

Molecular characterization of coat protein gene of *Garlic common latent virus* isolates from India: an evidence for distinct phylogeny and recombination

D. Pramesh · Virendra K. Baranwal

Received: 25 January 2013 / Accepted: 21 March 2013 / Published online: 4 April 2013
© Springer Science+Business Media New York 2013

Abstract The coat protein (CP) gene of five Indian *Garlic common latent virus* (GarCLV) isolates was sequenced and it was 960 bp long in all the five isolates, encoding a protein of 319 amino acids. Comparative nucleotide sequence analysis revealed diversity of 4.3 % among the Indian isolates and of 11.9 % among all isolates worldwide. Amino acid sequence comparison showed a significant variability in the N-terminal of CP of GarCLV. Various protein analysis tools identified thirteen conserved domains and motifs including *Carlavirus* and *Potexvirus*-specific Flexi CP and Flexi N CP. Phylogenetic analysis clustered GarCLV isolates in the subgroup II with isolates from Australia, Brazil, Japan, and South Korea. Intraspecies recombination study revealed that only one of the Indian isolates was a recombinant. Interspecies recombination study suggested the absence of genetic exchange from *Carlavirus* species to GarCLV; conversely, GarCLV was identified as a putative donor for at least two other *Carlavirus* species. This is the first report of molecular variability and recombination in GarCLV isolates.

Keywords GarCLV · Coat protein · Phylogeny · *Carlavirus* · Molecular variability · Recombination

Electronic supplementary material The online version of this article (doi:10.1007/s11262-013-0909-z) contains supplementary material, which is available to authorized users.

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

D. Pramesh
All India Coordinated Rice Improvement Project, Agricultural
Research Station, Gangavathi, University of Agricultural
Sciences, Raichur, Karnataka, India

Garlic common latent virus (GarCLV) is a member of the genus *Carlavirus* belonging to the family *Betaflexiviridae* of the order *Tymovirales*. GarCLV has flexuous, filamentous particles typical of carlaviruses [1], containing a single molecule of linear ssRNA of ~8.6 kb [2]. GarCLV is symptomless on garlic but it induces severe yellowing and mosaic when in mixture with other potyviruses such as *Onion yellow dwarf virus* and *Leek yellow stripe virus* [3–5]. Although, GarCLV has been reported from many countries around the world, its molecular characteristics such as genetic diversity, phylogenetic relationship, and intra and inter species recombination have not been sufficiently investigated to date. The present study reports the molecular characterization of five GarCLV isolates from India along with their phylogenetic relationships and evidence of recombination. For this study we have analyzed the sequence of CP since it is considered one of the most important region for species demarcation criteria in genus *Carlavirus* by ICTV [1].

Five isolates of GarCLV were selected from previously tested garlic accessions [6]: GarCLV-G1 (Northern India), GarCLV-RAU (Eastern India), GarCLV-Kolar (Southern India), GarCLV-JN and -Anand (Western India).

Primer GCLVCP-F (ATGTCAA/GC/TGAGTGAAC AGA) and primer GCLVCP-R(CTAGTCTGCATTGTTG GATCC) were designed and synthesized (Sigma Aldrich Co. India) for amplification of CP gene. Total RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN GmbH, Hilden, Germany) following the manufacturer's instructions. The first cDNA of GarCLV CP gene was synthesized using GCLVCP-R primer. The total 20 µl reverse transcription mixture containing 0.2 µM reverse primer (GCLVCP-R), 5 U MuLVRT (MBI, Fermentas, USA), 2 µl of 5X reaction buffer, and 0.3 mM dNTPs was incubated at 42 °C for 45 min. Then, the complete CP was

amplified following the PCR condition described previously [6].

