RESEARCH ARTICLE

Identification of transcription factor binding sites on promoter of *RNA dependent RNA polymerases (RDRs)* and interacting partners of RDR proteins through in silico analysis

Ved Prakash¹ · Supriya Chakraborty¹

Received: 10 May 2018 / Revised: 20 December 2018 / Accepted: 25 March 2019 / Published online: 6 May 2019
© Prof. H.S. Srivastava Foundation for Science and Society 2019

Abstract RNA silencing phenomenon in plants provides resistance to various pathogens and also, it maintains genome integrity. The process of RNA silencing is regulated by diverse proteins, among which RNA dependent RNA polymerases (RDRs) are very crucial for the amplification of small RNAs (sRNAs). Out of various RDR proteins present in plants, role of RDR1, RDR2 and RDR6 for providing resistance against various biotic stresses have been well documented. In contrast, very few information is available regarding the role of RDR3, RDR4 and RDR5 proteins in plant biology and stress response. Furthermore, the regulation of RDRs is not yet known. Here, we have carried out in silico studies for identification of the transcription factor (TF) binding sites on the promoter of *RDR1-6* genes of various plant species. Among the TFs predicted to bind on the promoter of *RDRs*, MYB44, AS1/AS2, WRKY1 are the major one. Furthermore, putative interacting protein partners of RDRs proteins of tomato and rice were also predicted by STRING database which suggests that DCL (Dicer-like) proteins are strong candidate proteins as the interacting partners of RDRs. The knowledge of regulation of RDRs and its interacting protein partners might help in developing resistant plants to biotic stresses.

Keywords Promoter · RDRs · Interacting partners · In silico · Small RNA · Gene regulation

Introduction

Plants protect themselves from the invading viruses and viroids through RNA silencing pathways. In addition, RNA silencing pathway also silences transposons, repetitive elements and endogenous genes, either transcriptionally or post-transcriptionally in a sequence-specific manner. RNA dependent RNA polymerases (RDRs) are one of the crucial proteins involved in RNA interference (RNAi) pathways which are required for amplification of the silencing signals (Devert et al. 2015; Vaucheret 2006; Prakash et al. 2017). The model plant *Arabidopsis thaliana* (*A. thaliana*) genome encodes for six RDR proteins, viz., RDR1-6 with distinct functions (Wassenegger and Krczal 2006). Among these RDRs, RDR1, RDR2 and RDR6 belong to the RDR α family (containing a typical DLDGD signature motif at the catalytic site) while RDR3, RDR4 and RDR5 belong to the RDR γ family (with a conserved DFDGD amino acid sequence at the active site). Several reports are available implicating the precise role of RDRs of the α family, however, the function of members of the RDR γ family remains elusive (Butterbach et al. 2014; Verlaan et al. 2013; Willmann et al. 2011).

Out of the six *Arabidopsis* RDRs, RDR1, RDR2 and RDR6 are the major antiviral proteins which restrict virus locally as well as systemically (Wang et al. 2010). Downregulation of *RDR1* and *RDR6* is linked to susceptibility of plants to certain viruses (Dalmay et al. 2000; Mourrain et al. 2000; Xie et al. 2001). *Nicotiana benthamiana* (*N. benthamiana*) harbours a natural mutant form of *RDR1* (Qu et al. 2005; Schwach et al. 2005).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12298-019-00660-w>) contains supplementary material, which is available to authorized users.

✉ Supriya Chakraborty
supriyachakrasls@yahoo.com

¹ Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India

Furthermore, downregulation of expression of the *NbRDR6* transcripts affects the efficiency of RNA silencing. A report suggests cross-talk between RDR1 and RDR6, since inhibition of *RDR1* transcripts accumulation also affects the expression of *RDR6* (Rakhshandehroo et al. 2009).

Several studies suggested induction of *RDR1* by various phytohormones and viruses. Xie et al. (2001) demonstrated increased level of *RDR1* transcripts following either application of salicylic acid (SA) or after an infection with *Tobacco mosaic virus* (TMV) in *Nicotiana tabacum* (*N. tabacum*), although, the mechanism of SA-mediated *RDR1* induction is not yet known. Interestingly, the RDR1 mutant *N. tabacum* showed susceptibility to both TMV and *Potato Virus Y* (PVY), however, resistance to these viruses was induced in these plants by exogenous application of SA (Xie et al. 2001). SA inducible *RDR1* genes have also been found in *A. thaliana*, *N. benthamiana* and *Cucumis sativus* (Leibman et al. 2018; Yang et al. 2004; Yu et al. 2003). In addition to SA, other phytohormones such as jasmonic acid (JA), abscisic acid (ABA) and ethylene (ET) are also known to induce the expression of *RDR1* gene, however, the mechanism of phytohormone mediated regulation of *RDR1* is not yet understood (Hunter et al. 2013; Pandey and Baldwin 2007; Pandey et al. 2008b; Prakash et al. 2017).

RDR2 is involved in the RNA directed DNA methylation (RdDM) pathway wherein cytosine bases of the transposons, repetitive sequences, retro-elements and telomeres are methylated leading to transcriptional gene silencing (TGS) (Willmann et al. 2011; Prakash et al. 2017). Antiviral role of RDR2 has been demonstrated against *Beet curly top virus* (BCTV) (Raja et al. 2008). RDR2, Dicer-like 3 (DCL3) and Argonaute 4 (AGO4), with the help of other proteins, produce 24 nucleotide (nt) long heterochromatic small-interfering RNA (hcsiRNA), which are the most abundant class of small-interfering RNAs (siRNAs) in plants (Vrbsky et al. 2010).

The RDR6 protein generates long double-stranded RNAs (dsRNAs) from siRNAs predominantly of 21 nt size class. In addition, Suppressor of Gene Silencing 3 (SGS3) protein is also prerequisite for dsRNA synthesis, since SGS3 interacts with RDR6 and also stabilises the RNA template for RDR6-mediated dsRNA synthesis (Lam et al. 2012; Mourrain et al. 2000). However, SGS3 is not required for RDR1 mediated dsRNA biogenesis (Cao et al. 2016). RDR6 is required for intracellular post-transcriptional gene silencing pathway (PTGS) as well as systemic spread of PTGS (Melnik et al. 2011; Devert et al. 2015).

In *Nicotiana attenuata* (*N. attenuata*), RDR3 is required for competitive growth in natural environment (Pandey et al. 2008a). In addition, expression of *RDR3* increases after *Cucumber mosaic virus* (CMV) infection (Shao and Lu 2014). Furthermore, *Ty1* and *Ty3* genes of tomato are allelic and belong to RDR γ family, and provide resistance

against geminivirus infection by participating in TGS pathway but not in PTGS (Butterbach et al. 2014; Verlaan et al. 2013).

Often, the expression of a gene is affected by various environmental, biological and chemical factors which initiate cellular signal transduction pathways. Such signal transduction pathways facilitate recruitment and interaction of various DNA binding proteins, called transcription factors (TFs), on the promoter region of target genes. This interaction in turn may either induce or repress gene expression. There are several TF families, for example, MYB, NAC and WRKY which are either directly or indirectly involved in various developmental and stress responses in plants (Guo et al. 2017; Mao et al. 2012; Mare et al. 2004; Oh et al. 2011).

AtMYB44 binds to the promoter of *WRKY70* (Shim et al. 2013). *WRKY70* performs crucial function during SA and JA antagonism. *WRKY70* transcription is induced by SA but suppressed by JA. Another TF, *A. thaliana* Asymmetric leaves 1 (AS1) protein (member of MYB like family transcription factors) in association with Asymmetric leaves 2 (AS2) protein function in epigenetic silencing during early leaf development (Guo et al. 2008; Iwasaki et al. 2013; Lodha et al. 2013; Phelps-Durr et al. 2005). A recent study has suggested the role of AS1/AS2 proteins in DNA methylation, however, the mechanism of AS1/AS2-mediated DNA methylation is not known yet (Nishimura 2018).

TFs from the NAC family (for example GmNAC81 and GmNAC30) regulates many biological processes and protects plants during various stresses through induction of several genes during stress conditions (Seki et al. 2003). It is also known that GmNAC81 enters into the nucleus and binds with GmNAC30, and the resulting heterodimer then either acts as transcriptional activator or repressor of various genes. GmNAC81/GmNAC30 activates Caspase-1 like vacuolar processing enzyme (VPE) gene, which is involved in program cell death (PCD), by binding to the promoter of *VPE* (Mendes et al. 2013).

