Molecular Plant Breeding 2022, Vol.13, No.18, 1-11
http://genbreedpublisher.com/index.php/mpb

Research Article Open Access

Genome-wide Identification and Bioinformatics Analysis of NRL Gene Family in Rice

Yangyang Hu, Linlin Sun, Duoxiang Wang, Dabing Zhang, Wenguo Cai
School of Life Sciences and Biotechnology, Shanghai Jiaotong University, Shanghai, 200240, China
Corresponding author email: wenguo.cai@sjtu.edu.cn
Molecular Plant Breeding, 2022, Vol.13, No.18 doi: 10.5376/mpb.2022.13.0018
Received: 18 May, 2022
Accepted: 29 May, 2022
Published: 15 Jun., 2022

Preferred citation for this article:
Hu Y.Y., Sun L.L., Wang D.X., Zhang D.B., and Cai W.G., 2022, Genome-wide identification and bioinformatics analysis of NRL gene family in rice, Molecular Plant Breeding, 13(18): 1-11 (doi: 10.5376/mpb.2022.13.0018)

Abstract NRL (NPH3/RPT2-Like) gene family is plant-specific and it is named after its two members in Arabidopsis, NPH3 (NONPHOTOTROPIC HYPOCOTYL 3) and RPT2 (ROOT PHOTOTROPISM 2), which are involved in phototropism. In this study, we identified 27 rice *NRL* genes distributed on 12 chromosomes except chromosome 10 with bioinformatic analysis. Gene duplication analysis revealed that whole genome duplication (WGD)/segmental duplication played a more vital role than tandem duplication in rice NRL gene family expansions. A total of 192 NRL gene family members from seven species were identified and subsequently they were divided into six groups in phylogenetic relationship analysis. Structural feature analysis showed that most members of rice NRL gene family contained specific domains of NRL family. Furthermore, based on RNA-Seq data and microarray data, expression analysis showed that rice NRL family genes were expressed in several tissues and some genes displayed higher expression in leaves, inflorescence and anther. Identification and analysis were conducted in NRL gene family from rice and these results provided some reference-able message for further functional research.
Keywords Rice; NRL; Gene family; Bioinformatics

In the process of plant growth and morphogenesis formation, it will respond to the changes in the wavelength, intensity and direction of external light signals, one of which is phototropism in plants, that is, the phenomenon that plant organs bend toward or away from a specific direction of light source (Holland et al., 2009). Arabidopsis phototropin PHOT1 (PHOTOTROPIN1) and PHOT2 (PHOTOTROPIN2) are involved in the regulation of phototropism, chloroplast movement, stomatal opening, leaf extension and localization (Christie et al., 2015). NPH3 (NONPHOTOTROPIC HYPOCOTYL3) and RPT2 (ROOT PHOTOTROPISM2) in *Arabidopsis thaliana* were identified in molecular genetic analysis of phototropic defect mutants. Studies have shown that NPH3 and RPT2 are involved in phototropism, leaf extension and location regulation mediated by phototropism, while RPT2 also plays a role in chloroplast movement (Christie et al., 2015, Suetsugu et al., 2016, Christie et al., 2018). NPH3 and RPT2 belong to a plant-specific NRL (NPH3/RPH2-Like) family, which mainly consists of three characteristic domains: the BTB (Broad complex, tramtrack, and bric à brac) domain at the N-terminal, the coiled helix domain at the C-terminal and the NPH3 domain in the middle (Motchoulski and Liscum, 1999, Sakai et al., 2000, Liscum et al., 2014). The NPH3 domain is common to the members of the NRL gene family, while some members of the NRL gene family lack the N-terminal BTB domain or the C-terminal coiled helix domain (Pedmale et al., 2010). In the functional studies of the three characteristic domains of NRL gene family, the BTB domain at the N-terminal was involved in the ubiquitination of phot1 mediated by E3 ubiquitin ligase $CRL3^{NPH3}$, and the crimp structure at the C-terminal was involved in the protein interaction between NPH3 and phot1. However, the function of NPH3 domain has not been reported (Motchoulski and Liscum, 1999, Roberts et al., 2011).

There are 33 NRL gene family members in *Arabidopsis thaliana*, including NPH3 and RPT2, but their functions are not limited to physiological activities such as plant phototropism, represented by NPH3 and RPT2 (Pedmale et al., 2010). AtNRL31 (AtNCH1/AtSR1P1) is involved in auxin-independent chloroplast accumulation and is also a positive regulator of plant immunity (Zhang et al., 2014, Suetsugu et al., 2016). AtNRL6 (AtNYP2), AtNRL7

(AtNYP4), AtNRL20 (At*MAB4/ENP/NPY1)*, AtNRL21 (AtNYP5) and AtNRL30 has functional redundancy in coordination with AGC kinase (cAMP-dependent protein kinase A, cGMP-dependent protein kinase G and phospholipid-dependent protein kinase C) in regulating auxin transport in organogenesis and root gravity response (Cheng et al., 2008, Li et al., 2011). AtNRL8 (AtSETH6) is involved in the regulation of pollen germination and pollen tube growth, while AtNRL23 (AtDOT3) is related to vein structure regulation pattern formation, vascular system development and reproductive development (Lalanne et al., 2004, Petricka et al., 2008). The study of *AtNPH3* homologous gene *CPT1* (*COLEOPTILE PHOTOTROPISM 1*) in rice showed that CPT1-mediated coleoptile phototropism was achieved by auxin lateral transport and subsequent growth redistribution (Haga et al., 2005). After overexpression of *WIN1* (*WINDING 1*) gene in rice, the helical phenotype and auxin were unevenly distributed in seedlings. Exogenous application of auxin polar transport inhibitor did not affect the helical phenotype, and overexpression of *WIN1* did not affect the phototropism (Cheng et al., 2017). *SlBTB5*, a member of NRL gene family in tomato cultivar M82, is highly expressed in roots and flowers and down-regulated under environmental stress. *StNRL1*, a member of NRL gene family in potato, is a sensitive factor in plant immunity caused by potato late blight and participates in plant immune regulation (Yang et al., 2016, Li et al., 2018). NRL gene family members are widely involved in plant growth and development and stress response. However, the function of this gene family member, especially in monocotyledonous plants, remains to be further studied.

At present, there are few studies on NRL gene family, only in *Arabidopsis thaliana*, while the whole genome analysis of NRL gene family in rice has not been reported. In this study, rice NRL gene family members were screened and analyzed by bioinformatics methods to provide reference and basis for further study of rice NRL gene family.

