

Characterization of siRNAs derived from cucumber green mottle mosaic virus in infected cucumber plants

Yongqiang Li¹ · Congliang Deng² · Qiaoxia Shang¹ · Xiaoli Zhao² · Xingliang Liu² · Qi Zhou³

Received: 3 August 2015 / Accepted: 13 November 2015 / Published online: 19 November 2015
© Springer-Verlag Wien 2015

Abstract Virus-derived small interfering RNAs (vsiRNAs) of cucumber green mottle mosaic virus (CGMMV), a member of the genus *Tobamovirus*, were characterised in cucumber plants by deep sequencing. CGMMV vsiRNAs of 21–22 nt in length predominated, suggesting that there might be a conserved mechanism of DCL2 and DCL4 involvement in the biogenesis of vsiRNAs, as well as a common RNA silencing pathway in CGMMV-infected cucumber plants. The 5′-terminal base of vsiRNAs was biased towards C/A/U, suggesting that CGMMV vsiRNAs might be loaded into diverse AGO-containing RISCs to disturb the gene expression of host plants. Possible targets for some of the vsiRNAs were also predicted.

RNA interference (RNAi) is a conserved mechanism in eukaryotes that protects genomes against aberrant endogenous or exogenous RNA molecules [3, 6, 7, 35]. This defence mechanisms involves the processing of double-stranded RNA (dsRNA) into small interfering (si)

RNAs by RNase III Dicer-like proteins (DCLs) and the formation of an RNA-induced silencing complex mediated by proteins of the Argonaute (AGO) family to facilitate the cleavage of target RNAs through base-pairing mechanisms [3, 4, 7, 12]. Virus-derived small interfering RNAs (vsiRNAs) are key elements in plant antiviral strategies [8].

During infection, plant viruses can activate the RNA silencing machinery in infected cells through the formation of viral dsRNA by the activity of virus-encoded RNA-dependent RNA polymerase (RdRp), intermolecular base pairing between plus- and minus-strand viral RNAs, and imperfect folding of self-complementary sequences within viral ssRNA [8], and the production of vsiRNAs requires DCLs, AGOs, and RdRps [5, 31]. There are at least four DCL proteins in *Arabidopsis* (DCL1–4), with distinct functions [19]. DCL4, DCL2 and DCL3 target viral genomes in a hierarchical fashion to yield vsiRNAs of 21, 22 and 24 nt, respectively. Multiple AGO genes might be involved in antiviral defense, and the 5′ nucleotide of the siRNA may dictate the binding of a specific AGO protein to siRNAs [20, 23].

Cucumber green mottle mosaic virus (CGMMV, genus *Tobamovirus*, family *Virgaviridae*) causes serious damage to cucurbit crops worldwide and can be transmitted mechanically and through seed [1, 18, 30]. It causes different symptoms depending on the host species, including severe mosaic with discoloration and deformation of leaves [2, 16]. Like other tobamoviruses, CGMMV has a single-stranded, positive-sense RNA genome with a 3′ tRNA-like structure instead of a poly(A) tail [1]. The genomes of tobamoviruses encode four polypeptides: a 124- to 132-kDa protein, a 181- to 189-kDa readthrough protein, a 28- to 31-kDa movement protein (MP) and a 17- to 18-kDa coat protein (CP) [1, 13]. Both large proteins are translated from the 5′-proximal start codon. The two small proteins are expressed from individual 3′ co-terminal subgenomic

Electronic supplementary material The online version of this article (doi:10.1007/s00705-015-2687-5) contains supplementary material, which is available to authorized users.

✉ Yongqiang Li
lyq047@163.com; lyq@bua.edu.cn

✉ Qi Zhou
zhouq@bjciq.gov.cn

¹ College of Plant Science and Technology, Beijing University of Agriculture, Beijing 102206, China

² Plant Laboratory of Beijing Entry-Exit Inspection and Quarantine Bureau, Beijing 101312, China

³ Chinese Society of Inspection and Quarantine, Beijing 100029, China

mRNAs. Very little is known about the RNA silencing pathways and virus-host interactions in cucumber. The vsiRNAs could mediate RNA silencing in specific antiviral immunity and, interestingly, a few recent reports have shown that some of these vsiRNAs can guide the degradation of homologous cellular transcripts using base-pairing mechanisms to create cellular conditions suitable for viral proliferation or to induce disease symptoms [26–28].

