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. Granados[2] and Tsuguo Matsumoto[1]

[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 *Bam*HI fragments, 18 of the 22 *Eco*RI fragments, and 6 of the 27 *Pst*I fragments. By restriction endonuclease and Southern blot analysis of cloned or genomic viral DNA fragments, a complete physical map of TnGV was constructed for *Bam*HI, *Eco*RI, *Pst*I and *Xho*I. Three interspersed homologous regions (*ihs*1–*ihs*3) 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 *ihs*2 is inverted relative to *ihs*1 and *ihs*3. 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 *ihs*1–*ihs*3, 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 *ihs*3 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 & Consigli, 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 leucotreta* (Cl) GV (Jehle & Backhaus, 1994*b*); 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 ClGV (Jehle & Backhaus, 1994*a*). 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), ClGV (Jehle *et al.*, 1992) and several isolates of *Artogeia rapae* (Ar) GV (Smith & Crook, 1988*a*, *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*-

* Author for correspondence. Fax +81 75 724 7760/7710. e-mail lmp@ipc.kit.ac.jp

† Present address: Faculty of Agriculture, Hokkaido University, Sapporo 060, Japan.

‡ Present address: Research Laboratories, Pharmaceuticals Group, Nippon Kayaku Co. Ltd, Tokyo 115, Japan.

The nucleotide sequences reported in this article have been deposited in the DDBJ DNA database under the accession numbers D58375, D58376, D58377 and D58378.

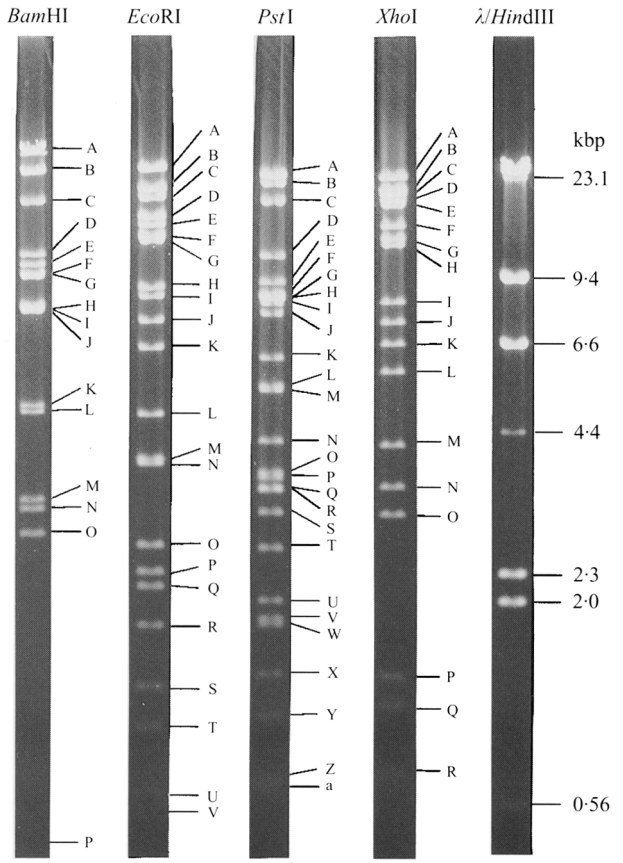

Fig. 1. Cleavage patterns of TnGV DNA using the restriction endonucleases *Bam*HI, *Eco*RI, *Pst*I and *Xho*I. Viral DNA was digested and the fragments were separated on a 0·7% agarose gel. The rightmost lane shows a *Hin*dIII digest of λ DNA with sizes indicated to the right (kbp). Letter assignments for each TnGV DNA fragment are shown.

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 (†) or pBluescriptII SK(+) (*).

	Size (kbp)			
Fragment	*Bam*HI	*Eco*RI	*Pst*I	*Xho*I
A	46·2	28·4	24·1	22·6
B	28·1†	19·85†	21·6	19·85
C	17·5†	18·3	17·1	18·6
D	11·1†	14·6	10·7	17·6
E	10·1	13·5†	9·2*	16·7
F	9·8†	12·5†	8·8*	13·3
G	9·5†	11·9	8·7	12·4
H	8·0	9·0†	8·7*	11·7
I	8·0†	8·4†	8·2	8·3
J	7·9†	7·6†	7·8	7·6
K	4·9†	6·5†	6·1	6·8
L	4·8†	4·7†	5·4	6·0
M	3·3†	3·9†	5·35	4·2
N	3·2†	3·85†	4·25	3·5
O	2·8†	2·65†	3·65	3·1
P	0·4†	2·45†	3·6	1·4
Q		2·2†	3·4	1·2
R		1·8†	3·4	0·75
S		1·35†	3·15	
T		1·05†	2·6*	
U		0·6†	2·1*	
V		0·5†	1·95	
W			1·9	
X			1·4	
Y			1·1	
Z			0·7	
a			0·65*	
Total	175·6	175·6	175·6	175·6

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 *hr*s may function as origins of DNA replication (Pearson *et al.*, 1992; Kool *et al.*, 1993*a*, *b*, 1994; Leisy & Rohrmann, 1993; Pearson & Rohrmann, 1995; Ahrens *et al.*, 1995; Xie *et al.*, 1995). The *hr*s 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 *hr*s of LdMNPV and CfMNPV. So far, no evidence of *hr*s 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 *hr*s, is present at three locations in the TnGV genome. The size, DNA homology and genome locations of TnGV *ihs*s are similar to those of the BmNPV *ihs*s 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-Na_2CO_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–isoamyl alcohol (25:24:1) and once

