In silico plum pox virus silencing via host-retrieved miRNAs in peach plant

Muhammad Naveed Shahid1,*, Syeda Shehnaz1, Muhammad Shehzad Iqbal2,5, Aneela Shabbeer1, Adil Jamal3 and Sana Khalid4

1Department of Botany, Division of Science and Technology, University of Education, Lahore, Pakistan
2Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan
3Science and Research College of Nursing, Ummal Qura University, Makkah, Kingdom of Saudi Arabia
4Department of Botany, Lahore College for Women University, Lahore, Pakistan
5Department of Biotechnology, Faculty of Life Sciences, University of Central Punjab, Lahore, Pakistan

Peach (Prunus persica) is a deciduous, edible, stone fruit producing plant, belonging to the family Rosaceae. The plant is prone to various pathogens and one of them is the plum pox virus (PPV). This is a lethal virus of peach plant causing pox disease of plum. Its attack results in 83%–100% yield loss in highly susceptible varieties of peach plant. The complete genome of PPV is 9791 base pairs with positive-sense single strand. The full-length genome of PPV encodes a large poly-protein initially, which cleaves proteolytically into ten mature proteins – coat protein, helper component proteinase, P1, P3, viral genome-linked protein, 6K1, 6K2, cylindrical inclusion protein, cylindrical inclusion protein b and NIa-pro-proteins. The objective of this study is to identify such sites in the PPV genome which can be targeted by PPV-derived miRNAs through target prediction computational tools/algorithms. A total of 214 mature miRNAs were chosen from the miRNA database to check their complementarity with the PPV genome. Minimum free energy, folding energy, seed pairing, target-site accessibility, pattern recognition and multiple target sites were the parameters considered for target prediction algorithms. Two out of 214 miRNAs were predicted as potential against plum pox virus by three of four tools used for target prediction. Thus, the results encourage generating transgenic, PPV-resistant peach plants by expression of predicted miRNAs.

Keywords: miRNAs, peach plant, plum pox virus, yield loss, target prediction algorithms.

Peach (Prunus persica) is a temperate, deciduous, edible, stone fruit-producing plant. Prunus is an important genus in the family Rosaceae which comprises of different fruit varieties1. Peach is an important commercial and agronomical plant as it provides vitamins, fibres, antioxidants and minerals for a healthy diet (http://faostat.fao.org/). The genus Prunus is infected by various viral attacks and one of the most devastating among them is the plum pox disease (PPV), also known as pox disease of plum. This disease is transmitted in stone fruits through PPV, a member of the family Potyviridae2. PPV is a single-stranded, positive-sense, filamentous virus. Its genome size is estimated as 9.9 kb and is 750 nm long3. Its transmission results in premature fruit drop and decreased fruit quality, resulting in agronomic and economic loss4. The virus attack also results in 83%–100% yield loss in peach production5,6.

Numerous strategies have been developed to control the multiplication of the virus in peach plants. Breeding programmes and genetic engineering approaches have also been used to develop resistance against PPV. These strategies do not provide complete information for resistance due to the specific nature of virus strain, long juvenile seedling period5 and unnecessary degradation of RNA-silencing pathways7–9. Computational strategy is another technique to control this disease. Plants respond to distinctive abiotic stress and viral infection. These are controlled both at the transcriptional and post-transcriptional levels. At the post-transcriptional level, quality controller miRNAs are utilized10. miRNAs are small, 21–22 nucleotides, endogenous and non-coding RNAs that regulate gene expression in plants and animals through their proteolytical activities. They bind to their complementary sequences in target regions of the viral genome and repress the transcription or cleave the transcript11. Peach plant has 180 mature miRNAs and 214 precursors (miRNA database). To identify miRNAs in peach plants which have the potential to inhibit infection through mRNA cleavage, five target prediction algorithms were used.

The aim of the present study is to identify the best host-retrieved potential miRNAs that can develop resistance against PPV using computational techniques. The predicted miRNAs can be a source to inhibit infection by cleaving the mRNA.

Materials and methods

Mature miRNAs of Prunus persica

In many living organisms miRNAs are endogenous, non-coding and small sequences of RNA that regulate gene
expression. To search for potential miRNAs, mature miRNAs were obtained from miRbase (http://www.mirbase.org/cgi-bin/browse.pl) and 214 mature miRNAs of P. persica were retrieved.

**PPV genome**

The complete PPV genome was downloaded from NCBI in FASTA format (https://www.Ncbi.nlm.nih.gov/nuccore) with accession number AY028309.2. It contains 9791 base pairs.

