Phylogenetic analysis of *Heliothis armigera* cytoplasmic polyhedrosis virus type 14 and a series of dwarf segments found in the genome

Yanqiu Li, Jiamin Zhang, Yang Li, Li Tan, Wuguo Chen, Haishan Luo and Yuanyang Hu

Laboratory of Insect Virology, State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, Hubei, China

Full-length nucleotide sequences for the genome segments (S1–S6) of *Heliothis armigera* cytoplasmic polyhedrosis virus type 14 (HaCPV-14) have been characterized. Each segment consists of a single open reading frame with conserved motifs AGAA and AGCU at the 5' and 3' ends, respectively. Comparison of the proteins of HaCPV-14 with those of other members of the family Reoviridae suggests that S1 encodes an RNA-dependent RNA polymerase (RdRp), whilst S2 encodes a major capsid protein of the virus. Phylogenetic analysis of RdRps from 16 viruses in the family Reoviridae reveals that the genera *Cypovirus* and *Oryzavirus* may have originated from a common insect virus ancestor. A series of viable dwarf segments originating from S5 of HaCPV-14 has been identified. Analysis of the predicted secondary structures for these dwarf segments suggests that the signals essential for replication and packaging are located within the terminal sequences of these segments.

**INTRODUCTION**

Cytoplasmic polyhedrosis viruses (CPVs) belong to the genus *Cypovirus* of the family *Reoviridae* (Mertens et al., 2004). Viruses belonging to the family *Reoviridae* are characterized by the presence of capsids made up of concentric layers of proteins organized in one, two or three shells containing 10–12 segments of linear, double-stranded RNA (dsRNA); a genome with nine segments of dsRNA also has been reported recently (Tyler & Fields, 1996; Hill et al., 1999; Attouï et al., 2005). CPVs are the only members of this family possessing single shells in their capsids. The segmented dsRNA genomes are never detected free in the cytoplasm and are transcribed and replicated within viral capsids by an RNA-dependent RNA polymerase (RdRp). Structural studies have indicated that each segment may exist as a tightly wound spiral around one of the 12 RdRp-capping complexes positioned at the fivefold axes of the core (Gouet et al., 1999).

Based on differences in the electrophoretic-migration patterns of genomic dsRNA segments, 16 CPV types have been identified (Payne & Rivers, 1976; Mertens et al., 1999, 2004; Belloncik, 1996). Recently, cypoviruses 18 and 19 have been reported from the winter moth, *Operophthera brumata* (Graham et al., 2006). The nucleotide sequences of the polyhedrin gene from type 1 CPVs, *Bombyx mori* CPV (BmCPV) and type 5 CPVs, including *Heliothis armigera* CPV (HaCPV), *Euxoa scandens* CPV and *Orygia pseudotsugata* CPV, have been determined. BmCPV-1, one of the most important members of the genus *Cypovirus*, has six structural proteins, VP1, VP2, VP3, VP4, VP6 and VP7, encoded by segments S1, S2, S3, S4, S6 and S7, respectively (Hagiwara & Matsumoto, 2000; Hagiwara et al., 2002). S1 encodes the major capsid protein, which assembles into single-shelled virus-like particles (Hagiwara & Naitow, 2003), whilst S2 encodes a putative RdRp (Rao et al., 2003). The products of segment S5 are likely to be autocleaved post-translationally into two non-structural proteins, one of the products being a 2A pro-like protease (Hagiwara et al., 2001). Sequences of segments S8 and S9, encoding these two non-structural proteins, have also been determined (Hagiwara et al., 1998a, b). All of the segments of BmCPV have five conserved nucleotides at the 5’ end and seven conserved nucleotides at the 3’ end (5’-AGUAA......GUUAGCC-3’).

CPVs are common insect pathogens detected routinely in insect colonies. These viruses have been isolated from more than 250 different insect species, both reared in laboratories and collected from the field. Observation of the electrophoretic-migration pattern of more than 10 dsRNA...
segments from extracts of infected insects reveals that infection with more than one CPV strain is common (Belloncik, 1996; Belloncik & Mori, 1998). Although deviations from equimolarity would support the conclusion of a multiple CPV infection, only one definite type of CPV has been isolated successfully from a mixture. Belloncik (1996) and colleagues reported separation of one type of CPV from a mixture of HaCPV types (Chinese strain) by using cell culture and have suggested that the HaCPV isolate belongs to type 14. Some CPVs show almost-identical dsRNA electrophoretic-migration patterns after several passages (Belloncik, 1989). However, the appearance of a viable dwarf gene (deletion-mutant gene) is noted occasionally for some CPVs after multiple passages in insects (Arella et al., 1988).

