

Complete genome sequence and genetic organization of a Garlic virus D infecting garlic (*Allium sativum*) from northern India

J. SINGH¹, T. N. TRUONG², D. M. AN², M. R. PRAJAPATI¹, A. MANAV¹, N. B. QUOC², K. RANJAN³,
A. KUMAR¹, P. KUMAR¹, R. KUMAR¹, V. K. BARANWAL⁴

¹College of Biotechnology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut, Uttar Pradesh, India-250110; ²Research Institute for Biotechnology and Environment, Nong Lam University, Ho Chi Minh City, Vietnam; ³College of Veterinary and Animal Sciences, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut, Uttar Pradesh, India-250110; ⁴Division of Plant Pathology, ICAR-Indian Agricultural Research Institute, New Delhi, India-110012

Received March 23, 2020; revised April 15, 2020; accepted August 11, 2020

Summary. – The present paper describes first full genome sequence of the Garlic virus D (GarV-D) from northern India with a genome size of 8425 bp long ssRNA. The infected leaves and bulbs of garlic variety Yamuna Safed (G-282) plants suspected for GarV-D infection were collected with the aim to identify contagion virus during March, 2018. The total RNA was extracted from the pooled garlic plants using TRIzol reagent and sequenced using an Illumina HiSeq 2000 platform. BLASTn search in the NCBI database identified contagion as GarV-D (MK518067). It shared 83.63–85.83% nucleotide sequence identities with other (GarV-D) isolates from Argentina (KF550407, KF555653, KR819505) and 83.15% with isolates from China (MF795136, MF363012).

Keywords: *Allium sativum*; Allexivirus; Garlic virus D; India

Introduction

Garlic (*Allium sativum*), a spice or condiment, is used as carminative and gastric stimulant throughout India. The allicin present in aqueous extract of garlic reduces cholesterol concentration in human blood and inhalation of garlic oil or garlic juice is recommended in cases of pulmonary tuberculosis, rheumatism, sterility, impotency, cough and red eyes (Bayan *et al.*, 2014). Garlic extracts also exhibit antibacterial, antifungal, antiviral and platelet activator properties in experimental conditions (Allison *et al.*, 2012; Adler *et al.*, 2002; Adetumbi *et al.*, 1986; Nai-Lan *et al.*, 1993).

India is a major garlic cultivating country and ranks second in total world production. It produces 1617.34

metric tons of garlic from an area of 280.95 thousand hectares with the average production of 5.09 tons/ha (Patidar *et al.*, 2018). Yamuna Safed (G-282) is one of the major varieties of garlic that has creamy white bulbs of about 4.5–6 cm in diameter with 15–16 cloves, mainly grown in Madhya Pradesh, Maharashtra, Haryana, Gujarat, Punjab, Rajasthan, Uttar Pradesh and Chhattisgarh states of India (Mishra and Vikram, 2017).

Production and productivity of garlic is limited by several biotic and abiotic factors, especially viral diseases that are responsible for decreasing yield and diminishing bulb and clove quality (Lunello *et al.*, 2007). Due to sexual sterility, the agamic propagation of garlic is preferred, which results in natural infections by multiple plant viruses. The mixed viral infection in garlic results in mosaic pattern, chlorotic streaking, mottling, and twisting with curling of leaves and stunting of plants, that ultimately leads to the formation of small bulbs and cloves, and reduction in the yield up to the 78% (Conci *et al.*, 2003; Lunello *et al.*, 2007). To control such viral diseases, their molecular level identification is mandatory. The complete

E-mail: jeets80@gmail.com; phone: +91-7017785669.

