Notas Científicas

New species emergence via recombination among isolates of the Brazilian tomato infecting *Begomovirus* complex

Alice Kazuko Inoue-Nagata[1], Darren Patrick Martin[2], Leonardo Silva Boiteux[1], Leonardo de Britto Giordano[1], Isabel Cristina Bezerra[1] and Antonio Carlos de Ávila[1]

[1]Embrapa Hortaliças, Caixa Postal 218, CEP 70359-970 Brasília, DF, Brazil. E-mail: alicenag@cnph.embrapa.br, boiteux@cnph.embrapa.br, giordano@cnph.embrapa.br, bezerra@cnph.embrapa.br, avila@cnph.embrapa.br [2]Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory, Cape Town 7925, South Africa. E-mail: darren_mcb_staff_sci_main_uct@mail.uct.ac.za

Abstract – Partial nucleotide sequences of five tomato infecting *Begomovirus* isolates were determined from DNA-A fragments, corresponding to the 5' region of the replication associated protein gene, the intergenic region and the 5' region of the coat protein gene. Isolate DFM shared 95% identity with Tomato mottle leaf curl virus (TMoLCV), isolates 34, PA-05, and Ta4 were 88% identical to Tomato yellow vein streak virus and isolate DF-BR3 shared 77% identity with TMoLCV. Recombination analysis indicated that isolate DF-BR3 was a chimaera, and it provided evidence that there is a complex and actively recombining population of tomato infecting begomoviruses in Brazil.

Index terms: begomovirus, geminivirus, virus recombination.

Emergência de nova espécie viral por recombinação entre isolados do complexo *Begomovirus* do tomateiro

Resumo – A seqüência de nucleotídeos parcial de cinco isolados de *Begomovirus* foi determinada do DNA-A, correspondente à região intergênica e à porção 5' do gene associado à replicação e da capa protéica. O isolado DFM apresentou identidade de 95% com Tomato mottle leaf curl virus (TMoLCV); os isolados 34, PA-05 e Ta4 foram 88% idênticos ao Tomato yellow vein streak virus (ToYVSV); e o isolado DF-BR3 mostrou 77% de identidade com TMoLCV. Análise de recombinação indicou que o isolado DF-BR3 seria uma quimera e evidencia que um complexo de espécies de begomovírus bipartidos está em formação no Brasil.

Termos para indexação: begomovírus, geminivírus, recombinação viral.

The genus *Begomovirus* in the family *Geminiviridae* contains many of the most economically important tomato infecting virus species currently known (Fauquet et al., 2003). Although *Tomato golden mosaic virus* (TGMV) was the only Brazilian tomato infecting species described prior to the 1990's (Flores et al., 1960), other less characterized species with the potential to infect tomato might have been present in the country since before the 1970's (Costa, 1974).

Since the introduction into Brazil of a new whitefly vector biotype (*Bemisia tabaci* Genn. biotype B or *B. argentifolii* Bellows & Perring) in the 1990's, ensuing geminivirus epidemics in tomatoes, which resulted in the identification of ten other tomato infecting Brazilian begomovirus species/provisional species (Ribeiro et al., 2003). These include Tomato yellow vein streak virus (ToYVSV), *Tomato rugose mosaic virus* (ToRMV), *Tomato severe rugose virus* (ToSRV), *Tomato chlorotic mottle virus* (ToCMoV), Tomato mottle leaf curl virus (TMoLCV), Tomato chlorotic vein virus, Tomato severe mosaic virus, Tomato infectious yellow virus, Tomato crinkle leaf yellow virus and Sida micrantha mosaic virus (Ribeiro et al., 2003; Calegario et al., 2004). Phylogenetic analyses of these and other ones found in the Americas have indicated that the Brazilian viruses form a distinct, albeit only weakly supported, monophyletic clade (Figure 1) (Ribeiro et al., 2003).

