

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA GOIANO
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM OLERICULTURA

ANÁLISE COMPUTACIONAL: DIVERSIDADE GENÉTICA
DE *Tomato chlorosis virus* INFECTANDO TOMATEIRO

Autor: Luis Henrique Boff
Orientador: Prof. Dr. Leonardo Cunha de Albuquerque

MORRINHOS – GO
2018

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA GOIANO
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM OLERICULTURA

ANÁLISE COMPUTACIONAL: DIVERSIDADE GENÉTICA
DE *Tomato chlorosis virus* INFECTANDO TOMATEIRO

Autor: Luis Henrique Boff

Orientador: Prof. Dr. Leonardo Cunha de Albuquerque

Dissertação apresentada, como parte das exigências para obtenção do título de MESTRE EM OLERICULTURA, no Programa de Pós-Graduação em Olericultura do Instituto Federal de Educação, Ciência e Tecnologia Goiano – Campus Morrinhos, Área de Concentração Olericultura.

Sistema desenvolvido pelo ICMC/USP
Dados Internacionais de Catalogação na Publicação (CIP)
Sistema Integrado de Bibliotecas - Instituto Federal Goiano

B673a Boff, Luis Henrique
Análise computacional: diversidade genética de
Tomato chlorosis virus infectando tomateiro / Luis
Henrique Boff;orientador Leonardo Cunha de
Albuquerque. -- Morrinhos, 2018.
35 p.

Dissertação (Mestrado em Programa de Pós-Graduação
Mestrado Profissional em Olericultura) -- Instituto
Federal Goiano, Campus Morrinhos, 2018.

1. Variabilidade genética. 2. ToCV. 3. Solanum
lycopersicum. I. Albuquerque, Leonardo Cunha de,
orient. II. Título.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA GOIANO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM OLERICULTURA

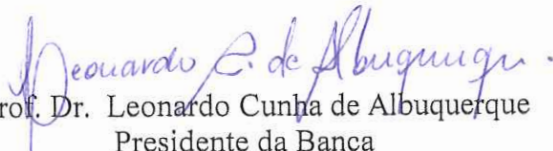
ANÁLISE COMPUTACIONAL: DIVERSIDADE DE *Tomato
chlorosis virus* INFECTANDO TOMATEIRO

Autor: Luís Henrique Boff

Orientador: Leonardo Cunha de Albuquerque

TITULAÇÃO: Mestre em Olericultura-Área de Concentração em Manejo
Fitossanitário em Olerícolas.

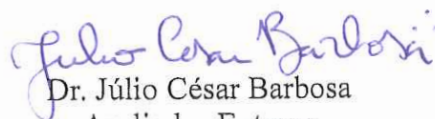
APROVADO em 25 de setembro de 2018.



Prof. Dr. Leonardo Cunha de Albuquerque
Presidente da Banca
IF Goiano – Campus Morrinhos



Prof. Dr. Nadson de Carvalho Pontes
Avaliador Interno
IF Goiano – Campus Morrinhos



Dr. Júlio César Barbosa
Avaliador Externo
Bayer S.A. - Crop Science Division

AGRADECIMENTOS

Ao Professor Dr. Leonardo Cunha de Albuquerque, por toda compreensão, dedicação, orientação, paciência, partilha de conhecimentos e prontidão.

Ao Instituto Federal Goiano, Embrapa Hortaliças, Instituto Federal Catarinense e aos profissionais destas instituições envolvidos nas atividades do Programa de Pós-Graduação em Olericultura.

Às colegas de pesquisa no laboratório de Virologia Vegetal (Fitopatologia) Ariela Rizo, Helena Mota, Laysla Coêlho, Tânia Carneiro e Thayssa Oliveira, pela acolhida e disponibilidade em compartilhar conhecimentos teóricos e práticos; à Sr.^a Maria do Carmo Rocha por toda gentileza e receptividade.

À Prefeitura municipal de Concórdia (Secretaria de agricultura), Epagri (Empresa de pesquisa agropecuária e extensão rural de Santa Catarina) e aos profissionais destas instituições que contribuíram para obtenção de informações e materiais.

Aos membros da banca, Dr. Júlio César Barbosa e Prof. Dr. Nadson de Carvalho Pontes, pela disponibilidade em avaliar e contribuir com este trabalho.

Aos familiares por todo auxílio, compreensão e incentivo.

A todos os colegas de pós-graduação pelo apoio, companheirismo e troca de conhecimentos e experiências.

A todos que de alguma forma contribuíram para a consecução deste trabalho.

Muito Obrigado!

BIOGRAFIA DO AUTOR

LUÍS HENRIQUE BOFF, filho de Eusébio Luis Boff e Neiva Lorensetti Boff, nasceu no ano de 1990, em Concórdia, Santa Catarina. De 2005 a 2007 cursou Técnico em Alimentos concomitante ao Ensino Médio, na Escola Agrotécnica Federal de Concórdia. De 2009 a 2011 cursou bacharelado em Engenharia da Computação na Universidade do Oeste de Santa Catarina, interrompido por motivos profissionais. De 2011 a 2013 cursou graduação em Análise e Desenvolvimento de Sistemas, no Instituto de Desenvolvimento Educacional do Alto Uruguai, tendo participado de atividades envolvendo computação científica (*cluster* e software livre). Em 2016 ingressou no Programa de Pós-graduação em Olericultura – Mestrado Profissional, no Instituto Federal Goiano – Campus Morrinhos, desenvolvendo atividades de pesquisa com virologia vegetal e computação científica (bioinformática), tendo defendido dissertação em 25 de setembro de 2018.

ÍNDICE

	Página
RESUMO.....	vii
ABSTRACT.....	ix
1. INTRODUÇÃO GERAL.....	1
2. REVISÃO DE LITERATURA.....	2
2.1. Cultura do tomateiro.....	2
2.2. Viroses.....	3
2.3. <i>Tomato chlorosis virus</i> (ToCV).....	4
2.4. Análise computacional de genes e genomas.....	7
2.5. Glossário.....	9
2.6. Referências.....	11
3. CAPÍTULO I.....	17
3.1. Introdução.....	18
3.2. Material e Métodos.....	19
3.3. Resultados e Discussão.....	21
3.4. Referências.....	23
4. CONCLUSÃO GERAL.....	31
APÊNDICES.....	32
Apêndice 1. Isolados correspondentes às sequências nucleotídicas das proteínas P22, RdRp e Hsp70h de <i>Tomato chlorosis virus</i> obtidos de amostras coletadas no Distrito Federal (DF) e no estado de Goiás (GO), Brasil.....	33
Apêndice 2. Sequências nucleotídicas das proteínas P22, RdRp e Hsp70h de isolados de <i>Tomato chlorosis virus</i> obtidas no <i>GenBank</i>	34

LISTA DE TABELAS

Tabela 1: Estrutura genética de populações de <i>Tomato chlorosis virus</i> , genes RdRp, P22 e Hsp70h.	27
Tabela 2. Diferenciação genética entre populações de <i>Tomato chlorosis virus</i>	27
Tabela 3. Teste da hipótese de neutralidade genética de cada proteína (RdRp, P22 e Hsp70) entre as populações.	27
Tabela 4. Relação entre substituições sinônimas e não sinônimas (d_{NS}/d_S) entre as sequências nucleotídicas dos genes das proteínas P22, Hsp70h e RdRp nas populações de <i>Tomato chlorosis virus</i>	27

LISTA DE FIGURAS

2.REVISÃO DE LITERATURA.....	2
Figura 1. Sintomas de Amarelão do tomateiro causado por ToCV: clorose internerval, deformação e necrose foliar	5
Figura 2. Diagrama da estrutura genômica do ToCV. Linhas representam a fita de RNA e os quadros as proteínas constituintes (ORFs) (Albuquerque et al., 2013).	6
Figura 3. Fluxograma das etapas que os dados (códigos genéticos dos isolados virais) passam até se tornarem conhecimento. Para obtenção de parâmetros descritores e árvores filogenéticas utilizam-se ferramentas computacionais (softwares) distintas.	9
3.CAPÍTULO I	17
Figura 1. Árvores filogenéticas construídas por inferência bayesiana, baseadas nas sequências nucleotídicas de isolados de <i>Tomato chlorosis virus</i> (ToCV) deste estudo (identificação na árvore inicia com ToC, provenientes do Brasil) comparados a isolados da China, Coreia do Sul, Espanha, Grécia e Estados Unidos. A barra de escala abaixo representa o número de substituições nucleotídicas por sítio. Cada árvore representa a análise de uma proteína: RdRp (A), P22 (B), Hsp70h (C).	28

RESUMO

BOFF, LUIS HENRIQUE BOFF. Instituto Federal Goiano – Campus Morrinhos, setembro de 2018. **Análise computacional: diversidade genética de *Tomato chlorosis virus* infectando tomateiro.** Orientador: Prof. Dr. Leonardo Cunha de Albuquerque.

Tomato chlorosis virus (ToCV) (gênero *Crinivirius*, família *Closteroviridae*) infecta cultivos de tomateiro (*Solanum lycopersicum*) no Brasil e em outros países, reduzindo a produtividade. Sintomas visíveis são manchas cloróticas e necrose das folhas. A transmissão ocorre pelas espécies de mosca *Bemisia tabaci*, *B. argentifolii*, *Trialeurodes vaporariorum* e *T. abutilonea*, de forma semipersistente. ToCV possui genoma bipartido, RNA1 (4 ORF) (*open reading frames*) e RNA2 (9 ORF). Objetivou-se adequar o uso de ferramentas de bioinformática, utilizando sequências genéticas das ORFs RdRp, P22 e Hsp70h de isolados de ToCV para determinar a variabilidade genética e obter árvores filogenéticas. As sequências foram obtidas de 16 isolados provenientes de amostras foliares de tomateiro coletadas no Brasil, em Brasília (Distrito Federal, n=7) e Goianápolis (Goiás, n=9), junto a sequências de 43 isolados obtidas no banco de dados *GenBank*. Quatro populações (Brasil n=17, China n=16, Coreia do Sul n=9 e Espanha n=15) foram analisadas usando métodos (*Muscle*, *ModelTest*, *Inferência Bayesiana* e *SLAC*) e ferramentas de bioinformática (*Mega*, *Paup*, *MrBayes*, *Figtree*, *DNAsp* e *DataMonkey*) para obtenção de descritores de variabilidade genética e árvores filogenéticas. Os resultados indicam que dentre as regiões genômicas dos isolados analisados, a RdRp apresenta maior diferenciação genética e a P22 apresenta maior

diversidade nucleotídica e pressão de seleção, comparadas às demais ORFs analisadas. Os isolados brasileiros analisados apresentam menor diversidade nucleotídica e maior neutralidade genética quando comparados às demais populações analisadas. Os isolados brasileiros analisados apresentam possível proximidade evolutiva com alguns isolados da Coreia do Sul.

PALAVRAS-CHAVE: Variabilidade genética; ToCV; *Solanum lycopersicum*.

ABSTRACT

BOFF, LUIS HENRIQUE BOFF. Goiano Instituto Federal – Campus Morrinhos, September 2018. **Computational analysis: genetic diversity of *Tomato chlorosis virus* infecting tomato plant.** Advisor: Prof. Dr. Leonardo Cunha de Albuquerque.

Tomato chlorosis virus (ToCV) (genus *Crinivirus*, family *Closteroviridae*) infects tomato (*Solanum lycopersicum*) crops in Brazil and in other countries, reducing productivity. Visible symptoms are chlorotic spots and leaf necrosis. Transmission occurs by *Bemisia tabaci*, *B. argentifolii*, *Trialeurodes vaporariorum* and *T. abutilonea* flies, in a semipersistent way ToCV has bipartite genome, RNA1 (4 ORF) (open reading frames) and RNA2 (9 ORF). The aim of this study was to adapt the use of bioinformatics tools using genetic sequences from the RdRp, P22 and Hsp70h ORFs of ToCV isolates to determine genetic variability and obtain phylogenetic trees. Sequences were obtained from 16 isolates from tomato leaf samples collected in Brazil, in Brasília (Federal District, n = 7) and Goianópolis (Goiás, n = 9), with sequences from 43 isolates obtained from the GenBank database. Four populations (Brazil n = 17, China n = 16, South Korea n = 9 and Spain n = 15) were analyzed using methods (*Muscle*, *ModelTest*, *Bayesian Inference* and *SLAC*) and bioinformatics tools (Mega, Paup, Figtree, DNAsp and DataMonkey) to obtain descriptors of genetic variability and phylogenetic trees. The results indicate that among the genomic regions of the analyzed isolates, RdRp presents higher genetic differentiation and P22 shows higher nucleotide diversity and selection pressure, compared to the other ORFs analyzed. The analyzed

Brazilian isolates showed lower nucleotide diversity and higher genetic neutrality when compared to the other analyzed populations. The analyzed Brazilian isolates present a possible evolutionary proximity with some South Korean isolates.

