ISSN 1607-6729, Doklady Biochemistry and Biophysics, 2008, Vol. 421, pp. 176–180. © Pleiades Publishing, Ltd., 2008.
Original Russian Text © T.V. Kapelinskaya, E.U. Martynova, A.L. Korolev, C. Schal, D.V. Mukha, 2008, published in Doklady Akademii Nauk, 2008, Vol. 421, No. 2, pp. 256–261.

BIOCHEMISTRY, BIOPHYSICS, AND MOLECULAR BIOLOGY

Transcription of the German Cockroach Densovirus BgDNV Genome: Alternative Processing of Viral RNAs

T. V. Kapelinskaya[a], E. U. Martynova[a], A. L. Korolev[a], C. Schal[b], and D. V. Mukha[a]

Presented by Academician I.G. Atabekov January 23, 2008

Received October 4, 2007

DOI: 10.1134/S1607672908040042

Parvoviruses are among the smallest DNA animal viruses: their single-stranded genome is 4–6 thousand nucleotides long. For this reason, alternative splicing of viral pre-mRNAs plays a major role in increasing the coding ability of their small genome. In all studied vertebrate parvoviruses, expression is regulated primarily at the transcriptional level. However, transcription and processing of RNAs of the majority of densoviruses has not yet been studied. In this work, we studied the consistent transcription patterns of the German cockroach densovirus BgDNV. This study provided new information on the expression of this densovirus, which has not been described for any other insect densovirus. These data are important not only for studying particular characteristics of the life cycle of viruses but also for understanding the fundamentals of regulation of eukaryotic gene expression. When BGE2 cockroach cell culture is infected with the BgDNV virus, several viral RNAs, formed as a result of alternative splicing and use of alternative transcription termination codons, are transcribed from each DNA strand of the viral genome. It was shown that splicing of RNAs encoding both VP and NS proteins is observed during infection development and that the viral RNAs for NS proteins undergo alternative splicing. We determined the size of the transcripts synthesized from two viral promoters and showed that the transcription from the NS promoter starts much earlier than the transcription from the VP promoter. We also showed that the transcripts synthesized from antiparallel DNA strands of the BgDNV virus genome considerably overlap as a result of RNA processing.

The Parvoviridae family consists of two subfamilies—Parvovirinae (infecting vertebrates) and Densovirinae (infecting arthropods, primarily insects). The Densovirinae subfamily includes five genera differing in the structural organization of the genome and expression strategies: Densovirus, Brevidensovirus, Iteravirus, and Bidensovirus; the densovirus of the smoky-brown cockroach *Periplaneta fuliginosa* was segregated into a special genus [2–4]. The BgDNV densovirus of the German cockroach *Blattella germanica* is a member of the Parvoviridae family. We have first isolated it from cockroaches of the laboratory strain P6, originated from the population of cockroaches of an American pig farm [1]. Similarly to vertebrate parvoviruses, densoviruses autonomously replicate only in dividing cells (e.g., in insect larvae or in tissues with rapidly changed cells, such as gastrointestinal tract epithelial cells or hemocytes), since they use the replicative machine of the cell. However, in contrast to parvoviruses, the majority of densoviruses cause fatal pathological disorders in infected cells and are highly species-specific. Possibly, the last feature is determined by their strong dependence of the replicative functions of hosts, which are much more diverse than in vertebrates. Currently, biological methods of control of abundance of several mosquito species (the *Aedes albopictus* virus AaDNV) [5], oil-palm aphids (the *Casphalia extranea* virus CaDNV) [6], and bee moth that causes significant harm to bee-farming (the *Galleria mellonella* virus GmDNV) [7] have been developed on the basis of densoviruses. The study of the German cockroach densovirus BgDNV [1] may be useful for designing biological pesticides of a new generation intended to control the abundance of this synanthropic parasite.