Recombination is now well established as a primary mechanism of begomovirus diversity generation (Padidam et al., 1999). Given the diverse assemblage of tomato infecting Brazilian begomoviruses and potentially high incidences of coinfection, there is a chance that recombination between different tomato infecting begomovirus species might currently be

contributing to the emergence of novel genotypes. This report describes the genetic diversity of, and recombination between, tomato begomoviruses sampled in the Brasília region of Brazil, in 2001.

Five begomovirus isolates were collected from tomato plants in the north and south regions of the Brasília green belt: isolates DF-BR3, DFM and 34 were collected at Embrapa Hortaliças in Brasília, DF; isolate Ta4 was collected in a commercial tomato field in Tabatinga, DF; and isolate PA-05 was collected in Ponte Alta, DF. These isolates were preliminary identified as begomoviruses due to the symptoms they caused in tomato plants, such as veinal yellowing, golden-mosaic, leaf distortion, and rugosity. They were kept in tomato plants by grafting or frozen as total DNA extracts.

Total DNA was extracted from individual plants and used as templates for PCR amplification of DNA-A fragments with the primers pALv1978 (Rojas et al., 1993) (annealing in the middle of the *Rep* gene), and pARc715 (Rojas et al., 1993) (annealing 400 nt downstream the *CP* gene start codon) for isolates 34, DF-BR3 and DFM (1.3 kb DNA fragment amplified), or CP2 (CCC CTG CAG AAC TTC CAA GTC TGG ACG) (annealing 200 nt downstream from the *CP* stop codon) for isolates PA-05 and Ta4 (1.9 kb DNA fragment amplified). Primer pair pALv1978 and CP2 only amplified a DNA-A fragment from PA-05 and Ta4. Therefore for isolates 34, DF-BR3 and DFM the primer pair pALv1978 and pARc715 was used. Amplified fragments were cloned into pGEM-T (Promega) and transformed into *Escherichia coli* XL-1 Blue. Sequencing reactions were carried out with both vector primers and the internal primer pARc496 (Rojas et al., 1993). Nucleotide sequences were obtained using an ABI 377 sequencer and were compiled using DNASIS (Hitachi). Final sequences were deposited in GenBank under accession numbers AY751742 (PA-05), DQ346649 (34), DQ346650 (Ta4), DQ346651 (DFM), and DQ34665 (DF-BR3).

BLASTN searches (www.ncbi.nih.nlm.gov\blast) and phylogenetic analyses (Figure 1) of ClustalW (http://www.ebi.ac.uk/clustalw/) aligned sequences indicated that (1) PA-05, 34, and Ta4 are most closely related to ToYVSV, collectively sharing ~88% identity with this species; (2) DFM and DF-BR3 are most closely related to TMoLCV. PA-05, 34, and Ta4 share between 95 and 97% nucleotide sequence identity suggesting that they are variants of the same species, and further analyses were therefore only carried out with PA-05. A comparison of the 5' regions of *CP* (251 nt) and *Rep* (666 nt) and of the entire intergenic region (IR) of PA-05, DFM as well as DF-BR3 was carried out with other Brazilian tomato begomovirus species. Whereas DF-BR3 shares 95% and 91% identity with PA-05 in the *CP* and IR regions, respectively, it shares only 74% identity with PA-05 in the *Rep* region. This observation as well as the fact that both isolates were obtained from the same geographic region suggested that either DF-BR3 or PA-05 might be recombinant. Similarly, whereas the DFM *Rep* and *CP* sequences shared identities higher than 95% with TMoLCV, the IR nucleotide sequence of DFM shared only 87% identity with that of TMoLCV, indicating that either DFM or TMoLCV might be recombinant.

