Virus Genes 28:3, 239–246, 2004

Complete Nucleotide Sequence of a Strawberry Isolate of *Beet pseudoyellows virus**

IOANNIS E. TZANETAKIS[1] & ROBERT R. MARTIN[2,†]

[1]*Molecular and Cellular Biology Program, Department of Botany and Plant Pathology, Oregon State University, Corvallis 97331, USA*

[2]*Horticultural Crops Research Laboratory, USDA-ARS, Corvallis OR 97330, USA*

Received September 9, 2003; Accepted November 23, 2003

Abstract. In our effort to identify the causal agent(s) of strawberry pallidosis we found a single pallidosis positive plant that did not give any amplicons after RT-PCR using primer sets representing multiple regions of *Strawberry pallidosis associated virus* (SPaV) genome and failed to react with antibodies directed against the recombinant coat protein (CP) of SPaV. DsRNA extracted from this plant showed a similar pattern to that of SPaV indicating that another crinivirus may infect strawberry. Sequence analysis of multiple cDNA clones corresponding to the heat shock 70 homolog gene (HSP70h) of the unknown virus indicated that it was *Beet pseudoyellows virus* (BPYV). Analysis of the complete nucleotide sequence of BPYV-strawberry revealed that this isolate has several distinct features when compared to *Cucumber yellows virus* (CuYV), a cucumber strain of BPYV, including an entire ORF not found in CuYV.

Key words: *Beet pseudoyellows virus*, *Closterovirus*, *Crinivirus*, dsRNA, nucleotide sequence, pallidosis, strawberry, *Strawberry pallidosis associated virus*

Introduction

Closteroviruses have the largest genome of all plant positive-strand RNA viruses, with genomes close to 20 kbp in size [1]. The viruses in the family are transmitted in a semi-persistent manner by aphids (genus *Closterovirus*), whiteflies (genus *Crinivirus*) or mealybugs (genus *Ampelovirus*) and are generally phloem-limited [2]. Viruses in the genera *Closterovirus* and *Ampelovirus* are monopartite while criniviruses have a divided genome. Closteroviruses infect plants of agriculturally important families such as Cucurbitaceae, Rosaceae, Solanaceae, Compositae, Vitaceae and Rutaceae, causing significant economical losses [3].

Pallidosis is a disease of strawberry (*Fragaria* × *anannasa*) first identified in 1957 [4] and may be the first disease described being caused by a *Crinivirus* [5]. *Strawberry pallidosis associated virus* (SPaV), a putative causal agent of the pallidosis disease was found in all but one of the 38 plants used in studies to characterize that virus [5]. This plant failed to give any positive results in RT-PCR tests using oligonucleotide primers developed against the HSP70h, coat protein (CP), minor coat protein (CPm) and p28 genes of SPaV (Tzanetakis and Martin, unpublished data) as well as tissue blot immunoassay using polyclonal antibodies derived against the recombinant major coat protein of the virus expressed in bacteria. The dsRNA pattern of this single isolate (M29) gave a large (~8 kbp) band and several others that most

*The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers AY 330918, 20 AY 330919, NC 005209 and NC 005210.

†Author for all correspondence:
E-mail: martinrr@science.oregonstate.edu

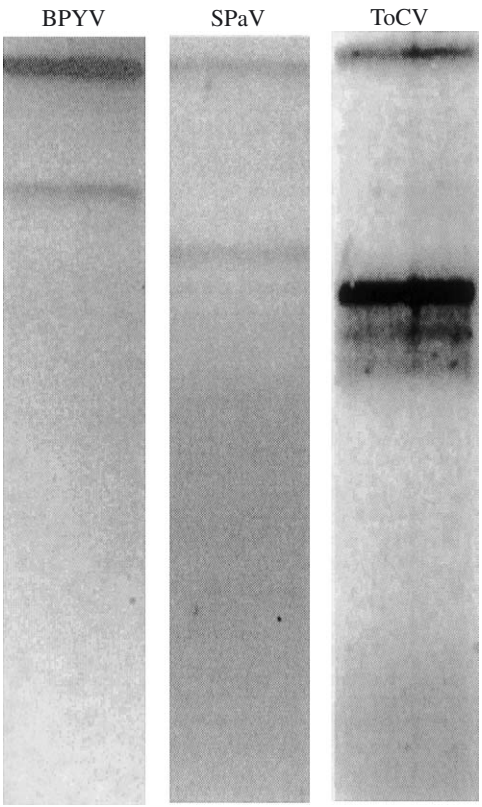

Fig. 1. DsRNA pattern of three Criniviruses, BPYV, SpaV and ToCV. The genomic sizes (top band) represent RNA 1 and RNA 2 and are close to 8 kbp for BPYV and SpaV and for ToCV (actual size unknown but larger than BPYV and SpaV). The size of the subgenomics is similar but the abundance of the different subgenomics varies between the viruses.

