Characterization of a New World monopartite begomovirus causing leaf curl
disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution

Tomas A. Melgarejo*†, Tatsuya Kon, Maria R. Rojas, Lenin Paz-Carrasco#‡, F. Murilo Zerbini# and Robert L. Gilbertson*

*Department of Plant Pathology, University of California, Davis, California 95616
†Departamento de Fitopatologia, Universidad Nacional Agraria La Molina, Lima, Peru
#Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, 36570-000, Brasil
‡ INIAP, EELS Enrique Ampuero Pareja, Casilla 09-01-7069, Guayaquil, Ecuador*

*Corresponding author. Mailing address: Department of Plant Pathology, University of California-Davis, 1 Shields Ave., Davis, CA 95616. Phone: (530) 752-3163. Fax: (530) 752-5674. E-mail: rlgilbertson@ucdavis.edu

Current address of second author: Plant Pathology Laboratory, Faculty of Agriculture, Iwate University, Morioka 020-8550, Japan

Running title: Tomato leaf curl disease in Peru is caused by a New World monopartite begomovirus
All characterized whitefly-transmitted geminiviruses (begomoviruses) with origins in the New World (NW) have bipartite genomes composed of a DNA-A and DNA-B component. Recently, a NW begomovirus lacking a DNA-B component was associated with tomato leaf curl disease (ToLCD) in Peru and named Tomato leaf deformation virus (ToLDeV). Here, we show that isolates of ToLDeV associated with ToLCD in Ecuador and Peru have a single, genetically diverse, genomic DNA that is most closely related to DNA-A components of NW bipartite begomoviruses. Agroinoculation of multimeric clones of the genomic DNA of three ToLDeV genotypes (two variants and a strain) resulted in the development of tomato leaf curl symptoms indistinguishable from those of ToLCD in Ecuador and Peru. Biological properties of these ToLDeV genotypes were similar to those of Old World (OW) monopartite tomato-infecting begomoviruses, including lack of sap-transmissibility, phloem limitation, a resistance phenotype in tomato germplasm with the Ty-1 gene and functional properties of the V1 (capsid protein) and C4 genes. Differences in symptom phenotypes induced by the ToLDeV genotypes in tomato and Nicotiana benthamiana plants were associated with a highly divergent left intergenic region and C4 gene. Together, these results establish that ToLDeV is an emergent NW monopartite begomovirus, which is causing ToLCD in Ecuador and Peru. This is the first report of an indigenous NW monopartite begomovirus, and evidence is presented that it emerged from the DNA-A component of a NW bipartite progenitor via convergent evolution and recombination.
The family *Geminiviridae* is a large and diverse group of plant pathogenic viruses that possess a small circular single-stranded (ss) DNA genome encapsidated within characteristic twinned, quasi-isometric virions (4, 29, 30, 60). Four genera of geminiviruses are recognized based upon genome structure, host range and type of insect vector (13). The genus *Begomovirus* contains viruses transmitted by the whitefly *Bemisia tabaci*, and many of these cause economically important diseases of dicotyledonous crop plants in tropical and sub-tropical regions throughout the world (45, 59, 66). Begomoviruses have the largest number of species in the family (and probably of any type of virus), and many have been described within the past approx. 20 years (13). This makes begomoviruses a remarkably successful group of emerging viruses (4, 59), and this can be attributed to a number of factors. First, indigenous begomoviruses are widely distributed in non-cultivated plants (e.g., weeds), and these serve as progenitors of crop-infecting viruses. Second, population explosions and the expanded geographical range of the invasive polyphagous B biotype of *B. tabaci* have facilitated the introduction of indigenous begomoviruses into crop plants. Third, the begomovirus genome has the capacity to evolve rapidly via mutation, pseudorecombination, recombination and acquisition of new DNA components and satellites (4, 59, 60, 65, 66). Finally, the worldwide cultivation of susceptible crop plants (e.g., cassava, cotton, cucurbits and tomato), often in monoculture and overlapping plantings, has resulted in the exposure of potential host plants to high populations of whiteflies and a diversity of begomoviruses (20, 66).

There are two types of begomoviruses: those with a monopartite genome (a single genomic DNA of approximately 2.9 kb), and those with a bipartite genome (two DNA components, each approximately 2.6 kb, referred to as DNA-A and DNA-B) (4, 13, 30, 60). Monopartite begomoviruses originated in the OW, and many are associated with satellite DNAs.
that are either required for development of typical disease symptoms (betasatellites) or have no apparent effect or modulate disease symptoms (alphasatellites) (4). In the absence of a betasatellite, many monopartite begomoviruses can infect and cause symptoms in the experimental host, *Nicotiana benthamiana*; however, only those that infect tomato cause typical disease symptoms in the natural host without a betasatellite (4). All indigenous New World (NW) begomoviruses have a bipartite genome, as do a relatively small number (approximately 15%) of OW viruses (24). Bipartite begomoviruses do not require satellite DNAs for induction of typical symptoms, but they can be found co-infecting plants (4, 60). Within the past approximately 20 years, the genomic complexity of begomoviruses in the NW has increased with 1) the introduction of the OW monopartite begomovirus *Tomato yellow leaf curl virus* (TYLCV) (40) and 2) recent reports of alphasatellites associated with bipartite begomoviruses in Latin America (15, 49). However, monopartite begomoviruses indigenous to the NW have not been reported (4).

The genomic DNA of monopartite begomoviruses and the DNA-A component of bipartite begomoviruses are homologous and have a similar genome organization that includes the V1/AV1, C1/AC1, C2/AC2, C3/AC3 and C4/AC4 genes encoding the capsid protein (CP), replication-associated protein (Rep), transactivator protein (TrAP), replication enhancer protein (REN) and C4/AC4 proteins, respectively (22, 30). However, there are also differences between these viral DNAs. The genomic DNA of monopartite begomoviruses is larger (approximately 2.9 kb) than the DNA-A components of NW bipartite begomoviruses (approximately 2.6 kb), and it has an additional virion-sense gene (V2), which encodes the pre-coat protein (note that DNA-A components of OW bipartite begomoviruses have a homologous AV2 gene). The CP of NW bipartite begomoviruses has a distinctive N-terminal PWRsMaGT motif (31). There are also
differences in the function of homologous gene products. The CP of monopartite begomoviruses is required for systemic infection, whereas CP is not required systemic infection by many bipartite species (60, 68). In monopartite begomoviruses, the C4 is a pathogenicity determinant that can play a role in movement and suppression of host defense responses. In many NW bipartite begomoviruses, the homologous AC4 does not play a role in pathogenicity (51, 60). The evolution of the bipartite genome, by the acquisition of the DNA-B component, probably occurred following the appearance of monopartite species. The DNA-B component has two genes, BV1 and BC1, that encode proteins involved in nucleocytoplasmic transport (nuclear shuttle protein, NSP) and cell-to-cell movement (movement protein, MP), respectively (60). The differences between genome structure and gene function of monopartite and bipartite begomoviruses is also reflected in their biological properties. Monopartite begomoviruses are phloem-limited, induce stunting and leaf curl symptoms and are not sap-transmissible; whereas many bipartite begomoviruses infect phloem and non-phloem tissues, induce leaf curling, crumpling and mosaic/mottling symptoms and are sap-transmissible.

The cultivated tomato (*Solanum lycopersicum*) has proven well suited for the emergence and evolution of bipartite and monopartite begomoviruses. Indeed, at least 67 species of tomato-infecting begomoviruses have been described, which is the largest number infecting any crop (13, 50). In the OW, monopartite begomoviruses predominate, although there are a few bipartite species; whereas in the NW, only bipartite species have been described (26, 45, 50, 67). However, the finding that DNA-A components of some bipartite tomato-infecting begomoviruses are infectious has blurred the distinction between the monopartite and bipartite genome. For example, when delivered by agroinoculation, the DNA-A component of the OW bipartite begomovirus *Tomato yellow leaf curl Thailand virus* (TYLCTHV) induces disease...
In the mid 1990s, a leaf curl disease of tomatoes in Peru appeared following outbreaks of *B. tabaci* (43, 45, 46). In 2003, a similar disease was described in Ecuador. Disease symptoms included stunting and distorted growth, upward curling of leaves and swelling and purpling of veins. A begomovirus was the suspected causal agent, but it was also noted that the symptoms were unusual with respect to those typically induced by NW begomoviruses, in particular, the absence of mottling, mosaic and yellowing. The DNA-A component of a putative new begomovirus species was associated with this disease (43, 46), and the name Tomato leaf deformation virus (ToLDeV) was proposed (43). However, Koch’s postulates were not fulfilled for ToLCD, nor was a DNA-B component identified.

Here, we show that ToLDeV isolates associated with ToLCD in Ecuador and Peru have a single, genetically divergent, genomic DNA, which is most closely related to DNA-A components of NW bipartite begomoviruses. The development of leaf curl symptoms in tomato plants, following agroinoculation with the cloned genomic DNA of ToLDeV, fulfilled Koch’s postulates for this disease and confirmed the monopartite nature of the viral genome. In addition, the biological properties and functional properties of the V1 (CP) and C4 genes are similar to...
those of OW monopartite tomato-infecting begomoviruses. Together, these results establish that ToLDeV is a bone fide indigenous NW monopartite begomovirus, and evidence is presented that it evolved from the DNA-A component of bipartite progenitor via convergent evolution and recombination.

MATERIALS AND METHODS

Virus sources, DNA extraction and detection of begomovirus DNA. In 1998, leaf samples were collected from tomato plants with stunting and upward leaf curling symptoms associated with a new virus-like disease, from two fields (98-1 and 98-2) in the Ica province of Peru. These samples were sent to UC Davis and tested for begomovirus infection by squash blot hybridization analysis (18) and PCR with degenerate begomovirus DNA-A and DNA-B primer pairs (58).

