

CAMILA CHAVES PINA DE BARROS

**Sequenciamento do Genoma Completo e Expressão Heteróloga da
Capa Protéica do Marafivirus Associado à Morte Súbita dos Citros**

Orientador: Tatsuya Nagata

**Brasília
2006**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

CAMILA CHAVES PINA DE BARROS

**Sequenciamento do Genoma Completo e Expressão Heteróloga da
Capa Protéica do Marafivirus Associado à Morte Súbita dos Citros**

“Dissertação apresentada ao Programa de Pós-Graduação ‘Stricto Sensu’ em Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília, como requisito para a obtenção do Título de Mestre em Ciências Genômicas e Biotecnologia”.

Orientador: Tatsuya Nagata

Brasília

2006

B277s Barros, Camila Chaves Pina de
Sequenciamento do genoma completo e expressão heteróloga da capa
protéica do marafivirus associado à morte súbita dos citros / Camila Chaves
Pina de Barros; orientador Tatsuya Nagata – 2006.
175 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade Católica de Brasília, 2006.

1. Laranja – doenças e pragas. 2. Vírus de plantas. 3. Genoma.
I. Nagata, Tatsuya, orient. II. Título.

CDU 632.937

Ficha elaborada pela Divisão de Processamento do Acervo do SIBI – UCB.

TERMO DE APROVAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação como requisito parcial para obtenção do Título de Mestre em Ciências Genômicas e Biotecnologia, defendida e aprovada em 24 de março de 2006, pela banca examinadora composta por:

Dr. Waldir Cintra de Jesus Junior
Universidade Federal do Espírito Santo
Examinador Externo

Dra. Eliane Ferreira Noronha
Universidade Católica de Brasília
Examinadora Interna

Tatsuya Nagata
Universidade Católica de Brasília
Orientador

Dedico a minha família, em especial a minha mãe e ao meu pai, por tudo que são e representam pra mim e ao meu querido Guilherme pela força e companheirismo.

Ofereço aos meus sobrinhos como forma de incentivo para o crescimento pessoal e profissional.

Agradecimentos

Aos meus pais, incentivadores, guias e mestres pelo apoio incondicional e incentivo durante o curso, e aos meus irmãos que sempre torceram pelo meu sucesso.

Ao Professor Dr. Tatsuya Nagata pela idealização do trabalho e orientação.

Ao Dr. Waldir Cintra de Jesus Junior por toda colaboração, conhecimento e incentivo.

A Professora Dra. Eliane Ferreira Noronha pelo auxílio e presteza em todos os momentos que precisei.

A todos os professores do Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília, pelo aprimoramento da minha formação acadêmica.

Ao meu amor Guilherme Filho pela compreensão, incentivo e carinho em todas as horas.

Aos técnicos e colegas do Laboratório de Biotecnologia da Universidade Católica de Brasília, Idacuy Mundim, André Luiz Cardoso Ramos, Alessandra Reis, Willian Baião Reis e Márcia Ferreira de Araújo pelo auxílio e agradável convivência.

A todos os funcionários da UCB, em especial a Francisco Fábio Gomes da Costa e Danielle Cordeiro, pelo auxílio e presteza de sempre.

A todos os amigos que conquistei durante o mestrado, Carol Villarinho, Lílian Silveira, Patrícia Brunetta, Alexandre Figueira, Sandra Carvalho, Suzana Neiva, Letícia Fagundes, Alexandre Póvoa, Adriane, Andréa Tagliaferro, Keisiane, Joaquim Silva, Alexandre Figueira, Luciana Retori e Elisa Pavin pelas alegrias, carinho e incentivo nas horas boas e nas horas difíceis.

À Fundecitrus, Fundo de Defesa da Citricultura, pelo suporte financeiro concedido para a realização deste curso.

A todos colegas do grupo de pesquisa em virologia da UCB, Thais, Luisa, Priscila, Célia, Ana Carolina, Marcela, Natália, Clarissa e Wesley.

A todos os meus amigos que acompanharam nesta etapa pela dedicação, amizade, alegrias e amparo que me deram sempre que precisei: Luiza, Silvia, Ariane, Natália, Marcela, Raquel, Michely, Grazielle, Cláudio e Ana Karina.

A todos que contribuíram de alguma forma para o êxito deste trabalho.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xviii
Lista de Símbolos e Abreviações	xx
Resumo	xxii
Abstract.....	xxiv
1. Introdução	1
1.1 A Morte Súbita dos Citros (MSC).....	2
1.1.1 O porta-enxerto limoeiro ‘Cravo’	3
1.1.2 Sintomas da Morte Súbita dos Citros	4
1.1.3. Distribuição da doença	7
1.1.4 Manejo da doença.....	9
1.1.5 Etiologia	10
1.1.5.1 A Família <i>Tymoviridae</i>	14
1.2 <i>Pichia pastoris</i> Como Vetor de Expressão.....	19
2. Hipótese	22
3. Objetivos.....	24
3.1 Objetivo Geral	25
3.2 Objetivos Específicos	25
4. Material e Métodos.....	26
4.1 Clonagem e Seqüenciamento do Genoma do CSDaV.....	27
4.1.1 Extração do RNA viral	27

4.1.2 RT-PCR.....	27
4.1.2.1 Síntese de cDNA	27
4.1.2.2 Amplificação do cDNA por PCR.....	28
4.1.2.3 Confirmação da amplificação em gel de agarose	29
4.1.3 Eluição do fragmento amplificado a partir do gel de agarose	29
4.1.4 Ligação do cDNA ao vetor pCR4 TOPO	30
4.1.4.1 Adenilação	31
4.1.4.2 Ligação	32
4.1.5 Transformação de bactérias <i>E. coli</i> por eletroporação	32
4.1.6 Extração de DNA plasmidial.....	33
4.1.7 Verificação da presença do inserto.....	33
4.1.8 Amplificação da extremidade 5' do genoma viral.....	34
4.1.8.1 PCR Touchdown	35
4.1.9 Sequenciamento automático.....	37
4.1.10 Análise computacional das seqüências de DNA obtidas.....	38
4.2 Expressão da Proteína da Capa do Vírus.....	40
4.2.1 Amplificação da região genômica codificadora da capa do vírus.....	40
4.2.1.1 PCR.....	40
4.2.1.2 Confirmação da amplificação em gel de agarose	41
4.2.2 Ligação da região CP ao vetor de expressão	41
4.2.2.1 Digestão com <i>Eco</i> RI e <i>Xba</i> I	42
4.2.2.2 Eluição do vetor e inserto digeridos	43
4.2.2.3 Ligação	44
4.2.3 Transformação em <i>E. coli</i>	44
4.2.4 Extração de DNA plasmidial.....	45

4.2.5	Verificação da presença do inserto	45
4.2.6	Linearização do vetor	46
4.2.7	Transformação de <i>Pichia pastoris</i>	46
4.2.7.1	Preparo das células para eletroporação.....	46
4.2.7.2	Transformação por eletroporação	47
4.2.8	PCR de colônia	47
4.2.8.1	Confirmação da amplificação em gel de agarose	48
4.2.9	Indução da expressão da proteína heteróloga	49
4.2.10	DIBA	50
4.2.11	Indução da expressão da proteína heteróloga em maior escala	51
5.	Resultados e Discussão	52
5.1	Clonagem e Seqüenciamento do Genoma do CSDaV.....	53
5.1.1	Análise computacional das seqüências obtidas	59
5.1.1.1	Análise filogenética	60
5.1.1.2	Seqüência e organização genômica	64
5.1.1.3	Comparação entre os clones de CSDaV sequenciados.....	70
5.1.2	Citrus sudden death-associated virus (CSDaV)	73
5.2	Expressão da Proteína da Capa do Vírus.....	76
6.	Perspectivas.....	82
7.	Referências Bibliográficas	84
8.	Apêndices.....	96
	APÊNDICE A	97
	APÊNDICE B.....	99
	APÊNDICE C.....	115

APÊNDICE D	116
APÊNDICE E	117
APÊNDICE F	120
APÊNDICE G	120
APÊNDICE H	122
APÊNDICE I	126
APÊNDICE J	133
APÊNDICE K	136

Lista de Figuras

- Figura 1. Vista aérea de plantios de laranjeira-doce ‘Valência’ enxertada em limoeiro ‘Cravo’, apresentando sintomas da morte súbita dos citros, e de laranjeira-doce ‘Valência’ enxertada em tangerineira ‘Cleópatra’, sadias (A). Em B verificam-se duas linhas de plantio de laranjeira-doce ‘Valência’ enxertada em tangerineira ‘Cleópatra’, sem sintomas da doença no meio de um plantio de laranjeira-doce ‘Valência’ enxertada em limoeiro ‘Cravo’, totalmente afetado pela doença.....3
- Figura 2. Laranjeiras com sintomas da morte súbita dos citros mostrando perda de brilho generalizada e coloração verde-pálida das folhas, contrastando com a cor normal das plantas sadias (A) e folhas coletadas em plantas sadia e doente, mostrando o contraste de coloração entre as mesmas (B).4
- Figura 3. Plantas sadia e com sintomas de desfolha parcial, um dos sintomas da morte súbita dos citros.5
- Figura 4. Coloração amarelada nos tecidos internos da casca do porta-enxerto limão ‘Cravo’, na região do floema funcional, contrastando com os tecidos brancos da casca da copa (A) e comparação da coloração supra citada em A entre planta sadia e doente (B), mostrando que esse sintoma é típico de plantas com morte súbita dos citros.5
- Figura 5. Detalhe do sistema radicular de planta com sintoma inicial de morte súbita dos citros, o qual apresenta grande quantidade de raízes mortas, normalmente com poucas radículas (A) e detalhe da podridão que se inicia da ponta para a base da raiz (B).....6
- Figura 6. Mapa da distribuição da morte súbita dos citros em SP e MG. Fonte: Jesus Junior *et al.*, 2004b.9
- Figura 7. Subenxertia realizada em planta com seis anos de idade através da técnica do T invertido, utilizando-se dois cavalinhos.10

- Figura 8. Organização e expressão do genoma do *Turnip yellow mosaic virus* (TYMV), a espécie tipo do gênero *Tymovirus* mostrando a posição relativa das ORFs. MP: proteína de movimento. 15
- Figura 9. Estrutura genômica de um isolado de *Grapevine fleck virus* (GFkV), a espécie tipo do gênero *Maculavirus*, mostrando a posição relativa das ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CP, capa proteica; p31 e p16, proteínas ricas em prolina. 16
- Figura 10. Os dois tipos de estrutura genômica do gênero *Marafivirus*, exemplificado pelo *Maize rayado fino virus* (MRFV), a espécie-tipo deste gênero, e *Oat blue dwarf virus* (OBDV), mostrando a posição relativa das ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CPs, capas proteicas; p43, proteína rica em prolina. 17
- Figura 11. Eletromicrografia de partículas de CSDaV encontradas em plantas com sintomas de MSC..... 27
- Figura 12. Desenho do *primer* degenerado C7-For baseado nas regiões mais conservadas verificadas através do alinhamento múltiplo entre diferentes vírus da família *Tymoviridae*: ELV (*Erysimum latent virus*); TYMV (*Turnip yellow mosaic virus*); KYMV (*Kennedya yellow mosaic virus*); CMTV (*Chayote mosaic virus*); EMV (*Eggplant mosaic virus*); OYMV (*Ononis yellow mosaic virus*); PnMV (*Poinsettia mosaic virus*); OBDV (*Oat blue dwarf virus*) e MRFV (*Maize rayado fino virus*). ... 28
- Figura 13. Mapa do vetor de clonagem pCR4 TOPO, o qual apresenta a origem de replicação pUC, promotor lac, sítios múltiplos de clonagem, gene letal *ccdB*, sítio de ligação do inserto com única deoxitimidina nas extremidades 3' e genes de resistência a ampicilina e canamicina. 31

- Figura 14. Esquema mostrando as etapas de amplificação da extremidade 5' da genoma do CSDaV de acordo com o kit GeneRacer (Invitrogen).....35
- Figura 15. Mapa do vetor de expressão pPICZ α A, mostrando a origem de replicação do plasmídeo em *E. coli*, pUC; fragmento contendo o promotor AOX1; sinal de secreção α -factor; região de múltiplos sítios de clonagem; myc epitopo tag para detecção da proteína através de anticorpos; cauda de 6x Histidina para detecção e purificação da proteína; terminador de transcrição AOX1 TT; promotor do *transcription elongation factor 1* (TEF1) que dirige a expressão do gene de resistência a zeocina, e terminador de transcrição CYC1 de *S. cerevisiae*.....42
- Figura 16. A) Esquema representativo mostrando a posição relativa dos *primers* utilizados na reação de PCR e dos fragmentos obtidos; B) Fotografia de gel de agarose evidenciando o produto amplificado na RT-PCR utilizando a combinação de *primers* C7For e C7Rev. (1) Marcador 1Kb plus (Invitrogen); (2) fragmento de 4560pb amplificado nesta reação; C) Fotografia de gel de agarose evidenciando o produto amplificado na RT-PCR utilizando a combinação de *primers* C1 e M4. (1) Marcador 1Kb plus (Invitrogen); (2)fragmento de 1600pb amplificado nesta reação.....53
- Figura 17. Digestão plasmidial dos clones selecionados para o cDNA de 1600pb. M: Marcador 1Kb plus (Invitrogen); 8, 9, 14, 15 e 16: clones digeridos com a enzima *Eco* RI; 8', 9', 14', 15' e 16': clones não digeridos com a enzima *Eco* RI.....55
- Figura 18. Digestão plasmidial dos clones selecionados para os cDNA de 4650pb e 6000pb. M: Marcador 1Kb plus (Invitrogen); 1, 2, 3, 8, 9, 13 e 16: clones digeridos com a enzima *Eco* RI; 1', 2', 3', 8', 9', 13' e 16': clones não digeridos com a enzima *Eco* RI. Os clones 1, 3, 8 e 16 foram selecionados quanto à presença do fragmento de interesse. O clone 1 revelou inserto de 4560pb, ao passo que os clones 3, 8 e 16 apresentaram inserto de 6000pb.56

- Figura 19. Esquema representando o processo de subclonagens do genoma do CSDaV.56
- Figura 20. Eletroforese em gel de agarose evidenciando o produto amplificado na RT-PCR 5' RACE. (1) Marcador 1Kb plus (Invitrogen); (2) fragmento de 1020pb amplificado utilizando-se o *primer* reverso TymoGSP 1 e o *primer forward* GeneRacer™ 5' nested.....57
- Figura 21. Digestão plasmidial dos clones selecionados para o fragmento de 5' RACE. M: Marcador 1Kb plus (Invitrogen); 1', 4', 5', 9', 10' e 13': clones não digeridos com a enzima *Eco* RI; 1, 4, 5, 9, 10 e 13: clones digeridos com a enzima *Eco* RI. Os clones 1, 4, 5, 9 e 13 foram selecionados quanto à presença do fragmento 5' RACE (~1020pb).58
- Figura 22. Esquema representando a obtenção do genoma completo do CSDaV através de três subclonagens.....59
- Figura 23. Resultado da busca por similaridade nucleotídica através da ferramenta BLASTn, utilizando um consenso de 750pb obtido a partir da leitura de parte da seqüência genômica do vírus isolado de plantas com sintomas de morte súbita dos citros..... 60
- Figura 24. Árvore filogenética gerada pelo programa Mega 3 pelo método Neigbor-Joining a partir da seqüência nucleotídica do genoma completo do CSDaV (DQ185573) e do TYMV (*Turnip yellow mosaic virus*), MRFV (*Maize rayado fino virus*), OBDV (*Oat blue dwarf virus*), GFkV (*Grapevine fleck virus*) e PnMV (*Poinsetia mosaic virus*), disponíveis no GenBank. Os números indicados são os valores de bootstrap (%) de 1000. As distâncias foram calculadas pelo método Kimura-2.62
- Figura 25. Árvore filogenética gerada pelo programa Mega3 pelo método Neigbor-Joining a partir da seqüência nucleotídica codificadora da CP do CSDaV (DQ185573) e do GAMaV (*Grapevine asteroid mosaic-associated virus*), OBDV (*Oat blue dwarf*

virus), MRFV (*Maize rayado fino virus*), GFkV (*Grapevine fleck virus*), TYMV (*Turnip yellow mosaic virus*), e APLV (*Andean potato latent virus*), disponíveis no GenBank. Os números indicados são os valores de bootstrap (%) de 1000. A, B e C mostram o agrupamento dos vírus do gênero *Marafivirus*, *Maculavirus* e *Tymovirus*, respectivamente. As distâncias foram calculadas pelo método Kimura-2. 64

Figura 26. Alinhamento da seqüência nucleotídica de 16pb que compõe o “marafibox” (A) e do sítio de iniciação de transcrição (B) do sgRNA do CSDaV (DQ185573) com as regiões correspondentes de outras espécies do gênero *Marafivirus*, GAMaV (AJ249357), OBDV (U87832) e MRFV (AF265566), e o “tymobox” do TYMV (J04373). 66

Figura 27. Sítios de clivagem da *papain-like protease* entre os domínios da Helicase (Hel) e da RdRp (A) e entre o domínio da RdRp e da CP de 22,5 kDa (B), com a posição relativa dos aminoácidos na seqüência da poliproteína. A baixa numeração para o GAMaV é devido ao sequenciamento ainda incompleto deste vírus. O número de acesso da seqüência utilizada para cada vírus está identificada na figura. Os resíduos conservados entre os vírus comparados estão marcados em cinza. 67

Figura 28. Esquema representativo do genoma do CSDaV. A: desenho esquemático do genoma do CSDaV, mostrando a disposição das três ORFs ao longo do genoma; B: desenho esquemático da ORF1, mostrando a disposição dos domínios de proteína ao longo da mesma. 69

Figura 29. Estruturas genômicas do MRFV, OBDV e CSDaV, mostrando a posição relativa das ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CPs, capa proteica; p43, proteína rica em prolina. 70

- Figura 30. Resultado do alinhamento entre as seqüências nucleotídicas do genoma DQ185573 e AY884005 visualizado através do recurso Jalview do ClustalW, mostrando regiões de maior e menor homologia ao longo do genoma. 71
- Figura 31. Eletroforese em gel de agarose evidenciando a região codificadora da capa do vírus amplificada na PCR. M: Marcador 1Kb plus (Invitrogen); CP: amplificação da região CP mostrando fragmento de 650pb. 76
- Figura 32. Confirmação da presença do fragmento da capa protéica do vírus nos plasmídios extraídos dos clones selecionados de *E. coli*. M: Marcador 1Kb plus (Invitrogen); 1, 2, 3 e 4: clones digeridos com as enzimas *Eco* RI e *Xba* I; 1', 2', 3' e 4': clones não digeridos com as enzimas *Eco* RI e *Xba* I. 77
- Figura 33. Eletroforese em gel de agarose do produto da PCR de colônia de 12 clones de *P. pastoris* transformados com vetor pPICZ α A ligado ao gene CP. M: Marcador 1Kb plus (Invitrogen). 78
- Figura 34. Membrana de DIBA. Esta membrana mostra os tempos 24, 48, 72 e 96 horas após o início da indução da expressão da CP em 19 clones identificados na membrana e do controle negativo (CN). 79

Lista de Tabelas

Tabela 1.	Combinações de <i>primers</i> utilizadas nas reações de PCR para amplificação do cDNA viral.....	28
Tabela 2.	Descrição da quantidade de cada componente utilizado na reação de PCR.....	29
Tabela 3.	Descrição da reação de adenilação, mostrando seus componentes e suas respectivas concentrações.	31
Tabela 4.	Descrição da quantidade de cada componente utilizado na reação de PCR <i>touchdown</i>	36
Tabela 5.	Programa utilizado para a realização da PCR <i>touchdown</i>	37
Tabela 6.	Nome, seqüência e sentido dos <i>primers</i> utilizados para o seqüenciamento do genoma do CSDaV.....	38
Tabela 7.	Descrição da quantidade de cada componente utilizado na PCR para a amplificação da região CP.	40
Tabela 8.	Descrição da quantidade de cada componente utilizado na reação de digestão dupla do produto da amplificação da região CP com as enzimas de restrição <i>Eco</i> RI e <i>Xba</i> I..	43
Tabela 9.	Descrição da quantidade de cada componente utilizado na reação de digestão dupla do vetor pPICZ α A com as enzimas de restrição <i>Eco</i> RI e <i>Xba</i> I.....	43
Tabela 10.	Meio de cultura LB <i>low salt</i> utilizado para crescimento de bactérias transformadas com pPICZ α A.....	44
Tabela 11.	Descrição da PCR de colônia.	48
Tabela 12.	Resultado do alinhamento múltiplo entre a seqüência genômica completa obtida do CSDaV e outros vírus da família <i>Tymoviridae</i> , realizado com o software ClustalW. 61	
Tabela 13.	Resultado obtido do alinhamento múltiplo entre a seqüência nucleotídica codificadora da CP do CSDaV e de outros vírus da família <i>Tymoviridae</i> , realizado com o software ClustalW.....	63

Tabela 14. Comparação entre a seqüência de aminoácidos de cada domínio protéico dos CSDaV seqüenciados: DQ185573 e AY884005.	72
---	----

Lista de Símbolos e Abreviações

%	Porcentagem
aa	Aminoácido
AOX	<i>Alcohol oxidase</i>
APLV	<i>Andean potato latent virus</i>
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BELV	<i>Bermuda grass etched-line virus</i>
BeMV	<i>Belladonna mottle virus</i>
BMGY	<i>Buffered glycerol-complex medium</i>
BMMY	<i>Buffered methanol-complex medium</i>
ccdB	<i>Controller of Cell Death B protein</i>
cDNA	DNA complementar
CIP	<i>Calf intestinal fosfatase</i>
cm	Centímetro
CP	<i>Coat protein</i>
CSDaV	Citrus sudden death-associated virus
CTV	<i>Citrus tristeza virus</i>
CVC	Clorose variegada dos citros
DIBA	<i>Dot immunobinding assay</i>
DNA	Ácido desoxirribonucléico
dNTP	Desoxynucleosídeotrifosfato
EBI	<i>European Bioinformatics Institute</i>
<i>Eco</i>	<i>Escherichia coli</i>
For	<i>Forward</i>
g	Gramma
GAMaV	<i>Grapevine asteroid mosaic-associated virus</i>
GFkV	<i>Grapevine fleck virus</i>
GRGV	<i>Grapevine red globe virus</i>
GRVfV	<i>Grapevine rupestris vein feathering virus</i>
GSP	<i>Gene specific primer</i>
Hel	Helicase
ICTV	<i>International Committee on Taxonomy of Viruses</i>
Kb	Kilobase
kDa	Kilodalton
mg	Miligrama
mL	Mililitro
mM	Milimolar
MP	Proteína de movimento
MRFV	<i>Maize rayado fino virus</i>
mRNA	RNA mensageiro
MSC	Morte Súbita dos Citros
Mtr	Metiltransferase
NBT	<i>Nitroblue tetrazolium</i>

NCBI	<i>National Center for Biotechnology Information</i>
ng	Nanograma
nm	Nanômetro
nt	Nucleotídeo
OBDV	<i>Oat blue dwarf virus</i>
°C	Graus Celsius
ORF	<i>Open reading frame</i>
PAMV	<i>Potato aucuba mosaic virus</i>
pb	Par de base
PBS	<i>Phosphate buffered saline</i>
PCR	<i>Polymerase chain reaction</i>
<i>Pme</i>	<i>Pseudomonas mendocina</i>
PnMV	<i>Poinsettia mosaic virus</i>
Pol	Polimerase
Pro	<i>Papain-like protease</i>
RdRp	<i>RNA-dependent RNA polymerase</i>
Rev	<i>Reverse</i>
RNA	Ácido ribonucléico
Rpm	Rotações por minuto
RT PCR	<i>Reverse transcription – polymerase chain reaction</i>
sgRNA	RNA subgenômico
TAP	<i>Tobacco acid pyrophosphatase</i>
<i>Taq</i>	<i>Thermus aquaticus</i>
tRNA	RNA transportador
TYMV	<i>Turnip yellow mosaic virus</i>
U	Unidade enzimática
x g	Gravidade
<i>Xba</i>	<i>Xanthomonas badrii</i>
YNB	<i>Yeast nitrogen base with ammonium sulfate without amino acids</i>
YPD	<i>Yeast extract peptone dextrose medium</i>
YPDS	<i>Yeast extract peptone dextrose sorbitol medium</i>
µL	Microlitro

Resumo

A Morte Súbita dos Citros (MSC) é uma doença que causa sérias perdas econômicas nos Estados de São Paulo e Minas Gerais. Esses estados plantam cerca de 200 milhões de laranjeiras, principalmente para o processamento de suco de laranja. Até o momento mais de 4 milhões de laranjeiras foram afetadas por essa doença e a maioria delas foi erradicada ou teve severo declínio. O agente etiológico dessa doença ainda não está confirmado, mas foi sugerida a associação de um novo vírus da família *Tymoviridae*, gênero *Marafivirus*, com a MSC, chamado Citrus sudden death-associated virus (CSDaV). Para contribuir com a elucidação da etiologia dessa doença a caracterização do vírus é muito importante. Por esta razão, o genoma completo deste vírus isolado de laranjeira com MSC foi seqüenciado. O genoma do CSDaV possui 6806 nucleotídeos, excluindo-se a cauda poli(A) na extremidade 3' e codifica uma poliproteína de aproximadamente 240 kDa (p240) com motivos de seqüência similares às proteínas associadas à replicação (metiltransferase, *papain-like protease*, helicase e polimerase) presentes nas proteínas não estruturais de outros vírus com genoma de RNA senso positivo. A extremidade 3' da *open reading frame* (ORF) que codifica a poliproteína também pode codificar uma proteína do capsídeo (CP) de 22,5 kDa que é presumivelmente clivada da região C-terminal da poliproteína de 240 kDa. Uma segunda CP de 21 kDa é traduzida de um RNA subgenômico (sgRNA) localizado na extremidade 3' do genoma do CSDaV. Uma possível pequena ORF na extremidade 3' do genoma foi detectada através de análise computacional e possivelmente codifica uma proteína de função desconhecida, cuja seqüência apresenta baixa similaridade com a proteína de movimento encontrada em outros vírus da família *Tymoviridae*. A CP de 22,5 kDa foi expressa na levedura metilotrófica *Pichia pastoris* para possibilitar a sua utilização posterior como antígeno na produção de anticorpos para o diagnóstico da doença. A região CP deste vírus foi inserida no genoma da *P. pastoris* através do vetor pPICZ α A. A CP recombinante foi expressa como proteína de fusão

com uma cauda de polihistidina para posteriormente ser purificada utilizando-se resinas com afinidade a esta cauda. Através do ensaio de DIBA, foi verificado que onze clones obtidos secretaram altos níveis da proteína recombinante.

Palavras-chave: Morte súbita dos citros (MSC), Citrus sudden death-associated virus (CSDaV), *Marafivirus*, *Pichia pastoris*.

Abstract

Citrus sudden death (CSD) is an orange tree disease that causes serious economic losses in São Paulo and Minas Gerais States of Brazil. These states have approximately 200 million of orange trees, principally for orange juice processing. Until now more than 4 million oranges were affected with the disease. Most of them were eradicated or had declined. The etiologic agent of this disease is still unconfirmed, but it was suggested the association of a new virus of *Marafivirus* genus within *Tymoviridae* family, with Citrus sudden death disease, tentatively named Citrus sudden death-associated virus (CSDaV). To contribute for the etiology of CSD, CSDaV complete genome was sequenced in this study. The genome had 6806 nucleotides in length, excluding the 3'-terminal poly(A) tail and encodes a putative approximately 240 kDa polyprotein (p240) with sequence motifs similar to the replication-associated proteins (methyltransferase, papain-like protease, helicase and polymerase) present in nonstructural proteins of other virus with positive strands RNA genomes. The 3' end of the open reading frame (ORF) that encodes p240 also encodes a 22,5 kDa capsid protein (CP) that is presumably cleaved from p240 polyprotein C-terminus. A second CP (21 kDa) can be also translated from a subgenomic RNA (sgRNA). A second small ORF at the 3' end of the CSDaV genome was found by computer analysis and possibly encodes a protein with unknown function but with limited sequence similarities to putative movement proteins of tymoviruses. The 22,5 kDa coat protein (CP) gene was expressed in the methylotrophic yeast *Pichia pastoris* to be used as antigen to produce antiserum against CSDaV CP for disease diagnostic. A cDNA encoding the CSDaV CP was inserted into the *P. pastoris* genome using the pPICZ α A vector. The recombinant CSDaV CP was expressed as a fusion protein with an N-terminal hexa-histidine tag to be purified using metal-chelating resins. By DIBA it was verified that eleven clones secreted high level of recombinant protein.

Key words: Citrus sudden death (CSD), Citrus sudden death-associated virus (CSDaV), *Marafivirus*, *Pichia pastoris*.

1. Introdução

1.1 A Morte Súbita dos Citros (MSC)

A morte súbita dos citros (MSC) é uma nova e extremamente destrutiva doença, cuja etiologia, bem como seus possíveis vetores e controle, ainda não foram esclarecidos. O nome atribuído a esta doença foi devido à rápida velocidade com que as plantas afetadas morrem. Todo o quadro sintomatológico da MSC pode ocorrer rapidamente, de um a mais de 12 meses, dependendo da época do ano e da variedade (Informativo Centro de Citricultura, 2001).

A MSC foi observada pela primeira vez em 1999 na cidade de Comendador Gomes, no sul do Estado de Minas Gerais, Brasil (Gimenes-Fernandes & Bassanezi, 2001). As primeiras observações da doença foram feitas em laranja 'Valência' enxertada sobre limoeiro 'Cravo' (*Citrus limonia*). Posteriormente, a mesma doença foi constatada em outras variedades ('Hamlin', 'Natal', 'Westin', 'Pêra', 'Rubi', 'Pineapple' e tangerineiras 'Cravo' e 'Ponkan') enxertadas sobre limoeiros 'Cravo' e 'Volkameriano' (*C. volkameriana*).

A MSC é uma doença de combinação entre laranjeiras-doces (*C. sinensis*) ou tangerineiras (*C. reticulata*) enxertadas em limoeiro 'Cravo' e 'Volkameriano' (Jesus Junior & Bassanezi, 2004a), visto que outras combinações de laranjeiras-doces sobre porta-enxertos de tangerineiras 'Cleópatra' (*Citrus reshni* Hort. ex Tanaka) e 'Sunki' (*Citrus sunki* Blanco), *Poncirus trifoliata* (L.) Rafinesque-schmhlitz, citrumeleiro 'Swingle' (*P. trifoliata* X *Citrus paradisi* Macfayden) e *Poncirus trifoliata* até o momento não apresentam sintomas da doença (Figura 1) (Gimenes-Fernandes *et al.*, 2002; Román *et al.*, 2004).

A MSC é uma das mais importantes doenças da citricultura brasileira, uma vez que os limoeiros 'Cravo' e 'Volkameriano' representam mais de 85% dos porta-enxertos utilizados em todo plantio brasileiro, sendo, portanto, uma das maiores ameaças à competitividade do agronegócio citrícola.



Figura 1. Vista aérea de plantios de laranja-doce ‘Valência’ enxertada em limoeiro ‘Cravo’, apresentando sintomas da morte súbita dos citros, e de laranja-doce ‘Valência’ enxertada em tangerineira ‘Cleópatra’, sadias (A). Em B verificam-se duas linhas de plantio de laranja-doce ‘Valência’ enxertada em tangerineira ‘Cleópatra’, sem sintomas da doença no meio de um plantio de laranja-doce ‘Valência’ enxertada em limoeiro ‘Cravo’, totalmente afetado pela doença.

1.1.1 O porta-enxerto limoeiro ‘Cravo’

A maior parte da citricultura brasileira se baseia em um único porta-enxerto, o limoeiro ‘Cravo’. Este fato se deve a algumas características vantajosas deste material, tais como rusticidade, precocidade, produtividade e, principalmente, resistência às condições de seca. Esta última característica é muito importante, já que permite o plantio sem irrigação, reduzindo bastante o custo de produção, o que torna competitiva a nossa citricultura no mercado internacional.

A decisão de se substituir o porta-enxerto limoeiro ‘Cravo’ por algum porta-enxerto tolerante a MSC deve considerar a inviabilidade da citricultura em locais onde não há disponibilidade de água para irrigação, pois os porta-enxertos tolerantes a MSC apresentam baixa produção em sistemas de plantio sem suprimento de água (Jesus Junior *et al.*, 2003).

1.1.2 Sintomas da Morte Súbita dos Citros

As plantas afetadas pela MSC apresentam sintomas de definhamento generalizado. Inicialmente, as folhas de toda a copa da planta afetada ficam sem brilho e com cor verde-pálida (Figura 2). Posteriormente, ocorre desfolha parcial, produção de poucas brotações e ausência de brotações internas, fato que se acentua com a evolução da doença (Figura 3).



Figura 2. Laranjeiras com sintomas da morte súbita dos citros mostrando perda de brilho generalizada e coloração verde-pálida das folhas, contrastando com a cor normal das plantas sadias (A) e folhas coletadas em plantas sadia e doente, mostrando o contraste de coloração entre as mesmas (B).

O sistema radicular apresenta poucas radículas e muitas raízes podres e mortas. Entretanto, o principal sintoma desta doença precede a morte das raízes e o quadro sintomatológico da copa. Ele se caracteriza por uma coloração entre amarelo e alaranjado presente nos tecidos internos da casca do porta-enxerto, na região do floema funcional, abaixo da região de enxertia até as raízes mais grossas, em contraste com a cor creme dos tecidos da casca da copa (Figura 4) (Jesus Junior *et al.*, 2003). Essas regiões de tecido amarelado correspondem a vasos do floema que se encontram obstruídos e degenerados (Gimenes-Fernandes & Bassanezi, 2001).

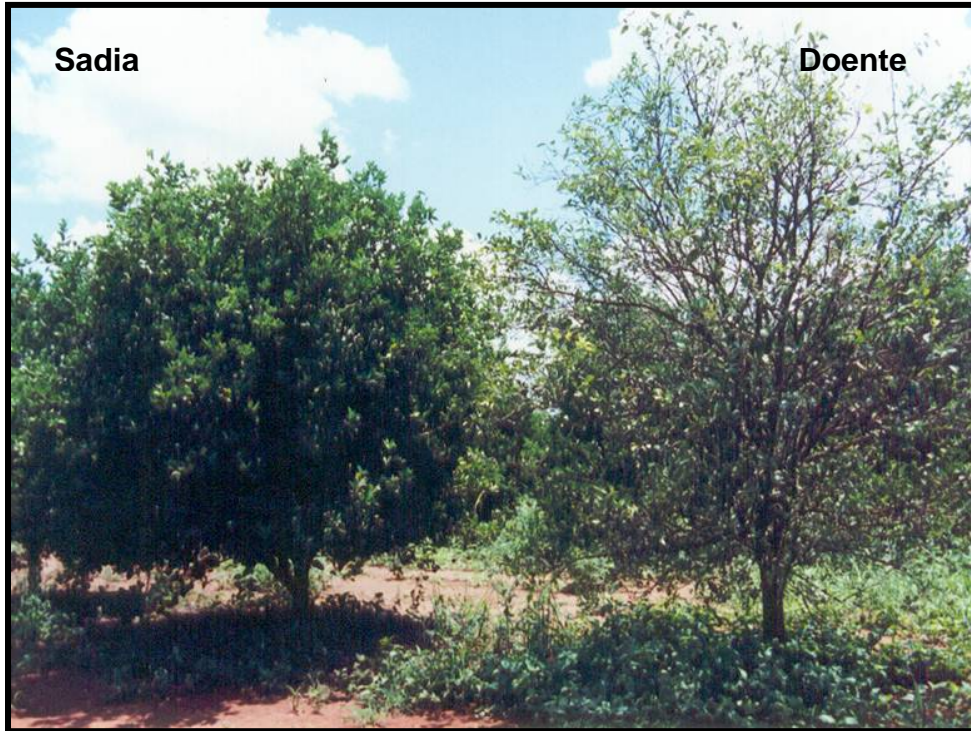


Figura 3. Plantas sadia e com sintomas de desfolha parcial, um dos sintomas da morte súbita dos citros.

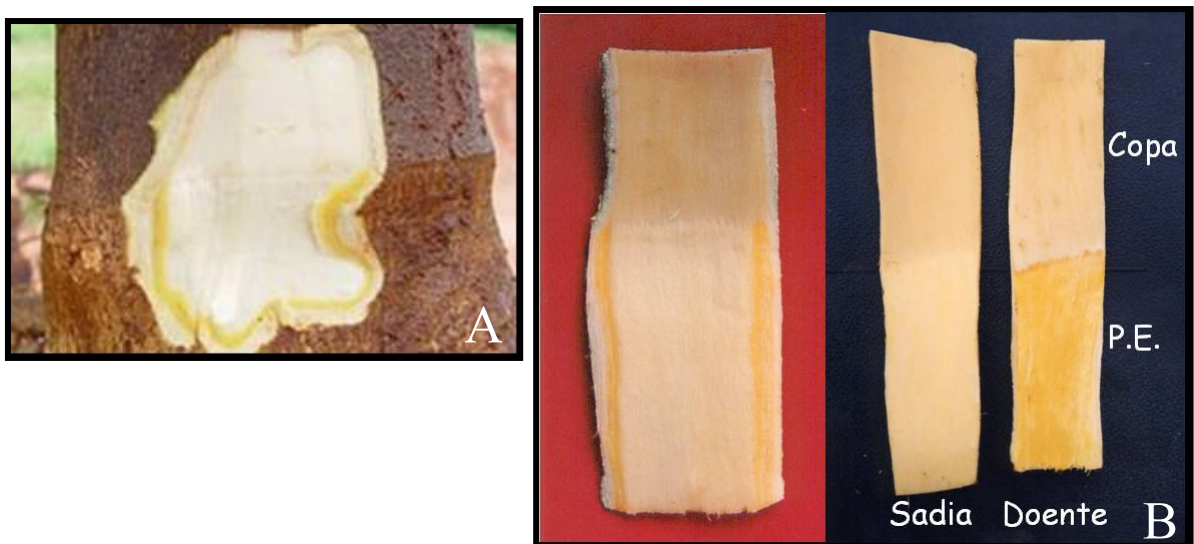


Figura 4. Coloração amarelada nos tecidos internos da casca do porta-enxerto limão ‘Cravo’, na região do floema funcional, contrastando com os tecidos brancos da casca da copa (A) e comparação da coloração supra citada em A entre planta sadia e doente (B), mostrando que esse sintoma é típico de plantas com morte súbita dos citros.

Esses sintomas se intensificam à medida que a doença se desenvolve e culmina com a morte da planta. O aparecimento dos sintomas e a morte da planta ocorrem quando a necessidade de absorção de água para os novos brotos e para a maturação dos frutos se torna alta. As raízes da laranjeira afetada apresentam morte de uma grande porção do sistema radicular, progredindo da ponta da raiz para a base, até que a raiz principal seja afetada e morra (Figura 5).

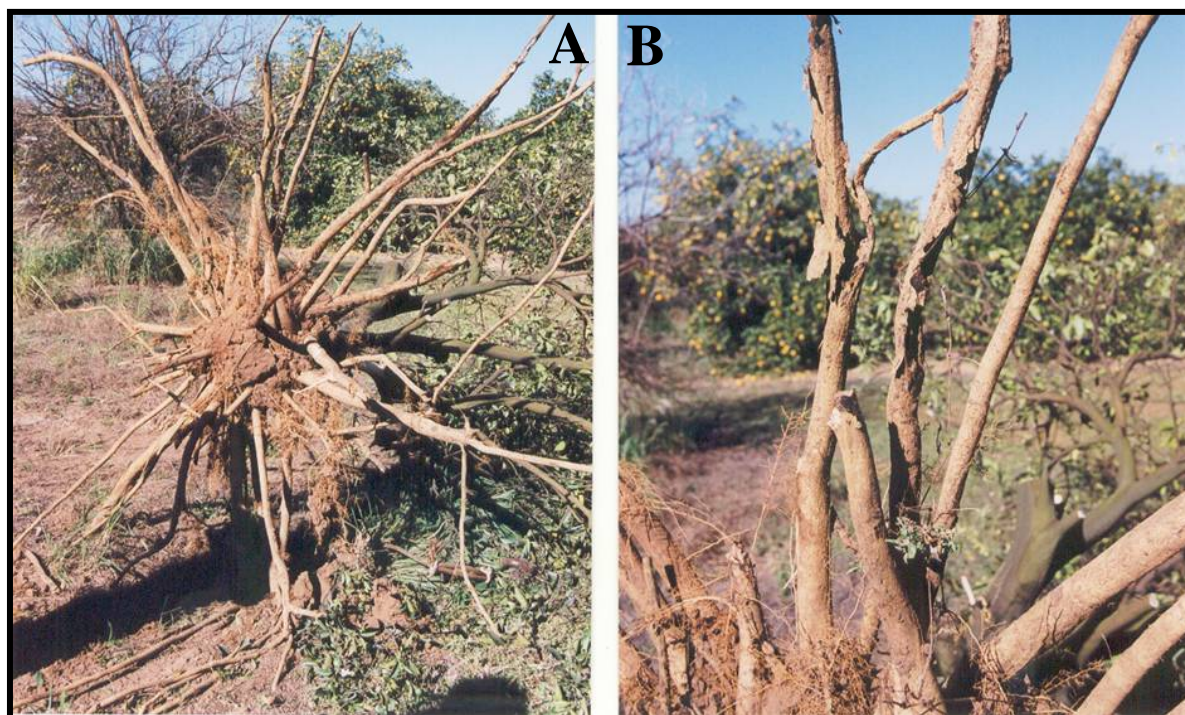


Figura 5. Detalhe do sistema radicular de planta com sintoma inicial de morte súbita dos citros, o qual apresenta grande quantidade de raízes mortas, normalmente com poucas radículas (A) e detalhe da podridão que se inicia da ponta para a base da raiz (B).

A produção das plantas que não entram em colapso repentino vai diminuindo gradativamente até a morte dessas plantas, sendo que quanto mais severos os sintomas da doença, maiores são as diferenças de peso, número e tamanho dos frutos colhidos. Plantas com sintomas iniciais de MSC apresentam uma produção, em média, 27,59%, 11,64% e 21,56% menor que a das plantas saudáveis em termos de peso total de frutos por planta, número total de frutos por planta e tamanho de fruto, respectivamente. Plantas com sintomas severos

de MSC reduzem, em média, 50,41% do peso total de frutos por planta, 25,64% do número total de frutos por planta e 40,67% do tamanho de frutos em relação às plantas saudáveis (Bassanezi *et al.*, 2005). Entretanto, em alguns casos, tem sido observada rápida mortalidade da planta após ter ocorrido uma grande produção, cujos frutos, nesses casos, são normais em número, peso e tamanho e não se desprendem da planta com a morte da laranjeira, devido à rapidez com que ocorre o colapso.

Adicionalmente, tem sido observada uma relação positiva entre produção no ano e número de plantas mortas na safra seguinte, ou seja, quanto maior a produção do pomar com MSC, maior será a mortalidade de plantas na próxima safra.

Todos esses sintomas podem aparecer rapidamente, no prazo de um a 12 meses, dependendo da época do ano (mais rápido na primavera) e da condição de produção da planta (Bassanezi *et al.*, 2002; Müller *et al.*, 2002). Até hoje, nenhuma planta com menos de 22 meses apresentou sintomas da MSC nos pomares de regiões de maior incidência da doença. A análise da frequência de talhões em diferentes idades, nos quais a doença está apenas se iniciando, permitiu deduzir que o período de incubação seja de dois a três anos, dados confirmados pelos testes de transmissão por enxertia (Yamamoto *et al.*, 2004). Este fato leva a pensar que o período de incubação da MSC deve ser de pelo menos dois anos em plantas jovens, e que as plantas que atualmente apresentam sintomas da MSC foram infectadas há algum tempo.

1.1.3. Distribuição da doença

O progresso e a disseminação da MSC em pomares afetados são bastante rápidos, podendo atingir de 30% a 70% das plantas em apenas seis meses (Bassanezi *et al.*, 2002; 2003a; 2003b).

Entre setembro de 2002 e dezembro de 2003, o número de plantas com sintomas de MSC encontradas no campo no Estado de São Paulo passou de 22,1 mil para 44,4 mil e outras 392,2 mil plantas foram erradicadas pela doença no mesmo período (Fundecitrus, 2004).

O número de laranjeiras com os sintomas de MSC aumentou rapidamente na região em que a doença foi detectada inicialmente: de 500 plantas em 1999 (Fundecitrus, 2002) para mais de 2.000.000 em dezembro de 2004 (Fundecitrus, 2004). De dezembro de 2001 a dezembro de 2003, a doença passou de sete municípios (Libanore *et al.*, 2002) para 30 municípios (Fundecitrus, 2004) (Figura 6), sendo 12 no Estado de Minas Gerais (Campo Florido, Comendador Gomes, Conceição das Alagoas, Fronteira, Frutal, Ituiutaba, Monte Alegre de Minas, Planura, Prata, São Francisco de Sales, Uberaba e Uberlândia) e 18 no Estado de São Paulo (Altair, Bálamo, Barretos, Bebedouro, Cajobi, Colômbia, Cosmorana, Embaúba, Guaraci, Ibirá, Ipiriguaná, Monte Azul Paulista, Nova Granada, Olímpia, Onda Verde, Paulo de Faria, Riolândia e Tanabi) (Jesus Junior *et al.*, 2004b).

Tendo em vista que ainda não há forma de diagnóstico precoce para a morte súbita dos citros, o número de plantas infectadas pode estar subestimado (Jesus Junior & Bassanezi, 2004a).

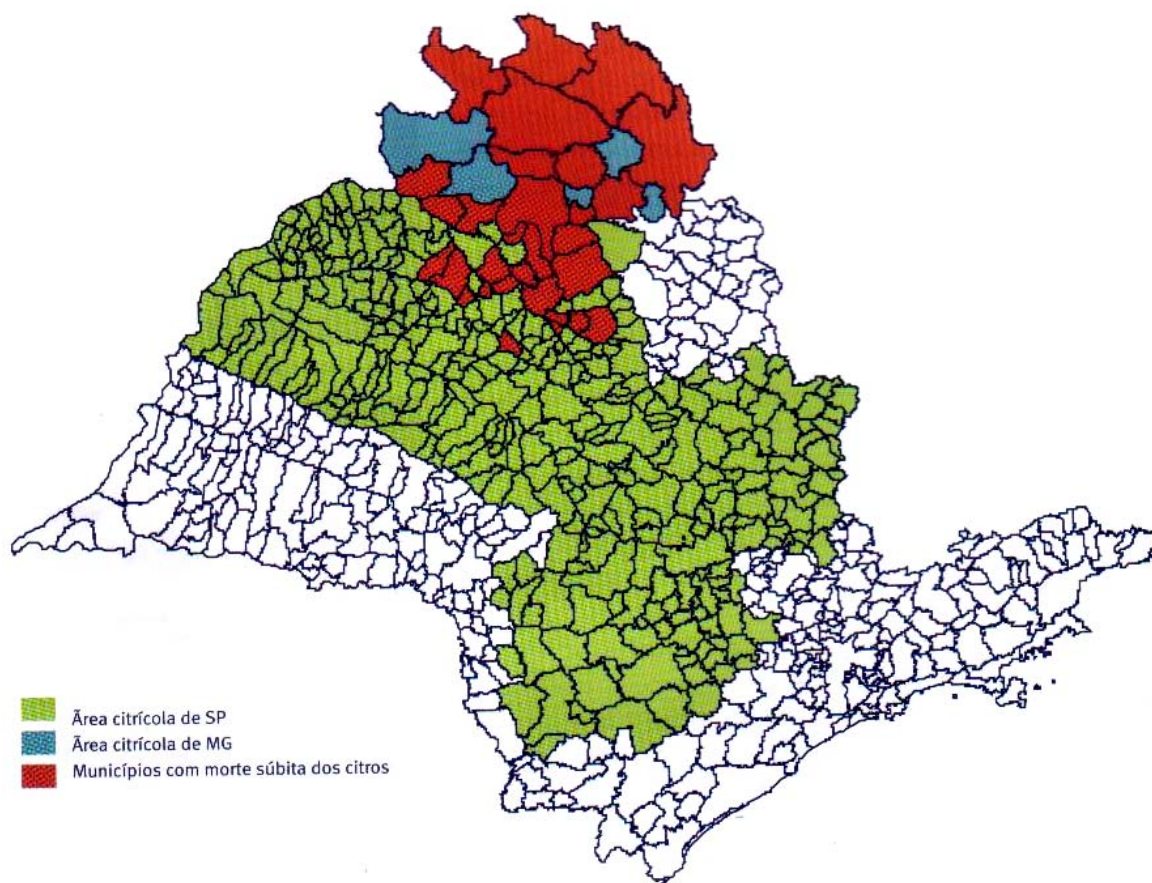


Figura 6. Mapa da distribuição da morte súbita dos citros em SP e MG. Fonte: Jesus Junior *et al.*, 2004b.

1.1.4 Manejo da doença

É um risco efetuar recomendações para o manejo ou prevenção da doença cuja etiologia ainda não está cientificamente comprovada, conforme veremos no item 1.1.5. As estratégias recomendadas até o momento são baseadas no que foi observado em campo, faltando, entretanto, comprovação técnica-científica. Dentre as medidas recomendadas, tem-se: 1) não transportar mudas, borbulhas e cavalinhos das regiões com a doença para aquelas onde a doença ainda não foi constatada; 2) evitar o plantio sobre limoeiros ‘Cravo’ e ‘Volkameriano’; 3) diversificação do plantio com porta-enxertos tolerantes; e, 4) subenxertia com porta-enxertos tidos como tolerantes (citrumeleiro ‘Swingle’, tangerineiras ‘Cleópatra’ e ‘Sunki’ e *Poncirus trifoliata*) em plantas enxertadas sobre limoeiros ‘Cravo’ e ‘Volkameriano’.

A subenxertia consiste no plantio de um novo porta-enxerto, conhecido por cavalinho, ao lado da planta existente e a enxertia deste no tronco acima da região da enxertia, utilizando-se a técnica de T invertido (Figura 7). Em plantas com até três anos de idade tem-se utilizado um único cavalinho, enquanto em árvores mais velhas são utilizados dois cavalinhos. A idade ideal do porta-enxerto a ser usado é de cinco a oito meses, dependendo da variedade, com altura aproximada de 45 cm e com tecido maduro (lenhoso) na altura da enxertia.



Figura 7. Subenxertia realizada em planta com seis anos de idade através da técnica do T invertido, utilizando-se dois cavalinhos.

