

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-
19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# Analysis of a hypervariable region in the $3^{\prime}$ non-coding end of the infectious bronchitis virus genome 

A.K. Williams, L. Wang, L.W. Sneed and E.W. Collisson<br>Department of Veterinary Pathobiology, Texas A\&M University, College Station, TX 77843, USA

(Received 13 August 1992; revision received 3 December 1992; accepted 4 December 1992)


#### Abstract

Summary

Previous studies on infectious bronchitis virus (IBV) cDNA have identified a region of about 184 bases in the $3^{\prime}$ non-coding terminus of both the U.S. prototype strain (Beaudette) and a Japanese strain (KB8523), that was not present in an antigenically closely related U.S. strain, Massachusetts (Mass) 41 (Boursnell et al., 1985; Sutou et al., 1988). In order to investigate the origin and function of this region and its occurrence in nature, the cDNA sequences of the $3^{\prime}$ non-coding regions of three additional strains of IBV, Gray, Arkansas (Ark) 99 and Holland (Holl) 52, were determined and compared to the sequences of the Beaudette, KB8523 and Mass41 strains. Not only was this U-rich sequence absent from the $3^{\prime}$ non-coding region of the Mass 41 strain, it was also highly variable, especially in comparison to the highly conserved $3^{\prime}$ non coding region downstream of this sequence. Computer analyses of the sequences adjacent to this hypervariable region (HVR) showed that the $3^{\prime}$ end of the IBV genome was highly conserved downstream of this region, with 94.3 to $97.8 \%$ similarity. However, the similarities for the HVR ranged from $53.2 \%$ between Holl52 and Ark99, to $92.8 \%$ between


[^0]Beaudette and Gray. The flanking sequences were not only conserved but these sequences upstream and downstream of the HVR also formed mirrored images.

Infections bronchitis virus; Hypervariable region; Sequencing; $3^{\prime}$ non-coding region

Avian infectious bronchitis virus (IBV) is a highly contagious pathogen of chickens with many strains varying widely in virulence, serotype, and pathotype (Darbyshire et al., 1979; Hopkins, 1974). For example, Beaudette is a highly attenuated strain of IBV, Holl52 is a vaccine strain, Gray and a Japanese strain (KB8523) are both known to be nephropathogenic but have also been associated with respiratory disease, and Ark99 and Mass 41 are generally thought to result in severe respiratory disease in the absence of gross kidney lesions (Beaudette and Hudson, 1937; Cumming, 1963; Darbyshire et al., 1979; Johnson and Marquardt, 1986; Hungherr et al., 1956; Sutou et al., 1988; Winterfield and Hitchener, 1962). ArkDPI strains (embryonating chicken egg passages 11 and 75 ) belong to the same serogroup as Ark99 (Sneed et al., 1989).

IBV was the first coronavirus to be completely sequenced and is the prototype of the family Coronaviridae (Boursnell et al., 1987). The IBV genome is 27.6 kb in length and consists of single-stranded RNA of positive polarity with a polyA tail (Schochetman et al., 1977). Viral replication takes place under the control of a viral RNA-dependent RNA polymerase produced by translation of the genomic RNA. The polymerase presumably binds to the $3^{\prime}$ end of the genomic RNA to initiate transcription, resulting in the production of negative sense RNA, either full length or subgenomic length, from which mRNAs can be transcribed (Boursnell et al., 1987, Sethna et al., 1989).

Previous studies on IBV cDNA for the $3^{\prime}$ non-coding terminus of both a U.S. strain (Beaudette) and a Japanese strain (KB8523) identified a region of 184 bp with a higher $U$ content that the remainder of the genomic RNA (Boursnell et al., 1985; Sutou et al., 1988). This region, however, was not present in Mass 41 , an antigenically closely related strain. The $3^{\prime}$ non-coding regions of additional strains were cloned and sequenced in order to investigate functional importance and to determine the extent of this U-rich sequence within the IBV group of Coronaviruses. A similar U-rich region of 184 bases was found in the sequence of the Ark99 genome immediately downstream of the nucleocapsid gene where it had been described in the Mass-related strains. A 40-base synthetic oligonucleotide was synthesized that was complementary to the middle of this Ark99 U-rich region, and was labelled at the $3^{\prime}$ end using ${ }^{32} \mathrm{P}$-dCTP and terminal deoxy transferase (Sambrook et al., 1989). Four IBV RNAs, Mass41, Holl52 and 2 Arkansas isolates (ArkDPI-11 and ArkDPI-75) were applied to and baked on a nitrocellulose membrane. The RNA was applied in four-fold dilutions, starting with $1 \mu \mathrm{~g}$, and dot blot hybridisation was performed at $42^{\circ} \mathrm{C}$ overnight in the presence of formamide (Sambrook et al., 1989).

