SHORT COMMUNICATION

Molecular characterization of an avian GA13-like infectious bronchitis virus full-length genome from Costa Rica

Ricardo A. Villalobos-Agüero[1] · Lisbeth Ramírez-Carvajal[2] · Rebeca Zamora-Sanabria[3,4] ·
Bernal León[2] · James Karkashian-Córdoba[1]

Abstract We describe the first whole-genome sequence of a GA13-like isolate of avian infectious bronchitis virus CK/CR/1160/16 (MN757859), obtained in 2016 in the province of Alajuela, Costa Rica. This virus caused an outbreak with great economic impact to the local poultry industry. The genome sequence is 27 696 bp in length, with the following genome organization 5′-UTR-Pol-S-3a-3b-E-4b-4c-M-5a-5b-N-6b-3′-UTR. The complete genome sequence has the highest sequence identity (94.03%) with DMV/1639/GA9977/2019 (MK878536) from Georgia, USA, and the lowest identity (86.03%) with ck/CH/LHLJ/08-6 (KX252788), from China. Analysis of the S1 subunit indicates that the Costa Rican isolate belongs to genotype I, lineage 17 (GI-17) and displays 96.89% identity with the S1 subunit of Ga-13/14255/14 (KM087780) (USA). Possible recombination events in genes S, E, M, 4b y 4c were detected, with Massachusetts, Connecticut, Arkansas and MA5 as potential parental types. This study highlights the importance of the epidemiological and molecular surveillance of avian infectious bronchitis.

✉ Ricardo A. Villalobos-Agüero
ricardo.villalobos@ucr.ac.cr

[1] Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica

[2] Laboratorio Nacional de Servicios Veterinarios (LANASEVE), Servicio Nacional de Salud Animal, Heredia, Costa Rica

[3] Escuela de Zootecnia, Universidad de Costa Rica, San José, Costa Rica

[4] Centro de Investigación en Nutrición Animal, Universidad de Costa Rica, San José, Costa Rica

Infectious bronchitis (IB) is a viral disease that affects chickens. It causes high morbidity and economic losses in the industry around the world [1, 10]. The etiological agent is the *Avian coronavirus* [8] or avian infectious bronchitis virus (IBV), member of the genus *Gammacoronavirus,* family *Coronaviridae,* order *Nidovirales* [2, 8, 13]. The virion has a lipid envelope and the genome is a positive-sense linear RNA of approx. 27.6 kb [1, 2, 10], with the following as the most common genome organization 5′-UTR-Pol-S-3a-3b-E-M-5a-5b-N-3′UTR [24, 35]. The first gene encodes proteins involved in replication and transcription, and it has two open reading frames (ORFs): 1a and 1b [1, 4, 13, 34] which are translated into polyprotein 1a and 1ab due to a change in the reading frame [1, 4, 34]. The last third of the genome has the structural genes for the spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins, with additional ORFs that codify for the accessory proteins 3a, 3b, 5a y 5b [24, 28, 34].

Different strains of IBV have been described around the world including Massachusetts, Beaudette, Holte, 4/91, Arkansas, Connecticut, D274, and QX-Like, among others [2, 5, 12]. The classification of IBV genotypes is based on the S1 sequence of the spike gene, and six genotypes which comprise 32 lineages have been described [1, 34]. Mass-like, Ark-like and Penn-like variants, plus one unassigned genotype designated as IBV-CR-53 [5, 23], have been reported in Costa Rica since 1990 [17, 23].

Starting in May 2016 until mid 2017, there was an IBV outbreak throughout farms in Costa Rica. Poultry exhibited mild respiratory infections and mortality. However, great economic losses due to carcass condemnation were

reported [30]. The IBV isolate associated to this outbreak was classified as a Georgia 13-like type (GA13-like), based on S1 gene sequence [30].

