Virus Genes 17:1, 63–83, 1998

Sequence Comparison of Avian Infectious Bronchitis Virus S1 Glycoproteins of the Florida Serotype and Five Variant Isolates from Georgia and California

KRISTI M. MOORE,[2] JOYCE D. BENNETT,[1] BRUCE S. SEAL,[1] & MARK W. JACKWOOD

Department of Avian Medicine, Poultry Diagnostic and Research Center, College of Veterinary Medicine, University of Georgia, Athens, GA 30602 mjackwoo@arches.ugq.edu [1]Southeast Poultry Research Laboratory, Athens, Georgia [2]Present address: Department of Animal and Food Science, University of Delaware, Newark, DE, 19717, USA

Received February 11, 1997; Accepted January 3, 1998

Abstract. The infectious bronchitis virus (IBV) spike glycoprotein S1 subunit is required to initiate infection and contains virus-neutralizing and serotype-specific epitope(s). Reported are the S1 gene nucleotide and predicted amino acid sequences for the Florida 18288 strain and isolates GA-92, CV-56b, CV-9437, CV-1686, and 1013. These sequences were compared with previously published gene sequences of IBV strains, and phylogenetic relationships are reported. The S1 amino acid sequence of Florida 18288 was 94.9% similar to the Connecticut strain, and GA-92 was 92.8% similar to the Arkansas 99 strain. S1 amino acid sequences of the California variants, CV-56b, CV-9437, and CV-1686, were 97.6–99.3% similar to one another and only 76.6%–76.8% similar to the Arkansas-type strains. Isolate 1013, also from California, was 84.0% similar to Ark DPI and 77.9% similar to CV-56b. When comparing 19 viruses isolated from the United States, sequence variations were observed between amino acids 55–96, 115–149, 255–309, and 378–395. Similar regions are reported to be involved in virus-neutralizing and/or serotype-specific epitopes. These data demonstrate that variant IBV strains continue to emerge, and unique variants may circulate among poultry in geographically isolated areas.

Key words: infectious bronchitis virus, spike glycoprotein, coronavirus, phylogenetic analysis

Introduction

Avian infectious bronchitis virus (IBV) causes an economically important upper respiratory tract disease in chickens that results in reduced weight gains and reduction of egg production. Infectious bronchitis is complicated by secondary bacterial infections such as *Mycoplasma* sp. and *Escherichia coli* that may increase mortality (1). Vaccination for infectious bronchitis is hindered because monovalent IBV vaccines do not provide complete protection against heterologous serotypes (2,3).

Infectious bronchitis virus, a member of the Coronaviridae family, contains four structural proteins. These are the small membrane protein, the integral membrane glycoprotein, the phosphorylated nucleocapsid protein, and the spike (S) glycoprotein (4,5). The S glycoprotein is cleaved posttranslationally by cellular proteases into the S1 and S2 subunits (6). The globular S1 subunit forms the tip of the spike, extending outwardly, whereas S2 anchors S1 to the viral membrane (7).

The S1 subunit is involved with infectivity and contains virus-neutralizing epitopes, serotype-specific sequences, and hemagglutinin activity (7–11). Different serotypes and subtypes of IBV are thought to be generated by nucleotide insertions, deletions, or point mutations in the S1 subunit made by the viral polymerase (12). Another mechanism for variability in the S1 subunit may be RNA recombination (13–16). Consequently, changes in the S1 subunit

The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers 1013 = AF027508, CV-9437 = AF027510, CV-1686 = AF027511, CV56-b = AF027509, FL18288 = AF027512.

can lead to emergence and proliferation of variant serotypes.

Sequence variations in S1 subunits have been identified among European IBV strains and within the Massachusetts (Mass) serotype. These are designated hypervariable region (HVR) I (residues 56–69) and HVR II (residues 117–131) (17,18). Isolates from chickens in the United States (US), other than the Mass serotype, have similar highly variable regions in the S1 subunit between residues 53–148 (15). Other variable regions have been detected between residues 250–310 for the strains H120, D207, D1466, V1397, Mass 41, Mass 42, and residues 269–365 (including the signal sequence) in the strains 6/82, 123/82, and 167/82 (12,19). These variable regions (HVR I, HVR II, and residues 250–365) are similar to sequence regions designated I, II, and III associated with three virus-neutralizing epitopes determined by selection of monoclonal antibody (Mab) neutralization-resistant (NR) mutants (20). HVR I was associated with a virus-neutralizing epitope and a hemagglutinin, whereas amino acids 250–386 are associated with a serotype-specific epitope (17,21).

The objectives of this investigation were to sequence the S1 gene of the Florida 18288 strain and variant IBV isolates GA-92, CV-56b, CV-9437, CV-1686, and 1013. Sequence comparison and phylogenetic analysis were used to determine the relationship of these viruses with other USA isolates of IBV and to further identify variable regions that may be virus-neutralizing epitopes on the S1 glycoprotein subunit.

Materials and Methods

Viruses and Histories

The S1 nucleotide and predicted amino acids sequences were determined from these viruses isolated in the United States; Florida 18288 (FL-18288), Georgia isolate 1992 (GA-92), and isolates from California CV-56b, CV-9437, CV-1686, and 1013.

FL-18288 strain, representing the Florida serotype, was isolated from layers in 1972, and was obtained from Dr. Pedro Villegas at the University of Georgia (22). GA-92 (Arkansas serotype) was isolated in 1992 at the Poultry Diagnostic and Research Center (PDRC) from a broiler flock in Georgia exhibiting respiratory distress (23). All isolates from California were from chickens showing respiratory distress. CV-56b (isolated in 1991 from broilers), CV-1686 (isolated in 1995 from broilers), and 1013 (isolated in 1995 from silkies) were obtained from the California Veterinary Diagnostic Laboratory System (Fresno, CA), whereas CV-9437 (isolated in 1995 from broilers) was submitted to PDRC. Isolation of IBV strains from California started in 1984 and are referred to as California variants. California variants in this study are represented by CV-56b. Data from serum neutralization studies suggest that these variant strains may compose a new serotype (23,24).

RNA Isolation and RT-PCR

All procedures for RNA isolation and reverse transcriptase-polymerase chain reaction (RT-PCR) were previously described (25). Briefly, RNA was isolated from IBV infective allantoic fluids after treatment with Proteinase K (250 μg/ml) and SDS (final 2% solution wt/wt), with incubation at 55°C for 5 min. Phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1) extractions were completed, and RNA was isolated from infective allantoic fluids using the RNAid Kit (Bio 101). Primers MIBVPCR and NIBVPCR were used to amplify a conserved region between the membrane and nucleocapsid genes for identification of IBV (26,27). The S1oligo5′ (5′-TGAAACTGAACAA-AAGAC-3′) and S1oligo3′ (5′-CATAACTAACAT-AAGGGCAA-3′) primers were used to amplify S1 gene as previously described (25,28). The S1 gene amplification product was used for subsequent restriction fragment length polymorphism (RFLP) or DNA sequencing.

Restriction Fragment Length Polymorphism Analysis

Amplification products were excised from an agarose gel and purified using Geneclean (Bio 101). The restriction endonucleases, *Bst* YI, *Hae* III, and *Xcm* I (New England Biolabs) were used to digest the S1 gene amplification product as previously described (25). RFLPS were determined following agarose gel electrophoresis (25).

DNA Sequencing and Sequence Analysis

The S1 gene polymerase chain reaction (PCR) products for FL-18288, GA-92, and CV-9437 were cloned into the TA cloning vector (Invitrogen Corp.), and ligated areas of the plasmids were sequenced with the M13 universal forward and reverse primers (Molecular Genetics Instrumentation Facility, University of Georgia). After initial sequence data was obtained from plasmid DNA, the S1 gene for each sample was amplified by RT-PCR in six or more tubes then combined and used as template for subsequent sequencing. Sequencing primers to various regions within the S1 gene for FL-18288, GA-92, CV-56b, CV-9437, CV-1686, and 1013 were designed using OLIGO™ version 4.0 software (National Biosciences, Inc.) and are available upon request.

Double-stranded DNA sequencing was conducted using automated sequencing with the Prism™ DyeDeoxy terminator cycle sequencing kit as recommended by the manufacturer (Perkin Elmer; 29,30). All samples were purified for DNA sequencing using Microcon™ 30 columns (Amicon, Inc.). Automated sequencing was conducted at the USDA Southeastern Poultry Research Laboratory (Athens, GA).

The nucleotide sequence of the S1 gene from the ATG start site to the cleavage site for the FL-18288, GA-92, CV-56b, CV-9437, CV-1686, and 1013 were compared with published sequences from Beaudette (accession # X02342), Mass 41 (accession # X04722), Connecticut (Conn) (accession # L18990), Arkansas (Ark) 99 (accession # M85244), Ark DPI (accession # AF006624), CU-T2 (accession # U04739), Pp14 (accession # M99483), Se17 (accession # M99484), Iowa 609 (Dr. Collisson, Texas A & M U., College Station, TX), Holte (accession # L18988), JMK (accession # L14070), Gray (accession # L14069), and De-072 (accession # U90751) (15,16,34–36). Alignments and phylogenetic analysis were conducted using CLUSTAL V method in MegAlign software version 1.03 (DNAStar Inc., Madison, WI). This method uses a multiple alignment algorithm, and unweighted pair group method with arithmetic mean algorithm (UPGMA) to derive a preliminary phylogeny (31,32). The final phylogeny is produced by applying the neighbor-joining method (33). The CLUSTAL V method was compared with UPGMA with MacDNASIS ProV3.0 software (Hitachi Software Engineering Co., Ltd) and IntelliGenetics GeneWorks 2.45™ (Intelligenetics, Inc. Mountain View, CA).

Results

IBV Isolate Identification and RFLP Analysis of Amplification Products

Reverse-transcriptase PCR in conjunction with RFLP were conducted for initial characterization of IBV strain FL-18288 and isolates, GA-92, CV-56b, CV-9437, CV-1686, and 1013 (25). All samples were identified as IBV by the amplification of a 1.02 kb PCR product with the MIBVPCR and NIBVPCR primers (data not shown). The S1 gene PCR product (1.7 kb) was subsequently amplified with the S1oligo5′ and S1oligo3′ primers for all isolates (data not shown).

Using the restriction enzymes *Bst* YI, *Hae* III, and *Xcm* I, the RFLP pattern for amplification products from FL-18288 corresponded to a pattern reported by Kwon *et al*. (25). The RFLP patterns for CV-9437 and CV-1686 amplification products were identical to one another and to patterns for other California variants represented by CV-56b. However, the RFLP pattern for GA-92 and 1013 differed from one another and from other IBV strains using the restriction enzyme, *Bst* YI (Fig. 1).

