Conserved motifs in the invertebrate iridescent virus 6 (IIV6) genome regulate virus transcription

Journal of Invertebrate Pathology
Yesilyurt, Aydin; Demirbag, Zihni; Oers, Monique M.; Nalcacioglu, Remziye
https://doi.org/10.1016/j.jip.2020.107496

This article is made publicly available in the institutional repository of Wageningen University and Research, under the terms of article 25fa of the Dutch Copyright Act, also known as the Amendment Taverne. This has been done with explicit consent by the author.

Article 25fa states that the author of a short scientific work funded either wholly or partially by Dutch public funds is entitled to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' project. In this project research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and / or copyright owner(s) of this work. Any use of the publication or parts of it other than authorised under article 25fa of the Dutch Copyright act is prohibited. Wageningen University & Research and the author(s) of this publication shall not be held responsible or liable for any damages resulting from your (re)use of this publication.

For questions regarding the public availability of this article please contact openscience.library@wur.nl
Conserved motifs in the invertebrate iridescent virus 6 (IIV6) genome regulate virus transcription

Aydin Yesilyurt a, b, Zihni Demirbag a, Monique M. van Oers b, Remziye Nalcacioglu a, * 

a Department of Biology, Faculty of Science, Karadeniz Technical University, 61080 Trabzon, Turkey  
b Laboratory of Virology, Wageningen University and Research, 6708 PB Wageningen, the Netherlands

**A R T I C L E   I N F O**

Keywords:  
Transcriptional analysis  
Promoter sequence  
Repressor  
Iridovirus

**A B S T R A C T**

Invertebrate iridescent virus 6 (IIV6) is the type species of the Iridoviridae genus in the Betairidovirinae subfamily of the Iridoviridae family. Transcription of the 215 predicted IIV6 genes is temporally regulated, dividing the genes into three kinetic classes: immediate-early (IE), delayed-early (DE), and late (L). So far, the transcriptional class has been determined for a selection of virion protein genes and only for three genes the potential promoter regions have been analyzed in detail. In this study, we investigated the transcriptional class of all IIV6 genes that had not been classified until now. RT-PCR analysis of total RNA isolated from virus-infected insect cells in the presence or absence of protein and DNA synthesis inhibitors, placed 113, 23 and 22 of the newly analyzed viral ORFs into the IE, DE and L gene classes, respectively. Afterwards, in silico analysis was performed to the upstream regions (200 bp) of all viral ORFs using the MEME Suite Software. The AA(A/T)(T/A)TG(A/G)A and (T/A/C)(T/G) promoter sequences were identified in the upstream region of IE and DE genes, respectively. These motifs were validated by luciferase reporter assays as crucial sequences for promoter activity. For the L genes two conserved motifs were identified for all analyzed genes: (T/G)(C/T)(A/C)A(T/G/C)(T/C)T(T/C) and (C/G/T)(G/C)T(T/A)ATGG motifs were identified in the upstream region of IE and DE genes, respectively. Conversely, the presence of these two motifs did not influence promoter activity. Independant mutations in either motif caused a severe decrease in luciferase expression. Information on temporal classes and upstream regulatory sequences will contribute to our understanding of the transcriptional mechanisms in IIV6.

1. Introduction

Invertebrate iridescent viruses (IIVs, family Iridoviridae, subfamily Betairidovirinae, genus Iridovirus) form icosahedral particles of 120–180 nm in diameter (Chinchar et al., 2017). Virions consist of a DNA/protein core surrounded by an internal lipid membrane, a protein capsid and in the case of those particles that bud out of cells, an outer viral envelope. IIVs have been reported to infect over 100 species of arthropods (Williams et al., 2017). Invertebrate iridescent virus 6 (IIV-6), also known as Chilo iridescent virus (CIV), is the type species of the Iridoviridae genus. The IIV6 genome consists of 212,482 bp of linear dsDNA (Jakob and Darai, 2002) with 215 non-overlapping and putative protein-encoding ORFs selected from the 468 computationally predicted ORFs (Eaton et al., 2007). Proteomic analysis has shown that IIV6 particles contain 54 structural, viral-encoded proteins (Ince et al., 2010). The replication of the IIV6 genome is presumed to be essentially similar to that of Frog virus 3 (FV3), the type species of the genus Ranavirus, in the subfamily Alphairidovirinae (Granoff, 1984; Williams and Ward, 2010). Viral genome replication starts in the nucleus and is followed by genome concatamerization and subsequent cleavage, particle assembly and maturation in the cytoplasm (Goorha, 1982). Since purified IIV6 DNA is not infectious, one or more virion-associated proteins are needed for the initiation of IIV gene transcription (Cerutti et al., 1989).

A previous study on IIV6 mRNAs detectable by northern blot analysis revealed three temporal transcript classes in IIV6 infections: immediate-early (IE), delayed-early (DE) and late (L) (D’Costa et al., 2001). Thirty
eight of the detected transcripts were synthesized in the presence of protein synthesis inhibitors and were classified in the IE class; thirty four transcripts were produced in the presence of DNA synthesis inhibitors and were classified in the DE class, while 65 five transcripts were detected only in the absence of inhibitors and were classified in the L class. However, as the transcripts were classified prior to genome sequencing, the relationship between the ORFs and their temporal classification was not previously established. In a later study, the S4 IIV6 structural virion protein genes were analyzed for their temporal expression, showing that the majority of these were expressed as IE genes (Ince et al., 2013).