**Target prediction tools**

A number of programs and algorithms have been introduced to search for target genes of miRNAs. Four algorithms were used for target prediction of miRNAs in the PPV genome, to analyse their accuracy and efficiency. In the virus genome, these algorithms predict the target sites. The most important miRNAs which blocked translation or cleaved the mRNA to prevent protein synthesis were selected using these four algorithms. The four algorithms chosen were miRanda, psRNATarget, RNA22 and RNAhybrid.

**miRanda:** For target prediction of plants and animals, miRanda is frequently used. To predict potential targets in viral genome, miRanda (V3.3a) was downloaded from the web source (http://cbio.mskcc.org/mirRNA2003/miranda.html). The algorithm was run after setting parameters with changing energy threshold. Different energy levels (E-15, E-20, E-22 and E-25 Kcal mol) were used to select the most suitable one. The parameters set for target prediction were gap open penalty = –9.0, gap extent penalty = –4.0, score threshold = 140, scaling parameter = 4.0 and energy threshold = –20.

**psRNATarget:** This software is used for target analysis, specifically for identification of target transcript of miRNAs using proven scoring schema (V1 and V2), analysing the reverse complementary matching between miRNA and target, and target site accessibility. V2 released in 2017, is an improved scoring schema which is capable of identifying more miRNA target pairs without an increase in the final output. This algorithm is available at http://plantgrn.noble.org/psRNATarget/. All parameters were set as by default then submitted and got the miRNAs which were more potential to target.

**RNA22:** This is a pattern-based algorithm used for target prediction. It is utilized in the identification of target regions of the mRNA sequence that have a higher likelihood to contain miRNA binding sites. RNA22 V2.0 can be accessed from https://cm.jefferson.edu/rna22/Interactive/. The 50 miRNAs sequences were uploaded to the RNA22 in FASTA format. The sensitivity and specificity values of RNA22 were set as 63% and 61% respectively. In seed/nucleus region seed size of 7 was selected to allow a maximum of one unpaired base, while the minimum number of paired bases in the heteroduplex was set at 12. Maximum number of G:U wobbles allowed in the seed region was set as ‘no limit’, while maximum folding energy was selected as –20 kcal/mol. Using these parameters, potential miRNAs which can target the genome of the virus (target sites) were predicted using the position of the targets, folding energy and P-values.

**RNAhybrid:** This is a tool used for easy and fast miRNA target prediction. It predicts the RNA secondary structure. RNAhybrid provides many useful parameters. The first parameter is hits per target and MFE (minimum free energy). This algorithm is available at http://bibiserv.techfak.unibielefeld.de/rnahybrid (ref. 18). It was used to eliminate all possible false-positive attachments shown by miRanda. The E-value was set –20 kcal/mol and other parameters were set to default.

**PPV gene retrieval and annotation**

Retrieved gene names from https://www.ncbi.nlm.nih.gov/nuccore/9626508 and CLC genomic workbench (version 11.0.1) were used for annotation of the genomic feature of PPV.

**Statistical analysis**

For statistical analysis, R studio and R language were used. R studio helps run the R language and in R studio readxl package was fed. Results obtained from four softwares about miRNAs were uploaded into the readxl in excel sheet and used ggplot 2 package. It gave the results in the form of a graph.

**miRNA genome binding site conservation analysis**

MEGAX was run for alignment purpose. The complete PPV genome sequences were retrieved from the NCBI nucleotide database with accession numbers AY028309.2, MF346290.1, MF346289.1, MF346288.1, MF346287.1, MF346286.1 and MF346285.1. ClustalW algorithm was used for alignment of the PPV genome with miRNAs. Viral genome and putative miRNA sequences were analysed using MEGAX to check whether the predicted miRNAs block other genomes as well.

**Results and discussion**

Plum pox disease commonly called Sharka disease is caused by PPV infection and can have adverse effects on...
the yield of peach plants. Various reports of peach infection with PPV are available around the world, especially in Europe resulting in severe economic loss. RNA interference (RNAi) is considered as a defence mechanism against foreign organisms and it regulates gene expression. miRNAs are small RNAi containing 21–24 nucleotides and are reported to enhance resistance in plants. miRNAs were reported to block the specific genes of murine cytomegalovirus (MCMV) and potato virus Y (PVY). This makes the plant transgenic and resistant to diseases caused by the viruses. 

miRNA target prediction in the PPV genome

Four algorithms were used for the prediction of miRNA target in the PPV genome. These algorithms helped eliminate the false-positive results and maximize the accuracy of target prediction in the viral genome. The parameters considered in miRANDA were score threshold, scaling parameter and energy threshold. psRNA-Target was specifically used to recognize mismatch-sensitive seed regions. RNAhybrid is an algorithm used for the secondary structure prediction of RNA. Parameters considered in this algorithm were hits per target and energy threshold. Figure 1 demonstrates all the genome positions targeted by P. persica miRNA using different calculations.

