Genome sequences of 10 new carnation mottle virus variants

Breit, T.M.; Leeuw, W. de; van Olst, M.; Ensink, W.A.; van Leeuwen, S.; Jonker, M.J.; Dekker, R.J.

Link to publication

Genome sequences of 10 new carnation mottle virus variants

Timo M. Breit,[1] Wim de Leeuw,[1] Marina van Olst,[1] Wim A. Ensink,[1] Selina van Leeuwen,[1] Martijs J. Jonker,[1] Rob J. Dekker[1]

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT Here, we report the genome sequences of 10 Carnation mottle virus variants. Six variants originated from a single proprietary carnation cultivar, and four were derived from four different proprietary cultivars. All variants showed nucleotide differences, but the last four did not show any variation at the amino acid level.

Carnation mottle virus (CarMV, genus: *Alphacarmovirus*, family: *Tombusviridae*) is a single-stranded positive-sense RNA virus with four open reading frames (ORFs) (1). Infection of the major host for CarMV, *Dianthus caryophyllus* (carnation), can result in mild mottling in young leaves but typically only faint chlorosis in mature leaves. Besides the CarMV genome reference sequence NC_001265.2, just 11 complete variant genome sequences are present in Genbank (2, 3).

To gain deeper understanding of the CarMV genetic variation, we conducted small-RNA sequencing (sRNA-seq) on 10 asymptomatic carnation plants from various undisclosed commercial cultivars sourced from an anonymous plant-breeding company in the Netherlands. In carnation experiment 1 (CE1), six out of eight plants belonging to the same confidential cultivar were found positive for CarMV enzyme-linked immunosorbent assay (ELISA). In carnation experiment 2 (CE2), four distinct confidential cultivars that tested positive for CarMV ELISA were used. All samples were collected in February 2020 from a greenhouse complex situated in the southern region of the North-Holland province in the Netherlands. Genomes sequences of possible CarMV variants were determined by examining the siRNA response of the plants (4).

The ELISA assay was performed using the Prime Diagnostics (Wageningen, the Netherlands) CarMV Kit (https://shop.wur.nl/primediagnostics/carnation-mottle-virus-carmv.html) following the protocol provided by the manufacturer. Small-RNA was isolated by grinding a flash-frozen ±1 cm² leaf fragment to fine powder using mortar and pestle, dissolving the powder in QIAzol Lysis Reagent (Qiagen) and enriching small-RNA using the miRNeasy Mini Kit (Qiagen). Barcoded small-RNAseq libraries were generated using Ion Total RNA-Seq Kit v2 and Ion Xpress RNA-Seq Bar Coding Kit (Thermo Fisher Scientific). Sequencing (150 bases) with an Ion Proton System using Ion PI-v3 chips and a Ion Chef System using Ion PI-Hi-Q-Chef Kit (Thermo Fisher Scientific) CE1 and CE2 yielded on average 5 (CE1) and 31 (CE2) million reads per sample.

The CarMV variant genomes were constructed using NC_001265.2 as a reference. First, reads were trimmed using trimmomatic v0.39 (5) [parameters: LEADING:3;TRAILING:3;SLIDINGWINDOW:4:15;MINLEN:19] and mapped to the reference using bowtie2-v2.4.1 (6). Next, Freebayes-v1.3.2 (7) was used to extract variations which were incorporated in the reference using gatk's FastaAlternateReferenceMaker-v4.2.5.0 (8) for an updated sequence. This process was repeated until no new variations were detected. Default parameters were used except where otherwise noted.

