![JVI logo] Journals.ASM.org

Brazilian Begomovirus Populations Are Highly Recombinant, Rapidly Evolving, and Segregated Based on Geographical Location

Carolina S. Rocha,[a*] Gloria P. Castillo-Urquiza,[a] Alison T. M. Lima,[a] Fábio N. Silva,[a] Cesar A. D. Xavier,[a] Braz T. Hora-Júnior,[a] José E. A. Beserra-Júnior,[a*] Antonio W. O. Malta,[a*] Darren P. Martin,[b] Arvind Varsani,[c,d,e] Poliane Alfenas-Zerbini,[a*] Eduardo S. G. Mizubuti,[a] F. Murilo Zerbini[a]

Dep. de Fitopatologia/BIOAGRO and National Research Institute for Plant-Pest Interactions (INCT-IPP), Universidade Federal de Viçosa, Viçosa, MG, Brazil[a]; Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory, Cape Town, South Africa[b]; Electron Microscope Unit, University of Cape Town, Rondebosch, Cape Town, South Africa[c]; School of Biological Sciences, University of Canterbury, Ilam, Christchurch, New Zealand[d]; Biomolecular Interaction Centre, University of Canterbury, Ilam, Christchurch, New Zealand[e]

The incidence of begomovirus infections in crop plants sharply increased in Brazil during the 1990s following the introduction of the invasive B biotype of the whitefly vector, *Bemisia tabaci*. It is believed that this biotype transmitted begomoviruses from noncultivated plants to crop species with greater efficiency than indigenous *B. tabaci* biotypes. Either through rapid host adaptation or selection pressure in genetically diverse populations of noncultivated hosts, over the past 20 years various previously unknown begomovirus species have became progressively more prevalent in cultivated species such as tomato. Here we assess the genetic structure of begomovirus populations infecting tomatoes and noncultivated hosts in southeastern Brazil. Between 2005 and 2010, we sampled and sequenced 126 DNA-A and 58 DNA-B full-length begomovirus components. We detected nine begomovirus species in tomatoes and eight in the noncultivated host samples, with four species common to both tomatoes and noncultivated hosts. Like many begomoviruses, most species are obvious interspecies recombinants. Furthermore, species identified in tomato have probable parental viruses from noncultivated hosts. While the population structures of five well-sampled viral species all displayed geographical subdivision, a noncultivated host-infecting virus was more genetically variable than the four predominantly tomato-infecting viruses.

T he family *Geminiviridae* is comprised of viruses with circular, single-stranded DNA genomes and particles structured as twinned imperfect icosahedra (1). The family is divided into four genera (*Mastrevirus*, *Curtovirus*, *Topocuvirus*, and *Begomovirus*) based on the type of insect vector, host range, genome organization, and phylogenetic relationships (2). The genus *Begomovirus* includes viruses with one or two genomic components which infect dicotyledonous plants and are transmitted by the whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae) (2). Begomovirus diseases are a major factor limiting crop yields in tropical and subtropical regions (3, 4). Tomatoes (*Solanum lycopersicum* L.) in particular are seriously affected by begomoviruses on a worldwide scale (5–7). In the Americas, diseases caused by begomoviruses have significantly impacted tomato production since the 1980s (8–10). With the exception of *Tomato yellow leaf curl virus* (TYLCV), which was first introduced to the Americas via tomato seedlings imported into the Dominican Republic from Israel (11) and has now spread to the United States, Mexico, and a number of countries in the Caribbean (12), all begomoviruses isolated and characterized from tomato plants from the American continents (North, Central, and South America) are native to these continents and have never been detected in other regions (for examples, see references 13, 14, and 15). In fact, the phylogeny of geminiviruses as a whole is highly structured around geographical distributions of its composite viruses (2). Based on genome organization, phylogenetic relationships, and geographical distribution, begomoviruses have been broadly divided into Old World (OW [Europe, Africa, Asia, and Australia]) and New World (NW [the Americas]) lineages (16).

In Brazil, 11 begomovirus species, all of which have been found only in that country, are currently known to infect tomatoes nat-

urally in the field: *Tomato golden mosaic virus* (TGMV) (17), *Tomato rugose mosaic virus* (ToRMV) (18), *Tomato chlorotic mottle virus* (ToCMoV) (19), *Tomato yellow spot virus* (ToYSV) (20), *Tomato severe rugose virus* (ToSRV) (21), *Tomato common mosaic virus* (ToCmMV), *Tomato mild mosaic virus* (ToMlMV) (15), *Tomato yellow vein streak virus* (ToYVSV) (22), *Tomato interveinal chlorosis virus* (ToICV), *Tomato mottle leaf curl virus* (TMoLCV), and *Tomato golden vein virus* (TGVV) (23). At least seven additional genetically distinct begomoviruses have been associated with diseases of tomato, but these viruses have not been completely characterized and their role in the etiology of these diseases remains to be established (10, 24, 25).

Currently, the best-supported hypothesis explaining the sudden emergence of tomato-infecting begomoviruses in Brazil assumes that indigenous viruses were transferred from noncultivated hosts to tomatoes after the introduction and dissemination

of the B biotype of *B. tabaci* in the early 1990s. This invasive biotype feeds on a greater variety of host species than indigenous *B. tabaci* biotypes and has probably facilitated transmission of viruses from a broader array of hosts into tomato (26). The biological characterization of some of the tomato viruses (ToRMV, ToCMoV, and ToYSV) has indeed confirmed that ubiquitous noncultivated species such as *Nicandra physaloides*, *Solanum nigrum*, and *Datura stramonium* are noncultivated alternative hosts (18–20). Moreover, begomoviruses originally found in common noncultivated species, such as *Sida mottle virus* (SiMoV) and *Sida micrantha mosaic virus* (SimMV), have also been found infecting tomatoes under field conditions (27, 28).

It is possible that noncultivated host-infecting viruses underwent a period of rapid host adaptation following productive infection of tomatoes (10, 19). Evolution of geminiviruses is driven mainly by mutation, genetic drift, and recombination (1, 29). Mutation frequencies and nucleotide substitution rates have been estimated for the begomoviruses TYLCV, *Tomato yellow leaf curl China virus* (TYLCCNV), and *East African cassava mosaic virus* (EACMV) and for the mastrevirus *Maize streak virus* (MSV) and have been found to be similar to those estimated for RNA viruses ($\sim 10^{-4}$ substitutions per site per year) (30–34). The presence of several begomoviruses in the field, all transmitted by the same vector, likely facilitates frequent mixed infections in which two or more virus species are simultaneously present within individual plants. This situation increases the probability of recombination and/or pseudorecombination (reassortment of genomic components) among viral genomic components, which could potentially accelerate host adaptation (19, 25, 35–39).

Management strategies for plant viral diseases are generally preventive measures which tend to work most efficiently when based on firm epidemiological and pathogen demographic data (4, 40). Although much has been done to characterize Brazilian tomato-infecting begomoviruses, viral population studies which might provide valuable clues on the potential of these viruses to evolve in response to management strategies are lacking.

We have carried out a large-scale study to determine the genetic structure of begomovirus populations associated with tomato crops and noncultivated species in five important tomato-producing regions of southeastern Brazil. Our results corroborate previous studies indicating the presence of several begomovirus species in the field and demonstrate that viruses originally detected in tomatoes can also be found in noncultivated species and vice versa. Viruses infecting noncultivated hosts are more genetically variable than the tomato-infecting viruses, and in both cases DNA-B is more variable than DNA-A. Surprisingly, even over the relatively small geographical area studied, phylogenetic analysis showed local division between populations. Recombination analysis confirmed the previously suggested pervasiveness of interspecies recombination among Brazilian begomoviruses, with viruses infecting noncultivated hosts often identified as recombinant parents of tomato viruses but not vice versa. Together, these results indicate that tomato-infecting begomoviruses in Brazil are evolving rapidly within sharply divided subpopulations and stress the need for a management strategy that should include, but must go beyond, deployment of resistant cultivars.

MATERIALS AND METHODS

Sample collection and storage. Foliar samples with virus-like symptoms such as yellow mosaic, leaf curl, and down-cupping were collected in tomato fields located near the cities of Paty do Alferes, Rio de Janeiro (RJ) state, Brazil (May 2005), Coimbra, Minas Gerais (MG) state, Brazil (July 2007), Florestal, Carandaí, and Jaíba, MG, Brazil (July 2008), and Viçosa, MG, Brazil (May 2010). For each sample the following information was recorded: plant species (noncultivated plants) or cultivar/hybrid (for the tomato samples), date of collection, GPS coordinates of the sampling location, and symptoms (description and digital image of the sample at the time of collection). Samples were either stored in an ultrafreezer ($-80°C$) as desiccated foliar material or press dried and stored at room temperature as herbarium-like samples until analyzed.

DNA amplification and cloning. Total DNA was extracted as described by Doyle and Doyle (41) and was used as a template for rolling-circle amplification (RCA) of viral genomes (42). RCA products were cleaved with ApaI, BamHI, ClaI, EcoRI, HindIII, KpnI, PstI, SacI, or SpeI and ligated to the pBluescript KS+ (Stratagene) plasmid vector, previously cleaved with the same enzyme. RCA products also were cleaved with MspI to check for the presence of satellite-like DNA molecules. Viral inserts were sequenced commercially (Macrogen Inc., Seoul, South Korea) by primer walking. All genome sequences were organized to begin at the nicking site in the invariant nonanucleotide at the origin of replication (5′-TAATATT//AC-3′).

Sequence comparisons and phylogenetic analysis. Sequences were initially analyzed with the BLAST*n* algorithm (43) to determine the viral species with which they shared greatest similarity. Specific sets of sequences were then prepared for each analysis that was performed. Besides the sequences determined in this study, reference sequences for each begomovirus from Brazil and selected begomoviruses from the Americas were retrieved from GenBank (see Table S1 in the supplemental material).

Multiple nucleotide sequence alignments used for recombination and phylogenetic analyses were prepared using the MUSCLE program (44). Phylogenies for each data set were reconstructed using maximum-likelihood (ML) and Bayesian approaches.

ML trees were inferred using PAUP v. 4.0 (45). The program MODELTEST v. 3.7 (46) was used to provide the nucleotide substitution model with the best fit for each data set. A heuristic search was initiated with a neighbor-joining tree using the tree-bisection-reconnection (TBR) algorithm to optimize the ML tree. The robustness of each internal branch was estimated using a nonparametric test (47) with 1,000 bootstrap replications. The Nearest Neighbor Interchange (NNI) algorithm was used to optimize the bootstrap replications of the ML tree.

Additional phylogenetic trees were constructed using Bayesian inference performed with MrBayes v. 3.0b4 (48), with the model selected by MrModeltest v. 2.2 (49) by the Akaike Information Criterion (AIC). The analyses were carried out running 10,000,000 generations and excluding the first 2,000,000 generations as burn-in. Trees were visualized using the TreeView program (50) and edited using CorelDraw X3.

A species tree was reconstructed using BEAST v. 1.6.1 (51), based on coding sequences only (coat protein [CP], replication-associated protein [Rep], *trans*-activating protein [Trap], and replication enhancer protein [Ren] open reading frames [ORFs]). MrModeltest v. 2.2 (49) was used to infer the best-fitting model of nucleotide substitution for each data set in the Akaike Information Criterion (AIC). On the basis of the results of Duffy and Holmes (30), we assumed a constant population size and a log-normal relaxed molecular clock. The Markov Chain Monte Carlo (MCMC) simulation was run for 100,000,000 generations and sampled at every 10,000 steps to produce a posterior tree distribution containing 10,000 trees. The maximum clade credibility tree was made using TreeAnnotator v. 1.6.1 and discarding the first 25% of the MCMC chains as burn-in with a posterior probability limit of 0.5. The maximum clade credibility tree thus obtained was edited in FigTree v. 1.3.

