Rescue of Tomato spotted wilt tospovirus entirely from cDNA clones, establishment of the first reverse genetics system for a segmented (-)RNA plant virus

Mingfeng Feng[a], Ruixiang Cheng[a], Minglong Chen[a], Rong Guo[a], Luyao Li[a], Zhike Feng[a], Jianyan Wu[a], Li Xie[b], Jian Hong[b], Zhongkai Zhang[c], Richard Kormelink[d] and Xiaorong Tao[a,1]

[a]Department of Plant Pathology, Nanjing Agricultural University, Nanjing 210095, P. R. China; [b]Analysis Center of Agrobiology and Environmental Sciences, Zhejiang University, Hangzhou 317502, P. R. China; [c]Yunnan Provincial Key Laboratory of Agri-Biotechnology, Institute of Biotechnology and Genetic Resources, Yunnan Academy of Agricultural Sciences, Kunming, Yunnan 650223, P. R. China; [d]Laboratory of Virology, Department of Plant Sciences, Wageningen University, 6708PB Wageningen, The Netherlands

Author contributions: M.F., Z.F. and X.T. conceived and designed the experiments and H.J., Z.Z. and R. K. provided input. M.F., R.C., M.C., G.R., L.L., Z.F., J.W., and X.L. performed the experiments. M.F., R. K. and X.T. wrote the manuscript.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: taoxiaorong@njau.edu.cn

Running title

Establishment of a reverse genetics system for TSWV

Abstract

The group of negative strand RNA viruses (NSVs) includes not only dangerous pathogens of medical importance but also serious plant pathogens of agronomical importance. Tomato spotted wilt tospovirus (TSWV) is one of those plant NSVs that cause severe diseases on agronomic crops and pose major threats to global food security. Its negative-strand segmented RNA genome has, however, always posed a major obstacle to molecular genetic manipulation. In this study, we report the complete recovery of infectious TSWV entirely from cDNA clones, the first reverse genetics (RG) system for a segmented plant NSV. First, a replication and transcription competent mini-genome replication system was established based on 35S-driven constructs of the $S_{(-)}$-genomic (g) or $S_{(+)}$-antigenomic (ag) RNA template, flanked by a 5' Hammerhead and 3' Ribozyme sequence of Hepatitis Delta virus, a nucleocapsid (N) protein gene and codon-optimized viral RNA dependent RNA polymerase (RdRp) gene. Next, a movement competent mini-genome replication system was developed based on $M_{(-)}$-gRNA, which was able to complement cell-to-cell and systemic movement of reconstituted ribonucleoprotein complexes (RNPs) of S RNA replicon. After further optimization, infectious TSWV and derivatives carrying eGFP reporters were successfully rescued *in planta* via simultaneous expression of full-length cDNA constructs coding for $S_{(+)}$-agRNA, $M_{(-)}$-gRNA and $L_{(+)}$-agRNA. Viral rescue occurred in the additional presence of various viral suppressors of RNAi, but TSWV NSs interfered with the rescue of genomic RNA. The establishment of a RG system for TSWV now allows detailed molecular genetic analysis of all aspects of tospovirus life

cycle and their pathogenicity.

Key words: Reverse genetics system, Tomato spotted wilt tospovirus, negative-strand RNA virus, tripartite RNA genome, mini-replicon, genome-length infectious cDNA clones

Significance

For many different animal-infecting segmented negative-strand viruses (NSVs), a reverse genetics system has been established that allows the generation of mutant viruses to study disease pathology and the role of *cis*- and *trans*-acting elements in the virus life cycle. In contrast to the relative ease to establish RG systems for animal-infecting NSVs, establishment of such system for the plant-infecting NSVs with a segmented RNA genome so far has not been successful. Here we report the first reverse genetics system for a segmented plant NSV, the Tomato spotted wilt tospovirus, a virus with a tripartite RNA genome. The establishment of this RG system now provides us with a new and powerful platform to study their disease pathology during a natural infection.

Introduction

Negative-stranded RNA viruses (NSVs) present of a large group of viruses that include well known members of medical importance such as Ebola (EBOV), Rabies (RV), Influenza A (FLUAV) and Rift valley fever virus (RVFV) (1, 2). Infections with these viruses may cause considerable morbidity and mortality in humans and form an important burden on national health care budgets. The group also contains plant viruses of agronomical importance such as Tomato spotted wilt virus (TSWV) and Rice stripe virus (RSV) that cause severe diseases on agronomic crops and pose major threats to global food security (3-10).

