

LAÍS CESCHINI MACHADO

**AVALIAÇÃO E REFINAMENTO DE UMA NOVA METODOLOGIA DE
AMOSTRAGEM ASSOCIADA A SEQUENCIAMENTO DE GENOMAS VIRAIS**

Dissertação apresentada ao curso de Mestrado Acadêmico em Biociências e Biotecnologia em Saúde do Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, como requisito para a obtenção do título de Mestre em Ciências.
Área de concentração: Eco-biologia de patógenos, vetores e hospedeiros.

Orientador: Dr. Gabriel da Luz Wallau

Co-orientador: Dr. Marcelo Henrique Santos Paiva

RECIFE

2020

LAÍS CESCHINI MACHADO

**AVALIAÇÃO E REFINAMENTO DE UMA NOVA METODOLOGIA DE
AMOSTRAGEM ASSOCIADA A SEQUENCIAMENTO DE GENOMAS VIRAIS**

Dissertação apresentada ao curso de Mestrado Acadêmico em Biociências e Biotecnologia em Saúde do Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, como requisito para a obtenção do título de Mestre em Ciências.
Área de concentração: Eco-biologia de patógenos, vetores e hospedeiros.

Aprovado em: 19/02/2020

BANCA EXAMINADORA

Dra. Flavia Figueira Aburjaile (Membro externo)
Universidade Federal de Pernambuco

Dra. Claudia Maria Fontes de Oliveira (Membro interno)
Instituto Aggeu Magalhães /- Fundação Oswaldo Cruz

Dr. Gabriel da Luz Wallau (Orientador)
Instituto Aggeu Magalhães - Fundação Oswaldo Cruz

The truth is more magical - in the best and most exciting sense of the world - than any myth or made-up mystery or miracle. Science has its own magic: the magic of reality. (Richard Dawkins, *The Magic of Reality: How We Know What's Really True*)

Dedico este trabalho às minhas
sobrinhas, minhas meninas, Esther e
Maria Eulalia.

Agradecimentos

Agradeço primeiramente a meus pais Elton e Rosélia por todo o amor e apoio em todas as minhas escolhas, e sem sombra de dúvidas são meus maiores exemplos de vida.

Aos meus colegas de grupo do Wallau-Lab e BioInfo, em especial a Luisa, Filipe e Alexandre, o apoio deles foi essencial para a finalização deste trabalho.

Aos colegas do laboratório de Entomologia. A meus colegas mais especiais e meus melhores amigos nessa jornada louca chamada pós graduação, Alexandre, Luísa, Filipe, Lari, Junior e Jade, sem eles eu não duraria nem até a qualificação. Obrigada por enxugar minhas lágrimas, me dar bronca, dividir a xícara de café e também amor e solidariedade em todos os momentos.

A Duschinka, Marcelo e Lari, sem ele este trabalho não seria possível.

Ao melhor orientador de todos (desculpe aos outros) Gabriel, obrigada por todos os ensinamentos, todas as conversas (acadêmicas ou não), pelas cervejas e pizzas e também por confiar em mim, quando nem eu mesma confiava e por sempre estar com a porta aberta para ouvir nossos desesperos.

A minha família Do Nada, que virou tudo, Aline, Lari, Jef, Alex, Filipe, Luísa, que foram essenciais neste último ano.

Por fim e mais importante, agradeço ao meu noivo Alexandre. Obrigada por ser minha calma, meu ponto seguro, por guardar meus sorrisos e minhas lágrimas. Não há palavras que descrevam minha gratidão, apenas que te amo mais que tudo e que é um privilégio dividir a maior aventura de todas contigo. PS: te amo para todo o sempre.

MACHADO, Laís Ceschini. **Avaliação e refinamento de uma nova metodologia de amostragem associada a sequenciamento de genomas virais 2020**. Dissertação (Mestrado Acadêmico em Biociências e Biotecnologia em Saúde) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2020.