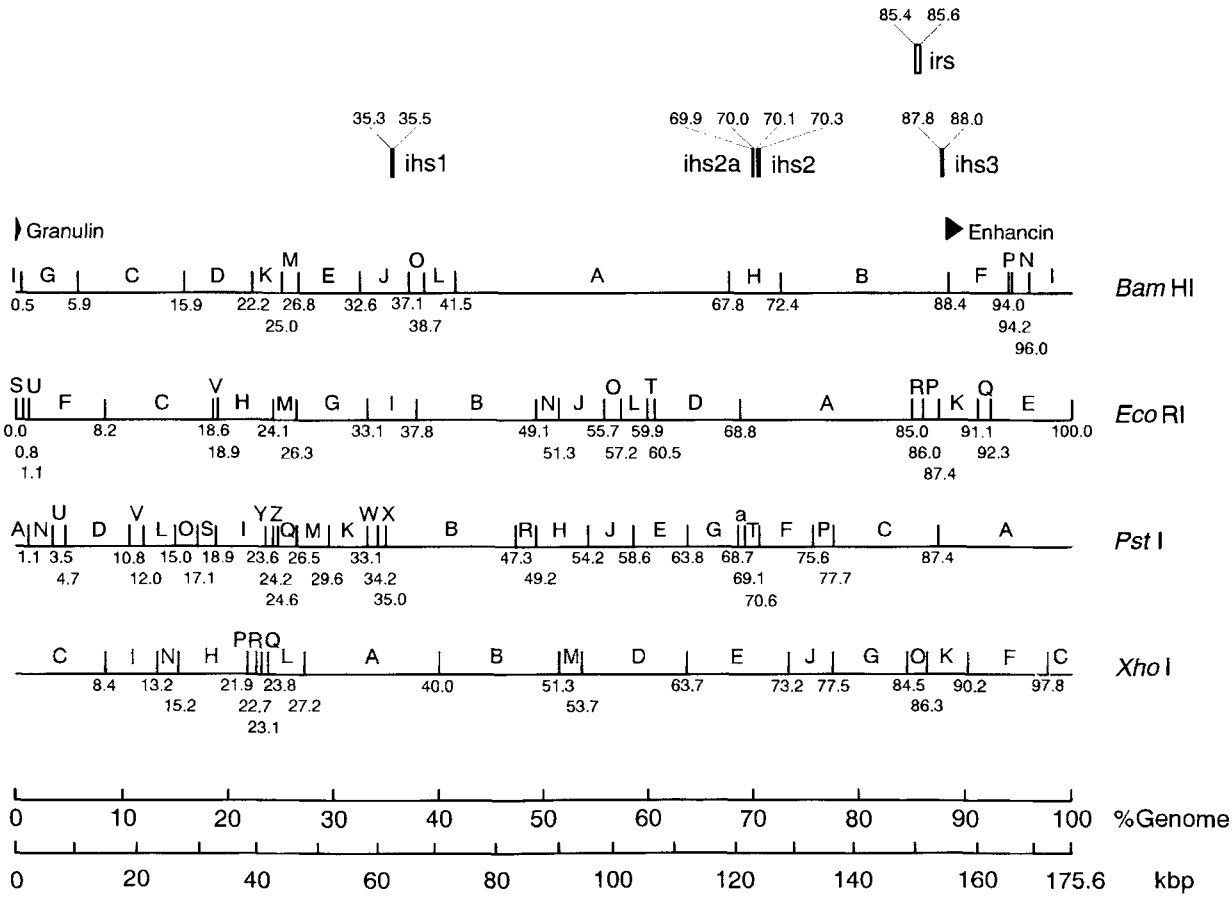

Fig. 2. Restriction map of TnGV genome for the enzymes *Bam*HI, *Eco*RI, *Pst*I and *Xho*I. Each cleavage site is indicated with a vertical line and numbered as a percentage of the genome. The positions and the orientations of the granulin and enhancin genes are shown by arrowheads. The map positions of interspersed homologous sequences (*ihs*s) 1, 2, 2a and 3, and an inverted repeat sequence (*irs*) are indicated at the top.

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 *Bam*HI, *Eco*RI, *Pst*I or *Xho*I 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 *Hin*dIII or *Hae*III 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 TnGV DNA fragments generated by digestion with the same enzyme.