Recombination analysis. Evidence of non-tree-like evolution (possibly indicative of recombination) was initially assessed using the Neighbor-Net method (52) implemented in the program SplitsTree4 v. 4.10 (53). Parental sequences and recombination breakpoints were then determined using Recombination Detection Program (RDP) v. 3.44 (54). The

TABLE 1 Viruses cloned from the tomato samples collected at five different locations in the states of Rio de Janeiro and Minas Gerais, Brazil[a]

Location	No. of clones (DNA-A and DNA-B) obtained for indicated virus									
	ToCmMV	ToCMoV	ToSRV	ToYVSV	ToLDV	ToMlMV	TMoLCV	ToYSV	SimMV	Total
Carandaí		6	26							32
Coimbra	19									19
Florestal		20	5							25
Jaíba	1		2				2	1	1	7
Paty do Alferes	19			39	2	7			3	70
Total	39	26	33	39	2	7	2	1	4	153

[a] ToCmMV, *Tomato common mosaic virus*; ToCMoV, *Tomato chlorotic mottle virus*; ToSRV, *Tomato severe rugose virus*; ToYVSV, *Tomato yellow vein streak virus*; ToLDV, *Tomato leaf distortion virus*; ToMlMV, *Tomato mild mosaic virus*; TMoLCV, *Tomato mottle leaf curl virus*; ToYSV, *Tomato yellow spot virus*; SimMV, *Sida micrantha mosaic virus*.

analyses were performed with default settings and a Bonferroni-corrected *P*-value cutoff of 0.05. Only the recombination events detected by more than four of the seven methods implemented in RDP were considered to be reliable.

General descriptors of the genetic structure of viral populations. Partition of genetic variability and inferences about population structure were based on Wright's F fixation index (55). Molecular variance analysis (AMOVA) was performed to estimate the Φ_{ST} parameter, using the program Arlequin v. 3.11 (56) with distance determined using the Kimura 2-parameter nucleotide substitution model and statistical significance estimated by a permutation test with 1,000 replications.

The program Structure v. 2.3.1 was used to test for evidence of genetic structure among subpopulations and to identify individuals that were admixed or had migrated among populations. To select the number of clusters that best represented population structure, one run of 1 to 10 subpopulations ($K = 1$ to 10) was performed using 1,000,000 Markov chain steps after a burn-in period of 100,000 steps. We compared the likelihood estimates for each K value based on maximum log probability of data ln $P(D)$, assayed to determine the best-supported number of subdivisions present in (i.e., the best-supported K values for) the populations.

The main descriptors of molecular variability were estimated for each population/subpopulation, including total number of segregating sites (*s*), total number of mutations (Eta), average number of nucleotide differences between sequences (*k*), nucleotide diversity (π), mutation frequencies, number of haplotypes (h), haplotype diversity (Hd), and Watterson's estimate of the population mutation rate based on the total number of segregating sites (θ-w) and on the total number of mutations (θ-Eta). Diversity indices were calculated using DnaSP v. 5.10 software (57).

Parameterization of evolutionary mechanisms. Three types of neutrality tests were used to test the hypothesis of selection occurring in populations: Tajima's *D* and Fu and Li's *D** and *F**. These analyses were performed using DnaSP v. 5.10, with different sets of data considering unique populations or subpopulations separated on the basis of geographical location.

Detection of positive and negative selection at amino acid sites. To detect sites under positive and negative selection, we analyzed the CP, Rep, Trap, and Ren data sets using three different ML-based methods implemented in DataMonkey (http://www.datamonkey.org): single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), and random-effects likelihood (REL) (58). These methods were applied using the best-fit model of nucleotide substitution determined in MODELTEST (46) by the Akaike Information Criterion (AIC). Bayes factors larger than 50 and *P* values smaller than 0.1 were used as thresholds for the REL and FEL methods. Analyses were performed on data sets comprised of nucleotide sequences of each ORF for all isolates of all species such that these analyses primarily amounted to an interspecific test of synonymous versus nonsynonymous substitution rates.

RESULTS

Viral detection and sequence comparisons. A total of 432 samples (326 tomato and 106 noncultivated species) were collected, and 219 (143 tomato and 76 noncultivated species) were preliminarily positive for the presence of a begomovirus, based on the detection of an ~2,600-bp band after digestion of the RCA products with restriction enzymes (data not shown). The use of nine different enzymes ensured that no virus(es) was left uncloned because it lacked one or more of the sites that were used (i.e., none of the samples had undigested forms for all enzymes). From these samples, 126 full-length DNA-A and 58 DNA-B components were cloned (see Table S2 in the supplemental material; that table lists only the samples from which full-length components were cloned). BLASTn analysis and pairwise sequence comparisons of cloned genome sequences with those deposited in GenBank indicated that each of the cloned genome sequences could be classified into one of the begomovirus species/tentative species previously described in Brazil. An exhaustive analysis of the RCA products after digestion with the restriction enzyme MspI (which has a 4-nucleotide recognition sequence) failed to identify fragments that added up to totals consistent with the presence of a satellite (i.e., a factor of 1.3 kb).

Among the tomato samples, four begomovirus species—ToCmMV, ToCMoV, ToSRV, and ToYVSV—accounted for 137 of 153 (90%) clones obtained (Table 1). However, one or two of these four different begomoviruses were most frequently cloned from samples collected at a given location: ToYVSV and ToCmMV in Paty do Alferes, ToCmMV in Coimbra (the only virus cloned from tomato samples collected at this location), and ToSRV and ToCMoV in Carandaí and Florestal (curiously, in opposite proportions at each location) (Table 1). Five other begomovirus species—SimMV, ToLDV, ToMlMV, ToMoLCV, and ToYSV—were cloned at a much lower frequency (Table 1). A small number of samples had detectable mixed infections with two viruses (see Table S2 in the supplemental material).

Eight begomoviruses were cloned from samples of noncultivated species (see Table S2 in the supplemental material), of which three (ToCMoV, ToMlMV, and ToSRV) were also cloned from the tomato samples. The most commonly detected virus from a noncultivated species was *Blainvillea yellow spot virus* (BlYSV) (15), which was cloned from all samples of *Blainvillea rhomboidea* and the single *Physalis* sp. sample (see Table S2 in the supplemental material).

Phylogenetic analyses. Phylogenetic relationships were ana-

lyzed based on complete DNA-A nucleotide sequences. A ML tree was constructed that included sequences of one isolate from each begomovirus species obtained in this study, together with representative sequences of all New World begomovirus species (Fig. 1). This analysis indicated that the Brazilian begomoviruses form seven clusters (Fig. 1, clusters I to VII). Clusters I to III contain mostly viruses originally isolated from cultivated hosts, and clusters IV to VII contain mostly viruses originally isolated from noncultivated hosts.

Of the nine species falling in clusters I to III, four (ToCMoV, TMoLCV, ToSRV, and ToYVSV) were detected in the samples analyzed here (Fig. 1; see also Table S2 in the supplemental material). Even though we have not performed infectivity studies comparing these viruses, it is reasonable to assume that the viruses in these three clusters are well adapted to cultivated plants, as ToCMoV, ToSRV, and ToYVSV were the species most frequently detected both here (Table 1) and in other studies (21, 24, 59), and *Soybean blistering mosaic virus* (SoBlMV) is widespread in soybean fields in northwestern Argentina (60). Although detected in only a small number of samples here, another member of this group, TMoLCV, is widespread in tomato fields in the Brazilian northeast (23).

A different situation is found in cluster IV, which contains mostly viruses from noncultivated hosts. The fact that ToLDV, ToMlMV, and ToYSV are included in this cluster (Fig. 1), coupled with the fact that these three species were detected in tomatoes at a very low frequency in our study, suggests that they could be "wild" viruses which have "spilled over" into tomatoes but are not as well adapted to tomatoes as many of the species in clusters I, II, and III. ToMlMV was detected in a single infection in three symptomatic *Sida urens* plants (see Table S2 in the supplemental material). Analysis of a larger number of samples of this host could indicate whether this plant species is the predominant "natural" host of this virus.

Cluster V includes two viruses which so far have been detected only in noncultivated hosts (BlYSV and *Sida mosaic Brazil virus* [SimBV]) and one virus (*Bean golden mosaic virus* [BGMV]) which is widespread in both noncultivated (e.g., *Macroptilium* spp. [61]) and cultivated hosts.

Clusters VI and VII, which contain EuYMV, SiYLCV, and ToCmMV, are part of a branch containing viruses from several countries in Central and North America (Fig. 1; see also Table S1 in the supplemental material). Thus, these three species are members of a lineage of New World begomoviruses which is distinct from other Brazilian begomoviruses. Interestingly, ToCmMV is one of the viruses which was most frequently found in our tomato samples (Table 1).

Bayesian inference was employed to reconstruct phylogenies of three viral populations for which location-based subdivision was suspected to occur: ToCmMV (including isolates from Coimbra, Jaíba, and Paty do Alferes), ToCMoV (including isolates from Carandaí and Florestal plus three previously sequenced isolates available in GenBank) and ToSRV (isolates from Carandaí, Florestal, Jaíba, and Viçosa plus five previously sequenced isolates available in GenBank) (Fig. 2). The results are consistent with the geographical subdivision hypothesis, as the isolates from all three species were clearly clustered according to sampling locations (Fig. 2).

Recombination analysis. Occurrence of non-tree-like evolution within populations of ToCmMV, ToCMoV, ToSRV, and ToYVSV, possibly attributable to recombination, was initially tested by Neighbor-Net analysis. The results did not indicate any significant evidence of deviation from tree-like evolution within the ToCmMV and ToYVSV populations (data not shown). However, clear cycles within the neighbor networks generated using ToCMoV and ToSRV sequences suggested that some of the analyzed sequences may have been recombinant (Fig. 3).

Strong evidence of non-tree-like evolution was also revealed by the Neighbor-Net analysis generated using all available Brazilian begomovirus sequences together with one sequence from each of the discrete begomovirus clusters detected in this study (Fig. 4). The strongest signals of non-tree-like evolution appeared to be associated with AbMBV, EuYMV, SiYLCV, ToCmMV, ToCMoV, ToRMV, and ToSRV sequences, the latter three previously having been identified as recombinants (19).

To further investigate evidence of recombination within these sequences and to locate recombination breakpoints and identify parental viruses, data sets that included either all Brazilian begomoviruses or all begomoviruses from the Americas, both of which included one isolate from each of the main begomovirus groups obtained in this study, were analyzed using the RDP3 package (Tables 2 and 3; data sets and RDP3 project files available on request). This analysis identified several unique recombination events, most of which involved breakpoints located within the Rep gene and the common region (CR), a finding consistent with previous studies which identified these regions as recombination hot spots (62–67). Interestingly, analysis based on both data sets indicated that a large number (12 of 19) of recombination events involving tomato viruses have viruses from noncultivated hosts as parents but not vice versa (the vast majority, 12 of 13, of the recombination events involving viruses from noncultivated hosts have other viruses from noncultivated hosts as parents). Although parental sequence identification is not always reliable (the probability of actual parental sequences being sampled in field studies is extremely low such that the parental sequences identified by RDP3 analyses are really just the sequences in the data set that most closely resemble the actual parental sequences), these results are an additional line of evidence indicating that tomato viruses probably evolved/emerged from viruses infecting noncultivated hosts. It is also consistent with the hypothesis that the genetic makeup of tomato-infecting viruses predominantly found in Brazil might be, at least, partially attributable to tomato-adapted viruses arising through recombination between species adapted to noncultivated hosts.