In July 2010, a survey of commercial tomato fields in two provinces of Peru, Ica (8 fields) and Lima (2 fields), was conducted for leaf curl symptoms and whiteflies. In each field, a minimum of 20 plants at 3-4 locations were examined for leaf curl symptoms and representative samples were collected. The presence and relative abundance (high, medium, or low) of adult whiteflies was estimated in each field by examining the underside of at least 10 leaves. Twenty-five samples with leaf curl symptoms were collected, and sap from each sample was prepared and applied onto Agdia absorption strips (Agdia, Elkhart, IN). The strips were dried and then transported to UC Davis. In Ecuador, surveys of commercial tomato fields were conducted in the provinces of Guayas and Santa Helena in August and September 2010. Sixty-eight samples with leaf curl symptoms were collected, dried in a plant press and stored at room temperature.
Total genomic DNA was extracted from leaf samples collected in Peru in 1998 (frozen tissue) and 2010 (dried plant sap on absorption strips) as described by Dellaporta et al. (8), whereas the method of Doyle and Doyle (10) was used for the samples collected in Ecuador. PCR with the degenerate DNA-A primer pair, PAL1v1978 and PAR1c496, was used to detect begomovirus infection (58). PCR-amplified DNA fragments were purified with the QIAquick gel extraction kit (Qiagen, Germantown, MD) and directly sequenced.

Three approaches were used to detect a begomovirus DNA-B component and satellite DNAs. First, PCR with various degenerate DNA-B primer pairs (PCRc1/PLB1v2040, PCRc1/PLB1v2039, PCRc154/PLB1v2040 and PCRc154/PLB1v2039) and the universal primer pair Beta01/Beta02 was used (3, 58). Second, restriction enzyme digestion analysis of circular DNA molecules generated by rolling circle amplification (RCA) with Φ-29 DNA polymerase (TempliPhi, GE Healthcare, Piscataway, NJ) was used. The third approach was DNA gel blot hybridization analysis of total genomic DNA with the cloned DNA-B component of the bipartite begomovirus *Bean dwarf mosaic virus* (BDMV) under conditions of low stringency. Controls were total genomic DNA extracted from leaves of uninfected and BDMV-infected *N. benthamiana* plants.

**Obtaining full-length begomovirus DNA clones.** DNA extracts in which begomovirus infection was detected by PCR were used for RCA as described (34). The RCA products were digested with various restriction enzymes to identify sites for obtaining full-length clones. Putative full-length clones were obtained by digesting DNA generated by RCA with restriction enzymes that linearized the viral DNA, purifying the approximately 2.6 kb fragment and cloning into pGEM-11Zf(+) (Promega, Madison, WI). Recombinant plasmids were identified by restriction enzyme digestion.
Sequencing and phylogenetic analyses. DNA sequencing was conducted with the ABI 3730 DNA Analyser (Applied Biosystems). Sequence alignments were performed with the MUSCLE program in Mega 5 (70). The aligned nucleotide (nt) sequences were edited and exported as nexus files. For each data set, phylogenies were constructed with a Bayesian approach performed with MrBayes 3.2 (62). The best-fit model of nt substitution for each data set was determined with the program ModelTest (52). The analyses were carried out by running 2 M generations, and trees were sampled every 100th generation resulting in 20,000 trees. After discarding the first 10% of samples as burn-in, the remaining 18,000 trees were used for calculating posterior probabilities in the consensus tree. Trees were generated in Treeview (48) and rooted with the curtovirus Beet severe curly top virus (BSCTV).

Detection of potential recombination events was carried out with the recombination detection program (RDP) version 3.0 (44) and various data sets including one with the complete nt sequences of 180 begomoviruses (13). Similarity plot analysis was performed with the SimPlot software 3.5.1 (42) using the Jukes–Cantor model, a window size of 200 nt and a step size of 20 nt.

Analysis of genetic diversity and selection pressure. Nucleotide diversity (\(\pi\)) was estimated with Tajima’s test (69) using DnaSP 5.10 (41). Codon-specific positive selection was determined by Bayesian posterior probability under the Ny98 fitness regime (47) (\(0<\omega_1<1\), \(\omega_2=1\), and \(\omega_3>1\)) as described and performed by MrBayes 3.2 (62). The distribution of synonymous and nonsynonymous mutations was analyzed with the SNAP program (www.hiv.lanl.gov) (38). The hypothesis of independent evolution in overlapping genes was tested as described by Zaaijer et al. (76). Variation was assessed by determining the entropy values for individual amino acids (aas) and nts with the BioEdit software (28). An alignment of AC4/C4 aa sequences was
generated with the MUSCLE program in Jalview (74). Jalview was used to edit the alignment and determine conserved regions.

**Production of multimeric clones for agroinoculation.** Multimeric clones were generated in pCAMBIA1300 (27) as previously described (36). For the PA98-1-1 and PA98-2-1 clones from samples collected in Peru in 1998, an approximately 1.6 kb EcoRI-HindIII fragment containing the intergenic region (IR) was cloned to generate a 0.6-mer. Full-length monomers were released with HindIII and cloned into HindIII-digested 0.6-mers to generate the 1.6-mers. For the PA10-3 clone from the sample collected in Peru in 2010, the same strategy was used except an approximately 2.4 kb EcoRI-HindIII fragment was cloned to generate a 0.9-mer. A 1.9-mer was generated following ligation with the full-length monomer. Recombinant plasmids with the multimeric clones were identified by restriction enzyme digestion, and transformed into competent *Agrobacterium tumefaciens* cells (strain C58C1) with the freeze-thaw method (75).

**Determination of infectivity, host range and response of tomato germplasm with the Ty-1 resistance gene.** Infectivity of the multimeric clones was initially assessed by agroinoculation of tomato (cv. Glamour) and *N. benthamiana* plants with the needle puncture method (33). Host range determination was performed by agroinoculation of tomato (cv. Glamour), *Datura stramonium*, *N. benthamiana*, *N. glutinosa*, *N. tabacum* cv. Samsum, *Phaseolus vulgaris* (cv. Topcrop), *Cucurbita pepo* (cv. Small sugar), *Chenopodium quinoa* and *C. amaranticolor* plants. Infection was determined 18 to 21 days post inoculation (dpi) based upon symptom development and detection of viral DNA by PCR analysis with the PAR1c496 and ToLC-v1730 primer pair, which directs the amplification of an approximately 1.1 kb DNA fragment from these clones.
Tomato germplasm with the Ty-1 begomovirus resistance gene (LA3473 [C.M. Rick Tomato Genetics Resource Center, UC Davis] and cv. Dominator [Seminis Vegetable Seeds Woodland, CA]) and without the gene (LA3474 and cv. Glamour) were agroinoculated as described above. Controls were uninfected plants and plants agroinoculated with the monopartite begomovirus TYLCV-Israel (TYLCV-IL from the Dominican Republic (TYLCV-IL[DO]) (37) and the bipartite begomovirus Tomato mosaic Havana virus (ToMHV) from Guatemala (Kon and Gilbertson, unpublished). Inoculated plants were maintained in a controlled-environment growth chamber (25°C), and disease symptoms were assessed 21 to 25 dpi.

**Assessment of viral DNA levels.** Total genomic DNA was extracted from leaf tissue, and DNA gel blot hybridization analysis was conducted as previously described (37). Newly emerged leaves were collected from infected plants 21 and 60 dpi, total genomic DNA was extracted and DNA gel blot hybridization analysis conducted with a probe composed of the recombinant plasmids pPA98-1-1, pPA98-2-1 and pPA10-3 labeled with [$\alpha^{32}$P] dCTP by nick translation.

**Sap transmission and immunolocalization studies.** Leaves of *N. benthamiana* plants at the 4-6 leaf stage were rub-inoculated with sap as previously described (18). Sap was prepared with frozen leaf tissue from the 1998 samples, and symptomatic leaf tissue from *N. benthamiana* plants agroinoculated with the multimeric PA10-3 clone. Controls were *N. benthamiana* plants inoculated with sap prepared with symptomatic leaf tissue from common bean plants infected with the sap-transmissible BDMV (positive control) and 0.1 M phosphate buffer (pH 7.2) alone (negative control). Three independent experiments were performed with five plants inoculated per treatment per experiment.
Immunolocalization experiments were conducted with leaf and petiole tissues collected from symptomatic tomato plants (cv. Glamour) four weeks after agroinoculation with the multimeric PA98-1-1 and PA98-2-1 clones. The antibody used for immunolocalization was raised against the *E. coli*-expressed CP of the bipartite begomovirus *Chino del tomate virus* (73). Controls were equivalent tissues from uninfected plants probed with the CP antisera, and from infected plants probed with preimmune antisera. Tissue preparation, sectioning and immunolocalization were performed as previously described (61). Sections were visualized with a phase contrast light microscope (Leica DM5000 B). Images were recorded with the Leica Application Suite (LAS) 3.6 software.

**Mutational analysis.** Mutations were introduced into the V1 (CP), C1 (Rep), C4 and C5 genes of the PA98-1-1, PA98-2-1 and PA10-3 clones. The recombinant plasmids pPA98-1-1.6, pPA98-2-1.6 and pPA10-3-1.9 were used as templates for site-directed mutagenesis with phosphorylated primers (primers sequences are available upon request). The mutations in the V1 and C1 genes were single nt deletions in the start codons (V1: ATG→A-G at nt 176 or 177 [PA10-3], and CAT→C-T at nt 2456 or 2458 [PA10-3]). The mutations in the C4 and C5 genes were single nt changes that abolished the start codons (C4: CAT→CGT at nt 2300 or 2301 [PA10-3] and C5: CAT→CAG at nt 824 [PA98-1-1], 825 [PA98-2-1] and 826 [PA10-3]), but were silent in the overlapping C1 and V1 genes, respectively (Fig. 1). The PCR reaction was performed according to the protocol of the Phusion site-directed mutagenesis kit (Finnzymes, Espoo, Finland). Following mutagenesis, the PCR-amplified DNA components were self-ligated, digested with HindIII and ligated into HindIII-digested pGEM-11Zf(+). Recombinant plasmids were recovered and the presence of the introduced mutation (as well as the lack of any other introduced
mutations) was verified by sequencing. Monomers confirmed to give the mutations were used to
generate multimeric clones for agroinoculation as described.

Replication was assessed in *N. benthamiana* leaf discs as previously described (37). Infectivity and symptomatology of mutants in agroinoculated tomato (cv. Glamour) and *N. benthamiana* plants were assessed as described. The presence of the introduced mutation was assessed for c4 and c5 mutants in selected tomato and *N. benthamiana* plants by sequencing PCR-amplified DNA fragments with the mutated region of the gene. The stability of the c4 mutants was assessed by grafting uninfected tomato scions onto infected rootstocks. Three weeks after grafting, the scions were assessed for symptoms and the presence of the mutation was determined as described.