Southern blot hybridization. After electrophoresis on 0·7% agarose gels, digests (approximately 1 μg) of TnGV DNA were transferred bidirectionally to Hybond-N+ (Amersham) and hybridized with ^{32}P-labelled TnGV DNA fragments. Probes were prepared by digesting 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-^{32}P]dCTP (NEN) using a random-primed DNA labelling kit (United States Biochemical). Prehybridization was carried out in 6 × SSC, 0·5% SDS, 5 × 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 15–18 h, the membranes were washed twice in 2 × SSC, 0·1% SDS at room temperature for 15 min, once in 0·1 × 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 *Pst*I–*Bam*HI (35·0–37·1 m.u.), *Pst*I-T (69·1–70·6 m.u.), *Eco*RI–*Bam*HI (87·4–88·4 m.u.) and *Eco*RI-R (85·0–86·0 m.u.), respectively. Plasmid DNA was deleted with exonuclease III and Mung bean nuclease for *ihs*1 and *ihs*3, and with nuclease *Bal* 31 for *ihs*2, *ihs*2a and *irs*. 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 *Bam*HI fragments and 18 of the 22 *Eco*RI fragments were cloned into pACYC184, and six of the 27 *Pst*I fragments were cloned into pBluescriptII SK(+) (Table 1). These clones together covered 89·3% of the viral genome; uncloned regions are *Bam*HI-E (26·8–32·6 m.u.) and *Pst*I-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 *Eco*RI-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 *Eco*RI-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 *Eco*RI–*Bam*HI fragment at 87·4–88·4 m.u. to blots of TnGV restriction digests showed that in addition to the expected fragments (*Bam*HI-B; *Eco*RI-K; *Pst*I-A; *Xho*I-K) the probe hybridized to other fragments (*Bam*HI-H, J; *Eco*RI-A, I; *Pst*I-B, T; *Xho*I-A, E) which do not correspond to the map position of the probe (Fig. 3). These results indicated that three *ihs* regions occur in the TnGV genome within the *Pst*I–*Bam*HI (35·0–37·1 m.u.), *Pst*I-T (69·1–70·6 m.u.) and *Eco*RI–*Bam*HI (87·4–88·4 m.u.) fragments.

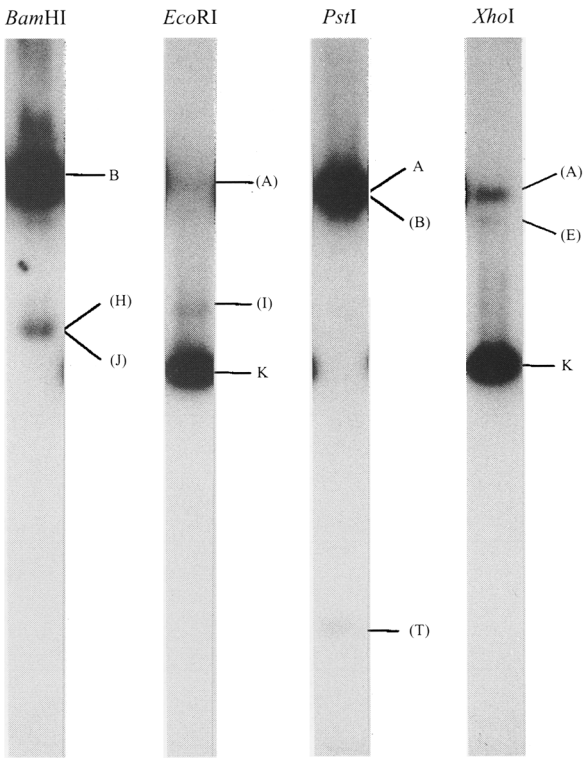

Fig. 3. Hybridization of a ^{32}P-labelled *Eco*RI–*Bam*HI fragment (87·4–88·4 m.u.) to Southern blots of TnGV DNA digested with *Bam*HI, *Eco*RI, *Pst*I and *Xho*I. Hybridizing fragments which do not correspond to the map position of the probe are in parentheses.

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*s (*ihs*1, 35·3–35·5 m.u.; *ihs*2, 70·1–70·3 m.u.; *ihs*3, 87·8–88·0 m.u.) and part of a fourth (*ihs*2a, 69·9–70·0 m.u.), showing significant identity to each other. Nucleotide sequences of the three TnGV regions containing the *ihs*s 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 *ihs*2a contained only the 'downstream' element (162 bp) of the

Nucleotide sequence from 35.2 m.u. to 35.6 m.u.

```
GCCCCTTGTA CTCCATAATT TGACGTTCGT AACACTGCTC CACGTCGCGC AATTTCGTTT CAAATTCAGC ACGACACATC ACCAGCTCGT TACGATGATC TGCTCTAATT GTAGCCAATT 120
GATTCTGCAA GGGTATAATT TCTTGCAGGT ATTTGTTTGC CCAATTGAGC ATGTTACAGT CTTGTTGTTT TTGTTGATTG TCTTGAATGG AGTACTTGCC GCTTCTCCTC AACTCAGGCA 240
GCACCTCTTC GAACAACCAA CGCTGGAACT CCTCGGCAGC AGGTAGCTTG GATCGCATAA TTAAAGCGTA AACTCCAGCT TCGCTGATAA AGACTGTGTT TGGTTGCCAG TTTACTGGTA 360
TATTGTCTTG AGCCAATGAC GTCACAAGGG GGCCCTGATT TAAGGCTCCC TTGATTTCTT CCCAATTCTT TCGCCAAGCC GGTTTAACGT GTTGCTGTAA AGCGTTTCGA GGCTTAACGT 480
                          ApaI(35.4m.u.)
ATCCTAAAAA TTGCGCTACA CCATGTCCGC CGTACATAAA TTTATCATTT TCCACTTCCA CAATCCAAAT TTCCGCAGAA ACTCCTCCAA TGTTGCACGA TTTTTTTACG AGAGACATTT 600
TGTTATTTTT GAAGGAAGAT GTGTGATAGC GACGAAATGA GCCGATAAAC TGTGCAGCAT CCCGAAATCG CTCGCT                                                  676
```

