

GISELLE AYRES RAZERA ROSSA

**Diversidade molecular dos genes codificadores das proteínas  
não-estruturais Nsp2 e protease Papáína-*like* e da proteína estrutural  
S1 de amostras brasileiras de *Coronavirus aviário***



São Paulo  
2014

**GISELLE AYRES RAZERA ROSSA**

**Diversidade molecular dos genes codificadores das proteínas não estruturais Nsp2 e protease Papáína-like e da proteína estrutural S1 de amostras brasileiras de *Coronavirus aviário***

Tese 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 obtenção do título de Doutor em Ciências

**Departamento:**

Medicina Veterinária Preventiva e Saúde Animal

**Área de Concentração:**

Epidemiologia Experimental Aplicada às Zoonoses

**Orientador:**

Prof. Dr. Paulo Eduardo Brandão

São Paulo

2014

Autorizo a reprodução parcial ou total desta obra, para fins acadêmicos, desde que citada a fonte.

#### DADOS INTERNACIONAIS DE CATALOGAÇÃO-NA-PUBLICAÇÃO

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

T.3005  
FMVZ

Rossa, Giselle Ayres Razera

Diversidade molecular dos genes codificadores das proteínas não-estruturais Nsp2 e protease Papaina-like e da proteína estrutural S1 de amostras brasileiras do *Coronavirus aviário* / Giselle Ayres Razera Rossa. -- 2014.  
76 f. : il.

Tese (Doutorado) - Universidade de São Paulo. Faculdade de Medicina Veterinária e Zootecnia. Departamento de Medicina Veterinária Preventiva e Saúde Animal, São Paulo, 2014.

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

Área de concentração: Epidemiologia Experimental Aplicada às Zoonoses.

Orientador: Prof. Dr. Paulo Eduardo Brandão.

1. Coronavirus aviário. 2. Vírus da bronquite infecciosa das galinhas. 3. Diversidade molecular. 4. Nsp2. 5. Nsp3. I. Título.

UNIVERSIDADE DE SÃO PAULO



FACULDADE DE MEDICINA VETERINÁRIA E ZOOTECNIA

Comissão de Ética no Uso de Animais

## CERTIFICADO

Certificamos que o Projeto intitulado "Diversidade molecular de genes codificadores das proteínas não-estruturais nsp2 e protease papaína-like e da proteína estrutural S1 de amostras brasileiras coronavírus aviário", protocolado sob o nº 2174/2011, utilizando amostras biológicas e de conteúdo entérico de 100 (cem) galináceos, sob a responsabilidade do Prof. Dr. Paulo Eduardo Brandão, está de acordo com os princípios éticos de experimentação animal da "Comissão de Ética no uso de animais" da Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo e foi aprovado em reunião de 04/5/2011.

We certify that the Research "Molecular diversity of nsp2 and papain-like protease non-structural protein genes and S1 structural protein gene, from Brazilian strains of avian coronavirus", protocol number 2174/2011, utilizing biological samples and enteric contents of 100 (one hundred) domestic fowl, under the responsibility Prof. Dr. Paulo Eduardo Brandão, agree with Ethical Principles in Animal Research adopted by "Ethic Committee in the use of animals" of the School of Veterinary Medicine and Animal Science of University of São Paulo and was approved in the meeting of day 05/04/2011.

São Paulo, 15 de julho de 2014.

Denise Tabacchi Fantoni  
Presidente

Av. Prof. Dr. Orlando Marques de Paiva, nº87  
Cidade Universitária "Armando de Salles Oliveira"  
São Paulo/SP – Brasil  
05508-270

Fone: + 55 11 3091-7671/7676/0904  
Fax: +55 11 3032-2224  
E-mail: ceuavet@usp.br  
<http://www.fmz.usp.br>

## FOLHA DE AVALIAÇÃO

Nome: AYRES, Giselle Razera Rossa

Título: Diversidade molecular dos genes codificadores das proteínas não estruturais Nsp2 e protease Papaína-like e da proteína estrutural S1 de amostras brasileiras de *Coronavirus aviário*

Tese 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 obtenção do título de Doutor em Ciências

Data: \_\_\_\_/\_\_\_\_/\_\_\_\_

### Banca Examinadora

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_ Julgamento: \_\_\_\_\_

*Vejo melhor os rios quando vou contigo,  
Pelos campos até à beira dos rios;  
Sentado a teu lado reparando nas nuvens  
Reparo nelas melhor –  
Tu não me tiraste a Natureza...  
Tu mudaste a Natureza...  
Trouxeste-me a Natureza para o pé de mim,  
Por tu existires vejo-a melhor, mas a mesma,  
Por tu me amares, amo-a do mesmo modo, mas mais,  
Por tu me escolheres para te ter e te amar,  
Os meus olhos fitaram-na mais demoradamente  
Sobre todas as cousas.  
Não me arrependo do que fui outrora  
Porque ainda o sou.*

*De O Pastor Amoroso – Alberto Caeiro  
(Fernando Pessoa)*

*Ao Guto,  
Meu grande amor,  
Companheiro de todas as horas,  
Meu melhor amigo,  
Dedico essa nossa realização.*

## **AGRADECIMENTOS ESPECIAIS**

À Deus, por todas as oportunidades que tive em minha vida e por ter saúde para aproveitá-las.

Ao meu marido, Luis Augusto Ferreira Rossa, por acreditar, por vezes mais do que eu, na realização dos meus sonhos; por me apoiar, compreender, incentivar, patrocinar e, principalmente, por me amar.

À minha mãe, Miriam Cristina de C. M. Ayres, pela presença e amor incondicionais.

Ao meu pai, Edevaldo Razera, por toda a torcida e amor.

Aos meus irmãos, Caroline Ayres Razera e Bruno Ayres Razera, pela paciência e amor.

À toda a minha querida família: os Ayres, os Razera e os Rossa, por esse grande amor e confusão que vivemos.

Ao Prof. Dr. José Ricardo de C. M. Ayres, meu querido tio, por todos os sábios conselhos (de quem está há mais tempo nesse caminho de incertezas) e, especialmente, por todo o apoio e amor.

Ao Prof. Dr. Paulo Eduardo Brandão,  
pela oportunidade, confiança e paciência.  
Por tudo que me ensinou e ainda ensinará.

## AGRADECIMENTOS

Agradeço a todos os que tenham colaborado para a execução desse trabalho, ou que tenham feito parte da minha vida nessa trajetória, com os quais eu possa ser injusta por não mencionar nesse momento.

Ao Prof. Dr. Ed. Hoffmann Madureira, Prof. Mario Binelli, Profa. Dra. Cláudia Bertan e Dr. Luis Augusto Ferreira Rossa, pela iniciação na carreira científica.

À Dra. Laura Villarreal, pela amizade, pelos conhecimentos compartilhados e pela indicação para cursar o mestrado sob orientação do Prof. Dr. Paulo Eduardo Brandão, abrindo o caminho para o presente trabalho.

Ao Prof. Dr. Leonardo José Richtzenhain, por me inspirar a ingressar na Medicina Veterinária Preventiva e a buscar a pesquisa e docência como carreira. Agradeço, especialmente, pelo agradável e enriquecedor convívio.

Ao Prof. Dr. Fabio Gregori, pela permanente disposição e disponibilidade para ajudar e sanar minhas dúvidas.

Ao Prof. Dr. Ricardo Augusto Dias, pela amizade, por seus valiosos conselhos e pelo auxílio imensurável, sem o qual o depósito da tese não teria sido possível.

Ao Prof. Dr. Ferando Ferreira e à Profa. Dra. Solange Maria Gennari pela dedicação ao programa de pós-graduação e apoio pessoal nas adversidades.

A todos os professores do VPS, pelo carinho e contribuição diária na minha formação pessoal e profissional.

Aos amigos dentre os funcionários do departamento, pela amizade e auxílio permanentes: Danival, Gisele, Sandra, Alexandre, Cristina, Virgínia, Carol, Rose, Cláudia, Priscilla, Eveline, Zenaide, Hilda, Pedro, Renato, Lilian, Sueli, Antônio, Jucélia, Bispo, Washington...

Agradeço especialmente ao amigo Danival, por todo o suporte dado nos momentos desafiadores durante essa jornada da pós-graduação, em especial para o depósito da tese.

A Carolina por todo o suporte à execução inicial desse trabalho, pela triagem das amostras positivas para o ACoV, tornando sua execução possível, mas especialmente, agradeço pela amizade.

As minhas irmãs: Sheila, Cintia, Sueli, Juliana, Iracema e Claudia, por tudo o que me ensinam e pelo amor verdadeiro, em todas as horas.

A nossa querida equipe do LABMAS por tornarem o dia-a-dia mais instrutivo e divertido: Nara, Laila, Fernanda, Bruna, Ekaterina, Karen, Patrícia, Luis, André, Ana Marcia e todos aqueles que passaram pelo laboratório.

Aos grandes amigos que encontrei em outros laboratórios, parceiros de todas as horas, queridos: Herbert, Marcos, Vasco, Roberta, Bárbara, Paula, Mariana, Bruno, Danilo, Thiago, Jonas, Amália, Júlia, Chico, Andréa, Igor, Sérgio, Vanessa, Aline, Ana Júlia, Raul, Fernando, Guilherme, Amane, Patrícia, Rui... Felizmente muitos!

Aos amigos de outros tempos, por permanecem na minha vida: Malheiros, Haila, Cintia, Francismar e Artemir.

À companheira de todas as horas, desde o primeira dia da graduação, Flavia, pela amizade verdadeira, presença constante (mesmo a distância) e apoio nos momentos difíceis.

A todos do VPS-FMVZ-USP, pela ajuda na realização desse trabalho.

Aos funcionários da Biblioteca Virginie Buff D' Ápice pelo carinho com o qual sempre me receberam e auxiliaram, sempre que precisei, desde a graduação. Em especial à Daiane, Sandra e Estela por ajudarem pessoalmente na adequação o da tese.

Aos funcionários do Depto. de Pós-Graduação da FMVZ-USP também pelo carinho com o qual sempre me receberam e auxiliaram, sempre que precisei, desde o mestrado.

A todo o corpo de funcionários e docentes da FMVZ-USP, por tudo, desde o primeiro dia da graduação.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pela bolsa de doutorado e financiamento de auxílio à pesquisa concedidos.

*Só sei que nada sei...*

*(Sócrates)*

## PREFÁCIO

A tese aqui apresentada foi redigida com a intensão de expor, de maneira sintetizada e objetiva, as motivações do projeto de pesquisa executado e os resultados relevantes alcançados.

Na introdução buscou-se apresentar o contexto econômico e científico no qual o presente trabalho está inserido, sendo que o cenário científico foi apresentado de maneira mais elaborada (mas sempre com vistas à concisão e à objetividade) no item “revisão de literatura”.

Nos itens “resultados” e “discussão” ativemo-nos a apresentar os resultados relevantes alcançados, que pudessem contribuir e dialogar com o conhecimento científico disponível até o momento da redação, evitando-se a apresentação de dados redundantes, que pudessem causar confusão e tornar a leitura desagradável.

Buscou-se, tanto quanto possível, uma redação que se aproximasse daquela empregada em divulgação científica (em artigos de revistas científicas). Espera-se, com isso, que o formato escolhido torne o trabalho inteligível e de agradável leitura.

## RESUMO

AYRES, G. R. R. **Diversidade molecular dos genes codificadores das proteínas não-estruturais Nsp2 e protease Papaína-like e da proteína estrutural S1 de amostras brasileiras de *Coronavirus aviário***. [Molecular diversity of Nsp2 and Papain-like protease and S1 structural protein coding genes in Brazilian isolates of *Avian coronavirus*]. 2014. 76 f. Tese (Doutorado em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2014.

Coronavírus, incluindo-se o *Coronavirus aviário* (ACoV), possuem o maior genoma composto por RNA conhecido entre os vírus. Aproximadamente dois terços desse genoma codificam proteínas não estruturais (Nsps), cujas funções parecem estar associadas à replicação e patogênese viral. Até o momento, esses alvos têm sido pouco explorados quanto a sua diversidade em diferentes linhagens de ACoV. O presente estudo teve como objetivo investigar a diversidade dos genes codificadores das proteínas não estruturais Nsp2 e protease Papaína-like (Plpro), utilizando-se linhagens brasileiras de ACoV. Para tanto, 10 linhagens de ACoV, isoladas em ovos embrionados, foram submetidas à RT-PCR direcionada aos genes codificadores de Plpro e Nsp2, seguindo-se o sequenciamento de DNA e a análise filogenética, juntamente com sequências homólogas obtidas no GenBank. Além disso, realizou-se a genotipagem por meio do sequenciamento parcial do gene codificador da proteína de espícula (região S1). Três das amostras virais obtidas e investigadas no presente trabalho apresentaram padrão de segregação discordante para os genes estudados. O isolado CRG I22 agrupou-se com linhagens virais pertencentes ao genótipo Massachusetts para S1 e com o grupamento de ACoVs brasileiros os genes da Nsp2 e Plpro. O isolado CRG I33 agrupou-se com linhagens virais pertencentes ao genótipo brasileiro para s1 e plpro e de maneira divergente para o gene da Nsp2. Para o isolado CRG I38, não foi obtida a genotipagem por s1, entretanto, similarmente ao observado para o isolado CRG I33, esse isolado agrupou-se com linhagens virais brasileiras para o gene plro e de maneira independente para o gene nsp2. As demais linhagens estudadas resultaram na formação de um grupamento especificamente brasileiro de ACoV, para os três genes estudados. Esses achados sugerem a ocorrência de recombinação nessas amostras discrepantes. Quanto às identidades médias entre as sequências nucleotídicas analisadas, a região de s1 analisada apresentou as menores identidades (73,75% ±16,78), seguido pelo gene plpro (88,06% ±5,7) e do gene nsp2 (92,28%

±4,37), em acordo com a literatura. Assim sendo, os alvos investigados podem constituir ferramentas úteis na epidemiologia molecular do ACoV e na investigação de linhagens recombinantes do vírus. O presente estudo é o primeiro a investigar a diversidade genética de genes codificadores de proteínas não-estruturais em linhagens brasileiras de ACoV. Os resultados aqui apresentados reforçam a existência de um genótipo brasileiro de ACoV, para os 3 genes estudados. Entretanto, discrepâncias pontuais encontradas no padrão genotípico para s1, nsp2 e nsp3 permitem inferir uma diversidade genética maior do que a conhecida até o momento, possivelmente resultante de eventos de recombinação entre ACOVs brasileiros, ACOVs vacinais e outros ainda desconhecidos. Os resultados obtidos auxiliam na compreensão dos padrões e evolução dos ACOVs.

Palavras-chave: *Coronavírus aviário*. Vírus da bronquite infecciosa das galinhas.  
Nsp2. Nsp3. Diversidade molecular.

## ABSTRACT

AYRES, G. R. R. **Molecular diversity of Nsp2 and Papain-like protease and S1 structural protein coding genes in Brazilian isolates of *Avian coronavirus*** [Diversidade molecular dos genes codificadores das proteínas não-estruturais Nsp2 e protease Papaína-*like* e da proteína estrutural S1 de amostras brasileiras de *Coronavirus aviário*]. 2014. 76 f. Tese (Doutorado em Ciências) – Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2014.

Coronaviruses, including *Avian coronavirus* (ACoV), have the largest known RNA genome. Nearly two thirds of its genome codes for non-structural proteins (Nsps), whose functions appear to be linked to viral replication and pathogenesis. Hitherto these targets have been poorly explored regarding the ACoV lineages diversity. The present study aimed to assess the diversity of non-structural protein 2 (nsp2), papain-like protease (plpro) and spike protein (S1 subunit) coding genes, in Brazilian ACoV strains. To this end, 10 ACoV strains, isolated in embryonated eggs, had its 3<sup>rd</sup> and 5<sup>th</sup> passages submitted to RT-PCR targeting nsp2, plpro and s1, followed by DNA sequencing and phylogenetic analysis, herewith homologous sequences obtained from GenBank. Three of the ACoV strains sequenced showed a discordant segregation pattern for target genes. CRG I22 strain clustered with Massachusetts genotype strains for S1, and with Brazilian cluster for nsp3 and plpro genes. CRG I33 strain, clustered with Brazilian strains for S1 and plpro genes, and was divergent for nsp2 gene. For CRG I38 strain, the S1 sequence was not obtained, however, similarly to what was observed for CRG I33, this strain grouped with the Brazilian lineage for plpro gene and was divergent for nsp2 gene. All the other ACoV here sequenced resulted in a specific Brazilian cluster for the three studied genes. Regarding the mean nucleotide identities measured, s1 gene showed the lowest identity (73.75%  $\pm$ 16.78), followed by plpro gene (88.06%  $\pm$ 5.7) and nsp2 gene (92.28%  $\pm$ 4.37), in accordance with previous reported data. Therefore, the targets of the present study are useful tools for ACoV molecular epidemiology studies and for the survey of recombinant ACoV strains. The presented study is the first one investigating the molecular diversity of non-structural proteins coding genes in Brazilian strains of ACoV. Results achieved herein reinforce the data over the circulation of ACoV Brazilian strains in this country, for the three investigated genes. However, divergences found between S1, nsp2 and plpro genetic patterns allow inferring a higher molecular diversity than previously known. It is possible that this divergence is due to recombination events between ACoV from vaccines, Brazilian field

strains and others still unknown. These results contribute on the comprehension over genetic patters and evolution of ACoV.

Keywords: *Avian coronavirus*. Infectious bronchitis virus . Nsp2. Nsp3. Molecular diversity.

