SEQUENCE COMPARISONS OF THE 3' END OF THE GENOMES OF FIVE STRAINS OF AVIAN INFECTIOUS BRONCHITIS VIRUS

Ellen W. Collisson[1], Anna K. Williams[1], Ray Vonder Haar[2], Wang Li[1], and Loyd W. Sneed[1]

[1]Department of Veterinary Microbiology
[2]Department of Biology
Texas A&M University
College Station, TX

INTRODUCTION

Avian infectious bronchitis virus (IBV) causes an acute, highly contagious respiratory disease of chickens characterized by tracheal rales, coughing and sneezing[1]. The disease was first described in 1931 by Schalk and Hawn[1] and since that time, many strains have been defined[2,3]. These strains vary widely in virulence and tissue tropism. A number of serologically distinguishable strains of infectious bronchitis virus have been isolated from poultry in the U.S.A. Of the strains included in this study, Beaudette, Conn. and Ark DPI 75 are vaccine strains, Gray and AustT are known to be nephropathogenic causing limited respiratory disease, and infections with Ark99, Mass41 and a Japanese strain (KB8523) are generally thought to result in severe respiratory disease in the absence of nephritis and nephrosis[4].

In this study, the 3' end of the Ark99 and Gray strains were compared with the published data for Mass41 and Beaudette[5] and a Japanese strain, KB8523[6]. An insert in the 3' noncoding region of the genome described in the Beaudette and Japanese strains, but absent in the Mass41 strain, was also found in the Ark99 and Gray strains.

MATERIALS AND METHODS

Viral Preparation

The Gray and Ark99 strains of IBV were purified in our lab by three terminal dilution cycles in embryonating chick embryos (ECE) and were propagated by allantoic sac inoculation into 11-day old specific pathogen free (SPAFAS) ECE. Virus was precipitated with polyethylene-glycol and banded on a 30-50% glycerol/potassium tartrate gradient. After concentrating by ultracentrifugation, virus was reconstituted and the virions were disrupted with Proteinase K and SDS. The RNA was extracted in phenol/chloroform/isoamyl and ethanol precipitated[7].

Cloning of The Gray and Ark99 Strains

First strand cDNA synthesis was carried out by reverse transcriptase using an oligo dT primer and second strand synthesis with DNA Poll and RNase H[8]. The double stranded cDNA was tailed with deoxy C's using

Coronaviruses and Their Diseases
Edited by D. Cavanagh and T.D.K. Brown
Plenum Press, New York, 1990

Table 1. Percent similarity among the four strains having the insert in the 3' non-coding region not present in the Mass41 strain.

	Beau	KB8523	Ark99	Gray
Beau	100	67.4	91.9	92.9
KB8523	67.4	100	69.4	69.9
Ark99	91.9	69.4	100	94.6
Gray	92.9	69.9	94.6	100

Table 2. Percent similarities among the 3' non-coding regions of the genomes of five strains of IBV.

	Mass41	Beau	KB8523	Ark99	Gray
Mass41	100	99.1	97.1	93.3	96.9
Beau	99.1	100	98.1	94.0	97.7
KB8523	97.1	98.1	100	93.2	95.1
Ark99	93.3	94.0	93.2	100	96.1
Gray	96.9	97.7	95.1	96.1	100

terminal deoxy transferase and annealed with oligo dG tails in the Pst1 site of the pUC9 plasmid[9]. Clones containing 1-2Kb of IBV cDNA were selected after transformation of E. coli JM109 cells. Dideoxy sequencing of plasmid cDNA[10] and of single-stranded cDNA following subcloning into M13[11] was performed, and the resulting sequences of Ark99 and Gray were compared with each other and with the published data for the Mass41 and Beaudette strains[5] and the Japanese strain, KB8523[6], using the University of Wisconsin Genetics Computor Group programs.

