A recombinant begomovirus resulting from exchange of the C4 gene

Yan Xie,1 Liling Zhao,1 Xiaoyang Jiao,1 Tong Jiang,1 Huanran Gong,1 Bi Wang,1 Rob W. Briddon2 and Xueping Zhou1

1State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, People’s Republic of China
2National Institute for Biotechnology and Genetic Engineering, Jhang Road, P.O. Box 577, Faisalabad, Pakistan

A begomovirus isolated from Malvastrum coromandelianum and tomato originating from Yunnan province (China) was shown to be representative of a new begomovirus species, for which the name tomato leaf curl Yunnan virus (TLCYnV) is proposed. TLCYnV has high levels of sequence identity to tomato yellow leaf curl China virus (TYLCCNV) across the whole genome, except for sequences encompassing the C4 gene. Agrobacterium-mediated inoculation showed TLCYnV to be highly infectious to a range of plant species but poorly infectious to M. coromandelianum. In contrast to TYLCCNV, TLCYnV was shown to infect tomato in the absence of a betasatellite. In field-collected samples, TLCYnV was identified most frequently in tomato in which it was not associated with a betasatellite. Transgenic expression in Nicotiana benthamiana showed that the C4 protein of TYLCCNV did not induce developmental abnormalities, whereas the C4 of TLCYnV induced severe developmental abnormalities, reminiscent of virus symptoms. The genome of TLCYnV was shown to be significantly less methylated in plants than that of TYLCCNV and the C4 protein of TLCYnV was shown to suppress post-transcriptional gene silencing and transcriptional gene silencing more effectively than the C4 of TYLCCNV. The results indicate that TLCYnV evolved from TYLCCNV by recombination, acquiring a more virulent C4, allowing it to dispense with the requirement for a betasatellite. The implications of these findings in relation to the evolution of monopartite begomoviruses are discussed.

INTRODUCTION

Viruses of the family Geminiviridae are small single-stranded DNA (ssDNA) phytopathogens encapsidated within geminate particles. They are divided into four genera based on their genome organization, host range and insect vector, with the genus Begomovirus being the largest (Fauquet et al., 2003; Rojas et al., 2005; Brown et al., 2011). Begomoviruses cause numerous diseases of economically important dicotyledonous crops across the world (Fauquet et al., 2008). The genomes of some begomoviruses consist of two DNA molecules, referred to as DNA A and DNA B, each of about 2.8 kb. However, across the Old World, most have genomes that consist of a single component (Brown et al., 2011).

Six genes are encoded by the genomes of monopartite begomoviruses and the DNA A components of bipartite begomoviruses originating from the Old World (Brown et al., 2011). The virion-sense strand encodes the coat protein (CP; also known as the V1 protein) involved in virus encapsidation and virus transmission by the whitefly vector, and the V2/AV2 protein involved in virus movement in plants and overcoming host defences triggered by double-stranded RNA [termed RNA interference (RNAi)] (Wartig et al., 1997; Rojas et al., 2001; Zrachya et al., 2007; Amin et al., 2011b). The complementary-sense strand encodes the replication-associated protein (Rep; C1/AC1 protein; Hanley-Bowdoin et al., 2004), the transcriptional activator protein (TrAP; C2/AC2 protein) that may be involved in overcoming RNAi symptom induction and upregulating late virion-sense genes (Wartig et al., 1997; Vanitharani et al., 2004; Yang et al., 2007; Amin et al., 2011b), the replication enhancer protein (REN; C3/AC3 protein) that provides a cellular environment suitable for virus replication (Settlage et al., 2005), and the C4/AC4 protein that may be involved in overcoming RNAi symptom induction (Rigden et al., 1994; Vanitharani et al., 2004). The DNA B component of...
bipartite begomoviruses encodes two proteins, the nuclear shuttle protein (BV1 protein) and the movement protein (BC1 protein), which are required for virus spread within and between cells (Rojas et al., 2005).

Many monopartite begomoviruses, such as tomato yellow leaf curl China virus (TYLCCNV), are associated with betasatellites. Betasatellites are circular ssDNA molecules about half the size (~1.4 kb) of the helper virus genome. They depend on the helper virus for replication, encapsidation, insect transmission and movement in plants (Saunders et al., 2000, 2004; Briddon et al., 2001; Jose & Usha, 2003; Zhou et al., 2003; Cui et al., 2004; Saeed et al., 2005). For some viruses, such as TYLCCNV, the betasatellite is essential for induction of typical disease symptoms in the plants from which they were isolated (Cui et al., 2004).

RNA interference (RNAi) is an evolutionarily conserved surveillance system that occurs in all euukaryotic organisms, including animals, fungi and plants. Geminivirus transcripts are subject to degradation by post-transcriptional gene silencing (PTGS). PTGS is a sequence-specific RNA degradation process that leads to the downregulation of targeted mRNAs, and consequently reduced expression of proteins (Bisaro, 2006; Raja et al., 2010, 2008). The TrAP, V2/AV2 and C4/AC4 proteins of a number of begomoviruses, as well as the βC1 proteins of betasatellites, have been shown to suppress PTGS (Stanley & Latham, 1992; Vanitharani et al., 2004; Zrchya et al., 2007; Glick et al., 2008; Amin et al., 2011a). In addition, geminiviruses replicate in the nucleus through double-stranded DNA intermediates that serve as replication and transcription templates and associate with cellular histones to form minichromosomes (Pilartz & Jeske, 1992, 2003). Minichromosomes are potential targets for epigenetic repression and plants employ RNA-directed methylation, leading to transcriptional gene silencing (TGS), as a defence mechanism against geminiviruses (Law & Jacobsen, 2010; Raja et al., 2008; Rodriguez-Negrete et al., 2009). As a counter-defensive measure, the Rep of begomoviruses, the TrAP of cabbage leaf curl virus (CaLCuV) and tomato golden mosaic virus, and the related C2 proteins of curtoviruses, inhibit methylation and suppress TGS (Buchmann et al., 2009; Zhang et al., 2011; Rodriguez-Negrete et al., 2013).

