LARISSA GOULART ZANARDO

CARACTERIZAÇÃO BIOLÓGICA, MOLECULAR E ANÁLISE DA VARIABILIDADE GENÉTICA DE *Cowpea mild mottle virus* (CPMMV) EM SOJA NO BRASIL

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2013

Ficha catalográfica preparada pela Seção de Catalogação e Classificação da Biblioteca Central da UFV

T

Z27c
2013

Zanardo, Larissa Goulart, 1987-
Caracterização biológica, molecular e análise da variabilidade genética de *Cowpea mild mottle virus* (CPMMV) em soja no Brasil / Larissa Goulart Zanardo. – Viçosa, MG, 2013.
vii, 93f. : il. (algumas color.) ; 29cm.

Texto em inglês
Orientador: Francisco Murilo Zerbini Júnior
Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. *Cowpea mild mottle virus*. 2. *Carlavirus*. 3. Vírus de plantas. 4. Soja - Doenças e pragas. 5. Filogenia. 6. Recombinação (Genética). I. Universidade Federal de Viçosa. Departamento de Fitotecnia. Programa de Pós-Graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 579.28

LARISSA GOULART ZANARDO

CARACTERIZAÇÃO BIOLÓGICA, MOLECULAR E ANÁLISE DA VARIABILIDADE GENÉTICA DE *Cowpea mild mottle virus* (CPMMV) EM SOJA NO BRASIL

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 25 de fevereiro de 2013.

Fábio Nascimento da Silva

Poliane Alfenas Zerbini

Claudine Márcia Carvalho
(Coorientadora)

Francisco Murilo Zerbini Júnior
(Orientador)

AGRADECIMENTOS

À Deus por guiar meus passos e abençoar a minha vida.

Aos meus pais e irmãos pelo amor, compreensão e apoio incondicional em todos os momentos.

À Universidade Federal de Viçosa (UFV) pela oportunidade de realização deste curso.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de estudos.

Ao meu orientador o Prof. Francisco Murilo Zerbini, pela orientação e oportunidade, pelos ensinamentos constantes, pela amizade e receptividade durante o meu mestrado.

À Prof^{a} Claudine Márcia Carvalho pela dedicação e confiança, pela orientação valiosa, pela oportunidade de aprendizado e amizade durante a realização dos trabalhos.

Ao Bruno pelo apoio, dedicação, amor e compreensão durante essa etapa.

Aos pesquisadores Alison Talis Martins Lima, Fábio Nascimento Silva e Gloria Patricia Castillo Urquiza os meus sinceros agradecimentos pela valiosa colaboração, pela ajuda constante e amizade.

Aos meus amigos de laboratório Adriana, Camila, Chaianne, Daniel, David, Diego, Hermano, Lenin, Marcelo, Márcio, Marcos, Pedro, Renan, Roberto, Sílvia e Tathiana pela amizade, brincadeiras, companheirismo e ajuda. Em especial aos amigos: André, César, Diogo e Fernanda.

A todos que contribuíram e torceram por mim durante toda essa etapa, minha sincera gratidão.

SUMÁRIO

RESUMO

ZANARDO, Larissa Goulart, M.Sc., Universidade Federal de Viçosa, Fevereiro de 2013. **Caracterização biológica, molecular e análise da variabilidade genética de *Cowpea mild mottle virus* (CPMMV) em soja no Brasil.** Orientador: Francisco Murilo Zerbini Júnior. Coorientadoras: Claudine Márcia Carvalho e Gloria Patricia Castillo Urquiza.

A partir do ano 2000 plantas de soja nos campos dos estados da Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais, Paraná e Tocantins foram descritas apresentando sintomas da doença da necrose da haste da soja. Os sintomas eram variados, alguns mais suaves outros mais severos. A doença foi associada ao *Cowpea mild mottle virus* (CPMMV, família *Betaflexiviridae*, gênero *Carlavirus*). Nesse estudo foi proposta a realização da caracterização biológica, molecular e análise da variabilidade genética de isolados de CPMMV, causando sintomas da necrose da haste em soja nos campos de diferentes estados produtores do Brasil. O estudo foi realizado com amostras coletadas nos estados da Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais e Pará. Os isolados causaram uma variedade de sintomas em soja cv. CD206, isolados brandos e severos foram observados. O genoma completo de 6 isolados foi sequenciado e adicionalmente a sequencia parcial de outros 12 isolados foi também determinada (ORF2-3'terminal). Nenhum isolado brasileiro de CPMMV, independente do hospedeiro, havia sido totalmente sequenciado até esse trabalho. As caracterizações biológica e molecular mostraram que os seis isolados brasileiros, cujos genomas foram completamente determinados, pertencem a uma nova estirpe de CPMMV, distinta daquela à qual pertence o único isolado de CPMMV previamente sequenciado, oriundo de Gana na Àfrica. A ORF1 (RdRp) desses seis isolados brasileiros apresentou valores de identidade de sequencia (60-61% para nt e 58-69% para aa), inferiores ao estabelecido pelo Comitê Internacional de Taxonomia de Vírus (ICTV), quando foram comparados com o isolado Gana de CPMMV. A ORF5 (CP), no entanto, apresentou valores de identidade (79% para nt e 95-96% para aa) superiores ao estabelecido pelo ICTV quando foram comparados com o isolado Gana. Ambas as proteínas são utilizadas para classificar isolados do gênero *Carlavirus* em uma mesma espécie. As comparações par-a-par e análises filogenéticas mostraram que os isolados brasileiros são altamente relacionados entre si e distintos de isolados de outras espécies do gênero *Carlavirus*. As árvores filogenéticas construídas com as sequencias parciais dos genomas não mostraram agrupamentos com base em região geográfica ou ano de coleta, porém agrupamentos com base nos sintomas foram observados para as árvores construídas com a sequência parcial (ORF2-3'terminal), ORF2 (TGB1), ORF5 (CP) e ORF6 (NABP). Além do relacionamento entre os isolados de CPMMV, foi demonstrado através do sequenciamento parcial, que existem variações entre os isolados brasileiros. Evidencias de duas possíveis estirpes de CPMMV no Brasil foram encontradas, com variações moleculares e biológicas entre os isolados de ambas as estirpes. Eventos de recombinação foram identificados ao longo do genoma dos isolados, e eles ocorreram principalmente na ORF1, região da polimerase, e com menor frequência em outras regiões genoma. Com esse trabalho foi verificado que o critério taxonômico, que define as espécies do gênero *Carlavirus*, pode ser falho em casos em que apenas a sequencia parcial é determinada, se apenas a ORF1 tivesse sido determinada durante o estudo poderíamos propor que os nossos isolados pertencem à uma nova espécie do gênero *Carlavirus.* Além disso, ficou clara a necessidade de se determinar a ocorrência de transmissão por semente dos isolados de CPMMV brasileiros, pois a transmissão por sementes e/ou a alta capacidade de dispersão e voo da mosca branca *Bemisia tabaci* podem ter contribuído para a dispersão do vírus nos diferentes estados produtores de soja. Esses fatos justificam o não agrupamento dos isolados com base em região geográfica ou ano de coleta.

ABSTRACT

ZANARDO, Larissa Goulart, M.Sc., Universidade Federal de Viçosa, February 2013. **Biological characterization, molecular and analysis of genetic variability of *Cowpea mild mottle virus* (CPMMV) in soybean in Brazil.** Advisor: Francisco Murilo Zerbini Júnior. Co-advisers: Claudine Márcia Carvalho and Gloria Patricia Castillo Urquiza.

Since 2000, field soybean plants in the states of Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais, Paraná and Tocantins have been described with symptoms of soybean stem necrosis disease. The symptoms are varied, some milder other severe. The disease has been associated with *Cowpea mild mottle virus* (CPMMV, family *Betaflexiviridae*, genus *Carlavirus*). This study is aimed at the biological characterization and molecular and genetic analyses of genetic variability of isolates of CPMMV that cause symptoms of stem necrosis in soybean fields of different producing Brazil states. The study was conducted with samples collected in the states of Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais and Pará. The isolates caused a variety of symptoms in soybean cv. CD206, and mild and severe isolates were observed. The complete genomes of six isolates were sequenced and additionally the partial sequence of another 12 isolates was also determined (ORF2-3'end). No Brazilian CPMMV isolate, from any host, had been entirely sequenced until now. Biological and molecular characterization showed that the six Brazilian isolates, whose genomes have been completely determined, belong to a new CPMMV strain distinct from that which belongs to the only isolated CPMMV previously sequenced from Ghana in Africa. The ORF1 (RdRp) of these six Brazilian isolates showed values of sequence identity (60-61% to 58-69% for nt and aa), less than that set by the International Committee on Taxonomy of Viruses (ICTV), when they were compared with a Ghanaian CPMMV isolate. ORF5 (CP), however, showed identity values (79% to 95-96% for nt and aa) greater than those established by the ICTV when they were compared with the Ghanaian isolate. Both proteins are used to classify isolates of the genus *Carlavirus* in the same species. Pairwise comparisons and phylogenetic analysis showed that Brazilian isolates are highly related to each other and distinct from isolates of other species of the genus *Carlavirus*. The phylogenetic trees constructed with partial sequences of the genomes did not show groupings based on geographic region or collection year, but groupings based on symptoms were observed for trees constructed with partial sequences (ORF2-3'end), ORF2 (TGB1), ORF5 (CP) and ORF6 (NABP). Besides the relationship between isolates CPMMV demonstrated by partial sequencing, there are variations between the Brazilian isolates. Evidence of two possible CPMMV strains were found in Brazil with biological and molecular variations between isolates of both strains. Recombination events were identified in genome of the isolates, and they occurred mainly in the ORF1 region of the polymerase, less frequently in other regions of the genome. With this study it was found that the taxonomic criteria, which define the genus *Carlavirus* may fail in cases where only a partial sequence is determined: if only the ORF1 had been determined during the study we could propose that our isolates belong to a new species of the genus *Carlavirus*. In addition, there is clearly the need to determine the occurrence of seed transmission in Brazilian CPMMV isolates, since the transmission by seed and/or whitefly *Bemisia tabaci*, with high dispersibility may have contributed to the spread of the virus in different soybean producing states. These facts justify the non-grouping of isolates based on geographic region or collection year.

INTRODUÇÃO GERAL

A família *Betaflexiviridae* é constituída por seis gêneros: *Capillovirus*, *Carlavirus*, *Citrivirus*, *Foveavirus*, *Trichovirus* e *Vitivirus* (King *et al.*, 2011). Vírions da família *Betaflexiviridae* são filamentosos com 12-13nm de diâmetro e 600-1000 nm de comprimento, variando de acordo com o gênero. São monopartidos, com genoma de RNA fita simples sentido positivo (ssRNA). Possuem na extremidade 5' terminal um 'cap' (m7GpppG) e na extremidade 3' terminal a cauda Poli(A). A tradução das fases abertas de leitura ('Open Reading Frames' - ORFs), é feita através de mRNAs subgenômicos, e até seis proteínas podem ser geradas desde a extremidade 5' até a 3' do genoma viral (Adams *et al.*, 2004; Martelli *et al.*, 2007; King *et al.*, 2011). Há uma grande diversidade no número e na natureza dos genes virais que são expressos próximos a região 3' através de mRNAs subgenômicos. Estes genes codificam proteínas do movimento viral como as da superfamília p30-like (gêneros *Capillovirus*, *Citrivirus*, *Trichovirus* e *Vitivirus*), proteínas do bloco triplo de genes (TGB) (gêneros *Carlavirus* e *Foveavirus*), proteínas de ligação a ácidos nucléicos (NABP), a proteína do capsídeo (CP) e outras proteínas encontradas em alguns membros da família (Figura 1) (Martelli *et al.*, 2007; King *et al.*, 2011). Apesar da variedade de proteínas encontradas para os gêneros dessa família a RNA polimerase dependente de RNA (RdRp) e a proteína do capsídeo (CP) formam uma unidade taxonômica coerente (Adams *et al.*, 2004).

A organização genômica típica varia entre os gêneros (Figura 1), e até entre espécies de um mesmo gênero. A primeira e maior ORF é sempre o gene da replicase, cujo tamanho varia de 190 a 250 kDa (King *et al.*, 2011). Nesta sequência está contido os domínios metiltransferase, RNA helicase e RNA polimerase dependente de RNA comuns a todos os vírus pertencentes á família (Figura 1)

(Koonin & Dolja, 1993; King *et al.*, 2011). Em seguida estão as ORFs responsáveis por codificar proteínas relacionadas ao movimento viral. Dois gêneros da família *Betaflexiviridae*, *Carlavirus* e *Foveavirus*, possuem um conjunto de três ORFs (ORF2, 3 e 4), que se sobrepõem parcialmente, conhecido como bloco triplo de genes (‘triple gene block’ - TGB), envolvido no movimento célula-á-célula e a longa distância (Morozov & Solovyev, 2003; Martelli *et al.*, 2007; King *et al.*, 2011). Adicionalmente, foi demonstrado que a TGBp1 (ORF2) do *Potato virus* M (PVM, gênero *Carlavirus*) suprime o silenciamento sistêmico (Senshu *et al.*, 2011). Os outros quatro gêneros possuem uma única ORF que codifica uma proteína da superfamília 30K. A última ou penúltima ORF (ORF5) é a responsável por codificar a proteína do capsídeo (CP), já que alguns gêneros, *Carlavirus*, *Trichovirus* e *Vitivirus,* o genoma viral apresenta uma sexta ORF, cujo produto é uma proteína rica em cisteína (‘cysteine-rich protein’ - CRP) com atividade de ligação à ácido nucleico (‘nucleic acid-binding activity’ - NABP) (Martelli *et al.*, 2007; King *et al.*, 2011). Foi demonstrado para o *Grapevine virus A* (GVA, gênero *Vitivirus*) (Zhou *et al.*, 2006) e para o PVM que a proteína NABP ou CRP é capaz de suprimir o silenciamento de RNA sistêmico e local (Senshu *et al.*, 2011).

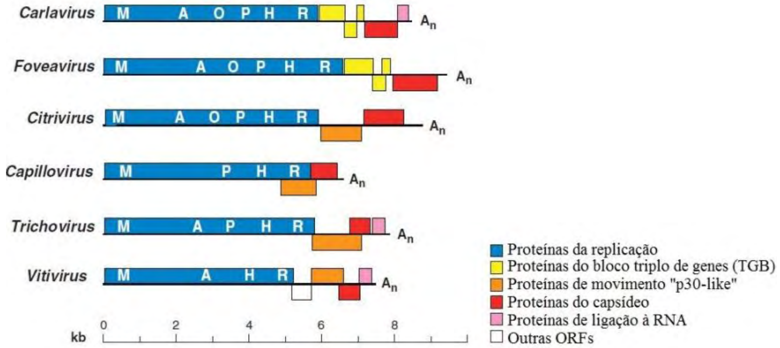


Figura 1: Representação esquemática dos genomas de espécies pertencentes aos gêneros que compõe a família *Betaflexiviridae*. As barras coloridas representam as proteínas codificadas por cada uma das possíveis fases de leitura aberta (‘Open Reading Frame’ - ORFs). Os genomas virais podem atingir até 9000 nucleotídeos e seis ORF’s. Domínios da replicase: M-metiltransferase, A-alkB, O-otu-like peptidase, P-papain-like protease, H-RNA helicase e R-RNA polimerase dependente de RNA. Reproduzido de Martelli et al. (2007) com modificações.

Os flexivírus de modo geral infectam uma variedade de hospedeiros selvagens e cultivados incluindo plantas herbáceas, dicotiledôneas lenhosas e com menos frequência monocotiledôneas (Martelli *et al.*, 2007). Os capillovírus, citrivírus, foveavírus, trichovírus e os vitivírus tendem a infectar plantas lenhosas, já os carlavírus infectam preferencialmente plantas herbáceas, podendo ser assintomáticos ou sintomáticos, com o mosaico um dos sintomas principais (Martelli *et al.*, 2007). Os membros da família *Betaflexiviridae* tem sido descritos em vários hospedeiros, mas a gama de hospedeiros é restrita para membros individuais de cada gênero (King *et al.*, 2011).

Todas as espécies da família podem ser transmitidas via inoculação mecânica. Os capillovírus, citrivírus e foveavírus não possuem vetor conhecido, porém são capazes de ser transmitidos via enxertia, material propagativo e apenas os capillovírus via sementes. Os carlavírus podem ser transmitidos por afídeos, na grande maioria, ou por mosca branca, no caso do *Cowpea mild mottle virus* (CPMMV) (Naidu *et al.*, 1998; Almeida *et al.*, 2005; King *et al.*, 2011), do *Melon yellowing-associated virus* (MYaV) (Nagata *et al.*, 2005) e do *Cucumber vein-clearing virus* (Menzel *et al.*, 2011). Os trichovírus são também transmitidos por material propagativo e enxertia e são os únicos transmitidos por ácaros (King *et al.*, 2011). Já os vitivírus podem ser transmitidos por material propagativo, cochonilhas e afídeos (King *et al.*, 2011).

Os carlavírus replicam-se nas células parênquimais do hospedeiro e por isso são mais facilmente transmitidos que os vírus restritos ao floema (Martelli *et al.*, 2007). Os sintomas induzidos por carlavírus variam de muito leve à grave, dependendo da variedade das plantas, das condições ambientais, das espécies ou estirpes de vírus, como verificado em especial para viroses da batata (Nie *et al.*, 2008). Muitas vezes pode ser observado a ocorrência de infecções assintomáticas (Martelli *et al.*, 2007).

No Brasil entre as espécies do gênero *Carlavirus* encontradas estão o *Potato virus S* propagado em batata (*Solanum tuberosum*) (Gaspar *et al.*, 2008; Duarte *et al.*, 2012), *Cole latent virus* que infecta couve (Gaspar *et al.*, 2008), o *Cowpea mild mottle virus* (CPMMV) em plantas de feijão (Costa *et al.*, 1983) e soja (Almeida *et al.*, 2003; Almeida *et al.*, 2005; Almeida, 2008; Gaspar *et al.*, 2008), *Garlic common latent virus* em alho (Fajardo *et al.*, 2001), *Lily symptomless virus* em lírio (Rivas, 2010) e o *Melon yellowing-associated virus* (MYaV) em plantas de melão (Nagata *et al.*, 2005). Até o ano de 2005 o CPMMV era considerado o único carlavírus transmitido pela mosca branca *Bemisia tabaci* (Munyappa & Reddy, 1983; Naidu *et al.*, 1998; Adams *et al.*, 2004). Nesse ano, porém o MYaV foi descrito como também sendo transmitido por esse vetor (Nagata *et al.*, 2005). A incidência e os danos causados por *B. tabaci* aumentaram exponencialmente a partir da década de 70, em associação ao grande aumento da área plantada com soja especialmente porque a soja é um excelente hospedeiro de *B. tabaci*. A não adoção de medidas de controle permite que as populações de insetos atinjam níveis altíssimos.

Na safra de 2000/01 plantas de soja na região de Morrinhos e Goiatuba, estado de Goiás, apresentavam sintomas de nanismo, queima do broto e necrose da haste. As necroses eram severas e levavam as plantas à morte. Os sintomas foram associados à necrose da haste, cujo agente etiológico é o CPMMV. Novamente, em 2001/02 sintomas similares apareceram em plantas de soja em Barreiras, estado da Bahia. Em 2002, um novo surto ocorreu devastando os campos do estado de Goiás, nas regiões de Acreúna, Quirinópolis e Porteirão. Cerca de 1000 ha apresentaram perdas totais, ocasionando um prejuízo de 600 mil dólares. Na safra 2002/2003 o

CPMMV foi encontrado em Sorriso (MT), Balsas (MA), Palotina (PR) e Goiânia, Luziânia e Vianópolis (GO) (Almeida *et al.*, 2003; Almeida *et al.*, 2005).

Os hospedeiros naturais do CPMMV incluem espécies da família *Fabaceae*. Sua ocorrência tem sido relatada na Ásia (Irã, Israel, Taiwan e Tailândia) (Antignus & Cohen, 1987; Tavasoli *et al.*, 2009), África (Costa do Marfim, Gana, Nigéria, Tanzânia) (Brunt & Kenten, 1973; Thouvenel *et al.*, 1982; Mink & Keswani, 1987; Menzel *et al.*, 2010), Brasil (Costa *et al.*, 1983; Almeida *et al.*, 2005). Na Argentina foi relatada a ocorrência em feijão e soja na Província de Salta (Rodríguez-Pardina *et al.*, 2004; Laguna *et al.*, 2006). Recentemente, foi encontrado infectando feijão de corda (*Vigna unguiculata* subsp. *Sesquipedalis*) na Venezuela (Brito *et al.*, 2012). No Brasil o vírus já foi encontrado em diversos estados produtores de soja: Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais, Paraná e Tocantins (Almeida, 2008). Entre os estados produtores apenas o Rio Grande do Sul ainda não apresentou relatos da doença.

Os sintomas causados por carlavírus variam de acordo com o hospedeiro que está sendo infectado e a época do ano. O CPMMV em feijão caupi (*Vigna unguiculata*) causa manchas cloróticas nas folhas primárias e distorção das folhas, em amendoim provoca lesões necróticas, anéis cloróticos, com posterior clorose sistêmica (Brunt & Kenten, 1973). Em soja e feijão comum (*Phaseolus vulgaris*) pode causar uma diversidade de sintomas como: clorose e mosaico nas folhas, necrose apical, distorção e nanismo (Brunt & Kenten, 1973; Iwaki *et al.*, 1982; Almeida *et al.*, 2003; Almeida, 2008; Tavasoli *et al.*, 2009). Em campo, os sintomas iniciais da virose da soja não chamam a atenção dos produtores, apenas próximo à floração e surgimento das vagens é que eles tornam-se mais evidentes (Almeida, 2008). É após esse período que são observadas a queima do broto e a necrose das hastes, que terminam na morte das plantas e consequente perda de produção (Almeida, 2008).

Alguns isolados causam sintomas severos (necroses) outros já mais brandos (mosaico e bolhas no limbo), podendo às vezes ser assintomáticos. Os sintomas distintos apresentados pelos isolados de CPMMV sugerem a existência de estirpes distintas ou mesmo mais de uma espécie de carlavírus.

A variabilidade genética para o CPMMV ainda não foi estudada, nem sua história evolutiva. Na verdade poucos isolados foram sequenciados parcialmente, e apenas um isolado teve seu genoma completamente determinado (Menzel *et al.*, 2010). A maior parte dos trabalhos envolvendo o CPMMV relata a identificação viral, sua caracterização biológica, o sequenciamento parcial do genoma e a identificação do vetor (Brunt & Kenten, 1973; Iwaki *et al.*, 1982; Thouvenel *et al.*, 1982; Costa *et al.*, 1983; Munyappa & Reddy, 1983; Jeyanandarajah & Brunt, 1993; Almeida *et al.*, 2005; Laguna *et al.*, 2006; Marubayashi *et al.*, 2010). Apesar de ter sido identificado há cerca de 40 anos atrás (Brunt & Kenten, 1973), muito pouco se sabe sobre esse vírus no Brasil e no mundo, porém o vírus tem sido encontrado com frequência nos campos brasileiros de soja. Levando em consideração que o Brasil é o segundo maior produtor de soja do mundo, que o vírus é disseminado pela mosca branca *B. tabaci* e que já foi encontrado em diversos campos de estado produtores de soja, maiores perdas econômicas poderão ocorrer se o vírus continuar a ser disseminado pelos campos de soja.

Evidências de variação genética em vírus de plantas foram reportadas em 1926, quando se observou possivelmente que diferentes variantes causariam diferentes sintomas, citado por (Garcia-Arenal *et al.*, 2001). As variações genéticas surgem de erros que ocorrem durante a replicação dos genomas virais (Garcia-Arenal *et al.*, 2001). As mutações representam a fonte primária de variação genética em que a seleção natural e a deriva genética atuam (Drake & Holland, 1999). Nesse contexto, a mutação corresponderiam ao processo de incorporação de nucleotídeos na fita recém formada, que durante a replicação do ácido nucléico estão ausentes na fita molde (Garcia-Arenal *et al.*, 2001). As taxas de mutação são importantes para o entendimento da estrutura genética de populações ao longo do tempo e consequentemente o curso da evolução. Os vírus de RNA possuem taxas de mutação superiores as determinadas para os vírus de DNA (Drake & Holland, 1999). Para vírus de DNA a taxa de mutação estimada é de 10^{-8} a 10^{-6} substituições por nucleotídeo por infecção celular (s/n/c) e para vírus de RNA a taxa de mutação estimada é de 10^{-6} a 10^{-4} s/n/c (Sanjuan *et al.*, 1999).

Novos variantes virais podem surgir também a partir da recombinação. No contexto da infecção viral, a recombinação seria o processo pelo qual segmentos do genoma são trocados entre segmentos de nucleotídeos de diferentes variantes genéticos durante o processo de replicação, resultando em trocas genéticas (Garcia-Arenal *et al.*, 2001). Compreender o papel da recombinação na evolução dos vírus de planta é altamente importante. Há evidências crescentes sobre a contribuição da recombinação da variabilidade genética em vírus de RNA, isso inclui uma variedade de vírus de senso positivo. As taxas de recombinação variam amplamente e as causas dessas variações ainda não são claras (Chare & Holmes, 2006).

Conhecer as características moleculares e biológicas de um dado vírus e a sua variabilidade genética é altamente importante, pois essas informações permitem conhecer características do patógeno que podem contribuir para a adoção de medidas preventivas contra a virose no campo. Assim, o estudo realizado nesse trabalho visa caracterizar biológica e molecularmente isolados de CPMMV obtidos de soja, e avaliar a variabilidade genética do CPMMV.

LITERATURA CITADA

ADAMS, M.J., ANTONIW, J.F., BAR-JOSEPH, M., BRUNT, A.A., CANDRESSE, T., FOSTER, G.D., MARTELLI, G.P., MILNE, R.G., ZAVRIEV, S.K. & FAUQUET, C.M. The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. Archives of Virology 149:1045-1060. 2004.

ALMEIDA, A.M.R. Viroses da soja no Brasil: sintomas, etiologia, controle. Série Documentos 306:1-62. 2008.

ALMEIDA, A.M.R., PIUGA, F.F., KITAJIMA, E.W., GASPAR, J.O., VALENTIN, N., BENATO, L.C., MARIN, S.R.R., BINECK, E., BELINTANI, P., NUNES JUNIOR, J., HOFFMANN, L. & MEYER, M.C. Necrose da haste da soja. Série Documentos 221:1-48. 2003.

ALMEIDA, A.M.R., PIUGA, F.F., MARIN, S.R.R., KITAJIMA, E.W., GASPAR, J.O., OLIVEIRA, T.G.D. & MORAES, T.G.D. Detection and partial characterization of a carlavirus causing stem necrosis of soybean in Brazil. Fitopatologia Brasileira 30:191-194. 2005.

ANTIGNUS, Y. & COHEN, S. Purification and some properties of a new strain of *Cowpea mild mottle virus* in Israel. Annals of Applied Biology 110:563-569. 1987.

BRITO, M., FERNANDEZ-RODRIGUEZ, T., GARRIDO, M.J., MEJIAS, A., ROMANO, M. & MARYS, E. First report of *Cowpea mild mottle Carlavirus* on yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis*) in Venezuela. Viruses 4:3804-3811. 2012.

BRUNT, A.A. & KENTEN, R.H. *Cowpea mild mottle*, a newly recognized virus infecting cowpeas (*Vigna unguiculata*) in Ghana. Annals of Applied Biology 74:67-74. 1973.

CHARE, E.R. & HOLMES, E.C. A phylogenetic survey of recombination frequency in plant RNA viruses. Archives of Virology 151:933-946. 2006.

COSTA, A.S., GASPAR, J.O. & VEGA, J. Mosaico angular do feijão jalo causado por um carlavírus transmitido pela mosca branca *Bemisia tabaci*. Fitopatologia Brasileira 8:325-327. 1983.

DRAKE, J.W. & HOLLAND, J.J. Mutation rates among RNA viruses. Proceedings of the National Academy of Sciences 96:13910–13913. 1999.

DUARTE, P.S.G., GALVINO-COSTA, S.B.F., RIBEIRO, S.R.R. & FIGUEIRA, A.R. Complete genome sequence of the first Andean strain of *potato virus S* from

Brazil and evidence of recombination between PVS strains. Archives of Virology 157:1357-1364. 2012.

FAJARDO, T.V.M., NISHIJIMA, M., BUSO, J.A., TORRES, A.C., ÁVILA, A.C. & RESENDE, R.O. Garlic viral complex: identification of Potyviruses and Carlavirus in Central Brazil. Fitopatologia Brasileira 26:619-626. 2001.

GARCIA-ARENAL, F., FRAILE, A. & MALPICA, J.M. Variability and genetic structure of plant virus populations. Annual Review Phytopathology 39:157-186. 2001.

GASPAR, J.O., BELINTANI, P., ALMEIDA, A.M. & KITAJIMA, E.W. A degenerate primer allows amplification of part of the 3'-terminus of three distinct carlavirus species. Journal of Virological Methods 148:283-285. 2008.

IWAKI, M., THONGMEEARKON, P., PROMMIN, M., HONDA, Y. & HIBI, J. Whitefly transmission and some properties of *Cowpea mild mottle virus* on soybean in Thailand. Plant Disease 66:265-268. 1982.

JEYANANDARAJAH, P. & BRUNT, A.A. The natural occurrence, transmission, properties and possible affinities of *Cowpea mild mottle virus*. Journal of Phytopathology 137:148-156. 1993.

KING, A.M.Q., ADAMS, M.J., CARSTENS, E.B. & LEFKOWITZ, E.J. (Eds.) Virus taxonomy. Ninth Report of the International Comittee on Taxonomy of Viruses. Elsevier Inc. 1272p. 2011.

KOONIN, E.V. & DOLJA, V.V. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Critical Reviews in Biochemistry and Molecular Biology 28:375-430. 1993.

LAGUNA, I.G., ARNEODO, J.D., RODRÍGUEZ-PARDINA, P. & FIORONA, M. *Cowpea mild mottle virus* infecting soybean crops in northwestern Argentina. Fitopatologia Brasileira 31:317-317. 2006.

MARTELLI, G.P., ADAMS, M.J., KREUZE, J.F. & DOLJA, V.V. Family *Flexiviridae*: a case study in virion and genome plasticity. Annual Review Phytopathology 45:73-100. 2007.

MARUBAYASHI, J.M., YUKI, V.A. & WUTKE, E.B. Transmissão do *Cowpea mild mottle virus* pela mosca branca *Bemisia tabaci* biótipo B para plantas de feijão e soja. Summa Phytopathologica 36:158-160. 2010.

MENZEL, W., ABANG, M.M. & WINTER, S. Characterization of *Cucumber vein-clearing virus*, a whitefly (Bemisia tabaci G.)-transmitted carlavirus. Archives of Virology 156:2309-2311. 2011.

