CLAUDIA CARRANZA CHAMORRO

Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil

São Paulo
2015

CLAUDIA CARRANZA CHAMORRO

Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil

Dissertação apresentada ao Programa de Pós-Graduação em Epidemiologia Experimental Aplicada às Zoonoses da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo para a obtenção do título de Mestre em Ciências

Departamento:

Medicina Veterinária Preventina e Saúde Animal

Área de concentração:

Epidemiologia Experimental Aplicada às Zoonoses

Orientador:

Prof. Dr. Antonio José Piantino Ferreira

São Paulo

2015

DADOS INTERNACIONAIS DE CATALOGAÇÃO-NA-PUBLICAÇÃO

(Biblioteca Virginie Buff D'Ápice da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo)

T.3229
FMVZ

Carranza Chamorro, Claudia
Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil / Claudia Carranza Chamorro. -- 2015.
38 f. il.

Dissertação (Mestrado) - Universidade de São Paulo. Faculdade de Medicina Veterinária e Zootecnia. Departamento de Medicina Veterinária Preventiva e Saúde Animal, São Paulo, 2015.

Programa de Pós-Graduação: Epidemiologia Experimental Aplicada às Zoonoses.

Área de concentração: Epidemiologia Experimental Aplicada às Zoonoses.

Orientador: Prof. Dr. Antonio José Piantino Ferreira.

1. dN/dS. 2. n-glicosilação. 3. BR-I. 4. BR-II. 5. Glicoproteína da espícula. I. Título.

CERTIFICADO

Certificamos que o Projeto intitulado "Diversidade genética do vírus da bronquite infecciosa isolado de aves de produção no Brasil. ", protocolado sob o CEUA nº 9565130215, sob a responsabilidade de **Antonio Jose Piantino Ferreira** *e equipe; Claudia Carranza Chamorro* - que envolve a produção, manutenção e/ou utilização de animais pertencentes ao filo Chordata, subfilo Vertebrata (exceto o homem), para fins de pesquisa científica (ou ensino) - encontra-se de acordo com os preceitos da Lei 11.794, de 8 de outubro de 2008, com o Decreto 6.899, de 15 de julho de 2009, com as normas editadas pelo Conselho Nacional de Controle da Experimentação Animal (CONCEA), e foi **aprovado** pela Comissão de Ética no Uso de Animais da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo (CEUA/FMZV) em reunião de 28/10/2015.

We certify that the proposal "Genetic Diversity of Avian Infectious Bronchitis Virus isolated from poultry in Brazil.", utilizing 100 Birds (100 males), protocol number CEUA 9565130215, under the responsibility of **Antonio Jose Piantino Ferreira** *and team; Claudia Carranza Chamorro* - which involves the production, maintenance and/or use of animals belonging to the phylum Chordata, subphylum Vertebrata (except human beings), for scientific research purposes (or teaching) - it's in accordance with Law 11.794, of October 8 2008, Decree 6899, of July 15, 2009, with the rules issued by the National Council for Control of Animal Experimentation (CONCEA), and was **approved** by the Ethic Committee on Animal Use of the School of Veterinary Medicine and Animal Science of São Paulo University (CEUA/FMZV) in the meeting of 10/28/2015.

Vigência da Proposta: de 02/2015 a 07/2015 Área: Vps/epidemiologia

Procedência: Granja de Aves SPF (Informação em anexo)

Espécie: Ave Gênero: Machos idade: 9 dias N: 100

Linhagem: ovo embrionado Peso: 46 g

São Paulo, 04 de novembro de 2015

Profa. Dra. Denise Tabacchi Fantoni
Presidente da Comissão de Ética no Uso de Animais
Faculdade de Medicina Veterinária e Zootecnia
Universidade de São Paulo

FOLHA DE AVALIAÇÃO

Autor: CARRANZA CHAMORRO, Claudia

Título: Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil

Dissertação apresentada ao Programa de Pós-Graduação em Epidemiologia Experimental Aplicada às Zoonoses da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo para a obtenção do título de Mestre em Ciências

Data: _____/_____/_____

Banca Examinadora

Prof. Dr._____

Instituição:_____ Julgamento:_____

Prof. Dr._____

Instituição:_____ Julgamento:_____

Prof. Dr._____

Instituição:_____ Julgamento:_____

A Isabel, Francisco, Leonor, Vilma, Pauli, Veco, Tata,
Diego, Morris, Petra y Nichi. Mi núcleo, fortaleza y motivación.

AGRADECIMENTOS

A mi madre por ser la persona que más fe en este mundo me ha tenido y me alentó en cada segundo de esta experiencia. A mi padre por ser mi ejemplo de esfuerzo, dedicación, constancia y equilibrio en todos los aspectos de la vida.

A mi tía Leo por ser mi segunda madre e incentivar y apoyar desde que tengo memoria mi desarrollo personal e intelectual. A mi abuela y mi Pauli porque su amor, pese a la distancia, fue tan tangible como si estuviera con ellas en Lima y me hizo extrañar menos mi hogar. A mi Tata, porque su buen humor contagiante era mi dosis de adrenalina.

A mi hermano por mantenerme en tierra. A Nichi por el apoyo incondicional y desinteresado, y enseñarme a lucharla hasta el final.

A mi orientador, el Dr. Antonio Piantino por la oportunidad de llegar al Lab. de Ornitopatología, la confianza para el desarrollo de los proyectos encomendados, la libertad para desarrollar proyectos paralelos y haber tenido la enorme paciencia de guiarme y aguantarme durante toda esta iniciación al mundo científico/académico. A Claudete por ser amiga, consejera y un gran apoyo durante estos años en Brasil.

A mis amigos incondicionales del cole que aguantaron mis conversas nocturnas o mis saludos matutinos para matar la nostalgia de estar lejos. En especial a Charlie por el ejemplo y la ayuda, y a mi soulmate Eveling por estar ahí 24/7.

A mis amigos del Lab. de Ornitopatología Luis y Silvana por su apoyo y guía constante y los buenos momentos que pasamos en Ecuador. A las meninas fit Natalia y Luciana por todas las locuras y los bate papos salvadores. A Mauricio por ayudarme a distraerme del estrés del día a día. . A Roberta, Vasco, Pedro, Paulinha, Rossi, Ricardo, Danival y toda la gente maravillosa del VPS por los momentos geniales. Y a mi familia adoptiva en SP, Seba (mi hermano chileno), Jason, Carlos, Pam, Karina, Carlitos y Nicolás, los adoro.

A Dios, por absolutamente todo.

“Expose yourself to your deepest fear; after that, fear has no power, and the fear of freedom shrinks and vanishes. You are free.”

Jim Morrison

RESUMO

CARRANZA CHAMORRO, C. **Diversidade genética do vírus da bronquite infecciosa isolado de aves de produção no Brasil.** [Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil]. 2015. 38 f. Dissertação (**Mestrado** em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2015.

O vírus da bronquite infecciosa das galinhas (IBV) é o agente causador de uma doença aviária economicamente importante. No Brasil, esta doença ocasiona problemas respiratórios, renais e reprodutivos em aves de todas as idades, apesar da vacinação constante com a cepa Massachusetts H120. Esta falha na proteção conferida pela vacina é ocasionada por mutações nos nucleotídeos do gene da glicoproteína da espícula, a qual está envolvida no processo de interação comas células do hospedeiro, a neutralização e a indução de imunidade protetora. As variantes brasileiras resultantes dessa mutação genética estão presentes desde os anos 80 e este estudo teve como objetivo analisar epidemiologicamente e caracterizar molecularmente os vírus variantes existentes durante 2010-2015 e realizar uma análise bioinformática das sequências disponíveis no GenBank em um período de 40 anos. Das 453 amostras analisadas, 61,4% foram positivas para IBV e 75,9% delas foram consideradas variantes e foram detectados em aves de todas as idades, distribuídos em todas as 5 regiões do Brasil. Um fragmento de 559-566 pb foi obtido a partir de 12 isolados, onde BR-I foi a variante predominante ao contrario que apenas um isolado pertencia ao genótipo BR-II. Análise bioinformática de 40 anos de variantes do IBV brasileiros revelou uma predominância de codões com as substituições não sinónimos no primeiro terço do gene S1 e uma relação dN / dS de 0,6757, indicando que esta porção do gene estava sob selecção negativa. Além disso a previsão de pontos de de N-glicosilação mostrou que a maioria das amostras variantes BR-I (entre o 2003 e início de 2014) apresentam um ponto adicional na posição 20, enquanto as variantes mais novas não apresentam esse ponto de n-glicosilação. Estes resultados sugerem que as variantes brasileiras teriam sofrido mutações provavelmente drásticas em alguns pontos do genoma, entre os anos de 1983 a 2003 e depois de atingir uma estrutura antigênica eficaz o suficiente para a

invasão e replicação em seus hospedeiros, o processo de seleção mudou para seleção negativa.

Palavras-chave: dN/dS. n-glicosilação. BR-I. BR-II. Glicoproteína da espícula.

ABSTRACT

CARRANZA CHAMORRO, C. **Genetic diversity of avian coronavirus infectious bronchitis detected from commercial poultry in Brazil.** [Diversidade genética do vírus da bronquite infecciosa isolado de aves de produção no Brasil.]. 2015. 38 f. Dissertação (Mestrado em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2015.

Infectious bronchitis virus (IBV) is the causative agent of an economically important disease of poultry. In Brazil this disease causes respiratory, renal and reproductive problems in birds of all ages, despite constant vaccination with the Massachusetts strain H120. This lack of immunological protection is known to be due the genetic variation in the spike glycoprotein of IBV, which is involved in host cell attachment, neutralization and the induction of protective immunity. Brazilian IBV variants resulting of this genetic variation are present since the 80s and this study aimed to epidemiologicaly analyze and molecularly characterize the existing variants during 2010-2015 and perform a bioinformatics analysis of the available sequences of IBV variants in a 40 year period. Of the 453 samples tested, 61.4% were positive for IBV and 75.9% of them were considered variants and were detected in birds of all ages, distributed in all five Brazilian regions. A fragment of 559-566 bp was obtained from 12 isolates, where BR-I was the predominant variant while only one isolate belonged to the BR-II genotype. Bioinformatics analysis of the sequences of 40 years of Brazilian IBV variants was performed and the ratio of non-synonymous substitutions per non-synonymous site (dn) to synonymous substitutions per synonymous site (ds) dN/dS was calculated. It revealed a predominance of codons with non-synonymous substitutions in the first third of the S1 gene and a dN/dS ratio of 0.6757, indicating that this portion of the gene was under negative selection. Additionally prediction of N-glycosilation sites showed that most of the BR-I variants (from 2003 to early 2014) present an extra site at animoacid position 20, while the newest ones lack this feature.Together these results suggest that IBV Brazilian variants had probably suffered drastic mutations in some points between the years 1983 to 2003 and after achieving an antigenic structure effective enough for invasion and replication in their hosts, the selection processes became silent.

Keywords: dN/dS. n-glycosilation. BR-I. BR-II. Spike glycoprotein.

SUMMARY

1 INTRODUCTION

The need for rapid and effective ways of providing nutrients for the increasing human population has led to an overwhelming rise in aviculture, turning it ubiquitous, also spreading the pathogens that birds are susceptible of. One of these pathogens is the Avian Coronavirus Infectious Bronchitis (IBV), which is a highly infectious agent of domestic birds and is characterized by systemic signs that affect the respiratory, reproductive and renal systems. Since its first discovery more than 80 years ago, the virus has been detected in every poultry-producing country in the world.

Globalization has significantly improved the connectivity of people and animals from distant areas, resulting in an enhanced possibility for pathogen exchange. Additionally, the genetic improvements of desirable characteristics of the several boiler and layer linages, as well as the industrialization of poultry have triggered the need for the implementation of rigorous biosecurity measures and vaccination programs. These in turn have increased the immunological pressure on pathogens, leading them to develop strategies to survive and persist in time.

The monitoring of pathogenic agents like IB is an extremely important epidemiologic tool for the development of more effective measures for disease prevention and control. In South America, where poultry production is such an important economic activity, the need for rapid identification and characterization of the virus variants as well as the determination of its geographical origin is a must, in order to ensure the successful continuity of the production chain.

2 GENETIC DIVERSITY AND EVOLUTION OF BRAZILIAN VARIANT ISOLATES OF AVIAN CORONAVIRUS INFECTIOUS BRONCHITIS

2.1 INTRODUCTION

Coronaviruses possess the largest genomes among RNA viruses. Their basic organization consists of non-segmented positive-sense, enveloped, single-stranded RNA genomes that range from 27 Kb to 32 Kb in length, (MASTERS, 2006; WOO et al., 2009). Three genera are described within the Coronavirinae subfamily. The first is the Alphacoronaviruses which mainly consist of the feline coronaviruses, human common cold coronaviruses, Rhinolophus bat coronavirus and porcine epidemic diarrhea virus. The second, the Betacoronavirus genus includes mouse hepatitis virus (MHV), bovine and equine coronaviruses, severe acute respiratory syndrome coronaviruses (human SARS-CoV and SARS-like bat coronaviruses) and most recently Middle East respiratory syndrome coronavirus (MERS-CoV). The third group is the Gammacoronavirus genus which includes the Beluga whale coronavirus SW1 and the avian coronaviruses, which group all known coronaviruses isolated from birds such as the infectious bronchitis virus (IBV) of poultry (Figure 1) (GONZÁLEZ et al., 2003; DE GROOT et al., 2013).

