ORIGINAL ARTICLE



Complete genome sequencing of cucumber mosaic virus from black pepper revealed rare deletion in the methyltransferase domain of 1a gene

K. A. Revathy^{1,2} · A. I. Bhat¹

Received: 1 April 2017/Accepted: 10 June 2017/Published online: 17 June 2017 © Indian Virological Society 2017

Abstract The complete genome of cucumber mosaic virus (CMV) from black pepper was sequenced and compared with CMV isolates from subgroups I and II reported worldwide. Percent identity and phylogenetic analyses clearly indicated that the CMV isolate from black pepper (BP) belongs to subgroup IB. Sequence analyses also showed the presence of a rare deletion of nine nucleotides in the putative methyltransferase domain of 1a gene which was observed only in two more isolates of CMV among one hundred 1a gene sequences of CMV for which sequence information is available in the database. Interestingly this deletion is not present in the black pepper isolate of CMV from China (WN1) and from Indian long pepper that is closely related to black pepper. Percent identity analyses showed that the 3'untranslated region (UTR) of the three RNAs of the BP isolate were conserved with 91% identity whereas the 5'UTR of three RNAs showed 52-80% identity. The level of gene conservation among the subgroups was highest in coat protein and lowest in 2b. The values of nucleotide diversity studies were further consistent with the above data. The ratio of non-synonymous to the synonymous substitution of the five genes of three RNAs was in the order 1a > 2a > 2b >3a > 3b and less than one for all the genes, indicating purifying selection. These clearly reflect that the protein

Electronic supplementary material The online version of this article (doi:10.1007/s13337-017-0386-4) contains supplementary material, which is available to authorized users.

A. I. Bhat aib65@yahoo.co.in

encoded by RNA1 is highly tolerant to amino acid changes followed by that of RNA2 and, RNA3 is the least tolerant correlating to its functional importance.

Keywords Cucumber mosaic virus · Black pepper · Complete genome · Phylogenetic analysis · Nucleotide diversity

Introduction

Black pepper (Piper nigrum L.) (Piperaceae) is one of the economically important spice crops of India mainly grown in the states of Kerala and Karnataka. Known as the King of spices, it originated in the tropical evergreen forests of Western Ghats of India [23]. Though, India has the largest area under black pepper cultivation, its productivity is often low due to biotic and abiotic stresses. Cucumber mosaic virus (CMV) (genus: Cucumovirus, family Bromoviridae) is one of the major production constraint of black pepper in India. Along with piper yellow mottle virus, it causes stunted disease in black pepper which is the third major disease of this spice crop. In a survey conducted in 2004 in India, high incidence and severity of the disease was reported from black pepper plantations located especially at high altitudes such as Idukki and Wayanad districts of Kerala where the mean incidence ranged from 29 to 45% respectively and, Kodagu and Hassan districts of Karnataka where the mean incidence was 14.9 to 5.2% [4]. Besides black pepper CMV infects other related species such as Indian long pepper and betel vine [12]. Association of CMV with the disease has also been reported from Brazil [10] and Sri Lanka [7, 8]. Typical symptoms include stunting of plant, mottling and mosaic on leaves, reduction of spike length and poor filling of spikes leading to yield



Division of Crop Protection, ICAR-Indian Institute of Spices Research, Marikunnu, Kozhikode, Kerala 673012, India

² University of Calicut, Malappuram, Kerala 673635, India

reduction. The virus is mainly spread through infected stem cuttings and more than 80 species of aphids are known to spread CMV in a non-persistent manner [21].

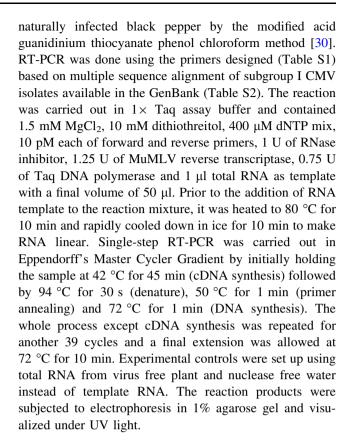
CMV is a tripartite positive sense RNA virus with three single stranded RNA molecules designated as RNA1, RNA2 and RNA3 enclosed in separate particles. RNA1 is monocistronic coding for 1a gene. RNA2 and RNA3 are dicistronic coding for 2a, 2b and 3a, 3b genes, respectively. Open reading frames 1a and 2a are the viral components of replicase [21]. N-terminal region of 1a protein has putative methyltransferase domain [20] and C-terminal is homologous to viral helicases [11]. The 2b gene expressed from subgenomic RNA4A acts as suppressor of post transcriptional gene silencing [2, 6, 16] besides its role in long distance movement [9, 14, 31, 35] and aphid transmission [36]. The 3a gene encodes the movement protein (MP) and 3b, the coat protein (CP) expressed through subgenomic RNA (RNA4). MP is involved in cell to cell movement of the virus and CP helps in encapsidation of viral nucleic acid and affects symptom expression [29, 33]. Based on serology, symptomology, host range data, peptide mapping of CP and nucleic acid hybridization, CMV strains have been classified into subgroups I and II [26]. Subgroup I is further divided into IA and IB based on the 3' UTR of RNA3 [27]. Subgroup IA and II is seen worldwide whereas IB is mainly restricted to Asian countries [28].

