Virus Genes (2015) 50:172–175
DOI 10.1007/s11262-014-1134-0

Complete genome sequencing of *Piper yellow mottle virus* infecting black pepper, betelvine, and Indian long pepper

K. P. Deeshma · A. I. Bhat

Received: 30 June 2014 / Accepted: 14 October 2014 / Published online: 21 October 2014

Abstract The complete genome of the *Piper yellow mottle virus* (PYMoV), a *Badnavirus* belonging to the family *Caulimoviridae*, was sequenced from three naturally infected hosts namely, black pepper, betelvine, and Indian long pepper. The genome length of the three virus strains (one from each of the three host species) varied from 7,559 to 7,584 nucleotides, and all the three strains possessed four open reading frames (ORFs) I to IV that potentially encode proteins of 15.67, 17.08, 218.6, and 17.22 kDa, respectively. ORF III encodes a polyprotein consisting of viral movement protein, trimeric dUTPase, zinc finger, aspartic protease, reverse transcriptase, and RNase H whereas ORF I, II, and IV encode proteins of unknown functions. The complete genome sequences at the nucleotide level were 89–99 % identical with one available sequence of PYMoV and 39–56 % identical with other badnaviruses, indicating that all three are strains of PYMoV. Nucleotide and amino acid sequences of ORF I–IV and of the intergenic region (IR) were 80–100 % identical among PYMoV strains. Phylogenetic analysis of ORF III amino acid sequences showed the PYMoV strains forming a distinct cluster well separated from other badnaviruses. Among other badnaviruses, *Fig badnavirus* 1 (FBV-1) was the one most closely related to PYMoV.

Keywords Full genome sequence · Sequence analysis · Phylogeny · Badnavirus · Open reading frame

Black pepper, betelvine, and Indian long pepper (family *Piperaceae*) are tropical perennial climbers cultivated in most of the South Asian and East Asian countries. Black pepper, the most important and widely used spice all over the world and hence sometimes referred to as 'the king of spices', is the mature dried berry of black pepper (*Piper nigrum*) plant. Black pepper has a characteristic pungency and aroma and is well known for medicinal properties [1, 2]. Betelvine (*Piper betle*) is grown for leaves which is valued as a mild stimulant and also has medicinal properties [2]. Indian long pepper (*Piper longum*) is cultivated for fruits which possesses diverse pharmacological applications [2].

Piper yellow mottle virus (PYMoV) is a *Badnavirus* (family *Caulimoviridae*) and infects black pepper, betelvine, and Indian long pepper [3–5]. The virus, which has been reported from Brazil and many Asian countries including India, Sri Lanka, Malaysia, Thailand, and Philippines [3–6], induces chlorotic mottling, vein clearing, leaf distortion, reduced plant vigor, and poor fruit set in affected plants [3, 4]. PYMoV is transmitted primarily through vegetative means (stem cuttings) and seeds and secondarily through various species of the mealybug [3, 4, 7, 8].

So far, the complete genome of only one PYMoV isolate has been sequenced and published [9]; the isolate, from Kozhikode in the southern Indian state of Kerala, infects black pepper 'Karimunda' and has 7,622 nucleotides with four open reading frames (ORFs). However, some information, although limited, is available about PYMoV strains from other *Piper* species and the variability of those strains. In the present study, complete genomes of PYMoV isolates from black pepper, betelvine, and Indian long pepper were sequenced and compared among themselves and with those of other badnaviruses.