probably represent subgenomic RNAs of the virus. The dsRNA pattern suggested that a second crinivirus may be present in strawberry (Fig. 1). We cloned and sequenced cDNA derived from dsRNA from this plant and compared the sequences acquired against those found in GenBank. The analysis indicated that the unknown virus had close to 99% identity with the published sequence of *Beet pseudoyellows virus* (BPYV) heat shock protein 70 homolog (HSP70h) [6].

BPYV, the crinivirus with the largest host range known today [7], was first identified in 1965 [8] and infects plants in the temperate and subtropical regions around the world [9]. The virus causes yellowing diseases in plants belonging to the families Cucurbitaceae, Solanaceae and Chenopodiaceae among others [8], and is solely transmitted by the greenhouse whitefly, *Trialeuroides vaporariorum,* with acquisition time of less than 1 h. The whiteflies remain virulent for 6 days, making the insect an excellent disseminator of the virus [7].

Recently, the complete nucleotide sequence of *Cucumber yellows virus* (CuYV), a strain of BPYV was published [10]. In this study, we present the complete nucleotide sequence of a strawberry isolate of BPYV (BPYV-strawberry) which bears unique features not found in CuYV, including an additional ORF.

Materials and Methods

Plant Material

A strawberry plant from Maryland initially thought to be infected with an isolate of SPaV, was the only isolate of a pallidosis study not infected with SPaV [5]. We re-grafted the M29 plant onto indicator plants *F. vesca* (''UC-4'' or ''UC-5'') and *F. virginiana* (''UC-10'' or ''UC-11''). *F. virginiana* plants showed typical pallidosis symptoms including leaf marginal chlorosis and cachexia (data not shown) while *F. vesca* plants remained asymptomatic. Thus, the plant was positive for pallidosis, based on the symptoms in indicator plants. After determining the presence of BPYV by cloning and sequencing from dsRNA template purified from the original M29 strawberry plant, the isolate was transferred to *Nicotiana benthamiana* using the whitefly vector of the virus, *Trialeuroides vaporariorum*, as a vector.

cDNA Synthesis and Cloning

Initial dsRNA extraction was performed with tissue from the M29 isolate as described previously [11]. After confirmation of infection of the plant with BPYV, dsRNA was extracted from *Nicotiana benthamiana* plants infected with the M29 isolate. Unless otherwise stated all enzymatic reactions described hereafter were performed using products of Invitrogen Corp. (Carlsbad, CA), according to the manufacturer's recommendations. cDNA was synthesized as described previously [12] using the Thermoscript® reverse transcriptase and was subsequently adenylated at the 3′ ends using *Taq* polymerase (1 unit), for 15 min at 72°C and purified utilizing the rapid PCR purification system (Marli-

gen Biosciences, Ijamsville, MD) prior to cloning. The products were then ethanol-precipitated, resuspended in 8 μl water and cloned into the pCR4 TOPO vector. Plasmids were purified, digested with *Eco* RI (New England Biolabs, Beverly, MA) and analyzed using agarose gel electrophoresis.

Amplification of BPYV Genome

Template for RT-PCR reactions was either dsRNA extracted as described above or ssRNA extracted using the Hughes and Galau method [13]. The original cDNA synthesis cloning provided sequence data that contained more than 14 kb of the viral genome but there were not a sufficient number of clones to build contigues other than nucleotide regions from approximately 1200–1800 and 4000–5600 of RNA 1 as well as nucleotide regions from about 60–1600 and 5000–7400 of RNA 2 which were assembled from clones derived from cDNA synthesis. These regions were represented by at least three clones per site. All the primers used for PCR amplification were developed after alignment of at least two cDNA clones. Reverse transcription was performed as described above while for PCR we used the *Platinum Taq* polymerase. The PCR program consisted of original denaturation for 5 min at 94°C followed by 40 cycles of 30 s at 94°C, 30 s at 50 or 55°C annealing, depending on the primer melting temperature and extension of 1–2 min at 68°C depending on the length of the amplified fragment. The reaction concluded after a final extension step at 68°C for 15 min. Each fragment was amplified twice and cloned into the pCR4 TOPO vector.