1.1.5 Etiologia

Com base nos resultados de estudos da evolução e da distribuição da MSC concluiu-se que a doença é infecciosa, o agente causal possui um vetor alado envolvido na sua disseminação (Bassanezi *et al.*, 2003), e a disseminação é bastante rápida e eficiente (Jesus-Junior & Bassanezi, 2004a). Entretanto, a morte súbita dos citros é uma doença de etiologia ainda desconhecida.

Sabe-se que estabelecimento da relação causal entre uma doença e um determinado agente patogênico só pode ser confirmado após o cumprimento de uma série de etapas, denominadas de postulados de Koch, que são:

- a) associação constante patógeno-hospedeiro;
- b) isolamento do patógeno;
- c) inoculação do patógeno e reprodução dos sintomas;
- d) reisolamento do patógeno.

Até o momento, estes postulados não foram concluídos para a MSC, visto que o isolamento dos patógenos candidatos desta doença (*Citrus tristeza virus*/CTV e CSDaV) é dificultado não ser transmitido por inoculação mecânica (CSDaV) ou ter baixa eficiência neste processo (CTV) e por não haver plantas indicadoras. Entretanto, enquanto isso não ocorrer, a sua etiologia permanecerá sujeita a várias hipóteses.

Vários trabalhos estão sendo desenvolvidos com o intuito de determinar o agente causal da MSC.

Na tentativa de entender a epidemiologia da doença, Bassanezi e colaboradores (2002) fizeram uma análise temporal e espacial da morte súbita dos citros para gerar hipóteses sobre a etiologia desta doença. Os resultados obtidos foram compatíveis àqueles de uma doença de causa biótica, por exemplo, um patógeno transmitido por um vetor.

Foi comprovado que a causa primária da MSC é de natureza biótica, pois foi verificado experimentalmente que a doença pode ser transmitida de uma planta doente para uma sadia através da inoculação por enxertia de borbulhas sob condições de ausência de insetos (Yamamoto *et al.*, 2003). Este resultado permite excluir como causa da doença os fatores abióticos, pragas, nematóides e fungos de solo, os quais não são transmitidos por enxertia. Assim, os agentes que podem ser transmitidos desta maneira são viróides, vírus ou bactérias endógenas (restritas ao xilema e floema).

Por ser o quadro sintomatológico da MSC resultante de alterações no sistema vascular das plantas, as buscas por bactérias e fitoplasmas que habitam e colonizam os vasos da planta foram realizadas intensivamente, principalmente em relação ao *Candidatus Liberibacter*

asiaticus, *Candidatus Liberibacter africanus* e *Candidatus Liberibacter americanus*, causadores do huanglongbing (greening), *Spiroplasma citri*, causador do *stubborn* e *Xylella fastidiosa*, causadora da clorose variegada dos citros (CVC). Os testes de diagnóstico por PCR em diferentes laboratórios do Brasil e do exterior não constataram a presença destes organismos em tecidos da casca do tronco e raiz de plantas doentes (I.P. Bedendo-ESALQ/USP; M.A. Machado-APTA Citros; J. Hartung-USDA/Estados Unidos; M. Garnier-INRA/França – *dados não publicados*).

Os viróides de citros são moléculas de RNA sem capsídeo que colonizam o floema das plantas infectadas. Apesar da sintomatologia da MSC ser distinta da observada para doenças causadas por viróides, foram realizados testes de PCR com *primers* específicos para detecção destes organismos em amostras de plantas com sintomas de MSC. Em todas as amostras, a presença de viróides foi negativa (N. Duran-Vila & P. Moreno – IVIA/Espanha – *dados não publicados*).

Excluídos fatores abióticos, pragas, nematóides, fungos, bactérias habitantes do xilema e do floema e viróides como possíveis agentes causais da MSC, a hipótese mais provável é que a MSC seja causada por um vírus.

Atualmente existem duas hipóteses quanto ao possível vírus causador desta doença. Os estudos têm apontado para o envolvimento do vírus da tristeza dos citros (CTV, família *Closteroviridae*, gênero *Closterovirus*) ou um novo vírus da família *Tymoviridae*, recentemente detectado em plantas doentes (Harakava, 2004; Barros *et al.*, 2004; Maccheroni *et al.*, 2005). Não se pode descartar, porém, a hipótese de haver um sinergismo entre estes dois vírus, hipótese ainda sem estudos.

A primeira hipótese é que a MSC seja causada por uma nova estirpe mais virulenta do CTV, dada a semelhança entre a MSC e a tristeza na sua forma de declínio rápido (Bové *et al.*, 2002), que entre 1939 e 1949 destruiu 90% das laranjeiras existentes em São Paulo

(Kimati & Galli, 1980). Essa hipótese se baseia em várias evidências, como a semelhança entre ambas as doenças em relação aos aspectos temporal e espacial na presença do pulgão preto (*Toxoptera citricida*) (Bassanezi *et al.*, 2003; Jesus Junior & Bassanezi, 2004a), associação constante das plantas com sintomas de MSC com o CTV e as semelhanças entre a MSC e a tristeza quanto à sintomatologia (Gimenes-Fernandes & Bassanezi, 2001), às alterações ultraestruturais (Román *et al.*, 2004) e por serem doenças de combinação copa/porta-enxerto (Gimenes-Fernandes & Bassanezi, 2001; Müller *et al.*, 2002), cuja remissão dos sintomas pode ser obtida com a utilização da subenxertia.

Além disso, somente CTV foi encontrado por microscopia eletrônica (E. W. Kitajima, dado não publicado), detectado por testes sorológicos (J. A. M. Rezende, M. A. Machado e M. Cambra, dado não publicado) e através de padrões de RNA dupla fita (P. Moreno, dado não publicado) em todas as amostras examinadas (Bassanezi *et al.*, 2002). Entretanto, até o momento nenhuma associação entre a MSC e o CTV foi comprovada experimentalmente (Derrick *et al.*, 2003; Targon *et al.*, 2004), visto que o CTV está presente tanto nas plantas com sintomas da MSC, quanto em plantas assintomáticas.

A segunda hipótese é que a MSC seja causada por um outro vírus que não o CTV, mas algum novo vírus da família *Tymoviridae*. Esta hipótese está baseada no fato de seqüências de RNA com semelhança a um vírus de RNA da família *Tymoviridae* terem sido detectadas em plantas afetadas pela MSC, mas não em árvores de regiões não infectadas (Reinach, 2003; Harakava, 2004; Barros *et al.*, 2004), por partículas virais desta família terem sido purificadas a partir de plantas com sintomas de MSC e visualizadas por microscopia eletrônica (Jesus Junior *et al.*, 2004c; Jesus Junior *et al.*, 2005) e um novo vírus denominado Citrus sudden death-associated virus (CSDaV, família *Tymoviridae*, gênero *Marafivirus*) ter sido detectado com grande frequência em plantas com sintomas da MSC e estar ausente em plantas sem a doença (Maccheroni *et al.*, 2005).

Adicionalmente, CSDaV foi detectado em afídeos coletados de pantas com sintomas da MSC através de PCR utilizando-se *primers* específicos a seqüência genômica do CSDaV (Jesus Junior, W. C. de, *comunicação pessoal*).

1.1.5.1 A Família *Tymoviridae*

Tymoviridae é uma família viral que foi recentemente aprovada oficialmente pelo *International Committee on Taxonomy of Viruses* (ICTV), em Junho de 2002. Essa família é caracterizada por vírions isométricos, não envelopados, com cerca de 30 nm de diâmetro, contorno arredondado com estruturas proeminentes na superfície, icosaédricos (T=3) com agrupamento de subunidades da capa protéica (CP) em pentâmeros e hexâmeros (Martelli *et al.*, 2002). Seu genoma é composto de RNA monopartido de fita simples, positiva, com cerca de 6 a 7,5 Kb e com conteúdo muito alto de citidina (de 32 a 50%). A extremidade 5' possui *Cap*, enquanto a extremidade 3' pode ter uma estrutura de tRNA ou uma cauda poli-A, dependendo do vírus (<http://www.dpvweb.net>). O genoma desta família contém uma grande ORF (*open reading frame*) que codifica proteínas relacionadas à replicação, possibilitando a clivagem proteolítica pós-traducional desta poliproteína codificada pela ORF1 por uma *papain-like protease* codificada pelo vírus e a expressão da proteína do capsídeo via RNA subgenômico (sgRNA) 3'-terminal. Há ORFs adicionais em alguns gêneros.

Esta família é composta por três gêneros de vírus vegetais:

- *Tymovirus* (Dreher *et. al.*, 2002), do qual deriva o nome da família;
- *Marafivirus* (Edwards M. C., 2000) e
- *Maculavirus* (Martelli *et al.*, 2002).

As características que distinguem cada gênero são:

- **Gênero *Tymovirus*:**

O RNA genômico (6,0 a 6,7Kb em tamanho) contém três ORFs (Figura 8). A ORF1 codifica uma proteína com 206 kDa com motivos de seqüência conservados de metiltransferase (Mtr), helicase (Hel), RNA polimerase (RdRp) e *papain-like protease* (Pro). A extremidade 3' dessa ORF é altamente conservada e contém uma seqüência de 16 nucleotídeos conhecida como “Tymobox”, que funciona como um promotor de RNA subgenômico (Ding *et al.*, 1990). A ORF2 sobrepõe quase inteiramente a ORF1 e codifica uma proteína de 50 a 80 kDa rica em prolina que é dispensável para replicação do vírus, porém necessária para o movimento do vírus célula-a-célula. A ORF3 codifica uma proteína da capa do vírus (20 kDa), a qual é expressa via RNA subgenômico. Os vírus deste gênero são transmitidos mecanicamente, infectam dicotiledôneas, acumulam-se nas células dos tecidos parenquimáticos, e possuem alguns coleópteros como vetores, por exemplo, *Chrysomelidae* e *Curculionidae*.

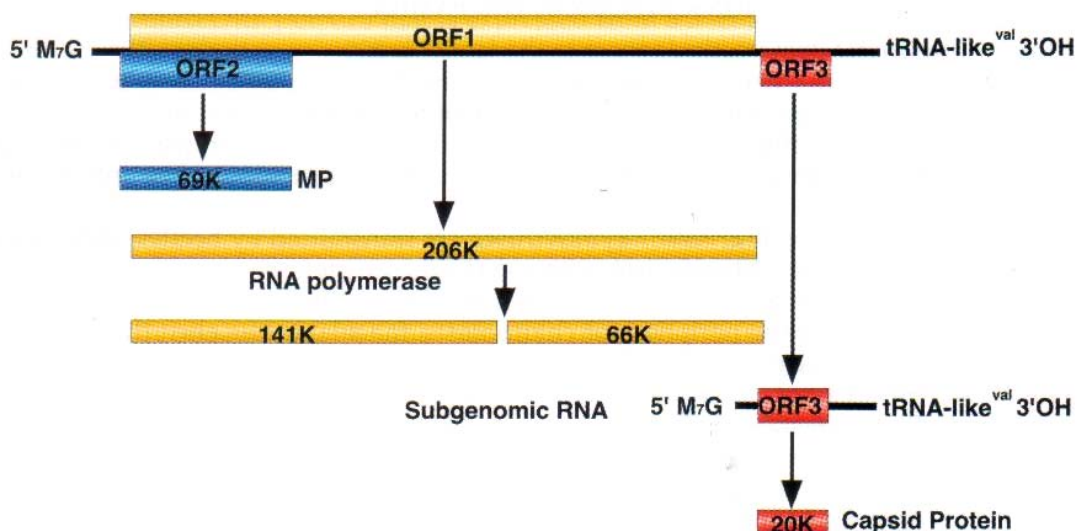


Figura 8. Organização e expressão do genoma do *Turnip yellow mosaic virus* (TYMV), a espécie tipo do gênero *Tymovirus* mostrando a posição relativa das ORFs. MP: proteína de movimento.

Algumas espécies do gênero *Tymovirus* são: *Andean potato latent virus* (APLV), *Belladonna mottle virus* (BeMV) e *Turnip yellow mosaic virus* (TYMV) (Dreher *et al.*, 2005).

- **Gênero *Maculavirus*:**

O RNA genômico do *Grapevine fleck virus* (GFkV) (7,5Kb em tamanho) é o maior desta família, contém *Cap 5'*, é poli-adenilado na extremidade 3' e consiste de quatro ORFs (Figura 9). A ORF1 é a maior delas e codifica um polipeptídeo de 215 kDa que contém motivos conservados de proteínas associadas à replicação (metiltransferase, helicase, RNA polimerase) e um domínio de *papain-like protease*, mas não possui seqüência conservada comparável ao “tymobox” ou “marafibox”. As outras três OFRs estão na extremidade 3' e se sobrepõem. A ORF2 codifica uma proteína do capsídeo de 24 kDa, enquanto as OFRs 3 e 4, codificam proteínas ricas em prolina, de 31 e 16 kDa, respectivamente, com relação distante com a proteína de movimento dos tymovírus. As três ORFs menores sobrepõem-se e são provavelmente traduzidas de dois ou mais RNAs subgenômicos.

Este gênero compreende um pequeno grupo de vírus limitado ao floema e não há nenhum vetor conhecido. Existem duas espécies conhecidas: *Grapevine fleck virus* (GFkV) e *Grapevine red globe virus* (GRGV) (Dreher *et al.*, 2005).

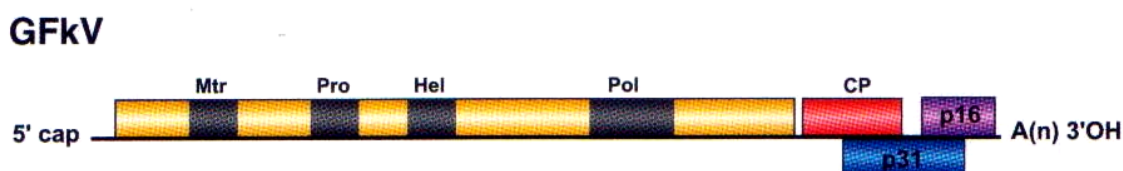


Figura 9. Estrutura genômica de um isolado de *Grapevine fleck virus* (GFkV), a espécie tipo do gênero *Maculavirus*, mostrando a posição relativa das ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CP, capa proteica; p31 e p16, proteínas ricas em prolina.

- **Gênero *Marafivirus*:**

A característica distintiva do genoma do marafivirus (6,3 a 6,5Kb em tamanho) é uma grande ORF que codifica um polipeptídeo (~225 kDa), consistindo de proteínas relacionadas à replicação e a maior (~25 kDa) das duas proteínas do capsídeo (CP) que este gênero possui. Esta ORF possui motivos conservados de proteínas associadas à replicação (Mtr, Hel, RdRp), um domínio de *papain-like protease* (Pro) e o “marafibox” (Izadpanah, 2002), uma sequência de 16 nucleotídeos comparável ao “tymobox”, do qual difere em três resíduos. Os vírions possuem duas subunidades da CP, de aproximadamente 25 kDa e 21 kDa (Edwards *et al.*, 1997), cujas ORFs estão localizadas na extremidade 3’ do genoma viral. A CP de aproximadamente 25 kDa é produzida inicialmente na região C-terminal do polipeptídeo e posteriormente clivada, enquanto a CP de aproximadamente 21 kDa é produzida a partir do sgRNA da região 3’-co-terminal.

Os vírus atualmente classificados como marafivirus exibem algumas diferenças entre si na sua estrutura genômica. O genoma do *Maize rayado fino virus* (MRFV) difere do *Oat blue dwarf virus* (OBDV) por possuir uma segunda ORF em fase sobrepondo a região 5’ da ORF maior, enquanto o OBDV possui cauda poli(A) na extremidade 3’ e o MRFV não a possui. Essa ORF codifica uma proteína de 43 kDa mostrando uma baixa similaridade de sequência com a proteína de movimento (MP) do *Tymovirus* (Figura 10).

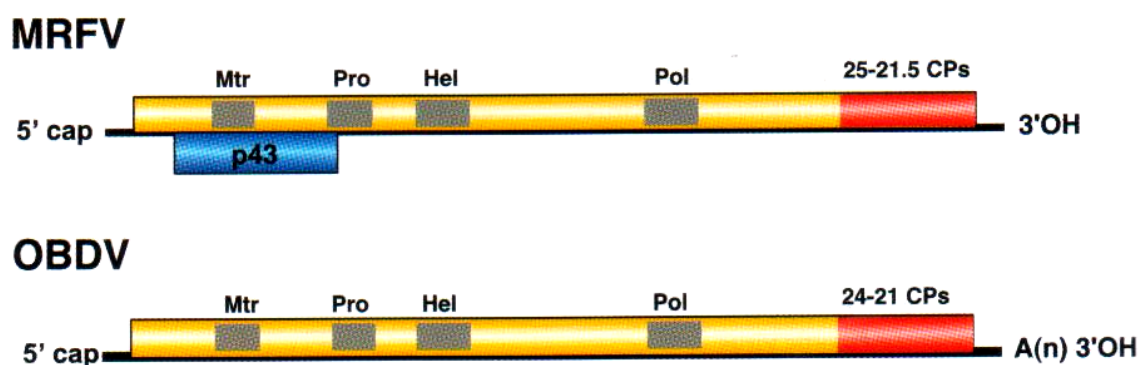


Figura 10. Os dois tipos de estrutura genômica do gênero *Marafivirus*, exemplificado pelo *Maize rayado fino virus* (MRFV), a espécie-tipo deste gênero, e *Oat blue dwarf virus* (OBDV), mostrando a posição relativa das

ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CPs, capas proteicas; p43, proteína rica em prolina.

A maioria das espécies deste gênero são confinadas ao floema do hospedeiro, não são transmitidas através de inoculação mecânica e nenhuma das espécies é transmitida através de sementes. Grapevine asteroid mosaic-associated virus (GAMaV) e Grapevine rupestris vein feathering virus (GRVFV) não possuem vetor conhecido. Outros membros do gênero, entretanto, são transmitidos por cigarrinhas, cada um por um gênero diferente, isto é, MRFV por *Dalbulus*, OBDV por *Macrosteles* e *Bermuda-grass etched-line virus* (BELV) por *Aconurella*. A transmissão é do tipo persistente-propagativa com replicação viral ocorrendo no inseto vetor.

As características que definem uma nova espécie neste gênero são:

- diferenças na estrutura da extremidade 3' e o número de ORFs
- seqüência da proteína do capsídeo com menos de 90% de identidade
- especificidade sorológica
- especificidade do vetor
- efeitos diferentes na ultraestrutura celular
- gama de hospedeiros diferente
- Marafibox diferente em seqüência do tymobox.

Este gênero compreende as espécies: *Bermuda grass etched-line virus* (BELV), *Maize rayado fino virus* (MRFV), *Oat blue dwarf virus* (OBDV), Grapevine asteroid mosaic-associated virus (GAMaV) e Grapevine rupestris vein feathering virus (GRVFV) (Dreher *et al.*, 2005).

1.2 *Pichia pastoris* Como Vetor de Expressão

Pichia pastoris é uma levedura metilotrófica (ou seja, capaz de utilizar o metanol como fonte de carbono e energia) que pode ser geneticamente modificada para expressar proteínas heterólogas (Higgins & Cregg, 1998). Durante os últimos 15 anos, ela tem se tornado um poderoso sistema de produção de uma variedade de proteínas heterólogas. O aumento da preferência por este sistema particular de expressão pode ser atribuído principalmente a quatro fatores: (1) a simplicidade da técnica necessária para a manipulação genética da *P. pastoris* e sua similaridade a *Saccharomyces cerevisiae*, um dos sistemas experimentais mais bem caracterizados na biologia moderna; (2) a habilidade de *P. pastoris* produzir proteínas heterólogas em altos níveis, intra e extracelularmente; (3) a capacidade de realizar várias modificações pós-traducionais características de eucariotos, tais como O e N-glicosilação, formação de pontes dissulfeto e processamento proteolítico; e (4) a disponibilidade deste sistema de expressão como kit comercial (Cereghino & Cregg, 2000).

Além disso, *P. pastoris* por si só não apresenta problemas como produção de endotoxina como ocorre com bactérias, nem problema de contaminação viral encontrados na expressão de proteínas produzidas em culturas de célula animal. Por causa destas características, algumas proteínas que não podem ser expressas eficientemente em bactérias, *S. cerevisiae* ou em baculovírus têm sido produzidas funcionalmente ativas com sucesso em *P. pastoris* (Lin Cereghino *et al.*, 2001).

P. pastoris pode crescer até atingir níveis muito altos de densidade celular (> 130 g peso seco celular por litro) (Jahic *et al.*, 2002; Wegner, 1990). Em culturas de alta densidade celular, o etanol, produto da fermentação da *S. cerevisiae*, rapidamente atinge níveis tóxicos, o qual limita o crescimento da levedura e a produção de proteína. Por sua preferência por crescimento respiratório, *P. pastoris* pode crescer em altos níveis e assim aumentar a

produção de proteínas secretadas, visto que a concentração da proteína no meio é proporcional à concentração de células na cultura (Cereghino & Cregg, 2000).

Leveduras metilotróficas oferecem avanços em relação às desvantagens freqüentemente encontradas em sistemas de expressão de proteína baseados em *S. cerevisiae*, por exemplo, instabilidade mitótica das cepas recombinantes, uma grande quantidade de superglicosilação indesejada e dificuldades em adaptar a produção a escala industrial (Hollenberg & Gellissen, 1997).

P. pastoris foi desenvolvida para ser um hospedeiro excepcional para a produção de proteínas heterólogas, visto que seu promotor da álcool oxidase foi isolado e clonado; essa transformação foi relatada pela primeira vez em 1985 (Cregg *et al.*, 1987; Brierley *et al.*, 1992). Esse promotor tem sido caracterizado e incorporado em uma série de vetores de expressão em *P. pastoris*.

O sistema de expressão em *P. pastoris* utiliza o promotor da álcool oxidase 1, rigorosamente regulado pelo metanol (Couderc & Baratti, 1980), o qual controla a expressão do gene que codifica a expressão da enzima álcool oxidase, a enzima que catalisa o primeiro passo no metabolismo do metanol. Ela não é detectada em células crescidas em meio com glicose, glicerol ou etanol, mas é dramaticamente induzida quando as células são transferidas para meio com metanol (Roggenkamp *et al.*, 1984). A expressão do gene *aox1* é controlada em nível de transcrição (Tschopp *et al.*, 1987a; Cregg *et al.*, 1989). Em células crescidas na presença do metanol, cerca de 5% do RNA poli(A) é derivado do *aox1*; entretanto, em células crescidas em meio com outras fontes de carbono, o RNA mensageiro do *aox1* não é detectado (Cregg *et al.*, 1988).

Genes heterólogos podem ser expressos em *P. pastoris* colocando-os sob controle do promotor AOX1. Isso tem se mostrado simples, fácil de ser realizado em grande escala e de custo compatível para a fermentação industrial (Cregg *et al.*, 1993).

Os vetores de expressão são integrados ao genoma da *P. pastoris* para maximizar a estabilidade das cepas. A maneira mais simples de se promover este evento é digerir o vetor de expressão em um único sítio antes de realizar a transformação. A terminação livre do DNA estimula a ocorrência de recombinação homóloga que resulta em um evento de integração baseado em um *crossing-over* simples (Cereghino & Cregg, 2000).

Várias proteínas heterólogas têm sido produzidas com sucesso utilizando a levedura *P. pastoris*, por exemplo, a enzima invertase (Tschopp *et al.*, 1987b), tetanus toxin fragment C (Clare *et al.*, 1991), antígeno de superfície da Hepatite B (Cregg *et al.*, 1987) e lisozima bovina c2 (Digan, *et al.*, 1989) entre outros.

2. Hipótese

O Citrus sudden death-associated virus (CSDaV), possível agente causal da Morte súbita dos citros (MSC), é um vírus da família *Tymoviridae*, gênero *Marafivirus*, possui genoma de RNA fita simples contendo *Cap* na extremidade 5' e cauda poli(A) na extremidade 3', e a proteína da capa (CP) deste vírus pode ser expressa na levedura *Pichia pastoris*.

3. Objetivos

3.1 Objetivo Geral

O objetivo geral deste trabalho é sequenciar o genoma completo do CSDaV (família *Tymoviridae*, gênero *Marafivirus*) e expressar a proteína da capa do vírus (CP) para possibilitar uma posterior produção de anticorpos a serem utilizados no diagnóstico do vírus.

3.2 Objetivos Específicos

Os objetivos específicos deste estudo são:

1. Sequenciamento do genoma completo do CSDaV;
2. Busca por similaridade a outros vírus em banco de dados e análise filogenética;
3. Expressão heteróloga da proteína da capa do vírus (CP) em *Pichia pastoris*.

4. Material e Métodos

4.1 Clonagem e Seqüenciamento do Genoma do CSDaV

4.1.1 Extração do RNA viral

Partículas virais de CSDaV (Figura 11) foram purificadas parcialmente a partir de plantas sintomáticas de MSC (Jesus Junior *et al.*, 2004c). A partir deste material foi feita a extração do RNA viral, segundo o protocolo do kit comercial Trizol LS Reagent (Invitrogen).

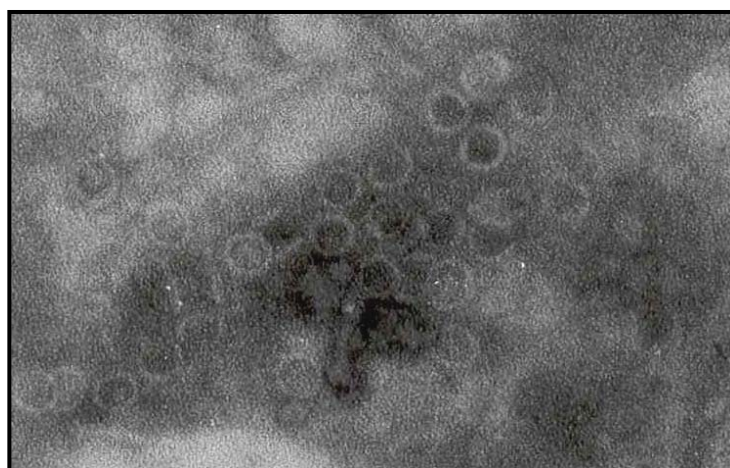


Figura 11. Eletromicrografia de partículas de CSDaV encontradas em plantas com sintomas de MSC.

4.1.2 RT-PCR

4.1.2.1 Síntese de cDNA

A partir do RNA viral foi sintetizado o DNA complementar (cDNA) ao RNA genômico por transcrição reversa (RT) usando *primer* oligo dT M4T 5'-GTTTTCCCAGTCACGAC(T)₁₅ – 3' (Chen *et al.*, 2001) e uma transcriptase reversa, *SuperScript II* (Invitrogen), conforme instruções do fabricante.

Posteriormente, o cDNA foi tratado com 2 unidades de *E. coli* RNase H (Invitrogen) durante uma hora a 37°C, para remover moléculas de RNA ligada ao cDNA. Em seguida a reação foi submetida a 70°C por 15 minutos em banho-maria, com a finalidade de inativar a mesma.

4.1.2.2 Amplificação do cDNA por PCR

Primers degenerados foram desenhados com base nas regiões mais conservadas verificadas através do alinhamento múltiplo entre os genomas de diferentes vírus da família *Tymoviridae*, preferencialmente para os vírus do gênero *Marafivirus* (Figura 12).

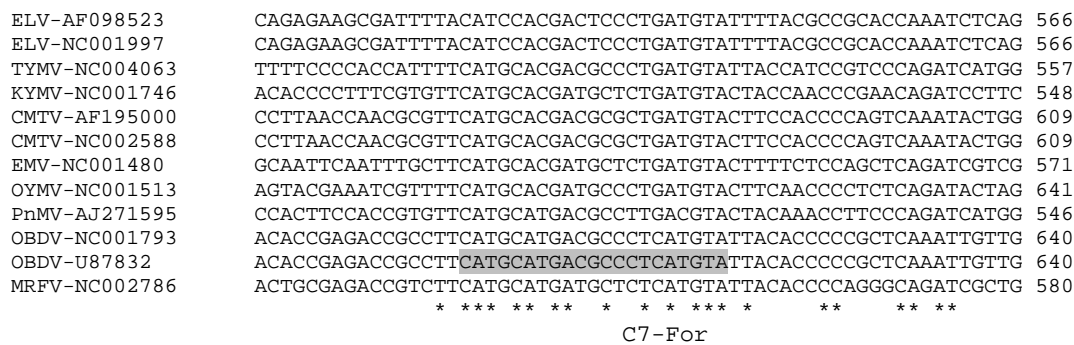


Figura 12. Desenho do *primer* degenerado C7-For baseado nas regiões mais conservadas verificadas através do alinhamento múltiplo entre diferentes vírus da família *Tymoviridae*: ELV (*Erysimun latent virus*); TYMV (*Turnip yellow mosaic virus*); KYMV (*Kennedya yellow mosaic virus*); CMTV (*Chayote mosaic virus*); EMV (*Eggplant mosaic virus*); OYMV (*Ononis yellow mosaic virus*); PnMV (*Poinsettia mosaic virus*); OBDV (*Oat blue dwarf virus*) e MRFV (*Maize rayado fino virus*).

O cDNA foi amplificado através da técnica de PCR, utilizando-se *primers* degenerados e a enzima Platinum *Taq* DNA polymerase High Fidelity (Invitrogen), que é uma enzima que possui *Hot Start* automático e atividade de *proofreading*.

As combinações de *primers* utilizadas para amplificação do cDNA viral através da reação de PCR são apresentadas na Tabela 1.

Tabela 1. Combinações de *primers* utilizadas nas reações de PCR para amplificação do cDNA viral.

<i>Primer Forward</i>	<i>Primer Reverse</i>	Produto PCR
5'-TTYGGCCCTCTCACBTGCATG-3' (C1)	5'-GTTTTCCAGTCACGACT-3' (M4)	1600pb
5'-CATGCATGACGCCCTSATGTA-3' (C7For)	5'-TCAGAGTTGTCGTCATAAGTGCC-3' (C7Rev)	4500pb

Os componentes da reação de PCR foram combinados como descrito na Tabela 2, com volume final de 25 μ L.

Tabela 2. Descrição da quantidade de cada componente utilizado na reação de PCR.

Água	18,0 μ L
Tampão 10X da enzima	2,5 μ L
dNTP (2,5mM)	2,0 μ L
MgSO ₄ (50mM)	1,0 μ L
<i>Primer Forward</i> 10 μ M	0,2 μ L
<i>Primer Reverse</i> 10 μ M	0,2 μ L
DNA <i>template</i> (cDNA)	1,0 μ L
Enzima Platinum <i>Taq</i> DNA polymerase High Fidelity (5U/ μ L)	0,1 μ L
Volume final	25,0 μ L

A reação foi submetida às seguintes condições em um termociclador GeneAmp PCR Systems 9700 (Applied Biosystems): *hot start* automático a 94°C por 1 min., 35 ciclos de desnaturação a 94°C por 30 seg., anelamento dos *primers* a 55°C por 1 min., extensão a 68°C por 2 min. e extensão final a 68°C por 20 min.

4.1.2.3 Confirmação da amplificação em gel de agarose

Gel de agarose a 0,8% foi utilizado para confirmação da amplificação do cDNA, utilizando 1Kb Plus DNA Ladder (Invitrogen) como marcador de peso molecular e brometo de etídeo para corar e visualizar os fragmentos de DNA sob a luz ultra-violeta.

4.1.3 Eluição do fragmento amplificado a partir do gel de agarose

Todo o produto da PCR foi submetido à eletroforese em gel de agarose 0,8% com violeta cristal (4mg/mL), o qual é utilizado no preparo do gel na proporção de 1:1000,

devendo a amostra ser preparada com 1 μ L de corante de DNA violeta cristal 6X para cada 5 μ L de amostra. Este corante permite a visualização direta do fragmento de DNA a olho nu. Desta forma é possível evitar quebras no DNA que poderiam ser causadas pela luz ultravioleta quando se visualiza o gel com brometo de etídeo.

A banda formada pelo produto da PCR foi cortada e eluída do gel de agarose, utilizando o kit de purificação *Perfectprep Gel Clean Up* (Eppendorf), conforme instruções do fabricante.

4.1.4 Ligação do cDNA ao vetor pCR4 TOPO

O vetor de clonagem *pCR4 TOPO TA sequencing vector* (Invitrogen) (Figura 13) permite a seleção direta das colônias recombinantes, pois possui um gene letal para *E. coli*, *ccdB* (Bernard *et al.*, 1994), que não é expresso quando é interrompido pela ligação do produto de PCR. Portanto, as colônias formadas são apenas as de bactérias transformadas com plasmídeo recombinante, pois as células que contêm o vetor sem o inserto expressam o gene *ccdB* e não sobrevivem.

Este plasmídeo linearizado possui genes de resistência a canamicina e ampicilina e extremidades coesivas com uma única deoxytimidina na extremidade 3' para facilitar a ligação e a clonagem direcional. Entretanto, é necessário garantir a presença de uma deoxyadenosina na extremidade 3' do cDNA a ser ligado neste vetor.

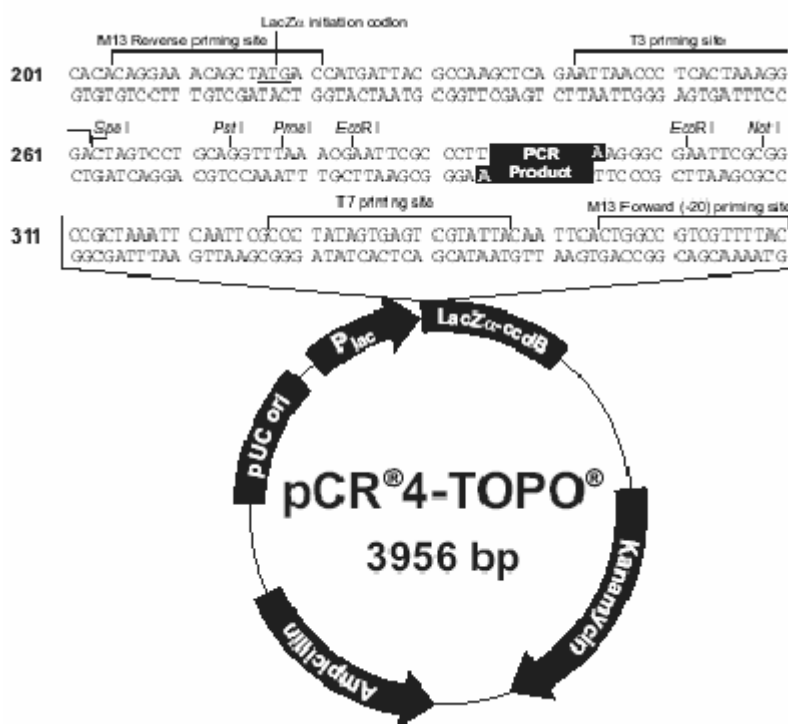


Figura 13. Mapa do vetor de clonagem pCR4 TOPO, o qual apresenta a origem de replicação pUC, promotor lac, sítios múltiplos de clonagem, gene letal ccdB, sítio de ligação do inserto com única deoxytimidina nas extremidades 3' e genes de resistência a ampicilina e canamicina.

4.1.4.1 Adenilação

A enzima Platinum *Taq* DNA polymerase High Fidelity que foi utilizada na amplificação do cDNA, diferentemente das outras *Taq* polimerases, não deixa deoxyadenosina na extremidade 3' do produto de PCR, pois esta é removida pela sua atividade *proofreading*. Por isso, antes de ligar o cDNA ao vetor, a adenilação (*A-overhanging*) foi realizada como descrito na Tabela 3.

Tabela 3. Descrição da reação de adenilação, mostrando seus componentes e suas respectivas concentrações.

<i>Taq</i> polymerase (Invitrogen) 5U/μL	0,20μL
dNTP 50mM	0,12μL
10X PCR Buffer	1,0μL
Produto de PCR	8,68μL
Volume final	10μL

Em um termociclador GeneAmp PCR Systems 9700 (Applied Biosystems) a reação foi submetida a 72°C durante 30 min.

4.1.4.2 Ligaçã

A ligaçã do cDNA adenilado ao vetor pCR4 TOPO foi realizada combinando-se os seguintes reagentes fornecidos pelo kit *TOPO TA cloning Kit for Sequencing* (Invitrogen): 1µL de dilute salt solution, 2µL de água estéril, 1µL de vetor (10ng) e 1µL de produto de PCR adenilado mantendo-se a proporçã molecular de vetor:inseto de 1:3, numa reaçã de volume final de 5µL. A reaçã foi homogeneizada moderadamente antes de se adicionar o vetor, incubada 10 min. no gelo e 30 min. a temperatura ambiente.

4.1.5 Transformaçã de bactérias *E. coli* por eletroporaçã

Para a realizaçã deste procedimento, foram utilizadas células de *E. coli* cepa DH5α (Invitrogen) competentes para eletroporaçã. Estas células foram previamente preparadas seguindo o protocolo de Sambrook & Russel (2001) e estocadas em um freezer a -80°C.

As bactérias foram transformadas por eletroporaçã sob 25µF de capacitância, 200Ω de resistênci e voltagem de 1,80KV em um eletroporador Bio-Rad Gene Pulser. Em um tubo de 1,5 mL foi colocado 1µL de cDNA ligado ao vetor com 40 µL de células *E. coli* DH5α eletro-competentes e transferido para uma curveta de eletroporaçã, Gene Pulser Curvette (Bio-Rad). Acoplada ao eletroporador, a curveta foi submetida ao pulso elétrico. Em seguida 1mL de meio SOC (Sambrook & Russel, 2001) foi adicionado à curveta, homogeneizado delicadamente e transferido para um tubo estéril, o qual foi incubado por uma hora a 37°C sob agitaçã de 240 rpm.

Após o período de incubaçã, as células transformadas foram centrifugadas em microcentrífuga a 22°C e 7000 rpm por 4 minutos. 840 µL do sobrenadante foram removidos

e as células foram ressuspensas no volume restante, o qual foi depositado em placa contendo meio LB-ágar (Sambrook *et al.*, 1989) com canamicina (50µg/mL) e espalhado por toda a placa com o auxílio de uma alça de Drigalski.

4.1.6 Extração de DNA plasmidial

Todas as colônias que cresceram na placa foram selecionadas e com o auxílio de um palito de dente estéril foram inoculadas em 5mL de meio LB (Sambrook *et al.*, 1989) com canamicina (50µg/mL). Os inóculos foram incubados a 37°C sob agitação de 240 rpm durante 16 horas para se fazer a extração de DNA plasmidial, utilizando-se o kit comercial *Wizard Plus SV Minipreps DNA purification System* (Promega), conforme recomendações do fabricante.

4.1.7 Verificação da presença do inserto

Os plasmídios foram digeridos com a enzima *Eco* RI (Amersham Pharmacia), visto que o vetor pCR4 TOPO possui este sítio de restrição nas regiões flanqueadoras dos insertos. A reação foi realizada com volume final de 10µL da seguinte forma: 3µL de plasmídio purificado (miniprep), 1µL de tampão H 10X (Amersham Pharmacia), 0,5µL da enzima *Eco* RI (12U/µL) e 5,5 µL de água destilada. A reação foi incubada a 37°C durante duas horas.

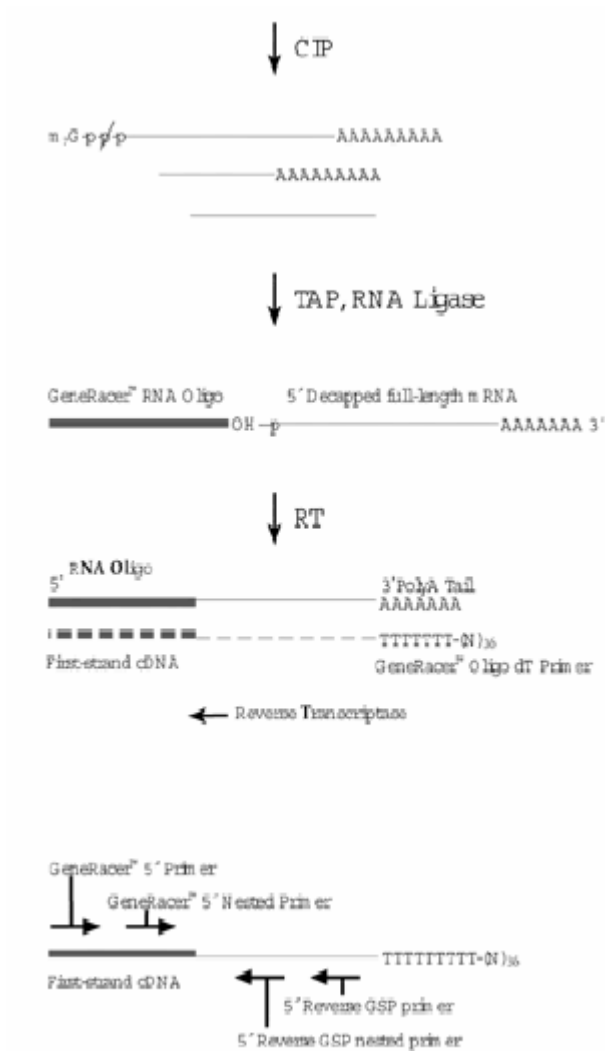
A liberação do inserto foi analisada em gel de agarose 0,8%, corado com brometo de etídeo, utilizando o marcador molecular *1Kb plus DNA ladder* (Invitrogen) como padrão de comparação.

4.1.8 Amplificação da extremidade 5' do genoma viral

Para se obter a seqüência completa do vírus, a extremidade 5' do genoma viral foi recuperada através da técnica 5' RACE (GeneRacer - Invitrogen), conforme as instruções do fabricante.

Essa técnica é baseada em uma enzima ligase (T4 RNA ligase/ Invitrogen) mediada por RNA e no método de rápida amplificação das extremidades do cDNA pela adição de oligonucleotídeo, e resulta na ligação seletiva de um oligonucleotídeo de RNA à extremidade 5' do mRNA cujo *Cap 5'* foi removido, usando *tobacco acid pyrophosphatase* (TAP) (Maruyama & Sugano, 1994; Schaefer, 1995).

Um *primer* homólogo a este oligonucleotídeo e um *primer* reverso específico são utilizados para amplificar a extremidade 5' do cDNA através de reação de PCR. Desta forma, apenas o mRNA que possui o GeneRacer RNA Oligo (tamanho: 44 nucleotídeos) ligado a extremidade 5' e que foi completamente transcrito em DNA será amplificado (Figura 14).



PASSO 1: O RNA total foi tratado com a enzima *calf intestinal fosfatase* (CIP) que remove o fosfato 5' das moléculas de RNA. Isso previne que mRNA truncado e outros tipos de RNA se liguem ao GeneRacer RNA Oligo subsequentemente.

PASSO 2: O RNA foi tratado com *tobacco acid pyrophosphatase* (TAP) que remove o Cap do mRNA, expondo o fosfato 5' e permitindo a ligação do GeneRacer RNA Oligo através de uma T4 RNA ligase

PASSO 3: O cDNA foi gerado por transcrição reversa usando Super Script II (como descrito acima) e um *primer* reverso específico: C7Rev (posição de anelamento no genoma: 5353 nt; seqüência: 5'-TCAGAGTTGTCGTCATAAGTGCC - 3')

PASSO 4: Foi feita PCR utilizando-se um *primer* interno específico TymoGSP1 e o *primer* GeneRacerTM 5' nested que é homólogo à região do GeneRacer RNA Oligo.

Figura 14. Esquema mostrando as etapas de amplificação da extremidade 5' da genoma do CSDaV de acordo com o kit GeneRacer (Invitrogen).

4.1.8.1 PCR Touchdown

A PCR foi realizada com volume final de 25µL como descrito na Tabela 4.

Tabela 4. Descrição da quantidade de cada componente utilizado na reação de PCR *touchdown*.

Água	18,3µL
Tampão 10X da enzima	2,5µL
dNTP (10mM)	0,5µL
MgSO ₄ (50mM)	1,0µL
<i>Primer Forward</i> 10µM (<i>GeneRacer</i> TM 5' nested primer)	1,5µL
<i>Primer Reverse</i> 10µM (TymoGSP1)	0,5µL
RT <i>template</i> (cDNA)	0,5µL
Enzima Platinum <i>Taq</i> DNA polymerase High Fidelity (5U/µL)	0,25µL
Volume final	25,0µL

As seqüências dos primers utilizados na reação acima são:

- GeneRacerTM 5' Nested Primer: 5' - GGACACTGACATGGACTGAAGGAGTA-3' (Posição: 15 do GeneRacerTM RNA Oligo)
- TymoGSP1: 5' – GAGCGGGGCTGAATGTAGTTGTG – 3' (Posição 991nt do cDNA amplificado na etapa de transcrição reversa com o *primer C7Rev*).

Para amplificar seletivamente o cDNA que está ligado ao GeneRacerTM RNA Oligo foi utilizada uma variação da PCR *touchdown* (Don *et al.*, 1991; Roux, 1995) explorando-se a alta temperatura de anelamento dos *primers* utilizados na reação: GeneRacerTM 5' nested *primer* (78°C) e o *primer* específico TymoGSP 1 (60,2°C). Ao se iniciar a reação com alta temperatura de anelamento, somente o cDNA ligado ao GeneRacerTM RNA Oligo é amplificado e uma amplificação eficiente é permitida reduzindo-se a temperatura de anelamento nos ciclos seguintes. Deste modo, a PCR *touchdown* aumenta a especificidade e reduz a presença de fragmentos não específicos.

A reação de PCR *touchdown* foi submetida ao programa descrito na Tabela 5, em um termociclador GeneAmp PCR Systems 9700 (Applied Biosystems).

Tabela 5. Programa utilizado para a realização da PCR *touchdown*.

Temperatura	Tempo	Ciclos
94°C	2 min.	1
94°C	30 seg.	4
72°C	1 min.	
94°C	30seg.	4
70°C	1 min.	
94°C	30 seg.	25
60°C	30 seg.	
68°C	1 min.	
68°C	20 min.	1

As etapas de confirmação da amplificação em gel de agarose, eluição do fragmento amplificado a partir do gel de agarose, adenilação, ligação do cDNA ao vetor pCR4 TOPO, transformação de bactérias *E. coli* por eletroporação, extração de DNA plasmidial e verificação da presença do inserto por digestão com a enzima de restrição *Eco* RI foram realizadas como descritas nos itens anteriores.

4.1.9 Sequenciamento automático

Com o objetivo de obter as seqüências nucleotídicas dos fragmentos clonados e de todo o genoma viral, foi realizado o sequenciamento automático dos plasmídios obtidos. Para tanto, utilizou-se o seqüenciador automático Perkin Elmer ABI-Modelo 377, o kit de sequenciamento através de di-deoxynucleotídeos marcados com fluoróforos DYEnamic ET Terminator Cycle Sequence (Amersham Biosciences) e os *primers* listados na Tabela 6.

Tabela 6. Nome, seqüência e sentido dos *primers* utilizados para o seqüenciamento do genoma do CSDaV.

Primer	Seqüência	Sentido
T3	5' - ATTAACCCTCACTAAAGGGA - 3'	For
T7	5' - TAATACGACTCACTATAGGG - 3'	Rev
P21	5' - TTTCAAAGCCCGAGTAC - 3'	For
P22	5' - GGGTCATCCTTGAAGTGG - 3'	For
P23	5' - TCCCGACTCTCCCCAGGAC - 3'	For
P26	5' - TGGATGGCGAGATCAATGTAG - 3'	Rev
P27	5' - ACCGGCTCAGGATGGCAAG - 3'	Rev
P20	5' - AACCTCCGCAGTCCGGCTT - 3'	For
P 28	5' - GTGAAGATGACACCAACCT - 3'	Rev
P29	5' - GGTGCCTGGGAGATAAGGAG - 3'	Rev
P30	5' - GAGCGGGGCTGAATGTAGTTGTG - 3'	Rev
P31	5' - GAAGAGGGAGATGGTGGTGAAG - 3'	Rev
P33	5' - CCGAATTCTCCTCTGCTCCCATCCTCACAC - 3'	For
P34	5' - GGGAACCTATTGTGGAACCAGTCA - 3'	Rev
P35	5' - ACCTTCAGGAACTTTAAGGTGGC - 3'	For
P36	5' - ACTCATTGTGGAACCAGTCATG - 3'	Rev
P37	5' - CACCACTACTCCCTCAGGTGCC - 3'	For
P42	5' - GTCCCCTGTGATCGTCTCTC - 3'	For
P43	5' - GACTTCACCACCATCTCCCTCTT - 3'	For
P44	5' - GGCTGTGGATGAAACAT - 3'	Rev
P45	5' - GAGTCCTCAGGAGCCATGC - 3'	For
P46	5' - CTCCGGCGAGCTCAATG - 3'	For
P47	5' - GCATGGATGTAAAGGTG - 3'	Rev
P48	5' - TGAGAAGAGGCTGGAGGATGA - 3'	Rev
P49	5' - GAAAAACGGTCTTTGGATCGAC - 3'	Rev

For = *Forward*; Rev = *Reverse*.

4.1.10 Análise computacional das seqüências de DNA obtidas

As seqüências e os cromatogramas obtidos após o seqüenciamento automático foram analisados com o auxílio do programa Staden Package (Staden *et al.*, 2003) com a retirada de bases com qualidade inferior a Phred 20, da região de poliadenilação e de seqüências do vetor.

Utilizando-se a ferramenta BLAST (Altschul *et al.*, 1990) disponível no endereço eletrônico www.ncbi.nlm.nih.gov/blast/Blast.cgi, as seqüências foram comparadas àquelas disponíveis nos bancos de dados de seqüências nucleotídicas e de aminoácidos na busca por similaridade com seqüências já descritas.

O alinhamento múltiplo entre as seqüências obtidas e seqüências depositadas nos bancos de dados foi realizado através do programa Clustal W (Thompson *et al.*, 1994; www.ebi.ac.uk) e o estudo filogenético foi feito por Neighbor-joining utilizando-se o programa Mega3 (www.megasoftware.net).

As seqüências nucleotídicas foram traduzidas em seqüências de aminoácidos utilizando-se as ferramentas da bioinformática oferecidas pelo ORF finder – NCBI (www.ncbi.nih.gov/gorf/gorf.html) e Transeq – EBI (www.ebi.ac.uk/emboss/transeq).

Buscas por domínios protéicos foram feitas com o auxílio de ferramentas como InterProScan (www.ebi.ac.uk/InterProScan), Pfam (www.sanger.ac.uk/Software/Pfam/search.shtml) e Motifs (<http://motif.genome.jp/>), e a predição do peso molecular das proteínas foi realizada com a ferramenta oferecida no Expasy (www.expasy.org/tools/pi_tool.html).