## LISTA DE ABREVIATURAS E SIGLAS

ACoV	<i>Coronavírus aviário (Avian Coronavirus)</i>
BIG	Bronquite Infecciosa das Galinhas
CoV	Coronavirus
cDNA	Ácido desoxirribonucleico complementar
DEPC	Dietil-pirocarbonato
DNA	Ácido desoxirribonucleico
dNTP	Deoxinucleosídeo-trifosfato
DO	Densidade óptica
et al.	e colaboradores
g	Grama(s)
h	Hora (s)
IBV	Vírus da Bronquite Infecciosa ( <i>Infectious Bronchitis Virus</i> )
ICTV	Comitê Internacional de Taxonomia de Vírus ( <i>International Comitee on Taxomony of Viruses</i> )
IL	Interleucina
Kb	Quilobase (s)
kD	QuiloDalton (s)
L	Litro(s)
m	Metro (s)
M	Molar
MHV	Vírus da hepatite murina ( <i>Murine Hepatitis Virus</i> )
min	Minuto de hora
mRNA	RNA mensageiro
Nsp	Proteína não estrutural ( <i>Non-structural protein</i> )
nsp	Gene codificador da proteína não estrutural
ORF	Janela aberta de leitura de genes ( <i>Open reading frameI</i> )
pb	Par (es) de base (s)
PCR	Reação pela polimerase em cadeia
PBS	Fosfato salino tamponado
PCR	Reação pela polimerase em cadeia
pH	Concentração de hidrogênio iônico
Plpro	Protease papaína- <i>like</i>
plpro	Gene codificador da protease papaína- <i>like</i>

qPCR	PCR quantitativa
q.s.p.	Quantidade suficiente para (um volume final de solução)
RNA	Ácido ribonucleico
RT	Transcrição reversa
s	Segundo (s) ou gene da proteína de espícula (conforme contexto)
S	Proteína de espícula ( <i>Spike protein</i> )
S1	Subunidade 1 da proteína de espícula
s1	Gene codificador da subunidade 1 da proteína de espícula
SARS	Síndrome Respiratória Severa Aguda ( <i>Severe Acute Respiratory Syndrome</i> )
TBE	TRIS Borato EDTA
TNF	Fator de necrose tumoral ( <i>Tumor necrosis factor</i> )
TRIS	Tris(hidroximetil)aminometano
x g	Aceleração da gravidade ( $9,8\text{m/s}^2$ )

## LISTA DE SÍMBOLOS

°C	Graus Celsius
+	Mais
®	Marca Registrada
μ	Micro
m	Mili
n	Nano
p	Pico
™	Marca registrada ( <i>Trademark</i> )
%	Porcentagem / Porcento
#	Número de acesso de sequencia gênica no GenBank

## SUMÁRIO

1	<b>INTRODUÇÃO</b> .....	21
2	<b>REVISÃO DE LITERATURA</b> .....	24
3	<b>OBJETIVO</b> .....	30
4	<b>MATERIAL E MÉTODO</b> .....	32
4.1	MEDIDAS DE CONTROLE À CONTAMINAÇÃO.....	32
4.2	CONTROLES POSITIVOS E NEGATIVOS.....	32
4.3	TRANSCRIÇÃO REVERSA, AMPLIFICAÇÃO E SEQUENCIAMENTO DOS GENES EM ESTUDO.....	33
4.4	AMOSTRAS SELECIONADAS PARA O ISOLAMENTO VIRAL.....	34
4.5	ISOLAMENTO DE ACoV EM OVOS EMBRIONADOS.....	36
4.6	AMPLIFICAÇÃO DOS GENES S1, NSP2 E NSP3.....	37
4.7	ANÁLISE DOS RESULTADOS.....	39
5	<b>RESULTADOS</b> .....	42
5.1	ISOLAMENTO DE ACoV.....	42
5.2	SEQUENCIAS NUCLEOTÍDICAS OBTIDAS.....	43
5.2.1	<b>Gene s1</b> .....	45
5.2.2	<b>Gene nsp2</b> .....	48
5.2.3	<b>Gene nsp3</b> .....	52
6	<b>DISCUSSÃO</b> .....	56
7	<b>CONCLUSÃO</b> .....	61
	<b>REFERÊNCIAS</b> .....	63
	<b>APÊNDICES</b> .....	68

## ***INTRODUÇÃO***

## 1 INTRODUÇÃO

O vírus da bronquite infecciosa das galinhas (IBV), mais recentemente denominado *Coronavirus aviário* (ACoV), é considerado endêmico no Brasil, apesar do amplo emprego de vacinas, devido à contínua emergência e disseminação de novos tipos virais, que compreendem diferentes patotipos, sorotipos e protectotipos (MONTASSIER, 2010).

Segundo a União Brasileira de Proteína Animal (UBPA, 2014), o setor avícola gera, direta e indiretamente, mais de 3,6 milhões de empregos no país, principalmente nos estados do Sul e do Sudeste, respondendo por quase 1,5% do Produto Interno Bruto (PIB) nacional. Desde 2004, o Brasil sustenta a posição de maior exportador mundial de carne de frango, tendo em 2011 atingido a marca histórica de 13,058 milhões de toneladas de carne de frango produzidas sendo dessas 3,9 milhões exportadas para mais de 150 países. Vale ressaltar, ainda, que cerca de 69% da produção avícola é consumida no mercado interno, com um consumo estimado de 39kg por pessoa ao ano (UBPA, 2014). Soma-se a isso o total acumulado de mais de 90 milhões de pintainhas alojadas para postura, em 2013 (OvoOnline, 2014), responsáveis por uma produção anual de mais de 2,745 bilhões de dúzias de ovos (IBGE, 2014).

Uma das principais causas de perdas econômicas na indústria avícola mundial são as infecções pelo ACoV, que compromete o desempenho produtivo das aves (CAVANAGH, 2007). Trata-se de um vírus epiteliotrópico, capaz de se replicar em toda a superfície epitelial respiratória, digestória, renal e nas gônadas (BRITTON; CAVANAGH, 2007), causando desordens clínicas diversas em aves reprodutoras, de postura e de corte, incluindo-se doença respiratória, doença entérica, nefrite, queda na produção de ovos e falhas reprodutivas em machos e fêmeas (CAVANAGH; GELB, 2008).

A manutenção e o aumento da lucratividade na indústria avícola dependem grandemente da sanidade do plantel em produção, sendo esse um grande desafio ao segmento, visto que os animais são usualmente confinados sob altas taxas de lotação, favorecendo a disseminação de agentes infecciosos.

A maior parte dos estudos de epidemiologia e caracterização molecular de ACoV concentram-se em genes codificadores de proteínas estruturais do vírus, especialmente as proteínas de espícula (S) e de nucleocapsídeo (N). Entretanto, a maior parte do genoma do ACoV codifica proteínas não estruturais, que são expressas pelo genoma do vírus nas células

por ele infectadas. Assim sendo, tem-se a caracterização molecular e, conseqüentemente, inferências epidemiológicas dela derivadas baseadas em uma pequena parte do genoma viral.

Por pouco se conhecer a diversidade dos genes codificadores de proteínas não estruturais em linhagens brasileiras de ACoV e, conseqüentemente, seus desdobramentos para o conhecimento dos ACOVs e sua epidemiologia molecular, o presente trabalho concentrou-se na caracterização molecular dos genes codificadores da proteína não estrutural 2 e da protease papaína-*like* (contida no gene da proteína não estrutural 3), em isolados de ACoV genotipados pelo sequenciamento da região codificadora da subunidade s1 da proteína de espícula viral.

***REVISÃO DE LITERATURA***

## 2 REVISÃO DE LITERATURA

O IBV, atualmente, está classificado na ordem *Nidovirales*, família *Coronaviridae*, subfamília *Coronavirinae*, gênero *Gamacoronavirus* (anteriormente denominado grupo 3 do gênero *Coronavirus*) e espécie *Coronavirus aviário* (ICTV, 2013). Assim, atendendo à nomenclatura atualmente adotada pelo ICTV, o presente trabalho adotará a designação *Coronavirus aviário* (ACoV) ao tratar do vírus causador da Bronquite Infecciosa das Galinhas (BIG).

Trata-se de um vírus altamente pleomórfico, envelopado com cerca de 120nm de diâmetro (CAVANAGH, 2007), composto por quatro proteínas estruturais: proteína de espícula (S), de envelope (E), de membrana (M) e de nucleocapsídeo (N) (MASTERS, 2006). Os principais alvos de estudos em ACoV são as proteínas S e N, e seus respectivos genes codificadores.

A proteína S é um dímero composto pelas subunidades S1 (amino-terminal) e S2 (carboxi-terminal), com cerca de 90kDa cada. S1 é responsável pela formação do bulbo da espícula viral, está diretamente associada à interação com receptores celulares e à resposta imune do hospedeiro, sendo, por isso, empregada para determinar o sorotipo/genotipo viral (CAVANAGH, 2007). A proteína N, por sua vez, liga-se ao genoma do vírus e participa dos processos de transcrição e tradução do RNA viral (MONTASSIER, 2010).

A análise dos genes codificadores de S1 e N tem se mostrado um método rápido e acurado para monitorar as variantes do vírus que emergem sob os pontos de vista fenotípico e epidemiológico, devido à participação das proteínas expressas pelos mesmos nos processos virais de replicação e imunidade, particularmente importantes na relação entre o vírus e seu hospedeiro (MONTASSIER, 2010).

Entretanto, por haver uma quantidade maior de estudos disponíveis sobre diversidade molecular de genes codificadores de proteínas estruturais, tem-se, atualmente, uma predição da evolução do ACoV baseada em uma pequena parte de seu genoma, já que a maior de seu genoma codifica proteínas não estruturais (MONDAL; CARDONA, 2004).

As proteínas não estruturais (Nsp) são expressas em células infectadas pelo ACoV. Embora suas funções ainda não estejam totalmente esclarecidas, sabe-se que os genes que as codificam (genes 1, 3a, 3b, 5a e 5b) compõe a maior parte de genoma viral (MASTERS, 2006).

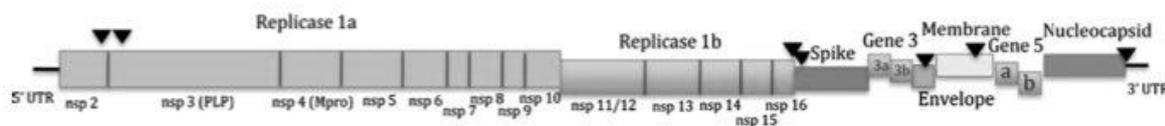
O genoma do ACoV é composto por RNA de fita simples e sentido positivo, com cerca de 27,6kb (CAVANAGH, 2007). Os vírus que possuem esse tipo de genoma expressam poliproteínas que são clivadas por proteases virais ou celulares, para formar proteínas maduras estruturais e não estruturais. Essa estratégia gera associações evolutivas, organizacionais e funcionais entre diferentes domínios proteicos (GRAHAM et al., 2005), originando alvos interessantes de investigação simultânea.

O processo de transcrição do genoma do ACoV, mais especificamente, ocorre de maneira descontínua e, durante o mesmo, são detectadas 6 espécies de RNAs mensageiros (mRNA): uma espécie que inclui a totalidade do genoma (27,6kb) e cinco mRNAs subgenômicos (2-7kb). Todos os 6 mRNAs possuem uma sequência líder não traduzida de 64 nucleotídeos (nt), idêntica à extremidade 5' do genoma viral (MONDAL; CARDONA, 2004). Este processo depende da ação da RNA polimerase dependente de RNA (*RdRp*), enzima viral cujo gene está localizado na região aberta de leitura 1 (ORF1) (GRAHAM et al., 2005; MASTERS, 2006).

A ORF1, também referida como gene da poliproteína replicase, compõe dois terços do genoma de todos os coronavírus (YU et al., 2012), e codifica as poliproteínas 1a e 1ab, que são processadas por proteases virais, durante e após o processo de transcrição, originando produtos proteicos intermediários e maduros não estruturais (CAVANAGH, 2007; GRAHAM et al., 2005).

Quanto à localização dos genes codificadores de domínios proteicos na ORF1 do ACoV, têm-se a seguinte distribuição: 10 domínios na ORF1a (nsp2 a nsp11) e 5 na ORF1b (nsp12 a nsp16) (SAWICKI; SAWICKI; SIDDELL, 2007; ZIEBUHR; SNIJDER, 2007). Em outras espécies de coronavírus, acrescenta-se na extremidade 5' da ORF1a o gene codificador da nsp1 (ZIEBUHR; SNIJDER, 2007), ausente no ACoV. Na figura 1 tem-se uma representação esquemática do genoma do ACoV.

Figura 1- Representação esquemática da organização do genoma do *Coronavirus aviário* (representação fora de escala)



Fonte: (JACKWOOD; HALL; HANDEL, 2012).

Legenda: A sigla nsp refere-se aos genes de proteínas não estruturais.  
A sigla PLP refere-se ao domínio codificador da protease papaína-like, contida no gene da Nsp3.  
A sigla Mpro refere-se ao domínio codificador da *Main* protease, contida no gene da nsp4.

A proteína não estrutural 2 (Nsp2) é, portanto, uma das primeiras proteínas do repertório genômico do ACoV e tem sido apontada como crucial para os primeiros estágios da infecção viral (YANG et al., 2009). Embora sua função exata permaneça desconhecida (YU et al., 2012) foi demonstrado que a Nsp2 possui resíduos hidrofóbicos, que parecem atuar no ancoramento do complexo replicase junto ao aparelho de Golgi da célula infectada (HAGEMEIJER et al., 2010).

Apesar de alguns estudos terem demonstrado que a Nsp2 não é necessária à replicação de coronavírus (GRAHAM et al., 2005), estudos mais recentes, conduzidos com o coronavírus causador da hepatite murina (MHV), um *Betacoronavirus*, demonstram que a Nsp2 desempenha papel fundamental na patogenia viral e que mutações no gene que a codifica levam a diferenças na virulência e patogênese de um mesmo genótipo viral inoculado em diferentes órgãos (ROTH-CROSS et al., 2009). Esses achados tornam a Nsp2 um alvo interessante para investigações genotípicas que possam colaborar para inferências fenotípicas também em ACoV.

Nas últimas duas décadas, novos domínios proteicos funcionais que atuam durante a replicação do ACoV foram reconhecidos ao longo do gene da poliproteína replicase (MONDAL; CARDONA, 2004), tendo-se destacado: a protease papaína-like (Plpro), localizada na Nsp3; a protease principal (*main proease* - Mpro), incluída na nsp5; e a RNA-polimerase RNA dependente (RdRp), localizada nos genes nsp11 e nsp12 (THOR et al., 2011).

Diversas proteases relacionadas a enzimas celulares semelhantes à papaína (proteases “papaína-like” - Plpro) foram descritas em vírus RNA de sentido positivo.

Por comparação de sequências gênicas, para o ACoV, foram descritos dois domínios ativos de P1pro: P11pro, localizado entre os nucleotídeos 4243 e 5019 (aminoácidos 1239 a 1497), que atuaria na clivagem da porção amino-terminal da poliproteína 1ab e, potencialmente, atuaria na regulação de reações na célula hospedeira à infecção; e P12pro, localizado entre os nucleotídeos 4681 e 5553 (aminoácidos 1385 e 1675), sem função descrita (LIM; NG; LIU, 2000; MONDAL; CARDONA, 2004).

Contudo, à medida em que diferentes espécies de coronavírus foram investigadas, verificaram-se diferenças importantes entre as P1pro e seus padrões de processamento proteolítico (THIEL, 2007).

Mais recentemente, em alguns *Betacoronavirus* e nos *Gammacoronavirus* (incluindo-se o ACoV), verificou-se a existência de uma única P1pro ativa, correspondente à P12pro das demais espécies de coronavírus. Em ACoV, especificamente, verificou-se a existência de sequências homólogas às da P11pro na P12pro, mas estas sequências foram consideradas codificantes de sítios inativos, por não incluírem porções essenciais à atividade proteolítica. Esses achados sugerem que devem haver diferenças na regulação dos processos mediados por P1pro entre diferentes linhagens filogenéticas de coronavírus (THIEL, 2007), ressaltando a importância dos mesmos como alvos para novos estudos.

Em SARS-CoV já se sabe que P1pro, juntamente com a Mpro, atua na clivagem das poliproteínas 1a e 1ab para a formação de suas 16 Nsps maduras, que incluem: helicase, nuclease, polimerase entre outras (TAN; LIM; HONG, 2005). Além disso, demonstrou-se a ação deubiquitinária P1pro em coronavírus (CLEMENTZ et al., 2010). A ubiquitina é uma pequena proteína regulatória de eucariotos, tem-se especulado que P1pro possa atuar na interação do ACoV com a célula infectada. Esses indícios somam-se a dados obtidos de ensaios *in vitro* que sugerem que essa protease é capaz de inibir a produção de interferon tipo 1 (ZHENG et al., 2008) e modular a expressão de TNF e  $\alpha$ -IL e IL-6 (ERIKSSON et al., 2008) em células infectadas pelo MHV. A multifuncionalidade verificada de P1pro, cuja importância se destaca pelo fato de incluir ações essenciais à replicação viral, tem feito com que a mesma despontasse como alvo de estudos de terapia antiviral (CHOU et al., 2012).

Quanto à variabilidade genética, em comparações realizadas entre linhagens heterólogas de ACoV, demonstrou-se que P1pro está dentre as regiões do genoma mais divergentes, comparando-se com outros domínios da ORF1ab, como Mpro e RdRp. Incluem-se nos achados de variabilidade no gene da replicase, achados de recombinação entre linhagens de campo e vacinais em ACoV. Essas diferenças na variabilidade genética verificada entre diferentes domínios da ORF1ab devem, provavelmente, decorrer de divergências na pressão

seletiva que incide sobre os esses genes e seus referidos produtos. Assim sendo, é possível que variações destes genes possam conferir vantagens adaptativas a linhagens de ACoV. Apesar de não se ter verificado uma associação entre a variabilidade gênica encontrada e aspectos geográficos, temporais ou sorotípicos, esses dados indicam a possibilidade de haver fenótipos do complexo replicase mais vantajosos ao processo de seleção viral (MONDAL; CARDONA, 2004).

Em um estudo que visou investigar o papel da ORF1ab nos processos de patogenia viral, um ACoV de linhagem virulenta, pertencente ao sorotipo/genótipo M41, teve um grande resíduo (8.4kb) da região final da OR1ab substituído pelo segmento correspondente, obtido de linhagem arquetípica de ACoV avirulenta Baudette. Foi possível demonstrar *in vivo* e *in vitro* uma redução da virulência do ACoV recombinado, comparando à linhagem virulenta M41, tendo-se apontado, portanto, o gene da replicase como um dos determinantes de patogenicidade em ACoV (ARMESTO; CAVANAGH; BRITTON, 2009).

Independentemente da discussão quanto ao seu papel na patogenicidade em diferentes linhagens de ACoV, análises de recombinação em sequencias completas de genoma de ACoV apontaram que os genes da Nsp2 e Nsp3 estão entre os alvos genômicos mais sujeitos a eventos de recombinação (THOR et al., 2011). Esses achados reforçam a importância de se investigarem essas regiões menos conservadas, a fim de levantar mais informações que colaborem para o conhecimento sobre a evolução e epidemiologia do ACoV.

É evidente, portanto, o estreito conhecimento sobre a diversidade molecular de genes codificadores de proteínas não estruturais de linhagens brasileiras de ACoV, e da escassez de estudos também em âmbito mundial. O presente trabalho pretende contribuir nesta seção da virologia básica do ACoV para que se possam compreender de um modo mais amplo as bases moleculares da evolução e epidemiologia viral, com o intuito de colaborar no delineamento e avaliação de medidas profiláticas à bronquite infecciosa das galinhas.

Assim sendo, esperou-se com o presente estudo ampliar o conhecimento sobre a diversidade genética do ACoV, sob um ponto de vista multigênico, com especial atenção aos genótipos brasileiros e suas particularidades.

***OBJETIVO***

### 3 OBJETIVO

O presente trabalho teve como objetivo investigar a diversidade molecular dos genes codificadores das proteínas Nsp2 e protease Papáína-*like* (contida na Nsp3), em isolados Brasileiros de ACoV, genotipados pelo sequenciamento do gene codificador da subunidade S1 da proteína de espícula viral.

## ***MATERIAL E MÉTODO***

## 4 MATERIAL E MÉTODO

Os materiais e os métodos empregados no presente trabalho estão apresentados nos tópicos a seguir.

### 4.1 MEDIDAS DE PREVENÇÃO À CONTAMINAÇÃO EM PCR

Todas as etapas anteriores à amplificação de DNA (preparação de amostras para o isolamento viral em ovos embrionados, extração de RNA, a transcrição reversa (RT), preparo da primeira PCR e de reagentes para *neste*-PCR) foram realizadas em laboratório exclusivo, empregando equipamento restrito, separadamente da manipulação de material amplificado (distribuição de produtos de PCR para reação de *neste*-PCR, eletroforese e purificação de *ampliações*), evitando-se contaminações.

### 4.2 CONTROLES POSITIVOS E NEGATIVOS

Para todas as reações, foram empregadas como controles positivos uma amostra vacinal do sorotipo Massachusetts e o isolado CRG I33 3<sup>a</sup> passagem (acervo do Laboratório de Biologia Molecular Aplicada à Sorologia – LABMAS). Como controle negativo, foi empregada água tratada com 0,1% dietilpirocarbonato (água DEPC), incluindo-se um controle negativo a cada 5 amostras investigadas.

#### 4.3 TRANSCRIÇÃO REVERSA, AMPLIFICAÇÃO E SEQUENCIAMENTO DOS GENES EM ESTUDO

As amostras virais investigadas foram submetidas à extração de RNA total com TRIzol<sup>®</sup> Reagent (Invitrogen<sup>™</sup>, California, USA) de acordo com as instruções do fabricante.