TYLCCNV is a monopartite begomovirus associated with the Tomato yellow leaf curl China betasatellite (TYLCCNB), which is required by the virus to induce typical disease symptoms in host plants (Zhou et al., 2003). In common with all betasatellites, TYLCCNB encodes a single protein known as βC1 that is a pathogenicity factor (Cui et al., 2004), a suppressor of both PTGS (Cui et al., 2005) and TGS (Yang et al., 2011b), and interferes with host auxin and jasmonate signalling pathways (Yang et al., 2008). The study described here has characterized a distinct, naturally occurring recombinant monopartite begomovirus originating from Yunnan, China. This virus shows the highest levels of nucleotide sequence identity with TYLCCNV, with most of the sequence variation between the two viruses occurring in the C4 gene. The C4 protein of the recombinant virus is shown to be multifunctional, acting as a pathogenicity factor and a suppressor of both PTGS and TGS. These findings highlight the functional equivalence between the C4 of the recombinant virus and βC1.

RESULTS

Genomic organization of viral DNA and affinities to other begomoviruses

Malvastrum coromandelianum samples Y193 and Y194 with yellow vein symptoms were found to be infected with malvastrum yellow vein virus (MYVV) (Guo et al., 2008). Partial sequencing of virus clones showed the two samples also to be infected with a TYLCCNV-like virus (data not shown), so the complete TYLCCNV-like DNA sequences from Y193 and Y194 (isolates Y193 and Y194) were determined. The sequences of Y193 and Y194 consist of 2742 and 2749 nt, respectively. The coding strategy for Y193 and Y194 is typical of begomoviruses, with two genes in the virion sense and four in the complementary sense, separated by an intergenic region (IR). Efforts to detect a possible DNA B component in M. coromandelianum plants, using specific primers PCR1 and PBL1v2040 (Rojas et al., 1993), were unsuccessful.

The full-length sequences of Y193 and Y194 share 98.7 % nucleotide sequence identity and 98.8–100 % amino acid sequence identities for the six gene products (Table S1, available in JGV Online), demonstrating that they represent isolates of a single begomovirus species; Y194 was chosen for further analysis. The sequence of Y194 shared the highest levels of nucleotide sequence identity (87.7 %) with TYLCCNV, and less than 79.7 % identity with other begomoviruses. However, when the sequence spanning the C4 gene was omitted from the analysis, the nucleotide sequence of Y194 showed 97.2 % identity with TYLCCNV. Closer inspection of alignments showed that the Y194 DNA sequence can be divided into two parts (see Fig. S1). Part 1 includes the V2, CP (V1), REn (C3) and TrAP (C2) genes as well as the 3′ end of Rep (C1) genes. The predicted amino acid sequences of the V2, CP, REn and TrAP show the highest amino acid sequence identity with TYLCCNV (97.4 %, 98.8 %, 97.8 % and 98.5 %, respectively). Part 2 includes the 5′ end of Rep, the C4 and IR, which show less sequence identity with TYLCCNV (77.6 % and 32.9 % amino acid sequence identities for Rep and C4 proteins, respectively, and 77.3 % nucleotide sequence identity for IR; Table S1, Fig. S1). The AC4/C4 gene of begomoviruses is contained entirely within the coding sequence of Rep, but in a different reading frame. This likely explains the low amino acid sequence identity (77.6 %) between the Rep proteins of Y194 and TYLCCNV. The C4 protein of Y194 is almost identical (84 out of 85
amino acids) to that of pepper yellow leaf curl China virus (PepYLCCNV; accession number KC149938) but shows relatively lower levels of sequence identity to the C4 proteins of other begomoviruses with the highest (83.5 %) being to the C4 of croton yellow vein mosaic virus (CYVMV). These findings indicate that Y194 likely originated by recombination between TYLCCNV and PepYLCCNV.

**Infectivity and symptoms induced by Y194**

Agrobacterium-mediated inoculation showed Y194 to be highly infectious to a range of plant species (Table 1). In *Nicotiana benthamiana*, the first leaf curling symptoms appeared at 6 days post-inoculation (days p.i.), and more severe upward leaf curling and vein thickening symptoms were observed several days later (Fig. 1a). In tomato (*Solanum lycopersicum*), the first symptoms were observed at 10 days p.i., and the full symptoms consisting of downward leaf curling and vein thickening developed within 20 days p.i. (Fig. 1b). For infections of *N. glutinosa*, no symptoms were evident on the upper leaf surface (Fig. 1c) whereas on the lower surface veins showed severe swelling. Y194 infection induced downward leaf curling symptoms in *N. tabacum* Samsum (Fig. 1d), and stunting, upward leaf curling and vein thickening on the underside of leaves in *Petunia hybrida* (Fig. 1e). The presence of viral DNA in *N. benthamiana* and tomato was confirmed by Southern blot (Fig. 2).