Currently there are no varieties or cultivars of black pepper resistant to CMV and no chemicals are available to control the virus. So far use of virus free planting materials of elite genotypes is the management strategy adopted. Due to the lack of virus resistant varieties in any of the black pepper growing countries, developing genetic resistance is the best management strategy for which sequence information of CMV from black pepper isolate is essential. From India many CMV strains have been reported but complete genome sequence is available from only five isolates. So far CP sequence of black pepper isolate from India is reported [5]. Here we report the complete genome sequence of this isolate that has a unique deletion in the methyltransferase domain of 1a gene, its phylogenetic relationships and nucleotide diversity studies.

Materials and methods

Virus source, total RNA extraction and amplification of CMV by RT-PCR

Black pepper naturally infected with CMV showing typical symptoms, collected from Kodagu District, Karnataka, India and maintained at the ICAR—Indian Institute of Spices Research, Kozhikode, Kerala, India was used for the study. Total RNA was extracted from 100 mg leaf tissue of



Cloning and sequencing

The RT-PCR products were gel purified and ligated into pTZ57R/T plasmid vector using InsTAcloneTM PCR Cloning kit (Fermentas, USA) according to the manufacturer's protocol. The ligated mixture was then used for the transformation of *Escherichia coli* DH5α competent cells. Luria Bertoni plates with ampicillin 100 μg/ml were used for screening of transformed colonies. Plasmid isolated from positive clones was sequenced at the automated sequence facility available at Eurofins, Bangalore, India. At least two clones were sequenced for each amplicon and each amplicon was sequenced from both strands to get the consensus sequence.

Sequence analyses

Sequences were assembled with Sequid Version 3.6 [22]. ORFs in the sequence data were predicted by ORF finder (www.ncbi.nlm.nih.gov/projects/gorf) and translated into amino acid residues using Expasy tool. The analyses were carried out using the complete genome of 29 isolates of CMV belonging to subgroup I and II and one outgroup (peanut stunt virus), retrieved from NCBI database (Table S2) in the form of nucleotide and translated amino acid sequences. Sequences were aligned using clustalX [34] and percentage identity was calculated using Clustal



Omega. Phylogenetic analyses were performed using Bayesian algorithm in MrBayes [25] and number of segregation sites and ratio of dN/dS for different genes was estimated using DnaSP version 5.10 [17].

Results

Complete genome sequence

Complete nucleotide [nt] sequence of the CMV-black pepper isolate comprised RNA1 (3349 nt), RNA2 (3049 nt) and RNA3 (2217 nt) (Table S3). RNA1 has a single ORF coding for 990 aa 1a protein. RNA2 and RNA3 each have two ORFs. The 2a and 2b genes codes for proteins with 858 and 111 amino acids respectively. ORF 3a codes for the movement protein with 279 amino acids and ORF 3b codes for 218 amino acids long coat protein.

Multiple sequence alignment

Multiple sequence alignment of the five genes revealed nine nucleotides deletion in the methyltransferase domain of 1a gene (Fig. 1) which was also reflected in the multiple sequence alignment of deduced amino acid sequence that showed three amino acids deletion. This deletion was observed only in two more CMV isolates (CLW2 from Malaysia infecting cucumber and HM3 from China infecting tomato) among one hundred 1a gene sequences analyzed including the CMV isolate infecting black pepper from China and Indian long pepper (a closely related species to black pepper) from India. The size of 1a gene in these three isolates was 2973 nt whereas in all other subgroup I isolates it was 2982 nt. A ten nucleotide deletion in the same region in peanut stunt virus (PSV), a distinct species under the genus *Cucumovirus* (Fig. 1) and two nucleotide deletion in the case of other two distinct species (gayfeather mild mottle virus and tomato aspermy virus) of the genus, *Cucumovirus* (not shown) are notable. In PSV another five nucleotide deletion compensates the frame shift made by the ten nucleotide deletion (Fig. 1).

