



Article

Genome-Wide Survey of the *RWP-RK* Gene Family in Cassava (*Manihot esculenta* Crantz) and Functional Analysis

Chenyu Lin ^{1,2,†}, Xin Guo ^{1,†}, Xiaohui Yu ¹, Shuxia Li ² , Wenbin Li ², Xiaoling Yu ², Feng An ³ ,
Pingjuan Zhao ^{2,*} and Mengbin Ruan ^{2,*}

- ¹ School of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, China; lcy1179456412@163.com (C.L.); guoxinyxh@hotmail.com (X.G.); xiaohuiyu@hainanu.edu.cn (X.Y.)
² Key Laboratory of Biology and Genetic Resources of Tropical Crops, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China; lishuxia@itbb.org.cn (S.L.); liwenbin@itbb.org.cn (W.L.); yuxiaoling@itbb.org.cn (X.Y.)
³ Hainan Danzhou Agro-Ecosystem National Observation and Research Station, Rubber Research Institute of Chinese Academy of Tropical Agricultural Sciences, Danzhou 571737, China; an-f@catas.cn
* Correspondence: zhaopingjuan@itbb.org.cn (P.Z.); ruanmengbin@itbb.org.cn (M.R.)
† These authors contributed equally to the work.

Abstract: The plant-specific RWP-RK transcription factor family plays a central role in the regulation of nitrogen response and gametophyte development. However, little information is available regarding the evolutionary relationships and characteristics of the RWP-RK family genes in cassava, an important tropical crop. Herein, 13 RWP-RK proteins identified in cassava were unevenly distributed across 9 of the 18 chromosomes (Chr), and these proteins were divided into two clusters based on their phylogenetic distance. The NLP subfamily contained seven cassava proteins including GAF, RWP-RK, and PB1 domains; the RKD subfamily contained six cassava proteins including the RWP-RK domain. Genes of the NLP subfamily had a longer sequence and more introns than the RKD subfamily. A large number of hormone- and stress-related cis-acting elements were found in the analysis of RWP-RK promoters. Real-time quantitative PCR revealed that all MeNLP1-7 and MeRKD1/3/5 genes responded to different abiotic stressors (water deficit, cold temperature, mannitol, polyethylene glycol, NaCl, and H₂O₂), hormonal treatments (abscisic acid and methyl jasmonate), and nitrogen starvation. MeNLP3/4/5/6/7 and MeRKD3/5, which can quickly and efficiently respond to different stresses, were found to be important candidate genes for further functional assays in cassava. The MeRKD5 and MeNLP6 proteins were localized to the cell nucleus in tobacco leaf. Five and one candidate proteins interacting with MeRKD5 and MeNLP6, respectively, were screened from the cassava nitrogen starvation library, including agamous-like mads-box protein AGL14, metallothionein 2, Zinc finger FYVE domain containing protein, glyceraldehyde-3-phosphate dehydrogenase, E3 Ubiquitin-protein ligase HUWE1, and PPR repeat family protein. These results provided a solid basis to understand abiotic stress responses and signal transduction mediated by RWP-RK genes in cassava.

Keywords: cassava; nodule inception-like proteins; RWP-RK domain proteins; abiotic stress; phylogenetic relationships; gene expression patterns; nitrogen starvation



Citation: Lin, C.; Guo, X.; Yu, X.; Li, S.; Li, W.; Yu, X.; An, F.; Zhao, P.; Ruan, M. Genome-Wide Survey of the *RWP-RK* Gene Family in Cassava (*Manihot esculenta* Crantz) and Functional Analysis. *Int. J. Mol. Sci.* **2023**, *24*, 12925. <https://doi.org/10.3390/ijms241612925>

Academic Editors: Alexander Betekhtin and Ewa Kurczyńska

Received: 27 July 2023

Revised: 12 August 2023

Accepted: 16 August 2023

Published: 18 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cassava (*Manihot esculenta* Crantz) is an important tropical and sub-tropical crop that provides food for over 800 million people worldwide because of its high starch production [1]. Cassava has a strong tolerance to drought and low-fertility soil environments because of its effective use of light, heat, and water resources [2]. However, the mechanism by which this crop displays high productivity under poor soil and abiotic stress conditions remains elusive. The *RWP-RK* transcription factor plays a central role in the absorption of nitrogen, which is an essential macronutrient for plant growth and development [3,4],

but reports on the *RWP-RK* family in cassava are lacking. Thus, a systematic study of the *RWP-RK* gene family in cassava may help to elucidate cassava's response to infertility and abiotic stresses.

The most prominent feature of the *RWP-RK* protein is a 60-amino-acid sequence of the *RWP-RK* domain that is conserved in all vascular plants, green algae, and slime molds [4]. The minus dominance (*MID*) protein is the first identified *RWP-RK* that is responsible for switching on and off the minus-plus programs in gamete differentiation [5]. Interestingly, the *MID* gene is also induced after N removal and activates gametogenesis [6]. Subsequently, the first nodule inception (*NIN*) protein is identified in the legume species *Lotus japonicus* as a crucial regulator that is required for the formation of infection threads and the initiation of primordia, and the inception of legume nodules is also dependent on the perception of N limitation in the plant [7]. Phenotypic analysis of pea (*Pisum sativum*) *sym35* (homologous gene of *PsNin*) mutants and *L. japonicus nin* mutants suggest that the *PsSym35* and *LjNin* genes have a similar function in early stages of root nodule formation [8]. To date, *RWP-RK* genes are widespread and preliminarily studied in algae, legumes, and non-legume plants [9–11].

RWP-RK proteins from different species have been classified into two subfamilies of *NIN*-like proteins (*NLPs*) and *RWP-RK* domain proteins (*RKDs*) depending on sequence [4]. *NLPs* mainly regulate the tissue-specific expression of genes involved in nitrogen, whereas *RKDs* mainly regulate the expression of genes involved in gametogenesis/embryogenesis [4,12,13]. In addition, *NLPs* have three typical characteristic domains, namely, *RWP-RK*, Phox and Bem1 (*PB1*), and *GAF*-like domains [4,13,14]. The *PB1* domain is a ubiquitous protein–protein interaction domain, and it comprises another regulatory layer for *NLPs* via the protein interactions within *NLPs* or with other essential components [4,15]. *GAF*-like domains are detected at the N-terminal region of the *NLP* subfamily and might be involved in signal transduction [15,16]. Moreover, the *RKD* family is found in the genomes of green algae, vascular plants, and amoebozoia [4,17]. The *RKD* proteins contain a conserved *RWP-RK* domain and some other optional motifs. Furthermore, *RKD* proteins can be divided into three groups based on their optional motif differences [4]. In general, *RKD* proteins are remarkably shorter than *NLP* proteins from the same species because the *PB1* and *GAF*-like domains may be absent [4,18].

The function mechanism of *RWP-RK* proteins has been extensively studied in different species including green algae, legumes, and vascular plants. *NLP* directly targets *NF-Y* subunit genes to control N-mediated symbiotic root nodule formation in *L. japonicus* [19]. *Arabidopsis thaliana* is the model plant used to study *RWP-RK* proteins [13,18]. *AtNLP7* plays an important role in early response to nitrate [19,20]. In addition, overexpression of *AtNLP7* enhances nitrogen and carbon assimilation and promotes plant growth under both restricted and sufficient nitrogen conditions [10]. *AtNLP8* is important to promote seed germination [21]. The nitrate-CPK-*NLP* signaling pathway and Ca^{2+} - Ca^{2+} sensor protein kinase-*NLP* signaling cascades have been identified, and they allow *NLPs* to have central roles in mediating the nitrate signaling pathway [16,22]. Nitrate leads to the accumulation of *NLP1* in the nucleus where it can suppress the ability of *NIN* to activate *CRE1* expression in *Nicotiana benthamiana* and *Medicago truncatula* [23]. *RKDs* controls germ cell differentiation during female gametophyte development in *Arabidopsis*, *Triticum aestivum*, and *Nicotiana tabacum* [11,18,24,25]. In summary, previous studies have highlighted the emerging roles played by *NLPs* in the N starvation response, N and P interactions [26–28], nodule formation in legumes [29–32], and root cap release in higher plants [32].

The increasing number of plant genomes sequenced has facilitated the study of the *RWP-RK* family in dicots and monocots, including *Zea mays* [33], *Brassica napus* [34], and *Malus domestica* [35]; however, information on *RWP-RK* in cassava has not yet been reported. In this study, cassava *RWP-RK* transcription factors were identified from the whole genome, and a phylogenetic tree based on the genome-wide gene duplication events was constructed. Gene structures, protein motifs, and the expression levels of *RWP-RK*

genes in different cassava tissues and leaves of different stresses were systematically studied in detail. Subcellular localization and interaction of two RWP-RK proteins were investigated. The present results will provide important information for further study of the regulation mechanism of cassava RWP-RK genes under abiotic stresses.

2. Results

2.1. Cassava RWP-RK Protein Family

A total of 21 candidate proteins corresponding to the RWP-RK domain were originally obtained by searching query sequences. However, candidate proteins that contained partial RWP-RK domains were excluded. The cassava genome encoded 13 proteins carrying an RWP-RK domain, and these proteins were unevenly distributed across 9 of the 18 cassava Chr, with three members on Chr 2, two members on Chr 1 and 11, and one on each on Chr 3, 5, 6, 7, 14, and 18 (Figure 1). These proteins were divided into two classes: NLPs with similarity to NIN along their entire length and RKDs sharing only the RWP-RK domain with NIN. Seven proteins in the former class and six smaller proteins in the latter were named MeNLP1-7 and MeRKD1-6, respectively, based on their position on different Chr. Among the 13 RWP-RK proteins, MeRKD1 was the smallest protein with 120 amino acids, whereas MeNLP2 was the largest one with 1002 amino acids. Eighteen splice variants were predicted in MeNLP5, seven were predicted in MeNLP6, three were predicted in MeNLP4, and two were predicted in MeRKD3/5 and MeNLP7; the remaining six genes each produced a single transcript. MeNLP5/6 with many alternative splices can create more proteins with a slightly different function to regulate cassava development and growth [36].

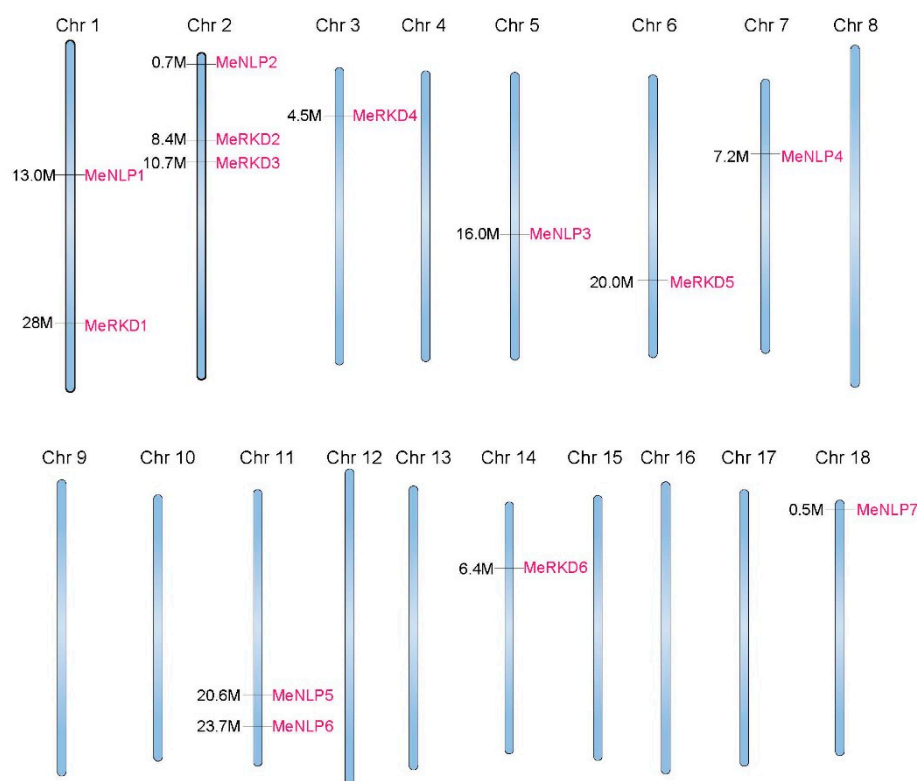


Figure 1. Location of 13 cassava RWP-RK proteins in chromosome.