We determined the genome sequence of the CarMV variant in each CarMV-positive plant with a 100% coverage breadth for each sample. Compared to the CarMV reference genome (1) (NC_001265.2), mutations were observed, ranging from 108 (2.7 %) to 141

TABLE 1 DNA and protein sequence mismatches of CarMV variants compared to the reference genome NC_001265.2. Abbreviations: SRA, sequence read archive; nt, nucleotides; aDOC, average depth of coverage; aa, amino acids; RdRp, RNA-dependent RNA polymerase; RAP, replicase-associated protein; MP2, movement protein 2; MP1, movement protein 1; CP, capsid protein; ORF, open reading frame.

|		Samples in carnation experiment 1 (CE1)								Samples in CE2			
		S01	S02	S03	S04	S05	S06	S07	S08	S01	S05	S11	S16
Genbank accession		OP506125		OP506126		OP506127	OP506128	OP506129	OP506130	OP506131	OP506132	OP506133	OP506134
SRA accession		SRR2175 1399	SRR2175 1398	SRR2175 1407	SRR2175 1406	SRR2175 1405	SRR2175 1404	SRR2175 1403	SRR217 51402	SRR2175 1401	SRR2175 1400	SRR2175 1397	SRR2175 1396
CarMV ELISA result		+	−	+	−	+	+	+	+	+	+	+	+
Total reads (10^6)		6.9	4.2	5.0	4.7	5.2	5.5	5.6	5.0	27.9	26.9	28.9	40.8
CarMV reads (10^6)		1.5		1.2		1.8	1.5	1.5	1.0	13.0	13.1	14.8	19.8
Genome													
Size (nt)		4,003		4,002		4,003	4,003	4,002	4,003	4,003	4,003	4,003	4,003
GC content (%)		49.1		49.1		49.2	49.1	48.8	48.8	49.0	48.9	49.0	48.8
Differences to NC_001265 (nt)		112		120		116	108	121	119	133	141	130	138
aDOC		7,700		6,100		9,500	7,900	7,900	5,500	68,300	68,900	77,400	104,000
Protein name	**Size (aa)**												
ORF1 RdRp	*245*	3		2		3	3	4	2	6	6	6	6
ORF1 RAP	*430*	11		8		11	10	10	8	5	5	5	5
ORF2 MP2	*84*	0		0		0	0	0	0	1	1	1	1
ORF3 MP1	*61*	1		1		1	1	1	1	2	2	2	2
ORF4 CP	*348*	6		8		8	6	6	7	4	4	4	4
All ORFs	*1,168*	21		19		23	20	21	18	18	18	18	18
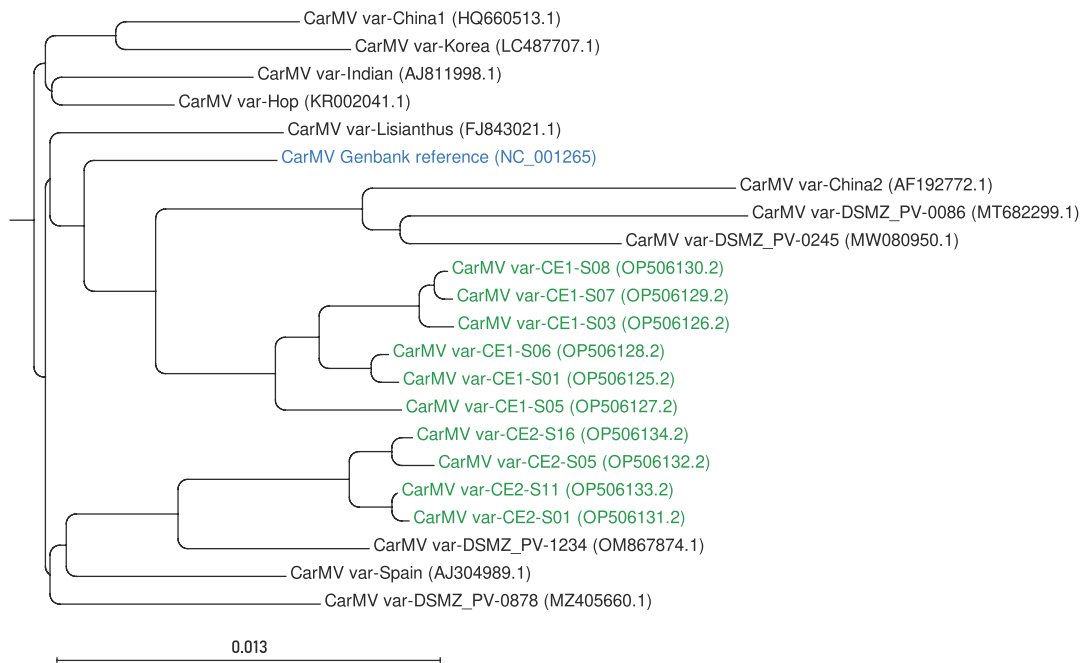
FIG 1 Phylogenetic analysis of CarMV variants. The study included 10 CarMV variants obtained from the current investigation (represented in green), along with 11 CarMV variants sourced from GenBank (represented in black). The CarMV reference genome sequence NC_001265 is shown in blue. The phylogenetic tree was constructed using CLC Genomics Workbench v21.0.5 software. The tree was generated by initially aligning the genome sequences of the CarMV variants, utilizing the following gap cost settings: open 10, extension 2, and end gap "As any other." Subsequently, the neighbor joining method was employed to construct the tree, and Kimura's two-parameter model (K80) was utilized for estimating genetic distances. To assess the robustness of the tree, bootstrapping was performed with 100 replicates.