Percent identity analyses

Percent identity analyses of CMV black pepper isolate revealed that all five genes at the deduced amino acid level showed maximum identity to subgroup IB strains with an identity of 91–97% in 1a, 94–98% in 2a, 77–95% in 2b, 85–97% in 3a and 95–98% in 3b. Whereas the identity of the corresponding genes to subgroup IA were 93, 92–94, 68–79, 91–95 and 95–96%, respectively. Among five genes at the amino acid level; 3b gene showed maximum identity among the subgroups (80–99%) and 2b showed the

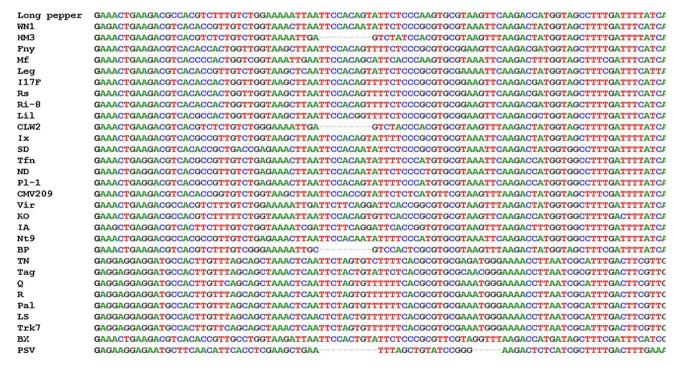


Fig. 1 Multiple nucleotide sequence alignment of a portion of cucumber mosaic virus (CMV) 1a gene of BP isolate with other CMV strains including black pepper (China), Indian long pepper and

outgroup peanut stunt virus showing deletion of nine bases aligned using clustalX. The details of isolates are provided in Table S2



312 K. A. Revathy, A. I. Bhat

Table 1 Percent identity in the 2b (below diagonal) and 3b (above diagonal) genes of BP isolate of cucumber mosaic virus at the amino acid level with other isolates of CMV subgroup I and II

Strain	BP	IB	IA	II
BP		94–98	95–96	80–82
IB	75–95		95–99	80-83
IA	68–79	65-84		80-85
II	51-52	43-55	39-52	

minimum (51–95%) (Tables 1 and 2). All the genes of BP isolate showed highest identity of 95–98% with KO isolate from Karnataka India, followed by CLW2 isolate from Malaysia sharing 89–98% identity. The percent identity with black pepper isolate WN1 from China ranged from 80–96% which is less as compared to the identity with KO isolate from India and CLW2 isolate from Malaysia. The 3' UTR of three RNAs of the current isolate shared more similarity of 91% compared to the 5' UTR (52–80%). Also, the 3' and 5' UTR of the three RNAs of BP isolate of CMV showed highest identity of 85–97% and 82–98% to CMV subgroup IB strains. The identity of the above regions of the BP isolate to subgroup IA strains was 86–92% and 81–92%. Percent identity analyses thus suggest that the BP belongs to subgroup IB.

Phylogenetic analyses

The phylogenetic analyses of the five genes at the nucleotide (not shown) and amino acid level (Fig. 2 and Fig.S1) also showed that the BP is closely clustered to subgroup IB strains. Both the trees generated using nucleotide and deduced amino acid were identical topologically. In all the trees BP isolate was seen in close association with KO and CLW2 strains which was consistent to the results of percent identity analyses.

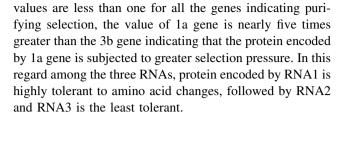
Evolutionary analyses

Though all the genes of the current isolate showed haplotype diversity value nearly one, the highest nucleotide diversity was seen in 2b while 3b had the least (Table 3). The ratio of non-synonymous to synonymous substitution is consistent with the above data with 3b gene having the

Table 2 Percent identity of 2a, 3a and 1a genes of BP isolate of cucumber mosaic virus at the amino acid level with other isolates of subgroup I and II done using Clustal Omega

