Genome Analysis of Infectious Pancreatic Necrosis Virus DRT Strain

Hyung-Hoan Lee, Hey-Kyung Chung, and Seong-Hun Lee

Department of Biology and Research Institute for Genetic Engineering, Konkuk University Seoul 143-701, Korea

Infectious pancreatic necrosis virus (IPNV) is a pathogen of economic interest to the world-wide fish-farming industry since it causes an acute, contagious disease in juvenile salmonids (Hill, 1982). IPNV is a member of the Birnaviridae, which is a newly established family (Brown, 1986), including infectious bursal disease virus (IBDV) of domestic fowl, Drosophila X virus (DXV) of *Drosophila melanogaster*, tellina virus (TV) and oyster virus (OV) of bivalve molluscus (Dobos *et al.*, 1979).

Several strains of IPNV have been isolated from salmonids and non-salmonids in North America, Europe and Asia (Wolf and Mann, 1980). Park *et al.* (1989) isolated a new serotype of IPNV from rainbow trout (*Salmo gairdneri*) in Daechung Dam, Korea. The new serotype, called the DRT strain, has been identified to be serologically distinguishable from three reference strains of IPNV; Ab, Sp and VR-299.

However the base sequence of the segment A of IPNV was known in the Jasper (Duncan and Dobos, 1986) and N 1 strains (Hävarstein *et al.*, 1990). In the case of segment B, the base sequences of the Jasper and Sp strains were known by Duncan *et al.*, 1991.

Here, we report the cDNA cloning and sequencing of the genome RNA segment A containing the VP2-NS-VP3 genes, and segment B containing the VP1 gene of the DRT strain of IPNV. The nucleotide and deduced amino acid sequences were compared to published sequences of the genes and the segment A and B of other IPNVs.

CHARACTERISTICS OF VIRION

IPNV consists of two double-stranded (ds) RNA segments (A and B) and of non-enveloped icosahedral virus particles of about 60nm in diameter (Dobos and Roberts, 1982). Genome segment A (approx. 3100 bp) of IPNV encodes three proteins as a polyprotein in a single large open reading frame, which is cleaved by a virus-coded protease (Duncan *et al.*, 1987).

The order of these polypeptides in the polyprotein is NpVP2-NS-VP3-C (Duncan and Dobos, 1986). VP2 is a major structural protein of 52~54 Kd which contains the antigenic regions responsible for serotype specificity and for neutralization of the virus particles; NS is a 29 kD, non-structural protein with which a protease activity has been associated; and VP3 is a 31 kD, minor structural protein which is an internal component of the capsid (Duncan *et al.*, 1987). Genome segment B (approx. 2800 bp) of IPNV encodes a single gene product (VP1), the putative RNA dependent RNA polymerase. This protein is present in virions as a free VP1 or as a genome-linked form or VPg (Calvert *et al.*, 1991; Persson and MacDonald, 1982).

MATERIALS AND METHODS

Virus and cells used

DRT strain (Park *et al.*, 1989) of IPNV and Chinook salmon embryo cell line (CHSE-214) (Lannan *et al.*, 1984) were used. The virus was multiplicated on CHSE-214 cells at 18°C.

Strategy of cDNA synthesis by RT-PCR

The strategy of cDNA synthesis of the IPNV-DRT is illustrated in Fig. 1 and 2. Oligonucleotide sequences for primers were deduced according to the published sequences of the Jasper (Duncan *et al.*, 1991; Duncan and Dobos, 1986) and N1 (Hävarstein *et al.*, 1990) strains. Purified genomic dsRNA was denatured by treatment with dimethyl sulfoxide (DMSO) was used as a template for the first strand cDNA. First strand cDNA was synthesized using the RNA segments A and B as templates, synthetic primers, and a cDNA

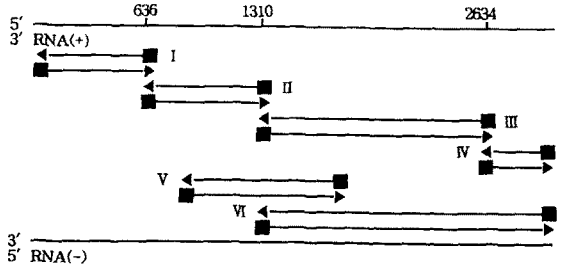

Fig. 1. Strategy of cDNA synthesis of the IPNV-DRT segment A. Arrows indicate the direction of synthesis of cDNA. Primers are indicated by filled boxes.

