The Complete Nucleotide Sequences of L3 and S7 Segments of Ibaraki Virus Encoding for the Major Inner Capsid Proteins, VP3 and VP7

Hiroyuki IWATA1), Sachio MANABE1), Ayumi YOSHIDA1), Eva Megumi Nara PEREIRA1) and Takeshi INOUE1)

1)Department of Veterinary Hygiene, Faculty of Agriculture, Yamaguchi University, 1677–1 Yoshida, Yamaguchi, Yamaguchi 753–8515, Japan

(Received 19 July 2000/Accepted 26 September 2000)

ABSTRACT. The complete nucleotide sequences of the genes encoding two of the major inner capsid proteins of Ibaraki virus (IBAV), belonging to epizootic hemorrhagic disease virus serotype 2 (EHDV-2) were determined. The L3 RNA segment is 2768 nucleotides in length which encodes VP3 polypeptides of 899 amino acid residues (M.W. 103 kDa). The S7 RNA segment, which encodes the VP7 core protein, is 1162 nucleotides in length and encodes 349 amino acids (M.W. 38 kDa). These RNA segments had the characteristic consensus motifs of Orbivirus RNA segments in termini, namely 5'-GUUAAA... and ...ACUUAC-3'. The comparison of the IBAV L3 and S7 sequences with those of other two EHDV-2 isolates revealed the higher homologies of 93% and 92% against EHDV-2 Australia isolate (EHDV-2AUS) and lower homologies of 80% and 81% against EHDV-2 North America isolate, respectively. The phylogenetic analysis based on L3 and S7 genes also indicated close relationships between IBAV and EHDV-2AUS.

KEY WORDS: dsRNA gene, Ibaraki virus, inner capsid, VP3, VP7.

Ibaraki virus (IBAV) is a strain of epizootic hemorrhagic disease virus serotype 2 (EHDV-2) belonging to the genus Orbivirus in the family Reoviridae [18, 19]. The virus is composed of seven structural proteins organized in a double-layered capsid containing the 10 dsRNA segments (L1-3, M4-6, S7-10) of the viral genome [22]. The inner capsid is composed of two major proteins, VP3 and VP7, and three minor structural proteins (VP1, 4 and 6). The segment L3 of viral genes encodes the inner capsid protein of VP3 which constitute the viral subcore [5]. This subcore is surrounded by 260 trimers of VP7 encoded by segment S7 to form the core particles, which has a knobby surface appearance in an icosahedral symmetry [13]. The inner capsid proteins were considered to be serogroup-specific and associated with infectivity to insect [9]. Previous phylogenetic analyses based on bluetongue virus (BTV) gene segment L3 indicated a geographical distribution of different genotypes, while the BTV S7 nucleotide sequences did not unequivocally display geographic distribution as seen in L3 [24]. The Nucleotide sequences of EHDV-2 L3 and S7 segments have been reported on the Australia (EHDV-2AUS) and North America isolates (EHDV-2NA) [2, 4, 10, 24]. The identity of L3 sequence between these two EHDV-2 subtypes was about 80%, but EHDV-2NA had a higher homology to EHDV-1 than that to EHDV-2AUS. Although partial nucleotide sequence of IBAV L3 segment has also been reported [4], the sequences of S7 and full-length L3 genes are still unknown. We previously described the nucleotide sequences of L2 and M5 genes encoding the major outer capsid proteins of VP2 and VP5 on IBAV [11, 12]. There are some genetic differences of L2 genes between Japan and North America isolates of EHDV-2 (70% nucleotide identity) in spite of the same serotype. Since L2 and/or M5 sequence for North America and Australia isolates has not been analysed, the genetic relationships among these EHDV-2 isolates have not been established. In this study, the complete nucleotide sequences of viral gene segment L3 and S7 of IBAV were analysed in order to investigate genetic relationships among EHDV-2 isolates with regard to the genes encoding inner capsid proteins.

MATERIALS AND METHODS

Viruses, cells and isolation of dsRNA: Ibaraki virus No.2 strain, which was kindly given from National Institute of Animal Health, Japan, was plaque-purified and propagated using monolayers of HmLu-1 cells in minimal essential medium (MEM; Nissui Pharmaceutical Corp., Ltd, Japan) supplemented with 5% fetal calf serum (FCS; M.A. Bioproducts, Inc., U.S.A.). The viral dsRNAs were purified from HmLu-1 cells infected with IBAV and the segment L3 and S7 RNAs of IBAV were separated on agarose gel electrophoresis and recovered as described elsewhere [14].

