[*Pearl * et al., 4(7): July, 2017*]

ISSN: 2349-5197
Impact Factor: 2.715

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

INSILICO DESIGN OF SMALL INTERFERENCE RNA (SIRNA) FOR PREVENTION OF VIRAL (BEGOMOVIRUS) DISEASES IN COMMERCIALLY VALUABLE PLANTS (TOMATO AND PULSES)

Madhu Pearl R*, Gopal Pandi & Sundara Baalaji Narayanan

*Laboratory of Structural Biology, Department of Bioinformatics, Bharathiar University, Coimbatore-641046. INDIA

Department of Plant Biotechnology, School of Biotechnology, Madurai Kamaraj University, Madurai- 625 021. INDIA

DOI: 10.5281/zenodo.831442

Keywords: Begomovirus, Gene Silencing, Monopartite Genome, Bipartite Genome, siRNA.

Abstract

Yellow vein mosaic or yellow leaf curl disease in tomato, okra and other pulses is transmitted by Bemisia tabaci (White fly), leading to loss of production and quality. Gene silencing is one of the powerful approaches to overcome viral based plant diseases. However it is of great challenge to design potential small interfering RNA (siRNA) molecules against genes responsible for the diseases. From the primary sequence information, using insilico approach, 15 siRNAs were designed for genes involved in pathogenicity like suppressor, movement and nuclear shuttle namely V2, C2 and C4 of monopartite genome and BC1 and BV1 of bipartite genome from the Geminiviridae family. Short listing based on ranking score, GC content, base pairs, secondary structure identified 2 siRNAs for V2, 1 siRNAs for C2 and 2 siRNAs for C4, 5 siRNAs for BV1 and 5 siRNAs for BC1 genes.

Introduction

Begomovirus infect a wide variety of vegetable crops and legumes [1]. *Bemisia tabaci* (White fly) act as a vector to transmit their viral genes to dicot plants and cause leaf curl diseases affecting crop yields. Begomoviruses have monopartite or bipartite genome and characteristic icosahedral geminate particles that encapsidate genome of circular single stranded DNA. The bipartite begomovirus genome is composed of two similarly sized DNA molecules that share no sequence similarity, except for a highly conserved (>90% identity) common region (CR) of ~200 nucleotide (nt) bases [3]. The CR region contains essential sequences for replication and transcription of both components, thereby maintaining the fidelity of the bipartite genome [4]. In bipartite begomoviruses, DNA-A component contains four open reading frames (ORFs), encoding proteins required for replication, transcription and encapsidation whereas DNA-B has two ORFs, encoding proteins involved in viral movement and symptom development [2]. Thus, both DNA components are essential for the establishment of an efficient systemic infection. As each component is individually encapsidated, distinct virions carrying the DNA-A or DNA-B molecule must both be delivered into a target cell to initiate a systemic, wild type infection. The capsid protein (CP) of the bipartite begomoviruses is not essential for wild-type systemic infection, indicating that a non-virion form can effectively move within the plant. However, this CP independent movement is host specific, primarily occurring in species to which the virus is well adapted [5]. Expression of the viral-sense genes, *CP* and *BV1*, is transactivated by TrAP in a non-virus-specific manner. The *AC2* encoded transcriptional activator protein (TrAP) is a nuclear protein with a basic N-terminus containing a nuclear localization signal (NLS), a central zinc finger domain and a C-terminal acidic domain [6]. Furthermore, TrAP appears to play a role in tissue specific expression of the CP promoter, in that it overrides the putative host repressor in phloem tissues [7]. Finally, TrAP also has been shown to act as a suppressor of gene silencing [8].

