COMPLETE SEQUENCE AND ORGANIZATION OF PERIPLANETA FULIGINOSA DENSOVIRUS GENOME

H. GUO, J. ZHANG, Y. HU*

Institute of Virology, Wuhan University, Wuhan, 430072, P.R. China

Received August 9, 2000; accepted September 15, 2000

Summary. – The replicative form (RF) of the Periplaneta fuliginosa densovirus (PfDNV) genome was cloned and its complete nucleotide sequence was determined. The PfDNV genome is 5454 nucleotides (nt) in length with distal 201-nt long inverted terminal repeats (ITRs). The first 122 nt at the 5′-end and the terminal 122 nt at the 3′-end of both strands are palindromes with identical sequences, of which each can fold into a typical U-shaped hairpin structure. The coding regions of PfDNV genome are evenly distributed in the 5′-halves of both strands. PfDNV genome contains seven major open reading frames (ORFs), four of which on the plus DNA strand may encode non-structural (NS) proteins, while the others on the minus DNA strand may encode structural proteins. Two potential functional promoters (P_3 and P_{97}) were found within the ITRs on both strands. The ORF2 polypeptide contains a highly conserved NTP-binding domain of NS proteins of parvoviruses. The ORF5 polypeptide has significant homology to the conserved PGY region of coat proteins of parvoviruses. The ORF6 polypeptide has distinct homology to the structural polypeptides of insect parvoviruses. The organization of PfDNV genome was compared to those of other parvoviruses.

Key words: Periplaneta fuliginosa densovirus; parvovirus; nucleotide sequence; amino acid sequence; genome organization

Introduction

All small (diameter of 18–26 nm) icosahedral non-enveloped viruses containing a monopartite linear single-stranded (ss) DNA genome (4–6 kb), belong to the *Parvoviridae* family (Van Regenmortel *et al.*, 2000). The members of the *Parvoviridae* family are further referred to in this article as "parvoviruses".This family is divided into two subfamilies, *Parvovirinae* (containing "vertebrate parvoviruses") and *Densovirinae* (containing "insect parvoviruses"). The *Densovirinae* subfamily includes three genera: *Densovirus*, *Iteravirus*, and *Brevidensovirus*.

The coding sequences of vertebrate parvoviruses are located on the same strand (viral or minus strand) and consist of two large non-overlapping ORFs, the left and the right one, coding for NS and structural proteins, respectively (Berns, 1990; Rhode and Iversen, 1990). The terminal non-coding sequence of the genome of all vertebrate parvoviruses can fold into palindromic, Y- or T-shaped structures (Astell *et al.*, 1983, 1985; Rhode and Paradiso, 1983; Srivastava *et al.*, 1983; Chen *et al.*, 1986; Shade *et al.*, 1986; Reed *et al.*, 1988; Bloom *et al.*, 1988, 1990; Ranz *et al.*, 1989; Vasudevacharya *et al.*, 1990; Deiss *et al.*, 1990; Kariatsumari *et al.*, 1991). Furthermore, adeno-associated viruses (AAVs)

*E-mail: wueml@whu.edu.cn; fax: +8627-87882661.

Abbreviations: aa = amino acid; AaDNV = *Aedes aegypti densovirus*; AlDNV = *Aedes albopictus densovirus*; AAV-2 = *Adeno-associated virus 2;* AAVs = adeno-associated viruses; B19V = *B19 virus;* BmDNV = *Bombyx mori densovirus*; bp = base pair; BPV = *Bovine parvovirus*; CPV = *Canine parvovirus*; DsDNV = Diatraea saccharalis densovirus; GmDNV = *Galleria mellonella densovirus*; ITRs = inverted terminal repeats; JcDNV = *Junonia coenia densovirus*; m.u. = map unit; MEV = *Mink enteritis virus*; NS = nonstructural protein; nt = nucleotide; ORF = open reading frame; PfDNV = Periplaneta fuliginosa densovirus; PPV = *Porcine parvovirus*; SDS-PAGE = polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate

and *B19 virus* (B19V) have ITRs. The genomic organization of insect parvoviruses has been reported to show obvious differences among individual viruses and in comparison to vertebrate parvoviruses (Bando *et al.*, 1987, 1990; Afanasiev *et al.*, 1991; Dumas *et al.*, 1992; Boublik *et al.*, 1994). As a type species of the *Iteravirus* genus, *Bombyx mori densovirus* (BmDNV) has two major ORFs located on the viral (minus) strand and a minor ORF located on the complementary (plus) strand. The left ORF (ORF1) encodes NS proteins while the right ORF (ORF2) encodes four capsid proteins. Whether the minor ORF is functional is unknown. Half of BmDNV virion population contains plus DNA strands and the other half minus DNA strands. *Aedes aegypti densovirus* (AaDNV) belongs to the *Brevidensovirus* genus. Its genome has three ORFs on the plus strand; the left and the mid ORFs (within the left ORF) encode two NS proteins, while the right ORF encodes two capsid proteins. There is also a minor ORF of unknown function on the minus strand. *Aedes albopisctus densovirus* (AlDNV), an insect parvovirus obtained from mosquitoes, shares 77.3% nucleotide sequence homology and 73%–78% amino acid sequence homology to AaDNV. Organization of both genomes is similar except that the minus DNA strand of AlDNV does not contain any potential ORF. The minus DNA strand is preferentially (by about 90%) encapsidated into virions in AaDNV and AlDNV. *Junonia coenia densovirus* (JcDNV), the type species of the *Densovirus* genus, reveals a unique genomic organization among parvoviruses, in that its coding sequences are evenly distributed in the 5′-halves of both DNA strands. On one strand, a major ORF (ORF1) encodes four structural proteins. On the complementary strand, three ORFs encode NS proteins. JcDNV encapsidates DNA strands of both polarities with equal frequency. The terminal structures of genomes of these insect parvoviruses are also diverse. Like AAVs and B19V genomes, BmDNV and JcDNV genomes possess also ITRs that can form panhandle structures. The BmDNV ITRs are 225-nt long and the complementary sequences in each terminus (175 nt) can generate a typical U-shaped hairpin structure. JcDNV ITRs are much longer (517 nt) and the first 96 nt of one extremity can fold into an Y-shaped hairpin structure. The AaDNV and the AlDNV genomes have no ITRs, their terminal sequences can form typical T-shaped hairpin structures. Those terminal structures are very important in the DNA replication process of these viruses (Astell *et al.*, 1985, 1990; Berns, 1990).

PfDNV is an agent causing densonucleosis of smoky-brown cockroach. It can infect its larvae and adults in which it replicates in almost all viscera except for midgut. The virions (22 nm in diameter) contain linear ssDNA of both polarities (1:1) and five structural proteins (VP1: 52 K, VP2: 56 K, VP3: 79 K, VP4: 82 K, and VP5: 105 K) (Hu *et al.*, 1991, 1994). This virus was classified as a tentative species of the genus *Brevidensovirus* of the subfamily *Densovirinae* of the family *Parvoviridae* (van Regenmortel *et al.*, 2000). In order to further investigate the molecular biology of PfDNV, its complete genome nucleotide sequence was identified and its genomic organization was compared with those of other parvoviruses.

Materials and Methods

Enzymes and reagents. Restriction endonucleases were purchased from Boehringer Mannheim Biochemicals or Takara Dalian Co., Ltd. The Klenow fragment of *E.coli* DNA polymerase I, T4 DNA ligase, vector pUC119, *E.coli* DH5α, and the *Exo*III-S1 Deletion Kit were purchased from Takara Dalian Co., Ltd. Oligonucleotides were synthesized by Sangon. The ABI PRISM® BigDye™ Sequencing Kit was purchased from Perkin-Elmer.

Virus. PfDNV was propagated in smoky-brown cockroach 4–5th instar larvae.

Virus and RF DNA purification. The virions and RF DNA were purified according to Hu *et al.* (1994).

Cloning of complete PfDNV genome. PfDNV RF DNA was inserted into the *Pst*I site of pUC19 vector. The recombinant vector PfDNV-pUC119 was obtained from *E. coli* by standard techniques (Sambrook *et al.*, 1989; Guo *et al.*, 1999).

DNA sequencing. The sequence of genomic PfDNV DNA was established from the recombinant PfDNV-pUC119 vector. A series of viral DNA fragments with approximately 400-bp interval were obtained by deleting the recombinant vector with the *Exo*III-S1 Deletion Kit (Guo *et al.*, 1999). The sequence was determined by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) in the Applied Biosystems Model 377 Sequencing System with universal or synthetic sequencing primers.

Sequence analysis. ORF and motif sequence search, analysis of secondary structure of ssDNA, sequence comparisons and homology search were performed using the DNASIS, DNASTAR, RNASTRUCTURE and BLAST softwares.

Results

Nucleotide sequence of PfDNV DNA

The complete nucleotide sequence of the genome of PfDNV was 5454 nt in length (it is accessible at the Genbank database under Acc. No. AF192260). The base composition of the plus DNA strand of the genome was as follows: 31.93% of A, 19.27% of C, 18.84% of G, and 29.94% of T; the G+C content was 38.13%. Both termini of PfDNV complementary strands contain 122-nt long palindromes which can fold over to form an U-shaped hairpin structure (Fig. 1). But in contrast to other reported parvoviruses these palindromes do not contain any internal palindromic sequences, so there are no Y- or T-shaped structures in this region. The sequence analysis of the extremities revealed

Fig. 1
Secondary structure of the 122 nts of the left-hand extremity of PfDNV genome
"Flip" orientation (A), putative "flop" orientation (B).

that the terminal 201 nt are ITRs except for 5 mismatch nucleotides.

ORFs

Analysis of PfDNV genome revealed 7 major ORFs on both DNA strands. Four ORFs (ORF1-ORF4) on the plus strand encode 33 K, 62 K, 13 K, and 31 K proteins. The minus strand contains 3 large ORFs (ORF5-ORF7) encoding 31 K, 63 K and 20 K proteins. In ORF5, the third in-frame ATG (nt 5253) would initiate a primary translation product of 26 K. All the ORFs on both strands cluster on their 5′-extremities and exist strand overlapping.