Recent high-throughput sequencing of vsiRNAs in different host-virus systems, along with functional characterization, has provided insights into the origin and composition of vsiRNAs and their potential role in virus-host interactions [24]. In the genus *Tobamovirus*, the global impact of oilseed rape mosaic virus infection on Arabidopsis siRNAs and their mRNA targets was determined using this method combined with transcriptome profiling [15]. The biochemical framework for small-RNA production and RNA silencing is conserved in plants [29]; thus, in this study we have characterized the vsiRNAs of CGMMV from infected cucumber plants using deep sequencing to investigate the interaction between CGMMV and its host plant.

Cucumber seedlings were kept at room temperature under insect-proof netting and inoculated with CGMMV (or phosphate buffer as a control) at the three-true-leaf stage. The CGMMV inoculum was prepared from infected cucumber leaves obtained from the Plant Laboratory of the Beijing Entry-Exit Inspection and Quarantine Bureau. Leaf samples were collected from CGMMV-inoculated or negative control plants at 14 dpi, and the presence of the virus was confirmed by RT-PCR using the primer pair 5'-GGCTTACAATCCGATCACACCT-3' and 5'-CTAAGCTTTCGAGGTGGTAGCCT-3'. Total RNA was extracted from CGMMV-infected plants using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Total RNA concentrations were estimated using a spectrophotometer (Nanodrop ND-2000, Thermo Fisher Scientific, Wilmington, DE, USA), and the RNA sample integrity was verified using a Bio-Analyser 2100 (Agilent Technologies, Waldbronn, Germany). Small RNA libraries were prepared using the Solexa protocol (<http://www.illumina.com>). Briefly, the total RNA was size-fractionated by 15 % polyacrylamide gel electrophoresis (PAGE), and small RNAs of 18–28 nt were isolated. Then, the purified small RNAs were ligated with 5' and 3' adaptors at both ends. After purification by PAGE, the final ligation products were reverse transcribed and a cDNA library was constructed. Finally, the small RNA libraries were sequenced using the Illumina HiSeq-q2000 platform. After removal of the adapter sequences and low-quality reads, the remaining 18- to 28-nt sRNAs were aligned with the CGMMV reference genome sequence (GenBank accession number D12505.1). Reads showing zero mismatches with the CGMMV reference

sequence were retained for subsequent analysis. Library characterization and determination of the CGMMV siRNA profile were performed using locally developed Perl scripts. Putative targets of the vsiRNAs were identified with the Patscan [11] program with the cucumber genome *Cucumis sativus v1.0* in Phytozome 10.3.

A total of 28,980,811 and 29,880,020 sRNAs were sequenced from the systemic leaves of CGMMV-infected and virus-free plants, respectively. Small RNAs with 21, 22 and 24 nt were dominant in both libraries, but in the CGMMV-infected library, 21- and 22-nt reads increased significantly, while 24-nt reads decreased (Fig. 1A), which was also observed in a previous report [32]. In total, 5,415,828 vsiRNA reads were identified in CGMMV-inoculated cucumber plants, accounting for 18.70 % of 18- to 28-nt reads. Conversely, only 1,012 reads matched the CGMMV genome in the mock-inoculated library, which corresponded to approximately 0.003 % of 18- to 28-nt reads. In CGMMV-infected cucumber plants, 21- and 22-nt vsiRNAs accumulated to high levels (Fig. 1B), representing 52.01 % and 30.53 % of total vsiRNAs, respectively, which suggested that the cucumber

