Identification and genome characterization of Heliothis armigera cypovirus types 5 and 14 and Heliothis assulta cypovirus type 14

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

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

Genomic characterization of Heliothis armigera cypovirus (HaCPV) isolated from China showed that insects were co-infected with several cypoviruses (CPVs). One of the CPVs (HaCPV-5) could be separated from the others by changing the rearing conditions of the Heliothis armigera larvae. This finding was further confirmed by nucleotide sequencing analysis. Genomic sequences of segments S10–S7 from HaCPV-14, S10 and S7 from HaCPV-5, and S10 from Heliothis assulta CPV-14 were compared. Results from database searches showed that the nucleotide sequences and deduced amino acid sequences of the newly identified CPVs had high levels of identity with those of reported CPVs of the same type, but not with CPVs of different types. Putative amino acid sequences of HaCPV-5 S7 were similar to that of the protein from Rice ragged stunt virus (genus Oryzavirus, family Reoviridae), suggesting that CPVs and oryzaviruses are related more closely than other genera of the family Reoviridae. Conserved motifs were also identified at the ends of each RNA segment of the same virus type: type 14, 5′-AGAAUUU...CAGCU-3′; and type 5, 5′-AGUU...UUGC-3′. Our results are consistent with classification of CPV types based on the electrophoretic patterns of CPV double-stranded RNA.

INTRODUCTION

Cytoplasmic polyhedrosis viruses (CPVs) belong to the genus Cypovirus in the family Reoviridae (Mertens et al., 2000). Viruses in the family Reoviridae have a capsid made up of concentric layers of proteins that are organized as one, two or three distinct capsid shells containing 10, 11 or 12 segments of linear double-stranded RNA (dsRNA); only CPVs have single-shelled particles. The main characteristic of CPV replication is the mass production of a major virus-coded protein (polyhedrin) that crystallizes in the cell cytoplasm (Payne & Mertens, 1983; Belloncik, 1989; Belloncik & Mori, 1998). Based on variation in the electrophoretic migration patterns of genomic dsRNA segments, 14 CPV types have been identified. Recently, a putative new electropherotype of Trichoplusia ni CPV 15 (TnCPV-15) has been identified with a genome consisting of 11 segments of dsRNA (Payne & Rivers, 1976; Mertens et al., 1989, 1999; Belloncik et al., 2003). These viruses have polyhedrin proteins with molecular masses of 27–31 kDa encoded by the smallest genome segment. The nucleotide sequences of the polyhedrin gene from type 1 CPVs, Bombyx mori CPV (BmCPV) and type 5 CPVs, including Heliothis armigera CPV (HaCPV), Euxou scandens CPV (EsCPV) and Orgyia pseudotsugata CPV (OpCPV), have been determined. No similarities have been found in the nucleotide sequence of the polyhedrin gene from distinct types; however, a slight sequence similarity has been found in three regions between the amino acid sequences of polyhedrin from EsCPV and BmCPV (Belloncik & Mori, 1998). 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). Sequences of segments S8 and S9, encoding the two non-structural proteins, have also been determined (Hagiwara et al., 1998a, b). S5 might be cleaved into two non-structural proteins, including a 2Apro-like protease (Hagiwara et al., 2001), by post-translational autocleavage. S2 encodes a putative RNA-dependent RNA polymerase (Rao et al., 2003) and S1 encodes the major capsid protein, which assembles into single-shelled virus-like particles (Hagiwara & Naitow, 2003). The three-dimensional structure of CPV particles has been determined by using cryoelectron microscopy. The current structure is reported at a resolution of 8 Å, which is up from 13 and 25 Å (Zhang et al., 1999; Nibert & Baker, 2003; Xia et al., 2003; Zhou et al., 2003). Another 25 Å cryoelectron microscopy reconstruction of a CPV is also reported (Hill et al., 1999). At a resolution of 8 Å, α-helices and β-sheets become distinguishable in the CPV structure.
This new information suggests that the virus particles function as stable machines for mRNA synthesis.

CPVs have been isolated from more than 250 insect species reared in a laboratory or collected from the field. These viruses are also very common pathogens found in insect colonies. Migration on gels of more than ten dsRNA segments from extracts of infected insects that have been collected from the field or reared in the laboratory demonstrate that infections with more than one CPV strain are common. Several experimental factors, such as insect strain used, larval instar, the route of infection, environmental conditions and timing of the virus infections, influence the nature of the interaction between the strains (Payne & Rivers, 1976; Belloncik, 1996; Belloncik & Morí, 1998). Although deviation from equimolarity would support the conclusion of a multiple CPV infection, only one successful separation of a definite type from a mixture of CPVs has been made (Belloncik et al., 1996). Belloncik and colleagues reported separation of one CPV from a mixture of HaCPV types (Chinese strain) using cell culture and they suggest that the HaCPV isolated belongs to type 14.