O RNA extraído foi, então empregado para a reação de transcrição reversa (RT) utilizando-se transcriptase reversa MMLV<sup>®</sup> (Invitrogen<sup>™</sup>- Life Technologies<sup>™</sup>, Foster City, CA) e Random Primers<sup>®</sup> (Invitrogen<sup>™</sup>- Life Technologies<sup>™</sup>, Foster City, CA), conforme as instruções do fabricante.

Nas reações de amplificação de DNA foi utilizado o Go Taq<sup>®</sup> Green Master Mix (Promega<sup>™</sup>, Madison, WI), de acordo com as instruções do fabricante.

A presença de produtos de amplificação foi verificada através de eletroforese em gel de agarose a 1,5%, corado com brometo de etídeo.

Nos casos em que não foi verificada a presença de produtos inespecíficos de amplificação, o produto amplificado foi purificado empregando-se o reagente EXoSap-IT<sup>®</sup> (Affymetrix<sup>™</sup>, USB<sup>®</sup> Products, Ohio, USA). Já para as reações em que bandas inespecíficas foram observadas, o produto amplificado de interesse foi cortado do gel de agarose, empregando-se lâmina estéril de bisturi, e conduziu-se sua purificação do gel de agarose com o kit Illustra GFX<sup>™</sup> (GE Healthcare<sup>™</sup>, Buckinghamshire, UK), sendo todas as etapas conduzidas em conformidade com as orientações dos respectivos fabricantes.

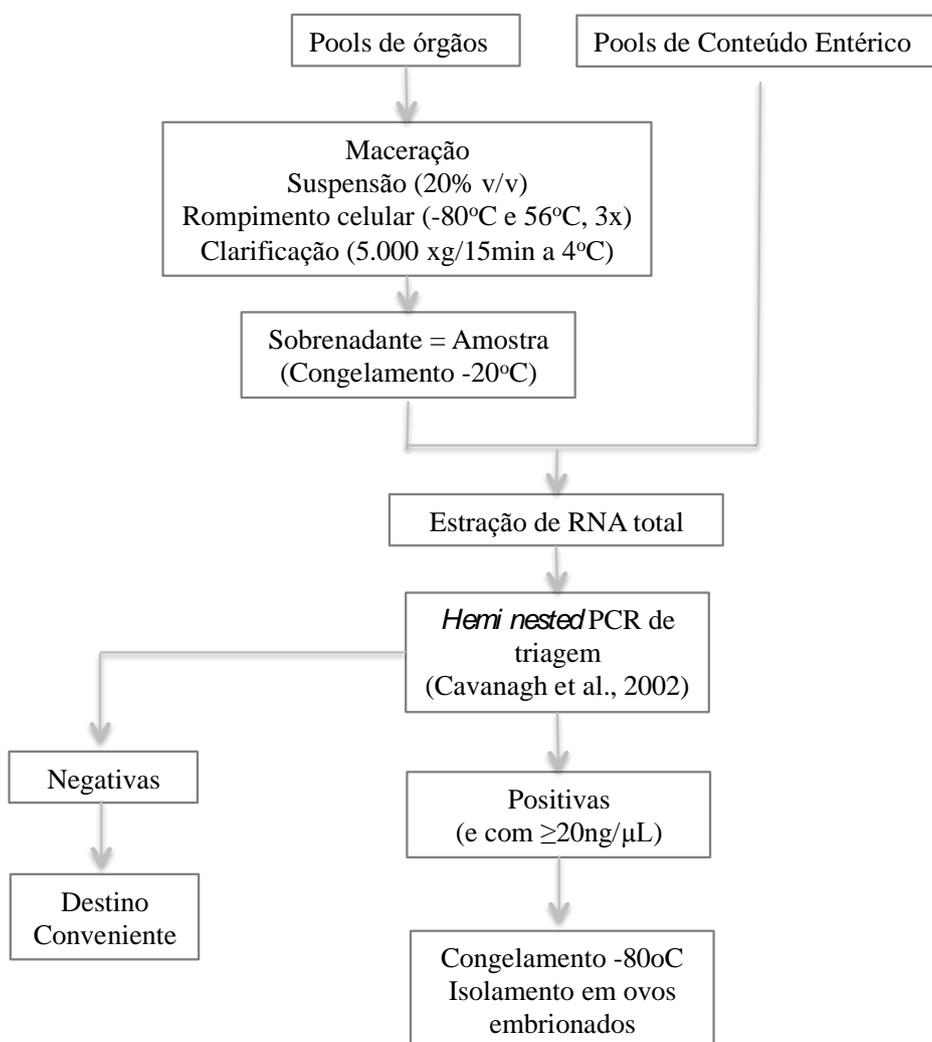
Os produtos de amplificação purificados foram, então, submetidos ao sequenciamento bidirecional, minimamente em duplicatas, empregando-se o Big Dye<sup>™</sup> 3.1 (Applied Biosystems<sup>™</sup>, Foster City, CA), no sequenciador automático ABI-3500<sup>®</sup> (Applied Biosystems<sup>™</sup>, Foster City, CA), também se atendendo as orientações do fabricante.

Para as amostras e/ou fragmentos cujo sucesso na obtenção de sequencias não foi satisfatório, alternativamente, realizou-se: extração de RNA total com o Direct-zol<sup>™</sup> (Zymon Research Coporation, Irvine, CA), reação de transcrição reversa com o Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, Pittsburg, PA) e PCR com a Phusion Green Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific, Pittsburg, PA). Como alternativa, ainda, foi realizada a RT empregando-se o primer anti-senso específico ao fragmento de difícil obtenção.

#### 4.4 AMOSTRAS SELECIONADAS PARA O ISOLAMENTO VIRAL

Foram empregadas nesse estudo amostras de campo, procedentes do Brasil, previamente diagnosticadas como positivas para a presença de RNA de ACoV, pela equipe do Prof. Dr. Paulo Eduardo Brandão do Laboratório de Biologia Molecular Aplicada e Sorologia (LABMAS-VPS-FMVZ-USP). O fluxograma apresentado na Figura 2 apresenta as principais etapas do processo de triagem realizado para as amostras utilizadas no presente estudo.

Figura 2- O fluxograma das principais etapas do processo de triagem realizado para as amostras utilizadas no presente estudo



As amostras foram, então, compostas de *pools* de órgãos (trato respiratório, rins e trato reprodutivo) e de conteúdo entérico, colhidos de 5 aves alojadas em um mesmo lote (compreendendo-se por lote o conjunto de animais alojados em um mesmo espaço físico). Foram testadas apenas as amostras mantidas sob adequada refrigeração (4°C) até sua chegada ao laboratório e que tenham sido recebidas bem acondicionadas, sem mistura de conteúdos de amostras distintas e devidamente identificadas.

Uma vez aprovadas quanto sua conservação, as amostras foram maceradas e utilizadas para o preparo de suspensões a 20% (volume/volume) em água-DEPC, em tubos livres de nucleases. As suspensões foram, então, submetidos a 3 ciclos de congelamento a -80°C e descongelamento a 56°C, e clarificadas a 5.000xg durante 15 minutos a 4°C, tendo-se tomado o sobrenadante como amostra para as etapas seguintes.

O processo de congelamentos e descongelamentos, utilizado para liberar os vírus das células possivelmente infectadas, não foi aplicado às amostras virais empregadas como controles positivos, nem às suspensões de conteúdo entérico, nem às amostras de vírus isolados (líquido alantoide obtido dos ovos embrionados empregados no isolamento). Nessas amostras o congelamento e descongelamento sucessivo é desnecessário, dada a ausência de tecido conjuntivo e outros tecidos, e poderia comprometer a integridade do RNA viral.

As suspensões de conteúdo entérico foram clarificadas conforme descrito para a suspensão de órgãos, também tomando-se o sobrenadante como amostra.

O cDNA (obtido conforme descrito no item 3.2) foi utilizado em uma reação de triagem, para detecção de RNA de *Gammacoronavirus*, previamente descrita (CAVANAGH et al., 2002).

Foram eleitas para a tentativa de isolamento em ovos embrionados aquelas amostras que resultaram em *amplicons* com mais altas concentrações de DNA (estimada com Low DNA Mass Ladder, Invitrogen™ by Life Technologies™, California, USA), apresentando minimamente 20ng de DNA por microlitro de reação. As amostras eleitas foram imediatamente acondicionadas a -80°C, até o momento de seu envio para isolamento viral.

O envio dessas amostras aos laboratórios particulares onde foram realizadas as tentativas isolamento viral em ovos embrionados foi feito em caixas térmicas empregando-se gelo reciclável previamente congelado a -80°C por 24 horas, tendo sido confirmado o recebimento das mesmas ainda congeladas na chegada ao seu destino, antes de iniciarem-se os procedimentos visando o isolamento do ACoV.

#### 4.5 ISOLAMENTO DE ACoV EM OVOS EMBRIONADOS

Visando aumentar as possibilidades de obtenção de ACoV isolados em altos títulos, optou-se por remeter as amostras eleitas para o isolamento em ovos embrionados a laboratórios particulares que sabidamente realizam esse procedimento rotineiramente.

Resumidamente, para o isolamento viral em ovos embrionados, cada amostra de campo (suspensão de *pool* de órgãos) foi inoculada em 5 ovos embrionados na 1ª passagem. Nas passagens subsequentes, os inóculos foram compostos pelo *pool* de conteúdo alantoide esgotado desses 5 ovos inoculados (ETERRADOSSI; BRITTON, 2011). O *pool* de fluido alantoide obtido da 3ª e ou 5ª passagem (de acordo com a disponibilidade) foi tomado como amostra para amplificação dos genes em estudo, sendo diretamente empregado na extração de RNA. O sucesso do isolamento de ACoV foi verificado pela detecção de RNA viral por RT-*hemi-nested* PCR (CAVANAGH et al., 2002).

Assim sendo, de acordo com a disponibilidade de recursos e visando a obtenção do maior número possível de isolados de ACoV, 31 amostras foram enviadas para tentativa de isolamento viral. As remessas de amostras foram feitas em 3 lotes distintos.

O 1º lote conteve 7 amostras (pools de órgãos macerados e congelados e remetidos a -80°C) a um laboratório do Estado de São Paulo.

O 2º e 3º lotes de amostras foram enviados a um novo laboratório localizado no Estado de Minas Gerais, cuja equipe técnica recomendou que as amostras fossem acondicionadas em glicerol absoluto e mantidas de -20°C a -80°C, para melhor preservar a viabilidade das partículas virais. Optou-se pelo acondicionamento em glicerol e congelamento a -80°C.

O 2º lote de amostras submetido ao laboratório conteve 13 amostras, cujas condições de conservação incluíram a conservação em glicerol a -80°C, imediatamente após a detecção do RNA viral. Por fim, o 3º lote conteve 11 amostras das quais 5 haviam sido colhidas nas mesmas condições previamente descritas para o 2º lote, e 6 foram colhidas frescas e conservadas em glicerol absoluto e mantidos a -80°C imediatamente após a necropsia (visando otimizar o sucesso de isolamento).

#### 4.6 AMPLIFICAÇÃO DOS GENES S, NSP2 E NSP3

Para a amplificação e sequenciamento parcial da subunidade S1 do gene de espícula (s) empregaram-se os *primers* previamente descritos (WORTHINGTON; CURRIE; JONES, 2008) e *primers* para a amplificação da região S1 completa, cujas sequências foram gentilmente cedidas pela equipe Prof. Dr. Mark Jackwood (Jackwood's Laboratory, Poultry Diagnostic and Research Center, University of Georgia, USA) sendo as sequências nucleotídicas dos *primers*: senso (5' TGA AAC TGA ACA AAA GAC 3'), e anti-senso (5' CCA TAA GTA ACA TAA GGR CRA 3'). Para esse último par de *primers* empregou-se o seguinte ciclo de PCR: 1 ciclo a 94°C por 5min, 11 ciclos a 94°C/30s – 47,5°C/30s – 68°C/5min, 24 ciclos a 94°C/30s – 45°C/30s – 68°C/5min e 1 ciclo a 68°C por 15min.

Para a amplificação dos genes da proteína não estrutural 2 (Nsp2) e da proteína não estrutural 3 (Nsp3), contendo o gene da protease papaína-like (Plpro), foram desenhados os *primers* apresentados no quadro 1.

Os *primers* foram desenhados a partir de sequências informativas e não redundantes obtidas do GenBank alinhadas com a ferramenta Clustal W incluído no software Bioedit 7.0.9.0 (HALL, 1999), buscando-se regiões genômicas conservadas para os alvos em estudo. As sequências de nucleotídeos eleitas para potencial ancoragem de *primers* foram então analisadas e escolhidas com o auxílio da ferramenta online NetPrimer (disponível em: <http://www.premierbiosoft.com/netprimer/>).

Para o *primers* apresentados no quadro 1 empregou-se o seguinte ciclo consensual de PCR: 1 ciclo a 94°C por 4min, 11 ciclos a 94°C/30s – 45°C/30s – 68°C/5min, 24 ciclos a 94°C/30s – 42°C/30s – 68°C/5min e 1 ciclo a 68°C por 15min.

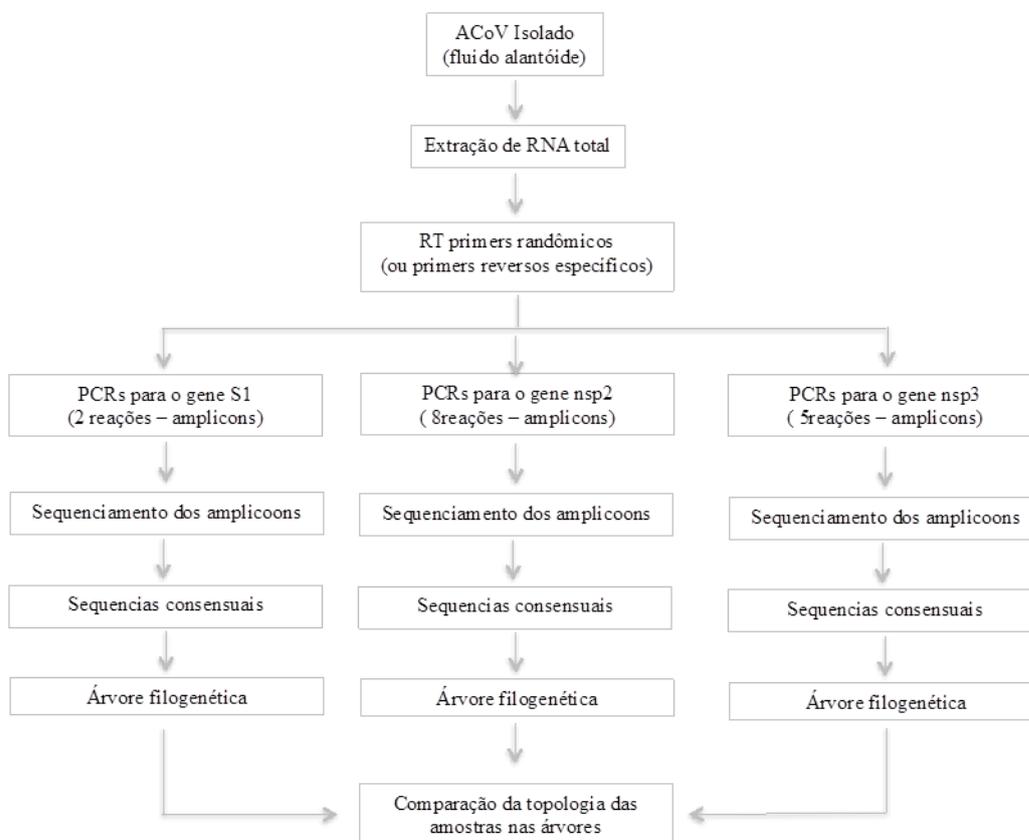
Quadro 1- *Primers* desenhados para amplificação dos genes das proteínas não estruturais 2 (Nsp2) e 3 (Nsp3) com suas respectivas posições 5'-3' (com relação à sequência de ACoV Beaudette, GenBank #NC\_001451) e o fragmento amplificado esperado, estimado por alinhamentos realizados no Clustal W incluído no software Bioedit 7.0.9.0 (HALL, 1999)

<b>Primer</b>	<b>Sequência ( 5' -3')</b>	<b>Posição (nt)</b>	<b>Fragmento (pb)</b>
nsp2-S1	GCAGACATACAATAGACAGTGAC	503-525	466
nsp2-A1	TGAAGCCAAAGTATTTTCAGT	949-968	
nsp2-S2	GAAGACATCTTTGGTGTCTC	820-839	443
nsp2-A2	CTTGCAAGACAAGTTCC	1246-1262	
nsp2-S3	CTGGCYATTTTTGARRATGT	1117-1136	449
nsp2-A3	TCACCACTCYTAAAGACATA	1546-1565	
nsp2-S4	GGCATGCGAAAYGAYYTAACD	1453-1473	440
nsp2-A4	TCAGTAARTGMAGACCAMACATC	1870-1892	
nsp2-S5	GTGAACCATTYAAGTTYTD	1740-1759	457
nsp2-A5	AGCARHAGTTGRAAACAR	2179-2196	
nsp2-S6	AGTYGTTGAYTTYTGTM	2071-2088	663
nsp2-A6	TCRTARAYCACAGHGAGC	2721-2733	
nsp2-A senso	CRGAGATAAATGATGGAC	1651- 1669	681
nsp2-A anti-senso	TCAAACCAAACCTTCATCGC	2312 - 2331	
nsp2-B senso	AYTGTTTTCAACTATTGCTGG	2185 - 2206	877
nsp2-B anti-senso	AGACMCTATAATCWTCATC	3041 - 3060	
nsp3-C senso	AGCATTGGATGARTTTAAAGAG	3810 - 3832	714
nsp3-C anti-senso	GCYTCTGCAAGAAAACC	4506 -4523	
nsp3-S1	GATGTTGAGGAYRAAGAARTYCTCT	4186-4210	506
nsp3-A1	TTTGWRTAYTGHRTYTAAAATG	4669-4691	
nsp3-D senso	GGTTTTCTTGCAGARGC	4506 - 4523	396
nsp3-D anti-senso	CTACAGCATTTTCATCAC	4884 - 4902	
nsp3-S2	CTTAAGARGYGYGTDWCNTGYAA	4573-4595	455
nsp3-A2	AGRTTRGMARCATCTTCYTTTA	5006-5027	
nsp3-S3	GCTTTTGAYAATYTKGCTMA	4870-4889	820
nsp3-A3	RGDGACANTYTYTTTCAAARTG	5667-5689	

Fonte: (AYRES, G. R. R., 2014).

A figura 3 apresenta um fluxograma dos procedimentos realizados para a obtenção das sequências dos genes nsp2, nsp3 e s1.

Figura 3- Fluxograma das principais etapas do processo de amplificação, sequenciamento e análise das sequências obtidas para os genes nsp2, nsp3 e s1, para cada um dos isolados de ACoV estudados



Fonte: (AYRES, G. R. R., 2014).

#### 4.7 ANÁLISE DOS RESULTADOS

As sequências nucleotídicas resultantes das reações de sequenciamento tiveram sua qualidade analisada com auxílio do aplicativo Phred (*online* em <http://asparagin.cenargen.embrapa.br/phph/>), prosseguindo na análise apenas sequências com escores superiores a 20 (indicando um erro inferior a 1 base a cada 100 sequenciadas) (EWING; GREEN, 1998). Adicionalmente, discrepâncias e possíveis erros de interpretação foram verificados com auxílio do programa FinchTV 1.5.0 (©Copyright 2013 PerkinElmer).

