

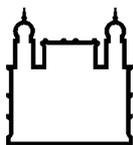
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Mestrado em Programa de Pós-Graduação em Biologia Parasitária

AVALIAÇÃO DA DIVERSIDADE DE QUASIESPÉCIES DO
HEPACIVÍRUS EQUINO A PARTIR DA REGIÃO DE ENVELOPE NO
CONTEXTO DA EVOLUÇÃO DA INFECÇÃO E LOCALIZAÇÃO
GEOGRÁFICA NO BRASIL

PEDRO PEREIRA LIRA FURTADO DE ALBUQUERQUE

Rio de Janeiro
Março de 2020



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Parasitária

PEDRO PEREIRA LIRA FURTADO DE ALBUQUERQUE

Avaliação da diversidade de quasiespécies do hepacivírus equino a partir da região de envelope no contexto da evolução da infecção e localização geográfica no Brasil

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Biologia Parasitária

Orientador: Prof. Dr. Marcelo Alves Pinto

Coorientadora: Dra. Andreza Soriano Figueiredo Nunes

RIO DE JANEIRO

Março de 2020

Pereira Lira Furtado de Albuquerque, Pedro .

Avaliação da diversidade de quasiespécies do hepacivírus equino a partir da região de envelope no contexto da evolução da infecção e localização geográfica no Brasil / Pedro Pereira Lira Furtado de Albuquerque. - Rio de Janeiro, 2020.

73 f.; il.

Dissertação (Mestrado) - Instituto Oswaldo Cruz, Pós-Graduação em Biologia Parasitária, 2020.

Orientador: Marcelo Alves Pinto.

Co-orientadora: Andreza Soriano Figueiredo Nunes.

Bibliografia: f. 64-70

1. Hepacivirus equino. 2. Quasiespécies. 3. Região Hipervariável. 4. Envelope. 5. Diversidade. I. Título.

Anexar a cópia da Ata que será entregue pela SEAC já assinada.

Dedico este trabalho aos alunos que não se sentem capazes. Não se deixem enganar.

v

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pelo auxílio financeiro.

À equipe do Laboratório de Desenvolvimento Tecnológico em Virologia, meus orientadores e meus colegas de bancada

Ao Instituto Oswaldo Cruz e ao curso de Biologia Parasitária

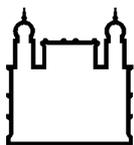
Aos membros da banca de defesa da dissertação.

Ao Centro de Apoio ao Discente

Aos colaboradores deste projeto, Gonzalo Belo e Suwellen Dias

À minha turma de mestrado desde março de 2018

Não se deixe viver no ideal, viva no real e no possível. Conheça seus limites.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

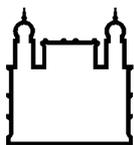
Avaliação da diversidade de quasiespécies do hepacivírus equino a partir da região de envelope no contexto da evolução da infecção e localização geográfica no Brasil

RESUMO

DISSERTAÇÃO DE MESTRADO EM BIOLOGIA PARASITÁRIA

Pedro Pereira Lira Furtado de Albuquerque

O *Hepacivirus A* (hepacivírus equino – EqHV) foi o primeiro vírus do gênero *Hepacivirus* (Família *Flaviviridae*) descrito para um hospedeiro mamífero não humano. Descrito em infecção de cavalos, cães e muares, possui uma ampla distribuição global, co-circulando como dois subtipos virais sem qualquer associação geográfica. O animal infectado pelo EqHV desenvolve quadros clínicos de hepatite e a infecção pode evoluir para cronicidade. No entanto, o desenvolvimento para infecção crônica constitui uma minoria dos casos. O genoma do EqHV possui regiões hipervariáveis na região do envelope viral e uma dinâmica populacional de quasiespécies. Essa complexa dinâmica é associada a muitas dificuldades clínicas no tratamento, desenvolvimento de vacinas e características biológicas do vírus. Estudar esta complexidade populacional no contexto de diferentes evoluções da infecção e localizações geográficas, isto é, de animais que eliminaram ou que evoluíram para a infecção crônica, pode elucidar características importantes da evolução e distribuição dos hepacivírus. Desta forma, este projeto teve por objetivo acessar a população de quasiespécies do EqHV em equinos que eliminaram a infecção ou evoluíram para infecção crônica, bem como estimar a diversidade genética entre equinos de diferentes regiões geográficas do Brasil. Para a análise das quasiespécies, foi realizada a amplificação, sequenciamento e clonagem molecular do envelope viral abrangendo a região hipervariável do EqHV (400 pb), e um total de 236 clones foram obtidos. As populações de quasiespécies do animal crônico mostraram-se significativamente diferentes em relação aos animais que eliminaram a infecção, inclusive quando analisadas as sequências majoritárias (sequenciamento direto do produto de PCR). As sequências dos produtos de PCR dos quatro animais que eliminaram a infecção foram idênticas entre si, e entre a maioria dos clones obtidos desses animais. Esse resultado não foi observado para nenhuma das coletas do animal com infecção crônica. Além disso, não houve relação entre a localização geográfica e a distância genética inter-populacional do EqHV. As variações não sinônimas se concentraram em regiões de conhecida hipervariabilidade, sendo mais frequentes no animal crônico. Ademais, os dois pontos de coleta do animal crônico mostraram dinâmica de população de quasiespécies acelerada. A análise filogenética não distinguiu populações de animais crônicos, de eliminação ou de diferentes localizações geográficas. Em conclusão, verificou-se associação entre diversidade viral e persistência da infecção em que uma população homogênea tende à eliminação viral, enquanto uma população de maior diversidade tende à persistência. Não houve relação entre a distância geográfica e a distância genética, sugerindo distribuição antropogênica do vírus. Entender a dinâmica populacional do EqHV envolve benefícios que podem abranger a área médica veterinária e humana, contribuindo também na agenda de eliminação das hepatites virais humanas da Organização Mundial da Saúde (OMS).



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

EVALUATION OF THE DIVERSITY OF EQUINE HEPACIVIRUS QUASIESPECIES FROM THE ENVELOPE REGION IN THE CONTEXT OF THE EVOLUTION OF INFECTION AND GEOGRAPHICAL LOCATION IN BRAZIL

ABSTRACT

MASTER DISSERTATION IN PARASITIC BIOLOGY

Pedro Pereira Lira Furtado de Albuquerque

Hepacivirus A (equine hepacivirus - EqHV) was the first virus of the genus *Hepacivirus* (Family *Flaviviridae*) described for a non-human mammal host. Described in infection of horses, dogs and muars, it has a wide global distribution, co-circulating as two viral subtypes without any geographical association. Animals infected by EqHV develop clinical symptoms of hepatitis and infection may evolve into chronicity. However, the development to chronic infection constitutes a minority of cases. EqHV genome has hypervariable regions in the viral envelope genes and a quasispecies population dynamics. Such complex dynamics is associated with many clinical difficulties in treatment, vaccine development and biological characteristics of the virus. Studying this population complexity in the context of different infection evolutions and geographical locations, i.e., animals that have eliminated or evolved to chronic infection, may elucidate important characteristics of hepacivirus evolution and distribution. Thus, this project aimed to assess the population of the EqHV quasispecies in horses from the same farm that either cleared the infection or progressed into chronic infection, as well as the genetic diversity between horses from different geographical regions in Brazil. For the analysis of the quasispecies, amplification, sequencing and molecular cloning of the viral envelope portion covering the hypervariable region of EqHV (400 bp) were performed, and a total of 236 clones were obtained. The populations of quasispecies of the chronic animal were significantly different from those that cleared the infection, even when the majority sequences were analyzed (direct sequencing of the PCR product). The sequences of the PCR products of the four animals that cleared the infection were identical among themselves, and among the most frequent clone population obtained from these animals. This result was not observed for any of the samples of the chronically infected animal. Moreover, there was no relationship between the geographical location and the inter-population genetic distance of EqHV. The non-synonymous variations were concentrated in regions of known hypervariability and were more frequent in the chronic animal. Moreover, the two samples of the chronically infected animal revealed accelerated quasispecies dynamics. Phylogenetic analysis did not distinguish populations of cleared, chronically infected animals or of different geographic locations. In conclusion, there was an association between viral diversity and persistence of infection in which a homogeneous population tends to viral elimination, while a population of greater diversity tends to persist. There was no relationship between geographic distance and genetic distance, suggesting anthropogenic distribution of the virus. Understanding the population dynamics of EqHV involves benefits that can cover the veterinary and human medical field, also contributing to the World Health Organization (WHO) agenda for the elimination of human viral hepatitis.

ÍNDICE

RESUMO	VIII
ABSTRACT	IX
1 INTRODUÇÃO	15
1.1 O Hepacivirus A	15
1.1.1 O genoma	17
1.1.2 A epidemiologia	19
1.1.3 A clínica	20
1.1.4 A diversidade	22
1.2 Quasiespécies	23
1.2.1 Conceito.....	23
1.2.2 Aplicações na Virologia.....	25
1.2.3 Como acessar a diversidade de quasiespécies	27
1.2.4 A Dinâmica de Quasiespécies dos hepacivírus	28
1.3 Justificativa	31
2 OBJETIVOS	32
2.1 Objetivo Geral	32
2.2 Objetivos Específicos	32
3 MATERIAL E MÉTODOS	33
3.1 População de estudo	33
3.2 Definição do alvo molecular e construção de oligonucleotídeos iniciadores	34
3.3 Extração de RNA viral e síntese de cDNA	37
3.4 Amplificação do alvo gênico E2 do hepacivírus A	37
3.5 Clonagem molecular do produto amplificado	38
3.5.1 Preparo de células competentes	38
3.5.2 Adenilação	38
3.5.3 Ligação	38
3.5.4 Transformação.....	39
3.5.5 Teste das colônias.....	39
3.5.6 Extração de material plasmidial	40
3.6 Sequenciamento	40

3.7	Análise das sequências nucleotídicas	40
3.7.1	Cálculo da distância genética	40
3.7.2	Análises estatísticas	40
3.7.3	Análise filogenética	41
4	RESULTADOS	42
4.1	Otimização e padronização do alvo molecular construído neste trabalho	42
4.2	Clonagem Molecular	43
4.2.1	Cálculo do erro intrínseco à técnica aplicada	43
4.2.2	Análise das populações virais obtidas	44
4.3	Análise Filogenética	53
5	DISCUSSÃO	56
6	CONCLUSÕES	62
7	PERSPECTIVAS	63
8	REFERÊNCIAS	64
9	APÊNDICES E/OU ANEXOS	71

ÍNDICE DE FIGURAS

Figura 1 Genoma do HCV	17
Figura 2. Sobrevivência do mais achatado.....	30
Figura 3: Esquema da amplificação por <i>semi-Nested</i> PCR da região do envelope viral do EqHV	36
Figura 4. Otimização da temperatura de hibridização em matriz de agarose	42
Figura 5. Gráfico de caixa da distância de Hamming.....	46
Figura 6 Gráfico de caixa da p-distância intra-populacional.....	47
Figura 7 Gráfico de caixa da p-distância intra-populacional entre localidades.	48
Figura 8 Gráfico de caixa da p-distância inter-populacional.....	49
Figura 9 Gráfico de balão da p-distância entre as sequências majoritárias	50
Figura 10. Gráficos de p-distância e localização genômica	51
Figura 11. Gráficos de p-distância e localização genômica no animal crônico.	52
Figura 12 Gráfico de p-distância e localização genômica das sequências majoritárias: Crônico v Eliminação	53
Figura 13. Árvore filogenética de populações de quasiespécies do EqHV.....	54
Figura 14: Árvore filogenética de populações majoritárias de EqHV	55

LISTA DE TABELAS

Tabela 1: População de estudo.....	34
Tabela 2: Carga viral (cópias/mL) do animal crônico.....	34
Tabela 3: Sequências do EqHV coletadas do GenBank.....	35
Tabela 4: Oligonucleotídeos iniciadores construídos para região de envelope viral do EqHV.....	36
Tabela 5. Distribuição de clones e variantes	44

LISTA DE SIGLAS E ABREVIATURAS

AST – Aspartato aminotransferase

BEAST – *Bayesian Evolutionary Analysis Sampling Trees*

CEUA - Comissão de Ética no Uso de Animais

CHV – hepacivirus canino

DNA – Ácido desoxirribonucleico (Desoxirribonucleic Acid)

EqHV – hepacivirus equino ou *Hepacivirus A*

HCV – *Vírus da Hepatite C (Hepatitis C vírus)* ou *Hepacivirus C*

HVR – Região Hipervariável (*Hypervariable Region*)

IBGE - Instituto Brasileiro de Geografia e Estatística

IRES – Região interna de entrada do ribossomo (Internal Ribosomal Entry Site)

Kb – Kilobases

LB – Lúria Berthani (Ágar)

MAVS - Proteína de sinalização antiviral mitocondrial (Mitochondrial Anti-viral Signaling protein)

MEGA - *Molecular Evolutionary Genetics Analysis*

NC – Região não-codificante do genoma

NGS – Sequenciamento de Nova Geração (New Generation Sequencing)

NPHV – hepacivirus não-primata

OMS – Organização Mundial da Saúde

ORF – Matriz de leitura aberta (Open Reading Frame)

PCR – Reação de Polimerização em Cadeia (Polymerase Chain Reaction)

RdRp – RNA polimerase RNA dependente (RNA dependent RNA polymerase)

RNA – Ácido ribonucleico (Ribonucleic Acid)

RNAm – RNA mensageiro

UFRRJ – Universidade Federal Rural do Rio de Janeiro

1 INTRODUÇÃO

O gênero *Hepacivirus*, da família *Flaviviridae*, é exemplo de um dos táxons favorecidos pelo desenvolvimento de novas tecnologias de sequenciamento e processamento de dados biológicos, denominados sequenciamento de nova geração (*New Generation Sequencing* - NGS), levando à descrição de novos vírus (1–3). A espécie tipo deste grupo foi descrita em 1989 (4) como um importante agente etiológico de hepatite não-A e não-B, posteriormente vindo a ser chamada de hepatite C, dando nome à espécie *Hepatitis C Virus* (HCV). Sua relevância clínica e epidemiológica trouxe consigo grandes avanços na biologia do vírus, embora uma vacina ainda seja elusivo (5). Por muitos anos, o humano foi o único hospedeiro representado neste gênero, até a detecção, por NGS, de um vírus com uma configuração genética muito semelhante, que a relação filogenética acusava grande proximidade (3) e o quadro clínico se assemelha ao HCV (6,7). Atualmente, são reconhecidas no gênero *Hepacivirus* 14 espécies virais, nomeadas de A a N, contendo uma gama de hospedeiros mamíferos (8) e até peixes cartilagosos (9).

1.1 O *Hepacivirus A*

Inicialmente o hepacivirus equino (EqHV) foi detectado na traqueia de cães albergados em dois canis independentes onde ocorriam surtos de infecção respiratória (3), recebendo o nome de hepacivírus canino (CHV). Posteriormente, em estudos epidemiológicos por via sorológica envolvendo equinos e caninos, o agente foi detectado em equinos, mas não em cães amostrados no estudo (10), passando a ser denominado de hepacivírus não primata (NPHV) e ampliando a possibilidade de hospedeiros para o vírus. A reorganização do gênero *Hepacivirus* resultou na classificação do vírus que infecta equinos e caninos como uma única espécie, o *Hepacivirus A*, e a renomeação do HCV (agora *Hepacivirus C* a partir de 2017) (8).

Técnicas de hibridização *in situ* revelaram a presença do ácido ribonucleico viral (EqHV-RNA) no citoplasma de células epiteliais da traqueia e alvéolos de cães, além do fígado, confirmando a presença do EqHV no trato respiratório de cães (3,11). A mesma técnica também identificou EqHV-RNA em células epiteliais dos brônquios e no baço de cavalos. Interessantemente, múltiplos *foci* (do latim, focos) de infecção foram detectados no fígado, comparados a *foci* únicos nos outros órgãos, demarcando

a maior presença da infecção neste órgão. De fato, o intermediário replicativo, a fita de RNA complementar à matriz de leitura do vírus que é traduzida, foi identificado apenas no fígado, levantando fortes indícios do hepatotropismo do EqHV (6,7).

O EqHV é frequentemente identificado em cavalos, ao passo que em cães foram relatados apenas dois casos desde a primeira descrição (11,12), sugerindo que o equino seria o hospedeiro natural deste vírus. Um destes estudos relatou um possível caso de transmissão entre um cão e um cavalo de fazenda que mantinham contato próximo, mostrando que a barreira entre espécies poderia ser quebrada (12). Outra evidência neste sentido foi a identificação de um isolado em soro equino comercial com alta similaridade ao vírus encontrado em cães (10). Análise de relógio molecular identificou um ancestral comum entre o vírus de hospedeiro canino e equino na década de 1970. Um ancestral tão recente reforça a hipótese da possibilidade de transmissão direta entre cães e equinos (13). O EqHV também foi comumente identificado em soro comercial equino (6,7,14), podendo ter sido a causa de transmissão entre estes dois hospedeiros, uma vez que produtos de origem equina são utilizados na produção de vacina da raiva e cinomose canina, patógenos de grande importância veterinária (13).

A diversidade de hospedeiros do EqHV seguiu crescendo quando dois trabalhos também detectaram o EqHV em muares (*Equus africanus asinus*), cujos isolados mesclavam-se junto com sequências de cavalos em análise filogenética, embora apresentassem maior diversidade (15,16). Demonstrou-se, assim que o vírus poderia infectar outra espécie equina sem muitas adaptações necessárias. Uma topografia que não separa o vírus de cavalos e muares e levando-se em consideração que as duas espécies equídeas divergiram a alguns milhões de anos, sugere uma história evolutiva recente para o EqHV. Consequentemente, diminuindo a possibilidade de um evento de infecção equino-humano ao longo da história evolutiva do vírus e diminuindo a probabilidade de que o HCV tenha se derivado de uma infecção humana pelo EqHV, uma vez que a evolução dos equídeos não é paralela à evolução do vírus (15,16).

De fato, o estudo de uma coorte de humanos em intenso contato com equinos, potencialmente configurando grupo de risco para a infecção pelo EqHV, indicou que seria pouco provável a transmissão equino-humano, uma vez que o grupo controle (humanos sem histórico de contatos com equinos) apresentou perfil similar ao grupo teste. Em ambos os grupos, foram encontradas amostras fracamente sororreativas aos antígenos de envelope do EqHV testados, mas nenhum doador apresentou

presença de anticorpos específicos para o HCV, sugerindo que haveria um ancestral comum ainda não identificado entre o HCV e o EqHV (17). A distância genética entre o HCV e o EqHV corrobora a existência de um ancestral comum entre eles e reforça a baixa probabilidade de um ter dado origem ao outro por salto entre espécies (13).

Se por um lado o EqHV já foi descrito em infecção natural de três espécies de hospedeiros diferentes, o HCV tem apenas humanos como hospedeiros naturais e chimpanzés como hospedeiros de infecção experimental (18). Estas similaridades e diferenças entre o EqHV e o HCV pedem uma análise mais minuciosa da biologia e evolução deste vírus recém descrito, trazendo novas perguntas e possibilidades para o combate de uma doença humana proeminente (18).

1.1.1 O genoma

O gênero *Hepacivirus* a que pertence o hepacivírus equino está contido dentro do domínio viral *Riboviria*, caracterizado pela presença de um domínio RNA-polimerase-RNA-dependente (do inglês, RdRp) (19) e na família *Flaviviridae*, cuja organização genômica é similar para todos os membros, com algumas particularidades para o gênero (8). O genoma de RNA fita simples de polaridade positiva dos hepacivírus possui aproximadamente 9.6 kilobases (kb) e pode ser diretamente traduzido no citoplasma celular, funcionando como um RNA mensageiro (RNAm). Produz uma única poliproteína de aproximadamente 3000 aminoácidos (aa), cujas clivagens co- e pós- traducionais geram dez proteínas maduras entre estruturais, que fazem parte da partícula viral, e não estruturais, que possuem função no ciclo replicativo (3,7,20,21) (Figura 1).