Genome transcription is a key stage in the life cycle of viruses, since this is the process of genetic information transduction from the virus to the translational apparatus of the host cell, which affects the synthesis of viral proteins required for replication of viral DNA and assembling of new viral particles. It is known that all vertebrate parvoviruses use several promoters, three reading frames, and alternative splicing in the course of formation of multiple RNAs, which are encoded by only one (plus) DNA strand [8]. In contrast to verte-

[a] *Vavilov Institute of General Genetics, Russian Academy of Sciences, ul. Gubkina 3, Moscow, 119991 Russia*
[b] *North Carolina State University, Raleigh, United States*

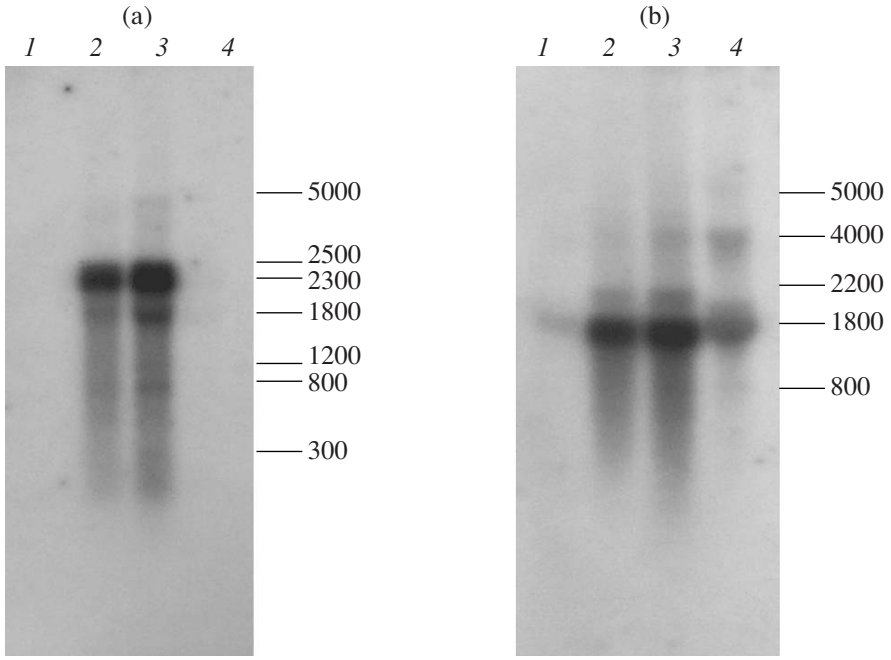

Fig. 1. Results of Northern blot hybridization of total RNAs isolated from BGE-2 German cockroach cell culture infected with the BgDNV densovirus. Hybridization was performed with probes complementary to different regions of the viral genome encoding (a) capsid proteins VP and (b) nonstructural proteins NS. Time elapsed after infection: (*1*) 3.5 h; (*2*) 20 days, V2 (passage 2); (*3*) 40 days, V4 (passage 4); (*4*) 120 days, V12 (passage 12). The sizes of hybridized transcripts (nt) are shown.

brate parvoviruses, densoviruses often have ambipolar genomes; their expression strategy is often based on the so-called leaky scanning mechanism, when ribosomes slide off some start ATG codons and initiate translation from internal start sites on viral RNAs [9]. However, some insect densoviruses also use splicing during transcription of either VP-RNA (viruses MpDNV [10] and PfDNV [3]) or NS-RNA (viruses GmDNV [9] and MlDNV [11]).

In this study, we used Northern blot hybridization of total RNA from BGE2 cells at different stages of infection. Northern blot hybridization was performed with labeled viral DNA fragments specific for the genomic regions corresponding to open reading frames for viral proteins. To obtain cDNA copies of viral RNAs and determine transcription start and termination sites, we used RT-PCR as well as 5'- and 3'-RACE with subsequent cloning of obtained fragments into the c pGEM T-Easy plasmid vector (Promega) and their sequencing.

The results of hybridization of viral RNAs with corresponding VP- and NS-coding viral DNA fragments are shown in Fig. 1. As can be seen in Fig. 1a, the major viral RNAs coding for VP proteins are approximately 2500 and 2300 nucleotides long. The same figure shows four additional minor transcripts approximately 1800, 1200, 800, and 500 nucleotides long. Viral RNAs coding for NS proteins are transcribed in the form of two major transcripts 2200 and 1800 nucleotides long (Fig. 1b). Figure 1b also shows the positions of two minor transcripts approximately 4000 and 800 nucleotides long. In addition, our results show that, up to the 12th passage of cells, the transcription of viral RNAs for NS proteins begins earlier and continues longer than the transcription of capsid RNAs.

Thus, Northern blot hybridization allowed us to determine the sizes of the major viral RNAs transcribed from two sides of the ambipolar genome of the BgDNV virus. However, their structure and ends can be determined precisely only by sequencing the cDNA copies obtained by cloning of RT-PCR and RACE products of viral transcripts. For this purpose, we ordered virus-specific primers to presumable sites of transcription initiation from viral promoters P1, P2, and P3 and to the starts and ends of open reading frames (ORFs) of the BgDNV virus, which had been shown earlier [1] to contain three presumable promoters and six ORFs (Fig. 2). In our work, we used oligonucleotide primers whose positions in the genome and nucleotide sequences are shown in Fig. 2.

