Virus Genes (2012) 44:323–328
DOI 10.1007/s11262-011-0683-8

Genetic variability in the coat protein genes of *Cymbidium mosaic virus* isolates from orchids

Ju-Yeon Yoon · Bong-Nam Chung · Gug-Seoun Choi · Seung-Kook Choi

Received: 18 July 2011 / Accepted: 8 October 2011 / Published online: 21 October 2011

Abstract The variability in the nucleotide (nt) and amino acid (aa) sequences of the coat protein (CP) of Cymbidium mosaic virus (CymMV), which naturally infects orchids worldwide, was investigated. The CP genes of 55 CymMV isolates originating from different locations in Korea were amplified using RT-PCR and sequenced. The encoded CP consists of 223 aa. The CP sequences of the Korean isolates were compared with those of previously published CymMV isolates originating from different countries at both nt and aa levels. The Korean isolates shared 74.9–98.3 and 52.7–100% CP homology with CymMV isolates from other countries at the nt and aa levels, respectively. No particular region of variability could be found in either grouping of viruses. In the deduced CymMV CP aa sequence, the C-terminal region was more divergent than the N-terminal. The phylogenetic tree analysis based on nt sequence diversity of CP genes of CymMV isolates supported the hypothesis that CymMV isolates were divided into two subgroups. However, these subgroups were not formed by phylogenetic tree analysis of CP aa sequences. There was no distinct correlation between geographical locations and specific sequence identity, while recombination analysis revealed that there were no intra-specific recombination events among CymMV isolates.

Keywords Coat protein · Cymbidium mosaic virus · Genetic variability · Orchid · RT-PCR · Sequence analysis

Ju-Yeon Yoon and Bong-Nam Chung have contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s11262-011-0683-8) contains supplementary material, which is available to authorized users.

J.-Y. Yoon
College of Natural Science, Seoul Women's University, Seoul 139-774, Republic of Korea

B.-N. Chung · G.-S. Choi
Virology Unit, Department of Horticultural Environment, National Institute of Horticultural and Herbal Science, RDA, Suwon 440-441, Republic of Korea

S.-K. Choi (✉)
Virology Unit, Department of Horticultural Environment, National Institute of Horticultural and Herbal Science, RDA, Suwon 440-310, Republic of Korea
e-mail: viroid73@yahoo.com

Introduction

Orchid is one of the most important commercial crops in the world's flower industry, because of their attractive flowers with diverse shapes and long vase life. In Korea, in 2007, the annual production of orchids was 9.7 million pots and the wholesale value was 1 billion US dollars (Ministry for Food, Agriculture, Forestry and Fisheries, Korea) [1]. Orchids have been reported to be infected with more than 50 viruses [2]. Among orchid viruses, *Cymbidium mosaic virus* (CymMV) and *Odontoglossum ring spot virus* (ORSV) are the most prevalent, economically important in orchid production [2, 3]. These viruses cause reductions in yield and quality of orchid flowers. Orchid diseases caused by these two viruses occur frequently in Korea [4] as well as in several countries [5–8]. CymMV and ORSV have attained worldwide distribution. CymMV is mechanically transmitted through infected sap on cutting tools and potting media. CymMV-infected orchids often show a variety of symptoms such as color-breaking in flowers, streak or stripe mosaic or necrotic spots, yellowing leaves, as well as symptomless conditions [3, 6–9]. CymMV as a member of

the genus *Potexvirus*, has a positive-sense single-stranded RNA genome of about 6.3 kb containing 5 ORFs flanked by 5′ and 3′ non-coding regions plus 3′ poly(A) tail [9, 10]. ORF 1 encodes a 160-kDa replicase containing a methyltransferase domain in the N-terminal region, a RNA helicase domain and the RNA-dependent RNA polymerase domain in the C-terminal region. ORFs 2–4, referred to as the triple gene block, are involved in cell-to-cell movement. ORF 5 encodes the 24-kDa viral coat protein (CP), which is indispensible for virus assembly and long-distance movement [11].