Com o auxílio da ferramenta Cap-Contig, incluída no programa Bioedit 7.0.9.0 (HALL, 1999), as sequências parciais de cada gene de cada amostra, foram alinhadas e comparadas,

obtendo-se sequências consensuais finais. Essas, por sua vez, foram verificadas com auxílio da ferramenta BLASTn (*online* em <http://www.ncbi.nlm.nih.gov/BLAST>).

As sequências finais obtidas foram, então, alinhadas e comparadas com sequências homólogas extraídas do GenBank, empregando-se o CLUSTAL W, incluído no programa Bioedit 7.0.9.0 (HALL, 1999), ou o Muscle, incluído no MEGA 6 (TAMURA et al., 2013), tendo-se escolhido os melhores alinhamentos obtidos.

Por fim, árvores filogenéticas das referidas sequências foram construídas segundo o algoritmo neighbor-joining de distância com o modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap, utilizando-se o programa MEGA6 (TAMURA et al., 2013). Valores de bootstrap superiores a 50 e o agrupamento de pelo menos 3 sequências foram tomados como necessários para o estabelecimento de um *cluster* ou linhagem de ACoV. As árvores obtidas foram analisadas individualmente e em conjunto, a fim de se investigar, sob o ponto de vista multigênico, a filogenia das sequências obtidas.

As identidades médias entre as sequências e seus respectivos desvios padrão foram calculados com o programa Microsoft® Office Excel® para Mac 2011 versão 14.4.3.

## ***RESULTADOS***

## 5 RESULTADOS

### 5.1 ISOLAMENTO DE ACoV

Não houve sucesso no isolamento de nenhuma das 7 amostras enviadas ao laboratório do Estado de São Paulo. Dentre as amostras remetidas ao laboratório do Estado de Minas Gerais, foi possível isolar ACoV em 2 das 13 amostras que compuseram o 2º lote e 2 das 11 amostras que compuseram o 3º lote.

No quadro 2 estão resumidas as informações relativas às amostras submetidas para isolamento de ACoV:

Quadro 2- Resumo das informações referentes às amostras que foram submetidas aos laboratórios especializados para isolamento viral em ovos embrionados

Amostra para tentativa de isolamento	Lote / Ano de colheita	Órgão	Laboratório	Conservação	Resultado
CRG 8	07/12 AR	Aparelho Reprodutivo	SP	Congelamento -80o C após a triagem	Não isolado
CRG 9	07/12PT	Pulmão e traquéia	SP	Congelamento -80o C após a triagem	Não isolado
CRG 10	07/12R	Rim	SP	Congelamento -80o C após a triagem	Não isolado
CRG 11	10/12P	Pulmão	SP	Congelamento -80o C após a triagem	Não isolado
CRG 12	10/12R	Rim	SP	Congelamento -80o C após a triagem	Não isolado
CRG 13	11/12P	Pulmão	SP	Congelamento -80o C após a triagem	Não isolado
CRG 14	11/12R	Rim	SP	Congelamento -80o C após a triagem	Não isolado
CRG 15	71/13	Pulmão e traquéia	MG	Congelamento -80o C após a triagem	Não isolado
CRG 16	71/13	Rim	MG	Congelamento -80o C após a triagem	Não isolado
CRG 17	73/13	Rim	MG	Congelamento -80o C após a triagem	Não isolado
CRG 18	73/13	Conteúdo entérico	MG	Congelamento -80o C após a triagem	Não isolado
CRG 19	81/13	Conteúdo entérico	MG	Congelamento -80o C após a triagem	Não isolado
CRG 20	81/13	Aparelho Reprodutivo	MG	Congelamento -80o C após a triagem	Não isolado
CRG 21	81/13	Rim	MG	Congelamento -80o C após a triagem	Não isolado
CRG 22	85/13	Pulmão e Traquéia	MG	Congelamento -80o C após a triagem	Isolado
CRG 23	85/13	Rim	MG	Congelamento -80o C após a triagem	Isolado
CRG 24	88/13	Rim	MG	Congelamento -80o C após a triagem	Não isolado
CRG 25	88/13	Traquéia	MG	Congelamento -80o C após a triagem	Não isolado
CRG 26	88/13	Aparelho Reprodutivo	MG	Congelamento -80o C após a triagem	Não isolado
CRG 27	91/13	Aparelho Reprodutivo	MG	Congelamento -80o C após a triagem	Não isolado
CRG 28	C12/13	Pulmão	MG	Imediatamente colocada em glicerol e -80oC	Não isolado
CRG 29	C02/13	Pulmão	MG	Imediatamente colocada em glicerol e -80oC	Não isolado
CRG 30	113/13	Pulmão	MG	Colocada em glicerol e -80oC após a triagem	Isolado
CRG 31	C15/13	Traquéia	MG	Imediatamente colocada em glicerol e -80oC	Não isolado
CRG 32	129/13	Traquéia	MG	Colocada em glicerol e -80oC após a triagem	Não isolado
CRG 34	122/13	Rim	MG	Colocada em glicerol e -80oC após a triagem	Não isolado
CRG 35	118/13	Rim	MG	Colocada em glicerol e -80oC após a triagem	Não isolado
CRG 36	C10/13	Rim	MG	Imediatamente colocada em glicerol e -80oC	Não isolado
CRG 37	C18/13	Intestino	MG	Imediatamente colocada em glicerol e -80oC	Não isolado
CRG 38	128/13	Tonsila	MG	Colocada em glicerol e -80oC após a triagem	Isolado
CRG 39	C05/13	Aparelho Reprodutivo	MG	Imediatamente colocada em glicerol e -80oC	Não isolado

Fonte: (AYRES, G. R. R., 2014).

Como é possível observar no quadro 2 houve um baixo sucesso no isolamento de amostras de ACoV, e, aparentemente, a conservação a  $-80^{\circ}\text{C}$  parece ter colaborado mais para o sucesso de isolamento do que o acréscimo do glicerol nesse processo.

Adicionalmente aos 4 isolados obtidos, o presente trabalho dispôs de 6 ACoV isolados do acervo do *Coronavirus Research Group* (CRG), liderado pelo Prof. Dr. Paulo Eduardo Brandão, sendo esses denominados: CRG I1, I2, I3, I4, I5 e I33. As informações acerca dos isolados estão resumidas no quadro 3.

Quadro 3- ACOVs isolados em ovos embrionados empregados no presente estudo

<b>Isolado</b>	<b>Procedência</b>	<b>Amostra empregada para o isolamento</b>	<b>Observação</b>
CRG-I1 3rd	Granja 1	Conteúdo entérico	Acervo CRG
CRG-I1 5th	Granja 1	Conteúdo entérico	Acervo CRG
CRG-I2 3rd	Granja 1	Pulmão e Traquéia	Acervo CRG
CRG-I2 5th	Granja 1	Pulmão e Traquéia	Acervo CRG
CRG-I3 3rd	Granja 1	Rim	Acervo CRG
CRG-I3 5th	Granja 1	Rim	Acervo CRG
CRG-I4 3rd	Granja 2	Conteúdo entérico	Acervo CRG
CRG-I4 5th	Granja 2	Conteúdo entérico	Acervo CRG
CRG-I5 3rd	Granja 2	Pulmão e Traquéia	Acervo CRG
CRG-I5 5th	Granja 2	Pulmão e Traquéia	Acervo CRG
CRG-I22 3rd	Granja 5	Pulmão e Traquéia	Isolado MG
CRG-I23 3rd	Granja 5	Rim	Isolado MG
CRG-I30 3rd	Granja 6	Pulmão	Isolado MG
CRG-I33 3rd	Granja 7	Aparelho Reprodutivo	Acervo CRG
CRG-I38 3rd	Granja 8	Tonsila	Isolado MG

Fonte: (AYRES, G. R. R., 2014).

## 5.2 SEQUÊNCIAS NUCLEOTÍDICAS OBTIDAS

Embora possam ocorrer variações na anotação precisa dos genes no genoma dos ACOVs, de acordo com as inserções e/ou deleções que podem ser encontradas nas diversas linhagens dos vírus, os genes em estudo estão anotados conforme o quadro 4 (tomando-se como referência a anotação do genoma completo do ACoV da linhagem Beaudette, com número de acesso no GenBank NC\_01451). Vale ressaltar ainda, que o posicionamento dos alinhamentos também pode variar de acordo com divergências entre as sequências alinhadas.

Empregando-se as sequências geradas no presente trabalho e sequências homólogas recuperadas do GenBank, foram construídas árvores filogenéticas para nucleotídeos e

aminoácidos (conforme descrito no item 4.7). Entretanto, apenas as árvores de nucleotídicas foram incluídas, por apresentarem topologia equivalente às de aminoácidos e por estarem as primeiras mais bem resolvidas e com maiores valores de bootstrap sustentando cada nó. Além disso, vale informar que estão inseridas no corpo do texto apenas as árvores mais relevantes para a apresentação dos resultados alcançados. Árvores complementares estão incluídas no tópico anexos (item 9 do presente documento).

Quadro 4- Anotação dos genes nsp2 e nsp3 (com seus respectivos domínios ativos conhecidos), com relação ao genoma do ACoV Beaudette (número de acesso no GenBank NC\_001451)

Gene	Domínio	nucleotídeo inicial	nucleotídeo final	Comprimento (nucleotídeos)	aminoácido inicial	aminoácido final	Comprimento (aminoácidos)
nsp2	-	529	2.547	2.019	177	849	673
nsp3	Domínio Acídico	2.548	3.027	480	850	1.009	160
	PL1	3.028	3.594	567	1.010	1.198	189
	ADRP	3.595	4.053	459	1.199	1.351	153
	PL2	4.054	4.944	891	1.352	1.648	297
	?	4.945	5.934	990	1.649	1.978	330
	Y	5.935	7.326	1.392	1.979	2.442	464
s1	-	20.368	22.020	1.653	6.614	7.165	551

Fonte: (AYRES, G. R. R., 2014).

No quadro 5 estão apresentadas as sequencias de genoma completo recuperadas do GenBank e que foram empregadas na reconstrução das árvores filogenéticas. Essas sequencias foram eleitas depois da reconstrução de diversas árvores filogenéticas para os 3 genes aqui estudados (s1, nsp2 e nsp3), tendo-se excluído aquelas redundantes e pouco informativas.

Quadro 5- Lista das sequencias recuperadas do GenBank e empregadas na reconstrução das árvores filogenéticas, com seus respectivos números de acesso

Número de acesso no GenBank (#)	Denominação	Observação
NC001451	Mass/Beaudette_ACoV_Ref	Linhagem Beaudette (arquetípica) de ACoV, sorotipo Massachusetts, sequencia referência do GenBank
GQ504724	Mass/Mass41/41	Linhagem virulenta de ACoV sorotipo Massachusetts
GQ504725	Mass41-attenuated	Linhagem vacinal de ACoV sorotipo Massachusetts
FJ904713	IBV/Mass/06	Linhagem de campo de ACoV sorotipo Massachusetts
GQ504720	Ark/Arkansas-DPI/81	Linhagem virulenta de ACoV sorotipo Arkansas
GQ504721	Ark-attenuated	Linhagem vacinal de ACoV sorotipo Arkansas
GQ504722	GA98/CWL047/98	Linhagem virulenta de ACoV sorotipo Georgia
GQ504723	GA98-attenuated	Linhagem vacinal de ACoV sorotipo Georgia
EU714028	CK/CK/ZJ971/91	Linhagem asiática de campo de ACoV sorotipo Massachusetts, associada a proventriculite.
GU393333	FL/FL18288/91	Linhagem de campo de ACoV sorotipo Florida
FJ904719	Conn/Conn46/91	Linhagem de campo de ACoV sorotipo Connecticut
AY641576	Peafowl/GD/KQ6/03	Linhagem de campo de ACoV sorotipo Massachusetts isolada de pavão
JF274479	ck/CH/LHJ/07VII	Linhagem asiática de campo de ACoV, sorotipo Massachusetts
FN430414	ITA/90254/05	Linhagem QX de ACoV
FN430415	NGA/A116E7?06	Linhagem QX de ACoV
EU637864	CK/CH/LSD/05I	Linhagem asiática de campo de ACoV
JF330898	ck/CH/LHB/100801	Linhagem asiática de campo de ACoV
DQ646405	TW2575/98	Linhagem asiática de campo de ACoV
NC_010800	TCoV_Ref	Coronavírus de perús, sequencia referência do GenBank
EU022526	TCoV-ATCC	Coronavírus de perús, linhagem ATCC

Fonte: (AYRES, G. R. R., 2014).

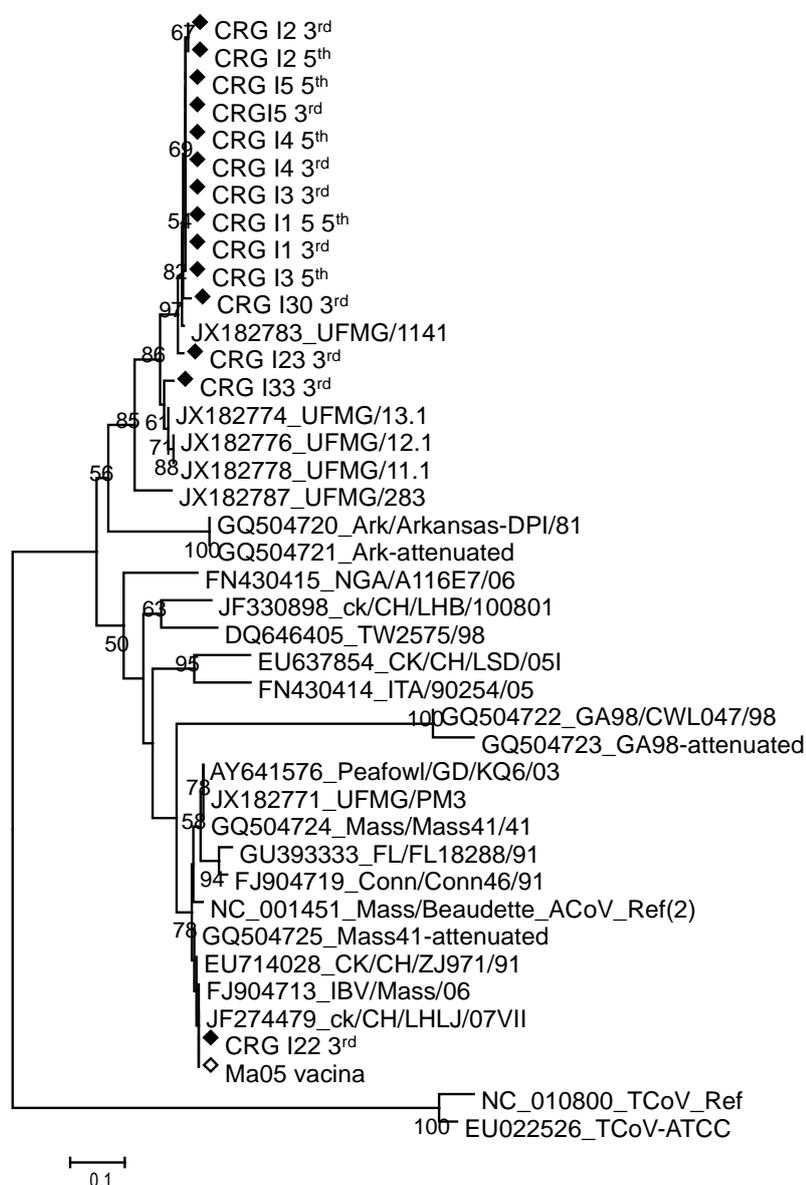
### 5.2.1 Gene s1

Foi possível realizar a genotipagem, pelo sequenciamento da região S1, de quase todos os isolados obtidos, exceto pelo isolado CRG I38 3<sup>a</sup> passagem. Destaca-se o isolado CRG I33 3<sup>a</sup> passagem, para o qual foi possível obter a 98,06% da sequencia completa de S1. Para a maioria dos isolados, foi possível obter sequencias do início e do final do gene s1, restando uma laguna intermediária para que se pudesse gerar uma sequencia única. Os resultados obtidos estão compilados no quadro 6.

As árvores filogenéticas reconstruídas demonstram que praticamente todos os isolados obtidos pertencem ao genótipo brasileiro de ACoV, previamente reportado (VILLARREAL et al., 2010). O isolado CRG I22 3<sup>a</sup> passagem foi o único que se agrupou com amostras do genótipo Massachusetts, incluindo-se a amostra vacinal empregada comercialmente aqui no Brasil (aqui denominada Ma05 vacina) sequenciada nesse trabalho.

A figura 4 apresenta a árvore filogenética referente ao segmento do gene s1 consensual a todas as amostras sequenciadas. Como as topologias se mantêm para os diversos segmentos sequenciados (correspondentes a diferentes segmentos do referido gene), árvores reconstruídas com segmentos maiores de S1 estão incluídas nos apêndices (Apêndices A, B e C).

Figura 4- Árvore filogenética par o gene s1, para os nucleotídeos 20.387 a 20.746 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

Para as sequências incluídas na árvore apresentada na figura 4, a identidade média entre os nucleotídeos foi de 73,75% ( $\pm 16,78\%$ ). Excluindo-se as sequências de coronavírus de perús (TCoV), bastante divergentes das sequências obtidas de ACoVs encontrados em galinhas, a identidade verificada se elevou para 77,58% (com desvio padrão de 12,65%).

Quando todas as amostras brasileiras foram consideradas, verificou-se uma identidade média de 81,43% ( $\pm 13,09$ ). E, por fim, quando apenas as sequências que compuseram o genótipo brasileiro foram analisadas, a identidade nucleotídica média foi de 82,09% ( $\pm 13,48$ ).

Quadro 6- Anotação das sequências obtidas para o gene s1, e as respectivas porcentagem de cobertura em relação ao gene completo, tomando-se por base o genoma do ACoV linhagem Beaudette (número de acesso no GenBank NC\_001451). Para alguns isolados obtiveram-se sequências da 3<sup>a</sup> (3rd) e 5<sup>a</sup> (5th) passagens em ovos embrionados

Gene	Isolado	nucleotídeo inicial	nucleotídeo final	Comprimento (pares de bases)	% de cobertura (em relação PL 2 completa)
s1	CRG-I1 3rd	20.354	21.094	740	43,95%
		21.150	21.984	834	50,48%
	CRG-I1 5th	20.354	21.094	740	43,95%
		21.150	21.993	843	51,03%
	CRG-I2 3rd	20.353	21.044	691	40,92%
		21.150	21.992	842	50,97%
	CRG-I2 5th	20.353	20.995	642	37,95%
		21.446	21.986	540	32,69%
	CRG-I3 3rd	20.351	21.014	663	39,10%
		21.329	21.993	664	40,19%
	CRG-I3 5th	20.353	21.017	664	39,29%
		21.150	21.992	842	50,97%
	CRG-I4 3rd	20.354	21.092	738	43,83%
		21.318	21.977	659	39,89%
	CRG-I4 5th	20.353	21.032	679	40,19%
		21.150	21.975	825	49,94%
	CRG-I5 3rd	20.353	21.121	768	45,58%
		21.293	21.986	693	41,95%
	CRG-I5 5th	20.346	20.988	642	37,53%
		21.150	21.976	826	50,00%
	CRG-I22 3rd	20.346	21.055	709	41,59%
		21.139	21.966	827	50,06%
	CRG-I23 3rd	20.345	21.141	796	46,79%
		21.246	21.989	743	44,98%
CRG-I30 3rd	20.386	20.683	297	17,98%	
CRG-I33 3rd	20.353	21.988	1.635	98,06%	
Ma05 (vacina)	20.353	21.142	789	46,85%	
	21.228	21.480	252	15,25%	

Fonte: (AYRES, G. R. R., 2014).