S1 Nucleotide Sequence Comparison among IBV Isolates

The S1 nucleotide sequence alignments and pair distances data are presented in Fig. 2 and Table 1 respectively. The S1 gene nucleotide sequence of FL-18288 was most similar to the Conn strain (97.3%). Isolates CV-9437, CV-1686, and CV-56b had S1 genes most similar (98.7%–99.7%) to one another, and 78.5% to 79.0% similar to Ark 99. The IBV isolate 1013 was most similar to Ark 99 (84.7%) and Ark DPI (85.1%), whereas GA-92 was most similar to Ark 99 (95.3%). All methods used to create phylogenetic trees produced similar dendrograms when comparing S1 genes and had similar topology to the trees based on S1 amino acid sequences as illustrated in Fig. 3.

S1 Amino Acid Sequence Comparison

A dendrogram was generated following alignment of the S1 glycoprotein sequences to determine phylogenetic relationships among IBV strains in the United States (Fig. 3). Similar topologies were obtained when

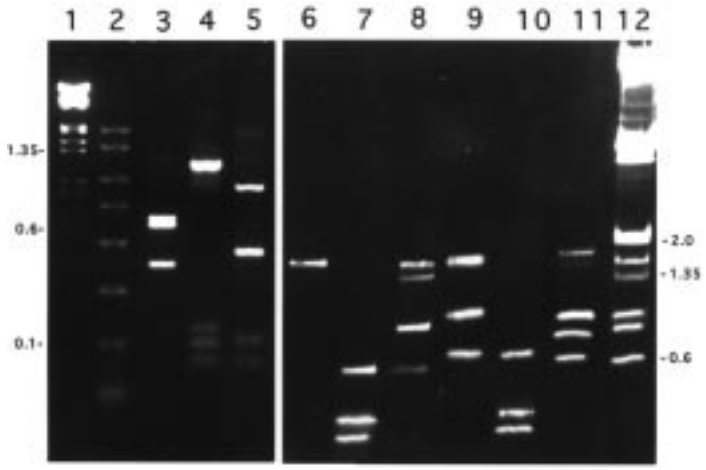


Fig. 1. Variation of infectious bronchitis virus S1 genes demonstrated by RFLP of RT-PCR amplification products using viral genomic RNA as a template. Lane 3 = GA-92 cut with *Bst* YI; Lane 4 = GA-92 cut with *Hae* III; Lane 5 = GA-92 cut with *Xcm* I; Lane 6 = CV-9437 cut with *Bst* YI; Lane 7 = CV-9437 cut with *Hae* III; Lane 8 = CV-9437 cut with *Xcm* I; Lane 9 = 1013 cut with *Bst* YI; Lane 10 = 1013 cut with *Hae* III; Lane 11 = 1013 cut with *Xcm* I. Numbers at the left and right are molecular-weight markers in kilobase pairs. Lanes 1 and 12 = molecular-weight marker, lambda cut with *Eco* RI and *Hind* III giving the fragments of the following sizes: 25.7 kb, 4.3 kb, 2.0 kb, 1.65 kb, 1.35 kb, 0.95 kb, 0.90 kb, and 0.60 kb; Lane 2 = molecular-weight marker (BioVenture, Inc., Murfreesboro, Tenn.).

generating phylogenetic trees using UPGMA, PAUP, and the neighbor-joining method. The S1 amino acid sequence alignments and pair distances data are presented in Fig. 4 and Table 2 respectively. The FL-18288 strain S1 protein sequence was 94.9% similar to that of the Conn strain. California variants, CV-56b, CV-9437, and CV-1686, were 97.6% to 99.3% similar to one another, and 76.6% to 76.8%

Table 1. S1 nucleic acid sequence alignment pair distances

Per cent Similarity

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**		
1		69.9	71.3	71.8	70.0	71.8	93.1	94.8	97.3	70.5	76.0	71.1	75.3	77.6	77.0	75.6	78.6	78.9	49.2	**1**	FL-18288
2	24.0		76.0	76.3	81.8	76.1	74.5	74.4	72.7	95.3	93.2	93.6	92.5	81.9	76.9	76.3	78.3	78.1	45.9	**2**	GA-92
3	23.1	23.3		98.7	81.1	98.8	76.1	75.8	73.9	78.5	78.6	78.1	77.2	76.1	77.1	74.5	75.4	75.5	46.4	**3**	CV-56b
4	22.7	23.0	1.2		81.3	99.7	76.3	75.9	74.3	79.0	79.0	78.5	77.7	76.5	77.2	74.8	75.5	75.9	46.7	**4**	CV-1686
5	24.3	16.9	18.0	17.8		81.1	76.6	75.3	72.7	84.7	85.1	84.2	82.1	80.4	78.3	77.2	78.2	78.3	46.0	**5**	1013
6	22.8	23.2	1.1	0.3	18.1		76.1	75.7	74.1	78.8	78.8	78.4	77.5	76.2	77.0	74.7	75.3	75.7	46.3	**6**	CV-9437
7	5.4	23.8	23.0	22.9	23.2	23.1		97.7	94.1	75.0	78.6	76.9	77.5	78.5	78.5	76.4	80.7	79.7	49.2	**7**	Beaudette
8	4.5	23.7	22.9	22.9	23.2	23.0	2.2		94.9	74.7	78.3	76.8	77.7	79.6	78.5	76.2	80.6	79.8	52.1	**8**	Mass 41
9	1.2	24.7	24.2	23.7	25.0	23.9	5.2	4.3		72.9	76.8	75.2	76.7	78.4	76.5	75.4	78.0	78.2	51.9	**9**	Conn
10	23.9	4.5	20.3	20.0	14.2	20.2	23.6	23.3	24.6		96.1	96.6	94.3	84.4	79.0	79.0	79.5	79.6	47.3	**10**	Ark 99
11	21.6	6.0	20.6	20.2	14.1	20.5	21.3	21.1	22.4	3.3		95.4	92.6	85.4	79.5	80.0	82.3	81.9	48.0	**11**	Ark DP1
12	23.3	5.3	20.6	20.6	14.6	20.8	22.9	22.9	23.6	2.7	4.1		93.7	84.3	79.1	79.2	79.7	79.6	48.2	**12**	CU-T2
13	22.2	7.1	21.6	21.2	17.5	21.5	22.1	21.7	22.3	5.3	7.0	6.0		83.3	79.0	77.9	80.2	80.1	47.9	**13**	PP14
14	19.9	18.6	23.9	23.7	19.5	24.0	20.3	19.7	20.1	15.8	15.0	15.9	17.4		80.9	79.7	85.6	85.8	45.8	**14**	SE17
15	21.2	23.9	22.5	22.5	21.9	22.7	21.4	21.1	21.7	21.9	21.5	21.8	22.2	20.6		79.0	80.1	79.8	47.7	**15**	Iowa 609
16	23.6	22.5	24.2	23.7	22.2	23.9	23.5	23.4	24.3	20.2	19.1	20.2	21.8	19.7	20.8		77.6	78.9	47.4	**16**	Holte
17	18.8	21.8	23.8	23.7	21.4	24.0	18.7	18.5	19.6	20.7	17.9	20.6	20.3	14.0	21.1	20.7		98.8	46.2	**17**	JMK
18	19.0	21.9	24.0	23.8	21.6	24.1	18.5	18.4	19.6	20.6	18.4	20.8	20.3	13.9	21.3	20.6	1.1		46.2	**18**	Gray
19	51.2	53.1	53.4	53.3	55.0	35.8	51.0	51.7	51.3	51.4	51.0	51.1	50.8	51.6	54.4	53.3	54.4	55.0		**19**	De-072
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**		

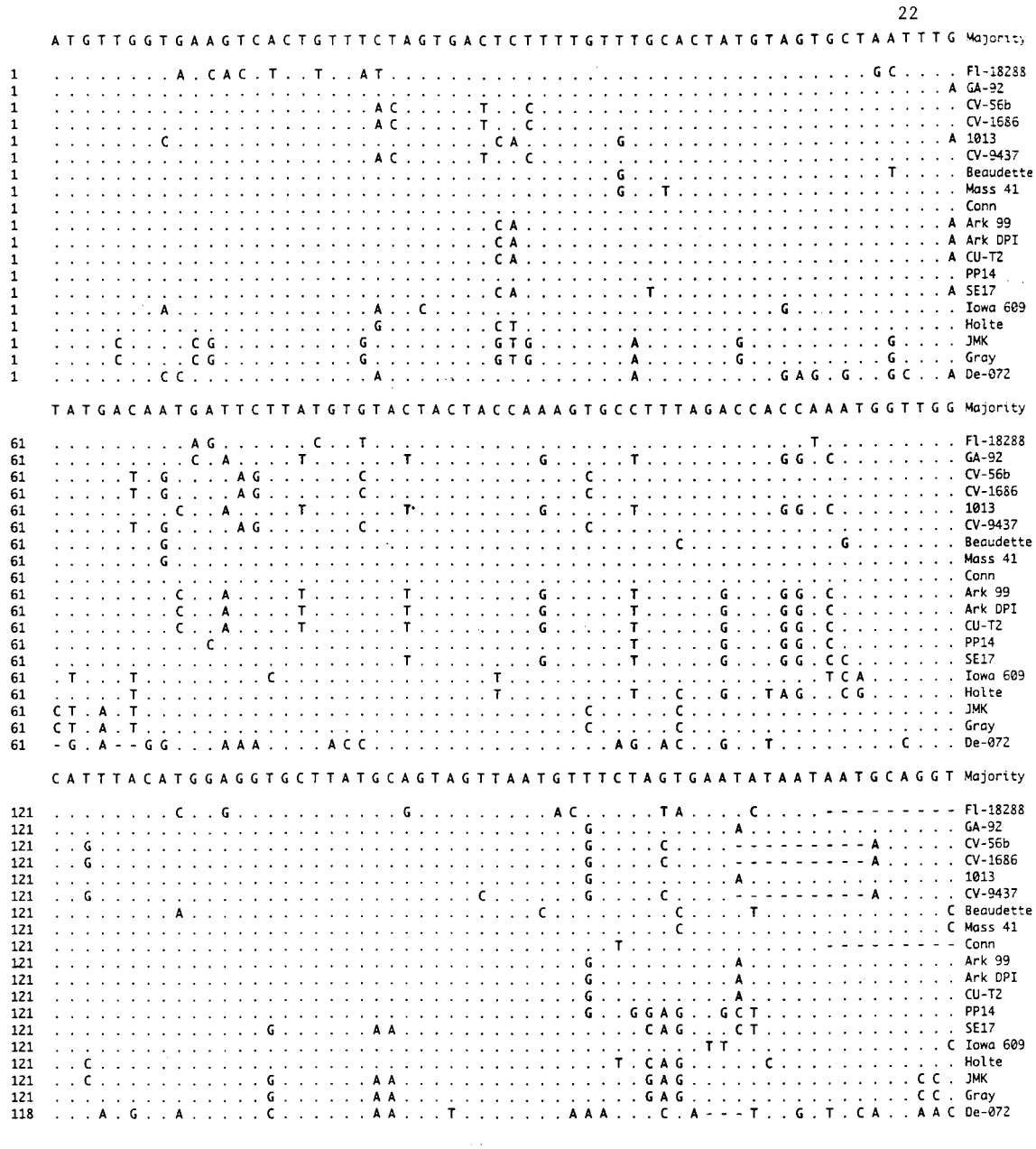


Fig. 2. Nucleotide sequence of the S1 glycoprotein gene region showing differences from the consensus for each of the virus strains listed. Dashes were introduced to align the sequences. ''X'' indicates unavailable sequences. Dots indicate nucleotides that match the consensus exactly.

