

Ana Paula Moreira Salles

**Caracterização molecular do vírus da febre amarela
circulante em São Paulo no surto de 2018-2019**

Dissertação apresentada à Faculdade de
Medicina da Universidade de São Paulo
para obtenção do título de Mestre em
Ciências

Programa de Ciências em
Gastroenterologia

Orientador: Prof. Dr. João Renato Rebello
Pinho

Coorientadora: Dra. Fernanda de Mello
Malta

São Paulo

2021

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Salles, Ana Paula Moreira
Caracterização molecular do vírus da febre amarela
circulante em São Paulo no surto de 2018-2019 / Ana
Paula Moreira Salles. -- São Paulo, 2021.
Dissertação (mestrado)--Faculdade de Medicina da
Universidade de São Paulo.
Programa de Ciências em Gastroenterologia.
Orientador: João Renato Rebello Pinho.
Coorientadora: Fernanda de Mello Malta.

Descritores: 1.Vírus da febre amarela
2.Sequenciamento completo do genoma
3.Sequenciamento de nova geração 4.Variantes
5.Nucleotídeos 6.Aminoácidos

USP/FM/DBD-259/21

Responsável: Erinalva da Conceição Batista, CRB-8 6755

DEDICATÓRIA

Dedico este trabalho a minha família,
aos meus colegas de profissão e aos pacientes
que aceitaram fazer parte deste estudo.

AGRADECIMENTOS

À **Deus**, por ter iluminado meu caminho, me preparado para enfrentar novos desafios e ter me dado os meios e as pessoas que me ajudaram a enfrentar os obstáculos e a combater toda a ansiedade e pensamentos pessimistas que surgiram ao longo do caminho. Definitivamente, ser pesquisadora no Brasil, não é uma tarefa fácil.

Ao meu **orientador, Dr. João Renato R. Pinho**, pela grande oportunidade de realizar o mestrado em seu laboratório, pelas palavras de apoio e sua grande ajuda durante todo este trabalho. Obrigada por confiar em mim e sempre me incentivar a novos desafios, e por acreditar em meu potencial.

À minha **coorientadora, Dra. Fernanda de Mello Malta**, que foi parte essencial dessa oportunidade maravilhosa. Fê, quero agradecer por toda confiança que você colocou em mim, por todo o apoio que você me deu, por todos os trabalhos que fizemos juntas, por ter sempre me tratado como igual e acreditado em mim e no meu potencial, em momentos que eu mesma não acreditei. Você, sem dúvidas, é a maior responsável pelo meu crescimento como pesquisadora e para que a realização deste trabalho fosse possível. Obrigada por ter estado disponível, mesmo com dois surtos no meio do caminho... febre amarela e agora, coronavírus. Quero te dizer, não só obrigada, mas também o quanto te admiro e me espelho em você. Parabéns pela profissional e pessoa maravilhosa que você é, sempre saiba o quanto você contribui para termos um mundo melhor e o quanto você é acolhedora e verdadeira.

À **Dra. Michele S. Gomes-Gouvêa** por ter estado sempre disponível e disposta a me ajudar em muitos momentos, principalmente nesta etapa final. Obrigada Mi, por me ensinar e auxiliar a enfrentar este desafio. Obrigada por todo o incentivo para seguir para a próxima etapa e por me incentivar sempre a buscar o melhor para mim.

À **minha amiga, Patrícia O. Gionda**, que foi uma grande amiga nesses dois anos e que com certeza vou levar para o resto da minha vida. Obrigada por todos os tubos que você me ajudou a marcar, por edições de texto e por todas as conversas que tivemos, seja dentro ou fora do laboratório.

À **minha amiga de graduação, Samira Chuffi**, em primeiro lugar por ter pensado em mim e me indicado para uma conversa com o Dr João, com certeza te devo todos os agradecimentos por ter sido uma amiga tão verdadeira e que me ajudou tanto no meu crescimento profissional. Obrigada Sá por todo o apoio, conversas, dicas e trocas de conhecimento.

À **minha amiga, Msc. Gessica Melo** por todo o apoio, discussões de trabalhos e almoços durante esses 2 anos. Obrigada pela sua companhia.

Às **Dras. Ana Catharina dos S. S. NASTRI e Luciana Vila Boas Casadio** por todo o apoio durante o caminho, por serem médicas impecáveis, super humildes e que fazem você se sentir bem ao estar ao lado delas. Sorte de cada paciente atendido por vocês.

Ao **Dr. Deyvid E. Amgarten**, que me ajudou muito na construção deste trabalho. Sem você, meus resultados não teriam saído, pois como é difícil bioinformática em? Obrigada pelas aulas de programação, pelo incentivo, e por ter tirado todas as minhas dúvidas conforme elas foram surgindo ao longo do caminho.

À secretária da pós-graduação do departamento de Gastroenterologia Clínica, **Vilma Libério**, por todo o apoio e suporte durante todo o caminho. Por estar sempre de prontidão a tirar as dúvidas desde a realização da matrícula até a defesa.

À **minha mãe**, por ter sido a melhor mãe e amiga que alguém poderia ter, me dando todo o amor que alguém poderia dar. Por ter lutado sempre por mim e por mim. Por noites incansáveis de trabalho para que nunca nos faltasse nada e por ter me ensinado como ser uma pessoa honesta, que possui valores e que o trabalho duro não mata. Não posso me alongar muito aqui, mas eu te amo com todo o meu coração e obrigada. Um adendo para agradecer aos integrantes da nossa família também: **Lucival, meus irmãos de coração Maria e Enrico e aos 12 gatos e 5 cachorros.**

Ao meu namorado **Thiago**, por ter me apoiado e acreditado em mim em momentos que foram difíceis. Por ter sido meu companheiro fiel e por ter me ajudado quando eu precisava aprender algo novo ou apenas desabafar. Obrigada por ter estado ao meu lado e por ter me lembrado que eu sou capaz.

Ao **Rafael, Carolina, Patrícia, Daniel, Larissa e Ana Laíse**, por serem os melhores amigos que alguém poderia sonhar em ter e por terem lido minha dissertação e demonstrarem interesse, mesmo não sendo da área de vocês.

Aos meus irmãos, **Luis Guilherme e Maria Eduarda**, agradeço pelo simples fato de vocês existirem.

Ao meu **pai e aos meus avôs Rose e Dirceu**, meu eterno agradecimento por todo o privilégio que vocês puderam me proporcionar, por sempre darem o seu melhor para que eu tivesse o melhor e por todo o amor e carinho. Embora em alguns momentos o porquê e a importância da minha escolha do mestrado não tenham feito muito sentido para vocês. E por fim, **a toda a minha família: tias, tio, primos, primas, Cardoso, Lúcia, Pamela, Caroline S.** (melhores edições de imagens) **e muitos outros.** Obrigada por todo o apoio!!!!!!

*“However difficult life may seem,
there is always something you can do
and succeed at”*

Stephen Hawking

LISTA DE ABREVIATURAS E SIGLAS

| | |
|----------------|---|
| A | aminoácido alanina |
| ALT | alanina aminotransferase |
| AST | aspartato aminotransferase |
| BAM | formato de arquivo de nucleotídeos alinhados (do inglês, <i>Binary Alignment Map</i>) |
| BLASTn | ferramenta de alinhamento de sequências (do inglês, <i>Basic Local Alignment Search Tool</i>) |
| C | proteína do capsídeo |
| Ct | ciclo de amplificação em relação ao threshold no PCR em tempo real (do inglês, <i>cycle threshold</i>) |
| D | ácido aspártico |
| DC | Células dendríticas (do inglês, <i>Dendritic Cells</i>) |
| dNTP | desoxirribonucleotídeos fosfatados |
| E | proteína do envelope |
| EYE | “Eliminação da Epidemia de Febre Amarela” (do inglês, <i>Eliminate Yellow Fever Epidemics Strategy</i>) |
| Elisa | ensaio sorológico para identificação de anticorpos (do inglês, <i>Enzyme-Linked Immunosorbent Assay</i>) |
| F ou L | <i>primer forward</i> |
| FA | Febre amarela (doença) |
| G | aminoácido glicina |
| GATK | Software chamador de variantes (do inglês <i>Genome Analysis Toolkit</i>) |
| H | aminoácido histidina |
| HCFMUSP | Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. |
| I | aminoácido isoleucina |
| IFN I | Interferon I |
| Indel | Inserção e/ou deleção de nucleotídeo(s) |
| K | aminoácido lisina |

| | |
|------------------|--|
| L | aminoácido leucina |
| M | aminoácido metionina |
| MAFFT | Software para alinhamento de sequências (do inglês, <i>Multiple Alignment using Fast Fourier transform</i>) |
| Max | máximo |
| Min | mínimo |
| ML | Máxima Verossimilhança (do inglês, <i>maximum likelihood</i>) |
| mRNA | RNA mensageiro |
| N | aminoácido asparagina |
| n | número de amostras |
| nM | nanomolar |
| NT | Nucleotídeo |
| NTPase | nucleotídeo trifosfatase |
| OMS | Organização Mundial da Saúde |
| OPAS/PAHO | Organização Pan-americana de Saúde/ <i>Pan American Health Organization</i> |
| P | aminoácido prolina |
| pb | pares de bases |
| PCR | Reação em cadeia da polimerase (do inglês, <i>Polymerase Chain Reaction</i>) |
| PNH | primata não humano |
| prM/M | proteína pré membrana/membrana |
| Qmap | qualidade de mapeamento |
| Qp | <i>phred score</i> |
| qPCR | PCR em tempo real |
| R | aminoácido arginina |
| r | <i>primer reverse</i> |

| | |
|----------------|--|
| RDC | República Democrática do Congo |
| RdRp | RNA polimerase RNA dependente |
| RE | Retículo Endoplasmático |
| RNA | Ácido Ribonucleico (do inglês, <i>Ribonucleic Acid</i>) |
| RT-qPCR | PCR em tempo real com uma etapa de transcrição reversa |
| RT | Transcrição reversa (do inglês, <i>reverse transcription</i>) |
| Sd | desvio padrão (do inglês, <i>standard deviation</i>) |
| SNPs | Polimorfismo de Nucleotídeo Único (do inglês, <i>Single Nucleotide Polymorphism</i>) |
| ssRNA | RNA de fita simples |
| SUS | Sistema Único de Saúde do Brasil |
| T | aminoácido treonina |
| TCLE | Termo de consentimento livre e esclarecido |
| UNICEF | Fundo das Nações Unidas para a Infância |
| UTI | unidade de terapia intensiva |
| UTR | região não codificante (do inglês, <i>untranslated region</i>) |
| V | aminoácido valina |
| VCF | formato de arquivo de dados após chamada de variantes (do inglês, <i>Variant Call Format</i>) |
| W | aminoácido triptofano |
| YFV | vírus da Febre Amarela (do inglês, <i>yellow fever virus</i>) |

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1: Ciclos de transmissão do YFV nas Américas e na África.----- | 20 |
| Figura 2: Representação esquemática do genoma dos flavivírus.----- | 21 |
| Figura 3: Esquema da replicação do Vírus da Febre Amarela.----- | 26 |
| Figura 4: Reemergência YFV no Brasil.----- | 30 |
| Figura 5: Distribuição dos casos autóctones e óbitos de YFV no Estado de São Paulo durante o ano de 2018, segundo semana epidemiológica.----- | 32 |
| Figura 6: Distribuição dos casos autóctones de YFV no Estado de São Paulo durante o surto ocorrido nos meses de janeiro a junho de 2019.----- | 32 |
| Figura 7: Fluxo de seleção das amostras em estudo.----- | 45 |
| Figura 8: Mapeamento dos prováveis locais de infecção.----- | 47 |
| Figura 9: Foto dos produtos de amplificação por PCR convencional.----- | 48 |
| Figura 10: Representação gráfica gerada durante a análise de qualidade das sequências, representando o escore de qualidade das leituras de sequenciamento.----- | 49 |
| Figura 11: Árvore filogenética realizada por máxima verossimilhança.----- | 57 |
| Figura 12: Árvore filogenética realizada por máxima verossimilhança referente às sequências de 2018 e 2019 obtidas neste estudo.----- | 59 |

LISTA DE TABELAS

- Tabela 1: Proteínas do vírus da febre amarela e suas possíveis funções.-----24
- Tabela 2: Descrição dos *primers* utilizados e suas posições de anelamento.-----37
- Tabela 3: Descrição das sequências utilizadas para gerar as árvores filogenéticas.----
-----43 e 44
- Tabela 4: Dados demográficos dos pacientes participantes do estudo.-----46
- Tabela 5: Dados das sequências obtidas descrevendo os Cts das amostras usadas durante a amplificação da PCR, o tamanho (pb) e a média de cobertura (x) referente aos genomas completos ou quase completos.-----50 e 51
- Tabela 6: Quantidade de variantes por amostra, após análises realizadas com os softwares *Freebayes* e *GATK*.-----52
- Tabela 7: Variações únicas de nucleotídeos detectados pela metodologia de NGS, Variantes não sinônimas estão indicadas com a troca de aminoácido correspondente. Frequências das variantes nos surtos de 2018-2019 estão listadas.-----53 a 55

SUMÁRIO

Lista de Abreviaturas e Siglas

Lista de Figuras

Lista de Tabelas

Resumo

Abstract

| | |
|---|----|
| 1. Introdução | 18 |
| 1.1. Febre Amarela: Breve histórico | 18 |
| 1.2. Transmissão | 19 |
| 1.3. Agente etiológico: o vírus da Febre Amarela | 21 |
| 1.3.1. Agente etiológico: Proteínas estruturais | 22 |
| 1.3.2. Agente etiológico: Proteínas não estruturais | 22 |
| 1.4. Replicação viral | 24 |
| 1.5. Diversidade genética viral | 26 |
| 1.6. Diagnóstico | 28 |
| 1.7. Epidemiologia no mundo | 28 |
| 1.8. Epidemiologia no Brasil | 29 |
| 2. Objetivos | 34 |
| 3. Materiais & Métodos | 35 |
| 3.1. Casuística | 35 |
| 3.1.1. Critérios de inclusão | 35 |
| 3.1.2. Critérios de exclusão | 35 |
| 3.1.3. Dados demográficos | 35 |
| 3.2. Extração do RNA viral | 36 |
| 3.3. qPCR quantitativo - Carga viral | 36 |
| 3.4. Amplificação do genoma completo do YFV | 36 |
| 3.5. Detecção do DNA amplificado | 38 |
| 3.6. Sequenciamento de nova geração (NGS) | 38 |
| 3.7. Análise das sequências dos genomas completos do YFV | 39 |
| 3.7.1. Controle de qualidade dos dados brutos de sequenciamento NGS, trimagem e filtros | 39 |
| 3.7.2. Montagem do genoma completo YFV | 39 |
| 3.7.2.1. Montagem contra sequência referência - <i>Spades</i> | 40 |
| 3.7.2.2. Montagem de novo - <i>Spades</i> | 40 |
| 3.7.2.3. Montagem de novo - <i>IVA assembler</i> | 40 |

| | |
|--|----|
| 3.7.2.4. Análise manual | 41 |
| 3.7.3. Chamada de variantes | 41 |
| 3.7.4. Análise de intersecção e união de variantes nos surtos de 2018 e 2019 | 42 |
| 3.7.5. Análise Filogenética | 42 |
| 4. Resultados | 45 |
| 4.1. Caracterização da amostra de estudo | 45 |
| 4.2. Amplificação do genoma completo | 48 |
| 4.3. Sequenciamento e montagem do genoma | 48 |
| 4.4. Análise de variantes | 52 |
| 4.5. Análise Filogenética | 56 |
| 5. Discussão | 60 |
| 6. Conclusões | 66 |
| 7. ANEXOS | 67 |
| 8. REFERÊNCIAS BIBLIOGRÁFICAS | 75 |

RESUMO

Salles APM. *Caracterização molecular do vírus da febre amarela circulante em São Paulo durante o surto de 2018-2019* [dissertação]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2021.

A febre amarela é uma febre hemorrágica, viral e endêmica de regiões tropicais na América Central e do Sul e no continente Africano, afetando principalmente humanos e primatas não-humanos, causada pelo vírus da Febre Amarela (YFV). O recente surto de Febre amarela no Brasil foi um dos mais graves nos tempos atuais disseminando-se para áreas com baixa cobertura vacinal. Durante os picos dos surtos, foi possível observar uma diferença na evolução clínica dos pacientes que trouxe o questionamento se o vírus poderia ter apresentado mutações que levassem a uma alteração no quadro clínico. O presente estudo avaliou a diversidade genética do vírus da Febre Amarela dos casos internados no Hospital das Clínicas da Faculdade de Medicina de São Paulo, que sobreviveram à infecção durante o surto ocorrido em São Paulo nos anos de 2018 e 2019 e determinou a carga viral bem como o genótipo da cepa viral circulante. Para a realização deste estudo foram utilizadas amostras de soro e/ou urina positivas para o vírus da Febre Amarela, que foram submetidas à extração do RNA viral e em seguida à amplificação de cDNA. O sequenciamento de nova geração foi realizado e posteriormente feita a análise de dados utilizando ferramentas de bioinformática próprias para gerar as informações sobre a qualidade e cobertura das sequências a fim de realizar a montagem do genoma viral, análises das variantes e análises filogenéticas. Utilizando essa estratégia, foram obtidos 39 genomas quase completos, sendo possível a análise da diversidade genética viral entre os grupos de 2018 e 2019. Foram encontradas 46 substituições nucleotídicas (nt), sendo 20 delas substituições não sinônimas, com destaque para as posições 195 nt (presente em 60% das amostras de 2019) e 218 nt no capsídeo (presente em 28,6% das amostras de 2018), 5055 nt na NS3 presente em 100% das amostras deste estudo e 10.104 nt na NS5 (42,85% de 2018 e 96% de 2019), proteína polimerase do vírus. A única substituição nucleotídica encontrada neste estudo previamente publicada foi na posição 5055 nt da proteína NS3. Analisando as árvores filogenéticas, os genomas foram classificados como genótipo Sul Americano I subgrupo E, sendo este resultado comparado com estudos previamente publicados e de acordo com o descrito. É possível observar que 85,7% da casuística deste estudo é composta por pacientes do sexo masculino. Mapeamentos geográficos do provável local de infecção dos pacientes incluídos no estudo, demonstraram que o surto de 2018 teve uma maior concentração na região da Grande São Paulo, chegando ao litoral paulista, e o surto de 2019 concentrou-se, em sua maior parte, no Vale do Ribeira. Este estudo sugere que os vírus circulantes nos surtos de 2018 e 2019 são distintos, porém para um melhor entendimento do impacto das variantes sobre a patogênese viral, é necessária uma análise *in silico*. Em geral, as análises aqui realizadas permitem um melhor entendimento do genoma do YFV circulante em São Paulo e contribui para futuros estudos virológicos e epidemiológicos.

Descritores: Vírus da febre amarela; Sequenciamento completo do genoma; Sequenciamento de nova geração; Variantes; Nucleotídeos; Aminoácidos

ABSTRACT

Salles APM. *Molecular characterization of yellow fever virus from 2018-2019 outbreak in São Paulo* [dissertation]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2021.

Yellow fever is a viral and endemic hemorrhagic fever from tropical regions in the Central and South Americas and Africa, affecting mainly humans and nonhuman primates, caused by the Yellow Fever virus (YFV). The recent outbreak of yellow fever in Brazil has been one of the most severe in the last decades, spreading to areas with low vaccine coverage. During the peak of the 2018-2019 outbreak, it was possible to observe a difference in the clinical evolution of the patients and it raised questions of whether the virus could have had mutations that led to milder clinical symptoms. The present study aimed to evaluate the genetic diversity of Yellow fever virus of Hospital das Clinicas da Faculdade de Medicina de Sao Paulo, hospitalized cases that survived the infection during the recent outbreak, in order to determine the viral load and genotype of the circulating viral strain in Sao Paulo. YFV-positive serum and / or urine samples were subjected to viral RNA extraction and then cDNA amplification. New generation sequencing and data analyses were carried out using bioinformatics tools for quality and coverage analysis for genome assembly, variant and phylogenetic analyses. Almost thirty-nine nearly complete YFV genomes were obtained and genetic diversity analysis comparing the circulating virus strains during 2018 and 2019, allowed us to find 46 nucleotides (nt) substitutions, and 21 of them were non-synonymous substitutions, highlighting the positions 195nt (60% of samples from 2019) and 218nt (28.6% of samples from 2018) in capsid protein, 5055nt (100% of samples) in NS3 protein and 10.104 in NS5 protein (42.85% of samples from 2018 and 96% samples from 2019). Phylogenetic analysis demonstrated that all YFV strains circulating during the outbreak form a unique major clade with the South American IE genotype. This finding was compared with previously published studies and is in accordance with what has been described so far. In conclusion, it was possible to describe that 85.7% of all patients enrolled in this study are male and the peak from 2018 YFV outbreak was concentrated in Greater Sao Paulo, going to the coast but the peak from 2019 outbreak was more concentrated in Vale do Ribeira. Our findings suggest that 2018 and 2019 circulating viruses were distinct from each other however for a better understanding of how these variants affect the viral pathogenesis it is necessary to perform an in silico analysis. Overall, our findings provide an overview of YFV genome characterization in Sao Paulo and contribute for future virological and epidemiological studies.

Descriptors: Yellow fever virus; Whole genome sequencing; Next generation sequencing; Variants; Nucleotides; Amino acids.

1. Introdução

A febre amarela (FA) é uma doença infecciosa causada por um vírus do gênero *Flavivirus*, pertencente à família *Flaviviridae* e transmitido pela picada de mosquitos fêmeas infectadas. O reservatório natural desse vírus são os primatas não humanos (PNH) que habitam florestas e matas tropicais (1). O Ministério da Saúde alerta quanto à importância de intensificar as ações de vigilância da FA no país durante o período sazonal da doença que ocorre entre os meses de dezembro a maio (2).

Nos seres humanos, a doença está associada a um amplo espectro de manifestações clínicas, tais como: assintomática (40-65% dos casos); leve a moderada (20-30%), em que a suspeita ocorre por meio de exames laboratoriais específicos como dosagens das enzimas hepáticas (ALT, AST); grave (10-20%) com a presença de febre e possivelmente icterícia. Esta última inclui as formas fulminantes (5-10%), com a presença dos sintomas clássicos (icterícia, dor abdominal, hepatite e hemorragia) e frequente evolução para óbito. O período de incubação médio é de 3 a 6 dias (3, 4).