Synthesis of cDNA: The L3 and S7 dsRNA was transcribed into cDNA by the Marathon™ cDNA amplification kit (Clontech Laboratories, Inc., U.S.A.) using a random primer according to the manufacturer's instruction. The cDNA blunt-ended with T4 DNA polymerase was ligated to the Marathon cDNA adaptors and was amplified by the polymerase chain reaction (PCR) method using adaptor primer or L3 specific primers designed from the sequence data of EHDV-2AUS (CSIRO439) and EHDV-2NA (sv 124) to obtain cDNA containing the whole ORF as follows; EHD2L3F: 5'-GTT AAA TTT CCA GAG CGA TG-3' and EHD2L3R: 5'-GCA AAT ATT AGT -3' (as a sense primer) and EHD2L3R: 5'-GCA AAT ATT AGT -3' (as an antisense primer). PCR conditions were 30 cycles of amplification with a denaturation temperature of 94°C for 1 min, an annealing temperature of 50°C for 1 min, and an extension temperature of 72°C for 3 min. The cDNA fragments were cloned into pMosBlueT
vector (Amersham Life Science, Inc., U.K.) or pGEM®-T Easy vector (Promega Corp., U.S.A.) by TA cloning procedure and introduced into *Escherichia coli* (*E. coli*) XL-1 Blue. The clones were randomly selected and screened by *Hind* I restriction pattern and sequence analyses of both termini of the insert.

5′- and 3′-terminal cDNA clones: To determine both terminal sequences, the terminal cDNAs were synthesized from purified L3 and S7 dsRNAs by 5′-rapid amplification of cDNA (RACE) method using 5′-Full RACE core set (TaKaRa, Japan) according to the manufacturer’s instruction. In brief, 5′-terminal phosphorylated primers were designed from the sequence data of cDNA obtained by the above method, which produced a single stranded (ss) cDNA of about 500 base from 5′ or 3′ terminus. After hybridization of the 5′-terminal phosphorylated primer to single stranded (ss) RNA, cDNA was extended to the 5′-terminus by reverse transcriptase and was digested with RNase H. Resulted ss cDNA was circularized or concatemerized using T4 RNA ligase and amplified by PCR method using TaKaRa Ex Taq™ (TaKaRa). The PCR product was cloned into pGEM®-T Easy vector by TA cloning procedure and transformed into *E. coli*.

Sequencing of cDNA clones: Appropriate cDNA fragments were subcloned into pUC118 and deletion derivatives were generated following the ExoII/Mung Bean Nuclease Deletion Kit protocol (Stratagene, U.S.A.). The sequences of both strands were determined by the cycle sequencing protocols using ABI Prism BigDye™ terminator cycle sequencing ready reaction kit (The Perkin-Elmer Corp., U.S.A.). The sequence analysis was carried out using ABI Prism 377 autosequencer. To confirm sequence data, at least two overlapped or PCR-derived clones were analysed. Sequence alignment and phylogenetic analysis were performed by a CLUSTALW program (DNA Data bank of Japan, National Institute of Genetics, Japan) [21].

**RESULTS**

**Complete nucleotide sequence of L3 cDNA:** The full-length cDNA of L3 segment was 2768 nucleotides in length (DDBJ/EMBL/GenBank Accession Number: AB041933) and the coding strand had a calculated base composition of 30.7% A, 17.8% C, 24.4% G, and 27.1% T. The characteristic conserved terminal hexanucleotides commonly found in Orbivirus genes, namely 5′-GTTAAA... and ...ACTTAC-3′, were present (Fig. 1). An open reading frame (ORF) of 2697 nucleotides, beginning with an ATG at position 18 to 20 and terminating at position 2698 to 2700 with a TAA codon, was flanked by 5′ non-coding region of 17 base pairs (bp) and 3′ non-coding region of 54 bp. The translated ORF generated a deduced amino acid sequence of 899 residues with a calculated molecular weight (M.W.) of 102,917 Da which contained 465 residues of hydrophobic amino acids (51.7%).

**Comparison of L3 gene and VP3 protein among Orbiviruses:** Comparison of L3 gene and VP3 protein sequences among Orbiviruses were performed by a CLUSTALW program [21]. Sequence analyses of IBAV L3 gene showed high identities of 93% with EHDV-2AUS, 80% with EHDV-2NA and EHDV-1 [8], while percent identities against BTV-10 and African horse sickness virus serotype 4 (AHSV-4) [6] were 63% and 53%, respectively. The alignment of VP3

---

Fig. 1. Comparison of the 5′ and 3′ non-coding region of S7 and L3 segments of Ibaraki virus (IBAV) with those of other EHDVs. The characteristic consensus motifs of Orbivirus RNA segments in termini are boxed. Start and stop codons are underlined.
amino acid sequence of IBAV with those of EHDV-2AUS, EHDV-2NA and EHDV-1 are shown in Fig. 2. The homologies of the deduced amino acids were 98% to EHDV-2AUS and 95% to EHDV-2NA and EHDV-1, whereas EHDV-2NA L3 gene had extremely high genotypic and phenotypic homologies of 98% and 99% to EHDV-1, respectively. The phylogenetic tree of the L3 sequence revealed that the AHSV, BTV and EHDV groups were very distinct (Fig. 4A).
In spite of the very low level of variability between the L3 sequences, it was evident from this analysis that EHDV-2NA and EHDV-1 sequences were unequivocally placed as a sister group, while IBAV showed closer relationship with EHDV-2AUS than EHDV-2NA. In addition, genetic relationships between IBAV and EHDV-2NA were more distant than those among BTV serotypes investigated to date.

