

**Understanding the role of glycine rich protein of *Nicotiana benthamiana* (NbGRP3) in geminiviral pathogenesis**

Thesis submitted to  
Jawaharlal Nehru University  
For the award of the degree of

**DOCTOR OF PHILOSOPHY**



Submitted by  
**RAGUNATHAN D**

School of Life Sciences,  
Jawaharlal Nehru University,  
New Delhi – 110067  
2019





School of Life Sciences  
Jawaharlal Nehru University  
New Delhi - 110067

### CERTIFICATE

This is to certify that the research work embodied in this thesis entitled, “**Understanding the role of glycine rich protein of *Nicotiana benthamiana* (NbGRP3) in geminiviral pathogenesis**” submitted for the award of the degree of doctor of philosophy has been carried out by **Mr. Ragunathan D** under the supervision of **Prof. Supriya Chakraborty** at the Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi. The work is original and has not been submitted in part or full for any degree or diploma in this or any other university or institute.

Ragunathan D

(Student)

12/3/19  
Prof. Supriya Chakraborty

**SUPRIYA CHAKRABORTY, Ph.D**  
(Supervisor)  
Professor  
School of Life Science  
Jawaharlal Nehru University  
New Delhi -110067

Prof. K. Natarajan 12/3/19

Dean, School of Life Sciences, JNU



प्रो. के. नटराजन / Prof. K. Natarajan  
डीन / Dean  
जीवन विज्ञान संस्थान / School of Life Sciences  
जवाहरलाल नेहरू विश्वविद्यालय  
Jawaharlal Nehru University  
नई दिल्ली / New Delhi-110067





## ***Acknowledgements***

*I convey my deep regards to all those who directly or indirectly helped me in accomplishing the research work I have performed. First and foremost, I would like to express my deepest gratitude & sincere thanks to **Prof. Supriya Chakraborty**, School of Life Sciences, Jawaharlal Nehru University (JNU), New Delhi for his patience, motivation, and immense mentorship.*

*I am thankful to current dean, Prof K. Natarajan and past deans, Prof. B. N. Mallick and Prof. S. K. Goswami, for providing the platform to use the best of the facilities. My sincere gratitude to the research advisory committee members Dr. Souvik Bhattacharjee, and Prof. A. K. Nandi for their critical comments for improving the research work. I thank University grant commission for providing me the Non-NET, Junior and Senior research fellowship during my PhD duration. I am grateful to the European Union for providing me the Erasmus Mundus Fellowship under the BRAVE project.*

*I am thankful to present and past lab members Dr. Vinoth, Dr. Veerendra, Dr. Soumik, Dr. Ashish, Dr. Nirbhay, Dr. Rajrani, Dr. Prabu, Dr. Kishore, Dr. Sneha, Dr. Vijay, Dr. Suprio, Dr. Arpita, Mansi, Ved, Tsewang, Neha, Manish, Jishan for their assistance. I am grateful to Prabhu K, who always helped me with the official procedures.*

*I am also thankful to Prof. Sven-Erik Behrens at the Martin-Luther University (MLU), Halle-Wittenberg for accepting me as a visiting PhD student and for giving access to the laboratory and research facilities. I am also thankful to the group members of Prof. Behrens for the kind cooperation and Dr. Torsten for offering me the bench space at MLU and Ms Cornelia for being a kind friend at MLU. My sincere thanks to Dr Ralph Golbik and Dr Tamilarasan for the active participations and discussions.*

*I thank my batch-mates for their informal milieu, lively atmosphere and proving a lot of memories of us at JNU. I am also thankful to all the SLS teachers and staffs for their help and support though out my research period.*

*I am so thankful to my family. They taught me from my childhood moral and social values. Overall, without the support of my family members, this research would not have been possible.*



## ABBREVIATION

nt	Nucleotide
IR	Intergenic region
LIR	large intergenic region
SIR	Small Intergenic region
CR	Common region
μg	Micro gram
μl	Micro liter
μM	Micro molar
mm	Milli molar
3-AT	3-Amino-1,2,4-triazole
TEMED	N,N,N',N',-Tetramethylethane-1.2-diamine
bp	base pair(s)
EDTA	Ethylenediaminetetraacetic acid
Kbp	kilobase pair(s)
KDa	Kilo Dalton
M	Molar
mg	Miligram(s)
mM	Millimole(s)
mRNA	Messenger RNA
N	Normal (concentrations)
ng	Nanogram(s)
RNA	Ribonucleic acid
DNA	Deoxyribonucleic acid
RNase A	Ribonuclease A
Tris	Tris (hydroxymethyl) aminomethane
%	Percentage
ICTV	International Committee on Taxonomy of Viruses
min	Minutes
ml	Milliliter
mm	Millimeter
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
OD	Optical density

rpm	Revolutions per minute
°C	Celsius
sec	Second
v/v	Volume by volume
w/v	Weight by volume

## TABLE OF CONTENTS

<b>Contents</b>	<b>Page no</b>
<b>1. Introduction</b>	<b>1-3</b>
<b>2. Review of literature</b>	<b>5-29</b>
<b>2.1 Geminivirus</b>	
<b>2.2 Genome organization</b>	
<b>2.3 DNA-A</b>	
<b>2.4 ORFs encoded by DNA-A</b>	
<b>2.4.1 AV1</b>	
<b>2.4.2 AV2</b>	
<b>2.4.3 AC1</b>	
<b>2.4.4 AC2</b>	
<b>2.4.5 AC3</b>	
<b>2.4.6 AC4</b>	
<b>2.4.7 AC5</b>	
<b>2.5 DNA-B</b>	
<b>2.6 ORFs encoded by DNA-B</b>	
<b>2.6.1 BV1</b>	
<b>2.6.2 BC1</b>	
<b>2.7 Satellite molecules</b>	
<b>2.7.1 Alphasatellites</b>	
<b>2.7.2 Betasatellites</b>	
<b>2.7.3 Deltasatellite</b>	
<b>2.8 Role of host factors in geminiviral pathogenesis</b>	



<p><b>2.8.1 Glycine rich proteins (GRPS)</b></p> <p><b>2.8.2 Classification of GRPs</b></p> <p><b>2.8.3 Class IVa GRP</b></p> <p>    <b>2.8.3.1 Structural organization of Class IVa GRP</b></p> <p>        <b>2.8.3.1.1 N-terminal RRM</b></p> <p>        <b>2.8.3.1.2 C-terminal GR domain</b></p> <p>    <b>2.8.3.2 Biophysical and Biochemical properties</b></p> <p>    <b>2.8.3.3 Role of GRP in biotic stress</b></p> <p>    <b>2.8.3.4 Role of GRP in abiotic stress</b></p> <p>    <b>2.8.3.5 Role of GRP in flowering</b></p> <p>    <b>2.8.3.6 Role of GRP in hormone signaling</b></p>	
<p><b>3. Materials and Methods</b></p> <p>    <b>3.1 Materials</b></p> <p>        <b>3.1.1 Seeds and plant material</b></p> <p>        <b>3.1.2 Infectious viral clones</b></p> <p>        <b>3.1.3 Bacterial strains</b></p> <p>        <b>3.1.4 Plasmid vectors</b></p> <p>        <b>3.1.5 Antibiotics and phenolic compound</b></p> <p>        <b>3.1.6 List of primers</b></p> <p>    <b>3.2 Methods</b></p> <p>        <b>3.2.1 <i>E. coli</i> competent cells preparation and their transformation</b></p> <p>        <b>3.2.2 <i>Agrobacterium tumefaciens</i> competent cell preparation and their transformation</b></p>	<p><b>31-44</b></p>





<p><b>3.2.3 Plasmid DNA extraction</b></p> <p><b>3.2.4 Restriction digestion of plasmid DNA</b></p> <p><b>3.2.5 Dephosphorylation of plasmid DNA</b></p> <p><b>3.2.6 Ligation</b></p> <p><b>3.2.7 Polymerase Chain Reaction (PCR)</b></p> <p><b>3.2.8 Prediction of disordered region in Glycine rich domain of NbGRP3</b></p> <p><b>3.2.9 Construction of expression clones of wild type <i>NbGRP3</i></b></p> <p><b>3.2.10 NbGRP3 expression and purification</b></p> <p><b>3.2.11 Circular Dichroism (CD) measurements</b></p> <p><b>3.2.12 Analytical ultracentrifugation</b></p> <p><b>3.2.13 Gel based DNA binding assays or Electrophoretic Mobility Shift Assays</b></p> <p><b>3.2.14 Intrinsic fluorescence assays</b></p> <p><b>3.2.15 Yeast two hybrid assay</b></p> <p><b>3.2.16 <i>NbGRP3</i> expression during viral infection</b></p> <p><b>3.2.17 Virus induced gene silencing of endogenous <i>NbGRP3</i></b></p> <p><b>3.2.18 Southern hybridization</b></p> <p><b>3.2.19 Analysis of C-terminal amino acid repeats in NbGRP3</b></p> <p><b>3.2.20 Cloning, expression and purification of NbGRP3 C-terminal deletion variant (NbGRP3<math>\Delta_{103-148}</math>)</b></p>	
<p><b>4. Results</b></p> <p><b>4.1 Structural domain analysis of NbGRP3</b></p> <p><b>4.2 Expression analysis of NbGRP3 during geminiviral infection</b></p> <p><b>4.3 Identification of interacting partners of NbGRP3</b></p> <p><b>4.4 Expression and purification of NbGRP3</b></p>	<p><b>45-74</b></p>



<p><b>4.5 Interaction of NbGRP3 with SCR of RaLCB</b></p> <p><b>4.6 Nucleic acid binding studies using intrinsic fluorescence</b></p> <p><b>4.7 Effect of transient silencing of <i>NbGRP3</i> on geminiviral pathogenesis</b></p> <p><b>4.8 Expression analysis of <i>NbGRP3</i> during TBSV infection</b></p> <p><b>4.9 Biochemical characterization of NbGRP3</b></p> <p><b>4.10 Sedimentation equilibrium analysis of NbGRP3 using analytical ultracentrifugation</b></p> <p><b>4.11 Identification of putative binding site of NbGRP3 in its 3'-UTR</b></p> <p><b>4.12 Amino acid repeats (AAR) in C-terminal</b></p> <p><b>4.13 Understanding repeats using different online servers</b></p> <p><b>4.14 Conservation of C-terminal repeats in Class IVa GRP from different species of plants</b></p> <p><b>4.15 Construction of C-terminal deletion construct of <i>NbGRP3</i></b></p> <p><b>4.16 <i>NbGRP3</i><math>\Delta_{103-148}</math> has impaired nucleic acid binding</b></p>	
<b>5. Discussion</b>	<b>75-79</b>
<b>6. Summary</b>	<b>81-82</b>
<b>7. References</b>	<b>83-91</b>
<b>8. Appendix</b>	<b>93-99</b>
<b>9. List of publications, conference abstracts and awards</b>	<b>101</b>



## LIST OF FIGURES

Figure no	Title	Page no
2.1	<b>Diagrammatic representation of genome organization of the genera classified under the family <i>Geminiviridae</i></b>	6
2.2	<b>Schematic representation of domain architecture of replication-associated protein (Rep)</b>	8
2.3	<b>Schematic representation of overview of recent advances in the understanding of host defense and counter defense exhibited by geminivirus-<math>\beta</math>C1</b>	17
2.4	<b>Diagrammatic representation of classification of GRPs into different classes</b>	20
2.5	<b>Overlay of solution structure of HvGRP1 and NtGRP1 solved by NMR</b>	22
4.1	<b>Prediction of disordered region using Globplot 2</b>	46
4.2	<b>Computational model of NbGRP3 generated by Robetta server</b>	46
4.3	<b>Expression analysis of <i>NbGRP3</i> during geminiviral infection</b>	47
4.4	<b>Cloning images of NbGRP3 into yeast two hybrid vectors</b>	49
4.5	<b>Yeast two-hybrid assay between NbGRP3 and ToLCNDV-A ORFs such as AC2, AV2 and AC4</b>	50
4.6	<b>Yeast two-hybrid assay between NbGRP3 and RaLCB-<math>\beta</math>C1</b>	50
4.7	<b>Construction of pET SUMOadapt-<i>NbGRP3</i> clone</b>	51
4.8	<b>Different steps involved in expression and purification of NbGRP3</b>	52
4.9	<b>Structural properties of purified full-length NbGRP3 using Far-UV CD spectrum</b>	53



<b>4.10</b>	<b>NbGRP3-RaLCB SCR interaction by EMSA</b>	<b>54</b>
<b>4.11</b>	<b>NbGRP3 – RaLCB SCR interaction using intrinsic fluorescence</b>	<b>55</b>
<b>4.12</b>	<b>Construction of pTRV2-<i>NbGRP3</i> clone for VIGS</b>	<b>56</b>
<b>4.13</b>	<b>Phenotype of uninoculated, TRV1+TRV2, TRV1+TRV2-<i>NbGRP3</i>, TRV1+TRV2-<i>PDS</i> infiltrated plants after 10 days post infiltration of VIGS constructs</b>	<b>56</b>
<b>4.14</b>	<b>Phenotype of geminivirus infiltrated plants on <i>NbGRP3</i> silenced plants 20 days post infiltration of geminivirus</b>	<b>57</b>
<b>4.15</b>	<b>Southern hybridization analysis demonstrating the effect of <i>NbGRP3</i> silencing in geminivirus DNA accumulation</b>	<b>58</b>
<b>4.16</b>	<b>Expression analysis of <i>NbGRP3</i> in TBSV infection samples at 5 dpi</b>	<b>59</b>
<b>4.17</b>	<b>Binding analysis of NbGRP3 with different homodeoxyribonucleotides using intrinsic fluorescence</b>	<b>59</b>
<b>4.18</b>	<b>Sedimentation equilibrium analysis of purified NbGRP3</b>	<b>61</b>
<b>4.19</b>	<b>Yeast two hybrid assay demonstrating the NbGRP3 as a monomer</b>	<b>61</b>
<b>4.20</b>	<b>Comparison of binding analysis of NbGRP3 with 3' UTR target sites form <i>NbGRP3</i> and <i>AtGRP7</i></b>	<b>63</b>
<b>4.21</b>	<b>Dotplot showing Low Complexity Repeats (LCR) in the C-terminal of NbGRP</b>	<b>64</b>
<b>4.22</b>	<b>Multiple sequence alignment of C-terminal disordered region of Class IVa GRP from different organisms using Clustal Omega</b>	<b>70</b>
<b>4.23</b>	<b>Multiple sequence alignment of C-terminal disordered region of Class IVa GRP from different species within the genera <i>Nicotiana</i> using Clustal Omega</b>	<b>71</b>
<b>4.24</b>	<b>Generation of C-terminal deletion variant NbGRP3<math>\Delta</math><sub>103-148</sub></b>	<b>72</b>
<b>4.25</b>	<b>Different steps involved in expression and purification of</b>	<b>73</b>





	<b>NbGRP3<math>\Delta</math><sub>103-148</sub></b>	
<b>4.26</b>	<b>Direct binding studies of NbGRP3<math>\Delta</math><sub>103-148</sub> with different nucleic acids showing the deletion has impaired nucleic acid binding</b>	<b>74</b>



## LIST OF TABLES

No	Table	Page no
3.1	List of different bacterial and yeast strains used in the current study	29
3.2	List of different plasmids used in the current study	30
3.3	List of antibiotics, phytohormones and phenolic compound used in the current study	31
3.4	List of the primers used in the current study	31
3.5	Reaction mixture for restriction digestion	34
3.6	Reaction mixture for alkaline phosphatase treatment	35
3.7	Reaction mixture for ligation	35
3.8	Reaction mixture for PCR	36
3.9	Details of different online servers used for analysis of C-terminal amino acid repeats	41
4.1	Results from online databases used for prediction of domains in NbGRP3	43
4.2	Conclusions from online servers used for understanding C-terminal repeats of NbGRP3	63
4.3	Dissection of C-terminal repeats by the server FAIR	64



# *Introduction*



## 1. Introduction

Plants are relentlessly exposed to threats by pathogens. Plant viruses pose a serious menace to agricultural productivity worldwide as they infect crops of significance to the economy. Geminiviruses, in particular, represent a potential threat to farming in tropical and subtropical regions of the world. Geminiviruses are the family of single-stranded DNA viruses transmitted by insects. Geminiviruses are classified under the family, *Geminiviridae* which includes nine different genera. The type member of this particular family is bean golden yellow mosaic virus—[Dominican Republic: 1987] which belong to the genus *Begomovirus* [1]. Ancestors of geminivirus were hypothesized to have evolved from phytoplasmal plasmids after the acquisition of protein-coding genes from ssRNA viruses infecting plants [2]. The genome of geminivirus can have either a single DNA component (referred as monopartite viruses) or two DNA components (referred as bipartite genome) encoding 5-7 proteins performing various functions such as replication, transcription, movement, etc. Geminiviruses, like any other plant viruses, are strictly obligate parasites; their dependency on the host machinery for their replication of genomes, expression of viral proteins, and pathogenesis is inevitable. In spite of their smaller genome and limited coding capacity, geminiviruses are known to be masters in reprogramming the host cellular processes for establishing a successful infection. Proteins encoded by geminivirus interact and interfere with numerous host cellular processes such as host cell cycle, host defense, cellular signaling and, hormone signaling, etc. [3]. The stated manipulations lead to physiological disorders in plants, which ultimately manifest as plant diseases.

Plants defend the viruses by restricting the viral replication and movement through approaches like plant immune signaling, gene silencing, triggering hypersensitive response, activation of defense hormones, protein degradation, etc. [4]. Pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) forms an essential part of plant immune signaling, and it is the first tier of plant defense. PTI is mediated by recognition of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRR). ETI involves recognition of specific virulence factors secreted by pathogens in the cell by intercellular receptors. *R* genes which are dominant resistant in nature are the best-studied examples of ETI. The consequence of both PTI

and ETI are activation of defense genes, induction of hypersensitive responses, salicylic acid (SA) accumulation, etc. [5]. Plants gain resistance against viruses when they lack the essential host factor required for the virus to complete its life cycle. Such host factors are governed by recessive resistance genes. Most of the recessive genes identified so far belong to eukaryotic initiation factors (eIFs) such as eIF4G and eIF4E. eIF4G and eIF4E mediated resistance were found to be effective against various groups of viruses [6, 7]. The main anti-viral strategy used by plants against viruses is RNA interference (RNAi) or gene silencing. Gene silencing is triggered by double-stranded RNA, and this process is evolutionarily conserved. Viruses have evolved mechanisms to circumvent this conserved process by encoding viral suppressors of gene silencing (VSR). The VSRS interfere at different steps of RNAi [8]. Defense hormones provide yet another layer of defense exhibited by plants against viruses. It was repeatedly observed that alteration of phytohormone levels were often associated with the viral infection and, there exists crosstalk among different phytohormones to mount an efficient antiviral defense against viruses [9]. One way of defense exhibited by plants against viruses is by eliminating the viral proteins by protein degradation via ubiquitin-proteasome system (UPS). UPS is like a double-edged sword, and this system is used by plants against viruses at the same time viral proteins target and hijack the members of UPS pathway to control their protein from degradation [10]. The examples cited indicate the essential role played by the host factors in viral pathogenesis. Upon further exploration more such examples may come up in the future.

Glycine rich proteins (GRP) are the superfamily of protein with high glycine content and are shown to perform varied functions. GRPs are classified into five different classes based on the conserved motifs and position of glycine rich region. Among the five different classes, class IVa deserves special attention. Structurally, class IVa GRP has two distinct domains: RNA recognition motif (RRM) at its N-terminal region and unstructured glycine rich region at its C-terminal [11, 12]. The proteins from this particular class have been demonstrated to be involved in various biological events such as seed germination, circadian rhythm, biotic, and abiotic stress responses [13]. Surplus reports confirm essentiality of Class IVa GRP in abiotic stress. Class IVa members from different organisms such as *Arabidopsis thaliana*, *Oryza sativa*, *Camelina sativa*, *Sorghum bicolor*, etc. are known to play a role in abiotic stress especially in conferring cold stress tolerance to plants upon its overexpression [14]. However, little efforts have gone in



understanding the role of class IVa GRP in biotic stress. *AtGRP7* a well-studied member in class IVa GRP was shown to influence plant defense directly and indirectly. *AtGRP7* could influence pathogenesis, both positively and negatively against diverse pathogens [15]. *AtGRP7* is the main target of the type III effector molecule HopU1 secreted by *Pseudomonas syringae* DC3000. HopU1, ADP-ribosylate *AtGRP7* at arginine in position 47 and 49 affecting the RNA binding property of *AtGRP7*, thereby making the plant susceptible to *P. syringae* [16]. *AtGRP7* overexpression had elevated levels of SA associated transcripts such as *PR* genes *PR1*, 2 and 5. JA/ET signalling associated transcripts such as *RAP2.3/ERF72/AtEBP*, *PDF1.1*, *PDF1.2a*, *RAP2.3/ERF72/AtEBP*, and *THI2.2* were down-regulated in overexpression line [17]. It was demonstrated that *AtGRP7* physically interact with *PDF1.2* regulating its expression post-transcriptionally [18]. In the context of viral pathogenesis, few clues are available in the literature. Overexpression of *AtGRP7* and its homolog *NgGRP* from *Nicotiana glutinosa* enhanced resistance towards tobacco mosaic virus (TMV) [15, 19], but so far no in-depth study was done linking class IVa GRP and viral pathogenesis. To have more insights into the role of class IVa GRP in viral pathogenesis, the current study was undertaken by using geminivirus and class IVa GRP (*NbGRP3*) from *Nicotiana benthamiana*. The following objectives were framed:

1. Identification of geminiviral factors interacting with *NbGRP3*.
2. Characterization of the interaction.
3. Identification of the role of *NbGRP3* in geminiviral pathogenesis.



# *Review of literature*



## 2. Review of Literature

### 2.1 Geminivirus:

Geminiviruses are the group of plant viruses defined by their small, single-stranded circular DNA of size 2.5 kb – 5.2 kb with non-enveloped capsid. The capsid contains single capsid protein forming twinned incomplete icosahedral shape giving its name geminivirus. They infect economically important crops causing severe damage to farmers and agriculture. Geminiviruses are transmitted by insect vectors. They are classified under the family *Geminiviridae* which include nine genera based on the insect vectors and genome organization namely *Becurtovirus*, *Curtovirus*, *Grablovirus*, *Mastrevirus*, *Turncurtovirus* (by leafhoppers), *Begomovirus* (by whiteflies), *Capulavirus* (by Aphids), *Eragrovirus*, *Topocuvirus* (by Treehopper). At present, the family *Geminiviridae* includes more than 440 species ([https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/ssdna-viruses/w/geminiviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/ssdna-viruses/w/geminiviridae)). The typical member of this family is *Bean golden yellow mosaic virus* which belongs to the genus *Begomovirus* [1]. The genus *Begomovirus* includes 288 species making it the largest genus in the viral taxonomy in the context of the number of species [20]. Begomoviruses infect a wide range of dicotyledonous plants of economic importance and their smaller genome size, making them among the early plant viruses to have its genome cloned and sequenced completely [21, 22].

### 2.2 Genome organization:

The genome organizations of the family *Geminiviridae* (except *Begomovirus*) are monopartite viruses (have a single genomic component). The genus *Begomovirus* have monopartite or bipartite (have two genomic components designated as DNA-A and DNA-B). In bipartite begomovirus, the DNA-A is equivalent to monopartite begomoviruses. Monopartite genome containing begomoviruses are common in old world countries such as Asia, Africa, Australia, and Europe. The bipartite begomoviruses are mostly in the new world countries like America, with exceptions [20]. Both the genomic components encode ORFs in bidirectional manner. ORFs partially overlap in some cases. Apart from the ORFs, every DNA component has a common region (CR) nested within intergenic region (IR) which is essential for replication and it is common for both DNA-A and DNA-B. CR region has Ori-region, a stem-loop region

containing a nonanucleotide sequence (5'-TAATATTAC-3'), and two promoters present bidirectionally for the RNA polymerase II to initiate transcription [23, 24].



**Figure 2.1: Diagrammatic representation of genome organization of the various genera classified under the family *Geminiviridae*.** The ORFs encoded by monopartite viruses are referred as V1, V2, V3, C1, C2 and C3 whereas in bipartite it is referred as AV1, AC1, AC2, AC3 etc. In bipartite viruses DNA-B encoded ORFs are BC1 and BV1. The ORFs are colored according to their functions. The hairpin loop structure has the site for origin of replication. CR – conserved region; LIR – long intergenic region; SIR – short intergenic region; cp – coat protein; rep, repA – replication associated protein; ren – replication enhancer protein; trap – transcription activator; mp – movement protein; nsp – nuclear shuttle protein [1].

### **2.3 DNA-A:**

The size of DNA-A is approximately 2.7 Kb whereas DNA-B is approximately 2.6 Kb. DNA-A encodes for 6 ORF in an overlapping manner such as AV1 and AV2 encoded from virion sense or sense strand and AC1, AC2, AC3, AC4, and AC5 encoded from complementary or antisense strand.

### **2.4 ORFs encoded by DNA-A:**

#### **2.4.1 AV1 (AR1/V1):**

It encodes for approximately 29 kDa protein named as coat protein (CP). It is the only structural gene encoded by geminiviruses, and it is a late gene which is highly conserved. It is involved in the assembly and packing of the geminiviral genome. Apart from functioning as a structural gene, it performs various other functions. The protein has N-terminal domain, central domain, and C-terminal domain. It is known to interact with ssDNA and dsDNA by its N-terminal domain [25]. The N-terminal domain also contains a nuclear localization signal (NLS) which helps in the shuttling of geminiviral DNA from the nucleus and the cytoplasm [26, 27]. Apart from nuclear shuttling, CP also helps in systemic movement of the virus [28]. CP being a late gene, it is not essential for geminiviral replication, but because of its multifunctional nature absence of CP results in the reduction of viral DNA accumulation [29].

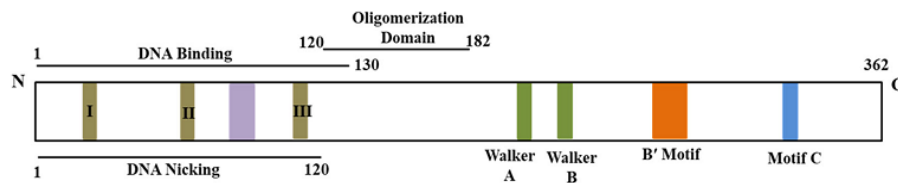
#### **2.4.2 AV2 (AR2/V2):**

This particular ORF encodes for a pre-coat protein which is of size approximately 13 kDa with its 3' end of ORF overlapping with 5' end of CP ORF. It is a pathogenicity determinant as well as silencing suppressor in *East African cassava mosaic Cameroon virus* (EACMCV), and its pathogenicity is dependent on a conserved protein kinase C [30]. In monopartite begomoviruses and in some bipartite begomovirus it helps in movement of the virus as it is evident by mutation studies, the mutation in AV2 causes reduction in ssDNA and dsDNA accumulation in plants not in protoplast and also localization studies with fusion proteins justify their role in movement [29, 31]. It also exhibits silencing suppressor activity which is mediated by its interaction with an important member in the RNA silencing pathway SGS3. V2 of *Tomato yellow leaf curl virus* (TYLCV) interacts with SISGS3 which is a homolog of AtSGS3. Mutation

at residues C84S/C86S in V2 abolishes AV2-SGS3 interaction as well as silencing suppressor activity of V2 [32]. Interestingly, V2 from *Tomato yellow leaf curl China virus* (TYLCCNV) act as VSR which doesn't interact with S1SGS3 but bind to 21 nt (nucleotide) siRNA duplex and 24 nt single-stranded siRNA [33]. Recent reports demonstrate that V2 is a suppressor of both TGS (Transcriptional Gene Silencing) and PTGS (Post-Transcriptional Gene Silencing). *Cotton leaf curl Multan virus* (CLCuMuV) V2 directly interacts with NbAGO4 and suppress RNA dependent DNA Methylation (RdDM) mediated by plant TGS machinery [34].

### 2.4.3 AC1 (AL1/C1):

AC1 encodes for Rep or replication associated protein, which is the most conserved protein as well as multitasking protein performing numerous functions during geminiviral pathogenesis. It is the first protein to be synthesized from geminivirus by the host. It is the essential protein which performs vital roles in replication and transcription. Structurally Rep has three distinct domains: N-terminal, central, and C-terminal domain. N-terminal domain is essential for DNA binding and DNA nicking activity, the central domain also referred as oligomerization domain which is involved in oligomerization of Rep. C-terminal domain is involved in ATPase activity as it has Walker A and Walker B motif essential for ATPase activity [35, 36].



**Figure 2.2: Schematic representation of the domain architecture of replication-associated protein (Rep).** Length of the Rep protein is 362 amino acids with three distinct domains: N-terminal, central, and C-terminal domain. N-terminal domain constitute DNA binding domain (1-130) and DNA nicking domain (1-120), conserved motifs I, II and III essential for initiation as well as termination of DNA replication; oligomerization domain (120-182) constitutes the central domain. The C-terminal domain possesses the ATPase domain, contain motifs such as walker A, walker B [37].

Rep was demonstrated to interact with numerous host factors involved in replication such as replication factor C (RFC), replication protein A (RPA), retinoblastoma-related protein



(RBR), proliferating cell nuclear antigen (PCNA), RAD51, RAD54, minichromosome maintenance protein 2 (MCM2), SUMO-conjugating enzyme (SCE-1) [36]. Rep also acts as silencing suppressor; it was shown to repress expression levels plant maintenance DNA methyltransferases of Methyltransferase 1 (MET1) and Chromomethylase 3 (CMT3) thereby interfering with the plant methylation cycle resulting in the reduction in CG methylation which is a powerful defense exhibited by plants against invading viruses [38]. Recently, it was demonstrated that Rep from *Chilli leaf curl virus* (ChiLCV) hijack host *N. benthamiana* ubiquitin machinery for enhanced transcription of viral genes. Rep interacts and co-localizes with ubiquitin-conjugating enzyme2 (NbUBC2) and histone monoubiquitination1 (NbHUB1) in the nucleus of the host. It also re-localizes NbUBC2 to the nucleoplasm from the cytoplasm. The above interaction results in monoubiquitination of histone 2B (H2B), making the ChiLCV minichromosome more transcriptionally active by trimethylation of histone 3 at lysine 4 residue (H3K4me3) [39].

#### **2.4.4 AC2 (AL2/C2):**

AC2 encodes a transcriptional activator protein (TrAP), as the name suggests TrAP known as transactivator protein controlling the promoter activity of viral genes AV1, BV1, and BC1 [40]. Initial studies with *Tomato golden mosaic virus* (TGMV) using protoplast transfection system revealed that deletion of AC2 resulted in the absence in coat protein formation and further studies with reporter assays revealed AC2 is essential for of coat protein (AV1) [41, 42]. AC2 does not interact with dsDNA strongly whereas, it exhibits strong, non-specific interaction with ssDNA. It was also found that binding of zinc with AC2 is essential for interaction with ssDNA. The transactivation activation of AC2 is mapped to 15 amino acids at its C-terminal [43]. It was later demonstrated that TrAP activates TGMV CP promoter by interacting with PEAPOD transcription factor and CP promoter driving the expression of CP gene [44]. AC2 is encoded by single cistron along with AC3 and the promoter mapped in AC1. AC2, along with AC1 synergistically activates the promoter in DNA-A driving expression of CP in case of *Mungbean yellow mosaic India virus* (MYMIV) [45]. AC2 localized in the nucleus, AC2 from EACMCV has NLS with sequence PRRR from position 28 to 32 where proline at 28th position is essential for nuclear localization but the entire stretch PRRR is essential for symptom

induction. At the same time, PRRR sequence is not required for transactivation role and interaction with adenosine kinase [46]. AC2 from *Bhendi yellow vein mosaic virus* (BYVMV) also has NLS at its N-terminal with 17-31 amino acids. NLS of C2 interacts with karyopherin –  $\alpha$  which is involved in the shuttling of molecules between nucleus and cytoplasm. C2 without NLS lead to localization of C2 both in nucleus and cytoplasm [47]. Moreover, BYVMV C2 is an essential for pathogenicity, C2 ORF with 2 stop codon resulted in a reduction in symptom development and viral titer. Recovery of symptoms appeared when the virus was inoculated with beta satellite [48].

The TrAP from different viruses acts as a VSR. TGMV AC2, *Beet curly top virus* (BCTV) C2, EACMCV C2 were all shown to be silencing suppressors. 1-100 amino acids are essential for silencing suppressor activity, and it is as effective as the full length in silencing suppressor activity. It mediates silencing suppressor activity by inhibiting Adenosine Kinase (ADK) [49]. *Beet severe curly top virus* (BSCTV) C2 interacts with S-adenosyl-methionine carboxylase 1 (SAMDC1) which is a positive regulator of viral pathogenesis and prevent the cellular degradations of SAMDC1 by 26S mediated proteasome complex. In addition, BSCTV C2 suppresses DNA methylation in viral DNA, mediates a decrease in DNA methylation in promotor of ACD6 and GSTF14 genes. ACD6 is the regulator of Salicylic acid (SA) signalling upstream whereas, GSTF14 is Glutathione-S-Transferase associated with numerous stress responses. Reduced accumulation of virus-associated siRNA observed in BSCTV without C2 when compared with BSCTV wild type [50, 51].

Apart from suppressing host silencing, suppression of host defense mediating transactivation, TrAP also propagates geminiviral transmission by attracting transmission vectors. C2 of *Tomato yellow leaf curl virus* (TYLCV) and *Papaya leaf curl China virus* (PaLCuCNV) interact with plant ubiquitin-40S ribosomal protein (RPS27A) which results in down-regulation of the inhibitor of jasmonic acid (JA) signaling i.e. JAZ1. Degradations JAZ1 is required for activation of JA signaling which mediates defense against insects by activating terpene biosynthesis. Down-regulation of JA signaling results in reduced expression of MYC2 regulated terpene synthesis. Reduced terpene synthesis attracts more insects [52].

#### **2.4.5 AC3 (AL3/C3):**

AC3 codes for an approximately 16 kDa protein referred as Replication Enhancer protein (REn). The virus with loss of AC3 results in a reduction in viral DNA replication in protoplast, leaf discs, whole plant, and compromised symptom development in whole plant [41, 53]. Heterologous recombination was also observed with AC3 from different organisms, i.e. AC3 from one virus can complement the function of AC3 to a different virus [54]. Their function as replication enhancer is mainly because of their interaction with geminivirus Rep and other replication associated host factors. REn can exhibit self-association leading to the formation of high order homo-oligomer through the hydrophobic domain at the central region of the protein and also with Rep enhancing the ATPase activity of Rep in vitro [55, 56]. The hydrophobic region of REn is also essential for the interaction with proliferative cell nuclear antigen (PCNA), which is a DNA clamp essential component of DNA replication. REn interacts with RBR via its N-terminal and C-terminal polar residues [55, 57, 58]. REn was shown to interact with a transcription factor NAC1 from tomato and its expression increased during viral infection. Both REn, and NAC1 co-localized in the nucleus, amino acids 1-70 comprising a putative  $\alpha$ -helix of REn is essential for the interaction with NAC1. Overexpression of NAC1 enhances viral replication [59]. Phage display studies with *Tomato leaf curl Kerala virus* (ToLCKeV) revealed several host factors involved in various metabolic pathways such as cell cycle and DNA replication, histone modification and RNAi pathway. Experimentally it was also shown that AC3 enhance gene silencing [60].

#### **2.4.6 AC4 (AL4/C4):**

AC4 is the ORF which is least conserved and located within AC1 ORF. It is a multifaceted protein serving as symptom determinant, helps in movement and a silencing suppressor. Disruption of C4 in *Tomato leaf curl virus* (ToLCV) in different host resulted in reduced symptom development but no impact on viral titer [61]. The role of AC4/C4 in symptom development was a much-explored area. Transgenic plant expressing BCTV C4 in *N. benthamiana* resulted in tissue distortion, development of enations containing a large clustered mass of unorganized cellular mass expressing C4 which justifies their role as symptom development [62]. Cotyledons and hypocotyledons from transgenic *Arabidopsis* expressing

BSCTV C4 under inducible promoter had extensive cell division with no clear demarcation of vascular bundles. With the increase in the concentration of the inducer, an increase in the abnormality observed. Interestingly, BR (Brassinosteroids) and ABA (Abscisic acid) could rescue the developmental abnormality to some extent, and the induced seedlings were more sensitive of GA (Gibberellic acid) and kinetin [63]. A similar transgenic study using BSCTV C4 in *Arabidopsis* showed that BSCTV C4 expressing transgenic plants had elevated levels of cell cycle associated transcripts. It had elevated levels of cyclins such as *cyc1*, *cyc2*, and *cyc2b*; cyclin-dependent kinases (CDK) such as *cdc2a*, *cdc2b*, *cdc25* and cyclin activated kinases (CAK) such as *cak1*, *cak2* and *cak3*. Some cell cycle inhibitors were suppressed in transgenic plants expressing C4 as well as during BSCTV infection [64].

Mutations studies with TYLCV C4 revealed the essentiality of C4 in systemic movement. TYLCV with mutated C4 could replicate in protoplast but not in tomato plants. The systemic movement in tomato was totally blocked, whereas in *N. benthamiana* systemic movement occurred but had compromised symptom development [65]. Similar observations were also made in BSCTV C4 where the BSCTV with two termination codon in the C4 ORF could replicate in protoplasts and leaf discs but could not establish symptoms in *Arabidopsis* and *N. benthamiana*. The newly-emerged leaves from the mutant virus infected plants had no viral DNA accumulation, whereas the exogenous application of BSCTV C4 could restore the movement function of mutated BSCTV C4. BSCTV C4 also shown to interact with ssDNA and dsDNA in a non-specific manner [66].

This ORF is also a VSR, first report demonstrating C4 as a VSR came in 2004. AC4 from ACMV synergistically suppress host PTGS along with EACMCV AC2 and enhanced EACMCV DNA accumulation by approximately 8 fold. The same study also reported AC4 from Sri Lankan cassava mosaic virus as PTGS suppressor [67]. ACMV C4 like that of other C4 induced abnormal development in transgenic *Arabidopsis* plants, interestingly upon further investigation, it was found the symptomology is associated with decreased miRNA accumulation. ACMV AC4 interacted with single-stranded siRNA and miRNA but not with any of double-stranded RNA forms. This displays the fact that AC4 blocks the PTGS at the mature stage [68]. Specific localization of AC4 seems to be essential for its silencing suppressor activity. EACMCV AC4

localizes in the plasma membrane, perinucleus, and cytoplasm. It has a consensus N-myristoylation site. Mutating glycine at 2 and cysteine at 3 to alanine abolishes its plasma membrane localization. AC4 without localization in plasma the membrane could not exhibit its silencing suppressor activity. It was shown that same AC4 interact with barely any meristem 1 (BAM1) which is a receptor-like kinase as well as a positive regulator for the cell to cell movement of silencing [69]. MYMV (*Mung bean yellow mosaic virus*) AC4 also targeted to the cellular membrane via S-palmitoylation despite the fact that it lacks transmembrane domain and hydrophilic in nature. MYMV AC4 also binds to 21-25 nucleotide siRNAs [70]. ToLCV C4 interacts with Shaggy like kinase from tomato in the nucleus region, and the interaction mapped to 12 amino acids at the C-terminal of C4. Deletion of the 12 amino acids abolished the interaction as well as silencing suppressor activity of C4 [71].

#### **2.4.7 AC5 (AL5/C5):**

This particular ORF is not prevalent like other ORF and least studied as well as it doesn't code for essential protein for geminivirus in most viruses studied and it is located downstream of AC3/C3 overlapping with CP [72-74], but there are exceptions. AC5 was found to be an essential protein for DNA replication in *Mungbean yellow mosaic India virus* (MYMIV) [75]. *Tomato leaf deformation virus* (ToLDeV) C5 mutant lacking C5 had a less severe symptom [74]. It is also a silencing suppressor and pathogenicity determinant in MYMIV. AC5 mutant of MYMIV resulted in a decrease in infectivity as well as viral DNA accumulation which was restored in AC5 transgenic plants. It can induce hypersensitive response when inoculated in Potato virus X (PVX) vector, but such phenotype was not observed with MYMIV infected plants. AC5 can suppress PTGS and N-terminal region of AC5 essential for PTGS suppressor activity. Interestingly, AC5 can also suppress TGS by suppressing the expression of Domain rearranged methyltransferase 2 (DRM2) which is a methyltransferase responsible for de novo and maintenance of CHH methylation. 33 amino acids at C-terminal of the protein found to be essential for TGS suppressor activity as well as symptom development [76].

## **2.5 DNA-B:**

DNA-B is similar to that of DNA-A in size and totally dependent on DNA-A for its replication. It encodes two ORFs: BR1/BV1 from the sense strand or virion strand and BL1/BC1 from the anti-sense or complementary strand. DNA-B also has IR region with CR region and nonnucleotide essential for replication.

## **2.6 ORFs encoded by DNA-B:**

### **2.6.1 BV1 (BR1):**

This ORF is oriented towards antisense direction in DNA-B of bipartite viruses. Its primary function is to export viral DNA out of the nucleus; therefore, it is commonly referred to as nuclear shuttle protein (NSP). NSP has a strong affinity towards nucleic acids [77-80]. Amino acids 177-198 in BV1 of *Squash leaf curl virus* (SLCV) contain leucine-rich region associated with nuclear export referred to as nuclear export signal (NES), and the region is conserved in other organisms. It was shown that the TFIIIA protein of *Xenopus* could complement the NES of BV1 [81]. The mechanism of nuclear export involves the interaction of NSP with histone 3 (H3) forming a complex of H3-NSP-viral DNA complex resulting in the translocation of viral DNA out of the nucleus [82]. BV1 is also pathogenicity determinant. It interacts with key host factors for successful pathogenesis and BV1 or NSP from *Cabbage leaf curl virus* (CaLCuV) is the much studied one towards that aspect. BV1 of DNA-B suppresses the host immunity by inducing expression of *Asymmetric leaves 2* (AS2) and translocating it out of the nucleus to the cytoplasm. AS2 activates decapping enzyme DCP2 in the cytoplasm resulting in compromising host defense against the virus [83]. BV1 of the same virus also mimics MYC2 transcription factor and suppress terpene biosynthesis making the host attractive to the vectors enhancing the chances of transmission of the virus [84]. It also interacts with a protein named as nuclear shuttle protein interactor (NSI) from *Arabidopsis* which generally acetylates histones in the host. The interaction mapped to 38 amino acids close to NES in NSP and 88 amino acids from position 107-194 in NSI. NSP-AtNSI interaction lead to acetylation of coat protein bound to viral DNA. Then, NSP-AtNSI complex replaces CP from viral DNA and exports it out of the nucleus. It was also shown that overexpression of *AtNSI* enhances viral titer. The virus with mutant NSP that

abolishes NSI-AtNSP interaction but not DNA binding, nucleocytoplasmic shuttling and CP interaction when expressed in plant show defective in symptom severity and delayed systemic spread of the virus. This stresses the essentiality of NSP-AtNSI interaction [79, 85, 86]. NSP interact with some group of kinases referred to as NIKs (NSP interacting kinases) NIK1, NIK2, and NIK3 which are leucine-rich repeat-receptor like kinases (LRR-RLK). NIKs are localized to membranes. NSP-NIK interaction leads to suppression of kinase activity of NIKs. Loss of function of NIKs supported viral replication of CaLCuV. Therefore one can assume that NSP compromise host immunity by interacting with NIKs [87]. Another host factor that interacts with NSP of CaLCuV is NsAK (NSP Associated Kinase), which is a proline-rich extension like receptor protein kinase (PERK). NsAK shown to phosphorylates NSP, loss of NsAK decrease viral titer [88]. *Tomato leaf curl New Delhi virus* (ToLCNDV) NSP shown to induce a hypersensitive response in *N. tabacum* and tomato through the N-terminal region of the protein whereas the same protein induces leaf curling symptoms similar to that of viral symptom in *N. benthamiana* [89].

### **2.6.2 BC1 (BL1):**

It also referred to as movement protein (MP) as it is involved in the cell-cell movement of viral DNA by increasing the size exclusion limit of plasmodesmata of mesophyll cells [78]. BC1 localized as small punctuate bodies in the periphery of the cells and around the nucleus. In the sink leaves of the host, BC1 forms a disc-like structure in the periphery. When BC1 inoculated with its cognate DNA-A and DNA-B it forms needle like structures in sink leaves. It was also shown it translocates nucleus localized NSP to plasma membrane, thereby assisting NSP to transport viral DNA from cell-cell [90, 91]. BC1 from different viruses have varied, and non-specific affinities for DNA, BC1 from SLCV and *Bean dwarf mosaic virus* (BDMV) exhibited weak affinity of single stranded DNA whereas BC1 from MYMIV displayed strong affinity to single-stranded DNA. BC1 shows preference to forms of specific forms of DNA [77, 80, 92].

Attempts to generate transgenic plants with BC1 always resulted in plants with an abnormal phenotype resembling the symptoms of the virus as well as resistance against viruses. BC1 from *Tomato mottle virus* (TMoV) and BDMV are examples of such attempts. Interestingly, BDMV infection is generally symptomless in tomato, whereas BDMV-BC1

overexpression induced viral symptoms like phenotype [93, 94]. The plants expressing mutated forms of TMoV BC1 which don't produce viral like symptoms show resistance against TMoV and CaLCuV which could be because of a transdominant negative effect [93]

## **2.7 Satellite molecules:**

Monopartite begomoviruses are often reported to be associated with satellite DNA molecules which include alphasatellites, betasatellites, and newly discovered deltasatellites. Satellite molecules rely on the helper viruses for their replication and movement [95, 96].

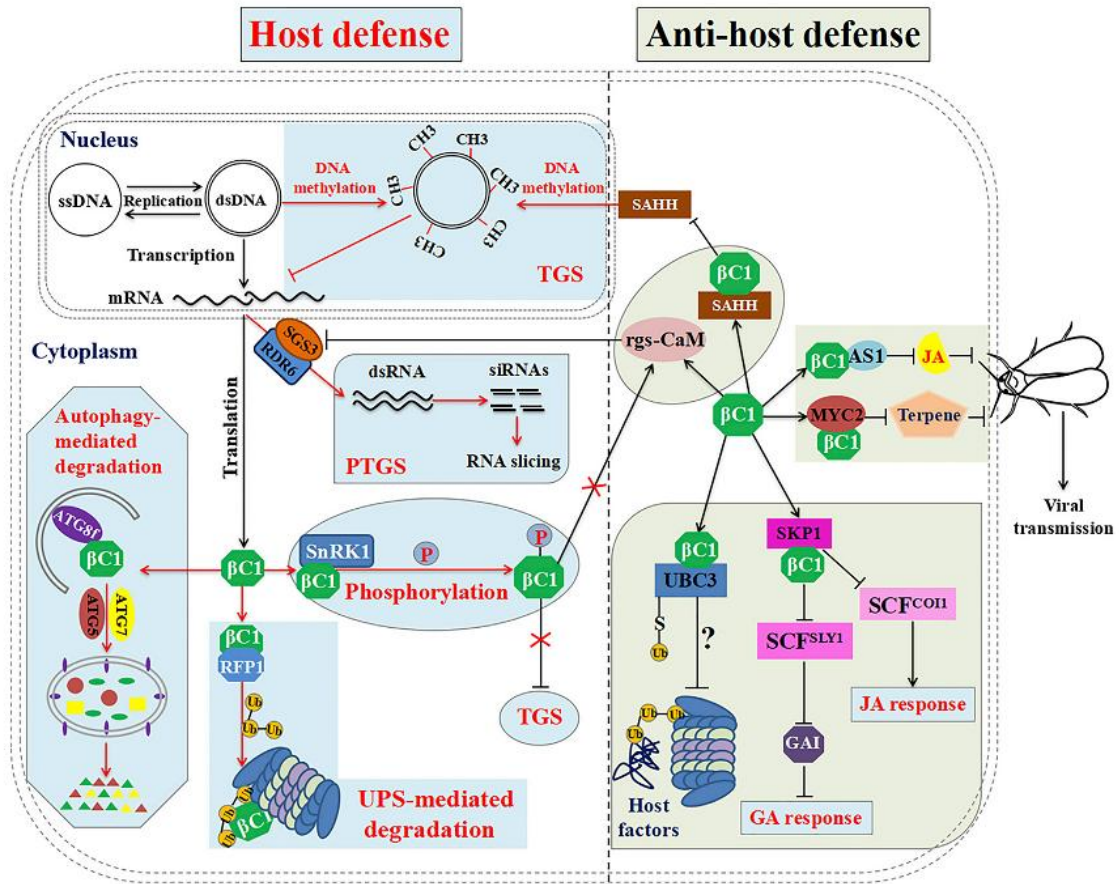
### **2.7.1 Alphasatellites:**

Alphasatellites are of size approximately 1375 nt with an A-rich region of size ranging from 150 to 200 nt, a hairpin loop containing the conserved nonanucleotide for replication and an ORF encoding Rep referred as alpha-rep of size approx 37 kDa. It is because alphasatellites encode their own rep they are not considered as strict satellites. Alphasatellites doesn't contribute to symptom development and pathogenicity [95].

### **2.7.2 Betasatellites:**

These satellites are essential for helper virus as they function as pathogenicity determinants during pathogenesis. These satellites are known to induce symptoms such as leaf curling, enations, yellowing etc. The sizes of these molecules are approximately 1350 nt and these molecules doesn't possess sequence similarity with the helper virus other than the conserved nonanucleotide. Betasatellites has highly conserved region named as satellite conserved region (SCR), an A-rich region and a single ORF called as  $\beta$ C1 encoded by the complementary strand.  $\beta$ C1 is a multitasking protein playing crucial a role during pathogenesis by suppressing host TGS and PTGS, suppressing host defense and promotes symptom development [95, 97].