### 5.2.2 Gene nsp2

Quanto às sequências obtidas para o gene nsp2, foi possível obter as sequências completas para 6 dos ACoV isolados estudados (Quadro 7).

Quadro 7- Anotação das sequências obtidas para o gene nsp2, e as respectivas porcentagem de cobertura em relação ao gene completo, tomando-se por base o genoma do ACoV linhagem Beaudette (número de acesso no GenBank NC\_001451). Para alguns isolados obtiveram-se sequências da 3<sup>a</sup> (3rd) e 5<sup>a</sup> (5th) passagens em ovos embrionados

Gene	Isolado	nucleotídeo inicial	nucleotídeo final	Comprimento (pares de bases)	% de cobertura (em relação ao gene completo)	
nsp2	CRG-I1 3rd	503	1.892	1.389	67,51%	
	CRG-I1 5th	503	3.035	2.532	100,00%	
	CRG-I2 3rd	503	2.295	1.792	87,47%	
	CRG-I2 5th	504	2.227	1.723	84,10%	
	CRG-I3 3rd	503	3.045	2.542	100,00%	
	CRG-I3 5th	503	1.565	1.062	51,31%	
	CRG-I4 3rd	503	2.251	1.748	85,29%	
	CRG-I4 5th	503	2.318	1.815	88,61%	
	CRG-I5 3rd	503	1.892	1.389	67,51%	
	CRG-I5 5th	503	2.909	2.406	100,00%	
	CRG-I22 3rd	503	2.988	2.485	100,00%	
	CRG-I23 3rd	503	3.034	2.531	100,00%	
	CRG-I30 3rd		515	969	454	21,79%
			1.117	2.880	1.763	70,83%
	CRG-I33 3rd	503	3.044	2.541	100,00%	
	CRG-I38 3rd		503	1.582	1.079	52,15%
			1.659	3.036	1.377	43,98%
	Ma05 (vacina)		503	1.572	1.069	51,66%
1.594			3.049	1.455	47,20%	

Fonte: (AYRES, G. R. R., 2014).

Em decorrência das múltiplas sequências homólogas obtidas para cada um dos isolados estudados, foi possível verificar a ocorrência de subpopulações virais para uma mesma linhagem isolada. Para que fosse atribuída uma degeneração em um sítio, foram estabelecidos os seguintes critérios: (1) só foram considerados nucleotídeos sequenciados com máximo escore Phred e cujo pico visualizado no eletroferograma (com auxílio do FichTV) estivesse resolvido e (2) apenas polimorfismos verificados em 2 ou mais sequências obtidas (e que atendam aos critérios de qualidade) foram considerados.

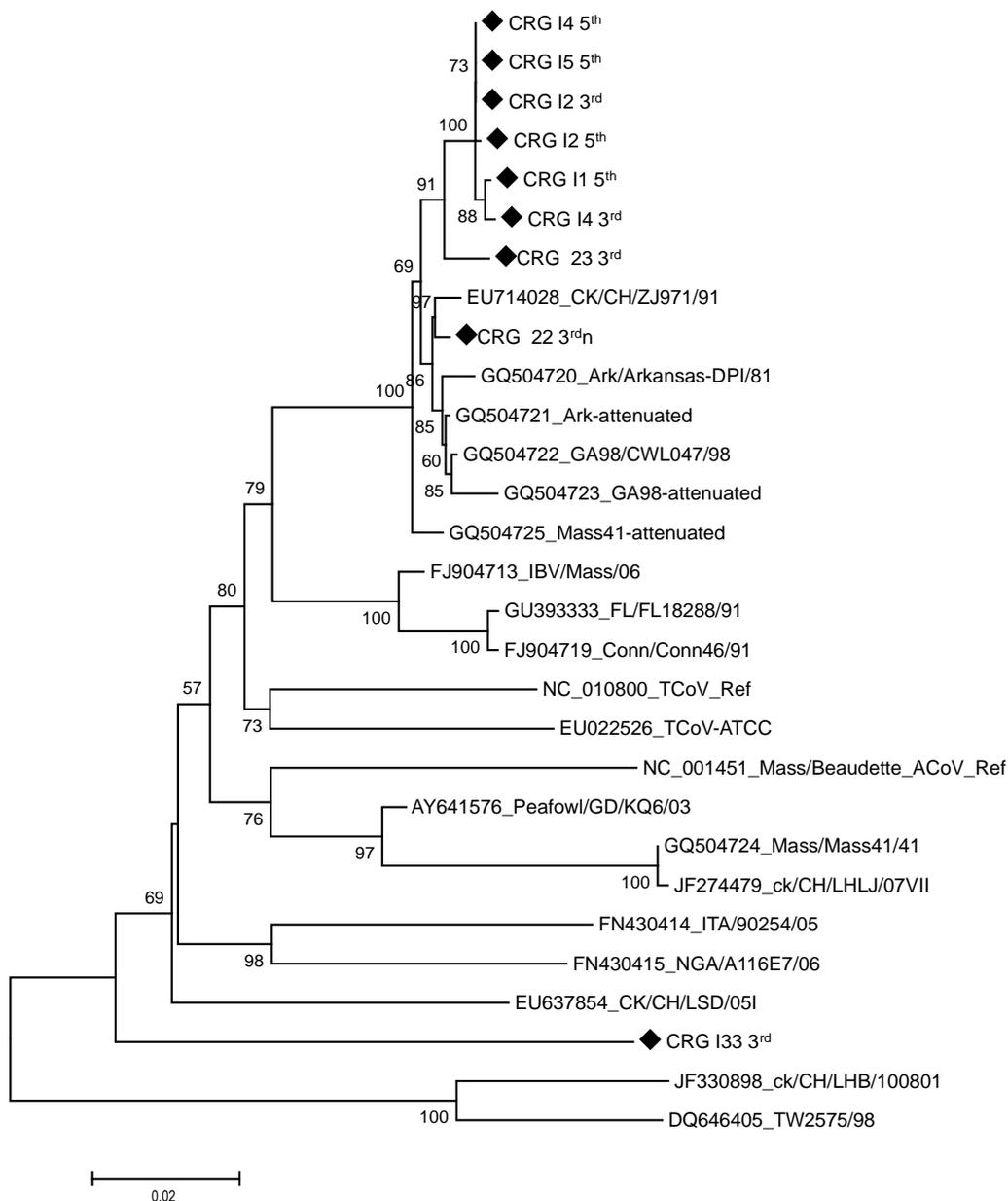
As sequências nucleotídicas contendo sítios de polimorfismo foram anotadas de acordo com as recomendações da IUPAC (*International Union of Pure and Applied Chemistry*) para indicação de degeneração de nucleotídeos. As sequências que informam sobre os sítios de degeneração serão empregadas para o depósito em bancos de dados.

Ainda com relação aos sítios de degeneração, foi possível verificar que, para esses sítios de polimorfismo, havia uma população viral dominante (ocorrendo em maior frequência), assim sendo, além da anotação com os nucleotídeos degenerados, foi realizada uma anotação apenas com nucleotídeos sequenciados em maior frequência. Essas sequências foram designadas como dominantes, e foram empregadas na reconstrução das árvores filogenéticas e para os cálculos de identidades entre as sequências.

A figura 5 apresenta a árvore filogenética obtida a partir das sequências completas obtidas para o gene nsp2. Para as sequências incluídas na reconstrução dessa árvore, a identidade nucleotídica média foi de 92,28% ( $\pm 4,37\%$ ). Quando apenas as sequências obtidas no presente trabalho foram consideradas, a identidade nucleotídica média foi de 96,96% ( $\pm 4,5\%$ ).

Adicionalmente, no item apêndices, estão apresentadas uma árvore filogenética incluindo todas as sequências obtidas para os isolados brasileiros de ACoV (Apêndice D) e uma árvore reconstruída a partir das maiores sequências obtidas (Apêndice E), que incluem parte da região 5' do genoma (anterior à ORF1a) e a porção inicial do gene nsp3 (correspondendo ao domínio ácido).

Figura 5- Árvore filogenética par o gene nsp2, para os nucleotídeos 529 a 2.547 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette - número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

Na árvore reconstruída, é possível verificar que a maioria dos isolados brasileiros de ACoV foram um grupamento exclusivo, suportado por um valor de bootstrap de 91. Esse grupamento constitui um subgrupo de um *cluster* maior de sequências (suportado por um valor de bootstrap de 100), que inclui o isolado CRG I22 3<sup>a</sup> passagem, linhagens vacinais e

atenuadas de ACoVs dos sorotipos Arkansas e Georgia, além da linhagem vacinal do sorotipo Massachusetts e a linhagem CK/CH/ZJ971/91 (#EU714028).

É interessante observar o agrupamento, para o gene *nsp2*, de linhagens de ACoV bastante distintas na genotipagem pelo gene *s1*, como as linhagens do sorotipo Georgia e Arkansas. Sob esse aspecto, cabe destacar o isolado CRG I22 3<sup>a</sup> passagem, que para *s1* agrupou-se com linhagens de ACoV pertencentes ao sorotipo Massachusetts, e para *nsp2*, a aproximou-se dos demais isolados brasileiros (o mesmo valendo para a sequência CK/CH/ZJ971/91, #EU714028).

Além disso, é possível observar que o isolado CRG I33 3<sup>a</sup> passagem, pertencente ao genótipo brasileiro de ACoV (à análise do gene *s1*), segregou-se de maneira independente com relação ao gene *nsp2*.

Na árvore apresentada no apêndice D, é possível observar o padrão de segregação do isolado CRG I38 3<sup>a</sup> passagem (para o qual não se obteve sequência de *S1*). É interessante observar que esse isolado também se segregou de maneira independente do *cluster* brasileiro, de modo semelhante ao observado para CRG I33 3<sup>a</sup> passagem.

Adicionalmente às sequências restritas ao gene da *Nsp2*, foi possível obter sequências contíguas à região inicial do gene *nsp3*, incluindo grande parte do domínio ácido de *nsp3*. Curiosamente, verificou-se uma deleção de 30 nucleotídeos (do nucleotídeo 2.918 ao 2.947) na região inicial de *nsp3*, para os isolados CRG I1 5<sup>a</sup> passagem, CRG I3 3<sup>a</sup> passagem, CRG I22 3<sup>a</sup> passagem, CRG I23 3<sup>a</sup> passagem e CRG I38 3<sup>a</sup> passagem. Essa mesma deleção foi também observada nas sequências de Ma05 vacina, Mass41-attenuated (#GQ504725), CK/CH/ZJ971/91 (#EU714028), Ark-attenuated (#GQ504721), Ark/Arkansas-DPI/81 (#GQ504720) e Peafowl/GD/KQ6/03 (#AY641576). Exceção dentre os isolados brasileiros foi o isolado CRG I33 3<sup>a</sup> passagem, que não apresentou essa deleção de 30 nucleotídeos. A árvore de distância construída para o segmento final de *nsp2* e inicial de *nsp3* (nucleotídeos 1.956 a 2.993, em relação ao genoma Mass/Beaudette-ACoV Ref, #NC\_0151) está apresentada no apêndice E. Também para essa região contígua aos genes *nsp2* e *nsp3*, foi possível verificar que o isolado CRG I33 3<sup>a</sup> passagem, mantém um padrão independente de segregação.

### 5.2.3 Gene nsp3

A região codificadora da protease Papaína-*like* constituiu o alvo, no gene nsp3, de maior interesse para o presente trabalho. Entretanto, justamente esse segmento do gene se revelou uma região bastante desafiadora para a obtenção de *amplicons* e para o sequenciamento.

Foi possível obter a sequência completa do domínio da protease papaína-*like* para os isolados CRG I1 para a 3<sup>a</sup> e 5<sup>a</sup> passagens (3<sup>rd</sup> e 5<sup>th</sup>), CRG I3 5<sup>a</sup> passagem, CRG I5 5<sup>a</sup> passagem, CRG I23 3<sup>a</sup> passagem e CRG I33 3<sup>a</sup> passagem. Embora não tenha havido o sequenciamento total do gene codificador da papaína-*like* para os outros isolados, foi possível sequenciar partes do mesmo, incluindo-se segmentos contíguos a ele, tanto anteriores (correspondendo ao domínio da ADP-ribose 1 fosfatase), quanto posteriores (correspondentes a uma região gênica de função desconhecida). O quadro 8 contém a anotação das sequências obtidas para o gene nsp3, destacando-se a porcentagem de cobertura para o domínio codificador da protease papaína-*like*.

Quadro 8- Anotação das sequências obtidas para o gene nsp3, com relação ao genoma do ACoV linhagem Beaudette (número de acesso no GenBank NC\_001451). Para alguns isolados obtiveram-se sequências da 3<sup>a</sup> (3rd) e 5<sup>a</sup> (5th) passagens em ovos embrionados. A porcentagem de cobertura do domínio codificador da protease papaína-*like* também está apresentado

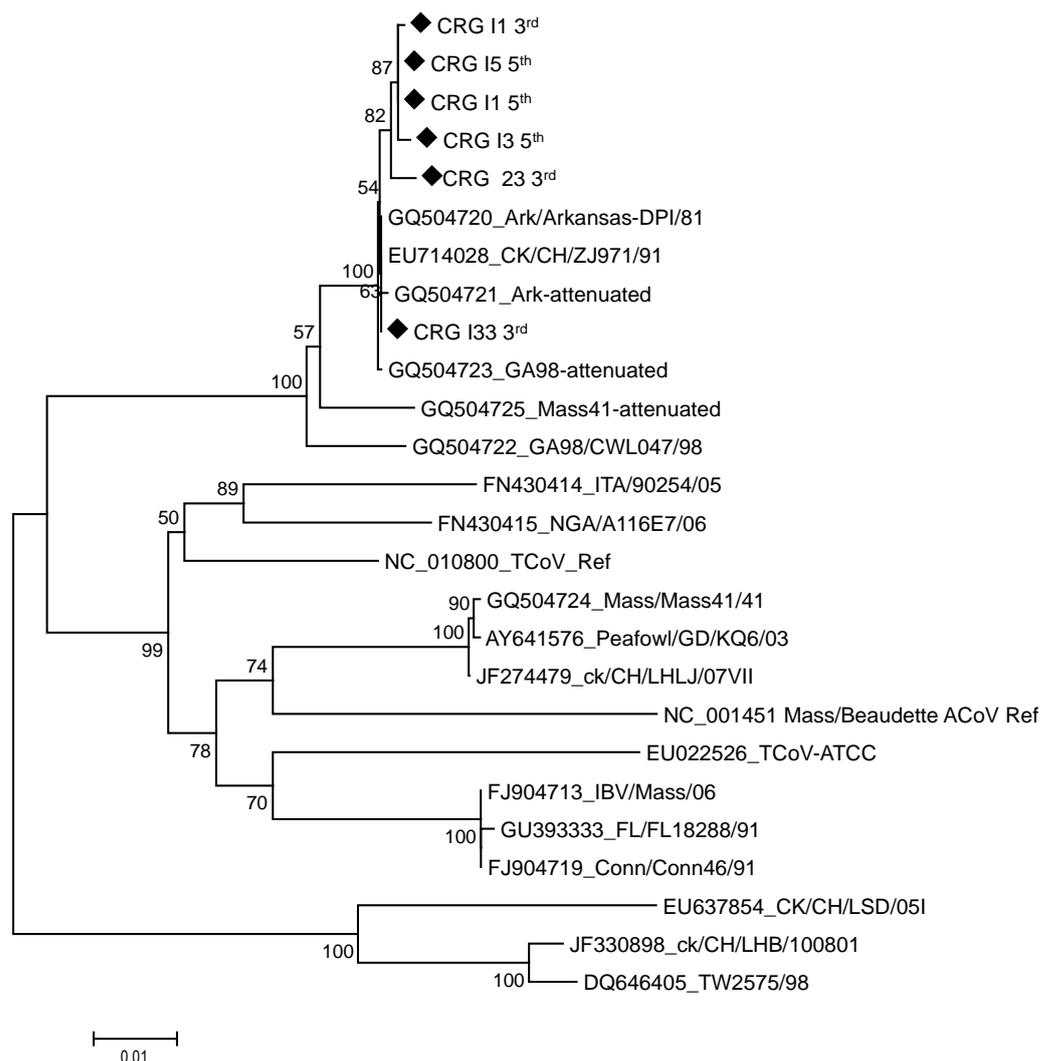
Gene	Isolado	nucleotídeo inicial	nucleotídeo final	Comprimento (pares de bases)	% de cobertura (em relação PL2 completa)
nsp3	CRG-I1 3rd	3.684	5.632	1.948	100,00%
	CRG-I1 5th	3.680	5.626	1.946	100,00%
	CRG-I2 3rd	4.180	5.664	1.484	85,75%
	CRG-I2 5th	3.658	4.807	1.149	84,51%
	CRG-I3 3rd	3.615	4.410	795	39,96%
		4.567	5.013	446	42,31%
	CRG-I3 5th	3.623	5.012	1.389	100,00%
	CRG-I4 3rd	3.761	4.335	574	31,54%
		4.570	5.630	1.060	41,98%
	CRG-I4 5th	4.402	5.617	1.215	60,83%
	CRG-I5 3rd	4.181	5.618	1.437	85,63%
	CRG-I5 5th	3.630	5.620	1.990	100,00%
	CRG-I22 3rd	4.182	5.660	1.478	85,52%
	CRG-I23 3rd	3.617	5.625	2.008	100,00%
	CRG-I30 3rd	4.569	5.019	450	42,09%
	CRG-I33 3rd	3.634	5.674	2.040	100,00%
	CRG-I38 3rd	3.664	4.377	713	36,25%
	Ma05 (vacina)	3.649	4.387	738	37,37%
4.403		5.020	617	60,72%	

Fonte: (AYRES, G. R. R., 2014).

A árvore filogenética reconstruída a partir das sequências completas do gene *plpro* está apresentada na figura 6. A identidade nucleotídica média verificada para as sequências incluídas nessa árvore foi de 88,6% (com um desvio padrão de 5,7).

A topologia observada para a árvore apresentada na figura 6, se mantém para as demais árvores construídas para o gene *nsp3*, e que estão apresentadas nos apêndices G, H e I. Dentre os apêndices, destaca-se o apêndice G, cuja árvore contempla um segmento de 1.939 nucleotídeos do gene *nsp3* (nucleotídeos 3.777 a 5.715, em relação ao primeiro nucleotídeo do genoma do ACoV linhagem Beaudette #NC\_00145).

Figura 6- Árvore filogenética para a porção codificadora da protease papaína-*like* contida no gene *nsp3*, para os nucleotídeos 4.054 a 4.944 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

Na árvore correspondente ao domínio plpro do gene nsp3 (Figura 6), é possível perceber que, assim como para o gene nsp2, a maioria dos isolados brasileiros de ACoV formam um grupamento exclusivo, suportado por um valor de bootstrap de 82. Esse grupamento constitui um subgrupo do *cluster* maior (suportado por um valor de bootstrap de 100), que inclui o isolado CRG I33 3<sup>a</sup> passagem, linhagens vacinais e atenuadas de ACoVs dos sorotipos Arkansas e Georgia, além da linhagem vacinal do sorotipo Massachusetts e a linhagem CK/CH/ZJ971/91 (#EU714028). Assim sendo, vale ressaltar que o isolado CRG I33 3<sup>a</sup> passagem, para o domínio plpro do gene nsp3, manteve o perfil topológico similar ao observado para a árvore do gene s1 (Figura 4).