Estudos genéticos demonstraram que o vírus da febre amarela (YFV) teve sua origem na África e possivelmente durante o século XVII tanto o vírus quanto o mosquito vetor de áreas urbanas (*Aedes aegypti*), foram trazidos para as Américas e Ásia durante o transporte dos escravos pelos navios negreiros e de especiarias (5, 6). Nessa época, as cidades não dispunham de saneamento básico e estavam infestadas de mosquitos. O resultado da migração transatlântica trouxe trágicas consequências para a saúde da população (5).

1.1. Febre Amarela: Breve histórico

Em 1685, foi reportada a primeira epidemia de FA no Brasil, em Recife, seguida por outras nos anos seguintes em diferentes estados do Nordeste onde foram instituídas medidas sanitárias de prevenção, como a higienização das casas e segregação dos doentes. Devido a essas ações, não houve relatos da doença durante um século (7).

Somente em 1823, surgiram novos relatos da doença quando um navio Francês atracou no porto da Bahia, mas só em 1849 houve de fato uma nova epidemia

atingindo os moradores do estado da Bahia (7). A partir de então, a FA alastrou-se para outras cidades portuárias como a do Rio de Janeiro.

Em 1903, Oswaldo Cruz, que estava acompanhando as pesquisas do médico cubano Carlos Finlay (responsável pela descoberta da forma de infecção da FA) (8), instituiu diversas medidas sanitárias, inclusive o combate ao vetor (9). Já em 1923, a Fundação Rockefeller, criada pelo exército Americano, iniciou sua atuação no Brasil e em 1932, por meio de estudos epidemiológicos no Vale do Canaã, no Estado do Espírito Santo, foi descoberto o ciclo silvestre da doença (10) o que tornou possível o mapeamento da provável rota de disseminação da onda de epizootias e epidemias da FA (7, 10) no Brasil.

Concomitantemente aos trabalhos da Fundação Rockefeller em solo brasileiro, uma equipe localizada na Nigéria, em 1927, conseguiu isolar o agente etiológico da FA. Os médicos da Fundação inocularam no macaco *rhesus* o sangue de um paciente de nome Asibi que havia apresentado a forma moderada da doença e em seguida por meio de novos experimentos confirmaram que o agente era um pequeno vírus filtrável (11).

Em 1936, após diversas passagens em macacos *rhesus*, com períodos de passagens intermitentes em *Aedes aegypti*, Max Theiler (ganhador do prêmio Nobel de Medicina em 1951 (11, 12) e Hugh H. Smith conseguiram atenuar o vírus e o propagaram em tecido embrionário de camundongo e cultura de embrião de galinha, onde sofreu novas passagens até que finalmente em 1937, a vacina eficaz contra YFV com o vírus atenuado (cepa 17D), foi desenvolvida (12).

Tão logo a vacina foi desenvolvida, ela foi trazida para o Brasil e no mesmo ano o Instituto Oswaldo Cruz desenvolveu a vacina 17DD, uma variação da 17D (7) e começou a produzi-la em larga escala (13) e por isso em 1942 foram registrados os últimos 3 casos de FA urbana no Estado do Acre (14). No entanto, no continente Africano, mesmo com os programas de vacinação instituídos, a transmissão da FA continua ocorrendo nos ciclos urbano e selvagem (15).

1.2. Transmissão

O modo de transmissão do YFV (Figura 1) só foi bem estabelecido em 1900, durante investigações que ocorreram principalmente em Cuba pelo cientista Carlos Finlay e depois que um grupo de militares americanos, do qual Walter Reed (6) médico

cirurgião fazia parte, se inteirou sobre as descobertas de Finlay, resolve testá-la e acaba corroborando-a (16).

Atualmente sabemos que na África ocorrem 3 tipos de ciclos de transmissão, o selvagem em áreas de densa vegetação entre o mosquito vetor e os PNHs; o intermediário em áreas peri-urbanas infectando tanto humanos quanto PNHs; e o urbano em áreas urbanizadas e diretamente entre o vetor de vida urbana e os humanos (17). Nas Américas, observou-se apenas os dois principais ciclos: selvagem, envolvendo a transmissão entre mosquitos vetores (*Haemagogus sp.* e *Sabethes sp.*) e PNH, podendo ocasionalmente afetar os humanos que adentram esse território; e o ciclo urbano que não ocorre no Brasil desde 1942, envolvendo mosquitos *Aedes aegypti* e/ou *Aedes albopictus* e o ser humano (18).

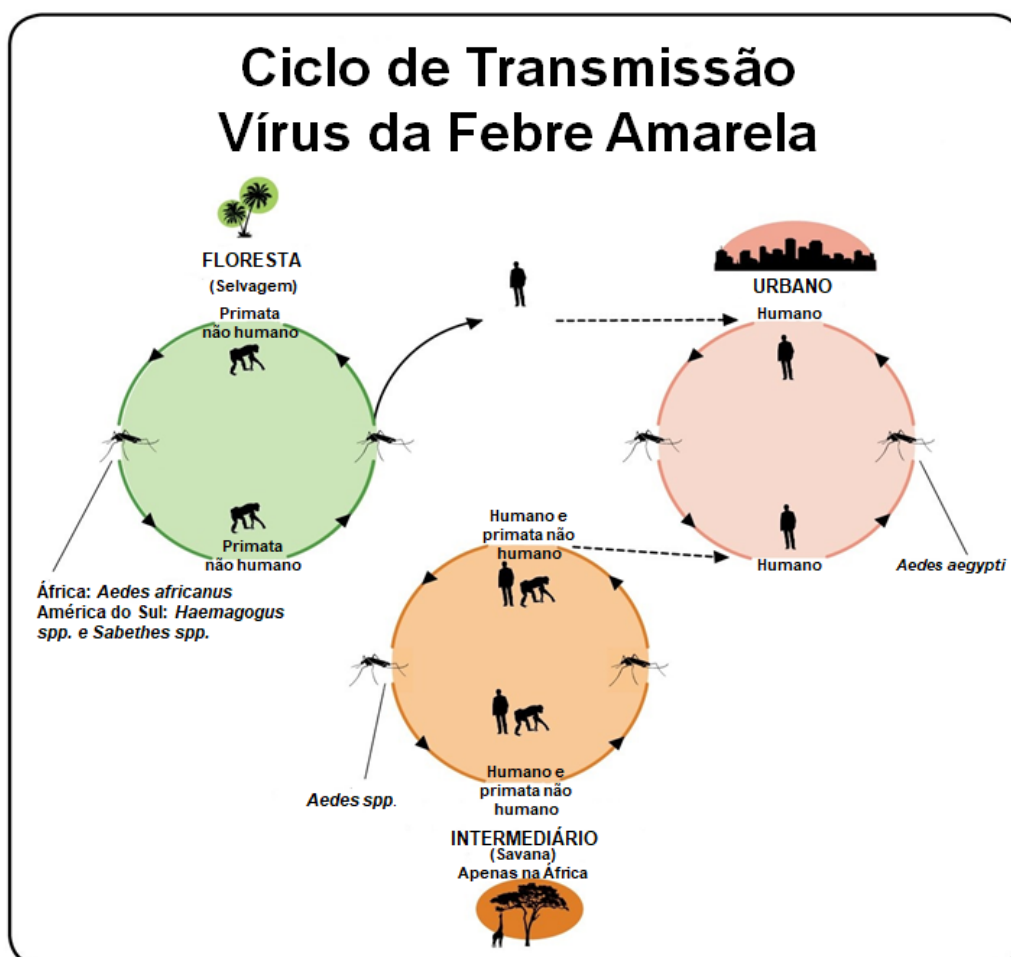


Figura 1: Ciclos de transmissão do YFV nas Américas e na África.

Representando os ciclos urbano e selvagem para as Américas e também, o ciclo intermediário exclusivo do continente Africano.

Fonte: Adaptado de *World Federation of Science Journalists (WFSJ)* (19)

1.3. Agente etiológico: o vírus da Febre Amarela

O YFV foi o primeiro vírus humano a ser isolado (20). Ele pertence ao gênero *Flavivirus* (21) e à família *Flaviviridae*. Nessa família existe um grande número de patógenos responsáveis por doenças humanas e de animais, dentre as quais temos o vírus da Dengue, Zika, Encefalite Japonesa e vírus do Oeste do Nilo (22).

O YFV é um vírus envelopado, com genoma RNA de fita simples e polaridade positiva de aproximadamente 11.000 nucleotídeos (Figura 2). A partícula viral é pequena, icosaédrica e envelopada. O nucleocapsídeo é formado pela proteína C que envolve o genoma de RNA fita simples de polaridade positiva (23).

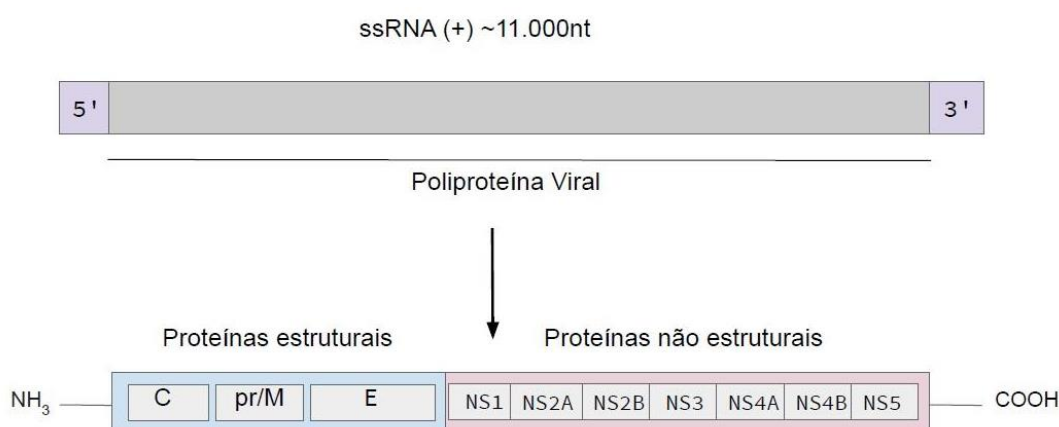


Figura 2: Representação esquemática do genoma dos flavivírus.

Fonte: Adaptado de Simmonds, P. et al., ICTV virus Report Consortium. 2017 (24)

Possui uma única região aberta de leitura (do inglês, *open reading frame* - ORF) que codifica uma poliproteína de aproximadamente 3.400 aminoácidos, que é clivada por proteases celulares e virais em proteínas estruturais (Capsídeo, C; Proteína de membrana, prM/M; Envelope, E) e proteínas não estruturais (NS1, NS2A-2B, NS3, NS4A-B, e NS5) (25).

A poliproteína é flanqueada por regiões não traduzidas (do inglês, *untranslated region* - UTR) conservadas e localizadas nas extremidades 5' UTR e 3' UTR do genoma (26), importantes no processo de regulação e replicação viral (27). A interação entre essas duas regiões é importante para o recrutamento e posicionamento da NS5 (RNA polimerase RNA dependente - RpRd) durante o processo de replicação (28).

A extremidade 5' é formada por estruturas complexas em forma de grampo (do inglês, *stem-loops*) e por uma estrutura denominada *cap* responsável pela

estabilidade e pela tradução da molécula de RNA (29). A extremidade 3' possui uma conformação mais complexa e devido a isso, seus nucleotídeos formam uma molécula muito estável, altamente conservada em forma de grampos ou de halteres (do inglês, *stem loops ou dumbbell*), entre essas estruturas observam-se sequências conservadas e sequências repetitivas, que servem como estabilizadores do genoma e promovem o sinal para que a tradução e a síntese de RNA sejam iniciadas. Entretanto, a extremidade 3' UTR não possui uma cauda poliadenilada comumente encontrada em outros flavivírus (23, 28, 30).

1.3.1. Agente etiológico: Proteínas estruturais

As proteínas estruturais (C, prM/M, E) são responsáveis por formar a estrutura básica da partícula viral e as suas possíveis funções estão resumidas na (Tabela 1).

A proteína C é altamente básica e faz parte do nucleocapsídeo viral (31). O contato do sistema imune com a proteína, desencadeia a indução de anticorpos neutralizantes pelo sistema imunológico (32). A proteína C será considerada madura quando o domínio proteico hidrofóbico, chamado de âncora do capsídeo, for clivado pela serino protease viral (NS3) (33).

A proteína precursora da proteína M é a prM que será translocada do retículo endoplasmático (RE) pelo terminal hidrofóbico C e, durante a saída do vírion pela via secretora, será clivada em M pela ação de uma enzima presente no organismo do hospedeiro, denominada furina ou furina-*like* (25). A proteína M será encontrada no vírion maduro e o segmento "pr" secretado no meio extracelular.

A proteína E é a principal proteína na superfície dos vírions e foi descrita como o principal alvo de anticorpos neutralizantes, por ser uma região de epítomos (31), e acredita-se que mutações nessa região estejam associadas à patogenicidade viral (25).

1.3.2. Agente etiológico: Proteínas não estruturais

Dentre as proteínas não estruturais temos as altamente conservadas NS1, NS3, NS5 e proteínas menores e hidrofóbicas: NS2A, NS2B, NS4A e NS4B (Tabela 1).

A NS1 é uma glicoproteína que possivelmente está associada com a virulência e à patogênese viral, importante para a inibição da cascata do sistema complemento no hospedeiro, participa da replicação do RNA viral, e os anticorpos específicos contra a NS1 podem desencadear imunidade protetora (23).

As proteínas NS2A, NS2B, NS4A, NS4B em geral, participam da montagem viral funcionando como cofatores para as proteínas NS1 e NS3 (23, 32).

A NS3 é uma proteína que possui atividade enzimática de serino protease na sua porção amino terminal (N) formando um complexo com a proteína NS2B que atuará como seu cofator durante o processamento da poliproteína viral. Na porção carboxi terminal (C), a NS3 atua como NTPase/helicase durante a replicação viral (23, 32, 33).

Entre as proteínas não estruturais, a NS5 também possui atividade enzimática, sendo a maior e mais conservada proteína entre os flavivírus. Ela contém duas atividades enzimáticas distintas: Metiltransferase (MTase) localizada na porção N terminal (34) que provavelmente está envolvida com a metilação da estrutura 5' *cap* RNA e na porção C terminal é uma RNA polimerase RNA dependente (RdRP - sigla em inglês) responsável pela replicação viral (35). Essa proteína foi associada com o antagonismo à expressão do Interferon I (IFN I) no hospedeiro, podendo ser um dos mecanismos de escape viral (36).

Tabela 1: Proteínas do vírus da febre amarela e suas possíveis funções.

| Proteína | Descrição da proteína | Possíveis funções |
|----------|-----------------------|---|
| C | Estrutural | formação do nucleocapsídeo; ligação ao RNA |
| prM/M | Estrutural | prM - no vírus imaturo; M - no vírus maduro; evita rearranjos na proteína E durante a fusão do vírus |
| E | Estrutural | medeia a ligação ao receptor celular; fusão do vírus com a membrana da célula hospedeira; epítopo viral |
| NS1 | Não Estrutural | participa da replicação viral; inibe a ativação da cascata complemento do hospedeiro |
| NS2A | Não Estrutural | montagem e liberação viral; pode estar associada com a maturação da proteína NS1 |
| NS2B | Não Estrutural | forma um complexo com a proteína NS3 |
| NS3 | Não Estrutural | serino protease; NTPase/helicase |
| NS4A | Não Estrutural | montagem e liberação viral; interage com a NS1 - forma um complexo de replicação |
| NS4B | Não Estrutural | pode estar associada com a NS3 nos sítios de replicação |
| NS5 | Não Estrutural | Metiltransferase; RNA polimerase; associada com o escape viral |

C: capsídeo; prM: pré membrana; M: membrana; E: envelope; NS: do inglês *non structural*

adaptado de: Chambers et al, 1990; Vasconcelos PFC, 2003 (3, 32).

1.4. Replicação viral

O processo de replicação dos *Flavivirus* envolve mecanismos complexos de interações entre proteínas virais e celulares, no caso do YFV, alguns desses mecanismos ainda não estão bem estabelecidos (25).

A replicação viral ocorre dentro do citoplasma das células susceptíveis ao YFV logo após a entrada do vírus (Figura 3). A proteína E do YFV liga-se de maneira inespecífica aos receptores de glicosaminoglicanos altamente sulfatados na superfície celular dos hepatócitos ou de células dendríticas (DCs - do inglês, *dendritic cells*) (37). A partir dessa interação, a entrada da partícula viral na célula ocorre via endocitose mediada por clatrina. Um rearranjo conformacional ocorre no endossomo com pH

baixo, facilitando a fusão. A eficiência dessa ligação depende da composição do lipídio da membrana alvo (o colesterol, por exemplo, facilita a fusão) (23, 38).

A liberação do nucleocapsídeo no citoplasma da célula hospedeira ocorre assim que a fusão das membranas é realizada (envelope viral e membrana das vesículas endossomais). Em seguida, o genoma viral é liberado e pode ser diretamente traduzido, uma vez que é uma molécula de RNA de fita simples (ssRNA) com polaridade positiva que funciona como RNA mensageiro (mRNA) (25).

A síntese do RNA viral inicia-se a partir de uma fita de RNA complementar com polaridade negativa que serve de molde para gerar a fita de RNA positiva (22, 39). A nova fita de RNA positiva poderá ser usada em um novo ciclo de tradução ou ser empacotada por proteínas estruturais gerando uma nova partícula viral que será liberada na superfície celular.

A montagem das partículas virais é pouco entendida. Acredita-se que durante esse processo, o nucleocapsídeo do YFV seja envolvido por membranas do RE cobrindo a glicoproteína E e a PrM. As partículas virais imaturas são direcionadas para o Complexo de Golgi pela via secretora, onde se tornam maduras ao passarem por um processo de glicosilação, o qual ocorre a clivagem da PrM em M pela protease furina (25, 40). O próximo passo será um rearranjo entre a proteína E e a subunidade M formando assim a partícula viral madura que será transportada para a membrana plasmática por meio de vesículas que se fundem com a membrana celular, liberando-as por meio da exocitose (25, 40, 41).

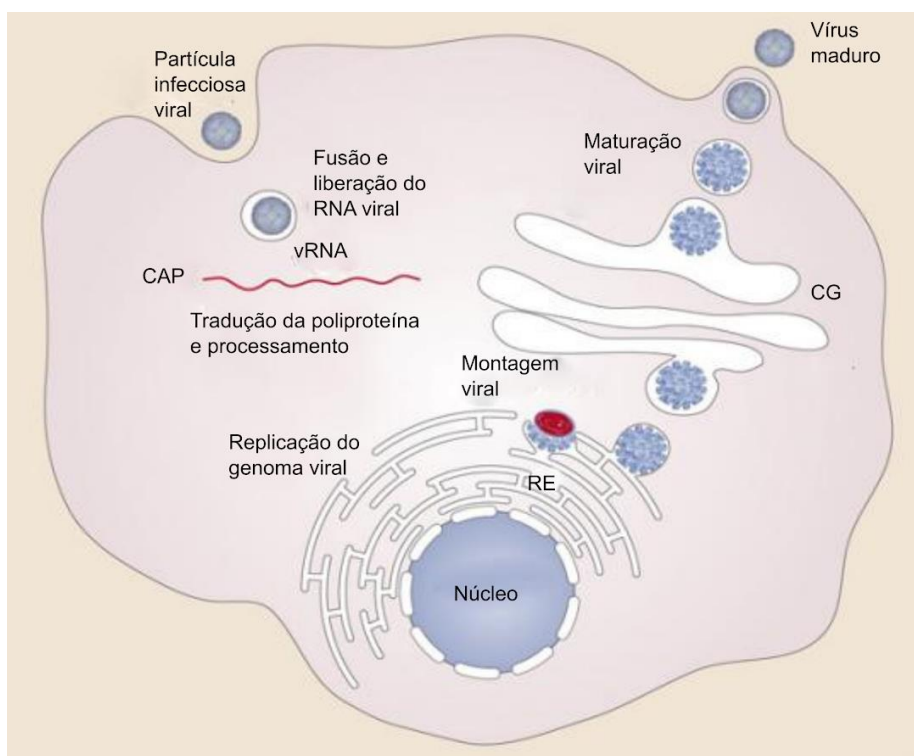


Figura 3: Esquema da replicação do Vírus da Febre Amarela.

O YFV liga-se à superfície da célula hospedeira e entra na célula via endocitose. Dentro da célula, o vírus funde-se com a membrana endossomal e libera seu material genético para o citoplasma. O RNA viral (vRNA) é traduzido numa única poliproteína que será traduzida em dez proteínas diferentes, com auxílio da enzima *CAP* (do inglês *capping enzyme*) responsável pela estabilidade da molécula de RNA durante a tradução e posteriormente ocorre a replicação do material genético. A montagem do vírus ocorre na superfície do retículo endoplasmático (RE). As partículas virais imaturas são transportadas através do Complexo de Golgi (CG), onde amadurecem e convertem a sua forma infecciosa. Os vírus maduros são liberados via exocitose.

Fonte: Adaptado de *Nature Publishing Group* Mukhopadhyay, S., Kuhn, R. J., & Rossmann M. G *Nature Reviews Microbiology* 3, 13–22 (2005) (40).

1.5. Diversidade genética viral

Estudos envolvendo análise filogenética sugerem a origem do YFV na região da África há 3.000 anos (42, 43) e demonstram que os genótipos de YFV africanos são mais heterogêneos em relação aos genótipos circulantes na América do Sul, o que corrobora a ideia que o vírus não é originário das Américas (42).

Entender a relação genética entre as cepas virais é de grande importância para o entendimento da epidemiologia molecular e dispersão global do vírus. (24). Estudos baseados nas análises das proteínas: prM/M, E e 3'UTR do genoma do YFV permitiram (44) a identificação de 7 genótipos diferentes, sendo 5 genótipos Africanos (Oeste Africano I e II, Leste Africano, Leste e Centro Africano, Angolano) e dois Americanos (América do Sul I e II) (45, 46).