Although Y194 was isolated from *M. coromandelianum*, agroinoculation of this species with Y194 did not result in infection (Table 1). Previous studies have shown that MYVV is not infectious to *M. coromandelianum* by agroinoculation, although plants could be infected by whitefly transmission (Guo et al., 2008). Whitefly-mediated transmission of Y194 from infected tomato plants to *M. coromandelianum* plants resulted in five plants (out of 15 inoculated) in which Y194 could be detected by PCR diagnostics using the specific primers Y194F and Y194R (data not shown). The five plants showed no discernible symptoms indicating asymptomatic infection.

TYLCCNV requires TYLCCNB to induce symptoms in *N. benthamiana* and tomato plants (Fig. 1f–i, Table 1; Cui et al., 2004). When Y194 was co-inoculated with TYLCCNB the symptoms in all inoculated species were enhanced in comparison to inoculation with only the virus (Fig. 1j–l). Southern blot analysis showed that the accumulation levels of Y194 DNA were enhanced in infected *N. benthamiana* and tomato plants in the presence of TYLCCNB (Fig. 2).

These findings indicate that Y194 and TYLCCNV have distinct pathogenicities in *N. benthamiana* and tomato; in contrast to TYLCCNV, Y194 has no requirement for a betasatellite to infect *N. benthamiana* and tomato symptomatically. This supports the contention that Y194 can be considered as a distinct species of begomovirus.

**Phenotypes induced in *N. benthamiana* resulting from constitutive and transient expression of C4 genes**

Transgenic *N. benthamiana* lines expressing Y194 or the TYLCCNV C4 gene under the control of the cauliflower

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>Plant species</th>
<th>No. of plants infected/no. of plants inoculated</th>
<th>Latent period* (days)</th>
<th>Symptoms†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y194</td>
<td><em>N. benthamiana</em></td>
<td>66/66</td>
<td>6</td>
<td>SULC, VT</td>
</tr>
<tr>
<td></td>
<td><em>S. lycopersicum</em></td>
<td>59/60</td>
<td>10</td>
<td>DLC, VT</td>
</tr>
<tr>
<td></td>
<td><em>N. glutinosa</em></td>
<td>31/32</td>
<td>8</td>
<td>VT</td>
</tr>
<tr>
<td></td>
<td><em>N. tabacum</em> Samsum</td>
<td>30/32</td>
<td>9</td>
<td>DLC</td>
</tr>
<tr>
<td></td>
<td><em>P. hybrida</em></td>
<td>32/32</td>
<td>7</td>
<td>ULC, VT, ST</td>
</tr>
<tr>
<td></td>
<td><em>M. coromandelianum</em></td>
<td>0/15</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Y194 + TYLCCNB</td>
<td><em>N. benthamiana</em></td>
<td>32/32</td>
<td>6</td>
<td>SULC, VT</td>
</tr>
<tr>
<td></td>
<td><em>N. glutinosa</em></td>
<td>32/32</td>
<td>7</td>
<td>DLC, YV, VT</td>
</tr>
<tr>
<td></td>
<td><em>S. lycopersicum</em></td>
<td>32/32</td>
<td>9</td>
<td>DLC, VT</td>
</tr>
<tr>
<td></td>
<td><em>M. coromandelianum</em></td>
<td>0/30</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TYLCCNV</td>
<td><em>N. benthamiana</em></td>
<td>16/20</td>
<td>–</td>
<td>Symptomless</td>
</tr>
<tr>
<td></td>
<td><em>S. lycopersicum</em></td>
<td>12/20</td>
<td>–</td>
<td>Symptomless</td>
</tr>
<tr>
<td>TYLCCNV + TYLCCNB</td>
<td><em>N. benthamiana</em></td>
<td>20/20</td>
<td>6</td>
<td>SDLC, VT, ST, YV, EN</td>
</tr>
<tr>
<td></td>
<td><em>S. lycopersicum</em></td>
<td>17/20</td>
<td>12</td>
<td>SDLC, ST, EN</td>
</tr>
<tr>
<td></td>
<td><em>M. coromandelianum</em></td>
<td>0/20</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*Time between inoculation and first appearance of symptoms.
† Symptoms are denoted as downward leaf curling (DLC), enation (EN), severe downward leaf curling (SDLC), stunting (ST), severe upward leaf curling (SULC), upward leaf curling (ULC), vein thickening (VT), yellow vein (YV).
**Fig. 1.** Symptoms induced by virus infection. *N. benthamiana* (a, f, h, j), *S. lycopersicum* (b, g, i, l), *N. glutinosa* (c, k), *N. tabacum* Samsum (d) and *P. hybrida* (e) plants were inoculated with Y194 (a–e), TYLCCNV (f, g), TYLCCNV plus TYLCCNB (h, i) or Y194 plus TYLCCNB (j–l). Photographs were taken at 30 days p.i.