In this paper, we report the sequences of genome segments S1–S6 of HaCPV-14 and characterize their predicted functions by using in silico analysis. Combined with our group’s recent report on sequencing of genome segments S7–S10 (Li et al., 2006), the present results complete the full genome sequence of HaCPV-14. Interestingly, a series of dwarf segments of about 750 nt mixed in the genome of HaCPV was discovered by electrophoresis (Fig. 1). Direct sequencing results coupled with comparative studies indicate clearly that the dwarf segments are viable deletion mutants originating from segment S5 of the HaCPV-14 genome.

METHODS

Source of virus and host. BmCPV was obtained from the China Center for General Viruses Culture Collection (CCGVCC). HaCPV (Chinese strain) was a gift from Dr Jiang Zhong (Fudan University, Shanghai, China).

Fig. 1. Comparison of the electrophoretic-migration patterns of CPV dsRNA on 1% agarose gel. Lanes: 1, HasCPV-14 (Heliothis assulta CPV-14); 2, HaCPV mixture [a faint line near 0.75 kb was identified as dwarf segments (arrow)]; M, molecular mass marker.

Purification of polyhedra. Polyhedra from infected larvae were purified by sucrose density-gradient centrifugation according to Ikeda et al. (2001).

Isolation of dsRNA and construction of a cDNA library. Genomic dsRNA was extracted from the purified polyhedra with phenol/chloroform, precipitated by using ethanol and finally separated by electrophoresis on a 1% agarose gel (Hagiwara et al., 1998b). Purified dsRNA was denatured at 100°C for 5 min prior to use. Synthesis of cDNA was carried out by using a Moloney murine leukemia virus cDNA synthesis kit from Takara (Japan). Total dsRNA was used as a template with random hexamers as primers. The reactions were set up following the manufacturer’s instructions for synthesis of cDNA. The resulting cDNA was cloned into the pMD18-T vector (Takara).

Determination of the sequence of segments S1–S6. Analysis and comparison of the partial HaCPV (Chinese strain) genome sequence revealed high identity to known segments of LdCPV-14. Forward and reverse primers designed on the basis of the terminal RNA sequences of each of the six segments of LdCPV-14 (segments 1–6) were used for the reverse transcription and amplification of each corresponding segment of HaCPV. Purified dsRNA was denatured at 100°C for 5 min and chilled rapidly on ice prior to RT-PCR amplification. The amplified cDNA was cloned into the pMD18-T vector (Takara).

Determination of 5’- and 3’-terminal sequences. The terminal sequence of each segment of the dsRNA was fixed by tailing with an anchor (5’-PO4-AGGTCTCGTAGACCGTGCACC-PO4-3’) using 50 U T4 RNA ligase (Takara). The tailed dsRNA was denatured by heating at 99°C for 1 min and cDNA synthesis was initiated by using two primers. The sequence of one of the primers was 5’CGTGAAGGTGCACGCTACGAGACCT-3’, which was partially complementary to the anchor sequence, whilst the second primer was designed based on the sequences of HaCPV-14 segments 1–6 determined in this study (see Supplementary Table S1, available in JGV Online). The amplicons were analysed by agarose-gel electrophoresis and ligated into the pMD18-T vector.

Cloning of the dwarf segments. The dwarf segments of HaCPV were cloned and sequenced by using the single-primer amplification technique as described previously (Lambden et al., 1992; Attoui et al., 2000a, b; Hagiwara et al., 2002). Briefly, the dsRNA segments were tailed with primer A (5’-PO4-AGGTTCTCGTAGACCGTGCACC-PO4-3’) using 50 U T4 RNA ligase (Takara). The tailed dsRNA was denatured by heating at 99°C for 1 min and cDNA synthesis was initiated by using two primers. The sequence of one of the primers was 5’CGTGAAGGTGCACGCTACGAGACCT-3’, which was partially complementary to the anchor sequence, whilst the second primer was designed based on the sequences of HaCPV-14 segments 1–6 determined in this study (see Supplementary Table S1, available in JGV Online). The amplicons were analysed by agarose-gel electrophoresis and ligated into the pMD18-T vector.