Nucleotide sequence from 69.8 m.u. to 70.4 m.u.

```
CCTAATCTCT GCAAAATGTC TCTCTCGTAA AAAAATCTTG CAACATGGAG GAGTTACCGC AGAAGTTTGG ATTGTGGAAG TGGAAAAGGA CAAATTTTTG TACGGAGGAC ATGGTATAGC 120
GCAGTTTTTG GGCTACACTA ATCCAAGAAA AGCAGTAAGA GATCATGTAA AACTAAAGTG GCGCAAAAAT TGGGAAGAAA TCTCAAGAGG AACGAATCGT TCCTCTTACG TGACGTCATG 240
TGAACAAACA GTACGATAAC GTGAAGATGG CTTACAAGCT TTTGTAAAAC CGAGCCGATG TTACTTATCA ATTTATAGGT CAAAGGTGAC GGTTTTATAA TGCGCGTTTT ACTCTCAGTC 360
                                    HindIII(70.0m.u.)
TAATCATTCG ACATCAAATA ATCTTTTGGT ACGCACGCAT CTAGCGCTTC ATTATGTCTC TCGTAAAAAA GTCTTGTAAT ATTGGAGGAG TTACTGCCGA AATTTGGATT GTAAAAGTAG 480
AAAAAGACAA TTTCATGTAT GGTGGACATG GTATTGCACA TGTTTTGGGC TATAAGCAAC CCAAAGACGC TATACGTAAT CACGTAAAAC CGCAATGGAA AACAAATTGG GAAGAGATTG 600
AGGGGGCGAT AAATCATCGC CCCCTTGTGA CGTCACTCGA TCAAGACAAT ATACCAGTAA ATTGGCAACC AAACACAGTG TTTATCAGTG AAGCGGGTGT TTACGCTTTG ATAATGAAGT 720
CAAAGCTACC AGCAGCAGAA GAATTTCAGC GTTGGTTGTT TGAAGAGGTT TTGCCTGAGT TGAGGAGAAG CGGCAAATAC TCTATTGAAA AAGATCAACA ACCGACGTCC ACTGATATTG 840
TAAATTATGA CAAAAAATTG GCAGAAGCAC AAATGGAAGC TATGCAGTTA AAATTAAAAC TATCTGAGGC TAACACCATA ATCGCTAATT ATAACACCAC TATTTCTGAA ATGAAG     956
```

Nucleotide sequence from 87.7 m.u. to 88.1 m.u.

```
AAGCTTGTAC TCTCGCTCCT TGTATTCGGC AATTTGACGT TCGTAGTTCC TCTGCATCTC TGACATGGCA ATATCATGAT TATGCTTAGC AACACTTAAA TCCAACTTTA ACTGCAAAGC 120
HindIII(87.7m.u.)                                                                                          HindIII(87.8m.u.)
TTCAATTTGA GCGTCAGCAG TTTTTTGTCG TAATTTACTA CATTAAAGTC TTTTTCCTTA TGCCGCGTAC TTTCCAGTAT TCCTTATTTC AGGCAGCACC TCTTCAAACA ACCAACGCTG 240
GAACTCCTCG GCAGCAGGTA GCTTTGATCT CATTATCAAA GCGTAAACTC AGCTTCGCTG ATAAAGACTG TGTTGGTGCC AATTACTGGT ATATTGCTAT GATCGAGTGA CGTCATAAGG 360
GTCGATTCGC GACGGTCTCC CACGTGGCCT TCCATTGAGG TTTTACGTGT TTGTGTATGC GTGCGAGTGT TTTTATAACC CAAAAACTCA GCCACACCGT GTCCACCGTA CATATACTTG 480
TCCTTTTCCA ATTCCACAAT CCAAATTTCC GCAGAAACTC CTCCAATGTT GCACGATTTT TTTACAAGAG TCATTTTGCA CGTTTACAAG AAATTTATTA CAAGATTAGC TGCTTGTGAT 600
AAAGGTCTGC ACGAGATGAG ATTCAAATAC GTAATGAGAA TTGCGTGATT TGCACGAGTT TATATAGCAT AATTTGCTAG                                              680
```

