



**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA CELULAR
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR
LABORATÓRIO DE VIROLOGIA VEGETAL**

**INTERAÇÕES TOSPOVÍRUS-
PLANTA: Caracterização estrutural de proteínas de
tospovírus e identificação de interações entre proteínas
virais e celulares**

Rayane Nunes Lima

**Tese de Doutorado apresentada ao Programa de Pós-
Graduação em Biologia Molecular da Universidade de
Brasília, para obtenção do título de Doutor em Biologia
Molecular.**

Orientador: Dr. Renato de Oliveira Resende

Brasília, 2018.



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA CELULAR
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR
LABORATÓRIO DE VIROLOGIA VEGETAL

**INTERAÇÕES TOSPOVÍRUS-PLANTA: Caracterização
estrutural de proteínas de tospovírus e identificação de
interações entre proteínas virais e celulares**

Capítulo I:
Introdução Geral

Capítulo II:
**Homology modeling and molecular dynamics provide structural
insights into tospovirus nucleoprotein**

Capítulo III:
**The complete genome of the tospovirus *Zucchini lethal chlorosis virus*
and the structural analysis of its nucleocapsid protein (NP)**

Capítulo IV:
**Predição da estrutura terciária, Expressão e Purificação da proteína
NS_S de GRSV**

Capítulo V:
Interações entre proteína virais e proteínas do hospedeiro

RAYANE NUNES LIMA

**INTERAÇÕES ORTHOTOSPOVÍRUS-
PLANTA: Caracterização estrutural de proteínas de
orthotospovírus e identificação de interações entre proteínas
virais e celulares**

Banca Examinadora:

Prof. Dr. Renato de Oliveira Resende (Orientador)
(CEL-UnB)

Prof. Dr. Tatsuya Nagata
(CEL-UnB)

Dra. Marília Santos Silva
(Embrapa-Recursos Genéticos e Biotecnologia)

Dra. Simone da Graça Ribeiro
(Embrapa-Recursos Genéticos e Biotecnologia)

Prof. Dr. Bergmann Morais Ribeiro (Suplente)
(CEL-UnB)

“We live in a society exquisitely dependent on Science and technology, in which, hardly anyone knows anything about Science and technology. This is a clear prescription for disaster.”

“Vivemos em uma sociedade extremamente dependente da ciência e tecnologia, na qual pouquíssimos sabem alguma coisa sobre ciência e tecnologia. Isto é uma clara prescrição para o desastre.” Carl Edward Sagan, 1990. *Why We Need To Understand Science. The Skeptical Inquirer*, Vol. 14.3.

Índice

Resumo.....	8
Abstract	9
Motivação.....	10
Justificativa.....	Erro! Indicador não definido.
Objetivos	11
Objetivo Geral	11
Objetivos Específicos.....	11
CAPÍTULO I: Introdução Geral	12
1.1 Ordem <i>Bunyavirales</i>	13
1.1.1 Taxonomia e Diversidade	13
1.1.2 Vírions e Genoma	16
1.2 Família <i>Tospoviridae</i>	19
1.2.1 Taxonomia e diversidade	19
1.2.1 Vírions e Genoma	21
1.2.3 <i>Groundnut ringspot orthotospovirus</i> (GRSV)	26
1.3 Ferramentas moleculares para o estudo das interações proteicas	27
1.3.1 Interações proteína-proteína para orthotospovírus	30
1.4 Ferramentas para predição in silico da estrutura tridimensional de proteínas.....	33
1.5 Referências.....	34
CAPÍTULO II: Homology modeling and molecular dynamics provide structural insights into orthotospovirus nucleoprotein	42
Homology modeling and molecular dynamics provide structural insights into orthotospovirus nucleoprotein.....	43
2.1 Abstract	44
2.1.1 Background	44
2.1.2 Results	44
2.1.3 Conclusions	44
2.2 Background.....	46
2.3 Results and Discussion	47
2.3.1 Three-dimensional model of GRSV NP and Oligomerization	47
2.3.2 RNA interaction	52
2.4 Conclusions.....	54
2.5 Methods	54

2.5.1 <i>In silico</i> Homology modeling and Model optimization	54
2.5.2 Molecular dynamics	55
2.6 Additional files:.....	56
Additional file 1: The sequence codes of the viruses used at this work.	56
Additional file 2: Ramachandran plot analysis of predicted structure of <i>Groundnut ringspot orthotospovirus</i> (GRSV) NP protein.....	56
2.7 Considerações.	57
2.8 References.....	58
CAPÍTULO III: The complete genome of the tospovirus <i>Zucchini lethal chlorosis orthotospovirus</i> and the structural analyses of its nucleocapsid protein (NP).	63
3.1 Abstract	65
3.1.1 Background.....	65
3.1.2 Findings	65
3.1.3 Conclusion	65
3.2 Body of text	66
3.2.1 Abbreviations	Erro! Indicador não definido.
3.2.2 Competing interests.....	Erro! Indicador não definido.
3.2.3 Authors' contribution.....	Erro! Indicador não definido.
3.2.4 Acknowledgments.....	Erro! Indicador não definido.
3.3 Análise da estrutura tridimensional da proteína do nucleocapsídeo (NP) do ZLCV	73
3.3.1 Figura Suplementar:.....	76
3.4 Referências.....	76
CAPÍTULO IV: Predição da estrutura terciária, Expressão e Purificação da proteína NS _S de GRSV.	79
4.1 Predição da estrutura da NSs.....	80
4.2 Expressão da proteína NS _S de GRSV.....	86
4.2.1 Metodologia.....	86
4.2.2 Expressão e Purificação.....	87
4.3 Material suplementar.....	91
4.3.1 Figura S1. Gráfico de Ramachandran para a proteína NS _S de GRSV	91
4.3.2 Mapa do vetor de expressão pet28a.	92
4.4 Referências.....	93
Capítulo V: Interações entre proteína virais e proteínas do hospedeiro.....	94
5.1 Busca por proteínas de plantas candidatas para interação com as proteínas N, NS _M e NS _S de GRSV.....	95

5.2 Metodologia	97
5.2.1 Amplificação dos genes	97
5.2.3 Clonagem dos insertos no vetor pGEM® -T easy (Promega).....	98
5.2.4 Digestão dos vetores recombinantes pGEM + insertos	98
5.2.5 Clonagem nos vetores pGADT7 AD e pGBKT7 para ensaio de Y2H	99
5.3 Resultados e Discussão	102
5.3.1 Amplificação dos genes das proteínas candidatas.....	102
5.3.2 Interações proteína-proteína	103
5.4 Referências.....	105
5.5 Figuras Suplementares	107
S.1 Mapas dos vetores pGADT7 AD e pGBKT7	107
I. Abreviações e Símbolos.....	110
II. Lista de Figuras.	11
Capítulo 1	11
Capítulo 2	11
Capítulo 3:	12
Capítulo 4	12
Capítulo 5	13
III. Lista de Tabelas.	14
Capítulo 1:	14
Capítulo 2:	14
Capítulo 3:	14
Capítulo 4:	14
Capítulo 5:	14
III. Lista de Figuras Suplementares.....	15
Capítulo 2	15
Additional file 1: The sequence codes of the viruses used at this work. 55	15
Additional file 2: Ramachandran plot analysis of predicted structure of <i>Groundnut ringspot virus</i> (GRSV) N protein.....	15
Capítulo 3	15
Capítulo 4	15
Figura S1: Gráfico de Ramachandran para a proteína NS ₅ de GRSV. 91	15
Figura S2: Mapa do vetor de expressão pet28a. 92.....	15
Capítulo 5	15

Figura S.1: Mapas dos vetores pGADT7 AD e pGBKT7. 106..... 15

I. Abreviações e Símbolos.

µg	microgramas
µg/ml	microgramas por mililitro
µl	microlitros
µm	micrômetros
µM	micromolar
aa	aminoácidos
Amp (R)	gene de resistência a ampicilina
att	do inglês, <i>attachement sites</i>
BiFC	Fluorescência Bimolecular Complementar
cDNA	DNA sintetizado a partir de um molde de RNA
d.p.i.	dias pós infiltração
DNA	desoxirribonucleato
EUA	Estados Unidos da América
g	grama
GP	poliproteína precursora das glicoproteínas
GRSV	<i>Groundnut ringspot orthospovirus</i>
H	hora
ICTV	Comitê Internacional de Taxonomia de Vírus
IR	do inglês, <i>intergenic region</i>
kb	quilo-pares de bases
kDa	quilodáltons (kilodaltons)
Km(R)	gene de resistência à canamicina
LB	meio de cultura Luria-Bertani
M	molar
min	minutos
ml	mililitros
mM	milimolar
M-MLV	do inglês, <i>Moloney Murine Leukemia Virus</i>
MSC	do inglês, <i>multiple cloning site</i>
ng	nanogramas

nm	nanômetros
NP	proteína do nucleocapsídeo
NS _M	proteína não estrutural do segmento M
NS _S	proteína não estrutural do segmento S
nt	nucleotídeos
°C	graus <i>Celsius</i>
ORF	do inglês, <i>open reading frame</i>
pb	pares de base
PDB	banco de dados proteicos (<i>protein data bank</i>)
pH	concentração de íon hidroxônio livre
RdRp	RNA polimerase RNA dependentes
RI	região intergênica
RNA	ácido ribonucleico
RNA _m	ácido ribonucleico mensageiro
RNP	ribonucleocapsídeo
s	segundos
ssDNA	desoxiribonucleato de fita simples
ssRNA	ácido ribonucleico de fita simples
Taq	<i>Thermus aquaticus</i>
UTR	região não traduzida
vcRNA	fita de RNA complementar à fita viral
vRNA	fita de RNA viral
x g	velocidade de sedimentação gravitacional
Y2HS	sistema de duplo híbrido em levedura
ZLCV	<i>Zucchini lethal chlorosis orthospovirus</i>

II. Lista de Figuras.

Capítulo 1

Figura 1: Árvore filogenética da Ordem <i>Bunyavirales</i>	14
Figura 2: Representação da partícula viral da Ordem <i>Bunyavirales</i>	15
Figura 3: Plantação de alface danificada por infecção de tospovírus.	19
Figura 4: Relação filogenética dos orthotospovírus.	21
Figura 5: Partícula e Genoma de orthotospovírus.	22
Figura 6: Estratégias ambisenso e de senso negativo para a expressão de proteínas de orthotospovírus.	24
Figura 7: Linha do tempo destacando alguns dos principais relatos de GRSV na África e nas Américas.	25
Figura 8: Representação esquemática do Sistema de duplo híbrido em levedura (Y2H).	27
Figura 9: Representação esquemática da técnica de Fluorescência Bimolecular Complementar (BiFC).	28
Figura 10: Representação esquemática da técnica <i>Pulldown</i>	29
Figura 11: Representação esquemática da rede de interações RNA-proteína e proteína-proteína identificada para orthotospovírus.	30
Figura 12: Esquema geral para predição <i>in silico</i> das estruturas tridimensionais de proteínas. Ao lado direito algumas sugestões de programas para serem utilizados em cada etapa.	33

Capítulo 2

Figure 1: <i>Groundnut ringspot virus</i> (GRSV) and <i>La crosse virus</i> (LACV) Nucleoproteins sequence alignment.	47
---	----

Figure 2: Monomeric and tetrameric structure of the <i>Groundnut ringspot virus</i> (GRSV) nucleoprotein (N).....	49
Figure 3: Sequence alignment of representative tospoviruses Nucleoproteins (N).	50
Figure 4: Molecular Dynamics of monomeric Nucleoprotein (N) of <i>Groundnut ringspot virus</i> (GRSV).....	51
Figura 5: Representação esquemática para comparação dos dados obtidos acerca da estrutura tridimensional das proteínas do nucleocapsídeo (NP) de <i>Groundnut ringspot orthotospovirus</i> (GRSV) e <i>Tomato spotted wilt orthotospovirus</i> (TSWV).	57

Capítulo 3:

Figure 1: The NS _M amino acid sequence alignment of ZLCV isolates.....	70
Figure 2: Maximum likelihood trees of tospoviruses based inferred using all viral encoded proteins.	72
Figura 3: Estrutura tridimensional da proteína NP de <i>Zucchini lethal chlorosis orthotospovirus</i> (ZLCV).....	74
Figura 4: Alinhamento entre as sequências de aminoácidos de <i>Groundnut ringspot orthotospovirus</i> (GRSV) e <i>Zucchini lethal chlorosis orthotospovirus</i> (ZLCV).....	75

Capítulo 4

Figura 1: Alinhamento entre as sequências da proteína NS _S de <i>Groundnut ringspot orthotospovirus</i> (GRSV) e a polifosfato kinase (Ppk) bacteriana (PDB: 2o8r).	81
Figura 2: Estrutura monomérica da proteína NS _S de <i>Groundnut ringspot orthotospovirus</i> (GRSV).	82
Figura 3: Representação esquemática da hipótese de interação do motivo GW/WG com a proteína Argonauta.	83

Figura 4: Motivos preditos de ligação à GTP/ATP da proteína NS _S de <i>Groundnut ringspot orthotospovirus</i> (GRSV).....	84
Figura 5: Alinhamento entre as sequências da proteína NS _S de algumas espécies de orthotospovírus.....	85
Figura 6: Expressão heteróloga da proteína NS _S de <i>Groundnut ringspot orthotospovirus</i> (GRSV).....	89

Capítulo 5

Figura 1: Representação esquemática da via de silenciamento gênico vegetal. . .	94
Figura 2: Amplificação dos genes das proteínas candidatas para interação com N, NS _M e NS _S	101
Figura 3: Digestão dos vetores recombinantes insertos + pGEM-T <i>easy</i>	102
Figura 4: Ensaio de Duplo híbrido em Levedura (Y2H) para interações proteína-proteína entre a NP, NS _M e NS _S de GRSV e <i>A. thaliana</i>	103

III. Lista de Tabelas.

Capítulo 1:

Tabela 1: Ordem *Bunyavirales*..... 13

Tabela 2: Diferenças entre os genomas na ordem *Bunyavirales*..... 16

Tabela 3: Hospedeiros e vetores da ordem *Bunyavirales*..... 17

Tabela 4: Espécies aceitas e espécies tentativas para a Família *Tospoviridae*.... 19

Capítulo 2:

Table 1: Pairs of interacting residues for GRSV N-N oligomerization..... 47

Capítulo 3:

Table1: Genome comparison of ZLCV isolates. 67

Table 2: Identity among ZLCV isolates proteins. 69

Table 3: GenBank accession number of the sequences used in this study. 71

Capítulo 4:

Tabela 1: Propriedades da proteína NS_S de GRSV. 88

Capítulo 5:

Tabela 1: Oligonucleotídeos utilizados para amplificar as proteínas candidatas. ..100

IV. Lista de Figuras Suplementares.

Capítulo 2

Additional file 1: The sequence codes of the viruses used at this work. 55

Additional file 2: Ramachandran plot analysis of predicted structure of *Groundnut ringspot virus* (GRSV) N protein. 55

Capítulo 3

Figura S1: Gráfico de Ramachandran para a proteína NP de *Zucchini lethal chlorosis orthotospovirus* (ZLCV). 76

Capítulo 4

Figura S1: Gráfico de Ramachandran para a proteína NS_S de GRSV. 91

Figura S2: Mapa do vetor de expressão pet28a. 92

Capítulo 5

Figura S.1: Mapas dos vetores pGADT7 AD e pGBKT7. 106

1 **Resumo**

2

3 O rápido progresso na compreensão dos mecanismos de enovelamento de proteínas e os
4 avanços no campo da bioinformática forneceram ferramentas confiáveis para predizer as
5 estruturas tridimensionais de proteínas de vírus de plantas. Por meio de Modelagem
6 Estrutural por Homologia, foi possível obter um modelo tridimensional para a proteína
7 do nucleocapsídeo (NP) e a proteína não estrutural do segmento S (NSs) de GRSV
8 (*Groundnut ringspot orthotospovirus*) e para a NP de ZLCV (*Zucchini lethal chlorosis*
9 *orthotospovirus*). Verificou-se que os monômeros GRSV NP e ZLCV NP são
10 organizadas em um domínio globular (33-223 aa) contendo uma cavidade carregada
11 positivamente para interação com RNA e com as duas cadeias terminais, formando um
12 braço N-terminal (1-32 aa) e um braço C-terminal (224-258 aa). Além disso, análise da
13 estrutura tridimensional da proteína NSs de GRSV sugeriu uma possível atividade
14 enzimática de fosfatase, para o metabolismo de nucleotídeos assim como um possível
15 domínio de interação com a proteína Argonauta 1. A proteína NSs foi expressa de forma
16 nativa e purificada para futuros ensaios de cristalização. Assim, as estruturas das NP e
17 NSs podem lançar luz sobre os mecanismos de formação de RNP e silenciamento
18 gênico e podem permitir a identificação de resíduos essenciais de aminoácidos como
19 possíveis alvos para estratégias de controle das doenças causadas pelos orthotospovírus.
20 Por fim, por meio da técnica de duplo híbrido em levedura, foi encontrada uma possível
21 interação entre a proteína AtCSN5a e as proteínas NP e NSs de GRSV, e a relevância
22 dessas interações para o sistema planta-vírus ainda serão investigadas.

23 **Palavras-chave:** orthotospovírus, sistema de duplo híbrido em levedura, modelagem
24 estrutural por homologia.

25 **Abstract**

26

27

28 The rapid progress in the understanding of protein folding mechanisms and the
29 advances in the bioinformatics field have provided reliable tools for modeling and
30 predict three-dimensional structures of plant virus proteins. Using Homology modeling
31 technique, it was possible to obtain three-dimensional models for both NP and NSs of
32 GRSV (*Groundnut ringspot orthospovirus*) and for ZLCV NP (*Zucchini lethal*
33 *chlorosis orthospovirus*). GRSV NP and ZLCV NP monomers have been organized
34 into a globular domain (33-223 aa) containing a positively charged cavity for RNA
35 interactions, and two terminal chains forming an N-terminal arm (1-32 aa) and a C-
36 terminal arm (224-258 aa). In addition, a three-dimensional structure of the GRSV NSs
37 protein revealed a possible phosphatase enzymatic activity for nucleotide metabolism
38 and a possible interaction domain with an Argonaute 1. Moreover, the NS_S protein was
39 natively expressed and purified for future crystallization assay. Thus, the proposed
40 models can shed light on the mechanisms of RNP formation and gene silencing and
41 may allow the identification of essential amino acid residues as possible targets for the
42 orthospovirus control strategy. Finally, using yeast two-hybrid technique, a possible
43 interaction between the AtCSN5a protein and the NP and NS_S of GRSV was found;
44 however, the relevance of these interactions for the plant-virus system remains to be
45 investigated.

46 **Keywords:** orthospovirus, yeast two hybrid system, homology modeling.

47

48

49

Motivação e Justificativa

50 Uma infecção viral eficiente requer interações entre proteínas virais e proteínas do
51 hospedeiro. Essas interações proteína-proteína alteram a rotina da célula para promover
52 todos os eventos do ciclo infectivo de um vírus como a transcrição e replicação,
53 formação dos ribonucleocapsídeos, movimento célula a célula e sistêmico,
54 sintomatologia viral, montagem da partícula viral e a supressão dos mecanismos de
55 defesa do hospedeiro e do vetor.

56 Comparado a outros vírus de plantas e apesar de várias espécies já terem sido
57 caracterizadas, pouco ainda se sabe a respeito das interações entre as proteínas dos
58 orthospovírus com seus vetores e seus hospedeiros e a ausência de um sistema de
59 genética reversa é um fator limitante para o avanço de estudos relacionados ao processo
60 infectivo desses vírus.

61 A adaptação e aplicação de técnicas *in vivo*, *in vitro* e *in silico* para a predição da
62 estrutura tridimensional e para a identificação de novas interações entre as proteínas dos
63 orthospovírus e seus hospedeiros podem fornecer conhecimentos relevantes sobre
64 novas funções biológicas e suas implicações para a infecção resultando em progressos
65 para as estratégias aplicadas de controle da infecção.

66 Ao contrário dos vírus da ordem *Bunyavirales* e de outros vírus vegetais, pouco
67 ainda é conhecido a respeito das interações proteicas de orthospovírus. Portanto torna-
68 se importante o estudo de suas estruturas, seus domínios funcionais e as interações com
69 proteínas das plantas hospedeiras para elucidar os mecanismos envolvidos na infecção
70 viral e dos mecanismos de interação planta-vírus. Com esse propósito, o presente
71 projeto visou estudar a proteína não estrutural do segmento S, NSs (supressão do
72 silenciamento gênico) e a proteína estrutural NP (proteína do nucleocapsídeo) de
73 espécies de orthospovírus. Esses estudos resultaram em conhecimentos que
74 contribuirão para o entendimento do movimento do vírus na planta, da formação da
75 partícula e da supressão do silenciamento gênico, fornecendo evidências do
76 comportamento molecular do vírus na célula vegetal e evidências para o
77 desenvolvimento de estratégias de controle da infecção.

78

79 **Objetivos**

80

81 **Objetivo Geral**

82 O presente trabalho intencionou investigar as estruturas tridimensionais e as
83 interações moleculares das proteínas do nucleocapsídeo (NP) de *Groundnut ringspot*
84 *orthospovirus* (GRSV) e de *Zucchini lethal chlorosis orthospovirus* (ZLCV), e da
85 proteína de supressão do silenciamento gênico (NS_S) de GRSV. Assim como expressou
86 e purificou a NSs de GRSV para a futura resolução da estrutura tridimensional da
87 mesma.

88 O presente trabalho também intenciona buscar possíveis proteínas candidatas da
89 planta hospedeira que possam interagir com as proteínas virais NP, NS_S e NS_M.

90

91 **Objetivos Específicos**

92

- 93 • Identificar proteínas celulares candidatas para a interação com as proteínas virais
94 NP, NS_S e NS_M utilizando técnicas como Sistema de duplo híbrido em levedura
95 (Y2H);
- 96 • Validar as interações proteína-proteína por técnicas de Fluorescência
97 Bimolecular Complementar (BiFC) e *Pull-down*;
- 98 • Prever a estrutura tridimensional das proteínas NP de GRSV, NP de ZLCV e
99 NS_S de GRSV por modelagem teórica e analisar os seus possíveis domínios
100 funcionais e de interações proteicas;
- 101 • Expressar e purificar a proteína NS_S em cepas de *Escherichia coli* em diferentes
102 condições;

CAPÍTULO I: Introdução Geral

103 **1.1 Ordem *Bunyavirales***

104 **1.1.1 Taxonomia e Diversidade**

105 A classificação padrão da família *Bunyaviridae* mudou consideravelmente em
106 2016 [1]. O 10º relatório online do Comitê Internacional de Taxonomia de Vírus
107 (https://talk.ictvonline.org/ictv-reports/ictv_online_report/) estabeleceu uma nova
108 ordem viral denominada *Bunyavirales*, a fim de incluir alguns vírus já reconhecidos
109 pelo ICTV, mas que ainda não foram atribuídos a uma família, bem como alguns
110 arbovírus recém-identificados em diversas bibliotecas metagenômicas provenientes de
111 artrópodes. De acordo com os autores da proposta, a manutenção da antiga classificação
112 em uma única família não retrataria todo o conhecimento sobre a biodiversidade e
113 as relações evolutivas entre os bunyavírus. O nome dessa nova ordem deriva do
114 *Bunyamwera virus* (BUNV), identificado pela primeira vez em 1943, na cidade de
115 Bunyamwera, Uganda [2]. As famílias filogeneticamente mais próximas aos bunyavírus
116 são a *Arenaviridae* e a *Orthomyxoviridae* [3].

117 A ordem *Bunyavirales* compreende treze gêneros distribuídos em nove famílias
118 virais que impactam economicamente setores da saúde e da agropecuária mundial
119 (Tabela 1).

120 Os gêneros *Orthobunyavirus*, *Orthohantavirus*, *Pheblovirus* e *Orthonairovirus*
121 infectam inúmeros vertebrados, causando doenças clínicas severas como a febre
122 hemorrágica e a encefalite, a Febre do Oropouche (OROV, *Orthobunyavirus*) e a
123 Síndrome Cardiopulmonar por Hantavírus (JUQV, *Orthohantavirus*) [4-6]. Além disso,
124 muitos bunyavírus são teratogênicos em animais domésticos gestantes como gado e
125 ovelha [7]. Os gêneros *Emaravirus*, *Tenuivirus* e *Orthospovirus* infectam hortaliças e
126 plantas ornamentais, além de muitas das *commodities* agrícolas como soja, milho e café

127 [8-10]. Por sua vez, *Herbesvirus*, *Orthoferavirus*, *Phasivirus*, *Goukovirus*,
 128 *Orthophasmavirus* e *Orthojonvirus* foram identificados apenas em insetos, a partir das
 129 tecnologias de sequenciamento genômico de alto desempenho [11-18].

130 Tabela 1. Ordem *Bunyvirales*.

Família	Gênero	Espécie tipo
<i>Feraviridae</i>	<i>Orthoferavirus</i>	<i>Ferak orthoferavirus</i>
<i>Fimoviridae</i>	<i>Emaravirus</i>	<i>European mountain ash ringspot-associated emaravirus</i>
<i>Hantaviridae</i>	<i>Orthohantavirus</i>	<i>Hantaan orthohantavirus</i>
<i>Jonviridae</i>	<i>Orthojonvirus</i>	<i>Jonchet virus</i>
<i>Nairoviridae</i>	<i>Orthonairovirus</i>	<i>Dugbe orthonairovirus</i>
<i>Peribunyaviridae</i>	<i>Orthobunyavirus</i>	<i>Bunyamwera Orthobunyavirus</i>
	<i>Herbesvirus</i>	<i>Herbert virus</i>
<i>Phasmaviridae</i>	<i>Orthophasmavirus</i>	<i>Kigluaik phantom orthophasmavirus</i>
<i>Phenuiviridae</i>	<i>Goukovirus</i>	<i>Gouléako virus</i>
	<i>Phasivirus</i>	<i>Phasi charoen virus</i>
	<i>Phlebovirus</i>	<i>Rift Valley fever virus</i>
	<i>Tenuivirus</i>	<i>Rice stripe virus</i>
<i>Tospoviridae</i>	<i>Orthotospovirus</i>	<i>Tomato spotted wilt orthotospovirus</i>

131 Embora esses gêneros apresentem variadas formas de infecção, estrutura e
 132 genoma, o estabelecimento da nova ordem fez-se imprescindível, pois as análises
 133 filogenéticas indicam que esses vírus podem ter uma origem evolutiva compartilhada
 134 (Fig.1). Além do mais, inúmeras espécies têm sido identificadas e investigadas como
 135 possíveis membros dessa ordem, mas ainda não foram nela incluídas devido à ausência
 136 da sequência genômica completa e/ou caracterização biológica [3, 19, 20].

137 Guterres et al. (2017) realizou uma análise detalhada das relações filogenéticas
 138 entres os membros da ordem *Bunyvirales* [21]. A inferência filogenética baseada nas
 139 sequências das proteínas RNA polimerase RNA dependentes (RdRp) de membros de
 140 cada gênero sugere a existência de pelo menos três supergrupos monofiléticos:

- 141 I- Composto por *Fimoviridae*, *Tospoviridae*, *Peribunyaviridae*,
 142 *Hantaviridae*, *Feraviridae*, *Jonviridae* e *Phasmaviridae*;
 143 II- Composto por *Phenuiviridae* e;

144

III- Composto por *Nairoviridae* e incluindo a família *Arenaviridae*.

145

A família *Arenaviridae* está mais filogeneticamente relacionada com os nairovírus

146

do que com outras famílias virais identificadas até o momento [3, 21]. Embora os

147

arenavírus possuam genoma segmentado de RNA fita simples e se agrupem com os

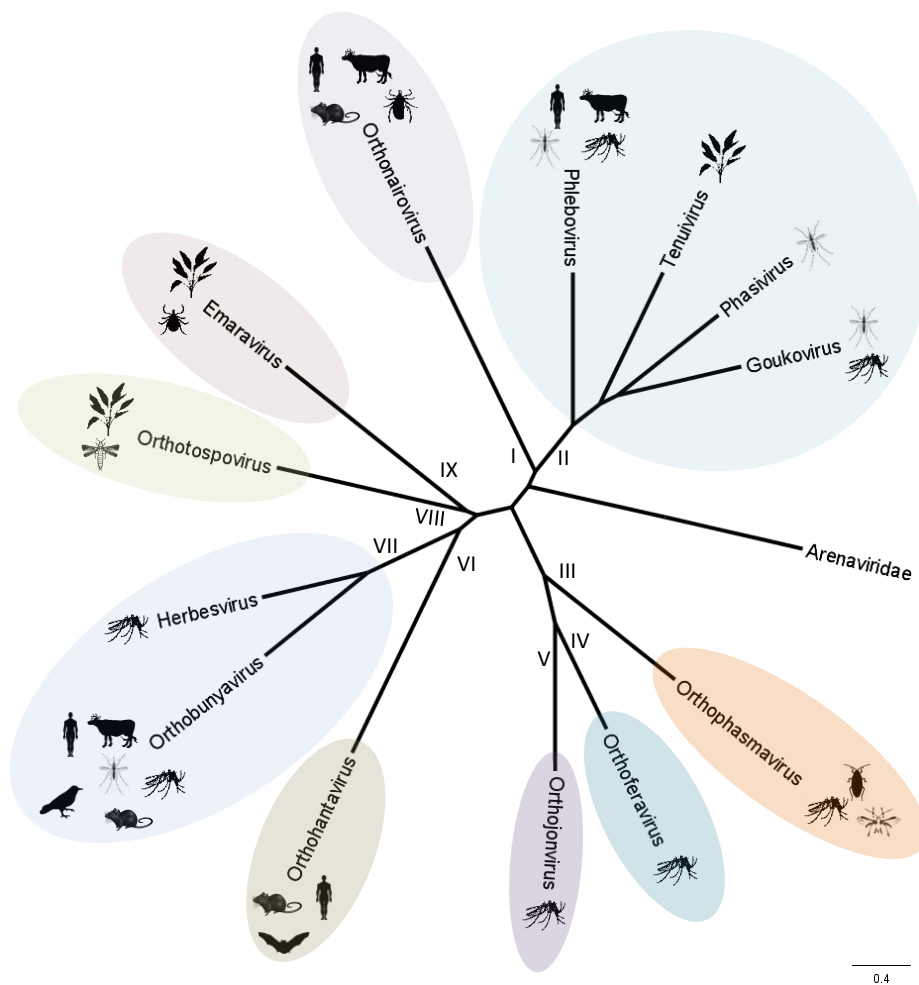
148

nairovírus, somente estudos futuros poderão ratificar a inclusão da família *Arenaviridae*

149

na ordem *Bunyavirales*.

150



151

Figura 1: **Árvore filogenética da Ordem *Bunyvirales*.** I: *Nairoviridae*; II: *Phenuiviridae*; III: *Phasmaviridae*; IV: *Feraviridae*; V: *Jonviridae*; VI: *Hantaviridae*; VII: *Peribunyaviridae*; VIII: *Tospoviridae*; IX: *Fimoviridae*. As lustrações representam alguns hospedeiros e vetores virais. Árvore não enraizada baseada na sequência de aminoácidos da RdRp de cada espécie tipo do gênero. As sequências foram alinhadas usando-se o MAFFT e as árvores de máxima verossimilhança foram inferidas usando-se o FastTree, ambos implementados no programa Geneious 9.1. Os valores de suporte foram suprimidos a fim de aperfeiçoar a visualização da imagem.

152 1.1.2 Vírions e Genoma

153 Os arbovírus da ordem *Bunyavirales* possuem em comum, partículas virais
154 pleomórficas, medindo de 80 a 200 nm (nanômetros) de diâmetro e envelope lipídico
155 com projeções glicoproteicas na superfície [22]. A exceção seria o gênero *Tenuivirus*
156 que possui vírions não envelopados medindo entre 3 e 10 nm de diâmetro [8].

157 O genoma viral é de RNA fita simples (ssRNA) segmentado de orientação
158 negativa ou ambisenso. Cada segmento está envolvido por proteínas do nucleocapsídeo
159 e moléculas de RdRp também estão presentes (Fig. 2) [22]. No entanto, esses vírus
160 apresentam diferenças em relação ao uso de vetores, ciclo de infecção, tamanho do
161 genoma, número de segmentos de ssRNA, na sequência consenso terminadora e na
162 expressão de proteínas não estruturais (Tabelas 2 e 3).

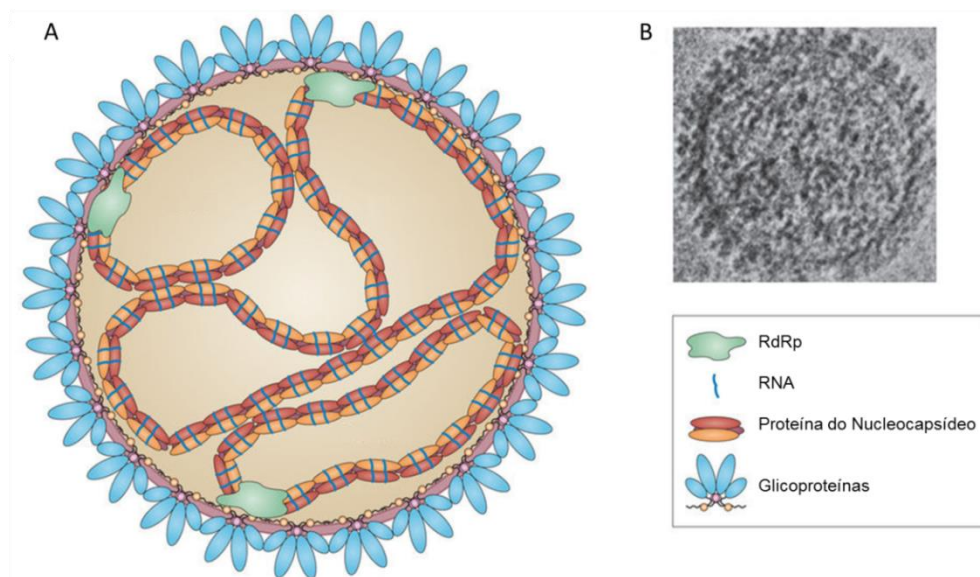


Figura 2: **Representação da partícula viral da Ordem *Bunyavirales*.** A: Ilustração esquemática da partícula de hantavírus mostrando a RdRp, as inclusões do complexo de Glicoproteínas G_N e G_C no envelope lipídico e o genoma trissegmentado envolvido pelas proteínas do Nucleocapsídeo. O diâmetro médio da partícula é de 135nm. B: Partícula de hantavírus observada por microscopia crioelétrica. Figura modificada de [doi:10.1038/nrmicro3066](https://doi.org/10.1038/nrmicro3066) [22].

163 Os arbovírus (*Arthropod-borne virus*) dependem de vetores artrópodes para sua
164 transmissão entre os hospedeiros. A capacidade do vetor em adquirir e transmitir um

165 arbovírus é referida como competência vetorial. Sabe-se que diferentes espécies de
 166 vetores e até mesmo populações, possuem vários graus de competências vetoriais
 167 distintas [23-25].

168 Os vírus trafegam pelo corpo do vetor e coevoluem constantemente com seus
 169 vetores [17]. Os arbovírus desenvolveram mecanismos específicos para permitir que
 170 eles sejam transportados através de vários tecidos e evadam as defesas vetoriais [17,
 171 26]. Em resposta, as espécies de vetores evoluíram de modo que nem todos os
 172 indivíduos dentro de uma espécie são suscetíveis à infecção, servindo como um vetor
 173 competente.

174 Os mecanismos das associações de artrópodes-vírus são variados e complexos, porém
 175 as novas tecnologias de sequenciamento de alto desempenho têm permitido avanços nas
 176 investigações desses mecanismos através da metagenômica, transcriptômica e
 177 proteômica [27-31]. Não só os componentes virais do processo de transmissão estão
 178 sendo identificados, como também estão sendo elucidados os componentes genéticos e
 179 fisiológicos dos vetores que determinam sua capacidade de serem utilizados com
 180 sucesso pelo vírus. Compreender os mecanismos subjacentes às interações arbovírus-
 181 vetor será fundamental para o desenvolvimento de novas estratégias de controle.