Genetic structure of viral populations. We focused on five species (BlYSV, ToCmMV, ToCMoV, ToSRV, and ToYVSV) for which sufficient sample sizes to carry out population genetics tests were available. The number of BlYSV DNA-A sequences was smaller than for the tomato viruses, but the number of DNA-B sequences was equivalent. Furthermore, since the BlYSV population was sampled in a smaller area, comparisons are valid.

In general, DNA-B sequences were more variable than DNA-A sequences for all five populations. For example, the average number of nucleotide differences between the seven BlYSV DNA-A sequences was 65.6, while for the seven DNA-B sequences it was 121.9 (Table 4). This was expected, as it has been demonstrated that the DNA-A and DNA-B components of bipartite begomoviruses can have distinct evolutionary histories, with the DNA-B being more variable (68).

Comparing the populations of each virus, BlYSV has a much
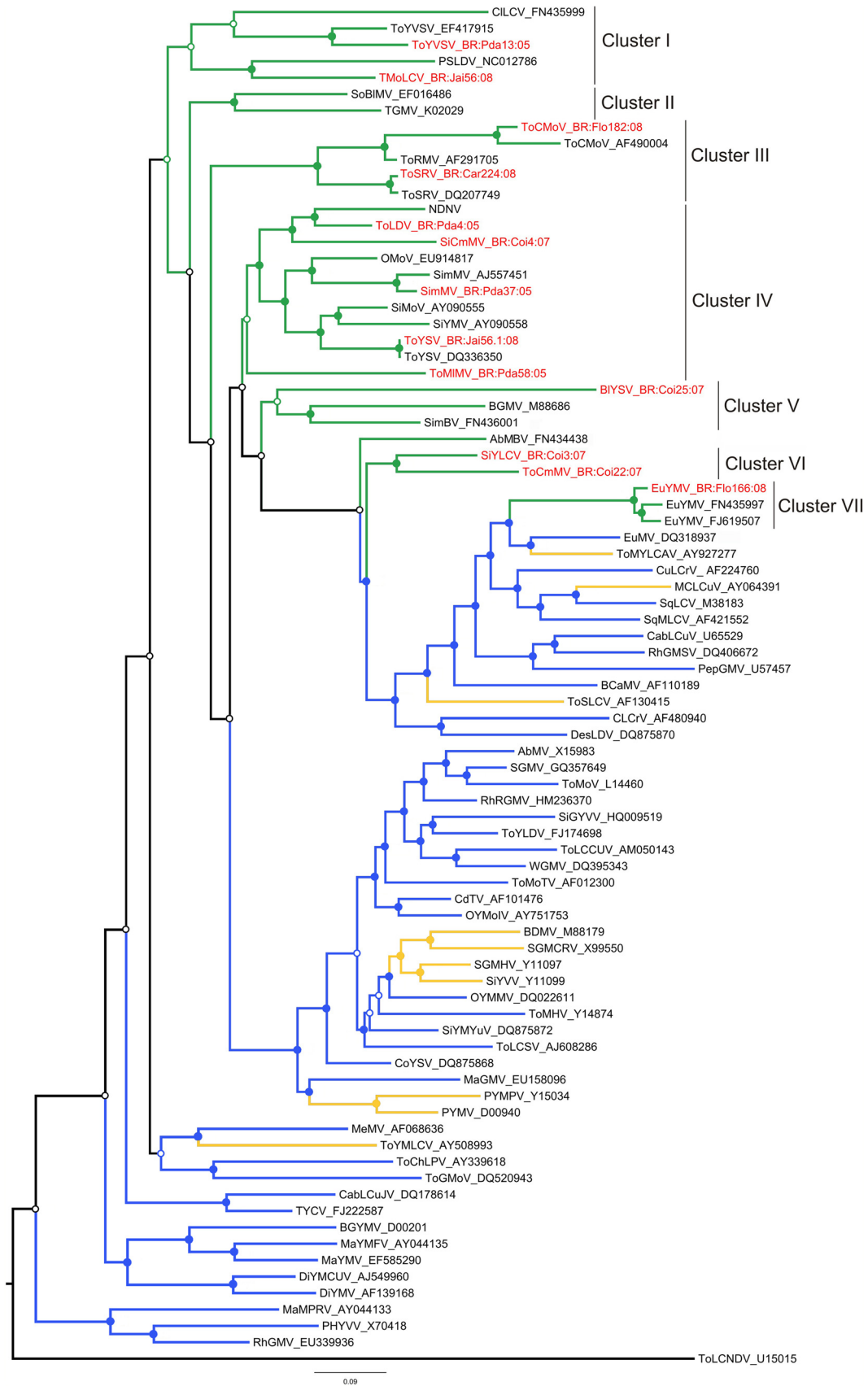
higher degree of genetic variability than the tomato viruses. For example, values for nucleotide diversity (DNA-A) are 0.02466 for BlYSV, 0.0143 for ToCmMV, 0.0071 for ToCMoV, 0.0102 for ToSRV, and 0.0021 for ToYVSV (Table 4). ToYVSV is the least diverse virus, with lower values for every descriptor.

Mutation frequencies were determined for the five populations (Table 4) and, except for ToCmMV subpopulations from Paty do Alferes, were found to be higher than those determined for other begomoviruses. The ToCmMV subpopulation from Paty do Alferes had a mutation frequency on the order of 10^{-4} for both the DNA-A and the DNA-B (Table 4). All other populations had mutation frequencies on the order of 10^{-3} (Table 4). Strikingly, the population of the noncultivated host-infecting BlYSV has an even higher frequency, on the order of 10^{-2} (Table 4). To verify if mutations are randomly distributed throughout the genome, we determined mutation frequencies for each coding region in the viral genome (CP, Rep, Trap, and Ren in the DNA-A, movement protein [MP] and nuclear shuttle protein [NSP] in the DNA-B) as well as in the intergenic (noncoding) common regions (CRs). In most cases, higher mutation frequencies were observed for the DNA-B than for the DNA-A, and for both DNA components the highest mutation frequencies were observed in the CR (Fig. 5).

Two of the five populations analyzed (ToCmMV and ToSRV) included enough DNA-A and DNA-B sequences from isolates collected at two locations to allow for a segregated analysis. For ToCmMV, both the DNA-A and DNA-B sequences could be divided into Coimbra and Paty do Alferes groups. The analysis indicated that the DNA-A and DNA-B sequences from Coimbra have a much greater genetic variability than the Paty do Alferes sequences (Table 4). In fact, values obtained for the Paty do Alferes group are equivalent to, and in many cases even lower than, those obtained for the ToYVSV population (which is comprised entirely of isolates from this same location). For ToSRV, DNA-A sequences could be divided into Carandaí and Florestal groups. The Florestal sequences are more diverse than the Carandaí ones, although in this case the result is not as clear due to the discrepancies in the sizes of the groups (19 sequences from Carandaí and only 5 from Florestal) (Table 4).

Differences between these groups of sequences in genetic variability indicate the existence of population subdivisions for both ToCmMV (Coimbra and Paty do Alferes) and ToSRV (Carandaí and Florestal). To verify subdivision of these populations and estimate variability within and among subpopulations, AMOVA and Fst and Nst tests were performed. Analyses of population differentiation using the Fst and Nst statistics for nucleotide diversity confirmed population subdivision for ToCmMV and ToSRV and also for ToCMoV (Table 5).

A cluster-based method (Structure) was used to identify individuals that were admixed or had migrated among tomato-infecting begomovirus populations as well as to infer the number of subpopulations making up each population (i.e., the optimal cluster number, or K, of each population). ToCMoV, ToCmMV,

ToSRV, and ToYVSV populations were estimated to be composed of three, two, two, and one subpopulations, respectively (Fig. 6), a result which is consistent with phylogenetic analyses of these species (Fig. 2). Two ToCmMV isolates from Coimbra (Coi20 and Coi21) appear as immigrants into the Paty do Alferes population (Fig. 2C, indicated by the arrows; Fig. 6B). Interestingly, ToSRV isolates from both Viçosa and Jaíba (which are located more than 600 km apart) group together with the Carandaí subpopulation, whereas at Florestal a completely distinct lineage of ToSRV was found (Fig. 6C). While Carandaí, Florestal, and Viçosa are relatively close (ca. 100 to 190 km apart), they are much (ca. 500 to 620 km) further away from Jaíba.

We could not perform the Structure analysis on BlYSV, since the number of sequences obtained was too small. To obtain such data for a virus infecting a noncultivated host, we used *Cleome leaf crumple virus* (ClLCrV), for which 16 complete DNA-A sequences are available: 15 from the state of Alagoas (69, 70) and 1 from the state of Mato Grosso do Sul (71). The analysis indicated the existence of two subpopulations, one comprised of isolates SC215 from Alagoas and FN435999 from Mato Grosso do Sul and the other comprised of the remaining isolates from Alagoas (Fig. 6E). It is noteworthy that Alagoas and Mato Grosso do Sul are more than 2,400 km apart.

Neutrality tests were used to assess what kind of selection is acting on the coding sequences of the BlYSV, ToCmMV, ToCMoV, ToSRV, and ToYVSV populations. Statistically significant deviations from neutrality were observed for different ORFs depending on the population being analyzed (Table 6). However, the absolute majority of these statistically significant values were negative, indicating that a large proportion of the genetic polymorphisms detected within these sequences are unique to individual sequences (as is reflected by the many multifurcating branches in the trees presented in Fig. 2). Such patterns of nucleotide diversity are expected (i) when most observable polymorphisms in a population are transient and are eventually removed by purifying selection operating at the nucleotide level or (ii) when a population expansion has occurred following a selective sweep that purged all but a few individuals. These results were confirmed by the detection of sites under negative/positive selection using three ML-based tests and taking recombination into account. For all ORFs, a much higher proportion of sites were detectably evolving under negative selection than were detectably under positive selection. For example, using the SLAC method (the most conservative of the three tests applied), 154 sites were found to be evolving under negative selection versus 0 sites evolving under positive selection for the CP, 163 versus 1 for Rep, 34 versus 4 for Trap, and 42 versus 4 for Ren (data not shown).