1 Results and Analysis

1.1 Identification of NRL gene family in rice

In this study, 30 rice NRL family genes were obtained by local BLAST and HMMER analysis methods, and the predicted rice NRL family genes were screened by NCBI online tool Batch CD-Search (https://www.ncbi.nlm.nih.gov/cdd/), SMART (http://smart.embl-heidelberg.de/), and Pfam (http://pfam.xfam.org/). A total of 27 rice NRL family genes were identified and named as *OsNRL1~OsNRL27* (Table 1). The length of CDS of rice NRL family genes ranged from 924 bp to 2 289 bp, and the amino acid sequence length ranged from 307 to 762. The predicted molecular weight of the protein was between 31.37~80.62 kD, and the theoretical isoelectric point was between 4.66~10.21. The NRL gene family members in rice were predicted to locate in the cytomembrane, cytoplasm, nucleus, plastid, peroxisome and endoplasm. In which, eleven NRL gene family members were predicted to locate in the nucleus, and six NRL gene family members were predicted to locate in the cytomembrane.

1.2 Chromosome distribution and gene duplication analysis of NRL gene family in rice

The location information of rice NRL family genes were obtained from RGAP (Rice genome annotation project) (http://rice.plantbiology.msu.edu/), and the chromosome distribution map of rice NRL family genes was drawn by TBtools software (Figure 1). The 27 rice NRL family genes were unevenly distributed on rice chromosomes except chromosome 10. Among them, the *NRL* genes on chromosome 3 were the most, with a total of seven genes, while only one *NRL* gene was distributed on chromosomes 5, 6 and 8, respectively.

MCScanX software was used to analyze rice tandem duplication events and whole genome duplication (WGD) or segmental duplication events. A total of seven gene duplication events occurred in the NRL gene family of rice, including two tandem duplication events and five WGD or segmental duplication events (Figure 1; Table 2). Gene pairs *OsNRL22-OsNRL23* and *OsNRL24-OsNRL25* were tandem replication, while gene pairs *OsNRL1-OsNRL15*, *OsNRL4-OsNRL12*, *OsNRL22-OsNRL24*, *OsNRL26-OsNRL8* and *OsNRL27-OsNRL9* were WGD or segmental duplication. DnaSP 6.0 software was used to calculate the Ka/Ks values of duplication gene pairs to evaluate the driving force of *NRL* gene evolution in rice. Ka/Ks > 1 means positive selection, Ka/Ks < 1 means purification selection, Ka/Ks = 1 means neutral selection. The Ka/Ks values of duplication gene pairs events in rice NRL family genes ranged from 0.107 0 to 0.546 5, indicating that they were purifying selection during evolution (Table 2).

Table 1 Information of NRL family genes in rice

Gene name	Gene ID	Length of CDS	Amino acid sequence length	MW (kD)	p*I*	Predicted protein location
OsNRL1	LOC_Os01g08130	924	307	31.37	7.58	Plastid
OsNRL2	LOC_Os01g57230	1 980	659	70.63	5.33	Peroxisome
OsNRL3	LOC_Os02g35970	2 289	762	80.62	7.73	Cytomembrane
OsNRL4	LOC_Os02g38120	1 962	653	69.62	7.04	Nucleus
OsNRL5	LOC_Os03g10800	1 818	605	66.66	8.31	Nucleus
OsNRL6	LOC_Os03g10880	1 881	626	68.03	8.2	Nucleus
OsNRL7	LOC_Os03g22600	2 016	671	71.52	9.22	Plastid
OsNRL8	LOC_Os03g41350	1 905	634	69.53	6.51	Nucleus
OsNRL9	LOC_Os03g43990	1 875	624	69.53	7.07	Plastid
OsNRL10	LOC_Os03g52880	2 004	667	72.93	6.17	Nucleus
OsNRL11	LOC_Os03g55830	2 052	683	74.63	5.91	Nucleus
OsNRL12	LOC_Os04g40100	1 884	627	66.34	9.57	Cytomembrane
OsNRL13	LOC_Os04g54400	1 935	644	70.04	4.66	Cytomembrane
OsNRL14	LOC_Os04g57800	1 701	566	61.74	8.42	Cytomembrane
OsNRL15	LOC_Os05g08530	1 506	501	51.33	10.21	Cytomembrane
OsNRL16	LOC_Os06g08550	1 935	644	70.02	5.97	Endoplasm
OsNRL17	LOC_Os07g36230	1 707	568	60.01	8.08	Cytomembrane
OsNRL18	LOC_Os07g39530	2 166	721	78.42	6.89	Nucleus
OsNRL19	LOC_Os08g03650	1 926	641	71.31	6.38	Nucleus
OsNRL20	LOC_Os09g09370	1 764	587	63.96	7.55	Cytoplasm
OsNRL21	LOC_Os09g25330	1 857	618	68.33	8.29	Nucleus
OsNRL22	LOC_Os11g02610	1 716	571	60.54	8.21	Nucleus
OsNRL23	LOC_Os11g02620	1 932	643	70.89	6.51	Plastid
OsNRL24	LOC_Os12g02530	1 716	571	60.68	8.47	Nucleus
OsNRL25	LOC_Os12g02540	1 956	651	71.85	6.74	Plastid
OsNRL26	LOC_Os12g39380	1 917	638	69.63	5.26	Endoplasm
OsNRL27	LOC_Os12g41910	1 953	650	72.09	5.85	Cytoplasm-nucleus