Tospoviruses belong to the NSV with a segmented (tripartite) RNA genome and rank among the most devastating plant viruses worldwide (11, 12). They are classified in the family of *Tospoviridae* within the order *Bunyavidales* (13). TSWV is the type member of the only genus *Orthotospovirus*, in the family of *Tospoviridae* (11, 13). TSWV has very broad host range infecting more than one thousand plant species over 80 families (14) and is transmitted by thrips in a persistent, propagative manner (6, 9, 15, 16). Crops losses due to this virus have been estimated more than one billion dollars annually (7, 14).

TSWV consists of spherical, enveloped virus particles (80-120 nm) that contain a tripartite genome consisting of a large- (L), medium- (M), and small-sized (S) RNA segment (11). The L segment is of entire negative polarity, whereas the S and M segments are ambisense. The L segment encodes the viral RNA-dependent RNA polymerase (RdRp, ~330 kDa) that is required for viral RNA replication and mRNA

89 transcription (17, 18). The viral (v) strand of the M segment encodes the precursor to
90 the glycoproteins (Gn and Gc, with n and c referring to the amino- and
91 carboxy-terminal end of the precursor, respectively) in the negative sense and a
92 nonstructural protein (NSm) in the positive sense. The glycoproteins are required for
93 particle maturation and are presented as spikes on the surface of the virus envelope
94 membrane (19, 20). They also play a major role as determinants for thrips vector
95 transmission (21). The NSm plays pivotal roles in cell-to-cell and long distance
96 movement of TSWV (22-26). The vRNA of the S segment codes for the nucleocapsid
97 protein (N) in the negative sense and a nonstructural protein (NSs) in the positive
98 sense. The N protein participates in the formation of ribonucleoprotein complexes
99 (RNPs) (27-29) and is required for viral intracellular movement (30, 31). The NSs
100 protein functions as a RNA silencing suppressor to defend against the plant innate
101 immunity system (32-34), and triggers a defence response and concomittant
102 programmed cell-death mediated by the dominant resistance gene *Tsw* from *Capsicum*
103 *chinense* (35-37).

104 As a virus documented for almost a century (7, 38), TSWV has served as an
105 important model for studying the molecular biology of tospovirus and other plant
106 NSVs with segmented genomes (6-8, 11). However, its negative-strand tripartite RNA
107 genome has posed a major obstacle to genetic manipulation of the virus. To initiate an
108 infection cycle with this virus requires at least RNPs, the minimal infectious units that
109 consist of viral RNA encapsidated by the N protein and associated with a few copies
110 of the viral RNA-dependent RNA polymerase (6, 11). TSWV RNPs can be

111 mechanically transferred from infected to healthy plants, however, transmission by
112 thrips requires RNPs to be enveloped and spiked with the glycoproteins (21).

113 The first animal-infecting and related counterpart of TSWV with a segmented
114 RNA genome to be rescued entirely from cDNA was from Bunyamwera virus in 1996
115 (39). Following this study, soon other segmented NSV were rescued from plasmid
116 DNA. The influenza virus, containing a genome of eight RNA segments, was
117 recovered in 1999 (40), while the first Arenavirus, with a bipartite RNA genome, was
118 recovered in 2006 (41). Just recently, the first nonsegmented plant NSVs from the
119 *Mononegavirales* have been rescued, *i.e.* the Sonchus yellow net nucleorhabdovirus
120 (SYNV) (42, 43). A recent study also reported on the establishment of a TSWV S
121 RNA-based mini-replicon in yeast (44), but in contrast to replication, no
122 transcriptional activity was observed.

123 In contrast to the relative ease to establish RG systems for animal-infecting NSVs,
124 reconstitution of infectious RNPs *in planta* for the plant-infecting viruses with a
125 segmented RNA genome seems particularly difficult. Not only have all DNA
126 constructs to be delivered into one and the same plant cell but for TSWV the RdRp is
127 exceptionally large (~330 kDa) compared to the RdRp of most other related
128 bunyaviruses (~240-260 kDa) and to typical open reading frames (ORFs) from the
129 plant genome. Expression of such a large protein gene may not only be very
130 inefficient but mRNA transcripts resulting from RNA polymerase II transcription of
131 35S promoter-constructs in the nucleus may also face splicing of cryptic splicing sites.
132 Moreover, achieving proper ratios of all three genome segments in plant cells is not

easily and consistently achieved by agrobacterium-mediated delivery of several constructs, which will affect the outcome of each individual experiment. All these obstacles may hamper the construction of a reverse genetics system for TSWV in plants.

In this study, we report the complete recovery of infectious TSWV entirely from cDNA clones in plants, the first reverse genetics system for a segmented plant NSV. The establishment of this system presents the start of a new research era for TSWV and provides us an entirely new and powerful platform to study the basic principles of the tospovirus life cycle and viral pathogenicity.