4.2 Expressão da Proteína da Capa do Vírus

Na etapa de expressão da proteína da capa do vírus (CP), a levedura *Pichia pastoris* foi utilizada como sistema de expressão, utilizando o Kit comercial *EasySelect™ Pichia Expression* (Invitrogen).

4.2.1 Amplificação da região genômica codificadora da capa do vírus

Primers específicos que flanqueiam a região do genoma viral codificadora da capa do vírus foram desenhados com a adição de sítios de restrição das enzimas *Eco* RI e *Xba* I: 5' – CCGAATTCTCCTCTGCTCCCATCCTCACAC – 3' e 5' – GGTCTAGAGCGAGAAGATTACCGGAGGGG – 3', respectivamente. Estas enzimas foram selecionadas por estarem presentes nos sítios de clonagem no vetor de expressão a ser utilizado (pPICZαA) e, ao mesmo tempo, por não existirem esses sítios de restrição dentro da região a ser amplificada codificadora da CP.

4.2.1.1 PCR

A PCR foi realizada combinando-se os componentes de acordo com a Tabela 7.

Tabela 7. Descrição da quantidade de cada componente utilizado na PCR para a amplificação da região CP.

Água	18,0µL
Tampão 10X da enzima	2,5µL
dNTP (2,5mM)	2,0µL
MgSO ₄ (50mM)	1,0µL
<i>Primer Forward</i> 10µM	0,2µL
<i>Primer Reverse</i> 10µM	0,2µL
DNA <i>template</i>	1,0µL
Enzima Platinum <i>Taq</i> DNA polymerase High Fidelity (5U/µL)	0,1µL
Volume final	25,0µL

A reação foi submetida às seguintes condições em um termociclador GeneAmp PCR Systems 9700 (Applied Biosystems): *hot start* automático a 94°C por 1 min., 35 ciclos de desnaturação a 94°C por 30 seg., anelamento dos *primers* a 55°C por 1 min., extensão a 68°C por 2 min. e extensão final a 68°C por 20 min.

4.2.1.2 Confirmação da amplificação em gel de agarose

Gel de agarose a 0,8% foi utilizado para confirmação da amplificação do cDNA, utilizando 1Kb Plus DNA Ladder (Invitrogen) como marcador de peso molecular e brometo de etídeo para corar e visualizar as bandas sob a emissão da luz ultra-violeta.

4.2.2 Ligação da região CP ao vetor de expressão

O vetor de expressão em *P. pastoris* pPICZ α A possui múltiplos sítios de clonagem com 10 sítios de restrição únicos que permitem uma eficiente inserção do gene no vetor de expressão. Este vetor possui também um sinal de secreção α -factor de *Saccharomyces cerevisiae* para possibilitar a secreção da proteína, uma cauda de polihistidina C-terminal que permite a detecção da proteína de fusão através de anticorpos anti-His tag e a purificação através de resina de níquel, além de um gene de resistência a zeocina. As características deste vetor de expressão estão sumarizadas na Figura 15.

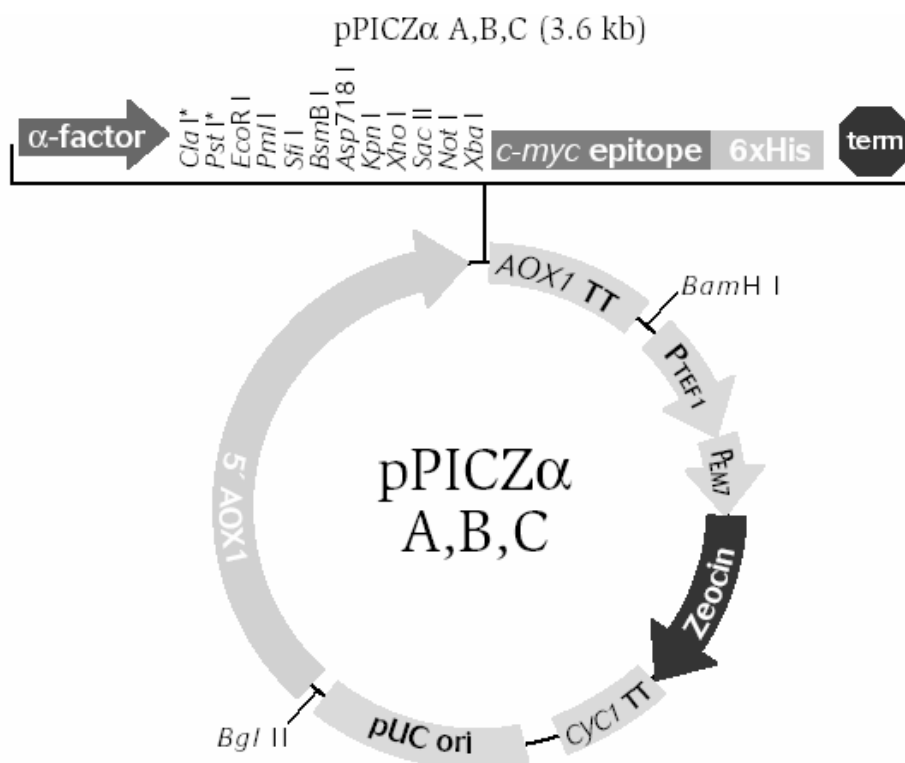


Figura 15. Mapa do vetor de expressão pPICZ α A, mostrando a origem de replicação do plasmídeo em *E. coli*, pUC; fragmento contendo o promotor AOX1; sinal de secreção α -factor; região de múltiplos sítios de clonagem; myc epitopo tag para detecção da proteína através de anticorpos; cauda de 6x Histidina para detecção e purificação da proteína; terminador de transcrição AOX1 TT; promotor do *transcription elongation factor 1* (TEF1) que dirige a expressão do gene de resistência a zeocina, e terminador de transcrição CYC1 de *S. cerevisiae*.

4.2.2.1 Digestão com *Eco* RI e *Xba* I

Tanto o fragmento amplificado quanto o vetor pPICZ α A foram digeridos com as enzimas *Eco* RI e *Xba* I para possibilitar a ligação do inserto ao vetor de clonagem.

A digestão dupla do produto de PCR foi submetida a 37°C durante 16 horas de acordo com a combinação de reagentes descrita na Tabela 8.

Tabela 8. Descrição da quantidade de cada componente utilizado na reação de digestão dupla do produto da amplificação da região CP com as enzimas de restrição *Eco* RI e *Xba* I.

DNA (produto de PCR) 100ng/ μ L	40 μ L
Tampão REact 2 (Invitrogen)	10 μ L
Enzima <i>Eco</i> RI (12U/ μ L)	2 μ L
Enzima <i>Xba</i> I (15U/ μ L)	2 μ L
Água destilada	46 μ L
Volume Final	100μL

A digestão dupla do vetor pPICZ α A foi submetida a 37°C durante 16 horas de acordo com a combinação de reagentes descrita na Tabela 9.

Tabela 9. Descrição da quantidade de cada componente utilizado na reação de digestão dupla do vetor pPICZ α A com as enzimas de restrição *Eco* RI e *Xba* I.

DNA (pPICZ α A)	10 μ L
Tampão REact 2 (Invitrogen)	10 μ L
Enzima <i>Eco</i> RI (12U/ μ L)	2 μ L
Enzima <i>Xba</i> I (15U/ μ L)	2 μ L
Água destilada	76 μ L
Volume Final	100μL

Após 16 horas de incubação, mais 2 μ L de cada enzima foram adicionados à reação, a qual foi submetida a mais 6 horas de incubação.

4.2.2.2 Eluição do vetor e inserto digeridos

O volume total da digestão do produto de PCR e do vetor pPICZ α A foi submetido a eletroforese em gel de agarose 0,8% com violeta cristal (4mg/mL), o qual é utilizado no preparo do gel na proporção de 1:1000, devendo a amostra ser preparada com 1 μ L de corante de DNA violeta cristal 6X para cada 5 μ L de amostra.

A eluição foi realizada utilizando-se o kit de purificação *Perfectprep Gel Clean Up* (Eppendorf), conforme instruções do fabricante.

4.2.2.3 Ligação

A ligação da região do DNA codificadora da CP ao vetor pPICZ α A foi realizada combinando-se os seguintes reagentes: 5 μ L de Tampão 10X, 180ng do vetor pPICZ α A digerido, 50ng da região CP digerida, 3 μ L de T4 DNA ligase (Amersham) (1U/ μ L) e 35 μ L de água destilada, numa reação com volume final de 50 μ L, a qual foi incubada a 4°C durante 16 horas. Paralelamente, foi realizada uma reação de controle negativo, onde a reação de ligação foi repetida, porém sem a adição do inserto.

4.2.3 Transformação em *E. coli*

A transformação foi realizada como descrito no item 4.1.5. Entretanto, 200 μ L do material transformado foi aliquoteado em placa contendo meio LB-ágar *Low Salt* (Tabela 10) com 25 μ g/mL de zeocina e espalhado por toda a placa com o auxílio de uma alça de Drigalski.

Tabela 10. Meio de cultura LB *low salt* utilizado para crescimento de bactérias transformadas com pPICZ α A.

Componentes	g/L
Triptona	10
NaCl	5
Extrato de levedura	5
<i>Ajustar o pH para 7,5</i>	
Ágar	15

4.2.4 Extração de DNA plasmidial

A extração de DNA plasmidial foi realizada utilizando-se o kit comercial *Wizard Plus SV Minipreps DNA purification System* (Promega), conforme recomendações do fabricante. Para isso, todas as colônias que cresceram na placa foram selecionadas e com o auxílio de um palito de dente estéril foram inoculadas em 5mL de meio LB *Low Salt* com 25µg/mL zeocina e incubadas a 37°C sob agitação de 240 rpm durante 16 horas.

O DNA plasmidial extraído foi quantificado em gel de agarose 0,8% com os padrões de quantidade DNA lambda de 20 e 100ng por poço, o marcador 1Kb Plus DNA Ladder (Invitrogen) como marcador de peso molecular e brometo de etídeo para corar e visualizar as bandas sob a emissão da luz ultra-violeta.

4.2.5 Verificação da presença do inserto

Os plasmídios foram digeridos com as enzimas *Eco* RI (Amersham Pharmacia) e *Xba* I (Amersham Pharmacia) para confirmar a presença do inserto nos clones selecionados. A reação foi realizada com volume final de 10µL da seguinte forma: 8,6µL de plasmídio purificado (miniprep), 1µL de tampão REact 2 (Invitrogen), 0,2µL da enzima *Eco* RI (12U/µL) e 0,2µL da enzima *Xba* I (15U/µL). A reação foi incubada a 37°C durante 16 horas.

A liberação do inserto foi analisada em gel de agarose 0,8%, corado com brometo de etídeo, utilizando o marcador molecular 1Kb plus DNA ladder como padrão de comparação.

Para verificar se o gene da CP ligou-se em fase no vetor de expressão, foi realizado o sequenciamento parcial da construção utilizando-se o *primer forward* α -factor (5'-TACTATTGCCAGCATTGCTGC-3') e o *primer reverse* 3' AOX1 (5'-GCAAATGGCATTCTGACATCC -3'), que se anelam em regiões do vetor que flanqueiam a seqüência do inserto.

4.2.6 Linearização do vetor

A probabilidade de integração do DNA plasmidial ao DNA genômico da *P. pastoris* aumenta quando o plasmídeo está linearizado. Para isso, foi realizada a linearização do plasmídeo com a enzima de restrição *Pme* I, pois há apenas um sítio de restrição desta enzima no vetor e nenhum dentro da seqüência da CP.

A reação de digestão com *Pme* I se deu a 37°C durante 16 horas com o volume final de 200µL e a seguinte combinação de reagentes: 20µL de Tampão NEB4 10X (New England Biolabs), 2µL da enzima *Pme* I (10.000U/mL), 70µL de DNA plasmidial e 108µL de água destilada.

Em seguida, toda a digestão foi submetida à eletroforese em gel de agarose 0,8% com violeta cristal, como descrito no item 4.2.3, e eluída utilizando-se o kit de purificação *Perfectprep Gel Clean Up* (Eppendorf), conforme instruções do fabricante, porém substituindo o tampão de eluição do fabricante por água destilada para evitar a presença de sais.

4.2.7 Transformação de *Pichia pastoris*

4.2.7.1 Preparo das células para eletroporação

As células de *P. pastoris* foram preparadas para eletroporação como descrito a seguir.

A cepa GS115 de *P. pastoris* foi inoculada em tubo de 50mL contendo 5mL de YPD (1% extrato de levedura; 2% peptona; 2% dextrose) e incubada a 30°C sob 250 rpm durante 21 horas. 100µL deste inóculo foram transferidos para 50mL de YPD em erlenmeyer de 500mL e deixados sob agitação de 250 rpm a 30°C até atingir OD₆₀₀ entre 1,3 e 1,5. O inóculo foi centrifugado a 1500 x g por 5 minutos a 4°C, o sobrenadante foi descartado e as células foram ressuspensas em 50mL de água gelada estéril. Uma nova centrifugação sob as

mesmas condições foi realizada, porém as células foram ressuspensas em 25mL de água gelada estéril. Este passo foi repetido uma vez, seguido de uma nova centrifugação, descarte do sobrenadante e ressuspensão das células em 20mL de sorbitol 1M estéril gelado. Uma nova centrifugação foi realizada, o sobrenadante foi descartado e as células foram ressuspensas em 1,5mL de sorbitol 1M estéril gelado.

As células foram mantidas no gelo e transformadas em seguida.

4.2.7.2 Transformação por eletroporação

Foram misturados 80µL das células eletro-competentes com 10µL do plasmídeo linearizado (10µg) e 320µL de sorbitol 1M, transferidos para uma curveta de eletroporação e incubados no gelo por 5 minutos. Em seguida, a curveta foi acoplada ao eletroporador *Bio-Rad Gene Pulser* e submetida às seguintes condições de eletroporação: voltagem de 1500 V; 400Ω de resistência e 25µF de capacitância. Imediatamente após o pulso elétrico, 1mL de sorbitol 1M gelado foi adicionado e, em um tubo de 15mL, as células eletroporadas foram incubadas a 30°C sem agitação durante 1 hora.

Após o período de incubação, 80, 100 e 120 µL das células transformadas foram depositadas em placas contendo meio YPDS com Agar (1% extrato de levedo; 2% peptona; 2% dextrose; 1M sorbitol; 2% Agar) e 100µg/mL de zeocina e espalhados por toda a placa com o auxílio de uma alça de Drigalski. As placas foram incubadas a 30°C durante 4 dias.

As colônias formadas foram repicadas em novas placas de YPDS com agar e zeocina e incubadas a 30°C durante 24 horas.

4.2.8 PCR de colônia

A PCR de colônia é uma forma simples de testar diretamente os clones de *P. pastoris* para verificar se o gene de interesse se integrou ao genoma da levedura (Linder *et al.*, 1996).

Para isto, com o auxílio de palito de dente autoclavado, colônias de *P. pastoris* que cresceram na placa de YPDS + zeocina após a transformação foram colocadas em um tubo com 20 μL de água destilada e fervidas em banho-maria durante 10 minutos para promover a lise das células. 2 μL deste material fervido foram utilizados numa PCR com volume final de 10 μL , como descrito na Tabela 11.

Tabela 11. Descrição da PCR de colônia.

10X Buffer	1,0 μL
dNTP 2,5mM	0,4 μL
<i>Primer forward</i> 10 μM	0,25 μL
<i>Primer reverse</i> 10 μM	0,25 μL
DNA (colônias de <i>P. pastoris</i> fervidas)	2,0 μL
<i>Taq</i> polymerase (Pht) (5U/ μL)	0,10 μL
Água destilada	6,0 μL
Volume final	10 μL

Os *primers* utilizados nesta reação foram os mesmos utilizados na PCR realizada para amplificação da região CP (item 4.2.1).

A reação foi submetida às seguintes condições em um termociclador *GeneAmp PCR Systems 9700* (Applied Biosystems): 94°C por 2 min., 35 ciclos de desnaturação a 94°C por 30 seg., anelamento dos *primers* a 52°C por 1 min., extensão a 72°C por 1 min. e extensão final a 72°C por 3 min.

4.2.8.1 Confirmação da amplificação em gel de agarose

Gel de agarose a 0,8% foi utilizado para confirmação da amplificação da região CP a partir dos clones de *P. pastoris* selecionados após a transformação, usando 1Kb Plus DNA

Ladder (Invitrogen) como marcador de peso molecular e brometo de etídeo para corar e visualizar as bandas sob a emissão da luz ultra-violeta.

Os clones selecionados foram submetidos à indução da expressão da proteína heteróloga e analisados quanto à secreção da CP.

4.2.9 Indução da expressão da proteína heteróloga

Os 40 clones crescidos nas placas de YPDS com agar e zeocina após a transformação, além de um controle negativo (cepa GS115 transformada com vetor sem o inserto CP) foram inoculados em 25mL de BMGY (1% extrato de levedura; 2% peptona; 100mM tampão fosfato de potássio, pH 6,0; 1,34% YNB; 4×10^{-5} biotina; 1% glicerol) em erlenmeyer de 250 mL com o auxílio de palito de dente esterilizado. Os inóculos foram incubados a 30°C sob agitação de 250 rpm até a cultura atingir OD₆₀₀ entre 1,3 e 1,5 (cerca de 24 horas) e centrifugados a 3000 x g durante 5 minutos a temperatura ambiente. O sobrenadante foi descartado e as células foram ressuspensas em meio com metanol BMMY (1% extrato de levedura; 2% peptona; 100mM tampão fosfato de potássio pH 6,0; 1,34% YNB; 4×10^{-5} biotina; 0,5% metanol) a OD₆₀₀ inicial de 1 com volume final de 50mL para iniciar a indução da expressão da proteína heteróloga. Neste momento foi recolhida uma alíquota de 5mL do inóculo correspondente ao tempo zero (T0) da indução da expressão da proteína heteróloga. O inóculo foi submetido à temperatura de 30°C e agitação de 250 rpm durante um período de 96 horas. A cada 12 horas foi adicionado metanol 100% a uma concentração final de 0,5% e a cada 24 horas foram recolhidas alíquotas de 5ml do inóculo para determinar em qual intervalo de tempo a proteína é expressa com maior eficiência (*Time course*). Cada alíquota de 5 mL do inóculo de indução foi centrifugada a 3000 x g durante 5 minutos a temperatura ambiente e o sobrenadante congelado a - 80°C.

4.2.10 DIBA

Para determinar se a CP foi expressa e qual o tempo (em horas) necessário para sua expressão em maior quantidade, foi realizado um imuno-ensaio (DIBA; *Dot immunobinding assay*) utilizando-se anticorpo monoclonal anti-poli-histidina (Sigma), o qual possui afinidade pela cauda de seis histidinas na extremidade C-terminal da proteína heteróloga. Para isso, ao fim da indução da expressão da proteína heteróloga em *P. pastoris*, 2 μ L do sobrenadante coletados em cada tempo do *time course* de cada clone, inclusive do controle negativo, foram aplicados diretamente em membrana de nitrocelulose Hybond C (Amersham Biosciences). Após secar completamente, a membrana foi bloqueada com 1X PBS mais 5% de leite em pó desnatado, sob agitação constante durante 3 horas. Em seguida, a membrana foi colocada em uma solução de 1X PBS, 5% de leite em pó desnatado e o anticorpo primário anti-poli-histidina na proporção de 1:3000 *overnight* a 4°C. A membrana foi lavada três vezes com 1X PBS e 0,1% Tween 20 sob agitação a temperatura ambiente por 3 minutos. Durante 2 horas a membrana foi colocada em uma solução de 1X PBS, 5% de leite em pó desnatado e anticorpo secundário conjugado à enzima fosfatase alcalina (Anti-mouse IgG, alkaline phosphatase conjugate/ Sigma) na proporção de 1:5000, sob agitação constante a temperatura ambiente. Em seguida, a membrana foi lavada três vezes com 1X PBS e 0,1% Tween 20 por 3 minutos a temperatura ambiente sob agitação constante. A membrana foi submetida a 10mL de tampão alcalino (0,1M Tris, 0,1M NaCl, 5mM MgCl₂, pH 9,5) e aos substratos da fosfatase alcalina: 40 μ L de NBT (10mg/mL) e 20 μ L de BCIP (10mg/mL) (Harlow & Lane, 1988), sendo que este deve ser adicionado por último. A fosfatase alcalina reage com seus substratos e, através de uma reação colorimétrica, há formação de produto colorido (roxo). A revelação foi realizada no escuro e deve ser acompanhada para verificar o aparecimento das marcas roxas na membrana, indicativas da presença da proteína de interesse.

4.2.11 Indução da expressão da proteína heteróloga em maior escala

Os clones com maiores níveis de expressão no DIBA foram selecionados para serem submetidos à nova indução em volume de 50mL de meio de cultura durante 96 horas. O protocolo seguido foi o mesmo descrito no item 4.2.10, porém sem a realização do *time course*.

5. Resultados e Discussão

5.1 Clonagem e Seqüenciamento do Genoma do CSDaV

A partir do semi-purificado de CSDaV fornecido pelo pesquisador Waldir Cintra de Jesus Junior (Jesus Junior *et al.*, 2004c) foi realizada RT-PCR (item 4.1.2). O *primer* reverso 'C1' se anela na posição 5315 (Figura 16A) e em combinação com o *primer* universal M4, gera um fragmento de aproximadamente 1600pb da extremidade 3' do genoma viral (Figura 16C). Já o *primer* degenerado 'C7 For' e o *primer* 'C7 Rev' (desenhado com base na seqüência obtida a partir do seqüenciamento do fragmento de C1-M4) se anelam aproximadamente na posição 790 e 5350 (Figura 16A), respectivamente, gerando, portanto, fragmentos com cerca de 4560pb do genoma viral através da reação de PCR (Figura 16B).

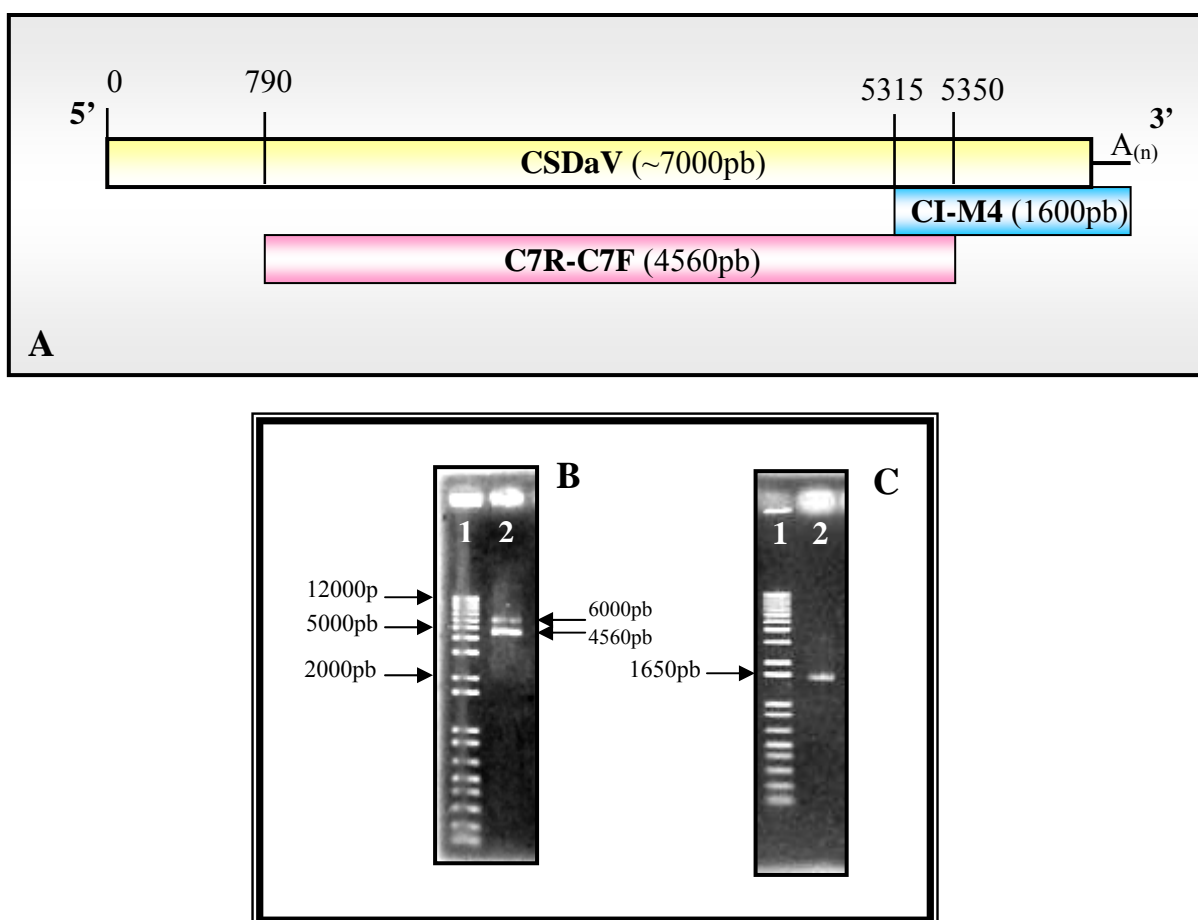


Figura 16. A) Esquema representativo mostrando a posição relativa dos *primers* utilizados na reação de PCR e dos fragmentos obtidos; B) Fotografia de gel de agarose evidenciando o produto amplificado na RT-PCR utilizando a combinação de *primers* C7For e C7Rev. (1) Marcador 1Kb plus (Invitrogen); (2) fragmento de 4560pb amplificado nesta reação; C) Fotografia de gel de agarose evidenciando o produto amplificado na RT-

PCR utilizando a combinação de *primers* C1 e M4. (1) Marcador 1Kb plus (Invitrogen); (2) fragmento de 1600pb amplificado nesta reação.

Entretanto, um fragmento inesperado de aproximadamente 6000pb (Figura 16B) foi amplificado na PCR com os *primers* C7For e C7Rev. Provavelmente, ao final da reação de transcrição reversa (RT), o *primer* M4T ainda estava presente em quantidade suficiente para que em combinação com o *primer* C7For amplificasse um fragmento de cerca de 6000pb. Essa hipótese foi confirmada após a clonagem e sequenciamento deste fragmento.

Após a RT-PCR, os fragmentos amplificados foram eluídos a partir do gel de agarose (item 4.1.3), ligados ao vetor pCR4 TOPO (item 4.1.4) e transformados em bactérias *E. coli* por eletroporação (item 4.1.5). Quinze clones de *E. coli* foram selecionados a partir da transformação utilizando o fragmento de 1600pb e 20 clones transformados com os fragmentos gerados por C7For e C7Rev. Os plasmídios dos clones selecionados foram extraídos (item 4.1.6) e digeridos com a enzima *Eco* RI (item 4.1.7) para verificar quais clones realmente possuíam o cDNA.

Dentre os 15 clones selecionados, 10 não continham o fragmento de tamanho esperado (1600pb) e apenas os clones 8, 9, 14, 15 e 16 tinham o fragmento do tamanho esperado (Figura 17). Na Figura 17 são mostrados os resultados obtidos para cinco dos 15 clones analisados. Nela é possível verificar a presença de duas bandas predominantes para os plasmídios digeridos, onde a banda correspondente ao fragmento maior possui o tamanho correspondente ao do vetor pCR4 TOPO (3956pb) linearizado sem o inserto e a outra banda possui o tamanho correspondente ao fragmento de 1600pb clonado.

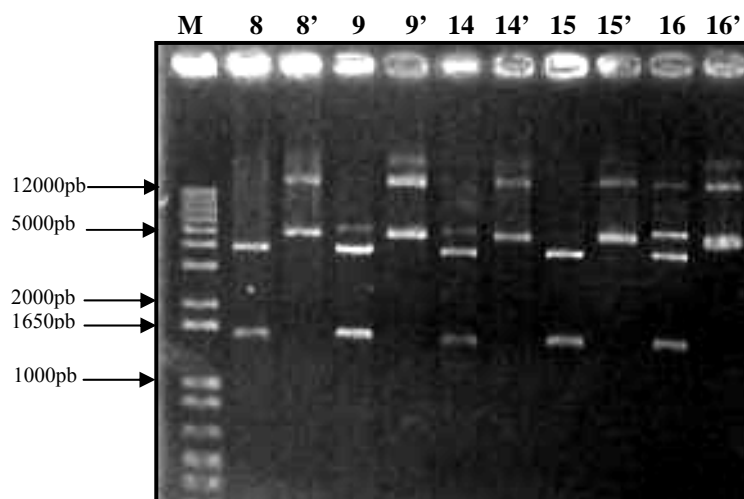


Figura 17. Digestão plasmidial dos clones selecionados para o cDNA de 1600pb. M: Marcador 1Kb plus (Invitrogen); 8, 9, 14, 15 e 16: clones digeridos com a enzima *Eco* RI; 8', 9', 14', 15' e 16': clones não digeridos com a enzima *Eco* RI.

Já dentre os 20 clones transformados com os fragmentos gerados por C7For e C7Rev, dez continham fragmentos muito menores que o esperado e foram descartados; três possuíam o fragmento de aproximadamente 4560pb e sete possuíam fragmento de cerca de 6000pb.

No gel ilustrado na Figura 18, sete clones foram avaliados. Apenas o clone 1 apresentou um fragmento de aproximadamente 4560pb e a digestão plasmidial dos clones 3, 8 e 16 revelou a presença do fragmento maior (~ 6000pb). Tendo em vista que a seqüência do inserto pode possuir sítios de restrição da enzima utilizada, era esperado que os clones positivos apresentassem bandas de diferentes tamanhos, desde que a soma do tamanho dos fragmentos fosse compatível com a esperada no caso da presença do fragmento de 4560pb ou 6000pb no vetor pCR4 TOPO (3956pb). Na Figura 18 é possível verificar a presença de três bandas para o clone 1 digerido. A banda correspondente ao fragmento maior possui o tamanho equivalente ao tamanho do vetor pCR4 TOPO (3956pb) linearizado sem o inserto, e o tamanho dos fragmentos das outras duas bandas correspondem, em soma, ao tamanho do fragmento clonado (~ 4560pb). Já para os clones 3, 8 e 16 digeridos, verifica-se a presença de basicamente quatro bandas. A banda correspondente ao fragmento maior é equivalente ao

tamanho do vetor pCR4 TOPO (3956pb) linearizado sem o inserto, enquanto o tamanho dos fragmentos das outras três bandas correspondem, em soma, ao tamanho do fragmento clonado (~ 6000pb).

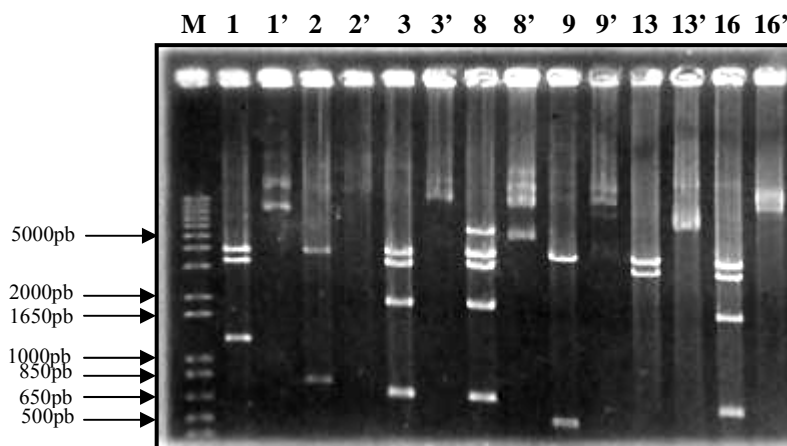


Figura 18. Digestão plasmidial dos clones selecionados para os cDNA de 4650pb e 6000pb. M: Marcador 1Kb plus (Invitrogen); 1, 2, 3, 8, 9, 13 e 16: clones digeridos com a enzima *Eco* RI; 1', 2', 3', 8', 9', 13' e 16': clones não digeridos com a enzima *Eco* RI. Os clones 1, 3, 8 e 16 foram selecionados quanto à presença do fragmento de interesse. O clone 1 revelou inserto de 4560pb, ao passo que os clones 3, 8 e 16 apresentaram inserto de 6000pb.

A partir dos clones obtidos com essas reações de RT-PCR foi possível sequenciar a maior parte do genoma do CSDaV, restando ainda a extremidade 5' (Figura 19).

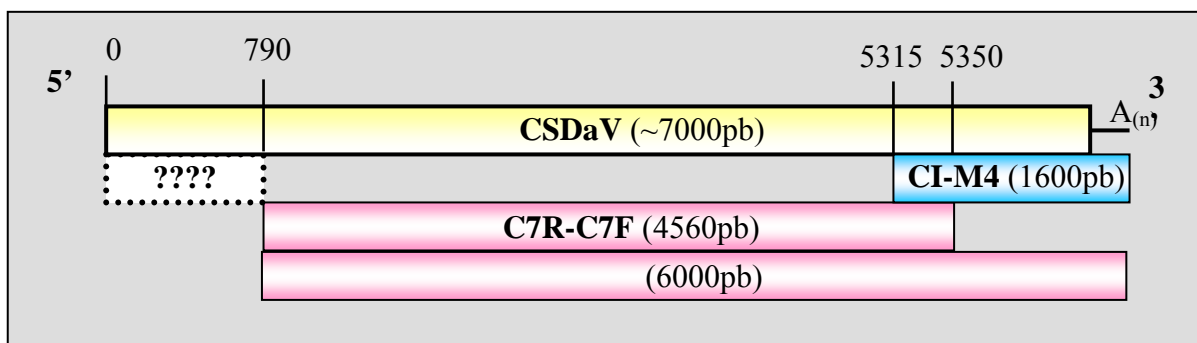


Figura 19. Esquema representando o processo de subclonagens do genoma do CSDaV.

Para amplificar a extremidade 5', portanto, foi realizada a técnica de 5'RACE (item 4.1.8) de acordo com as instruções do fabricante do Kit GeneRacer (Invitrogen). A transcrição reversa foi realizada com o *primer* C7Rev gerando cDNA estimado de cerca de 5350pb. A

partir deste cDNA, a PCR da extremidade 5' foi realizada com o *primer* reverso interno específico TymoGSP1 (posição nucleotídica: 991 do cDNA) e o *primer forward* GeneRacer™ 5' nested que é homólogo à região do GeneRacer RNA Oligo (posição nucleotídica: 15 do GeneRacer™ RNA Oligo), gerando um fragmento de aproximadamente 1020pb na reação de PCR (Figura 20).

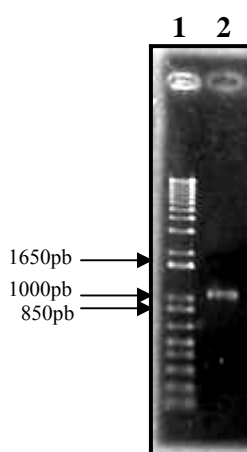


Figura 20. Eletroforese em gel de agarose evidenciando o produto amplificado na RT-PCR 5' RACE. (1) Marcador 1Kb plus (Invitrogen); (2) fragmento de 1020pb amplificado utilizando-se o *primer* reverso TymoGSP 1 e o *primer forward* GeneRacer™ 5' nested.

Após a RT-PCR do 5' RACE, os fragmentos amplificados foram eluídos a partir do gel de agarose (item 4.1.3), ligados ao vetor pCR4 TOPO (item 4.1.4) e transformados em bactérias *E. coli* por eletroporação (item 4.1.5). Vinte colônias de *E. coli* foram selecionadas e tiveram seus plasmídios extraídos (item 4.1.6) e digeridos com a enzima *Eco* RI (item 4.1.7) para verificar quais clones realmente possuíam o fragmento 5' RACE.

Dentre os 21 clones selecionados para o fragmento 5' RACE, oito não apresentaram o fragmento correspondente e 13 realmente o possuíam (Figura 21). Na Figura 21 apenas o clone 10 não apresentou o perfil esperado.

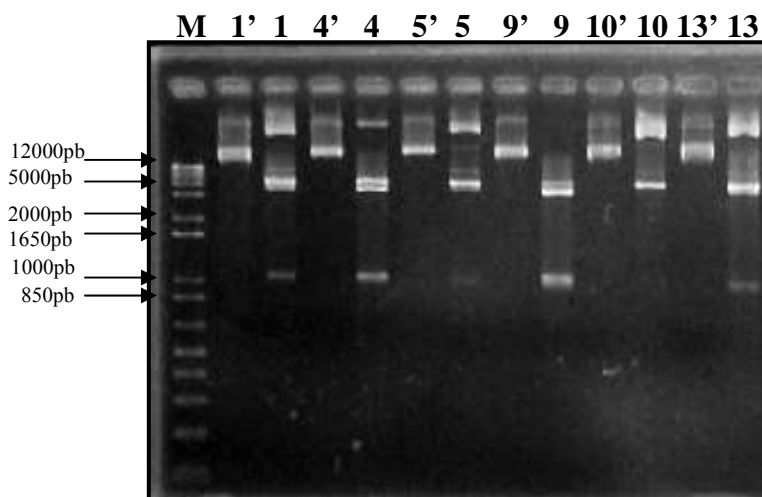


Figura 21. Digestão plasmidial dos clones selecionados para o fragmento de 5' RACE. M: Marcador 1Kb plus (Invitrogen); 1', 4', 5', 9', 10' e 13': clones não digeridos com a enzima *Eco* RI; 1, 4, 5, 9, 10 e 13: clones digeridos com a enzima *Eco* RI. Os clones 1, 4, 5, 9 e 13 foram selecionados quanto à presença do fragmento 5' RACE (~1020pb).

Com base na foto do gel apresentada na Figura 21 é possível verificar a presença de três bandas predominantes para os plasmídios digeridos. Pelo perfil observado é possível verificar que a digestão não foi completa (digestão parcial). Desta forma, a banda correspondente ao fragmento maior possui o tamanho equivalente ao das moléculas do vetor pCR4 TOPO (3956pb) com o inserto (1020pb) não digeridas; a banda do fragmento de tamanho intermediário possui o tamanho equivalente ao do plasmídio linearizado sem o inserto e a banda correspondente ao fragmento menor possui o tamanho equivalente ao do fragmento de 5' RACE (1020pb).

Com a recuperação da extremidade 5' do genoma viral, o genoma completo foi subclonado separadamente (Figura 22), possibilitando o sequenciamento do genoma completo do CSDaV.

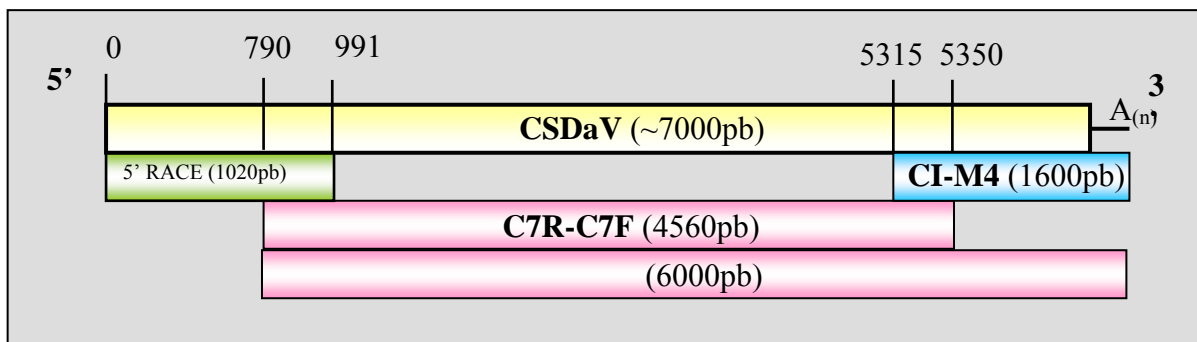


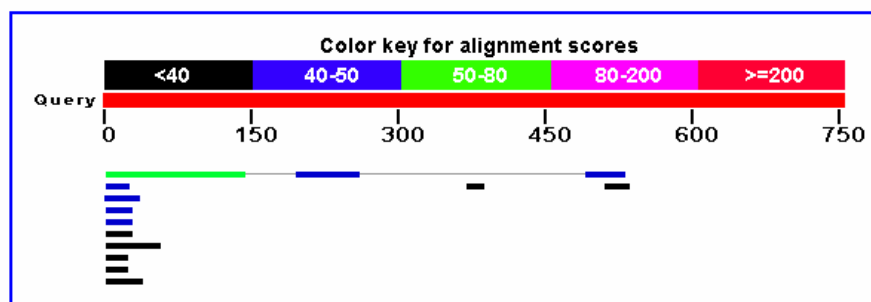
Figura 22. Esquema representando a obtenção do genoma completo do CSDaV através de três subclonagens.

5.1.1 Análise computacional das seqüências obtidas

A medida que as leituras das seqüências foram obtidas, as seqüências consensos foram submetidos a buscas por identidade, utilizando-se a ferramenta BLASTn (<http://www.ncbi.nih.gov/BLAST/Blast.cgi>), para verificar com quais outros vírus este se assemelha quanto a seqüência nucleotídica (Figura 23).

Com base nos resultados obtidos na análise foi verificado que as seqüências obtidas a partir do vírus isolado de plantas com sintomas de morte súbita dos citros tinham maior identidade a outros vírus da família *Tymoviridae* como GAMaV, OBDV, MRFV, GRGV, PnMV e TYMV.

Através da sobreposição de contigs, foi obtida a seqüência do genoma completo do CSDaV (Apêndice A), a qual foi submetida ao banco de dados de seqüências nucleotídicas GenBank, tendo recebido o número de acesso DQ185573.



Sequences producing significant alignments:	Score (Bits)	E Value
gi 10443132 emb AJ249358.1 GAS249358 Grapevine asteroid mosai...	<u>58.0</u>	9e-07
gi 10443189 emb AJ249362.1 GRE249362 Grapevine Red Globe viru...	<u>48.1</u>	9e-04
gi 1843579 gb U87832.1 OBU87832 Oat blue dwarf virus complete ge	<u>48.1</u>	9e-04
gi 2723934 gb AF035403.1 Turnip yellow mosaic Blue Lake isolate	<u>46.1</u>	0.003
gi 332244 gb J04373.1 MTYCLCGA Turnip yellow mosaic virus Club L	<u>46.1</u>	0.003
gi 6822203 emb AJ271595.1 PMO271595 Poinsettia mosaic virus comp	<u>38.2</u>	0.85
gi 14030259 gb AF265566.1 AF265566 Maize rayado fino virus isola	<u>38.2</u>	0.85
gi 62218 emb X16378.1 TYMVCG Turnip yellow mosaic virus complete	<u>38.2</u>	0.85
gi 62222 emb X07441.1 TYMVXX Turnip yellow mosaic virus genome	<u>38.2</u>	0.85
gi 59709783 gb AY751777.1 Scrophularia mottle virus putative...	<u>36.2</u>	3.4
gi 10443136 emb AJ249364.1 GFL249364 Grapevine fleck virus parti	<u>36.2</u>	3.4

Figura 23. Resultado da busca por similaridade nucleotídica através da ferramenta BLASTn, utilizando um consenso de 750pb obtido a partir da leitura de parte da seqüência genômica do vírus isolado de plantas com sintomas de morte súbita dos citros.

5.1.1.1 Análise filogenética

Após obter o seqüenciamento completo do genoma do CSDaV, foi realizado o alinhamento múltiplo entre a seqüência nucleotídica deste vírus e de outros vírus da família *Tymoviridae*, cujos RNA genômicos já foram seqüenciados e disponibilizados no banco de dados GenBank. Para isso foi utilizado o software Clustal W e as seqüências genômicas utilizadas neste alinhamento foram: MRFV (AF265566) e OBDV (U87832) como representantes do gênero dos *Marafivirus*, TYMV (J04373) como representante do gênero *Tymovirus*, GFkV (AJ309022) como representante do gênero *Maculavirus*, PnMV

(AJ271595), ainda não classificado dentro da família *Tymoviridae*, e a sequência obtida do CSDaV (DQ185573).

Com base no resultado do alinhamento múltiplo (Apêndice B) verificou-se uma variação de identidade entre 48% (GFkV) e 63% (OBDV) com o CSDaV. Assim verifica-se que o CSDaV se assemelha mais com os vírus do gênero *Marafivirus*, apresentando maior identidade com o OBDV (63%) e MRFV (60%) (Tabela 12), do que com o PnMV (52%), o tymovírus (49% TYMV) e com o maculavírus (48% GFkV).

Tabela 12. Resultado do alinhamento múltiplo entre a sequência genômica completa obtida do CSDaV e outros vírus da família *Tymoviridae*, realizado com o software ClustalW.

Vírus	Vírus	Identidade (%)
CSDaV	OBDV	63
CSDaV	MRFV	60
CSDaV	PnMV	52
CSDaV	TYMV	49
CSDaV	GFkV	48
OBDV	MRFV	61
OBDV	PnMV	52
OBDV	TYMV	47
OBDV	GFkV	58
MRFV	PnMV	51
MRFV	TYMV	50
MRFV	GFkV	50
PnMV	TYMV	48
PnMV	GFkV	48
TYMV	GFkV	47

A partir do alinhamento múltiplo obtido pelo ClustalW é possível gerar uma árvore filogenética utilizando-se o programa Mega 3, através da ferramenta Neighbor-Joining disponível neste programa (Figura 24).

A árvore filogenética obtida agrupou o CSDaV, o MRFV e o OBDV em um mesmo *cluster* e separou TYMV (*Tymovirus*), PnMV (*Maculavirus*) e GFkV (não classificado) em *clusters* diferentes. Desta forma, baseado na análise filogenética foi possível reafirmar a maior proximidade entre CSDaV e os vírus do gênero *Marafivirus*. Dentre os vírus do gênero

Marafivirus utilizados na construção desta árvore filogenética (MRFV e OBDV), o OBDV foi o que se mostrou mais próximo do CSDaV.

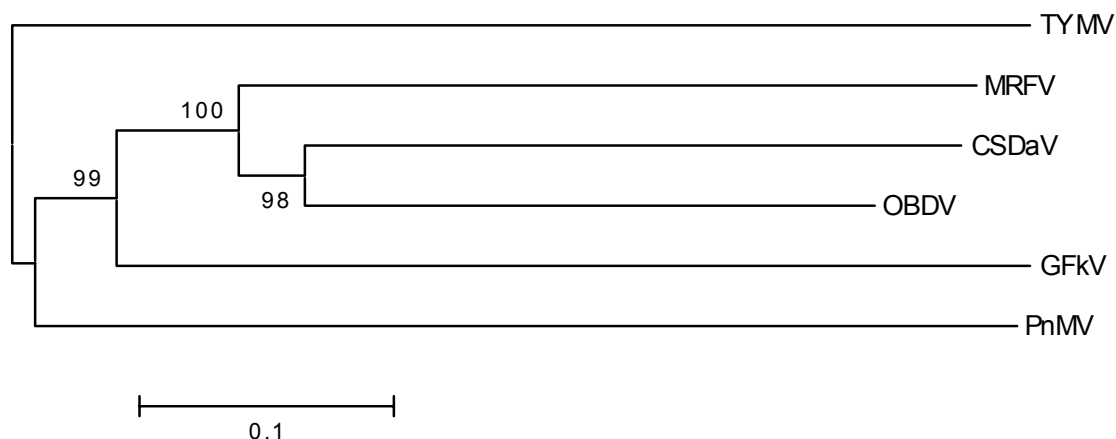


Figura 24. Árvore filogenética gerada pelo programa Mega 3 pelo método Neighbor-Joining a partir da seqüência nucleotídica do genoma completo do CSDaV (DQ185573) e do TYMV (*Turnip yellow mosaic virus*), MRFV (*Maize rayado fino virus*), OBDV (*Oat blue dwarf virus*), GFkV (*Grapevine fleck virus*) e PnMV (*Poinsetia mosaic virus*), disponíveis no GenBank. Os números indicados são os valores de bootstrap (%) de 1000. As distâncias foram calculadas pelo método Kimura-2.

Dentre os vírus do gênero *Marafivirus* apenas o OBDV e MRFV possuem suas seqüências genômicas completas disponíveis em bancos de dados. Entretanto, o vírus GAMaV já foi parcialmente seqüenciado e esta seqüência se encontra disponível no GenBank. Para incluir o GAMaV na busca por identidade e na análise filogenética do CSDaV, foi realizado um alinhamento múltiplo e elaborada uma árvore filogenética baseados na seqüência nucleotídica codificadora da CP do CSDaV e de outros vírus da família *Tymoviridae*. As seqüências utilizadas no alinhamento múltiplo e no desenvolvimento da árvore filogenética foram: GAMaV (CAC10493), MRFV (AF265566) e OBDV (U87832) como representantes do gênero *Marafivirus*; TYMV (J04373) e APLV (AF035402) como representantes do gênero *Tymovirus*; GFkV (AJ309022) como representante do gênero *Maculavirus*, e a seqüência obtida do CSDaV (DQ185573).

A identidade nucleotídica observada no alinhamento múltiplo realizado entre a seqüência CP do CSDaV e dos outros vírus da família *Tymoviridae* foi resumida na Tabela 13.

A seqüência nucleotídica da região CP do CSDaV mostrou uma variação de identidade entre 28% (TYMV) e 71% (GAMaV) e novamente revelou maior semelhança com os vírus do gênero *Marafivirus*, apresentando maior identidade com os vírus GAMaV (71%) e OBDV (70%), seguido do MRFV com 57% de identidade.

Tabela 13. Resultado obtido do alinhamento múltiplo entre a seqüência nucleotídica codificadora da CP do CSDaV e de outros vírus da família *Tymoviridae*, realizado com o software ClustalW.

Vírus	Vírus	Identidade (%)
CSDaV	GAMaV	71
CSDaV	OBDV	70
CSDaV	MRFV	56
CSDaV	APLV	46
CSDaV	GFkV	45
CSDaV	TYMV	28
GAMaV	OBDV	71
GAMaV	MRFV	63
GAMaV	APLV	43
GAMaV	GFkV	58
GAMaV	TYMV	46
OBDV	MRFV	56
OBDV	APLV	27
OBDV	GFkV	47
OBDV	TYMV	15
MRFV	APLV	29
MRFV	GFkV	45
MRFV	TYMV	5
TYMV	APLV	45
TYMV	GFkV	23
APLV	GFkV	43

A árvore filogenética baseada na região codificadora da CP (Figura 25) agrupou o CSDaV juntamente com os marafivirus GAMaV, MRFV e OBDV (Figura 25A), separadamente do APLV e TYMV do gênero *Tymovirus* (Figura 25C) e do maculavírus GFkV (Figura 25B). Com base na árvore filogenética para a região CP verifica-se que, dentre os vírus do gênero *Marafivirus* utilizados nesta análise (GAMaV, MRFV e OBDV), o CSDaV se assemelha mais com GAMaV do que com o OBDV.