MENZEL, W., WINTER, S. & VETTEN, H. Complete nucleotide sequence of the type isolate of *Cowpea mild mottle virus* from Ghana. Archives of Virology 155:2069-2073. 2010.

MINK, G.I. & KESWANI, C.L. First report of *Cowpea mild mottle virus* on bean and mung bean in Tanzania. Plant Disease 71:557. 1987.

MOROZOV, S.Y. & SOLOVYEV, A.G. Triple gene block: modular design of a multifunctional machine for plant virus movement. Journal of General Virology 84:1351-1366. 2003.

MUNYAPPA, V. & REDDY, D.V.R. Transmission of *Cowpea mild mottle virus* by *Bemisia tabaci* in a nonpersistent manner. Plant Disease 67:391-393. 1983.

NAGATA, T., ALVES, D.M., INOUE-NAGATA, A.K., TIAN, T.Y., KITAJIMA, E.W., CARDOSO, J.E. & DE AVILA, A.C. A novel melon flexivirus transmitted by whitefly. Archives of Virology 150:379-387. 2005.

NAIDU, R.A., GOWDA, S., SATYANARAYANA, T., BOYKO, V., REDDY, A.S., DAWSON, W.O. & REDDY, D.V. Evidence that whitefly-transmitted *Cowpea mild mottle virus* belongs to the genus *Carlavirus*. Archives of Virology 143:769-780. 1998.

NIE, X., BAI, Y., MOLEN, T.A. & DESJARDINS, D.C. Development of universal primers for detection of potato carlaviruses by RT-PCR. Journal of Virological Methods 149:209-216. 2008.

RIVAS, E.B. *Lily simptomless virus* no Brasil. Documento técnico 006 - Instituto Biológico - Agencia Paulista Tecnológica dos Agronegócios - APTA 1-5. 2010.

RODRÍGUEZ-PARDINA, P.E., ARNEODO, J.D., TRUOL, G.A., HERRERA, P.S. & LAGUNA, I.G. First Record of CpMMV in bean crops in Argentina Australasian Plant Pathology 33:129-130. 2004.

SANJUAN, R., NEBOT, M.R., CHIRICO, N., MANSKY, L.M. & BELSHAW, R. Viral mutation rates. Journal of Virology 84:9733-9748. 1999.

SENSHU, H., YAMAJI, Y., MINATO, N., SHIRAISHI, T., MAEJIMA, K., HASHIMOTO, M., MIURA, C., NERIYA, Y. & NAMBA, S. A dual strategy for the suppression of host antiviral silencing: two distinct suppressors for viral replication and viral movement encoded by *Potato virus M*. Journal of Virology 85:10269-10278. 2011.

TAVASOLI, M., SHAHRAEEN, N. & GHORBANI, S. Serological and RT-PCR detection of *Cowpea mild mottle Carlavirus* infecting soybean. Journal of General and Molecular Virology 1:7-11. 2009.

THOUVENEL, J.C., MONSARRAT, A. & FAUQUET, C. Isolation of *Cowpea mild mottle virus* from diseased soybean in the Ivory Coast. Plant Disease 66:336-337. 1982.

ZHOU, Z., DELL'ORCO, M., SALDARELLI, P., TURTURO, C., MINAFRA, A. & MARTELLI, G.P. Identification of an RNA-silencing suppressor in the genome of *Grapevine virus A*. Journal of General Virology 87:2387-2395. 2006.

CHAPTER 1

MOLECULAR AND BIOLOGICAL CHARACTERIZATION OF *Cowpea mild mottle virus* ISOLATES INFECTING SOYBEAN IN BRAZIL AND EVIDENCE OF RECOMBINATION

Zanardo, L.G., Silva, F.N., Bicalho, A.A.C., Castillo-Urquiza, G.P., Lima, A.T.M., Almeida, A.M.R., Zerbini, F.M. & Carvalho, C.M. Molecular and biological characterization of *Cowpea mild mottle virus* isolates infecting soybean in Brazil, with evidence of recombination. Plant Pathology, accepted for publication.

Molecular and biological characterization of *Cowpea mild mottle virus* isolates infecting soybean in Brazil and evidence of recombination

L.G. Zanardo[1]; F.N. Silva[1]; A.A.C. Bicalho[1]; G.P.C. Urquiza[1]; A.T.M. Lima[1]; A.M.R. Almeida[2]; F.M. Zerbini[1]; C.M. Carvalho[1*]

[1]Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa,
Viçosa, MG, Brazil, 36570-000

[2]Embrapa Soja, Londrina, PR, Brazil, 86001-970.

* Corresponding author: claudine.carvalho@ufv.br

Short title: Characterization of CpMMV isolates infecting soybean

Key words: CpMMV; *Carlavirus*; phylogeny; recombination.

Abstract

We report the biological and molecular characterization of six isolates of a new *Cowpea mild mottle virus* strain (CPMMV; *Carlavirus, Betaflexiviridae*). Soybean plants with mosaic and stem necrosis were collected in Bahia, Goiás, Mato Grosso and Minas Gerais states, Brazil. Complete genomes of the CPMMV isolates are 8,180-8,198 nucleotides (nt) long, excluding the 3'-polyadenylate tail, and have 67-68% nt sequence identity with a Ghanaian isolate of CPMMV, the only CPMMV isolate for which the genome previously has been sequenced. The replicase has only 60-61% nt sequence identity with the Ghanaian CPMMV isolate, and the coat protein is highly conserved (79% nt sequence identity and 95-96% amino acid sequence identity). The high CP identity and the phylogenetic analyses supported the classification of the Brazilian isolates as CPMMV. Biological and molecular differences with the Ghanaian CPMMV isolate were found and indicated that the six

isolates represent a distinct CPMMV strain denominated as CPMMV-BR. Furthermore, we show that recombination occurred mainly in the polymerase gene, and may occur less frequently in other regions of the CPMMV genome.

Introduction

Cowpea mild mottle virus (CPMMV, family *Betaflexiviridae*, genus *Carlavirus*) is a serious problem in Brazilian soybean (*Glycine max*), causing dwarfing, chlorosis, vein clearing/mosaic, leaf and stem necrosis (Iwaki *et al.*, 1982; Almeida *et al.*, 2003; Almeida, 2008; Tavasoli *et al.*, 2009). The virus was originally reported infecting cowpea (*Vigna unguiculata*) in Ghana (Brunt & Kenten, 1973; Tavasoli *et al.*, 2009) and was subsequently found causing mosaic and leaf crinkling in soybean in Thailand (Iwaki *et al.*, 1982) and the Ivory Coast (Thouvenel *et al.*, 1982). At around the same time, it was reported in Brazil in common bean (*Phaseolus vulgaris*) cv. Jalo (Costa *et al.*, 1983). Much later, in the 2000-2001 growing season, soybean plants showing symptoms of stem necrosis caused by CPMMV were observed in the state of Goiás. Over the next two years, new outbreaks occurred in the states of Mato Grosso, Bahia, Maranhão and Paraná (Almeida *et al.*, 2003), several thousands of kilometers apart.

CPMMV has a single-stranded, positive-sense RNA genome of 8,127 nucleotides polyadenylated at the 3'end, with a typical organization of genus *Carlavirus* in six open reading frames (ORFs) and a short UTR at the 5' and 3' termini (Martelli *et al.*, 2007; Menzel *et al.*, 2010; Adams *et al.*, 2012). ORF1 encodes the putative RNA-dependent RNA polymerase (RdRp). ORFs 2, 3 and 4 encode the triple gene block (TGB1, TGB2, TGB3, respectively), essential for virus movement (Martelli *et al.*, 2007; Adams *et al.*, 2012). ORF5 encodes the coat protein

(CP) while ORF6 encodes a nucleic acid-binding protein (NABP) with a zinc finger motif (Koonin *et al.*, 1991; Lukhovitskaya *et al.*, 2009). The filamentous virus particles can be found in the cytoplasm of palisade, mesophyll, parenchyma and epidermal cells in soybean and *Nicotiana clevelandii* (Brunt *et al.*, 1983). These particles form aggregates in the form of sheets or bundles and often brush-like inclusions (Brunt *et al.*, 1983; Almeida *et al.*, 2003; Almeida *et al.*, 2005). CPMMV is transmitted in a non-persistent manner by the whitefly *Bemisia tabaci* (Munyappa & Reddy, 1983; Jeyanandarajah & Brunt, 1993; Naidu *et al.*, 1998; Almeida *et al.*, 2005; Marubayashi *et al.*, 2010; Brito *et al.*, 2012). Seed transmission appears to be dependent on the viral isolate: for a Ghanaian isolate of CPMMV, seed transmission occurred in soybean, cowpea and with lower frequency in common bean (Brunt & Kenten, 1973); in Venezuela it was demonstrated that CPMMV can be transmitted by yardlong bean seeds (Brito *et al.*, 2012); but Almeida et al. (2005) reported that a Brazilian CPMMV isolate was not transmitted by soybean seeds.

To date, only one CPMMV isolate, from cowpea has been completely sequenced (Menzel *et al.*, 2010). This is probably the same isolate as that previously characterized biologically (Brunt & Kenten, 1973). Only partial sequences of CPMMV from soybean have been reported (Almeida *et al.*, 2005; Tavasoli *et al.*, 2009). Several partial sequences designated as CPMMV are available in GenBank, including the coat protein and/or NABP regions and movement proteins. Full molecular and biological characterization of Brazilian CPMMV isolates is important for a better understanding of this emerging pathogen in soybean crops. The aims of this study were to sequence the full genome of a range of CPMMV isolates collected from soybean, to determine their molecular and biological characteristics, and to compare the genetic variability between these isolates and the previously reported Ghanaian CPMMV isolate from cowpea.

Material and methods

Viral isolates and host range studies

The six CPMMV isolates used in this study were obtained from soybean plants that showed dwarfing and stem necrosis. Three isolates, CPMMV:BR:GO:01:1, CPMMV:BR:BA:02 and CPMMV:BR:MT:02:, were obtained from samples previously collected respectively in Goiatuba (Goiás state – 2000-2001 season, coordinates 18°0’40” S, 49°22’10” W), Barreiras (Bahia state – 2001-2002 season, coordinates 12°8’54” S, 44°59’33” W), and Sorriso, (Mato Grosso state – 2001-2002 season, coordinates 12°33’31” S, 55°42’51” W) (Almeida *et al.*, 2003). The other three isolates, CPMMV:BR:MG:09:2, CPMMV:BR:MG:09:3 and CPMMV:BR:GO:10:5, were obtained from samples collected in Capinópolis (coordinates 18°40’48” S, 49°33’58” W) and Tupaciguara (coordinates 18°36’12” S, 48°41’25” W), Minas Gerais state – 2008-2009 season; and in Cristalina (Goiás state – 2009-2010 season, coordinates 16°46’4” S, 47°36’47” W), respectively. All collected samples were stored in a freezer at -80° C, in order to preserve the original samples. For the study all six isolates were inoculated and maintained in a greenhouse by mechanical inoculation onto soybean cv. CD206 using 0.1 M phosphate buffer, pH 7.2, with 0.1% sodium sulfite.

To rule out the occurrence of mixed infections, the presence of begomoviruses and of *Soybean mosaic virus* (SMV, genus *Potyvirus*) was checked. The presence of begomoviruses was evaluated by total DNA extraction (Dellaporta *et al.*, 1983) followed by PCR amplification using the universal oligonucleotides

PBL1v2040/PCRc1 (Rojas *et al.*, 1993). Infection by SMV was tested by indirect ELISA using a specific polyclonal antiserum produced by one of the authors (AMRA).

For host range tests, the isolates were mechanically inoculated on to plants of 22 species/cultivars belonging to the families Amaranthaceae, Chenopodiaceae, Cucurbitaceae, Fabaceae and Solanaceae (Table 1). Inoculated plants were kept in a greenhouse with average daily temperatures of 26 ± 2^{o}C. Symptoms were recorded until 28 days post-inoculation (dpi). Systemic top leaves were used for indirect ELISA tests (Clark *et al.*, 1986) using a polyclonal antiserum (Carvalho *et al.*, 2013) to confirm CPMMV infection in these plants. Additionally, plants showing only local symptoms on inoculated leaves were used for indirect ELISA.

RT-PCR and cloning

Total RNA was extracted from 100 mg of leaf tissue of soybean cv. CD206 systemically infected with each of the six CPMMV isolates (21 dpi), using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. Reverse transcription (RT) was performed using Superscript III reverse transcriptase (Invitrogen), according to the manufacturer's protocol, using 500 ng of total RNA and the CPMMV primers (40 μM) (Suppl. Table S1). The primers were designed based on the alignment of carlavirus sequences available in GenBank and on the sequence of the Ghanaian isolate of CPMMV (Suppl. Table S2). PCR amplifications were carried out using Platinum *Taq* polymerase (Invitrogen) and forward and reverse primers (Suppl. Table S1). All amplifications consisted of 35 cycles of the following profile: 94ºC for 1 min, annealing at the temperatures listed in Suppl. Table S1, elongation at 72ºC for 1–2 min, depending on the expected size of the fragment to be amplified, and a final extension step at 72ºC for 10 min. Amplified PCR products were gel-purified using the Ilustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare), ligated into the pGEM-T Easy Vector (Promega) and transformed into *Escherichia coli* DH5α cells. Plasmid DNA was purified from recombinant clones using the QIAprep Spin Miniprep Kit (Qiagen) and sequenced in both directions with universal primers (M13F/M13R) at Macrogen (Seoul, South Korea). The sequence of the 5'-end of the viral genome was determined using the Rapid Amplification of complementary DNA (cDNA) Ends (RACE) kit, version 2.0 (Invitrogen) according to the manufacturer's protocol, using primer Race R (Suppl. Table S1).

Sequence analysis and phylogeny

The sequences were assembled using DNA BASER Sequence Assembler v.3.5 (Heracle Biosoft) and the ORFs were located using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The complete genomes were deposited in GenBank [accession no. KC884244 (CPMMV:BR:MG:09:2), KC884245 (CPMMV:BR:MG:09:3), KC884246 (CPMMV:BR:MT:02:1), KC884247 (CPMMV:BR:BA:02), KC884248 (CPMMV:BR:GO:01:1), KC884249 (CPMMV:BR:GO:10:5)]. The nucleotide (nt) sequences of the RdRp (ORF1) and CP (ORF5) were initially submitted to a BLAST search for preliminary species assignment based on the 72% threshold level established for the *Carlavirus* genus (Adams *et al.*, 2012). Additional pairwise nt sequence comparisons were performed by p-distance using MEGA v. 5 (Tamura *et al.*, 2011), and amino acid sequence comparisons were performed using DNAMAN v. 7.0 (Lynnon Biosoft) using the quick alignment method with default parameters. The transmembrane domains of the proteins were estimated using SMART (http://smart.embl.de/) (Letunic *et al.*, 2012). The prediction of putative nuclear localization signals were estimated using cNLS Mapper (http://nls-mapper.iab.keio.ac.jp) (Kosugi *et al.*, 2009).

Nucleotide sequences used in phylogenetic and recombination analyses were aligned using the MUSCLE (Edgar, 2004) module in MEGA v. 5. Phylogenetic analyses were performed using carlavirus sequences from GenBank and the six isolates obtained in this study (Suppl. Table S2). For the phylogenetic tree of the complete genomes, the 5’UTR, 3’UTR and intergenic regions were removed from the alignment and overlapping coding regions were maintained. Phylogenetic trees were constructed by Bayesian inference and Markov Chain Monte Carlo (MCMC) simulation implemented in MrBayes v. 3.0 (Ronquist & Huelsenbeck, 2003), with the evolution models selected by MrModeltest v. 2.2 (Nylander, 2004) using the Akaike Information Criterion (AIC). The MCMC simulation was run for 10 million generations and was sampled every 500 steps, resulting in 20,000 saved trees. Burn-in was set at 2 million generations, leaving 16,000 trees from which the 50% majority rule consensus trees and posterior probabilities were calculated. The trees were viewed using FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/)

Recombination analysis

Detection of potential recombinant sequences, identification of likely parental sequences and localization of possible recombination break points were performed using the Recombination Detection Program RDP v.3.44 (Martin *et al.*, 2010). A multiple comparison-corrected P-value cutoff of 0.05 was used throughout. The analysis included only full CPMMV nucleotide sequences. Only recombination events which were detected by three or more of the seven methods implemented in RDP 3.44 were considered.

Results

Biological properties of CPMMV isolates

After excluding the occurrence of begomoviruses and SMV in the plants in which the CPMMV isolates were being maintained (data not shown), we proceeded with the biological characterization. The Brazilian CPMMV isolates were able to infect a limited number of host plants (Table 1). None of the isolates infected the representatives of the families Amaranthaceae and Cucurbitaceae (Table 1). Of the seven cultivars tested, belonging to three species of family Fabaceae, three were systemically infected by all isolates (Table 1). One of these was soybean cv. CD206, in which a range of symptoms were observed: three isolates caused mosaic and/or crinkled leaves associated with lower disease severity (Figure 1C-E), while the other three isolates caused severe symptoms such as leaf and stem necrosis and bud blight (Figure 1F-K). Soybean cv. Pintado was also infected by all isolates but in this cultivar the infection was consistently symptomless. Meanwhile, common bean cv. Jalo was consistently infected, with mosaic symptoms, by all isolates (Table 1 and Figure 1L). Common bean cv. Manteigão was infected only by isolates CPMMV:BR:GO:01:1 and CPMMV:BR:MT:02:1, and cowpea cv. B7 Gurguéia was infected only by isolates CPMMV:BR:BA:02 and CPMMV:BR:GO:01:1, suggesting that they could be used as differential hosts (Table 1 and Figure 1M).

For the representatives of the Chenopodiaceae and Solanaceae families, only local symptoms in the inoculated leaves were observed (Table 1 and Figure 1A, B, N, O and P). Additionally, a biological characteristics exclusive to the Brazilian

CPMMV isolates was the induction of chlorotic local lesions in *N. benthamiana* (Figure 1N) and *N. debneyi* (Figure 1P), respectively. In the cases of CPMMV:BR:BA:02 and CPMMV:BR:GO:01:1, a further characteristic was the induction of necrotic local lesions and dwarfing in *N. glutinosa* (Table 1 and Figure 1O).

Genomic sequence of CPMMV

The complete genomic sequences of the Brazilian CPMMV isolates were assembled from overlapping clones of RT-PCR and RACE products. They range in size from 8,180 to 8,198 nucleotides (nt), excluding the 3’-terminal poly(A) (Table 2). The nt sequence identity values of the complete genomes among the Brazilian isolates are very high, ranging from 93 to 99%, but are far lower when compared to the Ghanaian isolate (HQ184471), at 67-68%.

The genomes are organized into six putative ORFs, as for other species in the genus *Carlavirus*. The 5’-untranslated region (UTR) is 72-94 nt in length, while the 3’ UTR is 36-89 nt (Table 2). The nucleotide identity values for the 5'-UTR and 3'-UTR regions range from 94-100% and 98-100%, respectively, among the Brazilian isolates, but again are lower when comparing Brazilian isolates with the Ghanaian isolate, at 79-80% and 89-91%, respectively.

ORF1 of all isolates encodes a putative replicase protein, with four conserved motifs: methyltransferase, C23 peptidase (cysteine endopeptidase of single-stranded RNA viruses), RNA helicase and RNA-dependent RNA polymerase. The RdRp is the most conserved ORF among Brazilian isolates (Figure 2). However, identity values between the Brazilian isolates and the Ghanaian isolate are lower than 72% (nt) and 80% (aa) (Figura 2).

ORF2 is separated from ORF1 by an intergenic region of 28 nt, except for isolate CPMMV:BR:GO:01:1. The ORF2 of this isolate shows two possible initiation codons. The first is situated within ORF1 and the second is separated from ORF1 by an intergenic region of 28 nt, with 234 aa residues (Table 2). ORF2 of the other five Brazilian isolates encodes a polypeptide of 231 aa. ORF3 and ORF4 of Brazilian CPMMV isolates showed the same size (Table 2). These three ORFs comprise the triple gene block (TGB). The TGB1 protein (ORF2) contains a P-loop NTPase domain and a RNA helicase domain in all Brazilian CPMMV isolates. For TGB2 protein (ORF3), two transmembrane domains (residues 12-34 and 69-91 or 75-92, in CPMMV:BR:GO:01:1) were predicted in all Brazilian isolates For TGB3 protein (ORF4) a transmembrane domain (residues 4-26) was predicted in all Brazilian CPMMV isolates. High levels of nt and aa identity were observed among the Brazilian CPMMV isolates, and lower levels between Brazilian and Ghanaian isolates, for TGB (Figure 2).

The CP (ORF5) contains a flexivirus domain that is required for genome encapsidation and the motif GLGVPTE at aa positions 120-126 in all Brazilian CPMMV isolates. This motif is conserved in all carlavirus species. ORF5 is separated from ORF4 by a second intergenic region of 14-15 nt for all isolates (Table 2). It is the most conserved region between the Ghanaian CPMMV isolate and all Brazilian isolates (identity values higher than 72% nt and 80% aa) (Figura 2). These results indicate that all six isolates reported in this study belong to the species *Cowpea mild mottle virus*.

ORF6 encodes a nucleic acid binding protein (NABP) with a motif for a putative C-4-type zinc finger and an adjacent putative nuclear localization signal (NLS) defined by YARKRR(A/S)KII with a basic motif (RKRR) for all Brazilian isolates. ORF6 showed the same size for all Brazilian CPMMV isolates except CPMMV:BR:GO:01:1 (Table 2). High identity values were found among Brazilian isolates, and lower values in the comparisons of Brazilian isolates with the Ghanaian isolate (Figura 2).

Sequence comparisons between CPMMV isolates and other carlaviruses showed low levels of identity for all ORFs (Suppl. Figure S1). The *Cucumber vein-clearing virus* (CuVCV) was the carlavirus that showed the highest sequence identity mainly to the ORF2 and ORF5 of CPMMV isolates (Suppl. Figure S1E).

Phylogenetic analyses of complete genomes showed that the Brazilian isolates cluster with the Ghanaian CPMMV isolate with a posterior probability of 1.0 (Figure 3), confirming their close relationship. Analyzing this cluster, it is clear that the Brazilian isolates are highly related to each other and that the isolate CPMMV:BR:GO:01:1 is most distinct amongst them. Additionally, phylogenies were reconstructed for each ORF, and in all cases the Brazilian CPMMV isolates were placed in the same cluster with the Ghanaian isolate (Suppl. Figure S2). The topology of the phylogenetic trees changed slightly depending on the ORF (Suppl. Figure S2), but the close relationship among CPMMV isolates was sustained. The Brazilian isolates formed a monophyletic group for all ORFs. The longer branch length separating CPMMV:BR:GO:01:1 from the other Brazilian isolates for all ORFs, except ORF1, suggests that these regions are more divergent for this isolate (Suppl. Figure S2). Overall, phylogenetic analyses indicated a close relationship between Brazilian and Ghanaian CPMMV isolates, supporting the idea that these isolates belong to the same viral species. The species of carlavirus closest to CPMMV was CuVCV, evidenced by the tree of ORF2, ORF3 and mainly ORF5 (Suppl. Figure S2B, C and E).

Recombination analyses

The analysis of the six Brazilian and the single Ghanaian CPMMV genomes revealed five putative recombination events (Figure 4). All detected recombination events had Brazilian isolates as possible recipients, but the Ghanaian isolate was identified as the putative donor (or a close relative) in one event. Recombination was detected mainly in ORF1 (RdRp). Only one event involved ORFs 2, 3 and 4 (TGB) and ORF5 (CP).

Two recombination events (events 3 and 4) involving a large portion of ORF1 of isolates CPMMV:BR:BA:02 and CPMMV:BR:GO:01:1 were supported by all methods in the RDP package (Figure 4). Recombination event 3 showed CPMMV:BR:MT:02:1 as the putative major parent and an unknown isolate as the minor parent, which led to recombinant isolate CPMMV:BR:BA:02. Recombination event 4 showed an unknown isolate as major parent and CPMMV:BR:MG:09:3 as the putative minor parent, which led to recombinant isolate CPMMV:BR:GO:01:1. The Ghanaian isolate (accession # HQ184471) showed up as a possible minor parent for recombination event 1, which occurred in all isolates except CPMMV:BR:GO:01:1 (Figure 4). It is important to note that the possible parents suggested by the RDP analyses, probably, can not be the actual donors. The parents suggested by recombination analysis may be most closely related to the true parental isolates.

Phylogenetic trees based on the recombinant or non-recombinant portions were constructed using Bayesian inference to determine whether the recombination events were supported. This analysis showed that topological incongruence occurred with all recombinant portions of the recombination events identified by RDP3, while for the non-recombinant portion the same groupings observed for the complete genome were maintained (data not shown). These results support our data and reinforce the occurrence of recombination events.

Discussion

Stem necrosis of soybean caused by CPMMV is a recent problem in Brazilian soybean. Although its causal agent was reported in common bean in the 1980's, the occurrence of CPMMV-associated stem necrosis was observed in soybean only 20 years later. Few studies have been carried out with CPMMV, and only one isolate from cowpea has been completely sequenced. Considering the economical importance of soybean in Brazil and the fact that CPMMV is transmitted by the widespread *B. tabaci*, stem necrosis could become an extremely important disease. This study was performed to gather information on soybean isolates of CPMMV.

Interestingly, different symptoms were observed in association with different CPMMV isolates. The virus can also cause a symptomless infection in some hosts. This ability to cause symptomless infections is a characteristic of carlaviruses (Martelli *et al.*, 2007). For cv. CD206, variable symptoms suggest the existence of necrotic and mild virus isolates. The different symptoms caused are an important aspect of these CPMMV isolates. It is too early to ascribe a genetic basis for the necrotic and mild phenotypes, but we suspect that the differences may be associated with the TGB proteins, for two reasons. First, in the case of *Pepino mosaic virus* (PepMV) (genus *Potexvirus*, family *Alphaflexiviridae*), a virus with a genomic organization similar to carlaviruses, mild isolates became necrotic following a point mutation at amino acid 67 of the TGB3 protein (Hasiów-Jaroszewska *et al.*, 2011). Secondly, the TGB was the region of the genome where we found the lowest identity values among our CPMMV isolates.

Overall, the symptoms induced by Brazilian CPMMV isolates were distinct from those induced by the Ghanaian isolate of CPMMV in plants of the family Fabaceae. The Ghanaian isolate, obtained from cowpea, was able to infect several soybean cultivars causing vein mosaic and leaf chlorosis occasionally followed by apical necrosis (Brunt & Kenten, 1973). The Brazilian isolates caused mild mosaic in cowpea cv. B7 Gurguéia, while infection by the Ghanaian isolate in cowpea was symptomless or characterized by mild chlorotic mottle or conspicuous chlorosis (Brunt & Kenten, 1973). In bean, the Brazilian isolates caused mosaic in cv. Jalo and a symptomless infection in cv. Manteigão, while the Ghanaian isolate caused a characteristic chlorotic spotting or faint chlorotic lesions in different cultivars (Brunt & Kenten, 1973).

The induction of local symptoms in representatives of the families Chenopodiaceae and Solanaceae and the inability to infect *G. globosa* (fam. Amaranthaceae) are generally in line with results from the Ghanaian CPMMV isolate (Brunt & Kenten, 1973). However, some relevant differences were also observed: the Ghanaian isolate is able to infect *Nicotiana clevelandii* and causes a local infection in *G. globosa* (Brunt & Kenten, 1973), and an Israeli CPMMV isolate was able to infect tomato, *Datura stramonium* and *N. glutinosa* (Antignus & Cohen, 1987), yet none of these hosts were infected by the Brazilian CPMMV isolates, except *N. glutinosa* that showed necrotic local lesions. Additionally, some Brazilian CPMMV isolates caused chlorotic local lesions in *N. benthamiana* and in *N. debneyi*. This symptom has not previously been described for infections caused by CPMMV.

Overall, the host range study indicated that the Brazilian CPMMV isolates are distinct from Israeli and Ghanaian CPMMV isolates previously described (Brunt & Kenten, 1973; Antignus & Cohen, 1987).

The genome organizations and functional domains of viral proteins of Brazilian and Ghanaian isolates are similar. However, some functional domains typically found in the carlavirus RdRp (alkB, otu-like peptidase and papain-like protease) were not detected in the RdRp of Brazilian CPMMV isolates characterized (Martelli *et al.*, 2007; Menzel *et al.*, 2010; Adams *et al.*, 2012). Among Brazilian CPMMV isolates the most divergent was CPMMV:BR:GO:01:1. The possible overlap between ORF1 and ORF2 observed for CPMMV:BR:GO:01:1 isolate, generated by the initiation codon at position 5,621-5,623, is not a common characteristic among carlaviruses, only *Potato virus S* (PVS) isolates have shown this overlap (GenBank access #: AJ863509, AJ863510, FJ813513, FJ813512 and HF571059). This, added to the fact that the second possible initiation codon, at position 5,681-5,683, generates a protein of the same size (234 aa) as the Ghanaian CPMMV isolate and of similar size as the other Brazilian CPMMV isolates (231 aa), suggests that the protein does not originate from a sequence overlap.

When comparing sequence identities among Brazilian isolates and with the Ghanaian isolate, we found considerably greater differences than we had expected, particularly for the polymerase (ORF1). According to the International Committee on Taxonomy of Viruses (ICTV), carlavirus isolates are considered to belong to the same species if they have greater than 72% nt identity (or 80% aa identity) in their polymerase or CP genes (Adams *et al.*, 2012). From our comparisons of Brazilian and Ghanaian CPMMV isolates, the identity values for polymerase were considerably below the threshold, at 60-61% nt and 58-61% aa, but identity values for the CP satisfy this criterion. A possible explanation for the low identity found in the polymerase would be recombination. This was observed for carlavirus *Chrysanthemum virus B* (CVB) (Singh *et al.*, 2012). The isolate CVB-S possesses identity values for the polymerase of 66-67% nt and 57-58% aa with other CVB isolates, and it was shown that the region encompassing nucleotides 538 to 4260 of the CVB-S polymerase was recombinant (Singh *et al.*, 2012). For *Lily symptomless virus* (LSV), recombination was detected in almost the entire RdRp or its C-terminal region, in the TGB and in almost the entire CP region (Singh *et al.*, 2008). We therefore examined the possibility of recombination among our CPMMV isolates.

Recombination events were indeed found in the polymerase (as well as one event encompassing TGB1, TGB2, TGB3 and CP). Only event 1 involved the Ghanaian isolate. An interesting fact was observed in the phylogenetic tree of ORF1 (RdRp) in which CPMMV:BR:GO:01:1 isolate showed a shorter branch length inside of the Brazilian isolates monophyletic group. This may reflect recombination event 4 in CPMMV:BR:GO:01:1, whose recombinant region, supposedly donated by CPMMV:BR:MG:09:3, encompasses almost the entire RdRp. For all other trees the CPMMV:BR:GO:01:1 isolate forms a separate branch from the other Brazilian isolates. Therefore, CPMMV:BR:GO:01:1 is likely a recombinant between a divergent, unidentified major parent (from which the region encompassing ORFs 2, 3, 4, 5 and 6 is derived), and CPMMV:BR:MG:09:03 (or a close relative), from which its RdRp is derived. This explains the phylogeny and the lower identity values between CPMMV:BR:GO:01:1 and the other Brazilian isolates for ORFs 2, 3, 4, 5 and 6. However, it does not explain the low identity values between the six Brazilian isolates and the Ghanaian isolate for the RdRp.

