Annotation of the *Hyposoter fugitivus* ichnovirus (HfIV) genome

Kohjiro Tanaka(1), Walter E. Barney(2), Andrea M. Makkay(2), Don B. Stoltz(2) and Bruce A. Webb(3)

(1) University of Kentucky, Department of Entomology, S-225 Agricultural Science Center North, Lexington, KY, (2) Dalhousie University, Department of Microbiology and Immunology, Halifax, Nova Scotia, Canada

Introduction

Certain endoparasitoid wasps carry polydnaviruses (PDV) which have genomes comprised of multiple circular segments of double-stranded DNA. These polydnaviruses induce physiological and immunological alterations in lepidopteran insects which enable wasp progeny to survive in the insect hemocoel. To reveal the polydnavirus genomes that produce these alterations, we sequenced the *Hyposoter fugitivus* ichnovirus (HfIV) genome. The HfIV genome is 279 kb in size and organized into 71 segments ranging in size from 2.5 to 8.8 kb. 151 potential open reading frames are identified with 93 of these ORFs homologous to other polydnaviruses. Six gene families were identified: repeat element genes, cysteine-motif genes, viral innoxins, viral ankryns, N genes, and polar residue rich proteins. In this study, we compare features of the HfIV genome and contents of gene family with those of 2 ichnoviruses (*Cotesia congregata* (CsIV) (1), *Cotesia malesiana* (TrIV) (2)).

Materials and Methods

**PDV genomic DNA preparation**

*Hyposoter fugitivus* polydnavirus was centrifuged by sucrose density gradients as described previously (3, 4). The viral genomic DNA was isolated from purified viruses by incubating with Proteinase K and Sarkosyl at 50°C for 2 hours and then extracted with phenol-chloroform and with chloroform prior to ethanol precipitation.

**Sequencing, assembly and analysis**

We constructed 10 libraries using a shotgun cloning approach. Nine were made by randomly transposon inserting into viral DNA segments. They are roughly ordered from shortest to longest segment. On the left side of each gray bar is the segment name. Predicted ORFs are represented by the colored area with each gene family as indicated in the box.

**Results and Discussions**

Naming of HfIV genome segments

The most conventional method for naming polydnavirus genome segments has been to name them alphabetically from shortest to longest. This method has worked well for genomes with less than 26 segments and with great diversity in segment length. However, in our case we sequenced 71 segments of HfIV with little variation in segment length (Fig. 2). Therefore, we decided to use an abbreviation for representing length of segments corresponding to an alphabetical listing. That is each letter indicates all segments within 1kb in length, starting with segment A at 2.000 kb. For example, all segments from B1-B26 are between 3.000 and 3.999 kb.

**Genomic features of HfIV**

Seven one segments were sequenced and assembled giving a total of 279.302bp. The genome has a G+C content of 44%. Segments are distributed from 2,500 to 8,851bp in size. Most of the segments (75%) are between 3 and 5 kb in length. The average segment size of HfIV is 3.9kb (Table 1). Compared with other ichnoviruses, CdIV and TrIV, and other bracoviruses, MdBV and CcBV, this genome is similar to CdIV, though segment size and number of segments are quite different. The G+C contents of HfIV is similar with these five polydnaviruses in that they all show a 3% variation between them.

**References**


![Figure 2](https://via.placeholder.com/150)

**Figure 2. Schematic representation of HfIV genomic DNA segments. Each gray bar represents a viral genomic DNA segment. They are roughly ordered from shortest to longest segment. On the left side of each gray bar is the segment name. Predicted ORFs are represented by the colored area with each gene family as indicated in the box.**

**Table 1. Comparison of genomic features among five polydnaviruses.**

<table>
<thead>
<tr>
<th>Species</th>
<th>Segments (No)</th>
<th>Segment size (kb)</th>
<th>Average segment Length (kb)</th>
<th>G+C Contents (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HfIV</td>
<td>71</td>
<td>2.5 - 8.8</td>
<td>3.9</td>
<td>279</td>
</tr>
<tr>
<td>CdIV</td>
<td>24</td>
<td>6.1 - 19.6</td>
<td>10.3</td>
<td>247</td>
</tr>
<tr>
<td>TrIV</td>
<td>20</td>
<td>2.3 - 10.1</td>
<td>6.05</td>
<td>121</td>
</tr>
<tr>
<td>MdBV</td>
<td>15</td>
<td>3.6 - 30.0</td>
<td>12.4</td>
<td>187</td>
</tr>
<tr>
<td>CcBV</td>
<td>30</td>
<td>4.9 - 11.5</td>
<td>8.18</td>
<td>568</td>
</tr>
</tbody>
</table>

* TdIV is a summary of 20 completely sequenced segments. 

* : For the CcBV, the method to define and find ORFs were different from the others. So comparisons of the number of unsassigned and other CcBV unique genes are based on those genes having significant similarities with other genes in the databases.

**Figure 3. Comparison of gene family compositions in five polydnaviruses (three ichnoviruses: HfIV, CdIV, TrIV and two bracovirus: MdBV and CcBV). Each colored and patterned box represents gene families as indicated above. Numbers in the bars represent amount of genes. Total ORFs of each species are represented above the bars.