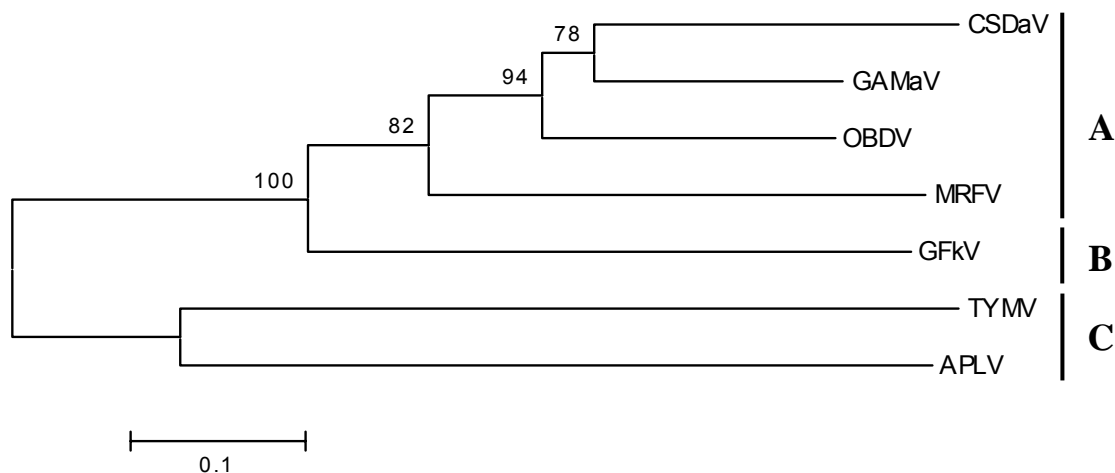


Figura 25. Árvore filogenética gerada pelo programa Mega3 pelo método Neigbor-Joining a partir da seqüência nucleotídica codificadora da CP do CSDaV (DQ185573) e do GAMaV (*Grapevine asteroid mosaic-associated virus*), OBDV (*Oat blue dwarf virus*), MRFV (*Maize rayado fino virus*), GFkV (*Grapevine fleck virus*), TYMV (*Turnip yellow mosaic virus*), e APLV (*Andean potato latent virus*), disponíveis no GenBank. Os números indicados são os valores de bootstrap (%) de 1000. A, B e C mostram o agrupamento dos vírus do gênero *Marafivirus*, *Maculavirus* e *Tymovirus*, respectivamente. As distâncias foram calculadas pelo método Kimura-2.

As análises filogenéticas da seqüência genômica e da região codificadora da CP do CSDaV com as seqüências correspondentes dos vírus representantes dos três gêneros da família *Tymoviridae* sustentam a classificação do CSDaV como um novo membro do gênero *Marafivirus*. De fato, o CSDaV se agrupou em um *cluster* separado dos gêneros *Maculavirus* e *Tymovirus*. Dentre os vírus do gênero *Marafivirus*, o CSDaV formou um ramo separado com o GAMaV e se assemelha mais ao OBDV que ao MRFV.

5.1.1.2 Seqüência e organização genômica

A seqüência completa do genoma do Citrus sudden death-associated virus apresentou 6806 nucleotídeos, excluindo-se a cauda poli(A) presente na extremidade 3'. Uma região 5' UTR com 108 nucleotídeos foi verificada e é compatível ao tamanho encontrado para outros vírus do mesmo gênero (OBDV, 114nt; MRFV, 96nt). A região 3' UTR está presente no genoma do CSDaV a partir do nucleotídeo 6679 e consiste de 128 nucleotídeos,

desconsiderando-se a cauda poli(A). Um trecho de pelo menos 40 adeninas foi observado, entretanto o tamanho real da cauda poli(A) é desconhecido. Essa característica coloca o CSDaV mais próximo dos *Marafivirus* OBDV (Edward *et al.*, 1997) e GAMaV (Ghanem-Sabanadzovic *et al.*, 2003) e do *Maculavirus* GFkV (Sabanadzovic *et al.*, 2001), do que dos *Tymovirus*, os quais possuem em sua extremidade 3' uma estrutura semelhante a de t-RNA (Dreher *et al.*, 2000).

A seqüência de nucleotídeos do CSDaV foi traduzida em seqüência de aminoácidos utilizando-se a ferramenta Transeq – EBI (www.ebi.ac.uk/emboss/transeq) e revelou uma seqüência de 2189 aminoácidos (Apêndice C). Em seguida, a mesma seqüência foi submetida à busca por ORFs através da ferramenta *ORF finder* disponível no NCBI. Assim como outros membros do gênero *Marafivirus*, a análise do genoma do CSDaV identificou uma grande ORF (Apêndice D) que codifica um polipeptídeo com motivos conservados das proteínas associadas à replicação (metiltransferase, helicase, RdRp), um domínio *papain-like protease* e uma capa protéica (CP) de aproximadamente 22,5 kDa codificada pela extremidade 3' desta ORF. Esta grande ORF possui ainda em sua seqüência uma região conservada de 16 nucleotídeos comparável ao “tymobox”, do qual difere apenas em dois ou três resíduos (Figura 26), chamada “marafibox” (Izadpanah *et al.*, 2002). O marafibox tem função de promotor de RNA subgenômico e é seguido de um sítio de transcrição localizado *downstream* à sua seqüência (Martelli *et al.*, 2002).

A grande ORF do CSDaV está compreendida entre os nucleotídeos 109 e 6678 e codifica uma proteína de 2189 aminoácidos com o peso molecular predito de aproximadamente 240,8 kDa (p240) (www.expasy.org/tools/pi_tool.html). Os domínios conservados das proteínas metiltransferase (posição nucleotídica: 487 ao 1119), *papain-like protease* (posição nucleotídica: 2797 ao 3114), helicase (posição nucleotídica: 3358 ao 4053), RdRp (posição nucleotídica: 4222 ao 6030), e de uma proteína de capsídeo (posição

nucleotídica: 6030 ao 6678) foram encontrados nas buscas por similaridade a proteínas através do Pfam (www.sanger.ac.uk/software/Pfam/serach.shtml) (Apêndice E), InterPro (www.ebi.ac.uk/InterProScan) (Apêndice F) e Blast X (<http://www.ncbi.nlm.nih.gov/blast/>) (Apêndice G).

A presença do motivo da enzima metiltransferase (Mtr) próximo à extremidade N-terminal da poliproteína do CSDaV, a presença de uma guanina na primeira posição do RNA genômico e a recuperação da extremidade 5' através da técnica 5' RACE (GeneRacer) sugerem a existência de uma *Cap* na extremidade 5' do genoma deste vírus.

O marafibox está na posição 5959 to 5974 do genoma do CSDaV, e um sítio de iniciação de transcrição CAAG provavelmente se encontra entre os nucleotídeos 5983 e 5986. A conservação das seqüências do promotor entre os vírus do gênero *Marafivirus* é completa, entretanto, a seqüência do marafibox do CSDaV possui uma única diferença em relação ao marafibox de outros marafivírus: na extremidade 5', uma Guanidina é substituída por uma Adenina (Figura 26).

CSDaV	DQ185573	5934	GATGACCCCGACACCACTCTGCTACAAGGTGAATTGCTTCAGCACTTCAAGTACCATTC	5994
GAMaV	AJ249357	983	GAAGATCCCAGCGTCACTCGGTACCAGGGTGAATTGCTTCAGCATCTCCAATCTGTCCCA	1043
OB DV	U87832	5523	GAAGACCCAGCCGTTTCTTCGCTTCAGGGTGAATTGCTTCAGACGTTCCATCTCCAATG	5583
MRFV	AF265566	5385	GAGGACCCTGAAGTTTCGCAGCTTCAGGGTGAATTGCTTCATCAGTTCCAATCGCTTCAC	5445
TYMV	J04373	5576	GAATCCATTGAGGTTTCCACGCTGAGTCTGAATTGCTTCACTATGTCCAATAGCAATCA	5636

Figura 26. Alinhamento da seqüência nucleotídica de 16pb que compõe o “marafibox” (A) e do sítio de iniciação de transcrição (B) do sgRNA do CSDaV (DQ185573) com as regiões correspondentes de outras espécies do gênero *Marafivirus*, GAMaV (AJ249357), OB DV (U87832) e MRFV (AF265566), e o “tymobox” do TYMV (J04373).

Na família *Tymoviridae*, a poliproteína traduzida a partir da grande ORF é processada através de uma atividade autoproteolítica separando o domínio RdRp dos domínios das proteínas metiltransferase, protease e helicase, os quais permanecem em um único polipeptídeo N-terminal (Martelli *et al.*, 2002a). Os membros do gênero *Marafivirus* têm uma proteólise adicional para separar o domínio RdRp do domínio CP, presente na extremidade C-

terminal do polipeptídeo codificado pela grande ORF. A identificação dos sítios de clivagem da poliproteína do CSDaV (Figura 27) foi realizada através de busca por similaridade na seqüência de aminoácidos dos sítios de clivagem do OBDV (Edwards *et al.*, 1997) e MRFV (Hammond & Ramirez, 2001).

			HEL RdRp		A
CSDaV	DQ185573	1365	GPRLTGA	ASPYSKV	1378
OBDV	U87832	1906	HNRLTGA	LAPAQPS	1919
MRFV	AF265566	1188	AVLLRGA	GYGIARP	1201
PnMV	AJ271595	1205	RGMLTGG	NTSGEDD	1218
TYMV	J04373	1253	SPKLNGA	TPSASPT	1266

			RdRp CP22.5		B
CSDaV	DQ185573	1968	LLPLTGG	SSAPILT	1981
GAMaV	AJ249357	356	LLPLKGG	SSAPQLT	369
OBDV	U87832	1828	SLPLFGG	LSPDNIL	1841
MRFV	AF265566	1783	HLPLRGG	HMPNPLA	1796

Figura 27. Sítios de clivagem da *papain-like protease* entre os domínios da Helicase (Hel) e da RdRp (A) e entre o domínio da RdRp e da CP de 22,5 kDa (B), com a posição relativa dos aminoácidos na seqüência da poliproteína. A baixa numeração para o GAMaV é devido ao sequenciamento ainda incompleto deste vírus. O número de acesso da seqüência utilizada para cada vírus está identificada na figura. Os resíduos conservados entre os vírus comparados estão marcados em cinza.

A busca por sítios proteolíticos próximos às posições esperadas na proteína p240 do CSDaV identificou duas regiões. Uma localizada entre os domínios da helicase (Hel) e da RdRp (Figura 27A), com resíduos de Alanina nas posições 1371 e 1372. O alinhamento dessa região do CSDaV com as regiões correspondentes do TYMV (Bransom *et al.*, 1996), do MRFV (Hammond & Ramirez, 2001), OBDV (Edwards *et al.*, 1997) e do PnMV (Bradel *et al.*, 2000) mostrou considerável conservação de seqüência entre os vírus, principalmente no lado esquerdo do sítio de clivagem (Figura 27A). O segundo sítio de clivagem foi localizado entre o domínio RdRp e CP de 22,5 kDa (Figura 27B), entre o resíduo de Glicina na posição 1974 e o resíduo de Serina na posição 1975. O alinhamento dessa região do CSDaV com as regiões preditas do MRFV (Hammond & Ramirez, 2001), OBDV (Edwards *et al.*, 1997) e do

GAMaV (Ghanem-Sabanadzovic *et al.*, 2003) mostrou que esta região é altamente conservada entre os marafivírus, especialmente entre CSDaV e GAMaV, que apresentaram apenas duas diferenças ao longo dos 14 aminoácidos. A Glicina da posição 1974 esteve presente em todos os vírus e a conservação no lado esquerdo do sítio de clivagem na região C-terminal do domínio da RdRp também se mostrou mais pronunciada (Figura 27B).

Além da CP codificada pela grande ORF, os *Marafivirus* têm uma segunda subunidade do capsídeo com cerca de 21 kDa que é codificada pela região 3'-co-terminal do RNA subgenômico, cujo promotor é o marafibox (Dreher *et al.*, 2000; Edwards, 2000; Hammond & Ramirez, 2001). A região codificadora da 21 kDa CP está compreendida entre os nucleotídeos 6084 e 6678 do genoma do CSDaV.

Análises posteriores revelaram uma pequena ORF na extremidade 3' do genoma do CSDaV (posição nucleotídica: 6263 e 6727) codificada pelo quadro de leitura +2 (Apêndice H). Essa possível ORF codificaria uma proteína de 154 aminoácidos com 37% de identidade com parte da proteína rica em prolina de 17 kDa do maculavírus GRGV (Ghanem-Sabanadzovic *et al.*, 2003), 38% de identidade com a proteína hipotética do maculavírus GFkV (Sabanadzovic *et al.*, 2001) e 88% de identidade com a proteína hipotética gp3 do CSDaV sugerida por Maccheroni e colaboradores (2005). Essas proteínas do GFkV e GRGV foram identificadas com base em uma baixa similaridade com proteínas de movimento descritas para o gênero *Tymovirus* (Sabanadzovic *et al.*, 2001). Neste gênero, a ORF codificadora desta proteína de movimento está localizada na extremidade 5' do genoma e codifica uma proteína rica em prolina de 50 a 80 kDa que é dispensável para a replicação do tymovírus, mas é necessária para a movimentação viral célula-a-célula, que seria fundamental para a dispersão sistêmica do vírus na planta. Entretanto, os marafivírus descritos até então são restritos ao floema, com exceção do MRFV, o qual é capaz de invadir as células parenquimáticas da planta hospedeira. Visto que a similaridade observada entre a proteína do

GFkV e GRGV com a proteína de movimento dos tymovírus não é significativa e a tradução desta proteína não foi verificada *in vivo* nem com o CSDaV, nem com os maculavírus, as implicações biológicas da conservação entre essas proteínas são desconhecidas.

Desta forma, o genoma do CSDaV pode ser resumidamente descrito conforme a Figura 28. A primeira e maior ORF (ORF1) é codificadora de uma poliproteína com domínios protéicos associados à replicação (metiltransferase, *papain-like protease*, helicase e polimerase) e uma proteína CP na sua extremidade C-terminal (Figura 28B).

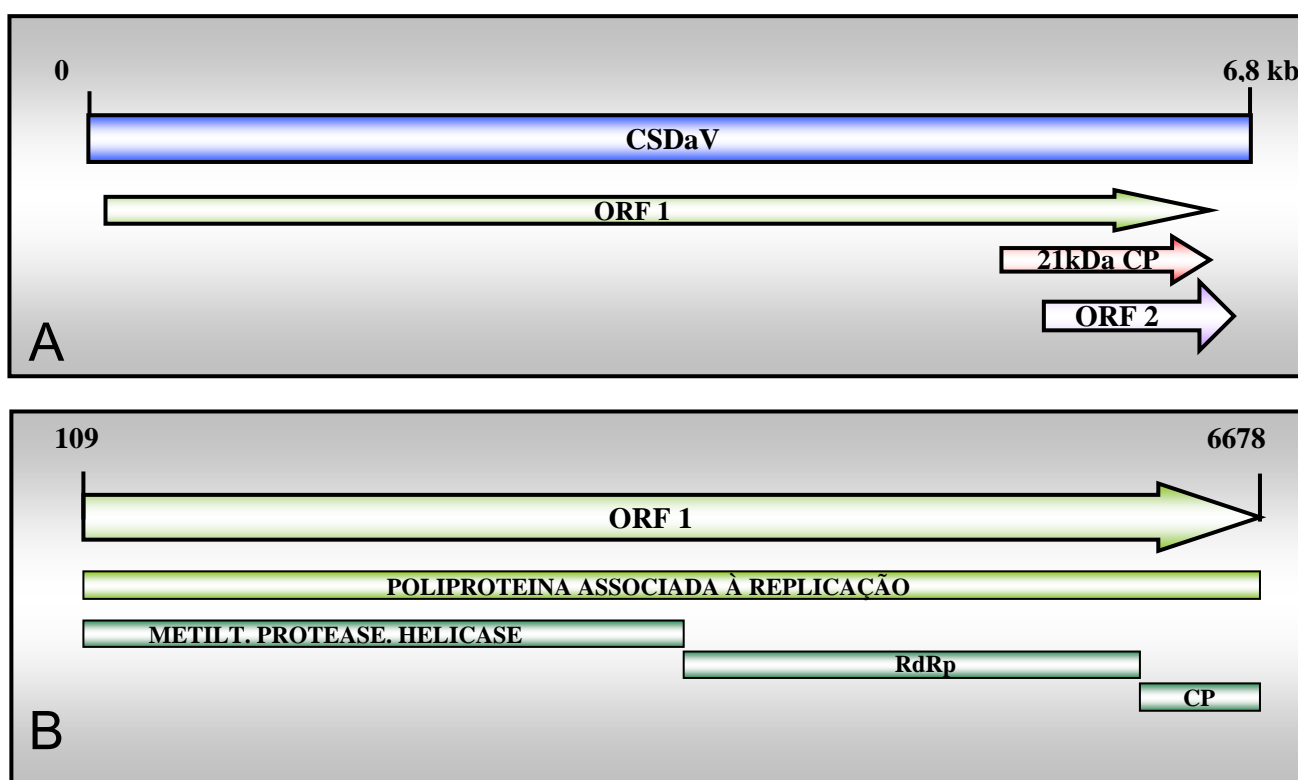


Figura 28. Esquema representativo do genoma do CSDaV. A: desenho esquemático do genoma do CSDaV, mostrando a disposição das três ORFs ao longo do genoma; B: desenho esquemático da ORF1, mostrando a disposição dos domínios de proteína ao longo da mesma.

No mesmo quadro de leitura da ORF1, a extremidade 3' do genoma do CSDaV codifica uma segunda CP de aproximadamente 21 kDa (Figura 28A) que é traduzida a partir de um RNA subgenômico, cujo promotor é uma seqüência altamente conservada entre os marafivírus, denominada “marafibox”. Codificada pelo quadro de leitura +2, possivelmente

há uma nova ORF (ORF2) sobrepondo a extremidade 3' codificadora da poliproteína e da CP de 21 kDa (Figura 28A). Esta ORF codificaria uma proteína de função desconhecida, entretanto com similaridade a proteínas ricas em prolina de vírus da mesma família.

Comparando-se a estrutura genômica do MRFV e OBDV, o CSDaV possui semelhanças com ambos os vírus (Figura 29). Entretanto, não é igual a nenhum deles, pois diferentemente do MRFV, o CSDaV possui cauda poli(A), fato semelhante ao que ocorre com o OBDV; ao passo que diferentemente do OBDV, e de acordo com o MRFV, o CSDaV possui uma ORF adicional, porém na extremidade 5' e não da extremidade 3' como ocorre no MRFV.

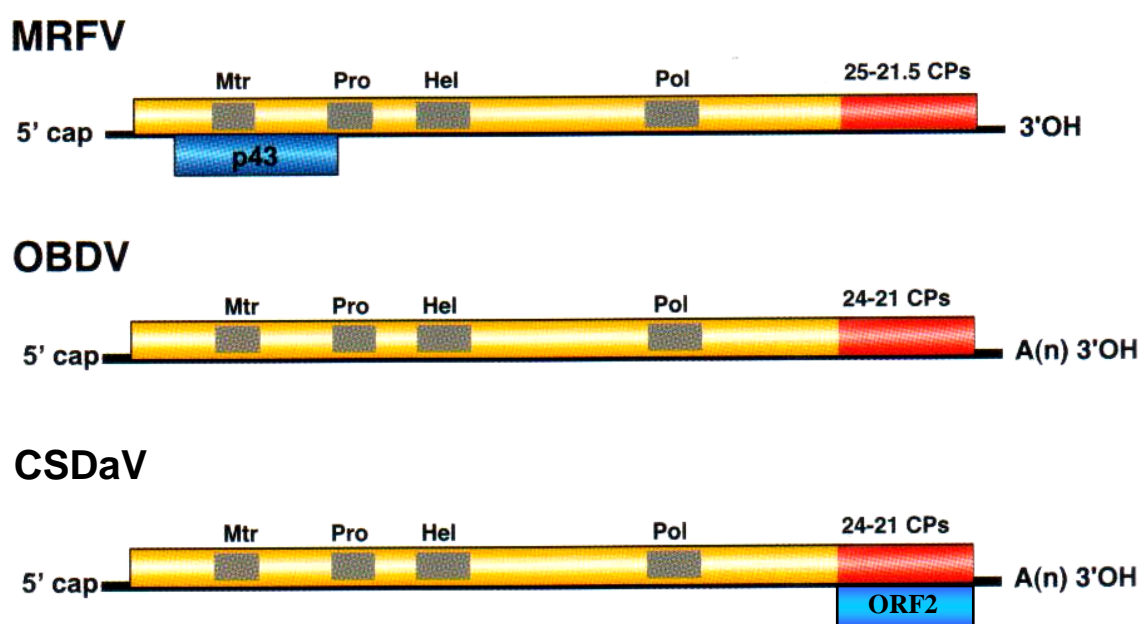


Figura 29. Estruturas genômicas do MRFV, OBDV e CSDaV, mostrando a posição relativa das ORFs e seus produtos de expressão. Mtr, metiltransferase; Pro, *papain-like protease*; Hel, helicase; Pol, polimerase (RdRp); CPs, capa proteica; p43, proteína rica em prolina.

5.1.1.3 Comparação entre os clones de CSDaV sequenciados

Maccheroni *et al.* (2005) também realizaram o sequenciamento de um vírus presente em plantas com MSC (GenBank: AY884005) e obtiveram uma seqüência de 6805 nucleotídeos, excluindo-se a cauda poli(A). Comparando-se a seqüência nucleotídica obtida

neste trabalho (DQ185573) com a seqüência AY884005 através do software ClustalW, a identidade encontrada foi de 89% (Apêndice I). Esse número aumenta para 96% quando essa comparação é feita em nível de aminoácidos (Apêndice J). Entretanto, é possível verificar ao longo da seqüência regiões de menor identidade e regiões com até 100% de identidade (Figura 30).

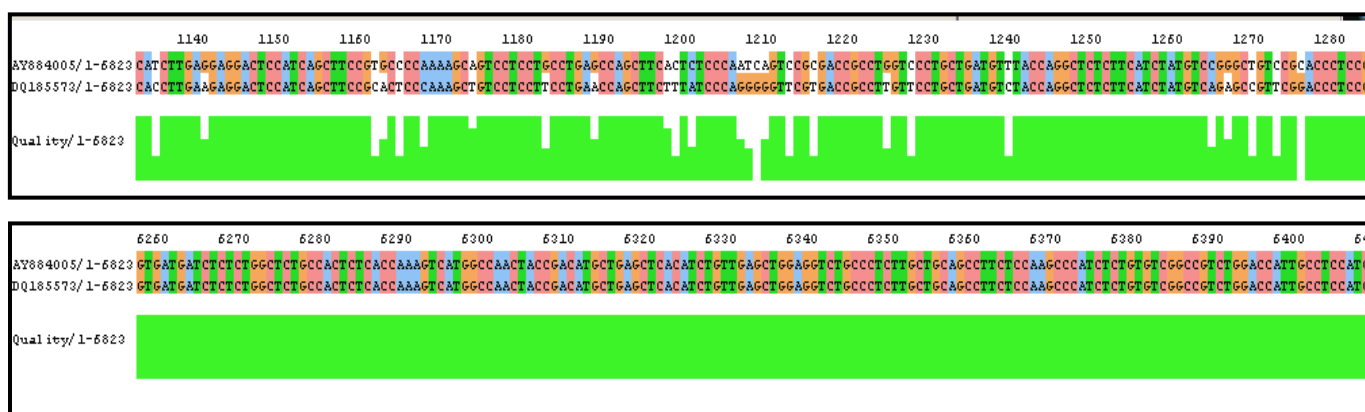


Figura 30. Resultado do alinhamento entre as seqüências nucleotídicas do genoma DQ185573 e AY884005 visualizado através do recurso Jalview do ClustalW, mostrando regiões de maior e menor homologia ao longo do genoma.

As regiões de maior homologia correspondem a região 3' do genoma, codificadoras das duas proteínas CP e parte da RdRp, ao passo que as regiões de menor homologia correspondem ao restante do genoma, codificador dos domínios da poliproteína relacionada a replicação (metiltransferase, protease, helicase e parte da RdRp).

Comparando-se a seqüência de aminoácido de cada domínio protéico entre os clones de CSDaV seqüenciados (DQ185573 e AY884005) obteve-se os resultados mostrados na Tabela 14.

Tabela 14. Comparação entre a sequência de aminoácidos de cada domínio protéico dos CSDaV seqüenciados: DQ185573 e AY884005.

Proteína	Identidade (%)	Similaridade	Gap	Tamanho (aa)
Poliproteína	96	2137/2189 (97,6%)	1/2189 (0,0%)	2189/2188
Proteína hipotética	88,3	137/154 (89%)	0/154 (0,0%)	154/154
CP 21 kDa	100	197/197 (100%)	0/197 (0,0%)	197/197
CP 22.5 kDa	100	215/215 (100%)	0/215 (0,0%)	215/215
RdRp	97	593/603 (98%)	1/603 (0,0%)	603/602
Mt.Pro.Hel	94	-	-	1372/1371

Ao se considerar a comparação da poliproteína inteira obteve-se 96% de identidade e 97,6% de similaridade. Entretanto, após sofrer a clivagem proteolítica, a mesma se divide em domínios Metiltransferase/Protease/Helicase que permanecem juntos no mesmo polipeptídeo, separando-se da RdRp e da CP 22,5 kDa. Quando se efetuaram comparações desses domínios, separadamente, o domínio da CP 22,5 kDa não apresentou nenhuma diferença entre os aminoácidos, mostrando 100% de identidade. Em seguida o domínio da RdRp apresentou 97% de identidade e 98% de similaridade e o Metiltransferase/Protease/Helicase apresentou 94% de identidade. Assim como o domínio da CP 22,5 kDa, a CP 21 kDa apresentou 100% de identidade, enquanto a proteína hipotética, codificada pela ORF2 foi o domínio que apresentou menor identidade entre os CSDaV, com 88,3 % de identidade e 89% de similaridade.

De acordo com esse resultado, pode-se concluir que as duas CPs são as regiões mais conservadas no CSDaV (100% de identidade), seguido da RdRp e Mt.Pro.Hel, e por último a proteína hipotética.

5.1.2 Citrus sudden death-associated virus (CSDaV)

A maioria das características dos *Marafivirus* podem ser encontradas no CSDaV, tais como a presença de uma grande ORF traduzida em uma poliproteína associada a replicação viral e atividade proteolítica, ligada ao domínio CP, uma seqüência promotora de sgRNA conhecida como “marafibox”, transcrição de um sgRNA 3’ coterminal e duas CPs, uma produzida pela clivagem C-terminal da poliproteína e outra traduzida a partir do sgRNA. Entretanto, algumas características encontradas no genoma deste vírus e sua epidemiologia devem ser discutidas:

1. Apesar de se assemelhar mais ao gênero *Marafivirus*, o CSDaV tem um genoma de 6,8 kb, enquanto o genoma de outros marafivírus estão compreendidos entre 6,0 e 6,5 kb, o genoma dos tymovírus estão entre 6,0 e 6,7 kb e dos maculavírus são de 7,5 kb (Dreher *et al.*, 2005).
2. Os vírus atualmente classificados como *Marafivirus* exibem algumas diferenças nas suas estruturas genômicas. Os dois tipos de estrutura genômica no gênero *Marafivirus* são exemplificados pelo MRFV e o OBDV (Dreher *et al.*, 2005). Uma das diferenças encontradas entre o genoma destes vírus é a possível presença de uma ORF no MRFV sobrepondo a extremidade 5’ da ORF codificadora da poliproteína associada a replicação, que codificaria uma proteína de 43 kDa rica em prolina (Hammond & Ramirez, 2001), ausente no OBDV. Neste ponto o CSDaV se assemelha ao MRFV, visto que o genoma do CSDaV também apresenta uma possível ORF codificadora de uma proteína rica em prolina, porém no CSDaV ela está localizada na extremidade 3’ do genoma e não na extremidade 5’ como verificado no MRFV. Outra diferença observada entre o genoma do MRFV e OBDV é a presença da cauda poli(A) na extremidade 3’ do OBDV (Edwards *et al.*, 1997), a qual está ausente no MRFV (Hammond & Ramirez, 2001). Quanto a

esta característica, o genoma do CSDaV se assemelha ao OBDV, pois a presença da cauda poli(A) foi confirmada, verificando-se um trecho de pelo menos 40 adeninas na extremidade 3' do seu genoma.

3. Com exceção do GAMaV e do GRVfV que não possuem vetor conhecido, o vetor natural dos marafivírus são as cigarrinhas, ocorrendo uma associação específica entre vírus e vetor, como por exemplo: MRFV é transmitido por cigarrinhas do gênero *Dalbulus*, OBDV pelo gênero *Macrostelus* e BELV por *Aconurella* (Dreher *et al.*, 2005). Entretanto, apesar de nenhum vírus da família *Tymoviridae* com vetor conhecido ser transmitido por afídeos, estudos epidemiológicos sugerem que o CSDaV tem os afídeos como prováveis vetores. Adicionalmente, o CSDaV já foi detectado em afídeos coletados em plantas com MSC (Maccheroni *et al.*, 2005; W. C. Jesus Junior, comunicação pessoal).
4. Além do CSDaV, o CTV, vírus da família *Closteroviridae*, causador da tristeza dos citros, também é considerado um possível agente causal do MSC. Caso se confirme que o afídeo é o vetor da MSC, não se pode desconsiderar a possibilidade de o CSDaV estar usando o CTV como *helper* para transmissão da doença (Román *et al.*, 2004), visto que o CTV é transmitido por várias espécies de afídeos, os quais são encontrados comumente em plantas cítricas. Desta forma a presença do CTV pode ser necessária para que o CSDaV seja transmitido por afídeo, como ocorre com *Potato aucuba mosaic virus* (PAMV), um potexvirus, que pode ser transmitido por afídeo a partir de plantas co-infectadas com um potyvirus, *Potato virus A* ou *Potato virus Y*, mas não a partir de plantas infectadas apenas com PAMV (Kassanis, 1961). Outro exemplo é o *Groundnut rosette virus* (gênero *Umbravirus*) que depende do *Groundnut rosette assistor virus* para ser transmitido por afídeo (Murant, 1990).

5. As espécies da família *Tymoviridae* transmitidas por insetos infectam principalmente *Graminea*, com exceção do OBDV que infecta linhaça (dicotiledônea). Apesar da maioria dos marafivírus infectarem principalmente monocotiledôneas, o CSDaV afeta laranjeiras, que são dicotiledôneas.

5.2 Expressão da Proteína da Capa do Vírus

A região do genoma viral codificadora da capa do vírus foi amplificada por PCR, utilizando-se *primers* com homologia às regiões flanqueadoras deste gene com a adição de sítios de restrição das enzimas *Eco* RI e *Xba* I. Este gene possui aproximadamente 650pb. Tal número foi compatível com o tamanho do fragmento amplificado na PCR (Figura 31).

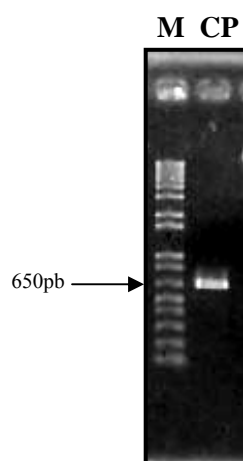


Figura 31. Eletroforese em gel de agarose evidenciando a região codificadora da capa do vírus amplificada na PCR. M: Marcador 1Kb plus (Invitrogen); CP: amplificação da região CP mostrando fragmento de 650pb.

Após a PCR, o fragmento amplificado foi ligado ao vetor pPICZ α A (item 4.2.2) e transformado em bactérias *E. coli* por eletroporação (item 4.2.3). Os plasmídios dos clones selecionados foram extraídos (item 4.2.4) e digeridos com as enzimas *Eco* RI e *Xba* I (item 4.2.5) para verificar quais clones realmente possuíam o fragmento CP de 650pb.

Na Figura 32 é mostrada a eletroforese em gel de agarose de quatro clones (clones 1, 2, 3 e 4) digeridos com as enzimas *Eco* RI e *Xba* I. Todos os clones analisados no gel da Figura 32 apresentaram o fragmento CP, visto que todos apresentaram uma banda de 650pb quando foram digeridos com as enzimas *Eco* RI e *Xba* I.

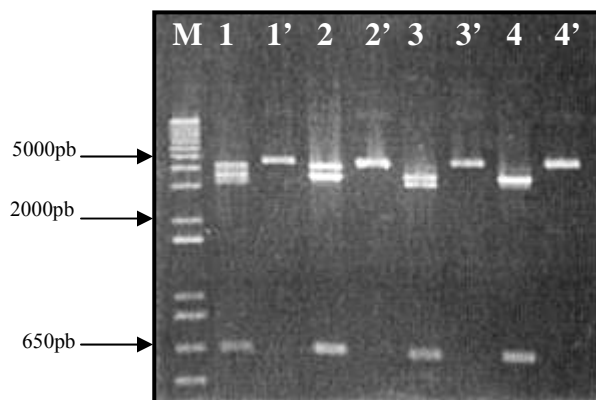


Figura 32. Confirmação da presença do fragmento da capa protéica do vírus nos plasmídios extraídos dos clones selecionados de *E. coli*. M: Marcador 1Kb plus (Invitrogen); 1, 2, 3 e 4: clones digeridos com as enzimas *Eco* RI e *Xba* I; 1', 2', 3' e 4': clones não digeridos com as enzimas *Eco* RI e *Xba* I.

Os plasmídios dos clones que foram positivos quanto a presença do inserto CP foram seqüenciados (item 4.2.5) e mostraram que o gene da CP ligou-se em fase no vetor de expressão. Estes plasmídios foram linearizados (item 4.2.6) e utilizados para transformados de *P. pastoris* (item 4.2.7). Através da PCR de colônia (item 4.2.8), as colônias de *P. pastoris* transformantes foram testadas para verificar se o gene de interesse realmente havia se integrado ao genoma da levedura.

No gel da Figura 33, dentre os 12 clones de *P. pastoris* analisados através da PCR de colônia, oito amplificaram o fragmento CP, visto que apresentaram uma banda de tamanho correspondente (650pb) (Figura 33). Entretanto, no total 40 clones foram analisados quanto a presença do fragmento CP através da PCR de colônia e apenas oito (Figura 33) foram positivos.

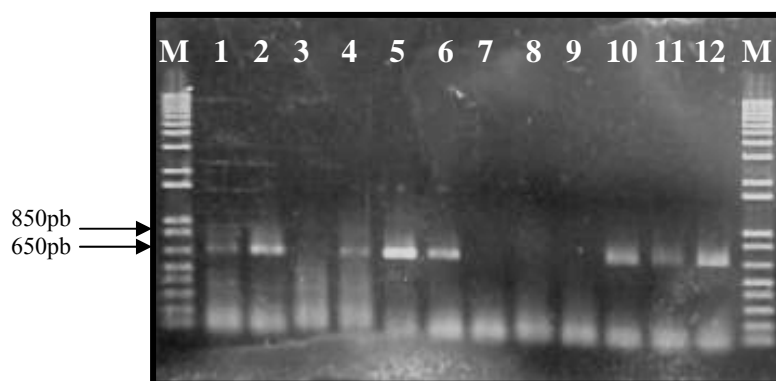


Figura 33. Eletroforese em gel de agarose do produto da PCR de colônia de 12 clones de *P. pastoris* transformados com vetor pPICZαA ligado ao gene CP. M: Marcador 1Kb plus (Invitrogen).

Visto que no teste de PCR de colônia muitos clones testados foram rejeitados e que este é um teste de baixa eficiência, todos os clones de *P. pastoris* que cresceram nas placas de YPDS + Zeocina após a transformação foram testados quanto à expressão da CP, para evitar a exclusão de possíveis clones positivos não detectados na reação de PCR de colônia.

Após a indução da expressão da proteína heteróloga (item 4.2.9), 2μL do sobrenadante coletados em cada tempo do *time course* de cada clone, inclusive do controle negativo, foram submetidos ao teste de DIBA (item 4.2.10) (Figura 34).

Dentre os 19 clones analisados foi possível verificar pelo teste de DIBA a presença da CP em cinco clones (clones 10, 12, 15, 39 e 41) (Figura 34), que aumenta a cada marca de tempo e atinge um pico 96 horas após o início da indução. É importante observar também que o controle negativo (CN; primeira coluna da membrana mostrada na Figura 34) não reagiu ao teste, como era esperado.

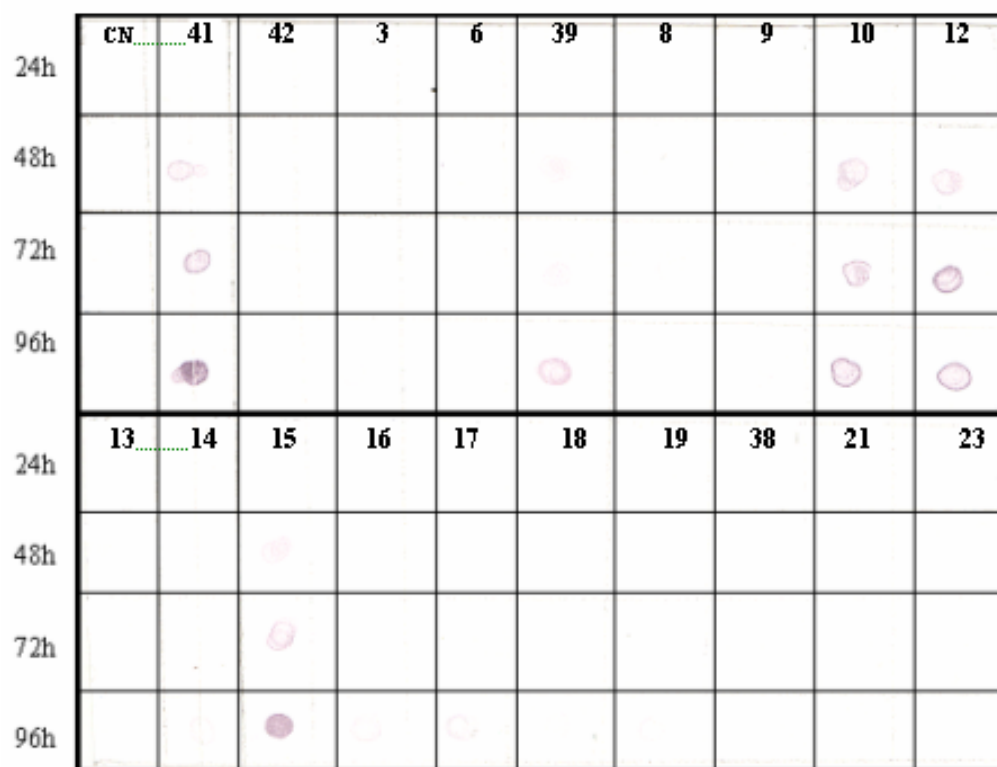


Figura 34. Membrana de DIBA. Esta membrana mostra os tempos 24, 48, 72 e 96 horas após o início da indução da expressão da CP em 19 clones identificados na membrana e do controle negativo (CN).

Dos 40 clones testados no total, 11 mostraram forte reação ao teste, 16 reagiram fracamente e 13 não reagiram. Três controles negativos foram utilizados neste imunoenensaio e nenhum reagiu ao teste. Foi possível observar que os clones que expressaram a proteína da capa do CSDaV apresentaram um gradiente de concentração da proteína no decorrer do tempo de indução: 24 horas após o início da indução não foi detectada a presença da proteína heteróloga; após 48 horas já foi possível detectar a CP, cuja concentração aumenta no tempo de 72 horas e atinge um pico após 96 horas.

Várias proteínas já foram expressas em *P. pastoris*, entretanto, até o momento não há relato de expressão de CP de qualquer vírus de planta nesta levedura para ser utilizada como antígeno para a produção de anticorpos. O sistema de expressão mais utilizado nestes casos é *E. coli*, pela sua fácil manipulação. Este é um organismo unicelular e sua simplicidade faz

dele barato e fácil de trabalhar. Seu ciclo rápido permite dobrar seu número de células em apenas uma hora. Assim como outras bactérias, além do seu próprio DNA circular a *E. coli* pode possuir plasmídeo. Os plasmídios são fáceis de isolar e manipular. Entretanto, esta bactéria também possui algumas desvantagens. Como todo procarioto, *E. coli* não tem organelas encontradas em eucariotos, nas quais as proteínas são modificadas após sua síntese. Essas modificações pós-traducionais envolvem diferentes formas de glicosilação. Várias proteínas podem ser produzidas em grande quantidade em *E. coli*, porém por não sofrerem as modificações pós-traducionais, permanecem não funcionais. Por causa dessas limitações encontradas quando se usa *E. coli* para produzir proteínas de eucariotos outros organismos têm sido estudados como vetores de expressão.

Células de mamíferos, insetos e leveduras são as mais adequadas para substituir as bactérias. Dentre estas as leveduras são as mais utilizadas, pois combinam fácil manipulação genética e crescimento rápido característicos dos procariotos com a maquinaria subcelular para realizar as modificações pós-traducionais das células eucarióticas (Cregg *et al.*, 1993). Um organismo alternativo para a produção de proteínas heterólogas tem sido a levedura *Pichia pastoris*. O protocolo para seu crescimento não precisou ser muito explorado, pois ela pode crescer em condições similares às de *Saccharomyces cerevisiae*. Além disso, *P. pastoris* tem um promotor forte que pode ser utilizado no controle de expressão da proteína recombinante, é capaz de realizar modificações pós-traducionais similares às dos eucariotos superiores e a purificação da proteína heteróloga é facilitada pelo fato da *P. pastoris* não secretar muitas proteínas nativas.

Visto que o CSDaV utiliza a maquinaria da célula vegetal pra expressar suas proteínas virais, a CP é naturalmente expressa sob condições celulares eucarióticas. Assim, anticorpos a serem utilizados no diagnóstico do vírus devem ser capazes de reconhecer a proteína nativa viral. Sabendo-se que *P. pastoris* é um sistema de expressão de proteína de grande eficiência e

reproduz as melhores condições de expressão de proteína de eucariotos superiores, este organismo foi selecionado neste estudo como o ideal para a produção de CP com a finalidade de utilizá-la como antígeno na produção de anticorpos para o diagnóstico sorológico do vírus.

Desta forma, este trabalho é pioneiro quanto a utilização de *P. pastoris* como vetor de expressão de proteína CP de vírus vegetal na tentativa de ser utilizada como antígeno na produção de anticorpos para o diagnóstico do vírus. Estudos posteriores são necessários para confirmar a eficiência da expressão de proteína viral vegetal em *P. pastoris* com tal finalidade.

6. Perspectivas

- Dentre os clones utilizados durante o sequenciamento do genoma do CSDaV foi observado que três apresentavam alta variação gênica na sua seqüência. Isso levantou a hipótese de que os vírus portadores dessa variação gênica podem ter os genótipos causadores da MSC. Dessa forma, é importante fazer o sequenciamento completo destes clones (em andamento), desenhar *primers* específicos para os possíveis genótipos do CSDaV que possibilitará verificar qual genótipo pode estar associado à MSC através de PCR e, assim, realizar a aplicação do RT-PCR para amostras do campo.
- Tendo a seqüência genômica completa do CSDaV é possível montar um clone infeccioso deste vírus e estudar a sua relação com o sintoma de MSC. Esta é uma forma de contornar a dificuldade de se isolar o CSDaV e assim tentar fechar os postulados de Koch necessários para o estabelecimento da relação causal entre a doença e o agente patogênico;
- A partir da CP que foi expressa em *P. pastoris* o próximo passo é purificar essa proteína e inoculá-la em coelhos para induzir a produção de anticorpos e assim possibilitar o diagnóstico sorológico do CSDaV com materiais de campo.

7. Referências Bibliográficas

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

Barros, C.C.P.; Jesus Junior, W. C. de; Teixeira, D. C.; Gomes, C. C.; Dutra, L. S.; Nagata, T. Molecular characterization of tympo-like virus isolated from orange trees with sudden death disease symptom. **Fitopatologia Brasileira**, v. 29, p. S278, 2004.

Bassanezi, R. B.; Busato, L. A.; Sanches, A. L.; Barbosa, J. C. Danos da Morte Súbita dos Citros sobre a produção de laranja. **Fitopatologia Brasileira**, v. 30, p. 497-503, 2005.

Bassanezi, R. B.; Bergamin Filho, A.; Amorim, L.; Gimenes-Fernandes, N.; Gottwald, T. R.; Bové, J. M. Spatial and temporal analysis of citrus sudden death as a tool to generate hypotheses concerning its etiology. **Phytopathology**, v. 93, p. 502-512, 2003a.

Bassanezi, R. B.; Gimenes Fernandes, N.; Yamamoto, P. T. Morte Súbita dos Citros. Araraquara: Fundecitrus/EECB, **Boletim Citrícola**, v. 24, 2003b.

Bassanezi, R. B.; Yamamoto, P. T.; Gimenes-Fernandes, N. Progresso dos sintomas de “morte súbita” em pomares de laranjeiras ‘Valência’ e ‘Pêra’. **Summa Phytopathologica**, v. 28, p. 23. 2002.

Bernard, P.; Gabant, P.; Bahassi, E. M.; Couturier, M. Positive selection vectors using the F plasmid ccdB Killer gene. **Gene**, v. 148, p. 71-74. 1994.

Bové, J. M.; Garnier, M.; Gimenes-Fernandes, N.; Bassanezi, R. B.; Yamamoto, P. T.; Ayres, A. J. Sudden death disease (Morte Súbita Doença) of sweet oranges trees on Ranpur lime: A hypothesis. **Report to Fundecitrus**, February 2002.

Bradel, B. G.; Preil, W. X.; Jeske, H. Sequence analysis and genome organization of *Poinsettia mosaic virus* (PnMV) reveal closer relationship to marafiviruses than to tymoviruses. **Virology**, v. 271, p. 289–297. 2000.

Bransom, K. L.; Wallace, S. E.; Dreher, T. W. Identification of the cleavage site recognized by the turnip yellow mosaic virus protease. **Virology**, v. 217, p. 404–406. 1996.

Brierley, R. A.; Bussineau, C.; Kosson, R.; Melton, A.; Siegel, R. S. Fermentation development of recombinant *Pichia pastoris* expression the heterologous gene: bovine lysozyme. **Annals New York Academy of Sciences**, p. 350-362, 1992.

Cereghino, L. G. P.; Sunga, A. J.; Cereghino, L. J.; Cregg, J. M. Expression of foreign genes in the yeast *Pichia pastoris*. *Quoted in Genetic Engineering: Principles and Methods*. Kluwer Academic. London, v. 23, p. 157-169, 2001.

Cereghino, L. J.; Cregg, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. **FEMS Microbiology Reviews**, v. 24, p. 45-66, 2000.

Chen, J.; Chen, J.; Adams, M. J. A universal PCR primer to detect members of the *Potyviridae* and its use to examine the taxonomic status of several members of the family. **Archives of Virology**, v. 146, n. 4, p. 757-766, 2001.

Clare, J. J.; Rayment, F. B.; Ballantine, S. P.; Sreekrishna, K.; Romanos, M. A. High level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. **Biotechnology**, v. 9, p. 455-460, 1991.

Couderc, R.; Baratti, J. Oxidation of methanol by the yeast *Pichia pastoris*: purification and properties of alcohol oxidase. **Agricultural and Biological Chemistry**, v. 44, p. 2279-2289, 1980.

Cregg, J. M.; Madden, K. R.; Barringer, K. J.; Thill, G. P.; Stillman, C. A. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. **Molecular and Cellular Biology**, v. 9, p. 1316-1323, 1989.

Cregg, J. M.; Madden, K.R. Development of methylotrophic yeast, *Pichia pastoris*, as a host system for the production foreign proteins. **Developments in Industrial Microbiology**, v. 29, p. 33-41, 1988.

Cregg, J. M.; Tschopp, J. F.; Stillman, C.; Siegel, R.; Akong, M.; Craig, W. S.; Buckholz, R. G.; Madden, L. R.; Kellaris, P. A.; Davis, G. R.; Smiley, B. L.; Cruze, J.; Torregrossa, R.; Velicelebi, G.; Thill, G. P. High-level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast, *Pichia pastoris*. **Biotechnology**, v. 5, p. 479-485, 1987.

Cregg, J. M.; Vedvick, T. S.; Raschke, W. C. Recent advances in the expression of foreign genes in *Pichia pastoris*. **Biotechnology**, v. 11, p. 905-910, 1993.

Derrick, K. S.; Beretta, M. J.; Barthe, G. A.; Kayim, M.; Harakava, R. Identification of strain of *Citrus tristeza virus* by subtraction hybridization. **Plant Disease**, v. 87, p.1355-1359, 2003.

Digan, M. E.; Lair, S. V.; Brierley, R. A.; Siegel, R. S.; Williams, M. E.; Ellis, S. B.; Kellaris, P. A.; Provow, S. A.; Craig, W. S.; Velicelebi, G.; Harpold, M. M. Continuous production of a novel lysozyme via secretion from the yeast, *Pichia pastoris*. **Biotechnology**, v. 7, p. 160-164, 1989.

Ding, S.; Howe, J.; Keese, P.; Mackenzie, A.; Meek, D.; Osorio-Keese, M.; Skotnicki, M.; Pattana, S.; Torronen, M.; Gibbs, A. The tymobox, a sequence shared by most *tymoviruses*: its use in molecular studies of *tymoviruses*. **Nucleic Acids Research**, v. 18, p. 181–1187, 1990.

Don, R. H.; Cox, P. T.; Wainwright, B. J.; Baker, K.; Mattick, J. S. “Touchdown” PCR to circumvent spurious priming during gene amplification. **Nucleic Acids Research**, v. 19, p. 4008, 1991.

DPV. Descriptions of plant virus. Disponível em: <<http://www.dpvweb.net>>. Acesso em 15 de setembro de 2005.

Dreher, T. W.; Haenni, A. L.; Mayo, M. A. Genus *Tymovirus*. In: Van Regenmortel, M. H. V.; Fauquet, C. M.; Bishop, D. H. L., Carstens, E.; Estes, M. K.; Lemon, S.; Maniloff, J.; Mayo, M. A., McGeoch, D., Pringle, C. R.; Wickner, R. B. (eds) **Virus Taxonomy**. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, p. 964–969. 2000.

Edwards, M. C. Genus Marafivirus. In: Van Regenmortel, M. H. V., Fauquet, M., Bishop, D. H. L. ; Carstens, E. ; Estes, M. K. ; Lemon, S. ; Maniloff, J.; Mayo, M. A.; McGeoch, D., Pringle, C. R. ; Wickner, R. B. (eds) **Virus Taxonomy**. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, p. 969–973, 2000.