Fig. 4. Nucleotide sequences of three regions of TnGV DNA containing the *ihs*s. Some restriction sites are indicated below the sequence.

```
ihs1  212 GTACTTGCCG CTTCTCCTCA ACTCAGGCAG CACCTCTTCG AACAACCAAC GCTGGAACTC CTCGGCAGCA GGTAGCTTGG ATCGCATAAT TAAAGCGTAA ACTCCAGCTT CGCTGATAAA 331
ihs2  800 GTATTTGCCG CTTCTCCTCA ACTCAGGCAA AACCTCTTCA AACAACCAAC GCTGAAATTC TTCTGCTGCT GGTAGCTTTG ACTTCATTAT CAAAGCGTAA ACACCCGCTT CACTGATAAA 681
ihs3  187 GTACTTTCCA GTATTCCTTA TTTCAGGCAG CACCTCTTCA AACAACCAAC GCTGGAACTC CTCGGCAGCA GGTAGCTTTG ATCTCATTAT CAAAGCGTAA ACTC-AGCTT CGCTGATAAA 306

          GACTGTGTTT GGTTGCCAGT TTACTGGTAT ATTGTCTTGA GCCAATGACG TCACAAGGGG GCCCTGATTT AAGGCTCCCT TGATTTCTTC CCAATTCTTT CGCCAAGCCG GTTTAACGTG 451
          CACTGTGTTT GGTTGCCAAT TTACTGGTAT ATTGTCTTGA TCGAGTGACG TCACAAGGGG GCGATGATTT ATCGCCCCCT CAATCTCTTC CCAATTTGTT TTCCATTGCG GTTTTACGTG 561
          GACTGTGTT- GGT-GCCAAT T-ACTGGTAT ATTGCTATGA TCGAGTGACG TCATAAGGGT C-GATTCGCG ACGGTCTCCC ACGTGGCCTT CCA-TTGAG- ---------- GTTTTACGTG 410
          ihs2a 173 GTTTTACATG 164

          TTGCTGTAAA GCGTTTCGAG GCTTAACGTA TCCTAAAAAT TGCGCTACAC CATGTCCGCC GTACATAAAT TTATCATTTT CCACTTCCAC AATCCAAATT TCCGCAGAAA CTCCTCCAAT 571
          ATTACGTATA GCGTCTTTGG GTTGCTTATA GCCCAAAACA TGTGCAATAC CATGTCCACC ATACATGAAA TTGTCTTTTT CTACTTTTAC AATCCAAATT TCGGCAGTAA CTCCTCCAAT 441
          TTTGTGTATG CGTGC--GAG TGTTTTTATA ACCCAAAACT CAGCCA-CAC CGTGTCCACC GTACATATAC TTGTCCTTTT CCAATTCCAC AATCCAAATT TCCGCAGAAA CTCCTCCAAT 527
          ATCTCTTACT GCTTTTCTTG GATTAGTGTA GCCCAAAAAC TGCGCTATAC CATGTCCTCC GTACAAAAAT TTGTCCTTTT CCACTTCCAC AATCCAAACT TCTGCGGTAA CTCCTCCA-T  45

          GTTGCACGAT TTTTTTACGA GAGA--CATT TTG                                                                                          602
          ATTACAAGAC TTTTTTACGA GAGA--CATA ATG                                                                                          410
          GTTGCACGAT TTTTTTACAA GAGT--CATT TTG                                                                                          558
          GTTGCAAGAT TTTTTTACGA GAGAGACATT TTG                                                                                           12
```

Fig. 5. Alignment of the sequences of *ihs*s. The nucleotide sequences of three *ihs*s and a part of *ihs*2a are presented, with gaps (-) inserted for optimal alignment. The sequences of *ihs*2 and *ihs*2a are presented in inverted orientation relative to the genetic map. The 'upstream' and 'downstream' regions of the *ihs*s showing high sequence similarity are underlined, and positions with a common nucleotide are indicated by a dot above them. Numbers correspond to nucleotide positions shown in Fig. 4.

core region which, like *ihs*2, was in inverted orientation relative to *ihs*1 and *ihs*3. The *ihs*s did not show any sequence identity to the *hr*s 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).

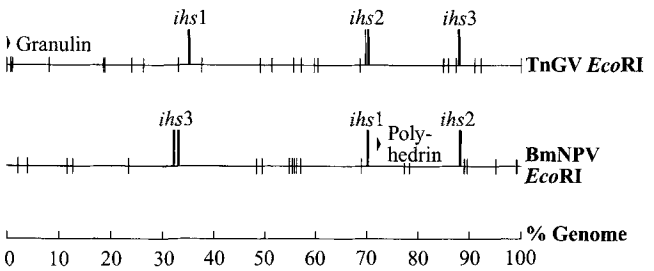

Fig. 6. Locations of the *ihs*s on *Eco*RI maps of TnGV and BmNPV genomes. The genomes are represented in m.u. (%) and the positions of the three BmNPV *ihs*s are aligned with those of TnGV.

Three *ihs*s have also been identified and sequenced in BmNPV isolates D1 and T3 (Hashimoto *et al.*, 1994). BmNPV D1 *ihs*1, *ihs*2 and *ihs*3 are each about 320 bp in size and located close to *hr*1, *hr*2 and *hr*3 on the genome. Together with the results in this paper, these data suggest that the *ihs*s in different baculovirus subgroups share certain similarities. The sizes of TnGV *ihs*s are similar to those of BmNPV D1 and T3. Comparing the DNA sequences of the *ihs*s 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*s and BmNPV *ihs*s was about 50%. The relative positions of *ihs*s on these viral genomes are similar (Fig. 6). These findings may indicate that baculovirus *ihs*s act as *cis*-elements to regulate replication, expression levels or packaging of the genome.

To see whether sequences similar to TnGV *ihs*s and BmNPV *ihs*s 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