The structure of the IBV is primarily spherical, although some pleomorphism can be observed. The genome consists of a multidomain replicase gene with two partially overlapping open reading rrames (ORFs 1a and 1b), that are expressed by ribosomal frameshifting. The ORF1 encodes 15 nonstructural proteins within the 5'-proximal end and occupies almost two thirds of the genome (ZIEBUHR, 2005; DE GROOT et al., 2012; NEUMAN et al., 2014). The rest of the genome encodes the following structural and accessory proteins (in this order from 5' to 3'): S, the spike protein, which binds to receptors and mediates membrane fusion; accessory proteins 3a and 3b with unknown function; E, envelope protein important for virus entry and assembly; M, the membrane glycoprotein, which spans the membrane three times and forms an internal core; accessory proteins 5a and 5b; and N, the nucleocapsid phosphoprotein which wraps the viral RNA genome (BRIAN; BARIC, 2005; MASTERS, 2006; MASTERS; PERLMAN, 2013).

IBV infection initiates via the respiratory tract regardless the tissue tropism of the strain, which may be respiratory, enteric, renal or reproductive (SCHALK; HAWN, 1931; GANAPATHY et al., 2012; TORRES et al., 2013). Then, viraemia occurs and the virus is widely disseminated to other tissues to replicate in the epithelial cells of the above mentioned systems (DE WIT, 2000). Broilers, layers and breeders of all ages seem to be susceptible to the pathogen (CHACON et al., 2011). However older birds are less susceptible to developing the disease. The virus can persist in the alimentary tract in asymptomatic animals.

The diagnosis of infectious bronchitis (IB) is based on many factors such as the clinical history, lesions, IBV titers, virus isolation and detection of viral RNA. This last factor, together with virus isolation, is the most frequently used technique for the diagnosis and epidemiologic analysis of IBV outbreaks (CAVANAGH; GELB, 2008). These techniques are used due to their versatility, low cost, high specificity andfor RNA molecular detection, their fast results and low absolute amount of material needed for viral detection (detection limit). However, for successful detection, knowledge of IBV pathogenesis is essential. Moreover, researchers should take into account the time between the beginning of the outbreak, the sampling time point and the level of immunity in the chicken at the moment of infection; other important factors include the collection of a sufficient number of samples, the sample choice and the sample quality (based on clinical signs) (DE WIT, 2000).

An IBV outbreak in vaccinated birds from vaccinated breeders, indicates that the virus has been able to overcome the host defenses. In an outbreak, both the innate and adaptive immune responses (the latter being the basis for the development of humoral and cell-mediated immune responses) failed to control the infectious capacity of the pathogenic field strain. This break in the host's immunity may be due to several factors, such as the misapplication of the vaccine or a change in the molecular structure of the field virus due to intense external immunological pressure. When facing outbreaks of IBV due to vaccine misapplication the problem can be rapidly overcome by certifying a more accurate vaccination process. However if the flocks are accurately vaccinated and the vaccination and biosecurity plans are carefully elaborated, then the problem may lay in the virus itself. The S1 region of the gene, contains antigenic sites (KANT et al., 1992) that are important for the development of humoral and cellular immune responses against the virus, and changes corresponding to as little as 5% in the S1 sub-unit have been able to alter a

vacine’s protection capacity (CASAIS et al., 2003). The importance of rapidly, efficiently and constantly detecting IBV and genotyping pathogenic strains highlights the need for the implementation of more effective prevention and control measures against a constantly evolving thread such as IBV.

2.2 MATERIAL AND METHODS

2.2.1 Field samples

A total of 453 samples from Brazilian broiler, breeder, layer and grandparent flocks were used for the detection, genotyping and evolutionary bioinformatics analyses of avian coronavirus infectious bronchitis. Samples were collected between January 2010 and July 2015 from flocks with suggestive signs of infectious bronchitis from all 5 Brazilian regions (central-west, north, north-east, south and south-east) and were sent to the Laboratory of Avian Diseases at the University of Sao Paulo for processing. Each sample consisted of a pool of a specific organ collected from 5 different birds from the same flock.

2.2.2 RNA extraction

Pooled samples from each tissue type and from each flock were homogenized separately and diluted in 1.5-mL microcentrifuge tubes containing 0.1 M sterile phosphate-buffered saline (PBS, pH 7.4) at a 1:1 ratio. The suspensions were frozen at -20 °C for 10 minutes and thawed at 56 °C for 1 minute, this process was repeated three times, including the homogenization step between the freeze-thaw cycles and vortexing for 20 seconds. For RNA extraction the samples were centrifuged for 30 minutes at 12.000 x g at 4 °C. RNA was extracted from the supernatants using TRIzol® (Invitrogen, Valencia, CA, USA) according to the manufacturer's instructions and stored at -80 ºC prior to use. Massachusetts live

vaccine H120 was used as the positive control and ultra-pure water was used as the negative control for the molecular reactions.

2.2.3 IBV detection

The extracted RNA from all 453 samples was tested for the presence of avian coronavirus IBV using the reaction conditions and primers previously describeb (CAVANAGH et al., 2002) which targeted the 3' untranslated (UTR) region of IBV.

2.2.4 IBV genotyping by multiplex RT-nested PCR

A previously described multiplex RT-nested PCR (CAVANAGH et al., 1999) targeting the S gene was used for the identification of IBV samples belonging to the genotypes Massachusetts strain H120, 793B and D274. Samples with no amplified fragments were considered non-typified IBV variants.

2.2.5 Amplification and partial S1 sequencing

A total of 39 samples that were considered non-typified variants according to the multiplex RT-nested PCR results (item 3.2.4) and with collection dates between 2010 and 2015, were submitted for amplification of a segment of the S1 gene. The reaction was performed in a final volume of 25 µL using 2.5 µL of 10X buffer, 4 µL of 1.25 mM dNTPs, 5 µL of each 10 pmol primer [Oligo 5' (HYUK MOO KWON; JACKWOOD, 1995) and CCCR (5'-CWARATCMCCRTTTARRTAHAC-3')], 1.25 of $MgCl_2$ 50 mM, 0.1 µL of Platinum Taq 1,25U (Invitrogen ™ by Life Technologies Brazil), 4.65 µL of ultra-pure water and 2.5 µL of cDNA. Amplified fragments 559 – 566 bp in length were purified directly from an agarose gel using GPX PCR DNA and Gel Band Purification Kit (GE Healthcare, Piscataway, NJ, USA) according to

manufacturer's instructions. A Thermo Scientific NanoDrop 2000c was used for DNA sample quantification. Each purified product was sequenced in the forward and reverse directions using a Big Dye Terminator version 3.1 Cycle Sequencing kit (Applied Biosystems by Life Technologies, Carlsbad, CA, USA). The sequencing reactions were run in an ABI 3730 DNA analyzer (Applied Biosystems by Life Technologies, Carlsbad, CA, USA).

2.2.6 Phylogenetic analysis

The obtained forward and reverse nucleotide sequences were assembled and the amino acid sequences were deduced using the software CLC Main Workbench v7.5 (CLC Bio-Qiagen, Aarhus, Denmark). The alignment was peformed with BioEdit v7.2.5 (Carlsbad, CA, USA) software using the CLUSTAL W method. The reliability of the alignment was verified with the p-distance method (THOMPSON; PLEWNIAK; POCH, 1999). Statistical selection of the best-fit model of amino acid substitution was performed with the JModel Test (POSADA, 2008) for the construction of maximum likelihood phylogenetic trees with 1000 bootstrap replicates, both tools were available in MEGA software v6.0 (TAMURA et al., 2013).

For the phylogenetic tree construction the first alignment was generated using the obtained sequences and the following partial S1 reference sequences from diverse IBV strains detected worldwide available in GenBank: GU383093 (USP 40), GU383093 (USP 56), JX182775 (UFMG/G), JX559817 (SB-A2398), JX559819 (SB-A2401), JX559821 (SB-A2960), JX559822 (SB-A2962), AF349620 (Quebec 16), AF349621 (Quebec mv), AJ457137 (Italy 02), AF227438 (Q1), DQ068701 (CK/CH/LDL/97I/97), AF093793 (4/91 attenuated), AF093794 (4/91 pathogenic), AF169859 (Ark/15C/96), AF210735 (JX), AF193423 (QX), KC795604 (QXIBV), X02342 (Beaudette), X04722 (M41), AF274438 (GA/2787/98), U77298 (Delaware) and DQ412629 (SARS CUHKtc55NS).

2.2.7 Evolution and bioinformatics analyses

A second alignment was generated using our obtained sequences and the following partial S1 Brazilian strain sequences available GenBank (1975 to 2010): USP-33 (GU383070), USP-34 (GU383071), USP-35 (GU383072), USP-36 (GU383073), USP-37 (GU383074), USP-38 (GU 383075), USP-39 (GU383076), USP-40 (GU383077), USP-41 (GU383078), USP-42 (GU383079), USP-43(GU383080), USP-44 (GU383081), USP-45 (GU 383082), USP-46 (GU383083), USP-47 (GU383084), USP-48 (GU383085), USP-49 (GU383086), USP-50 (GU383087), USP-51 (GU383088), USP-52 (GU 383089), USP-53 (GU383090), USP-54 (GU383091), USP-55 (GU383092), USP-56 (GU383093), USP-57 (GU383094), USP-58 (GU383095), USP-59 (GU 383096), USP-60 (GU383097), USP-61 (GU383098), USP-62 (GU383099), USP-63 (GU383100), USP-65 (GU383102), USP-66 (GU 383103), USP-67 (GU383104), USP-68 (GU383105), USP-69 (GU383106), USP-70 (GU383107), USP-71 GU383108), USP-72 (GU383109), USP-73 (GU383110), UFMG/G (JX182775), UFMG 29-78 (JX182790), UFMG 283 (JX182787) SB-A2398 (JX559817), SB-A2400 (JX559819), SB-A2960 (JX559821), SB-A2962 (JX559822) and H120 vaccine (M21970).

2.2.7.1 PHYLOGENY OF 40 YEARS OF BRAZILIAN IBV

To estimate the evolutionary divergence in S1 in IBV Brazilian variants over a span of 40 years, the percent nucleotide and amino acid identities were calculated using MEGA software v6.0 (TAMURA et al., 2013). A maximum likelihood phylogenetic tree with 1000 bootstrap replicates was also constructed.

2.2.7.2 SELECTION PRESSURE ANALYSIS

The prediction of codons under potential positive selection was calculated using MEGA software v6.0 (TAMURA et al., 2013) with the statistical test dN – dS, where dS was the number of synonymous substitutions per site (s/S) and dN was the number of nonsynonymous substitutions per site (n/N). A positive value for the test indicated an overabundance of nonsynonymous substitutions at each codon. Potential selective pressure along the partial S1 segment was also calculated as the ratio of non-synonymous substitutions per non-synonymous site (dN) to synonymous substitutions per synonymous site (dS) dN/dS.

2.2.7.3 PREDICTION OF N-GLYCOSYLATIONS

Changes in the potential N-glycosylation sites were predicted within the first third of the S1 glycoprotein. The analysis was performed with on-line tool available at http://www.cbs.dtu.dk/services/NetNGlyc and included two of the three hypervariable regions identified in the S1 subunit of IBV. The relationship between changes of N-glycosylation site and the first two antigenic sites located on the first third of the S1 gene: epitope D (residues 24 to 60-61) and epitope E (132 – 149), was also performed.

2.3 RESULTS AND DISCUSSION

2.3.1 IBV detection and genotyping

The detection of IBV in samples collected during a 5.5 years period, from flocks showing signs or lesions of the respiratory, renal, digestive or reproductive systems, is shown in table 1. Out of a total of 453 analyzed samples, 61.4% were typed as positive for IBV. The genotyping demonstrated that 75.9% of the IBV

positive samples were variant strains and as discussed in previous epidemiological surveys (VILLARREAL et al., 2010; CHACON et al., 2011; BALESTRIN et al., 2012) these variants were thought to be the primary reason for IBV outbreaks despite vaccination with live and/or inactivated Massachusetts strain vaccines.