Diseased plants showing typical symptoms of PYMoV infection were collected from black pepper (Panniyur 1 from Kodagu, Karnataka, India) and from betelvine and Indian long pepper (from Kozhikode). The DNA isolated

K. P. Deeshma · A. I. Bhat (✉)
Division of Crop Protection, Indian Institute of Spices Research, Marikunnu PO, Kozhikode 673 012, India
e-mail: aib65@yahoo.co.in

from partially purified virus preparations as described earlier [9] was used as a template for amplifying the viral genome. The complete genome of PYMoV was amplified using two sets of primers that generated overlapping sequences at the ends. Primers AIB 104 (5′CTATATGAA TGGCTAGTGATG3′) and AIB 105 (5′TTCCTAGG TTTGGTATGTATG3′), were designed based on multiple alignment of six PYMoV partial sequences available in GenBank whereas primers AIB 150 (5′GGAAGTAGAC CTAACAAAAGG3′) and AIB 151 (5′ATGAGTGATGAA GTAGTCAAC3′) were designed based on multiple alignment of one complete genome of PYMoV [KC808712] and other badnavirus sequences from GenBank. PCR was performed using LongAmp Taq DNA polymerase kit (New England Biolabs, UK) and 1 µl template DNA (0.1 µg). The thermocycler was programmed for initial denaturation at 94 °C for 30 s, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 47 °C (for primers AIB 104 and AIB 151) or 58 °C (for primers AIB 105 and AIB 151) for 1 min, synthesis at 65 °C for 5 min, and a final extension at 65 °C for 10 min. PCR products were analyzed by electrophoresis and specific DNA bands were eluted, cloned into vector pTZ57R/T (Fermentas, Glen Burnie, Maryland, USA), and sequenced at Chromous Biotech Pvt. Ltd, Bengaluru, India, by primer walking. Two sequencing reactions were performed for each of the clones and a third sequencing was carried out only for those regions where consensus sequence was not obtained in the first two reactions. The final sequence assembly was done using the consensus sequences from these three reactions and ORFs were identified using CLC sequence viewer. Multiple sequence alignment and percent identity was determined using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/). Phylogenetic analysis was performed using different methods such as Bayesian analysis (MrBayes version 3.1), maximum likelihood (GARLI version 2.0) and Minimum Evolution method (MEGA5) each for a bootstrap analysis of 10,000 replicates and the consensus tree was identified using consense package of Phylip 3.69.

PCR gave a 4 kb product with primers AIB 104/AIB 151 and a ~4.5 kb product with primers AIB 105/AIB150 for all the three isolates whereas PCR with samples from healthy plants showed no amplification. The sequences were assembled, analyzed, ORFs were predicted using CLC sequence viewer and submitted to GenBank and are available as accession numbers KJ873041, KJ873042, and KJ873043 for black pepper, betelvine, and Indian long pepper strains of PYMoV, respectively. The complete genomes of PYMoV from black pepper, betelvine, and Indian long pepper were 7584, 7559, and 7580 bp long, respectively, with a GC content of 43 % for black pepper, 45 % for betelvine, and 44 % for Indian long pepper.

Each of the strains contained four ORFs, namely I, II, III, and IV, with overlapping sequences at the ends. The numbers of nucleotides in ORF I (408), ORF II (465), ORF III (5778), and ORF IV (459) were the same in all the three strains and in the published PYMoV strain [9]. Although many badnaviruses have only three ORFs, a few, such as *Dracaena mottle virus* (DMV), *Cacao swollen shoot virus* (CSSV), *Citrus yellow mosaic virus* (CYMV), and *Sweet-potato badnavirus* a (SPBVa) and SPBVb contain four to seven ORFs [10, 12–14]. All the three PYMoV strains in the present study have a fourth ORF, as reported in PYMoV strain ISH-1 [9]. The TGA stop codon (nt 1112–1114, 1086–1089, and 1108–1110 in the PYMoV strains from black pepper, betelvine, and Indian long pepper respectively) of ORF I overlaps with the ATG start codon (nt 1111–1113, 1086–1088, and 1107–1109) of ORF II with a sequence ATGA, as in many other badnaviruses [12]. In all three strains, an overlap of 77 bp was seen between ORF II and ORF III whereas ORF IV overlapped largely with ORF III, with C-terminal portion (190 nt) extending outside of ORF III, as in PYMoV strain ISH-1

