



Bioinformatics of NS3 Gene and Inverted Terminal repeats (ITR) of *Bombyx Mori* Parvo-like Virus (China Zhenjiang Isolate)

Chen Sun

Institute of Life Sciences, Jiangsu University
301 Xuefu Road, Zhenjiang 212013, China
E-mail: boystrong001@163.com

Qin Yao (Corresponding author)

Institute of Life Sciences, Jiangsu University
301 Xuefu Road, Zhenjiang 212013, China
Tel: 86-511-879-1923 E-mail: yaoqin@ujs.edu.cn

Huijuan Yin

Institute of Life Sciences, Jiangsu University
301 Xuefu Road, Zhenjiang 212013, China

Yuanqing He

Institute of Life Sciences, Jiangsu University
301 Xuefu Road, Zhenjiang 212013, P. R. China
E-mail: yqhe@ujs.edu.cn

Keping Chen

Institute of Life Sciences, Jiangsu University
301 Xuefu Road, Zhenjiang 212013, China
E-mail: kpchen@ujs.edu.cn

The research is financed by the National Natural Science Foundation of China (30871826), China National "863" Project (No. 2008AA10Z145) and grants from Jiangsu Sci-Tech Support Project – Agriculture (No. BE2008379).

Abstract

Bombyx mori Parvo-like virus (China Zhenjiang isolate) had termed as *BmDENV-Z*. NS3 is one of the three nonstructural proteins that is responsible for the replication of virus DNA. To explore the nature of NS3 and determine how it differs from other NS3 sequence, we extracted amino acid sequences of NS3 from four sequenced DNV genomes, and carried out sequence-based phylogenetic and structural analyses. In addition, to investigate the importance of the ITR as a signal for some viral and /or cellular factors further, we performed a computer analysis for prediction of secondary structure. We found that Comparisons of protein sequence with those of the databases showed low homologies with NS-3 of JcDENV (30% identity), GmDENV (26% identity), MIDENV (26% identity). But zinc-finger motifs appear to be conserved greatly. Further the structural importance of the terminal sequence (CTS) common to VD1 and VD2 was also predicted by a DNA folding web.

Keywords: *BmDENV-Z*, NS3, Bioinformatics, ITR

1. Introduction

The family Parvoviridae includes two subfamilies: the Parvovirinae, members of which infect vertebrates, and the Densovirinae, members of which infect invertebrates (Berns 1990). Similar to other parvoviruses, densoviruses (DNVs) are small, nonenveloped particles of 18–26 nm diameter. Each virion contains two 4–6.5 kb single-stranded complementary DNA molecules, which are termed plus and minus strands, respectively. These strands will anneal when it is extracted in buffers under high salt concentrations. DNVs can infect a wide range of insect tissues and lead to death of the host in most cases; however, they cannot infect tissues of vertebrates, thus making them potential biological agents for controlling main agricultural pests (El-Far, Li et al. 2004).

Six DNV isolates have been obtained from silkworm and these six isolates are classified into two groups designated as Bombyx DNV type I and II (*Bm*DNV I and *Bm*DNV II) on the basis of their serological characteristics and genome structure (Lü 1998). *Bm*DNV I, such as Ina isolate, has a monosense genome of a little over 5 kb. *Bm*DNV II (Saku isolate, Yamanashi isolate and China isolate *Bm*DNV-Z), has two sets of genome (VD1 VD2) which are enveloped in different virions, respectively (Tijssen P 1995). Both DNAs of *Bm*DNV II have imperfect inverted terminal repeats, i.e. 270 nts for VD1 and about 530 nts for VD2, surprisingly; both are unable to form terminal hairpins. These unusual properties imply a unique replication mechanism different from the rolling circle replication of other densoviruses or vertebrate parvovirus. Furthermore, the genome of *Bm*DNV II is able to encode DNA polymerase itself (Bando, Kusuda et al. 1987; Bando, Kusuda et al. 1987; Bando, Choi et al. 1992; Bando, Hayakawa et al. 1995).

The terminal sequences of these viral genomes are believed to be important for specific interactions between the viral genome and the viral RNA polymerase and/or contribute to the encapsidation of the genome by the nucleocapsid protein. Furthermore, this structural feature has been observed in single stranded DNA virus, parvovirus, though the significance of the structure is still obscure. It must be noted here that all of parvovirus genomes which have been reported so far contain a terminal palindromic sequence of 50 to 70 nucleotides which play important roles during replication. However, the terminal palindrome could not be found in VD1 and 2.