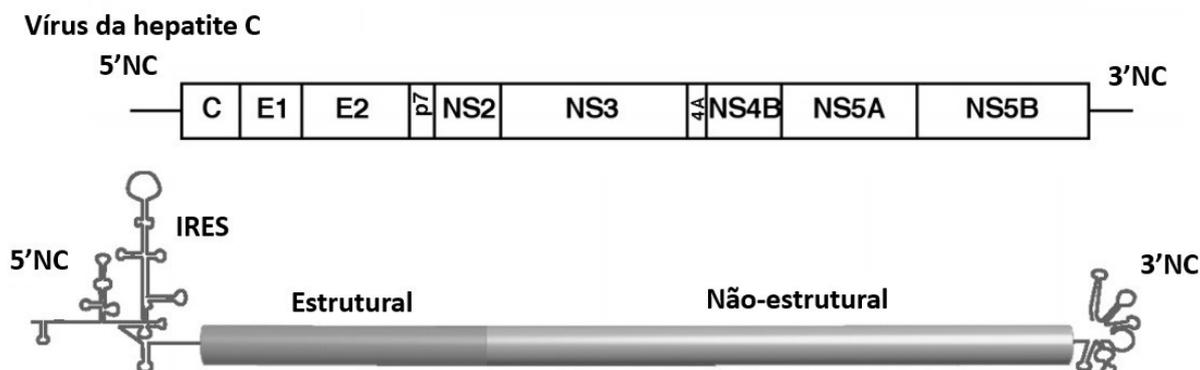


Figura 1 Genoma do HCV

Esquema do genoma do HCV mostrando a matriz de leitura aberta com os 10 genes que compõe a matriz de leitura aberta (ORF) flanqueada por duas regiões não codificantes (NC) nas posições 5' e 3' (acima). Detalhamento das estruturas 5' e 3' e divisão da ORF em região estrutural (C, E1 e E2) e não-estrutural (p7 até NS5B).

A fase de leitura aberta (do inglês, ORF) da poliproteína é flanqueada por duas regiões não-codificantes a 5' (5'NC) e 3' (3'NC), possuindo funções reguladoras importantes como a estrutura interna de entrada do ribossomo (do inglês, IRES) na ponta 5' que inicia a replicação cap-dependente viral. Na interface entre as proteínas estruturais e as não estruturais se encontra uma viroporina, p7, uma proteína de canal iônico importante para replicação e montagem da partícula viral. Esta porção da poliproteína é clivada por proteases celulares (3,20,21).

As proteínas não-estruturais consistem das proteínas NS2, NS3, NS4A, NS4B, NS5A e NS5B. A porção N-terminal desta região é clivada pela holoenzima NS2-NS3, enquanto a porção C-terminal é clivada pela holoenzima NS3-NS4A. A proteína NS4B está associada à formação do viroplasma, um ambiente lipídico propício para a replicação viral, a teia intramembranar, formada a partir da membrana do Retículo Endoplasmático. Os níveis de fosforilação da proteína NS5A já foram associados ao controle entre as etapas de replicação e montagem viral. Por fim, a NS5B é a responsável pela replicação viral em si, devido à atividade RNA polimerase dependente de RNA (RpRd), embora seja preciso a interação com outras regiões do genoma, como a 3'NC, para seu correto funcionamento (22).

Muitas destas informações foram inicialmente geradas em estudos para o HCV, mas a similaridade genética e a proximidade filogenética destes dois vírus permitem predição das funções por homologia (3,23). Algumas funções já foram confirmadas, como a presença de dois sítios micro-RNA122 e domínios de IRES na porção 5', importantes para tradução das proteínas virais (24,25), apesar de haver a separação de um componente comum entre os hepacivírus (a região contendo a IRES) e outro componente espécie específico (26). A proteína de capsídeo do EqHV, assim como para o HCV, também possui peptídeo sinal e auxilia na localização da partícula viral próxima a ilhas lipídicas (24).

A função proteolítica da holoenzima NS2-NS3 também já foi confirmada, inclusive sob mesmo mecanismo de dimerização (27). Quimeras das proteínas do HCV e EqHV mantêm-se funcionais, como é o caso da junção da proteína NS2 e a porção N-terminal da NS3 e da porção C-terminal da proteína p7 (27,28). Nosso grupo de pesquisa revelou por modelagem tridimensional da proteína NS5B do EqHV e *docking* molecular um mecanismo de ação similar à proteína do HCV e um potencial inibitório por anti-virais como o sofosbuvir e dasabuvir em análise *in silico* (29). Foi demonstrado que a holoenzima NS3-4A é capaz de clivar a proteína de sinalização antiviral mitocondrial (do inglês, MAVS) humana, um importante adaptador na cascata

de sinalização do ácido retinóico e ativação da imunidade mediada por interferon, mostrando que o EqHV pode compartilhar um importante mecanismo de escape da imunidade (7,30,31).

1.1.2 A epidemiologia

A epidemiologia do hepacivírus equino ainda não foi profundamente estudada, embora tenha sido detectado ao redor do globo (10,11,15,24,32–37) Estima-se que a soroprevalência global do EqHV gira em torno de 34,9% e a viremia em torno de 6,8% (15,20,38), mostrando sua ampla circulação na população equina.

No Brasil, foi encontrada uma viremia de 13,4% entre os estados do Rio de Janeiro, Espírito Santo e Mato Grosso do Sul (32) e 8,3% no estado do Pará (39). Dentro do estado do Rio de Janeiro, o EqHV foi encontrado em 9 municípios, havendo circulação do EqHV em 62,5% dos haras estudados com uma prevalência de 13,4% (40). Quando considerados os 4 estados, a prevalência do EqHV no Brasil é de 11,91% (83/697) em equinos *Equus ferus caballus*. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), a população de equídeos brasileira é de aproximadamente 5,5 milhões (acessado em <<https://sidra.ibge.gov.br/tabela/3939#resultado>> no dia 11/08/2019), o que corresponderia a mais de 650 mil equinos brasileiros potencialmente infectados pelo EqHV.

Mesmo cavalos nativos do Japão, nunca expostos a viagens internacionais, apresentaram soroprevalência entre 6,1% e 68,2% (41). Por mais de 100 anos, estes cavalos foram mantidos isolados em uma ilha remota ao sul do Japão e ainda assim, o EqHV era presente, levantando a pergunta de qual será a relação dos cavalos com o ambiente que possibilitariam infecção (41).

A sorologia é uma ferramenta epidemiológica importante para detecção de infecções passadas, no entanto, devido à baixa de anticorpos ao longo do tempo, pode haver uma subestimação da epidemiologia do EqHV na população equina (12). Já foi relatado, por exemplo, um caso de um animal com infecção persistente crônica, mas cuja sorologia era negativa. Não se sabe se isto se deve a uma característica natural da infecção crônica ou se foi o efeito de uma menor concentração antigênica (42), reforçando a subestimação da epidemiologia do EqHV.

A maior soroprevalência encontrada para o HCV se deu no Egito, atingindo 14,7% em estudo nacional de 2008 (43–45), atualmente, a distribuição global do HCV é 1,6% quanto a sorologia e 1,1% quanto a presença do RNA viral (46). Portanto, uma

circulação global do EqHV mais ampla do que do HCV poderia indicar vias de transmissão alternativas para o EqHV do que o HCV (12). A via parenteral foi comprovada por infecções experimentais em cavalos (42) e a via vertical foi relatada em infecção natural em um estudo de uma coorte de cavalos na Alemanha, onde o colostro apresentava anticorpos para o EqHV, mostrando transferência passiva de anticorpos, e o cordão umbilical era EqHV-RNA positivo (47). Uma terceira via de infecção por vetores artrópodes hematófagos foi testada, mas não confirmada (13,48). A transmissão entre hospedeiros é provável, uma vez que a análise filogenética da região conservada NS3 apresenta topografia poli ou parafilética para os vírus de diferentes hospedeiros dentro do gênero *Hepacivirus*, mas também pode indicar uma distribuição antropogênica, como já foi sugerido para o hepacivírus presente em gados (13,15,49,50).

Associações entre sexo, idade, raça e transporte foram testados em vários estudos, mas um padrão ainda não pôde ser desenhado (6,20,33,40,41,51,52). Um fator comum entre muitos estudos é a maior prevalência em equinos Puro-Sangue e de corrida, sendo o intenso tratamento veterinário uma das possíveis causas de transmissão (6,20,33,39,40,52).

1.1.3 A clínica

Desde as primeiras descrições do EqHV, a infecção do vírus animal parecia contrastar com a infecção pelo HCV em que este último é capaz de causar danos graves ao fígado, em alguns casos até mesmo em estados precoces da infecção, enquanto o primeiro dificilmente gerava dano hepático e eram relatados poucos casos de hepatite bioquímica ou subclínica (6,12,38,53). No entanto, já foram relatados casos de infecção natural por EqHV causando quadros clínicos severos (7,54,55). Um cavalo com evidente quadro de hepatite icterica e cuja intensidade do sintoma decrescia em paralelo com a carga viral, apesar do RNA viral ainda ser detectável ao longo de quase 5 meses, ilustrou a capacidade do EqHV de gerar dano hepático em fases mais precoces da infecção (54). Um estudo de caso também revelou uma infecção crônica do EqHV em um cavalo com quadro hepático severo, provavelmente multifatorial, resultando na recomendação de eutanásia do animal (55).

A associação entre hepatite e o EqHV também foi demonstrada em infecção experimental em cavalos, revelando também uma cinética de infecção e evolução da infecção similar ao do HCV (7,42). Análises histológicas de tecido hepático de cavalos infectados pelo EqHV revelaram múltiplos *foci* em processo de necrose distribuídas

aleatoriamente e circundados por células monocíticas inflamatórias, linfocitose portal hepática, hiperplasia biliar e evidências de perda e recuperação tissular hepática (6,7,42).

Na infecção pelo EqHV, a soroconversão ocorre junto à elevação das enzimas hepáticas, e à medida em que anticorpos são produzidos, a carga viral decai, evidenciando o papel da imunidade adaptativa no desenvolvimento da doença. Esta conclusão é reiterada pelo fato de que potros, cuja produção de determinadas imunoglobulinas é comprometida pelo sistema imune imaturo, e potros imunossuprimidos, que não produzem anticorpos, apresentam quadros clínicos mais brandos ou nenhum sintoma da doença quando comparado com cavalos adultos (42).

A cinética da resposta imune ao EqHV apresenta produção de anticorpos de forma tardia e queda da taxa de anticorpos ao longo do tempo, características até então similares ao HCV (6,7,12,42,56). Contudo, a taxa de animais com viremia e soroconversão concomitantes é de 22-31,4%, isto é, muitos animais têm sorologia positiva, mas o RNA viral não é detectável (6,10,12,37). Isto representa uma taxa de cronicidade menor da infecção pelo EqHV em cavalos do que a de HCV em humanos (6,10), uma característica clínica que deve ser explorada.

Em um estudo brasileiro conduzido em 2015, foram triados 202 cavalos entre os estados do Mato Grosso do Sul, Espírito Santo e Rio de Janeiro. Destes, apenas 27 eram positivos (13,4%). No ano seguinte, apenas dois animais permaneceram positivos, representando uma taxa de eliminação viral de 89,5% (57). Em uma coorte de 87 cavalos na Alemanha, houve um índice de eliminação viral de 60% em menos de 2 meses (6). Todos os quatro animais adultos inoculados em infecção experimental eliminaram o vírus até a 37^a semana (42), mas potros de infecção experimental e natural evoluíram para infecção crônica (38,42). Um estudo retrospectivo identificou um animal crônico há pelo menos 12 anos, mostrando que a infecção poderia persistir por ao menos uma década (7).

Portanto, a evolução da infecção do EqHV é em muitos aspectos similar ao processo do HCV: (i) é um vírus hepatotrópico; (ii) é capaz de causar infecção persistente, mesmo que em menor taxa; (iii) a carga viral sérica é similar (entre 10⁵ a 10⁶ cópias/mL); (iv) a resposta imune adaptativa é tardia e concomitante com o aumento de enzimas hepáticas; (v) quadro clínico de hepatite e alterações hepáticas similares (7). Este paralelo de características chama atenção à importância da infecção pelo EqHV em cavalos pelo ponto de vista da saúde humana e veterinária.

1.1.4 A diversidade

A diversidade dos hepacivírus não é homogênea ao longo do genoma, o que seria esperado, uma vez que incidem diferentes pressões seletivas em cada trecho (58,59). As regiões mais conservadas são os genes NS3, serino-protease e helicase viral, e NS5B, a replicase viral. Ambas possuem papel essencial no ciclo replicativo viral, logo sofrem forte pressão negativa (22). Ainda assim, não são tão conservadas quanto a porção 5'NC e, portanto, são mais filogeneticamente informativas. Por isso fragmentos destes dois genes são utilizados para classificação dos 8 genótipos do HCV e a diferenciação das 14 espécies do gênero *Hepacivirus* (8,60).

Análise da diversidade do EqHV através da NS5B, revelou a circulação de ao menos 2 subtipos mundialmente, mas não há relação geográfica entre elas (33–35,40,51). Outro estudo sugeriu a separação de 3 subtipos a partir da análise de genoma completo e reforçou a existência de um único genótipo para o EqHV (61). Este nível de diversidade é curioso para um vírus tão proximamente relacionado ao HCV, que possui 8 genótipos e 67 subtipos classificados e uma clara distribuição geográfica (59,60,62). A energia média de dobramento do genoma do EqHV é maior do que para o HCV (3,10,15). Em outras palavras, o estado natural da conformação espacial do genoma viral do EqHV estoca mais energia do que a conformação do HCV, o que significa maiores restrições a mutações devido à preservação de estruturas secundárias específicas do genoma viral. Todavia, tamanha diferença de diversidade entre o vírus humano e de cavalo não poderia ser unicamente explicada pela energia das estruturas genômicas, sendo a evolução dos hepacivírus ainda elusiva (10,15,63).

A diversidade de outros hepacivírus ainda precisa ser elucidada, mas uma diversidade baixa relativa ao HCV, embora maior relativa ao EqHV, foi encontrada para o hepacivírus bovino. Discussões atuais giram em torno da existência de 2 genótipos, suportadas por análise filogenética, mas não de acordo com a classificação taxonômica do HCV (64); ou a existência de um único genótipo e a subcategorização de seis subtipos (65). A maior diversidade dentro do gênero hepacivírus se encontra nos vírus de hospedeiros de roedores e quirópteros, que possuem mais de uma espécie viral (8). A recente expansão do gênero hepacivírus ressalta a peculiaridade da espécie *Hepacivirus C* que alberga uma enorme diversidade em uma única espécie viral, em que a distância entre genótipos seria suficiente para classificar espécies diferentes no gênero *Flavivirus*, por exemplo (13). Esta enorme diversidade torna-se

ainda mais interessante frente ao fato de que a espécie viral mais proximamente relacionada possui uma diversidade tão baixa quanto 2 ou 3 subtipos (20,40,61).

Para o hepacivírus equino, a maior diversidade é encontrada na região estrutural, com maior índice de mutações sinônimas (10). Quando analisados os genomas dos vírus de hospedeiro canino e equino, as regiões do capsídeo e envelope viral concentraram o maior número de mutações (13). Adicionalmente, a análise do genoma completo do EqHV mostrou que grande parte das mutações eram acumuladas na região de envelope, em especial na porção N-terminal do gene E2, revelando um padrão evolutivo similar ao do HCV (15). Apesar de ser uma das regiões mais variáveis do genoma, houve similaridade suficiente que permitisse o alinhamento satisfatório entre o HCV e o EqHV. O número e posicionamento de cisteínas sugerem uma estrutura terciária muito próxima entre eles (3). Estas semelhanças tornam a região de envelope do EqHV um importante alvo para se obter informação evolutiva para o EqHV.

1.2 Quasiespécies

1.2.1 Conceito

A teoria das quasiespécies veio como forma de estudar sistemas finitos de macromoléculas poliméricas, auto-catalíticas e auto-replicativas com alta incidência de erros e cujas variações fossem irreversíveis, de forma a escapar um equilíbrio em um estado de constante dinâmica de adaptação. Os melhores representantes destes sistemas são os códigos genéticos de DNA e RNA. Estes sistemas moleculares têm suas semelhanças às teorias de genética de população, mas suas dimensões lhe conferem particularidades. Em um sistema de genética de populações, variações de baixa frequência não afetam o curso de evolução do sistema e são logo eliminadas, ao passo que a presença do haplótipo de mais baixa frequência em um sistema de macromoléculas pode não só persistir mais consistentemente do que haplótipos de maior frequência, como definir interações e todo curso de evolução da população (66).

Estes sistemas de macromoléculas são visualizados por cadeias de caracteres de um alfabeto finito, onde o tamanho da população cresce geometricamente de acordo com o tamanho da cadeia. Em um alfabeto de 4 letras, como no código genético, uma cadeia (v) possui uma população de 4^v haplótipos. A taxa de mutação é tamanha que é pouco provável que uma molécula idêntica à parental imediata seja

produzida dentro de um mesmo sistema. O tamanho astronômico da população aliada a uma replicação propensa ao erro geram estas particularidades mencionadas no estudo da genética de população de macromoléculas (66,67).

Ainda que o sistema seja dinâmico e possua uma enorme diversidade, há um limite a que possa suportar para que a informação se preserve. Este limite, denominada catástrofe de erros, é o que define muitas características do sistema, como por exemplo a constante reposição de haplótipos de baixa frequência e que, portanto, não são extintos (68). Este fenômeno também tem ligação à “memória” do sistema. Há uma constante flutuação que faz com que populações majoritárias se tornem minoritárias e vice e versa. Por conta da catástrofe de erros, as populações, antes majoritárias, não se extinguem por completo e sua informação é mantida de forma que quando as condições voltem a se tornar favoráveis, esta população consegue se estabelecer mais uma vez como majoritária. Assim, o sistema mantém uma memória, pois não perde informação (66,67,69).

A teoria das quase espécies é determinística, atribuindo-se números e equações para calcular valores e realizar previsões com parâmetros bem definidos. Estas previsões podem ser representadas graficamente na sequência espaço. Entende-se por “*sequência espaço*” um plano abstrato dimensionado pelo número de caracteres diferentes entre sequências vizinhas (distância de Hamming). Considerando, por exemplo, as cadeias de caracteres: (a) “CA”, (b) “CAO” e (c) “CAOS”, as cadeias (a) e (c) estão mais afastadas na sequência espaço, pois possuem 2 caracteres diferentes entre si, logo uma maior distância de Hamming; e a cadeia (b) é equidistante, pois possui apenas 1 caractere diferente em relação às outras duas. Conhecendo a distribuição na sequência espaço de uma população de cadeias, seria possível definir com maior precisão como o sistema irá se comportar e entender como funciona sua dinâmica (66,67,70).

A dinâmica de quasiespécies possui padrões e regularidades singulares que afetam o sistema como um todo. Uma das regularidades mais fundamentais é a adaptação (ou *fitness*), à qual o raciocínio darwinista define limites e adiciona o acaso como fator evolutivo. Para um sistema de macromoléculas, o *fitness* pode ser definido como a capacidade replicativa de determinado haplótipo em determinado ambiente e condições. A sobrevivência também é uma regularidade definida na teoria darwinista, que quando aplicada a sistemas de macromoléculas se refere a um haplótipo que tenha frequência maior que zero. Desta forma sendo possível aplicar o famoso conceito da seleção natural darwinista a populações de macromoléculas (66).

Algumas alterações ainda precisam ser feitas ao raciocínio darwinista para que se aplique mais adequadamente a sistemas de macromoléculas. A alteração mais marcante é a unidade de seleção. Na genética de populações, a unidade de seleção é o indivíduo que carrega o genótipo e expressa o fenótipo que será selecionado. No entanto, quando colocado no contexto de populações de macromoléculas, frente ao tamanho populacional em um contexto de alta taxa de mutações, a unidade de seleção de uma quase espécie não é o haplótipo mais adaptado, mas sim o haplótipo com maior frequência (sequência mestre) junto às sequências de menor frequência e mais relacionadas. Em suma, quasiespécies consistem de um espectro de mutantes não-simétrica, replicando em diferentes taxas, que produzem um viés em massa para produção de novas partículas mutantes (66,67).

1.2.2 Aplicações na Virologia

A virologia veio como uma verificação experimental da teoria determinística, uma vez que vírus de RNA e DNA preenchem todos os requisitos para um sistema de quasiespécies (66,67,70). Vírus guardam informação no código genético de RNA e DNA, logo são moléculas poliméricas, autocatalíticas e autoreplicativas. A replicase viral de uma parte considerável de vírus não possui capacidade de correção durante o processo de replicação, logo as variações ou mutações são irreversíveis (71). A taxa de erro de incorporação de nucleotídeos varia na ordem de grandeza de 10^{-8} a 10^{-3} substituições nucleotídicas por sítio por célula infectada, sendo as maiores taxas para vírus de RNA, construindo um sistema dinâmico de constante adaptação (72–74).