To date, the complete nucleotide (nt) sequences of 11 CymMV isolates have been reported [9, 12–16]. There have been several reports about the molecular variability of CymMV isolates. In 2002, Ajjikuttira et al. [5] investigated the CP genes of 26 Asian isolates and found 89.1–99.7 and 93.2–100% identity at the nt and amino acid (aa) levels, respectively. They did not find any distinct regions of variability in any of the sequences. Moles et al. [17] examined the molecular variability of 30 CP genes from vanilla isolates with 55 CP genes of CymMV isolates from other plant sources and found that CymMV isolates could be divided into two subgroups. Interestingly, the two congruent monophyletic clusters were not observed by aa sequence analysis.

To assess the possible molecular diversity of CymMV isolates from a larger population of samples, we identified and determined the CP gene sequences from 55 CymMV isolates infecting different genera of orchids from geographically distinct areas in Korea. We analyzed the CP gene sequence of the CymMV isolates and compared these sequences with those of other isolates from geographically distinct locations, such as Korea, China, Singapore, Taiwan, India, French Polynesia, France, and USA.

Materials and methods

Virus isolates

Fifty-five virus isolates from different genera and species of orchids plants (*Cymbidium virescens*, *Cymbidium* sp., *Dendrobium* sp., and *Oncidium* sp.) were obtained from various locations in Korea (Supplementary Table S1). The presence of CymMV was confirmed by double-sandwich-enzyme-linked immunosorbent assay (DAS-ELISA) as described previously [18]. In brief, leaf tissues were homogenized in 100 mM phosphate buffer containing 0.02% NaN_3, 0.1% Tween 20, and 0.1% skim milk powder (pH 7.4) at a sample-to-buffer ratio of 1:3 (w:v), and 100 µl of extracted sap was loaded in duplicate onto microtiter plates. The primary antibody specific to CymMV purchased from Agdia (USA) was diluted to 1:100 in carbonate buffer (0.05 M sodium carbonate, pH 9.6) and the diluted antibody solution (1 µg/ml) was used for CymMV detection in microtiter plates. Subsequently, goat anti-rabbit IgG-conjugated alkaline phosphatase was used as a secondary antibody, according to the manufacturer's instructions (Promega, USA). Substrate, 4-nitrophenyl phosphate (0.6 mg/ml), was allowed to react at room temperature for 1 h (Sigma, USA). Absorbance values ($A_{405\ nm}$) of four times the healthy control reading were used as the positive threshold.

RT-PCR

For each sample, 1 mg of leaf tissue of infected orchid was frozen in liquid N_2 and ground to a fine powder. Total RNA was extracted using an RNeasy plant mini-kit, according to the manufacturer's instructions (Qiagen, USA). One microgram of each RNA preparation was reverse-transcribed and amplified using a primer set specific to CymMV. The AccuPower RT-PCR kit (Bioneer, Korea) was used to amplify the CymMV CP gene using a forward primer (5′-ACAATAATTTGAAATAATCATGGGA-3′, corresponding to nts 5,463–5,487, based on the sequence of GenBank accession number NC_001812 and a reverse primer (5′-AAAACCACACGCCTTATTAAGTTTG-3′, complementary to nts 6,156–6,180, based on NC_001812). The thermo-cycling conditions were as follows: 60 min at 42°C for RT, 2 min at 95°C (1 cycle), 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min (35 cycles), and a final extension at 72°C for 7 min.

