Identification and Molecular Characterization of a Novel Non-typical Carlavirus Infecting Rose Plants (*Rosa Chinensis* Jacq.)

Fei Xing

Chinese Academy of Agricultural Sciences Institute of Plant Protection

Dehang Gao

China Agricultural University

Nuredin Habili

Australian Wine Research Institute: The Australian Wine Research Institute Limited

Hongqing Wang

China Agricultural University

Zhixiang Zhang

Chinese Academy of Agricultural Sciences

Mengji Cao

Southwest University

Shifang Li (✉ SFLI@ippcaas.cn)

Chinese Academy of Agricultural Sciences Institute of Plant Protection <https://orcid.org/0000-0003-4614-1768>

Research Article

Keywords: rose plant, carlavirus, novel virus species, genome, molecular characteristics

Posted Date: June 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-498082/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

In the present study, systematic screening of public transcriptomic data followed by a shotgun assembly from *Rosa chinensis* Jacq. revealed a sequence of 8,332 nucleotides (nt) representing a potential novel virus, tentatively named rose virus C (RVC). The incidence of RVC in rose plants under our survey was low (5.4%, 5 out of 92). The complete nucleotide sequence of RVC was determined by Sanger sequencing. The genomic RNA of RVC consists of 8,386 nt, excluding the 3'-poly(A) tail, and contains five definite open reading frames (ORFs). BlastN analysis revealed that RVC had 67.3-71.3% nt identity with the carlaviruses with the maximum coverage of 19%. Phylogenetic analysis showed that RVC clustered with the carlaviruses and had a nt sequence identity of 48.3-50.0% based on its full-length genome. Replication protein had 48.7-51.0% nt sequence identity with other carlaviruses while the coat protein showed 39.5-45.6% nt identity being much below the species demarcation criteria of 72%. These all indicate that RVC is a distinct carlavirus. Overall, our data suggest that RVC is a novel non-typical virus species of the genus *Carlavirus*.

Background

Rose (*Rosa chinensis* Jacq.) is an important ornamental plant with a high popularity in cut-flower industry. Due to its vegetative propagation, mainly by grafting, rose plants have harbored many viruses over millennia. More than 24 viruses have been reported to infect rose plants worldwide, of these only six occur in China, including apple mosaic virus, arabis mosaic virus, apple stem grooving virus, blackberry chlorotic ringspot virus, prunus necrotic ringspot virus, and rose leaf rosette-associated virus (He et al. 2015; Diaz-Lara et al. 2020; Xing et al. 2020; Zhang et al. 2020).

Viruses in the genus *Carlavirus*, family *Betaflexiviridae* contain a single positive-sense linear RNA molecule with six open reading frames (ORFs). ORF1 encodes a polypeptide assumed to be the viral replicase. ORFs 2-4 form the triple gene block (TGB) proteins which allow cell-to-cell and long-distance movements. The coat protein (CP) is encoded by ORF5, while ORF6 encodes a cysteine-rich RNA-binding protein (RBP) which may facilitate aphid transmission or be involved in host gene transcription/ silencing and/or viral RNA replication (Adams et al. 2004; King et al. 2012). RNA-seq is now widely used to identify differentially expressed genes in higher plants which result in large volumes of data easily available in public database (Wang et al., 2009). So, the datasets can be mined to identify viruses/viroids using bioinformatics tools. In this paper, we identified a new carlavirus, rose virus C (RVC), in rose plants (*Rosa chinensis* Jacq.) using systematic BLAST-based data mining of public plant transcriptome data and Sanger sequencing. The complete genome sequence of RVC and its phylogenetic relationship are described.

To identify potential novel viruses/viroids, a systematic search for virus/viroid sequences in public plant transcriptome data was conducted using BLASTN analysis. The sequence of type member of each recognized virus/viroid species was used as query against all transcriptome shotgun assemblies (TSA) in GenBank. Using this method, a contig with 8332 nucleotides (nt) with distant homology to carlavirus was identified in the TSA from the RNA-Seq data of rose plants in China (Bioproject PRJNA546486).