A teoria de quasiespécies aplicada à virologia pode ser abordada de vários pontos de vista em diferentes campos do conhecimento. Para a física, é uma nuvem de informações preenchendo a sequência espaço. Para a química, é uma coleção de sequências nucleotídicas relacionadas, mas não idênticas. Para a biologia é a unidade de seleção da evolução viral. Para a medicina, é um espectro de mutantes que resulta em escape imunológico e resistência a antivirais (66,67). Trata-se, portanto, de um tema multidisciplinar e com amplas ramificações.

A dinâmica de quasiespécies é intrínseca à biologia dos vírus e vários conceitos da teoria são facilmente exemplificados por estas entidades biológicas. O conceito de “memória molecular”, por exemplo, pode ser ilustrado pela infecção de arboviroses em hospedeiros vertebrados e invertebrados, na qual ocorre adaptação aos dois hospedeiros de forma específica e generalizada (75). Isto é, ocorrem adaptações específicas para infecção em um hospedeiro invertebrado, mas são guardadas

informações no espectro de mutantes para que também possa infectar um hospedeiro vertebrado. Arboviroses exploram a sequência espaço de tal forma que mesmo havendo gargalos genéticos em eventos de troca de hospedeiros, o *fitness* viral entre populações que passaram pelo gargalo ou que se mantiveram em um mesmo hospedeiro não sofre grandes variações, ilustrando sua rápida adaptabilidade aos dois ambientes devido à memória molecular da população (75,76). Eventos frequentes de gargalo genético devido à troca de hospedeiros constante impacta negativamente a diversidade genética de arbovírus, considerada a menor dentre vírus de RNA. Isto parece ser consequência de sua dinâmica populacional e não de características biológicas dos vírus (75,77).

Outro conceito da teoria das quasiespécies que pode ser facilmente exemplificado na virologia é o da catástrofe de erros. O poliovírus parece ter evoluído de forma a se aproximar ao máximo ao limite da catástrofe de erros, uma vez que pequenos aumentos à taxa de mutação são altamente prejudiciais ao vírus, podendo chegar a uma perda de 95% da capacidade de infecção (78). Por outro lado, afastar o vírus do limite do erro gera populações com menor potencial patogênico, como mostram experimentos de um isolado viral com uma replicase de relativa maior fidelidade (79). Esta perspectiva também é explorada como estratégia anti-viral, como no caso do fármaco ribavirina. A ribavirina funciona como um análogo de nucleosídeo e agente mutagênico, lançando a taxa de mutação muito além do limite suportado pela catástrofe de erros (78,80).

A ampla capacidade de explorar a sequência espaço é responsável pelas maiores barreiras na prevenção e tratamento de doenças infecciosas de etiologia viral. A complexa dinâmica de populações virais torna mais fácil o surgimento de modificações nos sítios de ação dos fármacos ou nos epítomos de ação de linfócitos T-citotóxicos e anticorpos, levando a resistência ao tratamento com antivirais, dificuldade de desenvolver vacinas como método profilático e a eliminação do vírus pelo sistema imunológico (5,81,82). O surgimento de variantes de escape imunológico aumenta em função das taxas de mutação e eliminação viral. A eliminação viral é uma variável determinada pela eficiência de reconhecimento de epítomos e eliminação da célula marcada. Desta forma, o escape imunológico atinge seu máximo quando a carga viral é maior e as mutações são mais frequentes (81).

O estudo de quasiespécies virais, portanto, pode ser de grande auxílio para elucidar a evolução da infecção clinicamente, seu tratamento e a interação entre diferentes hospedeiros (83,84).

1.2.3 Como acessar a diversidade de quasiespécies

Para melhor acessar e analisar a diversidade do espectro de mutantes, as melhores técnicas envolvem direto acesso às sequências que formam a população, sendo o padrão ouro a metodologia da clonagem molecular (85,86). Para se obter acesso às sequências da população, no entanto, deve-se invariavelmente preceder uma etapa de transcrição reversa, para os vírus de RNA, e amplificação por polimerização em cadeia (PCR). Nesta etapa podem incidir erros devido à fidelidade da enzima empregada e que, portanto, podem mascarar a diversidade provinda da biologia do vírus. Assim, deve ser uma preocupação a escolha da enzima utilizada para a PCR (85).

Outra preocupação importante a se ter ao realizar qualquer metodologia que envolva uma etapa de amplificação anterior é o erro de reamostragem. Se não levado em consideração a diversidade resultante poderá estar grosseiramente subestimada (87). Considere uma população com 4 haplótipos ou sequências diferentes. Cada haplótipo está presente em diferentes frequências de valor “ x ”, “ $2x$ ”, “ $4x$ ” e “ $10x$ ”. Os eventos de amplificação das populações de maior frequência são mais numerosos, e ao final, será encontrada uma representatividade de apenas 2 haplótipos em face dos 4 que compõe a população. Desta forma, é importante conhecer a concentração de moléculas molde aplicadas na reação de amplificação, evitando-se amostras de concentração desconhecida.

Uma metodologia alternativa à clonagem que busca sanar o erro da reamostragem é a diluição limitante da PCR (86,87). Aplicando-se esta metodologia, a individualização dos haplótipos ocorre anteriormente à amplificação por um processo de diluição, eliminando, ou diminuindo substancialmente, o erro por reamostragem. A metodologia se baseia na probabilidade estatística de Poisson. Uma amostra é diluída serialmente e cada ponto da diluição é submetido à amplificação em replicatas. Sabe-se, através das equações de probabilidade de Poisson, que a diluição capaz de amplificar apenas 30% das replicatas, possui uma probabilidade alta de ter sido amplificada a partir de uma única molécula (87). Uma vez descoberta a diluição limitante para aquela amostra, é repetida sua amplificação com um número robusto de replicatas para se obter as quasiespécies.

Uma terceira alternativa mais moderna é a utilização das novas tecnologias de sequenciamento de NGS. O raciocínio destes sequenciamentos modernos envolve a individualização das sequências sendo processadas, de forma que as quasiespécies podem ser prontamente acessadas. Uma das maiores vantagens é a enorme

cobertura do fragmento analisado e quantidade de dados produzida, podendo ser acessadas frequências baixíssimas da população. Assim como as outras duas técnicas, esta também apresenta seus percalços. O custo deste procedimento ainda é alto e não é facilmente implementado em um laboratório. Esta técnica também adiciona um passo adicional de curadoria dos dados brutos o qual é necessário certo nível de conhecimento de bioinformática, podendo ser impeditivo (88–90).

Todas as três técnicas foram analisadas comparativamente e apresentam eficiências similares quanto à análise da diversidade viral (91,92). Quando a clonagem foi comparada à diluição limitante, foi demonstrado que os clones obtidos através das duas técnicas tinham origem na mesma população de quase espécie e não houve viés em qualquer uma das técnicas, embora um número adequado de clones deva ser analisado, na ordem entre 20 e 30 clones (85,92). A similaridade entre as populações obtidas por NGS e clonagem foi de 71%, não havendo diferenças na variabilidade entre elas, como calculado pela entropia de Shannon (91). No entanto, o processo de curadoria dos dados brutos provindos da NGS elimina sequências contendo deleções ou inserções, um viés não visto para a clonagem molecular (91). As populações de quasiespécies virais são constituídas por partículas defectivas e estas têm importante papel na dinâmica populacional, logo a técnica da clonagem possui essa vantagem sob a NGS (93).

1.2.4 A Dinâmica de Quasiespécies dos *hepacivírus*

O HCV possui complexa dinâmica de quasiespécies (94,95). O alvo mais comum para acessar a diversidade de quasiespécies do HCV é o gene de envelope E2. A porção N-terminal deste gene possui uma região hipervariável (HVR1) onde se localiza um epítipo para anticorpos neutralizantes, logo constituindo um sítio de forte seleção negativa (96–98). A alta diversidade da HVR1 está diretamente ligada à resposta imune, uma vez que pacientes hipogammaglobulínicos apresentam menor diversidade de HVR1 (95,98,99).

A literatura tenta conectar a severidade da hepatite causada pelo HCV, assim como sua habilidade de cronificação, à dinâmica e heterogeneidade das quasiespécies (98,100–103). No entanto, ainda existem discordâncias na literatura quanto ao fator preditivo clínico das quasiespécies. A literatura descreve flutuações de parâmetros dentro de um mesmo paciente ao longo do tempo e entre pacientes nos parâmetros, como a distância genética, razão de mutações sinônimas e não sinônimas e transversões e transições que podem explicar estas discordâncias (98).

Os dois perfis, de alta ou baixa diversidade da região HVR1, tem consequências clínicas que se sobrepõe, sendo compartilhadas entre os dois. Um perfil de maior diversidade da HVR1 já foi associado tanto a menor quanto maior severidade de doença hepática, ligada ao hepatocarcinoma celular, assim como à rápida resposta viral e à persistência da viremia (98,103–108). Igualmente, um perfil de menor diversidade da HVR1 foi associada ao progresso da doença, também ligada ao hepatocarcinoma celular ou hepatite fulminante e à eliminação viral (101,102,109,110). Este perfil mais homogêneo ocorre em pacientes recém transplantados (111). Quando analisadas como uma rede para melhor representação da dinâmica da população ao longo do tempo, percebem-se dois perfis de população para as quasiespécies do HCV. Uma população de deslize antigênico em que as populações ao longo do tempo divergem da população original e formam agrupamentos específicos para cada ponto de coleta analisado; e uma população estática de maior complexidade em que a população original é mantida em conjunto com populações de menor frequência (95,98). Pacientes de patogênese branda foram associados ao perfil de deslize antigênico, enquanto pacientes com evolução para quadro severo apresentam perfil estático (98). Durante a infecção crônica, a eficiência de neutralização dos anticorpos para a região de envelope decairia, levando a homogeneização da população (95). Na sequência espaço, a população de deslize antigênico ocuparia um pico acentuado com uma base menor (população 1 indicada na Figura 2), em que pequenas alterações no meio geram grandes consequências no *fitness* viral, enquanto a população estática ocuparia uma topografia mais achatada e de base maior (população 2 indicada na Figura 2), representando uma população mais robusta e resistente a variações. Teoria denominada sobrevivência do mais achatado (Figura 2) (*survival of the flattest*) (70).

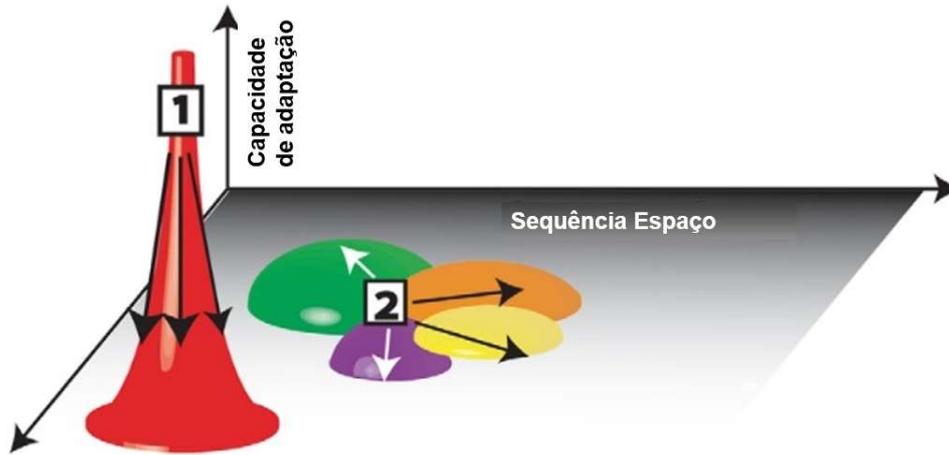


Figura 2. Sobrevivência do mais achatado.

Esquema da sequência espaço representando dois perfis de população. A população (1) em vermelho é altamente adaptada, mas qualquer mutação leva à perda de adaptabilidade ou *fitness*. A população (2) é composta por diversas sub-populações em lilás, verde, amarelo e laranja, com baixa adaptabilidade, mas grande flexibilidade na sequência espaço. Esta população é mais resistente às mutações. Adaptado de LAURING e ANDINO, 2010.

Evidências apontam para a possibilidade da circulação do EqHV com uma dinâmica de quasiespécies, tendo em vista um padrão evolutivo similar, o acúmulo de mutações na região equivalente à HVR1 e uma cinética de infecção parecida com o HCV (6,15,42,112). De fato, foi recentemente descrita uma distribuição de quasiespécies em duas populações de equinos independentemente infectadas, assim como a presença de ao menos 3 regiões hipervariáveis no gene E2 (112).

As duas populações analisadas apresentavam padrões de dinâmica diferentes entre si, discordando na relação entre a diversidade populacional e a persistência viral, assim como da progressão da patologia (112). Ao caracterizar a primeira população, havia uma clara associação entre o aumento da diversidade e a persistência da infecção, uma vez que animais com menor diversidade relativa eliminaram a infecção dentro de 5 semanas pós-inoculação. Da mesma forma, os animais com maior diversidade viral apresentaram uma patologia mais branda em relação aos de menor diversidade, reforçando a associação entre diversidade e patologia (42,112). A caracterização da segunda população, no entanto, contestou estas associações. A diversidade desta segunda população era significativamente menor à da primeira e se manteve estática, não sendo observadas variações ao longo do tempo. Mesmo assim, estes animais mantiveram infecção persistente e seu quadro clínico era branda, indo diretamente de encontro com as observações da primeira população (42,112). Desta forma, os autores concluíram que a dinâmica de quasiespécies não seria preciso

como fator preditivo da infecção pelo EqHV e que existem outras variáveis que influenciam na persistência da infecção (112).

O acúmulo de alterações de aminoácidos nas regiões de hipervariabilidade foi concomitante com o momento de soroconversão, indicando que estas regiões, assim como para o HCV, são regiões de epítipo para anticorpos neutralizantes (112). Contudo, a diversidade nas regiões hipervariáveis aumentava antes do processo de eliminação da infecção e não houve relação com a diminuição da carga viral. Logo, o processo para eliminação viral do EqHV possivelmente se dá através de uma rápida resposta imune generalizada, envolvendo epítipos de outras proteínas virais além da resposta para as regiões hipervariáveis (112).

Considerando as discrepâncias relatadas para o HCV quanto à associação da diversidade viral à clínica e persistência da infecção (98), a proximidade filogenética (20), uma patologia similar, porém com maior tendência à eliminação viral (6), um padrão evolutivo similar (15) e interações vírus/hospedeiro nas regiões de hipervariabilidade (112), o EqHV está inserido em um contexto ideal para a responder relevantes questões referentes ao HCV e ao gênero *Hepacivirus*, através da elucidação de suas características biológicas e evolutivas.

1.3 Justificativa

O animal infectado pelo EqHV desenvolve quadros clínicos de hepatite branda e severa. No curso de infecção natural, casos de hepatite branda são mais comumente descritos. O EqHV é capaz de estabelecer a infecção de forma crônica, no entanto, constitui uma minoria dos casos. O genoma do EqHV possui regiões hipervariáveis na região do envelope viral e apresenta uma variação ao longo do tempo que remete a uma dinâmica populacional de quasiespécies virais. Essa complexa dinâmica viral encontrada, por exemplo, no HCV, é associada a complicações clínicas, no tratamento e no desenvolvimento de vacinas. Entender esta complexidade populacional em contextos diferenciados da evolução da infecção, isto é, de animais que eliminaram ou que evoluíram para a infecção crônica, em especial em hospedeiros que têm maior tendência a eliminação, pode elucidar características importantes da evolução dos hepacivírus. Tendo em vista a proximidade filogenética e padrão evolutivo similar entre o EqHV e o HCV, um importante problema de saúde pública mundial, os benefícios deste tipo de estudo podem abranger a área médica veterinária e humana, contribuindo também na agenda de eliminação das hepatites virais humanas da Organização Mundial da Saúde (OMS).

2 OBJETIVOS

2.1 Objetivo Geral

Determinar a diversidade de quasiespécies do EqHV a partir da região de envelope E1/E2 em equinos (*Equus ferus caballus*) que eliminaram a infecção ou que evoluíram para a infecção crônica; e entre equinos de diferentes localizações geográficas

2.2 Objetivos Específicos

- Estimar a diversidade genética das populações e quasiespécies virais obtidas;
- Comparar a diversidade das populações de quasiespécies virais intra- e inter-populacional em equinos que eliminaram a infecção e o equino que evoluiu para infecção crônica.
- Comparar a diversidade genética das populações de quasiespécies virais inter-hospedeiro em equinos de diferentes localizações geográficas;

3 MATERIAL E MÉTODOS

3.1 População de estudo

Este trabalho foi composto por um total de nove equinos (*Equus ferus caballus*). Cinco deles pertencentes ao acervo de animais do Departamento de Veterinária da Universidade Federal Rural do Rio de Janeiro (UFRRJ), mesorregião Metropolitana, acompanhados de 10/2013 a 02/2017 (32,57) quanto à presença do EqHV circulante no sangue (licença da Comissão de Ética no Uso de Animais – CEUA/UFRRJ números 375/2013 e nº 1298251016). Além de dois equinos de coleta única das mesorregiões Central e Noroeste do estado do RJ (CEUA-IOC licença 047/2015), assim como um equino do estado do ES e um equino do estado de MS (CEUA/FIOCRUZ LW-1/12) (40).

O critério de inclusão foi amostras de soro cuja carga viral do EqHV superasse o limite de $> 10^4$ cópias/mL. Foram excluídos deste estudo os animais que apresentaram coinfeção com pegivírus equino (Família *Flaviviridae*, gênero *Pegivirus*).

Os animais foram agrupados conforme detalhado a seguir:

I- População de equinos de um mesmo haras do RJ, pertencentes à UFRRJ.

(a) Um equino com infecção crônica (animal U114) que se manteve EqHV-RNA-positivo desde 2014;

(b) Quatro equinos que apresentaram resolução espontânea da infecção entre outubro de 2013 e março de 2015 e mantiveram-se EqHV-RNA-negativos até o presente (animais: U64, U74, U78 e U82).

II – População de equinos de diferentes haras sem informação sobre o curso da infecção:

(a) Dois equinos de outros municípios do RJ: animais B25 (Central) e M42 (Noroeste);

(b) Dois equinos de diferentes estados brasileiros: animais 348 (Mato Grosso do Sul – MS) e 674 (Espírito Santo – ES);

A análise da diversidade genética viral intra-hospedeiro foi realizada comparando-se as populações de quasiespécies dentro de cada animal. A análise inter-hospedeiro foi realizada comparando-se as populações de quasiespécies entre cada animal pertencentes a um mesmo haras (UFRRJ) e entre localidades geográficas diferentes. Todos os nove animais tiveram as cargas virais previamente

determinadas por qPCR, como mostram a Tabela 1 e Tabela 2, conforme protocolo estabelecido por Figueiredo e colaboradores (2015) (32).

Tabela 1: População de estudo

Identificação Interna	Sexo	Idade (meses)	AST (226-366IU/L)	Estado/Região	Carga Viral ^a (cópias/mL)
U64	F	12	373	RJ/Metropolitana	1,65x10 ⁴
U74	F	86	355		4,67x10 ⁵
U78	F	12	328		1,34x10 ⁵
U82	F	10	325		3,46x10 ⁶
U114	M	55	491		Abaixo do limite ^b
B25	M	204	324	RJ/Central	4,2x10 ⁵
M42	M	8	405	RJ/Noroeste	8,13x10 ⁵
348	F	NA	212	MS	2,11x10 ⁵
674	M	48	127	ES	1,30x10 ⁵

Sexo: F – Fêmea; M – Macho. AST: Aspartato aminotransferase, valores referenciais entre parênteses. Valores que ultrapassaram a referência são destacados em vermelho. N/A – Informação não disponível. a – Cargas virais referentes à coleta 0 - out/2013 (U64, U74, U78, U82) e fev/2014 (U114). b - Limite de quantificação: 3,86x10³. Quantificação realizada por Figueiredo e colaboradores (2015 e 2018).

Tabela 2: Carga viral (cópias/mL) do animal crônico

	U114	
	AST (226-366IU/L)	Carga Viral
Coleta 0 (2014)	491	Abaixo do limite ^a
Coleta 1 (2015)	298	1,71x10 ⁵
Coleta 2 (2016)	342	1,86x10 ⁵
Coleta 3 (2017)	N/A	6,16x10 ⁴

AST: Aspartato aminotransferase, valores referenciais entre parênteses. Valores que ultrapassaram a referência são destacados em vermelho. N/A – Informação não disponível. a - Limite de quantificação: 3,86x10³.

