



UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE FITOPATOLOGIA

Felipe Fochat Silva Melo

**UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII* DESF. (COPAÍBA)  
NO BRASIL CENTRAL**

Brasília, DF - 2021

Felipe Fochat Silva Melo

**UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII* DESF. (COPAÍBA)  
NO BRASIL CENTRAL**

Trabalho de Conclusão de Curso como  
requisito para obtenção da Graduação em  
Engenharia Florestal, na Universidade de  
Brasília.

Orientadora Prof. PhD Rita de Cássia  
Pereira Carvalho

Brasília, DF - 2021

Ficha Catalográfica

Fochat Silva Melo, Felipe

FM528N UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII*  
DEF. (COPAÍBA) NO BRASIL CENTRAL / Felipe Fochat Silva  
Melo; orientador Rita de Cássia Pereira Carvalho. -- Brasília, 2021.  
43 p.

Monografia (Graduação - Engenharia Florestal) --  
Universidade de Brasília, 2021.

1. . I. de Cássia Pereira Carvalho, Rita , orient. II. Título.

*A ciência é um quebra-cabeças, cuja montagem só é possível de forma compartilhada.*

*Felipe Fochat*

## AGRADECIMENTOS

Gostaria de agradecer a Deus por ser meu guia e companhia eterna.

A minha mãe que se eu pudesse resumir em uma palavra seria a razão, e muito contribui para eu buscar a realidade das coisas e focar nos meus objetivos.

Ao meu Pai e meu Irmão com quem eu sei que sempre poderei contar.

À toda minha família, que contribuíram de alguma forma para minha educação e formação como ser humano, mesmo que inconscientemente e querer, ou nem saibam disto, saibam que todos estão guardados para sempre comigo e saibam que me inspiro muito em vocês.

Aos grandes amigos pela amizade durante todos estes anos e espero que ainda venham muitos pela frente. Espero que possamos sempre ter tempo para colocarmos a amizade em dia, obrigado sempre pelo apoio.

Gostaria muito de agradecer em especial a dois amigos Brasiliensis a Josiane por toda a motivação e ensinamentos, aprendizado e apoio no laboratório e ao Jorge um amigo que muito me ajudou, principalmente com as consultorias jurídicas e obrigado por estender isso a outros amigos meus, muito obrigado.

A minha Vó Maria pelo exemplo de fé e bondade na minha vida.

Em especial, a minha orientadora, Profa. Rita de Cássia pelo apoio, incentivo, paciência, inspiração e amizade.

A toda a Equipe do Laboratório de Virologia Vegetal: Luciane, Flávia, Maria Luíza, Caroline, Izaías, Henrique, Eduardo, Marcos, Juliana Silva, Juliana Isidório, Jordânia e Letícia.

Saudações florestais aos meus amigos do Curso em especial a Marta pelas parcerias.

Ao Dr. Fernando Lucas Melo, por sempre contribuir e ajudar nos esclarecimentos e dúvidas com bom humor, sempre estando à disposição.

A todo o Departamento de Engenharia Florestal, Fitopatologia e Professores desta Universidade que muito contribuíram para minha formação e a Universidade de Brasília.

A todos os funcionários da UnB que cuidam com tanto carinho e amor nossa instituição.

**UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII* DESF. (COPAÍBA)  
NO BRASIL CENTRAL**

**Estudante: Felipe Fochat Silva Melo**

**Matrícula: 150124929**

**Menção: SS**

---

Rita de Cássia Pereira Carvalho  
(Professora - Departamento de Fitopatologia)  
Orientadora

---

Cleber Furlanetto  
(Professor - Departamento de Fitopatologia)  
Examinador - membro interno

---

Rosana de Carvalho Cristo Martins  
(Professora - Departamento de Engenharia Florestal)  
Examinadora - membro externo

---

Josiane Goulart Batista  
(Doutora em Fitopatologia)  
Examinadora - Suplente

## SUMÁRIO

RESUMO GERAL .....	12
GENERAL ABSTRACT.....	13
1 INTRODUÇÃO GERAL .....	14
2 OBJETIVOS .....	16
2.1 Objetivo Geral .....	16
2.2 Objetivos Específicos.....	16
3 REVISÃO DE LITERATURA.....	17
3.1 <i>Hotspots</i> brasileiros: Cerrado e Mata Atlântica .....	17
3.2 O gênero <i>Copaifera</i> e a sua distribuição geográfica .....	18
3.3 Fitopatógenos em espécies arbóreas.....	19
3.4 <i>High Throughput Sequencing</i> : aplicações no diagnóstico de vírus de plantas .	20
3.4.1 <i>Sequence Read Archive</i> – SRA e <i>Transcriptome Shotgun Archive</i> - TSA ...	21
3.5 A família <i>Betaflexiviridae</i> e o gênero <i>Capillovirus</i> .....	22
4 MATERIAL E MÉTODOS .....	25
4.1 Obtenção das sequências .....	25
4.2 Anotação de ORF's e organização genômica viral .....	25
4.3 Análise das relações filogenéticas .....	25
5 RESULTADOS E DISCUSSÃO .....	27
5.1 Organização genômica de <i>copaifera langsdorffii associated virus 1</i> .....	27
5.2 <i>Sequence Demarcation Tool</i> utilizando a sequência de nucleotídeos de <i>copaifera langsdorffii associated virus 1 - CLaV1</i> e sequências de capilovírus obtidas do GenBank .....	28
5.3 Filogenia do isolado de <i>copaifera langsdorffii associated virus 1 - CLaV1</i> obtido mediante <i>Sequence Read Archive</i> - SRA .....	29
6 CONCLUSÕES.....	29
REFERÊNCIAS BIBLIOGRÁFICAS .....	30
Anexo 1 – Artigo em formato de Short Communications a ser submetido para <i>Virus Genes</i> : <i>Copaifera langsdorffi-associated virus 1: a new Capillovirus occurring in copaifera from Brasil</i> .....	38

## LISTA DE FIGURAS

- Figura 1.** Mapa de distribuição dos biomas brasileiros.....17
- Figura 2.** *Copaifera langsdorffi* A. Porte B. Ritidoma C. Folhas. Fotos: acervo próprio – Localização: Jardim Botânico Brasília – Data 11/09/2021 – Coordenadas geográficas: 15°52'26.31"S e 47°50'3.53"O..... 19
- Figura 3.** Figura 3. Representação da partícula viral e organização genômica típica dos membros de *Capillovirus* – família *Betaflexiviridae*. ORF1 - Poliproteína codificando: MT – metiltransferase, Pro – Proteinase, Hel-helicase, CP – capa proteica, RdRp -RNA polimerase dependente de RNA e na ORF sobreposta abaixo temos MP - proteína de movimento Fonte: ViralZone.....22
- Figura 4.** Pipeline utilizada na prospecção viral. *Copaifera langsdorffii* novel putative long non-coding RNAs: interspecies conservation analysis in adaptive response to different biomes ([Danilevicz et al. 2018](#)); bibliotecas SRA (repositório de sequências disponíveis publicamente); SRX6060682 ARF 1 e SRX6060683 CER 1(acessos para o *download* das bibliotecas de SRA utilizadas neste trabalho); Toolkit Linux (ferramenta para realizar o download das bibliotecas de SRA); maSPAdes (*software* utilizado para montagem De Novo); Kaiju (classificador taxônomico); microrganismos classificados pelo Kaiju para *hit* e não *hit* com sequências de genomas virais; montagem, alinhamento, *Blast* e filogenia (do genoma encontrado); IQTREE e SDT (*softwares* utilizados na classificação taxonômica). CLaV1 - *copaifera langsdorffii-associated virus 1* (nova espécie de *Capillovirus*) .....26
- Figura 5.** Organização genômica do vírus encontrado, Pol - Poliproteína codificando polymerase, replicase e capa proteica e Proteína 36K responsável pelo movimento viral.....27
- Figura 6.** Relações filogenéticas de *copaifera langsdorffii associated virus 1 - CLaV1* e análise SDT utilizando 8 representantes da família *Betaflexiviridae*. A. Árvore de máxima verossimilhança com base em sequências completas do genoma. Os números nos nós representam probabilidades posteriores. Sequências de *Divavirus* foram usadas como grupos externos. As sequências completas do genoma foram alinhadas usando MAFFT v 7.450, e a



árvore filogenética foi construída usando a árvore de máxima verossimilhança realizada com IQTREE , com substituição de aminoácidos selecionada por IQTREE v 2.1.3 GTR + F + I + G4. B. Matriz de pares em todo o genoma gerada por SDT v1.2. A ordem, o número de acesso e os nomes dos isolados (GenBank) são: *Capillovirus*: 1. CLaV1- “copaifera langsdorffi-associated virus 1” 2. KR185346 - Apple stem grooving virus 3. KX668488 - Apple stem grooving virus 4. KY445749 - Cherry virus A 5. KT7663043 - Currant virus A e Divavirus: 6. JX173276 - Diuris virus A 7. JX173277 - Diuris virus B e 8. HQ241409 - Hardenbergia virus A.....28

## LISTA DE TABELA

<b>Tabela 1.</b> Algumas espécies de vírus da família <i>Betaflexiviridae</i> relatadas em espécies arbóreas pelo mundo.....	27
--	----

## LISTA DE ABREVIATURAS

BLAST	Basic Local Alignment Search Tool
DDBJ	DNA Data Bank of Japan
FAO	Food and Agriculture Organization of United Nations
HTS	High Throughput Sequencing
ICTV	International Committee on Taxonomy of Viruses
NCBI	National Center for Biotechnology Information
ORF	Open Reading Frame
SRA	Sequence Read Archive
TNC	The Nature Conservancy
TSA	Transcriptome Shotgun Assembly
UNEP	United Nations Environmental Programme

Melo, FFS. UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII* DESF. (COPAÍBA) NO BRASIL CENTRAL 2021. 52 p. Trabalho de Conclusão de Curso - TCC em Fitopatologia) - Universidade de Brasília, Brasília, DF, Brasil.