O genótipo Angolano é extremamente divergente dos outros genótipos de origem Africana (17,1%-25,0%) sugerindo que tenha se desenvolvido independentemente (3, 26, 44). As cepas do Oeste Africano I possuem uma dispersão mais heterogênea em relação ao Oeste Africano II e em relação ao genótipo Leste e Centro africano, foi demonstrada, uma ampla distribuição e surtos menos intensos (47).

Estudos realizados por Von Lindemann e colaboradores., 2006, analisaram genomas completos de diversas sequências africanas e sugere que os cinco genótipos africanos estão dentro do parâmetro definido por Mutebi e colaboradores., 2001, ou seja, os cinco genótipos africanos possuem pelo menos 7% de divergência nucleotídicas entre si (26, 48). Quando comparados aos dois genótipos Sul-Americanos, foi encontrada uma divergência de aproximadamente 15% entre eles (49).

Em relação aos genótipos Americanos, estudos envolvendo análises filogenéticas sugerem que o genótipo Sul Americano I é mais encontrado no Brasil, Colômbia, Equador, Venezuela e Caribe. Por outro lado, o genótipo II é mais encontrado no Peru, Bolívia, com alguns isolados no Brasil e Trinidad e Tobago (26, 50). A divergência entre os aminoácidos dos genótipos Sul Americano I e II está próximo aos 7,6% (26).

O genótipo Sul Americano I apresenta cinco subgrupos denominadas: 1A,1B,1C,1D,1E. Esses subgrupos foram associados a epidemias que ocorreram no Brasil e outros países das Américas tropicais e subtropicais nos últimos anos (51).

No Brasil, nos anos 2000, apenas os subgrupos 1D e 1E foram detectados e desde 2008, apenas o subgrupo 1E vem sendo descrito durante os surtos (50).

1.6. Diagnóstico

O diagnóstico da infecção pelo YFV é baseado em testes moleculares e sorológicos realizados após o surgimento dos sintomas ou *post-mortem*.

O diagnóstico molecular consiste na detecção do RNA viral pela técnica de transcrição reversa (RT) seguida da reação em cadeia da polimerase em tempo real (qPCR), que é realizado em amostras de sangue e de urina, (56,57) denominado RT-qPCR. Esse teste é considerado rápido, sensível, e altamente específico na fase inicial de viremia por possuir *primers* que anelam em uma região altamente conservada entre os diferentes genótipos (5'UTR) do vírus selvagem, incluindo o vacinal e portanto, não podendo distinguir entre um e outro (52).

Entretanto, para os casos com relato de vacinação nos últimos 30 dias antes do início dos sintomas, são utilizados *primers* que anelam na região da proteína NS5 em consequência das alterações presentes nessa região, não detectam a presença do vírus selvagem devido a sua alta especificidade com o vírus vacinal (52, 53)

Segundo a Organização Pan-americana de Saúde (OPAS/PAHO), é possível realizar a detecção do RNA viral por métodos moleculares durante os 10 primeiros dias desde o início dos sintomas (54), porém, recentemente Casadio e colaboradores., 2019, demonstraram que o vírus pode ser detectado no sangue por até 28 dias e por até 47 dias na urina nos casos mais duradouros (55).

O diagnóstico sorológico, consiste na detecção de anticorpos da classe IgM pela técnica ELISA (do inglês, *Enzyme-Linked Immunosorbent Assay*), permitindo diagnosticar a infecção aguda pelo YFV (56), porém, nesse tipo de teste pode ser que ocorra reação cruzada com outros flavivírus como Dengue e Zika vírus (54).

Para os casos mais graves que evoluem a óbito existe o diagnóstico *post-mortem*, no qual é realizado o exame histopatológico a partir de cortes de fígado utilizando a técnica de imuno-histoquímica para a detecção do YFV no tecido, além da detecção do genoma viral pela técnica de PCR (57).

1.7. Epidemiologia no mundo

A FA é uma doença de notificação compulsória que ocorre no Brasil, principalmente nos meses de dezembro a maio, período de temperaturas elevadas com chuvas moderadas a intensas (14).

Anualmente, no mundo, aproximadamente 200.000 casos de FA são reportados com taxa de mortalidade em torno de 15% (aproximadamente 30.000 mortes) (58), sendo que 90% dos casos ocorrem no continente Africano (59). Nas últimas duas décadas, epidemias têm sido recorrentes na África subsaariana e na África do Sul tropical, sendo que entre os anos de 1985 e 2001, dos 25 surtos relatados: 22 ocorreram no Oeste Africano e 3 no Leste e Centro Africano (60).

Normalmente, os surtos de FA nos países Africanos estão associados à circulação do YFV pelo ciclo selvagem ou intermediário, porém ainda há relatos de ocorrência do ciclo urbano (61).

Em dezembro de 2015, os primeiros casos de FA foram identificados na Angola e em sua fronteira com a República Democrática do Congo (RDC), que declarou um surto epidêmico de FA com 90% dos seus casos confirmados na Província Central do Congo. Em resposta a essas epidemias, quase todo o estoque global de vacinação foi direcionado para Angola e em Abril de 2016 as doses já eram escassas (62, 63). Em Novembro de 2016, os surtos foram considerados pela Organização Mundial da Saúde (OMS) os maiores dos últimos 30 anos (64), com casos importados da Angola para outros países, como a China, demonstrando que a FA representa uma grave ameaça mundial em casos de baixa cobertura vacinal (65).

Portanto, foi criada sob direção da OMS, UNICEF (Fundo das Nações Unidas para a infância) e Gavi (Alianças para as Vacinas), a estratégia de “Eliminação da Epidemia de Febre Amarela” (do inglês – *Eliminate Yellow Fever Epidemics Strategy – EYE*), com o objetivo de proteger mais de 1 bilhão de pessoas, com a vacinação até 2026 (60, 65), sendo o Brasil, um dos 40 países que farão parte dessa estratégia.

1.8. Epidemiologia no Brasil

No Brasil, desde 2000, a preocupação das autoridades de vigilância em saúde aumentou devido a surtos fora da região amazônica. Até então a FA era restrita a essa região com casos esporádicos fora dela (66). De 2007 a 2009, o vírus atingiu as regiões Sudeste e Sul do país e causou mais de 100 casos da doença com letalidade de 51%, dando início a circulação do genótipo Sul-Americano IE (67).

Por esse motivo, o monitoramento do YFV tornou-se mais intenso nas regiões endêmicas (região amazônica) e extra-amazônica, registrando ocasionalmente períodos epidêmicos, caracterizando a reemergência do vírus no Brasil (1) (Figura 4).

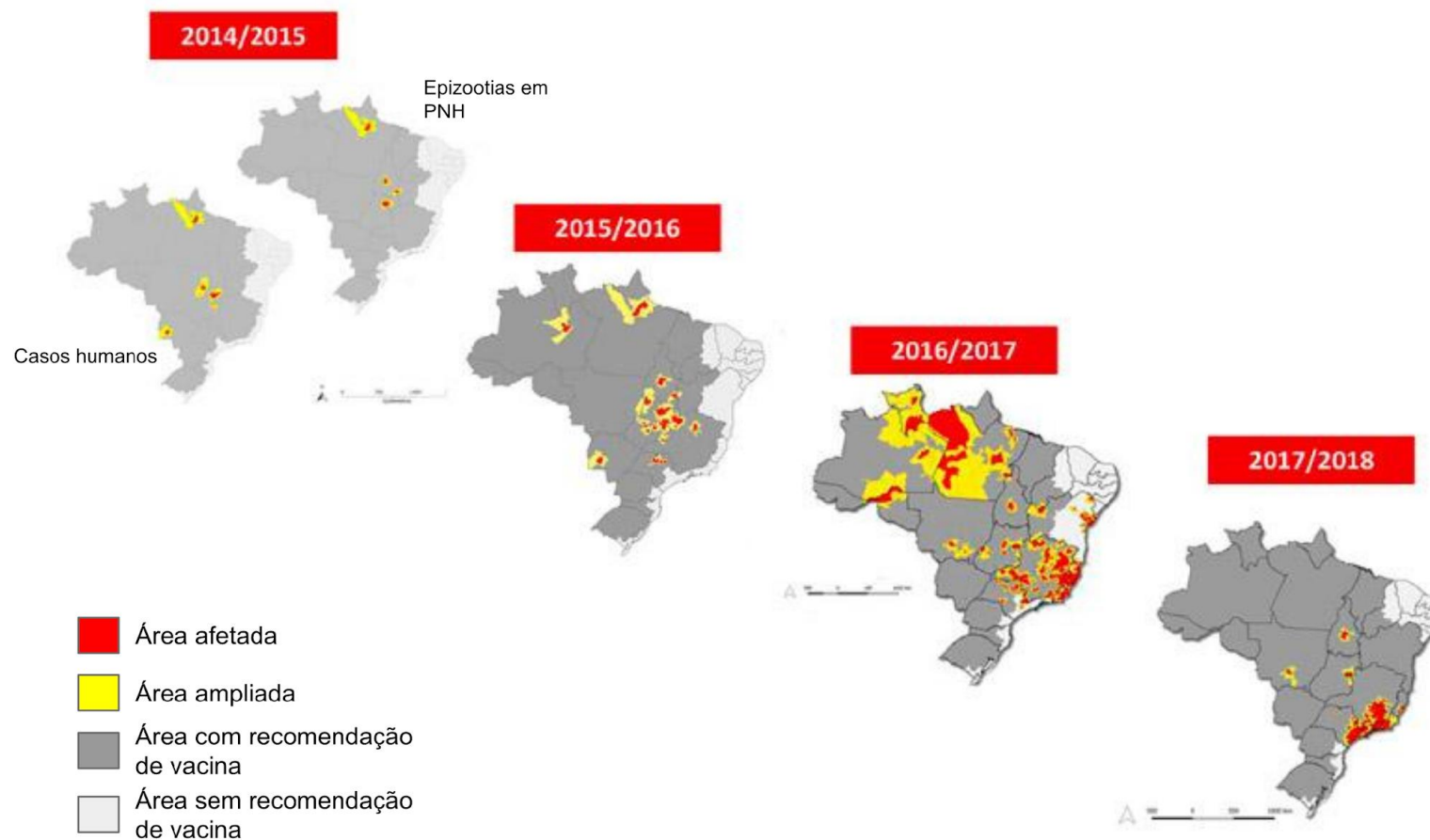


Figura 4: Reemergência YFV no Brasil.

Regiões endêmicas (região amazônica) e extra endêmicas estão representadas no mapa de cinza escuro (áreas com recomendação de vacina). Nota-se uma variação dessas regiões ao longo dos anos devido ao monitoramento dos surtos. Em 2017/2018, houve um aumento considerável de casos de FA em locais que não possuíam recomendação vacinal (área metropolitana de São Paulo, Minas Gerais, Rio de Janeiro, Espírito Santo).

Fonte: Adaptado de Ministério da Saúde, Febre Amarela.(1)

Alguns fatores de risco são levados em consideração para explicar o aumento no número de casos de FA fora das regiões endêmicas. Entre eles, podemos evidenciar o desmatamento, as mudanças climáticas com maiores períodos de chuvas, expansão da agricultura para áreas antes pertencentes às florestas e o aumento da atividade laboral para dentro das florestas ao redor de cidades (66, 68), principalmente entre homens de 15 a 35 anos que trabalham em área de mata ou próxima a elas uma vez que o trabalho desses indivíduos faz com que se exponham com maior frequência a áreas e situações de risco, ou seja, entrem em contato durante um maior período de tempo com o vetor (mosquito) e o hospedeiro PNH devido as exigências do trabalho (14).

Entre 2014 e 2016 novos casos surgiram afetando principalmente a região do Centro Oeste e desde final de dezembro de 2016 até 22 de fevereiro de 2017, um surto de FA atingiu o Brasil, com 1345 casos suspeitos no território nacional, dos quais 295 foram confirmados e 215 mortes foram reportadas para o Ministério da Saúde do Brasil, sendo Minas Gerais um dos estados brasileiros mais afetados durante esse período (69).

De acordo com a OPAS/PAHO, de Janeiro de 2017 a maio de 2019, a América Latina presenciou um aumento expressivo no número de epizootias e casos humanos, com notificações provenientes da Bolívia, Brasil, Peru, Colômbia e Guiana Francesa (70).

No Brasil, o surto recente mais expressivo afetou principalmente a região Sudeste, área mais populosa e mais povoada do país. Somente no Estado de São Paulo, foram confirmados 570 casos com 188 óbitos entre janeiro de 2018 (71) e junho de 2019 (72) com uma taxa de mortalidade de aproximadamente 34,2% em 2018 e 18,2% em 2019 (Figuras 5 e 6). O surto de 2018 ocorreu principalmente nas áreas ao redor da grande São Paulo, como por exemplo o município de Mairiporã, um dos mais atingidos, concentrando o maior número de casos (71).

Entretanto, em 2019, com as campanhas de vacinação em massa que já estavam em andamento e com o monitoramento intenso da movimentação dos vetores da doença e dos PNHs, o esperado era o aparecimento de novos casos na região mais ao sul do estado de São Paulo, conhecido como Vale do Ribeira, e segundo o mapeamento apresentado pelo boletim nº19 do Ministério da Saúde de São Paulo e Centro de vigilância epidemiológica (CVE), foi o que aconteceu para a

maioria dos casos, contando com alguns isolados em municípios da grande São Paulo (72).

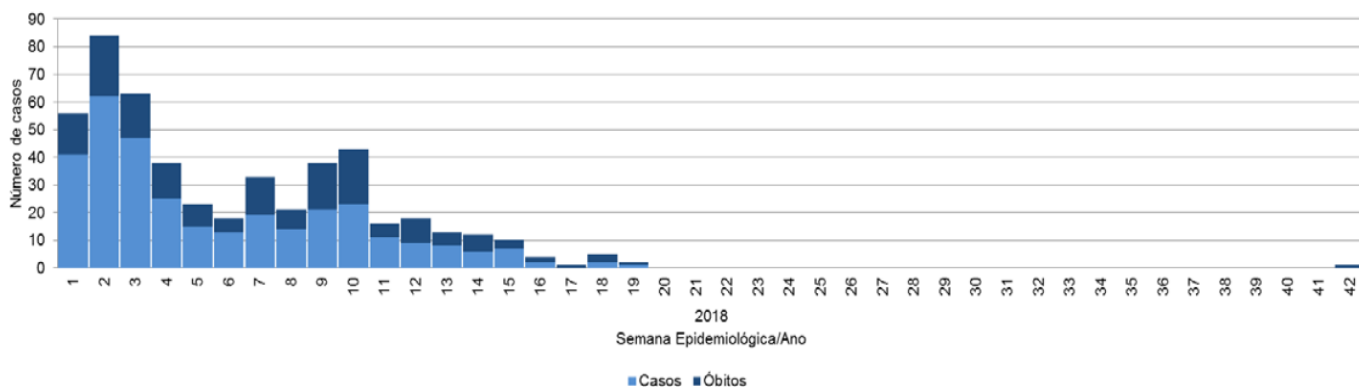


Figura 5: Distribuição dos casos autóctones e óbitos de YFV no Estado de São Paulo durante o ano de 2018, segundo semana epidemiológica.

Fonte: Ministério da Saúde, Centro de vigilância (71)

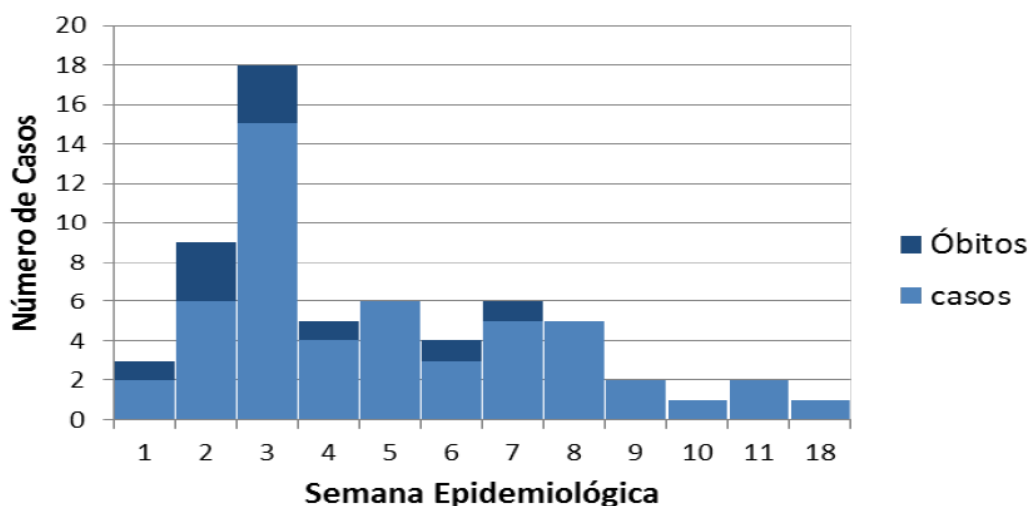


Figura 6: Distribuição dos casos autóctones de YFV no Estado de São Paulo durante o surto ocorrido nos meses de janeiro a junho de 2019.

Fonte: Ministério da Saúde, Centro de vigilância epidemiológica (CVE), Febre Amarela. Boletim nº 19 (72)

Em consequência aos atendimentos realizados pelo corpo clínico da Divisão de Moléstias Infecciosas do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP), durante as duas epidemias de FA que ocorreram em São Paulo em 2018 e 2019, o Laboratório de Gastroenterologia e Hepatologia Tropical (LIM07) vinculado ao departamento de Gastroenterologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) teve a oportunidade de realizar o diagnóstico molecular da infecção pelo YFV de todos os pacientes admitidos no HCFMUSP. Dessa maneira o corpo clínico assistencial obteve informações importantes que possibilitaram a agilidade da equipe médica no manejo desses pacientes. Vale ressaltar que estes eram pacientes graves (4) e muitos foram internados em UTIs e alguns foram submetidos a transplante hepático (73).

Esse contato direto com a equipe médica possibilitou a observação de um padrão de evolução clínica diferenciado entre os pacientes dos surtos de 2018 e 2019, sendo a taxa de mortalidade menor para o ano de 2019 (18,2% para 2019 e 34,2% para 2018), levantou a hipótese se os vírus circulantes nos surtos apresentam alguma diferença genética entre si. Utilizando metodologias de última geração, a análise do genoma completo do YFV fornecerá informações relevantes sobre a diversidade genética viral que poderá facilitar o monitoramento epidemiológico e tornar-se uma ferramenta para melhorar o entendimento do genoma viral e suas mutações.

2. Objetivos

Principal

Avaliar a diversidade genética do vírus da Febre Amarela de pacientes internados no HCFMUSP nos surtos de 2018 e 2019.

Específicos

Dentre os diferentes casos internados no HCFMUSP, pretendemos:

- Determinar a carga viral do YFV;
- Sequenciar o genoma completo do YFV;
- Determinar o genótipo do YFV;
- Realizar análise filogenética das cepas circulantes na Grande São Paulo;

3. Materiais & Métodos

3.1. Casuística

O estudo foi composto por 56 pacientes internados no Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) com diagnóstico de febre amarela. Após o paciente ser informado sobre o projeto de pesquisa e ter assinado o termo de consentimento livre e esclarecido (TCLE), as amostras de sangue e urina foram coletadas. Este trabalho teve aprovação do comitê de ética em pesquisa CAPesq do HCFMUSP (CAAE: 74535417.3.1001.0068).

3.1.1. Critérios de inclusão

- a) Pacientes admitidos no Hospital das Clínicas de São Paulo com o resultado de RT-qPCR positivo para YFV;
- b) Pacientes que assinaram o TCLE;
- c) Pacientes com resultado positivo de RT-qPCR para a infecção do YFV selvagem;
- d) Pacientes maiores de 18 anos;
- e) Pacientes que sobreviveram a infecção por YFV.

3.1.2. Critérios de exclusão

- a) Pacientes admitidos no Hospital das Clínicas de São Paulo com o resultado de RT-qPCR negativo para o YFV;
- b) Pacientes que não assinaram o TCLE;
- c) Pacientes com resultado positivo de RT-qPCR para a infecção do YFV vacinal;
- d) Pacientes menores de 18 anos;
- e) Pacientes que evoluíram a óbito.

3.1.3. Dados demográficos

Para o levantamento de dados demográficos, foram utilizados critérios como idade, sexo, carga viral, média de Cts e contagem de dias após início de sintomas em

relação à primeira coleta. O objetivo foi analisar se houve diferença estatisticamente significativa entre os anos de infecção e a diferença clínica observada. Além desses dados, também foram aferidos os locais de provável infecção para realizar o mapeamento geográfico dos casos durante os surtos de 2018 e 2019.

3.2. Extração do RNA viral

O RNA viral das amostras referente a primeira coleta, foi extraído a partir de 140µl de soro/plasma ou urina utilizando o QiaAmp® Viral RNA mini Kit (Qiagen, Hilden, Alemanha), de acordo com as instruções do fabricante. Após a extração, o RNA viral foi armazenado a -80°C até o uso.

3.3. qPCR quantitativo - Carga viral

A detecção do RNA viral nas amostras de soro e urina foi realizada por PCR em tempo real com uma etapa inicial de transcrição reversa (RT-qPCR) utilizando *primers* e sondas para o vírus selvagem e em caso de suspeita de infecção pelo vírus vacinal foram utilizados *primers* e sondas específicos para sua detecção. A quantificação viral seguiu o protocolo, conforme descrito por Casadio e colaboradores, 2019, utilizando uma curva sintética (55). Em ambas as situações, as amostras foram testadas em triplicatas na presença de controles negativos e positivos. As reações foram realizadas no equipamento 7500 *Fast Real Time PCR system* (Applied Biosystems™, Thermo Fisher Scientific Brand, Carlsbad, CA, USA) sob as seguintes condições de ciclagem: 50°C por 30 minutos, 95°C por 2 minutos e então 45 ciclos de 95°C por 15 segundos e 55°C por 1 minuto.

3.4. Amplificação do genoma completo do YFV

O RNA viral foi transcrito em cDNA utilizando a enzima M-MLV Transcriptase Reversa com 150 ng de *random primer* (Invitrogen™, Thermo Fisher Scientific Brand, Carlsbad, CA, USA) seguindo as instruções do fabricante.