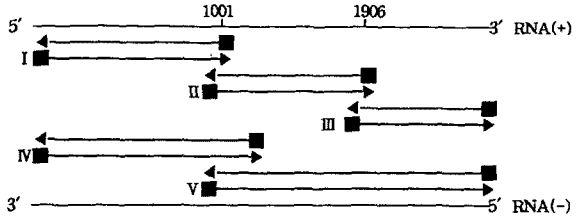

Fig. 2. Strategy of cDNA synthesis of genome segment B. ■: primer for synthesis cDNA. Arrows indicate the direction of synthesis of cDNA.

synthesis Kit (Amersham) with conditions as recommended by the manufacturer. First strand cDNA was used as a template for the second strand cDNA synthesis and amplification of ds-cDNA by PCR (Saiki, *et al.*, 1988).

Strategy of cDNA sequencing

Purified double-stranded cDNA was phosphorylated with T4 polynucleotide kinase and cloned into pBluescript SK(+) vector (Short *et al.*, 1988), and then the recombinants were transformed into competent *Escherichia coli* XL1-Blue cells (Maniatis *et al.*, 1989). The nucleotide sequences of the segments A and B of the IPNV-DRT were determined by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using the Sequenase Version 2.0 system (United States Biochemical). Both strands of the cDNA were sequenced completely except for 5' and 3' terminal primers. Each base has been confirmed from at least two independent clones. The nucleotide and amino acid sequence were analyzed using the DNASIS and PROSIS programs (Hitachi software, Pharmacia).

RESULTS

cDNA synthesis of the segment A

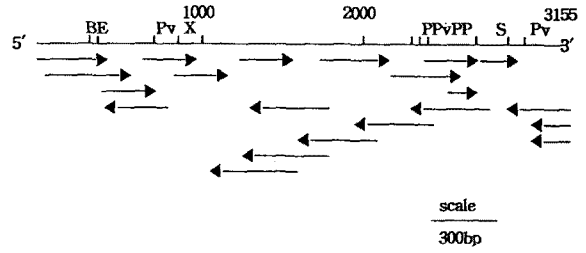

Fig. 3. Physical map and nucleotide sequencing strategy of the IPNV-DRT segment A. The partial physical map of segment the A is diagrammed. Arrows indicate the extent and direction of nucleotide sequence. Restriction enzyme sites: B, BamHI; E, EcoRI; P, PstI; Pv, PvuII; S, ScaI; X, XhoI.

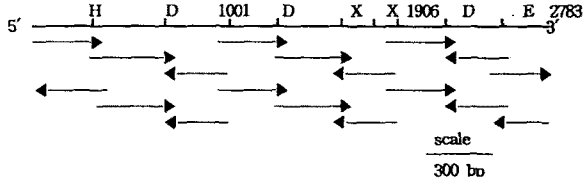

Fig. 4. Physical map and nucleotide sequencing strategy of the IPNV-DRT segment B. The partial physical map of segment B is diagrammed. Arrows indicate the extent and direction of nucleotide sequence. Restriction enzyme sites: D, DraII; E, EcoRI; H, HindIII; X, XhoI.

