HERMANO MONTEIRO DE BARROS PEREIRA

PHYLOGENY AND RECOMBINATION OF NEW WORLD BEGOMOVIRUSES

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

VIÇOSA
MINAS GERAIS - BRASIL
2017

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T
Pereira, Hermano Monteiro de Barros, 1988-
P436p Phylogeny and recombination of new world begomoviruses
2017 / Hermano Monteiro de Barros Pereira. – Viçosa, MG, 2017.
x, 83f. : il. (algumas color.) ; 29 cm.

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

1. Geminiviridae. 2. Begomovírus. 3. Filogenia.
4. Recombinação (Genética). 5. Vírus de plantas. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-graduação em Fitopatologia. II. Título.

CDD 22 ed. 572.8

HERMANO MONTEIRO DE BARROS PEREIRA

PHYLOGENY AND RECOMBINATION OF NEW WORLD BEGOMOVIRUSES

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

APROVADA: 24 de fevereiro de 2017.

Claudine Márcia Carvalho	Renan de Souza Cascardo

Francisco Murilo Zerbini Júnior
(Orientador

“O coração do homem planeja seu caminho, mas o Senhor lhe dirige os passos”
Provérbios 16:9.

A Deus,
Aos meus pais Eugênia Augusta e Victor Geraldo e a minha irmã Louisiana,

Dedico.

AGRADECIMENTOS

Primeiramente agradeço a Deus por guiar meus passos nesta caminhada, me dando discernimento nas minhas escolhas e colocando as melhores oportunidades em meu caminho. Agradeço também por me conceder a melhor família e as mais belas amizades.

Aos meus pais, Eugênia Augusta e Victor Geraldo, e a minha irmã Louisiana por me apoiarem em todas as situações, por se dedicarem a mim, me compreender nos mais diversos momentos e sempre me incentivar na busca da vitória e sucesso. Agradeço também a minha madrinha Rosana, que mesmo nas mais difíceis condições sempre se manteve presente e me apoiando. Obrigado pelo carinho, amor e confiança sempre depositados em mim, sem vocês eu nada seria. Essa conquista também é de vocês.

A todas minhas tias e primos, que sempre acreditaram em mim e me apoiaram na busca dos meus sonhos, o meu muito obrigado.

Ao Professor e orientador Francisco Murilo Zerbini, pela valiosa oportunidade de crescimento profissional, pela orientação neste trabalho, dedicação, respeito, atenção e pela amizade construída ao longo dos anos.

Aos mestres e amigos do Departamento de Fitopatologia da UFV, por proporcionarem excelentes condições de realização dos trabalhos, pelas amizades construídas e por todos os momentos vividos.

Aos amigos do Laboratório de Virologia Vegetal Molecular (LVVM), Alison, Angélica, Adriana, César, Professora Claudine, Cleydson, Diogo, Gabriela, Glória, Igor, João Paulo, Josi, Laíse, Larissa, Madelaine, Márcio, Marcos, Murilo, Osvaldo, Patrícia, Roberta, Roberto, Sarah, Sílvia, Talita, Tarsiane, Tathiana, Vitor e, em especial, à Camila pela amizade, compreensão nos momentos difíceis e valiosa participação no trabalho.

Aos amigos do Laboratório de Microbiologia Industrial, Ana, André, Fernanda Bruckner, Fernanda Pereira, Flávia, Laiane, Rafael e Professora Poliane.

A todos meus amigos da graduação e Pós-Graduação em Fitopatologia Rosana, Igor, Inaia, Nina, Mari, Maria Izabel, Tomás, Cileimar, Cris, Roberta, Maíra e Daiana.

A Larissa e Talita, que me ajudaram todos os dias fornecendo valiosas dicas e contribuições para o meu trabalho e ainda propiciaram uma agradável convivência no laboratório. A grata surpresa da amizade do Diogo e Sarah, pelas zueras sem fim de todos os dias. ENQVB.

Ao André, que sempre atentos aos detalhes, esticou a mão amiga e me impulsionou nesta conquista.

Ao Igor, meu verdadeiro irmão, a tantos momentos vividos conselhos e ensinamentos, a minha profunda gratidão e carinho.

Aos laços de amizades criados em Viçosa, em especial a Inaia, Nina, Rosana, Tomás, Talita, Larissa, Mari, Roberta, Sarah, Merielle e a tantos outros que me ajudaram nesta caminhada, deixando-a mais leve. Amizades sinceras que permanecerão junto a mim para sempre. Sou muito grato a vocês por me acolherem e compartilharem comigo seus momentos.

À UFV, que possibilitou a execução do trabalho fornecendo estrutura e boas condições assim permitindo alcançar meu título.

Agradeço também a todos que de alguma forma contribuíram para meu trabalho, tendo paciência de ensinar e fornecer inestimáveis dicas e conhecimentos que sempre me ajudaram.

A todos vocês, os meus sinceros agradecimentos!

BIOGRAFIA

Hermano Monteiro de Barros Pereira, filho de Eugênia Augusta Monteiro de Barros Pereira e Victor Geraldo Pereira, nasceu em 28 de setembro de 1988, em Itaúna, Minas Gerais.

Concluiu o curso técnico em agropecuária concomitante ao ensino médio em 2006 pela Universidade Federal de Viçosa, Campus Florestal, e em 2009 ingressou no curso de graduação em Agronomia pela Universidade Federal de Viçosa (UFV), campus Viçosa, vindo a graduar-se em agosto de 2014. Em março de 2015, iniciou o curso de Mestrado no Programa de Pós-Graduação em Fitopatologia da UFV, defendendo sua dissertação em fevereiro de 2017.

SUMÁRIO

RESUMO

PEREIRA, Hermano Monteiro de Barros, M.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **Phylogeny and recombination of New World begomoviruses.** Orientador: Francisco Murilo Zerbini Júnior. Coorientadora: Camila Geovana Ferro.

O gênero *Begomovirus* (família *Geminiviridae*) é constituído por vírus que apresentam um ou dois componentes genômicos de DNA circular de fita simples (ssDNA), infectam plantas dicotiledôneas e são transmitidos naturalmente pela mosca-branca *Bemisia tabaci* (Homoptera: Aleyrodidae). Os begomovírus constituem um importante grupo de patógenos de plantas responsáveis por perdas severas em diversas culturas de importância econômica, principalmente em regiões tropicais e subtropicais. Com base em relacionamento filogenético e organização do genoma, os begomovírus podem ser divididos em dois grupos: Velho Mundo (VM) e Novo Mundo (NM). Os begomovírus do NM possuem em sua maioria dois componentes genômicos, denominados DNA-A e DNA-B. Recentemente, a associação de alguns begomovírus do NM com DNAs satélites foi demonstrada. Begomovírus evoluem a taxas comparáveis às de vírus que possuem genoma de RNA. Embora se saiba que a mutação e a recombinação são os principais mecanismos geradores de variabilidade para esses vírus, ainda falta uma maior compreensão das forças evolucionárias que atuam sob as populações virais. Os objetivos deste estudo foram: (*i*) obter mais informações sobre os mecanismos evolucionários que influenciam a variabilidade genética dos begomovírus no NM; (*ii*) avaliar, por meio de filogenia, o efeito de recombinação na evolução dos begomovírus do NM; (*iii*) mensurar, por meio de análise filogeográfica, o efeito da migração na evolução dos begomovírus do NM. Sequências-referência de todos os DNA-A e DNA-B de begomovírus do NM foram obtidas do GenBank para montagem dos conjuntos de dados. Análises de recombinação evidenciaram um relacionamento entre begomovírus e alfassatélites do NM, sugerindo que a recombinação é um importante mecanismo evolucionário afetando a evolução do DNA-A, e em particular do gene *Rep* (presente em ambos os agentes). A retirada de blocos recombinantes mostrou-se desnecessária para a construção de filogenias, uma vez que não alterou o sinal filogenético. Foi encontrado um forte agrupamento entre as espécies baseado nos seus locais de origem, sugerindo que a introdução/migração de genomas oriundos de diferentes áreas contribui grandemente para o aumento dos eventos de recombinação e pseudo-recombinação, o que aumentaria a probabilidade do surgimento de novas variantes virais mais bem adaptadas que seus parentais. Resultados da filogeografia também indicam que DNA-A e DNA-B possuem histórias evolutivas distintas, já que as análises sugerem ancestrais diferentes. Em conjunto, os resultados indicam um importante papel da recombinação na diversificação dos begomovírus do NM.

ABSTRACT

PEREIRA, Hermano Monteiro de Barros, M.Sc., Universidade Federal de Viçosa, February, 2017. **Phylogeny and recombination of New World begomoviruses.** Advisor: Francisco Murilo Zerbini Júnior. Co-adviser: Camila Geovana Ferro.

The genus *Begomovirus* (family *Geminiviridae*) includes viruses with mono- and bipartite genomes of circular, single-stranded DNA (ssDNA), which infect dicot plants and are transmitted by the whitefly *Bemisia tabaci*. Begomoviruses constitute an important group of plant pathogens responsible for severe losses in several crops of economic importance, mainly in tropical and subtropical regions. Based on phylogenetic relationships and genomic organization, begomoviruses can be divided into New World (NW) and Old World (OW) groups. NW begomoviruses have mostly bipartite genomes, with the two components named DNA-A and DNA-B. Recently, the association of a small number of NW begomoviruses with alphasatellites was demonstrated. Begomoviruses evolve at high rates, compared to viruses with RNA genomes. Although it is well established that mutation and recombination are the main sources of genetic variability for these viruses, a better understanding of the evolutionary forces that act upon begomovirus populations is still lacking. The objectives of this study were: (*i*) to obtain more detailed information on the evolutionary mechanisms that influence the genetic variability of NW begomoviruses; (ii) to assess, phylogenetically, the effect of recombination on the evolution of NW begomoviruses; (iii) to measure, by phylogeographic analysis, the effect of migration on the diversification of NW begomoviruses. Datasets including all DNA-A and DNA-B reference sequences of NW begomoviruses were obtained from GenBank. Recombination analysis revealed a relationship between NW begomoviruses and alphasatellites, suggesting that recombination is an important evolutionary mechanism affecting DNA-A evolution, particularly of the *Rep* gene (present in both agents). Removal of recombinant blocks from the datasets was shown to be unnecessary, as it did not affect the phylogenetic signal when reconstructing phylogenies. Clusters among species were observed based on their locals of origin, suggesting that the introduction/migration of genomes from different areas greatly contributes to recombination and reassortment, which would increase the probability of the emergence of new variants better adapted than their parental viruses. Phylogeography results suggest that the DNA-A and DNA-B have distinct evolutionary histories, since they have different common ancestors. Together, the results indicate an important role of recombination in the diversification of NW begomoviruses.

INTRODUÇÃO GERAL

Os vírus classificados na família *Geminiviridae* possuem genoma composto de DNA fita simples (ssDNA) circular, envolto por um único tipo de proteína capsidial que confere à partícula uma morfologia icosaédrica geminada (Zerbini *et al.*, 2017). A família é composta por nove gêneros (*Begomovirus*, *Becurtovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus* e *Turncurtovirus*), definidos com base no tipo de inseto vetor, gama de hospedeiros, organização genômica e relacionamento filogenético (Varsani *et al.*, 2014; Varsani *et al.*, 2017; Zerbini *et al.*, 2017).

O gênero *Begomovirus* inclui vírus de plantas de grande relevância para a agricultura, devido à severidade das doenças por eles causadas, principalmente em regiões tropicais e subtropicais (Patil e Fauquet, 2009; Rocha *et al.*, 2013; Sattar *et al.*, 2013; Mabvakure *et al.*, 2016; Zaidi *et al.*, 2016). Os vírus classificados nesse gênero possuem um ou dois componentes genômicos, infectam plantas dicotiledôneas e são transmitidos naturalmente pela mosca-branca *Bemisia tabaci* (Homoptera: Aleyrodidae) (Brown *et al.*, 2015). Com base em estudos filogenéticos e características do genoma, os begomovírus podem ser divididos em dois grupos: Velho Mundo (VM; Europa, África e Ásia) e Novo Mundo (NM; as Américas) (Rybicki, 1994; Padidam *et al.*, 1999; Paximadis *et al.*, 1999; Briddon *et al.*, 2010).

A maioria dos begomovírus do VM possui genoma monossegmentados e estão normalmente associados a DNAs satélites (Mansoor *et al.*, 2003). Os begomovírus encontrados no Novo Mundo possuem genoma bissegmentado, com exceção do *Tomato leaf deformation virus* (ToLDeV), encontrado no Peru e no Equador (Melgarejo *et al.*, 2013; Sanchez-Campos *et al.*, 2013). Os componentes genômicos dos begomovírus bissegmentados são denominados DNA-A (que contém genes envolvidos na replicação, supressão de respostas de defesa da planta e encapsidação do genoma viral) e DNA-B (que contém genes envolvidos no movimento viral intra- e intercelular) (Rojas *et al.*, 2005).

Três tipos de DNAs satélites podem estar associados a begomovírus: alfassatélites (anteriormente denominados DNA-1), betassatélites (anteriormente DNA-β) e deltassatélites (Zhou, 2013; Lozano *et al.*, 2016). Esses agentes são dependentes do begomovírus auxiliar para movimento, encapsidação e transmissão pelo vetor e, no caso dos beta- e deltassatélites, também para a replicação (Zhou, 2013; Fiallo-Olivé *et al.*, 2016). Os alfassatélites ocorrem majoritariamente associados a complexos begomovírus/betassatélites no Velho Mundo (Leke *et al.*, 2015), entretanto, recentemente foram encontrados em associação com begomovírus bissegmentados do Novo Mundo (Brasil, Cuba e Venezuela) (Paprotka *et al.*, 2010; Romay *et al.*, 2010; Jeske *et al.*, 2014).

Os alfa- e betassatélites constituem-se de uma molécula de ssDNA de aproximadamente 1300 nucleotídeos. Ambos possuem organização genômica distinta dos begomovírus, entretanto possuem uma estrutura em forma de grampo, idêntica à dos begomovírus, que é fundamental para a replicação por círculo rolante. O genoma dos betassatélites compartilha apenas a sequência conservada presente na alça do grampo (5'-TAATATTAC-3') com o begomovírus auxiliar. Outras características do genoma desses agentes são uma região altamente conservada entre os betassatélites ("satellite conserved region", SCR), em posição análoga à região comum dos begomovírus; uma região rica em adenina; e um único gene, *betaC1*, codificado no sentido complementar (Briddon *et al.*, 2008). A proteína betaC1 atua como determinante de patogenicidade, supressora de silenciamento gênico transcricional e pós-transcricional (Cui *et al.*, 2004; Yang *et al.*, 2011; Li *et al.*, 2014). No genoma dos alfassatélites é encontrada uma estrutura em forma de grampo que contém o nanonucleotídeo 5'-TAGTATTAC-3', idêntico ao dos vírus classificados na família *Nanoviridae* (Zhou, 2013). Também é identificado um único gene, *alpha-Rep*, que codifica uma proteína essencial à replicação. Dessa forma, os alphassatélites são capazes de replicar de forma independente do begomovírus auxiliar, embora necessitem do vírus para encapsidação, movimento na planta e transmissão pelo vetor (Zhou, 2013).

A terceira classe de DNAs satélites associados aos begomovírus são os deltassatélites, constituídos por uma molécula de ssDNA com tamanho aproximado de 700 nucleotídeos. O genoma possui uma estrutura em forma de grampo com a sequência conservada 5'-TAATATTAC-3', sequências similares aos iterons de begomovírus, uma região rica em adenina e uma pequena região com alta identidade com a SCR dos betassatélites (Fiallo-Olivé *et al.*, 2012). Essa classe de satélite foi encontrada inicialmente em associação a begomovírus bissegmentados em plantas não-cultivadas em Cuba (Fiallo-Olivé *et al.*, 2012), e posteriormente foram relatados em batata-doce (*Ipomoea batatas*) e *Ipomoea indica* na Espanha e *Merremia dissecta* na Venezuela (Lozano *et al.*, 2016). Os possíveis efeitos da presença dessas moléculas na infecção por begomovírus é desconhecido.

Os geminivírus, incluindo os begomovírus, possuem elevada variabilidade genética, equivalente à dos vírus com genoma de RNA (Ge *et al.*, 2007; Prasanna *et al.*, 2010; Rocha *et al.*, 2013). Isso é devido principalmente aos frequentes eventos de recombinação (Padidam *et al.*, 1999; Lima *et al.*, 2013; Lefeuvre e Moriones, 2015), à pseudo-recombinção entre os vírus de genoma bissegmentado (Andrade *et al.*, 2006), e à alta taxa de mutação (Duffy e Holmes, 2008; Duffy e Holmes, 2009; Lima *et al.*, 2017).

A recombinação ocorre durante o processo de replicação viral e consiste na incorporação de um fragmento da fita de DNA ou RNA na fita de um indivíduo diferente

(Padidam *et al.*, 1999). Esse processo contribui grandemente para a diversidade genética dos begomovírus, aumentando seu potencial evolucionário e adaptação local (revisado por Lefeuvre e Moriones, 2015).

A estratégia de replicação dependente de recombinação (RDR) (Jeske *et al.*, 2001; Preiss e Jeske, 2003), em adição à replicação por círculo rolante (RCR) (Saunders *et al.*, 2001), pode explicar a elevada frequência de eventos de recombinação envolvendo os begomovírus. Além disso, as infecções mistas podem contribuir para uma maior variabilidade genética, principalmente em plantas não-cultivadas que constantemente são encontradas infectadas por vírus que apresentam um relacionamento filogenético distante (Torres-Pacheco *et al.*, 1996; Harrison *et al.*, 1997; Sanz *et al.*, 2000; Pita *et al.*, 2001; Ribeiro *et al.*, 2003; García-Andrés *et al.*, 2006; Davino *et al.*, 2009). Dessa forma, nessas plantas há uma maior probabilidade de recombinação entre espécies pouco relacionadas, o que pode gerar begomovírus com maior grau de adaptabilidade em comparação a seus progenitores (García-Andrés *et al.*, 2006; Alabi *et al.*, 2008; Silva *et al.*, 2012).

Os mecanismos precisos que controlam a recombinação em begomovírus permanecem desconhecidos (Lefeuvre e Moriones, 2015). No entanto, é conhecido que sítios recombinantes não são uniformemente distribuídos ao longo do genoma, com a existência de sítios frequentes ("hot spots") e não-frequentes ("cold spots") de recombinação (Stanley, 1995; Fauquet *et al.*, 2005; García-Andrés *et al.*, 2007; Lefeuvre *et al.*, 2009). Análises bioinformáticas para detectar vírus recombinantes ocorrendo naturalmente revelaram que a origem de replicação viral é um sítio frequente de recombinação (Gutierrez, 1999; Hanley-Bowdoin *et al.*, 1999).

Embora o papel da recombinação na geração de variabilidade de begomovírus esteja claro, uma análise detalhada de 15 conjuntos de dados de sequências de begomovírus, incluindo vírus mono- e bissegmentados no NM e VM, demonstrou que em todos os casos a mutação é o mecanismo que contribui em maior grau para a variabilidade genética (Lima *et al.*, 2017). Os autores concluíram também que não existe relação entre as freqûencias relativas desses dois processos e suas contribuições relativas para a variabilidade genética das populações de begomovírus.

A filogeografia pode ser utilizada com a finalidade de avaliar a evolução das espécies. Essa ferramenta avalia como os eventos históricos ajudaram a formar a distribuição geográfica atual dos genes, populações e espécies. Além disso, visa entender como as populações vêm comportando-se ao longo do tempo e inferir padrões de variação em seu tamanho efetivo para interpretá-los sob o ponto de vista evolutivo e biogeográfico. Deter o conhecimento e entender esses mecanismos é de grande importância quando se trata da conservação da diversidade genética das espécies (Avise, 2000; Templeton, 2004; Avise, 2009).

O objetivo deste estudo foi obter mais informações sobre os mecanismos evolucionários que influenciam a variabilidade genética dos begomovírus no NM. Análises de recombinação evidenciaram um relacionamento entre begomovírus e alfassatélites do NM, sugerindo que a recombinação é um importante mecanismo evolucionário afetando a evolução do DNA-A, e em particular do gene *Rep* (presente em ambos os agentes). Foi encontrado um forte agrupamento entre as espécies baseado nos seus locais de origem, sugerindo que a introdução/migração de genomas oriundos de diferentes áreas contribui grandemente para o aumento dos eventos de recombinação e pseudo-recombinação, o que aumentaria a probabilidade do surgimento de novas variantes virais mais bem adaptadas que seus parentais.

Literatura Citada

Alabi, O.J.; Ogbe, F.O.; Bandyopadhyay, R.; Kumar, P.L.; Dixon, A.G.O.; Hughes, J.D.; Naidu, R.A. Alternate hosts of African cassava mosaic virus and East African cassava mosaic Cameroon virus in Nigeria. **Archives of Virology,** v. 153, p. 1743-1747, 2008.

Andrade, E.C.; Manhani, G.G.; Alfenas, P.F.; Calegario, R.F.; Fontes, E.P.B.; Zerbini, F.M. *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from *Sida* sp., forms pseudorecombinants with begomoviruses from tomato but not from *Sida*. **Journal of General Virology,** v. 87, p. 3687-3696, 2006.

Avise, J.C. **Phylogeography: the history and formation of species**. Cambridge, MA, USA: Harvard University Press, 2000.

Avise, J.C. Phylogeography: retrospect and prospect. **Journal of Biogeography,** v. 36, p. 3-15, 2009.

Briddon, R.W.; Brown, J.K.; Moriones, E.; Stanley, J.; Zerbini, F.M.; Zhou, X.; Fauquet, C.M. Recommendations for the classification and nomenclature of the DNA-beta satellites of begomoviruses. **Archives of Virology,** v. 153, p. 763-781, 2008.

Briddon, R.W.; Patil, B.L.; Bagewadi, B.; Nawaz-ul-Rehman, M.S.; Fauquet, C.M. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. **BMC Evolutionary Biology,** v. 10, p. 97, 2010.

Brown, J.K.; Zerbini, F.M.; Navas-Castillo, J.; Moriones, E.; Ramos-Sobrinho, R.; Silva, J.C.; Fiallo-Olive, E.; Briddon, R.W.; Hernandez-Zepeda, C.; Idris, A.; Malathi, V.G.; Martin, D.P.; Rivera-Bustamante, R.; Ueda, S.; Varsani, A. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. **Archives of Virology,** v. 160, p. 1593-1619, 2015.

Cui, X.F.; Tao, X.R.; Xie, Y.; Fauquet, C.M.; Zhou, X.P. A DNA beta associated with *Tomato yellow leaf curl China virus* is required for symptom induction. **Journal of Virology,** v. 78, p. 13966-13974, 2004.

Davino, S.; Napoli, C.; Dellacroce, C.; Miozzi, L.; Noris, E.; Davino, M.; Accotto, G.P. Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. **Virus Research,** v. 143, p. 15-23, 2009.

Duffy, S.; Holmes, E.C. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. **Journal of Virology,** v. 82, p. 957-965, 2008.

Duffy, S.; Holmes, E.C. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. **Journal of General Virology,** v. 90, p. 1539-1547, 2009.

Fauquet, C.M.; Sawyer, S.; Idris, A.M.; Brown, J.K. Sequence analysis and classification of apparent recombinant begomoviruses infecting tomato in the nile and Mediterranean basins. **Phytopathology,** v. 95, p. 549-555, 2005.

Fiallo-Olivé, E.; Martinez-Zubiaur, Y.; Moriones, E.; Navas-Castillo, J. A novel class of DNA satellites associated with New World begomoviruses. **Virology,** v. 426, p. 1-6, 2012.

Fiallo-Olivé, E.; Tovar, R.; Navas-Castillo, J. Deciphering the biology of deltasatellites from the New World: maintenance by New World begomoviruses and whitefly transmission. **New Phytologist,** v. 212, p. 680-692, 2016.

García-Andrés, S.; Monci, F.; Navas-Castillo, J.; Moriones, E. Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: evidence for the presence of a new virus species of recombinant nature. **Virology,** v. 350, p. 433-442, 2006.

García-Andrés, S.; Tomas, D.M.; Sanchez-Campos, S.; Navas-Castillo, J.; Moriones, E. Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. **Virology,** v. 365, p. 210-219, 2007.

Ge, L.M.; Zhang, J.T.; Zhou, X.P.; Li, H.Y. Genetic structure and population variability of tomato yellow leaf curl China virus. **Journal of Virology,** v. 81, p. 5902-5907, 2007.

Gutierrez, C. Geminivirus DNA replication. **Cellular and Molecular Life Sciences,** v. 56, p. 313-329, 1999.

Hanley-Bowdoin, L.; Settlage, S.B.; Orozco, B.M.; Nagar, S.; Robertson, D. Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. **Critical Reviews in Plant Sciences,** v. 18, p. 71-106, 1999.

Harrison, B.D.; Zhou, X.; Otim Nape, G.W.; Liu, Y.; Robinson, D.J. Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. **Annals of Applied Biology,** v. 131, p. 437-448, 1997.

Jeske, H.; Kober, S.; Schäfer, B.; Strohmeier, S. Circomics of Cuban geminiviruses reveals the first alphasatellite DNA in the Caribbean. **Virus Genes,** v. 49, p. 312-324, 2014.

Jeske, H.; Lutgemeier, M.; Preiss, W. DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon mosaic virus*. **EMBO Journal,** v. 20, p. 6158-6167, 2001.

Lefeuvre, P.; Lett, J.M.; Varsani, A.; Martin, D.P. Widely conserved recombination patterns among single-stranded DNA viruses. **Journal of Virology,** v. 83, p. 2697-2707, 2009.

Lefeuvre, P.; Moriones, E. Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. **Current Opinion in Virology,** v. 10, p. 14-19, 2015.

Leke, W.N.; Mignouna, D.B.; Brown, J.K.; Kvarnheden, A. Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. **Agriculture & Food Security,** v. 4, p. 1, 2015.

Li, F.; Huang, C.; Li, Z.; Zhou, X. Suppression of RNA silencing by a plant DNA virus satellite requires a host calmodulin-like protein to repress *RDR6* expression. **PLoS Pathogens,** v. 10, p. e1003921, 2014.

Lima, A.T.M.; Silva, J.C.F.; Silva, F.N.; Castillo-Urquiza, G.P.; Silva, F.F.; Seah, Y.M.; Mizubuti, E.S.G.; Duffy, S.; Zerbini, F.M. The diversification of begomovirus populations is predominantly driven by mutational dynamics. **Virus Evolution,** v. 3, p. vex005, 2017.

Lima, A.T.M.; Sobrinho, R.R.; Gonzalez-Aguilera, J.; Rocha, C.S.; Silva, S.J.C.; Xavier, C.A.D.; Silva, F.N.; Duffy, S.; Zerbini, F.M. Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. **Journal of General Virology,** v. 94, p. 418-431, 2013.

Lozano, G.; Trenado, H.P.; Fiallo-Olive, E.; Chirinos, D.; Geraud-Pouey, F.; Briddon, R.W.; Navas-Castillo, J. Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, *Geminiviridae*): definition of a distinct class of begomovirus-associated satellites. **Frontiers in Microbiology,** v. 7, p. 162, 2016.

Mabvakure, B.; Martin, D.P.; Kraberger, S.; Cloete, L.; van Brunschot, S.; Geering, A.D.; Thomas, J.E.; Bananej, K.; Lett, J.M.; Lefeuvre, P.; Varsani, A.; Harkins, G.W. Ongoing geographical spread of *Tomato yellow leaf curl virus*. **Virology,** v. 498, p. 257-264, 2016.

Mansoor, S.; Briddon, R.W.; Zafar, Y.; Stanley, J. Geminivirus disease complexes: an emerging threat. **Trends in Plant Science,** v. 8, p. 128-134, 2003.

Melgarejo, T.A.; Kon, T.; Rojas, M.R.; Paz-Carrasco, L.; Zerbini, F.M.; Gilbertson, R.L. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. **Journal of Virology,** v. 87, p. 5397-5413, 2013.

Padidam, M.; Sawyer, S.; Fauquet, C.M. Possible emergence of new geminiviruses by frequent recombination. **Virology,** v. 265, p. 218-224, 1999.

Paprotka, T.; Metzler, V.; Jeske, H. The first DNA 1-like alphasatellites in association with New World begomoviruses in natural infections. **Virology,** v. 404, p. 148-157, 2010.

Patil, B.L.; Fauquet, C.M. Cassava mosaic geminiviruses: actual knowledge and perspectives. **Molecular Plant Pathology,** v. 10, p. 685-701, 2009.

Paximadis, M.; Idris, A.M.; Torres-Jerez, I.; Villarreal, A.; Rey, M.E.C.; Brown, J.K. Characterization of tobacco geminiviruses in the Old and New world. **Archives of Virology,** v. 144, p. 703-717, 1999.

Pita, J.S.; Fondong, V.N.; Sangare, A.; Otim-Nape, G.W.; Ogwal, S.; Fauquet, C.M. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. **Journal of General Virology,** v. 82, p. 655-665, 2001.

Prasanna, H.C.; Sinha, D.P.; Verma, A.; Singh, M.; Singh, B.; Rai, M.; Martin, D.P. The population genomics of begomoviruses: global scale population structure and gene flow. **Virology Journal,** v. 7, p. 220, 2010.

Preiss, W.; Jeske, H. Multitasking in replication is common among geminiviruses. **Journal of Virology,** v. 77, p. 2972-2980, 2003.

Ribeiro, S.G.; Ambrozevicius, L.P.; Ávila, A.C.; Bezerra, I.C.; Calegario, R.F.; Fernandes, J.J.; Lima, M.F.; Mello, R.N.; Rocha, H.; Zerbini, F.M. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. **Archives of Virology,** v. 148, p. 281-295, 2003.

Rocha, C.S.; Castillo-Urquiza, G.P.; Lima, A.T.M.; Silva, F.N.; Xavier, C.A.D.; Hora-Junior, B.T.; Beserra-Junior, J.E.A.; Malta, A.W.O.; Martin, D.P.; Varsani, A.; Alfenas-Zerbini, P.; Mizubuti, E.S.G.; Zerbini, F.M. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. **Journal of Virology,** v. 87, p. 5784-5799, 2013.

Rojas, M.R.; Hagen, C.; Lucas, W.J.; Gilbertson, R.L. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. **Annual Review of Phytopathology,** v. 43, p. 361-394, 2005.

Romay, G.; Chirinos, D.; Geraud-Pouey, F.; Desbiez, C. Association of an atypical alphasatellite with a bipartite New World begomovirus. **Archives of Virology,** v. 155, p. 1843-1847, 2010.

Rybicki, E.P. A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. **Archives of Virology,** v. 139, p. 49-77, 1994.

Sanchez-Campos, S.; Martinez-Ayala, A.; Marquez-Martin, B.; Aragon-Caballero, L.; Navas-Castillo, J.; Moriones, E. Fulfilling Koch's postulates confirms the monopartite nature of tomato leaf deformation virus: a begomovirus native to the New World. **Virus Research,** v. 173, p. 286-293, 2013.

Sanz, A.I.; Fraile, A.; García-Arenal, F.; Zhou, X.; Robinson, D.J.; Khalid, S.; Butt, T.; Harrison, B.D. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. **Journal of General Virology,** v. 81, p. 1839-1849, 2000.

Sattar, M.N.; Kvarnheden, A.; Saeed, M.; Briddon, R.W. Cotton leaf curl disease - an emerging threat to cotton production worldwide. **Journal of General Virology,** v. 94, p. 695-710, 2013.