Potential promoter domains and poly(A) sites

Seven potential promoter domains were found according to the characteristic of an eucaryotic promoter as described by Bensimhon *et al.* (1983) (Table 1). The latter includes an enabling sequence located 100 bp upstream from the cap site, a GC-rich activator sequence located 50–75 bp upstream from the cap site, and a TATA box located 30 bp upstream from the cap site. Promoter domains of P_3 (m.u.3) and P_{97} fulfill these criteria and are located similarly to the functional promoter of JcDNV genome(Dumas *et al.*, 1992).

Among other potential promoter domains within ORF1, ORF2, ORF5, and ORF6 (m.u. 6, 36, 93, and 73), P_{93} located within ORF5 was found 543 nt upstream from the first ATG codon; the first in-frame ATG codon (nt 2056 and nt 4809) of P_{36} and P_{73} is followed by two or three ATG codons, which terminate in other frames. Thus, unless there is splicing, ribosomes would have to bypass several upstream ATG codons to reach the first in-frame ATG. No large ORFs exist downstream from P_{20}. The first downstream ATG (nt 335) of P_6 can initiate translation of a polypeptide lacking six N-terminal amino acids as compared to that encoded by ORF1. In the absence of precise data about the transcription of PfDNV genome it is not possible at present to conclude which, if any, of these promoters are functional.

Three AATAAA boxes at nt 2732, 3784, and 4952 are present in the plus DNA strand of PfDNV genome. The second box downstream from the coding region is the only one to have a CAYTG sequence (nt 3816) around the AATAAA box and the G/T-rich sequence (nt 3819), both of which are necessary to fulfill the criteria of eucaryotic transcription terminators (Birnstiel *et al.*, 1985). Thus the second box appears to be the most likely one to act as potential polyadenylation site. The first box (nt 2732) is devoid of the surrounding consensus sequences. Interestingly, the TAA stop codon of ORF2 is part of the AATAAA box at nt 2734. The third box (nt 4952) lacks the surrounding consensus sequences and is almost 2000 nt downstream from the coding region. So it may be non-functional.

Six AATAAA boxes were found in the minus DNA strand. The first one (nt 4678) within ORF5 is probably non-functional; the second one (nt 4103) located within ORF7 lacks the CAYTG sequence and G/T-rich clusters; the third

Table 1. Analysis of potential promoter sequences

m.u.	Enabling sequence	Activator sequence	TATA box
Plus strand			
3	GGTGGGGT (96[a],-102[b])	GGTGGGGG (151,-47)	TATAAAA (168)
6	GGGGTGTA (272, -93)	GGGGTGTA (302,-63)	TATAAAA (335)
36	CCGCGAAGA (1902, -87)	CCCGTGTG (1936,-53)	TATATCT (1959)
Minus strand			
20	CCCGGCAT (1215,-127)	CTGGCTGC (1141,-53)	TATATCA (1118)
73	GGAGGAGC (4085,-122)	GAGGGAAC (4034,-69)	TATAAAT (3993)
93	GGGGACTC (5144,-118)	GCGTGAGC (5104,-78)	TATAGAA (5056)
97	GGTGGGGT (5359,-102)	GGTGGGGG (5304,-47)	TATAAAA (5287)

[a]Position of the first nucleotide in the genome.
[b]Position relative to the cap site.

Table 2. Location of 5′-donor and 3′-acceptor splicing sites

5′-donor site		3′-acceptor site	
Location (nt)	Sequence	Location (nt)	Sequence
Plus strand			
2491	CAGGTGGTG	3290	TCTGAACAGGT
2830	AAGGTAGTG		
3562	CAGGTGGAC		
Minus strand			
4533	CAGGTACGT	4399	CCTCTTTCAGGT
4400	CAGGTGACG	4180	TCTCACTCAGGT
4181	CAGGTGTAC		
4118	CAGGTGGTG		

Table 3. Putative polypeptides generated by spliced mRNAs transcribed from potential P3 and P97 promoters

5′-donor site (nt)	3′-acceptor site (nt)	Intron size (nt)	Size of putative protein (K)
Plus strand			
2491	3290	798	51, 31, 33
2830	3290	459	62, 31, 33
Minus strand			
4533	4399	133	85
4533	4180	352	77
4400	4180	219	26, 55

one (nt 3644), although being located within ORF6, has downstream G/T clusters and thus may act as poly(A) signal for ORF5 and ORF7; the fourth one (nt 2770) shares the TAA sequence with the termination codon of ORF6, but it lacks those two additional factors. The remaining two boxes (nt 2095 and nt 1111) downstream from ORF6 have downstream G/T clusters but that at nt 2095 has in addition a CAYTG sequence and thus appears potentially functional.

Potential splicing sites

The A/CAGGTA/GAGT and $(C/T)_nXCAGGC/T$ consensus sequences for 5′-donor and 3′-acceptor sites (Mount, 1982) were adopted to search the splicing sites within PfDNV genome. The results are listed in Table 2. Taking in account the two most favorable promoters (P_3 and P_{97}) and the theoretical splicing sites, a large number of transcripts can be generated. The sizes of the putative polypeptides resulting from the splicing are listed in Table 3.