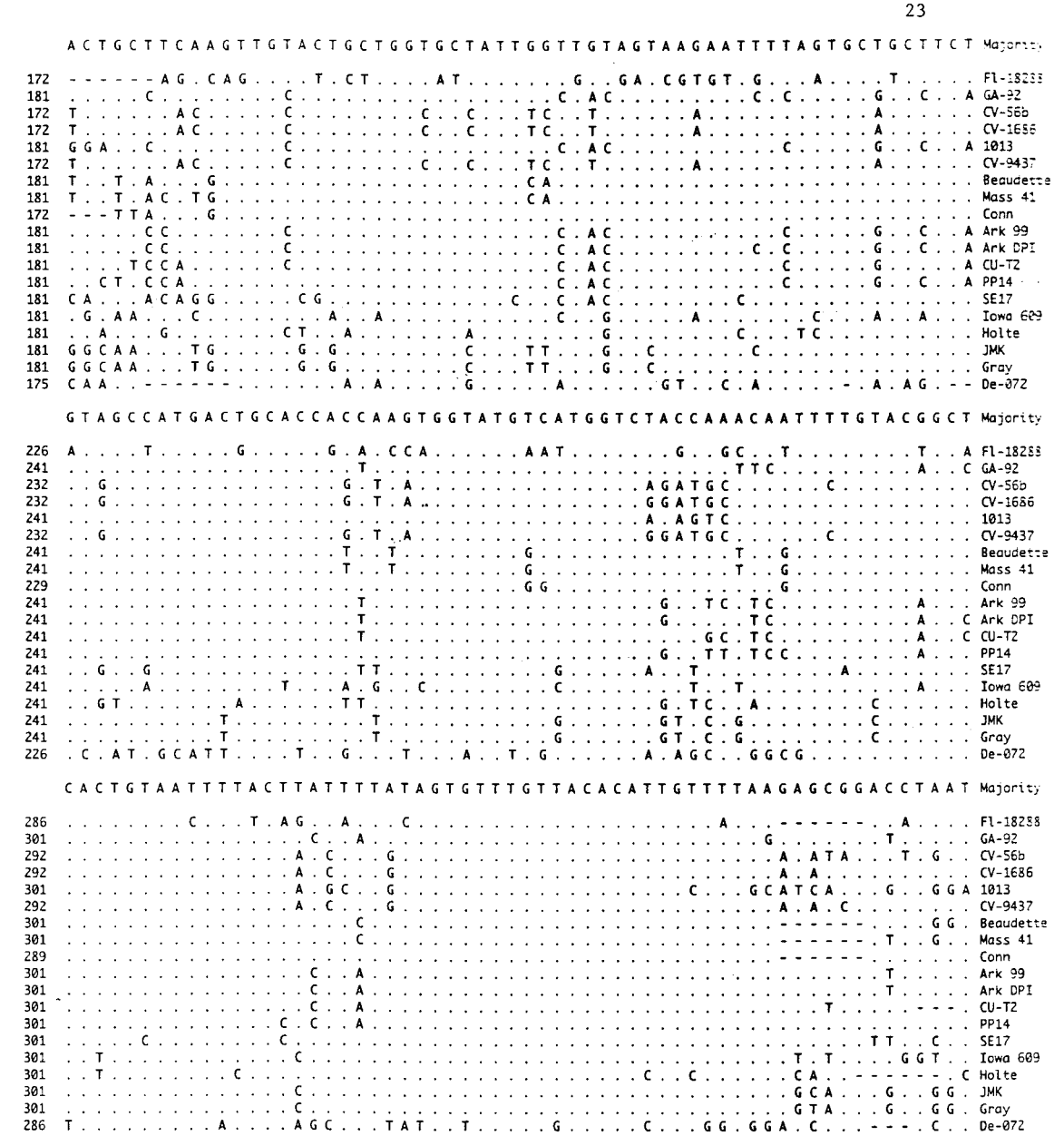


Fig. 2. (Continued)

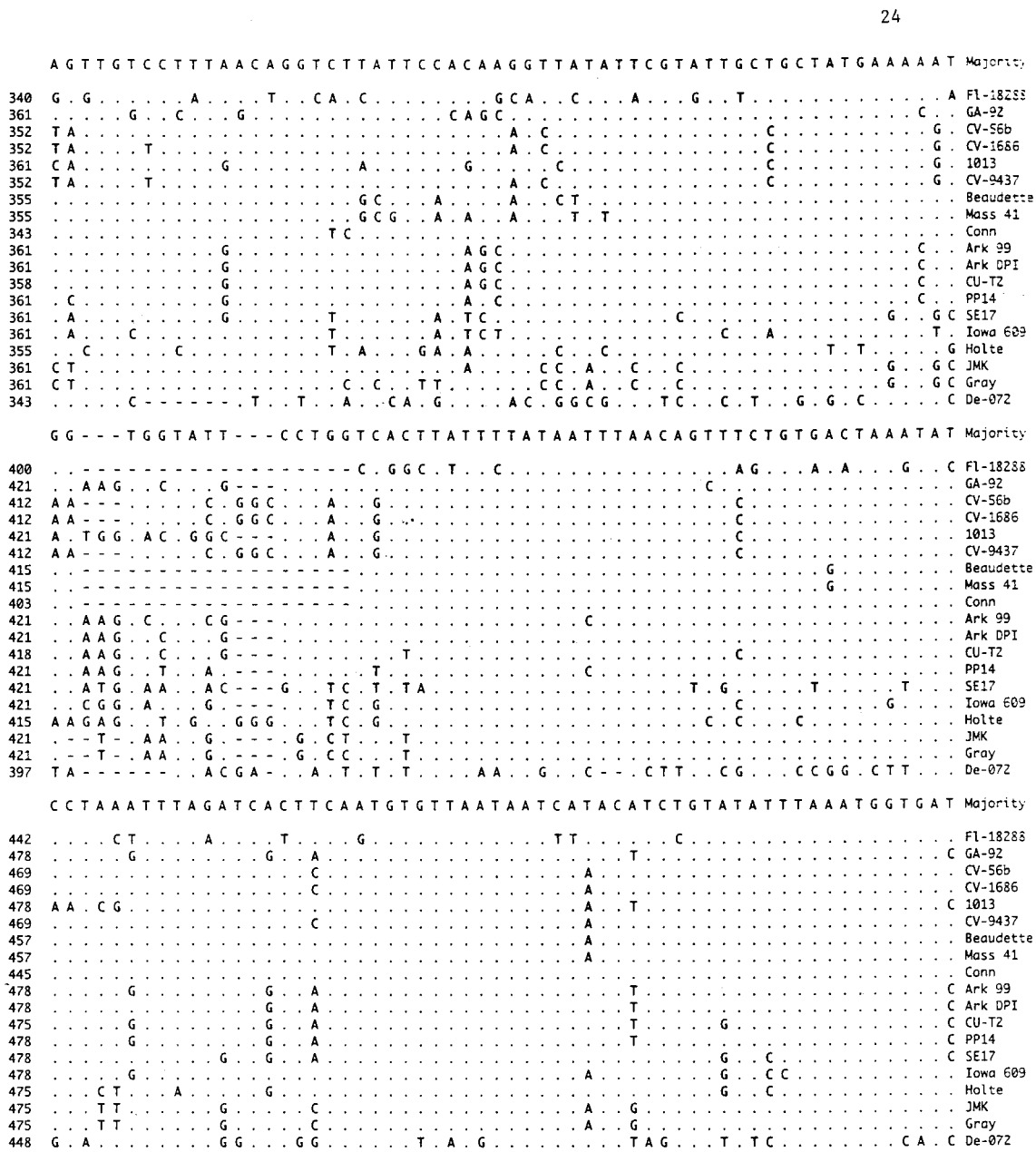


Fig. 2. (Continued)

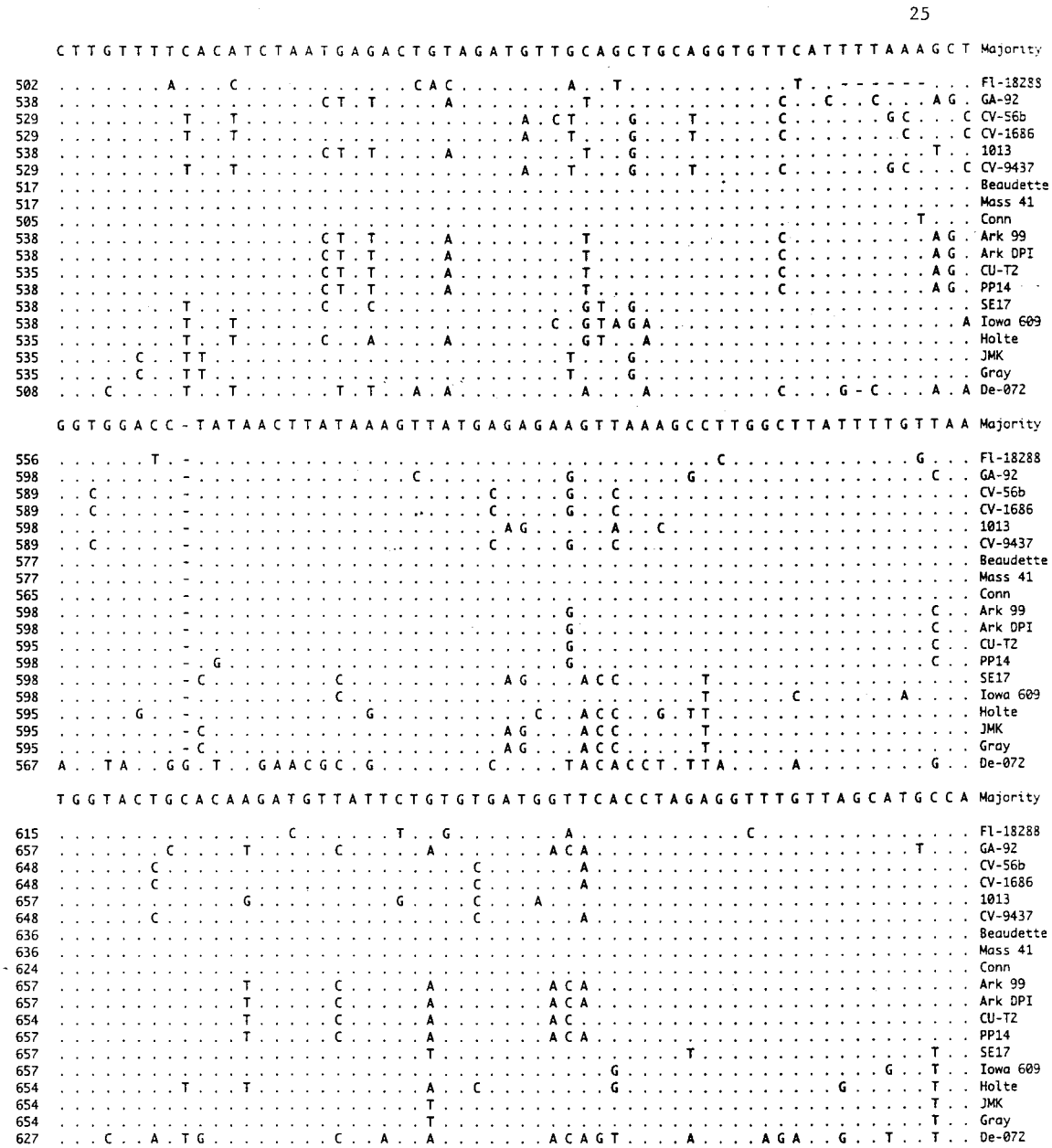


Fig. 2. (Continued)