Based on the contig sequences, two specific primer sets RVC-586F/-586R and RVC-699F/-699R were designed to confirm the presence of RVC under natural conditions. RVC-624R/-416R and RVC-7668F/-7938F were

designed for 5' RACE and 3' RACE, respectively, using the SMARTer RACE 5'/3' Kit (Clontech, USA) as outlined by the kit instructions. The specific primer sets RVC-F1/-R1 were used to amplify nearly full genome of RVC from the rose plant sample (Supplementary Fig. 1). All primers used in this study have been given in the Supplementary Table 1. The extracted RNA was treated with DNase I for 15 min at room temperature to remove DNA. The cDNAs were synthesized using the SMARTer RACE 5'/3' Kit. For the amplification of RVC genome, the expected PCR products were successfully amplified using Q5® High-Fidelity 2X Master Mix (NEB, USA), inserted into pTOPO vector using Zero Background pTOPO-Blunt Simple Vector Kit (Aidelab, China), and sequenced by Sangon Biotech (Shanghai) Co., Ltd.

In the present study, a total of 92 rose plants were analyzed, of which 5 tested positive for RVC. However, no association between virus-like symptoms in rose plants and RVC was found. Total nt length of the RVC genomic RNA isolated from the rose plant was 8,386, excluding the 3'-poly(A) tail (Fig. 1, accession no. MZ049538), showing 97.8% identical to the assembled contig. BlastN analysis revealed 67.3-71.3% nt identity of RVC with the carlaviruses with the highest query coverage of 19% in NCBI (<https://www.ncbi.nlm.nih.gov/>). The RVC genome contains five definite ORFs, namely replication protein, TGBs and CP, while no predicted RBP was found (Fig. 1). It appears that RVC is closer to a foveavirus or robigovirus than to the typical carlavirus which has six ORFs. The 5' and 3' untranslated regions (UTRs) have 57 nt and 94 nt, respectively (Fig. 1).

Phylogenetic analysis showed that the full genome of RVC clustered with the carlaviruses with 48.3-50.0% nt sequence identity as determined by MEGA version 6 (Tamura et al. 2013) using the neighbor-joining method (Saitou and Nei 1987) with 1000 bootstrap replications (Fig. 2). This was confirmed when the replication protein was analyzed (Supplementary Fig. 2). On the other hand, RVC may be considered as a new genus since it moved to a single group when the CP was used in the analysis (Supplementary Fig. 3). Here, the replication protein (6078 nt, 2025 amino acid [aa], 230.3 kDa) encoded by RVC shared 48.7-51.0% nt sequence identity with other carlaviruses, of which blueberry scorch virus showed the highest sequence identity (51%, accession no. AY941198) followed by ligustrum virus A (50.4%, accession no. MG865890) (Table 1). The replication protein contained the highly conserved motif sequence of glycine-aspartic acid-aspartic acid (GDD) at position 1900-1902 aa as in positive-strand RNA viruses. ORFs 2-4 encode TGB with proteins of 234 aa (26.0 kDa), 114 aa (12.3 kDa) and 78 aa (8.7 kDa), in tandem, which are involved in virus movement (Martelli et al., 2007). The CP (867 nt, 288 aa, 31.7 kDa) showed a relatively low nt sequence identity (39.5-45.6%,) with known carlaviruses, the highest sequence identity was with red clover carlavirus 1 (45.6%, accession no. MG596239) (Table 1). The sequence identity of the replication protein and CP, when compared to other members of carlaviruses, were much lower than the species demarcation criteria of 72% as set by the ninth report of the international committee on taxonomy of viruses (King et al., 2011). Overall, our data suggest that RVC is an atypical novel virus species of the genus *Carlavirus*.

Declarations

Acknowledgments

This work was supported by National Key R & D Program of China (2019YFD1001800) and key inter-governmental projects of the National Key Research and Development Program (2017YFE0110900).

Authors contributions

X.F. and L.S.F. designed the experiments and wrote the manuscript; X.F., G.D.H. performed the experiments; H.N. and W.H.Q. made comments and improved the manuscript. C.M.J conducted the systematic screen of public transcriptomic data. Funding was acquired by Z. Z.X. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest

All authors declare no conflict of interest.

