AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENOME REPLICATION AND REGULATION OF VIRAL GENE EXPRESSION

Short 5′ Untranslated Region Enables Optimal Translation of Plant Virus Tricistronic RNA via Leaky Scanning

Yuji Fujimoto,[a] Takuya Keima,[a] Masayoshi Hashimoto,[a] Yuka Hagiwara-Komoda,[b] Naoi Hosoe,[a] Shuko Nishida,[a] Takamichi Nijo,[a] Kenro Oshima,[c] Jeanmarie Verchot,[d] Shigetou Namba,[a] Yasuyuki Yamaji[a]

[a]Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan
[b]Department of Sustainable Agriculture, College of Agriculture, Food and Environment Sciences, Rakuno Gakuen University, Hokkaido, Japan
[c]Faculty of Bioscience, Hosei University, Tokyo, Japan
[d]Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas, USA

ABSTRACT Regardless of the general model of translation in eukaryotic cells, a number of studies suggested that many mRNAs encode multiple proteins. Leaky scanning, which supplies ribosomes to downstream open reading frames (ORFs) by readthrough of upstream ORFs, has great potential to translate polycistronic mRNAs. However, the mRNA elements controlling leaky scanning and their biological relevance have rarely been elucidated, with exceptions such as the Kozak sequence. Here, we have analyzed the strategy of a plant RNA virus to translate three movement proteins from a single RNA molecule through leaky scanning. The *in planta* and *in vitro* results indicate thatthe significantly shorter 5′ untranslated region (UTR) of the most upstream ORF promotes leaky scanning, potentially fine-tuning the translation efficiency of the three proteins in a single RNA molecule to optimize viral propagation. Our results suggest that the remarkably short length of the leader sequence, like the Kozak sequence, is a translational regulatory element with a biologically important role, as previous studies have shown biochemically.

IMPORTANCE *Potexvirus*, a group of plant viruses, infect a variety of crops, including cultivated crops. It has been thought that the three transition proteins that are essential for the cell-to-cell transfer of potexviruses are translated from two subgenomic RNAs, sgRNA1 and sgRNA2. However, sgRNA2 has not been clearly detected. In this study, we have shown that sgRNA1, but not sgRNA2, is the major translation template for the three movement proteins. In addition, we determined the transcription start site of sgRNA1 in flexiviruses and found that the efficiency of leaky scanning caused by the short 5′ UTR of sgRNA1, a widely conserved feature, regulates the translation of the three movement proteins. When we tested the infection of viruses with mutations introduced into the length of the 5′ UTR, we found that the movement efficiency of the virus was affected. Our results provide important additional information on the protein translation strategy of flexiviruses, including *Potexvirus*, and provide a basis for research on their control as well as the need to reevaluate the short 5′ UTR as a translational regulatory element with an important role *in vivo*.

KEYWORDS UTR, leaky scanning, plant virus, potexvirus, translation initiation

According to the general model of translation in eukaryotic cells, the translation machinery recognizes a single open reading frame (ORF) in a monocistronic mRNA. This recognition process has three steps: initiation, elongation, and termination. The first step, translation initiation, is a highly regulated process dependent on the 5′ cap structure of mRNA (1). The 5′ cap structure is recognized by the cap-binding protein eukaryotic translation initiation factor 4E (eIF4E), followed by recruitment of the 43S preinitiation complex (PIC) to the mRNA. The PIC then scans the mRNA downstream from the 5′ end until the first appropriate initiation codon is recognized (2).

Editor Anne E. Simon, University of Maryland, College Park

Address correspondence to Yasuyuki Yamaji, ayyamaji@g.ecc.u-tokyo.ac.jp.

The authors declare no conflict of interest.

Received 20 December 2021
Accepted 31 January 2022
Published 9 March 2022

Recognition of the initiation codon induces recruitment of the 60S ribosomal subunit to form the 80S ribosome complex, which transitions to polypeptide synthesis, referred to as the elongation step. Finally, in the termination step, the ribosomes and synthesized proteins are released from the mRNA after stopping at the termination codon.

Increasing experimental evidence has demonstrated that many eukaryotic mRNAs have a polycistronic structure: a single RNA molecule encoding multiple proteins. In fact, 44% of human mRNAs contain upstream ORFs (uORFs) located upstream of the start codons of annotated ORFs (3). Moreover, the start codon of uORFs is more conserved among various mammalian species than would be estimated for neutral evolution, suggesting that these uORFs play a particular biological role (4, 5). Similar to animals, approximately 35% of mRNAs in the model plant *Arabidopsis thaliana* contain uORF, and half of them have multiple uORFs (6). Translation from polycistronic mRNA is often accomplished by bypassing an initiation codon, referred to as leaky scanning. In leaky scanning, the PIC does not recognize the upstream AUG (uAUG) and reads past it, arriving at a downstream AUG (dAUG) to start the translation of the downstream ORF (dORF). Because the dAUG competes with the uAUG for the entry of PIC, increase of translation from the uAUG decreases translation from the dAUG, whereas decrease of translation from the uAUG increases translation from the dAUG (3, 6).

The efficiency of AUG codon recognition by PIC depends on the nucleotide sequence context near the AUG codon, especially the sequence from the −3 position to +4 position to the A of AUG, which is called the Kozak sequence and is known to have a significant effect on initiation codon recognition (7). In dicots, the optimal Kozak sequence is RNN AUG G (R, guanine [G] or adenine [A]) (8). The context of the uAUG is an important regulatory factor in dORF expression, because high efficiency of uAUG recognition promoted by a strong translation initiation sequence context disturbs the translation initiation from the dAUG. In fact, mutations in the translation initiation context of a uORF can result in serious physiological disorders at the organismal level, such as tumorigenesis in humans caused by natural variants in the Kozak sequence of a single gene (9). However, the regulatory mechanism of leaky scanning in eukaryotic cells is not fully understood.

Viruses are intracellular parasites that rely heavily on host machinery for their propagation. To synthesize viral proteins, viruses retain a sophisticated mechanism to exploit the host translation system (10 to 13). Presumably due to their limited genome size, RNA virus genomes are frequently polycistronic and are expressed by multiple strategies. The genus *Potexvirus* includes a group of monopartite positive-sense, single-stranded RNA viruses that infect a wide range of plant species. The potexviral genome with a 5′ cap and 3′ poly(A) tail possesses five ORFs encoding RNA-dependent RNA polymerase (RdRp); triple gene block proteins (TGBp) 1, 2, and 3; and coat protein (CP) (14, 15). TGBp1/2/3, which function as movement proteins (MPs) required for intracellular and intercellular movement of viruses in plant cells, are translated from three partially overlapping ORFs. Both the gene structure and the functions of TGBps are conserved in multiple genera of plant viruses other than potexviruses. Translation of these three proteins is considered to require two subgenomic RNAs (16), but the detailed mechanism of this translation system is not fully understood.

Here, we investigated the translation mechanism of TGBps and found that the three proteins are all translated from a single RNA molecule, sgRNA1, via leaky scanning. Through *in vitro* and *in planta* analyses of the mechanism that regulates the translation of TGBps from sgRNA1, we showed that, in addition to the Kozak sequence, the length of the uORF 5′ untranslated region (UTR) regulates leaky scanning. Thus, we propose to reevaluate the model named leaky scanning induced by a short 5′ UTR (LISH) as a mechanism for uORF-mediated translation control.

RESULTS

All three TGB proteins are translated mainly from sgRNA1. Although TGBp2 and TGBp3 of potexviruses have been suggested to be translated from sgRNA2 (16), we did not find any clear signal corresponding to deduced sgRNA2 in Northern blot analysis of *Nicotiana benthamiana* or *Arabidopsis thaliana* plants infected with the potexvirus

plantago asiatica mosaic virus (PlAMV) (17) (Fig. 1d, wild-type [WT] lane). Previous attempts by other research groups did not detect any clear putative sgRNA2 signal for other potexviruses (18, 19). To determine the transcription start sites (TSSs) of sgRNAs presumably transcribed from PlAMV genomic RNA, we used the 5′-rapid amplification of cDNA ends (5′-RACE) method with two gene-specific primers. The TSSs of sgRNA1 and sgRNA3 were successfully mapped to two major sites, G4224 and G5339, respectively, where the 5′ ends of all the cloned transcripts aligned (Fig. 1a). The same sites can be predicted to be TSSs based on consensus promoter sequences in viral genomic RNA (20). In contrast, no potential TSS was predicted for sgRNA2. The 5′ ends of cloned transcripts starting between the initiation codons of TGBp1 and TGBp2, which likely corresponded to the TSSs of sgRNA2, were consistently variable (Fig. 1b). These results indicated that a major TSS could not be defined for sgRNA2.

We hypothesized that the absence of sgRNA2 signals in Northern blot analysis may be due in part to lower sgRNA2 accumulation. We reasoned that another viral RNA, most likely sgRNA1, participates in the translation of TGBp2/3. We prepared an sgRNA1-deleted PlAMV mutant (Δsg1) in which several nucleotide substitutions were introduced into the sgRNA1 promoter sequence by following the procedure in a previous report (20) (Fig. 1c). To test whether the Δsg1 mutant produced TGBp1, TGBp2, and TGBp3, we inoculated the Δsg1 mutant into *N. benthamiana* leaves using the agroinoculation method. Total RNA extracted from the inoculated leaves at 1.5 days postinoculation (dpi) was analyzed using Northern blot hybridization. We confirmed that genomic RNA (gRNA) and sgRNA3 accumulated in Δsg1-inoculated leaves, but sgRNA1 did not (Fig. 1a). Immunoblot analysis revealed that TGBp2 and TGBp3, as well as TGBp1, were not detected in the Δsg1-inoculated leaves in which RdRp and CP accumulated to the same levels as in the WT (Fig. 1d). These results suggest that sgRNA1 is responsible for the translation of TGBp2 and -3 as well as TGBp1.

To further confirm the role of sgRNA1 in the production of TGBp2 and TGBp3 proteins, we examined whether TGBp2/3 were translated in sgRNA1-transfected protoplasts (Fig. 1e). Protoplasts isolated from *Arabidopsis* suspension culture cells were transfected with a plasmid expressing PlAMV sgRNA1 under the 35S promoter (sg1). Total protein extracted from the transfected protoplasts at 42 h was separated into soluble and insoluble fractions (S30 and P30, respectively). In sgRNA1-expressing protoplasts, TGBp1 was detected in both fractions, and both TGBp2/3 were specifically detected in the P30 fraction. This result together with the above-described findings indicates that sgRNA1 is a major template for the translation of TGBp2/3 in PlAMV.

To analyze whether the polycistronic nature of sgRNA1 is conserved in other groups of plant viruses with genomic structures similar to that of PlAMV, we examined the accumulation of TGBp1 and TGBp2 from sgRNA1 deletion mutants of viruses in the *Alphaflexiviridae* and *Betaflexiviridae* families. We introduced mutations into the sgRNA1 promoter sequences of three virus genus type species, potato virus X (PVX; genus *Potexvirus*) and lolium latent virus (LoLV; genus *Lolavirus*) in the family *Alphaflexiviridae* and potato virus M (PVM; genus *Carlavirus*) in the family *Betaflexiviridae*. Total RNA and protein were extracted from leaves inoculated with each Δsg1 mutant. We confirmed that sgRNA1 was undetectable in all Δsg1 mutants, and there was no significant difference in the accumulation of gRNA, sgRNA3, or CP between the WT and Δsg1 mutants of all viruses tested, showing that these mutations impaired sgRNA1 expression but not viral multiplication (Fig. 1f). Consistent with the above-described findings for PlAMV, TGBp1 and TGBp2 were not detected via immunoblotting in the Δsg1 mutants of all viruses. These results indicate that the role of sgRNA1 in translating the TGBp1/2/3 proteins is conserved among several distinct virus species in *Alphaflexiviridae* and *Betaflexiviridae*.

Translation of TGBp2/3 from sgRNA1 is not mediated by either an internal ribosome entry site (IRES) or reinitiation. Given the finding that TGBp1/2/3 are translated mainly from sgRNA1, we next analyzed the mechanism underlying the translation of TGBp1/2/3 from sgRNA1. Translation of dORFs in polycistronic mRNAs requires noncanonical mechanisms, such as leaky scanning (10). We hypothesized that the dORFs encoding TGBp2 and TGBp3 in sgRNA1 were regulated via leaky scanning for two