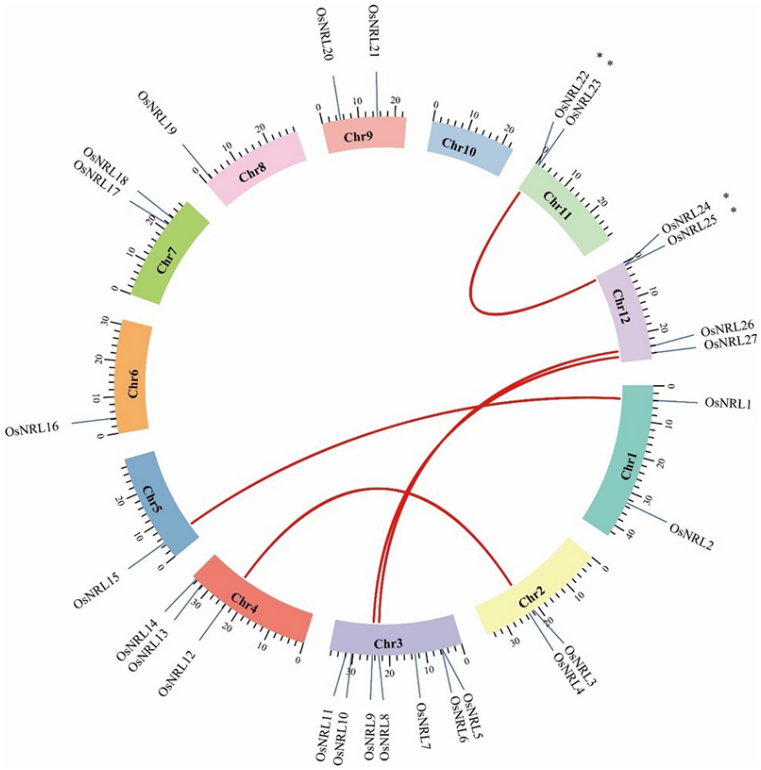

Figure 1 Chromosomal distribution and gene duplications of *OsNRLs*

Note: The length of chromosome is scaled in Megabase; Whole genome duplication (WGD)/segmental duplication gene pairs are linked by a red line; Tandem duplication gene pairs are marked with asterisks

Table 2 Ka/Ks values for duplication gene pairs of NRL family in rice

Gene I	Gene II	Value of Ka/Ks	Duplication type	Purifying selection
OsNRL22	OsNRL23	0.424 0	Tandem duplication	Yes
OsNRL24	OsNRL25	0.460 2	Tandem duplication	Yes
OsNRL1	OsNRL15	0.546 5	WGD or segmental duplication	Yes
OsNRL4	OsNRL12	0.504 2	WGD or segmental duplication	Yes
OsNRL22	OsNRL24	0.218 9	WGD or segmental duplication	Yes
OsNRL26	OsNRL8	0.107 0	WGD or segmental duplication	Yes
OsNRL27	OsNRL9	0.127 4	WGD or segmental duplication	Yes

Note: Ka: Nonsynonymous substitution rates; Ks: Synonymous substitution rates

1.3 Phylogenetic analysis of NRL gene family in rice

In order to further clarify the NRL gene family in rice, the protein sequences of dicotyledon Arabidopsis, grape and tomato, monocotyledon maize, barley, and Brachypodium were obtained from the online database Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html), and 33, 24, 27, 31, 26 and 24 NRL gene family members were identified, respectively, and the phylogenetic relationship was analyzed (Figure 2). The results showed that there were 192 members of NRL gene family in seven species, which could be divided into six groups (Group I~VI) with 8, 23, 42, 39, 54 and 26 members, respectively. *OsNRL3* (*OsCPT1*) in Group VI is homologous to *AtNPH3*, which is consistent with previous reports (Haga et al., 2005). In Group VI, there are two genes in maize and tomato that had similar evolutionary relationships with *OsNRL3* (*OsCPT1*) and *AtNPH3*, that is, *Zm00001d004440*, *Zm00001d016898*, *Solyc01g105680* and *Solyc10g047530*. While only *HORVU6Hr1G015480*, *Bradi3g46480* and *VIT_203s0038g00270* are homologous genes of *OsNRL3* (*OsCPT1*) and *AtNPH3* in barley, Brachypodium and grape, respectively (Figure 2). In Group IV, *OsNRL22*, *OsNRL24* and *Zm00001d005051*, *Zm00001d023314* are homologous genes of *AtRPT2*, while the homologous genes of *AtRPT2* in barley, Brachypodium, tomato and grape are *HORVU4Hr1G023260*, *Bradi4g25900*, *Solyc07g043130* and *VIT_206s0004g08230*, respectively. Group V contains 54 NRL gene family members, which is the largest group of NRL gene family members. There are 8 NRL gene family members in Arabidopsis and rice, among which AtNRL6, AtNRL7, AtNRL20, AtNRL21 and AtNRL30 have functional redundancy (Cheng et al., 2008, Li et al., 2011).

1.4 Genetic structure and conserved motif analysis of NRL family in rice

The results of genetic structure analysis of NRL family in rice showed that there were 4 exons and 3 introns in 17 rice NRL family genes, such as *OsNRL18*, *OsNRL7*, *OsNRL11* and *OsNRL20.* 5 rice NRL family genes contained 3 exons and 2 introns, that was, *OsNRL10*, *OsNRL14*, *OsNRL17*, *OsNRL22* and *OsNRL24.* 2 rice NRL family genes contained 5 exons and 4 introns, that was, *OsNRL3* and *OsNRL6.* While *OsNRL15* contained 2 exons and 1 intron, *OsNRL1* contained 1 intron. *OsNRL20* contained the largest number of exons/introns, with 6 exons and 5 introns (Figure 3B). Further conserved motif analysis showed that most rice NRL family members contained 14~15 conserved motifs (Figure 3C). The BTB domain was predicted to be composed of N-terminal motifs 2, 4 and 6, N-terminal motifs 1, 3, 5 ,7, 10, 11 and 14 were predicted to be composed of NPH3 domain, and C-terminal motif 13 was predicted to be composed of coiled-coil domain. Both BTB domain and coiled-coil domain can be involved in protein interaction. The absence of any domain may affect the interaction with potential specific proteins, while the specific function of NPH3 domain remains to be further studied (Motchoulski and Liscum, 1999, Inada et al., 2004, Roberts et al., 2011).