**Construction of chimeras.** Chimeras were generated between the PA98-1-1 and PA10-3 clones in the recombinant plasmids pPA98-1-1 and pPA10-3, respectively. The approximately 1.5 kb KpnI/EcoRV fragments with the left intergenic region (LIR) and C1 and C4 genes were exchanged (Fig. 1). This resulted in the generation of the chimeras: pPA98-1-1/PA10-3 and pPA10-3/PA98-1-1 (the second name indicates the source of approximately 1.5 kb fragment).

**RESULTS**

**ToLDeV is associated with ToLCD in Peru and Ecuador.** The survey of commercial tomato fields conducted in Peru in July 2010 revealed ToLCD incidences of 20% and 5% in the Ica and Lima provinces, respectively. Whitefly populations in these fields were low, with <10% of plants having adult insects. Surveys conducted in Ecuador in August and September 2010 revealed disease incidences of 5-10% and 20% in the Guayas and Santa Helena provinces, respectively,
and low whitefly populations. The relatively low incidence of ToLCD and whiteflies was attributed to these surveys being conducted outside of the main tomato-growing season (October to April).

The degenerate DNA-A primer pair PAL1v1978/PAR1c496 directed the amplification of the expected-size 1.1 kb fragment from the two samples collected from the Ica province in 1998, consistent with previous results of 1998 (data not shown). The 1.1 kb fragment was also amplified from 13 of 20 and 1 of 5 samples collected from the Ica and Lima provinces in 2010, respectively; and from 44 of 68 samples collected from Ecuador in 2010. Comparisons performed with sequences of representative 1.1 kb DNA fragments from these samples revealed highest identities with sequences of DNA-A components of NW bipartite begomoviruses (data not shown). Thus, many of the tomato plants with leaf curl symptoms from Ecuador and Peru were infected with a begomovirus, but the finding that some samples were not infected shows that these symptoms can be caused by other factors (7).

Several lines of evidence indicated that a DNA-B component and satellite DNAs were not present in the ToLCD samples from Ecuador and Peru. First, in PCR assays performed with degenerate DNA-B primer pairs and a universal betasatellite primer pair, no DNA fragments were amplified, whereas the expected-size fragments were amplified from the controls. Second, restriction enzyme digestion of DNA(s) generated by RCA from samples collected in Peru in 1998 and 2010 and Ecuador in 2010 revealed one or more fragments totaling approximately 2.6 kb, indicating a single begomovirus genomic DNA. For the 44 begomovirus-positive samples from Ecuador, digestion of RCA DNAs with the four base-cutting enzyme MspI generated with same pattern of fragments totaling approximately 2.6 kb; whereas fragments generated from an equivalent analysis of a tomato sample infected with a bipartite begomovirus from Brazil added
up to approximately 5.2 kb (data not shown). Finally, no DNA-B component was detected in DNA gel blot hybridization analyses of total genomic DNA from the ToLCD samples collected in Peru in 1998, whereas the expected hybridization signal was detected in total genomic DNA from *N. benthamiana* leaves infected with the bipartite begomovirus BDMV (data not shown). Together, these results indicated that ToLCD in Ecuador and Peru is associated with a single begomovirus genomic DNA, and are in agreement with results of Marquez-Martin et al. (43).

A single HindIII site, identified in the begomovirus genomic DNA in the ToLCD samples from Peru, was used to obtain five putative full-length clones from the two samples collected in Ica in 1998 (one from the 98-1 sample [PA98-1-1] and four from the 98-2 sample [PA98-2-1, -TM5, -TM8, -TM13]), and three clones each from a different sample collected in the same location in 2010 (PA10-3, -S8, -S9). A similar strategy was used to obtain 28 putative full-length clones from the Ecuador samples. The complete nt sequences of these clones were determined, and comparisons revealed identities ranging from 92%-100% with the DNA-A component of ToLDeV (ToLDV [PE:03], accession number GQ334472). These identities are above the threshold for defining distinct begomovirus species (89%) and within the range for strains (<94%) and variants (>94%) of a species (13). Thus, these are a collection of genetically diverse ToLDeV DNA-A clones.

**Relationships among ToLDeV isolates associated with ToLCD in Ecuador and Peru and other begomoviruses.** Comparisons performed with the complete nt sequence of the PA98-1-1 clone from Peru revealed identities of 98-99% with sequences of ToLDeV [PE:03] and all 28 clones from Ecuador, 95% with sequences of the four PA98-2 clones and 92% with those of the three PA10 clones (Table 1). Based upon the 94% threshold value for begomovirus strains, the PA98-1-1 clone, all 28 clones from Ecuador and ToLDeV [PE:03] are isolates of a ToLDeV
variant (named the PA98-1 variant); the four PA98-2 clones (with 99% identities) are isolates of another variant (named the PA98-2 variant); and the three clones from the 2010 samples (with 99% identities) are isolates of a strain (named the PA10 strain). Collectively, the variants and strain are referred to as the ToLDeV genotypes. Because of the high degree of nt sequence identity among the 28 isolates from Ecuador, the sequence of a representative isolate (EA10-LE3-5K) was used in subsequent analyses. The designations and GenBank accession numbers for these ToLDeV isolates are: EA10-LE3-5K (JX501510), PA98-1-1 (JX501502), PA98-2-1 (JX501503), PA98-2-TM8 (JX501507), PA98-2-TM13 (JX501508), and PA98-2-TM5 (JX501509), PA10-3 (JX501504), PA10-S8 (JX501505), and PA10-S9 (JX501506).

The genomic DNAs of the ToLDeV isolates ranged from 2589-2591 nts, which is typical for components of NW bipartite begomoviruses. The genome organization of these DNAs is similar to that of DNA-A components of NW bipartite begomoviruses, i.e., one virus-sense gene (V1) and no V2 gene, and four complementary-sense genes (C1, C2, C3 and C4) (Fig. 1). The CP aa sequence has the N-terminal motif PWrMaGT, which is characteristic of NW begomoviruses. An additional complementary-sense gene (C5) was identified, which encodes a predicted protein of 105 aa.

The IR sequences (between the start codons of the C1 and V1 genes) of the ToLDeV isolates are approximately 300 nt and have the characteristic geminivirus stem-loop structure with the conserved nonanucleotide sequence (TAATATT↓AC), the Rep high affinity binding site and the C1 TATA box (29). Interestingly, the ToLDeV genotypes have different Rep binding sites and inverted repeat sequences (Fig. 2A). The PA98-1 variant has an imperfect reiterated site (GGTGACTGGAGT, iteron underlined) and an ACACC inverted repeat, the PA98-2 variant has a GGGTAAAGGGGT reiterated site and AC CCC inverted repeat, and the PA10
strain has a **GGTAGAACGGTAA** reiterated site and TTACC inverted repeat (Fig. 2A). The ToLDeV genotypes also have different Rep iteron-related-domains (IRD). The PA98-1 variant has the GGTGT iteron and FKIN IRD motif, the PA98-2 variant has the GGGGT iteron and KRFKIS motif, and the PA10 strain has the GGTAG iteron and KRFQIS motif (Fig. 2B). These results were consistent with the ToLDeV variant/strain designations, and revealed an unusually high degree of genetic diversity in the IR sequences of these genotypes.

The Simplot analysis grouped the ToLDeV isolates into the PA98-1 and PA98-2 variants and PA10 strain, and revealed that the majority of the sequence divergence is in the IR and the C1 and C4 genes (Fig. 2C). This was in agreement with the divergent IR and IRD sequences (Fig. 2A and 2B), and was further supported by results of comparisons performed with IR and individual gene sequences (Table 1). Here, the right IR sequences were almost identical (>97% identity), whereas the left IR sequences were highly divergent (Table 1), especially for isolates of the PA10 strain (65-66% identity). The V1, C5, C2 and C3 genes were highly conserved among the three genotypes (≥95% nt and aa identities), whereas the C1 and C4 genes were more divergent (Table 1). In comparisons with the PA98-1 variant, the PA98-2 variant had C1/Rep and C4 nt/aa identities of 93/93% and 91/86%, respectively, whereas those of the PA10-3 variant were 88/86-87% and 86/78%, respectively.

Comparisons performed with the complete nt sequences of the ToLDeV isolates and those of other begomoviruses revealed highest identities with DNA-A components of NW bipartite begomoviruses from Latin America (Table 1). The highest identities were with Soybean blistering mosaic virus (SbBMV) from Argentina (79-86%), Tomato chino La Paz virus (ToChLPV) from Mexico (78-81%) and Merremia mosaic virus from Puerto Rico (77-81%). Similar results were obtained in comparisons made with IR and individual gene sequences.
(Table 1 and data not shown). An exception was with the C4 genes. Here, a BLAST search performed with the sequence of the C4 gene of the PA10 strain revealed highest identities (91-92%) with C4 genes of some OW monopartite tomato-infecting begomoviruses (e.g., Tomato leaf curl Mali virus, Tomato leaf curl Cameroon virus and Tomato curly stunt virus; data not shown) and slightly lower identities (89-90%) with AC4 sequences of NW bipartite begomoviruses (e.g., Rhynchosia golden mosaic virus and Sida golden mosaic virus). The C4 sequence of the PA98-1 variant had highest identities (88-89%) with AC4 and C4 sequences, whereas that of the PA98-2 variant had highest identities (89%) with AC4 sequences and slightly lower identities (88%) with C4 sequences (data not shown).

Results of the phylogenetic analyses performed with the complete and C4 nt sequences of the ToLDeV isolates and those of the most closely related begomovirus components/genomes are presented in Fig. 3. In the analysis with the complete nt sequences, NW and OW begomoviruses were placed in two strongly supported clades (Fig. 3A). The ToLDeV isolates were placed in a strongly supported clade (the ToLDeV clade), which was contained within the NW begomovirus clade. The ToLDeV clade had three groups: one with isolates of the PA98-1 variant, one with isolates of the PA98-2 variant, and the other with isolates of the PA10 strain. The sequence of the DNA-A component of ToChLPV was most closely related to those of the ToLDeV isolates, and was placed in a strongly supported group (100% credibility value) with the ToLDeV clade (Fig. 3A). The sequences of DNA-A components of other tomato- and weed-infecting bipartite begomoviruses, mostly from South America (Argentina, Bolivia, Brazil, and Venezuela), were the next most closely related and were placed in a large strongly supported group (100% credibility value) with the ToLDeV clade and ToChLPV (Fig. 3A).
The phylogenetic relationship revealed by the analysis of the AC4/C4 sequences was incongruent, as a clear division between NW and OW begomoviruses was not revealed. Moreover, the ToLDeV isolates were not placed into a single clade. The ToLDeV variants were placed in a strongly supported clade with two groups, one with PA98-1 isolates and the other with the PA98-2 isolates (Fig. 3B). The isolates of the PA10 strain were placed in a different clade, which included C4 sequences of OW monopartite tomato-infecting begomoviruses from Africa and the AC4 sequence of a cassava-infecting bipartite begomovirus from Burkina Faso (Fig. 3B). Although this clade had a lower credibility value (70%), this relationship was in agreement with the BLAST search results. The clades with the ToLDeV genotypes were included in a large strongly supported clade (99% credibility value) that included AC4 sequences of NW bipartite begomoviruses from South America (Fig. 3B). Another large and strongly supported clade (96% credibility value) included AC4/C4 sequences of other NW bipartite and OW monopartite begomoviruses (Fig. 3B).

