Sequence Comparison and Phylogeny of Nucleotide Sequence of Coat Protein and Nucleic Acid Binding Protein of a Distinct Isolate of *Shallot virus X* from India

S. Majumder · V. K. Baranwal

Received: 30 December 2010 / Accepted: 27 May 2011 / Published online: 11 June 2011
© Indian Virological Society 2011

Abstract *Shallot virus X* (ShVX), a type species in the genus *Allexivirus* of the family *Alfalexiviridae* has been associated with shallot plants in India and other shallot growing countries like Russia, Germany, Netherland, and New Zealand. Coat protein (CP) and nucleic acid binding protein (NB) region of the virus was obtained by reverse transcriptase polymerase chain reaction from scales leaves of shallot bulbs. The partial cDNA contained two open reading frames encoding proteins of molecular weights of 28.66 and 14.18 kDa belonging to Flexi_CP super-family and viral NB super-family, respectively. The percent identity and phylogenetic analysis of amino acid sequences of CP and NB region of the virus associated with shallot indicated that it was a distinct isolate of ShVX.

Shallot, *Allium cepa* L. var *aggregatum*, is a culinary herb endowed with properties similar to onion. It is preferred over onion in international cuisine for its milder and sweeter taste. It is mainly cultivated in Europe, China, and Southeast Asia. In India it is mainly grown in Tamil Nadu and other southern states. *Shallot virus X* (ShVX), a member of genus *Allexivirus* and family *Alfalexiviridae*, occurs in Russia, Germany, Netherland, India, and New Zealand [5, 8, 10, 11]. The virus particles are ~800 nm long flexuous rod. Two fully annotated whole genome

sequences and one coat protein gene sequence of ShVX are available in GenBank. Presence of an *Allexivirus* in shallot in India was confirmed previously through electron microscopy and sequencing of partial NB gene and untranslated region (UTR) and was identified as ShVX [5]. Coat protein (CP) and nucleic acid binding protein (NB) genes are components of viral pathogenesis apparatus and percent sequence similarity of CP gene is also used for species demarcation in *Allexivirus* [1]. The present study was undertaken for characterization of CP and NB genes and phylogenetic analysis of ShVX from India.

Shallot bulbs were obtained from National Horticultural Development Foundation, Dindugal, Tamil Nadu, India. For amplification of CP and NB genes, total RNA was extracted from 100 mg of scale tissue from shallot bulb using RNeasy Plant Minikit (Qiagen) according to the manufacturer's protocol. First strand of cDNA was prepared using a previously published reverse primer from UTR [2]. For reverse transcription polymerase chain reaction (RT-PCR) a forward primer, ShVXf 5'GGGTGA AGGTCTGTGTTAAGTT3' was designed from ORF4 using ShVX genome sequence (Acc. No. NC_00379). PCR was performed using 2 µl of RT reaction mixture, 5 U of *Taq* DNA polymerase (Promega, Madison, USA), 3 µl of 10× reaction buffer, 1.5 mM MgCl₂, primer pairs at a concentration of 0.2 µmol and 0.2 mM dNTPs. The temperature profile consisted of a denaturation step at 94°C for 5 min, then 30 cycles of 45 s at 94°C, 20 s at 60°C and 1 min at 72°C and one final extension step at 72°C for 10 min.

Ten microlitres of amplified products were separated by electrophoresis in a 1% agarose gel containing ethidium bromide at a concentration of 0.5 µg ml⁻¹ and photographed under UV illumination with an imaging system (Biorad XR documentation system). PCR products were

S. Majumder · V. K. Baranwal (✉)
Plant Virology Unit, Division of Plant Pathology, Indian
Agricultural Research Institute, New Delhi 110012, India
e-mail: vbaranwal2001@yahoo.com

Present Address:

S. Majumder
Department of Biotechnology, Sharda University,
Greater Noida 201306, UP, India

Fig. 1 Cluster dendrogram showing the relationship between CP gene of ShVX isolates and other virus species within the genus *Allexivirus* based on the amino acid sequence of the coat protein

Fig. 2 Cluster dendrogram showing the relationship between NB gene of ShVX isolates and other virus species within the genus *Allexivirus* based on the amino acid sequence of NB region

purified using PCR purification kit (Qiagen, Germany). Sequencing of two separate PCR products of ~1.6 Kb was done at Chromous Biotech, Bangalore, India.

Sequence obtained was subjected to BLAST analysis for deducing the sequence homology of nucleotides as well as amino acids. Bioedit (7.0.7.1) [3] was used for creating sequence identity matrix with sequences of twelve allexiviruses including four of ShVX isolates [ShVX sequences, GQ268322 (ShVX India) NC_00379 (ShVX Russia1), L76292 (ShVX Russia 2), EU835197 (ShVX New Zealand); *Garlic virus A* (GarV-A Japan) NC_003375; *Garlic virus B* (GarV-B), AB010301 (GarV-B Japan), AF543829 (GarV-B Korea); *Garlic virus C* (GarV-C Japan) AB010302; *Garlic virus D* (GarV-D Korea) AF519572; *Garlic virus E* (GarV-E China) AJ292230; *Garlic virus X*

(GarV-X), AJ551526 (GarVX China), NC 001800 (GarV-X Korea)]. Open Reading Frames (ORF) were determined and molecular weight of predicted proteins were estimated by using the programmes available at <http://www.bioinformatics.org>. The conserved domains in the predicted proteins were explored using conserved protein domain database (CDD) [6] and protein predictor database <http://www.predictprotein.org>.

Pair-wise and multiple alignments were done using CLUSTAL X2 [9] and was used for constructing un-rooted dendrograms using the full optimal alignment and neighbor-joining method options with 1,000 bootstrap replications contained within this software and TreeView [7].