A 960 bp product corresponding to the complete CP gene of GarCLV was amplified in all five isolates. Purified RT-PCR products were cloned into pGEM-T easy vector, according to manufacturer's instructions (Promega, Madison, USA) and confirmed by standard colony PCR with specific CP primers as well as by *EcoRI* digestion. Two clones of each isolate were sequenced using an AB 13130 Genetic analyzer (Xcelaris, Ahmadabad, India). The consensus sequences were deposited in GenBank bearing accession numbers JQ818255, JQ818256, JQ818257, JQ818258, and JQ818259 which correspond to JN, RAU, Kolar, Anand, and G1 isolates, respectively. In all isolates the length of the CP was found to be 960 nucleotide (nt) and the isolates shared nt and aa identity of 95.7–99.0 and 98.1–99.6 %, respectively, among them.

The CP sequences of 23 GarCLV isolates and a reference CP sequence for each of 37 *Carlavirus* species were retrieved from GenBank and used for sequence comparisons. The

comparison of 28 (5 from this study) GarCLV CP sequences revealed an identity of 88.1–99.8 % and 91.5–100 % at nt and aa level, respectively (Table 1). The comparison of CP sequences of 37 *Carlavirus* species with five GarCLV Indian isolates revealed an identity of 33–62 and 28–61 % at nt and aa level, respectively. The closest and farthest *Carlavirus* species of GarCLV were found to be *Coleus vein necrosis virus* (CVNV) and *Cole latent virus* (CoLV), sharing 61–62 and 33–34 % nt identity, respectively (Suppl. Table 1). This study shows the existence of an high sequence diversity (11.9 %) in CP gene among all GarCLV isolates worldwide. However, when the sole Indian isolates were considered, a diversity of only 4.3 % was observed in CP nucleotide sequences indicating distinctiveness of Indian isolates from other isolates of the world. Lack of substantial diversity among Indian isolates could be attributed to the propagative nature of the crop and the large diffusion of the single variety “G1” in the major garlic growing areas of India.

A multiple alignment of the CP aa sequences revealed that the major variability was located within the N-terminal

Table 1 Sequence identity (%) of coat protein (CP) gene of *Garlic common latent virus* isolates

S. no	Isolate name	Accession number in GenBank	Country	Nucleotide identity (%)	Amino acid identity (%)
1	JN	JQ818255	India	88.7–98.5	93.1–99.3
2	RAU	JQ818256	India	88.5–97.2	93.1–98.4
3	Kolar	JQ818257	India	88.2–98.8	91.5–96.8
4	Anand	JQ818258	India	88.8–98.8	93.1–99.6
5	G1	JQ818259	India	88.4–98.5	91.5–95.6
6	GAMS	AB004804	Japan	89.2–94.8	92.1–97.4
7	GHAV	AB004805	Japan	88.3–96.4	92.1–97.1
8	GF1	AB004566	Japan	88.6–90.1	92.1–95.9
9	K2	DQ520092	South Korea	88.9–96.8	93.7–99.3
10	GCLV-BZI	AF228416	Brazil	88.7–95.5	93.1–99.0
11	195-GCLV3	GQ475419	USA	88.3–99.8	92.5–100
12	366-GCLV3	GQ475420	USA	88.2–99.6	91.8–99.3
13	362-GCLV3	GQ475421	USA	88.3–99.8	92.1–99.6
14	365-GCLV3	GQ475422	USA	88.2–99.5	91.5–99.0
15	370-GCLV3	GQ475423	USA	88.2–99.4	93.1–99.3
16	341-GCLV3	HQ873852	USA	88.3–99.8	92.5–100
17	368-GCLV3	HQ873856	USA	88.3–99.8	91.5–99.0
18	369-GCLV3	HQ873860	USA	88.2–99.6	92.1–99.3
19	364-GCLV3	HQ873862	USA	88.5–99.8	92.8–98.1
20	367-GCLV3	HQ873863	USA	88.1–99.5	91.5–99.0
21	197-GCLV4	HQ873855	China	88.2–99.8	92.5–100
22	340-GCLV4	HQ873861	China	88.1–99.6	91.5–98.7
23	338-GCLV3	HQ873858	China	88.4–99.8	92.5–100
24	337-GCLV4	HQ873859	China	88.4–99.8	92.5–100
25	334-GCLV4	HQ873854	China	88.4–99.8	92.5–100
26	335-GCLV4	HQ873857	China	88.2–99.7	94.0–99.3
27	336-GCLV4	HQ873853	China	88.3–90.0	94.6–100
28	WA-1	JF320810	Australia	88.1–96.4	91.5–96.5