O genoma completo do YFV foi amplificado por PCR utilizando combinações de *primers* que anelam ao longo do genoma viral e geram fragmentos com sobreposição parcial entre si (150-200nt), cobrindo toda a sequência nucleotídica do genoma. Para isso, foram desenhados 11 pares de *primers* (Tabela 2) e também

foram utilizados *primers* previamente publicados na literatura por Faria e colaboradores., 2018 (74).

Tabela 2: Descrição dos *primers* utilizados e suas posições de anelamento

| Fragmentos | Posição no nt | Primer | Sequência 5' - 3' | Tamanho do Amplicon | Referência |
|------------|---------------|--------|------------------------|---------------------|-----------------------|
| F1 | 28 | 28F | CATTGGTCTGCAAATCGAGTTG | 1.120pb | Estudo atual |
| | 1148 | 1148R | TTCATTGAGGGGGTGCATGG | | |
| F2 | 1004 | 1004F | TYAAGACTGCACTGTARACAC | 1.131pb | Estudo atual |
| | 2135 | 2135R | TGATGCARGTGAARGTGCCA | | |
| F3 | 1929 | 1929F | TRCCYTCTTGTGCCACTG | 1.136pb | Estudo atual |
| | 3065 | 3065R | CYCGAGTGATAYATGGATGCAG | | |
| F4 | 2946 | 2946F | ATCCATGTTCCATTACCTCATG | 1.230pb | Estudo atual |
| | 4176 | 4176R | AATGGYTGAGGTGAGACTTGC | | |
| F5 | 3981 | 11L | TGACAATGGCTGAGGTGAGACT | 1.208pb | Faria N.R et al.,2018 |
| | 5189 | 13R | GTCTTGTTTTCCAGCTCCAGG | | |
| F6 | 5080 | 14L | AACTGAGGTGAAAGAGGAGGGG | 1.197pb | Faria N.R et al.,2018 |
| | 6277 | 16R | CTCAAAGCACCACCTTCGGTCA | | |
| F7 | 6191 | 6191F | CTCTCCACCYTTCTTGCCA | 1.150pb | Estudo atual |
| | 7341 | 7341R | GTGGCTGGAATCAATAACRGTG | | |
| F8 | 7182 | 7182F | TTCCTCRATGTCRACTGTTGGA | 1.175pb | Estudo atual |
| | 8357 | 8357R | GTGTGAAGGTGTTGGCTCCA | | |
| F9 | 8175 | 8175F | GGRCGCCTCATTCTYCTCAT | 1.289pb | Estudo atual |
| | 9464 | 9464R | ATGACACAGCTGGATGGGAC | | |
| F10 | 9237 | 9237F | GTTGGTGATGGTRTTCAARGC | 875pb | Estudo atual |
| | 10112 | 10112R | TGGCTGTTTCTTCTGCTGTTT | | |
| F11 | 9978 | 9978F | GTTTTGUCCTGCATGTGTGG | 989pb | Estudo atual |
| | 10967 | 10967R | TCTTCCCTGGCGTCAATATG | | |
| F11 D | 9114 | 25L | TGAATGAGGACCACTGGGCATC | 1.222pb | Faria N.R et al.,2018 |
| | 10336 | 27R | CTGCAGATCAGCATCCACAGAG | | |

nt: nucleotideo; F e L: primer forward; R: primer reverso

A reação de PCR foi realizada em 50µl de volume final composto por 35µl de água, 5µl de Buffer, 1µl de dNTPs (10nM), 1,5µl de MgCl₂ (50nM), 1µl do primer F (20nM), 1µl do primer R (20nM), 5U da *Platinum Taq DNA Polimerase* (Invitrogen™, Thermo Fisher Scientific Brand, Carlsbad, CA, USA) e 5µl do cDNA. O protocolo de ciclagem foi: 94°C por 5 minutos, seguido de 45 ciclos a 94°C por 30 segundos, 65°C por 30 segundos e 72°C por 90 segundos, e uma extensão final a 72°C por 10 minutos e incubação a 10°C.

3.5. Detecção do DNA amplificado

Os produtos de PCR foram submetidos à eletroforese em gel de agarose a 2% adicionado de 1X SYBR Safe (Invitrogen™, Thermo Fisher Scientific Brand, Carlsbad, CA, USA). Foram misturados 4µl do produto de PCR com 1µl do tampão 10X Bluejuice™ Gel Loading Buffer (Invitrogen™ Life Technologies, Carlsbad, CA, EUA). Como marcador de tamanho e peso molecular foi usado o padrão GeneRuler™ Express DNA Ladder (Invitrogen™ Life Technologies, Carlsbad, CA, EUA). Ao final, o produto amplificado foi visualizado em um transiluminador.

3.6. Sequenciamento de nova geração (NGS)

Os produtos de PCR foram quantificados utilizando o método fluorimétrico Qubit® (Qubit 4 Fluorometer™; Thermo Fisher Scientific, Waltham, Massachusetts, EUA). Os 11 produtos de PCR de cada amostra (11 fragmentos/amostra) foram ajustados para que ficassem em uma mesma concentração, e só então 5µL de cada diluição foram transferidos para um único tubo formando uma mistura (*pool*) de PCR, sendo esse processo realizado para cada amostra. Posteriormente, cada *pool* foi novamente quantificado e a concentração ajustada para 0,8 ng/µL.

Em seguida foi realizado o preparo da biblioteca de acordo com o protocolo *Nextera® XT DNA Sample Preparation* (Illumina Inc™., San Diego, CA, U.S.A.). A primeira etapa foi a de fragmentação e tagmentação do DNA, seguida de uma reação de PCR para a inserção dos adaptadores e indexes, que tem como função identificar cada amostra. Após esta etapa, o produto de PCR foi purificado utilizando *beads* magnéticas *AMPure XP®* (Beckman Coulter™; Life Sciences Division Headquarters; Indianapolis, USA), e então foram novamente quantificados e a concentração foi ajustada para 3 nM.

A próxima etapa foi adicionar 5µL de PCR purificado de cada amostra (a 3 nM) em um tubo único compondo a biblioteca, ou seja, uma mistura equimolar de todas as amostras. A biblioteca foi quantificada e a concentração foi ajustada para 2 nM, em seguida foi desnaturada utilizando uma solução de NaOH a 0,2 N (v/v) e em seguida diluída a 10 pM pela adição de 990µL de tampão HT1. A biblioteca desnaturada (600µL) foi adicionada ao cartucho de reagentes MiSeq V2 300 ciclos e iniciou-se o

sequenciamento na plataforma da Illumina, MiSeq® (Illumina Inc™., San Diego, CA, U.S.A.).

3.7. Análise das sequências dos genomas completos do YFV

3.7.1. Controle de qualidade dos dados brutos de sequenciamento NGS, trimagem e filtros

O primeiro passo da análise foi verificar a qualidade do sequenciamento, a partir dos dados brutos (.fastq.gz), verificando o tamanho mínimo das leituras de sequenciamento e o escore de qualidade (Qp; *phred score*). Foram utilizados o limiar de 50 pb (pares de bases) para o tamanho mínimo e Qp=30 para o escore mínimo de qualidade de leitura.

Adicionalmente, as leituras foram trimadas, ou seja, foram retirados os 9 nucleotídeos iniciais e 5 nucleotídeos finais, devido à queda de qualidade nas extremidades e sequências com “N” foram excluídas. Essas etapas foram realizadas utilizando o *software cutadapt* (75).

3.7.2. Montagem do genoma completo YFV

A montagem dos genomas completos de YFV foi realizada utilizando um conjunto de evidências geradas por três métodos de montagem diferentes, mais a análise e inspeção manual. Dois destes métodos utilizaram a abordagem *de novo*, onde os *contigs* são montados com base em sobreposição de grafos e abundâncias sem levar em consideração uma sequência de referência. Os *softwares Spades* (76, 77) e IVA (78) (parâmetros *default*, montagem *denovo*) foram utilizados nesta abordagem.

O terceiro método utilizou uma sequência referência para fazer a montagem dos genomas, o *software* utilizado foi o *Spades* (76, 77) (parâmetros *default* e a referência YFV MF538786.2/RJ104 como *trusted contigs*). Após a execução dos diferentes *pipelines*, foi realizada a análise manual para montar a melhor sequência possível. Detalhes adicionais são abordados nos tópicos abaixo.

3.7.2.1. Montagem contra sequência referência - *Spades*

A escolha da sequência referência de YFV MF538786.2 (RJ104) foi realizada fazendo uma análise dos resultados gerados pelo BLASTn (*Basic Local Alignment Search Tool*), uma ferramenta de busca de sequências similares utilizando as sequências do banco de dados do NCBI (*National Center for Biotechnology Information, website <http://www.ncbi.nlm.nih.gov/>*). Portanto, as sequências geradas neste estudo foram comparadas com o banco de dados do NCBI e foram analisados os parâmetros de similaridade (maior % identidade do alinhamento), melhor cobertura de alinhamento (99%-100%) e menor valor de e-value.

Após a escolha da referência, foi realizado o mapeamento das sequências obtidas contra o genoma humano GCF_000001405.12 (hg38) para a exclusão dessas leituras em ambos os pares (*forward* e *reverse*).

As leituras de sequenciamento já filtradas pelo *cutadapt* (75) e não mapeadas contra o genoma humano foram utilizadas para montar o *contig*, utilizando o *software Spades* (76, 77) e o parâmetro *trusted contigs*.

3.7.2.2. Montagem de novo - *Spades*

Na técnica de montagem *de novo* não há utilização de uma sequência referência, pois não é necessário uma orientação para realizar essa montagem (81). O genoma é montado a partir das leituras de sequenciamento, minimizando o risco de divergências decorrentes do arquivo de referência. Foi realizado o mesmo *pipeline* descrito no item 3.7.2.1, porém, sem anexar o genoma de referência viral (RJ104) e dessa forma, os *contigs* foram gerados sem interferência de uma referência molde.

3.7.2.3. Montagem de novo - IVA assembler

IVA é um montador de genomas virais que realiza a montagem pelo modo denominado *de novo*, que foi designado para montar genomas sem repetições internas, usando as leituras realizadas pelo sequenciador em populações com elevada e profunda variabilidade genética (78).

Sendo assim, o IVA foi utilizado como um *pipeline* adicional de forma a fornecer mais evidências para a montagem mais precisa, comparativa e que facilitasse a técnica da análise manual.

3.7.2.4. Análise manual

Com base nas evidências de montagem e nos *contigs* gerados pelas três técnicas descritas acima, os *scaffolds* foram selecionados de acordo com a completude e congruência do que seria esperado para o genoma de YFV. Adicionalmente, as leituras de sequenciamento foram mapeadas nos *scaffolds* com o *software* BWA (79) utilizando parâmetros *default* de forma a facilitar a inspeção e identificação de possíveis artefatos de montagem.

3.7.3. Chamada de variantes

Para realizar a chamada de variantes, foi necessário indexar e mapear a referência (RJ104), utilizando o *software* BWA (79) e o alinhamento em formato bam gerado pelo *bwa mem* (parâmetro *default*).

O *software samtools view* (80) foi utilizado para remover as leituras de sequenciamento com alinhamentos secundários, suplementar e com baixa qualidade de mapeamento (Qmap < 30) ou não mapeados frente ao arquivo .bam gerado após o mapeamento com o genoma referência. A ordenação do arquivo bam foi gerada com a ferramenta *samtools sort* e o *index* foi gerado utilizando o *samtools index*. Ao final deste processo, o arquivo de sequências gerado poderá ser visualizado no *Integrative Genomics Viewer (IGV)* para a verificação das variantes (81).

Para identificar os possíveis SNPs (do inglês, *single nucleotide polymorphism*) e pequenos Indels (inserção e/ou deleção de nucleotídeo), foram utilizados os softwares *Freebayes* (82) e *GATK haplotype caller* (83) para gerar os arquivos no formato .vcf contendo todas as variantes identificadas e suas coordenadas genômicas frente a referência MF538786.2. Desta forma, o arquivo final de variantes identificadas pode ser inspecionado em um editor de texto com o arquivo .vcf correspondente.

3.7.4. Análise de intersecção e união de variantes nos surtos de 2018 e 2019

Todos os arquivos .vcf gerados foram agrupados e analisados utilizando o *software* BCFtools (84) que separou as variantes encontradas em três grupos: exclusivas de 2018, exclusivas de 2019 e intersecções entre 2018 e 2019.

Além disso, para a análise de variantes, as sequências nucleotídicas do YFV foram alinhadas utilizando o *software* Clustal W (85) integrado ao *software* MEGA7 (86), em seguida as extremidades 5' e 3' UTR dos genomas foram cortadas restando apenas a sequência da poliproteína viral (região codificadora de proteínas) para a identificação e classificação das substituições.

A classificação das substituições nucleotídicas em sinônimas ou não sinônimas foi realizada a partir da tradução dos nucleotídeos em aminoácido utilizando o *software* MEGA7. Entre as substituições não sinônimas, a qual resulta na troca do aminoácido, também foi avaliado se houve alteração da classe do aminoácido, ou seja, se ocorreu alteração das propriedades químicas de suas cadeias laterais.

3.7.5. Análise Filogenética

Foram realizados alinhamentos múltiplos de sequências referências de YFV retiradas do NCBI e 39 sequências do presente estudo, utilizando o *software* MAFFT (87), pelo seu servidor online (*multiple alignment using fast Fourier transform, website* <https://mafft.cbrc.jp/alignment/software/>) modificando os parâmetros de direção das sequências nucleotídicas, os quais foram escolhidos para ajustar a direção de acordo com a primeira sequência. O método escolhido para a realização do alinhamento foi G-INS (método lento e progressivo para até 200 sequências homólogas) e optamos pela matriz de escore 20PAM/K=2.

Foram selecionadas 73 sequências referências com os números de acesso descritos na Tabela 3.

Tabela 3: Descrição das sequências utilizadas para gerar as árvores filogenéticas

| Origem | Ano de isolamento | Número de acesso no GenBank | Origem | Ano de isolamento | Número de acesso no GenBank |
|----------------------------|-------------------|-----------------------------|-----------|-------------------|-----------------------------|
| Bolívia | 1999 | KF907504.1 | Brasil/SP | 2018 | MK583154.1 |
| Costa do Marfim | 1982 | U54798 | Brasil/SP | 2018 | MK583168.1 |
| Costa do Marfim | 1973 | JX898869.1 | Brasil/SP | 2018 | MK583173.1 |
| Costa do Marfim | 1999 | AY603338.1 | Brasil/SP | 2018 | MK583159.1 |
| Gambia | 2001 | AY572535.1 | Brasil/SP | 2018 | MK583171.1 |
| Gana | 1927 | KF769016.1 | Brasil/SP | 2018 | MK583174.1 |
| Trinidad | 1979 | AF094612.1 | Brasil/SP | 2018 | MK583170.1 |
| Angola | 1971 | AY968064.1 | Brasil/SP | 2018 | MK583150.1 |
| China/ importado da Angola | 2016 | KU921608.1 | Brasil/SP | 2018 | MK583108.1 |
| Uganda | 2010 | JN620362.1 | Brasil/SP | 2018 | MK583169.1 |
| Etiópia | 1961 | DQ235229.1 | Brasil/SP | 2018 | MK583176.1 |
| Senegal | 1995 | JX898868.1 | Brasil/SP | 2018 | MK583156.1 |
| Holanda - Importado | 2018 | MK292067.1 | Brasil/SP | 2018 | MK583181.1 |
| Nigeria | - | MN211310.1 | Brasil/SP | 2018 | MK583162.1 |
| Nigeria | - | MN211306.1 | Brasil/SP | 2018 | MK583161.1 |
| Nigeria | - | MN211311.1 | Brasil/SP | 2018 | MK583182.1 |
| Uganda | 1948 | AY968065.1 | Brasil/SP | 2018 | MK583175.1 |
| 17DD - vacina | - | DQ100292.1 | Brasil/SP | 2018 | MK583160.1 |
| 17D - vacina | 1985 | NC_002031 | Brasil/SP | 2018 | MK583172.1 |
| Guinéa Bissau | 1965 | KU978765.1 | Brasil/SP | 2018 | MK583177.1 |
| Sudão | 1941 | KU978764.1 | Brasil/SP | 2018 | MK583164.1 |
| Brasil/RJ | 2017 | MF538784.2 | Brasil/SP | 2018 | MK583178.1 |

continua

Tabela 3: Descrição das sequências utilizadas para gerar as árvores filogenéticas

| | | | | | |
|-----------|------|------------|-----------|------|------------|
| Brasil/RJ | 2017 | MF538783.2 | Brasil/SP | 2017 | MK583152.1 |
| Brasil/RJ | 2017 | MF538782.2 | Brasil/SP | 2018 | MK583179.1 |
| Brasil/RJ | 2017 | MF538786.2 | Brasil/SP | 2018 | MK583148.1 |
| Brasil/ES | 2017 | KY885001.1 | Brasil/SP | 2018 | MK583157.1 |
| Brasil/ES | 2017 | MF370531.1 | Brasil/SP | 2018 | MK583167.1 |
| Brasil/PA | 1984 | JF912184.1 | Brasil/SP | 2018 | MK583163.1 |
| Brasil/MG | 2017 | MF465805.1 | Brasil/SP | 2018 | MK583149.1 |
| Brasil/RO | 1983 | JF912181.1 | Brasil/SP | 2018 | MK583166.1 |
| Brasil/MG | 2001 | MG969501.1 | Brasil/SP | 2018 | MK583153.1 |
| Brasil/MG | 2017 | MF370533.1 | Brasil/SP | 2018 | MK583158.1 |
| Brasil/GO | 2000 | JF912187.1 | Brasil/SP | 2018 | MK583165.1 |
| Brasil/RR | 2002 | JF912190.1 | Brasil/SP | 2018 | MK583151.1 |
| Brasil/GO | 2017 | MK728873.1 | Brasil/SP | 2018 | MK583147.1 |
| Brasil/SP | 2018 | MK333798.1 | Brasil/SP | 2018 | MK583155.1 |
| Brasil/SP | 2018 | MK333799.1 | - | - | - |

(-) sem informação

conclusão tabela 3

O alinhamento múltiplo foi submetido à análise de qualidade utilizando a plataforma online do Guidance v2 (*website*, <http://guidance.tau.ac.il/ver2/>) (88), o qual gerou um escore (0-1) de confiança possibilitando a verificação de regiões com artefatos e não informativas.

O próximo passo foi gerar uma árvore de máxima verossimilhança (ML, do inglês *maximum likelihood*) utilizando o *software* RAx-ML (do inglês, *Randomized Axelerated Maximum Likelihood*) (89). O modelo utilizado para essa análise foi o GTR gama com o parâmetro alfa estimado automaticamente, com número de árvores iniciais igual a cinco (algoritmo rápido), e a "semente" aleatória igual a 1.

4. Resultados

4.1. Caracterização da amostra de estudo

Dentre os 192 pacientes atendidos durante o surto de 2018 e 2019 no HCFMUSP, cujas amostras de sangue e urina foram encaminhadas ao Laboratório de Gastroenterologia e Hepatologia Tropical (LIM07), foram excluídos 63 casos recebidos (32,81%), pois após a análise do RT-qPCR, observou-se que os mesmos eram negativos para o YFV (Figura 7). Assim, permaneceram um total de 129 pacientes e dentre eles, optamos por trabalhar com os que sobreviveram a infecção. Posteriormente, devido a limitações durante a técnica de PCR convencional, foi necessário excluir pacientes que possuíam Ct muito alto (>34,0), uma vez que a baixa quantidade de vírus dificultou a amplificação do genoma viral. Assim, foram testados 56 pacientes.

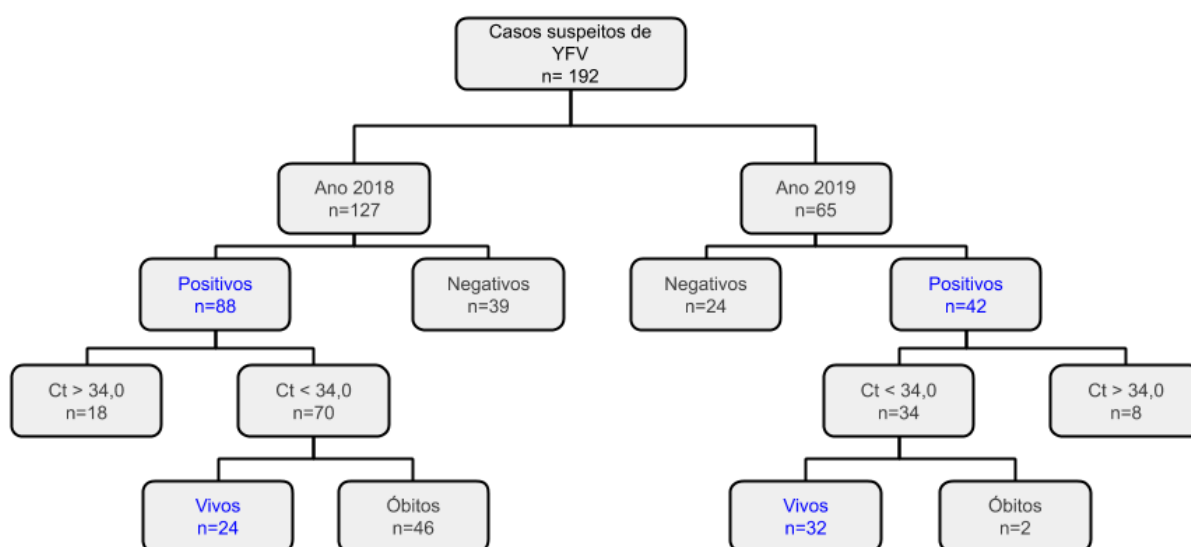


Figura 7: Fluxo de seleção das amostras em estudo.

Em relação a distribuição da amostra de acordo com o ano de infecção, os participantes que compuseram o estudo apresentaram características muito semelhantes entre si. A idade variou entre 19 a 88 anos e a maioria (85,7%) é pertencente ao gênero masculino. Todos os pacientes tiveram resultado positivo para o teste de YFV RT-qPCR (Tabela 4).