Cloning and sequence analysis

The synthesized RT-PCR products were purified using Qiaquick PCR purification (Qiagen, USA) and cloned into pGEM-T easy vector (Promega, USA), according to the manufacturer's instructions. All PCR products were sequenced using the Big-Dye Terminator Sequencing Kit (Applied Biosystems, USA), according to the manufacturer's instructions. The complete CP sequences of Korean isolates of the CymMV determined have been deposited in the NCBI GenBank under accession numbers AB541522–AB541573. Analysis of the nt and deduced aa sequences were done using BLAST search, the DNAMAN sequence analysis software (version 5.1, Lynnon Biosoft Co. Canada) and DNASTAR software (USA). For comparison, 108 CymMV CP gene sequences available from the GenBank database were added to our data set (Supplementary Table S1). The sequences from the database that were redundant, or smaller than the size of full-length CymMV CP gene, were omitted. Potato acuba mosaic virus (PAMV) CP gene (accession number S73580) was used as an out-group for sequence alignment and phylogenetic tree analysis.

Nt sequence identities were calculated using the Jukes and Cantor index and genetic diversity at synonymous positions was estimated using the Hasegawa–Kishino–Yano model in the MEGA 4.0 Software [19]. Model tests for determination of the evolutionary positions of CymMV CP genes were performed using in the MEGA 4.0 software. Basically, phylogenetic tree analysis with CymMV CP nt sequences was constructed using neighbor-joining (NJ) with Jukes and Cantor index or Hasegawa–Kishino–Yano model in the MEGA 4.0 software, DNAMAN, and DNASTAR software. To validate phylogenetic tree analysis from the NJ method, phylogenetic trees were produced using Minimum Evolution, Maximum Parsimony, and Maximum likelihood methods in the MEGA 4.0 software. No informative characters were deactivated to optimize tree statistics. Bootstrap analyses with 1,000 replicates were performed to evaluate the significance of the interior branches. The consensus tree built for the CymMV nt sequences was rooted with the PAMV CP sequence. Recombination sites in the aligned sequences were first detected by programs CHIMEARA, GENECONV, BootScan, MaxiChi, Chimaera, 3Seq, and SISCAN methods implemented in the RDP3.42 software, using default settings and a Bonferroni-corrected *P* value cut off of 0.05 (http://darwin.uvigo.es/rdp/rdp.html) [20]. The isolates identified as possible recombinants by the programs in RDP were rechecked using the program PHYLPRO (http://www.rsbs.anu.edu.au/ResearchGroups/GIG/Products/phylpro/index.html).

Results

ELISA showed that most of the orchids we collected from various locations were infected with CymMV, suggesting that this virus is one of the most widespread orchid viruses in Korea. Our previous results also showed that a number of ORSV isolates caused severe damage in orchid production [21]. Among the larger population of samples showing positive reactions, the nt sequences of the CP genes of 55 Korean isolates were further determined in this study.

We newly identified and determined CP sequences of 27 CymMV isolates from *Cymbidium* sp. (isolates CK13, NHRIK1, and CK15-CYK15), CP sequences of 18 CymMV isolates from *Dendrobium* sp. (isolates CK14 and DEK16–DEK32), CP sequence of one CymMV isolate from *Phalaenopsis* sp. (isolate NIHHSK2), and CP sequences of nine CymMV isolates from *Oncidium* sp. (isolates ONK18–ONK26) in this study (Supplementary Table S1). The amplified RT-PCR products were always of the expected size of 672 bp. All the encoded sequences of the CP of CymMV possessed 223 aa. This CP size was the same as that of previously published sequences. The CP gene nt sequences of the 55 Korean isolates analyzed shared 89.4–100% identity among them and 63.9–100% identity with those of other Korean CymMV isolates reported previously (data not shown). The nt sequences of the CP gene of the Korean isolates shared 74.9–98.3% identity with those of previously published CymMV isolates originating from different countries. The percentage of divergence with PAMV, representing the closest related virus species, ranged from 54.8 to 57.2% at the nt level.