Figure 3 shows the results of electrophoretic separation of RT-PCR products obtained from total cell RNAs isolated from infected cells or infected adult cockroaches with the use of different pairs of primers to the viral genome. Experiments with primers for obtaining VP-cDNA showed that two different cDNAs (2600 and 2400 bp long) were synthesized with one pair of primers (P1-start and ORF1-end). This finding suggests that the processing of viral RNAs from the side of the P1 promoter involves splicing (Fig. 3, lane *1*). These VP-cDNA fragments are comparable in size with the

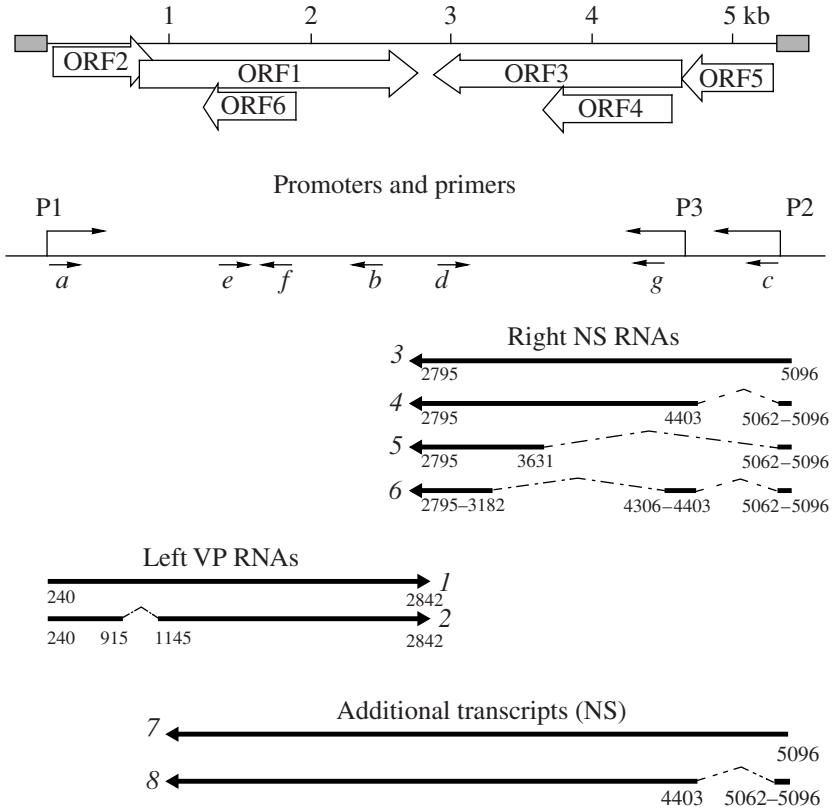

Fig. 2. Scheme illustrating the strategy of splicing of pre-RNA of German cockroach BgDNV densovirus. The BgDNV densovirus genome (5335 nucleotides) is shown on the top. Designations: gray rectangles, inverted terminal repeats (ITRs); large empty arrows, open reading frames (ORFs). Arrows *1–8* indicate the alternatively spliced viral RNAs. The dotted lines show the excised RNA fragments. Arrows indicate the direction of transcription. Numerals *1–8* are the numbers of transcripts. Numerals under arrows indicate the numbers of nucleotides in the viral genome brought together as a result of splicing as well as the starts and ends of transcripts. Arrows on the line of promoters and primers show the direction of transcription from promoters P1, P2, and P3, as well as the direction of amplification from the primers specified. Latin letters designate the names of primers and their sequences from the 5' end: (*a*) P1-start, agtatgtctacgtccggc; (*b*) ORF1-end, gatcgcccatatttctatgg; (*c*) P2-start, cctacgtctgctctcctcca; (*d*) ORF3-end, gtacatcaacacaagcatctg; (*e*) ORF6-end, acgatgtctgcaccagag; (*f*) ORF6-start, cttgggaatatgcatttatgtac; and (*g*) ORF3-start, gtgactccaaagcggctcca.