We isolated the virus associated to this outbreak in 9-to-11-day-old embryonated specific-pathogen-free (SPF) eggs [2]. For RNA isolation, allantoic fluid was collected and purified in a 30% sucrose cushion by ultracentrifugation at 27 800 rpm for 4 h [19]. The RNA from the pellet was extracted using TRIZOL (Ambion® CA, ES) in accordance with the manufacturer's instructions [35]. RNA was converted to double-stranded DNA using random hexamers, Superscript III and Klenow enzyme. The NGS library was prepared with Nextera^{TM} XT (Illumina®) reagents. RNA and dsDNA quantification and quality control were conducted using Quantus Fluorometer® (Promega) and a QIAxcel® (QIAGEN®). The library was sequenced on an Illumina MiSeq using a paired-end (2 × 250 bp) protocol. The quality control of the sequence run was analyzed with Sequence Analysis Viewer (SAV) (Illumina®). Short reads quality was analyzed using FASTQC and quality trimming was conducted using Trimmomatic [7]. De novo assembly was conducted using SPADES [6] and the longest contig was compared with viral sequences at NCBI using BLAST. Bowtie2 was used to align all short reads of the draft genome and Artemis [9] was used for visualization of the alignment. Automatic annotation was done using PROKKA [31] followed by manual curation using information from the ViPr database. Each CDS and genome feature was compared to existing sequences using BLASTP (nr database).

Possible recombination events were evaluated using the recombination detection program RDP4 V.4.95 [1, 2, 33, 35], to detect recombinations in at least five out of the seven possible methods [35]. The phylogenetic analysis was conducted by taking sequences of the S1 gene region and the IBV whole-genome sequences available in the GenBank database. The sequences were aligned using MAFFT algorithm available in *Guidance2* server [18, 32], and the software PartitionFinder2 was used to determine the best substitution model [20, 21]. The phylogenetic trees were made using the Bayesian inference with *Mr.Bayes* 3.2.6 [15]. Phylogenetic analyses were performed in the CIPRES Science Gateway Cluster [26].

The whole-genome sequence of isolate CK/CR/1160/16 was uploaded to the GenBank database under the accession number MN757859 and raw data were deposited in the SRA under accession number SRR10547950, BioProject number PRJNA592262, and BioSample number SAMN13419001. The complete sequence was 27 696 bp long, consistent with previously reported lengths [2, 29, 35], with thirteen ORFs, containing 9 genes, with two UTR regions, and a noncoding region between gene N and 6b (Table 1). The genome organization of isolate CK/

CR/1160/16 was 5′-UTR-Pol-S-3a-3b-E-4b-4c-M-5a-5b-N-6b-3′-UTR, which differs with the classic IBV gene distribution [24, 35], but has been previously reported [1, 27, 33].

The phylogenetic analysis of the S1 region shows that isolate CK/CR/1160/16 forms a cluster with sequences that belong to the genotype 1 lineage 17, where the variants from California and Pennsylvania isolates (1990s) are found [34] (Fig. 1a). Moreover, the Costa Rican isolate CK/CR/1160/16 shows a close phylogenetic relatedness to isolate GA-13/14,255/14 (KM087780) with a bootstrap value of 100 (Fig. 1a). The sequences from other regions in America form two different clusters: GI-11 (unique for South America, including sequences from Argentina and Brazil), and GI-16 (reported in Asia and Italy and including sequences from Argentina and Chile) [34] (Fig. 1a). The phylogeny obtained using the whole genome sequence shows that isolate CK/CR/1160/16 forms a cluster with isolates from USA, specifically with CAV/CAV56b/91 (GU393331) and Cal99/NE15172/95 (FJ904714) from California [25, 33] and with DMV/1639/GA9977/2019 (MK878536), from Georgia [14] (Fig. 1b). These sequences correspond to the GI-17 clasification, based on the S1 region.

Comparison of isolate CK/CR/1160/16 to other whole-genome sequences (Table 1) indicates that this isolate has the highest nucleotide sequence identity (94.03%) with DMV/1639/GA9977/2019, with which the nucleotide sequence identities for every gene were higher than 88%. The IBV sequence with the lowest identity (86.03%) was ck/CH/LHLJ/08-6. The S1 gene region is highly variable, with nucleotide sequence identities varying between 58.3–88.48% among the different IBV serotypes, due to a high mutation frequency as well as recombination events [2, 3, 11, 29]. For this reason, it is important to point out that the isolate in this study exhibits a very high sequence identity in the S1 gene (96.89%) with the Georgia 13 genotype (Ga-13/14255/14). Finally, ORF 6b shows the lowest sequence identity among all the coding regions in the genomes analyzed in this study.