KEY WORDS: Genetic variability; ToCV; *Solanum lycopersicum*.

1. INTRODUÇÃO GERAL

A cultura do tomateiro tem notável relevância na produção de hortaliças no Brasil, todavia doenças e pragas limitam o aumento da produtividade e causam severas perdas. As viroses transmitidas por mosca-branca estão entre as principais doenças do tomateiro. Patologias causadas por vírus do gênero *Crinivirus* têm surgindo como fator limitante à produção de tomate em diversos países (Tzanetakis et al., 2013).

No Brasil, o único *Crinivirus* infectando tomateiro ora relatado é a espécie *Tomato chlorosis virus* (ToCV), descrito em 2006 e 2007 no estado de São Paulo (Barbosa et al., 2008), posteriormente na Bahia, Distrito Federal, Espírito Santo, Goiás, Minas Gerais, Rio de Janeiro (Barbosa et al., 2013) e Paraná (Coêlho, 2017). A variabilidade dentre isolados de ToCV em amostras de tomate pode ser caracterizada por alta heterogeneidade e complexidade genéticas das populações virais (Lozano et al., 2009).

Objetivou-se adequar o uso de ferramentas de bioinformática, utilizando sequências genéticas das ORFs RdRp, P22 e Hsp70h de ToCV, para determinar a variabilidade genética entre isolados e obter árvores filogenéticas. Sequências genéticas de isolados de ToCV provenientes de amostras de folhas de tomateiro coletadas em Brasília, Distrito Federal e Goianópolis, Goiás (Brasil) foram comparadas as sequências de isolados de ToCV disponíveis no banco de dados *GenBank* (Clark et al., 2016).

Estudos de variabilidade genética e estruturas populacionais permitem alguma avaliação evolutiva da espécie e consequente acompanhamento da doença viral, verificando-se eventual estabilidade ou ocorrência de surtos decorrentes de maior mutabilidade viral, podendo subsidiar estratégias de manejo mais adequadas.

2. REVISÃO DE LITERATURA

2.1. Cultura do tomateiro

O tomate é uma hortaliça de importância econômica com amplas possibilidades de uso *in natura* ou industrializado e aceitação pelos mais variados perfis de consumidores (Filgueira, 2003). A produção de tomate no Brasil na safra de 2017 foi de aproximadamente 4,3 milhões (mi) de toneladas (t), em pouco mais de 64 mil hectares (ha), rendendo aproximadamente 67 t.ha⁻¹; foi a terceira hortaliça mais cultivada no país após a mandioca (20,6 mi t; 1,4 mi ha) e batata-inglesa (4,2 mi t; 140 mil ha) (IBGE, 2018). Em 2016, foi a nona maior produção mundial (precedido por China, Índia, Estados Unidos, Turquia, Egito, Itália, Irã e Espanha) (FAO, 2018).

As plantas ancestrais do tomateiro (*Solanum lycopersicum*) são originárias da América do Sul, de região situada entre o norte do Chile e sul da Colômbia, em meio a Cordilheira dos Andes e o Oceano Pacífico. Espécies com semelhanças genéticas são encontradas na Bolívia, Colômbia, Peru e Equador, inclusive Ilhas Galápagos (Filgueira, 2003; Peralta et al., 2005).

Mesmo não sendo cultivadas por indígenas andinos (*Aymarás*, Incas, *Quéchuas*), plantas silvestres chegaram à América Central e sul da América do Norte em migrações e intercâmbios das civilizações pré-colombianas. Provavelmente foi domesticado no México (Hanssen & Lapidot, 2012), passando por melhoramento genético mediante seleção e cruzamentos conforme características desejáveis. Quando os conquistadores espanhóis chegaram na região, o tomate já era plantado, consumido e comercializado pelos Astecas e povos vizinhos (Filgueira, 2003).

Introduzido na Espanha no início do século (séc.) XVI, o tomate inicialmente servia como planta ornamental e medicinal (Giordano & Ribeiro, 2000), não sendo consumido como alimento por temor da toxicidade presente em várias plantas da família das solanáceas. Assim, seu cultivo e consumo como hortaliça foi retardado até o final do séc. XVIII. No Brasil, a cultura foi introduzida por imigrantes europeus no final do séc. XIX (Filgueira, 2003).

A palavra tomate muito provavelmente originou-se de “*tumatl*”, termo do idioma do povo Asteca (*nauatl*) (Giordano & Ribeiro, 2000). Conforme o idioma, recebe diferentes denominações: tomate (português, espanhol, francês), *tomat* (indonésio), *faan ke'e* (chinês), *tomati* (africano ocidental), *tomatl* (*nauatle*), *jitomate* (espanhol mexicano), *nyanya* (*swahili*) (Naika et al., 2006). A palavra italiana *pomodoro* é exceção, pois nesta região as primeiras plantas introduzidas produziam frutos redondos alaranjados (*pomo d'oro*) (Giordano & Ribeiro, 2000; Filgueira, 2003).

O tomateiro foi classificado como pertencente ao gênero *Solanum* em 1753 (Linnaeus), reclassificado no gênero *Lycopersicon* em 1788 (Miller). Posteriormente a classificação inicialmente proposta por Linnaeus foi adotada na comunidade científica, considerando possibilidades de cruzamento e características morfológicas e moleculares (Lima, 2015; Ganaie et al., 2018).

Quando maduro o tomate (fruto do tomateiro) contém em média 94% de água e cerca de 6% de matéria seca (açúcares, proteínas, minerais, vitaminas e demais nutrientes). Tais valores são suscetíveis a variações conforme as cultivares e condições (climáticas, manejo, nutricionais, etc.) sob as quais a cultura se desenvolve (Filgueira, 2003; Nepa, 2011).

2.2. Viroses

Vírus são parasitas intracelulares compostos por ácidos nucleicos (DNA ou RNA) envoltos por proteínas (Lodish et al., 2014), têm propriedades de hereditariedade e mutação genética (Robertis & Hib, 2017), manifestadas quando da subversão de pelo menos parte dos mecanismos genéticos e estruturas celulares das células hospedeiras (Brown, 1999), replicando-se utilizando as organelas destas células (Junqueira & Carneiro, 2005).

Frequentes e intensivos ataques de doenças causadas por vírus podem comprometer a produtividade do tomateiro e causar severos prejuízos (Filgueira, 2003).

Estas doenças podem paralisar o crescimento da planta e levá-la à morte; em geral os sintomas são complexos, podendo apresentar folhas com mosaico, clorose, necrose e ou deformação foliar; hastes com descoloração, necrose e/ou encurtamento dos entrenós; frutos com mosaico, anéis cloróticos ou necróticos, manchas de necrose e deformações (Nagata et al., 2005; Ramalhete et al., 2007).

Em razão de sua natureza sistêmica, a planta doente não pode ser recuperada, e não existem “viricidas”, produtos químicos que combatam especificamente o vírus (Filgueira, 2003; Lima, 2015). As viroses são doenças de complexo e difícil controle, que depende principalmente de medidas preventivas que minimizem o contato com vírus. São exemplos de medida de manejo plantio em áreas com certa distância de lavouras mais velhas, rotação com culturas não suscetíveis, plantio em áreas de maior altitude, plantio de mudas sadias (produzidas em estufas protegidas, utilizando sementes tratadas), controle de vegetação espontânea hospedeiras do vírus, aplicação de inseticidas para diminuir a população dos vetores (insetos transmissores do vírus entre as plantas) (Colariccio & Vicente, 2000; Lima, 2015).

A utilização de cultivares com resistência genética é considerado o mais eficiente meio de controle a determinados vírus (Filgueira, 2003; Lima, 2015; Abreu et al., 2016), quando disponível cultivar com maior aptidão à região de plantio e resistência às doenças que lá ocorrem. No desenvolvimento de cultivares com tais propriedades demandam-se genes capazes de fornecer resistência à doença, buscados em plantas com semelhanças genéticas (García-Cano et al., 2010; Mansilla-Córdova et al., 2018).

2.3. *Tomato chlorosis virus* (ToCV)

O ToCV causa doença conhecida como Amarelão do tomateiro (Costa et al., 2009), provocando infecções sistêmicas a partir do floema da planta. Os sintomas muitas vezes confundem-se com deficiências nutricionais e compreendem manchas amareladas (cloróticas) irregulares, que se desenvolvem primeiro nas folhas do baixeiro, e gradualmente avançam em direção aos pontos de crescimento. Nas áreas amareladas internervosas desenvolvem-se manchas necróticas de cor vermelha ou marrom (Wisler et al., 1998) (Figura 1).



Figura 1. Sintomas de Amarelão do tomateiro causado por ToCV: clorose internerval, deformação e necrose foliar

É classificado no gênero *Crinivirus*, família *Closteroviridae*. Há duas espécies de *Crinivirus* que causam doença no tomateiro: *Tomato infectious chlorosis virus* (TICV) e *Tomato chlorosis virus* (ToCV), descritos nos Estados Unidos na década de 1990, com elevada semelhança sintomática e genética. TICV é transmitido somente pela espécie de mosca-branca *Trialeurodes vaporariorum*. ToCV também é transmitido pela *T. vaporariorum* e ainda por *T. abutilonea*, *Bemisia tabaci* e *B. argentifolii*, que transmitem o vírus a partir de outras plantas hospedeiras (Wisler et al., 1998; Wintermantel & Wisler, 2006). Ambos provocam manchas cloróticas internervosas, que iniciam nas folhas do baixeiro e podem evoluir para manchas necróticas e necrose das folhas, resultando em perdas produtivas.

O comprimento de partícula característico de ToCV fica entre 800 e 850 nm, tendo o genoma dividido em duas fitas de RNA (Wisler et al., 1998; Wintermantel & Wisler, 2006) (Figura 2). No RNA 1 há quatro ORFs (*open reading frames* – fases abertas de leitura) que codificam proteínas: ORF 1a: protease, metiltransferase e helicase, 221 kDa. ORF 1b: RdRp, “polimerase de RNA dependente de RNA”, 58 kDa.

ORF1a e ORF1b são associadas à replicação genômica, também presentes em outros vírus do gênero *Crinivirus* e da família *Closteroviridae* (Wintermantel et al., 2005; Karasev, 2000). ORF 2: P22, supressão de silenciamento gênico (Cañizares et al., 2008), 22 kDa. ORF 3: p5, localizada na terminação 3' do RNA 1, 5 kDa, domínio transmembrana (Wintermantel et al., 2005).

No RNA 2 há nove ORFs, que codificam proteínas relacionadas à encapsidação viral, movimento viral e transmissão pelo inseto vetor. São elas: ORF 1: p4, pequena proteína de 4 kDa com possível transmembrana. ORF 2: Hsp70h (*heat shock protein 70*), montagem e movimentação intercelular das partículas virais (cauda), presente em diversos vírus da família *Closteroviridae* e outros organismos (Martelli et al., 2009). ORF 3: p8, 8 kDa, função desconhecida. ORF 4: p59, 59 kDa, associada à cauda de partículas virais, possivelmente envolvida na movimentação intercelular. ORF 5: p9, 9 kDa, função desconhecida (Wintermantel et al., 2005). ORF 6: CP, 29 kDa, capa proteica, encapsidação de maior parte da partícula viral ORF 7: CPm, capa proteica menor, encapsidação da extremidade 5', 76 kDa (Satyanarayana et al., 2004; Alzhanova et al., 2001), possível envolvimento na transmissibilidade pelo vetor (Wintermantel & Wisler, 2006). CP e CPm envolvidas também na supressão do silenciamento gênico (Cañizares et al., 2008). ORF 8: p27, 27 kDa, função desconhecida, exclusiva dos *Crinivirus*. ORF 9: p7, 7 kDa, exclusiva do ToCV, domínio transmembrana) (Wintermantel et al., 2005).