Tabela 4: Dados demográficos dos pacientes participantes do estudo.

| Variável | 2018 (n=24) | 2019 (n=32) | p valor |
|--|--------------|--------------|---------|
| Sexo | | | |
| Masculino, n (%) | 19 (79,2) | 29 (90,6) | 0,268* |
| Feminino, n (%) | 5 (20,8) | 3 (9,4) | |
| Idade (anos) | | | |
| Média (Min-Max) | 43,7 (19-74) | 45,7 (19-88) | 0,596** |
| Dias após início de sintomas*** | | | |
| Média (± sd) | 5,6 (± 2,2) | 5,2 (± 2,2) | 0,536** |
| Carga Viral (log10) | | | |
| Média (± sd) | 6,3 (± 1,32) | 6,4 (± 1,4) | 0,308** |
| Ct | | | |
| Média (± sd) | 27,7 (± 4,9) | 27,4 (± 3,9) | 0,328** |

N: número de amostras, Min: mínimo, Max: máximo, sd: desvio padrão - do inglês *standard deviation*, Ct: ciclo em relação ao threshold, *Teste exato de Fisher, **Teste Mann Whitney, $p < 0,005$; ***após a primeira coleta

Foi realizado o levantamento dos dados referente ao provável local de infecção pelo YFV dos pacientes incluídos neste trabalho e foi realizado um mapeamento dessas áreas utilizando a ferramenta de geolocalização do Google Maps® (Figura 8).

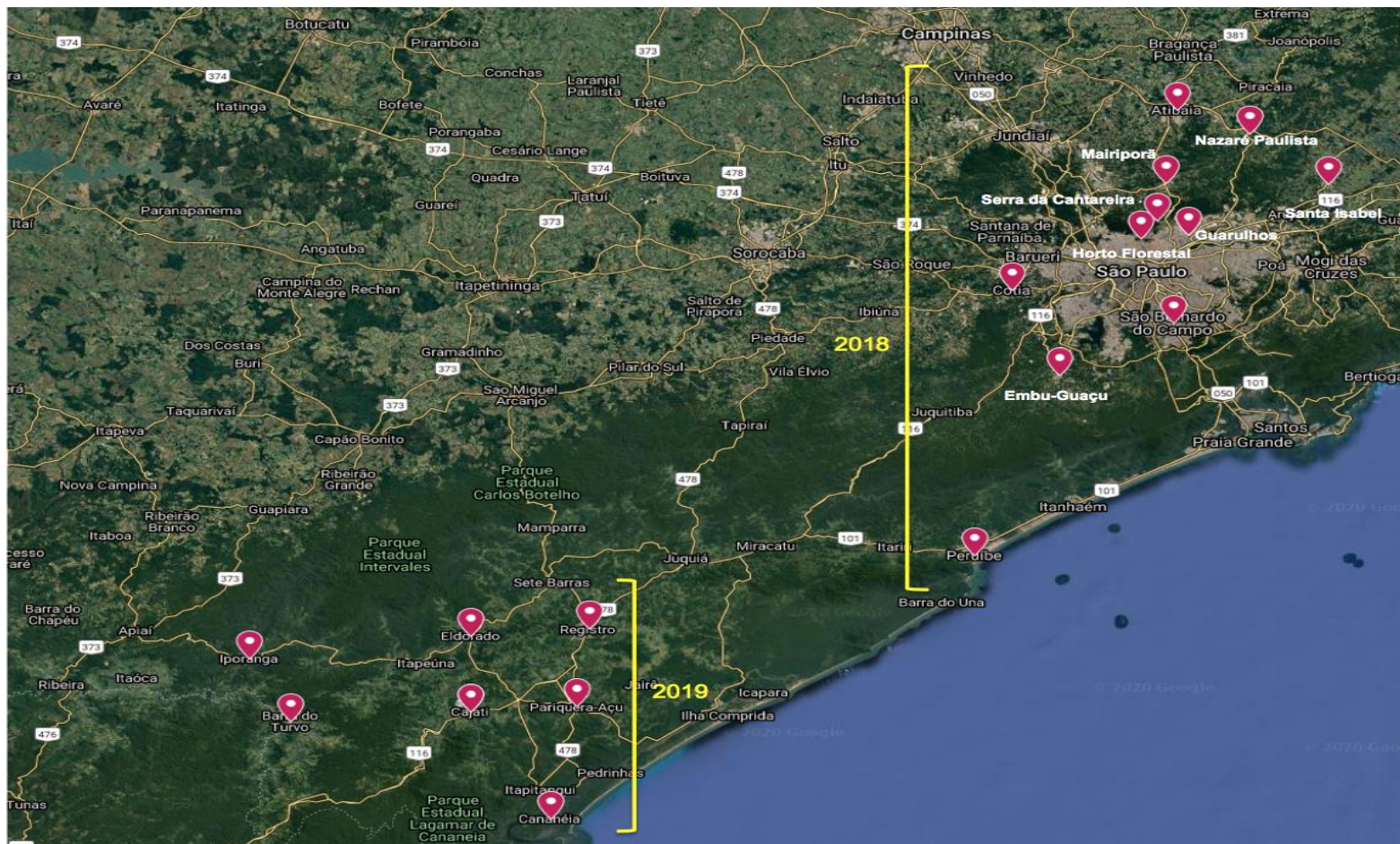


Figura 8: Mapeamento dos prováveis locais de infecção. Realizado com o auxílio da ferramenta fornecida pelo Google Maps®, destacando as cidades identificadas durante o levantamento dos dados dos pacientes participantes da casuística deste trabalho (identificadas por colchetes), indicando o ano correspondente à infecção. Apenas um paciente de 2019 indicou o local de infecção como sendo Mairiporã.

4.2. Amplificação do genoma completo

A técnica de PCR convencional foi aplicada para a amplificação do genoma completo do YFV utilizando onze pares de *primers* (Figura 9). Ao todo foi possível amplificar o genoma completo em 40 das 56 amostras (71,4%; média de Ct = 26,7). As 16 amostras restantes tiveram amplificação parcial ou ausência de amplificação; e em 12 amostras pelo menos um dos fragmentos foi amplificado (média de Ct = 29,2), entretanto em 4 amostras não observamos amplificação dos fragmentos (média de Ct = 30,8). A Figura 9 representa o resultado da amplificação dos 11 fragmentos de uma mesma amostra.

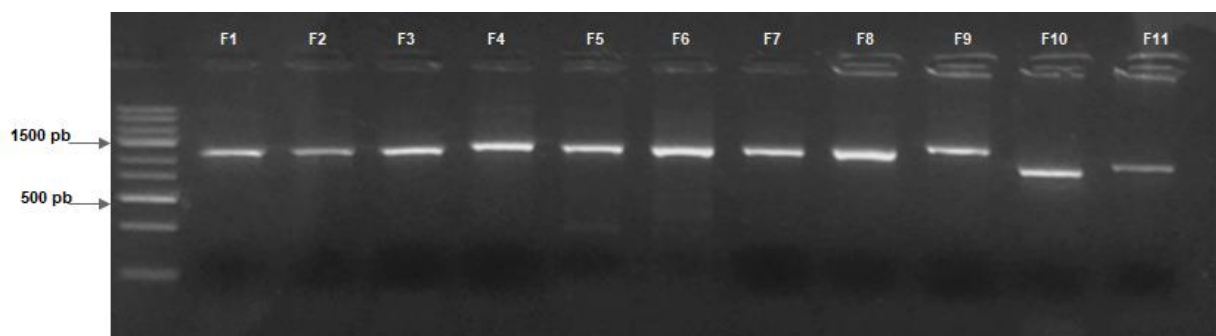


Figura 9: Foto dos produtos de amplificação por PCR convencional.

Para a análise dos 11 fragmentos foi utilizado como padrão de tamanho, no gel de agarose, o GeneRuler™ Express DNA Ladder (Invitrogen™ Life Technologies, Carlsbad, CA, EUA).

A região genômica mais difícil de amplificar foi a porção 3'UTR (correspondente ao *primer* F11) devido a sua alta complexidade estrutural, pois apresenta uma série de alças com pareamento entre as bases o que pode dificultar o anelamento dos primers. Como alternativa para isso utilizamos uma segunda opção de combinação de primers (F11D) que anela ao final da proteína NS5, gerando assim um fragmento menor e que não cobre a região da 3'UTR do genoma viral, mas sim o final da poliproteína viral.

4.3. Sequenciamento e montagem do genoma

As 40 amostras com amplificação do genoma completo do YFV foram submetidas ao procedimento de preparo de biblioteca e posterior sequenciamento de DNA (MiSeq Illumina). A qualidade das sequências geradas foi avaliada e as amostras

apresentaram sequências com valores acima do parâmetro preestabelecido Q30 (Figura 10) com exceção de uma amostra que apresentou baixa qualidade no sequenciamento, e devido a isso, seu genoma não seguiu para as análises posteriores, totalizando, portanto, 39 genomas montados. A cobertura média das leituras de sequências foi de 5.149x e a porcentagem do genoma coberto ficou entre 92,75% - 98,7%.

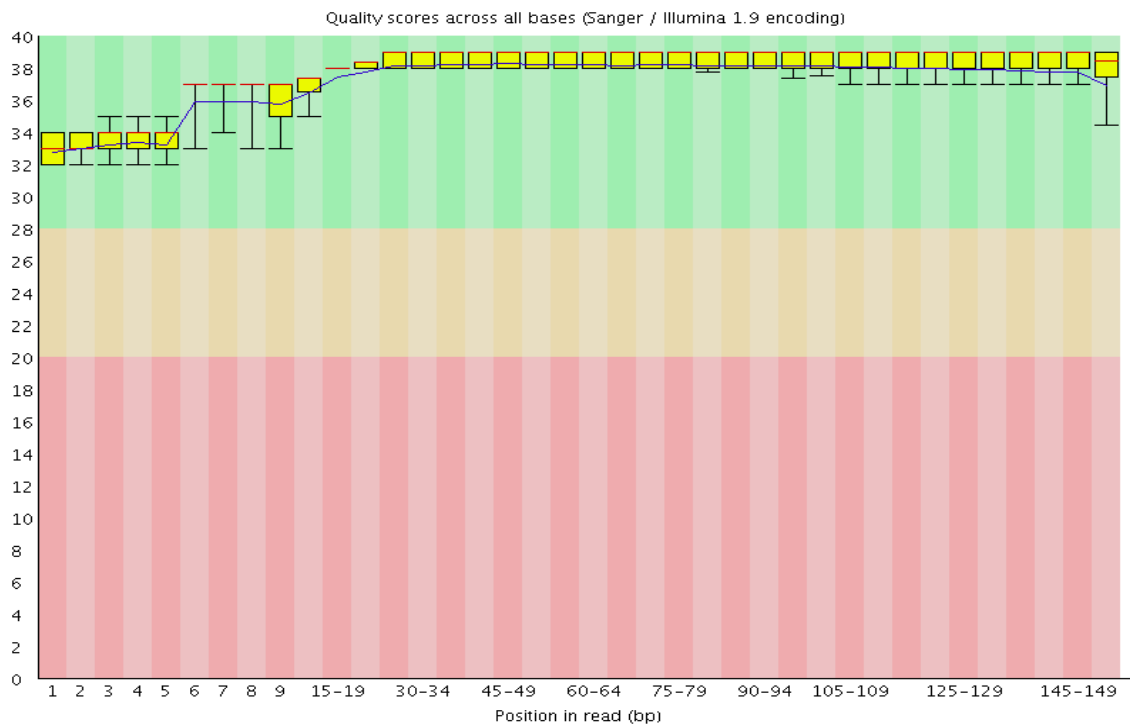


Figura 10: Representação gráfica gerada durante a análise de qualidade das sequências, representando o escore de qualidade das leituras de sequenciamento. O gráfico mostra um valor de qualidade (eixo Y) para todas as bases em cada uma das posições de leitura do sequenciamento (eixo X). A linha azul representa a qualidade média. O gráfico é dividido em três áreas: verde (boa qualidade), laranja (qualidade razoável) e vermelho (baixa qualidade).

Após a avaliação das métricas de qualidade, a média de cobertura do sequenciamento e a porcentagem do genoma montado, as sequências obtidas no estudo foram avaliadas em relação à sequência referência RJ104 (MF538786.2). Portanto, as análises apresentadas abaixo (Tabela 5) podem não refletir a completude real do genoma, mas sim a completude comparada com a referência RJ104 que

obteve 99,61% de identidade, *e-value*=0 e 100% de cobertura em relação às sequências geradas neste estudo.

Na tabela 5 é possível verificar que houve variação no tamanho do genoma entre as amostras, em decorrência da estratégia de amplificação utilizada. A combinação de *primers* F11 permitiu a montagem do genoma até a região 3'UTR (>10.600 pb). Entretanto, com a combinação de *primers* F11D o genoma montado ficou menor e não cobriu a região 3'UTR (<10.300 pb).

Tabela 5: Dados das sequências obtidas descrevendo os Cts das amostras usadas durante a amplificação da PCR, o tamanho (pb) e a média de cobertura (x) referente aos genomas completos ou quase completos.

| Sigla | Desfecho | Dias após início de sintomas | RT-qPCR (Ct) | Material | Cobertura do genoma pb (%) | Média de cobertura (x) |
|---------------|-----------------|-------------------------------------|---------------------|-----------------|-----------------------------------|-------------------------------|
| SP01/BRA/2018 | Alta | 5 | 20,6 | soro | 10279 (93,4) | 5.843 |
| SP02/BRA/2018 | Alta | 7 | 31,7 | soro | 10266 (93,3) | 5.279 |
| SP03/BRA/2018 | Alta | 4 | 29,9 | soro | 10269 (93,3) | 5.574 |
| SP04/BRA/2018 | Alta | 10 | 31,3 | soro | 10322 (93,8) | 4.305 |
| SP05/BRA/2018 | Alta | 6 | 28 | soro | 10268 (93,3) | 7.221 |
| SP06/BRA/2018 | Alta | 3 | 16,3 | soro | 10849 (98,5) | 5.899 |
| SP07/BRA/2018 | Alta | 3 | 23,9 | soro | 10279 (93,4) | 4.974 |
| SP08/BRA/2018 | Alta | 5 | 30,3 | soro | 10275 (93,4) | 5.522 |
| SP09/BRA/2018 | Alta | 6 | 33,1 | soro | 10852 (98,6) | 4.553 |
| SP10/BRA/2018 | Alta | 2 | 15 | soro | 10279 (93,4) | 4.791 |
| SP11/BRA/2018 | Alta | 4 | 29,9 | soro | 10187 (92,5) | 4.144 |
| SP12/BRA/2018 | Alta | 4 | 24,2 | soro | 10835 (98,4) | 6.397 |
| SP13/BRA/2018 | Alta | 9 | 29 | urina | 10839 (98,5) | 3.660 |
| SP14/BRA/2018 | Alta | 4 | 28 | soro | 10269 (93,3) | 4.414 |
| SP01/BRA/2019 | Alta | 3 | 28,4 | urina | 10266 (93,3) | 5.088 |
| SP02/BRA/2019 | Alta | 3 | 24,2 | soro | 10279 (93,4) | 4.418 |
| SP03/BRA/2019 | Alta | 10 | 20 | urina | 10279 (93,4) | 5.573 |
| SP04/BRA/2019 | Alta | 7 | 30,2 | soro | 10269 (93,3) | 4.910 |

*óbito após a alta; pb: pares de bases; (-) sem informação; RT-qPCR: transcrição reversa por PCR em Tempo Real; Ct: ciclo em relação ao threshold

continua

Tabela 5: Dados das sequências obtidas descrevendo os Cts das amostras usadas durante a amplificação da PCR, o tamanho (pb) e a média de cobertura (x) referente aos genomas completos ou quase completos.

| Sigla | Desfecho | Dias após início de sintomas | RT-qPCR (Ct) | Material | Cobertura do genoma pb (%) | Média de cobertura (x) |
|---------------|----------|------------------------------|--------------|----------|----------------------------|------------------------|
| SP05/BRA/2019 | Alta | 10 | 25,6 | urina | 10080 (91,6) | 5.472 |
| SP06/BRA/2019 | Alta | 4 | 22,5 | urina | 10269 (93,3) | 6.998 |
| SP07/BRA/2019 | Alta | 2 | 19,8 | soro | 10859 (98,6) | 6.557 |
| SP08/BRA/2019 | Alta | 4 | 31,6 | urina | 10269 (93,3) | 3.729 |
| SP09/BRA/2019 | Alta | 7 | 34,1 | urina | 10862 (98,7) | 7.348 |
| SP10/BRA/2019 | Alta | 7 | 28,2 | urina | 10860 (98,7) | 7.752 |
| SP11/BRA/2019 | Alta | 5 | 33,9 | urina | 10279 (93,4) | 6.153 |
| SP12/BRA/2019 | Alta | 7 | 26,8 | urina | 10277 (93,4) | 4.127 |
| SP13/BRA/2019 | Alta | 3 | 27,9 | soro | 10269 (93,3) | 3.675 |
| SP14/BRA/2019 | Alta | 6 | 19,6 | urina | 10855 (98,6) | 2.861 |
| SP15/BRA/2019 | Alta | 3 | 27,9 | urina | 10188 (92,5) | 5.728 |
| SP16/BRA/2019 | Alta | 3 | 27,1 | soro | 10855 (98,6) | 5.141 |
| SP17/BRA/2019 | Alta | 6 | 26,5 | urina | 10252 (93,1) | 1.675 |
| SP18/BRA/2019 | Alta | 6 | 29,1 | urina | 10257 (93,2) | 3.426 |
| SP19/BRA/2019 | Alta | 5 | 23,9 | urina | 10853 (98,6) | 2.678 |
| SP20/BRA/2019 | Alta | 4 | 29,2 | urina | 10283 (93,4) | 6.073 |
| SP21/BRA/2019 | Alta | 6 | 29 | urina | 10271 (93,3) | 4.840 |
| SP22/BRA/2019 | Alta | 4 | 30,8 | urina | 10853 (98,6) | 1.871 |
| SP23/BRA/2019 | Alta | 4 | 20,4 | urina | 10855 (98,6) | 6.267 |
| SP24/BRA/2019 | Alta* | 6 | 25,5 | urina | 10644 (96,7) | 4.108 |
| SP25/BRA/2019 | Alta* | 4 | 26,5 | urina | 10849 (98,6) | 7.769 |
| SP26/BRA/2019 | Alta | 5 | 30,2 | urina | - | - |

*óbito após a alta; pb: pares de bases; (-) sem informação, RT-qPCR: transcrição reversa PCR em Tempo Real; Ct: ciclo em relação ao threshold

conclusão tabela 5.

4.4. Análise de variantes

A partir dos dados de sequenciamento das 39 amostras com genoma montados, ou seja, que passaram no controle de qualidade (14 de 2018 e 25 de 2019), foi realizado o alinhamento dessas sequências contra a referência RJ104 (MF538786.2) no intuito de avaliar a diversidade genética entre as cepas virais dos surtos de 2018 e 2019.

Na análise geral realizada pelos *softwares FreeBayes* (82) e *GATK haplotype caller* (83) foi possível identificar todas as variantes presentes em todas as 39 sequências (Tabela 6).

Tabela 6: Quantidade de variantes por amostra, após análises realizadas com os softwares *Freebayes* e *GATK*.

| Amostra | Nº de variantes por amostra | Amostra | Nº de variantes por amostra |
|---------------|-----------------------------|---------------|-----------------------------|
| SP01/BRA/2018 | 55 | SP07/BRA/2019 | 27 |
| SP02/BRA/2018 | 42 | SP08/BRA/2019 | 39 |
| SP03/BRA/2018 | 33 | SP09/BRA/2019 | 39 |
| SP04/BRA/2018 | 48 | SP10/BRA/2019 | 34 |
| SP05/BRA/2018 | 37 | SP11/BRA/2019 | 37 |
| SP06/BRA/2018 | 48 | SP12/BRA/2019 | 48 |
| SP07/BRA/2018 | 46 | SP13/BRA/2019 | 42 |
| SP08/BRA/2018 | 52 | SP14/BRA/2019 | 42 |
| SP09/BRA/2018 | 42 | SP15/BRA/2019 | 43 |
| SP10/BRA/2018 | 36 | SP16/BRA/2019 | 47 |
| SP11/BRA/2018 | 45 | SP17/BRA/2019 | 35 |
| SP12/BRA/2018 | 34 | SP18/BRA/2019 | 38 |
| SP13/BRA/2018 | 55 | SP19/BRA/2019 | 53 |
| SP14/BRA/2018 | 38 | SP20/BRA/2019 | 28 |
| SP01/BRA/2019 | 42 | SP21/BRA/2019 | 37 |
| SP02/BRA/2019 | 36 | SP22/BRA/2019 | 34 |
| SP03/BRA/2019 | 36 | SP23/BRA/2019 | 41 |
| SP04/BRA/2019 | 35 | SP24/BRA/2019 | 38 |
| SP05/BRA/2019 | 39 | SP25/BRA/2019 | 43 |
| SP06/BRA/2019 | 43 | SP26/BRA/2019 | - |

(-): sem informação

Devido a um grande número de variantes e a complexidade da análise individual, optamos por fazer uma análise mais robusta nas variantes que seguissem um dos 3 parâmetros listados abaixo:

- a. Intersecção entre variantes das sequências de 2018 e 2019;
- b. Exclusivas de 2018;
- c. Exclusivas de 2019;

Aplicando esses parâmetros, foram encontradas 46 substituições nucleotídicas no total, sendo 20 delas substituições não sinônimas, o que resulta em uma mudança no aminoácido. Essas substituições não sinônimas foram observadas nas proteínas virais: capsídeo, envelope, NS1, NS2A, NS3, NS4A e NS5 (Tabela 7).