It is known that IIV6 transcripts possess generally short 5′ untranslated regions and lack poly A tails (Nalcacioglu et al., 2003). On the other hand, information regarding the promoter elements of IIV6 genes is rather limited. So far, potential promoter regions of only three IIV6 genes, exonuclease (012L, IE), DNApol (037L, DE) and major capsid protein gene (mcp) (274L, L), have been characterized in detail (Nalcacioglu et al., 2003; 2007; Dizman et al., 2012). These promoters have been identified by means of a luciferase reporter assay in conjunction with deletion mutagenesis of the sequences in the 5′ upstream region of the respective ORFs. In the current study, we investigated the transcriptional class of all as of yet unclassified IIV6 ORFs (170 transcripts) to compete the temporal classification and to be able to search for essential, conserved promoter motifs in IIV6 genes. Therefore, the upstream regions of all genes in a particular class (classified in this paper and in previous studies) were compared and analyzed for conserved sequence motifs. The identified conserved sequences were examined for promoter activity in insect cells using the luciferase reporter assay.

2. Material and methods

2.1. Cell line, virus and virus infections

Spodoptera frugiperda 9 (S9) cells were maintained in SF-900 II SFM (Gibco) supplemented with 5% fetal bovine serum (FBS, Sigma) at 28 °C as monolayer. Invertebrate iridescent virus 6 (IIV6) was propagated in these cells and the virus titer was determined in End Point Dilution Assays (EPDAs) (Cook et al., 1976). Virus infections were carried out with 2×10⁶ S9 cells in 6-well plates, infected at a multiplicity of infection (MOI) of 2. For the temporal classification of the genes, cultures were pre-treated 1 h before infection with cytosine-1-β-D-arabinofuranoside (Ara-C, 100 µg/ml) and cycloheximide (CHX, 150 µg/ml) to inhibit DNA synthesis and protein synthesis, respectively. The inhibitors remained present during the infection.

2.2. Reverse transcription PCR (RT-PCR)

To determine the temporal expression classes of IIV6 genes, RNA isolated as described below was subjected to RT–PCR. Forward and reverse gene specific primers were designed to amplify suitable regions from all viral genes (Table S1). Total RNA was isolated from infected and mock-infected S9 cells at 12 h post infection (p.i.) using Trizol Reagent (Sigma, T9424) following the manufacturer’s instructions. Isolated RNA samples were treated with DNase I (Sigma, AMPI-D1-KT) to remove any residual DNA and then extracted with phenol–chloroform. For cDNA synthesis, 1 µg of total RNA was mixed with 1 µl (10 µM) gene specific reverse primer and the total volume was adjusted to 12 µl with water. After incubation at 65 °C for 5 min, the samples were cooled on ice. Subsequently, 4 µl reaction buffer (5X), 1 µl Ribonuclease inhibitor (20 µ/µl), 2 µl dNTP mix (10 mM) and 1 µl reverse transcriptase (Thermo Scientific, RevertAid M–MuLV RT, 200 u/µl) were gently mixed in and reactions were incubated at 42 °C for 60 min. The cDNA synthesis was terminated by heating at 70 °C for 5 min. and the resulting cDNA mixture was then used as template for gene specific PCR amplifications with forward and reverse primers. PCR products were analyzed in a 1% agarose gel stained with ethidium bromide. PCR performed with cDNA, obtained from infected cells in absence of inhibitors, was used as positive control (PC).

2.3. Conserved sequence analyses

MEME (multiple expectation maximization for motif elicitation) (Bailey et al., 2009) software was used to search for conserved sequences in IIV6 noncoding sequences in the 200 nt regions upstream of the translation initiation codons. To that aim, the upstream sequences were categorized based on experimental data (IE, DE or L) to be able to compare upstream sequences within each expression class. Parameters were set to zero or one occurrence per sequence and we searched only the provided (coding) strand.

2.4. Plasmid construction

Upstream sequences of selected genes from each temporal group were investigated to determine whether conserved motifs, found with the MEME software, are indeed important for promoter activity. One gene was selected from each of the three groups. 193R, 126R and 259R were selected as models for IE, DE and L class genes, respectively. Upstream regions of these genes were tested for promoter activity in combination with a luciferase reporter system. Subsequently, upstream sequences of two additional L genes, 061R and 084L were also investigated with this system. Two different regions were amplified from each upstream region; one is the long one containing the conserved motif (wt) and the other is the short one missing the motif (del for E and DE or dell.1 + L2 in case of L genes). These DNA fragments were amplified from the viral genome using two different forward primers and a common reverse primer, for each gene (Table 1). The resulting PCR products, containing KpnI and HindIII sites at 5′ and 3′ ends, respectively, were ligated into the pJET1.2/blunt cloning vector (Thermo). Subsequently these fragments were cloned into upstream of the luciferase reporter ORF of the pSPLuc + vector (Promega), again using the restriction sites at the ends of the fragments.

2.5. Transfection and luciferase assay

S9 cells (2.5 × 10⁶ cells/well) in 6-well plates were infected with IIV6 for 2 h and then transfected using Cellfectin (Invitrogen) with plasmid DNA (1 µg) harboring the upstream sequences. The various putative promoter constructs were tested in parallel. At 24 h after transfection, cells were collected by centrifugation at 1000g for 5 min. Firefly luciferase activities were measured in cell extracts using the single luciferase reporter assay system (Promega) following the manufacturer’s instructions. Transfections were conducted in triplicate, and average values are reported.

2.6. Site directed mutagenesis

Highly conserved sequences, found in the upstream regions of the L gene 061R were mutated to understand the role of these sequences in determining promoter activity. Mutations were performed by PCR using primers specific for the upstream region of 061R, but carrying a number of mismatched nucleotides (Table 1) (Nalcacioglu et al., 2003). Amplified sequences were first cloned into the pJET1.2/blunt cloning vector and then transferred to the pSPLuc + vector, as described above.

