Location and Nucleotide Sequence of the *Orgyia pseudotsugata* Single Nucleocapsid Nuclear Polyhedrosis Virus Polyhedrin Gene

By D. Leisy, M. Nesson, M. Pearson, G. Rohrmann* and G. Beaudreau

Department of Agricultural Chemistry, Oregon State University, Corvallis, Oregon 97331, U.S.A.

(Accepted 12 February 1986)

**SUMMARY**

A restriction endonuclease map was determined for the *Orgyia pseudotsugata* single nucleocapsid nuclear polyhedrosis virus (SNPV) genome. The order of the fragments generated by the enzymes *BglII*, *BamHI* and *XbaI* was analyzed using double digestion of the total genome and digestion of isolated restriction fragments. The location of the polyhedrin gene was then determined using a cloned polyhedrin gene from the *O. pseudotsugata* multiple nucleocapsid NPV (MNPV) as a hybridization probe. A fragment containing this gene was cloned, mapped, subcloned and the nucleotide sequence of a 1.3 kb fragment was determined which contained the entire polyhedrin reading frame and some flanking sequences. This gene demonstrated 76% nucleotide sequence homology and 87% amino acid sequence homology to the *Autographa californica* MNPV polyhedrin sequence. A probable regulatory element was identified which is common to the 5' flanking region of all hypertranscribed late genes (polyhedrin and 10K proteins) which have been examined in baculoviruses.

**INTRODUCTION**

Two nuclear polyhedrosis viruses (NPVs) have been isolated from the Douglas fir tussock moth (*Orgyia pseudotsugata*) (Hughes & Addison, 1970). The OpSNPV has a single nucleocapsid per envelope whereas OpMNPV has multiple nucleocapsids per envelope. OpMNPV is closely related to the MNPV of *Autographa californica* (AcMNPV), and DNA–DNA hybridization has demonstrated substantial genomic sequence homology (Rohrmann et al., 1982b) and colinear genome organizations (Leisy et al., 1984). In contrast, OpSNPV shares only limited DNA sequence homology with either OpMNPV or AcMNPV (Rohrmann et al., 1982b). Also, the 162 kbp genome of OpSNPV (Schafer et al., 1979) is 25% larger than those of OpMNPV and AcMNPV (each 130 kbp). Although OpSNPV and OpMNPV demonstrate only limited genomic sequence homology, their polyhedrins, the major occlusion body proteins, are closely related; they cannot be readily distinguished with polyclonal antisera by radioimmunoassay (Rohrmann et al., 1981) or immunodiffusion (Rohrmann, 1977). Differences do exist, however, since they can be distinguished with monoclonal antibodies (Quant et al., 1984). In this report, we present a restriction map for the OpSNPV genome and describe the location, cloning and nucleotide sequence of the OpSNPV polyhedrin gene.

**METHODS**

Reagents, DNA sequencing, nick translation and Southern blots. All enzymes (restriction endonucleases, *Escherichia coli* DNA polymerase, T4 DNA ligase, Klenow fragment of *E. coli* DNA polymerase I) were purchased from Bethesda Research Laboratories (BRL), New England Biolabs or P-L Pharmacia and the buffers and protocols recommended by the manufacturers were used. Nick translation was done according to Maniatis et al. (1982). Cross-hybridization to Southern blots of heterologous DNAs was done under non-stringent conditions as previously described (Rohrmann et al., 1982b). M13 sequencing was done according to the methods of Sanger et al. (1980) and the Sequencing Manual supplied by BRL. using [32P]dATP. All radioactive nucleotides were purchased from New England Nuclear.
Restriction endonuclease mapping of the OpSNPV genome. To produce a restriction map, the viral genomic DNA was digested with a restriction enzyme and the fragments were separated on a 10 to 30% (w/v) sucrose gradient following the method described by De Fillipes (1982). Small fractions (five or ten drops) were collected directly in a 96 well microtitre plate. An aliquot of each fraction was electrophoresed in an agarose gel to allow identification of each restriction fragment. Where appropriate, these fractions were pooled, extracted with phenol, ethanol-precipitated and resuspended in TE buffer (10 mM-Tris-HCl pH 8.0, 1 mM-EDTA). Each isolated restriction fragment was then digested with a second restriction enzyme and these digests were electrophoresed alongside appropriate genomic double digests to allow identification of the equivalent fragments. Reciprocal double digestions were performed similarly.