RESUMO

Arbovírus são vírus que necessitam de um vetor artrópode, como os mosquitos e carrapatos, para serem transmitidos aos seres humanos. No Brasil, há uma grande variedade de arbovírus circulantes, que compõe um cenário epidemiológico extenso. Os métodos de vigilância para esses arbovírus são feitos de maneira que não se identifica o genótipo viral circulante, no entanto, estes vírus podem acumular mutações de maneira que podem ressurgir como novas epidemias. Tecnologias para o monitoramento viral estão sendo implementadas em diferentes cenários para a obtenção do genoma viral, como o sequenciamento direto de mosquitos infectados realizado com condições de campo. Uma nova tecnologia de monitoramento viral utilizando FTA Cards vem se mostrando robusta para realizar a captura de partículas virais direto da saliva de mosquitos infectados. Sendo assim, a proposta deste trabalho é realizar a combinação de duas técnicas, realizando o sequenciamento de larga escala a partir de FTA Cards para entender a robustez em capturar o genoma destes arbovírus. Para tanto, fêmeas de *Ae. aegypti* foram alimentadas artificialmente com o vírus Zika e FTAs foram utilizados para recuperar partículas virais. FTA Cards colocados em campo junto a armadilhas BR-OVT para diagnóstico de arbovírus em campo por RT-qPCR e PCR *Nested* convencional. Após estes serem confirmados como positivos, através de RT-qPCR, e sequenciados utilizando o kit Nextera XT (Miseq Illumina). Ao todo, foram sequenciados seis FTA Cards de Campo e seis FTA Cards de laboratórios positivos para o vírus Zika, onde estes apresentaram o genoma fragmentado. Foram diagnosticados 19 FTA Cards de Campo positivos para o vírus Zika no RT-qPCR. 18 FTA Cards foram positivos no PCR *Nested* convencional, dois para o vírus chikungunya e 16 para o vírus dengue 2. As análises filogenéticas mostraram que os vírus sequenciados são os circulantes na região Nordeste. Neste trabalho foi possível avaliar que o FTA Card tem um grande potencial para ser usado na vigilância contínua de arbovírus, porém ele não é capaz de recuperar o genoma completo utilizando a abordagem de sequenciamento por amplicons.

Palavras chave: FTA Card; sequenciamento de nova geração; genomas; arbovírus

MACHADO, Laís Ceschini. **Evaluation and refinement of a new sampling methodology associated with viral genome sequencing**, 2020. Dissertação (Mestrado Acadêmico em Biociências e Biotecnologia em Saúde) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2020.

ABSTRACT

Arboviruses are viruses that need an arthropod vector to infect humans. Brazil has numerous described arboviruses, such as dengue virus, Zika virus and chikungunya virus. Moreover, the country has many mosquito vector species, such as *Aedes aegypti*, *Aedes albopictus*, and *Culex* spp. The surveillance method used for these arboviruses does not identify their genomic profile. These viruses have a high mutation rate and are thus able to accumulate mutations and re-emerge as new epidemics. The classical identification of virus RNA sequences uses sequencing of clinical samples or cell cultures, but, due to the high mutation rate, the genetic profile of the virus sequence may not be the same as that of the circulating virus. The search for new technology for viral monitoring using FTA Card® has proven robust in performing the capture of viral particles directly from the saliva of mosquitoes, being able to detect the viruses through sensitivity techniques, such as Real-Time PCR. This study aims to perform two techniques in combination: high-throughput sequencing of FTA Cards samples (for understanding storage capacity) and total genome capture of these arboviruses. The protocol used was artificial feeding of RecLab lineage *Aedes aegypti* mosquitoes with Zika-virus infected blood. Subsequently, we carried out the capture of the virus particle using FTA Cards soaked in sucrose solution as feed. FTA Cards were placed in the field in BR-OVT traps for arbovirus diagnosis using RT-qPCR and conventional Nested PCR. After being confirmed using Real-Time PCR, the FTA Cards were sequenced using the Illumina platform. A total of 12 FTA Cards testing positive for Zika virus were sequenced—six from the field and six from the laboratory. Our results show the FTA Cards contained fragmented genomes. Nineteen FTA Cards tested positive using RT-qPCR for Zika virus from the field. Eighteen FTA Cards were positive using conventional Nested PCR. Two tested positive for chikungunya virus and 16 for dengue serotype 2 virus. Phylogenetic analysis demonstrated that sequenced viruses had circulated in the Northeast region. In the present study, it was possible to demonstrate that FTA Cards have great potential for continuous arbovirus surveillance. However, the FTA Card technology is not able to recover the complete genome using the amplicon sequencing approach.

Key-words: FTA Card; sequencing; arboviruses.