The monopartite begomovirus genome is homologous to the DNA-A component of bipartite members, likely reflecting its role as the progenitor. The complementary sense genes *C1*, *C2*, and *C3* are homologous of *AC1*, *AC2* and *AC3* respectively. An additional viral-sense ORF, *V1* acts as a pathogenicity determinant, possibly via the regulation of ss-DNA and ds-DNA levels [9]. *V1* gene enhances the CP mediated nuclear export of viral

DNA and exhibited limited capacity for cell-to-cell trafficking through mesophyll plasmodesmata (PD) [10]. The *V1* ORF encodes the multifunctional CP involved in systemic infection and insect transmission [11]. The accumulation of the CP within the nucleolus may well reflect compartmentalization that underlies the events involved in replication, encapsidation and export of viral nucleoprotein complex [12]. While *C2* functions in the suppression of gene silencing, *C4* may induce host cell proliferation. Expression of *C4* alone dramatically alters the pattern of host gene expression. Both the *C2* and *C4* genes function as pathogenicity determinants [13]. While in bipartite begomoviruses the DNA-B component contains two ORFs encoding BV1, the nuclear shuttle protein (NSP) and BC1, the cell-to-cell movement protein. It is well established that BV1 and BC1 coordinate the movement of viral DNA across nuclear and plasmadesmal boundaries respectively [14]. This functional importance has rationalized our study to design siRNAs against *V2, C2, C4* genes from monopartite genome and *BC1, BV1* genes from bipartite genome of begomoviruses.

Methodology

Sequence Retrieval
The sequence of the genes namely *V2, C2* and *C4; BC1* and *BV1* of Tomato yellow leaf *curl virus*, *mungbean yellow vein mosaic virus*, *bhendi yellow vein mosaic virus* were retrieved from NCBI (http://ncbi.nlm.nih.gov//) in FASTA format. The sequence accession number, Gene ID and length were noted.

siRNA Design
siRNAs were designed and selected for genes of interest using WHITEHEAD selection program [15]. This tool shows all the possible siRNAs for each gene with GC content and thermodynamic values.

Secondary Structure Prediction
GENEBEE (http://www.genebee.msu.su) was handled for the secondary structure prediction of designed siRNA. Using GOR IV (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html) secondary structure of the protein was predicted for the retrieved DNA sequences.

Results And Discussion

Sequence Retrieval
The sequences of the target genes *V2, C2* and *C4; BC1* and *BV1* of begomovirus were extracted from the NCBI database (Table 1). The sequences of the *AC5* and *AV1* gene of BYVMV were also retrieved to serve as a control via natural gene silencing pathway (Table 2, Figure 1, 2). Complementary base pairing of *AC5* and *AVI* genes are clearly seen, and this was proved by our manual screening of these gene sequences. Figure 2 shows the complementary base pairing between 348 - 705 nucleotide of V1 gene with the whole AC5 gene on antisense strand, hence there may be a chance of natural silencing of AC5 by AV1.

siRNA Designing For *Begomovirus*
siRNAs were designed for *V2, C2* and *C4*; BC1 and BV1 from *Begomovirus* genome using whitehead siRNA selection program. The result for V2 gene is discussed and the same has been obtained for other genes namely C2, C4, BC1, and BV1. From the 29 siRNA designed for *V2* gene using whitehead siRNA selection program, 19 siRNA were short listed depending up on the value of %GC and the thermodynamic values. %GC should be 32 - 50%. The thermodynamic values should be of negative value. Selection of siRNAs with positive energy difference is not recommended in our criteria. Because there could be a threshold beyond which a siRNA is unusable because of super stable end. A siRNA with one stable end and one super stable end might not be as good as one with one end much less stable and the other is a stable. These two criteria were used to filter the potential siRNA outputs from selection program. In the second step ranking of siRNAs depending upon the following nucleotide positional scoring [16] was carried out:

1. Highly functional siRNAs have G/C content between 30 and 52%
2. At least 3 A/U bases at positions 15 to 19
3. Absence of internal repeats
4. The presence of an A base at position 19 of the sense strand