2.2. Phylogenetic Analysis of the RWP-RK Family

We built a phylogenetic tree of RWP-RK proteins in *M. esculenta*, *O. sativa*, and *A. thaliana* to study the evolutionary relationships among the newly discovered RWP-RK proteins (Figure 2). Phylogenetic analysis indicated that RWP-RK proteins could be divided into four large groups. Among the four groups, group I contained the greatest number of RWP-RK proteins, and it was named the NLP subfamily because it included seven cas-

sava proteins (MeNLP1-7), nine *Arabidopsis* proteins (AtNLP1-9), and five rice proteins (OsNLP1-5). In addition, group II was named the RKD subfamily because they included six cassava proteins (MeRKD1-6), six *Arabidopsis* proteins (AtRKD1-6), and two rice proteins (OsRKD2/4). Meanwhile, only OsRKD1 and OsRKD3 existed separately in groups III and IV, respectively, thereby indicating that the RWP-RK family genes had species-specific expansion in rice.

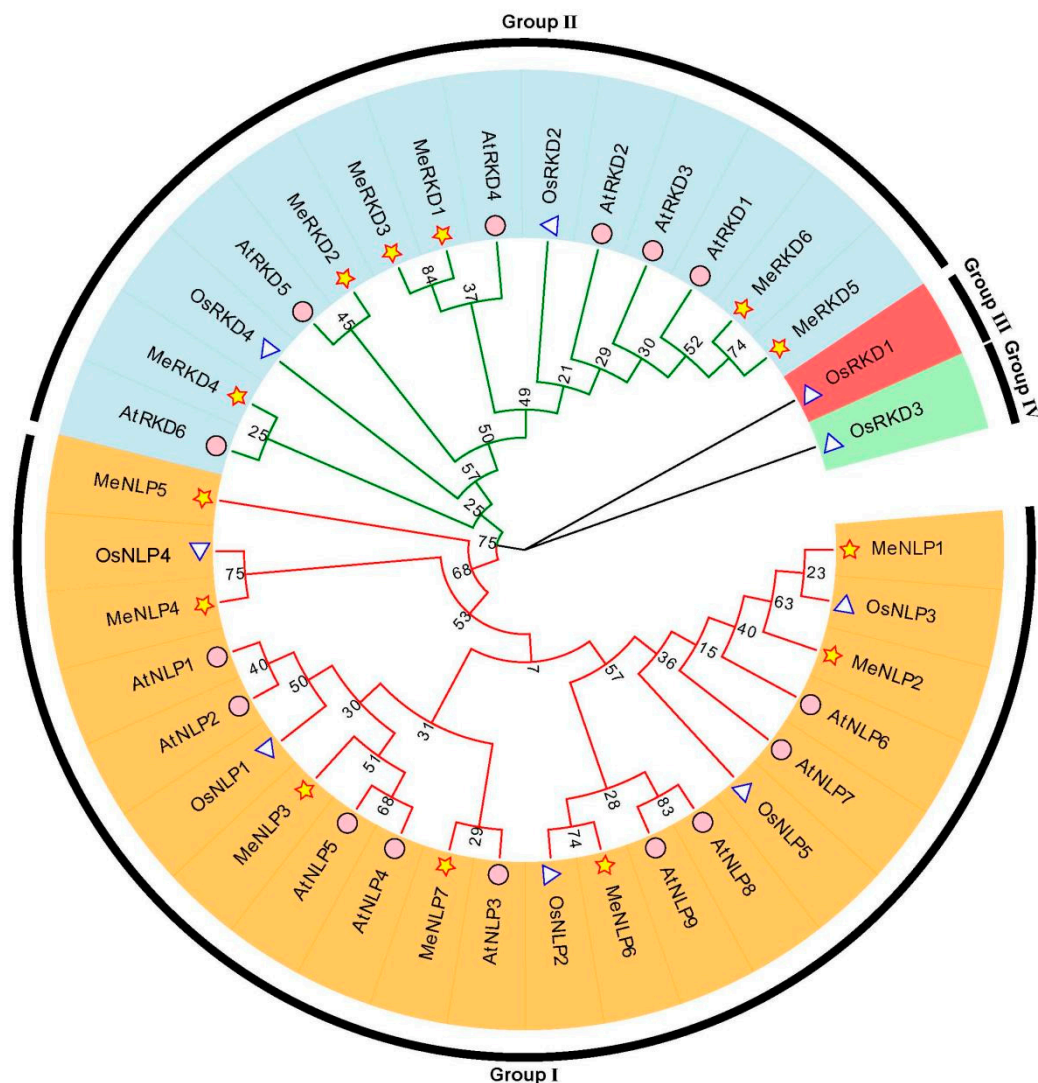


Figure 2. Phylogenetic evolution of RWP-RK proteins in *M. esculenta*, *A. thaliana* and *O. sativa*. The yellow pentagram represents cassava RWP-RK proteins, the white triangle represents rice RWP-RK proteins, and the purple circle represents *Arabidopsis* RWP-RK proteins. The numbers at nodes represent bootstrap values per 1000 replicate determined with the neighbor-joining (N-J) method.

2.3. Synteny Analyses of RWP-RK Genes

Tandem and segmental duplications are considered the primary driving forces that expand gene families during evolution [37]. Herein, six RWP-RK genes were clustered into five duplication event regions on cassava linkage groups 1, 2, 6, and 14 (MeNLP1/2, MeRKD1/3, MeRKD1/5, MeRKD1/6, and MeRKD5/6; Figure 3). These results indicated that some RWP-RK genes were possibly generated by gene duplication. A chromosomal region within 200 kb containing two or more homologous genes is defined as a tandem duplication event [38,39]. The findings support that cassava has a highly repetitive genome that contains gene pairs with a large number of segmental duplications. Segmental duplications are the main source of adding new members to the cassava RWP-RK gene family.

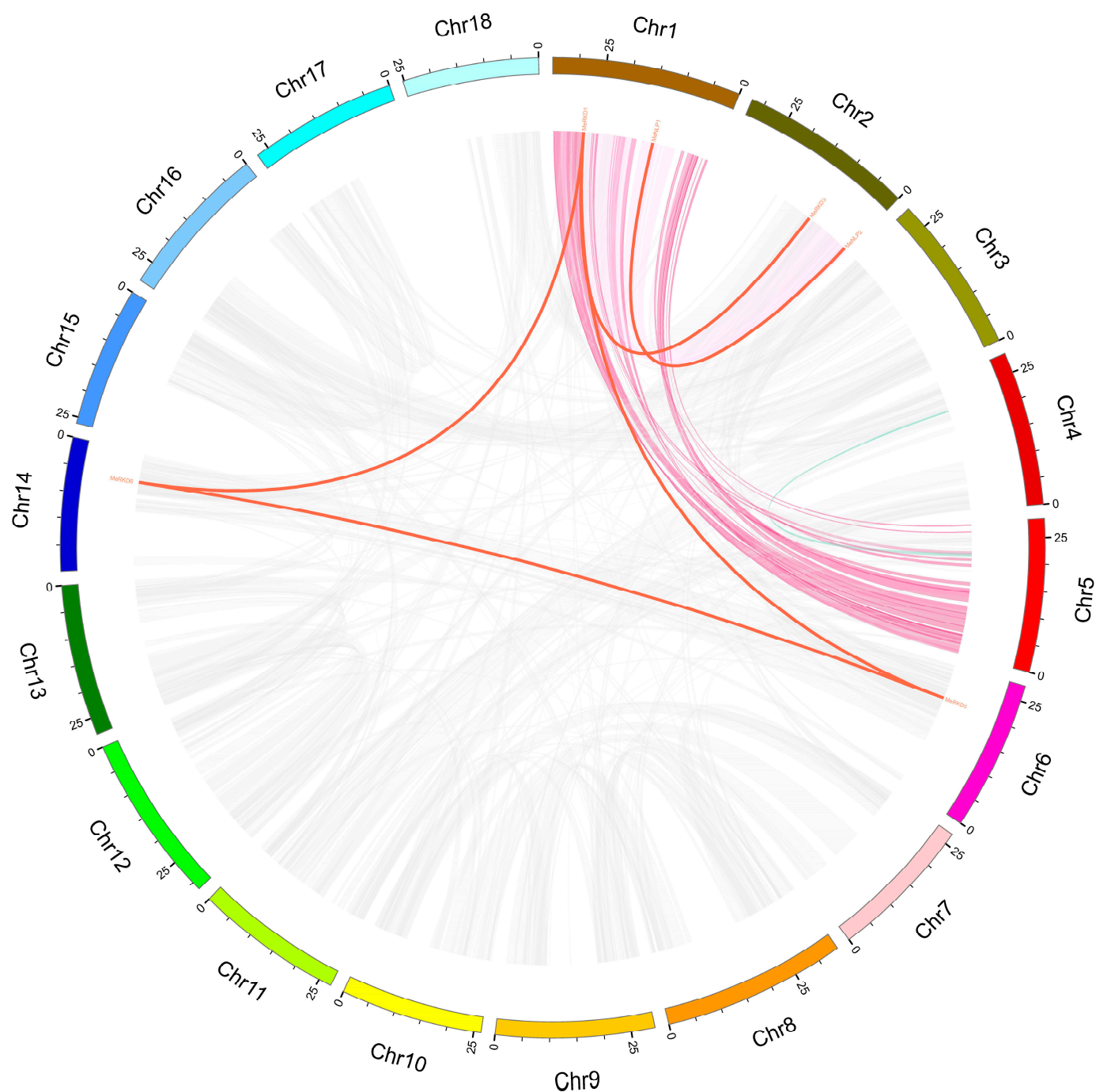


Figure 3. Schematic representations for the chromosomal distribution and inter chromosomal relationships of cassava *RWP-RK* genes. Gray lines indicate all syntenic blocks in the cassava genome, and the red lines indicate duplicated *RWP-RK* gene pairs. The chromosome number is indicated at the bottom of each chromosome.

A set of comparative syntenic graphs of cassava connected with dicots (*Arabidopsis*) and monocots (rice) was used to further infer the possible evolutionary implications of the cassava *RWP-RK* family. About 8 among the 13 identified *RWP-RK* genes showed no collinearity with either *Arabidopsis* or rice, suggesting that these homologous genes may have formed after the divergence of dicots and monocots or after the divergence of the two species. Nine and two cassava orthologous gene pairs were found in *Arabidopsis* and rice, respectively (Figure 4 and Supplementary Table S3). The two collinear genes of *MeRKD2* and *MeNLP7* were found between cassava and two other species (rice and

Arabidopsis, *MeRKD2/OsRKD4/AtRKD5*, *MeNLP7/OsNLP2/MeNLP4*, or *MeNLP5*), which indicate that these collinear genes may already have existed before the ancestors of monocot and dicot separation. More syntenic gene pairs were found between cassava and dicots than between cassava and monocots. For example, the collinear gene pairs of *MeNLP2/AtNLP6*, *MeRKD6/AtRKD1/2*, *MeRKD2/AtRKD5*, and *MeRKD5/AtRKD2/1* were not found between cassava and rice, which may indicate that these orthologous pairs formed after the divergence of dicotyledonous and monocotyledonous plants.