182

183 Tabela 2. Diferenças entre os genomas na ordem *Bunyvirales*.

Gênero	Nº de segmentos de RNA	Sequência terminadora 5' consenso	Presença de proteínas não estruturais
<i>Orthoferavirus</i>	3	AGUAGUA	NS _S e NS _M
<i>Emaravirus</i>	4 a 8	AGUAGUGAUCUCCC	Diversas
<i>Orthohantavirus</i>	3	UAGUAGUAUGC	X
<i>Orthojonvirus</i>	3	AGUAGUA	NS _S e NS _M
<i>Orthonairovirus</i>	3	UCUCAAGA	X
<i>Orthobunyavirus</i>	3	AGUAGUG	NS _S e NS _M
<i>Herbesvirus</i>	3	AGUAGUG	NS _S e NS _M
<i>Orthophasmavirus</i>	3	AGCAGCA	NS _S
<i>Goukovirus</i>	3	ACACA	X
<i>Phasivirus</i>	3	ACACAAAG	X
<i>Phlebovirus</i>	3	ACACAAAG	NS _S
<i>Tenuivirus</i>	3 a 4	ACACAAAG	Diversas
<i>Orthospovirus</i>	3	AGAGCAAU	NS _S e NS _M

184 Tabela 3. Hospedeiros e vetores da ordem *Bunyvirales*.

Gênero	Encontrado em:	
<i>Orthoferavirus</i>	Invertebrado <i>Anopheles gambiae</i> <i>Culex sp.</i> <i>Uranotaenia mashonaensis</i>	
<i>Emaravirus</i>	Vegetal (H) Invertebrado (V) <i>Aceria ficus</i> <i>Eriophyes pyri</i>	
<i>Orthohantavirus</i>	Vertebrado (H) Humanos Morcegos Roedores	
<i>Orthojonvirus</i>	Invertebrado <i>Culex sp.</i>	
<i>Orthonairovirus</i>	Vertebrado (H) Dromedários Humanos Morcegos Roedores Pombos	Invertebrado (V) <i>Carios sp.</i> <i>Dermacentor marginatus</i> <i>Hyalomma dromedarii</i> <i>Ixodes uriae</i> <i>Ornithodoros sp.</i>
<i>Orthobunyavirus</i>	Vertebrado (H) Caprinos Gado Guaxinim Humano Macaco Morcego Roedores	Invertebrado (V) <i>Aedeomyia squamipennis</i> <i>Aedes sp.</i> <i>Anopheles sp.</i> <i>Culicoides sp.</i> <i>Limatus flavisetosus</i> <i>Ochlerotatus communis</i> <i>Wyeomyia sp.</i>
<i>Herbesvirus</i>	Invertebrado <i>Culex sp.</i> <i>Culicidae sp.</i> Chrysopidae	
<i>Orthophasmavirus</i>	Invertebrado <i>Aedes sp.</i> <i>Anopheles sinensis</i> <i>Blattella germânica</i> <i>Chaoborus flavicans</i> <i>Chaoborus trivittatus</i> <i>Culex sp.</i> <i>Mansonia sp.</i> <i>Ostrinia furnacalis</i>	
<i>Goukovirus</i>	Invertebrado <i>Anopheles gambiae</i> <i>Aulacorthum magnoliae</i> <i>Culex sp.</i> <i>Culicidae sp.</i> <i>Uranotaenia mashonaensis</i>	
<i>Phasivirus</i>	Invertebrado <i>Aedes aegypti</i>	
<i>Phlebovirus</i>	Vertebrado (H) Cachorro Gado Gambá Humano	Invertebrado (V) <i>Ixodes sp.</i> <i>Phlebotomus sp.</i> <i>Rhipicephalus sp.</i> <i>Haemaphysalis sp.</i>
<i>Tenuivirus</i>	Vegetal (H) Invertebrado (V) <i>Laodelphax striatellus</i> <i>Peregrinus maidis</i>	
<i>Orthotospovirus</i>	Vegetal (H) <i>Solanaceae, Curcubitaceae, Asteraceae, Fabaceae, Iridaceae, Amaryllidaceae</i> Invertebrado (V) <i>Thrips sp.</i> <i>Frankliniella sp.</i>	

185 V= vetor;
186 H= hospedeiro.

187 **1.2 Família *Tospoviridae***

188 **1.2.1 Taxonomia e diversidade**

189 De acordo com os novos critérios estabelecidos pelo ICTV, o antigo gênero
190 *Tospovirus* foi renomeado para *Orthotospovirus* e inserido na Família *Tospoviridae*. O
191 nome do gênero é derivado da espécie de referência: “*TO*mato *SPO*tted wilt *VIRUS*”
192 (TSWV).

193 O critério de classificação taxonômica é baseado na morfologia da partícula, na
194 sorologia e análise comparativa da identidade das sequências das proteínas do
195 nucleocapsídeo (NP), e no espectro de hospedeiros de cada possível espécie. Sendo a
196 identidade de sequência da proteína NP entre dois orthotospovírus menor que 90%, os
197 isolados são considerados espécies distintas.

198 Atualmente já foram identificadas 29 espécies de orthotospovírus, sendo onze
199 espécies aceitas e dezoito ainda consideradas espécies tentativas (Tabela 4) [10, 32]. No
200 Brasil há uma grande diversidade de orthotospovírus, com a presença de isolados de
201 TSWV, GRSV, ZLCV, TCSV, CSNV, IYSV e BeNMV, identificados em campos em
202 diversas culturas em todo território nacional [10, 32-34].

203 Os orthotospovírus são encontrados em todos os continentes e infectam uma
204 grande variedade de cultivares de hortaliças, grãos e plantas ornamentais
205 economicamente importantes como crisântemo, íris, cebola, tomate, pimentão, alface,
206 abobrinha, melancia, batata, amendoim, soja e feijão [10, 33, 35].

207 Os sintomas iniciais apresentados pelas plantas infectadas são anéis cloróticos,
208 anéis necróticos, distorções folheares e bronzeamento. Os sintomas mais severos da
209 infecção resultam em necroses caulinares, redução da produtividade devido à morte da
210 planta (Fig. 3) ou perda do valor agregado dos frutos comerciais [10, 36]. O “Vira-

211 cabeça” do tomateiro e o “sapecá” da cebola são exemplos de doenças causadas pelos
 212 orthotospovírus nas condições brasileiras.

213 Tabela 4: Espécies aceitas e espécies tentativas para a Família *Tospoviridae*.

Acronym	Tospovirus	L RNA	M RNA	S RNA
ANSV	Alstroemeria necrotic streak orthotospovirus	-	-	NP: GQ478668
BeNMV	Bean necrotic mosaic orthotospovirus	JF417980	JN587269	JN587268
CaCV	Capsicum chlorosis orthotospovirus	NC_008302	NC_008303	NC_008301
CCSV	Calla lily chlorotic spot orthotospovirus	FJ822962	FJ822961	AY867502
CSNV	Chrysanthemum stem necrosis orthotospovirus	KM114546	KM114547	KM114548
GBNV	<i>Groundnut bud necrosis orthotospovirus</i>	AF025538	U42555	AY871098
GCFSV	Groundnut chlorotic fan- spot orthotospovirus	KP146140	KP146141	AF080526
GRSV	<i>Groundnut ringspot orthotospovirus</i>	KT9725901	NS _M : AF213673 GP: AY574055	NP: AF487516 NSs: JN571117
GYSV	<i>Groundnut yellow spot orthotospovirus</i>	-	-	HQ402596
HCRV	Hippeastrum chlorotic ringspot orthotospovirus	HG763861	JX833565	JX833564
INSV	<i>Impatiens necrotic spot orthotospovirus</i>	DQ425094	DQ425095	X66972
IYSV	<i>Iris yellow spot orthotospovirus</i>	FJ623474	FJ361359	AF001387
LNRV	Lisianthus necrotic ringspot orthotospovirus	-	-	AB852525
MSMV	Melon severe mosaic orthotospovirus	FJ157985	FJ157984	EU275149
MVBaV	Mulberry vein banding associated orthotospovirus	KM819698	KM819699	KM819701
MYSV	Melon yellow spot orthotospovirus	AB061774	AB061773	AB038343
PCSV	Pepper chlorotic spot orthotospovirus	KY315809	KY315810	KF383956
PNSV	Pepper necrotic spot orthotospovirus	-	-	HE584762
PoIRSV	<i>Polygonum ringspot orthotospovirus</i>	KJ541746	KJ541745	KJ541744
SVNaV	Soybean vein necrosis associated orthotospovirus	HQ728385	HQ728386	HQ728387
TCSV	<i>Tomato chlorotic spot orthotospovirus</i>	KX463272	KX463273	KX463274
TNRV	Tomato necrotic ringspot orthotospovirus	-	FJ947152	FJ489600
TNSV	Tomato necrotic spot orthotospovirus	-	-	KM355773
TSWV	<i>Tomato spotted wilt orthotospovirus</i>	NC_002052	NC_002050	DQ915948
TYRV	Tomato yellow ring orthotospovirus	JN560178	JN560177	AY686718
TZSV	Tomato zonate spot orthotospovirus	EF552435	EF552434	EF552433
WBNV	<i>Watermelon bud necrosis orthotospovirus</i>	GU735408	GU584185	GU584184
WSMoV	<i>Watermelon silver mottle orthotospovirus</i>	AY863200	DQ157768	AB042650
ZLCV	<i>Zucchini lethal chlorosis orthotospovirus</i>	KU641378	KU641379	KU641380

214



215

Figura 3: **Plantação de alface danificada por infecção de tospovírus.** Necroses foliares em plantas de alface na Califórnia, Estados Unidos (Foto do acervo pessoal).

216

217 Na árvore filogenética das espécies (Fig. 4) é possível diferenciar quatro clados
218 principais, um clado que representa o grupo dos vírus Americanos, um clado que
219 representa o grupo Euro-asiático e dois clados distintos: o clado com *Lisianthus necrotic*
220 *ringspot orthotospovirus* (LNRV) como único representante e o clado formado por
221 *Groundnut chlorotic fan-spot orthotospovirus* (GCFSV) e *Groundnut yellow spot*
222 *orthotospovirus* (GYSV) [32].

223 A classificação dos clados não corresponde totalmente à distribuição geográfica
224 dos orthotospovírus. TSWV, INSV, IYSV, dentre outros, estão distribuídos
225 mundialmente, provavelmente devido à ampla distribuição de seus vetores *Frankliniella*
226 *occidentalis* e *Thrips tabaci* e/ou intercâmbio de material vegetal infectado [37].

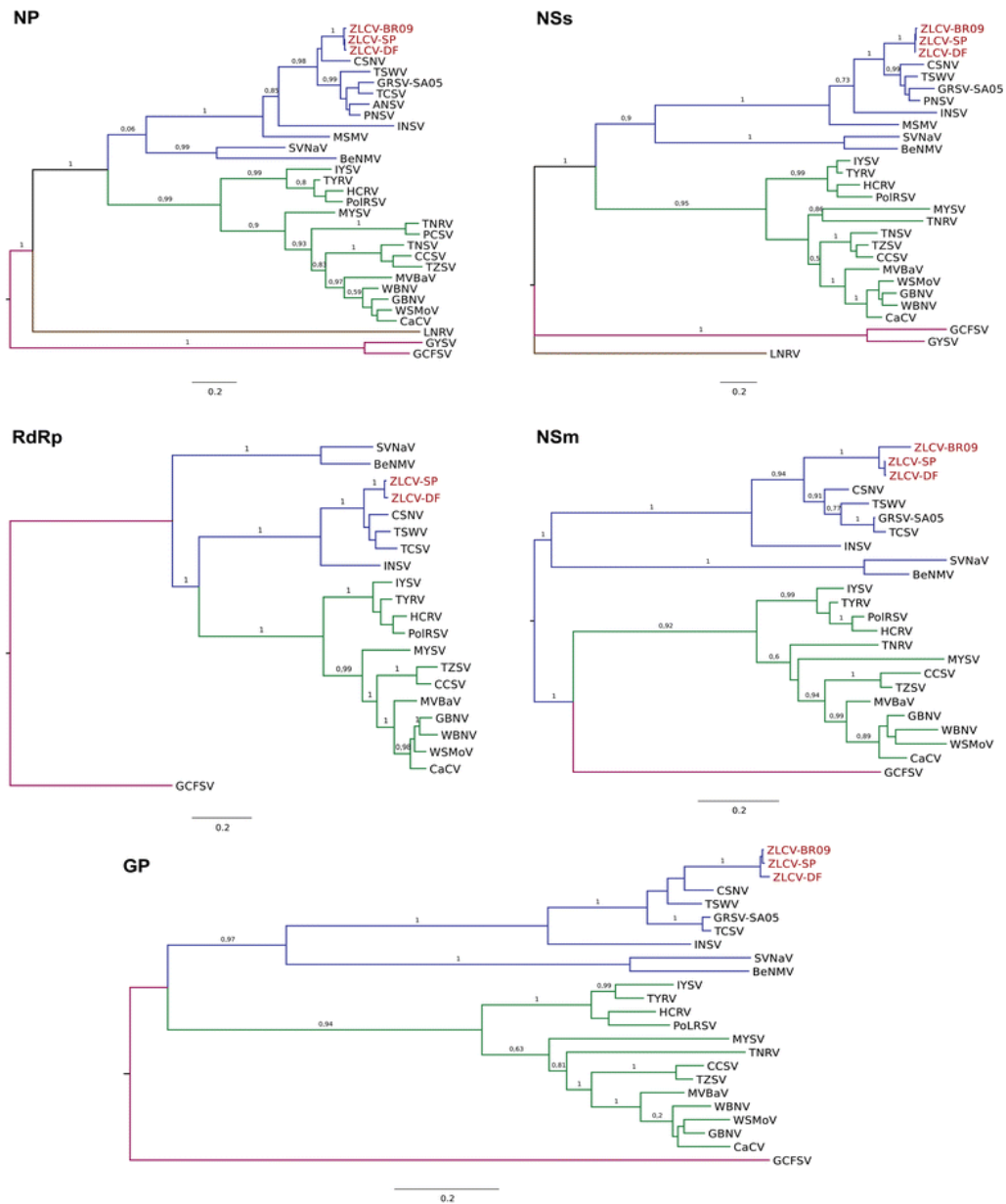
227 **1.2.1 Vírions e Genoma**

228 A partícula viral dos orthotospovírus é esférica e como o vírus “brota” a partir do
229 retículo endoplasmático e/ou do Golgi da célula vegetal, ele leva consigo um envelope
230 lipídico formado pela membrana dessas organelas. Posteriormente, as glicoproteínas são
231 exteriorizadas, envolvendo toda a partícula (Fig.5).

232 Cada partícula pode variar de 80 a 120 nm de diâmetro, de acordo com o número
233 de segmentos encapsidados. Para *Rift Valley fever virus* (RVFV) a proporção média de
234 segmentos L, M e S é de 1: 4: 4 respectivamente [38]; para orthotospovírus, a
235 estequiometria média dos segmentos encapsidados ainda é desconhecida.

236 O envelope lipídico, com as glicoproteínas imersas, circunda os três
237 ribonucleocapsídeos (RNPs) que são formados pelo RNA viral (L, M e S), envolvido e
238 protegido pela proteína NP com a presença de moléculas de RNA polimerase na
239 extremidade de cada ribonucleocapsídeo (Fig. 5). As extremidades 3' (UCUCGUUA) e
240 5' (AGAGCAAU) do RNA são conservadas e apresentam nucleotídeos
241 complementares, o que favorece a ligação não covalente dessas extremidades formando,

242 assim, uma molécula de ribonucleocapsídeo (RNP) pseudocircular, potencialmente
 243 importante no reconhecimento da polimerase viral para os processos de replicação e
 244 transcrição [39].
 245



246

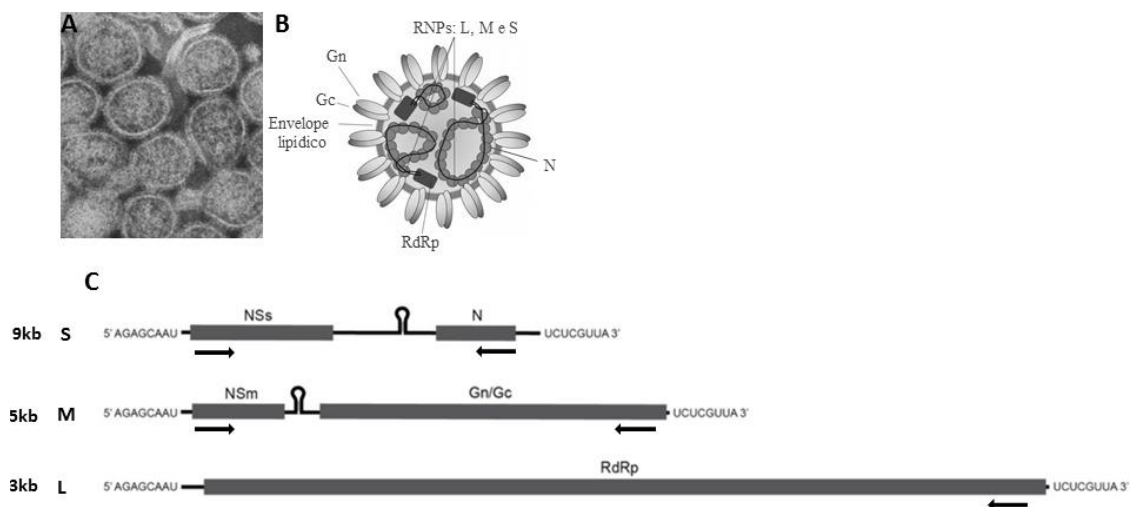
Figura 4: **Relação filogenética dos orthotospovírus.** Árvore filogenética baseada nas sequências disponíveis de aminoácidos das proteínas estruturais (NP, GP e RdRp) e não estruturais (NS_S e NS_M). Na árvore filogenética observa-se a formação de quatro clados principais, o clado de vírus euroasiáticos (verde), o clado de vírus americanos (azul), o clado com *Lisianthus necrotic ringspot orthotospovirus* (LNRV) como único representante (marrom) e, em rosa o clado formado por *Groundnut chlorotic fan-spot orthotospovirus* (GCFSV) e *Groundnut yellow spot orthotospovirus* (GYSV).

247

248 O genoma trissegmentado de RNA apresenta cinco fases abertas de leitura que
 249 codificam seis proteínas multifuncionais virais, sendo que quatro delas são estruturais,
 250 ou seja, estão presentes na partícula viral madura (NP, G_C, G_N e RdRp) e duas são
 251 proteínas não estruturais, pois não estão presentes na partícula (NS_S e a NS_M). O RNA L
 252 codifica a RNA polimerase dependente de RNA (RdRp); o RNA S codifica a proteína
 253 do nucleocapsídeo (NP) e a proteína com função de supressora do silenciamento gênico
 254 (NS_S) e o RNA M codifica a proteína precursora das glicoproteínas G_N/G_C e a NS_M,
 255 proteína de movimento viral célula-a-célula via plasmodesma e/ou floema (Fig. 5).

256 Os segmentos M e S apresentam uma região intergênica rica em nucleotídeos
 257 adenina e uracila e, por consequência, forma-se nessa região um *hairpin*, estrutura
 258 importante na regulação da transcrição e da replicação (Fig. 5).

259



260

261 **Figura 5: Partícula e Genoma de orthospovírus.** A: Microscopia eletrônica de transmissão mostrando
 262 partículas pleomórficas de orthospovírus (Foto do acervo pessoal de Renato de Oliveira Resende); B:
 263 Ilustração da partícula viral representando a disposição do envelope lipídico, dos ribonucleocapsídeos L,
 264 M e S e das proteínas estruturais N, Gn, Gc e RdRp (ilustração adaptada de <http://www.bunyavirus.org>);
 265 C: Esquema representativo com os tamanhos aproximados do genoma de orthospovírus, evidenciando
 266 as ORFs para as proteínas virais, a sequência intergênica com a formação do *hairpin* e os oito
 267 nucleotídeos complementares presentes nas extremidades dos segmentos.
 268

269 1.2.2 Transcrição e replicação

270 Apesar dos orthotospovírus poderem ser propagados para a planta por
271 inoculação mecânica, naturalmente, os orthotospovírus infectam a planta por meio de
272 um tripes. Ao entrar na célula vegetal, os ribonucleocapsídeos virais (RNPs) ficam
273 expostos no citoplasma, os processos de replicação, transcrição e tradução se iniciam e
274 novas partículas são formadas [40].

275 Os RNAs recém-replicados são envolvidos pela proteína NP juntamente com
276 pequenas quantidades de proteína L, formando novos RNPs. Estes RNPs são
277 transportados para as células vizinhas pela ajuda da NS_M, que forma estruturas tubulares
278 que adentram o plasmodesma facilitando assim a translocação de RNPs virais pela rota
279 simplástica da planta. Deste modo, a infecção viral se propaga para os tecidos vizinhos e
280 o vírus se estabelece em outras regiões da planta suscetível, resultando em uma infecção
281 sistêmica. As partículas dos orthotospovírus formadas são acumuladas em vesículas e
282 não são secretadas pela célula vegetal devido à barreira física rígida imposta pela parede
283 celular.

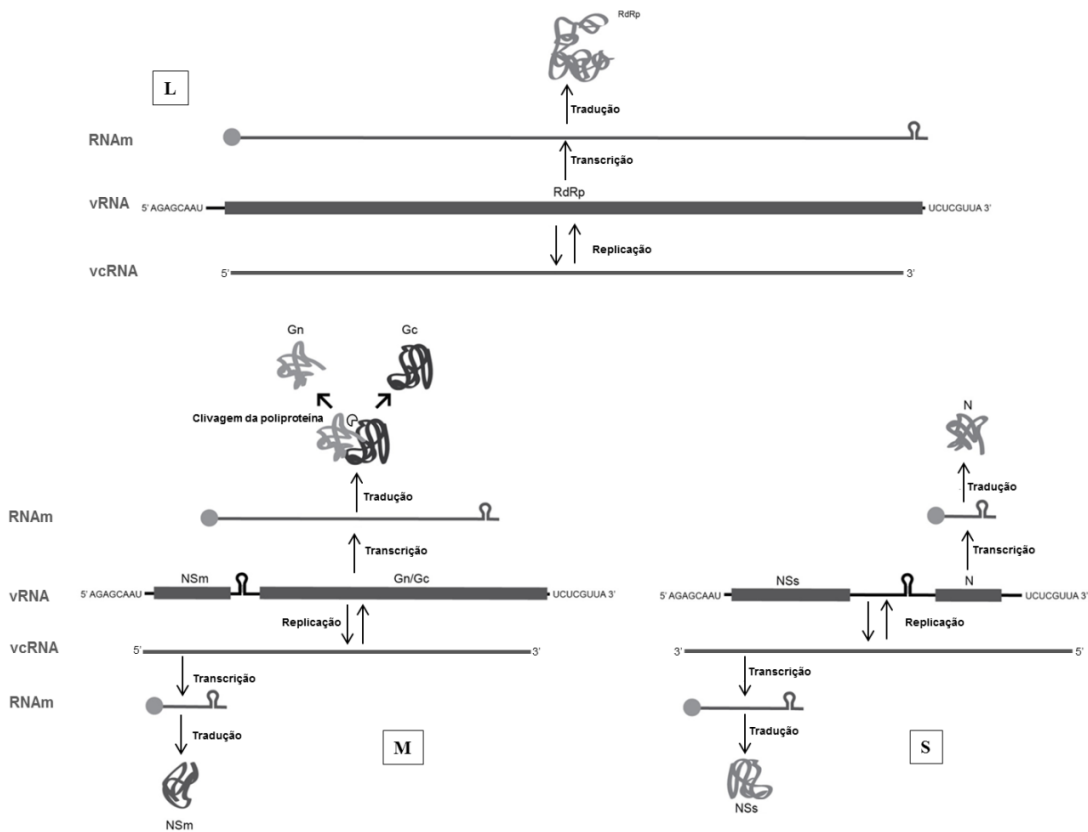
284 Os ribonucleocapsídeos são utilizados como moldes para a transcrição e
285 replicação, portanto, o RNA genômico não é infectivo se não estiver associado às
286 moléculas de NP e RdRp [40]. A proteína NP também regula o momento da passagem
287 da transcrição para replicação e vice versa, porém, o mecanismo pelo qual a RdRp
288 acessa o RNA viral encapsidado ainda é desconhecido.

289 A RdRp do orthotospovírus utiliza o mecanismo de *cap-snatching* para iniciar a
290 transcrição dos RNAm virais. A RdRp viral mobiliza o RNA mensageiro maduro
291 proveniente do núcleo da célula do hospedeiro e cliva-o entre 12 e 18 nucleotídeos
292 *downstream* à capa 5' metil-guanosina. Essas pequenas sequências são utilizadas como
293 iniciadores para a transcrição dos RNAs mensageiros do vírus [41, 42]. Esses RNAs
294 mensageiros virais não são poliadenilados, mas em suas extremidades 3'UTR existem

295 *hairpins* que conferem estabilidade e podem estar envolvidos com o término da
 296 transcrição [39, 43].

297 O segmento L possui apenas uma ORFs que codifica para a RNA polimerase
 298 viral no senso negativo; logo, é produzida uma molécula de RNA complementar de
 299 senso positivo que servirá de molde para a replicação. Os segmentos S e M possuem
 300 estratégia ambisenso: as ORFs para as proteína NS_S e NS_M estão no RNA viral
 301 complementar (vcRNA) enquanto que as ORFs para o precursor das glicoproteínas e
 302 para a NP estão presentes no RNA viral (vRNA). Após a tradução, o precursor das
 303 glicoproteínas é clivado proteoliticamente para produzir as duas glicoproteínas maduras
 304 G_N e G_C (fig.6).

305



306
 307

308 **Figura 6: Estratégias ambisenso e de senso negativo para a expressão de proteínas de**
 309 **orthotospovírus.** VRNA= RNA viral; vcRNA= fita de RNA complementar à vRNA; RNAm= RNA
 310 mensageiro viral. Os quadrados escuros representam as ORFs para as proteínas. No RNAm, o círculo
 311 cinza representa o 5'cap da região 5'UTR enquanto que na região 3'UTR está representado o *hairpin*. A
 312 semi-lua representa a protease do hospedeiro, que cliva a o precursor das glicoproteínas formando as
 313 glicoproteínas Gn e Gc. Por convenção, todos os RNAm virais estão posicionados no sentido 5'-3'.

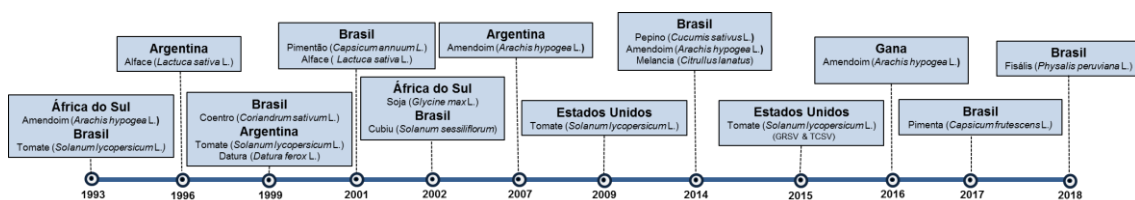
314 1.2.3 Groundnut ringspot orthotospovirus (GRSV)

315

316 As primeiras amostras de plantas infectadas com GRSV foram coletadas em 1990,
317 porém, acreditava-se que os sintomas eram causados por isolados distintos de TSWV,
318 que já era amplamente estudado na época [44]. Contudo, poucos anos depois, análises
319 moleculares desses isolados concluíram que, de fato, se tratava de espécies distintas de
320 orthotospovírus [45].

321 O GRSV foi caracterizado simultaneamente em amostras de amendoim
322 provenientes da África do Sul e de tomate provenientes do Brasil [46]. Posteriormente,
323 inúmeros relatos de infecção por GRSV foram sendo reportados em cultivares
324 economicamente importantes como alface, coentro, datura, pimentão, soja, cubiu,
325 berinjela, tomate mexicano e recentemente em melancia, pepino, pimenta e fisális [46-
326 57]. No período entre setembro de 2016 a janeiro de 2017, isolados de GRSV foram
327 detectados no Campo Experimental da Embrapa Hortalças, em soja (*Glycine max* (L.)
328 Merrill), ervilha (*Pisum sativum* L.), jiló (*Solanum aethiopicum* L. var. gilo Raddi),
329 almeirão (*Cichorium intybus* L.) e jurubeba vermelha (*Solanum stramonifolium* Jacq.)
330 [48]. Todos esses relatos de GRSV restringem-se, até o presente momento, aos
331 continentes Americano e Africano (Fig.7).

332



333

334 Figura7. Linha do tempo destacando alguns dos principais relatos de GRSV na África e nas Américas.

335

336 Em 2009, plantas de tomate com sintomas de infecção por orthotospovírus foram
coletadas na Flórida (EUA), e em 2015, identificou-se, pela primeira vez, a única

337 ocorrência de um orthospovírus pseudo-recombinante, com seu genoma composto
338 pelos segmentos L e S de GRSV e M de TCSV [47, 56, 59].

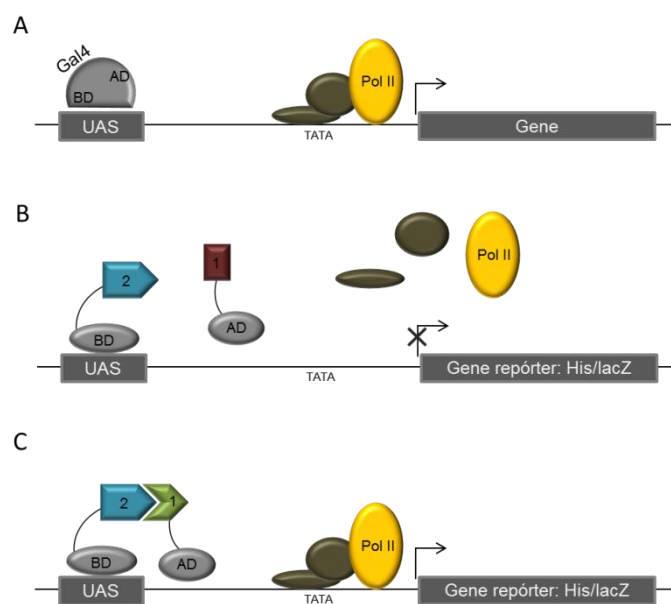
339 O GRSV não é transmitido verticalmente por semente e seus vetores são todos
340 tripes do gênero *Frankliniella*: *F. gemina*, *F. intonsa*, *F. occidentalis* e *F. schultzei* [10,
341 59, 60]. Esses vetores possuem diferentes níveis de competência para a transmissão do
342 GRSV, sendo *F. schultzei* o mais competente [56, 61]. *F. schultzei* se apresenta em duas
343 colorações, clara e escura. A forma clara se mostrou eficaz na transmissão de TCSV e
344 ineficaz na transmissão do GRSV, enquanto que a forma escura conseguiu transmitir
345 ambos os vírus [62]. Aliás, estudos morfológicos e moleculares recentes reforçam que
346 as formas escura e clara de *F. schultzei* se tratam, na verdade, de duas espécies distintas,
347 a escura é a *F. schultzei* e a clara seria renomeada para *F. sulphurea* [63]. Estudos
348 conclusivos acerca da situação taxonômica de *F. schultzei* ajudarão no processo de
349 compreensão da transmissão e do controle das infecções por GRSV e demais
350 orthospovírus no campo.

351 **1.3 Ferramentas moleculares para o estudo das interações proteicas**

352

353 As interações entre as próprias proteínas do vírus e as interações entre proteínas
354 virais e proteínas do hospedeiro são essenciais e obrigatórias para o sucesso de uma
355 infecção. Todos os vírus recrutam proteínas do hospedeiro para sua entrada, replicação,
356 tradução, evasão do sistema imune do hospedeiro e disseminação na célula, assim como
357 para a montagem do capsídeo e o movimento célula a célula [64-66]. Para a
358 identificação dessas interações proteína-proteína nos sistemas biológicos
359 orthospovírus/plantas e orthospovírus/vetores, diferentes técnicas moleculares vêm
360 sendo utilizadas, tais como: Sistema duplo híbrido em levedura (Y2H), *Pull-down* e
361 Fluorescência Bimolecular Complementar (BiFC).

362 O sistema de duplo híbrido em levedura é uma técnica para estudo, *in vivo*, da
 363 interação entre proteínas e consiste em transformar uma levedura com dois vetores de
 364 expressão de proteínas candidatas à interação com os dois domínios de ativação da
 365 proteína promotora da transcrição GAL4 (Fig.8). O domínio N-terminal da GAL4 é
 366 denominado BD (DNA *binding domain*) e o domínio C-terminal é denominado AD
 367 (*activating domain*). A proteína de estudo fundida ao BD é denominada “isca” e as
 368 proteínas da biblioteca de cDNAs fusionadas ao domínio AD são denominadas
 369 “presas”. O processo de transcrição não ocorre se as proteínas candidatas não
 370 interagem; porém, se as duas proteínas interagirem, os domínios AD e BD se
 371 aproximam, ficando funcionais e promovendo a transcrição seguida da expressão do
 372 gene repórter. O fenótipo da transcrição pode ser observado em placa de cultura com o
 373 fungo [67].



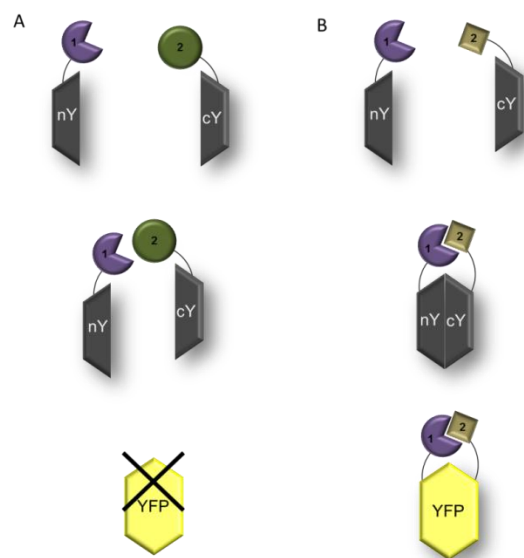
374

Figura 8. **Representação esquemática do Sistema de duplo híbrido em levedura (Y2H).** A organização esquemática mostrando o fator de transcrição Gal4 com seus domínios de ativação e ligação ao DNA para o recrutamento da RNA polimerase II (PolII); B, não recrutamento da PolII na ausência de interação proteína-proteína (1 e 2); C, a interação proteína-proteína aproxima os domínios do fator de transcrição Gal4 promovendo a transcrição.

375

376 A Fluorescência Bimolecular Complementar é uma técnica para estudo *in vivo*
377 que consiste na fusão de duas proteínas candidatas à interação a dois fragmentos não
378 fluorescentes individualmente, um N-terminal e outro C-terminal, de uma proteína
379 fluorescente (ex: GFP, EYFP) (Fig.9). Estes fragmentos são reunidos quando as duas
380 proteínas candidatas fusionadas a eles interagem-se, recuperando-se a fluorescência
381 [68].

382

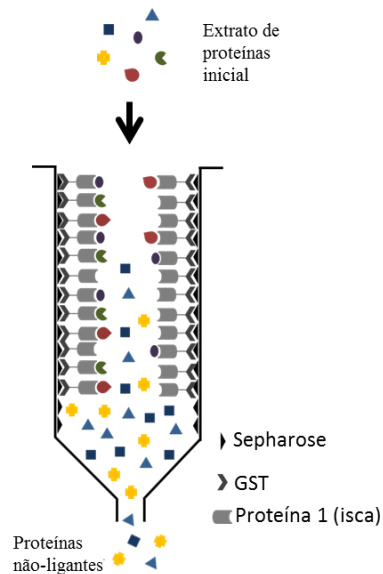


383

Figura 9. **Representação esquemática da técnica de Fluorescência Bimolecular Complementar (BiFC).** A, na ausência da interação proteína-proteína (1 e 2) os domínios N (nY) e C (cY) da molécula de YFP não emitem fluorescência; B, havendo a interação proteína-proteína (1 e 2) os domínios nY e cY são aproximados e emitem sinal fluorescente.

384

385 O *pulldown* é uma ferramenta de pesquisa *in vitro* que permite isolar as proteínas
386 das células hospedeiras que interagem com a proteína de interesse (Fig. 10). A “isca” é
387 feita fusionando-se a proteína de interesse a uma cauda de GST, HA ou cauda de 6xHis,
388 utilizando um vetor de expressão. Em seguida a proteína recombinante é purificada e
389 imobilizada em uma coluna de cromatografia específica, na qual o extrato de proteínas
390 celulares candidatas é colocado e transladado na coluna. A interação é visualizada por
391 meio de *Western blot* com os anticorpos específicos para cada proteína [69].