DISCUSSION

Begomoviruses became established in tomato crops in Brazil and other countries across South America after the introduction of the B biotype of *B. tabaci* in the mid-1990s. Since then, a large number

FIG 1 Maximum-likelihood tree based on the complete DNA-A nucleotide sequences of one isolate of each begomovirus obtained in this study (indicated in red font) plus sequences of all South American begomoviruses (green branches, including viruses from Brazil) and selected begomoviruses from Central America (orange branches), North America, and the Caribbean (blue branches). *Tomato leaf curl New Delhi virus* (ToLCNDV), an Old World bipartite begomovirus, was used as the outgroup. The seven clusters that include Brazilian begomoviruses are indicated at the right. Nodes to the right of branches with bootstrap support equal to or higher than 50% are indicated by filled circles and those with values lower than 50% by empty circles. See Table S1 in the supplemental material for full virus names.
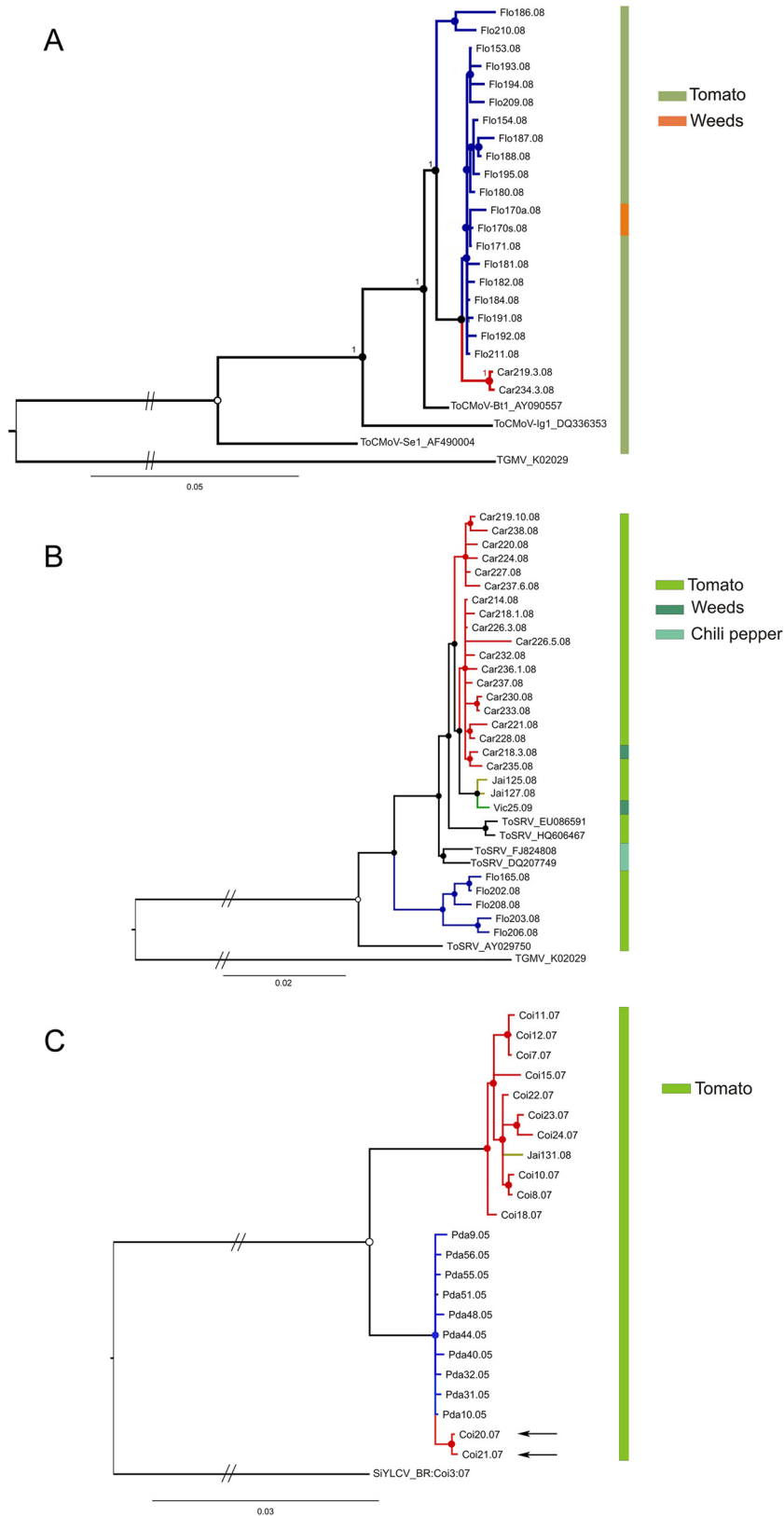
FIG 2 Bayesian 50% majority rule consensus tree based on the complete DNA-A nucleotide sequences of isolates of *Tomato chlorotic mottle virus* (ToCMoV) (A), *Tomato severe rugose virus* (ToSRV) (B), and *Tomato common mosaic virus* (ToCmMV) (C). Nodes to the right of branches with posterior probabilities equal to or higher than 0.5 are indicated by filled circles and those with values lower than 0.5 by empty circles. Branches with the same color indicate isolates collected at the same location. The color-coded vertical bars at the right of each tree indicate the host of each isolate.
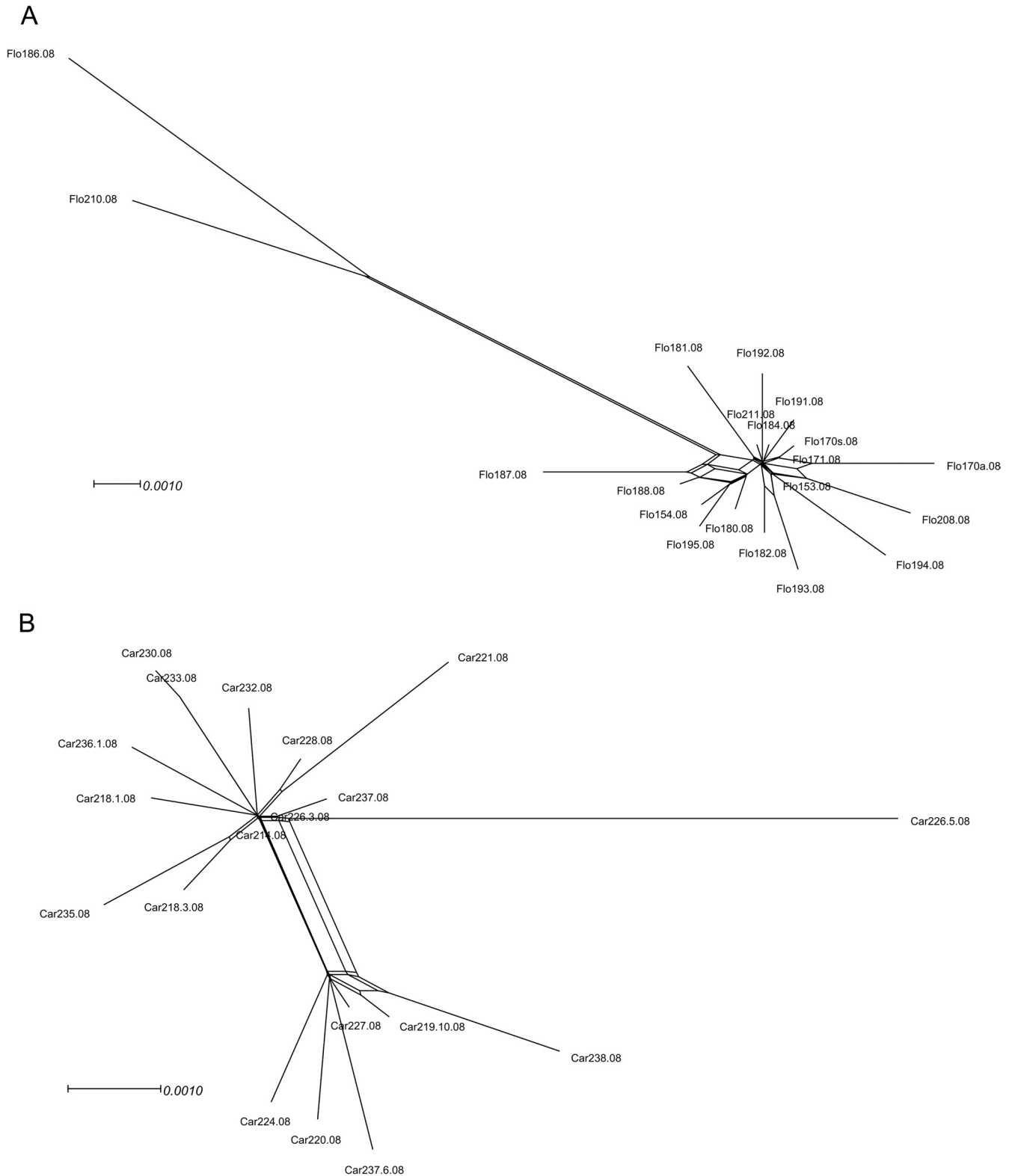
FIG 3 Phylogenetic evidence for recombination within populations of the begomoviruses *Tomato chlorotic mottle virus* (ToCMoV) (A) and *Tomato severe rugose virus* (ToSRV) (B). Neighbor-Net network analysis was performed using SplitsTree4. Formation of a reticular network rather than a single bifurcated tree is suggestive of recombination.
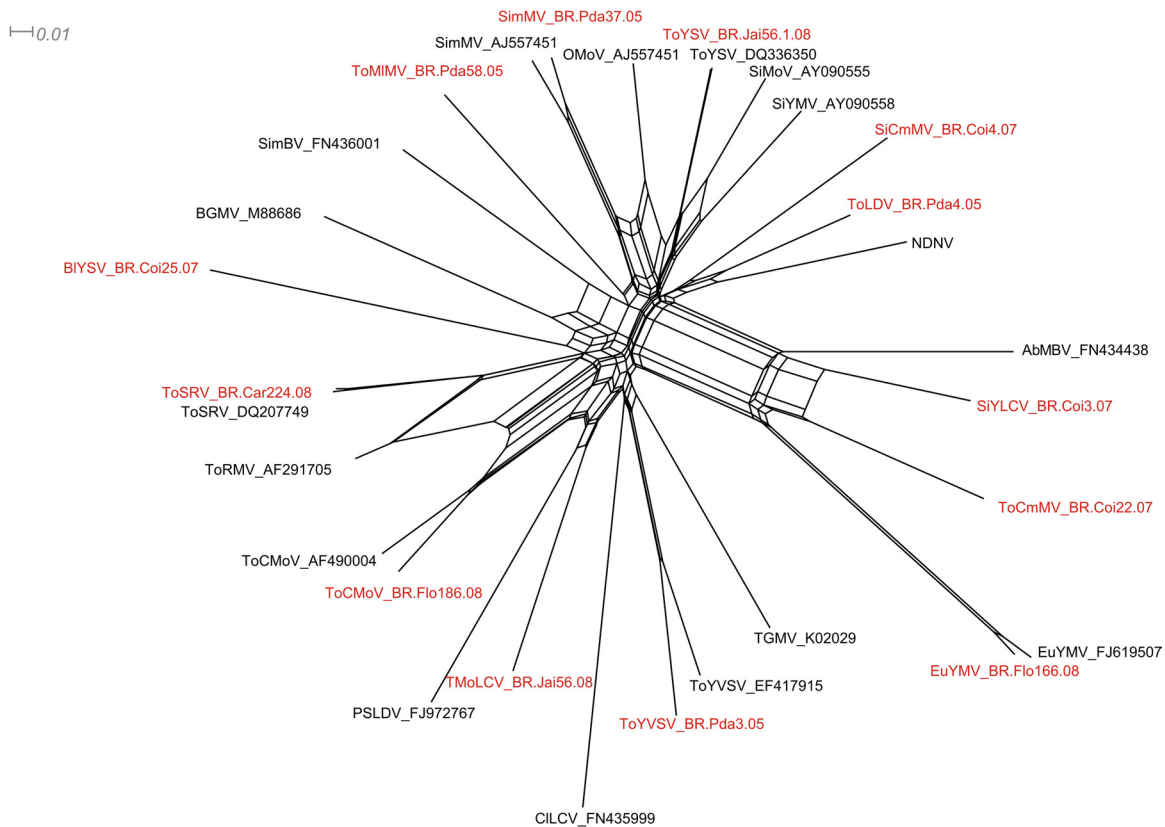
FIG 4 Phylogenetic evidence for recombination among begomoviruses from the Americas, including some of the isolates described in this study. Neighbor-Net network analysis was performed using SplitsTree4. Formation of a reticular network rather than a single bifurcated tree is suggestive of recombination. The isolates obtained in this study are indicated in red. See Table S1 in the supplemental material for full virus names.

of tomato-infecting begomovirus species have been described and characterized (10, 15, 18–20, 39). Despite intensive sampling and characterization of tomato-infecting viruses in North America, Central America, and the Caribbean, the Brazilian tomato-infect-

ing viral species never have been detected in these other regions, which strongly suggests that they are indigenous to South America and have emerged from noncultivated host species on this continent. Begomoviruses are notoriously prone to recombination (37,