Several studies indicate that *WRKY1* (a W-Box transcription factor) acts as a negative regulator of plant defense (Marchive et al. 2007; Mzid et al. 2007; Oh et al. 2008). In contrast, recent reports suggest that *WRKY1* acts as a positive regulator of defence in *Phytophthora infestans* (*P. infestans*) and *P. nicotianae* infected tomato and tobacco, respectively (Li et al. 2015a, b). *WRKY1* localizes to the nucleus where it binds to W box domain containing motif (Qiao et al. 2016).

Plant RDRs might interact with proteins either to maintain/amplify RNAi signals in cells or to regulate plant development. For example, NRPD1A interacts with RDR2 and, SGS3 interacts with RDR6 (Lam et al. 2012; Law et al. 2011). Despite several studies, information regarding

the interacting partners of RDR1, RDR3, RDR4 and RDR5 are not available.

Therefore, the present study was aimed to understand the role and regulation of *RDR* genes of various plant species. For this purpose, we performed in silico analyses to find out the TF binding sites on the putative promoters of *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6* of various plant species and probable interacting partners of RDR proteins of tomato and rice. Analysing the promoter sequences for TF binding sites might help in understanding the regulation of RDRs and their role in plant stress response.

Methods

Retrieval of mRNA sequences of plant *RDRs* (*RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6*)

Retrieval of mRNA sequences of *A. thaliana* *RDR1* (*AtRDR1*) (NM_101348.4), *Chenopodium quinoa* *RDR1* (*CqRDR1*) (LC146408.1), *Capsicum annuum* *RDR1* (*CaRDR1*) (XM_016691689.1), *Glycine max* *RDR1* (*GmRDR1*) (XM_014778083.2), *Gossypium hirsutum* *RDR1* (*GhRDR1*) (EF571899.1), *Helianthus annuus* *RDR1* (*HaRDR1*) (XM_022121758.1), *Medicago truncatula* *RDR1* (*MtRDR1*) (XM_013611146.2), *N. attenuata* *RDR1* (*NaRDR1*) (XM_019406857.1), *N. tabacum* *RDR1* (*NtRDR1*) (XM_016610634.1), *Oryza sativa* *RDR1* (*OsRDR1*) (XM_015771542.2), *Papaver somniferum* *RDR1* (*PsRDR1*) (XM_026547049.1), *Arachis hypogaea* *RDR1* (*AhRDR1*) (XM_025817081.1), *Populus trichocarpa* *RDR1* (*PtRDR1*) (XM_002311500.3), *Solanum tuberosum* *RDR1* (*StRDR1*) (XM_006366047.2), *Solanum lycopersicum* *RDR1* (*SIRDR1*) (NM_001247390.2), *Vitis vinifera* *RDR1* (*VvRDR1*) (XM_002284878.3), *AtRDR2* (NM_117183.3), *CqRDR2* (XM_021877960.1), *CaRDR2* (XM_016707984.1), *GmRDR2* (XM_003550696.4), *GhRDR2* (XM_016873182.1), *HaRDR2* (XM_022162556.1), *MtRDR2* (XM_022162556.1), *NaRDR2* (XM_019388169.1), *NtRDR2* (XM_016592572.1), *OsRDR2* (XM_015781398.2), *PsRDR2* (XM_026567068.1), *AhRDR2* (XM_025773370.1), *PtRDR2* (XM_002321546.3), *StRDR2* (XM_006344978.2), *SIRDR2* (XM_004236072.4), *VvRDR2* (XM_002280063.3), *Arabidopsis lyrata* *RDR3* (*AIRDR3*) (XM_021031768.1), *AhRDR3* (XM_025800335.1), *Beta vulgaris* *RDR3* (*BvRDR3*) (XM_010690005.2), *CqRDR3* (XM_021885732.1), *CaRDR3* (XM_016685137.1), *GmRDR3* (XM_006572877.3), *HaRDR3* (XM_022162235.1), *NtRDR3* (XM_016656409.1), *OsRDR3* (XM_015772870.2), *Physcomitrella patens* *RDR3* (*PpRDR3*) (XM_024522424.1), *Sorghum bicolor* *RDR3* (*SbRDR3*) (XM_002454957.2), *SIRDR3* (XM_010323

869.3), *AIRDR4* (XM_021031773.1), *OsRDR4* (XM_015768996.2), *PpRDR4* (XM_024520751.1), *Selaginella moellendorffii* *RDR4* (*SmRDR4*) (XM_024678182.1), *AIRDR5* (XM_021031772.1), *CqRDR5* (XM_021867874.1), *CaRDR5* (XM_016684957.1), *GhRDR5* (XM_016873588.1), *HaRDR5* (XM_022162237.1), *MtRDR5* (XM_013597709.2), *NaRDR5* (XM_019384318.1), *NtRDR5* (XM_016594094.1), *PsRDR5* (XM_026593126.1), *AhRDR5* (XM_025825937.1), *PtRDR5* (XM_024603284.1), *StRDR5* (XM_006353425.2), *SIRDR5* (XR_002026414.2), *VvRDR5* (XM_010657967.2), *AhRDR6* (XM_025817060.1), *AtRDR6* (NM_114810.3), *CqRDR6* (NM_114810.3), *CaRDR6* (XM_016716534.1), *GmRDR6* (XM_003522570.4), *GhRDR6* (NM_001327683.1), *HaRDR6* (XM_022174005.1), *MtRDR6* (XM_003603361.3), *NaRDR6* (XM_019404979.1), *NtRDR6* (FJ966891.1), *PsRDR6* (XM_026586016.1), *PtRDR6* (XM_002324259.3), *StRDR6* (XM_006346723.2), *SIRDR6* (XM_010320819.3), *VvRDR6* (XM_010650358.2) and *OsRDR6* (XM_015766751.2) were carried out from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). Amino acids and nucleotide sequences of the above mentioned *RDR* genes were used to construct the phylogenetic tree.

Multiple sequence alignment of RDRs and phylogeny analyses

Alignment of amino acid sequences of *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6* was performed by the Muscle function of the MEGA version 7 software with the default settings (Kumar et al. 2016). Bootstrap (1000 replicates), Jones-Taylor-Thornton (JTT) model, uniform rates among sites, and complete deletion were the parameters considered for the phylogenetic tree construction. Further, the coding sequences of respective RDRs were considered for generation of a DNA phylogenetic tree. DNA sequences were also aligned by Muscle's default settings and the maximum likelihood tree was constructed with the following parameters: bootstrap (1000 replicates), Tamura-Nei model, uniform rates among sites, and complete deletion.

Retrieval of promoter sequences

Promoters of all the *RDRs* from distinct plant species (16 for *RDR1*, 16 for *RDR2*, 16 for *RDR6*, 12 for *RDR3*, 4 for *RDR4* and 14 for *RDR5*) were identified from the NCBI database of respective plant species. Promoter sequences were retrieved by carrying out BLAST search using the specific organism's *RDR* mRNA sequences present in the genome. After locating mRNA start site in the chromosome, 1000 nucleotide bases upstream from the mRNA start site were selected as the putative promoter region for further analyses. For validation of promoter analysis,

Fig. 1 Phylogenetic analysis of RDR1, RDR2, RDR3, RDR4, RDR5 and RDR6 proteins of various plant species. Maximum likelihood tree was generated by MEGA 7 with bootstrap method as the test of phylogeny (1000 replications)

previously characterized promoter of the *SR2* gene (accession: DQ109992) from *Phaseolus vulgaris* (*PvSR2*) was used (Qi et al. 2007).

Prediction of TF binding site(s)

DNA sequences representing putative promoters (1.0 kb) were used for identification of nucleotide sequences for TF binding sites using ‘MatInspector’ and ‘Common TFs’ tool from Genomatix software suite (Cartharius et al. 2005).

Protein interactome network

Prediction of probable protein interacting partners of SIRDR1, SIRDR2, SIRDR3, SIRDR5, SIRDR6, OsRDR1, OsRDR2, OsRDR3, OsRDR4 and OsRDR6 were performed with the help of STRING database (version 10.5) (Szklarczyk et al. 2015). Minimum required interaction score was medium confidence (0.400) and the maximum number of interactors to display was set as 10.

Result

Phylogeny of plant RNA dependent RNA polymerases

Protein sequences and nucleotide sequences of the coding region of *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6* of a total of 21 plant species were retrieved from the NCBI database and subjected to phylogenetic analyses using MEGA7. The tree constructed with both amino acids and nucleotide sequences suggest that RDR α and RDR γ are phylogenetically distinct. In the phylogenetic dendrogram, RDR1, RDR2 and RDR6 clades are distinctly placed (Fig. 1; Fig S1). However, members of the RDR γ family formed a separate clade and no dedicated tree branches could be observed for either RDR3 or RDR4 or RDR5. Results of phylogenetic analyses suggest that unlike *RDR3*, *RDR4* and *RDR5*, members of the RDR α family viz., *RDR1*, *RDR2* and *RDR6* genes are more diverged and hence, formed separate clades (Fig. 1; Fig S1).

