Arch Virol (2007) 152: 2073–2078
DOI 10.1007/s00705-007-1042-x
Printed in The Netherlands

Archives of Virology

Brief Report

Complete genome sequences of Maize dwarf mosaic and Sugarcane mosaic virus isolates coinfecting maize in Spain

M. A. Achon[1], L. Serrano[1], N. Alonso-Dueñas[1], C. Porta[2]

[1] Area de Proteccio de Conreus, Centre UdL-IRTA, Lleida, Spain
[2] Warwick HRI, University of Warwick, Wellesbourne, Warwick, UK

Received 16 May 2007; Accepted 2 July 2007; Published online 6 August 2007

Summary

The genomes of Spanish isolates of *Maize dwarf mosaic virus* (MDMV-Sp) and *Sugarcane mosaic virus* (SCMV-Sp) were completely sequenced. Nucleotide sequence identities of SCMV-Sp to those of other SCMV isolates ranged from 79 to 90%. MDMV-Sp shared 85% nucleotide identity with the only other fully sequenced isolate of MDMV. MDMV-Sp and SCMV-Sp differed from each other by 31% in their nucleotide sequences. Phylogenetic analyses showed that SCMV isolates group by host rather than by geographical location. Two significant recombination signals were identified in the NIa and NIb regions of the SCMV-Sp genome.

*

Maize mosaic is the most widespread viral disease of maize in Europe and is caused by isolates of *Maize dwarf mosaic virus* (MDMV) and/or *Sugarcane mosaic virus* (SCMV) [1, 16]. In Spain MDMV is the prevalent virus [2]. SCMV has been detected either in single or mixed infections with MDMV and in the last few maize growing seasons a 10% increase in the occurrence of SCMV has been observed [2; Achon, unpublished]. Both viruses belong to the genus *Potyvirus* (family *Potyviridae*) and together with *Johnsongrass mosaic virus* (JGMV), *Sorghum mosaic virus* (SrMV), SCMV-MDB, and *Zea mosaic virus* (ZMV) [20, 21, 25] form the SCMV subgroup. All of these potyviruses infect maize and sorghum, with SrMV and some SCMV strains also infecting sugarcane. Members of the SCMV subgroup, like other potyviruses, have a monopartite, single-stranded messenger RNA genome encoding a single polyprotein and are non-persistently transmitted by aphids [22]. Complete sequence data are essential for understanding the biology, diversity and evolutionary history of a virus, for devising new control strategies and evaluating its risk. To date, the complete sequence of only one MDMV isolate from Bulgaria (MDMV-BG) is available [11], whilst ten SCMV isolates have been entirely sequenced, with nine of them originating from China [7, 8, 26]. In this study, we describe the complete genome sequence of the maize isolates MDMV-Sp and SCMV-Sp collected in the Northeast of Spain in 1992 and 1998, respectively.

Correspondence: Dr. Maria Angeles Achon, Area de Proteccio de Conreus, Centre UdL-IRTA, Rovira Roure 191, 25198 Lleida, Spain.
e-mail: achon@pvcf.udl.es

Viral RNAs were extracted from MDMV-Sp and SCMV-Sp particles as described earlier [1, 3]. Viral RNAs were reverse transcribed using Super-Script™ II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions and then treated with RNase H. Platinum Pfx DNA polymerase (Invitrogen) was used for PCR reactions. Three overlapping cDNA fragments were synthesized by RT-PCR of genomic RNA of MDMV and SCMV. The cDNA fragments corresponding to the 3′ termini of the viral genomes were amplified as described previously [1, 3] and were 2140 and 1700 bp in length for MDMV-Sp and SCMV-Sp, respectively. The sequences of these cDNAs were exploited to design specific antisense primers: TGG GAT CCC TAT CGA ATT GTG AAC $C^{7591-7573}$ and TTT AGC AAA CTG TTT AAA GTA TTT $TT^{8340-8325}$. These were respectively used in conjunction with a degenerate sense primer JF15: AAW TRW MAG AAT TCA AYA CAA CAT^{2-24} [15] and a specific sense primer TCA AAA GAG TGA TGG AGA TGT $TT^{213-236}$ to amplify cDNA fragments of 7589 and 8200 bp from the central genomes of MDMV-Sp and SCMV-Sp. The cDNA fragments corresponding to the 5′ termini of these viral genomes were amplified using the 5′RACE System (Invitrogen) in conjunction with the following primers: GTA GTG TTC TTG CGT TTC TTT $CC^{376-354}$ and ACG TCT GCA GAC ACT TCT TTG GTT GAA CCG $TGAT^{320-342}$ for MDMV, and TGT GGG GTT GTG GTC GCA CTT $TTCA^{368-392}$ and ACG TAT GCA TTG TGC TGG TTY AAY AAA CTC $AGC^{318-340}$ for SCMV. The 5′ terminal cDNA products were 390 and 370 bp long, respectively, for MDMV and SCMV. The PCR-amplified cDNA products were purified using the Geneclean II kit (BIO 101 Inc.). For cloning, adenosine residues were attached to the cDNAs by incubation with the Tfl DNA polymerase (Promega) in the presence of ATP. The resulting products were ligated into the pGEMT-Easy vector (Promega) and subsequently introduced by transformation into *E. coli* strain DH5α. Recombinant plasmids were purified using the Quantum Prep Plasmid Miniprep kit (Bio-Rad). Nucleotide sequence determination was performed twice on both strands, and the consensus sequence was obtained from at least three independent overlapping cDNA clones. DNA sequencing was done with an ABI Prism™ 3100 DNA Sequencer, using the BigDye™ Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer).