3. Results

3.1. Transcriptional classification of all IIV6 transcripts

To be able to categorize the whole set of genes in the IIV6 genome according to their transcriptional classes, we examined the expression of 170 IIV6 genes at the transcriptional level by RT-PCR. The other 45 genes in the IIV6 genome have previously been classified.
3

transcriptionally (Nalcacioglu et al., 2007; Ince et al., 2008; 2013; Dizman et al., 2012) and were not examined again, except for 012L (IE), 037L (DE) and 274L (L) that were used as positive controls in the current study. In order to classify the IIV6 genes, Sf9 cells were infected with IIV6 in the presence or absence of cycloheximide, which inhibits protein or DNA synthesis inhibitors, which means that viral protein DNA replication are not necessary for these transcripts and protein or DNA synthesis inhibitors, which means that viral protein synthesis and DNA replication are not necessary for these transcripts and therefore they are classified in the IE class (Fig. 1). The number of the late transcripts among the 170 newly tested genes, respective (Fig. 3 A-B). The program run for the upstream regions of IIE and DE genes. However, for the 35 scanned DE genes, the motif obtained was only observed in 20 genes. For late genes the program detected two conserved motifs, (T/G)(C/T)(A/C)(T/G)(T/C) (motif L1) and (C/G/T)(A/C)(T/A)(G/T)(G/T) (motif L2), with a similar and high probability in the upstream regions of all scanned late genes (Fig. 3C-D). The locations of all these upstream motifs respective to the translation start site varies for each gene.

4.1. Investigating the motifs for promoter activity

To analyze the influence of the conserved motifs on promoter activity, deletion mutagenesis was performed on the upstream regions of 193R, 126R and 259R ORFs, belonging to the IE, DE and L classes, respectively. Two fragments, one containing the motif and the other not, were prepared for each ORF. Reporter plasmids harboring the wild type (wt) or mutant sequence upstream of a firefly luciferase reporter ORF, were transfected into Sf9 cells. Cell lysates, obtained 24 h after transfection, were tested for luciferase activity.

The reporter plasmids that carried the wild type (wt) upstream regions for IE or DE genes (pSP193Rprom, pSP126Rprom), produced high luciferase activity, but the plasmids without the motif (pSP193Rdel, pSP126Rdel) produced a low activity (Fig. 4A-B). However, for L genes, the result was opposite. The plasmids containing both the L1 and L2 motifs produced a low luciferase activity (pSP259Rprom, pSP084Lprom), but the plasmids without these two motifs (pSP259RdelL1 + L2, pSP084LdelL1 + L2, pSP061RdelL1 + L2) produced a high activity. This result was validated with two additional late genes (061L, 084L) by preparing similar deletion mutations and testing the luciferase activity as mentioned above (Fig. 4C).

4.2. Site-directed mutations in conserved late gene motifs

To analyze these L motifs in more detail, we modified the L1 and L2 motifs in the upstream region of 061R individually by PCR using primers with mismatches. Reporter plasmids were prepared carrying both L1 and L2 motifs, but one unmutated sequence and the other mutated, to determine the impact of such changes on promoter activity. Mutation of motif L1 (mutation 1) did not affect the promoter activity, however mutation at motif L2 (mutation 2) increased promoter activity. This result demonstrates that motif L2 acts as a repressor on L gene promoter activity by a factor of over 90% (Fig. 5). The fact that expression levels were not fully restored to high levels by deleting L2, suggest that L1 is insufficient for a fully-functional promoter sequence.

The fact that the detected L motifs do not act as promoters led us to search for other conserved sequences that might have promoter activity.

### Table 1

<table>
<thead>
<tr>
<th>Primers</th>
<th>Tm (°C)</th>
<th>Primer sequences (5′-3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>193R-prom-Fw</td>
<td>46.4</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>193R-del-Fw</td>
<td>46.4</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>193R-prom-Rv</td>
<td>48.1</td>
<td>AAGCTTATATAAATTCACGATGTATCACAT</td>
</tr>
<tr>
<td>126R-prom-Fw</td>
<td>50.1</td>
<td>GGTAACGATAGCAGTTAAAACATTTGACGACTT</td>
</tr>
<tr>
<td>126R-del-Fw</td>
<td>45.2</td>
<td>GGTAACGATAGCAGTTAAAACATTTGACGACTT</td>
</tr>
<tr>
<td>126R-prom-Rv</td>
<td>47.4</td>
<td>AAGCTTATATAAATTCACGATGTATCACAT</td>
</tr>
<tr>
<td>259R-prom-Fw</td>
<td>50.4</td>
<td>GGTAACGATAGCAGTTAAAACATTTGACGACTT</td>
</tr>
<tr>
<td>259R-del-Fw</td>
<td>50.8</td>
<td>GGTAACGATAGCAGTTAAAACATTTGACGACTT</td>
</tr>
<tr>
<td>259R-prom-Rv</td>
<td>51.7</td>
<td>AAGCTTATATAAATTCACGATGTATCACAT</td>
</tr>
<tr>
<td>084L-prom-Fw</td>
<td>50.5</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>084L-del-Fw</td>
<td>49.5</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>061R-prom-Fw</td>
<td>46.9</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-del-Fw</td>
<td>45.9</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-prom-Rv</td>
<td>46.9</td>
<td>AAGCTTATATAAATTCACGATGTATCACAT</td>
</tr>
<tr>
<td>061R-mutL1-Fw</td>
<td>63.5</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL2-Fw</td>
<td>62.4</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL3-Fw</td>
<td>60.1</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL4-Fw</td>
<td>58.7</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
</tbody>
</table>

KpnI (GGTACC) and HindIII (AAGCTT) are shown in italicized and underlined.