392

Figura 10. **Representação esquemática da técnica Pull-down.** Exemplo baseado na interação sepharose-GST. A proteína de interesse (isca) é fusionada à GST que possui afinidade pela sepharose. O extrato proteico é adicionado e lavado diversas vezes para retirar as proteínas não ligantes. Por fim, as proteínas candidatas à interação são eluídas com o uso de uma solução salina.

393

394

395 1.3.1 Interações proteína-proteína para orthospovírus

396

397

Comparado aos vírus da mesma família e a outros vírus de plantas, poucas

398

proteínas do hospedeiro, que interagem com as proteínas dos orthospovírus, foram

399

caracterizadas (Fig. 11). Uhrig et al. (1999) mostraram, por ensaios com sistema de

400

duplo híbrido em levedura, que a proteína N se polimeriza formando multímeros e

401

Richmond et al. (1998) mostraram que N se liga a moléculas de RNA [70,71]. Soellick

402

et al. (2000) demonstraram *in vitro*, por meio do sistema de duplo híbrido em levedura,

403

que a proteína de movimento NSm de *Tomato spotted wilt orthospovirus* (TSWV)

404

interage com a proteína do nucleocapsídeo NP, com moléculas de RNA fita simples e

405

com uma proteína da família DnaJ (número de acesso do *GenBank*: AF191497) [72].

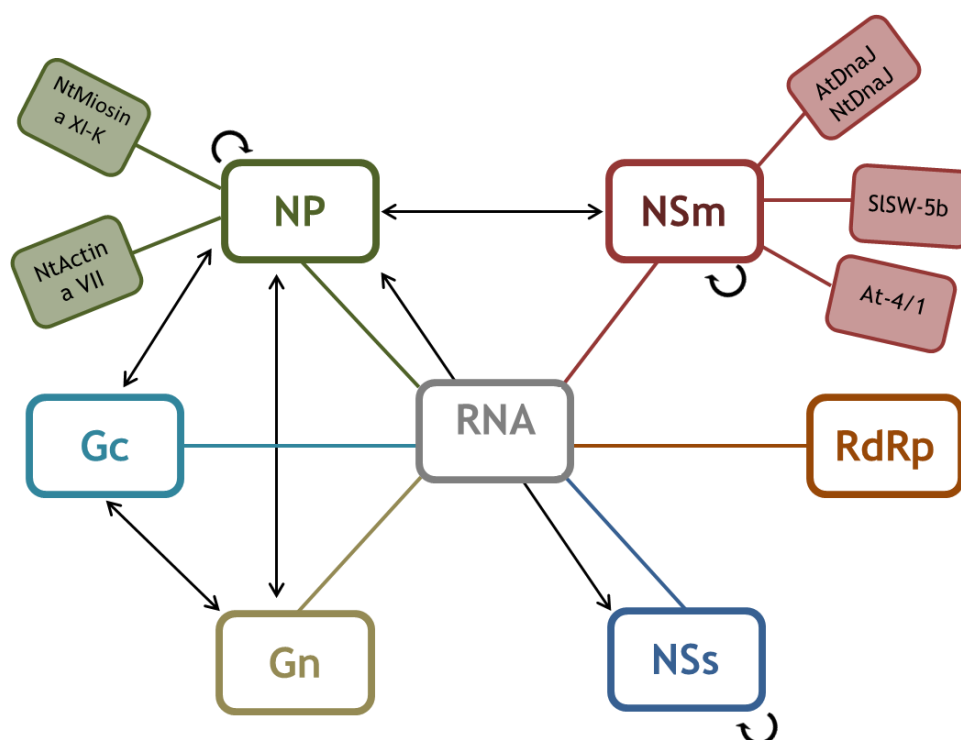
406

Essas associações são importantes para o movimento célula-a-célula do

407

ribonucleocapsídeo através dos plasmodesmas.

408 Também utilizando o sistema de duplo híbrido em levedura e microscopia
409 confocal, Paape et al. (2006) identificaram em *Arabidopsis thaliana* uma proteína
410 denominada At-4/1 (número de acesso do *GenBank*: AY212284) pertencente a uma
411 nova família de proteínas celulares, localizadas no plasmodesma, responsáveis pelo
412 movimento intra/intercelular de moléculas [73]. Essa proteína se associa à NS_M de
413 TSWV, fornecendo mais evidências para elucidar a via de dispersão viral célula-a-
414 célula.



415 **Figura 11. Representação esquemática da rede de interações RNA-proteína e proteína-**
proteína identificada para orthotospovírus. Todas as proteínas virais podem interagir com
moléculas de RNA. A seta em 270° indica homomultimerização e a seta bidirecional indica
heteromultimerização. Nt, *Nicotiana tabacum*; At, *Arabidopsis thaliana*; Sl, *Solanum*
lycopersicum.

416
417 Em dois estudos com o uso de microscopia confocal, Snippe et al. (2007) e
418 Ribeiro et al. (2009), demonstraram, *in vivo*, a co-localização entre as proteínas N e as
419 glicoproteínas Gn e Gc de TSWV, sugerindo uma possível interação entre essas
420 proteínas para a montagem das partículas virais. Porém, por BiFC, Dietzgen et al.
421 (2012) não encontraram interação entre as proteínas N, Gn e Gc de *Impatiens necrotic*

422 *spot orthospovirus* (INSV), deixando em aberto as discussões a respeito da natureza
423 dessas interações para a montagem das partículas [74-76].

424 Os ensaios de BiFC realizados por Dietzgen et al. (2012), com as proteínas de
425 INSV, não demonstraram interações proteína-proteína entre a NSs com ela mesma nem
426 com as outras proteínas virais. Também não demonstraram interações de
427 homo/heteromultimerização entre as glicoproteínas Gn e Gc ou entre as glicoproteínas e
428 as outras proteínas virais, mas confirmaram as interações NP-NP, NP-NS_M e NS_M-NS_M
429 de INSV em concordância com os trabalhos antecessores realizados com TSWV [76].

430 Recentemente, Widana Gamage & Dietzgen (2017), encontraram sublocalizações
431 e interações únicas entre as proteínas de uma espécie tentativa denominada *Capsicum*
432 *chlorosis orthospovirus* (CaCV). NS_S, G_N e G_C foram observadas no núcleo celular,
433 sendo que NSs e Gn foram observadas nesse compartimento celular pela primeira vez.
434 Assim como para TSWV, foram encontradas interações NP-G_N, NP-NP, NP-NS_M e
435 NS_M-NS_M; porém, de forma incomum, também foram observadas interações NS_S-N na
436 periferia celular e NS_S-NS_S no núcleo [77].

437 Feng et al. (2013) identificaram uma nova função de tráfego intracelular
438 independente de microtúbulo, para a proteína NP de TSWV [78]. Por meio de ensaios
439 de microscopia confocal a laser, foi observado que os agregados de proteína N estavam
440 em constante movimento pelo citoplasma. Esses agregados co-localizaram com
441 filamentos de actina, sugerindo uma interação direta entre essas duas proteínas para o
442 movimento intracelular com a atuação da miosina XI-K como proteína motora.

443 Por fim, Zhu et al. (2017) demonstraram por meio de ensaio de *pull-down* que o
444 gene de resistência Sw-5b de tomate interage fisicamente com a NS_M de TSWV [79].

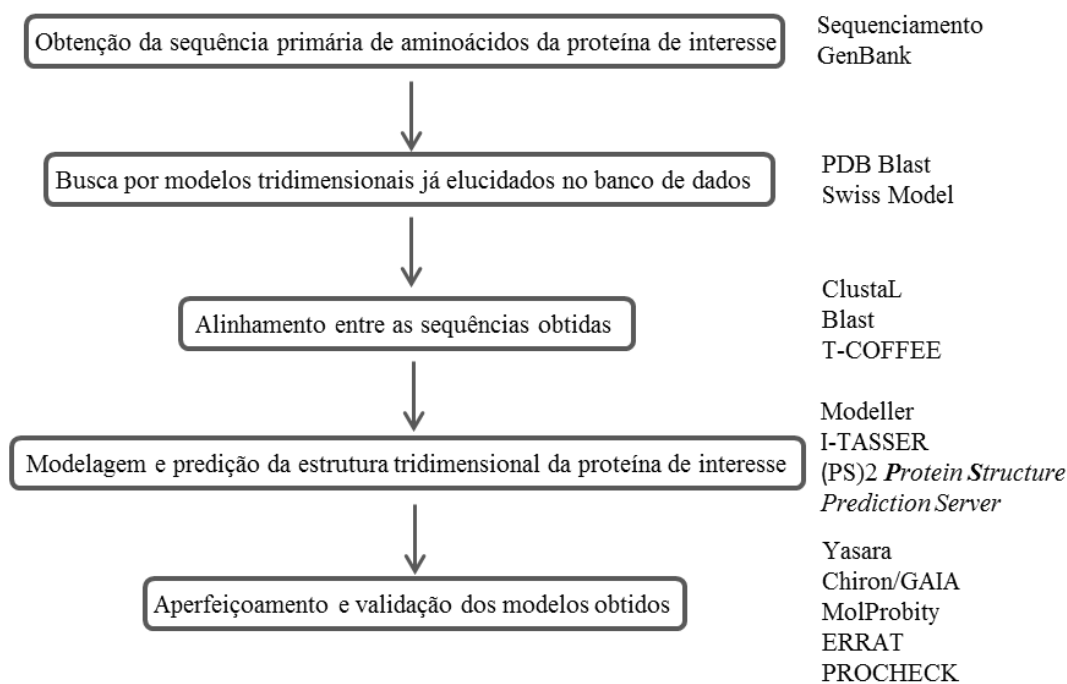
445

446 **1.4 Ferramentas para predição *in silico* da estrutura** 447 **tridimensional de proteínas** 448

449 A modelagem *in silico* de proteínas é uma importante ferramenta para obter
450 indícios da estrutura tridimensional de uma proteína não cristalizada, baseando-se em
451 uma estrutura de outra proteína já depositada no banco de dados.

452 As ferramentas de predição estrutural *in silico* têm sido determinantes para as
453 pesquisas em virologia, ajudando a caracterizar de forma rápida e barata, as possíveis
454 funções moleculares de proteínas virais distintas, e, dessa forma, reduzindo a lacuna
455 entre o sequenciamento e a identificação das funções [80-85]. Atualmente ainda
456 existem poucas proteínas com estrutura tridimensional e funções definidas, mesmo
457 diante de esforços globais para a cristalização das proteínas de diversos organismos.
458 Para os vírus em geral e principalmente para os vírus vegetais, as referências são
459 consideravelmente limitadas.

460 A modelagem computacional geralmente compreende quatro etapas principais: (1)
461 procura e seleção no banco de dados do maior número possível de estruturas de
462 proteínas já conhecidas (modelos) que são semelhantes à proteína de interesse; (2)
463 alinhamento entre a proteína de interesse e os modelos selecionados; [86] predição por
464 modelagem comparativa utilizando as variadas ferramentas de modelagem; e (4)
465 validação dos modelos gerados (Fig.12). As etapas 1 e 2 são consideradas as mais
466 críticas visto que a precisão dos modelos gerados costuma aumentar com a identidade
467 de sequência entre proteína de interesse e os modelos [87, 88].



468

Figura 12. Esquema geral para predição *in silico* das estruturas tridimensionais de proteínas. Ao lado direito algumas sugestões de programas para serem utilizados em cada etapa.

469

470

471

1.5 Referências

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

1. Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Mushegian AR *et al*: **Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017)**. *Archives of virology* 2017, **162**(8):2505-2538.
2. Smithburn KC, Haddow AJ, Mahaffy AF: **A neurotropic virus isolated from Aedes mosquitoes caught in the Semliki forest**. *The American journal of tropical medicine and hygiene* 1946, **26**:189-208.
3. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, Chen LJ, Qin XC, Xu J, Holmes EC, Zhang YZ: **Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses**. *eLife* 2015, **4**.
4. Guterres A, de Oliveira RC, Fernandes J, Maia RM, Teixeira BR, Oliveira FCG, Bonvicino CR, D'Andrea PS, Schrago CG, de Lemos ERS: **Co-circulation of Araraquara and Juquitiba Hantavirus in Brazilian Cerrado**. *Microbial ecology* 2017.
5. Travassos da Rosa JF, de Souza WM, Pinheiro FP, Figueiredo ML, Cardoso JF, Acrani GO, Nunes MRT: **Oropouche Virus: Clinical, Epidemiological, and**

- 491 **Molecular Aspects of a Neglected Orthobunyavirus.** *The American journal of*
492 *tropical medicine and hygiene* 2017, **96**(5):1019-1030.
- 493 6. Horne KM, Vanlandingham DL: **Bunyavirus-vector interactions.** *Viruses* 2014,
494 **6**(11):4373-4397.
- 495 7. Peperkamp NH, Lutikholt SJ, Dijkman R, Vos JH, Junker K, Greijdanus S,
496 Roumen MP, van Garderen E, Meertens N, van Maanen C *et al*: **Ovine and**
497 **Bovine Congenital Abnormalities Associated With Intrauterine Infection**
498 **With Schmallerberg Virus.** *Veterinary pathology* 2015, **52**(6):1057-1066.
- 499 8. He M, Guan SY, He CQ: **Evolution of rice stripe virus.** *Molecular phylogenetics*
500 *and evolution* 2017, **109**:343-350.
- 501 9. Mielke-Ehret N, Muhlbach HP: **Emaravirus: a novel genus of multipartite,**
502 **negative strand RNA plant viruses.** *Viruses* 2012, **4**(9):1515-1536.
- 503 10. Pappu HR, Jones RA, Jain RK: **Global status of tospovirus epidemics in diverse**
504 **cropping systems: successes achieved and challenges ahead.** *Virus research*
505 2009, **141**(2):219-236.
- 506 11. Junglen S, Marklewitz M, Zirkel F, Wollny R, Meyer B, Heidemann H, Metzger
507 S, Annan A, Dei D, Leendertz FH *et al*: **No Evidence of Gouleako and Herbert**
508 **Virus Infections in Pigs, Cote d'Ivoire and Ghana.** *Emerging infectious*
509 *diseases* 2015, **21**(12):2190-2193.
- 510 12. Marklewitz M, Handrick S, Grasse W, Kurth A, Lukashev A, Drosten C,
511 Ellerbrok H, Leendertz FH, Pauli G, Junglen S: **Gouleako virus isolated from**
512 **West African mosquitoes constitutes a proposed novel genus in the family**
513 **Bunyaviridae.** *Journal of virology* 2011, **85**(17):9227-9234.
- 514 13. Marklewitz M, Zirkel F, Rwego IB, Heidemann H, Trippner P, Kurth A, Kallies
515 R, Briese T, Lipkin WI, Drosten C *et al*: **Discovery of a unique novel clade of**
516 **mosquito-associated bunyaviruses.** *Journal of virology* 2013, **87**(23):12850-
517 12865.
- 518 14. Ballinger MJ, Bruenn JA, Hay J, Czechowski D, Taylor DJ: **Discovery and**
519 **evolution of bunyavirids in arctic phantom midges and ancient bunyavirid-**
520 **like sequences in insect genomes.** *Journal of virology* 2014, **88**(16):8783-8794.
- 521 15. Marklewitz M, Zirkel F, Kurth A, Drosten C, Junglen S: **Evolutionary and**
522 **phenotypic analysis of live virus isolates suggests arthropod origin of a**
523 **pathogenic RNA virus family.** *Proceedings of the National Academy of Sciences*
524 *of the United States of America* 2015, **112**(24):7536-7541.
- 525 16. Yamao T, Eshita Y, Kihara Y, Satho T, Kuroda M, Sekizuka T, Nishimura M,
526 Sakai K, Watanabe S, Akashi H *et al*: **Novel virus discovery in field-collected**
527 **mosquito larvae using an improved system for rapid determination of viral**
528 **RNA sequences (RDV ver4.0).** *Archives of virology* 2009, **154**(1):153-158.
- 529 17. Halbach R, Junglen S, van Rij RP: **Mosquito-specific and mosquito-borne**
530 **viruses: evolution, infection, and host defense.** *Current opinion in insect*
531 *science* 2017, **22**:16-27.
- 532 18. Pauvolid-Correa A, Solberg O, Couto-Lima D, Nogueira RM, Langevin S, Komar
533 N: **Novel Viruses Isolated from Mosquitoes in Pantanal, Brazil.** *Genome*
534 *announcements* 2016, **4**(6).

- 535 19. Xin M, Cao M, Liu W, Ren Y, Zhou X, Wang X: **Two Negative-Strand RNA**
536 **Viruses Identified in Watermelon Represent a Novel Clade in the Order**
537 **Bunyavirales**. *Frontiers in microbiology* 2017, **8**:1514.
- 538 20. Kuhn JH, Wiley MR, Rodriguez SE, Bao Y, Prieto K, Travassos da Rosa AP,
539 Guzman H, Savji N, Ladner JT, Tesh RB *et al*: **Genomic Characterization of the**
540 **Genus Nairovirus (Family Bunyaviridae)**. *Viruses* 2016, **8**(6).
- 541 21. Guterres A, de Oliveira RC, Fernandes J, de Lemos ER, Schrago CG: **New**
542 **bunya-like viruses: Highlighting their relations**. *Infection, genetics and*
543 *evolution : journal of molecular epidemiology and evolutionary genetics in*
544 *infectious diseases* 2017, **49**:164-173.
- 545 22. Vaheri A, Strandin T, Hepojoki J, Sironen T, Henttonen H, Makela S, Mustonen
546 J: **Uncovering the mysteries of hantavirus infections**. *Nature reviews*
547 *Microbiology* 2013, **11**(8):539-550.
- 548 23. Harris MC, Yang F, Jackson DM, Dotseth EJ, Paulson SL, Hawley DM: **La**
549 **Crosse Virus Field Detection and Vector Competence of Culex Mosquitoes**.
550 *The American journal of tropical medicine and hygiene* 2015, **93**(3):461-467.
- 551 24. Brustolin M, Talavera S, Nunez A, Santamaria C, Rivas R, Pujol N, Valle M,
552 Verdun M, Brun A, Pages N *et al*: **Rift Valley fever virus and European**
553 **mosquitoes: vector competence of Culex pipiens and Stegomyia albopicta (=**
554 **Aedes albopictus)**. *Medical and veterinary entomology* 2017.
- 555 25. Westmore GC, Poke FS, Allen GR, Wilson CR: **Genetic and host-associated**
556 **differentiation within Thrips tabaci Lindeman (Thysanoptera: Thripidae)**
557 **and its links to Tomato spotted wilt virus-vector competence**. *Heredity* 2013,
558 **111**(3):210-215.
- 559 26. Olson KE, Bonizzoni M: **Nonretroviral integrated RNA viruses in arthropod**
560 **vectors: an occasional event or something more?** *Current opinion in insect*
561 *science* 2017, **22**:45-53.
- 562 27. Badillo-Vargas IE, Rotenberg D, Schneweis DJ, Hiromasa Y, Tomich JM,
563 Whitfield AE: **Proteomic analysis of Frankliniella occidentalis and**
564 **differentially expressed proteins in response to tomato spotted wilt virus**
565 **infection**. *Journal of virology* 2012, **86**(16):8793-8809.
- 566 28. Schneweis DJ, Whitfield AE, Rotenberg D: **Thrips developmental stage-specific**
567 **transcriptome response to tomato spotted wilt virus during the virus**
568 **infection cycle in Frankliniella occidentalis, the primary vector**. *Virology*
569 2017, **500**:226-237.
- 570 29. Stafford-Banks CA, Rotenberg D, Johnson BR, Whitfield AE, Ullman DE:
571 **Analysis of the salivary gland transcriptome of Frankliniella occidentalis**.
572 *PloS one* 2014, **9**(4):e94447.
- 573 30. Zhang Z, Zhang P, Li W, Zhang J, Huang F, Yang J, Bei Y, Lu Y: **De novo**
574 **transcriptome sequencing in Frankliniella occidentalis to identify genes**
575 **involved in plant virus transmission and insecticide resistance**. *Genomics*
576 2013, **101**(5):296-305.

- 577 31. Bichaud L, de Lamballerie X, Alkan C, Izri A, Gould EA, Charrel RN:
578 **Arthropods as a source of new RNA viruses.** *Microbial pathogenesis* 2014,
579 **77**:136-141.
- 580 32. Lima RN, De Oliveira AS, Leastro MO, Blawid R, Nagata T, Resende RO, Melo
581 FL: **The complete genome of the tospovirus Zucchini lethal chlorosis virus.**
582 *Virology journal* 2016, **13**:123.
- 583 33. de Oliveira AS, Melo FL, Inoue-Nagata AK, Nagata T, Kitajima EW, Resende
584 RO: **Characterization of bean necrotic mosaic virus: a member of a novel**
585 **evolutionary lineage within the Genus Tospovirus.** *PloS one* 2012,
586 **7**(6):e38634.
- 587 34. Nagata T, Resende RdO, Kitajima EW, Costa H, Inoue-Nagata AK, de Ávila AC:
588 **First Report of Natural Occurrence of Zucchini Lethal Chlorosis Tospovirus**
589 **on Cucumber and Chrysanthemum Stem Necrosis Tospovirus on Tomato in**
590 **Brazil.** *Plant Disease* 1998, **82**(12):1403-1403.
- 591 35. Zhou J, Kantartzi SK, Wen RH, Newman M, Hajimorad MR, Rupe JC,
592 Tzanetakis IE: **Molecular characterization of a new Tospovirus infecting**
593 **soybean.** *Virus genes* 2011, **43**(2):289-295.
- 594 36. Srinivasan R, Abney MR, Culbreath AK, Kemerait RC, Tubbs RS, Monfort WS,
595 Pappu HR: **Three decades of managing Tomato spotted wilt virus in peanut in**
596 **southeastern United States.** *Virus research* 2017.
- 597 37. Rotenberg D, Jacobson AL, Schneeweis DJ, Whitfield AE: **Thrips transmission**
598 **of tospoviruses.** *Current opinion in virology* 2015, **15**:80-89.
- 599 38. Gaudiard N, Billecocq A, Flick R, Bouloy M: **Rift Valley fever virus noncoding**
600 **regions of L, M and S segments regulate RNA synthesis.** *Virology* 2006,
601 **351**(1):170-179.
- 602 39. van Knippenberg I, Goldbach R, Kormelink R: **Tomato spotted wilt virus S-**
603 **segment mRNAs have overlapping 3'-ends containing a predicted stem-loop**
604 **structure and conserved sequence motif.** *Virus research* 2005, **110**(1-2):125-
605 131.
- 606 40. Adkins S: **Tomato spotted wilt virus-positive steps towards negative success.**
607 *Molecular plant pathology* 2000, **1**(3):151-157.
- 608 41. van Knippenberg I, Goldbach R, Kormelink R: **In vitro transcription of Tomato**
609 **spotted wilt virus is independent of translation.** *The Journal of general*
610 *virology* 2004, **85**(Pt 5):1335-1338.
- 611 42. van Knippenberg I, Lamine M, Goldbach R, Kormelink R: **Tomato spotted wilt**
612 **virus transcriptase in vitro displays a preference for cap donors with**
613 **multiple base complementarity to the viral template.** *Virology* 2005,
614 **335**(1):122-130.
- 615 43. van Knippenberg I, Goldbach R, Kormelink R: **Purified tomato spotted wilt**
616 **virus particles support both genome replication and transcription in vitro.**
617 *Virology* 2002, **303**(2):278-286.
- 618 44. de Avila AC, Huguenot C, Resende Rde O, Kitajima EW, Goldbach RW, Peters
619 D: **Serological differentiation of 20 isolates of tomato spotted wilt virus.** *The*
620 *Journal of general virology* 1990, **71** (Pt 12):2801-2807.

- 621 45. de Avila AC, de Haan P, Smeets ML, Resende Rde O, Kormelink R, Kitajima
622 EW, Goldbach RW, Peters D: **Distinct levels of relationships between**
623 **tospovirus isolates.** *Archives of virology* 1993, **128**(3-4):211-227.
- 624 46. Webster CG, Reitz SR, Perry KL, Adkins S: **A natural M RNA reassortant**
625 **arising from two species of plant- and insect-infecting bunyaviruses and**
626 **comparison of its sequence and biological properties to parental species.**
627 *Virology* 2011, **413**(2):216-225.
- 628 47. Pietersen G, Morris J: **Natural Occurrence of Groundnut ringspot virus on**
629 **Soybean in South Africa.** *Plant Dis* 2002, **86**:1271.
- 630 48. Boari AJ, Maciel-Zambolim E, Lau DD, Lima GSA, Kitajima EW,
631 Brommonschenkel SH, Zerbini FM: **Detection and partial characterization of**
632 **an isolate of Groundnut ringspot virus in Solanum sessiliflorum.** *Fitopatol*
633 *Bras* 2002, **27**:249.
- 634 49. Webster CG, Perry KL, Lu X, Horsman L, Frantz G, Mellinger C, Adkins S: **First**
635 **report of Groundnut ringspot virus infecting tomato in south Florida.** *Plant*
636 *Health Prog* 2010.
- 637 50. Camelo-García VM, Lima EFB, Mansilla-Córdova PJ, Rezende JAM, Kitajima
638 EW, Barreto M: **Occurrence of Groundnut ringspot virus on Brazilian peanut**
639 **crops.** *J Gen Plant Pathol* 2014, **80**:282.
- 640 51. Leão EU, Spadotti DMA, Rocha KCG, Pantoja KFC, Rezende JAM, Pavan MA,
641 Krause-Sakate R: **Citrullus lanatus is a new natural host of Groundnut**
642 **ringspot virus in Brazil.** *J Phytopathol* 2014, **163**:1014.
- 643 52. Spadotti DMA, Leão EU, Rocha KCG, Pavan MA, Krause-Sakate R: **First report**
644 **of Groundnut ringspot virus in cucumber fruits in Brazil.** *New Dis Rep* 2014,
645 **29**:25.
- 646 53. de Breuil S, Abad JA, Nome CF, Giolitti FJ, Lambertini PL, Lenardon S:
647 **Groundnut ringspot virus: An emerging tospovirus inducing disease in**
648 **peanut crops.** *J Phytopathol* 2007, **155**:251.
- 649 54. Webster CG, Frantz G, Reitz SR, Funderburk JE, Mellinger HC, McAvoy E,
650 Turechek WW, Marshall SH, Tantiwanich Y, McGrath MT *et al*: **Emergence of**
651 **Groundnut ringspot virus and Tomato chlorotic spot virus in Vegetables in**
652 **Florida and the Southeastern United States.** *Phytopathology* 2015, **105**(3):388-
653 398.
- 654 55. Esquivel AF, Rezende JAM, Kitajima EW, Lima EFB, Diniz FO: **First report of**
655 **Groundnut ring spot virus on Physalis peruviana in Brazil.** *Plant Disease*
656 2018.
- 657 56. Webster, C. G., Turechek, W. W., Mellinger, H. C., Frantz, G., Roe, N., Yonce,
658 H., Vallad, G. E., and Adkins, S. **Expansion of Groundnut ringspot virus host**
659 **and geographic ranges in solanaceous vegetables in peninsular Florida.**
660 Online. *Plant Health Progress* 2011, doi:10.1094/PHP-2011-0725-01-BR.
- 661 57. Fontes MG: **Estudo da interação tospovírus – tomateiro: Análise**
662 **transcritômica, espectro da resistência no acesso ‘PI 203230’ e identificação**
663 **de novas hospedeiras de Groundnut ringspot virus (GRSV).** 2017.

- 664 58. Gilbertson RL, Batuman O, Webster CG, Adkins S: **Role of the Insect**
665 **Supervectors Bemisia tabaci and Frankliniella occidentalis in the Emergence**
666 **and Global Spread of Plant Viruses.** *Annual review of virology* 2015, **2(1):67-**
667 **93.**
- 668 59. Nagata T, Almeida ACL, Resende RO, DeÁvila AC: **The competence of four**
669 **thrips species to transmit and replicate four tospoviruses.** *Plant Pathology*
670 2004, **53(2):136-140.**
- 671 60. de Borbón CM, Gracia O, Píccolo R: **Relationships between Tospovirus**
672 **Incidence and Thrips Populations on Tomato in Mendoza, Argentina.** *Journal*
673 *of Phytopathology* 2006, **154(2):93-99.**
- 674 61. Webster CG, Frantz G, Reitz SR, Funderburk JE, Mellinger HC, McAvoy E,
675 Turechek WW, Marshall SH, Tantiwanich Y, McGrath MT *et al*: **Emergence of**
676 **Groundnut ringspot virus and Tomato chlorotic spot virus in vegetables in**
677 **Florida and the Southeastern United States.** *Phytopathology* 2015, **105:388.**
- 678 62. Wijkamp I, Almarza N, Goldbach R, Peters D: **Distinct levels of specificity in**
679 **thrips transmission of tospoviruses.** *Phytopathology* 1995, **85:1069-1074.**
- 680 63. Gikonyo MW, Niassy S, Moritz GB, Khamis FM, Magiri E, Subramanian S:
681 **Resolving the taxonomic status of Frankliniella schultzei (Thysanoptera:**
682 **Thripidae) colour forms in Kenya – a morphological-, biological-, molecular-**
683 **and ecological-based approach.** *International Journal of Tropical Insect Science*
684 2016, **37(2):57-70.**
- 685 64. Brizard JP, Carapito C, Delalande F, Van Dorselaer A, Brugidou C: **Proteome**
686 **analysis of plant-virus interactome: comprehensive data for virus**
687 **multiplication inside their hosts.** *Molecular & cellular proteomics : MCP* 2006,
688 **5(12):2279-2297.**
- 689 65. Pelaez P, Sanchez F: **Small RNAs in plant defense responses during viral and**
690 **bacterial interactions: similarities and differences.** *Frontiers in plant science*
691 2013, **4:343.**
- 692 66. Jin H, Zhu JK: **A viral suppressor protein inhibits host RNA silencing by**
693 **hooking up with Argonautes.** *Genes & Development* 2010, **24(9):853-856.**
- 694 67. Chen XH, Xiang HY, Wang Z, Zhang YJ, Han CG, Li DW, Yu JL, Cheng YQ:
695 **Studies on interaction of cucurbit aphid-borne yellow virus proteins using**
696 **yeast two-hybrid system and bimolecular fluorescence complementation.** *Acta*
697 *virologica* 2011, **55(3):235-241.**
- 698 68. Citovsky V, Gafni Y, Tzfira T: **Localizing protein-protein interactions by**
699 **bimolecular fluorescence complementation in planta.** *Methods (San Diego,*
700 *Calif)* 2008, **45(3):196-206.**
- 701 69. Huang YW, Hu CC, Liou MR, Chang BY, Tsai CH, Meng M, Lin NS, Hsu YH:
702 **Hsp90 interacts specifically with viral RNA and differentially regulates**
703 **replication initiation of Bamboo mosaic virus and associated satellite RNA.**
704 *PLoS pathogens* 2012, **8(5):e1002726.**
- 705 70. Uhrig JF, Soellick TR, Minke CJ, Philipp C, Kellmann JW, Schreier PH:
706 **Homotypic interaction and multimerization of nucleocapsid protein of**
707 **tomato spotted wilt tospovirus: identification and characterization of two**

- 708 **interacting domains.** *Proceedings of the National Academy of Sciences of the*
709 *United States of America* 1999, **96**(1):55-60.
- 710 71. Richmond KE, Chenault K, Sherwood JL, German TL: **Characterization of the**
711 **nucleic acid binding properties of tomato spotted wilt virus nucleocapsid**
712 **protein.** *Virology* 1998, **248**(1):6-11.
- 713 72. Soellick T, Uhrig JF, Bucher GL, Kellmann JW, Schreier PH: **The movement**
714 **protein NSm of tomato spotted wilt tospovirus (TSWV): RNA binding,**
715 **interaction with the TSWV N protein, and identification of interacting plant**
716 **proteins.** *Proceedings of the National Academy of Sciences of the United States of*
717 *America* 2000, **97**(5):2373-2378.
- 718 73. Paape M, Solovyev AG, Erokhina TN, Minina EA, Schepetilnikov MV,
719 Lesemann DE, Schiemann J, Morozov SY, Kellmann JW: **At-4/1, an interactor**
720 **of the Tomato spotted wilt virus movement protein, belongs to a new family**
721 **of plant proteins capable of directed intra- and intercellular trafficking.**
722 *Molecular plant-microbe interactions : MPMI* 2006, **19**(8):874-883.
- 723 74. Dietzgen RG, Martin KM, Anderson G, Goodin MM: **In planta localization and**
724 **interactions of impatiens necrotic spot tospovirus proteins.** *The Journal of*
725 *general virology* 2012, **93**(Pt 11):2490-2495.
- 726 75. Snippe M, Smeenk L, Goldbach R, Kormelink R: **The cytoplasmic domain of**
727 **tomato spotted wilt virus Gn glycoprotein is required for Golgi localisation**
728 **and interaction with Gc.** *Virology* 2007, **363**(2):272-279.
- 729 76. Ribeiro D, Borst JW, Goldbach R, Kormelink R: **Tomato spotted wilt virus**
730 **nucleocapsid protein interacts with both viral glycoproteins Gn and Gc in**
731 **planta.** *Virology* 2009, **383**(1):121-130.
- 732 77. Widana Gamage SMK, Dietzgen RG: **Intracellular Localization, Interactions**
733 **and Functions of Capsicum Chlorosis Virus Proteins.** *Frontiers in*
734 *microbiology* 2017, **8**:612.
- 735 78. Feng Z, Chen X, Bao Y, Dong J, Zhang Z, Tao X: **Nucleocapsid of Tomato**
736 **spotted wilt tospovirus forms mobile particles that traffic on an**
737 **actin/endoplasmic reticulum network driven by myosin XI-K.** *The New*
738 *phytologist* 2013.
- 739 79. Zhu M, Jiang L, Bai B, Zhao W, Chen X, Li J, Liu Y, Chen Z, Wang B, Wang C
740 *et al*: **The Intracellular Immune Receptor Sw-5b Confers Broad-spectrum**
741 **Resistance to Tospoviruses through Recognition of a Conserved 21-amino-**
742 **acid Viral Effector Epitope.** *The Plant cell* 2017.
- 743 80. Deeba F, Malik MZ, Naqvi IH, Haider MSH, Shafat Z, Sinha P, Ishrat R,
744 Ahmed A, Parveen S: **Potential entry inhibitors of the envelope protein (E2)**
745 **of Chikungunya virus: in silico structural modeling, docking and molecular**
746 **dynamic studies.** *Virusdisease* 2017, **28**(1):39-49.
- 747 81. Joseph J, Bhaskaran R, Kaliraj M, Muthuswamy M, Suresh A: **Molecular**
748 **Docking of Phytoligands to the viral protein receptor P. monodon Rab7.**
749 *Bioinformatics* 2017, **13**(4):116-121.
- 750 82. Ahmad N, Farman A, Badshah SL, Ur Rahman A, Ur Rashid H, Khan K:
751 **Molecular modeling, simulation and docking study of ebola virus**
752 **glycoprotein.** *Journal of molecular graphics & modelling* 2017, **72**:266-271.

- 753 83. Leyson CLM, Jordan BJ, Jackwood MW: **Insights from molecular structure**
754 **predictions of the infectious bronchitis virus S1 spike glycoprotein.**
755 *Infection, genetics and evolution : journal of molecular epidemiology and*
756 *evolutionary genetics in infectious diseases* 2016, **46**:124-129.
- 757 84. Ekins S: **Illustrating and homology modeling the proteins of the Zika virus.**
758 2016, **5**.
- 759 85. Lima RN, Faheem M, Barbosa J, Polêto MD, Verli H, Melo FL, Resende RO:
760 **Homology modeling and molecular dynamics provide structural insights**
761 **into tospovirus nucleoprotein.** *BMC Bioinformatics* 2016, **17**(Suppl 18).
- 762 86. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ,
763 Murray LW, Richardson JS, Richardson DC: **MolProbity: all-atom structure**
764 **validation for macromolecular crystallography.** *Acta crystallographica*
765 *Section D, Biological crystallography* 2010, **66**(Pt 1):12-21.
- 766 87. Dunbrack RL, Jr.: **Sequence comparison and protein structure prediction.**
767 *Current opinion in structural biology* 2006, **16**(3):374-384.
- 768 88. Xiang Z: **Advances in homology protein structure modeling.** *Current protein*
769 *& peptide science* 2006, **7**(3):217-227.
770

CAPÍTULO II: Homology modeling and molecular dynamics provide structural insights into orthospovirus nucleoprotein

Este capítulo foi publicado de forma modificada na revista *BMC Bioinformatics*.
Lima RN, Faheem M, Barbosa J, Polêto MD, Verli H, Melo FL, Resende RO:
**Homology modeling and molecular dynamics provide structural insights into
tospovirus nucleoprotein.** *BMC Bioinformatics* 2016, **17**(Suppl 18).

