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.

Sequence evidence for RNA recombination in field isolates of avian coronavirus infectious bronchitis virus

J.G. Kusters*†, E.J. Jager*, H.G.M. Niesters‡§ and B.A.M. van der Zeijst‡∞

Under laboratory conditions coronaviruses were shown to have a high frequency of recombination. In The Netherlands, vaccination against infectious bronchitis virus (IBV) is performed with vaccines that contain several life-attenuated virus strains. These highly effective vaccines may create ideal conditions for recombination, and could therefore be dangerous in the long term. This paper addresses the question of the frequency of recombination of avian coronavirus IBV in the field. A method was sought to detect and quantify recombination from sequence data. Nucleotide sequences of eight IBV isolates in a region of the genome suspected to contain recombination, were aligned and compared. Phylogenetic trees were constructed for different sections of this region. Differences in topology between these trees were observed, suggesting that in three out of eight strains in vivo RNA recombinant had occurred.

Keywords: RNA recombinant; coronavirus; infectious bronchitis virus

Introduction

Infectious bronchitis virus (IBV) causes considerable damage in the poultry industry by infections of the respiratory tract and the reproductive organs[1]. IBV is the prototype strain of the Coronaviridae, a family of positive stranded RNA viruses with a genome of about 28 kb. This genome contains the information for three structural proteins and a number of enzymes involved in virus replication (*Figure 2*). Although vaccines offer protection, they are often made ineffective by the continuous emergence of new serotypes[2,3]. Serotype-specific protective immunity is thought to be mediated by antibodies to a protein on the surface of the virion, the so-called peplomer protein[4–7]. The peplomer protein is synthesized as a precursor glycoprotein, which is cleaved into the subunits S1 and S2, derived from its N- and C-terminal half, respectively[8–10]. Point mutations are obviously involved in the generation of new serotypes. However, previous studies[11] suggested that RNA recombination is also involved in the generation of new antigenic variants. To address the question of the frequency of recombination in the generation of new field isolates the nucleotide sequences in five windows of homologous sequences of the genomes of eight IBV strains have been compared. To make this comparison the sequences of the S2 genes of D207 and D1466 were newly determined. Both strains are Dutch field isolates and are included in vaccines currently used in The Netherlands. The alignment of the sequences suggests that the Dutch field strain D207[12], the British field strain 6/82[13] and the Japanese field strain KB8523[14] result from RNA recombination events.

*Institute of Molecular Biology and Medical Biotechnology, University of Utrecht, PO Box 80.063, 3508 TD Utrecht, The Netherlands. †Present address. ‡Institute of Infectious Diseases and Immunology, University of Utrecht, PO Box 80.165, 3508 TD Utrecht, The Netherlands. §Present address: Stichting Samenwerking Delftse Ziekenhuizen, PO Box 5010, 2600 GA Delft, The Netherlands. ∞To whom correspondence should be addressed. (Received 2 April 1990; revised 1 June 1990; accepted 1 June 1990)

Materials and methods

The origin of the IBV strains as well as the procedures used for the cloning and sequencing of the peplomer genes have been published previously[3,11,15]. Sequence data were processed using the programs of Microgenic (Beckman Instruments, Ref. 16). Phylogenetic trees were derived from nucleotide difference-matrices using the program of Dr J. Felsenstein (University of Seattle) distributed as part of the PHYLIP package 2.6[17].

Results and discussion

The S2 sequences

The newly determined D207 and D1466 sequences are based on data from two or more independent cDNA clones. With strain D1466, two nucleotide differences between cDNA clones were observed. Neither of these differences resulted in an amino acid replacement. From the nucleotide sequences (submitted to the EMBL Genbank and DDBJ Nucleotide Sequence Databases), amino acid sequences have been deduced. In *Figure 1* the amino acid sequences are listed together with those of all other known S2 sequences.