Isolates from the current study (originated from India) are in bold

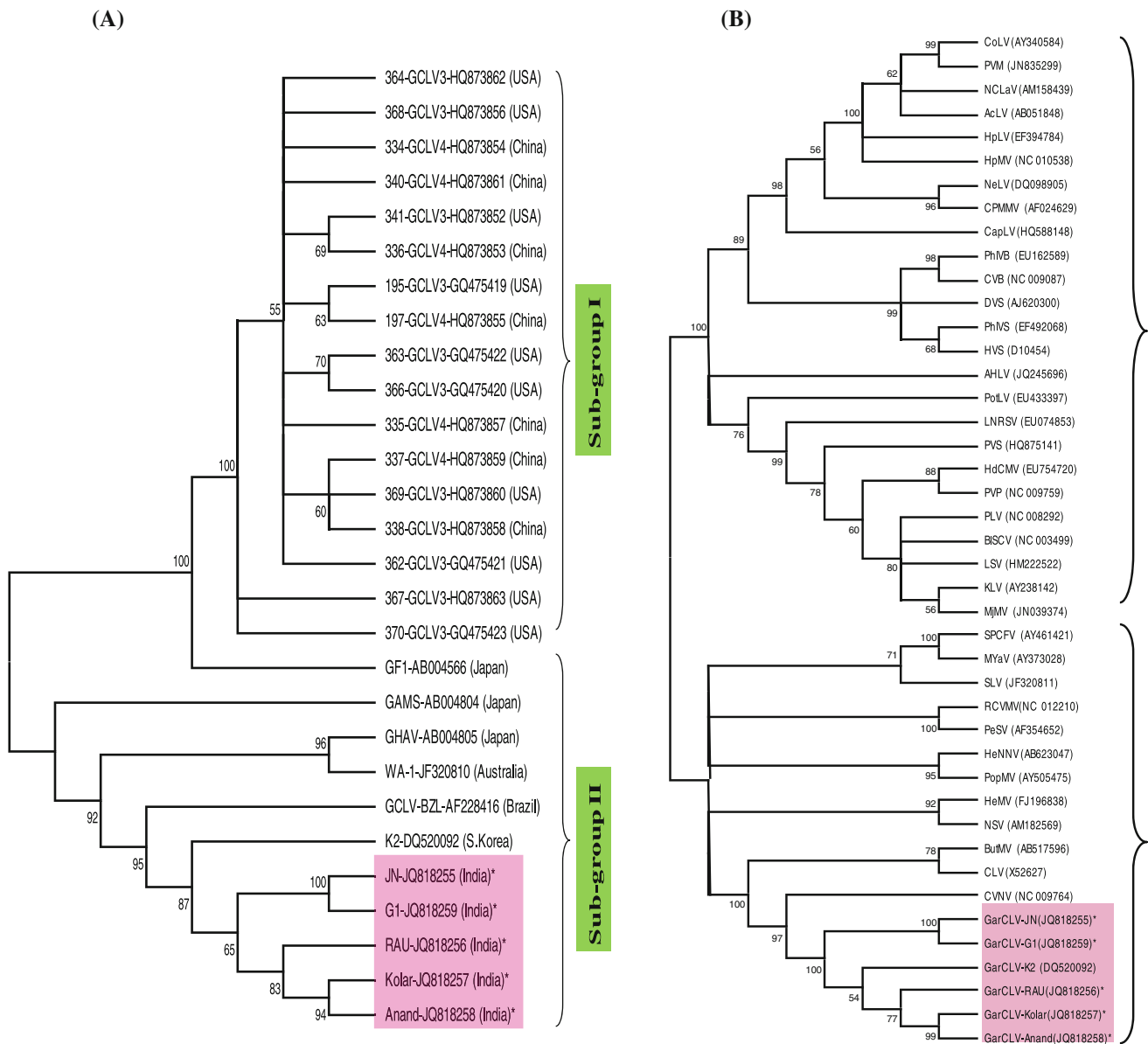


Fig. 1 a Phylogenetic analysis of coat protein (CP) of *Garlic common latent virus* (GarCLV) isolates from different parts of the world. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura-Nei model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was <100 or less than one-fourth of the total number of sites, the maximum parsimony method was used; otherwise BIONJ method with MCL distance matrix was used. The analysis involved 28 sequences. Codon positions included were 1st + 2nd + 3rd + Non-coding. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA4. **b** Phylogenetic analysis of coat protein (CP) of *Carlavirus* species infecting diverse plant species. The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1,000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50/bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches. The evolutionary distances were computed using the

p-distance method and are in the units of the number of base differences per site. The analysis involved 43 nucleotide sequences. All ambiguous positions were removed for each sequence pair. Evolutionary analyses were conducted in MEGA4. Full name of the viruses: NSV *Narcissus symptomless virus*, SPCFV *Sweet potato chlorotic fleck virus*, RCVNV *Red clover vein mosaic virus*, NCLV *Narcissus common latent virus*, HpMV *Hop mosaic virus*, CVNV *Coleus vein necrosis virus*, CoLV *Cole latent virus*, AcLV *Aconitum latent virus*, BISCV *Blueberry scorch virus*, CLV *Carnation latent