Analysis of TF binding sites on the promoters of *RDR1* of various plant species

For identification of potential TF binding sites, promoter analysis of the *RDR* genes was carried out and the results

Fig. 2 Schematic representation of transcription factor binding sites on the putative promoters of *StRDR1*, *SIRD1*, *CaRDR1*, *NiRDR1*, *NaRDR1*, *HaRDR1*, *GmRDR1*, *VvRDR1*, *PtRDR1*, *GhRDR1*, *CqRDR1*, *AtRDR1*, *MtRDR1*, *AhRDR1*, *PsRDR1* and *OsRDR1* based

on the stringent conditions, i.e., core and matrix similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('P\$' denotes plants and 'O\$' denotes general core promoter elements)

are presented below. For convenience, the analyses of the *RDRα* family are mentioned first followed by the *RDRγ* family.

RDR1 promoters of *A. thaliana*, *C. quinoa*, *C. annuum*, *G. hirsutum*, *H. annuus*, *N. attenuata*, *P. somniferum*, *P. trichocarpa*, *S. tuberosum*, *V. vinifera*, *N. tabacum*, *S. lycopersicum*, *A. hypogaea*, *M. truncatula*, *G. max* and *O. sativa* were used in the analysis. Analysis by MatInspector suggested a total of 118 matches for the binding sites of various TF matrix families on sequences of 16 *RDR1* promoters (Fig. S2). The detailed list of various TF matrix families predicted to bind on the putative promoters is presented in Table S1. This result was obtained based on the default settings of MatInspector, i.e., core similarity 0.75 and optimized matrix similarity. Results indicated presence of motifs of 17 different matrix families on the

promoter region of *RDR1* of each plant species used in the study (Table S1).

Furthermore, to narrow down the number of TF binding sites present in these promoter sequences, the threshold parameters, i.e., core similarity and matrix similarity were increased to 1.0 (stringent settings). As a result, binding sites for 22 different matrix families are found on all the promoters (Fig. 2, Table S1). The binding site of matrix family P\$MYBL (MYB like protein) ('P\$' denotes plants promoter elements) was predicted to be present on 10 *RDR1* promoters, i.e., *StRDR1*, *CqRDR1*, *AtRDR1*, *AhRDR1*, *PsRDR1*, *SIRD1*, *CaRDR1*, *HaRDR1*, *GmRDR1* and *VvRDR1* promoters (Table 1). AtMYB44 and AS1/AS2 repressor binding motifs are members of the P\$MYBL matrix family. The binding sites for P\$PCDR (factors involved in programmed cell death response) matrix family

Table 1 List of transcription factors matrix families commonly occurring on the promoters of different *RDRs*

TF matrix family	RDR1	RDR2	RDR3	RDR4	RDR5	RDR6
P\$MYBL	<i>Ah, At, Ca, Cq, Gm, Ha, Ps, Sl, St, Vv</i>	<i>Ca, Gh, Gm, Ha Na, Os, Pt Sl, St</i>	<i>Bv, Ha, Gm, Pp, Sl,</i>	<i>Al, Sm</i>	<i>Al, Ca, Cq, Gh, Mt, Na, Ps, Pt, Sl, St, Vv,</i>	<i>At, Ca, Cq, Gh, Gm, Ha, Mt, Na, Os, Ps, Pt, Sl,</i>
P\$PCDR	<i>Ah, Ca, Mt, Na, Nt, Os, Ps, Vv</i>	<i>At, Gh, Gm, Ha, Na, Nt, Ps, Pt, St, Vv</i>	<i>Al, Ca, Cq, Os, Sl</i>	<i>Al, Os, Pp, Sm</i>	<i>Al, Cq, Ha, Mt, Na, Nt, Pt, Vv</i>	<i>At, Ca, Gh, Ha, Mt, Na, Pt, Vv</i>
P\$WBXF	<i>At, Ha, Os</i>	<i>Ah, At, Gh, Ha, Mt, Na, Ps</i>	<i>Al, Cq, Nt, Pp, Sl</i>	<i>Os, Pp</i>	<i>Al, Na, Nt, Ps, Sl, St</i>	<i>At, St</i>
P\$DOFF	<i>Ah, Ca, Gm, Na, St</i>	<i>Ha</i>	–	<i>Os</i>	<i>Cq, Ha, Ps, Pt, St, Vv</i>	<i>Gh</i>
P\$LICM	<i>At, St</i>	–	<i>Os, Pp</i>	<i>Os, Pp</i>	<i>Ah, Ha, Na, Ps, St</i>	–
P\$MYBS	<i>Os, Ah</i>	–	<i>Sl</i>	–	<i>Ah, Pt</i>	–
P\$AHLF	<i>St</i>	<i>Pt, Sl, St</i>	<i>Al, Gm</i>	–	<i>Ah, Gh</i>	<i>Ca, Pt</i>
P\$C3HF	<i>Gm</i>	–	–	<i>Os</i>	<i>Ps</i>	–
P\$CCAF	<i>Ca</i>	–	<i>Sl</i>	–	–	<i>St</i>
P\$NACF	<i>Cq</i>	–	–	<i>Al</i>	–	<i>Os, Na</i>
P\$SBPD	<i>Sl</i>	<i>Ca</i>	–	–	–	–
P\$CAAT	<i>Cq, Gh, Ps</i>	<i>Ca, Mt</i>	<i>Bv, Cq, Nt, Sb</i>	–	–	<i>Ah, Ca</i>
P\$EPFF	–	<i>Ah</i>	<i>Ah</i>	–	–	–
P\$ABRE	–	–	–	<i>Os</i>	<i>St</i>	–
P\$SURE	–	–	–	–	<i>Ca</i>	<i>Ca, Vv</i>

Transcription factors (TFs) occurring on at least two or more plant RDR promoters are indicated. “–” indicates TF not found on promoters of respective *RDRs*. (‘P\$’ denotes plants and ‘O\$’ denotes general core promoter elements)

Ah = *Arachis hypogaea*, *Al* = *Arabidopsis lyrata*, *At* = *Arabidopsis thaliana*, *Bv* = *Beta vulgaris*, *Ca* = *Capsicum annuum*, *Cq* = *Chenopodium quinoa*, *Gh* = *Gossypium hirsutum*, *Gm* = *Glycine max*, *Ha* = *Helianthus annuus*, *Mt* = *Medicago truncatula*, *Na* = *Nicotiana attenuata*, *Nt* = *Nicotiana tabacum*, *Os* = *Oryza sativa*, *Pp* = *Physcomitrella patens*, *Ps* = *Papaver somniferum*, *Pt* = *Populus trichocarpa*, *Sb* = *Sorghum bicolor*, *Sl* = *Solanum lycopersicum*, *Sm* = *Selaginella moellendorffii*, *St* = *Solanum tuberosum*, *Vv* = *Vitis vinifera*

are present on *MtRDR1*, *AhRDR1*, *PsRDR1*, *OsRDR1*, *CaRDR1*, *NtRDR1*, *NaRDR1* and *VvRDR1* (Fig. 2, Table 1, Table S1). The heterodimer of NAC-domain transcription factors, GmNAC30 and GmNAC81 are members of the P\$PCDR matrix family. In addition, the *RDR1* promoters of *A. thaliana*, *O. sativa* and *H. annuus* possess the binding site for TFs belonging to P\$WBXF (W box family) matrix family. Zinc-dependent activator protein-1 and WRKY 1 are members of the P\$WBXF matrix family. *O. sativa* leaf and tiller angle increased controller (OsLIC) (member of P\$LICM), DNA-binding One Zinc Finger (DOF) (member of P\$DOFF matrix family) and AT-hook motif nuclear-localized (AHL) (members of P\$AHLF matrix family) TFs were also predicted to bind on *RDR1* promoters of several plants species. Interestingly, only *GhRDR1* promoter was found to contain the binding site for plant G-box/C-box bZIP protein TF belonging to the P\$GBOX family (Table 2, Table S1). We have used previously characterized promoter of the *PvSR2* gene to validate the presence of known TF binding sites on *PvSR2* promoter.