Assuming the conserved Arg-Arg-X-Arg-Arg-Ser sequence to be the cleavage site between S1 and S2[8], the Ser at position 546 (*Figure 1*) will become the N-terminal S2 residue. In contrast to S1, where many insertions/deletions were found, the variation in length of S2 is

0264-410X/90/060605-04

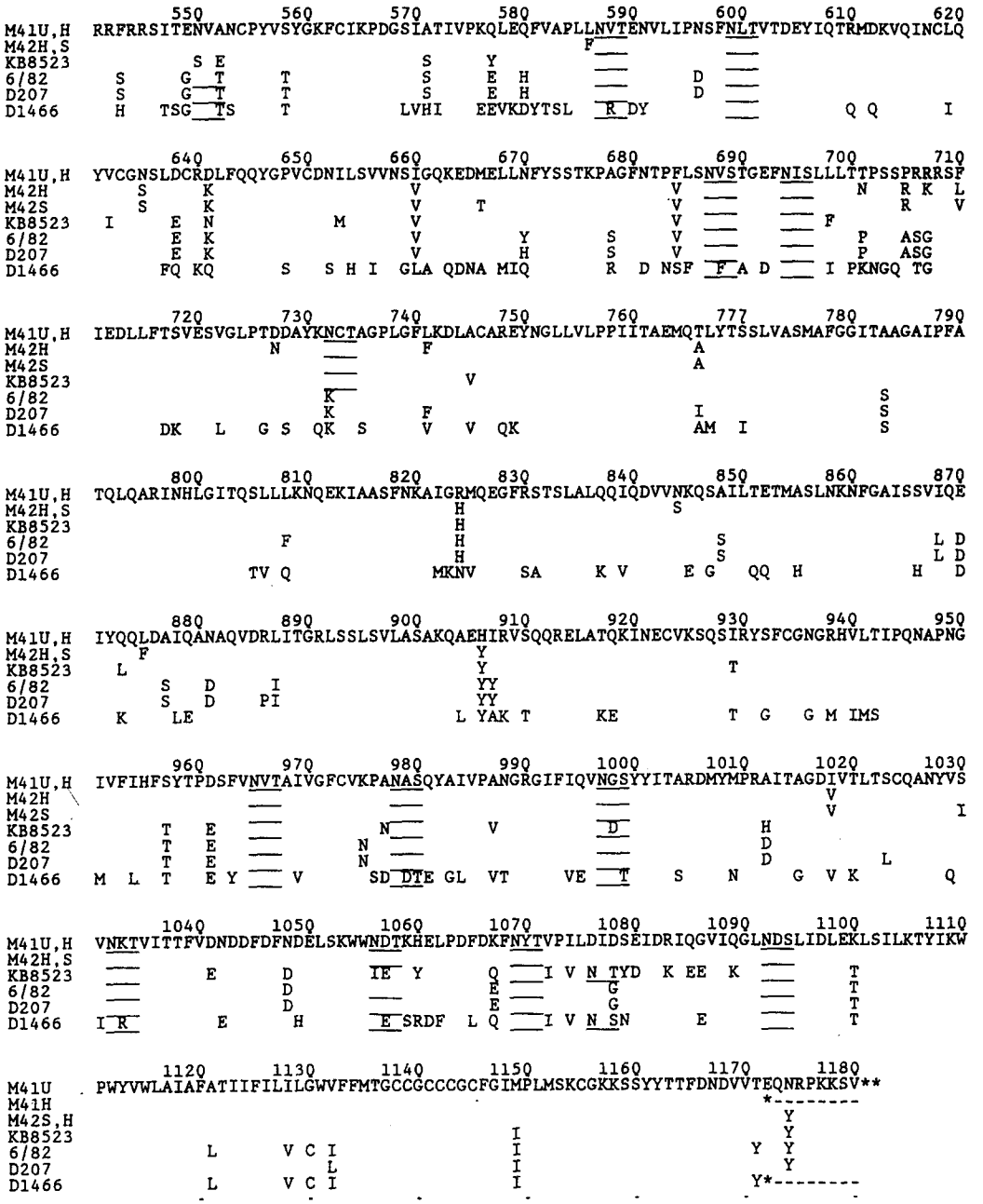

Figure 1 Amino acid sequences of IBV S2 proteins. The sequence of M41 variant M41U has been listed completely; from other strains only the differences with this sequence are shown. The sequences of M41U, and of the M42 variant M42S are from Kusters *et al.*[11]. The sequence of M42 variant M42H is from Binns *et al.*[20]. The sequences of M41 variant M41H and of strain 6/82 are from Binns *et al.*[13]. The sequence of KB8523 is from Sutou *et al.*[14]. Potential *N*-glycosylation sites (NXS or NXT, except when X=P) are underlined. Asterisks indicate the stop codons

caused solely by extension of the open reading frame. Substitution of the T in the TAA stop codon of strains M41H and D1466 by a G in all other strains resulted in an extension of the S2 protein by nine amino acid residues. Generally the amino acid sequences of the S2 proteins differ by less than 10%, strain D1466 being the exception (*Table 1*). However, 19 out of 20 cysteine residues and most glycosylation sites are completely conserved.

Analysis of nucleotide sequences for recombination