virus*, CVB *Chrysanthemum virus B*, CPMMV *Cowpea mild mottle virus*, HpLV *Hop latent virus*, LSV *Lily symptomless virus*, MYaV *Melon yellowing- associated virus*, PLV *Passiflora latent virus*, PeSV *Pea streak virus*, PotLV *Potato latent virus*(), PVM *Potato virus M*, PVS *Potato virus S* PVP *Potato virus P*, SLV *Shallot latent virus*(), AHLV *American hop latent virus*, DVS *Daphne virus S*, ButMV *Butterbur mosaic virus*, HeMV *Helleborus mosaic virus*, HeNNV *Helleborus net necrosis virus*, HdCMV *Hydrangea chlorotic mottle virus*, KLV *Kalanchoe latent virus*, LNRSV *Ligustrum necrotic ringspot virus*, MjMV *Mirabilis jalapa mottle virus*, PhIVB *Phlox virus B*, PhIVS *Phlox virus S*, PopMV *Poplar mosaic virus*, CapLV *Caper latent virus*, HVS *Helenium virus S*, NeLV *Nerine latent virus* and GarCLV *Garlic common latent virus* K2 isolate

48 aa, where 21 aa were highly variable among 28 isolates (Suppl. Fig. 1). Conversely, the central and the C-terminal regions (from 49 to 319 aa) are highly conserved in all isolates (data not shown). The presence of conserved domains such as Flexi_CP_N and Flexi_CP had been previously reported in CP of carlaviruses and potexviruses [7]. In the present study, the presence of 13 different conserved domains and of 3 motifs within the CP were predicted using SMART [8] and MEME [9] programs, respectively (Suppl. Table 2). Sequences of all the three motifs were subjected to TOMTON motif comparison web tool, which indicated that they showed similar signatures with transcriptional factors and a possible role in nucleic acid binding property.