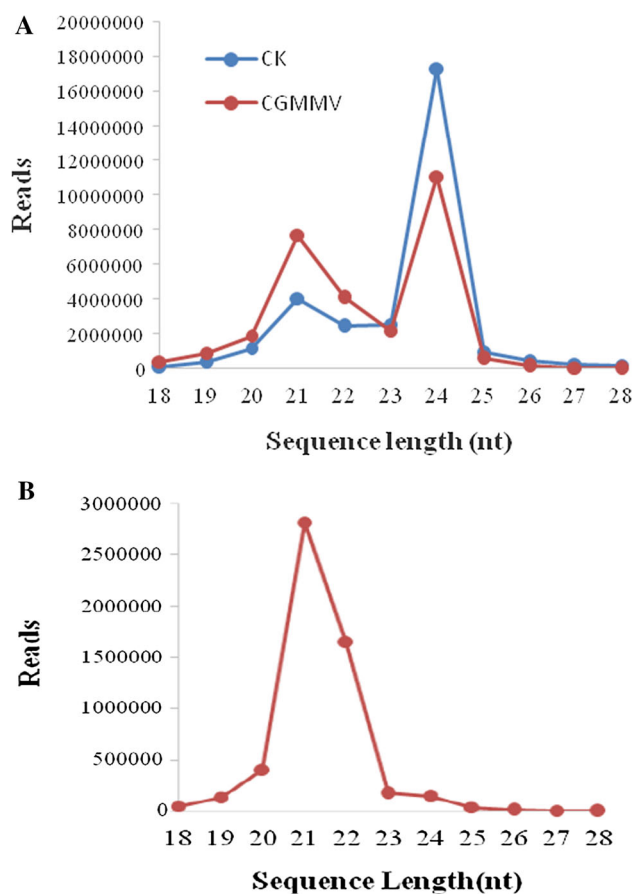


Fig. 1 Size distribution of total small RNAs in libraries prepared from either phosphate-buffer-inoculated (CK) or CGMMV-inoculated cucumber plants (A) and vsiRNAs in a CGMMV-inoculated library (B)

homologues of DCL4 and DCL2 may be the predominant Dicer-like ribonucleases involved in vsiRNA biogenesis [9]. The predominance of 21- to 22-nt sequences amongst vsiRNAs has been reported previously for several plant viruses [21, 33, 34] and suggests that the 21-nt siRNA is the predominant antiviral silencing component and DCL4 (responsible for the production of 21-nt siRNA) is the major producer of viral siRNAs [4, 9, 22, 25].

Previous studies have indicated that the 5'-terminal nucleotides of siRNAs have a pivotal role in directing the siRNAs to specific AGO complexes [20]. In contrast to the observations for diverse plant-virus-specific small RNAs, which display a clear tendency to begin with uracil (U) or adenine (A) [10], CGMMV siRNAs with a cytosine (C) at their 5' termini were the most abundant, accounting for 30.85 % of all vsiRNAs, while the percentage of A, U, and G residues at their 5' termini was 26.58 %, 25.64 %, and 19.93 %, respectively (Fig. 2A). To gain further insights into CGMMV siRNA biogenesis and sorting, the sequence complexity of these small RNAs was analyzed for different-sized species. For 21-, 22-, and 23-nt CGMMV siRNAs, a clear preference for C as the 5'-terminal nucleotide was observed, indicative of vsiRNAs with high binding affinity for the AGO5 homologue in cucumber; however, a strong bias for sequences beginning with a 5'-A was observed in the 24-nt vsiRNAs (Fig. 2A), indicative of the high binding affinity of AGO2 and AGO4 homologues for 24-nt sRNAs. The low proportion of vsiRNAs beginning with a G in our datasets is consistent with previous reports [10, 20, 21].

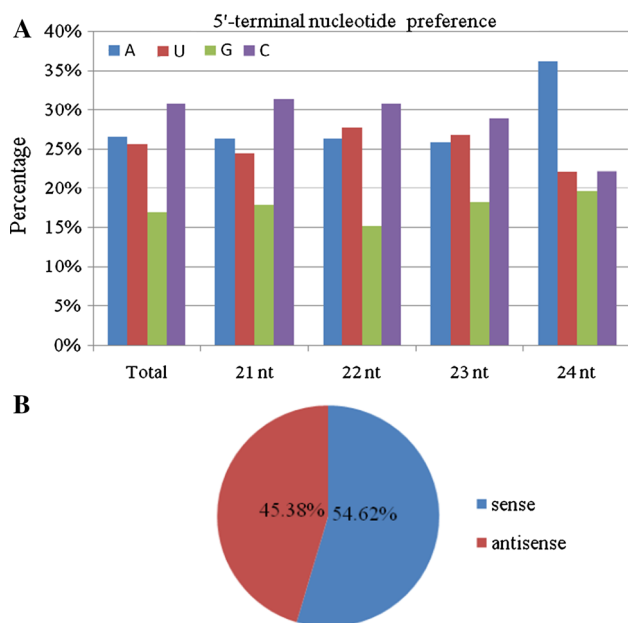


Fig. 2 Relative frequency of the 5'-terminal nucleotide of vsiRNAs (A) and the polarity distribution of vsiRNAs (B). The total represents all of the CGMMV-derived vsiRNAs

