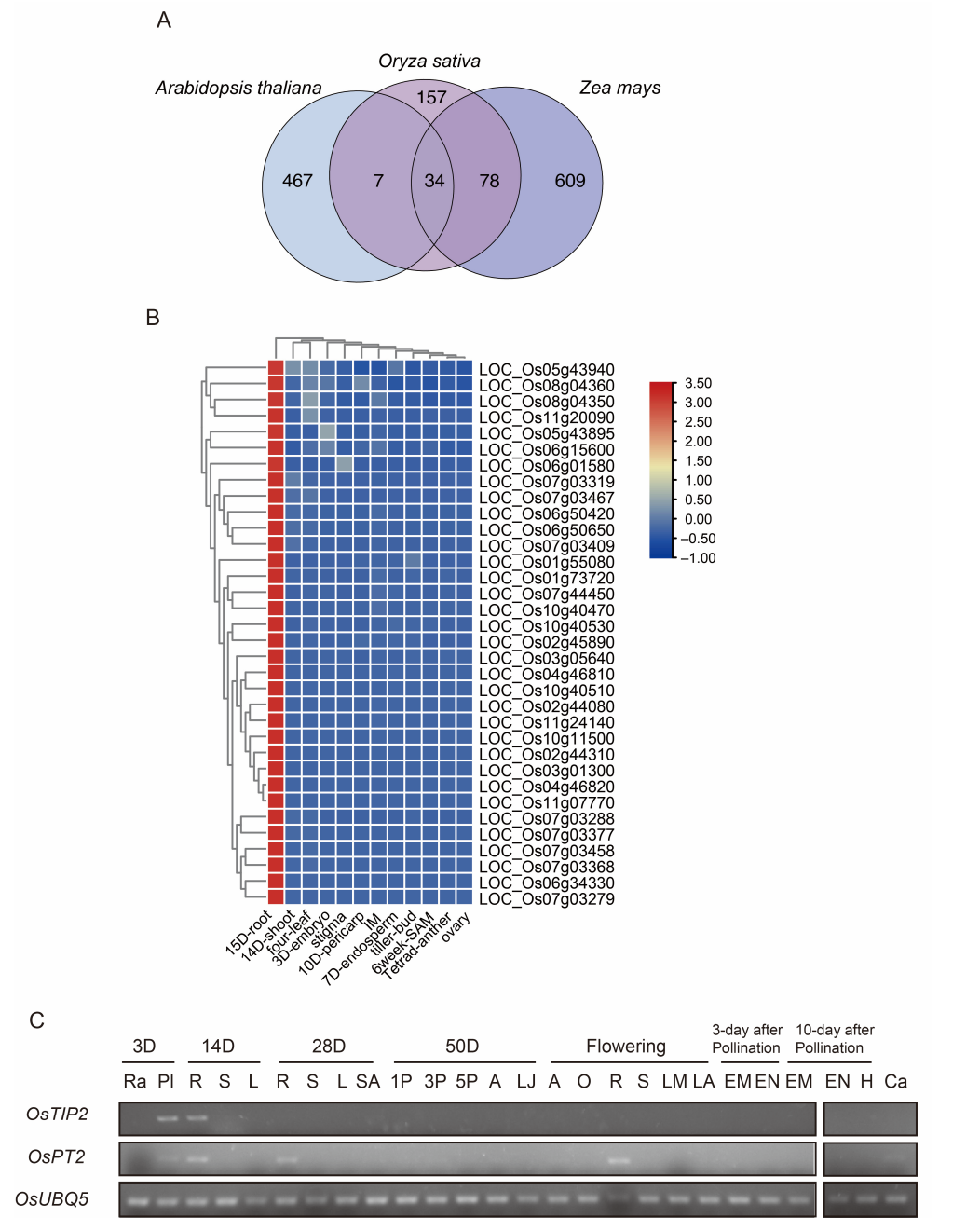

Figure 2. Identification of root tissue-specific expression genes. (**A**) Statistics of number of genes specifically expressed in root tissues of different species and number of conserved homologous genes. (**B**) Expression profiles of conserved root tissue-specific expression genes in rice. Thirteen different tissues from *Oryza sativa* used for profiling are mentioned on bottom of each column, and color bar on right side represents relative expression values. (**C**) Experimental validation of root tissue-specific expression of genes. *OsUBQ5* gene was used as internal control. Twenty-six tissues from different developmental stages were used. Ra: radicle, Pl: plumule, R: root, S: shoot, L: leaf, SA: shoot apical meristem, 1P: < 1 cm inflorescence meristem, 3P: 1–3 cm inflorescence meristem, 5P: 3–5 cm inflorescence meristem, A: anther, LJ: lamina joint, O: ovary, LM and LA: flag leaf at flowering stage (at 10 a.m. and 3 p.m.), EM: embryo, EN: endosperm, H: husk, Ca: callus. Different developmental stages are indicated above black line.