**Figure 2.3: Schematic representation of overview of recent advances in the understanding of host defense and counter defense exhibited by geminivirus- $\beta$ C1.** Arrows indicate positive regulation, lines with bars indicate negative regulation. GA – gibberellic acid, JA – jasmonic acid, TGS – transcriptional gene silencing, PTGS – post-transcriptional gene silencing [97].

Betasatellites impart typical symptoms like vein clearing, leaf crinkling, leaf curling, enation, and stunting, etc.  $\beta$ C1 being the only ORF encoded by betasatellites will be the obvious reasons for betasatellite associated symptoms. Few reports explore the mechanism of symptom induction by  $\beta$ C1 the symptom by phenocopying AS2, interacts with AS1 replacing AS2 interfering normal leaf development. AS1/AS2 interaction is essential for down regulation of miR165/166 as well as up-regulation of HD-ZIP III transcription factor which are essential for normal leaf development [98].  $\beta$ C1 also induce vein clearing or yellowing in veins, which is because of its interference with the chloroplast metabolism.  $\beta$ C1 from *Radish leaf curl virus* (RaLCV) was demonstrated to be localized in chloroplast, induces damage in chloroplasts and

alters the ultra-structure of chloroplasts as well the expression of chloroplast encoded proteins resulting in chloroplast with hampered photosynthesis. Moreover, it was also shown very recently that RaLCV  $\beta$ C1 interacts with oxygen evolving enhancer protein 2 (PsbP) and interfere with the binding of PsbP binding to RaLCV DNA. Binding of PsbP with RaLCV DNA is essential for plant immunity against the virus as it is clearly seen that silencing *PsbP* transiently enhances viral load [99, 100].  $\beta$ C1 mess up with host ubiquitin system for the symptom induction.  $\beta$ C1 interact with an ubiquitin conjugating enzyme from tomato designated as SIUBC3, as a result, the level of polyubiquitinated proteins in betasatellite infected plants goes down compared to that of wild type. The myristoylation like motif GMDVNE at the C-terminal of the  $\beta$ C1 is essential for the interaction with SIUBC3. *N. benthamiana* and *N. tabacum* plants infected with mutated  $\beta$ C1 at the myristoylation like motif had compromised symptom development which signifies the interaction between in  $\beta$ C1 in symptom induction [101].

$\beta$ C1 nullify antiviral defense of plants by suppressing TGS and PTGS.  $\beta$ C1 from *Tomato yellow leaf curl China Betasatellite* (TYLCCNB) shown to complement BCTV C2 and reverse TGS by interacting with an essential member of methyl cycle, SAHH (S-adenosyl homocysteine hydroxylase) which is replenishing methyltransferase cofactor S-adenosyl methionine (SAM).  $\beta$ C1-SAHH interaction leads to a decrease in SAM [102]. Later it was shown that Tyr at position 5 and 110 is essential for the reversal of TGS activity [103] In another way,  $\beta$ C1 suppress host silencing by up regulating the silencing suppressor, one such silencing suppressor is calmodulin like protein rgs-Cam which suppresses the expression of *RDR6*. Overexpression of *Nbrgs-Cam* resulted in a phenotype similar to that of  $\beta$ C1 infected plants [104].

$\beta$ C1 also shown to alter the proteasomal pathway in the host by which it can escape from proteasomal degradation by the host. TYLCCNB  $\beta$ C1 interact with a RING-finger protein from tobacco (NtRFP1) which is an E3 ubiquitin ligase targets  $\beta$ C1 to the 26S proteasomal pathway by ubiquitination [105]. TYLCCNB  $\beta$ C1 interact with Sucrose-non fermenting1-related kinase from tomato (SISnRK1). SISnRK1 phosphorylates  $\beta$ C1 at residues serine at 33 and threonine at position 78 resulting in inactivating  $\beta$ C1 mediated pathogenicity.  $\beta$ C1 with phosphomimic mutants at the abovementioned residues and overexpression of SISnRK1 resulted in the delay in viral symptom as well as decrease in viral titre [106]. SISnRK1 also phosphorylates tyrosine at

position 5 and 110 which results in attenuated disease symptoms, a weakened reversal of TGS as well as abolished interaction with AS1 [103]

Autophagic pathways were shown to impart antiviral defense in the host by targeting  $\beta$ C1. ATG8 an autophagy-related protein interacts with *Cotton leaf curl Multan Betasatellite* (CLCuMB)  $\beta$ C1 via amino acid residue Valine at position 32. V32A mutation at  $\beta$ C1 enhanced symptom severity and viral DNA accumulation. Interestingly, silencing of other autophagy-related proteins ATG5 and ATG7 made plants susceptible to different viruses [107].

### **2.7.3 Deltasatellite:**

These are a novel class of satellites found associated with new world bipartite begomovirus infecting plants from *Malvaceae* family. The size of the deltasatellites is 630-750 nt, which is nearly one-fourth of the begomoviral genome and half of the size of alpha and betasatellites. These satellites do not code for any ORF, has a stem loop structure with nonanucleotide and an A-rich region. These satellites possess an additional stem-loop structure between A-rich sequence and SCR [96, 108].

## **2.8 Role of host factors in geminiviral pathogenesis:**

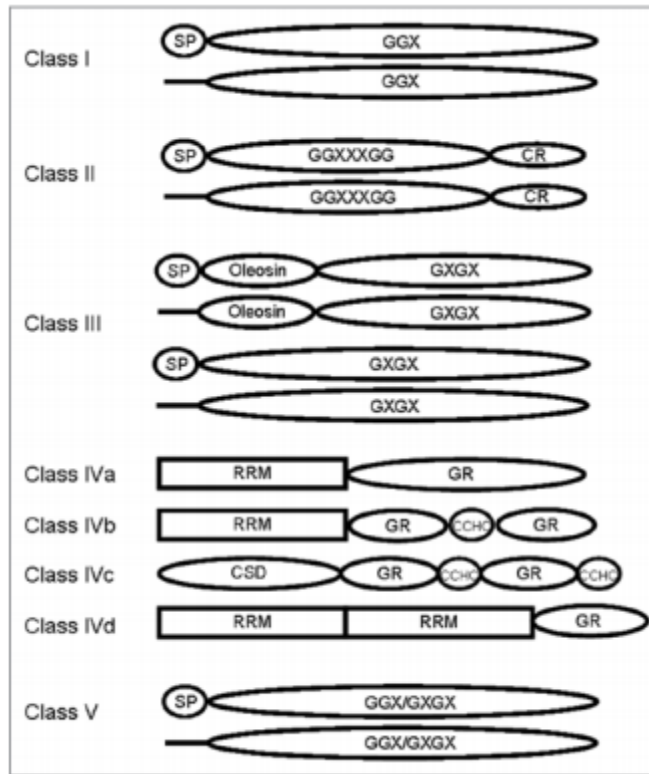
Like any other viruses, geminiviruses has limited coding capacities; therefore they are heavily relying on the host factors for their replication, suppression of host defense, assembly and release. For that aforesaid purpose viral proteins are multitasking in nature. Because of their artistry in subverting the host cellular processes, geminiviruses are refereed as masters in reprogramming and redirecting host cellular processes (Hanley-Bowdoin et al 2013). Role of host factors is crucial for successful infection. Few examples suggesting the importance of host factors in geminivirus pathogenesis and its impact in establishing either successful defense or successful infection was discussed in this review.

### **2.8.1 Glycine rich proteins (GRPs):**

Glycine rich proteins (GRPs) are superfamily of proteins characterized by the presence of semi-repetitive motifs that are rich in amino acid glycine. Initially, this protein was thought to be

involved in protein-protein interactions like that of keratins and other intermediary filament proteins [109].

### 2.8.2 Classification of GRPs:



**Figure 2.4: Diagrammatic representation of classification of GRPs into different classes.** SP - signal peptide; RRM – RNA Recognition Motif; CSD – Cold shock domain; GR – Glycine rich region; GX represents GR region where G is glycine and X is any amino acid. Source: [110].

The first GRP was reported in the year 1986 from petunia [111]. Since 1986 numerous publications reported with the name glycine rich protein but the proteins had different domain architectures, but the common feature of all the proteins was the enrichment of glycine in their amino acid sequence. The first categorization of plant GRPs was published in the year 2000 [109] and further elaborated in 2010 [110]. They attempted to categorize GRPs reported so far till the year 2000 into different categories, which led to the classification of GRPs into five classes. Class I GRP has simple sequence morphology with GGX repeats with or without signal peptide in its N-terminal. Class II GRP is similar to that of Class I with additional cysteine-rich

domain in its C-terminal. GR domain has amino acids repeated as GGXXXGG. Class III GRP has signal peptide followed by oleosin domain followed by GR motif with GXGX repeats. Class IV is quite complicated than other classes with four subclasses designated as a, b, c, and d. Class IV is essential as well as the members of this class were reported to interact with nucleic acids and shown perform numerous biological functions. Class IVa has single RNA Recognition Motif (RRM) followed by GR motif. Class IVb has single RRM followed by intermittent GR domain with Zinc finger motif. Class IVc has Cold Shock Domain (CSD) at its N-terminal followed by GR domain and Zinc finger motif present in tandem. Class IVd has two RRM followed by single GR domain. Class V is more like Class I; the only difference is in the complexity of the GR domain.

This particular review will focus only on this particular class of GRP as the gene under study belongs to this particular class. For the sake of convenience, class IVa GRP mentioned as GRP. Therefore, in the upcoming topics of this review GRP refers class IVa GRP.

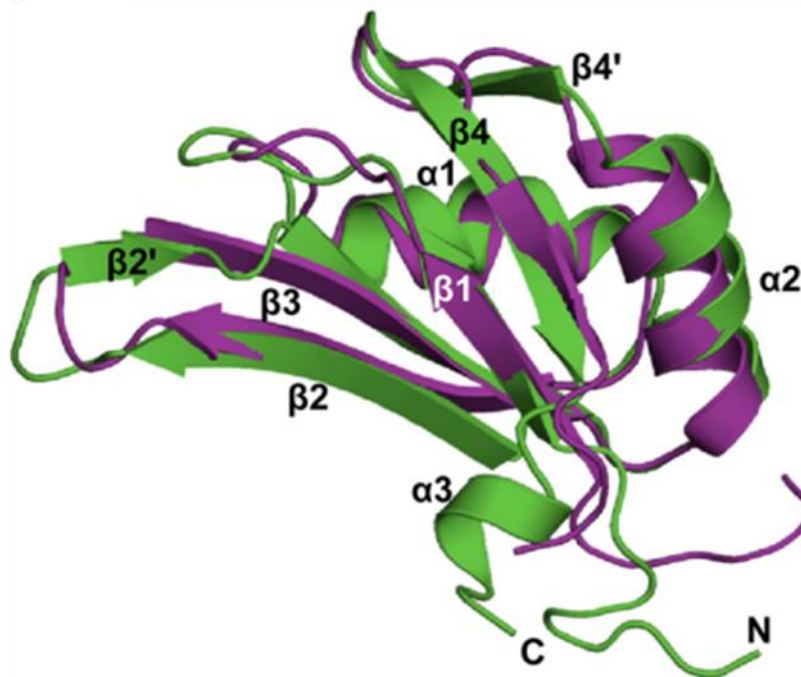
### **2.8.3.1 Structural organization of Class IVa GRP:**

As discussed earlier, in the beginning, Class IVa GRP structurally has two distinct domains: a structured N-terminal part containing RNA Recognition Motif (RRM)/RNA Recognition Domain (RRD) and unstructured C-terminal Glycine-Rich domain (GRD)/Glycine-Rich Motif (GRM).

#### **2.8.3.1.1 N-terminal RRM:**

It is one of the most commonly found domains in RNA binding proteins. It has 4 antiparallel beta sheets and two alpha helices arranged in the order  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$ , with beta sheet 3 and 1 interacting with nucleic acids referred to as ribonucleoprotein 1(RNP1) and ribonucleoprotein 2 (RNP2). The basic amino acid from RNP1 and aromatic amino acid from RNP2 shown to interact with the phosphate group and nitrogen base of RNA in case of U1A RRM [112]. The structure of Class IVa GRP was first published by two independent groups in the same year. Nuclear magnetic resonance (NMR) spectroscopy was used by both the groups to solve the structure of two GRPs from two different organisms: NtGRP1 from *N. tabacum* and HvGRP1 from *H. vulgare*. Interestingly, both the structures had similar properties. Like any

other RRM interaction with nucleic acids happens by base/aromatic acid interactions and electrostatic interaction with the phosphate group [11, 12]. Before that, the structure was computationally predicted and key residues identified using sequence alignment studies. Based on the sequence alignment with U1A and hnRNPA1 secondary structures mapped in AtGRP7 from which critical residues that were essential for RNA binding tracked to arginine at position 47 and 49 of RNP1 mutating them shown to reduce the RNA binding affinity [113, 114].



**Figure 2.5: Overlay of solution structure of HvGRP1 and NtGRP1 solved by NMR.** Structure of NtGRP1 (PDB accession no: 4C7Q) and HvGRP1 (PDB accession no: 2MPU) overlaid. HvGRP1 represented in green color and NtGRP1 represented in cyan color [11].

#### 2.8.3.1.2 C-terminal GR domain:

C-terminal is highly unstructured and flexible with nearly 30 % amino acid comprising of amino acid glycine. Even though this part is highly unstructured, but it has some meaningful repeats acknowledged by different groups. The repeats have glycine interspersed with aromatic and charged amino acids such as GGYGG, GYR, and RGG motifs [11, 115-117]. The length of the C-terminal region seems to vary among the different class IV proteins within *Arabidopsis*. AtGRP7 has lengthy C-terminal region than AtGRP4 and AtGRP2 [118]. Very recent studies in

different RNA binding proteins suggest that most RNA binding proteins have such intrinsically disordered proteins influencing RNA-protein interaction specifically and non-specifically. Such disordered regions are also influential in their post-transcriptional function and other function of proteins, disruption of such regions associated with diseases. In spite of such significance still, their precise nucleic acid binding properties, regulation, and biological significances remain poorly understood [119]. No study has done so far in Class IV GRP exploring the disordered region of Class IV GRP except a very recent study in phytopathogen *Magnaporthe oryzae*. *MoGRP1* is very essential for the normal development of the fungus as well for establishing pathogenesis. MoGRP1 has 4 RGG/RG motifs in the C-terminal of the protein which was found to be essential for the function of the protein. Deletion of the RGG/RG motifs from MoGRP1 resulted in lack of infection in rice along with developmental abnormalities. Deletion also altered the nuclear localization of the protein [120].

### **2.8.3.2 Biophysical and Biochemical properties:**

Class IVa GRPs from different plant species doesn't always behave the same way. It is a well-established fact that AtGRP7 and AtGRP8 from *Arabidopsis* bind to their own RNA specifically at polypyrimidine tract at 3'UTR and intron to regulate them [113]. It exhibits a strong affinity towards oligodeoxyribonucleotide (ODN) and also binds to oligoribonucleotide (ORN) without forming any secondary structures. AtGRP7 cannot bind to double-stranded form [121]. GRP from *Phycomitrella patens*, *Nicotiana sylvestris* and *Capsicum annum* show affinity to poly (U) and poly (G). GRP from *N. sylvestris* shows preferential binding to ssDNA at low salt concentrations whereas at high concentrations preference shifts to RNA molecules [116, 122, 123]. AtGRP4 exhibits non-specific binding towards nucleic acids, apart from nucleic acid binding it exhibits protein chaperone activity too. AtGRP4 has close sequence similarity to AtGRP7 except for the 25 amino acid disordered region at N-terminal region before RRM, Protein chaperone activity of AtGRP4 was attributed to the N-terminal disordered region [124]. An interesting feature of Class IVa GRP is the weak, transient intermolecular associations between two molecules. Such intermolecular associations were evident from NMR spectroscopy and EMSA. Both full-length HvGRP1 and NtGRP1 exhibited intermolecular associations resulting in aggregation at higher concentrations, which was evident from broadened NMR

signals. NMR signals also showed that the intermolecular interaction observed is between the RRM and GR region. The truncated protein lacking the 25 residues from the C-terminal GR region loses the intermolecular interaction. It was also found the interaction between GR and RRM is disrupted by varying temperatures and pH conditions, suggesting the nature of the interaction to be weak but dynamic. NtGRP1 form high state oligomers with the nucleic acids which was assumed because of inter and intramolecular interaction occur between GR and RRM. Interestingly, thioredoxin (which doesn't interact with nucleic acids) when fused with HvGR region shown to bind with ssRNA nonspecifically, which may be because of GR region and it is a important piece of evidence to show that GR region participate in the nucleic acid binding [11, 12].

### **2.8.3.3 Role of GRP in biotic stress:**

GRPs influence plant defense directly and indirectly. Studies show that RNA binding GRPs can influence the expression levels of defense-related genes positively and negatively [17, 18, 125]. It was shown how AtGRP7 could influence pathogenesis both in a positive and negative manner against diverse pathogens. Upon inoculation of necrotrophic bacteria *P. carotovorum* SCC1, AtGRP7 overexpression plants had less cfu (colony forming unit) at 3 dpi (days post inoculation) than the wild type and knock out lines in *Arabidopsis*. In contrast, vice versa was observed with necrotrophic fungus *B. cinerea* where overexpression line had more cfu than that of wild type and knockout line at the same 3 dpi. On the other hand, AtGRP7 overexpression lines in tobacco showed less necrotic lesions than that of wild type upon Tobacco mosaic virus infection. This particular study showed multifaceted roles of Class IVa GRP in the context of plant pathogenesis. An attempt to find the correlation between AtGRP7 with MAPK pathway was also made but AtGRP7 doesn't seem to alter the transcript level of MAPKs involved in plant defense (*MPK3*, *MPK4*, and *MPK6*); therefore AtGRP7 mediated defense is independent of MAPK pathway [15]. An important study by Staiger's group established an important role of AtGRP7 in salicylic acid and jasmonic acid mediated defense. It was found that overexpression line in *Arabidopsis* had the constitutive expression of salicylic acid-dependent transcripts such as *PR-1* (*Pathogenesis-Related-1*), *PR-2* and *PR-5* and conferred resistance against *Pseudomonas syringae* Pto3000 in *Arabidopsis*. Supporting their observation, AtGRP7



knock out lines were all susceptible and had *PR* transcript level similar to that of wild type whereas the overexpression lines were all restricting the pathogen [18]. AtGRP7 had an indirect influence of defense related genes that are having a role in JA/ET signalling such as *PDF1.1*, *PDF1.2a*, *RAP2.3/ERF72/AtEBP*, and *THI2.2* were down-regulated in overexpression line [17]. It was demonstrated that AtGRP7 physically interact with *PDF1.2* regulating its expression post-transcriptionally [18]. Recently, it was shown *AtPDF1.2* mediates defense against *P. carotovorum* subsp. *carotovorum* in *Arabidopsis* [126]. The mechanism of Class IVa GRP in HR was elucidated by [123]. Class IVa GRP from *Capsicum annum* CaGRP1 interacts with Receptor-like Cytoplasmic protein kinase (CaPIK1) from *C. annum* which is a positive regulator of HR-mediated defense. CaPIK1 interacts and phosphorylates CaGRP1, and the complex translocates to the nucleus which leads to the suppression of expression of PR genes and attenuation of CaPIK1 mediated cell death and defense responses. CaGRP1 promote the growth of *Xanthomonas campestris* pv. *vesicatoria* (Xcv) whereas CaPIK1 restrict the growth of Xcv. Silencing of *CaGRP1* made the plants resistance towards Xcv with increases electrolyte leakage, H<sub>2</sub>O<sub>2</sub> generation, and increased expression of defense related genes in pepper leaves [123].

To establish the suppression of host defense and establishment of successful infection, pathogens have adopted various strategies. One such strategy is the injection of effector molecules in the host, which target essential host factor and subvert them for successful colonization and infection [127]. HopU1 is a type III effector protein secreted by *Pseudomonas syringae* DC3000 capable of inducing a non-host hypersensitive response. HopU1 is a mono-ADP-ribosyl transferase protein. Targets of HopU1 identified so far were RNA binding proteins, and one such in the list is AtGRP7. Experiments demonstrate that HopU1 specifically interacts and targets arginine at position 47 and 49, ADP-ribosylates the residues. ADP-ribosylation AtGRP7 was unable to bind to its target RNA molecules [16]. Interestingly, the arginine at position 49 is the first amino acid at RNP-1 RRM at N-terminal of AtGRP7. RQ mutation at 49 impairs electrostatic interaction with nucleic acids, thereby affecting RNA binding property of AtGRP7 to seven-fold lesser than that of wildtype protein. Constitutive ectopic overexpression of *AtGRP7*-RQ mutants had defective alternative splicing [113]. *AtGRP7* knock out plants were all susceptible to *P. syringae*. This shows how an effector molecule secreted by a pathogen cleverly targets essential RNA binding proteins in the host and quells plant immunity by

impairing its RNA binding activity [16, 127]. The above examples stress the direct and indirect role of *AtGRP7* in plant defense.

#### **2.8.3.4 Role of GRP in abiotic stress:**

Because of numerous efforts on GRP, it is now a well-established fact that GRP has a tremendous role in abiotic stress. Majority of the class IVa GRPs reported so far have a role in imparting cold stress tolerance. Class IVa GRP, unlike class IVc GRP lack CSD but still provides cold stress tolerance by acting as RNA chaperone [12, 118, 128]. *E. coli* BX04 cells which lack 4 essential cold shock proteins when transformed with *AtGRP7* had better survival rate at 17 °C than vector transformed, and N-terminal RRM was found to be responsible for the cold stress tolerance [118]. Upon further exploration, *AtGRP7* conferred cold stress tolerance in *Arabidopsis* by increasing the stomatal aperture, whereas the reverse was observed for dehydration stress and saline stress. *AtGRP7* overexpression seedlings had poor germination and growth when subjected to drought and salinity stress [129]. Similar observations were reported in rice too. *OsGRP4* and *OsGRP6* could complement BX04 cold tolerance like that of *AtGRP7*. The genes could complement cold tolerance in *grp7* knockout lines as well. OsGRP4 and OsGRP6 promote the export of mRNA from the nucleus to cytoplasm in *grp7* knockout plants under cold stress [130].

Cold stress tolerance of GRP can also lead to potential applications. Overexpression of *AtGRP2* and *AtGRP7* in rice has improved the grain yield [131]. *Camelina sativus* L., which is an oilseed crop has many GRPs upregulated during cold stress, among the *CsGRPs*, *CsGRP7a* when overexpressed in *Arabidopsis* and *Camelina*, made plants tolerant to cold or freeze stress. [132, 133]. *Cucumis sativus* fruits upon pre-storage cold acclimatization had some GRPs upregulated, one among them was *CsGRP3* whose transcript level was significantly up-regulated, *CsGRP3* overexpression lines in *Arabidopsis* had better growth rate at ambient temperatures, the better survival rate at lower temperatures than that of wild type. The transgenic lines also had higher levels of ROS scavengers and lower level of ROS [134]. *LeGRP1a* from tomato was shown to enhance the protein content in tomato fruits under cold shock conditions [128].

### 2.8.3.5 Role of GRP in flowering:

Flowering in plants is mediated by photoperiodism, vernalization and autonomous pathway. RNA binding GRPs were initially identified as a slave oscillator in plant circadian rhythm, but later studies provided their versatile roles in numerous biological processes. Their role in flowering came up with the observation that loss of function mutants and overexpression lines of *AtGRP7* exhibited interesting flowering phenotypes in *Arabidopsis*. An *atgrp7* T-DNA insertion line and RNAi lines suppressing *AtGRP7* exhibited late-flowering phenotype in both long day and short day plants whereas the overexpression lines displayed early flowering phenotype compared to that of wild type plant. Interestingly, the T-DNA insertion line and RNAi lines had elevated levels of *FLC* (*Flowering Locus C*), a MADS-box transcription factor known for suppressor of flowering through a conserved motif. Overexpression lines had suppressed levels of *FLC* transcripts. Neither deletion nor overexpression of *AtGRP7* has not affected the photoperiodism and vernalization could reverse the late-flowering phenotype of *AtGRP7* compromised lines. These pieces of evidence support the role of *AtGRP7* in the autonomous pathway of flowering [135-137].

Vernalization is a process in which plants are exposed to cold to promote flowering. RNA binding GRPs were shown to influence vernalization. *TaVRN1*, an AP1 clade MADS-box transcription factor which is an activator of vernalization. Under non-vernalization conditions, *TaGRP2* (RNA binding GRP from wheat) bind to pre-mRNA of *TaVRN1*, thereby preventing their accumulation. It was shown that during vernalization increases O-GlcNAc signaling. O-GlcNAc signaling results in posttranslational modification of *TaGRP2* and O-GlcNAc Acylation of *TaGRP2* interact with Jacalin-lectin binding protein *VER2*. *VER2-GRP2* interaction results in translocation of *TaGRP2* from the nucleus to cytoplasm, making the pre-mRNA of *TaVRN1* free for mRNA accumulation [136]. Another study from the same group demonstrated the role of RNA-binding glycine rich protein in flowering. It was shown that a Jacalin-lectin from *Arabidopsis* Jacalin-lectin like1 (*AtJAC1*) influences the flowering by modulating the localization of *AtGRP7*. *AtGRP7* was shown to interact with the antisense strand of pre-mRNA of *FLC*. Loss of function of *AtGRP7* elevates the transcript abundance of *FLC*. *AtJAC1* modulates the localization pattern of *AtGRP7*. *AtJAC1* overexpression leads to localization of

AtGRP7 in the cytoplasm, whereas *AtJAC1* deletion leads to localization of AtGRP7 in the nucleus. Interestingly, *AtJAC1* overexpression plants had a delay in flowering whereas *AtJAC1* deletion lines had precocious flowering. From this, one can assume the significance of glycine rich proteins in shaping plant development [137].

#### **2.8.3.6 Role of GRP in hormone signaling:**

Numerous studies have shown the influence of GRP essential plant hormones. The first GRP that was isolated as salicylic acid-inducible gene [138]. The second report supporting the first observation on the role of RNA binding GRP in Salicylic acid (SA) signaling came in 1998, where 3 hours after the exogenous application of SA elevated the expression of *NgGRP* (a class IVa GRP from *Nicotiana glutinosa*) reaching the maximum at 6 hours [125]. Later studies revealed that defense-related genes such as *PR1*, *PR2*, and *PR5* which are associated with SA signaling were all upregulated in *AtGRP7* overexpression line (*AtGRP7-ox*) making them to restrict the growth of virulent pathogenic bacteria whereas the knockout line showed levels similar to that of wild type and susceptible the virulent pathogen [17, 18]. *AtGRP7-ox* also had increased levels of SA and SAG (SA O- $\beta$ -glucoside). Interestingly, *sid2-1 AtGRP7-ox* plants had no increase in elevation of *PR1* transcript level, but *PR2* and *PR5* level seem to be same to that of wild type, *SID2-1* gene is involved in the biosynthesis of SA. *PR2*, *PR5* levels compromised in *NahG AtGRP7-ox*. Upregulation of *PR* transcripts, elevation SA and SAG levels, resistance to the virulent pathogen in *AtGRP7-ox* plants clearly indicate the role of *AtGRP7* in SA pathway mediated defense. Interestingly there was no direct interaction of *AtGRP7* with *PR* transcripts observed *in vivo* whereas upregulation is mainly because of the increase in promoter activity of *PR* transcripts in *AtGRP7-ox* plants (Hackmann et al., 2014).

GRPs were also induced upon wounding [139], as it is a known fact that JA mediated defense is associated with wounding. *AtGRP7* overexpression lines had decreased levels of JA/ET associated transcript *PDF1.2*. *AtGRP7* exhibited direct interaction with *PDF1.2* transcript from which one can assume that *AtGRP7* being a protein known for their role in post-transcriptional modification influences *PDF1.2* levels post-transcriptionally [18]. Jasmonic acid induced transcript *PDF1.2a* did not recover to normal level in the *AtGRP7* overexpression line even after treatment with methyl jasmonate suggest the indirect influence of *AtGRP7* on JA signaling [17].

The most visible phenotype of *AtGRP7* overexpression was their stunted growth, which reverted to the normal height upon treatment with GA<sub>3</sub>. Global transcriptome studies in the overexpression lines revealed that the transcript level of enzymes involved in the biosynthesis of gibberellic acids was altered. The transcript level of enzymes *GA2* and *GAI* catalyzing the first two steps in GA biosynthesis were reduced. As a result, the level of bioactive GA<sub>4</sub> was reduced strongly in overexpression line, precursors of GA<sub>12</sub>, GA<sub>15</sub>, GA<sub>24</sub>, and GA<sub>9</sub> [140].

When we talk about plant hormones, one cannot ignore abscisic acid (ABA), which is an essential hormone in case of various environmental stresses such as salinity, drought, and freezing tolerance, etc. GRP from *Sorghum bicolor*, whose transcript level was upregulated 3 fold upon 2 hours of ABA treatment and 7 fold upon 6-6 hours of ABA treatment [141]. In contrast, the GRPs from other organisms showed a negative correlation. *AtGRP7* was repressed by ABA, T-DNA insertion line of *AtGRP7* displayed hypersensitivity to ABA with higher accumulation of ABA-stress inducible transcripts *RD29A* and *RAB18* [142], and *AtGRP7* overexpression lines had ABA-responsive genes with altered levels using microarray studies [17].



# *Materials and Methods*





### 3. Materials and Methods

#### 3.1 Materials:

##### 3.1.1 Seeds and plant material:

Seeds of *Nicotiana benthamiana* used in the current study was procured from Central Tobacco Research Institute (CTRI), Rajahmundry, Andhra Pradesh. Seeds of *N. benthamiana* were sown in the soilrite mix, grown in insect-proof glass-house/sterile culture room under 16h light/8h dark cycles of photoperiodic lighting.

##### 3.1.2 Infectious viral clones:

Two infectious viral clones used in the current study i.e., genomic components of *Tomato leaf curl New Delhi virus* DNA – A (ToLCNDV-A) (Accession no. U15015) and *Radish leaf curl betasatellite* (RaLCB) (Accession no. EF175734) were available in our laboratory.

##### 3.1.3 Bacterial strains:

The following bacterial strains were used in the present work; their genotype described below:

**Table 3.1:** List of different bacterial and yeast strains used in the current study

Strain	Genotype	Purpose
<i>Escherichia coli</i> DH5 $\alpha$	F <sup>-</sup> <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG</i> $\phi$ 80 <i>dlacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U169, <i>hsdR17</i> (r <sub>k</sub> <sup>-</sup> m <sub>k</sub> <sup>+</sup> ), $\lambda$ <sup>-</sup>	Cloning purpose.
<i>E. coli</i> BL21 CodonPlus (DE3)-RIPL	F <sup>-</sup> <i>ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>dcm</i> <sup>+</sup> Tet <sup>r</sup> <i>gal</i> $\lambda$ (DE3) <i>endA</i> Hte ( <i>argU proL Camr</i> ) ( <i>argU</i>	For protein expression and purification.

	<i>ileY leuW Strep/Spec<sup>r</sup></i>	
<i>Agrobacterium tumefaciens</i> EHA105	C58 (rif <sup>r</sup> ) Ti pEHA105 (pTiBo542DtDNA) L- succinamopine type, supervirulent	Innoculation/infiltration of infectious viral clones and leaf disc assay.
<i>Saccharomyces cerevisiae</i> AH109	<i>MATa trp1-901 leu2-3, 112</i> <i>ura3-52 his3-200 gal4 gal80</i> <i>URA3::MEL1UAS-</i> <i>MELITATA-LacZ</i> <i>LYS2::GAL1UAS-GALITATA-</i> <i>HIS3 GAL2UAS-GAL2TATA-</i> <i>ADE2</i>	Yeast two hybrid assay.

### 3.1.4 Plasmid vectors:

**Table 3.2:** List of different plasmids used in the current study

	Source	Purpose
pJET1.2	Thermo Fischer Scientific, USA	Cloning of genes from cDNA or genomic DNA
pET SUMOadapt	[143]	For expression and purification of NbGRP3 and its C-terminal deletion variant.
pGAD-C1	[144]	Yeast two hybrid assay
pGBD-C1	[144]	Yeast two hybrid assay

### 3.1.5 Antibiotics and phenolic compound:

**Table 3.3:** List of antibiotics, and phenolic compound used in the current study

	<b>Stock concentration</b>	<b>Working concentration</b>	<b>Solvent</b>
Kanamycin	50mg/ml	50 µg/ml	Water
Ampicillin	100mg/ml	100 µg/ml	Water
Rifampicin	30mg/ml	30 µg/ml	DMSO
Cefotaxime	100 mg/ml	250 mg/ml	Water
Acetosyringone	100 mg/ml	100 µg/ml	DMSO

### 3.1.6 List of primers:

**Table 3.4:** List of the primers used in the current study:

<b>Name</b>	<b>Sequence (5' – 3')</b>	<b>Length (nts)</b>	<b>Purpose</b>
NbGRP3 qRT FP	GAATACAGGTGCTTCGTCGG	20	For detection of <i>NbGRP3</i> transcript by Real-Time PCR
NbGRP3 qRT RP	TCCCTTCAATAGCGTCCCTC	20	
NbGRP3 del FP	TCTGAAGGAAACTGGAGGAG	20	For generation of variant
NbGRP3 del RP	ACCGCCACCACCACGGTAAC	20	NbGRP3 $\Delta_{103-148}$

NbGRP3 SUMO FP	GACCATGAAGACTCCTAGAGCCATG GCAGAAGTTGAATACAGG	43	For cloning of <i>NbGRP3</i> full length into pET SUMOadapt
NbGRP3 SUMO RP	TCCCCCGGGGA ACTAACTCCTCCAG TTTCCTTC	33	
NbGRP3 RNAi Xba FP	CTCAGAATTACTTTGATTCATGTTGT A	27	For generation of VIGS construct
NbGRP3 RNAi BamHI RP	GGATCCATTTCCGTTGCCTTTAGAT	25	

### 3.2 Methods:

#### 3.2.1 *E. coli* competent cells preparation and their transformation:

The *E. coli* strains used in this study were streaked on LB agar plates with appropriate selection to obtain isolated colony and were grown overnight in incubator shaker at 37 °C. Single colony was picked with a sterile toothpick, inoculated in liquid LB medium and grown overnight. One ml of overnight grown culture acts as inoculum for the secondary culture, which was grown till O.D<sub>600</sub> reaches 0.5. The culture transferred to ice for 15-30 min followed by centrifugation (6000 rpm for 5 min at 4 °C) to harvest the cells as a pellet. The pellet resuspended in 30 ml of an ice-cold solution of CaCl<sub>2</sub>- MgCl<sub>2</sub> containing 6 ml of CaCl<sub>2</sub> (0.1 M) and 24 ml of MgCl<sub>2</sub> (0.1 M), kept in ice for 30 min and centrifuged (6000 rpm for 5 min at 4 °C). The cellular pellet stored in -80 °C after dissolving the pellet in a solution containing 225 µl of 100 % glycerol and 1275 µl of 0.1 M CaCl<sub>2</sub>.

During the transformation, an aliquot of cells thawed in ice followed by incubation with DNA of interest in ice for 30 min. The cells were subjected to heat shock for 90 sec, followed by the snap chill in ice for 5 min. Nearly 750 µl of LB medium added and kept for recovery in incubator shaker maintained at 37 °C for 1 hour at 220 rpm followed by plating on LB agar with antibiotics.

### **3.2.2 *Agrobacterium tumefaciens* competent cell preparation and their transformation:**

A single colony of *A. tumefaciens* EHA105 inoculated into 2 ml of liquid LB medium with appropriate antibiotics and grown overnight at 28 °C with 220 rpm. 1 ml of overnight grown culture inoculated into 50 ml liquid LB for the secondary culture, which was grown at 28 °C with 220 rpm till O.D<sub>600</sub> reaches 0.5. The culture was incubated on ice for 5 min before centrifugation at 7000 rpm for 5 min at 4 °C. To the cellular pellet, 10 ml of ice-cold 0.15 M NaCl was added for resuspension followed by incubated in ice for 15-20 min. The cells centrifuged (7000 rpm for 5 min at 4 °C) followed by resuspension again in 1 ml of ice-cold 20 mM CaCl<sub>2</sub>. The cells finally aliquoted, snap frozen and stored at -80 °C.

Transformation was performed by the freeze-thaw method [145]. Competent cells thawed on ice followed by incubation with plasmid DNA for 30 min on ice. The cells were snap frozen using liquid nitrogen for a min, and then thawed at 37 °C followed by incubation in ice and kept for 5 min. The recovery step was similar to that of *E. coli* transformation except for the incubation temperature and time, which was 28 °C for 4-5 h at 220 rpm.

### **3.2.3 Plasmid DNA extraction:**

The Plasmid DNA extraction was carried out by the alkaline lysis method [146]. Overnight grown *E. coli* culture harboring plasmid DNA of interest was harvested by high-speed centrifugation for 2 mins, the supernatant was discarded. 100 µl of solution I (25 mM Tris-Cl, pH-8.0, 50 mM glucose and 10 mM EDTA) added to the pellet and resuspended followed by addition of 200 µl freshly prepared solution II (0.2 N NaOH and 1 % SDS) and mixed gently. 150 µl of solution III (3 M potassium acetate, pH 5.5) was added, mixed well and kept in ice for 5 min. The solution mixture was centrifuged at 13,000 rpm for 10 min at 4 °C, supernatant collected in fresh 1.5 ml microcentrifuge tube. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed well by inverting the tubes followed by centrifugation at higher rpm for 10 min at 4 °C. The upper aqueous phase transferred to the fresh tube followed by isopropanol precipitation using 0.8 volume of isopropanol. The samples were centrifuged at higher rpm for 20 min at 4 °C followed by a wash with 70 % ethanol. The DNA pellet was brief dried to remove traces of ethanol and dissolved in 30 µl of sterile double distilled water.

### 3.2.4 Restriction digestion of plasmid DNA:

The restriction digestion of plasmid DNA was performed according to the enzymes and buffer, incubation condition as prescribed by the manufacturer's instructions. In the case of digestion using two enzymes, it was either performed sequentially or together in a single reaction depending on the buffer and incubation temperature compatibility of the enzymes used. Sequential digestion involves the following steps: digestion of plasmid DNA with enzyme-1, heat inactivation followed by phenol:chloroform: isopropanol precipitation or gel extraction, and finally treatment with enzyme-2.

**Table 3.5:** Reaction mixture for restriction digestion

Components	Quantity
Reaction buffer	1X
Restriction enzyme	5-10 U
DNA	500 ng -- 1µg
Sterile distilled water	For adjusting the final volume

The abovementioned components were taken in 1.5 ml microcentrifuge tubes, incubated in a water bath or dry bath at the recommended temperature.

### 3.2.5 Dephosphorylation of plasmid DNA:

This step is generally necessary to avoid re-circularization of plasmid vector digested with a single restriction enzyme during the process of cloning. Linearized plasmids which were purified by phenol:chloroform: isopropanol precipitation or gel extraction was subjected to dephosphorylation by treating with Antarctic phosphatase (which is an alkaline phosphatase from New England Biolabs, USA) for 30 min at 37° C followed by heat inactivation at 80° C for 2 min.

**Table 3.6:** Reaction mixture for alkaline phosphatase treatment

Components	Quantity
10X Antarctic phosphatase reaction buffer	1X
Linearised plasmid	500 ng -- 1µg
Antartic phosphatase (5U/µl)	5 U
Sterile distilled water	For adjusting the final volume

### 3.2.6 Ligation:

The concentration of purified vector and insert was checked on 1% agarose in 1X TAE followed by quantification using Nanodrop (Thermo Fischer Scientific, USA). Vector molecule and insert molecule were taken either in 1:3 or 1:5 ratios for ligation using T4 DNA ligase (Thermo Fischer Scientific, USA). The ligation mixture incubated at 16 °C overnight in ligation bath (Multitemp III, GE healthcare). The ligation mixture was transformed into competent cells of *E. coli* DH5α.

**Table 3.7:** Reaction mixture for ligation

Component	Quantity
Insert:Vector	1:3 or 1:5
10X ligation buffer	1X
T4 DNA Ligase (10U/ µl)	5U
Sterile distilled water	For adjusting the final volume

### 3.2.7 Polymerase Chain Reaction (PCR):

PCR was carried out for amplifying the DNA fragment of interest from either cDNA or genomic DNA or plasmid DNA using the DNA (depending on the purpose of the experiment) oligonucleotide pairs referred to as primers. The DNA to be amplified taken as the template molecule for the reaction. It involves repetitive cycles of events such as template denaturation, primer annealing and extension of primers by DNA polymerases. Apart from these cycles, PCR cycles also include initial denaturation and final extension.

**Table 3.8:** Reaction mixture for PCR

Component	Quantity
10 X reaction buffer	1 X
10 $\mu$ M Forward and reverse primers	0.4 $\mu$ M
2.5 mM dNTP	0.1 mM
Template DNA	Plasmid DNA (20 ng – 100 ng) or plant genomic DNA (100 ng to 500 ng)
DNA polymerase	0.5 U
Sterile distilled water	For adjusting the final volume

### 3.2.8 Prediction of disordered region in Glycine rich domain of NbGRP3:

NbGRP3 amino acid sequence pasted into the online tool <http://globplot.embl.de/>; the disordered propensity sum was calculated using the Russell/Linding propensity tool. The results were obtained in the form of a plot demonstrating the globular domains as green and disordered regions as blue.



### **3.2.9 Construction of expression clones of wild type *NbGRP3*:**

The coding region of *NbGRP3* was amplified from *N. benthamiana* cDNA using the specific primer sequence. The amplified product was then cloned into pET SUMOadapt vector (Life Technologies, CA) using sites *BsaI* and *HindIII*. The putative clones were verified by PCR followed by restriction digestion finally by sequencing (Seqlab, Germany).

### **3.2.10 *NbGRP3* expression and purification:**

pET SUMOadapt-*NbGRP3* construct was transformed into different expression strains of *E. coli*. Retransformed colony was verified by PCR and was inoculated into LB medium with Kanamycin (50 µg/ml) as a primary inoculum and grown overnight. The overnight grown primary culture inoculated into fresh LB medium with Kanamycin as a secondary culture. The secondary culture was grown till O.D<sub>600</sub> reaches 0.5, induced with IPTG and grown for 4 hours. For purification of *NbGRP3*, 6g of induced cells of BL21-codon plus (DE3) RIPL were lysed using lysis buffer (50 mM Tris-Cl pH-7.5, 100 mM NaCl, 10 % glycerol, 1mM TCEP) followed by a French press. The lysate was spun at 40,000 rpm for one hour to separate clear supernatant containing the soluble fractions and cell debris. The clear supernatant was applied to Ni-NTA column for separation of His-SUMO-*NbGRP3* from other cellular proteins. The His-SUMO-*NbGRP3* was applied to anion exchange column to separate the other contaminants from His-SUMO-*NbGRP3*. The purified His-SUMO-*NbGRP3* then subjected to SUMO protease treatment to cleave His-SUMO tag and the cleaved product was subjected to gel permeation chromatography and second Ni-NTA to separate the cleaved His-SUMO tag and *NbGRP3*.

### **3.2.11 Circular Dichroism (CD) measurements:**

Far UV- Circular Dichroism spectrum for purified full-length *NbGRP3* was recorded using Jasco J-815 spectropolarimeter. Protein of concentration 263.6 µg/ml was taken for measurement in usual lysis buffer (50 mM Tris-Cl pH-7.5, 100 mM NaCl, 10 % glycerol, 1 mM TCEP). Spectrum from the range 190 – 260 nm recorded using 0.1 mm cuvette. The spectrum was analyzed by considering the achiral nature of glycine, which accounts for nearly 30 % of the protein.

### **3.2.12 Analytical ultracentrifugation:**

Sedimentation velocity and equilibration measurements were performed at 128794g (40000 rpm) and 5152g (8000 rpm), respectively, at 20 °C in an XL-A analytical ultracentrifuge (Beckman Instruments, Inc., Fullerton, CA) using double-sector cells and an An50Ti rotor. To determine the actual state of NbGRP3, analyses were conducted using the purified protein at concentrations 50 µM, 200 µM and 600 µM in usual lysis buffer (50 mM Tris-Cl pH-7.5, 100 mM NaCl, 10 % glycerol, 1 mM TCEP). To determine the influence of ligand on the state of NbGRP3, 30 µM protein titrated against different concentrations of (dC)<sub>12</sub> such as 0, 10, 20, 30, 40 µM.

### **3.2.13 Gel based DNA binding assays or Electrophoretic Mobility Shift Assays:**

EMSA was performed using the following protocol. DNA oligos used in this study were initially radiolabelled with [ $\gamma$ -<sup>32</sup>P] ATP using T4 polynucleotide kinase followed by purification of radiolabelled DNA oligos or probes using QIAquick Nucleotide removal kit (supplied by Qiagen). Radiolabelled probe (50 pM) was incubated with different concentrations of purified protein at 25 °C in binding buffer (20 mM Tris-Cl (pH – 7.5), 100 mM NaCl, 1mM EDTA, 0.02 % Tween 20 and 1mM DTT) for 1 hr. The incubated samples were then mixed with loading dye (50 mM Tris-Cl, pH-7.5, 10 mM EDTA, 0.002 % bromophenol blue, 0.002 % of xylene cyanol and 50 % glycerol), and were analyzed on 6% native PAGE. Detection of the free probe and protein-bound probe was carried out using phosphorimaging (Storm 860, Molecular Dynamics).

### **3.2.14 Intrinsic fluorescence assays:**

To measure intrinsic tryptophan fluorescence, 3-6 µM of purified protein taken in the lysis buffer (50 mM Tris-Cl pH-7.5, 100 mM NaCl, 10 % glycerol) without TCEP. The protein sample was excited at 295 nm, and fluorescence emissions were monitored till 450 nm. For direct binding 3-6 µM protein was titrated with increasing concentration of nucleic acids. The values plotted using kaleidagraph™ (Synergy software) to derive K<sub>d</sub> using the method as described in Reich, Kovermann [147].

### 3.2.15 Yeast two-hybrid assays:

The full-length coding sequence of the *NbGRP3* amplified from *N. benthamiana* cDNA using specific primer sequences designed with appropriate restriction sites (Table 3.4). The amplified product was successfully cloned in frame with GAL4 Activation domain (GAL4 AD) and GAL4 DNA-binding domain (GAL4 BD) of yeast two hybrid vectors pGAD-C1 and pGBD-C1 [144]. The clones were verified by PCR, restriction digestion and sequencing. Yeast two-hybrid assays were performed for two purposes; one was to identify the interacting partners of *NbGRP3* among the different proteins of ToLCNDV-A such as AC2, AC4, and AV2 and  $\beta$ C1 from RaLCB. The other purpose was to check whether *NbGRP3* appears as a dimer/monomer *in vivo*. The yeast two-hybrid clones of AC2, AC4, AV2, and  $\beta$ C1 were already available in the laboratory. Yeast two-hybrid assays were performed by co-transforming the yeast two-hybrid constructs into *Saccharomyces cerevisiae* AH109 by lithium acetate/single-stranded DNA carrier/polyethylene glycol method [148]. Co-transformed colonies selected for interaction by streaking on 3 DO (SD/-Trp/-Leu/-His) followed by 3 DO (SD/-Trp/-Leu/-His) + 2 mM 3-AT (3-Amino-1, 2, 4-triazole) to have the stringent selection. Yeast cells co-transformed with SV-40 large T-antigen (AD-T-Ag) and tumor suppressor protein p53 (BD-p53) serve as positive control as they both known to display strong interaction [149].

### 3.2.16 *NbGRP3* expression during viral infection:

To understand the role of *NbGRP3* during geminiviral infection 15 days old *N. benthamiana* seedlings were agroinoculated with the following constructs pCAMBIA2300, infectious clones of ToLCNDV-A alone, and infectious clones of ToLCNDV-A along with infectious clones of RaLCB [150]. Leaf samples were collected from infected samples at different dpi (days post inoculation) such as 3 dpi, 6 dpi, 9 dpi, 18 dpi, and 21 dpi. RNA was extracted from the harvested leaf samples, 1 $\mu$ g RNA converted to cDNA and qRT-PCR performed using SYBR green to quantify the transcript level of *NbGRP3*, *NtTubulin* used as an internal reference control [151]. In case of *NbGRP3* expression level during TBSV infection, qRT-PCR was performed as described above in RNA samples from TBSV infected plants at 5 dpi.

### 3.2.17 Virus induced gene silencing of endogenous *NbGRP3*:

A *Tobacco rattle virus* (TRV)-based gene silencing method was used in this study [152]. Vectors used in this study, such as pTRV1 (CD3-1039), pTRV2 (CD3-1040), and pTRV2-*NtPDS* (CD3-1045) were procured from the Arabidopsis Biological Resource Center (ABRC). 3' UTR of *NbGRP3* cloned into pTRV2 using sites *Xba*I and *Bam*HI. The vectors pTRV1, pTRV2, pTRV2-*NbGRP3* mobilized into *A. tumefaciens* EHA105. For silencing of *NbGRP3*, co-infiltration of *A. tumefaciens* EHA105 with constructs pTRV1+ pTRV2-*NbGRP3* was performed on the lower leaves of 3 weeks old *N. benthamiana* plants. pTRV1+ pTRV2 serve as a control. To know the effect of silencing of *NbGRP3* in geminiviral pathogenesis, agroinfiltration of the infectious clones of ToLCNDV-A alone and ToLCNDV-A along with infectious clone of RaLCB, samples harvested 20 days post infiltration of geminivirus, genomic DNA isolated as per the method described in [153] and processed for Southern hybridization to check the level of geminiviral DNA.