Sequence contigs were assembled after sequencing the PCR products; two clones in both directions using the Ml3 forward and reverse primers and clones of the original cDNA cloning reaction where applicable. For amplification of the 3′ ends of the virus we poly-adenylated dsRNA [14] and performed the reverse transcription step as described [12] utilizing random nucleotide hexamers and an oligo thymidine primer. Amplification of the 3′ ends was performed twice using the same parameters as above. For amplification of the 5′ termini we used two approaches. We developed oligonucleotide primers from the minus-strand sequence of the viral dsRNA and utilized the poly-adenylated dsRNA as template and two individual primers with the oligo thymidine primer. We also employed 5′ rapid amplification of cDNA ends (RACE) using ssRNA extracted as above as the reaction template. Each of the 5′ terminus amplification reactions was performed once.

Nucleotide Sequencing and Phylogenetic Analysis

All sequencing reactions were performed at the Macrogen Inc. facilities (Seoul, Korea) using an ABI 3700 DNA sequencer. The assembly of the virus genome was performed utilizing the ClustalW software (European Bioinformatics Institute http://www.ebi.ac.uk/clustalw/). The search for transmembrane domains in BPYV encoded proteins were performed using the TMHMM Server version 2.0 for prediction of transmembrane helices in proteins (http://www.cbs.dtu.dk/services/TMHMM/). The ORFs where verified using both the NCBI ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and the gene finder in viruses at http://www.softberry.com. The amino acid comparisons of the crinivirus ORFs were performed utilizing Vector NTI suite 6 (Informax, Bethesda, MD) For the RNA secondary structure prediction to investigate 3′ UTRs structure we used the software at http://www.genebee.msu.su (A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University). For identification of conserved domains the software at the protein motif website (Conserved Domain Architecture Retrieval Tool-hrtp://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi?cmd = rps) was used.

Results and Discussion

Closteroviridae is a very diverse virus family. The genome size and the genes encoded by members of the family varies greatly from that of *Citrus tristeza virus* [15] to CuYV [10]. The general organization of BPYV-strawberry resembles that of other sequenced criniviruses [16–19]. The overall nucleotide identity to CuYV for RNA 1 is 91% while for RNA 2 it reaches 98%, while the amino acid identities range from 94% to 100% (Table 1). Both RNAs of BPYV-strawberry have the same first 11 nucleotides while there is 91% nucleotide identities in the last 156 nucleotides at the virus 3′ UTRs which from stem-loops (data not shown)

Table 1. Amino acid identities and similarities (in parenthesis) between the polypeptides encoded by the strawberry isolate of BPYV compared to CuYV isolate of BPYV, CYSDV, SPCSV and LIYV ORFs

Gene/virus	CuYV	CYSDV	SPCSV	LIYV
1a	94 (96)	34 (50)	36 (51)	29 (45)
RNA1 1b (RdRp)	98 (98)	61 (74)	62 (78)	55 (71)
RNA1 p6	–	–	–	–
p9	97 (100)	12 (29)	16 (30)	31 (36)
HSP70h	99 (99)	62 (79)	63 (79)	49 (68)
p6	100	30 (43)	–	–
p59	98 (99)	39 (55)	38 (54)	32 (48)
p10	100	28 (42)	29 (39)	28 (29)
CP	99 (99)	40 (52)	29 (39)	26 (37)
CPm	97 (98)	26 (46)	23 (37)	19 (27)
p26	98 (98)	25 (46)	17 (35)	14 (29)

that may be a signal for the viral polymerase and facilitate replication [20].

RNA 1 of BPYV-strawberry consists of 8007 nucleotides and encodes three ORFs (Fig. 2). The first ORF of RNA1 starts at nucleotide 244 and terminates at nucleotide 6054. It encodes a protein of 1936 amino acids and the estimated molecular weight is 221 kDa. The overall amino acid identity with the CuYV strain is 94% (95% similarities). It encodes a multifunctional polyprotein with cysteine/papain protease, methyltransferase and helicase motifs found in all closteroviruses [2]. Protein database search revealed that the cysteine protease is encoded in the N′ terminus of the polyprotein, followed by the methyltransferase domain while the helicase domain is encoded in the C′ terminus of the protein.