Saunders, K.; Bedford, I.D.; Stanley, J. Pathogenicity of a natural recombinant associated with ageratum yellow vein disease: implications for geminivirus evolution and disease aetiology. **Virology,** v. 282, p. 38-47, 2001.

Silva, S.J.C.; Castillo-Urquiza, G.P.; Hora-Junior, B.T.; Assunção, I.P.; Lima, G.S.A.; Pio-Ribeiro, G.; Mizubuti, E.S.G.; Zerbini, F.M. Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. **Plant Pathology,** v. 61, p. 457-467, 2012.

Stanley, J. Analysis of African cassava mosaic virus recombinants suggest strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. **Virology,** v. 206, p. 707-712, 1995.

Templeton, A.R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Molecular Ecology,** v. 13, p. 789-809, 2004.

Torres-Pacheco, I.; Garzón-Tiznado, J.A.; Brown, J.K.; Becerra-Flora, A.; Rivera-Bustamante, R. Detection and distribution of geminiviruses in Mexico and the Southern United States. **Phytopathology,** v. 86, p. 1186-1192, 1996.

Varsani, A.; Navas-Castillo, J.; Moriones, E.; Hernández-Zepeda, C.; Idris, A.; Brown, J.K.; Zerbini, F.M.; Martin, D.P. Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. **Archives of Virology,** v. 159, p. 2193-2203, 2014.

Varsani, A.; Roumagnac, P.; Fuchs, M.; Navas-Castillo, J.; Moriones, E.; Idris, A.; Briddon, R.W.; Rivera-Bustamante, R.F.; Zerbini, F.M.; Martin, D.P. *Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. **Archives of Virology,** v. 162, p. 1819-1831, 2017.

Yang, X.L.; Xie, Y.; Raja, P.; Li, S.Z.; Wolf, J.N.; Shen, Q.T.; Bisaro, D.M.; Zhou, X.P. Suppression of methylation-mediated transcriptional gene silencing by betaC1-SAHH protein interaction during geminivirus-betasatellite infection. **PLoS Pathogens,** v. 7, p. e1002329, 2011.

Zaidi, S.S.; Martin, D.P.; Amin, I.; Farooq, M.; Mansoor, S. Tomato leaf curl New Delhi virus: a widespread bipartite begomovirus in the territory of monopartite begomoviruses. **Molecular Plant Pathology**, p. doi: 10.1111/mpp.12481., 2016.

Zerbini, F.M.; Briddon, R.W.; Idris, A.; Martin, D.P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Varsani, A.; ICTV Consortium. ICTV Virus Taxonomy Profile: *Geminiviridae*. **Journal of General Virology,** v. 98, p. 131-133, 2017.

Zhou, X. Advances in understanding begomovirus satellites. **Annual Review of Phytopathology,** v. 51, p. 357-381, 2013.

PHYLOGENY AND RECOMBINATION OF NEW WORLD BEGOMOVIRUSES

Pereira, H.M.B., Ferro, C.G., Nogueira, A.M., Ramos-Sobrinho, R., Zerbini, F.M. Phylogeny and recombination of New World begomoviruses. Virology Journal, *em preparação*.

PHYLOGENY AND RECOMBINATION OF NEW WORLD BEGOMOVIRUSES

Hermano M.B. Pereira[1], Camila G. Ferro[1], Angélica M. Nogueira[1], Roberto Ramos-Sobrinho[2], F. Murilo Zerbini[1]*

[1]Dep. de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[2]Departamento de Fitossanidade/CECA, Universidade Federal de Alagoas, Rio Largo, AL, 57100-000, Brazil

*Corresponding author (zerbini@ufv.br)

Summary

Begomoviruses are whitefly-transmitted, single-stranded DNA plant viruses that constitute one of the largest and most important groups of emerging plant viruses, growing in importance as the number of species characterized and their economic impact continues to increase. Most New World (NW) begomoviruses have two genomic components designated as DNA-A and DNA-B, and their genomes have the capacity to evolve quickly via mutation, reassortment and recombination. Datasets including all DNA-A and DNA-B reference sequences of NW begomoviruses were obtained from GenBank and analyzed to understand begomovirus evolution in the NW. Recombination was confirmed as a very important evolutionary mechanism acting on DNA-A evolution, particularly upon the *Rep* gene. The analysis revealed a relationship between NW begomoviruses and alphasatellites. Removal of recombinant blocks from the datasets was shown to be unnecessary, as it did not affect the phylogenetic signal when reconstructing phylogenies. Clustering of species was found based on their local of origin, suggesting that the introduction/migration of genomes from different areas greatly contributes to the increase of recombination and reassortment events, which would increase the probability of the emergence of new variants better adapted than their parents. Phylogeography results suggest that the DNA-A and DNA-B have distinct evolutionary histories, since they have different common ancestors.

Introduction

The family *Geminiviridae* includes plant viruses with a genome comprised of one or two molecules of circular, single-stranded DNA (ssDNA) encapsidated by a single structural protein in twinned quasi-icosahedral particles [1]. Members of this family infect a broad range of plants species causing devastating diseases, mainly in subtropical and tropical countries [2-6]. The family is divided into nine genera (*Begomovirus*, *Becurtovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus and Turncurtovirus*) based on the type of insect vector, host range, genome organization and phylogenetic relationships [1, 7].

Begomoviruses infect dicot plants and are transmitted by whiteflies of the *Bemisia tabaci* cryptic species complex (Hemiptera: Aleyrodidae). These viruses constitute the largest and one of the most important groups of plant viruses, growing in importance as the number of species characterized and their economic impact continues to increase [8, 9]. Members of this genus are further subdivided into two groups: Old World (OW; Europe, Africa, Asia and Oceania) and New World (NW; the Americas) based on genome features, phylogeny and geographical distribution [10-15].

Begomoviruses in the New World are mostly bipartite (DNA-A and DNA-B), except for *Tomato leaf deformation virus* (ToLDeV), an indigenous NW monopartite virus [16]. The DNA-A contains five open reading frames (ORFs) which encode proteins with functions in viral replication, suppression of host defenses and particle formation, while the two ORFs in the DNA-B encode proteins which are involved in intra- and intercellular movement in the plant, host range determination and suppression of host defenses [14, 17-19].

Commonly, two classes of circular single-stranded DNA molecules have been described in association with geminiviruses, the alphasatellites and the betasatellites [20]. Recently, a third class of DNA satellites, the deltasatellites, has been reported in association with begomoviruses [21].

Betasatellites have so far only been identified in the Old World, associated with mono- and bipartite begomoviruses [22]. Alphasatellites have also been found in association with NW bipartite begomoviruses, in Brazil, Cuba and Venezuela [23-26]. Both are ssDNA molecules of approximately half the size of begomovirus DNAs and code a single gene, βC1 and *alpha-Rep* respectively [20, 27]. These molecules are dependent on a helper virus for movement, encapsidation and transmission by the insect vector, and in the case of betasatellites also for replication [28, 29]. The deltasatellites have 700 nucleotides, do not encode proteins, and their effects in infections by begomoviruses are unknown [21].

Begomovirus genomes are capable of evolving rapidly through both mutation and recombination, having levels of genetic variability resembling those of viruses with RNA genomes. Recombination greatly contributes to the genetic diversity of begomoviruses, increasing their evolutionary potential and local adaptation (revised by Lefeuvre and Moriones [30]). Pseudo-recombination (or reassortment) is another source of genetic variability, which may occur by the exchange and acquisition of genomic components between bipartite viruses and satellites [10, 14, 31-33]. However, and despite the significant contribution of recombination and pseudo-recombination in the generation of genetic variability in begomoviruses, their diversification is driven primarily by mutational dynamics [34].

Mixed infections by begomoviruses are common in non-cultived hosts, and these hosts can contribute for greater genetic variability since they are often infected by viruses that present a distant phylogenetic relationship [26, 35-44]. Therefore, in these plants there is a greater probability of recombination between distantly related begomoviruses, which can generate new viruses with a greater degree of adaptability in comparison to its progenitors [37, 38].

Frequent recombination events, pseudo-recombination in bipartite viruses, and high mutation rates all contribute significantly to the emergence of new begomoviruses, increasing their potential of adaptation to different hosts and environmental conditions [33]. The expanded geographical range of *B. tabaci* populations [specially the Mediterranean (MED) and Middle East-Asia Minor 1 (MEAM1) species, which are more invasive and polyphagous] has facilitated the dissemination of begomoviruses beyond their normal geographical ranges [5], as well as the horizontal transfer of indigenous begomoviruses into crop plants [2, 45].

The recent description of a number of highly divergent gemini-like viruses in several countries, including Brazil [46-49], the detection of alphasatellites and the discovery of a new group of DNA satellites in association with NW begomoviruses [21, 23-26], all indicate that we are only beginning to recognize the full extent of the diversity of this group of pathogens. As much as we understand the evolutionary forces that act upon begomovirus populations, it is still not clear how these forces interact to drive the evolution of these viruses and of their associated DNA satellites.

The objective of this study was to obtain more detailed information about the evolutionary mechanisms that influence the genetic diversity of begomoviruses in the New World. Network analysis evidenced the relationship between NW begomoviruses and alphasatellites, suggesting that recombination directly affects the evolution of the DNA-A, particularly the *Rep* gene. We show that begomoviruses cluster phylogenetically based on their respective places of origin, suggesting that the introduction/migration of viruses from different areas greatly contributes to the increase of recombination and pseudo-recombination, which could increase the probability of the appearance of new viral variants better adapted than their progenitors.

Methods

Begomovirus datasets

All reference sequences of DNA-A and DNA-B of NW begomoviruses were obtained from GenBank (www.ncbi.nlm.nih.gov/refseq/). The dataset also included sequences of some OW bipartite begomoviruses which are more similar to NW than OW begomoviruses [50, 51]. To confirm assigned taxonomy, the sequences were analyzed using Species Demarcation Tool v. 1.2 [52].

Sequences were classified according to their sampling sites, using natural geographic features to divide the NW into five regions: Eastern South America (ESA; Argentina, Bolivia, Brazil, Paraguay and Uruguay), Western South America (WSA; Chile, Colombia, Ecuador and Peru), Central America (CA; Costa Rica, Guatemala, Honduras, Nicaragua and Panama), North America (NA; Mexico and the United States) and the Caribbean (CAR; Cuba, Jamaica, Puerto Rico and Venezuela).

Multiple sequence alignments were prepared for the full-length DNA-A and DNA-B and for the *CP*, *Rep*, *MP* and *NSP* genes of each viral species using the MUSCLE algorithm [53] implemented in MEGA6 [54].

Network and recombination analyses

Evidence of non-tree-like evolution was assessed for DNA-A, DNA-B, *CP*, *Rep*, *NSP* and *MP* datasets using the Neighbor-Net method [55] implemented in SplitsTree v. 4.10 [56]. Network images were edited using CorelDraw.

Recombination analysis was performed using the Rdp, Geneconv, Bootscan, Maximum Chi Square, Chimaera, SisterScan and 3Seq methods implemented in Recombination Detection Program (RDP) v. 4.77 [57]. Alignments were scanned with default settings for the different methods. Statistical significance was inferred by *p*-values lower than a Bonferroni-corrected cut-off of 0.05. Only recombination events detected by at least three different methods were considered to be reliable. To avoid the interference of recombination in the phylogenetic analysis, recombinant blocks in multiple sequence alignments were replaced by question marks (lost data) and later applied to the phylogeny in these data.

Phylogenetics analysis

Phylogenetic trees for the full-length DNA-A were inferred using the maximum likelihood (ML) method with RAxML v. 7.0.3 [58]. Phylogenetic trees for the full-length DNA-B and for the *CP*, *Rep*, *MP* and *NSP* genes were estimated using Bayesian inference with the program MrBayes v. 3.0b4 [59]. The models were generated with MrModeltest v. 2.2 [60] on the Akaike Information Criterion (AIC). Phylogenetic trees were visualized with the program FigTree (tree.bio.ed.ac.uk/software/figtree) and edited using CorelDraw.

Phylogeographic analysis

Spatial dispersion of NW begomoviruses was estimated separately for the DNA-A and DNA-B by Bayesian inference with the program BEAST v. 2.4.4 [61]. The geographic matrix was constructed based on the sampling locations of each species. The dataset was analyzed under demographic models of constant size, exponential growth and Bayesian Skyline. Strict and relaxed molecular clocks were used for all the models applied, assuming reversible substitutions. To analyze the effective size of the sample, the parameters used were analyzed using TRACER v. 1.4 (beast.bio.ed.ac.uk/Tracer), to verify if they converged. The MCMC tree was summarized using TreeAnnotator v. 2.4.4 (beast.bio.ed.ac.uk/TreeAnnotator), visualized with FigTree and edited with CorelDraw.

Results

Begomovirus diversity in the New World

Datasets containing a total of 125 and 97 full-length DNA-A and DNA-B sequences, respectively, were assembled, in which each sequence represents a NW begomovirus species (Suppl. Table S1). Although most begomoviruses naturally occurring in the NW have bipartite genomes, only the DNA-A is used to assign taxonomy [62]. Therefore, bipartite begomoviruses for which only the DNA-A sequence is available were included, leading to the higher number of DNA-A sequences retrieved. Using pairwise comparisons of the DNA-A sequences and the ≥91% nucleotide identity criterion established by *Geminiviridae* Study Group of the ICTV [8], we confirmed that all 125 DNA-A sequences corresponded to different species, with nucleotide identity percentages ranging from 65% to 90%. Percent identities for the full-length DNA-B sequences ranged from 65% to 98% (data not shown).

Non-tree-like evolution

Network analysis based on the DNA-A (Fig. 1a) revealed four main clusters (I, II, III and IV), with the remaining viruses constituting an undefined group. The clusters are mostly associated with geography: clusters I and II include viruses from CA/NA/CAR (plus two NW-like OW viruses, CoGMV and CoYVV), while cluster III is comprised entirely of viruses from ESA (*Oxalis yellow vein virus*, although reported in the United States, was found in ornamental plants originating from Brazil). The exception is cluster IV, which contains viruses from all regions.

A complex recombination pattern was observed, based on several ramifications in the networke, with strong evidence of a recombinant origin for the viruses in cluster IV (Fig. 1a). Interestingly, we also found evidence of a recombinant origin for viruses in cluster II (Fig.1a), which contains the two NW-like OW viruses. Cluster IV displays a higher degree of divergence compared to the other clusters, suggesting a strong recombination event that diverts the viruses in this cluster from the remaining ones in the dataset.

Networks based on the *CP* and *Rep* genes (Figs. 1b and 1c, respectively) are deeply incongruent, visualized across different patterns. The *Rep* network shows a similar pattern to the DNA-A, suggesting that this clustering is strongly influenced by recombination occurring mostly in the *Rep* gene. In the *Rep* network (Fig. 1c), the evidence of recombination in cluster II is stronger than in the network based on the DNA-A (Fig. 1a).

Due to the complex recombination pattern presented by cluster IV (Figs. 1a and 1c) and the fact that some of these viruses are associated with alphasatellites (EuYMV and MClMV), a new dataset was generated which included three sequences of NW alphassatelites. Analyzing the network, it is possible to recognize a distinct pattern of recombination between alphasatellites and divergent viruses (Fig. 1d), indicating that viral and alphasatellite *Rep* sequences share a common recombination event.

Network analysis based on the DNA-B indicated a weak recombination signal. The network reveled three main clusters (I, II and III), with support for cluster I (Fig. 1e). Cluster I was also observed in the DNA-A network (Fig. 1a). Networks based on the *NSP* and *MP* genes were highly congruent (Figs. 1f and 1g) and exhibited the same patterns visualized in the DNA-B network, with the same three clusters.

Recombination analysis

RDP results for the DNA-A dataset confirmed the complex pattern of recombination indicated by network analysis (100 events - *data not shown*), emphasizing the high divergence present in cluster IV. Fifteen recombination events were detected involving viruses in cluster IV and with breakpoints occurring mostly in the *Rep* gene (Table 1), confirming the involvement of this gene in the origin of these divergent viruses.

Because of the high divergence involving the 20 viruses in cluster IV, we prepared a dataset without their sequences. RDP results of this dataset (Suppl. Table S2) showed 70 recombination events involving sequences from clusters I, II and III. Viruses in cluster I share a recombination event with breakpoints located in the *CP* and *Rep* genes (events 5 and 50; Suppl. Table S2). Interestingly, all viruses belonging to cluster II share at least one recombination event with breakpoints in the *Rep* gene (events 13, 14, 20, 33, and 48; Suppl. Table S2). For the full-length DNA-B, 39 unique recombination events were observed, with breakpoints occurring in the *NSP* and *MP* genes (Suppl. Table S3).

Phylogenetics analysis

A maximum likelihood phylogenetic tree based on the DNA-A (without the 20 divergent viruses) shows the same clusters found in network analysis (Fig. 2a; clusters I, II, III). Cluster II indicates a strong phylogenetic relationship between NW viruses and the two NW-like viruses from Vietnam (CoGMV and CoYVV). Cluster III includes only viruses from ESA (as noted above, OxYVV is actually a ESA virus). On the contrary, cluster I contains viruses from all geographic regions except ESA.

Although the ML phylogenetic tree (Fig. 2a) exhibits the same topology observed in the network tree (Fig. 1a), the ML tree shows a very low bootstrap support for several branches, in particular the most basal ones. This is probably due to the several recombination events involving the DNA-A. Regardless, it complicates the interpretation of this tree.

To investigate the effects of recombination events in the DNA-A phylogeny, a ML phylogenetic tree based on an alignment without recombinants blocks was built (Fig. 2b). The recombination-free tree displays the same three clusters (I, II and III) observed in the original DNA-A tree (compare Figs. 2a and 2b), indicating that the absence of recombinant blocks in the alignment did not restore the phylogenetic signal. Bootstrap support was again low for several branches, indicating the unfeasibility of using the complete DNA-A for the construction of reliable phylogenetic trees of NW begomoviruses.

Using datasets comprised of *CP* and *Rep* gene nucleotide sequences, both without the 20 divergent sequences, phylogenetic trees were built using Bayesian inference. The *CP* and *Rep* phylogenetic trees (Figs. 2c and 2d, respectively) are deeply incongruent.

The *Rep* tree is similar to the DNA-A tree (compare Figs. 2a and 2d) and displays two well-supported clades (clusters I and II) that are not resolved in the *CP* tree.

A Baysean tree based on DNA-B sequences (Fig. 3a) displays three main clusters (I, II, III). The clustering observed in the phylogenetic tree is very similar to that found in the network analysis (compare Figs. 1e and 3a). Differently from the DNA-A tree, all clades were robustly supported in the DNA-B tree, indicating a high reliability of the analysis. Interestingly, the *NSP* and *MP* Baysean trees (Figs. 3b and 3c, respectively) are not incongruent and both are very similar to the DNA-B tree. All clades observed for the DNA-B tree were also present in the *NSP* tree, but with lower posterior probability values. Clades are very well supported in the *MP* tree.

Recombinants blocks detected by RDP analysis (Suppl. Table S3) were removed from the DNA-B alignment and a recombination-free phylogenetic tree was built by Bayesian inference (Fig. 3d). The same clusters (I, II, III) were found (compare Figs. 3a and 3d). In cluster III it is possible to observe a small difference involving SiMBoV1 and TRYLCV. These two sequences cluster together in the recombination-free phylogenetic tree (Fig. 3d), which is not observed in the original tree (Fig. 3a).

Phylogeography

Using Bayesian Skyline demographic models, phylogeographic relationships were constructed based on DNA-A (excluding the 20 divergent viruses) and DNA-B sequences. In the DNA-A tree (Fig. 4a), the same clusters observed in the ML phylogenetic tree are present (clusters I, II and III; compare Figs. 2a and 4a). Cluster I includes sequences from CA/NA/CAR, plus BDMV from WSA. This suggests migration among the different biogeographical regions, thus allowing a greater probability of recombination. The analysis also suggests that ESA viruses originated from NA/CAR viruses.

Phylogeographic analysis was also performed based on a recombination-free DNA-A alignment (Fig. 4b). The same clusters (I, II, and III) were found, with the same ancestor relationships (viruses from ESA originating from NA/CAR). However, the SA viruses were split into two clades (Fig. 4b; clusters IIIa and IIIb).

For the DNA-B, the trees based on the complete and recombination-free alignments (Figs. 4c and 4d, respectively) exhibited the same three clusters observed in the Bayesian phylogenetic tree (Fig. 3a). This great similarity can by explained based on the lower ocurrence of recombination in the DNA-B dataset compared with the DNA-A dataset. However, the common ancestor for the DNA-B was mapped to South America, differently from what was found for the DNA-A (compare Figs. 4a and 4c) and suggesting that the DNA-A and DNA-B have distinct evolutionary histories.

Discussion

Differently from most plant viruses, which group together based on host range, begomovirus populations segregate strongly based on geographical origin [2]. NW and OW begomoviruses form clearly distinct phylogenetic clades [63], and within these two major groups, viruses from the same geographical region are closer to each other, irrespective of their hosts, than to viruses infecting the same hosts in different regions (for example, begomoviruses from South America that infect solanaceous hosts are closer to other begomoviruses from SA that infect leguminous or malvaceous hosts than to begomoviruses that infect solanaceous hosts in the Indian subcontinent). Besides mutation and selection, recombination is an important generation of variability acting on the macroevolution of begomovirus populations [34, 64-66]. Here, we decided to assess the effect of recombination in NW begomovirus evolution in greater detail. Our results confirm that recombination is a very important mechanism affecting DNA-A evolution. More interestingly, we detected recombination in the *Rep* genes of begomoviruses and alphasatellites, suggesting that these two agents may share a common evolutionary history in the NW. Evidence of migration among the different biogeographical regions was also detected.

Pairwise comparisons indicated that the DNA-B displays a higher range of nucleotide identities. Molecular variability in the DNA-A has been associated with recombination events occurring mostly in the *Rep* gene [67, 68]. We observed also that this region is highly recombinant, and a much lower number of recombination events were observed in the DNA-B dataset. Therefore, the greater molecular variability of DNA-B components observed here cannot be explained by recombination.

Network analysis based on the DNA-A evidenced several recombination events, and identified a group of 20 divergent viruses most likely due to frequent recombination (cluster IV in Fig. 1a).

Network analysis of the DNA-B indicated that recombination is not responsible for the greater variability exhibited by this component. The higher variability of the DNA-B may result from a combination of factors, including the absence of overlapping genes and the smaller number of *cis*- and *trans*-interactions between its protein products. Thus, the DNA-B would be more permissive to variation, evolving mostly in response to the host. An alternative explanation is that the DNA-B could have a distinct origin from the DNA-A, initially arising as a satellite that was captured by a monopartite progenitor and later evolved to become an integral part of the genome [10, 69].

Network analysis using a dataset comprised of viral and alphasatellite *Rep* sequences indicated an intrinsic relationship between each other. The same group of 20 divergent begomoviruses (cluster IV) identified in the network analysis is also involved in the recombination with alphasatellites. Interestingly, this group includes two of the three viruses that have been recently associated with these subviral agents: *Euphorbia yellow mosaic virus* (EuYMV) and *Melon chlorotic mosaic virus* (MClMV). Previous studies have indicated that these viruses comprise a distinct phylogenetic lineage of NW begomoviruses, which remarkably is not associated with geographical location [2, 70]. While the group is comprised mostly of begomovirus found in CA and NA, it includes EuYMV and a few other viruses from ESA. The network analysis suggests a possible involvement of *alpha-Rep* in promoting speciation of this divergent group of NW begomoviruses. Interestingly, *alpha-Rep* is believed to have originated from the DNA-R component of nanoviruses (family *Nanoviridae*) [20], a group of viruses that has never been detected in the NW.

Phylogenetic trees exhibited a pattern similar to the clustering of network analysis, almost entirely coincidental with the places of origin of each virus. A relationship among CA and NA viruses is evident in cluster I. Also, the clade comprised of ESA viruses (cluster III) show a specific subclustering within this large region. Rocha *et al.* [2] reported equivalent results, showing evidence of a geographical segregation of Brazilian begomoviruses. However, the DNA-A phylogenetic tree displayed very low bootstrap scores, indicating insufficient phylogenetic signal in the dataset. This was observed even when a recombination-free dataset was analyzed (and even when the 20 divergent sequences from cluster IV were removed from the dataset), suggesting that the full-length DNA-A is inappropriate for phylogenetic reconstruction of NW begomoviruses. This could be due to the high genetic variability of these begomoviruses, which prevents a proper multiple sequence alignment.

The *Rep* and *CP* phylogenetic trees are incongruent, but their clades are very well suported. This suggests that these two genes may have different evolutionary histories, which could explain the low statistical support of the phylogenetic analysis based on the complete DNA-A. Interestingly, phylogenetic relationships inferred from *Rep*, but not from *CP*, reflect the phylogeny based on the complete DNA-A.

The *CP* gene is the most conserved among begomoviruses [64, 71]. This gene has been used for provisional classification of begomovirus species, and it is the only viral gene involved in insect vector transmission and in particle formation [72, 73]. These features could impose severe structural constraints on the CP protein, and in fact several studies indicate that begomovirus CPs are under strong negative selection [67, 71, 74]. Thus, the CP may be a poor surrogate for the detection of evolutionary relationships among begomoviruses.

The *Rep* gene, on the other hand, displays considerably greater genetic variability and is a recombination hotspot [64, 65]. The N-terminal region of the Rep protein includes conserved motifs essential for rolling-circle replication [75-77]. Conservation of the integrity of these elements is critical for successful infection cycles, despite the variation introduced by frequent recombination. Indeed, studies have provided evidence that the high nucleotide variability in the 5'-terminal portion of the *Rep* gene is accompanied by strong purifying selection that preserves the amino acid sequence [2, 67]. In our results, several of the recombination events detected did not involve the *Rep* 5'-terminal region, indicating the conservation of these elements. Nevertheless, the variability inserted by recombination in *Rep* may contribute to the diversification of the DNA-A.

Interesting results were obtained for the DNA-B datasets. Phylogenetic trees based on the full-length DNA-B and on *NSP* and *MP* sequences exhibited similar clustering, even though the DNA-B is more variable than the DNA-A [2]. Moreover, the recombination-free phylogenetic tree of the DNA-B exhibited the same clusters found in the original tree. Although this could be due to the low number of recombination events detected for this component, a more likely explanation is that the presence of recombinant blocks in the alignment does not affect the phylogenetic signal. Indeed, all clades were very well-supported with high bootstrap scores in both phylogenetic trees (with and without recombinant blocks).

Previous studies have suggested that the DNA-B can tolerate a greater number of mutations since it does not contain overlapping genes, and that the distinct origin of the DNA-B, possibly from a satellite molecule captured by a monopartite begomovirus, may contribute to an increase in variability [10, 78]. In spite of its greater variability (or maybe because of it), the great similarity among the trees based on the two coding regions of the DNA-B and the tree based on the entire component confers greater reliability for using a DNA-B dataset to reconstruct the phylogeny of NW begomoviruses.

Phylogeographic analysis provided additional evidence for the clustering of NW begomoviruses based on geographical location. The DNA-A tree without the 20 divergent sequences indicated a common ancestor from North America. All viruses from Eastern South America clustered together in the same clade, suggesting an ecological barrier effect, most likely of the Amazon rainforest and the Andes mountains. This is well illustrated by the fact that *Bean golden mosaic virus* (BGMV) and *Bean golden yellow mosaic virus* (BGYMV), two begomoviruses that infect the same hosts, have been kept separated since their description in the 1960's. BGMV is found from the eastern boundaries of the Amazon rainforest to Argentina, and BGYMV is found from southern Mexico to Colombia [79, 80].

Cluster I in the DNA-A tree contains viruses from all regions except Eastern South America. We believe it represents an evolutionary lineage separated from other NW begomoviruses. Therefore, begomoviruses placed in cluster I may represent a source of genetic variability not available (due to geographical isolation) to begomoviruses in South America. Introduction of these viruses could greatly affect begomovirus evolution in SA as they can recombine and introduce new features into South American begomoviruses.

Nevertheless, we found evidence that some long distance spread occurred between these areas. Evidence of migration can be observed for *Oxalis yellow vein virus* (OxYVV), which was originally described in North America [81] infecting a vegetatively-propagated ornamental plant which is native to South America [82, 83]. This virus clustered in the Eastern South America clade, indicating that it is related to viruses from that region rather than from North America.

The phylogeographic analysis indicated that recombination influences the spread of begomoviruses. The recombination-free tree showed a different topology compared to the original tree. Nevertheless, posterior probability values are lower in the recombination-free tree.

For the DNA-B, the analysis also indicated segregation based on geographical location, although forming distinct clades compared to the DNA-A tree. Segregation based on geography can be explained by the requisite relationship between DNA-A and DNA-B components of bipartite begomoviruses, both being necessary for the virus to establish a systemic infection and be transmitted under natural conditions [84]. Clade III of the DNA-B tree contains mostly viruses from Eastern South America, but some sequences from other regions are present (albeit with low posterior probability values). The same topology was observed in the recombination-free tree, suggesting that recombinantion does not affect the spread of the DNA-B.

This study confirmed the role of recombination as a mechanism of variability in NW begomoviruses. Nevertheless, the phylogenetic signal is maintained in the presence of recombinants blocks. Migration was detected to be very frequent into the different biogeographical regions, with its effect being highlighted by recombination, which was responsible for inserting new molecular features to indigenous begomoviruses and for mixing the genetic variability already present. Although migration and recombination deeply affect the evolution of NW begomoviruses, the maintenance of distinct clusters despite overlapping geographical regions and hosts was observed. These clusters represent sources of genetic variability, which can be transferred to begomoviruses in other clusters by migration and recombination.

References

1. Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas-Castillo J *et al.* ICTV Virus Taxonomy Profile: *Geminiviridae*. J Gen Virol 2017;98:131-133.
2. Rocha CS, Castillo-Urquiza GP, Lima ATM, Silva FN, Xavier CAD, Hora-Junior BT *et al.* Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. J Virol 2013;87:5784-5799.
3. Sattar MN, Kvarnheden A, Saeed M, Briddon RW. Cotton leaf curl disease - an emerging threat to cotton production worldwide. J Gen Virol 2013;94:695-710.
4. Patil BL, Fauquet CM. Cassava mosaic geminiviruses: actual knowledge and perspectives. Mol Plant Pathol 2009;10:685-701.
5. Mabvakure B, Martin DP, Kraberger S, Cloete L, van Brunschot S, Geering AD *et al.* Ongoing geographical spread of *Tomato yellow leaf curl virus*. Virology 2016;498:257-264.
6. Zaidi SS, Martin DP, Amin I, Farooq M, Mansoor S. Tomato leaf curl New Delhi virus: a widespread bipartite begomovirus in the territory of monopartite begomoviruses. Mol Plant Pathol 2016:doi: 10.1111/mpp.12481.
7. Varsani A, Roumagnac P, Fuchs M, Navas-Castillo J, Moriones E, Idris A *et al. Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. Arch Virol 2017;162:1819-1831.
8. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JC *et al.* Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Arch Virol 2015;160:1593-1619.
9. Varma A, Malathi VG. Emerging geminivirus problems: A serious threat to crop production. Ann Appl Biol 2003;142:145-164.
10. Briddon RW, Patil BL, Bagewadi B, Nawaz-ul-Rehman MS, Fauquet CM. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evol Biol 2010;10:97.
11. Harrison BD, Robinson DJ. Natural genomic and antigenic variation in whitefly transmitted geminiviruses (begomoviruses). Annu Rev Phytopathol 1999;39:369-398.

12. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. Virology 1999;265:218-224.
13. Paximadis M, Idris AM, Torres-Jerez I, Villarreal A, Rey MEC, Brown JK. Characterization of tobacco geminiviruses in the Old and New world. Arch Virol 1999;144:703-717.
14. Rojas MR, Hagen C, Lucas WJ, Gilbertson RL. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu Rev Phytopathol 2005;43:361-394.
15. Rybicki EP. A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. Arch Virol 1994;139:49-77.
16. Melgarejo TA, Kon T, Rojas MR, Paz-Carrasco L, Zerbini FM, Gilbertson RL. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. J Virol 2013;87:5397-5413.
17. Hanley-Bowdoin L, Bejarano ER, Robertson D, Mansoor S. Geminiviruses: Masters at redirecting and reprogramming plant processes. Nat Rev Microbiol 2013;11:777-788.
18. Mahajan N, Parameswari C, Veluthambi K. Severe stunting in blackgram caused by the Mungbean yellow mosaic virus (MYMV) KA27 DNA B component is ameliorated by co-infection or post-infection with the KA22 DNA B: MYMV nuclear shuttle protein is the symptom determinant. Virus Res 2011;157:25-34.
19. Brustolini OJ, Machado JP, Condori-Apfata JA, Coco D, Deguchi M, Loriato VA *et al.* Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. Plant Biotechnol J 2015;13:1300-1311.
20. Zhou X. Advances in understanding begomovirus satellites. Annu Rev Phytopathol 2013;51:357-381.
21. Lozano G, Trenado HP, Fiallo-Olive E, Chirinos D, Geraud-Pouey F, Briddon RW *et al.* Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, *Geminiviridae*) - definition of a distinct class of begomovirus-associated satellites. Front Microbiol 2016;7:162.
22. Briddon RW, Mansoor S, Bedford ID, Pinner MS, Saunders K, Stanley J *et al.* Identification of DNA components required for induction of cotton leaf curl disease. Virology 2001;285:234-243.
23. Paprotka T, Metzler V, Jeske H. The first DNA 1-like alphasatellites in association with New World begomoviruses in natural infections. Virology 2010;404:148-157.
24. Jeske H, Kober S, Schäfer B, Strohmeier S. Circomics of Cuban geminiviruses reveals the first alphasatellite DNA in the Caribbean. Virus Genes 2014;49:312-324.
25. Romay G, Chirinos D, Geraud-Pouey F, Desbiez C. Association of an atypical alphasatellite with a bipartite New World begomovirus. Arch Virol 2010;155:1843-1847.
26. Ferro CG, Silva JP, Xavier CAD, Godinho MT, Lima ATM, Mar TB *et al.* The ever increasing diversity of begomoviruses infecting non-cultivated hosts: new species from *Sida* spp. and *Leonurus sibiricus*, plus two New World alphasatellites. Ann Appl Biol 2017:10.1111/aab.12329.
27. Zhou X, Xie Y, Tao X, Zhang Z, Li Z, Fauquet CM. Characterization of DNAbeta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. J Gen Virol 2003;84:237-247.
28. Briddon RW, Bull SE, Amin I, Idris AM, Mansoor S, Bedford ID *et al.* Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. Virology 2003;312:106-121.

29. Briddon RW, Bull SE, Amin I, Mansoor S, Bedford ID, Rishi N *et al.* Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. Virology 2004;324:462-474.

30. Lefeuvre P, Moriones E. Recombination as a motor of host switches and virus emergence: Geminiviruses as case studies. Curr Opin Virol 2015;10:14-19.

31. Rey MEC, Ndunguru J, Berrie LC, Paximadis M, Berry S, Cossa N *et al.* Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in Southern Africa, including the South-West Indian Ocean islands. Viruses 2012;4:1753-1791.

32. Saunders K, Bedford ID, Stanley J. Adaptation from whitefly to leafhopper transmission of an autonomously replicating nanovirus-like DNA component associated with ageratum yellow vein disease. J Gen Virol 2002;83:907-913.

33. Seal SE, Van den Bosch F, Jeger MJ. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. Crit Rev Plant Sci 2006;25:23-46.

34. Lima ATM, Silva JCF, Silva FN, Castillo-Urquiza GP, Silva FF, Seah YM *et al.* The diversification of begomovirus populations is predominantly driven by mutational dynamics. Virus Evol 2017;3:vex005.

35. Torres-Pacheco I, Garzón-Tiznado JA, Brown JK, Becerra-Flora A, Rivera-Bustamante R. Detection and distribution of geminiviruses in Mexico and the Southern United States. Phytopathology 1996;86:1186-1192.

36. Sanz AI, Fraile A, García-Arenal F, Zhou X, Robinson DJ, Khalid S *et al.* Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. J Gen Virol 2000;81:1839-1849.

37. Pita JS, Fondong VN, Sangare A, Otim-Nape GW, Ogwal S, Fauquet CM. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. J Gen Virol 2001;82:655-665.

38. García-Andrés S, Monci F, Navas-Castillo J, Moriones E. Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: Evidence for the presence of a new virus species of recombinant nature. Virology 2006;350:433-442.

39. Davino S, Napoli C, Dellacroce C, Miozzi L, Noris E, Davino M *et al.* Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. Virus Res 2009;143:15-23.

40. Harrison BD, Zhou X, Otim Nape GW, Liu Y, Robinson DJ. Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. Ann Appl Biol 1997;131:437-448.

41. Silva SJC, Castillo-Urquiza GP, Hora-Junior BT, Assunção IP, Lima GSA, Pio-Ribeiro G *et al.* Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. Plant Pathol 2012;61:457-467.

42. Tavares SS, Ramos-Sobrinho R, Gonzalez-Aguilera J, Lima GSA, Assunção IP, Zerbini FM. Further molecular characterization of weed-associated begomoviruses in Brazil with an emphasis on *Sida* spp. Planta Dan 2012;30:305-315.

43. Fiallo-Olivé E, Navas-Castillo J, Moriones E, Martinez-Zubiaur Y. Two novel begomoviruses belonging to different lineages infecting *Rhynchosia minima*. Arch Virol 2010;155:2053-2058.

44. Fiallo-Olivé E, Navas-Castillo J, Moriones E, Martinez-Zubiaur Y. Begomoviruses infecting weeds in Cuba: Increased host range and a novel virus infecting *Sida rhombifolia*. Arch Virol 2012;157:141-146.

45. Ribeiro SG, Ávila AC, Bezerra IC, Fernandes JJ, Faria JC, Lima MF *et al.* Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. Plant Dis 1998;82:830.

46. Loconsole G, Saldarelli P, Doddapaneni H, Savino V, Martelli GP, Saponari M. Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family *Geminiviridae*. Virology 2012;432:162-172.

47. Basso MF, Silva JC, Fajardo TV, Fontes EP, Zerbini FM. A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. Virus Res 2015;210:27-33.

48. Ma Y, Navarro B, Zhang Z, Lu M, Zhou X, Chi S *et al.* Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease. J Gen Virol 2015;96:2421-2434.

49. Al-Rwahnih M, Dave A, Anderson MM, Rowhani A, Uyemoto JK, Sudarshana MR. Association of a DNA virus with grapevines affected by red blotch disease in California. Phytopathology 2013;103:1069-1076.

50. Ha C, Coombs S, Revill P, Harding R, Vu M, Dale J. Molecular characterization of begomoviruses and DNA satellites from Vietnam: Additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J Gen Virol 2008;89:312-326.

51. Ha C, Coombs S, Revill P, Harding R, Vu M, Dale J. Corchorus yellow vein virus, a New World geminivirus from the Old World. J Gen Virol 2006;87:997-1003.

52. Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM *et al.* A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family *Geminiviridae*). Arch Virol 2013;158:1411-1424.

53. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinf 2004;5:1-19.

54. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013;30:2725-2729.

55. Bryant D, Moulton V. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 2004;21:255-265.

56. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 2006;23:254-267.

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

58. Stamatakis A. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinf Appl Note 2006;22:2688-2690.

59. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19:1572-1574.

60. Nylander JAA. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University2004.

61. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 2007;7:214.

62. Brown JK, Fauquet CM, Briddon RW, Zerbini FM, Moriones E, Navas-Castillo J. Family *Geminiviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, editors. Virus Taxonomy - Ninth Report of the International Committee on Taxonomy of Viruses. London, UK: Elsevier Academic Press; 2012. p. 351-373.

63. Stanley J, Bisaro DM, Briddon RW, Brown JK, Fauquet CM, Harrison BD *et al.* Family *Geminiviridae*. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, editors. Virus Taxonomy - Eighth Report of the International Committee on Taxonomy of Viruses. San Diego: Elsevier Academic Press; 2005. p. 301-326.

64. Lefeuvre P, Lett JM, Varsani A, Martin DP. Widely conserved recombination patterns among single-stranded DNA viruses. J Virol 2009;83:2697-2707.

65. Lefeuvre P, Martin DP, Hoareau M, Naze F, Delatte H, Thierry M *et al.* Begomovirus 'melting pot' in the south-west Indian Ocean islands: Molecular diversity and evolution through recombination. J Gen Virol 2007;88:3458-3468.

66. Martin DP, Lefeuvre P, Varsani A, Hoareau M, Semegni JY, Dijoux B *et al.* Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. PLoS Pathog 2011;7:e1002203.

67. Lima ATM, Sobrinho RR, Gonzalez-Aguilera J, Rocha CS, Silva SJC, Xavier CAD *et al.* Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. J Gen Virol 2013;94:418-431.

68. Silva FN, Lima ATM, Rocha CS, Castillo-Urquiza GP, Alves M, Zerbini FM. Recombination and pseudorecombination driving the evolution of the begomoviruses *Tomato severe rugose virus* (ToSRV) and *Tomato rugose mosaic virus* (ToRMV): two recombinant DNA-A components sharing the same DNA-B. Virol J 2014;11:66.

69. Nawaz-Ul-Rehman MS, Fauquet CM. Evolution of geminiviruses and their satellites. Febs Letters 2009;583:1825-1832.

70. Castillo-Urquiza GP, Beserra Jr. JEA, Bruckner FP, Lima ATM, Varsani A, Alfenas-Zerbini P *et al.* Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. Arch Virol 2008;153:1985-1989.

71. De Bruyn A, Harimalala M, Zinga I, Mabvakure BM, Hoareau M, Ravigne V *et al.* Divergent evolutionary and epidemiological dynamics of cassava mosaic geminiviruses in Madagascar. BMC Evol Biol 2016;16:182.

72. Briddon RW, Pinner MS, Stanley J, Markham PG. Geminivirus coat protein gene replacement alters insect specificity. Virology 1990;177:85-94.

73. Brown JK, Idris AM, Torres-Jerez I, Banks GK, Wyatt SD. The core region of the coat protein gene is highly useful for establishing the provisional identification and classification of begomoviruses. Arch Virol 2001;146:1581-1598.

74. Ramos-Sobrinho R, Xavier CAD, Pereira HMB, Lima GSA, Assunção IP, Mizubuti ESG *et al.* Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. J Gen Virol 2014;95:2540-2552.

75. Ilyina TV, Koonin EV. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res 1992;20:3279-3285.

76. Koonin EV, Ilyina TV. Geminivirus replication proteins are related to prokariotic plasmid rolling circle DNA replication initiator proteins. J Gen Virol 1992;73:2763-2766.

77. Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibanez JT, Hanley-Bowdoin L. Functional analysis of a novel motif conserved across geminivirus Rep proteins. J Virol 2011;85:1182-1192.

78. Ho ES, Kuchie J, Duffy S. Bioinformatic analysis reveals genome size reduction and the emergence of tyrosine phosphorylation site in the movement protein of New World bipartite begomoviruses. PLoS ONE 2014;9:e111957.

79. Costa AS. Three whitefly-transmitted virus diseases of beans in São Paulo, Brazil. Plant Prot Bull FAO 1965;13:2-12.

80. Morales FJ. History and current distribution of begomoviruses in Latin America. Adv Virus Res 2006;67:127-162.

81. Herrera F, Aboughanem-Sabanadzovic N, Valverde RA. A begomovirus associated with yellow vein symptoms of *Oxalis debilis*. Eur J Plant Pathol 2015;142:203-208.

82. Luo S, Zhang D, Renner SS. *Oxalis debilis* in China: distribution of flower morphs, sterile pollen and polyploidy. Ann Bot 2006;98:459-464.

83. Lourteig A. *Oxalis* L. subgéneros *Monoxalis* (Small) Lourt., Oxalis y Trifidus Lourt. Bradea 2000;7:201-629.

84. Stanley J. Infectivity of the cloned geminivirus genome requires sequences from both DNAs. Nature 1983;305:643-645.

Table 1. Putative recombination events involving the DNA-A components of a set of 20 divergent New World begomoviruses.

Events	**Breakpoint**[a]		**Genomic region**	**Recombinants**	**Putative parents**		**Methods**[b]	***P*-value**[c]
	Begin	**End**			**Minor**	**Major**		
1	2625	2010	IR	AbMBV NC_016574	SiYBV NC_020254 SiMAlV NC_016573	CuLCrV NC_002984	RGBMCS3	1.03E-43
2	2023	2641	Rep	SiYLCV EU710750	ToCmMV NC_010835	SiYMV NC_004639	RGBMCS3	2.04E-40
3	25	1912	CP	DeLDV NC_008494	CoYSV NC_008492 MacGMV NC_010952 SiGMoV NC_014130	Unknown	RGBMCS3	3.38E-32
4	869	1944	CP/Rep	ToSLCV NC_004642	ToChLPV NC_005843	Unknown	RGBMCS3	1.83E-29
5	1972	2619	Rep	TRYLCV NC_020257	EuMV NC_008304 SMLCuV NC_004645 ToMYLCV NC_009490	TrYMV BR-Msj1-10A SiMoV NC_004637 SiYNV NC_020253	RGBMCS3	2.88E-26
6	1933	11	Rep	ToCmMV NC_010835	EuMV NC_008304 MCLCuV NC_003865	TrYMV BR-Msj1-10A OMoV NC_011181 SiYNV NC_020253	RGBMCS3	7.94E-26
7	1964	2491	Rep	CLCrV NC_004580	Unknown Unknown Unknown Unknown Unknown Unknown	BChMV NC_022005 AbMV NC_001928 BDMV NC_001931 OYMV NC_014066 SiGMCRV NC_004657 SiGMV NC_002046	RGBMCS3	1.70E-24
8	1928	24	Rep/CR	CabLCuV NC_003866 RhGMYuV NC_012481	CuLCrV NC_002984	Unknown	RGBMCS3	6.82E-25
9	1960	165	Rep/IR	BCaMV NC_003504 MCLCuV NC_003865	Unknown	BGYMV NC_001439	RGBMCS3	7.59E-26

Table 1 (cont.)

Events	Breakpoint		Genomic region	Recombinants	Putative parents		Methods	*P*-value
	Begin	End			Minor	Major		
10	1809	101	Rep/IR	MClMV NC_014380	Unknown	BleICV NC_019035	RGBMCS3	1.56E-21
11	1870	2572	Rep	PepGMV NC_004101	Unknown	ToChLPV NC_005843	RGBMCS3	4.43E-18
12	1801	2585	Rep	SqLCV NC_001936	Unknown	HeYMV BR-Mac7-10A	RGBMCS3	2.39E-23
13	2076	2440	Rep	EuYMV NC_012553 EuMV NC_008304 SqLCV NC_001936 SMLCuV NC_004645 ToMYLCV NC_009490 ToSLCV NC_004642	Unknown	DaChMV NC_018716	RGBMCS3	1.66E-11
14	601	1005	CP	SqLCV NC_001936	RhGMYuV NC_012481	MCLCuV NC_003865	RGBMCS3	9.64E-03
15	1032*	1145	CP	CabLCuJV DQ178614 CarYSYV NC_022007 RhGMYuV NC_012481	RhGMV NC_010294 MacMPRV NC_004097	BCaMV NC_003504	RGBMCS3	2.47E-04

---

[a] * breakpoint could not be precisely pinpointed.

[b] R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimera; S, SisScan; 3, 3SEQ.

[c] The reported *P* values are for the methods indicated in red, and they are the lowest *P* values calculated for the region in question.

Figure legends

Figure 1. Phylogenetic evidence for recombination based on **(a)** full-length DNA-A, **(b)** *CP*, **(c)** *Rep,* **(d)** *Rep* of begomoviruses and alphasatellites, **(e)** full-length DNA-B, **(f)** *NSP* and **(g)** *MP* nucleotide sequences of New World begomoviruses. Isolates from Eastern South America (ESA) are indicated in green, Western South America (WSA) in yellow, Central America (CA) in blue, North America (NA) in red, the Caribbean (CAR) in purple and Old World (OW) in black. Neighbor-Net network analysis was performed using SplitsTree4. Formation of a reticular network rather than a single bifurcated tree is suggestive of recombination.

Figure 2. Phylogenetic relationships among New World begomoviruses based on DNA-A sequences. **(a)** Maximum likelihood tree based on the full-length DNA-A, **(b)** Maximum likelihood tree based on a recombination-free alignment of DNA-A sequences **(c)** Bayesian tree based on the *CP* gene, **(d)** Bayesian tree based on the *Rep* gene. Isolates from Eastern South America (ESA) are indicated in green, Western South America (WSA) in yellow, Central America (CA) in blue, North America (NA) in red, the Caribbean (CAR) in purple and Old World (OW) in black.

Figure 3. Phylogenetic relationships among New World begomoviruses based on DNA-B sequences. Bayesian trees based on **(a)** full-length DNA-B, **(b)** *NSP* gene, **(c)** *MP* gene, **(d)** recombination-free alignment of DNA-B sequences. Isolates from Eastern South America (ESA) are indicated in green, Western South America (WSA) in yellow, Central America (CA) in blue, North America (NA) in red, the Caribbean (CAR) in purple and Old World (OW) in black.

Figure 4. Phylogeographic clustering of New World begomoviruses based on **(a)** DNA-A, **(b)** recombination-free DNA-A, **(c)** DNA-B, **(d)** recombination-free DNA-B. Isolates from Eastern South America (ESA) are indicated in green, Western South America (WSA) in yellow, Central America (CA) in blue, North America (NA) in red, the Caribbean (CAR) in purple and Old World (OW) in black.

Figure 1a

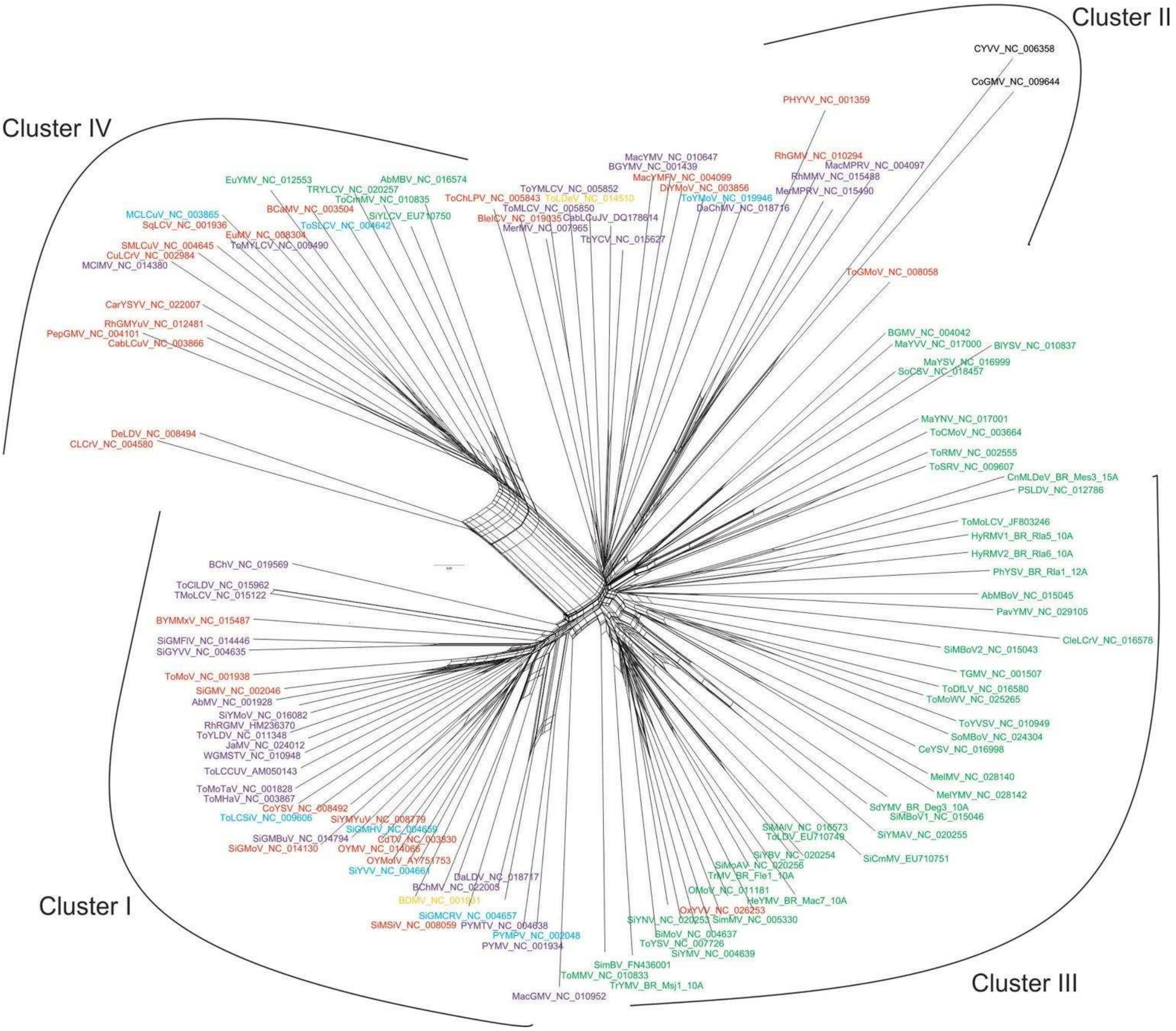
Cluster II
CYVV_NC_006358
CoGMV_NC_009644
PHYVV_NC_001359
Cluster IV
EuYMV_NC_012553
AbMBV_NC_016574
TRYLCV_NC_020257
ToCmMV_NC_010835
MCLCuV_NC_003865
SqLCV_NC_001936
BCaMV_NC_003504
ToSLCV_NC_004642
SiYLCV_EU710750
EuMV_NC_008304
SMLCuV_NC_004645
ToMYLCV_NC_009490
CuLCrV_NC_002984
MClMV_NC_014380
ToChLPV_NC_005843
ToLDeV_NC_014510
ToYMLCV_NC_005852
ToMLCV_NC_005850
BleICV_NC_019035
CabLCuJV_DQ178614
MerMV_NC_007965
TbYCV_NC_015627
MacYMV_NC_010647
BGYMV_NC_001439
MacYMFV_NC_004099
DiYMoV_NC_003856
ToYMoV_NC_019946
DaChMV_NC_018716
RhGMV_NC_010294
MacMPRV_NC_004097
RhMMV_NC_015488
MerMPRV_NC_015490
ToGMoV_NC_008058
CarYSYV_NC_022007
RhGMYuV_NC_012481
PepGMV_NC_004101
CabLCuV_NC_003866
DeLDV_NC_008494
CLCrV_NC_004580
BGMV_NC_004042
MaYVV_NC_017000
BlYSV_NC_010837
MaYSV_NC_016999
SoCSV_NC_018457
MaYNV_NC_017001
ToCMoV_NC_003664
ToRMV_NC_002555
ToSRV_NC_009607
CnMLDeV_BR_Mes3_15A
PSLDV_NC_012786
ToMoLCV_JF803246
HyRMV1_BR_Rla5_10A
HyRMV2_BR_Rla6_10A
PhYSV_BR_Rla1_12A
AbMBoV_NC_015045
PavYMV_NC_029105
CleLCrV_NC_016578
SiMBoV2_NC_015043
TGMV_NC_001507
ToDfLV_NC_016580
ToMoWV_NC_025265
ToYVSV_NC_010949
SoMBoV_NC_024304
CeYSV_NC_016998
MelMV_NC_028140
MelYMV_NC_028142
SdYMV_BR_Deg3_10A
SiMBoV1_NC_015046
SiMAIV_NC_016573
ToLDV_EU710749
SiYMAV_NC_020255
SiYBV_NC_020254
SiCmMV_EU710751
SiMoAV_NC_020256
TrMV_BR_Fle1_10A
OMoV_NC_011181
HeYMV_BR_Mac7_10A
OxYVV_NC_026253
SiYNV_NC_020253
SimMV_NC_005330
SiMoV_NC_004637
ToYSV_NC_007726
SiYMV_NC_004639
SimBV_FN436001
ToMMV_NC_010833
TrYMV_BR_Msj1_10A
Cluster III
MacGMV_NC_010952
BChV_NC_019569
ToClLDV_NC_015962
TMoLCV_NC_015122
BYMMxV_NC_015487
SiGMFlV_NC_014446
SiGYVV_NC_004635
ToMoV_NC_001938
SiGMV_NC_002046
AbMV_NC_001928
SiYMoV_NC_016082
RhRGMV_HM236370
ToYLDV_NC_011348
JaMV_NC_024012
WGMSTV_NC_010948
ToLCCUV_AM050143
ToMoTaV_NC_001828
ToMHaV_NC_003867
CoYSV_NC_008492
ToLCSiV_NC_009606
SiYMYuV_NC_008779
SiGMHV_NC_004659
SiGMBuV_NC_014794
CdTV_NC_003830
SiGMoV_NC_014130
OYMV_NC_014066
OYMoIV_AY751753
SiYVV_NC_004661
DaLDV_NC_018717
BChMV_NC_022005
BDMV_NC_001931
Cluster I
SiGMCRV_NC_004657
SiMSiV_NC_008059
PYMTV_NC_004638
PYMPV_NC_002048
PYMV_NC_001934

Figure 1b

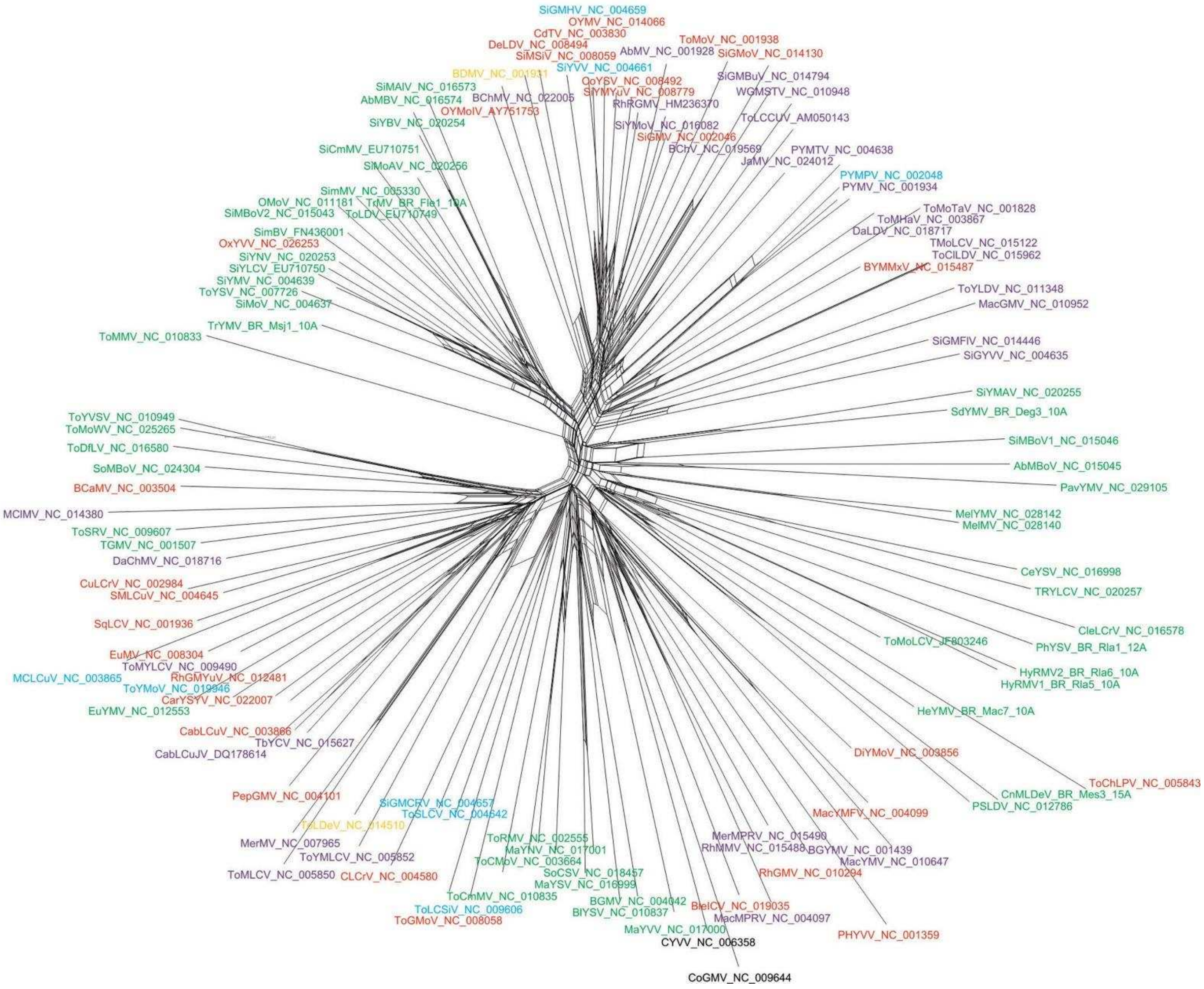
SiGMHV_NC_004659
OYMV_NC_014066
CdTV_NC_003830
DeLDV_NC_008494
SiMSiV_NC_008059
AbMV_NC_001928
ToMoV_NC_001938
SiGMoV_NC_014130
BDMV_NC_001931
SiYVV_NC_004661
SiGMBuV_NC_014794
SiMAlV_NC_016573
QoYSV_NC_008492
SiYMYuV_NC_008779
WGMSTV_NC_010948
AbMBV_NC_016574
BChMV_NC_022005
RhRGMV_HM236370
OYMolV_AY751753
ToLCCUV_AM050143
SiYBV_NC_020254
SiYMoV_NC_016082
SiGMV_NC_002046
BChV_NC_019569
PYMTV_NC_004638
SiCmMV_EU710751
JaMV_NC_024012
SiMoAV_NC_020256
PYMPV_NC_002048
PYMV_NC_001934
SimMV_NC_005330
OMoV_NC_011181
TrMV_BR_Fle1_10A
SiMBoV2_NC_015043
ToLDV_EU710749
ToMoTaV_NC_001828
ToMHaV_NC_003867
SimBV_FN436001
DaLDV_NC_018717
OxYVV_NC_026253
TMoLCV_NC_015122
ToClLDV_NC_015962
SiYNV_NC_020253
SiYLCV_EU710750
BYMMxV_NC_015487
SiYMV_NC_004639
ToYSV_NC_007726
ToYLDV_NC_011348
SiMoV_NC_004637
MacGMV_NC_010952
TrYMV_BR_Msj1_10A
ToMMV_NC_010833
SiGMFlV_NC_014446
SiGYVV_NC_004635
SiYMAV_NC_020255
SdYMV_BR_Deg3_10A
ToYVSV_NC_010949
ToMoWV_NC_025265
SiMBoV1_NC_015046
ToDfLV_NC_016580
AbMBoV_NC_015045
SoMBoV_NC_024304
PavYMV_NC_029105
BCaMV_NC_003504
MClMV_NC_014380
MelYMV_NC_028142
MelMV_NC_028140
ToSRV_NC_009607
TGMV_NC_001507
DaChMV_NC_018716
CeYSV_NC_016998
CuLCrV_NC_002984
TRYLCV_NC_020257
SMLCuV_NC_004645
SqLCV_NC_001936
CleLCrV_NC_016578
ToMoLCV_JF803246
PhYSV_BR_Rla1_12A
EuMV_NC_008304
ToMYLCV_NC_009490
HyRMV2_BR_Rla6_10A
MCLCuV_NC_003865
RhGMYuV_NC_012481
HyRMV1_BR_Rla5_10A
ToYMoV_NC_019946
CarYSYV_NC_022007
EuYMV_NC_012553
HeYMV_BR_Mac7_10A
CabLCuV_NC_003866
TbYCV_NC_015627
CabLCuJV_DQ178614
DiYMoV_NC_003856
ToChLPV_NC_005843
PepGMV_NC_004101
CnMLDeV_BR_Mes3_15A
SiGMCRV_NC_004657
PSLDV_NC_012786
ToSLCV_NC_004642
MacYMFV_NC_004099
ToLDeV_NC_014510
ToRMV_NC_002555
MerMPRV_NC_015490
MerMV_NC_007965
MaYNV_NC_017001
RhMMV_NC_015488
BGYMV_NC_001439
ToYMLCV_NC_005852
ToCMoV_NC_003664
MacYMV_NC_010647
ToMLCV_NC_005850
CLCrV_NC_004580
SoCSV_NC_018457
RhGMV_NC_010294
MaYSV_NC_016999
ToCmMV_NC_010835
BGMV_NC_004042
BleICV_NC_019035
ToLCSiV_NC_009606
BIYSV_NC_010837
MacMPRV_NC_004097
ToGMoV_NC_008058
MaYVV_NC_017000
PHYVV_NC_001359
CYVV_NC_006358
CoGMV_NC_009644