1,225 bp long nucleotide sequence was obtained from the overlapping amplified gene sequences. Two ORFs corresponding to ORF 5 (CP) and ORF 6 (NB) of ShVX (base 1–789 and base 789–1,160) could be obtained with the ORF finder programme. The nucleotide sequence of ORF 5 (CP) with 789 nucleotide showed 79–80% sequence identity to nucleotides of ORF 5 (encoding coat protein) of ShVX from New Zealand (EU835197) and Russia (NC_00379, L76292) and 66–76% with other allexiviruses. The nucleotide sequence of this ORF has a potential to encode a 28.66 kDa protein with 262 amino acids residues. The amino acid sequence for the predicted protein starts with AUG as start codon and ends with UGA as stop codon and shared maximum identity with capsid protein of ShVX sequences (85–88%). It shared identity of 54–61% with other allexiviruses capsid protein sequences. BLAST P analysis of CP indicated that it belonged to protein family pfam0286, a member of the Flexi_CP super-family cl02836 which included coat proteins from potexviruses and carlaviruses. It has a string of 14 amino acids (KFAGFDFFNAVLSDDSS) starting from the position 187–200, identical to signature sequence of coat proteins

(([RK][FYW]A [GAP] FD.F. {2} [LV]. {3}[GAST]{2}) for members of *Potexvirus*, *Carlavirus* and *Allexivirus*.

382 nucleotides extending from base 790 to base 1,161 in the ORF 6 (NB) showed nucleotide sequence identity of ~85% to that of two ShVX isolates from Russia (NC_003795 NB and L76292) and 68–76% with NB gene of other allexiviruses. Nucleotides in this ORF encoded a 14.18 kDa protein with 123 amino acid residues. The amino acid sequence of this predicted protein starts with AUG as start codon and ends with UAA as stop codon. BLAST P analysis revealed that it belongs to the protein family pfam 05515 of the viral_NABP super-family cl05203. This family includes nucleic acid binding protein from ssRNA positive-strand viruses. It has a cysteine rich protein with a distinct zinc ribbon motif, a common motif found in protein with nucleic acid binding property (65-CFDCGAYLYDNHVCKRFTSLSSDCLSVIH-94). It is a characteristic feature of viruses in *Allexivirus*, *Carlavirus*, *Mandarivirus*, and *Vitivirus*. The last Cys residue is, however, replaced by a His residue as has been observed in other ShVX isolates [4].

The cluster dendrogram based on the CP and NB amino acid sequences of allexiviruses (Figs. 1 and 2) showed that three major clusters are formed. GarV-B and GarV-X formed one cluster. GarV-A, GarV-E, and GarV-D were grouped in second cluster while isolates of ShVX formed a third cluster. GarV-C did not cluster with any of other allexiviruses and branched out separately. The Indian isolate of ShVX clustered with other ShVX isolates but branched out separately as a distinct isolate. That it is a distinct isolate can also be deduced from the percent identity matrix. The CP gene of Indian isolate of ShVX shared a maximum identity of 85–88% with ShVX isolates from Russia and New Zealand, and the latter two isolates shared identity of the order of 94–96% among themselves. The NB gene of Indian isolate of ShVX showed identity of ~85% with two Russian isolates which in turn shared identity of 96% with each other.

Coat protein plays an important role in virus infection while NB has been shown to play a role in suppressor of RNA silencing [12]. Therefore, knowledge of sequences of CP and NB would help develop diagnostics and management strategies of ShVX.

Acknowledgment The authors would like to thank Department of Science and Technology, Government of India, for financial assistance.

References

- Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Martelli GP, Milne RG, Zavriev SK, Fauquet CM. The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. *Arch Virol*. 2004;149:1045–60.
- Dovas CI, Hatzibukus E, Salomon R, Barg E, Shibolet Y, Katis NI. Comparison of methods for virus detection in *Allium* spp. *J Phytopathol*. 2001;149:731–7.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucl Acids Symp Ser*. 1999;41:95–98.
- Kanyuka KV, Vishnichenko VK, Levay KE, Kondrikov DYU, Ryabov EV, Zavriev SK. Nucleotide sequence of shallot virus X RNA reveals a 5'-proximal cistron closely related to those of potexviruses and a unique arrangement of the 3'-proximal cistrons. *J Gen Virol*. 1992;73:2553–60.
- Majumder S, Arya M, Pant RP, Baranwal VK. Shallot virus X in Indian shallot, a new virus report for India. *Plant Pathol*. 2007;57:2.
- Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH. CDD: a conserved domain database for interactive domain family analysis. *Nucl Acids Res*. 2007;35:237–40.
- Page RDM. Tree view: an application to display phylogenetic tree on personal computers. *Comp Appl Biosci*. 1996;12:357–8.
- Perez-Egusquiza Z, Ward LI, Clover GRG, Fletcher JD, van der Vlugt RAA. First report of *Shallot virus X* in shallot in New Zealand. *Plant Pathol*. 2009;58:407.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl Acids Res*. 1994;22:4673–80.
- Van Dijk P, van der Vlugt RAA. New mite-borne virus isolates from rakkyo, shallot and leek species. *Eur J Plant Pathol*. 1994;100:269–77.
- Vishnichenko VK, Konareva TN, Zavriev SK. A new filamentous virus in shallot. *Plant Pathol*. 1993;42:121.
- Zhou SZH, Dell'Orco M, Saldarelli P, Turturo C, Minafra A, Martelli GP. Identification of an RNA silencing suppressor in the genome of *Grapevine virus A*. *J Gen Virol*. 2006;87:2387–95.