(3.5%) nucleotides (Table 1). These differences are alike those between known Genbank CarMV variants (78–187 mismatch nucleotides). In a genetic relation analysis, the CarMV variants from our two experiments cluster together (Fig. 1). Additionally, CE1 and CE2 CarMV variants exhibit a closer genetic relationship to Asian and European variants, respectively.

The protein differences among all new CarMV variants ranged from 19 (1.6%) to 23 (2.0%) amino acids. Although there are nucleotide differences between the four CE2 CarMV variants, they are identical at the protein level, whereas the six CE1 CarMV variants are different at the genome and protein level.

ACKNOWLEDGMENTS

AUTHOR AFFILIATION

[1]RNA Biology & Applied Bioinformatics research group, Faculty of Science, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands

AUTHOR ORCIDs

Timo M. Breit http://orcid.org/0000-0002-6801-9969
Rob J. Dekker http://orcid.org/0000-0002-2662-8715

AUTHOR CONTRIBUTIONS

Timo M. Breit, Conceptualization, Supervision, Writing – original draft | Wim de Leeuw, Data curation, Formal analysis | Marina van Olst, Investigation | Wim A. Ensink, Investigation | Selina van Leeuwen, Investigation | Martijs J. Jonker, Conceptualization, Formal analysis, Investigation | Rob J. Dekker, Conceptualization, Supervision, Visualization, Writing – review and editing

DATA AVAILABILITY

The raw sequence reads have been deposited in the NCBI Sequence Read Archive under BioProject accession number PRJNA885393 (Table 1). CarMV variant sequences have been deposited under GenBank accession numbers OP506125, OP506126, OP506127, OP506128, OP506129, OP506130, OP506131, OP506132, OP506133, and OP506134.

REFERENCES

1. Guilley H, Carrington JC, Balàzs E, Jonard G, Richards K, Morris TJ. 1985. Nucleotide sequence and genome organization of carnation mottle virus RNA. Nucleic Acids Res 13:6663–6677. https://doi.org/10.1093/nar/13.18.6663
2. Vilar M, Esteve V, Pallas V, Marcos JF, Perez-Paya E. 2001. Structural properties of carnation mottle virus p7 movement protein and its RNA-binding domain. J Biol Chem 276:18122–18129. https://doi.org/10.1074/jbc.M100706200
3. Jo Y, Choi H, Cho WK. 2015. Complete genome sequence of a carnation mottle virus infecting hop plants. Genome Announc 3:e00416-15. https://doi.org/10.1128/genomeA.00416-15
4. 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
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
6. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923
7. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv preprint. https://doi.org/10.48550/arXiv.1207.3907
8. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. https://doi.org/10.1101/gr.107524.110