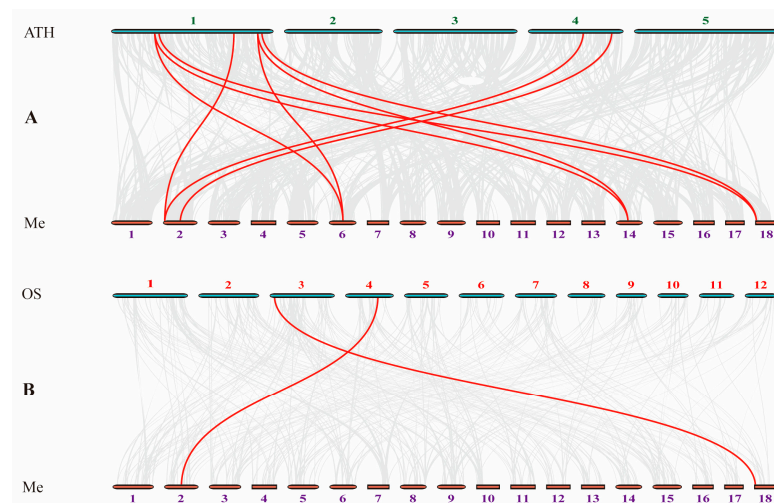


Figure 4. Synteny analysis of RWP-RK genes between cassava and two representative plant species. Gray lines in the background indicate the collinear blocks within cassava and other plant genomes, while the red lines highlight the syntenic cassava gene pairs with *A. thaliana* and *O. sativa*, respectively. The purple, red and green number represent cassava, rice and *Arabidopsis* chromosome, respectively.

2.4. RWP-RK Protein Structure and Conserved Motif Analysis

Thirteen proteins were divided into three categories by the MEME/MAST program to better understand the structural variability of RWP-RK proteins (Figure 5A). The amino acid sequences of 10 different motifs are shown in Figure 5B and Supplementary Table S4. Sequence alignment results showed that motifs #2/#8/#10 were different conserved RWP-RK domains that played a key role in the N-mediated control of development. Motifs #4/#7 were conserved PB1 domains, comprising approximately 80 amino acid residues. Sequence alignment showed that motifs #1/#3/#5/#7 were conserved GAF domains. Seven proteins that belonged to the first category contained 9 of all 10 motifs (except for *MeNLP5* which included 7 motifs). The second category only contained one protein of *MeRKD4* which only contained one RWP-RK motif. The last category included five proteins of *MeRKD1/2/3/5/6* which contained two RWP-RK motifs, except for *MeRKD5*, which included only one motif that was absent in *MeNLP1-7* and *MeRKD4* proteins. RWP-RK proteins in the same category shared similar motifs, thereby suggesting that similar RWP-RK members have conserved biological functions [34]. We aligned the coding sequences to analyze the exon/intron structures of RWP-RK genes (Figure 5C). Genes belonging to the first category had complete sequences with 4–6 exons and 3–5 introns, respectively, including UTR regions at both ends (except for *MeNLP5*). Gene sequences in the remaining category were incomplete with 2–5 exons and 1–4 introns, respectively, without a UTR region (except for *MeRKD1*). Therefore, the average number of exons was greater in the *NLPs* (average number of 4.7) than in the *RKDs* (average number of 3.3). The similar gene structures observed in the same RWP-RK subfamily were consistent with their phylogenetic relationships.

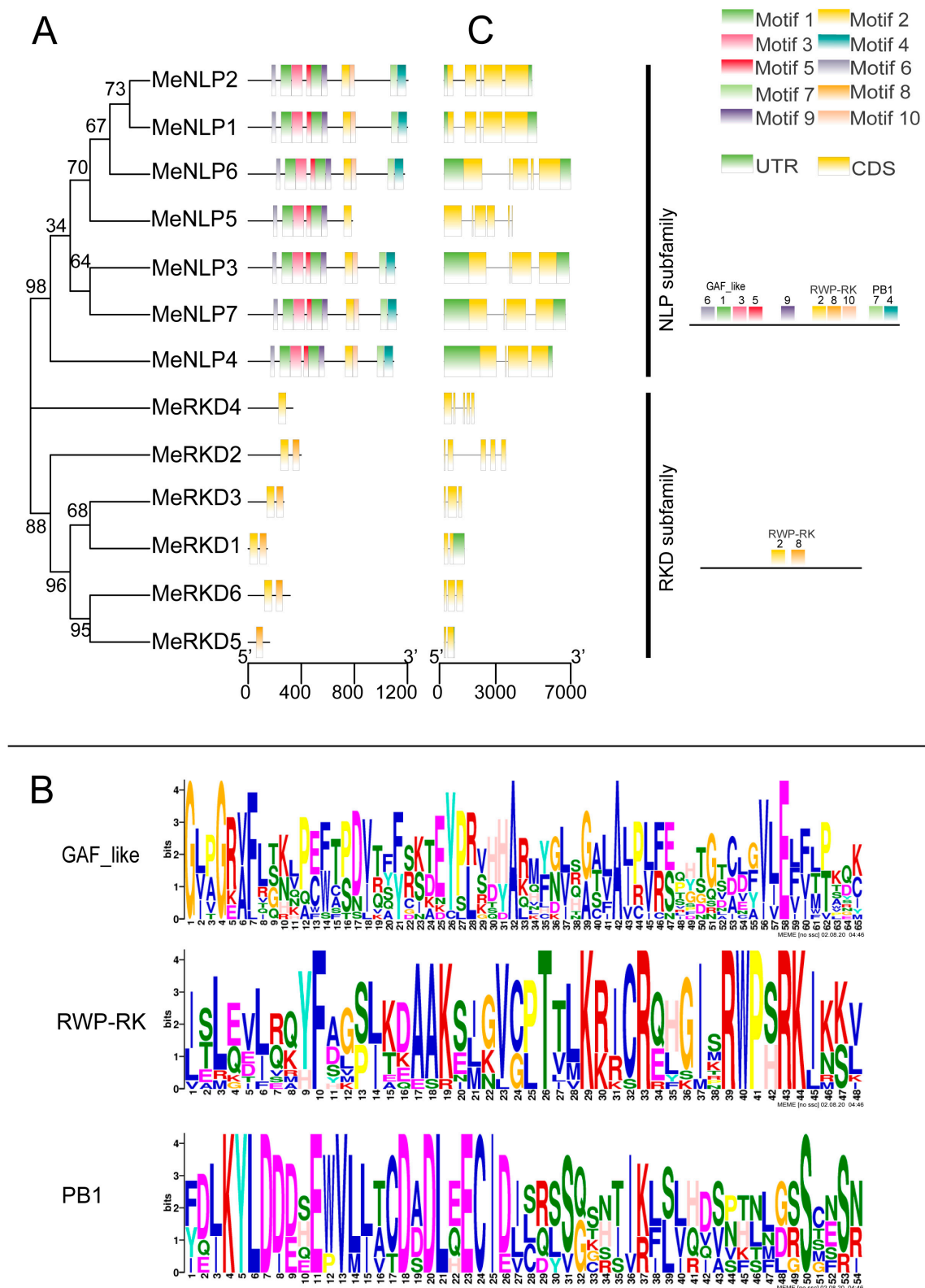


Figure 5. Motif and structure analysis of cassava RWP-RK family members. Note: (A), Protein motif analysis of cassava RWP-RK members. (B), Sequences of the three conserved motifs identified in this study. (C), Structure analysis of cassava RWP-RK genes. The green blocks represent the untranslated region (UTR), the yellow blocks represent exons, and the black lines represent introns.

2.5. Analysis of Cassava RWP-RK Gene Promoters

Cis-acting regulatory elements are important molecular switches involved in the transcriptional regulation of a dynamic network of gene activities controlling various biological processes, including hormone response, abiotic stress response, and development [40]. A large number of hormone-related cis-acting elements were identified through the analysis of cassava RWP-RK promoters (Figure 6), including ABA-responsive element (ABRE) in response to ABA [40,41], homeodomain-leucine zipper (HD-Zip) in response to auxin [42], and activation sequence-1 (as-1) in response to SA [43]. Furthermore, cis-acting elements related to adversity (e.g., low temperature, drought, and salt stress) were found, such as MYB, MYC, and MYC-like [44,45]. In addition, the TCT motif was required for the transcription of ribosomal protein gene promoters [46]. MYB and MYC recognition sites were present in almost all the genes analyzed (Figure 6), and MYB sites in a dehydration-responsive gene (RD22) promoter were proven to be cis-acting elements in the drought-inducible expression of RD22 in transgenic tobacco [47]. MYC sequence CATGTG played an important role in the dehydration-inducible expression of the *Arabidopsis* *EARLY RESPONSIVE TO DEHYDRATION STRESS* gene [45].

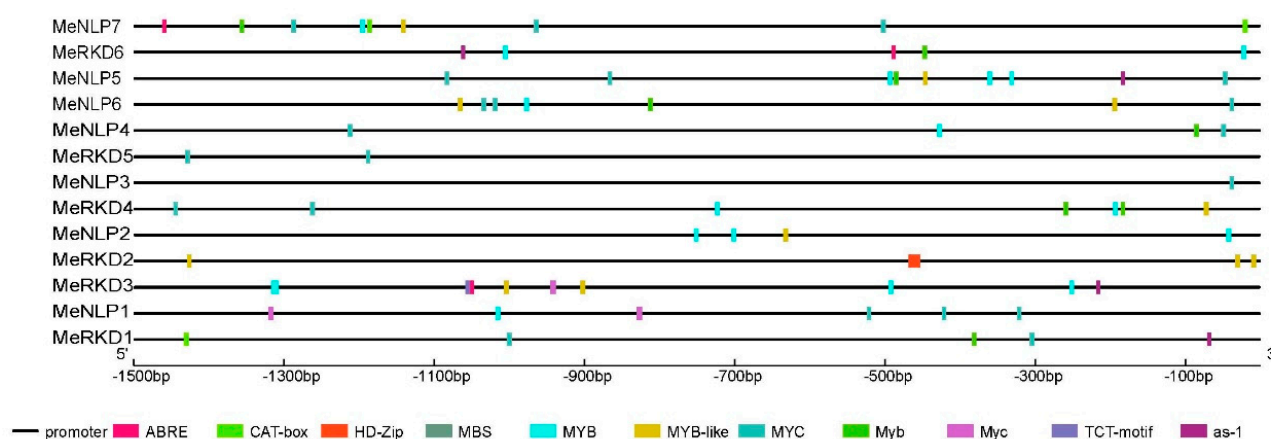


Figure 6. Prediction of cis-acting elements of RWP-RK gene promoter in cassava. Different cis-acting elements are represented in different colors. The same cis-acting elements of MYB, MYB-like, and Myb have different same because they derived from different plants, and MYC and Myc are also derived from different plants.