DF-BR3 shares only 78% nucleotide sequence identity with its nearest relative, TMoLCV, and initially appeared to be the most divergent of the isolates examined. However, the *CP* nucleotide sequence of DF-BR3 was 90% identical to that of ToYVSV, its *Rep* more than 98% identical to those of TMoLCV and DFM, and its IR sequence 91% identical to that of isolate PA-05. Again, it was considered probable that either PA-05/TMoLCV or DF-BR3 was a recombinant.

Identification of recombinant and likely parental sequences, and localization of possible recombination breakpoints was carried out using the RDP (Martin & Rybicki, 2000), GENECONV (Padidam et al., 1999), MAXIMUM CHI^2 (Smith, 1992), CHIMAERA (Martin et al., 2005b), RECSCAN (Martin et al., 2005a), and SISTER SCAN (Gibbs et al., 2000) methods as implemented in RDP2 (Martin et al., 2005b). The analysis was performed with default settings for the different detection methods with a Bonferroni corrected p value cutoff of 0.05.

It is probable ($p = 4.7x10^{-30}$) that isolate DF-BR3 is a recombinant (Figure 2 A). While it is clear that nucleotides -137 to 364 (position 0 at the origin of virion strand replication) originated from a sequence resembling isolate PA-05 (black region and right tree in Figure 2 B), the rest of the sequence originated from a virus closely resembling isolate DFM (white region and left tree in Figure 2 B). The genetic distances between the parental and recombinant viruses are very small (Figure 2 B), indicating that the recombination event has occurred relatively recently. Obvious recombination events (p values$<10^{-4}$) were also detected in four other sequences: TMoLCV, ToCMoV, ToRMV and TGMV (Figure 2 A). As anticipated from the distance analyses

of DFM and TMoLCV, it is apparent that part of the *Rep* and IR of TMoLCV has a recombinant origin ($p = 1.6x10^{-10}$) with DFM resembling a parental sequence.