3.2 Definição do alvo molecular e construção de oligonucleotídeos iniciadores

A construção de oligonucleotídeos iniciadores e avaliação destes foi realizada através da ferramenta virtual IDT *OligoAnalyzer* versão 3.1 disponível no endereço eletrônico: <https://www.idtdna.com/calc/analyzer>, a partir de um alinhamento de 49 sequências de genoma completo ou parcial do hepacivírus A coletadas em bancos de dados genéticos (Tabela 3). O alinhamento foi realizado utilizando o programa Clustal W (113) contido no software *Molecular Evolutionary Genetics Analysis* (MEGA) v.7.0 (114).

Tabela 3: Sequências do EqHV coletadas do GenBank

Número de Acesso	Referência
AB863589	(24)
JF74491	(3)
JQ434001-JQ434008	(10)
JX948116	(115)
KF177391	(54)
KJ472766	(42)
KP325401-KP325402	(7)
KP739814	(6)
KT880191-KT880193	(15)
KX056116-KX056117	(33)
KX064762-KX064784	(38)
KX421286- KX421287	(15)
KY124246- KY124248	(116)

Sequências disponíveis até Setembro de 2017

Os melhores sítios para hibridização de oligonucleotídeos iniciadores foram encontrados de forma que houvesse maior representação do espaço amostral disponível sem que houvesse perda de sensibilidade ao adicionar degenerações. A Tabela 4 ilustra as sequências encontradas para as reações de PCR e *Semi-Nested* PCR e suas localizações de acordo com a sequência referência de número de acesso KP325402.

Tabela 4: Oligonucleotídeos iniciadores construídos para região de envelope viral do EqHV

Identificação	Sequência (5'→3')	Pareamento	Extensão	Localização (KP325402)
EnvOS	CGCCATCTCTACCTTTGAGTTTA	EnvOS + EnvOR	499pb	1341
EnvOR	CACATCCACCCTGTCCTTATC			1819
EnvIR	GTCTGTGGCAAACCGACCTGT	EnvOS + EnvIR	443pb	1763

O produto amplificado foi definido de forma que abrangesse a porção C-terminal da proteína codificada pelo gene E1 e a porção N-terminal da proteína codificada pelo gene E2, onde se localiza a região hipervariável 1 (HVR1) (Figura 3) (112). Esta região foi escolhida devido ao potencial de estudo da diversidade do EqHV segundo trabalhos anteriores (15,112). O início do gene E2 foi determinado de acordo com Kapoor e colaboradores (2011), começando com o motivo YSGLVS, a partir do qual estabeleceu-se uma margem a montante e a jusante da sequência, considerando-se que a HVR-1 encontra-se nos 28 primeiros aminoácidos da proteína codificada pelo gene E2 (3,96).

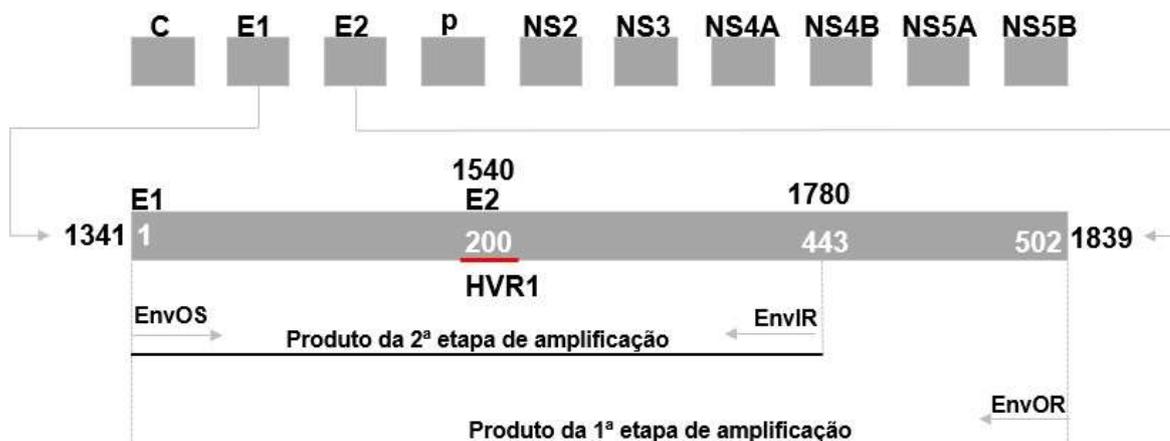


Figura 3: Esquema da amplificação por *semi-Nested* PCR da região do envelope viral do EqHV

Ilustração da região de interesse dentro do genoma do hepacivírus A. Dos 10 genes que formam o genoma do vírus, a porção 5' do gene E1 e 3' do gene E2 é delimitada. Números em preto do lado de fora da caixa cinza representam a localização na sequência referência KP325402. Os números em branco dentro da caixa cinza representam a numeração no produto amplificado. A linha vermelha demarca a região HVR1. Representação dos oligonucleotídeos EnvOS, EnvOR e EnvIR demarcam a extensão dos produtos amplificados nas duas etapas.

3.3 Extração de RNA viral e síntese de cDNA

O RNA total das amostras de soro equino foi extraído com o *QIAmp Viral RNA mini kit (Qiagen)* a partir de 140µL de soro e armazenado a -80°C. Um volume de 6µL de RNA total foi convertido a cDNA por transcrição reversa utilizando oligohexâmeros randômicos e transcriptase reversa *SuperScript™ III Reverse Transcriptase* em um volume final de 20µL, seguindo-se o protocolo recomendado pelo fabricante.

3.4 Amplificação do alvo gênico E2 do hepacivírus A

Seguiu-se a amplificação das amostras em duas etapas semi-aninhadas (*semi-nested PCR*) com os oligonucleotídeos descritos no item 3.2. As duas etapas da reação foram realizadas com a enzima *Platinum™ Taq DNA Polymerase High Fidelity (Thermo Fischer)* adicionando-se o tampão fornecido pela enzima a uma concentração final de 1x; desoxiribunucleotídeos trifosfatados (dNTP) a uma concentração final de 0,2mM; sulfato de magnésio (MgSO₄) a uma concentração final de 1,5mM; oligonucleotídeos iniciadores a uma concentração final de 0,3µM; 1 unidade de enzima por reação; água deionizada para completar 20µL; e 5µL de cDNA. As reações ocorreram com a ciclagem descrita a seguir: Desnaturação inicial a 94°C por 2 minutos; 35 repetições das temperaturas de desnaturação a 94°C por 30 segundos, hibridização dos oligonucleotídeos a 62°C por 30 segundos e extensão de amplificação a 68°C por 35 segundos; finalizando com uma extensão final a 68°C por 2 minutos. A segunda etapa de amplificação seguiu o mesmo protocolo, exceto com adição de água deionizada para 24µL, adição de 1µL da reação da primeira etapa e um tempo de extensão de amplificação de 30 segundos durante as repetições da ciclagem. O par de oligonucleotídeos iniciadores utilizados na primeira etapa foram EnvOS e EnvOR, e na segunda etapa, EnvOS e EnvIR (Tabela 4).

O produto amplificado foi visualizado através da eletroforese em gel de agarose 1,5% corado com brometo de etídeo.

3.5 Clonagem molecular do produto amplificado

3.5.1 Preparo de células competentes

Para individualização das sequências únicas e obtenção das populações de quasiespécies, optou-se pela técnica de clonagem molecular TOPO™ TA Cloning™, utilizando-se vetor pCR4™-TOPO™ (*Thermo Fischer*) e células competentes produzidas *in house* a partir de células estoque de *One Shot™ TOP10 Chemically Competent E. coli* (*Thermo Fischer*). Em uma placa de Petri contendo meio Lúria Berthani (LB) Ágar sólido, foram semeadas as bactérias estoque e incubadas por 16h à 37°C. Uma colônia foi coletada, inoculada em 5mL de meio LB líquido e incubada a 37°C sob agitação de 100 rotações por minuto (rpm) durante 20h. Um volume de 1,5mL desta solução de bactérias saturadas foi inoculado em 150mL de meio LB líquido e incubado a 37°C sob agitação de 150rpm. A densidade óptica (OD) a 600nm foi medida até atingir um valor entre 0,3 e 0,4 OD. A solução foi imediatamente levada ao gelo e incubada por 30 minutos, seguida de separação em tubos para centrifuga. Todas as etapas seguintes foram realizadas no gelo e com o material e equipamentos previamente resfriados. A solução foi centrifugada a 8000rpm por 15 minutos a 4°C, descartando-se o sobrenadante e ressuspensando o *pellet* em 25mL de CaCl₂ (0,1M) gelado. A solução foi novamente centrifugada nas mesmas condições que anteriormente, o sobrenadante descartado e o *pellet* ressuspensado desta vez com 2mL de solução de CaCl₂ (0,1M) e glicerol (25%). As alíquotas resultantes foram congeladas e guardadas à -80°C até o momento de uso.

3.5.2 Adenilação

Para garantir melhor eficiência da clonagem TA, e tendo-se em vista o uso de uma enzima de alta fidelidade na amplificação do produto, foi realizada uma etapa adicional de adição de uma cauda de adenina à ponta 3' dos produtos (adenilação). Após duas etapas de amplificação, o produto foi purificado utilizando o *QIAquick® PCR purification Kit* (*Qiagen*) e a 4μL foi adicionada adenina trifosfatada (dATP) a uma concentração final de 0,2mM, tampão contendo MgCl₂ a uma concentração final de 1.5mM, 1U de enzima GoTaq® DNA Polimerase (*Promega*) e água deionizada para um volume final de 10μL. Seguiu-se para incubação a 72°C por 20 minutos.

3.5.3 Ligaçao

Para próxima etapa da clonagem, a ligação do inserto ao vetor, seguiu-se as instruções do fabricante, adicionando-se 4µL do produto recém adenilado, 1µL de solução salina e 1µL de vetor pCR4 (ThermoFischer), incubando a solução por 20 minutos a temperatura ambiente.

3.5.4 Transformação

Todo o volume da ligação (6µL) foi inoculado nas células competentes já descongeladas. A transformação das células ocorreu através de um período de incubação de 20 minutos em banho de gelo, seguido de choque térmico a 42°C por 70 segundos e retorno imediato das células ao banho de gelo. Em seguida, adicionou-se 300µL de meio LB líquido às células transformadas, que foram deixadas sob agitação (220rpm) a 37°C por uma hora. Todo volume de bactérias obtido foi depositado em placas de Petri contendo meio de cultura LB ágar sólido e ampicilina (100µg/mL) e incubado em estufa a 37°C até crescimento adequado de colônias individualizadas.

3.5.5 Teste das colônias

Um total de 30 colônias individualizadas foram paralelamente transferidas (com a mesma alça microbiológica) para uma placa de Petri “espelho” contendo meio LB ágar sólido e ampicilina (100µg/mL) e submetidas a uma reação de amplificação com os alvos universais de plasmídeos M13 para verificação da presença ou ausência do inserto. Os plasmídeos que fossem positivos eram coletados da placa “espelho” para seguir o fluxograma dos experimentos. Esta amplificação foi realizada com a enzima GoTaq® *DNA Polymerase* (Promega) adicionando-se o tampão fornecido pela enzima a uma concentração final de 1x; desoxiribonucleotídeos trifosfatados (dNTP) a uma concentração final de 0,2mM; oligonucleotídeos iniciadores M13 senso e anti-senso a uma concentração final de 0,3µM; 1 unidade de enzima por reação; e água deionizada para completar 50µL. A reação ocorreu com a ciclagem descrita a seguir: Desnaturação inicial a 95°C por 10 minutos; 30 repetições das temperaturas de desnaturação a 95°C por 30 segundos, hibridização dos oligonucleotídeos a 55°C por 40 segundos e extensão de amplificação a 72°C por 1 minuto; finalizando com uma extensão final a 72°C por 7 minutos.

O produto amplificado foi visualizado através da eletroforese em gel de agarose 1,5% corado com brometo de etídeo.

3.5.6 Extração de material plasmidial

As colônias que continham o inserto foram incubadas em 4mL de meio LB líquido contendo 100µg/mL de ampicilina sob agitação (200rpm) a 37°C entre 14h a 16h. Destes 4mL de bactérias saturadas, 1mL foi separado em 20% de glicerol e o restante foi submetido à extração de material plasmidial utilizando-se o *QIAprep Spin Miniprep Kit* (QIAGEN). O material plasmidial obtido foi quantificado utilizando-se o espectrofotômetro NanoDrop™ 2000 (*Thermo Fischer*).

3.6 Sequenciamento

Todos os clones obtidos foram submetidos ao sequenciamento de Sanger, utilizando-se o *Big Dye Terminator v3.1 Cycle Sequencing kit* (*Life Technologies*) no sequenciador automático ABI3730xl (*Applied Biosystems*) da Plataforma de sequenciamento de DNA (RPT01A-VPPCB/FIOCRUZ) da Fundação Oswaldo Cruz.

3.7 Análise das sequências nucleotídicas

As sequências obtidas foram analisadas no programa MEGA v.7.0 (114). Para verificar a identidade das sequências, foi utilizada a ferramenta virtual *Basic Local Alignment Search Tool* (BLAST) da NCBI disponível no endereço eletrônico: http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.

3.7.1 Cálculo da distância genética

A distância genética (p-distância) e distância de Hamming dentro de cada população e entre populações foram calculadas utilizando-se os programas *SSE Alignment v.1.3* (117) e *MEGA v.7.0* (114), respectivamente. Para efeito de ilustração e localização das variações populacionais, gráficos foram construídos com auxílio do programa *SSE Alignment v.1.3* (117), Excel Microsoft.

3.7.2 Análises estatísticas

As análises estatísticas foram realizadas no *software R 3.5* incluindo os seguintes procedimentos: teste de Shapiro-Wilk para avaliação da normalidade de distribuição dos dados; teste de Kruskal-Wallis para comparar populações

independentes quando estas não possuem distribuição normal e correlação de Spearman. Todos os testes aplicados foram realizados com nível de significância de 5%.

3.7.3 *Análise filogenética*

Cálculo do melhor modelo evolutivo e construção da árvore filogenética foram realizados através do programa JModelTest (118) e MEGA v.7.0 (114) utilizando o algoritmo de Máxima Verossimilhança (1000 repetições) ou Inferência Bayesiana através do programa BEAST v.1.7 (119). A árvore filogenética foi visualizada com o programa FigTree v.1.4.3.

4 RESULTADOS

4.1 Otimização e padronização do alvo molecular construído neste trabalho

Os oligonucleotídeos iniciadores para o alvo gênico E2 foram construídos, testados, padronizados e o fragmento resultante confirmado por sequenciamento. A Figura 4 mostra o teste da temperatura de hibridização em gradiente a 56°C, 58°C e 60°C utilizando duas amostras de soro EqHV-RNA positivas e uma EqHV-RNA negativa para comprovação de especificidade da reação.



Figura 4. Otimização da temperatura de hibridização em matriz de agarose

Enumeração dos poços acima representam: PM100pb – Marcador de peso molecular de 100 pares de base; 1 a 3: PCR da amostra de soro EqHV positiva 1; 4 a 12: *Semi-nested* PCR da amostra de soro EqHV positiva 1; 13 a 15: PCR da amostra de soro EqHV positiva 2; 16 a 24: *Semi-nested* PCR da amostra de soro EqHV positiva 2; 25 a 27: PCR da amostra de soro EqHV negativa; 28 a 30: *Semi-nested* PCR da amostra de soro EqHV negativa; 31- água da PCR; 32- água do *Semi-nested* PCR; 33- branco da reação sem molde para amplificação. As temperaturas testadas foram 56°C, 58°C e 60°C, respectivamente, para PCR. Cada temperatura da PCR foi testada para as temperaturas de 56°C, 58°C e 60°C, respectivamente, na *Semi-nested*-PCR. De forma que, ao final, obtivessem nove amplificações visíveis para cada amostra de soro testada.

Os alvos foram específicos, uma vez que não houve amplificação de uma amostra negativa. Foi comprovada também a importância de realizar duas etapas de amplificação, uma vez que uma única etapa não foi suficiente para identificar as amostras positivas. As amostras apresentaram perfil de amplificação esperado segundo Tabela 4 na altura de aproximadamente 450pb. O sequenciamento destes produtos confirmou se tratar de um fragmento de envelope do hepacivírus equino. A fim de garantir a maior especificidade, a maior temperatura de hibridização foi escolhida.

4.2 Clonagem Molecular

4.2.1 Cálculo do erro intrínseco à técnica aplicada

4.2.1.1 Erro de reamostragem

A escolha pela técnica de clonagem molecular implica uma restrição no espaço amostral obtido se uma amostra de soro tiver baixa carga viral e a ocorrência do erro de reamostragem, quando uma cópia é amplificada diversas vezes, mascarando a verdadeira diversidade viral. A quantidade de cópias virais aplicadas na PCR variou de 57 (U64) a 12.110 (U82) cópias por reação. A amostra de soro U114 coleta 0 apresentava carga viral abaixo do limite de detecção, logo apenas a sequência majoritária foi analisada.

Exemplificando o cálculo com a amostra de soro do animal U64: Em 140 μ L de soro contendo uma carga de $1,65 \times 10^4$ cópias/mL são carregadas 2310 cópias virais ($16500 \times 0,14$). Estas foram diluídas em 60 μ L da eluição de RNA, atingindo concentração de 38,5 cópias/ μ L. Desta solução, 6 μ L (231 cópias), foram transferidas para solução de cDNA a um volume final de 20 μ L, com uma nova concentração de 11,55 cópias/ μ L. Para a reação de PCR, foram utilizados 5 μ L de cDNA, ou seja aproximadamente 57 cópias ($11,55 \times 5$) foram aplicadas na primeira etapa de amplificação.

4.2.1.2 Erros de amplificação

Para calcular o erro da técnica inerente à adição incorreta de nucleotídeos pela enzima empregada, foi feita clonagem de uma sequência já conhecida empregando-se o mesmo protocolo das amostras de soro estabelecido anteriormente, desde a amplificação até a clonagem e coleta de uma população de 30 clones. Todos os clones, em teoria, deveriam apresentar exatamente a mesma sequência e deveriam coincidir com a sequência original, uma vez que provinham de uma única sequência já estabelecida, e assim, qualquer erro encontrado terá sido artefato da metodologia aplicada. De um total de 29 clones, foi obtido um fragmento de 400pb em que uma única alteração nucleotídica foi detectada. Logo, de um total de 11.600 (29×400) nucleotídeos amplificados, um nucleotídeo errôneo significa uma taxa de erro de $8,62 \times 10^{-5}$ ($1 / 11.600$).

4.2.2 Análise das populações virais obtidas

A clonagem dos produtos amplificados gerou ao todo 236 clones como detalhado na Tabela 5. Foram coletados em média 23 clones por amostra de soro, sendo os animais U78 e U82 com o menor número de clones. As amostras de soro destes animais, no entanto, continham carga viral que resultou na introdução de 469 e 12.110 cópias, respectivamente, para primeira etapa de amplificação. Portanto, o número de variantes reduzido nas populações virais não necessariamente teve ligação com o menor campo amostral. Os animais U114 (pontos de coleta 2 e 3) e B25 apresentaram o maior número de variantes, sendo a amostra de soro U114.3 composta inteiramente por clones diferentes entre si. Um clone era considerado variante quando possuía ao menos um nucleotídeo diferente dos demais clones que compunham a população viral em análise.

Tabela 5. Distribuição de clones e variantes

Identificação	Nº de clones	Nº de variantes
U64	25	11
U74	22	6
U78	14	4
U82	14	2
U114.2	26	21
U114.3	25	25
B25	29	21
M42	27	4
348	25	8
674	29	11
Total	236	113

Através do cálculo da distância de Hamming, foi possível visualizar que uma mesma população viral de maior frequência foi compartilhada entre os animais que eliminaram a infecção (U64, U74, U78, U82), mas não foi possível detectar a mesma população em nenhum dos dois pontos de coleta do animal crônico (amostras de soro U114.2 e U114.3) dentro do mesmo haras. Isto é, provavelmente circulam populações virais diferentes entre os animais que eliminaram a infecção (U64, U74, U78, U82) e o animal crônico (U114) nos pontos de coleta 2 e 3. É possível que esta população não tenha sido detectada no animal U114 devido a um limite da técnica de clonagem

aplicada em acessar populações virais de menor frequência na população. A população viral como um todo dos animais que eliminaram a infecção (U64, U74, U78, U82) era constituída, em sua maior parte, por sequências com apenas uma variação em relação a de maior frequência. Para as coletas 2 e 3 do animal crônico (U114) houve maior número de variantes, sendo o terceiro ponto constituído inteiramente de variantes diferentes. Entre os dois pontos, as diferenças aconteceram em no mínimo quatro sítios e no máximo 14 sítios.