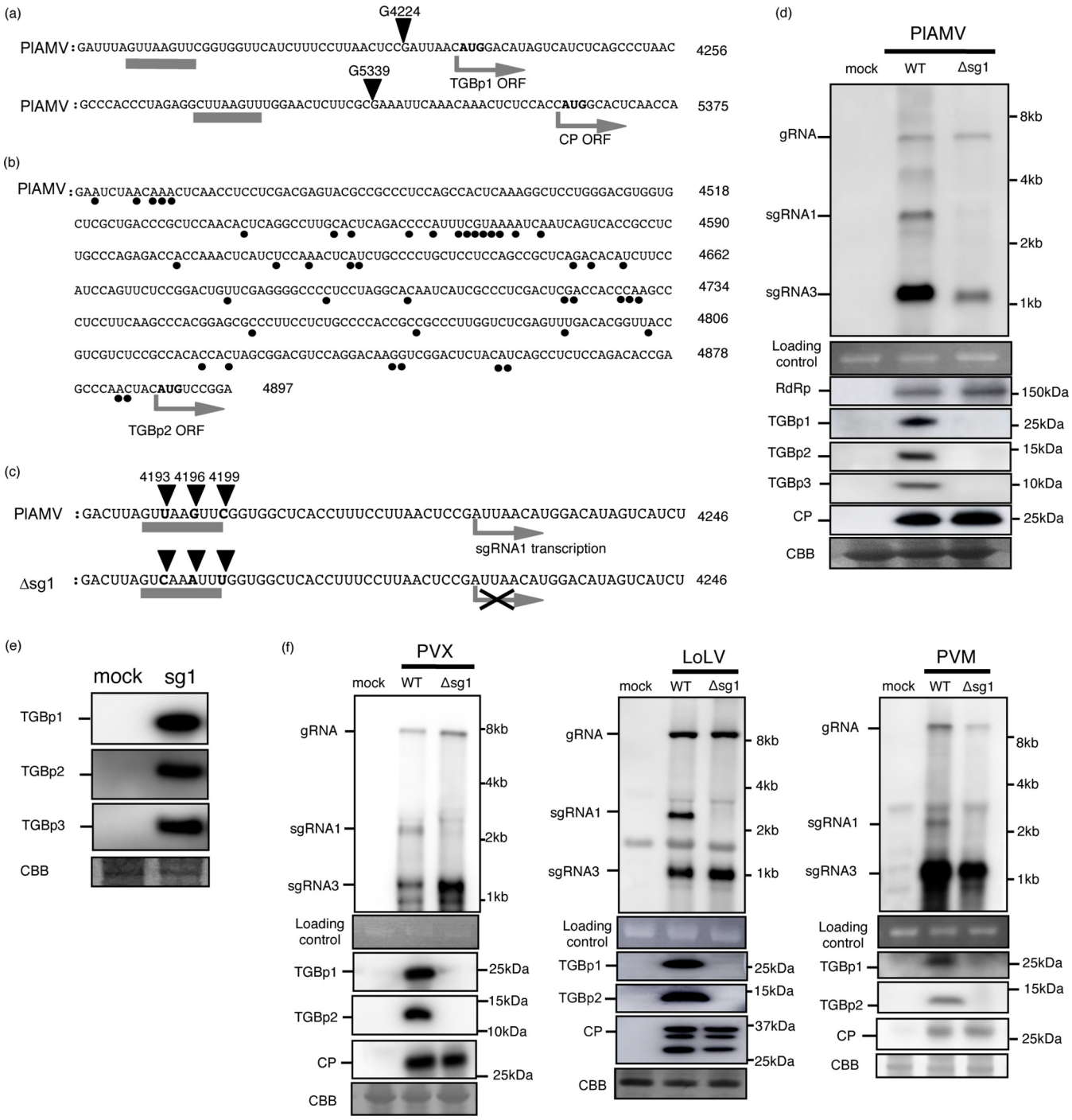

FIG 1 Flexivirus TGBps are translated from sgRNA1. (a) The sgRNA1 TSS and sgRNA3 TSS were uniquely determined at G4224 (upper black arrow) and G5339 (lower black arrow), respectively. The putative promoter sequences are underlined. The arrow indicates the TGBp1 ORF and CP ORF. Boldfaced AUG indicates the TGBp1 initiation codon and the CP initiation codon. (b) The sgRNA2 TSS was not uniquely determined. Black dots indicate the varied 5′ ends of cloned transcripts as determined by RACE analysis. There was no promoter-like sequence located in the region between the TGBp1 initiation codon and the TGBp2 initiation codon. (c) Schematic of PlAMV-Δsg1 constructs. The putative promoter sequences are underlined. U4193, G4196, and C4199 (boldfaced and indicated by black arrows in PlAMV) were changed to C4193, A4196, and U4199, respectively (boldfaced and indicated by black arrows in PlAMV-Δsg1). Transcription of sgRNA1 starts from G4224 as shown in panel a. (d) Northern blotting (upper window) and immunoblotting (lower windows) for the Δsg1 mutant of PlAMV. Total RNA and total protein samples extracted from noninoculated plants were used as mock samples. (e) Immunoblotting of TGBps extracted from protoplasts transfected with PlAMV-sgRNA1 (sg1). A nontransfected sample was used as a mock sample. (f) Northern blotting (upper window) and immunoblotting (lower windows) for the Δsg1 mutant of PVX (genus *Potexvirus*, *Alphaflexiviridae*), LoLV (genus *Lolavirus*, *Alphaflexiviridae*), and PVM (genus *Carlavirus*, *Betaflexiviridae*). Total RNA and total protein samples extracted from buffer-inoculated plants were used as mock samples.

reasons. First, according to the conventional potexviral TGBp expression model, leaky scanning regulates the translation of TGBp3 from sgRNA2 (16). Second, in the genome of PlAMV, there is no AUG triplet between the start codons of TGBp1 and TGBp2 or between the start codons of TGBp2 and TGBp3. This feature certainly favors leaky scanning to ensure adequate translation of TGBp2/3 from sgRNA1.

To exclude the possibility that a noncanonical translation mechanism other than leaky scanning regulates the translation of TGBp2/3 from sgRNA1, we tested the involvement of two potential noncanonical translation mechanisms, IRES and reinitiation.

IRES is the structural RNA sequence that permits cap-independent translation initiation for the dORF by directly recruiting the translation apparatus, including the PIC. We inserted a 40-nucleotide (nt) Kozak-stem-loop sequence (KS-sg1) to impair the migration of PIC scanning from the 5′ terminus (21, 22) (Fig. 2a). In addition, a green fluorescence protein (GFP)-coding sequence with a 21-nt 5′ leader sequence was fused to the 5′ end of sgRNA1 (GFP-sg1) to prevent the majority of the PIC from the 5′ terminus from reaching to TGBp initiation codons by making the PIC recognize the AUG of GFP. If an IRES was located upstream of the start codon of TGBp2 to allow translation of TGBp2/3, the level of TGBp2/3 accumulation with KS-sg1 or GFP-sg1 would be similar to that with WT sgRNA1. Immunoblot analysis showed no accumulation of TGBp1/2/3 in *Arabidopsis* protoplasts transfected with KS-sg1 and GFP-sg1 (Fig. 2b), suggesting that sgRNA1 does not contain a functional IRES. The expression of GFP was confirmed via immunoblotting (Fig. 2b).

To further evaluate the translational efficiency of TGBp1/2/3 in KS-sg1 and GFP-sg1 mutants, we performed a dual-luciferase assay. We constructed modified KS-sg1, GFP-sg1, and WT sgRNA1, in which the coding region of TGBp1, -2, or -3 was replaced with the *Renilla* luciferase (*Rluc*) gene (Fig. 2a). The *Rluc*-containing sgRNA1 constructs were transfected into *Arabidopsis* protoplasts, and Rluc activity was normalized to firefly luciferase (Fluc) activity to assess the activity levels and transformation efficiency in the protoplasts. For all TGBp proteins, the relative Rluc activities of each mutant (KS and GFP constructs) were significantly lower than that of each WT construct (Fig. 2c), supporting the notion that sgRNA1 does not contain a functional IRES.

Next, to examine the possibility of the involvement of reinitiation, we replaced the UGA stop codon of TGBp1 with a GGA triplet (Fig. 2a). The frequency of reinitiation depends on the distance between the stop codons of uORF and dAUG (21–24). If TGBp2 and TGBp3 were translated via reinitiation, replacement of the TGBp1 stop codon would result in clear separation between the initiation codon of TGBp2 and the end of the TGBp1 sequence, thereby reducing the level of TGBp2/3 expression. As expected, immunoblotting using anti-TGB1 antibody showed a band with a slightly higher molecular weight than that of the WT sg1-transfected protoplasts in the Δstp-transfected protoplasts, and a signal with a slightly lower molecular weight, presumably a degradation product, was also obtained (Fig. 2d). Both TGBp2 and TGBp3 accumulated in protoplasts transfected with Δstp to a level comparable that with WT-transfected samples. These results indicate that TGBp2 and TGBp3 are not translated from sgRNA1 via reinitiation.

TGBp2/3 are translated via leaky scanning through the TGBp1 initiation codon. We tested our hypothesis that TGBp2 and TGBp3 are translated from sgRNA1 via leaky scanning of the TGBp1 initiation codon. In leaky scanning, the uAUG competes with the dAUG for translation initiation by ribosomes: the nucleotide context of the uAUG affects the efficiency of translation initiation and subsequently the continuation of PIC scanning to reach the dAUG. In dicots, the optimal sequence context for translation initiation is RNN AUG G, which includes a purine residue (R; guanine [G] or adenine [A]) at the −3 position and a G at the +4 position relative to the first A in the initiation codon (8) (Fig. 3a). The initiation codon of WT TGBp1 has an optimal Kozak sequence with A at the −3 position. The plasmids containing KZ(−3A) variants were transfected to *Arabidopsis* protoplasts, and the accumulation of three TGB proteins was analyzed. When the A at −3 was replaced with the purine residue G [KZ(−3A)G], which produced another optimal

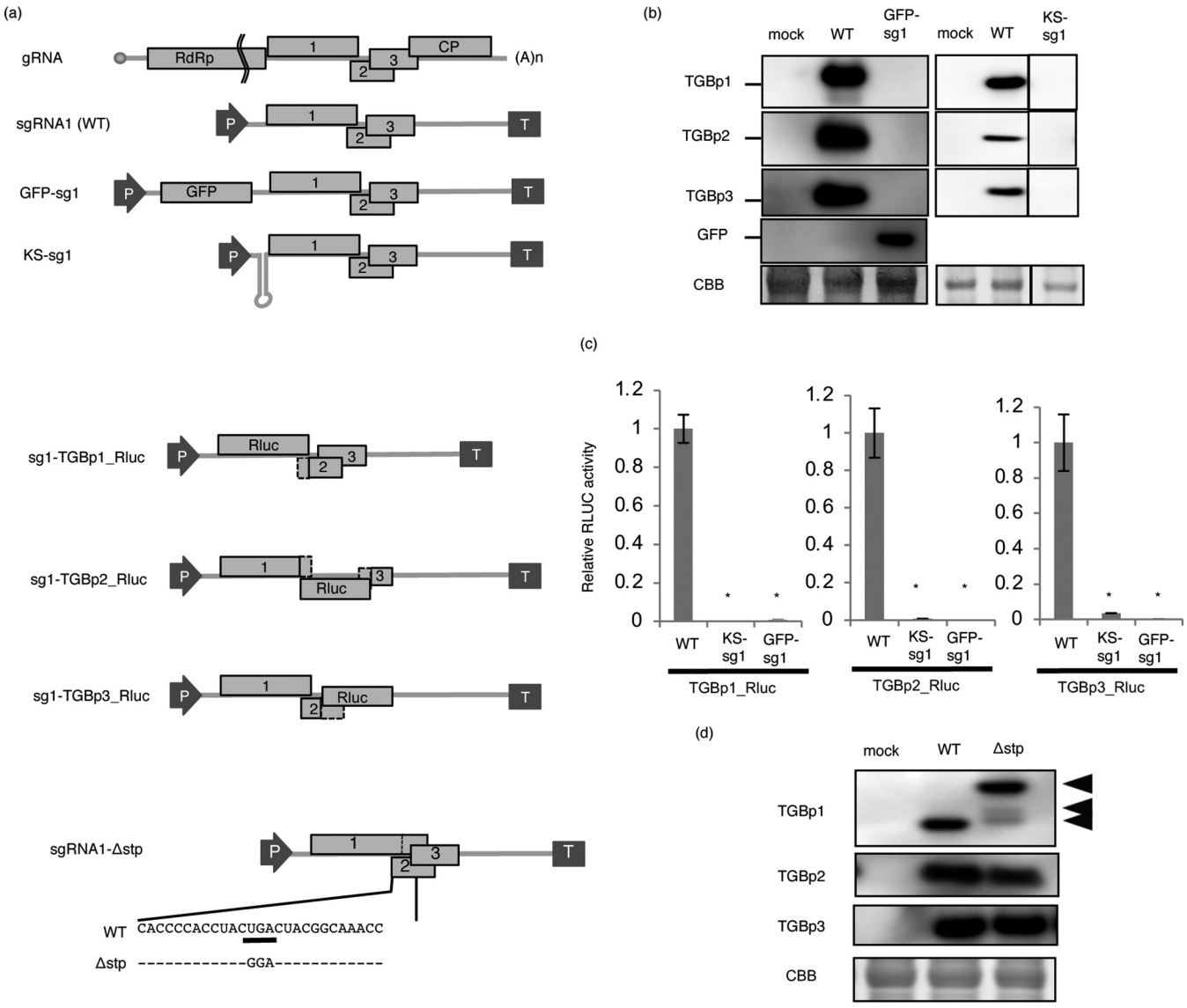

FIG 2 Translation of TGBp2/3 from sgRNA1 is not mediated by IRES or by reinitiation. (a) Schematic representation of PlAMV sgRNA1 cDNA constructs. Open box1, box2, and box3 represent TGBp1, TGBp2, and TGBp3 ORFs, respectively. The constructs contained full-length cDNA of sgRNA1 from a PlAMV-Pr isolate inserted between the cauliflower mosaic virus 35S promoter (P) and the NOS terminator (T). The 5′-terminal sequence of sgRNA1 is shown. A 40-nt Kozak stem-loop sequence (17) was inserted between the sgRNA1 leader sequence and the TGBp1 initiation codon (KS-sg1). A *GFP*-coding sequence with a 21-nt 5′ leader sequence (AGCUGUACAAGUAAGAUUAAC) was fused to the 5′ end of sgRNA1 (GFP-sg1) to promote the recognition of the upstream *GFP* ORF by the majority of PIC from the 5′ terminus, so as to impair the migration of scanning PIC. The TGBp1 stop codon (UGA, underlined) was changed to GGA, resulting in an extension of the TGBp1 ORF. Hyphens indicate that the nucleotide at that position was not changed (sgRNA1-Δstp). (b) Immunoblot analysis of TGBp accumulation in protoplasts transfected with KS-sg1 and GFP-sg1 constructs. (c) Translational efficiency of TGBps in KS-sg1 and GFP-sg1 variants as analyzed using a dual-luciferase assay. Rluc luminescence was normalized to Fluc luminescence. The mean values ± standard deviations (SD) from three independent experiments are shown. Asterisk indicates significant difference compared with the WT (Student's *t* test, $P < 0.05$). (d) Immunoblot analysis of TGBp accumulation with the Δstp construct. Modified TGBp1 with Δstp is indicated (black arrow).

sequence context, the accumulation of TGBp1, as well as TGBp2/3, was similar to that of the WT (Fig. 3b). In contrast, when we introduced pyrimidine (U and C) to create a poor sequence context at the −3 position of the TGBp1 AUG codon, TGBp1 accumulation was reduced, as expected (Fig. 3b). Strikingly, TGBp2/3 accumulation increased in accordance with the decrease in TGBp1 accumulation (Fig. 3b). These immunoblot results can be explained if TGBp2/3 are translated via leaky scanning of the TGBp1 initiation codon in sgRNA1.

Next, we used a dual-luciferase assay to analyze how mutations in the TGBp1 Kozak sequence influence TGBp2/3 translation efficiency from sgRNA1 via leaky scanning. We replaced each TGBp-coding region with the *Rluc* gene in the above-described sgRNA1