IBV is known to be a recurrent problem for the entire poultry industry. The detection age range for broilers, breeders, layers and grandparents observed in this study is detailed in table 2. Older birds are known to develop a degree of resistance to IB; however the data of this study showed that broilers could remain susceptible to the disease throughout almost their entire life spans. IBV was detected in samples from birds aging from 1 to 49 days old presenting clinical signs that were compatible with the disease. Layers, breeders and grandparents were likely to develop the disease throughout their entire productive life cycle (Table 2). In these cases IBV was detected in birds presenting respiratory, reproductive and digestive lesions and clinical signs between 5 to 52 weeks of age. With a morbidity of almost 100%, these results confirm that IBV represents a constant challenge for animal health and that age cannot be considered an excluding factor for the development for disease development and agent detection.

Table 1 – Total number of samples processed between 2010 and July 2015

Year of Detection	Processed Samples	IBV Positive Samples		IBV Variants	
		Quantity	%[a]	Quantity	%[b]
2010	33	18	54.5	6	33.3
2011	19	14	73.7	13	92.9
2012	213	170	79.8	126	74.1
2013	60	32	53.3	26	81.3
2014	88	32	36.4	28	87.5
2015	40	12	30.0	12	100.0
Total	453	278	61.4	211	75.9

[a]Percentage of IBV positive samples of the total of samples recieved per year

[b]Percentage of IBV variants of the total of IBV positive samples per year

Source: (CARRANZA, 2015)

Vertical transmission of the virus to the embryo has never been reported but the virus may be present on the shell surface of hatching eggs via shedding from the oviduct or alimentary tract (CAVANAGH, 2007; SJAAK DE WIT; COOK; VAN DER

HEIJDEN, 2011). In 2014, were detected positive samples for variant IBV strains from intestines and livers of 1 day-old broilers (data not shown) these results suggest that chicks may have been contaminated during the incubation period.

Table 2 – Detection and genotyping of IBV per type of bird

Type of bird	Age range	IBV positive samples	Genotyping			
			Mass strains	%[a]	Variant strains	%[a]
Broilers	1 - 49 days	114	49	43	65	57
Breeders	5 - 52 weeks	66	4	6	62	94
Layers	6 - 49 weeks	27	2	7	25	93
Grandparents	26 - 50 weeks	4	0	0	4	100
Not specified*	-	67	12	18	55	82
		278	67	24.1[b]	211	75.9[b]

[a]Percentage of IBV genotype of the total of IBV positive samples per type of bird

[b] Percentage of IBV genotype of the total of IBV positive samples

Source: (CARRANZA, 2015)

IBV tissue tropism is diverse and our results confirmed that the virus can be detected in several organs belonging to different systems. These findings were similar to other epidemiological studies of Brazilian isolates, such as the results of Balestrin et al. (2014) who found IBV-positive samples primarily in the digestive, respiratory and urinary/reproductive systems. Additionally this study detectec that IBV variants was present in the bursa of Fabricius, sinus, conjunctive tissue and trigeminal ganglion (Table 3); in these 4 cases the flocks presented only respiratory clinical signs; the virus was also detected in samples from the respiratory tract (tracheas) belonging to the same flocks (data not shown). In 2012 Ganapathy et al. (2012) reported the isolation of a QX-like infectious bronchitis virus in a case of proventriculitis in broiler chickens. The inoculation of this strain in embryonated eggs caused renal and reproductive lesions; however the proventricular damage was not reproduced. In the present study no cases of proventriculitis were described and only two variant IBV-positive flocks (6 samples) presented clinical signs compatible with gastrointestinal disorders. These results suggest that the pathogenic effects may be restricted to the respiratory, renal and/or reproductive systems even when IBV has tropism for a diversity of tissues. These differences in pathogenicity and tropism may

be related to the different expression levels of sialic acids in the tissue surface of the different organs (KUMLIN et al., 2008; ABD EL RAHMAN et al., 2009).

Table 3 – Viral detection in the different organs sampled

Sample origin	Positive Samples	%	Genotyping IBV Strains			
			MASS Strains	%[a]	Variant Strains	%[a]
Trachea	73	26.3	12	16.4	61	83.6
Kidney	66	23.7	10	15.2	56	84.8
Cecal Tonsils	54	19.4	5	9.3	49	90.7
Lungs	22	7.9	4	18.2	18	81.8
Intestine	21	7.6	12	57.1	9	42.9
Pancreas	12	4.3	12	100.0	0	0.0
Spleen	9	3.2	6	66.7	3	33.3
Liver	9	3.2	4	44.4	5	55.6
Oviduct	6	2.2	0	0.0	6	100.0
Trigeminal Ganglion	2	0.7	0	0.0	2	100.0
Bursa	2	0.7	2	100.0	0	0.0
Sinus	1	0.4	0	0.0	1	100.0
Conjunctive Tissue	1	0.4	0	0.0	1	100.0
	278	100	67	24.1[b]	211	75.9[b]

[a] Percentage of IBV genotype of the total of IBV positive samples per organ

[b] Percentage of IBV genotype of the total of IBV positive samples

Source: (CARRANZA, 2015)

The geographical distribution of variant IBV-positive samples (not belonging to the Massachusetts genotype) by the Multiplex-PCR reaction can be clearly observed in figure 1. Although the number of samples of this study is limited (453 samples analyzed), the data represents outbreaks that occurred in the Brazilian territory between the years 2010 and 2015, therefore this results show the present distribution of the virus variants in the country. Previous studies describe the presence of IBV variants in central-west, south, southeast and northeast regions (CHACON et al., 2011; BALESTRIN et al., 2012; FRAGA et al., 2013). This data confirms for the first time the presence of IBV variants in the northern region, showing that the virus and its variants are widely disseminated (Table 4) and represents an important pathogenic agent and a constant challenge for the Brazilian poultry industry.

Table 4 – Geographic origin and genotype of IBV samples

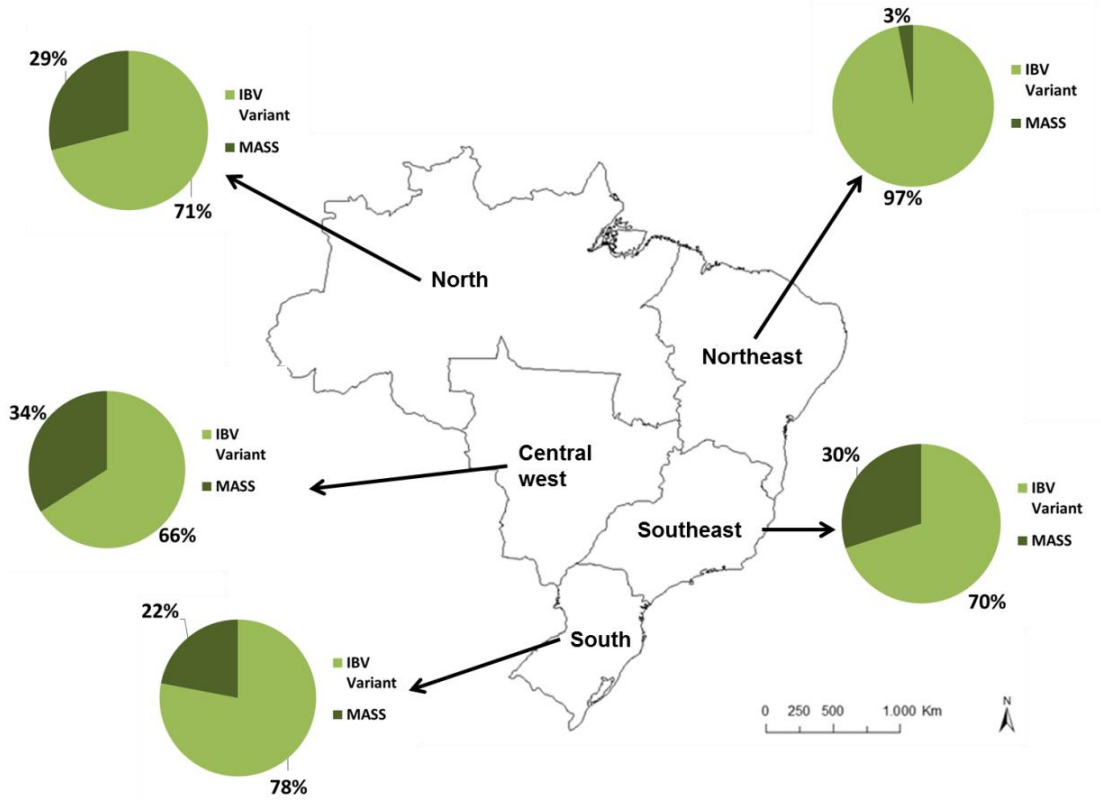
Geographic origin	IBV Positive samples	Genotyping			
		Mass Strains	%[a]	Variant Strains	%[a]
Central West	29	10	34	19	66
North East	29	1	3	28	97
North	7	2	29	5	71
South East	100	30	30	70	70
South	108	24	22	84	78
Not Specified*	5	0	0	5	100
	278	67	24.1[b]	211	75.9[b]

[a]Percentage of IBV genotype of the total of IBV positive samples per type of bird

[b] Percentage of IBV genotype of the total of IBV positive samples

Source: (CARRANZA, 2015)

Figure 1 – Distribution of IBV strains in the Brazilian territory



Source: (CARRANZA, 2015)

2.3.2 Partial S1 sequencing and phylogenetic analysis

A fragment of approximately 559 to 566 bp was obtained from 12 of the 39 Brazilian IBV isolates (30.8%) tested for sequencing, following a partial S1 gene amplification protocol designed in house. Amplified samples were collected from the southeast, central-west, northeast and northern Brazilian regions between the years 2011 to 2015, from broiler and layer flocks presenting mainly respiratory signs, and mortality. The low number of successfully amplified sequences may be due to RNA degradation during collection of samples, transport or processing. Considering that the optimal temperature during storage and transport of samples for a successful RNA isolation or amplification is 4ºC (KULIWABA; FAZZALARI; FINDLAY, 2005). The lack of adequate equipment and materials and the elevated temperatures in the field may have degraded the IBV RNA in some samples. Santos (2012) encountered the same difficulty when attempting to amplify a similar region of the S1 subunit of IBV directly from field samples, with only 15 sequences out of 58 isolates (25.9%) correctly amplified and sequenced. Other possible cause for an unsuccessful amplification can be a low viral concentration in the tissues. A common technique used to increase the viral load is the inoculation of the material into embryonated eggs. IBV successfully replicates regardless of the inoculation route, however the most favorable and commonly used technique is inoculation through the allantoic route, which results in extensive replication of the virus in the chorioallantoic membrane and high viral titers in the allantoic fluid (GUY, 2008).

The deduced amino acid deduced sequences were aligned and a phylogenetic tree was constructed that included 23 different IBV strains detected worldwide (Figure 2). The Brazilian isolates were divided into two clear unique clusters that were separated from strains detected outside of Brazil. Most of the isolates from this study were identified as belonging to the BR-I genotype (CHACON et al., 2011). This genotype appeared to be the most disseminated and predominant IBV variant causing disease in poultry, because our results and other epidemiological studies detected this genotype in different Brazilian regions (ABREU et al., 2010; FELIPPE et al., 2010) from birds with respiratory, renal, enteric or reproductive signs and mortality. Additionally only the isolate USP 555-6 (detected in tracheal tissue) clustered together with isolates belonging to the BR-II genotype

(Figure 2). This represents the first detection of this variant in the north region of the country, after Fraga *et al.* (2010) first description from caecal tonsil samples from the central-west region. This result suggests that despite the biosecurity measures and vaccination practices, this variant is slowly disseminating throughout the country. This analysis confirms the existence of two Brazilian variants circulating in all the poultry producing regions of Brazil, causing disease despite vaccination with the Massachusetts H120 strain.