771 **Homology modeling and molecular dynamics provide**
772 **structural insights into orthospovirus nucleoprotein**
773

774

775 Rayane Nunes Lima^{1¶}; Muhammad Faheem^{2,3¶}; João Alexandre Ribeiro Gonçalves
776 Barbosa^{2,3}; Marcelo Depólo Polêto⁴; Hugo Verli⁴; Fernando Lucas Melo¹; Renato
777 Oliveira Resende^{1*}.

778

779 ¹ Laboratório de Virologia Vegetal. Departamento de Biologia Celular, Instituto de Ciências Biológicas,
780 Universidade de Brasília, Brasília, DF, Brazil;

781 ² Laboratório de Biofísica. Departamento de Biologia Celular, Instituto de Ciências Biológicas,
782 Universidade de Brasília, Brasília, DF, Brazil;

783 ³ Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, DF, Brazil.

784 ⁴ Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

785

786

787

788 *Corresponding author

789 Email: rresende@unb.br

790

791 ¶These authors contributed equally to this work.

792 **2.1 Abstract**

793 **2.1.1 Background**

794 *Orthospovirus* is a plant-infecting genus within the order *Bunyvirales*, which also
795 includes four animal-infecting genera: *Orthohantavirus*, *Orthonairovirus*, *Phlebovirus*
796 and *Orthobunyavirus*. Compared to these members, the structures of *orthospovirus*
797 proteins still are poorly understood. Despite multiple studies have attempted to identify
798 candidate NP protein regions involved in RNA binding and protein multimerization for
799 *orthospovirus* using yeast two-hybrid systems (Y2HS) and site-directed mutagenesis,
800 the *orthospovirus* ribonucleocapsids (RNPs) remains largely uncharacterized at the
801 molecular level and the lack of structural information prevents detailed insight into
802 these interactions.

803 **2.1.2 Results**

804 Here we used the nucleoprotein structure of LACV (*La Crosse virus-Orthobunyavirus*)
805 and molecular dynamics simulations to access the structure and dynamics of the
806 nucleoprotein from *orthospovirus* GRSV (*Groundnut ringspot orthospovirus*). The
807 resulting model is a monomer composed by a flexible N-terminal and C-terminal arms
808 and a globular domain with a positively charged groove in which RNA is deeply
809 encompassed. This model allowed identifying the candidate amino acids residues
810 involved in RNA interaction and NP-NP multimerization. Moreover, most residues
811 predicted to be involved in these interactions are highly conserved among *ospoviruses*.

812 **2.1.3 Conclusions**

813 Crucially, the interaction model proposed here for GRSV NP is further corroborated by
814 the all available mutational studies on TSWV (*Tomato spotted wilt orthospovirus*)
815 NP, so far. Our data will help designing further and more accurate mutational and

816 functional studies of orthospovirus NP proteins. In addition, the proposed model may
817 shed light on the mechanisms of RNP shaping and could allow the identification of
818 essential amino acid residues as potential targets for orthospovirus control strategies.

819 **2.2 Background**

820

821 *Orthospovirus* is a thrips-borne plant-infecting genus within the order
822 *Bunyavirales*, which also includes four animal-infecting genera: *Hanta/Nairo/Phlebo-*
823 *and Orthobunyavirus* [1]. GRSV (*Groundnut ringspot orthospovirus*) is an emerging
824 orthospovirus, that has caused severe diseases in distinct vegetable crops in South
825 America and is phylogenetically close to the orthospovirus type-species TSWV
826 (*Tomato spotted wilt orthospovirus*) [2]. Like all orthospovirus, GRSV contain a
827 trisegmented negative single-stranded RNA (ssRNA) genome that encodes the viral
828 RNA-dependent RNA polymerase (RdRp), two glycoproteins (Gn/Gc), the movement
829 protein (NSm), the RNA silencing suppressor protein (NSs) and the nucleoprotein (NP)
830 [3]. NP is a multifunctional protein involved in RNA protection, particle assembly,
831 intracellular movement and might play a role in transcription/replication regulation [4-
832 14]. Multiple copies of the NP protein form oligomers that interact with the viral RNAs
833 to build ribonucleoprotein complexes (RNPs) that are proposed to be transported via
834 plasmodesmata and are functional templates for RNA replication and transcription [6,
835 15, 16].

836 Multiple studies have attempted to identify candidate NP protein regions involved
837 in RNA binding and protein multimerization for TSWV using yeast two-hybrid systems
838 (Y2HS) and site-directed mutagenesis [4, 6, 17, 18], but the orthospovirus RNPs
839 remains largely uncharacterized at the molecular level and the lack of structural
840 information prevents detailed insight into these interactions. The lack of a reverse
841 genetics system, which is available for other bunyaviruses, has hampered tospovirus
842 research. The NP protein crystal structures of related RNA virus families
843 (*Arena/Orthomyxo/Hantaviridae/Nairoviridae/Peribunyaviridae/Phenuiviridae*) have
844 been elucidated [8, 19-26] and despite different size and distinct NP-folding structures,

845 there are common features and architectural principles by which these proteins form
846 NP-NP multimers and NP–RNA complexes [27]. Therefore, these available structures
847 were used to predict a three-dimensional model for GRSV NP (the most important and
848 prevalent tospovirus in Brazil) using homology modeling.

849 **2.3 Results and Discussion**

850 **2.3.1 Three-dimensional model of GRSV NP and Oligomerization**

851

852 The GRSV NP and LACV NP have similar protein fold with the predicted GRSV
853 NP monomer forming thirteen helical segments and two small beta-sheets (Figs. 1, 2A
854 and 2E-F). The protein has a globular core domain (26-223 aa) containing a deep
855 positively charged groove with the two chain terminals forming an N-terminus arm (1-
856 25 aa) and a C-terminus arm (224-258 aa) (Fig. 2A-B and Fig.3). The N- and C-arms
857 extend outwards from the globular core domain and interacts with the globular core
858 domain of neighboring monomers to mediate the multimerization, supporting the “head-
859 to-tail” model proposed by [18].

860 Amino acids S2-V12 of the N-arm interact with the Q61-N82 of the core domain
861 of one neighboring monomer (Figs. 2C and 2E) while K227-K249 of the C-arm interact
862 with the K173-K198 of the core domain the other neighboring monomer (Figs. 2D and
863 2F). Specific residue-residue interactions have been listed in Table 1 for the two
864 independent interfaces. According to PISA, the intermolecular interactions were mainly
865 hydrogen bonds, but van der Waals and hydrophobic interactions also contribute to hold
866 the monomers together (data not shown). This interaction model is further corroborated
867 by the available mutational studies on TSWV NP [4, 17, 18].

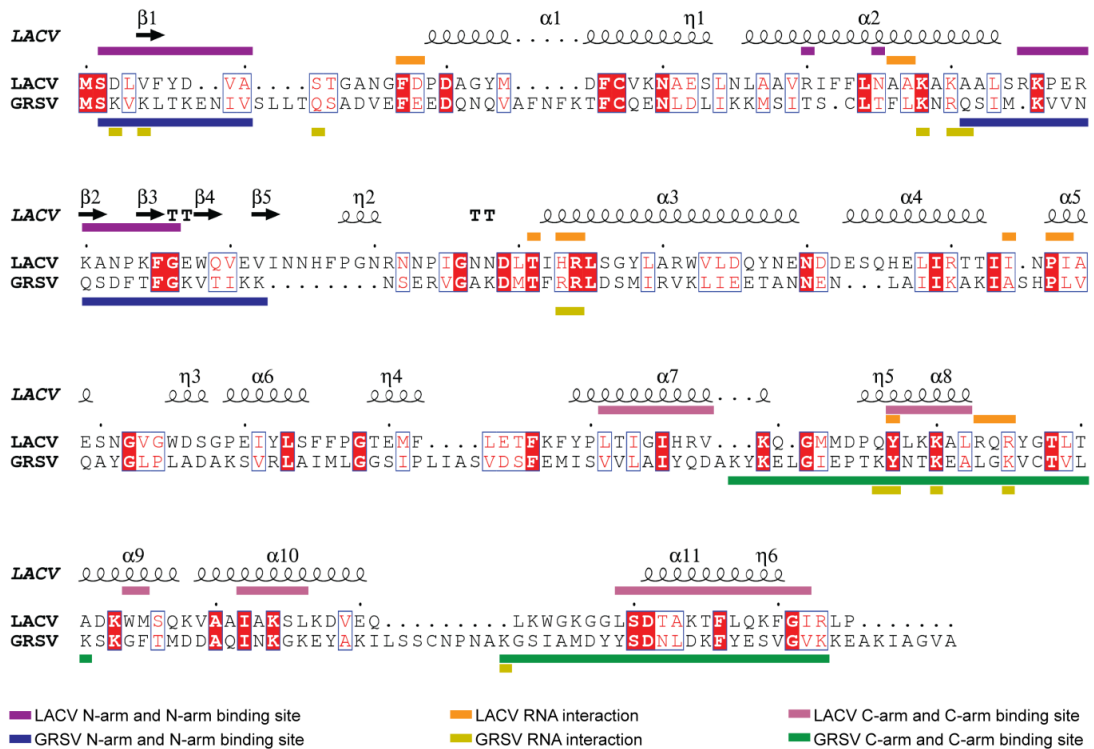


Figure 1. *Groundnut ringspot virus* (GRSV) and *La crosse virus* (LACV) Nucleoproteins sequence alignment. Key residues for GRSV NP and LACV NP oligomerization and for ssRNA binding are colored as indicated by the colored bars. The secondary structure of LACV NP is shown above, and every 10 residues are indicated with a dot (.). Strictly conserved residues are highlighted in red with white letter and highly conserved residues are displayed by red letters. GRSV NP GenBank accession number is AF251271 and LACV NP UniProt accession code is P04873.

868

869 Table 1. Pairs of interacting residues for GRSV NP-NP oligomerization.

N-arm ^a	N-arm binding site ^b	C-arm ^c	C-arm binding site ^b
S2 ^d	S83	A226	K183
V4	N82	S229	K183
T7	Q61	D233	T186
T7	S62	Y235	K183
K8	S62	N238	K198
N10	T73	Y243	N185
V12	T73	V246	K175
V12	G75	V248	K173
		K249	Y174

^a N-arm amino acids residues of GRSV NP

^b Interacting amino acids residues of GRSV NP globular core domain

^c C-arm amino acids residues of GRSV NP

^d Amino acids residues position in the GRSV NP sequence.

870 Actually, the first assay to map functional domains of TSWV NP, performing
871 Y2HS and random serial deletions, showed that both the N- (1-39 aa) and C-terminals
872 (233-248 aa) were important for NP-NP interaction [18], in clear agreement with the
873 structural results presented here. Furthermore, [17] identified three crucial
874 intermonomer binding regions: 42-56, 132-152 and 222-248 which have a clear
875 correspondence with the predicted interaction residues of GRSV NP located at N- and
876 C-arms, or buried in the core of the model (Fig. 3). Moreover, amino acids residues
877 located at the regions K103-A119 and L132-V135 are solvent accessible and therefore
878 are able to interact with NSm, glycoproteins, viral polymerase or host proteins [6, 7].
879 Recently, studies have been performed attempting to identify N-NSm interactions [28,
880 29] which results are in perfect congruence with the GRSV NP protein model. In both
881 cases, the model proposed here represents an efficient tool to assist in planning
882 experiments with mutations and deletion in the NP protein.

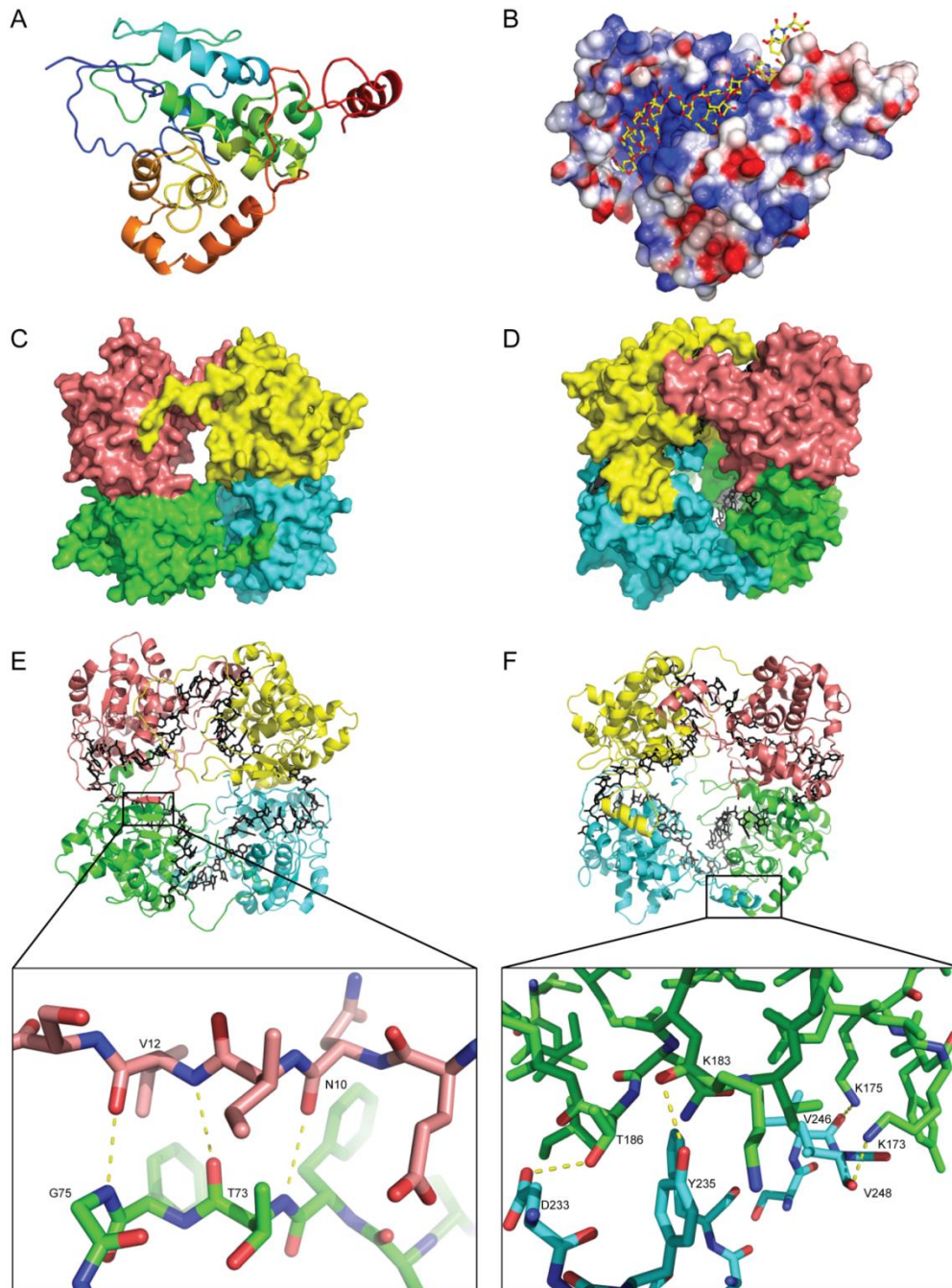


Figure 2. **Monomeric and tetrameric structure of the *Groundnut ringspot virus* (GRSV) nucleoprotein (NP).** (A) Cartoon representation of monomeric GRSV NP with rainbow coloring from N- (blue) to C-terminus (red). (B) Electrostatic surface of the GRSV NP with a positively charged groove in complex with RNA shown as yellow (carbons) and red (oxygens) sticks. Positive and negative charges are blue and red, respectively. (C) N-terminus interaction surface representation of four GRSV N monomers A, B, C, D shown in color pink, yellow, cyan and green, respectively. (D) C-terminus interaction surface representation of the GRSV NP tetrameric ring. The RNA is shown in black sticks deeply bound inside the tetrameric ring. (E) Cartoon representation with the N-arm oligomerization interface showing interacting residues. The N-terminal arm is in pink and the globular region is in green. The intermolecular hydrogen bonds are shown as yellow dotted lines. (F) 180° rotation of Figure 1E, C-arm oligomerization interface showing interacting residues. The C-terminal arm is in cyan and the globular region is in green.

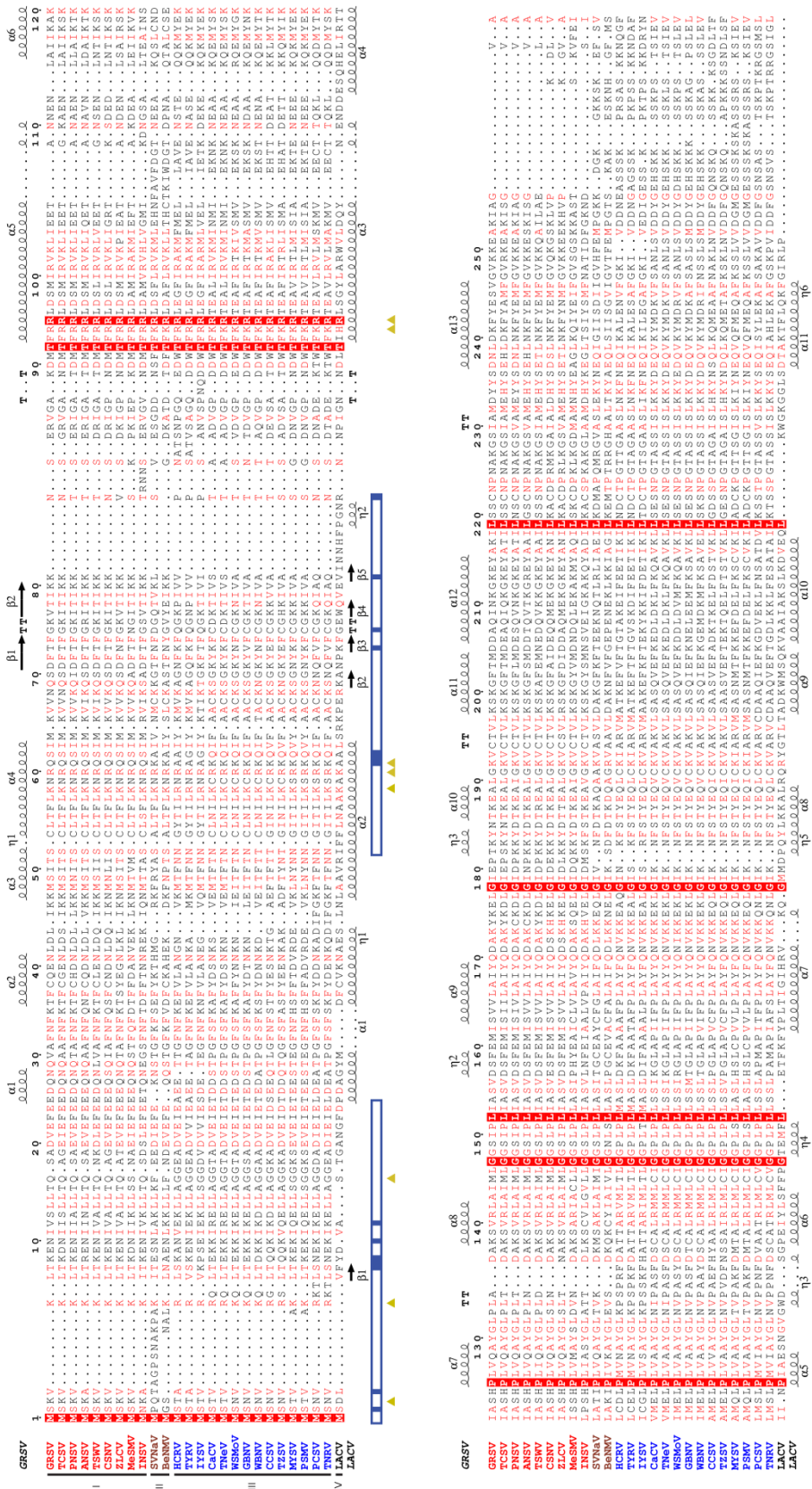


Figure 3. Sequence alignment of representative orthotospoviruses Nucleoproteins (NP). The secondary structure of *Groundnut ringspot virus* (GRSV) is shown above and of *La crosse virus* (LACV) is shown at the bottom. Key residues for GRSV NP ssRNA binding are marked with yellow triangles. GRSV NP and C-arms are marked with blue and green boxes respectively, with key residues for oligomerization highlighted. Strictly conserved residues are highlighted in red with white letter and highly conserved residues with red letter. I: *Tospovirus* American clade I; II: *Tospovirus* American clade II; III: *Tospovirus* Eurasian clade; IV: *Orthobunyavirus*. The sequence codes are supplied at the Additional file 1.

883 In addition, the obtained model for NP protein was submitted to molecular
 884 dynamics simulations in order to both refine the structure in aqueous solvent [30, 31]
 885 and access the protein conformational ensemble, further exploring its structural and
 886 functional roles. During the simulation time, the globular core domain did not reveal
 887 any loss of secondary structure, increase of radius of gyration or persistent increments
 888 on RMSD values, which supports the model quality. It is worthy to mention that RMSF
 889 calculations indicate the N-terminal arm (1-25 aa) as a very flexible region (Fig.4C).

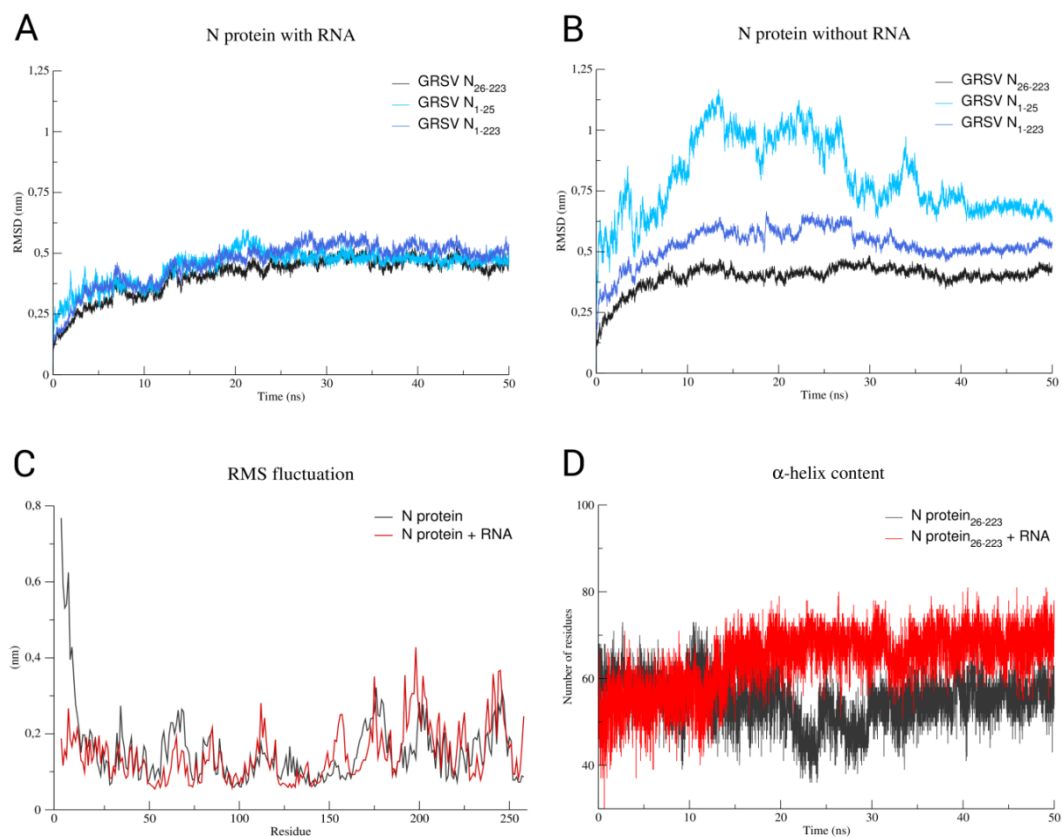


Figure 4. **Molecular Dynamics of monomeric Nucleoprotein (NP) of *Groundnut ringspot orthotospovirus* (GRSV).** Root Mean Square Deviation (RMSD) calculations for different set of atoms in both presence (A) and absence (B) of RNA. (C) Root Mean Square Fluctuations (RMSF) calculations for the entire NP in both presence (red) and absence (black) of RNA. (D) Plot of α -helix content as function of time in both presence (red) and absence (black) of RNA.

890 2.3.2 RNA interaction

891

892 According to the GRSV NP protein model, the RNA is primarily bound at the
 893 central RNA-binding groove (Fig. 2B), and the key residues for this interaction (K3,

894 K5, Q17, K58, R60, Q61, R94, R95, K183, Y184, K187, K192 and K227) are mainly
895 located in this positively charged groove. This positively charged groove is only
896 possible because residues F37, F56, F72, F74, I79, M91, F93 and L96 form a
897 hydrophobic core, which is indispensable to stabilize the protein folding and to correctly
898 orient the RNA interacting residues towards the groove. Importantly, these residues are
899 highly conserved among all tospoviruses (Fig. 3). Note that the N-terminal arm is also
900 involved in RNA binding and shielding RNA from the solvent (Figs. 2C-D). Residues
901 F23, L54, F56, L57 and F93 were observed to modulate the RNA nucleobases dynamics
902 during the performed simulation, while the N-terminal arm seems to play a stabilization
903 role during MD simulations of GRSV NP protein (Fig. 4A and B). In addition, the
904 content of alpha-helices in GRSV NP protein bound to RNA increased 25% during the
905 simulation in comparison to the free monomer (Fig. 4D), suggesting that, in the
906 simulated timescale, the monomeric state does not present a lack of conformational
907 stability in detriment of oligomeric states, as observed experimentally for other viruses
908 [32, 33].

909 Recently, the residues R60, R94, and R95 were confirmed to interact with RNA
910 [33], which also supports our results. RNA is strongly bent at each N–N interface and is
911 largely solvent-inaccessible in the tetramer (Fig. 2D). The dimensions of the groove can
912 accommodate ssRNA and PISA analysis showed that the majority of residue-nucleotide
913 interactions occur with the ribose and the phosphate moieties, suggesting a non-
914 sequence-specific RNA interaction. Indeed, Richmond et al. [4] carried out mutagenesis
915 and gel shift assay studies to identify NP regions important for ssRNA binding and
916 demonstrated that the NP-RNA complex is highly stable and non-sequence-specific,
917 further supporting these results.

918 **2.4 Conclusions**

919 Taken together, these data will help designing further and more accurate
920 mutational and functional studies of tospovirus NP proteins. In addition, the proposed
921 model may shed light on the mechanisms of RNP shaping and could allow the
922 identification of essential amino acid residues as potential targets for tospovirus control
923 strategies.

924 **2.5 Methods**

925 **2.5.1 *In silico* Homology modeling and Model optimization**

926

927 A template for modeling the GRSV NP was searched in expasy SWISS-MODEL
928 server [34] using the amino acid sequence of GRSV NP as a reference. Template crystal
929 structures of *Orthobunyavirus* genus were chosen due to their genetic relationship. The
930 LACV (*La Crosse virus-Orthobunyavirus*) NP tetrameric crystal structure in complex
931 with ssRNA (PDB ID 4BHH) was selected as the template [20], aligned with GRSV NP
932 using T-Coffee server [35] and the resulting alignment was manually improved using
933 BioEdit [36]. Aligned sequences were used with MODELLERv9.10 [37] to develop
934 high quality tetrameric models along with or without RNA.

935 Optimization of the models was achieved using energy minimization protocols
936 available at Yasara [38] and Chiron [39] servers. Quality of the 3D models were
937 evaluated with ERRAT (version 2.0) [40] and MOL probity [41]. Ramachandran plots
938 for the models were assessed and Ramachandran outlier residues were fixed with COOT
939 [42] and energy minimization. The highest quality model with 90.1% residues in
940 favored region and 8.4% in allowed region while 1.5% outlier at Ramachandran plot
941 was selected after visual inspection (see Additional file 2). The model was subjected to

942 the PISA program [43] for interface analysis at EBI-EMBL server and the retrieved
943 PISA data was analyzed for binding patterns using PyMOL [44].

944 **2.5.2 Molecular dynamics**

945 Molecular dynamics techniques were applied using GROMACS suite [45] in
946 order to evaluate the stability and consistency of the obtained NP monomeric model and
947 investigate GRSV NP- RNA interactions over time. Therefore, NP model was simulated
948 in the presence and absence of the modeled RNA, in two analytical systems.
949 Amber99SB-ILDN force field [46] was used to generate proper topologies. The models
950 were placed at the center of a dodecahedral box and solvated with TIP3P water model
951 [47]. Counterions were used to neutralize the net charge of the system, and 0.15M of
952 NaCl was added to the box in order to simulate cellular ionic environment.

953 After a minimization protocol using steepest descent and conjugate gradient to
954 eliminate possible clashes and bad contacts, NVT ensemble with restraint forces of
955 1000 kJ/mol was carried for 4 ns at 300 K. Moreover, five subsequent equilibration
956 steps in NPT ensemble were carried out at 1 bar with restraint forces of 800 kJ/mol on
957 heavy atoms, 600 Kcal/(mol x nm) and 400 kJ/mol on mainchain, 200 kJ/mol on
958 backbone and 100 kJ/mol on alpha-carbons, totalizing 13 ns. Finally, production runs
959 with no restraints were carried for 50 ns using an integration step of 2 fs and LINCS
960 algorithm [48]. Also, Particle Mesh Ewald method [49] was applied for Coulombic and
961 Lennard-Jones interactions longer than 1 nm.

962

963

964 **2.6 Additional files:**

965 **Additional file 1: The sequence codes of the viruses used at this work.**

966

Acronym	Name	Accession number ¹
ANSV	Alstroemeria necrotic streak virus	GQ478668 ^a
BeNMV	Bean necrotic mosaic virus	NC_018071 ^a
CaCV	Capsicum chlorosis virus	NC_008301 ^a
CCSV	Calla lily chlorotic spot virus	AY867502 ^a
CSNV	Chrysanthemum stem necrosis virus	AF067068 ^a
GBNV	Groundnut bud necrosis virus	NC_003619 ^a
GRSV	<i>Groundnut ringspot virus</i>	AF251271 ^a
HCRV	Hippeastrum chlorotic ringspot virus	KC290943 ^a
INSV	<i>Impatiens necrotic spot virus</i>	D00914 ^a
IYSV	<i>Iris yellow spot virus</i>	JQ973066 ^a
LACV	<i>La crosse virus-Orthobunyavirus</i>	P04873 ^b
MSMV	Melon severe mosaic virus	EU275149 ^a
MYSV	Melon yellow spot virus	NC_008300 ^a
PCSV	Pepper chlorotic spot virus	KF383956 ^a
PNSV	Pepper necrotic spot virus	HE584762 ^a
PSMV	Physalis severe mottle virus	AB038343 ^a
SVNaV	Soybean vein necrosis-associated virus	HQ728387 ^a
TCSV	<i>Tomato chlorotic spot virus</i>	JX244197 ^a
TNeV	Tomato necrosis virus	AY647437 ^a
TNRV	Tomato necrotic ringspot virus	FJ946835 ^a
TSWV	<i>Tomato spotted wilt virus</i>	JF808217 ^a
TYRV	Tomato yellow ring virus	DQ810195 ^a
TZSV	Tomato zonate spot virus	NC_010489 ^a
WBNV	<i>Watermelon bud necrosis virus</i>	EU249351 ^a
WSMoV	<i>Watermelon silver mottle virus</i>	NC_003843 ^a
ZLCV	<i>Zucchini lethal chlorosis virus</i>	KU641380 ^a

967 ^aGenBank accession number

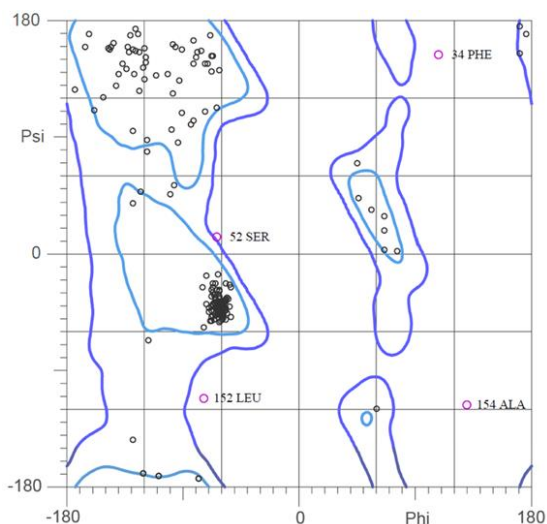
968 ^bUniProt accession number

969

970 **Additional file 2: Ramachandran plot analysis of predicted structure of**
 971 ***Groundnut ringspot orthospovirus (GRSV) NP protein.***

972 The regions covered by light blue lines show most favored regions, while the regions covered by dark
 973 blue lines show allowed regions. Other regions of the plot show the disallowed region. The pink dots
 974 show the outliers.

975



976

977 2.7 Considerações.

978 A técnica de Modelagem Molecular por Homologia Estrutural fornece inferências
979 para o estudo da estrutura tridimensional e das funções de uma proteína. Tais
980 inferências são feitas a partir da análise do nível de conservação dos domínios proteicos
981 e das propriedades físico-químicas dos aminoácidos individuais.

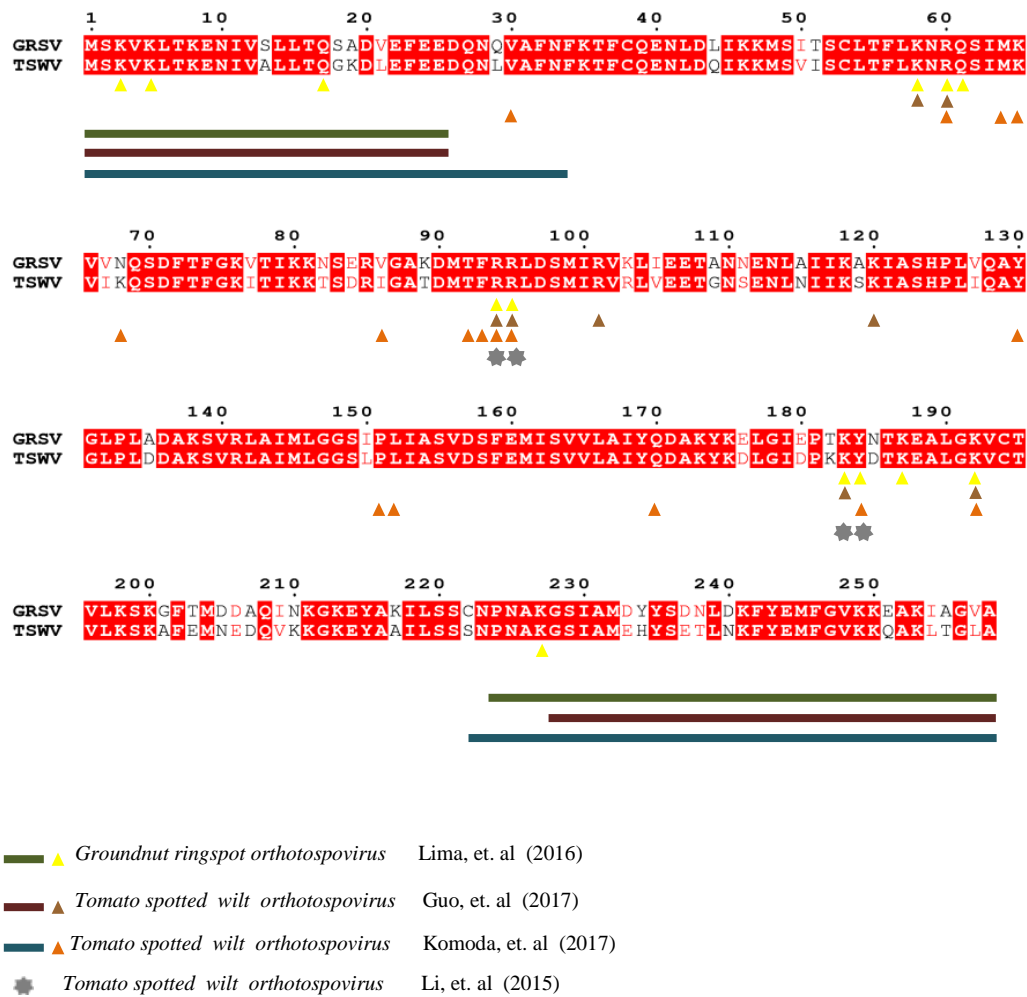
982 As pressões seletivas moldam a estrutura e as funções das proteínas virais, e não é
983 surpresa que os vírus utilizem uma variedade de estratégias durante a sua evolução. Um
984 exemplo é a interação proteína-proteína para formar oligômeros a fim de ampliar a
985 diversidade dos arranjos proteicos e com isso realizar o maior número de funções
986 possíveis. Em essência, o vírus codifica proteínas pequenas que podem se oligomerizar
987 em grandes estruturas tridimensionais dinâmicas; essas estruturas, então, interagem com
988 outros componentes virais e/ou celulares para constituir complexos com funções ainda
989 mais sofisticadas. Uma possível vantagem do uso desses complexos oligoméricos é
990 ajudar a limitar o tamanho do genoma, pois o vírus investe em uma única proteína
991 multifuncional e elimina a necessidade de se codificar diversas proteínas.