RESULTS AND DISCUSSION

Genome map and location of the polyhedrin gene

A restriction endonuclease map of the OpSNPV genome was produced using reciprocal double digestion with the enzymes BamHI, BglII and XbaI which produced six, five and 11 fragments respectively (Fig. 1). These three enzymes produced fragment profiles which were conducive to the separation and isolation of most fragments and allowed the production of a genomic map in which all regions were represented by fragments comprising 12% or less of the

![Fig. 1](image-url)
OpSNPV polyhedrin gene sequence

Table 1. Sizes* of OpSNPV restriction fragments (kbp)

<table>
<thead>
<tr>
<th>Single digests</th>
<th>Double digests</th>
<th>Triple digest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment</td>
<td>BamHI</td>
<td>BglII</td>
</tr>
<tr>
<td>A</td>
<td>52.5</td>
<td>59.7</td>
</tr>
<tr>
<td>B</td>
<td>35.9</td>
<td>34.3</td>
</tr>
<tr>
<td>C</td>
<td>26.3</td>
<td>27.7</td>
</tr>
<tr>
<td>D</td>
<td>23.7</td>
<td>28.2</td>
</tr>
<tr>
<td>E</td>
<td>12.3</td>
<td>11.3</td>
</tr>
<tr>
<td>F</td>
<td>10.6</td>
<td>11.9</td>
</tr>
<tr>
<td>G</td>
<td>10.3</td>
<td>7</td>
</tr>
<tr>
<td>H</td>
<td>9.6</td>
<td>8</td>
</tr>
<tr>
<td>I</td>
<td>8.0</td>
<td>9</td>
</tr>
<tr>
<td>J</td>
<td>6.6</td>
<td>10</td>
</tr>
<tr>
<td>K</td>
<td>4.8</td>
<td>11</td>
</tr>
<tr>
<td>L</td>
<td>12</td>
<td>4.8</td>
</tr>
<tr>
<td>M</td>
<td>13</td>
<td>4.5</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>3.9</td>
</tr>
<tr>
<td>O</td>
<td>15</td>
<td>2.8</td>
</tr>
<tr>
<td>P</td>
<td>16</td>
<td>1.8</td>
</tr>
<tr>
<td>Q</td>
<td>17</td>
<td>0.9</td>
</tr>
<tr>
<td>R</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

Total fragment size 161.3

* All fragment size values listed are derived from the sizes determined from gels of the triple digest, run with λ HindII and pBR322 HindI DNA digests as markers.

d genome. The total size of all fragments combined (161.3 kbp) (Table 1) was close to that (160 kbp) estimated previously by electron microscopy (Schafer et al., 1979).

Ideally, by digesting genomic DNA with one enzyme, isolating the fragments, digesting these fragments with a second enzyme, analysing the products on a gel, and then performing the reciprocal set reactions, a physical map of the genome can be deduced. However, in practice two problems arise: difficulty in resolving two or more fragments of similar size from the sucrose gradient and the inability to order internal fragments. Pairs of fragments imperfectly separated in gradient fractions could often be characterized by the performance of the second digestion on each of the gradient fractions spanning the region of overlap and running the resultant digests in adjacent lanes in an agarose gel. By comparing sets of double digests, fragment distributions were often observed to differ sufficiently to allow identification of their source. In this study, we were able to resolve all such ambiguities and produce a complete map by analysis of the fragments produced by each of the double or triple digests (Fig. 2).

In order to begin orientating the OpSNPV genome relative to the well characterized OpMNPV and AcMNPV genomes, we hybridized probes containing the AcMNPV 39K (Smith et al., 1982) and 10K (Kuzio et al., 1984) genes and the OpMNPV polyhedrin gene to blots of OpSNPV genomic DNA. A BglII restriction fragment of 11.3 kbp was identified which cross-hybridized with a recombinant plasmid containing the gene for the AcMNPV 39K protein (AcMNPV PstI K fragment) (Fig. 2). In addition, an attempt was made to locate a sequence homologous to the 10K gene of AcMNPV using an AcMNPV HindIII Q clone as a hybridization probe but no hybridization was detected. This indicates that the 10K protein has diverged much more rapidly than the 39K or polyhedrin genes (see below).