Analysis of TF binding sites on the promoters of *RDR2* of various plant species

RDR2 promoters of the above mentioned 16 plant taxa were used to find the TF binding sites present within. A total of 117 matrix family binding sites were found on the 16 *RDR2* promoter sequences by default settings of MatInspector (Fig S3, Table S2). A total of 14 different matrix families whose binding sites are present on each of the promoter sequence were used in this study. A total of 25 matrix family TFs are predicted to bind on all the *RDR2* promoter sequences based on stringent settings (Fig. 3, Table S2). Binding sites for P\$PCDR matrix family are present on *RDR2* promoter of *S. tuberosum*, *N. tabacum*, *N. attenuata*, *P. somniferum*, *G. max*, *G. hirsutum*, *V. vinifera*, *P. trichocarpa*, *A. thaliana* and *H. annuus* while matrix family P\$MYBL was predicted to be present on *S. tuberosum*, *S. lycopersicum*, *C. annuum*, *N. attenuata*, *O. sativa*, *G. max*, *G. hirsutum*, *P. trichocarpa* and *H. annuus* (Table 1). Moreover, binding sites of TFs belonging to the P\$WBXF matrix family were present on the *RDR2* promoter of *M. truncatula*, *N. attenuata*, *P. somniferum*, *A. hypogaea*, *G.*

Table 2 List of unique matrix families matching on a particular *RDR* promoter sequence of the specific plant

TF matrix family	<i>RDR1</i>	<i>RDR2</i>	<i>RDR3</i>	<i>RDR4</i>	<i>RDR5</i>	<i>RDR6</i>
P\$PNRE	–	–	–	–	–	<i>Ca</i>
P\$MYCL	–	–	–	–	<i>St, Ps</i>	–
P\$GLKF	–	–	–	–	<i>Na</i>	–
P\$GCCF	–	–	–	<i>Al</i>	–	–
P\$SRSF	–	–	–	<i>Os</i>	–	–
P\$TCPF	–	–	–	<i>Os</i>	–	–
O\$INRE	–	–	–	<i>Al</i>	–	–
P\$GARP	–	<i>Na</i>	–	–	–	–
P\$KAN1	–	<i>Nt</i>	–	–	–	–
P\$RAV5	–	<i>Gh</i>	–	–	–	–
P\$SEF4	–	<i>Pt</i>	–	–	–	–
P\$TALE	–	<i>Ah</i>	–	–	–	–
P\$TOEF	–	<i>Ca</i>	–	–	–	–
P\$WOXF	–	<i>Sl</i>	–	–	–	–
P\$ZFAT	–	<i>Gm</i>	–	–	–	–
O\$TF2B	–	–	<i>Os</i>	–	–	–
P\$FRSF	–	–	<i>Ah</i>	–	–	–
P\$IBOX	–	–	<i>Gm</i>	–	–	–
P\$RAV3	–	–	<i>Al</i>	–	–	–
P\$GBOX	<i>Gh</i>	–	–	–	–	–

Ah = *Arachis hypogaea*, *Al* = *Arabidopsis lyrata*, *Ca* = *Capsicum annuum*, *Gh* = *Gossypium hirsutum*, *Gm* = *Glycine max*, *Na* = *Nicotiana attenuata*, *Nt* = *Nicotiana tabacum*, *Os* = *Oryza sativa*, *Pt* = *Populus trichocarpa*, *Sl* = *Solanum lycopersicum*, *St* = *Solanum tuberosum* “–” indicates TF not found on promoters of respective *RDRs*. (‘P\$’ denotes plants and ‘O\$’ denotes general core promoter elements)

hirsutum, *A. thaliana* and *H. annuus*. *RDR2* promoters of some species also contain motif for the binding of DOF and AHL TFs (Table 1, Table S2). Binding sites of the TFs belonging to the P\$GARP, P\$KAN1, P\$RAV5, P\$SEF4, P\$TALE, P\$TOEF, P\$WOXF, and P\$ZFAT were also predicted to bind on the promoter of *NaRDR2*, *NtRDR2*, *GhRDR2*, *PtRDR2*, *AhRDR2*, *CaRDR2*, *SlRDR2* and *GmRDR2*, respectively (Table 2, Table S2).

Analysis of TF binding sites on the promoter *RDR6* of various plant species

For analysis, we used *RDR6* promoter from all the above-mentioned plants. MatInspector (default settings) suggested a total match of 121 matrix families on *RDR6* promoter sequences of these plants (Fig S4, Table S3). *RDR6* promoter of each plant possesses the presence of motif for the binding of TFs belonging to 14 different matrix families. Further, stringent analyses revealed presence of binding motifs of 19 different matrix families (Fig. 4, Table S3). *RDR6* promoter of *P. trichocarpa*, *G. hirsutum*, *A. thaliana*, *M. truncatula*, *G. max*, *S.*

lycopersicum, *C. annuum*, *N. attenuata*, *H. annuus*, *P. somniferum*, *C. quinoa* and *O. sativa* possess binding sites for P\$MYBL (Table 1). Moreover, motifs for P\$PCDR was present on *P. trichocarpa*, *G. hirsutum*, *A. thaliana*, *M. truncatula*, *C. annuum*, *N. attenuata*, *H. annuus* and *V. vinifera*. *RDR6* promoters of some species possess the binding of DOF and AHL TFs (Table 1, Table S3). Moreover, *CaRDR6* promoter possesses the binding site for the TF belonging to the P\$PNRE matrix family (Table 2).

Analysis of TF binding sites on the promoter of *RDR3* of various plant species

RDR3 gene and promoter sequence of some of the above-mentioned plant taxa, including *A. thaliana*, were not available in the NCBI database, therefore, we included *RDR3* promoter sequences of *A. lyrata*, *B. vulgaris*, *S. bicolor* and *P. patens* for promoter analysis. A total of 124 matrix family matches were obtained in all the *RDR3* promoter sequences and among them, 14 different matrix families matches were found to be common on each of the *RDR3* promoter sequences (Fig S5, Table S4). Further, we found a total of 22 matrix matches on all the promoters based on the stringent settings (Fig. 5, Table S4). Matrix match for P\$MYBL was found on *RDR3* promoters of *S. lycopersicum*, *P. patens*, *B. vulgaris*, *H. annuus* and *G. max* while P\$PCDR was present on *S. lycopersicum*, *C. quinoa*, *C. annuum*, *O. sativa* and *A. lyrata* promoters of this gene (Table 1). Furthermore, we found matrix matches of P\$WBXF in *S. lycopersicum*, *P. patens*, *C. quinoa*, *N. tabacum* and *A. lyrata*. *RDR3* promoter of few species possesses the binding OsLIC and AHL TFs. Moreover, promoters of *OsRDR3*, *AhRDR3*, *GmRDR3* and *AIRDR3* possess the binding sites for the TFs belonging to O\$TF2B (‘O\$’ denotes general core promoter elements), P\$FRSF, P\$IBOX and P\$RAV3, respectively (Table 2, Table S4).

Analysis of TF binding sites on the promoter of *RDR4* of various plant species

Full-length gene sequence of *RDR4* was available for only four plants species in the NCBI database (i.e., *A. lyrata*, *O. sativa*, *P. patens* and *S. moellendorffii*), therefore, we carried out *RDR4* promoter analysis of these four plant species. Based on default settings, MatInspector detected a total of 113 matrix family matches on all the promoter sequences used in the analysis, among them 25 different matrix matches were predicted to be common among the *RDR4* promoter sequences (Fig S6, Table S5). Furthermore, 21 matrix matches were present on all the promoter sequences based on stringent settings of the MatInspector (Fig. 6, Table S5). Interestingly, matrix match for P\$PCDR was present on all of the four *RDR4* promoter sequences

Fig. 3 Schematic representation of transcription factor binding sites on the putative promoters of *StRDR2*, *SIRD2*, *CaRDR2*, *NtRDR2*, *NaRDR2*, *HaRDR2*, *GmRDR2*, *VvRDR2*, *PtRDR2*, *GhRDR2*, *CqRDR2*, *AtRDR2*, *MtRDR2*, *AhRDR2*, *PsRDR2* and *OsRDR2* based

on the stringent conditions, i.e., core and matrix similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('P\$' denotes plants and 'O\$' denotes general core promoter elements)

(Table 1). Moreover, matrix matches of P\$WBXF and P\$LICM were present on the both *P. patens* and *O. sativa* *RDR4* promoter while P\$MYBL was present on *S. moellendorffii* and *A. lyrata* *RDR4* promoter sequences. *OsRDR4* promoter possesses the binding site for DOF and OsLIC TFs (Table 1). Furthermore, *AIRDR4* promoter possesses the binding sites for the TFs belonging to P\$GCCF and O\$INRE while *OsRDR4* promoter contains the binding sites for the TFs belonging to P\$SRSF and P\$TCPF matrix families (Table 2, Table S5).