The complete sequences of Spanish isolates of MDMV and SCMV were deposited into the

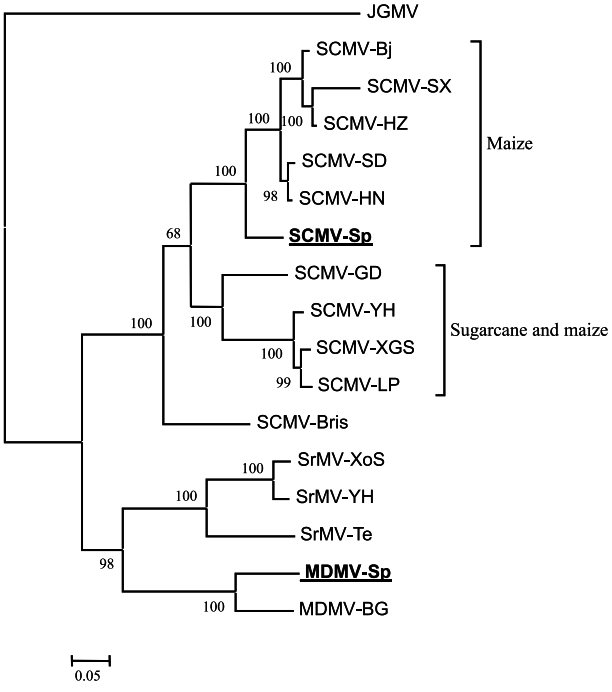

Fig. 1. Phylogenetic tree of MDMV-Sp, SCMV-Sp and the other viral species of the SCMV subgroup based on their complete nucleotide sequences. Trees were inferred using the maximum likelihood method (model HKY, with 1000 non-parametric bootstrap replicates) using PHYML and edited using MEGA3. Bootstrap values are shown at each node. The viruses used in this study, their sequence accession number and geographical origin were as follows: *Sorghum mosaic virus* (SrMV) isolates, SrMV-Te (U57358) from the USA, SrMV-XoS (AJ310197) and SrMV-YH (AJ310198) from China, *Johnsongrass mosaic virus*, (JGMV: Z26920) isolate from Australia; *Sugarcane mosaic virus* (SCMV) isolates, SCMV-HN (AF494510), SCMV-LP (AJ310102), SCMV-XGS (AJ310103), SCMV-YH (AJ310104), SCMV-GD (AJ310105), SCMV-HZ (AJ297628), SCMV-Bj (AY042184), SCMV-SD (AY149118), and SCMV-SX (AY569692), all from China, SCMV-Bris (AJ278405) from Australia, SCMV-Sp (AM110759) from Spain; and *Maize dwarf mosaic virus* (MDMV) isolates, MDMV-BG (AJ001691) from Bulgaria and MDMV-Sp (AM110758) from Spain

GenBank/EMBL/DDBJ database under accession numbers AM110758 (MDMV-Sp) and AM110759 (SCMV-Sp). The genome of MDMV-Sp is made up of 9414 nucleotides (nts) and that of SCMV-Sp has 9596 nts excluding the poly(A) tails. The genome of MDMV-Sp contains one open reading frame (ORF) encoding 3042 amino acids, flanked by 5′ and 3′ untranslated regions (UTRs) of 138 and 234 nts, respectively. The ORF of the SCMV-Sp genome codes for 3064 amino acids, and the 5′ and 3′ UTRs are 149 and 255 nts long, respectively. The sequences of the isolates of the sugarcane subgroup listed in the legend to Fig. 1 were phylogenetically compared using the neighbour-joining method implemented in MEGA3 [12]. SCMV-Sp was most closely related to other SCMV isolates from maize [7, 8, 26], and in particular to SCMV-SD from China with which it shares nucleotide and amino acid sequence identities of 90 and 96%, respectively. Nucleotide and amino acid sequence identities between MDMV-Sp and MDMV-BG [11] were 85 and 93%, respectively.