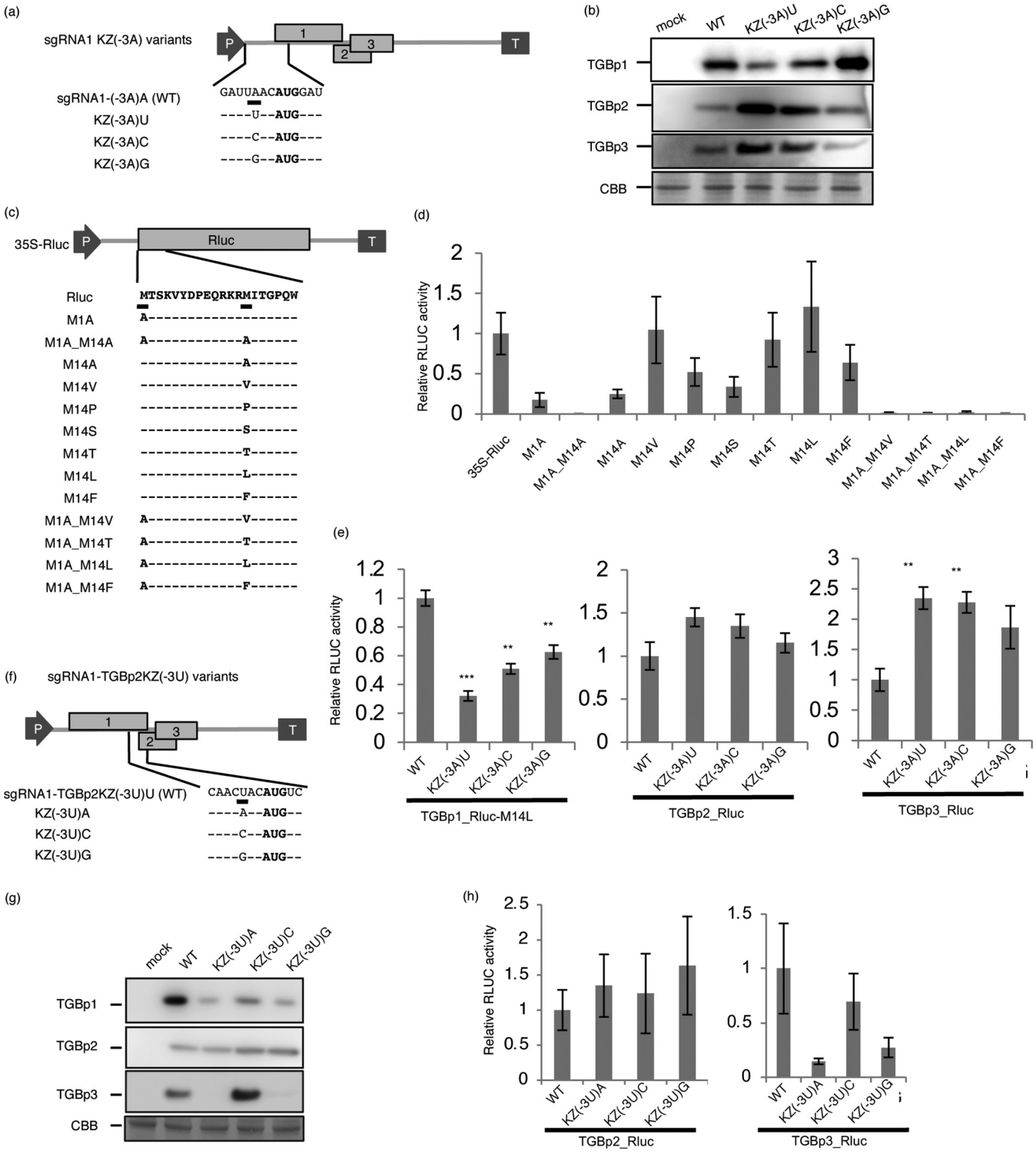

FIG 3 Translation of TGBp2/3 from sgRNA1 is mediated thorough leaky scanning. (a) KZ(−3A) variant constructs. The 5′-terminal sequence of sgRNA1 is shown. Boldfaced AUG is the initiation codon of TGBp1, and A at the −3 position (underlined) was changed to U, C, or G. Hyphens indicate that the nucleotide at that position was not changed. (b) Immunoblot analysis of TGBp accumulation in KZ(−3A) variants. (c) Rluc variant constructs. The N-terminal amino acid sequence of Rluc is shown. The first and fourteenth amino acids (M, underlined) were changed. (d) Rluc activity of the variants is shown. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SDs from three independent experiments are shown. (e) Dual-luciferase assay analysis of the translational efficiency of TGB proteins in KZ(−3A) variants. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SD from three independent experiments are shown. Asterisk indicates a significant difference compared with the WT (Student's *t* test, double asterisk, $P < 0.01$; triple asterisk, $P < 0.001$). (f) TGBp2KZ(-3U) variant constructs. Boldfaced AUG indicates the TGBp2 initiation codon, and U at the −3 position (underlined) was changed to A, C, or G. Hyphens indicate that the nucleotide at that position was not changed (sgRNA1-TGBp2KZ). (g) Immunoblot analysis of the accumulation of TGBps expressed by the TGBp2KZ variants. (h) Dual-luciferase assay analysis of the translational efficiency of TGBps expressed by TGBp2KZ variants. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SD from three independent experiments are shown.

Kozak −3 context variants and in WT sgRNA1. Note that we used an *Rluc* gene variant, Rluc-M14L, in which an in-frame second AUG triplet in the *Rluc* gene located 39 nt downstream of the canonical initiation codon was replaced with CUU to eliminate any possible interference from the potential translation of a partial gene product from the second AUG (Fig. 3c and d). As shown in Fig. 3e, the mutant with the purine substitution, sg1_KZ(−3A)G_TGBp1-Rluc-M14L, which had a potentially optimal Kozak sequence, exhibited higher translation efficiency from the TGBp1 initiation codon than did the mutants with the pyrimidine substitution, sg1_KZ(−3A)U_TGBp1-Rluc-M14L and sg1_KZ(−3A)C_TGBp1-Rluc-M14L. Moreover, the mutant with the purine substitution exhibited lower TGBp2/3 translation efficiency than did the mutants with pyrimidine substitution. These data support the notion that the translational efficiency of TGBp2/3 in sgRNA1 is dependent on the nucleotide context of the TGBp1 initiation codon, potentially through leaky scanning. In line with the conventional TGBp expression model (16), we also confirmed that TGBp3 translation is actually mediated by leaky scanning of the TGBp2 initiation codon in sgRNA1 (Fig. 3f to h).

To further test the leaky scanning hypothesis, we introduced six AUG codons throughout the TGBp1 ORF by substituting nucleotides without changing the TGBp1 amino acid sequence. In addition, the nucleotides located at position −3 of each introduced AUG were replaced with G to satisfy the optimal Kozak sequence unless the TGBp1 amino acid sequence were altered (Fig. 4a, midAUG). If TGBp2 and TGBp3 were translated from sgRNA1 by leaky scanning, insertion of several AUG codons preceding the AUG codons of TGBp2 and TGBp3 would reduce their translation efficiency by trapping the majority of PIC. The midAUG mutant and its *Rluc* derivatives were transfected into protoplasts, and expression of TGBp1/2/3 was analyzed via immunoblotting and luciferase assay. As expected, although immunoblotting showed that the accumulation of mutated TGBp1 did not differ from that of the WT, the levels of TGBp2 and TGBp3 and their translation efficiency, as quantified by the luciferase assay, were significantly reduced (Fig. 4b and c). These results also agree with our hypothesis that leaky scanning of the TGBp1 initiation codon regulates the translation of TGBp2/3 from sgRNA1.

As shown in Fig. 3f, g, and h, the translation of TGBp3 requires leaky scanning through the TGBp2 initiation codon. In summary, our results suggest that TGBp2 translation requires leaky scanning of the TGBp1 initiation codon, and TGBp3 translation requires leaky scanning of multiple upstream initiation codons in sgRNA1.

The TGBp1 ORF 5′-leader sequence length regulates the efficiency of leaky scanning in sgRNA1. Although our results indicate that the nucleotide context of the TGBp1 initiation codon regulates the efficiency of leaky scanning, the nucleotide context of the TGBp1 initiation codon is inherently optimal. This made us postulate that an unknown factor promotes leaky scanning of the TGBp1 initiation codon, because TGBp2 and TGBp3 are translated sufficiently. As shown in Fig. 1, the TGBp1 leader in sgRNA1 is remarkably short (7 nt). Due to the conserved genomic structure among flexiviruses, we predicted that the mechanism of TGBp2/3 translation from sgRNA1 would be conserved among other flexiviruses. To examine whether this short leader sequence of sgRNA1 is widely conserved among potexviruses and related viruses, we determined the TSSs of the sgRNA1s of several potexviruses, a lolavirus, and a carlavirus. The TSS of sgRNA1 was determined by analyzing RNA extracted from infected plants of six viruses: four viruses of the genus *Potexvirus* (PVX, hydrangea ringspot virus [HdRSV], pepino mosaic virus [PepMV], and white clover mosaic virus [WClMV]), LoLV of the genus *Lolavirus* of the same family, *Alphaflexiviridae*, as *Potexvirus*, and PVM of the genus *Carlavirus* of a different family, *Betaflexiviridae*. We found that the sgRNA1 5′ leaders of these viruses ranged from 1 (WClMV) to 8 (HdRSV and PVM) nt in length, supporting the notion that this short 5′-leader length in the translation of TGBp2/3 from sgRNA1 is conserved among these viruses (Fig. 5). Thus, we assumed that the short leader sequence of TGBp1 is another factor that regulates leaky scanning of sgRNA1.

To test this hypothesis, we analyzed the levels of TGBp1/2/3 accumulation from PlAMV sgRNA1 variants whose leader lengths were shortened to 5, 3, or 1 nt (Fig. 6a). Immunoblotting revealed that TGBp1 levels gradually decreased as the leader length became shorter; in contrast, accumulation of TGBp2 and TGBp3 increased

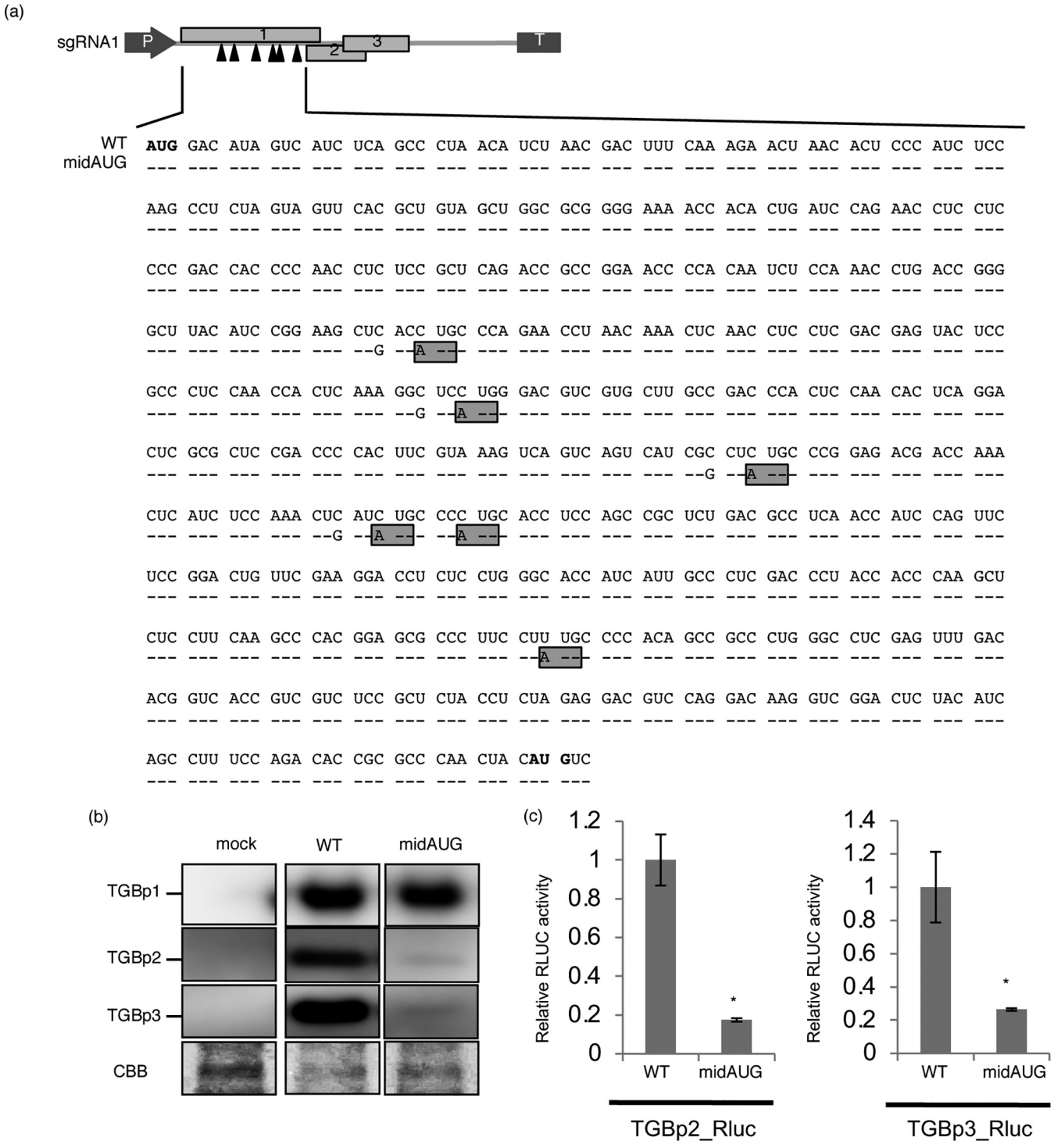

FIG 4 Translation of TGBp2/3 is dependent on leaky scanning of TGBp1 initiation codon. (a) Schematic of the midAUG construct. The sequence between the TGBp1 initiation codon and the TGBp2 initiation codon is shown. Boldfaced AUG indicates the initiation codon of TGBp1 (upper) and TGBp2 (lower). Six additional AUGs (black arrow) were inserted into the TGBp1 coding region via base substitution without changing the amino acid sequence (gray box). Hyphens indicate that the nucleotide at that position was not changed. (b) Immunoblot analysis of the accumulation of TGBps expressed by the midAUG construct. (c) Dual-luciferase assay analysis of the translational efficiency of TGB2/3 expressed by the midAUG construct. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SD from three independent experiments are shown. Asterisk indicates a significant difference compared with the WT (Student's *t* test, $P < 0.05$).

(Fig. 6b). Luciferase assay results also showed that the translational efficiency of TGBp2/3 increased in a stepwise manner as the 5′-leader sequence was shortened (7-nt length to 1-nt length), whereas that of TGBp1 decreased (Fig. 6c), indicating that shortening the leader sequence promoted the efficiency of leaky scanning. To

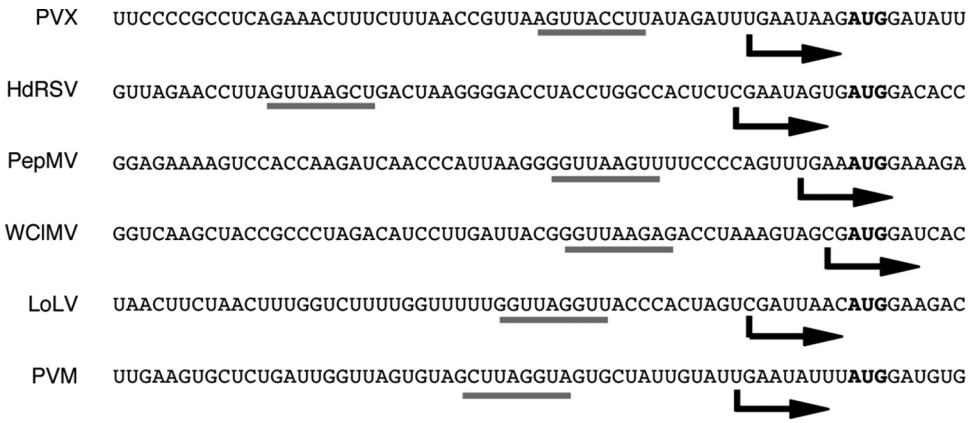

FIG 5 Short TGBp1 5′ leader sequence is highly conserved among several flexiviruses. PVX, HdRSV, WClMV, PepMV (genus *Potexvirus*, *Alphaflexiviridae*), LoLV (genus *Lolavirus*, *Alphaflexiviridae*), and PVM (genus *Carlavirus*, *Betaflexiviridae*) were analyzed. Deletion of sgRNA1 from PVX, LoLV (*Alphaflexiviridae*), and PVM (*Betaflexiviridae*) impairs TGB1/2 accumulation in a viral infection context, showing the relevance of sgRNA1 to the translation of all TGBps in other flexiviruses (Fig. 1). A survey of the TSSs of sgRNA1 of PVX, HdRSV, WClMV, PepMV, LoLV, and PVM showed that the length of the sgRNA1 5′ leader sequence was 7 nt in PVX and LoLV, 8 nt in HdRSV and PVM, 3 nt in PepMV, and 1 nt in WClMV. Putative promoter sequences are underlined. The arrow indicates the TSS of sgRNA1. Boldfaced AUG indicates the initiation codon of TGBp1. TSSs of each sgRNA1 are indicated by black arrows.

further validate this hypothesis, we reasoned that extension of the sgRNA1 5′-leader sequence would reduce the efficiency of leaky scanning. To test this possibility, we constructed three sgRNA1 variants: dp5U with a tandemly duplicated WT leader sequence (14 nt), 5Usub1 with a heterologous long leader sequence (42 nt), and 5Usub2 with a heterologous GC-rich leader sequence with the same length as the WT (7 nt) (Fig. 6d). Immunoblotting and luciferase assay results showed that the translation levels of TGBp2/3 in protoplasts transfected with sgRNA1 variants with a long leader sequence (dp5U and 5Usub1) were lower than those produced from WT sgRNA1 (sg1), whereas the accumulation of TGBp1 did not differ (Fig. 6e and f). In contrast, 5Usub2 had the same leader length as the WT but a higher GC content and was equivalent to the WT in terms of the accumulation levels of all TGBps (Fig. 6e and f). Taken together, these results indicate that leaky scanning induced by a short 5′ UTR (LISH) enables ribosomes to efficiently translate TGBp2 and TGBp3.