Two possible recombination points were detected, shown in at least six of the seven models using RDP4 software. The first event has a beginning breakpoint position at nucleotide 20,410 and an ending breakpoint position at 23,695 that is found in the S gene (Fig. 2a). In this case, the minor parent was inferred as a Massachusetts (FJ904722) strain, and the putative major parent was determined as a Connecticut type (KF696629). The second recombination event starts at nucleotide 24,291 and ends at position 25,518 in the sequence. The major parent belongs to an Arkansas type (EU418976), and the minor parent to Ma5 (KY6226045), comprising a part of the E, M, 4B and 4C genes (Fig. 2b). Recombination hotspots in region S of

Table 1 Viral genes, genome location, deduced protein length and nucleotide sequence identity (%) of the CK/CR/1160/16 isolate (Genbank accession MN757859), compared to full-length sequences of representative IBV strains

| Genes and ORFs	Genome location	Protein lenght	Cov	KM087780 GA-13/14,255/14	MK878536 DMV/1639/GA9977/2019	GU393331 Cal56b	FJ904713 Mass41	EU418976 ArkDPI11	GQ504720 Arkansas/DPI	KP868573 CK/CH/LJL/130,908	KY626045 MaSV	GQ504722 Georgia 1998pass8	KF377577 4.91.V	KX2527 88Ck/CH/LHLJ/086
5′ UTR	1–507		231	–	**94.03**	93.03	92.64	92.10	91.97	91.56	91.15	89.94	89.37	86.03
orf 1a	508–12,363	3951	292	–	**99.61**	98.82	99.21	**99.61**	96.84	96.98	99.21	98.82	96.06	93.90
orf 1ab	508–20,396	6628*	294	–	**93.45**	92.79	93.24	91.40	91.46	90.6	90.31	92.0	88.88	85.22
S	20,347–23,856	1169	324	–	**94.39**	93.87	93.80	93.01	92.92	92.53	91.84	93.37	90.30	86.80
S1	20,347–22,019	558	316	96.89**	**91.12**	86.76	83.43	83.55	83.49	83.21	83.09	66.84	81.36	80.32
S2	22,020–23,856	611	332	–	**88.48**	86.99	79.95	80.73	80.73	80.13	80.01	58.32	76.91	75.16
3a	23,856–24,029	57	403	–	93.52	87.93	**97.70**	90.81	91.38	84.48	84.48	93.10	90.81	81.61
3b	24,029–24,223	64	448	–	93.33	93.33	94.36	**95.39**	**95.39**	90.26	84.62	94.87	90.77	72.31
E	24,204–24,533	109	374	–	88.79	78.74	88.49	88.49	88.49	88.49	**95.46**	93.64	84.24	84.99
M	24,505–25,182	225	382	–	89.68	90.76	90.86	91.00	91.00	91.00	**97.49**	95.72	91.45	88.84
4b	25,183–25,467	94	456	–	88.07	91.23	84.91	84.91	84.91	84.91	**98.60**	92.63	84.91	85.26
4c	25,388–25,558	56	366	–	**97.66**	83.91	84.21	84.21	84.21	84.21	**97.66**	89.77	86.55	80.57
5a	25,542–25,739	65	243	–	**94.95**	92.93	91.92	91.92	91.92	91.92	90.40	91.92	89.90	85.35
5b	25,736–25,984	82	215	–	97.99	97.99	98.39	99.60	99.60	**100.00**	96.39	99.20	97.19	91.17
N	25,927–27,156	409	269	–	95.85	97.48	**97.64**	97.56	97.48	97.40	94.44	94.96	90.81	87.97
NCR	27,157–27,164		131	–	100	100	100	100	100	100	–	26.09	100	87.50
6b	27,165–27,389	74	105	–	96.89	97.33	96.44	**98.22**	**98.22**	**98.22**	–	57.69	92.44	67.56
3′UTR	27,390–27,696		50	–	99.28	97.84	**99.45**	98.27	97.82	98.24	–	97.09	98.60	98.60

Cov: NGS coverage

*Polyprotein 1ab frameshift

**IBV variant with the highest percentage of similarity in the S1 region

The boldface indicates the highest nucleotide sequence indentity in differents genes