### Table 2

<table>
<thead>
<tr>
<th>Primers</th>
<th>Tm (°C)</th>
<th>Primer sequences (5′-3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>061R-prom-Fw</td>
<td>46.9</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>061R-prom-Rv</td>
<td>45.9</td>
<td>GGTAACCGAGGATTTAAAAAAAGTTTATTAATTTAAA</td>
</tr>
<tr>
<td>061R-prom-Rv</td>
<td>46.9</td>
<td>AAGCTTATATAAATTCACGATGTATCACAT</td>
</tr>
<tr>
<td>061R-mutL1-Fw</td>
<td>63.5</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL2-Fw</td>
<td>62.4</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL3-Fw</td>
<td>60.1</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
<tr>
<td>061R-mutL4-Fw</td>
<td>58.7</td>
<td>GGTAACCGCTATTTCCTTCATTCAATTA</td>
</tr>
</tbody>
</table>

KpnI (GGTACC) and HindIII (AAGCTT) are shown in italicized and underlined.

Among the 138 IE genes identified in total, 61 have a known or a putative function according to gene ontology information obtained from the UniProt database. Eight of the 35 DE transcripts, have an identified putative function.

## 4. Conserved motifs in the upstream region of IIV6 genes

After grouping the genes in the three temporal classes, sequences upstream of the translational start codon of each gene were investigated for the presence of conserved and potentially important motifs for promoter activity. For each classified group of genes, motifs were generated by the MEME Suite database (Fig. 3). The AA(A/T)(T/A)TG(A/G)A and (T/A/C)(T/G/C)(T/T/C) ATGG sequences were identified with high probability as conserved motifs in the upstream regions of IE and DE genes, respectively (Fig. 3A-B). The program run for the upstream sequences of the 138 IE genes identified the conserved motif in all of these genes. However, for the 35 scanned DE genes, the motif obtained was conserved motifs, (T/G)(C/T)(A/C)(T/G)(T/C) (motif L1) and (C/G/T)(A/C)(T/A)(G/T)(G/T) (motif L2), with a similar and high probability in the upstream regions of all scanned late genes (Fig. 3C-D). The locations of all these upstream motifs respective to the translation start site varies for each gene.

4.1. Investigating the motifs for promoter activity

To analyze the influence of the conserved motifs on promoter activity, deletion mutagenesis was performed on the upstream regions of 193R, 126R and 259R ORFs, belonging to the IE, DE and L classes, respectively. Two fragments, one containing the motif and the other not, were prepared for each ORF. Reporter plasmids harboring the wild type (wt) or mutant sequence upstream of a firefly luciferase reporter ORF, were transfected into Sf9 cells. Cell lysates, obtained 24 h after transfection, were tested for luciferase activity.

The reporter plasmids that carried the wild type (wt) upstream regions for IE or DE genes (pSP193Rprom, pSP126Rprom), produced high luciferase activity, but the plasmids without the motif (pSP193Rdel, pSP126Rdel) produced a low activity (Fig. 4A-B). However, for L genes, the result was opposite. The plasmids containing both the L1 and L2 motifs produced a low luciferase activity (pSP259Rprom, pSP084Lprom), but the plasmids without these two motifs (pSP259RdelL1 + L2, pSP084LdelL1 + L2, pSP061RdelL1 + L2) produced a high activity. This result was validated with two additional late genes (061L, 084L) by preparing similar deletion mutations and testing the luciferase activity as mentioned above (Fig. 4C).

4.2. Site-directed mutations in conserved late gene motifs

To analyze these L motifs in more detail, we modified the L1 and L2 motifs in the upstream region of 061R individually by PCR using primers with mismatches. Reporter plasmids were prepared carrying both L1 and L2 motifs, but one unmutated sequence and the other mutated, to determine the impact of such changes on promoter activity. Mutation of motif L1 (mutation 1) did not affect the promoter activity, however mutation at motif L2 (mutation 2) increased promoter activity. This result demonstrates that motif L2 acts as a repressor on L gene promoter activity by a factor of over 90% (Fig. 5). The fact that expression levels were not fully restored to high levels by deleting L2, suggest that L1 is insufficient for a fully-functional promoter sequence.

The fact that the detected L motifs do not act as promoters led us to search for other conserved sequences that might have promoter activity.
Therefore, upstream sequences of all IIV6 L genes, from which the L1 and L2 motifs were in silico removed, were re-analyzed for the presence of potential promoter sequences using MEME Suite Software, resulting in two additional conserved sequences (motif L3 and motif L4) (Fig. 5). These motifs were individually mutated in PCR fragments that did not contain the L1 and L2 motifs and reporter analysis clearly showed that both motifs L3 and L4 contributed to L promoter activity (Fig. 5).