Together, these results indicated that the genome of ToLDeV is composed of a single genetically divergent genomic DNA that is most closely related to the DNA-A components of NW bipartite begomoviruses. However, these results did not reveal the genetic mechanism(s) underlying the diversity in the ToLDeV genome or resolve the etiology of ToLCD.

Recombination and mutation have driven the emergence and evolution of ToLDeV genome. The results of the Simplot analysis and sequence comparisons provided evidence of recombination in the ToLDeV genome, i.e., a divergent LIR/C1/C4 region embedded within the highly conserved remainder of the genome. Inspection of the IR/C1/C4 sequence alignment confirmed the high degree of sequence divergence in this region and revealed potential recombination junctions in the IR and C1 gene, respectively (Fig. S1). Furthermore, the
distribution of nt differences in this divergent region was correlated with the ToLDeV genotypes (Fig. S1). Considered together with the different Rep binding sites and IRDs (Fig. 2A and 2B), these results provided evidence of at least two recombination events in the evolution of these genotypes. Unexpectedly, the RDP3 analyses with different datasets (including one with 180 full-length sequences of bipartite and monopartite begomoviruses) did not reveal convincing evidence of recombination in the ToLDeV genome (i.e., $P$ values $>10^{-4}$). However, this may be due to the absence of parental sequences in the database (note the ToLDeV is the first begomovirus described from Peru).

Genetic diversity values revealed the greatest variation in the ToLDeV C4 gene followed by the IR, the C1 gene and the V1 gene (Table 2). A Bayesian analysis revealed strong constrained selection in the C1 and V1 genes, i.e., negative selection for 78 and 93% of sites, respectively, whereas the C4 gene showed less constrained selection with 50% of sites under negative selection (Table 2). The C4 gene showed the highest level of positive selection (20.6% of sites) compared with the C1 (10.4%) and V1 (4.2%) genes. Moreover, the finding that 20.6% of sites in the C4 gene were under positive selection compared with only 9.5% of sites in the overlapping region of the C1 gene (Table 2) revealed differential evolutionary pressure on these genes.

The overlapping C1/C4 region consists of 84 codons, with the first nt position in C1 codons corresponding to the third position in C4 codons (C1-1/C4-3), the second position in C1 codons corresponding to the first position in C4 codons (C1-2/C4-1) and the third position in C1 codons corresponding to the second position in C4 codons (C1-2/C4-3) (Fig 4A). Evolutionary constraints in this region were assessed by determining the entropy values for nt and aa positions (Fig. 4B). Variation at the Rep-2/C4-1 nt sites, which had the highest probability of altering the
aa sequences of both proteins, was the lowest (cumulative entropy value of 2.4) compared with values of 12.3 and 16.1 for the Rep-1/C4-3 and Rep-3/C4-2 positions, respectively. Furthermore, aa changes in the Rep protein were largely determined by nt substitutions at the C1-1/C4-3 sites, whereas those in the C4 protein were determined by substitutions at the C1-3/C4-2 sites (Fig. 4B). The SNAP analysis, which plots the cumulative occurrence of synonymous and nonsynonymous substitutions, showed that the C1 gene accumulated more synonymous mutations, whereas the C4 gene accumulated more nonsynonymous mutations (Fig. 4C). Together, these results revealed positive and negative selection in the overlapping sequences of the C4 and C1 genes, respectively, thereby allowing for more rapid evolution of the C4 protein while minimizing aa changes in the more conserved Rep protein.

Inspection of the AC4/C4 aa alignment revealed a high level of variability among these sequences, especially in the C-terminus (Fig S2). This is in agreement with a SNAP analysis, which showed a higher proportion of nonsynonomous sites in the 3’ region of the ToLDeV C4 gene (positions 50 to 84, data not shown). There was no evidence that the C4 genes of the ToLDeV genotypes (or AC4 genes of NW bipartite begomoviruses) were acquired by recombination with an OW begomovirus (i.e., extensive regions of shared aaas), and this is in agreement with RDP results that did not reveal such a recombination event (note that these potential OW begomovirus parental sequences were in some of the datasets used in these analyses). For example, although the C4 of the PA10 strain had slightly higher identities with C4 sequences of some OW monopartite begomoviruses, the alignment revealed a similar level of divergence (approximately 24% of residues) for OW C4 and NW AC4 sequences. The aa differences were also distributed throughout the sequence, and there were residues in the PA10 C4 sequence that were either unique (4 residues), found mostly in NW sequences (6 residues) or
found mostly in OW sequences (3 residues) (Fig. S2). Together, these results suggest that the
diversity in the AC4/C4 genes arose through independent mutational events and that the
relatively high level of aa sequence identity among the ToLDeV C4 proteins and those of OW
monopartite begomoviruses are an example of convergent or parallel evolution, as defined in
terms of the independent evolution of a given phenotype or trait (2).

The single genomic DNA of ToLDeV is infectious and induces leaf curl symptoms in
tomato and N. benthamiana. Tomato plants individually agroinoculated with the multimeric
clones of the PA98-1-1 and PA98-2-1 variants and the PA10-3 strain developed stunting and
upward leaf curling and swollen purple veins, whereas plants inoculated with the empty vector
control did not develop symptoms (Fig. 5A). Importantly, these symptoms were
indistinguishable from those associated with ToLCD in the field in Ecuador and Peru. N.
benthamiana plants agroinoculated with these clones developed stunting, leaf curling and vein
swelling (Fig. 5B). In both plant species and for all three clones, disease symptoms developed 9-
12 dpi and rates of infectivity were high (92-100%). The presence of viral DNA in leaves of
representative symptomatic plants was confirmed by PCR with the PAR1c496/ToLCv1730
primer pair. These results fulfilled Koch’s postulates for ToLCD and provided conclusive
evidence that ToLDeV is a monopartite begomovirus.

Distinct symptom phenotypes are induced in tomato and N. benthamiana plants by
the genomic DNAs of the ToLDeV genotypes. Tomato plants agroinoculated with the PA98-1-
1 variant developed severe leaf curl symptoms, but relatively little stunting. By approximately 30
dpi, the symptoms in newly emerging leaves were attenuated (Fig. 5A), and this recovery
phenotype continued for up to 60 dpi. Severe leaf curl symptoms also developed in tomato plants
agroinoculated with the PA98-2-1 variant, but these plants showed stunting and little or no
recovery (Fig. 5A). Tomato plants agroinoculated with the PA10-3 strain developed more severe stunting and leaf curl symptoms than plants inoculated with either variant, and infected plants did not recover (Fig. 5A). Similar results were observed in agroinoculated *N. benthamiana* plants; here, the symptoms induced by the PA10-3 strain were substantially more severe than those induced by either variant (Fig. 5B).

The recovery phenotype in plants infected with begomoviruses is a type of defense response and is associated with reduced levels of viral DNA (5, 25). Results of DNA gel blot hybridization analyses indicated that viral DNA levels were similar and relatively high for all three genotypes at 21 dpi when infected plants showed severe leaf curl symptoms (Fig. 5C). However, by 60 dpi, when plants infected with the PA98-1-1 variant had undergone recovery, viral DNA levels were reduced in these plants compared with those in plants infected with the PA98-2-1 variant and PA10-3 strain, which had not recovered (Fig. 5C). These results were observed in three independent experiments, and suggest that the recovery phenotype is associated with a reduced level of viral DNA.

Finally, the differential pathogenicity of these ToLDeV genotypes was quantified by determining the biomass of infected tomato plants. Regardless of the ToLDeV genotype, the biomass of infected plants was significantly less than that of the empty vector control plants. The biomass of plants infected with the highly virulent PA10-3 strain was significantly less than that of plants infected with the PA98-1-1 or PA98-2-1 variants (Fig. 5D).

**ToLDeV has a narrow host range and a resistant symptom phenotype in tomato germplasm with the Ty-1 gene.** The host range of the ToLDeV genotypes (the PA98-1-1 and PA98-2-1 variants and the PA10-3 strain) was identical (Table 3). Three types of symptom
phenotypes were observed: symptomatic infection, symptomless infection and no infection (no symptoms and no viral DNA detected in newly emerged leaves). Disease symptoms developed only in tomato and *N. benthamiana* plants, whereas symptomless infections were detected in the other solanaceous species (Table 3). *C. quinoa*, *C. amaranticolor*, *C. pepo* cv. Small sugar pumpkin and *P. vulgaris* cv. Topcrop plants were not infected. These results indicated that the host range of ToLDeV is limited to solanaceous plants.

The *Ty-1* gene confers resistance to monopartite begomoviruses, such as TYLCV, but not to many bipartite begomoviruses (77). In control experiments, cv. Dominator plants (with *Ty-1*) agroinoculated with TYLCV-IL [DO] did not develop symptoms, whereas all cv. Glamour plants (without *Ty-1*) developed yellow leaf curl symptoms. In contrast, all cv. Dominator and cv. Glamour plants agroinoculated with the bipartite begomovirus ToMHV developed leaf crumpling and yellow mottling symptoms. In experiments with the ToLDeV genotypes, LA3473 and cv. Dominator plants (with *Ty-1*) agroinoculated with each of the three ToLDeV genotypes did not develop symptoms (10-21 plants inoculated/genotype). In contrast, all LA3474 and cv. Glamour plants (without *Ty-1*) agroinoculated with the ToLDeV genotypes developed symptoms of stunting and leaf curl (10-16 plants inoculated/genotype). These results established that the *Ty-1* gene confers resistance to the three ToLDeV genotypes (i.e., none of the genotypes overcame this resistance), and supported the notion that this gene confers resistance to monopartite but not bipartite begomoviruses.