major transcripts shown in Fig. 1a and are cDNA copies of these transcripts. Note that the size of the upper fragment (2600 bp) on lane *1* precisely corresponds to the size of the viral genome fragment encompassed by these primers (Fig. 3, lane *2*). The last fragment was obtained by RT-PCR with the same pair of primers; the cockroach total RNA isolated at the terminal stage of infection with the BgDNV virus served as a template. At this stage of infection, cells accumulate high quantities of replicative forms of viral DNA, which is present as an admixture in this RNA preparation. The size of the lower 2400-bp cDNA fragment on lane *1* is reduced by approximately 200 bp, apparently as a result of excision of viral genomic sequence in the course of splicing. This assumption was confirmed by the results of sequencing of respective cloned cDNA fragments. Indeed, the nucleotide sequence showed that the ends of these two cDNAs are identical. Between ORF1 and ORF2 in the shorter cDNA, there is the site for splicing, as a result of which nucleotides 916–1144 are eliminated from the viral genome (in total, 229 nucleotides). The resultant new large ORF encodes a protein comprised of 778 amino acid residues. In Fig. 2, viral mRNAs encoding the capsid proteins (VP) are indicated with arrows with numbers 1 (pre-mRNA–VP) and 2 (mature viral VP-RNA). The results are consistent with the fact that many densoviruses contain only one large ORF encoding the VP1 capsid protein, whereas in other densoviruses it is also formed at the expense of pre-mRNA splicing between two ORFs [3, 9, 10].

NS-cDNA fragments were synthesized using pairs of primers corresponding to the given region of the virus genome. The results of RT-PCR shown in Fig. 3 (lanes *3* and *4*) were obtained with primers P2-start and ORF3-end. Several fragments whose sizes are smaller than the calculated size for the virus genome region encompassed between these primers (the upper fragment 2200 bp long on lane *4*) can be well seen in lane *3*. The brightest band on lane *3* in the gel, corresponding to the fragment approximately 1600 bp long, was extracted from agarose and sequences from both sides

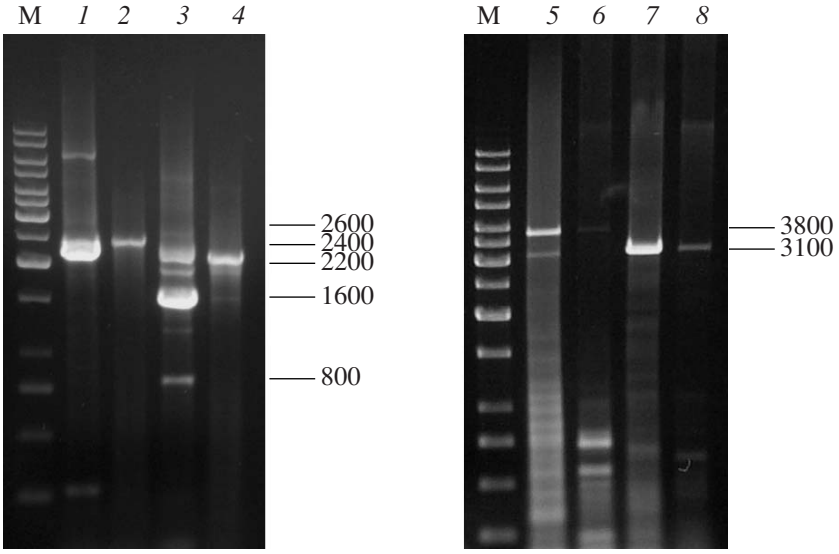

Fig. 3. Results of electrophoretic separation of RT-PCR products of total RNAs isolated from cell cultures infected with the BgDNV densovirus at two stages of infection development (V4 and V12) and from adult cockroaches two weeks after infection. RT-PCR was performed with the use of virus-specific primers. Lanes *1–8* show RT-PCR fragments obtained on (*1*) totRNA (V4) with primers P1-start and ORF1-end, (*2*) BgDNV DNA and infected cockroach totRNA with primers P1-start and ORF1-end, (*3*) totRNA (V4) with primers P2-start and ORF3-end, (*4*) BgDNV DNA and infected cockroach totRNA with primers P2-start and ORF3-end, (*5*) totRNA (V4) with primers P2-start and ORF6-end, (*6*) totRNA (V12) with primers P2- start and ORF6-end, (*7*) totRNA (V4) with primers ORF3-start and ORF6-end, and (*8*) totRNA (V12) with primers ORF3-start and ORF6-end. M is a 1-kb marker (Fermentas).