### RESUMO GERAL

O Cerrado brasileiro possui uma grande diversidade de espécies e um gênero popularmente conhecido é *Copaifera* (família Fabaceae). *Copaifera* tem como representantes espécies arbóreas que apresentam grande importância econômica, e ocorrem em alta frequência, não apenas no Bioma Cerrado, mas também em todos os outros biomas brasileiros. As espécies de copaíba são utilizadas para obtenção dos produtos madeireiros e não madeireiros. Nas indústrias cosmética e farmacêutica o óleo de copaíba por apresentar ação anti-inflamatória, antimicrobiana e antitumoral, vem ganhando espaço no mercado interno e apresenta, uma crescente demanda de exportações. A copaíba assim como outras espécies arbóreas, podem funcionar como nicho ecológico de diversos microrganismos, inclusive vírus. Características peculiares dos vírus como tamanho pequeno e pouca diversidade morfológica, dificultam sua detecção em espécies arbóreas, principalmente em ambientes naturais. Estima-se que menos de 1% da diversidade viral seja conhecida. Entretanto nos últimos anos, o surgimento de novas tecnologias moleculares aliadas a bioinformática e sequenciamento de alto rendimento, contribuíram para o avanço na descoberta e descrição de novos microrganismos incluindo vírus de plantas. Neste contexto, considerando a diversidade viral, pouco explorada, em espécies arbóreas de ambientes nativos do Brasil e a disponibilidade de informações em repositórios públicos, um levantamento inicial, exaustivo e minucioso de SRA, permitiu a obtenção de um genoma viral parcial em transcriptoma de *Copaifera langsdorffii* Desf. Posteriormente com o auxílio de *softwares* de bioinformática o genoma completo foi recuperado e as ORF's foram devidamente anotadas. Após análises filogenéticas e de predição taxonômica, utilizando critérios de demarcação de novas espécies estabelecidas pelo Comitê Internacional de Taxonomia de Vírus (ICTV), concluiu-se que o genoma obtido corresponde a uma nova espécie, denominada *copaifera langsdorffii-associated virus 1* (CLaV1) para o gênero *Capillovirus* e família *Betaflexiviridae*. Este é o primeiro relato de um vírus em copaíba.

Palavras-chave: espécies arbóreas, vírus, *High Throughput Sequencing*

Melo, FFS. UM NOVO CAPILOVÍRUS EM *COPAIFERA LANGSDORFFII* DESF. (COPAÍBA) NO BRASIL CENTRAL 2021. 52 p. Trabalho de Conclusão de Curso - TCC em Fitopatologia) - Universidade de Brasília, Brasília, DF, Brasil.

#### GENERAL ABSTRACT

The Brazilian Cerrado has a great diversity of species, and a popularly known genus is *Copaifera* (Fabaceae family), has as representatives tree species with great economic importance, and occurs in high frequency, not only in the Cerrado Biome, but also in all other Brazilian biomes. Copaiba species are used to obtain wood and non-wood products. In the cosmetic and pharmaceutical industries, copaiba oil has an anti-inflammatory, antimicrobial and antitumor action, that has been gaining space in the domestic market and presents a growing demand for exports. Copaiba, like other tree species, can function as an ecological niche for several microorganisms, including viruses. Peculiar characteristics of viruses, like small size and little morphological difference, difficult virus detection in arboreal species, especially in natural environments. It is estimated that less than 1% of viral diversity is known. However, in recent years, the emergence of new molecular technologies combined with bioinformatics and sequencing technologies have contributed to advances in the discovery and description of new microorganisms including plant viruses. In this context, considering the viral diversity, little explored, in arboreal species from native Brazilian environments and the availability of information in public repositories, an initial, exhaustive and thorough survey of SRA, allowed to obtain a partial viral genome in *Copaifera langsdorffi* Desf. Later, with the aid of bioinformatics software, the complete genome was recovered and the ORF's were noted. After phylogenetic analysis and taxonomic prediction, using criteria for demarcation of new species established by the International Committee on Taxonomy of Viruses (ICTV), it was concluded that the genome obtained, constitutes a new species named *copaifera langsdorffii*-associated virus 1 (CLaV1) for the genus *Capillovirus* and family *Betaflexiviridae*. This is the first report of a virus in copaiba.

Key words: tree species, viruses, high throughput sequencing

## 1 INTRODUÇÃO GERAL

Estima-se que existam 3,04 trilhões de árvores no planeta (Crowther et al. 2015). As florestas recobrem aproximadamente 31% da área terrestre (FAO e UNEP, 2020) com diversidade de domínios que vão do Boreal ao Tropical (Pan et al. 2013). Segundo Da Fonseca (2017), as florestas tropicais representam importantes centros globais da biodiversidade. O Brasil, por sua vez, é detentor da segunda maior cobertura florestal mundial (Higa, 2014), sendo composto por 6 diferentes biomas (Souza et al. 2020), destacando-se Cerrado e Mata Atlântica.

O Cerrado é um dos biomas mais importantes no continente Sul-Americano e o segundo maior do Brasil (MMA, 2021). Este bioma é considerado a savana mais rica do planeta devido à biodiversidade e elevado grau de endemismo. A flora do Cerrado possui em torno de 12 mil espécies catalogadas e destas, estima-se que 4 mil sejam endêmicas (ISPN, 2021).

Já a Mata Atlântica se estende por boa parte do litoral brasileiro, atingindo o interior austral do continente, onde abrange regiões do Paraguai e Argentina (Morellato; Haddad 2000; Ribeiro et al. 2009). Grandes compilados sobre abundância e distribuição de espécies têm sido realizados demonstrando uma elevada riqueza de espécies, alto nível de endemismo, e um alto impacto antrópico associado (Ribeiro et al. 2009). Para Myers et al. (2000), a Mata Atlântica é considerada uma região prioritária para conservação da biodiversidade.

Uma espécie arbórea muito comum encontrada nestes dois Biomas é *Copaifera lagsdorffii* Desf. Esta espécie tem grande importância econômica devido aos seus produtos madeireiros e não madeireiros (Junior e Pinto, 2002). Em uma revisão realizada por Pieri et al. (2009), foi demonstrado o grande potencial tecnológico e aplicações desta espécie principalmente pelo óleo que vem sendo utilizado pela medicina tradicional popular e silvícola há mais de 500 anos.

As espécies arbóreas funcionam como nichos ecológicos que abrigam diversos microrganismos (Gonthier e Nicolotti, 2013) e em decorrência desta interação podem acarretar prejuízos à planta hospedeira. Segundo Batista (2020) e Santos (2016), as espécies arbóreas podem ser acometidas por pragas e doenças (incluindo aquelas de etiologia viral) e

que podem representar ameaça em potencial quando transferidas de um ambiente nativo para lavouras e vice-versa (Rodríguez-Nevado et al. 2019). No entanto, devido à complexidade de detecção destes patógenos em ambientes naturais, existe uma escassez de informações na literatura (Santos, 2016). Acredita-se que menos de 1% da diversidade viral existente no planeta seja conhecida (Mokili et al. 2012) e há muitas interações vírus-planta que ainda não são muito bem elucidadas (Rossinck, 2011).

Nos últimos anos, com o surgimento de novas tecnologias moleculares, *pipelines* de trabalho, bioinformática, novos protocolos e tecnologias de sequenciamento de alto rendimento, contribuíram muito para descrição de diversas relações de microrganismos já conhecidos ou novos, em diferentes plantas e ambientes. Inclusive diversas espécies estão sendo descobertas por meio de *Sequence Read Archive* - SRA e *Transcriptome Shotgun Assembly* - TSA (Li et al 2019; Macchi et al. 2020; Marais et al. 2021).

O *High Throughput Sequencing* – HTS, cujo significado é tecnologia de sequenciamento de alto desempenho juntamente com a bioinformática, por exemplo, tem contribuído em um curto intervalo de tempo para a detecção de espécies virais em diferentes ambientes e diferentes hospedeiras. Diversas espécies de vírus e viróides associadas às espécies arbóreas já foram detectadas desta maneira (BARBA et al. 2014).

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

O objetivo do presente trabalho foi caracterizar molecularmente e registrar a ocorrência de vírus em SRA de *Copaifera langsdorffi* disponíveis ao público, provenientes de sequenciamento de alto rendimento (HTS).

### **2.2 Objetivos Específicos**

Identificar vírus ocorrendo em *Copaifera langsdorffi* por meio de arquivos SRA disponíveis ao público proveniente do sequenciamento de alto desempenho;

Montar o genoma viral, realizar anotação e alinhamento da sequência com genoma de referência e

Analisar as relações filogenéticas para a demarcação do genoma viral.



### 3 REVISÃO DE LITERATURA

---

#### 3.1 Hotspots brasileiros: Cerrado e Mata Atlântica

Os biomas brasileiros, ilustrados na **Figura 3**, abrigam uma porção significativa da biodiversidade mundial, constituindo importante centro de biodiversidade pela combinação de altos níveis de riqueza e endemismo (Aleixo et al. 2010).

Com a devastação ambiental e ameaças, em detrimento de atividades antrópicas, principalmente pela mudança do uso da terra onde paisagens naturais são convertidas em áreas para agricultura, empreendimentos imobiliários dentre outras atividades, há níveis com elevadíssimo risco à fauna e flora dos biomas, incluindo Mata Atlântica e Cerrado.



**Figura 1.** Mapa da distribuição dos biomas brasileiros **Fonte:** SNIF, 2020 (adaptado)

O Bioma Cerrado é uma savana neotropical e o segundo maior bioma da América do Sul, ocupando uma área cerca de 2 milhões km<sup>2</sup>, o equivalente a 22% do território nacional. A sua área contínua incide sobre os estados de Goiás, Tocantins, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Bahia, Maranhão, Piauí, Rondônia, Paraná, São Paulo e Distrito

Federal, e conseqüentemente, acaba por conectar 5 biomas brasileiros, além de se apresentar na forma de ilhas dentro de outros biomas nos estados no Amapá, Roraima e Amazonas (Ribeiro; Walter, 2008). Também, incorpora partes do território de países vizinhos como a Bolívia e o Paraguai (Ribeiro; Walter, 2008).

Em sua composição de fitofisionomias, o Cerrado abriga uma grande variedade de espécies, sendo reconhecido como a savana mais rica do mundo e estima-se que 1/3 da biota brasileira e cerca de 5% da fauna e flora mundial estão neste bioma (Ribeiro e Walter, 2008). Cerca de 12 mil espécies botânicas fazem parte do bioma Cerrado e estão representadas por plantas classificadas nas famílias Fabaceae, Melastomataceae, Vochysiaceae, Malpighiaceae, Clusiaceae, Erythroxylaceae e Myrtaceae. Deste total mais de cinco mil espécies foram consideradas endêmicas (Rezende et al. 2008).

A Mata Atlântica também abriga uma quantidade significativa de espécies. O Bioma abrange 17 estados brasileiros e devido a ocupação desordenada atualmente 70% da população brasileira, vive neste bioma que já acumula 2/3 de todas as espécies ameaçadas de extinção do país (TNC, 2021).

Uma espécie arbórea amplamente distribuída em todo o território brasileiro, incluindo os biomas Mata Atlântica e Cerrado é *Copaifera langsdorffii* Desf. Informações sobre esta espécie, incluindo importância econômica serão descritas a seguir.