The first major difference between BPYV-strawberry and CuYV is a 147 nucleotides insertion at nucleotide 3355, after the termination of the methyltransferase domain (Fig. 2). The nucleotide identities between CuYV and BPYV-strawberry before the insertion is 86% while after it, it reaches 94%. The polyprotein has two transmembrane domains in the region between the methyltransferase and helicase (amino acids 1194–1216 and 1298–1320) that could be the anchors for localization of the protein in the vesicles as previously shown for *Beet yellow virus* [21]. The termination signal for the 1a ORF is UUUGA which is also found in *Lettuce infectious yellow virus* (LIYV), *Sweet potato chlorotic stunt virus* (SPCSV) and (SpaV) (Tzanetakis I.E., unpublished data), while *Cucurbit yellow stunting disorder virus* (CYSDV) terminates with UUUAG [19]. In all the case the three uridines found near the stop codon of ORF 1a, may accommodate the +1 ribosomal frameshift by which the viral polymerase is expressed, a novelty found only in closteroviruses among all plant viruses. ORF 1b, the viral polymerase, starts at nucleotide 6052 and terminates at nucleotide 7570. It contains the signature RNA polymerase active site GDD motiff and its molecular weight is 58 kDa. All criniviruses have the sequence: LVSGDDSLIFS – in the active site

