Geminiviral genomes encode additional proteins with specific subcellular localizations and virulence function

Pan Gong[1]*, Huang Tan[2,3]*, Siwen Zhao[1], Hao Li[1], Hui Liu[4], Yu Ma[2,3], Xi Zhang[2,3], Junjie Rong[2,3], Xing Fu[2], Rosa Lozano-Durán[2,5#], Fangfang Li[1#], Xueping Zhou[1,4#]

[1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, China [2]Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 201602, China. [3]University of the Chinese Academy of Sciences, Beijing 100049, China. [4]State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, Zhejiang, 310058, China [5]Department of Plant Biochemistry, Centre for Plant Molecular Biology (ZMBP), Eberhard Karls University, D-72076 Tübingen, Germany.

*These authors contributed equally to this work.

[#]Co-corresponding authors: Rosa Lozano-Durán (rosa.lozano-duran@uni-tuebingen.de; lozano-duran@psc.ac.cn); Fangfang Li (lifangfang@caas.cn); Xueping Zhou (zzhou@zju.edu.cn).

ABSTRACT

Geminiviruses are plant viruses with limited coding capacity. Geminivirus-encoded proteins were identified applying a 10-kDa arbitrary threshold; however, it is increasingly clear that small proteins play relevant roles in biological systems, which calls for the reconsideration of this criterion. Here, we show that geminiviral genomes contain additional ORFs. Using tomato yellow leaf curl virus, we demonstrate that some of these novel ORFs are expressed during the infection, and that the encoded proteins display specific subcellular localizations. We prove that the largest of these new ORFs, which we name V3, is required for full viral infection, and that the V3 protein localizes in the Golgi apparatus and functions as an RNA silencing suppressor. These results imply that the repertoire of geminiviral proteins can be expanded, and that getting a comprehensive overview of the molecular plant-geminivirus interactions will require the detailed study of small ORFs so far neglected.

INTRODUCTION

Viruses are intracellular parasites that heavily rely on the host cell machinery to complete their life cycle. Most viruses have small genome sizes, with the concomitant limitation in coding capacity; in order to overcome the restrictions imposed by their reduced proteome, viruses have evolved to encode multifunctional proteins that efficiently target hub proteins in their host cells (reviewed in Brito and Pinney, 2017; King et al., 2018). Nevertheless, higher numbers of virus-encoded proteins might enable more sophisticated infection mechanisms, and therefore maximization of the coding space would be expected to be an advantage to the pathogen.

Geminiviruses are a family of plant viruses with circular, single-stranded (ss) DNA genomes causing devastating diseases in crops around the globe. This family includes nine genera, based on host range, insect vector, and genome structure: *Becurtovirus*, *Begomovirus, Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus*, *Turncurtovirus*, *Capulavirus*, and *Grablovirus* (Zerbini et al., 2017); most species described to date belong to the genus *Begomovirus*. Members of this family have small genomes, composed of one or two DNA molecules of less than 3 Kb each, in which the use of coding space is optimized by bidirectional and partially overlapping open reading frames (ORFs): in one <3 Kb molecule, geminiviruses contain up to 7 ORFs, with a known maximum of 8 viral proteins per species. The geminiviral infection cycle is complex, and multiple steps remain to be fully elucidated. Following transmission by an insect vector, the geminiviral DNA genome must be released from the virion and reach the nucleus, where it will be converted into a double-stranded (ds) DNA replicative intermediate; this dsDNA molecule will serve as template for the transcription of viral genes, including the replication-associated protein (Rep), which reprograms the cell cycle and recruits the host DNA replication machinery. Rolling-circle replication ensues, by which new ssDNA copies of the viral genome are produced. Eventually, the virus must move intracellularly, intercellularly, and systemically, invading new cells and making virions available for acquisition by the vector. In order to accomplish a successful infection, geminiviruses must tailor the cellular environment to favour their replication and spread; for this purpose, they modify the transcriptional landscape of the infected cell, re-direct post-transcriptional modifications, and interfere with hormone signalling, among other things (reviewed in Aguilar et al., 2020; Kumar, 2019; Liu et al., 2021), ultimately suppressing anti-viral defences, creating conditions favourable to viral replication, and manipulating plant development. Although geminivirus-encoded proteins are described as multifunctional, how the plethora of tasks required for a fruitful infection can be performed by only 4-8 proteins is an intriguing biological puzzle.

Whether members of this family encode additional small proteins, below the arbitrary 10 kDa threshold established following identification of the first geminivirus species, remains elusive.

Tomato yellow leaf curl virus (TYLCV) is a monopartite begomovirus, causal agent of the destructive tomato leaf curl disease (Basak, 2016). The TYLCV genome contains six known open reading frames (ORFs), encoding the capsid protein (CP)/V1 and V2 in the virion strand, and C1/Rep, C2, C3, and C4 in the complementary strand. Rep creates a cellular environment permissive for viral DNA replication and attracts the DNA replication machinery to the viral genome (reviewed in Hanley-Bowdoin et al., 2013); C2 suppresses post-transcriptional gene silencing (PTGS) (Luna et al., 2012), protein ubiquitination (Lozano-Duran et al., 2011) and jasmonic acid (JA) signalling (Lozano-Duran et al., 2011; Rosas-Diaz et al., 2016); C3 interacts with PCNA, NAC, and the regulatory subunits of DNA polymerases α and δ to enhance viral replication (Castillo et al., 2003; Settlage et al., 2005; Wu et al., 2020); C4 is a symptom determinant, interferes with the intercellular movement of PTGS, and hampers salicylic acid (SA)-dependent defences (Luna et al., 2012; Medina-Puche et al., 2020; Rosas-Diaz et al., 2018); the CP forms the viral capsid and is essential for the transmission by the insect vector, and shuttles the viral DNA between the nucleus and the cytoplasm (Azzam et al., 1994; Diaz-Pendon et al., 2010; Gotz et al., 2012; Kunik et al., 1998; Ohnesorge and Bejarano, 2009; Palanichelvam et al., 1998; Rojas et al., 2001; Rubinstein and Czosnek, 1997); V2 is a strong suppressor of PTGS as well transcriptional gene silencing (TGS), and it mediates the nuclear export of CP (Wang et al., 2014; Wang et al., 2018; Wang et al., 2020; Zhao et al., 2020; Zrachya et al., 2007). Interestingly, a recent report identified 21 transcription initiation sites within the TYLCV genome by taking advantage of cap-snatching by rice stripe virus (RSV) in the experimental *Solanaceae* host *Nicotiana benthamiana*, suggesting that transcripts beyond those encoding these known ORFs might exist (Lin et al., 2017). This idea is further indirectly supported by the fact that attempts at knocking-in tags in geminiviral genomes have so far been fruitless.

Here, we report that geminiviral genomes contain additional ORFs besides the canonical ones described to date. These previously neglected ORFs frequently encode proteins that are phylogenetically conserved. Using the geminivirus TYLCV as an example, we show that some of these ORFs are transcribed during the viral infection, and that the proteins they encode accumulate in the plant cell and show specific subcellular localizations and distinctive features. Moreover, we demonstrate that one of these novel ORFs, which we have named V3 and is conserved in begomoviruses, is essential for full infectivity in *N. benthamiana* and tomato, and encodes a Golgi-localized protein that acts as a suppressor of PTGS and TGS. Taken together,

our results indicate that geminiviruses encode additional proteins to the ones described to date, which may largely expand the geminiviral proteome and its intersection with the host cell.

RESULTS

The TYLCV genome contains additional conserved open reading frames

It is increasingly clear that small proteins (<100 aa) are prevalent in eukaryotes, including plants, and have biological functions (reviewed in Eguen et al., 2015; Hsu and Benfey, 2018; Murphy et al., 2012). In order to explore whether geminiviral genomes may contain additional ORFs encoding small proteins of predictable functional relevance, we designed a tool that we called ViralORFfinder; this tool uses the ORFfinder from NCBI (https://www.ncbi.nlm.nih.gov/orffinder/) to identify ORFs in an inputted subset of DNA sequences (geminiviral genomes, in this case) and creates a small database with the translated protein sequences, which can be used to BLAST a protein of choice, therefore assessing conservation among the selected species (Figure 1a). The distribution of the protein of interest is then displayed in a phylogenetic tree of the inputted viral species, generated based on the DNA sequences provided. Using ViralORFfinder, additional ORFs can be consistently predicted in geminiviruses of different genera, as illustrated for bipartite begomoviruses (Supplementary figure 1; Supplementary table 2), curtoviruses (Supplementary figure 2; Supplementary table 3), and mastreviruses (Supplementary figure 3; Supplementary table 4); the proteins encoded by some of these ORFs are conserved among species.

We then used ViralORFfinder to identify additional proteins encoded by monopartite begomoviruses causing tomato leaf curl disease isolated from different regions of the world (see Methods section; Supplementary figure 4; Supplementary table 5), and identify those present in TYLCV and conserved in other species. As shown in Figure 1b, 43 ORFs encoding proteins of >10 aa were identified in the TYLCV genome. Interestingly, a significant correlation can be found between the size of the encoded proteins and their representation in the selected subset of species, with the six larger proteins (>10 kDa) present in all of them (Supplementary figure 4b, pink). For further analyses, we selected the 6 ORFs that followed in size and prevalence (Supplementary figure 4b, blue), named ORF1-6; the position of these ORFs in the TYLCV genome is shown in Figure 1c. Of note, the proteins encoded by these ORFs are also conserved in other members of the *Begomovirus* genus, both bipartite and monopartite, infecting a broad range of hosts (Supplementary figures 5-7; Supplementary table 6; Figure 1d).