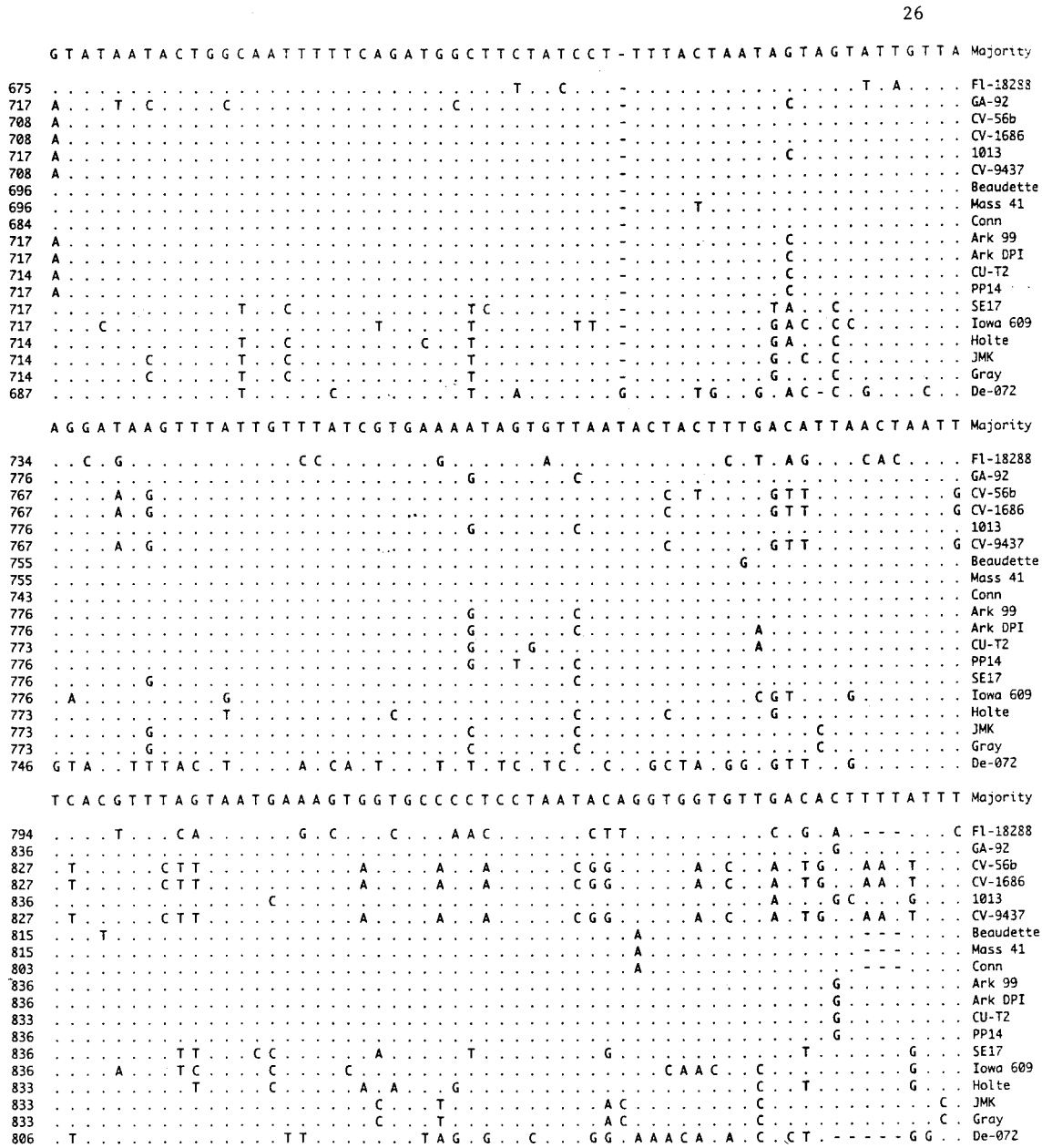


Fig. 2. (Continued)

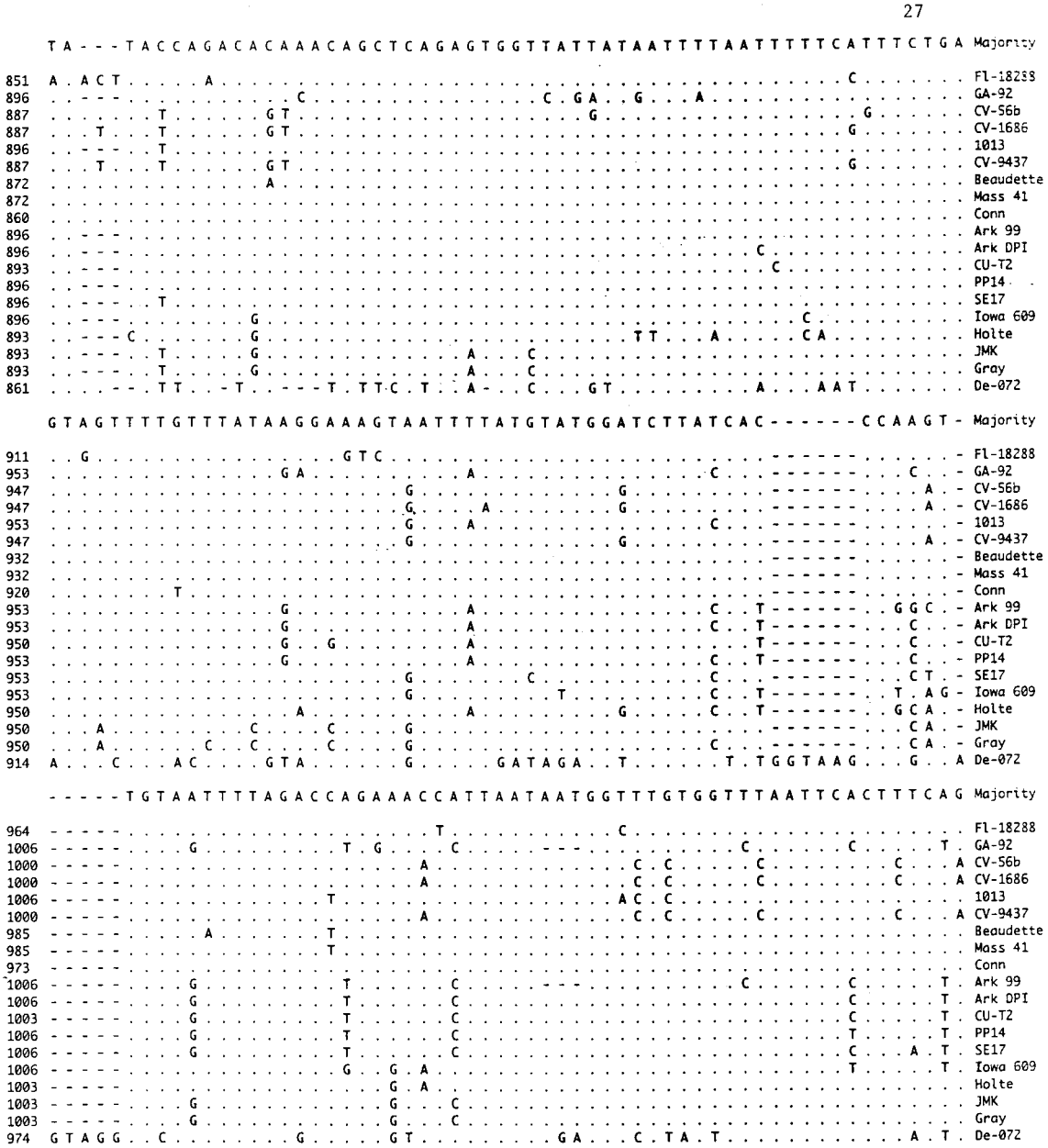


Fig. 2. (Continued)

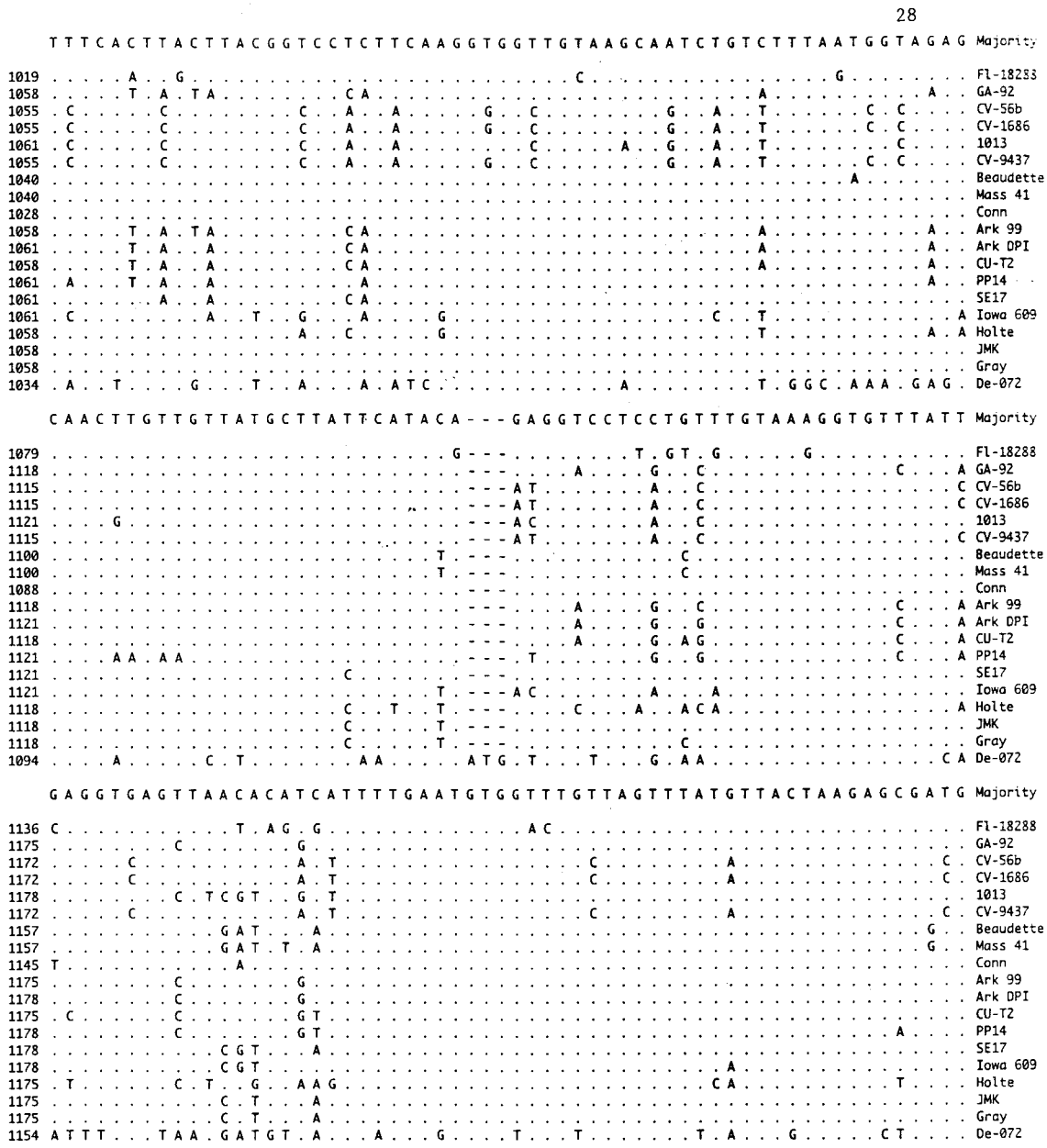


Fig. 2. (Continued)