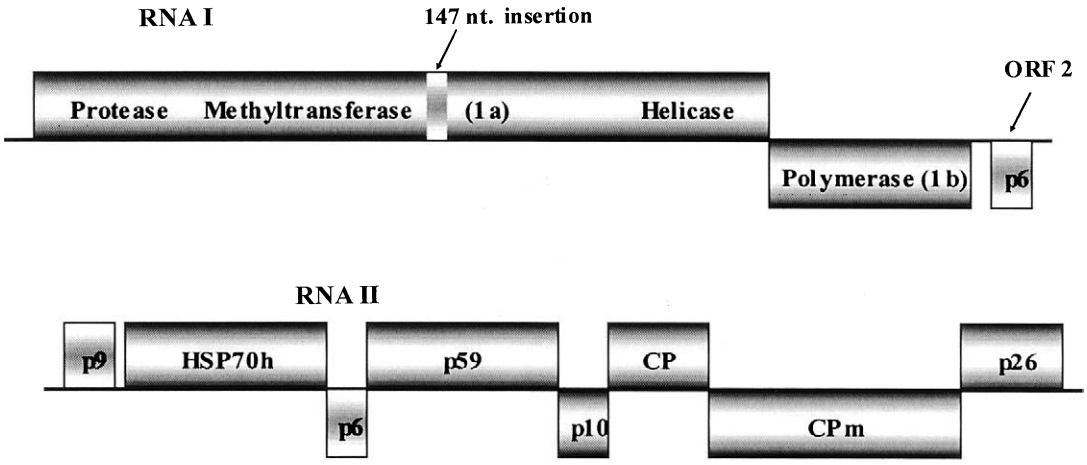

Fig. 2. Schematic representation of the genome organization of BPYV. The different shading indicates the main differences BPYV, CuYV and other criniviruses. The sizes of the ORFs are not to scale.

```
BPYV   -MQAKAGLDFGTTFSTISSFTNGEMKTLYVSNSPYIPTCLSISSEGDVIIGSAAQVIDES 59
CuYV   -MQAKVGLDFGTTFSTISSFTNGEMKTLYVNNSPYIPTCLSISSEGDVIIGSAAQVIDES 59
SPaV   MTEAKVGLDFGTTFSTISSYINNKMHVLKINDSPYIPTCLAISIDKDVIIGGAAQVLDSS 60
CYSDV  --MAKAGLDFGTTFSTISSYVNGVMKVLKLNETEFIPTCLAITSNNDVVVGGPAQVLSNS 58
SPCSV  -MEAKAGLDFGTTFSTISAYVGGTMKVLRINGSEFIPTCLSVTATGDVVVGGAAQVLDSS 59
ToCV   -MSIKAGLDFGTTFSTISCFYNNKLFSLKLNGTEYIPTCLSITPNNEVIVGGPSQVLEAS 59
LIYV   MRDCKVGLDFGTTFSTVSTLVNNSMYVLRLGDSAYIPTCIAITPGGEAIIGGAAEVLSGD 60
           * ********** *         *       ****           *    *

BPYV   EVKSCYFYDLKRWVGVDATNFLVIKEKIKPLYVVKLVGNDVYLTGVNKGFSCTYTVKQLI 119
CuYV   EVKSCYFYDLKRWVGVDATNFLVIKEKIKPLYVVKLVGNDVYYTGVNKGFSCTYTVKQLI 119
SPaV   EVANCYFYDLKRWVGVDKVNFENIKAKINPQYVAKLVNDDVMLTGVDRGYSCTYTVKQLI 120
CYSDV  DMPNCYFYDLKRWVGVDSINYNVIKTKINPVYVTELRGNDVYITGIDRGYTCTYTVKQLI 118
SPCSV  QLPHCYFYDLKRWVGVDRLSFEEIKRKISPQYTVRLEGNDVLITGISKGFSCTYTVKQLI 119
ToCV   ETPSCYFYDLKRWVGVTSVNYEVVKAKINPTYKTRLSNNKVYITGINKGFSTEFSVEQLI 119
LIYV   DTPHCFFYDLKRWVGVDDNTFKFAMNKIRPKYVAELVEGEVYLTGINKGFSIKLSVKQLI 120
           * **********          ** * *   *    *  **   *      * ***

BPYV   LLFIDTMVRLFSKTNNLNIISLNVSVPADYKCKQRMFMKSVCDSLNFSLRRIINEPSAAA 179
CuYV   LLFIDTMVRLFSKTNNLNIISLNVSVPADYKCKQRMFMKSVCDSLNFSLRRIINEPSAAA 179
SPaV   LLYIDTLVRLFSKTDNLNIISLNVSVPADYKCKQRMFMKSVCDSLNFSLRRIINEPSAAA 180
CYSDV  LLYIETLVRLFSKVESITITSLNVSVPADYKCKQRMFMKSVCDSLGFSLRRIINEPSAAA 178
SPCSV  LLYVDTLVRLFSNVEKLKILSLNVSVPADYKTKQRMFMKSVCESLGFPLRRIINEPSAAA 179
ToCV   LHYVNTLVRLFSKTENLKITDLNVSVPADYKSGQRLFMQAVCSSLGFNLRRIVNEPSAAA 179
LIYV   KAYIETIVRLLASSYSLRVIDLNQSVPADYKNAQRLAARSVLKALSFPCRRIINEPSAAA 180
            * ***           ** *******  **     *   * *  *** *******

BPYV   IYSVSKYPQHNYFIMYDFGGGTFDTSLITRDGQYVTVADTEGDSFLGGRDIDNEIQQFIV 239
CuYV   IYSVSKYPQHNYFIMYDFGGGTFDTSLITRDGQYVTVADTEGDSFLGGRDIDNEIQQFIV 239
SPaV   IYSVSKYPNYKYFLMYDFGGGTFDTSLIVRDGKVVTVADTEGDSFLGGRDIDNAISRFIV 240
CYSDV  IYFVSKYPQYNNFLMYDFGGGTFDSSLIVRDGKYVTVADTEGDSFLGGRDIDNAIADYIT 238
SPCSV  IYSISKHPGFDYFLVYDFGGGTFDTSLIAKDGKFVTVADTLGDSFLGGRDIDRAILSHIM 239
ToCV   IYCVSKYPQYAYFYIYDFGGGTFDTSLIVRYGKFVTVADTQGDSFLGGRDIDKTISKFIM 239
LIYV   VYCVSRYPNYNYFLVYDFGGGTFDVSLIGKYKSYVTVIDTEGDSFLGGRDIDKSIEDYLV 240
        *  *  *    *  ********* ***      *** ** ***********  *

BPYV   KSNNLSRPLPSDFLASIKEDCNTTGKSTFNVMDVDGKLLTIRFSREDLAACIEPYSKRSL 299
CuYV   KSNNLSRPLPSDFLASIKEDCNTTGKSTFNVMDVDGKLLTIRFSREDLAACIEPYSKRSL 299
SPaV   EKHSLPRPLSSDFLASIKEEVNNSSKSNFIALDTKGNIVNVSFNKDDLATCIQPFSVKSI 300
CYSDV  TTYGMKGGLSADVLASIKEDCNSKGRENFNVIDSSGKLHNVKFTRQDLSRCIEPFSKKSI 298
SPCSV  RTNSLQKPLSADSLAAIKEEVNSTGRSNFNVLDVDGNIIFVNFSGEELDKIVSKFTAKSL 299
ToCV   DKNALNAPLSADMLASIKEETNSTGRSSYNIISDDGSIINIQFTFDDLVKCVEPFARRSF 299
```

Fig. 3. Alignment of HSP70h proteins of BPYV-strawberry with the HSP70h of six other criniviruses. HSP70h accession nos. BPYV NP 940788, CuYV NP 821143, SpaV AAO 92347, SPCSV NP 689401, CYSDV NP 851572, ToCV AAD 01790, LIYV NP 619695.

of the polymerase while most closteroviruses have the consensus VSGDDSLI.