In this paper, another CPV, HaCPV-5, was separated from a mixture of HaCPVs (Chinese strain) and the complete nucleotide sequences of S7, S8, S9 and S10 of HaCPV-14, separated by Belloncik et al. (1996), and the complete sequences of S7 and S10 from HaCPV-5, separated in this study, are reported. Moreover, the complete nucleotide sequence of segment S10 of Heliothis assulta CPV 14 (HasCPV-14), isolated from Heliothis assulta larvae in a pure form, i.e. not associated with a virus mixture, was also determined.

**METHODS**

**Source of virus and host.** BmCPV was a gift from the China Center for General Viruses Culture Collection (CCGVCC). HaCPV (Chinese strain) and HasCPV were generously donated by Jiang Zhong (Fudan University, Shanghai, China) and Qilian Qing (Institute of Zoology, Chinese Academy of Sciences, Beijing, China). B. mori larvae and H. armigera larvae were fed in our laboratory.

**Purification of polyhedra.** Polyhedra from infected larvae were purified by sucrose density-gradient centrifugation according to Ikeda et al. (2001). Suspensions of 10^6 polyhedral inclusion bodies (PIB) ml^{-1} and 10^7 PIB ml^{-1} were used as the infectious materials.

**Insect rearing and infections.** H. armigera larvae were reared on an artificial diet. Eggs were surface-sterilized by immersion in 2 % formaldehyde for 15 min at room temperature, washed several times with tap water and finally rinsed with distilled water. Those eggs were allowed to air-dry on paper towels and placed in glass cups to formaldehyde for 15 min at room temperature, washed several times an artificial diet. Eggs were surface-sterilized by immersion in 2 % H.

Suspension (10^6 PIB ml^{-1}) on the artificial diet; the temperature was kept at 28°C. After 2 days, larvae were collected and homogenized. If necessary, this protocol would be repeated in order to get pure HaCPV-5. (iii) Newly moulted third-instar silkworm larvae were fed with fresh mulberry leaves that had been smeared with a high concentration of polyhedrin. One group had leaves smeared with HaCPV (10^7 PIB ml^{-1}). The other group had leaves smeared with BmCPV (10^7 PIB ml^{-1}).

**dsRNA isolation and cDNA library construction.** CPV-infected larvae were collected, washed with distilled water and homogenized. The filtrate of the homogenized water was centrifuged at 10,000 g for 10 min to sediment CPVs containing the polyhedra. After washing with PBS (pH 7.4), polyhedra were treated with 0.5 % SDS and shaking at room temperature, and then washed three times with PBS. Genomic dsRNA was extracted from the purified polyhedra with phenol/chloroform, precipitated with ethanol and separated by 1 % agarose-gel electrophoresis (Hagiwara et al., 1998b). Purified dsRNA was denatured at 100°C for 5 min. Total dsRNA was used as a template for cDNA synthesis with random hexamers using an M-MLV cDNA synthesis kit (Takara). The resulting cDNA was cloned into vector pMD18-T (Takara).

**Cloning of the dsRNA segment.** The genomic 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 sealed with primer A (5'-PO4-AGGTCTC-GTAGACCGTGACCC-PO4-3') using 50 U T4 RNA ligase (Takara). The tail dsRNA was purified by using a TRizol LS Reagent kit (Gibco), denatured by heating at 95°C for 1 min and the cDNA was processed by using primer B (5'-CGTGAGGTGCACGGTCTACGA-GACCT-3') and 15 U ThermoScript reverse transcriptase (Invitrogen). The cDNA was amplified by PCR using primer B and the amplicons were analysed by agarose-gel electrophoresis and ligated into the pMD18-T cloning vector.

**Sequencing and computer analysis of the sequenced data.** Insert sequences were determined with M13 universal primers and an ABI 377 automated DNA sequencer (Perkin Elmer). Sequence data were assembled and analysed with the programs EDITSEQ and SEQMAN (Lasergene, DNAstar). Protein analysis was performed by using the programs BIOEDIT 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 analysis was performed by using CLUSTAL_X version 1.83 (Thompson et al., 1997). Tree drawing was performed with the help of the program TREEVIEW (Page, 1996).