The full-length genome of PPV encoded into a large polyprotein. This polyprotein cleaved into ten proteins, viz. coat protein (CP), HC-Pro protein, P1, P3, VPg, 6K1, 6K2, CI, NIb and Nla-Pro-proteins (Figure 2). miRNAs of P. persica targeted these proteins at the different sites and blocked synthesis at the translational level.

P1 protein

This is the first protein of PPV. It involves in non-proteolytic functions like cell-to-cell transportation and viral infection. Suitable miRNAs for targeting the P1 gene have been predicted to be miR395a-3p, miR6257, miR6270, miR530, miR6274a, miR6263, miR395b-3p, miR858, miR169d, miR169e-5p, miR8131-3p and miR171e.

HC-pro protein: This is a multifunctional protein of PPV and is also called helper component protease. It is involved in aphid transmission, genome amplification and long-distance movement. It is also a plant viral suppressor of RNA silencing. HC-pro is considered to intercede the virus replication cycle at different steps. The HC-pro gene shows interaction with 22 miRNAs (miR6273, miR167c, miR156e, miR156c, miR482a-5p, 6282, miR530, miR8126-5p, miR482d-5p, miR6297b, miR403, miR6267b, miR164d, miR6288a, miR6291c, miR6275, miR6297a, miR393b, miR8133-3p, miR156d, miR6262 and miR6277).

CI protein: This protein is involved in virus replication and cell-to-cell movement. It forms pinwheel-shaped cytoplasmic inclusion bodies that are a unique feature of potyvirus infection. The maximum number of potential targets of P. persica miRNAs was for the C1 gene that predicts 57 targeting miRNAs (miR399h, miR160a, miR171g, irR8130-3p, miR399g, miR6273, miR2111a, miR399i, miR1710, miR2111d, miR399e, miR156g, miR390, miR393a, miR398b, miR8123-3p, miR6267c-5p, miR398a-3p, miR8129-5p, miR6260, miR8124-5p, miR482a-5p, miR8126-3p, miR6261, miR8130-5p, miR3627-3p, miR399n, miR2111c, miR397, miR156i, miR477-3p, miR7122a 5p, miR6284, miR3991, miR168, miR399d, miR156h, miR156a, miR399c, miR8128-3p, miR399i, miR171d-3p, miR6285, miR858, miR477a-3p, miR171a, miR8286, miR8127-3p, miR172c, miR399m, miR482c-5p, miR2111b, miR398a-5p, miR399k, miR156b and miR482b-5p).

6K1 and 6K2 protein: These are the two smallest proteins of PPV and both have the same molecular weight. The role of 6K1 is uncertain, but a small deletion in the sequence of 6K1 prevents viral replication. 6K2 is an integral membrane protein which induces the endoplasmic reticulum (ER)-originated replication vesicles that target the chloroplast for robust viral replication. These two proteins are targeted by the least number of miRNAs. For the 6K1 gene of PPV, ten targeting miRNAs were predicted (miR160a, miR172b, miR6292, miR160b, miR171h, miR8123-5p, miR172a-3p, miR828-3p, 172c and 172d). The 6K2 gene is targeted by six miRNAs (miR8130-3p, miR319b, miR482a-5p, miR5225-3p, miR482f and miR6277).

P3 protein: Little is known about the P3 protein. It is a membrane protein and localized in the ER membrane. The P3 protein gene is reported to prevent replication. It is targeted by more miRNAs after C1 and the CP. Forty-one miRNAs were predicted to target the P3 protein (miR169f, miR171g, miR169i, miR171c, miR395m, miR6288c-5p, miR5293, miR156e, miR482a-5p, miR167b, miR8124-3p, miR8130-5p, miR171h, miR6288b-5p, miR169h, miR169k, miR169g, miR171b, miR395a-5p, miR7125-5p, miR535a, miR482a-3p, miR6296, miR171f, miR8122-3p, miR6271, miR319a, miR6271, miR169i, miR6278, miR167d, miR167a, miR169d, miR169e-5p, miR171a, miR156d, miR6290, miR395b-5p, miR7121d-5p, miR3950 and miR8125).