LISTA DE FIGURAS

| | | |
|--------------------|--|-----------|
| Figura 1 - | Distribuição mundial dos principais arbovírus de importância médica e veterinária. | 22 |
| Figura 2 - | Organização estrutural do genoma viral de vírus da família <i>Bunyaviridae</i> . | 23 |
| Figura 3 - | Organização estrutural dos vírus do gênero <i>Alphavirus</i> | 24 |
| Figura 4 - | Organização do genoma dos vírus pertencentes à família <i>Flaviviridae</i> . | 25 |
| Figura 5 - | Distribuição de <i>Aedes aegypti</i> mundial. | 27 |
| Figura 6 - | Distribuição de <i>Aedes albopictus</i> mundial. | 28 |
| Figura 7 - | <i>FTA Cards (Flinders Technology Associates)</i> . | 33 |
| Figura 8 - | <i>Workflow</i> do sequenciamento Illumina. | 37 |
| Figura 9 - | Métodos de sequenciamento de genomas virais. | 39 |
| Figura 10 - | Fluxograma metodológico. | 45 |
| Figura 11 - | Células VERO infectada com ZIKV linhagem PE243. | 46 |
| Figura 12 - | Gaiolas contendo os grupos de mosquitos avaliados. | 48 |
| Figura 13 - | FTA Cards enriquecidos com mel e corante alimentício. | 49 |
| Figura 14 - | Desenho experimental do ensaio de infecção de mosquitos <i>A. aegypti</i> da linhagem Rec Lab pelo ZIKV. | 49 |
| Figura 15 - | Pontos de distribuição dos FTA Cards nos bairros de Recife-PE. | 52 |
| Figura 16 - | Distribuição de FTAs Cards positivos no experimento 1 com ZIKV. | 60 |

| | |
|--|-----------|
| Figura 17 - Distribuição de FTAs Cards positivos no experimento 2 com ZIKV. | 63 |
| Figura 18 - Pontos de FTA Cards do RT-qPCR positivos para a presença de ZIKV. | 65 |
| Figura 19 - Mapas genômicos comparativos dos genomas de ZIKV derivados de FTA Cards de laboratório sequenciado neste trabalho. | 68 |
| Figura 20 - Mapas genômicos comparativos dos genomas de ZIKV derivados de FTA Cards de Campo sequenciado neste trabalho. | 70 |
| Figura 21 - Gel de agarose 2% corado com brometo de etídeo. Amostras de FTA Cards de Campo de PCR <i>Nested</i> convencional para espécies de vírus da família <i>Alphavirus</i> . | 71 |
| Figura 22 - Pontos de FTA Cards positivos no PCR convencional <i>Nested</i> para arbovírus da família <i>Flavivirus</i> e <i>Alphavirus</i> . Pontos laranjas: FTA Cards Positivos para CHIKV, pontos azuis: FTA Cards positivos para DENV 2. | 72 |
| Figura 23 - Gel de agarose 2% corado com brometo de etídeo. Amostras de FTA Cards de Campo de PCR <i>Nested</i> convencional para vírus da família <i>Flavivirus</i> . | 73 |
| Figura 24 - Reconstrução filogenética do ZIKV sequenciados neste trabalho. | 75 |
| Figura 25 - Reconstrução filogenética do CHIKV utilizando alinhamento de genomas completos. | 78 |
| Figura 26 - Reconstrução filogenética do CHIKV utilizando alinhamento do gene NS1 | 81 |

- Figura 27** - Alinhamento utilizando MAFFT e visualizado no Aliview. **82**
Exemplificação dos SNPs dos contigs de CHIKV deFTA Cards de Campo.
- Figura 28** - Alinhamento utilizando MAFFT e visualizado no Aliview. **83**
Exemplificação dos SNPs dos contigs de DENV 2 de FTA Cards de Campo.
- Figura 29** - Reconstrução filogenética do DENV 2 utilizando alinhamento de **85**
genoma completo.
- Figura 30** - Reconstrução filogenética do DENV2 utilizando alinhamento do gene **87**
NS5.

LISTA DE TABELAS

| | | |
|-------------------|---|-----------|
| Tabela 1 - | Distribuição de FTAs Cards positivos no experimento 1. | 59 |
| Tabela 2 - | Valores de <i>Ct</i> e número de moléculas de RNA viral obtidos dos FTA Cards positivos no experimento 1 com ZIKV. | 61 |
| Tabela 3 - | Distribuição de FTAs Cards positivos no experimento 2 com ZIKV. | 62 |
| Tabela 4 - | Valores de <i>Ct</i> e número de partículas de RNA viral obtidos dos FTA Cards positivos no experimento 2 com o ZIKV. | 63 |
| Tabela 5 - | Endereço dos FTA Cards positivos para o ZIKV no RT-qPCR. | 64 |
| Tabela 6 - | Informações relativas aos FTA Cards utilizados no sequenciamento de nova geração (NGS). | 66 |