No apêndice H, é possível observar que o isolado CRG I38 3<sup>a</sup> passagem também agrupou-se no *cluster* dos demais isolados brasileiros, diferentemente do que foi observado para o gene da Nsp2 (Apêndice D).

## ***DISCUSSÃO***

## 6 DISCUSSÃO

Ao que se saiba, esse é o primeiro trabalho de investigação de genes codificadores de proteínas não estruturais, especificamente nsp2 e nsp3, que inclui linhagens de ACoV detectados na América Latina (especificamente no Brasil).

Análises genéticas prévias que tenham incluído a ORF1a têm sido realizados com linhagens Norte Americanas, Asiáticas, Africanas e Europeias de ACoV (PHILLIPS et al., 2012, MONDAL; CARDONA, 2004, MCKINLEY et al., 2011, THOR et al., 2011, LIU et al., 2013). Os resultados verificados para as análises filogenéticas das sequências obtidas do GenBank estão em acordo com aqueles previamente reportados, quanto à topologia das linhagens de ACoV nas árvores filogenéticas e as identidades nucleotídicas médias.

Foi possível notar que diferentes genes (e domínios gênicos) da primeira janela aberta de leitura (ORF1a) do genoma dos ACoV podem demonstrar padrões divergentes à análise filogenética.

Para a maior parte dos ACoV sequenciados no presente trabalho, foi possível observar uma marcante proximidade genética com ACoV isolados nos Estados Unidos (virulentos e atenuados), para as sequências de nsp2 e nsp3 (Figuras 5 e 6). Nos Estados Unidos, têm sido relatadas elevadas identidades genéticas entre sequências da ORF1ab (também conhecida como complexo de replicação e transcrição – RTC) obtidas de ACOVs isolados de amostras de campo e ACOVs vacinais, independentemente do soro-genotipo viral. Por essa razão, têm-se sugerido um padrão convergente para a evolução da ORF1ab. (MONDAL; CARDONA, 2004). Os resultados aqui apresentados também demonstram proximidade genética entre a linhagem vacinal aqui sequenciada (Ma05 vacinal) e as linhagens brasileiras de ACoV.

Estudos de sequências genômicas completas de ACOVs permitiram a detecção de prováveis eventos de recombinação entre vírus provenientes da Ásia e dos Estados Unidos. Os autores desse trabalho (THOR et al., 2011) minimizaram a possibilidade de os achados decorrerem propriamente de eventos de recombinação. Segundo eles, a proximidade genética observada para linhagens de ACoV dos Estados Unidos, China e Taiwan podem ser explicados pela evolução a partir de uma ancestral comum, possivelmente um ACoV do soro-genótipo Massachusetts (Mass) vacinal – atenuado.

Essa hipótese poderia explicar a similaridade verificada para sequências de nsp2 e nsp3 entre linhagens de ACoV aqui sequenciadas, americanas e CK/CH/ZJ971/97, uma vez que, no Brasil, apenas vacinas produzidas com vírus do sorotipo Massachusetts são autorizadas pelo Ministério da Agricultura Pecuária e Abastecimento (MAPA). Entretanto, essa hipótese não explica as discrepâncias no padrão de agrupamento verificado para as linhagens virais CRG I22 3<sup>a</sup> passagem, CRG I33 3<sup>a</sup> passagem e CRG I38 3<sup>a</sup> passagem.

O ACoV CRG I33 3<sup>a</sup> passagem agrupou-se com os demais vírus sequenciados no presente trabalho para os genes s1 e nsp3 (Figuras 4 e 6, respectivamente), mas apresentou um padrão independente de segregação para o gene nsp2 (Figura 5).

Por sua vez, o CRG I22 3<sup>a</sup> passagem agrupou-se no *cluster* de vírus brasileiros para nsp3 e nsp2 (Figuras 5 e 6, respectivamente), mas foi divergente para S1, agrupando-se com ACOVs pertencentes ao soro-genotipo Massachusetts, de linhagens vacinais e virulentas (Figura 4).

Quanto ao isolado CRG I38 3<sup>a</sup> passagem, embora a genotipagem pelo gene s não tenha sido obtida, foi possível observar padrões de segregação para os genes nsp2 e nsp3 semelhantes ao verificado para o isolado CRG I33 3<sup>a</sup> passagem (Apêndices D e H, respectivamente), tendo-se agrupado com os demais isolados brasileiros para a árvore do gene nsp3 e de maneira independente para o gene nsp2. Nesses casos, a possibilidade de terem ocorrido eventos de recombinação não podem ser descartadas.

O genes nsp2, 3 e 16 foram apontados, juntamente com o gene s, como *hotspots* para transferência de segmentos genômicos (THOR et al., 2011). Assim sendo, para CRG I33 3<sup>a</sup> passagem (e possivelmente para CRG I38 3<sup>a</sup> passagem), pode-se especular que se trate de uma linhagem viral resultante de recombinação, cuja maior parte do genoma seria proveniente de ACoV regionais (brasileiros), com possível evento de recombinação do gene nsp2 (herdado de um ACoV não definido, mas próximo a linhagens virais orientais – africanas e/ou asiáticas).

Para CRG I22 3<sup>a</sup> passagem, é possível fazer inferência semelhante, assumindo-se a recombinação de um genoma proveniente de linhagens virais regionais, com possível transferência do segmento S1, provavelmente proveniente de vírus do soro-genotipo Massachusetts.

Já foi descrita, em condições de campo, a ocorrência de recombinação natural entre ACOVs patogênicos e vacinais, em linhagens chinesas do soro-genotipo Massachusetts (LIU et al., 2013). Esses achados suportam a hipótese de as discrepâncias verificadas nas topologias das árvores para os diferentes genes serem decorrentes de eventos de recombinação.

Além disso, eventos de recombinação, somados às mutações pontuais e a seleção positivas já foram relatados como base para a diversidade molecular encontrada em linhagens brasileiras de ACoV (BRANDÃO et al., 2009).

Para que se comprovem os mecanismos envolvidos na emergência e disseminação dos ACOVs é necessário o sequenciamento de genomas completos de linhagens representativas desses vírus, entre outros estudos. Entretanto, a proximidade genética aqui verificada para linhagens de ACoV provenientes de continentes distintos pode ser decorrente tanto do acúmulo de mutações pontuais, a partir de um ancestral comum, quanto da ocorrência de eventos de recombinação entre vírus de campo e vacinais. Isso porque esse padrão reticulado de evolução tem sido descrito para os ACOVs (THOR et al., 2011; LIU et al., 2013; JACKWOOD; HALL; HANDEL, 2012).

O conhecimento sobre os padrões de evolução, que levam à diversidade verificada para ACOVs, tem sido apontado como uma ferramenta no desenvolvimento de estratégias de controle das infecções (LIU et al., 2013).

O sequenciamento do gene *s* tem constituído o principal método empregado para a genotipagem de linhagens de ACoV. Sabe-se que ACOVs pertencentes a um mesmo sorotipo podem apresentar elevada variabilidade em outras regiões genômicas, externas ao gene *s*, especialmente ao longo da ORF1a (AMMAYAPPAN et al., 2009; PHILLIPS et al., 2012). Uma vez que os genes *nsp2* e *3* já tenham sido apontados como regiões de frequente transferência genômica (THOR et al., 2011), o sequenciamento e análise multigênica, empregando-se os genes *s*, *nsp2* e *3*, podem constituir ferramentas úteis para a detecção de diversidade genética em ACOVs, quando o sequenciamento do genoma completo não estiver disponível.

Além disso, devido às potenciais consequências fenotípicas da ocorrência de variantes virais para os genes *nsp2* e *3*, o monitoramento da diversidade nesses alvos pode auxiliar na melhor compreensão acerca dos mecanismos de disseminação e patogênese viral.

Quanto às identidades médias e desvios padrões verificados para os genes estudados, os resultados obtidos estão em acordo com os previamente descritos. Apesar de os genes *nsp2* e *3* serem apontados como regiões de elevada variabilidade intra sorotípica (AMMAYAPPAN et al., 2009; PHILLIPS et al., 2012), na comparação entre linhagens de sorotipos divergentes o gene *s* continua a ser o alvo de maior variabilidade (PHILLIPS et al., 2012).

Não foi verificado nenhum indício de que os genes estudados possam servir como marcadores de tropismo para as linhagens de ACoV estudadas, por não se ter verificado diversidade gênica correspondente ao tipo de órgão do qual o ACoV foi isolado.

Entretanto, ficou clara a divergência de padrão filogenético encontrada para 3 das amostras aqui sequenciadas (CRG I22 3<sup>a</sup> passagem, CRG I33 3<sup>a</sup> passagem e CRG I38 3<sup>a</sup> passagem), quanto a um dos genes estudados. Esses achados podem ser indicadores da ocorrência de recombinação entre linhagens de ACoV de campo e linhagens vacinais, eventos já relatados anteriormente em outras regiões geográficas.

Além disso, foi possível verificar que, em acordo com a literatura previamente reportada, comparando-se diferentes soro-genotipos de ACoV, as sequências de gene s1 são as mais divergentes (gene s1: 82,09%  $\pm$ 13,48; gene nsp2: 96,96%  $\pm$ 4,5; gene plpro: 99,57%  $\pm$ 2,35).

Os resultados alcançados contribuem para um conhecimento mais amplo da evolução dos ACoV, e apontam os genes nsp2 e nsp3 como alvos também interessantes para estudos de monitoramento epidemiológico molecular desses vírus.

É possível que as discrepâncias encontradas para algumas linhagens de ACoV sejam resultantes de eventos de recombinação entre ACOVs brasileiros, ACOVs vacinais e outros ainda não conhecidos. Entretanto, a comprovação dessa inferência requer um maior conhecimento sobre a diversidade dos genes nsp2 e nsp3 de ACoV e, idealmente, sobre a diversidade de genomas completos dos ACOVs.

***CONCLUSÃO***

## 7 CONCLUSÃO

Os genes codificadores da proteína não estrutural 2 (Nsp2) e da protease papaína-*like* (Plpro - contida na Nsp3) e da proteína estrutural S1, segundo as regiões gênicas aqui estudadas, não são preditivos de tropismo e patogenicidade de amostras de ACoV. Entretanto, os resultados aqui apresentados reforçam a circulação de um genótipo brasileiro de ACoV, para os três genes estudados e apontam indícios de recombinação para estas regiões genômicas.

## ***REFERÊNCIAS***

## REFERÊNCIAS

AMMAYAPPAN, A., UPADHYAY, C., GELB JR, J., VAKHARIA, N. V. Identification of sequence changes responsible for attenuation of avian infectious bronchitis virus strain Arkansas DPI. **Archives of Virology**, v. 154, n. 3, p. 495-499, 2009.

ARMESTO, M.; CAVANAGH, D.; BRITTON, P. The replicase gene of avian coronavirus infectious bronchitis virus is a determinant of pathogenicity. **PLoS One**, v. 4, n. 10, e7384, 2009. Disponível em:

<<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0007384>>.

Acesso em: 15 de julho de 2010. doi:<10.1371/journal.pone.0007384>.

BRANDÃO, P. E.; SANDRI, T. L.; SOUZA, S. P.; KUANA, S. L.; RICHTZENHAIN, L. J.; VILLARREAL, L. Y. recombination, point mutations and positive selection on the basis of the molecular diversity of brazilian strains of avian infectious bronchitis virus. In: 6, 2009, Marburg, INTERNATIONAL SYMPOSIUM ON AVIAN CORONA AND PNEUMOVIRUSES, **Anais...** Marburg, 2009, p. 47-58.

BRITTON, P.; CAVANAGH, D. Avian coronavirus diseases and infectious bronchitis virus vaccine development. In: THIEL, V. **Coronaviruses: molecular and cellular biology**, Norfolk: Caister Academic Press, 2007. p. 161-180.

CAVANAGH, D.; MAWDITT, K.; WELCHMAN, 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.

CAVANAGH, D. Coronavirus avian infectious bronchitis virus. **Veterinary Research**, v. 38, p. 281-197, 2007.

CAVANAGH, D.; GELB, J. J. Infectious bronchitis. In: SAIF, Y.M.; FADLY, A.M.; GLISSON, J.R.; McDOUGALD, L.R.; NOLAN, L.K. SWAYNE, D.E. **Diseases of poultry**, 12. ed. Ames: Iowa State University Press, 2008. p. 117-135.

CHOU, Y. W.; CHENG, S. C.; LAI, H. Y.; CHOU, C. Y. Differential domain structure stability of the severe acute respiratory syndrome coronavirus papain-like protease. **Archives Of Biochemistry and Biophysics**, v. 502, n. 2, p. 74-80, 2012. Disponível em :<<http://www.sciencedirect.com/science/article/pii/S0003986112000732>>. Acesso em: 10 de julho de 2010. doi:<10.1016/j.abb.2012.02.15>.

CLEMENTZ, M. A.; CHEN, Z.; BANACH, B. S.; WANG, Y.; SUN, L.; RATIA, K.; BAEZ-SANTOS, Y. M.; WANG, J.; TAKAYAMA, J. Deubiquitinating and interferon antagonism

activities of coronavirus papain-like proteases. **Journal of Virology**, v. 84, p. 4619-4629, 2010.

ETERRADOSSI, N.; BRITTON, P. Infectious Bronchitis Virus. In: World Organization for Animal Health – OIE. **Manual of diagnostic tests and vaccines for terrestrial animals 2011**. Disponível em: <[http://www.oie.int/fileadmin/Home/eng/Health\\_standards/tahm/2.03.02\\_AIB.pdf](http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.03.02_AIB.pdf)> . Acesso em: 07 de janeiro de 2011.

ERIKSSON, K. K.; CERVANTES-BARRAGAN, L.; LUDEWIG, B., THIEL, V. Mouse hepatitis virus liver pathology is dependent on ADP-ribose-1<sup>st</sup>-phosphatase, a viral function conserved in the alpha-like subgroup. **Journal of Virology**, v. 82, p. 12325-12334, 2008.

EWING, B., GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome Research**, v. 8, p. 186–194, 1998.

GRAHAM, R. L.; SIMS, A. C.; BROCKWAY, S. M.; BARIC, R. S.; DENINSON, M. R. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. **Journal of Virology**, v. 79, n. 21, p. 13399-13411, 2005.

HAGEMEIJER, M. C.; VERHEIJE, M. H., ULASLI, M.; SHATIEL, I. A.; de VRIES, L. A.; REGGIORI, F.; ROTTIER, P. J.; DE HAAN, C. A. Dynamics of coronavirus replication-transcription complexes. **Journal of Virology**, v. 84, p. 2134-2149, 2010.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analyses program for Windows 95/98/NT. **Nucleic Acids Symposium**, n. 41, p. 95–98, 1999.

Instituto Brasileiro de Geografia e Estatística (IBGE). Disponível em: <[http://www.ibge.com.br/home/presidencia/noticias/noticia\\_visualiza.php?id\\_noticia=634;id\\_pagina=1](http://www.ibge.com.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia=634;id_pagina=1)>. Acesso em: 06 de junho de 2014.

ICTV- International Committee on Taxonomy of Viruses. **Virus Taxonomy**: 2013 Release. Disponível em: <<http://www.ictvonline.org/virusTaxonomy.asp>>. Acesso em 06 de julho de 2014.

JACKWOOD, M. W., HALL, D., HANDEL, A. Molecular evolution and emergence of avian gammacoronaviruses. **Infection, Genetics and Evolution**, v. 12, n. 6, p. 1305-1311, 2012.

LIM, K. P.; NG, L. F.; LIU, D. X. Identification of a novel cleavage activity of the first papain-like proteinase domain encoded by open reading frame 1a of the coronavirus avian infectious bronchitis virus and characterization of the cleavage products. **Journal of Virology**, v. 74, p. 1674-1685, 2000.

LIU, X., SHAO, Y., MA, H., SUN, C., ZHANG, X., LI, C., HAN, Z., YAN, B., KONG, X., LIU, S. Comparative analysis of four Massachusetts type strain and evidence of natural recombination in the genome. **Infection, Genetics and Evolution**, v. 14, p. 29-38, 2013

MASTERS, P. S. The molecular biology of coronaviruses. **Advances in Virus Research**, v. 66, p. 193-292, 2006.

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.

MONDAL, S. P.; CARDONA, C. J. Comparison of four regions in the replicase gene of heterologous infectious bronchitis virus strains. **Virology**, v. 324, p. 238-248, 2004.

MONTASSIER, H. J. Molecular epidemiology and evolution of avian infectious bronchitis virus. In: 2010, Campinas, WORKSHOP: INFECTIOUS BRONCHITIS (IB) IN BRAZILIAN POULTRY INDUSTRY, **Anais...** Campinas, 2010, v. 12, n. 12, p. 87-96.

OvoOnline: Somatória de Pintainhas Vermelhas e Brancas. Disponível em: <<http://www.ovoonline.com.br/?:=alojamento&tt=atd>>. Acesso em 06 de julho de 2014.

PHILLIPS, J. E., JACKWOOD, M. W., MCKINLEY, E. T., THOR, S. W., HILT, D. A., ACEVEDO N. D., WILLIAMS, S. M., KISSINGER, J. C., PATERSON, A. H., ROBERTSON, J. S., LEMKE, C. Changes in nonstructural protein 3 are associated with attenuation in avian coronavirus infectious bronchitis virus. **Virus Genes**, v. 44, n. 1, p. 63-74, 2012.

ROTH-CROSS, J. K.; STOKES, H.; CHANG, G.; CHUA, M. M.; THIEL, V.; WEISS, S. R.; GORBALENYA, A. N.; SIDDELL, S. G. Organ-specific attenuation of murine hepatitis virus strain A59 by replacement of catalytic residues in the putative viral cyclic phosphodiesterase ns2. **Journal of Virology**, v. 83, n. 8, p. 3743-3753, 2009.

SAWICKI, S. G.; SAWICKI, D. L.; SIDDELL, S. G. A contemporary view of coronavirus transcription. **Journal of Virology**, v. 81, n. 1, p. 20-29, 2007.

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

TAN, Y. J.; LIM, S. G.; HONG, W. Characterization of viral proteins encoded by the SARS-coronavirus genome. **Antiviral Research**, v. 65, p. 69-78, 2005.