**RESULTS**

**Diagnosis of disease**

Diseases developed in H. armigera larvae when they were infected with HaCPV. Some larvae became stunted, did not feed and died, usually in the third or fourth instar. It was observed that infected larvae died quickly if the rearing temperature was about 28°C. Under the same conditions, silkworms did not generally succumb to infection with HaCPV and had no obvious symptoms, but the majority of silkworms were dead if the infectious material was BmCPV.
Electrophoretic patterns of CPV mixtures and pure CPV dsRNA

The agarose-gel electrophoretic patterns of HaCPV dsRNA extracted from infected *H. armigera* larvae under different rearing conditions are shown in Fig. 1. Under rearing conditions at 20°C, electrophoretic patterns of HaCPV revealed more than 10 bands (Fig. 1), indicating a mixture of viruses. If the rearing temperature was changed to 28°C, HaCPV electrophoretic patterns revealed 10 bands, suggesting the presence of a single CPV type. A comparison of the electrophoretic patterns showed that HaCPV belonged to type 5. In addition, analysis of electrophoretic patterns showed that HasCPV dsRNA belonged to type 14 (Fig. 1). When the electrophoretic patterns of types 5 and 14 were compared to the mixture of HaCPVs, the same electrophoretic mobility was observed.

**Determination of partial sequences**

As HaCPVs isolated from natural sources were mixed with other CPVs, total dsRNA was used as template for cDNA synthesis with random hexanucleotides by RT-PCR. The cDNA fragments were cloned into vector pMD18-T. After screening of the recombinant clones, several inserted segments of HaCPV were identified and the nucleotide sequences of some larger inserted segments of HaCPV were determined. BLAST comparison of these sequences with GenBank revealed that these sequences had high similarity to the genome of *Lymantria dispar* CPV 14 (LdCPV-14) and the polyhedrin gene of CPV type 5.

**Determination of sequences of HaCPV-5 segments S7 and S10**

Sequences of genome segments S10 and S7 of HaCPV-5 were determined (GenBank accession nos DQ077912 and DQ077913, respectively). The two segments had conserved terminal sequences. They shared four conserved nucleotides at the 5′ end and four conserved nucleotides at the 3′ end (5′-AGUU...UUGC-3′; Table 1). Results from BLAST searching and sequencing comparisons revealed that the genome of HaCPV-5 S10 had high similarity to that of type 5 CPVs (EsCPV, OpCPV and HaCPV-5 South Africa strain). Deduced polyhedrin protein sequences were also similar to that of TnCPV-15, with 27% identity and 38% similarity. Plot analysis of the hydrophobicity of proteins encoded by BmCPV, HaCPV-5, HaCPV-14 and TnCPV-15 showed that there were some similar profiles in the N terminal of the protein, with three domains in the order hydrophobic–hydrophilic–hydrophobic (Fig. 3). Nucleotide sequence comparison of HaCPV-5 S7 showed that this sequence had no similarity to that of other CPVs. However, comparison of amino acid sequences encoded by HaCPV-5 S7 with sequences published in GenBank revealed that there were some similarities with type 1 CPVs, type 14 CPVs and, more interestingly, *Rice ragged stunt virus* (RRSV). Sequence comparison indicated that HaCPV-5 S7 shared 24% identity and 43% similarity with the S6-encoded protein of BmCPV, 27% identity and 47% similarity with HaCPV-14 proteins, and 25% identity and 45% similarity with S7 of RRSV. Results from comparison of nucleotide sequences encoding
Table 1. Properties of dsRNA segments of HaCPV-14, HaCPV-5 and HasCPV-14

For HaCPV-14, consensus sequences at the 5’ and 3’ NCRs are 5’-AGAAUUU and CAGCU-3’, respectively. For HaCPV-5, consensus sequences at the 5’ and 3’ NCRs are 5’-AGUU and UUGC-3’, respectively.