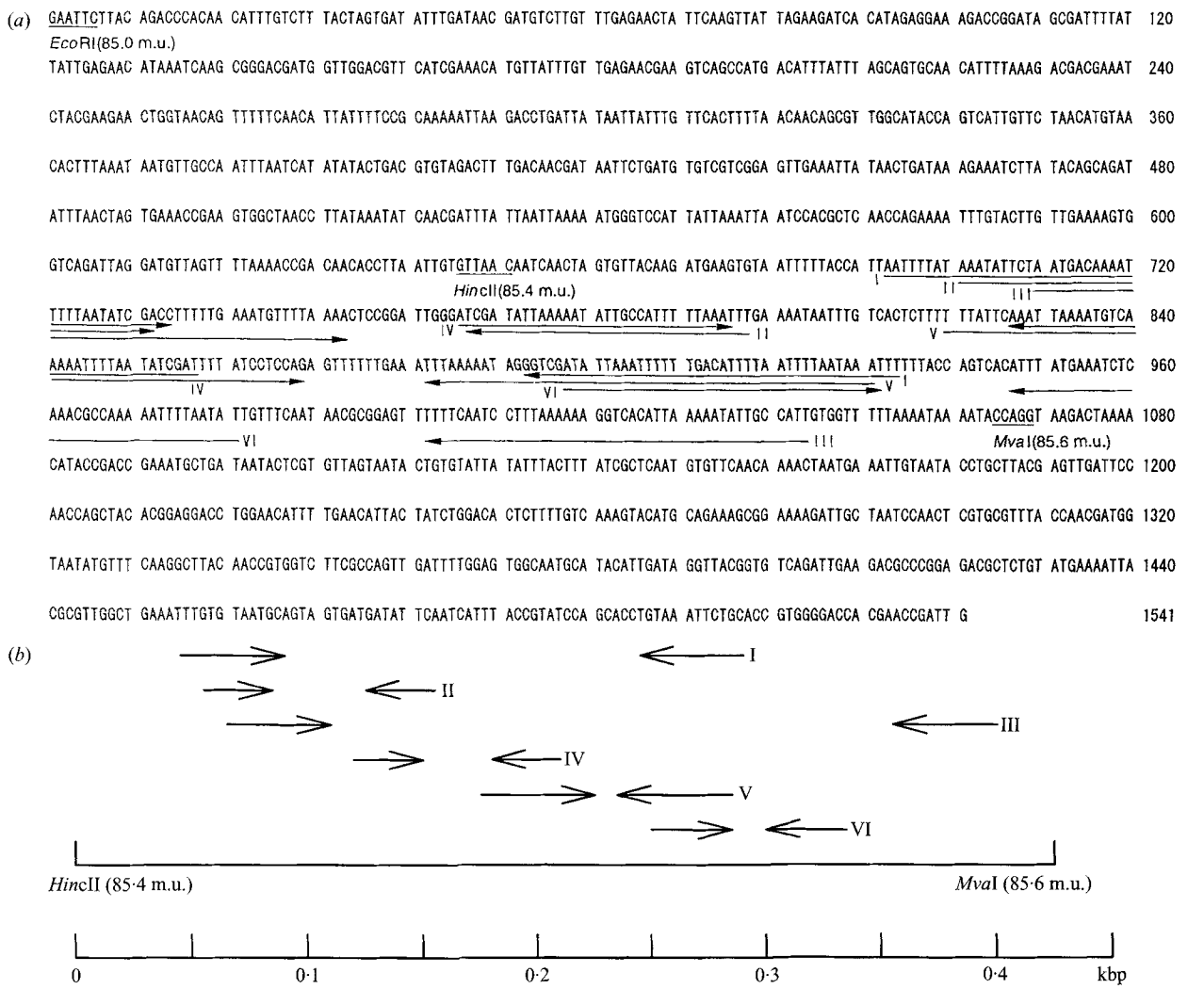

Fig. 7. Nucleotide sequence of the leftmost 1541 bp of *Eco*RI-R (85·0–86·0 m.u.) and the arms of six inverted repeats, I–VI, in the *irs* indicated by arrows along the sequence (*a*), and a schematic representation of the six inverted repeats in the *irs* (*b*).

region overlapping ORF2 in AcMNPV *Eco*RI-I. This region is not reiterated elsewhere in AcMNPV DNA, and appears to correspond to *ihs*1 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*s. ORFs were also detected within BmNPV *ihs*s. Ahrens & Rohrmann (1995) recently revealed that the *Hin*dIII-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*s 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*s are located close to *hr*1, *hr*2 and *hr*3 on the genome, we assumed that *hr* homologues of TnGV might be present in regions close to the *ihs*s; therefore we sequenced several restriction DNA fragments in the plasmid clones. When the nucleotide sequences of the *Eco*RI-P, R fragments were examined, a region of nearly 400 bp within *Eco*RI-R having multiple overlapping imperfect inverted repeats was found to the left of *ihs*3 and was designated an *irs*. The nucleotide sequence leftmost of *Eco*RI-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 *irs* were present elsewhere in the genomes of TnGV or BmNPV, *Eco*RI-R was hybridized to blots of genomic restriction digests of TnGV and BmNPV DNA. No hybridization signals were observed other than in TnGV *Eco*RI-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 > 80·00) 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 *hr*s. The *hr*s 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 *hr*s TnGV *irs* has neither DNA sequence identity nor replication activity in a heterologous replication system (Pearson *et al.*, 1992; Kool *et al.*, 1993*a*, *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*s 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*s 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

AHRENS C. H. & ROHRMANN, G. F. (1995). Identification of essential *trans*-acting regions required for DNA replication of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus: *lef-1* is an essential replication gene. *Virology* **27**, 417–428.

AHRENS, C. H., PEARSON, M. N. & ROHRMANN, G. F. (1995). Identification and characterization of a second putative origin of DNA replication in a baculovirus of *Orgyia pseudotsugata*. *Virology* **207**, 572–576.

AKIYOSHI, D., CHAKERIAN, R., ROHRMANN, G. F., NESSON, M. H. & BEAUDREAU, G. S. (1985). Cloning and sequencing of the granulin gene from the *Trichoplusia ni* granulosis virus. *Virology* **141**, 328–332.

ARIF, B. M. & DOERFLER, W. (1984). Identification and location of reiterated sequences in the *Choristoneura fumiferana* MNPV genome. *EMBO Journal* **3**, 525–529.

ARIF, B. M., GUANGYU, Z. & JAMIESON, P. (1986). A comparison of three granulosis viruses isolated from *Choristoneura* spp. *Journal of Invertebrate Pathology* **48**, 180–186.