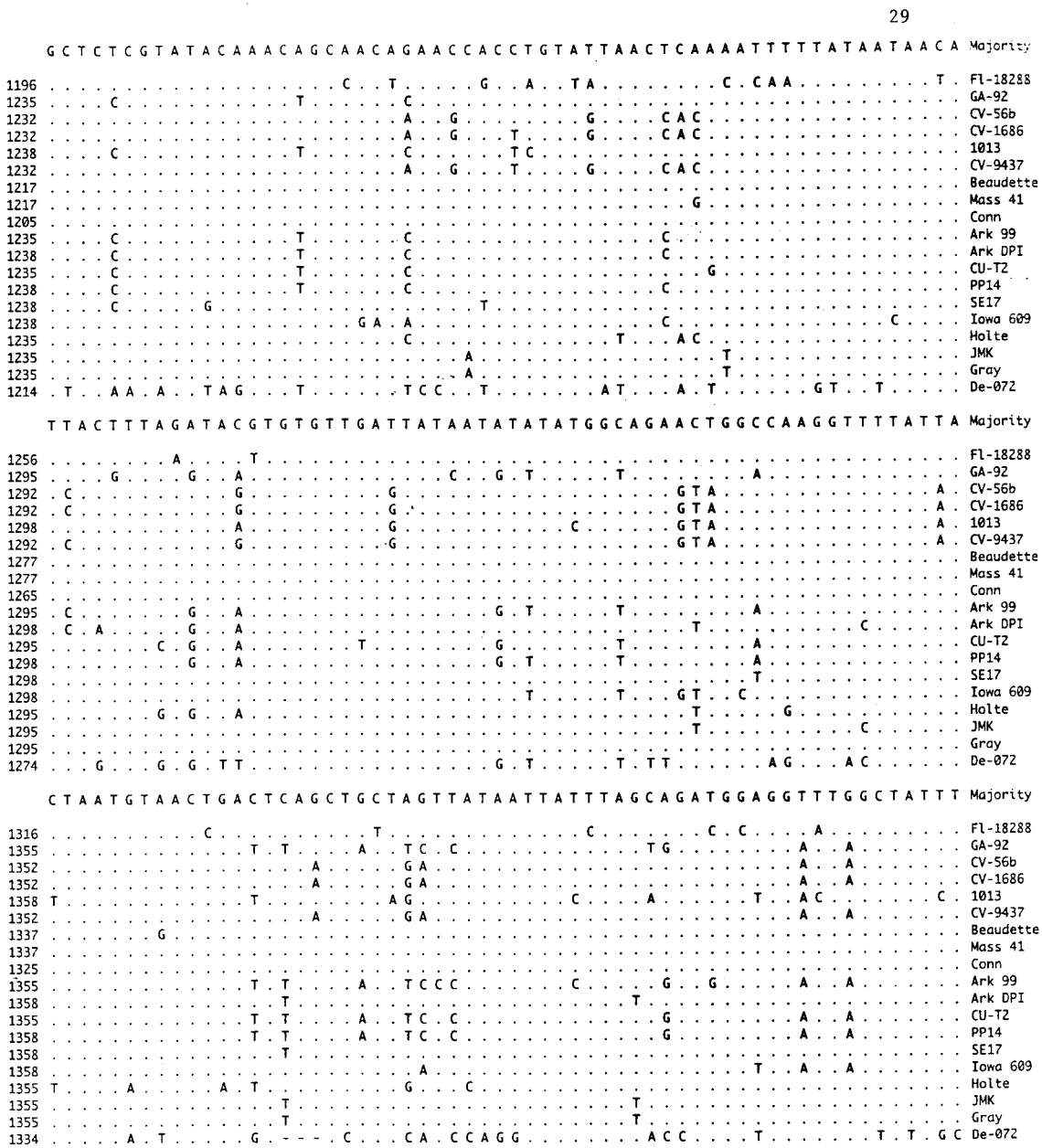


Fig. 2. (Continued)

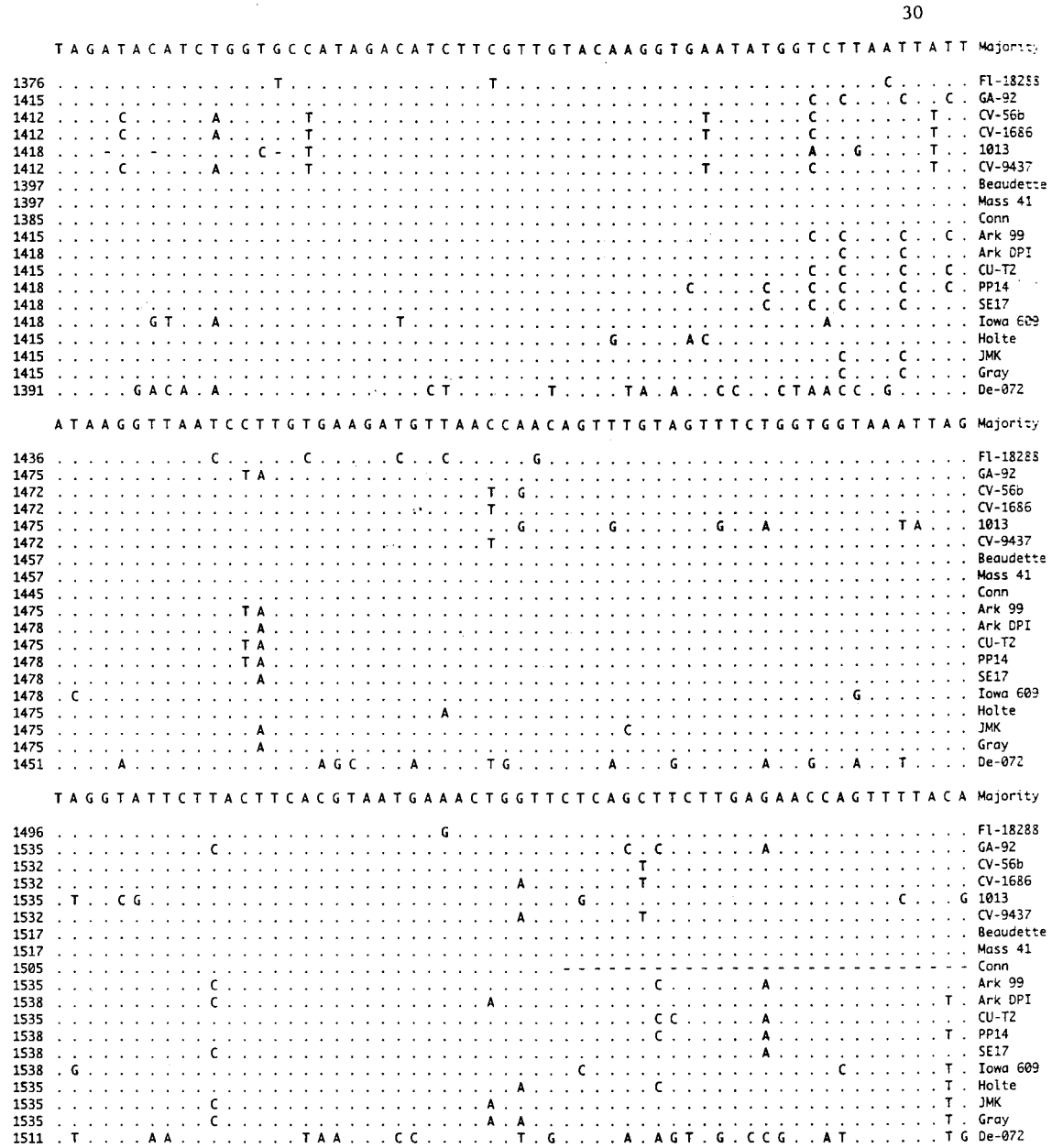


Fig. 2. (Continued)