TABLE 2 Putative recombination events detected among the tomato- and noncultivated host-infecting begomoviruses from Rio de Janeiro and Minas Gerais states, Brazil, based on a data set including only begomoviruses from Brazil

Virus (recombinant isolate)[a]	Recombination breakpoint[b]		Parent[c]		Methods[d]	P value[e]
	Initial	Final	Major	Minor		
EuYMV (Flo166:08)	1547	2537	*ClLCV*	Unknown	RGBMC**S**	9.3×10^{-18}
SiCmMV (Coi4:07)	1333	2285	*OMoV*	Unknown	RGBMC**S**3	8.2×10^{-15}
SiCmMV (Coi4:07)	1929	(?)	Unknown	*NDNV*	RGBMC**S**3	3.3×10^{-07}
SimMV (Pda37:05)	1619	2606	*OMoV*	Unknown	RBMC**S**3	3.2×10^{-21}
SimMV (Pda37:05)	1915	(?)	Unknown	*NDNV*	R**G**BMCS	3.3×10^{-07}
SiYLCV (Coi3:07)	66	1997	*EuYMV*	*SiMoV*	RGBMCS**3**	4.5×10^{-21}
ToCmMV (Coi22:07)	35	1856	*EuYMV*	*SiMoV*	RGBMCS**3**	4.5×10^{-21}
ToCmMV (Coi22:07)	1917	2516	Unknown	<u>*ToMlMV*</u>	RGBMC**S**	9.3×10^{-18}
ToLDV (Pda4:05)	(?)	921	*AbMBV*	*SimMV*	RGBMC**3**	1.7×10^{-07}
ToLDV (Pda4:05)	1582	2231	*OMoV*	Unknown	RGBMC**S**	8.2×10^{-15}
ToMlMV (Pda58:05)	1935	2635	<u>*ToCMoV*</u>	Unknown	RGBMC**S**	5.2×10^{-05}
TMoLCV (Jai13:08)	(?)	1159	*SiMoV*	<u>*ToSRV*</u>	RB**M**CS	6.2×10^{-04}
TMoLCV (Jai13:08)	1945	2330	Unknown	<u>*ToCMoV*</u>	R**G**BMCS	5.2×10^{-15}
ToSRV (Flo165:08)	1860	2179	<u>*ToCMoV*</u>	SimMV	R**G**BMCS3	1.7×10^{-14}
ToYSV (Jai56.1:08)	1898	2422	<u>*ToCMoV*</u>	Unknown	RGBMC**S**	5.2×10^{-05}
ToYVSV (Pda3:05)	831	1046	Unknown	<u>*ToSRV*</u>	RBMC**S**	1.3×10^{-06}

[a] For simplicity, only one isolate of each species is listed for each recombination event. See Table S1 in the supplemental material for full virus names.

[b] Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise. (?), breakpoint could not be precisely pinpointed.

[c] Tomato virus designations are underlined; designations of viruses from noncultivated hosts are indicated in italics.

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

[e] The reported P values are for the programs indicated in bold, underlined type in column 6 and are the lowest P values calculated for the region in question.

TABLE 3 Putative recombination events detected among the tomato- and noncultivated host-infecting begomoviruses from Rio de Janeiro and Minas Gerais states, Brazil, based on a data set including all begomoviruses from the Americas

Virus (recombinant isolate)[a]	Recombination breakpoint[b]		Parent[c]		Methods[d]	P value[e]
	Initial	Final	Major	Minor		
BlYSV (Vic04.1:09)	2099	2283	*RhGMV*	Unknown	RGBMC**S**	4.6×10^{-12}
EuYMV (Flo166:08)	1753	2519	*ToYMLCV*	Unknown	RGBMC**S**	1.1×10^{-20}
SiCmMV (Coi4:07)	203	614	AbMBV	SiMoV	RBMC**S**3	7.2×10^{-05}
SiCmMV (Coi4:07)	70	1445	*MaGMV*	SiMoV	R**B**MCS3	2.3×10^{-10}
SimMV (Pda37:05)	181	1619	OMoV	Unknown	RGBMCS**3**	2.7×10^{-19}
SimMV (Pda37:05)	1976	2234	*OYMolV*	ToRMV	R**G**BMC3	1.5×10^{-13}
SiYLCV (Coi3:07)	28	1999	SiMoV	*CabLCuV*	RGBMCS**3**	5.6×10^{-21}
ToCMoV (Flo182:08)	1790	2551	BGMV	PSLDV	RGBMC**S**3	6.7×10^{-20}
ToCmMV (Coi22:07)	1783	2501	SiMoV	*CabLCuV*	RGBMCS**3**	5.6×10^{-21}
ToLDV (Pda4:05)	270	929	AbMBV	SimMV	RGBMC**S**3	3.9×10^{-07}
ToLDV (Pda4:05)	592	1845	*MaGMV*	SiMoV	R**B**MC**S**3	2.3×10^{-10}
ToMlMV (Pda58:05)	502	2050	*RhGMV*	*TSLCV*	RGBMC**S**3	1.2×10^{-12}
TMoLCV (Jai13:08)	1608	2353	Unknown	PSLDV	RGBMC**S**	2.7×10^{-11}
ToSRV (Car224:08)	968	1907	*RhGMV*	*TSLCV*	RGBMC**S**3	1.2×10^{-12}
ToYSV (Jai56.1:08)	2561	(?)	SoBlMV	SiBV	RGM**C**3	4.8×10^{-08}
ToYVSV (Pda3:05)	(?)	2200	TGMV	Unknown	RGBMC**S**	7.8×10^{-05}

[a] For simplicity, only one isolate of each species is listed for each recombination event. See Table S1 in the supplemental material for full virus names.
[b] Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise. (?), breakpoint could not be precisely pinpointed.
[c] Designations of viruses from Brazil are underlined; designations of viruses from other countries in the Americas are indicated in italics.
[d] R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimera; S, SisScan; 3, 3SEQ.
[e] The reported P values are for the programs indicated in bold, underlined type in column 6 and are the lowest P values calculated for the region in question.

63, 66, 72) and have mutation rates (the basal rate at which mutations occur) and substitution rates (the rate at which mutations are retained within populations) comparable to those of RNA viruses (32, 73). The results described here, based on sequences of more than 200 viral genome components cloned from samples

collected over a 5-year period, provide further support for the hypothesis that tomato-infecting begomoviruses from Brazil evolved from indigenous viral populations infecting noncultivated hosts and indicates that the emergence of these viruses as tomato pathogens was probably facilitated by high rates of muta-

TABLE 4 Genetic structure of *Blainvillea yellow spot virus*, *Tomato chlorotic mottle virus*, *Tomato common mosaic virus*, *Tomato severe rugose virus*, and *Tomato yellow vein streak virus* populations from Rio de Janeiro and Minas Gerais states, Brazil[a]

Population	No. of sequences	Genome size (nt)	s	Eta	k	π	Mutation frequency	h	Hd	θ-w	θ-Eta
DNA-A											
BlYSV (Viçosa)	7	2,661	200	211	65.6	0.0247	1.1×10^{-2}	7	1.000	0.0307	0.0323
ToCmMV (Total)	22	2,560	103	104	36.6	0.0143	1.8×10^{-3}	20	0.987	0.0110	0.0111
Paty do Alferes	10	2,560	11	11	2.2	0.0009	4.3×10^{-4}	8	0.933	0.0015	0.0015
Coimbra	12	2,560	91	92	26.3	0.0103	3.0×10^{-3}	7	1.000	0.0312	0.0321
ToCMoV (Total)[b]	22	2,619	135	138	18.4	0.0070	2.4×10^{-3}	22	1.000	0.0141	0.0144
Florestal	20	2,619	120	122	16.6	0.0063	2.3×10^{-3}	20	1.000	0.0129	0.0131
ToSRV (Total)[c]	27	2,588	148	159	26.5	0.0102	2.3×10^{-3}	26	0.997	0.0148	0.0159
Carandaí	19	2,589	73	74	10.5	0.0040	1.5×10^{-3}	18	0.994	0.0080	0.0081
Florestal	5	2,592	37	37	19.0	0.0073	2.8×10^{-3}	5	1.000	0.0068	0.0068
ToYVSV (Paty do Alferes)	26	2,562	49	49	5.4	0.0021	7.4×10^{-4}	25	0.997	0.0050	0.0050
DNA-B											
BlYSV (Viçosa)	7	2,625	326	346	121.9	0.0464	1.9×10^{-2}	7	1.000	0.0506	0.0538
ToCmMV (Total)	16	2,500	205	214	60.5	0.0242	5.1×10^{-3}	14	0.975	0.0248	0.0258
Paty do Alferes	9	2,500	14	14	3.1	0.0012	6.2×10^{-4}	7	0.917	0.0021	0.0021
Coimbra	7	2,500	191	196	99.8	0.0400	1.1×10^{-2}	7	1.000	0.0312	0.0321
ToCMoV (Total)[d]	6	2,554	146	146	73.5	0.0288	9.5×10^{-3}	5	0.933	0.0250	0.0250
Carandaí	4	2,557	22	22	11.0	0.0043	2.2×10^{-3}	4	1.000	0.0046	0.0046
ToSRV (Carandaí)	7	2,568	50	50	15.9	0.0062	2.8×10^{-3}	7	1.000	0.0079	0.0079
ToYVSV (Paty do Alferes)	13	2,507	51	51	10.6	0.0042	1.6×10^{-3}	12	0.987	0.0066	0.0066

[a] BlYSV, *Blainvillea yellow spot virus*; ToCMoV, *Tomato chlorotic mottle virus*; ToCmMV, *Tomato common mosaic virus*; ToSRV, *Tomato severe rugose virus*; ToYVSV, *Tomato yellow vein streak virus*; s, total number of segregating sites; Eta, total number of mutations; k, average number of nucleotide differences between sequences (Tajima's estimate of the population mutation rate [θ[ρσθβ]]); π, nucleotide diversity; h, haplotype number; Hd, haplotype diversity; θ-w, Watterson's estimate of the population mutation rate based on the total number of segregating sites; θ-Eta, Watterson's estimate of the population mutation rate based on the total number of mutations.
[b] Data include two sequences from Carandaí.
[c] Data include two sequences from Jaíba and one sequence from Viçosa.
[d] Data include two sequences from Florestal.
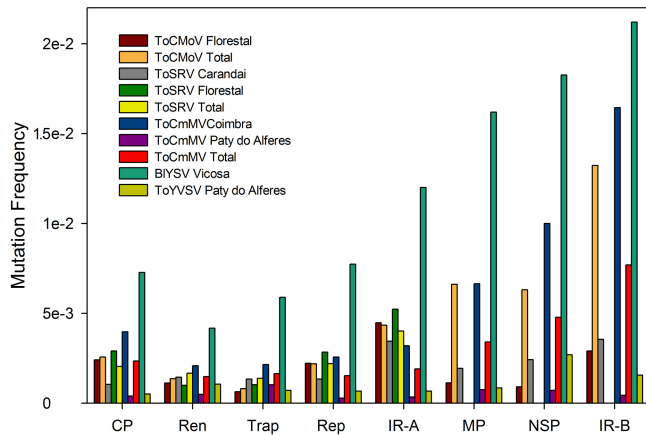
FIG 5 Mutation frequencies determined for each coding sequence and intergenic regions in the populations of *Blaivillea yellow spot virus* (BlYSV), *Tomato common mosaic virus* (ToCmMV), *Tomato chlorotic mottle virus* (ToCMoV), *Tomato severe rugose virus* (ToSRV), and *Tomato yellow vein streak virus* (ToYVSV). "ToCMoV Total" refers to the ToCMoV population from Florestal plus two isolates from Carandaí, "ToSRV Total" refers to the ToSRV populations from Carandaí and Florestal, and "ToCmMV Total" refers to the ToCmMV populations from Coimbra and Paty do Alferes. CP, coat protein; Ren, replication enhancer protein; Trap, *trans*-activating protein; Rep, replication-associated protein; MP, movement protein; NSP, nuclear shuttle protein; IR, intergenic or common region.

tion and recombination. This process has yielded a variety of viral species and lineages, but, interestingly, some of these species are much more prevalent in tomatoes in the field than others (21, 23, 24, 59), suggesting various degrees of adaptation to infect this host and/or to be transmitted by the B biotype of *B. tabaci*. For example, species such as ToCMoV, ToYVSV, and ToSRV have been detected at high prevalence in tomatoes in several field surveys (21, 24, 59), while others such as ToYSV, ToLDV, and ToMlMV have rarely been found infecting this host.