Abbreviations: CP = coat protein; GarV-D = Garlic virus D; NABP = nucleic acid binding protein; NTBP = viral nucleic acid binding protein, TGB = triple gene block protein

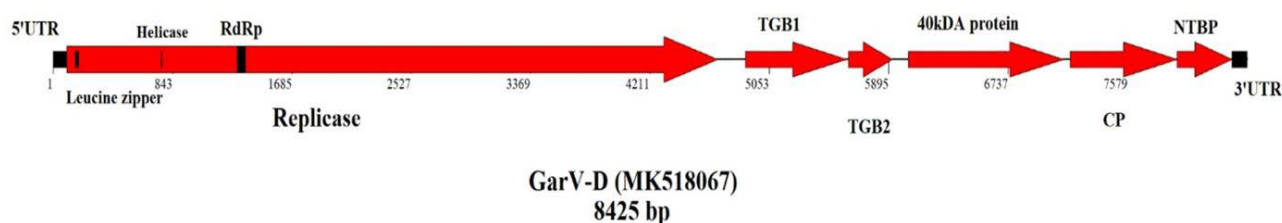


Fig. 1

Genome organization of GarV-D Yamuna safed-3 (MK518067) showing six predicted open reading frames and their corresponding products

Replicase, triple gene block (TGB1 and TGB2), coat protein (CP), nucleic acid binding protein (NTBP).

genome sequence-based study is one of the tools which can be employed for molecular characterization of Garlic virus D (GarV-D).

In India, six garlic viruses belonging to four different taxonomic groups namely, *Allexivirus* (Garlic virus X; Garv-X), *Potyvirus* (Onion yellow dwarf virus; OYDV and Leek yellow stripe virus; LYSV), *Carlavirus* (Garlic common latent virus; GarCLV and Shallot latent virus; SLV), *Tospovirus* (Iris yellow spot virus; IYSV) have been reported to infect garlic (Ghosh and Ahlawat, 1997; Majumder and Baranwal, 2009; Gawande *et al.*, 2010; Baranwal *et al.*, 2011; Gupta *et al.*, 2013). Recently, GarV-D was identified and characterized in India with Allex-NABP (for amplification with generic allexivirus primers) or DCPF/DCPR primers (for amplification with GarV-D specific primers) (Khan *et al.*, 2015, 2016; Chodorska *et al.*, 2014).

Materials and Methods

Sample collection. The samples of infected leaves and bulbs of garlic cultivar Yamuna Safed (G-282) plants suspected for GarV-D infection were collected from Horticulture Research Centre, S. V. Patel University of Agriculture and Technology, Meerut during March, 2018. The collected samples were processed to detect and identify viruses infecting garlic plants.

Virus genome sequencing. To determine the causal agent(s), total RNA was isolated using TRIzol reagent (Chomoczynski, Sacchi, 1987) from a pooled sample of garlic plants. The small RNA deep-sequencing technology (Kreuz *et al.*, 2009) was employed for virus identification. A small RNA library was prepared as described previously (Chen *et al.*, 2012) and libraries were sequenced using an Illumina HiSeq 2000 at NxGenBio Life Sciences, New Delhi. Complete sequence of GarV-D genome identified in this study was deposited in GenBank Acc. No. MK518067 and designated as GarV-D Yamuna safed-3 isolate.

Confirmation of GarV-D by PCR. To reconfirm the GarV-D infection in the garlic plants, PCR primers (GarV-D-7000F 5'-GGGGCATGGTTTGTGTTAAGTT-3' and GarV-DR-8424

(5'-GTCGCGTGGACATAAGTTGTT-3') targeting CP and NTBP region were designed from the sequence with Acc. No. MF363012, to generate ~1200 bp product in RT-PCR. The viral RNA was isolated with Thermo Scientific GeneJET plant RNA purification mini kit (Thermo Scientific, USA) and RT-PCR reaction was performed with Thermo Scientific Revert Aid reverse transcriptase (Thermo Scientific) and DreamTaq DNA polymerase (Thermo Scientific). Thermocycling conditions consisted of 30 min at 52°C, followed by denaturation for 30 s at 94°C, primer annealing for 45 s at 58°C and elongation for 1 min at 72°C for 35 cycles, and a final elongation step of 10 min at 72°C. The PCR amplicons were sequenced in the both directions by Sanger sequencing methods at DNA sequencing facility, Biokart India Pvt Ltd, Bangalore.

Bioinformatics data analysis. To identify possible viruses in garlic sample, ssRNA sequences were assembled and analyzed as previously described (Li *et al.*, 2012). The sequence reads were assembled using CLC Genomic workbench 7.0.4 (CLC Bio) and genome annotated using Blast2GO. The pair wise nucleotide sequence identities were calculated using BioEdit version 7.05.3 software package (Hall *et al.*, 1999) and phylogenetic analysis was done using MEGA X (Kumar *et al.*, 2018).