Leader sequence length is optimized for viral infection. Our results so far indicate that the amount of the three TGBp translations from sgRNA1 varies depending on the length of the leader sequence. Thus, to test whether this variation has biological relevance, we employed a *trans*-complementation system incorporating a GFP-labeled PVX mutant lacking sgRNA1 (PVX-GFP-Δsg1) in combination with transient transfection with PVX-sgRNA1 variants with various leader lengths (1, 4, 7, or 10 nt) (Fig. 7a). Using GFP-tagged PVX, the efficiency of viral spread from infected cells to neighboring healthy cells (cell-to-cell movement) was quantified by measuring the area of GFP fluorescence (17). GFP foci derived from the *trans*-complementation of sgRNA1 variants with different leader lengths were smaller than those resulting from WT leader-containing sgRNA1, suggesting that the alteration of LISH via the mutation of leader length modulated the cell-to-cell movement efficiency of PVX, likely due to an imbalance in the accumulation levels of TGBps (Fig. 7b and c). Similar results were obtained in PlAMV using a corresponding *trans*-complementation assay combining a GFP-labeled PlAMV mutant lacking sgRNA1 (PlAMV-GFP-Δsg1) and PlAMV-sgRNA1 variants with various leader lengths (1, 3, 5, 7, or 10 nt) (Fig. 7d). The efficiency of cell-to-cell movement of PlAMV-GFP increased as the leader sequence became longer, and, noticeably, the movement efficiency of the virus with a 10-nt leader was nearly the same as that of the WT (Fig. 7e and f). This may be because higher efficiency of cell-to-cell movement does not always maximize viral fitness (25).

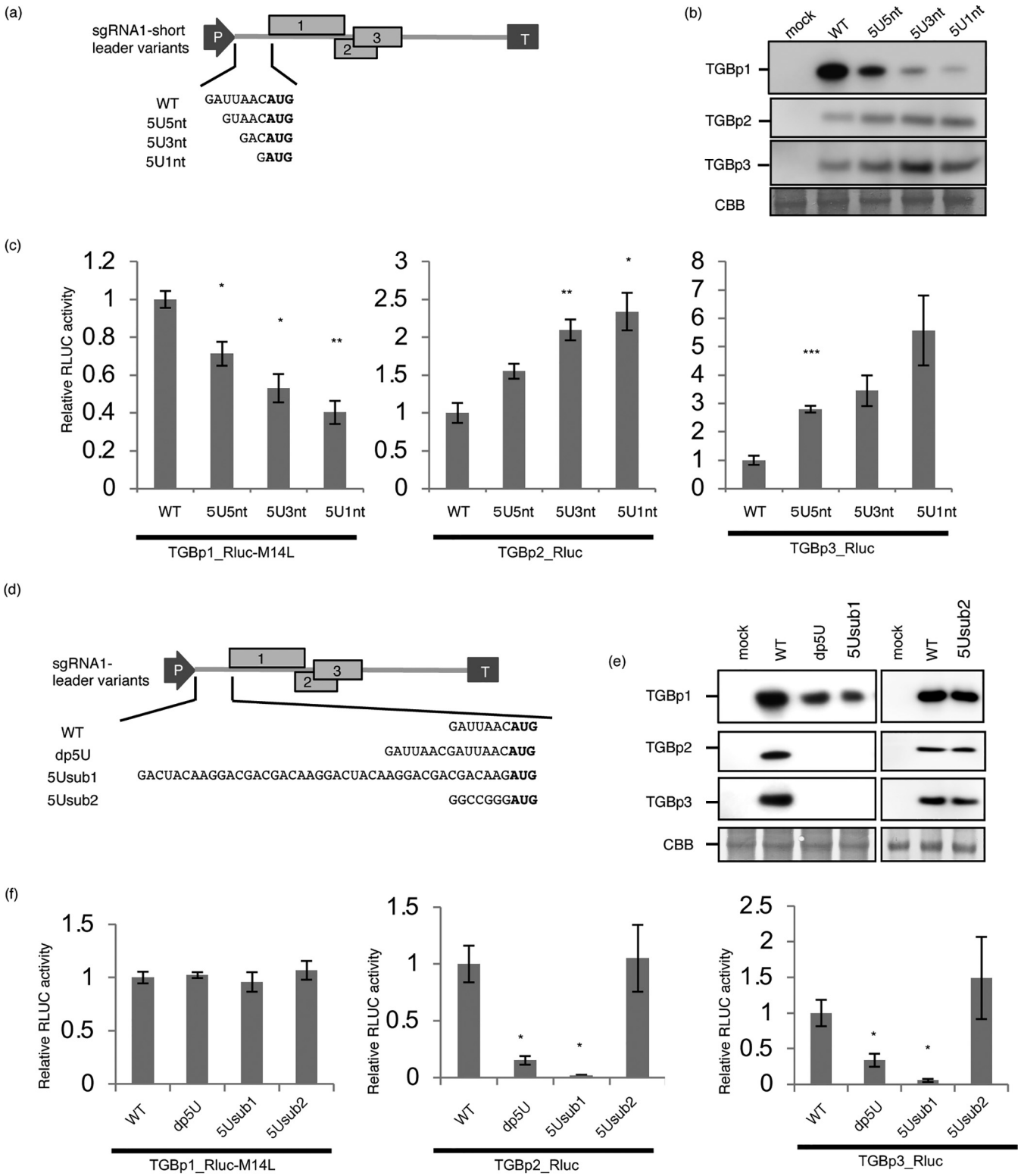

FIG 6 Efficiency of leaky scanning depends on the length of 5′ leader sequence. (a) Short leader variant constructs. The 5′-terminal sequence of sgRNA1 is shown. Boldfaced AUG indicates the initiation codon of TGBp1. (b) Immunoblot analysis of TGBp accumulation in short leader variants. (c) Dual-luciferase assay analysis of the translational efficiency of TGBp in short leader variants. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SD from three independent experiments are shown. Asterisks indicate a significant difference compared with WT (Student's *t* test, single asterisk, $P < 0.05$; double asterisk, $P < 0.01$; triple asterisk, $P < 0.001$). (d) Construct schematic. The 5′-terminal sequence of sgRNA1 is shown. Boldfaced AUG indicates the TGBp1 initiation codon, and the leader sequence was modified. (e) Immunoblot analysis of the accumulation of TGBps expressed by leader length variants. (f) Dual-luciferase assay analysis of the translational efficiency of TGBps expressed by leader length variants. Rluc luminescence was normalized to Fluc luminescence. Mean values ± SD from three independent experiments are shown. Asterisks indicate a significant difference compared with the WT (Student's *t* test, $P < 0.05$).

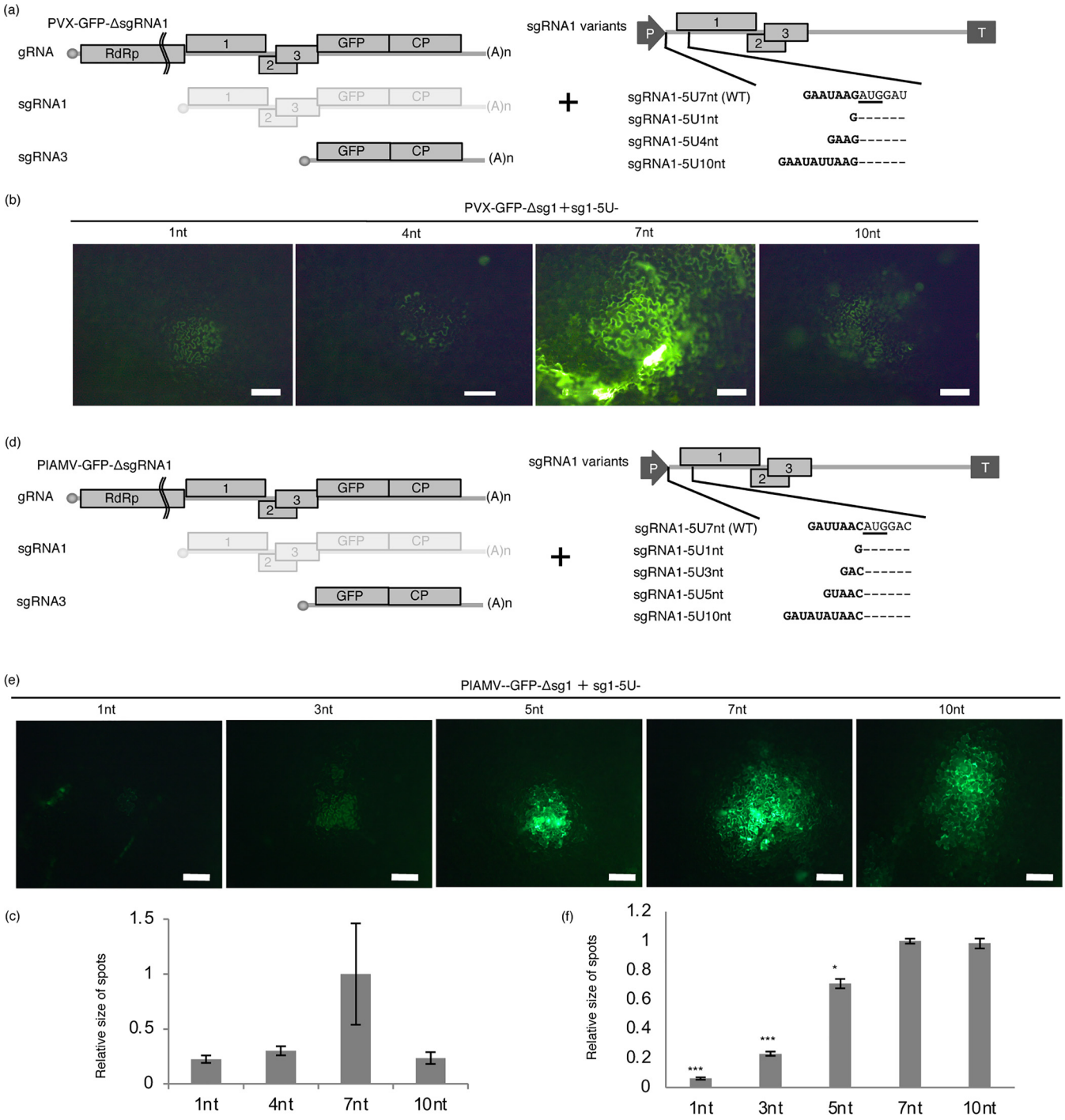

FIG 7 Leader-sequence length optimizes virus propagation. (a) 5′ UTR variant constructs. Boldfaced GAAUAAG indicates the WT PVX sgRNA1 leader sequence. Hyphens indicate that the nucleotide at that position was not changed from the WT. AUG (underlined) is the PVX TGBp1 initiation codon. (b) Images of the spread of PVX-GFP. Bars indicate 0.2 mm. (c) Quantification of the sizes of fluorescent foci in inoculated leaves. Fluorescence images with more than 10 foci in panel b were processed using ImageJ v1.40 software (NIH) to measure the size of viral infection foci. The sizes were normalized to those of 5U7nt (WT). (d) 5′ UTR variant constructs. Boldfaced GAUUAAC indicates the WT PlAMV sgRNA1 leader sequence. Hyphens indicate that the nucleotide at that position was not changed from the WT. AUG (underlined) is the PlAMV TGBp1 initiation codon. (e) Images of the spread of PlAMV-GFP. Bars indicate 0.2 mm. (f) Quantification of the sizes of fluorescent foci in inoculated leaves. Fluorescence images with more than 10 foci in panel b were processed using ImageJ v1.40 software (NIH) to measure the sizes of viral infection foci. The sizes were normalized to those of 5U7nt (WT). Asterisks indicate a significant difference compared with the WT (Student's *t* test, single asterisk, $P < 0.05$; triple asterisk, $P < 0.001$).

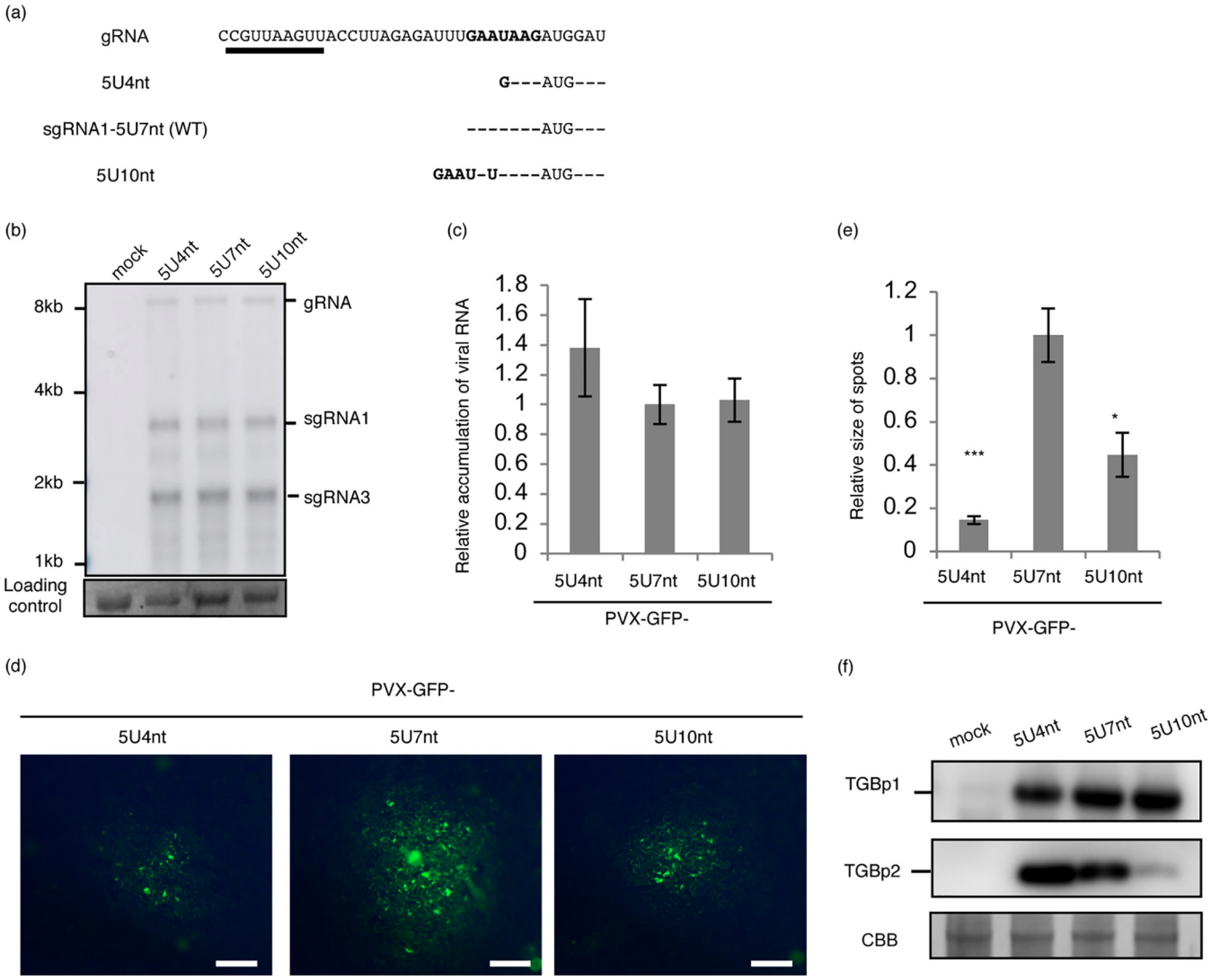

FIG 8 Length of leader sequence optimizes virus propagation. (a) 5′ UTR variant constructs. The putative sgRNA1 promoter sequence is underlined. Boldfaced GAAUAAG indicates the WT PVX sgRNA1 leader sequence. Hyphens indicate that the nucleotide at that position was not changed in the WT. (b) Northern blotting of RNA extracted from *N. benthamiana*. (c) Relative viral RNA accumulation as quantified by qRT-PCR. (d) Images of the spread of GFP signals produced by PVX-GFP. Bars indicate 0.2 mm. (e) Quantification of the size of the fluorescent foci in inoculated leaves. Fluorescence images with more than 10 foci in panel d were processed using ImageJ v1.40 software (NIH) to measure the sizes of viral infection foci. The sizes were normalized to those of 5U7nt (WT). Asterisks indicate a significant difference compared with the WT (Student's *t* test, single asterisk, $P < 0.05$; triple asterisk, $P < 0.001$). (f) Immunoblot analysis of TGBp accumulation in the 5′ UTR variant constructs.