### 3.2.18 Southern hybridization:

Southern hybridization [154] was performed to detect the accumulation of viral DNA from virus inoculated plants or leaf discs. 8.0 µg of genomic DNA isolated from virus inoculated plants or leaf discs was electrophoresed in 0.8 % agarose gel. The electrophoresed gel was treated with 0.25 M HCl for 15 min for depurination followed by 45 min treatment with denaturation solution (0.5 M NaOH and 1.5 M NaCl). The denatured gel was then finally transferred to neutralization solution containing 0.5 M Tris-HCl and 1.5 M NaCl. The DNA in the gel finally transferred onto a nylon membrane (MDI, India) by the capillary method using 10 X SSC solution (3.0 M NaCl and 0.3 M tri-sodium citrate) overnight. The DNA and membrane cross-linked using ultraviolet rays after transfer. ToLCNDV-AV1DNA was taken for synthesizing probe detect ToLCNDV-A and RaLCB-βC1 DNA was taken for synthesizing probe to detect RaLCB. Probes were synthesized using radiolabelled p<sup>32</sup>-dCTP which was hybridized with nylon membrane containing DNA in hybridization oven at 60 °C. Blots were washed to remove unincorporated probes with washing solution (2 X SSC and 0.1 % SDS) for 15 min to avoid nonspecific signals. Finally the blots were transferred to imaging plates and kept for

overnight. The radioactive signal was detected using phosphorimager (GE Amersham). Single blot used for probing both ToLCNDV-A and RaLCB. Initially, the blots probed for RaLCB first followed by stripping, washing and finally the blot probed for ToLCNDV-A.

### 3.2.19 Analysis of C-terminal amino acid repeats in NbGRP3:

NbGRP3 amino acid sequence was initially used for the construction of Dot plot using the server EMBOSS dotmatcher (<http://www.bioinformatics.nl/cgi-bin/emboss/help/dotmatcher>) followed by dissection of amino acids repeats using different online servers tabulated below,

**Table 3.9:** Details of different online servers used for the analysis of C-terminal amino acid repeats

Method	Source	Reference
REP	<a href="http://www.bork.embl.de/~andrade/papers/rep/search.html">http://www.bork.embl.de/~andrade/papers/rep/search.html</a>	[155]
TPRpred	<a href="https://toolkit.tuebingen.mpg.de/#/tools/tpred">https://toolkit.tuebingen.mpg.de/#/tools/tpred</a>	[156]
REPRO	<a href="http://www.ibi.vu.nl/programs/reprowww/">http://www.ibi.vu.nl/programs/reprowww/</a>	[157]
RADAR	<a href="http://www.ebi.ac.uk/Tools/Radar/">http://www.ebi.ac.uk/Tools/Radar/</a>	[158]
HHrepID	<a href="http://toolkit.tuebingen.mpg.de/hhrepid/">http://toolkit.tuebingen.mpg.de/hhrepid/</a>	[159]
SIMPLE	<a href="http://www.biochem.ucl.ac.uk/bsm/SIMPLE/">http://www.biochem.ucl.ac.uk/bsm/SIMPLE/</a>	[160]
REPPER	<a href="https://toolkit.tuebingen.mpg.de/repper/">https://toolkit.tuebingen.mpg.de/repper/</a>	[161]
FAIR	<a href="http://bioserver1.physics.iisc.ernet.in/cgi-bin/fair4/fair/indx.pl">http://bioserver1.physics.iisc.ernet.in/cgi-bin/fair4/fair/indx.pl</a>	[162]

### 3.2.20 Cloning, expression and purification of NbGRP3 C-terminal deletion variant (NbGRP3 $\Delta$ <sub>103-148</sub>):

Based on sequence alignment and analysis by different tools, it was decided to delete the repeat containing region from amino acid 103 to 148. To generate the C-terminal deletion variant

of (NbGRP3 $\Delta_{103-148}$ ), primers were designed flanking the repeats on either side in the coding sequence (from 307 – 444 bp). Entire pET SUMOadapt vector backbone along with NbGRP3 was amplified excluding the repeat, ligated to generate NbGRP3 $\Delta_{103-148}$ , expressed using the same conditions like that of full length. NbGRP3 $\Delta_{103-148}$  purified by the following methods: 6g of induced cells of BL21-codon plus (DE3) RIPL were lysed using lysis buffer (50 mM Tris-Cl pH-7.5, 100 mM NaCl, 10 % glycerol, 1mM TCEP) followed by a French press. The lysate spun at 40,000 rpm for one hour to separate clear supernatant containing the soluble fractions and cell debris. The clear supernatant was applied to Ni-NTA column for separation of His-SUMO-NbGRP3  $\Delta_{103-148}$  from other cellular proteins. The His-SUMO-NbGRP3 $\Delta_{103-148}$  was subjected to SUMO protease treatment followed by anion exchange column. Finally, the elution fractions of anion exchange chromatography were applied to heparin column to separate NbGRP3  $\Delta_{103-148}$  from other contaminants.

# *Results*





## 4. Results

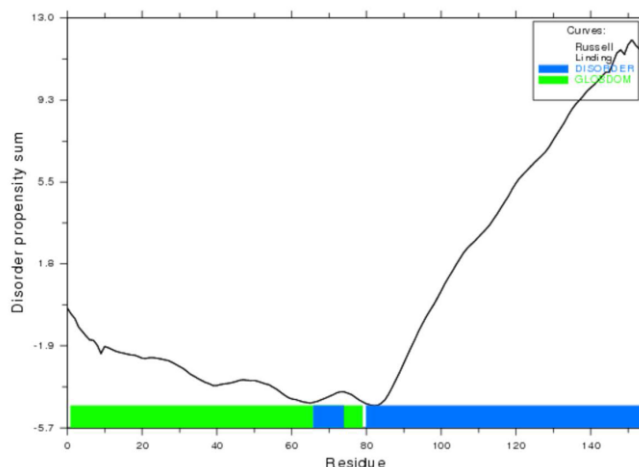
### 4.1 Structural domain analysis of NbGRP3:

The length of mRNA of *NbGRP3* was 1,394 bp with two exons and one intron. The size of coding sequence was 468 bp (including termination codon) that encode a protein of 155 amino acids. The amino acid sequence of NbGRP3 analyzed was using different databases for protein families such as NCBI conserved domains, Superfamily, Prosite, Smart and Pfam. All the databases predicted RNA Recognition Motif (RRM) at its N-terminal and none could predict the domains at the C-terminal region. The results from different online servers tabulated below;

**Table 4.1:** Results from the online databases used for prediction of domains in NbGRP3

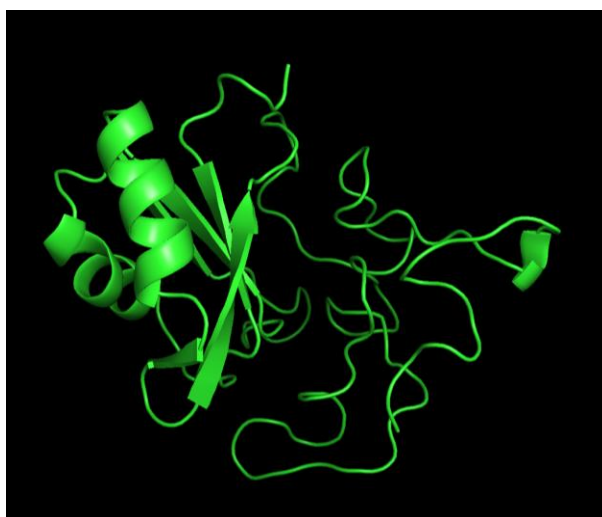
Online databases	Prediction of RRM
NCBI-Conserved domains	9-83
Superfamily	6-125
Expasy-Prosite	6-84
Smart	7-80
Pfam	9-77

The amino acid composition of the protein was found to have nearly 30 % of the amino acid glycine enriched in C-terminal region. The glycine content in the C-terminal alone accounts for 62.2 % making one to assume that the C-terminal is disordered. The disordered nature of NbGRP3 predicted using the program Globplot 2 [163]. Globplot 2 revealed that after amino acid positioned at 80, the entire stretch was highly disordered which is represented as blue color (Figure 4.1). NbGRP3 modeled using the online server Robetta with default parameters also predicted  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$  in the N-terminal RRM and random coiled structure at the C-terminal as expected (Figure 4.2).



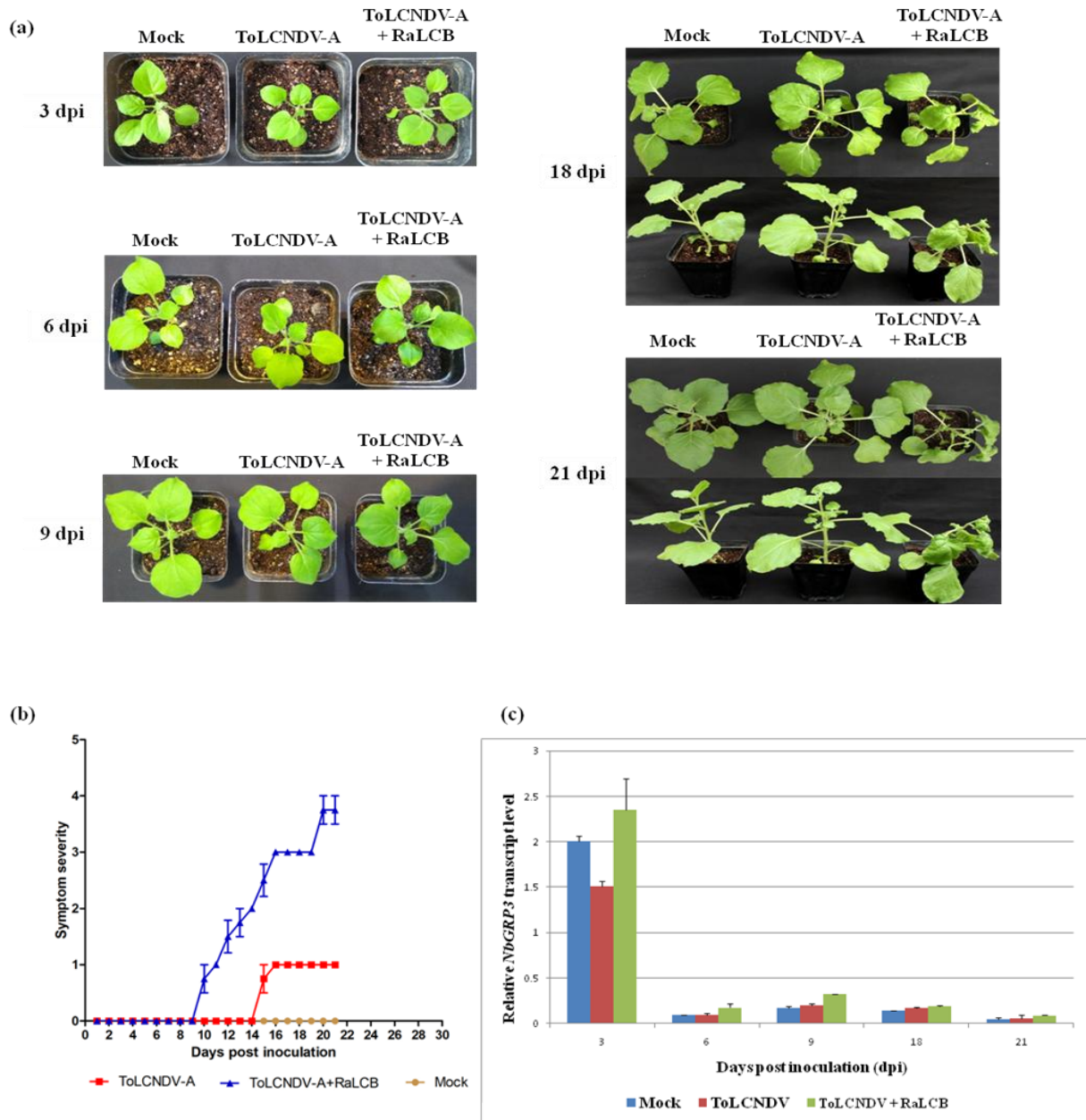
MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIINDRETGRSRGFGFVTFKDEKSMRDAIE  
 GMNGQDLLDGRNITV**NEA**QSRGSGGGGGGGGGYRGGGGGGYGGGRREGGYGGGGGYG  
**GGRREGGYGGGGGYGGGRREGGYGGGSEGNWRS**

**Figure 4.1: Prediction of disordered region using Globplot 2 online tool (<http://globplot.embl.de/>).** Green color represents the globular region of the protein whereas blue colour represents the disordered regions in the protein. The disordered region starts from amino acids “NEA”.



**Figure 4.2: Computational model of NbGRP3 generated by Robetta server (<http://robetta.bakerlab.org/>) displaying  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$  topology at the N-terminal and disordered region at the C-terminal.**

#### 4.2 Expression analysis of *NbGRP3* during geminiviral infection:

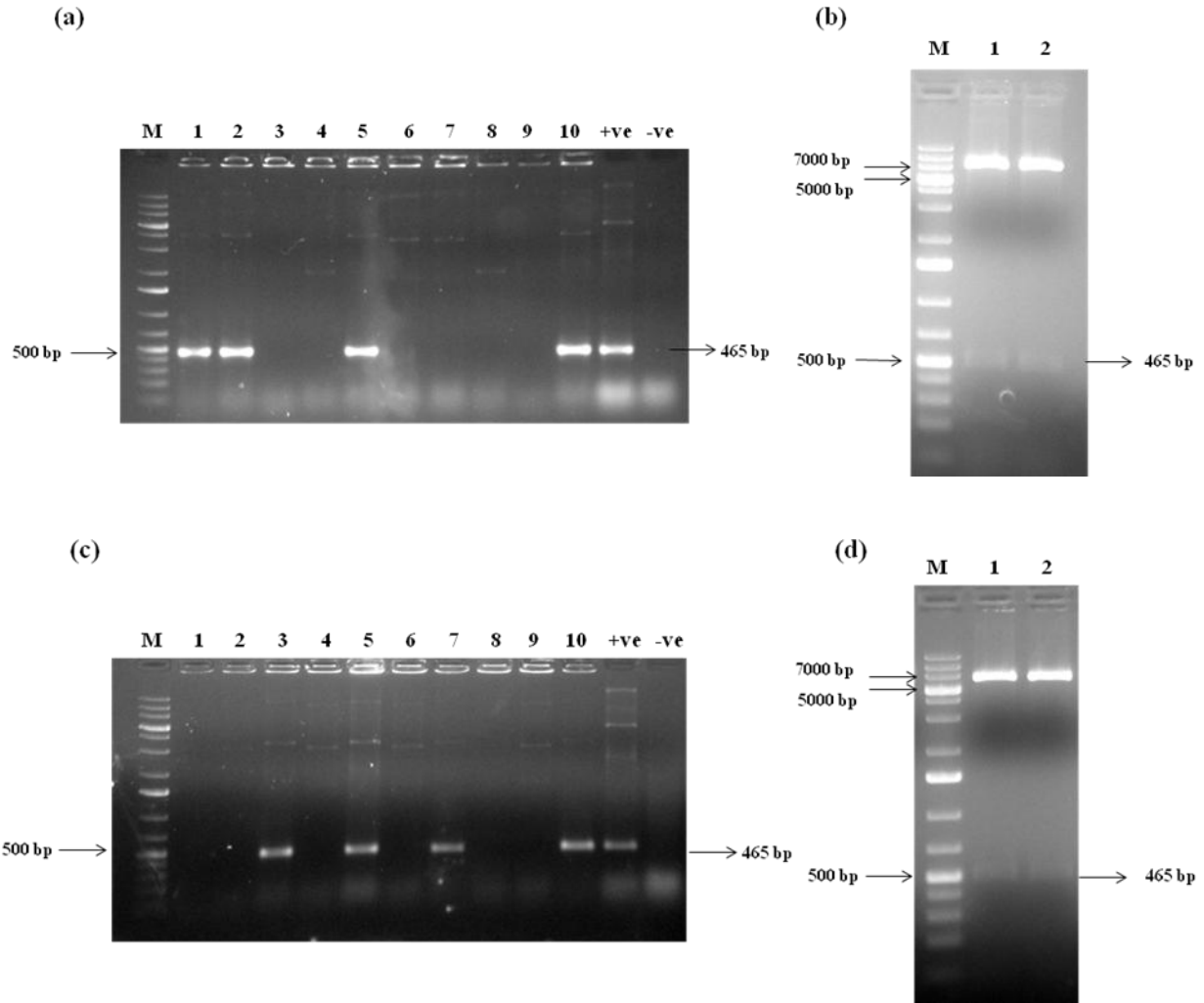


**Figure 4.3: Expression analysis of *NbGRP3* during geminiviral infection.** Two weeks old *N. benthamiana* plants agro-inoculated with infectious clones of ToLCNDV-A, ToLCNDV-A+RaLCB. The inoculated plants were daily scored for symptoms and samples harvested at different dpi such as 3, 6, 9, 18, and 21. RNA was isolated, transcript level of *NbGRP3* checked by RT-PCR (a) Phenotypes of plants used for the analysis at different dpi such as 3, 6, 9, 18, and 21 (b) Symptom severity graph of agro-inoculated plants. (c) Relative expression of *NbGRP3* in mock, ToLCNDV-A infected and ToLCNDV-A + RaLCB infected plants.

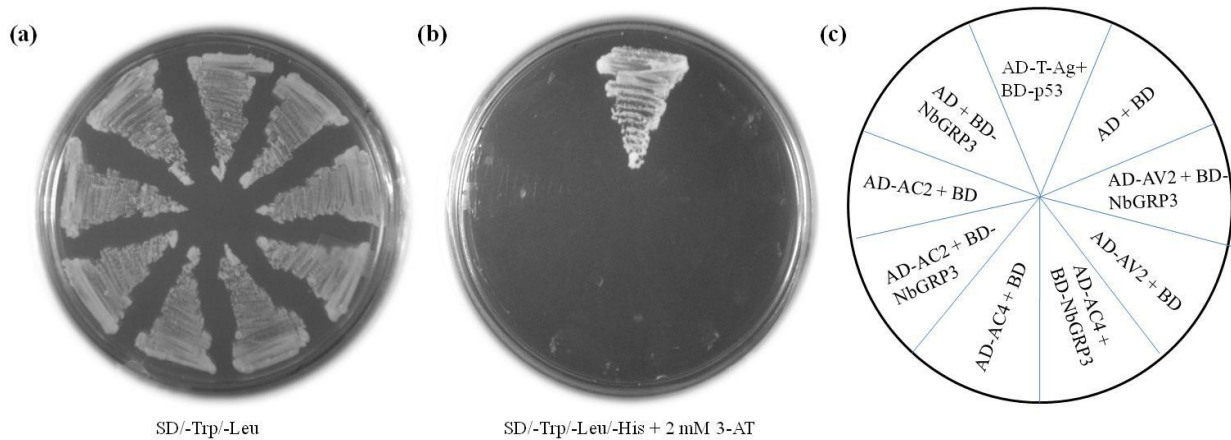
The first step towards understanding the role of *NbGRP3* in geminiviral pathogenesis was to study the expression of *NbGRP3* during geminiviral infection. We considered non-cognate combination of ToLCNDV and RaLCB because of their most prominent symptom induction than their cognate pairs [164] for the purpose above. ToLCNDV-A + RaLCB started displaying symptoms from 9-10 dpi onwards. ToLCNDV-A alone induced very mild leaf curling symptom at 18 and 21 dpi whereas at the same dpi ToLCNDV-A + RaLCB gave prominent symptoms such as veinal chlorosis, leaf curling, leaf crinkling, and stunting of the plant [164]. Expression of *NbGRP3* was studied at 3, 6, 9, 18, and 21dpi. From the expression analysis, the expression of *NbGRP3* could be correlated with the viral infection. The expression of *NbGRP3* was at the maximum at 3 dpi followed by gradual decrease in expression at subsequent dpi. Interestingly, at 6 and 9 dpi there was a noticeable increase in expression of *NbGRP3* in ToLCNDV-A + RaLCB infection when compared with mock and ToLCNDV-A alone infected plants. At 18 dpi, both ToLCNDV-A and ToLCNDV-A + RaLCB inoculated plants had a upregulation of *NbGRP3* than mock.

#### **4.3 Identification of interacting partners of NbGRP3:**

Recently, NgGRP from *Nicotiana glutinosa* has shown to be an endogenous silencing suppressor [19]. To understand whether NbGRP3 has any role in gene silencing during geminivirus, interaction of NbGRP3 was checked with the viral suppressors of RNA silencing (VSR) like AC2, AV2 and AC4 encoded by ToLCNDV and RaLCB- $\beta$ C1 [24]. The yeast two-hybrid clones for the abovementioned VSRS were already available in the laboratory whereas yeast two-hybrid clones of *NbGRP3* were generated as a part of this study. The coding sequence of *NbGRP3* was cloned into yeast two-hybrid vector pGAD-C1 and pGBD-C1 using the sites *EcoRI* and *BamHI* (Figure 4.4). Yeast two-hybrid assays indicated that NbGRP3 didn't interact with any of the ToLCNDV ORFs used. Other than the positive control, none of the transformed colonies grown in 3 DO (SD/-Trp/-Leu/-His) + 2 mM 3-AT (Figure 4.5).

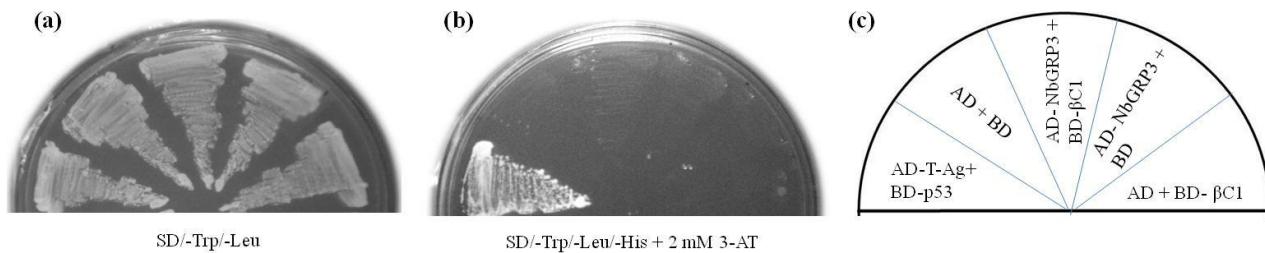


**Figure 4.4: Cloning of *NbGRP3* into yeast two-hybrid vectors pGAD-C1 and pGBD-C1.** (a) and (c) Agarose gel electrophoresis showing PCR confirmation of pGAD-C1-*NbGRP3* and pGBDC1-*NbGRP3*. Specific amplification of size 465 bp observed in both cases in few clones screened, which were further confirmed by restriction digestion and sequencing. (b) Agarose gel electrophoresis showing digestion confirmation of pGAD-C1-*NbGRP3* and pGBD-C1-*NbGRP3* with *EcoRI* and *Bam*HI. Release of appropriate size (465 bp) observed from vector backbone 6.6 kbp in case of pGAD-C1 and 5.8 kbp in case of pGBD-C1 after digestion.



**Figure 4.5: Yeast two-hybrid assays between NbGRP3 and ToLCNDV-A ORFs such as AC2, AV2 and AC4.** (a) Co-transformed yeast colonies in double dropout (SD/-Trp/-Leu) (b) Co-transformed yeast colonies in triple dropout (SD/-Trp/-Leu/-His). (c) Diagrammatic representation of co-transformed yeast colonies. Yeast cells co-transformed with AD-T-Ag + BD-p53 alone grown in double dropout and triple dropout with 2 mM 3-AT.

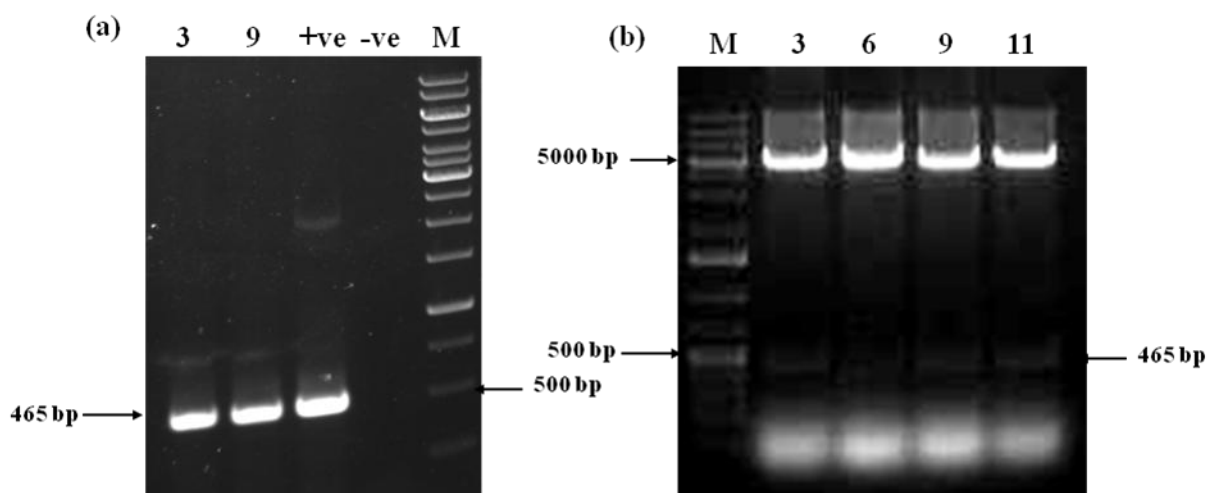
The only ORF encoded by RaLCB is  $\beta$ C1 which is a multitasking protein as well as symptom determinant and VSR [95]. The yeast two-hybrid clone of RaLCB- $\beta$ C1 was already available in the laboratory and same like ToLCNDV-A ORFs,  $\beta$ C1 also doesn't interact with NbGRP3 (Figure 4.6).



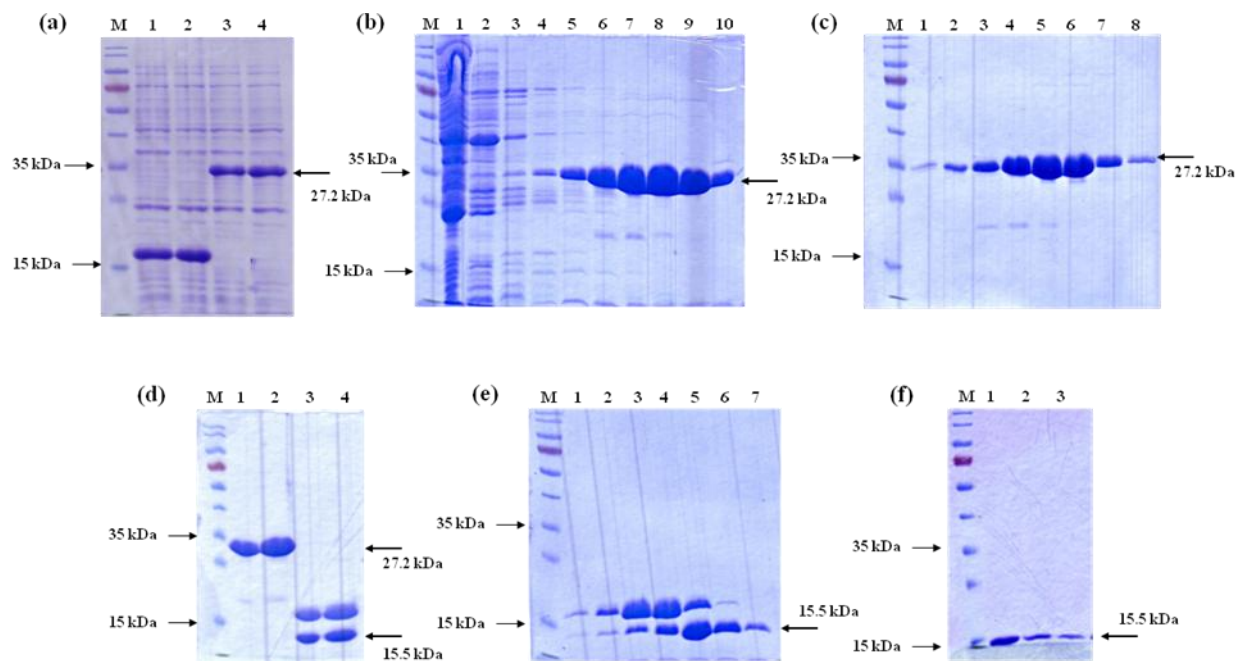
**Figure 4.6: Yeast two-hybrid assays between NbGRP3 and RaLCB- $\beta$ C1.** (a) Co-transformed yeast colonies in double dropout (SD/-Trp/-Leu) (b) Co-transformed yeast colonies in triple dropout (SD/-Trp/-Leu/-His). (c) Diagrammatic representation of co-transformed yeast colonies. Yeast cells co-transformed with AD-T-Ag + BD-p53 alone grown in double dropout and triple dropout with 2 mM 3-AT.

#### 4.4 Expression and purification of NbGRP3:

To have more information on *NbGRP3* purified protein was mandatory. Therefore, *NbGRP3* was cloned into pET SUMOadapt (modified version of the pET SUMO expression system) [143] between the sites *Bsa*I and *Hind*III (Figure 4.7). The clones were all verified by PCR, restriction digestion and finally a single clone by sequencing. The sequence was verified clone transformed into *E. coli* BL21 CodonPlus (DE3)-RIPL.



**Figure 4.7: Construction of pET SUMOadapt-*NbGRP3* clone.** *NbGRP3* coding sequence was cloned into pET SUMOadapt vector between the restriction sites *Bsa*I and *Hind*III (a) Agarose gel electrophoresis showing PCR confirmation of pET SUMOadapt-*NbGRP3*. (b) Agarose gel electrophoresis showing digestion confirmation of pET SUMOadapt-*NbGRP3* with *Bsa*I and *Hind*III. Release of appropriate size (465 bp) observed from vector backbone (5623 bp) after digestion.

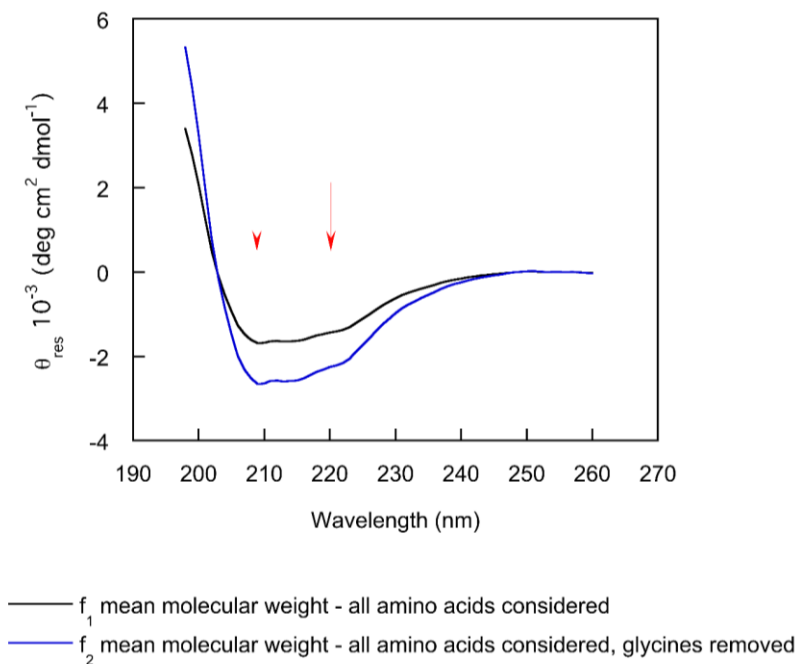


**Figure 4.8: Different steps involved in expression and purification of NbGRP3.** (a) Induction and solubilisation of pET SUMOadapt and pET SUMOadapt-NbGRP3 by 0.2 mM IPTG at 37 ° C for 4 hours. lane 1,2 – total fraction and soluble fraction of pET SUMOadapt; 3,4 - total fraction and soluble fraction of pET SUMOadapt-NbGRP3. (b) Different elute fractions after first Ni-NTA affinity chromatography. (c) Different elutes of 6xHis-SUMO-NbGRP3 (27.2 kDa) after anion exchange chromatography using Q-sepharose beads. (d) SUMO protease treatment: lane 1, 2 – before SUMO protease treatment; lane 3, 4 – post SUMO protease treatment. (e) Gel filtration chromatography of SUMO protease treated samples showing 6xHis-SUMO and NbGRP3 (15.5 kDa). (f) Final purification of NbGRP3 by second Ni-NTA affinity chromatography. The purified protein finally confirmed by mass spectrometry. Although the molecular weight of NbGRP3 was 15.5 kDa, in SDS-PAGE it was seen below 15 kDa which could be because of their amino acid composition.

NbGRP3 was purified by different steps of chromatography as shown in the figure 4.8 above. The purified protein finally confirmed by mass spectrometry. Although the molecular weight of NbGRP3 was 15.5 kDa, in SDS-PAGE it was seen below 15 kDa marker which could be because of their amino acid composition. The purified protein was characterized by CD spectroscopy which is an indispensable tool to study the secondary structure information of any protein. The glycine subtracted CD spectrum shed information on the secondary structure of structured part of NbGRP3 which indicated the presence of  $\alpha$ -helix and  $\beta$ -sheets. The spectrum



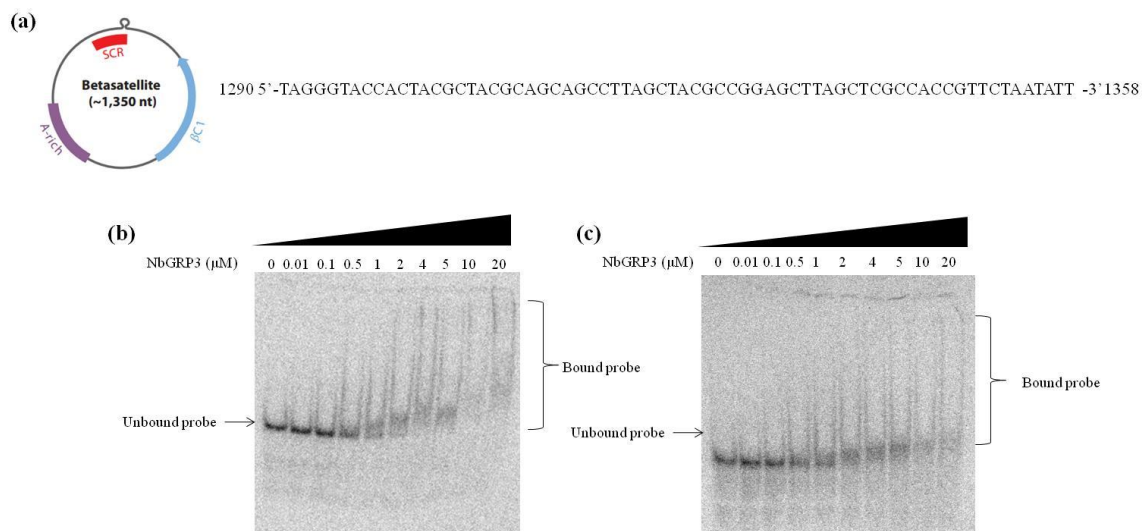
had a negative value near 222 nm and 208 nm which are the characteristics of  $\alpha$ -helix. It also had a negative value near 218 nm which is the characteristics peak for  $\beta$ -sheets (Figure 4.9).



**Figure 4.9: Structural properties of purified full-length NbGRP3 using Far-UV CD spectrum.** NbGRP3 (263.6  $\mu\text{g/ml}$ ) was taken for measurement in buffer 50 mM Tris-HCl, pH – 7.5, 100 mM NaCl, 10 % glycerol and 1 mM TCEP (tris(2-carboxyethyl) phosphine). The black line of the curve represents the complete amino acid sequence of NbGRP3 and the blue line represents the analyzed curve after subtracting glycines.

#### 4.5 Interaction of NbGRP3 with SCR of RaLCB:

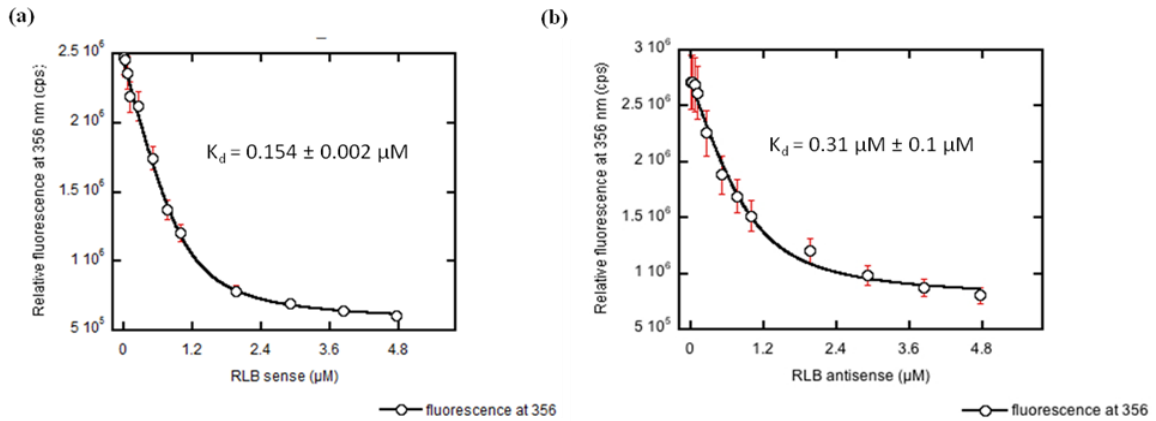
To have more insights into the role of NbGRP3 in RaLCB, EMSA was performed between purified NbGRP3 and sense strand of 69 nt region from SCR of RaLCB (Figure 4.10a). EMSA didn't show clear shift instead shift appeared as a smear. Similar smear observed with antisense strand of the same region to know whether the observed phenomenon was the general property of NbGRP3 (Figure 4.10b&c). Literature evidence denotes that its close homologs from other organisms like HvGRP1 and NtGRP1 also exhibit similar binding behavior. This confirms that NbGRP3 has a unique but unexplored binding behavior.



**Figure 4.10: NbGRP3-RaLCB SCR interaction by EMSA** (a) Nucleotides from 1290 to 1358 from the SCR region (highlighted in red) of RaLCB was considered for interaction studies. EMSA showing the interaction of NbGRP3 with sense strand (b) and antisense strand (c) RaLCB SCR region of size 69 nucleotides. NbGRP3 incubated with increasing concentration from 0  $\mu\text{M}$  till 20  $\mu\text{M}$  with 50 pM of the DNA. Shifts appear as a smear in both cases.

#### 4.6 Nucleic acid binding studies using intrinsic fluorescence:

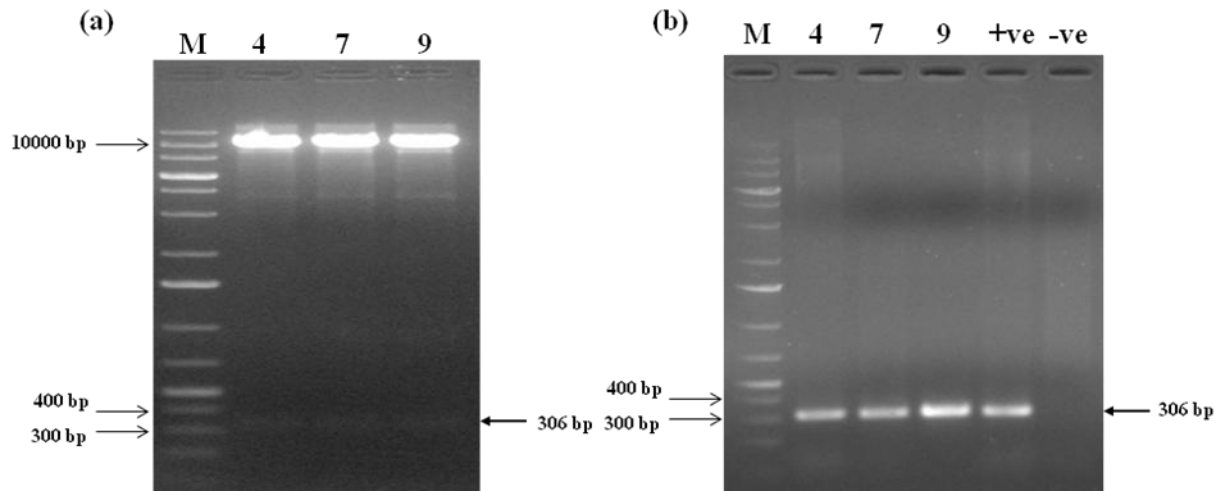
The NbGRP3 amino acid sequence has two tryptophan molecules: one at position 15 at N-terminal and another at position 153 and 8 tyrosine residue enriched at the C-terminal region. In a protein, the amino acids tryptophan, phenylalanine and tyrosine exhibit intrinsic fluorescence. Fluorescence of phenylalanine is negligible when compared with the other two amino acids. At the wavelength of 280nm both tryptophan and tyrosine get excited whereas at 295nm only tryptophan gets excited predominantly over other two amino acids, moreover the fluorescence properties of tryptophan are highly sensitive to the local environment making it more relevant to study protein-ligand interaction [165]. NbGRP3 has two tryptophan amino acids; one at N-terminal (position – 15) and second is at C-terminal (position – 153). Therefore, NbGRP3 was suitable for protein-ligand interaction studies using intrinsic fluorescence. Thereafter, all the NbGRP3-nucleic acid interaction was studied using fluorescence spectroscopy. By intrinsic fluorescence measurements the  $K_d$  measured for both the sense and anti-sense of SCR region taken for study (Figure 4.11).



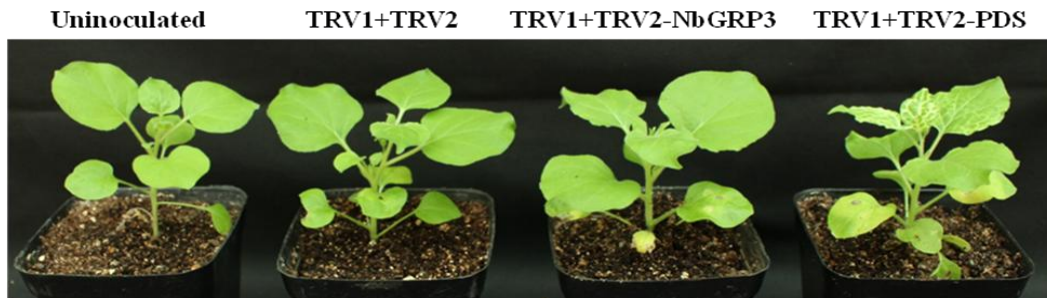
**Figure 4.11: NbGRP3 – RaLCB SCR interaction using intrinsic fluorescence.** (a) With sense strand having the  $K_d = 0.154 \pm 0.002 \mu\text{M}$  (b) With antisense strand having  $K_d = 0.31 \pm 0.1 \mu\text{M}$ .

#### 4.7 Effect of transient silencing of *NbGRP3* on geminiviral pathogenesis:

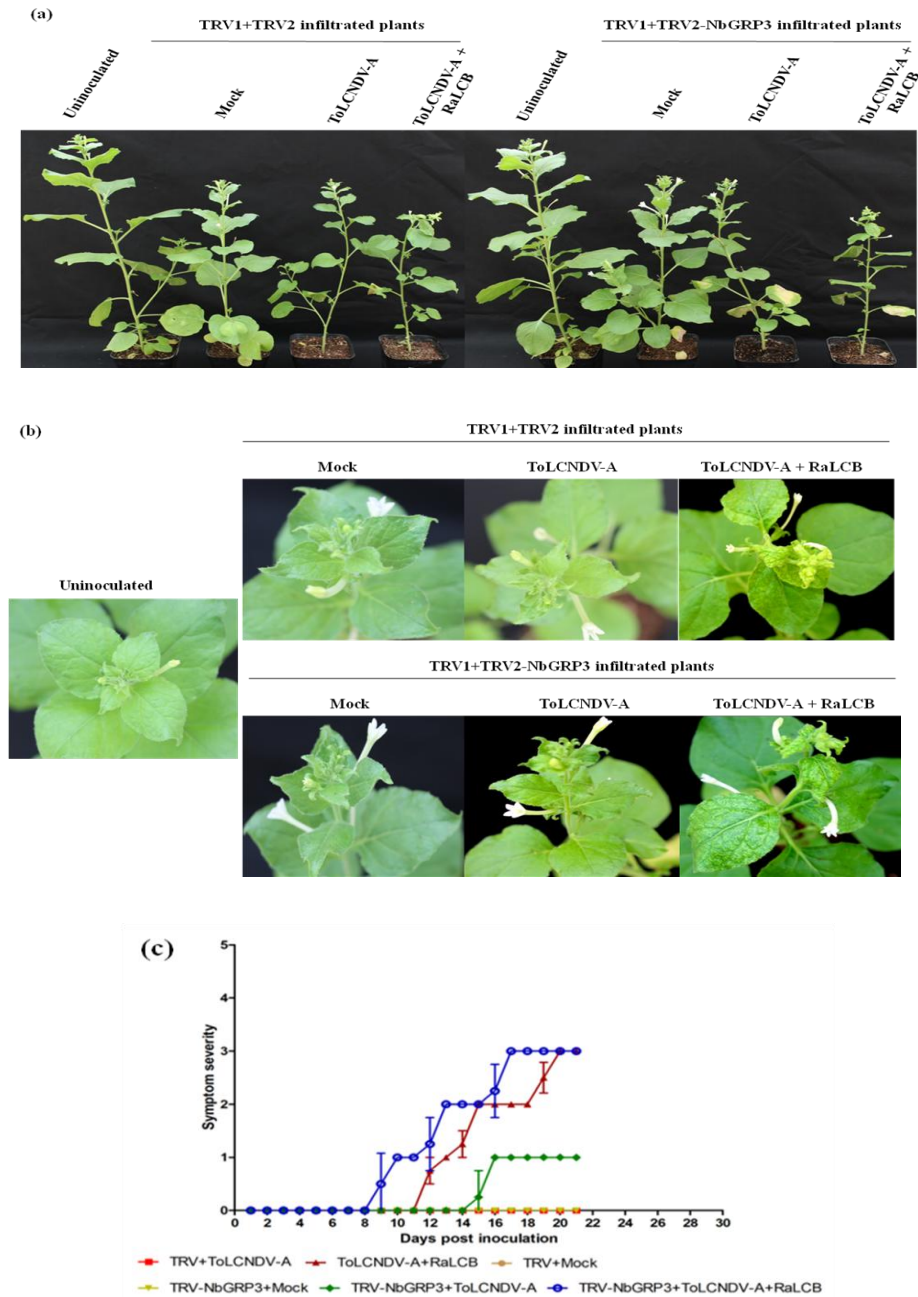
To have a better picture of role of *NbGRP3* in geminiviral pathogenesis, TRV-based VIGS was carried out to transiently silence *NbGRP3* expression. The 3'-UTR of *NbGRP3* was considered for silencing purpose, which was cloned into pTRV2 using *Xba*I and *Bam*HI. 3 weeks old *N. benthamiana* plants were initially infiltrated with pTRV2-*NbGRP3* silencing construct along with controls. ToLCNDV-A, ToLCNDV-A+RaLCB were infiltrated into TRV1+TRV2, TRV1+TRV2-*NbGRP3* plants 10 days post infiltration of the TRV virus when photobleaching manifest in TRV1+TRV2-*PDS* infiltrated plants. The geminivirus infiltrated plants were scored for symptoms post infiltration and leaf samples harvested at 20 dpi. ToLCNDV-A infiltrated plants developed mild leaf curling symptoms around 20 dpi in *NbGRP3* silenced plants whereas control plants didn't show any symptom. ToLCNDV-A+RaLCB infiltrated plants had visible symptoms like vein clearing, leaf curling and stunting in both *NbGRP3* silenced as well as control plants but *NbGRP3* silenced plants had developed symptoms early and had higher viral DNA accumulation than that of control plants (Figure 4.14c & 4.15). This particular observation gives us the clue that *NbGRP3* contributes to plant defense against geminivirus.



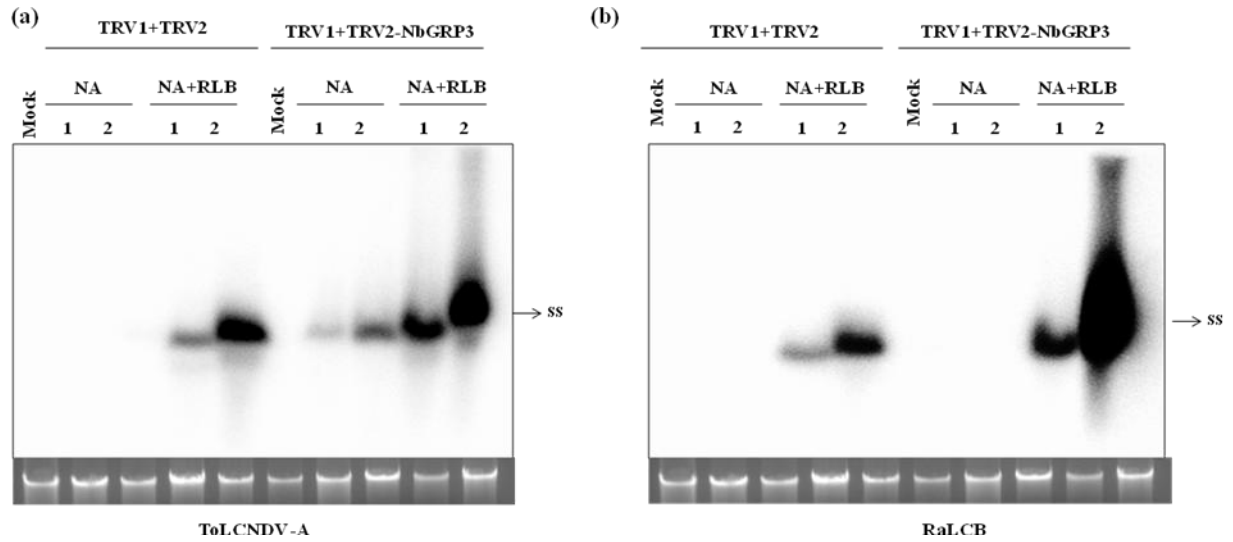
**Figure 4.12: Construction of pTRV2-*NbGRP3* clone for VIGS.** 3'UTR of *NbGRP3* of size 306 bp taken for cloning into pTRV2 vector (a) Agarose gel electrophoresis showing digestion confirmation of pTRV2-*NbGRP3* clone using enzymes *Xba*I and *Bam*HI. Fall out of 306 bp represent 3'UTR of *NbGRP3* (b) Agarose gel electrophoresis showing PCR confirmation of pTRV2-*NbGRP3* clone.