## Holl52



## Mass41

Fig. 1. Dot blot of RNA from 4 strains of IBV. RNA was immobilized onto a nitrocellulose membrane and probed with a ${ }^{32} \mathrm{P}$-labelled synthetic oligonucleotide complementary to a portion of the U-rich region in the $3^{\prime}$ non-coding region of Ark99.

Hybridization studies with this Ark99-derived probe detected similar sequences in genomic RNA from the ArkDPI-11 and ArkDPI-75 strains, but not in Mass 41 or Holl52 RNA, indicating, as expected, the presence of a similar U-rich sequence in ArkDPI-11 and ArkDPI-75 (Fig. 1). However, this probe also did not hybridize to Holl52, suggesting either that the strain did not contain this region or that it was too dissimilar to be recognized by the Ark99-derived probe. The cDNA sequences of two additional strains, Gray and Holl52, were then examined, and a comparable region was found in both.

Since the $3^{\prime}$ non-coding region of the IBV genome should be important in the recognition and/or binding of polymerase, we examined the sequences of this region of the genome. The viral isolates used in this study, previously described by Sneed et al. (1989), were propagated at $37^{\circ} \mathrm{C}$ for $36-48 \mathrm{~h}$ following allantoic sac inoculation into 11 -day-old specific pathogen-free embryonating chicken eggs. The eggs were then refrigerated at $4^{\circ} \mathrm{C}$ for 5 h , and the allantoic fluid collected and clarified by centrifugation at 10 K rpm for 20 min . Virus was precipitated overnight with $8 \%$ polyethylene-glycol, $2.33 \% \mathrm{NaCl}$, and concentrated by centrifugation at 10 K rpm for 20 min . Virus pellets were reconstituted in $1 / 50$ vol. 20 mM Tris, 1 mM EDTA, $100 \mathrm{mM} \mathrm{NaCl}, \mathrm{pH} 7.4$ and banded on a $30-50 \%$ glycerol/potassium tartrate gradient for 2 h at 28 K rpm . After concentrating by ultracentrifugation at 28 K rpm for 1 h , virus pellets were resuspended in $0.1 \mathrm{M} \mathrm{NaOAc}, \mathrm{pH} 5.2$, and the virions were disrupted with $1 \mathrm{mg} / \mathrm{ml}$ proteinase K and $1 \%$ SDS at $37^{\circ} \mathrm{C}$ for 1 h . The RNA was extracted first with phenol and then with chloroform/isoamyl alcohol followed by ethanol precipitation (Wang et al., 1988). Presence of IBV genomic RNA was confirmed by electrophoresis of a sample of the RNA on a $1 \%$ agarose gel and the remaining RNA was reprecipitated with ethanol.