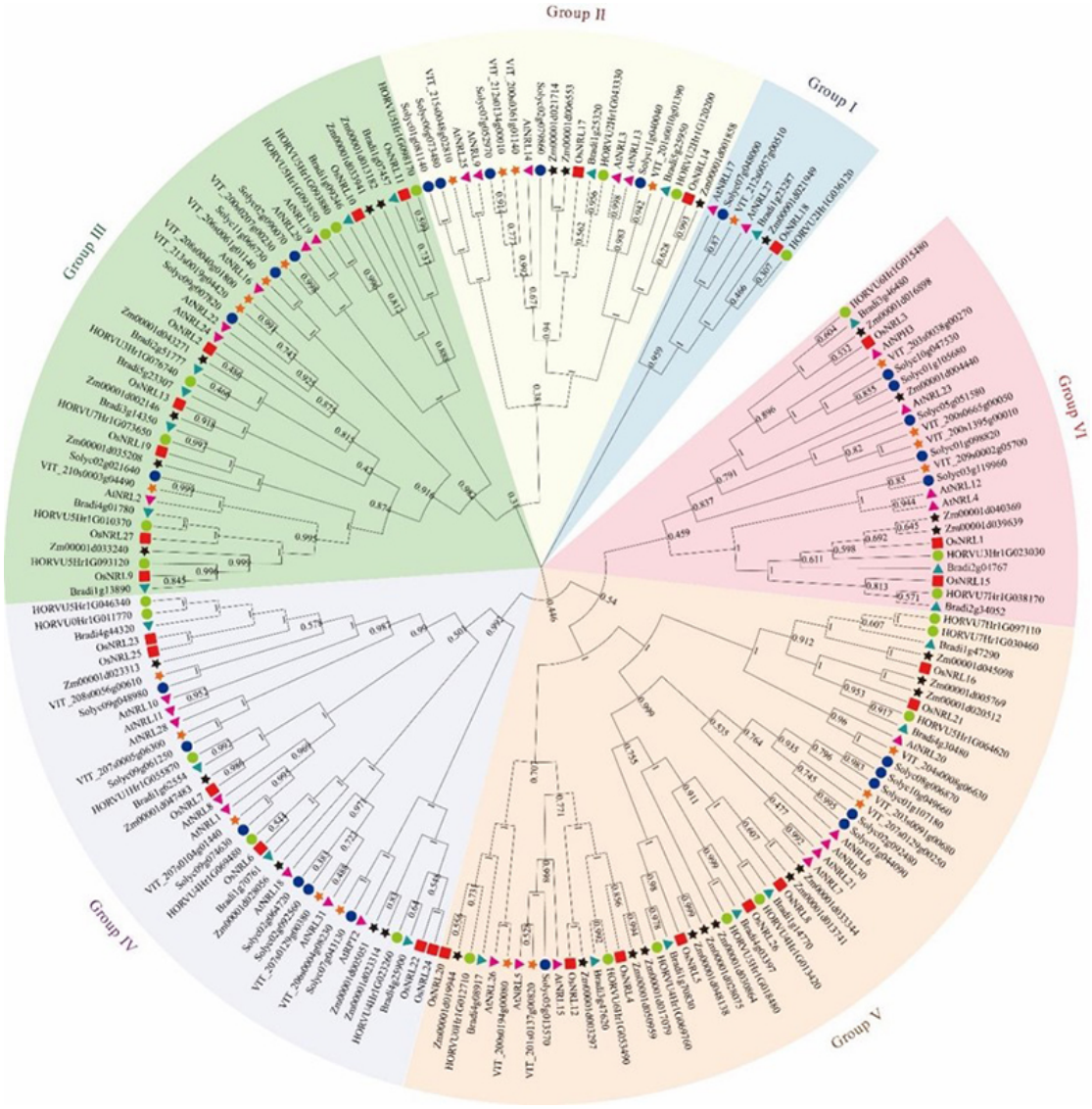

Figure 2 Phylogenetic relationship of NRL gene family in rice, maize, barely, Brachypodium, Arabidopsis, grape and tomato

Note: Neighbor-joining tree is constructed using MEGA 7.0 with 1000 bootstrap replicates; Rice, mazie, barely and Brachypodium are marked with red squares, black stars, green circles and cyan triangles; Arabidopis, grape and tomato are marked with magenta triangles, orange stars and blue circles

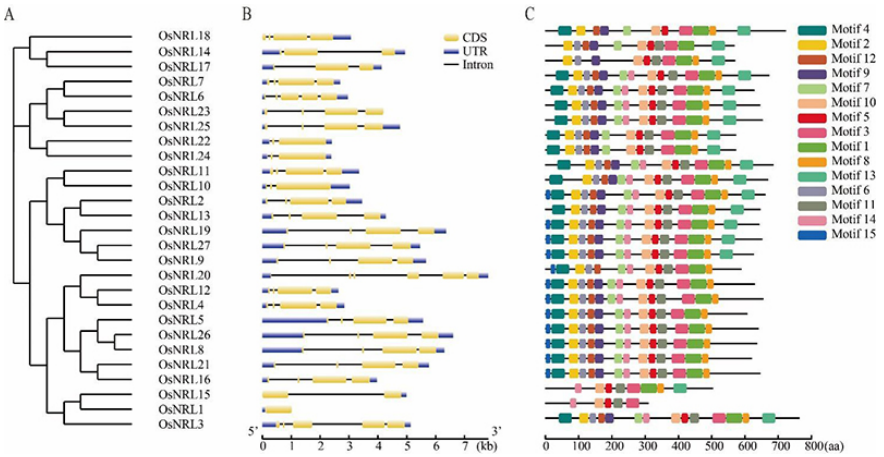

Figure 3 Phylogenetic tree, exon–intron structures and conserved motif analysis of NRL gene family in rice

Note: A: Phylogenetic tree of NRL family in rice; B: exon-intron structures analysis of NRL family in rice; C: Conserved motif analysis of NRL family in rice

1.5 Expression analysis of NRL gene family in rice

To further analyze the expression characteristics of rice NRL family genes in different tissues, we obtained RNA-Seq data of ten different tissues from RGAP (Rice genome annotation project) (http://rice.plantbiology.msu.edu/) and performed hierarchical clustering analysis on the expression data of rice NRL family genes (Figure 4). The data showed that NRL family genes in rice were expressed in different degrees in different rice tissues, and they could be divided into four groups according to the clustering analysis results. There were 7, 13, 4 and 3 genes in the first group to the fourth group, respectively (Figure 4). Genes in the first group, namely *OsNRL7*, *OsNRL8*, *OsNRL14*, *OsNRL16*, *OsNRL17*, *OsNRL19* and *OsNRL21*, were highly expressed in inflorescences and pistils before heading, and the expression levels of *OsNRL7* and *OsNRL14* were also high in seed-5 days after pollination. In the second group, 13 NRL family genes were highly expressed in the inflorescence before heading, and *OsNRL23*, *OsNRL25* and *OsNRL26* were also highly expressed in the leaves. *OsNRL20* was also highly expressed in the inflorescence after heading. In the third group, *OsNRL6*, *OsNRL12*, *OsNRL22* and *OsNRL24* were highly expressed in leaves, while in the fourth group, *OsNRL2*, *OsNRL13* and *OsNRL18* were highly expressed in inflorescence and anther after heading. It is worth noting that some duplication genes, such as *OsNRL22* and *OsNRL24*, are clustered in the same group, suggesting that some WGD or segmental duplication genes have the same or similar expression regulation sequences and have similar expression regulation patterns. Therefore, the expression of duplication genes is further analyzed globally.