Tabela 7: Variações únicas de nucleotídeos detectados pela metodologia de NGS, Variantes não sinônimas estão indicadas com a troca de aminoácido correspondente. Frequências das variantes nos surtos de 2018-2019 estão listadas.

| Posição referênciada (nt) | Proteína | Referência | Varição do alelo | Surto | Frequência | Troca de aminoácido |
|---------------------------|----------|------------|------------------|-----------|------------|---------------------|
| 195 | C | A | A/G | 2019 | (15/39) | K26R |
| 218 | C | A | A/G | 2018 | (4/39) | I34V |
| 297 | C | A | A/G | 2019 | (1/39) | K60R |
| 1021 | E | C | C/T | 2019 | (1/39) | H301Y |
| 1140 | E | C | C/T | 2019 | (1/39) | A341V |
| 1781 | E | A | A/G | 2019 | (1/39) | N555D |
| 1861 | E | C | C/T | 2018/2019 | (36/39) | - |
| 1908 | E | A | A/G | 2018 | (1/39) | D597G |
| 2975 | NS1 | T | T/C | 2018 | (1/39) | Y953H |
| 3097 | NS1 | T | T/C | 2018/2019 | (39/39) | - |
| 3099 | NS1 | T | T/C | 2019 | (1/39) | L994P |

nt: nucleotídeo; (-) sem informação; NGS: sequenciamento de nova geração - do inglês *new generation sequencing*; A: alanina; D: ácido aspártico; G: glicina; H: Histidina; I: isoleucina; K: lisina; L: leucina; M: metionina; N: asparagina; P: prolina; R: arginina; T: treonina; V: valina; W: triptofano

continua

Tabela 7: Variações únicas de nucleotídeos detectados pela metodologia de NGS. Variantes não sinônimas estão indicadas com a troca de aminoácido correspondente. Frequências das variantes nos surtos de 2018-2019 estão listadas.

| Posição referência (nt) | Proteína | Referência | Variação do alelo | Surto | Frequência | Troca de aminoácido |
|-------------------------|----------|------------|-------------------|-----------|------------|---------------------|
| 3317 | NS1 | G | G/A | 2019 | (1/39) | G1067R |
| 3502 | NS1 | A | A/G | 2018/2019 | (39/39) | - |
| 3744 | NS2A | C | C/T | 2019 | (1/39) | A1209V |
| 4948 | NS3 | T | T/C | 2018/2019 | (39/39) | - |
| 4966 | NS3 | C | C/T | 2018/2019 | (39/39) | - |
| 5055 | NS3 | A | A/C | 2018/2019 | (39/39) | N1646T |
| 5231 | NS3 | C | C/T | 2018/2019 | (39/39) | - |
| 5290 | NS3 | T | T/C | 2019 | (24/39) | - |
| 5335 | NS3 | T | T/C | 2018/2019 | (39/39) | - |
| 5362 | NS3 | C | C/T | 2018/2019 | (39/39) | - |
| 5452 | NS3 | G | G/A | 2018/2019 | (39/39) | - |
| 5533 | NS3 | T | T/C | 2018/2019 | (39/39) | - |
| 5595 | NS3 | C | C/T | 2018 | (1/39) | T1826M |
| 5839 | NS3 | C | C/T | 2018/2019 | (39/39) | - |
| 5976 | NS3 | C | C/A | 2018/2019 | (2/39) | P1953H |
| 5989 | NS3 | G | G/A | 2018/2019 | (39/39) | - |
| 6199 | NS3 | T | T/C | 2018/2019 | (39/39) | - |
| 6525 | NS4A | T | T/G | 2018/2019 | (5/39) | V2136G |
| 6526 | NS4A | GCTC | GCTC/TCTT | 2019 | (24/39) | - |
| 6528 | NS4A | T | T/C | 2018 | (1/39) | L2137P |
| 6586 | NS4A | G | G/A | 2018/2019 | (39/39) | - |
| 6623 | NS4A | C | C/T | 2019 | (24/39) | - |

nt: nucleotídeo; (-) sem informação; NGS: sequenciamento de nova geração - do inglês *new generation sequencing*; A: alanina; D: ácido aspártico; G: glicina; H: Histidina; I: isoleucina; K: lisina; L: leucina; M: metionina; N: asparagina; P: prolina; R: arginina; T: treonina; V: valina; W: triptofano

continua

Tabela 7: Variações únicas de nucleotídeos detectados pela metodologia de NGS. Variantes não sinônimas estão indicadas com a troca de aminoácido correspondente. Frequências das variantes nos surtos de 2018-2019 estão listadas.

| Posição referência (nt) | Proteína | Referência | Varição do alelo | Surto | Frequência | Troca de aminoácido |
|-------------------------|----------|------------|------------------|-----------|------------|---------------------|
| 6934 | NS4B | G | G/A | 2018/2019 | (39/39) | - |
| 7231 | NS4B | C | C/T | 2018/2019 | (39/39) | - |
| 7240 | NS4B | A | A/T | 2018/2019 | (39/39) | - |
| 7721 | NS5 | C | C/T | 2018 | (1/39) | R2535W |
| 7976 | NS5 | A | A/G | 2019 | (1/39) | M2620V |
| 8542 | NS5 | G | G/A | 2018/2019 | (39/39) | - |
| 8569 | NS5 | C | C/A | 2018/2019 | (39/39) | - |
| 8638 | NS5 | C | C/T | 2018/2019 | (39/39) | - |
| 8746 | NS5 | G | G/A | 2019 | (24/39) | - |
| 9564 | NS5 | C | C/T | 2018 | (1/39) | A3149V |
| 9775 | NS5 | C | C/T | 2018/2019 | (39/39) | - |
| 9982 | NS5 | T | T/C | 2018/2019 | (39/39) | - |
| 10104 | NS5 | C | C/T | 2018/2019 | (30/39) | T3329I |

nt: nucleotídeo; (-) sem informação; NGS: sequenciamento de nova geração - do inglês *new generation sequencing*; A: alanina; D: ácido aspártico; G: glicina; H: Histidina; I: isoleucina; K: lisina; L: leucina; M: metionina; N: asparagina; P: prolina; R: arginina; T: treonina; V: valina; W: triptofano
conclusão tabela 7.

Comparando as sequências das cepas de YFV dos surtos de 2018 e 2019, pudemos observar que a alteração de nucleotídeo na posição 195 esteve presente exclusivamente nas sequências do surto de 2019 (15/25; 60,0% das sequências). Essa substituição não sinônima resultou em uma mudança conservativa do aminoácido Lisina para Arginina, ou seja, a troca foi entre aminoácidos que apresentam propriedade físico-química similares.

Apenas um único polimorfismo de base única (SNP – do inglês, *single nucleotide polymorphism*) esteve presente em todas as amostras sequenciadas e está localizado na posição 5055(nt) do genoma. Essa substituição não sinônima ocorreu na proteína NS3 e a substituição foi entre famílias de aminoácidos que apresentam propriedades físico-químicas diferentes.

Na posição 10.104 do genoma, 76,9% das sequências (30/39) apresentaram substituição de nucleotídeo não sinônima com mudança na propriedade química da cadeia lateral do aminoácido, dessas, 6 sequências pertencem ao surto de 2018 e 24 ao surto de 2019. Em uma das sequências de 2019 não pudemos avaliar essa posição, pois o genoma montado não possui essa região devido à baixa cobertura de leitura

4.5. Análise Filogenética

Após realizado o alinhamento múltiplo, verificamos o valor do escore de qualidade gerado pela plataforma online Guidance v2 (92) e o valor obtido foi de 0,99, reforçando que o alinhamento era confiável e de ótima qualidade.

Em seguida, com a análise filogenética pudemos observar que todas as sequências são pertencentes ao genótipo Sul-Americano I subgrupo E, estando de acordo com outros estudos realizados durante o surto de 2018 na região Sudeste do Brasil (Figura 11) (94).

A análise da árvore filogenética (Figura 11) nos permite criar uma hipótese de que há uma maior variabilidade genética entre as cepas virais do surto de 2018, do que 2019, já que o clado formado pelas sequências de 2019, possui menos ramos e parece ser bem mais estabelecido, vindo de um único ancestral comum. Não foram encontradas diferenças significativas no agrupamento das sequências dos pacientes que foram a óbito (90) em relação aos vivos. Entretanto, podemos observar que a sequência referência MK583155.1 (90), agrupou-se com o clado das sequências de 2019.

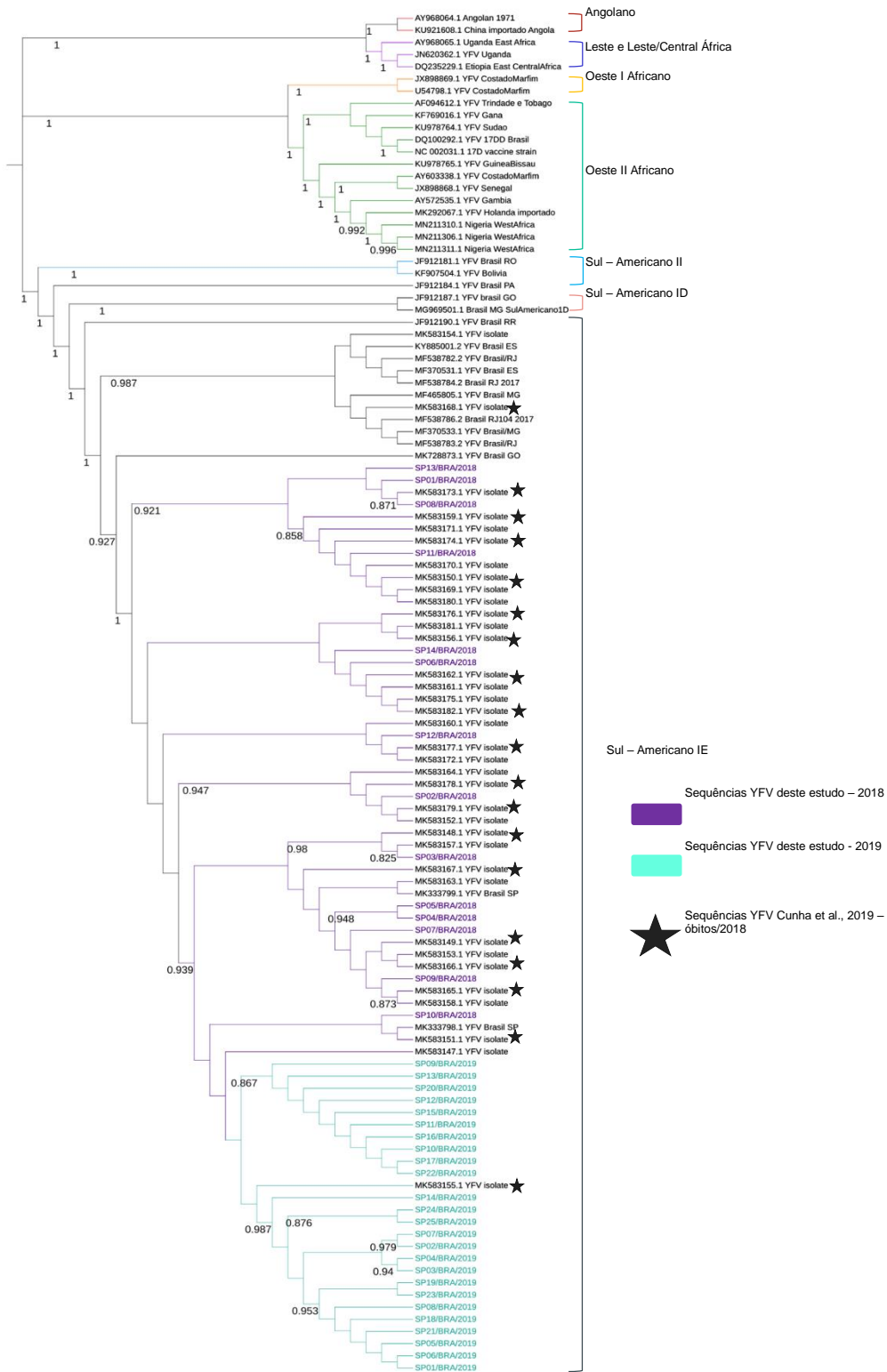


Figura 11: Árvore filogenética realizada por máxima verossimilhança.

Foram utilizadas sequências do genoma completo do YFV disponíveis no NCBI, as sequências dos pacientes que foram a óbito (90) e as sequências geradas no presente estudo.

Foi gerada uma nova árvore filogenética (Figura 12) no intuito de comparar as sequências obtidas neste estudo com as sequências referências: cepa 17D vacinal, RJ104 e SP_ 2018 (NC_002031; MF538786; MK333798, respectivamente).

Nas análises preliminares, as sequências de 2018 agruparam separadamente das sequências de 2019. As sequências de 2018 agruparam em 4 clados principais e o que chamou atenção durante as primeiras análises foram os subgrupos formados pelas sequências SP01/BRA/2018, SP08/BRA/2018, SP11/BRA/2018 e SP13/BRA/2018 que apresentam a mesma variação não sinônima no nucleotídeo 218 (região genômica do capsídeo viral); e as sequências SP03/BRA/2018, SP04/BRA/2018, SP05/BRA/2018, SP07/BRA/2018, SP09/BRA/2018, SP10/BRA/2018 também formaram um subgrupo separado e apresentam a variação não sinônima no nucleotídeo 10.104 (região genômica NS5 do YFV), assim como todas as sequências analisadas até essa região para o ano de 2019 (24/25, com exceção da SP05/BRA/2019 que não foi analisada por limitação de tamanho do genoma durante a montagem).

Em relação às sequências de 2019, dois principais clados foram observados, no agrupamento formado pelas sequências SP25/BRA/2019, SP24/BRA/2019, SP04/BRA/2019, SP03/BRA/2019, SP07/BRA/2019, SP02/BRA/2019, SP14/BRA/2019, SP19/BRA/2019, SP23/BRA/2019, SP05/BRA/2019, SP18/BRA/2019, SP21/BRA/2019, SP08/BRA/2019, SP01/BRA/2019 e SP06/BRA/2019 destaca-se a presença da variante não sinônima no nucleotídeo 195 da proteína do capsídeo que só está presente nas amostras citadas acima.

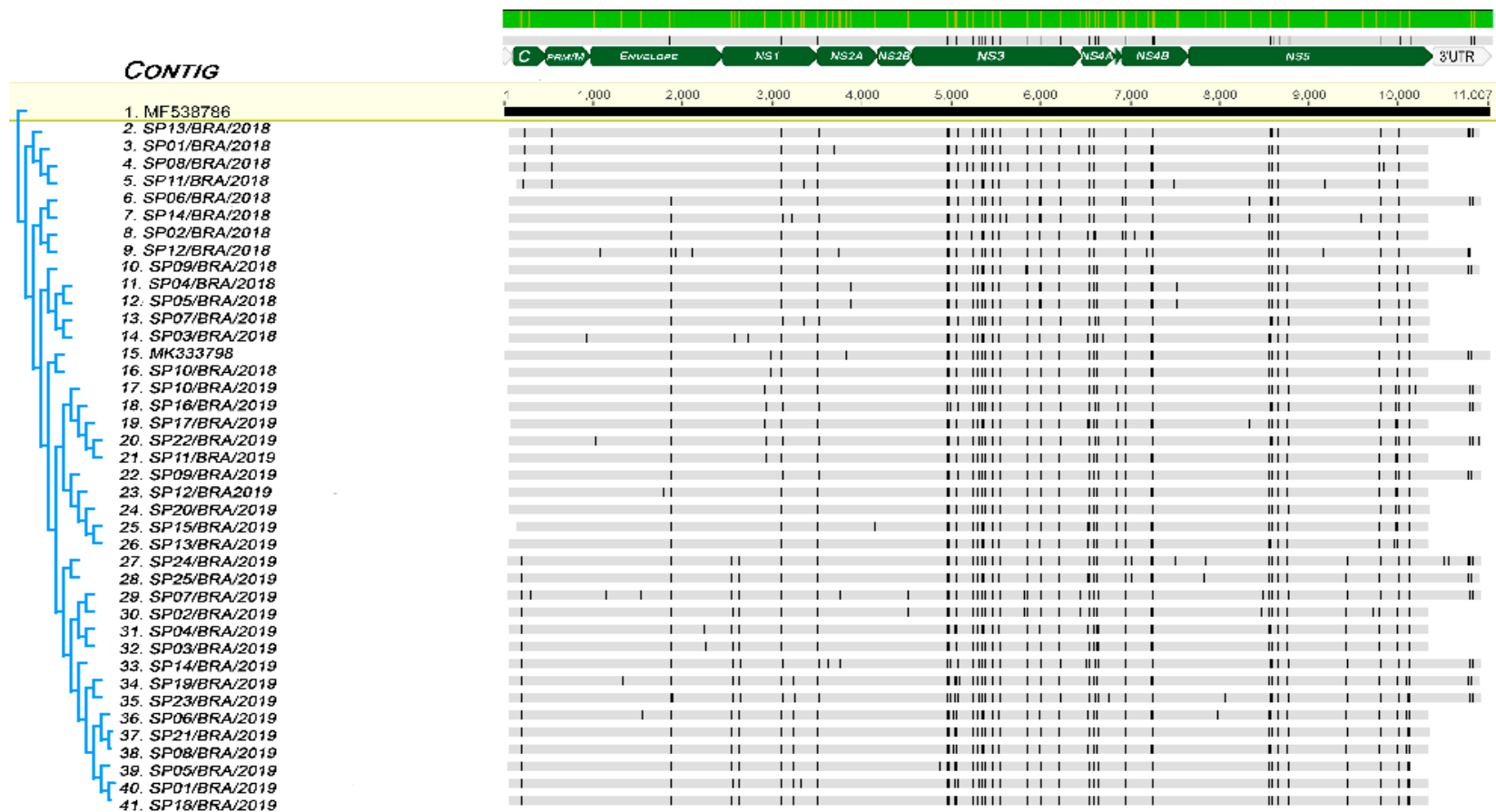


Figura 12: Árvore filogenética realizada por máxima verossimilhança referente às sequências de 2018 e 2019 obtidas neste estudo. Realizada com o intuito de avaliar a diversidade entre os vírus circulantes nos dois surtos e ao lado, podemos observar, de modo geral, todas as variantes encontradas ao longo dos 39 genomas, sejam elas sinônimas ou não sinônimas. Vale ressaltar, que por ser uma visão geral, muitas dessas variantes parecem estar no mesmo sítio nucleotídico, porém, isso ocorreu devido ao ajuste da escala gráfica para que todas as variantes fossem representadas em uma mesma imagem.

5. Discussão

Devido à grande quantidade de infectados mundialmente (aproximadamente 200.000 casos/ano) e por ser uma doença negligenciada que atinge, principalmente, regiões em desenvolvimento e pobres (África, América Central e do Sul), a infecção pelo YFV representa um grave problema de saúde pública (17, 91) e de tempos em tempos, atinge uma população não vacinada ou uma área de baixa cobertura vacinal. Em consequência a falta de medicamentos antivirais eficazes, a única solução é a vacinação em massa e o tratamento da sintomatologia dos infectados que desenvolveram a forma moderada ou grave (91)

O presente estudo analisou 56 amostras (soro ou urina) que apresentaram resultado positivo para o YFV no teste molecular RT-qPCR, essas as amostras foram coletadas de pacientes atendidos no HCFMUSP, considerado o hospital referência para tratamento de casos graves de febre amarela durante os surtos de 2018 e 2019; portanto, todos os casos deste trabalho são graves (4), e dentre eles, foram escolhidos apenas os que sobreviveram à infecção.

Dos 56 pacientes incluídos neste estudo, 85,7% é pertencente ao sexo masculino e este dado está de acordo com o monitoramento da doença realizado pelo ministério da saúde, relatando uma maior porcentagem entre os homens, pois é considerado reflexo das atividades laborais realizadas por eles em áreas de mata ou próximas a elas (3).

Entretanto, quando comparados os dados dos pacientes que sobreviveram a infecção em relação aos que evoluíram a óbito, referentes a sexo, idade, dias após início de sintomas, carga viral e média dos Cts dos pacientes separados pelo ano do surto, não foram encontrados resultados significativos. Estudos anteriores realizados por Casadio e colaboradores, também demonstrou não encontrar diferenças significativas nesses aspectos em questão, porém encontrou diferenças significativas em relação a fatores laboratoriais como o fator V de coagulação e aos níveis de lipase, considerando-os fatores importantes durante a evolução da doença severa de YFV (55).

Apesar disso, estudos realizados por Kallas, EG e colaboradores, analisando uma casuística diferente, encontrou diferenças estatisticamente significantes na

análise das variáveis: carga viral, idade, contagem de neutrófilos e a enzima AST, e concluíram que esses fatores, independentemente, são associados com o desfecho óbito (92).

Ao mapear os municípios de infecção dos pacientes, observa-se que o local de moradia relatado por eles, está de acordo com o que foi descrito nos boletins epidemiológicos do ministério da saúde para o respectivo ano do surto. Durante o pico do surto do ano de 2018, o boletim epidemiológico liberado no dia 28/12/2018, mapeou os casos e identificou uma maior concentração na região da grande São Paulo, com enfoque nos municípios de Mairiporã e Atibaia que foram considerados os responsáveis pela grande maioria dos casos (71) e aos poucos foi se espalhando para o litoral e mais ao final do ano, chegando ao Vale do Ribeira (72).

Portanto, ao realizar uma comparação dos boletins (71, 72) com o mapeamento realizado para este trabalho, podemos observar uma maior concentração dos casos durante o ano de 2018 na grande São Paulo e um caso no litoral. E ao analisarmos o surto do ano de 2019, percebemos uma concentração quase absoluta na região do Vale do Ribeira, com exceção de um caso que relatou moradia em Mairiporã.

Ao compararmos a taxa de mortalidade entre surtos de YFV, foi relatado que no Brasil, durante o surto ocorrido nos anos de 2008 e 2009, a taxa de mortalidade foi de 41,2% (93%). Em relação ao surto atual de YFV, destacam-se os casos ocorridos na Angola em 2016 que teve uma taxa de mortalidade de 13,7% e a República Democrática do Congo que atingiu uma taxa de cerca de 21,1% de óbitos segundo dados gerados pela OMS (94). Entretanto, os dados recentes referente a mortalidade para o Brasil no ano de 2018, indicam uma taxa de 35,1% e em 2019 de cerca de 17,1% (1).