**Figure 4.13: Phenotype of uninoculated, TRV1+TRV2, TRV1+TRV2-*NbGRP3*, TRV1+TRV2-*PDS* infiltrated plants after 10 days post infiltration of VIGS constructs.** Photobleaching because of silencing of *phytoene desaturase* (*PDS*) serve as control to monitor initiation of TRV - based silencing.



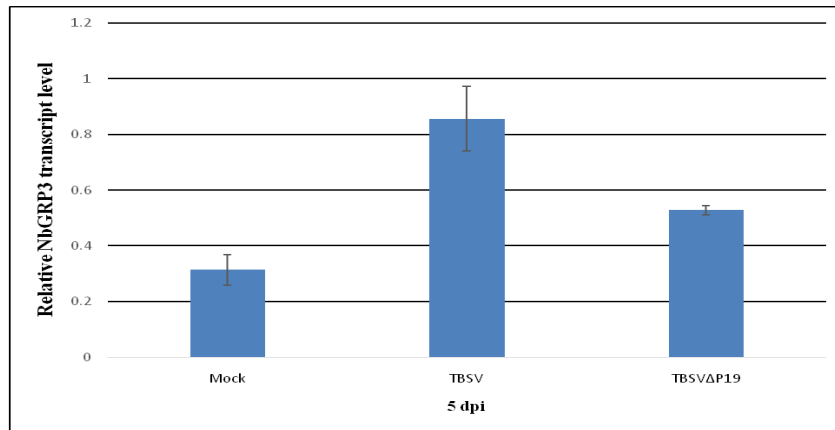
**Figure 4.14: Phenotype of geminivirus infiltrated plants on *NbGRP3* silenced plants 20 days post infiltration of geminivirus.** (a) Overview of plants infiltrated with ToLCNDV-A, ToLCNDV-A+RaLCB. (b) Close view of the apical region of plants infiltrated with ToLCNDV-A, ToLCNDV-A+RaLCB. Leaf curling and veinal chlorosis visible in ToLCNDV-A+RaLCB infiltrated plants. (c) Symptom severity chart for the plants used in this experiment. It can be observed that in *NbGRP3* silenced plants infiltrated with ToLCNDV-A+RaLCB symptoms started around 8 dpi which was early than non-silenced plants.



**Figure 4.15: Southern hybridization analysis demonstrating the effect of *NbGRP3* silencing in geminivirus DNA accumulation.** (a) DNA accumulation of ToLCNDV-A in TRV1+TRV2 and TRV1+TRV2-*NbGRP3* plants. Blot was probed with ToLCNDV-AV2. (a) DNA accumulation of RaLCB in TRV1+TRV2 and TRV1+TRV2-*NbGRP3* plants. Blot was probed with RaLCB- $\beta$ C1. It can be observed that *NbGRP3* silenced plants higher viral DNA accumulation than non-silenced plants. Ethidium bromide stained gel showing plant genomic DNA serve as loading control. NA – ToLCNDV-A, NA+RLB – ToLCNDV-A+RaLCB.

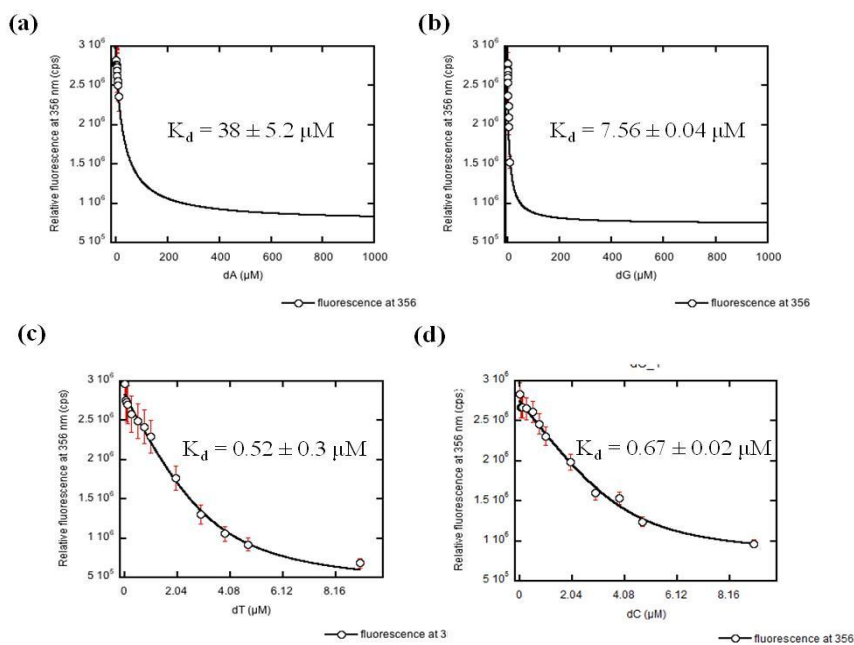
#### 4.8 Expression analysis of *NbGRP3* during TBSV infection:

From the experiments above demonstrated the role of *NbGRP3* in geminiviral pathogenesis. Geminiviruses are single stranded DNA viruses, to know whether the anti-viral role of *NbGRP3* limited only to single stranded DNA viruses, preliminary experiment performed. Expression of *NbGRP3* checked in *Tomato bushy stunt virus* (TBSV) infected samples. Expression studies revealed that TBSV infected samples had higher expression of *NbGRP3* than mock and interestingly, in samples infected with mutated version of TBSV ie. TBSV $\Delta$ p19 the level of *NbGRP3* quite compromised than that of wild type virus infected samples. . This stresses on the fact that role of *NbGRP3* in viral pathogenesis could be broad.



**Figure 4.16: Expression analysis of *NbGRP3* in TBSV infection samples at 5 dpi.** *NbGRP3* transcript accumulation checked in TBSV infected samples and TBSVΔp19. *NbGRP3* was higher in TBSV samples than that of mock and slightly compromised in TBSVΔP19.

#### 4.9 Biochemical characterization of NbGRP3:



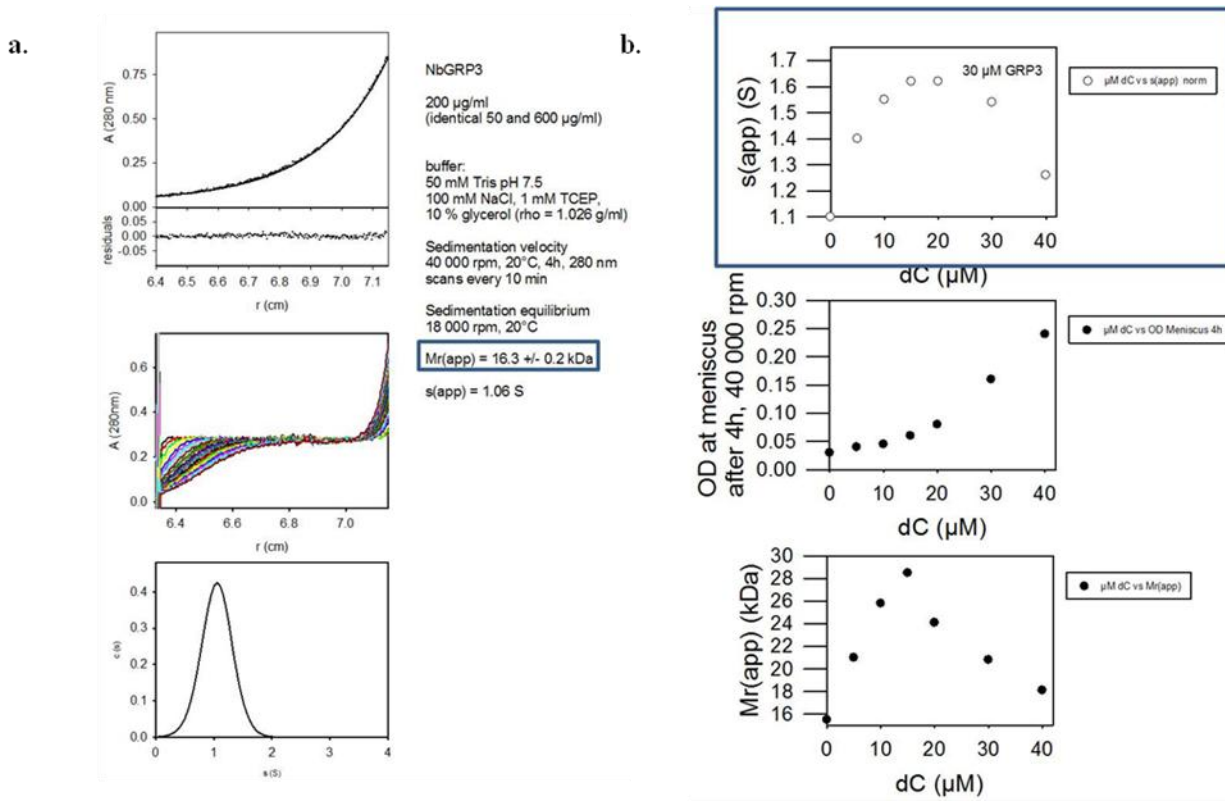
**Figure 4.17: Binding analysis of NbGRP3 with different homodeoxyribonucleotides using intrinsic fluorescence.** NbGRP3 exhibited weak affinity towards homodeoxypurines: (a) With  $(dA)_{12}$ , the  $K_d = 38 \pm 5.2 \mu\text{M}$ ; (b) With  $(dG)_{12}$ , the  $K_d = 7.5 \pm 0.04 \mu\text{M}$ . NbGRP3 exhibited strong affinity towards homodeoxypyrimidines (c) With  $(dT)_{12}$ , the  $K_d = 0.52 \pm 0.3$ ; and (d) With  $(dC)_{12}$ , the  $K_d = 0.67 \pm 0.02 \mu\text{M}$ .

After having smeary band shift in EMSA, it was realized the need for biochemical and biophysical characterization for NbGRP3. To start with biochemical characterization, NbGRP3 protein incubated with homodeoxyribonucleotides of size 12 nucleotides such as dA, dG, dT and dC, their binding affinities determined by fluorescence spectroscopy. NbGRP3 showed binding towards deoxypyrimidines i.e (dC)<sub>12</sub> and (dT)<sub>12</sub> but not to the deoxypurines (Figure 4.17).

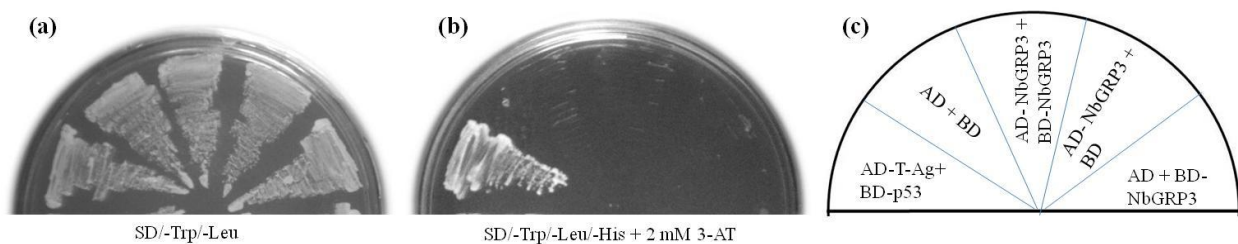
#### **4.10 Sedimentation equilibrium analysis of NbGRP3 using analytical ultracentrifugation:**

Analytical ultracentrifugation demonstrated the homogeneity of purified NbGRP3. The sedimentation velocity with 3 different concentrations of purified protein (50 µg/ml, 200 µg/ml and 600 µg/ml) corresponds to 16.3 +/- 0.2 kDa which was equivalent to theoretical mass of monomeric protein (15.56 kDa) (Figure 4.18a) and the yeast two-hybrid results also demonstrated that NbGRP3 doesn't interact with another NbGRP3 (4.19). Therefore, NbGRP3 purified protein as such is a monomer. To determine the state of NbGRP3 in the presence of ligands, 30 µM of NbGRP3 titrated against different concentrations of (dC)<sub>12</sub> such as 10 to 40 µM. The sedimentation velocity of NbGRP3 was at the maximum near 15 µM of (dC)<sub>12</sub>, beyond which the sedimentation velocity decreases. The inference once can assume from the above observation was 2 molecules of NbGRP3 binds either as two independent events or as a dimer to 1 molecule of ligand at optimum concentrations. When NbGRP3 was provided with a high concentration of ligands it shifts from dimeric state to monomeric state (Figure 4.18b).





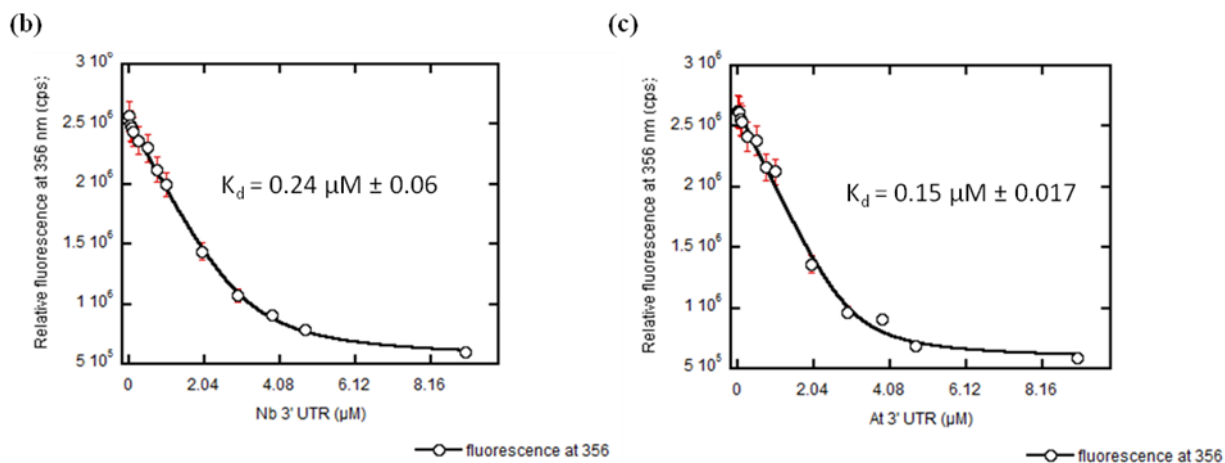
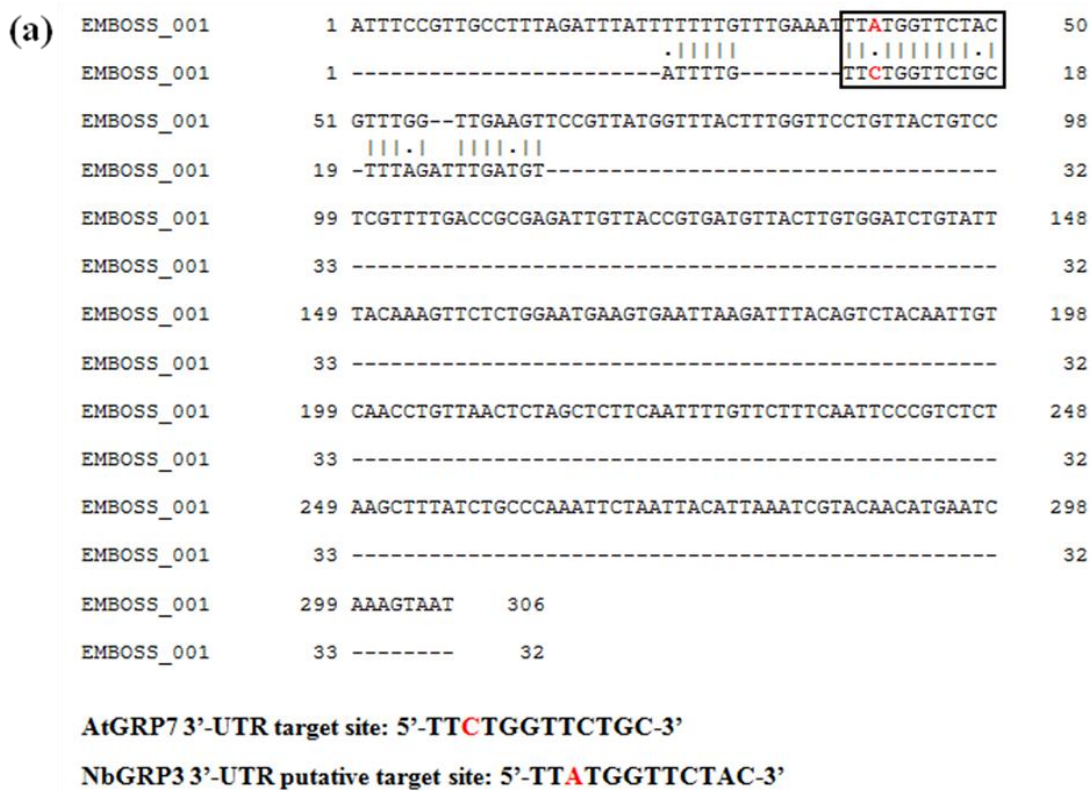
**Figure 4.18: Sedimentation equilibrium analysis of purified NbGRP3:** (a) purified NbGRP3 alone with the concentration of 200 µg/ml, identical results obtained with 50 µg/ml and 600 µg/ml. (b) Purified NbGRP3 titrated with different concentration of ligand - (dC)<sub>12</sub>.



**Figure 4.19: Yeast two hybrid assays demonstrating that the NbGRP3 as a monomer.** (a) Transformed yeast colonies in double dropout (SD/-Trp/-Leu) (b) Transformed yeast colonies in triple dropout (SD/-Trp/-Leu/-His) along with 2 mM 3-AT. (c) Diagrammatic representation of transformed yeast colonies streaked in double and triple dropout.

#### 4.11 Identification of putative binding site of NbGRP3 in its own 3'-UTR:

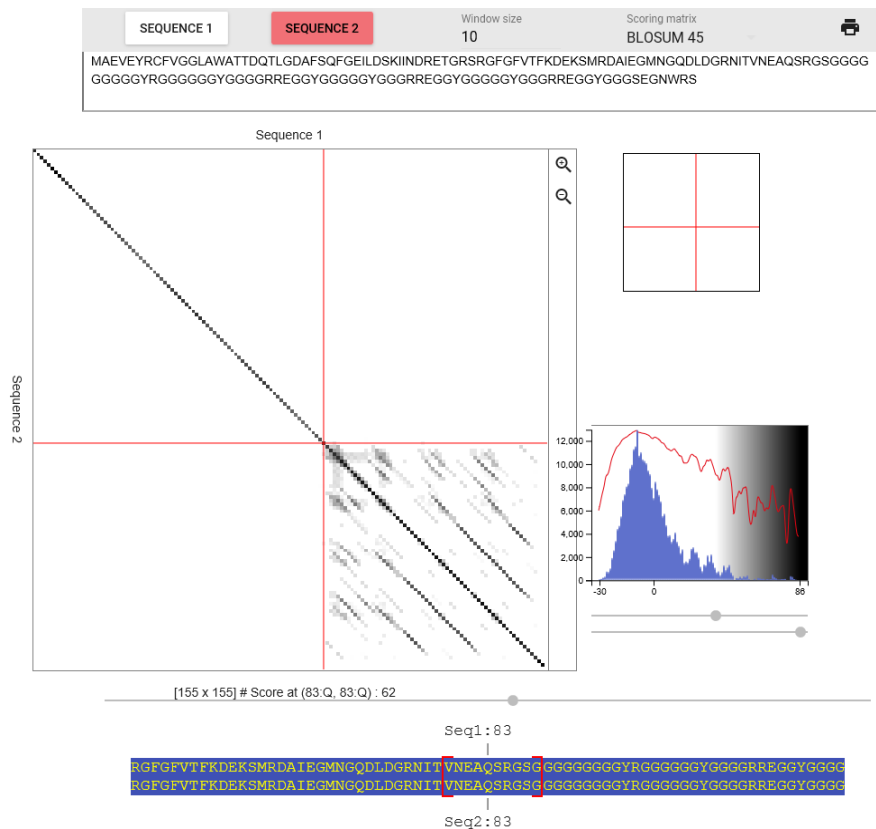
AtGRP7 a close homolog (68.7 % identity) of NbGRP3 was known to bind to its own 3'-UTR in a sequence-specific manner to regulate the expression of its own. The 3'-UTR target site of AtGRP7 was mapped to the sequence 5'-ATT TTG TTC TGC TTC TGG TTT AGA TTT GAT GT – 3' and especially 5'-TTC-3' at position 6-9 ntd (highlighted in red) was found to be essential for the binding of AtGRP7, deletion of 5'-TTC-3' reduces the nucleic acid affinity of AtGRP7 to nearly ten-fold [113, 121]. Two groups independently used the nucleotide stretch 5'-TTCTGG-3' as a ligand in NMR based structural studies of GRP from different organisms like NtGRP1 and HvGRP1 [11, 12]. It was found that the nucleotide stretch 5'-TTCTGGTTCTGC-3' was highly conserved in 3'-UTR of *NbGRP3* from position 38 to 50 nt but interestingly this stretch has two transversion. In 3'-UTR of NbGRP3, the nucleotide is “5'-TTA TGGTTCTGC-3'” instead of “5'-TTCTGGTTCTAC-3'”. The first transversion was C:A transversion is at the spot which was shown to be essential for the binding for AtGRP7 (Figure 4.20a). Binding studies revealed that NbGRP3 could bind to both 3'-UTR of *AtGRP7* as well as *NbGRP3* and of course yes, the C:A transversion had an effect on the binding as *AtGRP7* 3'-UTR binding site had higher affinity than *NbGRP3* 3'-UTR but not a drastic effect was noticed (Figure 4.20b&c).



**Figure 4.20: Comparison of binding analysis of NbGRP3 with 3'-UTR target sites from *NbGRP3* and *AtGRP7*** (a) Pairwise sequence alignment between *AtGRP7* 3'-UTR target site and *NbGRP3* 3'-UTR. The essential residue for binding of *AtGRP7* which was mutated in *NbGRP3* 3'-UTR highlighted as red. Direct binding studies using intrinsic fluorescence with *NbGRP3* putative target site, the  $K_d = 0.24 \pm 0.06 \mu\text{M}$  (b) with *AtGRP7* target site, the  $K_d = 0.15 \pm 0.017 \mu\text{M}$ .

## 4.12 Amino acid repeats (AAR) in C-terminal:

To gain deep understanding of the binding behavior of NbGRP3, the amino acid sequence was investigated. The binding properties of N-terminal region was sufficiently explored [166] but no exploration of C-terminal region was done except the NMR studies on HvGRP1 and NtGRP1 which stressed on the intra and inter-molecular associations exhibited by C-terminal GR domain and which could be reason for the unique binding behavior exhibited by NbGRP3. Therefore, our prime focus shifted towards C-terminal part. Dotplot analysis was performed using the online server EMBOSS dotmatcher (<http://www.bioinformatics.nl/cgi-bin/emboss/dotmatcher>). Dotplot identified the presence of low complexity repeats in the C-terminal disordered region which start from the residue 83 (Figure 4.21) and it was coincided with the disordered region of NbGRP3.



**Figure 4.21: Dotplot showing Low Complexity Repeats (LCR) in the C-terminal of NbGRP3.** From the dotplot, it can be observed that from the amino acid at position 83 onwards the C-terminal domain has several LCRs. Coincidentally, such repeats were in the disordered region of NbGRP3.

#### 4.13 Understanding repeats using different online servers:

To have more dissection of the repeats, different online servers used which are listed in the table below,

**Table 4.2:** Conclusions from online servers used for understanding C-terminal repeats of NbGRP3

Method	Source	Remarks
REP	<a href="http://www.bork.embl.de/~andrade/papers/rep/search.html">http://www.bork.embl.de/~andrade/papers/rep/search.html</a>	No conclusion
TPRpred	<a href="https://toolkit.tuebingen.mpg.de/#/tools/tpred">https://toolkit.tuebingen.mpg.de/#/tools/tpred</a>	No information on C-terminal repeats
REPRO	<a href="http://www.ibi.vu.nl/programs/reprowww/">http://www.ibi.vu.nl/programs/reprowww/</a>	No conclusion
RADAR	<a href="http://www.ebi.ac.uk/Tools/Radar/">http://www.ebi.ac.uk/Tools/Radar/</a>	No information on C-terminal repeats
HHrepID	<a href="http://toolkit.tuebingen.mpg.de/hhrepid/">http://toolkit.tuebingen.mpg.de/hhrepid/</a>	No hits
SIMPLE	<a href="http://www.biochem.ucl.ac.uk/bsm/SIMPLE/">http://www.biochem.ucl.ac.uk/bsm/SIMPLE/</a>	No results
REPPER	<a href="https://toolkit.tuebingen.mpg.de/repper/">https://toolkit.tuebingen.mpg.de/repper/</a>	No information on C-terminal repeats
FAIR	<a href="http://bioserver1.physics.iisc.ernet.in/cgi-bin/fair4/fair/indx.pl">http://bioserver1.physics.iisc.ernet.in/cgi-bin/fair4/fair/indx.pl</a>	Gave elaborate information on the repeats

Among the online servers used, FAIR performed a detailed dissection of the C-terminal repeats and the results are presented in table 4.3. The C-terminal has tandem of repeats from amino acid 103-148.

**Table 4.3:** Dissection of C-terminal repeats by server FAIR

S. No	Sequence of repeat	No. of residues	No. of repeats	Position	Location of repeat (highlighted in green)
1.	GGRREGGYGGGGYGG GRREGGYGGG	27	2	107-133; 122-148	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDRNI TVNEAQSRSGGGGGGGGYRGGGGGGYGGRRREG GGGGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRs
2.	GGRREGGYGGG	12	2	107-118; 137-148	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDRNI TVNEAQSRSGGGGGGGGYRGGGGGGYGGRRREG GGGGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRs
3.	GGGGGYGGG	9	3	100-108; 116-124; 131-139	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDRNI TVNEAQSRSGGGGGGGGYRGGGGGGYGGRRREG GYGGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRs
4.	GGGGGGG	8	2	88-95; 89-96	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDRNI TVNEAQSRSGGGGGGGGYRGGGGGGYGGRRREG GYGGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRs
5.	GGGGGGG	7	2	88-94; 90-96	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDRNI TVNEAQSRSGGGGGGGGYRGGGGGGYGGRRREG GYGGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRs

6.	GGGGGY	7	2	91-97; 99-105	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS GGG <b>GGGGGYRGGGGGY</b> GGGRRREG GYGGGGYGGRRREGYGGGGYGGRRREGGYGGGS EGNWRS
7.	GGYGGGG	7	3	103-109; 113-119; 128-134	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS GGGGGGGYRGGGG <b>GGYGGGGRRREG</b> <b>GYGGGG</b> GYGGRRRE <b>GGYGGGG</b> GYGGRRREGGYGGGS EGNWRS
8.	GGGGGG	6	5	88-93; 89-94; 90-95; 91-96; 99-104	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS <b>GGGGGGGGYRGGGGGG</b> YGGGGRRREG GYGGGGYGGRRREGYGGGGYGGRRREGGYGGGS EGNWRS
9.	GGGGGY	6	3	92-97; 116-121; 131-136	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS GGG <b>GGGGGYRGGGGGGY</b> GGGRRREG GY <b>GGGGGY</b> GGGRRREGY <b>GGGGGY</b> GGGRRREGGYGGGS EGNWRS
10.	GGYGGG	6	6	103-108; 113-118; 119-124; 128-133; 134-139; 143-148	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS GGGGGGGYRGGGG <b>GGYGGGGRRREG</b> <b>GYGGGGGYGGRRRE</b> <b>GGYGGGGGYGGRRREGGYGGGS</b> EGNWRS
11.	GGGGG	5	9	88-92; 89-93; 90-94; 91-95;	MAEVEYRCFVGGLAWATTDQTLGD AFSQFGEILDSKIIN DRETGRSRGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRS <b>GGGGGGGGYRGGGGGG</b> YGGGGRRREG GY <b>GGGGG</b> YGGRRREGY <b>GGGGG</b> YGGRRREGGYGGGS

					92-96; 99-103; 100-104; 116-120; 131-135	EGNWRS
12.	GGGG	4	14		88-91; 89-92; 90-93; 91-94; 92-95; 93-96; 99-102; 100-103; 101-104; 106-109; 116-119; 117-120; 131-134; 132-135	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRRGS GGGGGGGGYR GGGGGGY GGGGRRREG GY GGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRS
13.	GGG	3	22		88-90; 89-91; 90-92; 91-93; 92-94; 93-95; 94-96; 99-101; 100-102; 101-103; 102-104; 106-108;	MAEVEYRCFVGGGLAWATTDQTLGDAFSQFGEILDSKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDLDGRNI TVNEAQSRRGS GGGGGGGGYR GGGGGGY GGGGRRREG GY GGGGYGGRRREGGYGGGGYGGRRREGGYGGGS EGNWRS



				107-109; 116-118; 117-119; 118-120; 122-124; 131-133; 132-134; 133-135; 137-139; 146-148				
14.	GGY	3	4	95-97; 113-115; 128-130; 143-145			MAEVEYRCFVGGGLAWATTDQTLGD AFSQFGEILD SKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDL DGRNI TVNEAQSRGSGGGGGGGGGY RGGGGGGYGGGGRRREG GGYGGGGYGGRRREGGGYGGGGGGYGGRRREGGGYGGGS EGNWRS	
15.	SRG	3	2	46-48; 84-86			MAEVEYRCFVGGGLAWATTDQTLGD AFSQFGEILD SKIIN DRETGRSRGFGFVTFKDEKSMRDAIEGMNGQDL DGRNI TVNEAQSRGSGGGGGGGGGY RGGGGGGYGGGGRRREG GYGGGGGGYGGRRREGGYGGGGGGYGGRRREGGYGGGS EGNWRS	

#### 4.14 Conservation of C-terminal repeats in Class IVa GRP from different species of plants:

Domain analysis, prediction by Globplot, Dotplot, and other online tools predicted the amino acid repeats responsible for disordered nature starts around amino acids “QSR”. Therefore, the C-terminal region from amino acids “QSR” alone picked from Class IVa GRP of different genera of plants and aligned using Clustal Omega . The alignment results revealed that all the proteins taken for the study had identical repeats throughout the stretch such as “GGGYG”, “GYG/GYS/GYR”, “GGR” “RRE/RRD” and especially the extreme C-terminal residue for all the proteins had residues “SE/DGNWRS” (Figure 4.22). Intriguingly, the C-terminal part of Class IVa GRP is highly conserved among different species of genus *Nicotiana* (Figure 4.23).



**Figure 4.22: Multiple sequence alignment of C-terminal disordered region of Class IVa GRP from different organisms using Clustal Omega.** Nt – *Nicotiana tabacum*, Dc – *Daucus carota*, Ng – *Nicotiana glutinosa*, Ee – *Euphorbia esula*, Ta – *Triticum aestivum*, Sa – *Sinapis alba*, Ca – *Capsicum annum*, Hv – *Hordeum vulgare*, Os – *Oryza sativa*, At – *Arabidopsis thaliana*, Nb – *Nicotiana benthamiana*.

```

NgGRP      QSRGSGGGGGGYRGGRRGGGGGGYGGGGRREGGYGGGGGGYGGGRREGGYGGG--GGYGGG      58
NtGRP1     QSRGSG-GGGGGGGYRGGSGGGYGGGGRREGGYGGG-GGYGGGRREGGYGGGGGGYGGG      58
NsGRP      QSRGSG-GGGGGGGYRGGSGGGYGGGGRREGGYGGG-GGYGGGRREGGYGGGGGGYGGG      58
NbGRP3     QSRGSGGGGGGGGGYRGGGGGGYGGGGRREGGYGGG-GGYGGGRREGGYGGG--GGYGGG      57
NaGRP      QSRGSG-GGGGGGGYRGGGGGGYGGGGRREGGYGGG-GGYGGGRREGGYGGGGGGYGGG      58
          *****  **** * **_.*****  *****  *****

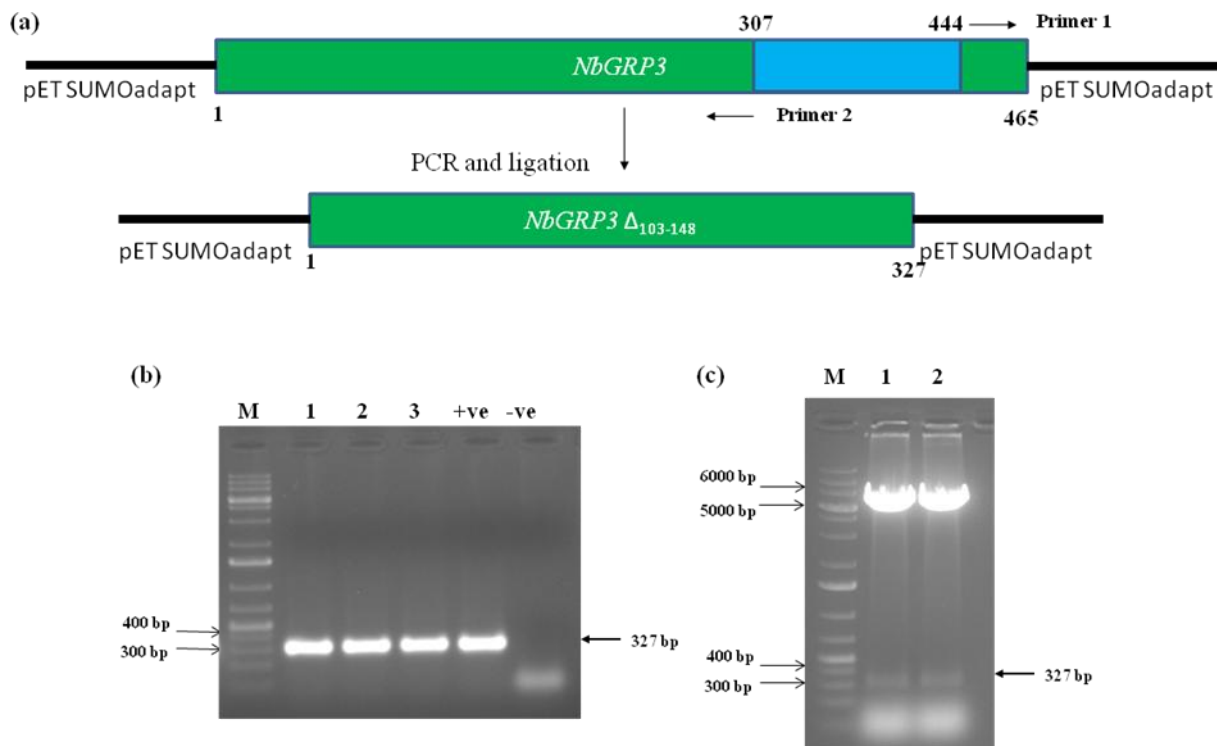
NgGRP      RREGGYGGGSEGNWRN      74
NtGRP1     RREGGYGGGSEGNWRS      74
NsGRP      RREGGYGGGSEGNWRS      74
NbGRP3     RREGGYGGGSEGNWRS      73
NaGRP      RREGGYGGGSEGNWRS      74
          *****.

```

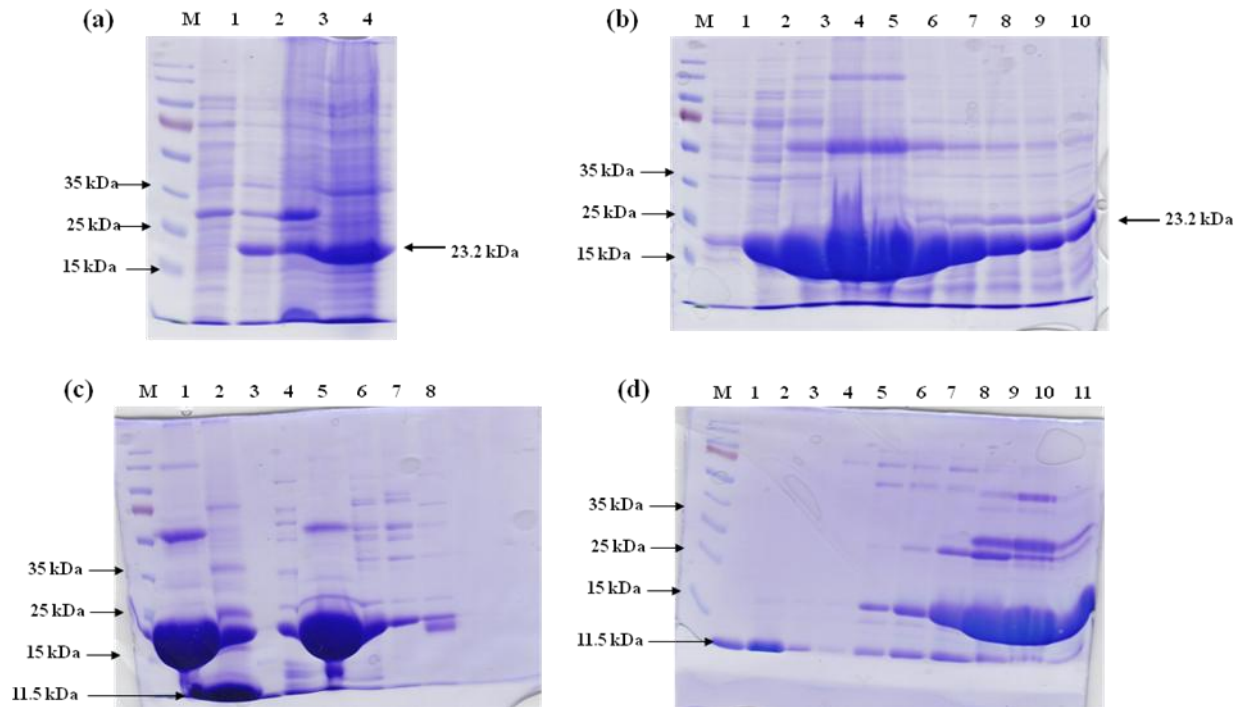
**Figure 4.23: Multiple sequence alignment of the C-terminal disordered region of Class IVa GRP from different species within the genera *Nicotiana* using Clustal Omega.** Alignment shows high conservation of the C-terminal repeats in tandem. Nt – *Nicotiana tabacum*, Ng – *Nicotiana glutinosa*, Nb – *Nicotiana benthamiana*, Ns – *Nicotiana sylvestris*, Na – *Nicotiana attenuate*.

#### 4.15 Construction of C-terminal deletion construct of *NbGRP3*:

From the detailed analysis from FAIR online server and alignment studies, we observed that the C-terminal region has tandem of repeats “GGYGGGRREGGYGG” repeated thrice from the position 103 to 148. Interestingly, abovementioned repeat was conserved in Class IVa GRP from different species of plants. In order to shed light on the repeats, the repeat region was deleted to generate the deletion variant NbGRP3 $\Delta_{103-148}$ .



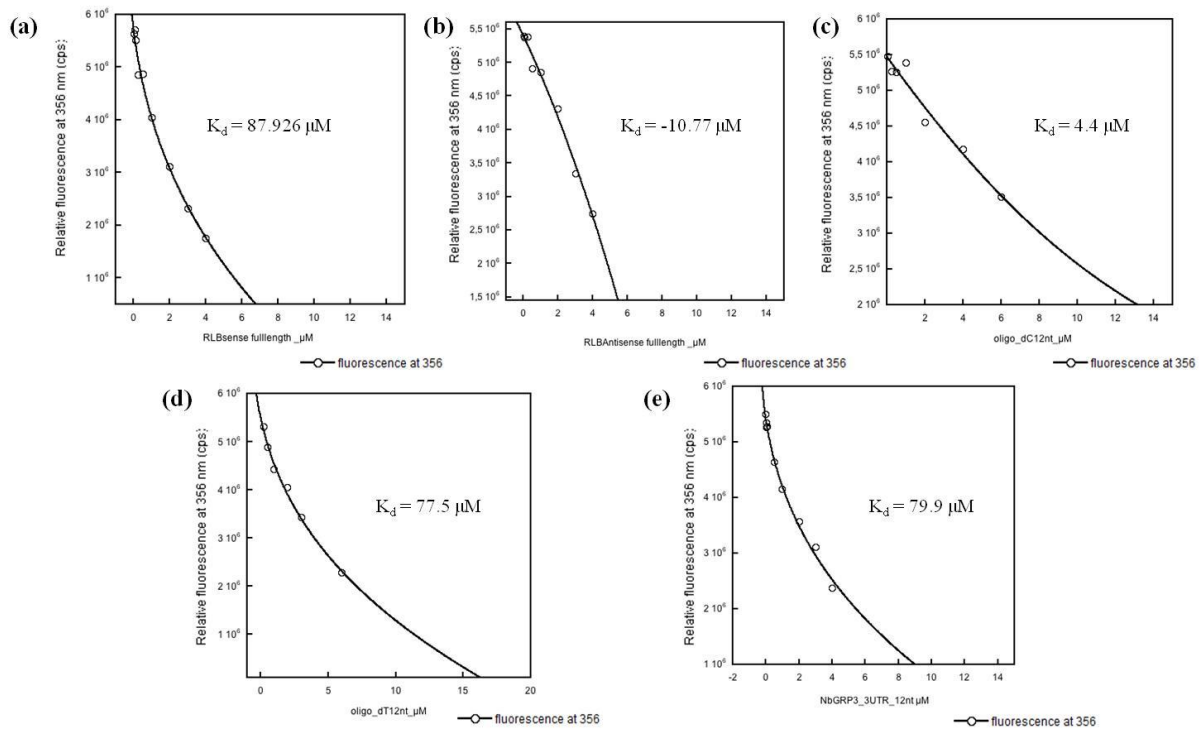
**Figure 4.24: Generation of C-terminal deletion variant  $NbGRP3\Delta_{103-148}$ .** (a) Strategy used for the generation of C-terminal deletion. Primers were designed flanking the repeats targeted. pET SUMOadapt  $NbGRP3$  construct were amplified, and ligated resulting in exclusion of 307-444 bp from full length  $NbGRP3$ .  $NbGRP3\Delta_{103-148}$  did not have amino acids 103-148. (b) PCR confirmation of the deletion construct. (c) Confirmation of deletion by restriction digestion confirmation using the enzyme *BsaI* and *HindIII*.



**Figure 4.25: Different steps involved in expression and purification of NbGRP3 $\Delta_{103-148}$ .** (a) Lane 1, 2 – un-induced and induced total cell extract of BL21 CodonPlus (DE3)-RIPL harboring pET SUMOadapt NbGRP3 $\Delta_{103-148}$ ; lane 3, 4 – pellet and supernatant fractions of IPTG induced lysate (b) Different elute fractions after first Ni-NTA affinity chromatography (c) Lane 1&2 – 6xHis-SUMO-NbGRP3 $\Delta_{103-148}$  (23.2 kDa) before and after SUMO protease treatment; lane 4-8 – fractions of anion exchange chromatography after SUMO protease treatment. (d) Fractions of heparin affinity chromatography: lane 1-4: purified NbGRP3 $\Delta_{103-148}$  (111.5 kDa) eluted as an unbound fraction; lane 5-11 – other elutes which containing impurities.

#### 4.16 NbGRP3 $\Delta_{103-148}$ has impaired nucleic acid binding:

NbGRP3 $\Delta_{103-148}$  when applied to heparin column, protein came out in the flow through itself, which itself was an indication that NbGRP3 $\Delta_{103-148}$  has impaired nucleic acid binding. NbGRP3 $\Delta_{103-148}$  still can exhibit intrinsic tryptophan fluorescence because still the two tryptophan amino acids at the N-terminal and C-terminal part remain untouched. Nucleic acid binding of NbGRP3 $\Delta_{103-148}$  measured using intrinsic fluorescence which revealed that NbGRP3 $\Delta_{103-148}$  lost its affinity towards nucleic acids tested. This signifies the importance of amino acid repeats in the C-terminal region.



**Figure 4.26: Direct binding studies of NbGRP3 $\Delta$ 103-148 with different nucleic acids showing the deletion has impaired nucleic acid binding (a) RaLCB SCR sense strand (b) RaLCB SCR anti-sense strand (c) With (dC)<sub>12</sub> (d) With (dT)<sub>12</sub> and (e) *NbGRP3* 3'UTR putative target site.**

# *Discussion*





## 5. Discussion

Extensive studies on AtGRP7 and partly on other GRP form class IVa GRP, it is now known that this particular class plays pivotal role in numerous biological processes of plants. At the biochemical level, it was shown to have a role in alternative splicing, miRNA biogenesis, mRNA transport, and RNA chaperone. The biological implications of these biochemical activities, it was awe-inspiring. Class IVa GRP was shown to have involved in plant immunity, stress adaptations, circadian rhythm, germination and, development, etc. A good volume of work had been done to understand the role of Class IVa GRP in abiotic stress, whereas not much was done in biotic stress in particular to plant viruses. Its influence on hormones involved in defense signaling makes it clear that this class of protein is essential in plant defense [13, 14]. In virus perspective, not much was done, hardly there were few reports demonstrating its role in viral pathogenesis [15, 19]. In the current study, an attempt was undertaken understand the role of Class IVa glycine rich protein from *N. benthamiana* in geminiviral pathogenesis and ended up in dissection of C-terminal repeats. The plant *N. benthamiana* commonly referred to as an experimental model for plant-virus interaction as it is susceptible for most plant viruses [167]. ToLCNDV-A and its non-cognate betasatellite RaLCB was used as a model virus for our studies as this non-cognate combination induce better symptom induction than the cognate combination [99, 164]. The Class IVa GRP was used for the current study was referred to as homolog of GRP3 from *Populus* sp. Therefore, it was referred to as *NbGRP3*. *NbGRP3* contains 155 amino acids, like any other Class IVa GRP it has two distinct domains N-terminal RRM and C-terminal glycine rich domain. Online disorder prediction tools and computational model predicted that amino acids after amino acid at position 80 were all disordered and are enriched by glycines interspersed by tyrosine, arginine and glutamic acid (Figure 4.1 & 4.2).

To begin with understanding the role of *NbGRP3* in geminiviral pathogenesis, 15 days old *N. benthamiana* plants infected with ToLCNDV-A and ToLCNDV-A along with RaLCB. *NbGRP3* transcript level was checked at 3, 6, 9, 18, and 21 dpi. At 3dpi, the highest expression was observed in both mock as well as virus infected samples and it can be correlated with the reports that Class IVa GRP induced several fold upon wounding [117] as the virus was delivered to the plants by wounding the veins. At further dpi especially at the initial dpi (6 and 9 dpi)

where symptom development has not started, there was upregulation of *NbGRP3* in ToLCNDV-A + RaLCB infected samples. At 18 dpi both ToLCNDV-A and ToLCNDV-A+ RaLCB infected samples had mild upregulation when compared to that of mock. At late dpi (21 dpi) where symptom development attains maximum severity, there was a decrease in transcript level of *NbGRP3* in mock as well as infected samples. Transcript analysis provided a clue on the role of *NbGRP3* in geminiviral pathogenesis (Figure 4.3). Yeast two-hybrid analysis performed to know whether NbGRP3 interact with any ORFs that were shown to be VSRs such as AC2, AV2, AC4 from ToLCNDV and  $\beta$ C1 the only ORF encoded by RaLCB. NbGRP3 didn't interact with any of the ORFs tested (Figure 4.5 & 4.6).

NbGRP3 failed to interact with viral proteins tested but it may interact with viral DNA, geminivirus being single-stranded DNA and NbGRP3 being RNA binding protein and its homologs from other species shown to bind with single-stranded DNA, double-stranded DNA as well apart from RNA [116, 120, 123, 168]. Therefore, *NbGRP3* was successfully cloned into pET SUMO expression system, expressed and purified. Expression studies during geminiviral infection, ToLCNDV + RaLCB had higher expression of *NbGRP3*. To have further insights into the role of NbGRP3 during RaLCB infection, 69 nucleotide region from satellite conserved region (SCR) taken to study its interaction with NbGRP3 using EMSA. Geminivirus being single-stranded DNA virus, single-stranded DNA of SCR was taken for EMSA. EMSA did not provide conclusive results as smeary gel shift observed in EMSA. The similar smeary pattern observed with anti-sense fragment of SCR (Figure 4.10). Literature survey denotes that such smeary pattern of band shift was observed with HvGRP1 and NtGRP1 isolated from *Hordeum vulgare* and *Nicotiana tabacum* respectively [11, 12]. NbGRP3 has two tryptophan molecules, one at N-terminal and another one at C-terminal, which made it ideal for DNA-protein interaction using intrinsic fluorescence by which dissociation constant ( $K_d$ ) was derived. The affinity for NbGRP3 towards sense strand was more than the anti-sense strand of SCR (Figure 4.11). To have more insight into the role of *NbGRP3* in geminiviral pathogenesis, TRV based VIGS of *NbGRP3* was carried out to see the effect of geminivirus with inadequate *NbGRP3* transcript levels. It was found that in NbGRP3 silenced plants had more ToLCNDV-A and RaLCB DNA accumulation than that of non-silenced plants which signifies the anti-viral role of NbGRP3 (Figure 4.15). To know whether the anti-viral role of NbGRP3 was restricted to

specific group of viruses, *NbGRP3* transcript level was checked in TBSV infected samples. Interestingly, the *NbGRP3* transcript level was also upregulated in TBSV infected samples and the transcript level of *NbGRP3* was compromised in TBSV infected samples without their VSR p19 (Figure 4.16). From these observations recorded in this study, it can be concluded that *NbGRP3* may have a broad anti-viral defense. This could be because of the fact that NbGRP3 could regulate defense-related genes. It is to be noted that AtGRP7 overexpression lines had upregulation of defense-related genes [17].