Sequencing and computer analysis of sequence data. The cDNA sequences cloned into the pMD18-T vector were sequenced by using M13 universal primers for amplification and an ABI 3730 automated DNA sequencer (Perkin Elmer). Sequence data were assembled and analysed by using EDITSEQ and SEQUEN (Lasergene). Protein analysis was performed by using the programs BIODIT and PROTEAN (Lasergene). Sequences were compared with those available from nucleic acid and protein databases by using BLAST (http://www.ncbi.nlm.nih.gov/blast/). Multiple sequence alignment was performed by using CLUSTAL_X version 1.83 (Thompson et al., 1997). Tree drawing was performed with the help of the programs PHYLP version 3.63 and TREEVIEW (Page, 1996). The RNAstructure 4.2 program was used to predict the secondary structure of dwarf-gene mRNA.
RESULTS

Determination of partial sequences

As HaCPVs isolated from natural sources were usually found mixed with other CPVs, total dsRNA was used as a template for cDNA synthesis with random hexanucleotides by RT-PCR. The cDNA fragments obtained were cloned into vector pMD18-T. A number of clones positive for segments of HaCPV were identified and the nucleotide sequences of some large segments of HaCPV cloned in the pMD18-T vector were determined. BLAST analysis of these sequences against GenBank revealed that the sequences showed high similarity to the genome of Lymantria dispar CPV 14 (LdCPV-14).

Determination of sequences of HaCPV-14 segments S1–S6

Nucleotide sequences of genome segments S1–S6 of HaCPV-14 were determined (GenBank accession numbers DQ242048, DQ388474, DQ388475, DQ388476, DQ388477 and DQ017080, respectively). Sizes of the RNA segments, putative proteins and 5’ and 3’ non-coding regions (NCRs) were identified for each segment of the genome and are listed in Table 1. Analysis of the sequencing data revealed that each of the six segments carried conserved sequences located at their terminals. The extreme 5’ and 3’ ends of the sense strand had the sequence 5’-AGAA...AGCU-3’. Comparison of the nucleotide sequences of S1–S6 of HaCPV-14 with those of other CPVs showed that there were high degrees of identity amongst the same virus type, but little nucleotide similarity amongst CPVs of different types. According to the results obtained from the BLAST search, putative proteins encoded by HaCPV-14 segments S1–S6 showed a significant match with those encoded by segments S1–S6 of LdCPV-14 (amino acid identities of 97–99%). HaCPV-14 S1 was the largest genome segment, containing 4329 nt, and the protein encoded by HaCPV-14 S1 contained the characteristic signature motifs for the RdRp of members of the Reoviridae: the conserved motif GKKQxGxxxD was found at position 530, the conserved motif SxDxxxMD was found at position 555, the conserved motif SG was found at position 649 and the conserved motif GDD was found at position 689. These findings are consistent with the hypothesis that the first genome segment of HaCPV-14 encodes the viral RdRp. The deduced amino acid sequence of the RdRp shows 43% identity and 60% similarity to that of BmCPV-1.

Sequences of the RdRp genes from representatives of 11 genera of the family Reoviridae, including BmCPV-1, HaCPV-14 (Cypovirus), Rice ragged stunt virus (RRSV; Oryzavirus), Rice black streaked dwarf virus (RBSDV), Fiji disease virus (FDV; Fijivirus), Eyach virus (EYAV; Coltivirus), grass carp reovirus (GCRV), Aquareovirus A (ARV-A; Aquareovirus), Bluettongue virus (BTV), Epizootic hemorrhagic disease virus (EHDV; Orbivirus), simian rotavirus A/SA11 (SRV-A), porcine rotavirus Cowden (PoRV-C; Rotavirus), Rice dwarf virus (RDV; Phytoreovirus), mammalian orthoreovirus 2-D5/Iones (MRV-2; Orthoreovirus), Banna virus (BAV; Seadornavirus) and Mycoreovirus 3 (MYRV; Mycoreovirus), were retrieved from GenBank and the evolutionary relationships between these viruses were evaluated by constructing a neighbour-joining tree for the RdRp genes (Fig. 2). Although the putative protein encoded by the gene on segment S2 of HaCPV-14 showed 99% identity to that of LdCPV-14, the open reading frame (ORF) of the gene on segment S2 of HaCPV was 74 aa longer than that of LdCPV-14, with the start codon beginning at nt 39, whereas in the case of LdCPV-14, translation initiates at nt 261. Comparison of the protein encoded by HaCPV-14 S2 with homologous proteins suggested that HaCPV-14 S2 encodes viral major capsid protein VP1. Homology searches using the BLAST program revealed that the HaCPV-14 S3 sequence was 3 nt