Purified RNA was pelleted and dried in a Savant vacuum drier. RNA was resuspended in $7 \mu \mathrm{l}$ deionized $\mathrm{H}_{2} \mathrm{O}\left(\mathrm{dH}_{2} \mathrm{O}\right)$ and denatured in 6 mM methyl mercuric hydroxide for 10 min at room temperature. First strand cDNA synthesis was carried out using 5 units of AMV reverse transcriptase at $42^{\circ} \mathrm{C}$ for 2 h in the presence of 14 mM 2 -mercaptoethanol ( 2 ME ), 50 mM Tris $\mathrm{pH} 8.3,50 \mathrm{mM} \mathrm{KCl}, 8$ $\mathrm{mM} \mathrm{MgCl}_{2}, 0.8 \mathrm{~mm} \mathrm{dNTP}, 20 \mu \mathrm{~g} / \mathrm{ml}$ of $3^{\prime}$ end primer ( $5^{\prime} \mathrm{GGATCCGCTC}$ TAACTCTATACTAGCCTAT $3^{\prime}$ ) and 20 units of RNasin (Promega, Madison, WI). Following phenol extraction and ethanol precipitation of the RNA/cDNA
hybrid, the RNA was eliminated by exposure to highly alkaline conditions at $37^{\circ} \mathrm{C}$ for 3 h . The cDNA was then neutralized, concentrated by ethanol precipitation, and amplified using the polymerase chain reaction (PCR). PCR was performed with the same $3^{\prime}$ end primer as above and a primer from the upstream intergenic region ( $5^{\prime}$ GAATTCCCGCGTGTACCTCTCTAGTA $3^{\prime}$ ) in the presence of 4 mM dNTP, $1 \times$ Taq buffer ( $1.5 \mathrm{mM} \mathrm{MgCl} 2,50 \mathrm{mM} \mathrm{KCl}, 10 \mathrm{mM}$ Tris- $\mathrm{HCl}, \mathrm{pH} 8$ ) and $0.5 \mu \mathrm{l}$ Taq polymerase (Perkin Elmer, Norwalk, CT). The PCR was performed with a primary denaturation step at $95^{\circ} \mathrm{C}$ for 3 min , annealing at $50^{\circ} \mathrm{C}$ for 30 s , extension at $72^{\circ} \mathrm{C}$ for 1 min , and subsequent denaturation steps at $95^{\circ} \mathrm{C}$ for 1 min . A total of 30 cycles were used with a final extension step of 3 min at $72^{\circ} \mathrm{C}$. Resulting PCR products were cloned into the pCR1000 vector, utilizing the single T overhang at the insertion site (Invitrogen, San Diego, CA) and sequenced using the Sequenase rit (USB, Cleveland, OH). Nucleotide sequences were processed using the GCG Seqed program and compared with the Gap and Bestfit programs. Sequences were aligned using Pileup and predicted secondary structures of the $3^{\prime}$ non-coding region were derived from Fold and Squiggles.

The nucleotide sequences of the $3^{\prime}$ non-coding regions of the Ark99, Gray and Holl52 strains of IBV were compared with the published sequences for Mass41, Beaudette and KB8523 (Boursnell et al., 1985; Sutou et al., 1988). Fig. 2 shows the alignment of these sequences and demonstrates the high degree of similarity among the strains.

However, the U-rich region upstream of the $3^{\prime}$ non-coding region, that was absent from Mass41, was found in Holl52, as well as Gray and Ark99, and was described as the $3^{\prime}$ hypervariable region (HVR) because of the extent of its variation among strains. This HVR ranged from 184 (Gray) to 196 (Holl52) nucleotides in length and the nucleotide sequence homologies among the strains containing this sequence ranged from $53.2 \%$ to $92.8 \%$, with the Holl42 strain showing the greatest divergence (Table 1). The greater variability of the Holl52 strain compared to the other strains in this region immediately downstream of the nuclcocapsid gene would explain why this region was not detected with the Ark99 probe (Fig. 1). In comparison, the non-coding region downstream of this HVR was highly conserved, ranging from $94.3 \%$ to $97.5 \%$ identity (Table 2). Fig. 3 shows a plot of the similarity of the nucleotide sequences of the $3^{\prime}$ non-coding regions of all 6 strains of IBV. The HVR is clearly seen immediately after the nucleocapsid gene and can be compared to the remaining $3^{\prime}$ non-coding region which has very high similarity scores. This HVR could be considered either an insertion or a deletion since it is not present in Mass41. However, the higher U content in comparison to the rest of the IBV genome suggests that it is of exogenous origin (Boursnell et al., 1985). The fact that it has at least $53.2 \%$ identity among various strains suggests a common ancestry for this sequence.

If the IBV genome acquired the HVR as exogenous RNA, this region could be a site for recombination. Sequences upstream and downstream of the HVR were analyzed to identify any distinctive patterns which would suggest that this region is of functional significance, such as being involved in recombination. Not only are the sequences flanking the $3^{\prime} \mathrm{HVR}$ highly conserved, the $5^{\prime}$ flanking sequence also

## 319 <br>  <br>      <br>  <br> 

65

## K42 <br> owtulitoz










析





Bea Hol 152









89929

篓
歌
4.
4
4
8
8
6
5
5
















Hol 152

## $\cdots$ 

## A


m




8



？

## 8

s32





 169 Mass41
Hol152

 тыстдaco дzстдаса
 ark geatrcagct traggtaga atgactitg a．g．．．．．．．．．．．．．．tat 19





Fig．2．Nucleotide sequence comparison of the $3^{t}$ non－coding region of 6 strains of IBV．Nucleotides are numbered from the stop codon for the