Edwards, M. C., Zhang, Z.; Weiland, J. J. *Oat Blue Dwarf Marafivirus* resembles the Tymoviruses in sequence, genome organization, and expression strategy. **Virology**, v. 232, p. 217 – 229, 1997.

Fundecitrus. Cresce morte súbita. **Revista do Fundecitrus**, v. 112, p. 4-6, 2002.

Fundecitrus. MSC evolui. **Revista do Fundecitrus**, v. 120, p. 6-7, 2004.

Ghanem-Sabanadzovic, N. A.; Sabanadzovic, S.; Martelli, G. P. Sequence analysis of the 3' end of three Grapevine Fleck virus-like viruses from grapevine. **Virus Gene**, v. 27, n. 1, p. 11-16, 2003.

Gimenes-Fernandes, N.; Bassanezi, R. B. Doença de causa desconhecida afeta pomares cítricos no norte de São Paulo e Sul do Triângulo Mineiro. **Summa Phytopatologica**, v. 27, p. 93, 2001.

Gimenes-Fernandes, N.; Bassanezi, R. B.; Yamamoto, P. T.; Ayres, A. J. Árvores de limoeiro 'Cravo' não mostram sintomas de "morte súbita" em área afetada. **Summa Phytopathologica**, v. 28, p. 72, 2002. (Resumo).

Hammond, R.W.; Ramirez, P. Molecular characterization of the genome of *Maize rayado fino virus*, the type member of the genus *Marafivirus*. **Virology**, v. 282, p. 338–347, 2001.

Harakava, R. Obtenção de seqüências de um vírus do gênero *Marafivirus* a partir de laranjeira com sintoma de morte súbita dos citros. **Summa Phytopahologica**, v. 30, p. 101, 2004.

Harlow, E.; Lane, D. **Antibodies: a laboratory manual**. Cold Spring, Harbor Laboratory, New York. p.709, 1988.

Higgins, D. R.; Cregg, J.M. *Methods in Molecular Biology: Pichia Protocols*. Totowa, New Jersey. Humana Press, 1998. *Quoted in* Lin Cereghino, G. P.; Lin Cereghino, J.; Ilgen, C.; Cregg, J. M. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. **Current Opinion in Biotechnology**. v 13, p. 329-332, 2002.

Hollenberg, C. P.; Gellissen, G. Production of recombinant proteins by methylotrophic yeasts. **Current Opinion in Biotechnology**, v. 8, p. 554–560, 1997.

Informativo Centro de Citricultura. Nova doença já causa prejuízos a pomares de SP e MG. **Informativo Centro de Citricultura**, v. 76, p. 1-4, 2001.

Izadpanah, K.; Zhang, Y. P.; Daubert, S.; Rowhani, A. Sequence of the coat protein gene of *Bermuda grass etched-line virus* and of the adjacent “marafibox” motif. **Virus Genes**, v. 24, p. 131–134, 2002.

Jahic, M.; Rotticci-Mulder, J.C.; Martinelle, M.; Hult, K.; Enfors, S. O. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. **Bioprocess and Biosystems Engineering**, v. 24, p. 385-393, 2002.

Jesus Junior, W. C. de; Teixeira, D. C.; Santos, M. A.; Polverari, F. S.; Martins, E. C.; Barros, C.C.P. ; Nagata, T. & Rezende, J. A. Detecção de um vírus da família *Tymoviridae* em plantas com Morte súbita dos citros. **Fitopatologia Brasileira**, v. 30, p.S183, 2005.

Jesus Junior, W. C. De; Bassanezi, R. B. Análise da dinâmica e estrutura de focos da morte súbita dos citros. **Fitopatologia Brasileira**, v. 29, p. 399-405, 2004a.

Jesus Junior, W. C. De; Bassanezi, R. B.; Bergamin Filho, A. Morte Súbita dos Citros ainda tem origem desconhecida. **Visão Agrícola**, v. 1, n. 2, p. 34-39, dez 2004b.

Jesus Junior, W. C. De; Teixeira, D. C.; Tanaka, F. A. O.; Martins, E. C.; Barros, C. C. P.; Nagata, T.; Rezende, J.A.M. Purification of Tymo-Like virus particles isolated from orange trees with Sudden Death. **XVI Conference of the International Organization of Citrus Virologists**, Monterrey, México, p.148., 2004c.

Jesus Junior, W.C. De; Yamamoto, P. T.; Bassanezi, R. B.; Bergamin Filho, A. Morte Súbita dos Citros. In: Zambolim, L. (ed.) **Manejo Integrado de Doenças e Pragas: Produção Integrada de Fruteiras Tropicais**. Suprema Gráfica e Editora, Visconde do Rio Branco, MG, p. 187-222, 2003.

Kassanis, B. The transmission of aucuba mosaic virus by aphid from plants also infected by potato virus A or Y. **Virology**, v. 13, p. 93-97, 1961.

Kimati, H.; Galli, F. Doenças dos citros. In: Galli, F. (Coord.) **Manual de Fitopatologia: doenças das plantas cultivadas**, São Paulo, Ed. Agronômica Ceres, v. 2, p.213-235, 1980.

Libanore, A. C.; Massari, C. A.; Bassanezi, R. B.; Yamamoto, P. T.; Gimenes-Fernandes, N.; Ayres, A. J. Levantamento da ocorrência da “morte súbita dos citros” no Estado de São Paulo e Sul do Triângulo Mineiro. **Summa Phytopathologica**, v. 28, p. 72-73, 2002. (Resumo).

Linder, S.; Schliwa, M.; Kube-Grandenath, E. Direct PCR screening of *Pichia pastoris* clones. **Biotechniques**, v. 20, p. 980-982, 1996.

Maccheroni, W.; Alegria, M. C.; Greggio, C. C.; Piazza, J. P.; Kamla, R.F.; Zacharias, P. R. A.; Bar-Joseph, M.; Kitajima, E. W.; Assumpção, L. C.; Camarotte, G.; Cardozo, J.; Casagrande, E. C.; Ferrari, F.; Franco, S. F.; Giachetto, P. F.; Girasol, A.; Jordão Junior, H.; Silva, V. H. A.; Souza, L. C. A.; Aguilar - Vildoso, C. I.; Zanca, A. S.; Arruda, P.; Kitajima, J. P.; Reinach, F. C.; Ferro, J. A.; Silva, A. C. R. dA. Identification and genomic characterization of a new vírus (*Tymoviridae Family*) associated with citrus sudden death disease. **Journal of Virology**, v. 79, n. 5, p. 3028-3037, 2005.

Martelli, G. P.; Sabanadzovic, S.; Abouh-Ghanem Sabanadzovic, N.; Saldarelli, P. Maculavirus, a new genus of plant viruses. **Archives of Virology**, v. 147, p. 1847–1853, 2002b.

Martelli, G. P.; Sabanadzovic, S.; Abouh-Ghanem Sabanadzovic, N.; Edwards, M. C.; Dreher, T. The family *Tymoviridae*. **Archives of Virology**, v. 147, p. 1837–1846, 2002a.

Maruyama, K.; Sugano, S. Oligo-Capping: A simple method to replace the cap structure of eucaryotic mRNAs with oligorribonucleotides. **Gene**, v. 138, p. 171-174, 1994.

Müller, G. W.; Negri, J. D.; Aguilar-Vildoso, C. I.; Mattos Júnior, D.; Pompeu Júnior, J.; Teófilo Sobrinho, J.; Carvalho, S. A.; Giroto, L. F.; Machado, M. A. Morte súbita dos citros: uma nova doença na citricultura brasileira. **Revista Laranja**, v. 23, n.2, p. 371-386, 2002.

Murant, A. F. Dependence of groundnut rosette virus on its satellite RNA as well as on groundnut rosette assistor luteovirus for transmission by *Aphis craccivora*. **Journal of General Virology**, v. 71, p. 2163–2166, 1990.

Reinach F., Alellyx, Applied Genomics, public conference on CSD, 14 October 2003.

Roggenkamp, R.; Janwiez, Z.; Stanikowski, B; Hollenberg, C. P. Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast *Hansenula polymorpha*. **Molecular & general genetics**, v. 194, p. 489-493, 1984. *Quoted in* Lin Cereghino, L; Cregg, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. **FEMS Microbiology Reviews**, v. 24, p. 45-66, 2000.

Román, M. P.; Cambra, M.; Juárez, J.; Moreno, P.; Duran-Vila, N.; Tanaka, F. A. O.; Alves, E.; Kitajima, E. W.; Yamamoto, P. T.; Bassanezi, R. B.; Teixeira, D. C.; Jesus Junior, W. C.; Ayres, A. J.; Gimenes-Fernandes, N.; Rabenstein, F.; Giroto, L. F.; Bové, J. M. Sudden death of citrus in Brazil: A graft-transmissible bud union disease. **Plant Disease**, v. 88, p. 453-467, 2004.

Roux, K. H. Optimization and troubleshooting in PCR. **PCR methods and applications**, v. 4, p. 5185-5194, 1995.

Sabanadzovic, S.; Ghanem-Sabanadzovic, N. A.; Saldarelli, P.; Martelli, G. P. Complete nucleotide sequence and genome organization of *Grapevine fleck virus*. **Journal of General Virology**, v. 82, p. 2009-2015, 2001.

Sambrook, J.; Fritsch, E. F.; Maniatis, T. **Molecular Cloning: a laboratory manual**. 2ed. Cold Spring Harbor Laboratory Press., v. 3, 1989.

Sambrook, J; Roussel, D. W. **Molecular Cloning: a laboratory manual**. 3 ed. Cold Spring Harbor Laboratory Press., 2001.

Schaefer, B. C. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. **Analytical biochemistry**, v. 227, p. 255-273, 1995.

Staden R.; Judge D. P.; Bonfield J. K. Managing sequencing projects in the GAP4 environment. In: Krawetz, S. A.; Womble, D. D. (eds) **Introduction to Bioinformatics. A Theoretical and practical approach**. Human Press Inc, Totawa, 2003.

Targon, M. L. P. N.; Astua-Monge, G.; Freitas-Astua, J.; Kishi, L.; Souza, A. A.; Santos, F. A.; Muller, G. W.; Machado, M. A. **Evaluation of CTV haplotypes in plants exhibiting citrus sudden death symptoms**. 10th ISC Congress, abstract P83, 2004.

Thompson, J. D.; Higgins, D. G; Gibson, T. J. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic acids research**, v. 22, p. 4673-4680, 1994.

Tschopp, J. F.; Brust, P. F.; Cregg, J. M.; Stillman, C. A.; Gingeras, T. R. Expression of the LacZ gene from two methanol-regulated promoters in *Pichia pastoris*. **Nucleic acids research**, v. 15, p. 3859-3876. 1987a.

Tschopp, J. F.; Sverlow, G.; Kosson, R.; Craig, W.; Grinna, L. High-level secretion of glycosylated invertase in the methylotrophic yeast, *Pichia pastoris*. **Biotechnology**, v. 5, p. 1305, 1987b.

Wegner, G. Emerging application of methelotrophic yeast. **FEMS Microbiology Reviews**, v. 87, p. 279-284, 1990.

Yamamoto, P. T.; Jesus Junior, W. C. de; Bassanezi, R. B.; Sanches, A. L.; Ayres, A. J.; Gimenes-Fernandes, N.; Bové, J. M.; Transmission of citrus sudden death agent by graft-inoculation under insect proof conditions. **Fitopatologia Brasileira**, v. 29, p. S206, 2004.

Yamamoto, P. T.; Jesus Junior, W. C.; Bassanezi, R. B.; Sanches, A. L.; Ayres, A. J.; Gimenes-Fernandes, N.; Bové, J. M. Transmission of the agent inducing symptoms of citrus sudden death by graft-inoculation under insect-proof conditions. (Abstr.) **Fitopatologia Brasileira** (suplemento) v. 28, p. S265, 2003.

8. Apêndices

APÊNDICE A

Seqüência nucleotídica do genoma completo do CSDaV, a qual foi submetida ao banco de dados de seqüências nucleotídicas GenBank e pode ser acessado pelo número DQ185573.

```

1  gtccccctgtg atcgtctctc ccgcctcatc gccggaaga tatttttggt ttaacttttc
61  tttgcacacc tacgccagat ctacgtgcct taggtcatct aagccgctat ggatcgcac
121 tctgccccga ttcccgtcgc gcccgttctt gccggcccga ccgagtatgt tccttacc
181 cacactaacc ccctcctacc ccgaggcgtt ttcacctccg ggcccatcca accccgcctt
241 catttccttc ctcacatgac ccaagatgac cccatccgct gctatcgacc cctcacctat
301 gccaacctac tccgctatga ccgctccgcc tcatcgetca agacttcgcc cgtcaagctc
361 ccgctgaccg gtgggtactct tgccgatgac atcctttccc tagctcccac cactcacctg
421 gataccattg ccacccccct tatggaagcc cttgctgaac cttaccgcca gtccttgacc
481 acctaccctt ggcacattcc aaccaatctt cagcccttcc tcacctctg tggaaatcacc
541 actgctggcc aaggctttaa ggcccaccct caccctgtgc acaagaccat tgagaccaat
601 ctccctctca atgtctggtc ccactacgcc accacccccca gtggcgctca gttcatgaaa
661 ccatccaagt tcgagaagct cagaatcaaa cagcctaact tctccaagct ctacaactac
721 cgcacacag ccaaggactc caccgttat cctccactt caccagacct acccactgag
781 gacacctgct ttatgcatga cgccctgatg tactactccc ccggccagat ctgtgacctc
841 ttcatctctc gccccagcct ccaaaagctc tatgcctctc tcgtgggtcc ccctgaaagt
901 gacttcacca ccatctccct cttccccgac ttgtatcgct acaggattga gaaggaccag
961 cttgtctatg agctggagca gaaccggcc cacaactaca ttcagcccag ctccgccatt
1021 gactggctga agaccaccac catccgctgc caagacctaa ccctcacaat ctcccgcctt
1081 gactcctggg gccagattca ctctcttttg atccaaagag gtagggcacc catccacctt
1141 gaagaggact ccatcagctt ccgcaactcc aaagctgtcc tccttctcga accagcttct
1201 ttatcccagg gggttcgtga ccgccttgtt cctgctgatg tctaccagcc tctcttcac
1261 tatgtcagag ccgttcggac cctccgtgtc actgaccccg ctggggttgt ccgaactcag
1321 atttcaaagc ccgagtactc ctgggtcacc tccttccgct gggacaactt tcccccttg
1381 gcccttgcca ctgctccaca cagaccccac accacttact tcctctcaa tccactgcy
1441 gcccggtgat cccactggtt ccgcacccat accttggtcc tgctctctgg cgcactgca
1501 gccgcccga gctcctcat gaccgccagt tggggcttcc gcgctctcat ttcttctc
1561 cttgtctccc tctccatctg caagcgctgg ctgagggccc ctccccatct cctttggccc
1621 gagaaggctc catggttcca actcacactg cggcctaaag tcaccggccc catgattgac
1681 ctgcccaccc tcggaccatt cagattgttc ccatccacct gtgctaaatt gggcgccaag
1741 tatcctgccc tcgccactt gttgcagca gcaccaggc ccctctggcc cctgaaagt
1801 ggccttgccc ttgccaccgt gcccgctgtt ttattcttgt ggcggaagt cattgtccc
1861 gactctcccc aggacatgca cgacagctac catgctatgt ttcacccaca gcctgggga
1921 ctctctctgg cccgcaaggc catttgctgt gacagagctc cattcctccc catccctctt
1981 gtgcccagct ctgacttcaa agccccacc acaccagcta cccctctact gacctccatc
2041 ccaatcaagg gtgtggagcc tcaagttcct ggagaaggag tacctccgca gtcggcttca
2101 tcaaccggcc cggcatcgga ctcccgtgc gccccgcaac cagcttcatc aactggtccg
2161 gatcagcccg ccaggacat gagtgtgctc cccaagctc ccacaaaatc caaagtact
2221 tttgctcaac cccttgagag cgtgacacct gtagttccag gagcaggaga acctccgag
2281 tcggcttcat caaccggccc ggcacgtgct tcccgtctg acccgcaaac ggcttcatca
2341 accgttcccg atgctcccac cctggacagc agtgtgacct ctccaccgat tatctacca
2401 gttgaccatc tcagaaatga cttcggcccc tgccgttgtt cactctgcca accgccccag
2461 ccagccccag ttccctctgc cctcccact gtctctgacc acaaggaatc ccaagatggt
2521 gaggccctcg cttccgcctt ccaagctctc gggctcgcct cgtctccacc agctcctctg
2581 tctcagaatc taacctcga gtcctcagga gccatgcatg cctcttctg ggtatcagctc
2641 tcttccccat catcggactg ggatccctcc cccgtggctc gcgacagctc tgccaccggt
2701 ccaccaggca tgtactctga tctcttccc gctccttatc tcccaggcac cggccaattc
2761 atcttccgct ccagagccaa cggccgggcc aacattccct accccgacat ggattgtctc
2821 ctactttcca tcgagcaggc tacccgctt cccaaggagg ccctctggga cactctctgt
2881 gccacctgcc ccgactctct gcttgacct gacaccatcc gccgagtcgg cctctccaca
2941 gaccattttg ttatcttctc ccaccactac tccctcaggt gccggttcca taccgcccac
3001 ggtgtcattg agctgggcat ggccgacgcc acttcatctt ttgatatcga ccacactgcc
3061 ggcaatcctg gccacttctc cctccggcag tctgccacgc cccggtcaa tgggtggcatt

```

3121 gctcaggacc ttgctgtggc cgccttaagg ttcaacattg acggcactct cctccccatc
 3181 cgctcagttc atgtctattc cacctggccc aagagagcca agaacctgtc ctcaaatacg
 3241 aaaaacggct ttgatggaat catggccaac attcatccca ccaagaccaaa tgaatccaga
 3301 gaaaagattt tggccctgga ctcccagett gatattgcag tccggagatc tgtccgcctc
 3361 atccacattg ccggctttcc aggctgcgge aagtccctcc ccattctccg cctcctccgc
 3421 actcccacct tcaggaactt taagggtggt gtgccaactg tcgagcttcg agctgaatgg
 3481 aaagcaatca ttggcctccc aacttctgag gcatggcgca tcggtacctg ggaatcctct
 3541 ctgctcaaat ctgcccgggt tctcgtcatt gatgaaatct acaagatgcc ccggggctac
 3601 attgatctcg ccattccactc tgatccaacc attgagatgg tcattgctct cggtgacca
 3661 ctccaaggcg agtaccactc cactcatctc tcttccacca actcccgcct tcttgctgaa
 3721 ccacaccacc tctccatgta tctcgactac tactgtatgt ggtcccatag agtccctcag
 3781 aatgctcgctg ctttcttcca catcaagacc acctccaagc agccaggctc cggccgctac
 3841 cagcgagagc ttcccaactc caggatctct gccaaactctc aaaatgcagg ccatactctt
 3901 cagcaatgcg gctatgccgc cgtcaccatt gcctcaagtc aggggttccac ttatgacaat
 3961 ggggcctgca tccatcttga ccgcaacagc tcctgtctct ccctgtctca ttccatggct
 4021 gctctcaccg gctccaagggt tgggtgtcatc ttcaccgggg atcctgcca gctgtccaat
 4081 gcccggagct ccaacagaat gttctctgaa ttcttctcag gccgcaccgg ccctcttcat
 4141 gactgggttc acaatgagtt cccaaaggcc actgtcctca ccgagccct caagactcgg
 4201 gggcccgcgc tcaccgggtc tgcctcacca tactccaagg tgttcccatc cggcaagat
 4261 tccaccccgg cctcaaacc tgatttccag ggggacgttc gttcttctcg gggccccatc
 4321 gttctgggct ccggcgagct caatgctcca cagggtctca ctcacttct tcttgagact
 4381 cgccgcccctc ttcactggga catcccctct gccatccctg agagtgccac caggcctgat
 4441 tccactgagc ccaccacttg ccactcctgag ccggtttacc caggggaaac ctttgaaaat
 4501 ctgtctgccc actttctccc ggctcacgat cccactgatc gagaaatcta tgggcagggc
 4561 cagctgcca accagttccc tcacatggac aaagagttcc acctggccg tcagccctg
 4621 agcctcctgg ctgcgattca tcaagagaaa caggacccca ctctgctccc agcttccatc
 4681 cagaagaggg tccgcttccg cccatcagac aggccttacc agatcactcc aaaagatgaa
 4741 atcctgggct aattgctctt tgaaggctc tgccgagcct atcatcggtc ccctttccac
 4801 actgaggcct ttgatcctgt tctgttcgcc gagggtcatca atctgaaatga attcgcccag
 4861 ctctcctcca agactcaggc caccatcatg ggcaatgctc gccgctctga tccagactgg
 4921 cgggtggagt ctgtccgat cttctccaag acccagcaca aagtaaatga aggggtccatc
 4981 tttggttcct ggaaggcatg tcagacttg gccctcatgc atgatgctgt tgtcctatc
 5041 ctgggcccgc tcaagaagta ccagagagtg tttgaccaga gagacagacc ccggcacctt
 5101 tacatccatg ccggcaacac tcttccacag atgagcaact ggtgccagca gcatctcacc
 5161 accgctgtca agctggccaa tgattacacc gcctttgatc aatctcagca tggggaggca
 5221 gttgttctcg aaagaaagaa gatggaaagg ctctccattc cacaggccct cattgacctc
 5281 cacatccatc tcaagaccca tgtctccacc cagttcggcc ccctcacttg catgcccctg
 5341 actggcgaac ctggcacctc tgatgacaat tctgactata atcttgctg gatccaactgt
 5401 gagtacatgg ctgccaacac tcccactatg gtctcaggcg acgactccct cctggactgt
 5461 gagcctccca ctgcacctga atgggtcatc ctccagcctc ttctcagtct ccgcttcaag
 5521 aaagaaaggg gtcggtagc caccttctgt ggctactacg cctcccatgt cggctgtgtc
 5581 cgctcccccg tggctctctt tgccaagctg gccatagctg tcgatgacgg ctccatctct
 5641 gacaaaatgg cctcatacct ctctgaattt gctcttggcc actcccttgg agaccatctc
 5701 tgggaagctt tgcccctcga ggccgttccc ttccaatctg cctgcttga ctcttctctg
 5761 cgccgggccc ccagacacct caagctctct ctcatgctcg gcgaggtccc agaattccatc
 5821 attgcccgca tcgggtcatc cttgaagtgg gcctctcatg ccactctacac cacactctcc
 5881 tctgccgctc gaggggccat tctgagatcc tcccgaaca gcagatccat gccagatgac
 5941 cccgacacca ctctgctaca aggtgaatgg ctccagcact ttcaagtacc attcgatgaa
 6001 tctgacactc tctgctctc cactgggtgt tctctgctc ccactctcac accagaagcc
 6061 ttctccacct ctctcgcttt ctccatggcc agcgatgccc aagcaggtcc ggccccagct
 6121 cgcgatgatc gcggtgaccg ccagcctcgc ctctctgctg ctctctcgtg tgcctgaagt
 6181 ggtctcaatg ccccgtcggc cgactaccgg ttccagtggt tcgctgcctc ctacgacgga
 6241 tcagaagcca agaacctaa gtagatctc tctggctctg ccactctcac caaagtcatg
 6301 gccaaactacc gacatgctga gctcacatct gttgagctgg aggtctgccc tcttgctgca
 6361 gccttctcca agcccactct tgtgtcggcc gtctggacca ttgcctccat ctccccagct
 6421 tccgctctcg aaacctccta ctatggcggt cgcctcttca ctgtcgggtg tccagctctc
 6481 atgtccagta ccaactcact tccagcagat ctaccgcc tcactctgt gctcaagggc
 6541 cctgtaagt acaccgact cccagattc tctactccg tctactcaa cggcggaaact
 6601 aagggacca acctctgcac catcatctc cggggagtgt tccgctcag cggcccctcc
 6661 ggtaatctcc tcgcttaggc gagccacttc aggtgaagga aaacacctcc tgggtctcagc
 6721 caggtaatga tgctaaacct ccccgcctca agcagcaatg cctaggggtg ccggctgatc
 6781 caaagaccgt ttttcttcta tttaat

APÊNDICE B

Alinhamento múltiplo entre a seqüência nucleotídica do CSDaV (DQ185573) e de outros vírus da família *Tymoviridae*, através do software Clustal W.

CLUSTAL W (1.82) multiple sequence alignment

```

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      GCACATTAGTACTGACCACCAGCCCGACTTTGACGCCACCCTGTCCAAGTCCGGCCCGTG 60
PnMV      -----
TYMV      -----

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      CTCACTCTTTTGTFTTTCTTTGTCTTTTCCCCTGCTCCCACGTTGACGACCTCGCGGC 120
PnMV      -----
TYMV      -----

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      CGTCCTCCTCGCTTCGTGCTCTGCATCCTTTAACTTCCTCCCCCTCTTCCTCCTCGCCCT 180
PnMV      -----
TYMV      -----

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      TTCCTTTCTCATCCACCTCCCTCCCAACCTCTCTTCCTCCCTTCACTCTCTTCTTCTTGC 240
PnMV      -----
TYMV      -----

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      -----GTCCCCTGTGATCGTCTCTCCCGCCTCATCGCCGAAAGATATTT 45
PnMV      -----
TYMV      -----
GFkV      TTTTCCCTTTCTCTTCTTTCTTTCTCCTCCTGCTCTCCCCATCCTCCCGCCATGGTGCTC 300

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      TTGTTTAACTTTTCTTTGACACCTACGCCAGATCTACGTGCCTTAGGTCATCTAAGCC 105
PnMV      -----
TYMV      -----
GFkV      ACGCTCTGCTTCCACTCGGACCGTGACCATCACTTCCTCTCCACCCTCCCCCTCGCCGAG 360

MRFV      -----
CSDaV     -----
OBDV      -----
GFkV      GCTATGGATCGCATCTCTGCCCGCATTCCCGTCGCGCCCGCTTCTGCCGGCCCGACCGAG 165
PnMV      -----
TYMV      -----
GFkV      ATCAGACGTCTCGCGCCAGGCGGCTTCTCCGTCTTCGTCCCTGAAGACCTCCTCCCCAC 420
    
```

PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	TATGTTCCCTTACCCCCACACTAACCCCTCCTACCCCGAGGCGTTTTCACCT--CCGGGC	223
OBDV	-----GTGTCCCAGTGTATTATTCCGCTC--AGTTTC	31
GFkV	CTCTCCCCCTCCCCCTCCTGCCCCACCCCCCTTCCCCGCGAGCTCACCTGGAGCGCC	480
PnMV	-----	
TYMV	-----	
MRFV	-----GTCGACGTGCGATTCTGCACCAGCTTTCGCTCGTCCAGAACCAG----	44
CSDaV	CCATCCAACCCCGCCTTCATTTCCCTCCTCACCATGCCAAGATGCCCCCATCCGCTGTCT	283
OBDV	AGATCTGCCGGAATTCTCCAAGCATCCCGCCCCAAAGCCGCTGCTTAAAATCTG-ATC	90
GFkV	GCTTGCAACCCCACTTCCCCCATCTCTTTGCGCCACCACCGCCACCCCCCTGC	540
PnMV	-----GGAA	4
TYMV	-----GTAATCAACTACCAATTC	19
MRFV	ATCATCCTTTTCG-----GCCTCTTTTCTTCGCACT--CTTCTTCGTTGCACTCTTCTAGA	97
CSDaV	ATCGACCCCTCACCTATGCCAACCATCTCCGCTATGACCGCTCCGCTCATCGCTCAAGA	343
OBDV	TTCTCCATCTTGTCAAGTGTGTTATGACCACATA--CGCCTTCCACCCGCTGTCCCCA	148
GFkV	ACCCGCCTTACCCCCGCCCCCCCCCCCCACTCCTGCCCCCTCCCTCAGCCTCCACCTTG	600
PnMV	ATACACATCTCCAGTCAGCTCCTTTACGCAACTGACTTTTATTTATCCATCATCTGTTG	64
TYMV	AGCTCTCTTTTGACAACCTGGTCTTATACCAACTTTCCGTACACTTGCAACCCTCGTAAGA	79
	*	
MRFV	T-----GTCTAGTTTCC--TGCCTGGTGGTCACTCTCCTTAGCGGAGTCGAGTCCCTT	147
CSDaV	CTTCGC-CCGTCAAGCTCCCGTACCCTGGTACTCTTGGCCGATGCCATCCTTTCCCTA	402
OBDV	CCCCGA-CCTCCTTCGCCACTATCACTGGGGGTGGTTGAAGGATGTTATCGAAACCCCTC	207
GFkV	CCCGTGTCTTCCACGGCGGCCCCCCAGCCGCTTCCAACCCGATCGACTTCTCT	660
PnMV	T-----TTGCCCCGGTATATCCCATGGCA--TTCCAAGACGCTTTCAACAACCTA	113
TYMV	C-----AATTGCAAATGAGTAATGGCC--TTCCAATTAGCATTGGACGCCCTT	125
	* * * * *	
MRFV	ACCCCCACCACCCATCGCGACACCATCACTGCACCC-ATTGTGCGAGTCCCTTGCCACTCC	206
CSDaV	GCTCCCACCACCTCACCGTGATACCATTGCCACCCCTTATGGAAGCCCTTGTGAACC	461
OBDV	TCGTCCACCATCCACAGAGACACGATCGCAGCACCC-CTCATGGAGACCCCTCGCCTCGCC	266
GFkV	CACAACACAATCCAGAAGGATACGATCGCCTCCTCC-ATCATTGCCGCCCTCAACCCCTTC	719
PnMV	TCCAGCACCGTCCACCGGACACGGTTGCCGCACCT-CTCTTAGAGAGCATTGCCCAACC	172
TYMV	GCACCCACGACTCACAGAGAT-CCCTCTCTGCATCCGATTCTCGAATCCACAGTAGATT	184
	*** ** * * * * * * * * *	
MRFV	TCTTCGCCGATCCCTCGAACGCTACCCTTGGTCCATCCCAAAGGAGTTCCTACTCGTTCT	266
CSDaV	TTACCGCCAGTCTTGGACACCTACCCTGGCACATTCCAACCAATCTTACGCCCTTCTCT	521
OBDV	TTACCGAGACTCCCTTCGCGACTTCCCTTGGGCGTCCCCGCTCCGCCCTGCCCTTCTCT	326
GFkV	CCTCACCTCCTCCTCACCTCTACCCTTACGCGTCCCCCGCGTGGCCCTCCGCCCT	779
PnMV	TCTCCGCGACTCTCTGAATTATATCCCTGGGCGATCAACGCCGAGAACTTGACCTGTCT	232
TYMV	GATTCGCTCCTCGATACAGACCTACCATGGTCCATTCCGAAGGAATTTGCCCCCTACT	244
	* * * * * * * * *	
MRFV	CCACACCTGCGGCGTCGACATCTCTGGCTTCGGCCATGCCGCTCACCCCCACCCGTTCA	326
CSDaV	CACCTCCTGTGGAATCACCACTGTGGCCAAGGCTTTAAGGCCACCCCTCACCTGTGCA	581
OBDV	CCAGGAATGTGGCATCACGGTCGCCGCCACGGTTTCAAAGCTCATCCCCACCCGTGTTCA	386
GFkV	CAACCAAGCCGGCATCCCCGCCACTTCTACGGCCACCAATCCCACCCCCACCAATCCA	839
PnMV	GAAGAAATTTGGCATTTCTGTCTCCGGCCTGGGTCAATCAAGCCCACCCCTCATCCGTTCCA	292
TYMV	CAACTCCTACGGCATCCCAACATCTGGTTTGGGAACATCCACCACCCCGCCGCCA	304
	* * * * * * * * *	
MRFV	TAAGACCATCGAGACCCACCTCCTTCTCGATGTCTGGCCCAACTACGCTCGGGGTCCGTC	386
CSDaV	CAAGACCATGAGACCAATCTCCTCTCTAATGTCTGGTCCCCTACGCCACCACCCCGAG	641
OBDV	CAAAACCATCGAGACCCACCTCCTCCACAAGGTTTGGCCCTCACTATGCCAAGTCCCTTC	446
GFkV	CAAGACCATGAAACCCATCTCCTCATGAGCATTGGGCCAACCGTGGCACTCTCCGCTC	899
PnMV	CAAGATCATCGAAACACATCTCCTTTTCCAACATTGGAGCCACAACCTGTGAGAGGACTC	352
TYMV	CAAGACAATCGAGACTTTTCTCCTTTGCACCCACTGGTCTTTCCAGGCCACCACTCCAG	364
	* * * * * * * * *	

MRFV	AGACGTCATGTTTCATCAAGCCTGAGAAGTTCGCCAAGCTCCA-----	428
CSDaV	TGGCGTCATGTTTCATGAAACCATCCAAGTTCGAGAAGCTCAG-----	683
OBDV	TTCCGTCTCTTCATGAAGCCCTCGAAGTTCGCCAAACTCCA-----	488
GFkV	CACGGTCATGTTTCATGAAGAGGTCGAAGTTCGACAAGCTCCGCGTCTCCAACGCGGCTCT	959
PnMV	TACAGTGTCTTTTCATGAAGCCCGAAAAATTCAAAACTCGC-----	394
TYMV	CTCCGTCTATGTTTCATGAAACCCAGCAAGTTCAACAACTTGC-----	406
	** * ***** ** ** *** ** **	
MRFV	-GTCCCAGACGCC--AACTTCGCCACCTCATCAACTACCGGCTCGTGCCTCAAAGACAC	485
CSDaV	-AATCAAACAGCCT--AACTTCCTCAAGCTCTACAACCTACCGCATCACAGCCAAGGACTC	740
OBDV	-GCGGGCAACGCC--AACTTCCTCCGACTCCACAACCTACCGCTACCGCCAAAGACAC	545
GFkV	CGTCAAATCGGCCTCAAACCTTCCCTCACCTCTCAACCCCATCCTCACCGCCGGGACGC	1019
PnMV	--CTCTTTCAACCCT-CACTTCAAGCATCTTCTCAACTACCGCTTTCACGCCAAGGACAC	451
TYMV	---CCAGGTGAACCTCAAACCTTTCGGGAATTGAAGAACTACCGCTGCACCCCAACGACAG	463
	* **** * *** * ** ***	
MRFV	CACCCGGTACCCC---TCCACTTCCACGAA--CCTCCCGACTGCGAGACCGTCTTTCAT	539
CSDaV	CACCCGTTATCCC---TCCACTTACCAGA--CCTACCCACTGAGGACACCTGCTTTAT	794
OBDV	CCC CGGTATCCT---AACACTTCAACCTC--TCTCCCGACACCGAGACCGCTTCAT	599
GFkV	CGACCGGTACACCACCTCCCCCTCCCCGACACCTCCCCTCCACCCTCTCTACTTCAT	1079
PnMV	AACCCGCTTTCCA---GAAACCGGTGCAG--TCTCCCGACACTTCCACCGTGTTCAT	505
TYMV	CACTCGTTACCCC---TTCACATACCAGA--CCTTCCCGTTTTCCACCATTTCAT	517
	** * * * * ** ** * ** **	
MRFV	GCATGATGCTCTCATGTATTACACCCAGGGCAGATCGCTGACCTCTTCTTCTCTGTCC	599
CSDaV	GCATGACGCCCTGATGTAATACTTCCCCGGCCAGATCTGTGACCTTTCATCTCTCGCCC	854
OBDV	GCATGACGCCCTCATGTATTACACCCCGCTCAAATFGTTGACCTGTTCTTTCTGCCC	659
GFkV	GCATCATTCCCTCATGTAATACTTCTCGCCAGCCAAATGCGGGCTCTTCTCGCCGCC	1139
PnMV	GCATGACGCCCTTGACGTACTACAACTTCCAGATCATGGATCTGTTCTCCGATTC	565
TYMV	GCACGACGCCCTGATGTAATACTTCCCTCCAGATCATGGACCTGTTCTTGAGAAACC	577
	*** *	
MRFV	CCAGCTCCAGAAGATCTATGCCTCCGTTGTGCTTCCGGCTGAGTCGAGCTTACGCATCT	659
CSDaV	CAGCTCCAAAAGCTCTATGCCTCTCTCGTGGTTCCTCCGTAAGTGACTTACCACCAT	914
OBDV	GAAGCTCGAGAACTGTACGCCCTCCTTGTGCTCCCCCGAGTCTCCTTACCTCTAT	719
GFkV	CTTCTCGAGAGGCTCTACGCCCTCTCTCGTCTCCCGGCCGAATCCACCATCGGCAGCCA	1199
PnMV	AAACCTCAACAGCCTCTACGCTTCCGTGGTGGTTCCTCCGCGGAGAGCTATTACCAGCA	625
TYMV	AAACCTCGAAGCTCTGTACGCCAGCTCGTAGTACCACCCGAGGCCATCTTTCCGACCA	637
	*** * * * * * * * * * * * * * * * * * *	
MRFV	CTCGCTCCATCCGAAATTTACCGCTTCCGTTTTTCAGGGTCTGACTTGTCTACGAGCC	719
CSDaV	CTCCTCTTCCCCGACTTGTATCGTACAGGATTGAGAAGGACCAGCTTGTCTATGAGCT	974
OBDV	CTCTCTCCATCCAGATCTTTACCGTTTTCGTTTTGACGGGACCGTTTGATTTATGAGTT	779
GFkV	CCCTTTCTTCCCCTCTGTGTACCGCTACCGACACCACCGCGGAGCACCTCCATTACGTGCT	1259
PnMV	TTCTTATACCCCGAGGTCTACCAGTATAAGACTATCAGGAGCCATCTCCACTACGTCCC	685
TYMV	ATCTTCTTCCCAGAGTTGTACACGTACACGACGCCCGCCACACTCTTCACTACGTCCC	697
	* * ** * * * * * * * * * * *	
MRFV	GGAGGGCAACCCAGCCGCAACTACACGCAGCCGCGTTCCGGCCCTCGACTGGCTCCAGAC	779
CSDaV	GGAGCAGAACCCGGCCACAACCTACATTCAGCCCCGCTCCGCCATTGACTGGCTGAAGAC	1034
OBDV	GGAGGGCAACCCGCCCACAACCTACACCAACCTCGATCCGCCCTCGACTGGCTCCGCAC	839
GFkV	CGAAGGCAACCCATCCTCTCTTACACCCAGCCCTCACCGCCACCCAGTGGCTCACCCAC	1319
PnMV	TGAGTCCAATCACTCCTCGGCTTACAATCAACCTGCCGACGCCCTTCTCTGGCTCGAGCT	745
TYMV	GGAAGGTCACGAAGCCGCGAGCTACAACCAACCTCCGACGCCACTCTTGGCTCCGAAT	757
	** * ** **** ** ** *** *****	
MRFV	CACCGTTTTACCGTTGGTTCACGAGTTCTTCTCCGTCACCTTCTCGACTCTTTCGGCCC	839
CSDaV	CACCACCATCCGCTGCCAAGACCTAACCTCACAATCTCCCGCTTACTCCTGGGGCCC	1094
OBDV	AACCACCATCCGCGGACAGGCGTTTCTCTCACCGTGTCCAGGCTCGACTCGTGGGGTCC	899
GFkV	CTCTCCATCACCGCCGGGACCTCCACCTCACCGTCACCGTCTCGAGTCATGGTTCTC	1379
PnMV	TTGACCATCCAGGGCCCTCCTTCTCCCTCACCGTACCAGGCTCGACTCTTGGGGCCC	805
TYMV	CAATTCAATTCGCCCTCGGCAACCACCACTCTCAGTGACGATCCTGGAATCCTGGGGCCC	817
	* * ** * * * * * * * * *	
MRFV	GGTCCATTCCCTCTCATCCAGCGTGGCCGCCCTCCCGTCTTTC-----	883
CSDaV	AGTTCACTCTCTTTTGTATCCAAAGAGGTAGGCCACCCATCCACC-----	1138
OBDV	CTGCCATTCCCTCTCATCCAGCGCGGCAATCCCCCATGCACG-----	943
GFkV	CGTGCACTCAATCCTCATCACCCGCGGTGTTTCGCCCCCTCGAGC-----	1423

MRFV	CACCGACCGTCTCAGGCTG----	TTCGGCTTCGACATCATCGGCC--	GGAGGTTCTGGC	1354
CSDaV	CATCTGCAAGCGCTGGCTGAGGGCCCT	CCCCATCTCTTTGGCCCGAGAAGGCTCCATG		1634
OBDV	CATCCAAGGCCTCGCGTA---	GCCCCCGGTGGC-TCATCACTCC-----	CCCTCAT	1413
GFkV	GGTCACCCAGCTCACCATC----	CTCGCCACCATTTCACCCCCC--	CAAGATCCTCG	1897
PnMV	AATCTGCTGCCCTCAGGCTCT---	TTGCCTAGCAACCTCGGTTTCCG-----	GAGTCTGGA	1318
TYMV	CATC-ACAGCCTTCCGAAGGGAGTCT	ACCAGAAAGGAGCCCCCACCACCCCTCGACG		1410
	* *	* *		
MRFV	-CTCGTCTCCCATTCATGGCCCCGAGCCA	AAGGTTTCTTTGGGAGACCCATCCAGCCTGC		1413
CSDaV	-GTTCCAACCTCACACTGCGGCC	TAAAGTCACCGGCCCATGATTGACCTGCCATCTCTCC		1693
OBDV	-CTCTCCGCCCCCTCCACCCCAAGCT	TCGCTCTCTTCT---CCAGCGCACT--	C	1466
GFkV	-CTCGCCTTCC---CGTCGCCCTCAAGCGCT	TAAATCCCAAAG-----CGCTGCTCCCC		1947
PnMV	-CGAGCACGTC-----AAGCTC	CAATGAAGGCCTCTCCGGGCTGGTTTTAGTTATACACA		1372
TYMV	TCTTCCATCTCCAGCAACACCTCCGCAAT	CACCACTCCGCGAT---CTCGCCGTACGCC		1467
	*			
MRFV	CGCCCCGCCAG--TGCTCTTTGCCGACT	CCGCC--TTTGAGTGCCAAAT-----		1457
CSDaV	GACCATTGAGATTGTTCCCATCCACTGTGCT	TAAATGGGCGCCAAGTATCTGCCCTCG		1753
OBDV	CACCGCCCCGGTCTTCTCCGTGGCTCCCGCC	---TCGAGTTTGAGC-----		1511
GFkV	CACCTCCCAT---CCCACCTCCGCC	CCCCCTTCTGGTCCCGGTCTTCA-----		1996
PnMV	CACTTCCCAA---ACCGCCAAGGCCTCGAGATT	TCTACCATAGGCGCGCTGAGCGATT		1429
TYMV	CAGCTTCCCA--CCCCACCAAGACTTCCACA	---CGCGCTCCAGAAAGCTG-----		1515
	* *			
MRFV	---CCTTGCTGG-----	TTTTGGC-----AAACC-GGTGTTCTCCCT		1489
CSDaV	CCACCTTGTTGCCAGCAGCACCAGGCC	CCTCTGGCCCTGAAAGTT-GGCCTTGCCCTT		1812
OBDV	---CTTCCCTTC-----	TCTCGCCCC--ACAAC-TCGCCGTGCTT		1547
GFkV	--CCCTCACCTTC-----	TCCGAGTCCCAAAGCCGGTTCCTCACCT		2038
PnMV	CAAAATTCATGTCCG-----	CCCTCGTTGTTCCATCTCCCTCTCTCTCC		1474
TYMV	---CATTGCTGCT-----	CCTCCGACCGATATCGCCCTCTTGACAGCG		1556
	* *			
MRFV	CCC-----CCTTCTGGTTCGCTCTGT-	-----TCCCCAC		1517
CSDaV	GCCACCGTGCCCGTCTGTTTATCTTGTGG	CGGAAGTTCATTGGTCCCGACTCTCCCCAG		1872
OBDV	TCCA-----TTCCTCGCTCGCCTTCTC-	-----CCCCAG		1575
GFkV	TCCCC-----ATCTCCGGCCAAACCCAC-	-----ACCCTCTCCG		2072
PnMV	CCTGCTTGCGTTCCGCATCTATTTCCAAGTGC	-----ACGCAGCCCATGCCCCCTCAG		1526
TYMV	ACCC--GTTCTTTCGGTCCGAACAGAAGTC-	-----CATGCTCCCG		1595
	* * *			
MRFV	GGCTCT-----CCGCCGTCTGGGTCGCCT	ACTCAGCCTTGG-----CCCTAGCA		1563
CSDaV	GACATGCACGACAGCTACCATGCTATGTTT	CATCCACAGCCCTGGGACTCTCTCTGGCC		1932
OBDV	AAAC-----CCATCGACCCCTGGGTCTG	CGGAGCCTCG-----CTGTCCG		1617
GFkV	ACAGTCCATGTCCCTGCCATCCTTTCGCC	CCCCAACGCCCCAGCGCCCCCTCATCTT		2132
PnMV	AGCCTTACAGATTCTACCACAATTACCAT	CATCCAAGCAAGTGG-----TCCTTGATG		1582
TYMV	AACGCCGAACTTTCATGGACCCTGAAGCG	TTCGCGCTGCCTTGGCAGGCCTCCCTAGTC		1655
		*		
MRFV	-GCCGTGCCGCTGGCTGCGTTGG-----	CTTTGCGCTGGTTTTACGGCCCGGACT		1612
CSDaV	CGCAAGGCCATTTGCTGTGACAGAGCTCC	ATTCTCCCATCCCTCTTGTGCCAGCTCT		1992
OBDV	-GTTGCTATACCCGCGCCTCCC-----	TCGCCGTTGCTGGTTCTTCGGCCCCGACA		1669
GFkV	CGCGGGCCTGCTCATTTGGGACTGTCCCT	GTTCTGTACGGGGCTACCGATGGTTTCG-TCT		2191
PnMV	GAGTCGCCGGCTCACCGTGGTCA-----	CTCCCGAGGCTTCTCTGCCTTTCGAACC		1633
TYMV	--CTCCTCTCTGTGCGAATCATC-----	CGTACTGCTTCAAACTGTTCTCCCCACC		1708
	*	*		
MRFV	--CTCCCCAGGCTCTTACGATCAG-TATCAT	GCCAC---CTTCCATCCCGACC----	C	1661
CSDaV	GACTTCAAAGCCCCACCCACACCAGCTAC	CCCTCTACTGACCTCCATCCCAATCAAGGGT		2052
OBDV	--CCCCCAAGCCATGCACGACCGA-TACC	ACCCAT---GTTCCACCCAGAGA-----		1718
GFkV	CCCGCTTCGACCCCCAGACGGTGTACA	ACCCTTACTCTGACCTTCTTACCGCCCCA--	C	2249
PnMV	--TGTTCAAATCCCTGCG-GACCAATCG	CGCCACAATGGACCCCGGCCCCCTC----	C	1686
TYMV	AACTCTCCAAGCCCAACACGAC-AC	TACCACCGACA---TCTTCAACCTGGATCC--	T	1761
	* * * *			
MRFV	GTGGACTCTCGA-----	CCTTCTCGTCGGCTCCG---CCGGTTTG		1699
CSDaV	GTGGAGCCTCAAGTTCCTGGAGAAGGAGT	ACCTCCGAGTCGGCTTCATCAACCGGCCCG		2112
OBDV	GTGGCGCTCAC-----	CCTGCCAGG-----GGCCCC		1746
GFkV	ATGGACCTCAC-----	CTTCGAGCGTG-----AGCCCC		2278