The phylogenetic tree deduced from the CP alignment is presented in Fig. 1. The NJ tree constructed from 163 CP sequences revealed a distinct segregation of the isolates into two monophyletic clusters. However, all of the isolates did not show any relationship according to geographical origins and isolation hosts, as expected by matrix of sequence homology. The NJ tree analysis using the nt sequences, regardless of the origin of the isolates being analyzed, produced nearly congruent result with the phylogenetic trees using minimum evolution, maximum parsimony, and maximum likelihood methods (Supplementary Fig. S1). Pairwise identities of CP genes between isolates from subgroups A and B ranged from 74.9 to 98.3%. The CP nt identities within clusters ranged from 74.2 to 100% for subgroup B and 72.6–100% for subgroup A.

The putative occurrence of recombination events in the CymMV population was examined using the RDP software. Although potential recombination was identified by MaxChi or Chimaera algorithms, none could be validated using other methods. The CP genes of the Korean isolates, as well as those of other isolates, did not have distinct recombination sites (data not shown). This result suggested that recombination has not occurred noticeably between CymMV isolates in the CP genes.

In the deduced aa sequences, the CP aa of the 55 Korean isolates analyzed shared at least 75.4–100% identity among them. Interestingly, while the nt sequence of the CP genes of the isolates was not completely identical, the translated CP sequences of 19 Korean isolates among the 55 isolates were identical at the aa level. Among the Korean isolates, including nine Korean isolates reported previously (Supplementary Table S1), the CP aa sequence showed 55.7–100% identity. NIHHSK2 shared the lowest sequence identity (55.7%) with KI. Including the 108 sequences downloaded from the aa databases, the level of aa identity between isolates ranged from 52.7 to 100%. Isolate KI shared the lowest sequence identity (52.7%) with PLM and Orissa. In addition, KI showed 55.4% identity with ZJXTHC1 and SCJJC7. The aa sequence matrix of all sequences suggested that the genetic variability of the deduced CP sequences did not have a relationship to the isolation host species (data not shown), as also observed from nt sequences. By pairwise alignment analysis, there