The novel ORFs in TYLCV encode proteins with predicted domains and specific subcellular localizations

With the aim of gaining insight into the properties of the proteins encoded by the new ORFs from TYLCV, we investigated the presence in their sequence of predicted domains or signals, namely transmembrane domains (TM), nuclear localization signal (NSL), and chloroplast transit peptide (cTP). As shown in Figure 2a, while none of these proteins contains an NLS, and only one of them contains a cTP, three of them contain a predicted TM, which is not present in any of the previously characterized proteins.

We then cloned ORF1-6, fused them to the GFP gene, and transiently expressed them in *N. benthamiana* leaves; confocal microscopy indicates that these fusion proteins present specific subcellular localizations (Figure 2b). Whereas the ORF1-encoded protein is mostly nuclear, co-expression with marker proteins or dyes unveils that the ORF2-encoded protein localizes in the endoplasmic reticulum (ER), as demonstrated by the co-localization with RFP-HDEL; the ORF3-encoded protein in mitochondria, as demonstrated by the co-localization with MitoTracker; the ORF4-encoded protein in the ER and the Golgi apparatus, as demonstrated by the partial co-localization with RFP-HDEL and SYP32-RFP; and the ORF5- and ORF6-encoded proteins mostly in Golgi, as demonstrated by the partial co-localization with SYP32-RFP (Figure 2c-g; Supplementary figure 8). The specific subcellular localization exhibited by each of these proteins suggests that, despite their small size (5.3-9.3 kDa; Figure 2a), either their sequence contains the appropriate targeting signals, or they interact with plant proteins that enable their precise targeting in the cell.

A prerequisite for these ORFs to have a biological function is their expression in the context of the viral infection. Therefore, we cloned the TYLCV genomic sequences upstream of each ATG (pORF1-6) before the GFP reporter gene and tested their promoter activity in transiently transformed *N. benthamiana* leaves. As shown in Supplementary figure 9, none of these sequences could drive GFP expression in the absence of the virus, but pORF1, pORF2, pORF4, and pORF5 could when the virus was present; the sequence upstream of the C4 ORF was used as positive control, and could activate GFP expression both in the presence and absence of TYLCV. The promoter activity of these upstream sequences in infected cells strongly suggests that at least ORF1, 2, 4, and 5 are expressed during the viral infection.

The novel V3 protein from TYLCV is a Golgi-localized silencing suppressor required for full infection

ORF6 is the largest of the newly described ORFs in the TYLCV genome, and the protein it encodes displays the highest degree of conservation in a selected subset of 26 representative begomoviruses (Supplementary figure 7). Therefore, we decided to further characterize this ORF as a proof-of-concept of the potential biological roles of novel ORFs. Hereafter, ORF6 will be referred to as V3, since it is the third ORF on the viral strand in the TYLCV genome.

The V3 ORF is located in positions 2350-2583 of the TYLCV genome (TYLCV-BJ; Figure 1d), and encodes a 77-amino acid protein. In different begomovirus species, the V3 ORF ranges from 87 to 234 nt, and the protein it encodes presents a high degree of similarity, with 6 residues completely conserved (Supplementary figure 7).

In order to determine whether V3 is transcribed during the viral infection, we checked if the corresponding transcript was present in TYLCV-infected samples: as shown in Figure 3a, the V3 transcript was found upon TYLCV infection, but not in uninfected plants. 5′ rapid amplification of cDNA ends (RACE) was then used to identify the transcriptional initiation site of V3, which was found to be located between 176 and 421 upstream of the start codon (Figure 3b). Interestingly, most of the sites identified by RACE (7 out of 11) are close to A2058, which was previously isolated by cap-snatching of RSV (Lin et al., 2017).

The finding that a transcript corresponding to the V3 ORF can be identified in infected samples strongly suggests that the sequence upstream of this ORF must act as a promoter. Since the 500 bp fragment previously tested (Supplementary figure 9) did not show promoter activity, we tested a larger, 833-nt sequence upstream of the V3 start codon, which was cloned before the GUS or GFP reporter genes. As shown in Figure 3c, d, this viral sequence could activate GUS expression, leading to detectable GUS activity, and it could also drive expression of GFP (Figure 3e-g), albeit more weakly than the 35S promoter. Taken together, these results confirm that the V3 ORF can be expressed *in planta* from its genomic context.

Given that protein function is tightly linked to its spatial location, we then decided to analyse the subcellular localization of the V3 protein in detail. As can be seen in Figure 3h, V3 is a Golgi-localized protein; this localization does not change in the presence of the virus (Supplementary figure 10a). A closer observation confirms that V3 is localized to cis-Golgi, as shown by the co-localization of V3-GFP with the markers Man49-mCherry and SYP32-RFP (Figure 3h; Supplementary figure 10b); nevertheless, the YFP fusions YFP-V3 and V3-YFP can also partially

co-localize with the endoplasmic reticulum (ER), as indicated by their partial co-localization with the ER marker mCherry-HDEL or RFP-HDEL (Figure 3h; Supplementary figure 10c, d), which may reflect the transition of the protein from the ER to the Golgi apparatus.

With the aim of assessing the biological relevance of the V3 protein for the TYLCV infection, we next generated a mutated infectious clone carrying a T2351C substitution in the V3 ATG, hence impairing the production of the V3 protein. Since the V3 ORF overlaps with the Rep/C1 and C4 ORFs, nt replacements in the start codon of the V3 ORF necessarily affect the protein sequence of the resulting Rep or C4 proteins; the chosen change results in a I89V substitution in the Rep/C1 protein, with no change in C4. This mutant infectious clone is hereafter referred to as TYLCV-mV3.

TYLCV-mV3 was then inoculated into *N. benthamiana* plants, and its performance compared to that of the wild-type (WT) virus (TYLCV-WT). At 10 days post-inoculation (dpi), TYLCV-mV3-infected plants displayed mild leaf curling symptoms and presented lower viral DNA load compared to plants inoculated with TYLCV-WT (Figure 4a, b), which correlated with a lower accumulation of CP (Figure 4c). These differences apparently result from a delay in the infection, measured as symptom appearance (Figure 4d). To evaluate the potential functional impact of the I89V substitution in Rep/C1 on the viral infection and disentangle this effect to that derived from the absence of V3, transgenic *N. benthamiana* lines expressing V3-YFP under a 35S promoter were generated, and a complementation assay was performed. Of note, the V3-expressing plants do not display obvious developmental abnormalities (Supplementary figure 11a); the expression of V3-YFP was confirmed by qRT-PCR and western blot (Supplementary figure 11b, c). As shown in Figure 4e and Supplementary figure 11d, transgenic expression of V3-YFP could fully complement the lack of V3 in the TYLCV-mV3 clone, measured as incidence and severity of symptom appearance, indicating that the lower progression of the infection observed upon inoculation with the V3 null mutant is due to the lack of this protein, and not to a suboptimal performance of Rep-I89V. Next, we evaluated the virulence of TYLCV-mV3 on tomato, the virus' natural host. As previously observed in *N. benthamiana*, the lack of V3 resulted in lower viral load and CP accumulation, and milder symptoms at 10 dpi (Figure 4f-h), confirming that V3 plays a relevant role in the viral infection that is not restricted to *N. benthamiana*.

Heterologous expression from a potato virus X (PVX)-derived vector and quantification of the impact on PVX pathogenicity is a widely used approach to test virulence activity of viral genes of interest. Confirming a contribution of V3 to virulence, the presence of this gene in PVX-V3 led to an exacerbation of disease symptoms compared to PVX alone at 10 and 30 dpi, with a

concomitant higher accumulation of the PVX CP at 30 dpi (Supplementary figure 12); βC1, a symptom determinant encoded by tomato yellow leaf curl China betasatelite, was used as positive control. Infection by PVX-V3, however, did not lead to H_2O_2 accumulation or cell death (Supplementary figure 12c).

It has been previously established that a high correlation exists between the ability of a viral protein to suppress RNA silencing and its capacity to enhance the severity of the PVX infection. RNA silencing is conserved in eukaryotes, and is considered the main anti-viral defence mechanism in plants (Ding, 2010; Ding et al., 2004; Jin et al., 2020). Supporting this notion, virtually all plant viruses described to date encode at least one protein with RNA silencing suppression activity (Burgyan and Havelda, 2011; Jin et al., 2020). With the aim to test if V3 can suppress post-transcriptional gene silencing (PTGS), we transiently expressed GFP from a 35S promoter in leaves of transgenic 16c *N. benthamiana* plants, harbouring a 35S:GFP cassette (Ruiz et al., 1998), in the presence or absence of V3; the well-described silencing suppressor P19 from tomato bushy stunt virus was used as positive control. At 4 dpi, fluorescence had already substantially decreased when no viral protein was co-expressed, but was maintained in the samples with P19 or Myc-V3 (Figure 5a). Western blot and qRT-PCR were used to confirm that both the GFP protein as well as the corresponding mRNA accumulated to higher levels in tissues expressing P19 or Myc-V3 (Figure 5b). At 20 dpi, systemic leaves of 16c plants inoculated with the constructs to express either of the viral proteins remained green, while fluorescence had disappeared in control plants as a result of systemic silencing (Figure 5a).