Figure 1c

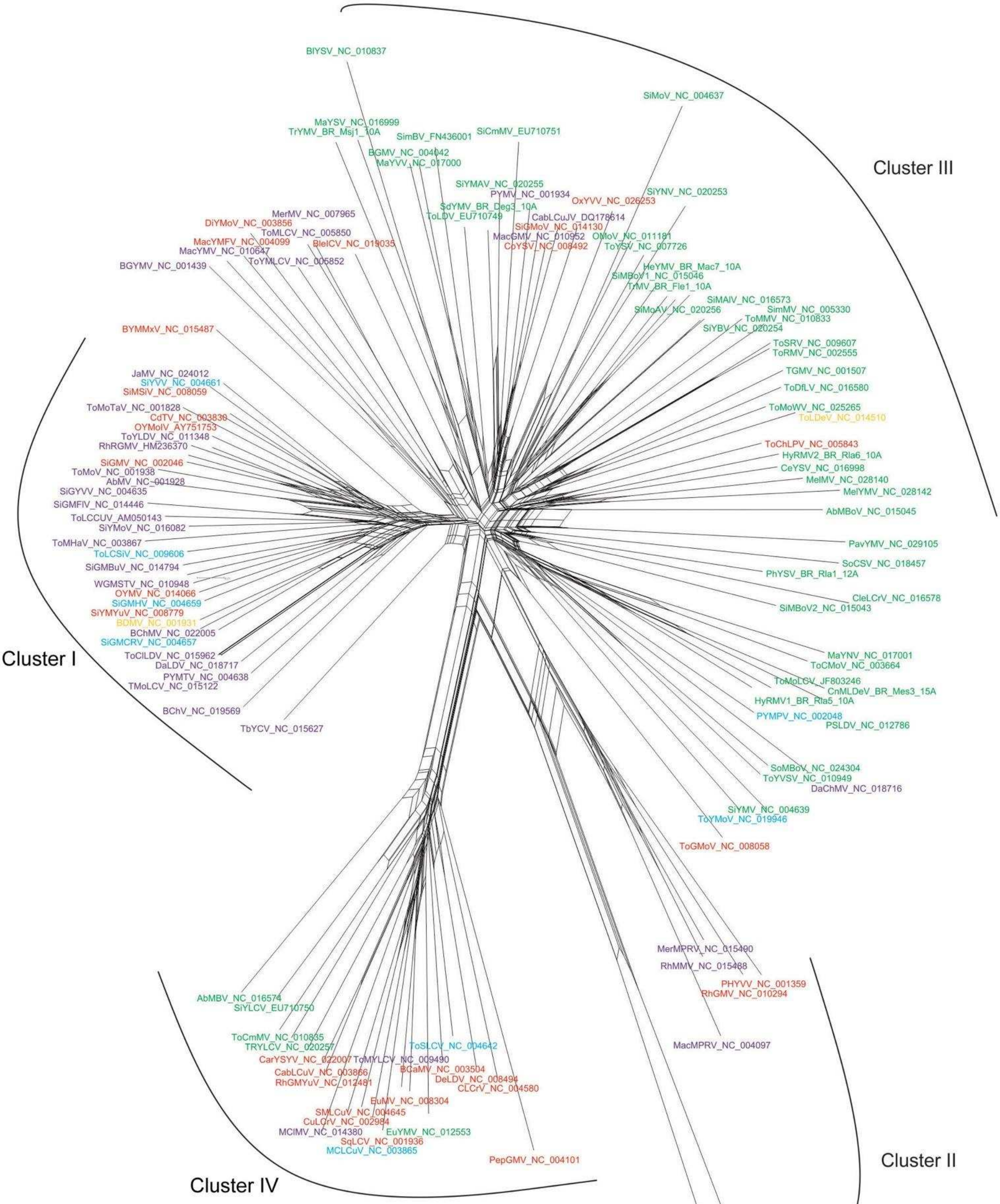
Cluster I
Cluster II
Cluster III
Cluster IV
BlYSV_NC_010837
SiMoV_NC_004637
MaYSV_NC_016999
TrYMV_BR_Msj1_10A
SimBV_FN436001
SiCmMV_EU710751
BGMV_NC_004042
MaYVV_NC_017000
SiYMAV_NC_020255
PYMV_NC_001934
OxYVV_NC_026253
SiYNV_NC_020253
SdYMV_BR_Deg3_10A
ToLDV_EU710749
CabLCuJV_DQ178614
MerMV_NC_007965
DiYMoV_NC_003856
ToMLCV_NC_005850
MacYMFV_NC_004099
BleICV_NC_019035
MacYMV_NC_010647
ToYMLCV_NC_005852
BGYMV_NC_001439
SiGMoV_NC_014130
MacGMV_NC_010952
OMoV_NC_011181
CoYSV_NC_008492
ToYSV_NC_007726
HeYMV_BR_Mac7_10A
SiMBoV1_NC_015046
TrMV_BR_Fle1_10A
SiMAIV_NC_016573
SiMoAV_NC_020256
SimMV_NC_005330
ToMMV_NC_010833
SiYBV_NC_020254
BYMMxV_NC_015487
ToSRV_NC_009607
ToRMV_NC_002555
TGMV_NC_001507
ToDfLV_NC_016580
ToMoWV_NC_025265
ToLDeV_NC_014510
JaMV_NC_024012
SiYVV_NC_004661
SiMSiV_NC_008059
ToMoTaV_NC_001828
CdTV_NC_003830
OYMolV_AY751753
ToYLDV_NC_011348
RhRGMV_HM236370
SiGMV_NC_002046
ToMoV_NC_001938
AbMV_NC_001928
SiGYVV_NC_004635
SiGMFlV_NC_014446
ToLCCUV_AM050143
SiYMoV_NC_016082
ToMHaV_NC_003867
ToLCSiV_NC_009606
SiGMBuV_NC_014794
WGMSTV_NC_010948
OYMV_NC_014066
SiGMHV_NC_004659
SiYMYuV_NC_008779
BDMV_NC_001931
BChMV_NC_022005
SiGMCRV_NC_004657
ToCILDV_NC_015962
DaLDV_NC_018717
PYMTV_NC_004638
TMoLCV_NC_015122
BChV_NC_019569
TbYCV_NC_015627
ToChLPV_NC_005843
HyRMV2_BR_Rla6_10A
CeYSV_NC_016998
MelMV_NC_028140
MelYMV_NC_028142
AbMBoV_NC_015045
PavYMV_NC_029105
SoCSV_NC_018457
PhYSV_BR_Rla1_12A
CleLCrV_NC_016578
SiMBoV2_NC_015043
MaYNV_NC_017001
ToCMoV_NC_003664
ToMoLCV_JF803246
CnMLDeV_BR_Mes3_15A
HyRMV1_BR_Rla5_10A
PYMPV_NC_002048
PSLDV_NC_012786
SoMBoV_NC_024304
ToYVSV_NC_010949
DaChMV_NC_018716
SiYMV_NC_004639
ToYMoV_NC_019946
ToGMoV_NC_008058
MerMPRV_NC_015490
RhMMV_NC_015488
PHYVV_NC_001359
RhGMV_NC_010294
MacMPRV_NC_004097
AbMBV_NC_016574
SiYLCV_EU710750
ToCmMV_NC_010835
TRYLCV_NC_020257
ToSLCV_NC_004642
CarYSYV_NC_022007
ToMYLCV_NC_009490
BCaMV_NC_003504
CabLCuV_NC_003866
DeLDV_NC_008494
RhGMYuV_NC_012481
CLCrV_NC_004580
EuMV_NC_008304
SMLCuV_NC_004645
CuLCrV_NC_002984
MCIMV_NC_014380
EuYMV_NC_012553
SqLCV_NC_001936
MCLCuV_NC_003865
PepGMV_NC_004101
CoGMV_NC_009644
CYVV_NC_006358

Figure 1d

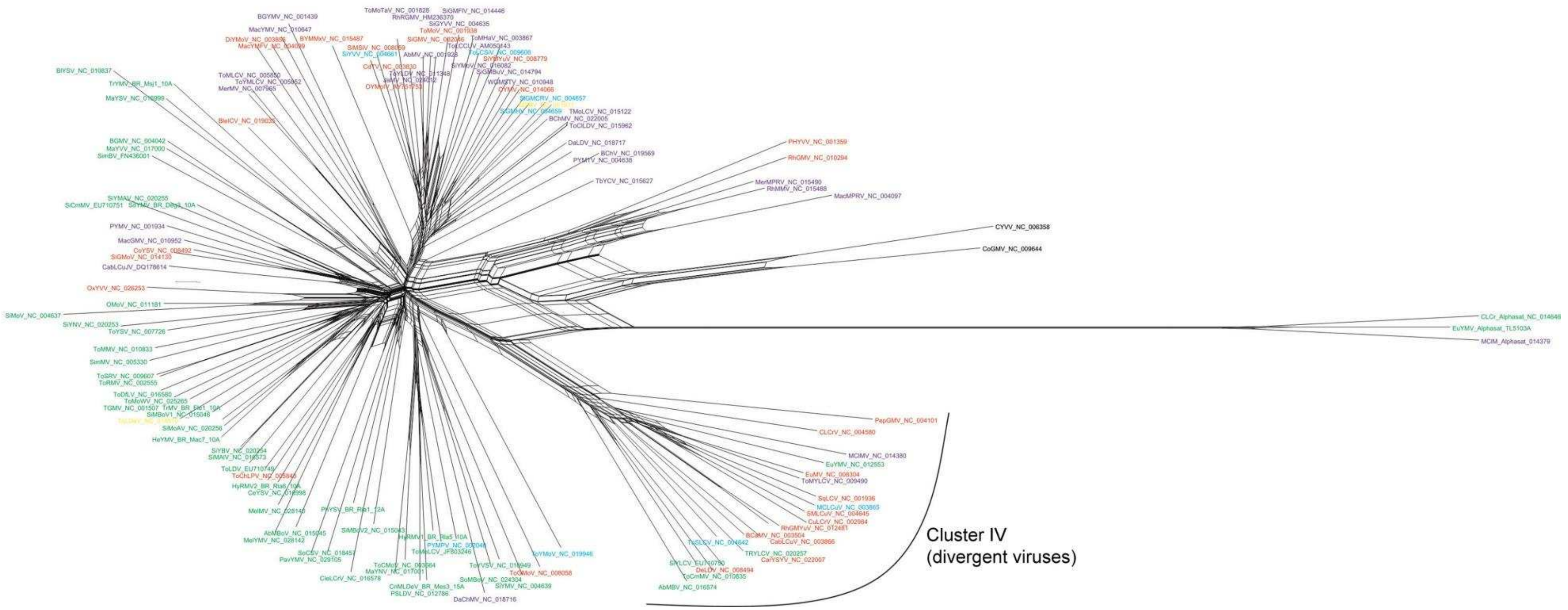
BGYMV_NC_001439
MacYMV_NC_010647
DiYMoV_NC_003856
MacYMFV_NC_004099
BYMMxV_NC_015487
ToMoTaV_NC_001828
SiGMFIV_NC_014446
RhRGMV_HM236370
SiGYVV_NC_004635
ToMoV_NC_001938
SiGMV_NC_002046
ToMHaV_NC_003867
SiMSIV_NC_008059
ToLCCUV_AM050143
SiYVV_NC_004661
AbMV_NC_001928
ToLCSiV_NC_009606
SiYMYuV_NC_008779
CdTV_NC_003830
SiYMoV_NC_016082
ToYLDV_NC_011348
SiGMBuV_NC_014794
JaMV_NC_024012
OYMoIV_NC_014066
WGMSTV_NC_010948
OYMV_NC_014066
SiGMCRV_NC_004657
SiGMHV_NC_004659
TMoLCV_NC_015122
BChMV_NC_022005
ToClLDV_NC_015962
DaLDV_NC_018717
BChV_NC_019569
PYMTV_NC_004638
TbYCV_NC_015627
ToMLCV_NC_005850
ToYMLCV_NC_005852
MerMV_NC_007965
BIYSV_NC_010837
TrYMV_BR_Msj1_10A
MaYSV_NC_016999
BleICV_NC_019035
BGMV_NC_004042
MaYVV_NC_017000
SimBV_FN436001
SiYMAV_NC_020255
SiCmMV_EU710751
SiYMV_BR_Deg2_10A
PYMV_NC_001934
MacGMV_NC_010952
CoYSV_NC_008492
SiGMoV_NC_014130
CabLCuJV_DQ178614
OxYVV_NC_026253
OMoV_NC_011181
SiMoV_NC_004637
SiYNV_NC_020253
ToYSV_NC_007726
ToMMV_NC_010833
SimMV_NC_005330
ToSRV_NC_009607
ToRMV_NC_002555
ToDfLV_NC_016580
ToMoWV_NC_025265
TGMV_NC_001507
TrMV_BR_Rie1_10A
SiMBoV1_NC_015046
SiMoAV_NC_020256
HeYMV_BR_Mac7_10A
SiYBV_NC_020254
SiMAIV_NC_016573
ToLDV_EU710749
ToChLPV_NC_005840
HyRMV2_BR_Ria6_10A
CeYSV_NC_016998
MelMV_NC_028140
PhYSV_BR_Ria1_12A
AbMBoV_NC_015045
MelYMV_NC_028142
SiMBoV2_NC_015043
SoCSV_NC_018457
PavYMV_NC_029105
HyRMV1_BR_Ria5_10A
PYMPV_NC_002048
ToMeLCV_JF803246
ClaLCrV_NC_016578
ToCMoV_NC_003664
MaYNV_NC_017001
CnMLDeV_BR_Mes3_15A
PSLDV_NC_012786
ToYVSV_NC_016949
ToGMoV_NC_008058
SoMBoV_NC_024304
SiYMV_NC_004639
DaChMV_NC_018716
ToYMoV_NC_019946
PHYVV_NC_001359
RhGMV_NC_010294
MerMPRV_NC_015490
RhMMV_NC_015488
MacMPRV_NC_004097
CYVV_NC_006358
CoGMV_NC_009644
CLCr_Alphasat_NC_014646
EuYMV_Alphasat_TL5103A
MCIM_Alphasat_014379
PepGMV_NC_004101
CLCrV_NC_004580
MCIMV_NC_014380
EuYMV_NC_012553
EuMV_NC_008304
ToMYLCV_NC_009490
SqLCV_NC_001936
MCLCuV_NC_003865
SMLCuV_NC_004645
CuLCrV_NC_002984
RhGMYuV_NC_012481
BCaMV_NC_003504
CabLCuV_NC_003866
ToSLCV_NC_004642
TRYLCV_NC_020257
CarYSYV_NC_022007
SiYLCV_EU710750
DeLDV_NC_008494
ToCmMV_NC_010835
AbMBV_NC_016574
Cluster IV
(divergent viruses)

Figure 1e

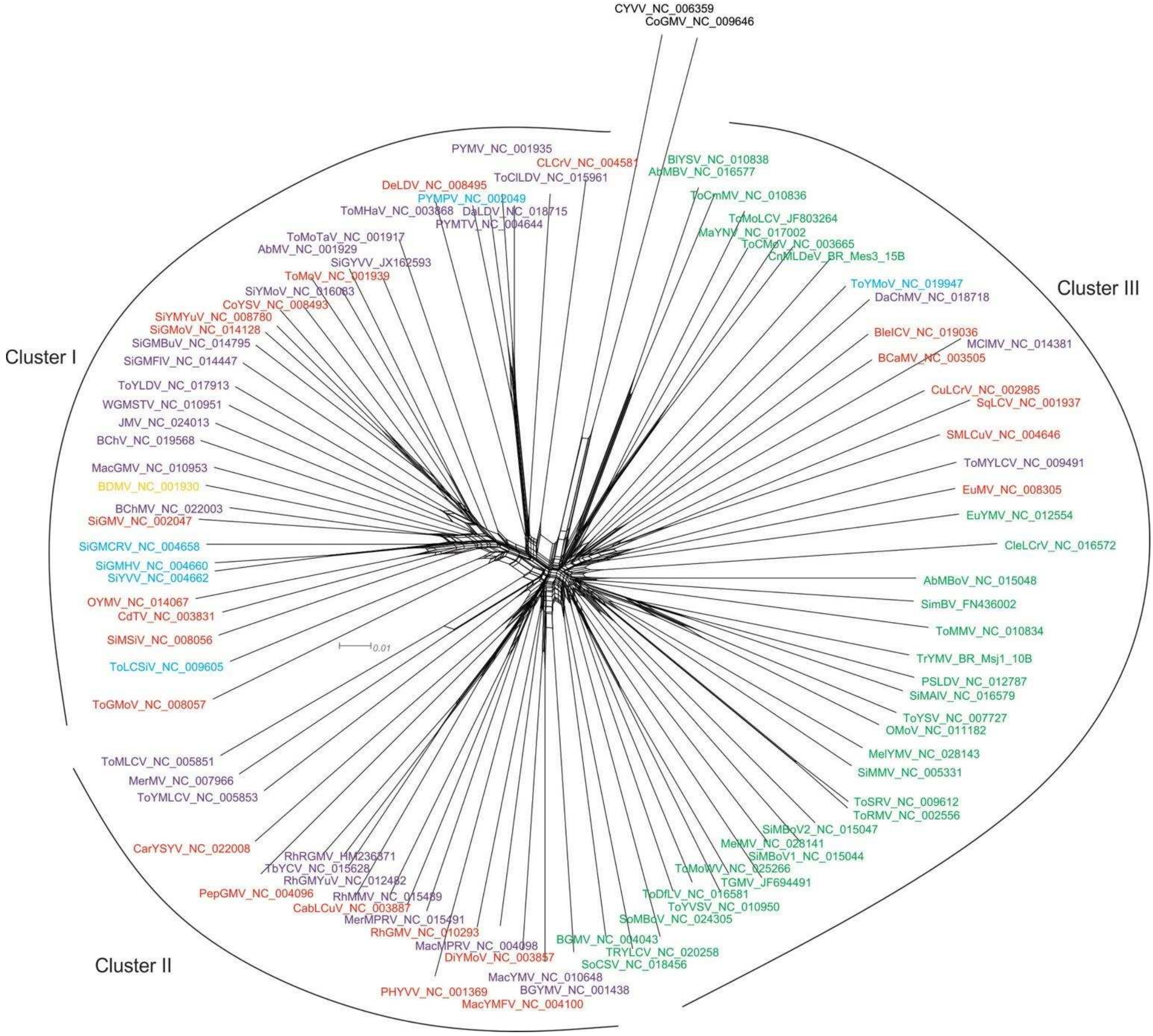
CYVV_NC_006359
CoGMV_NC_009646
PYMV_NC_001935
CLCrV_NC_004581
ToClLDV_NC_015961
DeLDV_NC_008495
PYMPV_NC_002049
ToMHaV_NC_003868
DaLDV_NC_018715
PYMTV_NC_004644
ToMoTaV_NC_001917
AbMV_NC_001929
SiGYVV_JX162593
ToMoV_NC_001939
SiYMoV_NC_016083
CoYSV_NC_008493
SiYMYuV_NC_008780
SiGMoV_NC_014128
SiGMBuV_NC_014795
Cluster I
SiGMFIV_NC_014447
ToYLDV_NC_017913
WGMSTV_NC_010951
JMV_NC_024013
BChV_NC_019568
MacGMV_NC_010953
BDMV_NC_001930
BChMV_NC_022003
SiGMV_NC_002047
SiGMCRV_NC_004658
SiGMHV_NC_004660
SiYVV_NC_004662
OYMV_NC_014067
CdTV_NC_003831
SiMSiV_NC_008056
0.01
ToLCSiV_NC_009605
ToGMoV_NC_008057
ToMLCV_NC_005851
MerMV_NC_007966
ToYMLCV_NC_005853
CarYSYV_NC_022008
RhRGMV_HM236371
TbYCV_NC_015628
RhGMYuV_NC_012482
PepGMV_NC_004096
RhMMV_NC_015489
CabLCuV_NC_003887
MerMPRV_NC_015491
RhGMV_NC_010293
MacMPRV_NC_004098
DiYMoV_NC_003857
Cluster II
MacYMV_NC_010648
BGYMV_NC_001438
PHYVV_NC_001369
MacYMFV_NC_004100
BGMV_NC_004043
TRYLCV_NC_020258
SoCSV_NC_018456
SoMBoV_NC_024305
ToYVSV_NC_010950
ToDfLV_NC_016581
TGMV_JF694491
ToMoWV_NC_025266
SiMBoV1_NC_015044
MelMV_NC_028141
SiMBoV2_NC_015047
ToRMV_NC_002556
ToSRV_NC_009612
SiMMV_NC_005331
MelYMV_NC_028143
OMoV_NC_011182
ToYSV_NC_007727
SiMAlV_NC_016579
PSLDV_NC_012787
TrYMV_BR_Msj1_10B
ToMMV_NC_010834
SimBV_FN436002
AbMBoV_NC_015048
CleLCrV_NC_016572
EuYMV_NC_012554
EuMV_NC_008305
ToMYLCV_NC_009491
SMLCuV_NC_004646
SqLCV_NC_001937
CuLCrV_NC_002985
BCaMV_NC_003505
MClMV_NC_014381
BleICV_NC_019036
DaChMV_NC_018718
ToYMoV_NC_019947
Cluster III
CnMLDeV_BR_Mes3_15B
ToCMoV_NC_003665
MaYNV_NC_017002
ToMoLCV_JF803264
ToCmMV_NC_010836
AbMBV_NC_016577
BlYSV_NC_010838

Figure 1f

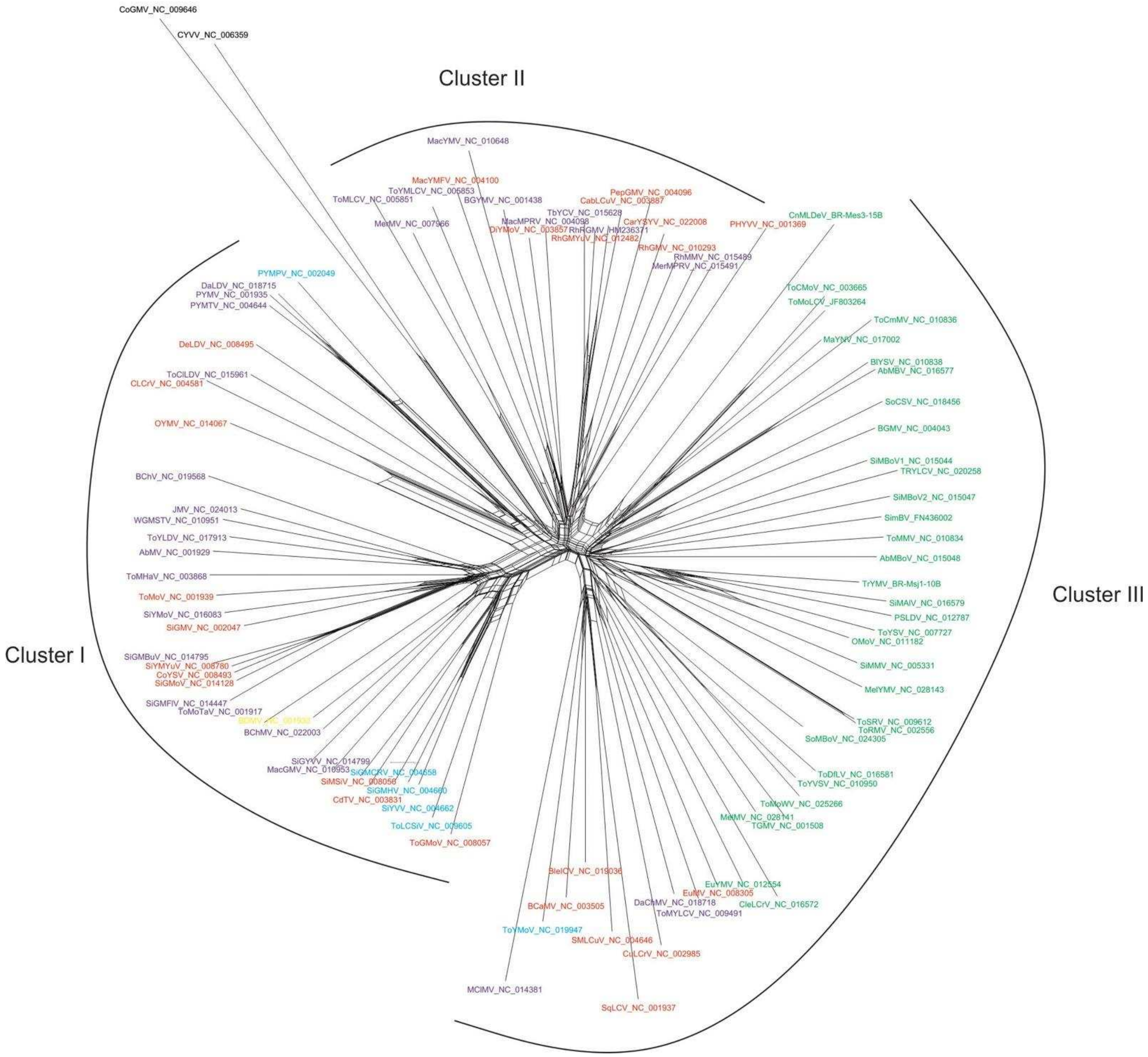
CoGMV_NC_009646
CYVV_NC_006359
Cluster II
MacYMV_NC_010648
MacYMFV_NC_004100
ToYMLCV_NC_005853
ToMLCV_NC_005851
BGYMV_NC_001438
PepGMV_NC_004096
CabLCuV_NC_003887
TbYCV_NC_015628
MacMPRV_NC_004098
MerMV_NC_007966
DiYMoV_NC_003857
RhRGMV_HM236371
CarYSYV_NC_022008
PHYVV_NC_001369
CnMLDeV_BR-Mes3-15B
RhGMYuV_NC_012482
RhGMV_NC_010293
RhMMV_NC_015489
MerMPRV_NC_015491
PYMPV_NC_002049
DaLDV_NC_018715
PYMV_NC_001935
PYMTV_NC_004644
DeLDV_NC_008495
ToClLDV_NC_015961
CLCrV_NC_004581
OYMV_NC_014067
BChV_NC_019568
JMV_NC_024013
WGMSTV_NC_010951
ToYLDV_NC_017913
AbMV_NC_001929
ToMHaV_NC_003868
ToMoV_NC_001939
SiYMoV_NC_016083
SiGMV_NC_002047
Cluster I
SiGMBuV_NC_014795
SiYMYuV_NC_008780
CoYSV_NC_008493
SiGMoV_NC_014128
SiGMFlV_NC_014447
ToMoTaV_NC_001917
BDMV_NC_001932
BChMV_NC_022003
SiGYVV_NC_014799
MacGMV_NC_010953
SiGMCRV_NC_004658
SiMSiV_NC_008056
SiGMHV_NC_004660
CdTV_NC_003831
SiYVV_NC_004662
ToLCSiV_NC_009605
ToGMoV_NC_008057
ToCMoV_NC_003665
ToMoLCV_JF803264
ToCmMV_NC_010836
MaYNV_NC_017002
BlYSV_NC_010838
AbMBV_NC_016577
SoCSV_NC_018456
BGMV_NC_004043
SiMBoV1_NC_015044
TRYLCV_NC_020258
SiMBoV2_NC_015047
SimBV_FN436002
ToMMV_NC_010834
AbMBoV_NC_015048
TrYMV_BR-Msj1-10B
Cluster III
SiMAlV_NC_016579
PSLDV_NC_012787
ToYSV_NC_007727
OMoV_NC_011182
SiMMV_NC_005331
MelYMV_NC_028143
ToSRV_NC_009612
ToRMV_NC_002556
SoMBoV_NC_024305
ToDfLV_NC_016581
ToYVSV_NC_010950
ToMoWV_NC_025266
MelMV_NC_028141
TGMV_NC_001508
BleICV_NC_019036
EuYMV_NC_012554
EuMV_NC_008305
DaChMV_NC_018718
CleLCrV_NC_016572
BCaMV_NC_003505
ToMYLCV_NC_009491
ToYMoV_NC_019947
SMLCuV_NC_004646
CuLCrV_NC_002985
MClMV_NC_014381
SqLCV_NC_001937