The putative fusion protein 1a/1b has a molecular weight of 276 kDa. Another major difference between CuYV and BPYV-strawberry is the presence of an additional ORF on RNA 1 (Fig. 2). This ORF encodes for a 54 amino acid peptide with hydrophobic residues found in transmembrane proteins and has a molecular weight of 6 kDa. This 6 kDa ORF does not show any

```
LIYV   GKYNIKKVIPATYLALIKEECNNTNKSIFTILFDDGSVQVVEFSKSELEKCVRPFVERSI 300
                    ** ***  *            *      *    *          *

BPYV   KILDNLVKRRKISSGALFLVGGSSLLSKVQQDVAAYASANNFECVIDKDLRCSVSFGCSM 359
CuYV   KILDNLVKRRKISSGALFLVGGSSLLSKVQQDVAAYASANNFECVIDKDLRCSVSFGCSM 359
SPaV   KILDNLVGRRKITNGALFLVGGSSLLKKIQQDVSSYARSKGLTCVIDEDLRCSVSFGCSM 360
CYSDV  ALLDNMVVRNITKDSAVFMVGGSSLLKKVQHDVMNYCARTKLECIIDKDLRSAVSFGCSM 358
SPCSV  KILKAIADRNKITSGALFLVGGSSLLRKVQLDVSNFAKSIGLTPIIDKDLRSAVSYGCSM 359
ToCV   SILRSLVSRNKTSNGALFLVGGSSLLRPIQNRADGFARNHGLALIIDPDLRAAVSFGCSM 359
LIYV   KLINDVVVRNKLTSGVIYMVGGSSLLQPVQDMVRSYASTKGLTLVADQDMRSAVSYGCSV 360
              *          *******   *               * * *  ** ***

BPYV   QHAQEDSGSMIYIDCNSHPLMDLTLFGNPKVVVRKPMSIPYSTSDSRTIRSHYSTAVNVY 419
CuYV   QHAQEDSGSMIYIDCNSHPLMDLTLFGNPKVVVRKPMSIPYSTSDSRTIRSHYSTAVNVY 419
SPaV   QHAQEDSGSMTYIDCNSHPLMDLLMYGNPKVVVRKPMPIPYTKYDTRTIRQHYNTVVNVY 420
CYSDV  SHAQEDTKNMIYIDCNSHPLMDISYFCSPKIIVRKPMAIPYTGVREETLTRHYTTILNVY 418
SPCSV  MHAQEDSGSMVYIDCNSHPLMDVSLFANPRVIIRKPMSVPFSYKTTRKVDRHMMTAVNVY 419
ToCV   LHAQEDSGNMTYIDCNSHPLMDLGLYCHPRIIIRKPMSVPYTHKIEREVTRFITTALNVY 419
LIYV   LHKLEDNKEIVYIDCNSHPLSDISFNCDPEPIIRKPMSIPYTHTVKMRHDRPLKTIVNIY 420
        *  **     ********* *      *    ****  *              *  * *

BPYV   EGSDIFTLNNDWLISANVRTSDHVNVGEDLVFIYKYNIDGILELYAKNGRTGAEKLLPNS 479
CuYV   EGSDIFTLNNDWLISANVRTSDHVNVGEDLVFIYKYNIDGILELYAKNGRTGVEKLLPNS 479
SPaV   EGSDLFVLNNDWLVSAKVNTSDHANVGEDLTFVYKYTIDGILELYAKNEKTGVEKLLPNT 480
CYSDV  EGSDPFVLNNDWLISANMQSNKYGEIGDTLQYLYKYNVDGILELVVRNKRTGKETVLPNS 478
SPCSV  EGSSLFVLDNDWLVSANVNTQDFVDLGQELSYVYKYNVDGILDLFVRNESTGVESLLPNS 479
ToCV   EGSDLFVLNNDWLISADVDYSKYAKMGETLVSVYKYTIDGILELSMANKTTGKSWVLPNT 479
LIYV   EGSNLFMPENDWLISSNINTTDFAKVGEEYSKVYEYDIDGIITLKIRNEVTGKMFTLPNS 480
       ***  *   **** *           *      * *  ***  *   *  **    ***

BPYV   FSLTEKIEKLNLKLTQLSNVDEIATILSILSAFDDSLTGLLKFVKTPSILEREVAKISTP 539
CuYV   FSLTEKIEKLNLKLTQLSNVDEIATILSILSAFDDSLTGLLKFVKTPSILEREVAKISTP 539
SPaV   FSLTEKINKLDLQLTQLSTIDESATLISIMSYFDDNYTRLLSLLRTPTILERELLKITST 540
CYSDV  FALTESIKKLDVNLTQLSNIDELATLVAIMSYYKPELKYLLTYVKTPTIFENEIKKFGSG 538
SPCSV  FALTEKVKKLDLNLTQLSSVDELATVVSILSHFDSSFSSLLKLVNTPSIFETSAAKFGDL 539
ToCV   FARSEKIVISDLTLTQLSNVDELATIVSILSYFDTTFNYLISMFNTPSIFEREVGKISDA 539
LIYV   FTKSDNIKPITFKLTQLSNTDDLATLTSLLGYHDKNFERFYGLFNVPTILIKEIDKLGGF 540
       *            *****  *  **                       * *      *

BPYV   KELLERLVKQNKNFS- 554
CuYV   KELLERLVKQNKNFS- 554
SPaV   KKLYSALCDVNKNFNN 556
CYSDV  EDLYKSLAALNKNFK- 553
SPCSV  RSLYERLIFVNKNFS- 554
ToCV   KGLYNRLVEQNRNFS- 554
LIYV   KTLYRRLKSMNANF-- 554
         *   *   * **
```