Results and Discussion

Garlic plants in field conditions might be infected with Garlic virus D (GarV-D) and several other Garlic viruses such as Garlic virus A (GarV-A), Garlic virus B (GarV-B), Garlic virus C (GarV-C), Garlic virus E (GarV-E), Garlic virus X (GarV-X) etc. (King *et al.*, 2012). In current study, upon high throughput sequencing sample yielded a total number of 34 873 264 reads and after trimming 199 718 reads were mapped to the reference sequence submitted from China (Acc. No. MF363012).

In this study, for GarV-D, 7809 bp long near complete genome in form of single contig was obtained. The 5' and 3'UTR sequences were amplified by primers designed from the previously available GarV-D genome using

Table 1. Pair wise comparison of nucleotide sequences of complete genome of GarV- D (Acc. No. MK518067) with other isolates of garlic viruses

	MK518067 GarV-D India	KF550407 GarV-Argentina	KF555653 GarV-Argentina	KR819505 GarV-Argentina	MF795136 GarV-D China: Xm	MF363012 GarV-D China-XM	MK503771 GarV-X India	AJ292229 GarV-X-YH China	JQ807994 GarV-X Australia	GVU89243 GarV-X Korea	JX429969 GarV-X SG4 Spain	NC003375 GarV-A Japan	AJ292230 GarV-E China	AB010302 GarV-C Japan	NC003795 ShVX Russia
MK518067 GarV-D India	ID														
KF550407 GarV-D- Argentina	85.5	ID													
KF555653 GarV-D Argentina	83.7	84	ID												
KR819505 GarV-D Argentina	83.3	86.3	88.8	ID											
MF795136 GarV-D China: Xm	82.9	82.4	86.1	83	ID										
MF363012 GarV-D China-Xm	82.9	82.4	86.1	83	1	ID									
MK503771 GarV-X India	59.2	59.9	58.9	59.3	59	59	ID								
AJ292229 GarV-X-YH China	60.6	60.7	60.3	60.9	60.5	60.5	85.4	ID							
JQ807994 GarV-X Australia	60	60.2	59.7	60.1	59.9	59.9	85.6	93.7	ID						
GVU89243 GarV-X Korea	59.6	59.5	58.8	59.6	59.4	59.4	83.8	87.9	86.5	ID					
JX429969 GarV-X SG4 Spain	60	60.1	59.7	60.3	59.9	59.9	83.6	91.8	90.6	90.7	ID				
NC003375 GarV-A Japan	61.7	61.3	61.8	61.8	61.6	61.6	55.3	56	55.6	55.3	55.7	ID			
AJ292230 GarV-E China	69.2	69.4	70	69.6	69.8	69.8	59.4	60	59.7	59.1	59.5	60.9	ID		
AB010302 GarV-C Japan	60.6	60.5	60.2	60.7	60.7	60.7	63.7	64.8	63.7	63.7	64	55.8	60.5	ID	
NC003795 ShVX Russia	57.5	57.2	57.7	57.3	57.3	57.3	53	54.1	53.5	53.2	53.6	63.1	56.9	54.7	ID

Table 2. Percentage amino acid sequence identity of pair wise combination of replicase, TGB 1, TGB 2, ORF 4 (serine rich protein), coat protein, ORF6 (nucleic acid binding protein) of GarV-D (Acc. No. MK518067) with other complete sequences of GarV-D isolates

