Physical mapping and identification of interspersed homologous sequences in the *Trichoplusia ni* granulosis virus genome

Yoshifumi Hashimoto,1* Kana Hayashi,1 Yuuki Okuno,1 Tohru Hayakawa,1† Atsuya Saimoto,1‡ Robert R. Granados2 and Tsuguo Matsumoto1

1 Department of Applied Biology, Kyoto Institute of Technology, Sakyo-ku, Kyoto 606, Japan and
2 Boyce Thompson Institute for Plant Research, Cornell University, Ithaca, NY 14850, USA

A restriction fragment library representing 89.3% of the genome of *Trichoplusia ni* granulosis virus (TnGV) was constructed. The library consisted of 13 of the 16 BamHI fragments, 18 of the 22 EcoRI fragments, and 6 of the 27 PstI fragments. By restriction endonuclease and Southern blot analysis of cloned or genomic viral DNA fragments, a complete physical map of TnGV was constructed for BamHI, EcoRI, PstI and XhoI. Three interspersed homologous regions (ihs1–ihs3) were identified from hybridization experiments and sequenced. Each TnGV ihs has an approximate size of 400 bp and shows homology to the other two. The orientation of ihs2 is inverted relative to ihs1 and ihs3. TnGV ihs regions do not have repetitive motifs or palindromic sequences, in contrast to homologous regions (hrs) of nuclear polyhedrosis viruses (NPVs). The genomic locations of TnGV ihs1–ihs3, represented in percentage map units, were very similar to those of ihs sequences previously reported in *Bombyx mori* NPV, suggesting that the ihs may be a novel type of cis-acting element common among baculoviruses. Additionally, an inverted repeat sequence, having overlapping multiple inverted repeats of 400 bp, was identified to the left of ihs3 on the linearized genome map of TnGV.

Introduction

Granulosis viruses (GVs), which constitute a genus of the family *Baculoviridae*, are singly enveloped viruses with a circular double-stranded DNA genome (Francki et al., 1991; Tanada & Hess, 1991). Gene expression and molecular mechanisms of virus replication have not been thoroughly studied in GVs because of the limited availability of susceptible insect cell lines (Naser et al., 1984; Dwyer & Granados, 1988; Funk & Consigl, 1992; Winstanley & Crook, 1993). Genes identified to date are the granulin genes of *Trichoplusia ni* (Tn) GV (Akiyoshi et al., 1985), *Pieris brassicae* (Pb) GV (Chakerian et al., 1985) and *Cryptophlebia leucoptera* (Cl) GV (Jehle & Backhaus, 1994b); the enhancin genes of TnGV (Hashimoto et al., 1991) and *Pseudaletia unipuncta* (Pu) GV (Corsaro et al., 1993); the iap gene of *Cydia pomonella* (Cp) GV (Crook et al., 1993); and the basic DNA-binding protein gene of C1GV (Jehle & Backhaus, 1994a). Physical maps have been constructed for *Pieris rapae* (Pr) GV (Dwyer & Granados, 1987), CpGV (Crook et al., 1985), *Xestia c-nigrum* (Xc) GV (Goto et al., 1992), TnGV (Hashimoto et al., 1991), C1GV (Jehle et al., 1992) and several isolates of *Arctiogea rapae* (Ar) GV (Smith & Crook, 1988a, b, 1993).

Homologous regions (hrs) have been identified in *Autographa californica* multinucleocapsid nuclear polyhedrosis virus (AcMNPV) (Cochran & Faulkner, 1983), *Bombyx mori* (Bm) NPV (Maeda & Majima, 1990), *Choristoneura fumiferana* (Cf) MNPV (Arif & Doerfler, 1984; Kuzio & Faulkner, 1984), *Lymantria dispar* (Ld) MNPV (Smith et al., 1988) and *Orgyia pseudotsugata* (Op) MNPV (Theilmann & Stewart, 1992). All five hrs have been sequenced in AcMNPV and BmNPV (Guarino & Summers, 1986; Guarino et al., 1986; Majima et al., 1993), two of eight hrs in LdMNPV (Pearson & Rohrmann, 1995), one of five hrs in OpMNPV (Theilmann & Stewart, 1992) and one of four hrs in CfMNPV (Xie et al., 1995); the sequences of these hrs have repetitive motifs. The evidence accumulated on baculovirus hrs indicates that they are involved in viral gene expression and replication. AcMNPV hrs cis-
enhance trans-activation of the early genes 39K, IE-N and p35 by the IE-1 gene product (Guarino & Summers, 1986; Carson et al., 1991; Rodems & Friesen, 1993). AcMNPV, LdMNPV, OpMNPV and CfMNPV hrs may function as origins of DNA replication (Pearson et al., 1992; Kool et al., 1993a, b, 1994; Leisy & Rohrmann, 1993; Pearson & Rohrmann, 1995; Ahrens et al., 1995; Xie et al., 1995). The hrs of OpMNPV have been shown to increase gene expression from the AcMNPV p39 promoter and the OpMNPV IE-2 promoter (Theilmann & Stewart, 1992). Enhancer activities and DNA sequences have not been elucidated for the hrs of LdMNPV and CfMNPV. So far, no evidence of hrs in GV genomes has been reported.