The ability of V3 to suppress PTGS was further confirmed by expressing this protein from a PVX-based vector; in this case, βC1, which also functions as silencing suppressor, was used as positive control. Transient co-transformation of a PVX infectious clone together with a 35S:GFP cassette in leaves of 16c plants led to weak fluorescence in the infiltrated tissues at 7 dpi, and systemic silencing at 20 dpi; in stark contrast, co-transformation with PVX-βC1 or PVX-V3 resulted in the maintenance of strong fluorescent signal at 7 dpi, and absence of systemic silencing (Figure 5d). Neither βC1 nor V3 enhanced the local accumulation of PVX, as indicated by the accumulation of the PVX CP protein (Figure 5e), hence ruling out an indirect effect of these proteins on the endogenous ability of PVX to suppress silencing. Therefore, our results demonstrate that V3 from TYLCV can effectively suppress PTGS in plants.

Another level of RNA silencing is transcriptional gene silencing (TGS), which acts through methylation of DNA at cytosine residues; TGS also acts as an anti-viral response in plants, and is particularly relevant against geminiviruses, which replicate their DNA genomes in the nucleus

of the infected cell (Jin et al., 2020; Wang et al., 2019). To investigate whether V3 can also act as a TGS suppressor, we inoculated 16-TGS plants, in which the GFP transgene is silenced due to methylation of the 35S promoter (Yang et al., 2011), with PVX or PVX-V3; PVX-βC1 was used as positive control, since βC1 can also act as a TGS suppressor. At 10 dpi, green fluorescence could be observed in systemic leaves of 16-TGS plants inoculated with PVX-βC1 and PVX-V3, as opposed to mock- or PVX-inoculated plants (Figure 5f); this fluorescence persisted at 28 dpi (Figure 5g). Visual assessment was confirmed at the molecular level by western blot (Figure 5h, i), indicating that V3 can also suppress TGS in the host plant.

To confirm the effect of V3 on DNA methylation, the level of genome-wide methylation in transgenic V3-YFP plants was examined by digestion with a methylation-dependent restriction enzyme, *Mcr*BC (Stewart et al., 2000). As presented in Figure 5j, genomic DNA from two independent V3-YFP lines was completely digested by the methylation-independent restriction endonuclease *Dra*I, but only partially digested by *Mcr*BC, in sharp contrast to the genomic DNA from WT plants. This indicates a lower level of DNA methylation in transgenic plants expressing V3, further supporting a function of this viral protein as TGS suppressor.

In summary, our results demonstrate that V3 is a newly described protein encoded by TYLCV, which preponderantly localizes in the cis-Golgi, significantly contributes to virulence, and functions as a suppressor of both PTGS and TGS.

DISCUSSION

Geminivirus-encoded proteins have so far been identified taking into consideration an arbitrary threshold of 10 kDa, below which potential proteins were discarded. However, it is increasingly clear that small proteins and peptides play relevant roles in biological systems, hence calling for the reconsideration of this criterion. Here, we show that geminiviral genomes contain additional ORFs beyond those previously described, at least some of which are conserved within members of a given genus, hinting at functional relevance. Using TYLCV as a model, we demonstrate that at least some of these conserved novel ORFs are expressed during the infection, and that the proteins they encode localize in specific subcellular compartments. Interestingly, in this subset of selected proteins novel localizations, not described for any of the other TYLCV-encoded proteins, are represented, including the Golgi apparatus (ORF4, 5, and 6) and mitochondria (ORF 3). Also of note, some of these proteins (ORF2, 4, and 6) harbor transmembrane domains, which are not present in any of the “canonical” proteins from TYLCV. These results indicate that the repertoire

of geminiviral proteins can be expanded, and that the new additions to the viral proteomes will most likely perform additional virulence functions and/or employ alternative molecular mechanisms to those exhibited by the previously described proteins. We might still be, therefore, far from getting a comprehensive overview of the plant-geminivirus molecular interaction landscape, which will require the detailed study of potentially multiple small ORFs that have so far been neglected.

In this work, we selected the largest of the new ORFs found in TYLCV, which we name V3, to be used as a proof-of-concept of the potential functionality of novel viral small proteins. Strikingly, the lack of V3 negatively impacted the viral infection in two different hosts, *N. benthamiana* and tomato, demonstrating that V3 has a biological function. However, V3 is not essential, since a V3 null mutant virus can still accumulate and establish a systemic infection.

The V3 protein is mostly localized in the cis-Golgi, a novel localization for a geminivirus-encoded protein. Remarkably, V3 can suppress both PTGS and TGS, an ability that may underlie its virulence-promoting effect on TYLCV and PVX. One intriguing question is how V3 can exert this effect from the Golgi apparatus. Interestingly, connections have been drawn between the endomembrane system and RNA silencing (reviewed in Kim et al., 2014). In the model plant *Arabidopsis thaliana*, electron microscopy unveiled an enrichment of the Argonaute protein AGO1, a central player in PTGS, in close proximity to Golgi (Derrien et al., 2012); another Argonaute protein, AGO7, which has been recently shown to play a role in anti-viral defence (Zheng et al., 2019), also co-purifies with membranes and concentrates in cytoplasmic bodies linked to the ER/Golgi endomembrane system (Jouannet et al., 2012). It seems therefore plausible that this subcellular localization is permissive for a direct targeting of RNA silencing, although further experiments will be necessary to uncover the exact molecular mechanism underlying this activity of V3.

PTGS and TGS are arguably the main plant defence mechanisms against geminiviruses. This idea is supported by the fact that, despite limited coding capacity, a given geminivirus species can produce several proteins that independently target these processes. TYLCV encodes at least three proteins capable of acting as PTGS suppressors, namely C2, C4, and V2 (Luna et al., 2012; Rosas-Diaz et al., 2018; Zrachya et al., 2007), and at least two, Rep and V2, capable of suppressing TGS (Rodriguez-Negrete et al., 2013; Wang et al., 2014; Wang et al., 2018; Wang et al., 2020); these proteins exert their functions through non-overlapping mechanisms. Only one of the viral proteins, V2, has been described as a simultaneous suppressor of PTGS and TGS, as observed for V3. All of these silencing suppressors encoded by TYLCV are essential for the

infection, although this may be due to their contribution to additional virulence activities, enabled by their multifunctional nature. Similarly, it is possible that V3 exerts additional, yet-to-be described functions during the viral infection.

Why a given geminivirus species needs multiple proteins targeting the same pathway is a thought-provoking question. Since the viral infection is a process, the temporal dimension must be considered: geminiviral genes can be classified as early or late, depending on the timing of their expression, with the strongest PTGS and TGS suppressor, V2, being a late gene, probably due to the requirement of another viral protein, C2, to activate its expression. The V3 promoter was active in the absence of the infection, which suggests that it can be expressed as an early gene. The early expression of V3 would guarantee the availability of a PTGS and TGS suppressor during the time between the synthesis of the dsDNA replicative intermediate and the expression of V2 later in the cycle, enhancing the effectiveness of viral accumulation and spread. Nevertheless, further work will be required to acquire a full understanding of the potential breadth of functions exerted by V3 and of the underpinning molecular mechanisms.

METHODS

Plant materials

N. benthamiana and *Solanum lycopersicum* (tomato) plants were grown in a growth chamber with 60% relative humidity and a 16 h:8 h light:dark, 25°C:18°C regime. The transgenic GFP 16c line was kindly provided by David C. Baulcombe (University of Cambridge, UK) (Ruiz et al., 1998); 16-TGS plants were described previously (Buchmann et al., 2009); the transgenic RFP-H2B line was kindly shared by Michael M. Goodin (University of Kentucky, USA) (Martin et al., 2009).

The *Agrobacterium tumefaciens* strain EHA105 containing the pEarleygate101:V3-YFP construct was used to generate 35S:*V3-YFP* transgenic *N. benthamiana* lines by leaf disc transformation as previously described (Li et al., 2015).

Agroinfiltration and viral inoculation

For *A. tumefaciens*-mediated transient expression in *N. benthamiana*, plasmids were transformed into the EHA105 strain by the freeze-thaw method (Figures 3a-h lower panel, 4, 5, Supplementary figure 10b, c, Supplementary figure 12) or to the GV3101 strain through electroporation (Figures 2, 3h upper panel, Supplementary figure 8, Supplementary figure 9, Supplementary figure 10a, d). *Agrobacterium* cultures were resuspended in infiltration buffer (10 mM $MgCl_2$, 10 mM MES

(pH 5.6), and 100 μM acetosyringone) to an OD_{600} = 0.1-0.5, then infiltrated into the adaxial side of four-week-old *N. benthamiana* leaves with a needle-less syringe. For viral inoculation, two-week-old *N. benthamiana* plants or tomato plants at the two-leaf stage were infiltrated with *Agrobacterium* cultures carrying the TYLCV-BJ (MN432609) infectious clone. For experiments that required co-infiltration, *Agrobacterium* suspensions carrying different constructs were mixed at 1:1 ratio before infiltration.

Plasmid construction

Viral open reading frames (ORFs) from the TYLCV (TYLCV-Alm, Accession No. AJ489258) genome were cloned in the pENTR™/D-TOPO® vector (Thermo Scientific) (for ORF1, ORF2, ORF4, ORF5, ORF6/V3) or the pDONR™/Zeo vector (Thermo Scientific) (for ORF3) without a stop codon. The binary plasmids to express GFP-fused viral proteins were generated by sub-cloning (Gateway LR reaction, Thermo Scientific) the viral ORFs from the corresponding entry vectors into pGWB505 (Nakagawa et al., 2007). To generate the constructs to express V3-YFP, YFP-V3, or Myc-V3, the full-length V3 ORF was obtained from TYLCV (TYLCV-BJ, Accession No. MN432609) and recombined into the binary destination vectors pEarleygate101, pEarleygate104, or pEarleygate203, respectively (Earley et al., 2006). Please note that the V3/ORF6 protein sequences from TYLCV-Alm and TYLCV-BJ are identical.

