INTRODUCTION

Begomoviruses belong to the family Geminiridae, whose members are important phytopathogens that cause great damage in many crops in tropical and subtropical areas (Polston & Anderson, 1997). They possess a circular, single-stranded DNA (ssDNA) genome consisting of one or two molecules of 2.3–2.9 kb each (reviewed by Hanley-Bowdoin et al., 2000; Rojas et al., 2005). Geminiviruses replicate by rolling-circle and recombination-dependent replication mechanisms (Alberter et al., 2005; Preiss & Jeske, 2003), using a double-stranded DNA intermediate that functions as a template for replication and transcription during infection. In contrast to most plant viruses, geminivirus replication and transcription occur in the nucleus of the host cell (Bass et al., 2000; Saunders et al., 1991). The geminiviral genome encodes only a few proteins, whose open reading frames (ORFs) extend away from an intergenic region where different regulatory elements important for replication and transcription are located (Fig. 1) (Argüello-Astorga et al., 1994; Eagle & Hanley-Bowdoin, 1997). Hence, the coding sequences are located on both strands of the double-stranded, transcriptionally active intermediate molecule, the virion strand and the complementary strand. This kind of genome organization is similar to that of some animal DNA viruses, such as simian virus 40 (SV40), and suggests the existence of a pattern of temporal gene expression. In this work, the transcription of pepper huasteco yellow vein virus (PHYVV) genes was studied. Green fluorescent protein replacements and RT-PCR analyses were used to monitor PHYVV gene expression chronologically in suspension cells and plant tissue. A model is proposed to describe the order of geminivirus gene expression, where the genes that encode Rep, TrAP and REn are expressed during an early stage of infection. The genes that encode the coat protein and the nuclear shuttle protein are expressed during the late stage of infection.

Early and late gene expression in pepper huasteco yellow vein virus

Harumi Shimada-Beltrán and Rafael F. Rivera-Bustamante

Departamento de Ingeniería Genética, Cinvestav-Unidad Irapuato, km 9.6 Libramiento Norte, Apartado Postal 629, 36500 Irapuato, GTO, Mexico

Viral infections usually take place in an orderly manner and can be divided into at least two phases: an early and a late stage. In geminiviruses, plant viruses with a circular, single-stranded DNA genome, expression of viral genes involves complex regulation strategies that suggest the existence of a pattern of temporal gene expression. In this work, the transcription of pepper huasteco yellow vein virus (PHYVV) genes was studied. Green fluorescent protein replacements and RT-PCR analyses were used to monitor PHYVV gene expression chronologically in suspension cells and plant tissue. A model is proposed to describe the order of geminivirus gene expression, where the genes that encode Rep, TrAP and REn are expressed during an early stage of infection. The genes that encode the coat protein and the nuclear shuttle protein are expressed during the late stage of infection.
with TGMV, Shung & Sunter (2007) demonstrated that, when Rep is downregulated, synthesis of the smaller complementary-sense transcripts that encode TrAP and REn is enhanced. Thus, geminiviruses have acquired an efficient way to compensate for their small-size genome. Currently, however, little is known about the temporal/spatial regulation of begomovirus gene expression.

In this study, we report the temporal expression analysis of pepper huasteco yellow vein virus (PHYVV), a bipartite begomovirus reported in Mexico (Garzón-Tiznado et al., 1993; Ruiz-Medrano et al., 1999; Torres-Pacheco et al., 1993). PHYVV gene expression was monitored in Nicotiana tabacum NT1 suspension cells by using PHYVV green fluorescent protein (GFP) replacements and by RT-PCR of infected plant tissues. Expression of Rep and MP was shown to occur at an early stage, whereas expression of CP and NSP occurs at a late stage. In addition, we suggest that TrAP and REn are expressed at an early and an early/late stage, respectively. A model of begomovirus temporal gene expression is proposed and discussed.

METHODS

Cell culture, viruses and plants. N. tabacum NT1 cells were used for transient assays. Viral sequences were obtained from PHYVV (GenBank accession nos NC001359, and NC001369). N. tabacum ‘Xanthi’ plants were used for biolistic inoculations.