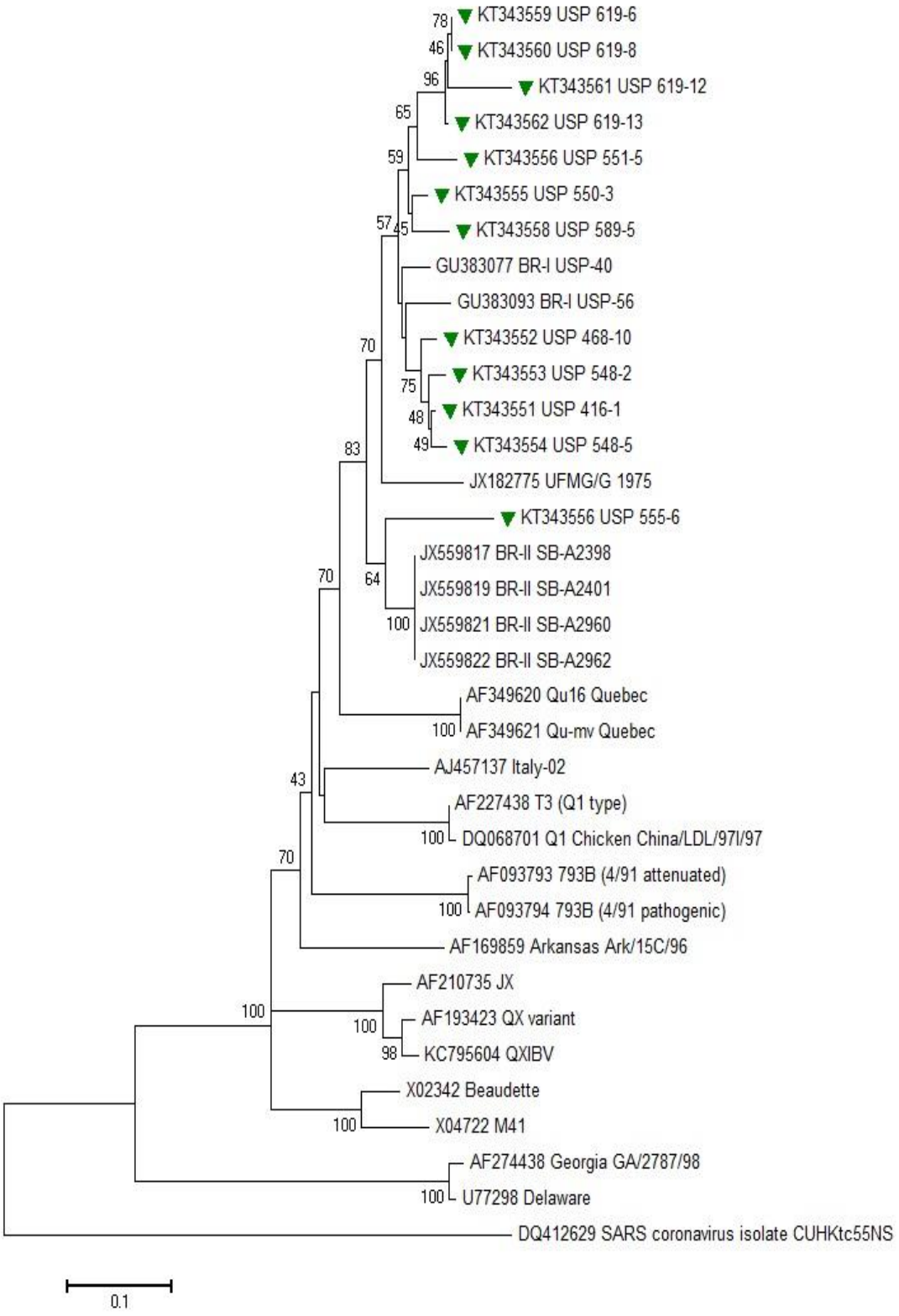
Table 5 - Epidemiological information from Brazilian IBV variants sequenced in the present study

Isolate Identification	Year Isolated	Geographical Origin (Brazil)	Type of Bird	Clinical Signs	Age	Vaccination	Genotype	GenBank Accession Number
USP 416-1	2011	Southeast	Layer	Respiratory signs	20 weeks	Not reported	BR-I	KT343551
USP 468-10	2012	Centralwest	Not reported	Not reported	Not reported	Not reported	BR-I	KT343552
USP 548-2	2014	Not reported	Broiler	Not reported	39 days	Not reported	BR-I	KT343553
USP 548-5	2014	Not reported	Broiler	Not reported	41 days	Not reported	BR-I	KT343554
USP 550-3	2014	Southeast	Layer	Runting stunting syndrome	12 weeks	Not reported	BR-I	KT343555
USP 551-5	2014	Northeast	Layer	Not reported	17 weeks	Yes	BR-I	KT343556
USP555-6	2014	North	Broiler	Mortality	39 days	No	BR-II	KT343557
USP 589-5	2014	Northeast	Layer	Respiratory signs	49 weeks	Not reported	BR-I	KT343558
USP 619-6	2015	Northeast	Broiler	Respiratory signs	42 days	Yes	BR-I	KT343559
USP 619-8	2015	Northeast	Broiler	Respiratory signs	42 days	Yes	BR-I	KT343560
USP 619-12	2015	Northeast	Broiler	Respiratory signs	42 days	Yes	BR-I	KT343561
USP 619-13	2015	Northeast	Broiler	Respiratory signs	42 days	Yes	BR-I	KT343562

Source: (CARRANZA, 2015)

GenBank accession numbers of the nucleotide data of the obtained sequences in this study (Table 5): USP 416-1 (KT343551), USP 468-10 (KT343552), USP 548-2 (KT343553), USP 548-5 (KT343554), USP 550-3 (KT343555), USP 551-5 (KT343556), USP 555-6 (KT343557), USP 589-5 (KT343558), USP 619-6 (KT343559), USP 619-8 (KT343560), USP 619-12 (KT343561) and USP 619-13 (KT343562).

Figure 2 – Phylogenetic analysis of partial S1 sequences of Brazilian and international IBV strains. Tree was generated by the Maximum Likelihood, using Jones-Taylor-Thorton amino acid substitution model with 1000 bootstrap repetitions.



Source: (CARRANZA, 2015)

2.3.3 Bioinformatics and evolution analyses

Coronaviruses possess the largest RNA genomes, ranging from 27 Kb to 32 Kb in length, (MASTERS, 2006; WOO et al., 2009). Having such a large and complex genome, coronavirus replication is error-prone. That leads to a high mutation rate (similar to those reported for Avian Influenza viruses), which may vary between 5.7×10^{-6} to 4×10^{-4} nucleotide substitutions per site per day (MCKINLEY et al., 2011; LICITRA et al., 2013). One of the regions more susceptible to mutations is the Spike protein gene which is comprised of subunits S1 and S2, and includes the hipervariable regions 1 and 2 (HVR1 and HVR2 respectively) (BRIAN; BARIC, 2005). This glycoprotein is the mediator of viral entry and receptor binding. Is also responsible for the viral fusion to the host's membrane and is a determinant of cell tropism (CAVANAGH, 2007b; WICKRAMASINGHE et al., 2014). In the specific case of Infectious Bronchitis Virus, Casais et al. (2003) showed that nucleotide substitutions as low as 5% in the S1 subunit diminish the protection conferred by vaccination, suggesting that the Spike protein also plays a role in the development of immune response and the induction of neutralizing antibodies (CASAIS et al., 2003; LIU et al., 2006). Therefore any change at this genome region —either by lack of proofreading ability or as an adaptive mechanism in response to the pressure associated with IBV intensive vaccination— may result in new viral variants capable of overcoming the immune response of the host and causing disease (SAIF, 2008).

To understand the evolution of Brazilian IBV variants, nucleotide analysis was performed with the 59 partial S1 sequences belonging to Brazilian IBV isolates collected within a time period of 40 years (1975 – 2015) and one sequence belonging to a Massachusetts vaccine strain H120.

The alignment showed that all IBV variant samples presented an insertion of 15 nucleotides at positions 430 – 444 compared with the Massachusetts strains, resulting in the addition of five amino acids at codons 144 – 148 of the S1 gene. Additionally insertion of the nucleotides TTA ACA at positions 355 – 360 was observed in all variant sequences with the except for USP 66 strain (collected in 2008 from Paraná, South of Brazil), which presented an insertion of TCG at positions 358 – 360. Among the most recent sequenced strains, USP 555-6 (that belongs to the BR-II genotype) showed a deletion of three nucleotides in the position 64 – 66 of the S1 gene, which was similar only to the 1983 isolate UFMG 283 (which clustered

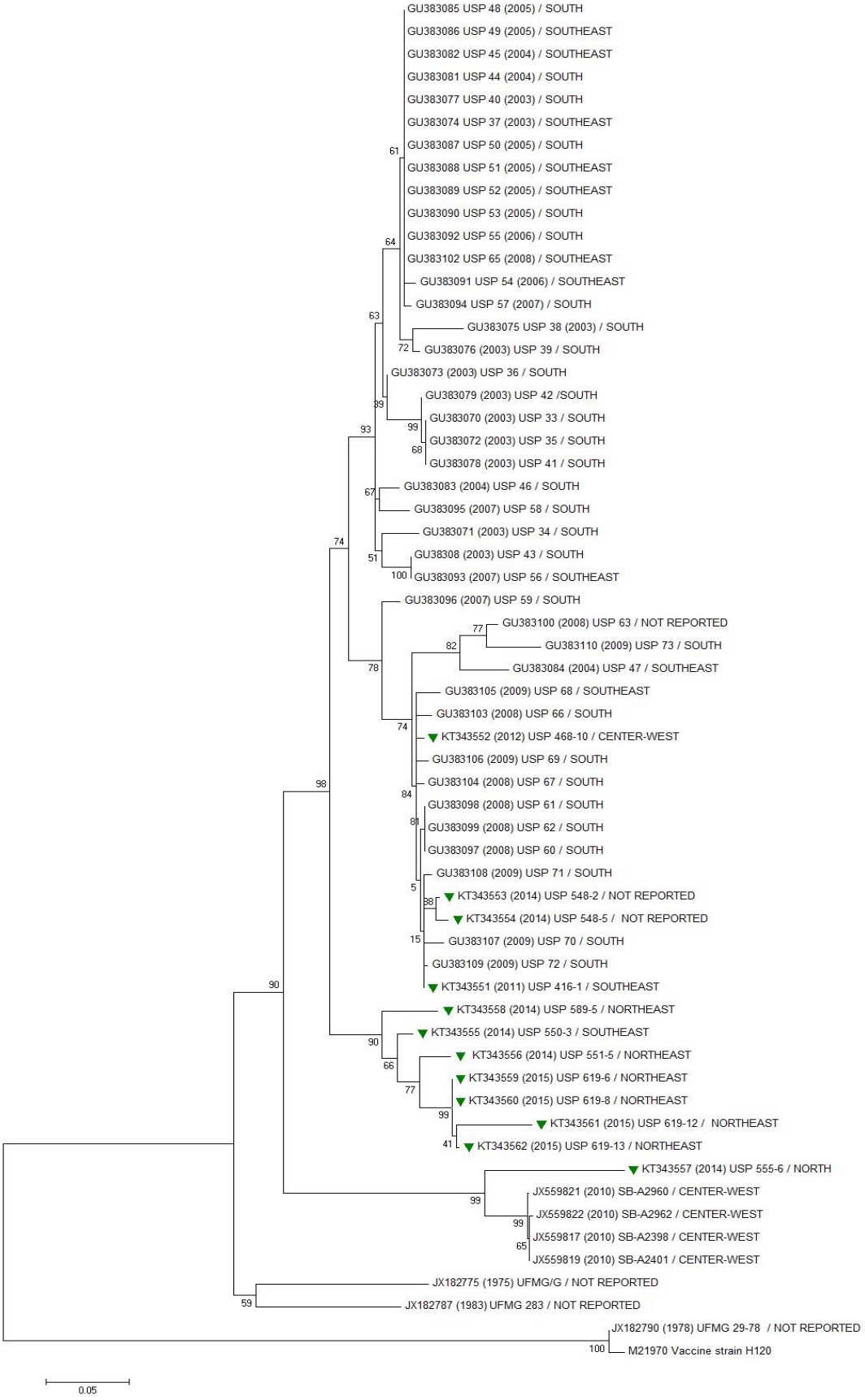
outside the BR-I and II genotypes but within the Brazilian variant group as shown in Figure 3). No other BR-II isolate had insertions or deletions that were similar to the newest detected BR-II isolate.

Brazilian IBV variants appear to be divided in 3 major groups (Figure 3). The first group comprises the older isolates, with variants that are no longer circulating in Brazil (UFMG/G and UFMG 283) and that share 86.5 – 86.9% of nucleotide and amino acid similarities (Table 6). The analyzed information about Brazilian variants suggests that changes in the S1 sequences occurred prior to the official introduction of the Massachusetts H120 vaccine in 1979 (SILVA, 2010) and 18 years after the first detection of IBV in Brazil (HIPÓLITO, 1957). Previous studies have detected the existence of IBV Brazilian variants by virus-neutralization tests in samples from 1995 from birds with respiratory, reproductive or digestive disorders (DI FABIO et al., 2000), however no molecular characterization was performed.

The second group comprises the BR-I genotype which is clearly divided in two sub-groups. The first sub-group comprises the oldest sequences of this genotype (from 2003 to early 2014), and the second sub-group includes the latest BR-I sequences obtained in this study (from the late 2014 to 2015). Abreu *et al.* (2010) suggested that dendograms were able to discriminate enteric from respiratory IBVs, however no differentiation was observed in the present dendograms analyzed probably due to the analysis of a fragment of the S1 subunit instead of the complete S gene.

The third group includes the genotype BR-II, first described in 2010 by *Fraga* et al. and includes one novel 2014 sample from the Northern region of Brazil. In general, the BR-I strains presented 62.8 – 75.2% nucleotide and amino acid similarities with the H120 vaccine sequence, whereas the BR-II strains presented 66.2 – 73.9%. These results are help understand what is happening in the field, where flocks are presenting signs compatible with IB disease despite being vaccinated with the Mass serotype (table 5). These results suggest that the vaccine may no longer promote an immune response that is sufficient to protect the flocks from the actual IBV field challenge.