Results

The 3' ends of the genomes of the Gray and Ark99 strains of IBV were cloned and sequenced. The 1712 bases of the cDNA for Gray included the entire nucleocapsid gene and 346 bases of the 3' non-coding region. Approximately 170 bases were missing from the 3' end of the Gray clone as determined by comparing the sequence with the other strains (Fig.1). Other Gray cDNA clones are currently being sequenced to complete this data. The cDNA of the Ark99 clone was 1255 bases in length and apparently included all of the 3' non-coding region but according to the data from the other four strains, Ark99 was missing 485 bases of the 5' end of the nucleocapsid structural gene (fig.1). Therefore, comparisons with the Ark99 strain were based on the 3' 742 bases of the nucleocapsid coding region (247 amino

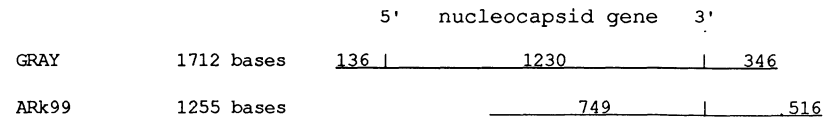

Fig.1. cDNA clones of the 3' ends of Gray and Ark99

acids) present in this clone.The complete open reading frame for the nucleocapsid gene of Gray, as with Beau, Mass41 and KB8523 contained 1227 bases, contain Kozac's consensus sequence at the AUG start codon and coded for a basic protein of 409 amino acids. The available amino acid sequencing data of the nucleocapsid proteins of Gray and Ark99 were compared to each other and with the amino acid sequences for the nucleocapsid proteins of Beaudette, Mass41 and KB8523. The similarities of the amino acid sequences of the nucleocapsid proteins of these 5 strains of IBV ranged from 90.7 to 96.3 with the Gray strain showing the least overall similarity to the other strains. There is an area of divergence in the Gray nucleocapsid protein sequence from residues 230-250 when compared to the other strains as can be seen in Fig.2 which shows the alignment comparison of the amino acids of the Gray and Beaudette strains. However, the significance of this apparent divergence is unknown.

The nucleotide sequences of the 3' non-coding regions of Gray and Ark99 were compared with each other and with the published data for Mass41, Beaudette and a Japanese strain KB8523. In the 3' end of the genome, 4 bases downstream from the stop codon for the nucleocapsid gene, there is a region that ranges from 184 to 187 bases in length that is present in the KB8523, Beaudette, Gray and Ark99 strains and is missing in the Mass41 strain. This region was from 67.4 to 94.6% similar among the strains containing this sequence with the Japanese strain being the most divergent (Table 1).

```
          .         .         .         .         .
  1 MASGKAAGKTDAPAPVIKLGGPKPPKVGSSGNASWFQAIKAKKLNTPPPK 50
    |||||| |||||||||||||||||||||||||||||||||||| | ||
  1 MASGKATGKTDAPAPVIKLGGPRPPKVGSSGNASWFQAIKAKKLNSPQPK 50
          .         .         .         .         .
 51 FEGSGVPDNENIKPSQQHGYCRRQARFKPGKGGRKPVPDAWYFYYTGTGP 100
    |||||||||||||| |||||| |||||||||||| ||||||||||||||||
 51 FEGSGVPDNENFKTSQQHGYWRRQARFKPGKGRRKPVPDAWYFYYTGTGP 100
          .         .         .         .         .
101 AADLNWGDTQDGIVWVAAKGADTKSRSNQVTRDPDKFDQYPLRFSDGGPD 150
    |||||||| |||||||||||||| |||||| ||||||||||||||||||||
101 AADLNWGDSQDGIVWVAAKGADVKSRSNQGTRDPDKFDQYPLRFSDGGPD 150
          .         .         .         .         .
151 GNFRWDFIPLNRGRSGRSTAASSAAASRAPSREGSRGRRSDSGDDLIARA 200
    |||||||||||||||||||||||| |||||||||||||||||||||||||
151 GNFRWDFIPLNRGRSGRSTAASSAASSRPPSREGSRGRRSGSEDDLIARA 200
          .         .         .         .         .
201 AKIIQDQQKKGSRITKAKADEMAHRRYCKRTIPPNYRVDQVFGPRTKGKE 250
    ||||||||||||||||||||||         |    ||   | | |||||
201 AKIIQDQQKKGSRITKAKADEMVIAGIASALFHLVIRLIKFLVPGTKGKE 250
          .         .         .         .         .
251 GNFGDDKMNEEGIKDGRVTAMLNLVPSSHACLFGSRVTPKLQLDGLHLRF 300
    |||||||||||||||||||||||||||||||||||||||||||| |||||||
251 GNFGDDKMNEEGIKDGRVTAMLNLVPSSHACLFGSRVTPKLQPDGLHLKF 300
          .         .         .         .         .
301 EFTTVVPCDDPQFDNYVKICDQCVDGVGTRPKDDEPKPKSRSSSRPATRG 350
    ||||||| ||||||||||||||||||||||||||||||||||||||||||
301 EFTTVVPRDDPQFDNYVKICDQCVDGVGTRPKDDEPKPKSRSSSRPATRT 350
          .         .         .         .         .
351 NSPAPRQQRPKKEKKLKKQDDEADKACTSDEERNNAQLEFYDEPKVINWG 400
     |||||||| ||||| |||||| ||| |||||||||||| |||||||||
351 SSPAPRQQRLKKEKRPKKQDDEVDKALTSDEERNNAQLEFDDEPKVINWG 400
          .
401 DAALGENEL* 410
    | |||||||
401 DSALGENEL* 410
```