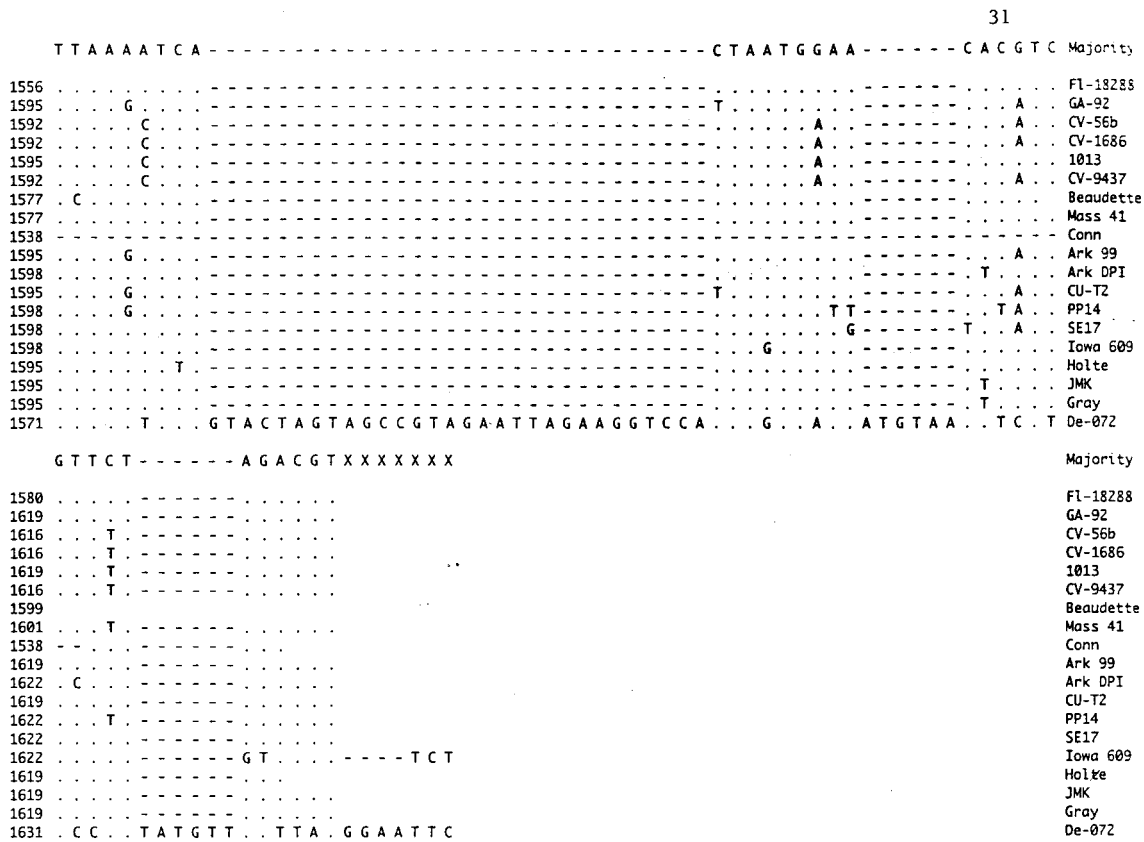


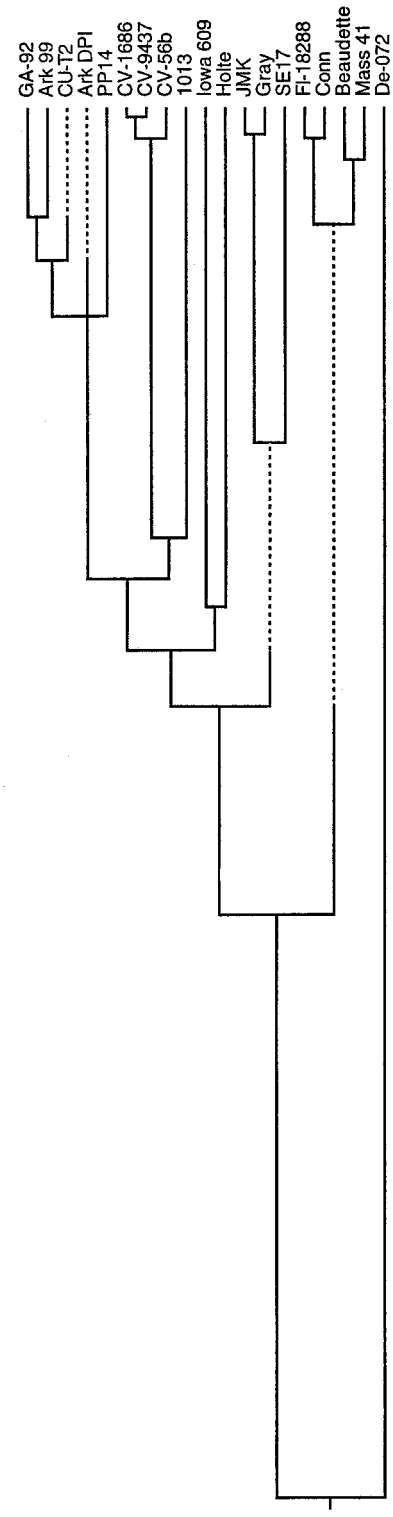
Fig. 2. (Continued)

similar to the S1 protein from Ark strains. Another isolate from California, 1013, had an S1 glycoprotein that was 84.0% similar to Ark DPI and only 77.9% similar to CV-56b. The GA-92 IBV isolates S1 protein was 92.8% similar to the Ark 99 glycoprotein.

The predicted S1 glycoprotein amino acid sequences for IBV isolates FL-18288, GA-92, CV-56b, CV-9437, CV-1686, and 1013 were aligned with previously published S1 glycoprotein sequences of other USA strains in the GenBank database. When comparing the S1 sequences, the most variations were observed between residues 55–96, 115–149, 255–309, and 378–395 (numbering is in reference to Mass 41 S1 and includes the signal sequence). When comparing strains Mass 41, Conn, and FL-18288 representing three serotypes with closely related sequences, variations occurred in the S1 subunit of all three strains at residues 91, 125, and 392 as reported in Table 3.

The S precursor glycoprotein is cleaved post-translationally marking the carboxyl terminus of the S1 subunit and the S2 subunit amino terminus (9). Several IBV isolates reported here had amino acid sequence variations at the cleavage site (residues

Fig. 3. Phylogenetic analysis demonstrating relationships among the S1 glycoprotein amino acid sequences of infectious bronchitis viruses isolated in the United States. The tree contains five major groups. One group contains the Arkansas strains, GA-92 (U16157), Ark 99 (M99482), Ark DPI (AF006624), CU-T2 (U04739), and Pp14 (M99483). Another includes the California variants, CV-1686, CV-9437, CV-56b. Gray (14069), JMK (L14070), Se17 (M99484), and Iowa 609 are grouped together, and Conn (L18990), Florida 18288, Beaudette (M95169), and Mass41 (X04722) form another variety. De-072 (U77298) is substantially different from other IBV strains, whereas Holte (18988) is placed between the California variants and Gray, with isolate 1013 between the Arkansas strains and the California variants.


GA-92
Ark 99
CU-T2
Ark DPI
PP14
CV-1686
CV-9437
CV-56b
1013
Iowa 609
Holte
JMK
Gray
SE17
FI-18288
Conn
Beaudette
Mass 41
De-072

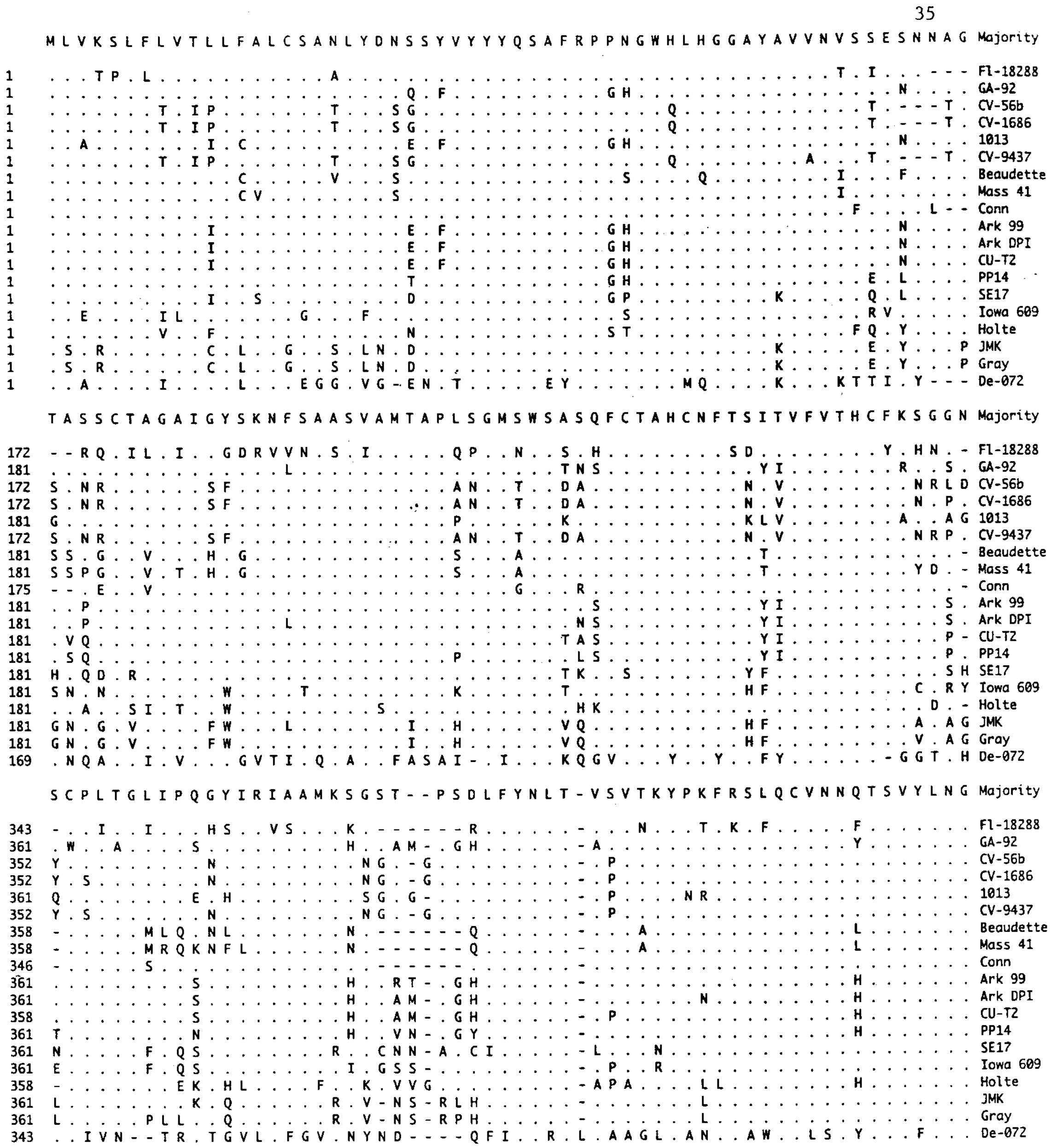


Fig. 4. Amino acid sequence of the S1 glycoprotein region showing differences from the consensus for each of the virus strains listed. Dashes were introduced to align the sequences. Dots indicate amino acids that match the consensus exactly.