Fig. 3. Continued

relationship with other viral proteins in the GenBank database when doing a BLASTp search and has no significant relationship to any protein in the database. The transmembrane domain spans from amino acids 29–46. This is the first of two such peptides encoded by BPYV-strawberry.

RNA 2 of BPYV-strawberry is 7904 nucleotides in length and has 8 ORFs (Fig. 2) in contrast to most criniviruses sequenced to date which encode 7 ORFs. All viruses of the family encode an array of five genes which are thought to be involved in cell-to-cell and long distance movement, which are found in RNA 2 of criniviruses [22,23]. The first ORF of BPYV-strawberry RNA 2 and one of the genes involved in closterovirus movement is a small hydrophobic protein similar in position and motifs to small hydrophobic proteins that all closteroviruses encode [2]. The starting codon is found in position 250 and termination signal at nucleotide 480. This peptide has molecular weight of 9 kDa. The transmembrane helix starts at amino acid 37 and terminates at amino acid 57. This larger than CuYV ORF 1 of RNA 2 is due to an additional cytosine found at position 285. The BPYV-strawberry peptide has a very strong transmembrane motif in contrast to that of CuYV which does not possess such motifs according to the software we utilized for the analysis. The next ORF which starts at nucleotide 697 encodes the 554-amino acid HSP70h gene of the virus (Fig. 3). The HSP70h gene has been shown to be associated with virions and plasmodesmata [24–26] for BYV, and is the second gene involved in viral movement. Downstream of the HSP70h there is a 53-amino acid encoding ORF. This ORF is homologous to ORFs in CYSDV (GenBank Accession No NP 851573) and SPaV (Tzanetakis and Martin, unpublished data) with 49% amino acid similarities in each case. This gene has homology to the C-terminus of *Little cherry virus-1* HSP70h with 28% similarities. The C-terminus of the heat shock protein superfamily is involved in substrate recognition [27] and one can only speculate on the function of this ORF. The fourth ORF starts at position 2536 and terminates at position 4089 encoding a gene of 59 kDa. This gene is also needed for virus movement and has recently been found associated with virions of BYV [21], being essential for assembly of the virions. RNA2 ORF 5 encodes a protein of 10 kDa. This is a gene that is found only in criniviruses and its' function is still unknown. The next two genes of RNA 2 of BPYV-strawberry encode the two CP of the virus. The major CP encapsidates ~95% of the genome while the CPm protects the remaining 5% [28], tentatively the 5′ end of the genome [29]. The major CP gene starts at nucleotide 4309 and encodes a protein of 252 amino acids with molecular weight of 28 kDa. The CPm of the virus begins at nucleotide 5057, 10 nucleotides before the termination signal for the CP, and is the virus largest gene after the 1a ORF, encoding a protein of 657 amino acids and 75 kDa. The last ORF of BPYV-strawberry RNA 2 encodes a protein of 230 amino acids and molecular weight of 26 kDa starting at 7027 and terminating at nucleotide 7719.

The diversity between BPYV strains is striking with the insertion of 49 amino acids in ORF la and the additional ORF in RNA1 found in the strawberry strain reported here compared to CuYV. We hope to further investigate the diversity of BPYV, by sequencing additional strains of the virus and identifying potential differences among them.