Although the RdRp of Brazilian CPMMV isolates showed low identity with the Ghanaian isolate the high CP identities support the classification of Brazilian isolates as CPMMV, the taxonomic criterion was satisfied. Thus, given the biological and molecular differences with the Ghanaian CPMMV isolate and the current species demarcation criteria for the genus *Carlavirus*, the Brazilian CPMMV isolates should be classified as a new CPMMV strain, denominated CPMMV-BR. The clear divergence of the CPMMV:BR:GO:01:1 isolate compared with other Brazilian isolates suggests subdivisions in the Brazilian strain. The taxonomic criterion clearly showed that RdRp or CP identity should be used for species demarcation in the genus *Carlavirus*. The results presented showed that the correct classification of Brazilian isolates as CPMMV species was only possible because the complete genomes were sequenced; if only RdRp had been sequenced, a new carlavirus species would have been proposed. The current taxonomic criterion can be risky in cases where only partial sequences have been determined, especially RdRp. Only one of these two genes should be considered and CP seems to be the most appropriate. Or, maybe, the time has come to use whole genomes sequences for species demarcation in the genus *Carlavirus*.

We have determined the complete sequences of the genomes of the six Brazilian isolates and biological characteristics. The results demonstrated significant differences between isolates infecting soybean and cowpea, and among the Brazilian isolates. The genetic and biological variability of Brazilian CPMMV isolates in addition with the transmission by the whitefly *B. tabaci* makes the causal agent of stem necrosis a potential threat to the soybean crop. This study increases our knowledge on the biological differences and genetic variability of Brazilian CPMMV isolates, and provides important information for the improved viral detection and disease management.

Acknowledgements

This work was funded by FAPEMIG (APQ-0992/09), CNPq (474112/2008-0) and Funarbe grants to CMC. LGZ was supported by a scholarship from CNPq.

References

Adams MJ, Candresse T, Hammond J, Kreuze JF, Martelli GP, Namba S, Pearson MN, Ryu KH, Saldarelli P, Yoshikawa N, 2012. Family *Betaflexiviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, eds. *Virus taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses.* San Diego: Elsevier Academic Press, 920-41.

Almeida AMR, 2008. Viroses da Soja no Brasil: Sintomas, Etiologia, Controle. *Série Documentos* **306**, 1-62.

Almeida AMR, Piuga FF, Kitajima EW*, et al.*, 2003. Necrose da haste da soja. *Série Documentos* **221**, 1-48.

Almeida AMR, Piuga FF, Marin SRR, Kitajima EW, Gaspar JO, Oliveira TGd, Moraes TGd, 2005. Detection and partial characterization of a carlavirus causing stem necrosis of soybean in Brazil. *Fitopatologia Brasileira* **30**, 191-4.

Antignus Y, Cohen S, 1987. Purification and some properties of a new strain of *Cowpea mild mottle virus* in Israel. *Annals of Applied Biology* **110**, 563-9.

Brito M, Fernandez-Rodriguez T, Garrido MJ, Mejias A, Romano M, Marys E, 2012. First report of *Cowpea mild mottle Carlavirus* on yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis*) in Venezuela. *Viruses* **4**, 3804-11.

Brunt AA, Atkey PT, Woods RD, 1983. Intracellular occurrence of *Cowpea mild mottle virus* in two unrelated plant species. *Intervirology* **20**, 137-42.

Brunt AA, Kenten RH, 1973. *Cowpea mild mottle*, a newly recognized virus infecting cowpeas (*Vigna unguiculata*) in Ghana. *Annals of Applied Biology* **74**, 67-74.

Carvalho SL, Silva FN, Zanardo LG, Almeida AMR, Zerbini FM, Carvalho CM, 2013. Production of polyclonal antiserum against *Cowpea mild mottle virus* coat protein and its application in virus detection. *Tropical Plant Pathology* **38**, 49-54.

Clark MF, Lister RM, Bar-Joseph M, Arthur Weissbach HW, 1986. ELISA techniques. In. *Methods in Enzymology*. Academic Press, 742-66.

Costa AS, Gaspar JO, Vega J, 1983. Mosaico angular do feijão jalo causado por um carlavírus transmitido pela mosca branca *Bemisia tabaci*. *Fitopatologia Brasileira* **8**, 325-7.

Dellaporta SL, Woud J, Hicks JB, 1983. A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter* **1**, 19-21.

Edgar R, 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BioMed Central Bioinformatics* **5**, 113.

Hasiów-Jaroszewska B, Borodynko N, Jackowiak P, Figlerowicz M, Pospieszny H, 2011. Single mutation converts mild pathotype of the *Pepino mosaic virus* into necrotic one. *Virus Research* **159**, 57-61.

Iwaki M, Thongmeearkon P, Prommin M, Honda Y, Hibi J, 1982. Whitefly transmission and some properties of *Cowpea mild mottle virus* on soybean in Thailand. *Plant Disease* **66**, 265-8.

Jeyanandarajah P, Brunt AA, 1993. The natural occurrence, transmission, properties and possible affinities of *Cowpea mild mottle virus*. *Journal of Phytopathology* **137**, 148-56.

Koonin EV, Boyko VP, Dolja VV, 1991. Small cysteine-rich proteins of different groups of plant RNA viruses are related to different families of nucleic acid-binding proteins. *Virology* **181**, 395-8.

Kosugi S, Hasebe M, Tomita M, Yanagawa H, 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 10171-6.

Letunic I, Doerks T, Bork P, 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Research* **40**, D302-5.

Lukhovitskaya NI, Ignatovich IV, Savenkov EI, Schiemann J, Morozov SY, Solovyev AG, 2009. Role of the zinc-finger and basic motifs of *Chrysanthemum virus B* p12 protein in nucleic acid binding, protein localization and induction of a hypersensitive response upon expression from a viral vector. *Journal of General Virology* **90**, 723-33.

Martelli GP, Adams MJ, Kreuze JF, Dolja VV, 2007. Family *Flexiviridae*: a case study in virion and genome plasticity. *Annual Review Phytopathology* **45**, 73-100.

Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P, 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462-3.

Marubayashi JM, Yuki VA, Wutke EB, 2010. Transmissão do *Cowpea mild mottle virus* pela mosca branca *Bemisia tabaci* biótipo B para plantas de feijão e soja. *Summa Phytopathologica* **36**, 158-60.

Menzel W, Winter S, Vetten H, 2010. Complete nucleotide sequence of the type isolate of *Cowpea mild mottle virus* from Ghana. *Archives of Virology* **155**, 2069-73.

Munyappa V, Reddy DVR, 1983. Transmission of *Cowpea mild mottle virus* by *Bemisia tabaci* in a nonpersistent manner. *Plant Disease* **67**, 391-3.

Naidu RA, Gowda S, Satyanarayana T, Boyko V, Reddy AS, Dawson WO, Reddy DV, 1998. Evidence that whitefly-transmitted *Cowpea mild mottle virus* belongs to the genus *Carlavirus*. *Archives of Virology* **143**, 769-80.

Nicolaisen M, Nielsen SL, 2001. Analysis of the triple gene block and coat protein sequences of two strains of *Kalanchoe latent carlavirus*. *Virus Genes* **22**, 265-70.

Nylander JAA, 2004. *MrModeltest v2. Program distributed by the author*.

Rojas MR, Gilbertson RL, Russel DR, Maxwell DP, 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease* **77**, 340-7.

Ronquist F, Huelsenbeck JP, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-4.

Singh A, Mahinghara B, Hallan V, Ram R, Zaidi A, 2008. Recombination and phylogeographical analysis of *Lily symptomless virus*. *Virus Genes* **36**, 421-7.

Singh L, Hallan V, Martin D, Ram R, Zaidi A, 2012. Genomic sequence analysis of four new *Chrysanthemum virus B* isolates: evidence of RNA recombination. *Archives of Virology* **157**, 531-7.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731-9.

Tavasoli M, Shahraeen N, Ghorbani S, 2009. Serological and RT-PCR detection of *Cowpea mild mottle Carlavirus* infecting soybean. *Journal of General and Molecular Virology* **1**, 7-11.

Thouvenel JC, Monsarrat A, Fauquet C, 1982. Isolation of *Cowpea mild mottle virus* from diseased soybean in the Ivory Coast. *Plant Disease* **66**, 336-7.

Figure Legends:

Figure 1: Symptoms induced in different hosts infected with the six Brazilian CPMMV isolates. **A.** Chlorotic local lesions in *C. amaranticolor;* **B.** Chlorotic local lesions in *C. quinoa*; **C.** Crinkled leaves, mosaic and vein clearing in soybean CD206 inoculated with CPMMV:BR:BA:02; **D.** Crinkled leaves in soybean CD206 inoculated with CPMMV:BR:GO:01:1; **E.** Mosaic and vein clearing in soybean CD206 inoculated with CPMMV:BR:MT:02:1; **F.** Leaf necrosis in soybean CD206 inoculated with CPMMV:BR:MG:09:2; **G.** Systemic necrosis and dwarfism in soybean CD206 inoculated with CPMMV:BR:MG:09:2; **H.** Leaf necrosis in soybean CD206 inoculated with CPMMV:BR:MG:09:3; **I.** Bud blight and stem necrosis in soybean CD206 inoculated with CPMMV:BR:MG:09:3; **J.** Leaf and stem necrosis in soybean CD206 inoculated with CPMMV:BR:GO:10:5; **K.** Bud blight, dwarfism, leaf and stem necrosis in soybean CD206 inoculated with CPMMV:BR:GO:10:5 **L.** Mosaic in common bean cv. Jalo; **M.** Mosaic in cowpea cv. B7 Gurguéia; **N.** Chlorotic local lesions in *N. benthamiana*; **O.** Necrotic local lesions in *N. glutinosa*; **P.** Chlorotic local lesions in *N. debneyi*. The symptoms described in A, B, L, M, N, O and P were induced by all Brazilian CPMMV isolates.

Figure 2: Two-dimensional plot representing the percent sequence identities between the Brazilian CPMMV isolates and the Ghanaian CPMMV isolate (HQ184471) for all open reading frames (ORFs). Percent nucleotide sequence identities are shown above the diagonal and percent amino acid sequence identities below. ORF1, RNA-dependent RNA polymerase (RdRp); ORFs 2-4, triple gene block (TGB1, TGB2 and TGB3, respectively); ORF 5, coat protein (CP); ORF6, nucleic acid binding protein (NABP).

Figure 3: Phylogenetic relationships, based on complete genome sequences, of Brazilian CPMMV isolates and other carlaviruses using Bayesian inference (implemented in MrBayes v. 3.1, with model GTR+I+G and 10 million generations). *Indian citrus ringspot virus* (ICRSV, genus *Mandarivirus*, family *Alphaflexiviridae*) was used as the outgroup. The 5’ UTR, 3’ UTR and intergenic regions were removed from the alignment and the overlapping coding regions were maintained. Support for the nodes is presented as filled circles (posterior probabilities from 0.95 to 1.0) or open circles (posterior probabilities from 0.85 to 0.94). The six Brazilian isolates are indicated in bold. The accession numbers of the sequences are shown next to their acronym.

Figure 4: A. Schematic representation of the recombination events identified among CPMMV isolates by the RDP3 program. Each box represents a viral isolate with the recombination events identified by numbers. **B.** Details of recombination events in the genomes of CPMMV isolates. Recombination detection methods are represented by letters: R=rdp; G=Genecov; B=Bootscan; M= Maximum χ^2 ; C= Chimaera; S= Sister scan; 3=3Seq. Only the lowest p-value is indicated for the underlined method. The CPMMV genome is shown at the top of the figure. The Ghanaian CPMMV isolate (HQ184471) is termed CPMMV HQ184471.

Supplementary figure S1: Two-dimensional plot representing the percent sequence identities between the Brazilian CPMMV isolates and the most closely related carlaviruses. Nucleotide sequence identities are shown above the diagonal and amino acid sequence identities below. The six Brazilian isolates are indicated in red. **A.** ORF1, RNA-dependent RNA polymerase (RdRp); **B, C** and **D.** ORFs 2-4, triple gene block (TGB1, TGB2 and TGB3, respectively); **E.** ORF5, coat protein (CP); **F.** ORF6, nucleic acid binding protein (NABP)**.**

Supplementary figure S2: Phylogenetic relationships, based on the sequences of individual ORFs, of the Brazilian CPMMV isolates with different carlaviruses, determined using Bayesian inference (implemented in MrBayes 3.1, with selection of models GTR+I+G for all ORFs and 10 million generations). *Indian citrus ringspot virus* (ICRSV, genus *Mandarivirus*, family *Alphaflexiviridae*) was used as the outgroup. Bayesian posterior probability values are given between nodes. The six Brazilian isolates are indicated in red. The accession numbers of the sequences are showed next to their acronym. **A.** ORF1, RNA-dependent RNA polymerase (RdRp); **B, C** and **D.** ORFs 2-4, triple gene block (TGB1, TGB2 and TGB3, respectively); **E.** ORF5, coat protein (CP); **F.** ORF6, nucleic acid binding protein (NABP).

Table 1: Symptoms induced in different host plants by the six CPMMV isolates described in this study.

Family	**Species**	**Symptoms* (Plants infected/inoculated)** $^{\text{no of experiments}}$					
		CPMMV:BR:BA:02	CPMMV:BR:GO:01:1	CPMMV:BR:MT:02:1	CPMMV:BR:MG:09:2	CPMMV:BR:MG:09:3	CPMMV:BR:GO:10:5
Amaranthaceae	*Gomphrena globosa*	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
Chenopodiaceae	*Chenopodium amaranticolor*	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}
	Chenopodium quinoa	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}
Cucurbitaceae	*Cucurbita pepo*	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}
Fabaceae	*Glycine max* cv CD 206	**Cl, M, Vc** (9/9)^{3}	**Cl** (9/9)^{3}	**M, Vc** (9/9)^{3}	**Bb, D, Ln, Sn** (9/9)^{3}	**Bb, D, Ln, Sn** (9/9)^{3}	**Bb, D, Ln, Sn** (9/9)^{3}
	G. max cv Pintado	**+** (6/6)^{2}	**+** (6/6)^{2}	**+** (6/6)^{2}	**+** (6/6)^{2}	**+** (6/6)^{2}	**+** (6/6)^{2}
	Phaseolus vulgaris cv Jalo	**M** (7/9)^{3}	**M** (9/9)^{3}	**M** (7/9)^{3}	**M** (7/9)^{3}	**M** (7/9)^{3}	**M** (9/9)^{3}
	P. vulgaris cv Manteigão	- (0/3)^{1}	**+** (2/3)^{1}	**+** (2/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	P. vulgaris cv Ouro Negro	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/6)^{2}
	Vigna unguiculata cv B7 Gurguéia	**M** (2/6)^{2}	**M** (2/6)^{2}	- (0/6)^{2}	- (0/6)^{2}	- (0/3)^{2}	- (0/3)^{2}
	V. unguiculata cv Pitiúba	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
Solanaceae	*Capsicum annuum*	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	Datura stramonium	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	Nicotiana benthamiana	**Cll** (4/6)^{2}	**Cll** (4/6)^{2}	**Cll** (4/6)^{2}	**Cll** (4/6)^{2}	**Cll** (6/6)^{2}	**Cll** (4/6)^{2}
	Nicotiana clevelandii	- (0/9)^{3}	- (0/9)^{3}	- (0/9)^{3}	- (0/9)^{3}	- (0/9)^{3}	- (0/9)^{3}
	Nicotiana glutinosa	**D, Nll** (2/3)^{1}	**D, Nll** (2/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	Nicotiana Debneyi	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}	**Cll** (6/6)^{2}
	Nicotiana tabacum cv Havana	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	N. tabacum cv TNN	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	N. tabacum cv White Burley	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	N. tabacum cv Xanthi	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}
	Solannum lycopersicum cv Rutgers	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}	- (0/3)^{1}

*Bb= Bud Blight Cl=Crinkled leaves, Cll=Chlorotic local lesions, D=Dwarfism, Ln=Leaf necrosis, M=Mosaic, Nll=Necrotic local lesions, Sn=stem necrosis; Vc=Vein clearing, + Symptomless infection; - no symptoms and negative by ELISA.

Table 2: Genomic organization of the six CPMMV isolates described in this study.

Isolate	Genomic organization*														Genome size (nt)
	5' UTR	ORF1		ORF2		ORF3		ORF4		ORF5		ORF6		3' UTR	
	# nt	nt position	# aa	nt position	# aa	nt position	# aa	nt position	# aa	nt position	# aa	nt position	# aa	# nt	
CPMMV:BR:BA:02	72	73-5,649	1,858	5,678-6,373	231	6,373-6,693	106	6,671-6,877	68	6,893-7,759	288	7,762-8,103	113	88	8,191
CPMMV:BR:GO:01:1	72	73-5,652	1,859	5,681-6,385	234	6,375-6,695	106	6,674-6,880	68	6,895-7,761	288	7,764-8,162	132	36	8,198
CPMMV:BR:MT:02:1	72	73-5,637	1,854	5,666-6,361	231	6,361-6,681	106	6,659-6,865	68	6,881-7,747	288	7,750-8,091	113	89	8,180
CPMMV:BR:MG:09:2	73	74-5,653	1,859	5,682-6,377	231	6,377-6,697	106	6,675-6,881	68	6,897-7,763	288	7,766-8,107	113	89	8,196
CPMMV:BR:MG:09:3	94	95-5,653	1,852	5,682-6,377	231	6,377-6,697	106	6,675-6,881	68	6,897-7,763	288	7,766-8,107	113	89	8,196
CPMMV:BR:GO:10:5	72	73-5,652	1,859	5,681-6,376	231	6,376-6,696	106	6,674-6,880	68	6,896-7,762	288	7,765-8,106	113	88	8,194
CPMMV HQ184471**	73	74-5,584	1,836	5,613-6,217	234	6,308-6,685	125	6,606-6,815	69	6,833-7,699	288	7,702-8,010	102	116	8,127

*ORF1: RNA-dependent RNA polymerase (RdRp); ORFs 2, 3 and 4: triple gene block (TGB1, TGB2 and TGB3, respectively); ORF5: coat protein (CP); ORF6: nucleic acid binding protein (NABP). ** Ghanaian CPMMV isolate (GenBank accession number: HQ184471).

Figure 1:

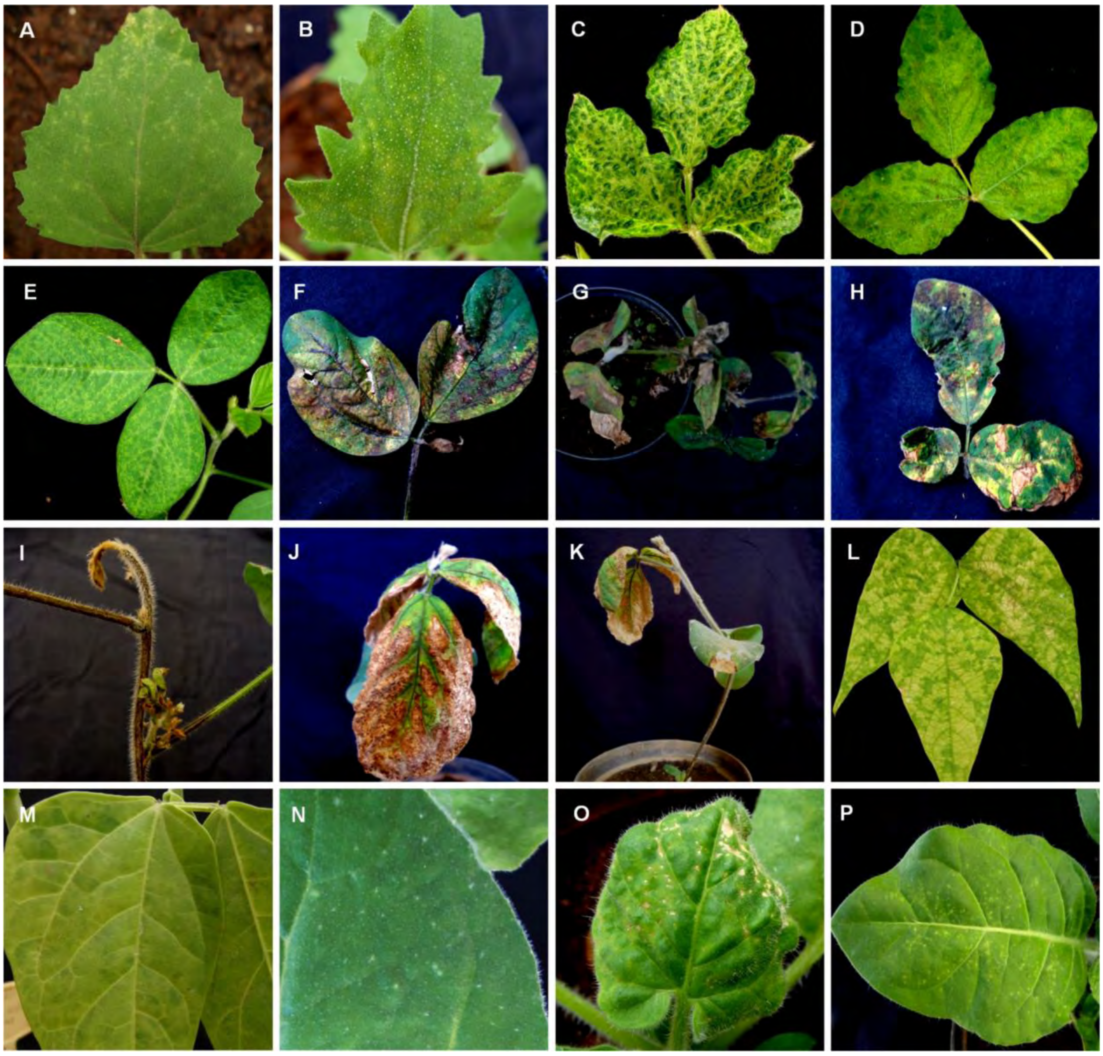
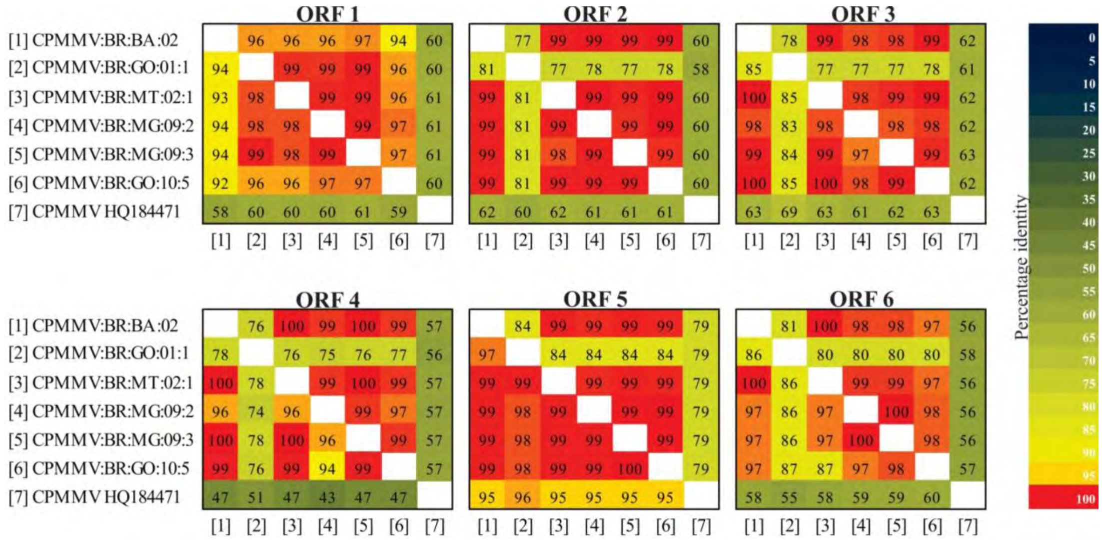


Figure 2



ORF 1

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		96	96	96	97	94	60
[2] CPMMV:BR:GO:01:1	94		99	99	99	96	60
[3] CPMMV:BR:MT:02:1	93	98		99	99	96	61
[4] CPMMV:BR:MG:09:2	94	98	98		99	97	61
[5] CPMMV:BR:MG:09:3	94	99	98	99		97	61
[6] CPMMV:BR:GO:10:5	92	96	96	97	97		60
[7] CPMMV HQ184471	58	60	60	60	61	59	

ORF 2

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		77	99	99	99	99	60
[2] CPMMV:BR:GO:01:1	81		77	78	77	78	58
[3] CPMMV:BR:MT:02:1	99	81		99	99	99	60
[4] CPMMV:BR:MG:09:2	99	81	99		99	99	60
[5] CPMMV:BR:MG:09:3	99	81	98	99		99	60
[6] CPMMV:BR:GO:10:5	99	81	99	99	99		60
[7] CPMMV HQ184471	62	60	62	61	61	61	

ORF 3

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		78	99	98	98	99	62
[2] CPMMV:BR:GO:01:1	85		77	77	77	78	61
[3] CPMMV:BR:MT:02:1	100	85		98	99	99	62
[4] CPMMV:BR:MG:09:2	98	83	98		98	98	62
[5] CPMMV:BR:MG:09:3	99	84	99	97		99	63
[6] CPMMV:BR:GO:10:5	100	85	100	98	99		62
[7] CPMMV HQ184471	63	69	63	61	62	63	

ORF 4

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		76	100	99	100	99	57
[2] CPMMV:BR:GO:01:1	78		76	75	76	77	56
[3] CPMMV:BR:MT:02:1	100	78		99	100	99	57
[4] CPMMV:BR:MG:09:2	96	74	96		99	97	57
[5] CPMMV:BR:MG:09:3	100	78	100	96		99	57
[6] CPMMV:BR:GO:10:5	99	76	99	94	99		57
[7] CPMMV HQ184471	47	51	47	43	47	47	

ORF 5

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		84	99	99	99	99	79
[2] CPMMV:BR:GO:01:1	97		84	84	84	84	79
[3] CPMMV:BR:MT:02:1	99	99		99	99	99	79
[4] CPMMV:BR:MG:09:2	99	98	99		99	99	79
[5] CPMMV:BR:MG:09:3	99	98	99	99		99	79
[6] CPMMV:BR:GO:10:5	99	98	99	99	100		79
[7] CPMMV HQ184471	95	96	95	95	95	95	

ORF 6

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] CPMMV:BR:BA:02		81	100	98	98	97	56
[2] CPMMV:BR:GO:01:1	86		80	80	80	80	58
[3] CPMMV:BR:MT:02:1	100	86		99	99	97	56
[4] CPMMV:BR:MG:09:2	97	86	97		100	98	56
[5] CPMMV:BR:MG:09:3	97	86	97	100		98	56
[6] CPMMV:BR:GO:10:5	97	87	87	97	98		57
[7] CPMMV HQ184471	58	55	58	59	59	60	

Figure 3:

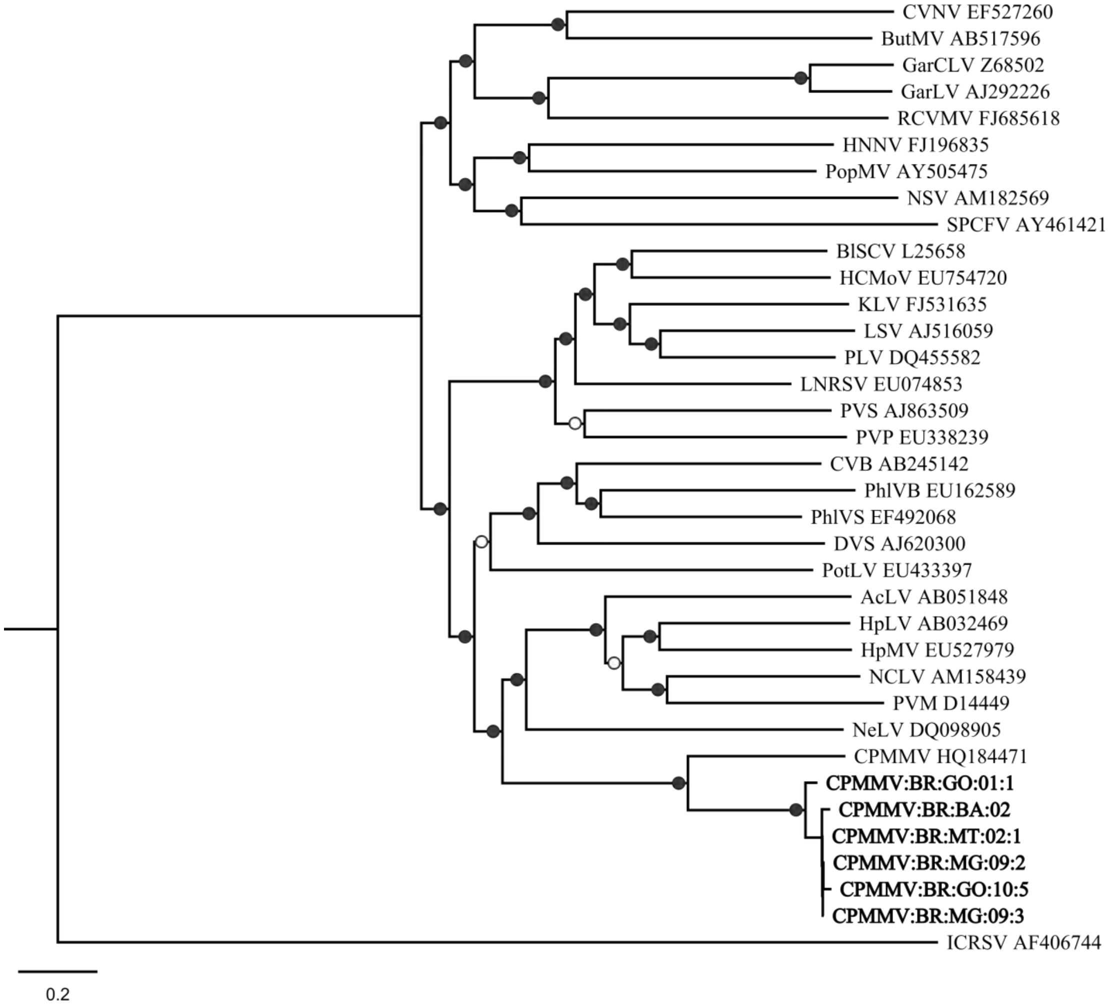

CVNV EF527260
ButMV AB517596
GarCLV Z68502
GarLV AJ292226
RCVMV FJ685618
HNNV FJ196835
PopMV AY505475
NSV AM182569
SPCFV AY461421
BlSCV L25658
HCMoV EU754720
KLV FJ531635
LSV AJ516059
PLV DQ455582
LNRSV EU074853
PVS AJ863509
PVP EU338239
CVB AB245142
PhlVB EU162589
PhlVS EF492068
DVS AJ620300
PotLV EU433397
AcLV AB051848
HpLV AB032469
HpMV EU527979
NCLV AM158439
PVM D14449
NeLV DQ098905
CPMMV HQ184471
CPMMV:BR:GO:01:1
CPMMV:BR:BA:02
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:2
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:3
ICRSV AF406744
0.2

Figure 4:

B

Recombinant	Event	Breakpoint	Major Parent	Minor Parent	Site of recombination	Event detected by methods shown	P-value
CPMMV:BR:BA:02	1	91 - 145	CPMMV:BR:GO:01:1	CPMMV HQ184471	ORF1	$\underline{R}$, G, M, C	2.1×10^{-08}
CPMMV:BR:BA:02	2	2965 - 3289	CPMMV:BR:MG:09:2	Unknown	ORF1	R, G, M, $\underline{C}$	2.0×10^{-09}
CPMMV:BR:BA:02	3	285 - 5654	CPMMV:BR:MT:02:1	Unknown	ORF1	R, G, B, M, C, S, $\underline{3}$	2.1×10^{-83}
CPMMV:BR:GO:01:1	4	262 - 5638	Unknown	CPMMV:BR:MG:09:3	ORF1	R, G, B, M, C, S, $\underline{3}$	1.1×10^{-82}
CPMMV:BR:MT:02:1	1	91 - 145	CPMMV:BR:GO:01:1	CPMMV HQ184471	ORF1	$\underline{R}$, G, M, C	2.1×10^{-08}
CPMMV:BR:MG:09:2	1	92 - 146	CPMMV:BR:GO:01:1	CPMMV HQ184471	ORF1	$\underline{R}$, G, M, C	2.1×10^{-08}
CPMMV:BR:MG:09:3	1	89 - 143	CPMMV:BR:GO:01:1	CPMMV HQ184471	ORF1	$\underline{R}$, G, M, C	2.1×10^{-08}
CPMMV:BR:GO:10:5	1	91 - 145	CPMMV:BR:GO:01:1	CPMMV HQ184471	ORF1	$\underline{R}$, G, M, C	2.1×10^{-08}
CPMMV:BR:GO:10:5	5	5961 - 7753	Unknown	CPMMV:BR:MG:09:2	ORF2, 3, 4, 5	B, M, $\underline{C}$	1.8×10^{-06}

Supplementary Table S1: Primers used in RT-PCR and predicted amplicon size for viral detection and cloning.