Table 1. Sixteen core motifs within the promoter sequences of root tissue-specific expression genes.

Motif Name	Motif Sequences	Motif Length	Species	Xstreme	Homer	Bioprospector	Weeder
REM1	AAAATTTGA	9	Arabidopsis	**	**	NA	NA
REM2	AATTAATC	8	Arabidopsis	NA	*	NA	NA
			Rice	NA	**	*	NA
REM3	AATTACG	7	Arabidopsis	**	*	*	NA
REM4	ACATGCA	7	Arabidopsis	NA	**	**	Y
REM5	ACGTGC	6	Arabidopsis	NA	**	NA	NA
REM6	AGCACG	6	Arabidopsis	*	**	NA	Y
REM7	CGACCG	6	Arabidopsis	*	*	NA	NA
REM8	AATCGAT	7	Maize	NA	**	*	NA
REM9	AGCTAGC	7	Maize	*	*	**	NA
REM10	AGCTAGCT	8	Maize	*	*	**	NA
REM11	AGCTAGCTA	9	Maize	*	*	**	NA
REM12	ATCGATCG	8	Maize	*	*	**	NA
REM13	ATCGATC	7	Maize	*	*	**	NA
REM14	ATACATA	7	Rice	NA	NA	*	NA
REM15	ATACATAT	8	Rice	*	**	**	NA
REM16	ATTAATTC	8	Rice	NA	**	**	NA

Four well-known bioinformatics tools, including Xstreme, Homer, BioProspector, and Weeder, were used for de novo motif discovery. * indicates that the sequence occurs either once in root embryo or root endosperm, ** indicates that the sequence occurs in both root embryo and root endosperm, and NA indicates that it does not occur in either case. "Y" is the motif identified by Weeder.
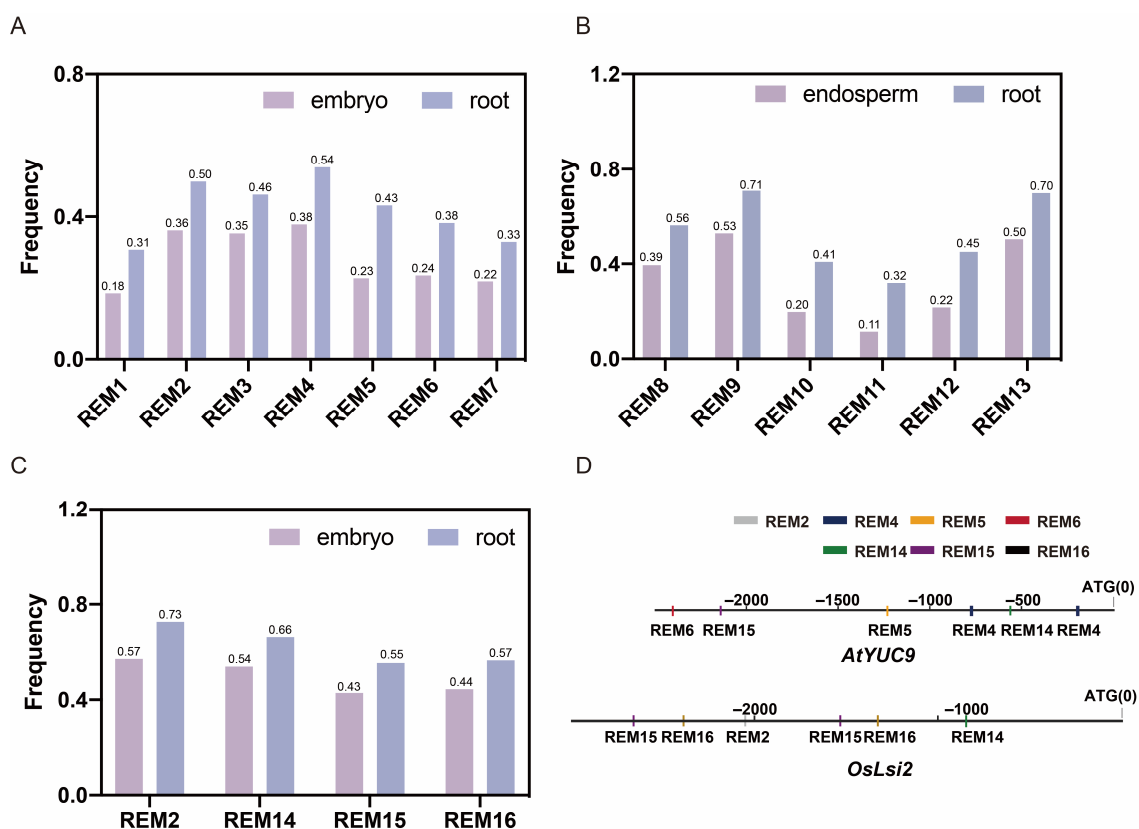