**Biological properties of ToLDeV are similar to those of monopartite begomoviruses.** Tomato and *N. benthamiana* seedlings rub-inoculated with sap prepared from leaf tissue from tomato plants with leaf curl symptoms collected in 1998 or from leaves of symptomatic *N. benthamiana* plants infected with the ToLDeV PA10-3 strain did not develop
symptoms, nor was viral DNA detected in newly emerged leaves (15 total plants inoculated in three independent experiments). In contrast, 14 of 15 N. benthamiana plants inoculated with sap from leaves infected with the sap-transmissible bipartite begomovirus BDMV developed stunting and leaf curling, crumpling and mosaic symptoms. Immunolocalization experiments performed with petioles and leaves from ToLDeV-infected tomato plants revealed virus only in phloem-associated cells, e.g., companion and phloem parenchyma cells (Fig. 6). No signal was detected in petiole and leaf tissues of infected plants probed with preimmune antisera, or in uninfected plants probed with CP antiserum (Fig. 6, data not shown). These results indicate that ToLDeV is phloem-limited, which is fully consistent with the lack of sap transmissibility.

Mutational analyses reveal functional homology of the ToLDeV V1 and C4 genes with monopartite begomovirus counterparts. In leaf disc replication assays, DNA gel blot hybridization analyses revealed replication of the wild-type genomic DNA of all three ToLDeV genotypes, including open circular double-stranded (oc) and single-stranded (ss) DNA forms. Similar levels of replication, as well as DNA forms, were detected in leaves agroinfiltrated with the v1, c4 and c5 mutants of all three genotypes (Fig. 7A). Replication was not detected for the c1 mutants, consistent with the requirement of Rep for viral DNA replication (29).

The c1 (Rep) and v1 (CP) mutants did not systemically infect tomato or N. benthamiana plants based upon lack of symptom development and failure to detect viral DNA in newly emerged leaves by PCR (Table 4). Tomato plants agroinoculated with the c4 mutants developed no symptoms or an attenuated symptom phenotype (Fig. 7B to D), and PCR analyses revealed that only plants with the attenuated symptoms were infected (Table 4). Sequencing of the PCR-amplified fragment with the C4 gene from selected plants with the attenuated symptom phenotype confirmed the presence of the mutation. N. benthamiana plants agroinoculated with
the c4 mutants developed no obvious symptoms (Fig. 7E to G), but PCR analyses revealed symptomless infections in most plants (Table 4). Mutation of the C5 gene had no effect on viral infectivity and symptom development, with the exception of the PA10-3 c5 mutant, which induced less severe symptoms in N. benthamiana plants (Table 4). Sequencing of the PCR-amplified fragment with the C5 gene from selected symptomatic tomato and N. benthamiana plants agroinoculated with the c5 mutant confirmed the presence of the mutation. Together, these results indicate that 1) Rep and CP are required for systemic infection, 2) C4 is a pathogenicity factor and 3) C5 does not play a major role in pathogenicity.

The stability of the c4 mutation was assessed by transmission through tomato plants following grafting. Uninfected scions (cv. Glamour) grafted onto rootstocks infected with the PA98-1-1 (14 plants) or PA10-3 c4 mutants (8 plants) became infected and developed the attenuated symptoms associated with the c4 mutation. Sequencing of the PCR-amplified C4 fragment from these scions confirmed the presence of the mutation (Table 4). In contrast, all scions grafted onto rootstocks infected with the PA98-2-1 c4 mutant (8 plants) developed wild-type symptoms (stunting and severe leaf curling), and sequencing of the PCR-amplified C4 fragment from these scions revealed that the c4 mutation had reverted to the wild-type sequence (Table 4). These results are consistent with the C4 functioning as a pathogenicity factor, and reveal positive selection pressure on the C4 gene during infection.

**Infectivity and symptomatology of chimeric components generated between two ToLDeV genotypes.** The role of the divergent region of the ToLDeV genome was further investigated by generating chimeras between the PA98-1-1 variant, which induces less severe symptoms and shows the recovery phenotype, and the PA10-3 strain, which induces severe symptoms and shows no recovery (Fig. 1 and 5A). The PA98-1-1/PA10-3 chimera composed of
the PA98-1-1 variant with the 1.5 kb fragment with the IR/C4/C1 sequences from the PA10-3 strain induced symptoms in tomato and *N. benthamiana* plants that were more severe than those induced by the PA98-1-1 variant, but less severe than those induced by the PA10-3 strain, and recovery was less pronounced (Fig. 8). The PA10-3/PA98-1-1 chimera induced symptoms that were less severe than those induced by the PA10-3 strain, but more severe than those induced by the PA98-1 variant. Plants infected with this chimera showed a partial recovery phenotype (Fig. 8). Together, these results support the hypothesis that this divergent region of the viral genome, including the C4 gene, plays a role in the pathogenicity of ToLDeV.

**DISCUSSION**

ToLCD in Ecuador and Peru is caused by an indigenous NW monopartite begomovirus. Based upon the known genome structure of NW begomoviruses, it was presumed that ToLDeV would have a bipartite genome (43). The finding that ToLDeV isolates associated with ToLCD in Ecuador and Peru have a single genomic DNA, with no associated DNA-B component or satellite DNAs, confirmed and extended the results of Marquez-Martin et al. (43) and indicated an unconventional etiology for the disease. Here, we establish that ToLDeV is an indigenous NW monopartite begomovirus, and is the causal agent of ToLCD in Ecuador and Peru. The induction of leaf curl symptoms in tomato, indistinguishable from those observed in the field, by the cloned genomic DNA of ToLDeV fulfilled Koch’s postulates for ToLCD and confirmed the monopartite nature of the viral genome. The finding that ToLDeV has biological properties similar to those of OW monopartite tomato-infecting begomoviruses, e.g., lack of sap transmission, phloem limitation, a resistance symptom phenotype in tomato germplasm with the
Ty-1 gene and functional properties of the V1 and C4 genes, provided further evidence for a monopartite genome. Several lines of evidence indicated that ToLDeV originated in the NW, rather than being an introduced exotic OW monopartite begomovirus, like TYLCV. First, the size (approximately 2.6 kb) and organization of the ToLDeV genome, including the lack of a V2 gene, are similar to DNA-A components of NW bipartite begomoviruses. Second, ToLDeV has the characteristic NW begomovirus motif, PWRsMaGT, in the N-terminus of the CP. Finally, sequence comparisons and phylogenetic analyses showed that ToLDeV was most closely related to DNA-A components of NW bipartite begomoviruses. Thus, these results established that ToLDeV is an indigenous NW monopartite begomovirus, which probably evolved from the DNA-A component of a local bipartite begomovirus progenitor.

Although the DNA-A components of some NW bipartite begomoviruses are infectious in the experimental host, N. benthamiana, these infections have not been found in nature and are not examples of NW monopartite begomoviruses. On the other hand, ToLDeV is a bona fide NW monopartite begomovirus because it causes a disease under natural conditions and possesses biological properties that are typical of monopartite begomoviruses. However, the identification of ToLDeV raises the question of what factor(s) led to the evolution of a NW monopartite begomovirus at this point in time. It is generally believed that the first begomoviruses were monopartite, with the bipartite genome appearing later, either through component duplication and diversification or acquisition and modification of satellite DNAs (4, 60, 63). Furthermore, because bipartite begomoviruses occur in NW and OW locations (4), it is likely the bipartite genome evolved before continental drift. The absence of indigenous monopartite begomoviruses in the NW can be explained by the predominance of bipartite progenitor(s) on the landmasses that gave rise to the NW, or a subsequent introduction event. Local evolution events then drove...
the emergence of new begomovirus species that reflected the genome structure of the virus(es) in
a given geographical area, i.e., bipartite species in the NW (14, 16, 19, 26, 59, 67) and
monopartite (36, 79) and, to a lesser extent, bipartite species in the OW (53). In addition, a few
evolutionary events in the OW have involved changes in the viral genome structure, e.g., the
aforementioned TYLCTHV, which can exist as a bipartite or monopartite virus, and the
monopartite Sri Lankan cassava mosaic virus, which acquired a DNA-B component genome and
became bipartite (65). However, the emergence of a monopartite begomovirus in the NW
represents a new evolutionary direction, which probably occurred due to a specific combination
of factors. First, extensive cultivation of tomato provided a highly susceptible host in which a
monopartite begomovirus could evolve in the absence of a betasatellite. Second, a suitable local
progenitor bipartite begomovirus existed in native or cultivated plants in Peru. Finally, the
invasive and polyphagous B biotype of *B. tabaci* mediated the introduction of this local
progenitor begomovirus into crop plants, such as tomato (45).

The evolution of ToLDeV likely started with the acquisition of indigenous bipartite
begomoviruses by the polyphagous whitefly vector, and delivery of virions into the phloem of
tomato plants. During this process, DNA-A-containing virions are delivered, via the phloem, to
replication-competent cells in the protophloem (73). Here, DNA-A infections are initiated, either
in the absence of the DNA-B component or possibly following the loss of this component. Viral
spread occurred via cell division or cell-to-cell movement mediated by DNA-A encoded proteins
(e. g., CP or C4). Nascent DNA-A containing virions were released into sieve elements, where
they were acquired and spread by whiteflies. Further evolution of the monopartite genome
occurred via mutation and recombination (see below). This evolutionary scenario gains support
from the aforementioned limited infectivity of some NW DNA-A components (16, 33), and fits
with the classical role of bottlenecks in the emergence and evolution of plant and animal virus populations during the process of within-host spread and host-to-host transmission (23).

The reason that more DNA-A components of NW bipartite begomoviruses have not evolved into a monopartite tomato-infecting begomoviruses may be due to the lack of movement-related functions encoded by the DNA-B component and/or the AV2 gene. The functional homology of the ToLDeV CP and C4 with monopartite begomovirus counterparts indicates a key role for these genes in the evolution of this NW monopartite begomovirus. The essential role of the ToLDeV CP in systemic infection is probably due to nucleocytoplasmic transport and long-distance movement functions. In bipartite begomoviruses, nucleocytoplasmic transport is provided by the NSP, whereas in monopartite viruses this function is provided by the CP (60, 61, 68). The likelihood that the ToLDeV CP provides this function is supported by the capacity of the CP of some NW begomoviruses to complement, enhance and mask NSP function (54, 60, 78). The CP also plays a key role in the long-distance movement of monopartite begomoviruses, and this has been associated with virion formation (35, 61). Thus, the CP probably plays a similar role in the long distance movement of ToLDeV. As a phloem-limited monopartite begomovirus, ToLDeV may not require the cell-to-cell movement function provided by the MP or, alternatively, this may be provided by C4 (35, 61). Finally, the findings that the C5 gene was not involved in pathogenesis and the V2 gene was absent indicated that neither gene played a role in the evolution of ToLDeV.