Coat protein: Capsid protein encapsulated single-stranded RNA is encoded by CP. It is involved in cell-to-cell and long-distance movement, genome amplification and aphid transmission. Suitable miRNAs for targeting the CP gene have been predicted (miR399h, miR171g, miR8130-3p, miR399g, miR399i, miR171c, miR399e, miR169c, miR172b, miR482d-3p, miR482a-5p, miR396a-3p, miR395a-5p, miR171g, miR169i, miR6292, miR160b, miR171h, miR8123-5p, miR172a-3p, miR828-3p, 172c and 172d). The 6K2 gene is targeted by six miRNAs (miR8130-3p, miR319b, miR482a-5p, miR5225-3p, miR482f and miR6277).
Figure 1. Target prediction results of miRNAs against PPV genome indicated target prediction obtained from (a) miRanda, (b) RNA22, (c) psRNATarget and (d) TapirHybrid.

Figure 2. Genome annotation of plum pox virus (PPV). Ten genes of PPV are represented along with their positions. Arrows indicate the position of open reading frames (ORFs). Genes were translated from these positions. Long arrow represents polyprotein, cleaved into ten mature proteins.
miR395f, miR8124-3p, miR395d, miR3627-3p, miR399n, miR164a, miR169b, miR8133-5p, miR3951, miR3991, miR399d, miR159, miR399c, miR399i, miR6260, miR172a-3p, miR395g, miR482e, miR395b-3p, miR395n, miR399a, miR171f, miR6275, miR395i, miR164b, miR395i, miR172c, miR395e, miR395k, miR399m, miR395c, miR399k, miR172d and miR395h).

*Nib* protein: The functional role of Nib protein is in RNA-dependent RNA polymerase and nuclear translocation activities\(^4\). Using bioinformatics tools, 37 miRNAs were found to target the Nib gene (miR8122-5p, miR3627-5p, miR394b, miR396a, miR319b, miR390, miR1827, miR7125-3p, miR6276, miR6269, miR3627-3p, miR6257, miR477b-3p, miR1511-3p, miR47b-5p, miR396b, miR6260, miR6266a, miR482e, miR395b-3p, miR6294, miR6280, miR6285, miR477a-5p, miR169d, miR477-5p, miR394a, miR8127-3p, miR6266c, miR6266b, miR6264, miR162, miR8129-3p, miR6274b-3p, miR6289, miR535b, miR399b and miR6267a).

*Nla-pro* protein: Nla-pro protein is necessary for proteolytic maturation of many viral proteins\(^4\). Bioinformatics tools were used to predict the miRNAs which target the Nla pro gene. The least number (17) of miRNAs was found to this gene (miR6287, miR3944b, miR169c, miR828-5p, miR8127-5p, miR169b, miR395a-5p, miR169a, miR6274b-5p, miR6272, miR477b-5p, miR169k, miR8132, miR477a-5p, miR394a, miR1511-5p and miR3950).

*VPg* protein: This protein plays an important role in viral infection. It is a viral genome link protein involved
in the replication and translation of the viral genome\textsuperscript{36}. Vpg may also serve as a primer for viral RNA replication\textsuperscript{36}. Through target prediction algorithms, only one miRNA (miR482b-3p) was found to target the VPG gene.

\textit{P. persica} screened miRNAs

Two miRNAs, viz. miR8122-3p and miR7125-5p were found to have greater potential to target PPV at multiple loci. They were predicted by at least three algorithms. These screened miRNAs were the most suitable \textit{P. persica} to the boost defence system in peach plants against PPV.

\textbf{miRNA-genome binding site conservation analysis}

For the binding of miRNA to the genome of viral strains, MEGAX software was used\textsuperscript{24} (Figure 3). The screened miRNAs targeted at various positions in different strains of the virus revealed details of the level of conservation in the binding sites. The secondary structure of predicted miRNAs precursors has been shown in Figure 4.

\textbf{Conclusion}

This study helped us to recruit the most suitable miRNAs by using \textit{in silico} approach and the aim of finding these potential miRNAs was to block the genome of the virus which was affecting the host plant by RNA interference. Two miRNAs (miR8122-3p and miR7125-5p) out of 124 were predicted as potential against the infection of plum pox virus in a peach plant using four different tools. These potential miRNAs can be used to develop transgenic peach plants with their higher copy number to block the infection of this virus.

\textbf{Conflict of interest:} The authors declare no potential conflict of interest.


ACKNOWLEDGEMENT. We thank the University of Education, Township, Lahore for providing the necessary facilities to conduct this study.

Received 21 August 2019; revised accepted 9 March 2021

doi: 10.18520/cs/v121/i10/1316-1322