To generate the construct to express SYP32-RFP (as a cis-Golgi marker), the gene encoding SYNTAXIN OF PLANTS 32 (SYP32) was amplified from *A. thaliana* cDNA, cloned into the pENTR™/D-TOPO® vector (Thermo Scientific), and sub-cloned into pGWB554 (Nakagawa et al., 2007) using a Gateway LR reaction (Thermo Scientific). The construct to express RFP-HDEL is from Liu et al., 2015, and the ones to express mCherry-HDEL and Man49-mCherry are from Nelson et al., 2007.

To generate the constructs used for the analysis of promoter activity, the 500-nt sequence upstream of the ORF1, ORF2, ORF4, ORF5, ORF6/V3, and C4 ATG, or the 570-nt sequence upstream of the ORF3 ATG, were PCR-amplified and cloned into the pDONR™/Zeo vector (Thermo Scientific). These sequences were then sub-cloned into pGWB504 (Nakagawa et al., 2007) by a Gateway LR reaction to generate pORF-GFP. In addition, the 833-nt sequence upstream of the V3/ORF6 ATG was PCR-amplified and cloned into pINT121-GUS digested with *Hind*III and *Bam*HI to generate pINT121-V3-GUS (pV3-GUS), or into pCHF3-GFP digested with *EcoR*I and *Sac*I to generate pCHF3-V3-GFP (pV3-GFP) using In-Fusion Cloning according to the manufacturer's instructions. The full-length V3 ORF was inserted into the PVX vector digested

with *Cla*I and *Sal*I to generate PVX-V3. The pCHF3-35S-GFP, pCHF3-p19, PVX-βC1, and a PVX-based expression vector for PTGS suppression assays have been described previously (Li et al., 2015; Xiong et al., 2009), as has the PVX-based expression vector PVX-βC1 for TGS suppression assays (Yang et al., 2011). All primers used in this study can be found in Supplementary table 1.

Sequence analysis

The ViralORFfinder platform was constructed by Shiny, an R package used to build interactive web applications (https://shiny.rstudio.com). The NCBI ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) was used to identify ORFs for each uploaded virus, and the R package Gviz (Hahne and Ivanek, 2016) was used for visualization; for Supplementary figures 1 and 3, 1.2-mer genomic sequences were used. To investigate the conservation of a ORF-encoded protein of choice, this tool creates a small database with the translated sequences from the inputted viral sequences, and BLASTp is used to identify proteins with high identity (e-value ≤ 0.05). Phylogenetic trees were obtained by the R package DECIPHER and visualized by ggtree. Multiple sequence alignments were constructed by ClustalW and visualized by the R package ggmsa (https://cran.r-project.org/web/packages/ggmsa/vignettes/ggmsa.html). Names and NCBI accession numbers of virus species used in this work are listed in Supplementary tables 2-6.

Prediction of domains or signals in protein sequences

The prediction of transmembrane domains (TM) was performed by TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and Phobius (https://phobius.sbc.su.se/). The prediction of nuclear localization signal (NSL) was performed by cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi; (Kosugi et al., 2009a; Kosugi et al., 2009b), The prediction of chloroplast transit peptide (cTP) was performed by ChloroP (http://www.cbs.dtu.dk/services/ChloroP/).

Confocal microscopy

Confocal microscopy was performed using a Leica TCS SP8 point scanning confocal microscope (Figure 2, 3h upper panel, Supplementary figures 8, 9, 10a, and 10d) or Zeiss LSM980 confocal microscope (Carl Zeiss) (Figure 3e, 3h lower panel, Supplementary figure 10b, c), with the preset settings for GFP (Ex: 488 nm, Em: 500-550 nm), RFP (Ex: 561 nm, Em: 570-620 nm), YFP (Ex :514 nm, Em: 515–570 nm), or mCherry (Ex: 594 nm, Em: 597–640 nm). For co-localization imaging, the sequential scanning mode was used.

413 **Mitochondrial staining**

414 To visualize mitochondria, staining with 250 nM MitoTracker® Red CMXRos (Invitrogen) was
415 used. The chemical was infiltrated 10-30 min before imaging. The stock solution (1 mM) was
416 prepared by dissolving the corresponding amount of MitoTracker® in dimethylsulfoxide (DMSO).
417 The working solution was prepared by diluting the stock solution in water or infiltration buffer (10
418 mM $MgCl_2$, 10 mM MES (pH 5.6), and 100 μM acetosyringone). The MitoTracker® red
419 fluorescence was imaged using a Leica TCS SP8 point scanning confocal microscope with the
420 following settings: Ex: 561 nm, Em: 570-620 nm.

421 **DNA and RNA extraction and qPCR/qRT-PCR**

422 Total DNA was extracted from infected plants using the CTAB method. Total RNA was extracted
423 from collected plant leaves using TaKaRa MiniBEST Universal RNA Extraction Kit (Takara,
424 Japan). 1 ug of total RNA was reverse-transcribed into cDNA using PrimeScript™ RT reagent Kit
425 with gDNA Eraser (Perfect Real Time) (Takara, Japan). qPCR or qRT-PCR was performed using
426 TB Green® Premix Ex Taq™ II (Takara, Japan). 25S RNA and *NbActin2* were used as internal
427 references for DNA and RNA normalization, respectively.

428 **5' rapid amplification of cDNA ends (RACE)**

429 Total RNA extracted from TYLCV (TYLCV-BJ, Accession No. MN432609)-infected *N.*
430 *benthamiana* plants was used for 5′ RACE with SMARTer RACE 5'/3' Kit (Takara, Japan)
431 according to the manual booklet.

432 **3,3′-diaminobenzidine (DAB) staining**

433 For DAB staining, systemic leaves of infected plants were incubated in DAB solution (1 mg/mL,
434 pH 3.8) for 10h at 25°C，then boiled for 5-10 min and decolorized in 95% ethanol.

435 **Protein extraction and western blotting**

436 Total protein was isolated from infiltrated leaf patches with protein extraction buffer (containing 50
437 mM Tris-HCl (pH 6.8), 4.5% (m/v) SDS, 7.5% (v/v) 2-Mercaptoethanol, 9 M carbamide).
438 Immunoblotting was performed with primary mouse polyclonal antibodies, followed by anti-mouse
439 IgG HRP-linked antibodies (1:5000; Cell Signaling Technology, USA); primary antibodies used
440 are as follows: anti-GFP (1:5000; ROCHE, USA), and custom-made anti-PVX CP (1:5000) and
441 anti-TYLCV CP (1:5000) (Wu and Zhou, 2005; Wu et al., 2012).

442

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (31930089), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB27040206), and the Shanghai Center for Plant Stress Biology from the Chinese Academy of Sciences. The authors thank all members of Rosa Lozano-Duran's lab and Alberto Macho's lab for fruitful discussions, Xinyu Jian, Aurora Luque, and the PSC Cell Biology Facility for technical assistance, Alberto Macho for critical reading of the manuscript, and Dr. Michael M. Goodin (University of Kentucky, USA) and Prof. David Baulcombe (University of Cambridge) for kindly sharing materials.

REFERENCES

Aguilar, E., Garnelo Gomez, B., and Lozano-Duran, R. (2020). Recent advances on the plant manipulation by geminiviruses. Current Opinion in Plant Biology *56*, 56-64.

Azzam, O., Frazer, J., de la Rosa, D., Beaver, J.S., Ahlquist, P., and Maxwell, D.P. (1994). Whitefly transmission and efficient ssDNA accumulation of bean golden mosaic geminivirus require functional coat protein. Virology *204*, 289-296.

Basak, J. (2016). Tomato yellow leaf curl virus: A serious threat to tomato plants world wide. Journal of Plant Pathollogy & Microbiology *7*, 4.

Brito, A.F., and Pinney, J.W. (2017). Protein-Protein Interactions in Virus-Host Systems. Frontiers in Microbiology *8*, 1557.

Burgyan, J., and Havelda, Z. (2011). Viral suppressors of RNA silencing. Trends in Plant Science *16*, 265-272.

Castillo, A.G., Collinet, D., Deret, S., Kashoggi, A., and Bejarano, E.R. (2003). Dual interaction of plant PCNA with geminivirus replication accessory protein (Ren) and viral replication protein (Rep). Virology *312*, 381-394.

Derrien, B., Baumberger, N., Schepetilnikov, M., Viotti, C., De Cillia, J., Ziegler-Graff, V., Isono, E., Schumacher, K., and Genschik, P. (2012). Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. Proceedings of the National Academy of Sciences of the United States of America *109*, 15942-15946.

Diaz-Pendon, J.A., Canizares, M.C., Moriones, E., Bejarano, E.R., Czosnek, H., and Navas-Castillo, J. (2010). Tomato yellow leaf curl viruses: menage a trois between the virus complex, the plant and the whitefly vector. Molecular Plant Pathology *11*, 441-450.

Ding, S.W. (2010). RNA-based antiviral immunity. Nature Reviews Immunology *10*, 632-644.