GFP-replacement constructs. All DNA techniques and molecular biology procedures were performed according to common protocols (Sambrook et al., 1989) unless stated otherwise. For construction of the GFP-replacement plasmids, a dimeric clone of PHYVV-A was used. In this construct, tandemly repeated copies of PHYVV-A are inserted into the HindIII site of pBluescript (Stratagene) (M. Bonilla & R. F. Rivera-Bustamante, unpublished data). To obtain Rep-GFP, primers were designed with BamHI and NcoI restriction sites at their ends to direct the amplification of the whole viral genome, except for the region of Rep that would be replaced by the GFP sequence (see Supplementary Table S1, available in JGV Online). The amplified product was inserted into TOPO-PCR4 (Invitrogen), excised with BamHI and XhoI and inserted into pBS-Rep using the same enzymes. Afterwards, the construct was rearranged by cutting with Ncol, religating and recutting with HindIII, and inserted into pBluescript. This construct was named Rep-GFP (Fig. 2). The constructs TrAP-GFP and REn-GFP (Fig. 2) were constructed similarly (see Supplementary Table S1, available in JGV Online). CP-GFP, which is a translational fusion where GFP replaces CP from codons 6 to 193, has been described previously (Méndez-Lozano et al., 2003).

GFP fluorescence analyses. GFP expression was monitored by using a fluorescence stereoscope (Leica MZ8) with the set of filters GFP-Plus (excitation filter: 480/40 nm; emission: 510 nm) and a fluorescence microscope (Leica DMRE) with an excitation filter (I3) BP 450–490 nm, a dichroic mirror at 510 nm and an observation filter at 515 nm. Images were generated by using a digital camera (Spot Diagnostic Instruments) and edited by using Adobe Photoshop CS software.

Transfection assays. Tungsten particles for microprojectile bombardment were prepared as described by Cabrera-Ponce et al. (1997). On the fourth day after subculture, 3 ml N. tabacum (NT1) cell suspension was collected on filter papers and placed in solid osmotic medium [MS medium, 30 g sucrose l⁻¹, 2 mg 2,4-dichlorophenoxyacetic acid (2,4-D) l⁻¹, pH 5.7, 0.25 M mannitol] for 3 h. DNA to be used in the transfection assays was digested with HindIII to liberate the vector from the PHYVV-A construct and with BamHI for the PHYVV-B construct. The cells were inoculated by using a biolistic DNA-delivery device (PDS1000-He; DuPont Bio-Rad) at a pressure of 800 p.s.i. (5.52 MPa). Afterwards, the filters with the cells were transferred to solid NT1 medium (MS medium, 30 g sucrose l⁻¹, 2 mg 2,4-D l⁻¹, pH 5.7) and incubated at 25 °C until monitored or collected at different times post-inoculation (p.i.). For RNA extraction, the cells were collected in liquid nitrogen and stored at –70 °C until extraction. At least five independent experiments were performed in each case, with reproducible results.

Plant inoculation. N. tabacum plants propagated for 3–4 weeks in MS medium were inoculated by using the biolistic DNA-delivery device at a pressure of 1200 p.s.i. (8.28 MPa). The plants were grown in growth chambers at 25 °C.

RNA isolation and RT-PCR. Total RNA was extracted by using TRIzol reagent (Invitrogen) according to the procedure recommended by the manufacturer, and was subjected to a second extraction with TRIzol reagent or to a CTAB wash as described by Camacho-Villasana et al. (2002). Prior to the RT-PCR analyses, the RNA was treated with DNase I (amplification grade; Invitrogen) for 1 h at 37 °C, followed by inactivation of the enzyme at 70 °C for 10 min. For the reverse-transcription step, 0.5–1.0 µg RNA and 100 ng primer were used. The conditions were as follows: 42 °C
(55 min), 65 °C (10 min) (SuperScript II, Invitrogen; RevertAid M-MuLV, Fermentas) and then amplification for 40 cycles at 94 °C (1 min), 55 °C (1 min) and 72 °C (1 min). Gel images were generated by using the Digi-Doc-it gel-documentation system (UVP Bioimaging systems).