Maximum likelihood phylogenetic tree constructed in MEGA4 [10] splits the GarCLV isolates into two major clusters, designated as Subgroup I and Subgroup II (Fig. 1a). Subgroup I included isolates from USA and China, whereas Subgroup II included isolates from Australia, Brazil, India, Japan, and South Korea. The five Indian isolates were grouped together with K2 isolate of South Korea and GCLV-BZL isolate of Brazil. On the basis of CP gene sequences, isolates of various virus species, such as, *Cucumber mosaic virus*, *Prunus necrotic ringspot virus*, *Turnip mosaic virus*, *Chrysanthemum virus B* (CVB), *Peanut stunt virus*, etc, have been grouped into 2–4 sub-groups [11–17]. In our study phylogenetic analysis of the CP revealed the distinct clustering pattern based on geographic origin of the GarCLV isolates. CP gene sequences could also group viruses of genus *Carlavirus* in two distinct phylogenetic subgroups. A similar sub grouping of *Carlavirus* species could also be shown when other genomic regions, i.e., nucleic acid binding protein and replicase genes, are used [7, 18, 19].

Phylogenetic analysis of GarCLV Indian isolates with representative isolates of other 37 *Carlavirus* species showed the formation of two phylogenetic subgroups (Fig. 1b). All the Indian isolates formed a single clustering pattern and grouped together with *Shallot latent virus* (SLV) and 10 other *Carlavirus* species. The remaining 26 *Carlavirus* species clustered in a separate subgroup.

Potential recombination events were identified using the RDP3 v.3.34 [20]. Intraspecies recombination analysis of 28 GarCLV isolates showed in the CP of the Anand isolate (breakpoint from nt position 167 to 5) a recombination event having JN and Kolar isolates as donors. The recombination was detected simultaneously by two algorithms of RDP3 (3Seq and SiScan) (Suppl. Fig. 2). Intraspecies recombination in the CP gene of the carlaviruses CVB [13, 21] and *Lily symptomless virus* has been already reported [22]. In our study we could detect intraspecies recombination in only one of the 28 isolates. Evidence for the lack of strong recombination signals in CP gene of

remaining GarCLV isolates has suggested that, the CP gene may be recombination cold spot as reported in other carlaviruses such as CVB [21], SLV, and *Potato virus S* [23]. The existence of recombinant isolate in India could be attributed to the existence of multiple isolates in the same garlic accessions which are brought together during the extensive multilocal trials and screening conducted before their release for commercial cultivation.

Interspecies recombination analysis of 37 *Carlavirus* species along with GarCLV in RDP3 program detected the recombination signals in the CP of 22 *Carlavirus* species. Interspecies recombination has played a major role in the evolution of members of ssDNA virus family *Geminiviridae* [24], whereas, in positive sense ssRNA viruses, it has been reported in the *Potyviridae* [25] and *Luteoviridae* [26]. Interspecies recombination analysis in the current study revealed the lack of extensive exchange of genetic material from other *Carlavirus* species to GarCLV, but still, provided an evidence for interspecies recombination events in the CP gene of 22 *Carlavirus* species (Suppl. Table 3). Interestingly, SLV, another *Carlavirus* infecting garlic neither contributed nor received the CP fragment from GarCLV (Suppl. Table 3), though both infect same hosts and are transmitted by the same vector.

In conclusion, molecular analysis of CP of GarCLV has clearly indicated the existence of two phylogenetic subgroups and the occurrence of a recombinant isolate in India.

Acknowledgments First author is thankful to PG School, Indian Agriculture Research Institute (IARI), New Delhi for financial assistance during Ph. D program. The help rendered by Dr. Subodh Joshi and Dr. Sabina Islam in providing garlic accessions is highly acknowledged. Authors are grateful to ICAR for financial support through Outreach program and Head, Plant Pathology, IARI for necessary laboratory facilities.