Ding, S.W., Li, H., Lu, R., Li, F., and Li, W.X. (2004). RNA silencing: a conserved antiviral immunity of plants and animals. Virus Research *102*, 109-115.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant Journal y *45*, 616-629.

Eguen, T., Straub, D., Graeff, M., and Wenkel, S. (2015). MicroProteins: small size-big impact. Trends in Plant Science *20*, 477-482.

Gotz, M., Popovski, S., Kollenberg, M., Gorovits, R., Brown, J.K., Cicero, J.M., Czosnek, H., Winter, S., and Ghanim, M. (2012). Implication of *Bemisia tabaci* heat shock protein 70 in begomovirus-whitefly interactions. Journal of Virology *86*, 13241-13252.

Hahne, F., and Ivanek, R. (2016). Visualizing genomic data using gviz and bioconductor. . Methods in Molecular Biology *1418, 335–351*.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D., and Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. Nature Reviews Microbiology *11*, 777-788.

Hsu, P.Y., and Benfey, P.N. (2018). Small but mighty: Functional peptides encoded by small ORFs in plants. Proteomics *18*, e1700038.

Jin, Y., Zhao, J.H., and Guo, H.S. (2020). Recent advances in understanding plant antiviral RNAi and viral suppressors of RNAi. Current Opinion in Virology *46*, 65-72.

Jouannet, V., Moreno, A.B., Elmayan, T., Vaucheret, H., Crespi, M.D., and Maizel, A. (2012). Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. EMBO Journal *31*, 1704-1713.

King, C.R., Zhang, A., Tessier, T.M., Gameiro, S.F., and Mymryk, J.S. (2018). Hacking the cell: Network intrusion and exploitation by adenovirus E1A. mBio *9*, e00390-18.

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H. (2009a). Six classes of nuclear localization signals specific to different binding grooves of importin alpha. Journal of Biological Chemistry *284*, 478-485.

Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009b). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proceedings of the National Academy of Sciences of the United States of America *106*, 10171-10176.

Kumar, R.V. (2019). Plant antiviral immunity against geminiviruses and viral counter-defense for survival. Frontiers in Microbiology *10*, 1460.

Kunik, T., Palanichelvam, K., Czosnek, H., Citovsky, V., and Gafni, Y. (1998). Nuclear import of the capsid protein of tomato yellow leaf curl virus (TYLCV) in plant and insect cells. Plant Journal *13*, 393-399.

Li, F., Xu, X., Huang, C., Gu, Z., Cao, L., Hu, T., Ding, M., Li, Z., and Zhou, X. (2015). The AC5 protein encoded by mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. New Phytologist *208*, 555-569.

Lin, W., Qiu, P., Jin, J., Liu, S., Ul Islam, S., Yang, J., Zhang, J., Kormelink, R., Du, Z., and Wu, Z. (2017). The cap snatching of segmented negative sense RNA viruses as a tool to map the transcription start sites of heterologous co-infecting viruses. Frontiers in Microbiology *8*, 2519.

Liu, X., Huang, W., Zhai, Z., Ye, T., Yang, C., and Lai, J. (2021). Protein modification: a critical modulator in the interaction between geminiviruses and host plants. Plant, Cell & Environment doi: 10.1111/pce.14008.

Liu, Y., Zhang, C., Wang, D., Su, W., Liu, L., Wang, M., and Li, J. (2015). EBS7 is a plant-specific component of a highly conserved endoplasmic reticulum-associated degradation system in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America *112*, 12205-12210.

Lozano-Duran, R., Rosas-Diaz, T., Gusmaroli, G., Luna, A.P., Taconnat, L., Deng, X.W., and Bejarano, E.R. (2011). Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in *Arabidopsis thaliana*. Plant Cell *23*, 1014-1032.

Luna, A.P., Morilla, G., Voinnet, O., and Bejarano, E.R. (2012). Functional analysis of gene-silencing suppressors from tomato yellow leaf curl disease viruses. Molecular Plant-Microbe Interactions *25*, 1294-1306.

Martin, K., Kopperud, K., Chakrabarty, R., Banerjee, R., Brooks, R., and Goodin, M.M. (2009). Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. Plant Journal *59*, 150-162.

Medina-Puche, L., Tan, H., Dogra, V., Wu, M., Rosas-Diaz, T., Wang, L., Ding, X., Zhang, D., Fu, X., Kim, C., *et al.* (2020). A defense pathway linking plasma membrane and chloroplasts and co-opted by pathogens. Cell *182*, 1109-1124 e25.

Murphy, E., Smith, S., and De Smet, I. (2012). Small signaling peptides in *Arabidopsis* development: how cells communicate over a short distance. Plant Cell *24*, 3198-3217.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., *et al.* (2007). Improved gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. Bioscience, Biotechnology, and Biochemistry *71*, 2095-2100.

Nelson, B.K., Cai, X., and Nebenfuhr, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. Plant Journal *51*, 1126-1136.

Ohnesorge, S., and Bejarano, E.R. (2009). Begomovirus coat protein interacts with a small heat-shock protein of its transmission vector (*Bemisia tabaci*). Insect Molecular Biology *18*, 693-703.

Palanichelvam, K., Kunik, T., Citovsky, V., and Gafni, Y. (1998). The capsid protein of tomato yellow leaf curl virus binds cooperatively to single-stranded DNA. Journal of General Virology *79*, 2829-2833.

Rodriguez-Negrete, E., Lozano-Duran, R., Piedra-Aguilera, A., Cruzado, L., Bejarano, E.R., and Castillo, A.G. (2013). Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytologist *199*, 464-475.

Rojas, M.R., Jiang, H., Salati, R., Xoconostle-Cazares, B., Sudarshana, M.R., Lucas, W.J., and Gilbertson, R.L. (2001). Functional analysis of proteins involved in movement of the monopartite begomovirus, tomato yellow leaf curl virus. Virology *291*, 110-125.

Rosas-Diaz, T., Macho, A.P., Beuzon, C.R., Lozano-Duran, R., and Bejarano, E.R. (2016). The C2 protein from the geminivirus tomato yellow leaf curl Sardinia virus decreases sensitivity to jasmonates and suppresses jasmonate-mediated defences. Plants *5*, 8.

Rosas-Diaz, T., Zhang, D., Fan, P., Wang, L., Ding, X., Jiang, Y., Jimenez-Gongora, T., Medina-Puche, L., Zhao, X., Feng, Z., *et al.* (2018). A virus-targeted plant receptor-like kinase promotes cell-to-cell spread of RNAi. Proceedings of the National Academy of Sciences of the United States of America *115*, 1388-1393.

Rubinstein, G., and Czosnek, H. (1997). Long-term association of tomato yellow leaf curl virus with its whitefly vector Bemisia tabaci: effect on the insect transmission capacity, longevity and fecundity. Journal of General Virology *78*, 2683-2689.

Ruiz, M.T., Voinnet, O., and Baulcombe, D.C. (1998). Initiation and maintenance of virus-induced gene silencing. Plant Cell *10*, 937-946.

Settlage, S.B., See, R.G., and Hanley-Bowdoin, L. (2005). Geminivirus C3 protein: replication enhancement and protein interactions. Journal of Virology *79*, 9885-9895.

Stewart, F.J., Panne, D., Bickle, T.A., and Raleigh, E.A. (2000). Methyl-specific DNA binding by McrBC, a modification-dependent restriction enzyme. Journal of Molecular Biology *298*, 611-622.

Wang, B., Li, F., Huang, C., Yang, X., Qian, Y., Xie, Y., and Zhou, X. (2014). V2 of tomato yellow leaf curl virus can suppress methylation-mediated transcriptional gene silencing in plants. Journal of General Virology *95*, 225-230.

Wang, B., Yang, X., Wang, Y., Xie, Y., and Zhou, X. (2018). Tomato yellow leaf curl virus V2 interacts with host histone deacetylase 6 to suppress methylation-mediated transcriptional gene silencing in plants. Journal of Virology *92*, e00036-18.