LISTA DE ABREVIATURAS E SIGLAS

CCHFV - vírus febre hemorrágica Crimean-Congo

CHIKV - Vírus da Chikungunya

Cr - *Cycle threshold*

DENV - Vírus da Dengue

EEEV - vírus da encefalite equina

env - envelope

FTA Card - *Flinders Technology Associates*

HIV - vírus da imunodeficiência humana

IGV - Integrated Genome Viewer

JEV - vírus da encefalite Japonesa

MEM - *Minimum Essential Media*

NGS - *New generation sequencing*

NPT - Núcleo de Plataforma tecnológica

ONNV - vírus o'nyong-nyong

ORFs - *Open read frame*

PCR - Reação em Cadeia da Polimerase

RRV - vírus Ross River

RT-qPCR - *Reverse transcription quantitative Polymerase chain reaction*

RVF - *virus Rift Valley Fever*

SBF - Soro fetal bovino

SFV - vírus Semliki Forest

SINV - vírus Sindbis

SLEV - vírus da Encefalite de Saint Louis

VEEV - vírus da Encefalite Equina Venezuelana

WEEV - vírus Encefalite Equina do Oeste

WNV - Vírus do Oeste do Nilo

YFV - vírus da Febre Amarela

ZIKV- Vírus Zika

SUMÁRIO

| | |
|--|-----------|
| INTRODUÇÃO | 16 |
| 2 MARCO TEÓRICO | 19 |
| 2.1 Vírus patogênicos ao homem | 19 |
| 2.1.1 Vírus de RNA | 19 |
| 2.1.2 Arbovírus e Arboviroses | 20 |
| 2.1.2.1 Família Bunyaviridae | 21 |
| 2.1.2.2 Gênero Alphavirus | 23 |
| 2.1.2.3 Gênero Flavivirus | 24 |
| 2.2.1 Gênero Aedes | 25 |
| 2.2.2 Gênero Culex | 27 |
| 2.3 Situação epidemiológica no Brasil | 28 |
| 2.3.1 DENV | 29 |
| 2.3.2 ZIKV | 30 |
| 2.3.3 CHIKV | 30 |
| 2.3.4 Técnicas de monitoramento viral | 31 |
| 2.3.4 FTA Cards (Flinders Technology Associates) | 32 |
| 2.4 Sequenciamento de ácidos nucleicos | 33 |
| 2.4.1 Sequenciamento de genomas virais | 37 |
| 4 PERGUNTA CONDUTORA | 41 |
| 5 HIPÓTESE | 42 |
| 6 OBJETIVOS | 43 |
| 6.1 Objetivo geral | 43 |
| 6.2 Objetivos específicos | 43 |
| 7 MATERIAIS E MÉTODOS | 44 |
| 6.1 Ensaio de Infecção | 45 |
| 7.2 Extração de RNA | 49 |
| 7.3 Amostras de campo | 50 |
| 7.4 Sequenciamento de nova geração (plataforma Illumina) | 55 |
| 7.5 Avaliação de qualidades das leituras e mapeamento/montagem de genomas virais | 56 |
| 7.6 Análises evolutivas | 57 |
| 8 RESULTADOS | 58 |
| 8.1 Utilização de FTA Cards em laboratório | 58 |
| 8.2 FTA Cards derivados de campo | 63 |

| | |
|--|------------|
| | 14 |
| 8.3 Sequenciamento de nova geração de amostras derivadas dos FTA Cards | 64 |
| 8.5 Análises filogenéticas dos FTA Cards | 72 |
| 9 DISCUSSÃO | 88 |
| 9.1 FTA Cards de laboratório | 88 |
| 9.2 FTA Cards de campo | 89 |
| 9.3 Análises filogenéticas | 92 |
| 10 CONCLUSÃO | 94 |
| REFERÊNCIAS | 96 |
| APÊNDICE A - SEQUÊNCIAS UTILIZADAS PARA RECONSTRUÇÃO FILOGENÉTICA DO VÍRUS ZIKA | 113 |
| APÊNDICE B - SEQUÊNCIAS UTILIZADAS PARA RECONSTRUÇÃO FILOGENÉTICA DO VÍRUS CHIKUNGUNYA. | 117 |
| APÊNDICE C - SEQUÊNCIAS UTILIZADAS PARA RECONSTRUÇÃO FILOGENÉTICA DO VÍRUS DENGUE | 120 |
| APÊNDICE D - ARTIGOS | 125 |