Furthermore, we constructed PVX-GFP mutants to express sgRNA1 with various leader lengths in *cis* (Fig. 8a). The overlap of the RdRp coding region and the leader sequence of sgRNA1 prevented the construction of PlAMV variants without changing the amino acid sequence of RdRp. We confirmed that the TSSs of sgRNA1 variants transcribed from these PVX mutants by RACE analysis were as intended. In addition, Northern blotting and quantitative reverse transcription-PCR (qRT-PCR) indicated that these mutations did not significantly affect viral RNA multiplication in protoplasts isolated from *Nicotiana* suspension culture cells (BY-2) (Fig. 8b and c). When GFP-tagged PVX variants expressing sgRNA1 with leader lengths of 4 or 10 nt were inoculated into *N. benthamiana*, the relative sizes of GFP fluorescence spots produced by the WT virus were significantly larger than those of mutants with shorter and longer leader sequences (Fig. 8d and e). We confirmed that the accumulation of TGB1 increased and that of TGB2 decreased as the length of the 5′ UTR increased (Fig. 8f). Taken together, our results suggest that the native leader length optimizes the efficiency of LISH to achieve optimal cell-to-cell movement.

DISCUSSION

In this study, we elucidated a novel regulatory mechanism involving leaky scanning. We found that the TGBps encoded by the genome of potexviruses are all translated from a single RNA molecule, sgRNA1, via leaky scanning. Moreover, the remarkable shortness of the leader sequence upstream of the first ORF in sgRNA1, TGBp1, is well conserved among potexviruses. Mutational analyses showed that the efficiency of leaky scanning was strictly dependent on the 5′ UTR length. We present here a mechanism for the regulation of leaky scanning in addition to the Kozak sequence.

All TGB proteins are translated from sgRNA1. We analyzed the molecular mechanism underlying the translation of three TGBps from sgRNA to reverify the potexviral TGBp expression model proposed previously (16). In the conventional model, TGBp1 is translated from sgRNA1, whereas TGBp2 and TGBp3 are translated from sgRNA2. Our study demonstrated that in an sgRNA1-deleted PlAMV mutant (Δsg1), the accumulation of TGBp2/3 as well as TGBp1 was impaired, despite a level of RdRp accumulation comparable to that in the WT. Consistent with this result, TGBp2/3 together with TGBp1 accumulated significantly in protoplasts transfected with only sgRNA1. These results indicate that TGBp2 and TGBp3 are substantially translated from sgRNA1. Furthermore, we did not find a clearly defined TSS for sgRNA2, in contrast to clearly defined TSSs for sgRNA1 and sgRNA3. It is true that there are several previous studies that have detected sgRNA2-like molecules from viruses belonging to the genus *Potexvirus* (26, 27), and the possibility that these viral RNAs have some function cannot be completely ruled out. Some of the degradation intermediates of viral RNAs accumulate in large amounts and function as noncoding RNAs (28), and sgRNA2-like molecules may also accumulate under certain conditions and have some function. However, the present results indicate that sgRNA1, not sgRNA2, is responsible for the conserved function as the major template of the translation of TGBp2 and TGBp3. Our experimental results also suggest that this TGBp translation model is conserved among *Alphaflexiviridae* and *Betaflexiviridae*. Therefore, our novel model of TGBp translation, in which all TGBps are translated from sgRNA1, may be employed across a range of plant viruses, including potexviruses and other related viruses.

5′ UTR lengths in viral gRNA and sgRNA. We found that the 5′ UTR of the flexiviral sgRNA1 is very short, ranging from 1 to 8 nt. Considering the 5′ UTR of flexiviral RdRp, which is translated directly from gRNA, is between 70 and 110 nt, the 5′ UTR of sgRNA1 is extremely short. Although plant virus MPs are, in many cases, translated from sgRNAs, the 5′ UTR of sgRNA1 of flexiviruses is much shorter than those of MP-encoding sgRNAs of other RNA viruses. For instance, the 5′ UTR length of the sgRNA encoding the MP of tobacco mosaic virus (TMV) is approximately 60 nt, which is much longer than that in flexiviruses, indicating that LISH is not required for the translation of TMV MP, in harmony with its single-ORF character. On the other hand, TGBp-type MPs are also found in several species in the genus *Benyvirus* and family *Virgaviridae*, but the 5′ UTRs of these MPs are all more than 100 nt long. This may be because TGBp2 in these viruses has been reported to be translated from its own sgRNA, not by leaky scanning of TGBp1. The control of leaky scanning by the length of the 5′ UTR is a unique mechanism in flexiviruses, and it has been reported that the length of the 5′ UTR plays an important role in the regulation of viral gene expression in viruses that are phylogenetically distant from potexvirus (12). It should be considered an important element in the analysis of viral gene expression control mechanisms.

Molecular mechanism underlying LISH. In this study, we demonstrated that the efficiency of leaky scanning is regulated by the length of the 5′ UTR. Although a previous study using an *in vitro* translation system and artificial mRNA showed that it is theoretically possible that the short 5′ UTR upregulates the efficiency of leaky scanning (29, 30), this is the first study to show that LISH is actually employed to regulate the efficiency of leaky scanning in viral and eukaryotic mRNAs in living cells. Our results raise a question about the molecular mechanism by which such short mRNA leader sequences enhance the efficiency of leaky scanning. A previous study reported that when the P-site of the mammalian 43S PIC binds to the AUG initiation codon, the PIC structurally

contacts multiple nucleotides proximal to the AUG codon, including the upstream 17 nt and the downstream 11 nt (31). Therefore, when the 5′ UTR length is shorter than 17 nt, the absence of RNA in the space inside the PIC normally occupied by the 5′ UTR sequence may alter the PIC's overall conformation, disrupt the efficiency of start codon recognition, and thereby increase the efficiency of leaky scanning. On the other hand, given that the efficiency of leaky scanning of sgRNA1 was also dependent on the Kozak sequence even when the 5′ UTR was 7 nt (Fig. 3), 43S PIC scanning seems to still function even under the LISH condition.

It should also be noted that the mechanism of LISH may be different from the mechanism commonly assumed as leaky scanning. In general, leaky scanning refers to inefficient recognition of start codons in poor contexts, such as suboptimal Kozak sequences or non-AUG start codons (32). However, although we have not yet analyzed how definitive this possibility is, it is possible that the PIC has passed through the AUG at the start of scanning in LISH, making recognition of the start codon less likely. In other words, if LISH is because scanning itself does not occur, it is a different mechanism from conventional leaky scanning, in which scanning occurs but its recognition efficiency is low.

Biological significance of LISH in viral infection. Using a *trans*-complementation assay, we showed that the infection efficiency of sgRNA1-deficient PVX was optimal when sgRNA1 with a WT 5′ UTR length was complemented. The infection efficiency decreased upon complementation with sgRNA1s containing elongated or shortened 5′ UTRs (Fig. 7). The same result was obtained when the intact genomic sequence of a PVX-GFP infectious clone was manipulated to generate virus mutants with elongated or shortened 5′ UTRs (Fig. 8). These results suggest that the length of the sgRNA1 leader sequence is optimal for regulating the efficiency of cell-to-cell movement, potentially through fine-tuning the accumulation levels of the three TGBps. In the region upstream of the TGBp3 initiation codon in PlAMV sgRNA1, there are only two AUG codons, including those of TGBp1 and TGBp2, implying evolutionary pressure to avoid unnecessary dAUGs that would impair the translation of TGBp2/3.

TGBps act as MPs, but their detailed characteristics differ slightly (33). TGBp1 is a multifunctional protein with RNA-binding and ATPase activities that acts as an MP and an RNA-silencing suppressor. TGBp1 modifies the plasmodesmata size exclusion limit and reorganizes the viral genome, proteins, host actin, and endomembranes to form the viral X body for cell-to-cell movement (33). Both TGBp2 and TGBp3 are transmembrane proteins, but they localize to different subdomains of the endoplasmic reticulum (34). Therefore, viruses may orchestrate the different functions of TGBps to adapt to the cellular environment of the host by optimizing the levels of the three MPs. In fact, the Kozak sequence of the AUGs of TGBp2 and TGBp3 are not optimal, and recognition of the TGBp1 AUG is restricted by LISH. Thus, accumulation of all three TGBps is suppressed to a certain degree by LISH together with Kozak sequence, regulating their translation initiation efficiency. This may be because the accumulation of TGBps can be stressful to plant cells. PVX TGBp1 triggers cell death resulting from endoplasmic reticulum stress in *N. benthamiana* (35). Moreover, potexviral TGBp3 drastically modifies the membrane structure of the endoplasmic reticulum (36), and its overexpression causes veinal necrosis (37). This is in agreement with the finding that TGBp3 of shallot virus X and lily virus X, both members of genus *Potexvirus*, generally has a non-AUG initiation codon with lower translation initiation efficiency than the normal AUG codon (38, 39). Furthermore, a simulation study clearly demonstrated that excessive cell-to-cell movement efficiency impairs optimal selection against defective genomes and deteriorates the quality of viral genomes (25). The highest efficiency of movement does not always maximize viral fitness. Suppression of MP translation may also increase viral fitness by suppressing movement efficiency. The accumulation of the sgRNA1-deficient mutant virus was not significantly different from that of the wild type at 1.5 days after inoculation. The need for TGB proteins in virus infection will take place during the advanced phase of infection, such as cell-to-cell transfer and countering

TABLE 1 List of antigens used to produce the antibodies

Target	Sequence
PVX-TGBp1	Full length
PVX-TGBp2	CGNNHSSH and SLPHGGAYRDGTK
LoLV-TGBp1	GFERTDEPLPSDS
LoLV-TGBp2	ADPHHPTV
LoLV-CP	Full length
PVM-TGBp1	LHKYKFERLNNKLA
PVM-TGBp2	GDRDHRLPHGGWY
PVM-CP	AKEAGTSQAAKGNR

plant defense responses. Members of the genus *Pelarspovirus* within the family *Tombusviridae* encode three proteins that are translated from a single sgRNA. These proteins are translated from poor Kozak context or non-AUG initiation codons, which suggests that leaky scanning is a basic strategy for RNA viruses (13).

In conclusion, this study revealed a protein translation strategy of plant viruses that utilizes LISH, a translation control mechanism involving leaky scanning dependent on the length of the 5′ UTR. This study confirms the first evidence that the phenomenon of LISH, which was previously confirmed in artificial environments, actually has biological significance.

MATERIALS AND METHODS

Plant materials and growth conditions. *Nicotiana benthamiana* plants were maintained in a growth chamber with a 16-h light, 25°C, 8-h dark, 20°C, cycle throughout the assays. Detailed conditions for *Arabidopsis* cell culture were described previously (40). Detailed conditions for *Nicotiana* cell culture were described previously (41).

Antibodies. Anti-PlAMV-TGBp1, -TGBp2, -TGBp3, -CP, -RdRp, and -PVX-CP antibodies were prepared as described previously (17, 42–45). Other antibodies were raised in rabbits using purified peptides as antigens (Eurofins Genomics, Tokyo, Japan). Detailed information on antigens is shown in Table 1.

Plasmid construction. The sequences of all the primers used in this study are shown in Table 2.

(i) Infectious cDNA clone. Infectious GFP-expressing PlAMV (pPlAMV-GFP) (46, 47) and PVX (pPVX-GFP) isolates were obtained as described previously. Infectious cDNA clones of PlAMV (46, 48), PVX-OS strain (pPVX) (45), LoLV (pLoLV); (49), hydrangea ringspot virus (HdRSV; pHdRSV) (47), and white clover mosaic virus (WClMV; pWClMV) (49) were constructed as described previously. Infectious clones of pepino mosaic virus (PepMV; pPepMV) and PVM (pPVM) were constructed as follows. Full-length cDNA sequences of PepMV (PV-0716; Leibniz-Institut) and PVM (MAFF number 307027) were amplified using the PepMV1F/GRR and PVM1F/GRR primer sets, respectively. DNA fragments harboring the 35S promoter and nopaline synthase (NOS) terminator from the pCAMBIA1301 vector backbone with 5′- and 3′-terminal sequences of PepMV and PVM were amplified from pPPVOu binary vectors (50) using the GRF/PepMV35SR and GRF/PVM35SR primer sets, respectively. Full-length viral cDNAs were assembled with the corresponding PCR product consisting of the 35S promoter, NOS terminator, and pCAMBIA1301 vector backbone using a GeneArt seamless cloning and assembly kit (Thermo Fisher Scientific, Waltham, MA, USA) to generate pPepMV and pPVM.

(ii) Δsg1 mutant virus. To construct the sgRNA1-deleted PlAMV mutant (Δsg1), several silent nucleotide substitutions were introduced into the sgRNA1 promoter while leaving the RdRp amino acid sequence unchanged, according to a previous report (20). Partial viral sequences containing nucleotide substitutions and a binary vector fragment were amplified from pPlAMV using the 35SF/PlAMVdsg1R, PlAMVdsg1F/NOSR, and NOSF/35SR primer sets. The three fragments were assembled using a GeneArt seamless cloning and assembly kit. PlAMV-GFP-Δsg1 was constructed using the same primer sets as those for PlAMV-Δsg1, with a distinct template, pPlAMV-GFP. LoLV-Δsg1 and PVM-Δsg1 were constructed as follows. Partial viral sequences were amplified from pLoLV or pPVM using the primer sets 35SF/LoLVdsg1R and LoLVdsg1F/NOSR or 35SF/PVMdsg1R and PVMdsg1F/NOSR, respectively. Fragments derived from each virus were assembled with the vector backbone generated from the NOSF/35SR primer set using a GeneArt seamless cloning and assembly kit. To generate PVX-Δsg1, we amplified a DNA fragment containing the nucleotide substitutions from pPVX using the OSdsg1F/OSdsg1R primer set. The amplified fragments were self-ligated using a ligation convenience kit (Nippon Gene, Tokyo, Japan). PVX-GFP-Δsg1 was constructed in the same manner as that for PVX-Δsg1, with pPVX-GFP used as the template.