992 As proteínas do nucleocapsídeo (NP) são um exemplo notável de como uma
993 pequena proteína viral pode atuar por meio de uma interação oligomérica para formar
994 estruturas complexas e variadas. Elas são únicas para os vírus e sua montagem envolve
995 geometrias bem diversificadas [50]. As NPs dos orthotospovírus possuem múltiplas
996 funções, das quais se destacam a interação com RNA e o tráfego via filamentos de
997 actina [4, 7, 18]. Li *et. al* (2015) demonstrou que as NP de TSWV são encontradas
998 formando diferentes estados oligoméricos: monômeros, dímeros, trímeros, tetrâmeros,
999 pentâmeros e hexâmeros [33], mas estudos ainda são necessários para identificar se há
1000 alguma correlação entre determinadas funções e estados oligoméricos específicos.

1001 A análise da estrutura tridimensional da NP de *Groundnut ringspot*
1002 *orthotospovirus* (GRSV) demonstrou que a proteína possui três grandes domínios: N-
1003 terminal (aminoácidos 1-25), C-terminal (aminoácidos 224-258) e um domínio globular
1004 (aminoácidos 26-223) que apresenta uma cavidade para interação com ácidos nucleicos.
1005 Além disso, foram identificados resíduos de aminoácidos que interagem com RNA fita
1006 simples: K3, K5, Q17, K58, R60, Q61, R94, R95, K183, Y184, K187, K192 e K227
1007 [51]. Li *et. al* (2015) demonstrou experimentalmente, por meio de ensaios de
1008 mutagênese, que substituir os resíduos de aminoácidos R94, R95, K183 e Y184 por
1009 alanina, diminuiu drasticamente a afinidade da NP de TSWV por RNA [33].

1010 Posteriormente, dois trabalhos elucidaram, independentemente, a estrutura
1011 tridimensional da NP de *Tomato spotted wilt orthotospovirus* (TSWV) [52, 53], o que
1012 ajudou a fundamentar os dados obtidos para GRSV (Fig.5). Juntos, os três estudos
1013 concluíram que os resíduos R60, R94, R95 e K192 interagem com ácidos nucleicos
1014 (Fig.5); porém, nota-se que Guo *et. al* (2017) e Komoda *et. al* (2017) identificaram
1015 resíduos distintos para interação com os ácidos nucleicos (Fig.5). Há também uma
1016 pequena divergência entre as extremidades N e C terminais propostas. Guo *et. al* (2017)
1017 propôs três domínios: N-terminal (aminoácidos 1-25), C-terminal (aminoácidos 228-
1018 258) e um domínio globular (aminoácidos 26-227) e Komoda *et. al* (2017) propôs três

1019 domínios: N-terminal (aminoácidos 1-32), C-terminal (aminoácidos 222-258) e um
 1020 domínio globular (aminoácidos 26-221) (Fig.5).
 1021



1022

1023

Figura 5. Representação esquemática para comparação dos dados obtidos acerca da estrutura tridimensional das proteínas do nucleocapsídeo (NP) de *Groundnut ringspot orthotospovirus* (GRSV) e *Tomato spotted wilt orthotospovirus* (TSWV). As barras indicam as delimitações das regiões N e C-terminais e os triângulos e a estrela indicam os resíduos de aminoácidos associados aos ácidos nucleicos.

1024

1025

1026 2.8 References

- 1027 1. Walter CT, Barr JN: **Recent advances in the molecular and cellular biology**
 1028 **of bunyaviruses.** *The Journal of general virology* 2011, **92**(Pt 11):2467-2484.

- 1029 2. de Avila AC, de Haan P, Kormelink R, Resende Rde O, Goldbach RW, Peters
1030 D: **Classification of tospoviruses based on phylogeny of nucleoprotein gene**
1031 **sequences.** *The Journal of general virology* 1993, **74 (Pt 2):**153-159.
- 1032 3. Pappu HR, Jones RA, Jain RK: **Global status of tospovirus epidemics in**
1033 **diverse cropping systems: successes achieved and challenges ahead.** *Virus*
1034 *research* 2009, **141(2):**219-236.
- 1035 4. Richmond KE, Chenault K, Sherwood JL, German TL: **Characterization of the**
1036 **nucleic acid binding properties of tomato spotted wilt virus nucleocapsid**
1037 **protein.** *Virology* 1998, **248(1):**6-11.
- 1038 5. Ribeiro D, Borst JW, Goldbach R, Kormelink R: **Tomato spotted wilt virus**
1039 **nucleocapsid protein interacts with both viral glycoproteins Gn and Gc in**
1040 **planta.** *Virology* 2009, **383(1):**121-130.
- 1041 6. Soellick T, Uhrig JF, Bucher GL, Kellmann JW, Schreier PH: **The movement**
1042 **protein NSm of tomato spotted wilt tospovirus (TSWV): RNA binding,**
1043 **interaction with the TSWV N protein, and identification of interacting plant**
1044 **proteins.** *Proceedings of the National Academy of Sciences of the United States*
1045 *of America* 2000, **97(5):**2373-2378.
- 1046 7. Feng Z, Chen X, Bao Y, Dong J, Zhang Z, Tao X: **Nucleocapsid of Tomato**
1047 **spotted wilt tospovirus forms mobile particles that traffic on an**
1048 **actin/endoplasmic reticulum network driven by myosin XI-K.** *The New*
1049 *phytologist* 2013.
- 1050 8. Ariza A, Tanner SJ, Walter CT, Dent KC, Shepherd DA, Wu W, Matthews SV,
1051 Hiscox JA, Green TJ, Luo M *et al*: **Nucleocapsid protein structures from**
1052 **orthobunyaviruses reveal insight into ribonucleoprotein architecture and**
1053 **RNA polymerization.** *Nucleic acids research* 2013, **41(11):**5912-5926.
- 1054 9. Guu TS, Zheng W, Tao YJ: **Bunyavirus: structure and replication.** *Advances*
1055 *in experimental medicine and biology* 2012, **726:**245-266.
- 1056 10. de Oliveira AS, Melo FL, Inoue-Nagata AK, Nagata T, Kitajima EW, Resende
1057 RO: **Characterization of bean necrotic mosaic virus: a member of a novel**
1058 **evolutionary lineage within the Genus Tospovirus.** *PloS one* 2012,
1059 **7(6):**e38634.
- 1060 11. Snippe M, Willem Borst J, Goldbach R, Kormelink R: **Tomato spotted wilt**
1061 **virus Gc and N proteins interact in vivo.** *Virology* 2007, **357(2):**115-123.
- 1062 12. Mir MA, Panganiban AT: **The bunyavirus nucleocapsid protein is an RNA**
1063 **chaperone: possible roles in viral RNA panhandle formation and genome**
1064 **replication.** *RNA (New York, NY)* 2006, **12(2):**272-282.
- 1065 13. Mir MA, Panganiban AT: **The hantavirus nucleocapsid protein recognizes**
1066 **specific features of the viral RNA panhandle and is altered in conformation**
1067 **upon RNA binding.** *Journal of virology* 2005, **79(3):**1824-1835.
- 1068 14. Brennan B, Welch SR, Elliott RM: **The consequences of reconfiguring the**
1069 **ambisense S genome segment of Rift Valley fever virus on viral replication**
1070 **in mammalian and mosquito cells and for genome packaging.** *PLoS*
1071 *pathogens* 2014, **10(2):**e1003922.

- 1072 15. Li W, Lewandowski DJ, Hilf ME, Adkins S: **Identification of domains of the**
1073 **Tomato spotted wilt virus NSm protein involved in tubule formation,**
1074 **movement and symptomatology.** *Virology* 2009, **390**(1):110-121.
- 1075 16. Singh P, Indi SS, Savithri HS: **Groundnut bud necrosis virus encoded NSm**
1076 **associates with membranes via its C-terminal domain.** *PloS one* 2014,
1077 **9**(6):e99370.
- 1078 17. Kainz M, Hilson P, Sweeney L, Derosé E, German TL: **Interaction Between**
1079 **Tomato spotted wilt virus N Protein Monomers Involves Nonelectrostatic**
1080 **Forces Governed by Multiple Distinct Regions in the Primary Structure.**
1081 *Phytopathology* 2004, **94**(7):759-765.
- 1082 18. Uhrig JF, Soellick TR, Minke CJ, Philipp C, Kellmann JW, Schreier PH:
1083 **Homotypic interaction and multimerization of nucleocapsid protein of**
1084 **tomato spotted wilt tospovirus: identification and characterization of two**
1085 **interacting domains.** *Proceedings of the National Academy of Sciences of the*
1086 *United States of America* 1999, **96**(1):55-60.
- 1087 19. Zheng W, Olson J, Vakharia V, Tao YJ: **The crystal structure and RNA-**
1088 **binding of an orthomyxovirus nucleoprotein.** *PLoS pathogens* 2013,
1089 **9**(9):e1003624.
- 1090 20. Reguera J, Malet H, Weber F, Cusack S: **Structural basis for encapsidation of**
1091 **genomic RNA by La Crosse Orthobunyavirus nucleoprotein.** *Proceedings of*
1092 *the National Academy of Sciences of the United States of America* 2013,
1093 **110**(18):7246-7251.
- 1094 21. Dong H, Li P, Elliott RM, Dong C: **Structure of Schmallenberg**
1095 **orthobunyavirus nucleoprotein suggests a novel mechanism of genome**
1096 **encapsidation.** *Journal of virology* 2013, **87**(10):5593-5601.
- 1097 22. Niu F, Shaw N, Wang YE, Jiao L, Ding W, Li X, Zhu P, Upur H, Ouyang S,
1098 Cheng G *et al*: **Structure of the Leanyer orthobunyavirus nucleoprotein-**
1099 **RNA complex reveals unique architecture for RNA encapsidation.**
1100 *Proceedings of the National Academy of Sciences of the United States of*
1101 *America* 2013, **110**(22):9054-9059.
- 1102 23. Raymond DD, Piper ME, Gerrard SR, Skiniotis G, Smith JL: **Phleboviruses**
1103 **encapsidate their genomes by sequestering RNA bases.** *Proceedings of the*
1104 *National Academy of Sciences of the United States of America* 2012,
1105 **109**(47):19208-19213.
- 1106 24. Carter SD, Surtees R, Walter CT, Ariza A, Bergeron E, Nichol ST, Hiscox JA,
1107 Edwards TA, Barr JN: **Structure, function, and evolution of the Crimean-**
1108 **Congo hemorrhagic fever virus nucleocapsid protein.** *Journal of virology*
1109 2012, **86**(20):10914-10923.
- 1110 25. Ferron F, Li Z, Danek EI, Luo D, Wong Y, Coutard B, Lantéz V, Charrel R,
1111 Canard B, Walz T *et al*: **The hexamer structure of Rift Valley fever virus**
1112 **nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein**
1113 **complexes.** *PLoS pathogens* 2011, **7**(5):e1002030.
- 1114 26. Brunotte L, Kerber R, Shang W, Hauer F, Hass M, Gabriel M, Lelke M, Busch
1115 C, Stark H, Svergun DI *et al*: **Structure of the Lassa virus nucleoprotein**

- 1116 revealed by X-ray crystallography, small-angle X-ray scattering, and
1117 electron microscopy. *The Journal of biological chemistry* 2011,
1118 **286**(44):38748-38756.
- 1119 27. Reguera J, Cusack S, Kolakofsky D: **Segmented negative strand RNA virus**
1120 **nucleoprotein structure**. *Current opinion in virology* 2014, **5**:7-15.
- 1121 28. Tripathi D, Raikhy G, Pappu HR: **Movement and nucleocapsid proteins coded**
1122 **by two tospovirus species interact through multiple binding regions in**
1123 **mixed infections**. *Virology* 2015.
- 1124 29. Leastro MO, Pallas V, Resende RO, Sanchez-Navarro JA: **The movement**
1125 **proteins (NSm) of distinct tospoviruses peripherally associate with cellular**
1126 **membranes and interact with homologous and heterologous NSm and**
1127 **nucleocapsid proteins**. *Virology* 2015, **478c**:39-49.
- 1128 30. Kairys V, Gilson MK, Fernandes MX: **Using protein homology models for**
1129 **structure-based studies: approaches to model refinement**.
1130 *TheScientificWorldJournal* 2006, **6**:1542-1554.
- 1131 31. Sellers BD, Nilmeier JP, Jacobson MP: **Antibodies as a model system for**
1132 **comparative model refinement**. *Proteins* 2010, **78**(11):2490-2505.
- 1133 32. Dong H, Li P, Bottcher B, Elliott RM, Dong C: **Crystal structure of**
1134 **Schmallenberg orthobunyavirus nucleoprotein-RNA complex reveals a**
1135 **novel RNA sequestration mechanism**. *RNA (New York, NY)* 2013, **19**(8):1129-
1136 1136.
- 1137 33. Li J, Feng Z, Wu J, Huang Y, Lu G, Zhu M, Wang B, Mao X, Tao X: **Structure**
1138 **and function analysis of nucleocapsid protein of tomato spotted wilt virus**
1139 **interacting with RNA using homology modeling**. *The Journal of biological*
1140 *chemistry* 2015, **290**(7):3950-3961.
- 1141 34. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F,
1142 Gallo Cassarino T, Bertoni M, Bordoli L *et al*: **SWISS-MODEL: modelling**
1143 **protein tertiary and quaternary structure using evolutionary information**.
1144 *Nucleic acids research* 2014, **42**(Web Server issue):W252-258.
- 1145 35. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and**
1146 **accurate multiple sequence alignment**. *Journal of molecular biology* 2000,
1147 **302**(1):205-217.
- 1148 36. Hall T: **BioEdit: a user-friendly biological sequence alignment editor and**
1149 **analysis program for Windows 95/98/NT**. *Nucleic Acids Symp Ser* 1999,
1150 **41**:95-98.
- 1151 37. Sali A, Blundell TL: **Comparative protein modelling by satisfaction of**
1152 **spatial restraints**. *Journal of molecular biology* 1993, **234**(3):779-815.
- 1153 38. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D,
1154 Karplus K: **Improving physical realism, stereochemistry, and side-chain**
1155 **accuracy in homology modeling: Four approaches that performed well in**
1156 **CASP8**. *Proteins* 2009, **77 Suppl 9**:114-122.
- 1157 39. Ramachandran S, Kota P, Ding F, Dokholyan NV: **Automated minimization of**
1158 **steric clashes in protein structures**. *Proteins* 2011, **79**(1):261-270.

- 1159 40. Colovos C, Yeates TO: **Verification of protein structures: patterns of**
1160 **nonbonded atomic interactions.** *Protein science : a publication of the Protein*
1161 *Society* 1993, **2**(9):1511-1519.
- 1162 41. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ,
1163 Murray LW, Richardson JS, Richardson DC: **MolProbity: all-atom structure**
1164 **validation for macromolecular crystallography.** *Acta crystallographica*
1165 *Section D, Biological crystallography* 2010, **66**(Pt 1):12-21.
- 1166 42. Emsley P, Lohkamp B, Scott WG, Cowtan K: **Features and development of**
1167 **Coot.** *Acta crystallographica Section D, Biological crystallography* 2010, **66**(Pt
1168 4):486-501.
- 1169 43. Krissinel E, Henrick K: **Inference of macromolecular assemblies from**
1170 **crystalline state.** *Journal of molecular biology* 2007, **372**(3):774-797.
- 1171 44. DeLano W: **The PyMOL molecular graphics system.** *DeLano Scientific, San*
1172 *Carlos, CA* 2002.
- 1173 45. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E:
1174 **GROMACS: High performance molecular simulations through multi-level**
1175 **parallelism from laptops to supercomputers.** *SoftwareX* 2015, **1–2**:19-25.
- 1176 46. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw
1177 DE: **Improved side-chain torsion potentials for the Amber ff99SB protein**
1178 **force field.** *Proteins* 2010, **78**(8):1950-1958.
- 1179 47. Jorgensen WL, Madura JD: **Quantum and statistical mechanical studies of**
1180 **liquids. 25. Solvation and conformation of methanol in water.** *Journal of the*
1181 *American Chemical Society* 1983, **105**(6):1407-1413.
- 1182 48. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM: **LINCS: A linear constraint**
1183 **solver for molecular simulations.** *Journal of Computational Chemistry* 1997,
1184 **18**(12):1463-1472.
- 1185 49. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG: **A**
1186 **smooth particle mesh Ewald method.** *The Journal of Chemical Physics* 1995,
1187 **103**(19):8577-8593.
- 1188 50. Mateu MG: **Assembly, stability and dynamics of virus capsids.** *Archives of*
1189 *biochemistry and biophysics* 2013, **531**(1-2):65-79.
- 1190 51. Lima RN, Faheem M, Barbosa J, Polêto MD, Verli H, Melo FL, Resende RO:
1191 **Homology modeling and molecular dynamics provide structural insights**
1192 **into tospovirus nucleoprotein.** *BMC Bioinformatics* 2016, **17**(Suppl 18).
- 1193 52. Komoda K, Narita M, Yamashita K, Tanaka I, Yao M: **The asymmetric**
1194 **trimeric ring structure of the nucleocapsid protein of Tospovirus.** *Journal of*
1195 *virology* 2017.
- 1196 53. Guo Y, Liu B, Ding Z, Li G, Liu M, Zhu D, Sun Y, Dong S, Lou Z: **A distinct**
1197 **mechanism for the formation of the ribonucleoprotein complex of the**
1198 **Tomato spotted wilt virus.** *Journal of virology* 2017.

1199

1200

1201 **CAPÍTULO III: The complete genome of the tospovirus**
1202 ***Zucchini lethal chlorosis orthospovirus* and the structural**
1203 **analyses of its nucleocapsid protein (NP).**

1204

Este capítulo foi publicado de forma modificada na revista *Virology Journal*. Lima RN, de Oliveira AS, Leastro MO, Blawid R, Nagata T, Resende RO, Melo FL: **The complete genome of the tospovirus *Zucchini lethal chlorosis tospovirus***. *Virology Journal* 2016, 13:123

1205

1206

1207 The complete genome of the tospovirus *Zucchini lethal chlorosis*
1208 *orthotospovirus* and the structural analyses of its nucleocapsid
1209 protein (NP).

1210

1211 Lima, R.N.¹; De Oliveira, A.S.¹; Leastro, M.O.¹; Blawid, R.¹; Nagata, T.¹; Resende, R.O.¹;

1212 Melo, F.L.^{1*}.

1213 ¹ Department of Cell Biology, University of Brasilia, Brasilia, Distrito Federal, Brazil.

1214

1215

1216

1217 *Corresponding author:

1218 Dr. Fernando Lucas Melo

1219 University of Brasilia

1220 Brasilia, DF, 70910-900

1221 e-mail: flmelo@unb.br

1222 **Keywords:** Zucchini, tospovirus, Illumina, genome, ZLCV.

1223 **3.1 Abstract**

1224

1225 **3.1.1 Background**

1226

1227 *Zucchini lethal chlorosis orthospovirus* (ZLCV) causes significant losses in the
1228 production of cucurbits in Brazil. This virus belongs to the genus *Orthospovirus*
1229 (family *Tospoviridae*) and seems to be exclusively transmitted by *Frankliniella zucchini*
1230 (Thysanoptera). Orthospoviruses have a tripartite and single-stranded RNA genome
1231 classified as S (Small), M (Medium) and L (Large) RNAs. Although ZLCV was
1232 identified as a member of the genus *Orthospovirus* in 1999, its complete genome had
1233 not been sequenced until now.

1234 **3.1.2 Results**

1235

1236 We sequenced the full-length genome of two ZLCV isolates named ZLCV-SP and
1237 ZLCV-DF. The phylogenetic analysis showed that ZLCV-SP and ZLCV-DF clustered
1238 with the previously reported isolate ZLCV-BR09. Their proteins were closely related,
1239 except the non-structural protein (NS_M), which was highly divergent (approximately
1240 90% identity). All viral proteins clustered similarly in our phylogenetic analysis,
1241 excluding that these ZLCV isolates have originated from reassortment events of
1242 different tospovirus species.

1243 **3.1.3 Conclusions**

1244

1245 Here we report for the first time the complete genome of two ZLCV isolates that were
1246 found in the field infecting zucchini and cucumber.

1247

1248 **3.2 Background**

1249

1250 *Zucchini lethal chlorosis virus* (ZLCV) is a member of the genus
1251 *Orthospovirus*, family *Tospoviridae* [1]. Although some orthospovirus species are
1252 notorious for their broad-host range [2], ZLCV mainly infects cucurbits and is only
1253 known to be transmitted by *Frankliniella zucchini* (Thysanoptera: Thripidae) [3]. So far
1254 ZLCV has just been reported in Brazil, naturally infecting sponge gourd, West Indian
1255 gherkin, cucumber, watermelon and several species of squash [4, 5]. Infection by ZLCV
1256 abrogate fruit production in zucchini plants and the only resistant cultivar is *Cucurbita*
1257 *maxima* cv. Exposição [6], making ZLCV one of the most economically important viral
1258 pathogen for cucurbits in Brazil.

1259 Orthospoviruses have a tripartite single-stranded RNA genome and each
1260 segment is named according to its size. The L (large) RNA has a negative polarity and
1261 encodes a RNA-dependent RNA polymerase (RdRp) [7]. The ambisense M (medium)
1262 RNA encodes the precursor of two viral glycoproteins (G_N/G_C) and a non-structural
1263 protein (NS_M) involved in viral cell-to-cell movement [8]. The S (small) RNA, which is
1264 also ambisense, encodes another non-structural protein (NS_S) with RNA silencing
1265 suppression activity and the nucleocapsid (NP) protein [9].

1266 There are currently 11 approved and 18 tentative tospovirus species, but only a
1267 small number of species were completely sequenced. Genome sequence data has the
1268 potential to solve key questions in tospovirus evolution, epidemiology and physiology,
1269 such as the occurrence and importance of interspecific reassortment [10] and the
1270 presence of potential undescribed genes [11]. Even though ZLCV was described in
1271 1999 [1] and some genes of one isolate has been sequenced [12-14], its complete
1272 genome is still unknown. Here, we report the complete genome of two ZLCV isolates

1273 found infecting zucchini (*Cucurbita pepo* cv. Caserta) and cucumber (*Cucumis sativus*
1274 L.) in Brazil and compared them to other tospoviruses.

1275 In 2010, a virus isolate (hereafter ZLCV-SP) from zucchini was found in a
1276 commercial field in São Paulo state and transmitted to *Datura stramonium* L. by *F.*
1277 *zucchini* as previously described [3]. Then, virus particles were propagated in *D.*
1278 *stramonium* by mechanical inoculation and infected leaves were used for
1279 ribonucleoprotein (RNP) purification following the protocol of De Avila et al. [15].
1280 Moreover, cucumber plants showing typical ZLCV symptoms were collected in
1281 Planaltina, Federal District, in 2015. Viral particles were semi-purified from leaves as
1282 previously described [16]. Briefly, 40g of leaf material were homogenized in PBS-
1283 EDTA plus 0.2% 2-mercaptoethanol. The plant extract was then filtered and centrifuged
1284 through a sucrose cushion at 33,000 x g for 2 h and the pellet resuspended in PBS.
1285 Genomic RNA was extracted from purified RNP_s of both isolates as previously
1286 described by De Oliveira et al. [17] and sequenced at Macrogen (South Korea) using
1287 Illumina HiSeq 2000 platform. The resulting paired-end reads were filtered and
1288 assembled *de novo* using CLC Genomics Workbench version 6.0.3. The contigs related
1289 to ZLCV were selected using Blastx against a RefSeq virus database. To determine if
1290 the entire length of each segment was included in the assembled contigs, the reads were
1291 mapped back to the ZLCV related contigs. All contigs from both isolates presented the
1292 consensus sequences AGAGCAAU and AUUGCUCU at the 5'- and 3'-terminal ends,
1293 but some contigs presented distal terminal bases that were trimmed off. These
1294 palindromic sequences are conserved among all tospoviruses. Moreover, the ZLCV
1295 segments derived from cucumber samples (hereafter ZLCV-DF) presented some
1296 unresolved gaps (L segment: 1 gap of 13 nucleotide and 1 of 2 nucleotide). The genome
1297 of both isolates were annotated and submitted to NCBI GenBank under the accession

1298 numbers no. KU641378-KU641380 (ZLCV-SP) and KU681010-KU681012 (ZLCV-
1299 DF)

1300 Despite few variations, both genomes presented the same characteristics as listed
1301 in Table 1.

1302 Table 1. Genome comparison of ZLCV isolates.

	ZLCV-SP	ZLCV-DF
L RNA		
L RNA full length (nt)	8885	8886
5'UTR (nt)	222 (1-222) *	223 (1-223)
L gene ORF (nt)	8631(223-8853)	8631(224-8854)
L protein (aa)	2877 (330.85 kDa)	2877 (331.16 kDa)
3'UTR (nt)	32 (8854-8885)	32 (8855-8886)
M RNA		
M RNA full length (nt)	4860	4829
5'UTR (nt)	113 (1-113)	113 (1-113)
NSm gene ORF (nt)	909 (114-1022)	909 (114-1022)
NSm protein (aa)	303 (34.4 kDa)	303 (34,37 kDa)
IGR (nt)	344 (1023-1366)	313 (1-1335)
GPs gene ORF (nt)	3408 (1367-4774)	3408 (1336-4743)
GP protein (aa)	1136 (127.58 kDa)	1136 (127.73 kDa)
3'UTR (nt)	86 (4775-4860)	86 (4744-4829)
S RNA		
S RNA full length (nt)	3524	3524
5'UTR (nt)	87(1-87)	87(1-87)
NSs gene ORF (nt)	1404 (88-1491)	1404 (88-1491)
NSs protein (aa)	468 (53.06 kDa)	468 (53.01 kDa)
IGR (nt)	1009 (1492-2500)	1009 (1492-2500)
NP gene ORF (nt)	783 (2501-3283)	783 (2501-3283)
NP protein (aa)	261 (29.22 kDa)	261 (29.24 kDa)
3'UTR (nt)	241 (3284-3524)	241 (3284-3524)

1303 nt= nucleotides
1304 *= position in the genome
1305 aa= amino acids
1306 kDa= kilodaltons
1307 IGR=Intergenic region
1308 GP= glycoproteins precursor

1309 The 5' and 3' UTR regions of both viruses presented the same size, except for
1310 one nucleotide in the 5'UTR of L segment of ZLCV-DF. The intergenic regions (IGR)
1311 presented small indels, for example, a duplication of 20 nucleotides in the M segment
1312 IGR of ZLCV-SP. Despite the importance of these regions as transcription termination
1313 signal [18], the impact of such deletions is still unclear.

1314 The L segments vary one nucleotide (as described above) between the isolates,
1315 8,885 nt for ZLCV-SP and 8,886 nt for ZLCV-DF and encode a RdRp of about 331 kDa
1316 (2877 aa). The M segment contains 4,860 nt for ZLCV-SP and 4,829 nt for ZLCV-DF,
1317 and presented the ambisense arrangement typical of tospoviruses. For both isolates, a
1318 NS_M protein of about 34 kDa (303 aa) and glycoprotein precursor (GP) protein of about
1319 127 kDa (1136 aa) are predicted. The S segment is 3,524 nt long and encodes the NS_S
1320 and NP proteins with predicted molecular sizes of about 53 kDa (468 aa) and 29 kDa
1321 (261 aa), respectively. Since some genes of other ZLCV isolate (ZLCV-BR09) have
1322 been previously sequenced [12-14], their encoded proteins were compared with ZLCV-
1323 SP and ZLCV-DF (Table 2).

1324 Pairwise sequence comparison of the NP amino acids indicates that ZLCV-SP is
1325 most closely related to the ZLCV-DF than ZLCV-BR09. The identities between the NP
1326 of ZLCV-SP/ZLCV-DF and between ZLCV-SP/ZLCV-BR09 are 99.23% and 98.46%,
1327 respectively, while the identity of ZLCV-DF/ZLCV-BR09 is 97.69%. The NS_M from
1328 ZLCV-BR09 displayed a significant variation in comparison with ZLCV-SP and
1329 ZLCV-DF (approximately 90% identity) (Table 2).

1330 Interestingly, most variation was located at the C-terminus with an increased
1331 number of non-conservative amino acid changes (Fig. 1), suggesting that this region
1332 may be under positive selection. Actually, it was recently demonstrated that this same

1333 region is implicated in NS_M and plasma-membrane interaction [19]. However, the
 1334 biological impact of these variations remains to be determined.

1335

1336 Table 2. Identity among ZLCV isolates proteins.

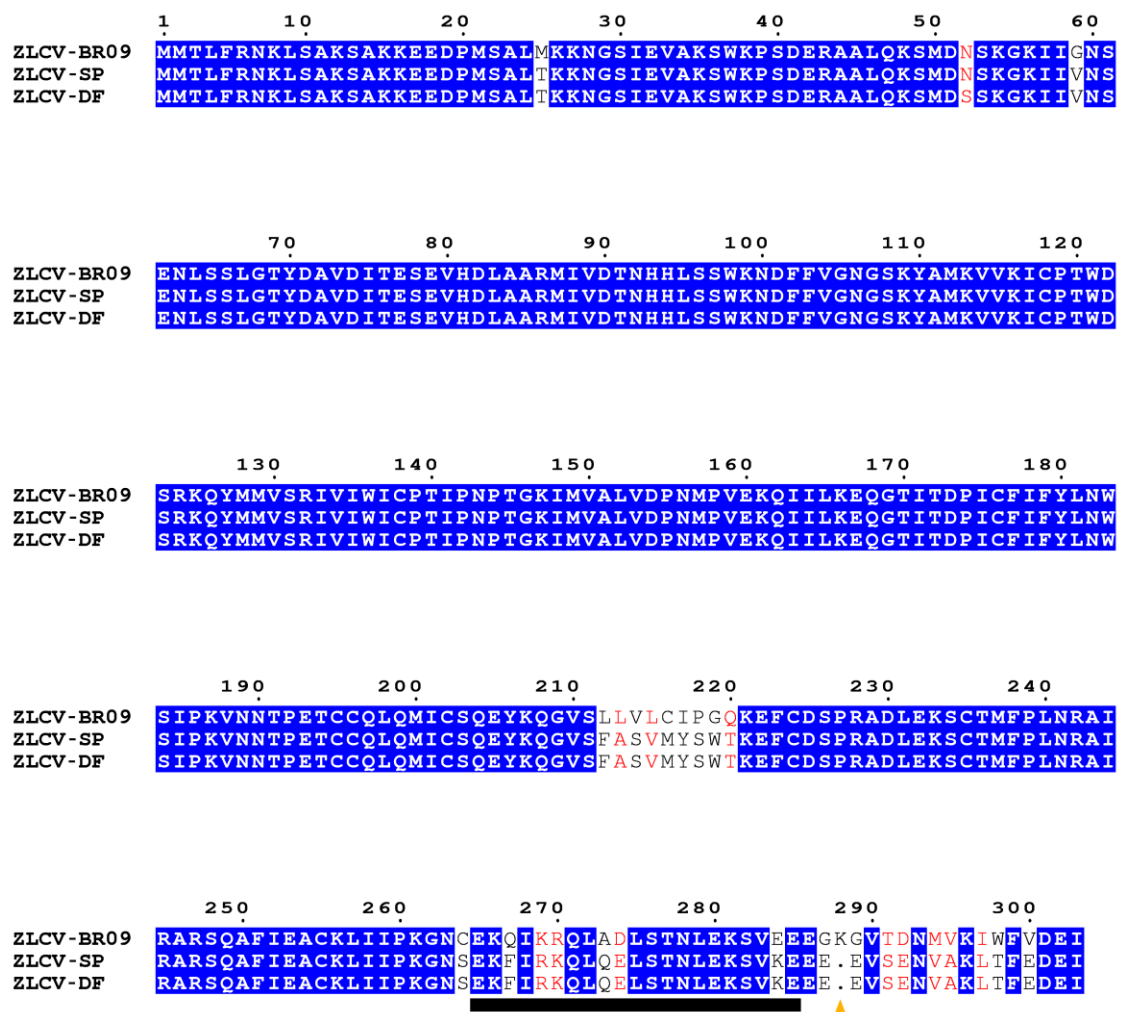
Identity of NP protein			
	ZLCV-BR09	ZLCV-DF	ZLCV-SP
ZLCV-BR09	100%	97.69%	98.46%
ZLCV-DF	97.69%	100%	99.23%
ZLCV-SP	98.46%	99.23%	100%
Identity of NSs protein			
	ZLCV-BR09	ZLCV-DF	ZLCV-SP
ZLCV-BR09	100%	98.72%	99,14%
ZLCV-DF	98.72%	100%	99.57%
ZLCV-SP	99,14%	99.57%	100%
Identity of GP protein			
	ZLCV-BR09	ZLCV-DF	ZLCV-SP
ZLCV-BR09	100%	97.53%	99.30%
ZLCV-DF	97.53%	100%	97.36%
ZLCV-SP	99.30%	97.36%	100%
Identity of NS _M protein			
	ZLCV-BR09	ZLCV-DF	ZLCV-SP
ZLCV-BR09	100%	90.73%	91.03%
ZLCV-DF	90.73%	100%	99.67%
ZLCV-SP	91.03%	99.67%	100%
Identity of RdRp protein			
	ZLCV-DF	ZLCV-SP	
ZLCV-DF	100%	98.02%	
ZLCV-SP	98.02%	100%	

1337

1338 To infer evolutionary relationships among tospoviruses, we compiled sequences
 1339 of each tospovirus protein. The final data sets contain 33, 28, 23, 23 and 20 sequences
 1340 for the NP, NSs, GP, NS_M and RdRp proteins, respectively. The GenBank accession
 1341 numbers are listed in Table 3.

1342 The sequences were aligned using MAFFT [20] and the maximum likelihood
 1343 trees were inferred using FastTree [21], both implemented in Geneious 9.1. As
 1344 previously observed, tospoviruses can be divided into two geographic groups, with
 1345 distinct viral species observed in Europe/Asia (Eurasian clade) and the Americas
 1346 (American Clade) (Fig. 2). However, the analysis of NP and NSs trees indicate that

1347 some tospoviruses are clearly distinct from any species of these two groups, such as the
 1348 *Lisianthus necrotic ringspot virus* (LNRV), Groundnut chlorotic fan-spot virus
 1349 (GCFSV) and the *Groundnut yellow spot virus* (GYSV) (Fig. 2). The ZLCV isolates
 1350 clustered within the American clade and no reassortment had occurred during the
 1351 evolution of these ZLCV isolates (Fig. 2).
 1352



1353

Figure 1. The NS_M amino acid sequence alignment of ZLCV isolates. Residues common to the three sequences are shown in white on blue and similar residues are shown in red. The yellow triangle indicates the single amino acid deletion and the black rectangle indicates the coiled-coil (E²⁶⁵-E²⁸⁵) predicted structure (Uniprot Accession code: Q91PB4). This figure was made with ESPrnt (<http://esprnt.ibcp.fr>).