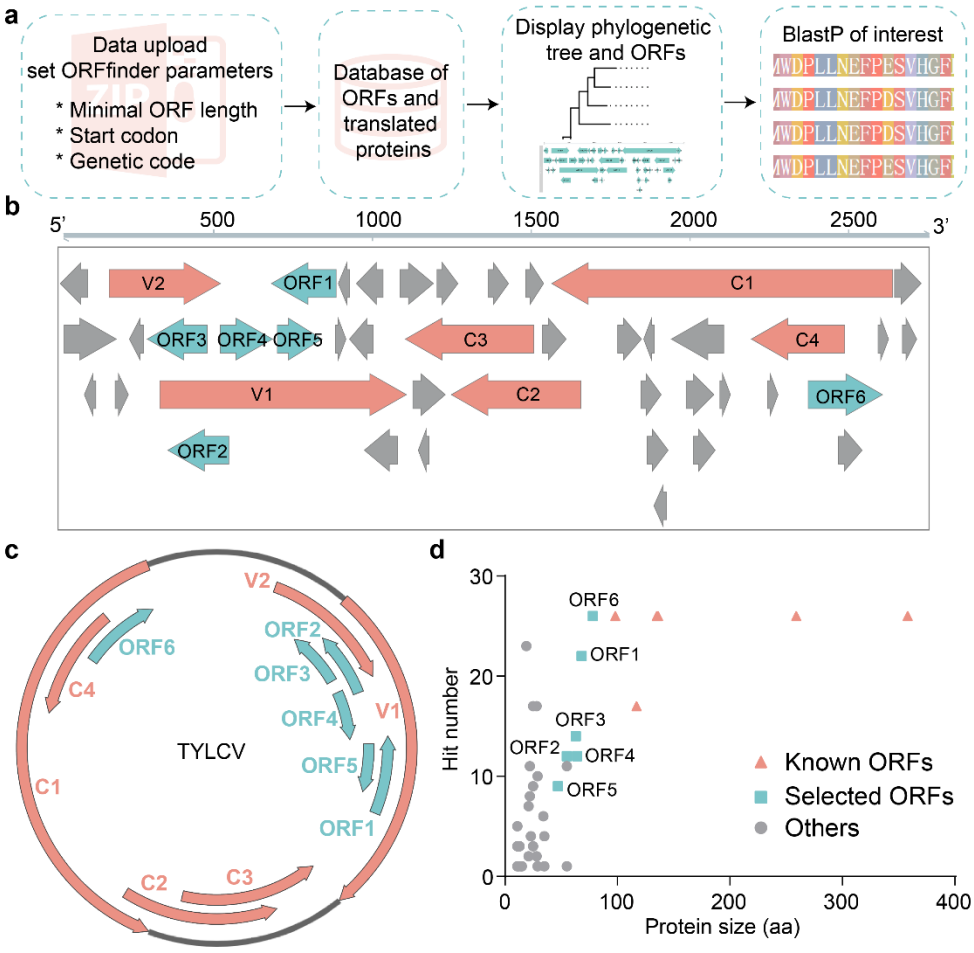
Wang, C., Wang, C., Zou, J., Yang, Y., Li, Z., and Zhu, S. (2019). Epigenetics in the plant-virus interaction. Plant Cell Reports *38*, 1031-1038.

Wang, L., Ding, Y., He, L., Zhang, G., Zhu, J.K., and Lozano-Duran, R. (2020). A virus-encoded protein suppresses methylation of the viral genome through its interaction with AGO4 in the Cajal body. eLife *9*, e55542.

588 Wu, J., and Zhou, X. (2005). Production and application of monoclonal antibodies against
589 potatovirus X. Journal of Zhejiang University B *31*, *608-612*.
590 Wu, J.X., Shang, H.L., Xie, Y., and Zhou, X.P. (2012). Monoclonal antibodies against the whitefly-
591 transmitted tomato yellow leaf curl virus and their application in virus detection. Journal of
592 Integrative Agriculture *11*, *263–268.*
593 Wu, M., W, H., Tan, H., Pan, S., Liu, Q., Bejarano, E.R., and Lozano-Durán, R. (2020). Plant DNA
594 polymerases alpha and delta mediate replication of geminiviruses. bioRvix
595 *https://doi.org/10.1101/2020.07.20.212167*
596 Yang, X., Xie, Y., Raja, P., Li, S., Wolf, J.N., Shen, Q., Bisaro, D.M., and Zhou, X. (2011).
597 Suppression of methylation-mediated transcriptional gene silencing by betaC1-SAHH
598 protein interaction during geminivirus-betasatellite infection. PLoS Pathogens *7*, e1002329.
599 Zerbini, F.M., Briddon, R.W., Idris, A., Martin, D.P., Moriones, E., Navas-Castillo, J., Rivera-
600 Bustamante, R., Roumagnac, P., Varsani, A., and Ictv Report, C. (2017). ICTV virus
601 taxonomypProfile: *Geminiviridae*. Journal of General Virology *98*, 131-133.
602 Zhao, W., Wu, S., Barton, E., Fan, Y., Ji, Y., Wang, X., and Zhou, Y. (2020). Tomato yellow leaf
603 curl virus V2 protein plays a critical role in the nuclear export of V1 protein and viral systemic
604 infection. Frontiers in Microbiology *11*, 1243.
605 Zheng, X., Fahlgren, N., Abbasi, A., Berry, J.C., and Carrington, J.C. (2019). Antiviral
606 ARGONAUTEs against turnip crinkle virus revealed by image-based trait analysis. Plant
607 Physiology *180*, 1418-1435.
608 Zrachya, A., Glick, E., Levy, Y., Arazi, T., Citovsky, V., and Gafni, Y. (2007). Suppressor of RNA
609 silencing encoded by tomato yellow leaf curl virus-Israel. Virology *358*, 159-165.
610

611

612 **FIGURES AND FIGURE LEGENDS**



613

614 **Figure 1. The TYLCV genome contains additional conserved open reading frames.** a.
615 Working pipeline of ViralORFfinder, a web-based tool for ORF prediction and protein conservation
616 analysis. b. Schematic view of predicted ORFs (≥ 30 nt) in the TYLCV genome. c. Genome
617 organization of TYLCV; arrows indicate ORFs. In b and c, pink arrows represent the six known
618 ORFs (C1, C2, C3, C4, V1, and V2), while blue arrows represent the six new ORFs described in
619 this work (ORF1-6). d. Correlation between the size of proteins encoded by the ORFs in the
620 TYLCV genome and their representation in the selected subset of begomoviruses (see
621 Supplementary table 6). The TYLCV isolate used in these experiments is TYLCV-Alm.

622

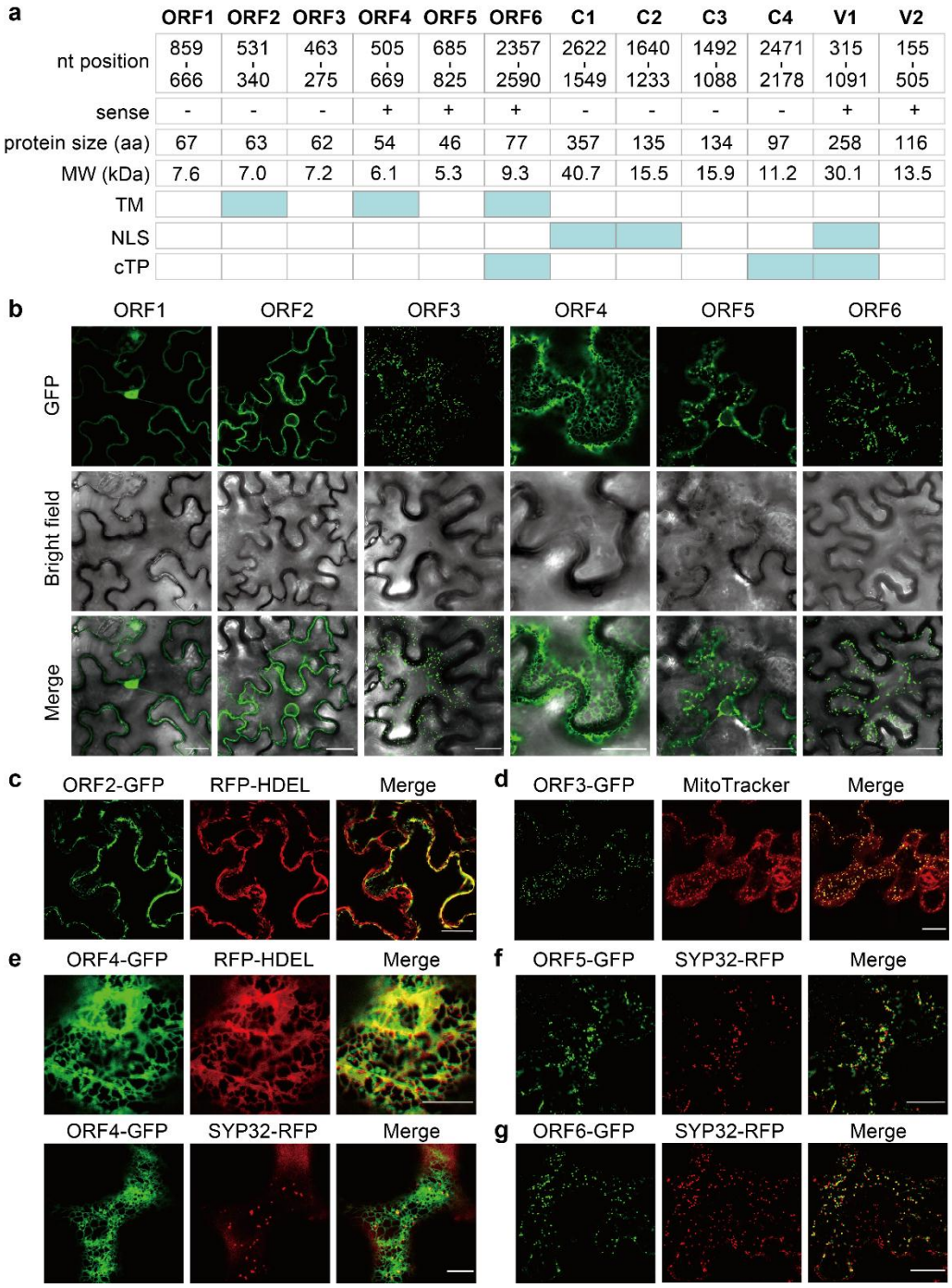


Figure 2. The novel open reading frames in the TYLCV genome encode proteins with predicted domains and specific subcellular localizations. a. Nucleotide position of the six novel ORFs and the six known ORFs in the TYLCV genome, size (in aa) and predicted molecular weight (MW; in kDa) of the corresponding encoded proteins, and domains or signals predicted in

629 the proteins sequence. TM: transmembrane domain; NLS: nuclear localization signal; cTP:
630 chloroplast transit peptide. b. Subcellular localization of the proteins encoded by ORF1-6 fused
631 to GFP at their C-terminus transiently expressed in *N. benthamiana* leaves. Scale bar: 25 μm. c.
632 Co-localization of ORF2-GFP with the ER marker RFP-HDEL. Scale bar: 25 μm. d. Co-localization
633 of ORF3-GFP with the mitochondrial stain MitoTracker Red. Scale bar: 25 μm. e. Co-localization
634 of ORF4-GFP with the ER marker RFP-HDEL and the cis-Golgi marker SYP32-RFP. Scale bar:
635 10 μm. f. Co-localization of ORF5-GFP with the cis-Golgi marker SYP32-RFP. Scale bar: 25 μm.
636 g. Co-localization of ORF6-GFP with the cis-Golgi marker SYP32-RFP. Scale bar: 25 μm. b-g.
637 These experiments were repeated at least three times with similar results; representative images
638 are shown. The TYLCV isolate used in these experiments is TYLCV-Alm.