Our groups had completed the genome analysis of *Bm*DNV-Z in 2005 (Genbank Accession Number: DQ017268; DQ017269). Sequence analysis showed that VD1 genome consisted of 6,543 nts and VD2 genome consisted of 6,022 nts (Wang, Yao et al. 2007). However, effects of the gene and replication mechanisms are still unclear. In this study, we collected amino acid sequences from five DNV genomes, followed by systematical analysis at three different levels with the GENEDOC and MEGA tools, including their overall phylogeny, domain structures, and identifiable motifs. After that, we predicted the secondary structure of ITR by (<http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/dna-form1.cgi>). The aim of this study is to infer the features and evolution of the NS3 of densoviruses using a phylogenetic approach, to provide information that could render some insights regarding NS3 and ITR functions during viral DNA replication.

2. Materials and methods

2.1 Protein sequence and Phylogenetics analysis

We used Inter Proscan (<http://www.ebi.ac.uk/InterProScan/index.html>) to analysis amino acid sequence. GENEDOC tool was also used to align amino acid sequence of NS3 from five DNV. We used the deduced amino acid sequences to reconstruct phylogenetic trees. The trees were constructed by using the neighbor-joining method (NJ) with JTT distances. The reliability of internal branches was assessed by using 1,000 bootstrap replicates, and sites with gaps were ignored in this analysis. NJ searches were conducted by using the computer program MEGA3.

2.2 Secondary structural analysis of ITR

We predicted the secondary structure of ITR by:

(<http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/dna-form1.cgi>.)

3. Results

3.1 Characterization of ORF2

ORF2 was found to span from nt 4761 to nt 5429, and the coding sequence of NS3 consisted of 222 aa residues starting at the ATG initiation codon at position 4761 and ending at the TAA stop codon at position 5429. This sequence presents some interesting features including two zinc-finger motifs [C:(X)2:C:X3:C (aa 152 to 173) and C:(X)2:C:H:X:C (aa 198 to 214)], six putative N glycosylation sites (NWSK/NRTT/NLST/NDSN/NNSN/NDSN, aa 74 to 77, 100 to 103, 130 to 133, 137 to 140, 164 to 167 and 219 to 222), and four putative phosphorylation site SvFD, TypD, SteE, SinD (aa 68 to 71, 103 to 106, 132 to 135 and 217 to 220). Comparisons of this sequence with those of the databases showed low homologies with NS-3 of MIDNV (26% identity), GmDNV (26% identity) and *Diatraea saccharalis* DNV (26% identity) (Fig. 1). Curiously, the DsDNV NS3 and *Bm*DNV-Z NS3 sequence is truncated of the 34 C-terminal amino acid sequence common to NS3 of JcDNV, MIDNV, and GmDNV. No further significant homologies could be detected with NS polypeptides of Densovirinae and Parvovirinae or with any protein of the data banks.

3.2 Phylogenetic analysis

To analyse the NS3 evolution in virus, we tried several methods (neighbor-joining, minimum evolution, and maximum parsimony) to assess the phylogeny of NS3 and achieved similar results. Based on the phylogenetic analysis, GmDENV, JcDENV and MIDENV, clustered together, indicating a common origin. However, *BmDENV-Z* also showed a low of relatedness to other. These results suggested a possible different origin for NS3.

3.3 Secondary structural analysis of ITR

Assuming that both virus DNAs (VD1 and 2) are necessary to complete replication of *BmDENV-Z*, it is not difficult to explain that VD1 and 2 share a common terminal sequence of 53 nucleotides(CTS)[Bando 1992], since all of the segmented genomic DNAs must contain the signals for recognition by a replication and /or for encapsidation by the viral components. To investigate the importance of the CTS as a signal for some viral and /or cellular factors further, we performed a computer analysis for prediction of secondary structure (<http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/dna-form1.cgi>.) Figure 3 shows the predicted structure of the terminal 150 nucleotides of VD1 and VD2 as their most stable secondary form. In the predicted structure (Fig. 3), CTS seems to be divided into 4 structural domains, A, B, C, and D. The structural importance of CTS was emphasized by the predicted localization of a Sp1-binding consensus motif as an exposed element on the stem-loop constituted by C and D. Moreover, D containing a Sp1-binding consensus motif also fold into a stem structure with a sequence out of the CTS which contains another Sp1-binding consensus motif, and the stem structure was also conserved in both VD1 and VD2

4. Discussion

The high level of sequence conservation of NS3 protein among the members of the Densovirus (Fig. 1) suggests that the biological function(s) of this protein is essential for the accomplishment of their life cycle. Zinc-finger motifs indicated the involvement of NS3 in the viral DNA replication. Our results constitute a first step toward the understanding of NS3 function(s) during the *BmDENV-Z* life cycle. However, the precise level at which this polypeptide acts and its specific biological function(s) remain to be elucidated. The lack of significant homology between the NS3 amino acid sequence and proteins of the data banks, including NS proteins of parvoviruses, implies that its function is either specific to this type of virus, i.e., of the genus *Bidensovirus*, or that it is somehow related to function(s) of the host cell necessary for their replication. At present, it is not possible to decide in favor of one of these hypotheses, which are not exclusive of one another. These differences might reflect specific requirements for cellular proteins functioning in partnership with NS and (or) VP proteins in order to achieve the viral replicative cycle, as has been recently demonstrated for vertebrate parvoviruses.