Underline indicates the lowest nucleotide identity in someone genes
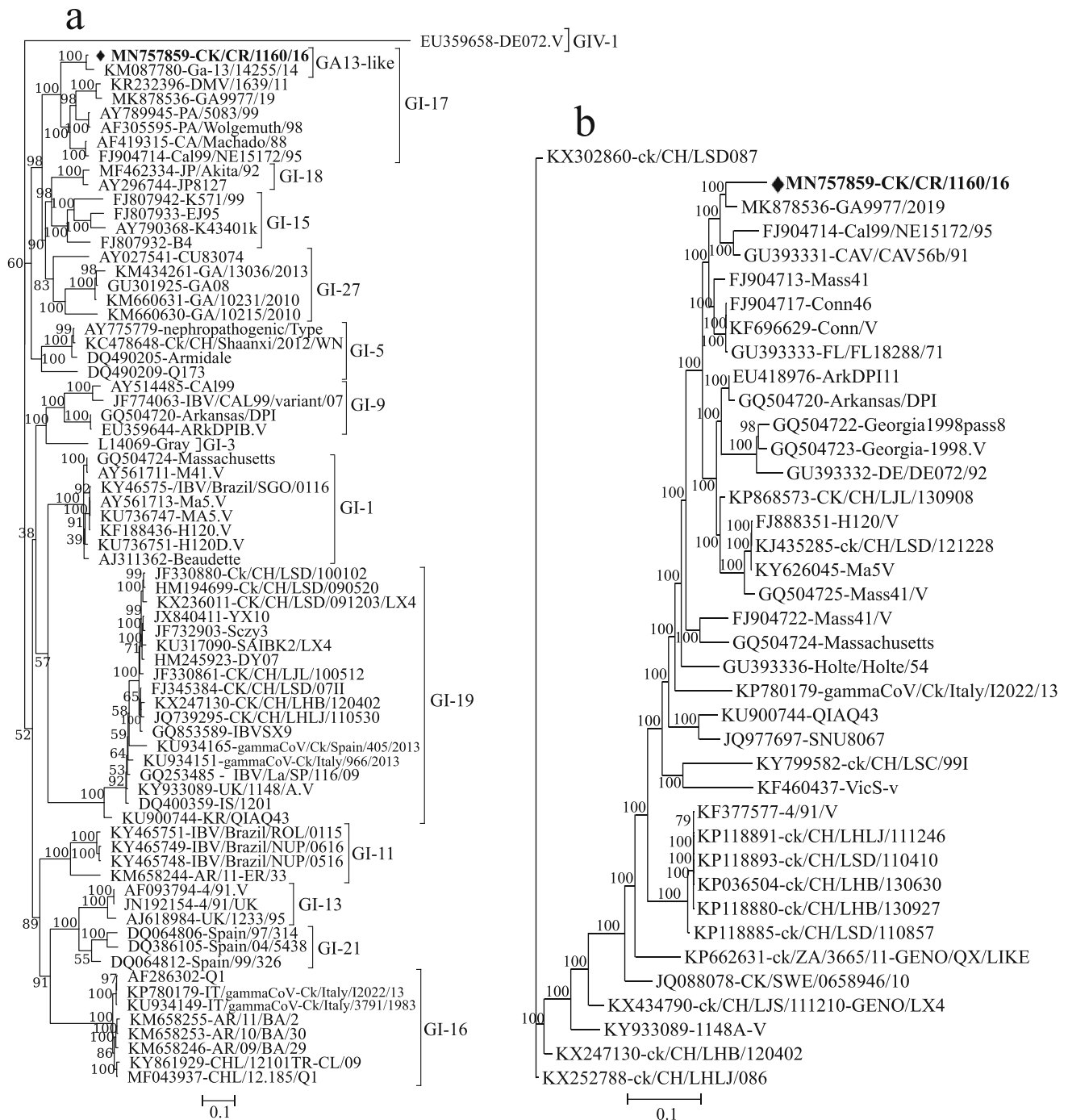
Fig. 1 Consensus phylogenetic trees constructed using (**a**) the S1 gene and (**b**) the full-length genome. Both trees were made using Bayesian Inference and four MCMC runs, 10,000,000 generations, and corroborated by the Maximun likelihood method, using Mega X. The CK/CR/1160/16 isolate (Genbank accession MN757859) is show in boldface

the genome, which are associated with the appearance of new virus variants have been described in the past [22, 33]. Recombination events in genes 3 and M have also been detected previously [22, 33].