In this paper we report that a novel type of interspersed homologous sequence (ihs), lacking repetitive DNA elements characteristic of the AcMNPV, BmNPV, OpMNPV and CfMNPV hrs, is present at three locations in the TnGV genome. The size, DNA homology and genome locations of TnGV ihs are similar to those of the BmNPV ihs previously reported (Hashimoto et al., 1994). Additionally, a unique region containing overlapping multiple imperfect inverted repeats, designated an inverted repeat sequence (irs), was identified.

### Methods

**Virus.** The origin of TnGV has been reported previously (Smith & Summers, 1978). TnGV was cloned by serial in vivo passages through T. ni neonates under conditions of limiting dilution (Huber & Hughes, 1984). Viral occlusion bodies (OBs) were produced by infection of fifth instar T. ni larvae with 10^8 OBs/larva, and purified as described by Dwyer & Granados (1987).

**Preparation of TnGV DNA.** To isolate viral DNA the OBs were dissolved in 0.1 M-NaCO_3, 0.17 M-NaCl, 1 mM-EDTA, pH 10.9 at room temperature for 30 min and centrifuged at 7000 g for 8 min to pellet undissolved material. The supernatant was layered on a 27-45 % (w/w) linear sucrose gradient and centrifuged at 90000 g for 40 min. The band of virus particles was removed, diluted with 3 vols sterile distilled water, and pelleted by centrifugation at 90000 g for 30 min. The virus preparation was incubated in 1% (w/v) SDS, 1 mM-EDTA, 0.5 M-NaCl, 0.2 mg/ml proteinase K (Merck), 10 mM-Tris-HCl (pH 7.4) at 37 °C for 4 h. Viral DNA was extracted twice with an equal volume of phenol-chloroform-isomyl alcohol (25:24:1) and once

### Table 1. Sizes of restriction endonuclease cleaved fragments of TnGV DNA

The sizes of restriction endonuclease fragments were determined as the mean of three independent measurements. The sizes of the fragments were determined by summing the sizes of the fragments generated by digestion with other restriction endonuclease. Fragments were cloned into pACYC184 (t) or pBluescriptII SK(+) (*).

<table>
<thead>
<tr>
<th>Fragment</th>
<th>BamHI</th>
<th>EcoRI</th>
<th>PstI</th>
<th>XhoI</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>46.2</td>
<td>28.4</td>
<td>24.1</td>
<td>22.6</td>
</tr>
<tr>
<td>B</td>
<td>28.1</td>
<td>19.85</td>
<td>21.6</td>
<td>19.85</td>
</tr>
<tr>
<td>C</td>
<td>17.5</td>
<td>18.3</td>
<td>17.1</td>
<td>18.6</td>
</tr>
<tr>
<td>D</td>
<td>11.1</td>
<td>14.6</td>
<td>10.7</td>
<td>17.6</td>
</tr>
<tr>
<td>E</td>
<td>10.1</td>
<td>13.5</td>
<td>9.2*</td>
<td>16.7</td>
</tr>
<tr>
<td>F</td>
<td>9.85</td>
<td>12.5</td>
<td>8.8*</td>
<td>13.3</td>
</tr>
<tr>
<td>G</td>
<td>9.5</td>
<td>11.9</td>
<td>8.7</td>
<td>12.4</td>
</tr>
<tr>
<td>H</td>
<td>8.0</td>
<td>9.0</td>
<td>8.7*</td>
<td>11.7</td>
</tr>
<tr>
<td>I</td>
<td>8.0†</td>
<td>8.4†</td>
<td>8.2</td>
<td>8.3</td>
</tr>
<tr>
<td>J</td>
<td>7.9†</td>
<td>7.6†</td>
<td>7.8</td>
<td>7.6</td>
</tr>
<tr>
<td>K</td>
<td>4.9†</td>
<td>6.5†</td>
<td>6.1</td>
<td>6.8</td>
</tr>
<tr>
<td>L</td>
<td>4.8†</td>
<td>4.7†</td>
<td>5.4</td>
<td>6.0</td>
</tr>
<tr>
<td>M</td>
<td>3.3†</td>
<td>3.9†</td>
<td>5.35</td>
<td>4.2</td>
</tr>
<tr>
<td>N</td>
<td>3.2†</td>
<td>3.85†</td>
<td>4.25</td>
<td>3.5</td>
</tr>
<tr>
<td>O</td>
<td>2.8†</td>
<td>2.65†</td>
<td>3.65</td>
<td>3.1</td>
</tr>
<tr>
<td>P</td>
<td>0.4†</td>
<td>2.45†</td>
<td>3.6</td>
<td>1.4</td>
</tr>
<tr>
<td>Q</td>
<td>2.2†</td>
<td>3.4</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>1.9†</td>
<td>3.4</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1.35†</td>
<td>3.15</td>
<td>2.6*</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1.05†</td>
<td>2.1*</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>0.6†</td>
<td>1.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.5†</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>0.65*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>0.65*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>175.6</td>
<td>175.6</td>
<td>175.6</td>
<td>175.6</td>
</tr>
</tbody>
</table>
with diethyl ether. The sample was then dialysed against TE buffer at 4 °C for 24 h.