The nucleotide and amino acid sequences were compared for possible recombination points. From this comparison we conclude that the Japanese field strain KB8523 is probably a recombinant. The S2 protein of this strain is almost identical to S2 of M41, except for the region corresponding to amino acids 929 to 1102 (*Figure 1*) where the sequence is remarkably similar to D1466. This suggests a recombination event with two crossovers in this region. The alignment of the S2 genes indicates a second potential recombinant: the British field isolate 6/82. Its S1 gene and most of the S2 gene are almost identical to those of strain D207. At the 3′ end of the S2 sequence (at the Lys-1105 codon) the percentage of identical nucleotides switches from 99% upstream to 57% downstream. At exactly the same position the nucleotide sequence of 6/82 becomes highly similar to the D1466 sequence: from 73% identity upstream to 97% downstream. The nucleotide identity between D207 and

Table 1 Difference matrix of S2 sequences of IBV strains[a]

Virus	M41U	M41H	M42H	M42S	KB8523	6/82	D207	D1466
M41U	–	1.4	2.9	2.6	6.7	7.0	6.9	26.7
M41H	1.4	–	4.2	3.8	8.0	8.3	8.2	25.6
M42H	2.9	4.2	–	1.1	7.8	7.5	6.9	27.4
M42S	2.6	3.8	1.1	–	7.5	7.5	7.2	27.4
KB8523	6.7	8.0	7.8	7.5	–	9.0	8.8	26.6
6/82	7.0	8.3	7.5	7.5	9.0	–	1.8	25.4
D207	6.9	8.2	6.9	7.2	8.8	1.8	–	26.6
D1466	26.7	25.6	27.4	27.4	26.6	25.4	26.6	–

[a]The figures represent the percentages of non-identical amino acids. Designation of strains is as in *Figure 1*