We began our analysis by identifying begomovirus species that were present and prevalent in each region analyzed. Minas Gerais state spans over 586,000 sq km, such that some of the sampling locations within the state were separated by over 600 km (moreover, Paty do Alferes in Rio de Janeiro state and Jaíba are over 780 km from each other). Confirming previous observations (21, 23, 24, 59), viruses belonging to only four species (ToCmMV, ToCMoV, ToSRV, and ToYVSV) accounted for more than 95% of the analyzed samples. Interestingly, the prevalences of each virus differed greatly among sampling locations. For example, whereas ToCmMV was the only begomovirus species detected in Coimbra, it was one of two predominant species in Paty do Alferes and was not found at other locations. ToYVSV was the other prevalent species in Paty do Alferes but was not found elsewhere. ToCMoV and ToSRV were the predominant species in Carandaí and Florestal. Thus, at any given geographical site across Brazil, different combinations of viruses are expected to predominate. Indeed, previous surveys carried out in the states of São Paulo (59, 74), Minas Gerais/Rio de Janeiro/Espírito Santo (24), Goiás (21), and Bahia/Pernambuco (23) indicated the prevalence of ToYVSV, ToCMoV, ToSRV, and ToMoLCV, respectively. These differences could reflect local variations in noncultivated host distributions at the sampling sites or the capacity of local, non-biotype B vector populations to transmit particular virus species. Alternatively, the

epidemiology of tomato-infecting viruses might be highly erratic, with, at any particular site, the predominant virus species that cause diseases in tomato changing from year to year due to introductions from other sites. The latter possibility might explain the extremely low degrees of genetic variability observed in the two begomovirus species (ToCmMV and ToYVSV) sampled in Paty do Alferes. The combined observations that the ToCmMV population from Coimbra has a higher degree of variability, that two isolates from Coimbra were identified as being very closely related to those sampled in Paty do Alferes, and that neither of these species was detected in tomatoes in Paty do Alferes in 1999 (24) all provide additional support for the hypothesis of a recent introduction of these species to Paty do Alferes. Obviously, once a given virus emerges and becomes established in a particular location, there is no particular reason why it should stay restricted to that location. Tomatoes are often transported over long distances in Brazil, with few or no interstate sanitary inspections (and, of course, no "in-state" inspections). Thus, it will be interesting to monitor prevalence of these viruses at each location over time. If the location-based segregation of virus species reported here is maintained over time, this would better support the hypothesis that variations in host noncultivated species and/or vector population transmission efficiencies are responsible for the differences in the species composition of tomato-infecting begomovirus populations. One additional layer of complexity is that different cultural practices are employed by growers of fresh-market and processing tomatoes. As pointed out by Fernandes et al. (21), these cultural practices, especially the fact that fresh-market tomatoes are grown year-round while processing-tomato growers adopt a 2-month tomato-free period, could greatly influence disease epidemiology. Year-round cultivation could allow the viruses to persist in the environment with no need for alternative hosts to act as reservoirs. Conversely, a tomato-free period could break the cycle of transmission from tomato to tomato, and a weed reservoir would therefore be essential to restart the cycle in the next growing season.

TABLE 5 Results of subdivision tests performed on the populations of *Tomato common mosaic virus*, *Tomato chlorotic mottle virus*, and *Tomato severe rugose virus* from Rio de Janeiro and Minas Gerais states, Brazil[a]

Population	Nst	Fst
ToCmMV (DNA-A)		
Paty do Alferes/Coimbra	0.742	0.741
ToCmMV (DNA-B)		
Paty do Alferes/Coimbra	0.358	0.358
ToCMoV (DNA-A)		
Carandaí/Florestal	0.638	0.640
ToCMoV (DNA-B)		
Carandaí/Florestal	0.958	0.958
ToSRV (DNA-A)	0.738	0.735
Carandaí/Jaíba	0.502	0.502
Carandaí/Florestal	0.743	0.740
Jaíba/Florestal	0.793	0.791

[a] ToCmMV, *Tomato common mosaic virus*; ToCMoV, *Tomato chlorotic mottle virus*; ToSRV, *Tomato severe rugose virus*. Values from 0 to 0.05 indicate little genetic differentiation; values from 0.05 to 0.15 indicate moderate differentiation; values from 0.15 to 0.25 indicate great differentiation; values > 0.25 indicate high differentiation.
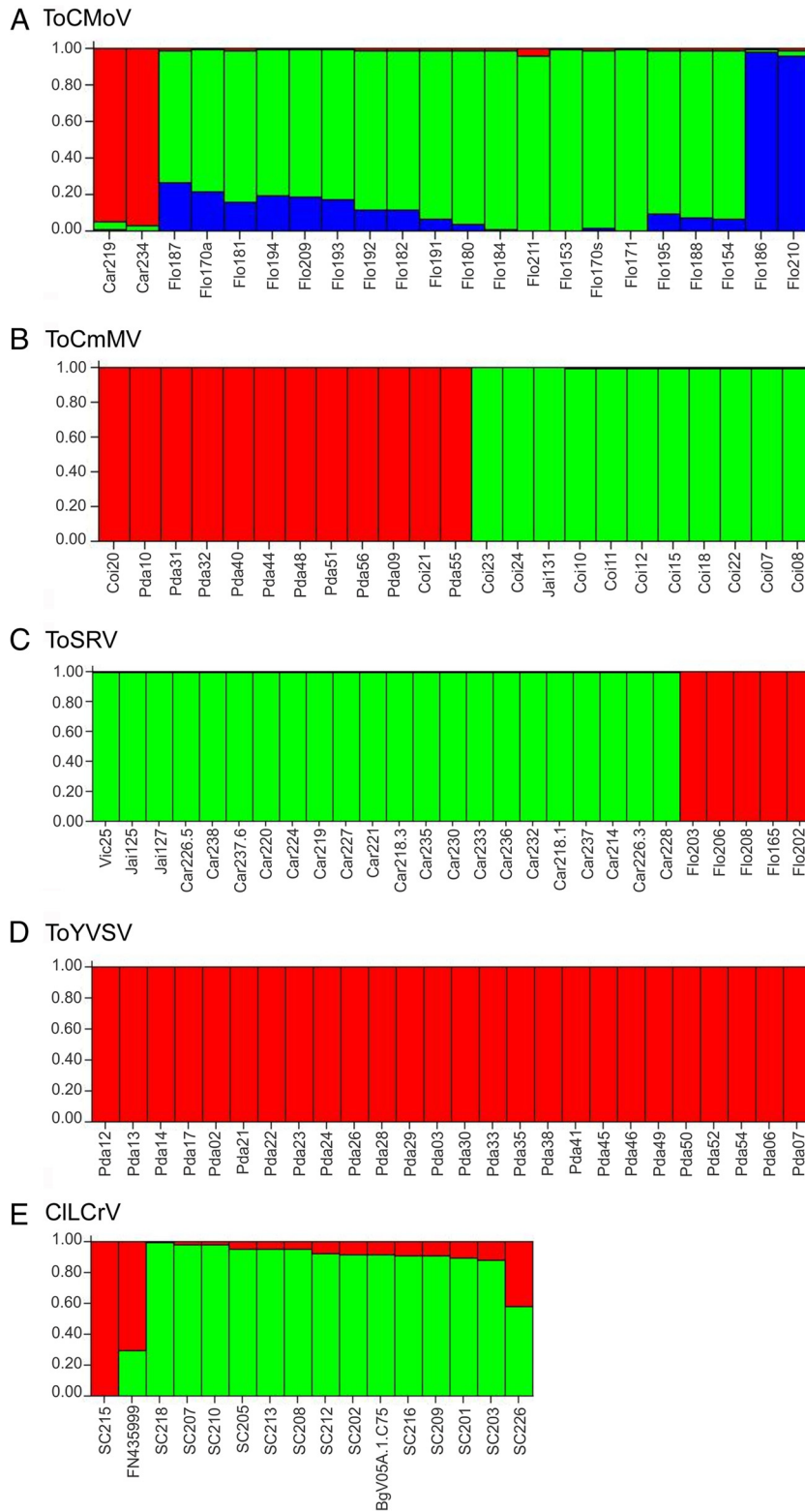
FIG 6 Cluster analysis of population subdivision using Structure. Each individual is represented by a vertical line divided into *K* colors, where *K* is the number of clusters assumed. Individuals are sorted according to *Q*. *Tomato chlorotic mottle virus* (ToCMoV), *K* = 3. *Tomato common mosaic virus* (ToCmMV), *K* = 2. *Tomato severe rugose virus* (ToSRV), *K* = 2. *Tomato yellow vein streak virus* (ToYVSV), *K* = 1. Car, Coi, Flo, Pda, and Vic correspond to isolates from Carandaí, Coimbra, Florestal, Paty do Alferes, and Viçosa, respectively. *Cleome leaf crumple virus* (ClLCrV), *K* = 2. Isolate names are according to references 69, 70, and 71.

TABLE 6 Results of the different neutrality tests for each open reading frame in the DNA-A and DNA-B of viral isolates comprising populations of *Blainvillea yellow spot virus*, *Tomato chlorotic mottle virus*, *Tomato common mosaic virus*, *Tomato severe rugose virus*, and *Tomato yellow vein streak virus* from Rio de Janeiro and Minas Gerais states, Brazil[a]

Population	ORF	Tajima's D	Fu and Li's D*	Fu and Li's F*
BlYSV (Viçosa)	CP	−1.517§	−1.458	−1.608
	Rep	−1.296	−1.306	−1.429
	Trap	−0.799	−0.825	−0.889
	Ren	−1.160	−1.161	−1.253
	NSP	−1.197	−1.131	−1.270
	MP	−0.505	−0.298	−0.382
ToCmMV (Paty do Alferes)	CP	−1.562	−1.784	−1.934
	Rep	−1.562	−1.784	−1.934
	Trap	−1.667	−1.916	−2.076
	Ren	−1.401	−1.587	−1.719
	NSP	−1.677	−1.881	−2.039
	MP	−1.728§	−1.943§	−2.107
ToCmMV (Coimbra)	CP	−0.604	0.993	0.654
	Rep	−0.772	0.259	−0.013
	Trap	0.027	1.036	0.881
	Ren	0.065	1.433§	1.226
	NSP	1.816	1.287	1.556
	MP	1.533	1.241	1.445
ToCMoV (Florestal)	CP	−1.973§	−2.2557	−2.529
	Rep	−2.182‡	−2.836§	−3.078†
	Trap	−1.608	−2.012	−2.191
	REn	−1.612	−0.687	−1.101
ToSRV (Carandaí)	CP	−2.282‡	−3.168†	−2.282†
	Rep	−2.302‡	−2.282†	−2.282†
	Trap	−2.010§	−2.326	−2.587§
	REn	−2.059§	−2.456§	−2.712§
	NSP	−1.314	−1.310	−1.443
	MP	−0.999	−0.904	−1.016
ToSRV (Florestal)	CP	0.708	0.708	0.749
	Trap	0.243	0.243	0.239
	REn	1.459	1.459	1.431
	Rep	0.812	0.812	0.865
ToYVSV (Paty do Alferes)	CP	−2.135§	−2.723§	−2.969§
	Rep	−2.156†	−3.341†	−3.483†
	Trap	−1.369	−2.042	2.143
	REn	−1.706	−2.400	−2.560
	NSP	−1.495	−1.917	−2.057
	MP	−1.830§	−2.060	−2.272

[a] BlYSV, *Blainvillea yellow spot virus*; ToCMoV, *Tomato chlorotic mottle virus*; ToCmMV, *Tomato common mosaic virus*; ToSRV, *Tomato severe rugose virus*; ToYVSV, *Tomato yellow vein streak virus*; CP, coat protein; Rep, replication-associated protein; Trap, *trans*-activating protein; Ren, replication enhancer protein; NSP, nuclear shuttle protein; MP, movement protein; §, significant value that rejects the null hypothesis of selective neutrality ($P < 0.05$); †, significant value that rejects the null hypothesis of selective neutrality ($P < 0.02$); ‡, significant value that rejects the null hypothesis of selective neutrality ($P < 0.01$).