Microarray data were used to analyze the expression patterns of WGD or segmental duplication gene pairs and tandem duplication gene pairs of NRL gene family in rice. The lack of partial expression data of *OsNRL1* and *OsNRL15*, *OsNRL27* and *OsNRL9* gene pairs was not included in the analysis (Figure 5). The tandem duplication genes of *OsNRL22* and *OsNRL23*, *OsNRL24* and *OsNRL25* showed different expression patterns. The WGD or segmental duplication genes of *OsNRL4* and *OsNRL12*, *OsNRL22* and *OsNRL24*, *OsNRL26* and *OsNRL8* showed similar expression patterns, respectively.

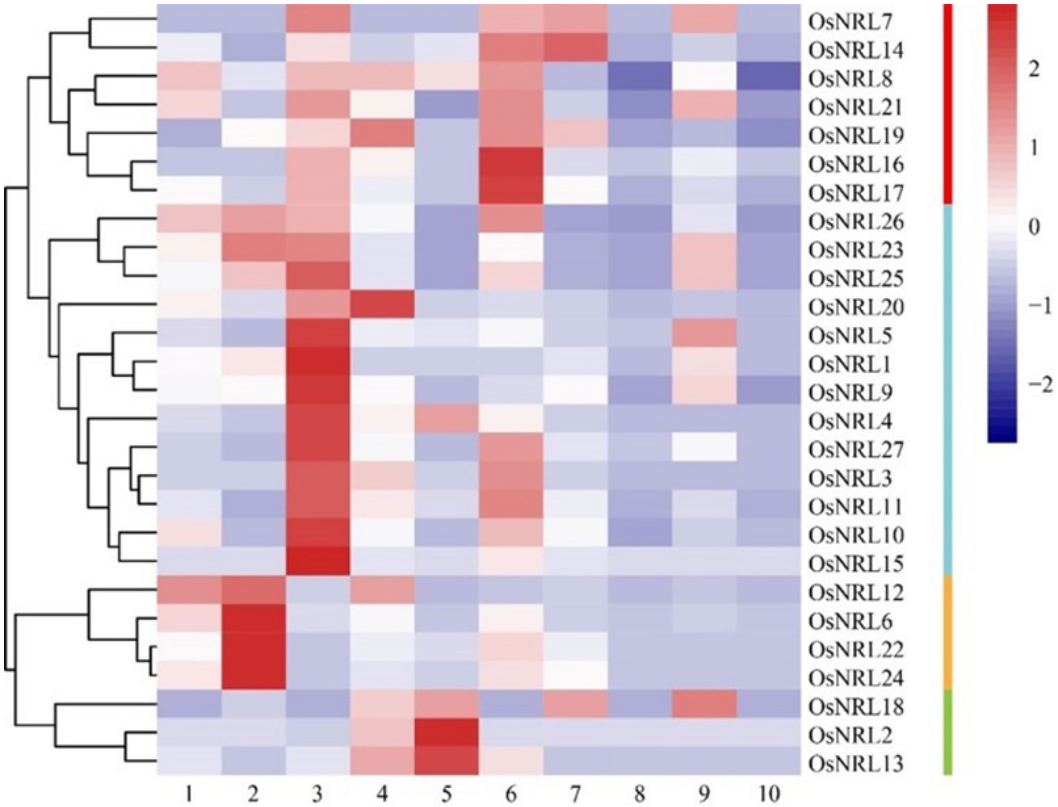

Figure 4 Expression of *OsNRLs* in different tissues and developmental stages

Note: 1: Shoots; 2: Leaves-20 days; 3: Pre-emergence inflorescence; 4: Post-emergence inflorescence; 5: Anther; 6: Pistil; 7: Seed-5 days after pollination; 8: Seed-10 days after pollination; 9: Embryo-25 days after pollination; 10: Endosperm-25 days after pollination