(iii) 35S-Fluc, 35S-Rluc, and amino acid substitution Rluc mutants. 35S-Fluc was constructed as follows. The coding region of the *FLUC* gene was amplified via PCR using the 35S-fluc-1F and 35S-fluc-1653R primer set. The plasmid fragment was obtained from 35S-sGFP (51) using the 35S-down1F and 35S-up1R primers. The PCR-amplified *FLUC* gene was inserted downstream of the 35S promoter in the plasmid fragments using a GeneArt seamless cloning and assembly kit. 35S-Rluc was constructed as

TABLE 2 List of primers used and their sequences

Primer name	Primer sequence
PepMV1F	GAAAACAAAACATAACACATAATATC
GRR	GGGGTACCGCTACGTAACGGCATGACAGTGTTTTTTTTTTTTTTTTTTTTTTTT
PVM1F	GATAAACAAACATACAATATCTGGACTTACACTGC
GRF	CCGTTACGTAGCGGTACCCCTCAAACATTTGGCAATAAA
PVM35SR	ATATTGTATGTTTGTTTATCCCTCTCCAAATGAAATGAAC
PepMV35SR	ATGTGTTATGTTTTGTTTTCCCTCTCCAAATGAAATGAAC
35SF	TCGGGAAACCTCCTCGGATTCCATTGCCCAGC
35SR	AATGGAATCCGAGGAGGTTTCCCGATATTACCC
NOSF	ATGATTAGAGTCCCGCAATTATACATTTAATACGC
NOSR	TGTATAATTGCGGGACTCTAATCATAAAAACCC
PlAMVdsg1F	CGGTTCGTGACTTAGTCAAATTTGGTGGCTCACCTTTCCTTAAC
PlAMVdsg1R	AAATTTGACTAAGTCACGAACCGTTACC
LoLVdsg1F	GTCTTTTGGTTTTTGGTTACCATACCCACTAGTCGATTAAC
LoLVdsg1R	GTTAATCGACTAGTGGGTATGGTAACCAAAAACCAAAAGAC
PVMdsg1F	GTTAGTGTAGCTTACCAAGTGCTATTGTATTGAATATTTATGG
PVMdsg1R	ATAGCACTTGGTAAGCTACACTAACCAATCAGAGC
OSdsg1F	TACCTTATAGATTTGAATAAGATGGATATTCTC
OSdsg1R	TGATAACGGTTAAAGAAAGTTTCTGAGGC
35S-fluc-1F	GACTCTAGAGGATCCATGGAAGACGCCAAAAACATAAAGAAAGGC
35S-fluc-1653R	AGCCGGGCGGCCGCTTTACACGGCGATCTTTCCGCCCTTC
35S-down1F	AGCGGCCGCCCGGCTGCAGATCG
35S-up1R	GGATCCTCTAGAGTCGACTG
M1AF	GCGACTTCGAAAGTTTATGATCCAGAAC
M14AF	GCGATAACTGGTCCGCAGTGGTGG
M14FF	TTCATAACTGGTCCGCAGTGGTGG
M14LF	CTTATAACTGGTCCGCAGTGGTGG
M14PF	CCAATAACTGGTCCGCAGTGGTGG
M14SF	AGCATAACTGGTCCGCAGTGGTGG
M14TF	ACTATAACTGGTCCGCAGTGGTGG
M14VF	GTTATAACTGGTCCGCAGTGGTGG
M14R	CCGTTTCCTTTGTTCTGGATC
35S-4224F	TCATTTGGAGAGGACGATTAACATGGACATAGTCATCTCAGCCC
35S-6102R	AGCCGGGCGGCCGCTGGCCCACCAGACTTTCACTGGTCG
35S-up96R	GTCCTCTCCAAATGAAATGAACTTCC
Rluc-1F	ATGACTTCGAAAGTTTATGATCCAG
Rluc-936R	TTATTGTTCATTTTTGAGAACTCGC
35S-TGB15U-Rluc-R	GAAGTCATGTTAATCGTCCTCTCCAAATGAAATGAACTTCC
TGB25U-Rluc-R	AACTTTCGAAGTCATGTAGTTGGGCGCGGTGTCTGG
TGB35U-Rluc-R	AACTTTCGAAGTCATGAAGGTTGTCGCCGACGTGGGGGAG
Rluc-TGB13U-F	CAAAAATGAACAATAACTACGGCAAACCCGTTCTCGCTGCTTCC
Rluc-TGB23U-F	CAAAAATGAACAATAAACCCGGCTGCACCATTGTGATCACAGGG
Rluc-TGB33U-F	CAAAAATGAACAATAACGCCCTCAACGCGATGACCTTCG
KSF	GATTAACGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCATGGACATAGTCATCTCAGCC
KSRlucF	GATTAACGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCATGACTTCGAAAGTTTATGATCCAG
dstpF	CACCTACGGACTACGGCAAACCCGTTCTCGC
PlAMV-4913R	GGGTGAGGTGGTGGGCTCCGGAC
KZUF	GATTTACATGGACATAGTCATCTCAGCCC
KZCF	GATTCACATGGACATAGTCATCTCAGCCC
KZGF	GATTGACATGGACATAGTCATCTCAGCCC
KZURlucF	GATTTACATGACTTCGAAAGTTTATGATCCAG
KZCRlucF	GATTCACATGACTTCGAAAGTTTATGATCCAG
KZGRlucF	GATTGACATGACTTCGAAAGTTTATGATCCAG
1ntR	CGTCCTCTCCAAATGAAATGAACTTCC
3ntR	GTCGTCCTCTCCAAATGAAATGAACTTCC
5ntR	GTTATCGTCCTCTCCAAATGAAATGAACTTCC
TGB1F	ATGGACATAGTCATCTCAGCCC
dpF	GATTAACGATTAACATGGACATAGTCATCTC
dpRlucF	GATTAACGATTAACATGACTTCGAAAGTTTATGATCCAG
sub1F	GACTACAAGGACGACGACAAGGACTACAAGGACGACGACAAGATGGACATAGTCATCTCAGCCC
sub1RlucF	GACTACAAGGACGACGACAAGGACTACAAGGACGACGACAAGATGACTTCGAAAGTTTATGATCCAG
sub2F	GGCCGGGATGGACATAGTCATCTCAGCCC
sub2RlucF	GGCCGGGATGACTTCGAAAGTTTATGATCCAG
35S-flag-GFP1F	GGAAGTTCATTTCATTTGGAGAGGACGACTACAAGGACGACGACAAGATGGTGAGCAAGGGCGAGGAG

(Continued on next page)

TABLE 2 (Continued)

Primer name	Primer sequence
GFP-5U-TGB1R	GGGCTGAGATGACTATGTCCATGTTAATCTTACTTGTACAGCTCGTCCATGC
GFP-5U-RlucR	CTGGATCATAAACTTTCGAAGTCATGTTAATCTTACTTGTACAGCTCGTCCATGC
midAUGF	CTATGCCCCACAGCCGCCCTGGG
RlucF	GGGCCAGATGTAAACAAATGAATG
RlucR	ATTTGCCTGATTTGCCCATAC
FlucF	CGGAGGAGTTGTGTTTGTGG
FlucR	ATCTTTCCGCCCTTCTTGG
actin2F	GCACCCTGTTCTTCTTACCG
actin2R	AACCCTCGTAGATTGGCACA
PVX5435F	AATACATATCTCAACGCAATCATACTTG
PVX-RT7	TAATACGACTCACTATAGGGATTTATATTATTCATACAATCAAACCAGAAAATACTATGAAAC
LoLV6611F	ATGTCAGAATCCAAAGCAGAGAC
LoLVRT7	TAATACGACTCACTATAGGGGCTTTGACGGCAAAACCGAGGG
PVM7531F	TCGCTTGAGGCACTGAGCAG
PVM-RT7	TAATACGACTCACTATAGGGAGGCTAAAAATAGTTAAAAACCTAGTTTTATTTATAGTAG
PrGRR	GGGGTACCGCTACGTAACGGCATGACAGTGTTTTTTTTTTTTTTTTTTTTTTTTTTGGCCCACCAGACTTTCACTGGTCG
35S-Prsg1-5U1ntF	GTTCATTTCATTTGGAGAGGGATGGACATAGTCATCTCAGCCCT
35S-Prsg1-5U3ntF	GTTCATTTCATTTGGAGAGGGACATGGACATAGTCATCTCAGCCCT
35S-Prsg1-5U5ntF	GTTCATTTCATTTGGAGAGGGTAACATGGACATAGTCATCTCAGCCCT
35S-Prsg1-5U7ntF	GTTCATTTCATTTGGAGAGGGATTAACATGGACATAGTCATCTCAGCCCT
35S-Prsg1-5U10ntF	GTTCATTTCATTTGGAGAGGGATATATAACATGGACATAGTCATCTCAGCCCT
OS-TGB1F	ATGGATATTCTCATCATTAGTTTGAAAAGTTTAGG
OS-sg1-5U4ntR	CAAACTAATGATGAGAATATCCATCTTCAAAACTATAAGGTAACTTAACGG
OS-sg1-5U10ntR	TCAAACTAATGATGAGAATATCCATCTTAATATTCAAATCTATAAGGTAACTT
35S-OS-5U1nt-F	TTTCATTTGGAGAGGGATGGATATTCTCATCATTAGTTTGAAAAG
35S-OS-5U4nt-F	TTTCATTTGGAGAGGGAAGATGGATATTCTCATCATTAGTTTGAAAAG
35S-OS-5U7nt-F	TTTCATTTGGAGAGGGAATAAGATGGATATTCTCATCATTAGTTTGAAAAG
35S-OS-5U10nt-F	TTTCATTTGGAGAGGGAATATTAAGATGGATATTCTCATCATTAGTTTGAAAAG
OSGRR	GGGGTACCGCTACGTAACGGCATGACAGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATATTATTCATACAATCAAACCAG
Pr5010R	ACGTGGGGGAGCGTGGACCTGGTGGCGG
Pr5603R	GCGAGGGAGATGGTGGCT
PVMsg1R	ATTAGCCTTATGACCTAAAGGAAGCACAACC

described previously (52). An amino acid substitution Rluc series was constructed via PCR with 35S-Rluc and M1A as templates using the primer sets shown in Table 3.

(iv) 35S-sg1. The sgRNA1 region of the PlAMV genome (46) was amplified via PCR using the 35S-4224F and 35S-6102R primers. The plasmid fragment was obtained from 35S-sGFP (51) using the 35S-down1F and 35S-up96R primers. The PCR-amplified sgRNA1 sequence was inserted downstream of the 35S promoter of the plasmid fragments using a GeneArt seamless cloning and assembly kit.

(v) sg1-TGBp1_Rluc, sg1-TGBp2_Rluc, sg1-TGBp3_Rluc, and sg1-TGBp1_Rluc-M14L. The Rluc coding region was amplified from 35S-Rluc via PCR using the Rluc-1F and Rluc-936R primers. Plasmid fragments for sg1-TGBp1_Rluc, sg1-TGBp2_Rluc, and sg1-TGBp3_Rluc were obtained from 35S-sg1 using the primer sets Rluc-TGB13UF/35S-TGB15U-RlucR, Rluc-TGB23UF/TGB25U-RlucR, and Rluc-TGB33UF/TGB35U-RlucR, respectively. The PCR-amplified *RLUC* gene was inserted into each plasmid fragment using a GeneArt seamless cloning and assembly kit. The Rluc-M14L coding region was amplified from the M14L mutant via PCR using the Rluc-1F and Rluc-936R primers and inserted into the plasmid fragment for sg1-TGBp1_Rluc to construct sg1-TGBp1_Rluc-M14L.

(vi) Leader sequence variants KS, Δstp, KZ(-3A)U, KZ(-3A)C, KZ(-3A)G, 5U5nt, 5U3nt, 5U1nt, dp5U, 5Usub1, and 5Usub2. Fragments were amplified via PCR from 35S-sg1, sg1-TGBp1_Rluc, sg1-TGBp1_Rluc-M14L, sg1-TGBp2_Rluc, and sg1-TGBp3_Rluc using the primer sets shown in Table 3. Amplified fragments were self-ligated using a ligation-convenience kit.

(vii) TGBp2-KZ variants. Fragments were amplified via PCR from 35S-sg1, sg1-TGBp2_Rluc, and sg1-TGBp3_Rluc using the primer sets shown in Table 3. Amplified fragments were self-ligated using a ligation-convenience kit.

(viii) GFP-containing variants. The GFP coding region was amplified from the 35S-sGFP plasmid via PCR using the 35S-flag-GFP1F and GFP-5U-TGB1R primers. The plasmid fragments were obtained from 35S-sg1, sg1-TGBp2_Rluc, and sg1-TGBp3_Rluc using the TGB1F/35S-up96R primer set and ligated with the GFP fragment, resulting in GFP-sg1, GFP-sg1-TGBp2_Rluc, and GFP-sg1-TGBp3_Rluc, respectively. GFP-sg1-TGBp1_Rluc was constructed in the same manner using the 35S-flag-GFP1F/GFP-5U-Rluc1R primer set for the 35S-sGFP plasmid and the Rluc1F/35S-up96R primer set for sg1-TGBp1_Rluc. A GeneArt seamless cloning and assembly kit was used for ligation.

(ix) midAUG. A partial sgRNA1 sequence that includes midAUG mutations (midAUGinsert) was synthesized by Eurofins Genomics (Tokyo, Japan). midAUGinsert was ligated with plasmid fragments obtained via PCR from 35S-sg1, sg1-TGBp2_Rluc, and sg1-TGBp3_Rluc using the midAUGF/35S-up96R primer set. The sequence data are shown in Table 4.