<table>
<thead>
<tr>
<th>Segment</th>
<th>GenBank accession no.</th>
<th>Length (bp)</th>
<th>Putative encoded protein</th>
<th>5’ NCR</th>
<th>3’ NCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>DQ242048</td>
<td>4329</td>
<td>1228</td>
<td>588 AGAACUUU</td>
<td>54 UUCAGCU</td>
</tr>
<tr>
<td>S2</td>
<td>DQ388474</td>
<td>4056</td>
<td>1303</td>
<td>38 AGAAUCAU</td>
<td>106 UUCAGCU</td>
</tr>
<tr>
<td>S3</td>
<td>DQ388475</td>
<td>3918</td>
<td>1266</td>
<td>42 AGAAUGAU</td>
<td>75 UUCAGCU</td>
</tr>
<tr>
<td>S4</td>
<td>DQ388476</td>
<td>3339</td>
<td>1075</td>
<td>15 AGAAUUUC</td>
<td>96 UUCAGCU</td>
</tr>
<tr>
<td>S5</td>
<td>DQ388477</td>
<td>3109</td>
<td>974</td>
<td>68 AGAAUUUC</td>
<td>116 UUCAGCU</td>
</tr>
<tr>
<td>S6</td>
<td>DQ017080</td>
<td>1783</td>
<td>550</td>
<td>30 AGAAUCUU</td>
<td>100 UUCAGCU</td>
</tr>
<tr>
<td>S7</td>
<td>DQ010326</td>
<td>1391</td>
<td>424</td>
<td>17 AGAAUUUC</td>
<td>99 UUCAGCU</td>
</tr>
<tr>
<td>S8</td>
<td>DQ010325</td>
<td>1250</td>
<td>379</td>
<td>35 AGAAUUUC</td>
<td>75 UUCAGCU</td>
</tr>
<tr>
<td>S9</td>
<td>DQ010324</td>
<td>1143</td>
<td>318</td>
<td>77 AGAAUUUC</td>
<td>109 UUCAGCU</td>
</tr>
<tr>
<td>S10</td>
<td>DQ010323</td>
<td>956</td>
<td>249</td>
<td>52 AGAAUUUU</td>
<td>154 UUCAGCU</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consensus</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

*Calculated theoretical molecular mass.

Table 1. Properties of dsRNA segments S1–S10, putative encoded proteins and 5’ and 3’ NCRs of HaCPV-14
shorter than LdCPV-14 S3, with 1 nt missing at each of positions 2332, 2349 and 2361. The resulting amino acid sequence from aa 764 to 774 encoded by HaCPV-14 S3 was therefore different from that of LdCPV-14. Deduced amino acid sequences of HaCPV-14 S3, S4, S5 and S6 showed low similarity to those of type 1 CPVs, showing 24, 31, 21 and 23 % identity to proteins encoded by genes on S3, S4, S5 and S6 of *Dendrolimus punctatus* CPV 1, respectively. Analysis of the sequences using the BioEdit motif-search program as described by Nibert & Kim (2004) revealed that the putative protein P63 encoded by HaCPV-14 S6 had two conserved sequence motifs for NTP binding: KgXgKxs and dSDxyG (upper-case letters indicate wholly conserved residues and lower-case letters indicate partially conserved residues), covering residues 346–352 and 374–379, respectively.