2.6. Transcription Profiles of RWP-RK Genes in Different Cassava Tissues

Specific primers were designed based on the conserved sequence to analyze the transcription activity of 13 RWP-RK genes in different tissues (Figure 7). *MeRKD2/6* genes were not expressed in all the tissues tested. The expression level of *MeRKD3/5* genes were lower in stems than in leaves, and the rest were close to leaves. Seeds and storage roots had more genes with low expressions levels than leaves, such as *MeRKD1/3/4/5* and *MeNLP2/3/6/7/* genes in seeds, and *MeRKD1/3* and *MeNLP1/2/3/5/7* genes in storage roots. Only *MeNLP1/5* genes were high in seeds, *MeRKD4/5* and *MeNLP6* genes were high in storage roots, and only *MeNLP4* in seed was close to the leaves. The expression levels of eight genes in flowers and fruits were similar; for example, *MeNLP3/6/7* and *MeRKD1/3/4* genes were low, *MeNLP1* and *MeRKD5* genes were high in two tissues, and only the *MeNLP2* gene was high in flower and low in fruit. Most genes with high expressions were found in roots, such as *MeNLP1/3/4/5/7* and *MeRKD1/5*. *MeNLP3/7* genes, which were an evolutionarily conserved gene pair, had a similar expression pattern in different tissues, but the expression levels of *MeNLP1/2* genes (an evolutionarily conserved gene pair) were completely different.

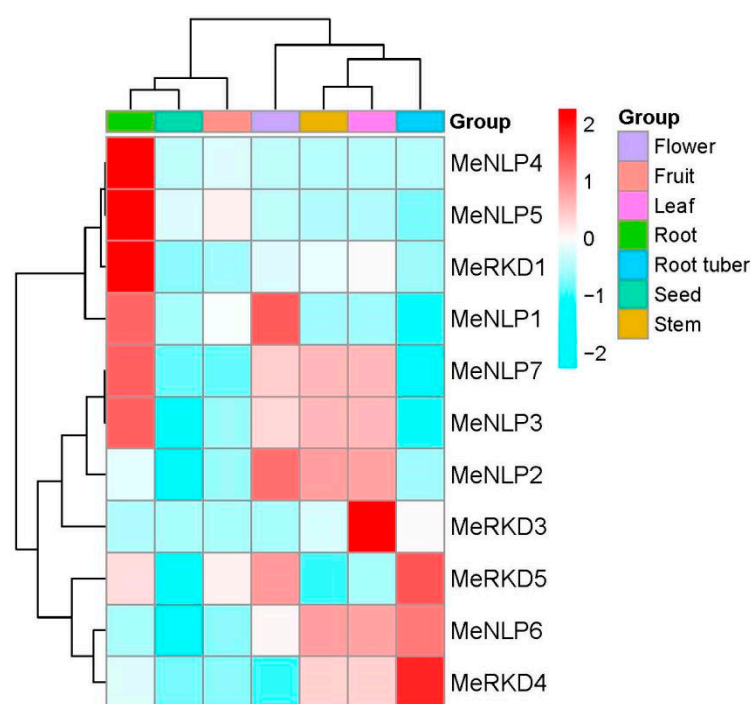
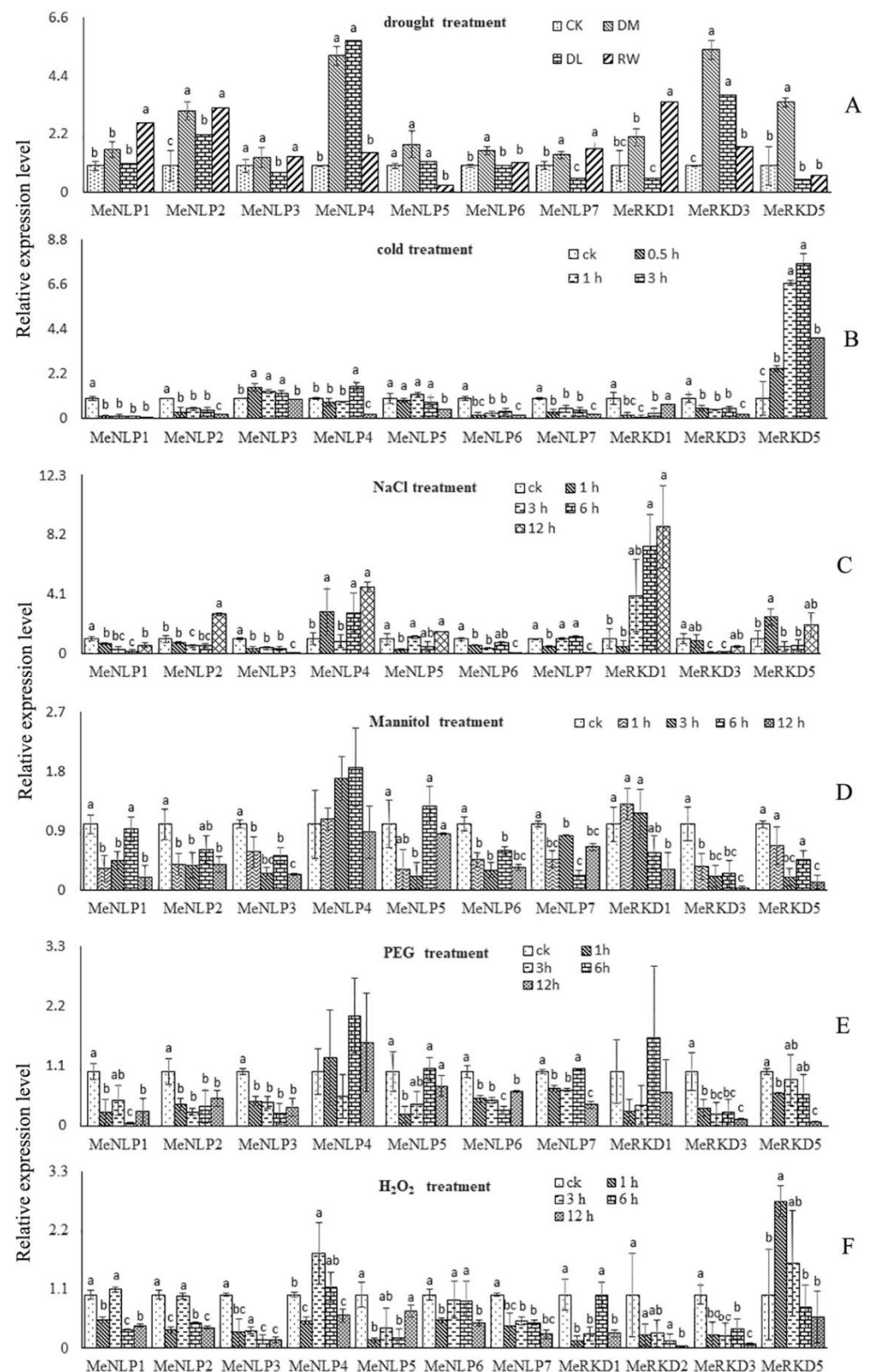


Figure 7. Tissue expression pattern of cassava *RWP-RK* genes in cassava. Log₂ transformed FPKM value was used to create the heat map. The scale represents the relative signal intensity of FPKM values.

2.7. Expression Patterns of *RWP-RK* Genes in Response to Different Stresses in Cassava

We hypothesized that *RWP-RK* genes may respond to stress and hormone treatments because their promoter regions contain hormone- and stress-related motifs. RNA was isolated from leaves treated with drought, cold, salt, osmotic, and H₂O₂ stress to identify the abiotic-responsive *RWP-RK* genes in the cassava genome. In addition, the relative expression patterns of 13 cassava *RWP-RK* genes were evaluated with qRT-PCR. Nevertheless, *MeRKD2/6* genes were not detected, and the *MeRKD4* gene showed no significant difference in all samples. The relative expression levels of 10 genes detected were significantly down- or up-regulated in one of the three treatment groups (DM, DL, and RW) compared with those in the control group. The gene expression levels of *MeNLP1/2/4/6* and *MeRKD1/3/5* were significantly higher than the control at some point during the drought period, the expression level of *MeNLP3/5* was significantly lower than that of the control, and *MeNLP7* and *MeRKD1* exhibited an increased and then depressed expression pattern (Figure 8A). *MeNLP3* and *MeRKD5* genes were up-regulated in one or all the processing points when cold treatment lasted from 0.5 to 12 h, whereas *MeNLP1/2/5/6/7* and *MeRKD3* were down-regulated in one or all the treatment groups. *MeNLP4* increased and then decreased, and *MeRKD1* showed the opposite trend. Of all the genes, *MeRKD5* gene expression increased the most significantly during the whole cold period (Figure 8B). Ten *RWP-RK* genes responded to NaCl treatment. *MeNLP1/3* genes were uniformly significantly down-regulated at all points when NaCl treatment lasted from 1 h to 12 h, *MeNLP4* and *MeRKD1/5* genes were significantly up-regulated at some point with the same treatment, and *MeNLP2/5/6/7* and *MeRKD3* genes exhibited a dynamic change (Figure 8C). Out of those ten genes, the responses of the nine *RWP-RK* genes to Man and PEG were highly similar. *MeNLP4/MeRKD1* and *MeNLP4* genes did not respond to PEG and Man during treatment, respectively, and the rest were uniformly significantly down-regulated in some time points when Man and PEG treatments lasted from 1 h to 12 h (Figure 8D,E).



We obtained the data of 11 *RWP-RK* genes for H_2O_2 treatment. *MeNLP3/7* and *MeRKD3* genes were uniformly significantly down-regulated in every time point, *MeNLP1/2/5/6* and *MeRKD1* genes were down-regulated at some point when H_2O_2 treatment lasted from 1 h to 12 h; only the *MeNLP4* gene exhibited dynamic changes, and the *MeRKD5* gene was significantly up-regulated at some point (Figure 8F).

2.8. Expression Patterns of *RWP-RK* Genes in Response to Different Hormone Treatments in Cassava

Herein, we obtained the data of 10 *RWP-RK* genes for ABA and MeJA treatments, but *MeRKD2/4/6* genes were not detected (Figure 9A,B). Our results showed that *MeNLP4* and *MeRKD1* genes were induced in some time points by ABA application, whereas *MeNLP1* and *MeRKD3/5* genes exhibited a depressed expression pattern at some time points. Meanwhile, *MeNLP2/3/5/6* genes demonstrated early up-regulation after 1 h, followed by reduced expression after 3 h. *MeNLP7* expression was lower after 1 and 3 h but promptly increased after 6–12 h (Figure 9A). Nine genes responded to MeJA treatment. *MeNLP1/3/6* genes were down-regulated at every point when MeJA treatment lasted for 1 h to 12 h. By contrast, *MeNLP4* and *MeRKD1* genes were up-regulated in every point, and *MeNLP2/7* genes were up-regulated only after 6 h and 3 h, respectively. *MeNLP5* was down-regulated after 1–3 h, and peaks occurred only after 6 h. *MeRKD3* was up-regulated after 3 h, and it continued to decline thereafter (Figure 9B).