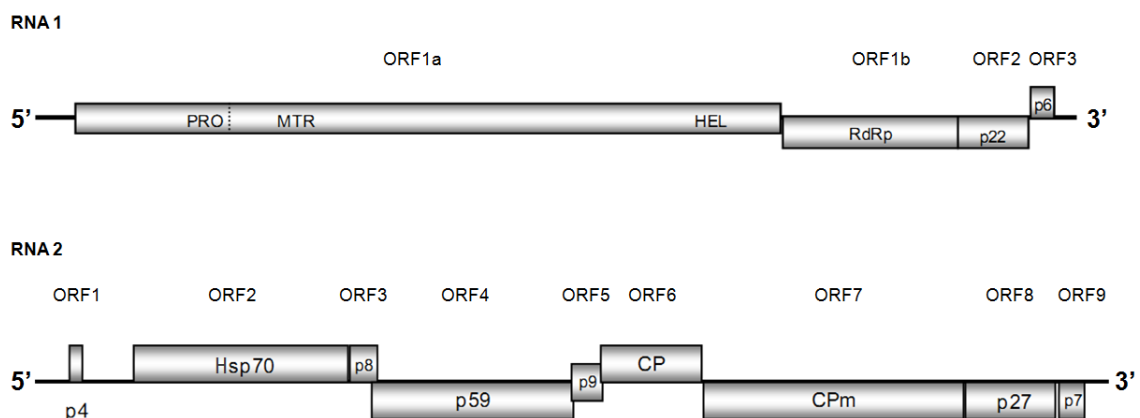


Figura 2. Diagrama da estrutura genômica do ToCV. Linhas representam a fita de RNA e os quadros as proteínas constituintes (ORFs) (Albuquerque et al., 2013).

Interações entre ToCV e TSWV (*Tomato spotted wilt virus*) pode diminuir a resistência do tomateiro a doença vira-cabeça do tomateiro (causada por TSWV) mesmo em plantas com resistência genética (García-Cano et al., 2006). Atenuar os danos através de fungicidas sistêmicos pode ser uma tentativa de controle (Guimarães et al., 2017). O uso do indutor de resistência *acibenzolar-S-methyl* (ASM) não apresentou efeitos significativos no controle desta doença (Vargas, 2016).

2.4. Análise computacional de genes e genomas

Amostras de plantas com sintomas de doenças viróticas podem ser analisadas por técnicas de biologia molecular, por exemplo PCR (reação em cadeia da polimerase), para detecção do agente causador. Uma vez identificado o vírus, é possível proceder o sequenciamento genético desta amostra, que pode auxiliar em estudos posteriores.

A partir de projeto para sequenciamento do genoma humano, iniciado em 1980 nos Estados Unidos da América, desenvolveram-se áreas de pesquisa que evoluem concomitantemente: a *bioinformática*, técnicas de aquisição, análise e armazenamento de dados sobre aspectos biológicos, em especial ácidos nucleicos e proteínas; e a *biologia molecular computacional*, desenvolvimento de algoritmos e softwares computacionais para obter informações significativas, bem como desenvolver e ou aperfeiçoar metodologias de trabalho para resolver problemas (Vasconcelos e Almeida, 2012), tendo-se maior precisão, em nível molecular e menor aleatoriedade no desenvolvimento de fármacos e agroquímicos (Jardine et al., 2014).

Em alguns casos, isso também possibilita contribuir para maior assertividade de programas de monitoramento e controle de doenças, manifestando-se através da *epidemiologia molecular*, área de conhecimento que essencialmente trata de dispor elementos da biologia molecular aos estudos epidemiológicos, bem como aplicar pesquisas e perspectivas epidemiológicas à interpretação das análises moleculares (Caldart et al., 2016).

Tais áreas compreendem conhecimentos multidisciplinares, de áreas como biologia, matemática, computação e estatística, buscando converter as informações contidas nas sequências genéticas (linguagem biológica) em conhecimentos bioquímicos e biofísicos (estruturas de proteínas e demais componentes, suas funcionalidades e mecanismos evolutivos). Consiste em tentar decodificar uma linguagem desconhecida (composta por *letras*: nucleotídeos e aminoácidos; *palavras*:

encadeamentos; *frases* – proteínas) para obter seu significado biológico, como verificar os impactos potencialmente consideráveis que alterações aparentemente insignificantes podem gerar no funcionamento do organismo (Vasconcelos & Almeida, 2012).

Pode ser feito um paralelo entre o código computacional e o código gênico. O primeiro, em tese foi desenvolvido por humanos, cuja programação demanda total compreensão da sintaxe e execução dos algoritmos. O segundo é inerente aos seres vivos, e por ora vem se tentando compreender os mecanismos de funcionamento do código, desconhecendo-se até então todos os pormenores de seu funcionamento.

Os notáveis progressos nas técnicas de sequenciamento têm possibilitando em tese estudos mais detalhados. Em razão da crescente quantidade de genomas, um desafio têm sido aplicar os recursos computacionais já existentes, mesmo que aperfeiçoados, em estudos sobre a dinâmica da atuação dos genes e dos genomas, para alcançar melhor entendimento de sua biologia (Vasconcelos & Almeida, 2012).

Desta forma, um conjunto de dados (sequências nucleotídicas) de interesse pode ser preparado e submetido a processamento em aplicativos computacionais que efetuam cálculos e fornecem parâmetros estatísticos, de modo buscar compreender quais mecanismos evolutivos estão atuando. Seleção positiva, seleção negativa, deriva genética, migração são as principais forças evolutivas atuantes em mutações (Lozano et al., 2009), bem como deriva genética e evolução neutra.

Observa-se clara relação de dados, processamento, informação e conhecimento. Na Figura 3, demonstra-se a sucessão de atividades que as sequências genéticas passam visando obtenção de informações. Tais dados tornar-se-ão conhecimento quando a partir destes forem feitas interpretações pelo autor ou por outrem, imediata ou posteriormente.

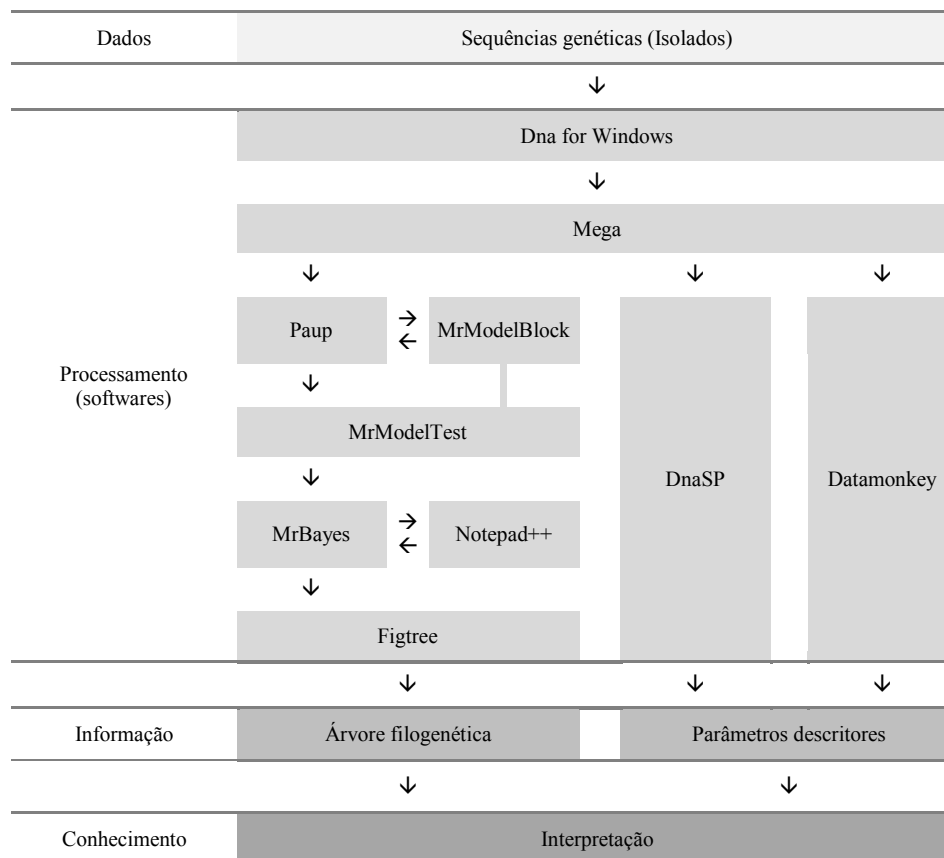


Figura 3. Fluxograma das etapas que os dados (códigos genéticos dos isolados virais) passam até se tornarem conhecimento. Para obtenção de parâmetros descritores e árvores filogenéticas utilizam-se ferramentas computacionais (softwares) distintas.

2.5. Glossário

Conhecimento: Entendimento obtido na inferência entre dados e informações, traduzindo a essência de qualquer elemento (Cruz, 2013).

Dado: parte fundamental do conhecimento, menor parte da informação, não pode ser entendido por seu conteúdo em si, somente tem significado quando a quem detiver o conhecimento necessário (Cruz, 2013).

Deriva genética: mutações aleatórias, absolutamente randômicas; quando ocorre troca de um aminoácido por outro em uma proteína, isso ocorre independentemente desta mudança vir a melhorar ou prejudicar a capacidade de funcionamento desta proteína. Não leva à adaptação, mas leva as mudanças nas frequências alélicas (Freeman & Herron, 2009).

DNA: Ácido desoxirribonucleico, macromolécula presente em todos os organismos vivos (com exceção dos vírus RNA), que contém o genoma – informação genética responsável por todas as características biológicas. É composto por duas

cadeias de nucleotídeos (bloco constituído por uma molécula de um açúcar, uma base nitrogenada e um grupo fosfato) (Almeida et al., 2004). Em geral os vírus de DNA replicam e agrupam-se no núcleo das células hospedeiras infectadas. Os vírus mais simples contêm DNA suficiente para codificar quatro proteínas; os mais complexos podem codificar de 100 a 200 proteínas (Rossetti, 2012; Lodish et al., 2014). Genomas virais podem ser de fita simples ou dupla (Lodish et al., 2014), com um ou dois filamentos (Brown, 1999).

Epidemiologia: “Estudo da frequência, da distribuição e dos determinantes dos estados ou eventos relacionados à saúde em específicas populações e a aplicação desses estudos no controle dos problemas de saúde” (Porta, 2014). Estudo da patologia em populações e dos fatores ambientais relacionados à sua quantidade e distribuição (Reis & Wardel, 2004). Estudo dos fatores que influenciam ocorrência da doença na população (Trigino et al., 2010).

Evolução neutra: a teoria neutralista da evolução molecular tem tido relevante sucesso como hipótese nula para a detecção de seleção positiva em evolução molecular. “A grande maioria das substituições de nucleotídeos que são fixadas nas populações é seletivamente neutra e que a evolução molecular resulta, em grande parte da deriva genética” (Freeman & Herron, 2009).

Informação: Resultado da transformação do dado em algo entendível, em algum determinado contexto, ou por conter significado formal (Cruz, 2013). Medida da alteração que os dados provocam numa estrutura de conhecimento (Araujo, 2018).

Migração: transferência de genes entre populações; o volume real de migração varia imensamente entre diferentes espécies, em função da mobilidade característica de cada uma, podendo ser um processo de homogeneização das populações (Freeman & Herron, 2009).

Nauatle: Idioma asteca, povo pré-colombiano nativo do México (Naika et al., 2006).

Processamento: transformar dados em informação pela execução de uma ou mais atividades eletrônicas (Cruz, 2013).

RNA: Ácido ribonucleico, macromolécula similar ao DNA, geralmente formada por cadeia simples de nucleotídeos (Almeida et al., 2004). Geralmente os vírus de RNA agrupam-se no citoplasma das células hospedeiras. Os vírus mais simples contêm RNA suficiente para codificar quatro proteínas; os mais complexos podem codificar de 100 a 200 proteínas (Rossetti, 2012; Lodish et al., 2014). Genomas virais

podem ser de fita simples ou dupla (Lodish et al., 2014), com um ou dois filamentos (Brown, 1999).

Seleção negativa ou purificadora: seleção natural contra mutações prejudiciais (eliminando-as através das gerações) (Freeman & Herron, 2009).