Ethical approval

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

References

1. King AM, Lefkowitz E, Adams MJ et al (2011) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier, Amsterdam.
2. Adams MJ, Antoniw JF, Bar-Joseph M et al (2004) Virology Division News: the new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. Arch Virol 149:1045-1060.
3. Xing F, Gao DH, Liu H et al (2020) Molecular characterization and pathogenicity test of prunus necrotic ringspot virus isolates from China rose (*Rosa chinensis* Jacq.). Arch Virol 165: 2479-2486.
4. Zhang XQ, Liu SY, Yang YZ et al (2020) Virus identification and detection by high-throughput sequencing and RT-PCR from rose plants in Beijing. Acta Phytopathologica Sinica, doi:10.13926/j.cnki.apps. 000481.
5. He Y, Yang ZK, Hong N et al (2015) Deep sequencing reveals a novel closterovirus associated with wild rose leaf rosette disease. Mol Plant Pathol 16: 449-458.
6. Wang Z, Gerstein M, Snyder M et al (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63.
7. Saitou N, Nei M et al (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.
8. Tamura K, Stecher G, Peterson D et al (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725-2729.
9. Martelli GP, Adams MJ, Kreuze JF et al (2007) Family *Flexiviridae*: a case study in virion and genome plasticity. Ann Rev Phytopathol 45:153-157.
10. Diaz-Lara A, Mollov D, Golino D et al (2020) Complete genome sequence of rose virus A, the first carlavirus identified in rose. Arch Virol 165: 241-244.

Table

Table 1

Percentage of nucleotide sequence identities of coat protein (upper right corner) and replication protein (lower left corner) among RVC and other carlaviruses

	RVC	BSV	CLV	CVB	CMMV	GLV	LVA	PRDV	PVH	PVM	PVS	RCC1
RVC		41.5	44.0	39.5	42.8	44.2	40.0	42.8	41.5	43.0	42.8	45.6
BSV	51.0		47.8	51.4	51.3	48.1	53.5	67.2	53.2	52.8	65.1	46.9
CLV	49.1	48.9		47.4	50.1	49.1	48.9	48.4	46.9	48.6	50.2	51.4
CVB	49.9	50.9	48.8		51.4	45.7	61.4	53.5	52.0	51.7	52.8	42.9
CMMV	49.7	50.2	49.0	54.3		45.0	52.5	51.1	53.8	55.5	50.1	47.1
GLV	48.7	49.2	48.1	47.6	48.8		47.5	48.0	46.6	46.4	46.4	51.7
LVA	50.4	52.0	49.6	54.7	52.5	49.5		52.0	55.6	57.5	51.7	45.6
PRDV	49.6	57.7	48.4	51.8	50.3	48.7	51.1		56.9	53.5	65.4	47.0
PVH	50.2	52.2	49.4	57.2	53.5	48.7	54.1	51.6		57.3	58.1	47.2
PVM	49.7	51.1	48.3	52.7	51.4	46.9	53.0	50.4	52.6		54.0	45.8
PVS	49.2	56.4	48.1	50.7	49.8	48.3	51.0	55.5	51.4	51.9		47.5
RCC1	50.2	50.3	50.6	49.8	48.6	50.9	50.4	50.7	49.7	50.0	49.7	

BSV: blueberry scorch virus (accession no. AY941198), CLV: carnation latent virus (accession no. MN450069), CVB: chrysanthemum virus B (accession no. NC009087), CMMV: cowpea mild mottle virus (accession no. MN908944), GLV: garlic latent virus (accession no. MN059197), LVA: ligustrum virus A (accession no. MG865890), PRDV: potato rough dwarf virus (accession no. EU020009), PVH: potato virus H (accession no. NC018175), PVM: potato virus M (accession no. NC001361), PVS: potato virus S (accession no. MN689479), RCC1: red clover carlavirus 1 (accession no. MG596239).

Figures

Figure 1

Schematic representation of the genomic organization of rose virus C (RVC).

Figure 2

Phylogenetic relationship of the RVC with other members of the subfamily Quinvirinae, in family Betaflexiviridae. The tree was constructed based on the complete nucleotide sequences of each virus genome by the neighbor-joining method with 1000 bootstrap replicates. RVC sequenced in this study is highlighted by the black triangle (accession no. MZ049538). The branch values lower than 70% are hidden.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RosevirusC.fasta](#)
- [supplementarymaterials.docx](#)