**Complete nucleotide sequence of S7 cDNA:** The full-length cDNA of S7 segment was 1162 nucleotides in length and the coding strand had a base composition of 30.6% A, 19.7% C, 24.3% G, and 25.4% T (DDBJ/EMBL/GenBank Accession Number: AB041934), which was similar to other Orbivirus data and also to IBAV L3 gene. The conserved terminal hexanucleotides were also present as seen in the L3 sequence (Fig. 1). An ORF of 1,047 nucleotides, beginning with an ATG at position 18 to 20 and terminating at position 1,065 to 1,067 with a TGA codon, was flanked by 5' and 3' non-coding region of 17 and 95 bp, respectively. The ORF was 3 nucleotides longer than that of EHDV-2AUS and the same length as those EHDV-2NA and EHDV-1 [8]. The translated ORF generated a deduced amino acid sequence of 349 residues with a calculated M.W. of 37,999 Da containing 194 residues of hydrophobic amino acids (55.4%).

**Comparison of S7 gene and VP7 protein among Orbiviruses:** The genotypic and phenotypic identities of S7 gene were also analysed as described above. The percent identities of S7 nucleotide sequence of IBAV against those of EHDV-2AUS, EHDV-2NA and EHDV-1 were 92%, 81% and 79%, respectively, while the percent identities against BTV-10 and AHSV-4 [16] were 63% and 57%, respectively. The alignment of VP7 amino acid sequence of IBAV with those of EHDV-2AUS, EHDV-2NA and EHDV-1 are shown in Fig. 3. The identities of the deduced amino acids were

![Fig. 3. Alignment of the deduced amino acid sequence of the Ibaraki virus (IBAV) VP7 protein with those of EHDV-2AUS (CSIRO439), EHDV-2NA (sv124) and EHDV-1. Alanine residue essential for the insoluble character at position 167 is boxed. Dots indicate identical amino acids to IBAV sequence and a hyphen indicates a gap.](image)

Fig. 3. Alignment of the deduced amino acid sequence of the Ibaraki virus (IBAV) VP7 protein with those of EHDV-2AUS (CSIRO439), EHDV-2NA (sv124) and EHDV-1. Alanine residue essential for the insoluble character at position 167 is boxed. Dots indicate identical amino acids to IBAV sequence and a hyphen indicates a gap.

In spite of the very low level of variability between the L3 sequences, it was evident from this analysis that EHDV-2NA and EHDV-1 sequences were unequivocally placed as a sister group, while IBAV showed closer relationship with EHDV-2AUS than EHDV-2NA. In addition, genetic relationships between IBAV and EHDV-2NA were more distant than those among BTV serotypes investigated to date.

**Complete nucleotide sequence of S7 cDNA:** The full-length cDNA of S7 segment was 1162 nucleotides in length and the coding strand had a base composition of 30.6% A, 19.7% C, 24.3% G, and 25.4% T (DDBJ/EMBL/GenBank Accession Number: AB041934), which was similar to other Orbivirus data and also to IBAV L3 gene. The conserved terminal hexanucleotides were also present as seen in the L3 sequence (Fig. 1). An ORF of 1,047 nucleotides, beginning with an ATG at position 18 to 20 and terminating at position 1,065 to 1,067 with a TGA codon, was flanked by 5' and 3' non-coding region of 17 and 95 bp, respectively. The ORF was 3 nucleotides longer than that of EHDV-2AUS and the same length as those EHDV-2NA and EHDV-1 [8]. The translated ORF generated a deduced amino acid sequence of 349 residues with a calculated M.W. of 37,999 Da containing 194 residues of hydrophobic amino acids (55.4%).