Figure 3. Frequency and positional distribution of 16 core motifs at promoters of tissue-specific expressed genes. Comparison of REM frequencies on promoters of different tissue-specific expressed genes in Arabidopsis (**A**), maize (**B**), and rice (**C**). (**D**) Distribution of REM positions in promoters of two root tissue-specific expression genes.

To further clarify the functions of REM1-16, we analyzed the promoters of some reported RTEGs. We found that two REM4 motifs, one REM5, 6, 14, and 15 were present in the promoter sequence of *AtYUC9* (AT1G0418), a reported Arabidopsis RTEG that promotes the synthesis of ethylene and auxin. On the other hand, *OsLsi2* (LOC_Os03g01700) is a rice RTEG encoding the silica transport protein. The predicted core motifs REM2 and REM14-16 all appeared in the promoter sequence of *OsLsi2*. The results indicate that REM1-16 may be the core cis-acting elements determining the spatial expression of the gene in the roots.

3. Discussion

In this study, 35 RNA-seq databases from 13 tissues in three species were used to identify the RTEGs on a large scale using the improved tissue specificity index (τ). We found 489, 276, and 728 root RTEGs in Arabidopsis, rice, and maize, respectively, of which 34 RTEGs in rice were found to be conserved in all three species. Furthermore, 16 newly reported core cis-acting elements, named REM1-16, were predicted by four well-known bioinformatics tools from the RTEGs of three plants, which may be useful in determining the root tissue expression patterns genes as their frequency of occurrence in the promoter sequence of RTEGs was much higher than that of other tissue-specific expressed genes. Importantly, REM2 is conserved not only in Arabidopsis, but also in rice. In future studies, the 16 core cis-acting elements deserve further experimental validation.

The conserved expression profiles of 34 evolutionarily conserved RTEGs suggest their potential involvement in fundamental regulatory pathways governing plant root development or signaling transduction. Notably, 29 of these 34 RTEGs have never been unreported before and are therefore priority targets for mechanistic studies on root-specific biological processes.

Although constitutive promoters have been widely adopted in plant biotechnology for strong and persistent transgene expression, their indiscriminate activity across all tissues and developmental stages often leads to unintended pleiotropic effects, such as impaired plant growth or metabolic burden. In contrast, gene regulatory systems employ cis-acting elements (e.g., conserved transcription factor-binding motifs) enable the precise spatiotemporal control of gene expression by integrating environmental or developmental signals. However, engineering such complexity in transgenic systems remains challenging. Synthetic gene circuits—modular assemblies of genetic components designed to process input signals and regulate output genes in a user-defined manner—offer a promising solution. By incorporating root-specific cis-acting motifs (identified in this study) as core regulatory elements, these circuits could dynamically fine-tune gene expression in a tissue-dependent manner, mimicking natural precision while avoiding the limitations of constitutive overexpression [45].