5. The presence of an A base at position 3 of the sense strand
6. The presence of a U base at position 10 of the sense strand
7. A base other than G or C at position 19 of the sense strand
8. A base other than G at position 13 of the sense strand;

Score one was given for satisfying each condition by the 19 siRNAs shortlisted from step 1. For each condition the ranking score were calculated, no siRNAs were found with the maximum score of 8. siRNAs with the score of 4-6 is given in Table 3. The region of siRNA should fall on the coding region where the protein should not posses regular secondary structure [15, 16]. Thus designed siRNA will get attached to the complementary region of mRNA and block the functional activity of the gene. siRNA positions were manually identified in the predicted secondary structure of V2 (Figure 3). Six siRNAs based on ranking scores lie in the protein secondary structure region. siRNA without proper secondary structure is the potential one for silencing [15]. The essential structural features of potential and non potential siRNA are taken from the secondary structure of six siRNAs predicted using GENEBEE tool, two were selected for further characterization.

mRNA Complementarity

The silencing action will be potential when anti-sense siRNA is binding with respective mRNA sequence. So the mRNA sequence was extracted for V2 and from the results, antisense_siRNAs were selected for the verification of siRNA binding affinity. The example for nucleotide complementarity between antisense siRNA and the mRNA region is given in figure5. Thus 2 potential siRNAs were designed and filtered with various criteria at gene and protein level against *V2* gene from the 29 siRNAs. Positions of two selected siRNAs were noted as 4-26 and 431-453. The designed siRNAs against other four genes are give in table 5.This strategy of designing potential siRNAs for the required genes will pave way for the development of reliable tools for gene therapy using custom designed functional siRNAs.

Conclusion

One of the advanced methods for treating viral diseases is silencing the RNA transcription by designing small interference RNAs (siRNAs). Potential siRNA has been identified against each investigated genes from begomovirus family. These potential siRNA satisfies necessary structural criteria. Further experimental evidences are needed to confirm siRNA-based treatments against *Begomoviruses* through gene silencing. These computational siRNA design strategies markedly reduce the costs of reagents and labor in early stage pharmaceutical research.

Table 1: List of retrieved begomoviral genes with accession number, ID, Total range in the genome, Length of the gene and the representing strand in monopartite and bipartite genome.

Gene	Virus	Acc No	Gene ID	Range (nt)	Length (bp)	Strand
V2	TYLCV(Malaga)	NC_004569.1	1482928	309-1082	774	Sense
C2	TYLCV(Almeri)	NC_004005.1	949137	1233-1640	408	Antisense
C4	TYLCV(Almeri)	NC_004005.1	949140	2178-2471	294	Antisense
BV1	MYMV	NC_001984.1	991128	538-1308	771	Sense
BC1	MYMV	NC_001984.1	991129	1334-2230	897	Anti sense

Table 2: List of retrieved begomoviral genes naturally silenced with accession number, ID and its range in the genome, Length of the gene and representing strand in BYVMV.

Gene	Virus	Acc No	Gene ID	Range (nt)	Length (bp)	Strand
AC5	BYVMV	NC_003418.1	935146	628-984	357	Anti sense
AV1	BYVMV	NC_003418.1	935148	280-1050	771	Sense

Table 3: List of potential siRNA sequences for V2 after first filtration with %GC,the thermodynamic values and nucleotide scores as observed from whitehead.