We have previously demonstrated substantial amino acid homology at the amino termini and considerable immunological cross-reactivity between polyhedrins from OpMNPV and OpSNPV (Rohrmann et al., 1981). Using a cloned polyhedrin gene from OpMNPV (Rohrmann et al., 1982a) as a hybridization probe to screen Southern blots of several OpSNPV restriction digests, the OpSNPV polyhedrin gene was located on BamHI–XbaI fragment 3 and BglII–XbaI.
Fig. 2. Restriction endonuclease map of OpSNPV DNA for BglII, BamHI and XbaI. The circular DNA of 161.3 kbp is presented in linear form with the polyhedrin gene at the zero point. Letters and numbers correspond to fragments described in Table 1. A BglII fragment having sequence homology to the AcMNPV 39K gene is indicated. Two EcoRI sites (E) bracketing the polyhedrin (PY) gene are indicated in the restriction site map, but no other EcoRI sites were mapped. Other restriction sites on this map are: G, BglII; X, XbaI and M, BamHI.

Fig. 3. Restriction map of the region containing the OpSNPV polyhedrin gene. The upper map is of the 3.3 kbp EcoRI fragment which was cloned into pBR325. The lower map is a detailed restriction map of the 1.3 kbp Smal–HindIII fragment. The thickened line indicates the polyhedrin reading frame. The arrows above the lower map indicate the location and direction of other potential open reading frames. The arrows below indicate the sequencing strategy. Cloning into plasmids and M13 was done according to Maniatis et al. (1982). Three sets of sequencing reactions were carried out using synthetic primers synthesized by an Applied Biosystems 380A DNA synthesizer and were used to confirm the regions identified by the arrows labelled 1, 2 and 3.

In addition, the polyhedrin gene was localized to a 3.3 kbp EcoRI fragment and cloned into pBR325. Sequencing of a subclone of this fragment which contained the polyhedrin gene (see below) revealed a BglII site a short distance 3' to the carboxy terminus. These results provided the polyhedrin gene orientation on the BglII–XbaI fragment (Fig. 2).
**OpSNPV polyhedrin gene sequence**