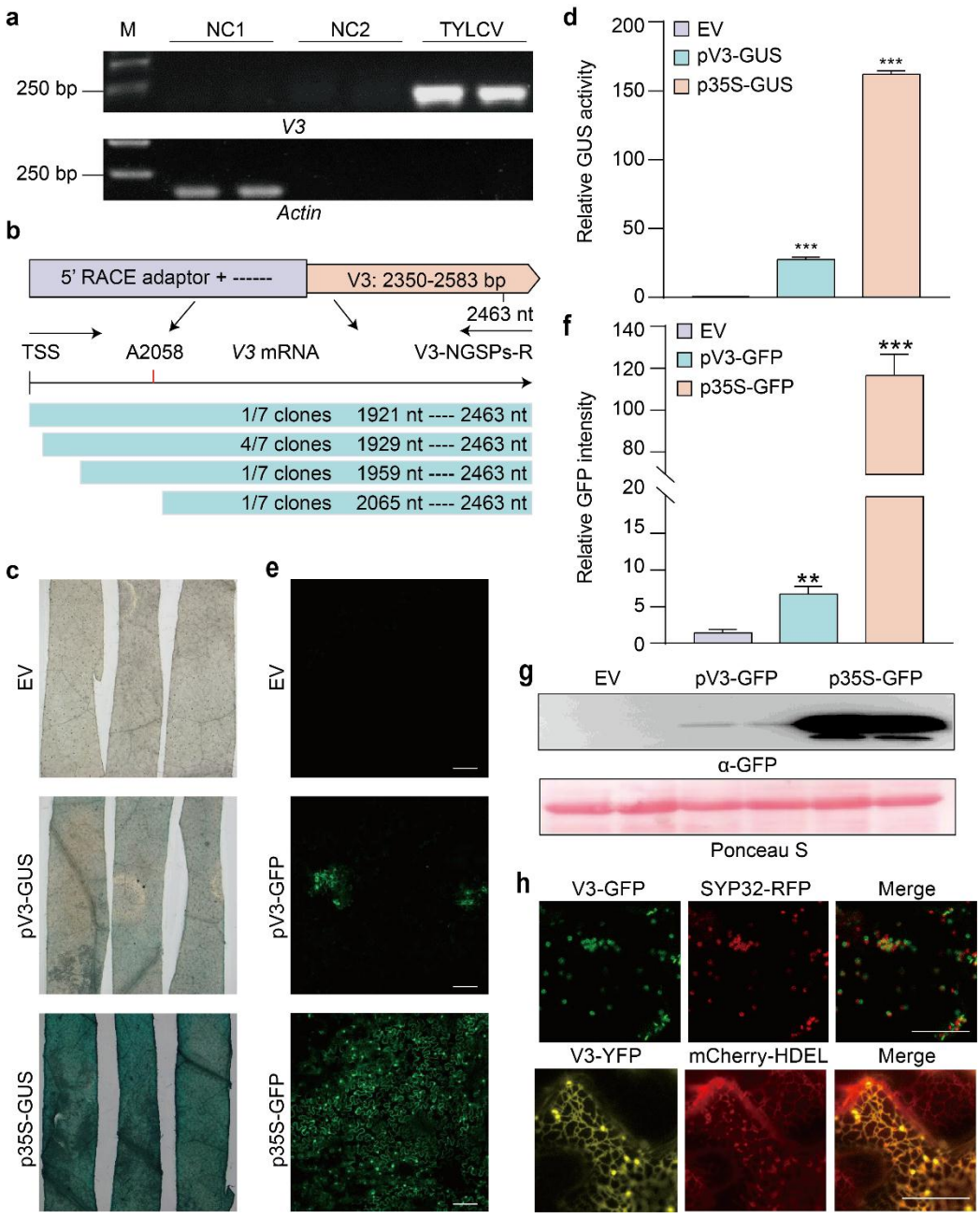


Figure 3. ORF6/V3 is expressed during the infection and encodes a Golgi-localized protein.

a. RT-PCR analysis of V3 transcripts from TYLCV-infected or uninfected *N. benthamiana* plants. M: DNA ladder marker. NC1: negative control 1 (reverse-transcription of total RNA extracted from uninfected plants with RT Primers). NC2: negative control 2 (reverse-transcription of total RNA extracted from uninfected plants with V3-specific primers). TYLCV: reverse-transcription of total RNA extracted from TYLCV-infected plants with V3-specific primers. b. Transcriptional start site

analysis of TYLCV V3 by 5' RACE. TSS: transcription start site. A2058: V3 TSS captured by RSV cap-snatching (Lin et al., 2017). c. Activity of pV3 promoter (and p35S promoter as positive control) in promoter-GUS fusions in transiently transformed *N. benthamiana* leaves at 2 dpi. EV: empty vector. d. Quantification of relative GUS activity in samples from (c). Error bars represent SD of n=3. Asterisks indicate a statistically significant difference according to Student's t-test, *** $p < 0.001$. e. Activity of pV3 promoter (and p35S promoter as positive control) in promoter-GFP fusions in transiently transformed *N. benthamiana* leaves at 2 dpi. Scale bar: 100 μm. EV: empty vector. f. Quantification of relative GFP intensity in samples from (e). Error bars represent SD of n=3. Asterisks indicate a statistically significant difference according to Student's t-test, ** $p < 0.01$, *** $p < 0.001$. g. Western blot analysis of GFP protein from (e). Ponceau S staining of the large RuBisCO subunit serves as loading control. EV: empty vector. h. Co-localization of V3-GFP with the cis-Golgi maker SYP32-RFP (upper panel) and co-localization of V3-YFP with the ER maker mCherry-HDEL (lower panel). Scale bar: 20 μm. The TYLCV isolate used in these experiments is TYLCV-BJ.