Smear band shift from EMSA led to characterize the binding behavior of NbGRP3 may differ from other nucleic acid binding proteins and the reports from NMR studies hinted on the dynamic behavior of C-terminal GR domain. As a step towards understanding it, direct binding with different homodeoxyribonucleotides of 12 nucleotides in size such as dA, dG, dT and dC performed, NbGRP3 showed affinity towards polypyrimidines (dC)<sub>12</sub> and (dT)<sub>12</sub> (Figure 4.17). This was in-contrast to other class IVa GRPs. Class IVa from *Phycomitrella patens*, *Capsicum annum* had affinity towards poly (G). The affinity of NbGRP3 towards polypyrimidines indicated that NbGRP3 have role involved in RNA metabolisms [169] like its homologs AtGRP7 shown to bind to its own pre-mRNA at its 3'-UTR to regulate its own transcript and the nucleotide "5'-TTCTGG-3'" is essential for the binding [121]. Interestingly, 3'-UTR of NbGRP3 contains "5'-TTATGG-3'" sequence. Removal of "5'-TTC-3'" reduces the affinity of AtGRP7 towards its 3'-UTR to great extent. To know whether the transversion affected the binding affinity of NbGRP3 to putative 3'-UTR binding site, directed binding performed with AtGRP7 3'-UTR target site and NbGRP3 3'-UTR the putative target site. The transversion, of course, affected the binding of NbGRP3 to NbGRP3 3'-UTR putative target site but not to a greater extent like that of AtGRP7 (Figure 4.20). This points out to the fact that although GRPs from different plant species share significant homology yet they vary in their biochemical properties.

NMR results from HvGRP1 and NtGRP1 pointed out intra and intermolecular association exhibited by the full-length protein. To have further insights, yeast two-hybrid analysis performed with AD-NbGRP3 and BD-NbGRP3 with appropriate control, which showed NbGRP3 does not exhibit self-interaction (Figure 4.19). Analytical ultracentrifugation performed

further with purified NbGRP3 which concluded that NbGRP3 exists in a monomeric state in the absence of nucleic acids. NbGRP3, upon titration with different concentration of (dC)<sub>12</sub>, it undergoes a shift from monomer to dimer to monomer according to the concentration of ligand (Figure 4.18), which very well explains the observation of smeary pattern in EMSA. This observation also supports the dynamism exhibited by NbGRP3 in presence of nucleic acids.

Amino acid repeats in glycine enriched glycine-rich protein was never a novel thing. In fact, one of the criteria considered for classification of the glycine-rich proteins were the presence of glycine repeats etc. [109]. The initial few reports on GRP were all from Class IVa GRP which has conserved RRM followed by glycine rich C-terminal region. Reports from maize, sorghum and carrot had a mention of “GGYGG” and “RRE”. Different scientific groups noticed repeats in C-terminal region, but no one explored the repeats [115, 117, 170]. For largely unknown reasons not much exploration gone into understanding of the C-terminal repeats but two reasons can be assumed, one reason could be the notion that the disordered regions exhibit lower conservation [171]. Another reason could be discouraging results obtained after deletion of C-terminal GR domains. Deletion of the C-terminal domain didn't affect nucleic acid binding in case of CaGRP1 [123] and did not promote growth under cold stress in AtGRP7 [118]. Interesting observations reported in case Class IVb and c GRP as it had a zinc finger motif at their C-terminal. As it was a well-established fact that zinc finger motif interacts with nucleic acids. Therefore, deletion of C-terminal domain affected the nucleic acid binding [172] but such exploration has not been undertaken in Class IVa GRP before 2014 which was nearly two and a half decade after the discovery of first Class IVa GRP. The significance of C-terminal part of Class IVa GRP came into notice only after NMR studies with Class IVa GRP from organisms such as *N. tabacum* and *H. vulgare*. NMR studies with nucleic acids indicated the dynamic role of glycine rich domain in nucleic acid binding. The C-terminal part exhibited dynamic and intra and intermolecular interaction. The interaction can be disturbed by varying temperatures and pH. Interesting data from HvGRP1 showed that the glycine rich region alone can interact with nucleic acids. NtGRP1 and HvGRP1 both exhibit a smeary pattern of a shift in EMSA and it is assumed to be because of dynamic behavior exhibited by glycine rich domain [11, 12]. Considering the aforesaid literature evidences and upon closer examination of the C-terminal domain revealed something which was interesting. C-terminal of NbGRP3 has

“GGYGGGRREGGYGGG” amino acid repeat arranged in tandem. Surprisingly, the repeats described above were conserved among different species of plants (Figure 4.22 & 4.23). To know the role of these repeats, *NbGRP3* in pET SUMOadapt modified using PCR based approach by designing the primers flanking the repeats in such a way that whole pET SUMOadapt -*NbGRP3* would be amplified excluding the repeats. The cloned variant has the amino acids from 103 to 148 excluded ( $NbGRP3\Delta_{103-148}$ ). The cloned variant was expressed and purified. During the purification process,  $NbGRP3\Delta_{103-148}$  could not bind to the heparin column (Figure 4.25d). This was a sign that  $NbGRP3\Delta_{103-148}$  lost nucleic acid binding property. Direct binding was performed using intrinsic fluorescence as the purified protein has two tryptophans. Binding studies revealed that the  $NbGRP3\Delta_{103-148}$  could not bind to any nucleic acid tested (Figure 4.26). This signifies the importance of the C-terminal glycine rich domain. Overall, in the current study, an attempt was made to understand the role of a class IVa GRP in viral pathogenesis and found the anti-viral role of *NbGRP3*. The biochemical studies made have further helped in understanding the role disordered region of an RNA binding protein in nucleic acid binding.



# *Summary*





## 6. Summary

Geminiviruses pose serious a threat to agricultural productivity as they infect economically important crops in tropical and sub-tropical regions. Geminiviruses are strict obligate parasites, and their genome encodes for 5-7 proteins performing various functions such as replication, transcription, movement, etc. Therefore, they rely heavily on the host machinery for the abovementioned functions [3]. Host factors play key role in viral pathogenesis. Host factors either support by host defense against the virus or support the virus in completing the pathogenesis [4].

Glycine rich proteins (GRP) are the superfamily of protein with high glycine content shown to perform varied functions. GRPs classified into five different classes based on the conserved motifs and position of glycine rich region. Among the five different classes, class IVa deserves special attention as it was shown to perform various biological events such as seed germination, circadian rhythm, biotic, and abiotic stress responses. Very less information was available in understanding the role of class IVa GRP in biotic stress, especially viral pathogenesis [15, 19]. To have a step towards understanding the role of class IVa GRP in viral pathogenesis, the current study was undertaken by using geminivirus and class IVa GRP from *N. benthamiana* which is considered as a model organism for host-virus interaction.

*NbGRP3* had elevated expression during geminiviral infection (ToLCNDV, RaLCB) and similar observation recorded with TBSV which is an RNA virus. *NbGRP3* failed to interact with VSRs encoded by ToLCNDV-A and RaLCB but *NbGRP3* bound to SCR region of RaLCB, SCR is the most conserved region among geminiviral beta satellites. To have more acuity, TRV-based VIGS performed to silence *NbGRP3* and to observe the effect of geminiviral pathogenesis under low level of *NbGRP3*. VIGS revealed that *NbGRP3* silenced plants had more viral DNA accumulation than that of control plants. These observations, along with literature evidences denote that *NbGRP3* may have an anti-viral role during viral pathogenesis in general.

The binding between SCR of RaLCB and *NbGRP3* in EMSA was smeary and such smeary protein-DNA complex observed with other class IVa GRPs from *H. vulgare* and *N. tabacum*. To have some understanding of the smeary protein-DNA complex on EMSA, some

biochemical characterization was carried out. It was found that NbGRP3 had more affinity towards deoxypyrimidines, which are the characteristics of proteins involved in RNA metabolism. The binding properties of NbGRP3 differed from other members of class IVa GRP.

Analytical ultracentrifugation analysis had shed light on the unusual binding behavior of NbGRP3. NbGRP3 without any ligand exist as monomer whereas in the presence of nucleic acid it undergoes the transition from monomer to dimer and again to monomer based on the concentration of the nucleic acids. This observation explains the dynamism exhibited by class IVa GRP during nucleic acid binding.

Upon closer scrutiny of sequence of NbGRP3, it was observed that class IVa GRP has conserved pattern of amino acid repeats at the C-terminal glycine rich region. In an attempt to understand the role of the repeats, C-terminal deletion variant of NbGRP3 (NbGRP3 $\Delta_{103-148}$ ) was constructed. NbGRP3 $\Delta_{103-148}$  lacked the amino acids from 103 to 148, it was the region where the enrichment of repeats as tandem was found. Interestingly, the NbGRP3 $\Delta_{103-148}$  when applied to heparin column, the protein came in flow through which itself was an early sign that NbGRP3 $\Delta_{103-148}$  has impaired nucleic acid binding and this observation supported by the direct binding experiments. The current study signifies the importance of amino acid repeats in the disordered region in nucleic acid binding.

# *References*



## 7. References

1. Zerbini, F.M., et al., *ICTV Virus Taxonomy Profile: Geminiviridae*. J Gen Virol, 2017. **98**(2): p. 131-133.
2. Krupovic, M., J.J. Ravantti, and D.H. Bamford, *Geminiviruses: a tale of a plasmid becoming a virus*. BMC Evolutionary Biology, 2009. **9**(1): p. 112.
3. Hanley-Bowdoin, L., et al., *Geminiviruses: masters at redirecting and reprogramming plant processes*. Nat Rev Microbiol, 2013. **11**(11): p. 777-88.
4. Calil, I.P. and E.P.B. Fontes, *Plant immunity against viruses: antiviral immune receptors in focus*. Ann Bot, 2017. **119**(5): p. 711-723.
5. Jones, J.D.G. and J.L. Dangl, *The plant immune system*. Nature, 2006. **444**(7117): p. 323-329.
6. Hashimoto, M., et al., *Recessive Resistance to Plant Viruses: Potential Resistance Genes Beyond Translation Initiation Factors*. Frontiers in microbiology, 2016. **7**: p. 1695-1695.
7. Nicaise, V., *Crop immunity against viruses: outcomes and future challenges*. Frontiers in plant science, 2014. **5**: p. 660-660.
8. Csorba, T., V. Pantaleo, and J. Burgyán, *Chapter 2 - RNA Silencing: An Antiviral Mechanism*, in *Advances in Virus Research*, G. Loebenstein and J.P. Carr, Editors. 2009, Academic Press. p. 35-230.
9. D Collum, T. and J. N Culver, *The impact of phytohormones on virus infection and disease*. Vol. 17. 2016. 25-31.
10. Luo, H., *Interplay between the virus and the ubiquitin-proteasome system: molecular mechanism of viral pathogenesis*. Curr Opin Virol, 2016. **17**: p. 1-10.
11. Tripet, B.P., et al., *Structural and biochemical analysis of the Hordeum vulgare L. HvGR-RBP1 protein, a glycine-rich RNA-binding protein involved in the regulation of barley plant development and stress response*. Biochemistry, 2014. **53**(50): p. 7945-60.
12. Khan, F., et al., *Structural basis of nucleic acid binding by Nicotiana tabacum glycine-rich RNA-binding protein: implications for its RNA chaperone function*. Nucleic Acids Res, 2014. **42**(13): p. 8705-18.
13. Ciuzan, O., et al., *The evolutionarily conserved multifunctional glycine-rich RNA-binding proteins play key roles in development and stress adaptation*. Physiol Plant, 2015. **153**(1): p. 1-11.
14. Czolpinska, M. and M. Rurek, *Plant Glycine-Rich Proteins in Stress Response: An Emerging, Still Prospective Story*. Front Plant Sci, 2018. **9**: p. 302.
15. Lee, H.J., et al., *Different roles of glycine-rich RNA-binding protein7 in plant defense against Pectobacterium carotovorum, Botrytis cinerea, and tobacco mosaic viruses*. Plant Physiol Biochem, 2012. **60**: p. 46-52.
16. Fu, Z.Q., et al., *A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity*. Nature, 2007. **447**(7142): p. 284-8.
17. Streitner, C., et al., *Global transcript profiling of transgenic plants constitutively overexpressing the RNA-binding protein AtGRP7*. BMC Plant Biol, 2010. **10**: p. 221.
18. Hackmann, C., et al., *Salicylic acid-dependent and -independent impact of an RNA-binding protein on plant immunity*. Plant Cell Environ, 2014. **37**(3): p. 696-706.
19. Huang, X., et al., *Identification and characterisation of a glycine-rich RNA-binding protein as an endogenous suppressor of RNA silencing from Nicotiana glutinosa*. Planta, 2019. **249**(6): p. 1811-1822.

20. Brown, J.K., et al., *Revision of Begomovirus taxonomy based on pairwise sequence comparisons*. Arch Virol, 2015. **160**(6): p. 1593-619.
21. Hamilton, W.D., et al., *Complete nucleotide sequence of the infectious cloned DNA components of tomato golden mosaic virus: potential coding regions and regulatory sequences*. Embo j, 1984. **3**(9): p. 2197-205.
22. Stanley, J. and M.R. Gay, *Nucleotide sequence of cassava latent virus DNA*. Nature, 1983. **301**(5897): p. 260-262.
23. Pradhan, B., et al., *Molecular Biology of Geminivirus DNA Replication*. 2017. p. 2-31.
24. Roshan, P., A. Kulshreshtha, and V. Hallan, *Genome Organization of Begomoviruses, in Begomoviruses: Occurrence and Management in Asia and Africa*, S. Saxena and A.K. Tiwari, Editors. 2017, Springer Singapore: Singapore. p. 11-32.
25. Liu, H., M.I. Boulton, and J.W. Davies, *Maize streak virus coat protein binds single- and double-stranded DNA in vitro*. J Gen Virol, 1997. **78 ( Pt 6)**: p. 1265-70.
26. Liu, H., et al., *A single amino acid change in the coat protein of Maize streak virus abolishes systemic infection, but not interaction with viral DNA or movement protein*. Mol Plant Pathol, 2001. **2**(4): p. 223-8.
27. Liu, L., et al., *Bean yellow dwarf virus RepA, but not rep, binds to maize retinoblastoma protein, and the virus tolerates mutations in the consensus binding motif*. Virology, 1999. **256**(2): p. 270-9.
28. Poornima Priyadarshini, C.G., et al., *Functional characterization of coat protein and V2 involved in cell to cell movement of Cotton leaf curl Kokhran virus-Dabawali*. PLoS One, 2011. **6**(11): p. e26929.
29. Padidam, M., R.N. Beachy, and C.M. Fauquet, *The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus*. Virology, 1996. **224**(2): p. 390-404.
30. Chowda-Reddy, R.V., et al., *Role of a geminivirus AV2 protein putative protein kinase C motif on subcellular localization and pathogenicity*. Virus Res, 2008. **135**(1): p. 115-24.
31. Rothenstein, D., et al., *Tissue and cell tropism of Indian cassava mosaic virus (ICMV) and its AV2 (precoat) gene product*. Virology, 2007. **359**(1): p. 137-45.
32. Glick, E., et al., *Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(1): p. 157-161.
33. Zhang, J., et al., *V2 protein encoded by Tomato yellow leaf curl China virus is an RNA silencing suppressor*. Virus Res, 2012. **163**(1): p. 51-8.
34. Wang, Y., et al., *Geminiviral V2 Protein Suppresses Transcriptional Gene Silencing through Interaction with AGO4*. Journal of Virology, 2019. **93**(6): p. e01675-18.
35. Orozco, B.M., et al., *Functional domains of a geminivirus replication protein*. J Biol Chem, 1997. **272**(15): p. 9840-6.
36. Rizvi, I., N.R. Choudhury, and N. Tuteja, *Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication*. Arch Virol, 2015. **160**(2): p. 375-87.
37. Ruhel, R. and S. Chakraborty, *Multifunctional roles of geminivirus encoded replication initiator protein*. Virusdisease, 2019. **30**(1): p. 66-73.
38. Rodriguez-Negrete, E., et al., *Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing*. New Phytol, 2013. **199**(2): p. 464-75.

39. Kushwaha, N.K., Mansi, and S. Chakraborty, *The replication initiator protein of a geminivirus interacts with host monoubiquitination machinery and stimulates transcription of the viral genome*. 2017. **13**(8): p. e1006587.
40. Haley, A., et al., *Regulation of the activities of African cassava mosaic virus promoters by the AC1, AC2, and AC3 gene products*. Virology, 1992. **188**(2): p. 905-9.
41. Sunter, G., et al., *Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication*. Virology, 1990. **179**(1): p. 69-77.
42. Sunter, G. and D.M. Bisaro, *Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression*. Virology, 1991. **180**(1): p. 416-9.
43. Hartitz, M.D., G. Sunter, and D.M. Bisaro, *The tomato golden mosaic virus transactivator (TrAP) is a single-stranded DNA and zinc-binding phosphoprotein with an acidic activation domain*. Virology, 1999. **263**(1): p. 1-14.
44. Lacatus, G. and G. Sunter, *The Arabidopsis PEAPOD2 transcription factor interacts with geminivirus AL2 protein and the coat protein promoter*. Virology, 2009. **392**(2): p. 196-202.
45. Shivaprasad, P.V., et al., *Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus*. J Virol, 2005. **79**(13): p. 8149-63.
46. Chowda-Reddy, R.V., et al., *Characterization of the cassava geminivirus transcription activation protein putative nuclear localization signal*. Virus Res, 2009. **145**(2): p. 270-8.
47. Chandran, S.A., et al., *Mapping of functional region conferring nuclear localization and karyopherin alpha-binding activity of the C2 protein of bhendi yellow vein mosaic virus*. J Gen Virol, 2012. **93**(Pt 6): p. 1367-74.
48. Chandran, S.A., C. Jeyabharathy, and R. Usha, *The C2 protein of Bhendi yellow vein mosaic virus plays an important role in symptom determination and virus replication*. Virus Genes, 2014. **48**(1): p. 203-7.
49. Wang, H., et al., *Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins*. J Virol, 2005. **79**(12): p. 7410-8.
50. Zhang, Z., et al., *BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in Arabidopsis*. Plant Cell, 2011. **23**(1): p. 273-88.
51. Yang, L.P., et al., *C2-mediated decrease in DNA methylation, accumulation of siRNAs, and increase in expression for genes involved in defense pathways in plants infected with beet severe curly top virus*. Plant J, 2013. **73**(6): p. 910-7.
52. Li, P., et al., *Plant begomoviruses subvert ubiquitination to suppress plant defenses against insect vectors*. 2019. **15**(2): p. e1007607.
53. Etessami, P., et al., *Mutational analysis of complementary-sense genes of African cassava mosaic virus DNA A*. J Gen Virol, 1991. **72 ( Pt 5)**: p. 1005-12.
54. Sunter, G., D.C. Stenger, and D.M. Bisaro, *Heterologous complementation by geminivirus AL2 and AL3 genes*. Virology, 1994. **203**(2): p. 203-10.
55. Settlage, S.B., A.B. Miller, and L. Hanley-Bowdoin, *Interactions between geminivirus replication proteins*. J Virol, 1996. **70**(10): p. 6790-5.
56. Pasumarthy, K.K., N.R. Choudhury, and S.K. Mukherjee, *Tomato leaf curl Kerala virus (ToLCKeV) AC3 protein forms a higher order oligomer and enhances ATPase activity of replication initiator protein (Rep/AC1)*. Virol J, 2010. **7**: p. 128.
57. Settlage, S.B., R.G. See, and L. Hanley-Bowdoin, *Geminivirus C3 Protein: Replication Enhancement and Protein Interactions*. Journal of Virology, 2005. **79**(15): p. 9885.
58. Settlage, S.B., et al., *Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator*. Virology, 2001. **279**(2): p. 570-6.

59. Selth, L.A., et al., *A NAC domain protein interacts with tomato leaf curl virus replication accessory protein and enhances viral replication*. *Plant Cell*, 2005. **17**(1): p. 311-25.
60. Pasumarthy, K.K., S.K. Mukherjee, and N.R. Choudhury, *The presence of tomato leaf curl Kerala virus AC3 protein enhances viral DNA replication and modulates virus induced gene-silencing mechanism in tomato plants*. *Virology*, 2011. **8**: p. 178.
61. Rigden, J.E., et al., *ORF C4 of tomato leaf curl geminivirus is a determinant of symptom severity*. *Virology*, 1994. **204**(2): p. 847-50.
62. Latham, J.R., et al., *Induction of plant cell division by beet curly top virus gene C4*. *The Plant Journal*, 1997. **11**(6): p. 1273-1283.
63. Mills-Lujan, K. and C.M. Deom, *Geminivirus C4 protein alters Arabidopsis development*. *Protoplasma*, 2010. **239**(1-4): p. 95-110.
64. Park, J., et al., *C4 protein of Beet severe curly top virus is a pathomorphogenetic factor in Arabidopsis*. *Plant Cell Rep*, 2010. **29**(12): p. 1377-89.
65. Jupin, I., et al., *Movement of tomato yellow leaf curl geminivirus (TYLCV): involvement of the protein encoded by ORF C4*. *Virology*, 1994. **204**(1): p. 82-90.
66. Teng, K., et al., *Involvement of C4 Protein of Beet Severe Curly Top Virus (Family Geminiviridae) in Virus Movement*. *PLOS ONE*, 2010. **5**(6): p. e11280.
67. Vanitharani, R., et al., *Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing*. *J Virol*, 2004. **78**(17): p. 9487-98.
68. Chellappan, P., R. Vanitharani, and C.M. Fauquet, *MicroRNA-binding viral protein interferes with Arabidopsis development*. *Proceedings of the National Academy of Sciences of the United States of America*, 2005. **102**(29): p. 10381-10386.
69. Fondong, V.N., et al., *The consensus N-myristoylation motif of a geminivirus AC4 protein is required for membrane binding and pathogenicity*. *Mol Plant Microbe Interact*, 2007. **20**(4): p. 380-91.
70. Carluccio, A.V., et al., *S-acylation mediates Mungbean yellow mosaic virus AC4 localization to the plasma membrane and in turns gene silencing suppression*. *PLOS Pathogens*, 2018. **14**(8): p. e1007207.
71. Dogra, S.C., et al., *A novel shaggy-like kinase interacts with the Tomato leaf curl virus pathogenicity determinant C4 protein*. *Plant Mol Biol*, 2009. **71**(1-2): p. 25-38.
72. Kheyr-Pour, A., et al., *Watermelon chlorotic stunt virus from the Sudan and Iran: Sequence Comparisons and Identification of a Whitefly-Transmission Determinant*. *Phytopathology*, 2000. **90**(6): p. 629-35.
73. Fontenelle, M.R., et al., *Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus Tomato chlorotic mottle virus*. *Virus Res*, 2007. **126**(1-2): p. 262-7.
74. Melgarejo, T.A., et al., *Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution*. *J Virol*, 2013. **87**(10): p. 5397-413.
75. Raghavan, V., et al., *The DNA-A Component of a Plant Geminivirus (Indian Mung Bean Yellow Mosaic Virus) Replicates in Budding Yeast Cells*. *Journal of Virology*, 2004. **78**(5): p. 2405-2413.
76. Li, F., et al., *The AC5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses*. *New Phytologist*, 2015. **208**(2): p. 555-569.
77. Pascal, E., et al., *The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus*. *Plant Cell*, 1994. **6**(7): p. 995-1006.



78. Noueir, A.O., W.J. Lucas, and R.L. Gilbertson, *Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport*. *Cell*, 1994. **76**(5): p. 925-32.
79. McGarry, R.C., et al., *A Novel Arabidopsis Acetyltransferase Interacts with the Geminivirus Movement Protein NSP*. *The Plant Cell*, 2003. **15**(7): p. 1605.
80. Rojas, M.R., et al., *Bean Dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner*. *Cell*, 1998. **95**(1): p. 105-13.
81. Ward, B.M. and S.G. Lazarowitz, *Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay*. *Plant Cell*, 1999. **11**(7): p. 1267-76.
82. Zhou, Y., et al., *Histone H3 interacts and colocalizes with the nuclear shuttle protein and the movement protein of a geminivirus*. *J Virol*, 2011. **85**(22): p. 11821-32.
83. Ye, J., et al., *Geminivirus Activates ASYMMETRIC LEAVES 2 to Accelerate Cytoplasmic DCP2-Mediated mRNA Turnover and Weakens RNA Silencing in Arabidopsis*. *PLOS Pathogens*, 2015. **11**(10): p. e1005196.
84. Li, R., et al., *Virulence Factors of Geminivirus Interact with MYC2 to Subvert Plant Resistance and Promote Vector Performance*. *The Plant Cell*, 2014. **26**(12): p. 4991-5008.
85. Carvalho, M.F. and S.G. Lazarowitz, *Interaction of the movement protein NSP and the Arabidopsis acetyltransferase AtNSI is necessary for Cabbage leaf curl geminivirus infection and pathogenicity*. *J Virol*, 2004. **78**(20): p. 11161-71.
86. Carvalho, M.F., R. Turgeon, and S.G. Lazarowitz, *The geminivirus nuclear shuttle protein NSP inhibits the activity of AtNSI, a vascular-expressed Arabidopsis acetyltransferase regulated with the sink-to-source transition*. *Plant Physiol*, 2006. **140**(4): p. 1317-30.
87. Fontes, E.P., et al., *The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity*. *Genes Dev*, 2004. **18**(20): p. 2545-56.
88. Florentino, L.H., et al., *A PERK-like receptor kinase interacts with the geminivirus nuclear shuttle protein and potentiates viral infection*. *J Virol*, 2006. **80**(13): p. 6648-56.
89. Hussain, M., et al., *The nuclear shuttle protein of Tomato leaf curl New Delhi virus is a pathogenicity determinant*. *J Virol*, 2005. **79**(7): p. 4434-9.
90. Frischmuth, S., et al., *The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast*. *Protoplasma*, 2007. **230**(1-2): p. 117-23.
91. Zhang, S.C., C. Wege, and H. Jeske, *Movement proteins (BC1 and BV1) of Abutilon mosaic geminivirus are cotransported in and between cells of sink but not of source leaves as detected by green fluorescent protein tagging*. *Virology*, 2001. **290**(2): p. 249-60.
92. Radhakrishnan, G.K., G.A. Splitter, and R. Usha, *DNA recognition properties of the cell-to-cell movement protein (MP) of soybean isolate of Mungbean yellow mosaic India virus (MYMIV-Sb)*. *Virus Res*, 2008. **131**(2): p. 152-9.
93. Duan, Y.P., et al., *Geminivirus Resistance in Transgenic Tobacco Expressing Mutated BC1 Protein*. *Molecular Plant-Microbe Interactions*, 1997. **10**(5): p. 617-623.
94. Hou, Y.-M., et al., *Transgenic Plants Expressing Geminivirus Movement Proteins: Abnormal Phenotypes and Delayed Infection by Tomato mottle virus in Transgenic Tomatoes Expressing the Bean dwarf mosaic virus BV1 or BC1 Proteins*. *Molecular Plant-Microbe Interactions*, 2000. **13**(3): p. 297-308.
95. Zhou, X., *Advances in understanding begomovirus satellites*. *Annu Rev Phytopathol*, 2013. **51**: p. 357-81.
96. Fiallo-Olive, E., et al., *A novel class of DNA satellites associated with New World begomoviruses*. *Virology*, 2012. **426**(1): p. 1-6.

97. Li, F., et al., *The betaC1 Protein of Geminivirus-Betasatellite Complexes: A Target and Repressor of Host Defenses*. Mol Plant, 2018. **11**(12): p. 1424-1426.
98. Yang, J.Y., et al., *betaC1, the pathogenicity factor of TYLCCNV, interacts with AS1 to alter leaf development and suppress selective jasmonic acid responses*. Genes Dev, 2008. **22**(18): p. 2564-77.
99. Bhattacharyya, D., et al., *A geminivirus betasatellite damages the structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis*. J Exp Bot, 2015. **66**(19): p. 5881-95.
100. Gnanasekaran, P., K. Ponnusamy, and S. Chakraborty, *A geminivirus betasatellite encoded betaC1 protein interacts with PsbP and subverts PsbP-mediated antiviral defence in plants*. 2019.
101. Eini, O., et al., *Interaction with a host ubiquitin-conjugating enzyme is required for the pathogenicity of a geminiviral DNA beta satellite*. Mol Plant Microbe Interact, 2009. **22**(6): p. 737-46.
102. Yang, X., et al., *Suppression of Methylation-Mediated Transcriptional Gene Silencing by  $\beta$ C1-SAHH Protein Interaction during Geminivirus-Betasatellite Infection*. PLOS Pathogens, 2011. **7**(10): p. e1002329.
103. Zhong, X., et al., *Mimic Phosphorylation of a  $\beta$ C1 Protein Encoded by TYLCCNB Impairs Its Functions as a Viral Suppressor of RNA Silencing and a Symptom Determinant*. Journal of Virology, 2017. **91**(16): p. e00300-17.
104. Li, F., et al., *Suppression of RNA Silencing by a Plant DNA Virus Satellite Requires a Host Calmodulin-Like Protein to Repress RDR6 Expression*. PLOS Pathogens, 2014. **10**(2): p. e1003921.
105. Shen, Q., et al., *Tobacco RING E3 Ligase NtRFP1 Mediates Ubiquitination and Proteasomal Degradation of a Geminivirus-Encoded betaC1*. Mol Plant, 2016. **9**(6): p. 911-25.
106. Shen, Q., et al., *Tomato SLSnRK1 Protein Interacts with and Phosphorylates  $\beta$ C1, a Pathogenesis Protein Encoded by a Geminivirus  $\beta$ -Satellite*. Plant Physiology, 2011. **157**(3): p. 1394-1406.
107. Haxim, Y., et al., *Autophagy functions as an antiviral mechanism against geminiviruses in plants*. 2017. **6**.
108. Lozano, G., et al., *Characterization of Non-coding DNA Satellites Associated with Sweepoviruses (Genus Begomovirus, Geminiviridae) - Definition of a Distinct Class of Begomovirus-Associated Satellites*. Front Microbiol, 2016. **7**: p. 162.
109. Sachetto-Martins, G., L.O. Franco, and D.E. de Oliveira, *Plant glycine-rich proteins: a family or just proteins with a common motif?* Biochim Biophys Acta, 2000. **1492**(1): p. 1-14.
110. Mangeon, A., R.M. Junqueira, and G. Sachetto-Martins, *Functional diversity of the plant glycine-rich proteins superfamily*. Plant Signal Behav, 2010. **5**(2): p. 99-104.
111. Condit, C.M. and R.B. Meagher, *A gene encoding a novel glycine-rich structural protein of petunia*. Nature, 1986. **323**(6084): p. 178-181.
112. Kielkopf, C.L., S. Lucke, and M.R. Green, *U2AF homology motifs: protein recognition in the RRM world*. Genes Dev, 2004. **18**(13): p. 1513-26.
113. Schoning, J.C., et al., *Auto-regulation of the circadian slave oscillator component AtGRP7 and regulation of its targets is impaired by a single RNA recognition motif point mutation*. Plant J, 2007. **52**(6): p. 1119-30.
114. Wang, S., et al., *Molecular characterization and expression analysis of a glycine-rich RNA-binding protein gene from Malus hupehensis Rehd*. Mol Biol Rep, 2012. **39**(4): p. 4145-53.
115. Gomez, J., et al., *A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein*. Nature, 1988. **334**(6179): p. 262-4.

116. Hirose, T., M. Sugita, and M. Sugiura, *cDNA structure, expression and nucleic acid-binding properties of three RNA-binding proteins in tobacco: occurrence of tissue-specific alternative splicing*. Nucleic Acids Res, 1993. **21**(17): p. 3981-7.
117. Sturm, A., *A Wound-Inducible Glycine-Rich Protein from Daucus carota with Homology to Single-Stranded Nucleic Acid-Binding Proteins*. Plant Physiol, 1992. **99**(4): p. 1689-92.
118. Kim, J.S., et al., *Cold shock domain proteins and glycine-rich RNA-binding proteins from Arabidopsis thaliana can promote the cold adaptation process in Escherichia coli*. Nucleic Acids Res, 2007. **35**(2): p. 506-16.
119. Jarvelin, A.I., et al., *The new (dis)order in RNA regulation*. Cell Commun Signal, 2016. **14**: p. 9.
120. Gao, X., et al., *A glycine-rich protein MoGrp1 functions as a novel splicing factor to regulate fungal virulence and growth in Magnaporthe oryzae*. Phytopathology Research, 2019. **1**(1): p. 2.
121. Schuttpelz, M., et al., *Changes in conformational dynamics of mRNA upon AtGRP7 binding studied by fluorescence correlation spectroscopy*. J Am Chem Soc, 2008. **130**(29): p. 9507-13.
122. Nomata, T., Y. Kabeya, and N. Sato, *Cloning and characterization of glycine-rich RNA-binding protein cDNAs in the moss Physcomitrella patens*. Plant Cell Physiol, 2004. **45**(1): p. 48-56.
123. Kim, D.S., N.H. Kim, and B.K. Hwang, *GLYCINE-RICH RNA-BINDING PROTEIN1 interacts with RECEPTOR-LIKE CYTOPLASMIC PROTEIN KINASE1 and suppresses cell death and defense responses in pepper (Capsicum annuum)*. New Phytol, 2015. **205**(2): p. 786-800.
124. Han, J.H., et al., *The RNA chaperone and protein chaperone activity of Arabidopsis glycine-rich RNA-binding protein 4 and 7 is determined by the propensity for the formation of high molecular weight complexes*. Protein J, 2013. **32**(6): p. 449-55.
125. Naqvi, S.M., et al., *A glycine-rich RNA-binding protein gene is differentially expressed during acute hypersensitive response following Tobacco Mosaic Virus infection in tobacco*. Plant Mol Biol, 1998. **37**(3): p. 571-6.
126. Hsiao, P.Y., et al., *The Arabidopsis defensin gene, AtPDF1.1, mediates defence against Pectobacterium carotovorum subsp. carotovorum via an iron-withholding defence system*. Sci Rep, 2017. **7**(1): p. 9175.
127. Deslandes, L. and S. Rivas, *Catch me if you can: bacterial effectors and plant targets*. Trends Plant Sci, 2012. **17**(11): p. 644-55.
128. Ruggieri, G.M., et al., *Overexpression of glycine-rich RNA-binding protein in tomato renders fruits with higher protein content after cold storage*. Biologia Plantarum, 2018. **62**(3): p. 501-510.
129. Kim, J.S., et al., *Glycine-rich RNA-binding protein 7 affects abiotic stress responses by regulating stomata opening and closing in Arabidopsis thaliana*. Plant J, 2008. **55**(3): p. 455-66.
130. Kim, J.Y., et al., *Glycine-rich RNA-binding proteins are functionally conserved in Arabidopsis thaliana and Oryza sativa during cold adaptation process*. J Exp Bot, 2010. **61**(9): p. 2317-25.
131. Yang, D.H., et al., *Expression of Arabidopsis glycine-rich RNA-binding protein AtGRP2 or AtGRP7 improves grain yield of rice (Oryza sativa) under drought stress conditions*. Plant Sci, 2014. **214**: p. 106-12.
132. Kwak, K.J., et al., *Diverse roles of glycine-rich RNA-binding protein 7 in the response of camelina (Camelina sativa) to abiotic stress*. Acta Physiologiae Plantarum, 2016. **38**(5): p. 129.
133. Kwak, K.J., et al., *Molecular cloning, characterization, and stress-responsive expression of genes encoding glycine-rich RNA-binding proteins in Camelina sativa L.* Plant Physiol Biochem, 2013. **68**: p. 44-51.
134. Wang, B., et al., *A Glycine-Rich RNA-Binding Protein, CsGR-RBP3, Is Involved in Defense Responses Against Cold Stress in Harvested Cucumber (Cucumis sativus L.) Fruit*. Front Plant Sci, 2018. **9**: p. 540.

135. Streitner, C., et al., *The small glycine-rich RNA binding protein AtGRP7 promotes floral transition in Arabidopsis thaliana*. Plant J, 2008. **56**(2): p. 239-250.
136. Xiao, J., et al., *O-GlcNAc-mediated interaction between VER2 and TaGRP2 elicits TaVRN1 mRNA accumulation during vernalization in winter wheat*. Nature Communications, 2014. **5**: p. 4572.
137. Xiao, J., et al., *JACALIN-LECTIN LIKE1 Regulates the Nuclear Accumulation of GLYCINE-RICH RNA-BINDING PROTEIN7, Influencing the RNA Processing of *FLOWERING LOCUS C* Antisense Transcripts and Flowering Time in Arabidopsis*. Plant Physiology, 2015. **169**(3): p. 2102-2117.
138. van Kan, J.A., B.J. Cornelissen, and J.F. Bol, *A virus-inducible tobacco gene encoding a glycine-rich protein shares putative regulatory elements with the ribulose biphosphate carboxylase small subunit gene*. Mol Plant Microbe Interact, 1988. **1**(3): p. 107-12.
139. Khan, F., et al., *Dynamics of mRNA of glycine-rich RNA-binding protein during wounding, cold and salt stresses in Nicotiana tabacum*. 2013.
140. Lohr, B., et al., *A glycine-rich RNA-binding protein affects gibberellin biosynthesis in Arabidopsis*. Mol Biol Rep, 2014. **41**(1): p. 439-45.
141. Aneeta, et al., *Salinity- and ABA-induced up-regulation and light-mediated modulation of mRNA encoding glycine-rich RNA-binding protein from Sorghum bicolor*. Biochem Biophys Res Commun, 2002. **296**(5): p. 1063-8.
142. Cao, S., et al., *AtGRP7 is involved in the regulation of abscisic acid and stress responses in Arabidopsis*. Cell Mol Biol Lett, 2006. **11**(4): p. 526-35.
143. Bosse-Doenecke, E., et al., *High yield production of recombinant native and modified peptides exemplified by ligands for G-protein coupled receptors*. Protein Expression and Purification, 2008. **58**(1): p. 114-121.
144. James, P., J. Halladay, and E.A. Craig, *Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast*. Genetics, 1996. **144**(4): p. 1425-36.
145. Hofgen, R. and L. Willmitzer, *Storage of competent cells for Agrobacterium transformation*. Nucleic Acids Res, 1988. **16**(20): p. 9877.
146. Birnboim, H.C. and J. Doly, *A rapid alkaline extraction procedure for screening recombinant plasmid DNA*. Nucleic Acids Res, 1979. **7**(6): p. 1513-23.
147. Reich, S., et al., *Initiation of RNA Synthesis by the Hepatitis C Virus RNA-Dependent RNA Polymerase Is Affected by the Structure of the RNA Template*. Biochemistry, 2014. **53**(44): p. 7002-7012.
148. Daniel Gietz, R. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method*, in *Methods in Enzymology*, C. Guthrie and G.R. Fink, Editors. 2002, Academic Press. p. 87-96.
149. Ali, S.H. and J.A. DeCaprio, *Cellular transformation by SV40 large T antigen: interaction with host proteins*. Semin Cancer Biol, 2001. **11**(1): p. 15-23.
150. Chattopadhyay, B., et al., *Infectivity of the cloned components of a begomovirus: DNA beta complex causing chilli leaf curl disease in India*. Archives of Virology, 2008. **153**(3): p. 533-539.
151. Sahu, P.P., et al., *Tomato cultivar tolerant to Tomato leaf curl New Delhi virus infection induces virus-specific short interfering RNA accumulation and defence-associated host gene expression*. Mol Plant Pathol, 2010. **11**(4): p. 531-44.
152. Liu, Y., M. Schiff, and S.P. Dinesh-Kumar, *Virus-induced gene silencing in tomato*. Plant J, 2002. **31**(6): p. 777-86.
153. Dellaporta, S.L., J. Wood, and J.B. Hicks, *A plant DNA miniprep: version II*. Plant molecular biology reporter, 1983. **1**(4): p. 19-21.

154. Southern, E.M., *Detection of specific sequences among DNA fragments separated by gel electrophoresis*. Journal of molecular biology, 1975. **98**(3): p. 503-517.
155. Andrade, M.A., et al., *Homology-based method for identification of protein repeats using statistical significance estimates*. J Mol Biol, 2000. **298**(3): p. 521-37.
156. Karpenahalli, M.R., A.N. Lupas, and J. Soding, *TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences*. BMC Bioinformatics, 2007. **8**: p. 2.
157. Heringa, J. and P. Argos, *A method to recognize distant repeats in protein sequences*. Proteins, 1993. **17**(4): p. 391-41.
158. Heger, A. and L. Holm, *Rapid automatic detection and alignment of repeats in protein sequences*. Proteins, 2000. **41**(2): p. 224-37.
159. Biegert, A. and J. Soding, *De novo identification of highly diverged protein repeats by probabilistic consistency*. Bioinformatics, 2008. **24**(6): p. 807-14.
160. Alba, M.M., R.A. Laskowski, and J.M. Hancock, *Detecting cryptically simple protein sequences using the SIMPLE algorithm*. Bioinformatics, 2002. **18**(5): p. 672-8.
161. Gruber, M., J. Söding, and A.N. Lupas, *REPPER—repeats and their periodicities in fibrous proteins*. Nucleic acids research, 2005. **33**(suppl\_2): p. W239-W243.
162. Senthilkumar, R., et al., *FAIR: A server for internal sequence repeats*. Bioinformatics, 2010. **4**(7): p. 271-275.
163. Linding, R., et al., *GlobPlot: exploring protein sequences for globularity and disorder*. Nucleic Acids Research, 2003. **31**(13): p. 3701-3708.
164. Singh, A.K., B. Chattopadhyay, and S. Chakraborty, *Biology and interactions of two distinct monopartite begomoviruses and betasatellites associated with radish leaf curl disease in India*. Virol J, 2012. **9**: p. 43.
165. Ghisaidoobe, A.B.T. and S.J. Chung, *Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques*. International journal of molecular sciences, 2014. **15**(12): p. 22518-22538.
166. Maris, C., C. Dominguez, and F.H.-T. Allain, *The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression*. The FEBS Journal, 2005. **272**(9): p. 2118-2131.
167. Goodin, M.M., et al., *Nicotiana benthamiana: its history and future as a model for plant-pathogen interactions*. Mol Plant Microbe Interact, 2008. **21**(8): p. 1015-26.
168. Kwak, K.J., Y.O. Kim, and H. Kang, *Characterization of transgenic Arabidopsis plants overexpressing GR-RBP4 under high salinity, dehydration, or cold stress*. J Exp Bot, 2005. **56**(421): p. 3007-16.
169. Sawicka, K., et al., *Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein*. Biochem Soc Trans, 2008. **36**(Pt 4): p. 641-7.
170. Cretin, C. and P. Puigdomenech, *Glycine-rich RNA-binding proteins from Sorghum vulgare*. Plant Mol Biol, 1990. **15**(5): p. 783-5.
171. Dunker, A.K., et al., *Function and structure of inherently disordered proteins*. Curr Opin Struct Biol, 2008. **18**(6): p. 756-64.
172. Park, S.J., et al., *The C-terminal zinc finger domain of Arabidopsis cold shock domain proteins is important for RNA chaperone activity during cold adaptation*. Phytochemistry, 2010. **71**(5-6): p. 543-7.



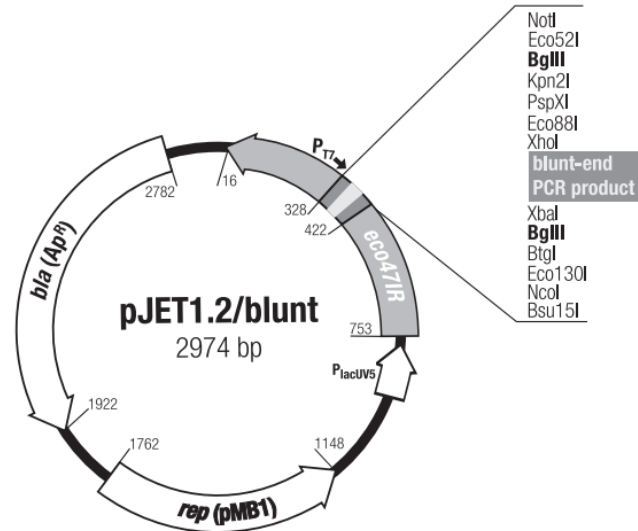
# *Appendix*



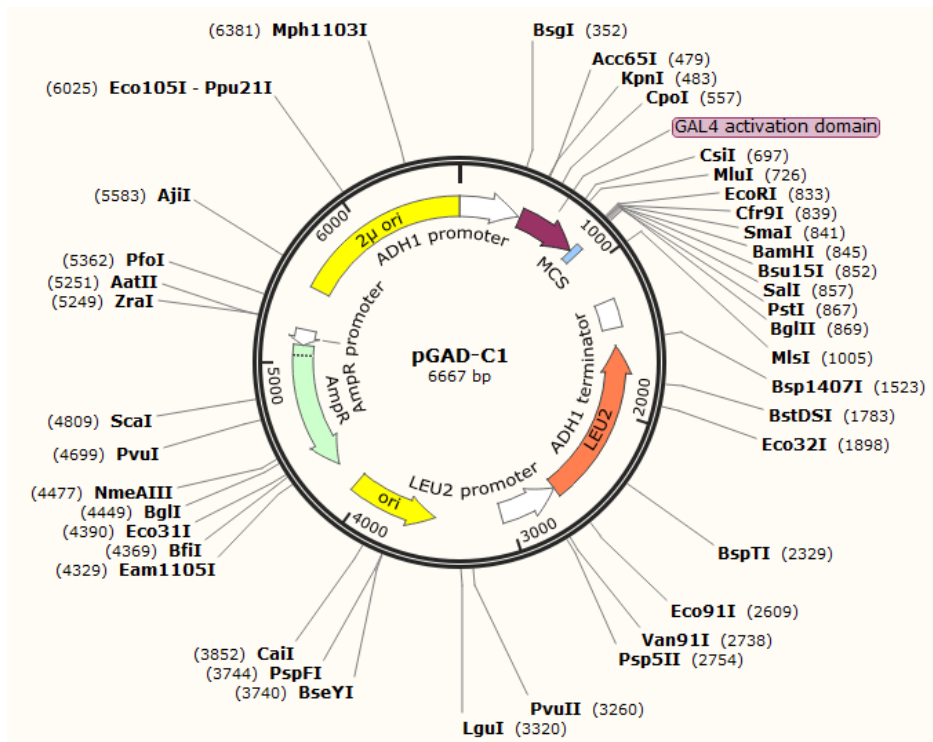


## 9. Appendix

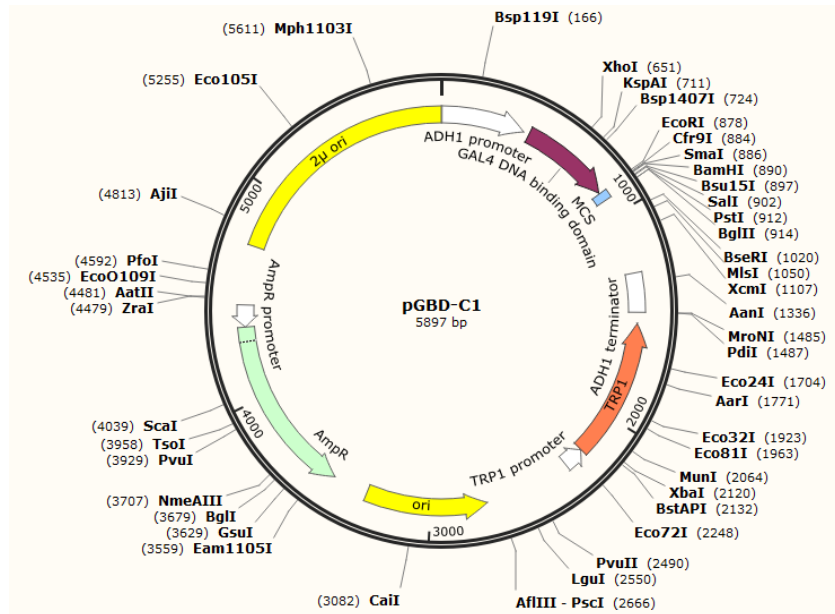
### pJET1.2 vector:



### pGAD-C1:

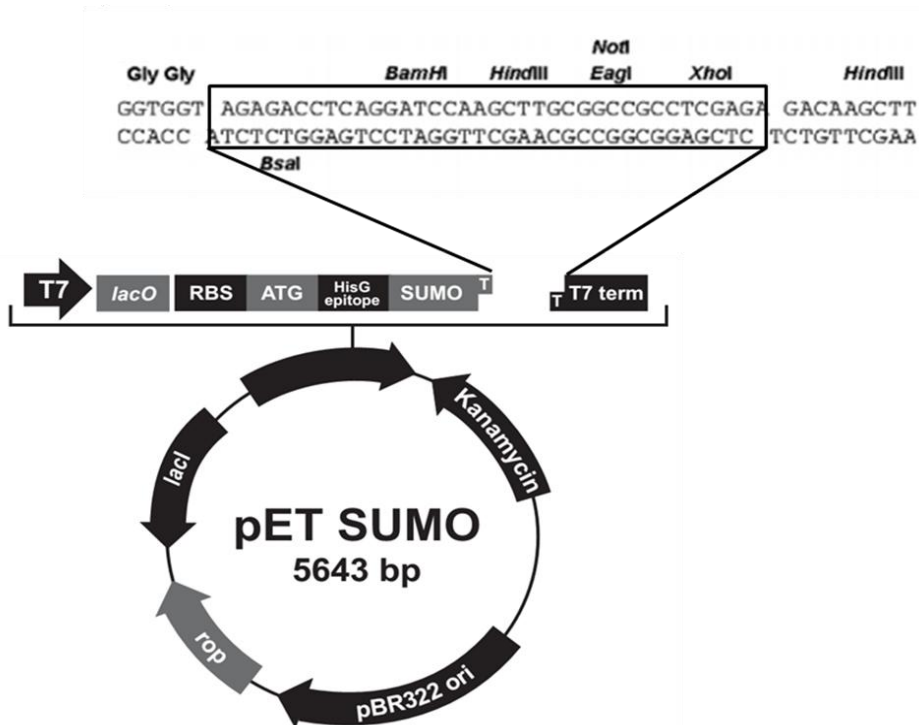


## pGBD-C1:

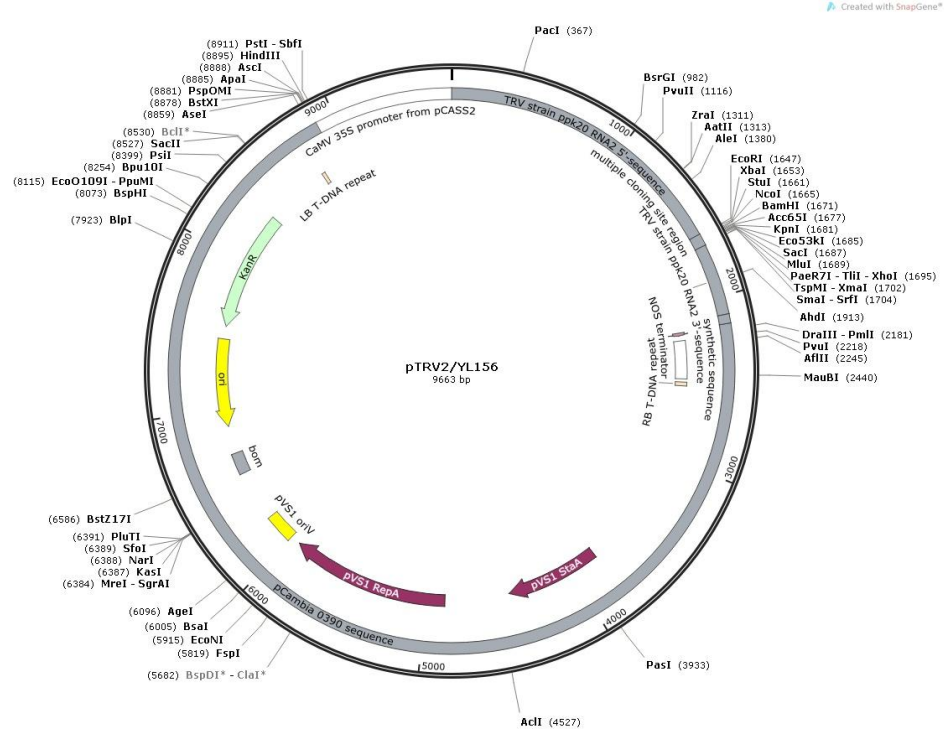


## pET SUMOadapt:

It is a modified version of pET SUMO by having sequences with restriction sites inserted between the T-overhangs to enhance efficiency of cloning.