MDMV-Sp and SCMV-Sp differed in nucleotide and amino acid sequences by 31 and 26%, respectively. When individual gene products were aligned, highest amino acid sequence identities between either the eleven SCMV isolates or the two MDMV isolates were found in CI (94–99 %) and HC-Pro (92–99%), whereas they were lowest (63–99%) in P1. With the exception of SCMV-GD (from maize) and SCMV-Bris (from sugarcane), the levels of similarity among SCMV isolates in P1, P3, 6K2, and the 5′ UTR are related to the host from which the viruses were obtained; the nucleotide sequence identities in these genomic regions are greater than 86% between isolates from the same host. The CP, followed by the 3′ UTR, are the regions that shared the highest identity between SCMV isolates from maize (>90%). An interesting feature observed in the 3% UTR of SCMV isolates is that maize isolates have 20 additional nts (ATAGT ACGTT CGTGA GGCCT) compared to sugarcane isolates.

Alignments of the complete nucleotide sequences were used as input data to construct phylogenetic trees by maximum likelihood (ML) with 1000 non-parametric bootstrap replicates in the HKY model using PHYML [10] and the neighbor-joining (NJ) distance method in the K2P model with 1000 bootstrap replicates using MEGA3 [12]. The topologies of both the ML and NJ trees were essentially identical. SCMV-Sp clustered with all the other SCMV isolates from maize, with the exception of isolate GD, which clustered with the three sugarcane SCMV isolates from China (Fig. 1). As expected, MDMV-Sp clustered with MDMV-BG. Phylogenetic relationships (ML and NJ) among SCMV isolates were also analysed using the nucleotide sequences of each of the twelve genome regions. This resulted in differential clustering properties for SCMV-Sp in the VPg and NIa-Pro regions; grouping in these regions was with sugarcane, and in all other regions, with maize isolates (trees not shown).

As a consequence of the above-mentioned observations, the nucleotide sequence alignment of SCMV isolates was examined for recombination using PhylPro v. beta 1.0 [24]. For SCMV-Sp recombination, signals were observed at positions 5604 (NIa-VPg), and 7866 (NIb) with low phylogenetic correlation coefficients (r) of 0.32 and 0.46, respectively, and at positions 6432 (NIa-Pro) and 8560 (CP) with $r > 0.6$. In addition, we detected the two recombination events reported by Zhong et al. [26] for isolate SX ($r < 0.26$), and other minor recombination signals for other SCMV isolates, the most remarkable of which was in the CP (at position 8826) for isolate GD ($r = 0.6$). To confirm recombination events, further analyses focusing exclusively on SCMV-Sp were conducted using recombination detection methods implemented in the RDP2 program [14] (RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA and SISTERSCAN). A multiple-comparison-corrected *P*-value cutoff of 0.05, 1000 bootscan replicates, and settings recommended in the RDP2 manual for each method were used throughout. Only two of the events described above were identified by RDP2 methods. One was in NIa-VPg (5151–5865) at position 5604, with a high degree of confidence (*P*-values: RDP $= 8.9518 \times 10^{-9}$, GENECONV $= 2.755 \times 10^{-10}$, BOOTSCAN $= 1.9905 \times 10^{-46}$ (Bootscan support = 98%), MAXIMUM CHI SQUARE $= 5.089 \times 10^{-24}$, CHIMAERA $= 1.675 \times 10^{-21}$ and SISTERSCAN =