1354

1355

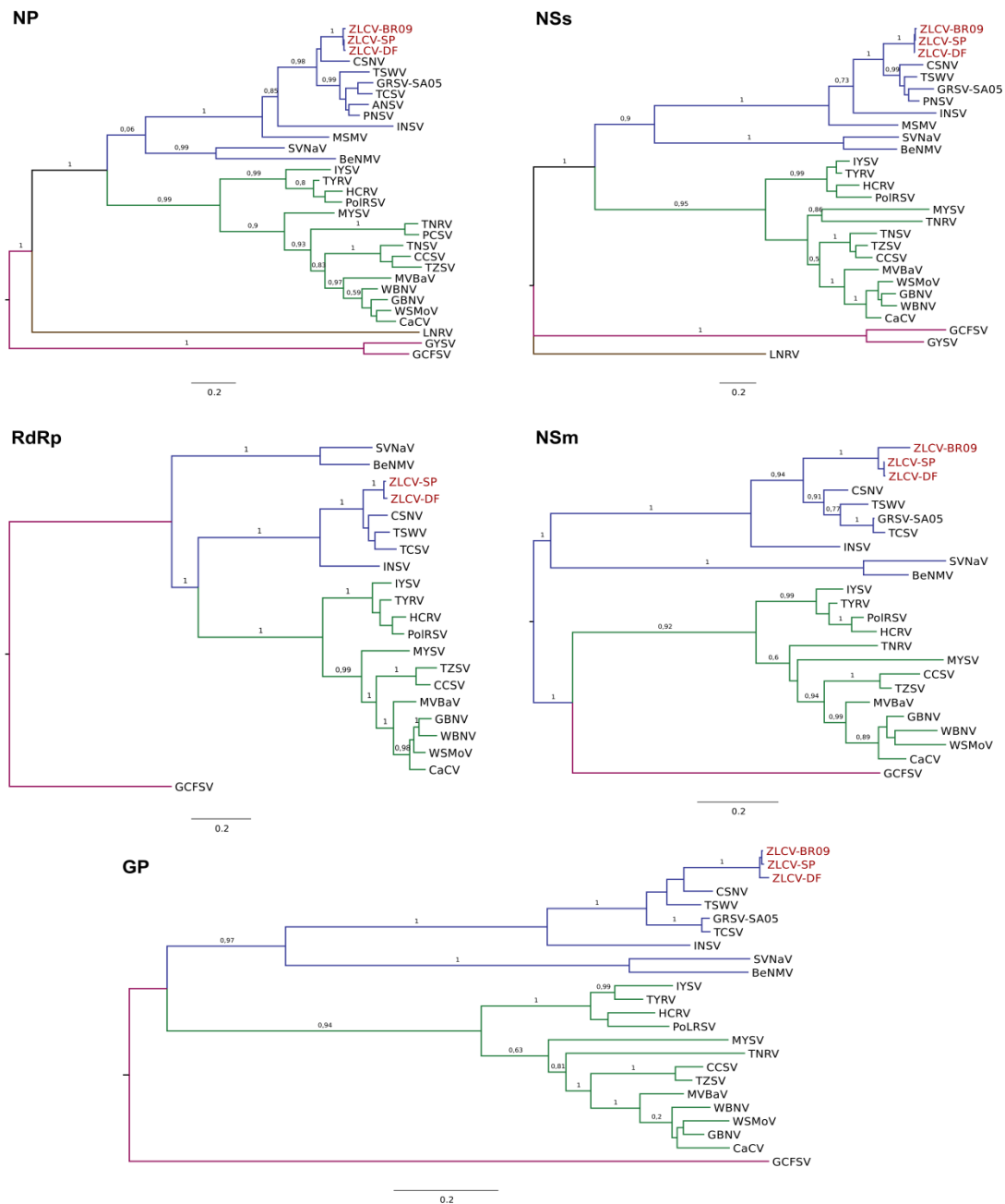
1356 Table 3. GenBank accession number of the sequences used in this study.

Acronym	Tospovirus	L RNA	M RNA	S RNA
ANSV	Alstroemeria necrotic streak virus	-	-	NP: GQ478668
BeNMV	Bean necrotic mosaic virus	JF417980	JN587269	JN587268
CaCV	Capsicum chlorosis virus	NC_008302	NC_008303	NC_008301
CCSV	Calla lily chlorotic spot virus	FJ822962	FJ822961	AY867502
CSNV	Chrysanthemum stem necrosis virus	KM114546	KM114547	KM114548
GBNV	<i>Groundnut bud necrosis virus</i>	AF025538	U42555	AY871098
GCFSV	Groundnut chlorotic fan- spot virus	KP146140	KP146141	AF080526
GRSV-SA05	<i>Groundnut ringspot virus</i>	-	NS _M : AF213673 GP: AY574055	NP: AF487516 NSs: JN571117
GYSV	<i>Groundnut yellow spot virus</i>	-	-	HQ402596
HCRV	Hippeastrum chlorotic ringspot virus	HG763861	JX833565	JX833564
INSV	<i>Impatiens necrotic spot virus</i>	DQ425094	DQ425095	X66972
IYSV	<i>Iris yellow spot virus</i>	FJ623474	FJ361359	AF001387
LNRV	Lisianthus necrotic ringspot virus	-	-	AB852525
MSMV	Melon severe mosaic virus	-	-	EU275149
MVBaV	Mulberry vein banding associated virus	KM819698	KM819699	KM819701
MYSV	Melon yellow spot virus	AB061774	AB061773	AB038343
PCSV	Pepper chlorotic spot virus	-	-	KF383956
PNSV	Pepper necrotic spot virus	-	-	HE584762
PolRSV	<i>Polygonum ringspot virus</i>	KJ541746	KJ541745	KJ541744
SVNaV	Soybean vein necrosis associated virus	HQ728385	HQ728386	HQ728387
TCSV	<i>Tomato chlorotic spot virus</i>	HQ700667	NS _M : AF213674 GP: AY574054	NP: AF282982
TNRV	Tomato necrotic ringspot virus	-	FJ947152	FJ489600
TNSV	Tomato necrotic spot virus	-	-	KM355773
TSWV	<i>Tomato spotted wilt virus</i>	NC_002052	NC_002050	DQ915948
TYRV	Tomato yellow ring virus	JN560178	JN560177	AY686718
TZSV	Tomato zonate spot virus	EF552435	EF552434	EF552433
WBNV	<i>Watermelon bud necrosis virus</i>	GU735408	GU584185	GU584184
WSMoV	<i>Watermelon silver mottle virus</i>	AY863200	DQ157768	AB042650
ZLCV-BR09	<i>Zucchini lethal chlorosis virus</i>	-	NS _M : AF213676 GP: AB274027	NP: AF067069 NSs: JN572104
ZLCV-DF	<i>Zucchini lethal chlorosis virus</i>	KU681010	KU681012	KU681011
ZLCV-SP	<i>Zucchini lethal chlorosis virus</i>	KU641378	KU641379	KU641380

1357

1358 Considering the high prevalence of tospoviruses worldwide, the complete
1359 genome of the ZLCV isolates is important for future surveillance and research.
1360 Additional investigations in important plant crops should keep being performed to
1361 extend the number of characterized species.

1362



1363

■ American Clade ■ Eurasian Clade ■ LNRV Clade ■ GYSV/GCFSV Clade

Figure 2. Maximum likelihood trees of tospoviruses based inferred using all viral encoded proteins. All trees were midpoint rooted for clarity purposes. Fast tree support values are shown at the branches. The ZLCV isolates are labeled in red.

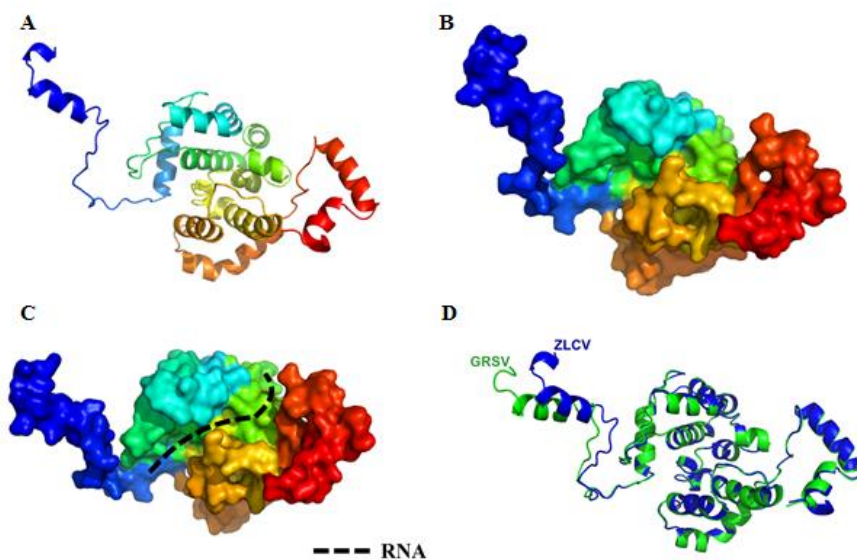
1364

1365 3.3 Considerações

1366 3.3.1 Análise da estrutura tridimensional da proteína do 1367 nucleocapsídeo (NP) do ZLCV

1368 O modelo tridimensional para a NP de ZLCV foi construído *in silico*, por
1369 Modelagem Estrutural por Homologia Molecular, a partir da estrutura da NP do GRSV
1370 e seguindo-se a mesma metodologia utilizada por Lima *et. al* (2016) [22]. O modelo
1371 obtido possui 93% de seus resíduos na região favorável e três resíduos na região
1372 desfavorável (L6, F74 e K76) de acordo com o gráfico de Ramachandran (Fig. S1). O
1373 modelo obteve score de -1.66 de acordo com as análises de minimização de energia
1374 Yasara.

1375 Assim como predito para a N de GRSV, a NP de ZLCV apresenta um braço N-
1376 terminal (1-25 aa), um braço C-terminal (224-260) e uma região globular com uma
1377 cavidade com alta afinidade por ácidos nucleicos (Fig. 3A-D).



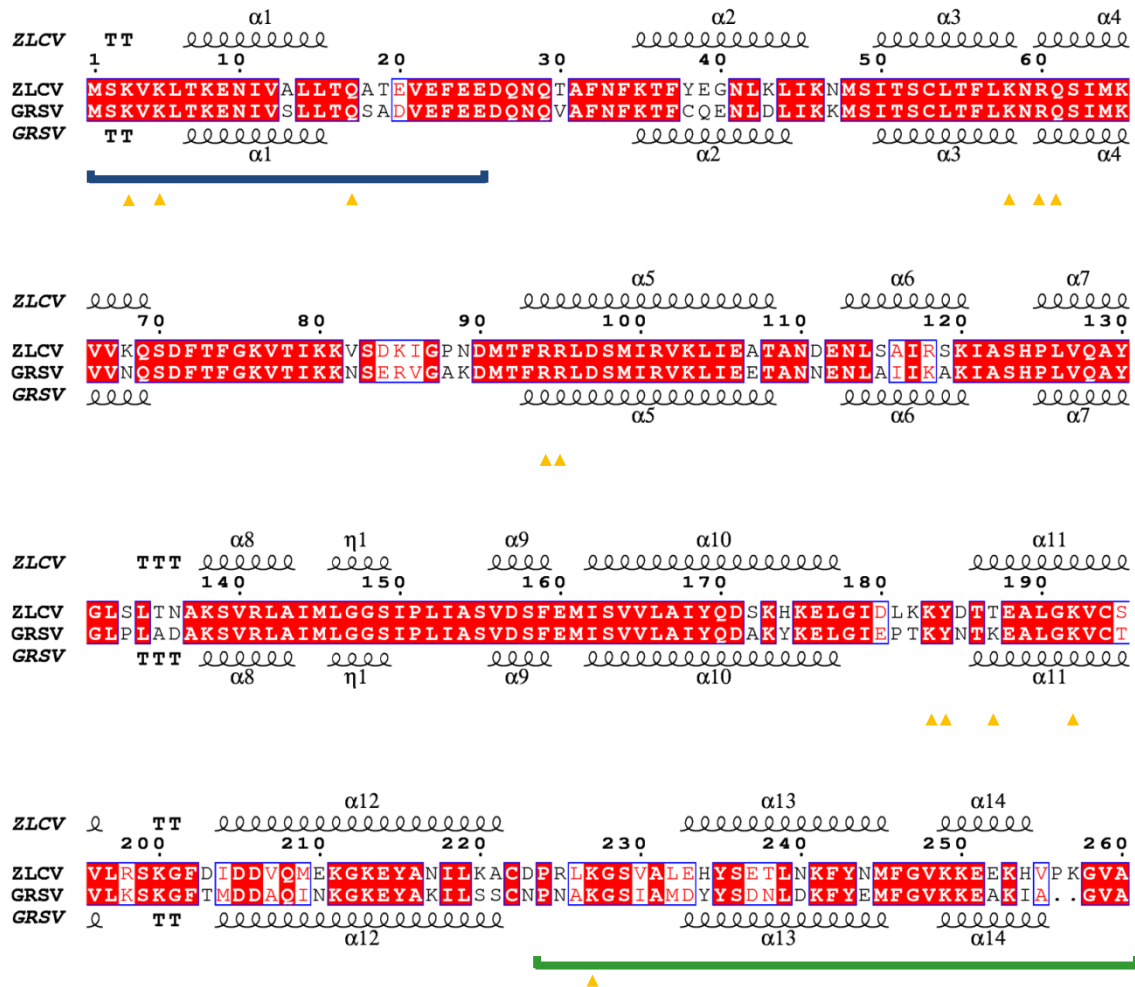
1378

1379 **Figura 3. Estrutura tridimensional da proteína NP de *Zucchini lethal chlorosis orthospovirus* (ZLCV).** **A:** representação monomérica mostrando as estruturas secundárias da proteína NP de ZLCV; **B:** representação monomérica de superfície; **C:** representação monomérica destacando a cavidade para encapsidação do RNA e **D:** sobreposição entre as estruturas secundárias das proteínas NP de ZLCV e de GRSV. Os braços N e C terminais da NP estão destacados em azul e vermelho, respectivamente.

1379

1380 O alinhamento entre sequências preditas para a proteína NP do ZLCV (260 aa) e
1381 NP do GRSV (258 aa) revelou um alto grau de conservação dos mecanismos estruturais
1382 de enovelamento e encapsidação do RNA, mesmo que mutações consideráveis sejam
1383 observadas no alinhamento primário (Fig. 4). É possível prever que a NP de ZLCV
1384 pode se associar para formar complexos oligoméricos complexos por meio das
1385 interações das extremidades N e C terminais com a região globular, da mesma forma

1386 que a predita para a NP de GRSV (Fig. 3 e 4). Os resíduos preditos para oligomerização
 1387 NP-NP na região N-terminal em GRSV: S2, V4, T7, K8, N10, V12, Q61, S62, T73,
 1388 G75 e S83 são conservados. Porém, para o resíduo N82, em ZLCV, encontra-se o V82.



1389

Figura 4. Alinhamento entre as seqüências de aminoácidos de *Groundnut ringspot orthotospovirus* (GRSV) e *Zucchini lethal chlorosis orthotospovirus* (ZLCV). As regiões N e C terminais, preditas para a NP de GRSV estão indicadas pelas barras azul e verde, respectivamente. Os resíduos de aminoácidos preditos para interação com RNA estão indicados com um triângulo amarelo. As letras vermelhas em fundo branco indicam mutações que conservam as propriedades físico-químicas dos aminoácidos.

1390 Os resíduos preditos para oligomerização NP-NP na região C-terminal em GRSV:
 1391 K 183, T186, K173, K175, S229, Y235, Y243, V248, K249 são conservados. Porém, os
 1392 seguintes resíduos são diferentes: Y174H^{ZLCV}, N185V^{ZLCV}, K198R^{ZLCV}, A226L^{ZLCV},
 1393 D233E^{ZLCV}, N238T^{ZLCV}. Destas diferenças, nota-se que há uma pressão seletiva para a
 1394 manutenção das propriedades físico-químicas dos aminoácidos tanto para manutenção
 1395 da estrutura secundária, quanto para manutenção das interações (Fig.2). Além do mais,
 1396 observa-se que os mesmos resíduos de aminoácidos aromáticos, hidrofóbicos e básicos

1397 preditos para interação com RNA e localizados na cavidade e no braço N-terminal em
1398 GRSV, são conservados em ZLCV, com exceção do resíduo T187^{ZLCV} (Fig. 2). No
1399 entanto, a treonina é um resíduo capaz de interagir com ácidos nucleicos, podendo essa
1400 mutação não interferir diretamente na encapsidação do RNA.

1401 3.2 Additional files.

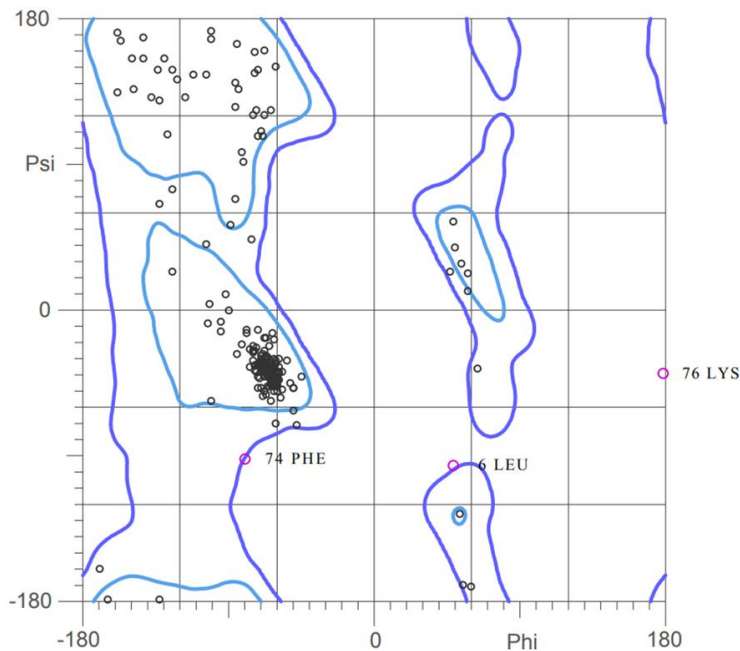


Figura S1: Gráfico de Ramachandran para a proteína NP de *Zucchini lethal chlorosis orthospovirus* (ZLCV). Delimitadas por traço azul claro estão as regiões favoráveis e em azul escuro as regiões aceitáveis. Círculos em preto representam cada resíduo de aminoácidos e em rosa, estão indicados os resíduos localizados na região desfavorável.

1402 3.4 References

1403

- 1404 1. Bezerra IC, Resende RD, Pozzer L, Nagata T, Kormelink R, De Avila AC.
1405 **Increase of tospoviral diversity in Brazil with the identification of two new**
1406 **tospovirus species, one from chrysanthemum and one from zucchini.** *Phytopathology*. 1999;89:823-830.
1407
- 1408 2. Pappu HR, Jones RAC, Jain RK. **Global status of tospovirus epidemics in**
1409 **diverse cropping systems: Successes achieved and challenges ahead.** *Virus*
1410 *Research* 2009;141:219-236.
- 1411 3. Nakahara S, Monteiro RC. **Frankliniella zucchini (Thysanoptera :**
1412 **Thripidae), a new species and vector of tospovirus in Brazil.** *Proceedings of*
1413 *the Entomological Society of Washington*. 1999;101:290-294.

- 1414 4. Yuki VA, Rezende JAM, Kitajima EW, Barroso PAV, Kuniyuki H, Groppo GA,
1415 Pavan MA. **Occurrence, distribution, and relative incidence of five viruses**
1416 **infecting cucurbits in the state of Sao Paulo, Brazil.** Plant Disease.
1417 2000;84:516-520.
- 1418 5. Camelo-García VM, Lima EFB, Rezende JAM. **Identification of natural hosts**
1419 **of Zucchini lethal chlorosis virus.** Tropical Plant Pathology. 2015;40:345-349.
- 1420 6. Giampan JS, Rezende JAM, Piedade SMS: **Yield loss caused by Zucchini**
1421 **lethal chlorosis virus (ZLCV) on zucchini squash 'Caserta'.** Summa
1422 Phytopathologica. 2009;35:223-225.
- 1423 7. De Haan P, Kormelink R, Resende RD, Vanpoelwijk F, Peters D, Goldbach R.
1424 **tomato spotted wilt virus-l rna encodes a putative rna-polymerase.** Journal
1425 of General Virology. 1991;72:2207-2216.
- 1426 8. Kormelink R, Storms M, Vanlent J, Peters D, Goldbach R. **Expression and**
1427 **subcellular location of the nsM protein of tomato spotted wilt virus (tsWV), a**
1428 **putative viral movement protein.** Virology. 1994;200:56-65.
- 1429 9. Takeda A, Sugiyama K, Nagano H, Mori M, Kaido M, Mise K, Tsuda S, Okuno
1430 T: **Identification of a novel RNA silencing suppressor, NSs protein of**
1431 **Tomato spotted wilt virus.** Febs Letters. 2002;532:75-79.
- 1432 10. Webster CG, Reitz SR, Perry KL, Adkins S: **A natural M RNA reassortant**
1433 **arising from two species of plant- and insect-infecting bunyaviruses and**
1434 **comparison of its sequence and biological properties to parental species.**
1435 Virology. 2011;413:216-225.
- 1436 11. Firth AE. **Mapping overlapping functional elements embedded within the**
1437 **protein-coding regions of RNA viruses.** Nucleic Acids Research.
1438 2014;42:12425-12439.
- 1439 12. Silva MS, Martins CR, Bezerra IC, Nagata T, de Avila AC, Resende RO.
1440 **Sequence diversity of NS(M) movement protein of tospoviruses.** Archives of
1441 Virology 2001;146:1267-1281.
- 1442 13. Nagata T, Carvalho KR, Sodre RDA, Dutra LS, Oliveira PA, Noronha EF,
1443 Lovato FA, Resende RDO, De Avila AC, Inoue-Nagata AK. **The glycoprotein**
1444 **gene of Chrysanthemum stem necrosis virus and Zucchini lethal chlorosis**
1445 **virus and molecular relationship with other tospoviruses.** Virus Genes.
1446 2007;35:785-793.
- 1447 14. Hallwass M, Leastro MO, Lima MF, Inoue-Nagata AK, Resende RO. **Sequence**
1448 **determination and analysis of the NSs genes of two tospoviruses.** 2012;
1449 157:591-596.
- 1450 15. Deavila AC, Dehaan P, Smeets MLL, Resende RD, Kormelink R, Kitajima EW,
1451 Goldbach RW, Peters D. **Distinct levels of relationships between tospovirus**
1452 **isolates.** Archives of Virology. 1993;128:211-227.
- 1453 16. Silva KN, Melo FL, Orilio AF, Nagata T, Silva MS, Fernandes CD, Frago
1454 RR, Dessaune SN, Resende RO. **Biological and molecular characterization of**

- 1455 **a highly divergent Johnsongrass mosaic virus isolate from Pennisetum**
1456 **purpureum.** Archives of Virology. 2016;Impress.
- 1457 17. de Oliveira AS, Melo FL, Inoue-Nagata AK, Nagata T, Kitajima EW, Resende
1458 RO. **Characterization of bean necrotic mosaic virus: a member of a novel**
1459 **evolutionary lineage within the Genus Tospovirus.** PLoS One.
1460 2012;7:e38634.
- 1461 18. van Knippenberg I, Goldbach R, Kormelink R. **Tomato spotted wilt virus S-**
1462 **segment mRNAs have overlapping 3 '-ends containing a predicted stem-**
1463 **loop structure and conserved sequence motif.** Virus Research. 2005;110:125-
1464 131.
- 1465 19. Singh P, Indi SS, Savithri HS. **Groundnut Bud Necrosis Virus Encoded NSm**
1466 **Associates with Membranes via Its C-Terminal Domain.** Plos One. 2014;
1467 9:e99370
- 1468 20. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software
1469 Version 7: **Improvements in Performance and Usability.** Molecular Biology
1470 and Evolution. 2013;30:772-780.
- 1471 21. Price MN, Dehal PS, Arkin AP. FastTree 2 - **Approximately maximum-**
1472 **likelihood trees for large alignments.** Plos One. 2010;5:e9490.
- 1473 22. Lima RN, Faheem M, Barbosa JA, Poletto MD, Verli H, Melo FL, Resende RO:
1474 **Homology modeling and molecular dynamics provide structural insights**
1475 **into tospovirus nucleoprotein.** BMC Bioinformatics 2016, 17(Suppl 18):489.

CAPÍTULO IV: Predição da estrutura terciária, Expressão e Purificação da proteína NS_S de GRSV.

1476 **4.1 Predição da estrutura da NSs**

1477

1478 **4.1.1 Metodologia**

1479 A sequência de aminoácidos da proteína NS_S de GRSV foi extraída do GenBank
1480 (número de acesso no GenBank: AEX32991) e submetida ao programa de predição
1481 PS2 (<http://ps2.life.nctu.edu.tw/>) [1,2]. Os alinhamentos entre a NSs e o modelo obtido
1482 foram realizados no programa online T-COFFEE e visualizados no programa ESPRIPT
1483 [3,4].

1484 O modelo obtido foi avaliado com auxílio dos parâmetros de validação dos
1485 programas ERRAT e MolProbity [5,6]. O modelo gerado foi submetido ao programa
1486 *Yasara Energy Minimization Server* [7].

1487 **4.1.2 Resultados e Discussão**

1488 O modelo tridimensional gerado nesse trabalho para a NS_S de GRSV possui
1489 89.2% de seus resíduos na região favorável e possui 8 resíduos na região desfavorável
1490 (G63, S94, G114, N189, S243, K255, V339 e M410) de acordo com o gráfico de
1491 Ramachandran (Fig. S1).

1492 Tal modelo foi predito a partir de uma Polifosfato quinase (Ppk) bacteriana (PDB:
1493 2o8r). O alinhamento das sequências de aminoácidos das duas proteínas mostrou que
1494 elas possuem uma identidade de aproximadamente 17.5% (Fig.1).

1495 Não foi surpresa, o fato de a NS_S apresentar uma similaridade estrutural com uma
1496 polifosfatase, uma vez que Lokesh et al. (2010) mostrou que a NSs de GBNV possui
1497 atividades de NTPase e fosfatase que podem estar envolvidas na remoção do 5' fosfato
1498 de moléculas de dsRNA, prevenindo, assim, a clivagem dessas moléculas em siRNAs
1499 pela atuação da enzima DICER [8].

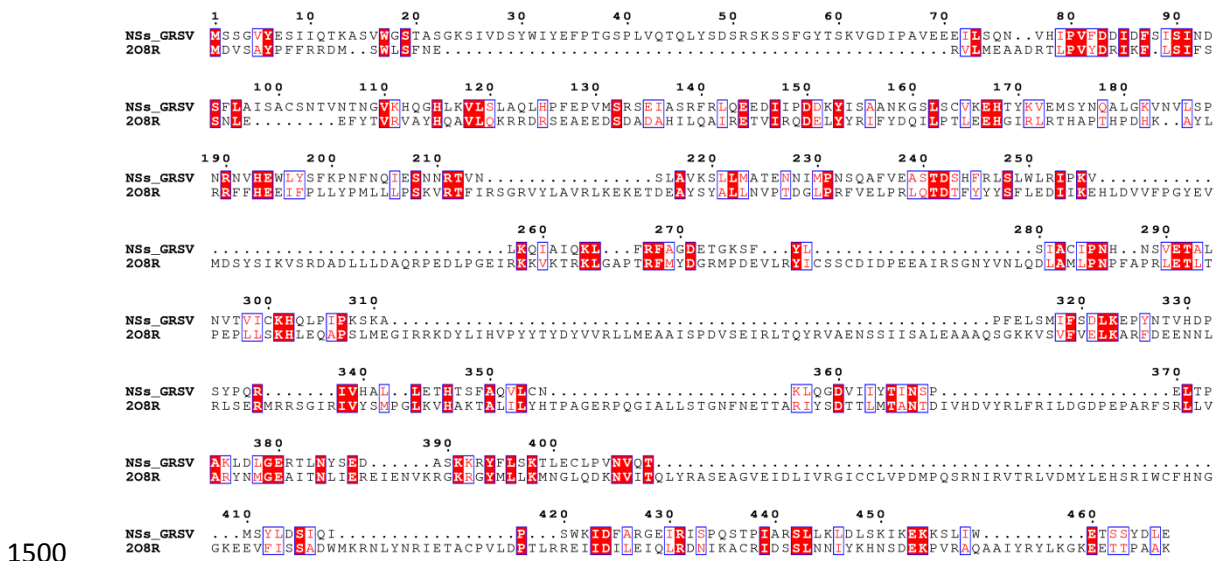


Figura 1: Alinhamento entre as seqüências da proteína NS_s de *Groundnut ringspot orthotospovirus* (GRSV) e a polifosfato kinase (Ppk) bacteriana (PDB: 2o8r). Resíduos altamente conservados entre as duas espécies estão sombreados em vermelho e escritos em branco e os resíduos com as mesmas características estão escritos em vermelho e destacados em quadrados azuis.

1501

1502 Análises das NSs de alguns orthotospovírus mostraram que essa proteína possui
 1503 afinidade para moléculas pequenas de RNA (*small* RNA) como siRNA (*small*
 1504 *interfering* RNA) e miRNA (microRNA) [9,10]. Além disso, foi demonstrado que a
 1505 NSs de GRSV possui afinidade para moléculas de RNA dupla fita (dsRNA) longas.
 1506 Uma análise da superfície eletrostática da proteína fornece regiões candidatas para
 1507 interação com essas moléculas de RNA (Fig. 2, C e D).

1508 A NSs de TSWV parece interferir na resposta do sistema imune inato da planta
 1509 sequestrando siRNAs e dsRNA longos, evitando, assim, o recrutamento do complexo
 1510 RISC [11]. De Rond et al., (2014) observou que a porção N-terminal (133 aminoácidos)
 1511 da NSs é importante para a função de silenciamento gênico e pode ser a região
 1512 candidata chave para a possível interação da NSs com a proteína Argonata1. Uma
 1513 mutação no motivo GW/WG, correspondentes aos aminoácidos 17 e 18, levou à perda
 1514 da atividade de silenciamento gênico em plantas expressando GFP [11].

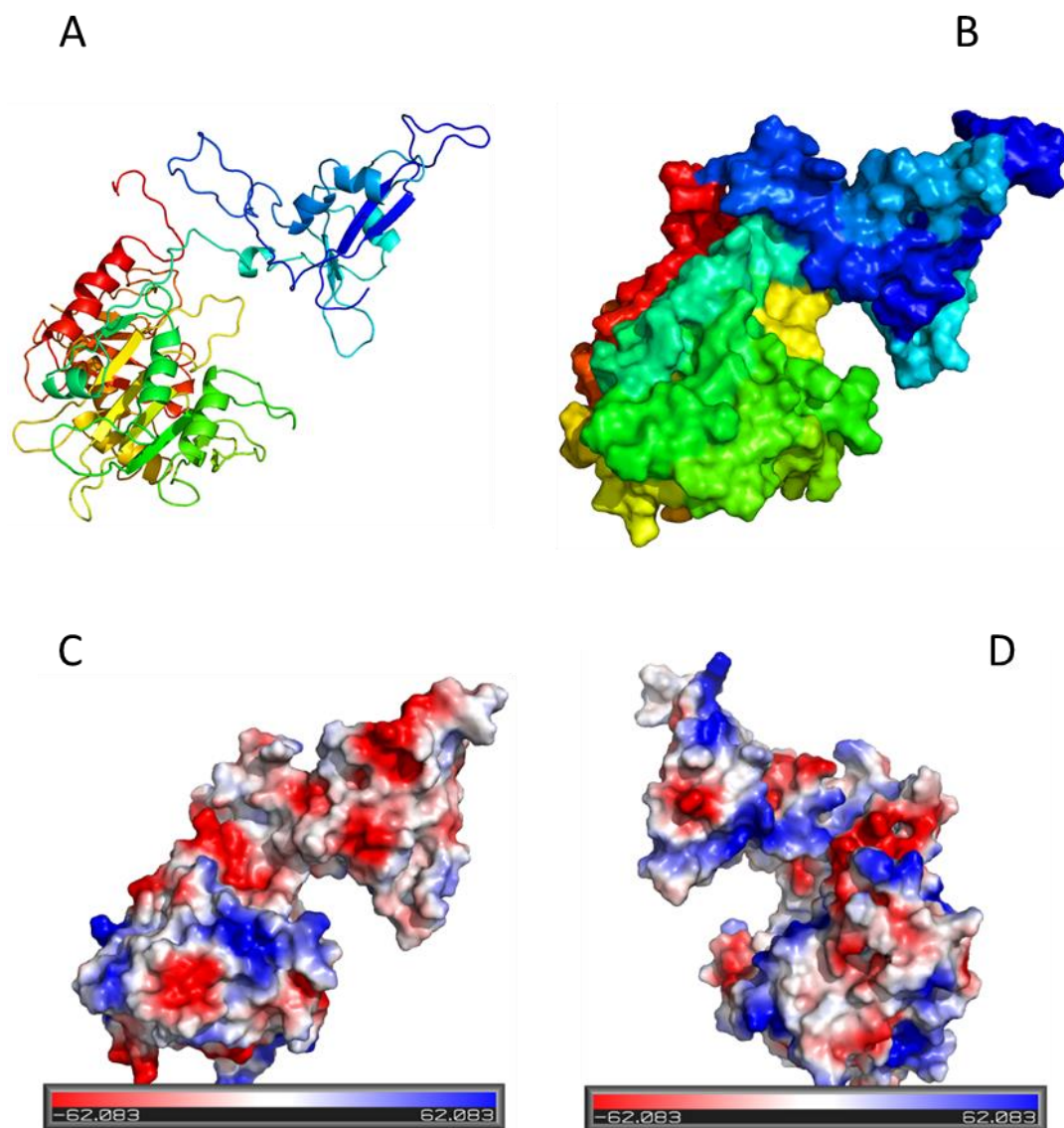


Figura 2: **Estrutura monomérica da proteína NS_s de *Groundnut ringspot orthotospovirus* (GRSV).** A representação das estruturas secundárias coloridas da região N-terminal (azul) para a região C-terminal (vermelha); B, representação da superfície da proteína com a mesma coloração em A; C, mapa da superfície eletrostática onde os resíduos de aminoácidos positivos, candidatos para interação com RNA, estão assinalados em azul e os resíduos negativos assinalados em vermelho; D, rotação de 180°C de C. As imagens foram geradas no programa PyMol.

1515 Acerca da possível interação com Argonauta 1, pelo modelo proposto por Jin e
 1516 Zhou (2010) uma proteína supressora viral pode interagir com a argonauta através do
 1517 motivo WG/GW (Fig. 3C). O grupo americano 1 de Tospovírus possui o motivo WG
 1518 (fig. 3A-B), porém o grupo americano II e o grupo euroasiático não possui esse motivo,

1519 apesar de a glicina estar altamente conservada. Talvez a NSs desses dois últimos grupos
 1520 tenham perdido a interação com a argonauta, suprimindo o silenciamento gênico apenas
 1521 pelo sequestro de dsRNA ou talvez a NSs do grupo americano I tenha adquirido a
 1522 novidade de interagir com a argonauta, além de sequestrar dsRNA. É interessante notar
 1523 que os aminoácidos Fenilalanina (F) e Tirosina (Y) que substituem o resíduo Triptofano
 1524 (W) também são aromáticos (Fig.3B).