Comparison of PfDNV to other parvoviruses on the basis of amino acid sequences

The putative viral polypeptides encoded by the genome of PfDNV were compared with the NCBI protein data using the BLAST program. As shown in Fig. 2, a significant homology was revealed between the 100-aa long C-terminal stretch of ORF2 of PfDNV genome (aa 379 to 478) and the NS-1 protein of both the vertebrate and invertebrate parvoviruses, but mainly of the latter. This stretch of amino acids corresponds to the highly conserved NTP-binding GKRN domain of NS-1 protein of all vertebrate parvoviruses (Chen *et al.*, 1986). A phylogenetic tree based on this region clearly indicates that insect parvoviruses are more closely related to each other than to vertebrate parvoviruses (Fig. 3). Unexpectedly, ORF2 of PfDNV genome has homology to ORF2 of BmDNV genome at amino acid level; ORF2 of BmDNV genome encodes a structural protein (Bando *et al.*, 1987).

Within the coding regions in the plus DNA strand of PfDNV genome, the deduced ORF1 protein has homology to the NS-3 protein of Diatraea saccharalis densovirus (DsDNV); the deduced ORF4 protein has homology to the

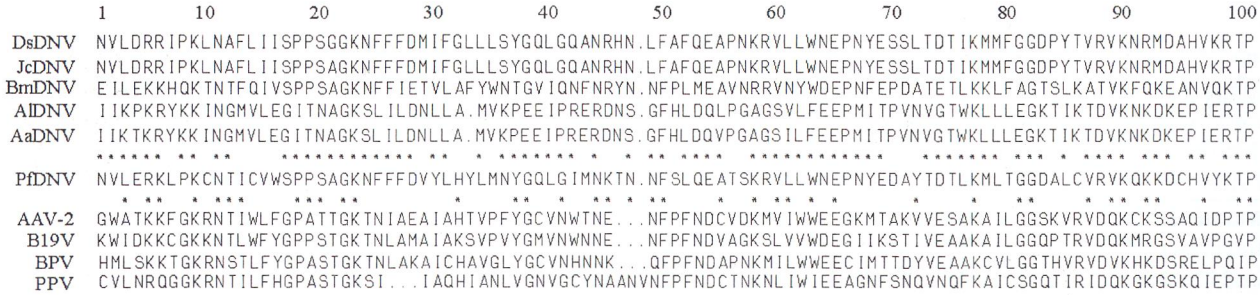

Fig. 2

Amino acid sequence homologies of the deduced proteins of PfDNV ORF2, JcDNV ORF2, BmDNV ORF2, left AlDNV ORF, left AaDNV ORF, DsDNV NS-1 protein, and the highly conserved GKRN region of NS-1 protein of vertebrate parvoviruses AAV-2, B19V, BPV, and PPV

(*) indicate identical residues between PfDNV and other parvoviruses

deduced ORF3 protein of JcDNV genome which may be a NS protein. The smallest ORF3 of PfDNV genome, encoding a putative hydrophobic polypeptide (Fig. 4), was not shown to share any homology to the known proteins of parvoviruses.

Within the coding region of the minus DNA strand of PfDNV genome, the deduced ORF5 protein (aa 148–241) has significant homology to the highly conserved PGY domain of structural protein VP-1 (Chen *et al.*, 1986) of DsDNV, JcDNV, GmDNV (Genbank, Acc. No. L32896), *Adeno-associated virus 2* (AAV-2), MEV (Kariatsumari *et al.*, 1991), PPV and CPV (Reed *et al.*, 1988) (Fig. 5). The deduced ORF6 protein (aa 212–370) of PfDNV shares high homology to the coat protein of JcDNV, DsDNV, and GmDNV (Fig. 6). Unexpectedly, the deduced ORF5 and ORF6 proteins of PfDNV have no homology to the proteins of BmDNV and AaDNV at amino acid level. No homology was found between the deduced ORF7 protein of PfDNV and any known proteins in the database.

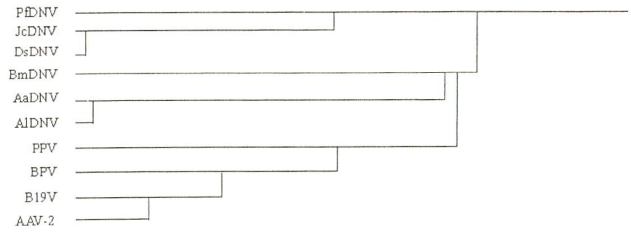

Fig. 3
Phylogenetic tree of insect and vertebrate parvoviruses according to Fig. 2

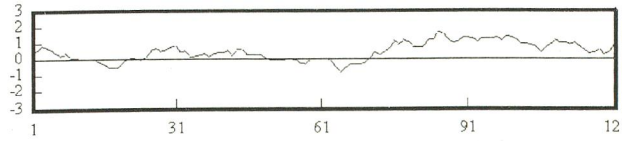

Fig. 4
Hydrophobicity analysis of the deduced ORF3 protein of PfDNV genome

Discussion

Up to date, about 30 insect parvoviruses were found, which constitute an increasingly diversified subfamily (the *Densovirinae* subfamily) among parvoviruses. However, far less data are available about the genome organization of insect parvoviruses as compared to vertebrate parvoviruses (Tijssen and Bergoin, 1995). Identification of the complete nucleotide sequence of PfDNV genome and analysis of genome organization provide direct evidence for categorizing and identifying this virus on molecular level.