RESULTS

Temporal expression of PHYVV GFP replacements

Several strategies (transcriptional or translational fusion to marker genes, RT-PCR or Northern blots assays, etc.) have been employed to study the temporal regulation of gene expression of an organism. In this report, we employed a combination of strategies to assure a comprehensive analysis. To follow the expression of PHYVV genes over time, we first used GFP as a reporter gene (Heim et al., 1995; Reichel et al., 1996). The Rep, TrAP, REn and CP ORFs of PHYVV were replaced partially or completely with the GFP coding region (Fig. 2). The resulting constructs were used to inoculate, via particle bombardment, N. tabacum NT1 suspension cells as described in Methods. Inoculated cells were observed under a fluorescence stereoscope and the GFP expression directed by the different constructs was determined. The kinetics of expression of each PHYVV GFP construct were determined by monitoring and counting the inoculated cells expressing GFP at several intervals. Initially, the observation time points were set at 0, 3, 5, 7, 9, 13, 15 and 25 h p.i., to verify the beginning and extent of fluorescence detection. When the cells were inoculated with the Rep-GFP construct, fluorescent cells were consistently observed as early as 3–4 h p.i. (Fig. 3a). Cells inoculated with the TrAP-GFP construct started to show GFP fluorescence at time points similar to the Rep-GFP construct. REn-GFP-inoculated cells, on the other hand, started to emit detectable fluorescence only after 5 h p.i. It is noteworthy to mention that these constructs (Rep-GFP, TrAP-GFP and REn-GFP)
do not replicate, as the proteins Rep or REn are either missing or truncated (Fig. 2). Nevertheless, the results shown could be considered as equivalent to transcriptional events happening at the early stage of infection, before replication takes place. In the case of cells inoculated with the CP-GFP construct (a rightward gene), strong fluorescence was detected by 9 h p.i. in the inoculated cells (Fig. 3a). This late and high-level expression has been reported previously for the CP gene in several geminiviral systems (Ruiz-Medrano et al., 1999; Shivaprasad et al., 2005; Sunter & Bisaro, 1991, 1997).

It was of interest to verify whether the pattern of expression of viral genes observed in this experimental system (suspension cells and bombardment with geminivirus gene replacements by GFP) could be correlated with those from previous reports, obtained by using other experimental systems (transgenic plants or protoplasts and different reporter genes) (Ruiz-Medrano et al., 1999; Shivaprasad et al., 2005; Sunter & Bisaro, 1991, 1997). Therefore, the intensity of fluorescence was also analysed. Cells bombarded with either Rep- or REn-GFP replacements emitted fluorescence with a similarly low intensity (Fig. 4a, b). In contrast, cells bombarded with TrAP-GFP consistently emitted up to 40% more fluorescence than cells bombarded with Rep-GFP (Fig. 4a, c).

Finally, cells inoculated with CP-GFP showed the highest fluorescence intensity (Fig. 4d). The expression of CP-GFP was, in some cases, as strong as that of a 35S-GFP construct used as a control (Fig. 4e). These GFP expression results are comparable to those reported for other begomoviruses (Brough et al., 1992; Shivaprasad et al., 2005).

Individual cells were observed throughout time to determine, in addition to the start point, the duration of GFP expression in independent cells. Again, cells inoculated with Rep-GFP and TrAP-GFP constructs began to express GFP at 3 h p.i., REn-GFP constructs at 5 h p.i. and CP-GFP constructs at 9 h p.i. However, whereas TrAP-GFP inoculated cells continued to express GFP after 32 h p.i., cells inoculated with REn-GFP constructs replacing the Rep and REn genes showed a noticeable decrease in fluorescence around 24 h p.i. (data not shown).

To characterize the regulation of PHYVV gene expression better, the effect of the presence of wild-type (wt) virus (and therefore viral proteins) on the expression of GFP directed from the different viral constructs was also analysed. NT1 suspension cells were bombarded independently with the different PHYVV GFP constructs with or without DNA of wt PHYVV-A. The fluorescent cells expressing GFP were monitored and counted at different time points as described above. Cells co-inoculated with TrAP-GFP plus PHYVV-A, REn-GFP plus PHYVV-A or CP-GFP plus PHYVV-A had differences neither in fluorescence intensity nor in the starting point of GFP expression, compared with cells inoculated with the GFP replacements alone (Fig. 3b). In contrast, no GFP expression was observed in cells co-inoculated with Rep-GFP plus PHYVV-A (Fig. 3b). Although this kind of Rep self-repression has been reported previously in other begomoviruses (Eagle et al., 1994; Sunter et al., 1993), it was unexpected to see a complete lack of expression when both DNAs (Rep-GFP and PHYVV-A) were co-inoculated. Overall, the fact that the leftward genes Rep, TrAP and REn exhibit different patterns of expression (i.e. varied fluorescence intensity and differences in self-regulation mechanism) suggests that these genes are regulated in a different manner and that these characteristics are maintained in the experimental system used here.