Seleção positiva: seleção natural a favor de mutações favoráveis (mantendo-as nas gerações seguintes) (Freeman & Herron, 2009).

Silenciamento gênico: mecanismo de proteção, degradação da atividade genética que resultaria em dano ao organismo quando em contato com agente externo, neutralizando o efeito nocivo (Loreto & Ferreira, 2012).

Swahili: Idioma pertencente à família Banto presente em muitos países da África Subsaariana (Britannica, 2016).

Transmembrana: Proteína (ou domínio proteico, parte funcional de uma proteína) (Pierce, 2017) que atravessa inteiramente a bicamada lipídica da célula (Junqueira & Carneiro, 2005).

2.6. Referências

Abreu PMV, Vaz AB, Fernandes PMB (2016) Enfoque biotecnológico para o controle de vírus de plantas. In: Resende RR (org.) Soccol CR, França LR (colab.). Biotecnologia aplicada à agro&indústria: fundamentos e aplicações. vol. 4. São Paulo. Blucher.

Albuquerque LC, Villanueva F, Resende RO, Navas-Castillo J, Barbosa JC, Nagata AKI (2013) Molecular characterization reveals Brazilian Tomato chlorosis virus to be closely related to a Greek isolate. *Tropical Plant Pathology* 38(4):332-336.

Almeida ME, Serafim RC, Eça LPM (2004) Apresentação de algumas técnicas utilizadas na biologia molecular. In: Eça LPM. *Biologia molecular: guia prático e didático*. Rio de Janeiro. Revinter.

Alzhanova DV, Napuli AJ, Creamer R, Dolja VV (2001) Cell-to-cell movement and assembly of a plant closterovirus: roles for the capsid proteins and Hsp70 homolog. *Embo Journal* 20:6997-7007.

Araujo CAA (2018) Um mapa da ciência da informação: história, subáreas e paradigmas. *Convergência em Ciência da Informação* 1(1) 45-70.

Avila AC (2010) Sintomas de ToCV em tomateiro. Available at: http://infopruga.cenargen.embrapa.br/ficha_praga.htm?idp=128726. Accessed on: October 13, 2018.

Barbosa JC, Teixeira APM, Moreira AG, Camargo LEA, Bergamin Filho A, Kitajima EW, Rezende JAM (2008) First report of Tomato chlorosis virus infecting tomato crops in Brazil. *Plant Disease* 92:1709.

Britannica (2016) Banto. Available at: <http://escola.britannica.com.br/article/487818/banto>. Accessed on August 17, 2018.

Brown TA (1999) *Genética: um enfoque molecular*. 3. ed. Rio de Janeiro. Guanaraba Koogan.

Caldart ET, Mata H, Canal CW, Ravazzolo AP (2016) Análise filogenética: conceitos básicos e suas utilizações como ferramenta para virologia e epidemiologia molecular. *Acta Scientiae Veterinariae* 44:1392

Cañizares MC, Navas-Castillo J, Moriones E (2008) Multiple suppressors of RNA silencing encoded by both genomic RNAs of the Crinivirus Tomato chlorosis virus. *Virology* 379(1):168-174.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. *Nucleic Acids Research* 44:D67-D72.

Coêlho LM (2017) *Diversidade genética de populações de Tomato chlorosis virus em tomateiro baseada no gene da proteína P22 (Dissertação de Mestrado)*. Morrinhos. Instituto Federal Goiano.

Colariccio A, Vicente M (2000) Principais viroses. In: São Paulo. *Manejo integrado de pragas e doenças do tomateiro: manual técnico*. Campinas. Secretaria de agricultura e abastecimento de SP.

Costa H, Ventura JA, Barbosa JC, Rezende JA (2009) *Amarelão do tomateiro (Documentos nº 175)*. Vitória. Incaper.

Cruz T (2013) *Sistemas, organização & métodos*. 4. ed. São Paulo. Atlas.

Edgar RC (2004) Muscle: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics* 5:1-19.

FAO (2018) Faostat: Data, Production, crops, all countries, production quantity, tomatoes, 2016. Available at: <http://www.fao.org/faostat/>. Accessed on: September 12, 2018.

Filgueira FAR (2003) *Solanáceas: agrotecnologia moderna na produção de tomate, batata, pimentão, pimenta, berinjela e jiló*. Lavras. UFLA.

Freeman S, Herron JC (2009) *Análise Evolutiva*. 4. ed. Porto Alegre. ArtMed.

Ganaie MM, Raja V, Reshi ZA, Verma V (2018) Family Solanaceae: Taxonomy and modern trends. *Annals of Plant Sciences* 7(8):2403-2414.

- García-Cano E, Navas-Castillo J, Moriones E, Fernández-Muñoz R (2010) Resistance to Tomato chlorosis virus in wild tomato species that impair virus accumulation and disease symptom expression. *Phytopathology* 100(6):582-592.
- García-Cano E, Resende RO, Fernández-Muñoz R, Moriones E (2006) Synergistic interaction between Tomato chlorosis virus and Tomato spotted wilt virus results in breakdown of resistance in tomato. *Phytopathology* 96(11):1263-1269.
- Giordano LB, Ribeiro CSC (2000) Origem, botânica e composição química do fruto. In: Silva JBC, Giordano LB. *Tomate para processamento industrial*. Brasília. Embrapa Hortaliças.
- Guimarães LRP, Nozaki DN, Moura MF, Spadotti DMA, Mituti T, Krause-Sakate R, Pavan MA (2017) Fungicide application can improve production of tomato coinfecting with Begomovirus and Crinivirus. *Pesquisa Agropecuária Brasileira* 52(6):435-442
- Hanssen IM, Lapidot M (2012) Major tomato viruses in the mediterranean basin. *Advances in virus research* 84:32
- Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting population subdivision. *Molecular Biology and Evolution* 9:138-151.
- IBGE (2018) Levantamento sistemático da produção agrícola: Tabela 2 - área, produção e rendimento médio - confronto das safras de 2017 e das estimativas para 2018 – Brasil. Available at: www.ibge.gov.br/estatisticas-novoportal/economicas/agricultura-e-pecuaria/9201-levantamento-sistemico-da-producao-agricola.html. Accessed on: September 12, 2018.
- Jardine JG, Neshich IAP, Mazoni I, Yano IH, Moraes FR, Salim JA, Borro L, Nishimura LS, Neshich G (2014) *Biologia computacional molecular e suas aplicações na agricultura*. In: Massruhá SMFS, Leite MAA, Luchiarini Junior A, Romani LAS (ed.) *Tecnologia da informação e comunicação e suas relações com a agricultura*. Campinas. Embrapa Informática Agropecuária.
- Junqueira LCU, Carneiro J (2005) *Biologia celular e molecular*. 8. ed. Rio de Janeiro: Guanabara Koogan.
- Karasev AV (2000) Genetic diversity and evolution of closteroviruses. *Annual review of phytopathology* 38, p. 293-324.
- Kimura M (1983) *The neutral theory of molecular evolution*. Cambridge University Press.
- Kumar S, Stecher G, Tamura K (2016) Mega7: molecular evolutionary genetics analysis. *Molecular Biology and Evolution* 33:1870-1874.
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452

- Lima JAA (org.) (2015) *Virologia essencial e viroses em culturas tropicais*. Fortaleza: UFC.
- Lodish H, Berk A, Kaiser CA, Krieger M, Bretscher A, Ploegh H, Amon A, Schott MP (2014) *Biologia celular e molecular*. 7. ed. São Paulo. Artmed.
- Loreto E, Ferreira HB (2012) Elementos genéticos móveis. In: Zaha A, Ferreira HB, Passaglia LMP (org.) *Biologia molecular básica*. 4. ed. Porto Alegre. Artmed.
- Lozano G, Grande-Perez A, Navas-Castillo J (2009) Populations of genomic RNAs devoted to the replication or spread of a bipartite plant virus differ in genetic structure. *Journal of Virology* 83(24):12973-12983.
- Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution* 7:377-394.
- Mansilla-Córdova PJ, Bampi D, Rondinel-Mendoza NV, Melo PCT, AL Lourenção, Rezende JAM (2018) Screening tomato genotypes for resistance and tolerance to Tomato chlorosis virus. *Plant Pathology* 67(5):1231-1237.
- Martelli GP, Agranovsky AA, Bar-Joseph M, Boscia D, Candresse T, Coutts RHA, Dolja VV, Hu JS, Jelkmann W, Karasev AV, Martin RR, Minafra A, Namba S, Vetten HJ (2009) Family Closteroviridae. In: King A, Adams MJ, Carstens EB, Lefkowitz E (eds.) *Virus Taxonomy*. 9 Report of the International Committee on Taxonomy of Viruses:987-1001.
- Nagata AKI, Ávila AC, Lopes CA (2005) Doenças viróticas. In: Lopes CA, Ávila AC (org.). *Doenças do tomateiro*. Brasília. Embrapa Hortaliças.
- Naika S, Jeude JVL, Goffau M, Hilmi M, Dam BV (2006) *A cultura do tomate: produção, processamento e comercialização*. Wageningen. Agromisa/CTA.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70(12):3321-3323.
- Nei M (1987) *Molecular Evolutionary Genetics*. New York. Columbia University Press.
- Nepa (núcleo de estudos e pesquisas em alimentação) (2011) *Tabela brasileira de composição de alimentos*. Campinas. Unicamp.
- Nylander JAA (2004) *MrModeltest v2*. Evolutionary Biology Centre, Uppsala University.
- Peralta I, Knapp S, Spooner DMS (2005) New species of wild tomatoes (*Solanum* Section *Lycopersicon*: Solanaceae) from northern Peru. *Systematic Botany* 30(2): 424–434.
- Pierce BA (2017) *Genética: um enfoque conceitual*. Rio de Janeiro. Guanabara Koogan.

Pond SLK, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* 22(5):1208-1222.

Porta M (ed.) (2014) A dictionary of epidemiology. 6. ed. New York. International Epidemiological Association/Oxford.

Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. *Systematic Biology*, 50:580-601.

Ramalhete ML, Cortez I, Furtado AJ (2007) Novas doenças da alface e o vírus do bronzeado do tomateiro na região Entre-Douro-e-Minho. *Mirandela. Direcção Regional de Agricultura e Pescas do Norte*.

Rambaut A (2016) Figtree: tree figure drawing tool. Version 1.4.3. Institute of Evolutionary Biology. University of Edinburgh.

Reis EM, Wardel Filho JA (2004) Princípios de epidemiologia. in: Reis EM (org.) *Previsão de doenças de plantas*. Universidade de Passo Fundo.

Robertis EM, Hib J (2017) *Biologia celular e molecular*. 16 ed. Rio de Janeiro. Guanabara Koogan.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2011) MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3):539–542.

Rossetti MLR (2012) A Célula e seus constituintes celulares. In: Zaha A, Ferreira HB, Passaglia LMP (org.) *Biologia molecular básica*. 4. ed. Porto Alegre. Artmed.

Satyanarayana T, Gowda S, Ayllón MA, Dawson WO (2004) Closterovirus bipolar virion: evidence for initiation of assembly by minor coat protein and its restriction to the genomic RNA 5' region. *Proceedings of the National Academy of Sciences of the United States of America* 101(3):799-804.

Swofford DL (2002) *Paup**: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates. Sunderland, Massachusetts.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3): 585–595.

Trigino R. N.; Windham, M. T.; Windham, A. S. (2010) *Fitopatologia: conceitos e exercícios de laboratório*. 2. ed. Porto Alegre. Artmed.

Tzanetakis IE, Martin RR, Wintermantel WM (2013) Epidemiology of criniviruses: an emerging problem in world agriculture. *Frontiers in microbiology* 4:119.

Vargas MRI (2016) Teste do indutor de resistência ASM como alternativa de controle da clorose do tomateiro causada pelo Tomato chlorosis virus (ToCV) e possíveis efeitos

sobre o vetor *Bemisia tabaci* biótipo B (Trabalho de conclusão de curso). Piracicaba. USP.

Vasconcelos ATR, Almeida DF (2012) Bioinformática na análise de genes e genomas. In: Zaha A, Ferreira HB, Passaglia LMP (org.) *Biologia molecular básica*. 4. ed. Porto Alegre. Artmed.

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Pond SL (2018) Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution* 35(3):773-777.