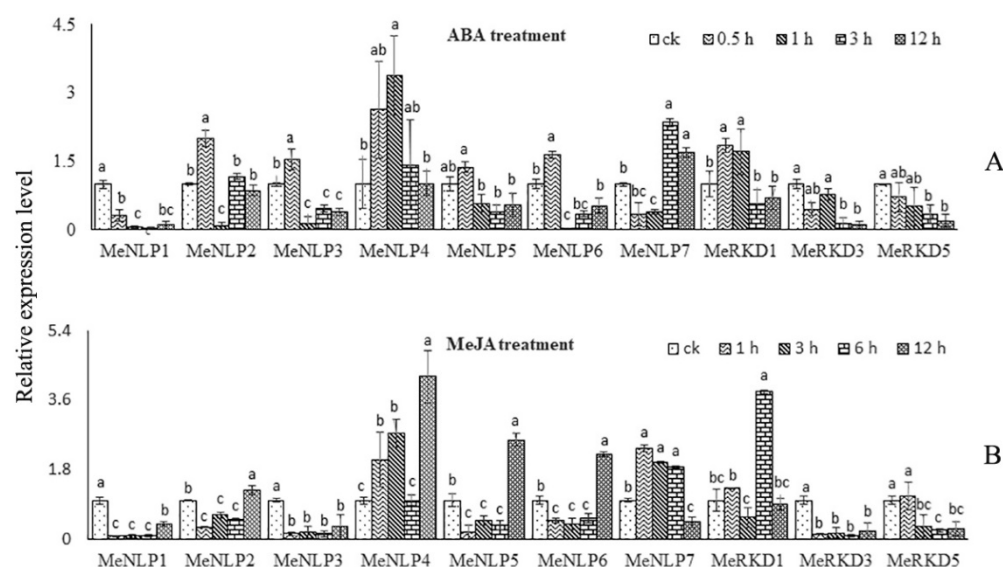


Figure 9. Expression analysis of *RWP-RK* genes in the leaves of cassava under different hormones. (A): ABA treatment. (B): JA treatment. The error bars represent the standard error of the means of three independent replicates. Values denoted by the same letter did not differ significantly at $p < 0.05$ according to one-way ANOVA.

2.9. Expression Patterns of *RWP-RK* Genes in Response to Nitrogen Starvation Treatments in Cassava

The response of *RWP-RK* family genes to nitrogen starvation was studied using the leaves and roots of three- and six-day transplanted seedlings in nitrogen starvation culture as samples. The levels of *MeNLP2/3/4/5/7* and *MRKD1* in the root were consistently up-regulated at all treatment samples. *MeNLP6* and *MeRKD3/5* were also up-regulated in three- and six-day samples, respectively; only *MeNLP1* expression was inhibited in three-day samples (Figure 10A,B). Meanwhile, the gene expression levels of *MeNLP1/2/3/5/6/7* and *MeRKD1/3* in leaves were up-regulated in six-day samples (Figure 10C). Notably, *MeNLP3/5/7* displayed prominent changes in expression levels in the root. Remarkably, *MeRKD3* and *MeRKD5* showed the most significant improvement in the root and leaf, respectively.

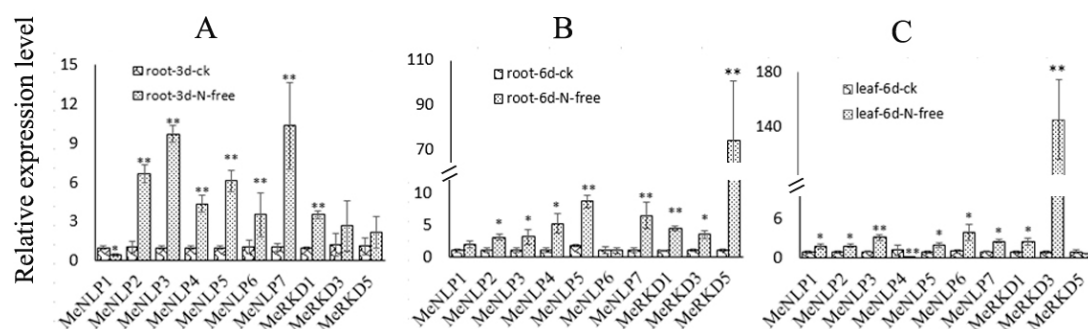


Figure 10. Expression profiles of different *RWP-RK* genes in the leaves and root of cassava under nitrogen starvation. (A): The third day of treatment in root. (B): The sixth day of treatment in root. (C): The sixth day of treatment in leaf. The error bars represent the standard error of the means of three independent replicates. * and ** represent different at $p < 0.05$ and $p < 0.01$ according to Duncan's multiple range tests.

2.10. Subcellular Localization of MeNLP6 and MeRKD5

Tobacco leaf epidermal cells expressing GFP showed cytoplasmic, plasma membrane, and nuclear localization (Figure 11A). However, the GFP signals in MeRKD5 and MeNLP6 fusion-expressing cells could be clearly detected in the nucleus with nuclear localization signal (Figure 11B,C). These observations confirmed that both MeRKD5 and MeNLP6 were localized to the nucleus.

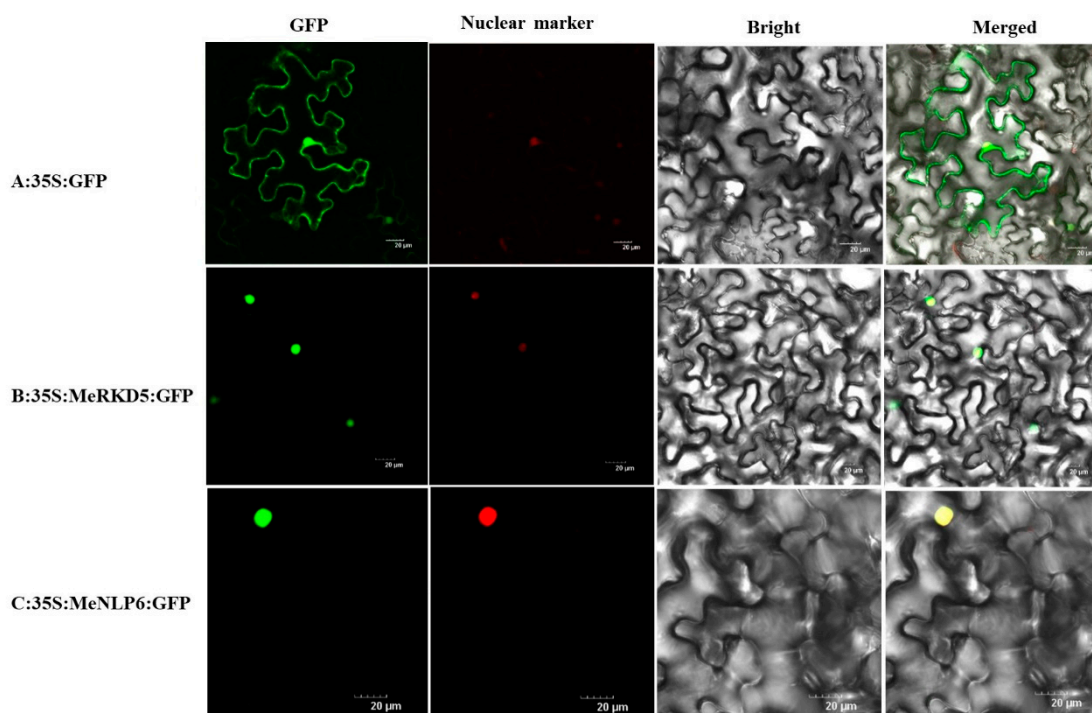


Figure 11. Fluorescent microscopic images of MeRKD5 and MeNLP6 in tobacco leaves. (A): 35S::green fluorescent protein (GFP). (B): 35S::MeNLP6::GFP. (C): 35S::MeRKD5::GFP.

2.11. Proteins Interacting with NLP6/RKD5 by Yeast Two-Hybrid System

The interaction proteins of MeNLP6 and MeRKD5 were screened in a cassava nitrogen starvation library with a yeast two-hybrid system. The results showed that five and one candidate proteins, respectively, were screened interacting with MeRKD5 and MeNLP6 (Table 1). The five proteins were agamous-like mads-box protein (MeAGL14), metallothionein-2 protein (MeMT2), Zinc finger FYVE domain containing protein (MeZF), glycera-

ldehyde-3-phosphate dehydrogenase (MeGAPD), and E3 Ubiquitin-protein ligase HUWE1 (MeUBA), and the one protein was PPR repeat family protein (MePPR).

Table 1. Screening of proteins interacting with NLP6/RKD5 by yeast two-hybrid system.

Protein Name	Identify	Gene Annotation	Amino Acid
MeNLP6	Manes.03G056700	PPR repeat family (PPR_2)	1498
	Manes.02G030100	Agamous-like mads-box protein AGL14	216
	Manes.01G174400	Metallothionein (Metallothio_2)	79
MeRKD5	Manes.08G023700	Zinc finger FYVE domain containing protein	343
	Manes.05G125200	Glyceraldehyde-3-phosphate dehydrogenase	453
	Manes.15G059000	E3 Ubiquitin-protein ligase HUWE1	3649

3. Discussion

A total of 13, 15, and 9 proteins containing the RWP-RK domain were identified in *M. esculenta*, *A. thaliana*, and *O. sativa*, respectively. There was one more RWP-RK protein than the 14 found in previous reports in *Arabidopsis* [4,14]. AT5G16100, which was discovered to include the RWP-RK motif, was named AtRKD6 (Supplementary Table S5). Previous reports on *O. sativa* revealed 7 and 16 RWP-RK proteins [4,14]. We identified nine proteins in rice, and they were named by comparing them with previous reported protein sequences (Supplementary Table S5). OsNLP1/4/5 and OsRKD1/2/4 followed the same names that were previously reported [48]. OsNLP2/3 in 2005 was later identified as the same protein named as OsNLP2, and OsRKD3 was matched with OsRKD8/9 in 2015; the rest of the rice RWP-RK proteins in 2014 were repeated or not matched.

A phylogenetic tree of RWP-RK proteins in *M. esculenta*, *O. sativa*, and *A. thaliana* indicated that RWP-RK proteins could be divided into four large groups. OsRKD2 and OsRKD4 existed separately in groups III and IV, respectively. RWP-RK proteins from cassava and *Arabidopsis*, two dicotyledonous plants, could be divided into groups I and II (Figure 2). On the basis of our analysis, MeNLP1-7, AtNLP1-9, and OsNLP1-5 proteins in group I were named as the NLP subfamily because these proteins harbored RWP-RK, PB1, and GAF domains; these results were consistent with the typical molecular structure of NLPs [14,28]. Meanwhile, MeRKD1-6, AtRKD1-5, and OsRKD2/4 proteins in group II were named as the RKD subfamily because these proteins harbored the RWP-RK domain and did not show any PB1 domains, and the GAF domains were random [14,48]. The number of NLPs was higher than that of RKDs in non-nodulating plants, such as nine NLPs and six RKDs in *Arabidopsis*, five NLPs and four RKDs in rice, and seven NLPs and six RKDs in cassava (Figures 1 and 5). The gene sequence of the NLP subfamily was longer and more complete with more introns as compared to that of the RKD subfamily. The average number of exons was greater in the NLPs, with 4.7 average exons (4–6), than in the RKDs, with only 3.3 average exons (2–5). The number of exons in 18 *Ta*NLP genes was generally 4 or 5, and most of the 19 *Ta*RKD genes did not differ much from 3 to 5 exons [11]. Similar gene structures observed in the same RWP-RK subfamily were consistent with their phylogenetic relationships.

Gene duplication helps expand the gene family [49]. Collinearity analysis indicated that *NLP1*, *RKD1*, *RKD3*, and *RKD5* had duplicate genes in cassava. Those segmental duplication events were important in the evolution of cassava RWP-RK genes (Figure 3). Syntenic maps of cassava associated with dicots (*Arabidopsis*) and monocots (rice) showed that two and five cassava RWP-RK genes, respectively, had a syntenic relationship with two rice and six *Arabidopsis* genes. Thus, some orthologous pairs formed after the divergence between dicotyledonous and monocotyledonous plants.