Table 1 The extent of shared identities in nucleotide sequences of the complete genome (below the diagonal) and intergenic region (IR) (above the diagonal) and deduced amino acid sequences of ORF I (below the diagonal), ORF II (above the diagonal), ORF III (below the diagonal), and ORF IV (above the diagonal) among different PYMoV strains

	Complete genome and IR				ORF I and ORF II				ORF III and ORF IV			
PYMoV strain	PN-ISH-1	PN-P1	PB	PL	PN-ISH1	PN-P1	PB	PL	PN-ISH1	PN-P1	PB	PL
PN-ISH-1[a]	–	90	84	80	–	100	83	84	–	100	91	91
PN-P1[b]	99	–	89	87	97	–	83	84	99	–	91	91
PB[b]	90	90	–	87	90	92	–	94	97	97	–	94
PL[b]	89	89	92	–	89	91	94	–	95	94	96	–

PN-ISH-1 PYMoV strain from black pepper Karimunda, *PN-P1* PYMoV strain from black pepper Panniyur-1, *PB* PYMoV strain from betelvine, *PL* PYMoV strain from Indian long pepper

[a] Sequence obtained from GenBank (KC808712)

[b] Sequences obtained in this study

[9]. The size and arrangement of ORF I, II, III, IV, and intergenic region (IR) of PYMoV followed those of other badnaviruses [11–13]. ORF I, II, and IV encode proteins of unknown functions with molecular weights of 15.6, 17.5, and 17.1 kDa, respectively. ORF III of PYMoV encodes a polyprotein with a molecular weight of 218.6 kDa, consisting of a viral movement protein (MP), trimeric dUTPase, zinc finger (ZnF), aspartic protease (AP), reverse transcriptase (RT), and RNase H, as reported in PYMoV strain ISH-1 and other badnaviruses [9–13].

The lengths of IR of the three PYMoV strains (black pepper, betelvine, and Indian long pepper) were 844, 819, and 840 nucleotides, respectively. The slight variation in length of genome of three strains was due to the variation in the length of the IR (a few deletions, additions, and repetitive sequences) although the number of nucleotides in each ORF remained the same in all the strains. IR of all three strains included an 18 nucleotide consensus plant cytoplasmic initiator methionine tRNA binding site (5′-TGGTATCAGAGCATGGTT-3′), which is highly conserved among all members of the *Caulimoviridae* family [10–15], and a putative poly(A) sequence (AATAAA) upstream of a $tRNA^{meth}$ binding site (staring from nt 7468, 7443, and 7464 in the PYMoV strains from black pepper, betelvine, and Indian long pepper respectively). However, unlike many badnaviruses [10–13], the TATA box promoter element upstream of the $tRNA^{met}$ binding site was not found in any of the PYMoV strains.

A comparison of the complete genome sequences revealed close identity (89–99 %) among all PYMoV strains including the one reported earlier (Table 1). The black pepper strain of PYMoV was 99 % identical with the previously reported PYMoV strain from black pepper [KC808712] and the strains from betelvine and Indian long pepper were 89–90 % identical with PYMoV strains from black pepper. The betelvine and Indian long pepper strains of PYMoV shared 92 % identity between them (Table 1). The corresponding ORFs (I, II, III, and IV) of all four strains of PYMoV were 90–98 %, 86–99 %, 90–99 %, and 91–100 % identical at the nucleotide level and 89–97 %, 84–100 %, 94–99 %, and 91–100 % identical at the amino acid level. The percent identity of individual proteins of ORF III such as MP, AP, and RT/RNase H ranged from 97 to 100 % among different strains of PYMoV. The IR sequences of the PYMoV strains were more variable (9–14 %) than the coding regions. Similar levels of sequence identities were reported among strains of CYMV infecting different species of citrus [16, 17] and among those of *Dioscorea sansibarensis bacilliform virus* (DsBV) infecting different species of *Dioscorea* [11]. All four strains of PYMoV were 39–58 % identical with other badnaviruses when complete genome sequences were considered (data not shown) and among them *Fig badnavirus* 1(FBV-1) was found to be the closest to PYMoV. A comparison among badnaviruses including PYMoV revealed that identities were shared to a greater extent in ORF III (45–60 %) than in ORF I and ORF II (25–50 %), indicating that ORF III is more conserved as reported by earlier workers [9, 12, 15].