Figure 3 - Phylogenetic tree generated by the Maximum Likelihood statistical method with 1000 bootstraps, using Tamura 3-parameter nucleotide substitution model.



Source: (CARRANZA, 2015)

Table 6 – Identity matrix of nucleotide and amino acid sequences from the partial S1 sequences of 40 years of Brazilian isolates

SEQUENCES		Percent aminoacid similarity																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 UFMG/G	1975		71.0	86.9	83.4	82.8	83.4	86.2	85.5	79.3	83.4	85.5	83.4	84.1	83.4	85.5	85.5	84.8	86.2	85.5	85.5	85.5
2 UFMG 29-78	1978	73.9		69.7	70.3	72.4	70.3	71.7	71.7	66.9	71.7	71.7	70.3	70.3	70.3	71.7	71.7	71.0	69.7	71.7	71.7	71.7
3 UFMG 283	1983	86.5	72.9		84.8	85.5	84.8	87.6	86.9	80.7	84.8	86.9	84.8	85.5	84.8	86.9	86.9	86.9	85.5	86.9	86.9	86.9
4 USP 33	2003	84.6	72.9	86.7		91.0	100.0	94.5	91.0	88.3	95.2	91.0	100.0	99.3	89.7	91.0	91.0	92.4	86.9	91.0	91.0	91.0
5 USP 34	2003	84.9	74.5	86.5	96.1		91.0	93.8	91.7	86.2	91.7	91.7	91.0	91.7	93.1	91.7	91.7	93.8	87.6	91.7	91.7	91.7
6 USP 35	2003	84.6	72.9	86.7	100.0	96.1		94.5	91.0	88.3	95.2	91.0	100.0	99.3	89.7	91.0	91.0	92.4	86.9	91.0	91.0	91.0
7 USP 36	2003	85.6	73.9	86.9	97.9	97.5	97.9		96.6	89.7	93.8	96.6	94.5	95.2	94.5	96.6	96.6	96.6	90.3	96.6	96.6	96.6
8 USP 37	2003	84.9	73.4	86.0	96.6	96.6	96.6	98.6		91.7	95.9	100.0	91.0	91.7	92.4	100.0	100.0	94.5	90.3	100.0	100.0	100.0
9 USP 38	2003	82.1	71.3	83.0	94.7	93.6	94.7	95.4	96.3		93.1	91.7	88.3	87.6	85.5	91.7	91.7	87.6	83.4	91.7	91.7	91.7
10 USP 39	2003	84.4	73.4	85.6	97.9	96.3	97.9	97.7	98.6	96.8		95.9	95.2	94.5	90.3	95.9	95.9	93.1	88.3	95.9	95.9	95.9
11 USP 40	2003	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6		91.0	91.7	92.4	100.0	100.0	94.5	90.3	100.0	100.0	100.0
12 USP 41	2003	84.6	72.9	86.7	100.0	96.1	100.0	97.9	96.6	94.7	97.9	96.6		99.3	89.7	91.0	91.0	92.4	86.9	91.0	91.0	91.0
13 USP 42	2003	84.9	72.9	86.9	99.8	96.3	99.8	98.2	96.8	94.5	97.7	96.8	99.8		90.3	91.7	91.7	93.1	87.6	91.7	91.7	91.7
14 USP 43	2003	84.9	73.2	85.8	95.9	96.8	95.9	97.5	96.6	93.3	96.1	96.6	95.9	96.1		92.4	92.4	95.2	91.7	92.4	92.4	92.4
15 USP 44	2004	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6		100.0	94.5	90.3	100.0	100.0	100.0
16 USP 45	2004	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0		94.5	90.3	100.0	100.0	100.0
17 USP 46	2004	84.9	72.9	86.2	96.6	96.8	96.6	98.2	97.2	94.0	96.8	97.2	96.6	96.8	97.2	97.2	97.2		91.0	94.5	94.5	94.5
18 USP 47	2004	84.4	73.9	85.6	90.8	92.0	90.8	92.4	92.7	89.7	92.2	92.7	90.8	91.1	93.1	92.7	92.7	92.7		90.3	90.3	90.3
19 USP 48	2004	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7		100.0	100.0
20 USP 49	2005	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0		100.0
21 USP 50	2005	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	
22 USP 51	2005	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	100.0
23 USP 52	2005	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	100.0
24 USP 53	2005	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	100.0
25 USP 54	2006	84.6	72.9	85.3	95.9	95.9	95.9	97.9	99.3	95.6	97.9	99.3	95.9	96.1	96.3	99.3	99.3	96.8	92.2	99.3	99.3	99.3
26 USP 55	2006	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	100.0
27 USP 56	2007	84.9	73.2	85.8	95.9	96.8	95.9	97.5	96.6	93.3	96.1	96.6	95.9	96.1	100.0	96.6	96.6	97.2	93.1	96.6	96.6	96.6
28 USP 57	2007	85.3	73.6	86.0	96.6	96.6	96.6	98.6	99.5	95.9	98.2	99.5	96.6	96.8	96.6	99.5	99.5	97.2	93.1	99.5	99.5	99.5
29 USP 58	2007	84.9	73.2	86.2	95.9	95.9	95.9	97.5	96.6	93.3	96.1	96.6	95.9	96.1	96.3	96.6	96.6	97.5	91.3	96.6	96.6	96.6
30 USP 59	2007	86.0	74.8	86.0	94.3	94.7	94.3	95.4	95.0	91.7	94.5	95.0	94.3	94.5	94.5	95.0	95.0	95.0	94.0	95.0	95.0	95.0
31 USP 60	2008	85.1	74.1	86.0	93.1	93.8	93.1	94.3	93.8	90.6	93.3	93.8	93.1	93.3	93.6	93.8	93.8	94.5	94.5	93.8	93.8	93.8
32 USP 61	2008	85.1	74.1	86.0	93.1	93.8	93.1	94.3	93.8	90.6	93.3	93.8	93.1	93.3	93.6	93.8	93.8	94.5	94.5	93.8	93.8	93.8
33 USP 62	2008	85.1	74.1	86.0	93.1	93.8	93.1	94.3	93.8	90.6	93.3	93.8	93.1	93.3	93.6	93.8	93.8	94.5	94.5	93.8	93.8	93.8
34 USP 63	2008	83.5	73.4	85.1	90.6	91.5	90.6	92.2	92.4	89.4	92.0	92.4	90.6	90.8	91.5	92.4	92.4	91.5	95.9	92.4	92.4	92.4
35 USP 65	2008	84.9	73.4	86.0	96.6	96.6	96.6	98.6	100.0	96.3	98.6	100.0	96.6	96.8	96.6	100.0	100.0	97.2	92.7	100.0	100.0	100.0
36 USP 66	2008	84.4	74.1	84.6	92.7	93.8	92.7	94.3	94.3	91.1	93.8	94.3	92.7	92.9	94.3	94.3	94.3	94.5	94.7	94.3	94.3	94.3
37 USP 67	2008	84.9	73.6	85.1	92.7	93.8	92.7	94.3	93.8	90.6	93.3	93.8	92.7	92.9	93.6	93.8	93.8	94.5	94.5	93.8	93.8	93.8
38 USP 68	2009	84.9	73.9	85.6	93.3	93.6	93.3	94.5	94.5	91.3	94.0	94.5	93.3	93.6	93.8	94.5	94.5	94.7	94.3	94.5	94.5	94.5
39 USP 69	2009	84.6	74.1	85.8	92.9	94.0	92.9	94.5	94.0	90.8	93.6	94.0	92.9	93.1	93.3	94.0	94.0	94.5	94.5	94.0	94.0	94.0
40 USP 70	2009	83.9	73.4	84.9	92.4	93.1	92.4	93.6	93.1	89.9	92.7	93.1	92.4	92.7	92.9	93.1	93.1	93.8	93.3	93.1	93.1	93.1
41 USP 71	2009	85.1	74.3	85.8	92.7	94.3	92.7	94.3	93.8	90.6	93.3	93.8	92.7	92.9	93.8	93.8	93.8	94.5	94.7	93.8	93.8	93.8
42 USP 72	2009	84.9	74.1	85.8	92.9	93.6	92.9	94.0	93.6	90.4	93.1	93.6	92.9	93.1	93.3	93.6	93.6	94.3	94.3	93.6	93.6	93.6
43 USP 73	2009	83.3	73.9	84.9	90.4	91.5	90.4	91.7	92.0	89.0	91.5	92.0	90.4	90.6	91.1	92.0	92.0	91.5	93.8	92.0	92.0	92.0
44 SB A2398	2010	80.7	71.8	82.6	86.7	85.8	86.7	86.5	86.2	83.5	86.7	86.2	86.7	86.5	85.8	86.2	86.2	86.5	86.2	86.2	86.2	86.2
45 SB A2401	2010	80.7	71.8	82.6	86.7	85.8	86.7	86.5	86.2	83.5	86.7	86.2	86.7	86.5	85.8	86.2	86.2	86.5	86.2	86.2	86.2	86.2
46 SB A2960	2010	80.7	71.6	82.6	86.7	85.8	86.7	86.5	86.2	83.5	86.7	86.2	86.7	86.5	85.8	86.2	86.2	86.5	86.2	86.2	86.2	86.2
47 SB A2962	2010	80.5	71.6	82.3	86.5	85.6	86.5	86.2	86.0	83.3	86.5	86.0	86.5	86.2	85.6	86.0	86.0	86.2	86.0	86.0	86.0	86.0
48 USP 416-1	2011	85.1	74.1	86.0	93.1	93.8	93.1	94.3	93.8	90.6	93.3	93.8	93.1	93.3	93.6	93.8	93.8	94.5	94.5	93.8	93.8	93.8
49 USP 468-10	2012	85.3	74.3	85.1	92.9	93.8	92.9	94.5	94.3	91.1	93.8	94.3	92.9	93.1	93.6	94.3	94.3	94.5	94.7	94.3	94.3	94.3
50 USP 548-2	2014	84.6	73.9	85.3	92.2	92.9	92.2	93.3	93.6	90.4	93.1	93.6	92.2	92.4	92.7	93.6	93.6	93.6	93.8	93.6	93.6	93.6
51 USP 548-5	2014	84.6	73.6	85.1	92.0	92.4	92.0	93.1	92.9	89.7	92.4	92.9	92.0	92.2	92.4	92.9	92.9	93.3	93.6	92.9	92.9	92.9
52 USP 550-3	2014	86.7	75.5	87.2	92.7	93.8	92.7	94.3	93.8	90.8	93.1	93.8	92.7	92.9	93.3	93.8	93.8	93.6	93.1	93.8	93.8	93.8
53 USP 551-5	2014	84.4	73.9	85.3	91.1	92.2	91.1	92.7	92.2	89.2	91.5	92.2	91.1	91.3	92.9	92.2	92.2	92.7	91.5	92.2	92.2	92.2
54 USP 555-6	2014	80.5	70.4	81.7	83.9	83.5	83.9	84.2	83.9	81.4	84.4	83.9	83.9	83.7	83.3	83.9	83.9	83.5	81.0	83.9	83.9	83.9
55 USP 589-5	2014	85.6	74.5	86.9	91.5	92.2	91.5	92.7	92.7	89.7	92.4	92.7	91.5	91.7	92.0	92.7	92.7	92.7	92.0	92.7	92.7	92.7
56 USP 619-6	2015	84.2	74.1	85.1	91.7	92.9	91.7	93.3	93.6	90.1	92.4	93.6	91.7	92.0	93.1	93.6	93.6	93.6	92.0	93.6	93.6	93.6
57 USP 619-8	2015	84.2	74.1	85.1	91.7	92.9	91.7	93.3	93.6	90.1	92.4	93.6	91.7	92.0	93.1	93.6	93.6	93.6	92.0	93.6	93.6	93.6
58 USP 619-12	2015	81.2	71.3	82.3	88.8	89.2	88.8	89.7	89.9	86.5	88.8	89.9	88.8	89.0	89.4	89.9	89.9	89.9	88.1	89.9	89.9	89.9
59 USP 619-13	2015	83.9	73.9	85.6	91.7	92.9	91.7	93.3	93.6	90.1	92.4	93.6	91.7	92.0	93.1	93.6	93.6	93.6	92.0	93.6	93.6	93.6
60 H120 vaccine	-	73.6	99.1	72.7	72.5	73.9	72.5	73.6	73.2	70.9	72.9	73.2	72.5	72.5	72.9	73.2	73.2	72.7	73.6	73.2	73.2	73.2