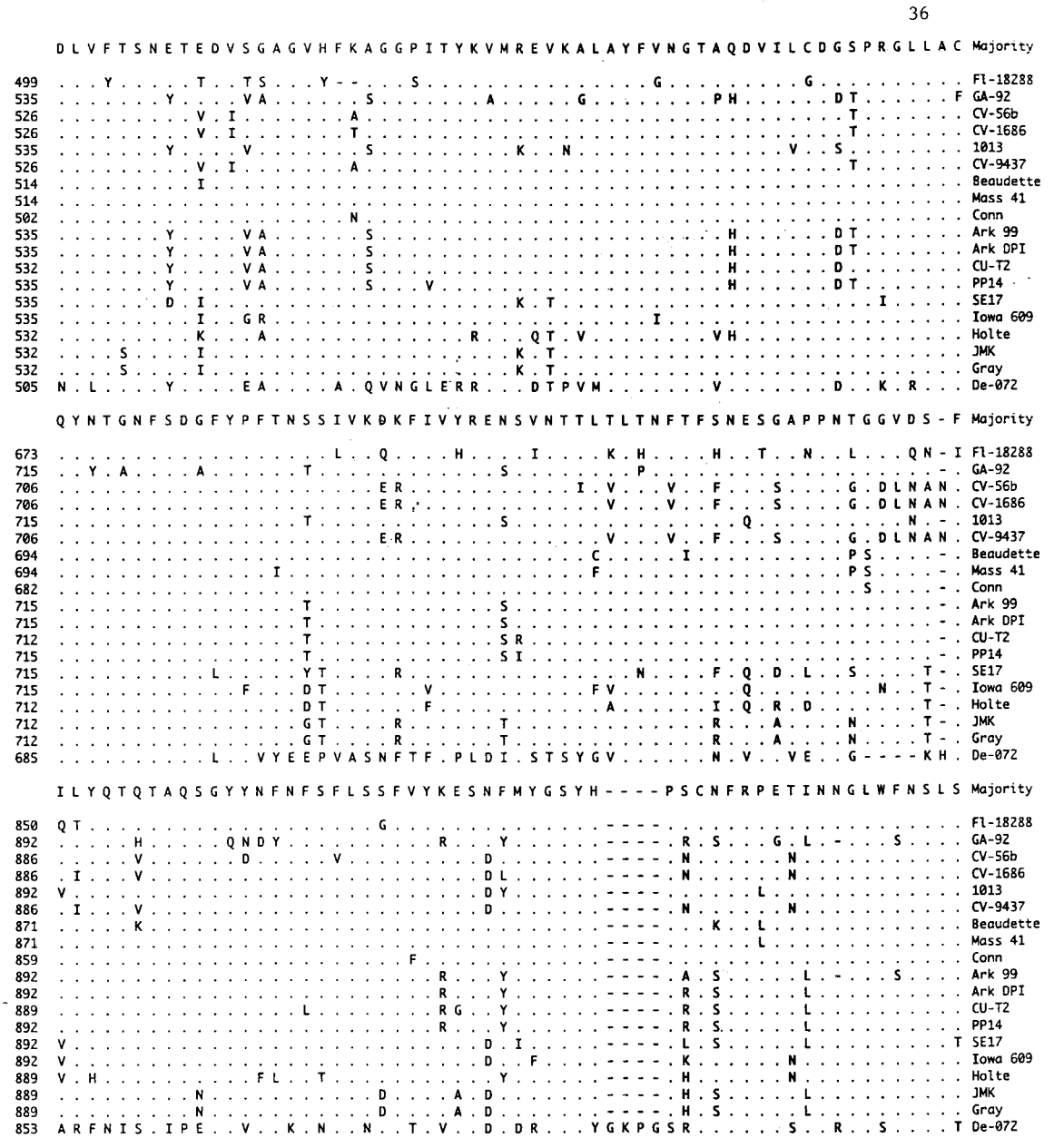


Fig. 4. (Continued)

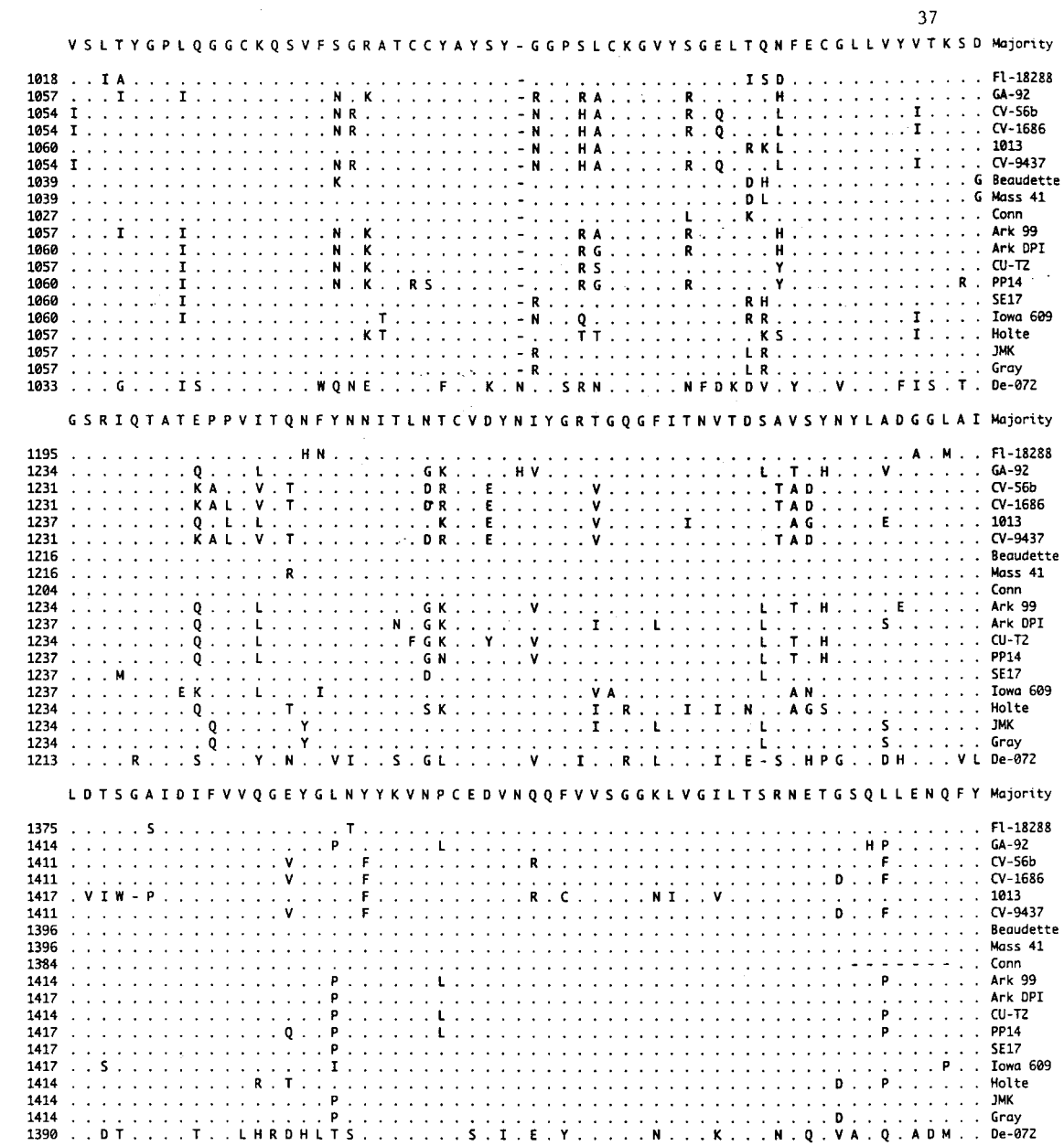


Fig. 4. (Continued)

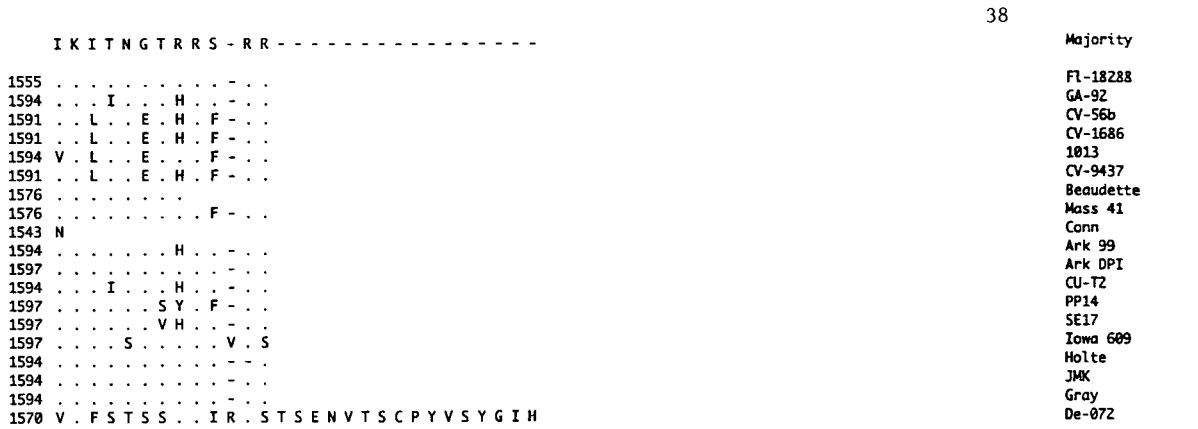


Fig. 4. (Continued)

542–546) conforming to the sequence pattern X_1-Arg-X_2-Arg-Arg. Strains Ark 99, CU-T2, GA-92, SE17 had the sequence His-Arg-Ser-Arg-Arg, whereas Ark DPI, Gray, JMK, Conn, and FL-18288 had the sequence Arg-Arg-Ser-Arg-Arg (34). Strains Beaudette, Mass 41, and 1013 had the sequence Arg-Arg-Phe-Arg-Arg, whereas CV-56b, CV-9437 and CV-1686 had the sequence His-Arg-Phe-Arg-Arg (6).

Discussion

Because of variant RFLP patterns using the RT-PCR/RFLP diagnostic test for IBV, GA-92, CV-56b, CV-9437, CV-1686, 1013 were selected for further characterization. The S1 gene nucleotide and predicted amino acid sequences of FL-18288 strain, GA-92, CV-56b, CV-9437, CV-1686, 1013 were determined, and compared with other IBV sequences.

Table 2. S1 amino acid sequence alignment pair distances

Percent Similarity

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
1		71.7	73.8	73.8	72.5	73.8	87.9	89.8	94.9	74.7	76.4	74.5	75.7	77.4	77.2	74.7	77.4	77.4	46.0	**1**	FL-18288
2	31.2		72.1	72.5	78.1	72.1	73.2	71.5	71.2	92.8	91.0	90.1	88.6	77.3	75.1	73.6	75.7	75.3	41.1	**2**	GA-92
3	28.2	31.6		97.6	77.9	98.0	73.9	73.4	73.9	76.6	76.6	75.8	75.5	74.5	76.2	73.4	74.0	74.0	43.4	**3**	CV-56b
4	28.5	31.1	2.4		78.2	99.3	73.9	73.0	73.9	77.1	77.1	76.4	76.2	74.7	76.4	73.8	74.2	74.5	43.4	**4**	CV-1686
5	29.8	24.2	23.6	23.1		77.9	74.7	74.1	72.2	83.1	84.0	82.9	80.7	78.8	79.6	76.2	77.7	77.5	43.3	**5**	1013
6	28.5	31.6	2.1	0.7	23.6		73.7	73.0	73.9	76.8	76.8	76.0	75.8	74.5	76.2	73.6	74.0	74.4	43.2	**6**	CV-9437
7	11.1	30.4	28.4	28.4	28.2	28.6		94.9	88.7	76.2	77.9	76.2	76.7	78.2	77.7	73.9	78.6	79.2	46.3	**7**	Beaudette
8	9.7	33.0	29.5	29.8	29.1	29.8	5.3		90.1	75.0	76.7	74.9	76.2	77.8	77.1	74.5	78.2	78.2	46.9	**8**	Mass 41
9	3.0	32.2	27.7	28.0	30.4	28.0	11.1	9.6		74.1	75.7	74.1	75.7	77.0	76.7	74.3	76.7	76.7	46.5	**9**	Conn
10	26.8	7.6	25.2	24.4	17.6	24.9	26.1	27.7	27.9		95.4	94.7	92.6	81.8	79.2	78.0	78.8	78.8	44.0	**10**	Ark 99
11	25.4	9.0	26.1	25.4	17.4	25.9	24.7	26.3	26.7	4.2		93.9	92.1	82.7	79.8	78.8	81.8	80.7	45.2	**11**	Ark DP1
12	27.8	9.2	26.2	25.4	17.9	25.9	27.1	29.0	28.7	4.2	5.7		92.3	81.6	79.4	78.4	78.8	78.8	44.6	**12**	CU-T2
13	26.5	11.8	27.7	26.7	21.7	27.2	26.3	27.1	26.7	7.2	8.4	7.6		80.7	78.5	76.9	79.4	79.4	43.9	**13**	PP14
14	23.8	26.2	29.1	28.8	24.2	29.1	24.2	24.8	24.5	20.2	19.7	20.5	22.4		81.1	76.6	84.0	84.2	44.7	**14**	SE17
15	24.1	29.4	26.7	26.4	23.2	26.7	25.0	25.8	25.0	23.7	23.6	23.4	25.4	21.9		77.9	79.7	79.9	44.8	**15**	Iowa 609
16	28.2	29.6	28.9	28.1	26.4	28.4	31.2	30.2	29.2	23.3	23.3	23.8	25.8	26.4	24.5		76.2	76.0	44.8	**16**	Holte
17	23.8	27.6	29.1	28.9	24.7	29.1	23.7	24.2	25.0	23.2	20.2	23.2	23.4	17.4	22.9	27.2		98.2	44.0	**17**	JMK
18	23.8	28.4	29.4	28.6	25.3	28.9	22.9	24.2	25.0	23.5	21.9	23.5	23.7	17.1	22.9	27.5	1.9		44.0	**18**	Gray
19	74.4	87.0	84.3	84.8	82.7	84.8	72.9	73.0	72.2	78.6	76.9	79.1	80.9	79.9	79.6	78.6	82.7	82.2		**19**	De-072
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**		