### **3.2 O gênero *Copaifera*: uso e distribuição geográfica**

O gênero *Copaifera* spp. (família Fabaceae) compreende espécies arbóreas que apresentam crescimento lento, mas que podem alcançar 40 metros de altura e viver em torno de 400 anos. São nativas da região tropical da América Latina e África Ocidental. O gênero possui 72 espécies e 16 delas são encontradas apenas do Brasil (Júnior e Pinto, 2002). As espécies botânicas deste gênero possuem grande importância econômica, devido aos produtos florestais madeireiros e não madeireiros obtidos, em especial o óleo-resina.

Este óleo-resina extraído das espécies de copaíba tem potencial terapêutico com comprovação pela ciência em humanos e animais (Pieri et al. 2009; Pieri et al. 2010), devido a ação anti-inflamatória, antimicrobiana, cicatrizante, diurética, inibidor tumoral e laxante.

Pode ser utilizado por via tópica e oral de acordo com o tipo terapêutico e a espécie de copaíba empregada (Quemel et al. 2021). É um excelente fixador para perfumes e possui propriedades emolientes, o que o faz ser usado como matéria prima na produção de cosméticos para cabelo e pele (Júnior e Pinto, 2002). Além disso em razão da viscosidade da resina produzida pela *C. langsdorffii*, o que confere alta durabilidade nos produtos gerados, a mesma é bastante aproveitada na produção de vernizes, tinturas e lacas (De Vasconcelos e Godinho, 2002; Tappin et al. 2004).

Devido ao potencial tecnológico da madeira, a copaíba é indicada para uso industrial na construção civil, naval, moveleira e energética (Pierre et al. 2009).



**Figura 2.** *Copaifera langsdorffii* A. Porte B. Ritidoma C. Folhas. **Fotos:** acervo próprio – Localização: Jardim Botânico Brasília – Data 11/09/2021 – Coordenadas geográficas: 15°52'26.31"S e 47°50'3.53"O.

### 3.3 Fitopatógenos em espécies arbóreas

Embora as doenças que acometem espécies arbóreas, na maioria dos casos são ocasionadas por fungos (Alfenas et al. 2004), oomicetos e bactérias, sabe-se que fitoplasmas, nematóides e vírus também podem afetar espécies arbóreas em ambientes naturais ou de cultivo. Um amplo e crescente trânsito de material vegetal infectado com vírus pode representar uma grande ameaça para espécies florestais em ecossistemas naturais e em áreas cultivadas em todo o mundo (Vicent et al. 2014). Apesar disto, pouca atenção tem sido dada

à ameaça que os vírus representam para espécies arbóreas e seus potenciais impactos na redução da biodiversidade global.

Cerca de 15% dos vírus que afetam plantas já foram descritos em associação natural com vegetação espontânea em diversas famílias (Brunt, 1997; Chaves et al. 2003).

Embora a importância sobre o assunto seja levantada, o viroma em espécies arbóreas é mais restrito a pomares comerciais. Em geral, estudos em espécies arbóreas sendo acometidas por novas espécies virais ou espécies já conhecidas ainda são emergentes, e a grande maioria limitadas a espécies nos gêneros de importância econômica, como: *Prunus*, *Pyrus*, *Malus*, *Citrus*, *Actinia*, *Diospyros*, *Morus* e *Vitis* (Nery et al. 2020).

Em ambientes naturais, entre outros ambientes, principalmente aqui no Brasil, os estudos ainda são incipientes e pouco descritos na literatura (Batista, 2014). A falta de informações de espécies virais ocorrendo em plantas arbóreas pode ser atribuída a algumas características peculiares apresentadas pelos vírus, como pequena diversidade morfológica e tamanho microscópico (Santos, 2016).

### **3.4 High Throughput Sequencing - HTS: Aplicações no diagnóstico de vírus de plantas**

A aplicação de *High-Throughput Sequencing* (HTS) tem se mostrado muito promissora para a descoberta de vírus para resolver a etiologia da doença em muitas culturas agrícolas (Maree et al. 2018; Rumbou et al. 2021). A identificação correta de espécies virais é um passo importante para adoção de estratégias de controle.

No passado várias dificuldades foram encontradas quanto à diagnose viral, atualmente, isso fica restrito na quantidade das informações geradas, ou na falta das informações para comparativo das sequências (Santos, 2016).

Em uma análise metagenômica realizada por Roosinck et al. (2010) em áreas de conservação na Costa Rica, foi possível identificar 344 vírus, dos quais 70% correspondiam a espécies novas virais presentes em 11 famílias de vírus: *Bromoviridae*, *Chrysoviridae*, *Totiviridae*, *Endornaviridae*, *Tymoviridae*, *Partitiviridae*, *Potyviridae*, *Closteroviridae*, *Luteoviridae*, *Narnaviridae* e *Caulimoviridae*.

Nos últimos anos, com o avanço destas tecnologias de sequenciamento e produção de dados com taxa de transferência significativamente maior, com menor custo e conforme esses projetos chegam à publicação, o conjunto de dados do sequenciamento bruto é enviado para o repositório *Sequence Read Archive* – SRA (Leinonen et al. 2010). Desta forma ficam disponíveis publicamente para serem usados por pesquisadores de qualquer área para serem analisados.

### **3.4.1 *Sequence Read Archive* – SRA e *Transcriptome Shotgun Archive* – TSA**

Os dados *Sequence Read Archive* (SRA) estão disponíveis por meio de vários provedores de nuvem e servidores NCBI, que corresponde ao maior repositório disponível publicamente de dados de sequenciamento de alto rendimento. O arquivo aceita dados de todos os ramos da vida, bem como pesquisas metagenômicas e ambientais. O SRA armazena dados de sequenciamento brutos e informações de alinhamento para aprimorar a reprodutibilidade e facilitar novas descobertas por meio da análise de dados (NCBI, 2021).

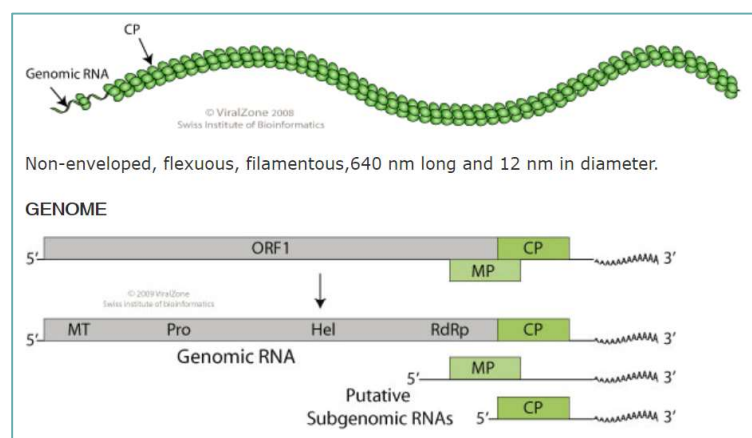
Segundo Leinonen (2010), arquivar dados experimentais é a chave para o progresso da ciência reproduzível. O SRA foi estabelecido como um repositório público para dados de sequência de próxima geração no Centro Nacional de Informações sobre Biotecnologia (NCBI), Instituto Europeu de Bioinformática (EBI) e o Banco de Dados de DNA do Japão (DDBJ). A SRA está acessível em [www.ncbi.nlm.nih.gov/sra](http://www.ncbi.nlm.nih.gov/sra) (NCBI), [www.ebi.ac.uk/ena](http://www.ebi.ac.uk/ena) (EBI) e [trace.ddbj.nig.ac.jp](http://trace.ddbj.nig.ac.jp) (DDBJ).

Já os arquivos *Transcriptome Shotgun Archive* - TSA são arquivos de sequências de transcrição montadas computacionalmente a partir de dados primários, com base em tecnologias de sequenciamento de última geração. As leituras de sequência de sobreposição de um transcriptoma completo são montadas em transcritos por métodos computacionais, em vez da clonagem tradicional e sequenciamento de cDNAs clonados. Os dados da sequência primária usados nas montagens devem ter sido determinados experimentalmente pelo mesmo remetente (NCBI, 2021).

Nos últimos anos, graças a estes repositórios, diversos trabalhos vêm sendo publicados utilizando de arquivos SRA e TSA pelo mundo para a descoberta de novas espécies, inclusive de vírus (Maachi et al. 2020; Kondo et al. 2013)

### 3.5 A família *Betaflexiviridae* e o gênero *Capillovirus*

Os vírus da família *Betaflexiviridae* estão classificados na ordem *Tymovirales* e apresentam partículas filamentosas flexuosas não envelopadas com tamanho em torno de 620-680 nm de comprimento com 12nm de largura e genoma de ssRNA (+) variando de 5,9 a 9,0 kb (Clover et al. 2003; Adams et al. 2004; Jo et al. 2016 e Macchi et al. 2020) (**Figura 5**). Um total de 108 espécies estão classificadas em 13 gêneros e em duas subfamílias. As subfamílias são: *Quinvirinae* (*Carlavirus*, *Foveavirus* e *Robigovirus*) e *Trivirinae* (*Capillovirus*, *Chordovirus*, *Citrivirus*, *Divavirus*, *Prunavirus*, *Ravavirus*, *Tepovirus*, *Trichovirus*, *Vitivirus* e *Wamavirus*) (Adams et al. 2004; ICTV, 2021). Enquanto representantes da subfamília *Quinvirinae* codificam para o triplo bloco gênico (TGB), os membros da subfamília *Trivirinae* codificam proteínas semelhantes a proteína de movimento de 30K (função de movimento viral nos dois casos) (Clover et al. 2003).



**Figura 3.** Representação da partícula viral e organização genômica típica dos membros de *Capillovirus* – família *Betaflexiviridae*. ORF1 - Poliproteína codificando : MT – metiltransferase, Pro – Proteinase, Hel- helicase, CP – capa proteica, RdRp -RNA polimerase dependente de RNA e na ORF sobreposta abaixo temos MP - proteína de movimento **Fonte:** ViralZone

A distribuição e gama de hospedeiras da família *Betaflexiviridae* tem se mostrado bastante variadas e apesar das espécies serem descritas principalmente em plantas cultivadas, existem relatos de espécies desta família ocorrendo em plantas lenhosas e arbustos (Rumbou et al. 2021). Na subfamília *Trivirinae* ocorre predominância de vírus infectando plantas lenhosas como videiras e árvores frutíferas. A transmissão viral ocorre por meio de vetores (ácaros, cochonilha e pulgões), enxertia, mecanicamente e por sementes (Gaafar e Ziebell 2021).

O apple stem grooving virus – ASGV (gênero *Capillovirus*) causa acanelamento do lenho da macieira, devido ao principal sintoma observado em plantas suscetíveis; e sua infecção pode levar à perda de qualidade e queda na produção de frutos. Também infecta uma ampla variedade de plantas hospedeiras, árvores frutíferas, arbustos e algumas plantas herbáceas (Souza 2016).