Fragments	Primer	Primer location**	Primer sequence 5'-3'	Annealing Temperature °C	PCR product size
5' end	Race R	518-539	CCAATGTTGCCCTGTGCCTAC	55	538
ORF1	ORF1 F1	143- 160	TTGCTTCCAAAGCTGCCT	57	1008
	ORF R1	1133-1150	TCTCGTTAGCTGAGGGTT[cD]		
ORF1	ORF1 F2	929-948	AGGTGCTGCCGTCACTTGT	50	1223
	ORF R2	2133-2152	CTGACTTAAGCTCATCTGG		
ORF1	ORF1 F3	1994-2012	TCAGATAAATGAGGGTGG	48	1618
	ORF1 R3	3593-3612	TCCAAGCAAGTCCCTATCT		
ORF1	ORF1 F4	3321-3341	AGGAAAGCCCTACTTGAGGA	50	1209
	ORF1 R4	4510-4530	CTTTACCGCCATAATGAACG		
ORF1	ORF1 F5	4386-4405	GGTTCGATTGTCTCAGATC	50	1347
	ORF1 R5	5713-5733	CTGCCCAGTCGAATGTAATT		
ORF2	ORF2 F	5653-5670	TCCTTTAGGTAGTGAGGC	45	938
	ORF2 R	6572-6590	AAGTTCGTGCCAGTTGAC[cD]		
ORF3	ORF3 F	6301-6321	CTTNATYTGCYTNACNAGGCA	45	552
	ORF3 R	6852-6837	TGTTCTCTNACCAAGT[cD]		
ORF4 - 3' end	ORF 4 F	6523-6540	TAYMRDGAYGGNACHAA*	45	1676
	ORF6 R	8183-8198	TAAAACCAGGAAATAAC[cD]		

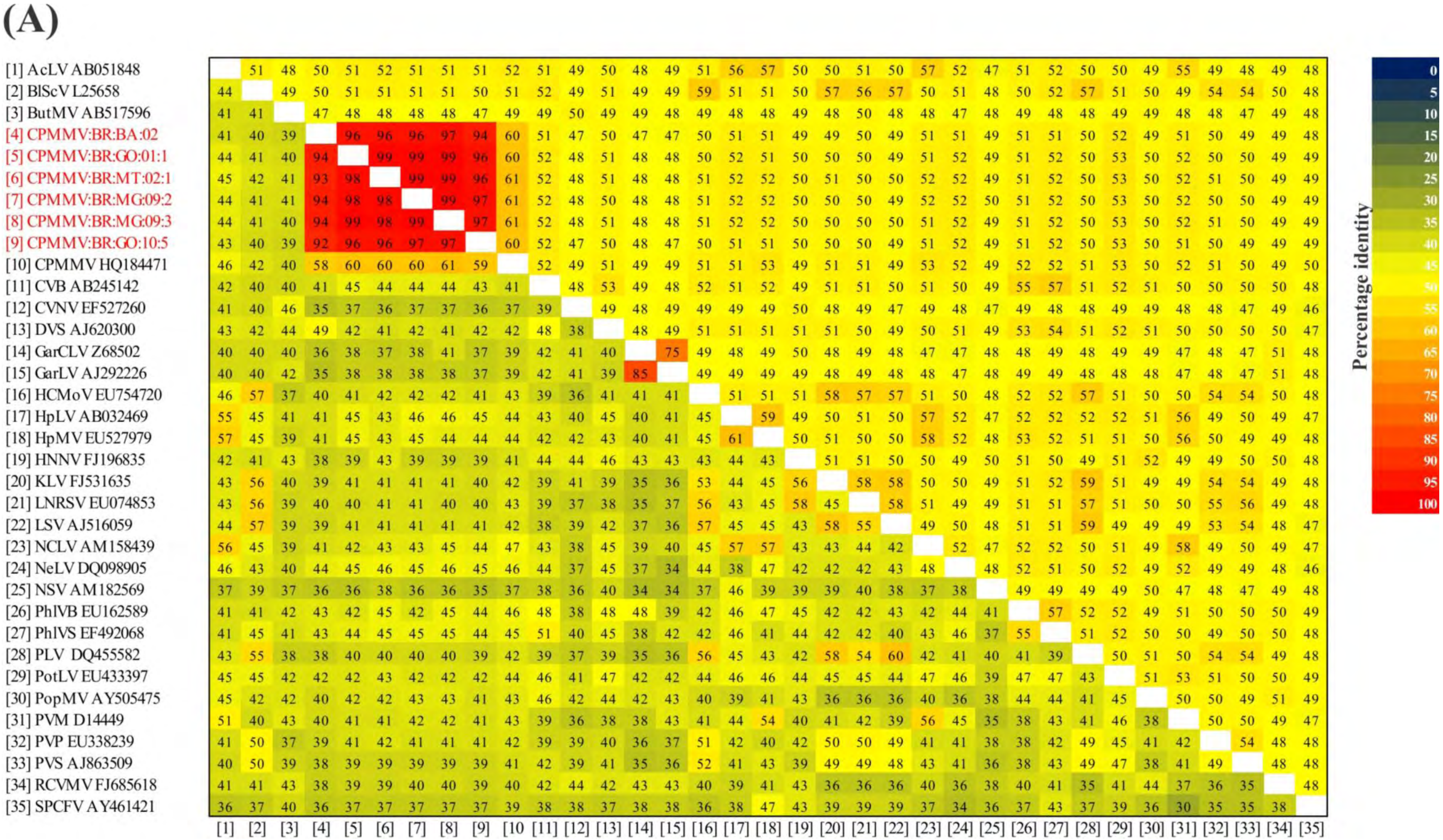
*Described by Nicolaisen and Nielsen (2001). [cD] Used for cDNA synthesis. **Based on the CPMMV:BR:GO:10:5 isolate.

Supplementary Table S2: Carlavirus sequences used for sequences comparisons, phylogenetic and recombination analyses.

Species	Acronym	GenBank accession #
Aconitum latent virus	AcLV	AB051848
Blueberry scorch virus	BlScV	L25658
Butterbur mosaic virus	ButMV	AB517596
Chrysanthemum virus B	CVB	AB245142
Cowpea mild mottle virus	CPMMV	HQ184471 KC884244 KC884245 KC884246 KC884247 KC884248 KC884249
Coleus vein necrosis virus	CVNV	EF527260
*Cucumber vein-clearing virus***	CuVCV	JN591720
Daphne virus S	DVS	AJ620300
Garlic common latent virus	GarCLV	Z68502
Garlic latent virus	GarLV	AJ292226
Helleborus net necrosis virus	HNNV	FJ196835
Hydrangea chlorotic mottle virus	HCMoV	EU754720
Hop latent virus	HpLV	AB032469
Hop mosaic virus	HpMV	EU527979
Kalanchoe latent virus	KLV	FJ531635
Ligustrum necrotic ringspot virus	LNRSV	EU074853
Lily symptomless virus	LSV	AJ516059
Narcissus common latent virus	NCLV	AM158439
Narcissus symptomless virus	NSV	AM182569
Nerine latent virus	NeLV	DQ098905
Passiflora latent carlavirus	PLV	DQ455582
Phlox virus B	PhlVB	EU162589
Phlox virus S	PhlVS	EF492068
Poplar mosaic virus	PopMV	AY505475
Potato latent virus	PotLV	EU433397
Potato virus M	PVM	D14449
Potato virus P	PVP	EU338239
Potato virus S	PVS	AJ863509
Red clover vein mosaic virus	RCVMV	FJ685618
Sweet potato chlorotic fleck virus	SPCFV	AY461421
*Indian citrus ringspot virus**	ICRSV	AF406744

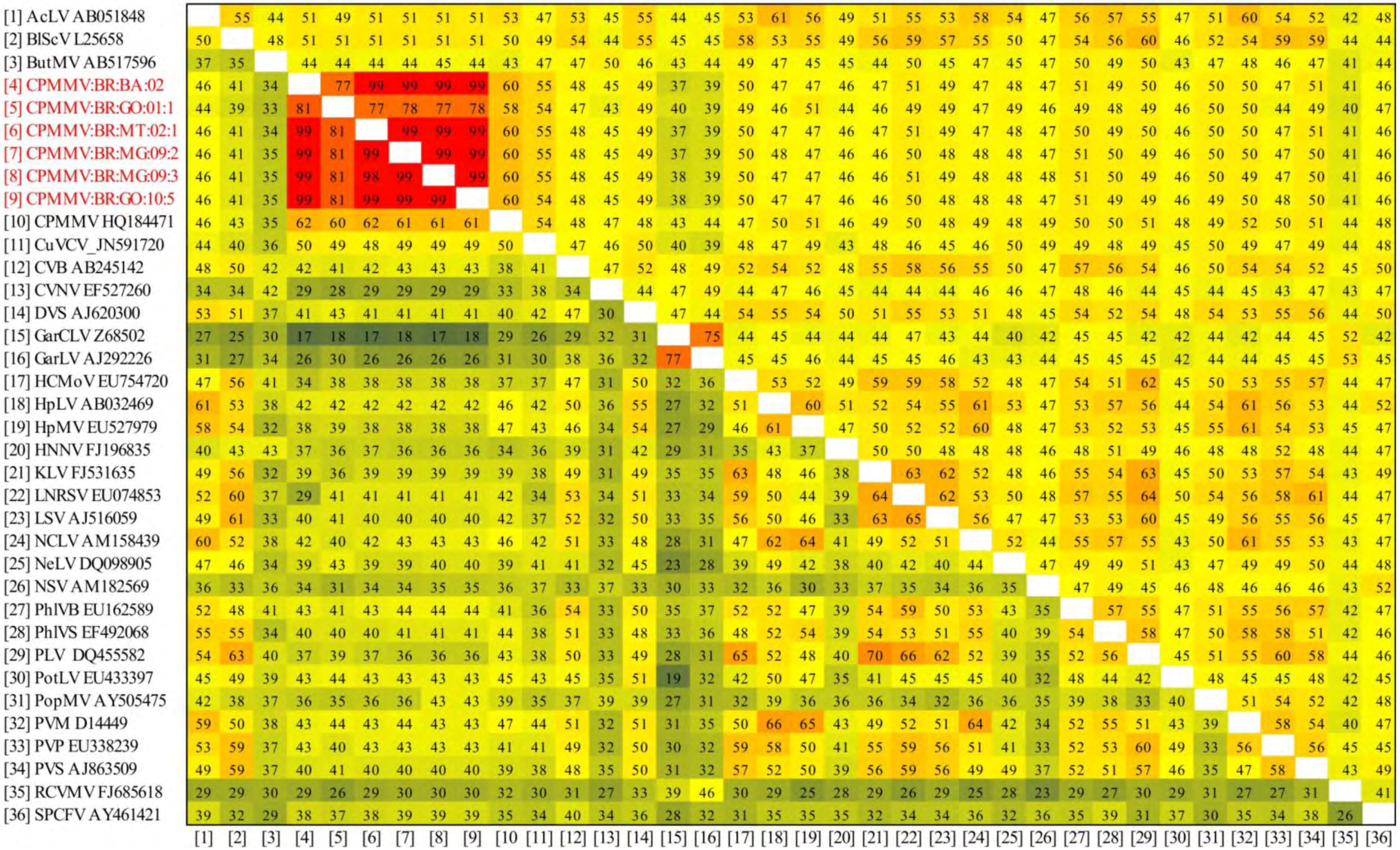
*****Used as outgroup in phylogenetic analysis. **Only carlavirus with partial genome used in the ORF2-6 analyses. The genbank accession # underlined correspond to sequenced isolates in the present study: CPMMV:BR:MG:09:2 (KC884244), CPMMV:BR:MG:09:3 (KC884245), CPMMV:BR:MT:02:1 (KC884246), CPMMV:BR:BA:02 (KC884247), CPMMV:BR:GO:01:1 (KC884248) and CPMMV:BR:GO:10:5 (KC884249). The recombination analysis included only full CPMMV nucleotide sequences.

Figure Supplementary S1:



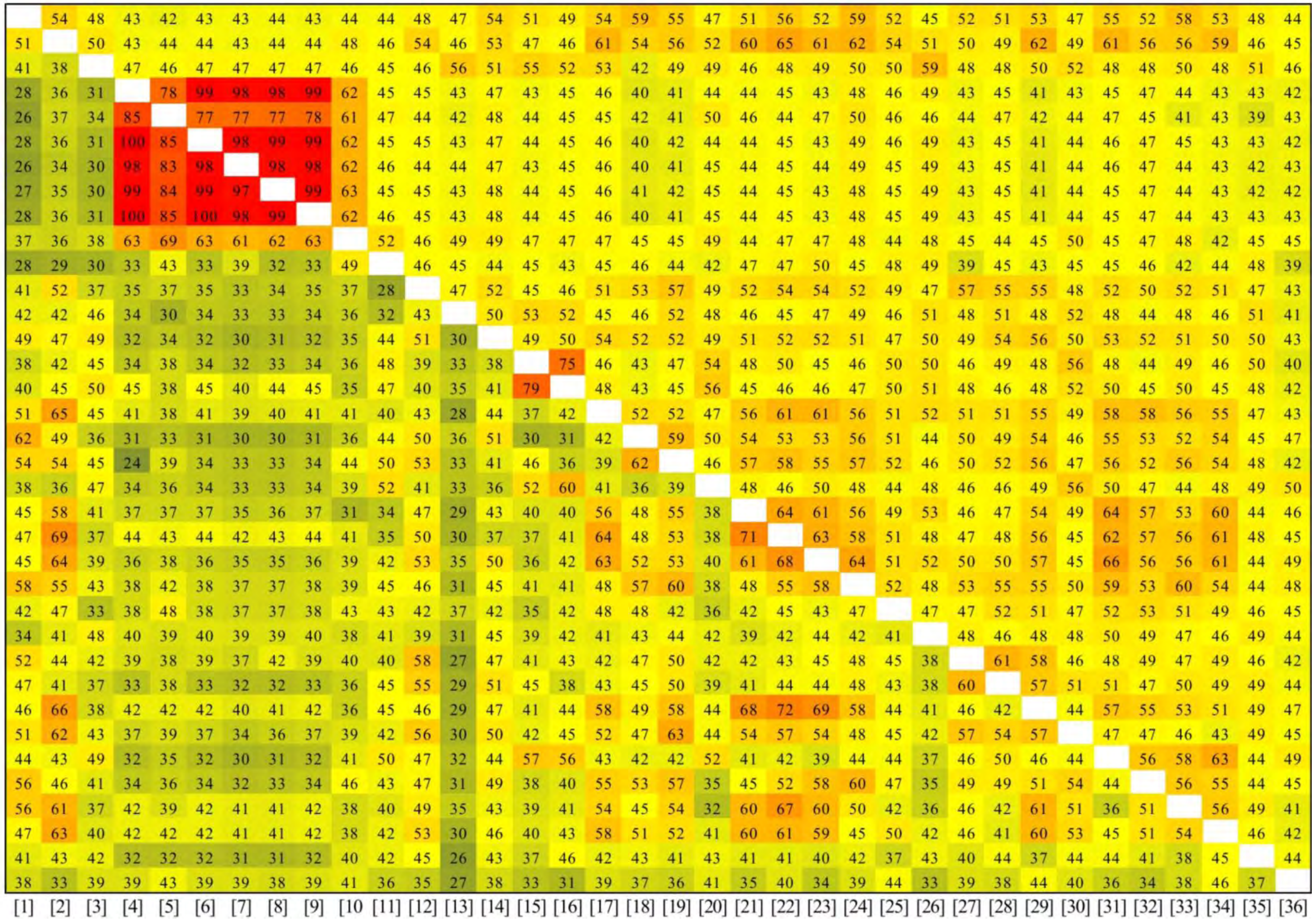
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]
[1] AcLV AB051848		51	48	50	51	52	51	51	51	52	51	49	50	48	49	51	56	57	50	50	51	50	57	52	47	51	52	50	50	49	55	49	48	49	48
[2] BlScV L25658	44		49	50	51	51	51	51	50	51	52	49	51	49	49	59	51	51	50	57	56	57	50	51	48	50	52	57	51	50	49	54	54	50	48
[3] ButMV AB517596	41	41		47	48	48	48	48	47	49	49	50	49	49	48	49	48	48	49	48	50	48	48	47	48	49	48	48	49	49	48	48	47	49	48
[4] CPMMV:BR:BA:02	41	40	39		96	96	96	97	94	60	51	47	50	47	47	50	51	51	49	49	50	49	51	51	49	51	51	50	52	49	51	50	49	49	48
[5] CPMMV:BR:GO:01:1	44	41	40	94		99	99	99	96	60	52	48	51	48	48	50	52	51	50	50	50	49	51	52	49	51	52	50	53	50	52	50	50	49	49
[6] CPMMV:BR:MT:02:1	45	42	41	93	98		99	99	96	61	52	48	51	48	48	51	52	52	50	51	50	50	52	52	49	51	52	50	53	50	52	51	50	49	49
[7] CPMMV:BR:MG:09:2	44	41	41	94	98	98		99	97	61	52	48	50	48	48	51	52	52	50	50	50	49	52	52	50	51	52	50	53	50	52	50	50	49	49
[8] CPMMV:BR:MG:09:3	44	41	40	94	99	98	99		97	61	52	48	51	48	48	51	52	52	50	50	50	50	51	52	49	51	52	50	53	50	52	51	50	49	49
[9] CPMMV:BR:GO:10:5	43	40	39	92	96	96	97	97		60	52	47	50	48	47	50	51	51	50	50	50	49	51	52	49	51	52	50	53	50	51	50	49	49	49
[10] CPMMV HQ184471	46	42	40	58	60	60	60	61	59		52	49	51	49	49	51	51	53	49	51	51	49	53	52	49	52	52	51	53	50	52	51	50	49	50
[11] CVB AB245142	42	40	40	41	45	44	44	44	43	41		48	53	49	48	52	51	52	49	51	51	50	51	50	49	55	57	51	52	51	50	50	50	50	48
[12] CVNV EF527260	41	40	46	35	37	36	37	37	36	37	39		49	48	49	49	49	49	50	48	49	47	49	48	47	49	48	48	49	49	48	48	47	49	46
[13] DVS AJ620300	43	42	44	49	42	41	42	41	42	42	48	38		48	49	51	51	51	51	51	50	49	50	51	49	53	54	51	52	51	50	50	50	50	47
[14] GarCLV Z68502	40	40	40	36	38	37	38	41	37	39	42	41	40		75	49	48	49	50	48	49	48	47	47	48	48	49	48	49	49	47	48	47	51	48
[15] GarLV AJ292226	40	40	42	35	38	38	38	38	37	39	42	41	39	85		49	49	49	49	48	49	48	48	47	48	49	49	48	48	48	47	48	47	51	48
[16] HCMoV EU754720	46	57	37	40	41	42	42	42	41	43	39	36	41	41	41		51	51	51	58	57	57	51	50	48	52	52	57	51	50	50	54	54	50	48
[17] HpLV AB032469	55	45	41	41	45	43	46	46	45	44	43	40	45	40	41	45		59	49	50	51	50	57	52	47	52	52	52	52	51	56	49	50	49	47
[18] HpMV EU527979	57	45	39	41	45	43	45	44	44	44	42	42	43	40	41	45	61		50	51	50	50	58	52	48	53	52	51	51	50	56	50	49	49	48
[19] HNNV FJ196835	42	41	43	38	39	43	39	39	39	41	44	44	46	43	43	43	44	43		51	51	50	50	49	50	51	50	49	51	52	49	49	50	50	48
[20] KLV FJ531635	43	56	40	39	41	41	41	41	40	42	39	41	39	35	36	53	44	45	56		58	58	50	50	49	51	52	59	51	49	49	54	54	49	48
[21] LNRSV EU074853	43	56	39	40	40	41	41	40	40	43	39	37	38	35	37	56	43	45	58	45		58	51	49	49	51	51	57	51	50	50	55	56	49	48
[22] LSV AJ516059	44	57	39	39	41	41	41	41	41	42	38	39	42	37	36	57	45	45	43	58	55		49	50	48	51	51	59	49	49	49	53	54	48	47
[23] NCLV AM158439	56	45	39	41	42	43	43	45	44	47	43	38	45	39	40	45	57	57	43	43	44	42		52	47	52	52	50	51	49	58	49	50	49	47
[24] NeLV DQ098905	46	43	40	44	45	46	45	46	45	46	44	37	45	37	34	44	38	47	42	42	42	43	48		48	52	51	50	52	49	52	49	49	48	46
[25] NSV AM182569	37	39	37	36	36	38	36	36	35	37	38	36	40	34	34	37	46	39	39	39	40	38	37	38		49	49	49	49	50	47	48	47	49	48
[26] PhlVB EU162589	41	41	42	43	42	45	42	45	44	46	48	38	48	48	39	42	46	47	45	42	42	43	42	44	41		57	52	52	49	51	50	50	50	49
[27] PhlVS EF492068	41	45	41	43	44	45	45	45	44	45	51	40	45	38	42	42	46	41	44	42	42	40	43	46	37	55		51	52	50	50	49	50	50	48
[28] PLV DQ455582	43	55	38	38	40	40	40	40	39	42	39	37	39	35	36	56	45	43	42	58	54	60	42	41	40	41	39		50	51	50	54	54	49	48
[29] PotLV EU433397	45	45	42	42	42	43	42	42	42	44	46	41	47	42	42	44	46	46	44	45	45	44	47	46	39	47	47	43		51	53	51	50	50	49
[30] PopMV AY505475	45	42	42	40	42	42	43	43	41	43	46	42	44	42	43	40	39	41	43	36	36	36	40	36	38	44	44	41	45		50	50	49	51	49
[31] PVM D14449	51	40	43	40	41	41	42	42	41	43	39	36	38	38	43	41	44	54	40	41	42	39	56	45	35	38	43	41	46	38		50	50	49	47
[32] PVP EU338239	41	50	37	39	41	42	41	41	41	42	39	39	40	36	37	51	42	40	42	50	50	49	41	41	38	38	42	49	45	41	42		54	48	48
[33] PVS AJ863509	40	50	39	38	39	39	39	39	39	41	42	39	41	35	36	52	41	43	39	49	49	48	43	41	36	38	43	49	47	38	41	49		48	48
[34] RCVMV FJ685618	41	41	43	38	39	39	40	40	39	40	42	44	42	43	43	40	39	41	43	36	36	36	40	36	38	40	41	35	41	44	37	36	35		48
[35] SPCFV AY461421	36	37	40	36	37	37	37	37	37	39	38	38	37	38	38	36	38	47	43	39	39	39	37	34	36	37	43	37	39	36	30	35	35	38	

(B)



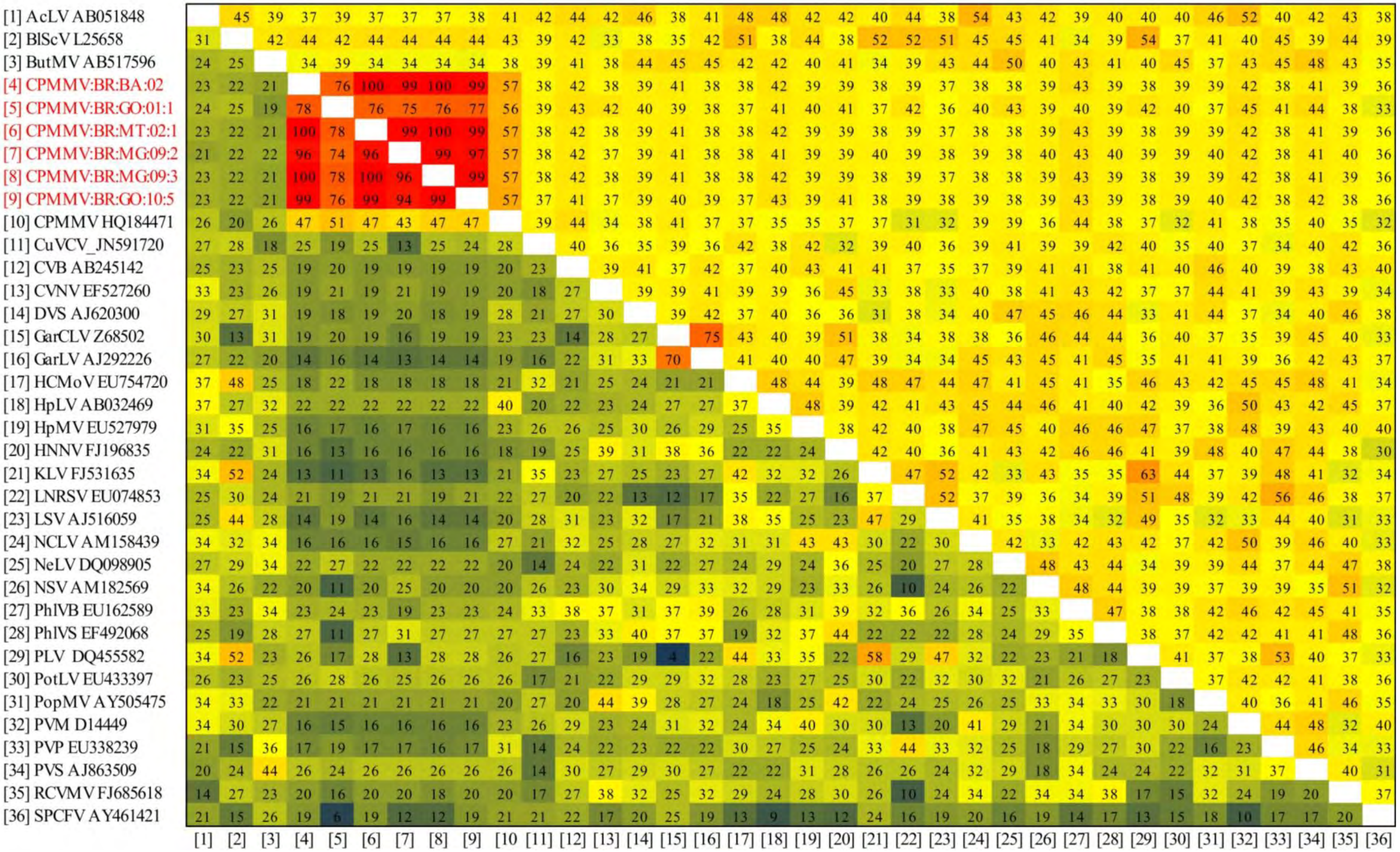
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]
[1] AcLV AB051848		55	44	51	49	51	51	51	51	53	47	53	45	55	44	45	53	61	56	49	51	55	53	58	54	47	56	57	55	47	51	60	54	52	42	48
[2] BlScV L25658	50		48	51	51	51	51	51	51	50	49	54	44	55	45	45	58	53	55	49	56	59	57	55	50	47	54	56	60	46	52	54	59	59	44	44
[3] ButMV AB517596	37	35		44	44	44	44	45	44	43	47	47	50	46	43	44	49	47	45	45	44	50	45	47	45	47	50	49	50	43	47	48	46	47	41	44
[4] CPMMV:BR:BA:02	46	41	34		77	99	99	99	99	60	55	48	45	49	37	39	50	47	47	46	47	51	49	47	48	47	51	49	50	46	50	50	47	51	41	46
[5] CPMMV:BR:GO:01:1	44	39	33	81		77	78	77	78	58	54	47	43	49	40	39	49	46	51	44	46	49	49	47	49	46	49	48	49	46	50	50	44	49	40	47
[6] CPMMV:BR:MT:02:1	46	41	34	99	81		99	99	99	60	55	48	45	49	37	39	50	47	47	46	47	51	49	47	48	47	50	49	50	46	50	50	47	51	41	46
[7] CPMMV:BR:MG:09:2	46	41	35	99	81	99		99	99	60	55	48	45	49	37	39	50	48	47	46	46	50	48	48	48	47	51	50	49	46	50	50	47	50	41	46
[8] CPMMV:BR:MG:09:3	46	41	35	99	81	98	99		99	60	55	48	45	49	38	39	50	47	47	46	46	51	49	48	48	48	51	50	49	46	50	49	47	50	41	46
[9] CPMMV:BR:GO:10:5	46	41	35	99	81	99	99	99		60	54	48	45	49	38	39	50	47	47	46	46	50	48	48	48	47	51	49	49	46	49	50	48	50	41	46
[10] CPMMV HQ184471	46	43	35	62	60	62	61	61	61		54	48	47	48	43	44	47	50	51	46	49	50	48	49	48	49	50	50	51	48	49	52	50	51	44	48
[11] CuVCV_JN591720	44	40	36	50	49	48	49	49	49	50		47	46	50	40	39	48	47	49	43	48	46	45	46	50	49	49	48	49	45	50	49	47	49	44	48
[12] CVB AB245142	48	50	42	42	41	42	43	43	43	38	41		47	52	48	49	52	54	52	48	55	58	56	55	50	47	57	56	54	46	50	54	54	52	45	50
[13] CVNV EF527260	34	34	42	29	28	29	29	29	29	33	38	34		44	47	49	44	47	46	45	44	44	44	46	46	47	48	46	44	45	44	45	43	47	43	47
[14] DVS AJ620300	53	51	37	41	43	41	41	41	41	40	42	47	30		47	44	54	55	54	50	51	55	53	51	48	45	54	52	54	48	54	53	55	56	44	50
[15] GarCLV Z68502	27	25	30	17	18	17	18	17	18	29	26	29	32	31		75	44	45	44	44	44	47	43	44	40	42	45	45	42	42	44	42	44	45	52	42
[16] GarLV AJ292226	31	27	34	26	30	26	26	26	26	31	30	38	36	32	77		45	45	46	44	45	45	46	43	43	44	45	45	45	42	44	44	45	45	53	45
[17] HCMoV EU754720	47	56	41	34	38	38	38	38	38	37	37	47	31	50	32	36		53	52	49	59	59	58	52	48	47	54	51	62	45	50	53	55	57	44	47
[18] HpLV AB032469	61	53	38	42	42	42	42	42	42	46	42	50	36	55	27	32	51		60	51	52	54	55	61	53	47	53	57	56	44	54	61	56	53	44	52
[19] HpMV EU527979	58	54	32	38	39	38	38	38	38	47	43	46	34	54	27	29	46	61		47	50	52	52	60	48	47	53	52	53	45	55	61	54	53	45	47
[20] HNNV FJ196835	40	43	43	37	36	37	36	36	36	34	36	39	31	42	29	31	35	43	37		50	50	48	48	48	46	48	51	49	46	48	48	52	48	44	47
[21] KLV FJ531635	49	56	32	39	36	39	39	39	39	39	38	49	31	49	35	35	63	48	46	38		63	62	52	48	46	55	54	63	45	50	53	57	54	43	49
[22] LNRSV EU074853	52	60	37	29	41	41	41	41	41	42	34	53	34	51	33	34	59	50	44	39	64		62	53	50	48	57	55	64	50	54	56	58	61	44	47
[23] LSV AJ516059	49	61	33	40	41	40	40	40	40	42	37	52	32	50	33	35	56	50	46	33	63	65		56	47	47	53	53	60	45	49	56	55	56	45	47
[24] NCLV AM158439	60	52	38	42	40	42	43	43	43	46	42	51	33	48	28	31	47	62	64	41	49	52	51		52	44	55	57	55	43	50	61	55	53	43	47
[25] NeLV DQ098905	47	46	34	39	43	39	39	40	40	39	41	41	32	45	23	28	39	49	42	38	40	42	40	44		47	49	49	51	43	47	49	49	50	44	48
[26] NSV AM182569	36	33	36	34	31	34	34	35	35	36	37	33	37	33	30	33	32	36	30	33	37	35	34	36	35		47	49	45	46	48	46	46	46	43	52
[27] PhlVB EU162589	52	48	41	43	41	43	44	44	44	41	36	54	33	50	35	37	52	52	47	39	54	59	50	53	43	35		57	55	47	51	55	56	57	42	47
[28] PhlVS EF492068	55	55	34	40	40	40	41	41	41	44	38	51	33	48	33	36	48	52	54	39	54	53	51	55	40	39	54		58	47	50	58	58	51	42	46
[29] PLV DQ455582	54	63	40	37	39	37	36	36	36	43	38	50	33	49	28	31	65	52	48	40	70	66	62	52	39	35	52	56		45	51	55	60	58	44	46
[30] PotLV EU433397	45	49	39	43	44	43	43	43	43	45	43	45	35	51	19	32	42	50	47	35	41	45	45	45	40	32	48	44	42		48	45	45	48	42	45
[31] PopMV AY505475	42	38	37	36	35	36	36	43	43	39	35	37	39	39	27	31	32	39	36	36	36	34	32	36	36	35	39	38	33	40		51	54	52	42	48
[32] PVM D14449	59	50	38	43	44	43	44	43	43	47	44	51	32	51	31	35	50	66	65	43	49	52	51	64	42	34	52	55	51	43	39		58	54	40	47
[33] PVP EU338239	53	59	37	43	40	43	43	43	43	41	41	49	32	50	30	32	59	58	50	41	55	59	56	51	41	33	52	53	60	49	33	56		56	45	45
[34] PVS AJ863509	49	59	37	40	41	40	40	40	40	39	38	48	35	50	31	32	57	52	50	39	56	59	56	49	49	37	52	51	57	46	35	47	58		43	49
[35] RCVMV FJ685618	29	29	30	29	26	29	30	30	30	32	30	31	27	33	39	46	30	29	25	28	29	26	29	25	28	23	29	27	30	29	31	27	27	31		41
[36] SPCFV AY461421	39	32	29	38	37	38	39	39	39	35	34	40	34	36	28	32	31	35	35	35	32	34	34	36	32	36	35	39	31	37	30	35	34	38	26	