Strain	BP	IB	IA	II
BP		93-95 (91-97)	92-95 (93)	75–76 (84–85)
IB	85-97		90-95 (92-97)	73–76 (82–85)
IA	91–95	82–98		75–80 (84–86)
II	81-82	75–85	78–93	

Above the diagonal is 2a, below the diagonal is 3a and values in brackets are 1a gene



least but 1a gene had the highest value of 0.5. Though these

Discussion

Cucumber mosaic virus with worldwide distribution is known to be the most devastating plant virus infecting more than 1200 species from 100 plant families [24]. Understanding genome of the viruses is essential for devising better diagnostic and management strategies by selecting the most appropriate gene for developing pathogen derived resistance. This study characterized for the first time the complete genome sequence of black pepper isolate of CMV (BP), and analysed its phylogeny with other isolates reported from different parts of the world. The presence of nine nucleotide deletion in the multiple nucleotide sequence of 1a gene in three of the isolates taken in the present study but absent from the black pepper isolate from China and Indian long pepper was notable. Direct sequencing of the PCR product amplified by RT-PCR using specific primers covering the region of deletion in RNA1 further confirmed the presence of deletion. The production of systemic mosaic symptoms in sap inoculated Nicotiana tabacum and chlorotic lesions in Chenopodium amaranticolor is an indication that the virus is replicating. However the functional importance of this deletion and whether the deletion is having any significant effects on the CMV strain is yet to be studied. The present isolate shared close identity with the KO strain of CMV from Karnataka followed by other subgroup IB isolates which was in compliance to the earlier studies where high level of conservation was reported among the CMV isolates infecting different hosts in a particular region [3]. High level of conservation was found in the percent identity analyses of 3'UTR of the three RNAs' suggesting a common ancestor. Similar observations were reported earlier [15, 24, 28]. High percent identity of the BP isolate towards the CLW2 isolate from Malaysia showed that these are genetically related which might be due to the fact that CMV move between the countries



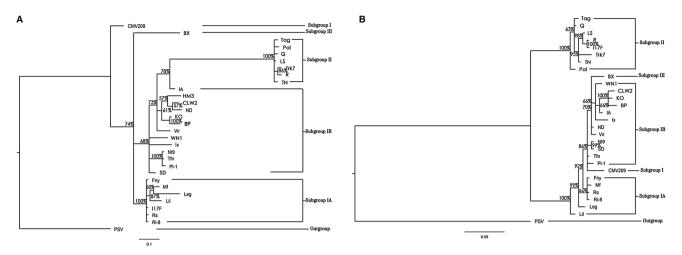


Fig. 2 Phylogenetic trees of 2b (A) and 3b (B) genes of BP isolate of cucumber mosaic virus (CMV) at the amino acid level with other isolates of CMV constructed using Bayesian algorithm in MrBayes. Designation used for the isolates is listed in Table S2

Table 3 Summary of nucleotide diversity studies of five genes of BP isolate of cucumber mosaic virus with other isolates of CMV subgroup I and II

Genes	n	S	Hd	π	nS	πS	nN	πΝ	ω
1a	2966	1107	1.0000	0.13643	637	0.25970	2311	0.13041	0.502
2a	2513	1086	1.0000	0.15187	576	0.366	1920	0.13651	0.372
2b	297	175	0.997	0.20663	67	0.57077	218	0.20276	0.36
3a	819	412	1.0000	0.12972	183	0.324	600	0.10554	0.33
3b	654	229	0.997	0.12255	168	0.4999	480	0.05913	0.12

n no of sites, S no of segregating sites, Hd measurement of haplotype diversity, π Nucleotide diversity, nS no of synonymous sites, πS synonymous substitution, nN no of nonsynonymous sites, πN nonsynonymous substitution, ω Ratio of $\pi N/\pi S$

or regions along with the infected materials as reported earlier [19].

Close clustering of the BP isolate towards subgroup IB strains in phylogenetic analyses of all the genes at the nucleotide and deduced amino acid level was in accordance with the previous study [28] where all the Asian strains were placed in subgroup IB. New isolates of viruses originate due to the adaptation to host plants and the environment. Specific adaptation of CMV to soybean plant was reported from Indonesia [13]. The less sequence identity between the two black pepper isolates and more identity of the BP isolate to KO and CLW2 strains suggests the adaptation of virus to new hosts and geographic locations, respectively. Earlier studies have reported high sequence identity among the Indian isolates of CMV [18, 32]. Previously based on coat protein gene sequence it was reported that the BP isolate belonged to subgroup I [5]. Phylogenetic analyses together with percent identity analyses in the present study confirmed the grouping of BP isolate in subgroup IB. Among the different genes at nucleotide and deduced amino acid level, 3b was the most conserved and 2b the most variable in percent identity and phylogenetic analyses as reported by many earlier workers [15, 28].