Alguns exemplos de espécies arbóreas sendo acometidas por espécies virais desta família: *Sambucus nigra* (Van Lent et al 1980; Kalinowska et al. 2013), *Hevea brasiliensis* (Gama et al. 1983; Li et al. 2019), *Senna macranthera* (Seabra et al. 2001; Lin et al. 1979), *Malus* spp. (Lister et al. 1965; Freitas-Astua et al. 2004; Rana et al. 2008; Jiao et al. 2021) *Prunus* spp. (Kinoti et al. 2017; Kinoti et al 2016; Çağlayan et al. 2019), *Sorbus aucuparia* (Sweet 1980), *Pyrus betulifolia* (Ma et al. 2021), *Populus* spp (Schmelzer, 1966; Navrátil e Boyer, 1968; Cooper e Edwards 1981, Cooper et al 1986, Smith and Campbell 2004, Smith et al.2004), *Aesculus hippocastanum* (Hentsch et al. 1997), *Ulmus laevis* (Eisold et al. 2017), *Bauhinia variegata* (Bhardwaj et al. 2015) e *Quercus* spp (Rana et al. 2008).

Na **Tabela 2**, encontra-se a representação de algumas espécies de vírus da família *Betaflexiviridae* ocorrendo em espécies arbóreas pelo mundo.

**Tabela 2.** Exemplos de vírus da família *Betaflexiviridae* relatadas em espécies arbóreas pelo mundo.

Gênero	Vírus	Hospedeiro	Distribuição	Referências
<b>Carlavirus</b>	Blueberry scorch virus	<i>Sambucus nigra</i>	Polônia	Kalinowska et al. (2013)
	Elderberry carlavirus		Holanda	Van Lent et al. (1980)
	-	<i>Hevea brasiliensis</i>	Brasil	Gama et al. (1983)
	-	<i>Senna macranthera</i>		Lin et al. (1979); Seabra et al. (2001)
	Poplar mosaic virus	<i>Populus</i> spp.	Europa	Schmelzer (1966); Navrátil e Boyer (1968) Cooper e Edwards (1981); Cooper et al. (1986); Smith e Campbell (2004); Smith et al. (2004)
<b>Cherry virus A</b>		<i>Prunus cerasifera</i>	Austrália	Kinoti et al. (2016)
<b>Rubber tree virus 1</b>	<i>Capillovirus</i>	<i>Hevea brasiliensis</i>	China	Li et al. (2019)
<b>Apple stem grooving virus</b>		<i>Bauhinia variegata</i>	Índia	Bhardwaj et al. (2015)
		<i>Pyrus betulifolia</i>	China	Ma et al. (2021)
	<i>Malus</i> spp.	Jiao et al. (2021)		
<b>Apple stem pitting virus</b>	<i>Foveavirus</i>	<i>Malus</i> spp.		Jiao et al. (2021)
		<i>Malus domestica</i>	Brasil	Freitas-Astua et al. (2004)
		<i>Sorbus aucuparia</i>	Reino Unido	Sweet et al. (1980)
		<i>Pyrus betulifolia</i>	China	Ma et al. (2021)
<b>Apricot vein clearing associated virus</b>	<i>Prunevirus</i>	<i>Prunus cerasifera</i>	Austrália	Kinoti et al. (2017)
<b>Chery virus turkey</b>	<i>Robigovirus</i>	<i>Prunus avium</i>	Turquia	Çağlayan et al. (2019)
		<i>Quercus</i> spp.	Índia	Rana et al. (2008)
		<i>Sorbus</i> spp.	Reino Unido	Sweet (1980)
		<i>Sorbus aucuparia</i>		Sweet et al. (1980)
		<b>Apple chlorotic leaf spot virus</b>	<i>Trichovirus</i>	<i>Malus</i> spp.
<i>Aesculus hippocastanum</i>	Alemanha			Hentsch et al. (1997)
<i>Prunus</i> spp.	Índia			Rana et al. (2008)
<i>Pyrus betulifolia</i>	China			Ma et al. (2021)
<i>Malus domestica</i>	Alemanha			Lister et al. (1965)



## **4 MATERIAL E MÉTODOS**

### **4.1 Obtenção das sequências**

Para este trabalho foram utilizados dois conjuntos de dados disponíveis publicamente baixados do diretório de SRA no repositório NCBI. Para download foi utilizada a ferramenta *Toolkit* de SRA para Linux em distribuição Ubuntu, de acordo com os respectivos códigos: bioproject SRX6060683: RNAseq de *Copaifera langsdorffii*: CER 1 com 83,2 milhões de *reads* correspondendo as amostras coletadas no bioma Cerrado e SRX6060682: RNAseq de *Copaifera langsdorffii*: ARF 1 84,3 milhões de *reads*, correspondendo a amostras coletadas no bioma Mata Atlântica, e ambas sequenciadas por meio de *Illumina* HiSeq 2500. Essas respectivas bibliotecas de RNA foram utilizadas por Danilevicz et al. (2018) para estudos de verificação de novos lncRNA's putativos de *Copaifera langsdorffii* para analisar a conservação interespecies em respostas adaptativas para diferentes biomas.

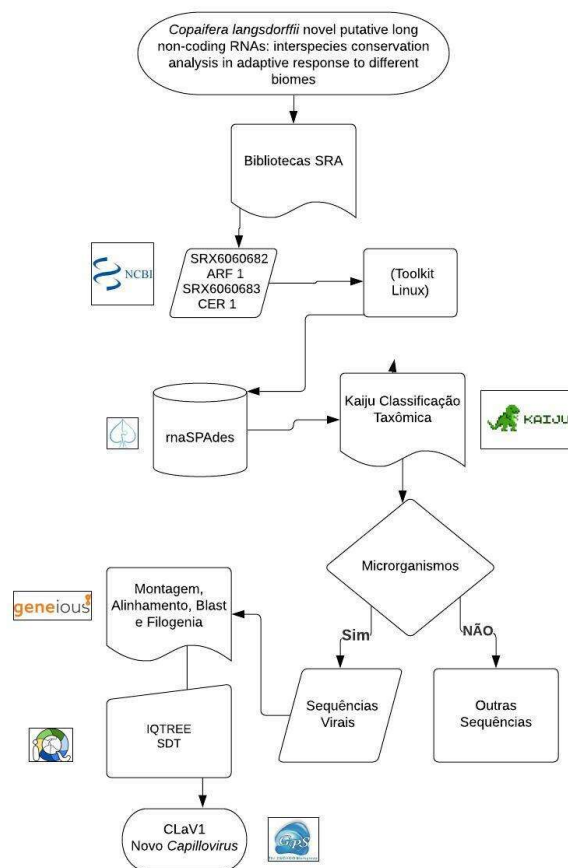
### **4.2 Anotação de ORF's e organização genômica viral**

As sequências de cada biblioteca foram montadas usando-se o programa rnaSPAdes v3.14.1 (Bushmanova et al. 2019). Após a montagem, as sequências passaram por predição taxonômica usando o Kaiju (Menzel et al. 2016), com os parâmetros padrões de classificação. A partir dessa análise, foram separadas aquelas de taxonomia predita como vírus, bactérias e fungos entre outros microrganismos e foram selecionadas as sequências de maior tamanho montadas. As sequências virais foram alinhadas contra o genoma de maior identidade extraído do banco de dados (Menzel et al. 2016) e utilizando-se o programa *Geneious* R11.1 (Kearse et al. 2012) que também permitiu realizar a montagem do genoma viral, anotação e o alinhamento das sequências montadas contra o genoma de referência utilizado.

### **4.3 Análise das relações filogenéticas**

Para a filogenia foi utilizado o *software* IQTREE (Minh et al. 2020), onde foram testados 277 modelos para construção das relações filogenéticas e seleção do melhor modelo. Para a demarcação de espécie, utilizou-se a ferramenta SDT 1.2 (Muhire et al. 2014). Ambas

as figuras foram geradas e compostas por meio do *Evolview* e submetida para edição no *Illustrator v 2021*.



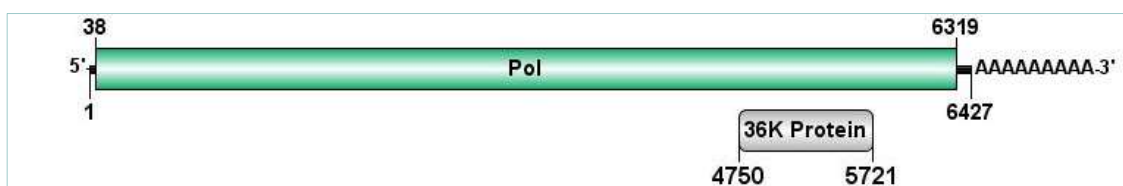
**Figura 4.** Pipeline utilizada na prospecção viral. *Copaifera langsdorffii* novel putative long non-coding RNAs: interspecies conservation analysis in adaptive response to different biomes (Danilevicz et al. 2018); bibliotecas SRA (repositório de sequências disponíveis publicamente); SRX6060682 ARF 1 e SRX6060683 CER 1 (acessos para o download das bibliotecas de SRA utilizadas neste trabalho); Toolkit Linux (ferramenta para realizar o download das bibliotecas de SRA); maSPAdes (software utilizado para montagem De Novo); Kaiju (classificador taxônomico); microrganismos classificados pelo Kaiju para hit e não hit com sequências de genomas virais; montagem, alinhamento, Blast e filogenia (do genoma encontrado); IQTREE e SDT (softwares utilizados na classificação taxonômica). *Copaifera langsdorffii*-associated virus 1 - CLaV1 (nova espécie de *Capillovirus*).

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Organização genômica de *copaifera langsdorffii*-associated virus 1

Por meio do arquivo SRX6060683 CER1 correspondente à dados de amostras foliares de *Copaifera langsdorffii* provenientes do bioma Cerrado, foi possível a recuperação de um genoma viral completo com 6427 nucleotídeos, que após ser submetido ao *Blastn* apresentou identidade em nucleotídeos de 70,23% com o genoma de referência NC\_001749 (Apple stem grooving virus, gênero *Capillovirus*). De acordo com os critérios de demarcação para novas espécies virais na família *Betaflexiviridae* estabelecidos pelo ICTV (2021), o genoma deve apresentar identidade de nucleotídeos menor que 72% e de aminoácidos menor que 80% para que seja considerada uma nova espécie viral. Identidade nucleotídica menor que 45% caracterizam um novo gênero. Neste trabalho, a identidade encontrada foi menor que 72% em nucleotídeos, sendo, portanto, possível a proposta de uma nova espécie dentro do gênero *Capillovirus*. A identidade em aminoácidos para as ORFs 1 e 2 foram 63,94% e 54,13 % também utilizando o genoma de referência NC\_001749. Os dados deste trabalho corroboram com um levantamento de vírus realizado em arquivos de SRA, porém com outra espécie botânica a *Phoenix dactylifera* (Tâmara), mas que também obtiveram resultado positivo para um novo vírus de *Betaflexiviridae* (Maachi et al. 2020).

Na **Figura 7** abaixo, pode ser observada a organização genômica deste novo vírus. Em verde temos a ORF 1 que codifica uma poliproteína com terminal 3' poliadenilado, e em cinza a ORF 2 que codifica para uma proteína responsável pelo movimento viral com tamanho de 971 nucleotídeos. O vírus foi nomeado *copaifera langsdorffii*-associated virus 1.



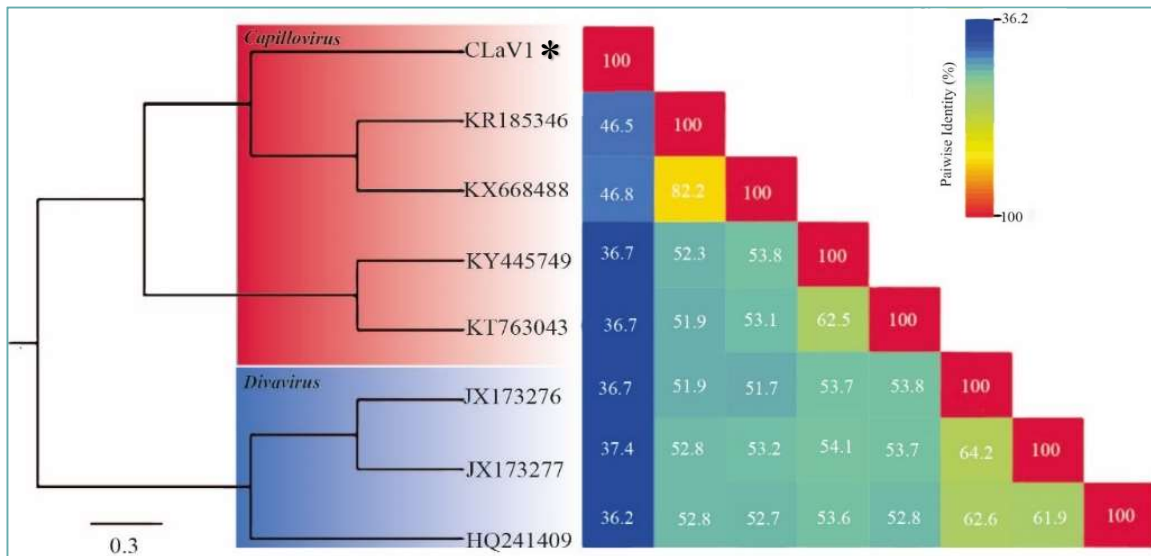
**Figura 5.** Organização genômica do vírus encontrado, Pol - Poliproteína codificando polymerase, replicase e capa proteica e Proteína 36K responsável pelo movimento viral.

Já por meio do arquivo SRX6060682 ARF1 correspondente à amostras foliares de *Copaifera langsdorffii* proveniente do bioma Mata Atlântica não foi possível detectar CLaV-

lou similar. Estudos futuros poderão ser realizados para detecção e montagem de outros vírus.

## 5.2 *Sequence Demarcation Tool* utilizando a sequência de nucleotídeos de copaifera langsdorffii virus 1 – CLaV1 e sequências de capilovírus e divavírus obtidas do *GenBank*

Resultado obtido através do *Sequence Demarcation Tool* – SDT v 1.2 (**Figura 8**), permitiu a classificação do isolado de copaifera langsdorffii virus 1 – CLaV1 com base na identidade dos pares de sequências utilizados. Observou-se identidades abaixo de 72% neste tipo de análise o que valida os resultados obtidos no *Blastn* e confirmam que se trata de uma nova espécie no gênero *Capillovirus* devido a alta relação com este gênero.



**Figura 6.** Relações Filogenéticas de CLaV1 copaifera langsdorffii associated virus 1 e análise SDT utilizando 8 representantes da família *Betaflexiviridae*. A. Árvore de máxima verossimilhança com base em sequências completas do genoma. Os números nos nós representam probabilidades posteriores. Sequências de *Divavirus* foram usadas como grupos externos. As sequências completas do genoma foram alinhadas usando MAFFT v 7.450, e a árvore filogenética foi construída usando a árvore de máxima verossimilhança realizada com IQTREE, com substituição de aminoácidos selecionada por IQTREE v 2.1.3 GTR + F + I + G4. B. Matriz de pares em todo o genoma gerada por SDT v1.2. A ordem, o número de acesso e os nomes dos isolados (GenBank) são: *Capillovirus*: 1. CLaV1- “copaifera langsdorffi-associated virus 1” 2. KR185346 - Apple stem grooving virus 3. KX668488 - Apple stem grooving virus 4. KY445749 - Cherry virus A 5. KT7663043 - Currant virus A e *Divavirus*: 6. JX173276 - Diuris virus A 7. JX173277 - Diuris virus B e 8. HQ241409 - Hardenbergia virus A.

### 5.3 Filogenia do isolado de *copaifera langsdorffii* associated virus 1 – ClaV1 obtido mediante *Sequence Read Archive* – SRA

A análise utilizando o IQTREE v 2.1.3 permitiu a seleção do melhor modelo pelo critério de informação bayesiano: GTR + F + I + G4. Em matriz de substituição nucleotídica foi construída a melhor árvore filogenética com base nos 277 modelos testados (**Figura 8**), com base no alinhamento MAFFT v 7.450. A análise filogenética de sequências de genoma completo de espécies representativas de *Capillovirus* e *Divavirus* (*outgroup*) e o genoma de “*copaifera langsdorffii*-associated virus 1” também indicaram alta relação com o gênero *Capillovirus*. Observou-se o agrupamento esperado de CLaV-1 com Apple stem grooving virus (KR185346) e Apple stem grooving virus (KX668488).

Com base na metodologia utilizada neste trabalho foi possível confirmar o benefício de se utilizar SRA. Como vantagens podem ser citadas baixo custo, tempo relativamente curto e acessibilidade por diferentes pesquisadores de diferentes áreas.

Outra vantagem de se utilizar de SRA é que os dados de sequência podem ser reanalisados à medida que os bancos de dados são expandidos além disso o pesquisador não precisa necessariamente usar computadores de alta performance.

## 6 CONCLUSÕES

Neste trabalho utilizando dados de SRA disponíveis publicamente foi possível identificar *copaifera langsdorffii*-associated virus 1 em *Copaifera langsdorffii*.

Foi possível realizar com o uso do *software Geneious*, anotação e a caracterização molecular bem como o alinhamento do genoma deste vírus com genomas de referência dentro da família *Betaflexiviridae*.

Através do estudo de relações filogenéticas foi possível verificar que o genoma encontrado se trata de uma nova espécie viral dentro da família *Betaflexiviridae*.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Adams IP, Glover RH, Monger WA, Mumford R, Jackeviciene E, Navalinskiene M, Samuitiene, M, Boonham N (2009) Next-Generation Sequencing and Metagenomic Analysis: A Universal Diagnostic Tool in Plant Virology. *Molecular Plant Pathology*, Oxford, 10: 537–545.
- Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Fauquet CM (2004) Virology Division News: The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. *Archives of Virology* 149:1045-1060. <https://doi.org/10.1007/s00705-004-0304-0>
- Aerts R, Honnay O (2011) Forest Restoration, Biodiversity and Ecosystem Functioning. *BMC Ecology* 11:11–29.
- Aleixo ALP, Albernaz ALKM, Grelle CEV, Vale MM, Rangel TF (2010). Mudanças climáticas e a biodiversidade dos biomas brasileiros: passado, presente e futuro. *Natureza e Conservação* 8:194-196.
- Alfenas AC, Zauza EAV, Mafia RG, de Assis TF (2004) Clonagem e doenças do eucalipto. UFV - Viçosa. p. 442.
- Almeida C, Leite G, Rocha SL, Machado M, Maldonado W (2006) Fenologia e artrópodes de *Copaifera langsdorffii* Desf. no cerrado. *Revista Brasileira de Plantas Mediciniais* 8:64-70.
- MMA – Ministério do Meio Ambiente (2021). <http://www.mma.gov.br/biomas> Consultado em 30 de Agosto de 2021.
- B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*. 37:1530-1534. <https://doi.org/10.1093/molbev/msaa015>
- Barba M, Czosnek H, Hadidi A (2014) Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses* 6:106–36.
- Batista JG (2014) Ocorrência natural de espécies virais e avaliação de potenciais hospedeiras experimentais de begomovirus, potyvirus e tospovirus em plantas arbóreas e berinjela (*Solanum melongena*). (Dissertação de mestrado). Universidade de Brasília 198 p.
- Batista, JG (2020) Viroma de plantas em áreas nativas e cultivadas do Distrito Federal 154 p.
- Biddle PG, Tinsley TW (1971) Poplar mosaic virus in Great Britain. *New Phytologist* 70: 61-66.

- Boccardo G, Luisoni E, Lovisolo D (1973) Characterization of an Italian iso late of poplar mosaic virus. In: Papers, 2nd International Congress of Plant Pathology, Minneapolis.
- Brunt A, Crabtree K, Dallwitz MJ, Gibbs AJ, Watson L (1997) Viruses of plants Wallingford UK CAB International.
- Bushmanova E, Antipov D, Lapidus A, Prjibelski, AD (2019) maSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* 8:giz100 <https://doi.org/10.1093/gigascience/giz100>
- Chaves, Alexandre LR et al. 2003. *Erigeron bonariensis*: an alternative host of Lettuce mosaic virus in Brazil. *Fitopatologia Brasileira* 28: 307-311.
- Çağlayan K, Roumi V, Gazel M, Elçi E, Acioglu M, Mavric Plesko I, Massart S (2019). Identification and characterization of a novel Robigovirus species from sweet cherry in turkey. *Pathogens* 8:57.
- Clover GRG, Pearson MN, Elliott DR, Tang Z, Smales TE, Alexander BJR (2003) Characterization of a strain of Apple stem grooving virus in *Actinidia chinensis* from China. *Plant Pathology* 52:371-378. <https://doi.org/10.1046/j.1365-3059.2003.00857.x>
- Costa FVD, Fagundes M, Nieves FDS (2010) Arquitetura da planta e diversidade de galhas associadas à *Copaiifera langsdorffii* (Fabaceae). *Ecología Austral* 20:9-17.
- Crowther TW, Glick HB, Covey KR, Bettigole C, Maynard DS, Thomas SM, Bradford, MA (2015) Mapping tree density at a global scale. *Nature* 525:201-205.
- Da Fonseca CR (2017) Diversidade de espécies arbóreas e sua relação com o histórico de perturbação antrópica em uma paisagem urbana da Floresta Atlântica. Tese Doutorado 90 p.
- Danilevicz MF, Moharana KC, Venancio TM, Franco LO, Cardoso SR, Cardoso M, Ferreira, PC (2018). *Copaiifera langsdorffii* novel putative long non-coding RNAs: interspecies conservation analysis in adaptive response to different biomes. *Non-coding RNA* 4:27. <https://doi.org/10.3390/ncrna4040027>
- De Vasconcelos AFF, Godinho OES (2002) Uso de métodos analíticos convencionados no estudo da autenticidade do óleo de copaíba. *Química Nova* 25:1057–1060. <https://doi.org/10.1590/S0100-40422002000700002>
- Eisold A-M (2017) Molecular characterization of a new Carlavirus detected in German *Ulmus laevis* (Pall.) by Illumina high throughput sequencing. *Communications in Agricultural and Applied Biological Sciences* 82/13.
- FAO and UNEP (2020) The State of the World's Forests 2020. Forests, biodiversity, and people. *Rome*.