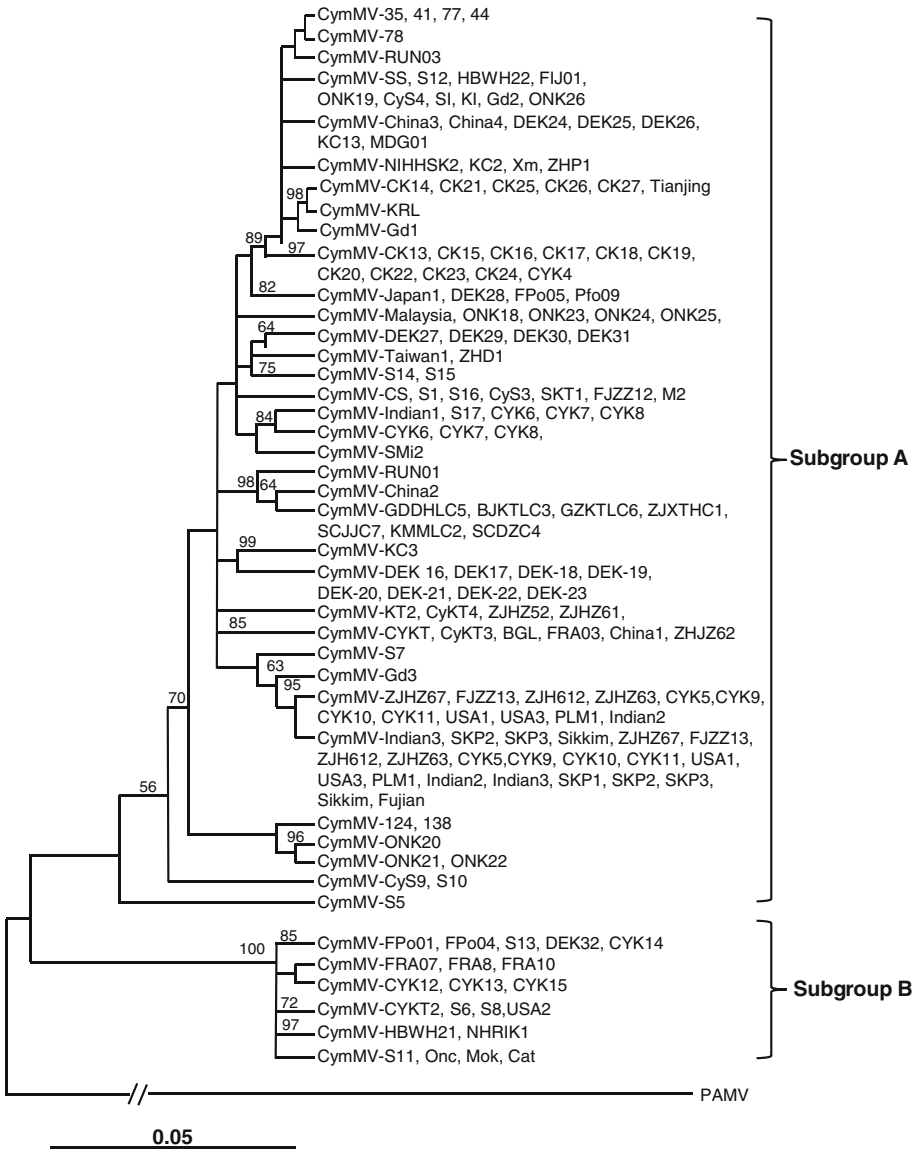

Fig. 1 Phylogenetic tree analysis of Korean CymMV isolates and other isolates based on deduced CP nt sequences. Unweighted NJ tree constructed from the sequence alignment of 647-nts of CymMV CP coding region for 163 isolates; 55 isolates from this study and 108 isolates from the Genbank database. Multiple sequence alignments were generated by using the MEGA4.0, and phylogenetic tree was constructed by the NJ algorithm, based on calculations from pairwise nt sequence distances for gene nt analysis derived from the multiple-alignment format. *Numbers above the lines* indicate bootstrap scores out of 1,000 replicates. The *scale bar* represents 0.05 divergence of the Jukes and Cantor dissimilarity index. Bootstrap analysis was done with 1,000 replicates of the starting tree. Bootstrap values are shown in each branch and only values greater than 50% are presented. The abbreviated names of the virus isolates are described in Supplementary Table S1. In addition, the CP sequence of PAMV (accession no. S73580), which represents the species whose members are most closely related to CymMV was used as the out-group, but this scale is not indicated

was no distinct variable region among the encoded CP ORFs of the 55 Korean isolates determined in this study or between the Korean isolates and other previously reported isolates. However, it was observed that the N-terminal sequence of CymMV CP was more conserved than the C-terminal sequence between the Korean isolates and other isolates reported previously. It is interesting that a distinct CymMV dichotomy was not observed when an aa phylogenetic tree of the Korean isolates was analyzed with the previously reported isolates (data not shown). The nt variations observed for the CP gene resulted in low aa divergence due to synonymous mutations (data not shown) and the isolates did not segregate inevitably in the aa phylogenetic analysis [5, 17]. The limits of intra-specific diversity of CymMV from different geographical origins were also supported by the conservation of the CP ORFs, as shown the result from RNA recombination analysis with the determined nt sequences of the Korean isolates and the nt sequences of other published isolates.