**Determination of sequences of a series of dwarf segments**

Sequences of a series of dwarf segments were determined and deposited in GenBank with accession numbers DQ677368–DQ677371. The four segments had conserved terminal sequences consisting of four conserved nucleotides at the 5' and 3' ends (5'-AGAA....AGCU-3'). Multiple sequence alignment using BLAST revealed that these segments were related closely to S5 of the HaCPV-14 genome. All four segments were approximately the same size and were named HaCPV14-737, HaCPV14-749, HaCPV14-760 and HaCPV14-773, depending on the length of the nucleotide sequence for a particular segment. The ORFs of all four segments start at the same position as that of HaCPV-14 S5, nt 68, and the stop codons are located at nt 565, 643, 502 and 529, respectively. HaCPV14-737 seems to have originated from the S5 gene by deletion of the internal 2372 nt linked by residues CGATTG, spanning nt 516–521. GATT can be found at two places: nt 517–520 at the 5' end and nt 2889–2892 at the 3' end. This gives rise to five possibilities for the deletion event, as shown in Fig. 3(a). HaCPV14-749 seems to have been formed by deletion of 2360 nt between the two Ts found at positions 486 and 2847 (Fig. 3b). HaCPV14-760 has possibly originated by the deletion of 2349 nt between two As positioned at nt 604 and 2954, between nt 605 and 2955, or between nt 606 and 2956, as shown in Fig. 3(c). All of the three possibilities result in the same dwarf segments for HaCPV14-760. Finally, HaCPV14-773 seems to have come into existence by the deletion of 2336 nt spanning a G at position 534 and a C at

---

**Fig. 2.** Phylogenetic analysis of the RdRp domains. Trees were constructed from the sequence alignment results of 16 RdRps. Numbers at nodes represent bootstrap values as percentages estimated from 100 replicates. Branch lengths are proportional to relatedness. GenBank accession numbers are shown on the tree. Bar, 0.1 nucleotide substitutions per site.

**Fig. 3.** Schematic representation of the generation of the four segments derived from S5 of HaCPV-14. (a) There were five possibilities for a junction to form HaCPV14-737, i.e. nt 516 and 2889, nt 517 and 2890, nt 518 and 2891, nt 519 and 2892, or nt 520 and 2893. (b) HaCPV14-749 was deleted from S5 by linking at two Ts at nt 486 and 2847. (c) HaCPV14-760 has possibly originated by the deletion of 2349 nt between two As positioned at nt 604 and 2954, between nt 605 and 2955, or between nt 606 and 2956, as shown in Fig. 3(c). All of the three possibilities result in the same dwarf segments for HaCPV14-760. Finally, HaCPV14-773 seems to have come into existence by the deletion of 2336 nt spanning a G at position 534 and a C at
nucleotide sequence of S2 was shorter than that of S1, the encoded the viral major capsid protein VP1. Although the S2 with homologous proteins indicated that HaCPV-14 S2 sequences showed that each segment contained a single sequence of HaCPV-14. Computer analysis for these sequences were found in all segments of the HaCPV-14 genome. Analysis of the results of the BLAST search for complementary sequences that generate a panhandle, (ii) a stem–loop and (iii) a hairpin (Patton & Spencer, 2000) – were seen for the modelled secondary structures of reoviruses – (i) 5'-9 segments indicated that the 5'-9 ends of the mRNAs could interact in cis via base pairing of complementary terminal sequences to form panhandle structures that might promote the synthesis of dsRNA (Fig. 5). The three major signals essential for packaging and replication of the virus were deletion mutants derived from S5 of HaCPV-14. The dwarf segments were extracted from purified virus particles. This indicates that the segments can undergo packaging and replication, despite deletion of part of the ORF. Thus, the signals essential for packaging and replication of the virus are probably located within the terminal sequences of the dwarf segments.

Modelling of the secondary structures of the deletion segments indicated that the 5' and 3' ends of the mRNAs could interact in cis via base pairing of complementary terminal sequences to form panhandle structures that might promote the synthesis of dsRNA (Fig. 5). The three major elements commonly observed for folded viral mRNA structures of reoviruses – (i) 5'–3' complementary sequences that generate a panhandle, (ii) a stem–loop and (iii) a non-base-paired 3' tail that forms the 3' essential replication signal (Patton & Spencer, 2000) – were seen for the modelled position 2871, as shown in Fig. 3(d). Whilst the ORF for HaCPV-14-760 remained the same as that of the original coding region of the S5 gene, the other three dwarf genes showed a shift in the reading frame and encoded different amino acids at the N terminus, as shown in Fig. 4. The secondary structures of the dwarf gene mRNAs (Fig. 5) were predicted by the RNAstructure 4.2 program.