Acknowledgments

We would like to thank Karen Keller for her excellent technical assistance. We could also like to thank William Wintermantel and Arturo Cortez (USDA-ARS, Salinas, CA) for transfering the virus isolate from strawberry to *N. benthamiana* and infected tissue of ToCV for dsRNA extraction. This project was funded by the North American Strawberry Growers' Association and USDA-ARS.

References

1. Martelli G.P., Agranovsky A.A., Bar-Joseph M., Boscia D., Candresse T., Coutts R.H., Dolja V.V., Falk B.W., Gonsalves D., Jelkmann W., Karasev A.V., Minafra A., Namba S., Vetten H.J., Wisler G.C., and Yoshikawa N., Arch Virol *147*, 2039–2044, 2002.
2. Karasev A.V., Ann Rev Phytopathol *38*, 293–324, 2000.
3. Dolja V.V., Karasev A.V., and Koonin E.V., Ann Rev Phytopathol *32*, 293–324, 1994.
4. Frazier, N.W. and Stubbs L.L., Plant Dis Rept *53*, 524–526, 1969.
5. Tzanetakis I.E., Halgren A.B., Keller K.E., and Martin R.R., Phytopathology *92*, S82, 2002.
6. Tzanetakis I.E., Wintermantel W.M., and Martin R.R., Plant Dis *87*, 1398, 2003.
7. Wisler G.C., Duffus J.E., Liu H.-Y., and Li R.H., Plant Dis *82*, 270–280, 1998.
8. Duffus J.E., Phytopathology *55*, 450–453, 1965.
9. Coutts R.H.A. and Coffin R.S., Virus genes *13*, 179–181, 1996.

10. Hartono S., Tomohide N., Genda Y., and Okuda S., J Gen Virol *84*, 1007–1012, 2003.
11. Yoshikawa N. and Converse R.H., Phytopathology *80*, 543–548, 1990.
12. Jelkmann W., Martin R.R., and Maiss E., Phytopathology *79*, 1250–1253, 1989.
13. Hughes D.W. and Galau G., Plant Mol Biol Reptr *6*, 253–257, 1988.
14. Sippel A.E., Eur J Biochem *37*, 31–40, 1973.
15. Karasev A.V., Boyko V.P., Gowda S., Nikolaeva O.V., Hilf M.E., Koonin E.V., Niblett C.L., Cline K., Gumpf D.J., Lee R.F. et al., Virology *208*, 511–520, 1995.
16. Kreuze J.F., Savenkov E.I., and Valkonen J.P., J Virol *76*, 9260–9270, 2002.
17. Livieratos I.C. and Coutts R.H., Virus Genes *24*, 225–230, 2002.
18. Klaassen V.A., Boeshore M.L., Koonin E.V., Tian T., and Falk B.W., Virology *208*, 99–110, 1995.
19. Aguilar J.M., Franco M., Marco C.F., Berdiales B., Rodriguez-Cerezo E., Truniger V., and Aranda M.A., J Gen Virol *84*, 2555–2564, 2003.
20. Yi M. and Lemon S.M., RNA *9*, 331–345, 2003.
21. Erokhina T.N., Vitushkina M.V., Zinovkin R.A., Lesemann D.E., Jelkmann W., Koonin E.V., and Agranovsky A.A., J Gen Virol *82*, 1983–1994, 2001.
22. Napuli A.J., Alzhanova D.V., Doneanu C.E., Barofsky D.F., Koonin E.V., and Dolja V.V., J Virol *77*, 2377–2384, 2003.
23. Dolja V.V., Molec Plant Pathol *4*, 91–98, 2003.
24. Peremyslov V.V., Hagiwara Y., and Dolja V.V., Proc Natl Acad Sci USA *96*, 14771–14776, 1999.
25. Medina V., Peremyslov V.V., Hagiwara Y., and Dolja V.V., Virology *260*, 173–181, 1999.
26. Napuli A.J., Falk B.W., and Dolja V.V., Virology *274*, 232–239, 2000.
27. Bukau B. and Horwich A.L., Cell *92*, 351–366, 1998.
28. Agranovsky A.A., Lesemann D.E., Maiss E., Hull R., and Atabekov J.G., Proc Natl Acad Sci USA *92*, 2470–2473, 1995.
29. Zinovkin R.A., Jelkmann W., and Agranovsky A.A., J Gen Virol *80*, 269–272, 1999.