AYRES, M. D., HOWARD, S. C., KUZIO, J., LOPEZ-FERBER, M. & POSSEE, R. D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**, 586–605.

BURGES, S. (1983). *Eco*RI restriction endonuclease fragment patterns of eight lepidopteran baculoviruses. *Journal of Invertebrate Pathology* **42**, 401–404.

CARSON, D. D., SUMMERS, M. D. & GUARINO, L. A. (1991). Transient expression of the *Autographa californica* nuclear polyhedrosis virus immediate-early gene, IE-N, is regulated by three viral elements. *Journal of Virology* **65**, 945–951.

CHAKERIAN, R., ROHRMANN, G. F., NESSON, M. H., LEISY, D. J. & BEAUDREAU, G. S. (1985). The nucleotide sequence of the *Pieris brassicae* granulosis virus granulin gene. *Journal of General Virology* **66**, 1263–1269.

COCHRAN, M. A. & FAULKNER, P. (1983). Location of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *Journal of Virology* **45**, 961–970.

CORSARO, B. G., GIJZEN, M., WANG, P. & GRANADOS, R. R. (1993). Baculovirus enhancing proteins as determinants of viral pathogenesis. In *Parasites and Pathogens of Insects*, vol. 2, pp. 127–145. Edited by N. Beckage, S. N. Thompson & B. A. Federici. San Diego: Academic Press.

CROOK, N. E. (1981). A comparison of the granulosis viruses from *Pieris brassicae* and *Pieris rapae*. *Virology* **115**, 173–181.

CROOK, N. E. (1986). Restriction enzyme analysis of granulosis viruses isolated from *Artogeia rapae* and *Pieris brassicae*. *Journal of General Virology* **67**, 781–787.

CROOK, N. E., SPENCER, R. A., PAYNE, C. C. & LEISY, D. J. (1985). Variation in *Cydia pomonella* granulosis virus isolates and physical maps of the DNA from three variants. *Journal of General Virology* **66**, 2423–2430.

CROOK, N. E., CLEM, R. J. & MILLER, L. K. (1993). An apoptosis-inhibiting gene with a zinc finger-like motif. *Journal of Virology* **767**, 2168–2174.

DWYER, K. G. & GRANADOS, R. R. (1987). A physical map of the *Pieris rapae* granulosis virus genome. *Journal of General Virology* **68**, 1471–1476.

DWYER, K. G. & GRANADOS, R. R. (1988). Mapping *Pieris rapae* granulosis virus transcripts and their *in vitro* translation products. *Journal of Virology* **62**, 1535–1542.

EASWARAMOORTHY, S. & CORY, J. S. (1990). Characterization of the DNA of granulosis viruses isolated from two closely related moths, *Chilo infuscatellus* and *C. saccariphagus indicus*. *Archives of Virology* **110**, 113–119.

FRANCKI, R. I. B., FAUQUET, C. M., KNUDSON, D. L. K. & BROWN, F. (editors) (1991). Classification and Nomenclature of Viruses. Fifth Report of the International Committee on Taxonomy of Viruses. *Archives of Virology*, Supplement 2.

FUNK, C. N. & CONSIGLI, R. A. (1992). Evidence for zinc binding by two structural proteins of *Plodia interpunctella* granulosis virus. *Journal of Virology* **66**, 3168–3171.

GOTO, C., MINOBE, Y. & IIZUKA, T. (1992). Restriction endonuclease analysis and mapping of the genomes of granulosis viruses isolated from *Xestia c-nigrum* and five other noctuid species. *Journal of General Virology* **73**, 1491–1497.

GUARINO, L. A. & SUMMERS, M. D. (1986). Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. *Journal of Virology* **60**, 215–223.

GUARINO, L. A., GONZALEZ, M. A. & SUMMERS, M. D. (1986). Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* **60**, 224–229.

HARVEY, J. P. & VOLKMAN, L. E. (1983). Biochemical and biological variation of *Cydia pomonella* (codling moth) granulosis virus. *Virology* **124**, 2134.

HARVEY, J. & TANADA, Y. (1985). Characterization of the DNAs of five baculoviruses pathogenic for the armyworm, *Pseudaletia unipuncta*. *Journal of Invertebrate Pathology* **46**, 174–179.

HASHIMOTO, Y., CORSARO, B. G. & GRANADOS, R. R. (1991). Location and nucleotide sequence of the gene encoding the viral enhancing factor of the *Trichoplusia ni* granulosis virus. *Journal of General Virology* **72**, 2645–2651.

HASHIMOTO, Y., KANAMORI, Y., HAYAKAWA, Y., KAMITA, S. G., MAEDA, S. & MATSUMOTO, T. (1994). Physical mapping of *Bombyx mori* nuclear polyhedrosis virus strain D1: identification of novel interspersed homologous regions. *Journal of Applied Entomology and Zoology* **29**, 442–447.

HUBER, J. & HUGHES, P. R. (1984). Quantitative bioassay in insect pathology. *Bulletin of the Entomological Society of America* **30**, 31–34.

JEHLE, J. A. & BACKHAUS, H. (1994*a*). Genome organization of the DNA-binding protein gene region of *Cryptophlebia leucotreta* granulosis virus is closely related to that of nuclear polyhedrosis viruses. *Journal of General Virology* **75**, 1815–1820.

JEHLE, J. A. & BACKHAUS, H. (1994*b*). The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylogenetic considerations. *Journal of General Virology* **75**, 3667–3671.