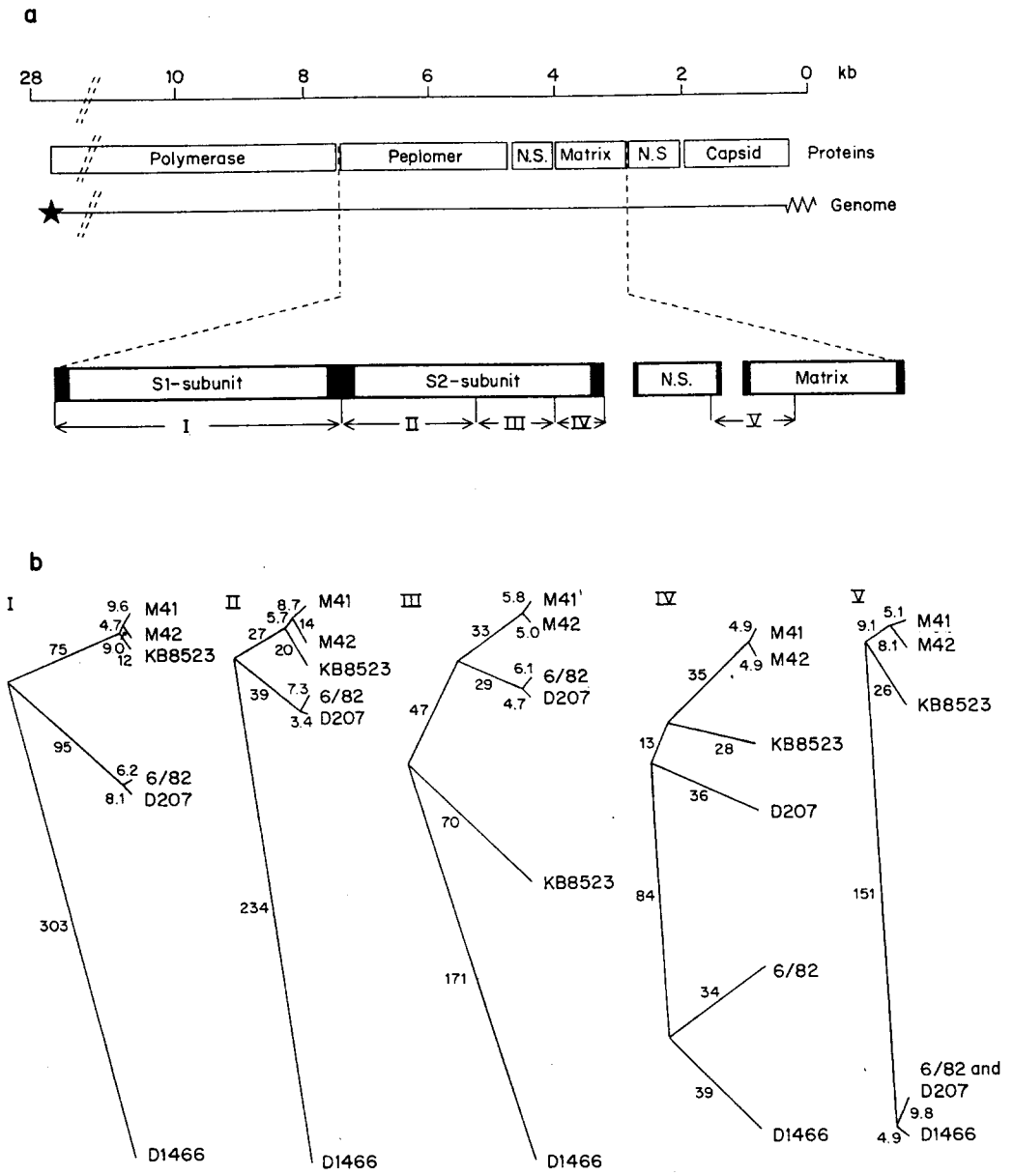

Figure 2 Phylogenetic relationships of IBV strains listed in *Figure 1* (A) Schematic representation of the genome of IBV. The roman numerals I–V indicate the regions used to construct the phylogenetic trees. (B) Phylogenetic trees for the regions I–V as indicated in (A). I, Tree of S1 sequences; II, tree of S2 sequences upstream residue 929 (*Figure 1*); III, tree of S2 sequences between residues 929 and 1105; IV, tree of S2 sequences downstream residue 1105 until the stop codon on position 1182; V, tree of sequences from the 5' part of the E1 gene and part of the preceding intergenic sequences[21]. The figures indicate the permillage of replaced nucleotides. The trees are unrooted, i.e. the position of the hypothetical ancestral IBV sequences are unknown. For the construction of the trees of this figure, only the M41H and M42U sequences were used. Replacing them by M41U and/or M42S, respectively, did not significantly alter the topology of the trees (data not shown)

D1466 is about the same on both sides of this putative crossover site: 73 and 67%, respectively.

The above comparisons are both time consuming and subjective. A method was sought to test and quantify these results more rigorously. This was achieved by comparing phylogenetic trees for each of five sections (windows) of the S2 gene and its flanking sequences (see *Figure 2a*). The S2 gene was divided into three sections.