Percent nucleotide similarity

Percent aminoacid similarity																			
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
85.5	85.5	85.5	84.8	85.5	83.4	86.9	84.8	86.9	86.2	86.2	86.2	85.5	85.5	84.8	84.8	85.5	85.5	84.1	85.5
71.7	71.7	71.7	71.0	71.7	70.3	72.4	71.7	72.4	71.0	71.0	71.0	68.3	71.7	72.4	70.3	70.3	71.0	69.0	71.7
86.9	86.9	86.9	86.2	86.9	84.8	86.9	86.9	86.9	86.2	86.2	86.2	85.5	86.9	84.8	84.8	86.9	86.2	85.5	86.2
91.0	91.0	91.0	90.3	91.0	89.7	91.0	91.7	93.1	93.1	93.1	93.1	87.6	91.0	90.3	91.0	93.1	91.0	91.0	90.3
91.7	91.7	91.7	91.0	91.7	93.1	91.7	93.1	93.8	92.4	92.4	92.4	86.9	91.7	92.4	91.7	92.4	91.7	91.0	93.1
91.0	91.0	91.0	90.3	91.0	89.7	91.0	91.7	93.1	93.1	93.1	93.1	87.6	91.0	90.3	91.0	93.1	91.0	91.0	90.3
96.6	96.6	96.6	95.9	96.6	94.5	96.6	95.9	95.2	95.2	95.2	95.2	91.0	96.6	93.8	94.5	95.2	94.5	93.8	93.8
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
91.7	91.7	91.7	91.0	91.7	85.5	90.3	86.9	86.2	86.2	86.2	86.2	84.1	91.7	86.2	85.5	87.6	86.2	85.5	85.5
95.9	95.9	95.9	95.2	95.9	90.3	94.5	92.4	91.7	91.7	91.7	91.7	88.3	95.9	91.7	91.0	93.1	91.7	91.0	91.0
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
91.0	91.0	91.0	90.3	91.0	89.7	91.0	91.7	93.1	93.1	93.1	93.1	87.6	91.0	90.3	91.0	93.1	91.0	91.0	90.3
91.7	91.7	91.7	91.0	91.7	90.3	91.7	92.4	93.8	93.8	93.8	93.8	88.3	91.7	91.0	91.7	93.8	91.7	91.7	91.0
92.4	92.4	92.4	91.7	92.4	100.0	92.4	93.1	93.1	93.1	93.1	93.1	89.0	92.4	93.8	92.4	93.1	92.4	91.7	93.1
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
94.5	94.5	94.5	93.8	94.5	95.2	94.5	95.9	94.5	95.9	95.9	95.9	90.3	94.5	94.5	95.2	95.9	95.2	94.5	94.5
90.3	90.3	90.3	89.7	90.3	91.7	91.7	89.7	91.7	93.1	93.1	93.1	94.5	90.3	93.1	91.7	92.4	92.4	91.0	92.4
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
100.0	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
	100.0	100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
100.0		100.0	99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
100.0	100.0		99.3	100.0	92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
99.3	99.3	99.3		99.3	91.7	97.9	93.1	92.4	92.4	92.4	92.4	89.7	99.3	92.4	91.7	93.8	92.4	91.7	91.7
100.0	100.0	100.0	99.3		92.4	98.6	93.8	93.1	93.1	93.1	93.1	90.3	100.0	93.1	92.4	94.5	93.1	92.4	92.4
96.6	96.6	96.6	96.3	96.6		92.4	93.1	93.1	93.1	93.1	93.1	89.0	92.4	93.8	92.4	93.1	92.4	91.7	93.1
99.5	99.5	99.5	98.9	99.5	96.6		93.8	94.5	94.5	94.5	94.5	91.7	98.6	94.5	93.8	94.5	94.5	93.8	93.8
96.6	96.6	96.6	96.3	96.6	96.3	96.6		94.5	93.8	93.8	93.8	88.3	93.8	92.4	93.1	93.8	93.1	92.4	93.1
95.0	95.0	95.0	94.7	95.0	94.5	95.4	95.2		97.2	97.2	97.2	91.0	93.1	94.5	95.2	95.2	95.2	94.5	95.9
93.8	93.8	93.8	93.6	93.8	93.6	94.3	94.0	97.2		100.0	100.0	93.8	93.1	97.2	97.9	97.9	97.9	97.2	97.2
93.8	93.8	93.8	93.6	93.8	93.6	94.3	94.0	97.2	100.0		100.0	93.8	93.1	97.2	97.9	97.9	97.9	97.2	97.2
93.8	93.8	93.8	93.6	93.8	93.6	94.3	94.0	97.2	100.0	100.0		93.8	93.1	97.2	97.9	97.9	97.9	97.2	97.2
92.4	92.4	92.4	91.7	92.4	91.5	92.9	90.4	93.6	95.0	95.0	95.0		90.3	92.4	92.4	93.1	93.8	91.7	92.4
100.0	100.0	100.0	99.3	100.0	96.6	99.5	96.6	95.0	93.8	93.8	93.8	92.4		93.1	92.4	94.5	93.1	92.4	92.4
94.3	94.3	94.3	94.0	94.3	94.3	94.7	94.0	96.3	98.6	98.6	98.6	95.4	94.3		96.6	95.2	96.6	94.5	95.9
93.8	93.8	93.8	93.6	93.8	93.6	94.3	94.0	96.8	99.1	99.1	99.1	95.0	93.8	98.6		95.9	97.2	95.2	96.6
94.5	94.5	94.5	94.3	94.5	93.8	94.5	94.3	96.3	98.6	98.6	98.6	95.0	94.5	98.2	98.2		96.6	96.6	95.9
94.0	94.0	94.0	93.3	94.0	93.3	94.5	93.8	96.6	98.9	98.9	98.9	95.2	94.0	98.4	98.9	97.9		96.6	97.9
93.1	93.1	93.1	92.9	93.1	92.9	93.6	93.3	96.1	98.4	98.4	98.4	93.8	93.1	97.5	97.9	97.5	98.2		95.9
93.8	93.8	93.8	93.6	93.8	93.8	94.3	94.0	96.8	99.1	99.1	99.1	95.0	93.8	98.6	99.1	98.2	99.3	98.4	
93.6	93.6	93.6	93.3	93.6	93.3	94.0	94.3	97.0	99.3	99.3	99.3	94.7	93.6	98.4	98.9	98.4	99.1	98.6	99.3
92.0	92.0	92.0	91.3	92.0	91.1	92.4	90.4	93.1	95.0	95.0	95.0	96.6	92.0	94.7	94.7	95.0	95.0	94.7	94.7
86.2	86.2	86.2	85.6	86.2	85.8	86.5	85.3	86.9	88.5	88.5	88.5	87.2	86.2	87.8	88.1	88.5	88.5	87.8	88.1
86.2	86.2	86.2	85.6	86.2	85.8	86.5	85.3	86.9	88.5	88.5	88.5	87.2	86.2	87.8	88.1	88.5	88.5	87.8	88.1
86.2	86.2	86.2	85.6	86.2	85.8	86.5	85.3	86.9	88.5	88.5	88.5	87.2	86.2	87.8	88.1	88.5	88.5	87.8	88.1
86.0	86.0	86.0	85.3	86.0	85.6	86.2	85.1	86.7	88.3	88.3	88.3	86.9	86.0	87.6	87.8	88.3	88.3	88.1	87.8
93.8	93.8	93.8	93.6	93.8	93.6	94.3	94.0	97.2	99.5	99.5	99.5	95.0	93.8	98.6	99.1	98.6	99.3	98.9	99.5
94.3	94.3	94.3	94.0	94.3	93.6	94.7	94.0	96.8	99.1	99.1	99.1	95.9	94.3	99.1	99.1	98.2	98.9	97.9	99.1
93.6	93.6	93.6	93.3	93.6	92.7	93.8	93.1	96.3	98.6	98.6	98.6	94.5	93.6	98.2	98.2	97.7	98.4	97.9	98.6
92.9	92.9	92.9	92.7	92.9	92.4	93.1	92.9	96.3	98.2	98.2	98.2	93.6	92.9	97.2	97.7	97.2	97.9	97.5	98.2
93.8	93.8	93.8	93.1	93.8	93.3	93.8	92.9	93.3	93.1	93.1	93.1	91.7	93.8	92.4	92.9	93.3	93.1	92.0	93.1
92.2	92.2	92.2	92.0	92.2	92.9	92.2	92.0	91.7	91.3	91.3	91.3	89.9	92.2	90.6	91.1	91.5	90.8	90.6	91.3
83.9	83.9	83.9	83.3	83.9	83.3	83.7	82.6	83.0	81.2	81.2	81.2	81.2	83.9	81.4	81.2	81.2	81.7	81.0	81.2
92.7	92.7	92.7	92.2	92.7	92.0	92.7	92.2	92.0	91.7	91.7	91.7	89.9	92.7	91.5	91.5	92.2	91.5	91.3	91.7
93.6	93.6	93.6	93.1	93.6	93.1	93.6	92.2	92.2	91.7	91.7	91.7	90.1	93.6	91.3	91.5	92.0	91.5	91.1	91.7
93.6	93.6	93.6	93.1	93.6	93.1	93.6	92.2	92.2	91.7	91.7	91.7	90.1	93.6	91.3	91.5	92.0	91.5	91.1	91.7
89.9	89.9	89.9	89.4	89.9	89.4	89.9	88.5	88.3	87.8	87.8	87.8	86.2	89.9	87.4	87.6	88.1	87.6	87.2	87.8
93.6	93.6	93.6	93.1	93.6	93.1	93.6	92.2	92.2	91.7	91.7	91.7	90.1	93.6	91.3	91.5	92.0	91.5	91.1	91.7
73.2	73.2	73.2	72.7	73.2	72.9	73.4	72.9	74.5	73.9	73.9	73.9	73.2	73.2	73.9	73.4	73.6	73.9	73.2	74.1
Percent nucleotide similarity																			