The representative of the genus *Iteravirus*, *BmDENV-1*, which lacks NS3, shares with the members of the genus *Densovirus* the property of having as natural hosts exclusively lepidopteran insects. His genome was only 5 kb in length, with J-shaped ITRs of 230bp. Moreover, the NS and VP genes were located on the same strand, as in the vertebrate parvoviruses. VP was located downstream of the NS ORF(Li, Zadori et al. 2001). *BmDENV-Z* VD1 genome consisted of 6 543 nt, including inverted terminal repeats (ITRs) of 224 nt. In the viral genome, three open-reading frames (NS ORF and VP) in the plus strand and one major ORF (ORF4) codes for DNA polymerases in the complementary strand were identified(Wang, Yao et al. 2007). The essential difference between *BmDENV-1* and *BmDENV-Z* VD1 at the genome level lies in ORF4 DNA polymerases. Computer analysis suggests that VD2 has not enough information to produce progeny virions by itself. So we predicted *BmDENV-Z* VD1 might have diverged from *BmDENV-1* and it obtained DNA polymerases from host cell during further evolution. VD2 come from other parvo-like maybe help infection and replication of VD1, while VD1 can provide necessary nonstructural proteins and VP proteins. Whether the *BmDENV-Z* is the mixture of two parvo-like viruses or a virus with a multipartite genome need further identification.

References

- Bando, H., H. Choi, et al. (1992). Structural analysis on the single-stranded genomic DNAs of the virus newly isolated from silkworm: the DNA molecules share a common terminal sequence. *Arch Virol*, 124(1-2): 187-93.
- Bando, H., T. Hayakawa, et al. (1995). "Analysis of the genetic information of a DNA segment of a new virus from silkworm. *Arch Virol* 140(6): 1147-55.
- Bando, H., J. Kusuda, et al. (1987). Organization and nucleotide sequence of a densovirus genome imply a host-dependent evolution of the parvoviruses. *J Virol*, 61(2): 553-60.
- Bando, H., J. Kusuda, et al. (1987). Molecular cloning and characterization of Bombyx densovirus genomic DNA. Brief report. *Arch Virol* 93(1-2): 139-46.
- Berns KI. (1990). Parvoviridae and their replication. *Virology*, Volume 2. New York: Raven Press, pp.1743–1763.
- El-Far, M., Y. Li, et al. (2004). Lack of infection of vertebrate cells by the densovirus from the maize worm *Mythimna loreyi* (MIDENV). *Virus Res*, 99(1): 17-24.

Lü, H. (1998). *Molecule Biology of Insect Virus*, China Agriculture Science and Technology Publishing Company Press.

Li, Y., Z. Zadori, et al. (2001). Genome organization of the densovirus from *Bombyx mori* (BmDENV-1) and enzyme activity of its capsid. *J Gen Virol* 82(Pt 11): 2821-5.

Tijssen P, B. M. (1995). Densonucleosis Viruses Constitute an increasingly diversified subfamily among the parvoviruses. *Virology* 6: 347-355.

Wang, Y. J., Q. Yao, et al. (2007). Characterization of the genome structure of *Bombyx mori* densovirus (China isolate). *Virus Genes* 35(1): 103-8.