A number of lines of evidence indicated that the ToLDeV C4 is a pathogenicity factor. First, c4 mutants induced attenuated or no symptoms and showed reduced infectivity in tomato. Second, distinct symptom phenotypes were associated with the C4 genes of the ToLDeV genotypes. Finally, the reappearance of wild-type symptoms after graft transmission of the
PA98-2 c4 mutant was associated with reversion of the mutation. Based on these results and studies of C4 function in OW begomoviruses (1, 9, 35, 55, 71, 72), the ToLDeV C4 could be involved in movement and suppression of host defense responses. The recovery phenotype in begomovirus-infected plants is a host defense response that involves methylation of the viral genome and a reduction in viral titer and symptom severity (25, 55, 57). The differences among ToLDeV genotypes in virulence, viral DNA levels and recovery phenotypes in infected plants suggests that the C4 functions as a suppressor of host defense responses. The high levels of aa identity among NW and OW AC4/C4 proteins (16 and Fig. 3B and S2) may indicate a common function(s) among these proteins (e.g., in defense suppression), and several lines of evidence indicated that some of these similarities have appeared via convergent evolution. However, it will be important to experimentally demonstrate a common function for these proteins.

Recombination and mutation have driven the emergence and evolution of the ToLDeV genotypes. The genetic diversity and differential virulence of the ToLDeV genotypes revealed evidence of viral evolution in the direction of increased virulence (e.g., the highly virulent PA10 strain associated with ToLCD in Peru in 2010). Low levels of diversity and purifying selection in the V1 and C1 genes reflected the conserved functions of the CP and Rep proteins, whereas the higher level of diversity and positive selection in the C4 gene indicated a role in the evolution of ToLDeV. The majority of the genetic diversity in the ToLDeV genome was in the LIR, C4 and C1 genes, which is a well-known recombinational hot-spot in the genomes of bipartite and monopartite begomoviruses (6, 32, 39). Evidence for recombination in this region of the ToLDeV genome came from the Simplot analysis (Fig. 2C), the different Rep binding sites and IRD sequences, and the identification of potential recombination junctions (Fig. S1). Recombination in this region may provide a selective advantage for ToLDeV genotypes,
e.g., the acquisition of a C4 with strong defense suppression properties. The failure of the RDP analysis to detect this recombinant region of the ToLDeV genome shows the importance of using multiple approaches to detect recombination in begomovirus genomes.

The relatively recent emergence of the ToLDeV (within the past 20 years) together with the relatively high degree of negative selection on begomovirus coding regions (17, 64) and the estimated mutation rates for these viruses (10^{-3}-10^{-4} substitution site^{-1} year^{-1}) (11, 12) suggests that the high level of genetic diversity in the ToLDeV genotypes was due to mutation and recombination. Thus, many of the nt differences in the divergent region were from mutations that occurred prior to the recombination events. However, random mutations would not explain the different diversity values and evolutionary constraints identified in the overlapping sequences of the C1 and C4 genes, which encode the conserved Rep and less conserved C4 proteins, respectively. The genetic analysis revealed that variation in the C4 protein was achieved through mutations that did not alter the Rep, whereas mutations that altered both proteins were rare (Fig. 4A and B). This finding was supported by SNAP and Bayesian analyses showing that positions in the 3’ end of the C4 gene, which were under adaptive evolution (positive selection), were under purifying selection in the overlapping portion of the C1 gene. This explains how convergent evolution in AC4 genes of NW bipartite begomoviruses can generate AC4/C4 proteins with functional similarity to C4 counterparts in OW monopartite begomoviruses, while maintaining the more conserved Rep protein sequence. This strategy, which allows for maximizing the viral genome size and maintaining differential rates of mutation in overlapping genes, is used by a range of animal and plant-infecting viruses (76).

The emergence of an indigenous NW monopartite begomovirus is a new evolutionary direction, and another example of the remarkable genetic flexibility of these viruses and their
capacity to exploit niches in agricultural ecosystems. Definitive evidence that ToLDeV is a bone
fide NW monopartite begomovirus came from the characterization of the viral genome and
biological properties, and the completion of Koch’s postulates for ToLCD. It will be of interest
to identify the progenitor(s) of the ToLDeV genotypes and to further investigate the function(s)
of the viral gene products. The finding that the Ty-1 gene confers resistance to ToLDeV and the
identification of a time of year when virus and whitefly incidence are low can be used to develop
an integrated pest management strategy for ToLCD can be developed (20).

ACKNOWLEDGMENTS

We thank the Servicio Nacional de Sanidad Agraria, Peru for assistance in carrying out
the July 2010 disease survey. This research was funded in part by a grant from the United States
Agency for International Development (USAID) as part of the Integrated Pest Management
Collaborative Research Support Program (IPM-CRSP) through Cooperative Agreement No.
EPPA-00-04-00016-00. T.A.M. was supported by fellowships from The Fulbright Program and
The Organization of American States.

REFERENCES

phenotypes produced in response to transient expression of genes encoded by four
distinct begomoviruses in Nicotiana benthamiana and their correlation with the levels of


FIGURE LEGENDS

FIG. 1. Organization of the genomic DNA of the PA98-1-1 (black) and PA10-3 (grey) isolates of Tomato leaf deformation virus (ToLDeV) cloned from plants with symptoms of tomato leaf curl disease collected in Peru in 1998 and 2010, respectively. The intergenic region (IR) and viral genes, including the V1, C1, C2 and C3 predicted to express the capsid (CP), replication-associated (Rep), transactivator (TrAP) and replication enhancer (REn) proteins, respectively, are shown. Chimeras generated between the genomic DNAs of these isolates by exchanging an approximately 1.5 kb EcoRV/KpnI fragment with the IR and C4 and C1 genes are shown beneath the line.

FIG. 2. Genetic diversity among Tomato leaf deformation virus isolates collected from Peru in 1998 (PA98-1-1, PA98-2-1, PA98-2-TM5, PA98-2-TM8 and PA98-2-TM13), 2003 (ToLDeV [PE:03] GenBank accession number GQ335572) and 2010 (PA10-3, PA10-S8 and PA10-S9) and Ecuador in 2010 (EA10-LE3-5K). (A) An alignment of intergenic region nucleotide (nt) sequences with the replication-associated protein (Rep)-binding site. The sequence and orientation of the Rep binding site iteron and inverted repeat are indicated with boxes and arrows, respectively; and the TATA box of the C1 (Rep) gene is shown with a horizontal bracket. Vertical brackets show isolates with identical iteron and inverted repeat sequences, and that correspond to the PA98-1 and PA98-2 variants and the PA10 strain of ToLDeV. (B) An alignment of N-terminal Rep protein amino acid sequences with the iteron-related domain (IRD). The IRD is shown in bold and underlined, and the Rep motifs I, II, and III and GRS are shown in boxes. Vertical brackets show isolates with identical IRD sequences, and that correspond to the
PA98-1 and PA98-2 variants and the PA10 strain of ToLDeV. (C) A similarity plot analysis of total nt sequences of the genomic DNA of isolates of ToLDeV collected from Peru in 1998 and 2010 and Ecuador in 2010. The short-dashed line is the PA98-1 variant (isolates PA98-1-1 and EA10-LE3-5K), the long-dashed line is the PA98-2 variant (isolates PA98-2-1, -TM5, -TM8 and -TM13) and the solid line is the PA10 strain (isolates PA10-3, -S8 and -S9).

FIG. 3. Phylogenetic consensus trees showing the relationships among Tomato leaf deformation virus (ToLDeV) isolates collected from Peru in 1998 (PA98-1-1, PA98-2-1, PA98-2-TM5, PA98-2-TM8 and PA98-2-TM13), 2003 (ToLDeV [PE:03] GenBank accession number GQ335572) and 2010 (PA10-3, PA10-S8 and PA10-S9) and Ecuador in 2010 (EA10-LE3-5K) and the most closely related genomes of New World (DNA-A component) and Old World begomoviruses based on alignments of (A) total nucleotide (nt) and (B) C4 gene nt sequences. Phylogenetic analyses were performed with MrBayes 3.2 (62). Branch strengths were evaluated by Bayesian posterior probabilities. Nodes with clade credibility values >60% are shown. The curtovirus Beet severe curly top virus (BSCTV) was used as an out-group. The scale bar indicates length of horizontal branches corresponding to the rate of substitution/nucleotide. The following sequences were obtained from GenBank and used for comparisons and phylogenetic analysis: Abutilon mosaic virus-[Hawaii] (AbMV-[US:HW]) and AbMV-India [AbMV-IN]; African cassava mosaic Burkina Faso virus (ACMVFV-[BF:OUa:08]); Cabbage leaf curl Jamaica virus-[Jamaica] (CabLCJV [JM:Cuc-3]); Merremia mosaic virus (MerMV)-[Puerto Rico] (MerMV-[PR]), MerMV-[PR:PR80-H3] and MerMV-Venezuela (MerMV-[VE:Trujillo]); Pepper leaf curl virus-[Thailand] (PepLCV-[TH]); Potato yellow mosaic virus-Trinidad (PYMV-[TT]) and PYMV-[VE]; Sida golden mosaic Liguanea virus-[Florida] (SiGMV-[US:FL:09]) and Sida golden mosaic virus-[Florida] (SiGMV-[US:FL]); Sida micrantha mosaic
FIG. 4. Mutations in the reading frames of the overlapping C1 and C4 genes in the Tomato leaf deformation virus (ToLDeV) genome result in different rates of amino acid changes in the conserved replication-associated (Rep) and more divergent C4 proteins, respectively. (A) In this overlapping region, a nucleotide (nt) substitutions in the third position of the C1 codon are likely to be synonymous mutations, whereas the same substitutions in the second position of the C4 codon are more likely to be nonsynonymous mutations (arrows). An example in which a nt substitution in codon 110 of the C1 gene results in a synonymous mutation, whereas the same mutation results in a nonsynonymous mutation in codon 58 of the C4 gene. (B) Levels of nt and aa variation relative to the nt positions of the codons in the reading frames of the overlapping C1 and C4 genes. Variation in nt of the three codons: C1-1/C4-3 (blue), C1-2/C4-1 (red) and C1-3/C4-2 (green), are shown in relation to the aa variation in the Rep and C4 proteins. (C) Results of a SNAP analysis (38) showing cumulative synonymous/nonsynonymous codon changes from the 5’ (N termini [left]) to the 3’ (C termini [right]) of the C1 (Rep) and C4 (C4) genes in the overlapping region. Nonsynonymous mutations in C4 codons, resulting in positive selection, are found at the 3’ end of the gene; whereas these mutations are synonymous in the overlapping codons of the C1 gene and show purifying selection or neutrality.