Percent aminoacid similarity																		
42	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
85.5	83.4	82.1	82.1	82.1	81.4	86.2	86.2	84.8	85.5	86.2	84.1	77.9	86.9	84.1	84.1	80.0	84.1	70.3
71.0	67.6	69.0	69.0	69.0	68.3	71.0	71.7	71.0	70.3	71.0	69.0	66.2	70.3	67.6	67.6	63.4	67.6	97.9
86.2	84.8	83.4	83.4	83.4	82.8	86.9	84.8	85.5	86.2	86.2	84.1	80.0	87.6	84.8	84.8	80.7	84.8	69.0
91.7	86.9	87.6	87.6	87.6	86.9	92.4	91.0	90.3	91.7	89.0	86.9	79.3	88.3	86.9	86.9	82.8	87.6	69.7
91.0	86.9	85.5	85.5	85.5	84.8	91.7	92.4	91.0	90.3	90.3	87.6	80.0	89.0	87.6	87.6	81.4	88.3	71.7
91.7	86.9	87.6	87.6	87.6	86.9	92.4	91.0	90.3	91.7	89.0	86.9	79.3	88.3	86.9	86.9	82.8	87.6	69.7
93.8	91.0	88.3	88.3	88.3	87.6	94.5	94.5	92.4	93.8	93.1	91.0	80.7	91.7	91.0	91.0	84.8	91.7	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
85.5	84.1	82.1	82.1	82.1	81.4	86.2	86.2	86.2	86.2	85.5	84.1	75.2	84.1	83.4	83.4	77.2	84.1	66.2
91.0	88.3	88.3	88.3	88.3	87.6	91.7	91.7	91.7	91.7	90.3	88.3	80.0	89.7	88.3	88.3	82.1	89.0	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
91.7	86.9	87.6	87.6	87.6	86.9	92.4	91.0	90.3	91.7	89.0	86.9	79.3	88.3	86.9	86.9	82.8	87.6	69.7
92.4	87.6	88.3	88.3	88.3	87.6	93.1	91.7	91.0	92.4	89.7	87.6	80.0	89.0	87.6	87.6	83.4	88.3	69.7
91.7	89.0	86.2	86.2	86.2	85.5	92.4	92.4	91.0	91.0	91.7	90.3	79.3	89.7	90.3	90.3	84.1	91.0	69.7
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
94.5	90.3	87.6	87.6	87.6	86.9	95.2	94.5	93.1	93.8	91.7	91.0	79.3	91.0	91.0	91.0	84.8	91.7	70.3
92.4	91.0	86.2	86.2	86.2	85.5	93.1	91.7	91.7	91.7	92.4	89.0	75.9	89.7	89.7	89.7	84.1	90.3	69.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
91.7	89.7	87.6	87.6	87.6	86.9	92.4	92.4	92.4	92.4	91.7	89.7	79.3	90.3	91.0	91.0	84.8	91.7	70.3
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
91.7	89.0	86.2	86.2	86.2	85.5	92.4	92.4	91.0	91.0	91.7	90.3	79.3	89.7	90.3	90.3	84.1	91.0	69.7
93.8	91.7	89.0	89.0	89.0	88.3	94.5	94.5	93.8	93.8	92.4	90.3	79.3	91.0	91.7	91.7	85.5	92.4	71.7
93.8	89.0	86.2	86.2	86.2	85.5	93.1	92.4	91.0	92.4	91.0	89.0	78.6	91.7	89.0	89.0	82.8	89.7	71.0
95.9	89.7	88.3	88.3	88.3	87.6	96.6	95.9	95.9	95.9	92.4	89.0	79.3	90.3	89.0	89.0	83.4	89.7	71.7
98.6	92.4	89.7	89.7	89.7	89.0	99.3	97.2	97.2	97.2	91.7	89.7	78.6	90.3	89.7	89.7	84.1	90.3	70.3
98.6	92.4	89.7	89.7	89.7	89.0	99.3	97.2	97.2	97.2	91.7	89.7	78.6	90.3	89.7	89.7	84.1	90.3	70.3
98.6	92.4	89.7	89.7	89.7	89.0	99.3	97.2	97.2	97.2	91.7	89.7	78.6	90.3	89.7	89.7	84.1	90.3	70.3
93.1	93.8	86.9	86.9	86.9	86.2	93.8	93.8	92.4	91.7	91.0	88.3	76.6	89.0	89.0	89.0	83.4	89.7	67.6
92.4	90.3	88.3	88.3	88.3	87.6	93.1	93.1	93.1	93.1	92.4	90.3	80.0	91.0	91.7	91.7	85.5	92.4	71.0
95.9	91.0	88.3	88.3	88.3	87.6	96.6	97.2	95.9	94.5	91.0	88.3	78.6	89.7	88.3	88.3	82.8	89.0	71.7
96.6	91.0	88.3	88.3	88.3	87.6	97.2	96.6	95.2	95.2	91.0	88.3	77.9	89.7	88.3	88.3	82.8	89.0	69.7
97.2	92.4	89.7	89.7	89.7	89.0	97.9	95.2	95.9	95.9	92.4	91.0	78.6	91.7	91.0	91.0	85.5	91.7	69.7
97.9	91.7	89.7	89.7	89.7	89.0	98.6	96.6	96.6	96.6	91.7	89.0	78.6	90.3	89.0	89.0	83.4	89.7	70.3
97.2	93.1	88.3	88.3	88.3	89.0	97.9	94.5	95.9	95.9	89.7	89.0	77.9	89.0	89.0	89.0	82.8	89.7	68.3
97.2	91.0	88.3	88.3	88.3	87.6	97.9	96.6	96.6	95.9	92.4	89.0	77.9	90.3	89.0	89.0	83.4	89.7	71.0
	92.4	89.0	89.0	89.0	88.3	99.3	95.9	97.2	97.2	91.0	89.0	77.2	89.7	89.0	89.0	83.4	89.7	70.3
94.7		85.5	85.5	85.5	86.2	92.4	91.7	91.0	90.3	89.7	88.3	75.2	88.3	89.0	89.0	82.8	89.7	66.9
88.3	86.9		100.0	99.3	99.3	89.7	87.6	88.3	89.7	84.8	83.4	86.2	84.8	83.4	83.4	78.6	84.1	69.0
88.3	86.9	100.0		99.3	99.3	89.7	87.6	88.3	89.7	84.8	83.4	86.2	84.8	83.4	83.4	78.6	84.1	69.0
88.3	86.9	99.8	99.8		98.6	89.7	87.6	88.3	89.7	84.8	83.4	86.2	84.8	83.4	83.4	78.6	84.1	68.3
88.1	87.2	99.8	99.8	99.5		89.0	86.9	87.6	89.0	84.1	82.8	85.5	84.1	82.8	82.8	77.9	83.4	68.3
99.8	95.0	88.5	88.5	88.5	88.3		96.6	97.9	97.9	91.7	89.7	77.9	90.3	89.7	89.7	84.1	90.3	70.3
98.9	95.4	88.1	88.1	88.1	87.8	99.1		97.2	94.5	92.4	89.0	78.6	90.3	89.0	89.0	83.4	89.7	71.0
98.9	94.5	88.1	88.1	88.1	87.8	99.1	98.6		96.6	91.0	88.3	77.9	89.0	88.3	88.3	82.8	89.0	70.3
98.4	93.6	87.8	87.8	87.8	87.6	98.6	97.7	99.1		90.3	88.3	77.9	89.0	89.7	89.7	84.1	90.3	69.7
92.9	92.2	86.2	86.2	86.2	86.0	93.1	92.9	92.2	92.0		92.4	78.6	93.8	93.1	93.1	87.6	93.8	70.3
91.1	90.8	84.2	84.2	84.2	83.9	91.3	91.1	90.4	90.1	96.3		75.2	91.7	93.8	93.8	87.6	93.1	68.3
81.0	81.0	91.3	91.3	91.3	91.1	81.2	81.7	81.2	80.5	83.7	81.7		79.3	75.9	75.9	71.0	75.9	65.5
91.5	90.6	85.6	85.6	85.6	85.3	91.7	92.0	91.7	91.1	95.6	93.8	83.3		92.4	92.4	86.9	91.7	69.7
91.5	91.3	84.9	84.9	84.9	84.6	91.7	91.7	91.1	91.1	96.3	96.6	82.1	94.3		100.0	93.8	99.3	66.9
91.5	91.3	84.9	84.9	84.9	84.6	91.7	91.7	91.1	91.1	96.3	96.6	82.1	94.3	100.0		93.8	99.3	66.9
87.6	87.4	81.4	81.4	81.4	81.2	87.8	87.8	87.2	87.2	92.2	92.4	79.4	90.1	95.9	95.9		93.1	62.8
91.5	91.3	84.9	84.9	84.9	84.6	91.7	91.7	91.1	91.1	96.3	96.1	81.9	93.8	99.5	99.5	95.9		66.9
73.9	73.6	71.6	71.6	71.3	71.3	73.9	74.1	73.6	73.4	75.2	73.6	70.6	74.1	73.9	73.9	71.1	73.6	
Percent nucleotide similarity																		

Bioinformatics analyses were performed to predict N-glycosylation points and to detect synonymous and non-synonymous substitutions, to understand the biological consequences of the nucleotide mutations that occurred in the Brazilian isolates.

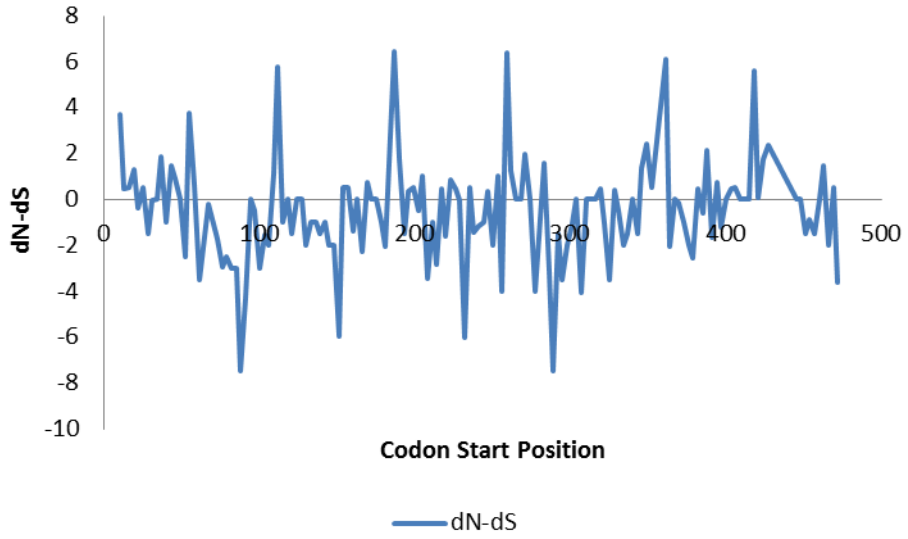
N-glycosylation is associated with cellular tropism and virulence (LI et al., 2000) and variation of the N-glycosylation sites of a viral protein affect the interaction with host receptors (SLATER-HANDSHY et al., 2004), In coronaviruses the N-glycosylation of spike protein and membrane proteins is involved in fusion, receptor binding and antigenic characteristics (DE HAAN et al., 2003; WISSING et al., 2004). Therefore any detected change in the N-glycosylation characteristics of the IBV sequences may help understand the evolution of this virus in Brazil.

The comparison of the estimated N-glycosylation sites between the isolates and the H120 vaccine showed that the Massachusetts H120 vaccine strain presented four N-glycosylation sites at the amino acids 48, 74, 100 and 141; all four sites were common N-glycosylation sites in all of the strains evaluated. The analysis showed that most of the BR-I variants (from 2003 to early 2014) presented an extra N-glycosylation site at position 20, which was very close of the start of antigenic site D (KANT et al., 1992) at amino acid 24. Surprisingly the newest 2014 isolates and all of the 2015 strains contained a deletion of the N-glycosylation site at the position 20 (this result corresponds with the dN-dS ratio results, where codon 20 appeared to have undergone non-synonymous substitutions). All four BR-II isolates from 2010 presented the same extra N-glycosylation site at position 20 and one extra site at position 138. This addition was located within antigenic site E. However, the BR-II isolate from late 2014 presented only the extra site at position 138, which corresponds with the loss of the N-glycosylation site at position 20 with the newest BR-I strains (from late 2014 to 2015). Small changes may alter the folding and conformation of the molecule (ABRO et al., 2012b). Therefore these extra features may influence the cross protection of the Mass vaccine when challenged in the field with the variant strains. In the case of the BR-II variants, these changes suggest that the N-glycosylation point in the 138 amino acid play a role in its antigenic characteristics, considering that the 2014 isolate was associated with mortality in broilers with 39 days of age.

To identify regions or specific positions with synonymous and non-synonymous substitutions within the first third of the S1 gene, the dN-dS statistical

test was used. Positive values indicated an overabundance of non-synonymous substitutions per site (n/N). Figure 4 plotted the dN-dS values obtained among the first 144 codons of the S1 gene. The positive and negative values were almost equally distributed among the first 144 amino acids of the S1 gene, in contrast with IBV variants from Spain and Sweden (DOLZ et al., 2008; ABRO et al., 2012a) where a predominance of codons with non-synonymous substitutions was found.

Figure 4 – dN-dS plotted per each amino acid position for the first 144 codons evaluated of the S1 gene of IBV



Source: (CARRANZA, 2015)

To infer the type of selective pressure over the entired analyzed region, the dN/dS ratio was calculated. This ratio was estimated to be 0.6757, indicating that the analyzed region was under negative selection. This result suggests that most of the mutations occurring in this part of the gene (including hypervariable regions 1 and 2) will not cause a change in the resulting amino acid when translated. This outcome is compatible with the phylogenetic and identity analyses, where both subgroups of the Brazilian variants (the BR-I and BR-II) have been maintained as the only two variants in the country and no other new genotype has yet emerged. This result may be due to the constant immunological pressure based on the approved used of only one vaccine serotype for prevention since 1980. Therefore, already mutated viruses have been successful in infecting and replicating within their hosts, and there is no need of

further dramatic nucleotide substitutions that could result in great amino acid changes. On the other hand a purifying selection may delete any changes that may affect the effective mechanism and the tropism of the IBV variant, as it may have happened with the loss of the N-glycosylation site 20. This entire Brazilian variant dynamic is expected to change with the recent approval and introduction of a vaccine based on the BR-I genotype.

Accordingly to Abro et al., (2012) strong positive selective pressure as a result of mutations in the S1 gene is most likely responsible for the genetic diversity and the appearance of new antigenic variants. The results shown in this study demonstrated that the IBV Brazilian variants likely suffered drastic mutations from 1983 to 2003. After achieving an antigenic structure sufficient for invasion and replication in their hosts, the selection processes became silent. Therefore the evolutionary strategy may have maintained the amino acid structure. With a divergence of 26.1 – 37.2% when compared with the Mass vaccine, these variant strains are no longer controlled by vaccination as many vaccinated flocks present the disease. This entire Brazilian variant evolutionary dynamic is expected to change with the recent approval and introduction of a vaccine based on the BR-I genotype. Further studies and epidemiological surveys should be performed to continue the surveillance of the evolution of the Brazilian IBV variants and to accurately define the most effective strategies for prevention and control.

REFERENCES

ABD EL RAHMAN, S.; EL-KENAWY, A.; NEUMANN, U.; HERRLER, G.; WINTER, C. Comparative analysis of the sialic acid binding activity and the tropism for the respiratory epithelium of four different strains of avian infectious bronchitis virus. **Avian Pathology**, v. 38, n. 1, p. 41–45, 2009.

ABREU, J.; MOURÃO, M.; SANTOS, C.; VELOSO, C.; RESENDE, J.; FLATSCHAR, R.; FOLGERAS-FLATSCHART, A.; JÚNIOR, S.; SANTORO, M.; MENDES, A.; FRANCO, G.; SILVA, A.; CAMPOS, A.; FERNANDEZ, S. Molecular studies of the Brazilian infectious bronchitis virus isolates. **Brazilian Journal of Poultry Science**, v. 12, n. 2, p. 107–110, 2010.