The evolutionary constraints on each coding gene might be the probable reason for the apparent differences in their phylogenetic trees, the 3b tree being the most compact and the 2b being the most variable. Thus the evolutionary forces acting on the different genes were also studied. Though all the genes of the current isolate showed haplotype diversity value nearly one, the highest nucleotide diversity was seen in 2b and 3b had the least (Table 3) which might be due to the fact that CP interacts with itself or viral RNA but little with the host and the interactions with the aphid vectors are also nonspecific and minimal [26]. The role of 2b gene in host specificity might be the reason for its high variability. The ratio of non-synonymous to synonymous substitution is consistent with the above data with 3b gene having the least but 1a had the highest, which is contradictory to the previous study [20]. The high nature of conservation of RNA 3 in the CMV strains from different geographic locations suggests its importance in viral RNA functions as reported earlier [1].

Acknowledgements Authors are thankful to Director, ICAR-Indian Institute of Spices Research, Kozhikode for the facilities. Senior author is thankful to the Department of Science and Technology, Government of India for funding under the WOS-A scheme.



K. A. Revathy, A. I. Bhat

References

- Andreev IA, Hyon Kim S, Kalinina NO, Rakitina DV, Fitzgerald AG, Palukaitis P, Taliansky ME. Molecular interactions between a plant virus movement and RNA: force spectroscopy investigation. J Mol Biol. 2004;339:1041–7.
- Beclin C, Berthome R, Palauqui JC, Tepfer M, Vaucheret H. Infection of tobacco or Arabidopsis plants by CMV counteracts systemic post-transcriptional silencing of nonviral (trans) genes. Virology. 1998;252:313–7.
- 3. Bhadramurthy V, Anju G, Bhat AI, Shiva KN. Coat protein gene sequence studies suggest that *Cucumber mosaic virus* infecting paprika (*Capsicum annuum* L.) in India belongs to subgroup IB. Arch Phytopathol Plant Prot. 2009;42:857–63.
- Bhat AI, Devasahayam S, Venugopal MN, Suseela Bhai R. Distribution and incidence of viral disease of black pepper (*Piper nigrum* L.) in Karnataka and Kerala, India. J Plant Crops. 2005;33:59–64.
- Bhat AI, Hareesh PS, Madhubala R. Sequencing of coat protein gene of an isolate of *Cucumber mosaic virus* infecting black pepper (*Piper nigrum*. L) in India. J Plant Biochem Biotechnol. 2005;14:37–40.
- Brigneti G, Voinnet O, Li WX, Ji LH, Ding SW, Baulcombe DC. Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. EMBO J. 1998;17:6739–46.
- de Silva DPP, Jones P, Shaw MW. Identification and transmission of *Piper yellow mottle virus* and *Cucumber mosaic virus* infecting black pepper (*Piper nigrum* L.) in Sri Lanka. Plant Pathol. 2002;51:537–45.
- de Silva DPP, Jones P, Shaw MW. Isolation, purification and identification of *Cucumber mosaic virus* in Sri Lanka on black pepper (*Piper nigrum* L.). Sri Lankan. J Agric Sci. 2001;38:17–27.
- Ding SW, Li WX, Symons RH. An interspecies hybrid RNA virus is significantly more virulent than either parental virus. Proc Natl Acad Sci USA. 1996;93:7470–4.
- Eiras M, Boari AJ, Colariccio A, Chaves AIR, Briones MRS, Figueira AR, Harakava R. Characterization of isolates of the cucumovirus *Cucumber mosaic virus* present in Brazil. J Plant Pathol. 2004;86:61–9.
- Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. 1989;17:4713–30.
- Hareesh PS, Madhubala R, Bhat AI. Characterization of Cucumber mosaic virus infecting Indian long pepper (Piper longum L.) and betel vine (Piper betle L.) in India. Indian. J Biotechnol. 2006;5:89–93.
- Hong JS, Masuta C, Nakano M, Abe J, Uyeda I. Adaptation of Cucumber mosaic virus soybean strains (SSVs) to cultivated and wild soybeans. Theor Appl Genet. 2003;107:49–53.
- 14. Ji LH, Ding SW. The suppressor of transgene RNA silencing encoded by *Cucumber mosaic virus* interferes with salicylic acid mediated virus resistance. Mol Plant Microbe Interact. 2001;14:715–24.
- Koundal V, Mohd Q, Haq R, Praveen S. Characterization, genetic diversity and evolutionary link of *Cucumber mosaic virus* strain New Delhi from India. Biochem Genet. 2011;49:25–38.
- Li HW, Lucy AP, Guo HS, Li WX, Ji LH, Wong SM, Ding SW. Strong host resistance targeted against a viral suppressor of the plant gene defence silencing mechanism. EMBO J. 1999;18:2683–91.
- Librado P, Rozas J. DnaSP v5: software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451–2.