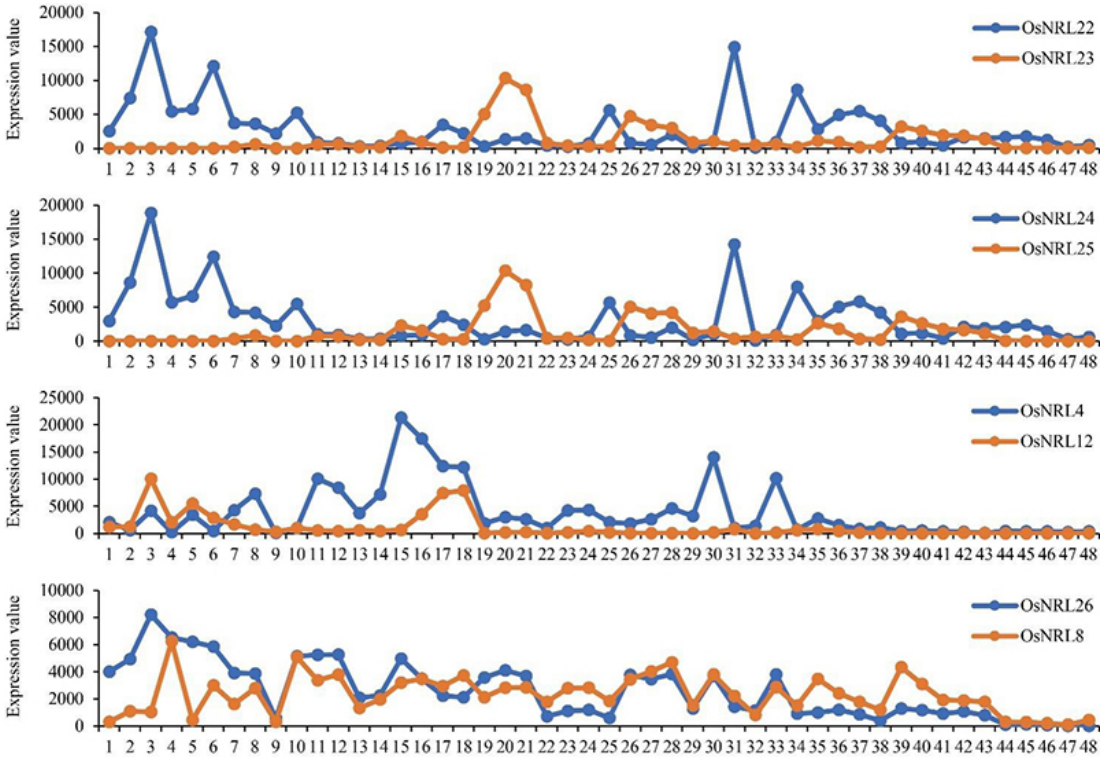

Figure 5 Expression patterns of duplication gene pairs of NRL family in rice

Note: The X-axis represents developmental stages of rice; The Y-axis represents the raw expression values obtained from the microarray data (GSE21396); 1~6: Leaf blade; 7~10: Leaf sheath; 11~14: Root; 15~18: Stem; 19~21: Inflorescence; 22~25: Anther; 26~28: Pistil; 29~31: Lemma; 32~34: Palea; 35~38: Ovary; 39~43: Embryo; 44~48: Endosperm; The matching Information between rice developmental stages and serial numbers obtains from the microarray data (GSE21396)

2 Discussion

NRL (NPH3/RPT2-Like) gene family is plant-specific. The functions of 10 members of 33 gene families in *Arabidopsis thaliana*, including NPH3 and RPT2, have been identified (Christie et al., 2018). It is believed that the NRL gene family is involved in the regulation of auxin-dependent and auxin-independent physiological responses mediated by members (such as phototropinphot1 and phot2) of the VIII subfamily of AGC kinases (Suetsugu et al., 2016, Christie et al., 2018). Only two members of NRL gene family, CPT1 and WIN1, were identified in rice. CPT1 has been proved to be involved in the phototropism of coleoptile and the lateral transport of auxin, while WIN1 is related to auxin transport, but does not participate in the phototropism of coleoptile and root (Haga et al., 2005, Cheng et al., 2017).

In this study, 27 NRL family genes were identified from rice genome by bioinformatics methods, and two NRL family genes, namely *OsNRL1* and *OsNRL15*, were added compared with previous reports (Gingerich et al., 2007). Compared with other rice NRL family genes, *OsNRL1* and *OsNRL15* have fewer exon/intron structures and shorter amino acid sequences, suggesting that they may be pseudogenes. However, unlike the general pseudogenes that are lower than homologous genes, the expression of *OsNRL1* and *OsNRL15* in many tissues is not lower than that of their homologous genes *OsNRL3* (*OsCPT1*) (Zou et al., 2009). Therefore, to determine whether *OsNRL1* and *OsNRL15* are pseudogenes needs to be combined with the results of future gene functional studies.

In this study, 27 rice NRL family genes were unevenly distributed on 12 chromosomes, and there were some duplication gene pairs, that is, 5 WGD or segmental duplication gene pairs and 2 tandem duplication gene pairs. Gene duplication is one of the driving forces of evolution, which is involved in the formation and functional differentiation of new genes, and there may be functional redundancy duplication genes (Panchy et al., 2016). It has been reported that tandem duplication and WGD or segmental replication play a major role in the expansion of a single gene family (Kong et al., 2019). In this study, there were seven gene duplication events in the rice NRL

gene family, indicating that gene duplication played an important role in the gene expansion of the rice NRL gene family, and WGD or segmental duplication played a more important role than tandem duplication. In the phylogenetic relationship analysis, the members of the NRL gene family of seven species, including rice, can be divided into six groups (Christie et al., 2018). *AtNPH3* and *AtRPT2* were in Group VI and Group IV, respectively. The rice homologous gene *OsNRL3* (*OsCPT1*) in the same group with *AtNPH3* was also involved in the phototropism response, and the rice homologous genes *OsNRL22* and *OsNRL24* in the same group with *AtRPT2* were not reported (Haga et al., 2005). *AtNPH3* and *AtRPT2* have seven and six homologous genes in the other five species, respectively, and these homologous genes may be similar to the function of *OsNRL3* (*OsCPT1*), and also participate in phototropism and other physiological responses of plants mediated by phototropism.

In this study, it was found that rice NRL family genes were expressed in many tissues, and some genes were highly expressed in some tissues, such as leaves, inflorescences before and after heading, anthers and pistils. *OsNRL3* (*OsCPT1*) is the first reported rice NRL family gene, which makes the phototropism of rice coleoptile nearly disappear completely after mutation (Haga et al., 2005). *OsNRL3* (*OsCPT1*) was highly expressed in inflorescence before heading and pistils, indicating that *OsNRL3* (*OsCPT1*) may also play a role in rice reproductive development. For *OsNRL7* (*OsWIN1*), on the one hand, the seedlings of its overexpressed strain *WIN1-OX* showed helical phenotype, while the coleoptile of this line and the T-DNA insertion mutant line $WIN1_{Act}$ showed normal phototropism response, indicating that some of the genes represented by *OsNRL7* (*OsWIN1*) in rice NRL family genes may have functional specificity. On the other hand, the seed size of the two lines was smaller than that of wild type seeds. However, *OsNRL7* (*OsWIN1*) was expressed in many rice tissues including seeds and embryos, indicating that *OsNRL7* (*OsWIN1*) may be involved in the regulation of rice seed size (Cheng et al., 2017). Combined with tissue expression analysis and evolutionary relationship analysis, it was found that genes similar to *OsNRL3* (*OsCPT1*), namely *OsNRL1* and *OsNRL15*, and *OsNRL7* (*OsWIN1*), which are *OsNRL6*, *OsNRL7*, *OsNRL22*, *OsNRL23*, *OsNRL24* and *OsNRL25* in rice NRL family genes may play a role in physiological activities such as phototropism and auxin transport in plants. In addition, from the RNA-Seq and microarray expression data, it can be seen that the expression patterns in the WGD or segmental duplication genomes of *OsNRL4* and *OsNRL12*, *OsNRL22* and *OsNRL24*, *OsNRL26* and *OsNRL8* are similar, especially the expression patterns of *OsNRL22* and *OsNRL24* duplication gene pairs are basically the same, and both of them are highly expressed in leaves, indicating that the composition of cis-acting elements in their promoter regions may be similar and their functions as well. The possible functional redundancy should be considered in their research.