PnMV	AGAGCTGTTCGAT-----CTCGCTCAGATG-----AGAAAC	1718
TYMV	ACAGTCTCCAGTG-----GGAGAGGACGCCATTGTTCGATTCGGA-GGACGACAGC	1810
	*	
MRFV	AGCGCGAGTCCCTCATGCG-----GACGGGTTCGGTCTCTTTGCCCCAGTCGCTCCCTC	1753
CSDaV	GCATCGGACTCCCGCTGCGCCCCGCAACCAGCTTCATCAACTGGTCCGGATCAGCCCGCC	2172
OBDV	ATCTCATGTGGCCGCT-----CCAGCTTCTCCCCCTTCCCCACCCACCTTCGC	1795
GFkV	TCAGCTGCTTCCCAACC-----CCCTTCTGCCCCACCCAGCTCCCACCCAGA	2328
PnMV	TCCTCTGGTGCCGGCAGCA-----CGGCGTCGTTACGGCCACCGACATCACCATA	1769
TYMV	ATTTCTTCTTTCACTCCCA-----CGACTTCAACAGCCCCCTCCGGACCACTCCGAA	1862
	*	
MRFV	CT---CCTGAAGGCTCTCTT-----CTCCCA-----GTTGAGCCT	1785
CSDaV	CAGGACATGAGTGTGCCCCCAAGCTCCCACAAAATCCAAAGTTACTTTTGCTCAACCC	2232
OBDV	C---CACTCCCGCTCCCGA-----CTCCCGA-----GCTGAACCC	1827
GFkV	CGAG-CGCGTCGGCTCCCCCCC---CTCCCGC-----CCGCCCCACC	2367
PnMV	ACAATCATGGCCTCCGTGTTTTCCCTGCCACA-----CCTACACCG	1811
TYMV	GC---CAGTCTCCCTCCCGCTTTCGCCTCCACC-----TCCGTTCC	1901
	* * * ** **	
MRFV	CCT-----CCTGT-----CCCAGCGATCCCG--A	1808
CSDaV	C'TTGAGAGCGTGACACCTGTAGTTCAGGAGCAGGAGAACCTCCGAGTCGGCTTCATCA	2292
OBDV	CTCCAGC-----CACCTCCG-----CTCCACCTCGACC-CACGA	1862
GFkV	CTTCCACC-----CCAGCCG-----CCACCGCCCCCCCCCCCCA	2402
PnMV	CTGAAGGG---ATGTCTATCA-----TCTGCACCGATTACTTGC	1847
TYMV	CGTCCACCT-----CCAGTGGCATCGAGCCTTGAGGCCAGCCTCTACGACTACGGCA	1955
	* * * * *	
MRFV	ACCAGCCC-----TCGAGCCGTGCCCC-CGAGCGGCTTCGGTCCCTGCCCCAG--	1855
CSDaV	ACCGGCCCGGCATCGGTCTCCCGTCTGACCCGCAAAACGGCTTCATCAACCGTTCGGAT	2352
OBDV	GCCGGCTC-----CCGCCGATTCGAGCCCCAAGCTCCTCCGGCCACGCCCC---	1910
GFkV	ACCATCCC-----CACACCCCCCTCTTCCCGCTTCAATCCCGTCGCC---	2447
PnMV	CTGATTTTCAG-----TGGAACAGAGAACCCTGGCCAACACCAAGCCGCTCATCTGA	1899
TYMV	GCGCCCCCGA-----CACCGATTGAACCCACCCAGCGGCTCATAAAATTCTGAC--	2006
	* * * *	
MRFV	---CCCCAGCTCTAGCTTTCAGAGC---CTCCTCCT-----TCTCCTGAGTCGGTCGCCCC	1904
CSDaV	GCTCCCACCCTGGACAGCAGTGTGACCCCTCCACCGATTATCTACCAGTTGACCATCTC	2412
OBDV	---CCAGACCGAGCCTCCGAGT---CCCGTGATCGAGCAAGAAGCGCTCCGAATCCCC	1963
GFkV	---TCCGCCCGCCCGTCTCACCGCCCCCC-----CCGCCACCTCCCCCGCTC	2495
PnMV	TGACCAAGAGGGCGAGGCGGAGC---CGCATCC-----TTGCCCTCTCGCGCCC	1948
TYMV	---CTCACGCTTGAAGTTCAAC---CCCAATTGAACCCCCCCCCACCCCATCCAATCT	2061
	* * *	
MRFV	T-----TCCCGTCGCCGTCTGCGGCCCGCCGTG-----CAGCT-GCGCGGGCT	1947
CSDaV	CAGAATGACTTCGGCCCTGCCGTTGTTCACTCTGCGAACCGCCCCAGCCAGCCCCAGTT	2472
OBDV	T-----TCCCGTCTGCCCCGTTTCTGCTCCA-----CCCCCTCCGCTTCCGCG	2010
GFkV	C-----ACCCCGCCCTCACTCCATCCCGGCGCGGAA--GACCGCCCCACCCCTTACC	2547
PnMV	T-----CTCAGAACCAAGGTGCGCCTGCTCAG-----CCCGCGGAGCCTTGAAA	1994
TYMV	CCGAC--ATCCCGCCTTCCGCCCCGTTCTTTTCCAGAAATCAACTCACCGCATCGTTT	2119
	* * *	
MRFV	CCGTCCCTTCTCC--AGCGCTTCTGGGCGCC-----GACCTTCGC	1986
CSDaV	CCCTCTGCCCCCTCC--ACTGTCTCTGACCACAAGGAATCCCAAGATGGTGAGGCCCTCGC	2531
OBDV	CCTTCACTTGCCCC--AACACCTCGGCCCC-----GAGCCTC-C	2048
GFkV	TTCCCTCCCCACC--CTCGTCTGAGCCC-----GACGCCCC	2586
PnMV	GTCTTTCAGCCCGC---CACCGCCAA---G-----AAGCCGA	2026
TYMV	TTCCCCAAACTTCCACACACCCGATTTCAACCC-----ACCCGCAC	2164
	* * * *	
MRFV	TTCGGGACCT---TCTCCTGTCTCTGCTTGGGACAGTGACCCCGAGATCTCGAAGCTT	2043
CSDaV	TTCCGCCCTCCAAGTCTCGGGCTCGCCCCGTCCACCAGCTCCTCTGTCTCAGAATCT	2591
OBDV	CTCGCCGACCG---CTTCCGAGCAGGCGCGTCCCTCATCCCTGCTCCCTCTTCCGCCCT	2105
GFkV	GTCACCGCCG---CCCCAGTCTTCCCTCGCCCCAGCCGCCCTTCTCCGAGCTC	2643
PnMV	TCCAGCACCC-----CAAAAACCCGAC---GTATTAACACCTACTCTGTGATGGGGCC	2078
TYMV	TTACCCCTCC--TTCACTTCGCATCAAGATTCGACTGACCCCGCGGACCCCTGATG	2222
	* * *	

MRFV	GGT---GAGTCCACCCAGGGCA-----CCGTCTTTGCGGTCACCCCGGCCCTCGCG	2092
CSDaV	AACCATCGAGTCCTCAGGAGCCATGCATGCCTCTTGTGGGATCAGCTCTCTCCCAT-	2650
OBDV	CGTCGTGGAGCCATCCGGCGTCGTCTCTGCCTCATCTTGGGGCGCCACCAACCAGCCGGC	2165
GFkV	TACCCCGGCCACTACGCCGATCATTAGGCTCCTTCTTCCCTCCAGCAGCCCTCTCGCGC	2703
PnMV	GACCA-AGTGTACCCGTCGATGA-----GCATGGTGACCTCTCCAATCTCCACGAGG	2132
TYMV	GGCTC---CCACCTTCTGCACCATT----CACTACCTGCACCTCCACCCACCCGCTTCA	2275
MRFV	--CTCCTGAGCCAGACACCGCCCGCTGGATGCTGATCCTTCAGCCAGCGGTCCAGTCA	2150
CSDaV	--CATCGGACTGGGATCCCTCCCCCGTGGCTCGCGACAGCTCTGCCACCGGTCCACCAGG	2708
OBDV	--CGATCAAGTC-GATGACTCCCTCTCGCTCGCGATCCCAGCGCTCCGGCCCCGTCCG	2222
GFkV	T-CCTCCGTCCCTACCCCGCCACTCGACTGCCTTCTGGTCTCTGTCCGCTGCAGCGG	2762
PnMV	--AGCTGTTCTCGAAATCAGCC----AGCTCATGGTTCGATCCACCGCCTATG-GCGGT	2185
TYMV	ATCTTCACAGCTCTTGCCCGCACCTTTGACAAACGACCCACCGGATCGGCCCGGTACT	2335
	* * *	
MRFV	GGAGTTTTCGTGAGCTTCAGAAGGGTGCTTACATCGAGCCGACTGGCGCTTCCCTCACTCG	2210
CSDaV	CATGTACTGATCTCTTCCCGGCTCCTTATCTCCAGGCACCGGCCAATTCATCTTCCG	2768
OBDV	CTTCTATCGAGACTCTTCCCGCCAACTACGCGGGTATTCCGGCACTTCGACTTCCG	2282
GFkV	AATCCCAAAGGAAGACTCTGGGCCAC---CCTTTGCCACATCTTTCCCCCAGTGATCT	2819
PnMV	CAGACTTCGCACACTTACCCCGGAACTATATAGGAACCTCCGCTGACTTCCCTCGCTCG	2245
TYMV	CCCCTTGAAGAACTCCACCCACGAGGTACCCCGAAAACACCGCCACTTTTCCCTACGAG	2395
	*	
MRFV	AGCCCGCAACTCGGTGTCTCCAGCATCCCTTACCCTACGAGGGCTGCCTGTCTCCTCGT	2270
CSDaV	CTCCAGAGCCAACGGCCGGCCAAACATTCCCTACCCCGACATGGAT---TGTCTCTACT	2825
OBDV	CGCCCGCGCCTCAGGCGCTCTCCACCCCATACCCCGCCATGGAT---TGCCTCCTCGT	2339
GFkV	CGTCTCCGACCTTGGCCTCTCCACCAACCATCTCACCGCCCTCGCC----TTCACCTACC	2875
PnMV	CCTCCGCAACGGCCCCCAAGCCGTGCCAATCCAGTGAACCAGAGTTGCCTCTTAGT	2305
TYMV	GCTCCGTTCACTTCTTCAAACCATCTACACAACCCACCCTGAAT---TGTCTCTCTC	2452
	* * * * *	
MRFV	GGCCGTCAAGCGACAGGGCTTCCACCCGCACTCTCTGGGCAGCCCTCTGCGCCAA	2330
CSDaV	TTCCATCGAGCAGGTAACCCGCTTCCCAAGGAGGCCCTCTGGGACACTCTCTGTGCCAC	2885
OBDV	CGCCACCGAGCAAGCCACCCGCATCTCTCGAGAGGCCTCTGGGACTGCCTCACAGCCAC	2399
GFkV	AGTGGCTGGTCA--CTCTCCGCTCAGGAGAGTTGGTCCAGCGCCACGGCCTCCTCTCCGC	2933
PnMV	GGCAATCAGCAAAGCCACCTCCATCTCGCTCGAAGAGTTGTGGAAGACCCTCGTGCCCT	2365
TYMV	TGCTGTCTCCGACCAAACCAAGGTTTCCGAGGATCACCTCTGGGAGTCCCTCAGACAAT	2512
	* * * * *	
MRFV	CCTTCCGGATTCCGTCTTGTAGCAGGGT-TCACTCGCCACCCTGGGCCTCACCACCGACC	2389
CSDaV	CTGCCCCGACTCTCTGCTTGACCTTGACACCATCCGCCGAGTCGG-CCTCTCCACAGACC	2944
OBDV	CTGCCCCGACTCATCTCTCGACCCCAAGAGCATCGCCAGCATGG-CCTCAGCACCAGATC	2458
GFkV	CCCTTTCGCCTTTGAGATCACCACACC-CCACCCGTTCCTCCCG---CCACCGTCGGGC	2989
PnMV	CCTGCCGATTCTCTCTGTTGCTTGAGGACATAACGACCCCGCG-CCTCTCCACGGACC	2424
TYMV	TCTCCAGACAGCCAACTCAGGAACGAAGAGATCAACTCTCTCGG-GCTTTCAACTGAAC	2571
	* * * * *	
MRFV	ACTTCGCCGTCTTGTCTCGCATCTTCTCCTTGAGATGTCGCTTCGTTCAG---CGAGCATG	2446
CSDaV	ATTTTGTATCTTTCGCCCACCACTACTCCCTCAGGTGCGGTTCCATAC---CGCCCATG	3001
OBDV	ACTTCGTCATCCTCGCTCATCGCTTTTCCCTATGTGCCAACTTCCACTC---CGCGAGC	2515
GFkV	ACTTCGCCCTCTCCGCCCCCTCACCCCCACCTCCGCTTGCCTCACCGGGGGCGCCCTT	3049
PnMV	ACTTCGTAGTCTTGGCTTCCGCCACTCTTTGAAGTGCATTTTCATCTC---CGGGGCGC	2481
TYMV	ACCTCACTGCGTTGGCCCATCTTTACAACCTCCAGGCAACCATCTACTC---CGATCGTG	2628
	* * * * *	
MRFV	GCGACGTTGAGCTCGGGCTTCATGACGCCACCTCTCGCTTACGATCCGCCATACCCCTG	2506
CSDaV	GTGTCATTGAGCTGGGCATGGCCGACGCCACTTCACTTTTGATATCGACCACACTGCCG	3061
OBDV	ACGTCATTAGCTCGGGATGGCCGATGCCACCTCCATTTTCATGATCAACCACACGGCTG	2575
GFkV	CCCCAGTCATCTCCGGCCCCAAGGCTCCGCTCCCTCCCCCGCGCCCGCTTTGGGCCCTC	3109
PnMV	TCGAGCTCGAGTTAGCCTCCACAACGCACACCCTCCTTCAATCAGGCACACAAGCT	2541
TYMV	GTCCCATCCTCTTTCGGCCCATCCGACACCATTAAGAGAATCGACATCACCCACACCACCG	2688
	* * * *	
MRFV	GC-----CACTTCGAGCTCGTGGCTGACAACCTTTTCTCTTCCCGC-TCT---CGTC	2553
CSDaV	GC-----AATCCTGGCCACTTCTCCCTCCGGCAGTCTGCCACGCC-CCG---GCTC	3108
OBDV	GCTCCGCGGGCCTCCCGGGCACTTCTCCCTCCGCTGGGTGACCAGCC-CCGTGCCCTC	2634
GFkV	GCCCTGAAG--CTCCCCGGCCTCCCTGACCCCCCTGGCTTCACTCCCATCACCGAGCC	3167

MRFV	GTCATCGATGAGGTGTACAAGATGCCTCGCGGTATCTTGATCTGGCCATCCATGCGGAC	3078
CSDaV	GTCATTGATGAAATCTACAAGATGCCCCGGGGCTACATTGATCTCGCCATCCACTCTGAT	3624
OBDV	GTGATCGATGAGATCTACAAGTTGCCCGGAGGGTACCTCGACCTAGCCATCCACTCCGAC	3150
GFkV	GTCATCGACGAGGTCTACAAGTCCCTCGCGGGTACCTCGACCTTGCCCTTCTCGCCGAC	3684
PnMV	GTTATTGATGAGGTGTACAAAATGCCCGCGGTTACGTGGATTGGCGGTGGCAGCTGAT	3116
TYMV	GTCATCGATGAAATCTACAAAATGCCAAGAGGGTACCTCGACCTTTCATCTCGTGAC	3272
	** ** *	
MRFV	CGCGCCATTAGTTTCGTGATCCTGCTCGGCGATCCGATTCAAGGCGAGTATCACTCCACA	3138
CSDaV	CCAACCATTGAGATGGTCATTGCTCTCGGTGACCCACTCCAAGGCGAGTACCCTCCACT	3684
OBDV	TCGTCCATCGAGTTTGTATCGCCCTGGGAGATCCTCTGCAAGGCGAGTATCACTCCACT	3210
GFkV	CCCACCGTCGAGTTCGTATCATCCTCGGCGACCCCTCCAAGGATCCTACAACCACCACC	3744
PnMV	CCTACCGTCCGTTTGTCTATCTTTTGGGTGACCCCTTCAAGGCGAGTATCACAGCTCC	3176
TYMV	CCCGCCCTCGAATCGTCATAATTTCTCGGTGATCCTCTCCAGGCGGAGTACCCTCTCAA	3332
	* *	
MRFV	CATCCCAGCTCCTCCAACGCCCGCCTTTTCTCCCGAGCACCGGTACCTCCGCCCGTATGTG	3198
CSDaV	CATCCTTCTTCCACCAACTCCCGCCTTCTTGCTGAACCACACCCTCTCCATGTATCTC	3744
OBDV	CATCCCAGCTCCTCCAACCTCTCGCTCATTCCCGAAGTCAGCCATCTCGCTCCCTACCTC	3270
GFkV	AACCCCGACTCTTCCAACCACCGCCTCATCCCCGAAGAAGACCCTCCGCCCGTTCATC	3804
PnMV	CATCCGAATCCAGAACAAGCGGCTCTCCTCGGAGATCCACAGGTGCTCAAGTACATC	3236
TYMV	TCCAAAGACTCATCCAATCACCGCCTTCCCTCCGAAACTCTCAGGTGCTACCATACATT	3392
	* *	
MRFV	GACTTCTACTGCTTCTGGAGCCGCCATCCCTCAGAATGTGGCCCGCTCCTCGACGTG	3258
CSDaV	GACTACTACTGTATGTGGTCCCATAGAGTCCCTCAGAATGTGCTGCTTTCTTCCACATC	3804
OBDV	GACTACTACTGCCTCTGGAGTTACCGGTCCCCCAAGACGTGCGCGCTTTCTTCCAGGTT	3330
GFkV	GACTTCTATTGCTCTGGACTCGCCGCTCCCGCCTCGTTCGCGGACTTCTTCCGGCGTA	3864
PnMV	GACTGCTATTGCCTCTGGTCTCACCGAATCCCTCAATGCGTTCGCGGCTGTTTCCGGGTT	3296
TYMV	GACATGTACTGCTGGTGGAGTTATCGCATTCCCAATGTATCGCCCGACTCTTCCAAATT	3452
	*** ** *	
MRFV	CCCACGACCTCGACTGAGATGGGTT---TTGCT--CGTACTCGCAGCAGTTCCCGTTC	3313
CSDaV	AAGACCACCTCCAAGCAGCCAGGCT---TCTGC--CGTACCAGCGAGAGTTCCCAACT	3859
OBDV	CAGAGCCACAACCCTGCTCTCGGGT---TTGCC--CGTCTCTCGAAGCAGTTTCCACGA	3385
GFkV	CCCACCACCAACCCACTCGCGGTACCTCGCCTTTCGCTCCCTCAACACCCTCAATCC	3924
PnMV	GAGAGCTCAAACGAGCACGAAGGTTACGTGGGCTTCAGCTCCTTTATCGCACCCAGGAC	3356
TYMV	CACAGCTTCAATGCCTGGCAGGGAATCATCGGCTCCGTTTCAACTCCCAGGATCAATCC	3512
	* *	
MRFV	CTGGAAGATCCTCATCTCAGCGCGGACTCCGCCAAGAGCCTGGCTGATTGTGGCTACC	3373
CSDaV	CCAGGA---TCCTGGCCAACTCTCAAAATGCAGGCCATACTCTTACGAATGCGGCTATG	3916
OBDV	CCGGGCGGTCCCTCACCAACTCACAGAACTCGATGCTTACCATGACGAGTGCAGGCTACT	3445
GFkV	CTC-----TCCTCGTCCCTCCGACTCCATGGCCGAGCCCTACTGTGGCGGCCACC	3979
PnMV	ATCA-----CCCTCGCTTGCTCGCAGAGCTCAGCCAAGACCTTGCAGCAGCCGGATACC	3411
TYMV	CCCG-----TTCTCACCAACAGTCATGCCTCATCTCTCACCTTCAACAGCCTGGGATATC	3567
	* *	
MRFV	ATGCTGTCACTATTGCCAGCAGCCAGGGTCCACCATTGCTGGTCCGGCCTATGTCCACC	3433
CSDaV	CCGCCGTACCATTTGCCTCAAGTCAGGGTCCACTTATGACAATGCGGCTGCATCCATC	3976
OBDV	CTGCCGTACCATTTGCCTCAAGCAGGGTCCACTTACAGCGGCGCCACGCACATCCACC	3505
GFkV	GCGCCATCACCTACGCCGCTCCCAAGGTTCCACTTACCCCGCCCCGTCACATCTTCC	4039
PnMV	CTGCCACCACCGTCAGCAGCTCTCAAGGTTCCACTTACCGACCCCAATCAATCTTCTCC	3471
TYMV	GCTCCTGCACGATCAGCTCTAGCCAAGGCTCACATTTCTGCGACCTGCCATCATCGTCC	3627
	* *	
MRFV	TCGATAACCATTCGAGGCGTCTCTCACACCAGCATTCTCTCGTCCCATCACTCGATCCA	3493
CSDaV	TTGACCGCAACAGCTCCCTGCTCTCCCTGCTCATTCCATGGTCTGCTCACCCTGCCA	4036
OBDV	TTGACCGCAACTCATCGCTCCTCTCCCTTTCGAATCCCTCGTCCGCTCACTCGCTCGA	3565
GFkV	TTGACCGCAATTTAACTCGTCCCAACCATGTCGCCCTCGTCCGCTTACCCGTTCCC	4099
PnMV	TCGACAGAAACAGTCGACAGCTTAGCAGCGCGTCTCACTGGTCCGCGCCACTCGTTCTA	3531
TYMV	TGACAACCTACCCAAGTGGCTCTCCTCGGCCAACGGCTCGTCCGCTCACCCTGATCCA	3687
	* *	
MRFV	AATCCGGCATAGTCTTCA---CCGGCGACAAGGCTGCAGCGGACGGCACCTCCTCCGCCA	3550
CSDaV	AGGTTGGTGTCTCTTCA---CCGGGGATCCTGCCAGCTGTCCAATGCCCGAGCTCCA	4093
OBDV	GAACCGGCGTGTCTTCT---CCGGGACCTGCTCTTCTCAACGGTGGTCCCAACTCCA	3622
GFkV	GCTCTGGCGTCCACTTCCGCCCCCGGCCAGGACCTCCCTCGCCACCCCGAGCACCTCT	4159

PnMV	CCATGGCCATAAACATGA---CGGGCGATCGGGATGTTCTGCAACAGGGCCACACAACA	3588
TYMV	GATCAGGTGTCCAATTCA---TGGGCCCTCTTCCTATGTGCGGGGAACCAACGGCTCTT	3744
	* * * *	*
MRFV	ATTCCTCTTCT-----CCGCTGTGCTC-----CTTGATCGGGCCTCT-----	3589
CSDaV	ACAGAATGTTCT-----CTGAATTCCTC-----TCAGGCCGCACCCGCC-----	4132
OBDV	ACCTCATGTTCT-----CTGCCTTCTTT-----CAGGGCAAAGTCTCGCC-----	3661
GFkV	TCACCGGTTCTACAAGTACGCGATCGACCTCCTCGCCACGAAGCCGATCCCTCCAAGC	4219
PnMV	ACCCGGTGTCT-----CTGCGCTTTAT-----AACGCCAATCTGATA-----	3626
TYMV	CTGCCATGTTTT-----CTGACGCCTTC-----AACACAGCCTCATCA-----	3783
	** *	*
MRFV	--CGGTGCGCTCCCTT-----TTCAGCGCGCTCCTGCCCTGCTGCCCGT	3631
CSDaV	--CTCTTCATGACTGG-----TTCCACAATGAGTTCCCAAAGGCCACTG	4174
OBDV	--ACATTCGCGCCTGG-----TTCCACCCTTTTCCCTACGGCCACTC	3703
GFkV	CCCGCCCCACCCCGTCGACGTACGCTGCTCTTCCAACAGCAGCTCCGCGGCCTCACCA	4279
PnMV	-TCCCTGCACAGCCTG-----TTCCCAACGCTTTTCGCAAAGTTGCCGA	3669
TYMV	--TCATGGATCGCTAC-----TTCCCATCCCTGTTCCCACAACTCAAGC	3825
	*	***
MRFV	TCGTCACGGAGCCTCCA-----ACGTCTC---GAGCCGTCTCT---TGCGCGGAGCTG	3679
CSDaV	TCCTCACCGAGCCCTC-----AAGACTC---GGGGCCCCGCC---TCACCGGTGCTG	4222
OBDV	TCCTCTTCTCCCCCTCCGCCAACGCCAC---AACCGCTCACTGGCGCCTCGCTCCCG	3760
GFkV	TCCTCCGTGACCCCTCCTTCTCCCGCATCACCGGTGGCGCCACCACGCCTTCTCGCCC	4339
PnMV	TCAT-----ACGCACC-----GCGATCACCA--GTCGTAGGGGCAT	3703
TYMV	TCATCACCTCCCCCTC-----ACAATC-----GCAGCCCCAAACTCAACGGGGCCAC	3874
	** *	* * *
MRFV	GCT-ACGGC-----ATTGCTCGGCCTCT---TCG-----CGCTC	3709
CSDaV	CCTCACCAT-----ACTCCAAGGTGTT---CCATC-----CGGCA	4255
OBDV	CCCAACCTTCCACCTCCTGCTCCTGACCTTC---CGAGCTCCCTCCTCTCCCGCCT	3817
GFkV	ACGCCCCACTTTCACCAACTCCACGGCTCCGCCACCACCTTCCCGGACAACCTTC	4399
PnMV	GCTGACCG-----GTGGCAACACCTCT-----	3725
TYMV	CCCCAGCGC-----ATCTCCACCCATCG-----CTCGC	3903
	*	*
MRFV	GCGA---C-GCCCCCTC-----TCGGCCCCGATTACGTTGGAGACGT-----	3749
CSDaV	AGGATTC-ACCCCGGCC-----TCAAACCTGATTTCCAGGGGACGTT-----	4299
OBDV	CCGGTCCCTACTCCCGCTCATTCCAGTTTCGATCTCGCTTCGCCGCGGCCGTCAGCCTT	3877
GFkV	CCACCGCCCCACCTACCTGGCCCGTTCCCTCCCTACCACCAGACCGA-GTCTACCCC	4458
PnMV	-----GGTGAGGACG-----	3735
TYMV	CAAACCTTCCACCTCCCCCA-----CACATTCCCTCTCTTATGATCGTG-----	3948
	*	
MRFV	-----CATTCT-TGACTCCTCCGCCCGATCCTTGGTGACGGATCCGCCAAC	3795
CSDaV	-----CGTTCT-TCTCGGGCCCCATCG-TTCTGGGCTCCGGCGAGCTCAAT	4344
OBDV	CCGACCGGTACAGCTCCTCTCGTGGGCCCTATCG-CCGTCCGTGACGGGAAACCAAC	3936
GFkV	ACCTCAGCCCTCCCGCACCGTGTCTTCCCGCCTCCACCACCGACTGGTCAGCCCGGAC	4518
PnMV	-----ACGTTATCATGATTCAAGACG--TCCGGGC-----CGCTGAAT	3773
TYMV	-----ATTCGTACCG--TCAACCCAACTCTCCCTGATCAGGGACCCGAA	3992
	* * *	*
MRFV	GC---TCCCCAGGTCAGCACTCACTTTCTTCCGGAGACGCGTCCGGCTCTGCATTTTCGAT	3852
CSDaV	GC---TCCACAGGTCTCAACTCACTTCTTCCCTGAGACTCGCCGCCCTTCTCACTGGGAC	4401
OBDV	GC---CCCTCGCATGACACCTCCTTCCCTGCCGAAACTCGCCGCCGCTTCATTTTGAT	3993
GFkV	GACCATCCCCGAGTCAACCCACCTTCGTGCGCGAGACCCGCCTCCCCCTCCAGAGTGAA	4578
PnMV	GC---TCCAGAGGTGACGATACTTCTTCCCGACTTGGCGACGCCCTTTCACCCGGAAT	3830
TYMV	AC-----AAGACTCGACACCCACTTCTCCACCTTCTCGGCTCCGCTTCATTTTCGAT	4046
	* ** * * ** * ** * ** *	
MRFV	ATTCCTTCCGCCGTCATCAAGTCGCGGATCACCCCTCGCGCCGATCACTCCGCCTGC	3912
CSDaV	ATCCCTTCTGCCATCCCTGAGAGTGCCACCAGGCCTGATTCCACTGAGCCACCCTTGC	4461
OBDV	CTTCCCTCGTTCCGCCCCCCAAGCCCCACCGCTCCCTCTGACCCAGCCCTTCTGGGACC	4053
GFkV	CTCGCCCCACCTTCCC---ATCCAACCAAGAACCTTCCCCGACCTACCCTCCCGCCG	4635
PnMV	CTTGCTCTGCTGTTTAC--AGCAATGCGCCGCCCAAT-CCACTACGTCACTCCGCGC	3887
TYMV	CTCCCACAGCTATCAC---CCCCCCCGATTTCCACAAGCGTCGACCCGCCACAAGCT	4103
	* * *	* *

MRFV	GCCATTGAGCCCGTATATCCCGGTGAGAGCTTCGAATCTCTCGCCTCTCTCTTTCTCCCC	3972
CSDaV	CATCCTGAGCCGGTTTACCAGGGGAAACCTTTGAAAATCTTGCTGCCACTTTCTCCCCG	4521
OBDV	GCCTTTGAGCCCGTTTACCCTGGCGGAAACCTTCGAAAATTTGGTCGCCCACTTCTCTCCG	4113
GFkV	ACCTTCGAGACCGTCTACCTGGAGTGGACGGCGAAGCCCTCGCCCGCACCTTCCCTTGCC	4695
PnMV	GCTATCACTGCGGTTTATCCTGGTGGAGATCTTTGAAAATCTCGCAGCACATTTCTCCCC	3947
TYMV	AAAGCTAGCCCCGTCTATCCAGGCAGTTCTTCGATCTCTGGCGGCTTCTTCTTACCA	4163
	* * * * * * * * * * * * * * * * * * * *	
MRFV	CCGACGGATGCCGAGAGCAAGGAACTTACTTCCGAGGGGAGATGTGAATCAATTCCC-	4031
CSDaV	GCTCACGATCCCCTGATCGAGAAATCTACTGGCAGGGCCAGCTGTCCAACCAGTTCCC-	4580
OBDV	GCTCACGACCCCTGACCCGCGAAATCCACTGGCGTCCGAGCTTTCCAACCAGTTTCC-	4172
GFkV	GCCACCGACCCGTTGGAGTGGAGATCTTCTCCGCAACAACCTGGTCCAACCAGTTCCC-	4754
PnMV	GCTCACGACCCCGAAATTAAGGAGATCTTGACCCGCGGAGTTTACGCGCGAGTTCCCCG	4007
TYMV	GCACACGACCCATCAACAAGGGAAGTACTCCACAAAGATCAATCTAGCAACCAGTTCCC-	4222
	* * * * * * * * * * * * * * * * * * * *	
MRFV	--CCATCTCGATAAGCCGTTCCAGTTGGGTGCGCAGACGTCGAGCCTCCTTGCGCCGCTT	4089
CSDaV	--TCACATGGACAAAGAGTTCCACCTGGCCGCTCAGCCCATGAGCCTCCTGGCTGCGATT	4638
OBDV	--CCATGTCGATAAGGAGTACCACCTCGCGGCTCAGCCAATGACGTCCTCGCTCCCATC	4230
GFkV	--CTTCATCAACCGCCCGACACCGTAGCCTGCAACCCTCTCACCCCTCGTCCGCCACC	4812
PnMV	ATTCTCCTGAATCAGGATTTCTCTCTCTCCGCGCAGTCTTCCACGTTGATTGCGGCTAAG	4067
TYMV	--TTGGTTCGACCGACCCCTTCCAGTTGTCTCTGCCAGCCCTCAAGTTTAAATTTCTGCCAAG	4280
	* * * * * * * * * * * * * * * * * * * *	
MRFV	CACAACTCCAAGCAGACCCCTACGCTCCTCCCGCCTCCATCGGGAAGAGATTGCGCTTC	4149
CSDaV	CATCAAGAGAAACAGGACCCCACTCTGCTCCAGCTTCCATCCAGAAGAGGCTCCGCTTC	4698
OBDV	CACGACTCCAAGCAGACCCCACTCTGCTCCAGCTTCCATCCAGAAGAGGCTCCGCTTC	4290
GFkV	CACAACCAAAAGCAAGACCCCACTCTCCACGCTTCCCTCGCCAAGCGTCTCCGCTTC	4872
PnMV	CATGACTCCAAGCGCAGCCGACGCTGCTCGTCCGCTCCATTACCAAGCGCCTCTGTTC	4127
TYMV	CATGACCCCAACCACGATCCGACCCCTTCTGCTGCTCCATCAATAAACGCTTGGGATTC	4340
	** *	
MRFV	CGCCATTCCGAGGCACCCTACGTCATAGCCCTCGCGACGAAATTTCTGGGTCACCTCTC	4209
CSDaV	CGCCATCAGACAGGCTTACCAGATCACTCCAAAAGATGAAATCCTGGGCCAATTGCTC	4758
OBDV	CGACCTCCGCTCTCCCTACCAGATCTCCCTCGTGACGAGCTGTGGCCAGCTCCTC	4350
GFkV	CGGACTCCACAGCCCCCTACCCATCACCGCCAAGGATCAAGCCTTAGGCTACATCCTT	4932
PnMV	CGCCGAGCGATTCTTCGTACCAGCTCTCTCCAAGGACCAATTTCTGGGCTTGACCCTC	4187
TYMV	AGACCCAGTGAAGCACCCCAAAATCACCGCAGCAGCGTGGTCTTAGCCTGCAACTC	4400
	* * * * * * * * * * * * * * * * * * * *	
MRFV	TACGAGGCCTGGTGTGAGCCTATCACCGTTCCCTCGGACGTTGAGCCGTTTGACCCA	4269
CSDaV	TTTGAAGGCCTCTGCGGAGCCTATCATCGGTCCCTTTCCACACTGAGGCCTTTGATCCT	4818
OBDV	TACGAGAGTCTCTGCGCGCGTATCATCGTTCCCAACCACCACCCACTTTTCGATGAG	4410
GFkV	TACCACTCCCTCAGCGCGCTACTGCGGTTCCCGGAGCCGTC---CCCTTCGACCCA	4989
PnMV	TACCACAGTTGGTGGCGGCTTACACCCGACCCGACCCGAGCAGTCTTCCATTTCGATGAG	4247
TYMV	TTCCACTCTCTGCGCGCCTACTCACGTCAACCCAAACATCACCGTTCCATTCAACCCCT	4460
	* * * * * * * * * * * * * * * * * * * *	
MRFV	GACCTTTATGCCGAGTGCATCAACCTCAACGAGTTTGGCGAGCTCTCTTCCAAGACCCAG	4329
CSDaV	GTTCGTTCGCCGAGTGCATCAATCTGAATGAATTCGCCAGCTCTCTTCCAAGACTCAG	4878
OBDV	GCCCTCTTCGTGAGTGTATCGACCTGAACGAATTCGCTCAACTCACCGCAAAACTCAG	4470
GFkV	GTCTCTTCGCGTCTGTCATCGCGGAGAAGACTTCGCCAGCTCACTCCAACCAACCA	5049
PnMV	TTGCTGTTTGGGAGTGCATCAACATTAATGAGTTTCGCGCAGTCTCGAACAAAGACCAAG	4307
TYMV	GAACTTTTTCGAGAATGTATCTCTCTGAATGAATACGCGCAGCTCAGTTCCAACCAAC	4520
	** *	
MRFV	GCAACCATCATGGCTAACGCCAACCGCAGCAGCCAGACTGGCGCTGGTCCGCCGTTCCG	4389
CSDaV	GCCACCATCATGGGCAATGCTCGCGCTCTGATCCAGACTGGCGGTGGAGTGTCTCCGC	4938
OBDV	GCCGTCATCATGGGCAACGCCCGCCGCTCTGACCCAGACTGGCGCTGGTCCGCCGTTCCG	4530
GFkV	GCCACCATCCAAGCCAACGCTTCCGCTCCGACCCGATTGGCGTCACACCTTCGTCGCG	5109
PnMV	AGCACGATCATAGCGAACGCCAGCCGATCGGACCTGATTGGCGCATCACAGCCGTCAG	4367
TYMV	TCCACCATAGTGGCAACGCTTCCGCTCCGACCCAGACTGGCGCACACACCACCGTCAAG	4580
	** *	
MRFV	ATCTTCGCCAAGACGCAGCATAAGGTCAACGAGGGCTCGCTTTTTCGGCTCCTGGAAAGCC	4449
CSDaV	ATCTTCTCCAAGACCCAGCACAAAGTAAATGAAGGGTCCATCTTTGGTTCTTGGAAAGCA	4998
OBDV	ATCTTCAGCAAAACCCAGCACAAAGTCAACGAAGTTTCGATCTTTGGAGCCTGGAAAGCT	4590
GFkV	ATCTTCTCGAAGACCCAGCACAAAGTCAACGAGAACTCTCTTTTCACTCTTGGAAAGCC	5169

MRFV	GACCAGTGGTCTGCCATTGCTCCCATGCTTGCTCTCACCTTCAAGAAGGAGCGTGGCCGT	4986
CSDaV	CCTGAATGGGTACATCCTCCAGCCTCTTCTCAGTCTCCGCTTCAAGAAGAAAGGGGTCCG	5535
OBDV	CCAGAGTGGGTGCGCCATCGAACCTCTTTTAGCCCTCCGCTTCAAGAAGGAGCGCGGTCTG	5127
GFkV	CCTTCTTGGCCTGCCCTCGCCCCCTCTTTCGCCCTCAAGCCCAAACCCGAGACCTCCCC	5706
PnMV	CGGAAGTGGCCTGCCATCGAAAAGCTGCTTGCTCTTCGCTTCAAACAGAAAGCACCAA	4964
TYMV	CACGACTGGCCCTCTGTTCTCAAACGCCTCCACCTCCGCTTTAAACTTGAACCTCACTTCT	5180
	*** * * ** ** **	
MRFV	TATGCCACCTTTTGGCGTACTACGTCGGCTTTACCGGTGCAGTTCGATCGCCCCGGCC	5046
CSDaV	TACGCCACCTTCTGTGGTACTACGCCTCCCATGTTCGGCTGTGTCCGCTCCCCCGTGGCT	5595
OBDV	TATGCCACCTTCTGCGGCTACTACGCCTCGCAGTTCGGCTGCGTCCGATCTCCCATCGCC	5187
GFkV	TTCGGCCTTTTCTGCGGTTACTTCTCGGCCCGCGCGCCGCTCCGCGCCCCCGGCC	5766
PnMV	CATCCCCTGTTCTGCGGATACTATGTAGGACCAGCCGGTGCATCAGGTCTCCAATCACC	5024
TYMV	CATCCCCTCTTTTGTGGTACTACGTCGGTCCAGCAGGCTGCATCCGCAACCCCTTGGCC	5240
	* ** * * * * ** * * ** *	
MRFV	CTTTTCGGAAGCTCATGATAGCGGTGATGACGGATCCATCTCTGACAAGCTCATCGCC	5106
CSDaV	CTCTTTGCCAAGCTGGCCATAGCTGTGATGACGGTCCATCTCTGACAAAATGGCCTCA	5655
OBDV	CTCTTCGCTAAGCTCGCCATCGCCGTCGACGACTCATCCATCTCCGACAAGCTCGCCGCA	5247
GFkV	CTTTTCGCAAGCTCGCGATCGCCCTCGAGGACGGCTCTCTGCCGAGAAAATAGCGTCT	5826
PnMV	CTTTTCGGAAGATCATGAGTCCGTCGACGATCTCTGATAGACGACAAAATGGCAAGT	5084
TYMV	CTTTTCTGCAAGCTCATGATCGCAGTGGACGATGACGCCCTCGACGACCGACGACTCAGC	5300
	** ** *** * * ** * * **	
MRFV	TATCTACGGAGTTCAGTGTGGCCATTCTCCGGGATGCCTTCTGGACCATCCTCCCA	5166
CSDaV	TACCTCTCTGAATTTGCTCTTGGCCACTCCCTTGGAGACCATCTCTGGGAAGCTTTGCC	5715
OBDV	TACCTCATGGAGTTCGCGGTGCGTCACTCTCTCGGCGACTCTCTTTGGTCCGCCCTCCC	5307
GFkV	TACGTCGCTGAGTTCGCGTGGTCACTCCCTCGGCGACTCCCTTTGGTCCCTCATCCC	5886
PnMV	TACCTGTCTGAGTTCGCGATCGGGCACTCTCTCGGCACGACTTTTGGGATCTTCTGCCA	5144
TYMV	TACCTCACCGAGTTCACCACGGACACTCCTTGGCGAATCACTATGGCACCTCCTCCCT	5360
	** * ** * * ** * * ** * ** *	
MRFV	GTCGAGGCTGTTCCTTACCAGAGCGCTGCTTCGACTTCTTCTGCCGTCGAGCTCCGGCA	5226
CSDaV	CTCGAGGCCGTTCCCTTCCAATCTGCCTGCTTTGACTTCTTCTGCCCGGGCCCCCAGA	5775
OBDV	CTGTCCGCGTCCCTTTTCAAGTCAAGCTGTTTCGATTTCTTCTGCCCGCGCTCCCGC	5367
GFkV	CCTGAAGTGGTCACTACCAGTCAAGCTGCTTTGACTCATCTGCCGCCACGCCCTTCCC	5946
PnMV	ATTGATCAAGTCCCTTACCAATCCGCGGTGTATGATTTCTTCTGCCGGCATGCCCTCGCG	5204
TYMV	GAAACCCAGTCCAGTATCAGTTCGCTTTGACTTCTTCTGCGAGCTTCTTCTGCGAGCTT	5420
	** * ** * * ** * * ** * * *	
MRFV	CAGCCAAAGTGTGTTGCGGCTTGGCGAGGCACCCGA-----GTCCCTGCTCTCCCTG	5280
CSDaV	CACCTCAAGCTCTCTCTCATGCTCGGCGAGGTCCAG-----ATCCATCATTGCC--	5827
OBDV	GATCTAAAGCTCGCCCTTACCTGGGCGAAGTCCCTGA-----AACCATCATCCAAC--	5419
GFkV	CAGCTAAAGCTCGCCCTCCGTCGCGTGGTGGTCCCGACTGGGGTCTCTCTCTCCAG	6006
PnMV	GAAACAGAGCAGATCCTCAACATCGGCGAAGTGCCTGA-----CACAGTCTATCAA	5256
TYMV	CACGAGAAGATGCTCCTCGATGATTCACACCCCACTC-----AGCCTCCTCGAACGA	5474
	* ** * * *	
MRFV	-----GCATTTGAGGGTC-----TGAAGTGGGCCTCC-----CACTCG-----	5313
CSDaV	-----GCATCGGGTCAATC-----TTGAAGTGGGCCTCT-----CATGCC-----	5862
OBDV	-----GCTTCTC-CCACC-----TCTCCTGGCTATCC-----CAGGCC-----	5451
GFkV	CTCAAGCTCCGCTTCTCACCAGCCCTCTTTCGCCCTCCTCGACGCTCACACTCGCGTC	6066
PnMV	-----GCATAGCTAGTCACT-----CAGGTTTGGCAGCC-----GAGCCC-----	5292
TYMV	A-----TCACTTCTTACC-----GAGGTGGCTCACCAG-----AACGCC-----	5510
	* ** * * *	
MRFV	---GTGTATGCCCTC---ATGAACTCCAGTCATC-----GACGCCAGCTGCTCCACA	5359
CSDaV	---ATCTACACCACA---CTCTCCTCTGCCGCTC-----GAG--TGGCCATTCTGAG	5906
OBDV	---GTCTACAGCCTC---CTCCCATCTCGCCTTC-----GCC--TCGCCATCCTTCA	5495
GFkV	ATGGTCCGACCCACAAGGCTCACCTCCTCCCTCCGGCCACGCCCTCCACCCCTCCACC	6126
PnMV	---TGTATTGTCTC---TTGCCGTCCTCAGTTC-----GCC--GGGCTCTCACATC	5335
TYMV	---ATGTACCTCCTC-----CCGCGCAAGCTCA-----GAC--TGGCTATCACCTC	5551
	* ** * *	
MRFV	G--C--TCTCGCCGCCCCGCTCTCT-----TCCCAGGACCCT--GAAGTTTTCGAG	5406
CSDaV	ATCC--TCCCAGCAACAGCAGATCCAT-----GCCAGATGACCCC--GACACCCTCTG	5955
OBDV	CAGC--TCACGCCAGCACCGTTCCCT-----CCCCGAAGACCCA--GCCGTTTCTTCG	5544
GFkV	GAGCCATTCTACTAGCCTACTTGCCTGTTGTTGTTTCCAATGACACCAGCTCGCCCTCCT	6186

PnMV	CAAG---AAGACAGTCACCGACGCTT-----TTGAAGATCCT--GATGTCTCGCAC	5381
TYMV	TC----TGTCTCAAACGCAATCTTTC-----CCAGAATCCATTGAGGTTTCCAC	5597
	* * * * *	
MRFV	C-----TTCAGGGTGAATT--GCTTCATCAGTTCCAATCGCTTACCTCC	5449
CSDaV	C-----TACAAGGTGAATT--GCTTCAGCACTTTCAAGTAC-----	5989
OBDV	C-----TTCAGGGTGAATT--GCTTCAGACGTTCCATGCTCCAATGCCCT	5587
GFkV	CGCCATCCTCGTCTTCTCATCGTCTGTTCTAGCCTCAGCCCTCTTCTGCGCC-CTCTTCG	6245
PnMV	A-----TGCAGGGTGAATT--GCTTCAGAACTTCCAAGAAT-----	5415
TYMV	G-----CTGAGTCTGAATT--GCTTCACTATGTCCAATAGC-----	5631
	* * * * *	
MRFV	CGCTCCGTGGTGGCCACATGCCAAA-----TCCGCTCGCCGCGCCCTTCA	5494
CSDaV	CATTC-----ATGC--AA-----TCTG-----ACACTCTCCT	6014
OBDV	CTCTCCCTT-CACTCCCCTCTTCG-----GCGGTCTA--TCTCCCGACA	5629
GFkV	CCCTTACCGTGTGCGCCATCCACCAGACCGTCTCTCTCCCGTCTCTCGCTCCCCCT	6305
PnMV	-----	
TYMV	-----	
MRFV	GGCTCCTTCAGCAAAGCAGCTCTCTTGGTCC-CACTTACGCAGTCGCCCTATCGC----	5549
CSDaV	GCCTCTCACTG----GTGGTTCCTCTGCTCC-CATCTCACACCAGAAGCCTTC-----	6063
OBDV	ACATCTCACTCCCCACGAGTTCGCGACCCG-CCTCTACGAAAGCTCCGCCTACCC----	5684
GFkV	CCCCCGTCCCGCTCTAGCATCCCCCCTCCGCATCCCCAAGCCCTCTCCGCCTCGTCT	6365
PnMV	-----CTAACCTAACCA--AGCTACGAGGTGCAGAC-----	5444
TYMV	-----AATCA--GCCCCGACATGGAAAT-----	5652
	*	
MRFV	-----TCGTGCT-----CCCCAGGTTCTCCGC----CCTC	5576
CSDaV	-----TCCACCT-----CTCTCGCTTTCTCC-----	6084
OBDV	-----TACTCCT-----CCAACTCTCCGACCT----CCAT	5711
GFkV	GATGAGCCTCCCCGCTGACCTCTCTTGGCGCCATCTCTCCCTCTCCGCAATCCCC	6425
PnMV	-----	
TYMV	-----	
MRFV	TATGGCCGACAATGCCACCCAA---GTCGGGCTGTTCTCTCCTCGCGACGACCGGTTGA	5633
CSDaV	-ATGCCAGCGATGCC---CAA---GCAGGTCCGGCCCCAGTCGCGATGATCGCGTTGA	6137
OBDV	GTCAGGAATCCATGCCTCGCAA---GTTGGTCCGGCCCCCGCCAGCGATGATCGCACTGA	5768
GFkV	CACGTCTGATGCCGCTCTCCGCTGCTGACCAGCCTGCTGTCTCTAGCTCTCGCTCTGA	6485
PnMV	-ATGTCAGACCAAGTC-----GTCAAACAGCTCGAATC-----CAACCACGTCGA	5488
TYMV	-----CGACAAAG-----AACTCGCCCCCAA-----GACCGCACCGT	5685
	* * * *	
MRFV	TC-----GCCAGC-CTCCTCTTCTGATCCTCCTCGTGT-GCTCGAGACGGCCCCGT	5683
CSDaV	CC-----GCCAGC-CTCGCCTTCTGCTGCTCCTCGCGTTGCTGAAGTTGGTCTCAA	6188
OBDV	CC-----GCCAGC-CTTCTCTTCTTCTGCTCCTCGTATTGTGGAGAGCTCTCTCGC	5819
GFkV	CTCTCGCTTGGTTCAGCGCTCCGCTCCTGCTGCGCACCCGCCCGACCGCTATTGCCCG	6545
PnMV	CT-----CCA---CTCCATCTCTCATCCACGCCCCCTTTCCAAGGTGGCAGTCG	5536
TYMV	CA-----CCGTCGCGACCGTTTTACCAGACTGTCCCCGGCCCCCACCTTTCACCAT	5736
	* * * * *	
MRFV	CGCACTTCC--TCGACCTCCCTTTCCAGTGGAAGGTACGGACTTACAGGATACGCCGC	5741
CSDaV	TGCCCGTTCGGTTCGACTACCCGTTCCAGTGGGTCTGCTCGCTCCTACGACGGATCAGAAGC	6248
OBDV	CGTGCCGCACGTCGACGTCCCCTTCCAATGGGCGCTCGCGTCTGACGCCGGAGACTCCGC	5879
GFkV	CAACCCTCGTGTAAAGCATCCATCTCCCCTTCCAGTTCTCTGTGGTATGACATCACGGGCAC	6605
PnMV	CACA-----GTGCCTTTCCA---GACTGTGCGAATGGACGTCGTCGCTGCCGG	5581
TYMV	CAAAC-----AACCGTTCCAGTCTGAAGTCTGTGTTGCTGGGACCAAAGATGC	5784
	* * * * *	
MRFV	CTACCACGGAACCGACGACCTGGTGCCTCCGCGGTACTGA-----CCACACTTTGCGC	5795
CSDaV	CAAGAACCTAAGTATGATCTCTTGGC-TCTGCCACTCTCA-----CCAAAGTCATGGC	6302
OBDV	CAAGTTCCTCACCGACGACCTCTCAGGA-TCTCTCACCTGA-----GCCGCTCACCAT	5933
GFkV	CGAGTCTCTACACCTCCCTGTCCATCGCCTCCGCCCCGTAAGTCTGACCGTCCGCCCG	6665
PnMV	CGGCAACGCAACTTTCAACCTAGCGGGCCACGTCCTCGCTTT-----CTGAGATCACTGC	5635
TYMV	CGAGGCCTC---TCTCACCATCGCCAACATCGACAGCGTTT-----CCACCCTCACCAC	5835
	* * * * *	