4. Materials and Methods

4.1. Plant Materials

Rice Zhonghua11 (ZH11, *Oryza sativa japonica*) plants were grown in Shanghai under standard paddy conditions.

To identify the expression pattern, 26 tissues from different developmental stages were collected, including the plumule and radicle (3 days after germination), root, shoot, and leaf (14 days, 28 days, and flowering stage), shoot apical meristem (SAM, 28-day-old plant), inflorescence meristem (less than 1 cm, 1–3 cm, and 3–5 cm), unmature pollen, lamina joint, anther, ovary, flag leaf at the flowering stage (at 10 a.m. and 3 p.m.), embryo (3 and 10 days after fertilization), husk (10 days after fertilization) and endosperm (3 and 10 days after fertilization), and callus.

4.2. Data Collection and RNA-Seq Analysis

A total of 35 pieces of high-quality RNA-seq data from 3 species of plants, including one dicotyledonous plant (*Arabidopsis thaliana*) and two monocots (*Oryza sativa*, *Zea mays*), were collected from 29 published papers [11–39]. In Supplementary Table S1, detailed information, e.g., species, tissue name, bioproject, sequence read archive (SRA), and publication, is listed.

RNA sequencing data were mapped to the reference genome (IRGSP-1.0 for rice, AGPv4 for maize, TAIR10 for Arabidopsis) using HISAT2, and StringTie was used to calculate the transcripts per million (TPM) values [46]. OrthoFinder was used to group genes into orthogroups [40].

TPM-matrix was further used as input for Principal Component Analysis (PCA), heatmap, and Gene Ontology (GO) enrichment analysis. FactoMineR, factoextra, and ggrepel (R package) were used to perform PCA and visualization [47]. We used TBtools to draw a heatmap [48]. In 15 November 2024, the agriGO website (http://systemsbiology. cau.edu.cn/agriGOv2/index.php) was used to perform GO enrichment analysis to identify enriched cellular components, molecular function, and biological process categories [49].

4.3. Tissue-Specific Expression Gene Analysis

The tissue specificity index (τ) [50] is defined as follows:

$$\tau = \frac{\sum_{i=1}^{N}(1 - \hat{x}_i)}{N - 1} \quad \hat{x}_i = log_{10}((x_i + 1)/log_{10}(x_{max} + 1)$$

N is the number of tissues, and $\hat{x}_i$ is calculated based on the expression value in different tissues i (x_i) and the maximal expression value (x_{max}). The tau score ranges from 0 to 1. To eliminate the effect of x_{max} and x_{sec} on the tissue specificity index (τ) under extreme conditions, tissue-specific expression gene screening is carried out based on the following criteria, where τ needs to be greater than 0.85 in all cases.

If the value of the second high expression (x_{sec}) is not equal to 0, $2 < x_{max} \leq 20$, and $x_{max}/x_{sec} \geq 4$;

If the value of the second high expression (x_{sec}) is not equal to 0, $20 < x_{max}$, and $x_{max}/x_{sec} \geq 10$;

If $x_{sec} = 0$, $x_{max} \geq 1$.

4.4. RNA Extract and RT-PCR

Total RNAs of different tissues mentioned above were extracted by Trizol reagent (Invitrogen, Carlsbad, USA) and used to synthesize cDNA through reverse transcription (Takara, Kusatsu, Japan). RT-PCR was conducted in a total volume of 20 μL containing 10 μL Taq mix, 6 μL cDNA (diluted 30 times), 1.5 μL primers (10 mM), and 2.5 μL double-distilled water. *UBQ5* gene was used as an internal control (primers: AC-CACTTCGACCGCCACTACT, ACGCCTAAGCCTGCTGGTT), and transcription levels of *OsPT2* and *OsTIP2* were examined using primers *OsPT2*-F, R (GGTATGCTCATGACGCT-GCT, TTGGGCGATCGCTTCTTGG), and *OsTIP2*-F, R (GGAAGCTTGGGTGACTCCTT, AAACGCCCACGAACAGGG). All examinations were conducted with three biological and technological replicates.