Wintermantel WM, Wisler GC (2006) Vector specificity, host range, and genetic diversity of Tomato chlorosis virus. *Plant Disease* 90(6):814-819.

Wintermantel WM, Wisler GC, Anchieta AG, Liu HY, Karasev AV, Tzanetakis IE (2005) The complete nucleotide sequence and genome organization of Tomato chlorosis virus. *Archives of Virology* 150:2287-2298.

Wisler GC, Li RH, Liu HY, Lowry DS, Duffus JE (1998) Tomato chlorosis virus: A new whitefly-transmitted, phloem-limited, bipartite closterovirus of tomato. *Phytopathology* 88:402-409.

3. CAPÍTULO I

Diversidade genética de *Tomato chlorosis virus* baseada nas sequências nucleotídicas dos genes das proteínas RdRp, p22 e Hsp70h

(Normas de acordo com a revista *Tropical Plant Pathology*)

RESUMO. *Tomato chlorosis virus* (ToCV), gênero *Crinivirius*, família *Closteroviridae*, provoca manchas cloróticas, necrose das folhas e consequente redução da produtividade em tomateiro (*Solanum lycopersicum*). É transmitido pelas espécies de mosca-branca *Bemisia tabaci*, *B. argentifolii*, *Trialeurodes vaporariorum* e *T. abutilonea* de forma semipersistente. Possui genoma bipartido. Utilizaram-se sequências genéticas das ORFs RdRp, P22 e Hsp70h de isolados ToCV para determinar a variabilidade genética e obter árvores filogenéticas. Os isolados foram obtidos de amostras foliares de tomateiro coletadas no Brasil (Brasília, Distrito Federal; Goianápolis, Goiás) juntamente a sequências obtidas no banco de dados *GenBank*. Quatro populações (Brasil, China, Coreia do Sul e Espanha) foram analisadas usando métodos (Muscle, ModelTest, Inferência Bayesiana, SLAC) e ferramentas de bioinformática (Mega, Paup, MrBayes, DNAsp, DataMonkey, Figtree) para obtenção de valores dos parâmetros descritores de variabilidade genética e construção de árvores filogenéticas. Os resultados indicam que dentre as regiões genômicas dos isolados analisados, a RdRp apresenta maior diferenciação genética e a P22 apresenta maior diversidade nucleotídica e pressão de seleção, comparadas às demais ORFs analisadas. Os isolados brasileiros analisados apresentam menor diversidade nucleotídica e maior neutralidade genética quando comparados às demais

populações analisadas. Os isolados brasileiros analisados apresentam possível proximidade evolutiva com alguns isolados da Coreia do Sul.

PALAVRAS-CHAVE: *Crinivirus*; *Closteroviridae*; RdRp; P22; Hsp70h; ToCV.

ABSTRACT. *Tomato chlorosis virus* (ToCV), genus *Crinivirus*, family *Closteroviridae*, causes chlorotic spots, leaf necrosis and consequent reduce tomato (*Solanum lycopersicum*) productivity. Transmitted by the whitefly species *Bemisia tabaci*, *B. argentifolii*, *Trialeurodes vaporariorum* and *T. abutilonea* in semipersistent way. It has a bipartite genome. Gene sequences of RdRp, P22 and Hsp70h ORFs from ToCV isolates were used to determine genetic variability and to obtain phylogenetic trees. The isolates were from tomato leaf samples collected in Brazil (Brasília, Distrito Federal; Goianópolis, Goiás) along with sequences obtained from the GenBank database. Four populations (Brazil, China, South Korea and Spain) were analyzed using methods (Muscle, ModelTest, Bayesian Inference and SLAC) and bioinformatics tools (Mega, Paup, MrBayes, DNAsp, DataMonkey, Figtree) to obtain parameter descriptors values of genetic variability and construction of phylogenetic trees. The results indicate that among the genomic regions of the analyzed isolates, RdRp presents higher genetic differentiation and P22 shows higher nucleotide diversity and selection pressure, compared to the other analyzed ORFs. The analyzed Brazilian isolates showed lower nucleotide diversity and higher genetic neutrality when compared to the other analyzed populations. The analyzed Brazilian isolates present a possible evolutionary proximity with some South Korean isolates.

KEY WORDS: *Crinivirus*; *Closteroviridae*; RdRp; P22; Hsp70h; ToCV.

3.1. Introdução

Tomato chlorosis virus (ToCV) é classificado no gênero *Crinivirus*, família *Closteroviridae*. No tomateiro, limita-se ao floema da planta, provocando sintomas de manchas cloróticas internervais que iniciam nas folhas do baixeiro, disseminam-se aos pontos de crescimento e evoluem para machas necróticas, resultando em perdas produtivas (Wisler et al., 1998; Wintermantel & Wisler, 2006).

Na natureza, é transmitido entre as plantas hospedeiras pelas espécies de mosca-branca *Bemisia tabaci*, *B. argentifolii*, *Trialeurodes vaporariorum* e *T. abutilonea* (Wisler et

al., 1998; Wintermantel & Wisler, 2006). Descrito inicialmente nos Estados Unidos (Wisler et al., 1998), foi relatado no Brasil entre 2006 e 2007 (Barbosa et al., 2008), ameaçando a produção nos diversos países em que há cultivo de tomateiro (Albuquerque et al., 2013);

O genoma de ToCV é bipartido, constituído por duas fitas de RNA (RNA1 e RNA2) de fita simples. O RNA 1 apresenta 4 ORF (*open reading frames* - fases abertas de leitura) e o RNA2 apresenta 9 ORFs (Wisler et al., 1998). Na ORF 1b (RNA1), codifica-se a *RNA-dependent RNA polimerase* (RdRp). Na ORF 2 (RNA1), codifica-se uma proteína (P22) que atua como supressora de silenciamento gênico (Cañizares et al. 2008). A Hsp70h (*heat shock protein 70*) é codificada na ORF 2 (RNA2), estando envolvida na montagem e movimentação célula a célula das partículas virais (Alzhanova et al., 2001; Peremyslov et al., 1999).

Mudanças ambientais e crescente mobilidade de bens, pessoas e vetores de doenças são elementos que facilitam a disseminação de doenças virais e criam condições potencialmente ideais para pandemias. Vírus podem evoluir rapidamente, variando seus genomas em resposta a pressões seletivas decorrentes de mecanismos de defesa dos organismos hospedeiros e ou intervenções de tratamento e controle (Ibrahim et al., 2018).

Objetivou-se determinar a variabilidade genética e obter árvores filogenéticas utilizando bioinformática a partir de sequências genéticas das ORFs RdRp, P22 e Hsp70h de isolados de ToCV.

3.2. Material e Métodos

Obtenção de amostras. As sequências genéticas de isolados de ToCV disponíveis para este estudo foram obtidas de isolados virais de amostras foliares de tomateiro coletadas em Brasília, Distrito Federal (DF) (n=4) e Goianápolis, Goiás (GO) (n=9), no ano de 2012. O RNA total das amostras foi extraído utilizando o reagente Trizol (Chomczynski & Sacchi, 1987, 2006; Chomczynski, 1993), conforme instruções do fabricante e a detecção de ToCV foi realizada por RT-PCR (*Reverse Transcriptase - Polymerase Chain Reaction*) (Dovas et al., 2002). Os produtos da PCR foram submetidos à eletroforese em gel de agarose a 1%, os fragmentos de tamanho esperado (720 pb) foram visualizados usando brometo de etídio em equipamento fotodocumentador. Os produtos da PCR foram clonados através do vetor pGEM-T Easy (Promega), introduzidos em células competentes de *Escherichia coli* e os insertos sequenciados pela empresa Macrogen Inc. (Seoul, South Korea). Em sua maior parte, as sequências da Hsp70h tinham 1665 pb (pares de base), da P22 582 pb e da RdRp 1515 pb.

Também, utilizaram-se sequências publicadas no banco de dados *GenBank* (Clark et al., 2016) (Brasil n=2, China n=18, Coreia do Sul n=17, Estados Unidos n=4, Espanha n=20 e Grécia n=2) (Apêndice 2). Foram analisadas as regiões genômicas correspondentes às ORFs RdRp, P22 (RNA1) e Hsp70h (RNA2).

Análises das sequências. As sequências obtidas do *GenBank* foram editadas usando DNA for Windows v.2.2 (Dixon, 1999) e colocadas em formato Fasta. Para o alinhamento das sequências foi utilizado o algoritmo *Muscle* (Edgar, 2004) incluído no Mega 7 (*Molecular evolutionary genetics analysis*) (Kumar et al., 2016).

Descritores de variabilidade. Análises de diferença entre as populações (polimorfismo) e inferência sobre estrutura e seleção populacional foram efetuadas usando o DnaSP v.5.10 (Librado & Rozas, 2009; Rozas et al., 2017). Os isolados foram separados em populações conforme origem geográfica, tendo quantificados seus principais parâmetros descritores da variabilidade genética: diversidade nucleotídica (π ; média de diferenças médias nucleotídicas entre sequências) (Nei, 1987; Lynch & Crease, 1990; Rozas et al., 2017), índice Tajima's D (teste da hipótese de neutralidade) (Tajima, 1989; Kimura, 1983; Rozas et al., 2017) e índice G_{ST} (estimativa de diferenciação entre populações) (Nei, 1973; Hudson et al., 1992; Rozas et al., 2017).

Pressão de seleção. A razão entre as substituições não sinônimas e sinônimas (d_{NS}/d_S) foi estimada com o método *Single Likelihood Ancestor Counting* (SLAC), para inferir a força evolutiva atuante em cada ORF, baseada na verossimilhança entre as sequências analisadas para reconstruir as ancestrais (Pond & Frost, 2005), disponível no Datamonkey (Weaver et al. 2018).

Análise filogenética. Para inferir as relações evolutivas entre os isolados analisados, as sequências nucleotídicas completas de cada ORF foram submetidas aos programas: Paup v.4.0b10 (Phylogenetic Analysis Using Parsimony, paup.phylosolutions.com), inferência de árvore filogenética considerando a menor mudança evolutiva como a mais plausível (Swofford, 2002); MrModeltest2, identificação de qual modelo evolutivo melhor se ajusta aos dados, utilizando testes de razão e verossimilhança (Nylander, 2004; Posada & Crandall, 2001; Salemi & Vandamme, 2003); MrBayes v.3.1, construção da árvore filogenética por inferência bayesiana (Ronquist et al., 2011); Figtree v.1.4.3, exibição e formatação da árvore filogenética (Rambaut, 2016). Durante a análise de inferência bayesiana, utilizou-se Notepad++ (Ho, 2016) (editor de textos adequado a arquivos com milhares de linhas) em ajustes de parâmetros possibilitando o final do processo.

3.3. Resultados e Discussão

Variabilidade genética. Foram analisadas populações do Brasil (BRA) (Brasília, DF e Goianópolis, GO) juntamente a populações da Coreia do Sul (KOR), Espanha (ESP) e China (CHN) disponíveis no *GenBank*. Referente a RdRp, a população do Brasil apresentou valores de diversidade nucleotídica (π) significativamente inferiores, aproximadamente 2 vezes menores aos observados em Coreia do Sul e Espanha (Tabela 1), na P22 isolados da Coreia tem quase 3 vezes e da Espanha 5 vezes maior diversidade nucleotídica. Possivelmente isso ocorra em razão da relativamente recente introdução do vírus no Brasil (Barbosa et al., 2008). Baixos valores de diversidade genética já foram encontrados para isolados brasileiros em estudo anterior (Barbosa et al., 2013).

Testes de neutralidade. Com o teste Tajima's D verifica-se a hipótese de neutralidade genética entre as populações. Valores de $D < 0$ indicam mutações negativas, $D = 0$ equilíbrio neutro e $D > 1$ seleção positiva (Equiarte et al., 2007; Tajima, 1983, 1989). Nota-se que a RdRp está sob possível seleção positiva nas populações provenientes da Espanha e Coreia do Sul; a população brasileira pode estar sob possível e intensa seleção negativa, possivelmente em razão de evento demográfico recente (contração ou expansão populacional). Na P22, há possível seleção negativa nas populações oriundas do Brasil, Espanha e Coreia do Sul. Na Hsp70h, aparenta-se possível seleção negativa nas populações China, Brasil e Espanha, e possível seleção positiva na Coreia do Sul. Avaliando-se as populações, no Brasil apresenta-se possível seleção negativa em todas ORFs avaliadas, enquanto Espanha apresenta possível seleção negativa em duas delas e Coreia do Sul possível seleção positiva em duas ORFs. China teve somente uma ORF avaliada, em possível seleção negativa (Tabela 3).