The first section (II in *Figure 2b*) contained the 5'-part of the S2 gene up to the putative KB8523 crossover site (in the codon for Ile-929 in *Figure 1*). A second S2 tree (III in *Figure 2b*) was constructed for the part of the S2 gene between the KB8523 and 6/82 crossover sites (from Ile-929 to Lys-1105). The 3'-border of this region did not exactly coincide with the second KB8523 crossover site (in the Ser-1102 instead of the Lys-1105 codon) but changing this border had no significant effect on the topology. The third S2 tree (IV in *Figure 2b*) was calculated for the region from the 6/82 crossover site (in the Lys-1105 codon) until the stop codon on position 1182.

Sequences upstream and downstream from the S2 gene were also used to construct phylogenetic trees. The tree based on S1 gene sequences (I in *Figure 2b*) has the same topology as an S2 tree based on region II in S2. KB8523 has shifted to the D1466 branch in a tree for region III of S2. Region IV of S2 yields a tree in which 6/82 has shifted from the D207 to the D1466 branch, while KB8523 is again near M41 and M42. Finally a tree (V) constructed from the E1 gene sequences shows that, compared to tree IV, also strain D207 has shifted towards the D1466 branch, and therefore may also be a recombinant.

Conclusions

With the murine coronavirus MHV a high frequency of recombination was found both *in vitro* and in mouse brain after infection with a *ts* mutant of MHV strain A59 and wild type JHM-virus[18,19].

Phylogenetic trees constructed from five different windows of the genomes of eight IBV strains were used to detect crossover events. A clear advantage of this method is that, due to its simplicity, large numbers of homologous genes can easily be screened. The data presented in this paper suggest that genomic recombination not only occurs under laboratory conditions but also plays an important role in the generation of new virulent strains; at least three IBV strains, all isolated from outbreaks, are potential recombinants. Previous data[11] already indicated that the Dutch field isolate V1397 is also the product of a recombination. Considering the limited number of sequences analysed so far, *in vivo* recombination of IBV seems to occur at a rather high frequency. This has implications for the use of IBV vaccines consisting of several serotypes of live attenuated IBV strains. These vaccines offer the best protection, but they may prove dangerous in the long term by the induction of new variants, not only by point mutations but also through recombination. In this respect vaccination protocols that combine IBV with other live attenuated viruses are equally suspect. They may create ideal conditions for non-homologous recombination which may result in the emergence of recombinant virus species with unwanted properties.

Acknowledgements

The authors are grateful to Ms G.A.W.M. Kremers and Ms K.A. Zwaagstra for experimental contributions and to Dr J.A. Lenstra for advice and help with computer programs. The use of the CAOS/CAMM computer facilities, under grant numbers SON-11-20-700 and STW-NCH44.0703, is gratefully acknowledged.

References

1 Wege, H., Siddell, S. and ter Meulen, V. The biology and pathogenesis of coronaviruses. *Curr. Top. Microbiol. Immunol.* 1982, **99**, 165–200

2 Cook, J.K.A. The classification of new serotypes of infectious bronchitis virus isolated from poultry flocks in Britain between 1981 and 1983. *Avian Pathol.* 1984, **13**, 733–741

3 Kusters, J.G., Niesters, H.G.M., Bleumink-Pluym, N.M.C., Davelaar, F.G., Horzinek, M.C. and van der Zeijst, B.A.M. Molecular epidemiology of infectious bronchitis virus in the Netherlands. *J. Gen. Virol.* 1987, **68**, 343–352

4 Cavanagh, D., Darbyshire, J.H., Davis, P. and Peters, R.W. Induction of humoral neutralising and haemagglutination-inhibiting antibody by the spike protein of avian infectious bronchitis virus. *Avian Pathol.* 1984, **13**, 573–583

5 Cavanagh, D., Davis, P.J., Darbyshire, J.H. and Peters, R.W. Coronavirus IBV; partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor of IBV strains Beaudette and M41. *J. Gen. Virol.* 1986, **67**, 1435–1442