Our results show that the 2016 outbreak of IBV in Costa Rica was caused by a virus that belongs to the GI-17 group, which includes strains native to the United States. More specifically, we confirmed that the outbreak was caused by a Ga-13/14255/14 strain similar to the one that circulated in the United States during in 2016 [16, 30]. Our whole genome analyses provide the first evidence that the isolate CK/CR/1160/16 may be the result of the recombination of at least four different variants (Mass, Connecticut, Arkansas and Ma5). Detection of recombination events supports
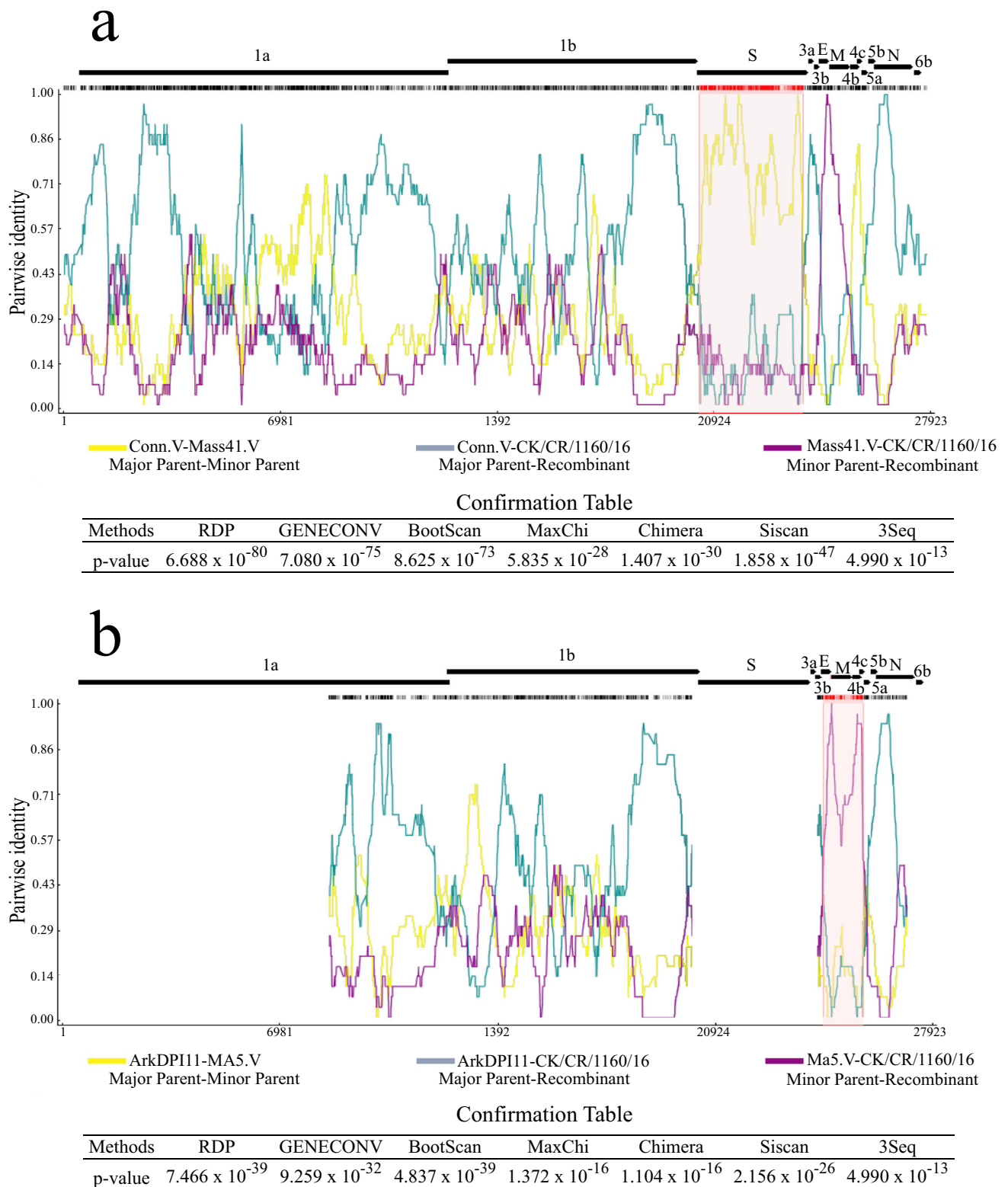
Fig. 2 RDP4 recombination plot and confirmation table of the recombination events detected on isolate CK/CR/1160/16. Putative recombination regions are shaded in red. **a** Breakpoint 20,410–23,695. The minor parent was inferred as Mass41.V and the putative major parent was determined as Conn.V. **b** Breakpoint 24,291–25,518. The major parent was inferred as ArkDPI11 and the putative minor parent was determined as Ma5.V