Curiosamente, a população viral do animal 674 (ES) era composta por duas populações de maior frequência. Uma destas era compartilhada com a população dos animais da UFRRJ que eliminaram a infecção (U64, U74, U78, U82) e a outra apresentava apenas um único nucleotídeo de diferença desta. Por outro lado, dentro da população viral deste mesmo animal, houve diferença de até 59 sítios de nucleotídeos. A população viral de maior frequência do animal 348 (MS) apresentou apenas dois nucleotídeos diferentes em relação às respectivas maiores frequências dos animais que eliminaram a infecção (U64, U74, U78, U82) e a 674 (ES). Em contraste, entre animais do estado do RJ, o número de nucleotídeos diferentes entre as populações virais foi de no mínimo 65 sítios entre a região Metropolitana (Animais U64, U74, U78, U82, U114) e as regiões Central (Animal B25) e Noroeste (Animal M42); e no mínimo de 42 sítios entre a região Central (Animal B25) e Noroeste (Animal M42). A Figura 5 ilustra as diferentes distâncias de Hamming entre os animais pertencentes à UFRRJ; entre os animais da UFRRJ e animais de outros estados; e entre os animais da UFRRJ e animais de outras regiões do RJ; As populações que circulam nos estados RJ (UFRRJ), ES e MS são mais similares, embora com mais valores discrepantes (*outliers*), quando comparadas às populações virais dentro do mesmo estado do RJ (Central e Noroeste).

de 0,167% (M42) até 2,26 (U114.3) e a nível de nucleotídeo de 0,072% (U82) até 1,63% (B25).

A distância intra-populacional entre as populações virais do animal crônico (U114) e as populações virais dos animais que eliminaram a infecção (U64, U74, U78, U82) foram significativamente diferentes (p valor $<0,0001$) (Figura 6 à direita e Tabela S 1) e o terceiro ponto de coleta (U114.3) mostrou-se significativamente mais elevado do que o segundo ponto (U114.2), indicando um processo evolutivo acelerado e dinâmico neste animal ao longo do tempo (Figura 6 à esquerda e Tabela S 1). Animais de diferentes regiões do mesmo estado (Metropolitana, Central e Noroeste do RJ) não apresentaram diferença significativa na distância intra-populacional de suas populações virais, mas houve diferença para as populações virais dos animais de estados diferentes, apesar de que para as populações virais dos animais U64 (RJ) e 348 (MS) não houve diferença (Figura 7, Tabela S 2 e Tabela S 3). A população dos animais U82 e M42 mostraram-se as mais conservadas.

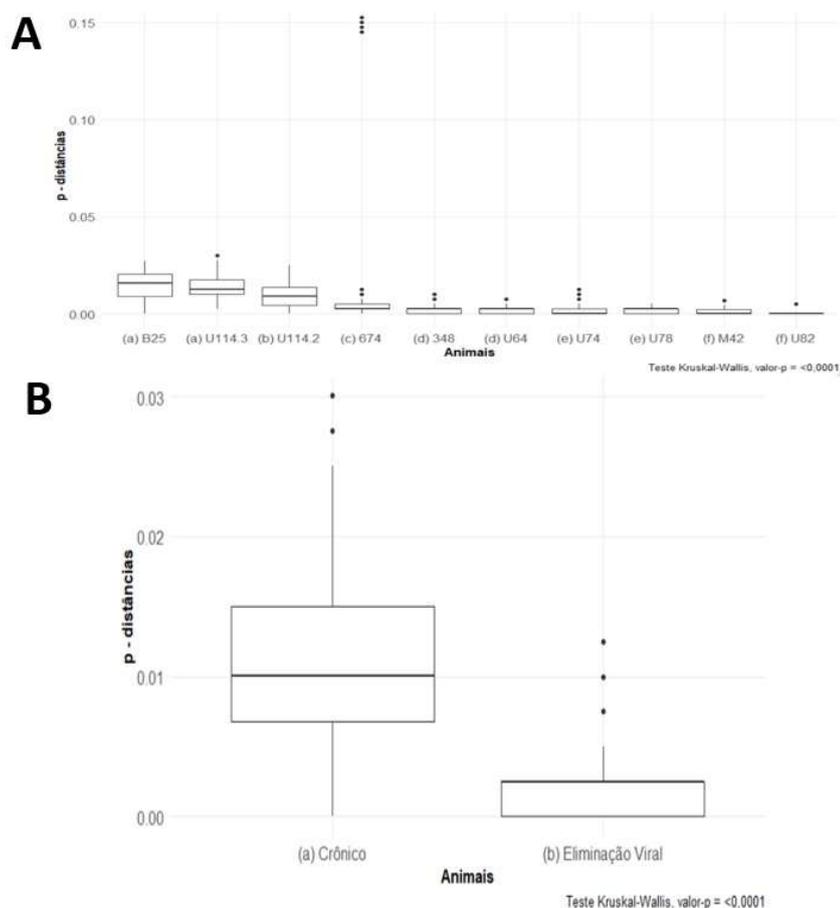


Figura 6 Gráfico de caixa da p-distância intra-populacional

Representação em gráficos de caixa das distâncias genéticas (p-distância) intra-populacionais de todos os animais separadamente (A) e entre as diferentes evoluções da infecção, eliminação/ crônico (B). Teste de Kruskal-Wallis, p -valor $<0,0001$.

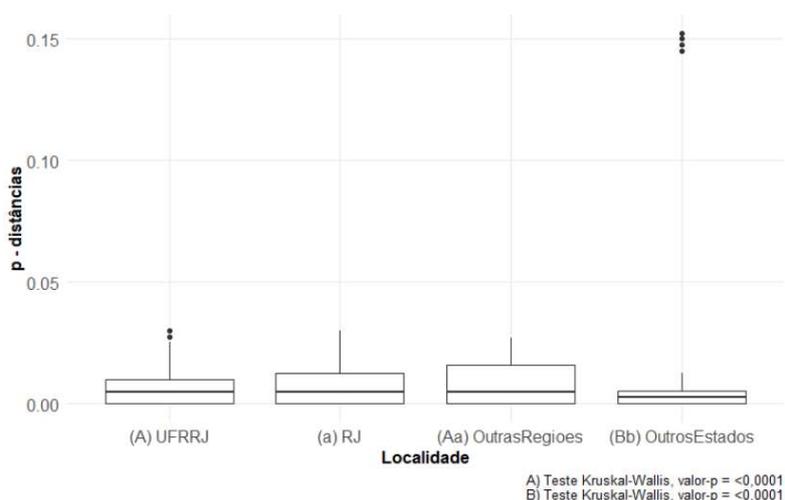


Figura 7 Gráfico de caixa da p-distância intra-populacional entre localidades

Representação em gráficos de caixa das distâncias genéticas (p-distância) intra-populacionais entre diferentes localidades. RJ compreende os animais U64, U74, U78, U82, B25 e M42; UFRRJ compreende os animais U64, U74, U78, U82; “Outros estados” compreende os animais 348 e 674; “Outras regiões” compreende os animais B25 e M42. Teste de Kruskal-Wallis, p-valor <0,0001 entre UFRRJ e Outras regiões (A) e entre RJ e outros estados (B).

Não houve correlação entre as distâncias genéticas intra-populacionais e os fatores de enzimas (AST) e carga viral (p-valor = 0,108 e 0,3677, respectivamente).

As distâncias inter-populacionais foram medidas através da p-distância entre todas as sequências de duas populações virais par a par, de forma que os 25 clones da população viral do animal U64 foram comparados aos 22 clones da população viral do animal U74 e assim sucessivamente. Desta forma, a maior variação inter-populacional ocorreu para as populações virais dos animais de regiões diferentes do RJ (B25 e M42) em relação a todas as outras populações (Figura 8, Tabela S 4). A população viral do animal do MS (348) não apresentou diferença inter-populacional significativa em relação às populações virais dos animais U64, U74 e U78 (grupo “w”) se mostrando equidistante entre elas. A mesma relação foi vista para a população viral do animal do ES (674) e as mesmas populações da UFRRJ, além das populações U114.2 e U82 (grupos “o”, “x” e “y”). As populações virais do ES e MS eram equidistantes em relação a duas populações do RJ, como é o caso da 348 (MS) e U82 (RJ/Metropolitana) equidistante da população viral B25 (RJ/Central) (grupo “j”) ou a 674 (ES) e as populações U78 e U64 equidistantes da população M42 (grupos “f” e “g”). Observa-se um padrão em que as populações de outros estados são mais próximas às populações virais do RJ/Metropolitana (UFRRJ), do que as populações desta região que são de outras localidades no mesmo estado (Figura 8).

As distâncias inter-populacionais entre as populações virais dos animais que eliminaram a infecção não foi significativa (grupo “z”), exceto entre as populações virais dos animais U64 e U82 (grupo “u”) (Figura 8, Tabela S 4). As distâncias inter-populacionais entre a população viral de cada coleta do animal crônico (U114.2 e U114.3) e os animais que eliminaram a infecção (U64, U74, U78, U82) não foram significativas (grupos “o”, “p” e “q” para a população U114.2; grupos “s” e “t” para a população U114.3 na (Figura 8, Tabela S 4). Isto é, a população viral de todos os animais que eliminaram a infecção está equidistante às populações virais dos dois pontos de coleta do animal crônico (U114) (Figura 8).

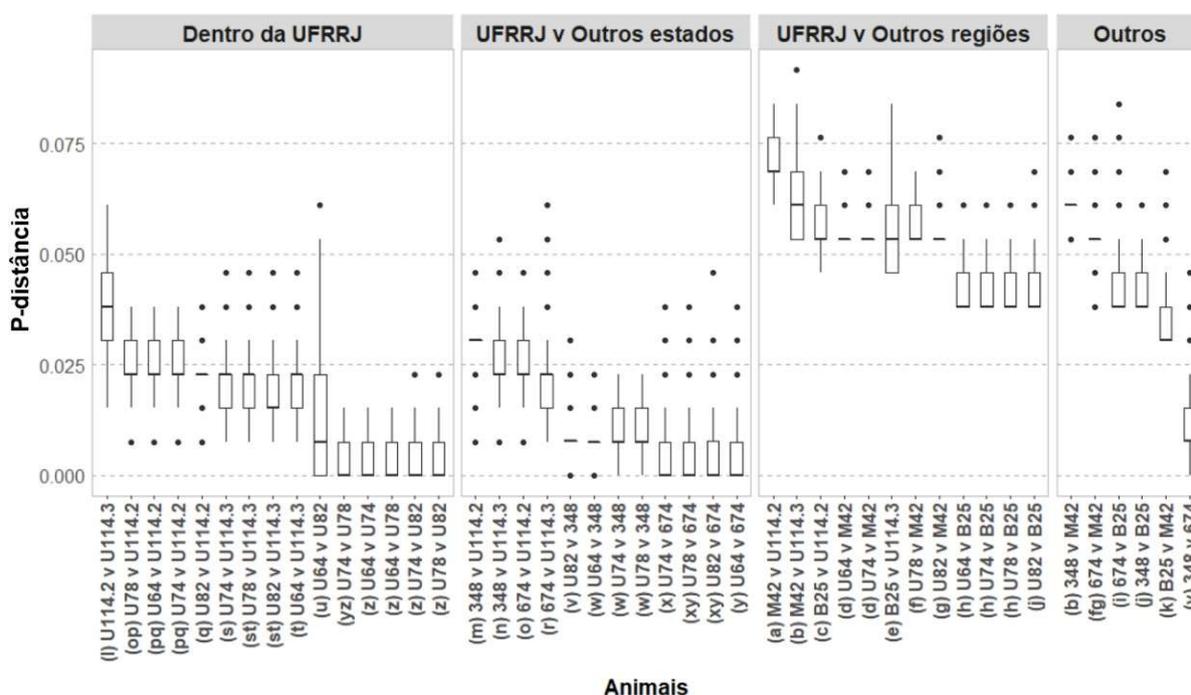


Figura 8 Gráfico de caixa da p-distância inter-populacional

Representação em gráficos de caixa das distâncias genéticas (p-distância) inter-populacionais entre todos os animais. UFRRJ compreende os animais U64, U74, U78, U82; “Outros estados” compreende os animais 348 e 674; “Outras regiões” compreende os animais B25 e M42; “Outros” compreende os animais 674, 348, B25 e M42. Teste de Kruskal-Wallis, p-valor <0,0001.

Quando analisadas as seqüências majoritárias (obtidas a partir do sequenciamento direto dos produtos de PCR) dentro do Haras UFRRJ, a distância entre os animais que eliminaram o vírus (U64, U74, U78 e U82) é nula sendo virtualmente idênticas. Entretanto entre estes e o animal crônico (U114) houve uma distância a nível de nucleotídeo de 1% (coletas 1 e 3) e 0.8% (coleta 2). Curiosamente, a distância a nível de amino ácido entre as amostras que eliminaram a infecção e o tempo de coleta 1 do animal crônico é nula, ao passo que a distância para os pontos seguintes é equalizada em 0.8% e entre os dois pontos crônicos é de 1.7%. Portanto, houve uma diferenciação mais expressiva entre os pontos de coleta 2 e 3 do animal

crônico (Figura 9). A p-distância calculada entre as sequências majoritárias do Brasil (deste trabalho) com as coletadas no GenBank (Tabela S 5) é de 3,54% a nível de aminoácido e 12,57% a nível de nucleotídeo, menor do que a maior distância entre as sequências do Brasil, em relação às majoritárias dos animais B25 e M42 (~5% e 17% para aminoácidos nucleotídeos).

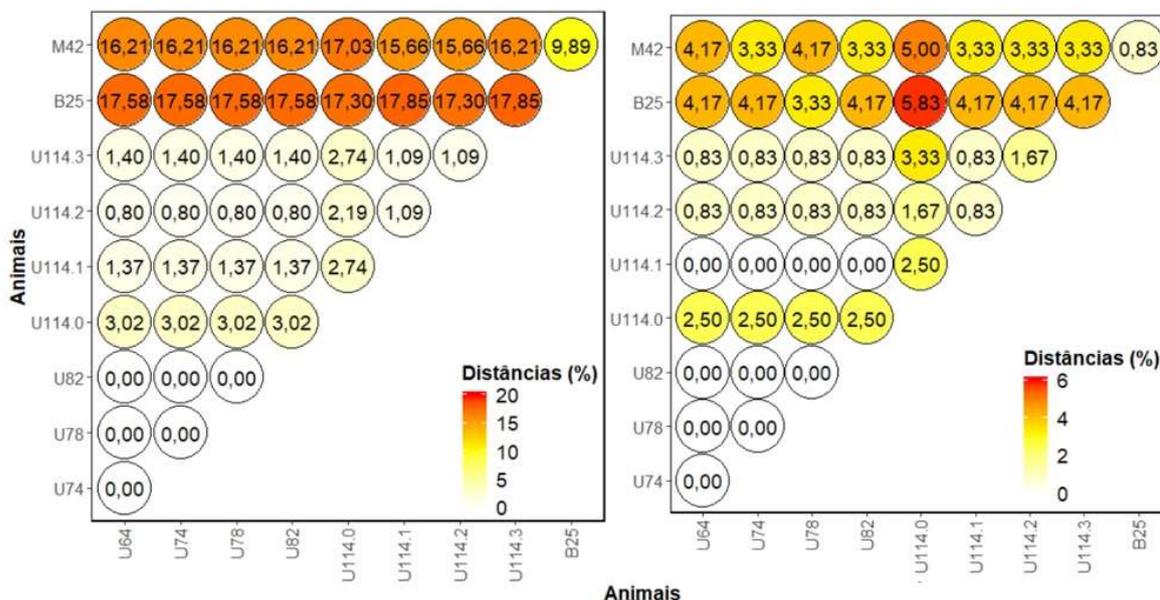


Figura 9 Gráfico de balão da p-distância entre as sequências majoritárias

Representação gráfica das p-distâncias entre as sequências majoritárias a nível de nucleotídeo (à esquerda) e aminoácido (à direita).

Quando as sequências majoritárias foram comparadas com as respectivas populações virais, o animal U114 apresentou a população majoritária mais distante da população viral obtida para os dois pontos de coleta analisados (1,3%). As populações dos animais que eliminaram a infecção (U64, U74, U78) se mostraram conservadas em relação às sequências majoritárias com médias de 0,09% a 0,12%, sendo a população do animal U82 a mais conservada (0,018%). A população do animal M42 também se mostrou conservada em relação à sequência majoritária (0,048%), enquanto o animal B25 apresentou maior distância entre a sequência majoritária e a população viral (1,04%).

A distância de Hamming demonstrou que dentro dos animais que eliminaram a infecção (U64, U74, U78, U82) as populações virais se distanciavam da sequência majoritária de 0 até 3 sítios, sendo a população de maior frequência coincidente com a sequência majoritária. O animal crônico (U114.2 e U114.3) apresentava distâncias de no mínimo 2 até 7 sítios diferentes entre as populações virais e as respectivas sequências majoritárias das coletas 2 e 3. Por sua vez, suas sequências majoritárias se distanciavam por 4 nucleotídeos entre si, mostrando que provavelmente circulavam populações diferentes entre os dois pontos de coleta. O animal M42 também mostrou

conservação das populações virais em relação à sequência majoritária, com no máximo 2 sítios diferentes. Assim como para os animais que eliminaram a infecção, a população de maior frequência do animal M42 coincidiu com a sequência majoritária. Por outro lado, a amostra B25 mostrou diferenças de 0 até 9 sítios entre as populações virais e a sequência majoritária, mostrando um perfil similar ao do animal crônico (U114). As amostras 348 (MS) e 674 (ES) não apresentaram sequenciamento satisfatório para as sequências majoritárias, logo não puderam ser incluídas na análise.

4.2.2.2 Localização das variações no genoma viral

Para determinar a região no fragmento amplificado onde se concentravam as variações, foram construídos gráficos com auxílio do programa SSE v.1.3 (117). A Figura 10 mostra a análise em janela deslizante da p-distância de aminoácidos e nucleotídeos para todas as dez populações virais analisadas, onde pode ser visto picos mais acentuados próximos a região de 200pb, que corresponde à região HVR1 descrita anteriormente. Os picos da distância intra-populacional acompanham em menor grau os picos da distância inter-populacional, mas este último é mais acentuado.

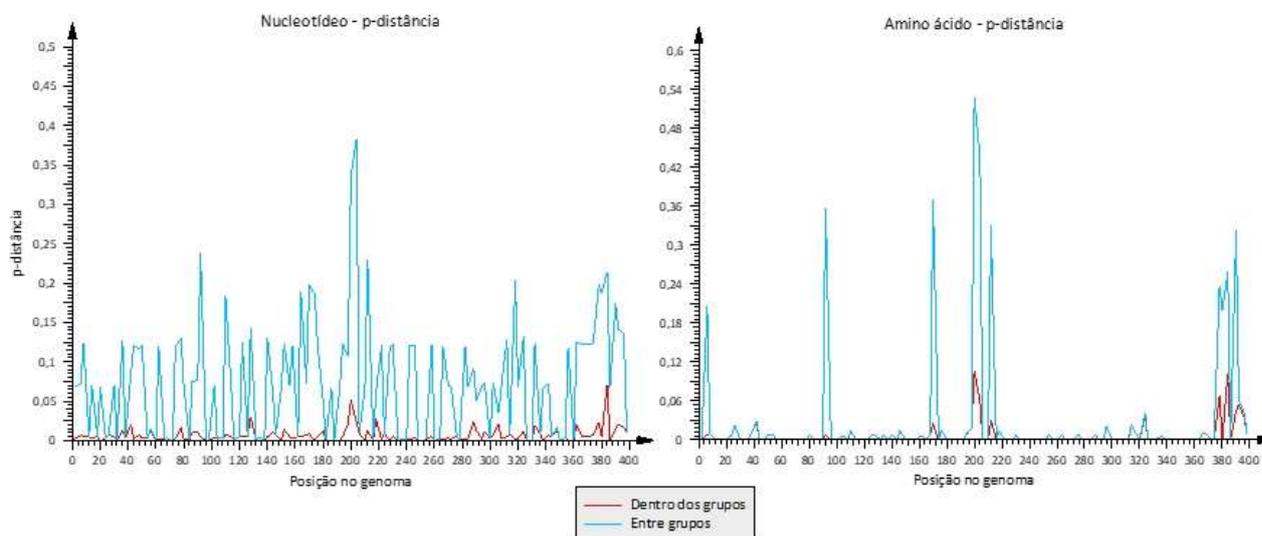


Figura 10. Gráficos de p-distância e localização genômica

Gráfico de p-distância de nucleotídeos (à esquerda) e aminoácidos (à direita) das 236 sequências analisadas neste trabalho. A linha vermelha demarca a distância dentro de cada grupo (intra-populacional); a linha azul demarca a distância entre os grupos (inter-populacional).