In this work, six fragments of cDNA covering the RNA segment A of IPNV-DRT were synthesized using a reverse transcriptase (RT)-polymerase chain reaction (PCR). The synthesized cDNA fragments,I (nt 1 to 721), II (nt 636 to 1332), III (nt 1310 to 2668), and IV (nt 2634 to 3097) were shown in Fig. 1. The cDNAs were cloned into pBluescript (SK+) and then named pDRTA1, pDRTA2, pDRTA3 and pDRTA4, respectively. To determine the nucleotide sequences of the primers (nt 1305 to 1336, nt 2634 to 2668), cDNA fragments V (nt 851 to 1787) and VI (nt 1310 to 3097) were also synthesized (Fig. 1) and then cloned into pBluescript (SK+).

Nucleotide sequence analysis of the segment A

To determine the nucleotide sequence of the segment A of IPNV-DRT, 19 cDNA clones were selected and sequenced. The nucleotide sequence of the segment A is 3, 155 bp long and contains a large ORF of 2,916 bp and a small ORF of 444 bp (Table 1). In the case of the segment A of IPNV, only the nucleotide sequences of the Jasper (Duncan and Dobos, 1986) and N1 strains (H varstein *et al.*, 1990) have previously been sequenced. The large ORF

of the DRT strain begins at nucleotide 120 and ends with a single TAA termination codon at nucleotide 3,035 and contains VP2-NS-VP3 genes (Table 1).

The large ORF of the IPNV-DRT is identical in length to the one already described for the Jasper and the N1 strains. We found 98.7% homology of the large ORF nucleotide sequence of the segment A between the IPNV-DRT and the IPNV-Jasper. A comparison with the segment A of the IPNV-N1 revealed a 79.1% homology(Table 1).

The nucleotide sequence of the segment A of the DRT strain contained a small ORF of 444 bp encoding 148 amino acids. The small ORF was perfectly conserved on the segment A of IPNVs. The 444-nucleotide DRT strain began at nucleotide 68 and terminates at a single TGA termination codon at nucleotide 511(Table 1).

In the case of DRT strain, the additional sequences at the 3' end of the plus strand of the segment A are longer than those of the Jasper and N1 strains by 58 bp and 51 nbp, respectively. The additional sequence at the 3'-termini of the STC-IBDV genome segment A has previously been reported (Kibenge *et al.*, 1990). The cDNA sequences of the segment A and B of IPNV-Jasper both start with a conserved pentanucleotide sequence (GGAAA) and terminate with a conserved tetranucleotide sequence.

cDNA synthesis of the segment B

To determine the nucleotide sequence of the genome segment B of IPNV-DRT strain, five cDNA fragments were synthesized using a reverse transcriptase (RT)-polymerase chain reaction (PCR). First of all, I (NT 1-1001), II (NT 982-1906) and III (NT 1880-2784) cDNA fragments were synthesized as illustrated in Fig. 2. And to determine the nucleotide sequences of the primers (NT 977-1011, NT 1880-1908), IV (NT 1-1252) and V (NT 982-2784) cDNA fragments were synthesized.

Nucleotide Sequence of the Segment B

We have determined the nucleotide sequence of IPNV-DRT genome segment B and deduced amino acid sequence of an encoded VP1 protein from nucleotide sequence. Above all things, we completely determined the nucleotide sequence of the RNA dependent RNA polymerase of the DRT strain of IPNV.

The cDNA sequence of the genome segment B of IPNV-DRT was 2,783 bp long and contained only a single ORF of 2,535 bp (Table 2). The DRT segment B sequence contained a single large ORF encoding the VP1 protein that was composed of 845 amino acids. The 2,535-nucleotide DRT ORF began at nucleotide 101 and terminated at a single TGA termination codon at nucleotide 2,636.