Figure 1g

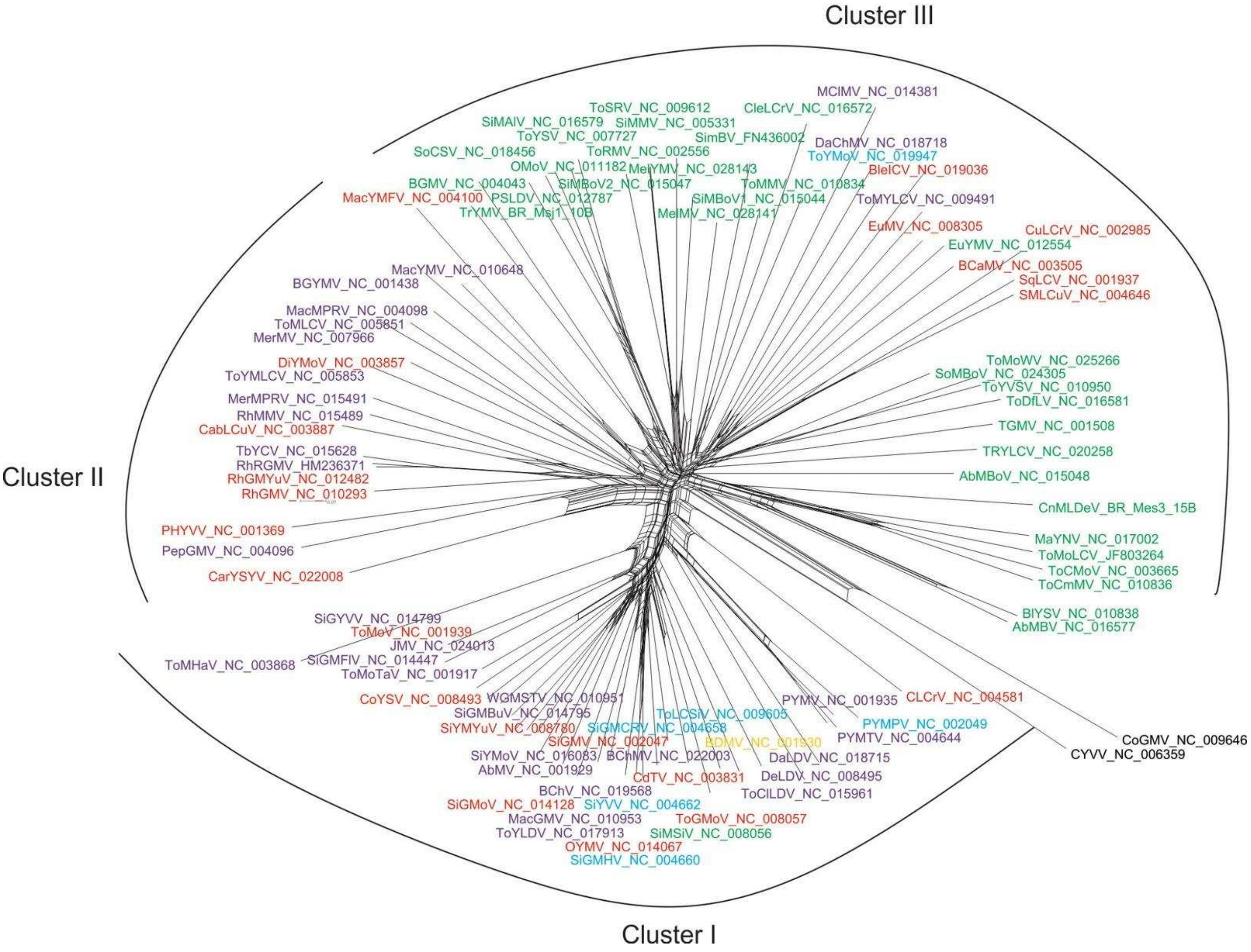
Cluster III
MClMV_NC_014381
ToSRV_NC_009612
CleLCrV_NC_016572
SiMAIV_NC_016579
SiMMV_NC_005331
ToYSV_NC_007727
SimBV_FN436002
DaChMV_NC_018718
SoCSV_NC_018456
ToRMV_NC_002556
ToYMoV_NC_019947
OMoV_NC_011182
MelYMV_NC_028143
BleICV_NC_019036
BGMV_NC_004043
SiMBoV2_NC_015047
ToMMV_NC_010834
MacYMFV_NC_004100
PSLDV_NC_012787
SiMBoV1_NC_015044
ToMYLCV_NC_009491
TrYMV_BR_Msj1_10B
MelMV_NC_028141
EuMV_NC_008305
CuLCrV_NC_002985
EuYMV_NC_012554
MacYMV_NC_010648
BCaMV_NC_003505
BGYMV_NC_001438
SqLCV_NC_001937
SMLCuV_NC_004646
MacMPRV_NC_004098
ToMLCV_NC_005851
MerMV_NC_007966
ToMoWV_NC_025266
DiYMoV_NC_003857
SoMBoV_NC_024305
ToYMLCV_NC_005853
ToYVSV_NC_010950
ToDfLV_NC_016581
MerMPRV_NC_015491
RhMMV_NC_015489
TGMV_NC_001508
CabLCuV_NC_003887
TbYCV_NC_015628
TRYLCV_NC_020258
RhRGMV_HM236371
AbMBoV_NC_015048
Cluster II
RhGMYuV_NC_012482
RhGMV_NC_010293
CnMLDeV_BR_Mes3_15B
PHYVV_NC_001369
MaYNV_NC_017002
PepGMV_NC_004096
ToMoLCV_JF803264
ToCMoV_NC_003665
CarYSYV_NC_022008
ToCmMV_NC_010836
BIYSV_NC_010838
SiGYVV_NC_014799
AbMBV_NC_016577
ToMoV_NC_001939
JMV_NC_024013
ToMHaV_NC_003868
SiGMFIV_NC_014447
ToMoTaV_NC_001917
CoYSV_NC_008493
WGMSTV_NC_010951
PYMV_NC_001935
CLCrV_NC_004581
SiGMBuV_NC_014795
ToLCSiV_NC_009605
PYMPV_NC_002049
SiYMYuV_NC_008780
SiGMCRV_NC_004658
CoGMV_NC_009646
SiGMV_NC_002047
BDMV_NC_001930
PYMTV_NC_004644
CYVV_NC_006359
SiYMoV_NC_016083
BChMV_NC_022003
DaLDV_NC_018715
AbMV_NC_001929
CdTV_NC_003831
DeLDV_NC_008495
BChV_NC_019568
ToClLDV_NC_015961
SiGMoV_NC_014128
SiYVV_NC_004662
MacGMV_NC_010953
ToGMoV_NC_008057
ToYLDV_NC_017913
SiMSiV_NC_008056
OYMV_NC_014067
SiGMHV_NC_004660
Cluster I

Figure 2a

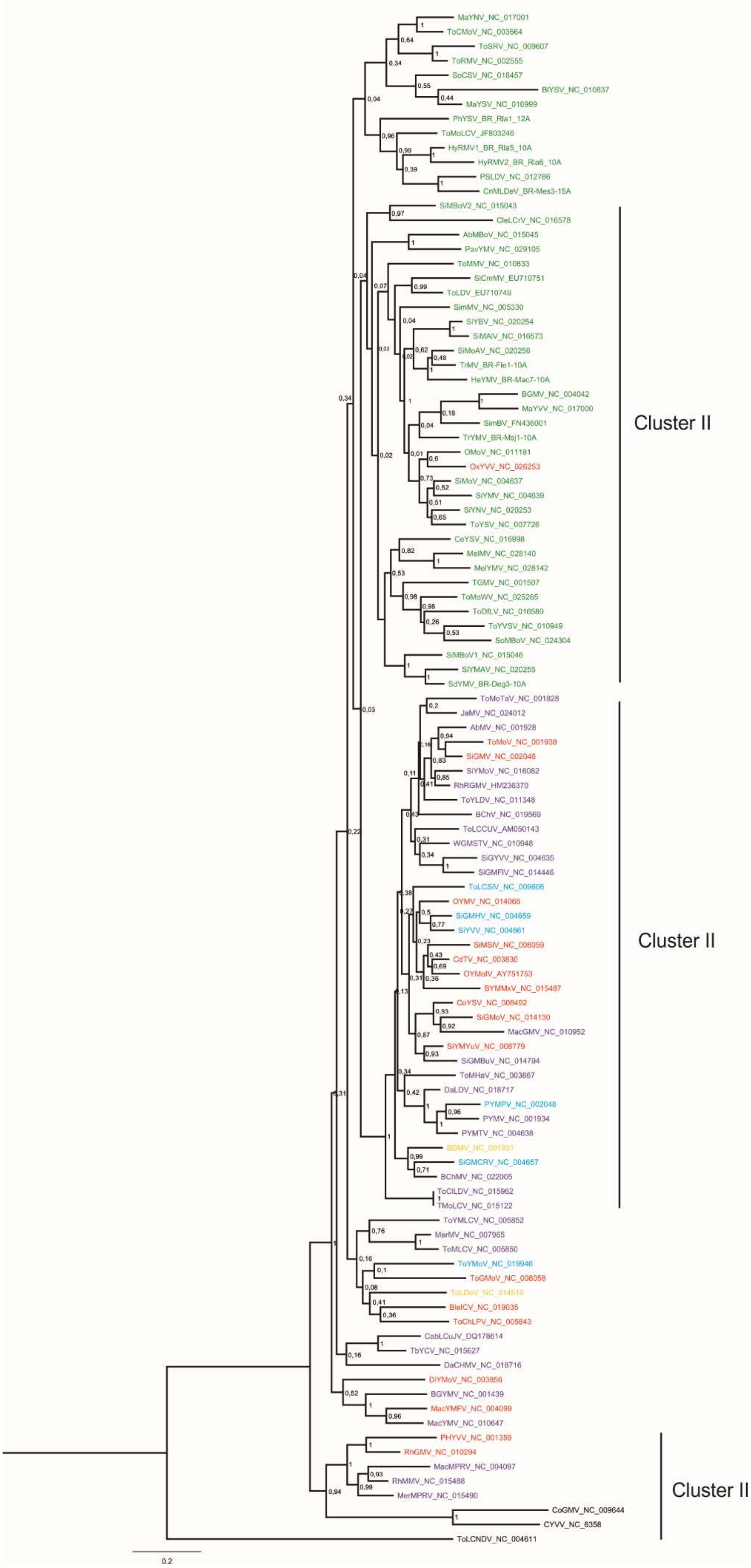
MaYNV_NC_017001
ToCMoV_NC_003664
ToSRV_NC_009607
ToRMV_NC_002555
SoCSV_NC_018457
BlYSV_NC_010837
MaYSV_NC_016999
PhYSV_BR_Rla1_12A
ToMoLCV_JF803246
HyRMV1_BR_Rla5_10A
HyRMV2_BR_Rla6_10A
PSLDV_NC_012786
CnMLDeV_BR-Mes3-15A
SiMBoV2_NC_015043
CleLCrV_NC_016578
AbMBoV_NC_015045
PavYMV_NC_029105
ToMMV_NC_010833
SiCmMV_EU710751
ToLDV_EU710749
SimMV_NC_005330
SiYBV_NC_020254
SiMAIV_NC_016573
SiMoAV_NC_020256
TrMV_BR-Fle1-10A
HeYMV_BR-Mac7-10A
BGMV_NC_004042
MaYVV_NC_017000
SimBV_FN436001
TrYMV_BR-Msj1-10A
OMoV_NC_011181
OxYVV_NC_026253
SiMoV_NC_004637
SiYMV_NC_004639
SiYNV_NC_020253
ToYSV_NC_007726
CeYSV_NC_016998
MelMV_NC_028140
MelYMV_NC_028142
TGMV_NC_001507
ToMoWV_NC_025265
ToDfLV_NC_016580
ToYVSV_NC_010949
SoMBoV_NC_024304
SiMBoV1_NC_015046
SiYMAV_NC_020255
SdYMV_BR-Deg3-10A
Cluster II
ToMoTaV_NC_001828
JaMV_NC_024012
AbMV_NC_001928
ToMoV_NC_001938
SiGMV_NC_002046
SiYMoV_NC_016082
RhRGMV_HM236370
ToYLDV_NC_011348
BChV_NC_019569
ToLCCUV_AM050143
WGMSTV_NC_010948
SiGYVV_NC_004635
SiGMFIV_NC_014446
ToLCSiV_NC_009606
OYMV_NC_014066
SiGMHV_NC_004659
SiYVV_NC_004661
SiMSiV_NC_008059
CdTV_NC_003830
OYMolV_AY751753
BYMMxV_NC_015487
CoYSV_NC_008492
SiGMoV_NC_014130
MacGMV_NC_010952
SiYMYuV_NC_008779
SiGMBuV_NC_014794
ToMHaV_NC_003867
DaLDV_NC_018717
PYMPV_NC_002048
PYMV_NC_001934
PYMTV_NC_004638
BDMV_NC_001931
SiGMCRV_NC_004657
BChMV_NC_022005
ToClLDV_NC_015962
TMoLCV_NC_015122
Cluster II
ToYMLCV_NC_005852
MerMV_NC_007965
ToMLCV_NC_005850
ToYMoV_NC_019946
ToGMoV_NC_008058
ToLDeV_NC_014510
BleICV_NC_019035
ToChLPV_NC_005843
CabLCuJV_DQ178614
TbYCV_NC_015627
DaCHMV_NC_018716
DiYMoV_NC_003856
BGYMV_NC_001439
MacYMFV_NC_004099
MacYMV_NC_010647
PHYVV_NC_001359
RhGMV_NC_010294
MacMPRV_NC_004097
RhMMV_NC_015488
MerMPRV_NC_015490
CoGMV_NC_009644
CYVV_NC_6358
ToLCNDV_NC_004611
Cluster II
0.2

Figure 2b

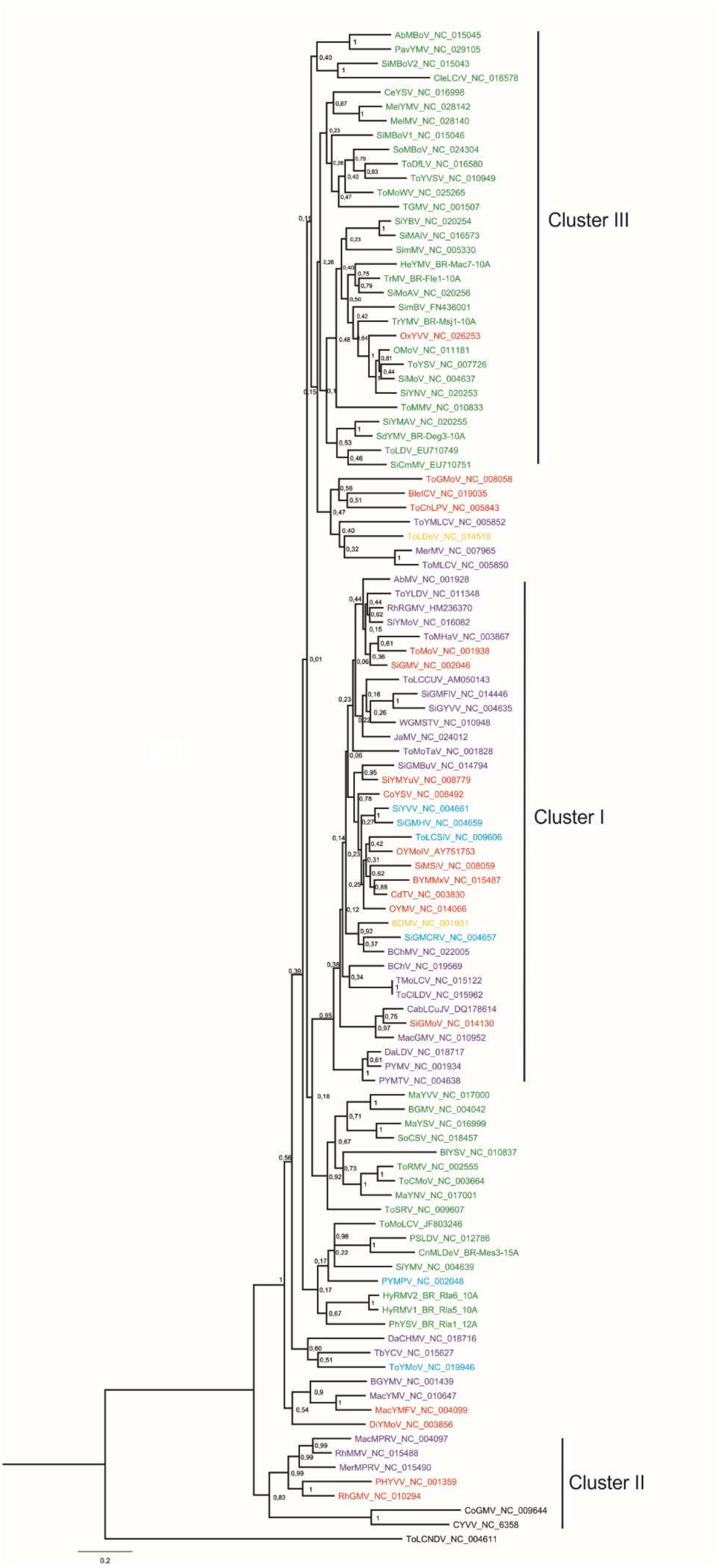
AbMBoV_NC_015045
PavYMV_NC_029105
SiMBoV2_NC_015043
CleLCrV_NC_016578
CeYSV_NC_016998
MelYMV_NC_028142
MelMV_NC_028140
SiMBoV1_NC_015046
SoMBoV_NC_024304
ToDfLV_NC_016580
ToYVSV_NC_010949
ToMoWV_NC_025265
TGMV_NC_001507
SiYBV_NC_020254
SiMAlV_NC_016573
SimMV_NC_005330
HeYMV_BR-Mac7-10A
TrMV_BR-Fle1-10A
SiMoAV_NC_020256
SimBV_FN436001
TrYMV_BR-Msj1-10A
OxYVV_NC_026253
OMoV_NC_011181
ToYSV_NC_007726
SiMoV_NC_004637
SiYNV_NC_020253
ToMMV_NC_010833
SiYMAV_NC_020255
SdYMV_BR-Deg3-10A
ToLDV_EU710749
SiCmMV_EU710751
Cluster III
ToGMoV_NC_008058
BleICV_NC_019035
ToChLPV_NC_005843
ToYMLCV_NC_005852
ToLDeV_NC_014510
MerMV_NC_007965
ToMLCV_NC_005850
AbMV_NC_001928
ToYLDV_NC_011348
RhRGMV_HM236370
SiYMoV_NC_016082
ToMHaV_NC_003867
ToMoV_NC_001938
SiGMV_NC_002046
ToLCCUV_AM050143
SiGMFlV_NC_014446
SiGYVV_NC_004635
WGMSTV_NC_010948
JaMV_NC_024012
ToMoTaV_NC_001828
SiGMBuV_NC_014794
SiYMYuV_NC_008779
CoYSV_NC_008492
SiYVV_NC_004661
SiGMHV_NC_004659
ToLCSiV_NC_009606
OYMolV_AY751753
SiMSiV_NC_008059
BYMMxV_NC_015487
CdTV_NC_003830
OYMV_NC_014066
BDMV_NC_001931
SiGMCRV_NC_004657
BChMV_NC_022005
BChV_NC_019569
TMoLCV_NC_015122
ToClLDV_NC_015962
CabLCuJV_DQ178614
SiGMoV_NC_014130
MacGMV_NC_010952
DaLDV_NC_018717
PYMV_NC_001934
PYMTV_NC_004638
Cluster I
MaYVV_NC_017000
BGMV_NC_004042
MaYSV_NC_016999
SoCSV_NC_018457
BlYSV_NC_010837
ToRMV_NC_002555
ToCMoV_NC_003664
MaYNV_NC_017001
ToSRV_NC_009607
ToMoLCV_JF803246
PSLDV_NC_012786
CnMLDeV_BR-Mes3-15A
SiYMV_NC_004639
PYMPV_NC_002048
HyRMV2_BR_Rla6_10A
HyRMV1_BR_Rla5_10A
PhYSV_BR_Rla1_12A
DaCHMV_NC_018716
TbYCV_NC_015627
ToYMoV_NC_019946
BGYMV_NC_001439
MacYMV_NC_010647
MacYMFV_NC_004099
DiYMoV_NC_003856
MacMPRV_NC_004097
RhMMV_NC_015488
MerMPRV_NC_015490
PHYVV_NC_001359
RhGMV_NC_010294
CoGMV_NC_009644
CYVV_NC_6358
Cluster II
ToLCNDV_NC_004611
0.2

Figure 2c

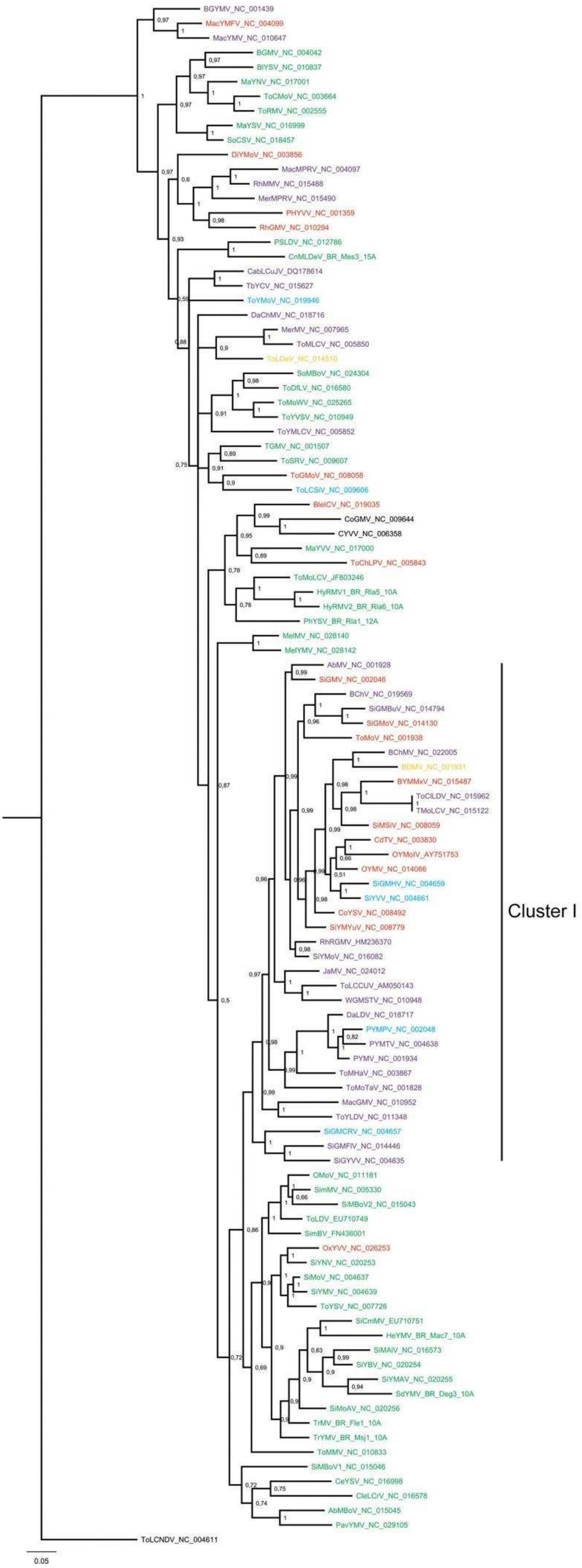
BGYMV_NC_001439
MacYMFV_NC_004099
MacYMV_NC_010647
BGMV_NC_004042
BlYSV_NC_010837
MaYNV_NC_017001
ToCMoV_NC_003664
ToRMV_NC_002555
MaYSV_NC_016999
SoCSV_NC_018457
DiYMoV_NC_003856
MacMPRV_NC_004097
RhMMV_NC_015488
MerMPRV_NC_015490
PHYVV_NC_001359
RhGMV_NC_010294
PSLDV_NC_012786
CnMLDeV_BR_Mes3_15A
CabLCuJV_DQ178614
TbYCV_NC_015627
ToYMoV_NC_019946
DaChMV_NC_018716
MerMV_NC_007965
ToMLCV_NC_005850
ToLDeV_NC_014510
SoMBoV_NC_024304
ToDfLV_NC_016580
ToMoWV_NC_025265
ToYVSV_NC_010949
ToYMLCV_NC_005852
TGMV_NC_001507
ToSRV_NC_009607
ToGMoV_NC_008058
ToLCSiV_NC_009606
BleICV_NC_019035
CoGMV_NC_009644
CYVV_NC_006358
MaYVV_NC_017000
ToChLPV_NC_005843
ToMoLCV_JF803246
HyRMV1_BR_Rla5_10A
HyRMV2_BR_Rla6_10A
PhYSV_BR_Rla1_12A
MelMV_NC_028140
MelYMV_NC_028142
AbMV_NC_001928
SiGMV_NC_002046
BChV_NC_019569
SiGMBuV_NC_014794
SiGMoV_NC_014130
ToMoV_NC_001938
BChMV_NC_022005
BDMV_NC_001931
BYMMxV_NC_015487
ToClLDV_NC_015962
TMoLCV_NC_015122
SiMSiV_NC_008059
CdTV_NC_003830
OYMoIV_AY751753
OYMV_NC_014066
SiGMHV_NC_004659
SiYVV_NC_004661
CoYSV_NC_008492
SiYMYuV_NC_008779
RhRGMV_HM236370
SiYMoV_NC_016082
JaMV_NC_024012
ToLCCUV_AM050143
WGMSTV_NC_010948
DaLDV_NC_018717
PYMPV_NC_002048
PYMTV_NC_004638
PYMV_NC_001934
ToMHaV_NC_003867
ToMoTaV_NC_001828
MacGMV_NC_010952
ToYLDV_NC_011348
SiGMCRV_NC_004657
SiGMFlV_NC_014446
SiGYVV_NC_004635
Cluster I
OMoV_NC_011181
SimMV_NC_005330
SiMBoV2_NC_015043
ToLDV_EU710749
SimBV_FN436001
OxYVV_NC_026253
SiYNV_NC_020253
SiMoV_NC_004637
SiYMV_NC_004639
ToYSV_NC_007726
SiCmMV_EU710751
HeYMV_BR_Mac7_10A
SiMAIV_NC_016573
SiYBV_NC_020254
SiYMAV_NC_020255
SdYMV_BR_Deg3_10A
SiMoAV_NC_020256
TrMV_BR_Fle1_10A
TrYMV_BR_Msj1_10A
ToMMV_NC_010833
SiMBoV1_NC_015046
CeYSV_NC_016998
CleLCrV_NC_016578
AbMBoV_NC_015045
PavYMV_NC_029105
ToLCNDV_NC_004611
0.05

Figure 2d

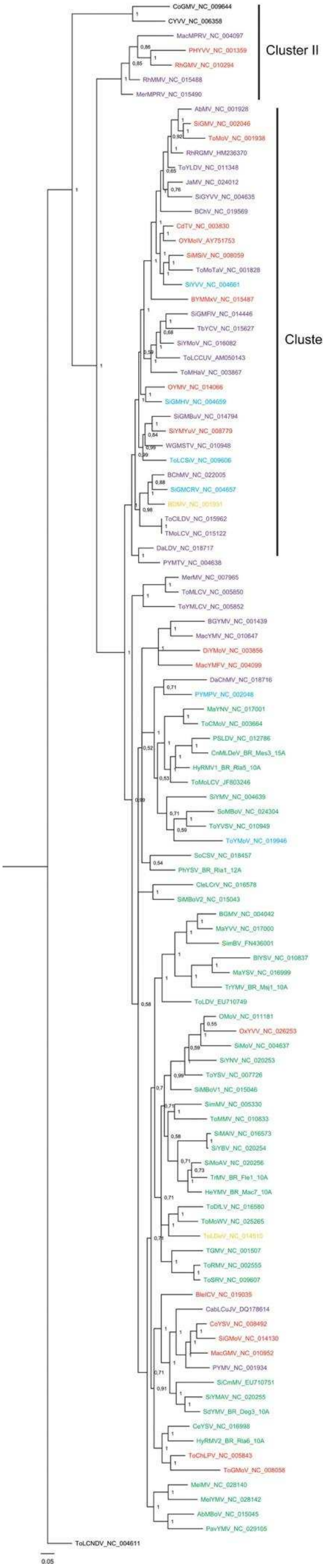
CoGMV_NC_009644
CYVV_NC_006358
MacMPRV_NC_004097
PHYVV_NC_001359
RhGMV_NC_010294
RhMMV_NC_015488
MerMPRV_NC_015490
Cluster II
AbMV_NC_001928
SiGMV_NC_002046
ToMoV_NC_001938
RhRGMV_HM236370
ToYLDV_NC_011348
JaMV_NC_024012
SiGYVV_NC_004635
BChV_NC_019569
CdTV_NC_003830
OYMoIV_AY751753
SiMSiV_NC_008059
ToMoTaV_NC_001828
SiYVV_NC_004661
BYMMxV_NC_015487
SiGMFlV_NC_014446
TbYCV_NC_015627
SiYMoV_NC_016082
ToLCCUV_AM050143
ToMHaV_NC_003867
OYMV_NC_014066
SiGMHV_NC_004659
SiGMBuV_NC_014794
SiYMYuV_NC_008779
WGMSTV_NC_010948
ToLCSiV_NC_009606
BChMV_NC_022005
SiGMCRV_NC_004657
BDMV_NC_001931
ToClLDV_NC_015962
TMoLCV_NC_015122
DaLDV_NC_018717
PYMTV_NC_004638
Cluster I
MerMV_NC_007965
ToMLCV_NC_005850
ToYMLCV_NC_005852
BGYMV_NC_001439
MacYMV_NC_010647
DiYMoV_NC_003856
MacYMFV_NC_004099
DaChMV_NC_018716
PYMPV_NC_002048
MaYNV_NC_017001
ToCMoV_NC_003664
PSLDV_NC_012786
CnMLDeV_BR_Mes3_15A
HyRMV1_BR_Ria5_10A
ToMoLCV_JF803246
SiYMV_NC_004639
SoMBoV_NC_024304
ToYVSV_NC_010949
ToYMoV_NC_019946
SoCSV_NC_018457
PhYSV_BR_Ria1_12A
CleLCrV_NC_016578
SiMBoV2_NC_015043
BGMV_NC_004042
MaYVV_NC_017000
SimBV_FN436001
BlYSV_NC_010837
MaYSV_NC_016999
TrYMV_BR_Maj1_10A
ToLDV_EU710749
OMoV_NC_011181
OxYVV_NC_026253
SiMoV_NC_004637
SiYNV_NC_020253
ToYSV_NC_007726
SiMBoV1_NC_015046
SimMV_NC_005330
ToMMV_NC_010833
SiMAIV_NC_016573
SiYBV_NC_020254
SiMoAV_NC_020256
TrMV_BR_Fle1_10A
HeYMV_BR_Mac7_10A
ToDfLV_NC_016580
ToMoWV_NC_025265
ToLDeV_NC_014510
TGMV_NC_001507
ToRMV_NC_002555
ToSRV_NC_009607
BleICV_NC_019035
CabLCuJV_DQ178614
CoYSV_NC_008492
SiGMoV_NC_014130
MacGMV_NC_010952
PYMV_NC_001934
SiCmMV_EU710751
SiYMAV_NC_020255
SdYMV_BR_Deg3_10A
CeYSV_NC_016998
HyRMV2_BR_Ria6_10A
ToChLPV_NC_005843
ToGMoV_NC_008058
MelMV_NC_028140
MelYMV_NC_028142
AbMBoV_NC_015045
PavYMV_NC_029105
ToLCNDV_NC_004611
0.05