FIG. 5. Disease symptoms induced by the cloned genomic DNAs of three genotypes of Tomato leaf deformation virus (ToLDeV). The PA98-1-1 and PA98-2-1 variants came from samples of tomato leaf curl disease collected in 1998, whereas the PA10-3 strain came from a sample collected in 2010. From leaf to right, empty vector (EV) control and symptoms induced by the PA98-1-1 and PA98-2-1 variants and PA10-3 strain in (A) tomato (cv. Glamour) and (B) Nicotiana benthamiana plants at 50 and 25 days after agroinoculation, respectively. (C) Viral DNA levels in newly emerged leaves of infected tomato plants (cv. Glamour) at 21 and 60 days.
post agroinoculation with the ToLDeV genotypes as determined by DNA gel blot hybridization with a probe composed of a mixture of the cloned DNA components of these genotypes. The numbers 1 and 2 represent individual plants. Abbreviations: oc, open circular double-stranded and ss, single-stranded forms of viral DNA; and M, the size marker (kb). (D) Biomass of tomato plants determined 50 days after agroinoculation with the EV control, the PA98-1-1 and PA98-2-1 variants and the PA10-3 strain. Biomass is presented as a percentage relative to the EV control. Bars with the same letters are not significantly different ($P$ value >0.05) according to Tukey’s test.

**FIG. 6.** Immunolocalization of Tomato leaf deformation virus (ToLDeV) in petiole tissues from tomato plants with leaf curl disease symptoms four weeks after agroinoculation with the cloned genomic DNAs of the PA98-1-1 and PA98-2-1 variants. (A) Section of a petiole from a tomato plant agroinoculated with the empty vector control and probed with begomovirus capsid protein (CP) antiserum. (B) and (C) Sections of a leaf petiole from a ToLDeV-infected plant, probed with CP antiserum, and showing signal over cells of the internal and external phloem. (D) Section of a leaf petiole from a ToLDeV-infected plant showing signal over companion cells and sieve elements. Abbreviations: cc, companion cells; ep, external phloem; ip, internal phloem; se, sieve element; x, xylem.

**FIG. 7.** Replication competence and infectivity of Tomato leaf distortion virus (ToLDeV) mutants. (A) Results of replication assays in *Nicotiana benthamiana* leaf discs for ToLDeV wild-type (wt) and v1, c1, c4 and c5 mutants of the genomic DNAs of three genotypes of ToLDeV: the PA98-1-1 and PA98-2-1 variants and the PA10-3 strain. Abbreviations: oc, open circular double-stranded and ss, single-stranded forms of viral DNA; and M, the size marker (kb). (B-G) Disease symptoms induced by c4 mutants of the three ToLDeV genotypes in tomato (cv.
Glamour) and *N. benthamiana* plants. From left to right, the empty vector control and symptoms induced by the c4 mutants and wild-type in tomato plants for (B) the PA98-1-1 variant, (C) PA98-2-1 variant and (D) PA10-3 strain; and in *N. benthamiana* plants for (E) the PA98-1-1 variant, (F) PA98-2-1 variant and (G) PA10-3 strain.

**FIG. 8.** Disease symptoms induced by the cloned DNA genome of the Tomato leaf deformation virus (ToLDeV) PA98-1-1 variant, PA10-3 strain and chimeras generated between these genotypes 21 and 32 days after agroinoculation of *Nicotiana benthamiana* and tomato plants, respectively. Symptoms induced in *N. benthamiana* plants by the empty vector control (A), the PA98-1-1 variant (B), the PA98-1-1/PA10-3 chimera (C), the PA10-3 strain (D) and the PA10-3/PA98-1-1 chimera (E). Symptoms induced in tomato plants by the empty vector control (F), the PA98-1-1 variant (G), the PA98-1-1/PA10-3 chimera (H), the PA10-3 strain (I) and the PA10-3/PA98-1-1 chimera (J). For a schematic representation of the wild-type and chimeric genomes see Fig. 1.
A

<table>
<thead>
<tr>
<th>Genotype of ToLeV</th>
<th>Nucleotide sequence</th>
<th>C1 (Rep) codon</th>
<th>C4 codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA98-1</td>
<td>..ACAC..</td>
<td>ACA = Thr</td>
<td>CAC = His</td>
</tr>
<tr>
<td>PA98-2</td>
<td>..ACTT..</td>
<td>ACT = Thr</td>
<td>CTT = Leu</td>
</tr>
<tr>
<td>PA10</td>
<td>..ACCC..</td>
<td>ACC = Thr</td>
<td>CCC = Pro</td>
</tr>
</tbody>
</table>

B

Rep-associated protein:

Amino acids

C1-1/C4-3

C4 protein:

Amino acids

C1-3/C4-2

C1-2/C4-1

Codon number in C4 gene

Cumulative nonsynonymous substitutions (SNAP)

C1 (Rep)  
C4

Cumulative synonymous substitutions (SNAP)

C1 (Rep)  
C4
TABLE 1 Nucleotide (nt) identities (%) for total, intergenic region (IR), left IR (LIR) and right IR (RIR) sequences, and nt and amino acid (aa) identities (similarities) for genes of Tomato leaf deformation virus (ToLDeV) isolates from Ecuador and Peru and those of most closely related begomoviruses\(^a\).

| ToLDeV isolate/ Begomovirus species | Year collected | Geographic location | Total nt nt IR CR IR LIR V1 C1 C2 C3 C4 C5 nt nt aa nt aa nt aa nt aa nt aa nt aa |
|------------------------------------|----------------|---------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| PA98-2-1                           | 1998           | PE                  | 95                  | 92                   | 85                   | 98                   | 98                   | 98                   | 99                   | 93                   | 93                   | 96                   | 94                   | 95                   | 97                   | 96                   | 98(97)               |
| PA98-2-TM5                         | 1998           | PE                  | 95                  | 92                   | 85                   | 98                   | 98                   | 98                   | 99                   | 93                   | 93                   | 96                   | 95                   | 97                   | 97                   | 98                   | 97(98)               |
| PA98-2-TM8                         | 1998           | PE                  | 95                  | 92                   | 85                   | 97                   | 98                   | 98                   | 99                   | 93                   | 93                   | 97                   | 95                   | 97                   | 97                   | 98                   | 97(98)               |
| PA98-2-TM13                        | 1998           | PE                  | 95                  | 92                   | 85                   | 98                   | 98                   | 98                   | 99                   | 93                   | 93                   | 96                   | 95                   | 97                   | 96                   | 97                   | 98                   |
| PA10-1                             | 2010           | PE                  | 92                  | 84                   | 65                   | 98                   | 98                   | 98                   | 99(100)             | 88                   | 87(92)               | 98                   | 95                   | 96                   | 97                   | 98                   | 78(84)               |
| PA10-SS                            | 2010           | PE                  | 92                  | 84                   | 66                   | 97                   | 98                   | 99(99)              | 88                   | 86(91)               | 98                   | 95                   | 98                   | 98(98)              | 86                   | 78(84)               |
| PA10-SS                            | 2010           | PE                  | 92                  | 84                   | 66                   | 97                   | 98                   | 99(99)              | 88                   | 86(91)               | 98                   | 95                   | 98                   | 98(98)              | 86                   | 78(84)               |
| SbBMV [NOA]                        | nd             | AR                  | 79                  | 73                   | nd                   | 84                   | 96(91)              | 82                   | 76(82)               | 78                   | 67(74)               | 84                   | 81(86)              | 84                   | 73(81)               | nd                   | nd                   |
| ToChLPV [MM1]                      | nd             | MX                  | 78                  | 66                   | nd                   | 78                   | 83(92)              | 82                   | 82(88)               | 83                   | 73(81)               | 86                   | 86(88)              | 87                   | 78(86)               | nd                   | nd                   |
| ToMMV [Ube]                        | nd             | BR                  | 78                  | 70                   | nd                   | 84                   | 90(94)              | 79                   | 79(88)               | 75                   | 61(72)               | 78                   | 75(85)              | 84                   | 70(77)               | nd                   | nd                   |
| ToMMV [Pac58]                      | 2005           | BR                  | 74                  | 66                   | nd                   | 80                   | 88(94)              | 79                   | 80(87)               | 81                   | 73(80)               | 84                   | 79(86)              | 87                   | 72(89)               | nd                   | nd                   |
| TGMV [YY]                          | nd             | BR                  | 77                  | 64                   | nd                   | 85                   | 92(96)              | 79                   | 79(87)               | 74                   | 61(72)               | 78                   | 75(85)              | 84                   | 65(74)               | nd                   | nd                   |
| TYMLCV [Mer]                       | nd             | BR                  | 77                  | 68                   | nd                   | 84                   | 92(96)              | 78                   | 77(86)               | 80                   | 73(83)               | 84                   | 81(86)              | 82                   | 62(72)               | nd                   | nd                   |
| MerMV [PR80]                       | nd             | PR                  | 77                  | 76                   | nd                   | 84                   | 93(97)              | 77                   | 74(83)               | 80                   | 63(71)               | 83                   | 78(86)              | 87                   | 69(80)               | nd                   | nd                   |
| ThYCV [Pepper]                     | 2007           | CU                  | 77                  | 63                   | nd                   | 83                   | 92(97)              | 78                   | 75(84)               | 76                   | 65(73)               | 80                   | 75(81)              | 84                   | 72(79)               | nd                   | nd                   |
| CabLCJV [Cuc-3]                    | nd             | JM                  | 77                  | 72                   | nd                   | 84                   | 91(96)              | 78                   | 78(88)               | 76                   | 67(74)               | 78                   | 75(84)              | 82                   | 76(76)               | nd                   | nd                   |