Analysis of TF binding sites on the promoter of *RDR5* of various plant species

RDR5 promoters of *A. lyrata*, *C. quinoa*, *C. annuum*, *G. hirsutum*, *H. annuus*, *N. attenuata*, *P. somniferum*, *P.*

trichocarpa, *S. tuberosum*, *V. vinifera*, *N. tabacum*, *S. lycopersicum*, *A. hypogaea* and *M. truncatula* plants were considered for the analysis. We found a total of 124 matrix matches on all of the promoter sequences, among them 16 different matrix matches were commonly present on each of the promoter sequence (Fig S7, Table S6). Based on the stringent settings, 21 TF matrix matches were present on all the promoter sequences (Fig. 7, Table S6). Motifs of P\$MYBL matrix family were present in all *RDR5* promoters except *RDR5* promoter of *G. hirsutum*, *A. hypogaea* and *H. annuus* (Table 1). Match for P\$PCDR was found on *RDR5* promoters of *N. tabacum*, *N. attenuata*, *A. lyrata*, *P. trichocarpa*, *M. truncatula*, *V. vinifera*, *C. quinoa* and *H. annuus*. Matches for P\$WBXF were present on *N. tabacum*, *N. attenuata*, *S. tuberosum*, *S. lycopersicum*, *A. lyrata* and *P. somniferum*. Binding sites for DOF, OsLIC and

Fig. 4 Schematic representation of transcription factor binding sites on the putative promoters of *StRDR6*, *SIRD6*, *CaRDR6*, *NiRDR6*, *NaRDR6*, *HaRDR6*, *GmRDR6*, *VvRDR6*, *PtRDR6*, *GhRDR6*, *CqRDR6*, *AtRDR6*, *MtRDR6*, *AhRDR6*, *PsRDR6* and *OsRDR6* based

on the stringent conditions, i.e., core and matrix similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('PS' denotes plants and 'OS' denotes general core promoter elements)

AHL were also found on the *RDR5* promoter of several of these plant species. In addition, promoters of *StRDR5* and *PsRDR5* contains the binding sites of TFs belonging to P\$MYCL while *NaRDR5* harbours the binding site for P\$GLKF matrix family (Table 2, Table S6). A summarized list of common and specific matrix families, whose binding sites are predicted to be present on the RDRs promoter is shown in Tables 1 and 2, respectively.

Construction of RDRs-responsive protein interactome networks in rice and tomato

Interaction studies of *SIRD1*, *SIRD2*, *SIRD3*, *SIRD5*, *SIRD6*, *OsRDR1*, *OsRDR2*, *OsRDR3*, *OsRDR4* and *OsRDR6* were performed with the help of STRING database to generate a comprehensive protein

interactome network. Since *SIRD4* and *OsRDR5* genes are not available in the database, we could not include these for analysis (Zong et al. 2009). Minimum required interaction score of 0.4 was considered for interactome analyses. The predicted interactome network of *SIRDs* and *OsRDRs* are shown in Fig. 8 and Fig S8, respectively. Based on the STRING, it appears that DCL protein may interact with *SIRD1* while the product of LOC_Os03g38740.1 locus may interact with *OsRDR1* with high confidence (Table 3, Table S7). Interestingly, DCL1 is predicted to interact with all the *SIRDs* and, DCL3B was predicted to interact with *OsRDR1*, *OsRDR2*, *OsRDR3* and *OsRDR4* with a high confidence score. In addition, AGO7 may interact with *OsRDR3* and *OsRDR6*. (Table 3, Table S7, Table S8). Further, interaction between of LEAFBLADELESS 1 and PINHEAD

Fig. 5 Schematic representation of transcription factor binding sites on the putative promoters of *SIRDR3*, *CaRDR3*, *NtRDR3*, *HaRDR3*, *GmRDR3*, *PpRDR3*, *CqRDR3*, *SbRDR3*, *OsRDR3*, *AIRDR3*, *AhRDR3*

and *BvRDR3* based on the stringent conditions, i.e., core and matrix similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('P\$' denotes plants and 'O\$' denotes general core promoter elements)

Fig. 6 Schematic representation of transcription factor binding sites on the putative promoters of *SmRDR4*, *PpRDR4*, *OsRDR4* and *AIRDR4* based on the stringent conditions, i.e., core and matrix

similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('P\$' denotes plants and 'O\$' denotes general core promoter elements)

with *OsRDRs* is suggested which needs further experimental validation. Interestingly, no interaction of *OsRDR4* with *LEAFBLADELESS 1* was seen. String data also suggests *SGS3* as an interacting partner of *SIRDR6*, which has already been proved earlier (Lam et al. 2012; Mourrain et al. 2000) thereby validating the STRING results obtained.

Discussion

Plants protect themselves from the invading pathogens like viruses and viroids through RNA silencing pathways. Additionally, RNA silencing pathways also contributes to maintaining genome integrity. Double stranded RNAs (dsRNA) precursors trigger the process of RNA silencing.

Fig. 7 Schematic representation of transcription factor binding sites on the putative promoters of *StRDR5*, *SIRD5*, *CaRDR5*, *NtRDR5*, *NaRDR5*, *HaRDR5*, *VvRDR5*, *PtRDR5*, *GhRDR5*, *CqRDR5*, *AIRDR5*, *MIRDR5*, *AhRDR5*, and *PsRDR5* based on the stringent conditions,

i.e., core and matrix similarity 1.0. *SR2* gene of *P. vulgaris* was used as a control ('PS' denotes plants and 'OS' denotes general core promoter elements)

The pathway begins when the DCLs cleave the dsRNA into 21–24 nt small RNAs (sRNAs) duplex (primary siRNAs). The primary siRNAs thus formed are loaded on the AGO proteins. Consequently, the resulting AGO-sRNA complex, in the form of RNA induced silencing complex (RISC) interacts with the mRNA transcripts leading to the mRNA cleavage or translational arrest. In addition to the silencing of RNA, DNA is also silenced by the process of RdDM (Simon and Meyers 2011).

RDRs convert primary siRNAs into long dsRNAs either in primer-dependent or -independent manner (Devert et al. 2015). Long dsRNAs, thus formed, act as the substrate for DCL2/4 mediated cleavage and produces 21–22 nt secondary siRNA. With the catalytic action of AGO proteins, these secondary siRNAs enhance the silencing signal (Devert et al. 2015; Vaucheret 2006). Amplification of the silencing signal is required for the maintenance of non-cell autonomous silencing of RNA in plants. Eukaryotes possess three major families of RDRs, i.e., RDR α , RDR β

(absent in plants) and RDR γ . Phylogenetic analysis suggests that plant RDRs of γ family are more closely related to each other as compared to the members of the RDR α family, which indicate that probably RDRs of α family duplicated earlier than the RDRs of γ family during the course of evolution.

Recently, the molecular mechanism of *RDR1* transcripts induction by *Rice stripe virus* (RSV) in *O. sativa* has been suggested (Wang et al. 2016). MADS-box protein binds to the promoter of *RDR1* to inhibit transcription of *OsRDR1*. miRNA444 (miR444) is induced upon RSV infection in *O. sativa* which in turn induces the expression of *OsRDR1* and enhances tolerance in rice plants against RSV infection. miR444 abolishes the MIKCC-type MADS-box mediated repression of *OsRDR1* promoter by binding and inhibiting the expression of MIKCC-type MADS-box transcript (Wang et al. 2016). Thus, it would be interesting to study, if there is any change in the level of miR444 transcripts after infection with other plant viruses. The presence of MADS-

Fig. 8 Interactome of a-SIRD1, b-SIRD2, c-SIRD3, d-SIRD5 and e-SIRD6 proteins with other proteins. Maximum degree of interaction was observed with the dicer like proteins (DCLs). Minimum required interaction score was medium confidence (0.400)

box TF (SEPALLATA3) binding motifs on all the *RDRs* promoters, including *OsRDR1*, were also observed in our analysis when the core similarity and the matrix similarity were set as 0.75 and optimal, respectively (Table S1, Table S2, Table S3, Table S4, Table S5 and Table S6). However, our stringent analysis parameters, i.e., core and matrix similarity 1.0, excludes the presence of *MADS box* motifs, suggesting that the MatInspector's default settings are indeed reliable for the promoter analysis. No other information is available regarding the regulation of *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6* in plants.