Mais especificamente em São Paulo, durante o surto de 2018, a taxa de mortalidade foi de 34,2% e em 2019, 18,2%, porcentagens bem altas e que nos mostram o quão grave podem ser os surtos de YFV e o quão necessária é a intervenção vacinal, o monitoramento da doença e a busca por tratamentos mais específicos, como os antivirais (71, 72).

Em relação à amplificação do genoma completo, foi possível visualizar os fragmentos virais no gel de agarose após a reação de PCR convencional, e foi possível determinar quais alvos eram condizentes com os tamanhos esperados em relação aos *primers* utilizados. O genoma foi dividido em 11 fragmentos e por isso

obtivemos 11 *amplicons* de tamanhos variáveis entre 850-1200pb para cada amostra. Como esperado, a região 3'UTR foi a mais complicada para amplificarmos, devido a sua alta complexidade estrutural (30) e como alternativa, foi usado uma combinação de *primers* que já foi descrita anteriormente (74) e eficiente para cobrir o final da proteína NS5 do vírus, mas não a 3'UTR.

Embora tenha sido realizada a PCR, tanto em amostras de RNA extraídas de soro quanto de urina, não foram encontradas diferenças em relação a melhor qualidade e eficiência entre os dois materiais utilizados durante a PCR. O único fator que levou a decisão da escolha do material a ser utilizado foi o menor Ct entre eles, ou seja, a maior carga viral.

A comparação das 39 sequências virais que foram obtidas foi realizada e durante as análises foram encontradas substituições nucleotídicas sinônimas e não sinônimas. As substituições não sinônimas, levam a troca de aminoácido, e essa troca pode ocorrer por um aminoácido de mesma propriedade bioquímica ou não.

Nove dessas substituições não sinônimas causam uma alteração da propriedade bioquímica do aminoácido. Interessantemente, a maioria dessas mudanças ocorrem nas duas proteínas virais mais importantes a NS3 que foi descrita como uma serino protease (95), RNA helicase/NTPase (22), responsável principalmente pela clivagem da poliproteína e a NS5, a maior e mais conservada proteína viral entre os *Flavivirus* e considerada a chave para replicação viral em consequência da sua atividade de polimerase (96).

Alterações de aminoácido nessas proteínas conservadas, podem ter um impacto na capacidade de infectividade viral, tanto em humanos e PNHs, quanto em mosquitos, o que pode levar a uma aceleração na dispersão do vírus entre hospedeiros (45).

Durante as análises das sequências deste estudo, foi encontrada, uma mutação não sinônima na posição nucleotídica 5055 (N1646T) na proteína NS3, que esteve presente em todas as sequências dos surtos de 2018 e 2019, podendo ser considerada característica dessa população viral e possivelmente, a assinatura da cepa viral responsável por esses surtos. Essa variante também foi descrita em um estudo realizado com os pacientes que foram a óbito em São Paulo, durante o surto de 2018 (90).

Quando comparamos as sequências do surto de 2018 com as de 2019, pudemos observar que na proteína do capsídeo, existe uma mutação não sinônima entre aminoácidos de mesma propriedade bioquímica no nucleotídeo 195 (K26R), que ocorre em 15/25 sequências no surto de 2019, entretanto não foi encontrada em nenhuma sequência de 2018. Estudos associam a proteína do capsídeo com o empacotamento do genoma viral e formação do nucleocapsídeo (97), porém para saber se essa alteração poderia ter algum efeito na patogênese viral, seriam necessários estudos funcionais que permitissem a avaliação (*in silico*) do real impacto dessa alteração.

Além disso, pudemos observar uma mutação não sinônima (T3329I) na proteína NS5 que está presente em apenas 6/14 sequências referente ao surto de 2018, mas ao analisarmos as sequências de 2019 levando a uma troca na propriedade bioquímica do aminoácido, polar para hidrofóbico. É importante dizer que cadeias laterais de aminoácidos polares possuem cargas positivas e negativas e são atraídos por água, enquanto que aminoácidos hidrofóbicos tendem a evitar água.

Vale ressaltar que a NS5 é uma proteína que possui um papel muito importante na replicação viral, possuindo a função de RNA polimerase RNA dependente. Mutações nesta proteína têm sido associadas com diferenças na replicação viral, resposta imune e interações entre as proteínas do genoma viral durante a replicação (98), além de ser uma estrutura muito conservada entre os flavivírus e um potencial alvo de drogas terapêuticas (99). Entretanto, assim como para todas as outras proteínas, afirmações sobre o impacto dessas alterações só podem ser realizadas após estudos *in silico*.

Para complementar, foram realizadas análises filogenéticas para identificarmos o genótipo da cepa viral circulante em São Paulo e para isso utilizamos diferentes sequências do genoma completo dos 7 diferentes genótipos descritos para YFV e adicionamos todas as sequências disponíveis dos pacientes atendidos no HCFMUSP, incluindo pacientes graves que sobreviveram à infecção (sequências geradas no presente estudo) e pacientes que evoluíram à óbito (sequências disponíveis no NCBI) (90).

Analisando a árvore filogenética da figura 11 pudemos inferir que todas as sequências se agruparam com as referências já anotadas como genótipo Sul

Americano I, subgrupo E. Esse resultado foi observado em outros trabalhos que sequenciaram o vírus, mesmo em outros estados Brasileiros (90, 100).

Conjuntamente, pudemos observar que não houve diferenças significativas no agrupamento das sequências dos pacientes que evoluíram a óbito (90) em relação aos que sobreviveram, e isso indica que não há uma variação genética viral alta entre os desfechos clínicos. Importante ressaltar que não possuímos nenhuma sequência com os dados brutos e/ou informação complementar dos pacientes que foram a óbito durante o ano de 2019, sendo assim, não foi possível realizar uma análise profunda dessas sequências.

É possível supor que a sequência referência MK728873.1 é o ancestral comum para as sequências provenientes de 2018, com um valor de confiança de 100%, e isso corrobora os estudos de Dellatorre, E e colaboradores (100), que realizaram análises com cepas virais de todo o sudeste Brasileiro e fizeram uma suposição que essa cepa viral, proveniente de Goiás, seja uma das portas de entrada do vírus durante o surto de 2018 e, seguindo as análises, observamos que a partir desse clado outros ramos foram formados, o que pode indicar uma pequena variabilidade entre uma mesma cepa viral e diferentes portas de entrada, principalmente na região da grande São Paulo, como foi descrito anteriormente por Cunha, M. P. e colaboradores (90).

Entretanto, ao analisar o clado formado pelas sequências de 2019, pudemos observar que temos um clado bem estabelecido, formando dois ramos com valores de confiança acima de 70, ou seja, possuem uma alta confiabilidade e provavelmente uma porta de entrada para o vírus, provavelmente oriundos das cidades previamente afetadas pelo YFV em 2018.

Interessante notar que a sequência identificada como MK583155.1 de janeiro de 2018, proveniente de um paciente que evoluiu a óbito (90), foi a única a agrupar-se fora do clado representante do surto de 2018, agrupando-se com o clado das sequências do surto de 2019, porém como não temos mais informações sobre a origem da sequência e nem do paciente, não há como fazer inferências mais significativas do porquê desse agrupamento distante do restante das sequências de 2018.

Ao analisarmos apenas as sequências referente a este estudo, ou seja, os pacientes que sobreviveram à infecção por YFV nos anos de 2018 e 2019, podemos observar alguma variabilidade entre elas. Foi possível observar a formação de clados

específicos de acordo com variantes não sinônimas presentes nas mesmas. Isso indica que os vírus são geneticamente diferentes entre si, mas não podemos afirmar se essas diferenças são o suficiente para levar a uma modificação estrutural de proteína, ou até mesmo se impactam no *fitness* viral e com isso explicar uma diferença na taxa de mortalidade entre os dois anos de surtos.

Entretanto, as análises realizadas neste estudo, permitem um melhor entendimento do genoma do YFV que esteve circulante em São Paulo nesses dois anos e os dados aqui gerados poderão contribuir com futuros estudos virológicos e epidemiológicos do vírus.

6. Conclusões

- ✓ Não houve diferenças estaticamente significantes para idade, dias após início de sintomas, média de carga viral e Ct, quando comparados entre os anos de 2018 e 2019;
- ✓ A grande maioria dos pacientes (85,7%) durante os dois surtos é pertencente ao sexo masculino;
- ✓ Dos pacientes pertencentes ao surto de 2018, quase todos ocorreram na grande São Paulo, com exceção de um paciente que era do litoral;
- ✓ Dos pacientes pertencentes ao surto de 2019, quase todos ocorreram no Vale do Ribeira, com exceção de um paciente que era de Mairiporã;
- ✓ Amostras com valor médio de Ct<30 apresentaram maior sucesso na amplificação e conseqüentemente no sequenciamento do genoma do YFV;
- ✓ Todas as cepas circulantes no surto de 2018 e 2019 foram classificadas como genótipo Sul Americano I (subgrupo 1E);
- ✓ As análises de filogenia e de variantes sugerem que os vírus circulantes em 2018 e 2019 são distintos, porém com pequenas variações entre eles;
- ✓ Para entender se as variantes encontradas possuem algum impacto sobre a patogênese viral, são necessárias análises *in silico*.

7. ANEXOS

ANEXO A – Termo de Consentimento Livre e Esclarecido (TCLE).

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

DADOS SOBRE A PESQUISA

TÍTULO DA PESQUISA: Caracterização das Hepatites Agudas em centros de referência da cidade de São Paulo

PESQUISADOR PRINCIPAL: Michele Soares Gomes Gouvêa

DEPARTAMENTO/INSTITUTO: Gastroenterologia/Instituto de Medicina Tropical da USP

Convidamos você para participar da pesquisa: “Caracterização das Hepatites Agudas em centros de referência da cidade de São Paulo”.

A hepatite é uma inflamação do fígado que pode ser provocada por diversas causas incluindo vírus, medicamentos, doenças autoimunes, dentre outras. As hepatites agudas de causa viral e medicamentosa são as mais frequentes. Dentre os vírus causadores de hepatite destacam-se aqueles que tem preferência pela infecção das células do fígado que são: vírus da hepatite A (HAV), vírus da hepatite B (HBV), vírus da hepatite C (HCV), vírus da hepatite Delta (HDV) e vírus da hepatite E (HEV).

Esta pesquisa tem por objetivo o mapeamento das causas de hepatite aguda entre os pacientes atendidos com esse quadro clínico em centros de referência da cidade de São Paulo. Além disso, nos casos de hepatite aguda causada por algum dos vírus citados acima o genoma do vírus detectado será analisado. Os resultados dessa análise nos permitirão sugerir a origem do vírus e nos ajudarão a entender como esses vírus estão se espalhando em nosso meio, o que nos ajudará a propor estratégias mais eficientes para combater o espalhamento dessas infecções. Em alguns pacientes a causa da hepatite aguda não é identificada ou é associada a uso de algum medicamento específico, nesses casos iremos aplicar métodos modernos de biologia molecular (estudo do viroma por sequenciamento de nova geração) que permitem a identificação de todos os possíveis vírus existentes na amostra analisada e dessa forma será possível associar ou descartar a participação de algum vírus no processo de inflamação aguda do fígado.

Pretendemos também identificar características genéticas e imunológicas individuais para avaliar possíveis associações com a evolução do quadro clínico.

Procedimentos que serão realizados e métodos que serão empregados: se você aceitar participar deste estudo, além dos exames de rotina comumente realizados para diagnóstico na instituição que você está sendo atendido, serão coletados 3 tubos de sangue (1 seco e 2 com EDTA) contendo cerca de 4 ml de sangue cada um e uma amostra de fezes (5 ml ou 5 gramas). A coleta de sangue será

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

continua

realizada por punção periférica da veia do antebraço e a coleta de fezes será realizada por você mesmo em frasco coletor de boca larga e tampa rosqueável.

Explicitação de possíveis desconfortos e riscos decorrentes da participação na pesquisa: a coleta de sangue não envolve risco importante para a sua saúde, no entanto, pode ocasionar desconforto em função da picada da agulha, que é levemente dolorosa. Também pode ocorrer discreto hematoma por extravasamento de sangue sob a pele; caso isto venha a ocorrer, esse problema desaparece em poucos dias sem nenhum cuidado especial.

Benefícios esperados para o participante: o benefício direto para você será a possibilidade de ter um diagnóstico diferencial que envolve a realização de exames que não são realizados na rotina pela instituição em que você está sendo atendido, como o exames para diagnóstico de infecção pelo vírus da hepatite E ou pelo vírus da hepatite Delta (nos casos com infecção pelo vírus da hepatite B). Se o seu caso for um daqueles em que não é possível identificar a causa da hepatite com os exames disponíveis para isso, você poderá contribuir para esclarecermos se há outros vírus, além dos conhecidos, envolvidos na causa de hepatites agudas.

No caso de cronificação de infecção por vírus das hepatites B ou C os exames de caracterização dos genes dos vírus permitirão a escolha de tratamento mais adequado, o que não é realizado rotineiramente na instituição em que você está sendo atendido.

Esclarecimento sobre a forma de acompanhamento e assistência a que terão direito os participantes da pesquisa: em qualquer etapa do estudo, você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. A principal investigadora é o Dra. MICHELE SOARES GOMES GOUVÊA, que pode ser encontrada no endereço Av. Dr. Enéas Carvalho de Aguiar n° 500 prédio 2, 2° andar Telefone(s) (11) 3061-8218. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) – Rua Ovídio Pires de Campos, 225 – 5º andar – tel: (11) 2661-7585, (11) 2661-1548, (11) 2661-1549; e-mail: cappesq.adm@hc.fm.usp.br. Você receberá uma via deste documento para poder consultar posteriormente as informações sobre o projeto e o telefone para contato.

É garantida a liberdade da retirada de consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade de seu tratamento na Instituição.

Direito de confidencialidade: as informações obtidas serão analisadas em conjunto com outros pacientes, não sendo divulgada a identificação de nenhum paciente.

Direito de ser mantido atualizado sobre os resultados parciais das pesquisas: você tem direito de ser mantido atualizado sobre os resultados parciais deste estudo assim que os pesquisadores tiverem conhecimento desses resultados.

Despesas e compensações: não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação.

continua

Destino do material coletado: após as análises propostas nesta pesquisa as amostras coletadas poderão ser armazenadas para possíveis pesquisas futuras. Se você autorizar esse armazenamento o material coletado será armazenado no laboratório de Gastroenterologia e Hepatologia Tropical – FMUSP sob responsabilidade da Dra. Michele Soares Gomes Gouvêa e caso seja útil para pesquisas futuras as amostras somente serão utilizadas após aprovação do novo projeto pelo comitê de ética e de forma anônima. Ressaltamos que o não consentimento para o armazenamento desse material não implica em exclusão do voluntário do estudo.

Autorizo o uso do material armazenado para o desenvolvimento de projetos futuros desde que seja feito de maneira anônima.

() SIM

() NÃO

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo " Caracterização das Hepatites Agudas em centros de referência da cidade de São Paulo". Eu discuti com o (a) Dr(a) _____ sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e estou ciente que poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste Serviço.

Assinatura do paciente/representante legal Data ____ / ____ / ____

Assinatura da testemunha Data ____ / ____ / ____

para casos de pacientes menores de 18 anos, analfabetos, semi-analfabetos ou portadores de deficiência auditiva ou visual.

(Somente para o responsável do projeto)

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste paciente ou representante legal para a participação neste estudo.

Rubrica do sujeito de pesquisa ou responsável _____

Rubrica do pesquisador _____

continua

Assinatura do responsável pelo estudo Data ____ / ____ / ____

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME:

DOCUMENTO DE IDENTIDADE Nº: SEXO: M F

DATA NASCIMENTO:/...../.....

ENDEREÇO Nº APTO:

BAIRRO: CIDADE:

CEP: TELEFONE: DDD (.....)

2. RESPONSÁVEL LEGAL

NATUREZA (grau de parentesco, tutor, curador etc.)

DOCUMENTO DE IDENTIDADE: SEXO: M F

DATA NASCIMENTO:/...../.....

ENDEREÇO: Nº APTO:

BAIRRO: CIDADE:

CEP: TELEFONE: DDD (.....)

conclusão anexo A.

ANEXO B – Parecer Consubstanciado do CEP.



USP - HOSPITAL DAS
CLÍNICAS DA FACULDADE DE
MEDICINA DA UNIVERSIDADE



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: Caracterização das Hepatites Agudas em centros de referência da cidade de São Paulo

Pesquisador: Michele Soares Gomes Gouvêa

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;);

Versão: 3

CAAE: 74535417.3.1001.0068

Instituição Proponente: Hospital das Clínicas da Faculdade de Medicina da USP

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.553.999

Apresentação do Projeto:

O diagnóstico de hepatite aguda de etiologia medicamentosa será definido pelo histórico de consumo ou exposição a determinada droga e pela exclusão de outras causas conforme critérios previamente estabelecidos (Benichou, 1990; Danan, 1993). Coleta de dados Durante o primeiro atendimento serão colhidos, em questionário específico (Anexo 2), dados demográficos, epidemiológicos e clínicos. Será colhida amostra de sangue para análise laboratorial, incluindo sorologias para HEV e HDV, detecção, quantificação e caracterização (genotipagem) dos respectivos genomas virais nos casos com evidência sorológica de infecção por de HAV, HBV, HCV, HDV ou HEV. Nessa amostra será também realizada a pesquisa de marcadores genéticos e imunológicos dos pacientes para estudo de associação com a evolução clínica da infecção viral e também para avaliação de fatores de risco genéticos associados a DILI. Será também colhida amostra de fezes para pesquisa de HAV e HEV por PCR (reação em cadeia da polimerase). As amostras de plasma e fezes também serão utilizadas para realização de estudo da metagenômica viral nos casos definidos como hepatite aguda de etiologia criptogênica ou com suspeita de DILI. Após a coleta as amostras serão adequadamente processadas e armazenadas e posteriormente encaminhadas para o Laboratório de Gastroenterologia e Hepatologia Tropical para análise. Haverá um seguimento agendado com os

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar

CEP: 05.403-010

UF: SP

Município: SAO PAULO

Telefone: (11)2661-7585

Fax: (11)2661-7585

E-mail: cappesq.adm@hc.fm.usp.br

Página 01 de 04

continua



Continuação do Parecer: 2.553.999

pacientes para conhecimento dos resultados dos exames laboratoriais, definir o desfecho do quadro agudo e conclusão do preenchimento da ficha de dados. O prontuário dos pacientes será consultado para obtenção dos resultados referentes à investigação da etiologia do quadro de hepatite aguda, resultados bioquímicos e dados clínicos relacionados à evolução do quadro.

Objetivo da Pesquisa:

Objetivo Primário: Caracterizar a etiologia dos casos de hepatites agudas atendidos em instituições de referência na cidade de São Paulo. Objetivo Secundário: • Avaliar as formas de transmissão dos diferentes agentes etiológicos infecciosos envolvidos; • Caracterizar os genótipos dos vírus hepatotrópicos identificados nos casos de hepatite aguda por esses agentes; • Entre os casos suspeitos de DILI e aqueles sem etiologia definida, avaliar o envolvimento de outros agentes virais ou infecção oculta por vírus hepatotrópicos através da análise de metagenômica viral utilizando sequenciamento de nova geração; • Descrever os principais agentes (fármacos, xenobióticos, ervas medicinais, suplementos dietéticos) e avaliar os fatores de risco genéticos e não genéticos envolvidos nos casos de lesão hepática aguda induzida por drogas; • Analisar as possíveis associações dos diferentes desfechos clínicos a variáveis demográficas, etiológicas, genéticas e imunológicas; • Comparar a etiologia e os diferentes desfechos clínicos observados entre pacientes coinfectados ou não pelo HIV;

Avaliação dos Riscos e Benefícios:

Riscos:

O desenvolvimento deste projeto não oferece riscos para os participantes. Para os profissionais envolvidos no desenvolvimento do projeto os riscos são aqueles inerentes às suas atividades de rotina, sejam elas clínicas no atendimento dos pacientes, sejam elas laboratoriais na análise das amostras para diagnóstico. Para eliminar ou minimizar esses riscos as regras de biossegurança serão estritamente seguidas.

Benefícios:

Com este projeto esperamos contribuir para gerar conhecimento mais detalhado acerca da etiologia das hepatites agudas atendidas em instituições de referência na cidade de São Paulo. Nos casos de etiologia por vírus hepatotrópicos a caracterização dos genótipos virais envolvidos e as análises filogenéticas poderão contribuir para inferências sobre origem da cepa viral, o que será de grande utilidade para que possamos ter uma visão mais concreta acerca da disseminação desses vírus na cidade de São Paulo, e para promoção de melhorias nas estratégias de prevenção e controles dessas infecções. Além disso, o estudo da metagenômica viral contribuirá para que possamos ter

Endereço: Rua Ovídio Pires de Campos, 225 5º andar
Bairro: Cerqueira Cesar CEP: 05.403-010
UF: SP Município: SAO PAULO
Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappesq.adm@hc.fm.usp.br



Continuação do Parecer: 2.553.999

um diagnóstico mais preciso nos casos de hepatite aguda criptogênica e também nos casos suspeitos de lesão hepática causada por drogas (DILI - do inglês: "Drug Induced Liver Disease"). Para os pacientes participantes da pesquisa o benefício direto será a possibilidade de ter um diagnóstico diferencial que envolve a realização de exames que não são realizados rotineiramente pelas instituições envolvidas, como os exames para diagnóstico de infecção pelo vírus da hepatite E ou pelo vírus da hepatite Delta. No caso de cronificação de infecção por vírus das hepatites B ou C os exames de caracterização dos genes dos vírus permitirão a escolha de tratamento mais adequado, o que também não é realizado rotineiramente nas instituições envolvidas.

Comentários e Considerações sobre a Pesquisa:

Os dois subprojetos são bastante relevantes e demonstram elementos necessários a sua execução.