5. Discussion

This study presents extensive information on the transcriptional regulation of invertebrate iridescent virus 6 (IIV6) genes. Transcriptional studies on iridovirids (members of the family Iridoviridae) have been reported previously for Frog virus 3 (Majji et al., 2009), Singapore grouper iridovirus (Chen et al., 2006; Teng et al., 2008), Red sea bream iridovirus (Lua et al., 2005; Dang et al., 2007; 2008), IIV6 (D’Costa et al., 2001, 2004; Ince et al., 2008; 2013; Nalcacioglu et al., 2003; Dizman et al., 2012) and IV9 (McMillan and Kalmakoff, 1994). The first transcriptional study on IIV6 genes identified 137 transcripts of which 38 corresponded to IE, 34 in DE and 65 in L temporal classes based on northern blot analysis. However, these authors did not clearly identify the ORFs in their study (D’Costa et al., 2004). Other transcriptional studies on IIV6 genes included temporal classification of a few genes: DNA polymerase (037L), major capsid protein (274L),
exonuclease (012L), inhibitor of apoptosis (193R), and putative myristoylated membrane protein (458R) (Nalcacioglu et al., 2003; Ince et al., 2008; Dizman et al., 2012; Aksu, 2017). A further study, performed on IV6, classified 41 of the 54 structural protein coding genes into temporal classes using a ligation-based amplification of cDNA ends (LACE) technique (Ince et al., 2013). In the present study the remaining IV6 genes that were not studied until now, have been classified according to their temporal expression class. Reverse transcription polymerase chain reaction (RT-PCR) has been successfully used to identify the gene transcripts using total RNA isolated from virus-infected cells in the presence of protein or viral DNA synthesis inhibitors. Since IV6 transcripts do not contain a polyA tail (Nalcacioglu et al., 2003), specific primer sets designed for the 170 genes were used to detect the gene specific transcripts. Of these, 158 genes were placed into either of the three temporal classes. Transcripts of the remaining 12 genes (069L, 121R, 146R, 148R, 201R, 212L, 236L, 238R, 315L, 414L, 426R, 463L), could not be detected in the RNA isolated at 12 h post infection. Transcripts that were not detected in the presence of the protein synthesis inhibitor (CHX) but were not affected by the DNA synthesis inhibitor are from DE genes (A). Genes that do not show transcripts in the presence of both inhibitors as categorized as L genes (B). ORF: open reading frame; Ara-C: DNA synthesis inhibitor (cytosine arabinoside); CHX: protein synthesis inhibitor (cycloheximide). PC: Positive control, infection without addition of inhibitors.

<table>
<thead>
<tr>
<th>ORF</th>
<th>RNA Class</th>
<th>Molecular function</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>006L</td>
<td>IE</td>
<td>Putative KIAA-N domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>009R</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>010R</td>
<td>IE</td>
<td>Transmembrane protein</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>012L</td>
<td>IE</td>
<td>Exonuclease activity, nucleic acid binding</td>
<td></td>
</tr>
<tr>
<td>019R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>022L*</td>
<td>IE</td>
<td>Helicase activity, ATP binding</td>
<td></td>
</tr>
<tr>
<td>029R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>030L</td>
<td>IE</td>
<td>DNA helicase activity</td>
<td>DNA repair, telomere maintenance</td>
</tr>
<tr>
<td>032R</td>
<td>L</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>034R*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>037L*</td>
<td>DE</td>
<td>DNA-directed DNA polymerase activity, DNA binding, nucleotide binding</td>
<td>DNA replication, viral DNA genome replication</td>
</tr>
<tr>
<td>041L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>042R</td>
<td>DE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>043L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>044R</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>045L</td>
<td>IE</td>
<td>DNA topoisomerase type II (ATP-hydrolyzing) activity, ATP binding, DNA binding, metal ion binding</td>
<td>DNA topological change</td>
</tr>
<tr>
<td>049L</td>
<td>IE</td>
<td>Nicotinamide riboside</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>050L</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>056R</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>060L</td>
<td>DE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>061R</td>
<td>L</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>062L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>065R</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>067R</td>
<td>L</td>
<td>Helicase family protein</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>069L</td>
<td>ND</td>
<td>Putative Bro-N domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>071L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>075L</td>
<td>IE</td>
<td>ATP binding, AAA-ATPase, similar to poxvirus A32, required for DNA packaging</td>
<td></td>
</tr>
<tr>
<td>077L</td>
<td>IE</td>
<td>Metal ion binding, Putative zinc finger protein</td>
<td></td>
</tr>
<tr>
<td>082L</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>083L</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>084L</td>
<td>L</td>
<td>Putative myristoylated protein</td>
<td>DNA replication, intein-mediated protein splicing</td>
</tr>
<tr>
<td>085L</td>
<td>IE</td>
<td>Endonuclease activity, ribonucleoside-diphosphosphate reductase activity, thioredoxin disulfide as acceptor</td>
<td></td>
</tr>
<tr>
<td>094L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>095L</td>
<td>IE</td>
<td>Metal ion binding, Putative RING finger protein</td>
<td></td>
</tr>
<tr>
<td>096L</td>
<td>L</td>
<td>Putative FAST domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>098R</td>
<td>DE</td>
<td>Serine/threonine protein kinase</td>
<td></td>
</tr>
<tr>
<td>100L</td>
<td>IE</td>
<td>NAD + ADP-ribosyltransferase activity</td>
<td></td>
</tr>
<tr>
<td>101L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>104L*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>106L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>107L</td>
<td>IE</td>
<td>DNA binding, DNA-directed 5′-3′ RNA polymerase activity</td>
<td>DNA-templated transcription</td>
</tr>
</tbody>
</table>

Fig. 2. Delayed-early (DE) and late (L) gene transcripts of IV6. Cells were infected with IV6 in the presence of DNA (Ara-C) or protein synthesis (CHX) inhibitors. ORF-specific RT-PCR was carried out on total RNA isolated at 12 h post infection. Transcripts that were not detected in the presence of the protein synthesis inhibitor (CHX) but were not affected by the DNA synthesis inhibitor are from DE genes (A). Genes that do not show transcripts in the presence of both inhibitors as categorized as L genes (B). ORF: open reading frame; Ara-C: DNA synthesis inhibitor (cytosine arabinoside); CHX: protein synthesis inhibitor (cycloheximide). PC: Positive control, infection without addition of inhibitors.