MRFV	GCCGTACCGCCATGCCGAGCTCCTCTACGTCGAGATCTCCGTGCGGCCGTGCCCCCTTC	5855
CSDaV	CAACTACCGACATGCTGAGCTCACATCTGTTGAGCTGGAGGCTGCCTCTTGCTGCAGC	6362
OBDV	CGGTATCGCCACGCCGAGCTCATCTCCGCCGAGCTCGAGTTCCGCCCTTGGCCGCCG	5993
GFkV	TCCCTACCGCCACGCCCGCCTCACCTCCCTTGAAGCCTTGTCCAACCCACTGCCTCCTC	6725
PnMV	ACCGTACCGCAAGGCGCGTTTGGCGGAACCTCAAAGCCATCGTCTGCCCCACCGTCCGTC	5695
TYMV	CTTCTATCGTCATGCCTCTCTGGAATCACTCTGGGTACCCATCCATCCTACCTTGCAAGC	5895
	* * * * * * * * *	
MRFV	GTTCTCCAAGCCCATCATGTTACCGTCTGGACTCCGGCCACTCTGTCCCCCGCGA	5915
CSDaV	CTTCTCCAAGCCCATCTCTGTGTCGGCCGTCTGGACCATTGCCTCCATCTCCCCAGCTTC	6422
OBDV	CTTCCGCAAGCCCATCTCCGTCACCCCGCTCTGGACCATAGCCATCCAGCCACGCCA	6053
GFkV	CGCCACCTACCCCGAGACCGTCTGCTGGACCATCGACTCCGGTACCCTCCGCTCCG	6782
PnMV	CTTCCAATCCCCGATCACTCTCGACTCGTCTGGAGCACCAACAATGTCATATTCACCGA	5755
TYMV	CCCAGCTTTCCCGACCACGGTTGGCGTTTGTGGGTACCCGCCAATCCCCAGTCACTCC	5955
	* * * * * * * * *	
MRFV	CGGCAAGGAGACCGACTACTATGGGGGCGGTACAGATCACCGTGGGCGGGCCTGTTATGCT	5975
CSDaV	CGCTCTGAAACCTCTACTATGGCGGTCCGCTCTTCACTGTCCGTTGGTCCAGTCCCTCAT	6482
OBDV	CACCACCGAGCTCCAGTACTACGGTGGCCGACTCCTCACCTCCGAGGCCCGCTCCCTCAT	6113
GFkV	CCGCTCTGAGATCCTCTCCGCTCTTCGGCGCTCAACGCATCGCGTGGGGCTCCGTCCACTT	6842
PnMV	CCTCCAGATACTTACGGTCTACGGGGCACCCGCTTGTCTATTGGCGGGCCCTCCCTCA-	5814
TYMV	CACCCAAATCACCAAGACCTACGGCGGCCAGATCTTCTGCATTGGAGGCGCCATCAACAC	6015
	* * * * * * * * *	
MRFV	GTCCAGCACCACCGCGTAC-CCGCTGACCTGGCCCGCATGAACCCCTTCATCAAGTCCA	6034
CSDaV	GTCCAGTACCACTCACCTTC-CAGCAGATCTCACCCGCTCAATCCTGTGCTCAAGGGCC	6541
OBDV	GGGCTCCGTCACCCGCATCC-CAGCCGACCTCACCCGCTCAACCCCGTCATCAAGACCG	6172
GFkV	CTCCGCCCCATCCTCCTCC-CGCGCGAGCTCTCCAGCCTCAACCCACCATCAAGGACT	6901
PnMV	--GCCACACGTATGAGCTCC-GCGCTGACCTCAGTCACTCAATCCTGTATCAAGGACT	5871
TYMV	TCTCTACCCCTCATTTGTCAAGTGCCCACTTGAATG-ATGAACCCCGGGTCAAAGATT	6074
	* * * * * * * * *	
MRFV	GCGTCTCCTACAACGACAGCCCCGCTGGACCATGTCCGTCCCCG--CCGTCACTGGT--	6090
CSDaV	CTGTCAAGTACACCGACTGCCCCAGATTCTCCTACTCCGTCTACT--CCAAC---GGC--	6594
OBDV	CCGTGGGCTTCACTGACTGCCCCCGCTTCACTACTCCGTCTATG--CCAAC---GGC--	6225
GFkV	CCGTCACTACACCGACTGCCCCCGCTCACCTGTGGCTTCTACC--GCAATGACGCCTG	6959
PnMV	CCGTCTCATACTAGACACCCCGAAGCTCACTCTCAACGCCTCTGACCCAATGGGTCT-	5930
TYMV	CAATTCAATACCTTGACTCGCCCAAACCTCCTCATCTCCATCACCG-CTAAC-CCACC--	6130
	* * * * * * * * *	
MRFV	-----GGGGACACCAAGATCCCGCTCGCCACCGCCTTCGTTGCGGGATCGTCCG	6140
CSDaV	-----GGAACTAAGGGCACCAACCTCTGCACCATCATCCTCCGGGGAGTTGTCCG	6644
OBDV	-----GGGTCCGCCAACACTCCTCTCATCACCGTCATGGTGCAGGAGTTATCCG	6275
GFkV	CGTCCGCTCGGTTCCCTCCGCCCCATCTGCGGCTCCATCCTCATCCGCGGCGTTATCGA	7019
PnMV	-----GGCTCAACCGCCACCACTGTCCGCCACCGTGTGGTGGAGCGGCAA-GCTGA	5979
TYMV	-----GCTCCCCCGCATCGACCTGCATAATAACTGTATCAGGAACCTCTCTCGA	6179
	* * * * * * * * *	
MRFV	CGTGCAGCCCCCTCTGGGGCTGCCACCCCTCTGCCTAATCCATCAGGGCAGCAATAAT	6200
CSDaV	CCTCAGCGGCCCTCCGG-----TAATCTCCTCGCTTA-----GGCGAGCCACTTC	6690
OBDV	CCTTCCGGCCCTTCGGGCAACACCGTCCACCGCCACCTAAGCCCTCACCGGTTTCAAC	6335
GFkV	GTGCTCCGCCCCATCAACCGCCCACTCCCTCGTCATGAACCGAGGCCCCCCCCCTCCGC	7079
PnMV	TGCTCCGACGCATTGGCCACACCCACACCCCTCGAGTAGGACCTTATTAGCAATGTT	6039
TYMV	TGCATTCTCCGCTCATCACGGACACTTCCACCTAAGTTCTCGATCTTTAAATCGTTAGC	6239
	* * * * * * * * *	
MRFV	GCGAGAAGGAAAACCTCGCCAGCCTTGA--GCTGGTAAGTTGTTAAAACC----TCCGCC	6254
CSDaV	AGGTGAAGGAAAACCTCCTGGTCTCA--GCCAGTAATGATGCTAAAC----CTCCCC	6744
OBDV	AGGAGTTTCTTCTCGTTCTTCTCCTGACGACCAATGAACGTTGCTTATC----CCCCCT	6391
GFkV	TCCCGCCCCCTCCAGCCCCCCCCGGCGTCCGCGTTCCCGGTCGCTTTCCTTTCCCC	7139
PnMV	TTGGGTAGTGGAGCCTCTAGCCTCCACCTTTTCTTTCCTCTGGCAATGGCGCCTAAAT	6099
TYMV	TCGCCAGTTAGCGAGGTCTGTCCCCACACGACAGATAATCGGGTGAACCTCCCGCCCCCT	6299
	* * * * * * * * *	
MRFV	---CATCCATGTTTGCAAACTAAGGTAGCAGGTCCGCCATAAGACCTGTGGGC-----	6305
CSDaV	---CGCTCAAGCAGCAATGCCTAGGGTTGCCGGTTCGATCCAAAGACCGTTTTTCTTTTAT	6801
OBDV	TCACATCCCTCCGTTTCCCCCTCCGTTTTCCTCTCTGTTCCATTCCCCCTCTCCCTCCCC	6451
GFkV	TCCCCCTCGCCCGCAACTCCCTCCCTCCGCTCCCCCGCCCCCTACCTGCACCCCC	7199

PnMV	-----	
TYMV	TTCCGAGGGTCATCGGAACC-----	6319
MRFV	-----	
CSDaV	TTAAT-----	6806
OBDV	GTCTCAGCAATGAGTAAGGTTCCAGGTCGATTCAAAGACCTGATGGGATTTTCCTCGG--	6509
GFkV	TCCTCGCCCGTCTCGCGCCCTTCGCCTCCGCCCGCTCCGGACCTCCCATCCCCCCGC	7259
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	TGTCCCCACCGCTCAGCCCCCCCCTCAGCCCCATCTCCCCCTTCACGCCCCGCACCCC	7319
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	CTCCCCACCCCGACCCCGTCTCTCTCCCGCGCTCTCCCTGGCTATCTCTCGCGCCGCTC	7379
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	CCGACCTCCTCCGCCTCCTTAGCCTCCTTTCCCCCCCCTCCCTCTTCTTCTTCAACC	7439
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	TCCTCTCCATCCACTTCTCCCCTTTCCCCATCTTCATCCTCCTTTCCCTCCTCTCTCC	7499
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	TTCAGTTTCCTCGTACCTAGCAATAGCTAGAGGTAGCCGGTTTCTCACTCCGGCCTTCT	7559
PnMV	-----	
TYMV	-----	
MRFV	-----	
CSDaV	-----	
OBDV	-----	
GFkV	TTCTG 7564	
PnMV	-----	
TYMV	-----	

APÊNDICE C

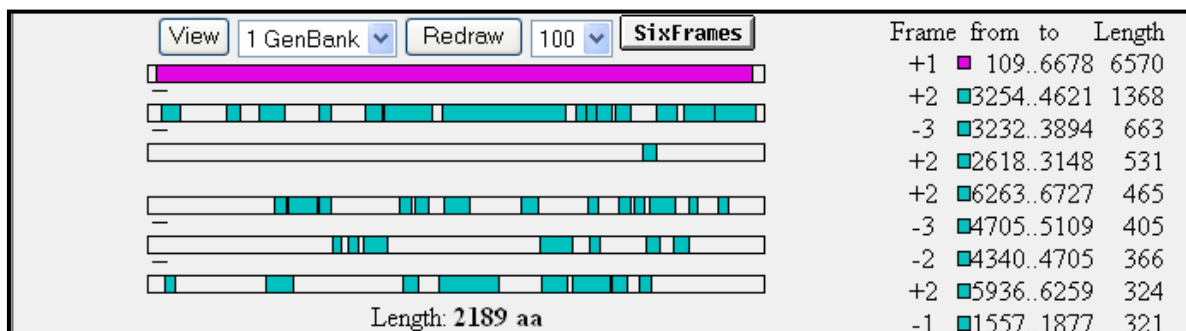
Seqüência de aminoácidos do CSDaV. Tradução realizada através da ferramenta

Transeq – EBI (www.ebi.ac.uk/emboss/transeq).

MDRISARIPVAPASAGPTEYVYPYHTNPLLPRGVFTSGPIQPRLHFLPHHAQDAPIRCYRPLTYANHL
RYDRSASSLKTSPVKLPLTGGTLADAILSLAPTTHRDTIATPLMEALAEPYRQSLSTYPWHIPTNLQP
FLTSCGITTAGQGFKAHPPVHKTIETNLLSNVWSHYATTPSGVMFMKPSKFEKLRKQPNFSKLYNY
RITAKDSTRYPSTSPDLPTEDTCFMHDALMYSPGQICDLFISRPSLQKLYASLVVPPESDFTTISLF
PDLYRYRIEKDQLVYELEQNPAHNYIQPRSAIDWLKTTTIRCQDLTLTISRLLDSWGPVHSLLIQRGRP
PIHLEEDSISFRTPKAVLLPEPASLSQGVDRDLVPADVYQALFIYVRAVRTLRVTDPAQFVRTQISKP
EYSWVTSFRWDFPPLALATAPHRPHTTYFLFNSTAARVSHWFRTHTLALLSGATAAAAASLLMTASWG
FRALISSHLVLSLICKRWLRAPPHELLWPEKAPWFQTLRKPVTGPMIDLPILRPFRLF PSTCAKLGAK
YPALATLLPAAPRPLWPLKVGLALATVPVCLFLWRKF IGPDSPQDMHDSYHAMFHPQPWGLSLARKAI
CCDRAPFLPIPLVPSDFKAPPTPATPLLLTSIPIKGVPEQVPGEGVPPQSASSTGPASDSRCAPQPAS
STGPDQPAQDMSAAPQAPTKSKVTFAQPLESVTPVVPGAGEPPQSASSTGPASVSRRDPQTASSTVPD
APTLDSVTPPPPIIYPVDHLQNDFGPCRCSLCEPPQAPVPSAPPTVSDHKESQDGEALASALQALGL
APAPPAPLSQNLTIESSGAMHASCWDQLSSPSSDWDPSVARDSSATGPPGMYSDLFPAPYLPGTGQF
IFRSRANRANIPYPMDCLLLSIEQATRLPKEALWDTLCATCPDSSLDPDTIRRVGLSTDHFVIFAH
HYSLRCRFHHTAHGVIELGMADATSSFDIDHTAGNPGHFSLRQSATPRLNGGIAQDLAVAALRFNIDGT
LLP IIRSVHVYSTWPKRKNLSSNMKNGFDGIMANIHPKTKNESREKILALDSQLDIARRSRVRLIHIA
GFPGCGKSFPI S RLLRTPTRNFKVAVPTVELRAEWKAI IGLPTSEAWRIGTWESSLLKSARVLVIDE
IYKMPRGYIDLAIHSDPTIEMVIALGDPLQGEYHSTHPSSTNSRLLAEPHLSMYLDYYCMWHRVVPQ
NVAAFFHIKTTSKQPGFCRYQRELPNSRILANSQNAGHTLQCCGYAAVTIASSQGSTYDNAACIHLDR
NSSLLSPAHSMAVALTRSKVGVIFTGDPQLSNAPSSNRMFSEFFSGRTRPLHDWFHNEFPKATVLTEP
LKTRGPRLTGAASPYSKVFPSGKDSTPALKPDFQGDVRRSRAPIVLGSSELNAPQVSTHFLPETRRPL
HWDIPSAIPESATRDPDSTEPTTCHPEPVYPGETFENLAAHFLPAHDPTDREIYWQQLSNQFPHMDKE
FHLAAQPMSLLAAIHQEKQDPTLLPASI QKRLRFRPSDRPYQITPKDEILGQLLFEGLCRAYHRSPFH
TEAFDPVLF AECINLNEFAQLSSKTQATIMGNARRSDPDWRWSAVRIFSKTQHKVNEGSIFGSWKACQ
TLALMHDAVVLIILGPVKKYQRVFDQRDRPRHLYIHAGNTPSQMSNWCQQHLLTAVKLANDYTAFDQSQ
HGEAVVLERKKMERLSIPQALIDLHIHLKTHVSTQFGPLTCMRLTGEPGTYYDDNSDYNLAVINCEYMA
ANTPTMVSGDSSLDRPPTREPEWVILQPLLSLRFKKERGRYATFCGYASHVGCVRSPVALFAKLAI
AVDDGSI SDKMASYLSEFALGHSLGDHLWEALPLEAVPFQSACFDFFCRRAPRHLKLSLMLGEVPESI
IARIGSSLKWASHAIYTTLSAARVAAILRSSRNSRSMPPDDPTTLLQGELLQHFQVPMQSDTLLPLT
GGSSAPILTPEAFSTSLAFSMASDAQAGPAPSRDDRVDQRPLPAAPRVAEVGLNAPSVDYPFQWVVA
SYDGSEAKNLSDDLSGSATLTKVMANYRHAELTSEVELEVCPLAAAFSKPISVSAVWTIASISPASASE
TSYYGGRLFTVGGPVLMSSTHLPADLTRLNPVLKGPVKYTDPCRFSYSVSYNGGKGTNLCTIILRG
VVRLSGPSGNLLA

APÊNDICE D

Resultado da busca por ORFs no genoma do CSDaV através da ferramenta *ORF finder* disponível no NCBI (www.ncbi.nih.gov/gorf/gorf.html), mostrando em rosa a grande ORF codificadora de uma poliproteína de 2189 aminoácidos.



frame +1: a partir do 109 ao 6678 (comprimento 6570)

























APÊNDICE E

Resultado da busca por similaridade a proteínas através da ferramenta PFam disponível no endereço eletrônico www.sanger.ac.uk/software/Pfam/serach.shtml

Trusted matches - domains scoring higher than the gathering threshold. Matches to Pfam-B

Domain	Start	End	Bits	Evalue	Alignment	Mode
Peptidase_C21	930	1035	185.90	8.4e-53	Align	ls
Viral_helicase1	1118	1350	194.70	1.9e-55	Align	ls
RdRP_2	1508	1924	-174.40	0.0013	Align	ls
Tymo_coat	2044	2212	52.60	1.1e-12	Align	ls

APÊNDICE F: Busca por similaridade a proteínas através do InterPro-Scan (www.ebi.ac.uk/InterProScan).

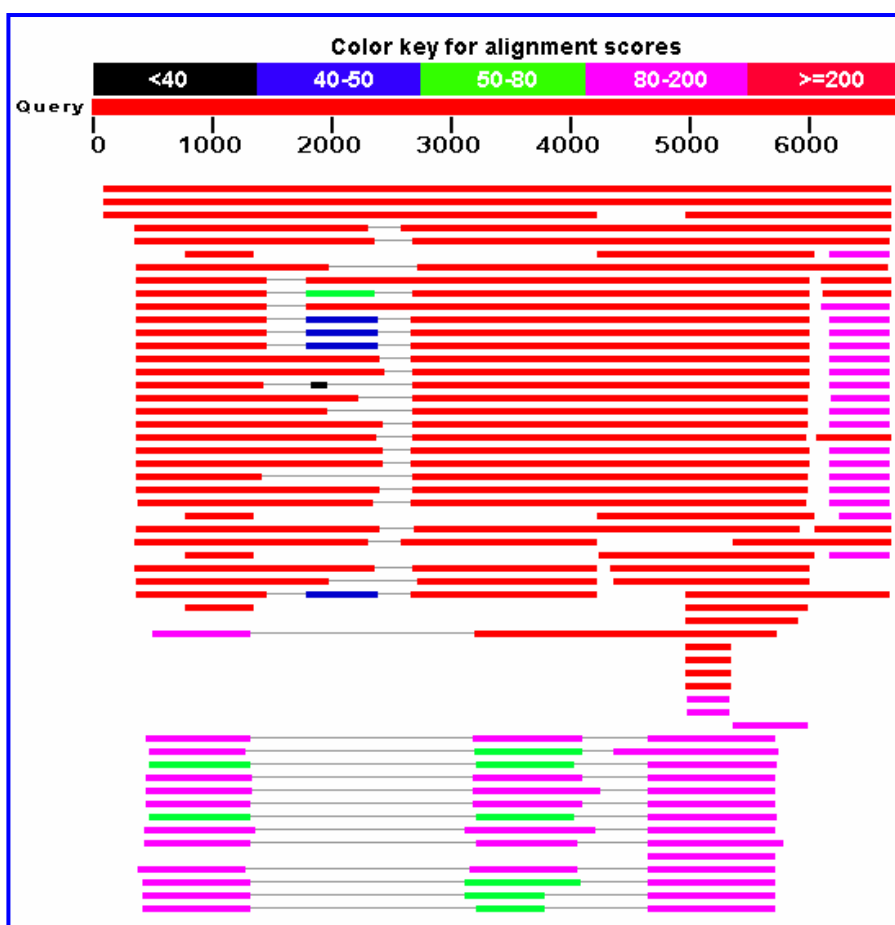
SEQUENCE: <u>Sequence 1</u> CRC64: 37D03902E96CC876 LENGTH: 2262 aa  	
InterPro IPR000574 Domain  	Tymovirus coat protein PD003886  PF00983.8  P89921_VVVVV_P89921; Tymovirus coat protein
InterPro IPR000606 Family  	Viral RNA helicase 1 PF01443.7  Viral (Superfamily 1) RNA helicase
InterPro IPR001093 Family  	IMP dehydrogenase/GMP reductase PF00478.12  IMP dehydrogenase / GMP reductase dom
InterPro IPR001788 Family  	RNA-dependent RNA polymerase PF00978.9  RNA dependent RNA polymerase
InterPro IPR007094 Domain  	RNA-directed RNA polymerase, positive-strand RNA eukaryotic virus PS50507  RDRP_POSITIVE
InterPro IPR007095 Domain  	RNA-directed RNA polymerase, double-strand and positive-strand RNA virus PS50521  RDRP_VIRAL
InterPro IPR008043 Family  	Peptidase C21 PF05381.2  Tymovirus endopeptidase

APÊNDICE G

Resultado da busca por similaridade a proteínas através da ferramenta Blast X disponível no endereço eletrônico <http://www.ncbi.nlm.nih.gov/blast/>.

Query=
Length=6806

Distribution of 186 Blast Hits on the Query Sequence



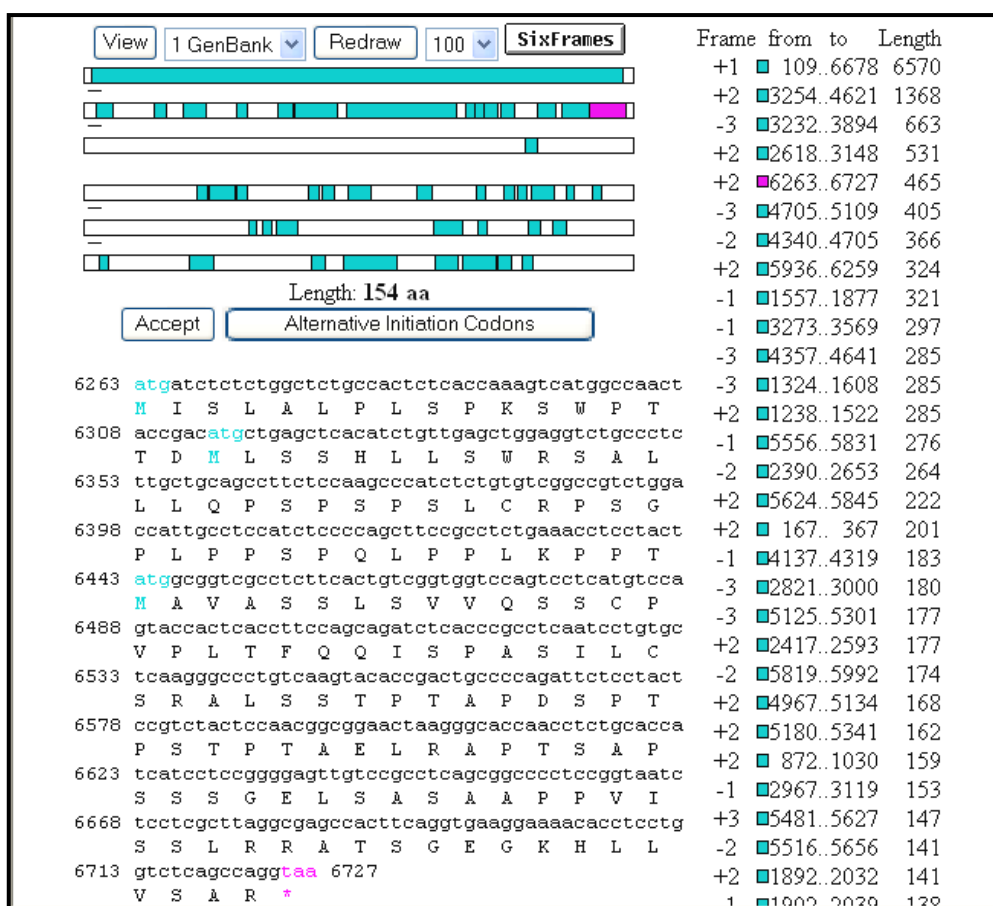
Sequences producing significant alignments:	Score (Bits)	E Value
gi 76577725 gb ABA54133.1 polyprotein [Citrus sudden death-asso	4212	0.0
gi 59609791 gb AAW88343.1 replicase-associated polyprotein [...	4067	0.0
gi 62327639 ref YP_224292.1 methyltransferase/protease/helic...	2448	0.0
gi 9629256 ref NP_044447.1 replicase-associated polyprotein ...	1869	0.0
gi 14030260 gb AAK52838.1 polyprotein [maize rayado fino vir...	1509	0.0

[gi|62327640|ref|YP_224293.1|](#) RNA-dependent RNA polymerase [Ci... [1201](#) 0.0
[gi|6822204|emb|CAB70969.1|](#) replicase-associated protein [Poin... [1172](#) 0.0
[gi|130555|sp|P20128|POLR_TYMVA](#) RNA replicase polyprotein >gi|... [1038](#) 0.0
[gi|9629158|ref|NP_044328.1|](#) replicase [Kennedy yellow mosaic... [1036](#) 0.0
[gi|2723936|gb|AAB92649.1|](#) replicase protein [Turnip yellow mosai [1031](#) 0.0
[gi|62220|emb|CAA34415.1|](#) unnamed protein product [Turnip yell... [1026](#) 0.0
[gi|130554|sp|P10358|POLR_TYMV](#) RNA replicase polyprotein [1026](#) 0.0
[gi|21686952|ref|NP_663297.1|](#) replicase/papain-like protease [... [1023](#) 0.0
[gi|58119514|gb|AAW65470.1|](#) replicase [Dulcamara mottle virus]... [1018](#) 0.0
[gi|59709789|gb|AAW88523.1|](#) polyprotein [Nemesia ring necrosis vi [1012](#) 0.0
[gi|9631137|ref|NP_047920.1|](#) replicase protein [Erysimum laten... [1011](#) 0.0
[gi|59709785|gb|AAW88520.1|](#) polyprotein [Scrophularia mottle viru [1001](#) 0.0
[gi|59709793|gb|AAW88526.1|](#) polyprotein [Plantago mottle virus] [991](#) 0.0
[gi|9627012|ref|NP_041257.1|](#) replicase [Ononis yellow mosaic v... [987](#) 0.0
[gi|11067739|ref|NP_067737.1|](#) replication protein [Chayote mos... [982](#) 0.0
[gi|67141|pir|RRWPEM](#) genome polyprotein - eggplant mosaic virus [978](#) 0.0
[gi|9626696|ref|NP_040968.1|](#) replicase [Eggplant mosaic virus]... [978](#) 0.0
[gi|3378685|emb|CAA76071.1|](#) replicase protein [Physalis mottle... [972](#) 0.0
[gi|59709797|gb|AAW88529.1|](#) polyprotein [Anagyris vein yellowing [961](#) 0.0
[gi|51014248|dbj|BAD35017.1|](#) RNA-dependent RNA polymerase [Bom... [933](#) 0.0
[gi|25013938|ref|NP_734079.1|](#) RNA-dependent RNA polymerase [oat b [926](#) 0.0
[gi|18073602|emb|CAC84400.1|](#) replicase [Grapevine fleck virus]... [922](#) 0.0
[gi|29335719|emb|CAC10493.2|](#) replicase-associated polyprotein ... [910](#) 0.0
[gi|59803285|gb|AAX07751.1|](#) replicase-associated polyprotein [... [881](#) 0.0
[gi|25013941|ref|NP_734076.1|](#) RNA-dependent RNA polymerase [maize [827](#) 0.0
[gi|25013937|ref|NP_734078.1|](#) methyltransferase/protease/helicase [703](#) 0.0
[gi|25013991|ref|NP_733998.1|](#) RNA-depedent RNA polymerase [Poinse [649](#) 0.0
[gi|24797130|ref|NP_733819.1|](#) RNA-dependent RNA polymerase [Turni [645](#) 0.0
[gi|37547398|gb|AAN52168.1|](#) replicase-associated polyprotein [... [600](#) 2e-170
[gi|25013940|ref|NP_734075.1|](#) NTPase/helicase [maize rayado fino [516](#) 3e-145
[gi|25013990|ref|NP_733997.1|](#) methyltransferase/protease/helicase [466](#) 3e-130
[gi|62327641|ref|YP_224294.1|](#) capsid protein [Citrus sudden death [426](#) 3e-118
[gi|3282630|gb|AAC25015.1|](#) replicase protein [Andean potato laten [395](#) 6e-109
[gi|59609792|gb|AAW88344.1|](#) 21 kDa capsid protein [Citrus sudd... [394](#) 2e-108
[gi|24797129|ref|NP_733818.1|](#) replicase/papain-like protease [Tur [393](#) 3e-108
[gi|33326816|gb|AAQ08825.1|](#) replicase RdRp [Grapevine Red Globe v [385](#) 7e-106
[gi|10443133|emb|CAC10494.1|](#) methyltransferase domain [Grapevi... [308](#) 1e-82
[gi|25013939|ref|NP_734080.1|](#) capsid protein [oat blue dwarf viru [267](#) 3e-70
[gi|9629257|ref|NP_044448.1|](#) 21 kDa capsid protein [Oat blue d... [264](#) 3e-69

APÊNDICE H

Resultado da análise genômica do CSDaV através do recurso ORF finder, mostrando em rosa uma possível ORF na extremidade 3' codificadora de uma proteína de 154 aminoácidos, que mostrou similaridade com a proteína hipotética gp3 do CSDaV (Maccheroni *et al.*, 2005), a proteína rica em prolina do GRGV e a proteína hipotética do GFkV, após análise pela ferramenta Blast.

ORF finder:



Blast:

Query=		Score	E
Length=154		(Bits)	Value
Sequences producing significant alignments:			
gb AAW88345.1	hypothetical protein CSdAVgp3 [Citrus sudden d...	260	2e-68
gb AAQ08827.1	17kDa proline rich protein [Grapevine Red Globe v	63.2	5e-09
emb CAC84402.1	hypothetical protein [Grapevine fleck virus] ...	61.2	2e-08
ref YP_294174.1	putative membrane protein [Emiliana huxleyi...	50.8	3e-05
ref NP_932693.1	unknown [Choristoneura fumiferana defective ...	50.4	4e-05
ref YP_053234.1	ICP4 [Gallid herpesvirus 2] >ref YP_053222.1...	50.1	5e-05
gb AAS01711.1	ICP4 [Gallid herpesvirus 2]	48.9	1e-04
gb AAG14284.1	RS1 immediate-early gene transactivator ICP4-1...	48.9	1e-04
ref NP_848394.1	unknown [Choristoneura fumiferana MNPV] >gb ...	48.9	1e-04
ref YP_294122.1	putative membrane protein [Emiliana huxleyi...	47.0	4e-04
ref NP_046248.1	unknown [Orgyia pseudotsugata multicapsid nu...	47.0	4e-04

Alignments

>[gb|AAW88345.1](#) hypothetical protein CSdAVgp3 [Citrus sudden death-associated virus]
[ref|YP_224220.1](#) hypothetical protein CSdAVgp3 [Citrus sudden death-associated virus]

Length=154

Score = 260 bits (664), Expect = 2e-68
 Identities = 136/154 (88%), Positives = 137/154 (88%), Gaps = 0/154 (0%)

```

Query 1  MISLALPLSPKSWPTTDM LSSHLLSWRSALLLQPSPPSLCRPSGPLPPSPQLPPLKPPT 60
          MISLALPLSPKSWPTTDM LSSHLLSWRSALLLQPSPPSLCRPSGPLPPS QLPPLKPPT
Sbjct 1  MISLALPLSPKSWPTTDM LSSHLLSWRSALLLQPSPPSLCRPSGPLPPSLQLPPLKPPT 60

Query 61 MAVASSLSVVQSSCPVPLTFQQISPASILCSRALSSTPTAPDSPTPSTPTAELRAPTSAP 120
          MAV SSL V SSCP P ISPASILCSRA SST TAPDSPTPSTP AE RAP SAP
Sbjct 61 MAVDSSLLAVLSSCPAPPISLLISPASILCSRAPSSTQTAPDSPTPSTPMAEPRAPISAP 120

Query 121 SSSGELSASAAPPVISSLRRATSGEGKHL LVSAR 154
          SSSGELSASAAPPVI SLRRA+SGEGKHL LVSAR
Sbjct 121 SSSGELSASAAPPVIFSLRRASSGEGKHL LVSAR 154
    
```

>[gb|AAQ08827.1](#) 17kDa proline rich protein [Grapevine Red Globe virus]

Length=170

Score = 63.2 bits (152), Expect = 5e-09
 Identities = 52/138 (37%), Positives = 64/138 (46%), Gaps = 11/138 (7%)

```

Query 7  PLSPKSWPT-TDMLSSHLLSWRSALLLQPSPPSLC--RPSGPLPPSPQLPPLKPPTMAV 63
          P SP S+P+ ML+ W S + P+P+ C P+G P SP P P A
Sbjct 26  PPSPPSFPSPAPMLTRSSNPWSSPCSQRTPTPTPCPSMPTGIPPQSPSPAPKSSPPTAA 85

Query 64 ASSLSVVQSSCPVPLTFQQISPASILCSRALSSTPTAPDSPTPST--PTAELR----APT 117
          S SVV S P P + Q S L SR L T T P PST PT +L +
Sbjct 86  PVSPSVVPSPLPTPSSQLTSAPPTLLSRTLCTPTTTPPSRLRPSTRIPTPQLSPSPQSS 145

Query 118 SAPSSSGELSASAAPPVI 135
          +APSSS L S APP +
    
```

Sbjct 146 TAPSSSAVL--SVAPPSV 161

Score = 41.2 bits (95), Expect = 0.022
 Identities = 32/104 (30%), Positives = 43/104 (41%), Gaps = 23/104 (22%)

Query 35 SPSPSLCRPSGPLPPSPQLPPLKPPTMAVASSLSVQSSCPVPLTFQQISPASILCSRAL 94
 SP+P P PPSP P P P+ + +P S CS+
 Sbjct 13 SPTPRPASPPWTWPPSPPSFP-----SPAPMLTRSSNPWSSPCSQRT 54
 Query 95 SSTPTAPDSPTPTPTA-ELRAPTSAPSSSGELSASAAPPVISS 137
 TPT P PS PT ++P+ AP SS +A +P V+ S
 Sbjct 55 PPTPT----PCPSMPTGIPPQSPSPAPKSSPPTAAPVSPSVVPS 94

>[emb|CAC84402.1](#) hypothetical protein [Grapevine fleck virus]
[ref|NP_542614.1](#) hypothetical protein [Grapevine fleck virus]
 Length=309

Score = 61.2 bits (147), Expect = 2e-08
 Identities = 54/140 (38%), Positives = 66/140 (47%), Gaps = 30/140 (21%)

Query 29 ALLLQSPS-----SPSLCRPSGPLPPSPQLP-----PLKPPTMAVASSLSV-- 69
 AL PSP SP L S PLPP P P PL PP + + S ++
 Sbjct 18 ALKSSPSPVPTATPASPLKPLSNPLPPPPTPRPSTSAGPSTPLPPPALRSSPSALNA 77
 Query 70 ---VQSSCPVPLTFQQISPASILCSRALSSTPTAPDSP-----TPSTPTAELRAPTSAP 120
 S+ P P + SPAS SR S TPTAP SP TP++P+ P++AP
 Sbjct 78 SRGAPSTSPPPSSPSPASTPPSRTPSPTPTAPASPVASTAMTPASPSVP-PPPSAAP 136
 Query 121 SSSGELSASAAPPVISSLRR 140
 SSS LS SA PP + L R
 Sbjct 137 SSSAALS-SAPPPSTAPLPR 155

Score = 38.5 bits (88), Expect = 0.15
 Identities = 30/106 (28%), Positives = 44/106 (41%), Gaps = 13/106 (12%)

Query 28 SALLLQSPSPSLCRPSGPLPPSPQLPPLKPP-----TMAVASSLSVQSSCPVPL---- 78
 +AL P PS + P PP P PPL+PP +VA L + + P+ L
 Sbjct 140 AALSSAPPSTAPLPRHEPRPPPLPPPLQPPPGVVRVRSVAFPLPLARELPLRLPPAP 199
 Query 79 ----TFQQISPASILCSRALSSTPTAPDSPTPTAELRAPTSAP 120
 +++P + L S P +P P +P + L AP P
 Sbjct 200 YLHPLLARLAPLRLRPPDLPSPLSPPLSPPLSPISPLHAPAPP 245

Score = 34.3 bits (77), Expect = 2.7
 Identities = 28/101 (27%), Positives = 43/101 (42%), Gaps = 9/101 (8%)

Query 34 PSPSPSLCRPS-GPLPPS-PQLPPLKPPTMAVASSLSVQSSCPVPLTFQQISPASILCS 91
 P PSP + S P+P + P PPLKP + + P P P++ L
 Sbjct 12 PCPSPPALKSSPSPVPTATPASPLKPLS-----NPLPPPPTPRPSTSAGPSTPLPP 64
 Query 92 RALSSTPTAPDSPTPTAELRAPTSAPSSSGELSASAAP 132
 AL S+P++ + + P+ +S PSS S P
 Sbjct 65 PALRSSPSALNASRGAPSTSPPPSSPSPASTPPSRTP 105

Score = 32.7 bits (73), Expect = 8.0
 Identities = 31/109 (28%), Positives = 47/109 (43%), Gaps = 10/109 (9%)

Query 2 ISLALPLSPKSWPTTDLSSHLLSWRSALLLQSPS-----SPSLCRPSGPLPP--SPQ 52
 + LA L P P L LL+ + L L+P P SP L P P+ P +P

Sbjct 184 LPLARELPPLRLPPAPYLHP-LLARLAPLRLRPPDDLPSPLSPPLSPPLSPISPLHAPA 242

Query 53 LPPLKPPTMAVASSLSVVQSSCPVPLTFQQISPASILCSRALSSTPTAP 101
PP P + A SL++ +++ + +SP S+ L S +P

Sbjct 243 PPPHPDPVLLPALSLAISRAAPDLLRLLSLLSPPSLFLLFTLLSIHFSP 291

APÊNDICE I

Alinhamento entre a seqüência nucleotídica do CSDaV obtida neste trabalho (DQ185573) com a seqüência AY884005 através do software ClustalW.

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	AY884005	6820	2	CSDaV	6806	89

CLUSTAL W (1.82) multiple sequence alignment

```

AY884005_      GTCCCCTGTGATCGTCTCTCCCGCCTCCAGCCGAAAGATATTTTGTCTTAACTTTTC 60
CSDaV_        GTCCCCTGTGATCGTCTCTCCCGCCTCATCGCCGAAAGATATTTTGTCTTAACTTTTC 60
*****

AY884005_      TTCGCACTCCTACGCCAGATCTACGTGCCTTAGGTCATCTAAGCCGCTATGGATCGCATC 120
CSDaV_        TTTGCACACCTACGCCAGATCTACGTGCCTTAGGTCATCTAAGCCGCTATGGATCGCATC 120
** ****

AY884005_      TCTGCCCGCATTCCTCGTCCGCGCCGCTTCCGCGGCCGACCGAGTACACTCCATACCCA 180
CSDaV_        TCTGCCCGCATTCCTCGTCCGCGCCGCTTCTGCCGCGGCCGACCGAGTATGTTCCCTTACCCC 180
*****

AY884005_      CACACTACCCACTCCTACCCCGAGGTGTCTTACCTCCGGGCCTATTCAACCCTGTCTC 240
CSDaV_        CACTAACCCTCCTACCCCGAGGCGTTTTTACCTCCGGGCCATCCAACCCCGCCTT 240
*****

AY884005_      CACTTCTTCCCTACCATGCCAAGATGCCCCATCCGCTGCTACAGACCCTCACCTTC 300
CSDaV_        CATTCCTTCCCTACCATGCCAAGATGCCCCATCCGCTGCTATCGACCCTCACCTAT 300
** *

AY884005_      GCCAACCATCTCCGCTATGACCGTCCGCTCATCGCTCAAGACTCCGCCGTCAAACTC 360
CSDaV_        GCCAACCATCTCCGCTATGACCGTCCGCTCATCGCTCAAGACTTCGCCCGTCAAGCTC 360
*****

AY884005_      CCACTGACCGGTGGTACCCTTGCCGATGCCATCCTTTCCTTGGCACCCACTCACCGC 420
CSDaV_        CCGCTGACCGGTGGTACTCTTGCCGATGCCATCCTTTCCTAGCTCCCACTCACCGT 420
** *****

AY884005_      GACACCATCGCCACCCCTCATGGAAGCCCTTGCTGAACCTTACCGCCAATCCTTGAGC 480
CSDaV_        GATACCATTGCCACCCCTTATGGAAGCCCTTGCTGAACCTTACCGCCAGTCCCTTGAGC 480
** *****

AY884005_      ACCTACCCATGGCACATTCCAACCAATCTTCAGCCCTTCCTCACCTCTTGCGGAATCACC 540
CSDaV_        ACCTACCCCTGGCACATTCCAACCAATCTTCAGCCCTTCCTCACCTCTGTGGAATCACC 540
*****

AY884005_      ACTGCTGGCCAAGGCTTCAAGGCCACCCTCACCCAGTGACAAGACCATCGAGACCAAT 600
CSDaV_        ACTGCTGGCCAAGGCTTCAAGGCCACCCTCACCTGTGACAAGACCATTGAGACCAAT 600
*****

AY884005_      CTCCTCACTAATGTCTGGCCCCACTACGCCACCCTCCTAGTGGCGTCATGTTTCATGAAA 660
CSDaV_        CTCCTCTCTAATGTCTGGTCCCCTACGCCACCACCCAGTGGCGTCATGTTTCATGAAA 660
*****

AY884005_      CCATCAAAGTTTGAGAAGCTCAAATCAAACAGCCCAACTTCTCCAAGCTCTACAACCTAC 720
CSDaV_        CCATCAAAGTTTGAGAAGCTCAGAATCAAACAGCCCAACTTCTCCAAGCTCTACAACCTAC 720
*****
    
```

AY884005_ CGCATCACAGCCAAGGACACCACCCGTTACCCCTCCACTTCCCCAGACTTGCCACCGAG 780
CSDaV_ CGCATCACAGCCAAGGACTCCACCCGTTATCCCTCCACTTACCAGACCTACCCACTGAG 780

AY884005_ GACACCTGCTTCATGCACGATGCCCTCATGTACTATTCCTGGACAGATCTGTGACCTC 840
CSDaV_ GACACCTGCTTTATGCATGACGCCCTGATGTACTACTCCCCGGCCAGATCTGTGACCTC 840

AY884005_ TTCCTCTCCCGCCCAGCCTCCAAAAGCTCTATGCTTCCCTTGTTGTCTCCGGAGAGC 900
CSDaV_ TTCATCTCTCGCCCCAGCCTCCAAAAGCTCTATGCTTCTCTCGTGGTTCCCCCTGAAAGT 900

AY884005_ GACTTCACCACCATCTCCCTCTTCCCAGATCTCTACCCTACCGGATCGAGAAAGATCAG 960
CSDaV_ GACTTCACCACCATCTCCCTCTTCCCAGACTTGTATCGCTACAGGATTGAGAAGGACCAG 960

AY884005_ CTCATCTACGAGCTCGAGCAGAACCCCGCCACAACACTACATCCAGCCTCGCTCTGCCATC 1020
CSDaV_ CTTGTCTATGAGCTGGAGCAGAACCCGGCCACAACACTACATTCAGCCCCGCTCCGCCATT 1020
**

AY884005_ GACTGGCTCAAGACCACCACCATCCGCTGCCAGGACCTCACTCTCACCATCTCCCGCTA 1080
CSDaV_ GACTGGCTGAAGACCACCACCATCCGCTGCCAAGACCTAACCCTCACAACTCTCCCGCTT 1080

AY884005_ GATTCCTGGGGCCAGTCCACTCTCTCTGATCCAAAGAGGCAAGCCCTATCCATCTT 1140
CSDaV_ GACTCCTGGGGCCAGTTCACTCTCTTTGATCCAAAGAGGTAGGCCACCCATCCACCTT 1140
**

AY884005_ GAGGAGGACTCCATCAGCTTCCGTGCCCAAAAGCAGTCTCTGCCTGAGCCAGCTTCA 1200
CSDaV_ GAAGAGGACTCCATCAGCTTCCGCACTCCCAAAGCTGTCTCTCTGAACCAGCTTCT 1200
**

AY884005_ CTCTCCAATCAGTCCGCGACCGCTGGTCCCTGCTGATGTTTACCAGGCTCTCTTCATC 1260
CSDaV_ TTATCCAGGGGTTTCGTGACCGCTTGTCTCTGCTGATGTCTACCAGGCTCTCTTCATC 1260
*

AY884005_ TATGTCCGGGCTGTCCGCACCCCTCCGTGTGACCGACCCGGCTGGCTTCGTTCCGACTCAG 1320
CSDaV_ TATGTCAGAGCCGTTTCGGACCCCTCCGTGTCACTGACCCCGCTGGGTTTGTCCGAATCAG 1320

AY884005_ ATCTCTAAGCCCGAGTACTCTTGGGTCACCTTCCCTCCGCTGGGACAATCTGGCCACTTT 1380
CSDaV_ ATTTCAAAGCCCGAGTACTCTTGGGTCACCTTCCCTCCGCTGGGACAATTTCCCCCTTG 1380
**

AY884005_ GCCTTGCCACAGCTCCGCACAGACCCACACCACCTACTTCCGTGTTCAACTCAACCCT 1440
CSDaV_ GCCCTTGCCACTGCTCCACACAGACCCACACCACCTACTTCCCTCTCAATTCCACTGCG 1440

AY884005_ GCTCGGGTGGCCATTTGGTTCCGCACTCATACCCTGGCTCCGCTCTCTGGCGCCACTGCT 1500
CSDaV_ GCCCGGGTATCCCACTGGTTCCGCACCCATAACCTTGGCCCTGCTCTCTGGCGCCACTGCA 1500
**

AY884005_ GCCGCCGCGAGCCTTCTCATGACCGCAGCTGGGGATTCCGTGCCATGATCTCTCTCAT 1560
CSDaV_ GCCGCCGCGAGCCTCTCATGACCGCAGTTGGGGCTTCCGCGCTCTCATTTCTTCTCAC 1560

AY884005_ CTTGTCTCCCTCTCCATCTGCAAGCGCTGGCTCAAAGCTCCTCCTCATCTCTCTGGCCC 1620
CSDaV_ CTTGTCTCCCTCTCCATCTGCAAGCGCTGGCTGAGGGCCCTCCCATCTCCTTTGGCCC 1620

AY884005_ GAGAAAGCTCCCTGGTTCCAGCTCACCTGAGGCCAAAGTCACTGGCCCTCTGATTGAC 1680
CSDaV_ GAGAAGGCTCCATGGTTCCAACCTCACTGCGGCCAAAGTCAACCGCCCCATGATTGAC 1680

AY884005_ CTGCCATTCTCCGACCCTTTAGGCTGTTCCCTTCCACATGCGCCAAACTGGCGCCAAG 1740
CSDaV_ CTGCCATCTCCGACCATTGAGATTGTTCCATCCACCTGTGCTAAATTGGCGCCAAG 1740

AY884005_ ATCTTCCGCTCCAGGGCCAATGGTCGGGCCAACATCCCTTATCCCGACATGGATTGCCTC 2820
CSDaV_ ATCTTCCGCTCCAGAGCCAACGGCCGGGCCAACATTCCTTACCCCGACATGGATTGTCTC 2820
***** ** ** ***** ** ** ***** **

AY884005_ TTGCTTTCATCGAGCAAGCCACCCGCTTCCCAAGGAGGCTCTCTGGGACACCTCTGT 2880
CSDaV_ CTACTTTCATCGAGCAGGCTACCCGCTTCCCAAGGAGGCCCTCTGGGACACTCTCTGT 2880
* ***** ** ***** ***** ***** *****

AY884005_ GCCACATGCCCCGACTCTCTCCTTGATCCTGATACCATTGCGCCGAGTCGGATTGTCCACT 2940
CSDaV_ GCCACCTGCCCCGACTCTCTGCTTGACCCTGACACCATCCGCGGAGTCGGCCTCTCCACA 2940
***** ***** ***** ***** ***** ***** * *****

AY884005_ GACCACTTTGCCATCCTGGCCACCCTACTCCCTCAGGTGCCGCTTTTACACCGCCCAT 3000
CSDaV_ GACCATTTTGTATCTTGCACCACCCTACTCCCTCAGGTGCCGCTTTCATACCGCCCAT 3000
***** ** * ***** ***** ***** ** ** *****

AY884005_ GGTGTCATTGAGCTCGGCATGGCTGATGCCACCTCCTCATTGACATCGACCACACTGCT 3060
CSDaV_ GGTGTCATTGAGCTGGGCATGGCCGACGCCACTTCATCTTTTGATATCGACCACACTGCC 3060
***** ***** ** ***** ** ** ** ** ***** *****

AY884005_ GGCAACCCCGGCCACTTCTCCCTCCGGCAATCTGCCACTCCGAGGCTAAATGGAGGAATT 3120
CSDaV_ GGCAATCCTGGCCACTTCTCCCTCCGGCAGTCTGCCACGCCCCGGTCAATGGTGGGATT 3120
***** ** ***** ***** ***** ** * ***** ** **

AY884005_ GCTCAAGATCTCGTGTGGCCGCTCTCAGGTTCAACATTGATGGCACTCTCCTCCCAATC 3180
CSDaV_ GCTCAGGACCTTGCTGTGGCCGCTTAAGGTTCAACATTGACGGCACTCTCCTCCCATC 3180
***** ** * ***** * ***** ***** ***** *****

AY884005_ CGCTCAGTTCATGTCTATTCCACTTGGCCAAAGAGAGCAAAGAACCCTGTCTCGAACATG 3240
CSDaV_ CGCTCAGTTCATGTCTATTCCACTTGGCCAAAGAGAGCAAAGAACCCTGTCTCAAATATG 3240
***** ***** ***** ***** ***** ***** ***** ** ** **

AY884005_ AAGAACGGCTTTGACGGCATCATGGCCAACATCCACCCACCAAGACCAATGAATCGAGA 3300
CSDaV_ AAAAACGGCTTTGATGGAATCATGGCCAACATTCATCCACCAAGACCAATGAATCCAGA 3300
** ***** ** ***** ***** ** ***** ***** *****

AY884005_ GAGAAGATCTTGGCACTCGATTTCGAGCTGGACATCGCTGTGAGGAGATCCGTCGGTCTG 3360
CSDaV_ GAAAAGATTTTGGCCCTGGACTCCAGCTTGATATTGAGTCCGGAGATCTGTCCGCTC 3360
** ***** ***** ** ** ** ***** ** ** * ***** ***** *****

AY884005_ ATCCATATTGCCGGTTCCAGGGTGGCCAAAGTCCCTTTCCCATCTCCGCCTCCTCCGC 3420
CSDaV_ ATCCACATTGCCGGCTTTCCAGGCTGGCCAAAGTCCCTTTCCCATCTCCGCCTCCTCCGC 3420
***** ***** ** ***** ***** ***** ***** ***** *****

AY884005_ ACTCCAACCTTCAGGAACCTTAAGGTGGCAGTTCCCACTGTTGAGCTCCGAGCCGAGTGG 3480
CSDaV_ ACTCCACCTTCAGGAACCTTAAGGTGGCTGTGCCAACTGTGAGCTTCGAGCTGAATGG 3480
***** ***** ***** ***** ** * ***** ***** ***** ** **

AY884005_ AAAACCATTAAGTCTCCCGGCTCAGAAGCCTGGCGCATCGGCACCTGGGAATCCTCT 3540
CSDaV_ AAAGCAATCATGGCCTCCCACTTCTGAGGCATGGCGCATCGGTACCTGGGAATCCTCT 3540
*** * ** * ** ***** * ** ** ***** ***** ***** *****

AY884005_ CTCTCAAGTCTGCCGGGTCTGGTTCATTGATGAAATCTACAAGATGCCAAGAGGCTAC 3600
CSDaV_ CTGCTCAAATCTGCCGGGTCTCGTTCATTGATGAAATCTACAAGATGCCCGGGGCTAC 3600
** ***** ***** ** ***** ***** ***** ***** * *****

AY884005_ ATTGATCTCGCCATCCACTCTGATCCCACCATTGAAATGGTCATTGCTCTCGGTGATCCA 3660
CSDaV_ ATTGATCTCGCCATCCACTCTGATCCAACCATTGAGATGGTCATTGCTCTCGGTGATCCA 3660
***** ***** ***** ***** ***** ***** ***** *****

AY884005_ CTCCAAGGAGGTACCCTCCACTCATCCTTCTTCCAACTCCCGCTTCTCTCTGAG 3720
CSDaV_ CTCCAAGGCGAGTACCCTCCACTCATCCTTCTTCCAACTCCCGCTTCTTGTGAA 3720
***** ***** ***** ***** ** ***** ***** *****