Testes da Pressão de Seleção. A razão entre o número de substituições não sinônimas (d_{NS}) e o de sinônimas (d_S) indica a pressão de seleção nos genes em análise. Valores de $d_{NS}/d_S > 1$ apontam pressão de seleção positiva, $d_{NS}/d_S = 1$ neutralidade e $d_{NS}/d_S < 1$ pressão de seleção negativa (Freeman & Herron, 2009). Valores de $d_{NS}/d_S < 1$ (Tabela 4) foram encontrados para todas as ORFs analisadas (RdRp, P22 e Hsp70h), denotando seleção negativa, que pode estar atuando na manutenção da integridade das sequências nas ORFs. A P22 apresentou aproximadamente três vezes mais variações sinônimas (0,67354) que a Hsp70h (0,301196), e seis vezes mais que a RdRp (0,119908); a RdRp é altamente conservada entre os membros do gênero *Crinivirus* (Wintermantel et al., 2005). Nota-se que a proteína P22 são demonstrados valores um tanto elevados, possivelmente em razão da

proximidade com a região 3' do RNA1, caracterizada como região com alta variabilidade genética (Dolja et al., 2006).

Relações filogenéticas. Para o descritor G_{ST} (estimativa de diferenciação entre as populações), valores mais próximos a 1 (um) representam maior divergência (Ridley, 2007). A comparação de todas as populações entre si resultou em valores (G_{ST}) abaixo de 0,1; as maiores diferenças foram encontradas na RdRp, com valores entre 0,09919 (Brasil x Coreia do Sul) e 0,7286 (Brasil x Espanha); na P22 os valores ficaram entre 0,05661 (Brasil x Coreia do Sul) e 0,01822 (Coreia do Sul x Espanha); na Hsp70h o maior valor encontrado foi de 0,0563 (Brasil x Coreia do Sul) e o menor 0,00962 (Brasil x China) (Tabela 2).

As árvores filogenéticas construídas por inferência bayesiana com base nas sequências nucleotídicas da Hsp70h, RdRp e P22 (Figura 1) demonstram agrupamentos com possível base na procedência geográfica entre partes dos isolados analisados e prevalência da divisão das populações em 2 grandes grupos (Tipo I e Tipo II) (Lozano et al., 2009; Lande-Rios et al., 2015), em que o Tipo I compõe-se somente de isolados espanhóis, e o Tipo II de isolados de outros países, exceto na P22, em que um isolado espanhol enquadra-se no Tipo I.

Os isolados brasileiros ToC121 e ToC221 constituem grupo a parte na RdRp (Figura 1A). Na P22, os isolados ToC121 e ToC221 também apresentam separação, derivando do isolado ToC201 (Figura 1B). Na RdRp, o isolado brasileiro JQ9525600 possui alguma proximidade com isolado grego (Figura 1A) (Albuquerque et al., 2013). Na Hsp70, os isolados ToC22H e ToC50H separam-se dos demais (Figura 1C). Os isolados Chineses estão agrupados somente na Hsp70h (Figura 1C), ORF em que esta população teve maior número de isolados. Há maior número de isolados sul-coreanos com sequências da RdRp e Hsp70h, constituindo grupos desta população.

Comparando a estrutura das árvores filogenéticas (Figura 1) aos valores do coeficiente de diferenciação gênica (G_{ST}) (Tabela 2) não se notam correlações. Observam-se algumas possíveis dispersões evolutivas na RdRp e P22 entre os isolados analisados da Coreia do Sul e China. Na Hsp70h, há clara constituição de grupos populacionais entre os isolados. Os isolados do Brasil possivelmente tenham o mesmo ancestral comum que dois isolados da Coreia do Sul e de um da Grécia.

Os resultados indicam que dentre as regiões genômicas dos isolados analisados, a RdRp apresenta maior diferenciação genética e a P22 apresenta maior diversidade nucleotídica e pressão de seleção, comparadas às demais ORFs analisadas. Os isolados brasileiros analisados apresentam menor diversidade nucleotídica e maior neutralidade genética quando comparados às demais populações analisadas. Os isolados brasileiros

analisados apresentam possível proximidade evolutiva com alguns isolados da Coreia do Sul. Futuros estudos podem verificar eventuais indicativos de rearranjos entre os isolados e possível surgimento de novas estirpes (Kang et al. 2018; Lee et al., 2018) e ou patologias (Ruiz et al., 2018).

3.4. Referências

Albuquerque LC, Villanueva F, Resende RO, Navas-Castillo J, Barbosa JC, Nagata AKI (2013) Molecular characterization reveals brazilian Tomato chlorosis virus to be closely related to a Greek isolate. *Tropical Plant Pathology* 38(4):332-336.

Alzhanova DV, Napuli AJ, Creamer R, Dolja VV (2001) Cell-to-cell movement and assembly of a plant Closterovirus: roles for the capsid proteins and HSP70H homolog. *European Molecular Biology Organization* 20:6997-7007.

Barbosa JC, Rezende JAM, Filho AB (2013) Low genetic diversity suggests a single introduction and recent spread of Tomato chlorosis virus in Brazil. *Journal of Phytopathology* 161(11-12):884-886.

Barbosa JC, Teixeira APM, Moreira AG, Camargo LEA, Bergamin Filho A, Kitajima EW, Rezende JAM (2008) First report of Tomato chlorosis virus infecting tomato crops in Brazil. *Plant Disease* 92:1709.

Cañizares MC, Navas-Castillo J, Moriones E (2008) Multiple suppressors of RNA silencing encoded by both genomic RNAs of the Crinivirus Tomato chlorosis virus. *Virology* 379(1):168-174.

Chomczynski P (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* 15(3):532-534, 536-537.

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* 162(1):156-159.

Chomczynski P, Sacchi N (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature Protocols* 1(2):581-585.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. *Nucleic Acids Research* 44:D67-D72.

Coêlho LM (2017) Diversidade genética de populações de Tomato chlorosis virus em tomateiro baseada no gene da proteína P22. Master dissertation. Instituto Federal Goiano. Morrinhos, GO, Brazil.

Dixon D (1999) Dna for Windows! v.2.2. Available at: dna-software.co.uk. Accessed on: October 20, 2018.

- Dolja VV, Kreuze JF, Valkonen JPT (2006) Comparative and functional genomics of closteroviruses. *Virus Research* 117:38-51.
- Dovas CI, Katis NI, Avgelis AD (2002) Multiplex detection of criniviruses associated with epidemics of a yellowing disease of tomato in Greece. *Plant Disease* 86:1345-1349.
- Edgar RC (2004) Muscle: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics* 5:1-19.
- Eguiarte LE, Souza V, Aguirre X (comp.) (2007) *Ecología molecular*. México D.F. Instituto nacional de ecología.
- Freeman S, Herron JC (2009) *Análise Evolutiva*. 4. ed. Porto Alegre. ArtMed.
- Gao LL (2016) Molecular identification of tomato chlorotic virus and establishment and application of fluorescent quantitative PCR detection system. Master dissertation. Shandong Agricultural University. Tai'an, Shandong, China.
- Ho D (2016) Notepad++. Available at: notepad-plus-plus.org. Accessed on: October 20, 2018.
- Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting population subdivision. *Molecular Biology and Evolution* 9:138-151.
- Ibrahim B, McMahon DP, Hufsky F, Beer M, Deng L, Mercier PL, Palmarini M, Thiel V, Marz M (2018) A new era of virus bioinformatics. *Virus Research* 251:86-90.
- Kang Y, Wang Y, Hsia, C, Tsai W, Huang L, Yeh S, Chen T (2018) Molecular characterization and detection of a genetically distinct tomato chlorosis virus strain in Taiwan. *Plant Disease* 102(3):600-607.
- Karwitha M, Feng Z, Yao M, Chen X, Zhang W, Liu X, Tao X (2014). The complete nucleotide sequence of the RNA 1 of a Chinese isolate of Tomato chlorosis virus. *Journal of phytopathology* 162:411-415.
- Kataya ARA, Stavridou E, Farhan K, Livieratos IC (2008) Nucleotide sequence analysis and detection of a Greek isolate of Tomato chlorosis virus. *Plant Pathology* 57(5):819-824
- Kimura M (1983) *The neutral theory of molecular evolution*. Cambridge University Press.
- Kumar S, Stecher G, Tamura K (2016) Mega7: molecular evolutionary genetics analysis. *Molecular Biology and Evolution* 33:1870-1874.
- Landeo-Rios YM, Navas-Castillo J, Moriones E, Canizares MC (2015) Genetic diversity and silencing suppression activity of the p22 protein of Tomato chlorosis virus isolates from tomato and sweet pepper. *Virus Genes* 51(2):283-289
- Lee Y, Kil E, Kwak H, Kim M, Seo J, Lee S, Choi H (2018) Phylogenetic characterization of Tomato chlorosis virus population in Korea: evidence of reassortment between isolates from different origins. *The Plant Pathology Journal* 34(3):199-207.

- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452
- Lozano G, Moriones E, Navas-Castillo J (2006) Complete nucleotide sequence of the RNA2 of the crinivirus tomato chlorosis virus. *Archives of Virology* 151(3):581-587
- Lozano G, Moriones E, Navas-Castillo J (2007) Complete sequence of the RNA1 of a European isolate of tomato chlorosis virus. *Archives of Virology* 152(4):839-841
- Lozano G, Grande-Perez A, Navas-Castillo J (2009) Populations of genomic RNAs devoted to the replication or spread of a bipartite plant virus differ in genetic structure. *Journal of Virology* 83(24):12973-12983.
- Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution* 7:377-394.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70:3321-3323.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press.
- Nylander JAA (2004) MrModeltest v2. Evolutionary Biology Centre, Uppsala University.
- Orilio AF, Fortes IM, Navas-Castillo J (2014) Infectious cDNA clones of the crinivirus Tomato chlorosis virus are competent for systemic plant infection and whitefly-transmission. *Virology* 464-465:365-374.
- Peremyslov VV, Hagiwara Y, Dolja VV (1999) HSP70H homolog functions in cell-to-cell movement of a plant virus. *Proceedings of the National Academy of Sciences of the United States of America* 96:14771-14776.
- Pond SLK, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* 22(5):1208-1222.
- Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. *Systematic Biology* 50:580-601
- Rambaut A (2016) Figtree: tree figure drawing tool. Version 1.4.3. Institute of Evolutionary Biology. University of Edinburgh. Available at: tree.bio.ed.ac.uk/software/figtree. Accessed on: October 20, 2018
- Ridley M (2007) *Evolução*. 3. ed. Porto Alegre. Artmed.
- Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2011) MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3):539-542.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution* 34:3299-3302.

Ruiz L, Simón A, García C, Velasco L, Janssen D (2018) First natural crossover recombination between two distinct species of the family Closteroviridae leads to the emergence of a new disease. *Plos One* 13(9).

Salemi M, Vandamme A (2003) *The phylogenetic handbook: a practical approach to DNA and protein phylogeny*. Cambridge University Press.

Swofford DL (2002) *Paup*: Phylogenetic Analysis Using Parsimony (*and other methods)*. Version 4. Sinauer Associates. Sunderland, Massachusetts.

Tajima F (1983) Evolutionary relationship of ADN sequences in finite populations. *Genetics* 105:437-460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Pond SL (2018) Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution* 35(3):773-777.

Wintermantel WM, Wisler GC, Anchieta AG, Liu HY, Karasev AV, Tzanetakis IE (2005) The complete nucleotide sequence and genome organization of tomato chlorosis virus. *Archives of Virology* 150:2287-2298.

Wisler GC, Li RH, Liu HY, Lowry DS, Duffus JE (1998) Tomato chlorosis virus: a new whitefly-transmitted, phloem-limited, bipartite Closterovirus of tomato. *Phytopathology* 88:402-409.

Zhao LM, Li G, Gao Y, Liu Y, Sun G, Zhu X (2014) Molecular detection and complete genome sequences of Tomato chlorosis virus isolates from infectious outbreaks in China. *Journal of Phytopathology* 162(10):627-634.