6 Koch, G., Hartog, L., Kant, A., van Roozelaar, D. and de Boer, G.F. Antigenic differentiation of avian bronchitis virus variant strains employing monoclonal antibodies. *Isr. J. Vet. Med.* 1986, **41**, 89–97

7 Mockett, A.P.A., Cavanagh, D. and Brown, T.D.K. Monoclonal antibodies to the S1 spike and membrane proteins of avian infectious bronchitis virus strain Massachusetts M41. *J. Gen. Virol.* 1984, **65**, 2281–2286

8 Cavanagh, D., Davis, P.J., Pappin, D.J.C., Binns, M.M., Boursnell, M.E.G. and Brown, T.D.K. Coronavirus IBV: virus retaining spike glycopolypeptide S2 but not S1 is unable to induce virus-neutralizing or haemagglutination-inhibiting antibody, or induce chicken tracheal protection. *Virus Res.* 1986, **4**, 133–143

9 De Groot, R.J., Luytjes, W., Horzinek, M.C., van der Zeijst, B.A.M., Spaan, W.J.M. and Lenstra, J.A. Evidence for a coiled-coil structure in the spike protein of coronaviruses. *J. Mol. Biol.* 1987, **196**, 963–966

10 Stern, D.F. and Sefton, B.M. Coronavirus proteins: biogenesis of avian bronchitis virus virion proteins. *J. Virol.* 1982, **44**, 794–803

11 Kusters, J.G., Niesters, H.G.M., Lenstra, J.A., Horzinek, M.C. and van der Zeijst, B.A.M. Phylogeny of antigenic variants of avian coronavirus IBV. *Virology* 1989, **169**, 217–221

12 Davelaar, F.G., Kouwenhoven, B. and Burger, A.G. Occurrence and significance of infectious bronchitis virus variant strains in egg and broiler production in the Netherlands. *Vet. Q.* 1984, **6**, 114–120

13 Binns, M.M., Boursnell, M.E.G., Tomley, F.M. and Brown, T.D.K. Comparison of the spike precursor sequences of coronavirus IBV strains M41 and 6/82 with that of IBV Beaudette. *J. Gen. Virol.* 1986, **67**, 2825–2831

14 Sutou, S., Sato, S., Okabe, T., Nakai, M. and Sakaki, N. Cloning and sequencing of genes encoding structural proteins of avian infectious bronchitis virus. *Virology* 1988, **165**, 589–595

15 Niesters, H.G.M., Lenstra, J.A., Spaan, W.J.M., Zijderveld, A.J., Bleumink-Pluym, N.M.C., Hong, F., van Scharrenburg, G.J.M. *et al.* The peplomer protein sequence of the M41 strain of coronavirus IBV and its comparison with Beaudette strains. *Virus Res.* 1986, **5**, 253–263

16 Queen, C. and Korn, L.L. A comprehensive sequence analysis program for the IBM personal computer. *Nucl. Acids Res.* 1984, **12**, 581–599

17 Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985, **39**, 783–791

18 Makino, S., Keck, J.G., Stohlman, S.A. and Lai, M.M.C. High-frequency RNA recombination of murine coronaviruses. *J. Virol.* 1986, **57**, 729–737

19 Keck, J.G., Matsushima, G.K., Makino, S., Fleming, J.O., Vannier, D.M., Stohlman, S.A. and Lai, M.M.C. In vivo RNA–RNA recombination of coronavirus in mouse brain. *J. Virol.* 1988, **62**, 1810–1813

20 Binns, M.M., Boursnell, M.E.G., Cavanagh, D., Pappin, D.J.C. and Brown, T.D.K. Cloning and sequencing of the gene encoding the spike protein of coronavirus IBV. *J. Gen. Virol.* 1985, **66**, 719–726

21 Cavanagh, D. and Davis, P.J. Evolution of avian coronavirus IBV: sequence of the matrix glycoprotein gene and intergenic region of several serotypes. *J. Gen. Virol.* 1988, **69**, 621–629