Restriction endonuclease digestion and gel electrophoresis. Purified viral DNA was digested with BsmHI, EcoRI, PstI or XhoI and electrophoresed on a 0.7% agarose gel in Tris-acetate buffer (Maniatis et al., 1982). DNA fragments used as standards for size determination were HindIII or HaelI digests of phage λ DNA. The sizes of fragments larger than 10 kbp were determined by summing the sizes of fragments generated by digestion with other enzymes. All enzymes were purchased from Takara and used as directed by the manufacturer.

Construction of genomic DNA libraries. DNA libraries were constructed by ligating TnGV DNA to pACYC184 or pBluescriptII SK(+) vectors using a DNA ligation kit (Takara). Recombinant plasmids were propagated in Escherichia coli HB101 or JM109, and purified by CsCl isopycnic ultracentrifugation or by alkaline lysis using standard procedures (Maniatis et al., 1982). Cloned DNA fragments were authenticated by comparing their migration in agarose gels with cloned DNAs with appropriate enzymes, separating the fragments on a gel, and purifying the insert DNA with Ultrafree C3GV (Millipore). The DNA was labelled with [α-32P]dCTP (NEN) using a random-primed DNA labelling kit (United States Biochemical). Prehybridization was carried out in 6 x SSC, 0.5% SDS, 5 x Denhardt's reagent and 100 μg/ml of denatured salmon sperm DNA at 68 °C for 2 h. The probe was denatured in 0.2 M-NaOH at room temperature for 5 min, and added to the prehybridization solution. After hybridization at 68 °C for 18 h, the membranes were washed twice in 2 x SSC, 0.1% SDS at room temperature for 15 min, once in 0.1 x SSC, 0.5% SDS at 65 °C for 30 min, and exposed to Kodak XRP-5 film (Eastman Kodak) with an intensifying screen at -80 °C.

DNA sequencing. Nucleotide sequences from TnGV genomic regions 35.2-35.6 map units (m.u.), 69.8-70.4 m.u., 87.7-88.1 m.u. and 85.0-85.9 m.u. were determined using plasmids which contained overlapping deletions from PstI-BsmHI (35.0-37.1 m.u.), PstI-T (69.1-70.6 m.u.), EcoRI-BsmHI (87.4-88.4 m.u.) and EcoRI-R (85.0-86.0 m.u.), respectively. Plasmid DNA was digested with exonuclease III and Mung bean nuclease for 1 h and 3 h, and with nuclease Bal31 for 1 h, 2 h and 3 h. The nucleotide sequence was determined from double-stranded DNA templates by the dideoxynucleotide chain termination method (Sanger et al., 1977) using a Sequenase version 2.0 kit (United States Biochemical). Nucleotide sequences from both DNA strands were obtained throughout this region.
Results and Discussion

Restriction endonuclease analysis and construction of genomic libraries of the TnGV genome

After electrophoresis of digested DNA (Fig. 1), the sizes of TnGV DNA restriction fragments (Table 1) were estimated by comparison with λ DNA digests. Restriction profiles did not show any noticeable submolar bands, indicating that the virus isolated by serial passages in vivo was genetically homogeneous. The electrophoretic profiles were essentially identical to those of TnGV DNA reported by Smith & Summers (1978) and were different from those reported for other GV genomes (Smith & Summers, 1978; Tweeten et al., 1980; Crook, 1981; Burges, 1983; Harvey & Volkman, 1983; Harvey & Tanada, 1985; Arif et al., 1986; Crook, 1986; Dwyer & Granados, 1987; Easwaramoorthy & Cory, 1990; Goto et al., 1992).