ABRO, S. H.; RENSTRÖM, L. H. M.; ULLMAN, K.; ISAKSSON, M.; ZOHARI, S.; JANSSON, D. S.; BELÁK, S.; BAULE, C. Emergence of novel strains of avian infectious bronchitis virus in Sweden. **Veterinary Microbiology**, v. 155, p. 237–246, 2012a.

ABRO, S. H.; ULLMAN, K.; BELÁK, S.; BAULE, C. Bioinformatics and evolutionary insight on the spike glycoprotein gene of QX-like and Massachusetts strains of infectious bronchitis virus. **Virology Journal**, v. 9, n. 211, 2012b.

BALESTRIN, E.; FRAGA, A. P.; IKUTA, N.; CANAL, C. W.; FONSECA, S. K.; LUNGE, V. R. Infectious bronchitis virus in different avian physiological systems — A field study in Brazilian poultry flocks. **Poultry Science**, v. 93, p. 1922–1929, 2012.

BRIAN, D. A.; BARIC, R. Coronavirus genome structure and replication. **Current Topics in Microbiology and Immunology**, v. 287, p. 1–30, 2005.

CASAIS, R.; DOVE, B.; CAVANAGH, D.; BRITTON, P.; CASAIS, R.; DOVE, B.; CAVANAGH, D.; BRITTON, P. Recombinant avian infectious bronchitis virus expressing a heterologoussSpike gene demonstrates that the spikep protein Is a determinant of cell tropism. **Journal of Virology**, v. 77, n. 16, p. 9084–9089, 2003.

CAVANAG, D.; Review article Coronavirus avian infectious bronchitis virus. **Veterinary Research**, v. 38, p. 281–297, 2007.

CAVANAGH, D.; MAWDITT, K.; BRITTON, P.; NAYLOR, C. J. Longitudinal field studies of infectious bronchitis virus and avian pneumovirus in broilers using type-specific polymerase chain reactions. **Avian Pathology**, v. 28, n. 6, p. 593–605, 1999.

CAVANAGH, D.; MAWDITT, K.; WELCHMAN, D. D. B.; BRITTON, P.; GOUGH, R. E. Coronaviruses from pheasants (Phasianus colchicus) are genetically closely related to coronaviruses of domestic fowl (infectious bronchitis virus) and turkeys. **Avian Pathology**, v. 31, n. 1, p. 81–93, 2002.

CHACON, J. L.; RODRIGUES, J. N.; ASSAYAG JUNIOR, M. S.; PELOSO, C.; PEDROSO, A. C.; FERREIRA, A. J. P. Epidemiological survey and molecular characterization of avian infectious bronchitis virus in Brazil between 2003 and 2009. **Avian Pathology**, v. 40, n. 2, p. 153–162, 2011.

DE GROOT, R. J. De; ZIEBUHR, J.; POON, L. L.; WOO, P. C.; ROTTIER, P. J. M.; HOLMES, K. V; GORBALENYA, A. E. Coronavirinae. In: KING, A. M. Q.; ADAMS, M. J.; CARSTENS, E. B.; LEFKOWITZ, E. J. (Ed.). **Virus taxonomy: eighth report of the international committee on taxonomy of viruses**. 9. ed. San Diego: Elsevier Inc., 2012. p. 806–828.

DE GROOT, R. J.; BAKER, S. C.; BARIC, R. S.; BROWN, C. S.; DROSTEN, C.; ENJUANES, L.; FOUCHIER, R. a M.; GALIANO, M.; GORBALENYA, A. E.; MEMISH, Z.; PERLMAN, S.; POON, L. L. M.; SNIJDER, E. J.; STEPHENS, G. M.; WOO, P. C. Y.; ZAKI, A. M.; ZAMBON, M.; ZIEBUHR, J. Middle East respiratory syndrome coronavirus (MERS-CoV); Announcement of the Coronavirus Study Group. **Journal of virology**, v. 5, p. 13–15, 2013

DE HAAN, C. A.; DE WIT, M.; KUO, L.; MONTALTO-MORRISON, C.; HAAGMANS, B. L.; WEISS, S. R.; MASTERS, P. S.; ROTTIER, P. J. The glycosylation status of the murine hepatitis coronavirus M protein affects the interferogenic capacity of the virus in vitro and its ability to replicate in the liver but not the brain. **Virology**, v. 312, n. 2, p. 395–406, 2003.

DE WIT, J. J. Detection of infectious bronchitis virus. **Avian Pathology**, v. 29, n. 2, p. 71–93, 2000.

DOLZ, R.; PUJOLS, J.; ORDÓÑEZ, G.; PORTA, R.; MAJÓ, N. Molecular epidemiology and evolution of avian infectious bronchitis virus in Spain over a fourteen-year period. **Virology**, v. 374, p. 50–59, 2008.

FELIPPE, P. a N.; DA SILVA, L. H. a; SANTOS, M. M. a B.; SPILKI, F. R.; ARNS, C. W. Genetic diversity of avian infectious bronchitis virus isolated from domestic chicken flocks and coronaviruses from feral pigeons in Brazil between 2003 and 2009. **Avian Diseases**, v. 54, n. 4, p. 1191–1196, 2010.

FRAGA, A. P.; BALESTRIN, E.; IKUTA, N.; FONSECA, A. S. K.; SPILKI, F. R.; CANAL, C. W.; LUNGE, V. R. Emergence of a new genotype of avian infectious bronchitis virus in Brazil. **Avian Diseases**, v. 57, n. 2, p. 225–32, 2013.

GANAPATHY, K.; WILKINS, M.; FORRESTER, A.; LEMIERE, S.; CSEREP, T.; MCMULLIN, P.; JONES, R. C. QX-like infectious bronchitis virus isolated from proventriculitis in commercial broilers in England. **Veterinary Record**, v. 171, p. 597, 2012.

GONZÁLEZ, J. M.; GOMEZ-PUERTAS, P.; CAVANAGH, D.; GORBALENYA, a. E.; ENJUANES, L. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. **Archives of Virology**, v. 148, n. 11, p. 2207–2235, 2003.

GUY, J. S. Isolation and propagation of coronaviruses in embryonated eggs. **Methods in Molecular Biology**, v. 454, p. 109–17, 2008.

HIPÓLITO, O. Isolamento e identificação do vírus da bronquite infecciosa das galinhas no Brasil. **Arquivo Escola Veterinaria Minas Gerais**, v. x, p. 131–163, 1957.

HYUK MOO KWON; JACKWOOD, M. W. Molecular cloning and sequence comparison of the S1 glycoprotein of the Gray and JMK strains of avian infectious bronchitis virus. **Virus Genes**, v. 9, n. 3, p. 219–229, 1995.

KANT, A.; KOCH, G.; VAN ROOZELAAR, D. J.; KUSTERS, J. G.; POELWIJK, F. a J.; VAN DER ZEIJST, B. a M. Location of antigenic sites defined by neutralizing monoclonal

antibodies on the S1 avian infectious bronchitis virus glycopolypeptide. **Journal of General Virology**, v. 73, n. 3, p. 591–596, 1992.

KULIWABA, J. S.; FAZZALARI, N. L.; FINDLAY, D. Stability of RNA isolated from human trabecular bone at post-mortem and surgery. **Biochimica et Biophysica ACTA Molecular Basis of Disease**, v. 1740, n. 1, p. 1–11, 2005.

KUMLIN, U.; OLOFSSON, S.; DIMOCK, K.; AMBERG, N. Sialic acid tissue distribution and influenza virus tropism. **Influenza and Other Respiratory Viruses**, v. 2, n. 5, p. 147–154, 2008.

LI, K.; SCHULER, T.; CHEN, Z.; GLASS, G. E.; CHILDS, J. E.; PLAGEMANN, P. G. Isolation nof lactate dehydrogenase-elevating viruses from wild house mice and their biological and molecular characterizatio. **Virus Research**, v. 67, n. 2, p. 153–162, 2000.

LICITRA, B. N.; MILLET, J. K.; REGAN, A. D.; HAMILTON, B. S.; RINALDI, V. D.; DUHAMEL, G. E.; WHITTAKER, G. R. Mutation in spike protein cleavage site and pathogenesis of feline coronavirus. **Emerging Infectious Diseases**, v. 19, n. 7, p. 1066–1073, 2013.

LIU, S. W.; ZHANG, Q. X.; CHEN, J. D.; HAN, Z. X.; LIU, X.; FENG, L.; SHAO, Y. H.; RONG, J. G.; KONG, X. G.; TONG, G. Z. Genetic diversity of avian infectious bronchitis coronavirus strains isolated in China between 1995 and 2004. **Archives of Virology**, v. 151, n. 6, p. 1133–1148, 2006.

MASTERS, P. S. The molecular biology of coronaviruses. **Advances in Virus Research**, v. 65, n. 06, p. 193–292, 2006.

MASTERS, P. S.; PERLMAN, S. Coronaviridae. In: KNIPE, D. M.; HOWLEY, P. M. (Ed.). **Fields Virology**. 6th ed. Alphen aan den Rijn: Wolters Kluwer, 2013. p. 825–859..

MCKINLEY, E. T.; JACKWOOD, M. W.; HILT, D. a.; KISSINGER, J. C.; ROBERTSON, J. S.; LEMKE, C.; PATERSON, A. H. Attenuated live vaccine usage affects accurate measures of virus diversity and mutation rates in avian coronavirus infectious bronchitis virus. **Virus Research**, v. 158, n. 1-2, p. 225–234, 2011.

NEUMAN, B. W.; CHAMBERLAIN, P.; BOWDEN, F.; JOSEPH, J. Atlas of coronavirus replicase structure. **Virus Research**, v. 194, p. 49–66, 2014.

POSADA, D. jModellTest: phylogenetic model averaging. **Molecular Biology and Evolution**, v. 25, n. 7, p. 1253–1256, 2008.

SANTOS, S., **Caracterização e comparação molecular de estirpes de referência e de campo do vírusda bronquite infecciosa das galinhas.** 2012. 63 f. Tese (Mestre em Ciências) - Faculdade de Medicina Veterinária, Universidade de São Paulo. 2012

SCHALK, A. F.; HAWN, M. C. An apparently new respiratory disease of baby chicks. **Journal of American Veterinary Medial Association**, v. 78, p. 413 – 422, 1931.

SILVA, E. Infectious bronchitis in Brazilian chickens: current data and observations of field service personnel. **Revista Brasileira de Ciência Avícola**, v. 12, n. 3, p. 197–203, 2010.

SJAAK DE WIT, J. J.; COOK, J. K. a; VAN DER HEIJDEN, H. M. J. F. Infectious bronchitis

virus variants: a review of the history, current situation and control measures. **Avian Pathology**, v. 40, n. 3, p. 223–235, 2011.

SLATER-HANDSHY, T.; DROLL, D. A.; FAN, X.; DI BISCEGLIE, A. M.; CHAMBERS, T. J. HCV E2 glycoprotein: mutagenesis of N-linked glycosylation sites and its effects on E2 expression and processing. **Virology**, v. 319, n. 1, p. 36–48, 2004.

TAMURA, K.; STECHER, G.; PETERSON, D.; FILIPSKI, A.; KUMAR, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.

THOMPSON, J. D.; PLEWNIAK, F.; POCH, O. A comprehensive comparison of multiple sequence alignment programs. **Nucleic Acids Research**, v. 27, n. 13, p. 2682–2690, 1999.

TORRES, C. A.; VILLARREAL, L. Y. B.; AYRES, G. R. R.; RICHTZENHAIN, L. J.; BRANDA, P. E. An avian coronavirus in quail with respiratory and reproductive signs. **Avian Diseases**, v. 57, p. 295–299, 2013.

VILLARREAL, L. Y. B.; SANDRI, T. L.; SOUZA, S. P.; RICHTZENHAIN, L. J.; DE WIT, J. J.; BRANDAO, P. E. Molecular epidemiology of avian infectious bronchitis in Brazil from 2007 to 2008 in breeders, broilers, and layers. **Avian Diseases**, v. 54, n. 2, p. 894–898, 2010.

WISSING, E. H.; KROESE, M. V.; MANESCHIJN-BONSING, J. G.; MEULENBERG, J. J.; VAN RIJN, P. A.; RIJSEWIJK, F. A.; ROTTIER, P. J. Significance of the oligosaccharides of the porcine reproductive and respiratory syndrome virus glycoproteins GP2a and GP5 for infectious virus production. **Journal of General Virology**, v. 85, n. 12, p. 3715–3723, 2004.

WOO, P. C. Y.; LAU, S. K. P.; LAM, C. S. F.; LAI, K. K. Y.; HUANG, Y.; LEE, P.; LUK, G. S. M.; DYRTING, K. C.; CHAN, K.-H.; YUEN, K.-Y. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. **Journal of Virology**, v. 83, n. 2, p. 908–917, 2009.

ZIEBUHR, J. The coronavirus replicase. **Current Topics in Microbiology and Immunology**, v. 287, p. 57–94, 2005.