Figure 3a

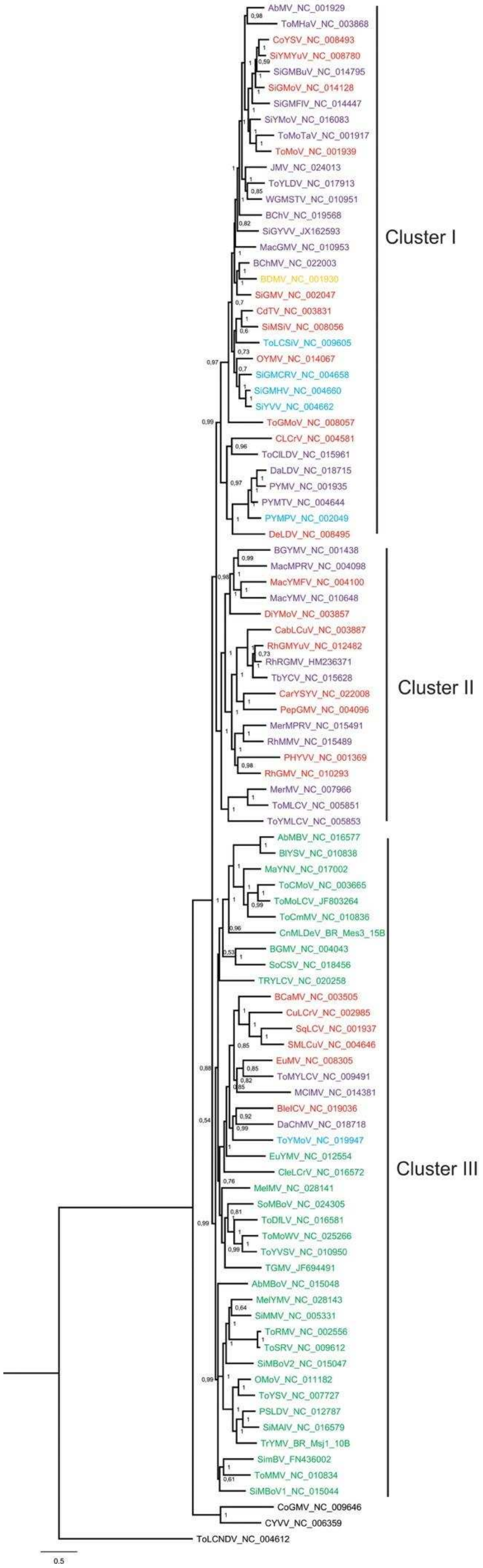
AbMV_NC_001929
ToMHaV_NC_003868
CoYSV_NC_008493
SiYMYuV_NC_008780
SiGMBuV_NC_014795
SiGMoV_NC_014128
SiGMFlV_NC_014447
SiYMoV_NC_016083
ToMoTaV_NC_001917
ToMoV_NC_001939
JMV_NC_024013
ToYLDV_NC_017913
WGMSTV_NC_010951
BChV_NC_019568
SiGYVV_JX162593
MacGMV_NC_010953
BChMV_NC_022003
BDMV_NC_001930
SiGMV_NC_002047
CdTV_NC_003831
SiMSiV_NC_008056
ToLCSiV_NC_009605
OYMV_NC_014067
SiGMCRV_NC_004658
SiGMHV_NC_004660
SiYVV_NC_004662
ToGMoV_NC_008057
CLCrV_NC_004581
ToClLDV_NC_015961
DaLDV_NC_018715
PYMV_NC_001935
PYMTV_NC_004644
PYMPV_NC_002049
DeLDV_NC_008495
Cluster I
BGYMV_NC_001438
MacMPRV_NC_004098
MacYMFV_NC_004100
MacYMV_NC_010648
DiYMoV_NC_003857
CabLCuV_NC_003887
RhGMYuV_NC_012482
RhRGMV_HM236371
TbYCV_NC_015628
CarYSYV_NC_022008
PepGMV_NC_004096
MerMPRV_NC_015491
RhMMV_NC_015489
PHYVV_NC_001369
RhGMV_NC_010293
MerMV_NC_007966
ToMLCV_NC_005851
ToYMLCV_NC_005853
Cluster II
AbMBV_NC_016577
BlYSV_NC_010838
MaYNV_NC_017002
ToCMoV_NC_003665
ToMoLCV_JF803264
ToCmMV_NC_010836
CnMLDeV_BR_Mes3_15B
BGMV_NC_004043
SoCSV_NC_018456
TRYLCV_NC_020258
BCaMV_NC_003505
CuLCrV_NC_002985
SqLCV_NC_001937
SMLCuV_NC_004646
EuMV_NC_008305
ToMYLCV_NC_009491
MClMV_NC_014381
BlelCV_NC_019036
DaChMV_NC_018718
ToYMoV_NC_019947
EuYMV_NC_012554
CleLCrV_NC_016572
MelMV_NC_028141
SoMBoV_NC_024305
ToDfLV_NC_016581
ToMoWV_NC_025266
ToYVSV_NC_010950
TGMV_JF694491
AbMBoV_NC_015048
MelYMV_NC_028143
SiMMV_NC_005331
ToRMV_NC_002556
ToSRV_NC_009612
SiMBoV2_NC_015047
OMoV_NC_011182
ToYSV_NC_007727
PSLDV_NC_012787
SiMAlV_NC_016579
TrYMV_BR_Msj1_10B
SimBV_FN436002
ToMMV_NC_010834
SiMBoV1_NC_015044
Cluster III
CoGMV_NC_009646
CYVV_NC_006359
ToLCNDV_NC_004612
0.5

Figure 3b

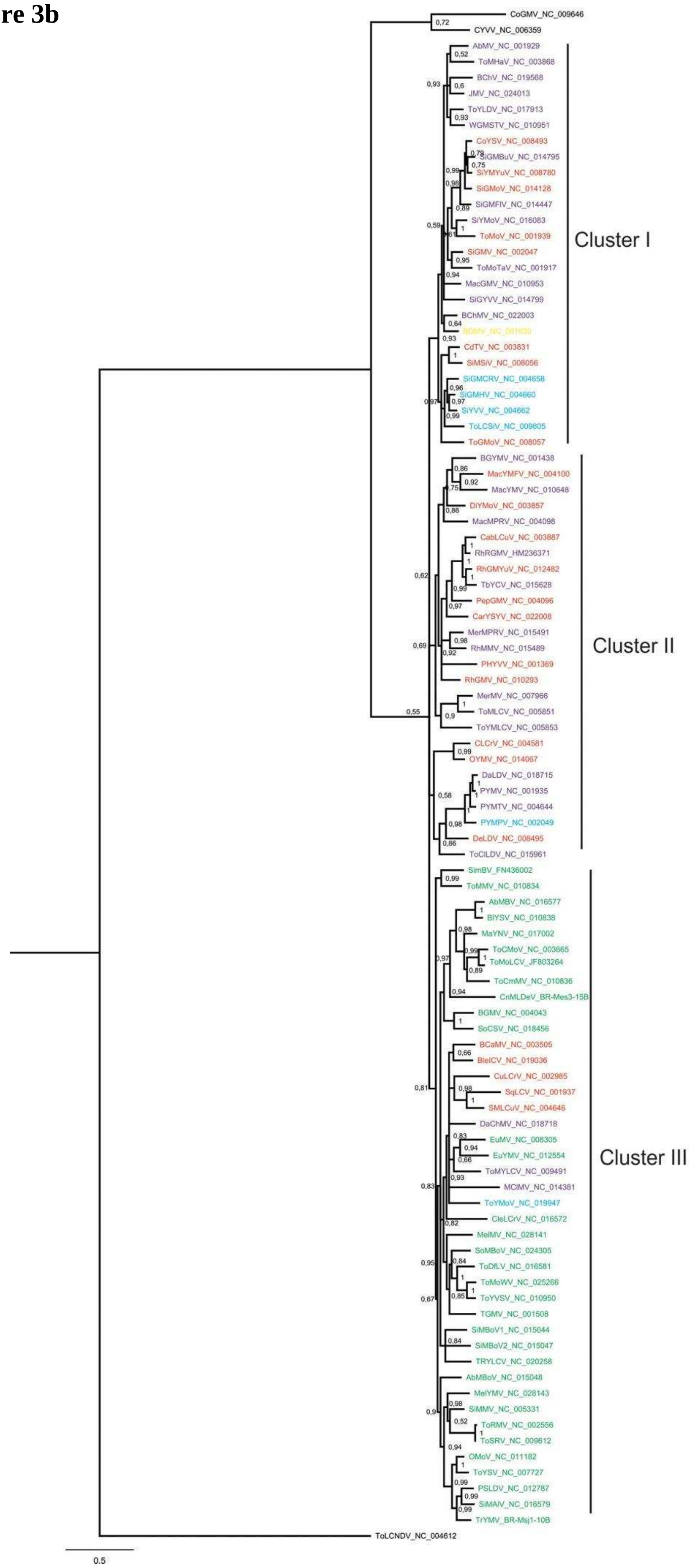
CoGMV_NC_009646
CYVV_NC_006359
AbMV_NC_001929
ToMHaV_NC_003868
BChV_NC_019568
JMV_NC_024013
ToYLDV_NC_017913
WGMSTV_NC_010951
CoYSV_NC_008493
SiGMBuV_NC_014795
SiYMYuV_NC_008780
SiGMoV_NC_014128
SiGMFIV_NC_014447
SiYMoV_NC_016083
ToMoV_NC_001939
SiGMV_NC_002047
ToMoTaV_NC_001917
MacGMV_NC_010953
SiGYVV_NC_014799
BChMV_NC_022003
BDMV_NC_001930
CdTV_NC_003831
SiMSiV_NC_008056
SiGMCRV_NC_004658
SiGMHV_NC_004660
SiYVV_NC_004662
ToLCSiV_NC_009605
ToGMoV_NC_008057
Cluster I
BGYMV_NC_001438
MacYMFV_NC_004100
MacYMV_NC_010648
DiYMoV_NC_003857
MacMPRV_NC_004098
CabLCuV_NC_003887
RhRGMV_HM236371
RhGMYuV_NC_012482
TbYCV_NC_015628
PepGMV_NC_004096
CarYSYV_NC_022008
MerMPRV_NC_015491
RhMMV_NC_015489
PHYVV_NC_001369
RhGMV_NC_010293
MerMV_NC_007966
ToMLCV_NC_005851
ToYMLCV_NC_005853
CLCrV_NC_004581
OYMV_NC_014067
DaLDV_NC_018715
PYMV_NC_001935
PYMTV_NC_004644
PYMPV_NC_002049
DeLDV_NC_008495
ToClLDV_NC_015961
Cluster II
SimBV_FN436002
ToMMV_NC_010834
AbMBV_NC_016577
BlYSV_NC_010838
MaYNV_NC_017002
ToCMoV_NC_003665
ToMoLCV_JF803264
ToCmMV_NC_010836
CnMLDeV_BR-Mes3-15B
BGMV_NC_004043
SoCSV_NC_018456
BCaMV_NC_003505
BleICV_NC_019036
CuLCrV_NC_002985
SqLCV_NC_001937
SMLCuV_NC_004646
DaChMV_NC_018718
EuMV_NC_008305
EuYMV_NC_012554
ToMYLCV_NC_009491
MClMV_NC_014381
ToYMoV_NC_019947
CleLCrV_NC_016572
MelMV_NC_028141
SoMBoV_NC_024305
ToDfLV_NC_016581
ToMoWV_NC_025266
ToYVSV_NC_010950
TGMV_NC_001508
SiMBoV1_NC_015044
SiMBoV2_NC_015047
TRYLCV_NC_020258
AbMBoV_NC_015048
MelYMV_NC_028143
SiMMV_NC_005331
ToRMV_NC_002556
ToSRV_NC_009612
OMoV_NC_011182
ToYSV_NC_007727
PSLDV_NC_012787
SiMAlV_NC_016579
TrYMV_BR-Msj1-10B
Cluster III
ToLCNDV_NC_004612
0.5

Figure 3c

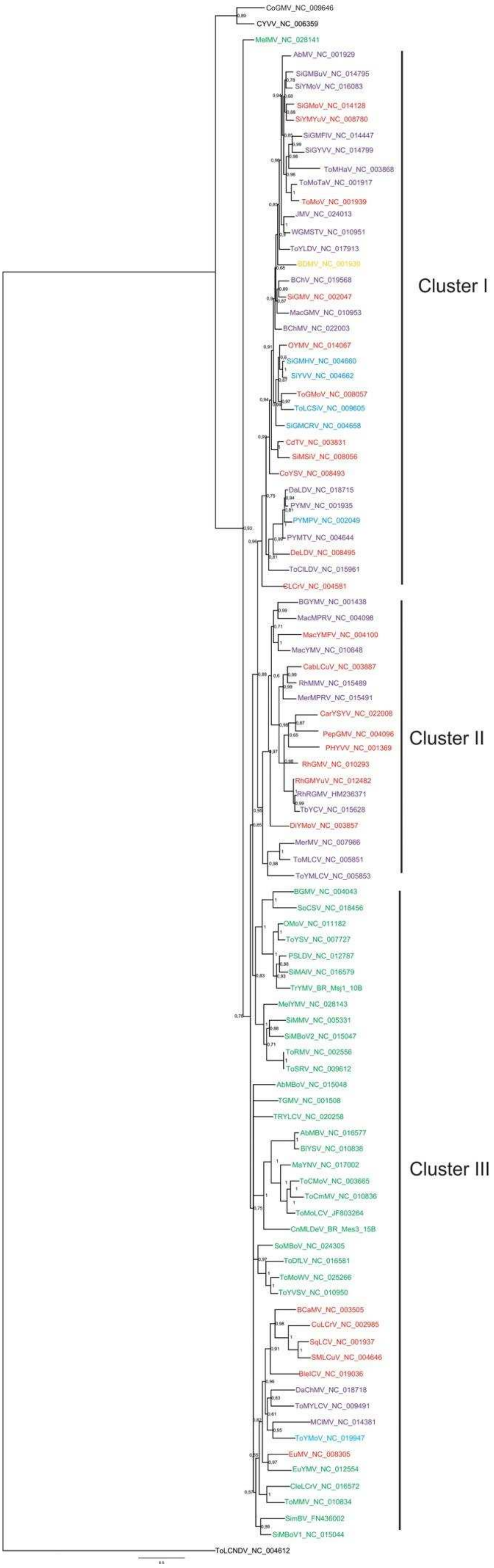
CoGMV_NC_009646
CYVV_NC_006359
MelMV_NC_028141
AbMV_NC_001929
SiGMBuV_NC_014795
SiYMoV_NC_016083
SiGMoV_NC_014128
SiYMYuV_NC_008780
SiGMFlV_NC_014447
SiGYVV_NC_014799
ToMHaV_NC_003868
ToMoTaV_NC_001917
ToMoV_NC_001939
JMV_NC_024013
WGMSTV_NC_010951
ToYLDV_NC_017913
SDMV_NC_001939
BChV_NC_019568
SiGMV_NC_002047
MacGMV_NC_010953
BChMV_NC_022003
OYMV_NC_014067
SiGMHV_NC_004660
SiYVV_NC_004662
ToGMoV_NC_008057
ToLCSiV_NC_009605
SiGMCRV_NC_004658
CdTV_NC_003831
SiMSiV_NC_008056
CoYSV_NC_008493
DaLDV_NC_018715
PYMV_NC_001935
PYMPV_NC_002049
PYMTV_NC_004644
DeLDV_NC_008495
ToClLDV_NC_015961
CLCrV_NC_004581
BGYMV_NC_001438
MacMPRV_NC_004098
MacYMFV_NC_004100
MacYMV_NC_010648
CabLCuV_NC_003887
RhMMV_NC_015489
MerMPRV_NC_015491
CarYSYV_NC_022008
PepGMV_NC_004096
PHYVV_NC_001369
RhGMV_NC_010293
RhGMYuV_NC_012482
RhRGMV_HM236371
TbYCV_NC_015628
DiYMoV_NC_003857
MerMV_NC_007966
ToMLCV_NC_005851
ToYMLCV_NC_005853
BGMV_NC_004043
SoCSV_NC_018456
OMoV_NC_011182
ToYSV_NC_007727
PSLDV_NC_012787
SiMAlV_NC_016579
TrYMV_BR_Msj1_10B
MelYMV_NC_028143
SiMMV_NC_005331
SiMBoV2_NC_015047
ToRMV_NC_002556
ToSRV_NC_009612
AbMBoV_NC_015048
TGMV_NC_001508
TRYLCV_NC_020258
AbMBV_NC_016577
BlYSV_NC_010838
MaYNV_NC_017002
ToCMoV_NC_003665
ToCmMV_NC_010836
ToMoLCV_JF803264
CnMLDeV_BR_Mes3_15B
SoMBoV_NC_024305
ToDfLV_NC_016581
ToMoWV_NC_025266
ToYVSV_NC_010950
BCaMV_NC_003505
CuLCrV_NC_002985
SqLCV_NC_001937
SMLCuV_NC_004646
BleICV_NC_019036
DaChMV_NC_018718
ToMYLCV_NC_009491
MClMV_NC_014381
ToYMoV_NC_019947
EuMV_NC_008305
EuYMV_NC_012554
CleLCrV_NC_016572
ToMMV_NC_010834
SimBV_FN436002
SiMBoV1_NC_015044
ToLCNDV_NC_004612
Cluster I
Cluster II
Cluster III

Figure 3d

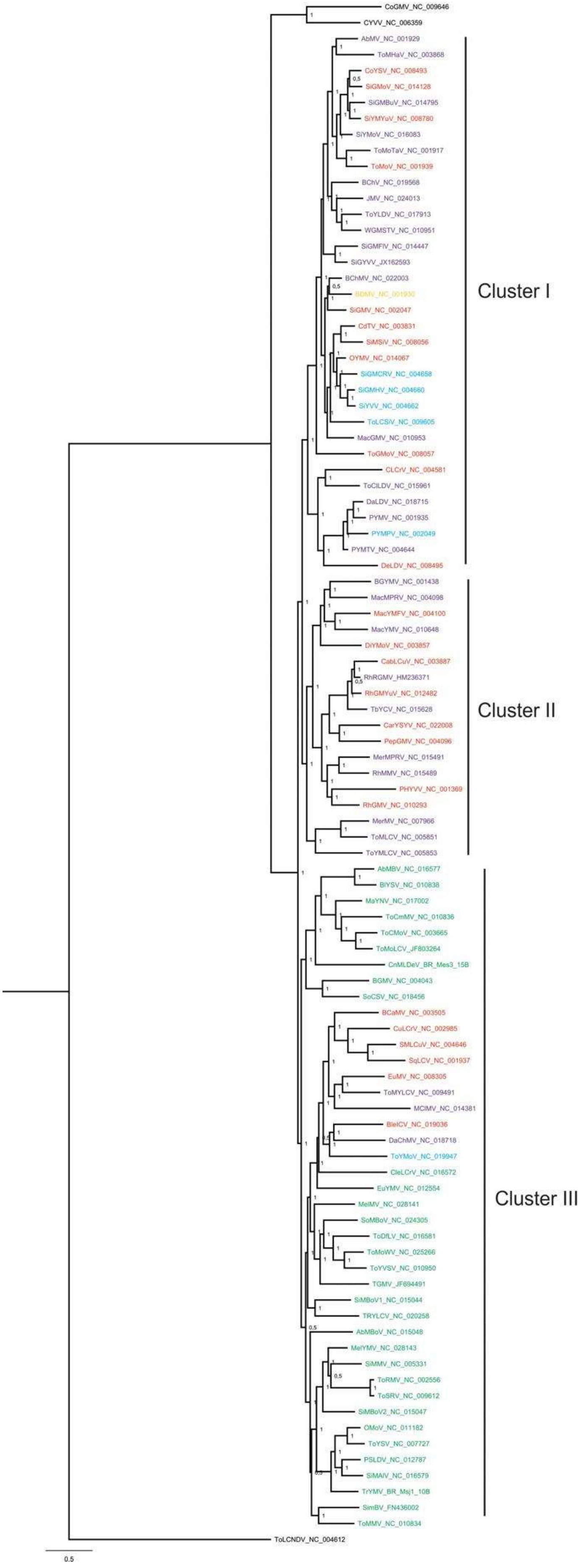
CoGMV_NC_009646
CYVV_NC_006359
AbMV_NC_001929
ToMHaV_NC_003868
CoYSV_NC_008493
SiGMoV_NC_014128
SiGMBuV_NC_014795
SiYMYuV_NC_008780
SiYMoV_NC_016083
ToMoTaV_NC_001917
ToMoV_NC_001939
BChV_NC_019568
JMV_NC_024013
ToYLDV_NC_017913
WGMSTV_NC_010951
SiGMFIV_NC_014447
SiGYVV_JX162593
BChMV_NC_022003
BDMV_NC_001930
SiGMV_NC_002047
CdTV_NC_003831
SiMSiV_NC_008056
OYMV_NC_014067
SiGMCRV_NC_004658
SiGMHV_NC_004660
SiYVV_NC_004662
ToLCSiV_NC_009605
MacGMV_NC_010953
ToGMoV_NC_008057
CLCrV_NC_004581
ToClLDV_NC_015961
DaLDV_NC_018715
PYMV_NC_001935
PYMPV_NC_002049
PYMTV_NC_004644
DeLDV_NC_008495
Cluster I
BGYMV_NC_001438
MacMPRV_NC_004098
MacYMFV_NC_004100
MacYMV_NC_010648
DiYMoV_NC_003857
CabLCuV_NC_003887
RhRGMV_HM236371
RhGMYuV_NC_012482
TbYCV_NC_015628
CarYSYV_NC_022008
PepGMV_NC_004096
MerMPRV_NC_015491
RhMMV_NC_015489
PHYVV_NC_001369
RhGMV_NC_010293
MerMV_NC_007966
ToMLCV_NC_005851
ToYMLCV_NC_005853
Cluster II
AbMBV_NC_016577
BlYSV_NC_010838
MaYNV_NC_017002
ToCmMV_NC_010836
ToCMoV_NC_003665
ToMoLCV_JF803264
CnMLDeV_BR_Mes3_15B
BGMV_NC_004043
SoCSV_NC_018456
BCaMV_NC_003505
CuLCrV_NC_002985
SMLCuV_NC_004646
SqLCV_NC_001937
EuMV_NC_008305
ToMYLCV_NC_009491
MClMV_NC_014381
BleICV_NC_019036
DaChMV_NC_018718
ToYMoV_NC_019947
CleLCrV_NC_016572
EuYMV_NC_012554
MelMV_NC_028141
SoMBoV_NC_024305
ToDfLV_NC_016581
ToMoWV_NC_025266
ToYVSV_NC_010950
TGMV_JF694491
SiMBoV1_NC_015044
TRYLCV_NC_020258
AbMBoV_NC_015048
MelYMV_NC_028143
SiMMV_NC_005331
ToRMV_NC_002556
ToSRV_NC_009612
SiMBoV2_NC_015047
OMoV_NC_011182
ToYSV_NC_007727
PSLDV_NC_012787
SiMAIV_NC_016579
TrYMV_BR_Msj1_10B
SimBV_FN436002
ToMMV_NC_010834
Cluster III
ToLCNDV_NC_004612
0.5

Figure 4a


AbMBoV_NC_015045_BOL_2007
PavYMV_NC_029105_BRA_2014
CeYSV_NC_016998_BRA_2009
MelMV_NC_028140_BRA_2014
MelYMV_NC_028142_BRA_2014
SdYMV_BR_Deg310A_BRA_2010
SiYMAV_NC_020255_BRA_2010
SiMBoV1_NC_015046_BOL_2007
SoMBoV_NC_024304_BOL_2007
ToDfLV_NC_016580_ARG_2008
ToMoWV_NC_025265_ARG_2008
ToYVSV_NC_010949_BRA_1983
TGMV_NC_001507_BRA_1984
BGMV_NC_004042_BRA_2007
MaYVV_NC_017000_BRA_2010
SimBV_NC_FN436001_BRA_2007
BlYSV_NC_010837_BRA_2007
MaYSV_NC_016999_BRA_2010
SoCSV_NC_018457_BRA_2010
MaYNV_NC_017001_BRA_2009
ToCMoV_NC_003664_BRA_1996
ToRMV_NC_002555_BRA_1996
ToSRV_NC_009607_BRA_2003
HeYMV_BR_Mac710A_BRA_2010
SiMoAV_NC_020256_BRA_2010
TrMV_BR_Fle110A_BRA_2010
SiMAIV_NC_016573_BRA_2010
SiYBV_NC_020254_BRA_2010
OMoV_NC_011181_BRA_2008
OxYVV_NC_026253_USA_2012
SiMoV_NC_004637_BRA_1999
SiYNV_NC_020253_BRA_2010
SiYMV_NC_004639_BRA_1999
ToYSV_NC_007726_BRA_1999
TrYMV_BR_Msj110A_BRA_2010
SimMV_NC_005330_BRA_2004
SiCmMV_NC_EU710751_BRA_2007
ToLDV_NC_EU710749_BRA_2005
ToMMV_NC_010833_BRA_2005
CleLCrV_NC_016578_BRA_2007
SiMBoV2_NC_015043_BOL_2007
CnMLDeV_BR_Mes315A_BRA_2015
PSLDV_NC_012786_BRA_2001
HyRMV1_BR_Rla510A_BRA_2010
HyRMV2_BR_Rla610A_BRA_2010
ToMoLCV_NC_JF803246_BRA_2004
PhYSV_BR_Rla112A_BRA_2012
Cluster III
AbMV_NC_001928_PRI_1989
SiGMV_NC_002046_USA_1998
ToMoV_NC_001938_USA_1989
RhRGMV_NC_HM236370_CUB_2009
SiYMoV_NC_016082_CUB_2009
BChV_NC_019569_VEN_2006
SiGMFlV_NC_014446_CUB_2009
SiGYVV_NC_004635_JAM_2008
JaMV_NC_024012_JAM_2004
ToLCCUV_NC_AM050143_CUB_2004
WGMSTV_NC_010948_JAM_2005
ToYLDV_NC_011348_CUB_2007
ToMHaV_NC_003867_CUB_1998
ToMoTaV_NC_001828_CUB_1997
BYMMxV_NC_015487_MEX_2006
CdTV_NC_003830_MEX_1983
OYMolV_NC_AY751753_MEX_2004
SiMSiV_NC_008059_MEX_2005
CoYSV_NC_008492_MEX_2005
MacGMV_NC_010952_JAM_2004
SiGMoV_NC_014130_USA_1997
SiGMBuV_NC_014794_JAM_2004
SiYMYuV_NC_008779_MEX_2005
OYMV_NC_014066_MEX_2004
SiGMHV_NC_004659_HND_2003
SiYVV_NC_004661_HND_2003
ToLCSiV_NC_009606_NIC_2002
BChMV_NC_022005_VEN_2007
SiGMCRV_NC_004657_CRI_2003
BDMV_NC_001931_COL_1987
DaLDV_NC_018717_VEN_2007
PYMPV_NC_002048_PAN_1997
PYMV_NC_001934_VEN_1991
PYMTV_NC_004638_TTO_2003
TMoLCV_NC_015122_VEN_2004
ToClLDV_NC_015962_VEN_2004
Cluster I
BleICV_NC_019035_MEX_2011
ToLDeV_NC_014510_PER_2003
ToChLPV_NC_005843_MEX_2002
ToGMoV_NC_008058_MEX_2005
ToYMoV_NC_019946_CRI_2003
MerMV_NC_007965_PRI_1997
ToMLCV_NC_005850_VEN_2004
ToYMLCV_NC_005852_VEN_2004
BGYMV_NC_001439_PRI_1996
MacYMFV_NC_004099_USA_1985
MacYMV_NC_010647_JAM_2006
DiYMoV_NC_003856_USA_1998
CabLCuJV_NC_DQ178614_JAM_2005
TbYCV_NC_015627_CUB_2007
DaChMV_NC_018716_VEN_2007
MacMPRV_NC_004097_PRI_1998
RhMMV_NC_015488_PRI_1997
MerMPRV_NC_015490_PRI_1998
PHYVV_NC_001359_MEX_1993
RhGMV_NC_010294_MEX_2007
CYVV_NC_006358_VNM_2004
CoGMV_NC_009644_VNM_2003
Cluster II
ToLCNDV_NC_004611_IND_1994
200.0
800
600
400
200

Figure 4b

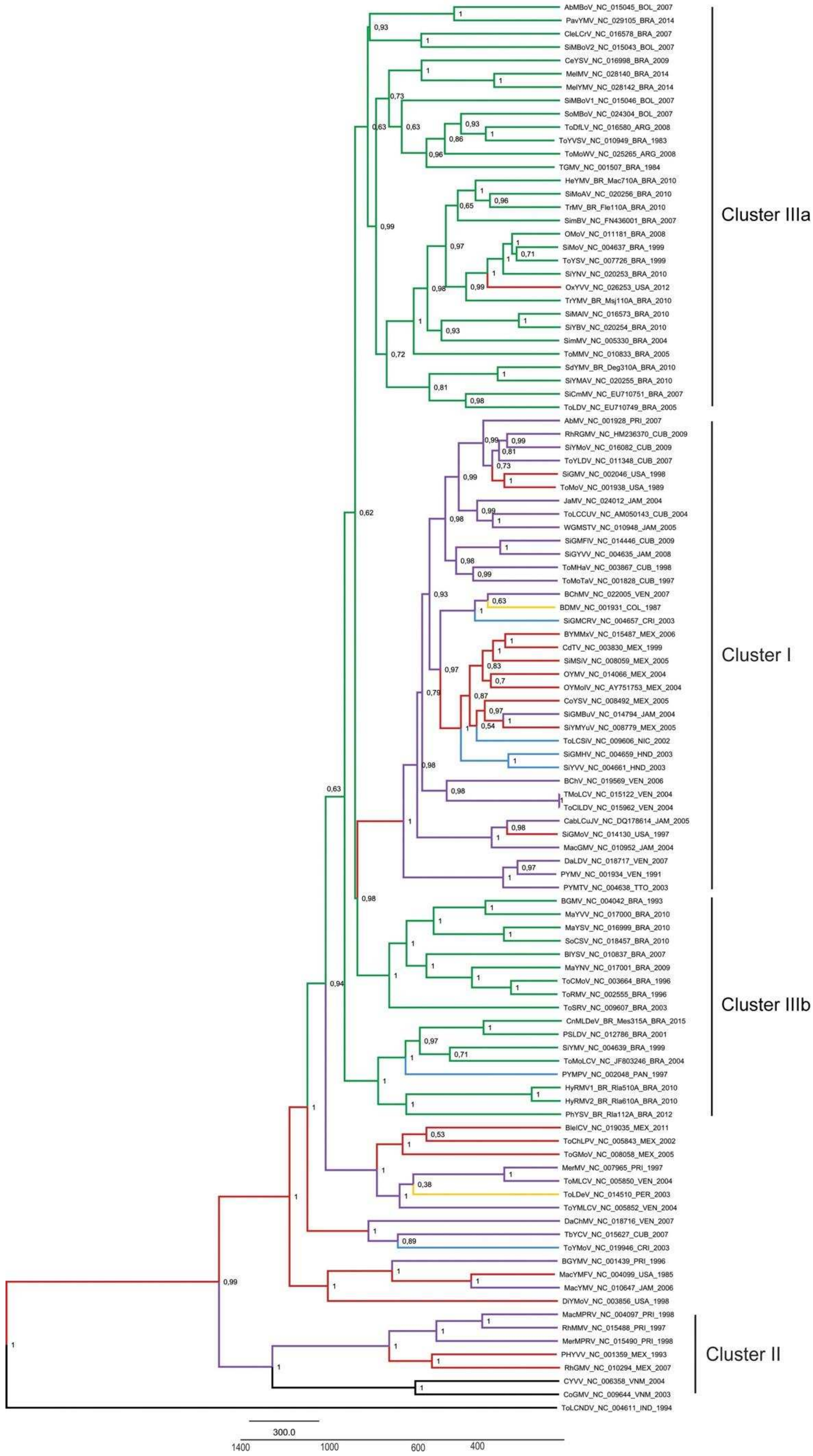
AbMBoV_NC_015045_BOL_2007
PavYMV_NC_029105_BRA_2014
CleLCrV_NC_016578_BRA_2007
SiMBoV2_NC_015043_BOL_2007
CeYSV_NC_016998_BRA_2009
MelMV_NC_028140_BRA_2014
MelYMV_NC_028142_BRA_2014
SiMBoV1_NC_015046_BOL_2007
SoMBoV_NC_024304_BOL_2007
ToDfLV_NC_016580_ARG_2008
ToYVSV_NC_010949_BRA_1983
ToMoWV_NC_025265_ARG_2008
TGMV_NC_001507_BRA_1984
HeYMV_BR_Mac710A_BRA_2010
SiMoAV_NC_020256_BRA_2010
TrMV_BR_Fle110A_BRA_2010
SimBV_NC_FN436001_BRA_2007
OMoV_NC_011181_BRA_2008
SiMoV_NC_004637_BRA_1999
ToYSV_NC_007726_BRA_1999
SiYNV_NC_020253_BRA_2010
OxYVV_NC_026253_USA_2012
TrYMV_BR_Msj110A_BRA_2010
SiMAIV_NC_016573_BRA_2010
SiYBV_NC_020254_BRA_2010
SimMV_NC_005330_BRA_2004
ToMMV_NC_010833_BRA_2005
SdYMV_BR_Deg310A_BRA_2010
SiYMAV_NC_020255_BRA_2010
SiCmMV_NC_EU710751_BRA_2007
ToLDV_NC_EU710749_BRA_2005
AbMV_NC_001928_PRI_2007
RhRGMV_NC_HM236370_CUB_2009
SiYMoV_NC_016082_CUB_2009
ToYLDV_NC_011348_CUB_2007
SiGMV_NC_002046_USA_1998
ToMoV_NC_001938_USA_1989
JaMV_NC_024012_JAM_2004
ToLCCUV_NC_AM050143_CUB_2004
WGMSTV_NC_010948_JAM_2005
SiGMFIV_NC_014446_CUB_2009
SiGYVV_NC_004635_JAM_2008
ToMHaV_NC_003867_CUB_1998
ToMoTaV_NC_001828_CUB_1997
BChMV_NC_022005_VEN_2007
BDMV_NC_001931_COL_1987
SiGMCRV_NC_004657_CRI_2003
BYMMxV_NC_015487_MEX_2006
CdTV_NC_003830_MEX_1999
SiMSiV_NC_008059_MEX_2005
OYMV_NC_014066_MEX_2004
OYMoIV_NC_AY751753_MEX_2004
CoYSV_NC_008492_MEX_2005
SiGMBuV_NC_014794_JAM_2004
SiYMYuV_NC_008779_MEX_2005
ToLCSiV_NC_009606_NIC_2002
SiGMHV_NC_004659_HND_2003
SiYVV_NC_004661_HND_2003
BChV_NC_019569_VEN_2006
TMoLCV_NC_015122_VEN_2004
ToCILDV_NC_015962_VEN_2004
CabLCuJV_NC_DQ178614_JAM_2005
SiGMoV_NC_014130_USA_1997
MacGMV_NC_010952_JAM_2004
DaLDV_NC_018717_VEN_2007
PYMV_NC_001934_VEN_1991
PYMTV_NC_004638_TTO_2003
BGMV_NC_004042_BRA_1993
MaYVV_NC_017000_BRA_2010
MaYSV_NC_016999_BRA_2010
SoCSV_NC_018457_BRA_2010
BIYSV_NC_010837_BRA_2007
MaYNV_NC_017001_BRA_2009
ToCMoV_NC_003664_BRA_1996
ToRMV_NC_002555_BRA_1996
ToSRV_NC_009607_BRA_2003
CnMLDeV_BR_Mes315A_BRA_2015
PSLDV_NC_012786_BRA_2001
SiYMV_NC_004639_BRA_1999
ToMoLCV_NC_JF803246_BRA_2004
PYMPV_NC_002048_PAN_1997
HyRMV1_BR_Rla510A_BRA_2010
HyRMV2_BR_Rla610A_BRA_2010
PhYSV_BR_Rla112A_BRA_2012
BleICV_NC_019035_MEX_2011
ToChLPV_NC_005843_MEX_2002
ToGMoV_NC_008058_MEX_2005
MerMV_NC_007965_PRI_1997
ToMLCV_NC_005850_VEN_2004
ToLDeV_NC_014510_PER_2003
ToYMLCV_NC_005852_VEN_2004
DaChMV_NC_018716_VEN_2007
TbYCV_NC_015627_CUB_2007
ToYMoV_NC_019946_CRI_2003
BGYMV_NC_001439_PRI_1996
MacYMFV_NC_004099_USA_1985
MacYMV_NC_010647_JAM_2006
DiYMoV_NC_003856_USA_1998
MacMPRV_NC_004097_PRI_1998
RhMMV_NC_015488_PRI_1997
MerMPRV_NC_015490_PRI_1998
PHYVV_NC_001359_MEX_1993
RhGMV_NC_010294_MEX_2007
CYVV_NC_006358_VNM_2004
CoGMV_NC_009644_VNM_2003
ToLCNDV_NC_004611_IND_1994
Cluster IIIa
Cluster I
Cluster IIIb
Cluster II
300.0
1400
1000
600
400