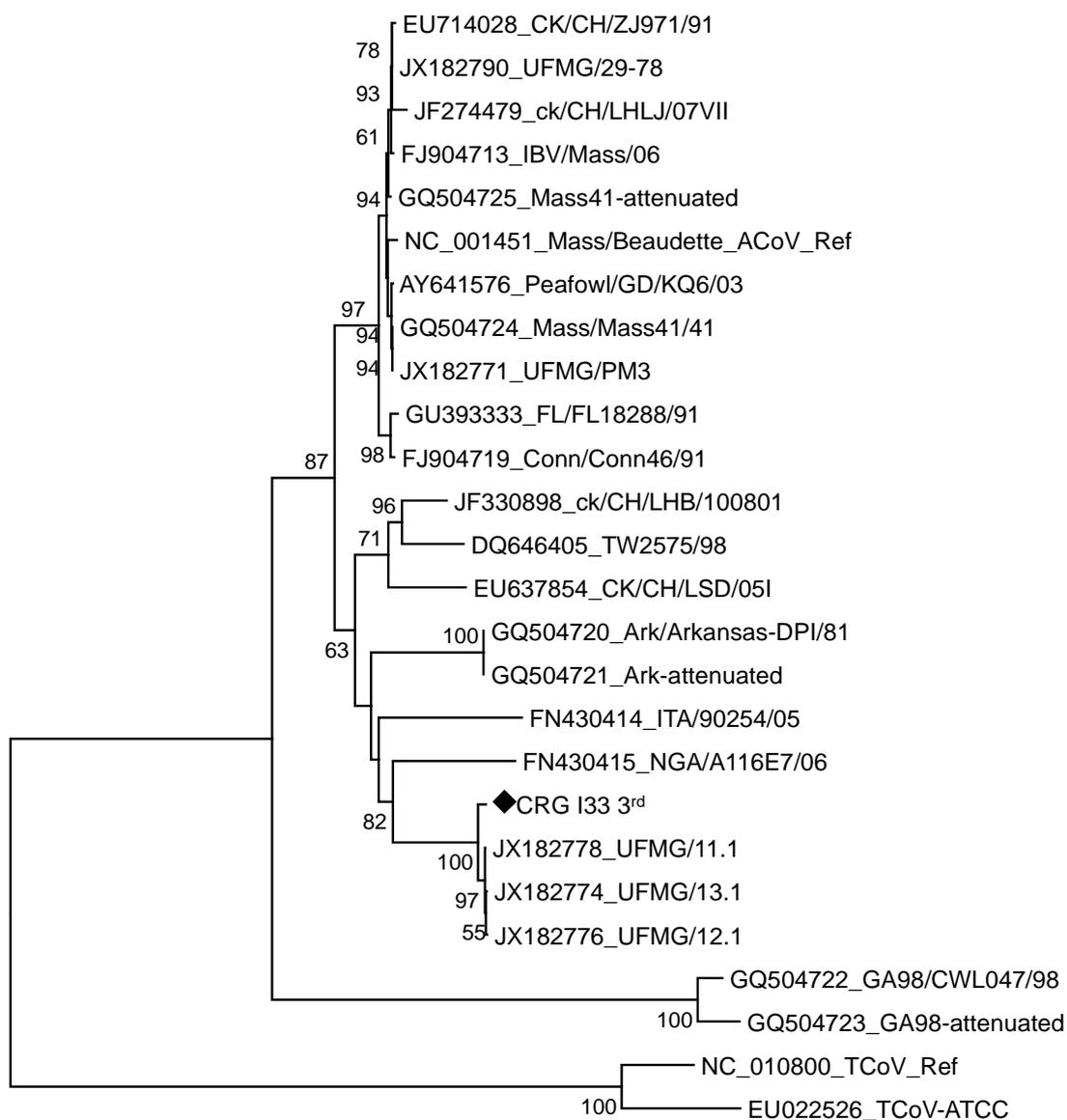
- THIEL, V. Reverse genetic analysis of coronavirus replication. In: THIEL, V. **Coronaviruses: molecular and cellular biology**. Norfolk: Caister Academic Press: 2007. p. 161-180.
- THOR, S. W.; HILT, D. A.; KISSINGER, J. C.; PATERSON, A. H.; JACKWOOD, M. W. Recombination in avian gamma-coronavirus infectious bronchitis virus. **Viruses**, v. 3, n. 9, p. 1777-1799, 2011.
- UNIÃO BRASILEIRA DE PROTEÍNA ANIMAL. Disponível em:  
<[http://www.ubabef.com.br/a\\_avicultura\\_brasileira/historia\\_da\\_avicultura\\_no\\_brasil](http://www.ubabef.com.br/a_avicultura_brasileira/historia_da_avicultura_no_brasil)>. Acesso em 08 de julho de 2014.
- VILLARREAL, L. Y.; SANDRI, S. P.; SOUZA, S. P.; RICHTZENHAIN, L. J.; de WIT, J. J.; BRANDÃO, P. E. Molecular epidemiology of avian infectious bronchitis in Brazil from 2007 to 2008 in breeders, broilers, and layers. **Avian Diseases**, v. 54, p. 894-898, 2010.
- WORTHINGTON, K. J., CURRIE, R. J., JONES, R. C. A reverse transcriptase-polymerase chain reaction survey of infectious bronchitis virus genotypes in Western Europe from 2002 to 2006. **Avian Pathology**, v. 37, p. 247-257, 2008.
- YANG, A.; WEI, L.; ZHAO, W.; XU, Y.; RAO, Z. Expression, crystallization and preliminary X-ray diffraction analysis of the N-terminal domain of nsp2 from avian infectious bronchitis virus. **Acta Crystallographica**, v. 65, p. 788-790, 2009.
- YU, K.; MING, Z.; LI, Y.; CHEN, C.; BAO, Z.; REN, Z.; LIU, B.; TAO, W.; RAO, Z.; LOU, Z. Purification, crystallization and preliminary X-ray analysis of nonstructural protein 2 (nsp2) from avian infectious bronchitis virus. **Acta Crystallographica, section F**, n. 68, p. 716-719, 2012.
- ZHENG, D.; CHEN, G.; GUO, B.; CHENG, G.; TANG, H. PLP2, a potent deubiquitinase from murine hepatitis virus, strongly inhibits cellular type I interferon production. **Cell Research**, v. 18, p. 1105-1113, 2008.
- ZIEBUHR, J.; SNIJDER, E. J. The coronavirus replicase gene: special enzymes for special viruses. In: THIEL, V.; **Coronaviruses: molecular and cellular biology**, Norfolk: Caister Academic Press, 2007. p. 161-180.

***APÊNDICES***

## APÊNDICES

### APÊNDICE A - ÁRVORE FILOGENÉTICA CONSTRUÍDA COM A SEQUÊNCIA DO GENE S1 DO ISOLADO CRG I33 3<sup>A</sup> PASSAGEM (FIGURA 7), COBRINDO 98,06% DO GENE COMPLETO.

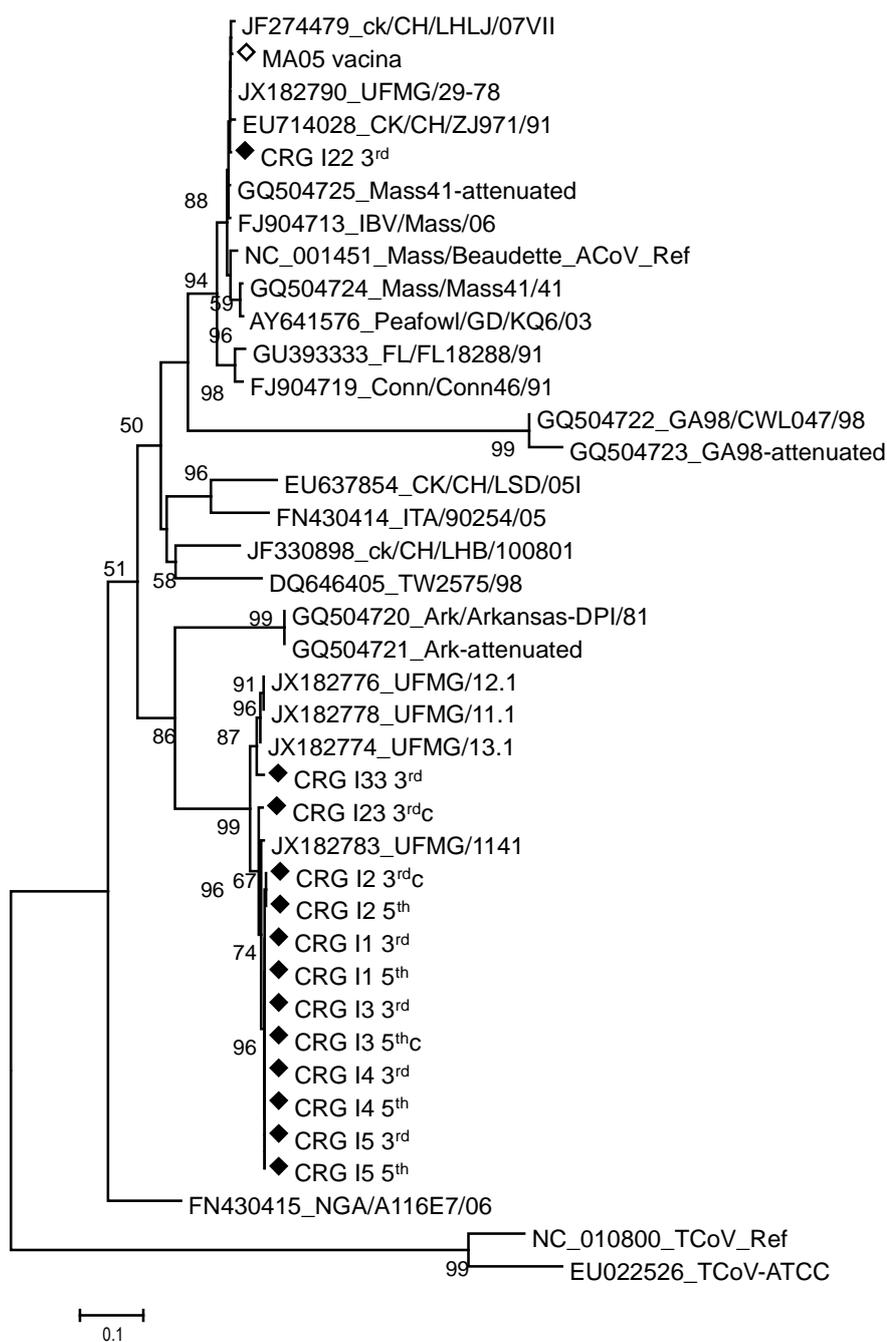
Figura 7. Árvore filogenética do gene s1, para os nucleotídeos 20.368 a 22.257 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE B - ÁRVORE FILOGENÉTICA CONSTRUÍDA COM SEQUENCIAS CONSENSUAIS PARA O INÍCIO DO GENE S1 (FIGURA 8).**

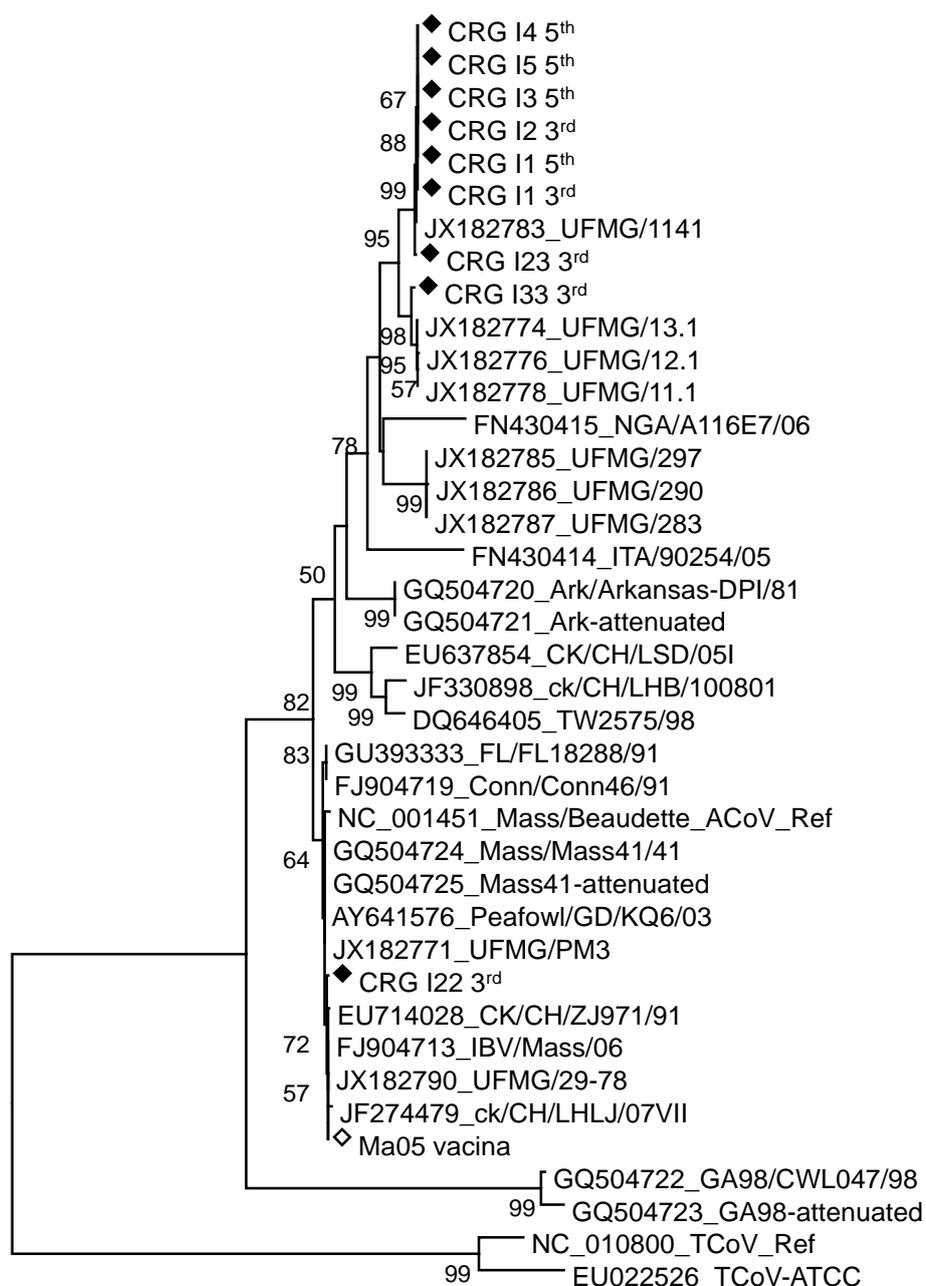
Figura 8. Árvore filogenética do gene s1, para os nucleotídeos 20.368 a 21.046 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE C - ÁRVORE FILOGENÉTICA CONSTRUÍDA COM SEQUENCIAS CONSENSUAIS PARA A PORÇÃO FINAL DO GENE S1 (FIGURA 9).**

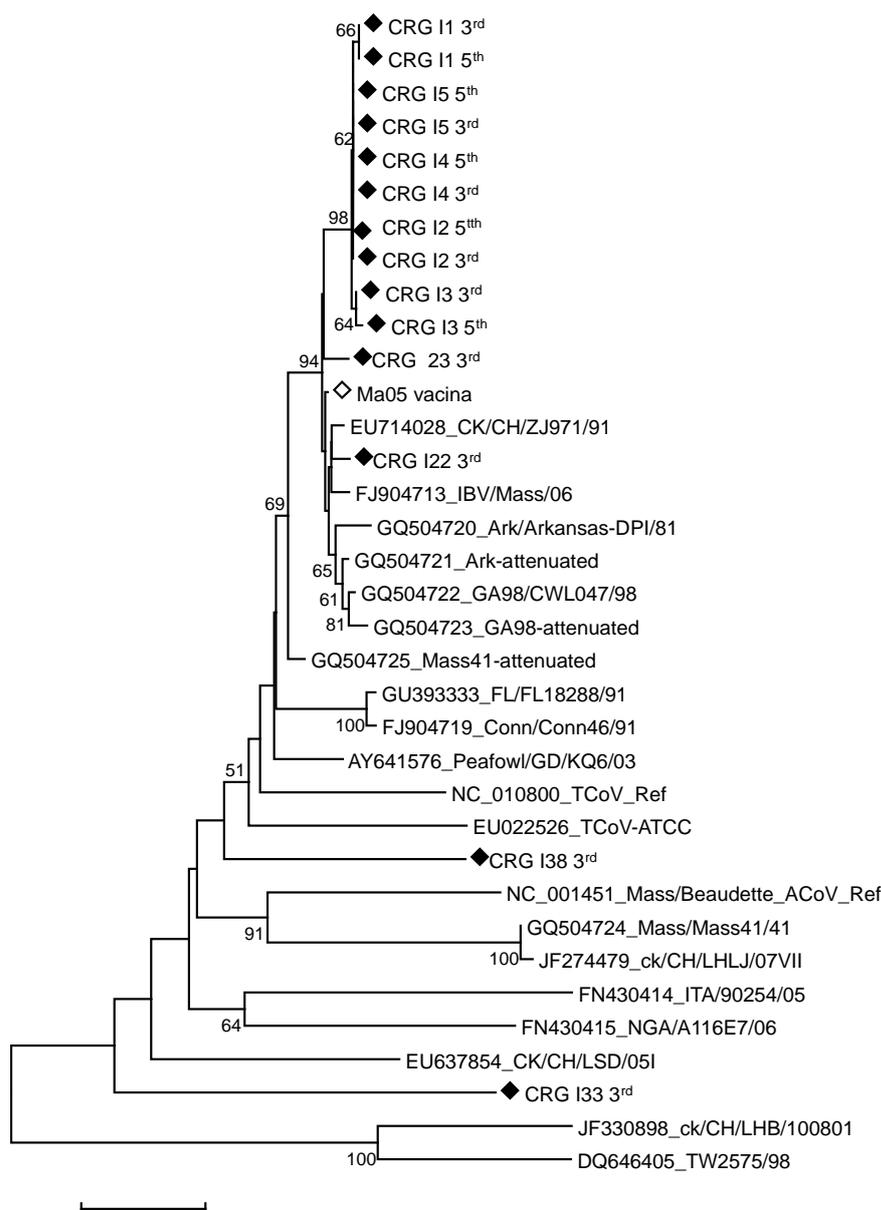
Figura 9. Árvore filogenética do gene s1, para os nucleotídeos 21.336 a 22.247 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE D - ÁRVORE FILOGENÉTICA PARA O GENE NSP2 RECONSTRUÍDA COM O SEGMENTO GÊNICO HOMÓLOGO A TODAS AS SEQUENCIAS OBTIDAS DOS ISOLADOS BRASILEIROS (QUE INCLUI PARTE DA REGIÃO 5' NÃO TRANSCRITA DO GENOMA E A PORÇÃO INICIAL DA NSP2), ABAIXO APRESENTADA NA FIGURA 10.**