PfDNV genome has typical terminal U-shaped hairpin structure at both extremities formed by identical 122-nt long sequences; the terminal 201 nt represent ITRs. The terminal palindromic structure of ssDNA is typical for parvoviruses, it plays a key role in parvovirus DNA replication. Two "rolling circle" models were proposed for replication of DNA of vertebrate parvoviruses and AAVs, respectively (Astell *et al.*, 1985, 1990; Berns, 1990). It is assumed that the 3′-terminus of the virion DNA strand folds back on itself to serve as the primer for initiating DNA synthesis, and both parental and progeny strands are resolved from a replicative intermediate by site-specific nucleases. The hairpin transfer mechanism should generate the "flip" and "flop" configuration at both extremities of the genome. The identical 3′-terminal hairpins of RF DNA of the PfDNV should initiate DNA synthesis at the same rate, which may explain the fact that PfDNV packages both strands at equal frequency. But identical ends are not required for the equal encapsidation of plus and minus strands of LuIII virus DNA (Diffoot *et al.*, 1989), so the mechanism how parvoviruses encapsidate ssDNAs of different polarity is still unclear.

```
          1        10        20        30        40        50        60        70        80        90       100
DsDNV     GLTVPGYKYLGPGNSLNRGPPTNEIDADAKEHDEAYSQSKTAQEVSKADNTFVNKALDHVVNAINLKESPSNTVGAIIGATGIGTKQAIEKHTGVIYPSVSG
JcDNV     GLTVPGYKYLGPGNSLNRGQPTNQIDEDAKEHDEAYDKAKTSQEVSQADNTFVNKALDHIVNAINLKETPGNAFGAAIGAIGIGTKQAIEKHSGVIYPSVSG
GmDNV     GLTVPGYKYLGPGNSLNRGQPINQIDEDAKEHDEAYDKVKTSQEVSRADNTFVNKALDHVVNAINFKETPGNAFGAAIGAIGIGTKQAIEKYSGVIYPSVSG
          *** *   ****** *    *    ** * *** **  ** *  *  **     ** **       *  **      ** **   * ***   *********
PfDNV     GLTYPFHHYLGPGNPLDNNEPVDRDDAIAEEHDKAYANAKSSIDVINAD....KKAIDHFSEDH...EKNGNLHSL.IGKTGLQIKTAIEQRFGVIYPSVSG
          *** *   ****** **  *** * ** * *******
AAV-2     GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYD
PPV       GLTLPGYKYLGPGNSLDQGEPTNPSDAAAKEHDEAYD
CPV       GLVPPGYKYLGPGNSLDQGEPTNPSDAAAKEHDEAYA
MEV       GLVPPGYKYLGPGNSLDQGEPTNPSDAAAKEHDEAYA
```

Fig. 5
Amino acid sequence homologies of the deduced ORF5 protein of PfDNV genome and the PGY conserved region of VP1 capsid proteins of DsDNV, JcDNV, GmDNV, AAV-2, PPV, CPV, and MEV

```
         1        10        20        30        40        50        60        70        80
DsDNV    LTTCLAEIPWQKIPLYMNQSEFDLLPPGSRIVECNVKVIFRSNRIAFET..SSTATKQATLNQISNLQTAVGLNKLGWGIDRSF
JcDNV    ITTCLAEIPWQKLPLYMNQSEFDLLPPGSRVVECNVKVIFRTNRIAFET..SSTATKQATLNQISNLQTAVGLNKLGWGIDRSF
GmDNV    LTTCLAEIPWQKLPLYMNQSEFDLLPPGSRVVECNVKVIFRTNRIAFET..SSTVTKQATLNQISNVQTAIGLNKLGWGINRAF
         *** *** *  *  ******* **** *  *    * *  *     * *  ***    *****  *    ***** * * **
PfDNV    LTTALAEVPVHKPVLYMNQSEYDLLPVGAEVLQVKVSVVQRNALLSFQTNASST..SLATLNQNKNGVYCIGLNKTGYGTNRRY

              90        100       110       120       130       140       150       160
DsDNV    TAFQSDQPMIPTASAPPKYASVSGANGYRGMIADYYGADSNNDIAFGNAGNYPHHQVGSFTFLQNYYCMYIQTERGTGGWPCL
JcDNV    TAFQSDQPMIPTATSAPKYEPITGTTGYRGMIADYYGADSTNDAAFGNAGNYPHHQVGSFTFIQNYYCMYQQTNQGTGGWPCL
GmDNV    TAFQSDQPMIPTATTAPKYEPVTGDTGYRGMIADYYGADSTNDTAFGNAGNYPHHQVSSFTFLQNYYCMYQQTNQGTGGWPCL
         ***     ***    ** ** *  * ** **  *  *   **  *       * ****  * * ** **  **   ***** *
PfDNV    TAFNATEKMIPEKCGPPVYAAV..AEGYEGMLEDLYGY..NNNVASFVTSL.PKHQVGMYTTLKNYFCM.TQTSLYTGGWPNL
```