Quando analisadas apenas as populações do haras na região metropolitana do RJ, percebe-se que os picos acentuados ocorrem por conta do animal U114, mostrando um padrão diferenciado de evolução para a amostra crônica (Figura 11). A

região terminal do fragmento, correspondente a HVR2 descrita anteriormente, é acentuada ao comparar-se as coletas de 2016 e 2017, sendo mutações não-sinônimas predominantes nesta região. Indicando que pode ter havido alterações da partícula viral e da interação com o sistema imune do animal neste intervalo de tempo.

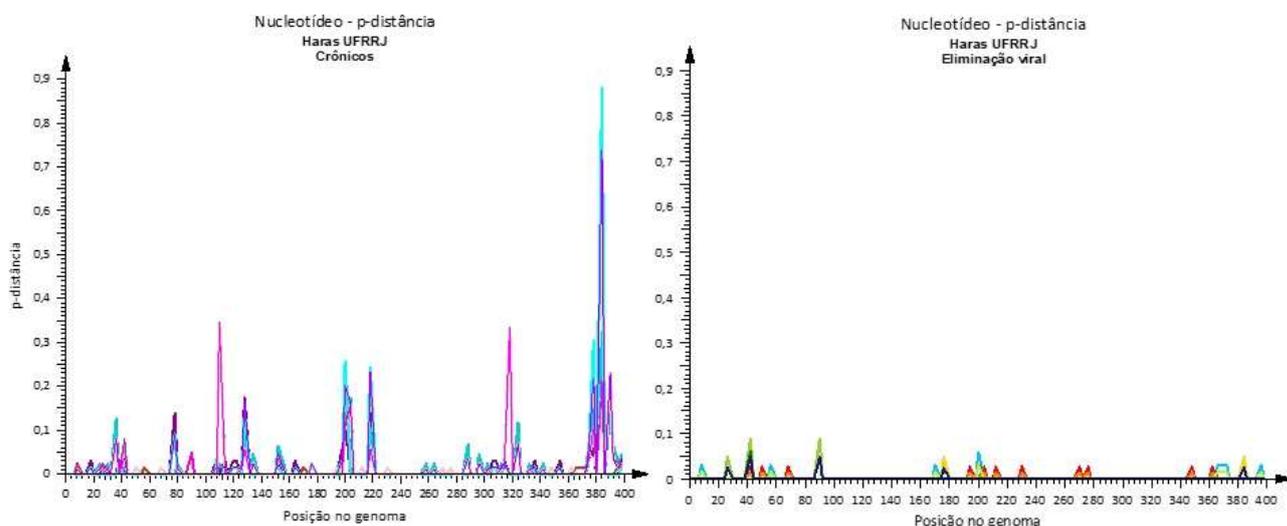


Figura 11. Gráficos de p-distância e localização genômica no animal crônico

Gráfico de p-distância nucleotídica no animal crônico (à esquerda) e nos animais que eliminaram a infecção (à direita). Animal crônico – U114, pontos de coleta 2 e 3; Animais de eliminação viral – U64, U74, U78 e U82

Como as sequências majoritárias dos animais que eliminaram a infecção (U64, U74, U78 e U82) eram as mesmas, apenas foi possível compará-las às sequências do animal crônico (U114,2 e U114.3) (Figura 12). O perfil de variações que ocorre dentro do grupo do animal crônico (linha vermelha) é o mesmo que entre os grupos de eliminação e crônico (linha azul), exceto por dois pontos que ocorrem apenas para mutações não sinônimas (à esquerda) entre os grupos. A média da distância entre estes dois grupos quando considerada apenas os sítios sinônimos é de 4,95% em comparação à média da distância calculada através de mutações não-sinônimas, de valor 0,49%.

Não houve diferença no perfil de distribuição de variações entre as quatro coletas do animal crônico (U114), seguindo o mesmo padrão que entre o animal crônico e os animais que eliminaram a infecção (Figura 12). As distâncias eram maiores quando considerados os sítios sinônimos (6,56%) em comparação aos sítios não-sinônimos (0,84%). Não houve temporalidade nas distâncias genéticas entre as coletas, uma vez que a distância entre as coletas 0 e 1 (1,19%/12,29%) é maior que entre as coletas 0 e 2 (0,774%/9,16%) tanto para sítios não-sinônimos quanto para sítios sinônimos.

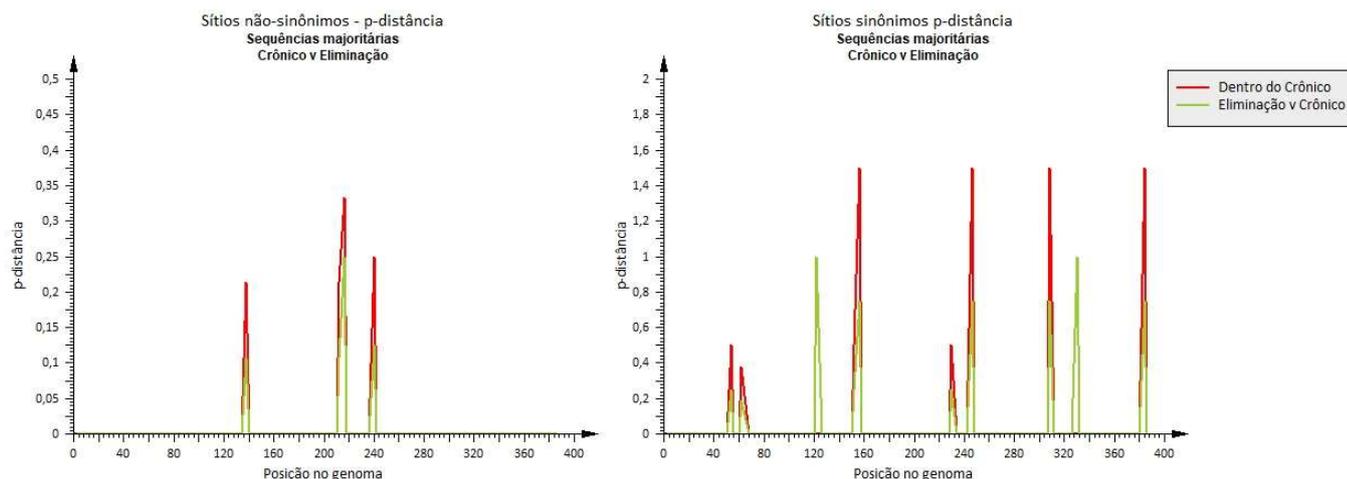


Figura 12 Gráfico de p-distância e localização genômica das sequências majoritárias: Crônico v Eliminação

Gráfico de p-distância de sítios não-sinônimos (à esquerda) e sítios sinônimos (à direita). O grupo “Crônico” (linha vermelha) compreende as sequências U114.0, U114.1, U114.2 e U114.3. O grupo “Eliminação” compreende as sequências U64, U74, U78 e U82. A linha verde demarca a distância entre os grupos “Crônico” e “Eliminação”.

4.3 Análise Filogenética

A análise filogenética por Máxima Verossimilhança das populações de quasiespécies, em conjunto com suas respectivas sequências majoritárias apresentou formação de três clados (Figura 13). O primeiro deles foi formado pelas populações virais dos animais que eliminaram a infecção (U64, U74, U78 e U82 – círculos azuis); dos dois pontos de coleta do animal crônico (U114.2 e U114.3 – quadrados verdes); das populações virais dos animais do ES e MS (674 e 348 – triângulos lilases); além das respectivas sequências majoritárias (marcadas em vermelho com os respectivos símbolos). O segundo e terceiro clado agruparam separadamente as populações virais dos animais das regiões Central e Noroeste do estado do RJ (B25 e M42 - losangos pretos), apesar de uma única sequência da população do animal 674 (triângulo lilás) agrupar com a população viral do animal B25. Não houve força estatística (*bootstrap* >75) para diferenciar as populações de diferentes evoluções da infecção, nem entre estados, mas populações de regiões diferentes do RJ se separaram, reforçando que distâncias geográficas não interferem na distância genética entre populações virais do EqHV.

Análise filogenética construída pelo algoritmo de Inferência Bayesiana das sequências majoritárias em conjunto com sequências coletadas no GenBank (Tabela S 5) formaram 2 cladogramas, um dos quais se subdivide em 2 subclados maiores (Figura 14). O primeiro clado separa amostras internacionais do EUA e Alemanha do resto das sequências. O segundo clado, diferencia um subclado contendo as sequências majoritárias dos animais M42 (RJ/Noroeste) e B25 (RJ/Central). Estas têm grupamento com uma sequência da Alemanha, as quais por sua vez são agrupadas por sequências do Japão e EUA. As sequências majoritárias dos animais que eliminaram a infecção (U64, U74, U78 e U82) e do animal crônico (U114) se encontram em ramos mais internos da árvore e agruparam com uma sequência da Alemanha. Apresentando topografia parafilética, o ponto de coleta U114.0 se separa com robustez estatística de um clado formado pelas três coletas consecutivas do animal crônico (U114.1, U114.2 e U114.3) e os animais que eliminaram a infecção (U64, U74, U78 e U82).

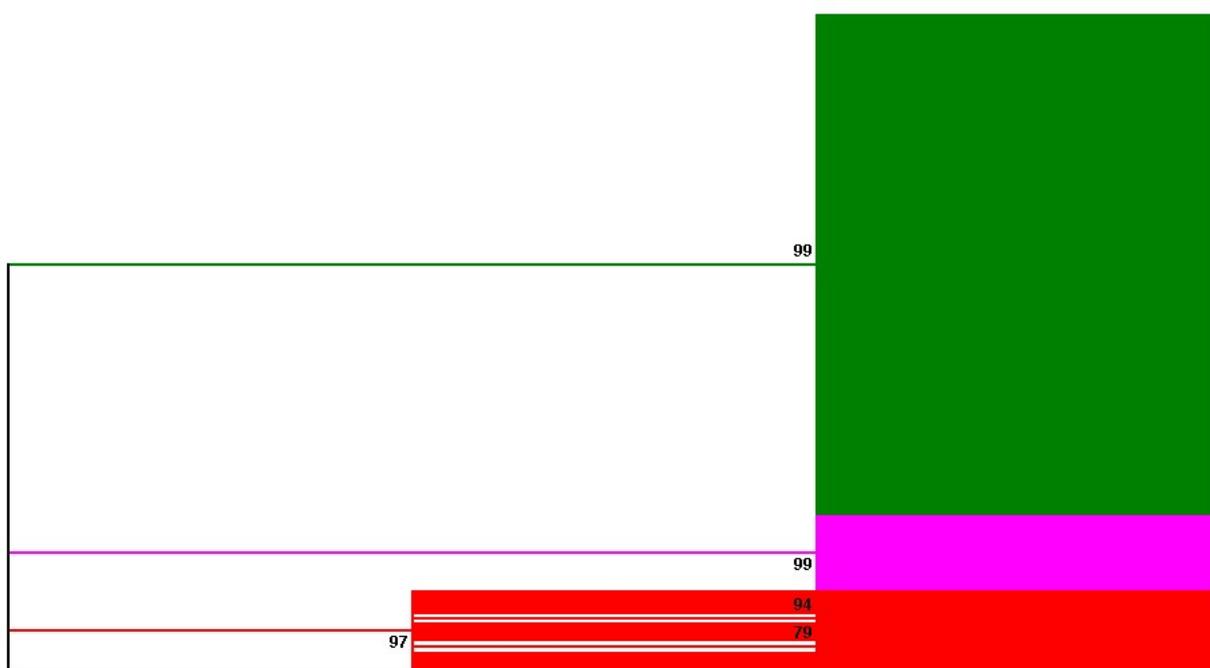


Figura 13. Árvore filogenética de populações de quasiespécies do EqHV

Árvore filogenética condensada (valor de corte 75) pelo método de Máxima Verossimilhança (1000 repetições) utilizando o modelo evolutivo Kimura 2 parâmetros + G + I, conforme calculado no programa MEGA v7.0. O clado verde contém as populações virais que eliminaram a infecção (U64, U74, U78, U82), as populações virais do animal crônico nos pontos de coleta 2 e 3 (U114.2 e U114.3), as populações virais dos animais dos estados de ES (674) e MS (348); Os cladogramas vermelho e rosa contêm as populações virais dos animais do estado do RJ na região Central (B25) e Noroeste (M42), respectivamente. O grupo externo desta análise foi uma sequência de HCV (AF009606). As sequências majoritárias e encontram junto às suas respectivas populações.

5 DISCUSSÃO

Este trabalho foi desenvolvido com intuito de explorar a variabilidade do EqHV sob a ótica da teoria das quasiespécies virais em contextos de diferentes evoluções da infecção. A premissa deste trabalho foi construída sob o fato de que equinos infectados pelo EqHV são capazes de estabelecer a infecção de forma crônica, mas que estes são minoria em relação aos animais que eliminam a infecção viral (6). Nossa hipótese, à luz da teoria das quasiespécies virais, é que a maior tendência à eliminação se dá por uma dinâmica populacional específica (58,67,93). Dissecar esta dinâmica populacional poderia trazer ramificações para entender porque algumas infecções tendem a evoluir para infecção crônica, como é o caso do vírus da hepatite C (120).

O alvo molecular escolhido para acessar a variabilidade do EqHV foi o envelope viral (genes E1 e E2), onde foram descritas regiões de hipervariabilidade (112) e que é muito utilizado em pesquisas similares para o HCV (95,98). Desta forma, construímos oligonucleotídeos iniciadores para esta região com base na descrição do genoma do EqHV por Kapoor e colaboradores (2011) e alinhamento múltiplo de sequências depositadas no banco de dados GenBank (Figura 3 e Tabela 3). O processo de padronização do alvo construído levou à utilização de uma temperatura de hibridização de alta estringência para se evitar a amplificação de inespecificidades que poderiam intervir com os processos até a obtenção das sequências únicas. Mesmo utilizando uma amplificação sob alta estringência e em condições sub-ótimas, populações de quasiespécies ainda podem ser acessadas sem que haja grande diferença nas populações obtidas, de forma que a escolha de temperatura de hibridização não interfere de forma impeditiva na análise (86).

A clonagem molecular é o padrão ouro para obtenção de populações de quasiespécies (85,86), mas está suscetível ao erro de reamostragem, logo a carga viral das amostras deve-se ser levada em consideração para que não haja viés nas análises resultantes (87). Conseqüentemente, a coleta 0 do animal U114 (Tabela 2) não pôde ser analisada quanto à população de quase espécie, mas foi mantida na população de estudo para análise da sequência majoritária. Segundo a fórmula para calcular a probabilidade de reamostragem obtida por Liu e colaboradores (1996), quando são aplicadas 65 cópias virais em uma reação de amplificação e são coletados 20 clones para análise da diversidade genética, a proporção de clones com sequências únicas deve ser de aproximadamente 90% (121,122). O número mínimo

de cópias virais aplicadas nas reações de PCR deste trabalho foi 57 para a amostra de soro do animal U64, bem próximo ao exemplo acima. Para esta amostra de soro, foram coletados 25 clones, sugerindo que o viés gerado por reamostragem é ao menos menor que 90%. O número de cópias virais aplicadas das outras amostras de soro deste trabalho foi superior e considerando a média de 23 clones coletados, o erro de reamostragem pode não afetar a diversidade viral apresentada.

Resultados preliminares com os animais que eliminaram a infecção (U64, U74, U78 e U82) mostravam baixa variabilidade para uma região de hipervariabilidade sugerindo uma rede de infecção compartilhada entre os animais presentes no haras UFRRJ. O cálculo de distância de Hamming mostrou que estes animais compartilhavam a mesma população de maior frequência e que a população como um todo era, em maior parte, constituída por clones que variavam em apenas um nucleotídeo (Figura 5). A análise da distância genética corrobora um processo evolutivo levemente diferente a nível intra-populacional, mas equidistante a nível inter-populacional entre estes animais. Estes dados corroboram a tendência à eliminação da infecção de populações homogêneas, como ocorre com o HCV (109).

Ainda assim, tendo em vista a ausência de trabalhos semelhantes em que se buscam quase-espécies em animais naturalmente infectados pelo EqHV, as baixas distâncias genéticas para a região de envelope levaram à preocupação de que a variabilidade visualizada era artificial, produto do erro de incorporação da polimerase utilizada no processo de amplificação, e não biológica consequente da replicação viral. A escolha da polimerase para acessar a diversidade de um vírus com altas taxas de mutação deve ser feita com cuidado, pois pode acarretar na hiperestimação da variabilidade estimada (85). Apesar de ser uma enzima de alta fidelidade, esta era composta por uma razão entre uma enzima de baixa fidelidade e uma enzima de alta fidelidade como forma de explorar tanto a processividade quanto a fidelidade de ambas (Platinum™ *Taq* DNA Polymerase High Fidelity - *Thermo Fischer*). A taxa de erro intrínseca aos nossos experimentos foi de $8,62 \times 10^{-5}$ erros por nucleotídeo amplificado, enquanto a taxa de erro do HCV é de $2,5 \times 10^{-5}$ ($1,6 - 6,2 \times 10^{-5}$) mutações por sítio por ciclo replicativo (modelo para este estudo, na ausência de dados do EqHV) (123), mas o alvo molecular estudado é uma região de alta variabilidade atingindo valores na ordem de grandeza de 10^{-2} e 10^{-3} (98,124), logo as variações encontradas neste trabalho foram seguramente consequência da biologia do vírus.

Uma maior variabilidade viral poderia ser alcançada levando-se em consideração populações virais de diferentes localizações geográficas, assim foram

adicionados animais infectados de outras regiões do estado do RJ e de outros estados do Brasil (ES e MS). Surpreendentemente, as populações de outros estados mostraram-se mais próximas (distâncias inter-populacionais) às do haras da UFRRJ do que as de dentro do mesmo estado (Figura 8). A distância de Hamming mostrou que uma das populações de maior frequência do animal 674 (ES) era a mesma que circulava nos animais que eliminaram a infecção no haras UFRRJ. Para o animal 348 (MS), houve diferenças entre 1 a 5 sítios. As distâncias inter-populacionais não eram significativamente diferentes entre os animais do haras UFRRJ e os animais 674 (ES) ou 348 (MS). Em contraste, para populações do mesmo estado, a distância de Hamming era de no mínimo 42 nucleotídeos diferentes, mostrando não haver relação entre a distância geográfica e a distância genética. Estes resultados demonstram a falta de associação com a localização geográfica e a circulação de populações virais idênticas em diferentes estados brasileiros, acessada pelo gene de envelope, reforçam a hipótese de uma distribuição antropogênica do EqHV, como já sugerido em trabalhos anteriores (15).

Em vista da baixa diversidade viral encontrada nos animais que eliminaram a infecção no haras da UFRRJ e de evidências da circulação de uma mesma população viral de maior frequência entre eles, seria esperado encontrar um perfil similar para o animal crônico. De fato, o número de variantes presentes nesta amostra era maior (Tabela 5), assim como a distância genética intra- e inter-populacional (Figura 6 e Figura 8). A população viral do animal crônico também foi responsável pelas variações mais expressivas ao longo do fragmento genômico de envelope (Figura 11) e mutações não-sinônimas faziam grande parte da diferenciação entre os dois pontos de coleta (Figura 12). Estes resultados apontam para uma dinâmica populacional mais acelerada no curso da infecção crônica, característica associada à persistência da infecção no HCV (124).