4.5. De Novo Motif Discovery

Promoter sequences of RTEGs were first extracted from maize, rice, and Arabidopsis. Among them, the 3 kb sequence of RTEGs before ATG was extracted from maize and rice, and the 2.5 kb sequence of RTEGs before ATG was extracted from Arabidopsis. Xstreme v5.4.1 [44], Weeder 2.0 [42], BioProspector [41], and Homer v4.11 [43] were used in this

study. Each bioinformatic tool was used to search for both strands with default parameters. The promoter sequence of either embryo tissue-specific expression gene or endosperm tissue-specific expression gene was used as background sequence.

Abbreviations

The following abbreviations are used in this manuscript:

RTEGs	Root tissue-specific expression genes
REM	Root expression motif

References

1. Hodge, A.; Berta, G.; Doussan, C.; Merchan, F.; Crespi, M. Plant root growth, architecture and function. *Plant Soil* **2009**, *321*, 153–187. [CrossRef]
2. Gho, Y.S.; Choi, H.; Moon, S.; Kim, S.R.; Ha, S.H.; Jung, K.H. Tissue-specific enhancement of OsRNS1 with root-preferred expression is required for the increase of crop yield. *J. Adv. Res.* **2022**, *42*, 69–81. [CrossRef] [PubMed]
3. Lee, S.; Woo, Y.M.; Ryu, S.I.; Shin, Y.D.; Kim, W.T.; Park, K.Y.; Lee, I.J.; An, G. Further characterization of a rice AGL12 group MADS-box gene. *Plant Physiol.* **2008**, *147*, 156–168. [CrossRef] [PubMed]
4. Ko, S.S.; Li, M.J.; Lin, Y.J.; Hsing, H.X.; Yang, T.T.; Chen, T.K.; Jhong, C.M.; Ku, M.S.B. Tightly Controlled Expression of bHLH142 is Essential for Timely Tapetal Programmed Cell Death and Pollen Development in Rice. *Front. Plant Sci.* **2017**, *8*, 1258. [CrossRef]
5. Li, Y.Y.; Liu, S.J.; Yu, Z.M.; Liu, Y.; Wu, P. Isolation and characterization of two novel root-specific promoters in rice (*Oryza sativa* L.). *Plant Sci.* **2013**, *207*, 37–44. [CrossRef]
6. Stark, R.; Grzelak, M.; Hadfield, J. RNA sequencing: The teenage years. *Nat. Rev. Genet.* **2019**, *20*, 631–656. [CrossRef]
7. Hashim, F.A.; Mabrouk, M.S.; Al-Atabany, W. Review of Different Sequence Motif Finding Algorithms. *Avicenna J. Med. Biotechnol.* **2019**, *11*, 130–148.
8. Liu, W.S.; Mazarei, M.; Peng, Y.H.; Fethe, M.H.; Rudis, M.R.; Lin, J.Y.; Millwood, R.J.; Arelli, P.R.; Stewart, C.N. Computational discovery of soybean promoter regulatory elements for the construction of soybean cyst nematode-inducible synthetic promoters. *Plant Biotechnol. J.* **2014**, *12*, 1015–1026. [CrossRef]
9. Mogno, I.; Vallania, F.; Mitra, R.D.; Cohen, B.A. TATA is a modular component of synthetic promoters. *Genome Res.* **2010**, *20*, 1391–1397. [CrossRef]
10. Aysha, J.; Noman, M.; Wang, F.W.; Liu, W.C.; Zhou, Y.G.; Li, H.Y.; Li, X.W. Synthetic Promoters: Designing the Regulatory Modules for Controlled Gene Expression. *Mol. Biotechnol.* **2018**, *60*, 608–620. [CrossRef]
11. Chettoor, A.; Givan, S.; Cole, R.; Coker, C.; Unger-Wallace, E.; Vejlupkova, Z.; Vollbrecht, E.; Fowler, J.; Evans, M. Discovery of novel transcripts and gametophytic functions via RNA-seq analysis of maize gametophytic transcriptomes. *Genome Biol.* **2014**, *15*, 414. [CrossRef] [PubMed]
12. Eveland, A.L.; Goldshmidt, A.; Pautler, M.; Morohashi, K.; Liseron-Monfils, C.; Lewis, M.W.; Kumari, S.; Hiraga, S.; Yang, F.; Unger-Wallace, E.; et al. Regulatory modules controlling maize inflorescence architecture. *Genome Res.* **2014**, *24*, 431–443. [CrossRef] [PubMed]
13. Liu, X.Y.; Zhou, S.L.; Wang, W.T.; Ye, Y.R.; Zhao, Y.; Xu, Q.T.; Zhou, C.; Tan, F.; Cheng, S.F.; Zhou, D.X. Regulation of Histone Methylation and Reprogramming of Gene Expression in the Rice Inflorescence Meristem. *Plant Cell* **2015**, *27*, 1428–1444. [CrossRef] [PubMed]
14. Lunardon, A.; Forestan, C.; Farinati, S.; Axtell, M.J.; Varotto, S. Genome-Wide Characterization of Maize Small RNA Loci and Their Regulation in the required to maintain repression6-1 (rmr6-1) Mutant and Long-Term Abiotic Stresses. *Plant Physiol.* **2016**, *170*, 1535–1548. [CrossRef]