**Fig. 2.** Southern blot analysis of viral DNA accumulation in plants infected with viruses by agroinoculation. Total nucleic acids (5 μg) were extracted at 30 days p.i. from individual *N. benthamiana* or *S. lycopersicum* plants agroinoculated with empty vector pBinPLUS (Mock, as a negative control), Y194, Y194 plus TYLCCNB, TYLCCNV or TYLCCNV plus TYLCCNB. The blot was hybridized with a Y194 CP probe. The positions of viral single-stranded (ss) DNA and subgenomic (sg) DNA are indicated. The total genomic DNA visualized by ethidium bromide staining is shown to provide a loading control.
mosaic virus (CaMV) 35S promoter were produced. Semiquantitative reverse transcriptase-PCR (sqRT-PCR) showed that C4 expression levels among the transgenic lines were similar as judged by sqRT-PCR amplification of the 3’UTR sequence (which is identical in the two C4 transcripts and the transcript produced by the empty vector control). The detection of similar levels of the endogenous housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH) in the cDNA samples from the three transgenic N. benthamiana lines indicated that approximately equal amounts of total mRNA were used (Fig. 3). Plants constitutively expressing Y194 C4 developed virus-like symptoms that consisted of severe leaf distortion, downward leaf curling and blistering (Fig. 4a). By contrast, transgenic N. benthamiana plants expressing TYLCCNV C4 developed normally and were indistinguishable from plants transformed with the vector (pBin438) lacking an insert. Although these results suggest that the two C4 proteins have differing effects on plant development when constitutively expressed, since we have not examined protein expression levels (by, for example, Western blotting), we cannot rule out the possibility that the differences are due to lower expression of the TYLCCNV C4 protein. The virus-like symptoms induced in transgenic N. benthamiana plants constitutively expressing the TYLCCNB βC1 gene (produced previously; Cui et al., 2004) consisted of upward leaf curling and some foliar distortion (Fig. 4a).

Infection of N. benthamiana plants with a PVX vector (pGR106), harbouring no insert, induced mild chlorotic spots on leaves but otherwise had no significant effect on plant development. By contrast, infection of plants with a PVX vector harbouring the C4 gene of TYLCCNV produced more severe chlorotic spots and foliar crumpling (distortion), although leaves expanded to near-normal size. PVX-mediated expression of the Y194 C4 induced severe upward leaf curling, vein thickening, and an etiolated appearance (longer internodes and petioles). These symptoms were similar to those induced in N. benthamiana plants infected with a PVX vector expressing the TYLCCNB βC1, which induced severe leaf curling, vein swelling and stunting but lacked the etiolation (Fig. 4b).

These findings show that the C4 proteins of both viruses are symptom (pathogenicity) determinants, in common with βC1, but also that the C4 proteins are distinct with respect to their effects on plant development when expressed from either PVX or constitutively in transgenic plants.

**Suppression of PTGS by C4 proteins**

The relative abilities of the C4 proteins of Y194 and TYLCCNV to suppress PTGS were assessed in N. benthamiana line 16c carrying a green-fluorescent protein (GFP) transgene under the control of the CaMV 35S promoter (Cui et al., 2005). For leaves of 16c plants co-infiltrated with Agrobacterium tumefaciens cultures harbouring the expression vector pBin438, with no insert, and a construct expressing a fragment of the GFP gene (35S-GFP), the initial green fluorescence gradually declined between 3 and 5 days p.i., indicative of PTGS of the GFP gene (Fig. 5a). However, for leaves co-infiltrated with constructs expressing either the tobacco etch virus (TEV) helper component protease (HC-Pro) or one of the C4 genes (Y194 C4 or TYLCCNV C4), together with 35S-GFP, green fluorescence remained in inoculated tissues at 5 days p.i., indicating that all three gene products suppress PTGS. The strongest effect was seen for HC-Pro and the weakest for TYLCCNV C4, with a very faint green fluorescence remaining in inoculated tissues at 5 days p.i. Northern blot analysis of total RNA extracted from inoculated tissues, hybridized for GFP mRNA, showed a correlation between GFP fluorescence and the levels of GFP mRNA detected (Fig. 5c). Although the levels of GFP mRNA decreased from 3 to 5 days p.i., the highest levels were detected with HC-Pro and the lowest for TYLCCNV C4. These results indicate that the Y194 C4 protein is a stronger suppressor of PTGS than the C4 protein of TYLCCNV.
Reversal of TGS by C4 proteins

The ability of the C4 proteins to reverse TGS was assessed using *N. benthamiana* line 16c-TGS (Buchmann et al., 2009), which contains a transcriptionally silenced GFP transgene driven by the 35S promoter. This assay has previously been used to demonstrate that the C2 protein encoded by beet curly top virus and the \( \beta \)C1 protein of TYLCCNB can suppress TGS (Buchmann et al., 2009; Yang et al., 2011b). The plasmids pGR106-Y194C4, pGR106-Y194mC4 and pGR106-TYLCCNVC4 were constructed to express Y194 C4, Y194 C4 mutant or TYLCCNV C4 using the PVX vector pGR106. 16c-TGS plants were inoculated with pGR106 (negative control), pGR106-\( \beta \)C1 (positive control), pGR106-Y194C4, pGR106-TYLCCNVC4 or pGR106-Y194mC4. Under UV light, plants in which GFP is not expressed appear red due to chlorophyll auto-fluorescence, whereas regions expressing GFP appear pale green. PVX infection had no impact on GFP transcriptional silencing. At 11 days p.i., plants infected by pGR106-\( \beta \)C1 or pGR106-Y194C4 exhibited a bright yellow-green fluorescence in shoots and newly emerging leaves. By contrast, the GFP fluorescence in plants inoculated with pGR106-TYLCCNVC4 was only visible in leaf petioles and upper leaves and at a much lower level than that elicited by pGR106-Y194C4. No green fluorescence was observed in plants inoculated with pGR106-Y194mC4 (Fig. 5b). ELISA showed that the PVX concentrations were similar in plants infiltrated with pGR106-Y194C4 or pGR106-TYLCCNVC4 (data not shown), indicating that weaker suppression of silencing mediated by TYLCCNV C4 is not due to its low expression by PVX. Northern blot analysis confirmed that TGS reversal was accompanied by the accumulation of GFP mRNA (Fig. 5d). The results indicate that both C4 proteins can reverse TGS, but that TGS reversal by Y194 C4 is stronger than that of TYLCCNV C4.