Figure 4c

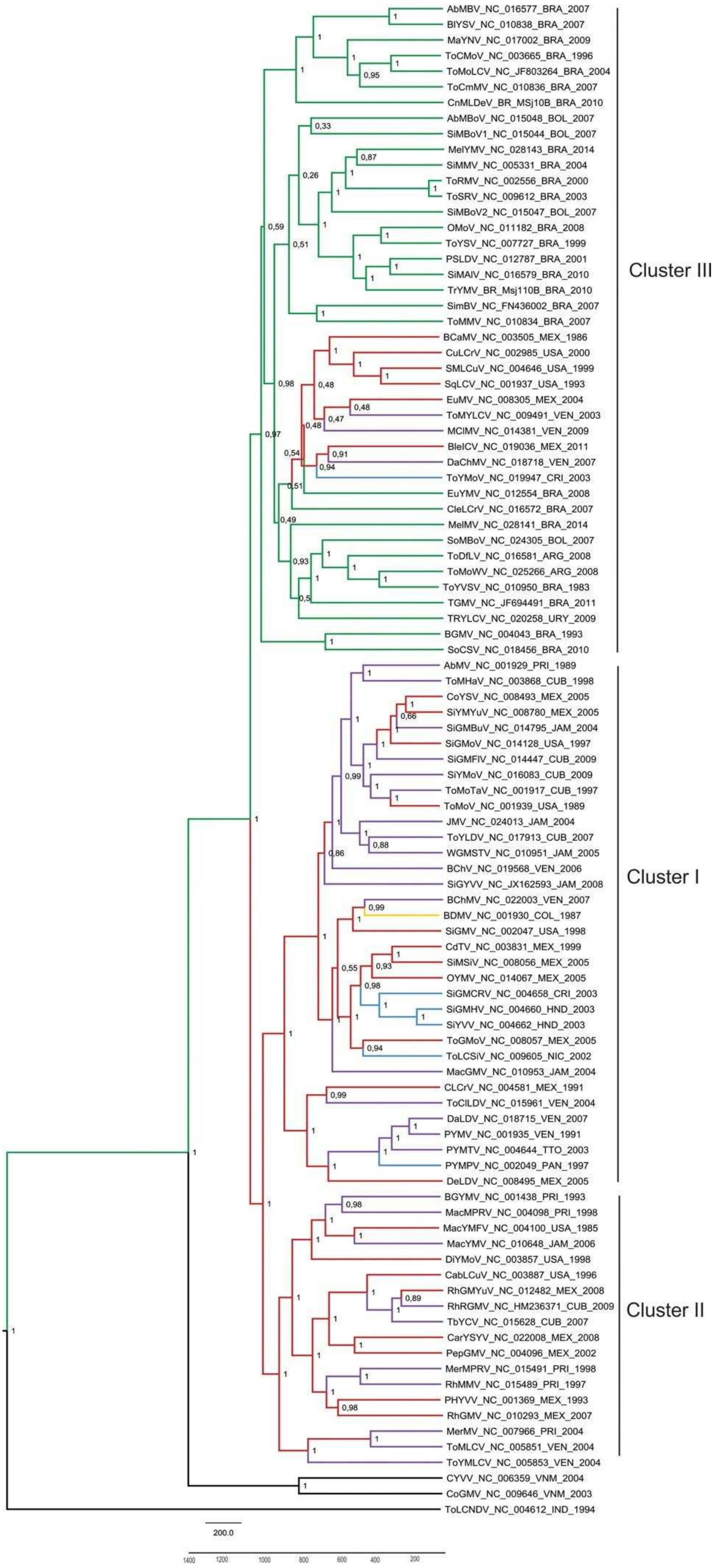
AbMBV_NC_016577_BRA_2007
BlYSV_NC_010838_BRA_2007
MaYNV_NC_017002_BRA_2009
ToCMoV_NC_003665_BRA_1996
ToMoLCV_NC_JF803264_BRA_2004
ToCmMV_NC_010836_BRA_2007
CnMLDeV_BR_MSj10B_BRA_2010
AbMBoV_NC_015048_BOL_2007
SiMBoV1_NC_015044_BOL_2007
MelYMV_NC_028143_BRA_2014
SiMMV_NC_005331_BRA_2004
ToRMV_NC_002556_BRA_2000
ToSRV_NC_009612_BRA_2003
SiMBoV2_NC_015047_BOL_2007
OMoV_NC_011182_BRA_2008
ToYSV_NC_007727_BRA_1999
PSLDV_NC_012787_BRA_2001
SiMAlV_NC_016579_BRA_2010
TrYMV_BR_Msj110B_BRA_2010
SimBV_NC_FN436002_BRA_2007
ToMMV_NC_010834_BRA_2007
BCaMV_NC_003505_MEX_1986
CuLCrV_NC_002985_USA_2000
SMLCuV_NC_004646_USA_1999
SqLCV_NC_001937_USA_1993
EuMV_NC_008305_MEX_2004
ToMYLCV_NC_009491_VEN_2003
MClMV_NC_014381_VEN_2009
BleICV_NC_019036_MEX_2011
DaChMV_NC_018718_VEN_2007
ToYMoV_NC_019947_CRI_2003
EuYMV_NC_012554_BRA_2008
CleLCrV_NC_016572_BRA_2007
MelMV_NC_028141_BRA_2014
SoMBoV_NC_024305_BOL_2007
ToDfLV_NC_016581_ARG_2008
ToMoWV_NC_025266_ARG_2008
ToYVSV_NC_010950_BRA_1983
TGMV_NC_JF694491_BRA_2011
TRYLCV_NC_020258_URY_2009
BGMV_NC_004043_BRA_1993
SoCSV_NC_018456_BRA_2010
Cluster III
AbMV_NC_001929_PRI_1989
ToMHaV_NC_003868_CUB_1998
CoYSV_NC_008493_MEX_2005
SiYMYuV_NC_008780_MEX_2005
SiGMBuV_NC_014795_JAM_2004
SiGMoV_NC_014128_USA_1997
SiGMFlV_NC_014447_CUB_2009
SiYMoV_NC_016083_CUB_2009
ToMoTaV_NC_001917_CUB_1997
ToMoV_NC_001939_USA_1989
JMV_NC_024013_JAM_2004
ToYLDV_NC_017913_CUB_2007
WGMSTV_NC_010951_JAM_2005
BChV_NC_019568_VEN_2006
SiGYVV_NC_JX162593_JAM_2008
BChMV_NC_022003_VEN_2007
BDMV_NC_001930_COL_1987
SiGMV_NC_002047_USA_1998
CdTV_NC_003831_MEX_1999
SiMSiV_NC_008056_MEX_2005
OYMV_NC_014067_MEX_2005
SiGMCRV_NC_004658_CRI_2003
SiGMHV_NC_004660_HND_2003
SiYVV_NC_004662_HND_2003
ToGMoV_NC_008057_MEX_2005
ToLCSiV_NC_009605_NIC_2002
MacGMV_NC_010953_JAM_2004
CLCrV_NC_004581_MEX_1991
ToClLDV_NC_015961_VEN_2004
DaLDV_NC_018715_VEN_2007
PYMV_NC_001935_VEN_1991
PYMTV_NC_004644_TTO_2003
PYMPV_NC_002049_PAN_1997
DeLDV_NC_008495_MEX_2005
Cluster I
BGYMV_NC_001438_PRI_1993
MacMPRV_NC_004098_PRI_1998
MacYMFV_NC_004100_USA_1985
MacYMV_NC_010648_JAM_2006
DiYMoV_NC_003857_USA_1998
CabLCuV_NC_003887_USA_1996
RhGMYuV_NC_012482_MEX_2008
RhRGMV_NC_HM236371_CUB_2009
TbYCV_NC_015628_CUB_2007
CarYSYV_NC_022008_MEX_2008
PepGMV_NC_004096_MEX_2002
MerMPRV_NC_015491_PRI_1998
RhMMV_NC_015489_PRI_1997
PHYVV_NC_001369_MEX_1993
RhGMV_NC_010293_MEX_2007
MerMV_NC_007966_PRI_2004
ToMLCV_NC_005851_VEN_2004
ToYMLCV_NC_005853_VEN_2004
Cluster II
CYVV_NC_006359_VNM_2004
CoGMV_NC_009646_VNM_2003
ToLCNDV_NC_004612_IND_1994
200.0
1400
1200
1000
800
600
400
200

Figure 4d


AbMBV_NC_016577_BRA_2007
BlYSV_NC_010838_BRA_2007
MaYNV_NC_017002_BRA_2009
ToCMoV_NC_003665_BRA_1996
ToMoLCV_NC_JF803264_BRA_2004
ToCmMV_NC_010836_BRA_2007
CnMLDeV_BR_Msj110B_BRA_2010
AbMBoV_NC_015048_BOL_2007
MelYMV_NC_028143_BRA_2014
SiMMV_NC_005331_BRA_2004
ToRMV_NC_002556_BRA_2000
ToSRV_NC_009612_BRA_2003
SiMBoV2_NC_015047_BOL_2007
OMoV_NC_011182_BRA_2008
ToYSV_NC_007727_BRA_1999
PSLDV_NC_012787_BRA_2001
SiMAIV_NC_016579_BRA_2010
TrYMV_BR_Msj110B_BRA_2010
SimBV_NC_FN436002_BRA_2007
ToMMV_NC_010834_BRA_2005
SiMBoV1_NC_015044_BOL_2007
TRYLCV_NC_020258_URY_2009
BCaMV_NC_003505_MEX_1986
CuLCrV_NC_002985_USA_2000
SMLCuV_NC_004646_USA_1999
SqLCV_NC_001937_USA_1993
EuMV_NC_008305_MEX_2004
ToMYLCV_NC_009491_VEN_2003
MClMV_NC_014381_VEN_2009
BleICV_NC_019036_MEX_2011
DaChMV_NC_018718_VEN_2007
ToYMoV_NC_019947_CRI_2003
CleLCrV_NC_016572_BRA_2007
EuYMV_NC_012554_BRA_2008
MelMV_NC_028141_BRA_2014
SoMBoV_NC_024305_BOL_2007
ToDfLV_NC_016581_ARG_2008
ToMoWV_NC_025266_ARG_2008
ToYVSV_NC_010950_BRA_1983
TGMV_NC_JF694491_BRA_2011
BGMV_NC_004043_BRA_1993
SoCSV_NC_018456_BRA_2010
AbMV_NC_001929_PRI_1989
ToMHaV_NC_003868_CUB_1998
CoYSV_NC_008493_MEX_2005
SiGMoV_NC_014128_USA_1997
SiGMBuV_NC_014795_JAM_2004
SiYMYuV_NC_008780_MEX_2005
SiYMoV_NC_016083_CUB_2009
ToMoTaV_NC_001917_CUB_1997
ToMoV_NC_001939_USA_1989
BChV_NC_019568_VEN_2006
JMV_NC_024013_JAM_2004
ToYLDV_NC_017913_CUB_2007
WGMSTV_NC_010951_JAM_2005
SiGMFlV_NC_014447_CUB_2009
SiGYVV_NC_JX162593_JAM_2008
BChMV_NC_022003_VEN_2007
SiGMV_NC_002047_USA_1998
BDMV_NC_001930_COL_1987
CdTV_NC_003831_MEX_1999
SiMSiV_NC_008056_MEX_2005
OYMV_NC_014067_MEX_2005
SiGMCRV_NC_004658_CRI_2003
SiGMHV_NC_004660_HND_2003
SiYVV_NC_004662_HND_2003
ToLCSiV_NC_009605_NIC_2002
MacGMV_NC_010953_JAM_2004
ToGMoV_NC_008057_MEX_2005
CLCrV_NC_004581_MEX_1991
ToClLDV_NC_015961_VEN_2004
DaLDV_NC_018715_VEN_2007
PYMV_NC_001935_VEN_1991
PYMTV_NC_004644_TTO_2003
PYMPV_NC_002049_PAN_1997
DeLDV_NC_008495_MEX_2005
BGYMV_NC_001438_PRI_1993
MacMPRV_NC_004098_PRI_1998
MacYMFV_NC_004100_USA_1985
MacYMV_NC_010648_JAM_2006
DiYMoV_NC_003857_USA_1998
CabLCuV_NC_003887_USA_1996
TbYCV_NC_015628_CUB_2007
RhGMYuV_NC_012482_MEX_2008
RhRGMV_NC_HM236371_CUB_2009
CarYSYV_NC_022008_MEX_2007
PepGMV_NC_004096_MEX_2002
MerMPRV_NC_015491_PRI_1998
RhMMV_NC_015489_PRI_1997
PHYVV_NC_001369_MEX_1993
RhGMV_NC_010293_MEX_2007
MerMV_NC_007966_PRI_2004
ToMLCV_NC_005851_VEN_2004
ToYMLCV_NC_005853_VEN_2004
CYVV_NC_006359_VNM_2004
CoGMV_NC_009646_VNM_2003
ToLCNDV_NC_004612_IND_1994
Cluster III
Cluster I
Cluster II
200.0
1800
1400
1000
600
200

Supporting Information

Suppl. Table S1. New World begomovirus sequences retrieved from GenBank's non-redundant nucleotide database.

		Acession #				
Nº	**Species**	**DNA-A**	**DNA-B**	**Acronym**	**Location**	**Date**
	Eastern South America					
1	*Abutilon mosaic Bolivia virus*	NC_015045	NC_015048	AbMBoV	Bolivia	2007
2	*Abutilon mosaic Brazil virus*	NC_016574	NC_016577	AbMBV	Brazil	2007
3	*Bean golden mosaic virus*	NC_004042	NC_004043	BGMV	Brazil	1993
4	*Blainvillea yellow spot virus*	NC_010837	NC_010838	BlYSV	Brazil	2007
5	*Centrosema yellow spot virus*	NC_016998	---	CeYSV	Brazil	2009
6	*Cleome leaf crumple virus*	NC_016578	NC_016572	CleLCrV	Brazil	2007
7	*Euphorbia yellow mosaic virus*	NC_012553	NC_012554	EuYMV	Brazil	2007
8	*Macroptilium yellow net virus*	NC_017001	NC_017002	MaYNV	Brazil	2009
9	*Macroptilium yellow spot virus*	NC_016999	---	MaYSV	Brazil	2010
10	*Macroptilium yellow vein virus*	NC_017000	---	MaYVV	Brazil	2010
11	*Melochia mosaic virus*	NC_028140	NC_028141	MelMV	Brazil	2014
12	*Melochia yellow mosaic virus*	NC_028142	NC_028143	MelYMV	Brazil	2014
13	*Okra mottle virus*	NC_011181	NC_011182	OMoV	Brazil	2008
14	*Passionfruit severe leaf distortion virus*	NC_012786	NC_012787	PSLDV	Brazil	2001
15	*Pavonia yellow mosaic virus*	NC_029105	---	PavYMV	Brazil	2014
16	*Sida Brazil virus*	FN436001	FN436002	SimBV	Brazil	2007
17	*Sida common mosaic virus*	EU710751	---	SiCmMV	Brazil	2007
18	*Sida micrantha mosaic virus*	NC_005330	NC_005331	SiMMV	Brazil	2004
19	*Sida mosaic Alagoas virus*	NC_016573	NC_016579	SiMAlV	Brazil	2010

Suppl. Table S1. (cont.)

		Acession #				
Nº	**Species**	**DNA-A**	**DNA-B**	**Acronym**	**Location**	**Date**
20	*Sida mosaic Bolivia virus 1*	NC_015046	NC_015044	SiMBoV1	Bolivia	2007
21	*Sida mosaic Bolivia virus 2*	NC_015043	NC_015047	SiMBoV2	Bolivia	2007
22	*Sida mottle Alagoas virus*	NC_020256	---	SiMoAV	Brazil	2010
23	*Sida mottle virus*	NC_004637	---	SiMoV	Brazil	1999
24	*Sida yellow blotch virus*	NC_020254	---	SiYBV	Brazil	2010
25	*Sida yellow leaf curl virus*	EU710750	---	SiYLCV	Brazil	2007
26	*Sida yellow mosaic Alagoas virus*	NC_020255	---	SiYMAV	Brazil	2010
27	*Sida yellow mosaic virus*	NC_004639	---	SiYMV	Brazil	1999
28	*Sida yellow net virus*	NC_020253	---	SiYNV	Brazil	2010
29	*Solanum mosaic Bolivia virus*	NC_024304	NC_024305	SoMBoV	Bolivia	2007
30	*Soybean chlorotic spot virus*	NC_018457	NC_018456	SoCSV	Brazil	2010
31	*Tomato chlorotic mottle virus*	NC_003664	NC_003665	ToCMoV	Brazil	1996
32	*Tomato common mosaic virus*	NC_010835	NC_010836	ToCmMV	Brazil	2007
33	*Tomato dwarf leaf virus*	NC_016580	NC_016581	ToDfLV	Argentina	2008
34	*Tomato golden mosaic virus*	NC_001507	NC_001508	TGMV	Brazil	1984
35	*Tomato leaf distortion virus*	EU710749	---	ToLDV	Brazil	2005
36	*Tomato mild mosaic virus*	NC_010833	NC_010834	ToMMV	Brazil	2005
37	*Tomato mottle leaf curl virus*	JF803246	JF803264	ToMoLCV	Brazil	2004
38	*Tomato mottle wrinkle virus*	NC_025265	NC_025266	ToMoWV	Argentina	2008
39	*Tomato rugose mosaic virus*	NC_002555	NC_002556	ToRMV	Brazil	1996
40	*Tomato rugose yellow leaf curl virus*	NC020257	NC_020258	TRYLCV	Uruguay	2009
41	*Tomato severe rugose virus*	NC_009607	NC_009612	ToSRV	Brazil	2003
42	*Tomato yellow spot virus*	NC_007726	NC_007727	ToYSV	Brazil	1999
43	*Tomato yellow vein streak virus*	NC_010949	NC_010950	ToYVSV	Brazil	1983
44	*Cnidoscolus mosaic leaf deformation virus*	BR-Mes3-15A	BR-Msj1-10B	CnMLDeV	Brazil	2015

Suppl. Table S1. (cont.)

Nº	Species	Acession # DNA-A	Acession # DNA-B	Acronym	Location	Date
45	*Herissantia yellow mosaic virus*	BR-Mac7-10A	---	HeYMV	Brazil	2010
46	*Herissantia rugose mosaic virus 1*	BR_Rla5_10A	---	HyRMV1	Brazil	2010
47	*Herissantia rugose mosaic virus 2*	BR_Rla6_10A	---	HyRMV2	Brazil	2010
48	*Physalis yellow spot virus*	BR_Rla1_12A	---	PhYSV	Brazil	2012
49	*Sidastrum yellow mosaic virus*	BR-Deg3-10A	---	SdYMV	Brazil	2010
50	*Triumfetta mosaic virus*	BR-Fle1-10A	---	TrMV	Brazil	2010
51	*Triumfetta yellow mosaic virus*	BR-Msj1-10A	BR-Msj1-10B	TrYMV	Brazil	2010
	Western South America					
52	*Bean dwarf mosaic virus*	NC_001931	NC_001930	BDMV	Colombia	1987
53	*Tomato leaf deformation virus*	NC_014510	---	ToLDeV	Peru	2003
	Central America					
54	*Melon chlorotic leaf curl virus*	NC_003865	---	MCLCuV	Costa Rica	1998
55	*Potato yellow mosaic Panama virus*	NC_002048	NC_002049	PYMPV	Panama	1997
56	*Sida golden mosaic Costa Rica virus*	NC_004657	NC_004658	SiGMCRV	Costa Rica	2003
57	*Sida golden mosaic Honduras virus*	NC_004659	NC_004660	SiGMHV	Honduras	2003
58	*Sida yellow vein virus*	NC_004661	NC_004662	SiYVV	Honduras	2003
59	*Tomato leaf curl Sinaloa virus*	NC_009606	NC_009605	ToLCSiV	Nicaragua	2002
60	*Tomato severe leaf curl virus*	NC_004642	---	ToSLCV	Guatemala	1996
61	*Tomato yellow mottle virus*	NC_019946	NC_019947	ToYMoV	Costa Rica	2003

Suppl. Table S1. (cont.)

Nº	Species	Acession # DNA-A	 DNA-B	Acronym	Location	Date
	Caribbean					
62	*Abutilon mosaic virus*	NC_001928	NC_001929	AbMV	Porto Rico	1989
63	*Bean chlorosis virus*	NC_019569	NC_019568	BChV	Venezuela	2006
64	*Bean chlorotic mosaic virus*	NC_022005	NC_022003	BChMV	Venezuela	2007
65	*Bean golden yellow mosaic virus*	NC_001439	NC_001438	BGYMV	Puerto Rico	1996
66	*Cabbage leaf curl Jamaica virus*	DQ178614	---	CabLCuJV	Jamaica	2005
67	*Dalechampia chlorotic mosaic virus*	NC_018716	NC_018718	DaChMV	Venezuela	2007
68	*Datura leaf distortion virus*	NC_018717	NC_018715	DaLDV	Venezuela	2007
69	*Jatropha mosaic virus*	NC_024012	NC_024013	JaMV	Jamaica	2004
70	*Macroptilium golden mosaic virus*	NC_010952	NC_010953	MacGMV	Jamaica	2004
71	*Macroptilium mosaic Puerto Rico virus*	NC_004097	NC_004098	MacMPRV	Puerto Rico	1998
72	*Macroptilium yellow mosaic virus*	NC_010647	NC_010648	MacYMV	Jamaica	2006
73	*Melon chlorotic mosaic virus*	NC_014380	NC_014381	MClMV	Venezuela	2009
74	*Merremia mosaic Puerto Rico virus*	NC_015490	NC_015491	MerMPRV	Puerto Rico	1998
75	*Merremia mosaic virus*	NC_007965	NC_007966	MerMV	Puerto Rico	1997
76	*Potato yellow mosaic Trinidad virus*	NC_004638	NC_004644	PYMTV	Trinidad and Tobago	2003
77	*Potato yellow mosaic virus*	NC_001934	NC_001935	PYMV	Venezuela	1991
78	*Rhynchosia mild mosaic virus*	NC_015488	NC_015489	RhMMV	Puerto Rico	1997
79	*Rhynchosia rugose golden mosaic virus*	HM236370	HM236371	RhRGMV	Cuba	2009
80	*Sida golden mosaic Buckup virus*	NC_014794	NC_014795	SiGMBuV	Jamaica	2004
81	*Sida golden mosaic Florida virus*	NC_014446	NC_014447	SiGMFlV	Cuba	2009
82	*Sida golden yellow vein virus*	NC_004635	NC_014799	SiGYVV	Jamaica	2008
83	*Sida yellow mottle virus*	NC_016082	NC_016083	SiYMoV	Cuba	2009

Suppl. Table S1. (cont.)

		Acession #				
Nº	Species	DNA-A	DNA-B	Acronym	Location	Date
84	*Tobacco leaf curl Cuba virus*	AM050143	---	ToLCCUV	Cuba	2004
85	*Tobacco yellow crinkle virus*	NC_015627	NC_015628	TbYCV	Cuba	2007
86	*Tomato chlorotic leaf distortion virus*	NC_015962	NC_015961	ToClLDV	Venezuela	2004
87	*Tomato mild yellow leaf curl Aragua virus*	NC_009490	NC_009491	ToMYLCV	Venezuela	2003
88	*Tomato mosaic Havana virus*	NC_003867	NC_003868	ToMHaV	Cuba	1998
89	*Tomato mosaic leaf curl virus*	NC_005850	NC_005851	ToMLCV	Venezuela	2004
90	*Tomato mottle leaf curl Zulia virus*	NC_015122	---	TMoLCV	Venezuela	2004
91	*Tomato mottle Taino virus*	NC_001828	NC_001917	ToMoTaV	Cuba	1997
92	*Tomato yellow leaf distortion virus*	NC_011348	NC_017913	ToYLDV	Cuba	2007
93	*Tomato yellow margin leaf curl virus*	NC_005852	NC_005853	ToYMLCV	Venezuela	2004
94	*Wissadula golden mosaic St Thomas virus*	NC_010948	NC_010951	WGMSTV	Jamaica	2005
	North America					
95	*Bean calico mosaic virus*	NC_003504	NC_003505	BcaMV	Mexico	1986
96	*Bean yellow mosaic Mexico virus*	NC_015487	---	BYMMxV	Mexico	2006
97	*Blechum interveinal chlorosis virus*	NC_019035	NC_019036	BleICV	Mexico	2011
98	*Cabbage leaf curl virus*	NC_003866	NC_003887	CabLCuV	USA	1996
99	*Capraria yellow spot Yucatan virus*	NC_022007	NC_022008	CarYSYV	Mexico	2007
100	*Chino del tomate virus*	NC_003830	NC_003831	CdTV	Mexico	1999
101	*Corchorus yellow spot virus*	NC_008492	NC_008493	CoYSV	Mexico	2005
102	*Cotton leaf crumple virus*	NC_004580	NC_004581	CLCrV	Mexico	1991
103	*Cucurbit leaf crumple virus*	NC_002984	NC_002985	CuLCrV	USA	2000
104	*Desmodium leaf distortion virus*	NC_008494	NC_008495	DeLDV	Mexico	2005

Suppl. Table S1. (cont.)

		Acession #				
Nº	**Species**	**DNA-A**	**DNA-B**	**Acronym**	**Location**	**Date**
105	*Dicliptera yellow mottle virus*	NC_003856	NC_003857	DiYMoV	USA	1998
106	*Euphorbia mosaic virus*	NC_008304	NC_008305	EuMV	Mexico	2004
107	*Macroptilium yellow mosaic Florida virus*	NC_004099	NC_004100	MacYMFV	USA	1985
108	*Okra yellow mosaic Mexico virus*	NC_014066	NC_014067	OYMV	Mexico	2004
109	*Okra yellow mottle Iguala virus*	AY751753	---	OYMoIV	Mexico	2004
110	*Oxalis yellow vein virus strain*	NC_026253	---	OxYVV	USA	2012
111	*Pepper golden mosaic virus*	NC_004101	NC_004096	PepGMV	Mexico	1996
112	*Pepper huasteco yellow vein virus*	NC_001359	NC_001369	PHYVV	Mexico	1993
113	*Rhynchosia golden mosaic virus*	NC_010294	NC_010293	RhGMV	Mexico	2007
114	*Rhynchosia golden mosaic Yucatan virus*	NC_012481	NC_012482	RhGMYuV	Mexico	2006
115	*Sida golden mosaic virus*	NC_002046	NC_002047	SiGMV	USA	1998
116	*Sida golden mottle virus*	NC_014130	NC_014128	SiGMoV	USA	1997
117	*Sida mosaic Sinaloa virus*	NC_008059	NC_008056	SiMSiV	Mexico	2005
118	*Sida yellow mosaic Yucatan virus*	NC_008779	NC_008780	SiYMYuV	Mexico	2005
119	*Squash leaf curl virus*	NC_001936	NC_001937	SqLCV	USA	1993
120	*Squash mild leaf curl virus*	NC_004645	NC_004646	SMLCuV	USA	1979
121	*Tomato chino La Paz virus*	NC_005843	---	ToChLPV	Mexico	2002
122	*Tomato golden mottle virus*	NC_008058	NC_008057	ToGMoV	Mexico	2005
123	*Tomato mottle virus*	NC_001938	NC_001939	ToMoV	USA	1989
	Old World and outgroup					
124	*Corchorus golden mosaic virus*	NC_009644	NC_009646	CoGMV	Vietnam	2003
125	*Corchorus yellow vein virus*	NC_006358	NC_006359	CoYVV	Vietnam	2000
126	*Tomato leaf curl New Delhi virus*	NC_004611	NC_004612	ToLCNDV	India	1994

Suppl. Table S2. Putative recombination events detected among begomoviruses from the New World, based on the full-length DNA-A.