15. Armenta-Medina, A.; Lepe-Soltero, D.; Xiang, D.Q.; Datla, R.; Abreu-Goodger, C.; Gillmor, C.S. miRNAs promote embryo pattern formation beginning in the zygote. *Dev. Biol.* **2017**, *431*, 145–151. [CrossRef]

16. Oka, R.; Zicola, J.; Weber, B.; Anderson, S.N.; Hodgman, C.; Gent, J.I.; Wesselink, J.J.; Springer, N.M.; Hoefsloot, H.C.J.; Turck, F.; et al. Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. *Genome Biol.* **2017**, *18*, 137. [CrossRef]

17. Yu, L.; Ma, T.F.; Zhang, Y.Q.; Hu, Y.; Yu, K.; Chen, Y.Y.; Ma, H.L.; Zhao, J. Identification and analysis of the stigma and embryo sac-preferential/specific genes in rice pistils. *BMC Plant Biol.* **2017**, *17*, 60. [CrossRef]

18. Pignatta, D.; Novitzky, K.; Satyaki, P.R.V.; Gehring, M. A variably imprinted epiallele impacts seed development. *PLoS Genet.* **2018**, *14*, e1007469. [CrossRef]

19. Schon, M.A.; Kellner, M.J.; Plotnikova, A.; Hofmann, F.; Nodine, M.D. NanoPARE: Parallel analysis of RNA 5′ ends from low-input RNA. *Genome Res.* **2018**, *28*, 1931–1942. [CrossRef]

20. Begcy, K.; Nosenko, T.; Zhou, L.Z.; Fragner, L.; Weckwerth, W.; Dresselhaus, T. Male Sterility in Maize after Transient Heat Stress during the Tetrad Stage of Pollen Development. *Plant Physiol.* **2019**, *181*, 683–700. [CrossRef]

21. Deforges, J.; Reis, R.S.; Jacquet, P.; Sheppard, S.; Gadekar, V.P.; Hart-Smith, G.; Tanzer, A.; Hofacker, I.L.; Iseli, C.; Xenarios, I.; et al. Control of Cognate Sense mRNA Translation by cis-Natural Antisense RNAs. *Plant Physiol.* **2019**, *180*, 305–322. [CrossRef] [PubMed]

22. Dong, Z.B.; Xiao, Y.G.; Govindarajulu, R.; Feil, R.; Siddoway, M.L.; Nielsen, T.; Lunn, J.E.; Hawkins, J.; Whipple, C.; Chuck, G. The regulatory landscape of a core maize domestication module controlling bud dormancy and growth repression. *Nat. Commun.* **2019**, *10*, 3810. [CrossRef] [PubMed]