(C)



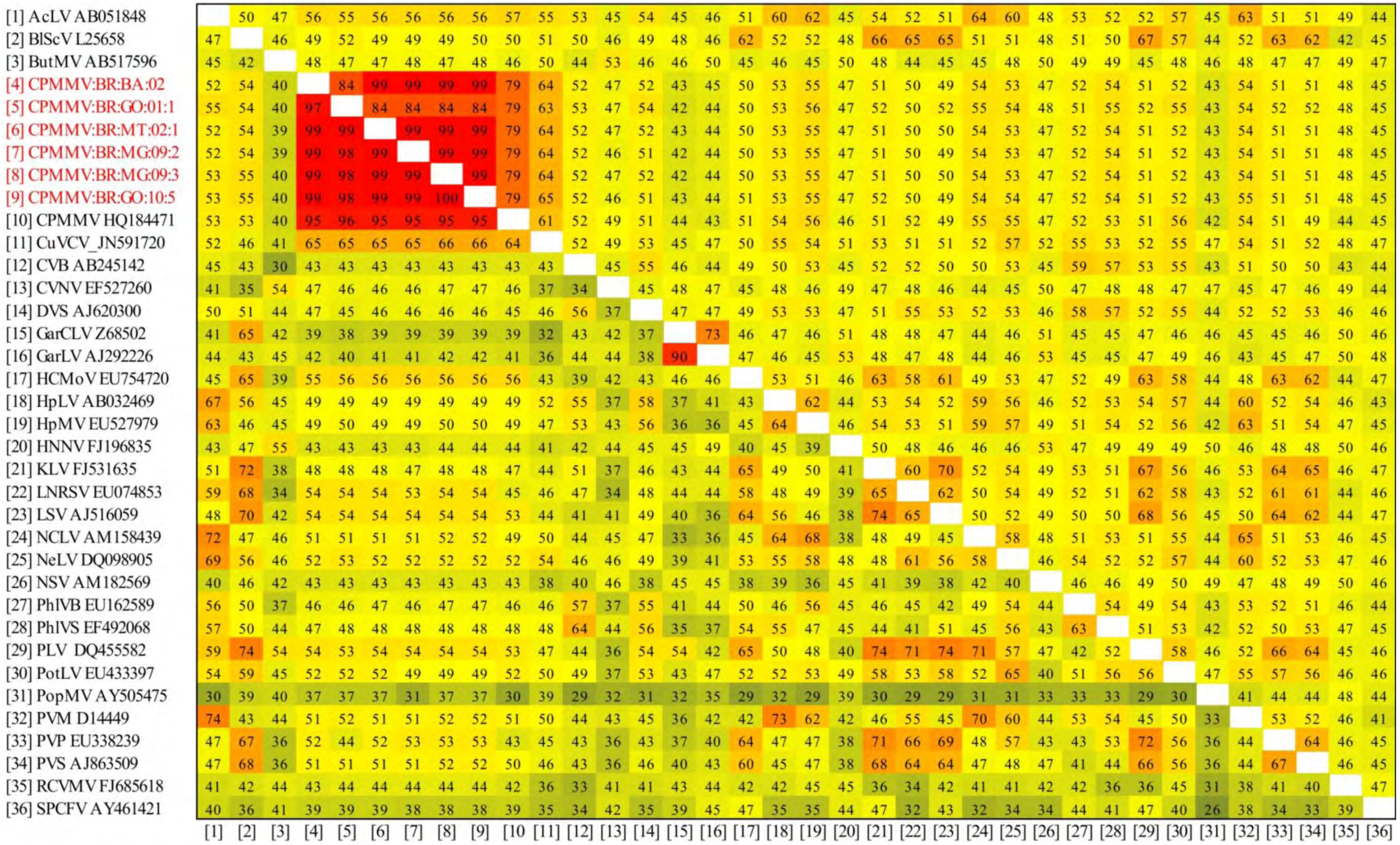
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]
[1] AcLV AB051848		54	48	43	42	43	43	44	43	44	44	48	47	54	51	49	54	59	55	47	51	56	52	59	52	45	52	51	53	47	55	52	58	53	48	44
[2] BlScV L25658	51		50	43	44	44	43	44	44	48	46	54	46	53	47	46	61	54	56	52	60	65	61	62	54	51	50	49	62	49	61	56	56	59	46	45
[3] ButMV AB517596	41	38		47	46	47	47	47	47	46	45	46	56	51	55	52	53	42	49	49	46	48	49	50	50	59	48	48	50	52	48	48	50	48	51	46
[4] CPMMV:BR:BA:02	28	36	31		78	99	98	98	99	62	45	45	43	47	43	45	46	40	41	44	44	45	43	48	46	49	43	45	41	43	45	47	44	43	43	42
[5] CPMMV:BR:GO:01:1	26	37	34	85		77	77	77	78	61	47	44	42	48	44	45	45	42	41	50	46	44	47	50	46	46	44	47	42	44	47	45	41	43	39	43
[6] CPMMV:BR:MT:02:1	28	36	31	100	85		98	99	99	62	45	45	43	47	44	45	46	40	42	44	44	45	43	49	46	49	43	45	41	44	46	47	45	43	43	42
[7] CPMMV:BR:MG:09:2	26	34	30	98	83	98		98	98	62	46	44	44	47	43	45	46	40	41	45	44	45	44	49	45	49	43	45	41	44	46	47	44	43	42	43
[8] CPMMV:BR:MG:09:3	27	35	30	99	84	99	97		99	63	45	45	43	48	44	45	46	41	42	45	44	45	43	48	45	49	43	45	41	44	45	47	44	43	42	42
[9] CPMMV:BR:GO:10:5	28	36	31	100	85	100	98	99		62	46	45	43	48	44	45	46	40	41	45	44	45	43	48	45	49	43	45	41	44	45	47	44	43	43	43
[10] CPMMV HQ184471	37	36	38	63	69	63	61	62	63		52	46	49	49	47	47	47	45	45	49	44	47	47	48	44	48	45	44	45	50	45	47	48	42	45	45
[11] CuVCV_JN591720	28	29	30	33	43	33	39	32	33	49		46	45	44	45	43	45	46	44	42	47	47	50	45	48	49	39	45	43	45	45	46	42	44	48	39
[12] CVB AB245142	41	52	37	35	37	35	33	34	35	37	28		47	52	45	46	51	53	57	49	52	54	54	52	49	47	57	55	55	48	52	50	52	51	47	43
[13] CVNV EF527260	42	42	46	34	30	34	33	33	34	36	32	43		50	53	52	45	46	52	48	46	45	47	49	46	51	48	51	48	52	48	44	48	46	51	41
[14] DVS AJ620300	49	47	49	32	34	32	30	31	32	35	44	51	30		49	50	54	52	52	49	51	52	52	51	47	50	49	54	56	50	53	52	51	50	50	43
[15] GarCLV Z68502	38	42	45	34	38	34	32	33	34	36	48	39	33	38		75	46	43	47	54	48	50	45	46	50	50	46	49	48	56	48	44	49	46	50	40
[16] GarLV AJ292226	40	45	50	45	38	45	40	44	45	35	47	40	35	41	79		48	43	45	56	45	46	46	47	50	51	48	46	48	52	50	45	50	45	48	42
[17] HCMoV EU754720	51	65	45	41	38	41	39	40	41	41	40	43	28	44	37	42		52	52	47	56	61	61	56	51	52	51	51	55	49	58	58	56	55	47	43
[18] HpLV AB032469	62	49	36	31	33	31	30	30	31	36	44	50	36	51	30	31	42		59	50	54	53	53	56	51	44	50	49	54	46	55	53	52	54	45	47
[19] HpMV EU527979	54	54	45	24	39	34	33	33	34	44	50	53	33	41	46	36	39	62		46	57	58	55	57	52	46	50	52	56	47	56	52	56	54	48	42
[20] HNNV FJ196835	38	36	47	34	36	34	33	33	34	39	52	41	33	36	52	60	41	36	39		48	46	50	48	44	48	46	46	49	56	50	47	44	48	49	50
[21] KLV FJ531635	45	58	41	37	37	37	35	36	37	31	34	47	29	43	40	40	56	48	55	38		64	61	56	49	53	46	47	54	49	64	57	53	60	44	46
[22] LNRSV EU074853	47	69	37	44	43	44	42	43	44	41	35	50	30	37	37	41	64	48	53	38	71		63	58	51	48	47	48	56	45	62	57	56	61	48	45
[23] LSV AJ516059	45	64	39	36	38	36	35	35	36	39	42	53	35	50	36	42	63	52	53	40	61	68		64	51	52	50	50	57	45	66	56	56	61	44	49
[24] NCLV AM158439	58	55	43	38	42	38	37	37	38	39	45	46	31	45	41	41	48	57	60	38	48	55	58		52	48	53	55	55	50	59	53	60	54	44	48
[25] NeLV DQ098905	42	47	33	38	48	38	37	37	38	43	43	42	37	42	35	42	48	48	42	36	42	45	43	47		47	47	52	51	47	52	53	51	49	46	45
[26] NSV AM182569	34	41	48	40	39	40	39	39	40	38	41	39	31	45	39	42	41	43	44	42	39	42	44	42	41		48	46	48	48	50	49	47	46	49	44
[27] PhlVB EU162589	52	44	42	39	38	39	37	42	39	40	40	58	27	47	41	43	42	47	50	42	42	43	45	48	45	38		61	58	46	48	49	47	49	46	42
[28] PhlVS EF492068	47	41	37	33	38	33	32	32	33	36	45	55	29	51	45	38	43	45	50	39	41	44	44	48	43	38	60		57	51	51	47	50	49	49	44
[29] PLV DQ455582	46	66	38	42	42	42	40	41	42	36	45	46	29	47	41	44	58	49	58	44	68	72	69	58	44	41	46	42		44	57	55	53	51	49	47
[30] PotLV EU433397	51	62	43	37	39	37	34	36	37	39	42	56	30	50	42	45	52	47	63	44	54	57	54	48	45	42	57	54	57		47	47	46	43	49	45
[31] PopMV AY505475	44	43	49	32	35	32	30	31	32	41	50	47	32	44	57	56	43	42	42	52	41	42	39	44	44	37	46	50	46	44		56	58	63	44	49
[32] PVM D14449	56	46	41	34	36	34	32	33	34	46	43	47	31	49	38	40	55	53	57	35	45	52	58	60	47	35	49	49	51	54	44		56	55	44	45
[33] PVP EU338239	56	61	37	42	39	42	41	41	42	38	40	49	35	43	39	41	54	45	54	32	60	67	60	50	42	36	46	42	61	51	36	51		56	49	41
[34] PVS AJ863509	47	63	40	42	42	42	41	41	42	38	42	53	30	46	40	43	58	51	52	41	60	61	59	45	50	42	46	41	60	53	45	51	54		46	42
[35] RCVMV FJ685618	41	43	42	32	32	32	31	31	32	40	42	45	26	43	37	46	42	43	41	43	41	41	40	42	37	43	40	44	37	44	44	41	38	45		44
[36] SPCFV AY461421	38	33	39	39	43	39	39	38	39	41	36	35	27	38	33	31	39	37	36	41	35	40	34	39	44	33	39	38	44	40	36	34	38	46	37	

(D)



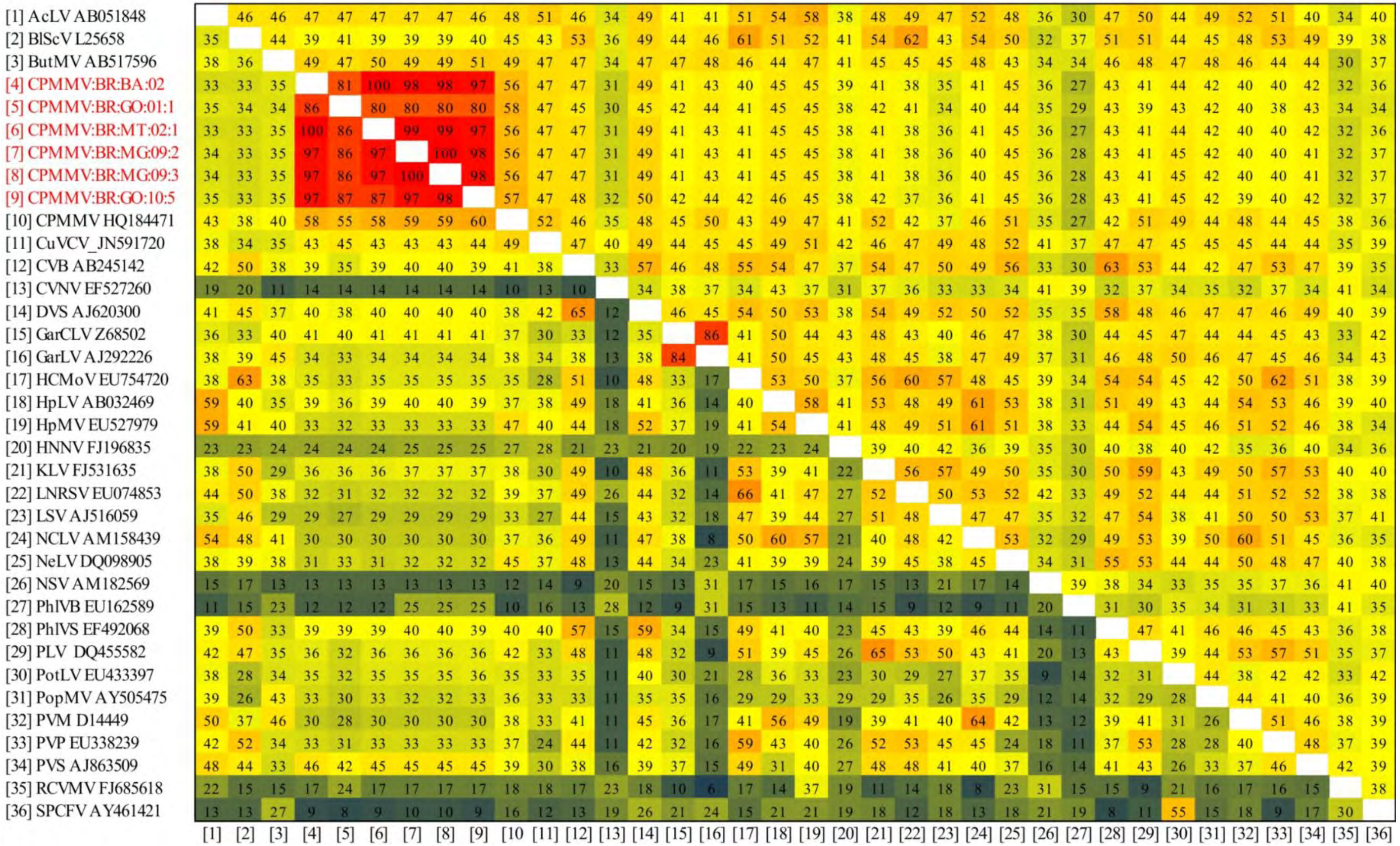
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]
[1] AcLV AB051848		45	39	37	39	37	37	37	38	41	42	44	42	46	38	41	48	48	42	42	40	44	38	54	43	42	39	40	40	40	46	52	40	42	43	38
[2] BlScV L25658	31		42	44	42	44	44	44	44	43	39	42	33	38	35	42	51	38	44	38	52	52	51	45	45	41	34	39	54	37	41	40	45	39	44	39
[3] ButMV AB517596	24	25		34	39	34	34	34	34	38	39	41	38	44	45	45	42	42	40	41	34	39	43	44	50	40	43	41	40	45	37	43	45	48	43	35
[4] CPMMV:BR:BA:02	23	22	21		76	100	99	100	99	57	38	42	38	39	41	38	38	42	39	39	38	39	37	38	38	39	43	39	38	39	39	42	38	41	39	36
[5] CPMMV:BR:GO:01:1	24	25	19	78		76	75	76	77	56	39	43	42	40	39	38	37	41	40	41	37	42	36	40	43	39	40	39	42	40	37	45	41	44	38	33
[6] CPMMV:BR:MT:02:1	23	22	21	100	78		99	100	99	57	38	42	38	39	41	38	38	42	39	39	38	39	37	38	38	39	43	39	38	39	39	42	38	41	39	36
[7] CPMMV:BR:MG:09:2	21	22	22	96	74	96		99	97	57	38	42	37	39	41	38	38	41	39	39	40	39	38	39	38	40	43	40	39	39	40	42	38	41	40	36
[8] CPMMV:BR:MG:09:3	23	22	21	100	78	100	96		99	57	38	42	38	39	41	38	38	42	39	39	38	39	37	38	38	39	43	39	38	39	39	42	38	41	39	36
[9] CPMMV:BR:GO:10:5	23	22	21	99	76	99	94	99		57	37	41	37	39	40	39	37	43	39	41	38	39	38	39	38	39	43	39	38	39	40	42	38	42	38	36
[10] CPMMV HQ184471	26	20	26	47	51	47	43	47	47		39	44	34	38	41	37	37	35	35	37	37	31	32	39	39	36	44	38	37	32	41	38	35	40	35	32
[11] CuVCV_JN591720	27	28	18	25	19	25	13	25	24	28		40	36	35	39	36	42	38	42	32	39	40	36	39	41	39	39	42	40	35	40	37	34	40	42	36
[12] CVB AB245142	25	23	25	19	20	19	19	19	19	20	23		39	41	37	42	37	40	43	41	41	37	35	37	39	41	41	38	41	40	46	40	39	38	43	40
[13] CVNV EF527260	33	23	26	19	21	19	21	19	19	20	18	27		39	39	41	39	39	36	45	33	38	33	40	38	41	43	42	37	37	44	41	39	43	39	34
[14] DVS AJ620300	29	27	31	19	18	19	20	18	19	28	21	27	30		39	42	37	40	36	36	31	38	34	40	47	45	46	44	33	41	44	37	34	40	46	38
[15] GarCLV Z68502	30	13	31	19	20	19	16	19	19	23	23	14	28	27		75	43	40	39	51	38	34	38	38	36	46	44	44	36	40	37	35	39	45	40	33
[16] GarLV AJ292226	27	22	20	14	16	14	13	14	14	19	16	22	31	33	70		41	40	40	47	39	34	34	45	43	45	41	45	35	41	41	39	36	42	43	37
[17] HCMoV EU754720	37	48	25	18	22	18	18	18	18	21	32	21	25	24	21	21		48	44	39	48	47	44	47	41	45	41	35	46	43	42	45	45	48	41	34
[18] HpLV AB032469	37	27	32	22	22	22	22	22	22	40	20	22	23	24	27	27	37		48	39	42	41	43	45	44	46	41	40	42	39	36	50	43	42	45	37
[19] HpMV EU527979	31	35	25	16	17	16	17	16	16	23	26	26	25	30	26	29	25	35		38	42	40	38	47	45	40	46	46	47	37	38	48	39	43	40	40
[20] HNNV FJ196835	24	22	31	16	13	16	16	16	16	18	19	25	39	31	38	36	22	22	24		42	40	36	41	43	42	46	46	41	39	48	40	47	44	38	30
[21] KLV FJ531635	34	52	24	13	11	13	16	13	13	21	35	23	27	25	23	27	42	32	32	26		47	52	42	33	43	35	35	63	44	37	39	48	41	32	34
[22] LNRSV EU074853	25	30	24	21	19	21	21	19	21	22	27	20	22	13	12	17	35	22	27	16	37		52	37	39	36	34	39	51	48	39	42	56	46	38	37
[23] LSV AJ516059	25	44	28	14	19	14	16	14	14	20	28	31	23	32	17	21	38	35	25	23	47	29		41	35	38	34	32	49	35	32	33	44	40	31	33
[24] NCLV AM158439	34	32	34	16	16	16	15	16	16	27	21	32	25	28	27	32	31	31	43	43	30	22	30		42	33	42	43	42	37	42	50	39	46	40	33
[25] NeLV DQ098905	27	29	34	22	27	22	22	22	22	20	14	24	22	31	22	27	24	29	24	36	25	20	27	28		48	43	44	34	39	39	44	37	44	47	38
[26] NSV AM182569	34	26	22	20	11	20	25	20	20	20	26	23	30	34	29	33	32	29	23	33	26	10	24	26	22		48	44	39	39	37	39	39	35	51	32
[27] PhlVB EU162589	33	23	34	23	24	23	19	23	23	24	33	38	37	31	37	39	26	28	31	39	32	36	26	34	25	33		47	38	38	42	46	42	45	41	35
[28] PhlVS EF492068	25	19	28	27	11	27	31	27	27	27	27	23	33	40	37	37	19	32	37	44	22	22	22	28	24	29	35		38	37	42	42	41	41	48	36
[29] PLV DQ455582	34	52	23	26	17	28	13	28	28	26	27	16	23	19	4	22	44	33	35	22	58	29	47	32	22	23	21	18		41	37	38	53	40	37	33
[30] PotLV EU433397	26	23	25	26	28	26	25	26	26	26	17	21	22	29	29	32	28	23	27	25	30	22	32	30	32	21	26	27	23		37	42	42	41	38	36
[31] PopMV AY505475	34	33	22	21	21	21	21	21	21	20	27	20	44	39	28	27	24	18	25	42	22	24	25	26	25	33	34	33	30	18		40	36	41	46	35
[32] PVM D14449	34	30	27	16	15	16	16	16	16	23	26	29	23	24	31	32	24	34	40	30	30	13	20	41	29	21	34	30	30	30	24		44	48	32	40
[33] PVP EU338239	21	15	36	17	19	17	17	16	17	31	14	24	22	23	22	22	30	27	25	24	33	44	33	32	25	18	29	27	30	22	16	23		46	34	33
[34] PVS AJ863509	20	24	44	26	24	26	26	26	26	26	14	30	27	29	30	27	22	22	31	28	26	26	24	32	29	18	34	24	24	22	32	31	37		40	31
[35] RCVMV FJ685618	14	27	23	20	16	20	20	18	20	20	17	27	38	32	25	32	29	24	28	30	26	10	24	34	22	34	34	38	17	15	32	24	19	20		37
[36] SPCFV AY461421	21	15	26	19	6	19	12	12	19	21	21	22	17	20	25	19	13	9	13	12	24	16	19	20	16	19	14	17	13	15	18	10	17	17	20	

(E)