Zhao LM, Li G, Gao Y, Zhu YR, Liu J, Zhu XP (2015) Reverse transcription loop-mediated isothermal amplification assay for detecting Tomato chlorosis virus. *Journal of virological methods* 213:93-97.

Zhao R, Wang N, Wang R, Chen H, Shi Y, Fan Z, Zhou T (2014) Characterization and full genome sequence analysis of a Chinese isolate of Tomato chlorosis virus. *Acta Virologica* 58:92-94.

Tabela 1: Estrutura genética de populações de *Tomato chlorosis virus*, genes RdRp, P22 e Hsp70h.

Sub população	RNA 1				RNA 2	
	RdRp		P22		Hsp70h	
	N	π (SD)	N	π (SD)	N	π (SD)
Brasil	14	0,00167 (0,00039)	14	0,00272 (0,00058)	15	0,00175 (0,00031)
Coreia	9	0,00352 (0,00045)	4	0,00716 (0,00137)	8	0,00341 (0,00108)
Espanha	5	0,00383 (0,00069)	11	0,01400 (0,00409)	5	0,00492 (0,00101)
China	-	-	-	-	15	0,00639 (0,00074)
Total	34	0,00742 (0,00134)	34	0,01692 (0,00187)	46	0,00872 (0,00089)

N: Número de isolados. π : Diversidade nucleotídica. SD: Desvio padrão

Tabela 2. Diferenciação genética entre populações de *Tomato chlorosis virus*

População 1	População 2	G_{ST}		
		RdRp	P22	Hsp70h
Brasil	Coreia	0,09919	0,05661	0,05263
Brasil	China	-	-	0,00962
Brasil	Espanha	0,07286	0,04638	0,02166
Coreia	China	-	-	0,04783
Coreia	Espanha	0,07995	0,01822	0,04926
China	Espanha	-	-	0,01813

G_{ST} : coeficiente de diferenciação gênica.

Tabela 3. Teste da hipótese de neutralidade genética de cada proteína (RdRp, P22 e Hsp70) entre as populações.

População	Tajima's D		
	RdRp	P22	Hsp70h
Brasil	-2,04167	-1,39890	-2,25629
Coreia	0,98520	-0,44640	0,67142
China	-	-	-1,69765
Espanha	0,70883	-0,68330	-1,07808
Total	-1,16576	-0,98267	-1,89718*

Tabela 4. Relação entre substituições sinônimas e não sinônimas (d_{NS}/d_S) entre as sequências nucleotídicas dos genes das proteínas P22, Hsp70h e RdRp nas populações de *Tomato chlorosis virus*.

ORF	d_{NS}/d_S
P22	0,673540
Hsp70h	0,301196
RdRp	0,119908

A (RdRp)

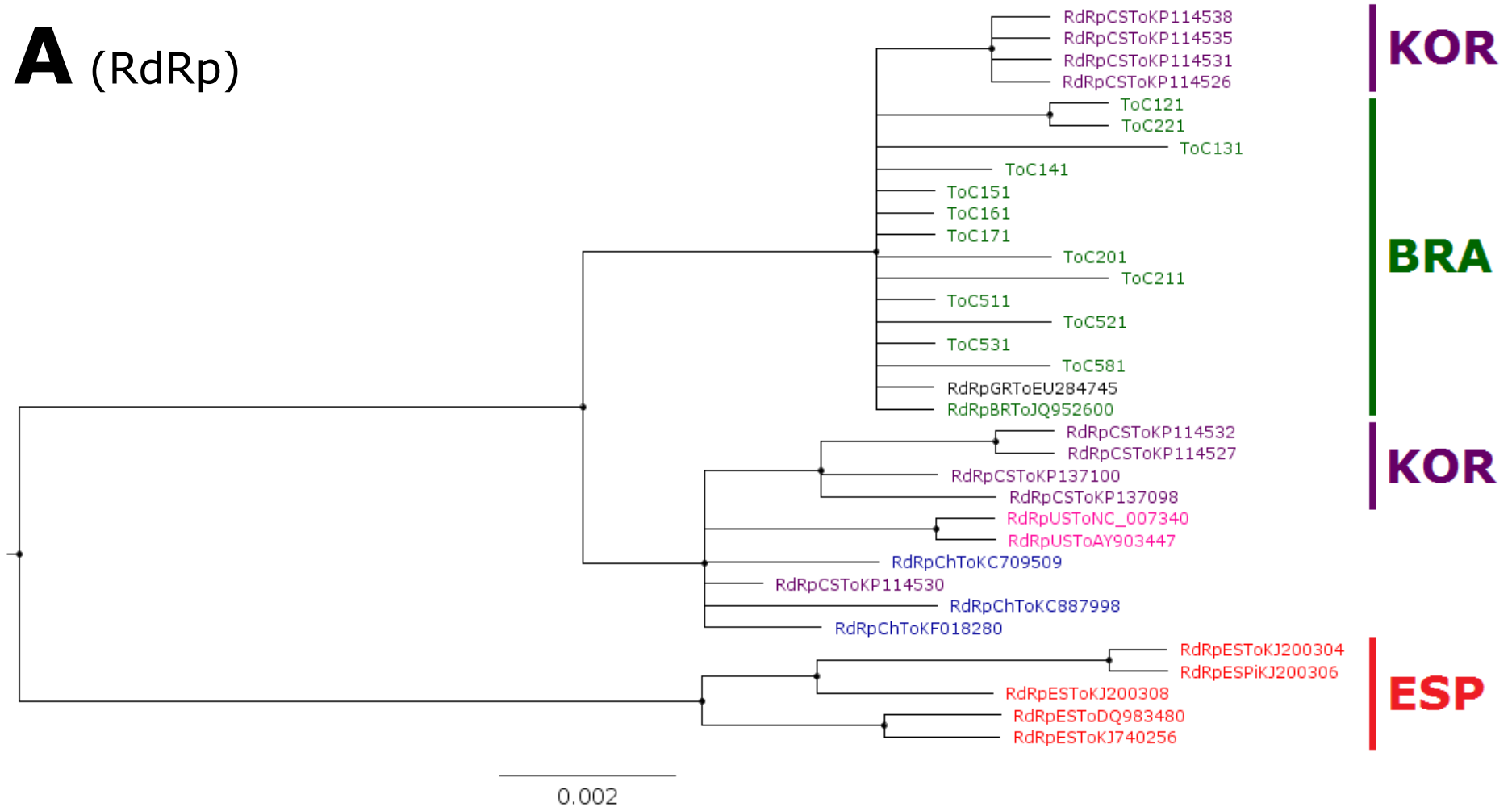
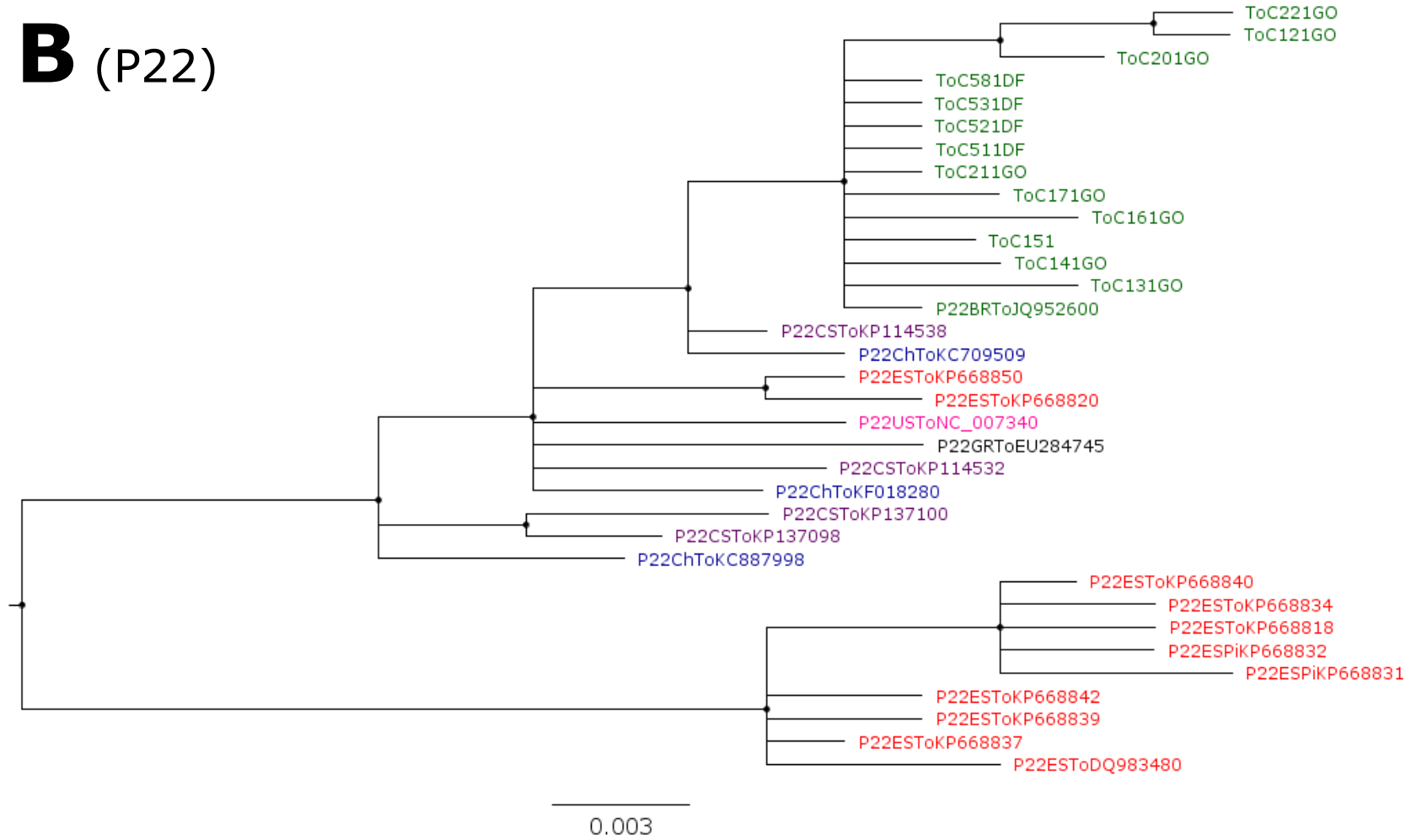
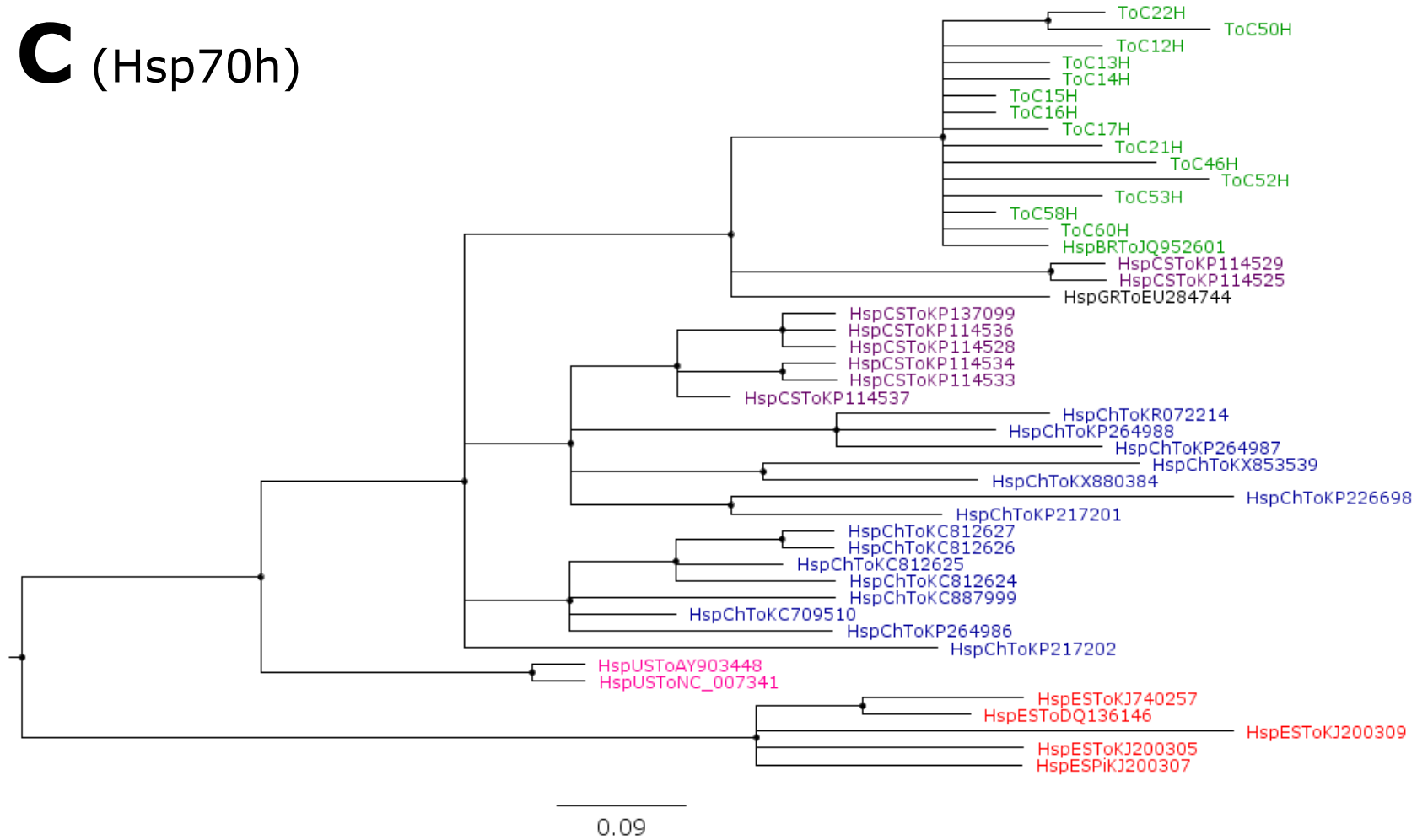