\(^a\)Dashed line separates ToLDeV isolates from the next closest begomovirus components, which were DNA-A components of New World bipartite begomoviruses. Underlined identities show divergent sequences that differentiate ToLDeV genotypes. Abbreviations are as in Fauquet at al. (12)

\(^\text{AR}=\text{Argentina, BR}=\text{Brazil, CU}=\text{Cuba, EC}=\text{Ecuador, JM}=\text{Jamaica, MX}=\text{Mexico, PE}=\text{Peru and PR}=\text{Puerto Rico} \)

\(^\text{nd} = \text{not determined} \)
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Nucleotide diversity (π)^a</th>
<th>Bayesian estimation of positively selected sites in coding regions^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intergenic region (IR)</td>
<td>0.07472</td>
<td>ωN = 1.000 (1.000), ω+ = 2.028 (1.750), ω- = 0.285 (0.237)</td>
</tr>
<tr>
<td>C4</td>
<td>0.09084</td>
<td>ωN = 1.000 (1.000), ω+ = 1.546 (1.343), ω- = 0.038 (0.033)</td>
</tr>
<tr>
<td>C1 (Rep)</td>
<td>0.07572</td>
<td>ωN = 1.000 (1.000), ω+ = 1.546 (1.343), ω- = 0.038 (0.033)</td>
</tr>
<tr>
<td>Overlapping C1 (Rep)/C4 region</td>
<td>0.08961</td>
<td>ωN = 1.000 (1.000), ω+ = 2.858 (1.934), ω- = 0.118 (0.111)</td>
</tr>
<tr>
<td>V1 (CP)</td>
<td>0.01577</td>
<td>ωN = 1.000 (1.000), ω+ = 5.817 (5.301), ω- = 0.022 (0.019)</td>
</tr>
</tbody>
</table>

^a The IR database was derived from 29 sequences, 5 from the GenBank, and 24 from this study, including those of the 9 full-length cloned components from Peru and Ecuador and 15 DNA fragments amplified from samples collected from Peru in 1998 and 2010 by PCR with the 1730v/496c primer pair. The databases for the C4, C1 (replication-associated protein, Rep), and V1 (capsid protein, CP) genes, as well as the overlapping region of the C4/C1 genes were derived from 11 sequences: two from the GenBank and those of the 9 full-length cloned components from this study.

^b Tajima’s test was implemented by DnaSP 5.10

^c Omega (ω) values (ratio of nonsynonymous [dN] and synonymous [dS] substitutions) for the negatively selected (dN/dS<1), neutral (dN/dS=1) and positively (dN/dS>1) selected sites ω−, ωN, and ω+, respectively. The proportions of the site categories (piN, pi+, pi-) associated with each omega value are presented. Median values are shown in parenthesis.
### TABLE 3

Infectivity, partial host range and symptomatology of three genotypes of Tomato leaf deformation virus (ToLDeV) from Peru.<sup>a</sup>

<table>
<thead>
<tr>
<th>Host plant</th>
<th>ToLDeV genotype&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Infected/inoculated (%) infected&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Symptoms&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solanum lycopersicum cv. Glamour</strong></td>
<td>PA98-1-1</td>
<td>49/50 (98)</td>
<td>dg, s, ulc, vp, vs</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>47/50 (94)</td>
<td>dg, s, ulc, vp, vs</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>46/50 (92)</td>
<td>dg, ss, ulc, vp, vs</td>
</tr>
<tr>
<td><strong>Nicotiana benthamiana</strong></td>
<td>PA98-1-1</td>
<td>28/28 (100)</td>
<td>s, dlc, vs</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>28/28 (100)</td>
<td>s, dlc, vs</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>28/28 (100)</td>
<td>s, svs, ulc</td>
</tr>
<tr>
<td><strong>N. tabacum cv. Samsun</strong></td>
<td>PA98-1-1</td>
<td>15/18 (83)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>10/18 (56)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>11/19 (58)</td>
<td>si</td>
</tr>
<tr>
<td><strong>N. glutinosa</strong></td>
<td>PA98-1-1</td>
<td>16/16 (100)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>14/14 (100)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>13/14 (93)</td>
<td>si</td>
</tr>
<tr>
<td><strong>Datura stramonium</strong></td>
<td>PA98-1-1</td>
<td>18/19 (95)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>15/18 (83)</td>
<td>si</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>10/19 (53)</td>
<td>si</td>
</tr>
<tr>
<td><strong>Chenopodium quinoa</strong></td>
<td>PA98-1-1</td>
<td>0/13 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/12 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/16 (0)</td>
<td>ni</td>
</tr>
<tr>
<td><strong>C. amaranticolor</strong></td>
<td>PA98-1-1</td>
<td>0/13 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/13 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/19 (0)</td>
<td>ni</td>
</tr>
<tr>
<td><strong>Cucurbita pepo cv. Small Sugar</strong></td>
<td>PA98-1-1</td>
<td>0/16 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/16 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/19 (0)</td>
<td>ni</td>
</tr>
<tr>
<td><strong>Phaseolus vulgaris cv. Topcrop</strong></td>
<td>PA98-1-1</td>
<td>0/12 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/12 (0)</td>
<td>ni</td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/17 (0)</td>
<td>ni</td>
</tr>
</tbody>
</table>

<sup>a</sup>Viruses were delivered by agroinoculation. Infectivity was determined 18 to 21 days post agroinoculation based on symptom development (tomato and *N. benthamiana*) and detection of viral DNA by PCR with the PAR1c496/ToLC-v1730 primer pair.

<sup>b</sup>Genotypes include the PA98-1 and PA98-2 variants, represented by the PA98-1-1 and PA98-2-1 isolates, respectively, and the PA10 strain, represented by the PA10-3 isolate.

<sup>c</sup>Results represent data of at least three independent experiments.

<sup>d</sup>Symptom abbreviations: dg, distorted growth; dlc, downward leaf curling; ni, no infection (no symptoms and no viral DNA detected in newly emerged leaves); si, symptomless infection; s, stunting; ss, severe stunting; ulc, upward leaf curling; vp, vein-purpling; vs, vein swelling and svs, severe vein swelling.
TABLE 4 Infectivity and symptomatology of mutants and the wild-type of three genotypes of Tomato leaf deformation virus (ToLDeV) from Peru.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>ToLDeV genotype</th>
<th>Symptomatic plants/total inoculated</th>
<th>Symptoms</th>
<th>Infected plants/total inoculated</th>
<th>Symptomatic plants/total inoculated</th>
<th>Symptoms</th>
<th>Infected plants/total inoculated</th>
<th>Grafting</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1</td>
<td>PA98-1-1</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>0/15 ni</td>
<td>0/15 ni</td>
<td>0/15</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>0/17 ni</td>
<td>0/17 ni</td>
<td>0/17</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/16 ni</td>
<td>0/16</td>
<td>0/13 ni</td>
<td>0/13 ni</td>
<td>0/13</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td>c1</td>
<td>PA98-1-1</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>0/15 ni</td>
<td>0/15 ni</td>
<td>0/15</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>0/14 ni</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/16 ni</td>
<td>0/16</td>
<td>0/14 ni</td>
<td>0/14 ni</td>
<td>0/14</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td>c4</td>
<td>PA98-1-1</td>
<td>0/23 si</td>
<td>23/23</td>
<td>20/45 ni or mlc and ms</td>
<td>30/45 ni or mlc and ms</td>
<td>30/45</td>
<td>0/14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>0/22 si</td>
<td>20/22</td>
<td>10/49 ni or mlc and ms</td>
<td>12/49 ni or mlc and ms</td>
<td>12/49</td>
<td>8/8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>0/17 si</td>
<td>17/17</td>
<td>15/45 ni or mlc and ms</td>
<td>15/45 ni or mlc and ms</td>
<td>15/45</td>
<td>0/8</td>
<td></td>
</tr>
<tr>
<td>c5</td>
<td>PA98-1-1</td>
<td>14/14 dlc, s, vs</td>
<td>14/14</td>
<td>9/9 dg, s, ulc, vp, vs</td>
<td>9/9 dg, s, ulc, vp, vs</td>
<td>9/9</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>15/15 dlc, s, vs</td>
<td>15/15</td>
<td>9/9 dg, s, ulc, vp, vs</td>
<td>9/9 dg, s, ulc, vp, vs</td>
<td>9/9</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>19/19 mnlc, s, vs</td>
<td>19/19</td>
<td>21/21 dg, ss, ulc, vp, vs</td>
<td>21/21 dg, ss, ulc, vp, vs</td>
<td>21/21</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>PA98-1-1</td>
<td>14/14 dlc, s, vs</td>
<td>14/14</td>
<td>14/15 dg, s, ulc, vp, vs</td>
<td>14/15 dg, s, ulc, vp, vs</td>
<td>14/15</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA98-2-1</td>
<td>14/14 dlc, s, vs</td>
<td>14/14</td>
<td>13/15 dg, s, ulc, vp, vs</td>
<td>13/15 dg, s, ulc, vp, vs</td>
<td>13/15</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA10-3</td>
<td>15/15 s, svs, ulc</td>
<td>15/15</td>
<td>14/15 dg, ss, ulc, vp, vs</td>
<td>14/15 dg, ss, ulc, vp, vs</td>
<td>14/15</td>
<td>nt</td>
<td></td>
</tr>
</tbody>
</table>

*Viruses were delivered by agroinoculation. Infectivity was determined 14 to 21 days post agroinoculation based on symptom development and detection of viral DNA by PCR with the PAR1c496/ToLC-v1730 primer pair. Results represent data of at least three independent experiments.

*Genotypes include the PA98-1 and PA98-2 variants, represented by the PA98-1-1 and PA98-2-1 isolates, respectively, and the PA10 strain, represented by the PA10-3 isolate.

*Symptom abbreviations: dg, distorted growth; dlc, downward leaf curling; mls, mild leaf curling; ms, mild stunting; mnlc, mild upward leaf curling; ni, no infection (no symptoms and no viral DNA detected in newly emerged leaves); si, symptomless infection; s, stunting; ss, severe stunting; ulc, upward leaf curling; vp, vein-purpling; vs, vein swelling and svs, severe vein swelling.

*Number of tomato plants with grafted scions in which symptoms reverted to wild-type; nt, not tested.