**Comparison of S7 gene and VP7 protein among Orbiviruses:** The genotypic and phenotypic identities of S7 gene were also analysed as described above. The percent identities of S7 nucleotide sequence of IBAV against those of EHDV-2AUS, EHDV-2NA and EHDV-1 were 92%, 81% and 79%, respectively, while the percent identities against BTV-10 and AHSV-4 [16] were 63% and 57%, respectively. The alignment of VP7 amino acid sequence of IBAV with those of EHDV-2AUS, EHDV-2NA and EHDV-1 are shown in Fig. 3. The identities of the deduced amino acids were

![Fig. 4. Phylogenetic relationships of Ibaraki virus (IBAV) based on L3 (A) and S7 (B) sequences from different EHDV, BTV, and AHSV serotypes. Branch lengths are proportional to genetic distance calculated by the CLUSTAL W program. This figure is of a bootstrap analysis where this species order was predicted with 100% confidence at each node.](image)
DISCUSSION

Epizootic hemorrhagic disease virus serotype 2 were isolated in Australia, North America and Japan. Although these isolates are expected to be closely related each other based on at least viral neutralization and hemagglutination test, genetic relationships are still unknown. We have previously described the viral genome sequences of outer capsid proteins, VP2 and VP5, of IBAV, which are the most variable proteins among Orbiviruses and have suggested that L2 has some extent of genetic distance between IBAV and North America isolates of EHDV-2 [11, 12]. However, L2 sequence of Australia isolate and M5 of Australia and North America isolates are not available for investigation of genetic relationship among these EHDV-2 isolates. In this study, we determined L3 and S7 sequence encoding for inner capsid proteins of VP3 and VP7, which were reported on Australia and North America isolates of EHDV-2 and genetic relationships among these three subtypes.

The 5’ and 3’ terminal sequences of L3 and S7 segments were found to contain the characteristic consensus sequences of Orbivirus RNA segments as seen in L2 and M5 genes [11, 12], which were considered to be important in transcription initiation, RNA replication, ribosome binding and translation of the mRNAs [15]. The deduced amino acid sequence of VP3 was 899 residues and had a low content of charged amino acids and a high content of hydrophobic amino acids similar to other Orbiviruses. In BTV, it was demonstrated that a methionine at residue 500 and an arginine at residue S02 were essential for CLP formation [20]. As expected from the low variability of VP3 amino acid sequences, the same structure was observed in IBAV. The VP7 amino acid sequence was 349 residues, which was one residue longer than EHDV-2AUS and the same as EHDV-2NA and EHDV-1, and also had a high hydrobobicity resulting in its highly insoluble nature. In VP7 sequence, the alanine residue at position 167 was reported to contribute to the insoluble character, which was also conserved in IBAV [1].

The comparison of four major capsid proteins among Orbiviruses indicated VP3 and VP7 were highly conserved and serogroup-specific, while the homologies of VP2 and VP5 between Orbiviruses were much lower, suggesting that VP2 and VP5 was serotype-specific [7]. These findings were also observed in the relationship between EHDV-1 and IBAV as well as other EHDV-2 isolates (data not shown). The comparison of L3 and S7 genes among three EHDV-2 isolates showed IBAV had a higher identity with EHDV-2AUS than EHDV-2NA. Although the homology of L3 among BTV serotypes was more than 89%, the homology between IBAV and EHDV-2NA was apparently lower (80%) in spite of the same serotype. Phylogenetic analysis based on L3 and S7 sequences also showed that IBAV had closer relationship against EHDV-2AUS than EHDV-2NA and the genetic distance between IBAV and EHDV-2NA was similar to that between the BTV serotypes. In addition, EHDV-2NA showed apparently closer relationship against EHDV-1 than other EHDV-2 isolates as reported previously [2, 4], resulting in different branching on the phylogenetic tree compared to S7 analysis. Previous phylogenetic analyses based on Orbivirus gene segment L3 indicated a geographical distribution of different genotypes [4], but not S7 gene [24]. The different branching in EHDV-2 isolates might be also due to these geographical evolution and/or to reassortment between EHDVs as seen in BTV-1 South Africa and Australia isolates [3]. Furthermore, we previously described the distinction of L2 sequence between IBAV and EHDV-2NA and suggested sero-reactivity might not be consistent to genetic relationship. The findings obtained in this study has also speculated this point. The inner core proteins, especially VP7 which is important for insect cell attachment, were considered to be important for infectivity to insect cells and the insect vector species might affect the evolution of inner capsid proteins [2, 9, 24]. Similarly, EHDV might have evolved in the same fashion.

As mentioned above, the complete nucleotide sequences of segment L3 and S7 genes of IBAV was determined. IBAV showed close evolutionary relationships to Australia isolate of EHDV-2 and some genetic difference against EHDV-2NA by phylogenetic analyses based on L3 and S7 genes, suggesting Japan and Australia isolates of EHDV-2 might form a subgroup of an EHDV-2 serotype. In BTV, it was reported that subunit vaccine using CLPs and VLPs synthesized by co-expression of structure proteins protected sheep from BTV challenge [17]. In order to develop subunit vaccine of IBAV and to investigate its structure and function, the data in this study might contribute to further studies such as co-expressions of several virus proteins using recombinant techniques.

REFERENCES

the complete nucleotide sequence of the capsid protein (VP3) of Australian epizootic haemorrhagic disease of deer virus (serotype 2) and cognate genes from other Orbiviruses. *Virus Res.* 21: 1–18.