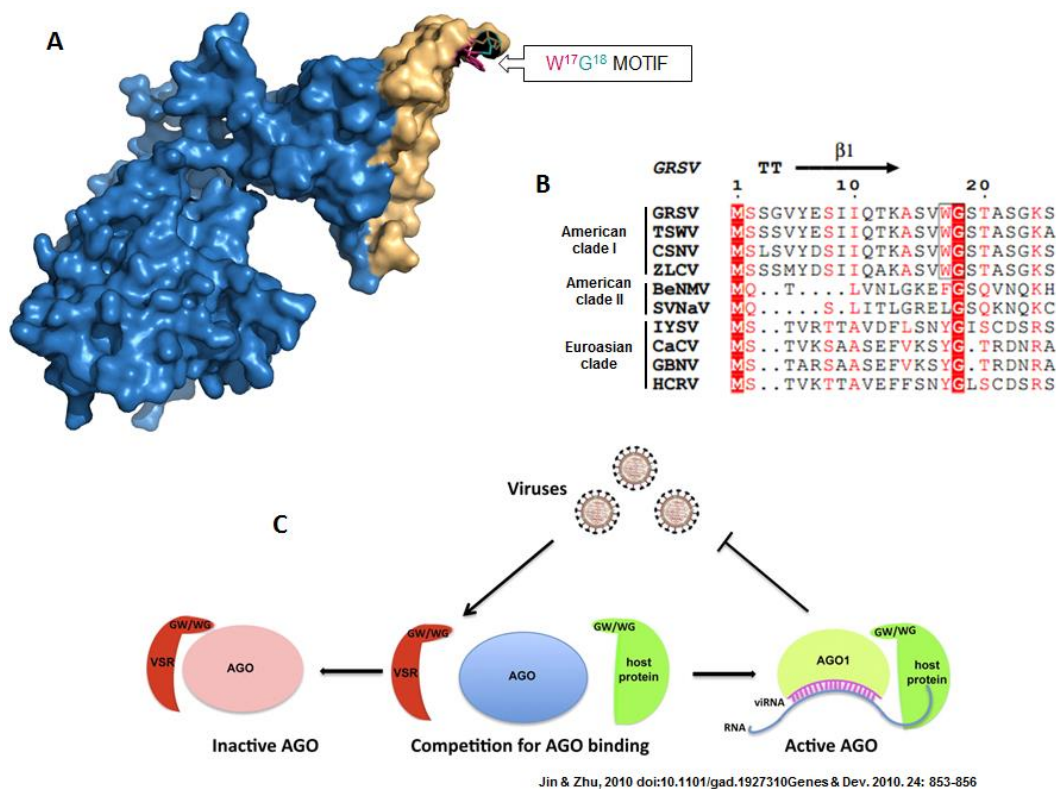


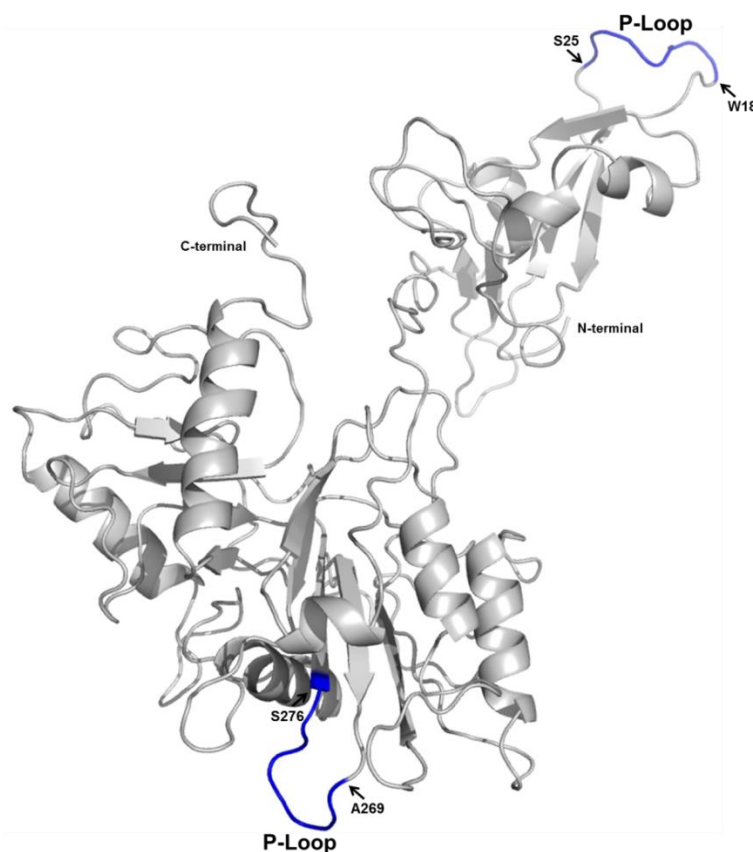
Figura 3: **Representação esquemática da hipótese de interação do motivo GW/WG com a proteína Argonauta.** A, modelo tridimensional predito da proteína NS_s salientando o motivo WG na sua porção N-terminal; B, alinhamento das sequências de aminoácidos N-terminais das proteínas NS_s de diversos orthotospovírus indicado as mutações em cada clado; C, ilustração retirada de [3] que fornece a hipótese de inibição da argonauta mediada pelo motivo GW/WG.

1525 Zhai et al. (2013), por meio de mutações de ponto sítio dirigidas na NSs de
 1526 TSWV e expressão transiente em plantas de *Nicotiana benthamiana* 16C, encontraram
 1527 dois motivos conservados entre os Tospovírus, envolvidos na supressão do
 1528 silenciamento gênico: GKV/T (aminoácidos 181-183) e YL (aminoácidos 412 e 413).

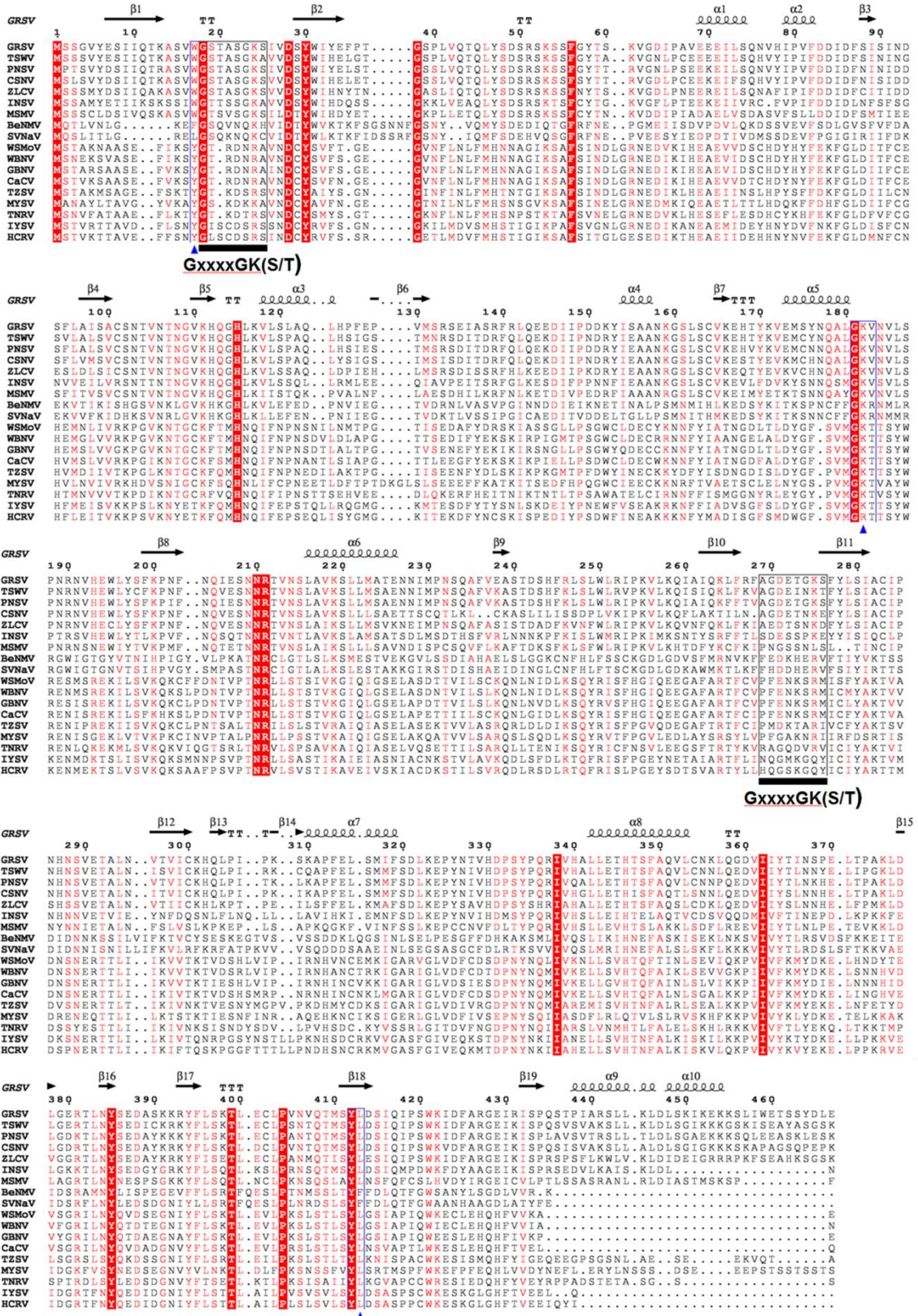
1529 As mutações K182A e L413A induzidas nesses domínios eliminaram a atividade
1530 supressora da NS_S [12]. Predições, *in silico*, apontam a presença de dois motivos para
1531 interação com ATP, denominados P-Loop ou Motif A ((GxxxxGK(S/T)) (Fig. 4 e 5). Essas
1532 duas regiões são fortes candidatas para o estudo da interação com RNA e da atividade
1533 de NTPase/fosfatase.

1534 A partir dos dados estruturais gerados neste trabalho, estudos futuros acerca
1535 desses ensaios mutagênicos poderão ser realizados para NS_S de GRSV. Análises dos
1536 domínios responsáveis pela fosforilação estão planejadas e poderão fornecer novas
1537 pistas sobre essa atividade enzimática.

1538



1539 Figura 4: **Motivos preditos de ligação à GTP/ATP da proteína NS_S de *Groundnut ringspot***
1540 ***orthotospovirus* (GRSV).** As alças para interação com GTP ou ATP, conhecidas como *P-Loop* ou *Motif*
1541 *A* estão destacadas em azul. As regiões foram preditas por meio do programa PROSITE [13].



1542

Figura 5: Alinhamento entre as sequências da proteína NS_s de algumas espécies de orthotospovírus. Resíduos altamente conservados entre as duas espécies estão sombreados em vermelho e escritos em branco e os resíduos com as mesmas características estão escritos em vermelho. Os P-Loops (GxxxxGK(S/T)) preditos estão delimitados por quadrados pretos, os motivos WG, GKV e YL estão delimitados por quadrados azuis. Os resíduos já mutados que resultaram na perda do bloqueio do silenciamento gênico estão indicados por triângulos azuis.

1543

1544 **4.2 Expressão da proteína NS_S de GRSV**

1545

1546 **4.2.1 Metodologia**

1547 A sequência de DNA que codifica a proteína NS_S havia sido previamente clonada
1548 no vetor pGEM® -T *Easy* (Promega), segundo as instruções do fabricante. Sítios de
1549 restrição foram inseridos em cada um dos genes: para *PciI* (ACATGT) na extremidade
1550 5' e para *XhoI* (CTCGAG) na extremidade 3'.

1551 Esse vetor foi então digerido com as enzimas de restrição *PciI* (NEB) e *XhoI*
1552 (NEB) utilizando-se o tampão NEB 3.1 de acordo com o protocolo de digestão dupla
1553 fornecido pelo aplicativo do fabricante ([https://www.neb.com/tools-and-
1554 resources/interactive-tools/double-digest-finder](https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder)).

1555 O vetor de expressão pet28a (material suplementar) foi digerido com as enzimas
1556 de restrição *NcoI* (CCATGG) (NEB) e *XhoI* (NEB) utilizando-se o tampão NEB 3.1,
1557 também de acordo com o protocolo de digestão dupla fornecido pelo fabricante. As
1558 enzimas *NcoI* e *PciI* possuem extremidades compatíveis, por isso não foi necessário
1559 sintetizar oligonucleotídeos e reamplificar os genes já clonados em pGEM® -T *Easy*.
1560 Todas as digestões ocorreram a 37°C por 16 horas.

1561 Essa estratégia permitiu a deleção das duas caudas de 6xhistidina que seriam
1562 fusionadas às porções N e C-terminais das proteínas. Dessa forma, o vetor foi
1563 modificado para expressão heteróloga somente das proteínas de interesse.

1564 Os produtos das digestões foram aplicados em gel de agarose (AMRESCO®) 1%,
1565 submetidos à eletroforese com tampão de corrida TBE 0,5x (45mM de Tris-borato;
1566 1mM de EDTA), corados em solução de brometo de etídeo (0,5µg/ml) e analisados no
1567 programa *Alpha Innotech AlphaImager®Mini*. Os fragmentos com tamanho esperado

1568 foram extraídos do gel com lâmina de bisturi e purificados seguindo instruções do kit
1569 comercial GFX PCR DNA *and Gel Band Purification* (GE Healthcare).

1570 O inserto (NS_S) e o vetor pET28a digeridos e purificados foram submetidos à
1571 reação de ligação, com a enzima T4 DNA ligase (Promega), de acordo com as
1572 instruções do fabricante. Os produtos da ligação foram utilizados para a transformação
1573 de células DH10BTM (Invitrogen) eletrocompetentes de acordo com o manual Sambrook
1574 & Russel (2000). Colônias positivas foram selecionadas em meio LB (Luria-Bertani)
1575 seletivo com o antibiótico canamicina (50 µg/ml).

1576 O vetor recombinante pET28a + NS_S foi isolados e purificado com o kit
1577 comercial Wizard® *Plus SV Minipreps DNA Purification Systems* (Promega) seguindo
1578 as instruções do fabricante. Em sequência, foi utilizado para a transformação de células
1579 de expressão Shuffle T7[®] (C3026-NEB) quimiocompetentes de acordo com o manual
1580 do fabricante.

1581 **4.2.2 Resultados e Discussão**

1582

1583 As colônias selecionadas foram inoculadas em 10 ml de meio LB, incubadas
1584 durante 16 horas, a 30°C, com o antibiótico canamicina (50 µg/mL), com agitação de
1585 180 rpm.

1586 Posteriormente, 5ml do pré- inóculo foi utilizado para inocular 100 ml de meio
1587 LB disposto em um frasco de Erlenmeyer de 1L. Os frascos seguiram sob agitação de
1588 180 rpm, a 30°C por aproximadamente duas horas até OD_{600nm} de 0,4. A expressão
1589 das proteínas foi induzida por meio da adição de IPTG (isopropil-β-
1590 tiogalactopiranosídeo) 0,5 mM e a temperatura foi abaixada até que o meio de cultura
1591 atingisse 16 °C.

1592 Alíquotas de 5ml do meio foram coletadas no momento da adição de IPTG (T₀h,
 1593 tempo-zero), quatro horas após a indução (T₄hs) e 16 horas após a indução. Após a
 1594 indução, as células foram coletadas em centrifugação a 6000 x g por 10 minutos, e
 1595 ressuspendidas em 10mL de tampão fosfato de sódio 20mM. As células foram
 1596 sonicadas no aparelho Bandelin Sonopuls (Sigma) no gelo em 04 ciclos de 30 segundos
 1597 com intervalos de 10 segundos. Centrifugou-se a amostra por 30 minutos a 6000 x g à
 1598 4°C. As frações celulares solúvel (sobrenadante) e insolúvel (precipitado) foram
 1599 separadas, fervidas a 98°C por 10 minutos com tampão de amostra Laemmli 5x e
 1600 submetidas ao SDS-PAGE.

1601 A corrida de eletroforese foi realizada por aproximadamente 150 minutos, à
 1602 temperatura ambiente, em corrente de 25 mA e voltagem de 130 V. Após a corrida, o
 1603 gel de poliacrilamida 12% foi colocado em uma solução de *Coomassie-blue* para
 1604 coloração por trinta minutos à temperatura ambiente. A remoção do excesso de corante
 1605 foi obtida em solução descorante até o aparecimento nítido das bandas de proteínas.

1606 As propriedades das proteínas foram preditas *in silico* para a escolha dos tampões,
 1607 do pH e das colunas de troca iônica (Tabela 1).

1608 Tabela 1. Propriedades da proteína NS_S de GRSV.

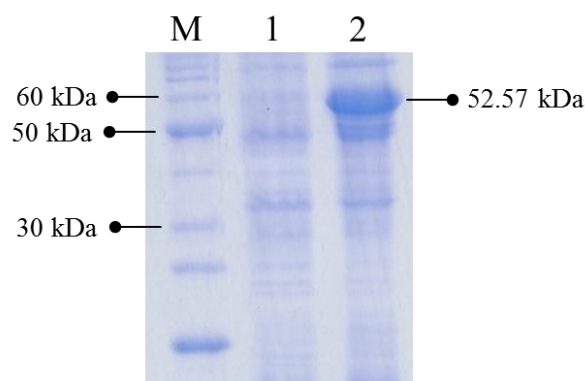
	Tamanho (pb)	Tamanho (aa)	Peso Molecular* (kDa)	Ponto Isoelétrico* (PI)
NS_S	1404	467	52.57	6,90

1609 * valores computados por meio do programa Compute pI/Mw (disponível em:
 1610 http://web.expasy.org/compute_pi/)
 1611
 1612

1613 A fração celular insolúvel (*pellet*) foi ressuspendida em 5ml de tampão fosfato
 1614 de sódio pH 6.0, e realizada subsequentes lavagens adicionando Triton x-100 1% (v/v),
 1615 e NaCl 1M e DTT 1mM, seguidas por sonicação de 2 ciclos de 30 segundos (Bandelin

1616 Sonopuls –Sigma), e centrifugação de 20 minutos a 13.000 x g à 4°C, com o descarte
1617 do sobrenadante, no intuito de retirar as proteínas insolúveis de *E.coli*. Seguiu-se a
1618 solubilização da proteína NSs do corpúsculo de inclusão adicionando-se 5ml de tampão
1619 fosfato de sódio pH 6.0 adicionado de uréia 6M, na ressuspensão do *pellet* obtido da
1620 última lavagem, e permanecendo 60 minutos em ultrasonicador (Branson 1510),
1621 seguido por centrifugação por 20 minutos a 13.000 x g à 4°C. O sobrenadante obtido
1622 (5ml) foi filtrado (0.22 µm) e aplicado em coluna de troca iônica Hitrap SP XI (GE
1623 Healthcare Life Sciences) equilibrada no tampão A (6M Uréia, Fosfato de sódio, pH
1624 6.0), fluxo de 1ml/min, utilizando o sistema cromatográfico automatizado “Akta
1625 purifier” (GE, Uppsala, Sweden). Proteínas não ligadas foram eluídas em 10 ml do
1626 tampão A, e proteínas ligadas foram eluídas em 20 ml do tampão B (1M NaCl, Fosfato
1627 de sódio, pH 6.0) em frações de 2 ml. Frações correspondentes aos picos de proteínas
1628 (280nm) foram analisadas em SDS-PAGE.

1629 A proteína NS_s foi expressa nas condições descritas acima (Fig. 6). Por
1630 conseguinte, as frações celulares solúvel (sobrenadante) e insolúvel (precipitado) foram
1631 analisadas. Em sua totalidade, as proteínas se encontram na fração celular insolúvel, ou
1632 seja, na forma de corpúsculos de inclusão (dados não mostrados).



1633

1634 **Figura 6: Expressão heteróloga da proteína NS_s de *Groundnut ringspot orthotospovirus***
1635 **(GRSV).** Gel SDS-PAGE 12% mostrando (1) fração não induzida, (2) fração induzida com 0,5 mM
de IPTG após 4 horas e (M) marcador de peso molecular BenchMark™ Protein Ladder (Ref.
10747012-Invitrogen™).

1636 As proteínas foram obtidas por expressão heteróloga, porém encontram-se na
1637 forma de corpúsculos de inclusão. Assim sendo, após o rompimento das células
1638 bacterianas e descarte da fração solúvel, a fração insolúvel foi ressuspensa várias
1639 vezes em tampão de lavagem, descrito na metodologia, para a eliminação de grande
1640 parte de proteínas que não são de interesse.

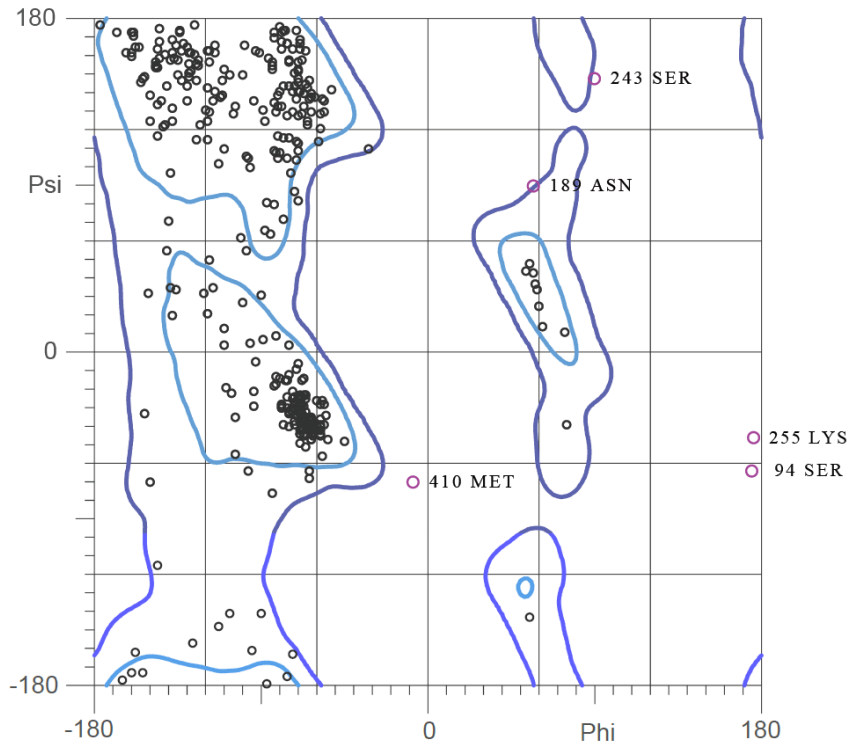
1641 Seguiu-se a solubilização da proteína NSs do corpúsculo de inclusão
1642 adicionando-se 10ml de tampão fosfato de sódio pH 6.0 adicionado de uréia 8M, na
1643 ressuspensão do pellet obtido da última lavagem, e permanecendo 60 minutos em
1644 ultrasonicador (Branson 1510), seguido por centrifugação por 30 minutos a 13.000 x g à
1645 4°C. O sobrenadante obtido (5ml) foi filtrado (0.22 µm) e aplicado em coluna de troca
1646 iônica Hitrap SP XI (GE Healthcare Life Sciences) equilibrada no tampão A (6M Uréia,
1647 Fosfato de sódio, pH 6.0), fluxo de 1ml/min, utilizando o sistema cromatográfico
1648 automatizado “Akta purifier” (GE, Uppsala, Sweden). Frações correspondentes ao pico
1649 (280nm) da proteína NS_S purificada foram analisadas em SDS-PAGE.

1650

1651 **4.3 Material suplementar.**

1652 **4.3.1 Figura S1. Gráfico de Ramachandran para a proteína NS_S de GRSV**

1653



1654

1655 **4.3.2 Mapa do vetor de expressão pet28a.**

1656

pET-28a-c(+) Vectors

TB074 12/98

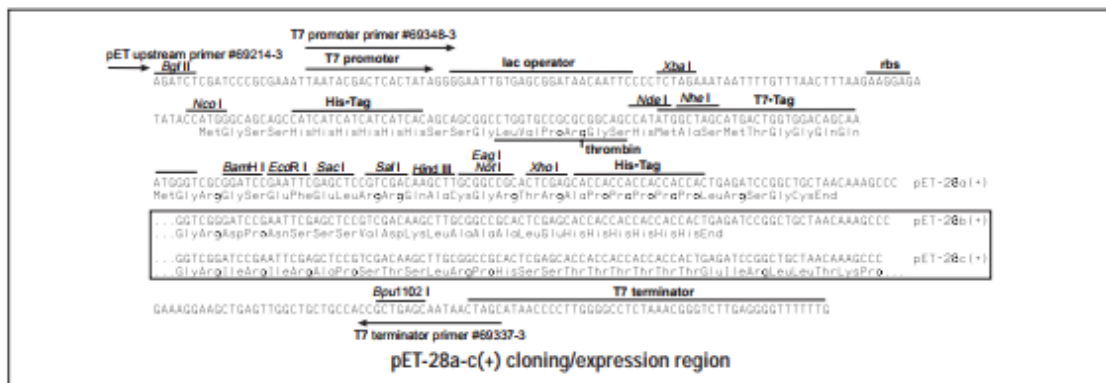
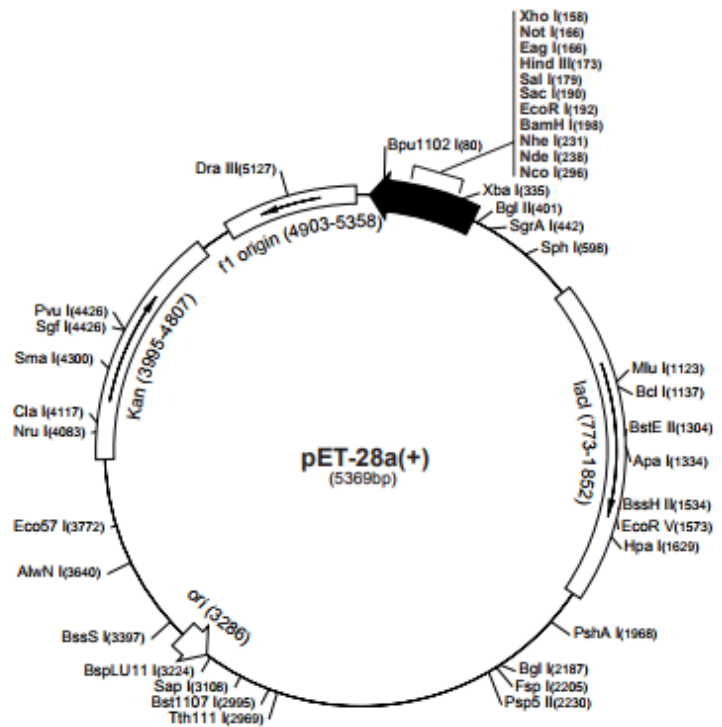
	Cat. No.
pET-28a DNA	69864-3
pET-28b DNA	69865-3
pET-28c DNA	69866-3

The pET-28a-c(+) vectors carry an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The fl origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 69337-3).

pET-28a(+) sequence landmarks

T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (<i>Bam</i> H I - <i>Xho</i> I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lacI</i> coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
fl origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.



1657

1658

1659 4.4 Referências

1660

1661

1662 1. Hallwass M, Leastro MO, Lima MF, Inoue-Nagata AK, Resende RO: **Sequence**
1663 **determination and analysis of the NSs genes of two tospoviruses.** *Archives of*
1664 *virology* 2012, **157**(3):591-596.

1665 2. Chen CC, Hwang JK, Yang JM: **(PS)2-v2: template-based protein structure**
1666 **prediction server.** *BMC Bioinformatics* 2009, **10**:366.

1667 3. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and**
1668 **accurate multiple sequence alignment.** *Journal of molecular biology* 2000,
1669 **302**(1):205-217.

1670 4. Robert X, Gouet P: **Deciphering key features in protein structures with the**
1671 **new ENDscript.** *Nucleic acids research* 2014, **42**(Web Server issue):W320-324.

1672 5. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ,
1673 Murray LW, Richardson JS, Richardson DC: **MolProbity: all-atom structure**
1674 **validation for macromolecular crystallography.** *Acta crystallographica*
1675 *Section D, Biological crystallography* 2010, **66**(Pt 1):12-21.

1676 6. Lahiri T, Singh K, Pal MK, Verma G: **Protein structure validation using a**
1677 **semi-empirical method.** *Bioinformatics* 2012, **8**(20):984-987.

1678 7. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D,
1679 Karplus K: **Improving physical realism, stereochemistry, and side-chain**
1680 **accuracy in homology modeling: Four approaches that performed well in**
1681 **CASP8.** *Proteins* 2009, **77 Suppl 9**:114-122.

1682 8. Lokesh B, Rashmi PR, Amruta BS, Srisathiyannarayanan D, Murthy MR,
1683 Savithri HS: **NSs encoded by groundnut bud necrosis virus is a bifunctional**
1684 **enzyme.** *PloS one* 2010, **5**(3):e9757.

1685 9. Schnettler E, Hemmes H, Huisman R, Goldbach R, Prins M, Kormelink R:
1686 **Diverging affinity of tospovirus RNA silencing suppressor proteins, NSs, for**
1687 **various RNA duplex molecules.** *Journal of virology* 2010, **84**(21):11542-
1688 11554.

1689 10. de Ronde D, Butterbach P, Lohuis D, Hedil M, van Lent JW, Kormelink R: **Tsw**
1690 **gene-based resistance is triggered by a functional RNA silencing suppressor**
1691 **protein of the Tomato spotted wilt virus.** *Molecular plant pathology* 2013,
1692 **14**(4):405-415.

1693 11. de Ronde D, Pasquier A, Ying S, Butterbach P, Lohuis D, Kormelink R:
1694 **Analysis of Tomato spotted wilt virus NSs protein indicates the importance**
1695 **of the N-terminal domain for avirulence and RNA silencing suppression.**
1696 *Molecular plant pathology* 2014, **15**(2):185-195.

1697 12. Zhai Y, Bag S, Mitter N, Turina M, Pappu HR: **Mutational analysis of two**
1698 **highly conserved motifs in the silencing suppressor encoded by tomato**
1699 **spotted wilt virus (genus Tospovirus, family Bunyaviridae).** *Archives of*
1700 *virology* 2013.

1701 13. Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V,
1702 Bairoch A, Hulo N: **PROSITE, a protein domain database for functional**
1703 **characterization and annotation.** *Nucleic acids research* 2010,
1704 **38**(suppl_1):D161-D166.

1705

1706 **Capítulo V: Interações entre proteína virais e proteínas do**
1707 **hospedeiro**

1708 As interações entre as próprias proteínas do vírus e as interações entre proteínas
1709 virais e proteínas do hospedeiro são essenciais e obrigatórias para o sucesso de uma
1710 infecção. Todos os vírus recrutam proteínas do hospedeiro para sua entrada, replicação,
1711 tradução, evasão do sistema imune do hospedeiro e disseminação na célula, assim como
1712 para a montagem do capsídeo e o movimento célula a célula.

1713 **5.1 Busca por proteínas de plantas candidatas para interação** 1714 **com as proteínas N, NS_M e NS_S de GRSV.**

1715

1716 Comparados aos outros membros da ordem *Bunyvirales*, uma das aquisições
1717 adaptativas dos orthotospovírus foi uma proteína de movimento (NS_M) cuja função é
1718 promover o movimento viral de uma célula a outra pelos plasmodesmas, uma vez que as
1719 células vegetais possuem parede celular que bloqueia os processos de endocitose e
1720 brotamento [1,2]. Estudos de sublocalização celular observaram que a NS_M está
1721 localizada nos plasmodesmas e se oligomeriza formando longos túbulos proteicos [3-9],
1722 mas ainda não foram identificadas interações com proteínas plasmodesma-localizadas
1723 da planta hospedeira. Diversas proteínas celulares já foram identificadas interagindo
1724 com proteínas de movimento virais [10-18]. Neste trabalho, as proteínas candidatas
1725 selecionadas para os testes de interação com a NS_M de GRSV foram a AtPDLP1,
1726 AtPDLP5, AtCalreticulina, AtDNAJ, AtHSPRO1.

1727 As proteínas localizadas ao plasmodesma (PDLPs) contribuem para a formação de
1728 túbulos dos vírus *Cauliflower mosaic virus* (CaMV) e *Grapevine fanleaf virus* (GFLV)
1729 [19-20]. As proteínas, como a PDLP1 e a calreticulina, mostraram uma interação com
1730 MPs virais e a proteína P6 de CaMV co-imunoprecipitou com PDLP1 [19,21].

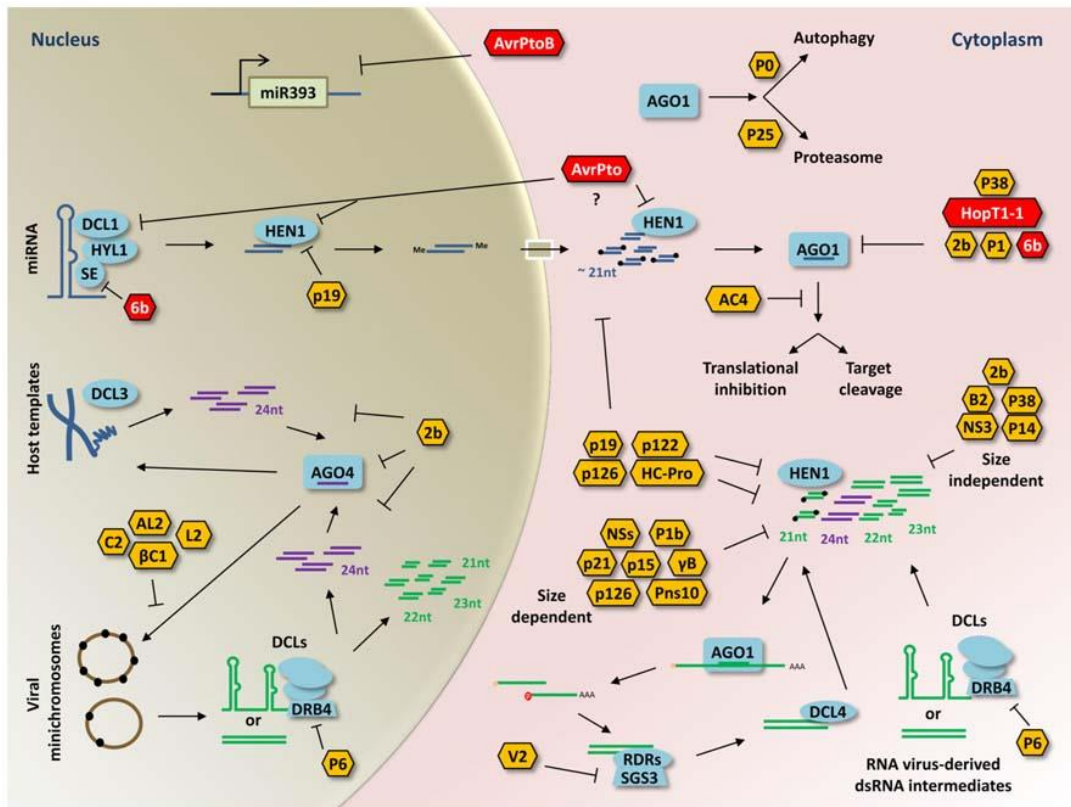
1731 As proteínas AtMiosina-IX e AtActina-VII foram selecionadas como candidatas
1732 porque foi demonstrado que a proteína N de TSWV trafega por essa via [22].
1733 AtHSPRO1 é uma proteína que promove resistência ao nematoide *Heterodera schachtii*
1734 [23] e foi selecionada porque altos níveis de transcrito desse gene foram encontrados

1735 em plantas infectadas com orthospovirus INSV [25]. Estas três proteínas candidatas
1736 serão testadas com as proteínas N, NS_M e NS_S.

1737 Outra aquisição importante dos orthospovirus foi uma proteína supressora de
1738 silenciamento gênico (NS_S), que permite ao vírus evadir as respostas de defesa das
1739 plantas hospedeiras e dos vetores para se replicar e sobreviver dentro do organismo.

1740 Acredita-se que a NS_S interfere na resposta do sistema imune da planta
1741 sequestrando siRNAs e dsRNA longos, evitando, assim, o recrutamento do complexo
1742 RISC. Também foi demonstrado que a porção N-terminal (133 aminoácidos) da NS_S é
1743 importante para a função de silenciamento gênico e pode ser a região candidata chave
1744 para a possível interação da NS_S com a proteína Argonauta1. Uma mutação no motivo
1745 GW/WG, correspondentes aos aminoácidos 17 e 18, levou à perda da atividade de
1746 silenciamento gênico em plantas expressando GFP [26,27].

1747 Baseando-se nessa hipótese de interação com Argonauta 1 e no fato de várias
1748 proteínas supressoras virais interagirem com proteínas da via de silenciamento gênico
1749 da planta (Fig.1), diversas proteínas foram escolhidas como candidatas para interação
1750 direta com a NS_S [24]: AtArgonauta1, AtArgonauta7, AtDCL1, AtDCL4, AtDRB4,
1751 AtHEN1, AtRDR6 e AtSGS3.



1752

Figura 1. Representação esquemática da via de silenciamento gênico vegetal. As proteínas da planta que atuam na via estão marcadas em azul enquanto que proteínas supressoras dessa via estão marcadas em amarelo. Figura retirada de [24].