**DISCUSSION**

Members of the family Reoviridae are grouped into 12 genera: Aquareovirus, Cypovirus, Idenovirus, Orthoreovirus, Rotavirus, Fijivirus, Mycoreovirus, Oryzavirus, Coltivirus, Orbivirus, Seadornavirus and Phytoenceirus (Mertens et al., 2004). The natural hosts of these viruses include vertebrates, invertebrates and plants. CPVs belong to the genus Cypovirus and the complete genome sequences of BmCPV, LdCPV-1, LdCPV-14 and Trichoplusia ni cypovirus 15 have been reported. The partial genome sequence of HaCPV-14, spanning segments S7–S10, was known previously. In the present study, the sequences of HaCPV-14 S1–S6 were determined to complete the genome sequence of HaCPV-14. Computer analysis for these sequences showed that each segment contained a single ORF, and a high identity (83–98%) was revealed among viruses classified within the same electropherotype; however, no nucleotide sequence identity was found amongst different types of CPV, despite some similarities in their amino acid sequences. Analysis of the protein sequences of the highly conserved RdRps shows amino acid sequence identity >20% within a single genus of the family Reoviridae. The amino acid sequence identity of the RdRp of HaCPV-14 to those of other cypoviruses falls in a wide range (42–99%). Hence, the RdRp gene was a potential candidate for comparison studies of different genera in the family, and this would serve to elucidate the evolutionary relationships between these viruses.

Comparison of the putative protein encoded by HaCPV-14 S2 with homologous proteins indicated that HaCPV-14 S2 encoded the viral major capsid protein VP1. Although the nucleotide sequence of S2 was shorter than that of S1, the ORF of S2 happened to be the largest within all viral proteins. Homology searches indicated that P63 encoded by S6 had motifs for NTP-binding sites, suggesting that P63 might play an important role in the cap formation of the genomic RNA.

Phylogenetic analysis of RdRps sequences from 16 viruses belonging to 11 different genera of the family Reoviridae also confirmed that HaCPV-14 was related more closely to RRSV (genus Oryzavirus) than to other members of different genera in the family Reoviridae. This result confirms previous conclusions that reoviruses, which exploit insects as their hosts, might have originated from a common insect virus ancestor.

In this study, a series of viable deletion-mutant genes was identified in the genome of HaCPV-14, including HaCPV14-737, HaCPV14-749, HaCPV14-760 and HaCPV14-773. Identical amino acids among the four proteins are indicated by asterisks; chemically similar amino acids are indicated by two dots.
mRNAs of all four deletion segments. In rotaviruses, the 3' consensus sequence included the 3' essential replication signal, a motif that was required for dsRNA synthesis. As the stem–loop differed for each of the mRNAs, the conserved motif at the 3' end was implicated in the packaging and assortment of the viral mRNAs. These highly conserved terminal motifs probably contain unique packaging signals that allow different segments to be distinguished during replication.

The generation of a terminally conserved mutant dsRNA gene by internal deletion events has been observed previously for the polyhedron gene of BmCPV, Wound tumor virus and rotavirus genomic RNA (Nuss & Summers, 1984; Arella et al., 1988; Taniguchi et al., 1996). A series of deletion-mutant genes in a CPV was identified for the first time in this study. The mechanism responsible for the generation of such deletion-mutant dsRNAs is at present unknown. Not only are the residues at the site of deletion of the four dwarf segments different, but also the exact locus of deletion cannot be identified. There is no evidence for the involvement of splicing events in the formation of these dwarf genes. Therefore, the plausible explanation for the cause of the deletions might be the result of the copy-choice mechanism, involving intermolecular ‘jumping’ of the viral replicase.

As the dwarf genes of HaCPV-14 could replicate and package themselves into virus particles, it would be interesting to determine the minimal sequence requirements for survival and propagation of the virus. This would throw some light on the underlying molecular mechanism of CPV genome replication and packaging. The deletions were not observed in all segments of the genome. It was unknown whether the dwarf segments could be mixed with a normal genome in the same particle or form a completely

Fig. 5. Predicted secondary structures of four dwarf segments: (a) HaCPV14-737; (b) HaCPV14-749; (c) HaCPV14-760; and (d) HaCPV14-773. Only the portion of the computed structure made up of the conserved 5'- and 3'-terminal sequences is shown as: (1) 5'-3' complementary sequences that generate a panhandle; (2) a stem–loop; and (3) a non-base-paired 3' tail.
different particle. Further studies at the molecular level are needed to investigate these intriguing questions.

REFERENCES