TABLE 1
Percent homologies of the $3^{\prime}$ HVR

|  | Ark99 | Gray | Beau | Holl52 | KB8523 |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Ark99 | 100 | 91.8 | 91.8 | 53.2 | 68.6 |
| Gray | 91.8 | 100 | 92.8 | 57.9 | 67.6 |
| Beau | 91.8 | 92.8 | 100 | 57.1 | 66.7 |
| Hol152 | 53.2 | 57.9 | 57.1 | 100 | 60.0 |
| KB8523 | 68.6 | 67.7 | 66.7 | 60.0 | 100 |

TABLE 2
Percent homologies of the $3^{\prime}$ non-coding region of IBV

|  | Ark99 | Gray | Beau | Holl52 | KB8523 | Mass41 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Ark99 | 100 | 97.2 | 97.5 | 94.7 | 97.2 | 96.3 |
| Gray | 97.2 | 100 | 97.8 | 94.3 | 96.8 | 96.6 |
| Beau | 97.5 | 97.8 | 100 | 94.7 | 97.2 | 97.5 |
| Holl52 | 94.7 | 94.3 | 94.7 | 100 | 95.3 | 94.3 |
| KB8523 | 97.2 | 96.8 | 97.2 | 95.3 | 100 | 94.7 |
| Mass41 | 96.3 | 96.6 | 97.5 | 94.3 | 94.7 | 100 |

resembles the cutting site ( $5^{\prime}$ TTAA $3^{\prime}$ ) of retrovius DNA and transposons Tn10 and Tn3 (Shoemaker et al., 1980). It was of special interest that a unique pattern, that may be of functional importance, was observed in the oligonucleotides flanking the HVR. Several nucleotides downstream of this $3^{\prime}$ HVR were mirrored in the sequences upstream. That is, the same sequences were seen in reverse orientation on either side of the $3^{\prime}$ HVR. An example can be seen in the Gray strain where the sequence TGAGAATG appears $5^{\prime}$ of the HVR and the sequence


Fig. 3. Similarity plot of the $3^{\prime}$ non-coding region of six strains of IBV. This plot starts from the end of the nucleocapsid gene and covers the $3^{\prime}$ HVR and conserved $3^{\prime}$ end.

TABLE 3
Mirrored sequences flanking the hypervariable region
$\left.\right|^{*} \mid$ indicates the position of the insert. Sequences mirrored on both sides of hypervariable region are underlined.

| ARK9 | TGAGAATGAACTTTGAA ${ }^{*} \mid$ AGTtCAATAGTAAGAGT |
| :---: | :---: |
| GRAY | TGAGAATGAACTCTGAGG\|*AGTTCAATAGTAAGAGT |
| BEAU | AgAGAATGAACTTTGAG\|* AATTCAATAGTAAGAGTT |
| KB8523 |  |
| MASS41 | gagantganctitgagt* AaAATtcastagtangag |
| HOLL52 | gGagagatgagttgital*\|gtatcantaagantcangg |

GIAAGAGT is in the $3^{\prime}$ flanking sequence (Table 3). The function of these mirrored sequences is unknown; however, they may be important in preventing folding of the RNA in this region, as mirrored sequences should prevent complementary base pairing.


Fig. 4. Phylogenetic tree of 6 strains of IBV based on the sequences of their $3^{\prime}$ non-coding regions.

The sequence relationships of the $3^{\prime}$ non-coding regions were examined with a phylogenetic tree (Fig. 4). The Holl52 and KB8523 strains were shown to be more distantly related to the other strains. The 3 ' non-coding sequences of Mass41, in spite of the absence of the HVR, appeared to be closely related to the Beaudette, and to a lesser extent to the Gray and Ark99 strains. The program used to determine this tree counts the absence of this whole HVR in Mass 41 as the result of one recombination event, so only the remainder of the $3^{\prime}$ non-coding region of the Mass 41 strain was considered for the construction of the phylogenetic tree.

The location of the $3^{\prime}$ non-coding HVR is conserved at a position immediately downstream from the stop codon of the nucleocapsid gene. There is no evidence as yet to show that this HVR has any effect on virulence or pathogenesis. There are no detectable protein products from the $3^{\prime} \mathrm{HVR}$ and no methionines with a preceding putative consensus ribosomal binding site. Since it is completely absent in the Mass41 strain, it would appear that it is not an absolute requirement for viral replication, at least for Mass41. Any specific function of this HVR apparently does not dictate that the sequences be highly conserved but its function may depend on the secondary structure of the RNA. The variability of this HVR as compared to the conservation of the remaining 3 ' non-coding sequences suggests that the primary sequence itself may not be as functionally important as its influence on the secondary structure of the $3^{\prime}$ end of the genome. It is possible that the presence of the HVR may increase the complexity of the secondary structure of the RNA which could in turn affect its association with replication machinery, such as RNA binding proteins.