Results

Development of a TSWV $S_{(-)}$-genomic RNA mini-replicon system in *Nicotiana benthamiana*

Prior to the rescuing of TSWV entirely from cDNA clones, a mini-replicon system based on the S RNA-template was established. To this end, a DNA copy of the TSWV $S_{(-)}$-genomic RNA (gRNA) was cloned and flanked with a self-cleaving hammerhead (HH) ribozyme at the 5'-terminus and a hepatitis delta virus (HDV) ribozyme at the 3'-terminus. For visual monitoring, quantification purposes, and discrimination between primary and secondary genome transcription, the NSs and N genes were replaced with mCherry and eGFP, respectively (Fig. 1*A*). The resulting $S_{(-)}$ mini-replicon reporter was cloned in a binary vector pCB301 downstream a double 35S promoter (2×35S) and denoted 35S:$SR_{(-)mCherry\&eGFP}$ (Fig. 1*A*), or a T7 promoter

and denoted T7: $SR_{(-)mCherry\&eGFP}$ (Fig. S1*A*). The RdRp and N ORFs were amplified from cDNA of TSWV infected tissue and cloned into pCambia 2300 binary vector downstream a double 35S promoter. Binary vector constructs of the RdRp, N and four viral RNA silencing suppressor genes (VSRs: NSs from TSWV, P19 from Tomato bushy stunt tombusvirus (TBSV), HcPro from Tobacco etch potyvirus (TEV) and γb from Barley yellow mosaic hordeivirus (BYMV)) were agroinfiltrated in *N. benthamiana* either with 35S:$SR_{(-)mCherry\&eGFP}$ or with T7:$SR_{(-)mCherry\&eGFP}$ and a 35S-driven T7 RNA polymerase gene construct and next monitored for eGFP fluorescence. However, during repeated experiments no eGFP fluorescence was observed from 35S:$SR_{(-)mCherry\&eGFP}$ nor T7:$SR_{(-)mCherry\&eGFP}$ (Fig. S1*B* and *C*).

The possibility of failures to establish a mini-replicon system for TSWV could be due to low (unstable) expression of the TSWV RdRp protein, therefore, the codon usage of the RdRp gene was optimized for *N. benthamiana* and potential intron splicing sites were removed. The optimized RdRp gene ($RdRp_{opt}$) was cloned in a binary, 35S-driven expression vector and next, again agroinfiltrated in *N. benthamiana* leaves together with binary expression constructs of the N gene, the four VSR gene constructs and the 35S:$SR_{(-)mCherry\&eGFP}$ mini-replicon reporter. At 5 days post infiltration (dpi), expression of the reporter genes was analyzed by monitoring for mCherry and eGFP fluorescence in the *N. benthamiana* leaves (Fig. 1*A* and *B*). Whereas no fluorescence was observed in the controls, *i.e.* leaves agroinfiltrated with 35S:$SR_{(-)mCherry\&eGFP}$ alone or co-expressing 35S:$SR_{(-)mCherry\&eGFP}$ with $RdRp_{opt}$ or N only, eGFP and mCherry fluorescence was consistently observed in leaves

177 agroinfiltrated with 35S:SR$_{(-)mCherry\&eGFP}$ and both N and RdRp$_{opt}$ (Fig. 1*C*). This was
178 confirmed by Western immunoblot analysis (Fig. 1*D*).

179 Northern blot analysis showed that both the genomic RNA and anti-genomic
180 RNA of the SR$_{(-)mCherry\&eGFP}$ mini-replicon were detected in the leaves that
181 co-expressed both N and RdRp (here after RdRp represent optimized RdRp) at 5 dpi,
182 but not in the leaves co-expressing RdRp or N only (Fig. 1*E*). In addition, using an
183 anti-sense eGFP probe genome length S RNA subgenomic-sized RNA, likely
184 presenting eGFP transcripts, were detected (Fig. 1*E*, upper panel). A time course
185 analysis showed that eGFP and mCherry fluorescence in *N. benthamiana* leaves was
186 visible from 3 dpi onwards and gradually increased to 12 dpi (Fig. S2*A*). This was
187 also confirmed by immunoblot analysis (Fig. S2*B*).