Recombination has contributed greatly to diversification of begomoviruses (63, 75) and has probably dramatically increased their evolutionary potential, particularly in the context of adaptation to new host species and vector biotypes (38, 63, 76, 77). Con-firming previous reports of frequent interspecies recombination among Brazilian begomoviruses (19, 25), we found widespread evidence of interspecies recombination among the viruses detected in this work. Most recombination events occurred at the N-terminal region of the *Rep* gene, the common region, and the intergenic region between the *CP* and *Ren* genes (Tables 2 and 3), all of which have been reported to be recombination hot spots (62, 64, 66, 67). Interestingly, whereas most recombination events detected in tomato-infecting begomoviruses had viruses from noncultivated hosts identified as putative parents, those detected among viruses infecting a noncultivated host did not have tomato-infecting viruses as parents.

The relevance of recombination notwithstanding, geminivirus evolution in general, and the emergence of novel virus species and lineages in particular, is ultimately driven by the occurrence and fixation of point mutations (73, 78). Especially relevant in this regard is that begomoviruses and other single-stranded DNA (ssDNA) viruses display basal rates of mutation (32) and substitution (30, 31) that are as high as those found in rapidly evolving RNA viruses. Although the mutation frequencies calculated for Brazilian begomoviruses cannot be directly compared with the mutation rates determined by Ge et al. (32) or the substitution rates determined by Duffy and Holmes (30, 31), it is noteworthy that the mutation frequency determined for the noncultivated host-infecting BlYSV is 1 order of magnitude higher than those determined for the tomato-infecting viruses ToCMoV, ToCmMV, ToSRV, and ToYVSV (Table 4). One possible explanation is that intraspecies begomovirus variability reflects the genetic variability of the host. We have recently observed similar mutation frequencies in a population of *Macroptilium yellow spot virus* (MaYSV), a novel bipartite begomovirus species described as infecting the ubiquitous noncultivated host *Macroptilium lathyroides* in northeastern Brazil (79). Further studies analyzing mutation frequencies of begomovirus populations (ideally of the same virus) infecting cultivated and noncultivated hosts must be carried out to verify this hypothesis.

New World begomoviruses have a bipartite genome, with a clear "division of labor" between the two DNA components: the DNA-A encodes all replication-related functions as well as the coat protein (which is also the only viral protein necessary for whitefly transmission), while the DNA-B encodes the movement-related functions. The DNA-B of all the viruses analyzed in our work was more variable than the DNA-A. For example, in BlYSV, seven DNA-A sequences have a total number of mutations equal to 211, while the corresponding value for the same number of DNA-B sequences is 326 (Table 4). This fact could be attributed to the nonspecific nature of the movement functions carried out by the DNA-B-encoded proteins, which would thus be more permissive of changes (80–82). An alternative explanation would be that the DNA-B had a separate origin from the DNA-A, possibly as a satellite that was captured by a parent monopartite virus and later evolved to become an integral part of the genome, as suggested by Nawaz-ul-Rehman and Fauquet (16) and Briddon et al. (68). These two hypotheses are not mutually exclusive; combined effects could lead to the greater variability in the DNA-B than in the DNA-A.

Both our phylogenetic and population structure analyses indicated that the ToCmMV, ToCMoV, and ToSRV populations were subdivided according to geographical location. Structure analysis indicated the same to be true for a population of ClLCrV from a

noncultivated host. Neutrality tests and dN/dS-based selection analyses (i.e., examining ratios of nonsynonymous and synonymous substitutions) indicated that all of these populations were evolving in a significantly nonneutral manner, with a tendency toward pervasive purifying selection operating at both the nucleotide and protein levels. All ORFs of the five populations analyzed contained many more sites evolving under negative selection than sites evolving under positive selection. These results broadly recapitulate those indicating that negative selection and population expansion are the major evolutionary forces acting both on TLCV in *Eupatorium makinoi* (83) and on *Tomato spotted wilt virus* (TSWV) in peanut (84).

Interestingly, we detected viruses which are normally associated with noncultivated hosts (e.g., SimMV isolates BR:Pda8:05, BR:Pda37:05, and BR:Pda43:05) in tomato samples as well as viruses which are normally associated with tomato infecting noncultivated hosts (e.g., ToSRV isolates BR:Car228:08 and BR:Vic25: 10, both found in *Sida* sp.). This suggests that at least some of the viruses infecting noncultivated hosts can infect tomatoes, if only at a low frequency, and that begomoviruses well adapted to cultivated species such as tomato can infect noncultivated hosts under field conditions. Moreover, ToMlMV, a virus which clusters with viruses infecting noncultivated hosts (Fig. 1), was found in both tomato (at a very low frequency) and noncultivated hosts (see Table S2 in the supplemental material). Based on these observations, we suggest that this begomovirus, and possibly ToYSV, is actually a noncultivated host-adapted virus which is capable of producing spillover infections in tomatoes. Such spillover infections are widely believed to be a first step in the host-range-switching process (4, 38, 85, 86), and it is conceivable that lineages within these species could in the future emerge as tomato pathogens. Infectivity assays in tomato and noncultivated hosts are necessary to provide experimental support to this hypothesis.

The genetic structure of begomovirus populations is the consequence of a complex interplay between evolutionary adaptation to various plant species and vector biotypes and is largely shaped by the geographical distribution and dispersal patterns of these other organisms. Crucially, subdivisions that arise within virus populations due to uneven host and vector distributions and/or limits on dispersal rates of vectors strongly influence the potential for further evolution and speciation. Uncovering patterns of genetic variability that have arisen within begomovirus populations is, therefore, a necessary first step in determining how these populations arose and for estimating their potential to overcome control measures. We have assessed the genetic structure and variability of Brazilian begomovirus populations and found that viruses comprising these populations have the innate potential to evolve very rapidly in response to changing vector and host populations. It is thus evident that, inasmuch as resistance-based approaches must be actively sought to allow economically feasible and environmentally friendly control, such strategies should not be used in the absence of other management strategies. If the strategies are used in isolation, inbred resistance could potentially fail in the face of continuous long-term exposure to diverse rapidly mutating and recombining begomovirus populations. This is exactly what happened in Pakistan in the 2000s, when the massive deployment of cotton cultivars with resistance to *Cotton leaf curl Multan virus* led to the quick emergence of the resistance-breaking recombinant virus *Cotton leaf curl Burewala virus*, which is now prevalent in that country (87, 88). Complementary control measures, such

as insecticide use, crop rotation, and implementation of a tomato-free period, should be investigated for concurrent use with resistant genotypes to obstruct the evolution and spread of resistance-breaking viral lineages.

REFERENCES

1. **Rojas MR, Hagen C, Lucas WJ, Gilbertson RL.** 2005. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu. Rev. Phytopathol. **43**:361–394.
2. **Brown JK, Fauquet CM, Briddon RW, Zerbini FM, Moriones E, Navas-Castillo J.** 2012. Family *Geminiviridae*, p 351–373. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus Taxonomy. 9th Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, United Kingdom.
3. **Patil BL, Fauquet CM.** 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. Mol. Plant Pathol. **10**:685–701.
4. **Seal SE, Van den Bosch F, Jeger MJ.** 2006. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. Crit. Rev. Plant Sci. **25**:23–46.
5. **Hanssen IM, Lapidot M, Thomma B.** 2010. Emerging viral diseases of tomato crops. Mol. Plant Microbe Interact. **23**:539–548.
6. **Lefeuvre P, Martin DP, Harkins G, Lemey P, Gray AJ, Meredith S, Lakay F, Monjane A, Lett JM, Varsani A, Heydarnejad J.** 2010. The spread of tomato yellow leaf curl virus from the Middle East to the world. PLoS Pathog. **6**:e1001164. doi:10.1371/journal.ppat.1001164.
7. **Reddy RVC, Colvin J, Muniyappa V, Seal S.** 2005. Diversity and distribution of begomoviruses infecting tomato in India. Arch. Virol. **150**:845–867.
8. **Morales FJ, Jones PG.** 2004. The ecology and epidemiology of whitefly-transmitted viruses in Latin America. Virus Res. **100**:57–65.
9. **Polston JE, Anderson PK.** 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. Plant Dis. **81**:1358–1369.
10. **Ribeiro SG, Ambrozevícius LP, Avila AC, Bezerra IC, Calegario RF, Fernandes JJ, Lima MF, de Mello RN, Rocha H, Zerbini FM.** 2003. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. Arch. Virol. **148**:281–295.
11. **Nakhla MK, Maxwell DP, Martinez RT, Carvalho MG, Gilbertson RL.** 1994. Widespread occurrence of eastern Mediterranean "strain" of tomato yellow leaf curl geminivirus in tomatoes in the Dominican Republic. Plant Dis. **78**:926.
12. **Duffy S, Holmes EC.** 2007. Multiple introductions of the Old World Begomovirus Tomato yellow leaf curl virus into the New World. Appl. Environ. Microbiol. **73**:7114–7117.
13. **Holguin-Pena RJ, Vazquez-Juarez R, Rivera-Bustamante RF.** 2005. A new begomovirus causes tomato leaf curl disease in Baja California Sur, Mexico. Plant Dis. **89**:341.
14. **Fiallo-Olivé E, Martinez-Zubiaur Y, Rivera-Bustamante RF.** 2009. Tomato yellow leaf distortion virus, a new bipartite begomovirus infecting tomato in Cuba. Plant Pathol. **58**:785.
15. **Castillo-Urquiza GP, Beserra JE, Jr, Bruckner FP, Lima AT, Varsani A, Alfenas-Zerbini P, Murilo Zerbini F.** 2008. Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. Arch. Virol. **153**:1985–1989.
16. **Nawaz-ul-Rehman MS, Fauquet CM.** 2009. Evolution of geminiviruses and their satellites. FEBS Lett. **583**:1825–1832.
17. **Matyis JC, Silva DM, Oliveira AR, Costa AS.** 1975. Purification and morphology of tomato golden mosaic virus. Summa Phytopathol. **1**:267–275. (in Portuguese.)
18. **Fernandes JJ, Carvalho MG, Andrade EC, Brommonschenkel SH, Fontes EPB, Zerbini FM.** 2006. Biological and molecular properties of *Tomato rugose mosaic virus* (ToRMV), a new tomato-infecting begomovirus from Brazil. Plant Pathol. **55**:513–522.
19. **Ribeiro SG, Martin DP, Lacorte C, Simões IC, Orlandini DRS, Inoue-Nagata AK.** 2007. Molecular and biological characterization of *Tomato*

chlorotic mottle virus suggests that recombination underlies the evolution and diversity of Brazilian tomato begomoviruses. Phytopathology **97**:702–711.