3 Materials and Methods

3.1 Identification of NRL gene family members in rice

The MSU 7.0 version of rice genome information was downloaded from the RGAP (http://rice.plantbiology.msu.edu/). HMM (Hidden markov model) (PF03000) of NPH3 domain and HMM (PF00651) of BTB domain were obtained from Pfam (http://pfam.xfam.org/). HMM and HMMER 3.0 software were used to perform the local HMM SEARCH on rice protein sequences to predict NRL family genes, and the E value was less than 1×10^{-10}. At the same time, BLAST 2.9.0 software was used for local BLASTP search, and the E value was less than 1×10^{-10}. Combined with the two results, a total of 30 possible members of rice NRL gene family were identified. The predicted rice NRL family genes were screened by NCBI online tool Batch CD-Search (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi), SMART (http://smart.embl-heidelberg.de/) and Pfam, and 27 NRL genes were identified as rice NRL family by removing genes with predicted domain length less than NPH3 domain or 50% of BTB domain. Gene information of NRL family genes was obtained from RGAP (http://rice.plantbiology.msu.edu/) (Huang et al., 2015; Chen et al., 2017). Plant-mSubP (http://bioinfo.usu.edu/Plant-mSubP/) was used for the prediction of protein subcellular localization.

3.2 Chromosome distribution and gene duplication analysis of NRL gene family in rice

The position information of genes on chromosomes in rice was obtained by RGAP (http://rice.plantbiology.msu.edu/), and the chromosome distribution information of NRL gene family in rice was screened out. MCScanX software was used to analyze the gene duplication events in rice, and the gene

duplication pairs in rice NRL gene family were screened and extracted. The chromosome distribution and gene duplication diagram of rice NRL gene were drawn by TBtools software (Chen et al., 2020). Ka/Ks (Nonsynonymous substitution rates/Synonymous substitution rates) was calculated with the help of DnaSP 6.0 software. And the selection pressure of duplication gene pairs was analyzed.

3.3 Phylogenetic analysis of NRL gene family

Identification of NRL gene family members for the obtained protein sequences of dicotyledons (*A. thaliana*, grape, tomato) and monocotyledons (maize, barley, Brachypodium) was performed with the help of online database Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). The neighbor-joining method was used to analyze the phylogenetic relationship of NRL gene family members by MEGA 7.0. In which, the model was set to Poisson model, Bootstrap was set to 1000, and the phylogenetic tree was drawn by online tool Evolview v2 (http://nar.oxfordjournals.org/content/44/W1/W236) at the same time. Using the IQ-tree component of TBtools software, the phylogenetic tree of rice NRL gene family was constructed by maximum likelihood method, with Bootstrap set to 1000 (Chen et al., 2020).

3.4 Analysis of NRL family gene structure and conserved motif in rice

Structure analysis of NRL gene family members in rice was performed with the help of the online website CSDS 2.0 (http://gsds.cbi.pku.edu.cn/). Conserved motif analysis of NRL gene family members in rice was performed with the help of the online website MEME (http://meme-suite.org/tools/meme), with default parameter and 15 motifs. And the conserved motif diagram was completed by software TBtools (Chen et al., 2020).

3.5 Expression analysis of NRL gene family in rice

The expression data of different rice tissues were downloaded from the RGAP (http://rice.plantbiology.msu.edu/). R language pheatmap was used to standardize the expression data of NRL family genes in rice by Z-score method, and hierarchical clustering analysis was carried out. The temporal and spatial expression microarray data (GSE21396) of various tissues/organs in rice during the whole growth process in the field obtained by NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) were used to analyze the expression pattern of duplication gene pairs in rice NRL family.

Authors' Contributions

HYY was the experimental designer and executor of this research. HYY completed the data analysis and manuscript writing. SLL and WDX participated in the experimental design, results analysis, manuscript writing and revision. ZDB and CWG conceived of the study, guided the experimental design, data analysis, writing and revision. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by the National Key R & D Program (2016YFD0100804), Agricultural Science and Technology Independent Innovation Fund of Jiangsu Province (JASTIF, CX(18)1001) and Science Foundation for Post Doctorate of China (Grant 2017M621451).

References

Chen C.J., Chen H., Zhang Y., Thomas H.R., Frank M.H., He Y.H., and Xia R., 2020, TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data, Mol. Plant, 13(8): 1194-1202
https://doi.org/10.1016/j.molp.2020.06.009
PMid:32585190

Chen X., Shi L., Zhu L., Shi J.X., and Xu J., 2017, Molecular evolution characteristics and expression pattern analysis of the heat shock protein 70(HSP70) gene superfamily in plant, Jiyinzuxue Yu Yingyong Shengwuxue (Genomics and Applied Biology), 36(10): 4282-4294

Cheng M.L., Lo S.F., Hsiao A.S., Hong Y.F., Yu S.M., and Ho T.D., 2017, Ectopic expression of *WINDING 1* leads to asymmetrical distribution of auxin and a spiral phenotype in rice, Plant & Cell Physiology, 58(9): 1494-1506
https://doi.org/10.1093/pcp/pcx088
PMid:28922746

Cheng Y.F., Qin G.J., Dai X.H., and Zhao Y.D., 2008, NPY genes and AGC kinases define two key steps in auxin-mediated organogenesis in Arabidopsis, Proc. Natl. Acad. Sci. USA, 105(52): 21017-21022
https://doi.org/10.1073/pnas.0809761106
PMid:19075219 PMCid:PMC2634868

Christie J.M., Blackwood L., Petersen J., and Sullivan S., 2015, Plant flavoprotein photoreceptors, Plant & Cell Physiology, 56(3): 401-413
https://doi.org/10.1093/pcp/pcu196
PMid:25516569 PMCid:PMC4357641

Christie J.M., Suetsugu N., Sullivan S., and Wada M., 2018, Shining light on the function of NPH3/RPT2-Like proteins in phototropin signaling, Plant Physiology, 176(2): 1015-1024
https://doi.org/10.1104/pp.17.00835
PMid:28720608 PMCid:PMC5813532

Gingerich D.J., Hanada K., Shiu S.H., and Vierstra R.D., 2007, Large-scale, lineage-specific expansion of a bric-a-brac/tramtrack/broad complex ubiquitin-ligase gene family in rice, Plant Cell, 19(8): 2329-2348
https://doi.org/10.1105/tpc.107.051300
PMid:17720868 PMCid:PMC2002615

Haga K., Takano M., Neumann R., and Iino M., 2005, The Rice *COLEOPTILE PHOTOTROPISM1* gene encoding an ortholog of Arabidopsis NPH3 is required for phototropism of coleoptiles and lateral translocation of auxin, Plant Cell, 17(1): 103-115
https://doi.org/10.1105/tpc.104.028357
PMid:15598797 PMCid:PMC544493