In general, RKDs are highly expressed in reproductive organs, thereby highlighting their regulatory importance in female gametophyte development in *Arabidopsis* [18]. In cassava, only the *MeRKD5* gene was highly expressed in flowers, whereas the other RKDs

had low expression in flowers and fruits. However, *MeRKD2/6* genes were not detected in any cassava tissues, and all the genes detected were lowest in seeds. The expression levels of *MeRKD1* in root, *MeRKD3* in leaf, and *MeRKD4/5* in root tuber were higher than those in the other parts (Figure 7). The promoter region of *RKDs* contained ABRE or/and SYB or/and HD-Zip elements in response to hormones or/and stresses (Figure 6). These results suggested that *MeRKDs* may also play a role in plants in response to stress. *MeNLP1/3/4/5/7* and *MeNLP6* genes showed the highest expression in root and storage roots, respectively. Strangely, the expression levels of *MeNLP1/2/3/7* were higher in flowers than in other parts.

Several studies on *NLP* gene function focused on nitrogen absorption [23,35,50]. In *Arabidopsis*, the expression levels of most *AtNLP* genes were induced by low nitro-gen/nitrate starvation in seedlings; for example, *AtNLP1/3* genes were induced by KNO_3 , and *AtNLP5/8/9* genes were repressed during the same treatment [14,51]. *OsNLP4* can affect the nitrate reductase activity and nitrate-dependent growth of rice by regulating gene expression within the nitrate signaling pathway [48,52]. Most *ZmNLP* genes demonstrated regulation with decreased expressions levels at 1 and 2 h after nitrate treatment and increased expression levels at 0.5 and 1.5 h [53]. Some experiments on N stress were also conducted in the study. Almost all of the *NLP* genes and *RKD1/3/5* gene in cassava root or leaf were extremely significantly induced in nitrogen-free culture, which indicates that participation in nitrogen stress remains the main function of *MeNLP* genes. Surprisingly, *MeRKD1/3/5* genes that play an important role in gamete formation may have the same function as *MeNLP* [12,17].

NLP genes can also respond to various stresses; for example, *AtNLP4/8* expression levels were induced by abiotic and chemical stresses and hormonal treatments [14], and *OsNLP4* expression in rice is repressed by drought/cold/heat abiotic stresses and induced by low phosphate availability [14]. In the present study, we completed seven different treatments on cassava leaves, and *MeRKD2/4/6* genes were not detected or no significant changes in all treatments were observed. Furthermore, *MeNLP4* expression after PEG and Man treatments and *MeRKD5* expression after MeJA treatment were not significantly changed. All the other cassava *RWP-RK* genes were significantly induced/repressed by multiple treatments in the experiments. The different variation trends in the expression levels of 10 genes at different periods of the same treatment indicated that the different *RWP-RK* genes had various responses to the same treatment. Interestingly, the transcript levels of *MeRKD1/3/5* genes were significantly different between the control and nine treatment groups. The results showed that these genes were associated not only with powdery spores but also with stresses.

As previously reported, *MeRKD5* and *MeNLP6* proteins were localized to the cell nucleus in tobacco leaf [12,21].

Five proteins interacting with *MeRKD5*, *MeAGL14*, *MeMT2*, and *MeZF* may interact with *MeRKD5* alone or together to modulate the expression of N metabolism genes by binding the NRE in their promoters [54–58]. *MeGAPD* was down-regulated while *MeUBA* was up-regulated in cassava leaves under drought stress [59]. This result indicates that *MeGAPD*, *MeUBA* and *MeRKD5* proteins may be involved in response to drought in cassava together. *UBA* was also found to interact with *GAPD* in wheat [60]. *MePPR* interacting with *MeNLP6* had not been identified in previous *RWP-RK* studies.

4. Materials and Methods

4.1. Identification of *RWP-RK* Family Members in Cassava

Herein, 15 *Arabidopsis* *RWP-RK* protein sequences were downloaded from NCBI by searching *RWP-RK* and previous reports [4,14]. The cassava protein database was obtained from JGI (<http://phytozome.jgi.doe.gov> (accessed on 30 May 2022)), *Manihot esculenta* V8.1), and rice (*Oryza sativa*) *RWP-RK* proteins were obtained by searching from EnsemblPlants (<http://plants.ensembl.org/RWP-RK> (accessed on 1 June 2022)) and previous reports [4,14,48]. Two different approaches were used to identify the cassava *RWP-RK* family proteins as follows: firstly, fifteen *Arabidopsis* *RWP-RK* amino acid sequences and

all cassava protein sequences were used as query sequences and a library for query sequences, respectively, and a local BLASTP search with E-value $\leq 1 \times 10^{-5}$ was performed; secondly, the query strategy based on HMM was derived from the PFAM database (<http://pfam.xfam.org/> (accessed on 2 June 2022)) to download the protein conserved domain RWP-RK file (PF02042) and execute the HMM search command E-value $\leq 1 \times 10^{-5}$. Cassava RWP-RK family members were finally obtained by confirming the RWP-RK domain with Batch CD-Search in NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml> (accessed on 3 June 2022)) and SMART (<https://smart.embl.de/> (accessed on 3 June 2022)). The location data of cassava RWP-RK proteins were also obtained from the genome information. MapChart2.32 (<https://www.wur.nl/en/show/mapchart.htm> (accessed on 5 June 2022)) was used to construct chr distribution map.

4.2. Phylogenetic Analysis of RWP-RK Family Members

The RWP-RK families containing the RWP-RK domain were identified to have 13, 15, and 9 proteins in *M. esculenta*, *A. thaliana*, and *O. sativa*, respectively (Supplementary Table S1). Cassava RWP-RK proteins were named based on their positions on Chrs and their included motifs (e.g., MeNLP1-7 and MeRKD1-6). RWP-RK proteins of *Arabidopsis* and rice were named by protein annotation (AtNLP1-9, AtRKD1-6, OsNLP1-5, and OsRKD1-4). Phylogenetic analysis was conducted using MEGA7 software (MEGA 7.0.21). The neighbor-joining method was used with 1000 bootstrap replications.

4.3. Cassava RWP-RK Gene Duplication

Gene duplication is ubiquitous and plays an important role in plant evolution. A gene duplication landscape was obtained using MCScanX (<http://chibba.pgml.uga.edu/mcscan2/> (accessed on 15 July 2022)). Each duplicate segment with RWP-RK family genes was selected, and the syntenic map was generated using CIRCOS (<http://circos.ca/software/download> (accessed on 15 July 2022)).

4.4. Cassava RWP-RK Gene Structure and RWP-RK Protein Motif Analysis

The gene structure data used in this study were obtained from the genome annotation file (<http://phytozome.jgi.doe.gov> (accessed on 20 July 2022)), *Manihot esculenta* V8.1) and the *Arabidopsis* Information Resource website (TAIR, <http://www.arabidopsis.org/browse/genefamily/RWP-RK.jsp> (accessed on 20 July 2022)). Gene structure analysis was conducted using Gene Structure Display Server version 2.0 (<http://gsds.cbi.pku.edu.cn/> (accessed on 20 July 2022)). The motifs of RWP-RK proteins were analyzed using the MEME program (<http://meme-suite.org/tools/meme> (accessed on 24 July 2022)) based on the amino acid sequences. The motif distribution type was zero or one occurrence per sequence, and only motifs with E-values > 0.05 were present.

4.5. Cassava RWP-RK Gene Promoter Analysis

A 1500 bp sequence of the RWP-RK gene translation initiation site upstream was extracted from the cassava genome as the promoter region. All cassava RWP-RK promoter sequences were submitted to plantCARE database (<http://bioinformatics.psb.ugent.be/webtools/PlantCare/HTML/> (accessed on 9 May 2023)) to predict the cis-acting elements of the promoter.

4.6. Transcription Profile of Cassava RWP-RK Genes in Different Tissues

Tissues of leaves, stems, roots, flowers, storage roots, and fruits were collected from field-grown South China 124 (SC124) plants (360 days old) for further RNA isolation to study the expression profiles of cassava RWP-RK genes in different tissues. SC124 was planted in Hainan China (20°02' N, 110°35' E).

4.7. Drought and Cold Treatment

To analyze the expression levels of *RWP-RK* genes in response to drought stress, SC124 plants were grown in a plastic pot (35 cm diameter × 25 cm height), and the substrate was made by mixing two parts latosolic red soil with sand (mixing 1:1). The plants were grown in a non-air-conditioned glass house with a temperature between 19 and 33 °C and sunlight, and the plants were watered two to three times every week and fertilized with a fertilizer (N:P:K = 15:15:15). Eighty-day-old plants were treated by withholding water for 14 and 18 days as drought in middle (DM) and drought in later (DL), respectively. Some plants were re-watered 24 h after 18 days of drought treatment as a re-watering sample (RW). Plants watered as normal were used as controls. Leaves were collected from the three plants. Three leaves (the second, third, and fourth leaves from the apex) were collected and mixed from each plant as a sample.

To analyze the *RWP-RK* gene expression levels in response to cold stress, 60 days old SC124 potted plants were kept at 4 °C (cold treatment) and 25 °C (control) with a 16 and 8 h light/dark cycle and 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity to study the expression profile of *RWP-RK* genes in response to cold stress. Treated leaves were collected as samples at 0.5, 1, 3, and 12 h after cold treatment with three biological replicates.

4.8. Different Salt, Osmotic, Oxidation and Hormonal Treatments

Mature leaves from SC124 potted plants (60 days old) were placed on sterile filter paper with 7% polyethylene glycol 6000 (PEG), 300 mmol mannitol (Man), 100 μmol abscisic acid (ABA), 100 μmol methyl jasmonate (MeJA), 10 mmol H_2O_2 , and 10 mmol NaCl solutions, and kept in a Petri dish at room temperature for 24 h to study the expression profile of *RWP-RK* genes in response to different treatments. The control leaves were placed on sterile filter paper with water. The samples were collected at 1, 3, 6, and 12 h after treatment with three biological replicates.

4.9. Nitrogen Starvation Treatments

One-month-old cv.60444 cassava seedlings cultured on MS medium were transplanted to MS medium in nitrogen-free conditions. Control seedlings were transplanted to MS medium. The root samples were collected for 3 and 6 days, and the leaf samples were collected for 6 days after treatment with three biological replicates.

4.10. RNA Extraction and qRT-PCR Analysis

All samples were collected, immediately preserved in liquid nitrogen, and stored at −80 °C for RNA extraction. Total RNA was extracted using the RNeasy Pure Plant Kit (DP432, Tiangen Biotech, Beijing, China) following the manufacturer's instructions. The cDNA used to verify exon/intron structure was synthesized using HiScript[®] II 1st Strand cDNA Synthesis Kit (R211-01, Vazyme Biotech Co., Nanjing, China). Herein, to detect transcription levels of the *RWP-RK* genes in different samples, specific primers for each gene were designed through NCBI Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (accessed on 20 July 2022), Supplementary Table S2), and PCR was conducted in triplicate for each cDNA sample by using an SYBR[®] Premix Ex Taq[™] II Kit FP205-02 (TIANGEN Biotech Co., Ltd., Beijing, China) on a StepOne[™] Real-Time PCR system (Thermo Fisher scientific, Waltham, MA, USA) with the following thermal cycling profiles: 95 °C for 10 min, 40 cycles of 95 °C for 10 s, and 60 °C for 30 s. The relative expression levels of the *RWP-RK* genes were calculated using the $2^{-\Delta\Delta\text{Ct}}$ method. Statistical difference was determined via one-way ANOVA.