- Freitas-Astua, J, Novaes QSD, Betti JA, Rezende JA, Kuniyuki H, Yuki, VA (2004) Detecção por DAS-ELISA do apple chlorotic leaf spot virus em pétalas desidratadas de macieira mantidas em geladeira por 16 anos. *Fitopatologia Brasileira* 29:572-572.
- GAAFAR YZA e ZIEBELL H (2021) Triviruses - Betaflexiviridae Encyclopedia of Virology Fourth Edition. Academic Press 805-817. <https://doi.org/10.1016/B978-0-12-809633-8.21533-1>.
- Gama, M, Kitagima EW, Ávila AC, Lim MT (1983) Um Carlavirus em seringueira (*Hevea brasiliensis*). *Fitopatologia Brasileira* 8:621.
- Gonthier P e Nicolotti G (2013) Infectious forest diseases Cabi Boston Estados Unidos da América. 641 p.
- Hansen EM, Goheen EM, (2000) *Phellinus weirii* and other nativeroot pathogens as determinants of forest structure and process in western North America. *Annual Review of Phytopathology* 38:515-39.
- Hentsch T, Fuchs E, Grüntzig M (1997) Das strawberry latent ringspot (?) nepovirus (SLRSV) an *Aesculus hippocastanum* L. *Archives of Phytopathology e Plant Protection* 31: 13-23.
- Higa RCV, Cardoso DJ, Andrade GDC, Zanatta, JA, Rossi LMB, Pulrolnik, K, Salis SD (2014) Protocolo de medição e estimativa de biomassa e carbono florestal. Colombo: Embrapa Florestas 1.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2017) UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* 35:518–522. <https://doi:10.1093/molbev/msx281>.
- Instituto Sociedade População e Natureza – ISPN (2021) <https://ispn.org.br/biomas/fauna-e-flora-do-cerrado/>. Consultado em 06 de setembro de 2021
- Jiao J, Kong K, Han J, Song S, Bai T, Song C, Zheng X (2021) Field detection of multiple RNA viruses/viroids in apple using a CRISPR/Cas12a-based visual assay. *Plant Biotechnology Journal*, 19: 394-405.
- Jo Y, Choi H, Kim SM, Kim SL, Lee BC, Cho WK (2016) Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for apple stem grooving virus. *BMC Genomics* 17:1-12. <https://doi.org/10.1186/s12864-016-2994-6>
- Junior VV, Pinto AC (2002). O gênero *copaifera* L. *Química Nova* 25:273-286. <https://doi.org/10.1590/S0100-40422002000200016>



- Kalinowska E, Paduch-Cichal, E, Chodorska M. (2013) First report of Blueberry scorch virus in elderberry in Poland. *Plant Disease* 97: 1515.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh, K. e Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* 30: 772-780. <https://doi.org/10.1093/molbev/mst010>
- Kinoti WM, Constable FE, Nancarrow N, Dann A, Plummer KM, e Rodoni, B. (2017) First Report of Apricot vein clearing-associated virus (AVCaV) in 2 Australia and in a new host, *Prunus cerasifera*. *Plant Disease* 101:1337.
- Kinoti WM, Constable FE, Nancarrow N, Plummer KM, Rodoni B (2018) The incidence and genetic diversity of Apple mosaic virus (ApMV) and Prune dwarf virus (PDV) in *Prunus* species in Australia. *Viruses* 10:136.
- Kircher M, Kelso J (2010) High-Throughput DNA Sequencing--Concepts and Limitations. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, Cambridge 32:524–536.
- Kondo H, Chiba S, Toyoda K, Suzuki N (2013) Evidência de infecção por vírus de RNA de fita negativa em fungos. *Virology* 435:201-209.
- Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database Collaboration (2010). The sequence read archive. *Nucleic Acids Research* 39:D19-D21.
- Li, Z Wang H, Zhao R, Zhang Z, Xia Z, Zhai J, Huang X (2019) Complete genome sequence of a novel Capillovirus infecting *Hevea brasiliensis* in China. *Archives of Virology*, p. 1-4,
- Lin MT, Kitajima EW, Cupertino FP, Costa CL (1979) From a cerrado native plant, cassia *sylvestris*. *Plant Disease Reporter* 63:501.
- Lister RM, Bancroft JB, Nadakavukaren MJ (1965) Some sap-transmissible 25 viruses from apple. *Phytopathology* 55:859.
- Luisoni E, Boccardo G, Milne RG (1976). Purification and some properties of an Italian isolate of poplar mosaic virus. *Journal of Phytopathology* 85: 65-73.
- Ma L, Zeng Q, Huang W, Wang S, Zhang Y, Cheng Y e Jiang F (2021) Incidence of major pome fruit tree viruses and viroids in commercial pear orchards in China and in *Pyrus betulifolia* seedlings *Plant Pathology*.00:1-9.

- Maachi A, Nagata T, Silva JMF (2020) Date palm virus A: first plant virus found in date palm trees. *Virus Genes* 56:792-795. <https://doi.org/10.1007/s11262-020-01801-0>
- Marais A, Murolo S, Faure C, Brans Y, Larue C, Maclot F, Candresse T (2021) Sixty years from the first disease description, a novel badnavirus associated with chestnut mosaic disease. *Phytopathology* PHYTO-09.
- Maree HJ, Fox A, Al Rwahnih M, Boonham N, Candresse T (2018) Application of HTS for routine plant virus diagnostics: state of the art and challenges. *Frontiers in plant science* 9:1082.
- Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Murrell B (2020) RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution* 7:veaa087 <https://doi.org/10.1093/ve/veaa087>
- Menzel P, NG KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257 [https://doi: 10.1038/ncomms11257](https://doi.org/10.1038/ncomms11257)
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37:1530-1534. <https://doi.org/10.1093/molbev/msaa015>
- Mokili JL, Rohwer F, Dutilh BE (2012) Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology* 2:63–77.
- Morellato LPC e Haddad CFB (2000) Introduction: The Brazilian Atlantic Forest. *Biotropica* 32: 786–792.
- Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* 9:e108277. doi:10.1371/journal.pone.0108277
- Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
- NCBI (2021) Sequence Read Archive. <https://www.ncbi.nlm.nih.gov/sra>. Acesso em 15/09/21.
- NCBI (2021) Banco de dados de sequência de montagem Transcriptome Shotgun <https://www.ncbi.nlm.nih.gov/genbank/tsa/>. Acesso em 15/09/21.
- Nery FMB, Melo FL, Boiteux LS, Ribeiro SG, Resende, RO, Orílio, AF, Pereira-Carvalho RC (2020). Molecular characterization of Hovenia Dulcis-Associated Virus 1 (HDaV1) and 2 (HDaV2): new tentative swithin the order Picornavirales. *Viruses* 12:950. <https://doi.org/10.3390/v12090950>

- Pan Y, Birdsey RA, Phillips OL, Jackson RB (2013) The structure, distribution, and biomass of the world's forests. *Annual Review of Ecology, Evolution, and Systematics*, 44:593-622.
- Pieri FA, Mussi MC, Fiorini JE, Schneedorf JM (2010) Efeitos clínicos e microbiológicos do óleo de copaíba (*Copaifera officinalis*) sobre bactérias formadoras de placa dental em cães. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia* 62:578-585. <https://doi.org/10.1590/S0102-09352010000300012>
- Pieri FA, Mussi, MC, Moreira MAS (2009) Óleo de copaíba (*Copaifera* sp.): histórico, extração, aplicações industriais e propriedades medicinais. *Revista Brasileira de Plantas Mediciniais*, 11:465-472.
- Pieri, FA, Mussi MC, Moreira, MAS (2009). Óleo de copaíba (*Copaifera* sp.): histórico, extração, aplicações industriais e propriedades medicinais. *Revista Brasileira de Plantas Mediciniais*, 11:465-472. <https://doi.org/10.1590/S1516-05722009000400016>
- Quemel GKC, Da Costa ABP, Teixeira IF, Machado IN, Machado TN, Machado VSN, De Oliveira NCL (2021) Propriedades medicinais do óleo da *Copaifera langsdorffii*: uma revisão integrativa da literatura. *Brazilian Journal of Health Review* 4:10490-10508. <https://doi.org/10.34119/bjhrv4n3-072>
- Ribeiro JF e Walter BMT (2008). As principais fitofisionomias do bioma Cerrado. *Cerrado: ecologia e flora*, 1:151-212.
- Ribeiro MC, Metzger JP, Martensen AC, Ponzoni, FJ, Hirota MM (2009) The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation* 142:1141–1153.
- Rocha G (2017) Begomovírus em áreas de cerrado: de plantas herbáceas cultivadas a arbóreas selvagens 70 p.
- Rodríguez-Nevado C, Gavilán R, Pagán I (2019) Host abundance and identity determine the epidemiology and evolution of a generalist plant virus in a wild ecosystem. *Phytopathology* 110:94-105.
- Roossinck MJ, Martin DP, Roumagnac P (2015) Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology* 105:716-27.
- Roossinck, M. J. (2012). Plant virus metagenomics: biodiversity and ecology. *Annual review of genetics* 46:359-369.