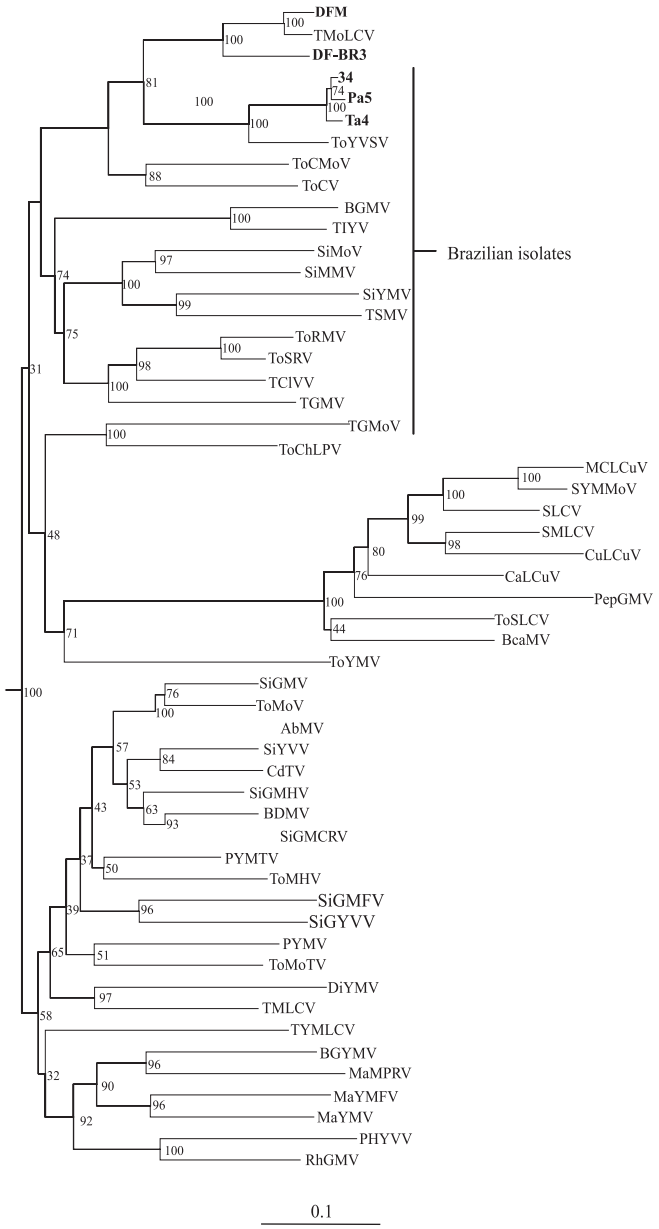

Figure 1. Phylogenetic relationships between five isolates described in this study (in bold) and previously described New World begomoviruses. The tree was constructed using the PHYLIP neighbor joining method with Jukes-Cantor distances and 1,000 bootstrap replicates (in the nodes). The tree is outgroup rooted using *African cassava mosaic virus* (not shown). A list of accession numbers for sequences in the tree and the multiple sequence alignment used in its construction are available at http://darwin.uvigo.es/RDP/BPFig1.zip.

This work confirms that, as elsewhere in Brazil, there is a tomato infecting begomovirus species complex in the central region of the country. Although only five

Isolate	*Rep*	IR	*CP*	Region	Minor parent (Black)	Major parent (White)	Methods	p-value
DF3–BR3				-137 / +364	Pa5	DFM	RGBMCS	4.7×10^{-30}
TMoLCV				-528 / -28	Unknown	DFM	RGBMCS	1.6×10^{-10}
ToCMoV				-589 / -261	DFM	Unknown	GBMCS	3.2×10^{-5}
				-140 / +39	DFM	Unknown	RGBMCS	3.0×10^{-9}
ToRMV				+89 / +435	ToCMoV	ToSRV	RGBMCS	7.2×10^{-26}
TGMV				-794 / -469	Unknown	ToRMV	MC	7.1×10^{-5}

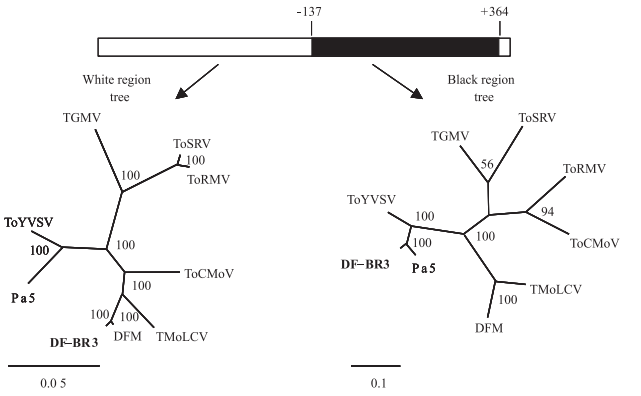

Figure 2. (A) Potential recombinant regions detected within Brazilian tomato infecting begomovirus sequences. Region coordinates are nucleotide positions relative to the origin of virion sense replication. Recombination events indicated were identified using the RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA and SISTER SCAN methods. The reported p value is for the method in bold type and is the best p value calculated for the region in question. A schematic representation of the analyzed genome fragment is shown above the recombinant region maps with the positions of partial *Rep*, partial *CP* and complete IR sequences indicated. (B) Phylogenetic evidence that isolate DF-BR3 is a recombinant virus. While nucleotides between -137 and +364 of the DF-BR3 genome are derived from a virus closely related to PA-05, the remainder of the analyzed DF-BR3 sequence most closely resembles that of DFM. Both trees were constructed using the neighbor joining method with Jukes Cantor distances. Numbers associated with branches indicate the percentage of 1,000 bootstrap replicates supporting the existence of that branch in the phylogeny. The schematic representation of the DF-BR3 sequence above the trees indicates the different genomic regions used to construct the trees.

isolates were studied, three distinct tomato virus genotypes were observed: (1) a genotype that seems to be a divergent strain of TMoLCV (DFM); (2) a group of isolates representing a potential new species (PA-05, Ta4 and 34); and (3) a recombinant between this new species and the TMoLCV like DFM (DF-BR3). These results are not surprising given the extraordinarily high prevalence of recombinant geminiviruses detected worldwide (Padidam et al., 1999; Rojas et al., 2005). However, the discovery of DF-BR3 supports the hypothesis that recombination amongst members of the tomato infecting begomovirus species complex in central Brazil is likely to contribute to the emergence of novel tomato infecting begomovirus genotypes.