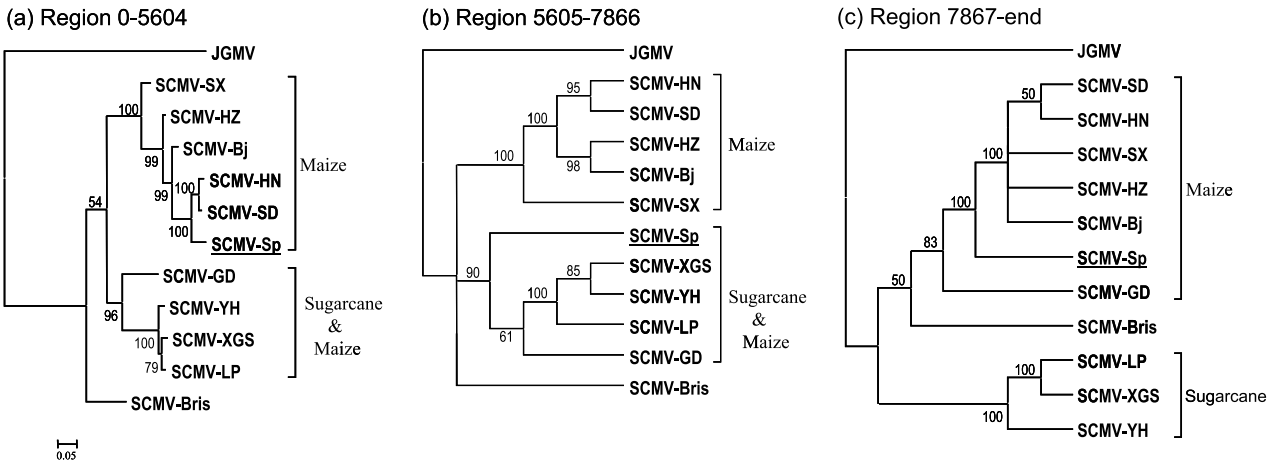

Fig. 2. Maximum-likelihood trees inferred from sequence alignments from either side of the two recombination break-points identified in the NIa and NIb genomic regions of SCMV-Sp by PhylPro and at least four of the six recombination detection methods implemented by the RDP2 software. 5′ terminal nucleotides before the first recombination signal at position 5604 (**a**), nucleotides between recombination signals at positions 5605 and 7866 (**b**), and 3′ terminal nucleotides following the second recombination signal at position 7866 (**c**). Trees were constructed using PHYML (model HKY, with 1000 non-parametric bootstrap replicates) and edited using MEGA3. JGMV was used as an out-group for rooting. Bootstrap values are shown at each node. Branches with <50% bootstrap support were collapsed. Abbreviations of virus isolates are those given in Fig. 1

2.2×10^{-33}). The other recombination event in NIb (7107–7953) at position 7866 was identified by four methods (*P*-values: RDP $= 2.279 \times 10^{-3}$, BOOTSCAN $= 1.79 \times 10^{-3}$ (Bootscan support = 88.7%), MAXIMUM CHI SQUARE $= 3.413 \times 10^{-21}$, SISTERSCAN $= 4.11 \times 10^{-2}$). In order to correlate recombination events with changes in tree topologies, we conducted phylogenetic analyses (ML) for sequence alignments of SCMV isolates on either side of the two recombination break-points identified in NIa and NIb (Fig. 2). The 5604 5′ terminal nucleotides of the SCMV-Sp genome appeared to have originated from maize isolates closely resembling the sequences of isolates HN and SD; the central 2280 nucleotides were similar to those of sugarcane isolates (LP, YH, and XGS) and maize isolate GD, and the 1730 3′ terminal nucleotides again resembled those of maize isolates including GD. The close relationship of the GD isolate in the 3′ terminal region with the other maize isolates was further analysed by RPD2 methods and by ML phylogenetic analysis of sequence alignments on either side of the CP recombination signal (8827). Nucleotides 7867–8826 closely resembled those of isolates from sugarcane (LP, YH and XGS) as did the first 7867 nucleotides of this genome, whilst its sequence from position 8826 onwards was more related to that of maize isolates. The recombination event for GD isolate was only supported by phylogenetic analysis (Fig. 3).

These results provide further evidence for the occurrence of natural recombinants for SCMV [26]. Although in this study we can not detect any recombination signal for MDMV-Sp and MDMV-BG due to the availability of two MDMV sequences, the fact that Chare and Holmes [6] identified a recombination signal in the CP of one MDMV isolate corroborates the significance of recombination for the evolution of viruses of the SCMV subgroup, other potyviruses [18], and RNA viruses in general [10].