No	Position (length=774bp)	siRNA	GC%	Thermodynamic values	score
1	4-26	UU CUUCGCUUGUCCGCUAUAU	47	-4.5(-10.0,-5.5)	5
2	9-31	UU CUUGUCCGCUAUAUGAUUA	37	-3.3 (-8.8, -5.5)	2
3	119-141	UU GCAGCUAGUACCUGAAUGU	47	-2.7 (-9.9, -7.2)	3
4	156-178	UU GCGCCUACAUGUCUUACAU	47	-4.6 (-12.0, -7.4)	3
5	158-180	UU GCCUACAUGUCUUACAUGU	42	-1.2 (-9.7, -8.5)	6
6	206-228	UU CCAGGAACAUUUCAGGUCA	47	-2.5 (-12.3, -9.8)	1
7	266-288	UU GCAACACAAUCACUACAUU	37	-1.0 (-8.2, -7.2)	3
8	289-311	UU CCCAAGACCAUAAUGUGUA	42	-3.2 (-10.9, -7.7)	3
9	290-312	UU CCAAGACCAUAAUGUGUAU	37	-2.6 (-10.0, -7.4)	5
10	431-453	UU GGCAUACCUUGAUCAGGAU	47	-0.2 (-9.8, -9.6)	4
11	519-541	UU CCCUAUCCAUGGUUCAUUA	42	-5.5 (-11.2, -5.7)	2
12	525-547	UU CCAUGGUUCAUUACUCCUU	42	-1.1 (-10.3, -9.2)	5
13	621-643	UU GGGUACAUCAAAUAUUAGU	32	-3.2 (-10.3, -7.1)	6
14	644-666	UU CUCGUCCGCUUCAUACUUU	47	-4.3 (-11.2, -6.9)	4
15	671-693	UU CUCUUACGCAACAACAACA	42	-1.3 (-8.9, -7.6)	3
16	682-704	UU CAACAACAUAUACCGUACA	37	-1.0 (-8.7, -7.7)	2
17	694-716	UU CCGUACAUGAGUACGAAGA	47	-3.6 (-11.1, -7.5)	3
18	696-718	UU GUACAUGAGUACGAAGAUU	37	-0.9 (-8.2, -7.3)	3
19	704-726	UU GUACGAAGAUUGGGUCACA	47	-0.5 (-9.3, -8.8)	1

Table 4: Potential siRNA selection based on secondary structure (Selected potential siRNA highlighted as bold) in V2

No	POSITION	ANTI SENSE -siRNA	SENSE -siRNA	PROPER 2^0 STRUCTURE
1	158-180	GCCUACAUGUCUUACAUGU	CGGAUGUACAGAAUGUACA	YES
2	621-643	GGGUACAUCAAAUAUUAGU	CCCAUGUAGUUUAUAAUCA	-
3	4-26	**CUUCGCUUGUCCGCUAUAU**	**GAAGCGAACAGGCGAUAUA**	**NO**
4	290-312	CCAAGACCAUAAUGUGUAU	GGUUCUGGUAUUACACAUA	-
5	431-453	**GGCAUACCUUGAUCAGGAU**	**CCGUAUGGAACUAGUCCUA**	**NO**
6	644-666	CUCGUCCGCUUCAUACUUU	GAGCAGGCGAAGUAUGAAA	-

ISSN: 2349-5197
Impact Factor: 2.715
INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Table 5: List of short listed siRNA after four steps of filtration against begomovirus family.

GENES	DESIGNED SiRNA	LOWEST THERMODYNAMIC & CORRECT GC%	SPECIAL CHARACTERISTICS	AFTER PREDICTING 2°STRUCTURE
V2	29	19	6	2
C2	21	9	6	1
C4	17	9	3	2
BV1	69	46	18	5
BC1	73	52	19	5

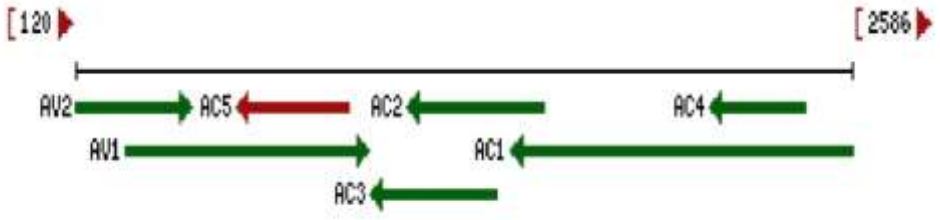

Figure 1: Pictorial representation of complementary genes AV1 and AC5 involved in natural gene silencing.