Inhibition of methylation by C4 proteins

To investigate whether Y194 C4 impacts methylation of the virus genome, bisulfite sequencing of a 354 nt fragment of TYLCCNV (including 13 CG, 11 CHG and 46 CHH) and a 366 nt fragment of Y194 (having 18 CG, 9 CHG and 53 CHH), each fragment spanning the N terminus of the Rep gene, the IR and the N terminus of the V2 gene, was performed to determine methylation status in plants. The result showed that the TYLCCNV genome is subject to DNA methylation in *N. benthamiana*, and that the presence of TYLCCNB reduced DNA methylation in all contexts, as shown previously by Yang et al. (2011b). The overall methylation of TYLCCNV in the presence of TYLCCNB was 7.9 % of cytosines, lower than the 10.3 % for TYLCCNV in the absence of TYLCCNB (Fig. 6). The Y194 genome was similarly found to be subject to DNA methylation but at a much lower level (1.3 %) than TYLCCNV. This suggests that Y194 C4 plays an important role in inhibiting DNA methylation.

Detection of the recombinant virus in field-collected samples

Naturally infected *M. coromandelianum* leaf samples showing yellow vein symptoms and tomato leaf samples showing leaf curling and vein thickening symptoms were tested for the presence of the recombinant virus. Fragments of about 2.8 kb were amplified with the specific primers...
Y194F and Y194R in 2 of 16 *M. coromandelianum* samples collected in 2004, 3 of 13 *M. coromandelianum* samples collected in 2011, and 29 of 31 tomato samples collected in 2011, indicating that Y194-like viruses are common in the field, particularly in tomato plants. Comparisons of partial sequences (~500 nt, which included part of the IR, V2 and CP genes) showed that these amplified fragments shared 97.8% to 100% nucleotide sequence identities (results not shown). The complete nucleotide sequences of virus isolates from tomato samples T1 and T5 were determined. T1 and T5 shared 99.7% nucleotide sequence identities with Y194, and the C4 of T1 and T5 shared 100% amino acid sequence identities with Y194 C4 (Table S1). No betasatellite was identified in infected tomato samples (data not shown), indicating that a betasatellite is not associated with Y194-like recombinant viruses in tomato.

![Image](https://example.com/image.png)

**Fig. 5.** Suppression of PTGS of GFP in leaves of *N. benthamiana* 16c (a), reversal of TGS in *N. benthamiana* line 16-TGS (b) and Northern blot assays probed for GFP mRNA (c, d). For (a), leaves were co-agroinfiltrated with 35S-GFP and one of the four different plasmids indicated above the images. Photographs were taken under UV light at 5 days p.i. For (b), plants were inoculated with pGR106 (PVX), pGR106-βC1 (PVX-βC1), pGR106-Y194C4 (PVX-Y194 C4), pGR106-TYLCCNV C4 (PVX-TYLCCNV C4) or pGR106-Y194mC4 (PVX-Y194 mC4). Photographs were taken under UV light at 11 days p.i. For (c) and (d), RNA samples run on gels were extracted from the leaves of *N. benthamiana* line 16c plants at 3 days p.i. and 5 days p.i. (c) or from leaves of 16-TGS plants at 11 days p.i. (d). For (c), samples were extracted from plants co-inoculated with an *Agrobacterium tumefaciens* culture harbouring 35S-GFP and cultures containing an empty expression vector (pBin438) or constructs for the expression of tobacco etch virus helper component protease (HC-Pro), TYLCCNV C4, or Y194 C4. For (d), samples were extracted from plants infected with the PVX vector or the PVX vector expressing TYLCCNB βC1, Y194 C4, TYLCCNV C4 or Y194 mC4. The 18S rRNA band on gels, visualized by ethidium bromide staining, is shown to confirm equal loading.
DISCUSSION

The results presented here indicate that Y193, Y194, T1 and T5 are representative of a newly identified species of begomovirus, based upon the presently applicable species demarcation threshold of 89% nucleotide sequence identity (Fauquet et al., 2003; Brown et al., 2011). In view of the virus being most commonly identified in tomato and being poorly infectious to M. coromandelianum, the name tomato leaf curl Yunnan virus (TLCYnV) is proposed for the new species.