Table 3. Infectious bronchitis virus S1 subunit variation at the same amino acid in strains representing three different serotypes

IBV serotype	Amino acid 91*	Amino acid 125	Amino acid 392
Mass 41	ALA (nonpolar)	MET (nonpolar)	ASP (uncharged polar)
Conn	GLY (nonpolar)	SER (uncharged polar)	LYS (charged polar)
FL-18288	ASN (uncharged polar)	ILE (nonpolar)	ILE (nonpolar)

*Amino acid numbering is in reference to Mass41 and includes the signal sequence.

Although GA-92 has a unique RFLP pattern, it is serologically related to the Arkansas serotype (23). Consistent with serologic data, we determined that the S1 glycoprotein sequence was 92.8% similar to the Ark 99 strain.

Initial characterization of isolates from California obtained from poultry between 1984 and 1993 (including CV-56b) indicated these variant strains (designated California variants) were serologically distinct from other IBV serotypes (23,24). The epidemiology of California variants is unknown. It is unlikely that these California variants originated directly from a live vaccine virus because the S1 glycoprotein sequences are not closely related to other IBV strains. Recombination of genomic RNA also is unlikely because no crossover sites were observed in the S1 gene. We observed a genetic drift from CV-56b (1991) to CV-9437 (1995) and CV-1686 (1995) with 98.0% and 97.6% similarity, respectively. Based on RFLP patterns and sequence analysis, we predict that CV-9437 and CV-1686 are serologically similar to other California variants represented by CV-56b. Another isolate from California, 1013, has a unique RFLP pattern, and the S1 glycoprotein sequence was more similar to Ark DPI (84.0%) than CV-56b (77.9%). Isolate 1013 may represent a new serotype or may belong to an existing serotype, such as the Ark serotype.

Following alignment of these newly reported S1 glycoprotein sequences with those previously published from other IBV strains, regions of sequence variation were detected between residues 59–96, 115–149, 255–309, and 378–395. Amino acid variations between 59–96 and 115–149 reported herein were similar to HVR I (residues 56–69) and HVR II (residues 117–131) of Mass and European strains as well as residues 53–148 in strains isolated from the United States (15,17). These variable regions were also similar to regions I (residues 38–67) and II (residues 97–141) associated with two separate virus-neutralizing, conformationally dependent epitopes (20). It appears that regions between residues 59–96 and 115–149 of the S1 glycoproteins reported in this study, may also be involved with epitopes that induce virus-neutralizing and serotype specific antibodies.

Two other regions of S1 glycoprotein variation were also observed. The first region between residues 255–309 corresponds to a previously reported hypervariable region (18,19), and the second domain between amino acids 378–395 has not been previously reported. Amino acid variation between residues 255–309 and 378–395 are similar to region III (274–387) associated with one virus-neutralizing epitope in European strains using monoclonal antibody resistant (MAR) mutants (20). In another study, MAR mutants of the Ark DPI strain had substitutions at Thr 304 and Arg 386 (Thr 306 and Arg 388 in this multiple alignment) of the S1 protein and were associated with one serotype-specific, virus-neutralizing epitope (21). During this investigation, amino acids surrounding 306 and 388 were identified as hypervariable. This suggests that sequences of the S1 glycoprotein between 255–309 and 378–395 may be important in forming serotype-specific, confirmationally dependent epitopes.

FL-18288 was 95.8% and 89.4% similar to Conn and Mass 41, respectively. This demonstrates that the S1 subunit may vary by a few amino acids giving rise to different serotypes. Amino acid variation in all three serotypes were observed at amino acids 91, 125, and 392. Amino acids 91 and 125 are within regions of variation, I and II, respectively. These regions are associated with two separate virus-neutralizing, conformationally dependent epitopes (20). Amino acid 392 is within region III that is associated with one virus-neutralizing, serotype-specific epitope.

Amino acid substitutions in the cleavage site among various IBV strains was detected during these studies. As reported previously, the cleavage site sequence for Beaudette and Mass 41 strains in

Arg-Arg-Phe-Arg-Arg (6,17), whereas Gray and JMK strains have Arg-Arg-Ser-Arg-Arg (34). In this study, differences were detected giving a cleavage site sequence of X1-Arg-X2-Arg-Arg. Herein X1 represents His or Arg and X2 represents Ser or Phe. The importance of these variations in the cleavage sequence to virulence is currently not known.

In summary, we observed S1 glycoprotein variations similar to others (residues 55–96, 115–149, and 255–309). However, a newly defined region of variation between amino acids 378–395 may interact with residues between 255–309 to form one conformationally dependent epitope with virus-neutralizing and serotype specificity. In addition, the S1 subunit may vary by 8% and remain the same serotype, whereas, a variation of only 4.2% may also change the serotype because of critical amino acid substitutions in the S1 glycoprotein. These data also demonstrate that variant IBV strains can circulate among poultry in geographically isolated areas.

Acknowledgments

We thank Ellen Collisson for providing the sequence of Iowa 609. We acknowledge Jessica A. Mulhern and Deborah A. Hilt for their assistance. Funding for these investigations was provided by the United States Poultry and Egg Association.

References

1. King D.J. and Cavanagh D. In Calnek B.W., Barnes H.J., Beard C.W., Reid W.M., and Yodes H.W., Jr. (eds). *Diseases of Poultry*, 9th ed. Iowa State University Press, Ames, Iowa, 1991, pp. 471–484.
2. Hofstad M.S., Avian Dis *25*, 650–654, 1981.
3. Roseberger J.K., Alphin R.L., and Krauss W.C., Avian Dis *20*, 199–201, 1975.
4. Spann W., Cavanagh D., and Horzinek M.C., J Gen Virol *59*, 2939–2952, 1988.
5. Cavanagh D., J Gen Virol *53*, 93–103, 1981.
6. Cavanagh D., Davis P.J., Pappin D., Binns M.M., Boursnell M., and Brown T., Virus Res *4*, 133–143, 1986.
7. Cavanagh D., J Gen Virol *64*, 2577–2583, 1983.
8. Cavanagh D. and Davis P.J., J Gen Virol *67*, 1443–1448, 1986.
9. Cavanagh D., Davis P.J., Darbyshire J.H., and Peters R.W., J Gen Virol *67*, 1435–1442, 1986.
10. Karaca K., Naqi S., and Gelb J.J., Avian Dis *36*, 903–915, 1992.
11. Koch G., Hartog L., Kant A., and van Roozelaar D.J., J Gen Virol *71*, 1929–1935, 1990.
12. 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., J Gen Virol *68*, 343–352, 1987.
13. Cavanagh D. and Davis P.J., J Gen Virol *69*, 621–629, 1988.
14. Jai W., Karaca K., Parrish C.R., and Naqi S.A., Arch Virol *140*, 259–271, 1995.
15. Wang L., Junker D., and Collission E.W., Virology *192*, 710–716, 1993.
16. Wang L., Junker D., Hock L., Ebiary E., and Collisson E.W., Virus Res *34*, 327–338, 1994.
17. Cavanagh D., Davis P.J., and Mockett A.P.A., Virus Res *11*, 141–150, 1988.
18. Kusters J.G., Niesters H.G., Lenstra J.A., Horzinek M.C., and van der Zeijst B.A.M., Virology *169*, 217–221, 1989.
19. Cavanagh D., Davis P.J., Cook J., Li D., Kant A., and Koch G., Avian Path *21*, 33–43, 1992.
20. Koch G., Kant A., Cook J.K.A., and Cavanagh D., World Veterinary Poultry Association, 154–160, 1991.
21. Moore K.M. and Jackwood M.W., Arch of Virol *142*, 1–8, 1997.
22. Winterfield R.W., Fadley A.M., and Hanley J.E., Avian Dis *15*, 305–311, 1972.
23. Gelb J.J., Personal communication.
24. Singbeil B., Personal communication.
25. Kwon H.M., Jackwood M.W., and Gelb J.J., Avian Dis *37*, 194–202, 1993.
26. Andreasen J.J., Jackwood M.W., and Hilt D.A., Avian Dis *35*, 216–220, 1991.
27. Jackwood M.W., Kwon H.M., and Hilt D.A., Avian Dis *36*, 403–409, 1992.
28. Jackwood M.W., Yousef N.M.H., and Hilt D.A., Avian Dis *41*, 105–110, 1997.
29. Smith L.M., Sanders J.Z., Kaiser R.J., Hughs P., Dodd C., Connell C.R., Heines C., Kent S.B.J., and Hood L.E., Nature (London) *321*, 673–681, 1986.
30. Sanger F., Nickelen S., and Coulson A.R., Proc Natl Acad Sci USA *74*, 5463–5467, 1977.
31. Higgins D.G. and Sharp P.M., CABIOS *5*, 151–153, 1989.
32. Sneath P.H.A. and Sokal R.R., *Numerical Taxonomy—The Principles and Practice of Numerical Taxonomy.* Freeman Press, San Francisco, CA, 1973.
33. Saitou N. and Nei M., Mol Biol Evol *4*, 406–425, 1987.
34. Kwon H.M. and Jackwood M.W., Virus Genes *9*, 219–229, 1995.
35. Binns M.M., Boursnell M., Cavanagh D., Pappin D., and Brown T., J Gen Virol *66*, 719–726, 1985.
36. Binns M.M., Boursnell M., Tomley F.M., and Brown T., J Gen Virol *67*, 2825–2831, 1986.