23. Gómez-Ariza, J.; Brambilla, V.; Vicentini, G.; Landini, M.; Cerise, M.; Carrera, E.; Shrestha, R.; Chiozzotto, R.; Galbiati, F.; Caporali, E.; et al. A transcription factor coordinating internode elongation and photoperiodic signals in rice. *Nat. Plants* **2019**, *5*, 358–362. [CrossRef] [PubMed]

24. He, J.; Jiang, Z.M.; Gao, L.; You, C.J.; Ma, X.; Wang, X.F.; Xu, X.F.; Mo, B.X.; Chen, X.M.; Liu, L. Genome-Wide Transcript and Small RNA Profiling Reveals Transcriptomic Responses to Heat Stress. *Plant Physiol.* **2019**, *181*, 609–629. [CrossRef]

25. Ishimoto, K.; Sohonahra, S.; Kishi-Kaboshi, M.; Itoh, J.; Hibara, K.I.; Sato, Y.; Watanabe, T.; Abe, K.; Miyao, A.; Nosaka-Takahashi, M.; et al. Specification of basal region identity after asymmetric zygotic division requires mitogen-activated protein kinase 6 in rice. *Development* **2019**, *146*, dev176305. [CrossRef]

26. Knauer, S.; Javelle, M.; Li, L.; Li, X.R.; Ma, X.L.; Wimalanathan, K.; Kumari, S.; Johnston, R.; Leiboff, S.; Meeley, R.; et al. A high-resolution gene expression atlas links dedicated meristem genes to key architectural traits. *Genome Res.* **2019**, *29*, 1962–1973. [CrossRef]

27. Miao, C.B.; Wang, Z.; Zhang, L.; Yao, J.J.; Hua, K.; Liu, X.; Shi, H.Z.; Zhu, J.K. The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nat. Commun.* **2019**, *10*, 3822. [CrossRef]

28. Poulev, A.; Heckman, J.R.; Raskin, I.; Belanger, F.C. Tricin levels and expression of flavonoid biosynthetic genes in developing grains of purple and brown pericarp rice. *PeerJ* **2019**, *7*, e6477. [CrossRef]

29. Xiong, Y.F.; Ren, Y.; Li, W.; Wu, F.S.; Yang, W.J.; Huang, X.L.; Yao, J.L. NF-YC12 is a key multi-functional regulator of accumulation of seed storage substances in rice. *J. Exp. Bot.* **2019**, *70*, 3765–3780. [CrossRef]

30. Doll, N.M.; Just, J.; Brunaud, V.; Caïus, J.; Grimault, A.; Depège-Fargeix, N.; Esteban, E.; Pasha, A.; Provart, N.J.; Ingram, G.C.; et al. Transcriptomics at Maize Embryo/Endosperm Interfaces Identifies a Transcriptionally Distinct Endosperm Subdomain Adjacent to the Embryo Scutellum. *Plant Cell* **2020**, *32*, 833–852. [CrossRef]

31. Hanamata, S.; Sawada, J.; Ono, S.; Ogawa, K.; Fukunaga, T.; Nonomura, K.I.; Kimura, S.; Kurusu, T.; Kuchitsu, K. Impact of Autophagy on Gene Expression and Tapetal Programmed Cell Death During Pollen Development in Rice. *Front. Plant Sci.* **2020**, *11*, 172. [CrossRef] [PubMed]

32. Ji, Y.Y.; Huang, W.T.; Wu, B.W.; Fang, Z.M.; Wang, X.L. The amino acid transporter AAP1 mediates growth and grain yield by regulating neutral amino acid uptake and reallocation in *Oryza sativa*. *J. Exp. Bot.* **2020**, *71*, 4763–4777. [CrossRef] [PubMed]