with the primers that were used in RT-PCR. The nucleotide sequence of this fragment did not correspond to the virus genomic sequences: the third ORF (ORF3) begins immediately after primer P2-start rather than the fifth ORF (ORF5), as could be expected on the basis of the scheme depicted in Fig. 2 (on the top). Thus, we showed that splicing of the primary transcript synthesized from promoter P2 results in excision of a genomic region containing ORF5, which encodes the NS3 nonstructural protein. The viral RNA formed as a result of splicing can code only nonstructural proteins NS1 (ORF3) and NS2 (ORF4). Then, all additional RT-PCR products obtained in this reaction (Fig. 3, lane *3*) were cloned and sequenced from flanks. These experiments revealed three splice variants of mRNA for NS proteins and one common pre-mRNA for them. The fragment approximately 1600 bp long, corresponding to the brightest band on lane *3*, is the first major splice variant of NS-mRNA; the lower band corresponds to the 800-bp fragment, which is the second splice variant.

We also cloned the smallest cDNA fragment (500 bp long), which is not seen in lane *3*. It appeared that this fragment corresponded to the third mRNA, in which two viral genome regions undergo splicing. All these mRNAs are read from the same promoter (P2) and are terminated at the end of ORF3 of the virus genome. In Fig. 2, these four mRNAs are shown to the right of the scheme (nos. 3–6). Protein NS3 (216 aa), encoded by ORF5, can be translated from the first pre-mRNA (2200 nt). The full-length protein NS1 (530 aa) can only be translated from mRNA no. 4. This is the most conservative protein of densoviruses. It contains the so-called Parvo-NS1 domain, which has several functions, including ATP-dependent DNA helicase and site-specific nickase activities, which are required for replication of all parvoviruses [12]. In the BgDNV densovirus, this domain is located between amino acid residues 343 and 500 (this region is encoded by ORF3). Therefore, shorter mRNAs (nos. *5* and *6*) may probably also encode functional splice variants of protein NS1. The NS1 protein is the best studied NS protein of densoviruses. However, little is known on two other noncapsid proteins—NS2 (262 aa) and NS3 (216 aa), which are encoded by ORF4 and ORF5, respectively. Interestingly, these proteins have no homology with NS proteins of vertebrate parvoviruses or with other protein sequences deposited in databases. It is only known that NS3 is also required for replication of viral DNA and production of viral particles [13]. The fact that NS3 is highly conservative confirms its important role in the virus life cycle. However, the mRNA encoding this protein is transcribed much less efficiently than the mRNA for the NS1 protein, as showed our results obtained by Northern blot hybridization and RT-PCR.

The presence of ORF6, encoding a protein comprised of 131 aa (Fig. 2), which is located near ORF1 but is oriented in the opposite direction, has not been shown earlier for any densovirus. It was of interest to determine its function and to test whether this nucleotide sequence is contained in any viral RNA. To answer these questions, we compared the amino acid sequence of the fragment encoded by ORF6 with the sequences deposited in the BLAST database but found no significant homology with any eukaryotic proteins.

However, the program revealed a weak homology with a bacterial protein functioning as a transcription regulator (GenBank acc. NP_106080.1). To determine the origin of long additional transcripts approximately 4000 bp long, shown in Fig. 1b, we decided to check the sizes of RT-PCR products obtained with the use of pairs of primers that comprise ORF6. Lanes *5* and *6* in Fig. 3 show two high-molecular-weight fragments 3800 and 3100 bp long (primers P2-start and ORF6-end); lanes *7* and *8*, only one fragment 3100 bp long (primers ORF3-start and ORF6-end), which were cloned and sequenced.

As expected, the start of the 3800-bp fragment coincided with the sequence of the P2-start fragment, which was followed by ORF5, and the end of this fragment coincided with the ORF6-end primer. The 3100-bp fragment also started with the sequence of the P2-start primer, which was followed by the ORF3 sequence, as in the case of NS-mRNA no. 4 in Fig. 2. The end of this fragment also coincided with the end of ORF6, indicating that the transcription from the P2 promoter may continue to this site in the genome. We assumed that an alternative polyadenylation site for NS-mRNA may be located at the end of ORF6. This assumption can be confirmed by the results of determination and sequencing of 3800- and 3100-bp transcripts, which are shown on the bottom of the scheme in Fig. 2 (nos. *7* and *8*, respectively), which will be the subject of our further studies.