## Acknowledgements

We thank Peggy M. Billingsley for her excellent technical assistance. This work was supported by the Southeastern Poultry and Egg Association, \#29 and \#460; Formula Animal Ilealth (section 1433), \#TEXO-6824 and TATRP \#999902-063.

## References

Beaudette, F.R. and Hudson, C.B. (1937) Cultivation of the virus of infectious bronchitis. J. Am. Vet. Med. Ass. 90, 51-60.
Boursnell, M.E.G., Binns, M.M., Foulds, I.J. and Brown, T.D.K. (1985) Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. J. Gen. Virol. 66, 573-580.
Boursnell, M.E.G., Brown, T.D.K., Foulds, I.J., Green, P.F., Tomley, F.M. and Binns, M.M. (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J. Gen. Virol. 68, 57-77.
Cumming, R.B. (1963) infectious avian nephrosis (uraemia) in Australia. Aust. Vet. J. 39, 145-147.
Darbyshire, J.H., Rowell, J.G., Cook, J.K.A. and Peters, R.W. (1979) Taxonomic studies on strains of avian infectious bronchitis virus using neutralization tests in tracheal organ cultures. Arch. Virol. 61, 227-238.
Hopkins, S.R. (1974) Serologic comparisons of strains of infectious bronchitis using plaque-purified isolates. Avian Dis. 18, 231-239.

Johnson, R.B. and Marquardt, W.W. (1976) Strains of infectious bronchitis virus on the Delmarva Peninsula and in Arkansas. Avian Dis. 20, 382-386.
Jungherr, E.L., Chomiak, T.W. and Luginbuhl, R.E. (1956) Immunologic differences in strains of infectious bronchitis virus. In: Proceedings of the 60th annual meeting of United States Livestock Sanitary Association" pp. 203-209.
Sambrook, S., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning - a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press.
Schochetman, G., Stevens, R.H. and Simpson, R.W. (1977) Presence of infectious polyadenylated RNA in the coronavirus avian infectious bronchitis virus. Virology 77, 772-782.
Sethna, P.B., Hung, S. and Brian, D.A. (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc. Natl. Acad. Sci. USA. 86, 5626-5630.
Shoemaker, C., Goff, S., Gilboa, E., Paskind, M., Sudha, W. and Baltimore, D. (1980) Structure of a cloned circular Maloney murine leukamia virus DNA molecule containing an inverted segment: implications for retrovius integration. Proc. Natl. Acad. Sci. USA 77, 3932-3936.
Sneed, L.W., Butcher, G.D., Wang, L., Parr, R. and Collisson, E.W. (1989) Comparisons of the structural proteins of avian infectious bronchitis virus as determined by Western blot analysis. Viral Immunol. 2(3), 221-227.
Sutou, S., Sato, S., Okabe, T., Nakai, M. and Sasaki, N. (1988) Cloning and sequencing of genes encoding structural proteins of avian infectious bronchitis virus. Virology 165,589-595.
Wang, L., Kemp, M.C., Roy, P. and Collisson, E.W. (1988) Tissue tropism and target cells of Bluetongue virus in the chicken embryo. J. Virol. 62, 887-893.
Williams, A.K., Wang, L., Sneed, L.W. and Collisson, E.W. (1992) Comparative analyses of the nucleocapsid genes of several strains of infectious bronchitis virus and other coronaviruses. Virus Res. 25, 213-222.
Winterfield, R.D. and Hitchener, S.B. (1962) Etiology of an infectious nephritis-nephrosis syndrome of chickens. Am. J. Vet. Res. 23, 1273-1279.


[^0]:    Correspondence to: E.W. Collisson, Department of Veterinary Pathobiology, Texas A\&M University, College Station, TX 77843, USA.
    The sequences reported in this paper have been submitted to Genbank and assigned the following accession numbers, with descriptions in parentheses: 106251 (Ark99 strain 507 bp ): L06252 (Gray strain 504 bp ): L06253 (Holl52 strain 528 bp ).