Fig.2. Optimal alignment of the amino acid sequences of Beaudette and Gray from the UWGCG GAP program.

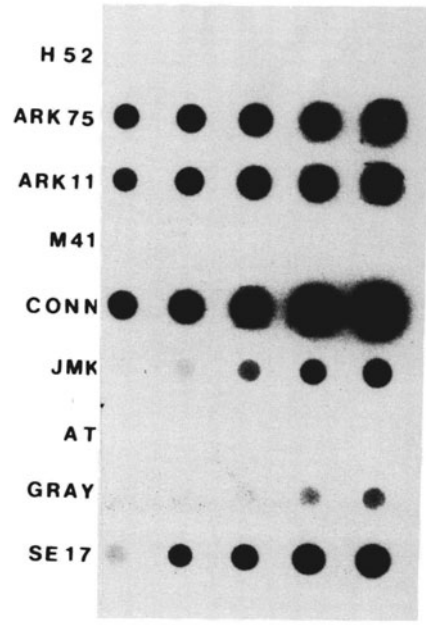

Fig.3. Northern blot analysis of 9 strains of IBV using 3'non-coding region "insert"-specific 40bp probe.

In order to determine the occurence of this 184bp region in various strains of IBV, a 40 base oligonucleotide was synthesized to the middle portion of the Ark insert and hybridized[12] to genomic RNA of several IBV strains (fig.3). In addition to the Ark99 and Gray strains, the Conn, JMK and SE17 strains had the 184 base insert; whereas the Mass41, Aust T and the Holl52 strains were missing this region. The remaining noncoding region was from 93.2 to 99.1% similar among the strains (Table 2) with the Ark99 strain showing somewhat less homology. The putative secondary structure of the non-coding 3' ends indicated that the Mass41 without the "insert" had two branches with hairpin loops and the other strains showed 4 hairpin loops with Beaudette, 5 with Gray, and 6 with the KB8523 and Ark99 strains.