References

CALEGARIO, R.F.; FERREIRA, S.S.; ANDRADE, C.; ZERBINI, F.M. Caracterização de um isolado do begomovírus Sida micrantha mosaic virus (SimMV) obtido de tomateiro. **Fitopatologia Brasileira**, v.29, p.150, 2004.

COSTA, A.S. Moléstias do tomateiro transmitidas por mosca-branca *Bemisia tabaci*. **Fitopatologia**, v.9, p.47, 1974.

FAUQUET, C.M.; BISARO, D.M.; BRIDDON, R.W.; BROWN, J.K.; HARRISON, B.D.; RYBICKI, E.P.; STENGER, D.C.; STANLEY, J. Revision of taxonomic criteria for species demarcation in the family *Geminiviridae*, and an updated list of begomovirus species. **Archives of Virology**, v.148, p.405-421, 2003.

FELSENSTEIN, J. PHYLIP phylogeny inference package (version 3.2). **Cladistics**, v.5, p.164-166, 1989.

FLORES, E.; SILBERSCHMIDT, K.; KRAMER, M. Observações de "clorose infecciosa" das malváceas em tomateiros do campo. **Biológico**, v.26, p.65-69, 1960.

GIBBS, M.J.; ARMSTRONG, J.S.; GIBBS, A.J. Sister scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. **Bioinformatics**, v.16, p.573-582, 2000.

MARTIN, D.; RYBICKI, E. RDP: detection of recombination amongst aligned sequences. **Bioinformatics**, v.16, p.562-563, 2000.

MARTIN, D.P.; POSADA, D.; CRANDALL, K.A.; WILLIAMSON, C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. **AIDS Research and Human Retroviroses**, v.21, p.98-102, 2005a.

MARTIN, D.P.; WILLIAMSON, C.; POSADA, D. RDP2: recombination detection and analysis from sequence alignments. **Bioinformatics**, v.21, p.260-262, 2005b.

PADIDAM, M.; SAWYER, S.; FAUQUET, C.M. Possible emergence of new geminiviruses by frequent recombination. **Virology**, v.265, p.218-225, 1999.

RIBEIRO, S.G.; AMBROZEVÍCIUS, L.P.; ÁVILA, A.C.; BEZERRA, I.C.; CALEGARIO, R.F.; FERNANDES, J.J.; LIMA, M.F.; MELLO, R.N. de; ROCHA, H.; ZERBINI, F.M. Distribution and genetic diversity of tomato infecting begomoviruses in Brazil. **Archives of Virology**, v.148, p.281-295, 2003.

ROJAS, A.; KVARNHEDEN, A.; MARCENARO, D.; VALKONEN, J.P.T. Sequence characterization of tomato leaf curl Sinaloa virus and tomato severe leaf curl virus: phylogeny of New World begomoviruses and detection of recombination. **Archives of Virology**, v.150, p.1281-1299, 2005.

ROJAS, M.R.; GILBERTSON, R.J.; RUSSEL, D.R.; MAXWELL, D.P. Use of degenerate primers in the polymerase chain reaction to detect whitefly transmitted geminiviruses. **Plant Disease**, v.77, p.340-347, 1993.

SMITH, J.M. Analyzing the mosaic structure of genes. **Journal of Molecular Evolution**, v.34, p.126-129, 1992.

Received on January 13, 2005 and accepted on April 16, 2006