188 Altogether the results indicated that in *N. benthamiana* the 35S replicon
189 transcript SR$_{(-)mCherry\&eGFP}$ was properly processed by the HH and RZ, and next used
190 by the RdRp as template for primary transcription of mRNA^{eGFP}, as well replicated
191 into SR$_{(-)mCherry\&eGFP}$ for secondary transcription of mRNA^{mCherry}. Furthermore, the
192 codon-optimized RdRp clearly exhibited full functionality in supporting viral genome
193 transcription and replication, while the wild type RdRp for unknown reasons didn't.
194 When the T7:SR$_{(-)mCherry\&eGFP}$ mini-replicon reporter was co-expressed with T7 RNA
195 polymerase, RdRp, N and four VSRs in *N. benthamiana* leaves, somewhat
196 unexpected, no mCherry or eGFP fluorescence was detected (Fig. S1*C* and *D*).

197

198 **The optimization of the concentration of N and RdRp, and VSRs on TSWV**

$SR_{(-)mCherry\&eGFP}$ mini-replicon

Having established a TSWV $S_{(-)}$-gRNA based mini-replicon system in *N. benthamiana*, attempts were made to further optimize the system. To this end, binary constructs of the TSWV $S_{(-)}$ mini-replicon were agroinfiltrated into *N. benthamiana* in the additional presence of varying amounts of N and RdRp gene expression constructs. During these experiments the amounts of *Agrobacterium* carrying the N expression construct were increased (from OD_{600} 0.2 to 0.8) while the *Agrobacterium* harboring the RdRp expression construct was kept fixed at OD_{600} 0.2, and vice versa. The results showed highest eGFP reporter gene expression from the 35S:$SR_{(-)mCherry\&eGFP}$ mini-replicon when *Agrobacterium* suspensions bearing N and RdRp expression constructs were both infiltrated at OD_{600} 0.2. When *Agrobacterium* harboring either N or RdRp was infiltrated onto *N. benthamiana* leaves at OD_{600} >0.4 the expression of eGFP from the $S_{(-)}$-mini-replicon greatly decreased (Fig. 2 *A*, *B* and *C*). Furthermore, at high concentrations (OD_{600} > 0.6) of *Agrobacterium*, the visible cell death was triggered in the infiltrated leaves (data not shown).

Next, in a similar approach and using the optimized setting, the effects of VSRs on the replication and transcription of the $SR_{(-)mCherry\&eGFP}$ mini-replicon were investigated. Without the addition of VSRs, mCherry and eGFP fluorescence was only observed in a small number of cells, but these numbers increased in the addition of TSWV NSs and/or the three VSRs P19, HcPro and γb. The largest number of cells showing eGFP expression from the $S_{(-)}$ mini-replicon, as monitored by fluorescence, were obtained when all four VSRs were added (Fig. 2*D* and *E*). These observations

were further confirmed by immunoblot assays (Fig. 2*E*).

The role of *cis*-acting sequences in transcription and replication of the TSWV SR$_{(-)mCherry\&eGFP}$ replicon

Using the optimized SR$_{(-)mCherry\&eGFP}$ mini-replicon system, the role of the 5'-untranslated region (5'-UTR), 3'-UTR and the non-coding A/U-rich intergenic region (IGR) between NSs and N genes (45, 46) in replication-transcription was examined. To this end, SR$_{(-)mCherry\&eGFP}$ derivatives were made from which the 5'-UTR, IGR or 3'-UTR, respectively, were deleted and next tested on transcription-replication using the mini-replicon assay (Fig. S3*A*). No eGFP and mCherry fluorescence was observed when the 5'-UTR or 3'-UTR of SR$_{(-)mCherry\&eGFP}$ was removed. However, eGFP reporter expression could still be detected when the IGR of SR$_{(-)mCherry\&eGFP}$ was deleted (Fig. S3*B*). Immunoblot analysis confirmed the expression of eGFP from SR$_{(-)mCherry\&eGFP\Delta IGR}$ (ΔIGR), and lack of expression from SR$_{(-)mCherry\&eGFP\Delta 5'UTR}$ (Δ5'UTR) and SR$_{(-)mCherry\&eGFP\Delta 3'UTR}$ (Δ3'UTR) (Fig. S3*C*). To analyze whether the lack of eGFP expression was a matter of translation or transcription-replication, samples from infiltrated *N. benthamiana* were collected and analyzed by Northern blot. The results showed that for SR$_{(-)mCherry\&eGFP\Delta 5'UTR}$ (Δ5'UTR) and SR$_{(-)mCherry\&eGFP\Delta 3'UTR}$ (Δ3'UTR), weak signals of genomic RNA could be detected but they are not similar to the signal strength obtained for the genome length and mRNA molecules as seen with the SR$_{(-)mCherry\&eGFP}$ replicon, while anti-genomic RNAs could not be detected (Fig. S3*D*). For SR$_{(-)mCherry\&eGFP\Delta IGR}$ (ΔIGR) both RNA strands were detected (Fig. S3*D*), suggesting that IGR is not essential for viral RNA synthesis,

while no signal is obtained for the mRNA length molecules as seen with the $SR_{(-)mCherry\&eGFP}$ replicon. Taken together, these findings suggest that the 5'-UTR and 3'-UTR play an essential role in viral transcription and replication of TSWV RNA segments.