JEHLE, J. A., BACKHAUS, H., FRITSCH, E. & HUBER, J. (1992). Physical map of the *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. *Journal of General Virology* **73**, 1621–1626.

KOOL, M., VAN DEN BERG, P. M. M. M., TRAMPER, J., GOLDBACH, R. W. & VLAK, J. M. (1993*a*). Location of two putative origins of DNA replication of *Autographa californica* nuclear polyhedrosis virus. *Virology* **192**, 94–101.

KOOL, M., VOETEN, J. T. M., GOLDBACH, R. W. & VLAK, J. M. (1993*b*). Identification of seven putative origins of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus DNA replication. *Journal of General Virology* **74**, 2661–2668.

KOOL, M., VOETEN, J. T. M., GOLDBACH, R. W. & VLAK, J. M. (1994). Functional mapping of regions of the *Autographa californica* nuclear polyhedrosis viral genome required for DNA replication. *Virology* **198**, 680–689.

KUZIO, J. & FAULKNER, P. (1984). Regions of repeated DNA in the genome of *Choristoneura fumiferana* nuclear polyhedrosis virus. *Virology* **139**, 185–188.

LEISY, D. J. & ROHRMANN, G. F. (1993). Characterization of the replication of plasmids containing *hr* sequences in baculovirus-infected *Spodoptera frugiperda* cells. *Virology* **196**, 722–730.

MAEDA, S. & MAJIMA, K. (1990). Molecular cloning and physical mapping of the genome of *Bombyx mori* nuclear polyhedrosis virus. *Journal of General Virology* **71**, 1851–1855.

MAJIMA, K., KOBARA, R. & MAEDA, S. (1993). Divergence and evolution of homologous regions of *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology* **67**, 7513–7521.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory.

NASER, W. L., MILTENBURGER, J. P., HARVEY, J. P. & HUGER, A. M. (1984). *In vitro* replication of the *Cydia pomonella* (codling moth) granulosis virus. *FEMS Microbiology Letters* **24**, 117–121.

PEARSON, M. N. & ROHRMANN, G. F. (1995). *Lymantria dispar* nuclear polyhedrosis virus homologous regions: characterization of their ability to function as replication origins. *Journal of Virology* **69**, 213–221.

PEARSON, M., BJORNSON, R., PEARSON, G. & ROHRMANN, G. (1992). The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* **257**, 1382–1384.

POSSEE, R. D., SUN, T.-P., HOWARD, S. C., AYRES, M. D., HILL-PERKINS, M. & GEARING, K. L. (1991). Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis 9·4 kbp *Eco*RI-I and -R (polyhedrin gene) region. *Virology* **185**, 229–241.

RODEMS, S. M. & FRIESEN, P. D. (1993). The *hr5* transcriptional enhancer stimulates early expression from the *Autographa californica* nuclear polyhedrosis virus genome but is not required for virus replication. *Journal of Virology* **67**, 5776–5785.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, USA* **74**, 5463–5467.

SMITH, G. E. & SUMMERS, M. D. (1978). Analysis of baculovirus genomes with restriction endonucleases. *Virology* **89**, 517–527.

SMITH, G. E. & SUMMERS, M. D. (1981). Application of a novel radioimmunoassay to identify baculovirus structural proteins that share interspecies antigenic determinants. *Virology* **39**, 125–137.

SMITH, I. R. L. & CROOK, N. E. (1988*a*). Physical maps of the genomes of four variants of *Artogeia rapae* granulosis virus. *Journal of General Virology* **69**, 1741–1747.

SMITH, I. R. L. & CROOK, N. E. (1988*b*). *In vivo* isolation of baculovirus genotypes. *Virology* **166**, 240–244.

SMITH, I. R. L. & CROOK, N. E. (1993). Characterization of new baculovirus genotypes arising from inoculation of *Pieris brassicae* with granulosis viruses. *Journal of General Virology* **74**, 415–424.

SMITH, I. R. L., VAN BEEK, N. A. M., PODGWAITE, J. D. & WOOD, H. A. (1988). Physical map and polyhedrin gene sequence of *Lymantria dispar* nuclear polyhedrosis virus. *Gene* **71**, 97–105.

TANADA, Y. & HESS, R. T. (1991). Baculoviridae. Granulosis Viruses. In *Atlas of Invertebrate Viruses*, pp. 227–257. Edited by J. R. Adams & J. R. Bonami. Boca Raton: CRC Press.

THEILMANN, D. A. & STEWART, S. (1992). Tandemly repeated sequence at the 3′ end of the IE-2 gene of the baculovirus *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus in an enhancer element. *Virology* **187**, 97–106.

TWEETEN, K. A., BULLA, L. A. & CONSIGLI, R. R. (1980). Restriction enzyme analysis of the genomes of *Plodia interpunctella* and *Pieris rapae* granulosis virus. *Virology* **104**, 514–519.

VLAK, J. M. & SMITH, G. E. (1982). Orientation of the genome of

Autographa californica nuclear polyhedrosis virus: a proposal. *Journal of Virology* **41**, 1118–1121.

WINSTANLEY, D. & CROOK, N. E. (1993). Replication of *Cydia pomonella* granulosis virus in cell cultures. *Journal of General Virology* **74**, 1599–1609.

XIE, W.-D., ARIF, B., DOBOS, P. & KRELL, P. J. (1995). Identification and analysis of a putative origin of DNA replication in the *Choristoneura fumiferana* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* **209**, 409–419.

(*Received 27 June 1995; Accepted 20 October 1995*)