Ark99
GGAT TCAGCTTT AGG TGA GAAT G AAC T TTGA GTAA||AGTT CAA TAG TAAG AGT TA AGG GAGA TAGG

Gray
AGG TGAGAATG AACT C TGA GTAA||AGT TCAA TA GTAAGAGT TAAGGA

Beaudette
GAGAATGA ACT TTGAG TAA||AAT TCA AT AGTAAG AG

Mass41
TAG GAGA GAAT GAA CTT TGA G TAA|AAT TCAAT AGT AAG AGT TAAG GAA GAT

K8523
GAA AA TGAA C TTTGA TTAA||AGTTT ATTGA AAGT T AAG GA

Fig.4. Sequences flanking the 3' non-coding region "insert." Underlined sequences represent mirrored sequences and the bold letters represent common sequences adjacent to the "insert."

However, these structures were computed only with the 3' non-coding region, and therefore the impact of the rest of the genome is not known. Also, the completion of the sequencing of the 3' end of Gray may result in an altered secondary structure for this strain. The significance of the differences resulting from the "insert" sequences on the function of this end of the genome is not known. This extemely AT rich "insert" appears to be in the same location in several strains of IBV (fig.4.) and mirrored sequences flanking this region also appear as loops in the secondary structures of the strain in which this sequence is absent. The sequence 5' of this 184bp insert is consistently TAA and the sequence 3' is either AATT or AGTT in the strains analyzed to date. Boursnell et al.[5] found the sequence AGTTTA to be repeated 6 times downstream of the "insert" in the 3' noncoding region of the Mass41 and Beau strains and the sequence TTTAGTTTAA repeated 3 times. Seven repeats of AGTTTA were found in the corresponding region of KB8523 and 5 in Ark99, and the latter sequence was repeated twice in the KB8523 and Ark99 strains. This portion of the 3'non-coding region of Gray was not represented in this data. The function(s) of the these mirrored and repeated sequences is unknown, but it is possible that they are important in producing a secondary structure facilitating more efficient polymerase and/or leader sequence binding and hence, more efficient transcription.

REFERENCES

1. A. F. Schalk, and M. C. Hawn, An apparantly new respiratory disease of baby chicks, JAVMA.78:418-422 (1931).
2. J. H. Darbyshire, J. G. Rowell, J. K. A. Cook, and R. W. Peters,Taxonomic studies on strains of avian infectious bronchitis virus using neutralization tests in tracheal organ cultures, Arch.Virol. 61:227-238 (1979).
3. S. R. Hopkins, Serologic comparisons of strains of infectious bronchitis using plaque-purified isolates, Avian Dis. 18:231-239 (1974).
4. R. B. Cumming, The etiology of uremia of chickens, Aust.Vet.J. 39:145-147 (1963).
5. M. Boursnell, M. Binns, I. Foulds, and T. Brown, Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus, J.Gen.Virol. 66:573-580 (1985).
6. S. Shizuyo, S. Seiji, T. Okabe, M. Nakai, and N. Sasaki, Cloning and sequencing of genes encoding structural proteins of avian infectious bronchitis virus, Virology 165: 589-595 (1988).
7. L. Wang, M. C. Kemp, P. Roy, and E. Collisson, Tissue tropism and target cells of Bluetongue virus in the chicken embryo. J.Virol. 62:887-893 (1988).
8. U. Gubler and B. J. Hoffman, A simple and very efficient method for generating cDNA libraries, Gene 25:263-269 (1983).
9. S. L. Berger and A. R. Kimmel, Guide to Molecular Cloning Techniques; in Meth. in Enzym. 152 (1987).
10. E. Y. Chen and P. H. Seeburg, Supercoil sequencing: A fast and simple method for sequencing plasmid DNA, DNA 4(2):165-170 (1985).
11. F. Sanger, S. Nicklen, A. R. Coulson, DNA sequencing with chain-terminating inhibitors. PNAS USA 74:5463-5467 (1977).
12. L. Sneed, G. Butcher, L. Wang, M. Kemp, and E. Collisson, Protein and RNA comparisons of several strains of IBV. Southern Conference on Avian Disease. (March 1987)