# pTRV2:



### Preparation of media, buffers and reagents:

Name	Preparation
------	-------------

#### Commonly used reagents

1 M Tris-Cl (Ph-8.0)	121 g of Tris base dissolved in 800 ml of double distilled water, pH adjusted to 8.0 using HCl and volume made up to 1000 ml using double distilled water. Solution sterilized by autoclaving.
5 N NaOH	200 g of NaOH pellets dissolved in 800 ml of water, volume adjusted to 1000 ml once the pellets were dissolved properly. Solution was stored at room temperature.
10 % SDS	100 g of electrophoresis grade SDS was dissolved 900 ml of double distilled water by constant stirring with magnetic stirrer. The final volume adjusted to 1000 ml with double distilled water.
0.5 M EDTA, pH 8.0	186.1 g of disodium EDTA. 2H <sub>2</sub> O added to 800 ml of double distilled water. It was stirred constantly on a magnetic stirrer and pH adjusted to 8.0.
50 X Tris acetate EDTA (TAE)	242 g of Tris base, 57.2 ml of glacial acetic acid, 0.5 M EDTA, pH – 8.0. Volume made up to 1000 ml with distilled water.
6 X loading dye	0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cynol FF, 30 % (v/v) glycerol. All the components dissolved in water and stored at 4 °C.
Ethidium bromide (10mg/ml)	1 g of ethidium bromide dissolved in 100 ml of distilled water, dissolved using magnetic stirrer for several hours and stored at room temperature in dark bottles.
LB media	1g of tryptone, 0.5 g of yeast extract, 1g of NaCl dissolved in water, pH adjusted to 7.0 and volume made up to 100 ml with distilled water. Sterilized by autoclaving.
LB agar	1.5 % of agar added to LB media
100 mM MgCl <sub>2</sub>	2.03 f of MgCl <sub>2</sub> .6H <sub>2</sub> O dissolved in double distilled water, volume made upto 100 ml with the same. Sterilized by

autoclaving.

100 mM CaCl<sub>2</sub> 1.47 g of CaCl<sub>2</sub>.2H<sub>2</sub>O dissolved in double distilled water, volume made upto 100 ml with the same. Sterilized by autoclaving.

1 M IPTG (Isopropyl thio- $\alpha$ -galactoside) 2.3 g of IPTG dissolved in 10 ml of double distilled water, filter sterilized through 0.22  $\mu$  filter and stored at -20 °C

### **Plasmid isolation reagents:**

Lysis Solution I 2.5 ml of 1M Tris-HCl (25 mM) pH – 8.0, 5.5 ml of 20% glucose (50 mM), 2.0 ml of 0.5 m EDTA (10 mM). The final volume made up to 100 ml using double distilled water

Lysis Solution II 2.0 ml of 10N NaOH (0.2N), 10.0 ml of 10% SDS. The volume was made up to 100 ml using double distilled water

Lysis Solution III 5M Potassium acetate – 60 ml (3M), glacial acetic acid – 11.5 ml, distilled water – 28.5 ml and the solution stored at 4 °C

### **Southern hybridization reagents:**

20 X SSC 175.3 g of NaCl, 88.2 g of sodicum citrate, dissolved in distilled water, pH adjusted to 7.0 by adding HCl and final volume made up to 1000 ml by distilled water

Denaturation solution 1.5 M NaCl, 0.5 N NaOH dissolved in distilled water.

Depurination solution 2.8 ml of HCl added to 197.2 ml of distilled water.

Neutralization solution 1.5 M NaCl, 1.0 M Tris-HCl (pH-7.2) dissolved and final volume made with distilled water.

50 X Denhardt's solution 1 % (w/v) Ficoll, 1 % (w/v) polyvinylpyrrolidone (PVP), 1 % (w/v) Bovine serum albumin dissolved in double distilled water.

10 mg/ml Herring sperm DNA 100 mg of DNA dissolved in 10 ml of double distilled water by vortexing and stored at -20 °C

Hybridization buffer 0.25 M sodium phosphate buffer (pH 7.2), 7% (w/v) SDS, 1 mM EDTA dissolved and final volume made with distilled

water.

**Yeast two hybrid reagents:**

10 X lithium acetate (LiAc)	1 M LiAc was prepared in double distilled water, pH adjusted to 7.5 with dilute acetic acid. Sterilized by autoclaving and stored at room temperature.
50 % (w/v) Poly ethylene glycol (PEG) 3350	50 g of PEG dissolved and volume made with double distilled water. Sterilized by autoclaving and stored at room temperature.
10 X Tris-EDTA (TE) buffer	100 mM Tris-Cl (pH 7.5), 10 mM EDTA (pH 8.0) solution sterilized by autoclaving and stored at room temperature.
YPAD broth	20 g of Difco peptone, 10 g of Yeast extract, dissolved in 960 ml of double distilled water, pH adjusted to 5.8. 40 mg adenine hemisulfate added to it. The media sterilized by autoclave and 2 % (w/v) of sterile glucose added to the autoclaved media when it is cooled to bearable warmth.
YPAD agar	20 g of Difco peptone, 10 g of Yeast extract, 15-20 g of agar, dissolved in 960 ml of double distilled water, pH adjusted to 5.8. 40 mg adenine hemisulfate added to it. The media sterilized by autoclave and 2 % (w/v) of sterile glucose added to the autoclaved media when it is cooled to bearable warmth.
Synthetic dropout media	6.7 g of Difco yeast nitrogen base without aminoacids dissolved in 860 ml of deionized water, pH adjusted to 5.8. The media sterilized by autoclave and 2 % (w/v) of sterile glucose added to the autoclaved media when it is cooled to bearable warmth.
Synthetic dropout agar	6.7 g of Difco yeast nitrogen base without aminoacids, 15-20 g of agar dissolved in 860 ml of deionized water, pH adjusted to 5.8. The media sterilized by autoclave and 2 % (w/v) of sterile glucose added to the autoclaved media when it is cooled to bearable warmth.

**Agroinfiltration:**

Infiltration buffer            10 mM MES buffer, 10 mM MgCl<sub>2</sub>, 100 μM acetosyringone, pH adjusted to 5.5-5.8.

**Protein expression and purification:**

1 M Tris-Cl (pH 6.8)            121.1 g of Tris base dissolved in 800 ml of distilled water, pH adjusted to 6.8 using HCl and final volume adjusted to 1000 ml using double distilled water.

1 M Tris-Cl (pH 8.8)            121.1 g of Tris base dissolved in 800 ml of distilled water, pH adjusted to 8.8 using HCl and final volume adjusted to 1000 ml using double distilled water.

30 % (w/v) Acrylamide solution    29 g of acrylamide, 1 g of N, N'-methylenebisacrylamide, dissolved and final volume made up to 100 ml with double distilled water.

0.5 M Tris(2-carboxy ethyl)phosphine (TCEP)    5.73 g of TCEP dissolved initially in 35 ml of cold molecular biology grade water, pH adjusted to 7.0 using NaOH. Final volume made to 100 ml using molecular biology grade water.

1M Dithiothreitol (DTT)    15.45 g of DTT dissolved in 100 ml of water

10 % (w/v) Ammonium per sulfate (APS)    1 g of APS dissolved in 10 ml of water and stored at 4 °C.

1 X Tris-glycine SDS gel running buffer    25 mM Tris, 192 mM glycine, dissolved in double distilled water, pH adjusted to 8.3. 0.1 % SDS added and final volume made using double distilled water.





*Publications and  
Conference abstracts and  
awards*



## 10. List of publications, conference abstracts and awards

### Publications:

R. Vinoth Kumar, H.C. Prasanna, A.K. Singh, **D. Ragunathan**, G.K. Garg, S. Chakraborty. Molecular genetic analysis and evolution of begomoviruses and betasatellites causing yellow mosaic disease of bhendi. *Virus genes* (2017) 53: 275-285

Ved Prakash, **Ragunathan Devendran**, Supriya Chakraborty. Overview of plant RNA dependent RNA polymerases in antiviral defense and gene silencing. *Ind J Plant Physiol* (2017) 22(4): 493-505

### Conferences:

Manish Kumar, R Vinoth kumar, **D. Ragunathan**, Supriya Chakraborty. Identification of begomovirus and associated satellite components in *Alcea Rosea L.* in India. International Geminivirus Symposium - 2016, New Delhi.

**Ragunathan D**, Golbik RP, Tamilarasan S, Behrens SE, Chakraborty S. Understanding the role of glycine rich RNA binding proteins in viral pathogenesis. *Intervirocon – 2018, Chandigarh (Won best poster award)*

### Awards:

Awarded best poster award in the conference “Intervirocon-2018“ held in Chandigarh from 12th – 14th November 2018

Awarded Erasmus Mundus Action plan 2 (BRAVE) scholarship to visit Martin-Luther University, Halle-Wittenberg, Germany for a period of 10 months from December 2016 – September 2017.





**8<sup>th</sup> International Geminivirus Symposium  
&  
6<sup>th</sup> International ssDNA Comparative Virology Workshop**

November 07-10, 2016



*This is to certify that Ragunathan Devendran, India has actively participated / made a poster / oral presentation in the 8<sup>th</sup> International Geminivirus Symposium & 6<sup>th</sup> International ssDNA Comparative Virology Workshop held during November 07-10, 2016 at New Delhi, India.*

*Arinban Roy*  
Secretary

*Basavaprabhu L. Patil*  
Joint Secretary

*Sunil Kumar Taksli*  
Sunil K. Mukherjee  
Co-Chair

*Indnil Dasgupta*  
Indranil Dasgupta  
Chair



**PA45****Molecular characterization of virus causing yellow mosaic disease in chayote**Sangeetha B, **Renukadevi P\***, Malathi V.G., Suganthi M and Alice DDept. of Plant Pathology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu-641003 (\*E-mail: [renucbe88@gmail.com](mailto:renucbe88@gmail.com))

Chayote (*Chow-chow-*Sechium edule**) is a cucurbitaceous vegetable crop grown in hilly regions of India in Karnataka, Tamil Nadu, West Bengal and Himachal Pradesh for its nutritious tasty fruits. It is grown in, Sirumalai, Thandikudi, Adallur, Kodaikanal hills of Dindigul District and Kothagiri of Nilgiris District of Tamil Nadu. The crop is infected by powdery mildew and downy mildew pathogens and a severe yellow mosaic disease. Among these, the yellow mosaic disease is widely prevalent and causes >90% yield loss. The leaves show mosaic mottling, yellowing and puckering. In extreme cases warts like dark green enations are observed on the under surface of the leaves. The fruits from symptomatic plants were found to be malformed and reduced in size. In the survey conducted it was observed that the seedlings do not express any symptom up to sixth month and only at the time of trailing when foliage gets exposed, symptoms are visible. Mandal *et al*, 2004 recorded a begomovirus closely related to ToLCNDV in yellow mosaic affected chayote in Sikkim and Darjeeling hills. An investigation was conducted on the etiological agent of the severe yellow mosaic disease in Tamil Nadu by analyzing the infected leaf, fruit samples from six locations in Dindigul district and kothagiri hills of Nilgiris district in Tamil Nadu. In dot immuno binding assay using antibody to African cassava mosaic virus and in PCR with Roja's primers (PAR1v772 5'GGNAAARATHGGATGGA3'; PAL1c1960 5'ACNGGNAARACNATGTGGGC3', Paximadis *et al*, 1997) positive results were obtained indicating the presence of a begomovirus in infected plants. Further characterization of the virus by Rolling circle amplification and cloning, the causal agent was identified as Tomato leaf curl New Delhi virus (ToLCNDV) since it showed 97.2% identity with ToLCNDV (GenBank accession no-KP191047) earlier described from chayote. However both the isolates from chayote showed only 92% identity with an isolate of ToLCNDV from ridge gourd (KJ1426903.1). Interestingly, the DNA B component exhibited very low identity with DNA B component of ToLCNDV (HM989846.1 and DQ169057.1), in fact only 65% of genome could be aligned and upstream of coding regions, ORFs BC1 and B1 did not find any identity with any of the begomoviruses. Preliminary study revealed the presence of virus in fruit tissues. Among seven seed samples analyzed virus was detected in seed coat, endosperm and in embryonic tissues. The virus was also detected in symptom free seedlings in 5 out of 17 seedlings studied in grow-out tests. From all the observations it can be inferred that ToLCNDV-Chayote, is seed borne and may play a role in disease spread. Accordingly efforts are being taken up to replace seed material with virus free seeds to ensure freedom from spread of the disease.

**PA46****Identification of begomovirus and associated satellite components in *Alcea rosea* L. in India****Manish Kumar**, R. Vinoth Kumar, D. Ragunathan, Supriya Chakraborty\*School of Life Sciences, Jawaharlal Nehru University, New Delhi-110067, India (\*E-mail: [mannuranchi@gmail.com](mailto:mannuranchi@gmail.com))

Hollyhock (*Alcea rosea* L.) is an annual ornamental plant grown in the tropical and subtropical regions in India. Geminiviruses are encapsidated in twinned icosahedral particle which infect large variety of crops around the world and causes a new threat to global food security. During February 2016, begomovirus-like symptoms were noticed on the hollyhock plants grown in a garden at the campus of Jawaharlal Nehru University, New Delhi. Since the hollyhock plants exhibited typical symptoms of yellow vein mosaic, we anticipated the possibility of association of begomoviruses in these plants. Rolling circle amplification was performed to clone the associated begomoviral genomic components from the total DNA isolated from these infected plants. We have identified an isolate of the Cotton leaf curl Rajasthan virus, Luffa associated betasatellite and Okra leaf curl alphasatellite in association with yellow mosaic disease infected hollyhock samples. The phylogenetic and recombination analysis of the cloned begomoviral genomic components were analyzed and will be presented.





Indian Virological Society & PGIMER, Chandigarh

# INTERVIROCON 2018

27<sup>th</sup> International Conference of Virology  
10<sup>th</sup> - 14<sup>th</sup> November 2018

Postgraduate Institute of Medical Education and Research, Chandigarh, India

## Participation Certificate

This is to certify that

Dr./Mr./Ms **Ragunathan D** .....

has participated as a Delegate / Presented Scientific Paper (Oral / Poster ) in the conference

**“Global Viral Epidemics: A Challenging Threat”**

Organized by the Department of Virology, Postgraduate Institute of Medical Education and Research, Chandigarh, India  
from 12<sup>th</sup> - 14<sup>th</sup> November 2018.

  
**Dr. G.P. Rao**

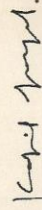
Secretary General  
Indian Virological Society



**Dr. RK Ratho**  
Chairperson  
INTERVIROCON 2018



**Dr. Mini P Singh**  
Organizing Secretary  
INTERVIROCON 2018



**Dr. Kapil Goyal**  
Joint Organizing Secretary  
INTERVIROCON 2018







# INTERNATIONAL CONFERENCE OF VIROLOGY

## “Global Viral Epidemics: A Challenging Threat”



Session		Plant Virology Poster	
Reg. No.	Poster No.	TOPIC	Speaker/Author
	PPP01	Current status of viral diseases of cucurbitaceous crops in India	Mahesha B.
	PPP02	Reverse Transcription-Loop Mediated Isothermal Amplification: A rapid diagnostic assay for detection of <i>Potato virus A</i> in potato and aphids	BaswarajRaigond
	PPP03	A comprehensive study on resistance to Tomato leaf curl virus against various combinations of Ty genes in tomato	Pardhasaradhi P
ON3	PPP04	Understanding the role of glycine rich RNA binding proteins in viral pathogenesis	Ragunathan Devendran
ON2	PPP05	Characterization of begomovirus and associated satellite components causing leaf curl disease in <i>Alcea rosea</i> L.	Kumar Manish
ON101	PPP06	Identification of Brassica yellows virus, a Polerovirus in Radish ( <i>Raphanus sativus</i> )	Suveditha S
ON8	PPP07	Identification of differentially expressed small RNAs following Tomato leaf curl virus infection in <i>N. benthamiana</i> .	Ved Prakash
	PPP08	Development of chimeric promoters from parareterovirus for strong expression of ectopic gene in plant system	Dipinte Gupta
	PPP09	Development of efficient synthetic Promoter from plant parareterovirus	Dipinte Gupta
	PPP10	Identification of chilli ( <i>Capsicum annum</i> L.) lines with enhanced resistance to cucumber mosaic virus under controlled and field conditions	Kavyashri V V
	PPP11	Emerging virus and phytoplasma diseases in Rajouri District of Jammu and Kashmir	Mohd Ashaq
ON245	PPP12	First report of Soybean Yellow Mottle Mosaic Virus on <i>Glycine max</i> in India	Ankita Tripathi
ON60	PPP13	Development of gene deletion or point mutation based constructs of croton yellow vein mosaic virus to understand their suitability as replicon vector	Gurpreet Kaur
	PPP14	Synonymous codon usage pattern of Citrus tristeza virus reveals high codon adaptability to <i>Citrus reticulata</i> , an evolutionary primitive citrus host	Supratik Palchoudhury
	PPP15	Development of single-tube multiplex polymerase chain reaction assay for simultaneous detection of Cotton leaf curl begomovirus and associated satellite molecules	Vivek Kumar Khare



programs and novel bio-technological interventions like genome editing (CRISPR/Cas9) techniques can be explored in the management of cucurbit viruses.

### Reverse transcription-loop mediated isothermal amplification: a rapid diagnostic assay for detection of *Potato virus A* in potato and aphids

Baswaraj Raigond<sup>\*1</sup>, Verma A<sup>1</sup>, Pathania S<sup>1</sup>, Jandrajupalli S<sup>2</sup>, Verma G<sup>1</sup>, Kochhar T<sup>1</sup> and Chakrabarti SK<sup>1</sup>

Email ID for Correspondence: raigond@gmail.com

<sup>1</sup>ICAR-Central Potato Research Institute, Shimla, Himachal Pradesh-171001, India; <sup>2</sup>ICAR-National Institute of Biotic Stress Management, Raipur, Chhattisgarh- 493225, India

*Potato virus A* is an important virus infecting potato, transmitted through infected tubers, vegetative planting material and by aphids. It's one of the major concerns for raising healthy seed potato. Under field conditions, window available for monitoring viral incidence, monitoring of viruliferous aphid etc., is minimal. Hence, draw special attention for developing rapid and sensitive diagnostic assay to taking-up timely management practices. Hence, we aimed to develop LAMP assay for specific detection of PVA. Total RNA isolation and RT-PCR based confirmation of PVA from pure culture and from experimentally prepared viruliferous aphids was conducted. The LAMP assay was optimized and confirmed its specificity and sensitivity. Extended for detection in tubers and in single aphids and validated by running across field collected potato and aphid samples. A RT-LAMP assay was optimized where, 61 °C for 60 min gave sharp ladder like amplification and total volume of the reaction mixture was successfully reduced to 10 µl. Assay was visualized by naked eyes by adding SYBR gold nucleic acid stain. It was found to be highly specific as there was no cross reaction with other viruses infecting potato. Sensitivity of the assay was equivalent to RT-PCR, were it was able to detect the virus at 10<sup>-2</sup> dilution. The assay can be successfully applied for detection of PVA in potato tubers. Squash print RT-LAMP assay was developed and found suitable to determine viruliferous nature of aphids. Finally, the optimized RT-LAMP and SP-RT-LAMP assays were successfully validated by running across potato samples and aphid vectors collected randomly from fields respectively. Developed RT-LAMP and SP-RT-LAMP assay for specific and rapid detection of PVA in potato plants and in single aphids respectively which can assist in raising healthy seed potato.

### A comprehensive study on resistance to tomato leaf curl virus against various combinations of Ty genes in tomato

Pardhasaradhi P<sup>1,2</sup>, Rakesh Kumar<sup>\*1</sup>, Sairam P<sup>1</sup>, Ramachandran E<sup>1</sup>, Santanu Acharya<sup>1</sup> and Rajashaker P<sup>2</sup>

Email ID for Correspondence: pardhasaradhi@jkagri.com  
JK Agri Genetics Ltd, Hyderabad; <sup>2</sup> JNTU, Hyderabad

Tomato Leaf Curl Virus (ToLCV), a begomovirus, is a major constraint for the production of tomato cause up to 100% crop loss. To overcome these problems, researchers identified Ty-genes from wild tomato species. In this study, we introgressed these genes into

commercial tomato lines and tested its efficacy against begomoviruses. ToLCV symptomatic samples were collected and isolated its genomic DNA. These samples were tested in PCR with universal primer, Deng A/B and confirmed the presence of begomovirus. However to identify viral strain, designed a specific primer, (TL12F/13R for mono and TL14F/15R for bipartite) using DNA-A genome of ToLCV. The amplified PCR products were sequenced and submitted in NCBI database. The samples amplified with primer TL12F/13R similar with tomato leaf curl Karnataka virus (Acc.No KY311884) and with TL14F/15R similar with Tomato leaf curl New Delhi virus (Acc.No. MH269415). These strains were maintained in greenhouse on tomato plants through whiteflies transmission. These strains were tested on breeding population with different combinations of Ty genes (Ty-1, Ty-2, Ty-3, ty-5, Ty-1 + Ty-2, Ty2 + Ty3, Ty2 + ty-5, Ty1 + Ty3, ty-5 + Ty-6 and Ty-1 + Ty-2 + Ty-3) along with susceptible control. The healthy whiteflies incubated on tomato plant infested with either mono or bipartite for 24 h then released on experimental tomato plants. After infestation, the data was recorded on 20th and 45th day using 0-4 disease scoring. The results after 45 days against monopartite showed that plants having Ty-2 + Ty-3 and Ty-1 + Ty-2 + Ty-3 were tolerant. However, against bipartite begomovirus, plants having Ty2 + Ty3 and Ty-1 + Ty-2 + Ty-3 were moderately tolerant after 20 dpi and became susceptible after 45 dpi. The Ty genes interrogation into commercial tomato lines is useful to alleviate losses due to ToLCV.

### Understanding the role of glycine rich RNA binding proteins in viral pathogenesis

Email ID for Correspondence: ragunathan.sci@gmail.com

Glycine rich proteins (GRP) represent a group of proteins, which are characterized by the conserved semi-repetitive glycine rich motif at its C-terminal. GRPs are involved in growth, development and other multiple cellular processes. GRPs are also induced during various abiotic and biotic stresses but the mechanism during biotic stress especially during viral pathogenesis is poorly observed. Structurally, the Class IV GRPs essentially consist of two domains, a structured N-terminus containing an RNA Recognition Motif (RRM) and a flexible C-terminus with numerous glycine making the part intrinsically disordered. Although key reports are available about the role of Class IV GRPs in biotic stress responses, a role of these proteins in the life cycle of viruses has not yet been intensively studied.

NbGRP3 transcript level in viral infected samples quantified by northern hybridization/qRT-PCR. NbGRP3 cloned and purified to study the interaction with viral DNA using EMSA/fluorescence binding. NbGRP3 (a Class IVa GRP from *N. benthamiana*, a model organism for plant-virus interaction studies) shows elevated transcript levels upon infection of Geminiviruses, i.e., Tomato leaf curl New Delhi virus (ToLCNDV) along with non-cognate Radish leaf curl betasatellite (RaLCβ) in *N. benthamiana*. Similar kinds of observations were obtained with Tomato bushy stunt virus (TBSV), which is an RNA virus. Upon further investigation, it was found that NbGRP3 interacts with Satellite Conserved Region (SCR) of Radish Leaf curl virus Beta satellite. Up-regulation of transcript level of NbGRP3 upon viral infection (both DNA and RNA viruses) and their interaction with viral DNA stress on a role of NbGRP3 during viral pathogenesis. Yet further experiments need to be performed to understand the precise role of NbGRP3 in viral pathogenesis.



Indian Virological Society & PGIMER, Chandigarh

# INTERVIROCON 2018

27<sup>th</sup> International Conference of Virology

## Global Viral Epidemics: A Challenging Threat Best Poster Presentation Award

in Medical/Plant/Veterinary

This is to certify that **Dr./Mr./Ms. RAJNINA THAN DEVENDRAN**.....  
is the recipient of the Best Poster Presentation Award conferred at the  
**27<sup>th</sup> International Conference of Virology "INTERVIROCON 2018"**  
from 12<sup>th</sup>-14<sup>th</sup> November, 2018 organized by the Department of Virology  
Postgraduate Institute of Medical Education and Research, Chandigarh, India

**Dr. G.P. Rao**  
Secretary General  
Indian Virological Society

**Dr. RK Ratho**  
Chairperson  
INTERVIROCON 2018

**Dr. Mini P Singh**  
Organizing Secretary  
INTERVIROCON 2018

**Dr. Kapil Goyal**  
Joint Organizing Secretary  
INTERVIROCON 2018



With the support of the Erasmus Mundus programme of the European Union

Athens, August 2016

## ERASMUS MUNDUS Action 2 programme

### BRAVE

(Plant Virology in the new era - Breeding for resistance)

### Acceptance Letter

I hereby declare that **Ragunathan DEVENDRAN**, with scholarship application number **20160429.3.1.PhD.15**, has been granted an ERASMUS MUNDUS Action 2 scholarship through the BRAVE project, Grant Agreement Number 2013-2536/001-001.

BRAVE cohort: **3**  
Country: **India**  
Target Group: **TG1**  
Mobility level: **PhD**  
Seeking for: **Credit Seeking**<sup>1</sup>  
Host University: **MLU**  
Max. Duration: **10 months**  
Monthly subsistence allowance: **1500 €**  
Start date: **Fall 2016**

This scholarship letter from the Coordinating Institution of the BRAVE programme shall be presented to official authorities for the visa application, jointly with the letter of invitation from the host institution.

In addition to the monthly subsistence allowance the student will have full health and insurance coverage and one return ticket paid by the programme.

The ERASMUS MUNDUS Action 2 BRAVE programme is financed by the European Commission.

To be granted the scholarship, the scholarship holder must agree to all the terms pertaining to the grant of an ERASMUS MUNDUS Action 2 scholarship and of the BRAVE scholarship agreement.

<sup>1</sup> For credit seeking, it is the responsibility of the grantee to enroll in an Indian University for obtaining the degree from.

The BRAVE Coordinator  
Assistant Prof. Andreas Voloudakis  
Department of Crop Science  
Agricultural University of Athens, Greece

This project has been funded by the European Commission. This communication reflects the views only of the author and the Commission cannot be held responsible for any use that may be made of the information contained therein.





# Molecular genetic analysis and evolution of begomoviruses and betasatellites causing yellow mosaic disease of bhendi

R. Vinoth Kumar<sup>1</sup> · H. C. Prasanna<sup>2</sup> · A. K. Singh<sup>1,2</sup> · D. Ragunathan<sup>1</sup> · G. K. Garg<sup>3</sup> · S. Chakraborty<sup>1</sup>

Received: 23 June 2016 / Accepted: 22 November 2016 / Published online: 28 November 2016  
© Springer Science+Business Media New York 2016

**Abstract** In India, Bhendi yellow vein mosaic disease (BYVMD) is one of the most economically important diseases of bhendi/okra and is caused by a complex of monopartite begomovirus (*Bhendi yellow vein mosaic virus*—BYVMV) and betasatellite (*Bhendi yellow vein betasatellite*—BYVB). In this study, we have analyzed the role of possible evolutionary factors involved in the evolution of BYVMV and BYVB isolates. Evidence of inter-species and inter-strain recombination events was detected among the viral isolates, and majority of these recombinant isolates possess microsatellites in their genome. Recombination analysis suggests that cotton-infecting and bhendi-infecting begomoviruses probably share a recent common ancestor. In addition to genetic differentiation and gene flow, high degree of genetic variability was detected among the viral population. A strong purifying selection seems to be acting on the viral coding regions. The nucleotide substitution rate of V1 gene (for BYVMV) and  $\beta$ C1 gene (for BYVB) was estimated to be  $7.55 \times 10^{-4}$  and  $2.25 \times 10^{-3}$

nucleotide substitutions/site/year, respectively. The present study underlines that the evolution of BYVMD-associated viral components is driven by selection acting on the genetic variation generated by recombination and mutation.

**Keywords** *Geminiviridae* · *Begomovirus* · Betasatellite · Microsatellite · Nucleotide diversity · Recombination

## Introduction

Okra (local name: bhendi; *Abelmoschus esculentus*, Malvaceae) is an important vegetable crop cultivated in the Africa and Indian sub-continent. India is the major producer of this crop, and bhendi yellow vein mosaic disease (BYVMD) is one of the major factors limiting its productivity [1]. This disease was for the first time documented in 1924 [2], and subsequently in 1940, the causal agent was reported to be a virus [3]. In recent times, the incidence of this disease has been spreaded throughout the bhendi-growing areas of the country [1, 4]. BYVMD is caused by a group of circular ssDNA viruses belonging to the genus, *Begomovirus* (Family: *Geminiviridae*) [1, 4]. Members of the *Geminiviridae* family have been categorized into seven genera based on genome organization, host range, and type of insect vector [5]. The whitefly-transmitted begomoviruses have been further sub-divided into either bipartite (with DNA-A and DNA-B components) or monopartite (having a genome consisting of only a single-component homologous to the DNA-A of bipartite viruses). Monopartite begomoviruses reported from the ‘Old World’ are often associated with group of satellite and satellite-like molecules, termed as betasatellites and alphasatellites, respectively [6–12].

Edited by Seung-Kook Choi.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11262-016-1414-y) contains supplementary material, which is available to authorized users.

✉ S. Chakraborty  
supriyachakrasls@yahoo.com

<sup>1</sup> Molecular Virology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

<sup>2</sup> Present Address: Indian Institute of Vegetable Research, Varanasi, Uttar Pradesh, India

<sup>3</sup> Krishidhan Research Foundation Private Limited, Jalna, Maharashtra, India

Microsatellites or simple sequence repeats (SSRs) are the iterations of repeat units of short DNA sequence which is generally of 1–6 bases in length. A compound microsatellite (cSSR) is defined as the occurrence of two or more adjacently located SSRs separated by non-repeat bases. The replication slippage has been reported to be mainly involved in microsatellite instability [13]. The SSRs have been extensively studied at the genome level of eukaryotes, prokaryotes, and several viruses [14–18]. The microsatellites have been reported to be mainly involved in transcription and gene regulation, and its application includes linkage-disequilibrium mapping studies, paternity testing, forensics, and the inference of demographic processes [17, 19, 20].

A total of six monopartite and two bipartite begomovirus species are known to infect bhendi crop worldwide [1, 4, 21–24]. In India, bhendi-infecting monopartite begomoviruses were mostly associated along with a specific betasatellite, *Bhendi yellow vein betasatellite* (BYVB) [1, 25]. These betasatellites depend on the helper begomoviruses for their replication and encapsidation, while betasatellite-encoded  $\beta$ C1 protein interacts with coat protein and assists in systemic movement [26]. These BYVBs have been reported to be pathogenicity determinant and found responsible for the characteristic yellow vein mosaic symptoms [1, 27]. This begomovirus–betasatellite complex thus forms a severe disease complex leading to significant crop losses. Majority of the begomovirus isolates causing BYVMD in India belong to the predominant begomovirus species, *Bhendi yellow vein mosaic virus* (BYVMV), and a specific betasatellite (BYVB) [4, 25]. However, no detailed studies on the factors driving evolution of any particular begomovirus–betasatellite complex causing BYVMD have been attempted.

The present study was carried out to identify the factors involved in evolution of BYVMD-associated viral components. Majority of these begomovirus components were found recombinants and the results presume that these begomoviruses could share a recent ancestor with cotton-infecting begomoviruses. The viral coding regions were found to be under purifying selection. The impact of various driving forces in the evolution of these begomoviruses and betasatellites has been discussed.

## Materials and methods

### Cloning of viral genomic components

The typical begomovirus-like symptoms such as vein thickening, yellow mosaic, and stunting were found on the bhendi plants from Bhubaneswar, Dhanbad, Jalna, Pune, and Varanasi regions. From each region, at least five

samples were collected from the infected bhendi plants. The genomic DNA was extracted from these samples, followed by the enrichment of circular viral DNAs through rolling circle amplification using  $\phi$ 29 DNA polymerases [28]. The complete genome of begomoviruses was PCR amplified using a set of primers designed from highly conserved region (FP 5'-GCCGGCGTCTGAACTTCG-3' and RP 5'-GCCGGCGTACTTTCGA-3'). Full-length betasatellites were also PCR amplified using standard primers [29]. The amplified products were cloned into pJET1.2 vector (Thermo Fisher Scientific, Waltham, USA), and based on the restriction pattern, one of the representative cloned begomoviruses and betasatellites from each location were sequenced at DNA sequencing facility of the University of Delhi South Campus, New Delhi, India. The details of sample collection and associated viral components along with their accession numbers are mentioned in Table S1.

### Analysis of phylogeny, recombination, and microsatellite

In addition to the sequences reported in this study, 62 BYVMV and 36 BYVB sequences were retrieved from the GenBank and their details are provided in Table S2. The sequences were aligned using the MUSCLE algorithm [30], and the per cent pairwise sequence identities between sequences were calculated using SDT version 1.0 [31]. The phylogenetic trees were generated after aligning the sequences using MUSCLE algorithm [30] available in MEGA6 [32]. The presence and position of microsatellites in the viral genome were analyzed by Imperfect Microsatellite Extractor tool [33]. Recombination breakpoints and putative parental sequences were detected using RDP3 [34] that contains RDP [35], GENECONV [36], Chimaera [37], MaxChi [37, 38], Bootscan [39], Siscan [40], and 3Seq [41] methods. To exclude the possibility of false-positive recombination detection, the recombination events were considered only when they were detected by at least 4 out of 7 methods used. These analyses were performed as described [42], and the begomovirus and betasatellite sequences used for RDP analysis are listed in Table S3.

### Population demography analysis

To study genetic variability among the viral population, parameters in DnaSP v. 5.10 [43] used were total number of segregating sites ( $s$ ), total number of mutations ( $\eta$ ), average number of nucleotide differences between sequences ( $k$ ), nucleotide diversity ( $\pi$ ; the average pairwise differences at all sites within the sequences), Watterson's estimate of the population mutation rate based on the total

number of segregating sites ( $\theta - w$ ), and Watterson's estimate of the population mutation rate based on the total number of mutations ( $\theta - \eta$ ). Neutrality tests were performed using Fu and Li's  $D^*$  (difference between the number of singletons and total number of mutations), Fu and Li's  $F^*$  (difference between the number of singletons and average number of nucleotide differences between paired sequences), and Tajima's  $D$  (nucleotide diversity with the proportion of polymorphic sites) tests available in DnaSP v. 5.10 [43].

### Selection pressure analysis

The selection pressure acting on the virus coding regions was performed using single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), random-effects likelihood (REL), and fast unbiased Bayesian approximation (FUBAR) methods available in the DataMonkey server [44, 45]. The ratio of non-synonymous substitutions per non-synonymous site to synonymous substitutions per synonymous site ( $d_N/d_S$ ) was calculated using SLAC method. The ratio  $d_N/d_S > 1$ ,  $d_N/d_S < 1$ , and  $d_N/d_S = 1$  indicates positive (diversifying), negative (purifying), and neutral selection pressure, respectively. All these selection analyses were performed as described by Kumar et al. [9].

### Genetic differentiation and gene flow

Genetic differentiation ( $K_s^*$  and  $Z^*$ ) and gene flow ( $Nm$  and  $F_{ST}$ ) analyses were performed using DnaSP 5.0 [43].  $Nm$  and  $F_{ST}$  were estimated using the nucleotide sequence of complete genomes and the ORFs (V1 and  $\beta$ C1) to assess the level of gene flow between the populations [46, 47]. The level of gene flow between the sub-populations was estimated by  $Nm$  (the product of effective size of individual populations,  $N$  and the rate of migration among them,  $m$ ) and  $F_{ST}$  (the inter-population component of genetic variation or the standardized variance in allele frequencies across populations). Gene flow between the populations is considered to be frequent, when the absolute value of  $Nm$  is  $>1$  or  $F_{ST}$  is  $<0.33$ . Permutation-based statistical tests,  $K_s^*$  and  $Z^*$ , were calculated to study the genetic differentiation between populations [48, 49]. For  $K_s^*$  and  $Z^*$  values,  $p < 0.05$  was used as the criterion to consider that there is a genetic differentiation among the populations tested.

### Estimation of nucleotide substitution rates

The V1 (for BYVMV) and  $\beta$ C1 (for BYVB) genes (excluding inter-species recombinants) were used for the estimation of nucleotide substitution rates using BEAST v.1.7 [50]. Both strict and relaxed molecular clocks were

utilized along with Bayesian skyline plot as demographic model for each dataset as described for *Chilli leaf curl virus* [9], *East African cassava mosaic virus* [51], and *Tomato yellow leaf curl virus* [52]. The Bayes factors ( $\log_{10}BF$ ) on the tree likelihoods were used to select the best-supported substitution (HKY vs. GTR) and molecular clock model (strict vs. relaxed). Markov Chain Monte Carlo analyses were run for 20–40 million steps with first 10% steps as burn-in values to achieve the effective sample size of at least 100. Each BEAST analysis was analyzed using Tracer v.1.5, and the nucleotide substitution rates were estimated as described [9, 51, 52].

### Statistical analysis

To identify the correlation between the association of recombination events and cSSR motifs,  $\chi^2$  contingency test was performed as described using GraphPad Prism version 5 [42].

## Results and discussion

### Sequence analysis of the cloned viral genomes

BYVMD is one of the economically important diseases which have been prevalent in India for almost a century [2–4]. A specific BYVMV–BYVB complex has been demonstrated to cause BYVMD in India [1]. In addition, an isolate of cotton-infecting begomovirus has been reported to be associated with this disease [23]. In the present study, BYVMD-infected plant tissues were collected from five different bhendi-growing regions in India, and five complete sequences (each of DNA- $\alpha$  and DNA- $\beta$ ) were cloned. The present begomovirus sequences from Jalna (KF471064), Dhanbad (HM590503), Pune (KF471062), and Varanasi (HM590505) samples share maximum nucleotide identities between 99.1 and 99.2% with an isolate of *Bhendi yellow vein mosaic virus* (BYVMV; EU589392). A sequence from Bhubaneswar region (KF471063) has the highest sequence identity of 99.3% with a “Haryana” strain of BYVMV (FN645923) and  $<96.6\%$  with other isolates of BYVMV. Therefore, according to begomovirus species demarcation threshold [11], they should be regarded as the isolates of BYVMV and BYVMV-Har, respectively (Table S1). However, all the betasatellites reported here (HM590504, HM590506, KF471034–KF471036) shared  $>95\%$  of nucleotide sequence identity with already known BYVBs. As per ICTV guidelines for betasatellite demarcation [53], these sequences belong to BYVB group (Table S1). In agreement with their sequence identity, these begomoviruses and betasatellites positioned themselves along with other

isolates of their respective strains/groups in the phylogenetic dendrogram (Fig. 1). Collectively, the previous reports [1, 4, 25] and our current analysis underscore that BYVMD in India is caused by a complex of monopartite begomovirus–betasatellite (BYVMV–BYVB).

### Evidence of recombination in the evolution of BYVMV and BYVB isolates

It is widely regarded that recombination is one of a driving force that contributes to the evolution of members of the *Geminiviridae* family [9, 36, 54–56]. Previous report using Split-tree and Structure analyses suggest probable involvement of recombination in their evolution [4]. Therefore, in order to detect recombination breakpoints and to identify putative parental sequences, we have analyzed a total of 67 BYVMV isolates of India. The results emphasize that 63 of 67 isolates were found to be recombinants; they altogether possess 21 unique recombination events (Fig. 1a; Table S4). Two different sets of recombination events were detected in the sequences of the begomoviruses cloned from Bhubaneswar (events 1 and 2), and Dhanbad, Jalna, Pune, and Varanasi samples (events 15 and 21). Inter-species and inter-strain recombination events were detected in 8 of 21 and 5 of 21 events, respectively. However, in 7 events, at least one of the parental sequences was likely belonging to unknown begomovirus species (Table S4). From the recombination analysis, ancestral progenitors of these begomoviruses were found belong to either bhendi (BYVMV)-infecting (17 of 20) or cotton (CLCuAV, CLCuBV, and CLCuKoV)-infecting (3 of 20) begomoviruses (Fig. 1a; Table S4). The recombination breakpoints were detected in the C-terminal Rep region (events 2, 4, 6, 11, and 21) of 22 isolates. Similarly, the recombination fragment in the sequences grouped with events 9, 13, 14 and 20 was detected in C2/C3 region; however, most of the other recombination events have been involved in the exchanges around V1/V2 regions (Fig. 1a). Results from our analysis indicate that V1/V2 and C1 regions were the recombination hotspot which is consistent with previous reports [9, 42].

Similar to helper begomoviruses (BYVMV), 40 of 43 betasatellites were found to be recombinants possessing 15 recombination events (Fig. 1b; Table S5). All the recombinant betasatellites have acquired recombination fragments belonging to the same betasatellite group, i.e., BYVB. Many of the recombinant fragments (5/15) were found to encompass  $\beta$ C1 ORF (including in the betasatellites identified in this study), and some of them (8/15) were located around SCR (Fig. 1b). It is evident from our analysis that recombination has played a significant role in the diversification of BYVMV isolates and its betasatellite

molecules. The nature of viral components at the time of BYVMD incidence remains unclear; however, on the basis of the information generated here, we therefore presume that cotton-infecting begomoviruses might have recombined with an unidentified ancestor species leading to the emergence of bhendi-infecting begomoviruses or vice versa. This recombinant nature of these viral populations could help them to acquire higher evolutionary potential to widen their host range.

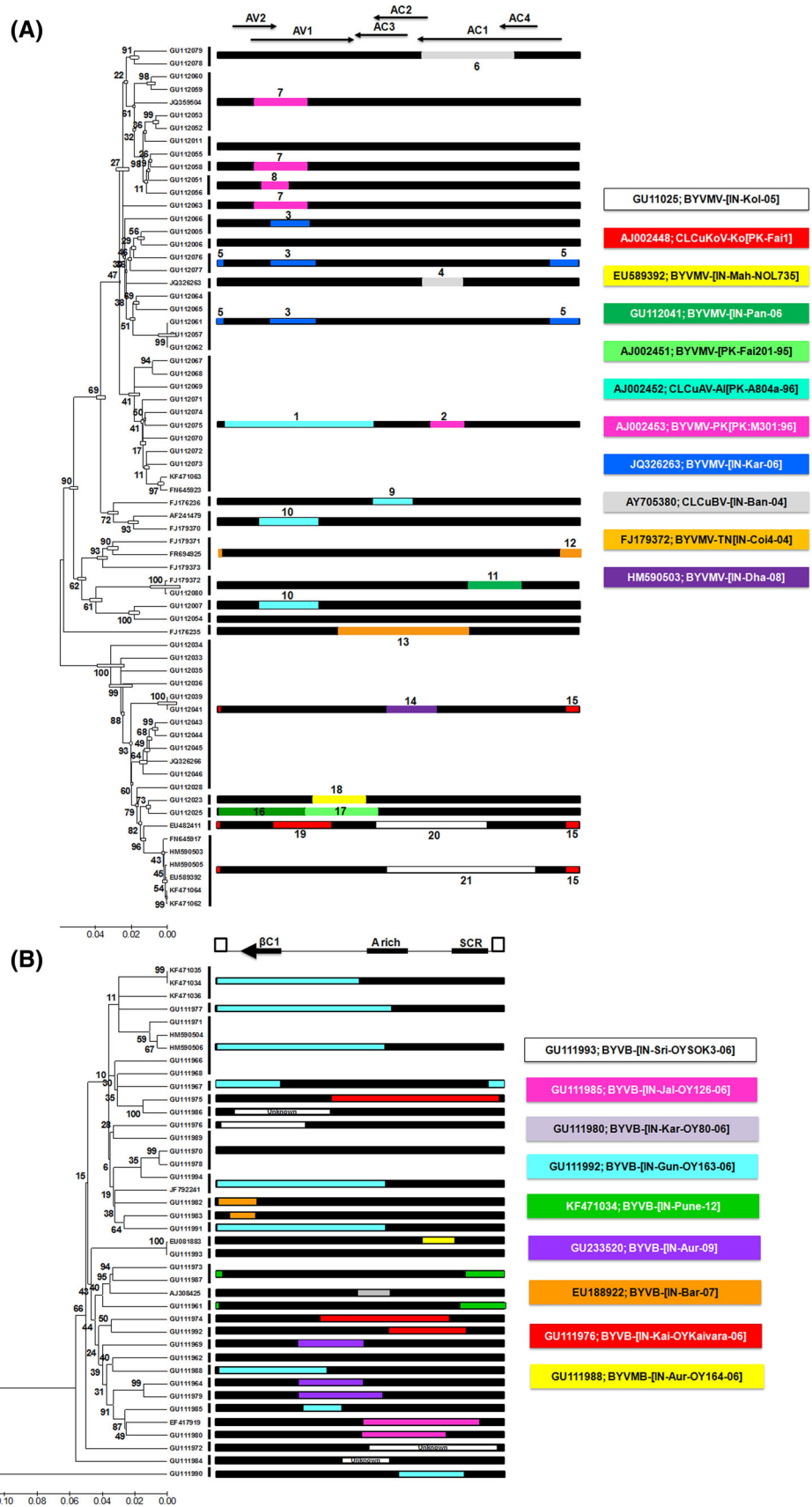
### Association between cSSR motifs and recombination events

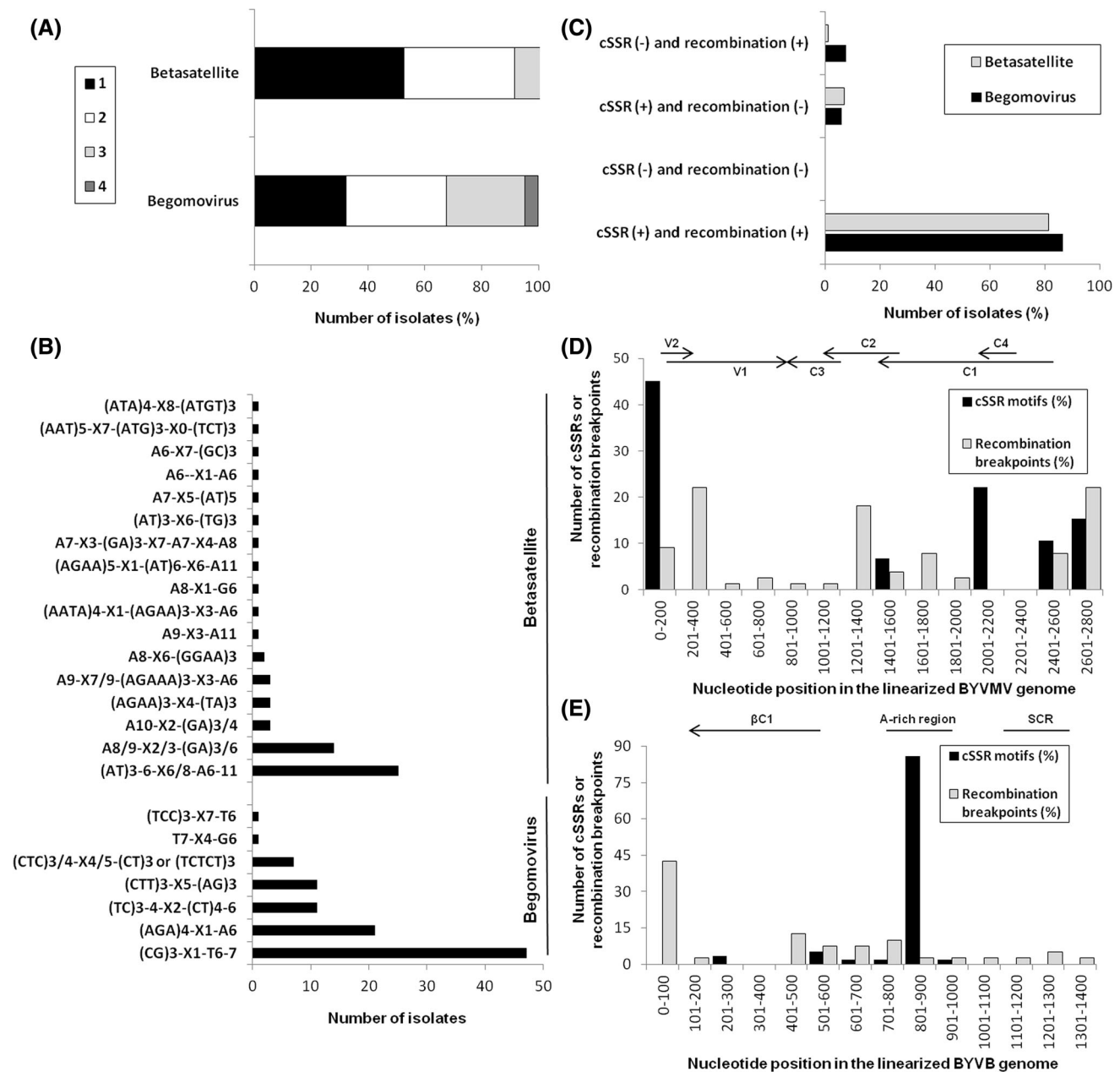
In recent times, the distribution of microsatellites among the members of various virus families have been reported [15, 18, 42]. Consequently, we have attempted to investigate the presence of cSSRs in BYVMV and BYVB isolates. The cSSRs were noticed in 62 of 67 begomoviruses, and 38 of 43 betasatellites analyzed. The presence of cSSRs in viral genome suggests that 32.2% of BYVMV isolates possess a single cSSR, while 2 or 3 cSSRs were detected in the remaining isolates (Fig. 2a). Similarly, one cSSR was present in 52.7% of betasatellites and the remaining sequences possess at least two cSSRs (Fig. 2a). In addition, the presence of frequently occurring cSSR motifs in viral sequences was analyzed. Two types of cSSRs, (CG)<sub>3</sub>-X<sub>1</sub>-T<sub>6-7</sub> and (AGA)<sub>4</sub>-X<sub>1</sub>-A<sub>6</sub>, were found to prevail in begomoviruses, and (AT)<sub>3-6</sub>-X<sub>6/8</sub>-A<sub>6-11</sub> and A<sub>8/9</sub>-X<sub>2/3</sub>-(GA)<sub>3/6</sub> were the most abundant cSSRs detected in betasatellites (Fig. 2b).