20. **Calegario RF, Ferreira SS, Andrade EC, Zerbini FM.** 2007. Characterization of *Tomato yellow spot virus*, (ToYSV), a novel tomato-infecting begomovirus from Brazil. Braz. J. Agric. Res. **42**:1335–1343.

21. **Fernandes FR, de Albuquerque LC, de Britto Giordano L, Boiteux LS, de Avila AC, Inoue-Nagata AK.** 2008. Diversity and prevalence of Brazilian bipartite begomovirus species associated to tomatoes. Virus Genes **36**:251–258.

22. **Albuquerque LC, Martin DP, Avila AC, Inoue-Nagata AK.** 2010. Characterization of tomato yellow vein streak virus, a begomovirus from Brazil. Virus Genes **40**:140–147.

23. **Albuquerque LC, Varsani A, Fernandes FR, Pinheiro B, Martin DP, de Tarso Oliveira Ferreira P, Lemos TO, Inoue-Nagata AK.** 2012. Further characterization of tomato-infecting begomoviruses in Brazil. Arch. Virol. **157**:747–752.

24. **Ambrozevicius LP, Calegario RF, Fontes EPB, Carvalho MG, Zerbini FM.** 2002. Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. Fitopatol. Bras. **27**:372–377.

25. **Inoue-Nagata AK, Martin DP, Boiteux LS, Giordano LD, Bezerra IC, de Avila AC.** 2006. New species emergence via recombination among isolates of the Brazilian tomato-infecting begomovirus complex. Braz. J. Agric. Res. **41**:1329–1332.

26. **Bedford ID, Briddon RW, Brown JK, Rosell RC, Markham PG.** 1994. Geminivirus transmission and biological characterization of *Bemisia tabaci* (Gennadius) biotypes from different geographical regions. Ann. Appl. Biol. **125**:311–325.

27. **Castillo-Urquiza GP, Beserra Junior JEA, Alfenas-Zerbini P, Varsani A, Lima ATM, Barros DR, Zerbini FM.** 2007. Genetic diversity of begomoviruses infecting tomato in Paty do Alferes, Rio de Janeiro state, Brazil. Virus Rev. Res. **12**:233.

28. **Cotrim MA, Krause-Sakate R, Narita N, Zerbini FM, Pavan MA.** 2007. Genetic diversity of begomoviruses in tomatoes in midwestern São Paulo state. Summa Phytopathol. **33**:300–303. (In Portuguese.)

29. **Varsani A, Shepherd DN, Monjane AL, Owor BE, Erdmann JB, Rybicki EP, Peterschmitt M, Briddon RW, Markham PG, Oluwafemi S, Windram OP, Lefeuvre P, Lett JM, Martin DP.** 2008. Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. J. Gen. Virol. **89**:2063–2074.

30. **Duffy S, Holmes EC.** 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. J. Virol. **82**:957–965.

31. **Duffy S, Holmes EC.** 2009. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J. Gen. Virol. **90**:1539–1547.

32. **Ge LM, Zhang JT, Zhou XP, Li HY.** 2007. Genetic structure and population variability of tomato yellow leaf curl China virus. J. Virol. **81**:5902–5907.

33. **Isnard M, Granier M, Frutos R, Reynaud B, Peterschmitt M.** 1998. Quasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars. J. Gen. Virol. **79**:3091–3099.

34. **Hall JS, French R, Morris TJ, Stenger DC.** 2001. Structure and temporal dynamics of populations within wheat streak mosaic virus isolates. J. Virol. **75**:10231–10243.

35. **Davino S, Napoli C, Dellacroce C, Miozzi L, Noris E, Davino M, Accotto GP.** 2009. 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. **143**:15–23.

36. **Zhou YC, Noussourou M, Kon T, Rojas MR, Jiang H, Chen LF, Gamby K, Foster R, Gilbertson RL.** 2008. Evidence of local evolution of tomato-infecting begomovirus species in West Africa: characterization of tomato leaf curl Mali virus and tomato yellow leaf crumple virus from Mali. Arch. Virol. **153**:693–706.

37. **García-Andrés S, Tomas DM, Sanchez-Campos S, Navas-Castillo J, Moriones E.** 2007. Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. Virology **365**:210–219.

38. **Monci F, Sanchez-Campos S, Navas-Castillo J, Moriones E.** 2002. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and *Tomato yellow leaf curl virus* exhibits a novel pathogenic

39. **Andrade EC, Manhani GG, Alfenas PF, Calegario RF, Fontes EPB, Zerbini FM.** 2006. *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*. J. Gen. Virol. **87**:3687–3696.

40. **Jones RAC.** 2009. Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. Virus Res. **141**:113–130.

41. **Doyle JJ, Doyle JL.** 1987. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

42. **Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T.** 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage phi 29 DNA polymerase. J. Virol. Methods **116**:209–211.

43. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

44. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113. doi:10.1186/1471-2105-5-113.

45. **Wilgenbusch JC, Swofford D.** 2003. Inferring evolutionary trees with PAUP*. Curr. Protoc. Bioinformatics **6**:6.4.21–6.4.28. doi:10.1002/0471250953.bi0604s00.

46. **Posada D, Crandall KA.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817–818.

47. **Efron B, Halloran E, Holmes S.** 1996. Bootstrap confidence levels for phylogenetic trees. Proc. Natl. Acad. Sci. U. S. A. **93**:13429–13434.

48. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.

49. **Nylander JAA.** 2004. MrModeltest v2, program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

50. **Page RDM.** 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12**:357–358.

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

52. **Bryant D, Moulton V.** 2004. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. **21**:255–265.

53. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23**:254–267.

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

55. **Weir BS.** 1996. Genetic data analysis II: methods for discrete population genetic data. Sinauer Associated Inc., Sunderland, MA.

56. **Excoffier L, Laval G, Schneider S.** 2007. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinformatics Online **1**:47–50.

57. **Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R.** 2003. DnaSP: DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19**:2496–2497.

58. **Kosakovsky Pond SL, Frost SDW.** 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. **22**:1208–1222.

59. **Della Vecchia MGS, Rosa DD, Bergamin Filho A, Amorim L, Rezende JAM, Ribeiro A.** 2007. Spatio-temporal pattern of a begomovirus disease caused by Tomato yellow vein streak virus in tomato in Campinas region, São Paulo. Summa Phytopathol. **33**:387–395. (In Portuguese.)

60. **Rodríguez-Pardina PE, Hanada K, Laguna IG, Zerbini FM, Ducasse DA.** 2010. Molecular characterisation and relative incidence of bean- and soybean-infecting begomoviruses in northwestern Argentina. Ann. Appl. Biol. **156**:69–78.

61. **De Fazio G.** 1985. Bean golden mosaic in Brazil. Fitopatol. Bras. **10**:41–48. (In Portuguese.)

62. **Stanley J.** 1995. 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 **206**:707–712.

63. **Padidam M, Sawyer S, Fauquet CM.** 1999. Possible emergence of new geminiviruses by frequent recombination. Virology **265**:218–224.

64. **Sanz AI, Fraile A, Gallego JM, Malpica JM, García-Arenal F.** 1999. Genetic variability of natural populations of cotton leaf curl geminivirus, a single-stranded DNA virus. J. Mol. Evol. **49**:672–681.

65. **Lefeuvre P, Lett JM, Reynaud B, Martin DP.** 2007. Avoidance of protein

fold disruption in natural virus recombinants. PLoS Pathog. **3**:e181. doi:10.1371/journal.ppat.0030181.

66. **Lefeuvre P, Martin DP, Hoareau M, Naze F, Delatte H, Thierry M, Varsani A, Becker N, Reynaud B, Lett JM.** 2007. Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. J. Gen. Virol. **88**:3458–3468.

67. **Hou YM, Gilbertson RL.** 1996. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. J. Virol. **70**:5430–5436.

68. **Briddon RW, Patil BL, Bagewadi B, Nawaz-ul-Rehman MS, Fauquet CM.** 2010. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evol. Biol. **10**:97.

69. **da Silva SJC, Castillo-Urquiza GP, Hora Júnior BT, Assunção IP, Lima GSA, Pio-Ribeiro G, Mizubuti ESG, Zerbini FM.** 2011. High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed *Cleome affinis* in northeastern Brazil. Arch. Virol. **156**:2205–2213.

70. **Wyant PS, Strohmeier S, Schafer B, Krenz B, Assuncao IP, Lima GSD, Jeske H.** 2012. Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. Virology **427**:151–157.

71. **Paprotka T, Metzler V, Jeske H.** 2010. The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. Virology **404**:148–157.

72. **van der Walt E, Rybicki EP, Varsani A, Polston JE, Billharz R, Donaldson L, Monjane AL, Martin DP.** 2009. Rapid host adaptation by extensive recombination. J. Gen. Virol. **90**:734–746.

73. **Duffy S, Shackelton LA, Holmes EC.** 2008. Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. **9**:267–276.

74. **Firmino AC, Yuki VA, Moreira AG, Rezende JAM.** 2009. Tomato yellow vein streak virus: relationship with *Bemisia tabaci* biotype B and host range. Sci. Agric. **66**:793–799.

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

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

77. **Berrie LC, Rybicki EP, Rey MEC.** 2001. Complete nucleotide sequence and host range of South African cassava mosaic virus: further evidence for recombination amongst begomoviruses. J. Gen. Virol. **82**:53–58.

78. **van der Walt E, Martin DP, Varsani A, Polston JE, Rybicki EP.** 2008. Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. Virol. J. **5**:104.

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

80. **Frischmuth T, Roberts S, von Arnim A, Stanley J.** 1993. Specificity of bipartite geminivirus movement proteins. Virology **196**:666–673.

81. **Rojas MR, Noueiry AO, Lucas WJ, Gilbertson RL.** 1998. Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. Cell **95**:105–113.

82. **Hou YM, Paplomatas EJ, Gilbertson RL.** 1998. Host adaptation and replication properties of two bipartite geminiviruses and their pseudorecombinants. Mol. Plant Microbe Interact. **11**:208–217.

83. **Yahara T, Ooi K, Oshita S, Ishii I, Ikegami M.** 1998. Molecular evolution of a host-range gene in geminiviruses infecting asexual populations of *Eupatorium makinoi*. Genes Genet. Syst. **73**:137–141.

84. **Kaye AC, Moyer JW, Parks EJ, Carbone I, Cubeta MA.** 2011. Population genetic analysis of *Tomato spotted wilt virus* on peanut in North Carolina and Virginia. Phytopathology **101**:147–153.

85. **Parrish CR, Holmes EC, Morens DM, Park EC, Burke DS, Calisher CH, Laughlin CA, Saif LJ, Daszak P.** 2008. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. **72**:457–470.

86. **Ooi K, Ohshita S, Ishii I, Yahara T.** 1997. Molecular phylogeny of geminivirus infecting wild plants in Japan. J. Plant Res. **110**:247–257.

87. **Mansoor S, Amin I, Iram S, Hussain M, Zafar Y, Malik KA, Briddon RW.** 2003. Breakdown of resistance in cotton to cotton leaf curl disease in Pakistan. Plant Pathol. **52**:784. doi:10.1111/j.1365-3059.2003.00893.x.

88. **Amrao L, Amin I, Shahid MS, Briddon RW, Mansoor S.** 2010. Cotton leaf curl disease in resistant cotton is associated with a single begomovirus that lacks an intact transcriptional activator protein. Virus Res. **152**:153–163.