4.11. Subcellular Localization of MeNLP6 and MeRKD5

To investigate the subcellular localization of MeNLP6 and MeRKD5, the 35S:MeNLP6-GFP and 35S:MeRKD5-GFP fusion constructs were produced by inserting the full length CDS of MeNLP6 and MeRKD5 into the pCambia1302-35S: GFP vector, and the resulting vectors were transformed into *Agrobacterium* LBA4404, transiently introduced into tobacco

leaf epidermal cells via *Agrobacterium* injection, and imaged under a confocal laser scanning microscope (Olympus FluoView FV1100).

4.12. Yeast One-Hybrid Assay (Y2H)

The cDNA of the nitrogen-free samples was obtained, and the cDNA and pGADT7 (AD)-Rec were co-transformed into yeast strain Y2HGold to directly construct the library and prepare competent cells. Two cDNA fragments encoding MeNLP6 and MeRKD5, respectively, were amplified and inserted into plasmid pGBKT7 Vector (BD) to get BD/MeNLP6 and BD/MeRKD5. The constructed vectors were co-transferred into yeast Y2HGold competent cells, and the cells were grown on SD-Trp-Leu medium for 3 days at 30 °C, then transferred to SD-Trp-Leu-His medium with 200 µM AbA⁺ at different dilutions. The yeasts were incubated at 30 °C for 5 days and the extent of yeast growth was determined. The AD and BD empty vector were used for negative control.

5. Conclusions

Genome-wide analysis of cassava revealed 13 *RWP-RK* family genes, which could be divided into *NLP* and *RKD* subfamilies including seven and six genes, respectively. The *RWP-RK* subfamily members had similar gene structures and motif compositions, and the *NLP* subfamily genes had more introns and motifs than the other subfamily. The evolution of cassava *RWP-RK* genes was affected by segmental duplication events. Six *MeNLP* genes showed higher expression levels in the root (including storage root) than in the leaf, but *RKD1/4/5* genes had high expressions levels in non-reproductive organs. A large number of hormone- and stress-related cis-acting elements, which were activated by a range of stimuli, were found through the analysis of *RWP-RK* promoters. Expression profile analyses revealed that almost all *MeNLP* and *MeRKD1/3/5* genes were susceptible to six abiotic stressors (drought, cold, salt, PEG, Man, and H₂O₂) and hormonal (ABA and MeJA) treatments. *MeNLP3/5/7* and *MeRKD3/5* genes, which displayed the most prominent changes in expression levels under nitrogen starvation, were the candidate genes to create an excellent new germplasm. The *MeRKD5* and *MeNLP6* proteins were found to localize to the cell nucleus in tobacco leaf. Five and one candidate proteins interacting with *MeRKD5* and *MeNLP6* were screened from a cassava nitrogen starvation library; these findings may be beneficial for studying *RWP-RK* gene function in cassava.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms241612925/s1>.

Author Contributions: Methodology, C.L., X.G., and P.Z.; formal analysis, X.G., C.L. and F.A.; resources, S.L., W.L., X.Y. (Xiaoling Yu) and X.G.; project administration, M.R. and X.Y. (Xiaohui Yu); supervision, F.A., M.R. and X.Y. (Xiaohui Yu); investigation, C.L., S.L., W.L., M.R. and P.Z.; writing—original draft, P.Z. and X.Y. (Xiaohui Yu); writing—review and editing, P.Z. and C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (1630052022008), and the National Natural Science Foundation of China (31960526), and the High-Level Talents Project of Hainan Provincial Natural Science Foundation of China (321RC1095).

Institutional Review Board Statement: Not applicable for this research.

Informed Consent Statement: Not applicable for this research.

Data Availability Statement: All data are available on reasonable request to the corresponding authors.

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

References

1. Khumaida, N.; Ardie, S.; Dianasari, M.; Syukur, M. Cassava (*Manihot esculenta* Crantz.) Improvement through Gamma Irradiation. *Procedia Food Sci.* **2015**, *3*, 27–34. [[CrossRef](#)]
2. Howeler, R.H. *Cassava Mineral Nutrition and Fertilization*; Centro Internacional de Agricultura Tropical CIAT: Cali, Colombia, 1981.

3. Zhao, L.; Liu, F.; Crawford, N.M.; Wang, Y. Molecular Regulation of Nitrate Responses in Plants. *Int. J. Mol. Sci.* **2018**, *19*, 2039. [[CrossRef](#)] [[PubMed](#)]
4. Chardin, C.; Girin, T.; Roudier, F.; Meyer, C.; Krapp, A. The plant RWP-RK transcription factors: Key regulators of nitrogen responses and of gametophyte development. *J. Exp. Bot.* **2014**, *65*, 5577–5587. [[CrossRef](#)] [[PubMed](#)]
5. Ferris, P.J.; Goodenough, U.W. Mating Type in *Chlamydomonas* Is Specified by mid, the Minus-Dominance Gene. *Genetics*. **1997**, *146*, 859–869. [[CrossRef](#)] [[PubMed](#)]
6. Lin, H.; Goodenough, U.W. Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, MID and MTD1. *Genetics*. **2007**, *176*, 913–925. [[CrossRef](#)] [[PubMed](#)]
7. Schauser, L.; Roussis, A.; Stiller, J.; Stougaard, J. A plant regulator controlling development of symbiotic root nodules. *Nature* **1999**, *402*, 191–195. [[CrossRef](#)] [[PubMed](#)]
8. Borisov, A.Y.; Madsen, L.H.; Tsyganov, V.E.; Umehara, Y.; Voroshilova, V.A.; Batagov, A.O.; Sandal, N.; Mortensen, A.; Schauser, L.; Ellis, N.; et al. The Sym35 gene required for root nodule development in pea is an ortholog of Nin from *Lotus japonicus*. *Plant Physiol.* **2003**, *131*, 1009–1017. [[CrossRef](#)]
9. Camargo, A.; Llamas, A.; Schnell, R.A.; Higuera, J.J.; Gonzalez-Ballester, D.; Lefebvre, P.A.; Fernandez, E.; Galvan, A. Nitrate signaling by the regulatory gene NIT2 in *Chlamydomonas*. *Plant Cell*. **2007**, *19*, 3491–3503. [[CrossRef](#)]
10. Yu, L.H.; Wu, J.; Tang, H.; Yuan, Y.; Wang, S.M.; Wang, Y.P.; Zhu, Q.S.; Li, S.G.; Xiang, C.B. Overexpression of Arabidopsis NLP7 improves plant growth under both nitrogen-limiting and -sufficient conditions by enhancing nitrogen and carbon assimilation. *Sci. Rep.* **2016**, *6*, 27795. [[CrossRef](#)]
11. Kumar, A.; Batra, R.; Gahlaut, V.; Gautam, T.; Kumar, S.; Sharma, M.; Tyagi, S.; Singh, K.P.; Balyan, H.S.; Pandey, R.; et al. Genome-wide identification and characterization of gene family for RWP-RK transcription factors in wheat (*Triticum aestivum* L.). *PLoS ONE* **2018**, *13*, e0208409. [[CrossRef](#)]
12. Koszegi, D.; Johnston, A.J.; Rutten, T.; Czihal, A.; Altschmied, L.; Kumlehn, J.; Wust, S.E.; Kirioukhova, O.; Gheyselinck, J.; Grossniklaus, U.; et al. Members of the RKD transcription factor family induce an egg cell-like gene expression program. *Plant J.* **2011**, *67*, 280–291. [[CrossRef](#)] [[PubMed](#)]
13. Castaings, L.; Camargo, A.; Pocholle, D.; Gaudon, V.; Texier, Y.; Boutet-Mercey, S.; Taconnat, L.; Renou, J.P.; Daniel-Vedele, F.; Fernandez, E.; et al. The nodule inception-like protein 7 modulates nitrate sensing and metabolism in Arabidopsis. *Plant J.* **2009**, *57*, 426–435. [[CrossRef](#)] [[PubMed](#)]
14. Schauser, L.; Wieloch, W.; Stougaard, J. Evolution of NIN-like proteins in Arabidopsis, rice, and *Lotus japonicus*. *J. Mol. Evol.* **2005**, *60*, 229–237. [[CrossRef](#)] [[PubMed](#)]
15. Konishi, M.; Yanagisawa, S. The role of protein-protein interactions mediated by the PB1 domain of NLP transcription factors in nitrate-inducible gene expression. *BMC Plant Biol.* **2019**, *19*, 90. [[CrossRef](#)] [[PubMed](#)]
16. Liu, K.H.; Niu, Y.; Konishi, M.; Wu, Y.; Du, H.; Sun Chung, H.; Li, L.; Boudsocq, M.; McCormack, M.; Maekawa, S.; et al. Discovery of nitrate-CPK-NLP signalling in central nutrient-growth networks. *Nature* **2017**, *545*, 311–316. [[CrossRef](#)] [[PubMed](#)]
17. Koi, S.; Hisanaga, T.; Sato, K.; Shimamura, M.; Yamato, K.T.; Ishizaki, K.; Kohchi, T.; Nakajima, K. An Evolutionarily Conserved Plant RKD Factor Controls Germ Cell Differentiation. *Curr. Biol.* **2016**, *26*, 1775–1781. [[CrossRef](#)]
18. Tedeschi, F.; Rizzo, P.; Rutten, T.; Altschmied, L.; Baumlein, H. RWP-RK domain-containing transcription factors control cell differentiation during female gametophyte development in Arabidopsis. *New Phytol.* **2017**, *213*, 1909–1924. [[CrossRef](#)]
19. Soyano, T.; Kouchi, H.; Hirota, A.; Hayashi, M. Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet.* **2013**, *9*, e1003352. [[CrossRef](#)]
20. Marchive, C.; Roudier, F.; Castaings, L.; Brehaut, V.; Blondet, E.; Colot, V.; Meyer, C.; Krapp, A. Nuclear retention of the transcription factor NLP7 orchestrates the early response to nitrate in plants. *Nat. Commun.* **2013**, *4*, 1713. [[CrossRef](#)]
21. Yan, D.; Easwaran, V.; Chau, V.; Okamoto, M.; Ierullo, M.; Kimura, M.; Endo, A.; Yano, R.; Pasha, A.; Gong, Y.; et al. NIN-like protein 8 is a master regulator of nitrate-promoted seed germination in Arabidopsis. *Nat. Commun.* **2016**, *7*, 13179. [[CrossRef](#)]
22. Marsh, J.F.; Rakocevic, A.; Mitra, R.M.; Brocard, L.; Sun, J.; Eschstruth, A.; Long, S.R.; Schultze, M.; Ratet, P.; Oldroyd, G.E. *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol.* **2007**, *144*, 324–335. [[CrossRef](#)]
23. Lin, J.S.; Li, X.; Luo, Z.; Mysore, K.S.; Wen, J.; Xie, F. NIN interacts with NLPs to mediate nitrate inhibition of nodulation in *Medicago truncatula*. *Nat. Plants* **2018**, *4*, 942–952. [[CrossRef](#)] [[PubMed](#)]
24. Zhao, J.; Xin, H.; Qu, L.; Ning, J.; Peng, X.; Yan, T.; Ma, L.; Li, S.; Sun, M.X. Dynamic changes of transcript profiles after fertilization are associated with de novo transcription and maternal elimination in tobacco zygote, and mark the onset of the maternal-to-zygotic transition. *Plant J.* **2011**, *65*, 131–145. [[CrossRef](#)] [[PubMed](#)]
25. Yokota, K.; Hayashi, M. Function and evolution of nodulation genes in legumes. *Cell. Mol. Life Sci.* **2011**, *68*, 1341–1351. [[CrossRef](#)] [[PubMed](#)]
26. Liu, C.W.; Breakspear, A.; Guan, D.; Cerri, M.R.; Jackson, K.; Jiang, S.; Robson, F.; Radhakrishnan, G.V.; Roy, S.; Bone, C.; et al. NIN Acts as a Network Hub Controlling a Growth Module Required for Rhizobial Infection. *Plant Physiol.* **2019**, *179*, 1704–1722. [[CrossRef](#)]
27. Mu, X.; Luo, J. Evolutionary analyses of NIN-like proteins in plants and their roles in nitrate signaling. *Cell. Mol. Life Sci.* **2019**, *76*, 3753–3764. [[CrossRef](#)]