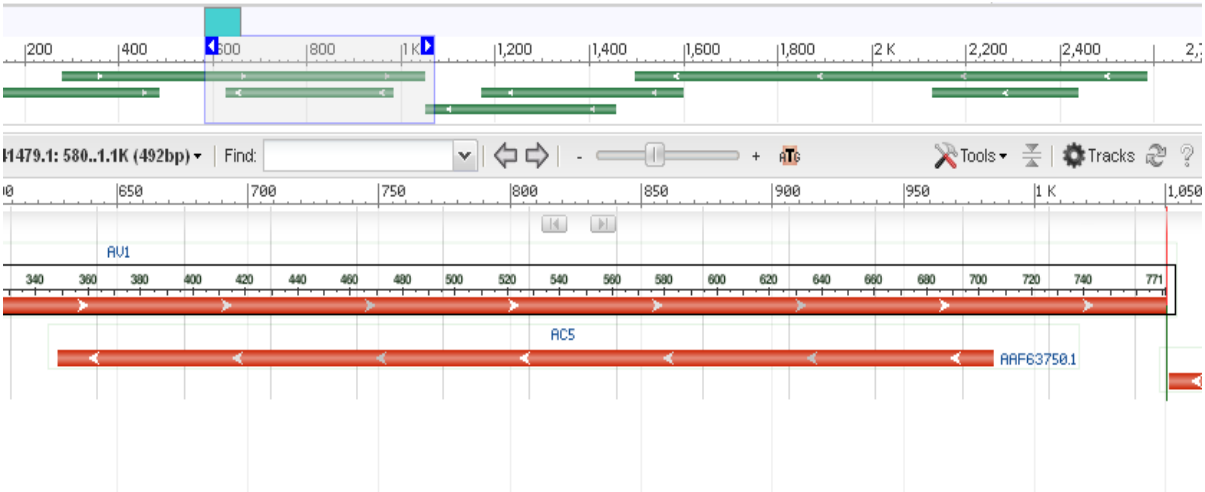

Figure 2: Degradation of AC5 gene by the sense strand of AV1gene. Natural Complementarity observed between the 348 – 705 nucleotide position of AV1 and full length of AC5.

```
MPKRTGDILISTPVSKVRRKLNFDSPYTSRAAAPTVQGIKRRSWTYRPMYRKPRMYRMYRSPDVPFGCEG
Ccccccceeecccchhhheeeccccccccccccceecceecccccccceecccceeeeeecccceeeccccc

PCKVQSYEQRDDVKHTGVVRCVSDVTRGSGITHRVGKRFCIKSIYILGKIWMDENIKKQNHTNQVIFFLV
Ccccccccccccceecceeeeeecccccccceeeecccccceeeeeeeccchhhhhccccccccccceeeeee

RDRRPYGTSPMDFGQVFNMFDNEPSTATVKNDLRDRYQVMRKFHATVVGGPSGMKEQCLLKRFFKVNTHV
Ccccccccccccccceeecccccccccceecccchhhhhhhhheeeeeeecccccchhhhhhheeecccceee

VYNHQEQAKYENHTENALLLYMACTHASNPVYATLKIRIYFYDAVTN
Eecchhhhhhhhchhhhhhhhhhccccccccccceeeeeeeeceeec
```

Figure 3: Positions of six siRNAs in the predicted protein secondary structure of V2 shown in cyan colour

Targeted region (cDNA): 5'-AACTGGACTTCCAGAAGAACATC-3'
mRNA region 5'- AACUGGACUUCCAGAAGAACAUC-3'

siRNA

Sense siRNA: 5'- CUGGACUUCCAGAAGAACA dTdT-3'
Antisense siRNA: 3'-dTdT GACCUGAAGGUCUUCUUGU-5'

Interaction

antisense siRNA: 3'-dTdT GACCUGAAGGUCUUCUUGU-5'
mRNA region 5'- AA CUGGACUUCCAGAAGAACAUC-3'

Figure 4 : Complementarity between antisense siRNA and mRNA region.