A análise da p-distância ao longo do fragmento de envelope também reforça a validade da variabilidade encontrada uma vez que são postas dentro de um contexto biológico. Isto é, não são mutações únicas em sítios aleatórios, mas tem significado biológico (85,101), visto que se concentraram em torno da posição de 200pb do fragmento amplificado, onde é localizada a região HVR1 descrita para o EqHV (3,112). Entre os dois pontos de coleta do animal crônico se destacou uma segunda região de alta variabilidade correspondente à HVR2 (112). Altas concentrações de mutações não sinônimas nestas regiões hipervariáveis também são descritas para o HCV e estão associadas à presença de epítomos para anticorpos neutralizantes (96). As

mutações acumuladas nestas regiões do fragmento amplificado neste trabalho são consistentes com a incidência de pressão imune semelhante e sugerem a presença de epítomos para anticorpos neutralizantes para o EqHV (15,112). Os diferentes perfis de população viral entre os animais que eliminaram a infecção e o que evoluiu para infecção crônica sugerem escape imunológico, gerando as diversidades aqui descritas. No entanto, o escape imunológico não é o único mecanismo de persistência viral, uma vez que foi relatado um animal com população homogênea que manteve infecção persistente durante o período observado (112).

As análises de sequências majoritárias, isto é, sequências obtidas diretamente após as duas etapas de amplificação, reforçam a circulação de uma mesma população viral entre os animais que eliminaram a infecção visualizada através da distância de Hamming nula, mas na amostra de soro do animal crônico obtida no ano seguinte (U114.1) e nos anos consecutivos (U114.2 e U114.3) mostram uma população diferente, onde a sequência majoritária não coincide com nenhuma variante das populações virais obtidas. A sequência majoritária do ponto de coleta U114.0 também não coincide com as dos animais que eliminaram a infecção, evidenciando que o processo de evolução pode ter ocorrido no espaço de alguns meses que separam as coletas destes animais. Evolução rápida em curtos períodos de tempo já foram descritas para o HCV, assim como para o EqHV (15,98,112). As distâncias genéticas máximas foram alcançadas dentro das sequências majoritárias do Brasil, atingindo até 18% a nível de nucleotídeo e 6% a nível de aminoácido para sequências do estado do RJ. As sequências do Brasil se distanciaram em 12,57% e 3,54% das sequências coletadas no GenBank. Apesar de ser uma região altamente variável, estes valores são inferiores às distâncias encontradas para o gene NS5B, um gene altamente conservado de replicação viral (32).

Quando comparadas as sequências majoritárias às populações virais respectivas, foi possível visualizar faixas com ordens de grandeza diferentes entre os animais que eliminaram a infecção (~0,1%) e os animais crônicos (1,3%), sugerindo que as populações virais entre estes dois grupos possuem dinâmica de evolução diferentes. A distância de Hamming mostrou variações de 2 a 7 sítios entre as sequências majoritárias e a população viral do animal crônico e de 0 até 3 sítios diferentes para os animais que eliminaram a infecção. Em outras palavras o grupo de animais que eliminou a infecção apresentava população viral mais próxima à sequência majoritária, enquanto o grupo crônico se distanciava. Esta dinâmica se aproxima do deslize antigênico, perfil de população associada a pacientes

cronicamente infectados pelo HCV que apresentavam quadros clínicos mais brandos de patologia hepática (95,98).

Seguindo análise das sequências majoritárias, as variações entre os grupos de eliminação viral e das coletas do animal crônico e a variação entre coletas segue um mesmo perfil ao longo do genoma, ocorrendo maiores índices de mutações sinônimas (Figura 11). Percebe-se, portanto, que não há um perfil diferenciado na distribuição das variações entre os dois grupos.

Não houve relação significativa entre os níveis da enzima AST e a carga viral sérica em relação às distâncias genéticas. Já foi relatado não haver correlação entre estes fatores nem em relação à severidade da doença hepática incidente pelo HCV (98,125), nem correlação entre a diversidade das quasiespécies e a carga viral para o EqHV (112).

Por último, pela análise filogenética observou-se não ocorrer a separação de populações virais de diferentes animais, nem diferentes localizações (Figura 13), tornando o cálculo de sinal temporal inviável, pois populações de diferentes pontos temporais são estatisticamente consideradas uma única população. A análise filogenética das sequências majoritárias mostrou topografia parafilética para as amostras da UFRRJ (Figura 14), havendo separação entre as sequências dos animais que eliminaram a infecção (U64, U74, U78, U82) e as sequências das amostras de soro do animal crônico (U114.1, U114.2 e U114.3) em dois clados que se agrupavam através da sequência da amostra de soro U114.0. Ramsay e colaboradores (2018) analisaram a população de quasiespécies de 7 animais experimentalmente infectados e também não foram capazes de diferenciar as populações mesmo em animais com contextos imunológicos diferentes (112). Igualmente, Walter e colaboradores (2017) utilizaram um burro cuja infecção com o EqHV era datada de 1979 para calcular o sinal temporal de suas amostras, mas também não obteve sucesso (15).

Nesta população experimentalmente infectada, as distâncias intra-populacionais a nível de aminoácido variavam de 0,37% para potros imunocomprometidos, chegando a 4,4% em potros imunocompetentes e 2,2% em cavalos adultos (112). Uma faixa de valores similar foi observada para as populações de quasiespécies deste trabalho de 0,165% a 2,26%, sendo os menores valores encontrados para os animais que eliminaram a infecção e os maiores para as populações dos animais que evoluíram para infecção crônica. Exceto para os potros imunocomprometidos, um perfil similar de maior diversidade em populações que evoluíram para a infecção crônica e menor diversidade para populações que eliminaram a infecção se repetiu

nesta população (112). No entanto, os animais adultos que eliminaram a infecção experimental possuíam diversidade intra-populacional similar ao cavalo de infecção crônica deste trabalho. De fato, Ramsay; Evanoff e Mealy (2018) não encontraram associação entre a diversidade de quasiespécies e a evolução da infecção ou estado imunológico, indo de frente com os dados produzidos neste trabalho. Reforça-se, portanto, a necessidade de realizar mais estudos como este.

A topografia da árvore das sequências majoritárias quando consideradas no contexto mundial apresentou topografia complexa com muitos ramos internos e mesclando as sequências deste trabalho com as sequências coletadas do Genbank distribuídas pelo globo sem que houvesse um padrão geográfico. Para o HCV, mesmo com os genes de envelope, é possível verificar separação dos 7 genótipos (126), o que não foi possível para o EqHV, pois a árvore apresentada neste trabalho é mais complexa do que a árvore para o gene da NS5B (32).

Portanto, este trabalho corrobora a associação entre diversidade viral e persistência da infecção em que uma população homogênea tende à eliminação viral, enquanto uma população de maior diversidade tende à persistência. Não houve relação entre a distância geográfica e a distância genética, pois populações virais de diferentes estados apresentaram populações mais próximas do que populações de dentro do mesmo estado, reforçando uma distribuição antropogênica do vírus. Entretanto, é necessária uma avaliação com amostragem mais ampla para aprofundar os resultados aqui apresentados

6 CONCLUSÕES

- Foi possível acessar a diversidade e dinâmica de quasiespécies do EqHV utilizando a técnica de padrão ouro da clonagem molecular, tomando os devidos cuidados para se evitar artefatos intrínsecos ao seu uso.
- A diversidade das sequências nucleotídicas do envelope viral, em sua maioria, apontaram para uma maior variação sinônima no fragmento como um todo, mas prevalência de variações não sinônimas nas regiões correspondentes à HVR1 e HVR2.
- Os resultados obtidos através da distância de Hamming e p-distância mostraram que animais que eliminaram a infecção e o animal de infecção crônica apresentaram dinâmicas populacionais virais diferentes.
- Verificou-se associação entre diversidade viral e persistência da infecção em que uma população homogênea tende à eliminação viral, enquanto uma população de maior diversidade tende à persistência.
- Não houve relação entre a distância geográfica e a distância genética, pois populações virais de diferentes estados apresentaram populações mais próximas do que populações de dentro do mesmo estado, sugerindo uma distribuição antropogênica do vírus.

7 PERSPECTIVAS

As perspectivas deste trabalho envolvem realizar análises populacionais em um maior número de animais crônicos, diminuir a escala de tempo para identificar com maior resolução a dinâmica de quasiespécies neste grupo de animais, incluir animais infectados de mais localizações geográficas e estudar as quasiespécies diretamente do fígado avaliando diferenças entre compartimentos. Adicionalmente, estudar como se comporta a distribuição de variações no genoma completo do EqHV através de técnicas de NGS.

8 REFERÊNCIAS

1. Radford AD, Chapman D, Dixon L, Chantrey J, Darby AC, Hall N. Application of next-generation sequencing technologies in virology. *J Gen Virol*. 2012 Sep 1;93(Pt_9):1853–68.
2. Scheel TKH, Simmonds P, Kapoor A. Surveying the global virome: identification and characterization of HCV-related animal hepaciviruses. *Antiviral Res*. 2015 Mar;115:83–93.
3. Kapoor A, Simmonds P, Gerold G, Qaisar N, Jain K, Henriquez JA, et al. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci*. 2011 Jul 12;108(28):11608–13.
4. Choo Q, Amy J, Lacy R, Daniel W. Isolation of a cDNA Clone Derived from a Blood-Borne Non-A, Non-B Viral Hepatitis Genome. 1989;
5. Liang TJ. Current progress in development of hepatitis C virus vaccines. *Nat Med*. 2013 Jul 8;19(7):869–78.
6. Pfaender S, Cavalleri JM, Walter S, Doerrbecker J, Campana B, Brown RJ, et al. Clinical course of infection and viral tissue tropism of hepatitis C virus-like nonprimate hepaciviruses in horses. *Hepatology*. 2015 Feb;61(2):447–59.
7. Scheel TKH, Kapoor A, Nishiuchi E, Brock K V., Yu Y, Andrus L, et al. Characterization of nonprimate hepacivirus and construction of a functional molecular clone. *Proc Natl Acad Sci*. 2015 Feb 17;112(7):2192–7.
8. Smith DB, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. *J Gen Virol*. 2016;97(11).
9. Shi M, Lin X-D, Vasilakis N, Tian J-H, Li C-X, Chen L-J, et al. Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. *Ou J-HJ*, editor. *J Virol*. 2016 Jan 15;90(2):659–69.
10. Burbelo PD, Dubovi EJ, Simmonds P, Medina JL, Henriquez J a., Mishra N, et al. Serology-Enabled Discovery of Genetically Diverse Hepaciviruses in a New Host. *J Virol*. 2012 Jun 1;86(11):6171–8.
11. El-Attar LMR, Mitchell JA, Brooks Brownlie H, Priestnall SL, Brownlie J. Detection of non-primate hepaciviruses in UK dogs. *Virology*. 2015 Oct;484:93–102.
12. Lyons S, Kapoor A, Schneider BS, Wolfe ND, Culshaw G, Corcoran B, et al. Viraemic frequencies and seroprevalence of non-primate hepacivirus and equine pegiviruses in horses and other mammalian species. *J Gen Virol*. 2014 Aug 1;95(Pt_8):1701–11.
13. Pybus OG, Thézé J. Hepacivirus cross-species transmission and the origins of the hepatitis C virus. *Curr Opin Virol* [Internet]. 2016 Feb;16:1–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1879625715001443>
14. Postel A, Cavalleri J-M V., Pfaender S, Walter S, Steinmann E, Fischer N, et al. Frequent presence of hepaci and pegiviruses in commercial equine serum pools. *Vet Microbiol*. 2016 Jan;182:8–14.
15. Walter S, Rasche A, Moreira-Soto A, Pfaender S, Bletsa M, Corman VM, et al. Differential Infection Patterns and Recent Evolutionary Origins of Equine Hepaciviruses in Donkeys. *Ou J-HJ*, editor. *J Virol*. 2017 Jan 1;91(1):e01711-16.
16. Schlottau K, Fereidouni S, Beer M, Hoffmann B. Molecular identification and characterization of nonprimate hepaciviruses in equines. *Arch Virol*. 2019 Feb 25;164(2):391–400.
17. Behrendt P, Cavalleri J-M V, Pfaender S, Engelmann M, Klawonn F, Wölk B, et al. Assessment of cross-species transmission of hepatitis C virus-related non-primate hepacivirus in a population of humans at high risk of exposure. *J Gen Virol*. 2015 Sep 1;96(9):2636–42.
18. Bukh J. Animal Models for the Study of Hepatitis C Virus Infection and Related Liver Disease. *Gastroenterology*. 2012 May;142(6):1279-1287.e3.
19. Siddell SG, Walker PJ, Lefkowitz EJ, Mushegian AR, Adams MJ, Dutilh BE, et al. Additional changes to taxonomy ratified in a special vote by the International Committee on Taxonomy of Viruses (October 2018). *Arch Virol*. 2019 Mar 1;164(3):943–6.
20. Pfaender S, Brown RJ, Pietschmann T, Steinmann E. Natural reservoirs for homologs of hepatitis C virus. *Emerg Microbes Infect*. 2014;3(3).

21. Chevaliez S, Pawlotsky J-M. HCV Genome and Life Cycle. *Hepatitis C Viruses: Genomes and Molecular Biology*. Horizon Bioscience; 2006.
22. Moradpour D, Penin F. Hepatitis C Virus Proteins: From Structure to Function. In 2013. p. 113–42.
23. Lauck M, Sibley SD, Lara J, Purdy M a, Khudyakov Y, Hyeroba D, et al. A Novel Hepacivirus with an Unusually Long and Intrinsically Disordered NS5A Protein in a Wild Old World Primate. *J Virol*. 2013 Aug 15;87(16):8971–81.
24. Tanaka T, Kasai H, Yamashita A, Okuyama-Dobashi K, Yasumoto J, Maekawa S, et al. Hallmarks of hepatitis C virus in equine hepacivirus. *J Virol*. 2014;88(22).
25. Stewart H, Walter C, Jones D, Lyons S, Simmonds P, Harris M. The non-primate hepacivirus 5' untranslated region possesses internal ribosomal entry site activity. *J Gen Virol*. 2013 Dec;94(PART 12):2657–63.
26. Tanaka T, Otoguro T, Yamashita A, Kasai H, Fukuhara T, Matsuura Y, et al. Roles of the 5' Untranslated Region of Nonprimate Hepacivirus in Translation Initiation and Viral Replication. Dermody TS, editor. *J Virol*. 2018 Jan 17;92(7).
27. Boukadida C, Fritz M, Blumen B, Fogeron M-L, Penin F, Martin A. NS2 proteases from hepatitis C virus and related hepaciviruses share composite active sites and previously unrecognized intrinsic proteolytic activities. Randall G, editor. *PLOS Pathog*. 2018 Feb 7;14(2):e1006863.
28. Walter S, Bollenbach A, Doerrbecker J, Pfaender S, Brown RJP, Vieyres G, et al. Ion Channel Function and Cross-Species Determinants in Viral Assembly of Nonprimate Hepacivirus p7. 2016;90(10):5075–89.
29. de Albuquerque PPLF, Santos LHS, Antunes D, Caffarena ER, Figueiredo AS. Structural insights into NS5B protein of novel equine hepaciviruses and pegiviruses complexed with polymerase inhibitors. *Virus Res* [Internet]. 2020 Mar;278:197867. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168170219308032>
30. Parera M, Martrus G, Franco S, Clotet B, Martinez MA. Canine Hepacivirus NS3 Serine Protease Can Cleave the Human Adaptor Proteins MAVS and TRIF. Li K, editor. *PLoS One*. 2012 Aug 1;7(8):e42481.
31. Patel MR, Loo Y-M, Horner SM, Gale M, Malik HS. Convergent Evolution of Escape from Hepaciviral Antagonism in Primates. Hurst LD, editor. *PLoS Biol*. 2012 Mar 13;10(3):e1001282.
32. Figueiredo AS, Lampe E, do Espírito-Santo MP, Mello FC do A, de Almeida FQ, de Lemos ERS, et al. Identification of two phylogenetic lineages of equine hepacivirus and high prevalence in Brazil. *Vet J*. 2015 Dec 1;206(3):414–6.
33. Kim H-S, Moon H-W, Sung HW, Kwon HM. First identification and phylogenetic analysis of equine hepacivirus in Korea. *Infect Genet Evol*. 2017 Apr;49:268–72.
34. Pronost S, Hue E, Fortier C, Foursin M, Fortier G, Desbrosse F, et al. Prevalence of Equine Hepacivirus Infections in France and Evidence for Two Viral Subtypes Circulating Worldwide. *Transbound Emerg Dis*. 2017 Dec 1;64(6):1884–97.
35. Lu G, Sun L, Xu T, He D, Wang Z, Ou S, et al. First Description of Hepacivirus and Pegivirus Infection in Domestic Horses in China: A Study in Guangdong Province, Heilongjiang Province and Hong Kong District. Melcher U, editor. *PLoS One*. 2016 May 16;11(5):e0155662.
36. Elia G, Lanave G, Lorusso E, Parisi A, Cavaliere N, Patruno G, et al. Identification and genetic characterization of equine hepaciviruses in Italy. *Vet Microbiol*. 2017 Aug;207:239–47.
37. Badenhorst M, Tegtmeyer B, Todt D, Guthrie A, Feige K, Campe A, et al. First detection and frequent occurrence of Equine Hepacivirus in horses on the African continent. *Vet Microbiol*. 2018 Sep;223(April):51–8.
38. Gather T, Walter S, Pfaender S, Todt D, Feige K, Steinmann E, et al. Acute and chronic infections with nonprimate hepacivirus in young horses. *Vet Res*. 2016 Dec 22;47(1):97.
39. Gemaque BS, Souza de Souza AJ, Soares M do CP, Malheiros AP, Silva AL, Alves MM, et al. Hepacivirus Infection in Domestic Horses, Brazil, 2011–2013. *Emerg Infect Dis*. 2014 Dec;20(12):2180–1.
40. Figueiredo AS, Lampe E, de Albuquerque PPLF, Chalhoub FLL, de Filippis AMB, Villar LM, et al. Epidemiological investigation and analysis of the NS5B gene and protein variability of non-