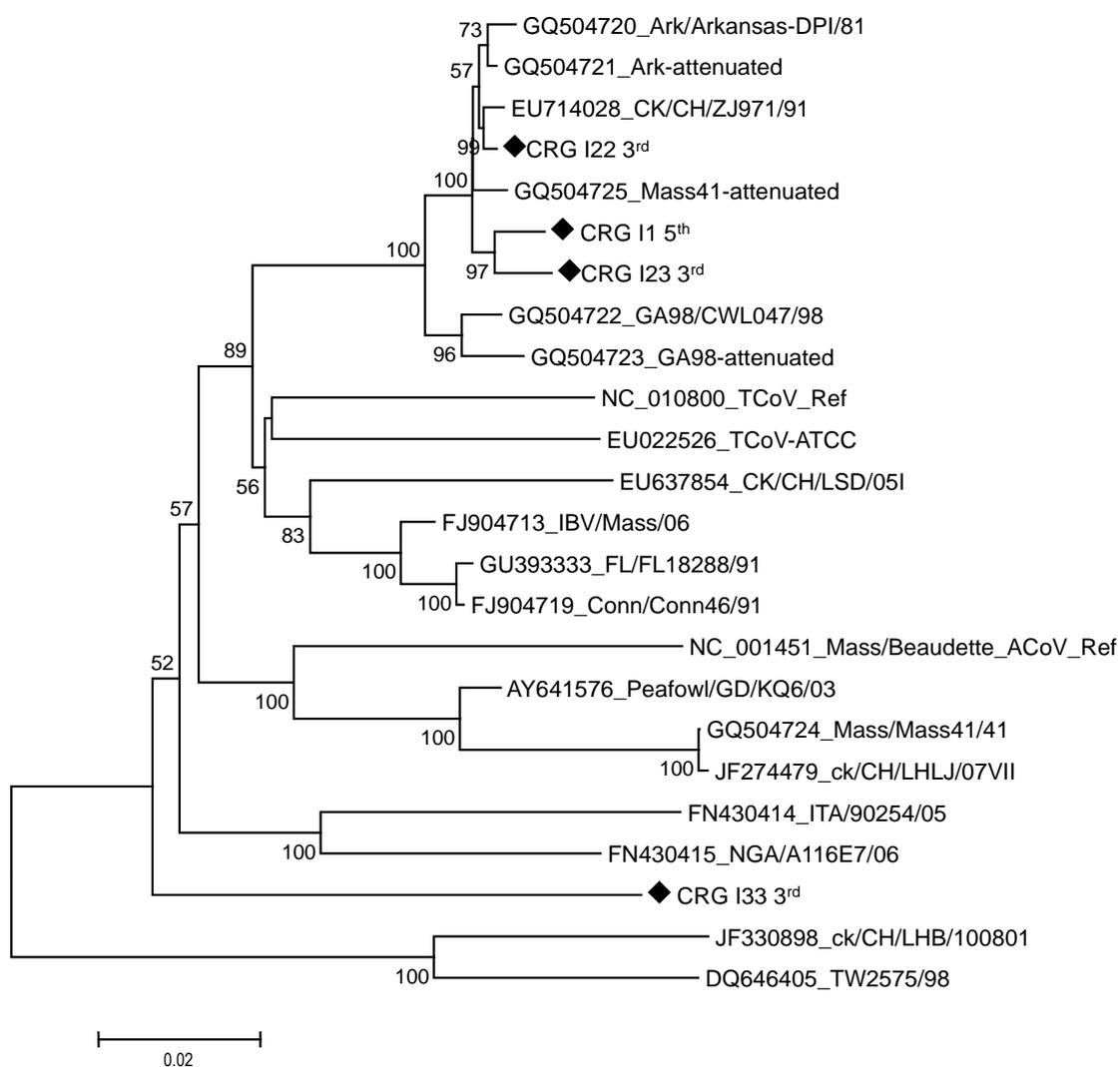
Figura 10. Árvore filogenética do gene *nsp2*, para os nucleotídeos 505 a 1.569 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (◊) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE E - A ÁRVORE FILOGENÉTICA INCLUINDO AS MAIORES SEQUÊNCIAS DE NUCLEOTÍDEOS PARA O GENE NSP2 (INCLUINDO OS SEGMENTOS CONTÍGUOS DA REGIÃO 5' NÃO TRANSCRITA DO GENOMA, ANTERIOR A NSP2, E O SEGMENTO INICIAL DE NSP3, CORRESPONDENTE AO DOMÍNIO ACÍDICO), ESTÁ APRESENTADA NA FIGURA 11.**

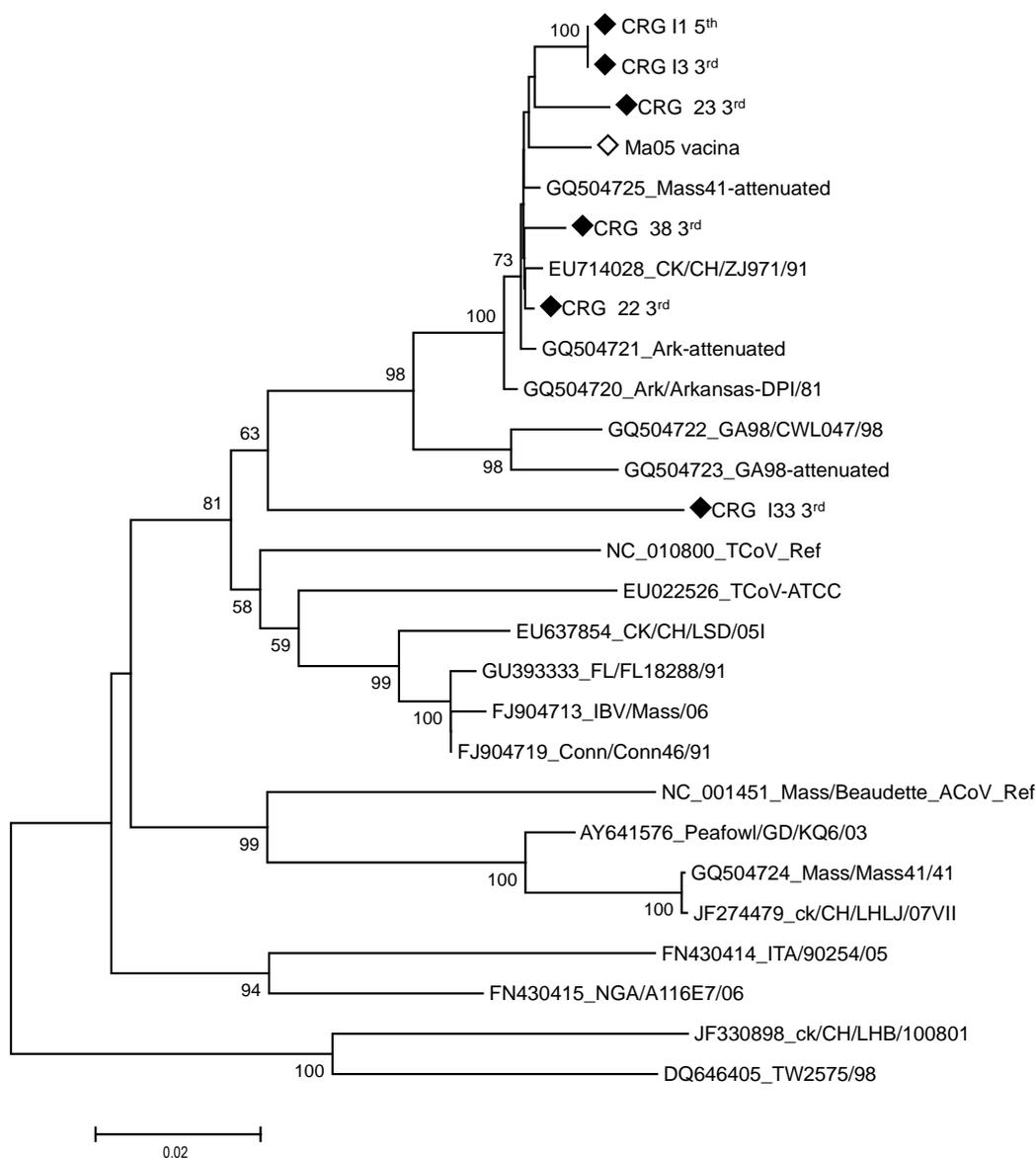
Figura 11. Árvore filogenética do gene nsp2, para os nucleotídeos 503 a 2.992 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE F - ÁRVORE FILOGENÉTICA (FIGURA 12) INCLUINDO TODAS AS SEQUÊNCIAS OBTIDAS PARA OS ISOLADOS BRASILEIROS DE ACOV E UMA ÁRVORE RECONSTRUÍDA A PARTIR DAS MAIORES SEQUÊNCIAS OBTIDAS, QUE INCLUEM PARTE DA REGIÃO 5' DO GENOMA (ANTERIOR À ORF1A), A TOTALIDADE DO GENE NSP2 E A PORÇÃO INICIAL DO GENE NSP3 (CORRESPONDENDO AO DOMÍNIO ACÍDICO).**

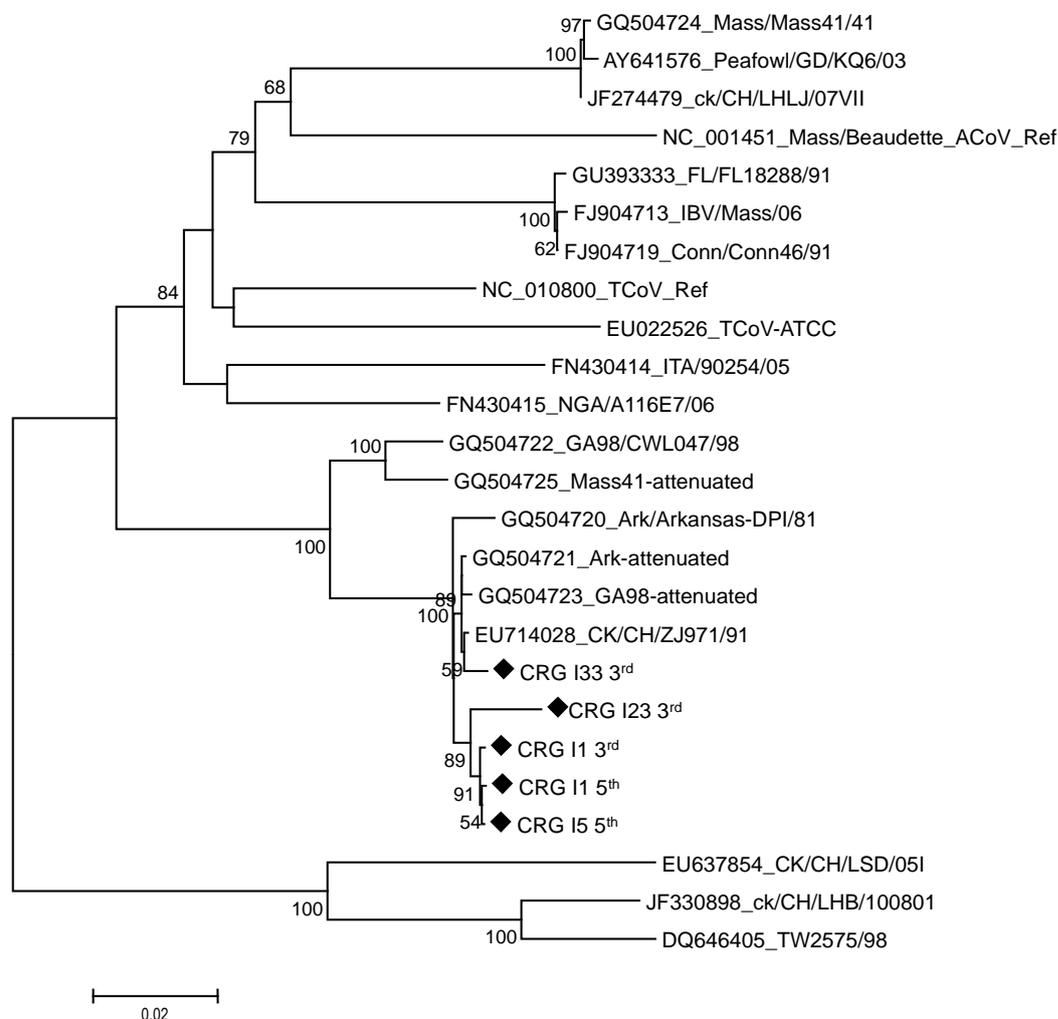
Figura 12. Árvore filogenética para a região contínua dos genes nsp2 e nsp3, para os nucleotídeos 1.956 a 2.993 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE G - ÁRVORE RECONSTRUÍDA COM AS MAIORES SEQUENCIAS OBTIDAS PARA NSP3, INCLUINDO PARTE DO DOMÍNIO CODIFICADOR DA FOSFATASE 1 ADP-RIBOSE, A TOTALIDADE DO DOMÍNIO CODIFICADOR DA PROTEASE PAPAÍNA-LIKE E PARTE DA REGIÃO GÊNICA DE FUNÇÃO DESCONHECIDA (FIGURA 13).**

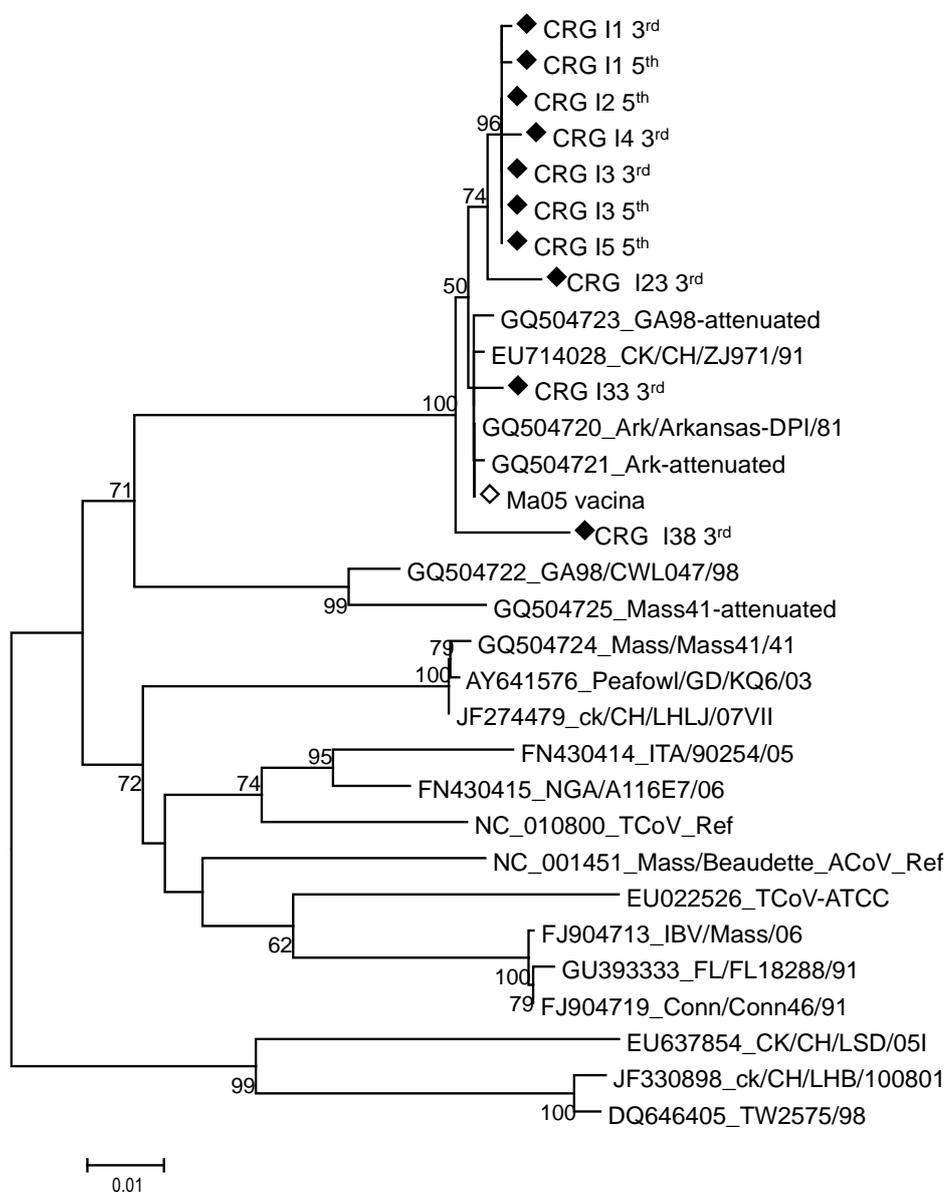
Figura 13. Árvore filogenética para o gene *nsp3*, para os nucleotídeos 3.777 a 5.715 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE H - ÁRVORE RECONSTRUÍDA COM SEQUENCIAS OBTIDAS PARA NSP3, INCLUINDO PARTE DO DOMÍNIO CODIFICADOR DA FOSFATASE 1 ADP-RIBOSE E PORÇÃO INICIAL DO DOMÍNIO CODIFICADOR DA PROTEASE PAPAÍNA-LIKE (FIGURA 14).**

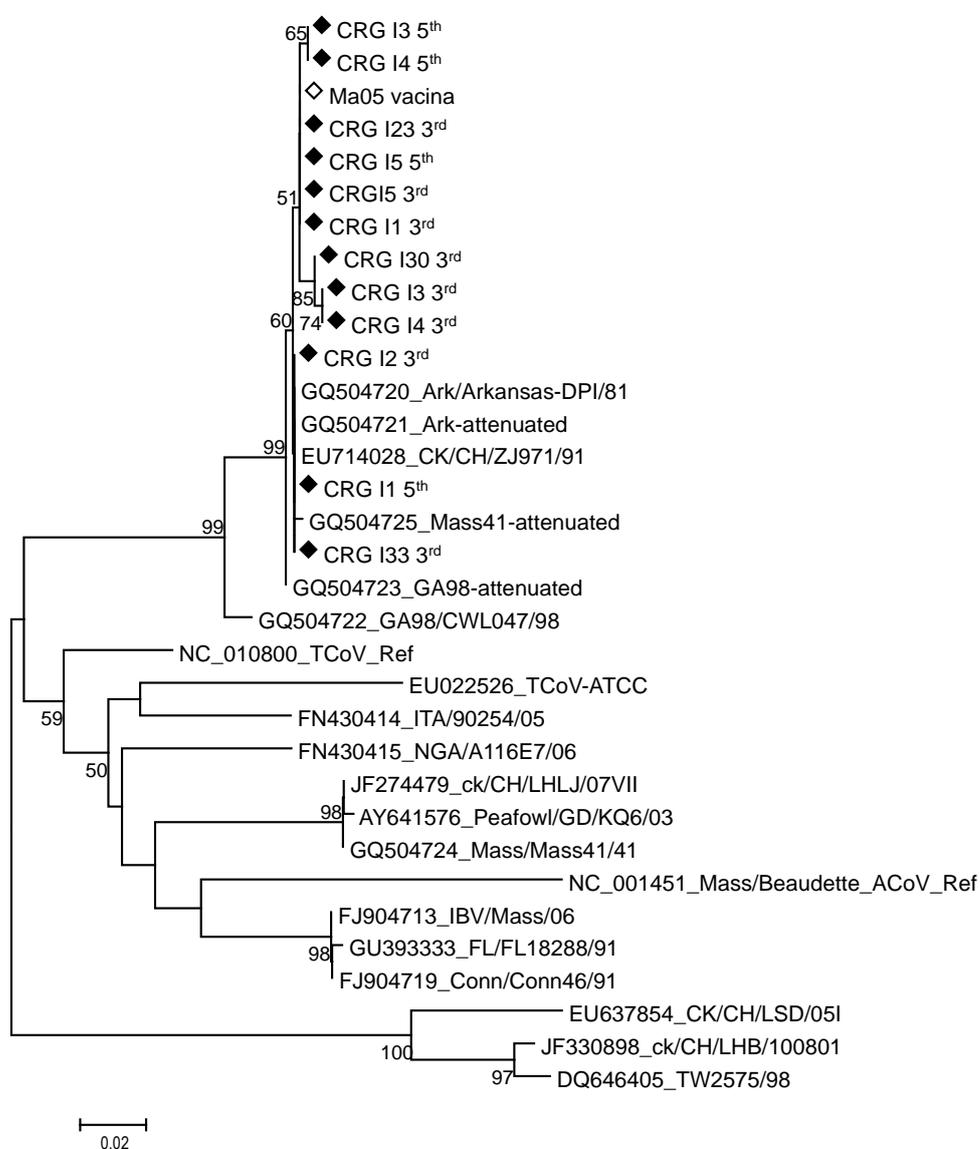
Figura 14. Árvore filogenética para o gene *nsp3*, para os nucleotídeos 3.854 a 4.428 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).

**APÊNDICE I - ÁRVORE RECONSTRUÍDA A PARTIR DAS SEQUENCIAS OBTIDAS PARA NSP3, INCLUINDO PARTE DO DOMÍNIO CODIFICADOR DA PROTEASE PAPAÍNA-LIKE E PARTE DA REGIÃO GÊNICA DE FUNÇÃO DESCONHECIDA (FIGURA 15).**

Figura 15. Árvore filogenética para o gene nsp3, para os nucleotídeos 4.660 a 5.108 (em relação ao primeiro nucleotídeo do genoma ACoV Beaudette número de acesso no GenBank NC\_001451) construída pelo algoritmo de neighbor-joining e modelo evolutivo Maximum Composite Likelihood, com 1000 repetições de bootstrap. Os losangos (♦) indicam as amostras sequenciadas no presente trabalho



Fonte: (AYRES, G. R. R., 2014).