Figura 1. Árvores filogenéticas construídas por inferência bayesiana, baseadas nas seqüências nucleotídicas de isolados de *Tomato chlorosis virus* (ToCV) deste estudo (identificação na árvore inicia com ToC, provenientes do Brasil) comparados a isolados da China, Coreia do Sul, Espanha, Grécia e Estados Unidos. A barra de escala abaixo representa o número de substituições nucleotídicas por sítio. Cada árvore representa a análise de uma proteína: RdRp (A), P22 (B), Hsp70h (C).

B (P22)



C (Hsp70h)



4. CONCLUSÃO GERAL

Os resultados obtidos demonstram possível estruturação baseada na localização geográfica, embora não seja possível inferir padrão em alguns casos. Possivelmente tenham ocorrido subdivisões populacionais entre isolados de regiões geograficamente próximas (p. ex. China e Coreia do Sul), que possivelmente compartilham um ou mais ancestrais comuns.

Embora nota-se alguma variação maior na Coreia do Sul (RdRp e Hsp70h) e Espanha (RdRp), na maior parte dos isolados sugere-se seleção purificadora como força evolutiva atuante, sugerindo maior preservação das sequências genéticas, sugerindo que o vírus analisado ora apresenta certa estabilidade.

Estudos desta natureza podem auxiliar no monitoramento das doenças virais, verificando a estabilidade viral em cultivos com incidência confirmada de ToCV, ou mesmo o surgimento de surtos repentinos, contribuindo à tomada de decisões que possam mitigar danos à coletividade de produtores e consumidores.

APÊNDICES

Apêndice 1. Isolados correspondentes às sequências nucleotídicas das proteínas P22, RdRp e Hsp70h de *Tomato chlorosis virus* obtidos de amostras coletadas no Distrito Federal (DF) e no estado de Goiás (GO), Brasil.

P22	RNA 1		RNA2 Hsp70h	Isolado	Origem
	RdRp				
>ToC121	>ToC121	-	-	12	Goianápolis, GO
-	-	>ToC12H	-	12	Goianápolis, GO
>ToC131	>ToC131	-	-	13	Goianápolis, GO
-	-	>ToC13H	-	13	Goianápolis, GO
>ToC141	>ToC141	-	-	14	Goianápolis, GO
-	-	>ToC14H	-	14	Goianápolis, GO
>ToC151	>ToC151	-	-	15	Goianápolis, GO
-	-	>ToC15H	-	15	Goianápolis, GO
>ToC161	>ToC161	-	-	16	Goianápolis, GO
-	-	>ToC16H	-	16	Goianápolis, GO
>ToC171	>ToC171	-	-	17	Goianápolis, GO
-	-	>ToC17H	-	17	Goianápolis, GO
>ToC201	>ToC201	-	-	20	Goianápolis, GO
>ToC211	>ToC211	-	-	21	Goianápolis, GO
-	-	>ToC21H	-	21	Goianápolis, GO
>ToC221	>ToC221	-	-	22	Goianápolis, GO
-	-	>ToC22H	-	22	Goianápolis, GO
-	-	>ToC46H	-	46	Brasília, DF
-	-	>ToC50H	-	50	Brasília, DF
>ToC511	>ToC511	-	-	51	Brasília, DF
>ToC521	>ToC521	-	-	52	Brasília, DF
-	-	>ToC52H	-	52	Brasília, DF
>ToC531	>ToC531	-	-	53	Brasília, DF
-	-	>ToC53H	-	53	Brasília, DF
>ToC581	>ToC581	-	-	58	Brasília, DF
-	-	>ToC58H	-	58	Brasília, DF
-	-	>ToC60H	-	60	Brasília, DF

Apêndice 2. Sequências nucleotídicas das proteínas P22, RdRp e Hsp70h de isolados de *Tomato chlorosis virus* obtidas no *GenBank*.

RNA 1 P22	RNA 1 RdRp	RNA2 Hsp70h	Isolado	Pais de origem	Código de acesso <i>GenBank</i>	Referência
>P22BRToJQ952600	>RdRpBRToJQ952600	-	ToC-Br2	Brasil	JQ952600	Albuquerque et al. (2013)
-	-	>HspBRToJQ952601	ToC-Br2	Brasil	JQ952601	Albuquerque et al. (2013)
-	-	>HspChToKP217201	HBCZ	China	KP217201	Gao (2016)
-	-	>HspChToKP217202	HBLF	China	KP217202	Gao (2016)
-	-	>HspChToKP264986	HNAYHX	China	KP264986	Não publicado
-	-	>HspChToKP264987	HNZZZM1	China	KP264987	Não publicado
-	-	>HspChToKP264988	HNZZZM2	China	KP264988	Não publicado
-	-	>HspChToKX880384	JLCC	China	KX880384	Não publicado
>P22ChToKF018280	>RdRpChToKF018280	-	NJ	China	KF018280	Karwitha et al. (2014)
-	-	>HspChToKC812626	SDLC	China	KC812626	Zhao LM et al. (2014)
>P22ChToKC709509	>RdRpChToKC709509	-	SDSG	China	KC709509	Zhao et al. (2015)
-	-	>HspChToKC709510	SDSG	China	KC709510	Zhao LM et al. (2014)
-	-	>HspChToKC812624	SDTADP	China	KC812624	Zhao LM et al. (2014)
-	-	>HspChToKC812625	SDT AFC	China	KC812625	Zhao LM et al. (2014)
-	-	>HspChToKC812627	SDTAMZ	China	KC812627	Zhao LM et al. (2014)
-	-	>HspChToKX853539	SXJZ	China	KX853539	Não publicado
-	-	>HspChToKP226698	TJ6	China	KP226698	Não publicado
>P22ChToKC887998	>RdRpChToKC887998	-	ToCV-BJ	China	KC887998	Zhao R et al. (2014)
-	-	>HspChToKC887999	ToCV-BJ	China	KC887999	Zhao R et al. (2014)
-	-	>HspChToKR072214	zb	China	KR072214	Não publicado
-	>RdRpCSToKP114530	-	HP	Coréia do Sul	KP114530	Não publicado
-	-	>HspCSToKP114537	HP	Coréia do Sul	KP114537	Não publicado
>P22CSToKP137098	>RdRpCSToKP137098	-	HS	Coréia do Sul	KP137098	Não publicado
-	-	>HspCSToKP137099	HS	Coréia do Sul	KP137099	Não publicado
-	>RdRpCSToKP114535	-	IS17	Coréia do Sul	KP114535	Não publicado
-	-	>HspCSToKP114525	IS17	Coréia do Sul	KP114525	Não publicado
>P22CSToKP114538	>RdRpCSToKP114538	-	IS29	Coréia do Sul	KP114538	Não publicado
-	-	>HspCSToKP114529	IS29	Coréia do Sul	KP114529	Não publicado
>P22CSToKP137100	>RdRpCSToKP137100	-	JJ	Coréia do Sul	KP137100	Não publicado
P22CSToKP114532	>RdRpCSToKP114532	-	JJ3	Coréia do Sul	KP114532	Não publicado
-	-	>HspCSToKP114533	JJ3	Coréia do Sul	KP114533	Não publicado
-	>RdRpCSToKP114527	-	JJ5	Coréia do Sul	KP114527	Não publicado
-	-	>HspCSToKP114534	JJ5	Coréia do Sul	KP114534	Não publicado
-	>RdRpCSToKP114531	-	JN1	Coréia do Sul	KP114531	Não publicado

-	-	>HspCSToKP114536	JN1	Coréia do Sul	KP114536	Não publicado
-	>RdRpCSToKP114526	-	YG	Coréia do Sul	KP114526	Não publicado
-	-	>HspCSToKP114528	YG	Coréia do Sul	KP114528	Não publicado
-	>RdRpESToKJ200304	-	2.5	Espanha	KJ200304	Não publicado
-	-	>HspESToKJ200305	2.5	Espanha	KJ200305	Não publicado
>P22ESToKP668820	-	-	AT103/09	Espanha	KP668820	Landeo-Rios et al. (2015)
>P22ESToKP668842	-	-	AT146/06	Espanha	KP668842	Landeo-Rios et al. (2015)
>P22ESPiKP668832	-	-	AT152/08	Espanha	KP668832	Landeo-Rios et al. (2015)
>P22ESPiKP668831	-	-	AT155/08	Espanha	KP668831	Landeo-Rios et al. (2015)
>P22ESToKP668837	-	-	AT158/06	Espanha	KP668837	Landeo-Rios et al. (2015)
>P22ESToKP668834	-	-	AT189/06	Espanha	KP668834	Landeo-Rios et al. (2015)
>P22ESToKP668818	-	-	AT2/09	Espanha	KP668818	Landeo-Rios et al. (2015)
>P22ESToKP668850	-	-	AT44/09	Espanha	KP668850	Landeo-Rios et al. (2015)
>P22ESToKP668840	-	-	AT60/09	Espanha	KP668840	Landeo-Rios et al. (2015)
>P22ESToKP668839	-	-	AT66/08	Espanha	KP668839	Landeo-Rios et al. (2015)
-	-	>HspESToDQ136146	AT80/99	Espanha	DQ136146	Lozano et al. (2006)
>P22ESToDQ983480	>RdRpESToDQ983480	-	AT80/99	Espanha	DQ983480	Lozano et al. (2007)
-	>RdRpESToKJ740256	-	AT80/99-IC	Espanha	KJ740256	Orilio et al. (2014)
-	-	>HspESToKJ740257	AT80/99-IC	Espanha	KJ740257	Orilio et al. (2014)
-	>RdRpESPiKJ200306	-	MM8	Espanha	KJ200306	Não publicado
-	-	>HspESPiKJ200307	MM8	Espanha	KJ200307	Não publicado
-	>RdRpESToKJ200308	-	PI-1-2	Espanha	KJ200308	Não publicado
-	-	>HspESToKJ200309	PI-1-2	Espanha	KJ200309	Não publicado
-	>RdRpUSToAY903447	-	Florida	Estados Unidos	AY903447	Wintermantel et al. (2005)
-	-	>HspUSToAY903448	Florida	Estados Unidos	AY903448	Wintermantel et al. (2005)
>P22USToNC_007340	>RdRpUSToNC_007340	-	Florida	Estados Unidos	NC_007340	Wintermantel et al. (2005)
-	-	>HspUSToNC_007341	Florida	Estados Unidos	NC_007341	Wintermantel et al. (2005)
-	-	>HspGRToEU284744	Gr-535	Grécia	EU284744	Kataya et al. (2008)
>P22GRToEU284745	>RdRpGRToEU284745	-	Gr-535	Grécia	EU284745	Kataya et al. (2008)