Sequence comparisons show TLCYnV to be closely related to TYLCCNV with the majority of sequence differences between the two viruses being confined to sequences encompassing the C4 gene. The C4 of Y194 is almost identical (98.8% amino acid sequence identity; Table S1) to PepYLCCNV C4. This indicates that TLCYnV has a recombinant origin with TYLCCNV as the major parent and PepYLCCNV as the donor of the C4 gene and the IR. Recombination is an important mechanism of evolution for geminiviruses, driven by the frequent occurrence of multiple infections in the field (Martin et al., 2011). In China, numerous begomoviruses have been reported which likely have arisen by recombination (Xie & Zhou, 2003; Xiong et al., 2007; Yang et al., 2011a).

Despite the fact that the coding sequence of C4/AC4 lies entirely within the well-conserved Rep gene, the C4/AC4 sequences of all dicot-infecting begomo-, topocu- and curtoviruses are the least conserved of all the genes. The precise function of the AC4 of bipartite begomoviruses remains unclear since mutation has no effect on DNA replication, symptoms or host range (Etessami et al., 1991; Pooma & Petty, 1996). However, the fact that the AC4 gene of bipartite begomoviruses has an effect on plant development when expressed from a PVX vector (Fondong et al., 2007; Amin et al., 2011b), as well as its maintenance across all begomo- curto- and topocuviruses, suggests that it should be functional. The AC4 of African cassava mosaic virus (ACMV) is a suppressor of gene silencing that binds single-stranded microRNA and small interfering RNA, the effectors of gene silencing (Vanitharani et al., 2004; Chellappan et al., 2005). The function of the C4 may differ across monopartite begomoviruses. Disruption of the C4 results in attenuated symptoms and reduced infectivity, suggesting that it is involved in symptom development (Rigden et al., 1994; Iqbal et al., 2012). For tomato yellow leaf curl virus, the subcellular localization of C4–GFP fusion proteins at the cell periphery suggested a role in virus movement (Rojas et al., 2001). The C4 of curtoviruses is an important symptom determinant (Dogra et al., 2009). It induces cell division (Latham et al., 1997), interacts with the brassinosteroid signalling pathway (Piroux et al., 2007) and upregulates host RING finger protein to interfere with the cell cycle (Lai et al., 2009). The results presented here show, for the first time, that a monopartite begomovirus C4 suppresses TGS by countering viral genome methylation. Previously only the Rep of begomoviruses, the TrAP of a bipartite begomovirus, the C2 of curtoviruses and the bC1 of TYLCCNB have been shown to suppress TGS activity (Buchmann et al., 2009; Yang et al., 2011b; Zhang et al., 2011; Rodriguez-Negrete et al., 2013).

The differences between the sequences of TLCYnV and TYLCCNV strongly implicate C4 as the mediator of the
biological differences between the two viruses, although it is not possible, based on the studies presented here, to rule out the possibility that other sequences also play a part. The pathogenicities of the two viruses are different. For TYLCCNV, the betasatellite TYLCCNB extends the host range of the virus to include *N. benthamiana* and tomato (Cui et al., 2004). The ability of betasatellites to extend the host ranges of monopartite begomoviruses and even to allow the DNA A components of bipartite begomoviruses to infect plant species (in the absence of DNA B) that are not part of their usual host range is well documented and is attributed to the βC1 product of betasatellites (Saunders et al., 2002; Saeed et al., 2007; Saeed, 2010a, b). This indicates that TYLCCNV is lacking some essential function required to infect tomato and *N. benthamiana* which can be complemented by the βC1 of TYLCCNB.

TYLCCNV infections of *N. benthamiana* in the absence of TYLCCNB are associated with very low viral DNA levels consisting of mostly ssDNA (Cui et al., 2004). This suggests that although the virus can spread out of inoculated tissues into the phloem, it does not have the ability to infect cells at the top of the plant or infects a small number of phloem-associated cells but cannot spread from these cells. For both βC1 and the C4 proteins of monopartite begomoviruses evidence has previously been presented suggesting that they may be involved in virus movement (Jupin et al., 1994; Rojas et al., 2001; Saeed et al., 2007). The work of Rojas et al. (2001) suggested that monopartite begomovirus V2 and/or C4 proteins may be analogues of the bipartite begomovirus movement protein and facilitate the cell-to-cell movement of viral DNA. However, the fact that these proteins, and the βC1 of betasatellites, are also suppressors has led to the suggestion that, rather than acting as classical movement proteins, they instead counter an RNAi-based host resistance to virus movement (Amin et al., 2010, 2011a).

Monopartite begomoviruses are essentially of three types. The first are the true monopartite begomoviruses which do not require a betasatellite and, in the field, are not found associated with betasatellites. This group includes viruses such as tomato yellow leaf curl Sardinia virus and tomato leaf curl virus (Kheyr-Pour et al., 1991; Dry et al., 1993). The second group has a more facilitative relationship with betasatellites. These viruses do not require the betasatellite to infect plants but, in the field, are often (but not always) associated with betasatellites (Li et al., 2005). The third group consists of viruses which have an absolute requirement for the betasatellite to infect the plant species from which they were isolated, such as the begomoviruses causing cotton leaf curl disease and TYLCCNV (Briddon et al., 2001; Cui et al., 2004). The results presented here suggest that a betasatellite-dependent monopartite begomovirus can become independent of a betasatellite by the acquisition of a distinct C4 protein or, conversely, that a monopartite begomovirus with a ‘deficient’ C4 can be ‘rescued’ by the acquisition of a betasatellite. Moreover, the differences between the C4 of TYLCCNV and that of TLCYnV suggest that the ability to dispense with the betasatellite arises from an overlap in the activities of C4 and βC1, with the betasatellite (or more specifically the βC1 protein encoded by betasatellites) possibly being required to compensate for a C4 protein with weak TGS/PTGS activity. It is also interesting to note that the C4 has recently been shown to play no part in the maintenance of a betasatellite by a monopartite begomovirus (Iqbal et al., 2012).