Our findings suggest several TFs which may bind to the promoter of *RDR* genes and further experimental validation of these is required in the future. Transcription factor families MYB, NAC and WRKY might influence the expression of *RDR* genes during various stresses. AtMYB44, which is a member of the P\$MYBL matrix family, was predicted to bind on *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6* promoters of many plants, but not all (Table 1). There are many members of P\$MYBL matrix

family, however, as per our findings, *RDRs* promoters have the binding sites for either AtMYB44 or the AS1/AS2 repressor protein from P\$MYBL family. MYB TFs possess several imperfect repeats of a 52 amino acid motif, which is known as MYB domain. MYB TFs are found in both animals and plants. MYB TFs of plants are classified under R2R3-MYB family. The role of *Arabidopsis* R2R3-MYB TFs have been suggested in biotic as well as abiotic stress responses (Seo and Park 2010; Dubos et al. 2010). JA induces the expression of AtMYB44 through CORONATINE INSENSITIVE 1 (COI1). SA also induces the expression of AtMYB44. Overexpression of AtMYB44 induces the expression of WRKY70 and PRI leading to enhanced defence response against biotrophic pathogens, however, they negatively regulate defence response against necrotrophic pathogens by downregulating anti-microbial defensin PDF1.2 and VEGETATIVE STORAGE PROTEIN 1 (VSP1) (Shim et al. 2013). Recently it has been demonstrated that JA and SA induces the expression of *RDR1* (Hunter et al. 2013). Taken together, our result

Table 3 Selected protein partners of RDR1, RDR2 and RDR6 of tomato and rice as predicted by STRING

	Tomato			Rice		
	Predicted partner	Function	Score	Predicted partner	Function	Score
RDR1	DCL	Protein DCL (224 aa)	0.906	4,333,337/ LOC_Os03g38740.1	Dicer (1410 aa)	0.929
	Solyc11g008520.1.1	Uncharacterized protein (1428 aa)	0.798	DCL3B	Dicer (1571 aa)	0.854
	Solyc06g048960.2.1	Uncharacterized protein (1399 aa)	0.798	4,324,864/ LOC_Os01g68120.1	DCL3, putative, expressed (1597 aa)	0.827
	Solyc07g005030.2.1	Uncharacterized protein (1536 aa)	0.797	4,336,991/ LOC_Os04g52540.1	Retrotransposon protein (1034 aa)	0.782
	DCL1	Endoribonuclease Dicer homolog 1 (1888 aa)	0.789	4,351,724/ LOC_Os12g09580.1	Leafbladeless1, (609 aa)	0.768
	LOC544268	Chloroplast-specific ribosomal protein (311 aa)	0.7	4,342,127/ LOC_Os06g51190.1	Lysine ketoglutarate reductase trans-splicing related 1 (376 aa)	0.763
	LOC543942	Regulator of gene silencing (198 aa)	0.695	4,331,440/ LOC_Os03g02970.1	Dicer (1883 aa)	0.716
	CLEB3J9	Thylakoid lumenal 29 kDa protein, chloroplastic (345 aa)	0.695	4331253/ LOC_Os02g58490.1	PINHEAD (1011 aa)	0.698
	AO	Ascorbate oxidase (578 aa)	0.695	4336362/ LOC_Os04g43050.1	Dicer, putative, expressed (1591 aa)	0.696
	SGS3	SUPPRESSOR OF GENE SILENCING 3 (633 aa)	0.648	4330281/ LOC_Os02g45070.1	PINHEAD (1082 aa)	0.634
RDR2	Solyc11g008520.1.1	Uncharacterized protein (1428 aa)	0.802	4331440/ LOC_Os03g02970.1	Dicer (1883 aa)	0.869
	Solyc06g048960.2.1	Uncharacterized protein (1399 aa)	0.802	DCL3B	Dicer (1571 aa)	0.866
	Solyc07g005030.2.1	Uncharacterized protein (1536 aa)	0.801	4333337/ LOC_Os03g38740.1	Dicer (1410 aa)	0.865
	DCL1	Endoribonuclease Dicer homolog 1 (1888 aa)	0.798	4351724/ LOC_Os12g09580.1	Leafbladeless1 (609 aa)	0.799
	Solyc03g120940.2.1	Uncharacterized protein (304 aa)	0.718	4331253/ LOC_Os02g58490.1	PINHEAD (1011 aa)	0.79
	SGS3	SUPPRESSOR OF GENE SILENCING 3 (633 aa)	0.667	4330281/ LOC_Os02g45070.1	PINHEAD (1082 aa)	0.774
	Solyc11g008540.1.1	Uncharacterized protein (1352 aa)	0.587	4336690/ LOC_Os04g47870.1	PINHEAD (1118 aa)	0.773
	Solyc11g008530.1.1	Uncharacterized protein (1317 aa)	0.587	4333232/ LOC_Os03g33650.1	AGO7 (1048 aa)	0.764
	DCL3	Endoribonuclease Dicer homolog 3a (1430 aa)	0.587	4336362/ LOC_Os04g43050.1	Dicer (1591 aa)	0.76
	Solyc01g109970.2.1	Uncharacterized protein (907 aa)	0.574	4324864/ LOC_Os01g68120.1	DCL3 (1597 aa)	0.728
RDR6	SGS3	SUPPRESSOR OF GENE SILENCING 3 (633 aa)	0.816	4336362/ LOC_Os04g43050.1	Dicer (1591 aa)	0.926
	Solyc11g008520.1.1	Uncharacterized protein (1428 aa)	0.803	4333232/ LOC_Os03g33650.1	AGO7 (1048 aa)	0.913
	Solyc06g048960.2.1	Uncharacterized protein (1399 aa)	0.803	4351724/ LOC_Os12g09580.1	Leafbladeless1 (609 aa)	0.902
	Solyc07g005030.2.1	Uncharacterized protein (1536 aa)	0.803	DCL3B	Dicer (1571 aa)	0.866
	DCL1	Endoribonuclease Dicer homolog 1 (1888 aa)	0.801	4333337/ LOC_Os03g38740.1	Dicer (1410 aa)	0.866
	Solyc11g008540.1.1	Uncharacterized protein (1352 aa)	0.589	4331698/ LOC_Os03g06440.1	RNA helicase SDE3 (959 aa)	0.733

Table 3 continued

Tomato			Rice		
Predicted partner	Function	Score	Predicted partner	Function	Score
Solyc11g008530.1.1	Uncharacterized protein (1317 aa)	0.589	4324864/ LOC_Os01g68120.1	DCL3 (1597 aa)	0.728
DCL3	Endoribonuclease Dicer homolog 3a (1430 aa)	0.587	4331440/ LOC_Os03g02970.1	Dicer (1883 aa)	0.723
Solyc06g054050.1.1	Uncharacterized protein (911 aa)	0.558	4336690/ LOC_Os04g47870.1	PINHEAD (1118 aa)	0.665
Solyc06g054020.2.1	Uncharacterized protein (956 aa)	0.558	4330281/ LOC_Os02g45070.1	PINHEAD (1082 aa)	0.665

suggests that MYB44 may bind to promoters of *RDRs* including *RDR1* diverse plants species and this pathway may be induced by both SA and JA.

Another TF, AS1, a nuclear protein with MYB domain, was also predicted to interact with *RDR* promoters, in association with AS2. AS1/AS2 TFs are also members of P\$MYBL matrix family (Cartharius et al. 2005). A recent report suggests the role of *RDR1* in TGS (Basu et al. 2018). Higher level of *RDR1* has been demonstrated to induce symptom remission against *Tomato leaf curl Gujarat virus* (*ToLCGV*) in tobacco. Further, *RDR1* was also found to enhance methylation of *ToLCGV* promoter (Basu et al. 2018). Validating the interaction of AS1 and AS2 proteins with the *RDR1* promoters might expose the link between *RDR1* and DNA methylation during TGS. Also, it will be interesting to understand the regulation of *RDR1* following geminivirus infection and the role of AS1/AS2 in regulating *RDR1* in recovered plant tissues.

Our analysis suggests that GmNAC81/GmNAC30 heterodimer, the members of the P\$PCDR matrix family, bind to the *RDRs* promoter of many plants (Table 1). The role of GmNAC81/GmNAC30 TFs have been implicated in PCD (Mendes et al. 2013). The binding of GmNAC81/GmNAC30 heterodimer and the expression of *RDRs* might be crucial for providing resistance to plants against biotic and abiotic stresses. Role of WRKY1 in plant's biology is not precisely understood, since in some plant-pathogen interaction WRKY acts as the positive regulator of defence while in other cases, it acts as a negative regulator of defence (Li et al. 2015a, b; Marchive et al. 2007; Mzid et al. 2007; Oh et al. 2008). Further studies would be needed to understand the link between WRKY1 and *RDR* proteins.

The role of DOF family of transcription factors have been implicated in light responses and development of plants (Noguero et al. 2013). Our in silico study suggests that DOF3.4 may bind on the promoters of *RDR1*, *RDR2*, *RDR4*, *RDR5* and *RDR6* of many plant species (except *RDR3*), implicating regulation of these *RDRs* during plant

development (Table 1, Table S1). OsLIC TF is required for the rice architecture, i.e., tillering number and angle, internodes elongation, panicle morphology and leaf angle (Wang et al. 2008). The role of *RDRs* in development is also indicated by our study since OsLIC TF was predicted to bind on the promoters *RDR1*, *RDR3*, *RDR4* and *RDR5* of various plant species (Table 1, Table S1). AHL TFs regulate the homeostasis of various plant hormones (Matsushita et al. 2007). AHL20 was predicted to bind on the promoters of all *RDRs* (except *RDR4*) of various species, indicating the role of AHL20 in maintaining hormone homeostasis in plants.