the need to maintain epidemiological surveillance, monitor the variants present in Latin America and optimize vaccination schemes, as outbreaks usually originate from variants not covered by vaccine serotypes [35].

The raw data of isolate CK/CR/1160/16 of the GA13-like strain has been deposited in the Sequence Read Archive (SRA) under number SRR10547950, BioProject number PRJNA592262, and BioSample number SAMN13419001 of the NCBI. The whole-sequence of the genome, has been uploaded in the GenBank database under accession number MN757859.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in animals were in accordance with the ethical standards of the "Comité Institucional de Cuido y Uso de Animales" (CICUA) of the University of Costa Rica.

References

1. Abozeid HH, Paldurai A, Khattar SK, Afifi MA, El-Kady MF, El-Deeb AH, et al. Complete genome sequences of two avian infectious bronchitis viruses isolated in Egypt: evidence for genetic drift and genetic recombination in the circulating viruses. Infect Genet Evol. 2017;53:7–14.

2. Abro SH, Renström LH, Ullman K, Belák S, Baule C. Characterization and analysis of the full-length genome of a strain of the European QX-like genotype of infectious bronchitis virus. Arch Virol. 2012;157:1211–5.

3. Amarasinghe A, Senapathi UDS, Abdul-Cader MS, Popowich S, Marshall F, Cork SC, et al. Comparative features of infections of two Massachusetts (Mass) infectious bronchitis virus (IBV) variants isolated from Western Canadian layer flocks. BMC Vet Res. 2018;14:1–12.

4. Ammayappan A, Upadhyay C, Gelb J, Vakharia VN. Complete genomic sequence analysis of infectious bronchitis virus Ark DPI strain and its evolution by recombination. Virol J. 2008;5:157.

5. Bande F, Arshad SS, Omar AR, Hair-Bejo M, Mahmuda A, Nair V. Global distributions and strain diversity of avian infectious bronchitis virus: a review. Animal Heal Res Rev. 2017;18:70–83.

6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.

7. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

8. Carstens EB. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2009). Arch Virol. 2010;155:133–46.

9. Carver T, Harris SR, Berriman M, Parkhill J, Mcquillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 2012;28:464–9.

10. Cavanagh D. Coronavirus avian infectious bronchitis virus. Vet Res. 2007;38:281–97.

11. DeFraga AP, Gräf T, Pereira CS, Ikuta N, Fonseca AS, Kazantzi F, et al. Phylodynamic analysis and molecular diversity of the avian infectious bronchitis virus of chickens in Brazil. Infect Genet Evol. 2018;61:77–83.

12. DeWit JJ, Cook JK, Van der Heijden HM. Infectious bronchitis virus variants: a review of the history, current situation and control measures. Avian Pathol. 2011;40:223–35.

13. Durães-Carvalho R, Caserta LC, Barnabé AC, Martini MC, Simas PV, Santos MM, et al. Phylogenetic and phylogeographic mapping of the avian coronavirus spike protein-encoding gene in wild and synanthropic birds. Virus Res. 2015;201:101–12.

14. Goraichuk IV, Kulkarni AB, Williams-coplin D, Suarez DL, Afonso CL. First complete genome sequence of currently circulating inectious bronchitis virus strain DMV/1639 of the GI-17 lineage. Microbiol Resour Announc. 2019;8:e00840-e919.

15. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001;17:754–5.

16. Jackwood MW. Impact of respiratory diseases with special emphasis to emerging infectious bronchitis virus. In: Proceedings of 65th western poultrt disease conference. 2016. p. 116–7.

17. Jiménez C, Cortés R, Ramírez M, Prendas J, Arrieta E, Lindahl J, et al. Diagnosis and epidemiology of IBV infections in Costa Rica. In: IV symposium avian corona-and pneumovirus infect Rauischholzhausen, Germany 20–23 June 2004. 2004. p. 39–49

18. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.

19. Keep SM, Bicherton E, Britton P. Partial purification of IBV and subsequente isolation of viral RNA for next-generation sequencing. In: Maier H, Brikerton E, Britton P, editors. Coronaviruses methods. New York: Humana Press; 2015. p. 109–12. https://doi.org/10.1007/978-1-4939-2438-7_11.

20. Lanfear R, Calcott B, Ho YWS, Guindon S. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 2012;29:1695–701.

21. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2016;34:772–3.

22. Lee CW, Jackwood MW. Evidence of genetic diversity generated by recombination among avian coronavirus IBV. Arch Virol. 2000;145:2135–48.

23. Lindahl J. Infectious bronchitis virus and infectious bursal disease virus; a study performed at the Universidad Nacional of Costa Rica. Ph.D. Thesis. Sveriges lantbruksuniversite. Upssala. 2004.

24. Mardani K, Noormohammadi AH, Hooper P, Ignjatovic J, Browning GF. Infectious bronchitis viruses with a novel genomic organization. J Virol. 2008;82:2013–24.

25. McKinley ET, Jackwood MW, Hilt DA, Kissinger JC, Robertson JS, Lemke C, et al. Attenuated live vaccine usage affects accurate measures of virus diversity and mutation rates in avian coronavirus infectious bronchitis virus. Virus Res. 2011;158:225–34.

26. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES science gateway for inference of large phylogenetic trees. Gatew Comput Environ Work. 2010;2010:1–8. https://doi.org/10.1109/GCE.2010.5676129.

27. Mousavi FS, Ghalyanchilangeroudi A, Hosseini H, Fasaei BH, Ghafouri SA, Abdollani H, et al. Complete genome analysis of Iranian IS-1494 like avian infectious bronchitis virus. Virusdisease. 2018;29:390–4.

28. Nafaji H, Langeroudi AG, Hashemzadeh M, Karimi V, Madagar O, Ghafouri SA, et al. Molecular characterization of infectious bronchitis viruses isolated from broiler chicken farms in Iran, 2014–2015. Arch Virol. 2016;161:53–62. https://doi.org/10.1007/s00705-015-2636-3.

29. Naguib MM, Höper D, Arafa A, Setta AM, Abed M, Monne I, et al. Full genome sequence analysis of a newly emerged QX-like infectious bronchitis virus from Sudan reveals distinct spots of recombination. Infect Genet Evol. 2016;46:42–9.

30. Roman T, Silva M, Barbosa J, Gómez L, Hernandez D, Zamora R, et al. Effectiveness of a vaccine program based on the Protectotype® concept against an infectious bronchitis variant virus strain challenge (GA13) in Costa Rica. In: Proccedings of American Association of Avian Pathologists Annual Meeting. 2018. p. 3.

31. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

32. Sela I, Ashkenazy H, Katoh K, Pupko T. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. Nucl Acids Res. 2015;43:W7-14.

33. Thor SW, Hilt DA, Kissinger JC, Paterson AH, Jackwood MW. Recombination in avian gamma-coronavirus infectious bronchitis virus. Viruses. 2011;3:1777–99.

34. Valastro V, Holmes EC, Britton P, Fusaro A, Jackwood MW, Cattoli G, et al. S1 gene-based phylogeny of infectious bronchitis virus: an attempt to harmonize virus classification. Infect Genet Evol. 2016;39:349–64. https://doi.org/10.1016/j.meegid.2016.02.015.

35. Zhao Y, Liu XY, Cheng JL, Wu YP, Zhang GZ. Molecular characterization of an infectious bronchitis virus strain isolated from northern China in 2012. Arch Virol. 2014;159:3457–61.