Discussion

We compared CP sequence of 55 Korean CymMV isolates and 108 previously reported isolates originated from different host ranges and locations. The CP sequences of CymMV isolates from *Dendrobium* sp. *Cymbidium* sp., *Phalaenopsis* sp., and *Oncidium* sp. were determined. To our knowledge, our study is the first report of CymMV isolates from *C. virescens*. The analysis of CP aa sequences revealed that the overall diversity of CymMV isolates is low. In addition, our CP analysis with 163 CymMV isolates confirmed the highly divergent C-terminal CPs of CymMV (data not shown), as previously suggested [17]. The CP nt

analysis of Korean CymMV isolates clearly showed CymMV isolates could be segregated to two subgroups, as shown by the congruent phylogenetic tree analyses from CP and partial RNA-dependent RNA polymerase sequences [17]. Approximately 13% of CymMV isolates (21 of 163 isolates) belonged to subgroup B, sufficient to be distinguished from the remaining 142 CymMV isolates forming subgroup A. Population subgroups in many plant virus species based on the observation of molecular genetic diversity have shown that virus subgroups from nt sequence analysis was similar to aa sequence analysis and correlated with biological or geographical features of the isolates. Since the phylogenetic tree analysis from the deduced CP sequence formed the subgrouping of CymMV isolates based on nt sequences, it provokes raising questions about the nature of CymMV subgroups. It is not likely that the CymMV subgrouping based on nt sequences reflects specific host-plant features, since phylogenetic tree analysis showed no correlation between isolates of each subgroup and plant species that they were isolated from or between isolates of each subgroup and the geographical origins. In accordance with these observations, the clustering of CymMV isolates confirmed previous CP analysis data for the trees obtained from the CP alignment of CymMV or other potexviruses [22–24]. Some isolates of CymMV showed significantly variability at the C-termini of the CP. A similar observation was found in the CP analysis of Hosta virus X (HVX), which showed that the C-terminal region of the CP in HVX isolates is the most variable region of this protein [24].

A possible criticism of our study is that this analysis considers only a limited part of the entire genome-coding regions. Nevertheless, by comparing the complete CP nt sequences of available strains from other CymMV isolates, including Asian isolates and Western strains, we observed that CymMV isolates could be divided clearly into two subgroups (Fig. 1). This suggests that CymMV isolates originated from two moderate distant ancestors that can genetically recombine. In addition, this might be in part explained by the fact that the tubers from the same origin are disseminated for multiplication during orchid production, resulting in the high degree of CP similarities. Based on the determination of only CP sequences of virus isolates and various analyses of the determined sequences observed from a few plant viruses, some studies showed no geographical variation of the analyzed viruses [25–29]. These studies support the validity of our data that the analysis of CP gene sequences is sufficient for investigating the geographical variation of CymMV isolates.

Interestingly, the CymMV isolates analyzed in this study indicated that that no significant recombination has occurred among the analyzed isolates, in particular, based on nt sequences of the CP genes between the Korean isolates and other previously reported isolates. This result was supported by two independent studies reported by Ajjikuttira et al. [5] and Moles et al. [17]. In contrast, recombination events were found at nt position 5,617 in the CP gene of Hawaiian isolate 1, which represented a recombinant between isolates 2 and 3 [13] and at positions 4,972–6,226 of one Indian CymMV isolate between an Indian isolate and a Korean isolate [14]. Therefore, it remains to be determined whether CymMV populations contain recombinant events using complete genome sequencing with a great numbers of worldwide CymMV isolates.

Acknowledgments This study was supported by funded by the Next Generation BioGreen21 program (Project no. PJ007984032011) and by 2011 Postdoctoral Fellowship Program of NIHHS, Rural Development Administration, Republic of Korea.