Thirteen of the 16 BamHI fragments and 18 of the 22 EcoRI fragments were cloned into pACYC184, and six of the 27 PstI fragments were cloned into pBluescriptII SK(+) (Table 1). These clones together covered 89.3% of the viral genome; uncloned regions are BamHI-E (26-8-32-6 m.u.) and PstI-G (63-8-68-7 m.u.).

Physical mapping of the TnGV genome

Restriction maps of the TnGV genome were determined using recombinant plasmids or uncloned DNA fragments, extracted from agarose gels, as probes for Southern blot hybridization to electrophoretically separated restriction digests of TnGV DNA. The circular genome is shown linearized at the EcoRI-E/S junction in Fig. 2, so that the smallest mapped fragment containing the granulin gene (Akiyoshi et al., 1985) is situated at the leftmost end of the map in accordance with the convention proposed by Vlak & Smith (1982). Positions of restriction sites on this map are shown in m.u., and the map is oriented in the direction of granulin gene transcription. The enhancin gene (Hashimoto et al., 1991) is located at 88-1-89-6 m.u. within EcoRI-K.

Locations and nucleotide sequences of ihs regions

In the hybridization experiments, some probes also hybridized to non-collinear regions of the genome, suggesting that TnGV DNA possesses interspersed homologous sequences. For example, hybridization of the EcoRI–BamHI fragment at 87-4-88-4 m.u. to blots of TnGV DNA digested with BamHI, EcoRI, PstI and XhoI. Hybridizing fragments which do not correspond to the map position of the probe are in parentheses. These results indicated that three ihs regions occur in the TnGV genome within the PstI–BamHI (35-0-37-1 m.u.), PstI-T (69-1-70-6 m.u.) and EcoRI–BamHI (87-4-88-4 m.u.) fragments.

To determine the nucleotide sequence of the ihs regions, the above fragments were subcloned and used to generate nested deletions. Hybridizations were then carried out to locate more specifically the ihs regions within the clones. DNA sequencing revealed the presence of three ihs sequences (ihs1, 35-3-35-5 m.u.; ihs2, 70-1-70-3 m.u.; ihs3, 87-8-88-0 m.u.) and part of a fourth (ihs2a, 69-9-70-0 m.u.), showing significant identity to each other. Nucleotide sequences of the three TnGV regions containing the ihs sequences are shown in Fig. 4. Each ihs covers approximately 400 bp and has neither repetitive motifs nor palindromic structures, in contrast to the hr regions of NPV genomes (Guarino et al., 1986; Theilmann & Stewart, 1992; Majima et al., 1993; Pearson & Rohrmann, 1995; Xie et al., 1995). The ihs regions show sequence identities to each other of 88% in the ‘upstream’ 179 bp; 73% in the ‘downstream’ 105 bp; and 35% in the central 109 bp (Fig. 5). TnGV ihs2a contained only the ‘downstream’ element (162 bp) of the

---

Fig. 3. Hybridization of a 32P-labelled EcoRI–BamHI fragment (87-4-88-4 m.u.) to Southern blots of TnGV DNA digested with BamHI, EcoRI, PstI and XhoI. Hybridizing fragments which do not correspond to the map position of the probe are in parentheses.
core region which, like ihs2, was in inverted orientation relative to ihs1 and ihs3. The ihs2 did not show any sequence identity to the hrs of AcMNPV, BmNPV, LdMNPV, OpMNPV or CfMNPV (Guarino et al., 1986; Theilmann & Stewart, 1992; Majima et al., 1993; Pearson & Rohrmann, 1995; Xie et al., 1995).
Three ihs have also been identified and sequenced in BmNPV isolates D1 and T3 (Hashimoto et al., 1994). BmNPV D1 ihs1, ihs2 and ihs3 are each about 320 bp in size and located close to hr1, hr2 and hr3 on the genome. Together with the results in this paper, these data suggest that the ihs in different baculovirus subgroups share certain similarities. The sizes of TnGV ihs are similar to those of BmNPV D1 and T3. Comparing the DNA sequences of the ihs of BmNPV with those of TnGV, there is a high degree of identity in the ‘upstream’ 100–180 bp and the ‘downstream’ 110 bp. The identity between TnGV ihs and BmNPV ihs was about 50%. The relative positions of ihs on these viral genomes are similar (Fig. 6). These findings may indicate that baculovirus ihs act as cis-elements to regulate replication, expression levels or packaging of the genome.