**SmaI**

```
CCCGGGTATATCGTTTIGACCACITGCAACGTGATCGCGGGCGGGTCTACAAACCGCTCGTCTAGCCCCA
TAAGACAGCCGGAATTGGAAGIATCITCGTCATATCCATTATTGTCACACGTGTTACTACTCTCG
GAAATGTCAATAATACTAATGCGCGTTATCTCAAACTACTGTAATAATATTGTGTCGATATCCGCTGAG
GTATTCGATTTTCGCTGAGAATTTCACCACACCTCAGAAATTTTTTCTCTTCTGTAAAAACAT
1GTGAATTCTCCATACCCATATTGTACCTGTAACAGCTACAGTTACCTGCATTACGCTATTACTAT

<table>
<thead>
<tr>
<th>Met</th>
<th>Tyr</th>
<th>Thr</th>
<th>Arg</th>
<th>Tyr</th>
<th>Ser</th>
<th>Tyr</th>
<th>Asn</th>
<th>Pro</th>
<th>Ser</th>
<th>Leu</th>
<th>Gly</th>
<th>Arg</th>
<th>Thr</th>
<th>Tyr</th>
<th>Val</th>
</tr>
</thead>
</table>

Fig. 4. Nucleotide sequence of the OpSNPV polyhedrin gene coding and flanking regions. Underlined, starting at nt 321, is a sequence of 12 nt which is highly conserved in the 5' regions of all baculovirus hypertranscribed late genes so far examined. Also underlined, starting at nt 1127, is a potential polyadenylation signal in the 3' flanking region.

**Polyhedrin sequence**

The 3.3 kbp cloned EcoRI fragment containing the polyhedrin gene was mapped with restriction enzymes by the analysis of double digests (Fig. 3). Further analysis revealed that cross-hybridization with the OpMNPV polyhedrin probe was limited to a 1.3 kbp HindIII-Smal fragment within the 3.3 kbp EcoRI fragment (data not shown). This smaller fragment was cloned into M13 mp8 and mapped with restriction endonucleases. A variety of subfragments were cloned into M13 vectors (Sanger *et al.*, 1980) for DNA sequence analysis. The subclones and sequencing strategy are indicated in Fig. 3.

After sequencing of the SmaI–HindIII fragment, a computerized search was used to locate open reading frames greater than 100 nucleotides (nt). A large open reading frame was
Fig. 5. Comparison of the OpSNPV polyhedrin amino acid sequence with the sequence of the AcMNPV polyhedrin (Hooft van Iddekinge et al., 1983). The dots indicate homology of the OpSNPV sequence with that of AcMNPV. Standard single letter abbreviations are used for the amino acids.

identified from nt 374 to nt 1111 (Fig. 4) which demonstrated an N-terminal amino acid sequence homologous to that which we previously determined by N-terminal sequence analysis of the OpSNPV polyhedrin protein (Rohrmann et al., 1979). Three additional open reading frames all of which spanned the SmaI site were also found (see Fig. 3).

The OpSNPV polyhedrin gene demonstrates 76% nucleotide and 87% amino acid (Fig. 5) sequence homology to AcMNPV polyhedrin. Most striking is the conservation of 12 nt in the 5' flanking sequence located about 50 nt upstream of the initiation codon (Fig. 4). This sequence appears to be conserved throughout the Baculoviridae: related sequences have been described upstream of granulin genes (Akiyoshi et al., 1985; Chakerian et al., 1985), AcMNPV 10K (Kuzio et al., 1984) and OpMNPV 10K (D. Leisy et al., unpublished observations) protein genes and the polyhedrin genes of AcMNPV (Hooft van Iddekinge et al., 1983), Bombyx mori NPV (BmNPV) (Iatrou et al., 1985) and OpMNPV (D. Leisy et al., unpublished observations). All hypertranscribed baculovirus genes examined (polyhedrin and 10K), have mRNA initiation sites at or very near this conserved sequence (Hooft van Iddekinge et al., 1983; Kuzio et al., 1984; D. Leisy et al., unpublished observations). No other sequence appears to be regularly conserved in the 5' flanking region and this sequence may therefore play a prominent role in the regulation of either transcription or translation of these hypertranscribed genes.

A computerized comparison of the 3' and 5' (above nt 320) flanking regions of the OpSNPV polyhedrin gene revealed no homologies with the reported flanking regions of the BmNPV polyhedrin gene (Iatrou et al., 1985) or the OpMNPV polyhedrin gene (D. Leisy et al., unpublished observations). This suggests either a more rapid rate of evolution of these sequences relative to polyhedrin or the occurrence of genomic rearrangement during the evolutionary divergence of these viruses (such that the flanking sequences do not correspond to one another).

From our comparison of polyhedrin sequences, it is not clear whether SNPVs and MNPVs represent true divergent evolutionary lines (Rohrmann, 1986). However, sequencing a more rapidly evolving gene may provide such evidence. SNPVs are found in a number of insect orders, whereas the presence of MNPVs appears to be confined to the Lepidoptera (Vlak & Rohrmann, 1985). The evolutionary advantage of SNPV compared to MNPV morphotypes is unclear and some lepidopteran NVPs demonstrate differing degrees of aggregation of their nucleocapsids. For example, BmNPV nucleocapsids appear singly or in aggregations of up to five per envelope (Iatrou et al., 1985). In addition, the presence of multiple nucleocapsids per envelope of the granulosis virus of Cydia pomonella has been reported (Falcon & Hess, 1985). The latter report indicated that the proximity of virus during assembly may influence the extent of aggregation and the authors suggested that the formation of multiple nucleocapsids per envelope may be influenced by host or environmental factors. A genetic explanation has also been proposed. Rohrmann (1986) suggested that the occurrence of MNPVs may have been selected for during evolution because they infect cells as multiple nucleocapsids (Granados &
Lawler, 1981); because of this, virions which had been inactivated by u.v. light or other phenomena might be enabled to survive via complementation. DNA recombination leading to a complete functional genome may occur under such situations.

We thank P. Faulkner and Gale Smith for the AcMNPV 10K and 39K clones respectively. This project was supported by NIH grant AI 21973. This is technical report No. 7710 from the Oregon State Experiment Station.

REFERENCES


(Received 22 November 1985)