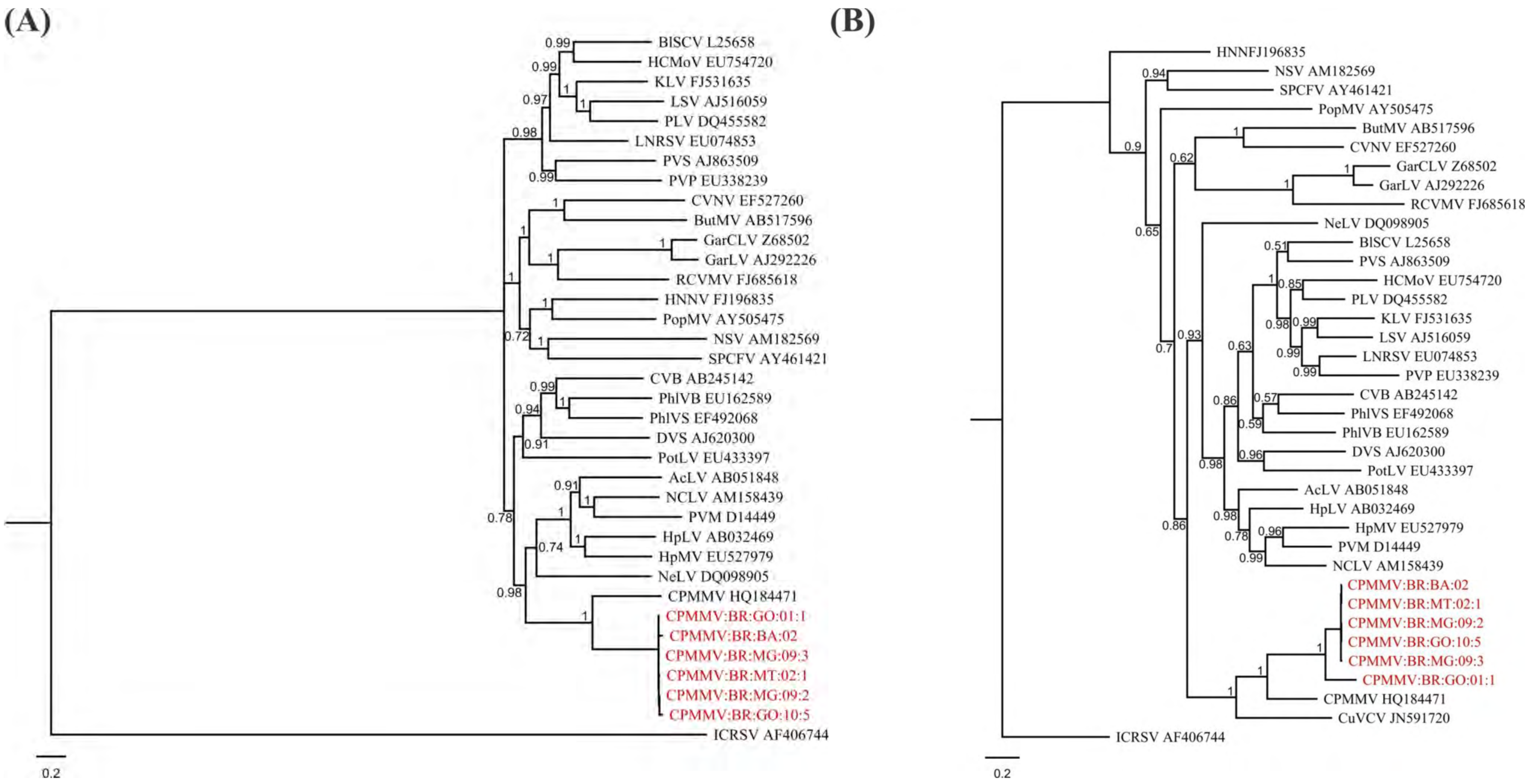
[1] AcLV AB051848		50	47	56	55	56	56	56	56	57	55	53	45	54	45	46	51	60	62	45	54	52	51	64	60	48	53	52	52	57	45	63	51	51	49	44
[2] BlScV L25658	47		46	49	52	49	49	49	50	50	51	50	46	49	48	46	62	52	52	48	66	65	65	51	51	48	51	50	67	57	44	52	63	62	42	45
[3] ButMV AB517596	45	42		48	47	47	48	47	48	46	50	44	53	46	46	50	45	46	45	50	48	44	45	45	48	50	49	49	45	48	46	48	47	47	49	47
[4] CPMMV:BR:BA:02	52	54	40		84	99	99	99	99	79	64	52	47	52	43	45	50	53	55	47	51	50	49	54	53	47	52	54	51	52	43	54	51	51	48	45
[5] CPMMV:BR:GO:01:1	55	54	40	97		84	84	84	84	79	63	53	47	54	42	44	50	53	56	47	52	50	52	55	54	48	51	55	52	55	43	54	52	52	48	45
[6] CPMMV:BR:MT:02:1	52	54	39	99	99		99	99	99	79	64	52	47	52	43	44	50	53	55	47	51	50	50	54	53	47	52	54	51	52	43	54	51	51	48	45
[7] CPMMV:BR:MG:09:2	52	54	39	99	98	99		99	99	79	64	52	46	51	42	44	50	53	55	47	51	50	49	54	53	47	52	54	51	52	43	54	51	51	48	45
[8] CPMMV:BR:MG:09:3	53	55	40	99	98	99	99		99	79	64	52	47	52	42	44	50	53	55	47	51	50	50	54	53	47	52	54	51	52	43	54	51	51	48	45
[9] CPMMV:BR:GO:10:5	53	55	40	99	98	99	99	100		79	65	52	46	51	43	44	51	53	55	47	52	50	49	54	54	47	52	54	51	52	43	55	51	51	48	45
[10] CPMMV HQ184471	53	53	40	95	96	95	95	95	95		61	52	49	51	44	43	51	54	56	46	51	52	49	55	55	47	52	53	51	56	42	54	51	49	44	45
[11] CuVCV_JN591720	52	46	41	65	65	65	65	66	66	64		52	49	53	45	47	50	55	54	51	53	51	51	52	57	52	55	53	52	55	47	54	51	52	48	47
[12] CVB AB245142	45	43	30	43	43	43	43	43	43	43	43		45	55	46	44	49	50	53	45	52	52	50	50	53	45	59	57	53	55	43	51	50	50	43	44
[13] CVNV EF527260	41	35	54	47	46	46	46	47	47	46	37	34		45	48	47	45	48	46	49	47	48	46	44	45	50	47	48	48	47	47	45	47	46	49	44
[14] DVS AJ620300	50	51	44	47	45	46	46	46	46	45	46	56	37		47	47	49	53	53	47	51	55	53	52	53	46	58	57	52	55	44	52	52	53	46	46
[15] GarCLV Z68502	41	65	42	39	38	39	39	39	39	39	32	43	42	37		73	46	47	46	51	48	48	47	44	46	51	45	45	47	46	46	45	45	46	50	46
[16] GarLV AJ292226	44	43	45	42	40	41	41	42	42	41	36	44	44	38	90		47	46	45	53	48	47	48	44	46	53	45	45	47	49	46	43	45	47	50	48
[17] HCMoV EU754720	45	65	39	55	56	56	56	56	56	56	43	39	42	43	46	46		53	51	46	63	58	61	49	53	47	52	49	63	58	44	48	63	62	44	47
[18] HpLV AB032469	67	56	45	49	49	49	49	49	49	49	52	55	37	58	37	41	43		62	44	53	54	52	59	56	46	52	53	54	57	44	60	52	54	46	43
[19] HpMV EU527979	63	46	45	49	50	49	49	50	50	49	47	53	43	56	36	36	45	64		46	54	53	51	59	57	49	51	54	52	56	42	63	51	54	47	45
[20] HNNV FJ196835	43	47	55	43	43	43	43	44	44	44	41	42	44	45	45	49	40	45	39		50	48	46	46	46	53	47	49	49	49	50	46	48	48	50	46
[21] KLV FJ531635	51	72	38	48	48	48	47	48	48	47	44	51	37	46	43	44	65	49	50	41		60	70	52	54	49	53	51	67	56	46	53	64	65	46	47
[22] LNRSV EU074853	59	68	34	54	54	54	53	54	54	45	46	47	34	48	44	44	58	48	49	39	65		62	50	54	49	52	51	62	58	43	52	61	61	44	46
[23] LSV AJ516059	48	70	42	54	54	54	54	54	54	53	44	41	41	49	40	36	64	56	46	38	74	65		50	52	49	50	50	68	56	45	50	64	62	44	47
[24] NCLV AM158439	72	47	46	51	51	51	51	52	52	49	50	44	45	47	33	36	45	64	68	38	48	49	45		58	48	51	53	51	55	44	65	51	53	46	45
[25] NeLV DQ098905	69	56	46	52	53	52	52	52	52	52	54	46	46	49	39	41	53	55	58	48	48	61	56	58		46	54	52	52	57	44	60	52	53	47	46
[26] NSV AM182569	40	46	42	43	43	43	43	43	43	43	38	40	46	38	45	45	38	39	36	45	41	39	38	42	40		46	46	49	50	49	47	48	49	50	46
[27] PhlVB EU162589	56	50	37	46	46	47	46	47	47	46	46	57	37	55	41	44	50	46	56	45	46	45	42	49	54	44		54	49	54	43	53	52	51	46	44
[28] PhlVS EF492068	57	50	44	47	48	48	48	48	48	48	48	64	44	56	35	37	54	55	47	45	44	41	51	45	56	43	63		51	53	42	52	50	53	47	45
[29] PLV DQ455582	59	74	54	54	53	54	54	54	54	53	47	44	36	54	54	42	65	50	48	40	74	71	74	71	57	47	42	52		58	46	52	66	64	45	46
[30] PotLV EU433397	54	59	45	52	52	52	49	49	49	52	50	49	37	53	43	47	52	52	53	49	58	53	58	52	65	40	51	56	56		47	55	57	56	46	46
[31] PopMV AY505475	30	39	40	37	37	37	31	37	37	30	39	29	32	31	32	35	29	32	29	39	30	29	29	31	31	33	33	33	29	30		41	44	44	48	44
[32] PVM D14449	74	43	44	51	52	51	51	52	52	51	50	44	43	45	36	42	42	73	62	42	46	55	45	70	60	44	53	54	45	50	33		53	52	46	41
[33] PVP EU338239	47	67	36	52	44	52	53	53	53	43	45	43	36	43	37	40	64	47	47	38	71	66	69	48	57	43	43	53	72	56	36	44		64	46	45
[34] PVS AJ863509	47	68	36	51	51	51	51	52	52	50	46	43	36	46	40	43	60	45	47	38	68	64	64	47	48	47	41	44	66	56	36	44	67		46	45
[35] RCVMV FJ685618	41	42	44	43	44	44	44	44	44	42	36	33	41	41	43	44	42	42	45	45	36	34	42	41	41	42	42	36	36	45	31	38	41	40		47
[36] SPCFV AY461421	40	36	41	39	39	39	38	38	38	39	35	34	42	35	39	45	47	35	35	44	47	32	43	32	34	34	44	41	47	40	26	38	34	33	39	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]

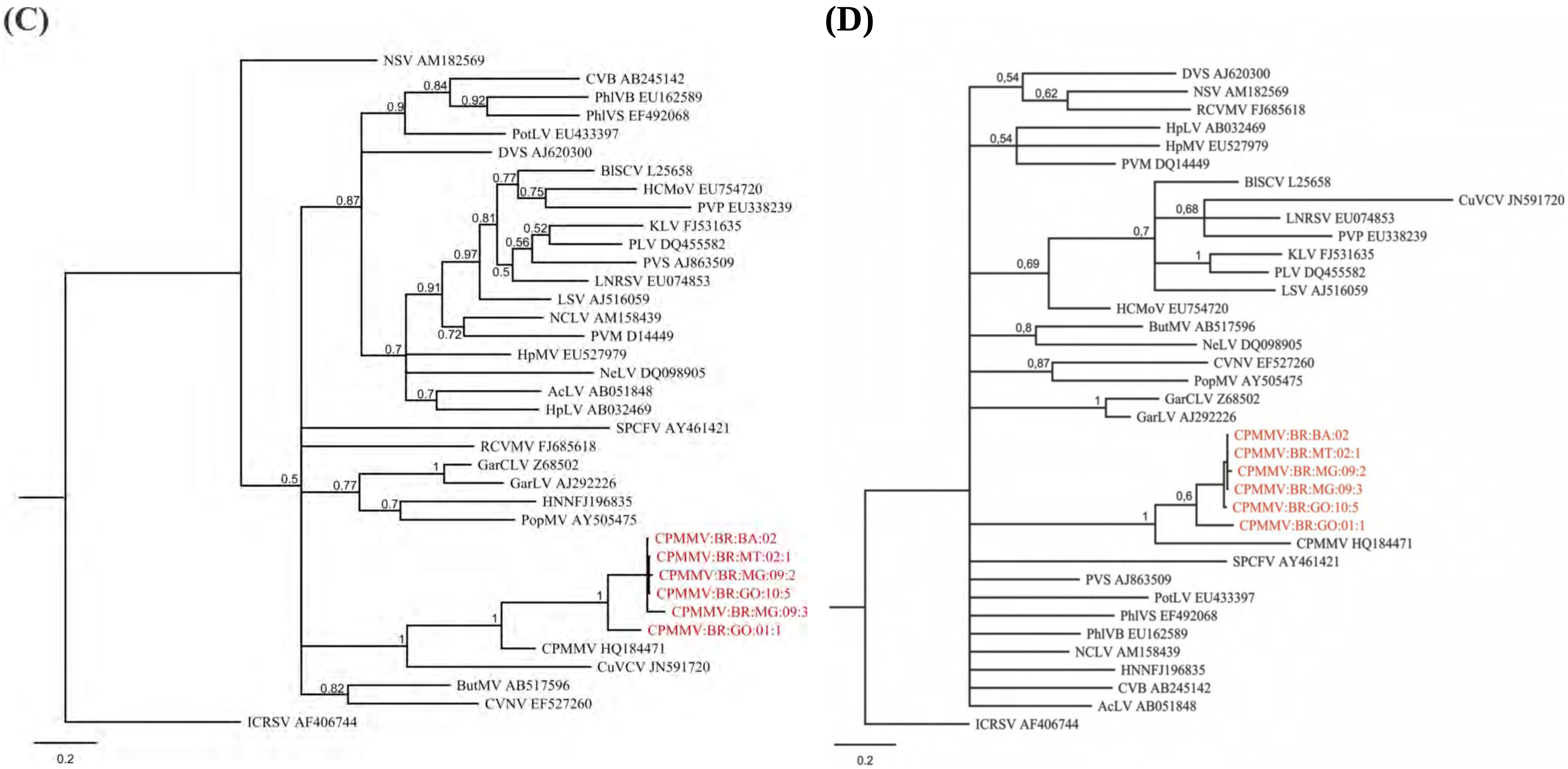
(F)

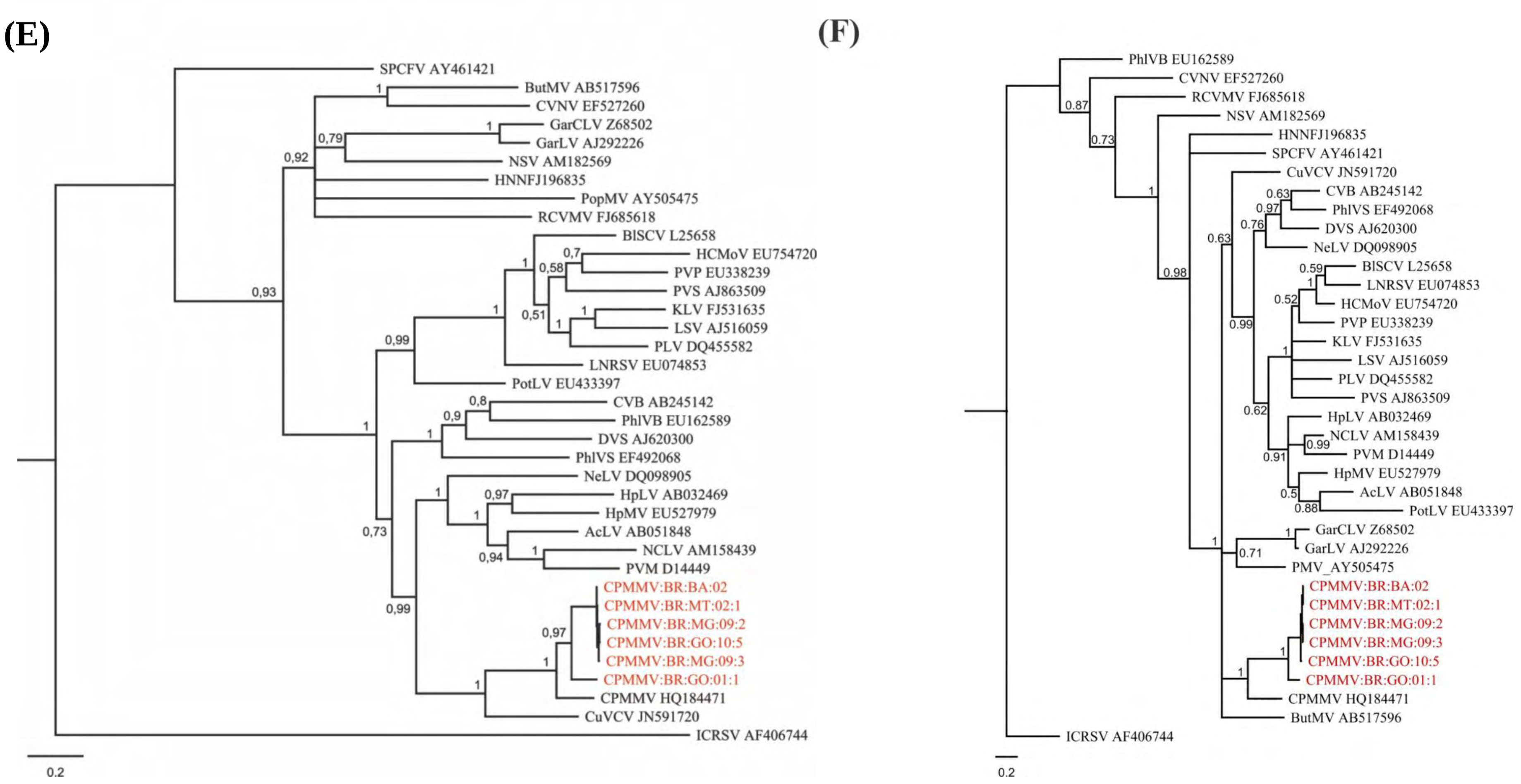


	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]
[1] AcLV AB051848		46	46	47	47	47	47	47	46	48	51	46	34	49	41	41	51	54	58	38	48	49	47	52	48	36	30	47	50	44	49	52	51	40	34	40
[2] BlScV L25658	35		44	39	41	39	39	39	40	45	43	53	36	49	44	46	61	51	52	41	54	62	43	54	50	32	37	51	51	44	45	48	53	49	39	38
[3] ButMV AB517596	38	36		49	47	50	49	49	51	49	47	47	34	47	47	48	46	44	47	41	45	45	45	48	43	34	34	46	48	47	48	46	44	44	30	37
[4] CPMMV:BR:BA:02	33	33	35		81	100	98	98	97	56	47	47	31	49	41	43	40	45	45	39	41	38	35	41	45	36	27	43	41	44	42	40	40	42	32	36
[5] CPMMV:BR:GO:01:1	35	34	34	86		80	80	80	80	58	47	45	30	45	42	44	41	45	45	38	42	41	34	40	44	35	29	43	39	43	42	40	38	43	34	34
[6] CPMMV:BR:MT:02:1	33	33	35	100	86		99	99	97	56	47	47	31	49	41	43	41	45	45	38	41	38	36	41	45	36	27	43	41	44	42	40	40	42	32	36
[7] CPMMV:BR:MG:09:2	34	33	35	97	86	97		100	98	56	47	47	31	49	41	43	41	45	45	38	41	38	36	40	45	36	28	43	41	45	42	40	40	41	32	37
[8] CPMMV:BR:MG:09:3	34	33	35	97	86	97	100		98	56	47	47	31	49	41	43	41	45	45	38	41	38	36	40	45	36	28	43	41	45	42	40	40	41	32	37
[9] CPMMV:BR:GO:10:5	35	33	35	97	87	87	97	98		57	47	48	32	50	42	44	42	46	45	38	42	37	36	41	45	36	28	43	41	45	42	39	40	42	32	37
[10] CPMMV HQ184471	43	38	40	58	55	58	59	59	60		52	46	35	48	45	50	43	49	47	41	52	42	37	46	51	35	27	42	51	49	44	48	44	45	38	36
[11] CuVCV_JN591720	38	34	35	43	45	43	43	43	44	49		47	40	49	44	45	45	49	51	42	46	47	49	48	52	41	37	47	47	45	45	45	44	44	35	39
[12] CVB AB245142	42	50	38	39	35	39	40	40	39	41	38		33	57	46	48	55	54	47	37	54	47	50	49	56	33	30	63	53	44	42	47	53	47	39	35
[13] CVNV EF527260	19	20	11	14	14	14	14	14	14	10	13	10		34	38	37	34	43	37	31	37	36	33	33	34	41	39	32	37	34	35	32	37	34	41	34
[14] DVS AJ620300	41	45	37	40	38	40	40	40	40	38	42	65	12		46	45	54	50	53	38	54	49	52	50	52	35	35	58	48	46	47	47	46	49	40	39
[15] GarCLV Z68502	36	33	40	41	40	41	41	41	41	37	30	33	12	35		86	41	50	44	43	48	43	40	46	47	38	30	44	45	47	44	44	45	43	33	42
[16] GarLV AJ292226	38	39	45	34	33	34	34	34	34	38	34	38	13	38	84		41	50	45	43	48	45	38	47	49	37	31	46	48	50	46	47	45	46	34	43
[17] HCMoV EU754720	38	63	38	35	33	35	35	35	35	35	28	51	10	48	33	17		53	50	37	56	60	57	48	45	39	34	54	54	45	42	50	62	51	38	39
[18] HpLV AB032469	59	40	35	39	36	39	40	40	39	37	38	49	18	41	36	14	40		58	41	53	48	49	61	53	38	31	51	49	43	44	54	53	46	39	40
[19] HpMV EU527979	59	41	40	33	32	33	33	33	33	47	40	44	18	52	37	19	41	54		41	48	49	51	61	51	38	33	44	54	45	46	51	52	46	38	34
[20] HNNV FJ196835	23	23	24	24	24	24	25	25	25	27	28	21	23	21	20	19	22	23	24		39	40	42	36	39	35	30	40	38	40	42	35	36	40	34	36
[21] KLV FJ531635	38	50	29	36	36	36	37	37	37	38	30	49	10	48	36	11	53	39	41	22		56	57	49	50	35	30	50	59	43	49	50	57	53	40	40
[22] LNRSV EU074853	44	50	38	32	31	32	32	32	32	39	37	49	26	44	32	14	66	41	47	27	52		50	53	52	42	33	49	52	44	44	51	52	52	38	38
[23] LSV AJ516059	35	46	29	29	27	29	29	29	29	33	27	44	15	43	32	18	47	39	44	27	51	48		47	47	35	32	47	54	38	41	50	50	53	37	41
[24] NCLV AM158439	54	48	41	30	30	30	30	30	30	37	36	49	11	47	38	8	50	60	57	21	40	48	42		53	32	29	49	53	39	50	60	51	45	36	35
[25] NeLV DQ098905	38	39	38	31	33	31	32	32	32	45	37	48	13	44	34	23	41	39	39	24	39	45	38	45		34	31	55	53	44	44	50	48	47	40	38
[26] NSV AM182569	15	17	13	13	13	13	13	13	13	12	14	9	20	15	13	31	17	15	16	17	15	13	21	17	14		39	38	34	33	35	35	37	36	41	40
[27] PhlVB EU162589	11	15	23	12	12	12	25	25	25	10	16	13	28	12	9	31	15	13	11	14	15	9	12	9	11	20		31	30	35	34	31	31	33	41	35
[28] PhlVS EF492068	39	50	33	39	39	39	40	40	39	40	40	57	15	59	34	15	49	41	40	23	45	43	39	46	44	14	11		47	41	46	46	45	43	36	38
[29] PLV DQ455582	42	47	35	36	32	36	36	36	36	42	33	48	11	48	32	9	51	39	45	26	65	53	50	43	41	20	13	43		39	44	53	57	51	35	37
[30] PotLV EU433397	38	28	34	35	32	35	35	35	36	35	33	35	11	40	30	21	28	36	33	23	30	29	27	37	35	9	14	32	31		44	38	42	42	33	42
[31] PopMV AY505475	39	26	43	33	30	33	32	32	33	36	33	33	11	35	35	16	29	29	33	29	29	35	26	35	29	12	14	32	29	28		44	41	40	36	39
[32] PVM D14449	50	37	46	30	28	30	30	30	30	38	33	41	11	45	36	17	41	56	49	19	39	41	40	64	42	13	12	39	41	31	26		51	46	38	39
[33] PVP EU338239	42	52	34	33	31	33	33	33	33	37	24	44	11	42	32	16	59	43	40	26	52	53	45	45	24	18	11	37	53	28	28	40		48	37	39
[34] PVS AJ863509	48	44	33	46	42	45	45	45	45	39	30	38	16	39	37	15	49	31	40	27	48	48	41	40	37	16	14	41	43	26	33	37	46		42	39
[35] RCVMV FJ685618	22	15	15	17	24	17	17	17	17	18	18	17	23	18	10	6	17	14	37	19	11	14	18	8	23	31	15	15	9	21	16	17	16	15		38
[36] SPCFV AY461421	13	13	27	9	8	9	10	10	9	16	12	13	19	26	21	24	15	21	21	19	18	12	18	13	18	21	19	8	11	55	15	18	9	17	30	

Suplementary Figure S2




(C)
NSV AM182569
CVB AB245142
PhlVB EU162589
PhlVS EF492068
PotLV EU433397
DVS AJ620300
BlSCV L25658
HCMoV EU754720
PVP EU338239
KLV FJ531635
PLV DQ455582
PVS AJ863509
LNRSV EU074853
LSV AJ516059
NCLV AM158439
PVM D14449
HpMV EU527979
NeLV DQ098905
AcLV AB051848
HpLV AB032469
SPCFV AY461421
RCVMV FJ685618
GarCLV Z68502
GarLV AJ292226
HNNFJ196835
PopMV AY505475
CPMMV:BR:BA:02
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:2
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:3
CPMMV:BR:GO:01:1
CPMMV HQ184471
CuVCV JN591720
ButMV AB517596
CVNV EF527260
ICRSV AF406744
0.84
0.9
0.92
0.87
0.77
0.75
0.81
0.52
0.56
0.97
0.5
0.91
0.72
0.7
0.7
1
0.5
0.77
0.7
1
1
1
0.82
0.2
(D)
DVS AJ620300
NSV AM182569
RCVMV FJ685618
HpLV AB032469
HpMV EU527979
PVM DQ14449
BlSCV L25658
CuVCV JN591720
LNRSV EU074853
PVP EU338239
KLV FJ531635
PLV DQ455582
LSV AJ516059
HCMoV EU754720
ButMV AB517596
NeLV DQ098905
CVNV EF527260
PopMV AY505475
GarCLV Z68502
GarLV AJ292226
CPMMV:BR:BA:02
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:3
CPMMV:BR:GO:10:5
CPMMV:BR:GO:01:1
CPMMV HQ184471
SPCFV AY461421
PVS AJ863509
PotLV EU433397
PhlVS EF492068
PhlVB EU162589
NCLV AM158439
HNNFJ196835
CVB AB245142
AcLV AB051848
ICRSV AF406744
0,54
0,62
0,54
0,68
0,7
1
0,69
0,8
0,87
1
0,6
1
0.2


(E)
SPCFV AY461421
ButMV AB517596
CVNV EF527260
GarCLV Z68502
GarLV AJ292226
NSV AM182569
HNNFJ196835
PopMV AY505475
RCVMV FJ685618
BlSCV L25658
HCMoV EU754720
PVP EU338239
PVS AJ863509
KLV FJ531635
LSV AJ516059
PLV DQ455582
LNRSV EU074853
PotLV EU433397
CVB AB245142
PhlVB EU162589
DVS AJ620300
PhlVS EF492068
NeLV DQ098905
HpLV AB032469
HpMV EU527979
AcLV AB051848
NCLV AM158439
PVM D14449
CPMMV:BR:BA:02
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:2
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:3
CPMMV:BR:GO:01:1
CPMMV HQ184471
CuVCV JN591720
ICRSV AF406744
1
0,79
0,92
0,93
1
0,7
0,58
1
0,51
1
1
0,99
1
0,8
0,9
1
1
0,97
1
0,73
1
0,94
0,99
0,97
1
1
0.2
(F)
PhlVB EU162589
CVNV EF527260
RCVMV FJ685618
NSV AM182569
HNNFJ196835
SPCFV AY461421
CuVCV JN591720
CVB AB245142
PhlVS EF492068
DVS AJ620300
NeLV DQ098905
BlSCV L25658
LNRSV EU074853
HCMoV EU754720
PVP EU338239
KLV FJ531635
LSV AJ516059
PLV DQ455582
PVS AJ863509
HpLV AB032469
NCLV AM158439
PVM D14449
HpMV EU527979
AcLV AB051848
PotLV EU433397
GarCLV Z68502
GarLV AJ292226
PMV_AY505475
CPMMV:BR:BA:02
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:3
CPMMV:BR:GO:10:5
CPMMV:BR:GO:01:1
CPMMV HQ184471
ButMV AB517596
ICRSV AF406744
0.87
0.73
1
0.98
0.63
0.97
0.76
0.63
0.59
1
0.52
0.99
1
0.62
0.99
0.91
0.5
0.88
1
0.71
1
1
1
0.2

CHAPTER 2

MOLECULAR VARIABILITY OF *Cowpea mild mottle virus* INFECTING SOYBEAN IN BRAZIL

Zanardo, L.G., Silva, F.N., Lima, A.T.M., Castillo-Urquiza, G.P., Milanesi, D.F.M., Almeida, A.M.R., Zerbini, F.M., Carvalho, C.M. Molecular variability of *Cowpea mild mottle virus* infecting soybean in Brazil. Archives of Virology, *Submitted*.

Molecular variability of *Cowpea mild mottle virus* infecting soybean in Brazil

L. G. Zanardo[1]; F. N. Silva[1]; A. T. M. Lima[1]; D. F. Milanesi[1]; G. P. Castilho-Urquiza[1]; A.M.R. Almeida[2]; F. M. Zerbini[1]; C. M.Carvalho[1*]

[1]Departamento de Fitopatologia /BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, Brazil, 36570-000

[2]Embrapa Soja, Londrina, PR, Brazil, 86001-970.

* Corresponding author: claudine.carvalho@ufv.br

Phone: (+55-31) 3899-1087; Fax: (+55-31) 3899-2240;

Keywords: CPMMV, *Carlavirus*, soybean and molecular variability

Abstract

We report the molecular variability of eighteen isolates of *Cowpea mild mottle virus* (CPMMV, *Carlavirus* genus, *Betaflexiviridae* family) infecting soybean fields of different Brazilian states (Bahia, Goiás, Maranhão, Mato Grosso, Minas Gerais and Pará) during the years of 2001 and 2010. The isolates showed a variety of symptoms in soybean cv. CD206, ranging from mild (crinkle/blistering leaves, mosaic and vein clearing) to severe (bud blight, dwarfism, leaves and stem necrosis). Recombination analysis showed that only one CPMMV isolate had a recombinant portion among the eighteen evaluated. Pairwise comparisons and phylogenetic analysis were performed for partial genomes (ORF2-3'terminus) and for each ORF individually (ORF2, 3, 4, 5 and 6), showing the isolates to be distinct. The phylogenetic tree did not show clustering based on the year of collection or geographical origin; some groupings were based on

symptoms. Additionally, the phylogenetic analysis made clear the existence of two distinct strains of the virus, (CPMMV-BR1 and CPMMV-BR2), with molecular variability between these. This is the first study of the molecular variability of CPMMV and is the first time that a large number of CPMMV isolates have been sampled and sequenced.

Introduction

Soybean stem necrosis disease is caused by the carlavirus *Cowpea mild mottle virus* (CPMMV, Family *Betaflexiviridae*, Genus *Carlavirus*). Symptoms of viral infection in soybean plants are variable and include bud blight, necrosis of stem and petiole, dwarfism, mosaic and foliar deformation, with the presence of blisters [3]. The genome of carlaviruses is composed of a single-stranded positive sense RNA molecule (7.8-8.9 Kb) encapsidated in flexuous filamentous particles (10-15 x 650-700 nm) [3, 26, 28]. The non-segmented viral RNA possess a cap structure [m7GpppG] linked to the 5’-terminus region and a poly(A) tail at its 3′ -terminus. It contains usually six Open Reading Frames (ORFs): ORF1 encodes the putative RNA-dependent RNA polymerase (RdRp); ORFs 2, 3 and 4 encode the triple gene block (TGB1-3), essential for virus movement; ORF5 encodes the coat protein (CP); and ORF6 encodes a nucleic acid-binding protein (NABP) containing a zinc finger motif [1, 31, 34]. Recently, a new virus, the sweet potato C6 virus, was described, with genomic organization typical of the genus *Carlavirus*, but without the ORF6 encoding the cysteine-rich protein. Instead a predicted protein was found with no similarity to any known protein [14].

CPMMV is transmitted by the whitefly *Bemisia tabaci* [4, 27, 36, 37] and was originally described infecting cowpea (*Vigna unguiculata*) [7]. In Brazil, although CPMMV was first reported in common bean (*Phaseolus vulgaris*) [13], only in the 2000/01 season was it reported infecting soybean plants, in Goiás state [3]. In addition to the diversity of symptoms caused in soybean, recent studies suggest a rapid spread of this virus in Brazilian soybean fields. Three years after its first identification in soybean, the virus was described infecting soybean fields in the states of Bahia (Barreiras), new localities of Goiás (Acreúna, Luzitania, Porteirão, Quirinópolis and Vianópolis), Mato Grosso (Sorriso), Maranhão (Balsas) and Paraná (Palotina) [2-4]. By 2008, the virus was present in several regions of the states of Paraná, Minas Gerais, Goiás, Mato Grosso, Bahia, Tocantins and Maranhão, infecting soybean plants [2]. The diversity of symptoms indicates the existence of variability among CPMMV isolates.