To see whether sequences similar to TnGV ihs and BmNPV ihs were present in AcMNPV, we searched the entire sequence of AcMNPV isolate C6 (Possee et al., 1991; Ayres et al., 1994). Identity was detected in a

---

Fig. 6. Locations of the ihs on EcoRI maps of TnGV and BmNPV genomes. The genomes are represented in m.u. (%) and the positions of the three BmNPV ihs are aligned with those of TnGV.

---

![Diagram of genome locations](image)

---

Fig. 7. Nucleotide sequence of the leftmost 1541 bp of EcoRI-R (85.0–86.0 m.u.) and the arms of six inverted repeats, I–VI, in the irsts indicated by arrows along the sequence (a), and a schematic representation of the six inverted repeats in the irsts (b).
region overlapping ORF2 in AcMNPV EcoRI-I. This region is not reiterated elsewhere in AcMNPV DNA, and appears to correspond to ihs1 of BmNPV isolates. Using the program Search for Open Reading Frames (GENETIX, version 6.2.0), ORFs were detected within or overlapping with TnGV ihs. ORFs were also detected within BmNPV ihs. Ahrens & Rohrmann (1995) recently revealed that the HindIII-N fragment of OpMNPV DNA has a binary function as a replication origin and a coding region of lef-1. These observations suggest that the ihs may also play roles as cis- and trans-acting regions in virus replication.

A unique region containing multiple overlapping imperfect inverted repeats in the TnGV genome

Since three BmNPV ihs are located close to hr1, hr2 and hr3 on the genome, we assumed that hr homologues of TnGV might be present in regions close to the ihs; therefore we sequenced several restriction DNA fragments in the plasmid clones. When the nucleotide sequences of the EcoRI-P, R fragments were examined, a region of nearly 400 bp within EcoRI-R having multiple overlapping imperfect inverted repeats was found to the left of ihs3 and was designated an irs. The nucleotide sequence leftmost of EcoRI-R is shown in Fig. 7(a). TnGV irs is an AT-rich region located at 85.4–85.6 m.u. To see whether sequences homologous to ihs were present elsewhere in the genomes of TnGV or BmNPV, EcoRI-R was hybridized to blots of genomic restriction digests of TnGV and BmNPV DNA. No hybridization signals were observed other than in TnGV EcoRI-R and restriction fragments sharing the same map location (data not shown). Analysis of the irs with the program Search Inverted & Complementary Repeats (GENETIX, version 6.2.0) (minimum length = 30 bp, maximum length = 50 bp, matching percentage > 80) showed the presence of 22 inverted repeats, composed of six sets of overlapping ones. The inverted repeats together with the maximum size of each group are shown in Fig. 7(b). Owing to the overlapping nature of these palindromes, they may be able to form thermodynamically diverse configurations, as may occur in the hrs. The hrs and irs both possess multiple inverted repeats, but they differ with respect to whether each inverted repeat has a motif sequence with a similar size and whether they are arranged in tandem or overlap. Analysis of the irs with the program Search for Open Reading Frames (GENETIX, version 6.2.0), showed no significant ORFs in this region.

Preliminary experiments showed that plasmids containing complete TnGV ihs or irs sequences were unable to replicate in Spodoptera frugiperda cells infected with AcMNPV (data not shown). Further evaluation using cell lines susceptible to TnGV should demonstrate whether these elements can act as origins of DNA replication. Thus, when compared with NPV hrs TnGV irs has neither DNA sequence identity nor replication activity in a heterologous replication system (Pearson et al., 1992; Kool et al., 1993a, b; Leisy & Rohrmann, 1993). Nonetheless, there are similar structural peculiarities in both types of element, suggesting that DNA sequences common to NPV hr and TnGV irs may provide some basic information on replication mechanisms in baculoviruses. Examination of the DNA sequences of other GV genomes, to see whether regions similar to TnGV irs and ihs occur, would also be of interest.

We thank Dr Ian R. L. Smith for careful review of this manuscript and Masamitsu Shikata and Takashi Minakata for reviewing nucleotide sequences of TnGV ihs and irs. We also thank Dr Susumu Maeda for providing DNA sequence from BmNPV isolate T3. This work was partly supported by grants from the Ministry of Education, Science and Culture of Japan (05806008, 02044087).

References


(Received 27 June 1995; Accepted 20 October 1995)