**METHODS**

**Virus sources.** Thirty-one tomato samples showing leaf curling, leaf yellowing and vein thickening, and 29 *M. coronandelianum* samples showing vein yellowing were collected from Yuanmou (Yunnan Province, China) in 2004 and 2011, respectively.

**Total DNA extraction, PCR amplification, cloning and sequencing.** Total DNA was extracted from the leaves of plants using the method of Xie et al. (2002). Degenerate primer pair FA/PB was used to amplify part of the IR and V2 gene of viral DNA. PCR was carried out as described by Zhou et al. (1997). The PCR products were then sequenced. Based on the partial sequences determined, primers Y194F and Y194R were designed and used to amplify the complete viral genome. PCR products were cloned into pGEM-T Easy Vector (Promega) and sequenced with the automated model 377 DNA-sequencing system (PerkinElmer). The sequences of primers used for PCR and cloning are given in Table S2.

**Sequence analysis.** Sequences were assembled and analysed using Lasergene software version 6.0 (DNASTAR Inc.) and MEGA version 4 (Tamura et al., 2007). Sequence alignments were performed using the CLUSTAL V algorithm in Lasergene.

**Production of constructs for Agrobacterium-mediated inoculation.** The production of constructs for the inoculation of TYLCCNV (pBinPLUS-TYLCCNV-1.7A) and TYLCCNB (pBinPLUS-TYLCCNV-2β) has been described previously (Cui et al., 2004). For the production of a construct for the agroinoculation of Y194, the full-length genome was amplified using primers Y194EF/Y194ER (Table S2), which overlap at a unique EcoRI site. The amplified fragment was cloned into pGEM-T (Promega), to yield pGEM-T-Y194 and sequenced. Based on the partial sequences determined, primers Y194F and Y194R were designed and used to amplify the complete viral genome. PCR products were cloned into pGEM-T Easy Vector (Promega) and sequenced with the automated model 377 DNA-sequencing system (PerkinElmer). The sequences of primers used for PCR and cloning are given in Table S2.

**Sequence analysis.** Sequences were assembled and analysed using Lasergene software version 6.0 (DNASTAR Inc.) and MEGA version 4 (Tamura et al., 2007). Sequence alignments were performed using the CLUSTAL V algorithm in Lasergene.

**Production of constructs for Agrobacterium-mediated inoculation.** The production of constructs for the inoculation of TYLCCNV (pBinPLUS-TYLCCNV-1.7A) and TYLCCNB (pBinPLUS-TYLCCNV-2β) has been described previously (Cui et al., 2004). For the production of a construct for the agroinoculation of Y194, the full-length genome was amplified using primers Y194EF/Y194ER (Table S2), which overlap at a unique EcoRI site. The amplified fragment was cloned into pGEM-T (Promega), to yield pGEM-T-Y194 and sequenced. The BamHI-EcoRI fragment of pGEM-T-Y194, encompassing the IR of the virus, was introduced into the binary vector pBinPLUS (van Engelen et al., 1995) to yield pBinPLUS-Y194-0.4A. The full-length EcoRI insert of pGEM-T-Y194 was then inserted into the unique EcoRI site of pBinPLUS-Y194-0.4A to yield pBinPLUS-Y194-1.4A, containing a 1.4-mer tandem repeat of Y194. *Agrobacterium tumefaciens* strain EHA105 was transformed with pBinPLUS-Y194-1.4A by triparental mating.

**Agrobacterium-mediated inoculation.** The production of *A. tumefaciens* cultures for inoculation of plants has been described previously (Cui et al., 2004). *M. coronandelianum, N. benthamiana, N. glutinosa, N. tabacum* Samsun, *P. hybrida* and *S. lycopersicum* plants were inoculated using a syringe with a fine needle as described previously (Cui et al., 2004). Inoculated plants were grown in an insect-free cabinet at 25 °C with supplementary lighting corresponding to a 16 h day length.

**Whitefly-mediated transmission.** Whitefly-mediated virus transmission was carried out as described by Guo et al. (2008) using *S. lycopersicum* plants experimentally infected with Y194 as the source of virus inoculum. Four weeks after whitefly transmission, plants were evaluated for infection by symptom observation and by PCR using the specific primers Y194F and Y194R.
Production of expression constructs. A construct for the expression of the BCI gene of TYLCCNB, under the control of a duplicated CaMV 35S promoter has been described previously (Cui et al., 2004). The C4 genes of TYLCCNV and Y194 were amplified by PCR using specific primer pairs Y10AC4F/Y10AC4R and Y194AC4F/ Y194AC4R, respectively, cloned in pGEM-T and sequenced. Each of the C4 fragments was then cloned into the binary vector pBin438 at BamHI and Sall sites to yield pBin438-TYLCCNV-C4 and pBin438-Y194-C4, respectively. The pBin438 vector lacking an insert and pBin438 constructs were introduced into A. tumefaciens EHA105 by triparental mating.