Holland J.J., Roberts D., and Liscum E., 2009, Understanding phototropism: from Darwin to today, J. Exp. Bot., 60(7): 1969-1978
https://doi.org/10.1093/jxb/erp113
PMid:19357428

Huang Y., Zhu L., Luo R.H., Zhang D.B., and Xu J., 2015, Evolutionary and expression profile analysis of the non-specific lipid transfer proteins genes superfamily in *Arabidopsis thaliana* and rice (*Oryza sativa*), Jiyinzuxue Yu Yingyong Shengwuxue (Genomics and Applied Biology), 34(7): 1510-1521

Inada S., Ohgishi M., Mayama T., Okada K., and Sakai T., 2004, RPT2 is a signal transducer involved in phototropic response and stomatal opening by association with phototropin 1 in *Arabidopsis thaliana*, Plant Cell, 16(4): 887-896
https://doi.org/10.1105/tpc.019901
PMid:15031408 PMCid:PMC412863

Kong W.L., An B.G., Zhang Y., Yang J., Li S.M., Sun T., and Li Y.S., 2019, Sugar transporter proteins (STPs) in gramineae crops: comparative analysis, phylogeny, evolution, and expression profiling, Cells, 8(6): 560
https://doi.org/10.3390/cells8060560
PMid:31181814 PMCid:PMC6628381

Lalanne E., Michaelidis C., Moore J.M., Gagliano W., Johnson A., Patel R., Howden R., Vielle-Calzada J.P., Grossniklaus U., and Twell D., 2004, Analysis of transposon insertion mutants highlights the diversity of mechanisms underlying male progamic development in Arabidopsis, Genetics, 167(4): 1975-1986
https://doi.org/10.1534/genetics.104.030270
PMid:15342534 PMCid:PMC1471024

Li J.H., Su X.X., Wang Y.L., Yang W., Pan Y., Su C.G., and Zhang X.G., 2018, Genome-wide identification and expression analysis of the BTB domain-containing protein gene family in tomato, Genes & Genomics, 40(1): 1-15
https://doi.org/10.1007/s13258-017-0604-x
PMid:29892895

Li Y.T., Dai X.H., Cheng Y.F., and Zhao Y.D., 2011, NPY genes play an essential role in root gravitropic responses in Arabidopsis, Mol. Plant, 4(1): 171-179
https://doi.org/10.1093/mp/ssq052
PMid:20833732 PMCid:PMC3026118

Liscum E., Askinosie S.K., Leuchtman D.L., Morrow J., Willenburg K.T., and Coats D.R., 2014, Phototropism: growing towards an understanding of plant movement, Plant Cell, 26(1): 38-55
https://doi.org/10.1105/tpc.113.119727
PMid:24481074 PMCid:PMC3963583

Motchoulski A., and Liscum E., 1999, *Arabidopsis* NPH3: A NPH1 photoreceptor-interacting protein essential for phototropism, Science, 286(5441): 961-964
https://doi.org/10.1126/science.286.5441.961
PMid:10542152

Panchy N., Lehti-Shiu M., and Shiu S.H., 2016, Evolution of gene duplication in plants, Plant Physiology, 171(4): 2294-2316
https://doi.org/10.1104/pp.16.00523
PMid:27288366 PMCid:PMC4972278

Pedmale U.V., Celaya R.B., and Liscum E., 2010, Phototropism: mechanism and outcomes, Arabidopsis Book, 8: e0125
https://doi.org/10.1199/tab.0125
PMid:22303252 PMCid:PMC3244944

Petricka J.J., Clay N.K., and Nelson T.M., 2008, Vein patterning screens and the *defectively organized tributaries* mutants in *Arabidopsis thaliana*, Plant Journal, 56(2): 251-263
https://doi.org/10.1111/j.1365-313X.2008.03595.x
PMid:18643975

Roberts D., Pedmale U.V., Morrow J., Sachdev S., Lechner E., Tang X.B., Zheng N., Hannink M., Genschik P., and Liscum E., 2011, Modulation of phototropic responsiveness in *Arabidopsis* through ubiquitination of phototropin 1 by the CUL3-Ring E3 ubiquitin ligase $CRL3^{NPH3}$, Plant Cell, 23(10): 3627-3640
https://doi.org/10.1105/tpc.111.087999
PMid:21990941 PMCid:PMC3229139

Sakai T., Wada T., Ishiguro S., and Okada K., 2000, RPT2. A signal transducer of the phototropic response in Arabidopsis, Plant Cell, 12(2): 225-236
https://doi.org/10.1105/tpc.12.2.225
PMid:10662859 PMCid:PMC139760

Suetsugu N., Takemiya A., Kong S.G., Higa T., Komatsu A., Shimazaki K., Kohchi T., and Wada M., 2016, RPT2/NCH1 subfamily of NPH3-like proteins is essential for the chloroplast accumulation response in land plants, Proc. Natl. Acad. Sci. USA, 113(37): 10424-10429
https://doi.org/10.1073/pnas.1602151113
PMid:27578868 PMCid:PMC5027436

Yang L., McLellan H., Naqvi S., He Q., Boevink P.C., Armstrong M., Giuliani L.M., Zhang W., Tian Z.D., Zhan J.S., Gilroy E.M., and Birch P.R., 2016, Potato NPH3/RPT2-Like protein StNRL1, targeted by a *Phytophthora infestans* RXLR effector, is a susceptibility factor, Plant Physiology, 171(1): 645-657
https://doi.org/10.1104/pp.16.00178
PMid:26966171 PMCid:PMC4854710

Zhang L., Du L.Q., Shen C.J., Yang Y.J., and Poovaiah B.W., 2014, Regulation of plant immunity through ubiquitin-mediated modulation of Ca^{2+}-calmodulin-AtSR1/CAMTA3 signaling, Plant Journal, 78(2): 269-281
https://doi.org/10.1111/tpj.12473
PMid:24528504

Zou C., Lehti-Shiu M.D., Thibaud-Nissen F., Prakash T., Buell C.R., and Shiu S.H., 2009, Evolutionary and expression signatures of pseudogenes in Arabidopsis and rice, Plant Physiology, 151(1): 3-15
https://doi.org/10.1104/pp.109.140632
PMid:19641029 PMCid:PMC2736005