- Madhubala R, Bhadramurthy V, Bhat AI, Hareesh PS, Retheesh ST, Bhai RS. Occurrence of *Cucumber mosaic virus* on vanilla (*Vanilla planifolia* Andrews) in India. J Biosci. 2005;30:339–50.
- Mazidah M, Yusoff K, Habibuddin H, Tan YH, Lau WH. Characterization of Cucumber mosaic symptom on *Catharanthus roseus* (L.) G. Don in Malaysia. Petranika J. Trop. Agric Sci. 2012;35:41–53.
- Mikhail NR, Eugene VK, Alexander EG. Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbislike' supergroup of positive-strand RNA viruses. J Gen Virol. 1992;73:2129–34.
- Palukaitis P, Roossinck MJ, Dietzgen RG, Francki RIB. Cucumber mosaic virus. Adv Virus Res. 1992;41:281–348.
- Peltola H, Soderlund H. Ukkonen E SEQAID: a DNA sequence assembling program based on a mathematical model. Nucleic Acids Res. 1984;12:307–21.
- Ravindran PN. Black pepper (*Piper nigrum*). Amsterdam: Harwood Academic Publishers; 2000.
- Reenu K, Pooja B, Lakhmir S, Aijaz AZ, Vipin H. Biological and molecular characterization of *Cucumber mosaic virus* Subgroup II isolate causing severe mosaic in cucumber. Indian J Virol. 2013;24:27–34.
- Ronquist F, Huelsenbeck JP. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19:1572–4.
- Roossinck MJ. Cucumber mosaic virus, a model for virus evolution. Mol Plant Pathol. 2001;2:59–63.
- Roossinck MJ, Zhang L, Hellwald KH. Rearrangements in the 5' nontranslated region and phylogenetic analyses of *Cucumber mosaic virus* RNA 3 indicate radial evolution of three subgroups. J Virol. 1999;73:6752–8.
- Roossink MJ. Evolutionary history of Cucumber mosaic virus deduced by phylogenetic analyses. J Virol. 2002;76:3382–7.
- 29. Shintaku MH, Lee Z, Palukaitis P. A single amino-acid substitution in the coat protein of *Cucumber mosaic virus* induces chlorosis in tobacco. Plant Cell. 1992;4:751–7.
- Siju S, Madhubala R, Bhat AI. Sodium sulphite enhances RNA isolation and sensitivity of *Cucumber mosaic virus* by RT- PCR in black pepper. J Virol Methods. 2007;141:107–10.
- Soards AJ, Murphy AM, Palukaitis P, Carr JP. Virulence and differential local and systemic spread of *Cucumber mosaic virus* (CMV) in tobacco are affected by the CMV 2b protein. Mol Plant Microbe Interact. 2002;15:647–53.
- Srivastava A, Raj SK. High molecular similarity among Indian isolates of *Cucumber mosaic virus* suggests a commom origin. Curr Sci. 2004;87:1126–31.
- Suzuki M, Kuwata S, Masuta C, Takanami Y. Point mutations in the coat protein of *Cucumber mosaic virus* affect symptom expression and virion accumulation. J Gen Virol. 1995;76:1791–9.
- 34. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997;24:4876–82.
- Wang Y, Tzfira T, Gaba V, Citovsky V, Palukaitis P, Gal-On A. Functional analysis of the *Cucumber mosaic virus* 2b protein: pathogenicity and nuclear localization. J Gen Virol. 2004;85:3135–47.
- 36. Ziebel H, Murphy AM, Groen SC, Tungadi T, Westwood JH, Lewsey MG, Moulin M, Kleczkowski A, Smith AG, Stevens M, Powell G, Carr JP. *Cucumber mosaic virus* and its 2b RNA silencing suppressor modify plant aphid interactions in tobacco. Sci Rep. 2011;1:187.