RDRs are crucial proteins involved in the plant defence, and its expression is induced by plant viruses and many plant hormones (Hunter et al. 2013). However, the mechanism of its induction remains to be studied. Till today, except for *RDR2* and *RDR6*, interacting protein partners of other *RDRs* are not yet identified (Lam et al. 2012; Law et al. 2011). STRING data suggests that DCL proteins, DCL1 and DCL3b may interact with *RDR* proteins. In addition, AGO7, which is required for the generation of trans-acting small interfering RNAs (ta-siRNAs), might regulate the function of OsRDR3 and OsRDR6 (Jouannet et al. 2012).

In summary, the present study highlights putative TFs that can bind to the *cis-acting* elements of promoters of *RDRs*. Results suggests that the TFs binding motifs of MYB family (AtMYB44 and AS1), WRKY family (WRKY1), MADS box family (SEPALLATA3) etc. are commonly present on the *RDRs* promoters. In addition, in silico interaction studies for identification of probable interacting partners of tomato and rice *RDRs* indicated various probable partners (DCL1, DCL3b, AGO7, LEAF-BLADELESS, PINHEAD etc.) of *RDR* proteins. The present in silico study shall provide some clue for the identification of TFs which bind to the *RDRs* promoter and the identification of interacting protein partners of *RDR* proteins in the future.

Funding Funding was provided by UPE-II, UGC (Grant No. JNU/UPE-II/SLS/SC/13).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing financial interests.

References

- Basu S, Kushwaha NK, Singh AK, Sahu PP, Kumar RV, Chakraborty S (2018) Dynamics of a geminivirus-encoded pre-coat protein and host RNA-dependent RNA polymerase 1 in regulating symptom recovery in tobacco. *J Exp Bot* 69(8):2085–2102. <https://doi.org/10.1093/jxb/ery043>
- Butterbach P, Verlaan MG, Dullemans A, Lohuis D, Visser RGF, Bai Y, Kormelink R (2014) Tomato yellow leaf curl virus resistance by Ty-1 involves increased cytosine methylation of viral genomes and is compromised by cucumber mosaic virus infection. *Proc Natl Acad Sci USA* 111(35):12942–12947. <https://doi.org/10.1073/pnas.1400894111>
- Cao JY, Xu YP, Li W, Li SS, Rahman H, Cai XZ (2016) Genome-wide identification of dicer-like, Argonaute, and RNA-dependent RNA polymerase gene families in Brassica species and functional analyses of their Arabidopsis homologs in resistance to *Sclerotinia sclerotiorum*. *Front Plant Sci* 7:1614. <https://doi.org/10.3389/fpls.2016.01614>
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M et al (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 21:2933–2942. <https://doi.org/10.1093/bioinformatics/bti473>
- Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101(5):543–553. [https://doi.org/10.1016/S0092-8674\(00\)80864-8](https://doi.org/10.1016/S0092-8674(00)80864-8)
- Devert A, Fabre N, Floris M, Canard B, Robaglia C, Crt P (2015) Primer-dependent and primer-independent initiation of double stranded RNA synthesis by purified *Arabidopsis* RNA-dependent RNA polymerases RDR2 and RDR6. *PLoS ONE* 10(3):e0120100. <https://doi.org/10.1371/journal.pone.0120100>
- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15:573–581. <https://doi.org/10.1016/j.tplants.2010.06.005>
- Guo M, Thomas J, Collins G, Timmermans MC (2008) Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. *Plant Cell* 20:48–58. <https://doi.org/10.1105/tpc.107.056127>
- Guo H, Wang Y, Wang L, Hu P, Wang Y et al (2017) Expression of the MYB transcription factor gene BpMYB46 affects abiotic stress tolerance and secondary cell wall deposition in *Betula platyphylla*. *Plant Biotechnol J* 15:107–121. <https://doi.org/10.1111/pbi.12595>
- Hunter LJR, Westwood JH, Heath G, Macaulay K, Smith AG et al (2013) Regulation of RNA-dependent RNA polymerase 1 and isochorismate synthase gene expression in Arabidopsis. *PLoS ONE* 8:e66530. <https://doi.org/10.1371/journal.pone.0066530>
- Iwasaki M, Takahashi H, Iwakawa H, Nakagawa A, Ishikawa T et al (2013) Dual regulation of ETTIN (ARF3) gene expression by AS1-AS2, which maintains the DNA methylation level, is involved in stabilization of leaf adaxial-abaxial partitioning in Arabidopsis. *Development* 140:1958–1969. <https://doi.org/10.1242/dev.085365>
- Jouannet V, Moreno AB, Elmayan T, Vaucheret H, Crespi MD, Maizel A (2012) Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *EMBO J* 31:1704–1713. <https://doi.org/10.1038/emboj.2012.20>
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Lam P, Zhao L, McFarlane HE, Aiga M, Lam V et al (2012) RDR1 and SGS3, components of RNA-mediated gene silencing, are required for the regulation of cuticular wax biosynthesis in developing inflorescence stems of Arabidopsis. *Plant Physiol* 159:1385. <https://doi.org/10.1104/pp.112.199646>
- Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE (2011) SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet* 7(7):e1002195. <https://doi.org/10.1371/journal.pgen.1002195>
- Leibman D, Kravchik M, Wolf D, Haviv S, Weissberg M et al (2018) Differential expression of cucumber RNA-dependent RNA polymerase 1 genes during antiviral defence and resistance. *Mol Plant Pathol* 19(2):300–312. <https://doi.org/10.1111/mpp.12518>
- Li JB, Luan YB, Liu Z (2015a) Overexpression of SpWRKY1 promotes resistance to *Phytophthora nicotianae* and tolerance to salt and drought stress in transgenic tobacco. *Physiol Plant* 155:248–266. <https://doi.org/10.1111/pp.12315>
- Li JB, Luan YB, Liu Z (2015b) SpWRKY1 mediates resistance to *Phytophthora infestans* and tolerance to salt and drought stress by modulating reactive oxygen species homeostasis and expression of defense-related genes in tomato. *Plant Cell Tissue Org Cult* 123:67–81. <https://doi.org/10.1007/s11240-015-0815-2>
- Lodha M, Marco CF, Timmermans MC (2013) The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes Dev* 27:596–601. <https://doi.org/10.1101/gad.211425.112>
- Mao X, Zhang H, Qian X, Li A, Zhao G, Jing R (2012) TaNAC2, a NAC-type wheat transcription factor conferring enhanced multiple abiotic stress tolerances in Arabidopsis. *J Exp Bot* 63:2933–2946. <https://doi.org/10.1093/jxb/err462>
- Marchive C, Mzid R, Deluc L, Barrieu F, Pirrello J et al (2007) Isolation and characterization of a *Vitis vinifera* transcription factor, VvWRKY1, and its effect on responses to fungal pathogens in transgenic tobacco plants. *J Exp Bot* 58:1999–2010. <https://doi.org/10.1093/jxb/erm062>
- Mare C, Mazzucotelli E, Crosatti C, Francia E, Stanca AM, Cattivelli L (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Mol Biol* 55:399–416. <https://doi.org/10.1007/s11103-004-0906-7>
- Matsushita A, Furumoto T, Ishida S, Takahashi Y (2007) AGF1, an AT-hook protein, is necessary for the negative feedback of *AtGA3ox1* encoding GA 3-oxidase. *Plant Physiol* 143(3):1152–1162. <https://doi.org/10.1104/pp.106.093542>
- Melnyk CW, Molnar A, Baulcombe DC (2011) Inter-cellular and systemic movement of RNA silencing signals. *EMBO J* 30:3553–3563. <https://doi.org/10.1038/emboj.2011.274>
- Mendes GC, Reis PAB, Calil IP, Carvalho HH, Arago FJL, Fontes EPB (2013) GmNAC30 and GmNAC81 integrate the endoplasmic reticulum stress- and osmotic stress-induced cell death responses through a vacuolar processing enzyme. *Proc Natl Acad Sci USA* 110(48):19627–19632. <https://doi.org/10.1073/pnas.1311729110>
- Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C et al (2000) Arabidopsis SGS2 and SGS3 genes are required for