**Temporal expression of PHYVV-A transcripts**

Although the GFP analysis provided information about a temporal pattern for PHYVV gene expression, there is

---

**Fig. 4.** Expression of GFP in NT1 suspension cells inoculated with each of the PHYVV GFP gene replacements. The GFP gene-replacement constructs (a) Rep-GFP (exposure time, 2 s), (b) REn-GFP (2 s), (c) TrAP-GFP (0.5 s), (d) CP-GFP (0.5 s) and (e) 35S-GFP (0.5 s) were bombarded individually into NT1 suspension cells. GFP expression was observed by using a fluorescent microscope (x20). Bars, 50 μm.
always a possibility that the GFP substitutions eliminated some regulatory elements influencing transcription. Therefore, we corroborated the GFP data with RT-PCR experiments using cells inoculated with wt PHYVV DNAs A and B, co-inoculated into N. tabacum NT1 suspension cells. Samples were taken at different time points and total RNA extracted from the inoculated cells was used for RT-PCR assays, employing pairs of primers that direct the amplification of segments of the Rep, TrAP/REn and CP ORFs [see Methods, Fig. 1 and Supplementary Table S1 (available in JGV Online)]. By using the set of Rep-specific primers, Rep/leftward transcripts were detected as early as 2 h p.i. (Fig. 5a). Likewise, a PCR product was detected at 2 h p.i. by using primers located in the TrAP/REn ORFs. In contrast, no amplification product was detected in the RT-PCR assays until 6 h p.i. when using the set of primers located in the CP ORF (Fig. 5a). The RT-PCR assays were more sensitive than the GFP assays and the CP mRNA was detected earlier; nevertheless, the overall temporal expression observed for PHYVV genes using both analytical methods (GFP expression and RT-PCR assays) showed a good correlation in the experiments using NT1 suspension cells.

It was also of interest to verify whether the regulation of PHYVV gene expression observed in tobacco suspension cells reflected that in infected leaf tissue. To analyse the kinetics of PHYVV-A transcripts in plant tissue, similar RT-PCR assays were performed using RNA extracts from inoculated leaves. N. tabacum plants were co-inoculated with PHYVV-A and -B components via bombardment, and samples were harvested from 0 to 24 h p.i. As observed with NT1 suspension cells, transcripts from the coding regions of Rep and TrAP/REn were detected as early as 2 h p.i. (Fig. 5b). To determine whether these primers targeted the same or different transcripts, we analysed the transcripts by real-time RT-PCR (qRT-PCR). The qRT-PCR assays showed that products on both regions could be detected as early as 2 h p.i. In addition, there was more product amplified from the TrAP/REn region, suggesting a higher level of transcription (see Supplementary Fig. S1, available in JGV Online). On the other hand, CP transcripts were detected only after 4 h p.i. (Fig. 5b). The minor differences observed with the different experimental systems and analysis techniques (GFP/suspension cells, RT-PCR/suspension cells or RT-PCR/leaf tissue) could be due to physiological variations. Nevertheless, the same overall pattern of early and late gene expression was observed in all cases. No amplification products were detected in RT-PCR assays using RNA extracted from healthy N. tabacum plants (data not shown). Compared with infected cells, CP transcription was detected earlier, but still a bit later than that of the leftward transcripts.

Temporal expression of PHYVV-B transcripts

The two genes of component B of bipartite begomoviruses encode an NSP and an MP (Hanley-Bowdoin et al., 2000; Hehnle et al., 2004; Sanderfoot & Lazarowitz, 1996), which are oriented in the viral sense (NSP) and the complementary sense (MP) (Fig. 1). To investigate whether the expression of MP and NSP genes in PHYVV-B is also regulated temporally, RT-PCR analyses were carried out. N. tabacum NT1 suspension cells were co-inoculated by bombardment with wt PHYVV DNA (PHYVV -A and -B). Total RNA was then extracted from samples taken at different time points. NSP and MP transcripts were
detected at 8 h p.i. (Fig. 5c). A faint band of the amplification product was occasionally seen as early as 2 h p.i. for both NSP and MP. However, no transcripts were detected at 4 or 6 h p.i. It is possible that this signal is due to an elevated basal expression, due to a sporadic exceptionally high concentration of inoculum DNA in some cells.