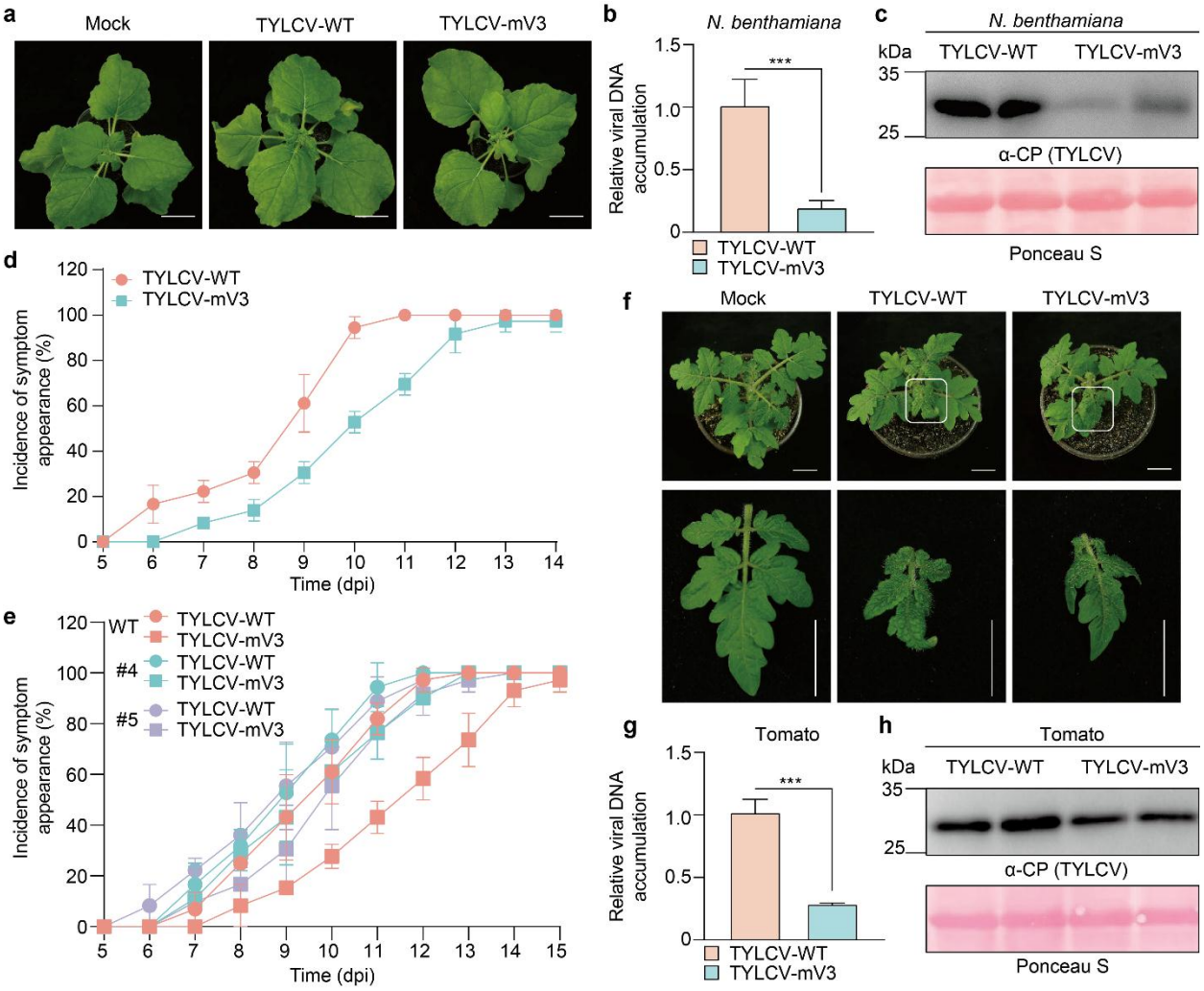
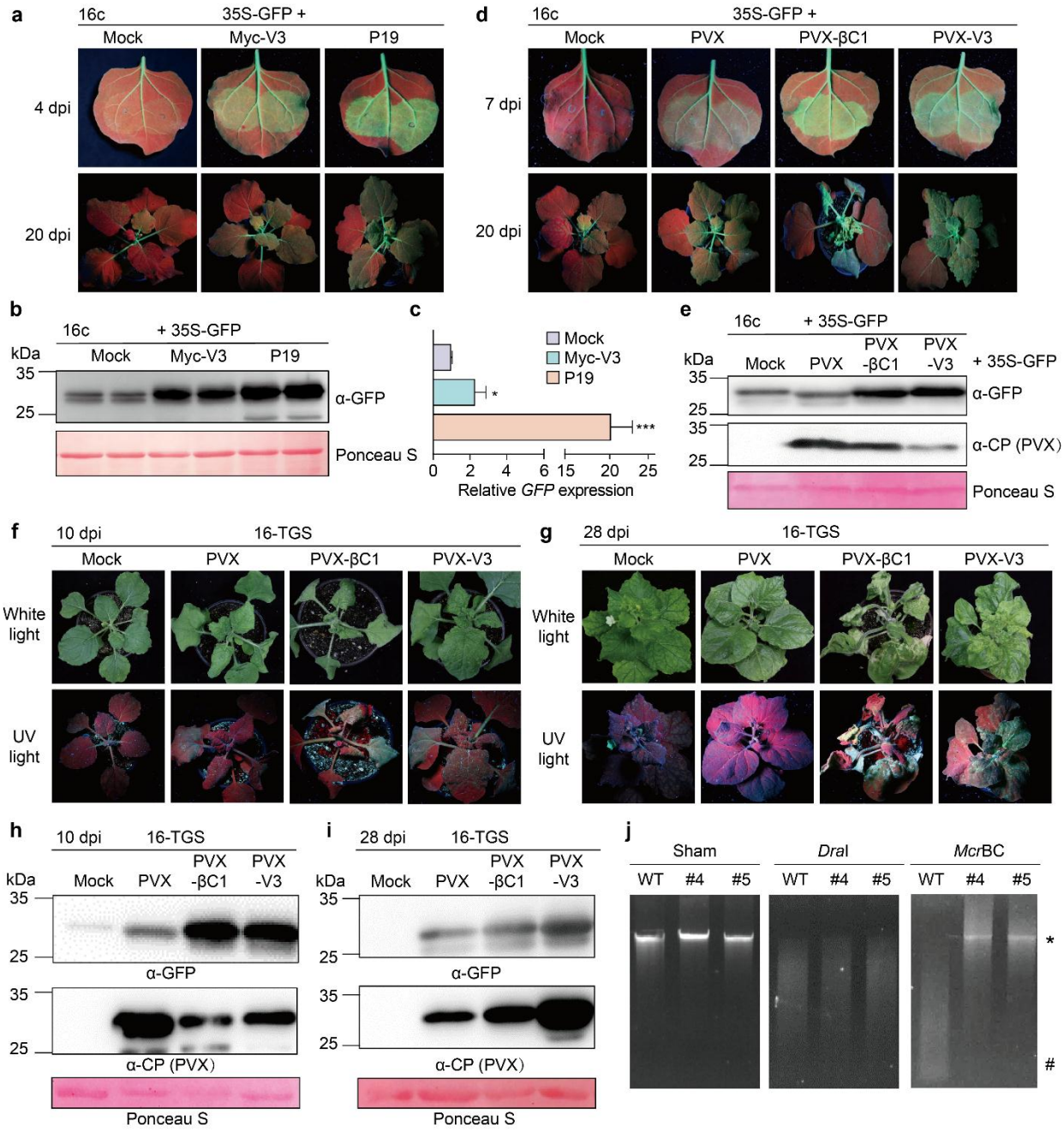


Figure 4. V3 is required for full TYLCV infection in *N. bentamiana* and tomato. a. Symptoms of *N. benthamiana* plants inoculated with wild-type TYLCV (TYLCV-WT), a V3 null mutant TYLCV (TYLCV-mV3), or mock-inoculated (pCAMBIA2300 empty vector), at 10 dpi. Bar = 2 cm. b. Viral DNA accumulation in TYLCV-WT- and TYLCV-mV3-infected plants in (a), measured by qPCR. Error bars represent means ± SD of n=3. Asterisks indicate a statistically significant difference according to Student's t-test, *** p<0.001. 25S RNA was used as internal reference. c. Western blot showing TYLCV CP accumulation in systemic leaves of TYLCV-WT- and TYLCV-mV3-infected plants from (a). Ponceau S staining of the large RuBisCO subunit serves as loading control. d, e. Incidence of symptom appearance in WT (d, e) or V3 transgenic (e) *N. benthamiana* plants infected with TYLCV-WT or TYLCV-mV3. Error bars represent means ± SD from three independent experiments; at least 8 plants were used per viral genotype and experiment. For

images of symptoms in the V3 transgenic lines, see Supplementary figure 10d. f. Symptoms of tomato plants inoculated with WT TYLCV (TYLCV-WT), a V3 null mutant TYLCV (TYLCV-mV3), or mock-inoculated (pCAMBIA2300 empty vector), at 10 dpi. Bar = 2 cm. g. Viral DNA accumulation in TYLCV-WT- and TYLCV-mV3-infected plants from (f), measured by qPCR. Error bars represent means ± SD of n=3. Asterisks indicate a statistically significant difference according to Student's t-test, *** p<0.001. 25S RNA was used as internal reference. h. Western blot showing TYLCV CP accumulation in systemic leaves of TYLCV-WT- and TYLCV-mV3-infected plants from (f). Ponceau S staining of the large RuBisCO subunit serves as loading control. The TYLCV isolate used in these experiments is TYLCV-BJ.



683

684 **Figure 5. V3 functions as a suppressor of PTGS and TGS.** a. Transgenic 16c *N. benthamiana*
685 plants co-infiltrated with constructs to express GFP (35S-GFP) and Myc-V3, P19 (as positive
686 control), or mock (empty vector, as negative control) at 4 dpi (upper panel) or 20 dpi (lower panel)
687 under UV light. b. Western blot showing the GFP accumulation in inoculated leaves from (a) at 4
688 dpi. The corresponding Ponceau S staining of the large RuBisCO subunit serves as a loading
689 control. c. Relative GFP mRNA accumulation in inoculated leaves from (a) at 4 dpi measured by
690 qRT-PCR. Error bars represent means ± SD of n=3. Asterisks indicate a statistically significant

691 difference according to Student's t-test, * p<0.5, *** p<0.001. *NbActin2* was used as the internal
692 reference. d. Transgenic 16c *N. benthamiana* plants co-infiltrated with constructs to express GFP
693 (35S-GFP) and PVX, PVX-V3, PVX-βC1 (as positive control), or mock (infiltration buffer, as
694 negative control) at 7 dpi (upper panel) or 20 dpi (lower panel) under UV light. e. Western blot
695 showing the GFP accumulation in inoculated leaves from (d) at 7 dpi. The corresponding Ponceau
696 S staining of the large RuBisCO subunit serves as a loading control. f, g. Symptoms of 16-TGS
697 *N*. *benthamiana* plants infected with PVX, PVX-V3, PVX-βC1 (as positive control), or mock-
698 inoculated under white light or UV light at 10 dpi (f) and 28 dpi (g). h. i. Western blot showing
699 accumulation of GFP and PVX CP in systemically infected leaves from (f) and (i). The
700 corresponding Ponceau S staining of the large RuBisCO subunit serves as loading control. j. DNA
701 methylation analysis by restriction enzyme digestion in V3-YFP transgenic *N. benthamiana* plants.
702 Genomic DNA extracted from WT *N. benthamiana* or two independent V3-YFP transgenic lines
703 (#4 and #5) was digested with the methylation-dependent restriction enzyme *Mcr*BC and the
704 methylation-insensitive enzyme *Dra*I. 'Sham' indicates a mock digestion with no enzyme added.
705 The positions of undigested input genomic DNA is indicated with an asterisk; the position of the
706 *Mcr*BC-digested products is indicated with a hashtag.

707

708

SUPPLEMENTARY MATERIAL

- Supplementary figures 1-12
- Supplementary tables 1-6