References

1. H.H. Kim, Y.J. Kyung, K. Ohkawa, C.H. Pak, B.H. Kwack, Acta Hortic. **482**, 407–414 (1999)
2. F.W. Zettler, N.J. Ko, G.C. Wisler, M.S. Elliot, S.M. Wong, Plant Dis. **74**, 621–626 (1990)
3. D.D. Jensen, A.H. Gold, Phytopathology **41**, 648–653 (1951)
4. K.H. Ryu, W.M. Park, Plant Dis. **79**, 321 (1995)
5. P.A. Ajjikuttira, C.L. Lim-Ho, M.H. Woon, K.H. Ryu, C.A. Chang, C.S. Loh, S.M. Wong, Arch. Virol. **147**, 1943–1954 (2002)
6. S.M. Wong, C.G. Chng, Y.H. Lee, Y. Tan, F.W. Zettler, Crop Prot. **13**, 235–239 (1994)
7. A.R. Sherpa, V. Hallan, P. Pathak, A. Zaidi, J. Phytopathol. **154**, 275–280 (2006)
8. G.C. Wisler, F.W. Zettler, L. Mu, Plant Dis. **71**, 1125–1129 (1987)
9. S.M. Wong, P.H. Mahtani, K.C. Lee, H.H. Yu, Y. Tan, K.K. Neo, Y. Chan, M. Wu, C.G. Chng, Arch. Virol. **142**, 383–391 (1997)
10. M.J. Adams, J.F. Antoniw, M. Bar-Joseph, A.A. Brunt, T. Candresse, G.D. Foster, G.P. Martelli, R.G. Milne, S.K. Zavriev, C.M. Fauquet, Arch. Virol. **149**, 1045–1460 (2004)
11. J. Verchot-Lubicz, C. Ye, D. Bamunusinghe, J. Gen. Virol. **88**, 1643–1655 (2007)
12. J.D. Kim, Y.B. Koo, M.U. Chang, Mol. Cells **8**, 181–188 (1998)
13. S.P. Vaughan, M. Grisoni, M.H. Kumagai, A.R. Kuehnle, Arch. Virol. **153**, 1185–1189 (2008)
14. A.R. Serpa, V. Hallan, P. Pathak, A.A. Zaidi, J. Biosci. **32**, 663–669 (2007)
15. Z. He, D. Jiang, A. Liu, L. Sang, W. Li, S. Li, Virus Genes **42**, 440–443 (2011)
16. H.L. Wang, W.C. Liu, Plant Pathol. Bull. **13**, 61–68 (2004)
17. M. Moles, H. Delatte, K. Farreyrol, M. Grisoni, Arch. Virol. **152**, 705–715 (2007)
18. W.M. Park, K.B. Shim, S.J. Kim, K.H. Ryu, Plant Pathol. J. **14**, 130–135 (1998)
19. K. Tamura, J. Dudley, M. Nei, S. Kumar, Mol. Biol. Evol. **24**, 1596–1599 (2007)
20. L. Heath, E. van der Walt, A. Varsani, D.P. Martin, J. Virol. **80**, 11827–11832 (2006)
21. J.Y. Yoon, B.N. Chung, S.K. Choi, Virus Genes **42**, 261–267 (2011)

22. B.A. Cox, R.A.C. Jones, Arch. Virol. **155**, 1349–1359 (2010)
23. E.R. Chare, E.C. Holmes, J. Gen. Virol. **85**, 3149–3157 (2004)
24. O.L. Fajolu, R.H. Wen, A.S. Windham, M.T. Windham, J.K. Moulton, M.R. Hajimorad, Arch. Virol. **154**, 1909–1916 (2009)
25. A. Fraile, F. Escriu, M.A. Aranda, J.M. Malpica, J.M. Gibbs, F. Garcia-Arenal, J. Virol. **71**, 8316–8320 (1997)
26. A. Fraile, J.L. Alonso-Prados, M.A. Aranda, J.J. Bernal, J.M. Malpica, F. Garcia-Arenal, J. Virol. **71**, 934–940 (1997)
27. D. Fargette, G. Konate, C. Fauquet, E. Miller, M. Peterschmitt, J.M. Thresh, Annu. Rev. Phytopathol. **44**, 235–260 (2006)
28. K. Komatsu, S. Kagiwada, S. Takahashi, T. Mori, Y. Yamaji, Virus Genes **31**, 99–105 (2005)
29. V. Aritua, E. Barg, E. Adipala, R.W. Gibson, D.E. Lesemann, H.J. Vetten, Plant Dis. **93**, 87–93 (2009)