AY884005_ CCCACACCTCTCCATGTACCTTACTTCTACTGCTTGTGGTCCCACCGGTTCCGCAG 3780
CSDaV_ CCACACCACCTCTCCATGTATCTCGACTACTACTGTATGTGGTCCCATAGAGTCCCTCAG 3780
** ** * ***** ** ***** ***** ***** ***** * ** ** **

```

AY884005_   AACGTGGCCGCTTCTTCCATGTCAAGACCACCTCCAAACAGCCTGGCTTCTGCCGCTAC 3840
CSDaV_      AATGTCGCTGCTTTCTTCCACATCAAGACCACCTCCAAGCAGCCAGGCTTCTGCCGCTAC 3840
* * * * *

AY884005_   CAGAGAGAGCTGCCGAAGTCCAGAATCCTGGCCAACTCTCAGAATGCAGGCCATAACCTC 3900
CSDaV_      CAGCGAGAGCTTCCCAACTCCAGGATCCTGGCCAACTCTCAAATGCAGGCCATACTCTT 3900
*** * * * *

AY884005_   CAGCAGTGTGGCTACGCTGCCGTCACCATTGCCCTCAGTCAGGGCTCCACCTATGAAAA 3960
CSDaV_      CAGCAATGCGGCTATGCCGCCGTCACCATTGCCCTCAGTCAGGGTTCACCTTATGACAAT 3960
***** * * * *

AY884005_   GCGGCTGCATTACCTGGACCGAAACAGCTCCTTGCTCTCCCTGCTCACTCCATGGTT 4020
CSDaV_      GCGGCTGCATCCATCTTGACCGCAACAGCTCCCTGCTCTCCCTGCTCATTCCATGGTC 4020
***** * * * *

AY884005_   GCTCTCACTCGCTCAAAGGTTGGTGTTCATCTTCACCGGGGATCCC GCCAGCTCTCCAAT 4080
CSDaV_      GCTCTCACCCGCTCCAAGGTTGGTGTTCATCTTCACCGGGGATCTGCCAGCTGTCCAAT 4080
***** * * * *

AY884005_   GCTCCAAGCTCCAACCGAATGTTCTCAGAGTTCCTTCTCAGGCCGCACCCGCCCTCTTCAT 4140
CSDaV_      GCCCGAGCTCCAACAGAATGTTCTCTGAATTCCTTCTCAGGCCGCACCCGCCCTCTTCAT 4140
* * * * *

AY884005_   GACTGGTTCACAATGAGTTCCCAAAGGCCACTGTCCTCACCAGCCCTCAAGACTCGG 4200
CSDaV_      GACTGGTTCACAATGAGTTCCCAAAGGCCACTGTCCTCACCAGCCCTCAAGACTCGG 4200
*****

AY884005_   GGGCCCCGCTCACCAGTGTGCTGCCTCACCATACTCCAAGGCTGTCCAATCCGCCAAGCC 4260
CSDaV_      GGGCCCCGCTCACCAGTGTGCTGCCTCACCATACTCCAAGGTTGCCATCCGCCAAGGAT 4260
***** * * * *

AY884005_   TCCACCCAGCTCTCAAGCCTGATTTCCAAGGGGACGTCA--TAATCTCAG-CACCCATA 4317
CSDaV_      TCCACCCGGCCCTCAAACCTGATTTCCAGGGGACGTTCGTTCTTCTCGGGCCCCCATC 4320
***** * * * *

AY884005_   GTTCTCGGCTCCGGCGAGCTCAATGCCCTCAAGTCTCCTCTCACTTCCTCCCCGAGACT 4377
CSDaV_      GTTCTGGGCTCCGGCGAGCTCAATGCTCCACAGGTCTCAACTCACTTCCTTCCCTGAGACT 4380
***** * * * *

AY884005_   CGCCGTCTCTCCACTGGGACATTCATCTGCCATCCTGAGAGTGCCACCAGACCGGAC 4437
CSDaV_      CGCCGCCCTCTTCACTGGGACATCCTTCTGCCATCCTGAGAGTGCCACCAGGCCTGAT 4440
***** * * * *

AY884005_   TCCACTGAGCCCACCACCTCCCATCCAGAGCCAGTCTACCCGGGGAAACTTTTGAGAA 4497
CSDaV_      TCCACTGAGCCCACCACCTTCCATCCTGAGCCGGTTTACCCAGGGGAAACCTTTGAAAAT 4500
***** * * * *

AY884005_   CTTGCTGCCCACTTTCTCCCTGCCACGACCCAACCGATCGTGAGATCTACTGGCAGGGT 4557
CSDaV_      CTTGCTGCCCACTTTCTCCCGCTCACGATCCCAGTATCGAGAAATCTACTGGCAGGGC 4560
***** * * * *

AY884005_   CAGCTGTCCAACCAGTTCACACATGGACAAGGAATTCATTGGCTGCACAACCCATG 4617
CSDaV_      CAGCTGTCCAACCAGTTCCTTCCATGGACAAGGAATTCACCTGGCCGCTCAGCCCATG 4620
***** * * * *

AY884005_   AGTCTCCTGGCTGCCGTTTCATCAAGAGAAGCAAGATCCCAGTCTACTGCCAGCTTCAATC 4677
CSDaV_      AGCTCCTGGCTGCCGATTTCATCAAGAGAAGCAAGATCCCAGTCTACTGCCAGCTTCCATC 4680
* * * * *

AY884005_   CAAAAGAGACTCCGCTTCCGCCCCCTCCGACAAGCCCTACCAGATCACCCAAAAGATGAA 4737
CSDaV_      CAGAAGAGGCTCCGCTTCCGCCCCATCAGACAGGCCTTACCAGATCACTCCAAAAGATGAA 4740
* * * * *

AY884005_   ATCCTGGGCCAGTCTCTTTGAAGGCTCTGCCGAGCTACCACAGATCTCCATTTCCAC 4797
CSDaV_      ATCCTGGGCCAATTGCTCTTTGAAGGCTCTGCCGAGCTATCATCGGTCCCCTTTCCAC 4800
***** * * * *
    
```

AY884005_ ACTGAGGCCTTTGATCCCGTGCTTTTCGCCGAGTGCATCAATCTCAATGAGTTCGCCAG 4857
CSDaV_ ACTGAGGCCTTTGATCCTGTTCGTTCGCCGAGTGCATCAATCTGAATGAATTGCCAG 4860
***** ** * *****

AY884005_ CTCTCGTCCAAGACCCAGGCTACTATATATGGGCAATGCTCGCCGCTCAGACCCTGATTGG 4917
CSDaV_ CTCTCCTCCAAGACTCAGGCCACCATCATGGGCAATGCTCGCCGCTCTGATCCAGACTGG 4920
***** ***** ** * *****

AY884005_ CGGTGGAGCGCAGTTCGCATCTTCTCCAAGACCCAACACAAGGTGAATGAAGGGTCCATT 4977
CSDaV_ CGGTGGAGTGTCTGCCGATCTTCTCCAAGACCCAGCACAAAGTAAATGAAGGGTCCATC 4980
***** ** * *****

AY884005_ TTCCGCTCCTGGAAGGCCTGCCAACTTTGGCTCTCATGCATGATGCTGTTGTTCTAATC 5037
CSDaV_ TTTGGTTCTGGAAGGCATGTGACTTTGGCCCTCATGCATGATGCTGTTGTTCTTATC 5040
** * ***** ** * *****

AY884005_ CTGGGCCCTGTCAAGAAGTACCAGCGAGTCTTTGATCAGAGAGACCCGACCCCGACACCTT 5097
CSDaV_ CTGGGCCCTGTCAAGAAGTACCAGAGAGTGTGGTACCAGAGAGACCCCGGCACCTT 5100
***** ***** ** * *****

AY884005_ TACATCCATGCAGGCAACACTCCATCACAATGAGCAACTGGTGTCAACAGCATCTCACT 5157
CSDaV_ TACATCCATGCCGCAACACTCCTTACAGATGAGCAACTGGTGCAGCAGCATCTCACC 5160
***** ***** ** * *****

AY884005_ ACTGCCGTCAAGTTGGCCAATGACTACACTGCCTTCGACCAGTCTCAGCATGGTGAAGCG 5217
CSDaV_ ACCGCTGTCAAGCTGGCCAATGATTACACCGCCTTTGATCAATCTCAGCATGGGGAGGCA 5220
** * ***** ***** ** * *****

AY884005_ GTCGTCCTTGAAAGAAAGAAAATGGAAAGACTCTCCATCCCCAGGCTCTCATTGATCTT 5277
CSDaV_ GTTGTCTCGAAAGAAAGAAATGGAAAGGCTCTCCATTCACAGGCCCTCATTGACCTC 5280
** * * ***** ***** ** * *****

AY884005_ CACATCCATCTCAAAACCCATGTTTCCACCCAGTTTGGCCCCCTCACATGCATGCGCCTG 5337
CSDaV_ CACATCCATCTCAAGACCCATGTCTCCACCCAGTTTCGGCCCCCTCACTTGCATGCGCCTG 5340
***** ***** ***** ***** *****

AY884005_ ACTGGCGAGCCTGGCACTTATGATGATAACTCTGACTACAATCTTGCAGTTGTCAACTGT 5397
CSDaV_ ACTGGCGAACCCTGGCACCTATGATGACAATTCGACTATAATCTTGGGGTATCAACTGT 5400
***** ***** ***** ** ***** *****

AY884005_ GAGTACATGGCTGCCAACACTCCACTATGGTCTCAGGCGACGACTCCCTCCTGGATCGT 5457
CSDaV_ GAGTACATGGCTGCCAACACTCCACTATGGTCTCAGGCGACGACTCCCTCCTGGATCGT 5460

AY884005_ GAGCCTCCACTCGCCCTGAATGGGTATCCTCCAGCCTCTTCTCAGTCTCCGCTTCAAG 5517
CSDaV_ GAGCCTCCACTCGCCCTGAATGGGTATCCTCCAGCCTCTTCTCAGTCTCCGCTTCAAG 5520

AY884005_ AAAGAAAGGGGTTCGGTACGCCACCTTCTGTGGCTACTACGCCTCCCATGTTCGGCTGTGTC 5577
CSDaV_ AAAGAAAGGGGTTCGGTACGCCACCTTCTGTGGCTACTACGCCTCCCATGTTCGGCTGTGTC 5580

AY884005_ CGCTCCCCGTGGCTCTCTTTGCCAAGCTGGCCATAGCTGTTCGATGACGGCTCCATCTCT 5637
CSDaV_ CGCTCCCCGTGGCTCTCTTTGCCAAGCTGGCCATAGCTGTTCGATGACGGCTCCATCTCT 5640

AY884005_ GACAAAATGGCCTCATACTCTCTGAATTTGCTCTTGGCCACTCCCTTGGAGACCATCTC 5697
CSDaV_ GACAAAATGGCCTCATACTCTCTGAATTTGCTCTTGGCCACTCCCTTGGAGACCATCTC 5700

AY884005_ TGGGAAGCTTTGCCCTCGAGGCCGTTCCCTTCCAATCTGCCTGCTTGGACTTCTTCTGC 5757
CSDaV_ TGGGAAGCTTTGCCCTCGAGGCCGTTCCCTTCCAATCTGCCTGCTTGGACTTCTTCTGC 5760

AY884005_ CGCCGGGCCCCCAGACACCTCAAACCTCTCTCATGCTCGGCGAGGTCCAGAATCCATC 5817
CSDaV_ CGCCGGGCCCCCAGACACCTCAAACCTCTCTCATGCTCGGCGAGGTCCAGAATCCATC 5820

AY884005_ ATTGCCCGCATCGGGTCATCCTTGAAGTGGGCCTCTCATGCCATCTACACCACACTCTCC 5877
 CSDaV_ ATTGCCCGCATCGGGTCATCCTTGAAGTGGGCCTCTCATGCCATCTACACCACACTCTCC 5880

AY884005_ TCTGCCGCTCGAGTGGCCATTCTGAGATCCTCCCGCAACAGCAGATCCATGCCAGATGAC 5937
 CSDaV_ TCTGCCGCTCGAGTGGCCATTCTGAGATCCTCCCGCAACAGCAGATCCATGCCAGATGAC 5940

AY884005_ CCCGACACCACACTCTGCTACAAGGTGAATTGCTTCAGCACTTTCAAGTACCATTTCATGCAA 5997
 CSDaV_ CCCGACACCACACTCTGCTACAAGGTGAATTGCTTCAGCACTTTCAAGTACCATTTCATGCAA 6000

AY884005_ TCTGACACTCTCCTGCCTCTCACTGGTGGTTCCCTCTGCTCCCATCCTCACACCAGAAGCC 6057
 CSDaV_ TCTGACACTCTCCTGCCTCTCACTGGTGGTTCCCTCTGCTCCCATCCTCACACCAGAAGCC 6060

AY884005_ TTCTCCACCTCCCTCGCCTTCTCCATGGCCAGCGATGCCAAGCAGGTCCGGCCCCCAGT 6117
 CSDaV_ TTCTCCACCTCTCTCGCTTCTCCATGGCCAGCGATGCCAAGCAGGTCCGGCCCCCAGT 6120
 ***** *****

AY884005_ CGCGATGATCGCGTTGACCGCCAGCCTCGCCTTCCTGCTGCTCCTCGGTTGCTGAAGTT 6177
 CSDaV_ CGCGATGATCGCGTTGACCGCCAGCCTCGCCTTCCTGCTGCTCCTCGGTTGCTGAAGTT 6180

AY884005_ GGTCCTCAATGCCCGTCGGTGCAGTACCCGTTCCAGTGGGTTCGTCGCTCCTACGACGGA 6237
 CSDaV_ GGTCCTCAATGCCCGTCGGTGCAGTACCCGTTCCAGTGGGTTCGTCGCTCCTACGACGGA 6240

AY884005_ TCAGAAGCCAAGAACCTAAGTGATGATCTCTCTGGCTCTGCCACTCTACCAAAGTCATG 6297
 CSDaV_ TCAGAAGCCAAGAACCTAAGTGATGATCTCTCTGGCTCTGCCACTCTACCAAAGTCATG 6300

AY884005_ GCCAACTACCGACATGCTGAGCTCACATCTGTTGAGCTGGAGGTCTGCCCTCTTGCTGCA 6357
 CSDaV_ GCCAACTACCGACATGCTGAGCTCACATCTGTTGAGCTGGAGGTCTGCCCTCTTGCTGCA 6360

AY884005_ GCCTTCTCCAAGCCCATCTCTGTGTCGGCCGCTCTGGACCATTGCCTCCATCTCTCCAGCT 6417
 CSDaV_ GCCTTCTCCAAGCCCATCTCTGTGTCGGCCGCTCTGGACCATTGCCTCCATCTCTCCAGCT 6420
 ***** *****

AY884005_ TCCGCCTCTGAAACCTCCTACTATGGCGGTCGACTCTTCACTGTTGGCGGTCCTGTCTC 6477
 CSDaV_ TCCGCCTCTGAAACCTCCTACTATGGCGGTCGACTCTTCACTGTTGGCGGTCCTGTCTC 6480
 ***** ***** ** *****

AY884005_ ATGTCCAGCACCACCATCTCCCTGCTGATCTCACCCGCCTCAATCCTGTGCTCAAGGGC 6537
 CSDaV_ ATGTCCAGTACCACCTCACCTTCCAGCAGATCTCACCCGCCTCAATCCTGTGCTCAAGGGC 6540
 ***** ** * * * * * *****

AY884005_ CCCGTCAAGTACACAGACTGCCCCAGATTCTCCTACTCCGTCTACTCCAATGGCGGAACC 6597
 CSDaV_ CCTGTCAAGTACACAGACTGCCCCAGATTCTCCTACTCCGTCTACTCCAACGGCGGAACC 6600
 ** ***** *****

AY884005_ AAGGGCACCAATCTCTGCACCATCATCCTCCGGGGAGTTGTCCGCCTCAGCGGCCCTCC 6657
 CSDaV_ AAGGGCACCAACCTCTGCACCATCATCCTCCGGGGAGTTGTCCGCCTCAGCGGCCCTCC 6660
 ***** *****

AY884005_ GGTAATCTTCTCGCTTAGGCGAGCCTCTCAGGTGAAGGAAAACACCTCTGGTCTCAGC 6717
 CSDaV_ GGTAATCTCCTCGCTTAGGCGAGCCTCTCAGGTGAAGGAAAACACCTCTGGTCTCAGC 6720
 ***** *****

AY884005_ CAGGTAATGATGCTAAACCTCCCCGCTCAAGCAGCAATGCCTAGGGTTGCCGGTCGATC 6777
 CSDaV_ CAGGTAATGATGCTAAACCTCCCCGCTCAAGCAGCAATGCCTAGGGTTGCCGGTCGATC 6780

AY884005_ CAAAGACCGTTTTTCTTTATTATTTAATAAAAAAAAAAAAAAAAAA 6820
 CSDaV_ CAAAGACCGTTTTTCTT--TTATTTAAT----- 6806
 ***** *****

APÊNDICE J

Alinhamento entre a seqüência de aminoácidos do CSDaV obtida neste trabalho (DQ185573) com a seqüência AY884005 através do software ClustalW.

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	CSDaV	2189	2	AY884005	2188	96

CLUSTAL W (1.82) multiple sequence alignment

CSDaV AY884005	MDRISARIPVAPASAGPTEYVYPYHTNPLLPRGVFTSGPIQPRLHFLPHHAQDAPIRCYR MDRISARIPVAPASAGPTEYTPYPTHPLLRGVFTSGPIQPCPLHFLPHHAQDAPIRCYR *****:*****	60 60
CSDaV AY884005	PLTYANHLRYDRSASSLKTSPVKLPLTGGTLADAILSLAPTTHRDTIATPLMEALAEPYR PLTFANHLRYDRSASSLKTTPVKLPLTGGTLADAILSLAPTTHRDTIATPLMEALAEPYR ***:*****	120 120
CSDaV AY884005	QSLSTYPWHIPTNLQPFLTSCGITTAGQGFKAHHPVHKTIETNLLSNVWSHYATTPSGV QSLSTYPWHIPTNLQPFLTSCGITTAGQGFKAHHPVHKTIETNLLTNVWPHYATTPSGV *****:***	180 180
CSDaV AY884005	MFMKPSKFEKLRKIQPNFSKLYNYRITAKDSTRYPSTSPDLPTEDTCFMDALMYSPGQ MFMKPSKFEKLRKIQPNFSKLYNYRITAKDTTRYPTSPDLPTEDTCFMDALMYSPGQ *****:*****:*****	240 240
CSDaV AY884005	ICDLFISRPSLQKLYASLVVPPESDFTTISLFPDLRYRIEKDQLVELEQNPAHNYIQP ICDLFISRPSLQKLYASLVVPPESDFTTISLFPDLRYRIEKDQLIYELEQNPAHNYIQP *****:*****:*****	300 300
CSDaV AY884005	RSAIDWLKTTTIRCQDLTLTISRLDVSWGPVHSLLIQRGRPIHLEEDSISFRTPKAVLLP RSAIDWLKTTTIRCQDLTLTISRLDVSWGPVHSLLIQRGKPIHLEEDSISFRAPKAVLLP *****:*****:*****	360 360
CSDaV AY884005	EPASLSQGVDRDLVPADVQALFIYVRAVRTLRVTDPAQFVRTQISKPEYSWVTSSFRWDN EPASLSQSVDRDLVPADVQALFIYVRAVRTLRVTDPAQFVRTQISKPEYSWVTSSAWDN *****:*****	420 420
CSDaV AY884005	FPPLALATAPHRPHTTYFLFNSTAARVSHWFRTHTLALLSGATAAAASLLMTASWGFRAL LAHFALATAPHRPHTTYFLFNSTAARVAHWFRTHTLAPLSGATAAAASLLMTASWGFRAM : :*****:*****:*****	480 480
CSDaV AY884005	ISSHLVSLSICKRWLRAPPHELLWPEKAPWFQLTLRPKVTGPMIDLPILRPFRFPSTCAK ISSHLVSLSICKRWLRAPPHELLWPEKAPWFQLTLRPKVTGPLIDLPILRPFRFPSTCAK *****:*****:*****	540 540
CSDaV AY884005	LGAKYPALATLLPAAPRPLWPKVGLALATVPVCLFLWRKFIGPDSQDMHDSYHAMFHP LGAKHPALATLLPAAPRPTWPKVGLALAAVPVCLFLWRKFIGPDSQDMHDSYHAMFHP ****:*****:*****	600 600
CSDaV AY884005	QPWGLSLARKAICCDRAPFLPIPLVPSSDFKAPPTPATPLLTSIPIKGVPEQVPEGVPP QPWGLTLTRKAICCDRAPFLPIPVVPSDFKAPPTPATPLLTSIPIKGVPEQVSGEGVPP *****:*****:*****	660 660
CSDaV AY884005	QSASSTGPASDRCAPQPASSTGPDQPAQDMSAAPQAPTKSKVTFAQPLESVTPVVPVGGAG QSASSTGPASDRRAPQPASSTGPDPTQNTSAAPQPPIESKVTFAQPIESVAPVVPVGGAG *****:*****:*****	720 720

```

*****
CSDaV      EPPQASSTGPASVSRDDPQTASSTVPDAPTLTSSVTPPPIIYPVDHLQNDFGPCRCSLC 780
AY884005   EPPQASSTGPASVSRDDPQVASSSTTPDAPTLTSSVTPPKTIYPIDHLQNDFGPCRCRSCV 780
*****

CSDaV      EPPQPAPVPSAPPTVSDHKESQDGEALASALQALGLAPAPPAPLSQNLTISSGAMHASC 840
AY884005   EPLQPAPVPSTPLTVSDHKEAQDAEALSSALQALGLAPTTPAPQSNLTVSSGAMHASS 840
** *****: * *****: * *****: * *****: * *****: * *****: *

CSDaV      WDQLSSPSSDWDPSVARDSSATGPPGMYSDLFPAPYLPGTGQFIFRSRANRANIPYPD 900
AY884005   WDQLSSPSSDWDPSPLARDSSASGPPGMYSDLFPAPYLPGTGQFIFRSRANRANIPYPD 900
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      MDCLLSIEQATRLPKEALWDTLTCATCPDSSLDPDTIRRVGLSTDHFVIFAHHYSRLCRF 960
AY884005   MDCLLSIEQATRLPKEALWDTLTCATCPDSSLDPDTIRRVGLSTDHFVIFAHHYSRLCRF 960
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      HTAHGVIELGMADATSSFDIDHTAGNPGHFSLRQSATPRLNGGIAQDLAVAALRFNIDGT 1020
AY884005   HTAHGVIELGMADATSSFDIDHTAGNPGHFSLRQSATPRLNGGIAQDLAVAALRFNIDGT 1020
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      LLPIRSVHVYSTWPKRAKNLSSNMKNGFDGIMANIHPKTNESREKILALDSQLDIARR 1080
AY884005   LLPIRSVHVYSTWPKRAKNLSSNMKNGFDGIMANIHPKTNESREKILALDSQLDIARR 1080
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      SVRLIHIAGFPGCGKSFPIRLLRTPTRFNFKVAVPTVELRAEWKAIIGLPTSEAWRIGT 1140
AY884005   SVRLIHIAGFPGCGKSFPIRLLRTPTRFNFKVAVPTVELRAEWKAITGLPASEAWRIGT 1140
*****: * *****:*****:*****:*****:*****:*****:*****

CSDaV      WESSLLKSARVLVIDEIKMGRYIDLAIHSDPTIEMVIALGDPLQGEYHSTHPSSTNSR 1200
AY884005   WESSLLKSARVLVIDEIKMGRYIDLAIHSDPTIEMVIALGDPLQGEYHSTHPSSTNSR 1200
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      LLAEPHLLSMYLDYCMWSHRVQNVAFFHIKTTSKQPFGCRYQRELPNSRILANSQNA 1260
AY884005   LLSEPHLLSMYLDYCLWSHRVQNVAFFHVKTTSKQPFGCRYQRELPNSRILANSQNA 1260
** : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

CSDaV      GHTLQCCGYAAVTIASSQGSTYDAAACIHLDRNSSLLSPAHSMAVALTRSKVGVIFTDGPA 1320
AY884005   GHTLQCCGYAAVTIASSQGSTYENAAACIHLDRNSSLLSPAHSMAVALTRSKVGVIFTDGPA 1320
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      QLSNAPSSNRMFSEFFSGRTRPLHDWFHNEFPKATVLTPELTKTRGPRLTGAASPYKAVP 1380
AY884005   QLSNAPSSNRMFSEFFSGRTRPLHDWFHNEFPKATVLTPELTKTRGPRLTGAASPYKAVP 1380
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      SGKDSTPALKPDFQGDVRSRAPIVLGSSELNAPQVSTHFLPETRRPLHWDIPSAIPESA 1440
AY884005   IRQASTPALKPDFQGDVRIIS-APIVLGSSELNAPQVSSHFLPETRRPLHWDIPSAIPESA 1439
: ***** * *****:*****:*****:*****:*****:*****:*****

CSDaV      TRPDSTEPTTCHPEPVYPGETFENLAAHFLPAHDPTDREIYWQGQLSNQFPMDKEFHLLA 1500
AY884005   TRPDSTEPTTSHPEPVYPGETFENLAAHFLPAHDPTDREIYWQGQLSNQFPMDKEFHLLA 1499
*****.*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      AQPMSLLAAIHQEKQDPTLLPASIQRRLRFRPSDRPYQITPKDEILGQLLFEGLCRAYHR 1560
AY884005   AQPMSLLAAVHQEKQDPTLLPASIQRRLRFRPSDKPYQITPKDEILGQLLFEGLCRAYHR 1559
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      SPFHTEAFDPVLFACINLNEFAQLSSKTQATIMGNARSDPDWRWSAVRIFSKTQHKVN 1620
AY884005   SPFHTEAFDPVLFACINLNEFAQLSSKTQATIMGNARSDPDWRWSAVRIFSKTQHKVN 1619
*****:*****:*****:*****:*****:*****:*****:*****

CSDaV      EGSIFGSKACQTLALMHDAVVLILGVPVKYQRFVDRDRPRHLYIHAGNTPSQMSNWCQ 1680
AY884005   EGSIFRSWKACQTLALMHDAVVLILGVPVKYQRFVDRDRPRHLYIHAGNTPSQMSNWCQ 1679
***** *****:*****:*****:*****:*****:*****:*****:*****

CSDaV      QHLTAVKLANDYTAFDQSQHGEAVVLERKKMERLIPQALIDLHIHLKTHVSTQFGLPT 1740
AY884005   QHLTAVKLANDYTAFDQSQHGEAVVLERKKMERLIPQALIDLHIHLKTHVSTQFGLPT 1739

```

```
*****
CSDaV          CMRLTGEPGTYDDNSDYNLAVINCEYMAANTPTMVSGDSSLDDREPPTREPWILQPLLS 1800
AY884005      CMRLTGEPGTYDDNSDYNLAVNCEYMAANTPTMVSGDSSLDDREPPTREPWILQPLLS 1799
               *****;*****
CSDaV          LRFKKERGRYATFCGYASHVGCVRSPVALFAKLAIAVDDGSISDKMASYLSEFALGHSL 1860
AY884005      LRFKKERGRYATFCGYASHVGCVRSPVALFAKLAIAVDDGSISDKMASYLSEFALGHSL 1859
               *****
CSDaV          GDHLWEALPLEAVPFQSAFCDFFCRRAPRHLKLSLMLGEVPESIIARIGSSLKWASHAIY 1920
AY884005      GDHLWEALPLEAVPFQSAFCDFFCRRAPRHLKLSLMLGEVPESIIARIGSSLKWASHAIY 1919
               *****
CSDaV          TTLSSAARVAILRSSRNSRSMPPDDPTLLQGELLQHFQVPFMQSDTLLPLTGGSSAPIL 1980
AY884005      TTLSSAARVAILRSSRNSRSMPPDDPTLLQGELLQHFQVPFMQSDTLLPLTGGSSAPIL 1979
               *****
CSDaV          TPEAFSTSLAFSMASDAQAGPAPSRDDRVRQPRLPAAPRVAEVLNAPSVDYPFQWVVA 2040
AY884005      TPEAFSTSLAFSMASDAQAGPAPSRDDRVRQPRLPAAPRVAEVLNAPSVDYPFQWVVA 2039
               *****
CSDaV          SYDGSEAKNLSDDLSGSATLTKVMANYRHAELTSVELEVCPAAAFSKPISVSAVWTIAS 2100
AY884005      SYDGSEAKNLSDDLSGSATLTKVMANYRHAELTSVELEVCPAAAFSKPISVSAVWTIAS 2099
               *****
CSDaV          ISPASASETSYGGRLFTVGGPVLMSSTTHLPADLTRLNPVLKGPVKYTDPCRFSYSVYS 2160
AY884005      ISPASASETSYGGRLFTVGGPVLMSSTTHLPADLTRLNPVLKGPVKYTDPCRFSYSVYS 2159
               *****
CSDaV          NGGTKGTNLCTIILRGVVRLSGPSGNLLA 2189
AY884005      NGGTKGTNLCTIILRGVVRLSGPSGNLLA 2188
               *****
```


APÊNDICE K**Two steps whole genome rescue of new Marafivirus associated with Citrus sudden death disease using GeneRacer and its genome characterization**

Barros, C. P. B.¹; Jesus Junior, W. C.²; Santos, M. A.³; Teixeira, D.³; Nagata, T.¹

¹ Graduate Program in Genomic Science and Biotechnology, Universidade Católica de Brasília, Campus II – SGAN 916 módulo B, Brasília 70790-160 DF, Brazil.

² Universidade Federal do Espírito Santo – UFES

³ Fundo de Defesa da Citricultura – Fundecitrus

Corresponding author:

Tatsuya Nagata

Graduate Program in Genomic Science and Biotechnology, Universidade Católica de Brasília, Campus II – SGAN 916 módulo B, Brasília 70790-160 DF, Brazil.

Phones: office: 55-61-34487213

Fax: 55-61-33474797

e-mail: tatsuya@ucb.br

Keywords: Citrus sudden death (CSD), citrus sudden death-associated virus (CSDaV), *Marafivirus*, GeneRacer.

Citrus sudden death (CSD) is a new disease of unknown etiology that affects sweet oranges (*Citrus sinensis*) and some mandarins (*C. reticulata*) grafted on Rangpur lime (*C. limonia*) or Volkamer lemon (*C. volkameriana*)(1, 2). The first report of this disease was published in 2001 (3), based on symptomatic plants identified in 1999 in the States of Minas Gerais and São Paulo, Brazil. The first observations of the disease were made in 'Valencia' sweet orange grafted on Rangpur lime in Comendador Gomes, Minas Gerais, Brazil (4). Later, the same disease were identified in other plantings of sweet oranges ('Hamlin', 'Natal', 'Westin', 'Pera', 'Rubi', 'Pineapple') and mandarins ('Cravo' and 'Pokan') grafted on Rangpur lime and Volkamer lemon.

The CSD is associated with the combination of sweet orange grafted on Rangpur lime or Volkamer lemon, since sweet orange grafted on mandarins 'Cleopatra' and 'Sunki' (*C. sunki* Hort. ex Tanaka), *Poncirus trifoliata* (L.) Raf. and citrumelo 'Swingle' [*P. trifoliata* (L.) Raf. X *C. paradise* Macf] does not present the symptoms of the disease (5).

The number of dead trees affected by the disease dramatically increased since it was detected from 500 trees in 1999 to more than 2 millions in 2004, that represents an estimated loss of 40 millions dollars. This number would be even greater if we considered those infected but still asymptomatic.

In the State of São Paulo, the largest sweet orange growing region in the world, 85% of 200 million orange trees are grafted on Rangpur lime, because this rootstock can be grown without irrigation (6). In view of the importance of Rangpur lime in Brazil, the discovery that sweet orange trees on this rootstock were rapidly declining and dying came as a shock (3). This new disease, which has taken on

alarming proportions (7), was called Citrus Sudden Death because of the rapid decline and death of the trees (8).

The causal agent of CSD is still unknown, but the most probable hypothesis is that CSD is caused for a virus disseminated by a vector, in the same manner of the quick-decline form of citrus tristeza (1), caused by *Citrus tristeza virus* (CTV, *Closteroviridae* family). The sudden appearance of CSD symptoms and the rapid disease progress have similarities with the citrus tristeza, quick-decline (10, 11). The anatomic modifications (6), spatial and temporal patterns (4, 2) of CSD dissemination has been remarkably similar to those seen with the spread of CTV. The similarities between the CTV decline and CSD lead researchers to suggest that the new disease is probably caused by a new strain of CTV, but several attempts to associate CSD with CTV mutant have failed (12, 13). Alternatively, a new virus of *Tymoviridae* family could be the causative agent of CSD, since it has been frequently detected in plants with CSD symptoms.

In the present study, we have cloned the complete genome of this new virus by the method of rapid amplification of cDNA ends (RACE), sequenced and compared with other tymoviruses.

The whole genome rescue of new *Marafivirus* associated with the CSD was made in two step of cloning: cloning a 3' end fragment of approximately 6000 bp and a 5' end fragment of an approximately 1000 pb generated by GeneRaceR technique (Figure 1).

From the P15 isolate, a virus of *Tymoviridae* family from infected trees in Comendador Gomes coutry, Minas Gerais State, Brazil, the RNA was extracted in accordance with the instruction of Trizol LS Reagent (Invitrogen). cDNA was

synthesized using primer oligo dT, M4T (14), and a transcriptase reverse, SuperScript II (Invitrogen), as the manufacture's instruction.

Degenerated primers designed on the basis of other marafivirus and tymoviruses combined with primer M4 (14) were used to amplify cDNA by 3' RACE, using *Platinum Taq DNA polymerase High Fidelity* (Invitrogen). A-overhang was added to the amplified fragment, which was ligated to a T-vector, *pCR4 TOPO TA sequencing vector* (Invitrogen) and electroporated into *Escherichia coli* DH5 α (Invitrogen).

The purified plasmid using Wizard Miniprep (Promega) was sequenced by dideoxynucleotide termination method and an automatic sequencer (Pekin Elmer ABI 377 Applied Biosystems).

The 5' end was rescued by 5' RACE with the technique GeneRacer (Invitrogen). In this technique the RNA was treated with tobacco acid pyrophosphatase (TAP) (15, 16) which removed the cap from capped RNA, exposing the 5' phosphate and permitting ligation of the GeneRacerTM RNA oligo, with T4 RNA ligase. A cDNA template was generated by reverse transcription using SuperScriptTM III and a specific primer (C7Rev: 5' - TCAGAGTTGTCGTCATAAGTGCC - 3'), which sequence was obtained by the sequencing of 3' end clone. 5' end cDNA were PCR amplified from these DNA templates with a primer to the GeneRacer RNATM oligo and the gene specific primer 1 (TymoGSP1: 5' - GAGCGGGGCTGAATGTAGTTGTG - 3').

The complete sequencing was obtained using vector universal primers and primers designed by primer walking. The sequences and eletropherograms of this shotgun sequencing were analyzed by Package Staden program. The search for homology in described sequences was made by Blast X (www.ncbi.nlm.nih.gov). The

multiple alignments were generated using ClustalW (www.ebi.ac.uk) and the phylogenetic study was made by neighbor-joining method (17) from Mega3 (www.megasoftware.net).

The complete CSDaV genome sequence was 6806 nucleotides in length, excluding the poly(A) tail. The poly(A) tail containing 63 adenosines was detected by synthesis of first strand cDNA primed with M4T (T₁₅) but its actual length remained undetermined.

As the other marafiviruses, the genome of CSDaV had a large single ORF encoding a polypeptide that contains the conserved signature motifs of the replication-associated proteins (methyltransferase, helicase and RdRp), a papain-like protease domain and the “marafibox” (18), a conserved 16 nucleotide stretch comparable to “tymobox”, from which differs by two or three residue changes (19). The marafibox is followed downstream by a transcription site and functions as subgenomic RNA promoter (20).

The single ORF was comprehended between 109 and 6678 nucleotides and encoded a protein of 2189 aminoacids with a predicted molecular weight of approximately 240 kDa (p240) (www.expasy.org/tools/pi_tool.html). Conserved signatures domains for methyltransferase (nucleotide position: 487 to 1119), papain-like protease (nucleotide position: 2797 to 3114), helicase (nucleotide position: 3358 to 4053), RdRp (nucleotide position: 4222 to 6030), replicase (nucleotide position: 5347 to 6675), and a coat protein (nucleotide position: 6030 to 6678) was found in searches for protein similarities by Pfam: (www.sanger.ac.uk/software/Pfam/serach.shtml) and InterPro (www.ebi.ac.uk/InterProScan).

The putative marafibox was present in CSDaV at positions 5959 to 5974, and a transcription initiation site is predicted at nucleotides 5983 to 5986. Excepted CSDaV that had a single discrepancy at the 5' end, where a guanidine was exchanged for an adenine, the sequence conservation among marafiviruses was complete in the promoter region.

Marafiviruses have two CP subunits that cistrons are located at the 3' end of the polypeptide, of which they are an integral part. The genome expression in *Marafivirus* genus is the result of the posttranslational autocatalytic cleavage of the largest ORF by a papain-like virus-encoded protease, and synthesis and translation of a 3'-co-terminal subgenomic RNA for a predicted 21 kDa CP expression (nucleotide position: 6084 to 6678) (21, 22, 23).

Comparing the CSDaV nucleotide sequence (DQ185573) with reported sequence available on GenBank (AY884005) the identity found was 89%. This number grows up to 96% in aminoacids analyses (ClustalW).

The highest levels of identity observed between the genomic sequence of the CSDaV and other tymovirus were to *Oat blue dwarf virus* (OBDV) with 63% and to *Maize rayado fino virus* (MRFV) with 60%. Phylogenetic analyses of the CSDaV genomic sequence revealed that CSDaV was closer to *Marafivirus* genus than others virus of *Tymoviridae* family (Figure 2).

A comparison of CP domain of CSDaV to corresponding sequence from other virus within *Tymoviridae* family confirmed the similarity to *Marafivirus*. The nucleotide sequence of the CP region of the virus had identity of 71% with *Grapevine asteroid mosaic-associated virus* (GAMaV), 71% with OBDV and 57% with MRFV. However, phylogenetic tree revealed that CSDaV was more related to GAMaV than OBDV (Figure 3).

In both comparisons, identities were higher between CSDaV and members of the genus *Marafivirus* (GAMaV, OBDV and MRFV) than with the representatives of the other genus of the family. These results support the classification of CSDaV as a new member of the genus *Marafivirus*.

The complete CSDaV nucleotide sequence is available on GenBank and the nucleotide sequence accession number is DQ185573.

Acknowledgements

We thank Fundecitrus (Fund for Citrus Plant Protection) that gave experimental material and supported financially this work.

References

- 1) Jesus Junior, W. C. De & Bassanezi, R.B. Análise da dinâmica e estrutura de focos da morte súbita dos citros. *Fitopatologia Brasileira*. 29:399-405. 2004.
- 2) Maccheroni, W., Alegria, M. C., Greggio, C. C., Piazza, J. P., Kamla, R.F., Zacharias, P. R. A., Bar-Joseph, M., Kitajima, E. W., Assumpção, L. C., Camarotte, G., Cardozo, J., Casagrande, E. C., Ferrari, F., Franco, S. F., Giachetto, P.F., Girasol, A., Jordão Junior, H., Silva, V. H. A., Souza, L. C. A., Aguilar - Vildoso, C. I., Zanca, A. S., Arruda, P., Kitajima, J. P., Reinach, F. C., Ferro, J.A. & Silva, A. C. R. Da. Identification and Genomic characterization of a new vírus (*Tymoviridae* Family) associated with citrus sudden death disease. *Journal of Virology*. v.79, n.5, p.3028-3037. 2005.
- 3) Gimenes-Fernandes, N. & Bassanezi, R. B. Doença de causa desconhecida afeta pomares cítricos no norte de São Paulo e Sul do Triângulo Mineiro. *Summa Phytopatologica*. 27:93. 2001.
- 4) Bassanezi, R. B., Bergamin Filho, A., Amorim, L., Gimenes-Fernandes, N., Gottwald, T. R. & Bové, J. M. Spatial and temporal analysis of citrus sudden death as a tool to generate hypotheses concerning its etiology. *Phytopathology*. 93:502-512. 2003.
- 5) Gimenes-Fernandes, N., Bassanezi, R. B., Yamamoto, P. T. & Ayres, A. J. Árvores de limoeiro 'Cravo' não mostram sintomas de "morte súbita" em área afetada. *Summa Phytopathologica*. 28:72. 2002. (Abstract).

- 6) Román, M. P., Cambra, M., Juárez, J., Moreno, P., Duran-Vila, N., Tanaka, F. A. O., Alves, E., Kitajima, E. W., Yamamoto, P. T., Bassanezi, R. B., Teixeira, D. C., Jesus Junior, W. C., Ayres, A. J., Gimenes-Frnandes, N., Rabenstein, F., Giroto, L. F. & Bové, J. M. Sudden Death of citrus in Brazil: A graft-transmissible bud union disease. *Plant Disease*. 88:453-467. 2004.
- 7) Muller, G. W., de Negri, J. D., Aguilar-Vildoso, C. I., Mattos Junior, D., Pompeu Junior, J., Teófilo Sobrinho, J., Carvalho, S. A., Giroto, L. F. & Machado, M. A. Morte súbita dos citros: uma nova doença na citricultura brasileira (Citrus sudden death: a new disease in the Brazilian citriculture). *Laranja*. 23:371-386. 2002.
- 8) Informativo Centro de Citricultura. Nova doença já causa prejuízos a pomares de SP e MG. *Informativo Centro de Citricultura*. 76:1-4. 2001.
- 9) Bassanezi, R. B., Yamamoto, P. T. & Gimenes-Fernandes, N. Progresso dos sintomas de "morte súbita em pomares de laranjeiras 'Valência' e 'Pera'". *Summa Phytopathologica*. 28:23. 2002. (Abstract).
- 10) Bar-Joseph, M., Marcus, R., and Lee, R. F. The continuous challenge of citrus tristeza virus control. *Annual review of phytopathology*. 27:292-316. 1989.
- 11) Rocha-Peña, M. A., Lee, R. F., Lastra, R., Niblett, C. L., Ochoa-Corona, F. M., Garnsey, S. M., & Yokomi, R. K. Citrus tristeza virus and its aphid vector

- Toxoptera citricida*. Threats to citrus production in the Caribbean and Central and North America. *Plant Disease*. 79:437-445. 1995.
- 12) Derrick, K. S., Berreta, M. J., Barthe, G. A. & Kaymin, M. Identification of strains of *Citrus tristeza virus* by subtraction hybridization. *Plant Disease*. 87:1355-1359. 2003.
- 13) Targon, M. I. P. N., Astaua-Monge, G., Freitas-Astaua, L. Kishi, Souza, A. A., Santos, F. A., Muller, G. W. & Machado M. A. Evaluation of CTV haplotypes in plants exhibiting citrus sudden death symptoms. 10th ISC Congress, abstract P83. 2004.
- 14) Chen, J., Chen, J. & Adams, M.J. A universal PCR primer to detect members of the *Potyviridae* and its use to examine the taxonomic status of several members of the family. *Archives of Virology* 146, 757-766. 2001
- 15) Maruyama, K. & Sugano, S. Oligo-Capping: A simple method to replace the cap structure of eucaryotic mRNAs with oligorribonucleotides. *Gene* 138, 171-174. 1994.
- 16) Schaefer, B. C. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Analytical Biochemistry*. 227, 255-273. 1995.
- 17) Page, R.D.M., & Holmes, E.C. *Molecular evolution: a phylogenetic approach*. Blackwell Science, Oxford, United Kingdom. 1998.

- 18) Ding S., Howe J., Keese P., Mackenzie A., Meek D., Osorio-Keese M., Skotnicki M., Pattana S., Torronen M. & Gibbs A. The tymobox, a sequence shared by most *tymoviruses*: its use in molecular studies of *tymoviruses*. *Nucleic Acids Research*. 18: 181–1187. 1990.
- 19) Martelli, G. P., S. Sabanadzovic, N. Abouh-Ghanem Sabanadzovic, and M. C. Edwards, and T. Dreher.. The family *Tymoviridae*. *Archives of Virology*. 147:1837–1846. 2002.
- 20) Izadpanah K, Zhang YP, Daubert S, Rowhani A. Sequence of the coat protein gene of bermuda grass etched-line virus and of the adjacent “marafibox” motif. *Virus Genes* 24: 131–134.2002.
- 21) Dreher T.W., Haenni A.L. & Mayo M.A. Genus *Tymovirus*. In: Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens E, Estes MK, Lemon S, Maniloff J, Mayo MA, McGeoch D, Pringle CR, Wickner RB (eds) *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego, 964–969. 2000.
- 22) Edwards M. C. Genus Marafivirus. In: Van Regenmortel M.H.V., Fauquet M., Bishop D.H.L., Carstens E., Estes M.K., Lemon S., Maniloff J., Mayo M.A., McGeoch D., Pringle C.R., & Wickner R.B. (eds) *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego, 969–973. 2000.

- 23) Hammond RW, Ramirez P. Molecular characterization of the genome of *Maize rayado fino virus*, the type member of the genus *Marafivirus*. *Virology* 282: 338–347. 2001.

Figure Legends

Figure 1. Electrophoresis of DNA fragment amplified by RT-PCR **A)** with the *GeneRacerTM 5' nested* and TymoGSP1 primers. (1) 1Kb plus ladder (Invitrogen); (2) 1000 pb fragment amplified in this reaction **B)** with C7For (5'-CATGCATGACGCCCTSATGTA-3') and M4 (5'-GTTTTCCCAGTCACGAC-3') primers; (1) 1Kb plus ladder (Invitrogen); (2) 6000 pb fragment amplified in this reaction.

Figure 2. Phylogenetic analysis of CSDaV genomic sequence with related tymoviruses. CSDaV, Citrus sudden death-associated virus (DQ185573); OBDV, *Oat blue dwarf virus* (U87832); MRFV, *Maize rayado fino virus* (AF265566); APLV, *Andean potato latent virus* (AF035402); GFkV, *Grapevine fleck virus* (AJ309022); TYMV, *Turnip yellow mosaic virus* (J04373).

Figure 3. Phylogenetic analysis of CP nucleotide sequence of CSDaV with related tymoviruses. Numbers are the bootstrap values (%) of 1000 replicates. CSDaV, Citrus sudden death-associated virus (DQ185573); GAMaV, *Grapevine asteroid mosaic-associated virus* (CAC10493); OBDV, *Oat blue dwarf virus* (U87832); MRFV, *Maize rayado fino virus* (AF265566); APLV, *Andean potato latent virus* (AF035402); GFkV, *Grapevine fleck virus* (AJ309022); TYMV, *Turnip yellow mosaic virus* (J04373). A, B and C showed the clustering of virus in *Marafivirus*, *Maculavirus* and *Tymovirus*, respectively.

Figure 1

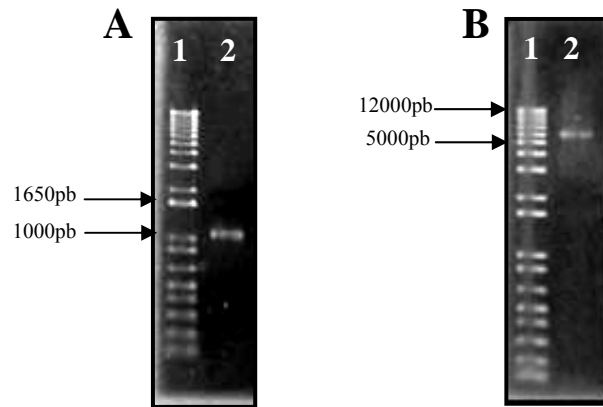


Figure 2

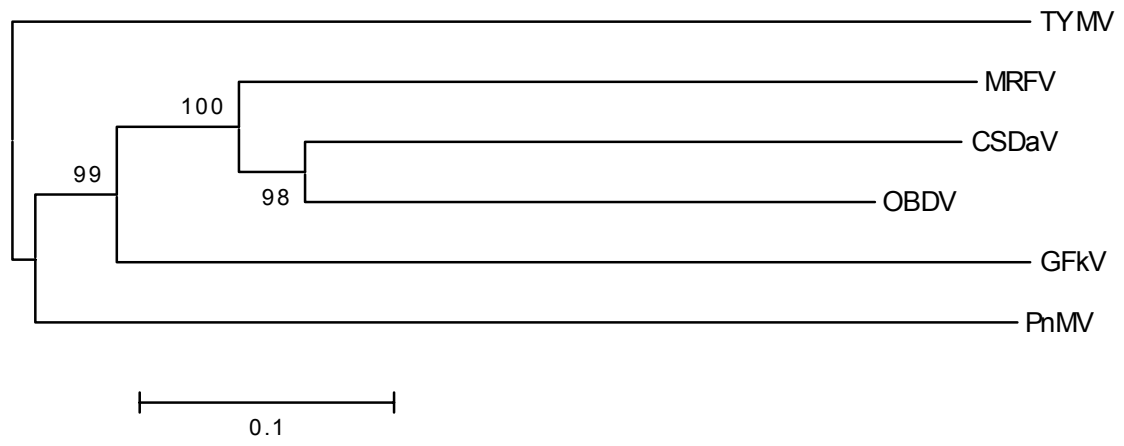
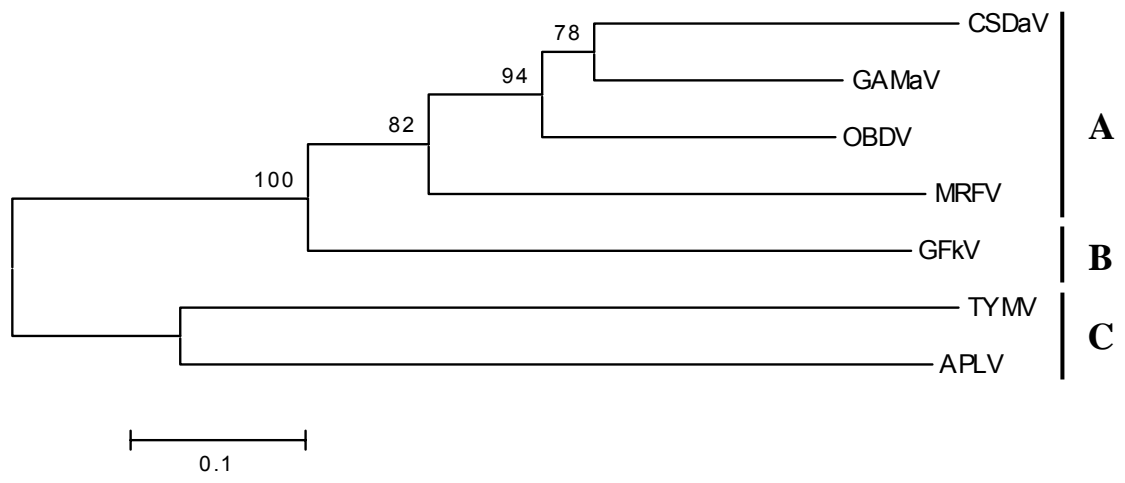


Figure 3



Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)