Garlic virus D (GarV-D) isolates	MK518067 GarV-D India	KF550407 GarV-D-SW9 Argentina	KF555653 GarV-D-SW10 Argentina	KR819505 GarV-D Argentina	MF363012 GarV-D China	MF795136 GarV-D China: Xiamen	MK518067 GarV-D India	KF550407 GarV-D-SW9 Argentina	KF555653 GarV-D-SW10 Argentina	KR819505 GarV-D Argentina	MF363012 GarV-D China	MF795136 GarV-D China: Xiamen
	replicase						triple gene block protein 1					
MK518067 GarV-D India	ID						ID					
KF550407 GarV-D-SW9 Argentina	92.9	ID					86.4	ID				
KF555653 GarV-D-SW10 Argentina	91.5	91.5	ID				83.9	86.9	ID			
KR819505 GarV-D Argentina	91.4	90.9	95.7	ID			86.9	91.5	86.9	ID		
MF363012 GarV-D China	91.4	91.8	91.4	91.6	ID		84.8	86	89	86.9	ID	
MF795136 GarV-D China: Xiamen	91.4	91.8	91.4	91.6	100	ID	84.8	86	89	86.9	100	ID
	triple gene block protein 2						ORF 4 (serine rich protein)					
MK518067 GarV-D India	ID						ID					
KF550407 GarV-D-SW9 Argentina	89.3	ID					78.0	ID				
KF555653 GarV-D-SW10 Argentina	90.2	88.3	ID				80.7	75.5	ID			
KR819505 GarV-D Argentina	89.3	96.1	91.2	ID			76.6	95.0	75.2	ID		
MF363012 GarV-D China	88.3	90.2	88.3	90.2	ID		79.3	73.6	88.4	73.3	ID	
MF795136 GarV-D China: Xiamen	88.3	90.2	88.3	90.2	100	ID	79.3	73.6	88.4	73.3	100	ID
	coat protein						ORF6 (nucleic acid binding protein)					
MK518067 GarV-D India	ID						ID					
KF550407 GarV-D-SW9 Argentina	96.8	ID					91.4	ID				
KF555653 GarV-D-SW10 Argentina	95.2	95.6	ID				89.8	92.9	ID			
KR819505 GarV-D Argentina	96.8	96.4	96.0	ID			85.9	91.4	87.5	ID		
MF363012 GarV-D China	96.8	96.4	96.0	100	ID		89.8	94.5	93.7	89.0	ID	
MF795136 GarV-D China: Xiamen	94.8	95.2	96.8	94.8	94.8	ID	89.8	94.5	93.7	89.0	100	ID

protein (NABP) genes (Table 2). Similarly, GarV-D Yamuna safed-3 (MK518067) also showed amino acid identities of 91.4–92.9% in replicase, 83.9–86.4% in TGB1, 88.3–90.2% in TGB2 and 76.6–80.7% in serine rich protein regions with other GarV-D viruses (Table 2).

The 5' and 3'UTR sequence of GarV-D Yamuna safed-3 shared 97.3% nucleotide sequence identity with other GarV-D isolates from Argentina (KR819505, KF550407); China (MF795136, MF363012) and UK (L38892); 96.4% with Korea (AF519572), Argentina (KF555653) and Japan (AB010303) isolates and 89.3% with Spain (KF632716), Australia (JX997951), and Japan (AB010300) isolates. In 5' and 3'UTR regions highly conserved regions were perceived. Phylogenetic analysis of 5' and 3'UTR sequences indicates very low diversity.

From the preliminary sequence assemblies and analyses, in addition to a near complete sequence of GarV-D, several other viruses, including Garlic virus X (GarV-X), Garlic common latent virus (GarCLV) and Leek yellow stripe virus (LYSV) were also identified. Owing to mixed infection, the disease symptoms observed could not be attributed to a single virus alone.

The garlic plants, which were tested positive for GarV-D, were again tested for GarV-D using primers set GarV-D-7000F/ GarV-D-8424 R and revealed an amplification of ~1200 bp. The nucleotide sequences of CP and NTBP region (1223 bp) were submitted to GenBank under the Acc. No. MN996259. In Blast analysis, the sequence showed maximum nucleotide identity of 81.10% with 93.5% query coverage to GarV-D isolate-591 from Spain (KX889819), GarV-D isolate 619 from China (KX889806) and 87.2% identity with GarV-D from Korea (AF519572) with 97% query coverage. GarV-D, Meerut (MN996259) showed an identity of 76.2–76.6% at nucleotide level and 89.9% at amino acid level with partial sequence of CP and NTBP region submitted from Ludhiana, Punjab, India under Acc. No. KP862055 and KR534889, respectively.