Fig. 6
Amino acid sequence homologies of the deduced ORF6 protein of PfDNV genome and the VP1 proteins of DsDNV, JcDNV and GmDNV

When a recombinant plasmid containing a deletion within the 201-bp long PfDNV ITRs was transfected into the host body, virus identical to the wild type could be rescued (Guo H., unpublished data), which implies that a correction mechanism similar to AAVs exists (Samulski *et al.*, 1982, 1983, 1987): the 3′-hairpin structure serving as a self-primer and the intact extremity acting as a template for repair of the other, altered end within ITRs.

According to the Seventh Report of the ICTV, PfDNV is a tentative species of the *Brevidensovirus* genus (Van Regenmortel *et al.*, 2000). We have shown similarities and differences between genomes of PfDNV, AAV-2 (taken as a representative vertebrate parvovirus), and JcDNV, BmDNV, and AaDNV (taken as type species of *Densovirus*, *Iteravirus*, and *Brevidensovirus* genera, respectively) (Fig. 7). Both strands of PfDNV genome have nearly equal capability to encode proteins. The coding regions span evenly in the 5′-half of each DNA strand. This phenomenon is similar in JcDNV (Dumas *et al.*, 1992), GmDNV (Gross and Tal, 2000), and DsDNV, but different in other densoviruses, such as BmDNV (Bando *et al.*, 1987), AaDNV (Afanasiev *et al.*, 1991), and AlDNV (Boublik *et al.*, 1994), whose coding regions are exclusively or predominantly located on one DNA strand.

The analysis of potential promoter domains and poly(A) sites suggested that two major mRNAs were transcribed from

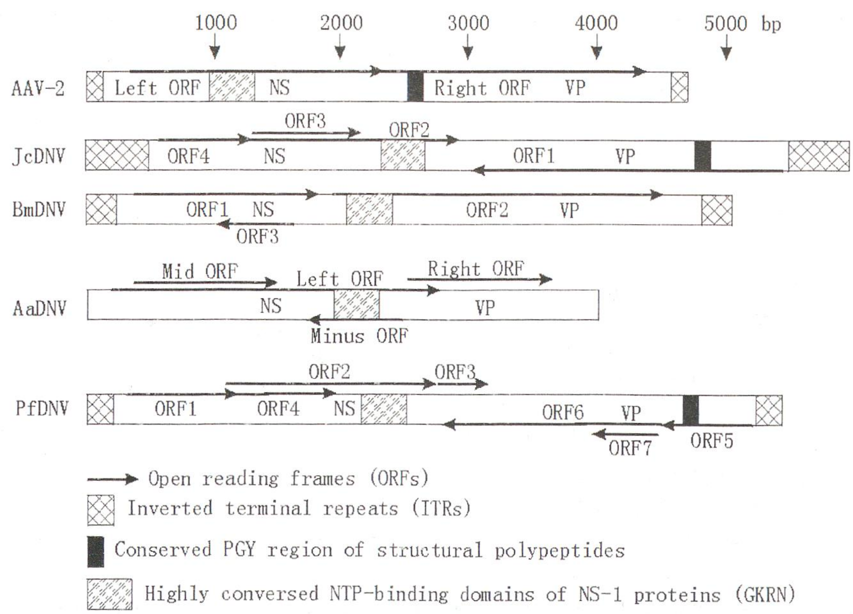

Fig. 7
Organization of the genomes of PfDNV, JcDNV, BmDNV, AaDNV, and AAV-2

the plus and minus strand of PfDNV genome, respectively, similarly to JcDNV (Dumas *et al.*, 1992) and GmDNV (Gross and Tal, 2000). It is worth to mention that both promoter domains are located symmetrically in the ITRs of PfDNV genome. These two major mRNAs might re-assemble by potential splicing but without frame shifting. For example, the potential splicing of the 5′-donor (nt 4533) and the 3′-acceptor (nt 4399) in the minus DNA strand would result in a new ORF connecting ORF5 to ORF6 in frame, which would encode an 85 K protein. It implies that PfDNV genome has a potential to be post-transcriptionally modified. Parvoviruses extensively utilize alternative splicing to increase the coding capacity of overlapping ORFs in their compact genomes (Pintel *et al.*, 1995).