The start and end transcription sites of viral RNAs were determined by 3'- and 5'-RACE using the SMART-RACE cDNA reagent kit (Clontech) using the protocol provided by the manufacturer. We determined the transcription start site from the P1 promoter in the virus genomic region corresponding to nucleotide 240 and from the P2 promoter in the genome region corresponding to nucleotide 5096. We detected no transcripts starting from the putative promoter P3; apparently, this promoter lacks functional activity. Similarly, we determined the 3' ends of the transcripts, which proved to be located on nucleotide 2842 in the case of VP-cDNA and on nucleotide 2795 in the case of NS-cDNA.

Interestingly, VP- and NS-transcripts overlap by 48 nucleotides at their 3' ends, which may result in the formation of double-stranded viral RNAs (dsRNAs) in the host cells and induce defense response of the infected cell to the viral infection by the RNA interference mechanism [14]. If the transcription from the P2 promoter continued to the end of ORF6, the transcripts of the BgDNV densovirus will overlap by over 1500 nucleotides, which may be important in the regulation of the life cycle of the virus. Currently, the problems of regulation of gene expression associated with the involvement in these processes of overlapping sense and antisense transcripts, which were found in all eukaryotes, are actively discussed. The genes whose transcripts overlap are usually expressed together and regulated by means of antisense RNAs [15]. The overlapping of transcripts, observed during the formation of viral RNAs, may also influence the coordinated expression of capsid and noncapsid proteins of the BgDNV densovirus, making this virus a promising model for studying the mechanisms of regulation of gene expression.

It is noteworthy that our data are of interest for densovirus taxonomy. The virus that was studied in this work has a slight nucleotide and amino acid homology with the PfDNV densovirus of *P. fuliginosa* but differs from the latter in the structural organization of the genome and the expression strategy. The genome of the PfDNV densovirus contains seven major open reading frames (three for capsid and four for nonstructural proteins) [4]. The BgDNV genome contains six open reading frames: two for capsid and three for nonstructural proteins. ORF6 has no homology with any ORF of other densoviruses. The strategy of BgDNV expression has certain similarity with the strategy of expression of other densoviruses; however, it has also some unique features. On the basis of the results of this study, it can be postulated that the German cockroach densovirus BgDNV can be regarded as a representative of an independent genus within its subfamily.

REFERENCES

1. Mukha, D.V., Chumachenko, A.J., Dykstra, M.J., et al., *J. Gen. Virol.*, 2006, vol. 87, no. 6, pp. 1567–1575.
2. Mayo, M.A., *Arch Virol.*, 2004, vol. 150, pp. 189–198.
3. Yamagishi, J., Hu, Y., Zheng, J., and Bando, H., *Arch. Virol.*, 1999, vol. 144, no. 11, pp. 2111–2124.
4. Guo, H., Zhang, J., and Hu, Y., *Acta Virol.*, 2000, vol. 44, no. 6, pp. 315–322.
5. Wei, W., Shao, D., Huang, X., et al., *Am. J. Trop. Med. Hyg.*, 2006, vol. 75, no. 6, pp. 1118–1126.
6. Fediere, G., Li, Y., Zadori, Z., et al., *Virology*, 2002, vol. 292, no. 2, pp. 299–308.
7. Tal, J. and Attathom, T., *Arch. Insect Biochem. Physiol.*, 1993, vol. 22, pp. 345–356.
8. Berns, K.I., *Microbiol. Rev.*, 1990, vol. 54, no. 3, pp. 316–329.
9. Tijssen, P., Li, Y., El-Far, M., et al., *J. Virol.*, 2003, vol. 77, no. 19, pp. 10357–10365.
10. Van Munster, M., Dullemans, A.M., Verbeek, M., et al., *J. Gen. Virol.*, 2003, vol. 84, pp. 165–172.
11. Fediere, G., El-Far, M., Li, Y., et al., *Virology*, 2004, vol. 320, pp. 181–189.
12. Tijssen, P. and Bergoin, M., *Semin. Virol.*, 1995, vol. 6, pp. 347–355.
13. Abd-Alla, A., Jousset, F.X., Li, Y., et al., *J. Virol.*, 2004, vol. 78, no. 2, pp. 790–797.
14. Wang, X-H., Aliyari, R., Li, W.X., et al., *Science*, 2006, vol. 312, pp. 452–454.
15. Munroe, S.H. and Zhu, J., *Cell. Mol. Life Sci.*, 2006, vol. 63, pp. 2102–2118.