<table>
<thead>
<tr>
<th>Virus/segment</th>
<th>Putative encoded protein</th>
<th>5’ NCR</th>
<th>3’ NCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length (bp)</td>
<td>Length (aa)</td>
<td>Mass (Da)</td>
</tr>
<tr>
<td>HaCPV-14</td>
<td>9</td>
<td>1143</td>
<td>7325</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1250</td>
<td>7325</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1391</td>
<td>7325</td>
</tr>
</tbody>
</table>

Determination of the sequence of the HasCPV-14 polyhedrin gene

The sequence of the HasCPV-14 polyhedrin gene was determined (GenBank accession no. DQ077914). The S10 segment of HasCPV-14 consisted of 956 nt and contained a single long ORF of 249 aa, starting with the ATG codon at nt 53 and ending with a TAA stop codon at nt 802. Comparison of the nucleotide sequences of S10 of HasCPV-14 and HaCPV-14 showed that the genomes of these viruses were the same size and had nucleotide sequence identities of 92%. Although more than 70 nucleotide variations were identified between HasCPV-14 S10 and HaCPV-14 S10, there was only one amino acid change. Sequences of the polyhedrin genes from three type 1 CPVs (BmCPV, LdCPV-1 and DpCPV-1), four type 5 CPVs (EsCPV-5, OpCPV-5 and HaCPV-5 Chinese strain and South Africa strain), three type 14 CPVs (HaCPV-14, HasCPV-14 and LdCPV-14) and TnCPV-15 were retrieved from GenBank and the evolutionary relationships among these CPVs were evaluated by constructing a neighbour-joining tree for the polyhedrin genes (Fig. 4).

DISCUSSION

Members of the family Reoviridae are grouped into nine genera (Mertens et al., 2000). The natural hosts of these viruses include vertebrates, invertebrates and plants, and the virulence that they exhibit towards these hosts differs widely. The current classification was established on the basis of three main criteria: RNA–RNA hybridization, electropherotype analysis and antigenic properties (Payne & Mertens, 1983; Mertens et al., 1999, 2000). Although CPV sequence data can be found in GenBank, determination of the molecular relationships among CPVs has been hampered by the paucity of the available genetic information. In this study, the complete sequences of HaCPV-14 S10–S7, HaCPV-5 S10 and S7 and HasCPV S10 were determined. Computer analysis showed that each segment of the three CPVs contains a single ORF. HaCPV-14 segments had seven conserved nucleotides at the 5’ end and five at the 3’ end (5’-AGAAUUU...CAGCU-3’). The same conserved terminal sequences were found in HasCPV S10. In the two segments of HaCPV-5 that were sequenced, there were four conserved nucleotides at the 5’ end and four at the 3’ end (5’-AGUU...UUGC-3’). Such conserved motifs are detected in many viruses containing multisegmented genomes and it has been suggested that they play an important role in transcription, replication and packaging of RNA, and virus maturation (Anzola et al., 1987; Xu et al., 1989; Attoui et al., 1997). However, when different types of CPVs are compared, the conserved terminal sequences differ, suggesting that the conserved terminal sequences are the same only when the viruses belong to the same electrophoretic type (Kuchino et al., 1982). In CPVs, no nucleotide sequence identity was found between different types. Among the same types, evolution of the nucleotide sequences also differed. BLAST comparison of CPV genomes revealed that nucleotide sequences of types 1 (BmCPV, LdCPV-1 and DpCPV-1) and 14 (HaCPV-14 and HasCPV-14) mutate easily. Nucleotide sequence comparison in this study showed that S10 of type 1 CPVs, BmCPV, LdCPV-1 and DpCPV-1, shared about 90% identity. Our study also showed that type 14 CPVs have 92% identity. However, in different strains of type 5, the nucleotide sequences of S10 have high identity. OpCPV and EsCPV were isolated from the Pacific northwest of the USA and eastern Canada, respectively, whereas HaCPV isolates were from South Africa and China (Fossiez et al., 1989; Galinski et al., 1994). The geographical sites of their isolation are distant and, therefore, evolution of these viruses might have occurred over the years. However, the
nucleotide sequences of type 5 CPVs are possibly more conserved than those of other types.

Comparison of the nucleotide sequences of these viruses in GenBank showed that different types of CPVs have no sequence similarity. However, if amino acid sequences are compared, some similarities have been identified. In particular, the protein encoded by S7 of HaCPV-5 is similar to the proteins of RRSV. The same result has been reported by Ikeda et al. (2001), i.e. amino acid sequences of BmCPV S4 show 22% identity and 39% similarity with an S2-encoded P2 protein of RRSV. It has been suggested that the family Reoviridae could be divided into two distinct groups on the basis of the structures of the inner protein layers. One group, including orthoreoviruses, aquareoviruses, oryzaviruses and fijiviruses, is characterized by an incomplete $T=13$ layer attached loosely to the outside of a protein shell composed of 120 copies of a capsid protein and penetrated by substantial turret-like spike structures at the fivefold axes (Dryden et al., 1998; Shaw et al., 1996). CPVs are also turreted, but lack the $T=13$ layer. The other group includes orbiviruses, rotaviruses, phytoreovirus and coltiviruses.