The C4 genes of Y194 and TYLCCNV were amplified using primer pairs Y194AC4PVX-F/Y194AC4PVX-R and Y10AC4PVX-F/ Y10AC4PVX-R, respectively, and then cloned into the PVX vector pGR106 (Chapman et al., 1992) at Ncol and SalI sites to produce pGR106-Y194C4 and pGR106-TYLCCNV-C4, respectively. Mutation was introduced into the Y194 C4 gene by inserting a stop codon (TAA) after the second start codon of Y194 C4 by PCR with primer pair Y194A4PVX-mF/Y194A4PVX-R. After confirming mutation by sequencing, the same construction strategy was performed to produce pGR106-Y194mC4 harbouring the Y194 C4 mutant. The resulting plasmids and pGR106 lacking an insert were transformed into A. tumefaciens strain GV3101 by electroporation.

Production of transgenic plants. A. tumefaciens cultures harbouring pBin438-TYLCCNV-C4 and pBin438-Y194-C4 were used to transform leaf explants of N. benthamiana as described previously (Pascal et al., 1993). Transformants were selected on medium containing kanamycin (100 mg l\(^{-1}\)) and carbenicillin (250 mg l\(^{-1}\)). Antibiotic-resistant shoots were collected, placed on rooting medium, grown to a height of 5–6 cm, and transferred to soil. The presence of C4 genes in transgenic plants was verified by PCR using the primer pairs used to make the expression constructs. C4 transcript levels were determined by sqRT-PCR using the Superscript III RT-PCR system (Invitrogen) and the primers Y10AC4F/Y10AC4R and Y194AC4F/ Y194AC4R, for the TYLCCNV and Y194 C4 genes, respectively. Fragments of the 3’ UTR sequence in the pBin438 vector which is identical in constructs pBin438-TYLCCNV-C4 and pBin438-Y194C4, and the endogenous glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNAs were amplified using primer pairs 3’UTRF/ 3’UTRRI and GAPDHF/GAPDHR, respectively. PCR was carried out for 25 cycles. The sequences of primers are given in Table S2.

Serological tests. Quantification of PVX by ELISA was performed as described previously (Wu et al., 2007).

Suppression of RNA silencing. To assess the ability of C4 proteins to suppress PTGS, A. tumefaciens strain C58C1 cultures harbouring pBin438-TYLCCNV-C4 or pBin438-Y194-C4 were infiltrated into N. benthamiana line 16c plants as described previously (Xiong et al., 2009). A plasmid for the expression of the TEV HC-Pro gene under the control of the CaMV 35S promoter (p35S-HC-Pro) and the empty vector pBin438 were used as positive and negative controls, respectively.

The ability of the C4 proteins to suppress TGS was assessed by inoculating A. tumefaciens cultures harbouring pGR106-Y194C4, pGR106-TYLCCNV-C4, pGR106-Y194mC4 or pGR106 to N. benthamiana line 16-TGS plants as described previously (Yang et al., 2011b).

Nucleic acid extraction and hybridization. Genomic DNA was extracted from plants using a DNA Easy Plant mini kit (Qiagen). DNA (5 µg) was separated on 1.0% agarose gels and transferred to Hybond-N + membranes (GE Healthcare) and UV cross-linked. Membranes were hybridized to [\(^{32}\)P]dCTP-labelled Y194 CP, which is conserved between Y194 and TYLCCNV (98.8% amino acid sequence identity).

Total RNA was extracted from N. benthamiana leaves with Trizol reagent (Invitrogen) according to the manufacturer’s instructions. Total RNA (10 µg) was separated on 1.2% formaldehyde-agarose gels and transferred to Hybond-N + membranes and UV cross-linked. Northern hybridization was performed as described previously (Cui et al., 2005), and signals were detected using a Typhoon 9200 imager (Amersham Pharmacia).

Bisulfite sequencing. DNA (1 µg) extracted from infected, agroinoculated N. benthamiana plants was digested with EcoRI (which cuts outside the IR of the viruses) followed by overnight treatment with proteinase K. Bisulfite modification was performed in a PCR machine using the DNA Methylation Gold kit (Zymo Research). Bisulfite-modified DNA was purified using a Zymo-Spin IC column and dissolved in 10 µl of Elution Buffer (Zymo Research). PCR amplification was carried out with Zymo Taq and primers BiF and BiR (Table S2). PCR products were cloned into the pGEM-T Easy plasmid (Promega), and 25 individual clones for each treatment were sequenced using an automated model 3730 DNA Sequencer (PerkinElmer). To ascertain if bisulfite modification was complete, plasmid containing TYLCCNV DNA was added to mock-inoculated plant DNA extracts and used for bisulfite treatment.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Natural Science Foundation of China (31070128 and U1136606) and the National Key Basic Research and Development Program of China (2012CB114004). R.W.B. was supported by the Higher Education Commission (Govt. of Pakistan) under the ‘Foreign Faculty Hiring Scheme’. We thank Professor David BAulcombe for providing the PVX vector pGR106 and Dr J. C. Carrington for providing TEV HC-Pro.

REFERENCES


required for symptom induction.