TABLE 3 List of template and primer combinations used for PCR

Name	Template	F	R
Leader sequence variant of sgRNA1			
Δstp	35S-PlAMV-sg1	dstpF	PlAMV-4913R
KS-sg1	35S-PlAMV-sg1	KSF	35S-up96R
KZ(-3A)U	35S-PlAMV-sg1	KZUF	35S-up96R
KZ(-3A)C	35S-PlAMV-sg1	KZCF	35S-up96R
KZ(-3A)G	35S-PlAMV-sg1	KZGF	35S-up96R
5U5nt	35S-PlAMV-sg1	TGB1-1F	5ntR
5U3nt	35S-PlAMV-sg1	TGB1-1F	3ntR
5U1nt	35S-PlAMV-sg1	TGB1-1F	1ntR
dp5U	35S-PlAMV-sg1	dpF	35S-up96R
5Usub1	35S-PlAMV-sg1	sub1F	35S-up96R
5Usub2	35S-PlAMV-sg1	sub2F	35S-up96R
Leader sequence variant of sgRNA1-TGBp1_Rluc			
KS	sgRNA1-TGBp1_Rluc	KSRlucF	35S-up96R
KZ(-3A)U	sgRNA1-TGBp1_Rluc_M14L	KZURlucF	35S-up96R
KZ(-3A)C	sgRNA1-TGBp1_Rluc_M14L	KZCRlucF	35S-up96R
KZ(-3A)G	sgRNA1-TGBp1_Rluc_M14L	KZGRlucF	35S-up96R
5U5nt	sgRNA1-TGBp1_Rluc_M14L	Rluc-1F	5ntR
5U3nt	sgRNA1-TGBp1_Rluc_M14L	Rluc-1F	3ntR
5U1nt	sgRNA1-TGBp1_Rluc_M14L	Rluc-1F	1ntR
dp5U	sgRNA1-TGBp1_Rluc_M14L	dpRlucF	35S-up96R
5Usub1	sgRNA1-TGBp1_Rluc_M14L	sub1RlucF	35S-up96R
5Usub2	sgRNA1-TGBp1_Rluc_M14L	sub2RlucF	35S-up96R
Leader sequence variant of sgRNA1-TGBp2_Rluc			
KS	sgRNA1-TGBp2_Rluc	KSF	35S-up96R
KZ(-3A)U	sgRNA1-TGBp2_Rluc	KZUF	35S-up96R
KZ(-3A)C	sgRNA1-TGBp2_Rluc	KZCF	35S-up96R
KZ(-3A)G	sgRNA1-TGBp2_Rluc	KZGF	35S-up96R
5U5nt	sgRNA1-TGBp2_Rluc	TGB1-1F	5ntR
5U3nt	sgRNA1-TGBp2_Rluc	TGB1-1F	3ntR
5U1nt	sgRNA1-TGBp2_Rluc	TGB1-1F	1ntR
dp5U	sgRNA1-TGBp2_Rluc	dpF	35S-up96R
5Usub1	sgRNA1-TGBp2_Rluc	sub1F	35S-up96R
5Usub2	sgRNA1-TGBp2_Rluc	sub2F	35S-up96R
Leader sequence variant of sgRNA1-TGBp3_Rluc			
KS	sgRNA1-TGBp3_Rluc	KSF	35S-up96R
KZ(-3A)U	sgRNA1-TGBp3_Rluc	KZUF	35S-up96R
KZ(-3A)C	sgRNA1-TGBp3_Rluc	KZCF	35S-up96R
KZ(-3A)G	sgRNA1-TGBp3_Rluc	KZGF	35S-up96R
5U5nt	sgRNA1-TGBp3_Rluc	TGB1-1F	5ntR
5U3nt	sgRNA1-TGBp3_Rluc	TGB1-1F	3ntR
5U1nt	sgRNA1-TGBp3_Rluc	TGB1-1F	1ntR
dp5U	sgRNA1-TGBp3_Rluc	dpF	35S-up96R
5Usub1	sgRNA1-TGBp3_Rluc	sub1F	35S-up96R
5Usub2	sgRNA1-TGBp3_Rluc	sub2F	35S-up96R
TGBp2KZ variant of sgRNA1			
KZ(-3U)A	35S-PlAMV-sg1	TGB2ATGF	TGB2KZTAR
KZ(-3U)C	35S-PlAMV-sg1	TGB2ATGF	TGB2KZTCR
KZ(-3U)G	35S-PlAMV-sg1	TGB2ATGF	TGB2KZTGR
TGBp2KZ variant of sgRNA1-TGBp2_Rluc			
KZ(-3U)A	sgRNA1-TGBp2_Rluc	Rluc-1F	TGB2KZTAR
KZ(-3U)C	sgRNA1-TGBp2_Rluc	Rluc-1F	TGB2KZTCR
KZ(-3U)G	sgRNA1-TGBp2_Rluc	Rluc-1F	TGB2KZTGR
TGBp2KZ variant of sgRNA1-TGBp3_Rluc			
KZ(-3U)A	sgRNA1-TGBp3_Rluc	TGB2ATGF	TGB2KZTAR
KZ(-3U)C	sgRNA1-TGBp3_Rluc	TGB2ATGF	TGB2KZTCR
KZ(-3U)G	sgRNA1-TGBp3_Rluc	TGB2ATGF	TGB2KZTGR

(Continued on next page)

TABLE 3 (Continued)

Name	Template	F	R
Amino acid substitution Rluc mutants			
M1A	35S-Rluc	M1AF	35S-up1R
M14A	35S-Rluc	M14AF	M14R
M14F	35S-Rluc	M14FF	M14R
M14L	35S-Rluc	M14LF	M14R
M14P	35S-Rluc	M14PF	M14R
M14S	35S-Rluc	M14SF	M14R
M14T	35S-Rluc	M14TF	M14R
M14V	35S-Rluc	M14VF	M14R
M1A_M14A	M1A	M14AF	M14R
M1A_M14F	M1A	M14FF	M14R
M1A_M14L	M1A	M14LF	M14R
M1A_M14T	M1A	M14TF	M14R
M1A_M14V	M1A	M14VF	M14R
Leader sequence variants for agroinoculation			
PlAMV-sgRNA1-5U1nt	pPlAMV	35S-1ntF	GRR
PlAMV-sgRNA1-5U3nt	pPlAMV	35S-3ntF	GRR
PlAMV-sgRNA1-5U5nt	pPlAMV	35S-5ntF	GRR
PlAMV-sgRNA1-5U7nt	pPlAMV	35S-7ntF	GRR
PlAMV-sgRNA1-5U10nt	pPlAMV	35S-10ntF	GRR
PVX-sgRNA1-1nt	pPVX	35S-PVX1ntF	GRR
PVX-sgRNA1-4nt	pPVX	35S-PVX4ntF	GRR
PVX-sgRNA1-7nt	pPVX	35S-PVX7ntF	GRR
PVX-sgRNA1-10nt	pPVX	35S-PVX10ntF	GRR
PVX-GFP-5U4nt	pPVX-GFP	35SF	PVX4ntR
PVX-GFP-5U10nt	pPVX-GFP	35SF	PVX10ntR

(x) Leader sequence variants of PlAMV-sgRNA1 and PVX-sgRNA1 for agroinoculation (PlAMV-sgRNA1-5U1nt, -3nt, -5nt, -7nt, and -10nt; PVX-sgRNA1-1nt, -4nt, -7nt, and -10nt). The sgRNA1 sequences of pPlAMV and pPVX were amplified via PCR using the primer sets shown in Table 3. Plasmid fragments were obtained from pPlAMV using the GRF/35SR primer set. The fragments were ligated using a GeneArt seamless cloning and assembly kit.

(xi) Leader-sequence variants of PVX-GFP for agroinoculation (PVX-GFP-5U4nt and -10nt). The sgRNA1 sequence was amplified from pPVX-GFP via PCR using the primer sets shown in Table 3. The plasmid fragments were obtained from pPVX-GFP using the PVXTGB1F/35SR primer set. The fragments were ligated using a GeneArt seamless cloning and assembly kit.

Inoculation with viral isolates. Agroinoculation of PlAMV-GFP, PVX-GFP, PlAMV, PVX, LoLV, PepMV, WClMV, HdRSV, and PVM was performed as described previously (53). The leaf area with a GFP signal was quantitated using ImageJ v1.40 software (National Institutes of Health [NIH], Bethesda, MD, USA).

qRT-PCR. Total RNA was extracted from transfected protoplasts using Sepasol-RNA I solution (Nacalai Tesque, Kyoto, Japan). Total RNA was subjected to DNase I treatment (Roche, Basel, Switzerland) followed by reverse transcription using a high-capacity cDNA reverse transcription kit (Thermo Fisher Scientific). For quantitative reverse transcription-PCR (qRT-PCR), viral RNA was amplified using a thermal cycler dice real-time system (TaKaRa, Shiga, Japan) with SYBR Premix *Ex Taq* II (TaKaRa). We used specific primers, RlucF and RlucR for *Rluc* mRNA and PVXrtF and PVXrtR for PVX RNA. The accumulation of PVX RNA was normalized to *Rluc* mRNA to assess the activity levels and transformation efficiency in the protoplasts.

Northern blot analysis. Total RNA (1 μg) was analyzed using the digoxigenin (DIG) system (Roche). A DIG-labeled probe for PlAMV RNA detection was produced as described previously (17). Probes for PVX, LoLV, and PVM RNA detection were transcribed with T7 RNA polymerase from the DNA fragments amplified from each viral cDNA clone using the primer sets PVX5435F/PVX-RT7, LoLV6611F/LoLVRT7, and PVM7531F/PVMRT7, respectively.

Protoplast preparation and transfection. *Arabidopsis* suspension culture cells (54) were kindly provided by S. Hasezawa (University of Tokyo). Protoplast isolation from *Arabidopsis* suspension cells and transfection were performed as described by Abel and Theologis (55), with slight modifications. First, 20 mL of cell suspension was collected by centrifugation and incubated with 10 mL of enzyme solution (1% cellulase Onozuka RS, 0.1% pectolyase Y-23, 0.4 M mannitol, and 5 mM MES-KOH [pH 5.6] dissolved in cell culture) for 120 min at 25°C. The cells were filtered through 100-μm nylon mesh to separate the protoplasts and washed three times with W5 buffer (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, and 1.5 mM MES-KOH [pH 5.6]). The separated protoplasts were incubated on ice for 30 min before transfection. The protoplasts were counted in a hemocytometer and prepared at a density of 5×10^6 protoplasts/mL. The protoplasts were collected and resuspended in the same volume of MaMg solution (0.4 M mannitol, 15 mM $MgCl_2$, and 5 mM MES-KOH [pH 5.6]), and then 300 μL of protoplast solution was mixed with plasmid DNA (10 μg for the luciferase assay or 30 μg for immunoblotting). Next, 300 μL

TABLE 4 Sequence of midAUGinsert

midAUGinsert sequence
GGAAGTTCATTTCATTTGGAGAGGACGATTAACATGGACATAGTCATCTCAGCCCTAACATCTAACGACTTTCAAAGAACTAACACTCCCATCTCCAAGCCTCTA GTAGTTCACGCTGTAGCTGGCGCGGGGAAAACCACACTGATCCAGAACCTCCTCCCCGACCACCCCAACCTCTCCGCTCAGACCGCCGGAACCCCACAATCTCC AAACCTGACCGGGGCTTACATCCGGAAGCTGACATGCCCAGAACCTAACAAACTCAACCTCCTCGACGAGTACTCCGCCCTCCAACCACTCAAAGGGTCATGGG ACGTCGTGCTTGCCGACCCACTCCAACACTCAGGACTCGCGCTCCGACCCCACTTCGTAAAGTCAGTCAGTCATCGGCTATGCCCGGAGACGACCAAACTCATCT CCAAACTGATATGCCCATGCACCTCCAGCCGCTCTGACGCCTCAACCATCCAGTTCTCCGGACTGTTCGAAGGACCTCTCCTGGGCACCATCATTGCCCTCGACCC TACCACCCAAGCTCTCCTTCAAGCCCACGGAGCGCCCTTCCTATGCCCCACAGCCGCCCTGGG

of polyethylene glycol-CHS solution (0.4 M mannitol, 0.1 M $CaCl_2$, and 40% polyethylene glycol 4000) was added to the protoplast-plasmid mixture, which was incubated for 30 min at room temperature. The transfected protoplasts were washed with 5 mL of W5 buffer, resuspended in 2 mL of W5 buffer, and incubated in the dark at 23°C until subsequent analysis.

Protoplast isolation from *Nicotiana* suspension cells and transfection were performed as follows. First, 30 mL of cell suspension was collected by centrifugation and washed with 20 mL 0.4 M mannitol solution. Cells were incubated with 10 mL of BY-2 enzyme solution (1% cellulase Onozuka RS, 0.1% pectolyase Y-23, 0.4 M mannitol, adjusted to pH 5.5 by adding HCl) for 120 min at 25°C. The cells were washed three times with W5 buffer. The separated protoplasts were incubated on ice for 30 min before transfection. The protoplasts were counted in a hemocytometer and prepared at a density of 2×10^6 protoplasts/mL. The protoplasts were collected and resuspended in the same volume of MaMg solution, and then 300 μL of protoplast solution was mixed with 30 μg plasmid DNA. Next, 300 μL of polyethylene glycol solution [0.4 M mannitol, 0.1 M $Ca(NO_3)_2$, and 40% polyethylene glycol 4000] was added to the protoplast-plasmid mixture, which was incubated for 30 min at room temperature. The transfected protoplasts were washed with 5 mL of W5 buffer, resuspended in 2 mL of W5 buffer, and incubated in the dark at 25°C for 3 days until RNA extraction.

Immunoblotting. Agroinfiltrated leaves were harvested at 4 dpi, and total protein was extracted using radioimmunoprecipitation assay (RIPA) buffer (50 mmol/liter Tris-HCl [pH 8.0], 150 mmol/liter NaCl, 0.5, wt/vol%, sodium deoxycholate, 0.1, wt/vol%, sodium dodecyl sulfate [SDS], 1.0, wt/vol%, NP-40, and 100 mM dithiothreitol [DTT]). The transfected protoplasts were collected at 42 h postinfection (hpi), and total protein was extracted using buffer A (50 mM Tris-HCl [pH 7.5], 15 mM $MgCl_2$, 120 mM KCl, 10 mM DTT, 1 tablet Complete Mini protease inhibitor cocktail [Roche, Basel, Switzerland]/10 mL, and 20, vol/vol%, glycerol). The sample was centrifuged at 1,000 $\times$ *g* and 4°C for 10 min, and the supernatant was centrifuged again at 2,000 $\times$ *g* and 4°C for 10 min to remove cell debris. The collected supernatant was centrifuged at 30,000 $\times$ *g* and 4°C for 30 min to generate the supernatant fraction (S30) and the membrane-containing pellet fraction (P30). Protein samples were denatured in gel sample buffer (50 mM Tris-HCl [pH 6.8], 2%, wt/vol, SDS, 10% glycerol, 100 mM DTT) and separated on a 3 to 8% Tris-acetate NuPAGE gel (Thermo Fisher Scientific) for RdRp, a 4 to 12% Bis-Tris gel for TGB1 and CP, and a 12% Bis-Tris gel for TGBp2 and TGBp3 of each virus. After electrophoresis, proteins were blotted onto a polyvinylidene fluoride membrane and detected with Can Get Signal (Toyobo, Osaka, Japan). The membrane was stained with Coomassie brilliant blue as a loading control.

Luciferase assay. Transfected protoplasts were collected at 19 hpi, and luciferase protein was extracted using extraction buffer (0.1 M phosphate [pH 7.0] and 5 μM DTT). Luciferase activity was measured using a dual-luciferase reporter assay system (TOYO B-Net, Tokyo, Japan) according to the manufacturer's instructions.

5′ RACE. The TSSs of PlAMV sgRNA and PVM sgRNA1 were detected in 5′ RACE analysis using a GeneRacer kit (Invitrogen, Carlsbad, CA, USA) by following the manufacturer's instructions. Pr5010R was used as a specific primer for sgRNA1 and sgRNA2 of PlAMV, Pr5603R was used as a specific primer for sgRNA3 of PlAMV, and PVMsg1R was used as a specific primer for sgRNA1 of PVM.

The TSSs of sgRNA1 of PVX, HdRSV, PepMV, WCIMV, and LoLV were identified via sequencing on a MiSeq sequencing platform (Illumina, San Diego, CA, USA). Prior to the construction of the cDNA library, total RNA extracted from infected *N. benthamiana* was treated with CIP and TAP (Invitrogen). The cDNA libraries were constructed using the TruSeq small RNA sample prep kit (Illumina) by following a previous study (56). The 5′ end regions of the libraries were sequenced on MiSeq using a MiSeq reagent kit v2 (50 cycles) with a 50-bp single-end read protocol. Mapping of sequence reads to the virus genomes was conducted by following a previous study (56).

Data availability. PepMV and PVM, with accession number PV-0716 (Leibniz-Institut) and MAFF number 307027, respectively, were used for the analysis.