To explore the origin of the vsiRNAs, their polarity distribution was further characterized. There was a slight bias towards CGMMV siRNAs derived from the sense strand (54.62 %) of the viral RNA (Fig. 2B). This is broadly similar to the results obtained for turnip mosaic virus [14] and sugarcane mosaic virus [32], where siRNAs were produced equally from the two strands, supporting the hypothesis that viral dsRNAs formed during replication are the substrates of DCL proteins. One exception is tobacco rattle virus, in the same family as CGMMV, for which vsiRNAs were reported to be derived predominantly from the viral positive-strand RNA [9], but the reason behind this observed asymmetry is not clear. Additionally, single-base resolution maps of all redundant CGMMV siRNAs along with the genomes were created using Bowtie tools [17] and in-house Perl scripts. Our results showed that CGMMV siRNAs covered almost the entire virus genome; however, there were two hotspots located at the 5' and 3' termini (Fig. 3A and B, respectively).

The predicted host target genes of CGMMV vsiRNAs that were represented in the virus-infected library by more than 200 reads totaled 92 targets predicted using the Patscan software (Table S1). These predicted targets are involved in a broad range of biological processes, such as biological regulation, cellular structure, cellular processes, localisation, metabolic processes (such as fructose and mannose metabolism, folate biosynthesis), and responses to stimuli, suggesting that vsiRNAs may play important roles due to their interactions with their targets during CGMMV infection.

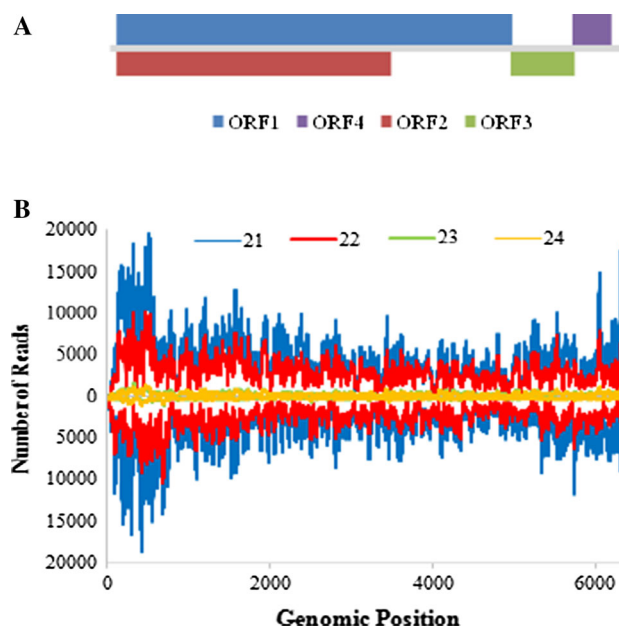


Fig. 3 Genome map of CGMMV (A) and maps of 21- to 24-nt vsiRNAs from CGMMV-inoculated cucumber plants at single-nucleotide resolution (B). The greyish line represents the CGMMV genome, and boxes with different colors represented the different ORFs of CGMMV shown in panel A

Acknowledgment This research was financially supported by the support program of General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China (201310068), National Natural Science Foundation of China (Grant No. 31401711) and a grant from Beijing University of Agriculture (BNRC&YX201415).