References

1. A.M.Q. King, M.J. Adams, E.B. Carstens, E.J. Lefkowitz, *Ninth Report of the International Committee on Taxonomy of Viruses* (Elsevier Academic Press, San Diego, 2012)
2. S.J. Wylie, H. Luo, H. Li, M.G. Jones, *Arch. Virol.* **157**(2), 271–284 (2012)
3. D.G.A. Walkey, M.J.W. Ebb, C.J. Ballooned, A. Millar, *J. Hortic. Sci.* **62**, 211–220 (1987)
4. M. Takaichi, M. Yamamoto, T. Nagakubo, K. Oeda, *Plant Dis.* **82**, 694–698 (1998)
5. Y.W. Lee, S. Yamazaki, T. Osaki, T. Inouye, *Ann. Phytopathol. Soc. Jpn.* **45**, 727–734 (1979)
6. D. Pramesh, S. Islam, J. Singh, V.K. Baranwal, *Med. Plant* **4**(1), 17–22 (2012)
7. J. Chen, J.P. Chen, S.A. Langeveld, A.F.L.M. Derks, M.J. Adams, *J. Phytopathol.* **151**, 26–29 (2003)
8. I. Letunic, T. Doerks, P. Bork, *Nucleic Acids Res.* **37**, D229–D232 (2009)

9. T.L. Bailey, N. Williams, C. Misleh, W.W. Li, *Nucleic Acids Res.* **34**, W369–W373 (2006)
10. K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol. Biol. Evol.* **24**, 1596–1599 (2007)
11. A. Srivastava, S.K. Raj, *Curr. Sci.* **87**, 1126–1131 (2004)
12. K. Ohshima, Y. Yamaguchi, R. Hirota, T. Hamamoto, K. Tomimura, Z. Tan, T. Sano, F. Azuhata, J.A. Walsh, J. Fletcher, J. Chen, A. Gera, A.J. Gibbs, *Gen. Virol.* **83**, 1511–1521 (2002)
13. L. Singh, V. Hallan, N. Jabeen, A.K. Singh, R. Ram, D.P. Martin, A.A. Zaidi, *Arch. Virol.* **152**, 405–413 (2007)
14. N. Verma, A.K. Singh, L. Singh, S. Kulshrestha, G. Raikhy, V. Hallan, R. Ram, A.A. Zaidi, *Plant Dis.* **88**, 1161 (2004)
15. N. Verma, A.K. Singh, L. Singh, S. Kulshrestha, G. Raikhy, V. Hallan, R. Ram, A.A. Zaidi, *Aus. Plant Pathol.* **34**, 119–120 (2005)
16. R.W. Hammond, *Arch. Virol.* **14**, 1195–1210 (2003)
17. Z. Xu, C. Higgins, K. Chen, R.G. Dietzgen, Z. Zhang, L. Yan, X. Fang, *Plant Dis.* **82**, 992–998 (1998)
18. K.C. Eastwell, K.L. Druffel, *Arch. Virol.* (2012). doi: [10.1007/s00705-012-1312-0](https://doi.org/10.1007/s00705-012-1312-0)
19. M.G. Bernardy, C.R. Dubeau, A. Braun, C.E. Harlton, A. Bunckle, L.A. Wegener, D.T. Lowery, C.J. French, *Can. J. Plant Pathol.* **27**, 581–591 (2005)
20. D.P. Martin, P. Lemey, M. Lott, V. Moulton, D. Posada, P. Lefevre, *Bioinformatics* **26**(19), 2462–2463 (2010)
21. L. Singh, V. Hallan, D.P. Martin, R. Ram, A.A. Zaidi, *Arch. Virol.* **157**, 531–537 (2012)
22. A.K. Singh, B.K. Mahinghara, V. Hallan, R. Ram, A.A. Zaidi, *Virus Genes* **36**, 421–427 (2008)
23. E.R. Chare, E.C. Holmes, *Arch. Virol.* **151**, 933–946 (2006)
24. M. Padidam, S. Sawyer, C. Fauquet, *Virol.* **265**, 218–225 (1999)
25. C. Desbiez, H. Lecoz, *Arch. Virol.* **149**, 1619–1632 (2004)
26. G.R. Smith, Z. Borg, B.E.L. Lockhart, K.S. Braithwaite, M.J. Gibbs, *J. Gen. Virol.* **81**, 1865–1869 (2000)