- posttranscriptional gene silencing and natural virus resistance. *Cell* 101:533–542. [https://doi.org/10.1016/S0092-8674\(00\)80863-6](https://doi.org/10.1016/S0092-8674(00)80863-6)
- Mzid R, Marchive C, Blancard D, Deluc L, Barrieu F et al (2007) Overexpression of VvWRKY2 in tobacco enhances broad resistance to necrotrophic fungal pathogens. *Physiol Plant* 131:434–447. <https://doi.org/10.1111/j.1399-3054.2007.00975.x>
- Nishimura T (2018) Gene body methylation involved in leaf development. *Plant Cell Physiol* 59:1288–1289. <https://doi.org/10.1093/pcp/pcy109>
- Noguero M, Atif RM, Ochatt S, Thompson RD (2013) The role of the DNA-binding One Zinc Finger (DOF) transcription factor family in plants. *Plant Sci* 209:32–45. <https://doi.org/10.1016/j.plantsci.2013.03.016>
- Oh SK, Baek KH, Park JM, Yi SY, Yu SH et al (2008) *Capsicum annuum* WRKY protein CaWRKY1 is a negative regulator of pathogen defense. *New Phytol* 177:977–989. <https://doi.org/10.1111/j.1469-8137.2007.02310.x>
- Oh JE, Kwon Y, Kim JH, Noh H, Hong SW, Lee H (2011) A dual role for MYB60 in stomatal regulation and root growth of *Arabidopsis thaliana* under drought stress. *Plant Mol Biol* 77:91–103. <https://doi.org/10.1007/s11103-011-9796-7>
- Pandey SP, Baldwin IT (2007) RNA-directed RNA polymerase 1 (RdR1) mediates the resistance of *Nicotiana attenuata* to herbivore attack in nature. *Plant J* 50:40–53. <https://doi.org/10.1111/j.1365-3113X.2007.03030.x>
- Pandey SP, Gaquerel E, Gase K, Baldwin IT (2008a) RNA-directed RNA polymerase 3 from *Nicotiana attenuata* is required for competitive growth in natural environments. *Plant Physiol* 147(3):1212–1224. <https://doi.org/10.1104/pp.108.121319>
- Pandey SP, Shahi P, Gase K, Baldwin IT (2008b) Herbivory-induced changes in the small-RNA transcriptome and phytohormone signaling in *Nicotiana attenuata*. *Proc Natl Acad Sci USA* 105(12):4559–4564. <https://doi.org/10.1073/pnas.0711363105>
- Phelps-Durr TL, Thomas J, Vahab P, Timmermans MC (2005) Maize rough sheath2 and its Arabidopsis orthologue ASYMMETRIC LEAVES1 interact with HIRA, a predicted histone chaperone, to maintain knox gene silencing and determinacy during organogenesis. *Plant Cell* 17:2886–2898. <https://doi.org/10.1105/tpc.105.03547>
- Prakash V, Devendran R, Chakraborty S (2017) Overview of plant RNA dependent RNA polymerases in antiviral defense and gene silencing. *Indian J Plant Physiol* 22:493–505. <https://doi.org/10.1007/s40502-017-0339-3>
- Qi X, Zhang Y, Chai T (2007) characterization of a novel plant promoter specifically induced by heavy metal and identification of the promoter regions conferring heavy metal responsiveness. *Plant Physiol* 143:50. <https://doi.org/10.1104/pp.106.080283>
- Qiao Z, Li CL, Zhang W (2016) WRKY1 regulates stomatal movement in drought stressed *Arabidopsis thaliana*. *Plant Mol Biol* 91:53–65. <https://doi.org/10.1007/s11103-016-0441-3>
- Qu F, Ye X, Hou G, Sato S, Clemente TE, Morris TJ (2005) RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in *Nicotiana benthamiana*. *J Virol* 79:15209–15217. <https://doi.org/10.1128/JVI.79.24.15209-15217.2005>
- Raja P, Sanville BC, Buchmann RC, Bisaro DM (2008) Viral genome methylation as an epigenetic defense against geminiviruses. *J Virol* 82(18):8997–9007. <https://doi.org/10.1128/JVI.00719-08>
- Rakshandehroo F, Takeshita M, Squires J, Palukaitis P (2009) The influence of RNA-dependent RNA polymerase 1 on potato virus Y infection and on other antiviral response genes. *Mol Plant Microbe Interact* 22:1312–1318. <https://doi.org/10.1094/MPMI-22-10-1312>
- Schwach F, Vaistij FE, Jones L, Baulcombe DC (2005) An RNA-dependent RNA polymerase prevents meristem invasion by potato virus X and is required for the activity but not the production of a systemic silencing signal. *Plant Physiol* 138:1842. <https://doi.org/10.1104/pp.105.063537>
- Seki M, Kamei A, Yamaguchi-Shinozaki K, Shinozaki K (2003) Molecular responses to drought, salinity and frost: common and different paths for plant protection. *Curr Opin Biotechnol* 14:194–199. <https://doi.org/10.1105/tpc.006130>
- Seo PJ, Park CM (2010) MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in Arabidopsis. *New Phytol* 186:471–483. <https://doi.org/10.1111/j.1469-8137.2010.03183.x>
- Shao F, Lu S (2014) Identification, molecular cloning and expression analysis of five RNA-dependent RNA polymerase genes in *Salvia miltiorrhiza*. *PLoS ONE* 9(4):e95117. <https://doi.org/10.1371/journal.pone.0095117>
- Shim JS, Jung C, Lee S, Min K, Lee Y, Choi Y, Lee JS, Song JT, Kim J, Choi YD (2013) AtMYB44 regulates WRKY70 expression and modulates antagonistic interaction between salicylic acid and jasmonic acid signaling. *Plant J* 73:483–495. <https://doi.org/10.1111/tpj.12051>
- Simon SA, Meyers BC (2011) Small RNA-mediated epigenetic modifications in plants. *Curr Opin Plant Biol* 14:148–155. <https://doi.org/10.1016/j.pbi.2010.11.007>
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D et al (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43:D447–D452. <https://doi.org/10.1093/nar/gku1003>
- Vaucheret H (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev* 20:759–771. <https://doi.org/10.1101/gad.1410506>
- Verlaan MG, Hutton SF, Ibrahim RM, Kormelink R, Visser RGF, Scott JW et al (2013) The tomato yellow leaf curl virus resistance genes *Ty-1* and *Ty-3* are allelic and code for DFDGD-class RNA-dependent RNA polymerases. *PLoS Genet* 9(3):e1003399. <https://doi.org/10.1371/journal.pgen.1003399>
- Vrbsky J, Akimcheva S, Watson JM, Turner TL, Daxinger L, Vyskot B et al (2010) siRNA-mediated methylation of Arabidopsis telomeres. *PLoS Genet* 6(6):e1000986. <https://doi.org/10.1371/journal.pgen.1000986>
- Wang L, Xu Y, Zhang C, Ma Q, Joo SH, Kim SK et al (2008) OsLIC, a novel CCCH-type zinc finger protein with transcription activation, mediates rice architecture via brassinosteroids signaling. *PLoS ONE* 3(10):e3521. <https://doi.org/10.1371/journal.pone.0003521>
- Wang XB, Wu Q, Ito T, Cillo F, Li WX, Chen X, Yu JL, Ding SW (2010) RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 107:484–489. <https://doi.org/10.1073/pnas.0904086107>
- Wang H, Jiao X, Kong X, Hamera S, Wu Y et al (2016) A signaling cascade from mir444 to RDR1 in rice antiviral RNA silencing pathway. *Plant Physiol* 170:2365–2377. <https://doi.org/10.1104/pp.15.01283F>
- Wassenegger M, Krczal G (2006) Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Sci* 11:142–151. <https://doi.org/10.1016/j.tplants.2006.01.003>
- Willmann MR, Endres MW, Cook RT, Gregory BD (2011) The functions of RNA-dependent RNA polymerases in Arabidopsis. *Arabidopsis Book* 9:e0146. <https://doi.org/10.1199/tab.0146>
- Xie Z, Fan B, Chen C, Chen Z (2001) An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proc Natl Acad Sci USA* 98:6516–6521. <https://doi.org/10.1073/pnas.111440998>
- Yang SJ, Carter SA, Cole AB, Cheng NH, Nelson RS (2004) A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by *Nicotiana*

- benthamiana*. Proc Natl Acad Sci USA 101:6297–6302. <https://doi.org/10.1073/pnas.0304346101>
- Yu D, Fan B, MacFarlane SA, Chen Z (2003) Analysis of the involvement of an inducible Arabidopsis RNA-dependent RNA polymerase in antiviral defense. Mol Plant Microbe Interact 16:206–216. <https://doi.org/10.1094/MPMI.2003.16.3.206>
- Zong J, Yao X, Yin J, Zhang D, Ma H (2009) Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene 447:29–39. <https://doi.org/10.1016/j.gene.2009.07.004>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.