References

- Adams MJ, Heinze C, Jackson AO, Kreuzer JF, Macfarlane SA, Torrance L (2012) Family *Virgaviridae*. In: King AMO, Lefkowitz E, Adams MJ, Carstens EB (eds) *Virus taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier, London, pp 1139–1160
- Ali A, Natsuaki T, Okuda S (2004) Identification and molecular characterization of viruses infecting cucurbits in Pakistan. *J Phytopathol* 152:677–682
- Baulcombe D (2004) RNA silencing in plants. *Nature* 431:356–363
- Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park HS, Vazquez F, Robertson D, Meins F Jr, Hohn T, Pooggin MM (2006) Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res* 34:6233–6246
- Brodersen P, Voinnet O (2006) The diversity of RNA silencing pathways in plants. *Trends Gen* 22:268–280
- Csorba T, Pantaleo V, Burgyan J (2009) RNA silencing: an antiviral mechanism. *Adv Virus Res* 75:35–71
- Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313:68–71
- Ding SW, Voinnet O (2007) Antiviral immunity directed by small RNAs. *Cell* 130:413–426
- Donaire L, Barajas D, Martinez-Garcia B, Martinez-Priego L, Pagan I, Llave C (2008) Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. *J Virol* 82:5167–5177
- Donaire L, Wang Y, Gonzalez-Ibeas D, Mayer KF, Aranda MA, Llave C (2009) Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology* 392:203–214
- Dsouza M, Larsen N, Overbeek R (1997) Searching for patterns in genomic data. *Trends Genet* 13:497–498
- Elbashir SM, Lendeckel W, Tuschl T (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* 15:188–200
- Goelet P, Lomonosoff GP, Butler PJ, Akam ME, Gait MJ, Karn J (1982) Nucleotide sequence of *Tobacco mosaic virus* RNA. *Proc Natl Acad Sci USA* 79:5818–5822
- Ho T, Wang H, Pallett D, Dalmay T (2007) Evidence for targeting common siRNA hotspots and GC preference by plant Dicer-like proteins. *FEBS Lett* 581:3267–3272
- Hu Q, Hollunder J, Niehl A, Kørner CJ, Gereige D, Windels D, Arnold A, Kuiper M, Vazquez F, Pooggin M, Heinlein M (2011) Specific impact of tobamovirus infection on the Arabidopsis small RNA profile. *PLoS One* 6:e19549
- Kim OK, Mizutani T, Natsuaki T, Lee KW, Soe K (2010) First report and the genetic variability of *Cucumber green mottle mosaic virus* occurring on bottle gourd in Myanmar. *J Phytopathol* 158:572–575
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25
- Lee SY, Win NKK, Cho DM, Lee SH, Jung HY (2012) *Cucumber green mottle mosaic virus* (CGMMV) can induce hair-like tissues on genus *cucumis* seeds. *Sci Hortic* 146:76–80
- Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM (2006) The evolution and diversification of Dicers in plants. *FEBS Lett* 580:2442–2450
- Mi SJ, Cai T, Hu YG, Chen YM, Hodges E, Ni FR, Wu L, Li S, Zhou HY, Long CZ, Chen S, Hannon GJ, Qi YJ (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5'-terminal nucleotide. *Cell* 133:116–127
- Mitter N, Koundal V, Williams S, Pappu H (2013) Differential expression of *Tomato spotted wilt virus*-derived viral small RNAs in infected commercial and experimental host plants. *PLoS One* 8(10):e76276
- Molnár A, Csorba T, Lakatos L, Varallyay E, Lacomme C, Burgyan J (2005) Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. *J Virol* 79:7812–7818
- Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of argonaute 7-miR390 interaction and dual functionality in TAS3 transacting siRNA formation. *Cell* 133:128–141
- Prabha K, Baranwal VK, Jain RK (2013) Applications of next generation high throughput sequencing technologies in characterization, discovery and molecular interaction of plant viruses. *Ind J Virol* 24:157–165
- Qi XP, Bao FS, Xie ZX (2009) Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS One* 4:e4971
- Ruiz-Ferrer V, Voinnet O (2009) Roles of plant small RNAs in biotic stress responses. *Annu Rev Plant Biol* 60:485–510
- Shimura H, Pantaleo V, Ishihara T, Myojo N, Inaba J, Sueda K, Burgyan J, Masuta C (2011) A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. *PLoS Pathog* 7:e1002021
- Smith NA, Eamens AL, Wang MB (2011) Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathog* 7:e1002022
- Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. *Genes Dev* 17:49–63
- Varveri C, Vassilakos N, Bem F (2002) Characterization and detection of *Cucumber green mottle mosaic virus* in Greece. *Phytoparasitica* 30:493–501
- Wang XB, Wu QF, Ito T, Cillo F, Li WX, Chen X, Yu JL, Ding SW (2010) RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 107:484–489
- Xia ZH, Peng J, Li YQ, Chen L, Li S, Zhou T, Fan ZF (2014) Characterization of small interfering RNAs derived from *Sugarcane mosaic virus* in infected maize plants by deep sequencing. *PLoS One* 9:e97013
- Yan F, Zhang HM, Adams MJ, Yang J, Peng JJ, Antoniw JF, Zhou YJ, Chen JP (2010) Characterization of siRNAs derived from *Rice stripe virus* in infected rice plants by deep sequencing. *Arch Virol* 155:935–940
- Yang J, Zheng SL, Zhang HM, Liu XY, Li J, Li JM, Chen JP (2014) Analysis of small RNAs derived from *Chinese wheat mosaic virus*. *Arch Virol* 159:3077–3082
- Zamore PD (2002) Ancient pathways programmed by small RNAs. *Science* 296:1265–1269