Our analysis shows that the relationships between SCMV isolates are more correlated to the host than to the geographical origin. With the exception of the Bris isolate, the 3′ terminal region (CP and 3′ UTR) is the most host-related sequence, followed by P1, P3, 6K2 and 5′ UTR. It is also remarkable that the 3′ UTRs of maize isolates of SCMV contain an additional stretch of 20 nts. In a previous analysis of the 3′ terminal region of SCMV-Sp, which involved the alignment of a large number of CP sequences available from the Gen-

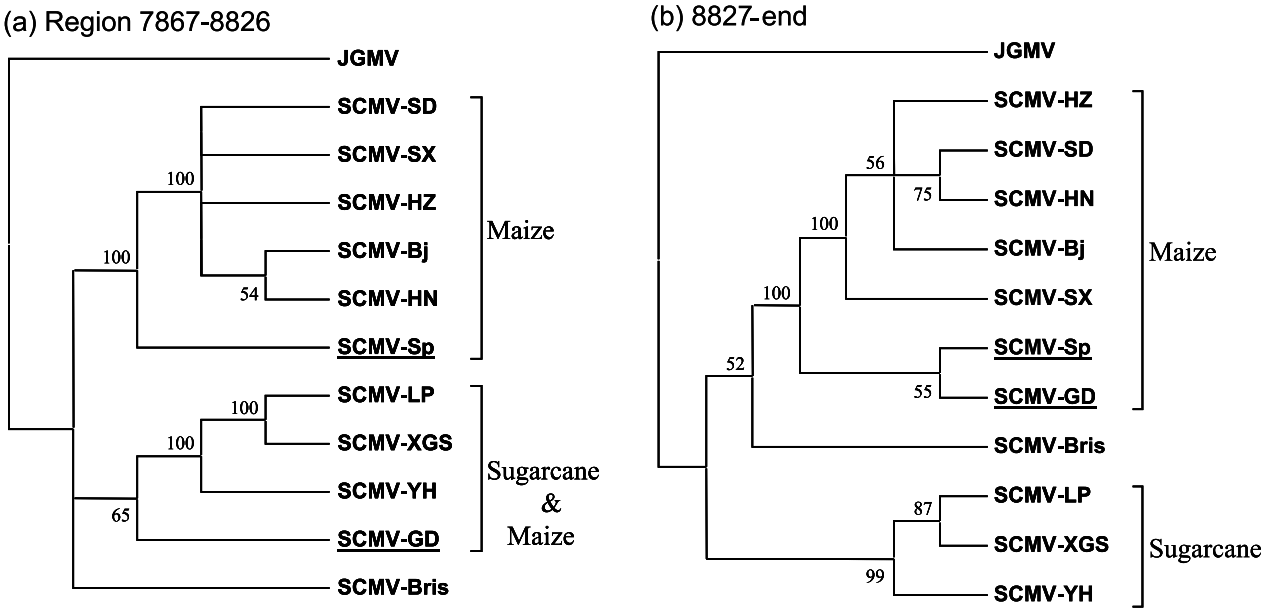

Fig. 3. Maximum-likelihood trees inferred from sequence alignments from either side of the recombination break-points identified in the CP of SCMV-GD by PhylPro: nucleotide sequences between the recombination signals identified in the NIb of SCMV-Sp and in the CP of SCMV-GD (**a**), nucleotide sequences after the recombination breakpoint at position 8827 in SCMV-GD (**b**). Trees were constructed, rooted and bootstrapping performed as described in Fig. 2. Abbreviations of virus isolates are those given in Fig. 1

Bank, we had already shown that the CP variability was related to the host from which isolates had been obtained [3]. Alegria et al. [5] also demonstrated that similarity levels in the CPs of SCMV isolates varied more between hosts than between geographical locations. Host-related sequences, mainly in P1, P3 and the CP could be attributed to the fact that these proteins are involved in symptomatology, pathogenesis and movement [19, 23], whilst the 3′ UTR has in turn been implicated in symptom attenuation [20]. The high levels of identity observed between isolates obtained from distant geographic areas (SCMV isolates from Spain and China) could be explained by movement of maize germplasm since SCMV can be seed transmitted [13, 17].

This is the first description of the complete genome sequence of an SCMV isolate from Europe and the second determination of the entire genomic sequence of a European isolate of MDMV. This study also demonstrates the occurrence of recombination in SCMV and suggests a common origin of European and Asia SCMV isolates from maize and sugarcane.

Acknowledgments

This study was supported by the project AGL2003-04077 of MCYT from Spain. We thank R. Alves (from UdL of Spain) for his help with recombination analysis.