- Rumbou A, Vainio EJ, Büttner C (2021) Towards the Forest Virome: High-Throughput Sequencing Drastically Expands Our Understanding on Virosphere in Temperate Forest Ecosystems. *Microorganisms* 9:1730. <https://doi.org/10.3390/microorganisms9081730>
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, Washington 74:5463–5467.
- Santos, FMBD (2016) Análise do viroma em espécies arbóreas. 95 p.
- Seabra PV, Rivas EB, Duarte LML, Gatteti SR, Alexandre MAV (2001) Detecção de Carlavirus em *Senna macranthera*. *Revista Brasileira de Horticultura Ornamental* 7:67-71.
- Slik, JWF. et al. (2015) An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences* 112:7472–7477.
- Souza Jr. CM, Shimbo JZ, Rosa MR, Parente LL, Alencar AA, Rudorff BFT, Hasenack H, Matsumoto M, Ferreira LG, Souza-Filho PWM, Oliveira SW, Rocha WF, Fonseca AV, Marques CB, Diniz CG, Costa D, Monteiro D, Rosa ER, Vélez-Martin E, Weber EJ, Lenti FEB, Paternost FF, Pareyn FGC, Siqueira JV, Viera JL, Ferreira Neto LC, Saraiva MM, Sales MH, Salgado MPG, Vasconcelos R, Galano S, Mesquita VV, Azevedo T (2020) Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine Remote Sensing 12:2735. <https://doi.org/10.3390/rs12172735>
- De Souza, EB (2016) Caracterização biológica e molecular de isolados de Apple stem grooving virus e construção de um clone infeccioso. 67 p.
- Subramanian B, Gao S, Lercher MJ, Hu S, Chen, WH (2019) Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Research*. <https://doi:10.1093/nar/gkz357>
- Sweet, J. B. (1980) Fruit tree virus infections of woody exotic and indigenous plants in Britain. XI International Symposium on Fruit Tree Virus Diseases 94 Anais.
- Tappin MRR, Pereira JFG, Lima LA, Siani AC, Mazzei JL, Ramos MFS (2004). Análise química quantitativa para a padronização do óleo de copaíba por cromatografia em fase gasosa de alta resolução. *Química Nova* 27:236–240. <https://doi:10.1590/s0100-40422004000200012>
- TNC - The Nature Conservancy (2021) Mata Atlântica [https://www.tnc.org.br/sobre-a-tnc/onde-trabalhamos/mata-atlantica/?gclid=CjwKCAjwzaSLBhBJEiwAJSRokrF-4\\_Z4L85T9JblX0WPsMkTMbP5HzRL8ppBawQSQHtCSKfct4fq\\_BoCqPkQAvD\\_BwE](https://www.tnc.org.br/sobre-a-tnc/onde-trabalhamos/mata-atlantica/?gclid=CjwKCAjwzaSLBhBJEiwAJSRokrF-4_Z4L85T9JblX0WPsMkTMbP5HzRL8ppBawQSQHtCSKfct4fq_BoCqPkQAvD_BwE)

- Van Lent JWM, Wit AJ, Dijkstra J (1980) Characterization of a Carlavirus in *Elderberry* (*Sambucus* spp.). *Netherlands Journal of Plant Pathology* 86:117-134.
- Vincent SJ, Coutts BA, Jones RAC (2014) Effects of Introduced and Indigenous Viruses on Native Plants: Exploring their disease causing potential at the agro-ecological interface. *PLoS ONE* 9: e91224.
- Walter, BMT (2006) *Fitofisionomias do bioma Cerrado: síntese terminológica e relações florísticas*. Embrapa Recursos Genéticos e Biotecnologia. 389p.

## Anexo 1 – Artigo a ser submetido a revista *Virus Genes*

### Copaifera langsdorffii-associated virus 1: a new *Capillovirus* occurring in copaifera from Brazil

Felipe Fochat Silva Melo<sup>1</sup> Juliana Gabrielle Isidorio da Silva<sup>1</sup> Rita de Cássia Pereira Carvalho<sup>1</sup> Fernando Lucas Melo<sup>1</sup>

<sup>1</sup> Departamento de Fitopatologia, Universidade de Brasília, Laboratório de Virologia Vegetal, Brasília-DF 70910-900, Brazil.

#### Resumo

Uma sequência viral completa foi recuperada de transcriptoma de *Copaifera langsdorffii* coletada em Brasília (Brasil) usando *Sequence Read Archive* – SRA disponível publicamente no National Center for Biotechnology Information (NCBI). O genoma viral apresentou 6427 nucleotídeos (nt) e com organização genômica típica de membros da família *Betaflexiviridae*. Duas ORFs (Open Reading Frames) denominadas ORF1 e ORF2, codificando uma poliproteína e uma proteína de movimento MP, respectivamente foram recuperadas. A análise filogenética utilizando o genoma completo mostrou que o vírus se agrupou com membros do gênero *Capillovirus*. O novo vírus foi denominado “copaifera langsdorffii-associated virus 1”. Este consiste no primeiro vírus relatado em uma árvore de copaíba no mundo. Este trabalho estabelece uma base para a prospecção viral em espécies arbóreas que tem dados escassos na literatura e do potencial de arquivos SRA disponíveis publicamente para prestação de informações e descoberta de novas espécies virais.

**Palavras-chave:** *Betaflexiviridae*, *Fabaceae*, new virus, SRA

A família *Betaflexiviridae* é composta por duas subfamílias com 13 gêneros: *Quinvirinae* (*Carlavirus*, *Foveavirus* e *Robigovirus*) e *Trivirinae* (*Capillovirus*, *Chordovirus*, *Citrivirus*, *Divavirus*, *Prunavirus*, *Ravavirus*, *Tepovirus*, *Trichovirus*, *Vitivirus* e *Wamavirus*). Juntas, as duas subfamílias possuem 108 espécies [1]. As subfamílias *Quinvirinae* e *Trivirinae* diferem pelos tipos de suas proteínas de movimento viral. Enquanto os vírus da subfamília *Quinvirinae* codificam proteínas de movimento pelo triplo bloco gênico (TGB), os vírus da subfamília *Trivirinae* codificam proteínas de movimento semelhantes à proteína de 30K de TMV [2]. Os membros de *Betaflexiviridae* possuem vírus com partículas filamentosas flexuosas com um comprimento de 620-680 nm e largura 12nm não envelopados [3,4] e genoma de ssRNA (+) variando de 5,9 a 9,0 kb.[5].

Os vírus da família *Betaflexiviridae* apresentam distribuição geográfica e gama de hospedeiras variadas. Em uma revisão recente foram listados 11 vírus da família *Betaflexiviridae* infectando árvores e arbustos florestais em ecossistemas temperados [6]. Uma nova espécie no gênero *Capillovirus*, por exemplo, foi descrito em *Betula* sp. provenientes de amostras da Alemanha [7]. Por outro lado, *Apple stem grooving virus*, espécie tipo do gênero *Capillovirus*, possui isolados que ocorrem em árvores frutíferas, rosáceas, rutáceas do gênero *citrus*, plantas de lírio e, também, no kiwi, além de plantas utilizadas em ensaios biológicos como

indicadoras de vírus, incluindo: *Chenopodium*, *Cucumber*, *Nicotiana*, *Phaseolus* e *Vigna* [4]. Em 2015, foi descrito na Índia a ocorrência natural do apple stem grooving virus - ASGV em *Bauhinia variegata* (família Fabaceae) apresentando sintomas severos como clorose, manchas cloróticas e necróticas, manchas, distorção foliar e caneluras na haste [8].

*Copaifera* spp. são árvores pertencentes a família Fabaceae, que apresentam crescimento lento podendo alcançar 40 metros de altura e viver em torno de 400 anos. As espécies são nativas da região tropical da América Latina e África Ocidental. O gênero possui 72 espécies, sendo que 16 delas são encontradas apenas do Brasil [9]. Uma espécie que se destaca no gênero é a *Copaifera langsdorffii* Desf. (Brasil, Argentina e Paraguai), por estar amplamente distribuída em todo o território brasileiro e por ser uma das plantas mais estudadas do gênero [9,10], além de possuir grande importância econômica pelos produtos florestais madeireiros e não madeireiros, em especial o óleo-resina.

O óleo-resina da copaíba possui efeitos medicinais comprovados pela ciência em humanos e animais [11,12] devido a ação anti-inflamatória, antimicrobiana, cicatrizante, diurética, inibidor tumoral e laxante. Este óleo pode também ser utilizado por via tópica e oral de acordo com o tipo terapêutico e a espécie de copaíba empregada [10], sendo um excelente fixador para perfumes e possui propriedades emolientes, o que faz ser usado como matéria prima na produção de cosméticos para cabelo e pele [9]. Em razão da viscosidade da resina produzida pela *C. langsdorffii*, o que confere alta durabilidade nos produtos gerados, a mesma é bastante aproveitada na produção de vernizes, tinturas e lacas [13,14]. Devido ao potencial tecnológico da madeira é indicada para uso industrial na construção civil, naval, moveleira e energética [11].

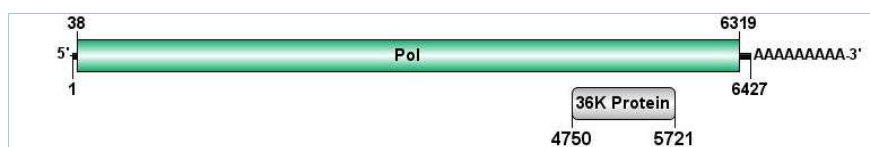
Em geral, estudos em espécies arbóreas sendo acometidas por novas espécies virais ou espécies já conhecidas ainda são emergentes e na grande maioria limitadas a espécies nos gêneros de importância econômica, como: *Prunus*, *Pyrus*, *Malus*, *Citrus*, *Actinia*, *Diospyros*, *Morus* e *Vitis* [15]. Na literatura não existe, até o momento, relato de vírus ocorrendo em *C. langsdorffii*

Para este estudo, foram utilizados dois conjuntos de dados SRA disponíveis publicamente, provenientes de uma pesquisa em amostras de *C. langsdorffii* para avaliar lncRNA's putativos e analisar a conservação interespecies em respostas adaptativas para diferentes biomas. Foram analisados 83,2 milhões de *reads*, correspondentes ao bioma Cerrado do BioProject SRX6060683 - CER 1, e 84,3 milhões de *reads*, correspondentes a amostras do bioma Mata Atlântica BioProject SRX6060682 - ARF 1, utilizando como ferramenta de download (SRA Toolkit para Linux em distribuição Ubuntu), baseados no sequenciamento de Illumina HiSeq 2500 [16].