<p>| Table 2: Temporal classification of CIV gene transcripts based on inhibitor studies. |
|-----------------|----------------------------------------|-----------------|</p>
<table>
<thead>
<tr>
<th>ORF</th>
<th>RNA Class</th>
<th>Molecular function</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>006L</td>
<td>IE</td>
<td>Putative KIAA-N domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>009R</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>010R</td>
<td>IE</td>
<td>Transmembrane protein</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>012L</td>
<td>IE</td>
<td>Exonuclease activity, nucleic acid binding</td>
<td></td>
</tr>
<tr>
<td>019R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>022L*</td>
<td>IE</td>
<td>Helicase activity, ATP binding</td>
<td></td>
</tr>
<tr>
<td>029R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>030L</td>
<td>IE</td>
<td>DNA helicase activity</td>
<td>DNA repair, telomere maintenance</td>
</tr>
<tr>
<td>032R</td>
<td>L</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>034R*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>037L*</td>
<td>DE</td>
<td>DNA-directed DNA polymerase activity, DNA binding, nucleotide binding</td>
<td>DNA replication, viral DNA genome replication</td>
</tr>
<tr>
<td>041L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>042R</td>
<td>DE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>043L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>044R</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>045L</td>
<td>IE</td>
<td>DNA topoisomerase type II (ATP-hydrolyzing) activity, ATP binding, DNA binding, metal ion binding</td>
<td>DNA topological change</td>
</tr>
<tr>
<td>049L</td>
<td>IE</td>
<td>Nicotinamide riboside</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>050L</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>056R</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>060L</td>
<td>DE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>061R</td>
<td>L</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>062L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>065R</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>067R</td>
<td>L</td>
<td>Helicase family protein</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>069L</td>
<td>ND</td>
<td>Putative Bro-N domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>071L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>075L</td>
<td>IE</td>
<td>ATP binding, AAA-ATPase, similar to poxvirus A32, required for DNA packaging</td>
<td></td>
</tr>
<tr>
<td>077L</td>
<td>IE</td>
<td>Metal ion binding, Putative zinc finger protein</td>
<td></td>
</tr>
<tr>
<td>082L</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>083L</td>
<td>IE</td>
<td>Unknown</td>
<td>Integral component of membrane</td>
</tr>
<tr>
<td>084L</td>
<td>L</td>
<td>Putative myristoylated protein</td>
<td>DNA replication, intein-mediated protein splicing</td>
</tr>
<tr>
<td>085L</td>
<td>IE</td>
<td>Endonuclease activity, ribonucleoside-diphosphosphate reductase activity, thioredoxin disulfide as acceptor</td>
<td></td>
</tr>
<tr>
<td>094L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>095L</td>
<td>IE</td>
<td>Metal ion binding, Putative RING finger protein</td>
<td></td>
</tr>
<tr>
<td>096L</td>
<td>L</td>
<td>Putative FAST domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>098R</td>
<td>DE</td>
<td>Serine/threonine protein kinase</td>
<td></td>
</tr>
<tr>
<td>100L</td>
<td>IE</td>
<td>NAD + ADP-ribosyltransferase activity</td>
<td></td>
</tr>
<tr>
<td>101L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>104L*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>106L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>107L</td>
<td>IE</td>
<td>DNA binding, DNA-directed 5′-3′ RNA polymerase activity</td>
<td>DNA-templated transcription</td>
</tr>
</tbody>
</table>

(continued on next page)
### Table 2 (continued)