As with component A, component B transcripts were also analysed in plant-leaf tissue. RT-PCR was performed on RNA extracted from leaves inoculated via particle bombardment with both PHYVV components (A and B). As shown in Fig. 5(d), the RT-PCR product of NSP was detected at 4 h p.i., as observed with the CP transcripts (Fig. 5b). In contrast, the RT-PCR product of MP was detected as early as 2 h p.i. In some cases, the MP product was detected at the same time point as NSP (4 h p.i.) (data not shown).

**DISCUSSION**

Most viruses show a temporal expression of their genes during infection. Usually, the first viral genes to be expressed are those that encode proteins involved in viral replication and in the transactivation of viral and/or host genes, whereas structural proteins (e.g. coat proteins) are typically late proteins (Berk, 1986; Cook & Coen, 1996; Zhang & Wagner, 1987). A classical example is SV40, in which the coding regions have been divided into an early region that includes the multifunctional proteins large T and small t antigens, involved in activation of transcription and replication, and a late region that contains the structural proteins (Balakrishnan & Milavetz, 2006; Fanning & Knippers, 1992; Farrell & Mertz, 2002; Hartzell et al., 1984; Kelly & Wildeman, 1991; Wiley et al., 1993; Zuo & Mertz, 1995).

In this work, we have used two different types of analysis (RT-PCR and GFP gene replacements) in two different experimental systems, cultured NT1 cells and infected leaf tissue, to study the temporal expression of PHYVV genes. We chose a cell-line suspension to have a uniform population of replicating cells inoculated at similar growth stages to facilitate the detection of transcriptionally active virus genes. In addition, we inoculated leaf tissue, which may better reflect a natural infection. However, it has the disadvantages of a low number of inoculated cells and the presence of naturally fluorescent molecules that potentially complicate the analysis of GFP experiments. RT-PCR is a well-accepted method to detect RNA as an indication of the transcriptional regulation of a given gene. It is sensitive, it can be quantified and several transcripts can be detected simultaneously. However, the overlapping arrangement of the complementary-sense genes of geminivirus DNA A limits the potential of an RT-PCR-only approach to study gene expression. As no reliable immunological-detection methods for PHYVV complementary sense-encoded proteins are available, we instead used gene replacements by GFP. GFP is a widely used reporter protein that can be analysed in live tissue and it can be fused to peptides without modifying their expression or regulation significantly (Heim et al., 1995; Reichel et al., 1996). A limitation of gene replacements is that potential regulatory elements within the genes to be analysed are lost in the GFP replacements. Therefore, we chose both approaches, RT-PCR and GFP gene replacements, to compensate for their respective disadvantages. In the case of the PHYVV GFP replacements, the fact that the different constructs showed different patterns of expression (i.e. differences in intensity and temporal appearance of GFP fluorescence, as well as the distinct response when co-inoculated with wt PHYVV) suggested strongly that GFP detection can indeed be considered a valid reflection of the actual expression of the viral genes. The correlation of the GFP and RT-PCR data for Rep and CP expression also supports the validity of the results.

By using both analysis systems, we demonstrated that the genes of the begomovirus PHYVV are indeed expressed in a temporally regulated manner. PHYVV transcription takes place in at least two phases: first, an early phase in which Rep, TrAP and REn genes are expressed from component A, whilst the MP gene of component B is expressed soon after. The CP and NSP genes, on the other hand, are expressed at a second and later stage in the virus replication cycle.