The deduced ORF2 protein of PfDNV genome has significant homology to the GKRN domain of NS-1 protein of parvoviruses, which is reported to have NTP-binding and helicase activities (Im and Muzyczka, 1990; Li and Rhode, 1990; Vanacker and Rommelaere, 1995). Thus, considering important functions of NS-1 polypeptide in the replication process of vertebrate parvovirus genome, ORF2 of PfDNV genome may encode an equivalent of NS-1 protein. In addition, a putative metal-binding domain of DNA-binding proteins was also found in the N-terminus of the deduced ORF2 protein of PfDNV (aa 126-140) using the consensus sequence H:X:H:$(X)_2$:H:X:C:$(X)_6$:C (Berg, 1986), which was found also in the deduced ORF2 protein of JcDNV. It is worth to mention that the deduced ORF2 protein of PfDNV genome has significant homology to the structural protein of BmDNV. Two potential metal-binding motifs "CXXC" (aa 145–148 and 197–200) were found in the deduced ORF1 protein of PfDNV genome; these motifs show significant homology to the NS-3 protein of DsDNV. It is assumed that the plus strand of PfDNV genome may encode NS proteins, while the functions of the deduced ORF3 and ORF4 proteins are still unknown.

The deduced proteins of ORF5 and ORF6 in the minus strand of PfDNV genome have significant homologies to VP1 proteins of other parvoviruses, which indicates that the minus DNA strand of PfDNV genome encodes structural proteins. Theoretically, the maximum size of the deduced protein from the minus strand of PfDNV genome is 85 K, which is much less than 105 K of VP5 estimated by SDS-PAGE. This contradiction could be explained by several assumptions: (1) the protein is N-glycosylated, it contains several N:X:S or N:X:T potential glycosylation sites, (2) VP5 is a stable dimeric protein of VP1 (52 K); (3) the virions probably package also host proteins. Further work on the mode of transcription of PfDNV genome is under way.

Acknowledgement. This work was supported by grant No. 39470032 from the National Natural Science Foundation of China.

References

Afanasiev BN, Galyov EE, Buchatsky LP, Kozlov YV (1991): Nucleotide sequence and genomic organization of Aedes densonucleosis virus. *Virology* **185**, 323–336.

Astell CR (1990): Terminal hairpins of parvovirus genome and their role in DNA replication. In Tijssen P (Eds): *Handbook of Parvoviruses*. Vol. 1, CRC Press, Boca Raton.

Astell CR, Chow MB, Ward DC (1985): Sequence analysis of termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling hairpin model for autonomous parvovirus DNA replication. *J. Virol.* **54**, 171–177.

Astell CR, Thomson M, Merchlinsky M, Ward DC (1983): The complete DNA sequence of minute virus of mice, an autonomous parvovirus. *Nucleic Acids Res.* **11**, 999–1018.

Bando H, Choi H, Ito Y, Kawase S (1990): Terminal structure of a densovirus implies a hairpin transfer replication which is similar to the model for AAV. *Virology* **179**, 57–63.

Bando H, Kusuda J, Gojobori T, Maruyama T, Kawase S (1987): Organization and nucleotide sequence of a densovirus genome imply a host-dependent evolution of the parvoviruses. *J. Virol.* **61**, 553–560.

Bensimhon MJ, Gabarro-Arpa J, Ehrlich R, Reiss C (1983): Physical characteristics in eucaryotic promoters. *Nucleic Acids Res.* **11**, 4521–4540.

Berg JM (1986): Potential metal-binding domains in nucleic acid binding proteins. *Science* **232**, 485–487.

Bergoin M, Tijssen P (1998): Biological and molecular properties of densoviruses and their use in protein expression and biological control. In Miller LK, Ball LA (Eds): *The Insect Viruses*. Plenum Press, New York.

Berns KI (1990): Parvoviridae and their replication. In Fields BN, Knipe DM (Eds): *Virology*. 2nd ed., Raven Press, New York.

Birnstiel ML, Busslinger M, Strub K (1985): Transcription termination and 3' processing: the end is in site. *Cell* **41**, 349–359.

Bloom ME, Alexandersen S, Garon CF, Mori S, Wei W, Perryman S, Wolfinbarger JB (1990): Nucleotide sequence of he 5'-terminal palindrome of aleutian mink disease parvovirus and construction of an infectious molecular clone. *J. Virol.* **64**, 3551–3556.

Bloom ME, Alexandersen S, Perryman S, Lechner D, Wolfinbarger JB (1988): Nucleotide sequence and genomic organization of aleutian mink disease parvovirus (ADV): Sequence comparisons between a nonpathogenic and a pathogenic strain of ADV. *J. Virol.* **62**, 2903–2915.

Boublik Y, Jousset F-X, Bergoin M (1994): Complete nucleotide sequence and genomic organization of the *Aedes albopictus* parvovirus (*Aa*PV) pathogenic for *Aedes aegypti* larvae. *Virology* **200**, 752–763.

Chen KC, Shull BC, Moses EA, Lederman M, Stout ER, Bates RC (1986): Complete nucleotide sequence and genome organization of bovine parvovirus. *J. Virol.* **60**, 1085–1097.

Deiss V, Tratschin J-D, Weitz M, Siegl G (1990): Cloning of the human parvovirus B19 genome and structural analysis of its palindromic termini. *Virology* **175**, 247–254.

Diffoot M, Shull BC, Chen KC, Stout ER, Lederman M, Bates RC (1989): Identical ends are not required for the equal encapsidation of plus- and minus- strand parvovirus LuIII DNA. *J. Virol.* **63**, 3180–3184.