The virus isolates identified from different regions of the globe showed that all known isolates have diverged from a common ancestor isolates. In due course of time, they might have spread to different geographical regions by different means such as transport of biological material of plant origin, etc. (Bereda *et al.*, 2015, 2017). GarV-D in garlic plants was earlier reported in Pacific Northwest (Gieck *et al.*, 2009), Democratic Republic of Congo (Majumder *et al.*, 2018), Poland (Chodorska *et al.*, 2013; Bereda *et al.*, 2015) and Australia (Wylie *et al.*, 2014).

Based on phylogenetics, genomics analysis and available information in scientific databases, to the best of our knowledge, this is first report of the complete genome sequence of Garlic virus D from *Allium sativum* variety Yamuna Safed (G-282) from India. This report will encourage researchers to investigate impact of this virus

on garlic production in India and reinforce the need of a garlic-seed certification programme to improve the health conditions of the crop. However, biological studies are needed to identify mechanisms of transmissions and to assess effects of single/mixed infection on garlic plants.

Acknowledgments. The authors acknowledge the Vice Chancellor, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250110, Uttar Pradesh, Bioinformatics facility, Department of Biotechnology, India and ASEAN-India Science, Technology & Innovation Cooperation, Department of Science & Technology, India for providing financial support and the facilities to carry out this research work.

References

- Adetumbi M, Javor GT, Lau BH (1986): *Allium sativum* (garlic) inhibits lipid synthesis by *Candida albicans*. *Antimicrob. Agents Chemother.* 30, 499–501. <https://doi.org/10.1128/AAC.30.3.499>
- Adler BB, Beuchat LR (2002): Death of *Salmonella*, *Escherichia coli* O157:H7, and *Listeria monocytogenes* in garlic butter as affected by storage temperature. *J. Food Prot.* 65, 1976–1980. <https://doi.org/10.4315/0362-028X-65.12.1976>
- Allison GL, Lowe GM, Rahman K (2012): Aged garlic extract inhibits platelet activation by increasing intracellular cAMP and reducing the interaction of GPIIb/IIIa receptor with fibrinogen. *Life Sci.* 91, 1275–1280. <https://doi.org/10.1016/j.lfs.2012.09.019>
- Altschul S, Gish W, Miller W, Myers EW, Lipman DJ (1990): Basic Local Alignment Search Tool. *J. Mol. Biol.* 215, 403–441. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Baranwal VK, Singh P, Jain RK, Joshi S (2011): First report of garlic virus X infecting garlic in India. *Plant Dis.* 95, 1197. <https://doi.org/10.1094/PDIS-03-11-0198>
- Bayan L, Koulivand PH, Gorji A (2014): Garlic: a review of potential therapeutic effects. *Avicenna J. Phytomed.* 4(1), 1–14.
- Bereda M, Kalinowska E, Paduch-Cichal E, Szyndel MS (2015): Low genetic diversity of a natural population of Garlic virus D from Poland. *Eur. J. Plant Pathol.* 142, 411–417. <https://doi.org/10.1007/s10658-015-0615-x>
- Bereda M, Paduch-Cichal E, Dabrowska E (2017): Occurrence and phylogenetic analysis of allel viruses identified on garlic from China, Spain and Poland commercially available on the polish retail market. *European J. Plant Pathol.* 149, 227–237. <https://doi.org/10.1007/s10658-017-1171-3>
- Celli MG, Perotto MC, Buraschi D and Conci VC (2016): Biological and molecular characterization of Garlic virus D and its effects on yields of garlic. *Acta Hort.* 1143, 193–200. <https://doi.org/10.17660/ActaHortic.2016.1143.28>
- Chen YR, Zheng Y, Liu B, Zhong S, Giovannoni J, Fei Z (2012): A cost-effective method for Illumina small RNA-