RNA viruses exhibit a high genetic variability [16, 22, 23, 25]. The evolutionary mechanisms of RNA virus evolution include mutation, recombination and genome reassortment, which act differently in each family of virus [16, 22, 23, 25]. The evolution of RNA viruses is driven by high rates of mutation [25], due to the error-prone replication attributed to the absence of proofreading activity in RNA-dependent RNA polymerases (RdRp) and a short generation time [18, 23, 25].

Recombination and genome reassortment can play an important role in the generation of genetic variability, increasing the potential for evolutionary changes [11]. Recombination most frequently takes place within a viral population in the same host cell, although it also occurs between different viral strains or different viruses [49]. It probably occurs when RdRp ‘jumps’ from the donor RNA templates to the acceptor template during the strand synthesis, remaining bound to the nascent RNA strand, so the hybrid molecule is produced [25, 46]. This process is known as copy-choice replication [25, 46].

So far, no study has been conducted to assess the genetic variability of CPMMV isolates. In fact, only one isolate has had its genome completely sequenced [34], and few isolates have been partially sequenced [4, 5, 24, 37, 51]. Although the virus has been known since 1973 [7], little is known about its variability in Brazil and globally. Considering that Brazil is the second largest producer of soybean in the world, that CPMMV is spread by the whitefly *B. tabaci* and that it has been found in several soybean fields, a greater understanding of the virus is needed. Therefore, this study was performed to assess the genetic variability of Brazilian CPMMV isolates infecting soybean.

Materials and methods

Sampling, detection and characterization of CPMMV isolates

A total of 65 samples of soybean plants showing symptoms of stem necrosis, dwarfism and bud blight were collected between 2009 and 2010 in soybean growing regions of the Brazilian states of Minas Gerais (35 samples collected in January 2009), Mato Grosso (3 samples collected in January 2009), Pará (22 samples collected in March 2009) and Goiás (5 samples collected in January 2010). Additionally, five symptomatic samples (collected between 2001 and 2002 in the Brazilian states of Bahia, Goiás, Maranhão, Mato Grosso and Pará) were provided by one of the authors (AMRA).

All collected samples were stored in a freezer at -80° C, in order to preserve the original samples.

The collected samples were tested by indirect ELISA [12] using a CPMMV polyclonal antiserum [9]. Positive samples were used as inoculum for mechanical inoculation of soybean plants cv. CD206 using 0.1M phosphate buffer, pH 7.2, with 0.1% sodium sulfite. Inoculated plants were maintained in a greenhouse, with average daily temperatures of 26 ± 2^{o}C, for 40 days post-inoculation (dpi) to check the infection and onset of symptoms. CPMMV infection of inoculated plants was again confirmed through indirect ELISA.

Additionally, the presence of begomoviruses and *Soybean mosaic virus* (SMV, genus *Potyvirus*) in mixed infection with CPMMV was checked for all samples. Infection by SMV was tested by indirect ELISA using a specific polyclonal antiserum produced by one of the authors (AMRA), and the presence of begomoviruses was evaluated by total DNA extraction [15] followed by PCR amplification using the universal oligonucleotides PBL1v2040/PCRc1 [44].

RT-PCR and molecular cloning

Leaf tissue of systemically CPMMV-infected soybean plants (positive in biological and serological tests) were submitted to total RNA extraction using the RNeasy Plant Mini Kit (Qiagen), according to the manufacturer’s instructions. RT-PCR was performed from 500 ng of total RNA, using Superscript III reverse transcriptase (Invitrogen) and Platinum *Taq* DNA polymerase (Invitrogen), according to the manufacturer’s protocol. The CPMMV primers (40 μM) used in RT-PCR are described in Supplementary Table S1 and amplifying the portion of the genome that includes from ORF2 to the 3 'terminal portion (ORF2-3'end). All amplifications consisted of 35 cycles of the following profile: 94ºC for 1 min, annealing at 45ºC, elongation at 72ºC for 1–2 min (depending on the size of amplicon) and a final extension step at 72ºC for 10 min. The amplicons were gel-purified using the Ilustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare), ligated into the pGEM-T Easy Vector (Promega), transformed into *Escherichia coli* DH5α cells and sequenced by Macrogen Inc (Seoul, South Korea).

Pairwise comparisons and recombination analysis

Partial sequences of CPMMV isolates, including ORF2 to 3’-end (fragment of 2514-2519 nt), were assembled using DNA BASER Sequence Assembler v.3.5 (Heracle Biosoft). Pairwise nt comparisons were performed by p-distance using MEGA v. 5 [50], and amino acid sequence comparisons were performed in using the quick alignment option and default settings in DNAMAN v. 7.0 (Lynnon Biosoft).

Multiple sequence alignments were performed using the Muscle module [19] in MEGA v. 5. Detection of potential recombinant sequences was performed using Recombination Detection Program (RDP) v.3.44 [32]. Default settings and a multiple comparison-corrected P-value cutoff of 0.05 were used throughout. Only those recombination events detected by three or more methods cited above were considered. The recombination analysis only involved isolates of this study.

Phylogenetic analysis

Phylogenetic trees were constructed in two ways: (i) for the complete amplified fragment (ORF2-3'-end); or (ii) for each of the five ORFs individually (ORF2 to ORF6). Phylogenetic relationships were inferred using Bayesian inference (BI) in MrBayes v. 3.0 (Ronquist & Huelsenbeck, 2003), with the evolution models selected by MrModeltest v. 2.2 [39] using the Akaike Information Criterion (AIC). The MCMC simulation was run for 20 million generations and sampled once in every 1000 generations. Burn-in was set at 4 million generations resulting in 16000 saved trees. The visualization of the trees was performed on the FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

Description of the CPMMV molecular variability

Descriptors of molecular variability were estimated using the DnaSP software v.5.10 [45]. The following descriptors estimated: (i) total number of segregating sites (*s*); (ii) average number of nucleotide differences between sequences (*k*); (iii) nucleotide diversity (π); (iv) number of haplotypes (h); (v) haplotype diversity (Hd) and (vi) Watterson's estimate of the population mutation rate based on the total number of segregating sites (θ-w). The π statistic was also calculated using a sliding window of 100 bases, with a step size of 10 bases for the purpose of estimate the nucleotide diversity throughout the length of ORF2 to ORF6.

Site-specific selection analysis

The gene- and site-specific selection pressures were measured for each of five ORFs evaluated (ORF2-6). The detection of sites under negative and positive selection in the genes was determined using four different maximum-likelihood-based algorithms, Single Likelihood Ancestor Counting (SLAC), Fixed Effects Likelihood (FEL), Random Effects Likelihood (REL) and Partitioning for Robust Inference of Selection (PARRIS) within the HyPhy software package (http://www.hyphy.org/) implemented in the Datamonkey server (www.datamonkey.org) with default conditions. The SLAC algorithm was also used to estimate the mean non-synonymous to synonymous substitutions ratio (dN/dS). The nucleotide substitution model incorporated was the Hasegawa-Kishino-Yano (HKY), only ORF6 incorporated the General Reversible substitution (REV) model in the analysis. Phylogenetic trees corrected for recombination were inferred by GARD (available at the Datamonkey server) and used as input for the selection analysis.

Results

Assessment of symptoms and mixed infections

We evaluated a total of seventy soybean samples collected in different Brazilian states and different years (2001-2010). CPMMV infection was confirmed in thirty of these plants by serological test (data not shown). Of the thirty positive samples, only eighteen induced symptoms in soybean cv. CD206 inoculated in greenhouse and were used in study. Six of them have been characterized previously (Table 1) [52]. Soybean samples did not show mixed infection with begomoviruses and SMV (data not shown).

The symptoms caused by eighteen CPMMV isolates in soybean plants cv. CD206 were highly variable in the greenhouse. The first symptoms were observed from 14 to 28 days post inoculation (dpi). Eight CPMMV isolates caused severe symptoms (bud blight, dwarfism, leaves and stem necrosis), and ten CPMMV isolates caused mild symptoms (crinkled/blistering leaves, mosaic and vein clearing) (Table 1 and Figure 1). Mosaic symptoms accompanied by vein clearing showed variable intensity (Figure 1A-C).

Sequence comparisons and recombination analysis

Viral RNA sequences were obtained from eighteen symptomatic soybean cv. CD206 samples. Nucleotide (nt) and amino acid (aa) sequence comparisons were performed for the ORF5 (coat protein - CP) amongst all seventeen viral isolates and other representative carlaviruses (data not shown), confirming the taxonomic classification of these isolates as a single viral species (CPMMV). All identity values were higher than 72% for nt and 80% for aa. The eighteen CPMMV isolates showed high nt and aa sequence identity in the different ORFs analyzed (ORF2 to ORF6) (Supple. Figure S1). Based on the pairwise comparisons of five ORFs, among all eighteen CPMMV, two distinct groups were evident: the first one composed of three isolates, CPMMV:BR:GO:01:1, CPMMV:BR:PA:02 and CPMMV:BR:GO:10:4 (group 1), that were very similar to each other, and the second composed of the other fifteen isolates (group 2) (Supple. Figure S1).

Sequence comparisons among the isolates of group 1 and isolates of group 2 showed that the identity values were smaller than the values obtained for each group

(Supple. Figure S1). ORF5 showed the highest sequence identity between the two groups of isolates CPMMV (Supple. Figure S1). The identity values among the isolates within each group were higher than 91% (nt and aa) for all ORFs evaluated (Supple. Figure S1), except for ORF3.

Additionally, we verify that the isolates sequenced in this study showed some variations in the number of aa for some predicted proteins based on nt sequences. ORF2 of CPMMV:BR:GO:01:1 and CPMMV:BR:PA:02 isolates supposedly encodes a protein of 234 aa, while for the other isolates a protein of 231 aa was predicted. ORF3 of CPMMV:BR:GO:10:4 isolate encodes a putative protein of 107 aa, for the other isolates 106 aa. Curiously, ORF6 of isolate CPMMV:BR:GO:10:4 showed a putative protein with 103aa, the CPMMV:BR:GO:01:1 and CPMMV:BR:PA:02 isolates supposedly encodes a protein of 132 aa and the other isolates a protein with 113 aa. These, associated with lower levels of identity with isolates found in group 2, suggest that the three isolates of group 1 are apparently distinct.

Partial genome sequences of CPMMV Brazilian isolates (encompassing five ORFs: 2, 3, 4, 5, 6 and the 3 '-end, were used in the recombination analysis. A single putative recombination event was identified among all eighteen partial CPMMV sequences evaluated. This event was identified in a small portion of ORF3 in the CPMMV:BR:MG:09:4 isolate (Breakpoints position: 787-847), and was supported by four analysis methods (*P*-values: RDP = 4.962 x 10^{-16}, Genecov = 6.088 x 10^{-08}, Max chi = 1.793 x 10^{-03}, Chimaera = 8.642 x 10^{-04}). The recombinant CPMMV:BR:MG:09:4 isolate had an unknown virus as a possible minor parent and CPMMV:BR:MG:09:05 isolate as putative major parent.

Phylogenetic analysis

We performed phylogenetic analysis to determine the relationships amongst seventeen CPMMV isolates (the recombinant CPMMV:BR:MG:09:4 isolate was excluded from this analysis). Phylogenetic relationships were reconstructed by BI based on different portions of the genome. This included a partial genome (ORF2 to 3’-end) and each one of the five ORFs individually (ORF 2, 3, 4, 5 and 6) (Figure 2).

The analyses of all resulting trees showed that two clades were formed [posterior probability (pp) of 1]. This was evidenced by a significant genetic distance separating the two groups (Figure 2). Clade 1 includes isolates CPMMV:BR:GO:01:1, CPMMV:BR:PA:02 and CPMMV:BR:GO:10:4, and clade 2 the other fourteen CPMMV isolates of this study (Figure 2).

Although the seventeen CPMMV isolates were collected in different years and regions, several thousands of kilometers apart within Brazil, they did not group in any significant way by these parameters in the phylogenetic trees. The phylogenetic tree analysis showed that isolates collected in Goiatuba and Cristalina, Goiás state, in 2001 and 2010, respectively, and the isolate collected in Paragominas (PA) in 2002 were clustered in clade 1 in all trees (Figure 2). Similarly, in clade 2 of all trees, the isolates collected in Barreiras (BA), Balsas (MA) and Sorriso (MT) in 2002 clustered with isolates collected in different regions of Minas Gerais state in 2009 and in Cristalina (GO) in 2010 (Figure 2, Table 2). We also observed that isolates collected in the same region (Cristalina, GO) clustered in different clades.

The clustering in term of the phylogenetic trees indicated a relation with the symptoms developed in soybean plants cv. CD206. Within clade 1, the two isolates that induce crinkle leaves and blistering (CPMMV:BR:GO:01:1 and CPMMV:BR:PA:02) clustered together in all trees, with pp greater than 0.75, except in the ORF6 tree (Figure 2). Within clade 2, the trees constructed using partial genomes (ORF2-3'-end), ORF5 and ORF6 showed the clustering of severe CPMMV isolates (inducing necrosis, bud blight and dwarfism) (Figure 2A, E and F), and it was well supported in the trees of partial genome and ORF6 (with pp equal to 0.92 and 0.99, respectively) (Figure 2A and 2F). For the ORF2 tree the grouping occurred among the mild CPMMV isolates that induce mosaic and vein clearing with pp equal to 0.79 (Figure 2B). The trees of ORFs 3 and 4 did not show any separation based on symptoms among the isolates from clade 2 (Figure 2C-D).

Additionally, we performed a second clustering analysis with the program STRUCTURE v.2.3.1 [43]. Although this program is used for inferring population structure using genotype data it implements a model-based clustering method. Confirming our results, the analysis showed that the clusters obtained by STRUCTURE v.2.3.1 recovered the same grouping of the phylogeny (data not shown).

Genetic variability of CPMMV isolates

To evaluate the molecular variability of CPMMV isolates we considered two data sets and analyzed the five ORFs individually (ORFs 2 to 6). The first dataset comprised seventeen CPMMV isolates and the second one fourteen, excluding the most distant CPMMV:BR:GO:01:1, CPMMV:BR:PA:02 and CPMMV:BR:GO:10:4 isolates. Additionally, the recombinant CPMMV:BR:MG:09:4 isolate was also excluded from both datasets.

The descriptors for the dataset including seventeen CPMMV isolates indicated higher genetic variability than the dataset with fourteen CPMMV isolates, represented by a higher number of segregating sites (*S*), nucleotide diversity (π), haplotype number (h) and haplotype diversity (Hd) (Table 2). This was verified for all CPMMV ORFs analyzed. The higher values obtained for all indexes were probably a consequence of the inclusion of the three divergent isolates (CPMMV:BR:GO:01:1, CPMMV:BR:PA:02 and CPMMV:BR:GO:10:4). The mutation rate (θ-W) estimated for all datasets was in the order of 10^{-2} (Table 2). In general the nucleotide diversity (π) was lower than 0.09 for all the ORFs of both datasets (Table 2). ORF5 was the region with lowest π value (0.05616 ± 0.01806), and ORF3 the region with highest π value (0.07802 ± 0.02789) in the dataset with seventeen isolates. However, in the dataset with fourteen isolates, ORF4 was the gene with the lowest π value (0.00621 ± 0.00248) and ORF6 the region with the highest (0.01806 ± 0.00287).

Additionally, we evaluated π values throughout the length of ORF2, ORF3, ORF4, ORF5 and ORF6 using the two datasets. We observed that the tendency for π values observed in the graphs is similar for all ORFs, independent of the dataset analyzed (Figure 3) although, as expected, higher π values were found along the lengths of the ORFs in the dataset with sixteen isolates (Figure 3). Within ORF2, the central region seems to be more variable than other regions (Figure 3). In ORF3 the 3’ region of the genes seems to be more variable than the 5’ region, the opposite of that observed for ORF6. In ORF4 there is little variability among CPMMV isolates analyzed, especially for the dataset with thirteen CPMMV isolates (Figure 3). ORF5 is variable along the sequence, and the portion that extends from the 5’ region to the nucleotide position 500 of gene seems to be the most diverse.

Analysis of site-specific selection

Initially the recombinant CPMMV:BR:MG:09:4 isolate was used in the analysis of selection, since the trees phylogenetic corrected for recombination were inferred by GARD and used as input in the analysis. However, the recombination event identified by RDP was not detected by GARD. In the presence of the recombinant isolate, some sites under positive selection were found in ORF3, located exactly in the recombinant portion (data not shown). Thus, we opted to eliminate the recombinant isolate of dataset before proceeding with the selection analysis.

We evaluated the effect of positive and negative selection at each site of ORFs 2 to 6 of the seventeen isolates. All ORFs evaluated showed dN/dS ratios (ϖ) lower than 1, indicating purifying selection (Table 3). The dN/dS ratios for the ORF5 (ϖ=0.1057) was the lowest of the viral ORFs analyzed (Table 3), showing this to be the most constricted region. The highest dN/dS ratios was for ORF6 (ϖ=0.2116).

Selection analyses showed that most of the sites are under negative selection (Table 3). In TGB using the SLAC method, five sites under negative selection were identified in ORF2 (positions 3, 150, 191, 201 and 210), three sites in ORF3 (positions 45, 84 and 96), and no site under negative or positive selection was found in ORF4 (Table 3). The SLAC method identified seven sites under negative selection in ORF5 (positions 68, 80, 97, 120, 140, 180 and 187) and two sites in ORF6 (positions 48 and 65). The FEL method only found sites under negative selection in all ORFs evaluated (Table 3). The REL method showed that all sites in ORFs 2 to 6 were under negative selection. The PARRIS method did not identify any site under positive selection in any of the ORFs analyzed (Table 3).

Discussion

The first identification of CPMMV was about 40 years ago [7] and it was first reported in Brazil 30 years ago [13]. However, our knowledge of the molecular variability of CPMMV and evolutionary aspects is limited. There are few genomic sequences of CPMMV available in public databases (thirteen nucleotide sequences in GenBank), and only one corresponds to a complete genome of an isolate infecting cowpea in Ghana. Additionally, six complete CPMMV sequences were obtained recently [52]. So far, no study has attempted to assess the molecular variability of CPMMV.

The symptom diversity observed in soybean fields drew our attention to CPMMV. Initially, we could not exclude the possibility of mixed infections with other viruses that infect soybean in Brazil, or having distinct species of carlaviruses infecting soybean plants. We did not detect begomoviruses or SMV in the analyzed samples, and comparisons of the nt and aa CP sequences showed that the isolates belong to the CPMMV species. Thus, the different symptoms that have been observed in the soybean fields are caused by different CPMMV isolates. In fact, we demonstrated that the eighteen CPMMV isolates analyzed in this study caused assorted symptoms in soybean cv. CD206 in greenhouse.

To verify molecular variability we analyzed eighteen partial sequences of CPMMV, spanning the ORFs 2, 3, 4, 5, 6 and the 3'- end, of isolates sampled from soybean plants in Brazil. Pairwise comparisons and phylogenetic analysis clearly showed the existence of two groups of isolates, with variations in the sequences of some ORFs. Within each group, CPMMV isolates causing varied symptoms were collected from different Brazilian states in different years, but none of these factors seems to define the clustering of phylogenetic trees, indicating that both groups are widespread and are established in the field. Thus, we can only affirm at the moment that we identified two strains of CPMMV infecting Brazilian soybean fields, denominated CPMMV-BR1 (isolates from group 1, in pairwise comparisons, and clade 1, in phylogenetic trees) and CPMMV-BR2 (isolates from group 2 and clade 2). These strains exhibit different biological characteristics based on soybean cv. CD206. In phylogenetic analysis of ORF 2, 5, 6 and a partial genome a second grouping in clades based on the symptoms is evident, thus asserting that variations among isolates of the same strain also exist.

Two hypotheses are proposed to explain the presence of the same viral strain in geographically distant locations. The first involves the high dissemination efficiency of CPMMV by *B. tabaci* and its flight capacity. It has been demonstrated that CPMMV transmission from soybean to soybean plants by *B. tabaci* occurs in a relatively short time at 15 minutes and a single insect vector is able to transmit CPMMV with transmission rate of 16.7% [33, 36]. Additionally, Byrne (1999) [8] affirmed that the largest long distance flight by *B. tabaci* was 7 km, passively in a stream of air. This type of flying facilitates migration of whitefly populations to distant sites and the consequent colonization of other crops and fields [20], allowing the virus to spread across regions. The insect vector can also be transported easily long distances along transport routes by vehicles.

Flying long distances also allows gene flow among whitefly populations [20]. Currently, there are reports of more than 24 biotypes in the world, and this variability suggests that *B. tabaci* is a complex of species or biotypes [6, 21, 41]. In Brazil, it has been demonstrated that intra- and interpopulation variability exists in *B. tabaci* B-biotype [20, 30]. Additionally, Lima et al. (2012) [30] demonstrated that population differentiation of whitefly occurred mainly according to the plant host, rather than geographical region. Thus, a species or a different biotype or even a different variant of whitefly can be found at a higher frequency in a given region. It may be that these vector biotypes are associated with different strains of the virus, which would strongly influence their dissemination.

The second possibility to explain why isolates geographically and temporally distant are so close molecularly and phylogenetically is the occurrence of seed transmission. However, the transmission by seeds is still controversial and depends on the viral isolate. The Ghanaian CPMMV isolate described by Brunt and Kenten (1973) [7] was transmitted by seeds in soybean, cowpea and with lower frequency in common bean and in Venezuela was demonstrated that CPMMV can be transmitted by yardlong bean seeds [5]. In contrast, Almeida et al. (2005) [4] showed that a Brazilian CPMMV isolate (CPMMV:BR:BA:02) was not transmitted by soybean seeds. If CPMMV transmission by seeds is confirmed for other Brazilian CPMMV isolates, this could explain the distribution of the virus in Brazil. It cannot be excluded that the two situations described above (transmission by whitefly and seeds) may be occurring simultaneously.

Recombination can be an important factor in viral evolution and can result in genetic exchange. Sequence analyses of various RNA and DNA plant viruses provide evidence that recombination may be a major source of variation [22, 23, 29, 35, 46, 49]. In this study we found a single recombinant isolate (CPMMV:BR:MG:09:4) with breakpoints located in ORF3 (TGB2). The recombinant isolate was eliminated from our analysis, because recombination can interfere with phylogeny [42]. In fact, we verified that besides phylogeny, recombination affects the assessment of variability and selection analysis. When the recombinant isolate was used, sites under positive selection were found in recombinant region of ORF3, in the absence of recombinant isolate no site under positive selection was found in this ORF.

Recombination events have previously been reported in different carlavirus species. Analysis involving six isolates of *Chrysanthemum virus B* (CVB) showed sixteen recombination events, thirteen of them involving the RdRp and the other three events involving TGB, CP and NABP [48]. In a study of isolates of *Lily symptomless virus* (LSV) two recombination events were confirmed, and in both events the RdRp had recombinant portions [47]. For *Potato virus S* (PVS) the recombination analyses showed a portion of ORF2 of an isolate as recombinant [17]. In study involving six CPMMV Brazilian isolates, five recombination events were detected, four of these located in the RdRp (ORF1) and one involving others regions of the genome (ORF2-5) [52]. These facts suggest that recombination events in the portion that extends from ORF2 to 3' end are not found frequently.

Analysis of variability descriptors performed for the two datasets showed different results. The dataset with seventeen CPMMV isolates was more variable than the one with fourteen CPMMV isolates. This probably occurred because isolates belonging to different strains were analyzed together. The values of nucleotide diversity (π) were less than 0.09 for all regions of the genome evaluated in the two datasets, in concordance with some π values found by García-Arenal et al. (2001) [22]: the CP of dataset with seventeen isolates showed had a nucleotide diversity (0.05616) similar to the CP of a global population of *Citrus tristeza virus* (CTV, *Closterovirus* genus) (π=0.068). For the dataset with fourteen CPMMV isolates the nucleotide diversity of the CP (π=0.01101) showed a value close to that found for the CP of African populations of *Groundnut rosette assistor virus* (GRAV, *Luteoviridae* family) (π=0.018). Pagan et al. [40] showed for *Pepino mosaic virus* (PepMV, genus *potexvirus*) a π value for TGB (0.0075) similar to the ones we found for the genes that comprising the CPMMV TGB (TGB1: 0.00689; TGB2: 0.00876; and TGB3: 0.00621) of the dataset with fourteen CPMMV isolates. ORF6 could not be compared due to lack of data. Our dataset was also obtained from different geographical regions and with a smaller number of isolates than the studies cited above, but the values were very close to those found by different authors.

The analysis of selection showed that the different regions of the CPMMV genome are under negative selection. Sequence analysis of other viruses shows that in most instances selection is negative, purifying [22]. Selection is associated with every factor in the life cycle of the virus, including the maintenance of structural features and functional activity of proteins [22]. ORF5 (CP) was the region under strongest negative selection of the viral ORFs analyzed, showing it to be the most constricted region. This situation is typical for arthropod-vectored viruses [10].

The results presented here advance our understanding of the molecular variability of CPMMV Brazilian isolates infecting soybean. We have shown that CPMMV isolates collected from different regions of Brazil in different years and causing different symptoms belong to two different strains of CPMMV. This is the first study of the molecular variability of CPMMV, genus *Carlavirus*. Future studies are necessary to clarify the mode of transmission and spread of Brazilian CPMMV isolates.

Such additional information can help in the development and adoption of preventive control measures.

Acknowledgements

This work was funded by FAPEMIG (APQ-0992/09), CNPq (474112/2008-0) and Funarbe grants to CMC. LGZ was supported by a scholarship from CNPq.

References

1. Adams MJ, Candresse T, Hammond J, Kreuze JF, Martelli GP, Namba S, Pearson MN, Ryu KH, Saldarelli P, Yoshikawa N (2012) Family *Betaflexiviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, pp 920-941

2. Almeida AMR (2008) Viroses da soja no Brasil: sintomas, etiologia, controle. Série Documentos 306:1-62

3. Almeida AMR, Piuga FF, Kitajima EW, Gaspar JO, Valentin N, Benato LC, Marin SRR, Bineck E, Belintani P, Nunes Junior J, Hoffmann L, Meyer MC (2003) Necrose da haste da soja. Série Documentos 221:1-48

4. Almeida AMR, Piuga FF, Marin SRR, Kitajima EW, Gaspar JO, Oliveira TGd, Moraes TGd (2005) Detection and partial characterization of a carlavirus causing stem necrosis of soybean in Brazil. Fitopatol Bras 30:191-194

5. Brito M, Fernandez-Rodriguez T, Garrido MJ, Mejias A, Romano M, Marys E (2012) First report of *Cowpea mild mottle Carlavirus* on yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis*) in Venezuela. Viruses 4:3804-3811

6. Brown JK, Frohlich DR, Rosell RC (1995) The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? Annu Rev Entomol 40:511-534

7. Brunt AA, Kenten RH (1973) *Cowpea mild mottle*, a newly recognized virus infecting cowpeas (*Vigna unguiculata*) in Ghana. Ann Appl Biol 74:67-74

8. Byrne DN (1999) Migration and dispersal by the sweet potato whitefly, *Bemisia tabaci*. Agric For Meteorol 97:309-316

9. Carvalho SL, Silva FN, Zanardo LG, Almeida AMR, Zerbini FM, Carvalho CM (2013) Production of polyclonal antiserum against *Cowpea mild mottle virus* coat protein and its application in virus detection. Trop Plant Pathol 38:49-54

10. Chare ER, Holmes EC (2004) Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission. J Gen Virol 85:3149-3157

11. Chare ER, Holmes EC (2006) A phylogenetic survey of recombination frequency in plant RNA viruses. Arch Virol 151:933-946

12. Clark MF, Lister RM, Bar-Joseph M, Arthur Weissbach HW (1986) ELISA techniques. In: Methods in Enzymology. Academic Press, pp 742-766

13. Costa AS, Gaspar JO, Vega J (1983) Mosaico angular do feijão jalo causado por um carlavírus transmitido pela mosca branca *Bemisia tabaci*. Fitopatol Bras 8:325-327

14. De Souza J, Fuentes S, Savenkov EI, Cuellar W, Kreuze JF (2013) The complete nucleotide sequence of sweet potato C6 virus: a carlavirus lacking a cysteine-rich protein. Arch Virol. doi: 10.1007/s00705-013-1614-x

15. Dellaporta SL, Woud J, Hicks JB (1983) A plant DNA minipreparation: Version II. Plant Mol Biol Report 1:19-21

16. Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. Annu Rev Microbiol 51:151-178

17. Duarte PSG, Galvino-Costa SB, Ribeiro SRP, Figueira AR (2012) Complete genome sequence of the first Andean strain of *Potato virus S* from Brazil and evidence of recombination between PVS strains. Arch Virol 157:1357-1364

18. Duffy S, Shackelton LA, Holmes EC (2008) Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9:267-276

19. Edgar R (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BioMed Cent Bioinforma 5:113

20. Fontes FVHM, Colombo CA, Lourenção AL (2010) Caracterização molecular e divergência genética de *Bemisia tabaci* (Genn.) (Hemiptera: *Aleyrodidae*) em diferentes culturas e locais de cultivo. Neotrop Entomol 39:221-226

21. Frohlich DR, Torres-Jerez I, Bedford ID, Markham PG, Brown JK (1999) A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. Mol Ecol 8:1683-1691

22. Garcia-Arenal F, Fraile A, Malpica JM (2001) Variability and genetic structure of plant virus populations. Annu Rev Phytopathol 39:157-186

23. Garcia-Arenal F, Fraile A, Malpica JM (2003) Variation and evolution of plant virus populations. Int Microbiol 6:225-232

24. Gaspar JO, Belintani P, Almeida AM, Kitajima EW (2008) A degenerate primer allows amplification of part of the 3'-terminus of three distinct carlavirus species. J Virol Methods 148:283-285

25. Holmes EC (2009) The evolutionary genetics of emerging viruses. Annu Rev Ecol Evol Sist 40:353–372

26. Iwaki M, Thongmeearkon P, Prommin M, Honda Y, Hibi J (1982) Whitefly transmission and some properties of *Cowpea mild mottle virus* on soybean in Thailand. Plant Dis 66:265-268

27. Jeyanandarajah P, Brunt AA (1993) The natural occurrence, transmission, properties and possible affinities of *Cowpea mild mottle virus*. J Phytopathol 137:148-156

28. Laguna IG, Arneodo JD, Rodríguez-Pardina P, Fiorona M (2006) *Cowpea mild mottle virus* infecting soybean crops in northwestern Argentina. Fitopatol Bras 31:317-317

29. Lima AT, Sobrinho RR, Gonzalez-Aguilera J, Rocha CS, Silva SJ, Xavier CA, Silva FN, Duffy S, Zerbini FM (2013) Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. J Gen Virol 94:418-431

30. Lima LHC, Campos L, Moretzsohn MC, Návia D, Oliveira MRV (2002) Genetic diversity of *Bemisia tabaci* (Genn.) populations in Brazil revealed by RAPD markers. Genet Mol Biol 25:217-223

31. Martelli GP, Adams MJ, Kreuze JF, Dolja VV (2007) Family *Flexiviridae*: a case study in virion and genome plasticity. Annu Rev Phytopathol 45:73-100

32. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462-2463

33. Marubayashi JM, Yuki VA, Wutke EB (2010) Transmissão do *Cowpea mild mottle virus* pela mosca branca *Bemisia tabaci* biótipo B para plantas de feijão e soja. Summa Phytopathol 36:158-160

34. Menzel W, Winter S, Vetten H (2010) Complete nucleotide sequence of the type isolate of *Cowpea mild mottle virus* from Ghana. Arch Virol 155:2069-2073

35. Milgroom MG, Peever TL (2003) Populations biology of plant pathogens. Plant Dis 87:608-617

36. Munyappa V, Reddy DVR (1983) Transmission of *Cowpea mild mottle virus* by *Bemisia tabaci* in a nonpersistent manner. Plant Dis 67:391-393

37. Naidu RA, Gowda S, Satyanarayana T, Boyko V, Reddy AS, Dawson WO, Reddy DV (1998) Evidence that whitefly-transmitted *Cowpea mild mottle virus* belongs to the genus *Carlavirus*. Arch Virol 143:769-780

38. Nicolaisen M, Nielsen SL (2001) Analysis of the triple gene block and coat protein sequences of two strains of *Kalanchoe latent carlavirus*. Virus Genes 22:265-270

39. Nylander JAA (2004) MrModeltest v2. Program distributed by the author.

40. Pagan I, Del Carmen Cordoba-Selles M, Martinez-Priego L, Fraile A, Malpica JM, Jorda C, Garcia-Arenal F (2006) Genetic structure of the population of *Pepino mosaic virus* infecting tomato crops in Spain. Phytopathology 96:274-279

41. Perring TM (2001) The *Bemisia tabaci* species complex. Crop Prot 20:725-737

42. Posada D, Crandall KA (2002) The effect of recombination on the accuracy of phylogeny estimation. J Mol Evol 54:396-402

43. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

44. Rojas MR, Gilbertson RL, Russel DR, Maxwell DP (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis 77:340-347

45. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497

46. Simon-Loriere E, Holmes EC (2011) Why do RNA viruses recombine? Nat Rev Microbiol 9:617-626

47. Singh A, Mahinghara B, Hallan V, Ram R, Zaidi A (2008) Recombination and phylogeographical analysis of *Lily symptomless virus*. Virus Genes 36:421-427

48. Singh L, Hallan V, Martin D, Ram R, Zaidi A (2012) Genomic sequence analysis of four new *Chrysanthemum virus B* isolates: evidence of RNA recombination. Arch Virol 157:531-537

49. Sztuba-Solinska J, Urbanowicz A, Figlerowicz M, Bujarski JJ (2011) RNA-RNA recombination in plant virus replication and evolution. Annu Rev Phytophathol 49:415-443

50. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-2739.