References

1. Achon MA, Medina V, Shanks M, Markham P, Lomonossoff GP (1994) Characterisation of a maize-infecting potyvirus from Spain. Eur J Plant Pathol 100: 157–165
2. Achon MA, Sobrepere M (2001) Incidence of potyviruses in commercial maize fields and their seasonal cycles in Spain. Z PflKrankh Pfl Schutz 108: 399–406
3. Achon MA, Sobrepere M, Minguell R (2003) Molecular and biological properties of a Sugarcane mosaic potyvirus isolate from Spain. Z Pflkrankh Pflshutz 110: 324–331
4. Adams MJ, Antoniw JF, Fauquet CM (2005) Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol 150: 459–479

5. Alegria OM, Royer M, Bousalem M, Chatener M, Peterschmitt M, Girard JC, Rott P (2003) Genetic diversity in the coat protein coding region of sugarcane eighty-six mosaic virus isolates from eight countries particularly from Cameroon and Congo. Arch Virol 148: 357–372
6. Chare ER, Holmes EC (2006) A phylogenetic survey of recombination frequency in plant RNA viruses. Arch Virol 151: 933–946
7. Chen J, Chen J, Adams MJ (2002) Charaterisation of potyviruses from sugarcane and maize in China. Arch Virol 147: 1237–1246
8. Fan ZF, Cheng HY, Liang XM, Li HF (2003) Complete sequence of the genomic RNA of the prevalent strain of a potyvirus infecting maize in China. Arch Virol 148: 773–782
9. Gough KH, Shukla DD (1993) Nucleotide sequence of Johnsongrass mosaic potyvirus genomic RNA. Intervirology 36: 181–192
10. Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704
11. Kong P, Steinbiss HH (1998) Complete nucleotide sequence and analysis of the putative polyprotein of maize dwarf mosaic virus genomic RNA (Bulgarian isolate). Arch Virol 143: 1791–1794
12. Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for Molecular evolutionary Genetic Analysis and sequence aligment. Brief Inform 5: 150–163
13. Li L, Wang X, Zhou G (2007) Analyses of maize embryo invasion by sugarcane mosaic virus. Plant Sci 172: 131–138
14. Martin DP, Williamson C, Posada D (2005) RDP2: Recombination detection and analysis from sequence alignments. Bioinformatics 21: 260–262
15. Nicolas O, Laliberte JF (1992) The complete nucleotide sequence of turnip mosaic potyvirus RNA. J Gen Virol 73: 2785–2793
16. Oertel U, Schubert JE, Fuchs E (1997) Sequence comparison of the 3′-terminal parts of the RNA of four German isolates of sugarcane mosaic potyvirus (SCMV). Arch Virol 142: 675–687
17. Oertel U, Fuchs E, Hohmann F (1999) Differentiation of isolates of sugarcane mosaic potyvirus (SCMV) on the basis of molecular, serological and biological investigations. Z PflKrankh Pfl Schutz 106: 304–313
18. Revers F, Le Gall O, Le Romancer M, Dunez J (1996) Frequent occurrence of recombinant potyvirus isolates. J Gen Virol 77: 1953–1965
19. Revers F, Le Gall O, Candresse T, Maule AJ (1999) New advances in understanding the molecular biology of plant/potyvirus interactions. Mol Plant Microbe Interact 12: 367–376
20. Rodriguez-Cerezo E, Gamble-Klein P, Shaw JG (1991) A determinant of disease symptom severity is located in the 3′-terminal noncoding region of the RNA of a plant virus. Proc Natl Acad Sci 88: 9863–9867
21. Seifers DL, Salomon R, Marie-Jeanne V, Alliot B, Signoret P, Haber S, Loboda A, She YM, Standing KG (2000) Characterization of a novel potyvirus isolated from maize in Israel. Phytopathology 90: 505–513
22. Shukla DD, Ward CW, Brunt AA (1994) The sugarcane mosaic virus subgroup. In: The potyviridae. CAB International, Wallingford, pp 360–371
23. Urcuqui-Inchima S, Haenni AL, Bernadi F (2001) Potyvirus proteins: a wealth of functions. Virus Res 74: 157–175
24. Weiller GF (1998) Phylogenetic profiles: a graphical method for detecting genetic recombination in homologous sequences. Mol Biol Evol 15: 326–335
25. Yang ZN, Mirkov TE (1997) Sequence and relationships of sugarcane mosaic and sorghum mosaic virus strains and development of RT-PCR-based RFLPs for strain discrimination. Phytopathology 87: 932–939
26. Zhong Y, Guo A, Li C, Zhuang B, Lai M, Wei C, Luo J, Li Y (2005) Identification of a naturally occurring recombinant isolate of sugarcane mosaic virus causing maize dwarf mosaic disease. Virus Genes 30: 75–83