Duncan *et al.* (1986) have reported the nucleotide sequence of the genome segment B of the IPNV-Jasper and - Sp strains. The Jasper sequence contained 2,784 bp and the Sp sequence consisted of 2,630 bp. Each sequence contained a single large ORF encoding the 845-amino-acid Jasper or 844-amino-acid Sp VP1 proteins. The

Table 1. Composition of RNA genome A segment sequence among three IPNV Viruses

Items for comparison	Homology among IPNV (%)*		
	DRT and Jasper	DRT and N1	Jasper and N1
Nucleotide sequence homology			
Whole genome segment	97.7	79.5	79.5
Large ORF (2916)	98.7	79.1	78.9
5' Non-coding sequences	97.8	85.2	86.7
3' Non-coding sequences	90.3	82.8	93.1
Small ORF (444)	98.4	84.6	84.5
Amino acid sequence homology			
Large ORF (972)	99.0	84.0	84.0
VP2[a]	99.2	88.7	88.5
NS	97.9	78.5	79.4
VP3	99.6	79.2	78.8
Small ORF (148)	95.3	64.9	64.2

*Values indicate percentage of homology; figures in parentheses indicate length of nucleotide(bp) or amino acid. [a]The large ORFs were divided into three segments corresponding to the gene products VP2, NS and VP3. The regions of VP2, NS and VP3 are from Duncan *et al.*(1987).

Table 2. Comparison of genome segment B of the IPNV strains

	DRT	Jasper*	Sp*
Segment B (nucleotides)	2,783	2,784	2,630
VP1 ORF (nucleotides)	101~2,635	101~2,635	94~2,625
VP1 ORF (codons)	845	845	844
Predicted molecular weight	94,426	94,441	94,064
Termination codon	TGA	TAA	TAA

*Sequence data from Duncan *et al.*

comparison result of genome segment B of three IPNV strains was summarized in Table 3. A unique feature of the genome segment B sequence of the DRT strain was the deletion of the 2,645th nucleotide in the Jasper strain.

Amino Acid Sequence Analysis of the Segment A Encoding VP3-NS-VP2 Proteins

The predicted amino acid sequence of the segment A of the IPNV-DRT was made. The segment A of the IPNV-DRT contains two overlapping ORFs, positioned in different reading frames. The large ORF of the IPNV-DRT encodes 972 amino acid precursor polyprotein with a predicted molecular weight of 106,641 daltons (Table 2). Kozak reported that the optimal sequence for translation initiation by eukaryotic ribosomes has been determined to be CCACCATGG. The start codon of large ORF of the IPNV-DRT at position 120 contains the important purine in the −3 position (A), but lacks a G in position +4. The polyproteins of three IPNVs are more closely related than the nucleotide sequence homology would indicate. Most of the nucleotide mismatches are in third base codon positions, the majority of which do not cause amino acid changes.

The exact cleavage sites of the polyprotein are unknown, but cleavage sites are predicted by partial proteolysis mapping with N-chlorosuccinimide. Dibasic residues (Tyr-Leu) at positions 491 and 725 have been postulated as possible sites (Duncan *et al.*, 1987). Azad *et al.* (1987) have shown that expression in *Escherichia coli* of the large ORF of the IBDV segment A results in the processing of the polyprotein. They also suggested that NS (VP4), which is thought to be a protease, is involved in the processing of the precursor polyprotein. In comparison with amino acid sequence of NS (VP4) of the IPNV-Jasper and the IPNV-DRT was different only in five residues. In the case of the VP3, a single amino acid was different between the DRT and the Jasper strains.

The segment A of the IPNV-DRT contains a small ORF coding capacity of 148 amino acid polypeptide with a predicted protein of molecular weight of 17,369 daltons, and overlapped with N-terminal part of polyprotein (nt 68 to 511). The initiation sequence for small ORF at position 68 does not match with the Kozak rules (1987).

The small ORF in segment A of birnavirus has been mentioned by Azad *et al.* (1987) and Duncan *et al.* (1987). The existence of a 17 kD polypeptide has never been corroborated. However Hävarstein *et al.* (1990) implicated that the 17 kD band is probably of viral origin when [^{35}S] methionine labelled purified virus was analyzed by SDS-PAGE gel electrophoresis. We found that the size of small ORF was perfectly same between strains.