Recently, we have described the existence of a possible link between recombination events and the presence of cSSRs among begomoviruses [42]. To further authenticate this relationship, we have analyzed the association of cSSRs with recombination events detected among these viral genomes. Our results indicate that both the cSSR and recombination event were noticed in 86.5% of begomoviruses and 81.3% of betasatellites. Nevertheless, no cSSR motif was found in the non-recombinant sequences (Fig. 2c). In addition, the statistical correlation between occurrence of cSSR and recombination events among isolates of BYVMV and BYVB was tested using  $\chi^2$  test. After performing recombination analysis, parameter '1' was assigned for the presence of cSSR or recombination event and '0' for the absence of cSSR or recombination event as described [42]. The *p* value for the correlation of cSSR and recombination event detected in BYVMD-associated begomoviruses and betasatellites was found to be <0.001, which further underscores a significant correlation between the presence of cSSRs and recombination events. Among the BYVMV and BYVB isolates, at least one cSSR motif was found with each recombination events.



**Fig. 1** Phylogenetic relationship and recombination events of BYVMD-associated begomoviruses (a) and betasatellites (b) distributed in India. The bootstrap values of 1000 pseudo-replicates are placed at each node of the trees. The schematic representation depicting the detected recombination breakpoint(s) in each of the sequences along with recombination event number is provided at the right side of phylogenetic tree. The details of recombination events numbered here are listed in Tables S4 and S5, respectively. For comparison, the linearized viral genomic components are also provided at the top of each tree. The abbreviated names include *BYVB* Bendi yellow vein betasatellite, *BYVMV* Bendi yellow vein mosaic virus, *BYVMV-PK* Bendi yellow vein mosaic virus—Pakistan, *BYVMV-TN* Bendi yellow vein mosaic virus—Tamil Nadu, *CLCuBV* Cotton leaf curl Bangalore virus, *CLCuAIV-AI* Cotton leaf curl Alabad virus—Alabad, and *CLCuKoV-Ko* Cotton leaf curl Kokhran virus—Kokhran





**Fig. 2** Association between microsatellites and recombination events detected in Indian BYVMD-associated begomoviruses and betasatellites. **a** Distribution of the total number of cSSRs detected in each of the begomoviruses and betasatellites. The digits (1–4) denote the number of microsatellites present in the given sequence. **b** The frequency of cSSR motifs noticed among the viral genomic components. **c** The correlation between the cSSRs and recombination events

### Physical proximity of cSSRs and recombination breakpoints

To identify the location of compound microsatellites and recombination breakpoints, they were mapped onto the linearized viral genome. The result indicated that recombination breakpoints were distributed throughout the viral

genome, whereas cSSRs were mostly confined to intergenic region and N-terminal Rep of begomoviruses, and A-rich regions of betasatellites (Fig. 2d, e). Moreover, the physical proximity of cSSR motifs and recombination breakpoints were analyzed by calculating the distance (in bases) between end of each cSSR and the closest recombination breakpoint in a given sequence which possess both

cSSR and recombination events as described [42]. The results emphasize that majority of cSSRs (33 of 59 begomoviruses and 20 of 36 betasatellites) are located within 100 nucleotides from the recombination breakpoints (Fig. 3). Altogether, this result strengthens our previous assumption of a relationship between the compound microsatellites and recombination events in the begomovirus genomic components. In congruence with studies from prokaryotes and eukaryotes [19, 42, 57, 58], the speculative reason for this association is that these microsatellites may act as a signal for the recombination machinery to act or these repeat sequences somehow trigger crossing-over which enables in the exchange of DNA fragments. However, these assumptions need support from experimental evidences.

### Population demography analysis

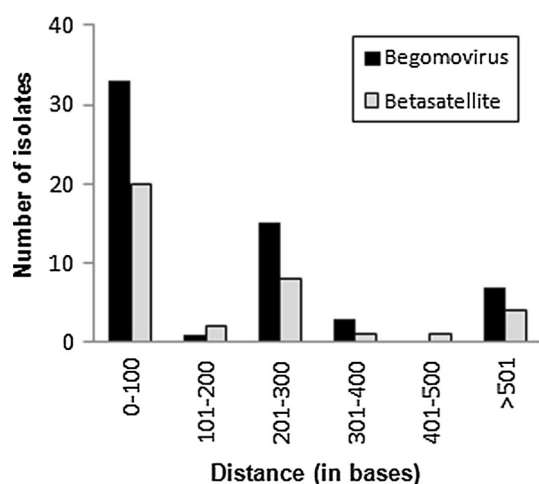
Recombination and population genetic structure are significant aspects in influencing virus evolution [9, 51, 52, 59]; however, the report on genetic variability among individual begomovirus species infecting cultivated crops in India is limited. Therefore, the genetic structure among the isolates of BYVMV and BYVB was evaluated. High genetic variability was estimated in the begomovirus ( $\pi = 0.08$ ) and betasatellite ( $\pi = 0.09$ ) isolates (Table 1), and the values were found to be several folds higher than the other begomoviruses reported [59, 60]. In addition, the values of neutrality tests (Tajima's  $D$ ; Fu and Li's  $D^*$ , and Fu and Li's  $F^*$ ) were detected negative and statistically significant. Moreover, we have also established the genetic variability of all ORFs encoded by these begomovirus–betasatellite complexes (Table 1). The betasatellite-

encoded  $\beta C1$  was found to be less variable, whereas the nucleotide diversity of V1, C1, and C4 ORFs were found to be higher (Table 1). Recombination analysis also suggests that the regions around C1 and V1/V2 ORFs were indeed recombination hotspots (Fig. 1a). In agreement with the previous reports [59, 60], the possible scenario for this higher nucleotide diversity observed could be the result of higher frequency of recombination breakpoints exchanging these coding regions.

To identify any region-specific population structure among the isolates distributed throughout the country, we have divided the virus population into 3 regions: northern, central, and southern sub-populations. The results indicate that isolates belonging to the central region were less variable than the isolates identified from either northern or southern regions (Table 1). The genetic differentiation result implies that a significant genetic differentiation ( $p < 0.001$ ) exists among the sub-populations within north-central and central-south regions (Table 2). Similarly, the possibility of occurrence of gene flow among this sub-population was tested, and the results suggest that gene flow occurred only among north and south sub-populations (Table 2). These results presume that low genetic differentiation noticed among north–south sub-populations could be the effect of frequent gene flow between them. However, we cannot rule out the possibility of less number of isolates available for the analysis as a reason for low genetic differentiation among the central India sub-population. However, among betasatellite sub-populations, no significant genetic differentiation was observed, nevertheless a strong genetic flow was evidenced (Table 2).

### Diversifying/positive selection

The type of natural selection acting on the amino acids of viral proteins (coat protein, Rep, and  $\beta C1$ ) was analyzed using various methods available at DataMonkey server (Table 3). The nature of selection (diversifying or purifying) acting on a particular amino acid was considered only when detected by at least 3 out of 4 methods tested. Amino acids 31 and 146 in the coat protein were found to be selected positively (Table 3). Interestingly, amino acid 146 in the coat protein has been reported to be involved in vector transmission; even a single amino acid change around this region may alter the efficiency of insect transmission which subsequently enables them to invade alternate host(s). Further experimental evidence is required to prove this assumption of expanding host adaptation. Evidence of positive selection was found in 9 and 3 amino acid residues of the Rep and  $\beta C1$  proteins (Table 3). These residues in the Rep are mostly residing in the N-terminal



**Fig. 3** Physical proximity of the compound microsatellite (cSSRs) and recombination breakpoints. Distance between the end of each cSSR and the closest recombination breakpoint in viral sequences which possess both cSSR and recombination event is mentioned

**Table 1** Neutrality tests, haplotype, and nucleotide diversity of begomoviruses and betasatellites associated with BYVMD in India

Sequences analyzed	Viral component	Regions/genes	Genetic variability <sup>a</sup>						Neutrality tests		
			Length	<i>s</i>	$\eta$	$\pi$	<i>k</i>	Hd	<i>D</i>	<i>D</i> *	<i>F</i> *
Complete genome sequence	BYVMV	North ( <i>n</i> = 18)	2619	781	943	0.089	233.40	1.000 ± 0.010	-0.766	-0.814	-0.908
		Central ( <i>n</i> = 9)	2727	364	390	0.047	128.39	1.000 ± 0.063	-0.803	-0.939	-0.939
		South ( <i>n</i> = 40)	2692	841	1010	0.070	190.83	0.996 ± 0.007	-0.737	-0.971	-1.057
		All ( <i>n</i> = 67)	2587	1130	1473	0.080	208.02	1.000 ± 0.020	-1.101	-2.750*	-2.477*
	BYVB	North ( <i>n</i> = 11)	1085	505	603	0.119	128.41	1.000 ± 0.034	-1.689	-1.888	-2.010
		Central ( <i>n</i> = 7)	1313	228	242	0.060	79.29	0.952 ± 0.096	-1.159	-1.062	-1.202
		South ( <i>n</i> = 24)	1218	624	814	0.090	110.07	1.000 ± 0.013	-2.044*	-2.298	-2.605
		All ( <i>n</i> = 42)	1012	715	1078	0.098	99.65	0.998 ± 0.006	-2.234 <sup>#</sup>	-2.744*	-3.053**
		Complete sequence of ORFs	BYVMV	V1 ( <i>n</i> = 67)	768	236	278	0.038	27.99	0.995 ± 0.002	-1.821*
		V2 ( <i>n</i> = 67)	366	388	107	0.031	11.40	0.979 ± 0.011	-1.691	-3.690 <sup>#</sup>	-3.475 <sup>#</sup>
		C1 ( <i>n</i> = 64)	1092	487	637	0.123	123.59	0.997 ± 0.004	-0.293	-2.003	-1.576
		C2 ( <i>n</i> = 67)	450	137	165	0.048	20.28	0.990 ± 0.006	-1.437	-3.438 <sup>#</sup>	-3.162*
		C3 ( <i>n</i> = 67)	405	105	122	0.038	15.25	0.987 ± 0.007	-1.393	-4.043 <sup>#</sup>	-3.585 <sup>#</sup>
		C4 ( <i>n</i> = 66)	306	222	100	0.139	30.78	0.996 ± 0.013	0.563	-0.856	-0.352
	BYVB	βC1 ( <i>n</i> = 42)	232	136	176	0.065	15.02	0.960 ± 0.020	-2.315 <sup>#</sup>	-3.818 <sup>#</sup>	-3.902 <sup>#</sup>

\*  $p < 0.05$ , <sup>#</sup>  $p < 0.02$ <sup>a</sup> Details of the abbreviations used are described in Methods**Table 2** Genetic differentiation and gene flow between the viral sub-population identified from various regions in India

Genome	Population	Genetic differentiation		Gene flow	
		<i>K</i> <sub>S</sub> *	<i>Z</i> *	<i>N</i> <sub>m</sub>	<i>F</i> <sub>ST</sub>
BYVMV	North-Central	5.222 <sup>#</sup> (3.491)	4.341 <sup>#</sup> (4.702)	0.41 (-51.10)	0.377 (-0.004)
	Central-South	5.114 <sup>#</sup> (3.098 <sup>#</sup> )	5.769 <sup>#</sup> (5.986 <sup>#</sup> )	0.36 (5.83)	0.408 (0.041)
	North-South	5.250 (3.178)	6.347 (6.389)	8.06 (9.25)	0.030 (0.026)
BYVB	North-Central	4.345 (2.418)	4.176 (4.105)	348.51 (69.00)	0.001 (0.004)
	Central-South	4.383 (2.778)	5.082 (5.109)	21.15 (19.08)	1.012 (0.013)
	North-South	4.467 (2.501)	5.386 (5.427)	9.20 (3.98)	0.026 (0.059)

The values in the parentheses were calculated using V1 ORF of BYVMV and βC1 ORF of BYVB

<sup>#</sup>  $p < 0.001$ 

region comprising of various conserved motifs essential for virus replication [63]. Amino acids 37 to 74 of *Tomato leaf curl China betasatellite*-βC1 protein have been found to be essential for nuclear localization and RNA silencing suppression [64]. Recently, sites under positive selection have been shown to modulate silencing suppression activity of a viral protein [65]. In this context, positively selected amino acids 44 and 61 of βC1 may have an effect on its biological function. Collectively, our data imply that these positively

selected residues may have an impact on the viral DNA replication and/or vector transmissibility.

### Purifying/negative selection

The number of negatively selected sites was found to be large in comparison to positively selected sites in the viral coding regions (Table 3). Our findings underline that the Rep protein possesses large number of sites that evolved

**Table 3** List of positively and negatively selected amino acid residues encoded by the isolates of BYVMV and BYVB

Viral genes	Methods used for selection pressure analysis <sup>c</sup>					Hierarchical Bayesian method FUBAR <sup>b</sup>	Number of positively selected sites	Number of negatively selected sites <sup>d</sup>
	Codon-based maximum likelihood methods							
	$d_N/d_S^a$	SLAC	FEL	REL				
V1	0.113	<b>31, 146</b>	<b>31, 146</b>		<b>31, 111, 146, 171</b>	<b>146</b>	04 (2)	93/101/48/114
C1	0.250	<b>11, 39, 73, 85, 349, 362, 363</b>	<b>10, 11, 39, 43, 48, 53, 55, 68, 73, 77, 80, 85, 112, 145, 148, 175, 182, 214, 320, 328, 334, 337, 338, 348, 349, 351, 360, 362, 363</b>		<b>10, 11, 39, 53, 73, 85, 182, 214, 337, 338, 349, 360, 362, 363</b>	<b>10, 11, 39, 53, 73, 85, 349, 360, 363</b>	29 (9)	118/162/147/152
$\beta$ C1	0.596	<b>24, 41, 64</b>	<b>6, 24, 31, 41, 52, 64, 128, 145</b>		<b>12, 21, 22, 23, 24, 26, 30, 41, 64, 69</b>	<b>12, 24, 41, 64</b>	10 (3)	23/24/2/4

<sup>a</sup>  $d_N/d_S$  ratio was calculated using SLAC method

<sup>b</sup> Pervasive diversifying selection; The FUBAR method is used to identify individual codon sites evolving under either diversifying or negative selection throughout the entire evolutionary history

<sup>c</sup> The positively selected amino acid residues detected by at least three of four methods are given by bold letters, and the total number of sites is provided in the parentheses of next column

<sup>d</sup> The number of negatively selected sites as detected by methods, SLAC, FEL, REL, and FUBAR, respectively

**Table 4** The rate of nucleotide substitutions for the V1 and  $\beta$ C1 genes encoded by various begomoviruses and betasatellites

Viral components	Viral genes	Number of sequences used	Chain length (millions)	Mean substitution rate (nucleotides/site/year)	HPD substitution rate (nucleotides/site/year)	References
DNA-A	BYVMV-V1	60	20	$7.55 \times 10^{-4}$	$1.76 \times 10^{-4}$ to $1.37 \times 10^{-3}$	In this study
	ChiLCV-V1	52	20	$2.60 \times 10^{-3}$	–	[9]
	EACMV-AV1	71	15	$1.37 \times 10^{-3}$	$4.69 \times 10^{-4}$ to $2.32 \times 10^{-3}$	[51]
	TYLCV-V1	54	20	$4.63 \times 10^{-4}$	$6.62 \times 10^{-5}$ to $8.85 \times 10^{-4}$	[52]
DNA-B	EACMV-BV1	46	20	$2.77 \times 10^{-4}$	$9.18 \times 10^{-6}$ to $6.46 \times 10^{-4}$	[52]
DNA- $\beta$	BYVB- $\beta$ C1	41	20	$2.25 \times 10^{-4}$	$6.18 \times 10^{-8}$ to $6.96 \times 10^{-4}$	In this study
	ChiLCB- $\beta$ C1	25	40	$2.57 \times 10^{-4}$	–	[9]
	ToLCBDB- $\beta$ C1	31	40	$5.22 \times 10^{-4}$	–	[9]

under negative selection followed by the coat protein. Further,  $d_N/d_S$  ratio of the coat protein and the Rep protein were calculated as 0.113 and 0.250, respectively, indicating the strongest purifying selection pressure to prevent any structural or functional constraints for amino acid changes. Nevertheless,  $\beta$ C1 ORF was apparently evolving under weak negative ( $d_N/d_S$  ratio of 0.596) selection pressure (Table 3). Consistent with earlier reports [9, 52, 56, 60],

our observations emphasize the role of negative selection pressure acting on the begomovirus-encoded ORFs.

**Estimation of nucleotide substitution rates**

In order to assess the selection pressure acting on the V1 and  $\beta$ C1 genes, we have also estimated the nucleotide substitution rate. The overall rate of evolution of the V1



and  $\beta$ C1 genes has been estimated as  $7.55 \times 10^{-4}$  (95% HPD  $1.76 \times 10^{-4}$  to  $1.37 \times 10^{-3}$ ) and  $2.25 \times 10^{-4}$  (95% HPD  $6.18 \times 10^{-8}$  to  $6.96 \times 10^{-4}$ ) nucleotide substitutions/site/year, respectively (Table 4). This high nucleotide substitution rates could be due to the effect of rapid mutations. In addition, the translesion synthesis polymerases have been shown to perform geminivirus DNA replication [66]. In this context, the involvement of error-prone DNA polymerases in virus replication is most likely resulting in the higher rate of nucleotide substitution of begomovirus-encoded ORFs [66].

In summary, the isolates of monopartite begomovirus–betasatellite complex causing BYVMD in India were found to be highly recombinant. The results of recombination analysis further provided evidence for the existence of an association between recombination events and microsatellites. The evolution of BYVMD-associated viral components is driven by selection acting on the genetic variation generated by recombination and nucleotide diversity. The finding of this study on evolutionary factors that shape the emergence and spread of BYVMD could help in the development and implementation of effective biological and/or management strategies to curb this economically important disease.

**Acknowledgements** SC gratefully acknowledges the Department of Biotechnology, Government of India, for the financial support through a Research Grant (BT/SBIRI/571/23-B11/2009).

**Author's contributions** RV and SC conceived the study, analyzed data, and drafted the article. RV, DR, and AKS performed the experiments. SC, HCP, and GKG contributed reagents/materials/analysis tools. All the authors had read and approved the article.

#### Compliance of ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

1. J. Jose, R. Usha, *Virology* **305**, 310–317 (2003)
2. G.S. Kulkarni, *Poona Agric. Coll. Mag.* **16**, 6–12 (1924)
3. B.N. Uppal, P.M. Varma, S.P. Kapoor, *Curr. Sci.* **9**, 2278 (1940)
4. V. Venkataravanappa, H.C. Prasanna, C.N.L. Reddy, M.K. Reddy, *Plant. Pathol.* **64**, 508–518 (2014)
5. A. Varsani, J. Navas-Castillo, E. Moriones, C. Hernandez-Zepeda, A. Idris, J.K. Brown, F.M. Zerbini, D.P. Martin, *Arch. Virol.* **159**, 2193–2203 (2014)
6. B. George, R.V. Kumar, S. Chakraborty, *Virus Genes* **48**, 397–401 (2014)
7. S. Mansoor, S.H. Khan, A. Bashir, M. Saeed, Y. Zafar, K.A. Malik, R.W. Briddon, J. Stanley, P.G. Markham, *Virology* **259**, 190–199 (1999)
8. M.S. Nawaz-ul-Rehman, C.M. Fauquet, *FEBS Lett.* **583**, 1825–1832 (2009)
9. R.V. Kumar, A.K. Singh, A.K. Singh, T. Yadav, S. Basu, N. Kuswaha, B. Chattopadhyay, S. Chakraborty, *J. Gen. Virol.* **96**, 3157–3172 (2015)
10. P. Ranjan, A.K. Singh, R.V. Kumar, S. Basu, S. Chakraborty, *Virus Genes* **48**, 334–342 (2014)
11. J.K. Brown, F.M. Zerbini, J. Navas-Castillo, E. Moriones, R. Ramos-Sobrinho, J.C.F. Silva, E. Fiallo-Olive, R.W. Briddon, C. Hernandez-Zepeda, A. Idris, V.G. Malathi, D.P. Martin, R. Rivera-Bustamante, S. Ueda, A. Varsani, *Arch. Virol.* **160**, 1593–1619 (2015)
12. J.K. Brown, C.M. Fauquet, R.W. Briddon, F.M. Zerbini, E. Moriones, J. Navas-Castillo, in *Virus Taxonomy 9th Report of the International Committee on Taxonomy of Viruses*, ed. by A.M.Q. King, M.J. Adams, E.B. Carstens, E.J. Lefkowitz (Elsevier Academic Press, London, 2012), pp. 351–373
13. G. Levinson, G.A. Gutman, *Nucleic Acids Res.* **15**, 5323–5338 (1987)
14. B. Valota, J. Knappa, G. Umhang, F. Grenouilleta, L. Millona, *Infect. Genet. Evol.* **32**, 338–341 (2015)
15. C.M. Alam, A.K. Singh, C. Sharfuddin, S. Ali, *Infect. Genet. Evol.* **21**, 287–294 (2014)
16. C.P. Hong, Z.Y. Piao, T.W. Kang, S.J. Kwon, J.S. Kim, T.J. Yang, B.S. Park, Y.P. Lim, *Mol. Cell.* **23**, 349–356 (2007)
17. J. Mrazek, X. Guo, A. Shah, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8472–8477 (2007)
18. C.M. Alam, B. George, C. Sharfuddin, S.K. Jain, S. Chakraborty, *Gene* **521**(2), 238–244 (2013)
19. Y. Kashi, D.G. King, *Trends Genet.* **22**, 253–259 (2006)
20. H. Ellegren, *Nat. Rev. Genet.* **5**, 435–445 (2004)
21. C. Hernandez-Zepeda, T. Isakeit, A. Scott Jr., J.K. Brown, *Plant Dis.* **94**, 924 (2010)
22. T. Kon, M.R. Rojas, I.K. Abdourhamane, R.L. Gilbertson, *J. Gen. Virol.* **90**, 1001–1013 (2009)
23. V. Venkataravanappa, C.N.L. Reddy, A. Devaraju, S. Jalali, M.K. Reddy, *Indian J. Virol.* **24**, 188–198 (2013)
24. V. Venkataravanappa, C.N.L. Reddy, S. Jalali, M.K. Reddy, *Eur. J. Plant Pathol.* **136**, 811–822 (2013)
25. V. Venkataravanappa, C.N.L. Reddy, P. Swaranalatha, S. Jalali, R.W. Briddon, M.K. Reddy, *Virology J.* **8**, 555 (2011)
26. P. Kumar, R. Usha, A. Zrachya, Y. Levy, H. Spanov, Y. Gafni, *Virus Res.* **122**, 127–136 (2006)
27. P. Gopal, P.P. Kumar, B. Sinilal, J. Jose, A.K. Yadunandama, R. Usha, *Virus Res.* **123**, 9–18 (2007)
28. S.L. Dellaporta, J. Wood, J.B. Hicks, *Plant Mol. Biol. Report.* **1**, 19–21 (1983)
29. R.W. Briddon, S.E. Bull, S. Mansoor, I. Amin, P.G. Markham, *Mol. Biotechnol.* **20**, 315–318 (2002)
30. R.C. Edgar, *BMC Bioinform.* **5**, 1–19 (2004)
31. B.M. Muhire, A. Varsani, D.P. Martin, *PLoS ONE* **9**, e108277 (2014)
32. K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, *Mol. Biol. Evol.* **28**, 2731–2739 (2013)
33. S.B. Mudunuri, H.A. Nagarajaram, *Bioinformatics* **23**, 1181–1187 (2007)
34. D.P. Martin, P. Lemey, M. Lott, M. Vincent, D. Posada, P. Lefevre, *Bioinformatics* **26**, 2462–2463 (2010)
35. D.P. Martin, E. Rybicki, *Bioinformatics* **16**, 562–563 (2000)
36. M. Padidam, S. Sawyer, C.M. Fauquet, *Virology* **265**, 218–225 (1999)
37. D. Posada, K.A. Crandall, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3757–3762 (2001)
38. J.M. Smith, *J. Mol. Evol.* **34**, 126–129 (1992)
39. D.P. Martin, D. Posada, K.A. Crandall, C. Williamson, *AIDS Res. Hum. Retroviruses* **21**, 98–102 (2005)

40. M.J. Gibbs, J.S. Armstrong, A.J. Gibbs, *Bioinformatics* **16**, 573–582 (2000)
41. M.F. Boni, D. Posada, M.W. Feldman, *Genetics* **176**, 1035–1047 (2007)
42. B. George, C.M. Alam, R.V. Kumar, G. Prabu, S. Chakraborty, *Virology* **148**, 41–50 (2015)
43. P. Librado, J. Rozas, *Bioinformatics* **25**, 1451–1452 (2009)
44. W. Delport, A.F.Y. Poon, S.D.W. Frost, S.L.K. Pond, *Bioinformatics* **26**, 2455–2457 (2010)
45. S.L.K. Pond, S.D.W. Frost, *Mol. Biol. Evol.* **22**, 1208–1222 (2005)
46. S. Wright, *Ann. Eugen.* **15**, 323–354 (1951)
47. B.S. Weir, C.C. Cockerham, *Evolution* **38**, 1358–1370 (1984)
48. R.R. Hudson, *Genetics* **155**, 2011–2014 (2000)
49. R.R. Hudson, D.D. Boos, N.L. Kaplan, *Mol. Biol. Evol.* **9**, 138–151 (1992)
50. A.J. Drummond, M.A. Suchard, D. Xie, A. Rambaut, *Mol. Biol. Evol.* **29**, 1969–1973 (2012)
51. S. Duffy, E.C. Holmes, *J. Gen. Virol.* **90**, 1539–1547 (2009)
52. S. Duffy, E.C. Holmes, *J. Virol.* **82**, 957–965 (2008)
53. R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, F.M. Zerbini, X. Zhou, C.M. Fauquet, *Arch. Virol.* **153**, 763–781 (2008)
54. P. Lefeuve, J.M. Lett, A. Varsani, D.P. Martin, *J. Virol.* **83**, 2697–2707 (2009)
55. P. Lefeuve, E. Moriones, *Curr. Opin. Virol.* **10**, 14–19 (2015)
56. S. Kraberger, S.G. Kumari, A.A. Hamed, B. Gronenborn, J.E. Thomas, M. Sharman, G.W. Harkins, B.M. Muhire, D.P. Martin, A. Varsani, *Infect. Genet. Evol.* **29**, 203–215 (2015)
57. M. Brandstrom, A.T. Bagshaw, N.J. Gemmell, H. Ellegren, *Mol. Biol. Evol.* **25**, 2579–2587 (2008)
58. A.J. Jeffreys, J. Murray, R. Neumann, *Mol. Cell* **2**, 267–273 (1998)
59. R.S. Sobrinho, C.A.D. Xavier, H.M.B. Pereira, G.S.A. Lima, I.P. Assuncao, E.S.G. Mizubuti, S. Duffy, F.M. Zerbini, *J. Gen. Virol.* **95**, 2540–2552 (2014)
60. C.S. Rocha, G.P. Castillo-Urquiza, A.T.M. Lima, F.N. Silva, C.A.D. Xavier, B.T. Hora-Junior, J.E.A. Beserra-Junior, A.W.O. Malta, D.P. Martin, A. Varsani, P.A. Zerbini, E.S.G. Mizubuti, F.M. Zerbini, *J. Virol.* **87**, 5784–5799 (2013)
61. M. Hohnle, P. Hofer, I.D. Bedford, R.W. Briddon, P.G. Markham, T. Frischmuth, *Virology* **290**, 164–171 (2001)
62. A. Kheyr-Pour, K. Bananej, G.A. Dafalla, P. Caciagli, E. Noris, A. Ahoonmanesh, H. Lecoq, B. Gronenborn, *Phytopathology* **90**, 629–635 (2000)
63. L. Hanley-Bowdoin, E.R. Bejarano, D. Robertson, S. Mansoor, *Nat. Rev. Microbiol.* **11**, 777–788 (2013)
64. X. Yang, W. Guo, X. Ma, Q. An, X. Zhou, *Appl. Environ. Microbiol.* **77**, 3092–3101 (2011)
65. D. Sereme, S. Lacombe, M. Konate, M. Bangratz, A. Pinel-Galzi, A. Fargette, A.S. Traore, G. Konate, C. Brugidou, *J. Gen. Virol.* **95**, 213–218 (2014)
66. K.S. Richter, M. Gotz, S. Winter, H. Jeske, *Virology* **488**, 137–148 (2016)







# Overview of plant RNA dependent RNA polymerases in antiviral defense and gene silencing

Ved Prakash<sup>1</sup> · Ragunathan Devendran<sup>1</sup> · Supriya Chakraborty<sup>1</sup>

Received: 11 November 2017 / Accepted: 21 November 2017 / Published online: 22 December 2017  
© Indian Society for Plant Physiology 2017

**Abstract** RNA silencing is a mechanism by which plants not only protect themselves from viruses and viroids but also keep transposons and endogenous genes silenced epigenetically. Among a myriad of proteins involved in RNA silencing pathway, RNA dependent RNA polymerases (RDRs) are of utmost importance. RDR proteins are central in the amplification of small interfering RNAs (siRNAs) which are crucial for efficient silencing of the nucleic acids. These proteins convert single-stranded RNAs (ssRNAs) into double-stranded RNAs (dsRNAs). These siRNAs can silence the expression of genes locally or systematically. Understanding the function of RDR proteins can be useful in developing crops which would be resistant to biotic and abiotic agents using biotechnological tool. Thus, in this review, we shall focus on the evolution, function and regulation of RDR proteins in plants.

**Keywords** RNA dependent RNA polymerase · Gene silencing · Viruses · Salicylic acid · Systemic acquired resistance

## Introduction

With the ongoing scenario of climatic change, it has become imperative to strengthen the ‘fight-back’ potential of plants towards pathogens attack. Because of immense research done in the area of plant-pathogen interaction, now it is becoming feasible to develop pathogen-resistant

plants. Delivery or transfer of viral sequences into the plant genome is a tool for developing virus resistant plants. Coat-Protein-Mediated Resistance (CP-MR) was the first experimental evidence that came in support of the aforesaid hypothesis where the over expression of coat protein of *Tobacco Mosaic Virus* (TMV) in tobacco enabled plants to show delay in symptom development towards TMV (Powell et al. 1986). Further evidence came in 1993 when the transgenic tobacco expressing the Tobacco etch virus (TEV) coat protein showed resistance to TEV infection (Lindbo et al. 1993). Initially, the transgenic plant showed symptoms upon TEV infection but recovered after 3–5 weeks and the newly emerging leaves were TEV free. The resistance was not broad-spectrum rather specific to TEV. The mechanism involved is known as RNA silencing. There are many proteins effectors, for example, Dicer-like proteins (DCLs), Argonautes (AGOs), RNA dependent RNA polymerases (RDRs) which constitute the major components of the RNA silencing pathways. Among various effector proteins of RNA silencing pathways, RDRs play a crucial role in the amplification of small RNA. Plant genome contains multiple copies of RDR proteins with specialized functions which are involved in regulating development, physiology, abiotic and biotic stress responses (Borges and Martienssen 2015; Zvereva and Pooggin 2012) (Table 1).

RNA silencing can be triggered by double-stranded RNA precursors (either intra-molecular fold-back structures or inter-molecular dsRNA). These dsRNAs are cleaved by RNase III-type endoribonucleases (DICER) in animals (Bernstein et al. 2001) and DCL in plants (Carmell and Hannon 2004) into short RNA duplex of 21–24 nucleotide (nt) in length, which interact and load onto AGO proteins. These short 21–24 nt long RNA duplexes can be either siRNA or micro-RNA (miRNA) (Henderson

✉ Supriya Chakraborty  
supriyachakrasls@yahoo.com

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

**Table 1** Number of RDR homologs identified in various plant species

Plants	RDRs	References
<i>Arabidopsis thaliana</i>	6	Wassenegger and Krczal (2006)
<i>Solanum lycopersicum</i>	6	Bai et al. (2012)
<i>Zea mays</i>	5	Qian et al. (2011)
<i>Oryza sativa</i>	5	Kapoor et al. (2008)
<i>Brassica napus</i>	16	Cao et al. (2016)
<i>Setaria italica</i>	11	Yadav et al. (2015)
<i>Sorghum bicolor</i>	7	Liu et al. (2014)
<i>Cucumis sativus</i>	8	Gan et al. (2016)
<i>Glycine max</i>	7	Liu et al. (2014)
<i>Vitis vinifera</i>	5	Zhao et al. (2015)
<i>Coffea canephora</i>	8	Fernandes-Brum et al. (2017)

et al. 2006; Margis et al. 2006; Willmann et al. 2011). After loading onto AGO, one of the strand which is known as passenger strand will be degraded while the other strand which is referred as guide strand will bind to the complementary sequences in mRNA leading to cleavage of the transcript or translation inhibition. Genomic DNA can also be silenced by guide small RNA loaded on to AGO in a process known as RNA directed DNA methylation (RdDM) (Henderson et al. 2006; Margis et al. 2006; Simon and Meyers 2011).

In plants, DCL mediated cleavage produces primary siRNA which can be used by RDRs for amplification and production of secondary siRNAs either in primer-independent or primer-dependent manner (Devert et al. 2015). These secondary siRNAs can spread systematically in plants thereby amplifying the signal leading to the systemic spread of silencing.

Detection of the RDR activity in plants was first observed in the early 1970s. In 1971, Astier-Manificier and Cornuet first demonstrated the RDR activity in Chinese cabbage. Later, the RDR activity was shown in many plants such as in cauliflower, tobacco, tomato, cow pea and cucumber (Schiebel et al. 1998). However, the first biochemical and in vitro characterization of RDR was done in tomato in 1993 (Schiebel et al. 1993a, b). The current review aims to discuss our current understanding on the evolution, classification, function and regulation of RDRs in plants.

### Evolution of RDRs

Evolution of RDRs is yet a puzzle and very little is known in this area. RDRs might have been significantly important during the initial period of evolution after ancient RNA world transition. Before the discovery of viral RNA

dependent RNA polymerase (RdRP), it was believed that the replication of RNA viruses occurs inside the nucleus of the host cell by host-encoded DNA dependent RNA polymerase. The first report of RNA dependent RNA polymerase used by RNA virus to replicate their RNA genome was published in 1962 by Baltimore and Franklin (Baltimore and Franklin 1962). Three decades later, RDR from tomato was characterized which was the first RDR to be characterized from the eukaryotes (Schiebel et al. 1993a) followed by discovery of QDE-1 protein from a filamentous fungus *Neurospora crassa* which possesses plant RDR like activity and was found to be involved in PTGS (which was referred as quelling) (Cogoni and Macino 1999). Subsequently, EGO-1, RRF-1, and RRF-3 in *Caenorhabditis elegans* (nematode) (Smardon et al. 2000), SGS2/SDE1 in *Arabidopsis thaliana* (flowering plant) (Dalmay et al. 2000) and RrPA in *Dictyostelium discoideum* (slime mold) (Martens et al. 2002) were found to have RDR activity.

Evolutionarily, eukaryotic RDR and bacterial DNA dependent RNA polymerase (DDRP or RNAP) (and its orthologues in archaea and eukaryotes) are correlated with the presence of double psi beta barrel (DPBB) which is part of the catalytic core for both the enzymes. DPBB consists of 6- $\beta$  sheets. The bacterial core RNAP consists of a single copy of  $\beta$ ,  $\beta'$  and  $\omega$  and two copies of  $\alpha$  subunits (Burgess 1969). The  $\beta$  and  $\beta'$  are the largest subunits of RNAP and contains single DPBB in both subunits. Within a single  $\beta\beta'$  type polypeptide, RDR of *Neurospora crassa* contains two DPBBs. Since both the proteins (DDRPs and RDRs) can be of  $\beta\beta'$  type, therefore, DDRPs might have evolved from RDRs (Iyer et al. 2003). Even the amino acid sequences in the catalytic core of DDRPs and RDRs share similarities. RDRs have conserved amino acid sequences DXDGD (X can be any amino acid) in the catalytic core and DDRP has similar amino acid sequence NADFDGD. In both the polymerases negatively charged three aspartic acid residues (D) within the signature motif tend to bind a  $Mg^{2+}$  ion (Iyer et al. 2003).

### Classification of RDRs

In 2009, a group of researchers used the available genome sequences from public databases and performed a phylogenetic analysis of eukaryotic RDR gene homologs present in protists, plants, fungi and animals. This group explained the evolutionary relationship of RDR genes in different kingdoms of life and proposed a model for their evolution. Their study is also based on the similarity of the shared and conserved signature amino acid sequence DXDGD at the catalytic site. Multiple RDR homologs were found to be present in a wide group of plants including angiosperms (in *Arabidopsis* which is a eudicot and rice, a monocot), fern

(*Selaginella moellendorffii*) and moss (*Physcomitrella patens*). Similarly, all the three major groups of fungi i.e., ascomycetes, basidiomycetes and zygomycetes genomes possess *RDR* homologs (most of the fungus have 3–4 *RDR* homologs). Although, *RDRs* have been identified in plant, fungi, protists and in a few lower animals [For example, *RDR* gene homologs were found in *Dictyostelium* (protists) *Branchiostoma floridae* (Chordate), *Nematostella vectensis* (sea anemone, Cnidaria) and in *Ixodes scapularis* (Arthropod) genomes (Zong et al. 2009)] but most of the animals including the highly evolved humans and widely used model organism *Drosophila* lack *RDRs* (Stein et al. 2003), even though these organisms harbor the phenomenon known as RNA interference (RNAi). One explanation for the existence of RNAi in these animals might be through the horizontal transfer of viral *RDRs* which can complement for the endogenous *RDRs*. There is no evidence of the presence of *RDR* in bacteria and archaea so far (Burton 2014).

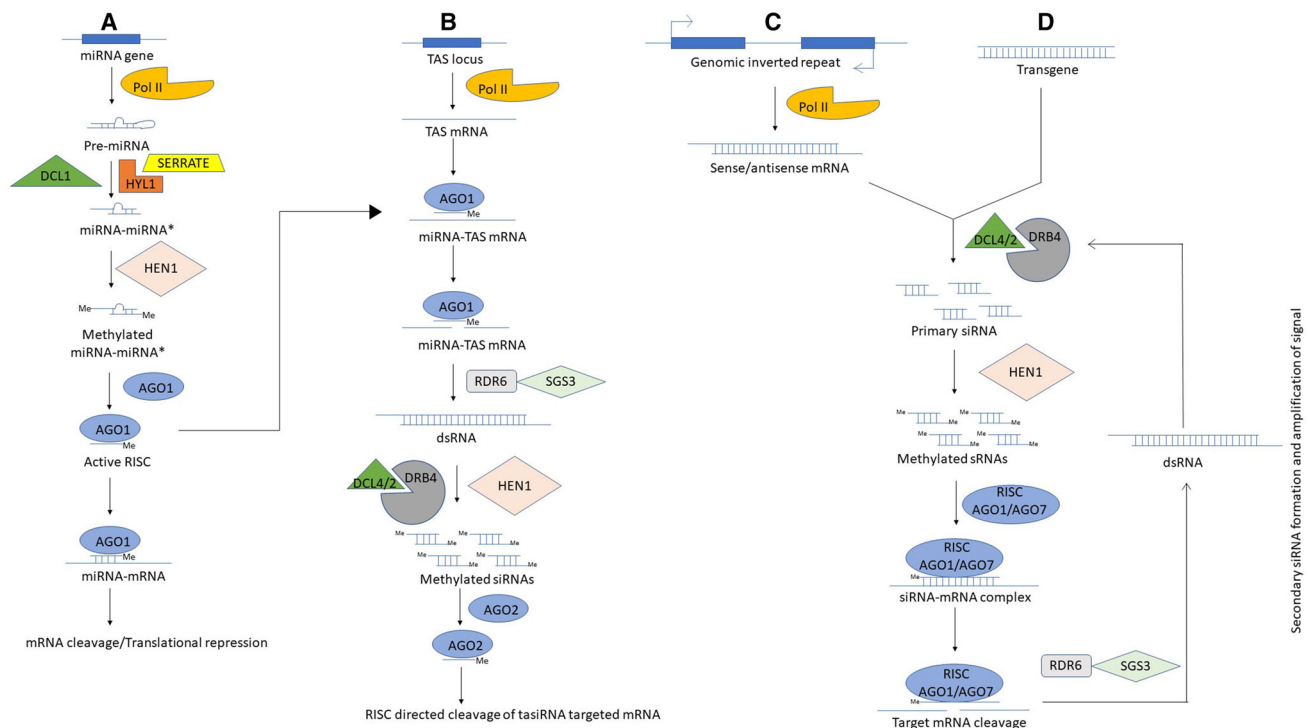
Eukaryotic *RDRs* are divided into three major clades—*RDR* $\alpha$ , *RDR* $\beta$  and *RDR* $\gamma$ . The most recent common ancestor of fungi, plants and animals must have possessed at least one copy of *RDR* from each of this group. All the three kingdoms i.e., Plants, animals and fungi possess *RDR* $\alpha$  genes. *RDR* $\beta$  genes are present in animals and fungi while *RDR* $\gamma$  genes are contained in the genomes of plants and fungi. Thus, during the course of evolution *RDR* $\beta$  genes were disappeared from plant kingdom while *RDR* $\gamma$  genes were lost from animal kingdom (Zong et al. 2009). The model plant *Arabidopsis* genome encodes for six *RDR* proteins, viz., *RDR1*, *RDR2*, *RDR3*, *RDR4*, *RDR5* and *RDR6*. Out of six, *RDR1*, *RDR2* and *RDR6* share a typical DLDGD signature motif sequence at the C-terminal catalytic site and belongs to *RDR* $\alpha$  clade. Many *RDRs* identified from other plants also belong *RDR* $\alpha$  clade. While other three, i.e., *RDR3*–*RDR5* belongs to *RDR* $\gamma$  clade and share an atypical DFDGD motif at the catalytic site (Wassenegger and Krczal 2006). Until today, the function of *RDR3*, *RDR4* and *RDR5* have been elusive while there are many reports on the functional analysis of *RDR1*, *RDR2* and *RDR6* in *Arabidopsis* as well as in other plant species. Recently it was found that tomato allelic *Ty-1* and *Ty-3* genes code for *RDR*, belong to *RDR* $\gamma$  clade and contain an atypical DFDGD motif. Moreover, *Ty-1* is the first geminivirus resistant gene to be identified in tomato (Verlaan et al. 2013). We will focus on plant *RDRs* from now onwards.

## Function, regulation and regional expression of *RDR1* and *RDR6*

### *RDR1* and *RDR6* in antiviral silencing: post-transcriptional gene silencing

The facts that *RDR1* and *RDR6* are important components of RNA silencing and function in antiviral activity have been implicated in many plants. RNA silencing was affected when the accumulation of *RDR6* transcript was inhibited in *N. benthamiana* (contains a natural mutant variant of *RDR1*) (Schwach et al. 2005; Qu et al. 2005). Furthermore, the plants showed susceptibility to certain viruses tested when the expression of *RDR1* and *RDR6* were affected (Dalmay et al. 2000; White 1979; Xie et al. 2001; Mourrain et al. 2000), suggesting the involvement of *RDR1* and *RDR6* in antiviral defense. Rakshandehroo et al. (2009) demonstrated that the expression of *RDR6* transcript is affected when the accumulation of *RDR1* mRNA is inhibited, however, the mechanism behind this is not known. Moreover, *RDR1* and *RDR6* are the two-main antiviral *RDRs* out of the six *RDRs* in *Arabidopsis*, which restrict the pathogen at both the single cell and systemic level (Wang et al. 2010). *RDR6* along with many other cellular factors are required for intracellular PTGS as well as for the efficient cell to cell movement of siRNA signal leading to systemic PTGS during VIGS (virus-induced gene silencing) (Smith et al. 2007; Searle et al. 2010; Melnyk et al. 2011; Qin et al. 2012). However, role of other factors, if any involved in this intercellular movement of antiviral siRNA signal needs to be elucidated.

In *Arabidopsis*, there are specialized classes of sRNAs, DCL and AGO proteins and the phenomena of TGS and PTGS depends on a set of combinations of each of them. 21–22 nt endogenous sRNAs may be categorized into miRNA, ta-siRNA, and natural antisense siRNA (nat-siRNA), depending on their mode of formation, which leads to the PTGS pathway while 24-nt siRNA, hcsiRNA (heterochromatic siRNA)/rasiRNA (repeat-associated siRNA) directs TGS (Chapman and Carrington 2007; Vaucheret 2006). 22-nt siRNAs are processed by DCL2, a less well studied DCL and is involved in the formation of nat-siRNA (Borsani et al. 2005; Bouche et al. 2006) (Fig. 1C). Moreover, DCL2 is also involved in the secondary siRNA mediated transitive silencing of transgene-mediated by *RDR6* (Mlotshwa et al. 2008; Parent et al. 2015) (Fig. 1D). The 22-nt siRNA is used as a substrate by *RDR6* to synthesize longer dsRNA and subsequently, DCL4 competes with DCL2, culminating in comparatively lower abundance of 22-nt siRNAs than 21-nt siRNAs (Parent et al. 2015). The 21-nt miRNA produced by DCL1 (Fig. 1A) and 22-nt siRNA produced by DCL2 are the



**Fig. 1** Schematic representation of post-transcriptional gene silencing pathways in plants. **A** miRNA pathway: miRNA generation requires the synthesis of pre-miRNA transcripts by RNA Pol II. Presence of complementary bases allows pre-miRNA transcripts to form internal stem and loop structure which is cleaved by DCL1 into 21-nt mature miRNA. HYL1 and SERRATE which are AGO1 binding proteins assist the DCL1 and HEN1 methylate the 3' ribose sugars for stabilization of miRNA protecting from uridylation. miRNAs are loaded onto AGO1 and form active miRNA-RISC complex which can suppress the expression of genes. **B** tasiRNA pathway: active miRNA-RISC complex can also target TAS transcripts which are transcriptional product of RNA Pol II. Following miRNA mediated cleavage of TAS transcripts, RDR6 with the help of SGS3 can synthesize long dsRNAs. DCL4-DRB4/DCL2

triggers for the secondary siRNA production (Wang et al. 2010). More recent finding in *Arabidopsis* indicates that DCL2 is a key component for RDR6 dependent systemic PTGS (Taochy et al. 2017).