REFERENCES

1. Jackson RJ, Hellen CUT, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol 11:113–127. https://doi.org/10.1038/nrm2838.
2. Mouilleron H, Delcourt V, Roucou X. 2016. Death of a dogma: eukaryotic mRNA can code for more than one protein. Nucleic Acids Res 44:14–23. https://doi.org/10.1093/nar/gkv1218.
3. Kochetov AV, Sarai A, Rogozin IB, Shumny VK, Kolchanov NA. 2005. The role of alternative translation start sites in the generation of human protein diversity. Mol Genet Genomics 273:491–496. https://doi.org/10.1007/s00438-005-1152-7.
4. Kochetov AV, Prayaga PD, Volkova OA, Sankararamakrishnan R. 2013. Hidden coding potential of eukaryotic genomes: nonAUG started ORFs. J Biomol Struct Dyn 31:103–114. https://doi.org/10.1080/07391102.2012.691367.
5. Churbanov A, Rogozin IB, Babenko VN, Ali H, Koonin EV. 2005. Evolutionary conservation suggests a regulatory function of AUG triplets in 5′ UTRs

of eukaryotic genes. Nucleic Acids Res 33:5512–5520. https://doi.org/10.1093/nar/gki847.
6. von Arnim AG, Jia Q, Vaughn JN. 2014. Regulation of plant translation by upstream open reading frames. Plant Sci 214:1–12. https://doi.org/10.1016/j.plantsci.2013.09.006.
7. Kozak M. 1989. The scanning model for translation: an update. J Cell Biol 108:229–241. https://doi.org/10.1083/jcb.108.2.229.
8. Kozak M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292. https://doi.org/10.1016/0092-8674(86)90762-2.
9. Xu H, Wang P, You J, Zheng Y, Fu Y, Tang Q, Zhou L, Wei Z, Lin B, Shu Y, Zhu Y, Hu L, Kong X. 2010. Screening of Kozak-motif-located SNPs and analysis of their association with human diseases. Biochem Biophys Res Commun 392:89–94. https://doi.org/10.1016/j.bbrc.2010.01.002.
10. Ryabova LA, Pooggin MM, Hohn T. 2006. Translation reinitiation and leaky scanning in plant viruses. Virus Res 119:52–62. https://doi.org/10.1016/j.virusres.2005.10.017.
11. Simon A, Miller WA. 2013. 3 cap-independent translation enhencers of plant viruses. Annu Rev Microbiol 67:21–42. https://doi.org/10.1146/annurev-micro-092412-155609.
12. Gao F, Alekhina OM, Vassilenko KS, Simon AE. 2018. Unusual dicistronic expression from closely-spaced initiation codons of overlapping open reading frames in Pea enation mosaic virus 2. Nucleic Acids Res 46:11726–11742. https://doi.org/10.1093/nar/gky871.
13. Castaño A, Ruiz L, Hernández C. 2009. Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus. Virology 386:417–426. https://doi.org/10.1016/j.virol.2009.01.017.
14. Sonenberg N, Shatkin AJ, Ricciardi RP, Rubin M, Goodman RM. 1978. Analysis of terminal structures of RNA from potato virus X. Nucleic Acids Res 5:2501–2512. https://doi.org/10.1093/nar/5.7.2501.
15. Verchot-Lubicz J, Ye CM, Bamunusinghe D. 2007. Molecular biology of potexviruses: recent advances. J Gen Virol 88:1643–1655. https://doi.org/10.1099/vir.0.82667-0.
16. Verchot J, Angell SM, Baulcombe DC. 1998. In vivo translation of the triple gene block of potato virus X requires two subgenomic mRNAs. J Virol 72:8316–8320. https://doi.org/10.1128/JVI.72.10.8316-8320.1998.
17. Keima T, Hagiwara-Komoda Y, Hashimoto M, Neriya Y, Koinuma H, Iwabuchi N, Nishida S, Yamaji Y, Namba S. 2017. Deficiency of the eIF4E isoform nCBP limits the cell-to-cell movement of a plant virus encoding triple-gene-block proteins in Arabidopsis thaliana. Sci Rep 7:39678. https://doi.org/10.1038/srep39678.
18. Nijo T, Neriya Y, Koinuma H, Iwabuchi N, Kitazawa Y, Tanno K, Okano Y, Maejima K, Yamaji Y, Oshima K, Namba S. 2017. Genome-wide analysis of the transcription start sites and promoter motifs of phytoplasmas. DNA Cell Biol 36:1081–1092. https://doi.org/10.1089/dna.2016.3616.
19. González-Jara P, Atencio FA, Martínez-García B, Barajas D, Tenllado F, Díaz-Ruíz JR. 2005. A single amino acid mutation in the plum pox virus helper component-proteinase gene abolishes both synergistic and RNA silencing suppression activities. Phytopathology 95:894–901. https://doi.org/10.1094/PHYTO-95-0894.
20. Kim KH, Hemenway CL. 1999. Long-distance RNA-RNA interaction and conserved sequence elements affect potato virus X plus-strand RNA accumulation. RNA 5:636–645. https://doi.org/10.1017/s1355838299982006.
21. Pooggin MM, Rajeswaran R, Schepetilnikov MV, Ryabova LA. 2012. Short ORF-dependent ribosome shunting operates in an RNA picorna-like virus and a DNA pararetrovirus that cause rice tungro disease. PLoS Pathog 8:e1002568. https://doi.org/10.1371/journal.ppat.1002568.
22. Kozak M. 1989. Circumstances and mechanisms of inihibition of translation by secondary structure in eucaryotic mRNAs. Mol Cell Biol 11:5134–5142.
23. Kozak M. 2001. Constraints on reinitiation of translation in mammals. Nucleic Acids Res 29:5226–5232. https://doi.org/10.1093/nar/29.24.5226.
24. Luukkonen BG, Tan W, Schwartz S. 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. J Virol 69:4086–4094. https://doi.org/10.1128/JVI.69.7.4086-4094.1995.
25. Miyashita S, Kishino H. 2010. Estimation of the size of genetic bottlenecks in cell-to-cell movement of soil-borne wheat mosaic virus and the possible role of the bottlenecks in speeding up selection of variations in trans-acting genes or elements. J Virol 84:1828–1837. https://doi.org/10.1128/JVI.01890-09.
26. Hammond J, Reinsel MD, Maroon-Lango CJ. 2006. Identification and full sequence of an isolate of Alternanthera mosaic potexvirus infecting *Phlox stolonifera*. Arch Virol 151:477–493. https://doi.org/10.1007/s00705-005-0646-2.
27. Lee CC, Wang JW, Leu WM, Huang YT, Huang YW, Hsu YH, Meng M. 2019. Proliferating cell nuclear antigen suppresses RNA replication of Bamboo mosaic virus through an interaction with the viral genome. J Virol 93:e00961-19. https://doi.org/10.1128/JVI.00961-19.
28. Shrestha N, Bujarski JJ. 2020. Long noncoding RNAs in plant viroids and viruses: a review. Pathogens 9:765. https://doi.org/10.3390/pathogens9090765.
29. Kozak M. 1991. A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. Gene Expr 1:111–115.
30. Brito Querido J, Sokabe M, Kraatz S, Gordiyenko Y, Skehel JM, Fraser CS, Ramakrishnan V. 2020. Structure of a human 48S translational initiation complex. Science 369:1220–1227. https://doi.org/10.1126/science.aba4904.
31. Pisarev AV, Kolupaeva VG, Yusupov MM, Hellen CU, Pestova TV. 2008. Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. EMBO J 27:1609–1621. https://doi.org/10.1038/emboj.2008.90.
32. Kozak M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene 299:1–34. https://doi.org/10.1016/S0378-1119(02)01056-9.
33. Park MR, Jeong RD, Kim KH. 2014. Understanding the intracellular trafficking and intercellular transport of potexviruses in their host plants. Front Plant Sci 5:60. https://doi.org/10.3389/fpls.2014.00060.
34. Ju HJ, Samuels TD, Wang YS, Blancaflor E, Payton M, Mitra R, Krishnamurthy K, Nelson RS, Verchot-Lubicz J. 2005. The potato virus X TGBp2 movement protein associates with endoplasmic reticulum-derived vesicles during virus infection. Plant Physiol 138:1877–1895. https://doi.org/10.1104/pp.105.066019.
35. Aguilar E, Del Toro FJ, Brosseau C, Moffet P, Canto T, Tenllado F. 2019. Cell death triggered by the P25 protein in Potato virus X-associated synergisms results from endoplasmic reticulum stress in Nicotiana benthamiana. Mol Plant Pathol 20:194–210. https://doi.org/10.1111/mpp.12748.
36. Chou YL, Hung YJ, Tseng YH, Hsu HT, Yang JY, Wung CH, Lin NS, Meng M, Hsu YH, Chang BY. 2013. The stable association of virion with the triple-gene-block protein 3-based complex of Bamboo mosaic virus. PLoS Pathog 9:e1003405. https://doi.org/10.1371/journal.ppat.1003405.
37. Jang C, Seo EY, Nam J, Bae H, Gim YG, Kim HG, Cho IS, Lee ZW, Bauchan GR, Hammond J, Lim HS. 2013. Insights into alternanthera mosaic virus TGB3 functions: interactions with Nicotiana benthamiana PsbO correlate with chloroplast vesiculation and veinal necrosis caused by TGB3 overexpression. Front Plant Sci 4:5. https://doi.org/10.3389/fpls.2013.00005.
38. Lezzhov AA, Gushchin VA, Lazareva EA, Vishnichenko VK, Morozov SY, Solovyev AG. 2015. Translation of the shallot virus X TGB3 gene depends on non-AUG initiation and leaky scanning. J Gen Virol 96:3159–3164. https://doi.org/10.1099/jgv.0.000248.
39. Chen J, Shi YH, Adams MJ, Chen JP. 2005. The complete sequence of the genomic RNA of an isolate of Lily virus X (genus Potexvirus). Arch Virol 150:825–832. https://doi.org/10.1007/s00705-004-0441-5.
40. Oda Y, Mimura T, Hasezawa S. 2005. Regulation of secondary cell wall development by cortical microtubules during tracheary element differentiation in Arabidopsis cell suspensions. Plant Physiol 137:1027–1036. https://doi.org/10.1104/pp.104.052613.
41. Nagata T, Nemoto Y, Hasezawa S. 1992. Tobacco BY-2 cell line as the "HeLa" cell in the cell biology of higher plants. Int Rev Cytol 132:1–30. https://doi.org/10.1016/S0074-7696(08)62452-3.
42. Komatsu K, Hashimoto M, Maejima K, Shiraishi T, Neriya Y, Miura C, Minato N, Okano Y, Sugawara K, Yamaji Y, Namba S. 2011. A necrosis-inducing elicitor domain encoded by both symptomatic and asymptomatic Plantago asiatica mosaic virus isolates, whose expression is modulated by virus replication. Mol Plant Microbe Interact 24:408–420. https://doi.org/10.1094/MPMI-12-10-0279.
43. Ozeki J, Hashimoto M, Komatsu K, Maejima K, Himeno M, Senshu H, Kawanishi T, Kagiwada S, Yamaji Y, Namba S. 2009. The N-terminal region of the Plantago asiatica mosaic virus coat protein is required for cell-to-cell movement but is dispensable for virion assembly. Mol Plant Microbe Interact 22:677–685. https://doi.org/10.1094/MPMI-22-6-0677.
44. Okano Y, Senshu H, Hashimoto M, Neriya Y, Netsu O, Minato M, Yoshida T, Maejima K, Oshima K, Komatsu K, Yamaji Y, Namba S. 2014. In planta recognition of a double-stranded RNA synthesis protein complex by a potexviral RNA silencing suppressor. Plant Cell 26:2168–2183. https://doi.org/10.1105/tpc.113.120535.
45. Kagiwada S, Yamaji Y, Komatsu K, Takahashi S, Mori T, Hirata H, Suzuki M, Ugaki M, Namba S. 2005. A single amino acid residue of RNA-dependent RNA polymerase in the Potato virus X genome determines the symptoms in Nicotiana plants. Virus Res 110:177–182. https://doi.org/10.1016/j.virusres.2004.12.006.
46. Komatsu K, Yamaji Y, Ozeki J, Hashimoto M, Kagiwada S, Takahashi S, Namba S. 2008. Nucleotide sequence analysis of seven Japanese isolates of Plantago asiatica mosaic virus (PlAMV): a unique potexvirus with

significantly high genomic and biological variability within the species. Arch Virol 153:193–198. https://doi.org/10.1007/s00705-007-1078-y.
47. Minato N, Komatsu K, Okano Y, Maejima K, Ozeki J, Senshu H, Takahashi S, Yamaji Y, Namba S. 2014. Efficient foreign gene expression in planta using a plantago asiatica mosaic virus-based vector achieved by the strong RNA-silencing suppressor activity of TGBp1. Arch Virol 159:885–896. https://doi.org/10.1007/s00705-013-1860-y.
48. Ozeki J, Takahashi S, Komatsu K, Kagiwada S, Yamashita K, Mori T, Hirata H, Yamaji Y, Ugaki M, Namba S. 2006. A single amino acid in the RNA-dependent RNA polymerase of Plantago asiatica mosaic virus contributes to systemic necrosis. Arch Virol 151:2067–2075. https://doi.org/10.1007/s00705-006-0766-3.
49. Yusa A, Neriya Y, Hashimoto M, Yoshida T, Fujimoto Y, Hosoe N, Keima T, Tokumaru K, Maejima K, Netsu O, Yamaji Y, Namba S. 2019. Functional conservation of EXA1 among diverse plant species for the infection by a family of plant viruses. Sci Rep 9:5958–5958. https://doi.org/10.1038/s41598-019-42400-w.
50. Maejima K, Himeno M, Netsu O, Ishikawa K, Yoshida T, Fujita N, Hashimoto M, Komatsu K, Yamaji Y, Namba S. 2014. Development of an on-site plum pox virus detection kit based on immunochromatography. J Gen Plant Pathol 80:176–183. https://doi.org/10.1007/s10327-014-0504-8.
51. Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J. 1996. Engineered GFP as a vital reporter in plants. Curr Biol 6:325–330. https://doi.org/10.1016/S0960-9822(02)00483-9.
52. Hashimoto M, Neriya Y, Keima T, Iwabuchi N, Koinuma H, Hagiwara-Komoda Y, Ishikawa K, Himeno M, Maejima K, Yamaji Y, Namba S. 2016. EXA1, a GYF domain protein, is responsible for loss-of-susceptibility to plantago asiatica mosaic virus in *Arabidopsis thaliana*. Plant J 88:120–131. https://doi.org/10.1111/tpj.13265.
53. Yamaji Y, Maejima K, Ozeki J, Komatsu K, Shiraishi T, Okano Y, Himeno M, Sugawara K, Neriya Y, Minato N, Miura C, Hashimoto M, Namba S. 2012. Lectin-mediated resistance impairs plant virus infection at the cellular level. Plant Cell 24:778–793. https://doi.org/10.1105/tpc.111.093658.
54. Mathur J, Koncz C. 1998. Establishment and maintenance of cell suspension cultures. Methods Mol Biol 8:27–30.
55. Abel S, Theologis A. 1994. Transient transformation of Arabidopsis leaf protoplasts: a versatile experimental system to study gene expression. Plant J 5:421–427. https://doi.org/10.1111/j.1365-313x.1994.00421.x.
56. Hsu Y-H, Chen H-C, Cheng J, Annamalai P, Annamali P, Lin B-Y, Wu C-T, Yeh W-B, Lin N-S. 2006. Crucial role of the 5′ conserved structure of bamboo mosaic virus satellite RNA in downregulation of helper viral RNA replication. J Virol 80:2566–2574. https://doi.org/10.1128/JVI.80.5.2566-2574.2006.