Amino Acid Sequence Analysis of the Segment B Encoding VP1 Proteins

The VP1 protein is the presumptive birnavirus RdRp. The predicted amino acid sequences of the IPNV-DRT, -Jasper and -Sp strains VP1 proteins were aligned with that of the IBDV VP1. The comparison of the IPNV VP1 sequences with the 878 amino acid IBDV sequence revealed the conserved 372 amino acids located all four proteins. We found the GTP-binding sequence in VP1 protein of the DRT strain. Comparative analysis of the amino acid sequences of the nucleotide binding protein and putative RdRps has revealed the presence of the conserved GTP-binding motif in four VP1 proteins.

Dever *et al.* (1987) reported the GTP-binding consensus sequence present in nucleotide binding proteins (e.g. elongation factors, ras p21 protein, phosphoenolpyruvate carboxykinase). The consensus sequence is composed of three consensus elements GXXXXGK, DXXG and NKXD with consensus spacings of either 40~80 or 130~170 amino acid residues between the first and second elements and 40~80 amino acid residues between the second and third sequence elements. The same sequence is present in IPNV VP1 between residues 248 and 255 (GLPYIGKT). The corresponding region of IBDV VP1 is GLPYVGRT. As such, this region represents a potential GTP binding site in the VP1 of birnaviruses.

ACKNOWLEDGEMENT

This work was supported by a grant for genetic engineering research from the Ministry of Education.

Korea, and also by the Research Center for Molecular Microbiology, Seoul National University (SRC).

CONCLUSION

The cDNA of the infectious pancreatic necrosis virus (IPNV) DRT strain was synthesized using the reverse transcriptase (RT)-polymerase chain reaction and its cDNA nucleotide sequence was determined. The cDNA nucleotide sequence of the segment A was 3,155 base pairs (bp) in length and contained two overlapping open reading frames (ORFs). The large ORF was 2,916 bp and predicted a polyprotein of a molecular weight of 106,641 daltons, whereas the small ORF was a 444 bp and predicted a protein of a molecularweight of 17,369 in IPNV-DRT. The large ORF contains three genes, VP2-NS-VP3, which make a polyprotein. The DRT segment B was 2,783 bp long and contained only a single long of 2,535 bp in length. This ORF nucleotides encoded the VP1 protein, the putative RdRp of IPNV. The VP1 protein was consisted of 845 amino acids. The molecular weight of the RdRp, as deduced from the nucleotide sequence, is 94,426. The conserved GTP-binding motif was detected in VP1 protein. The nucleotide sequence data of the segment A and B in this paper has been deposited with the GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases and has been assigned accession number D26526.