33. Martínez-Fernández, I.; de Moura, S.M.; Alves-Ferreira, M.; Ferrándiz, C.; Balanzà, V. Identification of Players Controlling Meristem Arrest Downstream of the FRUITFULL-APETALA2 Pathway. *Plant Physiol.* **2020**, *184*, 945–959. [CrossRef]

34. Pan, S.Y.; Li, K.E.; Huang, W.; Zhong, H.; Wu, H.H.; Wang, Y.; Zhang, H.; Cai, Z.W.; Guo, H.W.; Chen, X.M.; et al. DXO1 possesses deNADding and exonuclease activities and its mutation affects defense-related and photosynthetic gene expression. *J. Integr. Plant Biol.* **2020**, *62*, 967–983. [CrossRef]

35. Singh, R.R.; Verstraeten, B.; Siddique, S.; Tegene, A.M.; Tenhaken, R.; Frei, M.; Haeck, A.; Demeestere, K.; Pokhare, S.; Gheysen, G.; et al. Ascorbate oxidation activates systemic defence against root-knot nematode *Meloidogyne graminicola* in rice. *J. Exp. Bot.* **2020**, *71*, 4271–4284. [CrossRef]

36. Tan, J.Y.; Wu, Y.; Guo, J.P.; Li, H.M.; Zhu, L.L.; Chen, R.Z.; He, G.C.; Du, B. A combined microRNA and transcriptome analyses illuminates the resistance response of rice against brown planthopper. *BMC Genom.* **2020**, *21*, 144. [CrossRef]

37. Jones, D.S.; John, A.; VanDerMolen, K.R.; Nimchuk, Z.L. CLAVATA Signaling Ensures Reproductive Development in Plants across Thermal Environments. *Curr. Biol.* **2021**, *31*, 220–227. [CrossRef]

38. Lv, B.S.; Wei, K.J.; Hu, K.Q.; Tian, T.; Zhang, F.; Yu, Z.P.; Zhang, D.J.; Su, Y.H.; Sang, Y.L.; Zhang, X.S.; et al. MPK14-mediated auxin signaling controls lateral root development via ERF13-regulated very-long-chain fatty acid biosynthesis. *Mol. Plant* **2021**, *14*, 285–297. [CrossRef]

39. Zhang, M.J.; Zhao, T.Y.; Ouyang, X.K.; Zhao, X.Y.; Dai, X.R.; Gao, X.Q. Pollen-specific gene enhances growth of pollen tubes in the transmitting tract in *Arabidopsis*. *J. Exp. Bot.* **2022**, *73*, 696–710. [CrossRef]

40. Emms, D.M.; Kelly, S. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **2015**, *16*, 157. [CrossRef]

41. Liu, X.; Brutlag, D.L.; Liu, J.S. BioProspector: Discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac. Symp. Biocomput.* **2001**, 127–138.

42. Pavesi, G.; Mereghetti, P.; Mauri, G.; Pesole, G. Weeder Web: Discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res.* **2004**, *32*, W199–W203. [CrossRef]

43. 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] [PubMed]

44. Grant, C.E.; Bailey, T.L. XSTREME: Comprehensive motif analysis of biological sequence datasets. *bioRxiv* **2021**. [CrossRef]

45. Lloyd, J.P.B.; Ly, F.; Gong, P.; Pflüger, J.; Swain, T.; Pflüger, C.; Fourie, E.; Khan, M.A.; Kidd, B.N.; Lister, R. Synthetic memory circuits for stable cell reprogramming in plants. *Nat. Biotechnol.* **2022**, *40*, 1862–1872. [CrossRef] [PubMed]

46. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650–1667. [CrossRef]

47. Le, S.; Josse, J.; Husson, F. FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* **2008**, *25*, 1–18. [CrossRef]

48. Chen, C.J.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.H.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

49. Tian, T.; Liu, Y.; Yan, H.Y.; You, Q.; Yi, X.; Du, Z.; Xu, W.Y.; Su, Z. agriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **2017**, *45*, W122–W129. [CrossRef]

50. Lüleci, H.B.; Yilmaz, A. Robust and rigorous identification of tissue-specific genes by statistically extending tau score. *BioData Min.* **2022**, *15*, 31. [CrossRef]