Dumas B, Jourdan M, Pascaud A-M, Bergoin M (1992): Complete nucleotide sequence of the infectious genome of *Junonia coenia* densovirus reveals an organization unique among parvoviruses. *Virology* **191**, 202–222.

Gross O, Tal J (2000): Expression of the insect parvovirus *Gm*DNV *in vivo*: the structural and nonstructural proteins are encoded by opposite DNA strands. *J. Invertbr. Pathol.* **75**, 126–132.

Gross O, Weinberg D, Tijssen P, Tal J (1990): Expression of densonucleosis virus genome in *Galleria mellonella* larvae: characterization of viral RNA transcripts and of their *in vitro* translation products. *J. Invertbr. Pathol.* **56**, 175–180.

Guo H, Zhang J, Zhu L, Zhang X , Hu Y (1999): Construction of the DNA clones, subclones and deletion clones of the *Periplaneta fuliginosa* densovirus genome. *J. Wuhan Univ.* **6**, 883–886.

Hu Y, Liang D, Huang Y (1991): A small spherical virus of cockroach, *Periplaneta fuliginosa*. *J. Wuhan Univ.* **1**, 114.

Hu Y, Zheng J, Iizuka T, Bando H (1994): A densovirus newly isolated from the smoky-brown cockroach *Periplaneta fuliginosa*. *Arch. Virol.* **138**, 365–372.

Im DS, Muzyczka N (1990): The AAV originbinding protein Rep 68 is an ATP dependent site specific endonuclease with DNA helicase activity. *Cell* **61**, 447–457.

Jourdan M, Jousset FX, Gervais M, Skory S, Bergoin M, Dumas B (1990): Cloning of the genome of the genome of a densonucleosis virus and rescue of infectious virions from recombiant plasmid in the insect host *Spodoptera littoralis*. *Virology* **179**, 403–409.

Kariatsumari T, Horiuchi M, Hama E, Yaguchi K, Ishigurio N, Goto H, Shinagawa M (1991): Construction and nucleotide sequence analysis of an infectious DNA clone of the autonomous parvovirus, mink enteritis virus. *J. Gen. Virol.* **72**, 867–875.

Li X, Rhode SL III (1990): Mutation of lysine 405 to serine in the parvovirus H-1 NS1 abolishes its functions for viral DNA replication, late promoter trans-activation and cytotoxicity. *J. Virol.* **64**, 4654–4660.

Mount SM (1982): Acatalogue of splice junction sequences. *Nucleic Acids Res.* **10**, 459472.

Pintel DJ, Gersappe A, Haut D, Pearson J (1995): Determinants that govern alternative splcing of parvovirus pre-mRNAs. *Sem. Virol.* **6**, 283–290.

Ranz AI, Manclus JJ, Diaz-Aroca E, Casal JI (1989): Porcine parvovirus: DNA sequence and genome Organization. *J. Gen. Virol.* **70**, 2541–2553.

Reed AP, Jones EV, Miller TJ (1988): Nucleotide sequence and genome organization of canine parvovirus. *J. Virol.* **62**, 266–276.

Rhode Sl III, Iverson P (1990): Parvovirus genome: DNA sequences. In Tijssen P (Eds): *Handbook of Parvoviruses*. Vol. 1, CRC Press, Boca Raton, FL, pp. 31–57.

Rhode SL III, Paradiso PR (1983): Nucleotide sequence of H-1 and mapping of its genes by hybrid-arrested translation. *J. Virol.* **45**, 173–184.

Sambrook J, Fritsch EF, Maniatis T (1989): *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Samulski RJ, Berns KI, Tan M, Muzyczka N (1982): Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. USA* **79**, 2077–2081.

Samulski RJ, Chang LS, Shenk T (1987): A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. *J. Virol.* **61**, 3096–3101.

Samulski RJ, Srivastava A, Berns KI, Muzyczka N (1983): Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. *Cell* **33**, 135–143.

Sanger F, Nicklen S, Coulson AR (1977): DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

Shade RO, Blundell MC, Cotmore SF, Tattersall P, Astell CR (1986): Nucleotide sequence and genome organization of human parvovirus B19 isolated from the serum of a child during aplastic crisis. *J. Virol.* **58**, 921–936.

Srivastava A, Lusby EW, Berns KI (1983): Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol.* **45**, 555–564.

Tijssen P, Bergoin M (1995): Densonucleosis viruses constitute an increasingly diversified subfamily among the parvoviruses. *Sem. Virol.* **6**, 347–355.

Vanacker J-M, Rommelaere J (1995): Non-structural proteins of autonomous parvoviruses from cellular effects to molecular mechanism. *Sem. Virol.* **6**, 291–297.

Vasudevacharya J, Basak S, Srinivas RV, Compans RW (1990): The complete nucleotide sequence of an infectious clone of porcine parvovirus, strain NADL-2. *Virology* **178**, 611–616.

Van Regenmortel MHV, Fauquet CM, Bishop DHL (2000): *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego.