

**SECRETARIA DE AGRICULTURA E ABASTECIMENTO DO ESTADO
DE SÃO PAULO**

**AGÊNCIA PAULISTA DE TECNOLOGIA DOS AGRONEGÓCIOS
INSTITUTO BIOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM SANIDADE, SEGURANÇA
ALIMENTAR E AMBIENTAL NO AGRONEGÓCIO**

**VARIABILIDADE GENÔMICA E EMERGÊNCIA DE PEQUENAS ORFs NO
RNA2 DE PASSION FRUIT GREEN SPOT VIRUS**

ISABELA GUARNIERI FURIATO GAROFALO LEÃO

Dissertação apresentada para a obtenção do
título de Mestre em Sanidade, Segurança
Alimentar e Ambiental no Agronegócio.
Área de concentração: Segurança
Alimentar e Sanidade no Agroecossistema.

SÃO PAULO

2025

ISABELA GUARNIERI FURIATO GAROFALO LEÃO

**VARIABILIDADE GENÔMICA E EMERGÊNCIA DE PEQUENAS ORFs NO
RNA2 DE PASSION FRUIT GREEN SPOT VIRUS**

Dissertação apresentada para a obtenção do
título de Mestre em Sanidade, Segurança
Alimentar e Ambiental no Agronegócio.

Área de concentração: Segurança
Alimentar e Sanidade no Agroecossistema.

Orientadora: Juliana Freitas-Astúa.

Coorientador: Pedro Luis Ramos-
González.

SÃO PAULO

2025

Eu **Isabela Guarnieri Furiato Garofalo Leão**, autorizo o Instituto Biológico (IB-APTA), da Secretaria de Agricultura e Abastecimento do Estado de São Paulo, a disponibilizar gratuitamente e sem ressarcimento dos direitos autorais, o presente trabalho acadêmico de minha autoria, no portal, biblioteca digital, catálogo eletrônico ou qualquer outra plataforma eletrônica do IB para fins de leitura, estudo, pesquisa e/ou impressão pela Internet desde que citada a fonte.

Assinatura: Isabela Leão Data 17/11/25

Dados Internacionais de Catalogação na Publicação (CIP)
Secretaria de Agricultura e Abastecimento do Estado de São Paulo
Núcleo de Informação e Documentação – IB

Leão, Isabela Guarnieri Furiato Garofalo.

Variabilidade genômica e emergência de pequenos ORFs no RNA2 passion fruit green spot vírus. / Isabela Guarnieri Furiato Garofalo Leão. - São Paulo, 2025.

62 p.

doi: 10.31368/PGSSAAA.2025D.IL04

Dissertação (Mestrado). Instituto Biológico (São Paulo). Programa de Pós-Graduação.

Área de concentração: Segurança Alimentar e Sanidade no Agroecossistema.

Linha de pesquisa: Sanidade, gestão ambiental e qualidade de alimentos, produtos e processos na produção agropecuária sustentável.

Orientador: Juliana Freitas-Astúa

Coorientador: Pedro L. Ramos-González

Versão do título para o inglês: Genomic variability and emergence of small ORFs in the RNA2 of passion fruit green spot virus.

1. Vírus transmitidos por *Brevipalpus* 2. *Cilevirus* 3. Pequenas ORFs
4. ORFs órfãs I. Leão, Isabela Guarnieri Furiato Garofalo II. Freitas-Astúa, Juliana III. Instituto Biológico (São Paulo) IV. Título.

IB/Bibl./2025/04

AGRADECIMENTOS

Em primeiro lugar, à minha família: minha mãe Angélica, meu pai Álvaro, meus irmãos Felipe e Gabriel, meus avós Nerci e Sebastiao, minhas cunhadas Camila e Lindsay, por todo o amor, compreensão e incentivo incondicional. Aos meus sobrinhos, Thiago, Enzo, Heloísa e João Felipe, que, mesmo sem compreenderem exatamente o que eu faço, sempre foram um refúgio de paz nos momentos em que mais precisei.

Ao Instituto Biológico de São Paulo, minha gratidão pela estrutura, acolhimento e oportunidades proporcionadas, bem como à Coordenação de Aperfeiçoamento de Pessoa de Nível Superior (CAPES) pela apoio financeiro por meio da concessão da bolsa de mestrado (maio/2023 – abril/2025, processo 88887.847580/2023-00), fundamental para a realização deste trabalho.

Aos meus orientadores Dra. Juliana Freitas-Astúa e Dr. Pedro L. Ramos-González, meu sincero agradecimento pela paciência, pela partilha generosa de conhecimento e orientação valiosa durante todas as etapas desta pesquisa.

Aos amigos e colegas de laboratório, deixo um agradecimento especial. Obrigada por estarem ao meu lado em tantos momentos, dividindo os desafios, as conquistas e os dias difíceis. Por cada risada compartilhada, pelos ótimos cafezinhos com bolo, pelos abraços sinceros e pelas conversas que aliviaram o peso da rotina. A amizade de vocês foi um dos pilares mais bonitos desta caminhada, e levo comigo cada lembrança com profundo carinho.

RESUMO

GAROFALO-LEÃO, Isabela Guarnieri Furiato. **Variabilidade genômica e emergência de pequenas ORFs no RNA2 de passion fruit green spot virus.** Dissertação (Mestrado em Sanidade, Segurança Alimentar e Ambiental no Agronegócio) – Instituto Biológico, Agência Paulista de Tecnologia dos Agronegócios, Secretaria da Agricultura e Abastecimento do Estado de São Paulo, São Paulo, 2025.

O passion fruit green spot virus (PfGSV), pertencente ao gênero *Cilevirus* (família *Kitaviridae*), é um vírus fitopatogênico associado a surtos históricos na cultura do maracujazeiro e capaz de infectar também espécies ornamentais. Estudos recentes indicam que o extremo 5' do RNA2 do PfGSV representa uma região de plasticidade genômica, abrigando pequenas ORFs cuja emergência pode estar relacionada a processos adaptativos envolvendo tanto o hospedeiro quanto o vetor *Brevipalpus*. Neste trabalho, buscou-se caracterizar a variabilidade genética dessa região, com ênfase na prevalência, organização e possíveis funcionalidades das ORFs *p13*, *p11*, $\Delta p11$ e *p7*. Para tanto, foram analisados 27 isolados coletados entre 2013 e 2023, abrangendo 23 amostras do Brasil, duas da Colômbia e duas do Paraguai, a partir de sequências obtidas por sequenciamento Sanger, dados prévios de HTS (high throughput sequencing) gerados pelo grupo de pesquisa e informações depositadas no GenBank. As análises compreenderam detecção molecular por RT-PCR, clonagem, e sequenciamento Sanger. Abordagens bioinformáticas foram utilizadas para a montagem e alinhamento das sequências, buscas de homologia, predições de domínios funcionais, modelagem estrutural por AlphaFold e avaliação do contexto de Kozak para as diferentes ORFs. Os resultados revelaram que o alelo *p13* foi o mais prevalente, detectado em 21 isolados distribuídos em diferentes regiões e hospedeiros, enquanto *p11* ocorreu em apenas um isolado proveniente da Bahia e $\Delta p11$ em três isolados, dois do Paraguai e um de São Paulo. As análises demonstraram que esses três alelos representam uma série de truncamentos que interromperam a ORF, sugerindo relação evolutiva entre eles. No caso da ORF *p7* foram identificados seis alelos, com variação de comprimento decorrente de mutações no códon de início e de mutações sem sentido que introduziram códons de parada prematuros. A investigação dos contextos de Kozak indicou padrões fortemente favoráveis para *p7* e eficiência moderada para *p15*.

e *p13*, reforçando a hipótese de que a tradução dessas pequenas ORFs poderia ser modulada de maneira adaptativa. Predições funcionais apontaram para a presença de sítios conservados de glicosilação na ORF *p13* e seus alelos, além de similaridade remota de *p11* com domínios nucleases, sugerindo potenciais papéis funcionais na interação vírus-hospedeiro. De forma geral, os achados indicam que o extremo 5' do RNA2 do PfGSV é uma região com alta variabilidade, podendo atuar como reservatório para o surgimento de ORFs emergentes. A caracterização dessa plasticidade genômica reforça a importância de pequenas proteínas acessórias na adaptação e evolução viral, apontando para possíveis mecanismos de manutenção do ciclo infeccioso e de otimização da transmissão pelo vetor. O trabalho contribui para a compreensão da diversidade genômica do PfGSV, fornecendo subsídios relevantes para estudos futuros de epidemiologia e evolução dos cilevirus.

Palavras-chave: vírus transmitidos por *Brevipalpus*, *Cilevirus*, pequenas ORFs, ORFs órfãs.

ABSTRACT

GAROFALO-LEÃO, Isabela Guarnieri Furiato. **Genomic variability and emergence of small ORFs in the RNA2 of passion fruit green spot virus.** Thesis (Masters in Plant and Animal Health, Food and Environmental Safety in the Agribusiness) – Instituto Biológico, Agência Paulista de Tecnologia dos Agronegócios, Secretaria da Agricultura e Abastecimento do Estado de São Paulo, São Paulo, 2025.

Passion fruit green spot virus (PfGSV), belonging to the genus Cilevirus (family Kitaviridae), is a plant pathogenic virus associated with historical outbreaks in passion fruit crops and capable of infecting ornamental species. Recent studies indicate that the 5' end of PfGSV RNA2 represents a region of genomic plasticity, harboring small ORFs whose emergence may be related to adaptive processes involving both the host and the *Brevipalpus* vector. This study sought to characterize the genetic variability of this region, with an emphasis on the prevalence, organization, and possible functions of the ORFs p13, p11, Δ p11, and p7. Twenty-seven isolates collected between 2013 and 2023 were analyzed, including 23 samples from Brazil, two from Colombia, and two from Paraguay. These isolates were based on Sanger sequencing, previous HTS (High-Throughput Sequencing) data generated by the research group, and information deposited in GenBank. The analyses included molecular detection by RT-PCR, cloning, and Sanger sequencing. Bioinformatic approaches were used for sequence assembly and alignment, homology searches, functional domain predictions, structural modeling by AlphaFold, and Kozak context evaluation for the different ORFs. The results revealed that the p13 allele was the most prevalent, detected in 21 isolates distributed across different regions and hosts, while p11 occurred in only one isolate from Bahia and Δ p11 in three isolates, two from Paraguay and one from São Paulo. The analyses demonstrated that these three alleles represent a series of truncations that interrupted the ORF, suggesting an evolutionary relationship between them. In the case of ORF p7, six alleles were identified, with length variation resulting from mutations in the start codon and nonsense mutations that introduced premature stop codons. The investigation of Kozak contexts indicated strongly favorable patterns for p7 and moderate efficiency for p15 and p13, reinforcing the hypothesis that the translation of these small ORFs could be adaptively modulated. Functional predictions indicated the presence of conserved glycosylation sites in the p13 ORF and its alleles, as well as remote similarity of p11 to nuclease domains, suggesting

potential functional roles in virus-host interactions. Overall, the findings indicate that the 5' end of PfGSV RNA2 is a region of high variability and may act as a reservoir for the emergence of emerging ORFs. The characterization of this genomic plasticity reinforces the importance of small accessory proteins in viral adaptation and evolution, pointing to possible mechanisms for maintaining the infectious cycle and optimizing transmission by the vector. This work contributes to the understanding of PfGSV genomic diversity, providing relevant information for future studies on the epidemiology and evolution of cileviruses.

Keywords: *Brevipalpus*-transmitted viruses, *Cilevirus*, small ORFs, orphan ORFs.

LISTA DE ILUSTRAÇÕES

Figura 1. A – Sintomas da infecção por PfGSV em orquídeas do gênero <i>Eria</i> (isolado PfGSV_Lim01). B – Sintomas de infecção por PfGSV em folhas, ramos e frutos de maracujazeiro (Ramos-González <i>et al.</i> , 2020). C – Micrografia de <i>Brevipalpus yotheri</i> coletado em uma amostra de orquídea originária de Limeira infectada com PfGSV_Lim01 (colaboração com Dra. Aline Tassi). D-a , D-b e D-c – Partículas baciliformes observadas em cisternas do retículo endoplasmático de células de maracujazeiro infectadas pelo PfGSV (Kitajima <i>et al.</i> , 2003a; Antonioli-Luizon, 2009).	14
Figura 2. Representação esquemática da organização genômica de PfGSV. Adaptado de Ramos-González <i>et al.</i> , 2020; 2021.	15
Figura 3. Esquema da região 5' do RNA2 de PfGSV, indicando a localização relativa dos iniciadores e os tamanhos esperado dos fragmentos amplificados.	24
Figura 4. Distribuição geográfica das amostras de PfGSV coletadas entre 2013 e 2023, incluindo 23 isolados provenientes do Brasil, 2 isolados da Colômbia e 2 isolados do Paraguai. O conjunto abrange amostras recentemente sequenciadas por Sanger neste trabalho, sequências previamente geradas por HTS por nosso grupo de pesquisa (ASU03, Prb11, Cmp02 e Lim01) e sequências depositadas no GenBank.	28
Figura 5. Alinhamento da sequência de aminoácidos das proteínas Δ P11, P11 e P13 dos isolados de PfGSV. O alinhamento foi realizado utilizando o software Vector NTI, evidenciando as regiões conservadas. Os aminoácidos destacados em amarelo dizem respeito às posições altamente conservadas em todos os isolados; em azul, estão representados os aminoácidos com conservação parcial e divergências entre os isolados; em verde, estão destacadas as substituições pontuais entre os isolados.	30
Figura 6. Alinhamento da sequência de nucleotídeos das ORFs <i>p11</i> _BJL01 e <i>p13</i> _Snp01, realizado com o software Vector NTI. As regiões conservadas entre as sequências estão destacadas em amarelo. O ponto em que ocorre a adição de uma uracila, responsável pela introdução de um códon de parada prematuro, está destacado na imagem.	34
Figura 7. Alinhamento das sequências de nucleotídeos das ORFs <i>p13</i> _Snp01 e Δ <i>p11</i> _Lim01, realizado no Vector NTI. As regiões conservadas são estacadas em amarelo, observa-se na sequência de Δ <i>p11</i> _Lim01 a presença de uma mutação sem sentido que introduz um códon de parada prematuro.	35

Figura 8. Alinhamento das sequências de nucleotídeos das ORFs *p13_Snp01* e *Δp11_ASU01*, realizado no Vector NTI. As regiões conservadas estão destacadas em amarelo, e o ponto de adição de uma adenina, responsável pela introdução de um códon de parada prematuro na sequência de *Δp11_ASU01*, está assinalado na imagem. 36

Figura 9. Alinhamento das sequências de aminoácidos das proteínas P7 de PfGSV, realizado no Vector NTI. Os resíduos destacados em amarelo correspondem às posições altamente conservadas entre todos os isolados; em azul, indicam conservação parcial e divergências entre os isolados; e em verde, estão assinaladas substituições pontuais entre os isolados. 38

Figura 10. Modelos estruturais preditos das proteínas P7 de diferentes isolados do PfGSV, gerados com o AlphaFold 3. As representações ilustram as estruturas tridimensionais previstas das proteínas P7 e os respectivos comprimentos em aminoácidos. As cores correspondem ao nível de confiança da predição. 39

Figura 11. Alinhamento das sequências de nucleotídeos de alelos da ORF *p7*, realizado no Vector NTI. **A** – Comparação entre os isolados *p7_ASU01* e *p7_Snp01*, destacando mutações pontuais que introduziram códon de início e de parada. **B** – Comparação entre os isolados *p7_Pr07* e *p7_Snp01*, evidenciando mutações pontuais responsáveis pela adição de códon de início e de parada. As regiões conservadas estão indicadas em amarelo, e as alterações específicas estão assinaladas em suas respectivas posições. 40

Figura 12. WebLogos criados a partir dos nucleotídeos -6 e +6 flanqueadores dos códon de início das ORFs no genoma de PfGSV. **A, B, F, G e H**: criados a partir da sequência de 11 isolados (CYPe-246, Snp01, NFO01, Snp02, BJL01, BSB01, RsTHr-269, Prb11, Lim01, ASU03 e Cmp02). **C, D e E**: criados a partir da sequência dos 27 isolados descritos no trabalho. 42

Figura 13. Representação esquemática das seis organizações genômicas identificadas no extremo 5' do RNA2 entre as ORFs p15 e p61 nos 27 isolados analisados. As variações refletem diferentes combinações dos alelos p13, p11 e *Δp11*, bem como das formas p7 e *Δp7*, resultando em distintos arranjos genéticos. 43

LISTA DE TABELAS

Tabela 1. Iniciadores específicos utilizado para a detecção de PfGSV e amplificação da região de interesse. Os iniciadores referentes ao RNA1 incluem regiões

como *RdRp* e *p29*, enquanto os do RNA2 abrangem *MP*, *p24* e o extremo 5' do RNA2 de PfGSV. 23

Tabela 2. Lista de isolados de PfGSV analisados neste estudo. A nomenclatura dos isolados refere-se ao local de coleta de cada uma das amostras. 26

Tabela 3. Mutações sinônimas e não sinônimas identificadas nas sequências da ORF *p13*. Os códons marcados por um “*” e dois “**” (asteriscos) representam mais de uma mutação na mesma trinca em diferentes isolados. Os códons foram numerados a partir do códon de início AUG (1° códon). Estão sinalizados, em vermelho, os nucleotídeos que sofreram mutações. 29

Tabela 4. Percentual de identidade nucleotídica entre os alelos $\Delta p11$ nos isolados ASU01, ASU03 e Lim01. 32

Tabela 5. Prevalência dos alelos da ORFs *p7* entre os isolados de PfGSV. 37

Sumário

1	Introdução	5
2	Revisão Bibliográfica	9
2.1	O maracujazeiro: diversidade biológica, economia e desafios na produção....	9
2.2	Gênero <i>Cilevirus</i> : caracterização genômica e relevância fitopatológica	10
2.2.1	Ácaros do gênero <i>Brevipalpus</i> como vetores de fitovírus	11
2.3	Passion fruit green spot virus: biologia, epidemiologia e impactos agrícolas	12
2.4	Pequenas proteínas acessórias virais e suas funções acessórias adaptativas .	16
2.5	Dinâmica evolutiva das ORFs virais.....	18
3	Objetivo geral	21
3.1	Objetivos específicos	21
4	Material e métodos.....	22
4.1	Amostras vegetais: origem, conservação e processamento	22
4.1.1	Processamento de amostras vegetais <i>in natura</i>	22
4.1.2	Processamento de material vegetal de maracujá conservados em ultrafreezer	22
4.2	Reação de transcrição reversa	22
4.3	Deteção de passion fruit green spot virus	23
4.4	Clonagem molecular e sequenciamento.....	23
4.5	Sequenciamento Sanger dos fragmentos clonados	24
4.6	Análises bioinformáticas.....	25
4.7	Isolados analisados.....	26
5	Resultados.....	28
5.1	Caracterizações de novos isolados de PfGSV.....	28
5.2	Variabilidade genética das ORFs <i>p13</i> , <i>p11</i> e $\Delta p11$ de PfGSV	28
5.2.1	O alelo <i>p13</i> é o prevalente entre o conjunto de amostras analisadas	29
5.2.2	Alelo <i>p11</i> está presente em um único isolado.....	31

5.2.3	Alelo $\Delta p11$	32
5.3	Estrutura da ORF entre os alelos $p13$, $p11$ e $\Delta p11$	33
5.4	ORFs $p7$: variabilidade genética e possível conservação funcional	36
5.5	Análises de sequência Kozak	41
6	Discussão	43
7	Considerações finais	49
8	Conclusões	51
9	Referências	52

1 INTRODUÇÃO

O Brasil é o maior produtor mundial de maracujá, fruto de relevância econômica, ambiental e alimentar. Sua importância decorre tanto de suas múltiplas formas de consumo quanto ao papel do cultivo do maracujazeiro (*Passiflora* spp.), planta do gênero *Passiflora*, família Passifloraceae, em diversas regiões do país. Contudo, a produção enfrenta desafios significativos devido à incidência de diferentes patógenos. Entre as doenças fúngicas, destacam-se a antracnose, causada por *Colletotrichum* sp., que gera lesões marrons em folhas e frutos, e a fusariose, provocada por *Fusarium* sp., que causa murcha e morte das plantas. A mancha-bacteriana, causada por *Xanthomonas* sp., manifesta-se como manchas translúcidas que necrosam folhas e frutos (Viana *et al.*, 2003; Machado *et al.*, 2017). A principal virose, o endurecimento dos frutos, é causada pelo cowpea aphid-borne mosaic virus (CABMV, *Potyvirus vignae*), levando ao endurecimento do pericarpo e à redução da polpa (Santos-Filho *et al.*, 2017). Outras viroses, como o definhamento precoce ou a pinta verde, embora menos frequentes, são ameaças aos pomares brasileiros (Kitajima *et al.*, 2003a).

Os cilevírus (gênero *Cilevirus*, família *Kitaviridae*) são fitovírus transmitidos de maneira persistente por ácaros do gênero *Brevipalpus* (Ramos-González *et al.*, 2023). Estes vírus possuem genomas segmentados de RNA de fita simples com polaridade positiva, e são encapsulados em partículas baciliformes curtas. Entre os cilevírus descritos, o passion fruit green spot virus (PfGSV, *Cilevirus passiflorae*) foi responsável por surtos de doenças que devastaram pomares de maracujazeiros durante a década de 1990, associados a sintomas como lesões cloróticas, necroses e manchas verdes nos frutos maduros (Santos-Filho *et al.*, 1999; Kitajima *et al.*, 2003a). A doença compromete a aparência e o valor comercial dos frutos, além de afetarem a produtividade das plantas. Apesar da importância fitopatológica do PfGSV, muitos aspectos de sua biologia, epidemiologia e interação com o hospedeiro permanecem pouco compreendidos.

O genoma do PfGSV é composto por dois segmentos, RNA1 e RNA2. O RNA1 contém duas ORFs (fases de leitura aberta, do inglês *open reading frames*), a *RdRp*, responsável pela replicação viral, e a *p29*, responsável por codificar a provável capa proteica. O RNA2 codifica proteínas que participam da interação vírus-hospedeiro e possivelmente vírus-vetor. No RNA2, cinco ORFs são codificadas: *p15*, *p13*, *p61*, *p32/MP* e *p24*. O trecho genômico entre as ORFs *p15* e *p61*, de aproximadamente 1.500 nucleotídeos, vinha sendo considerado uma região intergênica não codificante até a

recente descrição de alelos da ORF *p13* em PfGSV (Ramos-González *et al.*, 2020). A análise de cinco isolados de diferentes estados brasileiros demonstrou a presença de dois alelos: *p13*, de 330 nucleotídeos e, *p11*, de 282 nucleotídeos, o que apontou como incorreto descrever a região como intergênica. Os ensaios de hibridização por *Northern blot* revelaram a presença de uma banda extra, com tamanho compatível com o que poderia ser uma molécula de RNA subgenômico a partir do qual ocorreria a tradução da proteína P13 (Ramos-González *et al.*, 2020).

Avanços em biologia molecular destacaram a importância de pequenas proteínas virais, antes consideradas secundárias, na funcionalidade e adaptação dos vírus (Pinto e Lamb, 2005; Walker *et al.*, 2011; DiMaio, 2014; Morozov e Solovyev, 2020; Gao *et al.*, 2022; Volic *et al.*, 2023; Zandi *et al.*, 2025). Essas proteínas, frequentemente associadas a genomas virais compactos, desempenham papéis essenciais na replicação, evasão imunológica e patogenicidade. No vírus influenza A (*Alphainfluenzavirus influenzae*), por exemplo, a proteína M2 atua como canal iônico, facilitando o desnudamento viral e o brotamento dos vírions (Pinto *et al.*, 1992). O vírus da imunodeficiência humana 1 (*Lentivirus humimdefl*), codifica a proteína Vpu que promove a evasão imunológica ao regular negativamente CD4 e tetherin, além de formar canais iônicos (Neil *et al.*, 2008). Nos papilomavírus, a proteína E5 atua como fator oncogênico, alterando receptores de crescimento celular e contribuindo para a oncogênese (DiMaio e Petti, 2013). Já nos poxvírus, várias pequenas proteínas transmembrana participam da biogênese e maturação dos vírions (DiMaio, 2014). Em vírus de plantas como os potyvirus, as proteínas 6K1 e 6K2 são cruciais para a formação do complexo de replicação viral e para a permeabilidade da membrana celular, facilitando a infecção (Cui e Wang, 2016; Chai *et al.*, 2024). As proteínas BMB1 e BMB2, codificadas pelo kitavirus hibiscus green spot virus (HGSV, *Higrevirus waimanalo*), interagem de forma cooperativa para mediar o transporte célula a célula do genoma viral, a BMB1 atua direcionando a BMB2 aos plasmodesmos (Atabekova *et al.*, 2022).

A emergência de novas ORFs em genomas virais representa um fenômeno essencial para a adaptação e evolução desses organismos, e ocorre por meio de mecanismos altamente dinâmicos e diversificados (Guerra-Almeida e Nunes-da-Fonseca, 2020). Duas hipóteses principais explicam a origem dessas ORFs: a divergência rápida de genes pré-existentes e o surgimento *de novo* a partir de regiões não codificantes. Na primeira hipótese, genes ancestrais sofrem diversas mutações, gerando uma nova ORF,

que ao longo do tempo, perde a similaridade com o gene prévio. Já o surgimento *de novo*, regiões anteriormente consideradas não codificantes, sofrem mutações que levam a formação de novas ORFs, que passam a ser transcritas e traduzidas (Singh e Syrkin Wurtele, 2020; Valkirilis *et al.*, 2020). Além disso, os RNAs não codificantes também podem desempenhar papéis cruciais na emergência de novos genes. Esses transcritos longos podem abrigar pequenas ORFs, que, por meio de uma tradução pervasiva, dão origem a “microproteínas” (ou micropeptídeos) (Guerra-Almeida e Nunes-da-Fonseca, 2020). Os proto-genes gerados por esses fenômenos são favorecidos pelos mecanismos evolutivos ao fornecerem vantagens adaptativas. A fixação dessas novas variantes geralmente envolve a aquisição de elementos regulatórios, como sequências promotoras eficientes, regiões 5’ e 3’ estáveis e sinais de tradução, que aumentam a eficiência e a estabilidade do RNA mensageiro (Carvunis *et al.*, 2012). Os processos de mutação e recombinação são fundamentais no processo de surgimento de novas ORFs (Park e Hahn, 2021; Pavessi e Romerio, 2022; Ji *et al.*, 2023). Mutações pontuais, inserções ou deleções podem criar fases de leituras ou modificar aquelas existentes, enquanto eventos de recombinação genômica facilitam o rearranjo de segmentos genéticos. Essas modificações criam funcionalidades, permitindo que os vírus se adaptem de forma ágil às mudanças ambientais (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012).

Em monkeypox virus (*Orthopoxvirus monkeypox*), durante o surto de 2022, foram identificadas mutações genômicas significativas que incluíram deleções extensas e interrupções em ORFs, em um total de 25 amostras clínicas contendo mutações disruptivas (Sereewit *et al.*, 2022). Essas alterações ocorreram principalmente em pequenos genes, não essenciais para a replicação, mas que desempenham um papel crucial na modulação da resposta do hospedeiro, potencialmente reduzindo a funcionalidade de proteínas imunomoduladoras. Além disso, as análises identificaram a predominância de mutações $G \rightarrow A$ e $C \rightarrow T$, indicando uma provável atuação do sistema de edição APOBEC3 do hospedeiro humano, conhecido por induzir essas modificações como mecanismo de defesa antiviral. De forma notável, as mutações detectadas afetaram genes não essenciais para a replicação viral, mas relacionados com a adaptação ao hospedeiro, demonstrando a pressão seletiva durante a dispersão viral. A identificação das mutações e deleções também reforça a importância de monitoramento para melhores estratégias de controle e diagnóstico (Sereewit *et al.*, 2022).

A crescente compreensão do papel das proteínas virais de pequeno tamanho e a dinâmica evolutiva dos vírus, incluindo mutações e recombinações genômicas, revelam sua importância na adaptação e patogenicidade viral. No caso do PfGSV, a emergência de pequenas proteínas em seu genoma, como o caso da ORF *p13*, não encontrada em outros membros do gênero *Cilevirus*, poderia ter uma ação fundamental na sua evolução e interação com o hospedeiro e/ou vetor. Esses processos permitem que os vírus se adequem rapidamente a mudanças ambientais, facilitando a evasão dos mecanismos de defesa dos seus hospedeiros e a propagação viral. Os estudos dessas pequenas proteínas e das mudanças genéticas nos genomas virais, como em PfGSV, é essencial para o entendimento das relações vetor-vírus-hospedeiro e para o desenvolvimento de estratégias de controle.

A exploração das ORFs emergentes é fundamentada na possibilidade de que elas desempenhem papéis cruciais na interação vírus-hospedeiro, como a modulação de resposta antiviral ou a indução de alterações fisiológicas no hospedeiro para favorecer a replicação viral ou o movimento célula a célula. Além disso, considera-se que essas proteínas emergentes também poderiam contribuir para a eficiência da transmissão pelo vetor *Brevipalpus*, otimizando e impulsionando a transmissão. Entender essas dinâmicas é essencial, pois características genômicas únicas, como as pequenas proteínas codificadas por ORFs recentemente descritas, frequentemente proporcionam vantagens evolutivas aos vírus, permitindo sua adaptação a novos desafios ecológicos (DiMaio, 2014; Storz *et al.*, 2014). Assim, a caracterização dessas ORFs pode esclarecer mecanismos determinantes para a epidemiologia do PfGSV, fornecendo subsídios valiosos para o desenvolvimento de estratégias de controle mais eficazes.

2 REVISÃO BIBLIOGRÁFICA

2.1 O maracujazeiro: diversidade biológica, economia e desafios na produção

O termo “maracujá”, que possui origem Tupi-Guarani e significa “alimento em forma de cuia”, dá nome aos frutos das plantas do gênero *Passiflora*, família Passifloraceae. Suas utilidades são amplas, destacando-se na indústria alimentícia, cosmética e farmacêutica, devido às suas propriedades calmantes, antioxidantes e anti-inflamatórias (Rai *et al.*, 2022). Além disso, é valorizado como planta ornamental pela exuberância de suas flores, que variam de branco a tons de rosa, vermelho, azul e roxo (Faleiro *et al.*, 2017).

Com mais de 500 espécies descritas, a maioria nativa da América do Sul, principalmente Brasil e Colômbia, os primeiros registros de plantas do gênero *Passiflora* datam do século XVI, segundo o Tratado Descritivo do Brasil, de 1587 (de Jesus e Faleiro, 2016). São plantas trepadeiras, herbáceas ou lenhosas, de crescimento vigoroso, atingindo de 5 a 10 metros de comprimento, com flores hermafroditas que se abrem uma única vez ao dia (de Jesus e Faleiro, 2016). A polinização é realizada majoritariamente por abelhas-mamangavas, além de morcegos e beija-flores em menor escala (Abrahamczyk *et al.*, 2014).

O Brasil, líder mundial na produção de maracujá, é responsável por mais de 70% da produção global (Faleiro, 2022). Em 2024, a produção nacional de maracujá atingiu 736.585 toneladas, resultando em um valor de R\$ 2,56 bilhões, com destaque para os estados do Ceará e Bahia (IBGE, 2024). As principais espécies cultivadas comercialmente são *Passiflora edulis* Sims. e *Passiflora alata* Curtis. Apesar de sua relevância econômica e produtividade, a cultura do maracujá enfrenta diversos desafios fitossanitários (Machado *et al.*, 2017). Entre as doenças fúngicas mais prejudiciais estão a antracnose, causada por *Colletotrichum* sp., e a fusariose, causada por *Fusarium* sp., que comprometem folhas, frutos e levam as plantas à morte. No caso de doenças bacterianas, a mancha-bacteriana causada por bactérias *Xanthomonas* sp., provoca lesões nas folhas e frutos (Viana *et al.*, 2003; Machado *et al.*, 2017). A principal virose, o endurecimento dos frutos, causada por CABMV, reduz drasticamente a produção de polpa nos frutos (Santos-Filho *et al.*, 2017). Além do CABMV, relatos recentes também escrevem a ocorrência de Ugandan passiflora virus (UPV, gênero *Potyvirus*) e East Asian passiflora distortion virus (EAPDV, gênero *Potyvirus*), ambos associados ao endurecimento dos frutos no Quênia

(Munguti et al., 2025). Outras viroses, como o definhamento precoce e a pinta verde, causadas pelo passion fruit green spot virus (PfGSV, *Cilevirus passiflorae*), embora de baixa incidência, permanecem como ameaças potenciais à produção brasileira. Recentemente, o PfGSV também foi detectado em amostras no Quênia (Munguti et al., 2025) representando ameaça à produção local. A ampla gama de patógenos virais ressaltam a necessidade de manejo adequado para garantir a viabilidade da cultura (Santos-Filho et al., 1999; Kitajima et al., 2003a; Antonioli-Luizon, 2010; Ramos-González et al., 2020).

2.2 Gênero *Cilevirus*: caracterização genômica e relevância fitopatológica

O gênero *Cilevirus* inclui vírus de partículas baciliformes curtas envelopadas, com dimensões de 50-60 nm por 110-120 nm, que são frequentemente associadas ao retículo endoplasmático das células infectadas. Nas células infectadas são observadas inclusões eletrodensas conhecidas como viroplasmas, que provavelmente correspondem aos locais de replicação viral (Kitajima et al., 2003a). Atualmente oito espécies no gênero são aceitas pelo ICTV (International Committee on Taxonomy of Viruses): *Cilevirus australis*, *Cilevirus colombiaense*, *Cilevirus leprosis*, *Cilevirus ligustri*, *Cilevirus oahuense*, *Cilevirus passiflorae*, *Cilevirus solani* e *Cilevirus pistaciae* (Olmedo-Velarde et al., 2021, 2022; Ramos-González et al., 2022). O vinca ringspot virus (ViRSV) foi recentemente considerado membro de uma nova espécie (*Cilevirus chilense*) do gênero¹. As plantas acometidas por cilevírus desenvolvem manchas cloróticas e/ou necróticas nas folhas, frutos e ramos, localizadas nos locais de alimentação do ácaro *Brevipalpus* spp., resultado de infecções não sistêmicas.

O genoma dos cilevírus é composto por duas moléculas de RNA de fita simples positiva (+) ssRNA, contendo uma cauda poli-A nos extremos 5' e seus membros podem ser classificados em três subgrupos com base na organização genômica (Freitas-Astúa et al., 2018; Ramos-González et al., 2023). O subgrupo 1 inclui o vírus da leprose dos citros C (CiLV-C, *Cilevirus leprosis*), o vírus da leprose dos citros C2 (CiLV-C2, *Cilevirus colombiaense*) e o PfGSV (*Cilevirus passiflorae*). O CiLV-C é o agente etiológico prevalente da leprose dos citros (LC), principal doença viral da citricultura brasileira

1

https://ictv.global/system/files/proposals/pending/2025/Plant%20virus%20%28P%29%20proposals/2025.004P.N.v2.Kitaviridae_13nsp.docx

(Ramos-González *et al.*, 2018). Na Colômbia, o CiLV-C2 também foi relatado causando LC em plantas de laranja doce (Roy *et al.*, 2013). O RNA1 desses vírus possui entre 8,7 e 8,9 kb e contém duas ORFs: *RdRp*, que codifica um polipeptídio único envolvido na replicação viral, e *p29*, que é responsável pela provável codificação da proteína de capa. O RNA2 é menor, com um tamanho de 4,7 a 5 kb, e codifica quatro ORFs: *p15*, *p61*, *MP* e *p24* (Freitas-Astúa *et al.*, 2018). A proteína P15 atua como um possível supressor de silenciamento gênico; a P61 é provavelmente uma glicoproteína e efector viral; P24 é uma putativa proteína estrutural do vírion e a mais conservada entre os cilevirus; e a MP é a proteína de movimento (Ramos-González *et al.*, 2021). A região que separa as ORFs *p15* e *p61* é conhecida como região intergênica e abriga ORFs órfãs específicas de cada vírus.

O subgrupo 2 é constituído pelo *Solanum violifolium* ringspot virus (SvRSV, *Cilevirus solani*), *Ligustrum leprosis* virus (LigLV, *Cilevirus australis*), *Ligustrum chlorotic spot* virus (LigCSV, *Cilevirus ligustri*) (Ramos-González *et al.*, 2023), pistachio virus Y (PisVY, *Cilevirus pistaciae*) e ViRSV. A principal diferença em comparação com o subgrupo 1 é a falta de 1,0 a 1,5 kb no extremo 5' do RNA2, que tem entre 3,5 e 3,6 kb e agrupa ORFs ortólogas aos genes *p61*, *p32* (*MP*) e *p24*. O RNA2 dos cilevirus do subgrupo 2 não apresentam o ortólogo a ORF *p15*. Por fim, o subgrupo 3 é representado por uma única espécie, *Cilevirus oahuense* (hibiscus yellow blotch virus, HYBV) (Ramos-González *et al.*, 2023), até o momento detectada apenas no Havaí, EUA. O RNA1 do HYBV é o menor entre os cilevírus descritos, com 8,4 kb, e contém duas ORFs, *RdRp* e ORF *p10*, que estão parcialmente sobrepostas. Diferentemente dos cilevírus dos subgrupos 1 e 2, em que a ORF *p29* está localizada no extremo 3' do RNA1, no HYBV essa ORF encontra-se posicionada no extremo 3' do RNA2, após a ORF *p24* (Olmedo-Velarde *et al.*, 2021). Por sua organização genômica, o HYBV tem sido considerado um elo evolutivo entre cilevírus e higrevírus, outro gênero da família *Kitaviridae* também encontrado no Havaí (Olmedo-Velarde *et al.*, 2021; Ramos-González *et al.*, 2023).

2.2.1 Ácaros do gênero *Brevipalpus* como vetores de fitovírus

O gênero *Brevipalpus* (Acari: Tenuipalpidae) agrupa cerca de 300 espécies de ácaros conhecidos como falso ácaro-aranha, distribuídos mundialmente, de hábito cosmopolita e polífago (Childers e Rodrigues, 2011; Castro *et al.*, 2020; Santos de Mendonça *et al.*, 2025). São ácaros avermelhados achatados dorso-ventralmente e que apresentam quatro pares de pernas na vida adulta. As colônias de algumas espécies são formadas em grande maioria por fêmeas haploides que se reproduzem a partir de

partenogênese telítica devido à presença da bactéria endossimbionte *Cardinium* sp. (Weeks *et al.*, 2001). Em 2022, o Ministério da Agricultura e Pecuária declarou os ácaros *Brevipalpus* como pragas de alto risco para a agricultura brasileira (MAPA, 2022).

Os ácaros do gênero *Brevipalpus* atuam como vetores de fitovírus em culturas de significativa importância agrônômica, como *Citrus sinensis*, *Coffea arabica*, *Passiflora* sp. e plantas ornamentais, como orquídeas de diferentes gêneros (Kitajima *et al.*, 2003b, 2010). Os vírus transmitidos por esses ácaros, denominados vírus transmitidos por *Brevipalpus* (VTBs), pertencem a três gêneros: *Cilevirus* e *Higrevirus* (família *Kitaviridae*, ordem *Martellivirales*) (Locali-Fabris *et al.*, 2006; Melzer *et al.*, 2012; Rodrigues, 2022) e *Dichorhavirus* (família *Rhabdoviridae*, ordem *Mononegavirales*) (Dietzgen *et al.*, 2018). Os cilevírus e higrevírus são conhecidos como VTBs citoplasmáticos, uma vez que realizam a replicação viral no citoplasma celular (Ramos-González *et al.*, 2023). Por outro lado, os dichorhavirus possuem RNA de fita simples de sentido negativo, dividido em duas moléculas, e por replicarem no núcleo das células infectadas, são chamados de VTBs nucleares (Dietzgen *et al.*, 2018; Freitas-Astúa *et al.*, 2018).

2.3 Passion fruit green spot virus: biologia, epidemiologia e impactos agrícolas

Duas importantes doenças impactaram os pomares de maracujá no Brasil na década de 1990. O definhamento precoce do maracujazeiro (DPM), identificado em 1994 na Bahia, caracterizava-se por intensa necrose nos ramos, que os circundava, levando à morte prematura das plantas, o que deu nome à doença (Santos-Filho *et al.*, 1999). Nas folhas, observavam-se anéis e manchas cloróticas ao redor das nervuras, que posteriormente se tornavam verdes. Simultaneamente, a pinta verde do maracujazeiro (PVM) foi descrita em 1997 na região de Alta Paulista, São Paulo, manifestando-se principalmente por manchas verdes nos frutos maduros e clorose ao redor das nervuras foliares, que se tornavam verdes em folhas senescentes (Kitajima *et al.*, 2003a). As plantas afetadas também desenvolviam necrose nos ramos, embora de forma menos intensa que no DPM. Na época, ambas as condições foram associadas à infestação por ácaros do gênero *Brevipalpus* e à observação de partículas virais baciliformes curtas no citoplasma das células infectadas (Kitajima *et al.*, 2003a). Durante a primeira década dos anos 2000, sintomas similares foram observados em outras regiões do Brasil, como Minas Gerais, Rio de Janeiro, Sergipe, Rondônia, Maranhão, Rio Grande do Norte, Paraíba, Pará e Distrito Federal (Kitajima *et al.*, 2003; Moraes *et al.*, 2006). Em 2009, o sequenciamento

parcial do PfGSV identificou este vírus como o agente etiológico das duas doenças (Antonioli-Luizon, 2010). Mais recentemente, sua ocorrência foi confirmada na Colômbia, onde plantas de maracujá-amarelo (*Passiflora edulis f. flavicarpa*) exibindo sintomas de manchas verdes em folhas senescentes e clorose ao redor das nervuras foliares foram diagnosticadas com infecção por PfGSV. A detecção foi confirmada por RT-PCR e sequenciamento de alto desempenho (HTS), que revelaram elevada identidade nucleotídica (>99%) com isolados brasileiros de PfGSV, representando o primeiro relato do vírus na Colômbia (Roy *et al.*, 2023).

Embora o maracujazeiro amarelo seja o principal hospedeiro do PfGSV, o vírus é capaz de infectar pelo menos 35 gêneros botânicos (Kitajima *et al.*, 2017; Saito, 2020; da Costa-Rodrigues *et al.*, 2022; Pires, 2022). Recentemente, 16 espécies foram identificadas como hospedeiras experimentais através da inoculação com ácaros virulíferos do gênero *Brevipalpus*, incluindo: *Brunfelsia uniflora* (Pohl) D. Don, *Cajanus cajan* (L.) Millsp., *Capsicum annuum* L., *Cataranthus roseus* L., *Dolichos lablab* L., *Hedera helix* L., *Hibiscus sinensis* Linn, *Hoya carnosa* (L.) R. Br., *Ligustrum sinense* Lour., *Lonicera japonica* Thunb., *Malvaviscus arboreus* Cav., *Mucuna pruriens* (L.) DC., *Nicotiana megalosiphon* Van Heurck & Muell. Arg., *Pelargonium* sp., *Phaseolus vulgaris* L. e *Vigna unguiculata* (L.) Walp (Pires, 2022). Além disso, quatro espécies (*Antirrhinum majus* L., *Chenopodium quinoa* Willd., *Datura stramonium* L., e *Nicotiana tabacum* L.) não desenvolveram sintomas quando inoculadas com ácaros *Brevipalpus*, mas foram infectadas pelo PfGSV por meio de inoculação mecânica, indicando que são hospedeiras do vírus, mas não dos ácaros vetores (Pires, 2022).

A presença do PfGSV nas plantas é evidenciada pelo surgimento de manchas verdes nos frutos maduros. Nas folhas, a infecção promove o aparecimento de anéis cloróticos e/ou necróticos e o amarelamento das nervuras, especialmente em folhas senescentes. Já nos ramos, observa-se o desenvolvimento de necroses intensas, que frequentemente coalescem, circundando-os e culminando na morte das plantas (Figura 1 A e B). O PfGSV é transmitido pelo ácaro *Brevipalpus yothersi* (Figura 1 C), de forma persistente, mas não há nenhum estudo avaliando a possibilidade de replicação deste vírus no vetor. A caracterização do genoma de cinco isolados permitiu a classificação do PfGSV no gênero *Cilevirus*, família *Kitaviridae* (Ramos-González *et al.*, 2020). Em análises de microscopia eletrônica de transmissão, é possível observar partículas virais baciliformes curtas envelopadas (50-70 nm x 100-120 nm) no citoplasma das células infectadas, além

de viroplasmas eletrodensos, associados ao retículo endoplasmático (Figura 1D-a, 1D-b e 1D-c) (Kitajima *et al.*, 2003a; Pires, 2022).

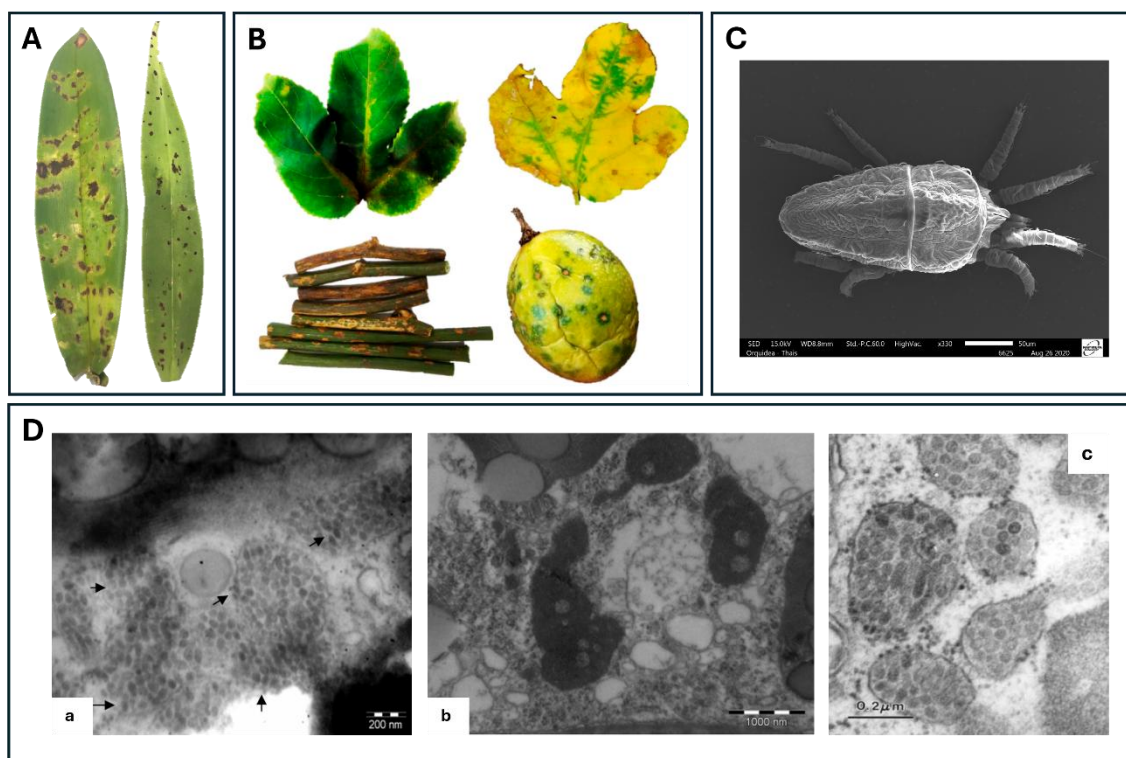


Figura 1. **A** – Sintomas da infecção por PfGSV em orquídeas do gênero *Eria* (isolado PfGSV_Lim01). **B** – Sintomas de infecção por PfGSV em folhas, ramos e frutos de maracujazeiro (Ramos-González *et al.*, 2020). **C** – Micrografia de *Brevipalpus yothersi* coletado em uma amostra de orquídea originária de Limeira infectada com PfGSV_Lim01 (colaboração com Dra. Aline Tassi). **D-a, D-b e D-c** – Partículas baciliformes observadas em cisternas do retículo endoplasmático de células de maracujazeiro infectadas pelo PfGSV (Kitajima *et al.*, 2003a; Antonioli-Luizon, 2009).

O genoma do PfGSV é composto por duas moléculas de RNA de fita simples com polaridade positiva [(+) ssRNA], e cauda poli-A nos extremos 3' (Figura 2) (Ramos-González *et al.*, 2020). O RNA1 possui aproximadamente 9 kb e contém duas ORFs: a que codifica a proteína RNA polimerase dependente de RNA (*RdRp*), envolvida na replicação viral e que contém os domínios de metiltransferase, helicase e *RdRp*, e a ORF *p29*, que codifica a provável proteína de capa. O RNA2 é menor, com cerca de 5 kb, e contém cinco ORFs: *p15*, *p13/p11*, *p61*, *MP* e *p24*. A ORF *p15* codifica uma provável proteína de 15 kDa, exclusiva dos cilevírus, que provavelmente desempenha um papel adaptativo crucial, relacionado com a regulação da interação vírus-planta, atuando como supressor de silenciamento gênico e/ou interagindo com outras proteínas virais ou do hospedeiro (Ramos-González *et al.*, 2021). A ORF *p13* codifica uma proteína putativa de 13 kDa, sem homologia conhecida em bancos de dados públicos, com função ainda não elucidada. A ORF *p61* codifica uma provável glicoproteína que atua como um efector viral.

A ORF *MP* codifica a proteína de movimento pertencente à superfamília 30K, e a ORF *p24* codifica a putativa proteína estrutural do vírion, sendo a proteína P24 a mais conservada entre os cilevírus (Ramos-González *et al.*, 2020).

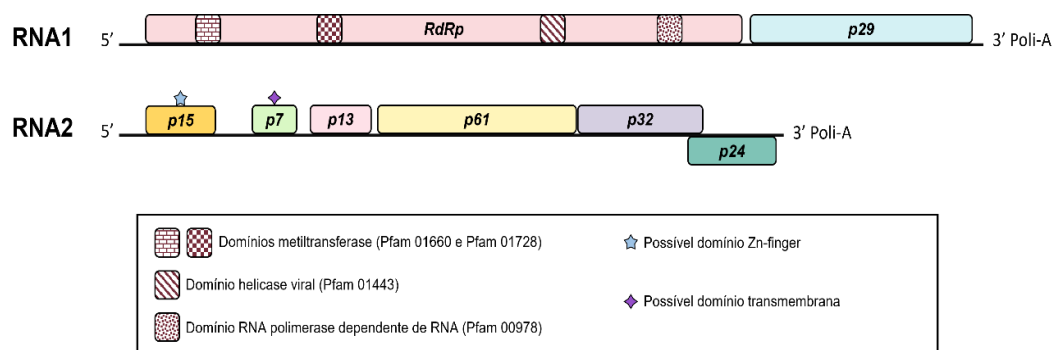


Figura 2. Representação esquemática da organização genômica de PfGSV. Adaptado de Ramos-González *et al.*, 2020; 2021.

A região genômica entre as ORFs *p15* e *p61* do RNA2 de cilevírus foi denominada região intergênica (RI) por ser considerada não codificante. No entanto, essa denominação é inadequada para o PfGSV, uma vez que foram constatadas evidências de regiões codificadoras, além de ser uma RI menor que as observadas em outros cilevírus. A análise das sequências de cinco isolados identificou um alelo da ORF *p13*, denominado *p11* (Ramos-González *et al.*, 2020). Esse alelo foi detectado em apenas um isolado, coletado em Bom Jesus da Lapa, Bahia, em 2008 (PfGSV_BJL01), obtido em uma amostra de maracujazeiro. O alelo *p11* codifica uma proteína putativa de 11,1 kDa, que compartilha os primeiros 40 resíduos com a proteína codificada pelo alelo *p13*. A proteína codificada pela ORF *p13* apresenta 10 resíduos de arginina e um de lisina, o que sugere uma carga positiva, possivelmente favorecendo interações com moléculas de carga negativa, como os ácidos nucleicos. Além disso, a proteína contém um domínio de 15 aminoácidos (aa), com quatro resíduos de cisteínas e dois de histidina, que compartilha similaridade com motivos encontrados no inibidor antiapoptose induzido por citocina 1 (CIAPIN1; PF05093) e no domínio PGC7/Stella/Dppa3 (PGC7_Stella; PF15549). Esses 15 aa fazem parte de uma região maior que apresenta semelhanças com o domínio FDB, descrito em proteínas vegetais contendo domínios FBox e BRCT (Doerks *et al.*, 2002; Ramos-González *et al.*, 2020). Na região C-terminal da P13, foi identificada uma sequência de 26 aa com moderada similaridade ao domínio de ligação à quitina tipo 3 (InterPro IPR003610). Essas características sugerem que a proteína P13 pode

desempenhar funções auxiliares, potencialmente mediando interações moleculares cruciais para o ciclo de multiplicação viral, como a ligação a ácidos nucleicos e outras moléculas essenciais à patogênese viral.

O padrão de RNAs subgenômicos (sgRNAs) de PfGSV apresentou similaridade com o descrito para CiLV-C (Ramos-González *et al.*, 2020). Análises por Northern blot utilizando sondas específicas para as ORF *p29* e *p24* permitiram a identificação de duas bandas: gRNA1 (RNA genômico 1) e sgRNA1, o qual codifica a proteína P29. Para o RNA2, foram detectadas cinco bandas distintas, correspondendo a gRNA2, sgRNA2, sgRNA3, sgRNA4 e sgRNA5. Os sgRNAs 2 e 3 possivelmente codificam as proteínas P61 e MP, respectivamente. O sgRNA4 foi o mais abundante e corresponde a ORF *p24*. A quinta banda, referente ao sgRNA5, foi identificada como uma banda de baixa intensidade, sugerindo a necessidade de mais estudos para sua confirmação. O sgRNA5 demonstrou-se tão abundante quanto o sgRNA2, com um peso molecular ligeiramente superior, e pode estar associado à expressão da ORF *p13* (Ramos-González *et al.*, 2020).

2.4 Pequenas proteínas acessórias virais e suas funções acessórias adaptativas

No surgimento das técnicas bioinformáticas para a anotação de genomas foram aplicados limites de tamanhos arbitrários para a detecção de ORFs putativas e as funções de suas respectivas proteínas (Storz *et al.*, 2014). Nesse contexto, proteínas menores foram muitas vezes negligenciadas, sendo consideradas erros de predição ou sem função relevante (Hellens *et al.*, 2016). No entanto, no caso particular dos vírus, que possuem genomas compactos, mecanismos alternativos foram selecionados para codificar proteínas de pequeno tamanho e com funções biológicas cruciais. Nos últimos anos, pequenas fases de leitura abertas (sORFs, do inglês *small* ORFs), que codificam proteínas menores de 100 aa com papéis cruciais foram identificadas (DiMaio, 2014).

A proteína M2 do vírus influenza A (*Alphainfluenzavirus influenzae*, família *Orthomyxoviridae*), por exemplo, é uma proteína integral de membrana que forma tetrâmeros com canais iônicos, facilitando a entrada de prótons no vírion, essencial para o desnudamento viral, processo bloqueado por antivirais como a amantadina (Sugrue e Hay, 1991; Pinto *et al.*, 1992; Pinto e Lamb, 2005). Além disso, sua porção citoplasmática participa do brotamento do vírion maduro. No vírus da imunodeficiência humana 1 (*Lentivirus humimdefl*, família *Retroviridae*), a proteína transmembrana Vpu regula negativamente CD4 e tetherin, promovendo evasão imunológica e disseminação viral,

além de formar canais iônicos (Neil *et al.*, 2008; McNatt *et al.*, 2009; Strebel, 2014). Os papilomavírus (família *Papillomaviridae*) codificam a proteína E5, um fator oncogênico com domínio transmembrana, que ativa receptores de crescimento celular. No caso do papilomavírus humano 16, a proteína E5 além de modular a atividade do fator de crescimento epidérmico, também é capaz de oligomerizar formando canais iônicos, atuando assim como viroporina (DiMaio e Petti, 2013; DiMaio, 2014). Nos poxvírus (família *Poxviridae*), muitas proteínas transmembrana pequenas são codificadas, como a O3L, que facilita a entrada viral na célula hospedeira, e pelo menos outras oito proteínas participam da biogênese e maturação do vírion, embora algumas funções específicas permaneçam desconhecidas (DiMaio, 2014). Já as pequenas proteínas hidrofóbicas (SH), encontradas em *Paramyxoviridae* e *Pneumoviridae*, atuam como viroporinas e modulam a resposta imune, inibindo apoptose e vias imunológicas, sendo essenciais para a evasão do sistema imune, apesar de não serem necessárias para a replicação viral (DiMaio, 2014; Brynes e Williams, 2024).

Pequenas proteínas acessórias também são encontradas em vírus que infectam plantas. Os vírus da família *Potyviridae*, a maior família de vírus de plantas com genoma de RNA, codificam duas proteínas auxiliares de pequeno porte essenciais para a proliferação viral, denominadas 6K1 e 6K2 (Cui e Wang, 2016; Yang *et al.*, 2021). A proteína 6K2 atua como uma proteína integral de membrana, desempenhando um papel crucial na formação do complexo de replicação viral ao interagir com o retículo endoplasmático das células hospedeiras infectadas. A proteína 6K1 foi caracterizada como uma viroporina, formando pentâmeros com canais iônicos centrais, o que resulta no aumento da permeabilidade da membrana e facilita a passagem de íons potássio (Chai *et al.*, 2024).

A caracterização da proteína P6 do barley yellow striate mosaic virus (BYSMV, *Cytorhabdovirus hordei*) revelou outro papel adicional essencial para a infecção viral: a modulação da resposta do hospedeiro com o objetivo de atrair insetos vetores (Gao *et al.*, 2022). A proteína P6_BYSMV demonstrou a capacidade de inibir a sinalização mediada por ácido jasmônico, o principal hormônio associado à resposta de plantas contra herbivoria. Especificamente, a proteína P6 bloqueia a ativação do complexo E3 ubiquitina-ligase, responsável pela degradação de proteínas reguladoras, como as da família Jasmonate ZIM-domain (JAZ). O acúmulo de JAZ, por sua vez, resulta em uma redução na resposta mediada por ácido jasmônico, tornando as plantas infectadas menos

repelentes para os insetos vetores e, conseqüentemente, promovendo a dispersão viral (Gao *et al.*, 2022).

As proteínas BMB1 e BMB2, codificadas pelo HGSV, gênero *Higrevirus*, desempenham papéis complementares no transporte célula a célula do material genético viral por meio dos plasmodesmos (Atabekova *et al.*, 2022). A BMB1 é uma proteína com domínios NTPase/helicase com capacidade de ligação a RNA, enquanto a BMB2 é uma pequena proteína hidrofóbica associada às membranas do retículo endoplasmático. A BMB2 atua induzindo constrições nos túbulos do ER e direcionando BMB1 para os “corpos de membranas associados aos plasmodesmos”, a partir da interação entre a região C-terminal de BMB1 e a N-terminal de BMB2 (Atabekova *et al.*, 2022). O sistema funcional formado pelas proteínas BMB1 e BMB2 é comparável ao “triple gene block” (TGB) codificado por vírus de diferentes famílias (Morozov e Solovyev, 2003), porém com adaptações para um “dual gene block” (DGB) funcional. Além disso, a BMB2 atua aumentando a permeabilidade dos plasmodesmos, permitindo a passagem de complexos de ribonucleoproteínas por meio de canais dilatados, atuando de forma semelhante a proteínas viroporinas. Essas funções cooperativas entre as duas proteínas são essenciais para o transporte eficiente do genoma viral entre células adjacentes, destacando a modularidade e a especialização evolutivo do sistema DGB (Atabekova *et al.*, 2022).

2.5 Dinâmica evolutiva das ORFs virais

No passado considerava-se que novas proteínas poderiam surgir apenas de regiões que previamente codificavam proteínas (Singh e Syrkin Wurtele, 2020). Atualmente, duas principais hipóteses são aceitas: divergência rápida de genes pré-existentes e surgimento *de novo* a partir de sequências não codificantes (Tautz e Domazet-Lošo, 2011; Singh e Syrkin Wurtele, 2020; Xia *et al.*, 2025). No caso da divergência rápida é proposto que uma sequência codificante pré-existente sofre mutações intensas que eliminam a homologia com os genes ancestrais. Esse fenômeno é mediado pelo processo de “pseudogenização” parcial, no qual o gene ancestral perde sua funcionalidade, permitindo que o gene novo sofra processos evolutivos livremente, adquirindo novas funções (Singh e Syrkin Wurtele, 2020). Em contraste, o mecanismo de surgimento *de novo* envolve a emergência de genes a partir de regiões não codificantes, como regiões intergênica, íntrons ou fases de leitura alternativas dentro de genes existentes (Carvunis *et al.*, 2012; Van Oss e Carvunis, 2019). Essas regiões, anteriormente não codificantes, podem ser transcritas, gerando protogenes que atuam como reservatório de diferenciação genética

(Van Oss e Carvunis, 2019). Além disso, RNAs não codificantes, RNAs longos não codificantes (*lncRNA*) também poderiam atuar no surgimento de novos genes. Esses transcritos contêm ORFs, que podem ser traduzidas por meio de tradução pervasiva, gerando peptídeos primitivos que, sob pressões seletivas, evoluem para genes codificantes estáveis (Carvunis *et al.*, 2012; Van Oss e Carvunis, 2019; Guerra-Almeida e Nunes-da-Fonseca, 2020).

O surgimento de novas ORFs em genomas virais está diretamente relacionado a plasticidade genômica característica desses organismos, que possuem altas taxas de mutação e recombinação genômica, além de estarem submetidos a pressões seletivas intensas (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012). Eventos de mutação, como substituições nucleotídicas, inserções ou deleções, frequentemente resultam na emergência de ORFs *de novo*, especialmente em regiões intergênicas ou sobrepostas a ORFs pré-existentes. Mutações que alteram o quadro de leitura podem levar ao aparecimento de novas proteínas, algumas das quais, com funções adaptativas vantajosas, são favorecidas pela seleção positiva, como evidenciado no aumento da replicação ou na capacidade de evasão das defensas hospedeiras (Guerra-Almeida e Nunes-da-Fonseca, 2020). Além disso, mecanismos de recombinação e rearranjos genômicos desempenham papéis centrais no surgimento de novas ORFs em genomas virais. Esses processos permitem a reorganização de segmentos genéticos e criam um ambiente propício para diversidade genômica que pode ser explorado pela seleção natural (Carvunis *et al.*, 2012).

A fixação de certas variantes genômicas virais ocorre de forma complexa e pode depender de uma combinação de fatores adaptativos e estocásticos. Variantes que conferem vantagens adaptativas, como maior eficiência de replicação ou evasão de defesas do hospedeiro, são rapidamente favorecidas pela seleção natural e ao longo do tempo são estabelecidas no genoma viral. A fixação de novas ORFs é amplamente favorecida pela aquisição de elementos reguladores que aprimoram a eficiência do processo de tradução e aumentam a estabilidade do RNA mensageiro (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012). Sinais regulatórios, como sequências de ligação ao ribossomo, elementos de estabilização de extremidade 3', e a formação de contextos favoráveis para o início da tradução, como a sequência Kozak, contribuem para o sucesso dessas variantes, favorecendo sua integração/fixação nos genomas virais. Por outro lado, a eliminação de variantes ocorre quando mutações deletérias ou regiões genômicas redundantes são alvos de seleção purificadora, otimizando o genoma viral para maximizar

sua eficiência de replicação. Esses processos dinâmicos de emergência, fixação e deleção de variantes modulam continuamente a organização dos genomas virais, promovendo adaptações a mudanças nos ambientes e diferentes pressões seletivas (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012; Guerra-Almeida e Nunes-da-Fonseca, 2020; Gao *et al.*, 2022).

Uma recente análise dos genomas do monkeypox virus (*Orthopoxvirus monkeypox*) relacionados ao surto de 2022 revelou alterações genéticas relevantes, incluindo mutações que interrompem ORFs em um total de 25 amostras clínicas (Sereewit *et al.*, 2022). Entre os achados, destacaram-se mutações sem sentido em genes não essenciais, envolvidos na modulação da resposta imune do hospedeiro, resultando na interrupção das proteínas codificadas. Além disso, foram documentadas grandes deleções que afetam múltiplos genes, deletando completamente ORFs e provocando o truncamento de outras. Também foi identificada uma mutação de inserção, induzindo uma mudança de quadro de leitura e alterando significativamente o produto proteico. As mutações identificadas causaram interrupção de determinados genes, provavelmente gerando proteínas não funcionais, e completa deleção de outros. As análises indicaram que as mutações identificadas foram predominantemente transições $G \rightarrow A$ e $C \rightarrow T$, sugerindo um possível envolvimento do sistema de edição enzimática APOBEC3 (Sereewit *et al.*, 2022). Esses resultados fornecem evidências importantes sobre os mecanismos de adaptação do vírus e os efeitos da pressão seletiva em hospedeiros humanos.

Nas ORFs, o contexto Kozak é definido pela sequência consenso que flanqueia o códon de iniciação AUG, e é um dos principais determinantes da eficiência traducional em eucariotos, regulando o reconhecimento do códon de início pelos ribossomos e influenciando diretamente nos níveis de expressão proteica (Kozak, 1986; 1987). Em vírus, a variação nesse contexto constitui uma estratégia adaptativa, permitindo modular a tradução de proteínas de acordo com o momento e as necessidades do ciclo infeccioso e do hospedeiro. Essa plasticidade favorece a emergência e a fixação de novas ORFs, uma vez que sinais regulatórios eficientes, como a presença de sequências Kozak, aumentam a probabilidade de estabilização e expressão funcional de protogenes virais (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012; Guerra-Almeida e Nunes-da-Fonseca, 2020). Nesse sentido, o estudo dos contextos de Kozak é essencial para compreender como mecanismos de regulação traducional contribuem para a adaptação e evolução viral.

3 OBJETIVO GERAL

Caracterizar a variabilidade genética da região do extremo 5' do RNA2 do cilevírus passion fruit green spot virus (PfGSV, *Cilevirus passiflorae*, *Kitaviridae*).

3.1 Objetivos específicos

- Determinar a organização das ORFs emergentes no RNA2 de PfGSV, com ênfase na variabilidade nucleotídica.
- Identificar potenciais características funcionais e filogenéticas das proteínas codificadas pelas ORFs emergentes de PfGSV.

4 MATERIAL E MÉTODOS

4.1 Amostras vegetais: origem, conservação e processamento

4.1.1 Processamento de amostras vegetais *in natura*

As amostras foliares com sintomas característicos da infecção por PfGSV tiveram as lesões isoladas com uso de lâminas estéreis. Um fragmento de cerca de 100 mg de tecido de cada lesão foi submerso em nitrogênio líquido (N₂) e macerado em almofariz e pistilo. O RNA total de amostras foliares foi extraído mediante o protocolo de Trizol™, de acordo com as instruções do fabricante (Thermo Scientific, Waltham, EUA). O tecido sintomático excedente foi armazenado em tubos estéreis em ultrafreezer a -80°C.

4.1.2 Processamento de material vegetal de maracujá conservados em ultrafreezer

A Unidade Laboratorial de Referência em Biologia Molecular Aplicada (ULR-BMA), no Instituto Biológico de São Paulo, SP, possui uma coleção de RNA de amostras com suspeita de infecção por VTBs conservada a -80°C. A maioria das amostras utilizadas neste estudo foi obtida dessa compilação de RNAs extraídos entre os anos de 2012 e 2022. Conforme protocolo estabelecido no laboratório, as amostras recebidas são catalogadas, e informações detalhadas, como a espécie vegetal, a data e o local de coleta, são registradas. Seguindo um protocolo padrão estabelecido, as lesões nas amostras são isoladas utilizando lâminas estéreis, armazenadas em tubos estéreis, e posteriormente mantidas a -80°C.

4.2 Reação de transcrição reversa

A transcrição reversa (RT, do inglês *reverse transcription*) ou síntese de DNA complementar (cDNA) foi realizada com o kit GoScript Reverse Transcriptase (Promega, Madison, EUA), de acordo com as instruções do fabricante. A reação foi realizada para um volume final de 20 µL, usando inicialmente aproximadamente 400 ng de RNA e 8 µL de H₂O livre de RNase para a desnaturação. Em seguida, foram adicionados 2 µL de enzima, 4 µL de buffer e 4 µL de H₂O livre de RNase. Os ciclos térmicos foram: 5 min a 70°C para a desnaturação e, 25°C por 5 min, 42°C por 1h, 70°C por 15 min para a síntese do cDNA.

4.3 Detecção de passion fruit green spot virus

Para a confirmação da infecção por PfGSV foi realizada a reação em cadeia da polimerase (PCR, do inglês *polymerase chain reaction*) com iniciadores específicos (Tabela 1).

Tabela 1. Iniciadores específicos utilizado para a detecção de PfGSV e amplificação da região de interesse. Os iniciadores referentes ao RNA1 incluem regiões como *RdRp* e *p29*, enquanto os do RNA2 abrangem *MP*, *p24* e o extremo 5' do RNA2 de PfGSV.

	Iniciador	ORF/região	Sequência (5'-3')	Tamanho do fragmento (pb)	Referência
RNA1	C13	<i>RdRp</i>	F: ATTCATGCGTTTCACGGTTA R: CGAATGCCTCTGACACAACT	322	Antonioli-Luizon, 2010
	C16	<i>p29</i>	F: ACACCAAGAGTACTATCGATC R: CATCAAGTGGAGCAAGTTC	452	Ramos-González <i>et al.</i> , 2020
RNA2	C6	<i>MP</i>	F: CGATATTTGATCAATCCGTT R: CACCTTAAATTCGAGGGTT	245	Antonioli-Luizon, 2010
	C8	<i>p24</i>	F: TTCATCGCAAGTTCGTATACCT R: CTGTTGTGCCAAATCATCAA	299	Antonioli-Luizon, 2010
	1Fb/1R	Extremo 5' do RNA2	F: GATAAAAATACAAAACCTTTTAGTGATATC R: GGTTTTCACCTCACCTA	1.432	Ramos-González <i>et al.</i> , 2020
	1Fd		F: TAAAAGTCTGGGTGTCATC R: GGTTTTCACCTCACCTA	1.328	

Para amplificação dos fragmentos genômicos de interesse, as PCRs foram montadas usando o kit GoTaq G2 Green Master Mix (Promega, Madison, EUA), seguindo as orientações do fabricante, para uma reação final de 10 µL. Foram utilizados 5 µL de mix, aprox. 100ng de cDNA, 200 nM de cada iniciador e 3,6 µL de H₂O estéril.

4.4 Clonagem molecular e sequenciamento

Para a clonagem dos fragmentos de interesse, a região entre as ORFs *p15* e *p61* do RNA2 de PfGSV foi amplificada utilizando iniciadores específicos (Tabela 1 e Figura 3). As reações de amplificação foram conduzidas com o kit Taq Pol (-) MgCl₂ (Cellco, São Paulo, Brasil), seguindo as recomendações do fabricante, em um volume final de 60 µL. Os produtos amplificados, de aproximadamente 1.500 pb, foram visualizados em gel de agarose a 0,8%, corado com brometo de etídio, e as bandas correspondentes foram recortadas do gel com lâminas estéreis, seguidas de purificação utilizando o kit Wizard SV Gel and PCR Clean-Up System (Promega, Madison, EUA).

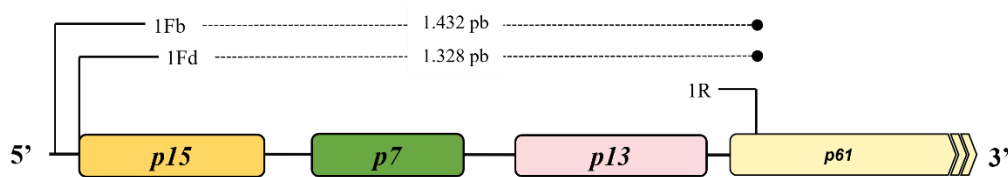


Figura 3. Esquema da região 5' do RNA2 de PfGSV, indicando a localização relativa dos iniciadores e os tamanhos esperado dos fragmentos amplificados.

Os fragmentos purificados foram quantificados utilizando um espectrofotômetro NanoDrop ND 8000 (Thermo Fisher Scientific, EUA) e posteriormente ligados ao plasmídeo pGem-T Easy Vector System (Promega) por meio da enzima T4 DNA ligase (Promega), conforme as instruções do fabricante. Os produtos das reações de ligação foram transformados em células competentes de *Escherichia coli* 10 β via eletroporação (Hanahan, 1983). Em seguida, as células foram incubadas em agitação a 37°C por 1 hora em meio de cultura líquido LB (Luria-Bertani) (Bertani, 1951). O meio LB foi preparado com 10 g/L de triptona, 5 g/L de extrato de levedura e 10 g/L de NaCl. Placas de Petri contendo meio sólido LB, preparado com adição de 15 g/L de ágar, foram suplementadas com ampicilina (100 μ g/mL), IPTG (0,5 μ g/mL) e X-Gal (40 μ g/mL) para a seleção das colônias transformadas. A seleção foi realizada por meio da técnica de coloração branco/azul, onde as colônias de coloração branca foram submetidas à PCR de colônia utilizando o kit GoTaq G2 Green Master Mix (Promega) e iniciadores universais M13 para o vetor. As colônias que apresentaram padrão de amplificação coerente com o fragmento clonado foram submetidas à extração de DNA plasmidial via lise alcalina (Birnboim e Doly, 1979). Por fim, a presença do inserto foi confirmada por digestão com enzima de restrição EcoRI (Promega). O produto da digestão foi observado por eletroforese em gel de agarose 0,8% com adição de brometo de etídio.

4.5 Sequenciamento Sanger dos fragmentos clonados

Após confirmação da inserção do fragmento nas construções, o DNA de três colônias independente de cada transformação foi purificado mediante o procedimento PEG 6000 (Schmitz e Riesner, 2006), quantificado em espectrofotômetro NanoDrop ND 8000 (Thermo Fisher Scientific, Waltham, EUA) e encaminhado para sequenciamento

Sanger, na ULRBMA do Instituto Biológico de São Paulo, em colaboração com o Dr. Ricardo Harakava.

4.6 Análises bioinformáticas

Para as análises bioinformáticas foram utilizados diversos softwares especializados. Para a montagem de contigs e o alinhamento de sequências, empregaram-se ferramentas como o Vector NTI, incluindo ClustalW, ContigExpress e AlignX (Thermo Fisher Scientific) (Lu e Moriyama, 2004) e o software MEGA v. 11.0.13 (Tamura *et al.*, 2021). As buscas de homologia e similaridade de sequências foram conduzidas utilizando os programas BLAST e BLASTX (NCBI, EUA) (Altschul *et al.*, 1990), bem como HHpred e HHblits (MPI Bioinformatics Toolkit, Tübingen, DE) (Zimmermann *et al.*, 2018; Gabler *et al.*, 2020).

As análises *in silico* das proteínas incluíram a utilização de ferramentas específicas para diferentes aspectos: NetNGlyc 1.0 (Gupta e Brunak, 2002) e NetOGlyc 4.0 (Steentoft *et al.*, 2013) foram utilizados para a predição de sítios de glicosilação; DeepLoc 2.1 (Ødum *et al.*, 2024) para predição da localização subcelular. Adicionalmente, os softwares DeepTMHMM 1.0 (Hallgren *et al.*, 2022) e TMHMM 2.0 (Sonnhammer *et al.*, 1998; Krogh *et al.*, 2001) foram empregados na identificação de domínios transmembrana. Já os softwares SignalP 4.1 (Petersen *et al.*, 2011) e TargetP 2.0 (Armenteros *et al.*, 2019) foram utilizados para a predição de peptídeos sinal e para a determinação do destino subcelular das proteínas com base nos peptídeos sinal, respectivamente. Por fim, o AlphaFold 3 foi empregado para previsão da estrutura terciária das proteínas (Jumper *et al.*, 2021; Mirdita *et al.*, 2022; Varadi *et al.*, 2024).

As regiões que flanqueiam ao códon de iniciação AUG foram analisadas a partir das ORFs de PfGSV, considerando uma janela de 6 nucleotídeos a montante e 6 nucleotídeos a jusante. A frequência das bases (conservação nucleotídica e os padrões) no contexto Kozak foram avaliados mediante WebLogos versão 2.8.2 (software disponível em www.weblogo.berkeley.edu) (Schneider e Stephens, 1990; Crooks *et al.* 2004), utilizando os parâmetros padrão de frequência relativa de cada nucleotídeo. A composição e arquitetura do contexto Kozak foram classificadas segundo critérios previamente descritos, considerando ótima a presença de uma base purina (A/G) na posição -3 e de

guanina (G) na posição +4, moderada quando apenas uma das posições foi atendida a esses critérios, e fraca quando ambas as bases estavam ausentes (Kozak, 1986; 1987).

4.7 Isolados analisados

Foram analisadas sequências do RNA 2 de 27 isolados de PfGSV, correspondentes à região entre as ORFs *p15* e *p61*. Dentre essas, 16 sequências foram obtidas e descritas no presente trabalho, representando novos isolados provenientes de diferentes estados brasileiros e do Paraguai. Quatro isolados adicionais (Prb11, Cmp02, ASU03 e Lim01), previamente sequenciados por meio de HTS por nosso grupo de pesquisa, também foram incluídos nas análises. Por fim, sete sequências complementares foram recuperadas de estudos anteriores e do banco de dados GenBank, totalizando o conjunto de 27 isolados analisados neste estudo (Tabela 2).

Tabela 2. Lista de isolados de PfGSV analisados neste estudo. A nomenclatura dos isolados refere-se ao local de coleta de cada uma das amostras.

Isolado	Gênero vegetal	Local de coleta	Ano	Nºacesso GenBank	Método de sequenciamento	Referência
Crd01	<i>Hibiscus</i>	Cordeirópolis - SP	2013	PQ762178	Sanger	Presente trabalho
ASU01	<i>Hibiscus</i>	Assunção - Paraguai	2015	PQ773497	Sanger	
Prb04	<i>Miltonia</i>	Piracicaba - SP	2014	PQ762185	Sanger	
Prb02	<i>Hibiscus</i>	Piracicaba - SP	2016	PQ762186	Sanger	
Prb05	<i>Nicotiana</i>	Piracicaba - SP	2016	PQ773496	Sanger	
Prb01	<i>Hibiscus</i>	Piracicaba - SP	2012	PQ762187	Sanger	
Cmp01	<i>Hibiscus</i>	Campinas - SP	2017	PQ762182	Sanger	
Prb06	<i>Hibiscus</i>	Piracicaba - SP	2019	PQ762188	Sanger	
SPa02	<i>Passiflora</i>	São Paulo - SP	2019	PQ762181	Sanger	
Lon01	<i>Hibiscus</i>	Londrina - PR	2020	PQ762179	Sanger	
Con01	<i>Ligusturm</i>	Conchas - SP	2021	PQ762180	Sanger	
Prb07	<i>Ligusturm</i>	Piracicaba - SP	2021	PQ762189	Sanger	
Prb09	<i>Tibouchina</i>	Piracicaba - SP	2023	PQ762190	Sanger	

Prb10	<i>Tibouchina</i>	Piracicaba - SP	2023	PQ762191	Sanger	
ASP01	<i>Thunbergia</i>	Águas de São Pedro - SP	2023	PQ762192	Sanger	
Cas01	<i>Hibiscus</i>	Castanhal - PA	2023	PQ753211	Sanger	
Prb11	<i>Ligustrum</i>	Piracicaba – SP	2014	-	HTS	Chabi-Jesus, dados não publicados
Cmp02	<i>Thunbergia</i>	Campinas – SP	2017	-	HTS	
ASU03	<i>Ligustrum</i>	Assunção - Paraguai	2017	PQ762184	HTS	
Lim01	<i>Eria</i>	Limeira - SP	2020	PQ762183	HTS	
Snp1	<i>Passiflora</i>	Sinop - MT	2016	NC_055652.1	HTS	Ramos-González <i>et al.</i> , 2020
Snp2	<i>Passiflora</i>	Sinop - MT	2016	MN746811.1	HTS	
NFo1	<i>Passiflora</i>	Nova Fronteira - MT	2018	MN746813.1	HTS	
BSB1	<i>Passiflora</i>	Brasília - DF	2003	MK804174.1	HTS	
BJL01	<i>Passiflora</i>	Bom Jesus da Lapa - BA	2008	MK804176.1	HTS	
CYPe246	<i>Passiflora</i>	Colômbia	2020	OP564896.1	HTS	Roy <i>et al.</i> , 2022
RsTHr269	<i>Hibiscus</i>	Colômbia	2022	PP115954.1	HTS	Roy <i>et al.</i> , 2024

5 RESULTADOS

5.1 Caracterizações de novos isolados de PfGSV

Foram sequenciados fragmentos de aproximadamente 1.500 nucleotídeos de sete novos isolados de PfGSV, provenientes de amostras coletadas entre 2013 e 2023 em diferentes estados brasileiros, como São Paulo e Pará, além de uma amostra oriunda do Paraguai (Figura 4). Esses isolados foram obtidos de oito gêneros vegetais, incluindo *Hibiscus*, *Ligustrum*, *Tibouchina* e *Thunbergia*.



Figura 4. Distribuição geográfica das amostras de PfGSV coletadas entre 2013 e 2023, incluindo 23 isolados provenientes do Brasil, 2 isolados da Colômbia e 2 isolados do Paraguai. O conjunto abrange amostras recentemente sequenciadas por Sanger neste trabalho, sequências previamente geradas por HTS por nosso grupo de pesquisa (ASU03, Prb11, Cmp02 e Lim01) e sequências depositadas no GenBank.

5.2 Variabilidade genética das ORFs *p13*, *p11* e $\Delta p11$ de PfGSV

Dois alelos da ORF *p13*, designados como *p13* e *p11*, foram previamente identificados (Ramos-González *et al.*, 2020). Um terceiro alelo, denominado $\Delta p11$, foi encontrado em três isolados, um de Limeira, SP (PfGSV_Lim01) e dois do Paraguai (PfGSV_ASU01 e PfGSV_ASU03). O alelo *p13* contém 330 nucleotídeos e codifica uma

proteína tentativa de 110 aa, enquanto o alelo *p11* possui 282 nucleotídeos e codifica uma proteína de 94 aa. O alelo $\Delta p11$, por sua vez, apresenta 177 nucleotídeos e codifica uma proteína de 59 aa.

5.2.1 O alelo *p13* é o prevalente entre o conjunto de amostras analisadas

O alelo *p13*, de 330 nucleotídeos, é o prevalente nos isolados estudados. *p13* foi identificado em 23 isolados, coletados nos estados brasileiros de São Paulo, Mato Grosso, Paraná, Pará e Distrito Federal, além de isolados da Colômbia, entre os anos de 2003 e 2023. Os isolados nos quais o alelo *p13* foi identificado são provenientes de amostras pertencentes aos gêneros *Hibiscus*, *Ligustrum*, *Miltonia*, *Nicotiana*, *Thibouchina* e *Thunbergia*, além de *Passiflora*.

A sequência de nucleotídeos da ORF *p13* é altamente conservada entre os isolados, apresentando valores de identidade que variam de 97,6 a 100%. Foram identificadas 18 mutações na comparação da sequência de nucleotídeos da ORF *p13*, sendo nove sinônimas, que não alteram o aminoácido codificado, e nove não sinônimas, nas quais ocorre a substituição do aminoácido (Tabela 3).

Tabela 3. Mutações sinônimas e não sinônimas identificadas nas sequências da ORF *p13*. Os códons marcados por um “*” e dois “**” (asteriscos) representam mais de uma mutação na mesma trinca em diferentes isolados. Os códons foram numerados a partir do códon de início AUG (1º códon). Estão sinalizados, em vermelho, os nucleotídeos que sofreram mutações.

Mutações sinônimas						
Localização códon	Códon	Aminoácido	Isolados	Códon mutado	Aminoácido	Isolados
9º	AAU	Asp	22 isolados	AAC	Asp	Prb07, Prb09 e Prb10
10º	UAC	Tyr	22 isolados	UAU	Tyr	Sn01
43º	GAG	Glu	22 isolados	GAA	Glu	CYPe-246 e RsTHr-269
49º	CAU	His	22 isolados	CAC	His	Cas01
63º	UCU	Ser	22 isolados	UCG	Ser	Sn02
80º	AUA	Ile	22 isolados	AUU	Ile	Prb07
88º	GCG*	Ala	22 isolados	GCA	Ala	Con01
89º	AAU	Asn	22 isolados	AAC	Asn	Crd01
98º	AGA**	Arg	16 isolados	AGG	Arg	BSB01, Prb01, Prb02, Prb04, Prb05, Prb06
Mutações não sinônimas						

3°	AUA	Ile	22 isolados	GUA	Val	Cas01
34°	GUG	Val	22 isolados	AUG	Met	Sn02
40°	GUA	Val	22 isolados	AUA	Ile	Cas01
46°	CGU	Arg	21 isolados	UGU	Cys	ASP01 e NF001
56°	AUU	Ile	21 isolados	GUU	Val	Prb06
65°	AGA	Arg	21 isolados	GGA	Gly	Sn02
88°	GCG*	Ala	21 isolados	ACG	Thr	Cas01
96°	GGC	Gly	22 isolados	GAC	Asp	Lon01
98°	AGA**	Arg	16 isolados	AAA	Lys	ASP01

A proteína P13 é altamente conservada entre os isolados (Figura 5). A identidade na sequência dos aminoácidos varia entre 95,5% e 100%, sendo os menores valores entre os isolados de Sn02/Cas01 e ASP01/Cas01. Os resultados da análise de predição de homologia realizada no HHpred (Zimmermann *et al.*, 2018; Gabler *et al.*, 2020) sugerem que a proteína P13 possui similaridade remota com proteínas que desempenham funções regulatórias e estruturais. A principal correspondência é uma proteína com estrutura de dedo de zinco, frequentemente associada à ligação ao DNA e regulação transcricional (PDB 1M36_A), sugerindo um possível papel na regulação da expressão gênica. Outras possíveis relações de identidade incluem proteínas com domínios LIM, conhecidos por mediar interações proteína-proteína, geralmente envolvidas na arquitetura do citoesqueleto e na sinalização celular. Esses resultados fornecem uma base para hipóteses sobre as possíveis funções biológicas da sequência-alvo, com plausíveis papéis em regulação genética, interação celular e resposta imune.

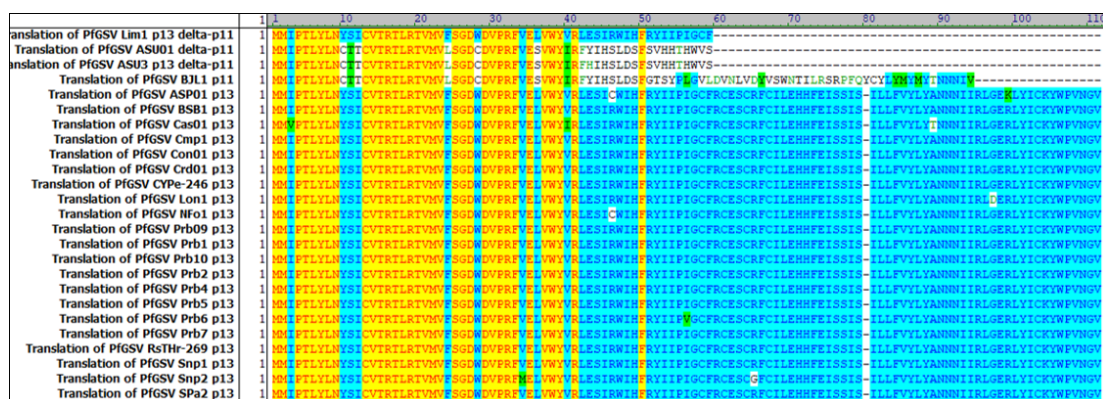


Figura 5. Alinhamento da sequência de aminoácidos das proteínas ΔP11, P11 e P13 dos isolados de PfGSV. O alinhamento foi realizado utilizando o software Vector NTI, evidenciando as regiões conservadas. Os aminoácidos destacados em amarelo dizem respeito às posições altamente conservadas em todos os isolados; em azul, estão representados os aminoácidos com conservação parcial e divergências entre os isolados; em verde, estão destacadas as substituições pontuais entre os isolados.

A análise dos resultados de predição de localização subcelular das proteínas P13 foi realizada utilizando DeepLoc-2.1 (Ødum *et al.*, 2024). A proteína predita P13 possui uma distribuição tentativa predominantemente no citoplasma e na mitocôndria, com valores de probabilidade de predição variando entre 0,40 até 0,50 em ambas as localizações celulares. No entanto, os valores para núcleo e meio extracelular também foram consideráveis, entre 0,30 e 0,40. O perfil transmembrana de todas as proteínas P13 é consistente, apresentando o valor de probabilidade de 0,70 em todas as sequências. Os valores de probabilidade de solubilidade das proteínas P13 também foi consistente, variando de 0,56 a 0,61 em todas as 21 sequências.

Os resultados obtidos pela análise preditiva de N-glicosilação utilizando o software NetNGlyc 1.0 (Gupta e Brunak, 2022) indicam uma alta probabilidade de glicosilação para P13. Em todas as sequências, o sítio de glicosilação NYSI foi recorrentemente detectado, e a concordância entre os nove métodos de predição utilizados pelo software foi elevada, com o valor de consenso estatístico de 0,79. Por outro lado, a análise preditiva de O-glicosilação, utilizando o software NetOGlyc 4.0 (Steentoft *et al.*, 2013), não previu nenhum sítio de possível O-glicosilação para P13.

5.2.2 Alelo *p11* está presente em um único isolado

O alelo *p11* foi identificado em apenas um isolado, coletado em Bom Jesus da Lapa, na Bahia, no ano de 2008. Esse alelo tem 282 nucleotídeos e codifica uma proteína prevista de 94 aa. Os resultados da análise de homologia, obtidos a partir do software HHpred, indicam que P11 apresenta uma similaridade remota com o "Fator de estabilidade de mRNA e regulador da meiose 1" (Marf1), uma proteína envolvida na divisão celular e metabolismo do mRNA. Essa proteína possui o domínio NYNI, característico de nucleases, o que sugere uma função catalítica relacionada ao processamento de ácidos nucleicos. A associação com atividades de hidrólise e estabilidade de mRNA indica um possível papel na regulação pós-transcricional, essencial para o controle da expressão gênica durante a meiose e outros processos celulares.

A localização subcelular prevista pelo DeepLoc 2 demonstrou que P11 possivelmente localiza-se no citoplasma e núcleo com valores de probabilidade de predição de 0,43 e 0,52, respectivamente. Além disso, o caráter solúvel de P11 foi

semelhante ao observado pelas proteínas P13. No caso da P11_BJL01, a probabilidade para o perfil transmembrana foi 0,44, menor do que aquele calculado para P13. A análise de previsão de N-glicosilação indicou uma alta probabilidade (0,77), assim como P13, no entanto, apontou o sítio de glicosilação NCTT ao invés de NYSL.

5.2.3 Alelo $\Delta p11$

O alelo $\Delta p11$ foi identificado em três isolados: Lim01, ASU01 e ASU03. O Lim01 é oriundo de uma amostra de orquídea *Eria* sp., coletada em Limeira, São Paulo, no ano de 2020. O isolado ASU01 foi identificado em uma amostra de *Hibiscus* sp., coletada em Assunção, no Paraguai, em 2015, enquanto o ASU03 é uma amostra de *Ligustrum* sp., coletada em 2017, no mesmo município. A identidade da sequência nucleotídica das ORFs $\Delta p11$ de cada um dos isolados é alta, os isolados ASU01 e ASU03 possuem uma maior identidade entre si do que quando comparados ao isolado Lim01 (Tabela 4).

Tabela 4. Percentual de identidade nucleotídica entre os alelos $\Delta p11$ nos isolados ASU01, ASU03 e Lim01.

$\Delta p11$	ASU01	ASU03	Lim01
ASU01	-	99,4	90,4
ASU03	99,4	-	90,4
Lim01	90,4	90,4	-

O alelo $\Delta p11$ é composto de 177 nucleotídeos e codifica uma proteína prevista de 59 aa. A análise de similaridade realizada no HHpred revelou correspondências remotas com proteínas envolvidas principalmente na regulação da transcrição e estabilidade do RNA. Dentre as proteínas identificadas, observou-se homologia com subunidades do fator de iniciação de transcrição IIA (TFIIA), que estão associadas ao complexo TFIIA/TBP/DNA, essencial para o recrutamento da RNA polimerase e o início da transcrição. Por fim, certa similaridade remota foi observada com um regulador de meiose e fator de estabilidade de mRNA, responsável por modular a estabilidade de transcritos específicos durante a meiose, o que coincide com os resultados para P11.

A previsão de localização subcelular, realizada pelo DeepLoc 2.1, das proteínas $\Delta P11$ dos isolados ASU01 e ASU03 foi mais semelhante entre si quando comparadas ao isolado Lim01. A previsão para $\Delta P11_{ASU01}$ e $\Delta P11_{ASU03}$ apresentou valores de probabilidades mais altos para localizações celulares como citoplasma e mitocôndria,

0,54 e 0,66, respectivamente, quando comparados aos valores de $\Delta P11_Lim01$ para as mesmas localizações celulares, 0,39 e 0,57. A predição de N-glicosilação, realizada no software NetNGlyc 1.0, apontou alta probabilidade para as proteínas $\Delta P11$, assim como para P13 e P11. Para a proteína $\Delta P11_Lim01$ foi previsto o sítio de glicosilação NYSI, assim como para todas as P13. No entanto, para as proteínas $\Delta P11$ dos isolados ASU01 e ASU03 foi previsto o sítio de glicosilação NCTT, assim como para a proteína P11_BJL01.

5.3 Estrutura da ORF entre os alelos *p13*, *p11* e $\Delta p11$

A análise comparativa dos três alelos revela uma alta conservação nucleotídica entre eles. O alelo $\Delta p11$ apresenta-se como uma versão truncada do alelo *p11*, que, por sua vez, aparenta ser uma forma truncada do alelo *p13*. Os alelos dos 27 isolados apresentam uma alta conservação, principalmente na primeira metade da ORF, sugerindo que essa região, se traduzida, provavelmente codifica um trecho de função crucial para a proteína.

Quando se compara o trecho genômico de BJL01 correspondente ao alelo *p13_Snp01*, é possível identificar a inserção de uma uracila na posição 285, que cria um códon de parada âmbar (UAG) inexistente em Snp01 (Figura 6), interrompendo a leitura e originando a ORF *p11*. Segundo os dados do HTS, a sequência do isolado BJL01 tem cobertura de 1.000x, o que fornece evidências consistentes para a presença da inserção na posição 285. A sequência de *p11_BJL01* tem 87,8% de identidade de nucleotídeo quando comparado a *p13_Snp01*, incluindo o códon de parada de *p13_Snp01*, que está conservado na sequência de BJL01, apesar a perturbação da leitura das trincas.

		Section 1				
	(1)	1	10	20	30	40 50
BJL01 p11 +55 nts*	(1)	AUGAUGAUACCAACUUU	GUAUCUUAAC	UGUACCA	CAUGUGUCACCAGAAC	
Snpl01 p13 stop*	(1)	AUGAUGAUACCAACUUU	AUAUCUUAAC	UACUCCA	U AUGUGUCACCAGAAC	
		Section 2				
	(51)	51	60	70	80	90 100
BJL01 p11 +55 nts*	(51)	UCUUAGGACCGUAAUGGUGUU	AGUGGAGACUG	UGAUGUCCCACGCUUUG		
Snpl01 p13 stop*	(51)	UCUUAGGACCGUAAUGGUGUU	AGUGGAGACUG	GGAUGUCCCACGCUUUG		
		Section 3				
	(101)	101	110	120	130	140 150
BJL01 p11 +55 nts*	(101)	UGGAAU	CGUUUGGUAA	UAAGGUUU	UAUCCAUCGUUGGAUUCAUU	
Snpl01 p13 stop*	(101)	UGGAAU	UAUUUGGUAA	UAAGGUUA	G-UCCAUCGUUGGAUUCAUU	
		Section 4				
	(151)	151	160	170	180	190 200
BJL01 p11 +55 nts*	(150)	-	CGGUACAUCAUACCCA	UUGGGUGU	CUAGAUGUGAAUCUUGUAGAUAU	
Snpl01 p13 stop*	(150)	U	CGGUACAUCAUACCCA	UUGGGUGU	UUAGAUGUGAAUCUUGUAGAUAU	
		Section 5				
	(201)	201	210	220	230	240 250
BJL01 p11 +55 nts*	(199)	SUAUCUUGGAACACCAU	UUUGAGAUUCUGUCCA	UUCAAUAUUG	UUUAUU	
Snpl01 p13 stop*	(200)	SUAUCUUGGAACACCAU	UUUGAGAUUCUGUCCA	UUCAAUAUUG	CUUAUU	
		Section 6				
	(251)	251	260	270	280	290 300
BJL01 p11 +55 nts*	(249)	SUAUA	UGUAUA	UGUAUA	CGAAUAACAAUAUAGUU	UAGAUAUAGCGAGACA
Snpl01 p13 stop*	(250)	SUAUA	----CU	UGUAUG	CGAAUAACAAUAU	CAUU-AGUAUAGCGAGAGA
		Section 7				
	(301)	301	310	320	330	
BJL01 p11 +55 nts*	(299)	CUGUAUAU	UGUAAAUACUGGCCA	GUUCCUGGCGUGUAA		
Snpl01 p13 stop*	(295)	CUGUAUAU	UGUAAAUACUGGCCG	GUCAACGGCGUGUAA		

S

Figura 6. Alinhamento da sequência de nucleotídeos das ORFs *p11*_BJL01 e *p13*_Snpl01, realizado com o software Vector NTI. As regiões conservadas entre as sequências estão destacadas em amarelo. O ponto em que ocorre a adição de uma uracila, responsável pela introdução de um códon de parada prematuro, está destacado na imagem.

Observa-se que os nucleotídeos na posição 1 a 177 de *p13*_Snpl01 são 100% idênticos aos do alelo $\Delta p11$ _Lim01. A troca de um único nucleotídeo na posição 178 na sequência de Lim01 cria um códon de parada opala (UGA), interrompendo o alelo $\Delta p11$ _Lim01. O nucleotídeo 178 de *p13*_Snpl01 é uma adenina, que no caso de $\Delta p11$ _Lim01 foi mutado para uracila, formando o códon de parada UGA (Figura 7). A mutação em $\Delta p11$ _Lim01 é sustentada por uma cobertura de mais de 6.000x, eliminando a possibilidade de que esta substituição seja um artefato técnico. Em análises de mutações *in silico*, ao trocar de U para A, a sequência após o códon de parada de $\Delta p11$ _Lim01 tem 99,7% de identidade com *p13*_Snpl01, o que demonstra que a sequência para a informação da proteína P13 está presente na sequência do isolado Lim01, no entanto, está interrompida por uma única mutação não sinônima. Além disso, o códon de parada de *p13*_Snpl01 (nucleotídeos 331-333) está presente na sequência de Lim01.

						Section 1	
	(1)	1	10	20	30	48	
Lim01 delta_p11 +160*	(1)	AUGAUGAUACCAACUUUAUUCUUAAUUAUACUCCAUUGUGUCACCAGA					
Snp01 p13 stop*	(1)	AUGAUGAUACCAACUUUAUUCUUAAUUAUACUCCAUUGUGUCACCAGA					
						Section 2	
	(49)	49	60	70	80	96	
Lim01 delta_p11 +160*	(49)	ACUCUUAGGACCGUAAUGGUGUUUAGUGGAGACUGGGAUGUCCACGC					
Snp01 p13 stop*	(49)	ACUCUUAGGACCGUAAUGGUGUUUAGUGGAGACUGGGAUGUCCACGC					
						Section 3	
	(97)	97	110	120	130	144	
Lim01 delta_p11 +160*	(97)	UUUGUGGAAUUAUGUUUGGUAUGUAAGGUUAGAGUCCAUUCGUUGGAUU					
Snp01 p13 stop*	(97)	UUUGUGGAAUUAUGUUUGGUAUGUAAGGUUAGAGUCCAUUCGUUGGAUU					
						Section 4	
	(145)	145	150	160	170	180	192
Lim01 delta_p11 +160*	(145)	CAUUUUCGGUACAUCUAUACCAUUGGGUGUUUUAGAGUGGAAUCUUGU					
Snp01 p13 stop*	(145)	CAUUUUCGGUACAUCUAUACCAUUGGGUGUUUUAGAGUGGAAUCUUGU					
						Section 5	
	(193)	193	200	210	220	230	240
Lim01 delta_p11 +160*	(193)	AGAUCUCUGUAUCUUGGAACACCAUUUUGAGAUCUCGUCCAUUCUCAAUA					
Snp01 p13 stop*	(193)	AGAUCUCUGUAUCUUGGAACACCAUUUUGAGAUCUCGUCCAUUCUCAAUA					
						Section 6	
	(241)	241	250	260	270	288	
Lim01 delta_p11 +160*	(241)	UUGCUAUUCGUUAUACAUGUAUGCGAAUAACAAUAUCAUUAAGGUC					
Snp01 p13 stop*	(241)	UUGCUAUUCGUUAUACAUGUAUGCGAAUAACAAUAUCAUUAAGGUC					
						Section 7	
	(289)	289	300	310	320	333	
Lim01 delta_p11 +160*	(289)	GAGAGACUGUAUUAUUUGUAAUACUGGCCGGUCAACGGCGUGUAA					
Snp01 p13 stop*	(289)	GAGAGACUGUAUUAUUUGUAAUACUGGCCGGUCAACGGCGUGUAA					

Figura 7. Alinhamento das seqüências de nucleotídeos das ORFs *p13_Snp01* e *Δp11_Lim01*, realizado no Vector NTI. As regiões conservadas são estacadas em amarelo, observa-se na seqüência de *Δp11_Lim01* a presença de uma mutação sem sentido que introduz um códon de parada prematuro.

Nas comparações entre as ORFs *p13_Snp01* e *Δp11_ASU01* foi possível identificar um fenômeno diferente daquele que gerou o códon de parada em *Lim01*. No caso do isolado *ASU01*, identificou-se que a adição de um nucleotídeo de adenina na posição 130 da seqüência de *Δp11_ASU01* alterou o marco de leitura de códons, formando um códon de parada âmbar (UAG) entre os nucleotídeos 178-180 (Figura 8). Na seqüência de *p13_Snp01*, os três nucleotídeos que formam o códon de parada de *Δp11_ASU01* estão presentes, porém, em códons diferentes (₁₇₅UUU AGA₁₈₀). As seqüências nucleotídicas de *p13_Snp01* e *Δp11_ASU01* possuem 91% de identidade dos nucleotídeos 1 ao 177, onde encontra-se o códon de parada de *Δp11_ASU01*. A identidade de seqüência após o códon de parada de *Δp11_ASU01* é de 88,2%, o que demonstra que a informação de codificação de uma proteína P13 tentativa, se fosse traduzida, está quase completa na seqüência do isolado *ASU01*. Da mesma forma, na simulação de mutações *in silico* com caráter preditivo na seqüência de *Δp11_ASU01*, na qual se retira a adenina no nucleotídeo 130, a seqüência passa a ter 90,2% de identidade com *p13_Snp01* e 97,1% com *p11_BJL01*. Nesse caso, ocorre uma segunda mutação na seqüência do isolado *ASU01* nos nucleotídeos 256 e 257, onde ocorre a deleção de uma citosina e uma uracila presentes em *p13_Snp01*, que novamente perturba o marco de leitura dessa ORF em *ASU01*, criando um códon de parada nos nucleotídeos 265-267 (UAA). Os nucleotídeos

desse suposto códon de parada estão presentes na sequência do isolado Snp01, no entanto, estão em trincas distintas.

		Section 1				
ASU01 delta_p11 +160 nts*	(1)	1	10	20	30	45
Snp01 p13 stop*	(1)	AUGAUGAUACCAACUUUGUAUCUUAACUGUAACCAUGUGUCACC				
		Section 2				
ASU01 delta_p11 +160 nts*	(46)	46	60	70	80	90
Snp01 p13 stop*	(46)	AGAACUCUUAGGACCGUAAUGGUGUUAGUGGAGACUGUGAUGUC				
		Section 3				
ASU01 delta_p11 +160 nts*	(91)	91	100	110	120	135
Snp01 p13 stop*	(91)	CCACGCUUUGUGGAAUCGUAUAGGUUUAUAJCCAU				
		Section 4				
ASU01 delta_p11 +160 nts*	(136)	136	150	160	170	180
Snp01 p13 stop*	(135)	UCGUUGGAUUCAUUUUCGGUACAUCUACCAUUGGGUGUCUAG				
		Section 5				
ASU01 delta_p11 +160 nts*	(181)	181	190	200	210	225
Snp01 p13 stop*	(180)	AUGUGAAUCUUGUAGAUUAUGUACUUGGAACACCAUUUUGAGAU				
		Section 6				
ASU01 delta_p11 +160 nts*	(226)	226	240	250	260	270
Snp01 p13 stop*	(225)	CUCGUCCAUCUCAAAUUGUUAUUUGUAUA--UGUAUACGAUAA				
		Section 7				
ASU01 delta_p11 +160 nts*	(271)	271	280	290	300	315
Snp01 p13 stop*	(270)	CAAUAUAGUUAGAUAUGACGAGACACUGUAUAUUGUAAAUAUCU				
		Section 8				
ASU01 delta_p11 +160 nts*	(316)	316	335			
Snp01 p13 stop*	(314)	GGCCAGUUCUUGGCGUGUAA				

Figura 8. Alinhamento das sequências de nucleotídeos das ORFs *p13_Snp01* e $\Delta p11_ASU01$, realizado no Vector NTI. As regiões conservadas estão destacadas em amarelo, e o ponto de adição de uma adenina, responsável pela introdução de um códon de parada prematuro na sequência de $\Delta p11_ASU01$, está assinalado na imagem.

5.4 ORFs *p7*: variabilidade genética e possível conservação funcional

Ainda na região considerada como intergênica, entre as ORFs *p15* e *p61*, à montante dos alelos das ORFs *p13/p11/Δp11* foram observados alelos de uma segunda ORF órfã (ORFan, do inglês *open reading frame orphan*) conservada nos 27 isolados analisados, denominada *p7*. Foi possível identificar seis alelos da ORF *p7* (Tabela 5). Apesar da divergência entre os tamanhos dos alelos entre os isolados, é possível identificar uma alta conservação na região central entre todos os alelos, o que poderia apontar uma região de importância para a funcionalidade da tentativa proteína expressada.

Tabela 5. Prevalência dos alelos da ORFs *p7* entre os isolados de PfGSV.

Alelo da ORF <i>p7</i> (nts)	195	141	126	111	108	63
Isolados	Snp01	NFo01				
		ASP01				
		BSB01				
		Cas01				
		Cmp01				
		Con01				
		Crd01				
		CYPe-246				
		Prb01				ASU01
		Prb02	Lon01	Prb07	Lim01	ASU03
		Prb04				BJL01
		Prb05				
		Prb06				
		Prb09				
		Prb10				
		Prb11				
		Cmp02				
		RsTHr-269				
		Snp02				
		Spa02				

O alelo predominante apresenta uma sequência de 141 nucleotídeos, codificando uma proteína tentativa composta por 47 resíduos, sendo identificado em 20 dos 27 isolados analisados. O segundo alelo mais frequente possui 63 nucleotídeos, está presente em 3 isolados e codifica uma proteína tentativa de 21 aa. Os demais alelos foram observados em apenas um isolado cada, sendo eles: um alelo com 195 nucleotídeos, um com 126 nucleotídeos, um com 111 nucleotídeos e outro com 108 nucleotídeos (Figura 9).

	1	10	20	30	40	50	60
Translation of PfGSV Snp1 p7	1	MVVLNFTDSCVIRTLRTV	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV ASP01 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV BSB1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Cas01 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Cmp1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Con01 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Crd01 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV CYPe-246 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Nfo1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb09 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb10 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb2 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb4 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb5 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb6 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV RsThr-269 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Snp2 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV SPa2 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Lon1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Prb7 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV Lim1 p7	1	-----	MVLSGRWDAPHLLSVCLVRLISLLVCHPVVLLYWEFDWSDHLMSSSL				
Translation of PfGSV ASU01 p7	1	-----	MFCALDFAVSLSFSGSTILGI				
Translation of PfGSV ASU03 p7	1	-----	MFCALDFAVSLSFSGSTILGI				
Translation of PfGSV BJL01 p7	1	-----	MFRALDLAVSLSFSGSTILGI				

Figura 9. Alinhamento das sequências de aminoácidos das proteínas P7 de PfGSV, realizado no Vector NTI. Os resíduos destacados em amarelo correspondem às posições altamente conservadas entre todos os isolados; em azul, indicam conservação parcial e divergências entre os isolados; e em verde, estão assinaladas substituições pontuais entre os isolados.

As proteínas P7 tentativas apresentaram ampla diversidade em termos de características físico-químicas e estruturais. Quanto à composição de resíduos, observou-se uma predominância de resíduos hidrofóbicos, como alanina, leucina e isoleucina, em proteínas maiores, enquanto as menores apresentaram maior proporção de resíduos polares, como serina e treonina. As proteínas com maior quantidade de resíduos hidrofóbicos, como aquelas com 47 aa ou mais, podem estar associadas a ambientes lipídicos, enquanto as proteínas com pI (ponto isoelétrico) mais baixos, como as dos isolados ASU01 e ASU03, tendem a possuir maior carga negativa, o que poderia facilitar interações em meio aquoso. Nas previsões de estruturas de proteínas geradas a partir do AlphaFold2, observa-se um padrão estrutural predominante helicoidal, com a presença de α -hélices em todas as proteínas (Figura 10).

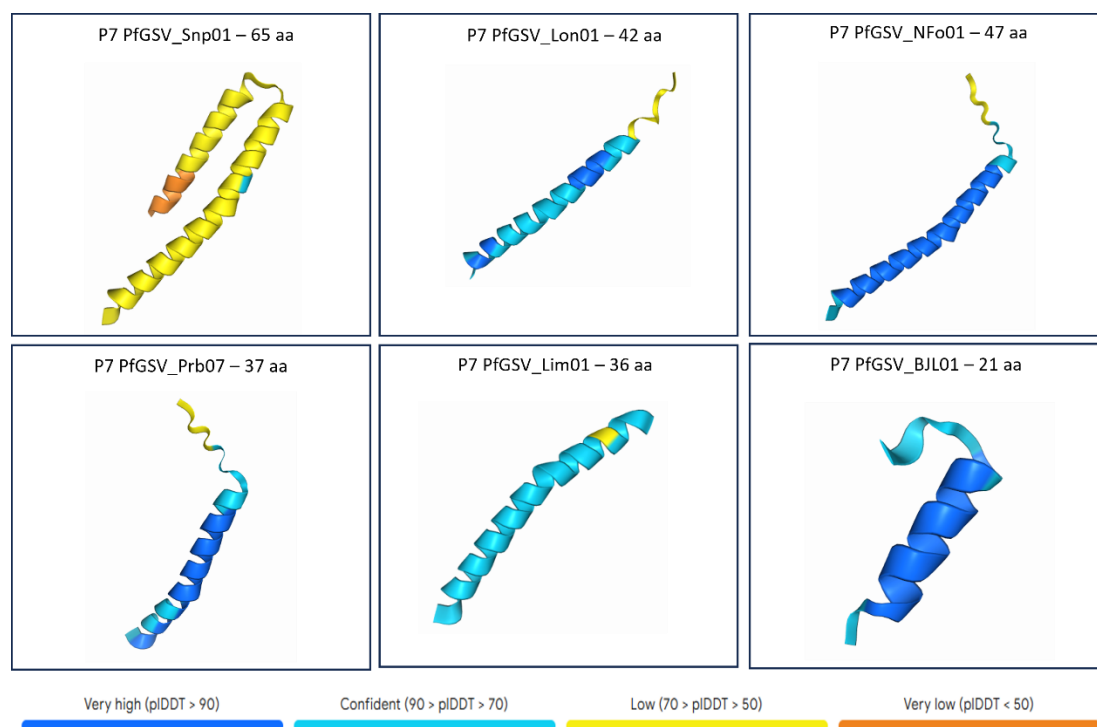


Figura 10. Modelos estruturais preditos das proteínas P7 de diferentes isolados do PfGSV, gerados com o AlphaFold 3. As representações ilustram as estruturas tridimensionais previstas das proteínas P7 e os respectivos comprimentos em aminoácidos. As cores correspondem ao nível de confiança da predição.

A análise preditiva de domínios transmembrana demonstrou que a maioria das proteínas P7 apresenta um domínio transmembrana (TM) típico, geralmente localizado entre os resíduos 12 e 34, com regiões intracelulares e extracelulares adjacentes ao domínio TM. No entanto, em três alelos foram observadas ligeiras variações nas previsões. Para o alelo *p7_Snp01*, o domínio TM foi identificado entre os aminoácidos 30 e 52; no caso de *p7_Lon01*, entre os resíduos 15 e 37; e no isolado *p7_Lim01*, entre os aminoácidos 5 até 25, com probabilidades (*p*) de previsão que variaram entre 0,70 e 0,98. Além disso, o domínio transmembrana previsto para *p7_Lim01* é composto por 21 aa, enquanto para as demais proteínas P7, os domínios TM previstos têm 23 resíduos. Em contrapartida, no três alelos correspondentes aos isolados ASU01, ASU03 e BJL01, não apresentaram regiões TM previstas, sendo esses os alelos de menor comprimento, o que poderia sugerir uma possível perda de funcionalidade.

Os resultados obtidos na predição de localização subcelular indicam que a maioria das proteínas analisadas apresenta uma alta probabilidade de estar associada ao sistema de membranas celulares. As probabilidades atribuídas para a associação com outros compartimentos, como o citoplasma e a membrana externa, são significativamente

menores, mas ainda relevantes para algumas proteínas específicas, como ASU03 e BJLO1, que apresentaram valores de probabilidade de 0,40 e 0,16 para o citoplasma, respectivamente, e Crd01, com 0,21 para membranas celulares. As probabilidades relacionadas à parede celular ou espaço extracelular foram consistentemente muito baixas ($<0,003$), indicando pouca ou nenhuma associação dessas proteínas com essas regiões.

Ao comparar o alelo *p7_Snp01* com o alelo *p7_ASU01*, observa-se que a diferença entre o maior alelo (195 nucleotídeos) e a menor versão (63 nucleotídeos) pode ser explicada por uma mutação pontual que resultou na substituição do códon de início de *p7_Snp01* (AUG → AUA) (Figura 11A). Essa mutação impede a tradução da ORF na posição inicial observada em *Snp01*, fazendo com que a ORF *p7* dos isolados ASU01, ASU03 e BJL01 seja iniciada após 98 nucleotídeos, no próximo códon AUG, que também está presente na sequência de *p7_Snp01*. No caso do isolado Prb07, o alelo *p7* inicia 54 nucleotídeos após o início de *p7_Snp01* e termina 30 nucleotídeos antes. Isso ocorre devido a duas mutações distintas: uma que altera o códon de início de *Snp01* em *Prb07* e outra que resulta em uma mutação sem sentido, gerando um códon de parada prematuro (UGA) na sequência de *Prb07* (Figura 11B).

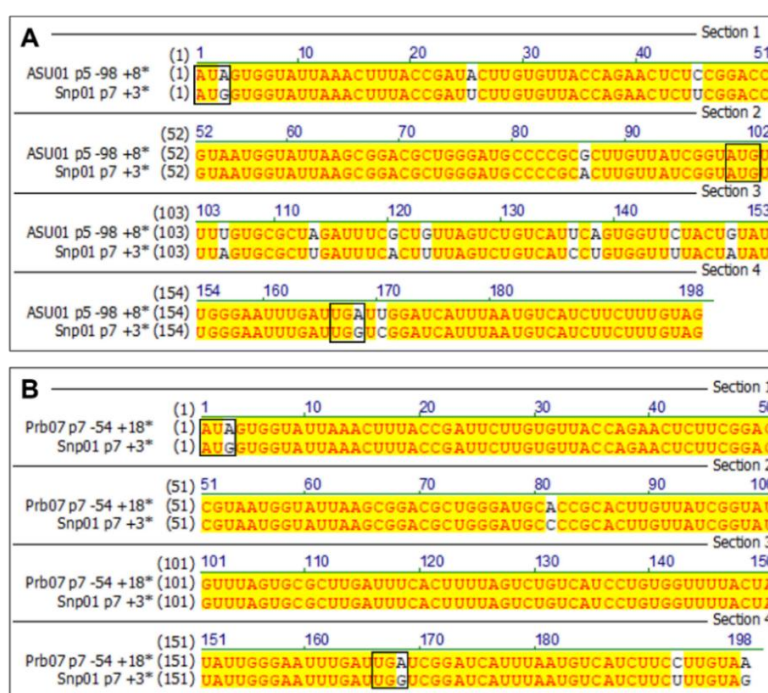


Figura 11. Alinhamento das sequências de nucleotídeos de alelos da ORF *p7*, realizado no Vector NTI. **A** – Comparação entre os isolados *p7_ASU01* e *p7_Snp01*, destacando mutações pontuais que introduziram códon de início e de parada. **B** – Comparação entre os isolados *p7_Pr07* e *p7_Snp01*, evidenciando mutações pontuais responsáveis pela adição de códon de início e de parada. As regiões conservadas estão indicadas em amarelo, e as alterações específicas estão assinaladas em suas respectivas posições.

5.5 Análises de sequência Kozak

As análises das sequências Kozak das ORFs *RdRp* e *p29*, realizadas a partir de 11 isolados com genoma completo disponíveis, revelaram alta conservação (Figura 12 A e B). Ambos os conjuntos de sequências apresentam o modelo canônico de Kozak, o que sugere uma alta eficiência no início de tradução – característica compatível com ORFs virais essenciais. O perfil de conservação nucleotídica da ORF *p15* (Figura 12C) demonstra contexto Kozak funcional, com predominância de A na posição -3, fator que favorece a eficiência traducional. Embora a posição +4 exiba predominância de T, o contexto geral sugere capacidade moderada de iniciação, possivelmente refletindo uma regulação da expressão genica durante o ciclo viral.

A ORF *p7* apresenta um padrão distinto (Figura 12D), com conservação significativa das G nas posições -3 e +4, configurando um contexto Kozak fortemente favorável à tradução. A análise da ORF *p13* (Figura 12E) revelou alto grau de conservação nos 15 nucleotídeos flanqueadores da sequência Kozak. A presença de A na posição -3 sugere reconhecimento ribossômico eficiente, enquanto a A na +4, embora neutra ou levemente favorável, não oferece o mesmo nível de eficiência que a presença da G. Dessa forma, o contexto Kozak da ORF *p13* pode ser classificado como uma eficiência intermediária de início de tradução, sugerido uma estratégia viral durante o ciclo de infecção.

As ORFs *p61*, *mp* e *p24* exibem contextos Kozak altamente conservados (Figura 12 F-H). Particularmente as ORFs *MP* e *p24* apresentam a combinação A em -3 e G em +4, considerada a configuração ótima para início de tradução. A ORF *p61* tem presença de G tanto em -3 quanto +4, configuração que, embora ligeiramente menos eficiente que o padrão canônico, ainda mantém capacidade traducional elevada.

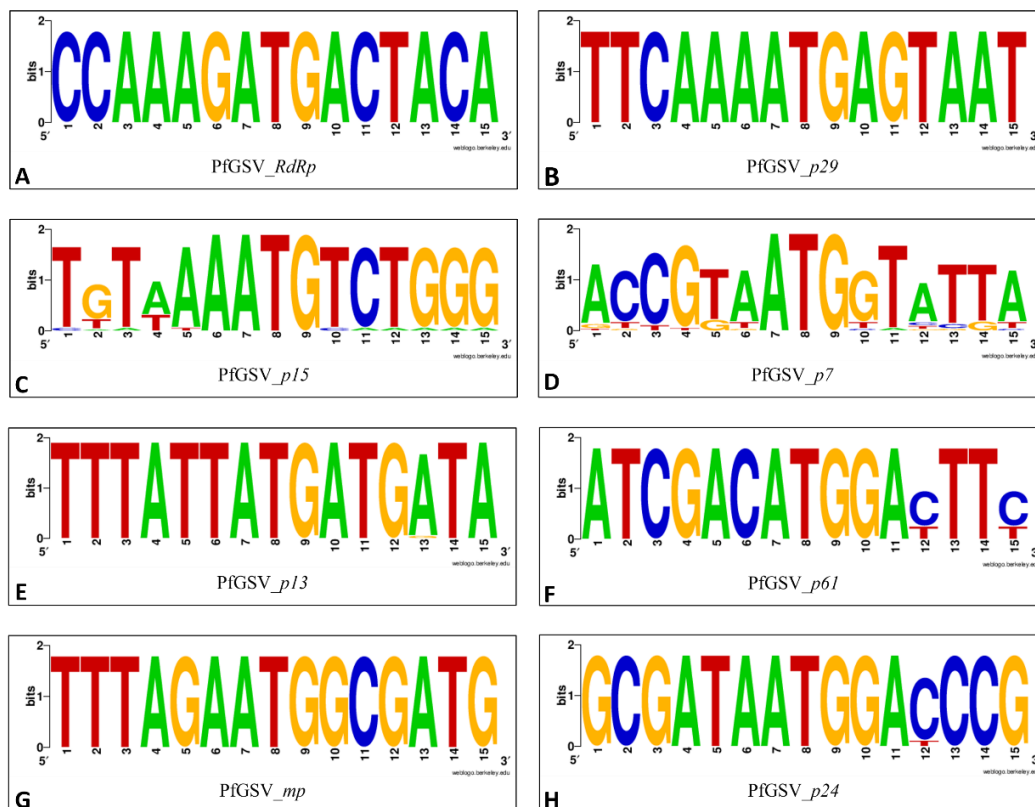


Figura 12. WebLogos criados a partir dos nucleotídeos -6 e +6 flanqueadores dos códons de início das ORFs no genoma de PfGSV. **A, B, F, G e H:** criados a partir da sequência de 11 isolados (CYPe-246, Snp01, Nf001, Snp02, BJL01, BSB01, RsTHr-269, Prb11, Lim01, ASU03 e Cmp02). **C, D e E:** criados a partir da sequência dos 27 isolados descritos no trabalho.

6 DISCUSSÃO

O presente estudo acerca da região 5' do RNA2 de passion fruit green spot virus contribui ao elucidar e levantar hipóteses sobre a dinâmica genômica e possíveis interações bióticas deste cilevírus. Os resultados obtidos com a análise de novos isolados de PfGSV confirmam a diversidade genômica do vírus, especialmente na região entre as ORFs *p15* e *p61* (Ramos-González *et al.*, 2020). Essa região, anteriormente considerada não codificante, aparenta atuar como um “berço” para novas ORFs no genoma de PfGSV, demonstrando potencial para gerar novos elementos genéticos que poderiam contribuir com a aptidão viral (Ramos-González *et al.*, 2021). Essa análise revelou a presença de alelos distintos (*p13*, *p11*, $\Delta p11$), reforçando o possível papel dessa região na adaptação e interação com seus hospedeiros vegetais. Essa característica sugere que o RNA2 de PfGSV pode atuar como um reservatório adaptativo, favorecendo a interação com múltiplos hospedeiros (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012; Guerra-Almeida e Nunes-da-Fonseca, 2020). A identificação de diferentes alelos de ORFs emergentes em distintos isolados sugere que essa região genômica representa um *hotspot* de plasticidade, possivelmente sob influência de pressões seletivas associadas a diferentes hospedeiros ou ambientes (Ramos-González *et al.*, 2020; Morozov e Solovyev, 2020; Gao *et al.*, 2022; Ramos-González *et al.*, 2023). Entre os 27 isolados analisados, foram identificadas seis organizações distintas no extremo 5' do RNA2, variando conforme a combinação dos alelos presentes nas ORFs *p13/p11/ $\Delta p11$* e *p7*. A organização mais frequente corresponde àquela que contém a variante $\Delta p7$ de 141 nucleotídeos e *p13*, de 330 nucleotídeos, observada em 20 isolados, indicando que essa configuração genômica é predominante entre as sequências avaliadas (Figura 13).