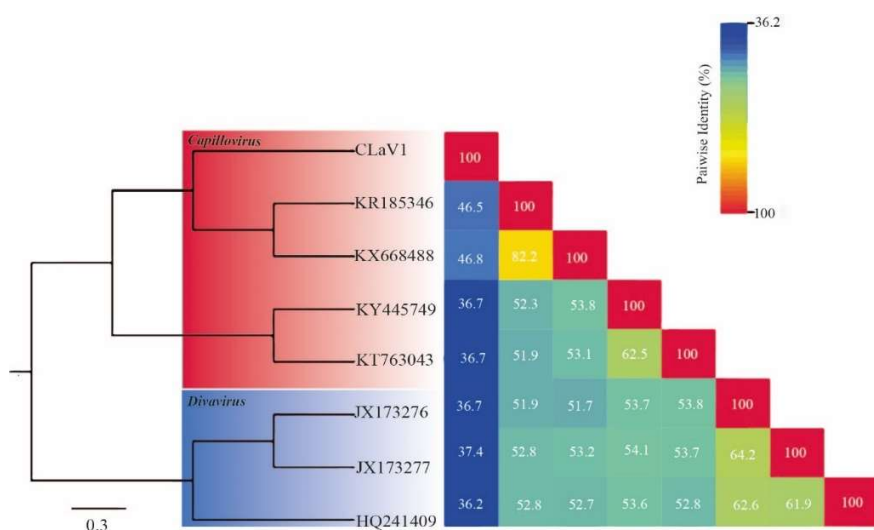
As sequências de cada biblioteca foram montadas usando o programa rnaSPAdes v 3.14.1 [17]. Após a montagem as sequências passaram por predição taxonômica usando o Kaiju v 1.8 [18], com os parâmetros padrões de classificação. A partir dessa análise, foram separadas aquelas de taxonomia predita como vírus e selecionadas as sequências de maior tamanho. As sequências virais selecionadas foram submetidas ao BLASTn e BLASTx, utilizando o programa Geneious v R11.1 que também permitiu realizar a montagem do genoma viral, anotação e o alinhamento MAFFT v 7.450 [19] da sequência montada contra os genomas completos de referência.

Na filogenia foi utilizado o software IQTREE v 2.1.3 [20] onde foram testados 277 modelos – Model finder [21] para a construção das relações filogenéticas e seleção do melhor modelo na construção da árvore filogenética utilizando um intervalo de confiança de 95% UFBoot2 [22] com bootstrap de 1000 repetições, para a demarcação de espécie foi utilizada a ferramenta SDT v 1.2 [23] e Evolview v 3[24] para relacionar os dados e geração da **Figura 2**. Também foi utilizando o RDP5 para verificar eventos de recombinação[25].

Após a montagem e anotação do genoma pelo Geneious v R11.1, **Figura 1**, foi recuperado um genoma completo de 6427 nt com estrutura e características típicas de *Capillovirus*, verificado por meio de BLASTn e BLASTx, com base no critério de demarcação de novas espécies estabelecido pelo ICTV para família Betaflexiviridae e para verificação do status taxonômico deste genoma, apresentou identidade 64,43% em nt e identidade 63,94 % em aa (ORF1-Poliproteína) e identidade 52.13% em aa (ORF 2-Proteína de movimento viral do tipo 30K), utilizando o genoma de referência NC\_001749, e por isso o mesmo foi qualificado como uma nova espécie viral. A análise utilizando o IQTREE v 2.1.3 [20] selecionou o melhor modelo BIC: GTR + F + I + G4 em matriz de substituição nucleotídica e construiu a árvore filogenética neste modelo com base no alinhamento MAFFT v 7.450 [19]. Análise filogenética usando sequências de genoma completas de espécies representativas de *Capillovirus* e *Divavirus* e o genoma de “copaifera langsdorffi-associated virus 1” indicou alta relação com o gênero *Capillovirus*. Não foram detectados eventos fortes de recombinação em RDP5 [25] no genoma de “copaifera langsdorffi-associated virus 1”.



**Figura 1.** Organização genômica do vírus encontrado, Pol - Poliproteína codificando polymerase, replicase e capa proteica e Proteína 36K responsável pelo movimento viral.





**Figura 2.** Relações Filogenéticas de CLaV1 copaifera langsdorffii associated virus 1 e análise SDT utilizando 8 representantes da família *Betaflexiviridae*. A. Árvore de máxima verossimilhança com base em sequências completas do genoma. Os números nos nós representam probabilidades posteriores. Sequências de *Divavirus* foram usadas como grupos externos. As sequências completas do genoma foram alinhadas usando MAFFT v 7.450 , e a árvore filogenética foi construída usando a árvore de máxima verossimilhança realizada com IQTREE , com substituição de aminoácidos selecionada por IQTREE v 2.1.3 GTR + F + I + G4. B. Matriz de pares em todo o genoma gerada por SDT v1.2 . A ordem, o número de acesso e os nomes dos isolados (GenBank) são: *Capillovirus*: 1. CLaV1- “copaifera langsdorffii-associated virus 1” 2. KR185346 - Apple stem grooving virus 3.KX668488 - Apple stem grooving virus 4. KY445749 - Cherry virus A 5. KT7663043 - Currant virus A e *Divavirus*: 6. JX173276 - Diuris virus A 7. JX173277 - Diuris virus B 8 e HQ241409 - Hardenbergia virus A.

Neste trabalho utilizando dados de SRA disponíveis publicamente foi possível identificar copaifera langsdorffii-associated virus 1 em *Copaifera langsdorffii*.

Foi possível realizar com o uso do software Geneious, anotação e a caracterização molecular bem como o alinhamento do genoma deste vírus com genomas de referência dentro da família *Betaflexiviridae*.

Através do estudo de relações filogenéticas foi possível verificar que o genoma encontrado se trata de uma nova espécie viral dentro da família *Betaflexiviridae*.

## Referências

1. Fu S, Zhang T, He M, Sun B, Zhou X, Wu, J (2021) Molecular characterization of a novel wheat-infecting virus of the family *Betaflexiviridae*. *Archives of Virology* 166:2875-2879. <https://doi.org/10.1007/s00705-021-05175-y>
2. Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Fauquet CM (2004) *Virology Division News: The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation*. *Archives of virology* 149:1045-1060. <https://doi.org/10.1007/s00705-004-0304-0>
3. Clover GRG, Pearson MN, Elliott DR, Tang Z, Smales TE, Alexander BJR (2003) Characterization of a strain of Apple stem grooving virus in *Actinidia chinensis* from China. *Plant Pathology* 52:371-378. <https://doi.org/10.1046/j.1365-3059.2003.00857.x>
4. Jo Y, Choi H., Kim SM, Kim SL, Lee BC, Cho WK (2016) Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for apple stem grooving virus. *BMC Genomics* 17:1-12. <https://doi.org/10.1186/s12864-016-2994-6>
5. Maachi A, Nagata T, Silva JMF (2020) Date palm virus A: first plant virus found in date palm trees. *Virus Genes* 56:792-795. <https://doi.org/10.1007/s11262-020-01801-0>
6. Rumbou A, Vainio EJ, Büttner C (2021) Towards the forest virome: High-Throughput Sequencing drastically expands our understanding on virosphere in temperate forest ecosystems. *Microorganisms* 9:1730. <https://doi.org/10.3390/microorganisms9081730>
7. Rumbou, A., Candresse, T., Marais, A., Svanella-Dumas, L., Landgraf, M., von Bargen, S., e Büttner, C. (2020). Unravelling the virome in birch: RNA-Seq reveals a complex of known and novel viruses. *Plos One* 15: e0221834. <https://doi:10.1371/journal.pone.0221834>

8. Bhardwaj P, Ram R, Zaidi AA, Hallan V (2015) Natural occurrence of Apple stem grooving virus on *Bauhinia variegata*. *Trees* 29:1415-1422. <https://doi.org/10.1007/s00468-015-1219-y>
9. Junior VV, Pinto AC (2002). O gênero *Copaifera* L. *Química Nova* 25:273-286. <https://doi.org/10.1590/S0100-40422002000200016>
10. Quemel GKC, Da Costa ABP, Teixeira IF, Machado IN, Machado TN, Machado VSN, De Oliveira, NCL (2021) Propriedades medicinais do óleo da *Copaifera langsdorffii*: uma revisão integrativa da literatura. *Brazilian Journal of Health Review* 4:10490-10508. <https://doi.org/10.34119/bjhrv4n3-072>
11. Pieri, FA, Mussi MC, Moreira, MAS (2009). Óleo de copaíba (*Copaifera* sp.): histórico, extração, aplicações industriais e propriedades medicinais. *Revista Brasileira de Plantas Medicinai*s, 11:465-472. <https://doi.org/10.1590/S1516-05722009000400016>
12. Pieri FA, Mussi MC, Fiorini JE, Schneedorf JM (2010). Efeitos clínicos e microbiológicos do óleo de copaíba (*Copaifera officinalis*) sobre bactérias formadoras de placa dental em cães. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia* 62:578-585. <https://doi.org/10.1590/S0102-09352010000300012>
13. De Vasconcelos AFF, Godinho OES (2002) Uso de métodos analíticos convencionados no estudo da autenticidade do óleo de copaíba. *Química Nova* 25:1057–1060. <https://doi.org/10.1590/S0100-40422002000700002>
14. Tappin MRR, Pereira JFG, Lima LA, Siani AC, Mazzei JL, Ramos MFS (2004). Análise química quantitativa para a padronização do óleo de copaíba por cromatografia em fase gasosa de alta resolução. *Química Nova* 27:236–240. <https://doi.org/10.1590/s0100-40422004000200012>
15. Nery FMB, Melo FL, Boiteux LS, Ribeiro SG, Resende, RO, Orílio, AF, Pereira-Carvalho RC (2020). Molecular characterization of Hovenia Dulcis-Associated Virus 1 (HDaV1) and 2 (HDaV2): new tentative swithin the order Picornavirales. *Viruses* 12:950. <https://doi.org/10.3390/v12090950>
16. Danilevicz MF, Moharana KC, Venancio TM, Franco LO, Cardoso SR, Cardoso M, Ferreira, PC (2018). *Copaifera langsdorffii* novel putative long non-coding RNAs: interspecies conservation analysis in adaptive response to different biomes. *Non-coding RNA* 4:27. <https://doi.org/10.3390/ncrna4040027>
17. Bushmanova E, Antipov D, Lapidus A, Prjibelski, AD (2019) rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* 8:giz100 <https://doi.org/10.1093/gigascience/giz100>
18. Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257 <https://doi.org/10.1038/ncomms11257>
19. Katoh, K., e Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772-780. <https://doi.org/10.1093/molbev/mst010>
20. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37:1530-1534. <https://doi.org/10.1093/molbev/msaa015>

21. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. <https://doi:10.1038/nmeth.4285>
22. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2017) UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35:518–522. <https://doi:10.1093/molbev/msx281>
23. Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* 9(9): e108277. <https://doi:10.1371/journal.pone.0108277>
24. Subramanian B, Gao S, Lercher MJ, Hu S, Chen, WH (2019) Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Research*. <https://doi:10.1093/nar/gkz357>
25. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Murrell B (2020) RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution* 7:veaa087 <https://doi:10.1093/ve/veaa087>