Fig. 2. Multiple sequence alignment of proteins from BmCPV S6, LdCPV-1 S6, LdCPV-14 S6, HaCPV-5 S7 and RRSV S7. Black and shaded regions represent identical and similar amino acids, respectively. Identical amino acids among the five viruses are indicated with asterisks; chemically similar amino acids are indicated with one or two dots. The four highly conserved regions of the amino acid sequences are underlined. GenBank accession numbers of the sequences of BmCPV S6, LdCPV-1 S6, LdCPV-14 S6 and RRSV S7 are AB030014, AF389467, AF389457 and U66713, respectively.
The three-dimensional structures of two different types of CPV, BmCPV-1 (Zhou et al., 2003) and HaCPV-5 (Hill et al., 1999), were determined at almost the same time. Interestingly, although no identity is evident in the nucleotide sequences of these two types, their three-dimensional structures are very similar. Our results are consistent with these data. Amino acid sequences encoded by HaCPV-5 S7 are similar to those of S6-encoded VP4 of BmCPV-1 and S6-encoded P6 of HaCPV-14. VP4 is one of the structural proteins of BmCPV-1 (Hagiwara & Matsumoto, 2000). Based on these data, it is suggested that the structural proteins of different types of CPVs are conservative, with the exception of a non-structural protein encoded by RRSV S7 (Upadhyaya et al., 1997). Evolution among different types in the genera and among the different genera in the same family also differs. It was also observed that CPVs and oryzaviruses are related more closely than other genera in the family Reoviridae.

Mixed infections of insect colonies by several types of CPVs are common in nature (Payne & Rivers, 1976; Belloncik et al., 1996). In this study, components of a mixture of HaCPVs have been identified. Even though native CPV infection takes place only in the insect midgut, it has been demonstrated that insect cells from different origins are permissive or resistant to CPV infection in vitro. This difference in virus susceptibility permits the separation of CPV strains from mixed CPVs. Belloncik et al. (1996) first reported the successful separation of one CPV (HaCPV-14) from a mixture of HaCPV types by using cell culture. Cell cultures infected with HaCPV can be subcultured for more than 1 year (over 100 subcultures) without disappearance of polyhedra, showing that HaCPV type A (type 14) infection is more persistent in cell culture than EsCPV (Belloncik, 1989). EsCPV and HaCPV-5 belong to the same type and their nucleotide sequences have high identities. Therefore, HaCPV types 14 and 5 can easily be separated by cell culture. At the same time, HaCPV-5 is recognized to be more lethal for the insect than HaCPV-14. In our study, changing the rearing conditions resulted in successful separation of HaCPV-5 from the virus mixture. Type 14 CPV has only been found in dual infections before and, therefore, there might be some synergism in the dual infections. In this study, for the first time, a pure type 14 CPV from a natural source was isolated, suggesting that synergism is not necessary for infection of type 14 CPVs.

H. armigera is one of the most serious pests in China. As an economic polyphagous pest, it has caused considerable economic loss to many vegetable and field crops, such as cotton, corn, tobacco and wheat. Many years of chemical pest control have led to serious resistance and environmental pollution. Heliothis armigera nuclear polyhedrosis virus has been used as a commercial bioinsecticide for many years because of its specificity and harmlessness to other organisms, as well as the environment (Chen et al., 2000).
comparison with baculoviruses, the use of CPVs as an insecticide up to now is relatively limited, mainly due to their pathogenesis for the insects. CPV infection of an insect population is, in general, chronic rather than epidemic and CPV is recognized to be less lethal for the insect than a baculovirus. Therefore, the use of CPVs in stable ecosystems such as forests is more suitable, because of their persistence. *Dendrolimus spectabilis* CPV has been commercialized in Japan. According to several reports, some CPV types are more lethal than others, e.g. OpCPV is extremely virulent (Belloncik, 1996; Belloncik & Mori, 1998). Our results are consistent with these findings. OpCPV and HaCPV are both type 5 CPVs and it is possible that some type 5 CPVs are more virulent than other types. HaCPV has the potential to be used as a bioinsecticide in the future.

ACKNOWLEDGEMENTS

The authors are grateful to Professor Serge Belloncik for giving some advice and the critical reading of the manuscript.

REFERENCES