Events	**Breakpoints**[a]		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**[b]	***P*-values**[c]
	Begin	**End**			**Minor**	**Major**		
1	1517	2599	Rep	ToRMV_NC_002555	ToSRV_NC_009607	ToCMoV_NC_003664	RGBMCS3	4.87E-62
2	2631	1708	Rep/CP	SiYMV_NC_004639	SiMoV_NC_004637	Unknown	RGBMCS3	1.02E-52
3	1952	76	Rep/IR	HyRMV1_BR_Rla5_10A	PSLDV_NC_012786	HyRMV2_BR_Rla6_10A	RGBMCS3	6.09E-43
4	1911	104	Rep/IR	PYMV_NC_001934	Unknown	DaLDV_NC_018717	RGBMCS3	2.86E-41
5	117	1762	CP/Rep	PYMPV_NC_002048	DaLDV_NC_018717	Unknown	RGBMCS3	5.89E-34
6	19	1838	CP/Rep	CabLCuJV_DQ178614	TbYCV_NC_015627	CoYSV_NC_008492 SiGMoV_NC_014130	RGBMCS3	3.96E-32
7	1	1908	CP/Rep	SiGMoV_NC_014130	SiGMBuV_NC_014794	MacGMV_NC_010952	RGBMCS3	4.48E-33
8	1837	2473	Rep	MaYSV_NC_016999	BlYSV_NC_010837	SoCSV_NC_018457 MaYNV_NC_017001	RGBMCS3	2.52E-23
9	1749	24	Rep	CoYSV_NC_008492	MacGMV_NC_010952	OYMV_NC_014066	RGBMCS3	1.10E-22
10	1907	2140	Rep	ToSRV_NC_009607 ToRMV_NC_002555	SimMV_NC_005330 ToLDV_EU710749 ToMMV_NC_010833	SoMBoV_NC_024304 ToYVSV_NC_010949	RGBMCS3	2.08E-20
11	1891	2115	Rep	SiYMoV_NC_016082 ToLCCUV_AM050143	TbYCV_NC_015627	RhRGMV_HM236370 AbMV_NC_001928 CdTV_NC_003830 OYMoIV_AY751753	RGBMCS3	1.54E-18

Suppl. Table S2(cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
						SiGMV_NC_002046 ToYLDV_NC_011348		
12	1768	129	Rep/IR	DaLDV_NC_018717	BChMV_NC_022005 BDMV_NC_001931 SiGMCRV_NC_004657	PYMTV_NC_004638	RGBMCS3	9.01E-17
13	1950	157	Rep/IR	MacMPRV_NC_004097	Unknown	RhMMV_NC_015488	RGBMCS3	3.54E-16
14	1959	2418	Rep	RhMMV_NC_015488 MerMPRV_NC_015490	BChMV_NC_022005 AbMV_NC_001928 BDMV_NC_001931 DaLDV_NC_018717 JaMV_NC_024012 OYMoIV_AY751753 SiGMCRV_NC_004657 SiGMHV_NC_004659 SiGMV_NC_002046 SiYVV_NC_004661 TMoLCV_NC_015122 ToClLDV_NC_015962 ToMHaV_NC_003867 ToMoV_NC_001938 WGMSTV_NC_010948	Unknown	RGBMCS3	5.01E-16
15	2217	2493	Rep	MelYMV_NC_028142	Unknown	MelMV_NC_028140 CeYSV_NC_016998 SdYMV_BR_Deg3_10A	RGBMCS3	5.06E-16

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
15						SiMBoV1_NC_015046 SiMoAV_NC_020256 ToMoWV_NC_025265 ToYSV_NC_007726		
16	1973	2222*	Rep	SiMoV_NC_004637 CeYSV_NC_016998 HyRMV2_BR_Rla6_10A OMoV_NC_011181 OxYVV_NC_026253	Unknown	MaYNV_NC_017001 ToCMoV_NC_003664	RGBMCS3	1.70E-16
17	1878	2593	Rep	SimBV_FN436001	Unknown	HeYMV_BR_Mac7_10A SiMAlV_NC_016573 SiMoAV_NC_020256 SiYBV_NC_020254	RGBMCS3	6.97E-16
18	1925	2397	Rep	ToYMLCV_NC_005852	BChMV_NC_022005 BDMV_NC_001931 SiGMCRV_NC_004657 SiYVV_NC_004661 TMoLCV_NC_015122 ToClLDV_NC_015962 WGMSTV_NC_010948	Unknown	RGBMCS3	1.73E-15
19	1653	2597	Rep	BYMMxV_NC_015487	Unknown	CdTV_NC_003830	RGBMCS3	2.57E-15
20	1955	2373	Rep	RhGMV_NC_010294	BChMV_NC_022005	Unknown	RGBMCS3	4.67E-15
					BDMV_NC_001931	Unknown		
					CdTV_NC_003830	Unknown		

Suppl. Table S2 (cont.)

Events	**Breakpoints**		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**	***P*-values**
	Begin	**End**			**Minor**	**Major**		
20					DaLDV_NC_018717	Unknown		
					OYMoIV_AY751753	Unknown		
					OYMV_NC_014066	Unknown		
					PYMTV_NC_004638	Unknown		
					RhRGMV_HM236370	Unknown		
					SiGMBuV_NC_014794	Unknown		
					SiGMCRV_NC_004657	Unknown		
					SiGMFlV_NC_014446	Unknown		
					SiGMHV_NC_004659	Unknown		
					SiGYVV_NC_004635	Unknown		
					SiYMYuV_NC_008779	Unknown		
					SiYVV_NC_004661	Unknown		
					TbYCV_NC_015627	Unknown		
					TMoLCV_NC_015122	Unknown		
					ToClLDV_NC_015962	Unknown		
					ToLCSiV_NC_009606	Unknown		
					ToMHaV_NC_003867	Unknown		
					ToMLCV_NC_005850	Unknown		
					ToYLDV_NC_011348	Unknown		
					WGMSTV_NC_010948	Unknown		
21	1912	2419	Rep	ToCMoV_NC_003664	PSLDV_NC_012786	BGMV_NC_004042	RGBMCS3	6.01E-17
					CnMLDeV_BR_Mes3_15A	MaYVV_NC_017000		
22	2498	1838	Rep/CP	ToMoWV_NC_025265	ToYVSV_NC_010949	ToLDV_EU710749	RGBMCS3	3.25E-21
23	1891	2308	Rep	ToMLCV_NC_005850	BChMV_NC_022005	Unknown	RGBMCS3	2.72E-14
				MerMV_NC_007965	BChV_NC_019569			
					BDMV_NC_001931			

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
23					DaLDV_NC_018717 OYMV_NC_014066 RhRGMV_HM236370 SiGMCRV_NC_004657 SiGMFlV_NC_014446 SiGMHV_NC_004659 SiYVV_NC_004661 ToLCSiV_NC_009606 WGMSTV_NC_010948			
24	2167	2540	Rep	TrYMV_BR_Msj1_10A	Unknown	HeYMV_BR_Mac7_10A SiMoAV_NC_020256 TrMV_BR_Fle1_10A	RGBMCS3	6.73E-14
25	2308	28	Rep/IR	SiYNV_NC_020253	TrMV_BR_Fle1_10A SiMoAV_NC_020256	ToYSV_NC_007726	RGBMCS3	5.12E-13
26	953	1403	CP/Rep	MacGMV_NC_010952	Unknown Unknown	SiYMAV_NC_020255 SdYMV_BR_Deg3_10A	RGBMCS3	3.69E-13
27	2207	27	Rep/IR	ToLCNDV_NC_004611	ToMMV_NC_010833 AbMBoV_NC_015045 BChMV_NC_022005 BChV_NC_019569 BGYMV_NC_001439 CdTV_NC_003830 CleLCrV_NC_016578	CoGMV_NC_009644 CYVV_NC_006358	RGBMCS3	3.89E-13

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
27					DaChMV_NC_018716 DaLDV_NC_018717 DiYMoV_NC_003856 HeYMV_BR_Mac7_10A JaMV_NC_024012 MacGMV_NC_010952 MacYMFV_NC_004099 MacYMV_NC_010647 MelMV_NC_028140 OYMoIV_AY751753 OYMV_NC_014066 PavYMV_NC_029105 PhYSV_BR_Rla1_12A PHYVV_NC_001359 PYMTV_NC_004638 RhRGMV_HM236370 SdYMV_BR_Deg3_10A SiCmMV_EU710751 SiGMBuV_NC_014794 SiGMHV_NC_004659 SiMAlV_NC_016573 SiMBoV1_NC_015046 SiMBoV2_NC_015043 SimMV_NC_005330 SiMoAV_NC_020256 SiYMoV_NC_016082 SiYVV_NC_004661 SoCSV_NC_018457 TbYCV_NC_015627			

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
27					ToDfLV_NC_016580 ToLCCUV_AM050143 ToLCSiV_NC_009606 ToLDeV_NC_014510 ToLDV_EU710749 ToMHaV_NC_003867 ToMoLCV_JF803246 ToMoV_NC_001938 ToSRV_NC_009607 ToYLDV_NC_011348 ToYSV_NC_007726 TrMV_BR_Fle1_10A			
28	1915	2186	Rep	SoMBoV_NC_024304 ToYVSV_NC_010949	Unknown	ToDfLV_NC_016580	RGBMCS3	7.37E-13
29	2019	2412	Rep	TbYCV_NC_015627 SiGMFlV_NC_014446	ToYLDV_NC_011348 BChV_NC_019569 CdTV_NC_003830 JaMV_NC_024012 OYMoIV_AY751753 RhRGMV_HM236370 SiGMBuV_NC_014794 SiGMV_NC_002046 SiGYVV_NC_004635 SiMSiV_NC_008059 SiYMYuV_NC_008779 SiYVV_NC_004661	ToYMoV_NC_019946	RGBMCS3	1.73E-16

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
29					ToLCSiV_NC_009606			
					ToMoTaV_NC_001828			
					ToMoV_NC_001938			
					WGMSTV_NC_010948			
30	69	1000	CP	ToYLDV_NC_011348	BChV_NC_019569	SiYMoV_NC_016082	RGBMCS3	2.08E-11
						RhRGMV_HM236370		
31	2371	20	Rep/IR	MerMV_NC_007965	JaMV_NC_024012	Unknown	RGBMCS3	3.94E-11
				ToMLCV_NC_005850	MacGMV_NC_010952	Unknown		
					SiMSiV_NC_008059	Unknown		
					ToLCCUV_AM050143	Unknown		
					ToMoTaV_NC_001828	Unknown		
32	2013	2250*	Rep	SiYMV_NC_004639	CYVV_NC_006358	SiYNV_NC_020253	RGBMCS3	5.87E-14
				CnMLDeV_BR_Mes3_15A	CoGMV_NC_009644	AbMBoV_NC_015045		
						AbMV_NC_001928		
						CoYSV_NC_008492		
						DaLDV_NC_018717		
						HeYMV_BR_Mac7_10A		
						MacGMV_NC_010952		
						MelYMV_NC_028142		
						PYMTV_NC_004638		
						RhRGMV_HM236370		
						SiCmMV_EU710751		
						SiGMBuV_NC_014794		
						SiGMCRV_NC_004657		
						SiGMHV_NC_004659		

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
32						SiGMoV_NC_014130 SiGMV_NC_002046 SiMBoV1_NC_015046 SiMoAV_NC_020256 ToClLDV_NC_015962 ToLDeV_NC_014510 ToMMV_NC_010833 TrMV_BR_Fle1_10A		
33	1986	2338	Rep	PHYVV_NC_001359	BDMV_NC_001931 BChMV_NC_022005 DaLDV_NC_018717 OYMV_NC_014066 PYMTV_NC_004638 SiGMBuV_NC_014794 SiGMCRV_NC_004657 SiGMHV_NC_004659 SiYMYuV_NC_008779 TMoLCV_NC_015122 ToClLDV_NC_015962 ToLCSiV_NC_009606	Unknown Unknown Unknown	RGBMCS3	6.71E-11
34	1954	2104	Rep	CoYSV_NC_008492 CabLCuJV_DQ178614 MacGMV_NC_010952 PYMV_NC_001934 SiGMoV_NC_014130	MaYVV_NC_017000 BGMV_NC_004042 SimBV_FN436001	JaMV_NC_024012 AbMV_NC_001928 BChV_NC_019569	RGBMCS3	6.24E-11

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
34						CdTV_NC_003830 OYMoIV_AY751753 RhRGMV_HM236370 SiGMV_NC_002046 SiGYVV_NC_004635 SiMSiV_NC_008059 SiYVV_NC_004661 ToMoTaV_NC_001828 ToMoV_NC_001938 ToYLDV_NC_011348		
35	1920	2248	Rep	SoCSV_NC_018457	PhYSV_BR_Rla1_12A	MaYVV_NC_017000 BGMV_NC_004042 SimBV_FN436001	RGBMCS3	8.01E-11
36	1814	2042	Rep	SiGMHV_NC_004659	SiMAlV_NC_016573 SiYBV_NC_020254	SiYVV_NC_004661 CdTV_NC_003830 JaMV_NC_024012 OYMoIV_AY751753 RhRGMV_HM236370 SiGMV_NC_002046 SiMSiV_NC_008059 ToYLDV_NC_011348	RGBMCS3	1.08E-10
37	1465	1555	Rep	SiCmMV_EU710751	Unknown	CeYSV_NC_016998 HyRMV2_BR_Rla6_10A MelMV_NC_028140 MelYMV_NC_028142	RGBMCS3	1.32E-10

Suppl. Table S2 (cont.)

Events	**Breakpoints**		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**	***P*-values**
	Begin	**End**			**Minor**	**Major**		
37						OMoV_NC_011181 OxYVV_NC_026253 PavYMV_NC_029105 SdYMV_BR_Deg3_10A SiMBoV1_NC_015046 SiMBoV2_NC_015043 SiMoAV_NC_020256 SiMoV_NC_004637 SoMBoV_NC_024304 TrMV_BR_Fle1_10A TrYMV_BR_Msj1_10A		
38	184	1642	CP	OMoV_NC_011181	SimMV_NC_005330	SiMoV_NC_004637 OxYVV_NC_026253 SiYMV_NC_004639 ToYSV_NC_007726	RGBMCS3	1.46E-10
39	2231*	2616	Rep	ToYSV_NC_007726	MelMV_NC_028140	OMoV_NC_011181 OxYVV_NC_026253	RGBMCS3	8.19E-11
40	2453	78	Rep/IR	BleICV_NC_019035	SiMSiV_NC_008059 MacGMV_NC_010952	Unknown Unknown	RGBMCS3	5.13E-10
41	1881	2088	Rep	SiYMYuV_NC_008779 SiGMBuV_NC_014794	Unknown Unknown	SiGMV_NC_002046 AbMV_NC_001928	RGBMCS3	6.49E-10
42	27	622	CP	SiYMAV_NC_020255 MelMV_NC_028140 MelYMV_NC_028142 SdYMV_BR_Deg3_10A	Unknown	SiMSiV_NC_008059 OYMV_NC_014066 SiGMHV_NC_004659 SiGMoV_NC_014130	RGBMCS3	1.42E-09

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
43	2024	2178	Rep	SiMBoV1_NC_015046 TGMV_NC_001507	ToMMV_NC_010833	Unknown Unknown	RGBMCS3	4.68E-09
44	1426	1887*	Rep	RhGMV_NC_010294	Unknown	ToYMoV_NC_019946	RGBMCS3	5.27E-15
45	1447	1709*	Rep	BChV_NC_019569	Unknown	CoYSV_NC_008492 CdTV_NC_003830 OYMoIV_AY751753 OYMV_NC_014066 SiGMBuV_NC_014794 SiGMHV_NC_004659 SiGMoV_NC_014130 SiGMV_NC_002046 SiMSiV_NC_008059 SiYMYuV_NC_008779 SiYVV_NC_004661 ToLCSiV_NC_009606 ToMHaV_NC_003867	RGBMCS3	9.94E-08
46	2416	2542	Rep	SiMSiV_NC_008059 ToMoTaV_NC_001828	JaMV_NC_024012 CabLCuJV_DQ178614 CoYSV_NC_008492	SiGMV_NC_002046 ToMHaV_NC_003867 ToMoV_NC_001938	RGBMCS3	2.66E-08
47	1831	2566	Rep	OYMV_NC_014066	Unknown	CdTV_NC_003830	RGBMCS3	2.45E-10
48	1247	1915	Rep	RhMMV_NC_015488 MacMPRV_NC_004097	CYVV_NC_006358	ToChLPV_NC_005843 AbMV_NC_001928 BGMV_NC_004042	RGBMCS3	7.83E-10

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
48						BlYSV_NC_010837		
						CleLCrV_NC_016578		
						CnMLDeV_BR_Mes3_15A		
						HeYMV_BR_Mac7_10A		
						JaMV_NC_024012		
						MacYMFV_NC_004099		
						MacYMV_NC_010647		
						MaYVV_NC_017000		
						MelMV_NC_028140		
						MelYMV_NC_028142		
						OYMoIV_AY751753		
						PavYMV_NC_029105		
						PhYSV_BR_Rla1_12A		
						PYMTV_NC_004638		
						RhRGMV_HM236370		
						SdYMV_BR_Deg3_10A		
						SiGMFlV_NC_014446		
						SiGMV_NC_002046		
						SiGYVV_NC_004635		
						SiMBoV1_NC_015046		
						SiMBoV2_NC_015043		
						SimMV_NC_005330		
						SiMoAV_NC_020256		
						SiYBV_NC_020254		
						SiYNV_NC_020253		
						SiYVV_NC_004661		
						TbYCV_NC_015627		
						TGMV_NC_001507		
						ToClLDV_NC_015962		
						ToDfLV_NC_016580		

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
48						ToLDeV_NC_014510		
						ToLDV_EU710749		
						ToMHaV_NC_003867		
						ToMMV_NC_010833		
						ToMoTaV_NC_001828		
						ToMoV_NC_001938		
						ToSRV_NC_009607		
						ToYLDV_NC_011348		
						ToYSV_NC_007726		
						TrMV_BR_Fle1_10A		
						WGMSTV_NC_010948		
						ToLCSiV_NC_009606		
49	2212*	2305	Rep	MaYVV_NC_017000	Unknown	OxYVV_NC_026253	RGBMCS3	2.68E-07
				BGMV_NC_004042	Unknown	BChV_NC_019569		
				SimBV_FN436001	Unknown	CdTV_NC_003830		
						JaMV_NC_024012		
						OYMoIV_AY751753		
						PYMTV_NC_004638		
						RhRGMV_HM236370		
						ToChLPV_NC_005843		
						ToMoTaV_NC_001828		
50	1957	2102*	Rep	JaMV_NC_024012	Unknown	ToClLDV_NC_015962	RGBMCS3	3.32E-07
				AbMV_NC_001928		TMoLCV_NC_015122		
				BChV_NC_019569				
				CdTV_NC_003830				
				OYMoIV_AY751753				
				RhRGMV_HM236370				
				SiGMV_NC_002046				
				ToYLDV_NC_011348				

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
50				SiYVV_NC_004661 ToMoTaV_NC_001828 ToMoV_NC_001938				
51	350	622*	CP	SiYMAV_NC_020255	SiMAlV_NC_016573 SiYBV_NC_020254	SdYMV_BR_Deg3_10A SiMBoV1_NC_015046	RGBMCS3	4.11E-07
52	1967	2082	Rep	ToGMoV_NC_008058	BYMMxV_NC_015487	ToChLPV_NC_005843	RGBMCS3	6.02E-07
53	2144	2494	Rep	MacYMV_NC_010647	Unknown	MacYMFV_NC_004099	RGBMCS3	7.52E-07
54	2446	2599	Rep	SiYMoV_NC_016082 RhRGMV_HM236370 ToYLDV_NC_011348	Unknown	SiGMV_NC_002046	RGBMCS3	9.28E-07
55	170*	256	IR/CP	TGMV_NC_001507	RhGMV_NC_010294	SiGMFlV_NC_014446	RGBMCS3	9.32E-07
56	2177*	2461	Rep	TrMV_BR_Fle1_10A HeYMV_BR_Mac7_10A SiMoAV_NC_020256 SiYNV_NC_020253	ToMoV_NC_001938 SiGMCRV_NC_004657 SiGMV_NC_002046	Unknown	RGBMCS3	1.56E-06
57	2504	2616*	Rep	SdYMV_BR_Deg3_10A SiYMAV_NC_020255	JaMV_NC_024012 MacGMV_NC_010952	Unknown Unknown	RGBMCS3	2.74E-07
58	756	1074	CP	ToMoWV_NC_025265	ToSRV_NC_009607	AbMBoV_NC_015045	RGBMCS3	5.57E-06

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
59	2486*	2635	Rep	TrMV_BR_Fle1_10A	Unknown	SiGMFlV_NC_014446	RGBMCS3	5.25E-10
				HeYMV_BR_Mac7_10A		SiGYVV_NC_004635		
				SiMBoV1_NC_015046		ToMHaV_NC_003867		
				SiMoAV_NC_020256				
60	245	1685	CP/Rep	ToLDV_EU710749	OxYVV_NC_026253	Unknown	RGBMCS3	1.93E-18
61	1357	1516	Rep	RhRGMV_HM236370	Unknown	SiYBV_NC_020254	RGBMCS3	1.07E-05
				AbMV_NC_001928	Unknown	OxYVV_NC_026253		
				BChMV_NC_022005				
				BDMV_NC_001931				
				BYMMxV_NC_015487				
				CdTV_NC_003830				
				CoYSV_NC_008492				
				DaLDV_NC_018717				
				JaMV_NC_024012				
				OYMoIV_AY751753				
				OYMV_NC_014066				
				PYMTV_NC_004638				
				SiGMCRV_NC_004657				
				SiGMFlV_NC_014446				
				SiGMHV_NC_004659				
				SiGMV_NC_002046				
				SiMSiV_NC_008059				
				SiYMoV_NC_016082				
				SiYMYuV_NC_008779				
				SiYVV_NC_004661				
				ToLCCUV_AM050143				
				ToMHaV_NC_003867				
				ToMoTaV_NC_001828				

Suppl. Table S2 (cont.)

Events	Breakpoints		Genomic regions	Recombinants	Putative parents		Methods	*P*-values
	Begin	End			Minor	Major		
62	2252*	35	Rep/IR	ToMoLCV_JF803246	CnMLDeV_BR_Mes3_15A PSLDV_NC_012786	HyRMV2_BR_Rla6_10A	RGBMCS3	3.86E-06
63	1940	2196	Rep	AbMBoV_NC_015045	ToLDV_EU710749	PavYMV_NC_029105	RGBMCS3	7.73E-07
64	403	1062	CP	MacYMFV_NC_004099 BGYMV_NC_001439 MacYMV_NC_010647	Unknown	PavYMV_NC_029105 AbMBoV_NC_015045	RGBMCS3	1.40E-05
65	2520	2619*	Rep	MacYMFV_NC_004099	SiGYVV_NC_004635	Unknown	RGBMCS3	1.40E-07
66	2486	11	Rep/IR	CnMLDeV_BR_Mes3_15A PSLDV_NC_012786	SiGYVV_NC_004635	BlYSV_NC_010837	RGBMCS3	1.72E-07
67	2407	2536	Rep	JaMV_NC_024012	MacGMV_NC_010952 PYMV_NC_001934	SiGMV_NC_002046	RGBMCS3	5.79E-07
68	1739	1964*	Rep	ToYLDV_NC_011348 BChV_NC_019569 CdTV_NC_003830 RhRGMV_HM236370 ToMoTaV_NC_001828	SiYMV_NC_004639	ToMHaV_NC_003867	RGBMCS3	5.18E-05
69	264*	791	CP	SimMV_NC_005330	SiMBoV2_NC_015043	Unknown	RGBMCS3	1.93E-06
70	166	249	IR/CP	SiGMFlV_NC_014446	Unknown	JaMV_NC_024012	RGBMCS3	4.02E-04

a *, breakpoint could not be precisely pinpointed.

b R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimera; S, SisScan; 3, 3SEQ.

c The reported - values are for the methods indicated in red, and they are the lowest *P* values calculated for the region in question.

Suppl. Table S3. Putative recombination events detected among begomoviruses from the New World, based on the full-length DNA-B**.**

Events	**Breakpoint**[a]		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**[b]	***P*-value**[c]
	Begin	**End**			**Minor**	**Major**		
1	79	667	NSP	OMoV_NC_011182	SiMMV_NC_005331	ToYSV_NC_007727	RGBMCS3	2.31E-31
2	592	1356	NSP	OYMV_NC_014067	CLCrV_NC_004581	BChV_NC_019568	RGBMCS3	4.29E-29
3	2406	2522	IR	PYMV_NC_001935	CoYSV_NC_008493 SiGMoV_NC_014128	PYMTV_NC_004644	RGBMCS3	6.91E-16
4	2391	27	IR	RhGMYuV_NC_012482	ToMYLCV_NC_009491	TbYCV_NC_015628	RGBMCS3	4.31E-14
5	2170	2393*	IR	PYMTV_NC_004644	PYMPV_NC_002049	PYMV_NC_001935	RGBMCS3	1.09E-15
6	2603	73	IR	BChV_NC_019568	SiMBoV2_NC_015047	BChMV_NC_022003	RGBMCS3	4.71E-11
7	102	1416	NSP/MP	SiGMFlV_NC_014447	CoYSV_NC_008493	Unknown	RGBMCS3	6.55E-11
8	2446*	2544	IR	DaLDV_NC_018715	Unknown	PYMTV_NC_004644	RGBMCS3	2.18E-12
9	2074	129	MP/IR	CabLCuV_NC_003887	Unknown	TbYCV_NC_015628	RGBMCS3	5.94E-10
10	288	1538	NSP/MP	SiYMoV_NC_016083	Unknown	CoYSV_NC_008493	RGBMCS3	6.92E-11
11	103	345	IR	SiGMBuV_NC_014795	Unknown	SiYMYuV_NC_008780	RGBMCS3	4.63E-13
12	29	62*	IR	ToRMV_NC_002556	Unknown	ToSRV_NC_009612	RGBMCS3	3.55E-09
13	2137	2298	MP	DaLDV_NC_018715	Unknown	PYMPV_NC_002049	RGBMCS3	4.38E-09
14	1645	1978	MP	CoYSV_NC_008493	DeLDV_NC_008495	SiGMoV_NC_014128	RGBMCS3	1.94E-07

Suppl. Table S3 (cont.)

Events	**Breakpoint**		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**	***P*-value**
	Begin	**End**			**Minor**	**Major**		
15	2643	80	IR	SiMBoV1_NC_015044	SiMAlV_NC_016579	Unknown	RGBMCS3	1.68E-09
16	2102	2414	MP	WGMSTV_NC_010951	BDMV_NC_001930	ToYLDV_NC_017913	RGBMCS3	4.06E-07
17	381	2122	NSP/MP	TbYCV_NC_015628	CarYSYV_NC_022008	PYMTV_NC_004644	RGBMCS3	4.27E-07
18	2387	2491	MP/IR	SiMSiV_NC_008056	Unknown	CdTV_NC_003831	RGBMCS3	5.14E-07
19	69	179	IR	TrYMV_BR_Msj1_10B	MelMV_NC_028141	Unknown	RGBMCS3	4.26E-06
20	2350	28*	IR	MerMV_NC_007966	ToMLCV_NC_005851	Unknown	RGBMCS3	4.82E-11
21	697	2452	NSP/MP	MelYMV_NC_028143	Unknown	OMoV_NC_011182	RGBMCS3	9.70E-27
22	2470*	26	IR	MacYMFV_NC_004100	WGMSTV_NC_010951	CarYSYV_NC_022008	RGBMCS3	9.76E-06
23	736	1345	NSP	SiGMV_NC_002047	BDMV_NC_001930	SiGMBuV_NC_014795 CoYSV_NC_008493 SiYMYuV_NC_008780	RGBMCS3	3.87E-09
24	2459	33	IR	JMV_NC_024013	ToMLCV_NC_005851	Unknown	RGBMCS3	5.51E-06
25	2427*	72*	IR	OYMV_NC_014067	SiMMV_NC_005331	SiYVV_NC_004662 SiGMHV_NC_004660	RGBMCS3	1.04E-05
26	2462	306	MP/IR	ToYVSV_NC_010950	SiMBoV2_NC_015047	Unknown	RGBMCS3	2.49E-12
27	2390	60	IR	TbYCV_NC_015628	ToMoV_NC_001939	Unknown	RGBMCS3	5.50E-05

Suppl. Table S3 (cont.)

Events	**Breakpoint**		**Genomic regions**	**Recombinants**	**Putative parents**		**Methods**	***P*-value**
	Begin	**End**			**Minor**	**Major**		
28	1542	2003	MP	CabLCuV_NC_003887	RhMMV_NC_015489	RhRGMV_HM236371 RhGMYuV_NC_012482	RGBMCS3	6.62E-05
29	2464	13	IR	CnMLDeV_BR_Mes3_15B	ToYLDV_NC_017913	CLCrV_NC_004581	RGBMCS3	3.90E-05
30	1045	1285	NSP	SiGMHV_NC_004660 SiGMCRV_NC_004658 SiYVV_NC_004662	RhMMV_NC_015489	SiGMoV_NC_014128 CoYSV_NC_008493 SiGMBuV_NC_014795 SiYMYuV_NC_008780	RGBMCS3	5.86E-04
31	85	222	IR/NSP	MelMV_NC_028141	MelYMV_NC_028143	ToYMoV_NC_019947	RGBMCS3	1.62E-04
32	2409	293*	IR	ToDfLV_NC_016581	SiMBoV2_NC_015047	Unknown	RGBMCS3	2.63E-06
33	2278	16*	MP/IR	ToCmMV_NC_010836	TRYLCV_NC_020258	BlYSV_NC_010838	RGBMCS3	6.51E-05
34	1645	2117*	MP	SiGYVV_JX162593	Unknown	SiGMBuV_NC_014795	RGBMCS3	1.17E-05
35	1105	1593	NSP	WGMSTV_NC_010951	AbMV_NC_001929	ToYLDV_NC_017913	RGBMCS3	7.40E-04
36	1114	2090*	MP	ToLCSiV_NC_009605	PYMPV_NC_002049	Unknown	RGBMCS3	3.64E-05
37	42*	1062*	NSP	PYMTV_NC_004644	OMoV_NC_011182 ToYSV_NC_007727	SiGMHV_NC_004660 SiYVV_NC_004662	RGBMCS3	2.04E-04
38	1328	2113	MP	DeLDV_NC_008495	ToMoV_NC_001939	ToYVSV_NC_010950	RGBMCS3	3.33E-06
39	1019	1274	NSP	RhMMV_NC_015489	BChV_NC_019568	Unknown	RGBMCS3	1.29E-04

a * breakpoint could not be precisely pinpointed.

b R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimera; S, SisScan; 3, 3SEQ.

c The reported *P*-values are for the methods indicated in red, and they are the lowest *P* values calculated for the region in question.