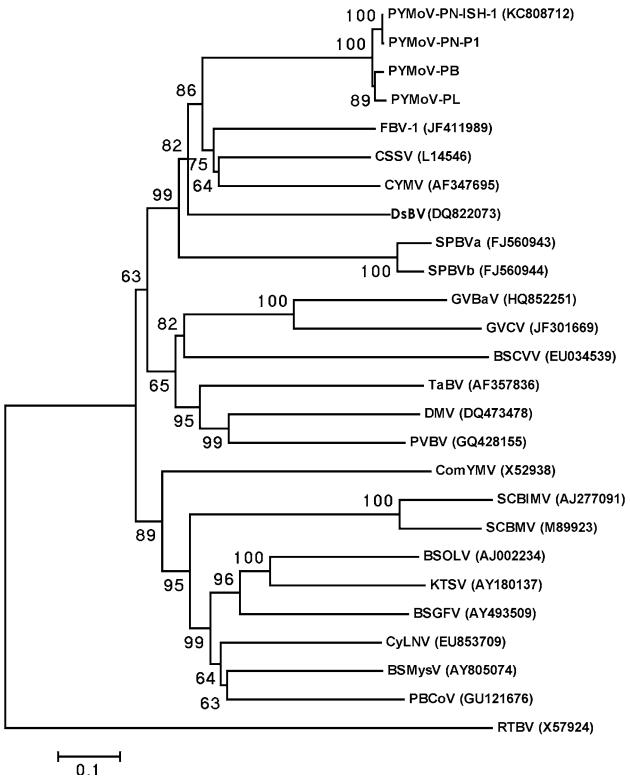

Fig. 1 Phylogenetic tree drawn using consense package of Phylip showing the relationships between the deduced amino acid sequences of ORF III of PYMoV strains and those of other badnaviruses. *Banana streak GF virus* (BSGFV), *Banana streak Mysore virus* (BSMysV), *Banana streak OL virus* (BSOLV), *Bougainvillea spectabilis chlorotic vein-banding virus* (BSCVV), *Cacao swollen shoot virus* (CSSV), *Citrus yellow mosaic virus* (CYMV), *Commelina yellow mottle virus* (ComYMV), *Cycad leaf necrosis virus* (CyLNV), *Dioscorea sansibarensis bacilliform virus* (DsBV), *Dracaena mottle virus* (DMV), *Fig badnavirus* 1 (FBV-1), *Gooseberry vein banding virus* (GVBaV), *Grapevine vein-clearing virus* (GVCV), *Kalanchoe top-spotting virus* (KTSV), *Pelargonium vein banding virus* (PVBV), *Pineapple bacilliform comosus virus* (PBCoV), *Sugarcane bacilliform IM virus* (SCBIMV), *Sugarcane bacilliform Mor virus* (SCBMV), *Sweetpotato badnavirus A* (SPBVa), *Sweetpotato badnavirus B* (SPBVb), *Taro bacilliform virus* (TaBV). *PYMoV-PN-ISH-1* PYMoV strain from black pepper Karimunda, PYMoV-*PN-P1* PYMoV strain from black pepper Panniyur-1, *PYMoV-PB* PYMoV strain from betelvine, *PYMoV-PL* PYMoV strain from Indian long pepper. Bootstrap values are shown at the nodes. *Rice tungro bacilliform virus* (RTBV) was taken as an outgroup

The putative amino acids of ORFs of all four PYMoV strains were aligned with those of FBV-1 (the virus close to PYMoV), GVCV (the badnavirus most distant to PYMoV based on alignment of ORF III amino acids), and *Commelina yellow mottle virus* (ComYMV), which is the type

species for the genus *Badnavirus*. Conserved motifs of 4–25 amino acids between ORF III of PYMoV and FBV numbered 89; those between PYMoV and GVCV numbered 46; and those between PYMoV and CoYMV numbered 48. Of these motifs, 18 were found conserved in all four, of which the following eight were located in the RT-RNase H region, the highly conserved genomic region among members of *Calimoviridae*: KSTIDAEI, RSWLG, FIAVYIDDILVFS, FQRKMD, LYEWLVMPFG, WTAF, SKFDLKSGF, and DQYSLPGI [12].