Acknowledgement

Jebastin for assistance during MS preparation.

References

1. Bradeen J.M, Timmermans M.C.P and Messing J, Dynamic genome organization and gene evolution by positive selection in geminivirus (*Geminiviridae*). Mol. Biol. **14**, 1114–1124 (1997)
2. Dong X, vanWezel R, Stanley J and Hong Y 2, Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. J. Virol. **77**, 7026-7033 (2003)
3. Baulcombe, D. C , Gene silencing: RNA makes no protein. Curr Biol. **9**, 599–601(1999)
4. Kunik T, Salmon R, Zamir D, Navot N, Zeidan M, Michelson I, Gafni Y and Czosnek H,Tansgenic tomato plants expressing the tomato yellow leaf curl virus capsid protein are resistant to the virus. Biotechnol. **12**, 500- 504 (1994)
5. Azzam, O., Frazer, J., De La Rosa, D., Beaver, J. S., Ahlquist, P. and Maxwell, D. P, Whitefly transmission and efficient ssDNA accumulation of bean golden mosaic geminivirus require functional coat protein. Virology. **204**, 289–296 (1994)
6. Stanley, J., Frischmuth, T. and Ellwood, S,Defective viral DNA ameliorates symptoms of geminivirus infection in transgenic plants. Proc. Natl. Acad. Sci. **87**, 6291-6295(1990)
7. Sunter G and Bisaro DM,Transactivation of geminivirus *AR1 and BR1* gene expression by the viral *AL2* gene product occurs at the level of transcription. Plant Cell. **4**, 1321-31(1992)
8. Voinnet O, Pinto YM and Baulcombe DC, Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proc. Natl. Acad. Sci. **96**, 14147-14152(1999)
9. Wartig L, Kheyr-Pour A, Noris E, De Kouchkovsky F, Jouanneau F, et al, Genetic analysis of the monopartite tomato yellow leaf curl geminivirus: roles of *V1, V2*, and *C2* ORFs in viralpathogenesis. Virology. **228**, 132-140(1997)
10. Wezel R.V, Tien P, Stanley J, and Hong Y *et al*, Gene *C2* of the Monopartite Geminivirus *Tomato yellow leaf curl virus*- China Encodesa Pathogenicity Determinant That Is Localized in the Nucleus. Mol Plant Microbe Interact. **9**, 1125-1128 (2001)
11. Rojas MR, Jiang H, Salati R, Xoconostle- Cazares B, Sudarshana MR, *et al*, Functional analysis of proteins involved in movement of the monopartite begomovirus, *Tomato yellow leaf curl virus.* Virology. **291**, 110-125 (2001)
12. Rigden JE, Dry IB, Mullineaux PM and Rezaian MA, Mutagenesis of the virion- sense open reading frames of tomato leaf curl *geminivirus* . Virology. **193**, 1001- 1005 (1993)
13. Selth LA, Randles J and Rezaian A,Host responses to transient expression of individual genes encoded by *Tomato leaf curl virus.* Mol. Plant-Microbe Interaction. **17** 27-33(2004)
14. Gafni Y and Epel BL,The role of host and viral proteins in intra and inter-cellular trafficking of geminiviruses. Physiol Mol Plant Path. **60**, 231-241(2002)
15. Choudhary M, and Sahi S, *Insilico* designing of insecticidal small interfering RNA (siRNA) for Helicoverpa armigera control. Indian J Exp Biol. **49**, 469-474(2011)
Reynolds A, Leake D, Boese Q, Scaringe S, William S Marshall and Khvorova A Rational siRNA design for RNA interference. Nature Biotechnol. **22**, 326-330 (2004)