1753

1754 5.2 Metodologia

1755

1756 5.2.1 Amplificação dos genes

1757

1758 Folhas de *Arabidopsis thaliana* sadias e agroinfiltradas com vetor viral foram
 1759 utilizadas para extração de RNA total de acordo com as recomendações do fabricante
 1760 usando o reagente comercial TRIzol® da Invitrogen™ (Cat. No. 15596-018). Os
 1761 cDNAs para os genes das proteínas candidatas para interação foram sintetizados
 1762 utilizando a enzima transcriptase reversa Super Script III (Invitrogen) de acordo com as
 1763 recomendações do fabricante. Os genes foram amplificados através da reação em cadeia

1764 de polimerase (PCR) utilizando-se a Platinum® Taq DNA Polymerase (Invitrogen™)
1765 para os genes de até 2500 pb e LongAmp (NEB) , com a utilização dos pares de
1766 oligonucleotídeos específicos listados na Tabela 1. Adaptando a técnica de clonagem ao
1767 vetor de destino, a cada gene foram inseridos nucleotídeos para manutenção da ordem
1768 de leitura e sítios para enzimas de restrição.

1769 Os produtos das PCR foram aplicados em gel de agarose (AMRESCO®) 1% e
1770 submetido à eletroforese com tampão de corrida TBE 0,5x (45mM de Tris-borato; 1mM
1771 de EDTA), corados em solução de brometo de etídeo (0,5µg/ml). Os fragmentos com
1772 tamanho esperado foram extraídos do gel com lâmina de bisturi e purificados seguindo
1773 instruções do kit GFX PCR DNA and Gel Band Purification (GE Healthcare).

1774

1775 **5.2.3 Clonagem dos insertos no vetor pGEM® -T easy (Promega)**

1776

1777 Os genes eluídos foram ligados no vetor de entrada para produtos de PCR
1778 pGEM® -T Easy (Promega) de acordo com as instruções do fabricante. A ligação entre
1779 os insertos o vetor pGEM® -T Easy foi feita com a utilização da T4 DNA Ligase
1780 (Promega) permanecendo overnight a 4°C. *E. coli* da linhagem DH10B
1781 eletrocompetentes foram transformadas com os plasmídeos recombinados.

1782 **5.2.4 Digestão dos vetores recombinantes pGEM + insertos**

1783

1784 Esses vetores foram então digeridos com as enzimas de restrição da NEB,
1785 respeitando os sítios adicionados nos oligonucleotídeos, utilizando-se o tampão NEB de
1786 acordo com o protocolo de digestão dupla fornecido pelo aplicativo do fabricante
1787 (<https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder>). Os

1788 produtos das digestões foram aplicados em gel de agarose (AMRESCO®) 1% e
1789 submetidos à eletroforese com tampão de corrida TBE 0,5x (45mM de Tris-borato;
1790 1mM de EDTA) e corados em solução de brometo de etídeo (0,5µg/ml). Os fragmentos
1791 com tamanho esperado foram extraídos do gel com lâmina de bisturi e purificados
1792 seguindo instruções do kit GFX PCR DNA and Gel Band Purification (GE Healthcare).
1793

1794 **5.2.5 Clonagem nos vetores pGADT7 AD e pGBKT7 para ensaio de** 1795 **Y2H**

1796

1797 Os vetores foram então digeridos com as enzimas de restrição da NEB,
1798 respeitando os sítios adicionados nos insertos, utilizando o tampão NEB de acordo com
1799 o protocolo de digestão dupla fornecido pelo aplicativo do fabricante
1800 (<https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder>). Os
1801 produtos das digestões foram aplicados em gel de agarose (AMRESCO®) 1% e
1802 submetidos à eletroforese com tampão de corrida TBE 0,5x (45mM de Tris-borato;
1803 1mM de EDTA), corados em solução de brometo de etídeo (0,5µg/ml). Os fragmentos
1804 com tamanho esperado foram extraídos do gel com lâmina de bisturi e purificados
1805 seguindo instruções do kit GFX PCR DNA and Gel Band Purification (GE Healthcare).

1806 Os insertos eluídos foram ligados nos vetores de acordo com as instruções do
1807 protocolo da T4 DNA Ligase (Promega) permanecendo overnight a 16°C. *E. coli* da
1808 linhagem DH10B eletrocompetentes foram transformadas com as ligações de acordo
1809 com Sambrook & Russel (2000).

1810 Os ensaios de duplo híbrido em levedura foram realizados de acordo com o
1811 protocolo fornecido no Curso De Treinamento: Sistema de Duplo Híbrido em Levedura

1812 para o estudo das interações proteína-proteína (Ciba-Embrapa). O Protocolo foi
1813 traduzido e está em anexo no final deste capítulo (Material suplementar 1).

Gene	Código de acesso no GenBank	Peso molecular do gene (pb)	Sítios de restrição adicionados aos oligonucleotídeos	Oligonucleotídeo Senso	Oligonucleotídeo antissenso
<i>AtActina VII</i>	NM_121018	1134	<i>NdeI</i> e <i>XhoI</i>	FWACT7: CATATGGCGGATGGTGAGGATATTCAG	REVACT7: CTCGAGTCATTACACCAATTCCTCATCTGTAAATTT
<i>AtArgonauto 1</i>	NM_001198240	3153	<i>XmaI</i> e <i>XhoI</i>	F_AG1XM: CCCGGGATGGTGAGAAAGAGAAGACGGATGCTCCA	R_AG1XH: CTCGAGTCAGCAGTAGAACATGACACGCCTTCACATTT
<i>AtArgonauto 7</i>	NM_105611	2973	<i>NdeI</i> e <i>SalI</i>	F_AG7ND: CATATGGAAGAAAAAATCATCATCATCATCACAGTACC	R_AG7ND: GTCGACTCAGCAGTAAAAACATGAGATTCCTTGACATTT
<i>AtCalreticulina</i>	NM_104513	1278	<i>NdeI</i> e <i>XhoI</i>	FWCRT: CATATGGCGAAACTAAACCCATAATTCATCTCT	REYCR1: CTCGAGTTAGGCTCGTCTCATGGCGGCATCGGTTTC
<i>AIDCL1</i>	JN661702	5661	<i>NdeI</i> e <i>XmaI</i>	F_DCAIND: CATATGGTAATGGAGGATGAGCCTAGAGAGCCACAA	R_DCAIXM: CCCGGGTCAAGAAAAAAGTTTTATTTAAAAAGCTCAAG
<i>AIDCL4</i>	NM_001203419	5067	<i>NdeI</i> e <i>SalI</i>	FW_YH_AIDC4: CATATGCGGTGACGAAAGTTGACTT	RV_YH_AIDC4: GTCGACTCAGCAAAAGGAAATCCAGAAT
<i>AIDnaJ</i>	NM_118990	1047	<i>NdeI</i> e <i>XhoI</i>	FWDNI:CATATGGGAGTTGATTATTACAAGGTTTACAGGHTG	REVDNI:CTCGAGTCACCCAATAAGCTTCTTAAACCCAGCTTT
<i>AIDrb4</i>	AY150509	1068	<i>NdeI</i> e <i>XhoI</i>	DR4_FWN: CATATGGATCATGTATACAAAAGGTCAACTGCAA	DR4_REX: CTCGAGTTATGGCTTCCAAAGACGATAGGCTATAA
<i>AIHEN1</i>	NM_001203853	2829	<i>EcoRI</i> e <i>XhoI</i>	FW_YH_AIHEN1: CCCGGGTATGCCCGGTGGTGGGAAGCA	RV_YH_AIHEN1: CCCGGGTCAAAGATCAGTCTTTTTCT
<i>AIMosina IX</i>	NM_122056	4638	<i>Clat</i> e <i>BamHI</i> (<i>pGAD</i>) <i>BamHI</i> e <i>PstI</i> (<i>pGBK</i>)	FWMYOGAD:ATCGATACATGAAGAAACTGAGGAAATGCCGAGTGT FWMYOGBK:GGATCCGATGAAGAAACTGAGGAAATGCCGAGTGT	REVMYOGAD: GGATCCTTACGATGACTGCCTTCTTTACGTTGAG REVMYOGBK:CTGGCAGTTACGATGACTGCCTTCTTTACGTTGAG
<i>AIPdip1</i>	NM_123765	912	<i>NdeI</i> e <i>BamHI</i>	FWPDI: CATATGAAACTCACCTATCAATTTCAFTCTCTGG	REVPDI: GGATCCTCAATAAGCATCAATTTTACTTCTTCTTC
<i>AIPdip5</i>	NM_105737	900	<i>NdeI</i> e <i>XhoI</i>	PD5_FND:CATATGATCAAGACAAAGACGACGTCCTCTCTCT PD5_FBH: GGATCCATATGAAACTCACCTATCAATTTCTCACTTC	PD5_RXH: CTCGAGTCATTTACACCAATTTCTCATCTTGTAAATTT
<i>AIRDR6</i>	NM_114810	3591	<i>EcoRI</i> e <i>XhoI</i>	F_ATRD6EIGAATTCTATGGGGTCAGAGGAAATATGAAGAAGTCG	R_ATRD6XH: CTCGAGTTAGACGCTGAGCAAGAAACTTAGCCAA
<i>AISGS3</i>	NM_122263	1878	<i>NdeI</i> e <i>XhoI</i>	FW_YH_SL5G3: CATATGAGTTTCAGCAAAATGGGGTGGGA	RV_YH_SL5G3: GGATCCTCACTGGTGCCACTGCTAT

1816 **5.3 Resultados e Discussão**

1817

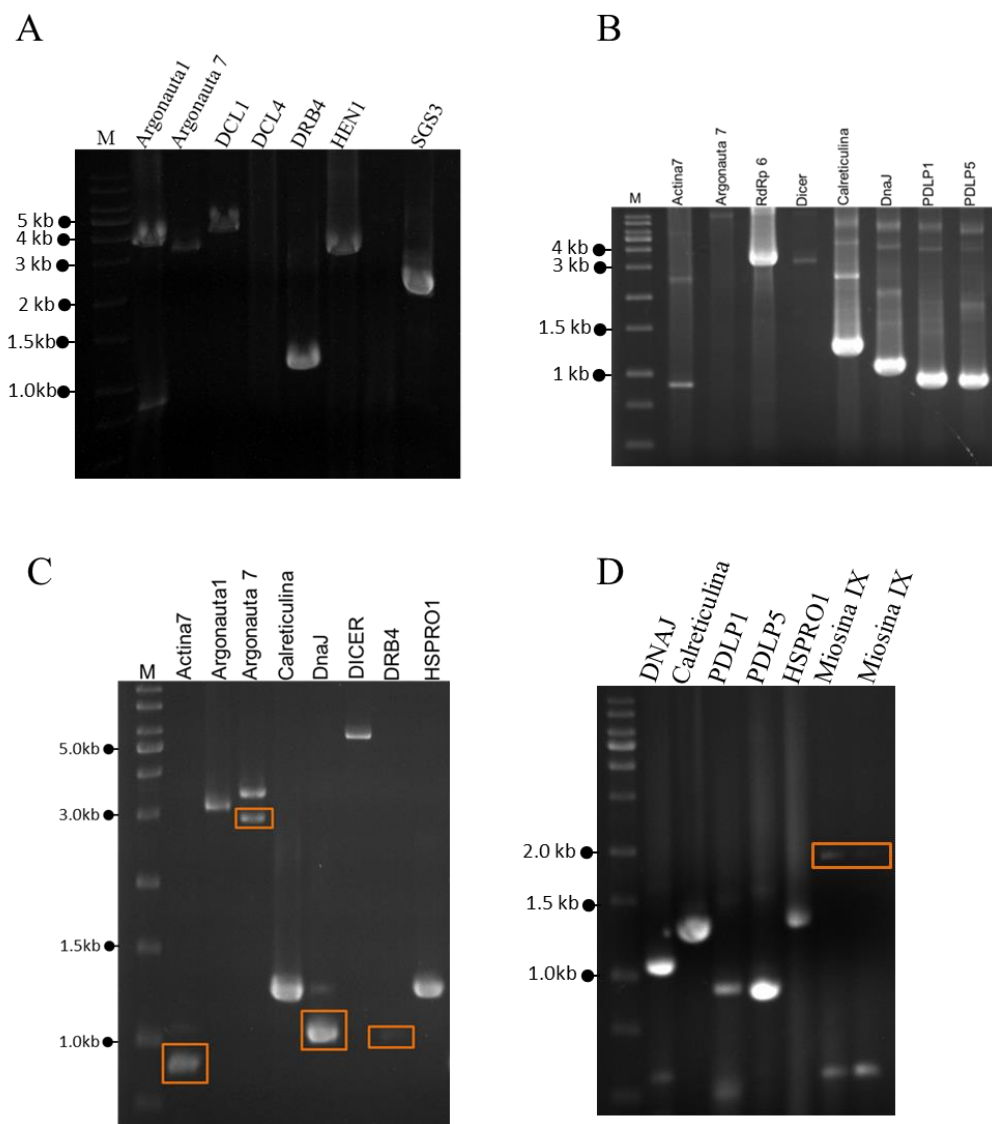
1818 **5.3.1 Amplificação dos genes das proteínas candidatas**

1819

1820 Por meio de RT-PCR seguida de PCR, os fragmentos dos genes: Argonauta 1,

1821 Argonauta 7, Calreticulina, DCL1, Dnaj, Drb4, Hspro1, PDPL1, PDPL5, Hen1 e

1822 SGS3 foram amplificados (Fig.2).



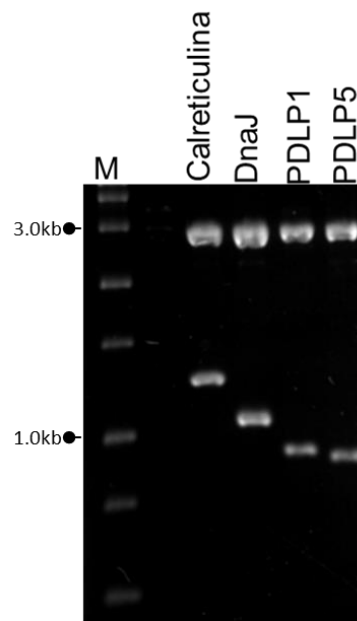
1823

Figura 2. **Amplificação dos genes das proteínas candidatas para interação com N, NS_M e NS_S.** Marcado em alaranjado estão as bandas escolhidas para eluição e clonagem. Em B e C Dicer corresponde à DCL1. Gel de agarose 1%. Marcador 1 kb ladder Ludiwig Biotec.

1824

1825 Os genes DCL4 e Miosina IX ainda não foram amplificados e diferentes
1826 protocolos e enzimas estão sendo testados.

1827 Os genes Calreticulina, DnaJ, PDLP1 e PDLP5 foram inseridos no vetor de
1828 entrada para produtos de PCR pGEM-T *easy* (3015pb) (Fig.3). Os insertos liberados
1829 serão utilizados para ligação nos vetores pGADT7 AD e pGBK T7 (Fig. S1).



1830 Figura 3. **Digestão dos vetores recombinantes inseridos + pGEM-T *easy***. Gel de agarose 1%.
1831 Marcador 1 kb *ladder* Ludiwig Biotec. Calreticulina (1278 pb), DnaJ (1047 pb), PDLP1 (912pb),
1832 PDLP5 (900pb) e pGEM-T *easy* (3000 pb).

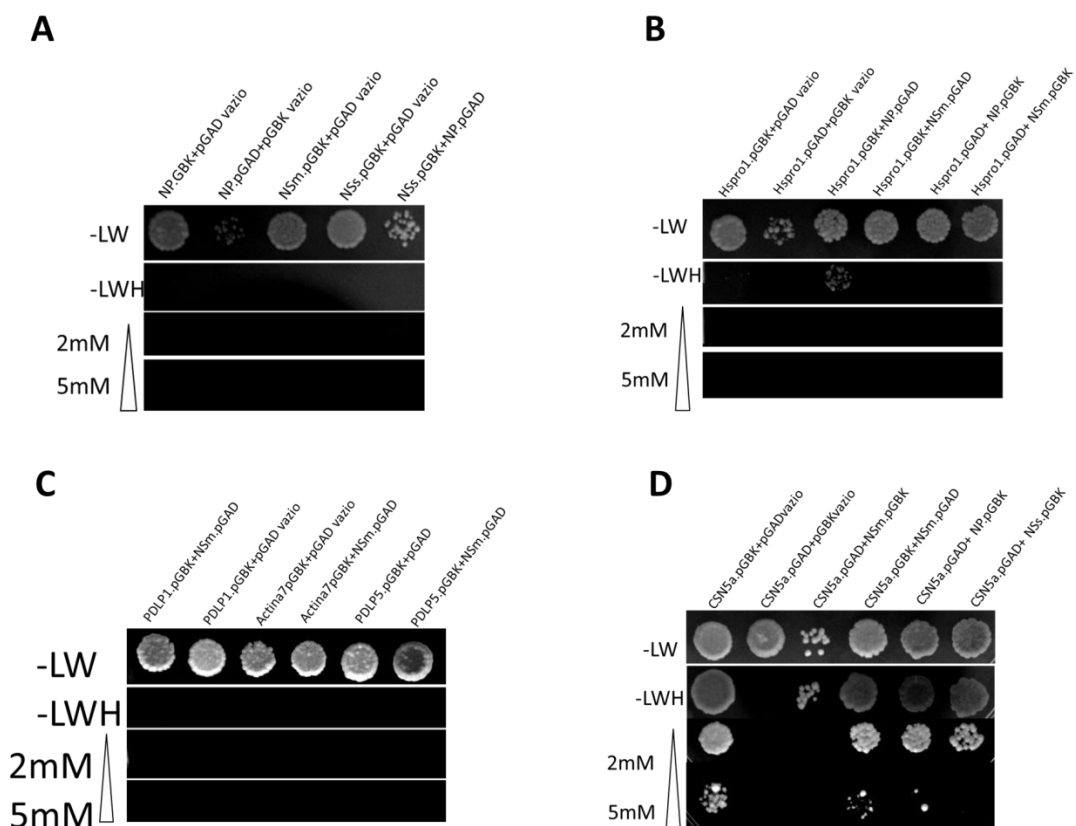
1831
1832

1833 5.3.2 Interações proteína-proteína

1834

1835 De acordo com o encontrado para TSWV, as proteínas NP e NSs não interagem
1836 com as proteínas do hospedeiro testadas neste ensaio (Fig.4A). As proteínas HSPRO1,
1837 PDLP1 e PDLP5 e Actina 7 não interagiram com as proteína NP, NSm e NSs de GRSV,
1838 visto que não foi observada nenhuma colônia visível nos meios seletivos contendo 3AT
1839 (2mM e 5 mM) (Fig.4 B-C). As demais proteínas presentes na Tabela 1, até o presente
1840 momento, ainda estão em fase de teste (dados não mostrados).

1841 Em relação à possível interação entre NSm.pGBK e CSN5a.pGAD, houve
 1842 crescimento de colônias no meio seletivo na ausência de histidina, (-LWH) porém, no
 1843 meio seletivo com 3AT (2mM e 5mM) não foi verificado crescimento, concluindo-se
 1844 que se trata de um falso positivo. No entanto, vale ressaltar que possíveis interações
 1845 entre a NP.pGBK e a CSN5a.pGAD e entre a NSs.pGBK e a CSN5a.pGAD foram
 1846 identificadas pelo crescimento de colônias de levedura nos meios seletivos contendo
 1847 3AT (Fig. 4D).



1848

1849 **Figura 4. Ensaio de Duplo híbrido em Levedura (Y2H) para interações proteína-proteína entre a**
 1850 **NP, NSm e NSs de GRSV e *A. thaliana*.** Meios seletivos: -LW, ausente em leucina e Triptofano; -LWH,
 1851 ausentes em leucina, Triptofano e histidina; 2mM, -LWH com presença de 3AT na concentração de
 1852 2mM; 5mM, -LWH com presença de 3AT na concentração de 5mM.

1853

1854 O gene da CSN5a está localizado no cromossomo 1 de *A. thaliana* e a proteína
 1855 codificada atua no desenvolvimento morfológico e nas respostas de defesa da planta por
 1856 meio da degradação de proteínas via ubiquitinação e produção de ácido jasmônico [28,

1857 29]. Por ser uma proteína que ativa as defesas da planta, a ausência de interação com a
1858 proteína NSm pode ser um resultado importante para ser analisado. Estudos futuros
1859 poderão confirmar se, de fato, essas interações são relevantes no contexto do sistema
1860 vírus/planta.

1861 5.4 Referências

- 1862 1. Feng Z, Xue F, Xu M, Chen X, Zhao W, Garcia-Murria MJ, Mingarro I, Liu Y,
1863 Huang Y, Jiang L *et al*: **The ER-Membrane Transport System Is Critical for**
1864 **Intercellular Trafficking of the NSm Movement Protein and Tomato**
1865 **Spotted Wilt Tospovirus**. *PLoS pathogens* 2016, **12**(2):e1005443.
- 1866 2. Zhao W, Jiang L, Feng Z, Chen X, Huang Y, Xue F, Huang C, Liu Y, Li F, Liu
1867 Y *et al*: **Plasmodesmata targeting and intercellular trafficking of Tomato**
1868 **spotted wilt tospovirus movement protein NSm is independent of its**
1869 **function in HR induction**. *The Journal of general virology* 2016, **97**(8):1990-
1870 1997.
- 1871 3. Leastro MO, Pallas V, Resende RO, Sanchez-Navarro JA: **The movement**
1872 **proteins (NSm) of distinct tospoviruses peripherally associate with cellular**
1873 **membranes and interact with homologous and heterologous NSm and**
1874 **nucleocapsid proteins**. *Virology* 2015, **478c**:39-49.
- 1875 4. Tripathi D, Raikhy G, Goodin MM, Dietzgen RG, Pappu HR: **In vivo**
1876 **localization of iris yellow spot tospovirus (Bunyaviridae)-encoded proteins**
1877 **and identification of interacting regions of nucleocapsid and movement**
1878 **proteins**. *PloS one* 2015, **10**(3):e0118973.
- 1879 5. Tripathi D, Raikhy G, Pappu HR: **Movement and nucleocapsid proteins coded**
1880 **by two tospovirus species interact through multiple binding regions in**
1881 **mixed infections**. *Virology* 2015.
- 1882 6. Lewandowski DJ, Adkins S: **The tubule-forming NSm protein from Tomato**
1883 **spotted wilt virus complements cell-to-cell and long-distance movement of**
1884 **Tobacco mosaic virus hybrids**. *Virology* 2005, **342**(1):26-37.
- 1885 7. Li W, Lewandowski DJ, Hilf ME, Adkins S: **Identification of domains of the**
1886 **Tomato spotted wilt virus NSm protein involved in tubule formation,**
1887 **movement and symptomatology**. *Virology* 2009, **390**(1):110-121.
- 1888 8. Leastro MO, Pallas V, Resende RO, Sanchez-Navarro JA: **The functional**
1889 **analysis of distinct tospovirus movement proteins (NSM) reveals different**
1890 **capabilities in tubule formation, cell-to-cell and systemic virus movement**
1891 **among the tospovirus species**. *Virus research* 2017, **227**:57-68.
- 1892 9. Storms MM, Kormelink R, Peters D, Van Lent JW, Goldbach RW: **The**
1893 **nonstructural NSm protein of tomato spotted wilt virus induces tubular**
1894 **structures in plant and insect cells**. *Virology* 1995, **214**(2):485-493.
- 1895 10. Geng C, Cong QQ, Li XD, Mou AL, Gao R, Liu JL, Tian YP: **Developmentally**
1896 **regulated plasma membrane protein of Nicotiana benthamiana contributes**
1897 **to potyvirus movement and transports to plasmodesmata via the early**

- 1898 **secretory pathway and the actomyosin system.** *Plant physiology* 2015,
1899 **167(2):394-410.**
- 1900 11. Jiang J, Patarroyo C, Garcia Cabanillas D, Zheng H, Laliberte JF: **The Vesicle-**
1901 **Forming 6K2 Protein of Turnip Mosaic Virus Interacts with the COPII**
1902 **Coatomec Sec24a for Viral Systemic Infection.** *Journal of virology* 2015,
1903 **89(13):6695-6710.**
- 1904 12. Levy A, Zheng JY, Lazarowitz SG: **The tobamovirus Turnip Vein Clearing**
1905 **Virus 30-kilodalton movement protein localizes to novel nuclear filaments**
1906 **to enhance virus infection.** *Journal of virology* 2013, **87(11):6428-6440.**
- 1907 13. Li Z, Zhang Y, Jiang Z, Jin X, Zhang K, Wang X, Han C, Yu J, Li D: **Hijacking**
1908 **of the nucleolar protein fibrillarlin by TGB1 is required for cell-to-cell**
1909 **movement of Barley stripe mosaic virus.** *Molecular plant pathology* 2017.
- 1910 14. Niehl A, Amari K, Gereige D, Brandner K, Mely Y, Heinlein M: **Control of**
1911 **Tobacco mosaic virus movement protein fate by CELL-DIVISION-CYCLE**
1912 **protein48.** *Plant physiology* 2012, **160(4):2093-2108.**
- 1913 15. Park MR, Jeong RD, Kim KH: **Understanding the intracellular trafficking**
1914 **and intercellular transport of potexviruses in their host plants.** *Frontiers in*
1915 *plant science* 2014, **5:60.**
- 1916 16. Rodriguez A, Angel CA, Lutz L, Leisner SM, Nelson RS, Schoelz JE:
1917 **Association of the P6 protein of Cauliflower mosaic virus with**
1918 **plasmodesmata and plasmodesmal proteins.** *Plant physiology* 2014,
1919 **166(3):1345-1358.**
- 1920 17. Verchot J: **Cellular chaperones and folding enzymes are vital contributors to**
1921 **membrane bound replication and movement complexes during plant RNA**
1922 **virus infection.** *Frontiers in plant science* 2012, **3:275.**
- 1923 18. Zhou Y, Rojas MR, Park MR, Seo YS, Lucas WJ, Gilbertson RL: **Histone H3**
1924 **interacts and colocalizes with the nuclear shuttle protein and the movement**
1925 **protein of a geminivirus.** *Journal of virology* 2011, **85(22):11821-11832.**
- 1926 19. den Hollander PW, Kieper SN, Borst JW, van Lent JWM: **The role of**
1927 **plasmodesma-located proteins in tubule-guided virus transport is limited to**
1928 **the plasmodesmata.** *Archives of virology* 2016, **161:2431-2440.**
- 1929 20. Amari K: **A Family of Plasmodesmal Proteins with Receptor-Like**
1930 **Properties for Plant Viral Movement Proteins.** 2010, **6(9).**
- 1931 21. Ishikawa K, Hashimoto M, Yusa A, Koinuma H, Kitazawa Y, Netsu O, Yamaji
1932 Y, Namba S: **Dual targeting of a virus movement protein to ER and plasma**
1933 **membrane subdomains is essential for plasmodesmata localization.** *PLoS*
1934 *pathogens* 2017, **13(6):e1006463.**
- 1935 22. Feng Z, Chen X, Bao Y, Dong J, Zhang Z, Tao X: **Nucleocapsid of Tomato**
1936 **spotted wilt tospovirus forms mobile particles that traffic on an**
1937 **actin/endoplasmic reticulum network driven by myosin XI-K.** *The New*
1938 *phytologist* 2013.
- 1939 23. Thureau T, Kifle S, Jung C, Cai D: **The promoter of the nematode resistance**
1940 **gene Hs1pro-1 activates a nematode-responsive and feeding site-specific**
1941 **gene expression in sugar beet (Beta vulgaris L.) and Arabidopsis thaliana.**
1942 *Plant molecular biology* 2003, **52(3):643-660.**
- 1943 24. Pelaez P, Sanchez F: **Small RNAs in plant defense responses during viral and**
1944 **bacterial interactions: similarities and differences.** *Frontiers in plant science*
1945 2013, **4:343.**
- 1946 25. Senthil G, Liu H, Puram VG, Clark A, Stromberg A, Goodin MM: **Specific and**
1947 **common changes in Nicotiana benthamiana gene expression in response to**

- 1948 **infection by enveloped viruses.** *The Journal of general virology* 2005, **86**(Pt
1949 9):2615-2625.
- 1950 26. Schnettler E, Hemmes H, Huisman R, Goldbach R, Prins M, Kormelink R:
1951 **Diverging affinity of tospovirus RNA silencing suppressor proteins, NSs, for**
1952 **various RNA duplex molecules.** *Journal of virology* 2010, **84**(21):11542-
1953 11554.
- 1954 27. de Ronde D, Pasquier A, Ying S, Butterbach P, Lohuis D, Kormelink R:
1955 **Analysis of Tomato spotted wilt virus NSs protein indicates the importance**
1956 **of the N-terminal domain for avirulence and RNA silencing suppression.**
1957 *Molecular plant pathology* 2014, **15**(2):185-195.
- 1958 28. Gusmaroli G, Feng S, Deng XW: **The Arabidopsis CSN5A and CSN5B**
1959 **Subunits Are Present in Distinct COP9 Signalosome Complexes, and**
1960 **Mutations in Their JAMM Domains Exhibit Differential Dominant**
1961 **Negative Effects on Development.** *The Plant cell* 2004, **16**(11):2984-3001.
- 1962 29. Stratmann JW, Gusmaroli G: **Many jobs for one good cop - the COP9**
1963 **signalosome guards development and defense.** *Plant science : an*
1964 *international journal of experimental plant biology* 2012, **185-186**:50-64.
- 1965
- 1966

1967 **5.5 Figuras Suplementares**

1968

1969 **S.1 Mapas dos vetores pGADT7 AD e pGBKT7**

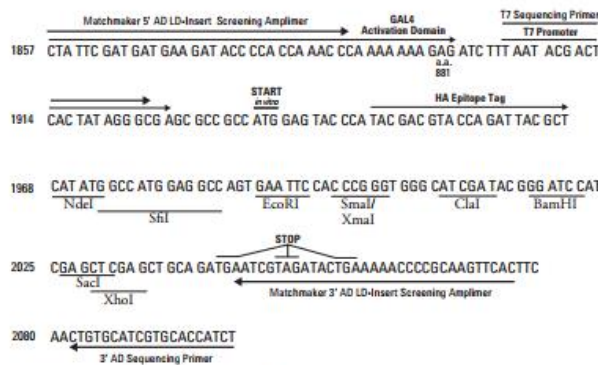
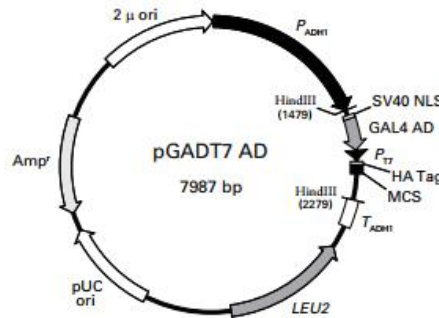
1970

1971

pGADT7 AD Vector Information

PT3249-5

Cat. Nos. 630442
630489
630491



pGADT7 AD Vector Map and Multiple Cloning Site (MCS).

Description

pGADT7 AD is a yeast expression vector that is designed to express a protein of interest fused to a GAL4 activation domain (AD; amino acids 768–881). Transcription of the GAL4 AD fusion is driven by the constitutively active *ADH1* promoter (P_{ADH1}), and is terminated at the *ADH1* transcription termination signal (T_{ADH1}). The GAL4 AD fusion contains an N-terminal SV40 nuclear localization signal (SV40 NLS; 1) that targets the protein to the yeast nucleus, and a hemagglutinin epitope tag (HA Tag), located between the GAL4 AD and the protein of interest, that allows the protein to be easily detected with HA-tag antibodies.

The T7 promoter (P_{T7}), located just upstream of the HA tag sequence, allows *in vitro* transcription and translation of the HA-tagged protein of interest (without the GAL4 AD and the SV40 NLS). pGADT7 AD replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 μ ori, respectively. The vector also contains an ampicillin resistance gene (*Amp^r*) for selection in *E. coli* and a *LEU2* nutritional marker for selection in yeast.

(010312)



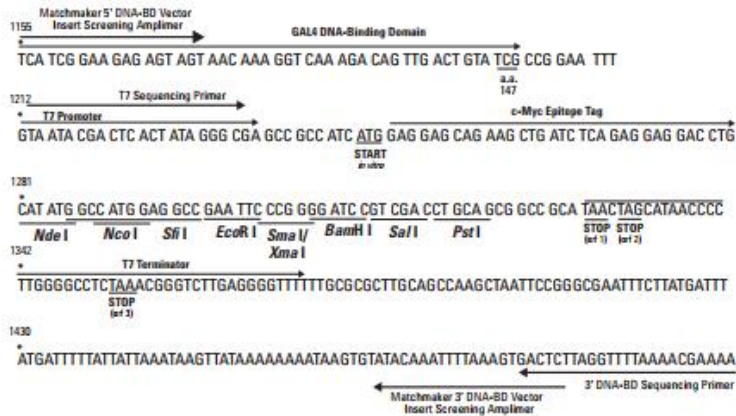
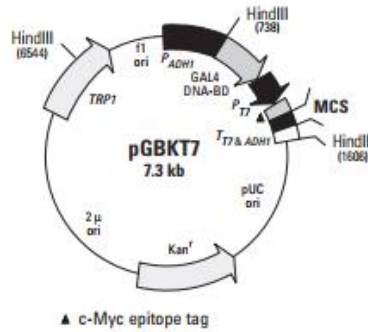
United States/Canada
800.552.2595
Asia Pacific
+1.650.919.7300
Europe
+33.(0)1.3904.6880
Japan
+81.(0)77.543.6116

Clontech Laboratories, Inc.
A Takara Bio Company
1290 Terra Bella Ave.
Mountain View, CA 94043
Technical Support (US)
E-mail: tech@clontech.com
www.clontech.com

1972
1973

pGBKT7 Vector Information

PT3248-5
 Cat. No. 630489
 630443



Restriction Map and Multiple Cloning Site (MCS) of pGBKT7. Unique restriction sites are in bold.

Description:

The pGBKT7 vector expresses proteins fused to amino acids 1–147 of the GAL4 DNA binding domain (DNA-BD). In yeast, fusion proteins are expressed at high levels from the constitutive *ADHI* promoter (*P_{ADHI}*); transcription is terminated by the T7 and *ADHI* transcription termination signals (*T_{T7 & ADHI}*). pGBKT7 also contains the T7 promoter, a c-Myc epitope tag, and a MCS. pGBKT7 replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 μ ori, respectively. The vector carries the Kan^r for selection in *E. coli* and the *TRP1* nutritional marker for selection in yeast. Yeast strains containing pGBKT7 exhibit a higher transformation efficiency than strains carrying other DNA-BD domain vectors (1).

(PRBY2643; published 5 Dec 2008)



Clontech

United States/Canada
 800.562.2566
 Asia Pacific
 +1.850.919.7300
 Europe
 +33.(0)1.3904.8880
 Japan
 +81.(0)77.543.6116

Clontech Laboratories, Inc.
 A Takara Bio Company
 1290 Terra Bella Ave.
 Mountain View, CA 94043
 Technical Support (US)
 E-mail: tech@clontech.com
 www.clontech.com

1974

1975

1976 **Conclusão geral e Perspectivas**

1977

1978 Por meio desse trabalho foi possível identificar as estruturas tridimensionais para
1979 as proteínas NP e NS_S de GRSV e para a NP de ZLCV, pois os principais motivos
1980 proteicos para oligomerização e interação com ácidos nucleicos foram preditos. Esses
1981 resultados abrem caminho para futuros estudos de mutações sítio dirigidas para
1982 compreensão da formação do ribonucleocapsídeo, formação das partículas e interação
1983 com demais proteínas do vírus e dos hospedeiros.

1984 Esse trabalho permitiu que a técnica de duplo híbrido em levedura fosse
1985 estabelecida no laboratório de Virologia Vegetal da Universidade de Brasília-UnB, o
1986 que possibilitará, não só os estudos de interação proteína-proteína para orthospovírus,
1987 como para os demais vírus estudados pelo nosso grupo de pesquisa.

1988 Por meio desse trabalho também foi possível estabelecer e aperfeiçoar um sistema
1989 de expressão de proteína nativa para a NSs de GRSV, passo essencial para a
1990 cristalização proteica visando a resolução da estrutura tridimensional dessa e de outras
1991 proteínas de orthospovírus.