Considerações sobre os Termos de apresentação obrigatória:

Os termos de apresentação obrigatória estão claros e bem descritos.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

Os dois braços da pesquisa propostos estão aprovados e terão como executantes as alunas: Samira Chuffi e Ana Paula Moreira Salles, que desenvolverão esses braços da pesquisa como projetos de pós graduação (Mestrado) no programa de Ciências em Gastroenterologia do departamento de Gastroenterologia da FMUSP sob orientação do Dr. João Renato Rebello Pinho. Os títulos propostos para esses projetos são os seguintes: 1-"Estudo molecular e filogenético dos casos de Hepatite A circulantes na cidade de São Paulo" Executante: Samira Chuffi 2-"Caracterização molecular do vírus da Febre Amarela circulante em São Paulo no surto de 2017-2018" Executante: Ana Paula Moreira Salles. A inclusão das referidas alunas como executantes da pesquisa está aprovada.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|--------------------------------|---------------------------------------|------------------------|-------|----------|
| Informações Básicas do Projeto | PB_INFORMAÇÕES_BÁSICAS_1094780_E2.pdf | 15/03/2018 11:15:51 | | Aceito |

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar

CEP: 05.403-010

UF: SP

Município: SAO PAULO

Telefone: (11)2661-7585

Fax: (11)2661-7585

E-mail: cappesq.adm@hc.fm.usp.br



Continuação do Parecer: 2.553.999

| | | | | |
|---|---|------------------------|--------------------------------|--------|
| Outros | Projeto_Mestrado_Ana_Paula_Febre_Amarela.pdf | 15/03/2018 11:13:56 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Projeto_HAV_Mestrado_Samira.pdf | 15/03/2018 11:12:36 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Carta_emenda_2_CAAE_74535417_3_1001_0068.pdf | 15/03/2018 11:11:48 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Carta_emenda_CAAE_74535417_3_0000_0068.pdf | 10/10/2017 15:24:08 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Cadastronline_Michele.PDF | 25/08/2017 03:26:28 | Michele Soares Gomes Gouvêa | Aceito |
| Folha de Rosto | Folha_de_rosto_plataformaBrasilassinada_Michele.PDF | 25/08/2017 03:25:41 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Carta_de_anuencia_Dr_Joao_Renato_I_MT.pdf | 17/08/2017 15:38:40 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Termo_de_Anuencia_Dra_Cassia_Mendes_Dep_MI.pdf | 17/08/2017 15:38:12 | Michele Soares Gomes Gouvêa | Aceito |
| Outros | Questionario_hepatites_Agudas.pdf | 17/08/2017 15:37:22 | Michele Soares Gomes Gouvêa | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | Anexo_1_TCLE_Hepatites_Agudas_FINALcomcorrecoes.pdf | 17/08/2017 15:36:58 | Michele Soares Gomes Gouvêa | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_Hepatites_Agudas_v20170714.pdf | 17/08/2017 15:36:21 | Michele Soares Gomes Gouvêa | Aceito |

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO PAULO, 21 de Março de 2018

Assinado por:
ALFREDO JOSE MANSUR
(Coordenador)

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar CEP: 05.403-010

UF: SP Município: SAO PAULO

Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappesq.adm@hc.fm.usp.br

8. REFERÊNCIAS BIBLIOGRÁFICAS

1. Saúde Md. Febre Amarela: sintomas, tratamento, diagnóstico e prevenção: Ministério da Saúde; 2018 [Available from: [http://www.saude.gov.br/saude-de-a-z/febre-amarela-sintomas-transmissao-e-prevencao.](http://www.saude.gov.br/saude-de-a-z/febre-amarela-sintomas-transmissao-e-prevencao)] (Acessado em agosto/2018)
2. OPAS/PAHO. Intensificação da vigilância de febre amarela durante o período sazonal OPAS/PAHO2013 [Available from: [https://www.paho.org/bra/index.php?option=com_content&view=article&id=3414:intensificacao-da-vigilancia-da-febre-amarela-durante-o-periodo-sazonal-3&Itemid=812.](https://www.paho.org/bra/index.php?option=com_content&view=article&id=3414:intensificacao-da-vigilancia-da-febre-amarela-durante-o-periodo-sazonal-3&Itemid=812)] (Acessado em agosto/2018)
3. Vasconcelos PF. [Yellow Fever]. Rev Soc Bras Med Trop. 2003;36(2):275-93.
4. Ho YL, Joelsons D, Leite GFC, Malbouisson LMS, Song ATW, Perondi B, et al. Severe yellow fever in Brazil: clinical characteristics and management. J Travel Med. 2019;26(5).
5. Chippaux JP, Chippaux A. Yellow fever in Africa and the Americas: a historical and epidemiological perspective. J Venom Anim Toxins Incl Trop Dis. 2018;24:20.
6. Clements AN, Harbach RE. History of the discovery of the mode of transmission of yellow fever virus. J Vector Ecol. 2017;42(2):208-22.
7. Costa ZGA, Elkhoury ANM, Romano APM, Flannery B. Evolução histórica da vigilância epidemiológica e do controle da Febre Amarela no Brasil. Rev Pan Amazônica de Saúde. 2011:11-26.
8. Chaves-Carballo E. Carlos Finlay and yellow fever: triumph over adversity. Mil Med. 2005;170(10):881-5.
9. Philadelphia TCoPo. The history of vaccines 2016 [Available from: <https://www.historyofvaccines.org/timeline/yellow-fever>] (Acessado em agosto/2018).
10. SOPER FL, PENNA H, CARDOSO E, SERAFIM JJ, FROBISHER MJ, PINHEIRO J. Yellow Fever without *Aedes Aegypti*. Study of a rural epidemic in the Valle Do Chanaan, Espírito Santo, Brazil, 1932. *American Journal of Epidemiology*. 1933;18(3):555–87.
11. Stokes A, Bauer JH, Hudson NP. The transmission of yellow fever to *Macacus rhesus*. 1928. Rev Med Virol. 2001;11(3):141-8.

12. Tan SY, Pettigrew K. Max Theiler (1899-1972): Creator of the yellow fever vaccine. *Singapore Med J.* 2017;58(4):223-4.
13. Camacho LA, Freire MaS, Leal MaL, Aguiar SG, Nascimento JP, Iguchi T, et al. Immunogenicity of WHO-17D and Brazilian 17DD yellow fever vaccines: a randomized trial. *Rev Saude Publica.* 2004;38(5):671-8.
14. Litvoc MN, Novaes CTG, Lopes M. Yellow fever. *Revista Da Associacao Medica Brasileira.* 2018;64(2):106-13.
15. WHO. P&R: Surto de febre amarela em Angola e República Democrática do Congo: WHO; 2016 [Available from: <https://www.who.int/emergencies/yellow-fever/mediacentre/qa/pt/>] (Acessado em agosto/2018).
16. Army NMU. A great army deserves a great great museum: Major Walter Reed and the eradication of Yellow Fever 2015 [Available from: <https://armyhistory.org/major-walter-reed-and-the-eradication-of-yellow-fever/>] (Acessado em agosto/2018).
17. WHO. Yellow Fever WHO2019 [Available from: <https://www.who.int/news-room/fact-sheets/detail/yellow-fever>] (Acessado em agosto/2018).
18. Possas C, Lourenço-de-Oliveira R, Tauil PL, Pinheiro FP, Pissinatti A, Cunha RVD, et al. Yellow fever outbreak in Brazil: the puzzle of rapid viral spread and challenges for immunisation. *Mem Inst Oswaldo Cruz.* 2018;113(10):e180278.
19. Journalists WFoS. Yellow Fever - Transmission cycle 2017 [
20. Frierson JG. The yellow fever vaccine: a history. *Yale J Biol Med.* 2010;83(2):77-85.
21. Wachtman L, Mansfield K. Flavivirus - Yellow Fever. *Viral Diseases of Nonhuman Primates. Elsevier Public Health Emergency Collection* 2012. p. 104.
22. Brand C, Bisailon M, Geiss BJ. Organization of the Flavivirus RNA replicase complex. *Wiley Interdiscip Rev RNA.* 2017;8(6).
23. Gardner CL, Ryman KD. Yellow Fever: A Reemerging Threat. *Clinics in Laboratory Medicine.* 2010;30(1):237-+.
24. Simmonds P, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. ICTV Virus Taxonomy Profile: Flaviviridae. *J Gen Virol.* 2017;98(1):2-3.
25. Douam F, Ploss A. Yellow Fever Virus: Knowledge Gaps Impeding the Fight Against an Old Foe. *Trends Microbiol.* 2018;26(11):913-28.

26. von Lindern JJ, Aroner S, Barrett ND, Wicker JA, Davis CT, Barrett AD. Genome analysis and phylogenetic relationships between east, central and west African isolates of Yellow fever virus. *J Gen Virol*. 2006;87(Pt 4):895-907.
27. Rice C, Lenches E, Eddy S, Shin S, Sheets R, Strauss J. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science*. 1985;229(0036-8075):726-33.
28. Ng WC, Soto-Acosta R, Bradrick SS, Garcia-Blanco MA, Ooi EE. The 5' and 3' Untranslated Regions of the Flaviviral Genome. *Viruses*. 2017;9(6).
29. Picard-Jean F, Tremblay-Létourneau M, Serra E, Dimech C, Schulz H, Anselin M, et al. RNA 5'-end Maturation: A Crucial Step in the Replication of Viral Genomes. In: Romanowski V, editor. *Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications* 2013. p. 29.
30. Brinton MA, Fernandez AV, Dispoto JH. The 3'-nucleotides of flavivirus genomic RNA form a conserved secondary structure. *Virology*. 1986;153(1):113-21.
31. Lindenbach B, Thiel HJ, Rice CM. Flaviviridae: The Viruses and Their Replication. *Fields Virology*. 2007:1101-51.
32. Chambers TJ, Hahn CS, Galler R, Rice CM. FLAVIVIRUS GENOME ORGANIZATION, EXPRESSION, AND REPLICATION. *Annual Review of Microbiology*. 1990;44:649-88.
33. Droll DA, Krishna Murthy Hm Fau - Chambers TJ, Chambers TJ. Yellow fever virus NS2B-NS3 protease: charged-to-alanine mutagenesis and deletion analysis define regions important for protease complex formation and function. *Virology*. 2000;275(0042-6822 (Print)):335-47.
34. Issur M, Geiss BJ, Bougie I, Picard-Jean F, Despins S, Mayette J, et al. The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *RNA*. 2009;15(12):2340-50.
35. Godoy AS, Lima GM, Oliveira KI, Torres NU, Maluf FV, Guido RV, et al. Crystal structure of Zika virus NS5 RNA-dependent RNA polymerase. *Nat Commun*. 2017;8:14764.
36. Laurent-Rolle M, Morrison J, Rajsbaum R, Macleod JML, Pisanelli G, Pham A, et al. The interferon signaling antagonist function of yellow fever virus NS5 protein is activated by type I interferon. *Cell Host Microbe*. 2014;16(3):314-27.

37. Germi R, Crance JM, Garin D, Guimet J, Lortat-Jacob H, Ruigrok RW, et al. Heparan sulfate-mediated binding of infectious dengue virus type 2 and yellow fever virus. *Virology*. 2002;292(1):162-8.
38. Mackenzie J. Wrapping things up about virus RNA replication. *Traffic*. 2005;6(11):967-77.
39. Pastorino B, Nougairède A, Wurtz N, Gould E, de Lamballerie X. Role of host cell factors in flavivirus infection: Implications for pathogenesis and development of antiviral drugs. *Antiviral Res*. 2010;87(3):281-94.
40. Mukhopadhyay S, Kuhn RJ, Rossmann MG. A structural perspective of the flavivirus life cycle. *Nat Rev Microbiol*. 2005;3(1):13-22.
41. Apte-Sengupta S, Sirohi D, Kuhn RJ. Coupling of replication and assembly in flaviviruses. *Curr Opin Virol*. 2014;9:134-42.
42. Lepiniec L, Dalgarno L, Huong VT, Monath TP, Digoutte JP, Deubel V. Geographic distribution and evolution of yellow fever viruses based on direct sequencing of genomic cDNA fragments. *J Gen Virol*. 1994;75 (Pt 2):417-23.
43. Zlotoff PM, Gould EA, Gao GF, Harvey PH. Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(2):548-53.
44. Wang E, Weaver SC, Shope RE, Tesh RB, Watts DM, Barrett AD. Genetic variation in yellow fever virus: duplication in the 3' noncoding region of strains from Africa. *Virology*. 1996;225(2):274-81.
45. Bonaldo MC, Gómez MM, Dos Santos AA, Abreu FVS, Ferreira-de-Brito A, Miranda RM, et al. Genome analysis of yellow fever virus of the ongoing outbreak in Brazil reveals polymorphisms. *Mem Inst Oswaldo Cruz*. 2017;112(6):447-51.
46. Monath TP, Vasconcelos PFC. Yellow fever. *Journal of Clinical Virology*. 2015;64:160-73.
47. Barrett ADT, Higgs S. Yellow fever: A disease that has yet to be conquered. *Annual Review of Entomology*. Annual Review of Entomology. 52. Palo Alto: Annual Reviews; 2007. p. 209-29.
48. Mutebi JP, Wang H, Li L, Bryant JE, Barrett AD. Phylogenetic and evolutionary relationships among yellow fever virus isolates in Africa. *J Virol*. 2001;75(15):6999-7008.

49. Tabachnick WJ. Climate Change and the Arboviruses: Lessons from the Evolution of the Dengue and Yellow Fever Viruses. *Annu Rev Virol.* 2016;3(1):125-45.
50. Nunes MR, Palacios G, Cardoso JF, Martins LC, Sousa EC, de Lima CP, et al. Genomic and phylogenetic characterization of Brazilian yellow fever virus strains. *J Virol.* 2012;86(24):13263-71.
51. de Souza RP, Foster PG, Sallum MAM, Coimbra TLM, Maeda AY, Silveira VR, et al. Detection of a New Yellow Fever Virus Lineage Within the South American Genotype I in Brazil. *Journal of Medical Virology.* 2010;82(1):175-85.
52. Domingo C, Patel P, Yillah J, Weidmann M, Méndez JA, Nakouné ER, et al. Advanced yellow fever virus genome detection in point-of-care facilities and reference laboratories. *J Clin Microbiol.* 2012;50(12):4054-60.
53. Barban V, Girerd Y, Aguirre M, Gulia S, Pétiard F, Riou P, et al. High stability of yellow fever 17D-204 vaccine: a 12-year retrospective analysis of large-scale production. *Vaccine.* 2007;25(15):2941-50.
54. PAHO/WHO. Laboratory Diagnosis of Yellow Fever Virus Infection. PAHO/WHO [Internet]. 2018:[1-8 pp.]. Available from: https://www.paho.org/hq/index.php?option=com_docman&view=download&category_slug=guidelines-5053&alias=46877-laboratory-diagnosis-of-yellow-fever-virus-infection&Itemid=270 (Acessado em agosto/2018).
55. Casadio LVB, Salles APM, Malta FM, Leite GF, Ho YL, Gomes-Gouvêa MS, et al. Lipase and factor V (but not viral load) are prognostic factors for the evolution of severe yellow fever cases. *Mem Inst Oswaldo Cruz.* 2019;114:e190033.
56. Domingo C, Escadafal C, Rumer L, Méndez JA, García P, Sall AA, et al. First international external quality assessment study on molecular and serological methods for yellow fever diagnosis. *PLoS One.* 2012;7(5):e36291.
57. Domingo C, Charrel RN, Schmidt-Chanasit J, Zeller H, Reusken C. Yellow fever in the diagnostics laboratory. *Emerg Microbes Infect.* 2018;7(1):129.
58. Shearer FM, Moyes CL, Pigott DM, Brady OJ, Marinho F, Deshpande A, et al. Global yellow fever vaccination coverage from 1970 to 2016: an adjusted retrospective analysis. *Lancet Infect Dis.* 2017;17(11):1209-17.
59. Barnett ED. Yellow fever: Epidemiology and prevention. *Clinical Infectious Diseases.* 2007;44(6):850-6.

60. WHO. "EYE - Eliminate Yellow fever epidemics by 2026" Who2017 [20/01]. Available from: <https://www.who.int/csr/disease/yellowfev/eye-strategy-one-pager.pdf>. (Acessado em agosto/2018)
61. Gershman MD, Staples JE. CDC - Yellow Book. 2019. In: CDC - Yellow Book 2020 [Internet]. Oxford University Press Travel-Related Infectious Diseases.
62. Otshudiema JO, Ndakala NG, Mawanda EK, Tshapenda GP, Kimfuta JM, Nsibu LN, et al. Yellow Fever Outbreak - Kongo Central Province, Democratic Republic of the Congo, August 2016. *MMWR Morb Mortal Wkly Rep.* 2017;66(12):335-8.
63. Barrett AD. Yellow Fever in Angola and Beyond--The Problem of Vaccine Supply and Demand. *N Engl J Med.* 2016;375(4):301-3.
64. Zhao S, Stone L, Gao D, He D. Modelling the large-scale yellow fever outbreak in Luanda, Angola, and the impact of vaccination. *PLoS Negl Trop Dis.* 2018;12(1):e0006158.
65. OPAS/PAHO. Folha informativa - Febre Amarela OPAS/PAHO2019 [Available from: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5578:folha-informativa-febre-amarela&Itemid=875. (Acessado em agosto/2018)
66. Jentes ES, Pomeroy G, Gershman MD, Hill DR, Lemarchand J, Lewis RF, et al. The revised global yellow fever risk map and recommendations for vaccination, 2010: consensus of the Informal WHO Working Group on Geographic Risk for Yellow Fever. *Lancet Infect Dis.* 2011;11(8):622-32.
67. Vasconcelos PF. Yellow fever in Brazil: thoughts and hypotheses on the emergence in previously free areas. *Rev Saude Publica.* 2010;44(6):1144-9.
68. Saúde Md. Monitoramento do Período Sazonal da Febre Amarela Brasil – 2017/2018. In: Saúde, editor.: Ministério da Saúde; 2017/2018. p. 1-12.
69. Goldani LZ. Yellow fever outbreak in Brazil, 2017. *Braz J Infect Dis.* 2017;21(2):123-4.
70. OPAS/PAHO. OPAS/OMS divulga alerta epidemiológico sobre febre amarela para as Américas. OPAS/PAHO; 2017 [Available from: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5329:opas-oms-divulga-alerta-epidemiologico-sobre-febre-amarela-para-as-americas&Itemid=812. Acessado em ago/2018

71. SAÚDE SDED, DOENÇAS CDCD, VRANJAC CDVEPA, ZOONOSES DD. Boletim Epidemiológico Febre Amarela. In: Saúde, editor.: Governo do Estado de São Paulo; 2018. p. 1-9.
72. Paulo GdEdS, Vranjac CdvEPA. Boletim Epidemiológico da Febre Amarela. In: Saúde, editor. 2019. p. 1-6.
73. Song ATW, Abdala E, de Martino RB, Malbouisson LMS, Tanigawa RY, Andrade GM, et al. Liver Transplantation for Fulminant Hepatitis Attributed to Yellow Fever. *Hepatology*. 2019;69(3):1349-52.
74. Faria NR, Kraemer MUG, Hill SC, Goes de Jesus J, Aguiar RS, Iani FCM, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science*. 2018;361(6405):894-9.
75. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnetjournal*. 2011;17(1):10-2.
76. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-77.
77. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol*. 2013;20(10):714-37.
78. Hunt M, Gall A, Ong SH, Brener J, Ferns B, Goulder P, et al. IVA: accurate de novo assembly of RNA virus genomes. *Bioinformatics*. 2015;31(14):2374-6.
79. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.
80. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-9.
81. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14(2):178-92.
82. Garrison E, Marth GT, editors. Haplotype-based variant detection from short-read sequencing 2012.
83. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-8.

84. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011;27(21):2987-93.
85. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22(22):4673-80.
86. 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.
87. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*. 2019;20(4):1160-6.
88. Penn O, Privman E, Ashkenazy H, Landan G, Graur D, Pupko T. GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res*. 2010;38-23-8.
89. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312-3.
90. Cunha MdP, Duarte-Neto AN, Pour SZ, Ortiz-Baez AS, Černý J, Pereira BBdS, et al. Origin of the São Paulo Yellow Fever epidemic of 2017–2018 revealed through molecular epidemiological analysis of fatal cases. *Scientific reports*. 2019;9(1).
91. Beck AS, Barrett AD. Current status and future prospects of yellow fever vaccines. *Expert Rev Vaccines*. 2015;14(11):1479-92.
92. Kallas EG, D'Elia Zanella LGFA, Moreira CHV, Buccheri R, Diniz GBF, Castiñeiras ACP, et al. Predictors of mortality in patients with yellow fever: an observational cohort study. *Lancet Infect Dis*. 2019;19(7):750-8.
93. Saúde Md. Emergências em Saúde Pública de Importância Nacional (ESPIN) de Febre Amarela Silvestre em São Paulo e no Rio Grande do Sul e a Situação Epidemiológica Atual no Brasil (2008/2009). In: Saúde, editor. 2009. p. 10.
94. Brasil OO. Angola e República Democrática do Congo declaram fim do surto de febre amarela OPAS/OMS2017 [Available from: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5350:angola-e-republica-democratica-do-congo-declaram-fim-do-surto-de-febre-amarela&Itemid=812].

95. Chen S, Wu Z, Wang M, Cheng A. Innate Immune Evasion Mediated by Flaviviridae Non-Structural Proteins. *Viruses*. 2017;9(10).
96. Baleotti FG, Moreli ML, Figueiredo LT. Brazilian Flavivirus phylogeny based on NS5. *Mem Inst Oswaldo Cruz*. 2003;98(3):379-82.
97. Patkar CG, Jones CT, Chang YH, Warriar R, Kuhn RJ. Functional requirements of the yellow fever virus capsid protein. *J Virol*. 2007;81(12):6471-81.
98. Barbosa CM, Di Paola N, Cunha MP, Rodrigues-Jesus MJ, Araujo DB, Silveira VB, et al. Yellow Fever Virus RNA in Urine and Semen of Convalescent Patient, Brazil. *Emerg Infect Dis*. 2018;24(1).
99. Dubankova A, Boura E. Structure of the yellow fever NS5 protein reveals conserved drug targets shared among flaviviruses. *Antiviral Res*. 2019;169:104536.
100. Delatorre E, de Abreu FVS, Ribeiro IP, Gómez MM, Dos Santos AAC, Ferreira-de-Brito A, et al. Distinct YFV Lineages Co-circulated in the Central-Western and Southeastern Brazilian Regions From 2015 to 2018. *Front Microbiol*. 2019;10:1079.