- primate hepacivirus in several horse cohorts in Rio de Janeiro state, Brazil. *Infect Genet Evol.* 2018 Apr;59(February):38–47.
41. HAYASHI S, TANAKA T, MORIISHI K, HIRAYAMA K, YAMADA A, HOTTA K. Seroepidemiology of non-primate hepacivirus (NPHV) in Japanese native horses. *J Vet Med Sci.* 2018;80(1):186–9.
 42. Ramsay JD, Evanoff R, Wilkinson TE, Divers TJ, Knowles DP, Mealey RH. Experimental transmission of equine hepacivirus in horses as a model for hepatitis C virus. *Hepatology.* 2015;61(5):1533–46.
 43. Elgharably A, Gomaa AI, Crossey MME, Norsworthy PJ, Waked I, Taylor-Robinson SD. Hepatitis C in Egypt ‐ past, present, and future. *Int J Gen Med.* 2016 Dec;Volume 10:1–6.
 44. El-Zanaty F, Way A. Egypt demographic and health survey 2008 [Internet]. 2008 [cited 2019 Aug 12]. Available from: <http://www.measuredhs.com/pubs/pdf/FR220/FR220.pdf>
 45. Abu-Raddad, L, DeWolfe M. Evidence of intense ongoing endemic transmission of hepatitis C virus in Egypt. *Qatar Found Annu Res Forum Proc* [Internet]. 2010 Dec 13 [cited 2019 Jul 26];2010(2010):BMP3. Available from: <http://www.qscience.com/doi/abs/10.5339/qfarf.2010.bmp3>
 46. World Health Organization. Guidelines for the care and treatment of persons diagnosed with chronic hepatitis C virus infection [Internet]. Geneva: Licence: CC BY-NC-SA 3.0 IGO.; 2018 [cited 2019 Aug 12]. Available from: <https://apps.who.int/iris/bitstream/handle/10665/273174/9789241550345-eng.pdf?ua=1>
 47. Gather T, Walter S, Todt D, Pfaender S, Brown RJP, Postel A, et al. Vertical transmission of hepatitis C virus-like non-primate hepacivirus in horses. *J Gen Virol.* 2016 Oct 13;97(10):2540–51.
 48. Badenhorst M, de Heus P, Auer A, Rümenapf T, Tegtmeyer B, Kolodziejek J, et al. No Evidence of Mosquito Involvement in the Transmission of Equine Hepacivirus (Flaviviridae) in an Epidemiological Survey of Austrian Horses. *Viruses.* 2019 Nov 1;11(11):1014.
 49. Corman VM, Grundhoff A, Baechlein C, Fischer N, Gmyl A, Wollny R, et al. Highly Divergent Hepaciviruses from African Cattle. *Ou J-HJ*, editor. *J Virol.* 2015 Jun 1;89(11):5876–82.
 50. Baechlein C, Fischer N, Grundhoff A, Alawi M, Indenbirken D, Postel A, et al. Identification of a Novel Hepacivirus in Domestic Cattle from Germany. 2015 [cited 2019 Oct 22]; Available from: <http://dx.doi.org/10.1128>
 51. Matsuu A, Hobo S, Ando K, Sanekata T, Sato F, Endo Y, et al. Genetic and serological surveillance for non-primate hepacivirus in horses in Japan. *Vet Microbiol.* 2015;179(3):219–27.
 52. Reichert C, Campe A, Walter S, Pfaender S, Welsch K, Ruddat I, et al. Frequent occurrence of nonprimate hepacivirus infections in Thoroughbred breeding horses – A cross-sectional study for the occurrence of infections and potential risk factors. *Vet Microbiol.* 2017 May;203:315–22.
 53. Pawlotsky J-M, Negro F, Aghemo A, Berenguer M, Dalgard O, Dusheiko G, et al. EASL Recommendations on Treatment of Hepatitis C 2018. *J Hepatol.* 2018 Aug;69(2):461–511.
 54. Reuter G, Maza N, Pankovics P, Boros Á. Non-primate hepacivirus infection with apparent hepatitis in a horse — Short communication. *Acta Vet Hung.* 2014 Sep 19;62(3):422–7.
 55. Tegtmeyer B, Echelmeyer J, Pfankuche VM, Puff C, Todt D, Fischer N, et al. Chronic equine hepacivirus infection in an adult gelding with severe hepatopathy. *Vet Med Sci.* 2019 Jul 3;vms3.181.
 56. Park SH, Reherrmann B. Immune responses to HCV and other hepatitis viruses. Vol. 40, *Immunity.* 2014. p. 13–24.
 57. Figueiredo AS, Lampe E, Espírito Santo MP, Mello FCA, Almeida FQ, Lemos ERS, et al. VV103 - Prevalence, clearance and nucleotide diversity of the equine hepacivirus (NPHV) in Brazil. In: Silva EE, Spilki FR, editors. *Virus Reviews and Research* [Internet]. Florianópolis, SC: Journal of the Brazilian Society for Virology; 2015 [cited 2019 Aug 8]. p. 221. Available from: <https://sbv.org.br/sbv/eventos-anteriores/#1527533417399-0-0>
 58. BUKH J, MILLER RH., PURCELL RH. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis.* 1995;15:41–63.
 59. Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, et al. Expanded classification

- of hepatitis C virus into 7 genotypes and 67 subtypes: Updated criteria and genotype assignment web resource. *Hepatology*. 2014 Jan;59(1):318–27.
60. Borgia SM, Hedskog C, Parhy B, Hyland RH, Stamm LM, Brainard DM, et al. Identification of a Novel Hepatitis C Virus Genotype From Punjab, India: Expanding Classification of Hepatitis C Virus Into 8 Genotypes. *J Infect Dis*. 2018 Oct 20;218(11):1722–9.
 61. Lu G, Ou J, Sun Y, Wu L, Xu H, Zhang G, et al. Natural recombination of equine hepacivirus subtype 1 within the NS5A and NS5B genes. *Virology*. 2019 Jul;533(May):93–8.
 62. Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, et al. Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology*. 2015 Jan 1;61(1):77–87.
 63. Simmonds P, Tuplin A, Evans DJ. Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implications for virus-evolution and host persistence. *RNA*. 2004 Sep;10(9):1337–51.
 64. Da Silva MS, Weber MN, Baumbach LF, Cibulski SP, Budaszewski RF, Mósena ACS, et al. Highly divergent cattle hepacivirus N in Southern Brazil. *Arch Virol*. 2019 Dec 28;164(12):3133–6.
 65. Lu G, Ou J, Zhao J, Li S. Presence of a Novel Subtype of Bovine Hepacivirus in China and Expanded Classification of Bovine Hepacivirus Strains Worldwide into 7 Subtypes. *Viruses*. 2019 Sep 11;11(9):843.
 66. Eigen M, McCaskill J, Schuster P. Molecular quasi-species. *J Phys Chem*. 1988 Dec;92(24):6881–91.
 67. Domingo E, Sheldon J, Perales C. *Viral Quasispecies Evolution*. 2012;76(2).
 68. Biebricher CK, Eigen M. The error threshold. *Virus Res*. 2005;107(2):117–27.
 69. Ruiz-Jarabo CM, Arias A, Baranowski E, Escarmis C, Domingo E. Memory in Viral Quasispecies. *J Virol*. 2000 Apr 15;74(8):3543–7.
 70. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog*. 2010;6(7):1–8.
 71. Steinhauer DA, Domingo E, Holland JJ. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*. 1992 Dec 15;122(2):281–8.
 72. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral Mutation Rates. *J Virol*. 2010 Oct 1;84(19):9733–48.
 73. Drake JW, Charlesworth B, Charlesworth D, Crow JF. *Rates of Spontaneous Mutation*. 1998.
 74. Domingo E, Holland JJ. RNA VIRUS MUTATIONS AND FITNESS FOR SURVIVAL. *Annu Rev Microbiol*. 1997 Oct;51(1):151–78.
 75. Weaver SC, Brault AC, Kang W, Holland JJ. Genetic and Fitness Changes Accompanying Adaptation of an Arbovirus to Vertebrate and Invertebrate Cells. Vol. 73, *JOURNAL OF VIROLOGY*. 1999.
 76. Ciota AT, Kramer LD. Insights into Arbovirus Evolution and Adaptation from Experimental Studies. *Viruses*. 2010 Dec 2;2(12):2594–617.
 77. Li H, Roossinck MJ. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. *J Virol*. 2004 Oct 1;78(19):10582–7.
 78. Crotty S, Cameron CE, Andino R. RNA virus error catastrophe: Direct molecular test by using ribavirin. *Proc Natl Acad Sci*. 2001 Jun 5;98(12):6895–900.
 79. Pfeiffer JK, Kirkegaard K. Increased Fidelity Reduces Poliovirus Fitness and Virulence under Selective Pressure in Mice. *PLoS Pathog*. 2005;1(2):e11.
 80. Crotty S, Maag D, Arnold JJ, Zhong W, Lau JYN, Hong Z, et al. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med*. 2000 Dec 1;6(12):1375–9.
 81. Woo HJ, Reifman J. A quantitative quasispecies theory-based model of virus escape mutation under immune selection. *Proc Natl Acad Sci U S A*. 2012;109(32):12980–5.
 82. Rong L, Dahari H, Ribeiro RM, Perelson AS. Rapid emergence of protease inhibitor resistance in hepatitis C virus. *Sci Transl Med*. 2010;2(30).

83. Aurora R, Donlin MJ, Cannon NA, Tavis JE. Genome-wide hepatitis C virus amino acid covariance networks can predict response to antiviral therapy in humans. *J Clin Invest.* 2009;119(1):225–36.
84. Morishima C, Polyak SJ, Ray R, Doherty MC, Di Bisceglie AM, Malet PF, et al. Hepatitis C Virus–Specific Immune Responses and Quasi-Species Variability at Baseline Are Associated with Nonresponse to Antiviral Therapy during Advanced Hepatitis C. *J Infect Dis.* 2006 Apr;193(7):931–40.
85. Malet I, Belnard M, Agut H, Cahour A. From RNA to quasispecies: A DNA polymerase with proofreading activity is highly recommended for accurate assessment of viral diversity. *J Virol Methods.* 2003;109(2):161–70.
86. Ramachandran S, Xia G liang, Ganova-Raeva LM, Nainan O V., Khudyakov Y. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: Performance under optimal and suboptimal conditions. *J Virol Methods.* 2008;151(2):217–24.
87. RODRIGO AG, GORACKE PC, ROWHANIAN K, MULLINS JI. Quantitation of Target Molecules from Polymerase Chain Reaction-Based Limiting Dilution Assays. *AIDS Res Hum Retroviruses.* 1997 Jun 10;13(9):737–42.
88. Scholz MB, Lo CC, Chain PSG. Next generation sequencing and bioinformatic bottlenecks: The current state of metagenomic data analysis. Vol. 23, *Current Opinion in Biotechnology.* 2012. p. 9–15.
89. Beerenwinkel N, Günthard HF, Roth V, Metzner KJ. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Vol. 3, *Frontiers in Microbiology.* Frontiers Research Foundation; 2012.
90. Longmire AG, Sims S, Rytsareva I, Campo DS, Skums P, Dimitrova Z, et al. GHOST: global hepatitis outbreak and surveillance technology. *BMC Genomics.* 2017 Dec 6;18(S10):916.
91. Ramírez C, Gregori J, Buti M, Tabernero D, Camós S, Casillas R, et al. A comparative study of ultra-deep pyrosequencing and cloning to quantitatively analyze the viral quasispecies using hepatitis B virus infection as a model. *Antiviral Res.* 2013 May 1;98(2):273–83.
92. Jordan MR, Kearney M, Palmer S, Shao W, Maldarelli F, Coakley EP, et al. Comparison of standard PCR/cloning to single genome sequencing for analysis of HIV-1 populations. *J Virol Methods.* 2010 Sep;168(1–2):114–20.
93. Perales C, Lorenzo-Redondo R, López-Galíndez C, Martínez MA, Domingo E. Mutant spectra in virus behavior. *Future Virol.* 2010 Nov;5(6):679–98.
94. Martell M, Esteban JI, Quer J, Genescà J, Weiner A, Esteban R, et al. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J Virol.* 1992 May;66(5):3225–9.
95. Palmer BA, Schmidt-Martin D, Dimitrova Z, Skums P, Crosbie O, Kenny-Walsh E, et al. Network Analysis of the Chronic Hepatitis C Virome Defines Hypervariable Region 1 Evolutionary Phenotypes in the Context of Humoral Immune Responses. *J Virol.* 2015;90(7):3318–29.
96. Weiner AJ, Brauer MJ, Rosenblatt J, Richman KH, Tung J, Crawford K, et al. Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology.* 1991;180(2):842–8.
97. Zibert A, Schreier E, Roggendorf M. Antibodies in human sera specific to hypervariable region 1 of hepatitis C virus can block viral attachment. *Virology.* 1995 Apr 20;208(2):653–61.
98. Curran R, Jameson CL, Craggs JK, Grabowska AM, Thomson BJ, Robins A, et al. Evolutionary trends of the first hypervariable region of the hepatitis C virus E2 protein in individuals with differing liver disease severity. Vol. 83, *Journal of General Virology.* 2002.
99. Kumar U, Monjardino J, Thomas HC. Hypervariable Region of Hepatitis C Virus Envelope Glycoprotein (E2/NSI) in an Agammaglobulinemic Patient. *Gastroenterology.* 1994;106(4):1072–5.
100. Forns X, Purcell RH, Bukh J. Quasispecies in viral persistence and pathogenesis of hepatitis C virus. *Trends Microbiol.* 1999 Oct 1;7(10):402–10.
101. Farci P. The Outcome of Acute Hepatitis C Predicted by the Evolution of the Viral Quasispecies. *Science (80-).* 2000 Apr 14;288(5464):339–44.

102. Sullivan DG, Bruden D, Deubner H, McArdle S, Chung M, Christensen C, et al. Hepatitis C Virus Dynamics during Natural Infection Are Associated with Long-Term Histological Outcome of Chronic Hepatitis C Disease. *J Infect Dis.* 2007;196(2):239–48.
103. Honda M, Kaneko S, Sakai A, Unoura M, Murakami S, Kobayashi K. Degree of diversity of hepatitis C virus quasispecies and progression of liver disease. *Hepatology.* 1994;20(5):1144–51.
104. Farci P, Quinti I, Farci S, Alter HJ, Strazzeri R, Palomba E, et al. Evolution of hepatitis C viral quasispecies and hepatic injury in perinatally infected children followed prospectively. *Proc Natl Acad Sci.* 2006;103(22):8475–80.
105. Fan X, Mao Q, Zhou D, Lu Y, Xing J, Xu Y, et al. High diversity of hepatitis C viral quasispecies is associated with early virological response in patients undergoing antiviral therapy. *Hepatology.* 2009;50(6):1765–72.
106. Ray SC, Wang Y-M, Laeyendecker O, Ticehurst JR, Villano SA, Thomas DL. Acute Hepatitis C Virus Structural Gene Sequences as Predictors of Persistent Viremia: Hypervariable Region 1 as a Decoy. Vol. 73, *JOURNAL OF VIROLOGY.* 1999.
107. Hayashi J, Kishihara Y, Yamaji K, Furusyo N, Yamamoto T, Pae Y, et al. Hepatitis C viral quasispecies and liver damage in patients with chronic hepatitis C virus infection. *Hepatology.* 1997 Mar;25(3):697–701.
108. Holmes JA, Chung RT. Viral hepatitis: HCV compartmentalization in HCC: Driver, passenger or both? Vol. 13, *Nature Reviews Gastroenterology and Hepatology.* Nature Publishing Group; 2016. p. 254–6.
109. Farci P, Strazzeri R, Alter HJ, Farci S, Degioannis D, Coiana A, et al. Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. *Proc Natl Acad Sci.* 2002 Mar 5;99(5):3081–6.
110. Schvoerer E, Soulier E, Royer C, Renaudin A-C, Thumann C, Fafi-Kremer S, et al. Early Evolution of Hepatitis C Virus (HCV) Quasispecies after Liver Transplant for HCV-Related Disease. *J Infect Dis.* 2007 Aug 15;196(4):528–36.
111. Feliu A, Gay E, García-Retortillo M, Saiz JC, Forn X. Evolution of hepatitis C virus quasispecies immediately following liver transplantation. *Liver Transplant.* 2004 Sep;10(9):1131–9.
112. Ramsay JD, Evanoff R, Mealey RH. Hepacivirus A Infection in Horses Defines Distinct Envelope Hypervariable Regions and Elucidates Potential Roles of Viral Strain and Adaptive Immune Status in Determining Envelope Diversity and Infection Outcome. Ou J-HJ, editor. *J Virol.* 2018 Jul 5;92(18).
113. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007 Nov 1;23(21):2947–8.
114. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 2016;33(7):1870–4.
115. Lyons S, Kapoor A, Sharp C, Schneider BS, Wolfe ND, Culshaw G, et al. Nonprimate Hepaciviruses in Domestic Horses, United Kingdom. *Emerg Infect Dis.* 2012 Dec;18(12):1976–82.
116. Pfaender S, Walter S, Grabski E, Todt D, Bruening J, Romero-Brey I, et al. Immune protection against reinfection with nonprimate hepacivirus. *Proc Natl Acad Sci.* 2017;114(12):E2430–9.
117. Simmonds P. SSE: a nucleotide and amino acid sequence analysis platform. *BMC Res Notes.* 2012 Dec 20;5(1):50.
118. Posada D. jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol.* 2008 Apr 3;25(7):1253–6.
119. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012 Aug;29(8):1969–73.
120. Bailey JR, Barnes E, Cox AL. Approaches, Progress, and Challenges to Hepatitis C Vaccine Development. *Gastroenterology.* 2019 Jan 1;156(2):418–30.
121. Liu S-L, Rodrigo AG, Shankarappa R, Learn GH, Hsu L, Davidov O, et al. HIV Quasispecies and Resampling. *Science (80-).* 1996 Jul 26;273(5274):413c – 417.
122. Rodrigo AG, Hanley EW, Goracke PC, Learn GH. Sampling and Processing HIV Molecular

- Sequences: A Computational Evolutionary Biologist's Perspective. In: Computational and Evolutionary Analysis of HIV Molecular Sequences. Boston, MA: Springer US; 2002. p. 1–17.
123. Ribeiro RM, Li H, Wang S, Stoddard MB, Learn GH, Korber BT, et al. Quantifying the Diversification of Hepatitis C Virus (HCV) during Primary Infection: Estimates of the In Vivo Mutation Rate. Wilke CO, editor. PLoS Pathog. 2012 Aug 23;8(8):e1002881.
 124. Schmidt-Martin D, Crosbie O, Kenny-Walsh E, Fanning LJ. Intensive temporal mapping of hepatitis C hypervariable region 1 quasispecies provides novel insights into hepatitis C virus evolution in chronic infection. J Gen Virol. 2015;96(8):2145–56.
 125. Naito M, Hayashi N, Moribe T, Hagiwara H, Mita E, Kanazawa Y, et al. Hepatitis C viral quasispecies in hepatitis C virus carriers with normal liver enzymes and patients with type C chronic liver disease. Hepatology. 1995;22(2):407–12.
 126. Forni D, Cagliani R, Pontremoli C, Pozzoli U, Vertemara J, De Gioia L, et al. Evolutionary Analysis Provides Insight Into the Origin and Adaptation of HCV. Front Microbiol [Internet]. 2018 May 1 [cited 2019 Aug 12];9:854. Available from: <http://journal.frontiersin.org/article/10.3389/fmicb.2018.00854/full>

9 APÊNDICES E/OU ANEXOS

Tabela S 1 Teste de diferença mínima significativa para diversidade intra-populacional

Animais	Média dos postos	Grupo
U114.3	2259	a
B25	2230	a
U114.2	1916	b
674	1275	c
348	1101	d
U64	1023	d
U74	916	e
U78	906	e
M42	612	f
U82	602	f

Médias seguidas de mesma letra não diferem estatisticamente entre si.

Tabela S 2 Teste de diferença mínima significativa entre localidades diferentes

Animais	Média dos postos	Grupo
Outras regiões	1480,01	A
UFRRJ	1461,96	A
Outras estados	1200,742	B

Médias seguidas de mesma letra não diferem estatisticamente entre si

Tabela S 3 Teste de diferença mínima estatística entre localidades parte 2

Animais	Média dos postos	Grupo
Outras regiões	1858,51	a
RJ	1838,73	a
Outras estados	1519,00	b

Médias seguidas de mesma letra não diferem estatisticamente entre si

Tabela S 4 Teste de diferença mínima estatística para diversidade genética inter-populacional

Animais	Média dos postos	Grupo
M42 v U114.2	24936,38	a
M42 v U114.3	23271,42	b
348 v M42	23138,88	b
B25 v U114.2	22531,95	c
U74 v M42	21948,35	d
U64 v M42	21797,36	d
B25 v U114.3	21261,83	e
U78 v M42	20333,10	f
674 v M42	20121,36	fg
U82 v M42	19969,07	g
U78 v B25	17375,39	h
U74 v B25	17213,43	h
U64 v B25	17179,42	h
674 v B25	16890,56	i
348 v B25	16297,69	j
U82 v B25	16274,00	j
B25 v M42	15430,23	k
U114.2 v U114.3	14690,73	l
348 v U114.2	12874,67	m
348 v U114.3	11655,20	n
674 v U114.2	10815,99	o
U78 v U114.2	10612,38	op
U64 v U114.2	10464,80	pq
U74 v U114.2	10385,70	pq
U82 v U114.2	10224,01	q
674 v U114.3	9880,66	r
U74 v U114.3	9558,31	s
U78 v U114.3	9508,04	st
U82 v U114.3	9315,72	st
U64 v U114.3	9244,45	t
U64 v U82	6875,72	u
348 v 674	6412,48	v
U82 v 348	6131,30	v
U78 v 348	5034,80	w
U74 v 348	4989,03	w
U64 v 348	4875,32	w
U74 v 674	3307,08	x
U82 v 674	3191,76	xy
U78 v 674	3151,45	xy
U64 v 674	2997,16	y
U74 v U78	2842,15	yz
U64 v U74	2685,65	z
U74 v U82	2653,05	z
U64 v U78	2521,28	z
U78 v U82	2495,23	z

Médias seguidas de mesma letra não diferem estatisticamente entre si.

Tabela S 5 Sequências majoritárias coletadas no GenBank para análise filogenética

País	Número de Acesso	Referência
	JF74491	(3)
	JQ434001-JQ434008	(10)
EUA	KJ472766	(42)
	KP325401-KP325402	(7)
	KR349928	(11)
China	MF152651- MF152652	(35)
Coréia	KX056116-KX056117	(33)
Escócia	JX948116	(115)
França	KT880191-KT880193	(15)
Japão	AB863589	(24)
	KF177391	(54)
	KP739814	(6)
Alemanha	KX064762-KX064784	(38)
	KX421286- KX421287	(15)
	KY124246- KY124248	(116)