- Seq library preparation using T4 RNA ligase 1 adenylated adapters. *Plant Methods* 8, 41. <https://doi.org/10.1186/1746-4811-8-41>
- Chen J, Adams MJ (2001): Molecular characterisation of a complex mixture of viruses in garlic with mosaic symptoms in China. *Archives Virol.* 146, 1841-1853. <https://doi.org/10.1007/s007050170037>
- Chodorska M, Paduch-Cichal E, Kalinowska E, Szyndel MS (2013): Occurrence of the viruses belonging to the *Allexivirus* genus on garlic plants in Poland. *Prog. Plant Prot.* 53, 605-609. <https://doi.org/10.14199/ppp-2013-070>
- Chodorska M, Paduch-Cichal E, Kalinowska E, Szyndel MS (2014): Assessment of *allexiviruses* infection in garlic plants in Poland. *Acta Sci. Pol.* 13, 179-186.
- Chomczynski P, Sacchi N (1987): Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156-159. [https://doi.org/10.1016/0003-2697\(87\)90021-2](https://doi.org/10.1016/0003-2697(87)90021-2)
- Conci VC, Canavelli A, Lunello P (2003): Yield losses associated with virus-infected garlic plants during five successive years. *Plant Dis.* 87, 1411-1415. <https://doi.org/10.1094/PDIS.2003.87.12.1411>
- Gawande SJ, Khar A, Lawande KE (2010): First report of Iris yellow spot virus on garlic in India. *Plant Dis.* 94, 1066. doi:10.1094/PDIS-94-8-1066C. <https://doi.org/10.1094/PDIS-94-8-1066C>
- Ghosh DK, Ahlawat YS (1997): Filamentous viruses associated with mosaic disease of garlic in India. *Indian Phytopathol.* 50, 266-267.
- Gieck, S L, Hamm PB, David NL, Pappu HR (2009): First report of Garlic virus B and Garlic virus D in garlic in the Pacific northwest. *Plant Dis.* 93, 431. <https://doi.org/10.1094/PDIS-93-4-0431A>
- Gupta N, Prabha K, Islam S, Baranwal VK (2013): First report of Leek yellow stripe virus in India from garlic. *J. Plant Pathol.* 95, 69-77.
- Hall TA (1999): BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95-98.
- Khan I, Sharma A, Dhatt AS, Kang SS, Kaur G (2015): First report of Garlic virus D infecting onion in India. *J. Plant Pathol.* 97, S69.
- Khan I, Sharma A, Kang SS, Kaur G, Dhatt A (2016): Molecular characterization and identification of Garlic virus D associated with garlic in India. *J. Plant Pathol.* 98, 607-609. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012): *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses.* London, Waltham: Academic Press.
- Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, Simon R (2009): Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388, 1-7. <https://doi.org/10.1016/j.virol.2009.03.024>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018): MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547-1549. <https://doi.org/10.1093/molbev/msy096>
- Li R, Gao S, Hernandez AG, Wechter WP, Fei Z, Ling K-S (2012): Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLoS One* 7, e37127. <https://doi.org/10.1371/journal.pone.0037127>
- Lunello P, Di Rienzo J, Conci VC (2007): Yield loss in garlic caused by leek yellow stripe virus Argentinean isolate. *Plant Dis.* 91, 153-158. <https://doi.org/10.1094/PDIS-91-2-0153>
- Majumder S, Baranwal VK (2009): First report of Garlic common latent virus in garlic from India. *Plant Dis.* 93, 106. <https://doi.org/10.1094/PDIS-93-1-0106C>
- Majumder S, Mbay K, Singh J (2018): First report of garlic virus D in garlic from D.R. Congo. *J. Plant Pathol.* 100, 143. <https://doi.org/10.1007/s42161-018-0033-y>
- Mishra TD, Vikram B (2017): Evaluation of Garlic (*Allium sativum* L.) germplasms for yield potential and quality characters under Allahabad agro-climatic conditions. *J. Pharmacogn. Phytochem.* 6, 433-436.
- Nai-Lan G, Cao-Pei L, Woods GL, Reed E, Gui-Zhen Z, Li-Bi Z, Waldman RH (1993): Demonstration of antiviral activity of garlic extract against human cytomegalovirus in vitro. *Chin. Med. J.* 106, 93-96.
- Patidar PK, Khan N, Kumar S (2018): An economic analysis of Garlic cultivation in Ratlam district of Madhya Pradesh. *Int. J. Agric. Environ. Biotechnol.* 11, 371-377.
- Wylie SJ, Li H, Saqib M, Jones MG (2014): The global trade in fresh produce and the vagility of plant viruses: a case study in garlic. *PLoS One* 9, e105044. <https://doi.org/10.1371/journal.pone.0105044>