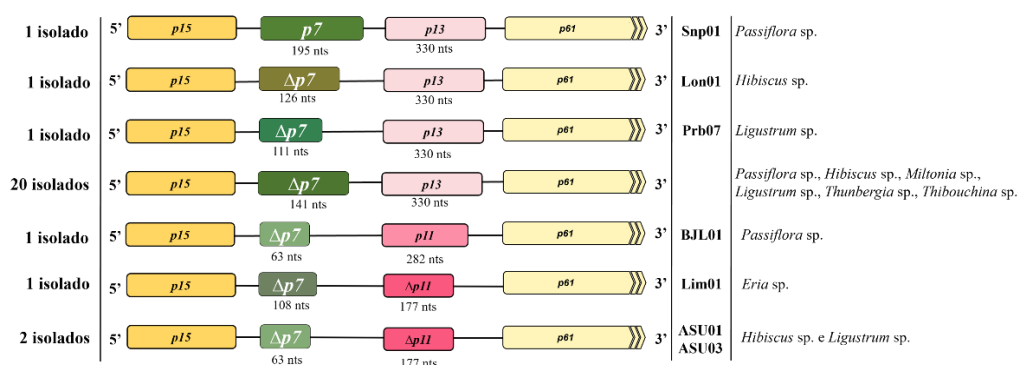


Figura 13. Representação esquemática das seis organizações genômicas identificadas no extremo 5' do RNA2 entre as ORFs *p15* e *p61* nos 27 isolados analisados. As variações refletem diferentes combinações dos alelos *p13*, *p11* e $\Delta p11$, bem como das formas *p7* e $\Delta p7$, resultando em distintos arranjos genéticos.

Os alelos *p13* foram detectados em 21 isolados provenientes de diversas regiões geográficas, como os estados brasileiros de São Paulo, Mato Grosso, Distrito Federal, Paraná, Pará, além da Colômbia. Ainda, esses alelos estão presentes em isolados virais coletados de hospedeiros de diferentes gêneros vegetais, incluindo *Hibiscus*, *Ligustrum*, *Miltonia*, *Nicotiana*, *Tibouchina*, *Thunbergia* e *Passiflora*, indicando sua distribuição geográfica abrangente e a sua capacidade de infectar uma ampla gama de hospedeiros. Por outro lado, o alelo *p11* foi encontrado em apenas um isolado coletado em Bom Jesus da Lapa, na Bahia, sugerindo uma distribuição geográfica mais restrita. Esse fato indica que o alelo *p11* poderia representar uma adaptação local ou mutações únicas que interromperam a leitura da ORF, porém, não impediram a infecção viral. No entanto, maior número de amostras é necessário para a avaliação dessa hipótese.

Por sua vez, o alelo $\Delta p11$ foi identificado em três isolados oriundos de Limeira, São Paulo, Brasil e Assunção, no Paraguai, localidades separadas por uma distância de 1.078 km em linha reta. Esse fato sugere que as mutações que levaram à sua forma truncada pode ter surgido de maneira independente ou que isolados com esse alelo truncado podem ter se espalhado para essas localizações. Essa observação pode indicar a dinâmica evolutiva do PfGSV e sua provável habilidade de gerar variantes funcionais adaptadas a contextos ecológicos específicos. Mutações não sinônimas pontuais, como aquelas que introduzem códons de parada prematuros, podem interromper ORFs existentes e prejudicar a funcionalidade de proteínas resultantes, demonstrando ao mesmo tempo a fragilidade e a dinâmica evolutiva do genoma viral. Por outro lado, mutações não sinônimas em códons de parada podem originar um produto proteico mais extenso, uma vez que o códon de parada se torna um códon de leitura, e a tradução continua até o próximo códon de parada.

A emergência dos alelos *p11* e $\Delta p11$, identificados em um e três isolados, respectivamente, sugere que eventos evolutivos como mutações pontuais e alterações no quadro de leitura desempenham um papel significativo na geração de pequenas proteínas funcionais em genomas compactos, como genomas virais (Guerra-Almeida e Nunes-da-Fonseca, 2020). A alta conservação na região 5' da ORF *p13* entre os alelos de todos os isolados aponta que essa porção do genoma poderia codificar uma estrutura essencial para a funcionalidade das proteínas resultantes. Em contrapartida, as diferenças observadas nas extremidades C-terminais das proteínas codificadas podem refletir adaptações específicas a hospedeiros ou condições ambientais.

O sequenciamento de fragmentos genômicos de novos isolados de PfGSV também possibilitou o aprofundamento acerca da ORF *p7*, localizada a montante dos ORFs *p13/p11/Δp11*. Os alelos da ORF *p7* sugere a existência de outros elementos genéticos funcionalmente relevantes no RNA2 de PfGSV. A alta conservação da região central das proteínas P7, mesmo com tamanhos variados entre os isolados, reforça a hipótese de que esta é uma região funcionalmente crítica (Morozov e Solovyev, 2003; Storz *et al.*, 2014). Os padrões estruturais predominantes, como a presença de domínios TM, indicam que as proteínas P7 podem desempenhar papéis cruciais nas interações do vírus com o hospedeiro, por exemplo com o sistema de membranas celulares (DiMaio, 2014; Cui e Wang, 2016; Atabekova *et al.*, 2022; Chai *et al.*, 2024).

Esse padrão de diversidade e reorganização gênica é compatível com observações feitas em outros vírus do gênero *Cilevirus*, como o CiLV-C e o CiLV-C2, nos quais regiões intergênicas também apresentam ORFs adicionais de origem incerta, provavelmente resultantes de processos de duplicação, recombinação ou seleção positiva. No entanto, o PfGSV se destaca pelo número e variedade de alelos identificados, o que pode indicar uma trajetória evolutiva mais diversificada da que observada pelo cilevirus CiLV-C (Chabi-Jesus *et al.*, 2021). Embora o RNA2 dos cilevírus compartilhe uma organização genômica geral semelhante, existem diferenças marcantes quanto à presença e diversidade dessas pequenas ORFs adicionais. No CiLV-C, por exemplo, destaca-se a *p8*, localizada a jusante da *p15*, cuja função ainda é desconhecida, mas a conservação entre os isolados sugere um possível papel adaptativo; no CiLV-C2, essa posição é ocupada pela *p7*, também de função não elucidada, mas preservada entre diferentes isolados, indicando relevância biológica (Ramos-González *et al.*, 2021). Já no PfGSV, além da manutenção da *p15*, a região equivalente abriga um conjunto mais amplo de ORFs emergentes, incluindo *p13*, *p11*, *Δp11* e *p7*, que variam em presença, sequência e contexto de iniciação, configurando um cenário de maior complexidade e plasticidade evolutiva. A presença de sinais conservados de iniciação da tradução em algumas dessas ORFs reforça a hipótese de que possam ser funcionais e expressas durante a infecção viral. O padrão de diversificação mais acentuado do PfGSV poderia estar associado à geração de novas ORFs e consequentemente à adaptação a diferentes pressões seletivas.

A emergência de pequenas ORFs funcionais em vírus de plantas tem sido cada vez mais documentada. A presença de múltiplas ORFs anteriormente não identificadas no genoma do tomato yellow leaf curl Thailand virus (TYLCTHV, *Begomovirus*

solanumflavusthailandense), cuja expressão foi mapeada por perfis de iniciação de tradução, foi associada diretamente à virulência viral (Chiu *et al.*, 2022). Fatores como o contexto de Kozak e o mecanismo de *leaky scanning* são determinantes para a expressão dessas ORFs, o que poderia ser semelhante ao observado nos genomas dos isolados de PfGSV. Assim, é possível propor que essas ORFs emergentes de PfGSV possam desempenhar papéis importantes na infecção viral, ainda que seus produtos proteicos não tenham sido caracterizados experimentalmente.

No genoma de sonchus yellow net virus (SYNV, *Betanucleorhabdovirus retesonchi*) três pequenas ORFs conservadas foram identificadas e foi demonstrado que são efetivamente traduzidas e possuem funções importantes para a virulência (Gao *et al.*, 2025). A deleção dessas ORFs reduziu a acumulação viral e os sintomas de infecção nas plantas hospedeiras. Tais achados reforçam que pequenas ORFs não devem ser ignoradas em análises genômicas de vírus vegetais uma vez que podem representar estratégias de compactação e inovação funcional. No contexto do PfGSV, a observação de múltiplos alelos, alguns com elementos estruturais conservados, sugere que as pequenas proteínas a serem produzidas podem desempenhar papéis importantes para a infecção viral.

A caracterização de sequências codificantes em genomas virais requer uma avaliação criteriosa dos mecanismos de iniciação de tradução, particularmente no que diz respeito a identificação de códons de início funcionais. O contexto nucleotídeo flanqueador do códon de iniciação, conhecido como sequência Kozak, representa um elemento regulador fundamental na iniciação de tradução eucariótica (Kozak, 1987). Estudos demonstraram que os nucleotídeos nas posições -3 (relativo à adenina do AUG, considerada como +1) e +4 exercem influência determinante na eficiência do recrutamento ribossômico (Kozak, 1986). Em particular, a presença de uma base purina (preferencialmente adenina) na posição -3 e de uma guanina na posição +4 é considerado o arranjo ótimo para maximizar a eficiência da iniciação traducional (Kozak, 1986; 1987).

A análise dos contextos Kozak das ORFs do genoma de PfGSV revela um provável padrão estratégico de regulação traducional, possivelmente associado à evasão da resposta imune do hospedeiro e controle da replicação (Kozak, 1986; 1987; Neil *et al.*, 2008; Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012; DiMaio e Petti, 2013; DiMaio, 2014; Morozov e Solovyev, 2020; Park e Hahn, 2021; Ji *et al.*, 2022; Pavessi e Romerio, 2022; Gao *et al.*, 2022; Brynes e Williams, 2024;). Enquanto as ORFs essenciais, como

RdRp e *p29*, apresentam sequências de Kozak canônicas (A-3/G+4), garantindo alta eficiência na tradução, outras ORFs, como *p15*, *p7* e *p13*, exibem contextos moderados ou subótimos, sugerindo um controle fino da expressão gênica durante a infecção viral. Além disso, a presença de sequências de Kozak fortemente conservadas em genes essenciais reforça a necessidade da síntese proteica dessas ORFs durante a infecção, enquanto ORFs com eficiência intermediária poderiam refletir adaptações para controlar a tradução dessas ORFs (Kozak, 1987; 1989; Carvunis *et al.*, 2012; Morozov e Solovyev, 2020). Assim, a variabilidade nos contextos de Kozak sugere uma estratégia viral, na qual a regulação traducional atua como um mecanismo que permitiria ao vírus modular sua expressão gênica de acordo com as demandas do ciclo infeccioso (Kozak, 1987; 1989; Neil *et al.*, 2008; Carvunis *et al.*, 2012; DiMaio, 2014; Guerra-Almeida e Nunes-da-Fonseca, 2020; Gao *et al.*, 2022). Além disso, a diferença na conservação dos contextos de iniciação da tradução pode indicar uma hierarquia funcional, na qual a tradução de certas ORFs poderia acontecer sob condições específicas ou em momentos diferentes da infecção (Tautz e Domazet-Lošo, 2011; Morozov e Solovyev, 2020). Essa organização também é coerente com a lógica de economia genômica dos vírus, permitindo modular a expressão gênica sem aumentar significativamente o tamanho do genoma (Storz *et al.*, 2014; DiMaio, 2014).

O presente estudo também apresenta evidências que ressaltam a plasticidade genômica de PfGSV. As mutações detectadas, como adição ou deleção de nucleotídeos que alteram os quadros de leitura, são fenômenos chave na geração de diversidade genômica viral. Esses eventos permitem aos vírus “explorar” novos contextos ecológicos, proporcionando vantagens adaptativas em resposta a pressões seletivas impostas por diferentes ambientes ou hospedeiros vegetais (Tautz e Domazet-Lošo, 2011; Carvunis *et al.*, 2012; Guerra-Almeida e Nunes-da-Fonseca, 2020; Park e Hahn, 2021; Pavessi e Romerio, 2022; Ji *et al.*, 2023). Essa possibilidade de adaptação, em conjunto com a sabida habilidade dos cilevírus, e em particular de PfGSV, de infectar uma ampla gama de hospedeiros, reflete a relevância biológica e epidemiológica desses patógenos emergentes (Kitajima *et al.*, 2003a; 2010; da Costa-Rodrigues *et al.*, 2022; Pires, 2022; Roy *et al.*, 2022).

Apesar da relevância dos achados, este estudo apresenta limitações importantes. A ausência de validação experimental da expressão das proteínas codificadas pelas ORFs emergentes impede uma conclusão definitiva sobre sua funcionalidade. Futuras

abordagens serão fundamentais para verificar o papel dessas proteínas no ciclo de vida do PfGSV. Além disso, investigações estruturais e de interação molecular com proteínas do hospedeiro ou vetor poderiam revelar funções específicas associadas à virulência, movimentação viral ou supressão de vias de defesa. Portanto, os dados aqui apresentados, somados às evidências da literatura recente, sustentam a hipótese de que a diversidade observada no RNA2 do PfGSV, especialmente nas ORFs emergentes, não é aleatória. Ao contrário, pode representar um mecanismo adaptativo e funcionalmente ativo, relevante para a evolução e a adaptação deste vírus em diferentes contextos ecológicos.

7 CONSIDERAÇÕES FINAIS

O presente estudo apresentou uma análise detalhada sobre o extremo 5' do RNA2 do passion fruit green spot virus, destacando avanços na compreensão de sua estrutura genômica, diversidade molecular e sugerindo um papel na interação com hospedeiros. O Brasil, como maior produtor mundial de maracujá, enfrenta desafios significativos na manutenção da produtividade devido à ação de patógenos, entre os quais o PfGSV se destaca pela capacidade de afetar múltiplos gêneros botânicos, que podem atuar como reservatórios do vírus, e pela complexidade de sua biologia.

Os resultados demonstram que o genoma do PfGSV, particularmente o extremo 5' do RNA2, é altamente variável, como evidenciado pela presença de ORFans no RNA2, incluindo os alelos *p13*, *p11*, $\Delta p11$ e a ORF *p7*, que apresentam variações significativas entre os isolados analisados. As diferenças entre os alelos refletem processos evolutivos, como mutações pontuais, mudanças de quadro de leitura e deleções, que reforçam a plasticidade genômica característica de vírus com genomas de RNA. A alta conservação de determinadas regiões genômicas, especialmente nos trechos correspondentes às proteínas P13 e P11, sugere funções cruciais na patogenicidade viral e na interação vírus-hospedeiro. Além disso, os novos alelos identificados, como $\Delta p11$ e as variantes de *p7*, ampliam o repertório funcional conhecido do PfGSV, indicando possíveis papéis na adaptação a diferentes hospedeiros e ambientes.

As análises bioinformáticas revelaram semelhanças remotas entre as proteínas estudadas e domínios conhecidos de regulação transcricional, estabilidade de RNA e interação proteína-proteína, levantando hipóteses sobre seus potenciais papéis no ciclo de multiplicação viral. A presença de sítios conservados de glicosilação em proteínas como P13, P11 e $\Delta P11$ sugere que modificações pós-traducionais poderiam ser críticas para a funcionalidade dessas moléculas, influenciando sua estabilidade, localização subcelular e interação com outras proteínas ou ácidos nucleicos.

A identificação de seis alelos para a ORF *p7* reforça a diversidade funcional dentro da região genômica previamente classificada como intergênica. A variabilidade no comprimento e na composição das proteínas codificadas por esses alelos aponta para um potencial processo de perda ou aquisição de função ao longo da evolução do PfGSV, possivelmente modulada por pressões seletivas impostas pelos diferentes hospedeiros ou ecossistemas.

Os resultados obtidos neste trabalho fornecem uma base sólida para futuros estudos sobre a biologia molecular do PfGSV, com implicações na compreensão da interação vírus-vetor-hospedeiro e no desenvolvimento de estratégias de manejo de doenças associadas. A caracterização de novas ORFans e a análise comparativa entre isolados de diferentes origens geográficas e hospedeiros são avanços significativos para a elucidação dos mecanismos de patogenicidade e adaptação viral. Esses resultados reforçam a necessidade de estudos contínuos para identificar possíveis alvos moleculares que contribuam para o controle efetivo do PfGSV, prevenindo seu ressurgimento como um obstáculo à produção de maracujá.

8 CONCLUSÕES

O presente trabalho permitiu avançar na compreensão da variabilidade genômica do extremo 5' do RNA2 do Cilevirus passion fruit green spot virus, destacando a importância dessa região como reservatório para o surgimento de pequenas ORFs emergentes. Os resultados obtidos revelaram a predominância do alelo *p13* entre os isolados analisados, bem como a ocorrência de alelos menos frequentes, como *p11* e $\Delta p11$ e múltiplas variantes da ORF *p7* entre os isolados, reforçando a plasticidade do genoma viral. A detecção de vários truncamentos, mutações em códons de início e/ou parada, além de contextos Kozak variáveis, sugere que a tradução dessas pequenas proteínas poderia ser modulada de forma adaptativa, contribuindo para a manutenção do ciclo infeccioso e para a interação com diferentes hospedeiros e vetores.

As análises bioinformáticas forneceram indícios de potenciais papéis funcionais das proteínas codificadas por essas ORFs, como a presença de sítios conservados glicosilação e similaridades remotas com domínios nucleases. Tais achados apontam para funções que poderiam envolver a estabilidade das partículas virais, a modulação de respostas do hospedeiro e até mesmo a eficiência da transmissão pelo ácaro *Brevipalpus*. Assim, a caracterização dessa região contribui para consolidar o entendimento de que pequenas proteínas acessórias poderiam desempenhar papéis importantes na adaptação e evolução viral.

Apesar dos avanços obtidos, o estudo apresenta limitações, uma que vez que não foi possível confirmar experimentalmente a expressão e a funcionalidade das ORFs emergentes. Nesse sentido, futuras pesquisas devem incluir análises funcionais e estruturais mais aprofundadas, de modo a elucidar o papel dessas proteínas na patogenicidade e na transmissão do PfGSV. Além disso, a ampliação dos estudos comparativos entre diferentes isolados e hospedeiros poderá revelar padrões evolutivos mais amplos.

Em síntese, este trabalho contribui de forma significativa para o conhecimento da diversidade de PfGSV, oferecendo novas perspectivas para a compreensão de seus mecanismos adaptativos e evolutivos e, para o desenvolvimento de abordagens de controle da virose na cultura do maracujazeiro, além de plantas ornamentais.

9 REFERÊNCIAS

ABRAHAMCZYK, S.; SOUTO-VILARÓS, D.; RENNER, S. S. Escape from extreme specialization: Passionflowers, bats and the sword-billed hummingbird. *Proceedings of the Royal Society B: Biological Sciences*, v. 281, n. 1795, 2014. DOI: <https://doi.org/10.1098/rspb.2014.0888>.

ALTSCHUL, S. F. et al. Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, n. 3, p. 403-410, 1990. DOI: 10.1016/S0022-2836(05)80360-2

ANTONIOLI-LUIZON, R. Sequenciamento parcial do PfGSV e desenvolvimento de métodos para diagnóstico. 2010. Dissertação (Mestrado) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba. 2010.

ARMENTEROS, J. J. A. et al. Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, v. 2, n. 5, 2019. DOI: <https://doi.org/10.26508/LSA.201900429>.

ATABEKOVA, A. K. et al. Interaction between movement proteins of Hibiscus green spot virus. *Viruses*, v. 14, n. 12, 2022. DOI: <https://doi.org/10.3390/v14122742>.

BERTANI G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol.* 1951 Sep;62(3):293-300. doi: 10.1128/jb.62.3.293-300.1951

BIRNBOIM, H. C.; DOLY, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, v. 7, 1979. DOI: 10.1093/nar/7.6.1513

BRYNES, A.; WILLIAMS, J. V. Small hydrophobic (SH) proteins of Pneumoviridae and Paramyxoviridae: small but mighty. *Journal of Virology*, v. 98, n. 9, 2024. DOI: <https://doi.org/10.1128/jvi.00809-24>.

CARVUNIS, A. R. et al. Proto-genes and de novo gene birth. *Nature*, v. 487, n. 7407, p. 370-374, 2012. DOI: <https://doi.org/10.1038/nature11184>.

CASTRO, E. B. et al. A newly available database of an important family of phytophagous mites: Tenuipalpidae Database. *Zootaxa*, v. 4868, n. 4, p. 577-583, 2020. DOI: <https://doi.org/10.11646/zootaxa.4868.4.7>.

CHAI, M. et al. The 6-kilodalton peptide 1 in plant viruses of the family Potyviridae is a viroporin. *Proceedings of the National Academy of Sciences of the United States of America*, v. 121, n. 21, 2024. DOI: <https://doi.org/10.1073/pnas.2401748121>.

CHILDERS, C. C.; RODRIGUES, J. C. V. An overview of Brevipalpus mites (Acari: Tenuipalpidae) and the plant viruses they transmit. *Zoosymposia*, v. 6, n. 1, p. 180-192, 2011. DOI: <https://doi.org/10.11646/zoosymposia.6.1.28>.

CHIU, C. W. et al. Translation initiation landscape profiling reveals hidden open-reading frames required for the pathogenesis of tomato yellow leaf curl Thailand virus. *Plant Cell*, v. 34, n. 5, p. 1804-1821, 2022. DOI: <https://doi.org/10.1093/plcell/koac019>.

CROOKS, G. E. et al. WebLogo: A sequence logo generator. *Genome Research*, v. 14, n. 6, p. 1188-1190, 2004. DOI: <https://doi.org/10.1101/gr.849004>.