<table>
<thead>
<tr>
<th>ORF RNA</th>
<th>Molecular function</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>110R DE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>111R L</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>113L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>115R DE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>116L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>117L* DE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>118L* IE</td>
<td>Putative myristoylated protein</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>120L ND</td>
<td>Putative RING finger protein, metal ion binding</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>122R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>123R* IE</td>
<td>Putative metalloendopeptidase</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>124R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>125L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>126R DE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>127L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>128R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>129L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>130R DE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>131L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>132L IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>133R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>134R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>135R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>136R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>137R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>138R* IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>139R ND</td>
<td>Putative RING finger protein, metal ion binding</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>140R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>141R IE</td>
<td>Unknown</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>142R* IE</td>
<td>Ribonuclease III activity, RNA binding</td>
<td>RNA processing</td>
</tr>
<tr>
<td>143R IE</td>
<td>Nucleoside kinase activity, ATP binding</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>144R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>145L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>146R ND</td>
<td>Putative MSV199 domain-containing protein</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>148R ND</td>
<td>Putative MSV199 domain-containing protein</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>149* DE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>150L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>151R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>152L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>153R DE</td>
<td>Metalloendopeptidase</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>154L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>155L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>156R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>157L IE</td>
<td>Metal ion binding</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>158R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>159L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>160L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>161L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>162L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>163R DE</td>
<td>Putative metalloendopeptidase activity, zinc ion binding</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>164L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>165R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>166R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>167L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>168L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>169L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>170L IE</td>
<td>Putative RING finger protein, metal ion binding</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>171L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>172L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>173R IE</td>
<td>Putative RING finger protein, metal ion binding</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>174L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>175R* IE</td>
<td>Protein kinase activity</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>176R DE</td>
<td>Helicase activity, DNA primase activity, ATP binding, D5 family NTPase</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>177L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>178L IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>179R* IE</td>
<td>Protein kinase activity</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>180R* DE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>181R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>182R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>183R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>184R DE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>185R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>186R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>187R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>188R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>189R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>190R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>191R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>192R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>193R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>194R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>195R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
<tr>
<td>196R IE</td>
<td>Unknown</td>
<td>DNA replication, termination</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>ORF</th>
<th>RNA Class</th>
<th>Molecular function</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>293R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>295L*</td>
<td>IE</td>
<td>Vaccinia virus early transcription factor</td>
<td></td>
</tr>
<tr>
<td>300R</td>
<td>IE</td>
<td>Transmembrane protein</td>
<td></td>
</tr>
<tr>
<td>301L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>302L</td>
<td>IE</td>
<td>Putative zinc finger protein, metal ion binding, nucleic acid binding</td>
<td></td>
</tr>
<tr>
<td>305L</td>
<td>IE</td>
<td>Putative SWIB domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>307L</td>
<td>L</td>
<td>Uvr/REP helicase</td>
<td>integral component of membrane</td>
</tr>
<tr>
<td>308L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>309L*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>310L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>311R</td>
<td>DE</td>
<td>Immediate-early protein ICP-47</td>
<td></td>
</tr>
<tr>
<td>312R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>313L</td>
<td>IE</td>
<td>Putative KilA-N domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>315L</td>
<td>ND</td>
<td>Probable lipid hydrolase, metal ion binding</td>
<td></td>
</tr>
<tr>
<td>317L*</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>318L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>319L*</td>
<td>IE</td>
<td>Putative zinc finger protein, nucleic acid binding, zinc ion binding</td>
<td></td>
</tr>
<tr>
<td>320R</td>
<td>DE</td>
<td>DNA-templated transcription elongation factor S-II-like protein</td>
<td></td>
</tr>
<tr>
<td>322R</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>325L</td>
<td>IE</td>
<td>Immediate-early protein ICP-47</td>
<td></td>
</tr>
<tr>
<td>326L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>329R*</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>330L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>331L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>332L</td>
<td>IE</td>
<td>Uncharacterized RING finger protein, metal ion binding</td>
<td></td>
</tr>
<tr>
<td>333L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>335L</td>
<td>ND</td>
<td>Probable lipid hydrolase, metal ion binding</td>
<td></td>
</tr>
<tr>
<td>336L</td>
<td>IE</td>
<td>DNA-templated transcription elongation factor S-II-like protein</td>
<td></td>
</tr>
<tr>
<td>337L*</td>
<td>DE</td>
<td>Putative membrane protein, Myristylated membrane protein</td>
<td>virion membrane, integral component of membrane</td>
</tr>
<tr>
<td>340R</td>
<td>DE</td>
<td>DRBM domain-containing protein, RNA binding</td>
<td></td>
</tr>
<tr>
<td>342R*</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>343L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>345L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>346L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>347L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>348R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>349L</td>
<td>IE</td>
<td>Immediate-early protein ICP-47</td>
<td></td>
</tr>
<tr>
<td>350L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>352L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>355R</td>
<td>IE</td>
<td>Putative CTD phosphatase-like protein, phosphoprotein phosphatase activity</td>
<td></td>
</tr>
<tr>
<td>357R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>358L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>359L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>361L*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>362L</td>
<td>IE</td>
<td>Putative RING finger protein, metal ion binding</td>
<td></td>
</tr>
<tr>
<td>364L</td>
<td>L</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>365L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>366R</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>368L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>369L</td>
<td>IE</td>
<td>Putative CTD phosphatase-like protein, phosphoprotein phosphatase activity</td>
<td></td>
</tr>
<tr>
<td>373L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>374R*</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>375R</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>376L</td>
<td>IE</td>
<td>Metal ion binding, deoxyribonucleotide diphosphate reductase activity, thioredoxin disulfide as acceptor</td>
<td></td>
</tr>
<tr>
<td>378R*</td>
<td>DE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>380L*</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>381L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>382L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>383L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>384L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>388L</td>
<td>IE</td>
<td>Unknown</td>
<td></td>
</tr>
</tbody>
</table>

Stars (*) indicate that previously determined ORFs. Hash (#) symbols show the transcripts which are detected in the presence of CHX, but not determined in the presence of Ara-C, so those are classified as immediate early genes. IE, immediate early; DE, delayed early; L, late; ND, not determined. Molecular functions and biological process were obtained from UniProt database according to gene ontology. Grey shaded ORFs are core genes among all iridovirids.

RT-PCR. Bioinformatics analysis showed that most of the 12 remaining unclassified genes in the IIV6 genome contain domains associated with DNA binding: Bro-N (069L, 201R), RING finger (121R), MSV199 (146R, 148R, 212L, 238R) and KilA-N (315L). Also, two of these unclassified genes, 414L and 463L, have hydrolase activity.

Five ORFs (216R, 235L, 253L, 411L, 413R) were detected in the presence of the protein synthesis inhibitor (consistent with their presence of Ara-C, so those are classified as immediate early genes. IE, immediate early; DE, delayed early; L, late; ND, not determined. Molecular functions and biological process were obtained from UniProt database according to gene ontology. Grey shaded ORFs are core genes among all iridovirids.

RT-PCR. Bioinformatics analysis showed that most of the 12 remaining unclassified genes in the IIV6 genome contain domains associated with DNA binding: Bro-N (069L, 201R), RING finger (121R), MSV199 (146R, 148R, 212L, 238R) and KilA-N (315L). Also, two of these unclassified genes, 414L and 463L, have hydrolase activity.

Five ORFs (216R, 235L, 253L, 411L, 413R) were detected in the presence of the protein synthesis inhibitor (consistent with their presence of Ara-C, so those are classified as immediate early genes. IE, immediate early; DE, delayed early; L, late; ND, not determined. Molecular functions and biological process were obtained from UniProt database according to gene ontology. Grey shaded ORFs are core genes among all iridovirids.
classification as IE genes), but not in the presence of the DNA synthesis inhibitor (consistent with their classification as L genes). This result was therefore not conforming the defined temporal groups. However, since these transcripts were detected in the presence of protein synthesis inhibitor like IE genes, they were also accepted as the products of IE genes. These genes may be transactivated early in infection by already existing proteins in the virion structure (IE), but at the same time be expressed in the late stage so that their gene products could be incorporated in the virion and may themselves assist in the transactivation process. A similar result was obtained for 219L and 295L by Ince et al., (2013).