51. Tavasoli M, Shahraeen N, Ghorbani S (2009) Serological and RT-PCR detection of *Cowpea mild mottle Carlavirus* infecting soybean. J Gen Mol Virol 1:7-11

52. Zanardo LG, Silva FN, Bicallho AAC, Castillo-Urquiza GP, Lima ATM, Almeida AMR, Zerbini FM, Carvalho CM (2013) Molecular and biological characterization of *Cowpea mild mottle virus* isolates infecting soybean in Brazil and evidence of recombination. Plant Pathol. *In press.*

Figure legends:

Fig. 1 Symptoms observed in soybean cv. CD206 inoculated with different CPMMV isolates. **A.** Weak mosaic and vein clearing in soybean plants inoculated with CPMMV:BR:MG:09:5 isolate; **B.** Intermediate mosaic and vein clearing in soybean plants inoculated with CPMMV:BR:MG:09:11; **C.** Strong mosaic, vein clearing and crinkled leaves in soybean plants inoculated with CPMMV:BR:BA:02; **D.** Stem necrosis and bud blight in soybean plants inoculated with CPMMV:BR:MG:09:6; **E.** Stem and leaf necrosis in soybean plants inoculated with CPMMV:BR:MG:09:5 and **F.** Blistering and crinkled leaves in soybean plants inoculated with CPMMV:BR:GO:01:1. Similar symptoms have been observed for other isolates from the study.

Fig. 2 Phylogenetic relationships, based on the sequences of partial genome (ORFs2-3'end) and individual ORFs, of the seventeen Brazilian CPMMV isolates using Bayesian inference (implemented in MrBayes 3.1, with 20 million generations). Bayesian posterior probability values are given between nodes. **A.** Partial genome (ORFs 2-3'end) with selection of model GTR+G; **B.** ORF 2, triple gene block (TGB1) with selection of model GTR+G; **C-D.** ORF3-4 (TGB2-3) with selection of model HKY+G; **E.** ORF5, coat protein (CP) with selection model HKY+I; **F.** ORF6, nucleic acid binding protein (NABP) with selection model GTR+G**.** The blue bars show the grouping of isolates causing severe or mild symptoms in soybean plants in clade 2.

Fig. 3 Average pairwise number of nucleotide differences per site (nucleotide diversity, π) calculated on a sliding window across the ORF2, 3, 4, 5 and 6. Sequences from data set with seventeen CPMMV isolates (blue line), and data set with fourteen CPMMV isolates (red line). **A.** ORF 2, triple gene block (TGB 1); **B.** ORF 3 (TGB 2); **C.** ORF4 (TGB 3); **D.**ORF5, coat protein (CP); **E.** ORF6, nucleic acid binding protein (NABP).

Supplementary Figure S1 Two-dimensional plot representing the percent sequence identities between the seventeen Brazilian CPMMV isolates for five ORFs [ORF 2, triple block genes (TGB 1); ORF 3 (TGB 2); ORF 4 (TGB 3); ORF5, coat protein (CP) and ORF 6, nucleic acid binding protein (NABP)] evaluated. Nucleotide sequence identities are above the diagonal and amino acid sequence identities below.

Table 1 Isolates of CPMMV obtained from samples of soybean collected in different Brazilian states, and the symptoms induced in soybean plants cv. CD206 inoculated in greenhouse.

Isolates of CPMMV	Location	Coordinates	Year	Symptoms in Soybean CD206[a]	Genbank accession #
CPMMV:BR:BA:02[b]	Barreiras, BA	12°8’54” S, 44°59’33” W	2002	Cl, Ms, Vc	KC884247
CPMMV:BR:GO:01:1[b]	Goiatuba, GO	18°0’40” S, 49°22’10” W	2001	Cl, B	KC884248
CPMMV:BR:GO:10:4	Cristalina, GO	16°46’4” S, 47°36’47” W	2010	Bb, D, Ln, Sn	-
CPMMV:BR:GO:10:5[b]	Cristalina, GO	16°46’4” S, 47°36’47” W	2010	Bb, D, Ln, Sn	KC884249
CPMMV:BR:MA:02	Balsas, MA	07º31’57” S, 46º02’08” W	2002	Cl, Mw, Vc	-
CPMMV:BR:MG:09:1	Tupaciguara, MG	18°36’12” S, 48°41’25” W	2009	Bb, D, Ln, Sn	-
CPMMV:BR:MG:09:2[b]	Capinópolis, MG	18°40’48” S, 49°33’58” W	2009	Bb, D, Ln, Sn	KC884244
CPMMV:BR:MG:09:3[b]	Tupaciguara, MG	18°36’12” S, 48°41’25” W	2009	Bb, D, Ln, Sn	KC884245
CPMMV:BR:MG:09:4	Tupaciguara, MG	18°36’12” S, 48°41’25” W	2009	Mw, Vc	-
CPMMV:BR:MG:09:5	Tupaciguara, MG	18°36’12” S, 48°41’25” W	2009	Bb, D, Ln, Sn	-
CPMMV:BR:MG:09:6	Tupaciguara, MG	18°36’12” S, 48°41’25” W	2009	Bb, D, Ln, Sn	-
CPMMV:BR:MG:09:7	Capinópolis, MG	18°40’48” S, 49°33’58” W	2009	Cl, Mi, Vc	-
CPMMV:BR:MG:09:11	Pontal do Triângulo, MG	19°41’19”S, 50°41’45” W-	2009	Mi, Vc	-
CPMMV:BR:MG:09:12	Pontal do Triângulo, MG	19º31’47”S, 45º57’59” W	2009	Bb, D, Ln, Sn	-
CPMMV:BR:MG:09:15	Uberaba, MG	19º44’54” S, 47º55’55” W	2009	Ms, Vc	-
CPMMV:BR:MG:09:16	Uberaba, MG	19º44’54” S, 47º55’55” W	2009	Mi, Vc	-
CPMMV:BR:MT:02:1[b]	Sorriso, MT	12°33’31” S, 55°42’51” W	2002	Mw, Vc	KC884246
CPMMV:BR:PA:02	Paragominas, PA	02°59’51” S, 47°21’13” W	2002	Cl, B	-

[a] B=Blistering, Bb= Bud Blight, Cl=Crinkled leaves, D=Dwarfism, Ln=Leaf necrosis, Mw=Weak mosaic, Mi=Intermediate mosaic, Ms=Strong mosaic, Nll=Necrotic local lesions, Sn=Stem necrosis; Vc=Vein clearing. [b] Previously sequenced isolates [52]

Table 2 Descriptors of variability for *Cowpea mild mottle virus* (CPMMV) populations obtained from soybean plans in different states of Brazil.

Genome region	Number of isolates*	Region lengh (nt)	S^a	K^b	π^c	H^d	Hd^e	θ-W^f
ORF2 (TGB1)	17	697	190	52.978	0.07623 ± 0.02778	13	0.949	0.08086
	14		29	4.791	0.00689 ± 0.00145	11	0.934	0.01312
ORF3 (TGB2)	17	321	88	25.044	0.07802 ± 0.02789	14	0.971	0.08109
	14		18	2.813	0.00876 ± 0.00168	12	0.967	0.01763
ORF4 (TGB3)	17	207	55	15.787	0.07626 ± 0.02817	7	0.662	0.07859
	14		9	1.286	0.00621 ± 0.00248	5	0.505	0.01367
ORF5 (CP)	17	867	184	48.632	0.05616 ± 0.01806	17	1.0	0.06285
	14		55	9.549	0.01101 ± 0.00116	14	1.0	0.01995
ORF6 (NABP)	17	342	81	21.949	0.07035 ± 0.02104	13	0.949	0.07679
	14		26	6.176	0.01806 ± 0.00287	10	0.923	0.02391

[a] Total number of segregating sites.
[b] Average number of nucleotide differences between sequences.
[c] Nucleotide diversity.
[d] Haplotype number.
[e] Haplotype diversity.
[f] Watterson’s estimate of the population mutation rate based on the total number of segregating sites.
* The CPMMV:BR:MG:09:4 isolate recombinant was excluded from the analysis.

Table 3 Analysis of selection for ORF2, ORF3, ORF4, ORF5 and ORF6 of CPMMV Brazilian isolates implemented in Datamonkey server.

Genome region	Number of Isolates	dN/dS	SLAC[a]		FEL[b]		REL[c]		PARRIS[d]
			PS	NS	PS	NS	PS	NS	PS
ORF2	17	0.1552	-	5	-	59	-	+	-
ORF3	17	0.1668	-	3	-	27	-	+	-
ORF4	17	0.1905	-	-	-	11	-	+	-
ORF5	17	0.1057	-	7	-	89	-	+	-
ORF6	17	0.2116	-	2	-	27	-	+	-

(PS) Sites under positive selection; (NS) sites under negative selection; (-) no site under selection; (+) all sites under selection. [a,b,c,d]Codon-based maximum-likelihood algorithms. [a]Single Likelihood Ancestor Counting (SLAC); [b]Fixed Effects Likelihood; (FEL); [c]Random Effects Likelihood (REL) and [d]Robust Inference of Selection (PARRIS).

Figure 1

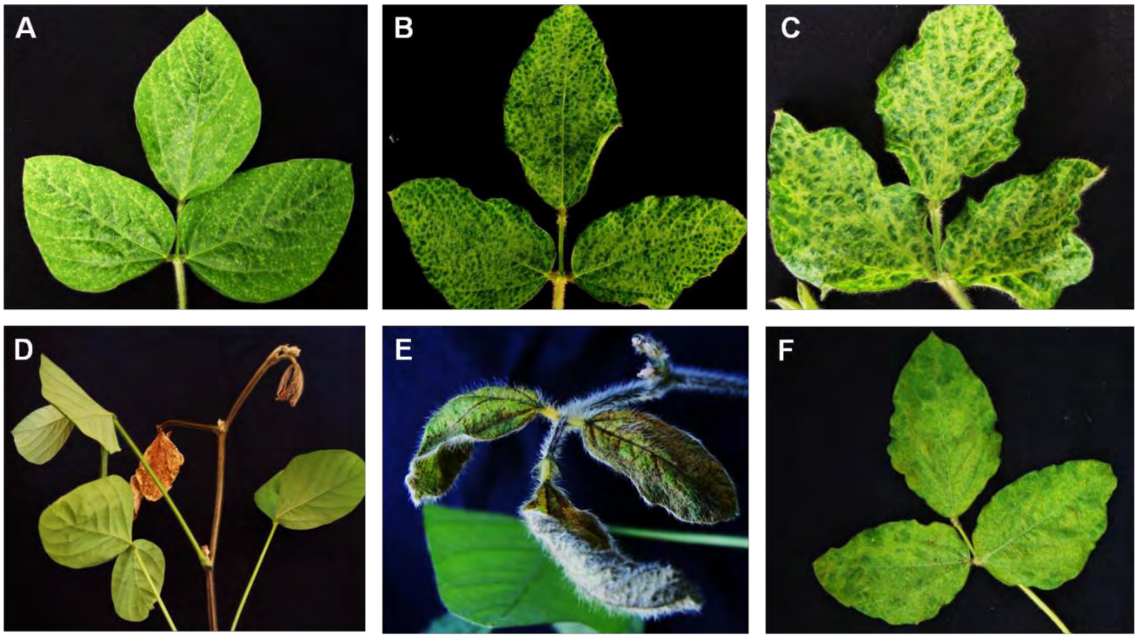


Figure 2

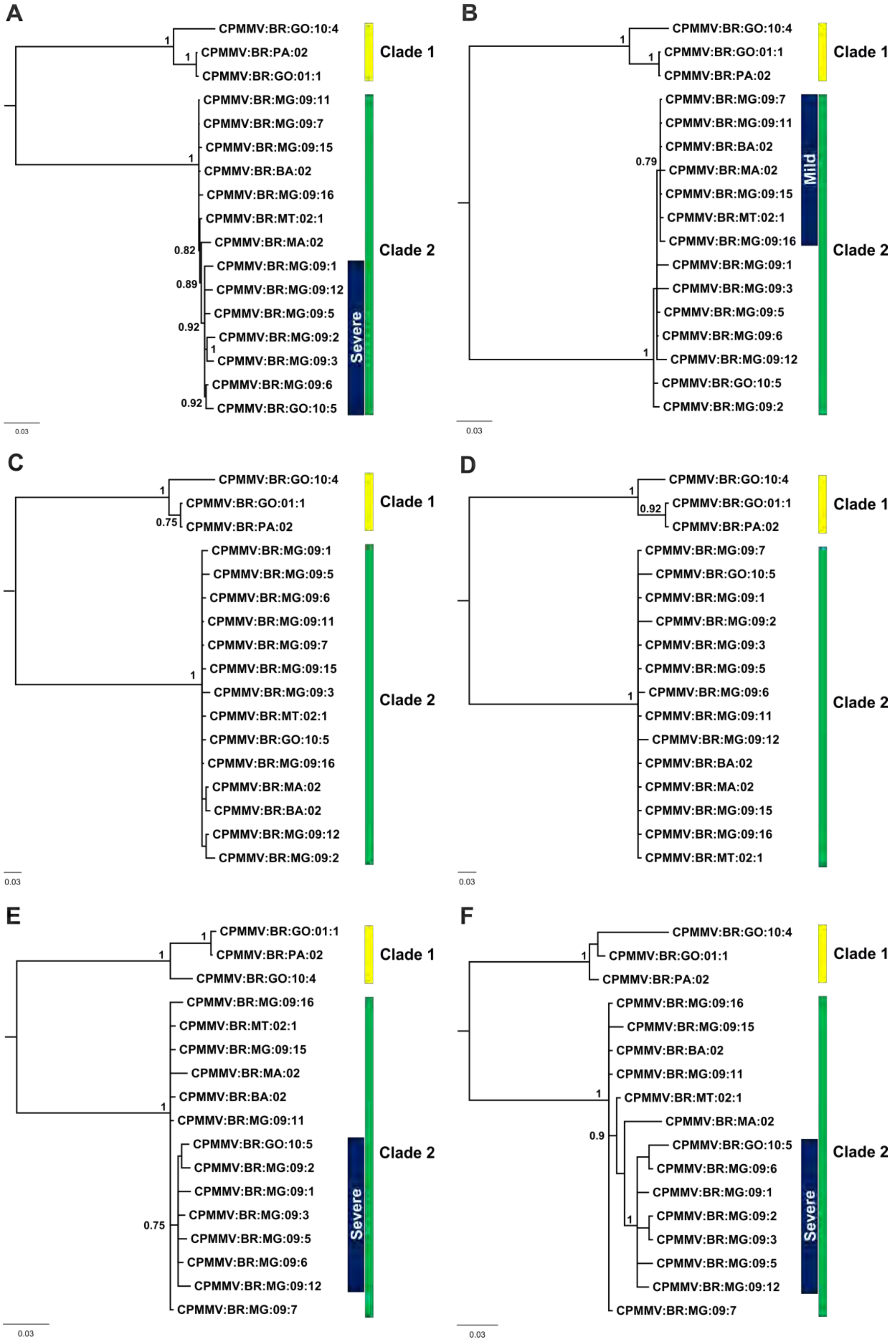
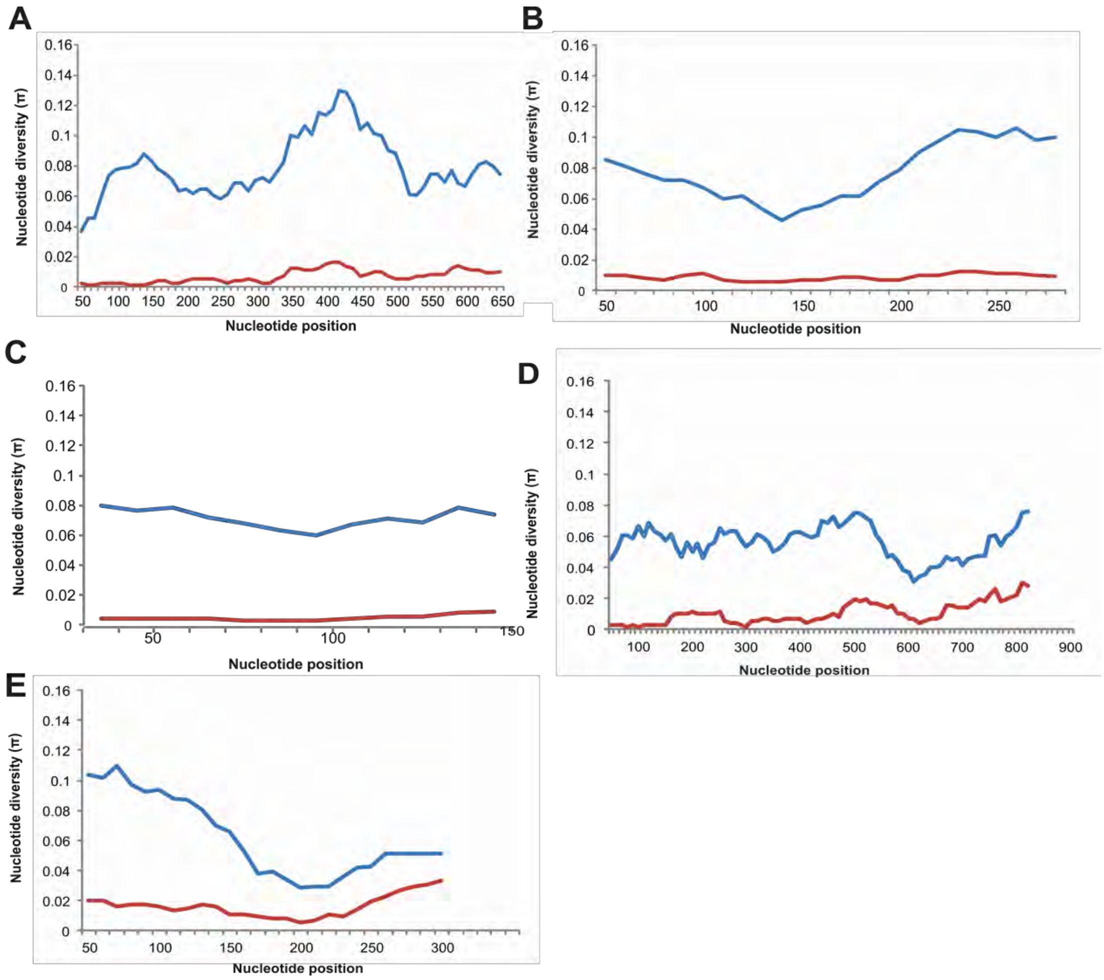

A
CPMMV:BR:GO:10:4
CPMMV:BR:PA:02
CPMMV:BR:GO:01:1
CPMMV:BR:MG:09:11
CPMMV:BR:MG:09:7
CPMMV:BR:MG:09:15
CPMMV:BR:BA:02
CPMMV:BR:MG:09:16
CPMMV:BR:MT:02:1
CPMMV:BR:MA:02
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:12
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:3
CPMMV:BR:MG:09:6
CPMMV:BR:GO:10:5
1
1
1
0.82
0.89
0.92
1
0.92
Clade 1
Clade 2
Severe
0.03
B
CPMMV:BR:GO:10:4
CPMMV:BR:GO:01:1
CPMMV:BR:PA:02
CPMMV:BR:MG:09:7
CPMMV:BR:MG:09:11
CPMMV:BR:BA:02
CPMMV:BR:MA:02
CPMMV:BR:MG:09:15
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:16
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:3
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:6
CPMMV:BR:MG:09:12
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:2
1
1
0.79
1
Clade 1
Mild
Clade 2
0.03
C
CPMMV:BR:GO:10:4
CPMMV:BR:GO:01:1
CPMMV:BR:PA:02
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:6
CPMMV:BR:MG:09:11
CPMMV:BR:MG:09:7
CPMMV:BR:MG:09:15
CPMMV:BR:MG:09:3
CPMMV:BR:MT:02:1
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:16
CPMMV:BR:MA:02
CPMMV:BR:BA:02
CPMMV:BR:MG:09:12
CPMMV:BR:MG:09:2
1
0.75
1
Clade 1
Clade 2
0.03
D
CPMMV:BR:GO:10:4
CPMMV:BR:GO:01:1
CPMMV:BR:PA:02
CPMMV:BR:MG:09:7
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:3
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:6
CPMMV:BR:MG:09:11
CPMMV:BR:MG:09:12
CPMMV:BR:BA:02
CPMMV:BR:MA:02
CPMMV:BR:MG:09:15
CPMMV:BR:MG:09:16
CPMMV:BR:MT:02:1
1
0.92
1
Clade 1
Clade 2
0.03
E
CPMMV:BR:GO:01:1
CPMMV:BR:PA:02
CPMMV:BR:GO:10:4
CPMMV:BR:MG:09:16
CPMMV:BR:MT:02:1
CPMMV:BR:MG:09:15
CPMMV:BR:MA:02
CPMMV:BR:BA:02
CPMMV:BR:MG:09:11
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:3
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:6
CPMMV:BR:MG:09:12
CPMMV:BR:MG:09:7
1
1
1
0.75
Clade 1
Clade 2
Severe
0.03
F
CPMMV:BR:GO:10:4
CPMMV:BR:GO:01:1
CPMMV:BR:PA:02
CPMMV:BR:MG:09:16
CPMMV:BR:MG:09:15
CPMMV:BR:BA:02
CPMMV:BR:MG:09:11
CPMMV:BR:MT:02:1
CPMMV:BR:MA:02
CPMMV:BR:GO:10:5
CPMMV:BR:MG:09:6
CPMMV:BR:MG:09:1
CPMMV:BR:MG:09:2
CPMMV:BR:MG:09:3
CPMMV:BR:MG:09:5
CPMMV:BR:MG:09:12
CPMMV:BR:MG:09:7
1
1
0.9
1
Clade 1
Clade 2
Severe
0.03

Figure 3

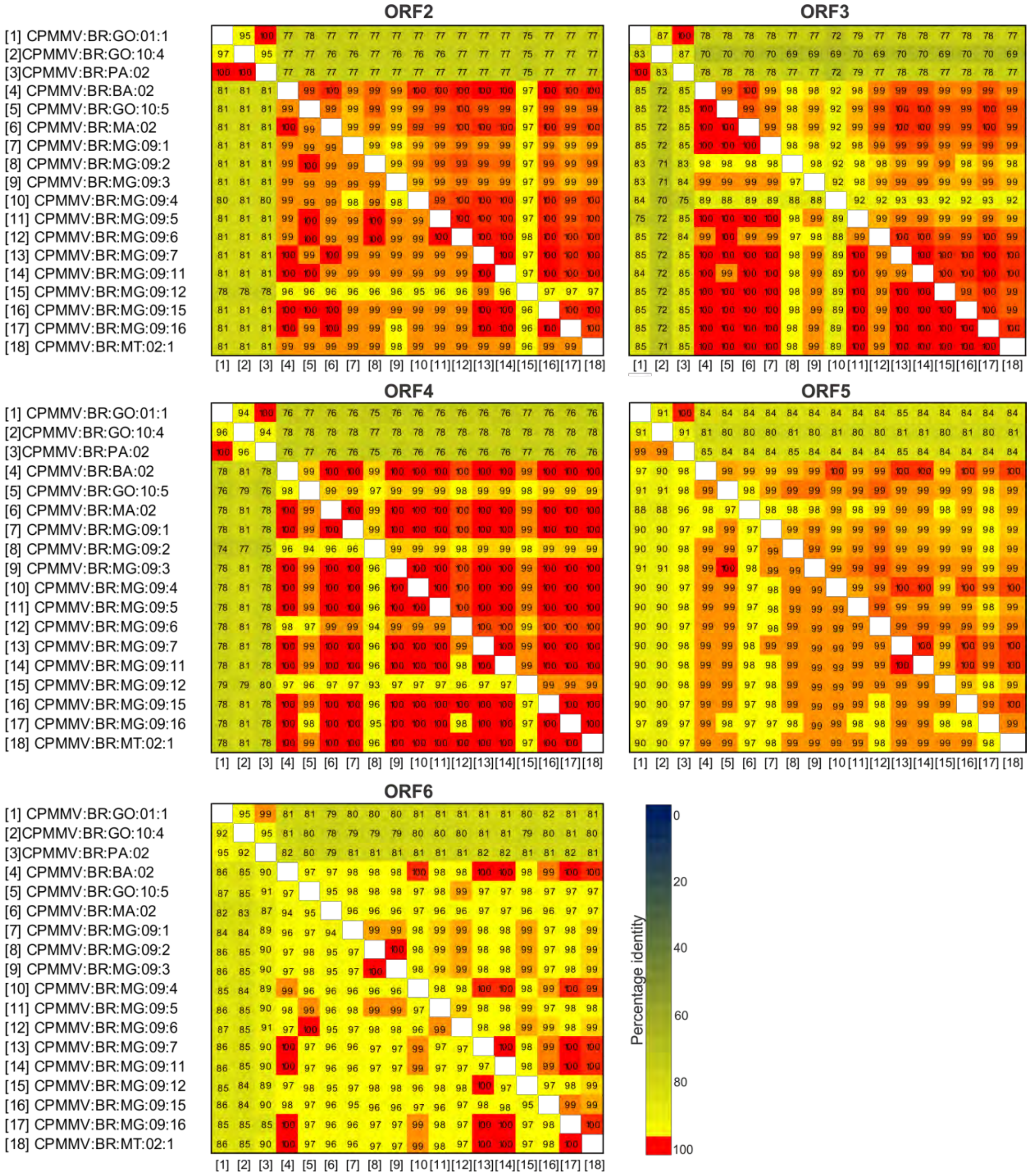

A
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Nucleotide diversity (π)
50 100 150 200 250 300 350 400 450 500 550 600 650
Nucleotide position
B
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Nucleotide diversity (π)
50 100 150 200 250
Nucleotide position
C
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Nucleotide diversity (π)
50 100 150
Nucleotide position
D
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Nucleotide diversity (π)
100 200 300 400 500 600 700 800 900
Nucleotide position
E
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Nucleotide diversity (π)
50 100 150 200 250 300
Nucleotide position

Supplementary Table S1: Primers used in RT-PCR and predicted amplicon size for viral detection and cloning.

Fragments	**Primer**	**Primer sequence 5'-3'**	**PCR product size**
ORF2	ORF2 F	TCCTTTAGGTAGTGAGGC	938
	ORF2 R	AAGTTCGTGCCAGTTGAC^{cD}	
ORF3	ORF3 F	CTTNATYTGCYTNACNAGGCA	552
	ORF3 R	TGTTCTCTNACCAAGT^{cD}	
ORF4 - 3' end	ORF 4 F	TAYMRDGAYGGNACHAA*	1676
	ORF6 R	TAAAACCAGGAAATAAC^{cD}	

*Described by Nicolaisen and Nielsen [38]. The other oligonucleotides were described by Zanardo et al. (2013) [52]. ^{cD} Used for cDNA Synthesis.

Supplementary Figure S1



CONCLUSÕES GERAIS

- Foram sequenciados completamente 6 isolados de CPMMV e outros 12 isolados parcialmente.
- Foi a primeira vez que a sequencia completa de isolados brasileiros de *Cowpea mild mottle virus* (CPMMV) oriundos de soja foi obtida.
- Os isolados virais coletados em soja no Brasil apresentaram características moleculares e biológicas variadas entre si e com o único isolado de CPMMV sequenciado completamente e caracterizado, oriundo de feijão caupi em Gana na África, evidenciando que os isolados brasileiros são de uma estirpe diferente.
- Eventos de recombinação foram encontrados ao longo do genoma viral especialmente na região correspondente à polimerase viral.
- O estudo deixou claro que os sintomas relacionados à doença da necrose da haste da soja eram causados pelo CPMMV e que os isolados virais causam sintomas variados em soja cv. CD206.
- A análise filogenética para os isolados sequenciados parcialmente mostrou que os agrupamentos não se baseavam na origem geográfica ou ano de coleta dos isolados, de fato evidencias de agrupamentos com base nos sintomas causados em soja cv. CD206 foram observados**.**
- Foi demostrada a existência de duas estirpes de CPMMV infectando os campos de soja brasileiros, isso foi evidenciado pela comparação par-à-par de sequencias e através das análises filogenéticas. Os isolados pertencentes a cada estirpe ocasionaram sintomas variados em soja cv. CD206.