REFERENCES

1. Azad, A. A., M. N. Jagadish, M. A. Brown, and P. J. Hudson. 1987. *Virol.* **161**, 145-152.
2. Bernard, J. 1980. Drosophila X virus RNA polymerase: tentative model for *in vitro* replication of the double stranded virion RNA. *J. Virol.* **33**, 717-723.
3. Brown, F. 1986. The classification and nomenclature of viruses: Summary of results of meetings of the International Committee on Taxonomy of Viruses in Sendai, September 1984. *Intervirology* **25**, 141-143.
4. Calvert, J. G., E. Nagy, M. Soler, and P. Dobos. 1991. Characterization of the VPg-dsRNA linkage of infectious pancreatic necrosis virus. *J. Gen. Virol.* **72**, 2563-2567.
5. Chung, H. K., S. H. Lee, H. H. Lee, D. S. Lee and Y. S. Kim. 1994. Nucleotide sequence analysis of the VP2-NS-VP3 genes of infectious pancreatic necrosis virus DRT strain. *Mol. and Cells* **4**, 349-354.
6. Chung, H. K., S. H. Lee, S. Y. Kim and H. H. Lee. 1994. Nucleotide sequence analysis of the RNA-dependent RNA polymerase gene of infectious pancreactic necrosis virus DRT strain. *J. Microbiol. Biotech.* **4**(4), 264-269.
7. Dever, T. E., M. J. Glynias, and W. C. Merrick. 1987. GTP-binding domain: three consensus sequence elements with distinct spacing. *Pro. Natl. Acad. Sci. USA.* **84**, 1814-1818.
8. Dobos, P., B. J. Hill, R. Hallett, D. T. C. Kells, H. Becht, and D. Teninges. 1979. Biophysical and biochemical characterization of five animal viruses with bisegmented double-stranded RNA genomes. *J. Virol.* **32**, 593-605.
9. Dobos, P. and T. E. Roberts. 1983. The molecular biology of infectious pancreatic necrosis virus: a review. *Can. J. Microbiol.* **29**, 377-384.
10. Duncan, R. and P. Dobos. 1986. The nucleotide sequence of infectious pancreatic necrosis virus (IPNV) dsRNA segment A reveals one large ORF encoding a precursor polyprotein. *Nucleic Acids Res.* **14**, 5934.
11. Duncan, R., E. Nagy, P. J. Krell, and P. Dobos. 1987. Synthesis of the infectious pancreatic necrosis virus polyprotein, detection of a virus-encoded protease and fine structure mapping of genome segment A coding regions. *J. Virol.* **61**, 3655-3664.
12. Duncan, R., C. L. Mason, E. Nagy, J. A. Leong, and P. Dobos. 1991. Sequence analysis of infectious pancreatic necrosis virus genome segment B and its encoded VP1 protein: a putative RNA-dependent RNA polymerase lacking the Gly-Asp-Asp motif. *Virol.* **181**, 541-552.
13. Hävarstein, L. S., K. H. Kalland, K. E. Christie, and C. Endresen. 1990. Sequence of the large double-stranded RNA segment of the N1 strain of infectious pancreatic necrosis virus: a comparison with other *Birnaviridae. J. Gen. Virol.* **71**, 299-308.
14. Hill, B. J. 1982. Infectious pancreatic necrosis virus and its virulence. In *Microbial Diseases of Fish*, pp.91-114. Edited by R. J. Roberts. New York and London: Academic Press.
15. Kozak, M. 1987. *Nucleic acids Research* **15**, 8125-8148.
16. Lannan, C. N., J. R. Winton, and J. L. Fryer. 1984. Fish cell lines: establishment and characterization of nine cell lines from salmonids. *In Vitro* **20**, 671-676.
17. Mertens, P.P.C. and P. Dobos. 1982. Messenger RNA of infectious pancreatic necrosis virus is polycistronic. *Nature* **297**, 243-246.
18. Morgan, M. M., I. G. Macreadie, V. R. Harley, P. J. Hudson, and A. A. Azad. 1988. Sequence of the small double-stranded RNA genomic segment of infectious bursal disease virus and deduced 90-kDa product. *Virol.* **163**, 240-242.
19. Park, J. W., J. J. Lee, G. J. Jeong, and Y. C. Hah. 1989. Characterization of the infectous pancreatic necrosis virus (IPNV) isolated from pan-cultured rainbow trout in Korea. *Kor. J. Microbiol.* **27**, 225-230.
20. Persson, R. H. and R. D. MacDonald. 1982. Evidence that infectious pancretic necrosis virus has a genome-linked pro-

tein. *J. Virol.* **44**, 437-443.

21. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.
22. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Pro. Natl. Acad. Sci. USA.* **74**, 5463-5467.
23. Short, Jm M., J. M. Fernandez, J. A. Sorge, and W. D. Huse. 1988. λ ZAP: A bacteriophage λ expression vector with *in vivo* excision properties. *Nucleic Acids Res.* **16**, 7583-7600.
24. Somogyi, P. and P. Dobos. 1980. Virus-specific RNA synthesis in cells infected by infectious pancreatic necrosis virus. *J. Virol.* **33**, 129-139.
25. Spies, U. and H. Müller. 1990. Demonstration of enzyme activities required for cap structure formation in infectious bursal disease virus, a member of the birnavirus group. *J. Gen. Virol.* **71**, 977-981.
26. Wolf, K. and J. A. Mann. 1980. Poikilothermic vertebrate cell lines and viruses: A current listion for fishes. *In Vitro* **16**, 168-179.