CUI, H.; WANG, A. Plum Pox Virus 6K1 protein is required for viral replication and targets the viral replication complex at the early stage of infection. *Journal of Virology*, v. 90, n. 10, 2016. DOI: <https://doi.org/10.1128/jvi.00024-16>.

DA COSTA-RODRIGUES, M. et al. First report of Passion fruit green spot virus infecting hibiscus plants. *New Disease Reports*, v. 45, n. 2, 2022. DOI: <https://doi.org/10.1002/ndr2.12080>.

DE JESUS, O. N.; FALEIRO, F. G. Maracujá: o produtor pergunta, a Embrapa responde. Brasília: Embrapa, 2016.

DIETZGEN, R. G. et al. Dichorhavirus in their host plants and mite vectors. *Advances in Virus Research*, v. 102, p. 119-148, 2018. DOI: <https://doi.org/10.1016/BS.AIVIR.2018.06.001>.

DIMAIO, D. Viral miniproteins. *Annual Review of Microbiology*, v. 68, p. 21-43, 2014. DOI: <https://doi.org/10.1146/annurev-micro-091313-103727>.

DIMAIO, D.; PETTI, L. The E5 proteins. *Virology*, v. 28, n. 1, 2013. DOI: <https://doi.org/10.1016/j.virol.2013.05.006>.

DOERKS, T. et al. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Research*, v. 12, n. 1, p. 47-56, 2002. DOI: <https://doi.org/10.1101/gr.203201>.

FALEIRO, F. G. Maracujá, fruta nativa do Brasil para o mundo. Campo & Negócios, Anuário HF, 2022.

FALEIRO, F. G. et al. Maracujá. San José: Instituto Interamericano de Cooperación para la Agricultura (IICA), 2017.

FREITAS-ASTÚA, J. et al. Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? Current Opinion in Virology, v. 33, p. 66-73, 2018. DOI: <https://doi.org/10.1016/j.coviro.2018.07.010>.

GABLER, F. et al. Protein sequence analysis using the MPI bioinformatics toolkit. Current Protocols in Bioinformatics, v. 72, n. 1, 2020. DOI: <https://doi.org/10.1002/cpbi.108>.

GAO, D. M. et al. A rhabdovirus accessory protein inhibits jasmonic acid signaling in plants to attract insect vectors. Plant Physiology, v. 190, n. 2, p. 1349-1364, 2022. DOI: <https://doi.org/10.1093/plphys/kiac319>.

GAO, M. et al. Functional characterization of conserved small open reading frames on the genomic RNA of sonchus yellow net virus. Phytopathology Research, v. 7, n. 1, 2025. DOI: <https://doi.org/10.1186/s42483-025-00333-5>.

GUERRA-ALMEIDA, D.; NUNES-DA-FONSECA, R. Small open reading frames: how important are they for molecular evolution? Frontiers in Genetics, v. 11, 2020. DOI: <https://doi.org/10.3389/fgene.2020.574737>.

GUPTA, R.; BRUNAK, S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing, v. 7, p. 310-322, 2002. DOI: 10.1142/9789812799623_0029

HALLGREN, J. et al. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv, 2022. DOI: <https://doi.org/10.1101/2022.04.08.487609>.

HANAHAH, D. Studies on transformation of Escherichia coli with plasmids. Journal of Molecular Biology, v. 166, 1983. DOI: 10.1016/s0022-2836(83)80284-8.

HELLENS, R. P. et al. The emerging world of small ORFs. Trends in Plant Science, v. 21, n. 4, p. 317-328, 2016. DOI: <https://doi.org/10.1016/j.tplants.2015.11.005>.

IBGE. Produção agropecuária – Maracujá. Disponível em: <https://www.ibge.gov.br/explica/producao-agropecuaria/maracuja/br>. Acesso em: 15 de outubro de 2025.

IBGE. Pesquisa: Produção Agrícola – Lavoura Permanente – Brasil. Disponível em: <https://cidades.ibge.gov.br/brasil/pesquisa/15/0>. Acesso em: 15 de outubro de 2025.

Ji, Y. F. et al. Frameshift mutations in the microevolution and macroevolution of viruses. *bioRxiv*, 2022. DOI: <https://doi.org/10.1101/2022.07.20.500745>.

JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, v. 596, n. 7873, p. 583-589, 2021. DOI: <https://doi.org/10.1038/s41586-021-03819-2>.

KITAJIMA, E. W.; REZENDE, J. A. M.; RODRIGUES, J. C. V. Passion fruit green spot virus vectored by *Brevipalpus phoenicis* (Acari: Tenuipalpidae) on passion fruit in Brazil. *Experimental and Applied Acarology*, v. 30, n. 1-3, p. 225-231, 2003a. DOI: <https://doi.org/10.1023/B:APPA.0000006551.74604.84>.

KITAJIMA, E. W.; CHAGAS, C. M.; RODRIGUES, J. C. V. *Brevipalpus*-transmitted plant virus and virus-like diseases: cytopathology and some recent cases. *Experimental and Applied Acarology*, v. 30, n. 1-3, p. 135-160, 2003b. DOI: <https://doi.org/10.1023/B:APPA.0000006546.55305.e3>.

KITAJIMA, E. W.; CARLOS, J.; RODRIGUES, V.; FREITAS-ASTÚA, J. An annotated list of ornamentals naturally found infected by *Brevipalpus* mite-transmitted viruses. *Scientia Agricola*, n. 3, 2010. DOI: <https://doi.org/10.1590/S0103-90162010000300014>

KITAJIMA, E. W.; TASSI, A. D.; SAITO, M. S. G.; SALAROLI, R. B.; CHABI-JESUS, C.; FREITAS-ASTUA, J. Novas espécies hospedeiras do vírus da leprose do ligustro (*Ligustrum leprosis virus* – LigLV). 40º Congresso Paulista de Fitopatologia, Instituto Agronômico, Campinas, São Paulo, 2017.

KOZAK, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, v. 44, 1986. DOI: [10.1016/0092-8674\(86\)90762-2](https://doi.org/10.1016/0092-8674(86)90762-2)

KOZAK, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*, v. 15, 1987. DOI: <https://academic.oup.com/nar/article/15/20/8125/2377809>.

KROGH, A. et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, v. 305, n. 3, p. 567-580, 2001. DOI: <https://doi.org/10.1006/jmbi.2000.4315>.

LOCALI-FABRIS, E. C. et al. Complete nucleotide sequence, genomic organization and phylogenetic analysis of Citrus leprosis virus cytoplasmic type. *Journal of General Virology*, v. 87, n. 9, p. 2721-2729, 2006. DOI: <https://doi.org/10.1099/vir.0.82038-0>.

LU, G.; MORIYAMA, E. N. Vector NTI, a balanced all-in-one sequence analysis suite. *Briefings in Bioinformatics*, v. 5, n. 4, p. 378-388, 2004. DOI: <https://doi.org/10.1093/bib/5.4.378>

MACHADO, C. de F. et al. Guia de identificação e controle de pragas na cultura do maracujazeiro. 2017. Disponível em: www.embrapa.br/fale-conosco/sac.

MCNATT, M. W. et al. Species-specific activity of HIV-1 Vpu and positive selection of tetherin transmembrane domain variants. *PLoS Pathogens*, v. 5, n. 2, 2009. DOI: <https://doi.org/10.1371/journal.ppat.1000300>.

MELZER, M. J. et al. Characterization of a virus infecting Citrus volkameriana with citrus leprosis-like symptoms. *Phytopathology*, v. 102, n. 1, p. 122-127, 2012. DOI: <https://doi.org/10.1094/PHYTO-01-11-0013>.

MIRDITA, M. et al. ColabFold: making protein folding accessible to all. *Nature Methods*, v. 19, n. 6, p. 679–682, 2022. DOI: <https://doi.org/10.1038/s41592-022-01488-1>.

MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (Brasil). Hierarquização de pragas de maior risco fitossanitário do Brasil. Secretaria da Defesa Agropecuária, Departamento de Sanidade Vegetal e Insumos Agrícola. Brasília: MAPA, 2022. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/noticias/mapa-divulga-lista-com-hierarquizacao-de-pragas-de-maior-risco-fitossanitario>. Acesso em: 15 de outubro de 2025.

MORAES, F. H. R.; BELO, W. R. F.; MORAES, G. J.; KITAJIMA, E. W. Ocorrência do Vírus da Pinta Verde em maracujá no Estado do Maranhão, Brasil. *Fitopatologia Brasileira*, v. 31, n. 1, 2006. DOI: <https://doi.org/10.1590/S0100-41582006000100019>

MOROZOV, S. Y.; SOLOVYEV, A. G. Triple gene block: modular design of a multifunctional machine for plant virus movement. *Journal of General Virology*, v. 84, n. 6, p. 1351–1366, 2003. DOI: <https://doi.org/10.1099/vir.0.18922-0>.

MOROZOV, S. Y.; SOLOVYEV, A. G. Small hydrophobic viral proteins involved in intercellular movement of diverse plant virus genomes. *AIMS Microbiology*, v. 6, n. 3, p. 305–329, 2020. DOI: <https://doi.org/10.3934/microbiol.2020019>.

MUNGUTI F, NYABOGA EN, HOLTON T, KREUZE J, MAINA S. Genomics insight on passion fruit viral disease complexity. *Microbiology Spectrum*. 2025. DOI: [10.1128/spectrum.00344-25](https://doi.org/10.1128/spectrum.00344-25)

NEIL, S. J. D.; ZANG, T.; BIENIASZ, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, v. 451, n. 7177, p. 425–430, 2008. DOI: <https://doi.org/10.1038/nature06553>.

ØDUM, M. T. et al. DeepLoc 2.1: multi-label membrane protein type prediction using protein language models. *Nucleic Acids Research*, v. 52, W1, p. W215–W220, 2024. DOI: <https://doi.org/10.1093/NAR/GKAE237>.

OLMEDO-VELARDE, A.; HU, J.; MELZER, M. J. A virus infecting *Hibiscus rosa-sinensis* represents an evolutionary link between Cileviruses and Higreviruses. *Frontiers in Microbiology*, v. 12, 2021. DOI: <https://doi.org/10.3389/fmicb.2021.660237>.

OLMEDO-VELARDE, A. et al. First report of the Hibiscus strain of Citrus Leprosis Virus C2 infecting passionfruit (*Passiflora edulis*). *Plant Disease*, v. 106, n. 9, p. 2539, 2022. DOI: <https://doi.org/10.1094/PDIS-10-21-2314-PDN>.

PARK, D.; HAHN, Y. Rapid protein sequence evolution via compensatory frameshift is widespread in RNA virus genomes. *BMC Bioinformatics*, v. 22, n. 1, p. 1–16, 2021. DOI: <https://doi.org/10.1186/S12859-021-04182-9/FIGURES/3>.

PAVESI, A.; ROMERIO, F. Extending the coding potential of viral genomes with overlapping antisense ORFs: a case for the de novo creation of the gene encoding the

antisense protein ASP of HIV-1. *Viruses*, v. 14, n. 1, 2022. DOI: <https://doi.org/10.3390/v14010146>.

PETERSEN, T. N.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, v. 8, n. 10, p. 785–786, 2011. DOI: <https://doi.org/10.1038/nmeth.1701>.

PINTO, L. H.; HOLSINGER, L. J.; LAMB, R. A. Influenza virus M2 protein has ion channel activity. *Cell*, v. 69, 1992. DOI: [10.1016/0092-8674\(92\)90452-i](https://doi.org/10.1016/0092-8674(92)90452-i)

PINTO, L. H.; LAMB, R. A. The M2 proton channels of influenza A and B viruses. *Journal of Biological Chemistry*, v. 281, n. 14, p. 8997–9000, 2005. DOI: <https://doi.org/10.1074/jbc.R500020200>.

PIRES, G. R. R. Gama de hospedeiros naturais e experimentais de Passion fruit green spot virus e interações com o vetor *Brevipalpus yothersi*. 2022. Dissertação (Mestrado) – Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba. 2022.

RAI, S.; NAGAR, J. C.; MUKIM, M. Pharmacological and medicinal importance of *Passiflora edulis*: a review. *International Journal of Research and Review*, v. 9, n. 4, p. 341–349, 2022. DOI: <https://doi.org/10.52403/ijrr.20220442>.

RAMOS-GONZÁLEZ, P. L. et al. Citrus leprosis: a unique multi-etiological disease. *Citrus in the Americas*, v. 1, n. 1, p. 4–19, 2018. DOI: doi.org/10.6084/m9.figshare.13513857.v1

RAMOS-GONZÁLEZ, P. L. et al. Passion Fruit Green Spot Virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-end of the RNA2 segment across Cileviruses. *Frontiers in Microbiology*, 2020. DOI: <https://doi.org/10.3389/fmicb.2020.00206>.

RAMOS-GONZÁLEZ, P. L. et al. Poorly conserved P15 proteins of Cileviruses retain elements of common ancestry and putative functionality: a theoretical assessment on the evolution of Cilevirus genomes. *Frontiers in Plant Science*, v. 12, 2021. DOI: <https://doi.org/10.3389/fpls.2021.771983>.

RAMOS-GONZÁLEZ, P. L. et al. A novel lineage of Cile-like viruses discloses the phylogenetic continuum across the family Kitaviridae. *Frontiers in Microbiology*, v. 13, 2022. DOI: <https://doi.org/10.3389/fmicb.2022.836076>.

RAMOS-GONZÁLEZ, P. L. et al. Kitaviruses: a window to atypical plant viruses causing nonsystemic diseases. *Annual Review of Phytopathology*, 2023. DOI: <https://doi.org/10.1146/annurev-phyto-021622>.

ROY, A. et al. A Novel Virus of the Genus Cilevirus causing symptoms similar to Citrus Leprosis. *Virology*, v. 103, n. 5. 2013. DOI: doi.org/10.1094/PHYTO-07-12-0177-R

ROY, A. et al. First report of Passion Fruit Green Spot Virus in yellow passion fruit (*Passiflora edulis* f. *flavicarpa*) in Casanare, Colombia. *Plant Disease*, v. 107, n. 7, p. 2270, 2022. DOI: [10.1094/PDIS-09-22-2267-PDN](https://doi.org/10.1094/PDIS-09-22-2267-PDN).

ROY, A. et al. Meta-Transcriptomic Analysis Uncovers the Presence of Four Novel Viruses and Multiple Known Virus Genera in a Single Hibiscus rosa-sinensis Plant in Colombia. *Viruses* 2024, 16, 267. DOI: [10.3390/v16020267](https://doi.org/10.3390/v16020267).

SAITO, M. S. G. O PfGSV isolado de ligustro: suas relações com o ácaro vetor *Brevipalpus*. 2020. Dissertação (Mestrado) - Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba. 2020.

SANTOS FILHO, H. P. et al. Definhamento precoce do maracujazeiro. Embrapa – Comunicado Técnico, n. 60, p. 1–5, 1999.

SANTOS-FILHO, H. P.; LARANJEIRA, F. F.; HADDAD, F. Maracujá: do cultivo à comercialização. In: *Maracujá: do cultivo à comercialização*, 1. ed. 2017. Disponível em: www.embrapa.br/fale-conosco/sac.

DE MENDONÇA, R. S.; FERRAGUT, F.; DE OLIVEIRA, I. C. S.; TASSI, A. D.; FILENI, F.; OCHOA, R.; NAVIA, D. Revisiting the systematics of *Brevipalpus* flat mites (Tenuipalpidae): phylogeny, species groups and cryptic diversity. *Peer Community Journal*, Volume 5 (2025), article no. e92. DOI: doi.org/10.24072/pcjournal.615

SCHMITZ, A.; RIESNER, D. Purification of nucleic acids by selective precipitation with polyethylene glycol 6000. *Analytical Biochemistry*, v. 354, n. 2, p. 311–313, 2006. DOI: <https://doi.org/10.1016/j.ab.2006.03.014>.

SCHNEIDER, T. D.; STEPHENS, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research*, v. 18, n. 20, p. 6097, 1990. doi: 10.1093/nar/18.20.6097

SEREEWIT, J. et al. ORF-interrupting mutations in Monkeypox virus genomes from Washington and Ohio, 2022. *Viruses*, v. 14, n. 11, 2022. DOI: <https://doi.org/10.3390/v14112393>.

SINGH, U.; SYRKIN WURTELE, E. How new genes are born. *ELife*, v. 9, 2020. DOI: <https://doi.org/10.7554/eLife.55136>.

SONNHAMMER, E. L. L.; VON HEIJNE, G.; KROGH, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *American Association for Artificial Intelligence*, v. 1, p. 175–182, 1998. Disponível em: www.aaai.org.

STEENTOFT, C. et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO Journal*, v. 32, n. 10, p. 1478–1488, 2013. DOI: <https://doi.org/10.1038/EMBOJ.2013.79>.

STORZ, G.; WOLF, Y. I.; RAMAMURTHI, K. S. Small proteins can no longer be ignored. *Annual Review of Biochemistry*, v. 83, p. 753–777, 2014. DOI: <https://doi.org/10.1146/annurev-biochem-070611-102400>.

STREBEL, K. HIV-1 Vpu - An ion channel in search of a job. *Biochimica et Biophysica Acta - Biomembranes*, v. 1838, n. 4, p. 1074–1081, 2014. DOI: <https://doi.org/10.1016/j.bbamem.2013.06.029>.

SUGRUE, R. J.; HAY', A. J. Structural Characteristics of the M2 Protein of Influenza A Viruses: Evidence That It Forms a Tetrameric Channel. *Virology*, v. 180, p. 617–624, 1991. DOI: 10.1016/0042-6822(91)90075-m.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, v. 38, n. 7, p. 3022–3027, 2021. DOI: <https://doi.org/10.1093/molbev/msab120>.

TAUTZ, D.; DOMAZET-LOŠO, T. The evolutionary origin of orphan genes. *Nature Reviews Genetics*, v. 12, n. 10, p. 692–702, 2011. DOI: <https://doi.org/10.1038/nrg3053>.

VAKIRLIS, N.; CARVUNIS, A. R.; MCLYSAGHT, A. Synteny-based analyses indicate that sequence divergence is not the main source of orphan genes. *ELife*, v. 9, 2020. DOI: <https://doi.org/10.7554/eLife.53500>.

VAN OSS, S. B.; CARVUNIS, A. R. De novo gene birth. *PLoS Genetics*, v. 15, n. 5, 2019. DOI: <https://doi.org/10.1371/journal.pgen.1008160>.

VARADI, M. et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Research*, v. 52, D1, p. D368–D375, 2024. DOI: <https://doi.org/10.1093/NAR/GKAD1011>.

VIANA, F. M. P.; FREIRE, F. das C. O.; CARDOSO, J. E.; VIDAL, J. C. Principais doenças do Maracujazeiro na região Nordeste e seu controle. 2003.

VOLCIC, M.; WIESMÜLLER, L.; KIRCHHOFF, F. Small but Highly Versatile: The Viral Accessory Protein Vpu. *Annual Review of Virology*, v. 10, n. 1, p. 243–259, 2023. DOI: doi.org/10.1146/ANNUREV-VIROLOGY-111821-100816/CITE/REFWORKS.

WALKER, P. J.; DIETZGEN, R. G.; JOUBERT, D. A.; BLASDELL, K. R. Rhabdovirus accessory genes. *Virus Research*, v. 162, n. 1–2, p. 110–125, 2011. DOI: <https://doi.org/10.1016/J.VIRUSRES.2011.09.004>.

WEEKS, A. R.; MAREC, F.; BREEUWER, J. A. J. A mite species that consists entirely of haploid females. *Science*, v. 292, n. 5526, p. 2479–2482, 2001. DOI: <https://doi.org/10.1126/science.1060411>.

Xia, S., Chen, J., Arsala, D. et al. Functional innovation through new genes as a general evolutionary process. *Nature Genetics* 57, 295–309 (2025). <https://doi.org/10.1038/s41588-024-02059-0>.

YANG, X.; LI, Y.; WANG, A. Research Advances in Potyviruses from the lab bench to the field. *Annual Review of Phytopathology*, 2021. DOI: <https://doi.org/10.1146/annurev-phyto-020620-114550>.

ZANDI, M.; SHAFATI, M.; KALANTAR-NEYESTANAKI, D.; POURGHADAMYARI, H.; FANI, M.; SOLTANI, S.; KALEJI, H.; ABBASI, S. The role of SARS-CoV-2 accessory proteins in immune evasion. *Biomedicine &*

Pharmacotherapy, v. 156, 113889, 2022. DOI:
<https://doi.org/10.1016/J.BIOPHA.2022.113889>.

ZIMMERMANN, L.; STEPHENS, A.; NAM, S. Z.; RAU, D.; KÜBLER, J.;
LOZAJIC, M.; GABLER, F.; SÖDING, J.; LUPAS, A. N.; ALVA, V. A Completely
Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core.
Journal of Molecular Biology, v. 430, n. 15, p. 2237–2243, 2018. DOI:
<https://doi.org/10.1016/J.JMB.2017.12.007>.