DCL4 cleaves endogenous RDR6 dependent dsRNA precursors into 21-nt long tasiRNA (Fig. 1B) and pha-siRNA which silence the endogenous mRNA post-transcriptionally and the process is known as PTGS (Xie et al. 2005; Liu et al. 2014; Komiya 2017). During virus infection, RDR1 and RDR6 derived dsRNA precursors are also cleaved by DCL2 and DCL4 into 22nt and 21-nt vsiRNA (viral siRNA) to silence viral gene post-transcriptionally and the process is known as VIGS (Fig. 2) (Qu et al. 2005; Liu et al. 2014). However, it is the DCL4 which is required for the potent antiviral silencing through secondary vsiRNA but not DCL2 (Garcia-Ruiz et al. 2010). Efficient antiviral silencing could not be accompanied by the 22-nt viral secondary siRNAs produced by DCL2 (Garcia-Ruiz

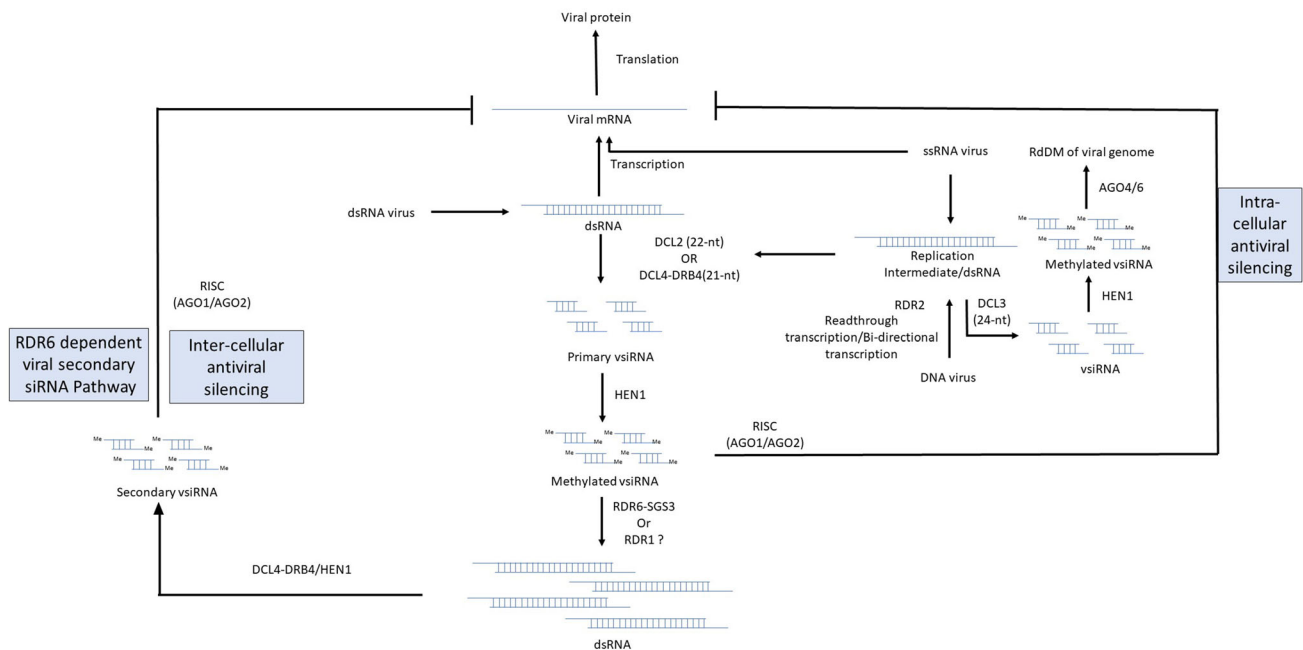
et al. 2010). The siRNAs produced in a virus-infected cell can spread to adjacent cells and further to the distal parts of the plant through non-cell autonomous VIGS and this process makes these cells resistant to further viral invasion (Schwach et al. 2005). DCL4 and 21-nt siRNAs produced by DCL4 are needed for this cell to cell systemic spread in *Arabidopsis* (Dunoyer et al. 2005). Recent findings in *N. benthamiana* suggest that DCL4 enhances while DCL2 reduces the intercellular spread of VIGS (Qin et al. 2017). Furthermore, dsRNA binding protein (DRB4) interacts with DCL4 and is required for DCL4 activity at least in seedling extract since a mutation in DRB4 abolishes DCL4 function (Fukudome and Fukuhara 2017; Nakazawa et al. 2007; Hiraguri et al. 2005).

The extent of RNA silencing varies in different organs of the plants. When *Arabidopsis* transgenic plant was silenced using sense PTGS, root tissues exhibited lower RNA silencing than leaves (Wilde et al. 2001). Likewise,

et al. 2010). The siRNAs produced in a virus-infected cell can spread to adjacent cells and further to the distal parts of the plant through non-cell autonomous VIGS and this process makes these cells resistant to further viral invasion (Schwach et al. 2005). DCL4 and 21-nt siRNAs produced by DCL4 are needed for this cell to cell systemic spread in *Arabidopsis* (Dunoyer et al. 2005). Recent findings in *N. benthamiana* suggest that DCL4 enhances while DCL2 reduces the intercellular spread of VIGS (Qin et al. 2017). Furthermore, dsRNA binding protein (DRB4) interacts with DCL4 and is required for DCL4 activity at least in seedling extract since a mutation in DRB4 abolishes DCL4 function (Fukudome and Fukuhara 2017; Nakazawa et al. 2007; Hiraguri et al. 2005).

The extent of RNA silencing varies in different organs of the plants. When *Arabidopsis* transgenic plant was silenced using sense PTGS, root tissues exhibited lower RNA silencing than leaves (Wilde et al. 2001). Likewise,





**Fig. 2** A model representing steps involved in VIGS during RNA and DNA virus infection in plants. dsRNA produced by dsRNA viruses or by ssRNA viruses or by DNA viruses during bidirectional transcription acts as a trigger for silencing process. 21–22 nt vsiRNA production mediated by DCL2/4 inhibits the expression of viral transcripts locally by forming vsiRNA-RISC complex. While RDR6

(RDR1?)-DCL4 mediated amplification produces secondary vsiRNA which move to distal parts of the plant and mediate systemic silencing of viral transcripts. The double stranded RNA formed during DNA virus transcription can be elicitor of TGS in which DCL3 cleaves the dsRNA into abundant 24-nt vsiRNAs. 24-nt vsiRNA load on to AGO4/6 and mediate RdDM of DNA genome of viruses

in *N. tabacum*, silencing of endoplasmic reticulum  $\omega$ -3 fatty acid desaturase (NtFAD3) by sense transgene could be observed in leaves but not in roots, even though the presence of transgene-derived siRNA was found in both root and leaf tissues (Tomita et al. 2004). However, inverted repeat (IR) transgene (which are engineered to make dsRNAs) mediated silencing of the target gene is equally effective in both leaf and root tissues (Fusaro et al. 2006; Marjanac et al. 2009). Unlike IR transgene-mediated silencing, dsRNA formation from ssRNA is required in the sense transgene-mediated silencing process. Therefore, it has been proposed that lower extent of sense-PTGS in root might be because of less activity of RDR6 which convert ssRNA into dsRNA or less efficiency of DCL protein in processing RDR6 dependent dsRNA into siRNA (Andika et al. 2016). Moreover, the similar expression level of core genes involved in RNA silencing pathways has been observed in leaves and roots of *Arabidopsis*, *N. benthamiana* and rice (Kapoor et al. 2008; Nakasugi et al. 2013). Therefore, it is not clear why sense transgene silencing operates at different rates between roots and leaves. Recent studies suggest that non-cell autonomous sRNAs (mobile sRNAs) mediate RdDM of many loci in root cells (Molnar et al. 2010; Lewsey et al. 2016).

Both the RDRs, RDR1 and RDR6 specifically recognize and amplify the tripartite positive strand RNA genome of

*Cucumber mosaic virus* (CMV). siRNA from 5'-terminal of each of the three viral genomic RNAs of CMV are preferentially targeted by RDR1 for amplification, on the other hand, RDR6 is appeared to amplify the 3' half of RNA 3 in the *rdr1* mutant (Wang et al. 2010). Suppressor of gene silencing 3 (SGS3), which is a RNA binding protein, does interact with RDR6 and is indispensable for RDR6 mediated dsRNA biogenesis which might take place in the cytoplasmic bodies (Mourrain et al. 2000; Jouannet et al. 2012; Lam et al. 2012). However, SGS3 is not required for RDR1 mediated siRNA biogenesis (Cao et al. 2014). More recently it has been found that polyadenylated tail of canonical endogenous mRNAs block RDR6 from converting them into dsRNA and aberrant RNAs lacking poly (A) tail are used as the templates over endogenous mRNAs. This phenomenon can be viewed as a way to protect plant from PTGS against endogenous mRNA (Baeg et al. 2017). Moreover, attempts to understand the subcellular-localization of RDR1 has been made but till today RDR1's localization remains a mystery (Lam et al. 2012). Furthermore, it is a well-established fact that virus infection triggers the induction of *RDR1* transcripts, however, the molecular mechanism of activation of *RDR1* transcript has been studied only recently in rice. Researchers have shown that the expression of monocot-specific miR444 is enhanced upon *Rice stripe virus* (RSV) infection.

Overexpression of miR444 strengthens the rice resistance against RSV by inducing the expression of *OsRDR1*. miR444 induces expression of *OsRDR1* by negatively regulating the MIKCC-type MADS-box proteins (which binds to the promoter of *OsRDR1*) (Wang et al. 2016).

The greater number of siRNAs produced during viral infection, known as viral “secondary” siRNAs, are the products of amplification by either RDR1 or RDR6 in *Arabidopsis* (Wang et al. 2010; Garcia-Ruiz et al. 2010). Reports suggest that vsiRNAs not only silence viral genes but can also inhibit the expression of host genes (Zheng et al. 2007). Interestingly, recently it has been confirmed that RDR1 is not only involved in production of vsiRNAs but also generate siRNAs from the host genes (known as virus-activated siRNA or vasiRNA) during virus infection which are supposed to target the host genes itself. However, this response of host i.e., vasiRNA generation upon virus infection is virus-specific because CMV 2b protein, which is a viral suppressor of gene silencing (VSR), inhibits RDR1 dependent vasiRNA production while VSR of *Turnip mosaic virus* (TuMV) does not affect its production (Cao et al. 2014).

Apart from the role of RDR1 in PTGS, a recent study in rice showed that RDR1 may also be involved in the TGS by RdDM and depends on 21-nt sRNA for silencing the genomic loci (Wang et al. 2014), however, further studies are required to strengthen this notion.

### Regulation of RDR1, RDR6 by phytohormones

Xie et al. (2001) established a link between RDR1 and salicylic acid (SA). They found that there is an increase in *RDR1* transcript level after treatment with SA [or with TMV] in tobacco. They also reported that *NtRDR1* down-regulated lines of tobacco showed susceptibility towards TMV and *Potato virus X* (PVX) and application of exogenous SA to these plants could induce resistance against these viruses. Later, SA inducible *RDR1* was also found in *Arabidopsis*. In the same study, it was found that the *AtRDR1* mutant plants also had no effect on SA dependent viral resistance (Yu et al. 2003) indicating host specific roles and regulation of RDR1. *N. benthamiana* harbors a natural mutant variant of *RDR1* gene with loss of function which is also SA inducible. Interestingly, overexpression of *M. truncatula RDR1* (*MtRDR1*) in *N. benthamiana* showed an increase in resistance against TMV (Yang et al. 2004). Contrary to *MtRDR1*, *NtRDR1* from tobacco acts as a suppressor of RNA silencing and increases infection of virus in *N. benthamiana* (Ying et al. 2010). However, *N. benthamiana* transformed with *MtRDR1* transgene inhibit virus spread to meristem and fatal symptom development during TMV infection and this inhibition is enhanced by SA treatment. In a recent study, a

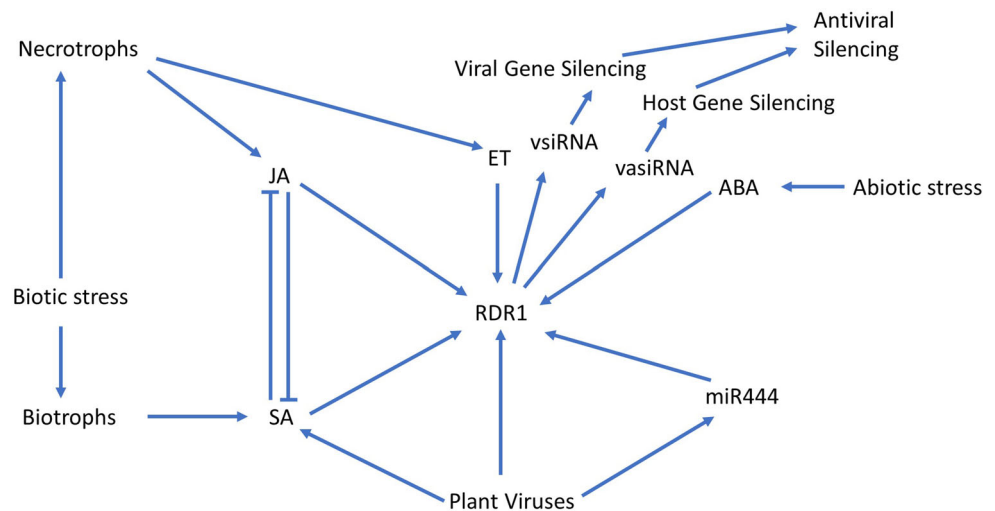
small gene family of four *RDR1* genes (*CsRDR1a*, *CsRDR1b* and *CsRDR1c1/c2*) from cucumber and their relationship with SA was studied. The *CsRDR1a* and *CsRDR1b* express at a higher level in a healthy plant while *CsRDR1c1/c2* were almost undetectable. After virus infection, the expressions of all four *RDR1* genes were induced and transcript level of *CsRDR1b* was found to be increased 10–20 folds in virus-resistant transgenic lines. While in virus susceptible lines of cucumber, the transcript level of *CsRDR1c1/c2* was enhanced by 25–1300 folds. Furthermore, all four *RDR1* genes were induced by ectopic application of SA (Leibman et al. 2017).

Recent reports suggest that *RDR1* is not only induced by SA but also by a wide range of phytohormones including jasmonic acid (JA) and abscisic acid (ABA). *N. attenuata RDR1* is induced by JA which is known to be induced by necrotrophic pathogens like insects and is considered as antagonistic to SA (Pandey and Baldwin 2007; Thaler et al. 2004). Silencing of *N. attenuata RDR1* using inverted repeat (*irRDR1*) showed susceptibility towards insects both in the glasshouse and in the fields (Pandey and Baldwin 2007). JA and ethylene (ET) are central to the defense machinery in plants against necrotrophs. The level of JA was found to be less accumulated in *irRDR1* plants, in contrast, ethylene accumulation was enhanced in *irRDR1* compared to wild-type plants (Pandey et al. 2008). *ZmRDR1* from Maize was also induced by MeJA, SA and infection with *Sugarcane mosaic virus* (He et al. 2010). Despite the role of SA in *RDR1* induction, the aforesaid observations implicate link between JA and RDR1 as well. Other than SA and JA, ABA and ET might also trigger *RDR1* expression, however, this induction is plant specific (Hunter et al. 2013). For example, ABA induces *NgRDR6* and *OsRDR6* but not *NgRDR1* and *OsRDR1* respectively which are induced by SA and several viruses (Liu et al. 2009; Yang et al. 2008). Furthermore, downregulating the expression of *NtRDR1* gene in transgenic tobacco showed reduced expression of many defense-related genes including *NtRDR6* and alternative oxidase 1a (*Aox1*) (Hunter et al. 2013). Therefore, based on above-mentioned evidences it's obvious that *RDR1* is regulated by many factors at various levels during plant defense (Fig. 3). Nonetheless, further research is required to understand the regulation of *RDR1*.

### Regional expression of RDR1 and RDR6

There is no general pattern of RDR1 expression in the tissues of different parts of plants. For instance, RDR1 in *N. benthamiana*, which is a natural loss of function mutant, expresses strongly in flowers (Yang et al. 2004). Comparative transcript analysis of the expression pattern of RDR1 in various tissues of maize and rice also showed

**Fig. 3** Networking of the factors regulating RDR1 gene expression which is involved in antiviral silencing. Various plant hormones such as SA, JA, ABA and ET have been found to induce the RDR1 expression in various plant species, however, this response is plant specific. Many plant viruses also enhance RDR1 expression in plants. In rice, viral infection triggers the increased expression of mir444 which positively regulate RDR1 expression



the similar results (Kapoor et al. 2008; He et al. 2010). However, in *Arabidopsis*, RDR1 expresses intensely in the vascular tissues and particularly the phloem cell layers of root showed strong expression (Xu et al. 2013). Cotton RDR6 has been found to express constitutively highest in the root, lower in leaves and lowest in the stem (Wang et al. 2012). The knowledge about expression pattern of RDR1 is limited to conclude its significance in antiviral resistance.

## Function, regulation and regional expression of RDR2

### Role of RDR2: RNA directed DNA methylation and transcriptional gene silencing

The term ‘changes of phase’ was coined by Barbara McClintock to explain how transposons change between active and inactive forms and from that time there has been a lot of interest among researchers to understand the heritable epigenetic changes by TGS through RdDM. The phenomenon of TGS is brought about by the methylation of cytosine bases in the DNA (RdDM) and selected modifications of the histone proteins which either stabilizes or destabilizes the chromatin architecture. RdDM operates and silences highly repetitive sequences, pericentric regions, telomeres, retroelements and transposons (Willmann et al. 2011; Chan et al. 2005).

*RDR2 is a main component of canonical RdDM pathway*

Current opinion about the TGS and RdDM comes from research on *Arabidopsis*. Like RDR1 and RDR6, RDR2 also belongs to RDR $\alpha$  clade and possesses a typical C-terminal DLDGD signature motif at the catalytic site

(Wassenegger and Krczal 2006). While RDR1 and RDR6 are main participants of the PTGS pathway, RDR2 functions in TGS pathway in plants and have orthologues in many species of plants. RDR2 protein not only functions in the antiviral and transgene silencing but also in the various aspect of plant’s physiology and development by modifying chromatin structure leading to the phenomena of imprinting and paramutation. The most abundant endogenous siRNAs in plants i.e. hcsiRNA which are 24-nt in length, are the product of the activity of RDR2 and DCL3 along with the assistance of many other proteins and are the major player to guide the methylation process. hcsiRNA mediated methylation and histone modification at the *Arabidopsis* telomere are dependent on the RDR2-mediated RdDM thus heterochromatinizing and switching off the expression of telomeric genes (Vrbsky et al. 2010). Notably, it is a common feature of telomere of many organisms to be in heterochromatinized state (Vrbsky et al. 2010).

*Canonical RNA directed DNA methylation* A typical RdDM pathway involves two consecutive steps: (1) generation of 24-nt hcsiRNA/rasiRNA which require DNA dependent RNA polymerase IV (Pol IV), RDR2 and DCL3. (2) de novo methylation which needs DNA dependent RNA polymerase V (Pol V)–dependent scaffold non-coding RNAs, 24-nt hcsiRNAs/rasiRNAs loaded AGO-4, -6 or -9, DNA methyltransferase Domains Rearranged Methyltransferase 2 (DRM2), as well as many other proteins. Generation of single-stranded transcripts from the heterochromatic region by Pol IV (which is plant specific enzyme) and CLASSY1 marks the initiation of TGS process. The ssRNA produced in the nucleus is used as a template by RDR2 which convert into dsRNA. Recent evidences suggest that Pol IV and RDR2 physically associate with each other (Haag et al. 2012, 2014; Law et al.

2011). Analysis of dsRNA produced by Pol IV-RDR2 in *dcl2/3/4* mutants suggests that Pol IV-RDR2 generated dsRNA could be of several hundreds of nucleotides in length and many siRNAs could be generated from them (Li et al. 2015). DCL3 cleaves the dsRNA into 24-nt hcsiRNAs/rasiRNAs which are then methylated by HEN1 (Xie et al. 2004; Lu et al. 2006). Apart from the role of DCL3 in generating 24-nt siRNAs, generation of 23–25 nt miRNAs from long miRNA precursors have also been reported by DCL3 (Vazquez et al. 2008). hcsiRNAs/rasiRNAs generated in the nucleus are exported to the cytoplasm. In the cytoplasm hcsiRNAs/rasiRNAs load onto AGO-4, -6 or -9 and import back to the nucleus where they bind with complementary non-coding RNA transcribed by Pol V. Chromatin remodeling DDR complex [Defective In RNA-Directed DNA Methylation 1 (DRD1), Defective In Meristem Silencing 3 (DMS3) and RNA-Directed DNA Methylation 1 (RDM1)] assist the generation of pol V mediated non-coding RNA transcripts (Kanno et al. 2004, 2008; Wierzbicki et al. 2008). The non-coding RNA transcribed by Pol V acts as a scaffold to bring AGO-4, -6 or -9 bound hcsiRNA/rasiRNA to the genomic loci and marks the complementary sequences for methylation both in cis- and trans- positions (Zheng et al. 2007; Havecker et al. 2010; Olmedo-Monfil et al. 2010; Wierzbicki et al. 2009; Law and Jacobsen 2010). DNA transferases such as Domains Rearranged Methylase1 and 2 (DRM1 and DRM2) are recruited by the AGO-4, -6 or -9 bound hcsiRNA/rasiRNA at the target genomic loci by an unknown process and carry out de novo methylation at cytosine in all sequence contexts (CG, CHG and CHH, where H is A, C or T) (Law and Jacobsen 2010; Borges and Martienssen 2015). Pol V and chromatin remodeling DDR complex are crucial for the heterochromatin organization since mutation in these proteins disrupts the heterochromatin formation during interphase while mutation in Pol IV, RDR2 or DCL3 does not affect so (Pontes et al. 2009). Interestingly, dsRNA generated by Pol IV-RDR2 can also induce methylation without being cleaved into siRNAs (Ye et al. 2016; Yang et al. 2016).

A recent study suggests that JMJ24, a JmjC domain-containing protein interacts physically with RDR2 and is responsible for basal level transcription of heterochromatic silenced loci by counteracting H3K9me to initiate RdDM (Deng et al. 2015).

**Non-canonical RNA directed DNA methylation** Recent evidences indicate presence of another RNA dependent methylation pathway which links with PTGS and does not involve Pol IV, RDR2 and DCL3. Pol II-RDR6 pathway has been observed at young transposons and some TAS loci, which are known to generate tasiRNA, indicating a shift from PTGS to TGS pathway (Nuthikattu et al. 2013;

Panda and Slotkin 2013; Marí-Ordóñez et al. 2013). The pathway starts with the synthesis of transcript by Pol II which is further used by RDR6 as a template to convert it to dsRNA substrate for DCL2/DCL4. The resulting 21–22 nt siRNAs are possibly loaded onto AGO-4 or -6 and further initiate Pol V mediated RdDM. Alternatively, 21–22 nt siRNAs might enter the typical PTGS pathway by loading onto AGO1 and thus, targeting the mRNA transcripts. The requirement for RDR2-Independent DNA Methylation (NERD), a plant-specific GW protein, is known to bind AGO and is involved in RDR6 mediated RdDM (Pontier et al. 2012).

### Regulation of RDR2 by phytohormones

Although, Campos et al. (2014) showed that RDR2 can be induced significantly by SA and its metabolic derivative gentisic acid (GA) in tomato, much more research is needed to fully understand the regulation of RDR2 in plants.

### Regional expression of RDR2

*Arabidopsis thaliana* Tiling Array Express (At-TAX) is a tool to study the gene expression analysis and using this tool RDR2 has been found to be expressed almost uniformly in various parts of the plant. The study also showed that various stress treatments, such as salt, osmotic, heat, cold and ABA treatments do not affect RDR2 expression (Laubinger et al. 2008; Willmann et al. 2011). However, in cucumber, CsRDR2 expressed more in roots during dehydration treatment (Gan et al. 2017). Apart from phytohormone SA, *Citrus exocortis viroid* (CEVd) infection can also induce its expression (Campos et al. 2014).

### Function, regulation and regional expression of RDR3, 4 and 5

Despite many efforts, function, regulation and regional expression of RDRs of RDR $\gamma$  clade remains elusive (Willmann et al. 2011). There are very few evidences suggesting the role of RDR $\gamma$  proteins in plant defence and plant physiology. For example, *Salvia miltiorrhiza smRDR3* is induced after CMV infection suggesting its role in the antiviral silencing (Shao and Lu 2014). Moreover, in the natural environment, NaRDR3 may be needed for competitive growth in *Nicotiana attenuata* (Pandey et al. 2008). At-TAX tiling microarray experiments suggest that AtRDR3 and 5 are expressed strongly at inflorescence apex while AtRDR5 expresses almost uniformly in different parts of *Arabidopsis* plant (Willmann et al. 2011). *Ty1/Ty3* gene from tomato encodes for RDR similar to AtRDR3, 4 and 5 which displayed a new class of antiviral resistance.



*Ty1* and *Ty3* are alleles of the same gene. Cultivars of tomato with *Ty1/Ty3* displayed resistance against geminivirus *Tomato yellow leaf curl virus* (TYLCV) and displayed enhanced siRNA generation and cytosine methylation of geminiviral genome, interestingly such plants were susceptible to RNA virus like CMV. This shows that *Ty1/Ty3* is involved in siRNA amplification required for TGS but not involved in PTGS pathway thereby making the plant carrying allele *Ty1/Ty3* resistant to geminivirus, susceptible to RNA virus and mixed infections (Verlaan et al. 2013; Butterbach et al. 2014).

## Conclusion and future perspective

It is quite clear that most of the plants possess many homologs of RDR protein and many have redundant while others have specified functions. In this scenario the overall picture describing the specific functions of RDRs is not clear. For example, almost negligible information is available about RDR3, 4 and 5 and loss of function mutants could be promising in understanding the function of these proteins. Even the function of RDR1 is not fully understood. Reports have already suggested that both Pol IV and RDR2 are equally needed for the formation of dsRNA in *dcl3* mutant plants. Interestingly, affinity purified Pol IV has been shown to be transcriptionally active in vitro from *rdr2* null mutant plants (Haag et al. 2012). Similarly, recombinant RDR2 is also transcriptionally active in vitro suggesting that physical interaction between Pol IV and RDR2 is dispensable for their fundamental catalytic activities (Wendte and Pikaard 2017). Thus, at present, it is unknown that, in order to synthesize either strand of siRNA precursors why both the proteins are needed in vivo. Furthermore, there is almost absolute lack of information about how the expression of different RDRs is mechanistically regulated by phytohormones, viruses, transgenes or abiotic agents.

**Acknowledgements** The work was supported by the UPOE-II Grant of JNU to SC.

## Compliance with ethical standards

**Conflict of interest** We declare that the review is not written with any conflict of interest.

## References

Andika, I. B., Kondo, H., & Sun, L. (2016). Interplays between soil-borne plant viruses and RNA silencing-mediated antiviral defense in roots. *Frontiers in Microbiology*, 7, 1458.

- Astier-Manificier, S., & Cornuet, P. (1971). RNA-dependent RNA polymerase in Chinese cabbage. *Biochimica et Biophysica Acta*, 232, 484–493.
- Baeg, K., Iwakawa, H. O., & Tomari, Y. (2017). The poly (A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing. *Nature Plants*, 3, 17036.
- Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., et al. (2012). Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene*, 501(1), 52–62.
- Baltimore, D., & Franklin, R. M. (1962). Preliminary data on a virus-specific enzyme system responsible for the synthesis of viral RNA. *Biochemical and Biophysical Research Communications*, 9, 388–392.
- Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409, 363–366.
- Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology*, 16, 727–741.
- Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R., & Zhu, J. K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, 123, 1279–1291.
- Bouche, N., Laussergues, D., Gascioli, V., & Vaucheret, H. (2006). An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO Journal*, 25, 3347–3356.
- Burgess, R. R. (1969). Separation and characterization of the subunits of ribonucleic acid polymerase. *Journal of Biological Chemistry*, 244, 6168–6176.
- Burton, Z. F. (2014). The old and new testaments of gene regulation: Evolution of multi-subunit RNA polymerases and co-evolution of eukaryote complexity with the RNAP II CTD. *Transcription*, 5, e28674.
- Butterbach, P., Verlaan, M. G., Dulleman, A., Lohuis, D., Visser, R. G. F., Bai, Y., et al. (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. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 12942–12947.
- Campos, L., Granell, P., Tarraga, S., Lopez-Gresa, P., Conejero, V., Belles, J. M., et al. (2014). Salicylic acid and gentisic acid induce RNA silencing-related genes and plant resistance to RNA pathogens. *Plant Physiology and Biochemistry*, 77, 35–43.
- Cao, M., Du, P., Wang, X., Yu, Y. Q., Qiu, Y. H., Li, W., et al. (2014). Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. *Proceedings of the National Academy of Sciences*, 111(40), 14613–14618.
- Cao, J. Y., Xu, Y. P., Li, W., Li, S. S., Rahman, H., & Cai, X. Z. (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*. *Frontiers in Plant Science*, 7, 1614.
- Carmell, M. A., & Hannon, G. J. (2004). RNase III enzymes and the initiation of gene silencing. *Nature Structural & Molecular Biology*, 11, 214–218.
- Chan, S. W., Henderson, I. R., & Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews Genetics*, 6, 351–360.
- Chapman, E. J., & Carrington, J. C. (2007). Specialization and evolution of endogenous small RNA pathways. *Nature Reviews Genetics*, 8, 884–896.

- Cogoni, C., & Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature*, *399*, 166–169.
- Dalmay, T., Hamilton, A., Rudd, S., Angell, S., & Baulcombe, D. C. (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*, 543–553.
- Deng, S., Xu, J., Liu, J., Kim, S.-H., Shi, S., & Chua, N.-H. (2015). JM124 binds to RDR2 and is required for the basal level transcription of silenced loci in *Arabidopsis*. *The Plant Journal*, *83*, 770–782.
- Devert, A., Fabre, N., Floris, M., Canard, B., Robaglia, C., & Crete, 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.
- Dunoyer, P., Himber, C., & Voinnet, O. (2005). DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nature Genetics*, *37*, 1356–1360.
- Fernandes-Brum, N. C., Rezende, M. P., Ribeiro, C., et al. (2017). A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. *PLoS ONE*, *12*(4), e0176333.
- Fukudome, A., & Fukuhara, T. J. (2017). Plant dicer-like proteins: Double-stranded RNA-cleaving enzymes for small RNA biogenesis. *Journal of Plant Research*, *130*, 33.
- Fusaro, A. F., Matthew, L., Smith, N. A., Curtin, S. J., Dedic-Hagan, J., Ellacott, G. A., et al. (2006). RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Reports*, *7*, 1168–1175.
- Gan, D. F., Liang, D. D., Wu, J., Zhan, M. D., Yang, F., Xu, W. J., et al. (2016). Genome-wide identification of the Dicer-like, Argonaute, and RNA-dependent RNA polymerase gene families in cucumber (*Cucumis sativus* L.). *Journal of Plant Growth Regulation*, *35*, 135–150.
- Gan, D., Zhan, M., Yang, F., et al. (2017). Expression analysis of argonaute, Dicer-like, and RNA-dependent RNA polymerase genes in cucumber (*Cucumis sativus* L.) in response to abiotic stress. *Journal of Genetics*, *96*, 235.
- Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., et al. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during turnip mosaic virus infection. *The Plant Cell*, *22*, 481–496.
- Haag, J. R., Brower-Toland, B., Krieger, E. K., Sidorenko, L., Nicora, C. D., Norbeck, A. D., et al. (2014). Functional diversification of maize RNA polymerase IV and v subtypes via alternative catalytic subunits. *Cell Reports*, *9*, 378–390.
- Haag, J. R., Ream, T. S., Marasco, M., Nicora, C. D., Norbeck, A. D., Pasa-Tolic, L., et al. (2012). In vitro transcription activities of pol IV, pol v, and RDR2 reveal coupling of pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Molecular Cell*, *48*, 811–818.
- Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., et al. (2010). The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *The Plant Cell Online*, *22*, 321–334.
- He, J., Dong, Z., Jia, Z., Wang, J., & Plant, G. (2010). Isolation, expression and functional analysis of a putative RNA-dependent RNA polymerase gene from maize (*Zea mays* L.). *Molecular Biology Reports*, *37*, 865–874.
- Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., et al. (2006). Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genetics*, *38*, 721–725.
- Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., et al. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB family dsRNA-binding proteins in *Arabidopsis thaliana*. *Plant Molecular Biology*, *57*, 173–188.
- Hunter, L. J., Westwood, J. H., Heath, G., Macaulay, K., Smith, A. G., Macfarlane, S. A., et al. (2013). Regulation of RNA-dependent RNA polymerase 1 and isochorismate synthase gene expression in *Arabidopsis*. *PLoS ONE*, *8*, e66530.
- Iyer, L. M., Koonin, E. V., & Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Structural Biology*, *3*, 1–23.
- Jouanet, V., Moreno, A. B., Elmayer, T., Vaucheret, H., Crespi, M. D., & Maizel, A. (2012). Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *EMBO Journal*, *31*, 1704–1713.
- Kanno, T., Bucher, E., Daxinger, L., Huettel, B., Bohmdorfer, G., Gregor, W., et al. (2008). A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nature Genetics*, *40*, 670–675.
- Kanno, T., Mette, M. F., Kreil, D. P., Aufsatz, W., Matzke, M., & Matzke, A. J. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Current Biology*, *14*, 801–805.
- Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., et al. (2008). Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*, *9*, 451.
- Komiya, R. (2017). Biogenesis of diverse plant phasiRNAs involves an miRNA-trigger and Dicer-processing. *Journal of Plant Research*, *130*(1), 17–23.
- Lam, P., Zhao, L., McFarlane, H. E., Aiga, M., Lam, V., Hooker, T. S., 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 Physiology*, *159*(4), 1385–1395.
- Laubinger, S., Zeller, G., Henz, S. R., Sachsenberg, T., Widmer, C. K., Naouar, N., et al. (2008). At-TAX: A whole genome tiling array resource for developmental expression analysis and transcript identification in *Arabidopsis thaliana*. *Genome Biology*, *9*(7), R112.
- Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, *11*, 204–220.
- Law, J. A., Vashisht, A. A., Wohlschlegel, J. A., & Jacobsen, S. E. (2011). SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genetics*, *7*, e1002195.
- Leibman, D., Kravchik, M., Wolf, D., Haviv, S., Weissberg, M., Ophir, R., et al. (2017). Differential expression of cucumber RNA-dependent RNA polymerase 1 genes during antiviral defence and resistance. *Molecular Plant Pathology*. <https://doi.org/10.1111/mpp.12518>.
- Lewsey, M. G., Hardcastle, T. J., Melnyk, C. W., Molnar, A., Valli, A., Ulrich, M. A., et al. (2016). Mobile small RNAs regulate genome-wide DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, E801–E810.

- Li, S., Vandivier, L. E., Tu, B., Gao, L., Won, S. Y., Li, S., et al. (2015). Detection of pol IV/RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis. *Genome Research*, 25, 235–245.
- Lindbo, J. A., Silva-Rosales, L., Proebsting, W. M., & Dougherty, W. G. (1993). Induction of a highly specific antiviral state in transgenic plants: Implications for regulation of gene expression and virus resistance. *The Plant Cell*, 5, 1749–1759.
- Liu, Y., Gao, Q., Wu, B., Ai, T., & Guo, X. (2009). NgRDR1, an RNA dependent RNA polymerase isolated from *Nicotiana glutinosa*, was involved in biotic and abiotic stresses. *Plant Physiology and Biochemistry*, 47, 359–368.
- Liu, X., Lu, T., Dou, Y., Yu, B., & Zhang, C. (2014). Identification of RNA silencing components in soybean and sorghum. *BMC Bioinformatics*, 15, 4.
- Lu, C., Kulkarni, K., Souret, F. F., MuthuValliappan, R., Tej, S. S., Poethig, R. S., et al. (2006). MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. *Genome Research*, 16, 1276–1288.
- Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., et al. (2006). The evolution and diversification of Dicers in plants. *FEBS Letters*, 580, 2442–2450.
- Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., & Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. *Nature Genetics*, 45, 1029–1039.
- Marjanac, G., Karimi, M., Naudts, M., Beeckman, T., Depicker, A., & De Buck, S. (2009). Gene silencing induced by hairpin or inverted repeated sense transgenes varies among promoters and cell types. *New Phytologist*, 184, 851–864.
- Martens, H., Novotny, J., Oberstrass, J., Steck, T. L., Postlethwait, P., & Nellen, W. (2002). RNAi in Dictyostelium: The role of RNA-directed RNA polymerases and double-stranded RNase. *Molecular Biology of the Cell*, 13, 445–453.
- Melnyk, C. W., Molnar, A., & Baulcombe, D. C. (2011). Intercellular and systemic movement of RNA silencing signals. *EMBO Journal*, 30, 3553–3563.
- Mlotshwa, S., Pruss, G. J., Peragine, A., Endres, M. W., Li, J., Chen, X., et al. (2008). DICER-LIKE2 plays a primary role in transitive silencing of transgenes in Arabidopsis. *PLoS ONE*, 3, e1755.
- Molnar, A., Melnyk, C. W., Bassett, A., Hardcastle, T. J., Dunn, R., & Baulcombe, D. C. (2010). Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science*, 328, 872–875.
- Mourrain, P., Beclin, C., Elmayer, T., Feuerbach, F., Godon, C., Morel, J.-B., et al. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural resistance. *Cell*, 101, 533–542.
- Nakasugi, K., Crowhurst, R. N., Bally, J., Wood, C. C., Hellens, R. P., & Waterhouse, P. M. (2013). De novo transcriptome sequence assembly and analysis of RNA silencing genes of *Nicotiana benthamiana*. *PLoS ONE*, 8, e59534.
- Nakazawa, Y., Hiraguri, A., Moriyama, H., & Fukuhara, T. (2007). The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 in vivo and functions in the trans-acting siRNA pathway. *Plant Molecular Biology*, 63, 777–785.
- Nuthikattu, S., McCue, A. D., Panda, K., Fultz, D., DeFraia, C., et al. (2013). The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. *Plant Physiology*, 162, 116–131.
- Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimanelli, D., et al. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. *Nature*, 464, 628–632.
- Panda, K., & Slotkin, R. K. (2013). Proposed mechanism for the initiation of transposable element silencing by the RDR6-directed DNA methylation pathway. *Plant Signaling & Behavior*, 8, e25206.
- Pandey, S. P., & Baldwin, I. T. (2007). RNA-directed RNA polymerase 1 (RdR1) mediates the resistance of *Nicotiana attenuata* to herbivore attack in nature. *The Plant Journal*, 50, 40–53.
- Pandey, S. P., Shahi, P., Gase, K., & Baldwin, I. T. (2008). Herbivory-induced changes in the small-RNA transcriptome and phytohormone signaling in *Nicotiana attenuata*. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 4559–4564.
- Parent, J. S., Bouteiller, N., Elmayer, T., & Vaucheret, H. (2015). Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing. *The Plant Journal*, 81, 223–232.
- Pontes, O., Costa-Nunes, P., Vithayathil, P., & Pikaard, C. S. (2009). RNA polymerase V functions in Arabidopsis interphase heterochromatin organization independently of the 24-nt siRNA-directed DNA methylation pathway. *Molecular Plant*, 2, 700–710.
- Pontier, D., Picart, C., Roudier, F., Garcia, D., Lahmy, S., et al. (2012). NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis. *Molecular Cell*, 48, 121–132.
- Powell, A. P., Nelson, R. S., De, B., Hoffmann, N., Rogers, S. G., Fraley, R. T., et al. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*, 232, 738–748.
- Qian, Y. X., Cheng, Y., Cheng, X., Jiang, H. Y., Zhu, S. W., & Cheng, B. J. (2011). Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Reports*, 30, 1347–1363.
- Qin, C., Li, B., Fan, Y., Zhang, X., Yu, Z., Ryabov, E., et al. (2017). Roles of Dicer-like Proteins 2 and 4 in intra- and intercellular antiviral silencing. *Plant Physiology*, 174(2), 1067–1081.
- Qin, C., Shi, N., Gu, M., Zhang, H., Li, B., Shen, J., et al. (2012). Involvement of RDR6 in short-range intercellular RNA silencing in *Nicotiana benthamiana*. *Scientific Reports*, 2, 467.
- Qu, F., Ye, X., Hou, G., Sato, S., Clemente, T. E., & Morris, T. J. (2005). RDR6 has a broad-spectrum but temperature dependent antiviral defense role in *Nicotiana benthamiana*. *Journal of Virology*, 79, 15209–15217.
- 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. *Molecular Plant-Microbe Interactions*, 22, 1312–1318.
- Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., & Sanger, H. L. (1993a). RNA-directed RNA polymerase from tomato leaves. I. Purification and physical properties. *Journal of Biological Chemistry*, 263, 11851–11857.
- Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., & Sanger, H. L. (1993b). RNA-directed RNA polymerase from tomato leaves. II. Catalytic in vitro properties. *Journal of Biological Chemistry*, 263, 11858–11867.
- Schiebel, W., Pelissier, T., Riedel, L., Thalmeier, S., Schiebel, R., Kempe, D., et al. (1998). Isolation of an RNA-directed RNA polymerase-specific cDNA clone from tomato. *Plant Cell*, 10, 2087–2101.
- Schwach, F., Vaistij, F. E., Jones, L., & Baulcombe, D. C. (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 Physiology*, 138, 1842–1852.
- Searle, I. R., Pontes, O., Melnyk, C. W., Smith, L. M., & Baulcombe, D. C. (2010). JM14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in Arabidopsis. *Genes & Development*, 24, 986–991.

- Shao, F., & Lu, S. (2014). Identification, molecular cloning and expression analysis of five RNA-dependent RNA polymerase genes in *Salvia miltiorrhiza*. *PLoS ONE*, *9*, e95117.
- Simon, S. A., & Meyers, B. C. (2011). Small RNA-mediated epigenetic modifications in plants. *Current Opinion in Plant Biology*, *14*, 148–155.
- Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N., & Maine, E. M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Current Biology*, *10*, 169–178.
- Smith, L. M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F. K., Herr, A. J., et al. (2007). An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell*, *19*, 1507–1521.
- Stein, P., Svoboda, P., Anger, M., & Schultz, R. M. (2003). RNAi: Mammalian oocytes do it without RNA-dependent RNA polymerase. *RNA*, *9*, 187–192.
- Taochy, C., Gursansky, N. R., Cao, J., Fletcher, S. J., Dressel, U., Mitter, N., Tucker, M. R., Koltunow, A. M., Bowman, J. L., Vaucheret, H., & Carroll, B. J. (2017). DCL2 promotes systemic PTGS in *Arabidopsis*. *Plant Physiol*, pp. 01181.2017.
- Thaler, J. S., Owen, B., & Higgins, V. J. (2004). The role of the jasmonate response in plant susceptibility to diverse pathogens with a range of lifestyles. *Plant Physiology*, *135*, 530–538.
- Tomita, R., Hamada, T., Horiguchi, G., Iba, K., & Kodama, H. (2004). Transgene overexpression with cognate small interfering RNA in tobacco. *FEBS Letters*, *573*, 117–120.
- Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: Mechanisms and regulations. *Genes & Development*, *20*, 759–771.
- Vazquez, F., Blevins, T., Ailhas, J., Boller, T., & Meins, F. (2008). Evolution of *Arabidopsis* MIR genes generates novel microRNA classes. *Nucleic Acids Research*, *36*, 6429–6438.
- Verlaan, M. G., Hutton, S. F., Ibrahim, R. M., Kormelink, R., Visser, R. G., Scott, J. W., 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 Genetics*, *9*(3), e1003399.
- Vrbsky, J., Akimcheva, S., Watson, J. M., Turner, T. L., Daxinger, L., Vyskot, B., et al. (2010). siRNA-mediated methylation of *Arabidopsis* telomeres. *PLoS Genetics*, *6*(6), e1000986. <https://doi.org/10.1371/journal.pgen.1000986>.
- Wang, H., Jiao, X., Kong, X., et al. (2016). A signaling cascade from miR444 to RDR1 in rice antiviral RNA silencing pathway. *Plant Physiology*, *170*(4), 2365–2377.
- Wang, M., Li, S., Yang, H., Gao, Z., Wu, C., & Guo, X. (2012). Characterization and functional analysis of GhRDR6, a novel RDR6 gene from cotton (*Gossypium hirsutum* L.). *Bioscience Reports*, *32*, 139–151.
- Wang, X.-B., Wu, Q., Ito, T., et al. (2010). RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 484–489.
- Wang, N., Zhang, D., Wang, Z., Xun, H., Ma, J., Wang, H., et al. (2014). Mutation of the RDR1 gene caused genome-wide changes in gene expression, regional variation in small RNA clusters and localized alteration in DNA methylation in rice. *BMC Plant Biology*, *14*, 177. <https://doi.org/10.1186/1471-2229-14-177>.
- Wassenegger, M., & Krzczal, G. (2006). Nomenclature and functions of RNA-directed RNA polymerases. *Trends in Plant Science*, *11*, 142–151.
- Wendte, J. M., & Pikaard, C. S. (2017). The RNAs of RNA-directed DNA methylation. *Biochimica et Biophysica Acta*, *1860*, 140–148.
- White, R. F. (1979). Acetylsalicylic acid (aspirin) induces resistance to tobacco mosaic virus in tobacco. *Virology*, *99*, 410–412.
- Wierzbicki, A. T., Haag, J. R., & Pikaard, C. S. (2008). Noncoding transcription by RNA Polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, *135*(4), 635–648.
- Wierzbicki, A. T., Ream, T. S., Haag, J. R., & Pikaard, C. S. (2009). RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genetics*, *41*, 630–634.
- Wilde, C. De, Podevin, N., Windels, P., & Depicker, A. (2001). Silencing of antibody genes in plants with single-copy transgene inserts as a result of gene dosage effects. *Molecular Genetics and Genomics*, *265*, 647–653.
- Willmann, M. R., Endres, M. W., Cook, R. T., & Gregory, B. D. (2011). The functions of RNA-dependent RNA polymerases in *Arabidopsis*. *Arabidopsis Book*, *9*, 1–20.
- Xie, Z., Allen, E., Wilken, A., & Carrington, J. C. (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 12984–12989.
- Xie, Z., Fan, B., Chen, C., & Chen, Z. (2001). An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 6516–6521.
- Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., et al. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biology*, *2*(5), E104.
- Xu, T., Zhang, L., Zhen, J., Fan, Y., Zhang, C., & Wang, L. (2013). Expressional and regulatory characterization of *Arabidopsis* RNA-dependent RNA polymerase 1. *Planta*, *237*(6), 1561–1569.
- Yadav, C. B., Muthamilarasan, M., Pandey, G., & Prasad, M. (2015). Identification, characterization and expression profiling of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in foxtail millet. *Plant Molecular Biology Reporter*, *33*(1), 43–55.
- Yang, S. J., Carter, S. A., Cole, A. B., Cheng, N. H., & Nelson, R. S. (2004). A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by *Nicotiana benthamiana*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 6297–6302.
- Yang, J. H., Seo, H. H., Han, S. J., Yoon, E. K., Yang, M. S., & Lee, W. S. (2008). Phytohormone abscisic acid control RNA-dependent RNA polymerase 6 gene expression and post-transcriptional gene silencing in rice cells. *Nucleic Acids Research*, *36*, 1220–1226.
- Yang, D. L., Zhang, G., Tang, K., Li, J., Yang, L., Huang, H., et al. (2016). Dicer-independent RNA-directed DNA methylation in *Arabidopsis*. *Cell Research*, *26*, 66–82.
- Ye, R., Chen, Z., Lian, B., Rowley, M. J., Xia, N., Chai, J., et al. (2016). A dicer-independent route for biogenesis of siRNAs that direct DNA methylation in *Arabidopsis*. *Molecular Cell*, *61*, 222–235.
- Ying, X.-B., Dong, L., Zhu, H., Duan, C.-G., Du, Q.-S., et al. (2010). RNA-dependent RNA polymerase 1 from *Nicotiana tabacum* suppresses RNA silencing and enhances viral infection in *Nicotiana benthamiana*. *The Plant Cell*, *22*, 1358–1372.
- Yu, D., Fan, B., MacFarlane, S. A., & Chen, Z. (2003). Analysis of the involvement of an inducible *Arabidopsis* RNA-dependent RNA polymerase in antiviral defence. *Molecular Plant-Microbe Interactions*, *16*, 206–216.
- Zhao, H. L., Zhao, K., Wang, J., Chen, X., Chen, Z., Cai, Y. H., et al. (2015). Comprehensive analysis of Dicer-like, Argonaute, and RNA dependent RNA polymerase gene families in grapevine

- (*Vitis Vinifera*). *Journal of Plant Growth Regulation*, 34, 108–121.
- Zheng, X., Zhu, J., Kapoor, A., & Zhu, J. K. (2007). Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO Journal*, 26, 1691–1701.
- 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.
- Zvereva, A. S., & Pooggin, M. M. (2012). Silencing and innate immunity in plant defense against viral and non-viral pathogens. *Viruses*, 4, 2578–2597.





## Turnitin Originality Report

Thesis by Rangunathan D...

From M.Phil/Ph.D 2019 (M.Phil/Ph.D)



- Processed on 10-Jul-2019 16:45 IST
- ID: 1150719969
- Word Count: 19465

## Similarity Index

10%

## Similarity by Source

## Internet Sources:

6%

## Publications:

8%

## Student Papers:

8%

**sources:**

- 1 1% match (student papers from 12-Jul-2016)  
[Submitted to Jawaharlal Nehru University \(JNU\) on 2016-07-12](#)
- 2 1% match (student papers from 11-Apr-2017)  
[Submitted to Jawaharlal Nehru University \(JNU\) on 2017-04-11](#)
- 3 1% match (student papers from 12-Jul-2016)  
[Submitted to Jawaharlal Nehru University \(JNU\) on 2016-07-12](#)
- 4 < 1% match (Internet from 22-Mar-2018)  
<https://www.frontiersin.org/articles/10.3389/fpls.2016.00042/full>
- 5 < 1% match (publications)  
["Begomoviruses: Occurrence and Management in Asia and Africa", Springer Nature, 2017](#)
- 6 < 1% match (publications)  
[Nirbhay Kumar Kushwaha, Mansi Bhardwaj, Supriya Chakraborty. "The replication initiator protein of a geminivirus interacts with host monoubiquitination machinery and stimulates transcription of the viral genome", PLOS Pathogens, 2017](#)
- 7 < 1% match (Internet from 30-May-2018)  
<http://mbio.asm.org/content/9/2/e00419-18.full>
- 8 < 1% match (student papers from 25-Apr-2017)  
[Submitted to Jawaharlal Nehru University \(JNU\) on 2017-04-25](#)