As the genes on the complementary strand of component A overlap, it is difficult to discriminate differences in gene expression by using RT-PCR assays. For example, early expression of the Rep gene is detected easily by using PCR primers located in the 5′ end of the ORF. However, primers located in the 3′ end will not discriminate between the long transcript reported for Rep and shorter transcripts described for TrAP and REn of other begomoviruses. However, sets of primers designed for qRT-PCR assays with PHYVV showed a higher amount of mRNA when using the TrAP/REN primers compared with the 5′ Rep primers at the times analysed (2 and 11 h p.i.; see Supplementary Fig. S1, available in JGV Online). This suggests that at least two transcripts are produced from the leftward region: a long one that includes the entire Rep ORF, and one (or more) transcript(s) covering the second half of the Rep ORF. Thus, RT-PCR assays only confirmed the presence of at least two leftward transcripts expressed as early as 2 h p.i. Attempts to identify the initiation site(s) of the small transcript(s) by using a 5′ RACE (rapid amplification of cDNA ends) protocol produced inconclusive results (data not shown). The differences observed with the TrAP- and REn-GFP constructs may suggest the existence of two distinct transcripts, as cells inoculated with TrAP-GFP showed earlier and higher fluorescence than cells inoculated with REn-GFP. Also, TrAP-GFP-inoculated cells remained fluorescent for a longer period. However, we cannot exclude completely the possibility of one multicistronic transcript with a complex translational regulation (Hanley-Bowdoin et al., 1989; Shung et al., 2006).
The classification of CP and NSP genes as late is based on the facts that their expression requires the presence of TrAP (Ruiz-Medrano et al., 1999; Shivaprasad et al., 2005; Sunter & Bisaro, 1991, 1992) and that their detection was observed several hours later in both types of assay.

The classification of MP as an early or late gene was more complicated. In PCR assays using plant tissue, MP mRNA was detected as early as 2 h p.i. However, in the cell-suspension assays, MP expression was detected as late as NSP expression. Therefore, we cannot rule out the possibility that MP expression in plants occurs at an early stage.

It is not clear whether the type of inoculated cell, suspension cells as opposed to leaf tissue, influences the expression of a protein expected to interact with components of plasmodesmata. The variation in the detection times observed between leaf tissue and suspension cells might reflect the particular characteristics of the respective cell types.

The temporal regulation of viral gene expression observed here is consistent with previously reported regulatory mechanisms in geminivirus infections. This concerns the downregulation of Rep (Haley et al., 1992; Sunter et al., 1993) and the resulting enhanced expression of TrAP and REn (Shung & Sunter, 2007). Also, the fact that TrAP transactivates the expression of the CP and NSP genes (Ruiz-Medrano et al., 1999; Sunter & Bisaro, 1991, 1992) requires its expression prior to that of CP and NSP.

Based on the demonstrated temporal expression and taking into account the different functions of the viral proteins, PHYVV infection follows the generally accepted model of geminivirus infection (Fig. 6). First, the virus enters the host cell; it is uncoated and the ssDNA genome is converted by the host machinery into double-stranded DNA to form the replicative form (RF) that is also the template for transcription. The early genes Rep and TrAP are expressed (Fig. 6a). They prepare the host cell for replication and block some defence pathways, respectively. REn is also expressed and, together with Rep, initiates viral replication (Fig. 6b). At some stage, Rep represses its own expression, and the synthesis of the shorter complementary-sense transcript(s) encoding TrAP and REn is maintained at a high level or is even increased (c). At this stage, MP is also expressed. TrAP expression is maintained, as it is still required as a pathogenicity determinant and a host susceptibility enhancer (Sunter et al., 2001). Then, TrAP activates expression of the CP and NSP genes (Fig. 6c).

Understanding the fine tuning of viral gene expression may also provide further knowledge about gene regulation of the host and may ultimately lead to novel strategies for geminivirus control.

**ACKNOWLEDGEMENTS**

We are very grateful to Drs Bruno Gronenborn and Cecilia Ramírez for invaluable and helpful discussions about the work. We also thank Drs Julio Vega-Arreguin and Jimena Carrillo-Tripp for useful
comments on the manuscript, Dr Cristina Reynaga for technical assistance in microscopy and Rosa Maria Adame for her support with the cell culture. The CP-GFP construct pJMHPVAGFP was kindly provided by J. Méndez-Lozano (Departamento de Ingeniería Genética, Cinvestav-Unidad Irapuato, Irapuato, Mexico) and C. Reichel (Scripps Research Institute, La Jolla, CA, USA). H. S.-B. was supported by a PhD fellowship from Conacyt-México. This work was supported by a grant from Conacyt-SAGARPA, Mexico (2004-115), to R. F. R.-B.

REFERENCES