So, more than half of the IIV6 genome can be immediately transcribed, most likely due to the action of transactivators incorporated in the viral particles. Compared to the other iridovirids (Majji et al., 2009; Chen et al., 2006; Teng et al., 2008; Dang et al., 2007), IIV6 has the highest number of ORFs in the IE class.

Among the 98 putative ORFs of Frog virus 3 (FV3), 31 have corresponding homologous in IIV6 (Tan et al., 2004). However, the temporal classes of these 31 IIV6 ORFs are mostly inconsistent with their counterparts in the FV3 transcriptome (Majji et al., 2009). For most of these homologues these inconsistencies differed by only one temporal class. Whether IIV6 has a closer resemblance in respect to transcriptional classes of its genes to invertebrate iridoviruses needs to be determined. Furthermore, the fact that genes are classified as IE genes, does not imply that a number of these genes may not continue to be expressed in the DE phase, or are expressed subsequently as late class transcripts, as implied that a number of these genes may not continue to be expressed in the late stage so that their gene products could be incorporated in the virion and may themselves assist in the transactivation process. A similar result was obtained for 219L and 295L by Ince et al., (2013).

Two conserved motifs without promoter activity were detected in the upstream sequences of IIV6 L genes. Motif L2, ending with the sequence (T/A)(T/G)(T/C) served as cis-acting repressor element for the ORF downstream, at least for the three ORFs (259R, 061L, 084L) tested in this study. The promoter region of the mcp has been analyzed previously (Nalcioglu et al., 2003) and contains several repeats of motif L2. A mcp promoter construct starting at –53 relative to the ATG start codon, contained such a motif and had a high promoter activity, but the construct –23 that lacked this motif had reduced promoter activity in the luciferase reporter assay (Nalcioglu et al., 2003). According to that result we may conclude that this motif does not function as a repressor for all IIV6 L genes, and it may be the context in which motif L2 is present is also important. Cis-acting repressor elements have been reported for a few viruses including Human cytomegalovirus (HCMV) (Biegelke, 1995; Lashmit et al., 1998), Equine herpesvirus 1 (EHV-1) (Ahn et al., 2010), Human Immunodeficiency Virus Type 1 (HIV-1) (Huffman and Arrigo, 1997) but to our knowledge this is the first identification of such an element in an insect virus. The in silico identified upstream motif is indeed present within the experimentally identified promoter region of 012L. The IIV6 DNA polymerase gene (037L) was reported previously as a DE class gene and it was shown experimentally that an AAAAT motif was essential for promoter activity (Nalcioglu et al., 2007). The most prevalent motif, (T/A/C)(T/G/C)/T (T/A)ATGG, identified for DE gene promoters is not present in the upstream sequence of 037L.

The obtained information on the temporal classes of almost all predicted IIV6 genes provided the basis for identifying whether ORFs in the same temporal class contained common upstream regulatory regions. The MEME Suite program identified the AA(A/T)(T/A)TG(A/G)/A motif in the upstream sequence of all identified IE genes of IIV6. Previously, Dizman et al. (2012) mapped the limits of the putative promoter region of the IIV6 IE gene 012L using reporter gene assays. The in silico identified upstream motif is indeed present within the experimentally identified promoter region of 012L. The IIV6 DNA polymerase gene (037L) was reported previously as a DE class gene and it was shown experimentally that an AAAAT motif was essential for promoter activity (Nalcioglu et al., 2007). The most prevalent motif, (T/A/C)(T/G/C)/T (T/A)ATGG, identified for DE gene promoters is not present in the upstream sequence of 037L.
Luciferase activity (RLU/sec)

The authors of the review suggested that the MEGA-box has been used in cytoplasmic large DNA viruses (NCLDV), including iridovirids, and that the identified motifs are illustrated in red text. Luciferase activities of deletion and site directed-mutated promoter plasmids were compared to the wild type ORF 061R. The mutated motifs are indicated with yellow boxes. The mutations in the four identified motifs in the upstream region are exemplified for Fig. 5.

A specific 8 nt conserved motif, AAAATTGA, has been found in nearly 50% of the IIV6 IE motif is remarkable. In a review by Oliveira et al. (2017b), the similarity of this motif with the IIV6 IE motif is remarkable. In a review by Oliveira et al. (2017b), the current available information about promoter regions in nucleocytoplasmic large DNA viruses (NCLDV), including iridovirids, was discussed. The fact that some promoter sequences found in one family are very similar to those found in their relatives suggests a common origin for NCLDV genes and a promoter motif (TATAA/ATT/GA) with the name ‘MEGA-box’ was proposed for their common ancestor. The authors of the review suggested that the MEGA-box has been gradually evolved by nucleotide gain and loss into the functional promoters in the present-day NCLDVs. The resemblance of the IIV6 IE motif found in our study to the MEGA-box sequence reinforces this hypothesis.

Here, we have presented a study including both experimental and in silico analysis on transcriptional regulators of IIV6. Two hundred and twenty-five ORFs in the IIV6 genome have now been classified and include 138 IE genes, 35 DE genes and 30 I genes. Motifs detected in the upstream regions of the genes further delineate the temporal classes. The fact that all IE class genes contain the motif detected by the computer program increases the probability that this motif will be essential for promoter activity for all IIV6 IE genes. This information increases our understanding of the molecular mechanisms of IIV6 infection.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgement**

This study was supported by a 2214/a scholarship (Project no: 2142172) to Aydin Yesilyurt from the Scientific and Technological Research Council of Turkey (TÜBİTAK) allowing him to do part of his studies at Wageningen University, the Netherlands and a research grant from the Karadeniz Technical University (Project no. 5839).

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jip.2020.107496.

**References**