28. Wu, Z.; Liu, H.; Huang, W.; Yi, L.; Qin, E.; Yang, T.; Wang, J.; Qin, R. Genome-Wide Identification, Characterization, and Regulation of RWP-RK Gene Family in the Nitrogen-Fixing Clade. *Plants* **2020**, *9*, 1178. [\[CrossRef\]](#)
29. Clavijo, F.; Diedhiou, I.; Vaissayre, V.; Brottier, L.; Acolatse, J.; Moukouanga, D.; Crabos, A.; Auguy, F.; Franche, C.; Gherbi, H.; et al. The Casuarina NIN gene is transcriptionally activated throughout Frankia root infection as well as in response to bacterial diffusible signals. *New Phytol.* **2015**, *208*, 887–903. [\[CrossRef\]](#)
30. Liu, M.; Soyano, T.; Yano, K.; Hayashi, M.; Kawaguchi, M. ERN1 and CYCLOPS coordinately activate NIN signaling to promote infection thread formation in *Lotus japonicus*. *J. Plant Res.* **2019**, *132*, 641–653. [\[CrossRef\]](#)
31. Liu, J.; Bisseling, T. Evolution of NIN and NIN-like Genes in Relation to Nodule Symbiosis. *Genes* **2020**, *11*, 777. [\[CrossRef\]](#)
32. Zhang, T.T.; Kang, H.; Fu, L.L.; Sun, W.J.; Gao, W.S.; You, C.X.; Wang, X.F.; Hao, Y.J. NIN-like protein 7 promotes nitrate-mediated lateral root development by activating transcription of TRYPTOPHAN AMINOTRANSFERASE RELATED 2. *Plant Sci.* **2021**, *303*, 110771. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Ge, M.; Liu, Y.; Jiang, L.; Wang, Y.; Lv, Y.; Zhou, L.; Liang, S.; Bao, H.; Zhao, H. Genome-wide analysis of maize NLP transcription factor family revealed the roles in nitrogen response. *Plant Growth Regul.* **2017**, *84*, 95–105. [\[CrossRef\]](#)
34. Liu, M.; Chang, W.; Fan, Y.; Sun, W.; Qu, C.; Zhang, K.; Liu, L.; Xu, X.; Tang, Z.; Li, J.; et al. Genome-Wide Identification and Characterization of NODULE-INCEPTION-Like Protein (NLP) Family Genes in *Brassica napus*. *Int. J. Mol. Sci.* **2018**, *19*, 2270. [\[CrossRef\]](#)
35. Feng, Z.Q.; Li, T.; Wang, X.; Sun, W.J.; Zhang, T.T.; You, C.X.; Wang, X.F. Identification and characterization of apple MdNLP7 transcription factor in the nitrate response. *Plant Sci.* **2022**, *316*, 111158. [\[CrossRef\]](#)
36. Reddy, A.S.; Marquez, Y.; Kalyna, M.; Barta, A. Complexity of the alternative splicing landscape in plants. *Plant Cell.* **2013**, *25*, 3657–3683. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Kuo, Y.T.; Chao, Y.T.; Chen, W.C.; Shih, M.C.; Chang, S.B. Segmental and tandem chromosome duplications led to divergent evolution of the chalcone synthase gene family in *Phalaenopsis* orchids. *Ann. Bot.* **2019**, *123*, 69–77. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Holub, E.B. The arms race is ancient history in *Arabidopsis*, the wildflower. *Nature* **2001**, *1*, 516–527. [\[CrossRef\]](#)
39. Zhang, P.; Gu, Z.L.; Li, W.H. Different evolutionary patterns between young duplicate genes in the human genome. *Genome Biol.* **2003**, *4*, R56. [\[CrossRef\]](#)
40. Yamaguchi-Shinozaki, K.; Shinozaki, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Sci.* **2005**, *10*, 88–94. [\[CrossRef\]](#)
41. Gomez-Porras, J.L.; Riano-Pachon, D.M.; Dreyer, I.; Mayer, J.E.; Mueller-Roeber, B. Genome-wide analysis of ABA-responsive elements ABRE and CE3 reveals divergent patterns in *Arabidopsis* and rice. *BMC Genom.* **2007**, *8*, 260. [\[CrossRef\]](#)
42. He, G.; Liu, P.; Zhao, H.; Sun, J. The HD-ZIP II Transcription Factors Regulate Plant Architecture through the Auxin Pathway. *Int. J. Mol. Sci.* **2020**, *21*, 3250. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Garretton, V.; Carpinelli, J.; Jordana, X.; Holuigue, L. The as-1 promoter element is an oxidative stress-responsive element and salicylic acid activates it via oxidative species. *Plant Physiol.* **2002**, *130*, 1516–1526. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Ruan, M.B.; Guo, X.; Wang, B.; Yang, Y.L.; Li, W.Q.; Yu, X.L.; Zhang, P.; Peng, M. Genome-wide characterization and expression analysis enables identification of abiotic stress-responsive MYB transcription factors in cassava (*Manihot esculenta*). *J. Exp. Bot.* **2017**, *68*, 3657–3672. [\[CrossRef\]](#)
45. Tran, L.S.; Nakashima, K.; Sakuma, Y.; Simpson, S.D.; Fujita, Y.; Maruyama, K.; Fujita, M.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell.* **2004**, *16*, 2481–2498. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Parry, T.J.; Theisen, J.W.; Hsu, J.Y.; Wang, Y.L.; Corcoran, D.L.; Eustice, M.; Ohler, U.; Kadonaga, J.T. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev.* **2010**, *24*, 2013–2018. [\[CrossRef\]](#)
47. Abe, H.; Yamaguchi-Shinozaki, K.; Urao, T.; Iwasaki, T.; Hosokawa, D.; Shinozaki, K. Role of *Arabidopsis* MYC and MYB Homologs in Drought- and Absciscic Acid-Regulated Gene Expression. *Plant Cell.* **1997**, *9*, 1859–1868. [\[CrossRef\]](#)
48. Jagadhesan, B.; Sathee, L.; Meena, H.S.; Jha, S.K.; Chinnusamy, V.; Kumar, A.; Kumar, S. Genome wide analysis of NLP transcription factors reveals their role in nitrogen stress tolerance of rice. *Sci. Rep.* **2020**, *10*, 9368. [\[CrossRef\]](#)
49. Magadum, S.; Banerjee, U.; Murugan, P.; Gangapur, D.; Ravikesavan, R. Gene duplication as a major force in evolution. *J. Genet.* **2013**, *92*, 155–161. [\[CrossRef\]](#)
50. Konishi, M.; Yanagisawa, S. *Arabidopsis* NIN-like transcription factors have a central role in nitrate signalling. *Nat. Commun.* **2013**, *4*, 1617. [\[CrossRef\]](#)
51. Hruz, T.; Laule, O.; Szabo, G.; Wessendorp, F.; Bleuler, S.; Oertle, L.; Widmayer, P.; Gruissem, W.; Zimmermann, P. Genevestigator v3: A reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinform.* **2008**, *2008*, 420747. [\[CrossRef\]](#)
52. Wang, M.; Hasegawa, T.; Beier, M.; Hayashi, M.; Ohmori, Y.; Yano, K.; Teramoto, S.; Kamiya, T.; Fujiwara, T. Growth and Nitrate Reductase Activity Are Impaired in Rice Oslp4 Mutants Supplied with Nitrate. *Plant Cell Physiol.* **2021**, *62*, 1156–1167. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Ge, M.; Wang, Y.; Liu, Y.; Jiang, L.; He, B.; Ning, L.; Du, H.; Lv, Y.; Zhou, L.; Lin, F.; et al. The NIN-like protein 5 (ZmNLP5) transcription factor is involved in modulating the nitrogen response in maize. *Plant J.* **2020**, *102*, 353–368. [\[CrossRef\]](#) [\[PubMed\]](#)

54. Wu, J.; Zhang, Z.S.; Xia, J.Q.; Alfatih, A.; Song, Y.; Huang, Y.J.; Wan, G.Y.; Sun, L.Q.; Tang, H.; Liu, Y.; et al. Rice NIN-LIKE PROTEIN 4 plays a pivotal role in nitrogen use efficiency. *Plant Biotechnol. J.* **2021**, *19*, 448–461. [[CrossRef](#)]
55. Higuera, J.J.; Fernandez, E.; Galvan, A. Chlamydomonas NZF1, a tandem-repeated zinc finger factor involved in nitrate signalling by controlling the regulatory gene NIT2. *Plant Cell Environ.* **2014**, *37*, 2139–2150. [[CrossRef](#)]
56. Mekawy, A.M.M.; Assaha, D.V.M.; Ueda, A. Constitutive overexpression of rice metallothionein-like gene OsMT-3a enhances growth and tolerance of Arabidopsis plants to a combination of various abiotic stresses. *J. Plant Res.* **2020**, *133*, 429–440. [[CrossRef](#)]
57. Perez-Ruiz, R.V.; Garcia-Ponce, B.; Marsch-Martinez, N.; Ugartechea-Chirino, Y.; Villajuana-Bonequi, M.; de Folter, S.; Azpeitia, E.; Davila-Velderrain, J.; Cruz-Sanchez, D.; Garay-Arroyo, A.; et al. XAANTAL2 (AGL14) Is an Important Component of the Complex Gene Regulatory Network that Underlies Arabidopsis Shoot Apical Meristem Transitions. *Mol. Plant.* **2015**, *8*, 796–813. [[CrossRef](#)]
58. Zhang, J.; Zhang, M.; Tian, S.; Lu, L.; Shohag, M.J.; Yang, X. Metallothionein 2 (SaMT2) from *Sedum alfredii* Hance confers increased Cd tolerance and accumulation in yeast and tobacco. *PLoS ONE.* **2014**, *9*, e102750. [[CrossRef](#)] [[PubMed](#)]
59. Zhao, P.; Liu, P.; Shao, J.; Li, C.; Wang, B.; Guo, X.; Yan, B.; Xia, Y.; Peng, M. Analysis of different strategies adapted by two cassava cultivars in response to drought stress: Ensuring survival or continuing growth. *J. Exp. Bot.* **2015**, *66*, 1477–1488. [[CrossRef](#)] [[PubMed](#)]
60. Li, X.; Wei, W.; Li, F.; Zhang, L.; Deng, X.; Liu, Y.; Yang, S. The Plastidial Glyceraldehyde-3-Phosphate Dehydrogenase Is Critical for Abiotic Stress Response in Wheat. *Int. J. Mol. Sci.* **2019**, *20*, 1104. [[CrossRef](#)]

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