Phylogenetic tree analysis using ORF III amino acid sequences of 23 badnaviruses including PYMoV strains showed close clustering of all four PYMoV strains (Fig. 1). The phylogenetic tree analysis could be divided into three major groups. The first group included 7 badnaviruses consisting of all strains of PYMoV along with FBV-1, CSSV, CYMV, DsBV SPBVa, and SPBVb (Fig. 1). The phylogenetic tree analysis confirmed that FBV-1 is the most closely related virus to PYMoV. We therefore conclude that the virus isolates from black pepper, betelvine, and Indian long pepper are all strains of PYMoV. Overall, there was not much diversity (1–11 %) among the three strains of PYMoV and the already reported PYMoV strain ISH-1.

Acknowledgments The authors are thankful to the Distributed Information Sub Centre (DISC) and the Director, Indian Institute of Spices Research, Kozhikode, for providing necessary facilities. The first author is thankful to Kerala State Council for Science Technology and Environment, Thiruvananthapuram, for a fellowship.

References

1. P.N. Ravindran, *Black pepper: Piper-nigrum, series: medicinal and aromatic plants—industrial profiles*, 1st edn. (Harwood Academic Publishers, Amsterdam, 2000), pp. 1–13
2. K.R. Kirtikar, B.D. Basu, *Indian medicinal plants*, 2nd edn. (Lalit Mohan Basu Publications, Allahabad, 1933), pp. 2128–2135
3. B.E.L. Lockhart, K.A. Kirtisak, P. Jones, D.D. Padmini, N.E. Olszieewski, N. Lockhart, D. Nuarchan, J. Sangalang, Eur. J. Plant Pathol. **103**, 303 (1997)
4. P.S. Hareesh, A.I. Bhat, Indian J. Virol. **19**, 160 (2008)
5. S. Siju, A.I. Bhat, P.S. Hareesh, J. Plant Biochem. Biotech. **17**, 73 (2008)
6. D.P.P. de Silva, P. Jones, M.W. Shaw, Plant Pathol. **51**, 537 (2002)
7. A. I. Bhat, S. Devasahayam, P. S. Hareesh, N. Preeti, T. Thomas, Entomon. 85 (2005)
8. P.S. Hareesh, A.I. Bhat, J. Plant. Crop. **38**, 62 (2010)
9. U. Hany, I.P. Adams, R. Glover, A.I. Bhat, N. Boonham, Arch. Virol. **159**, 385 (2014)
10. L. Su, S. Gao, Y. Huang, C. Ji, D. Wang, Y. Ma, R. Fang, X. Chen, Virus Genes **35**, 423 (2007)
11. S. Seal, E. Muller, Arch. Virol. **152**, 819 (2007)
12. Q. Huang, J.S. Hartung, Gen. Virol. **82**, 2549 (2001)
13. L.S. Hagen, M. Jacquemond, A. Lepingle, H. Lot, M. Tepfer, Virology **196**, 619 (1993)
14. J.F. Kreuze, A. Perez, M. Untiveros, D. Quispe, S. Fuentes, I. Barker, R. Simon, Virology **388**, 1 (2009)
15. I.C. Yang, G.J. Hafner, J.L. Dale, R.M. Harding, Arch. Virol. **148**, 937 (2003)
16. B.K. Borah, A.M.A. Johnson, D.V.R. Sai, Gopal, I. Dasgupta. Virus Genes **39**, 137 (2009)
17. A.M.A. Johnson, B.K. Borah, D.V.R. Sai Gopal, I. Dasgupta, Virus Genes **45**, 600 (2012)