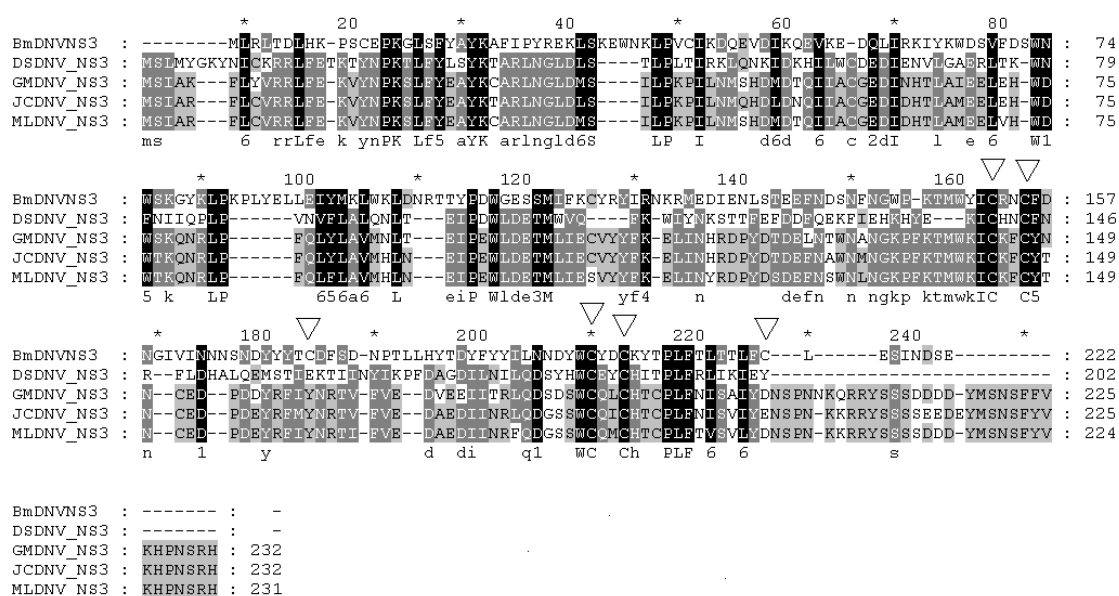


Figure 1. Alignment of amino acid sequence of NS3 from five DNV by using the Gene doc program Amino acid numbers begin with the start codon. Identical amino acids among the NS3s are marked by asterisks (*). Putative zinc-finger motifs are marked by ▽. Abbreviations: BmDENV-Z(GenBank accession number DQ017269), *Bombyx mori* densovirus (China isolate); JcDENV(GenBank accession number NP_694826), *Junonia coenia* densovirus; GmDENV (GenBank accession number NC_004286), *Galleria mellonella* densovirus; And DsDENV (GenBank accession number NC_001899), *Diatraea saccharalis* densovirus; MIDNV(GenBank accession number NP_958098), *Mythimna loreyi* densovirus

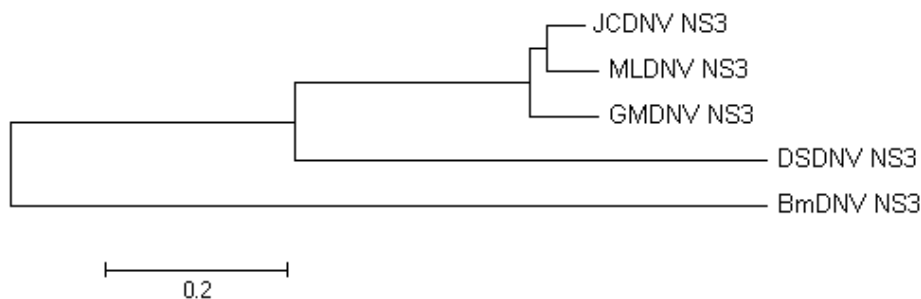


Figure 2. Phylogenetic relationship of the NS3 groups. Numbers in bootstrap consensus tree indicated degree of confidence. Sequences above are from GenBank

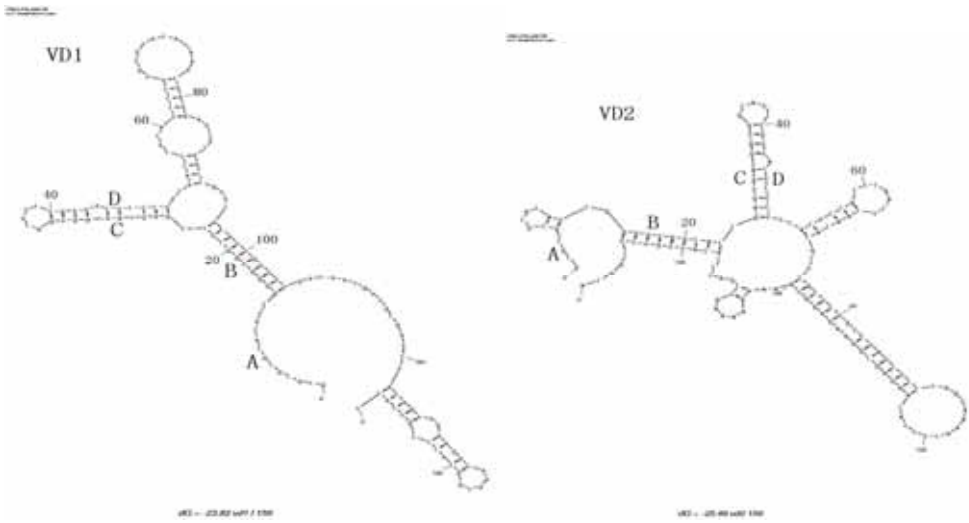


Figure 3. Predicted secondary structure of the terminal regions of VD1 and VD2.