Development of a TSWV $S_{(+)}$-agRNA based mini-replicon system

For many reverse genetics systems of NSV, mini-replicons have initially been established based on gRNA (vRNA). Here, we managed to develop a first system for TSWV based on antigenomic (ag)RNA (vcRNA). In order to analyze whether a system could be developed based on agRNA, a $S_{(+)}$-agRNA mini-replicon was constructed similarly to the one based on $S_{(-)}$-gRNA but in which the N gene was maintained and only the NSs gene was replaced by eGFP, denoted $SR_{(+)eGFP}$ (Fig. 3*A*). In analogy to the replicon assays with $SR_{(-)mCherry\&eGFP}$, *N. benthamiana* leaves were agro-infiltrated with binary expression constructs of $SR_{(+)eGFP}$, four VSRs and either N or RdRp separately or together, respectively, and monitored for eGFP fluorescence. Whereas eGFP fluorescence, resulting from primary transcription of the replicon transcript by viral RdRp, was not detected when $SR_{(+)eGFP}$ was expressed alone or in the additional presence of N, eGFP fluorescence was observed when $SR_{(+)eGFP}$ was co-expressed with both RdRp and N, but also when $SR_{(+)eGFP}$ was co-expressed with RdRp alone (Fig. 3*B*). This strongly indicated that a certain (residual) amount of $SR_{(+)eGFP}$ transcripts, resulting from 35S transcription, did not become fully processed by the HH and RZ and remained functional in translation, thereby giving rise to N

265 protein. This was confirmed by Western immunoblot analysis (Fig. 3*C*). Northern blot
266 analysis furthermore showed that samples from the replicon assays performed in the
267 presence of RdRp and N or RdRp alone, besides eGFP mRNA transcripts, also
268 contained agRNA and gRNA of $SR_{(+)eGFP}$, indicating the occurrence of replication
269 (Fig. 3*D*). Altogether, these results demonstrate that the N protein can also be
270 expressed from the $SR_{(+)eGFP}$ replicon to support its transcription and replication. This
271 provides an attractive alternative to the $S_{(-)}$-gRNA based mini-replicon as additional
272 binary expression constructs for N do not have to be supplied anymore.
273
274 **Development of a $M_{(-)}$-gRNA based mini-replicon for cell-to-cell movement of**
275 **TSWV in *N. benthamiana***
276 As a first step towards development of a reverse genetics system to rescue TSWV
277 virus entirely from cDNA, a movement competent mini-replicon was also established.
278 To this end, a TSWV $M_{(-)}$-gRNA based mini-replicon was constructed, similar as the
279 ones made for $S_{(-)}$ and $S_{(+)}$.Within this construct the NSm cell-to-cell movement
280 protein gene was maintained but the GP ORF was exchanged for eGFP, resulting in a
281 mini-replicon designated as $MR_{(-)eGFP}$ (Fig. 4*A*). After *Agrobacterium*-mediated
282 delivery into *N. benthamiana*, no eGFP fluorescence was observed in leaves
283 containing the $MR_{(-)eGFP}$ replicon with RdRp or N. However, in the presence of both
284 RdRp and N, eGFP fluorescence was observed in many cells that connected each
285 other (Fig. 4*B* and *C*). In comparison to the eGFP fluorescence always expressed in a
286 single plant cells from $SR_{(-)mCherry\&eGFP}$ or $SR_{(+)eGFP}$ reporter, the results suggested that

287 the $MR_{(-)eGFP}$ mini-replicon has moved from cell-to-cell in *N. benthamiana* leaves.
288 Northern blot analysis confirmed the synthesis of gRNA, agRNA and
289 (subgenomic-length) eGFP mRNA transcripts of the $MR_{(-)eGFP}$ replicon in the
290 presence of both RdRp and N, but not with RdRp or N only (Fig. 4*D*).

291 To further substantiate the findings on possible cell-to-cell movement of the
292 $MR_{(-)eGFP}$ mini-replicon, a stop codon was introduced immediately downstream the
293 start codon of NSm and the construct designated $MR_{(-)eGFP\&NSmMut}$) (Fig. 4*A*). When
294 the $MR_{(-)eGFP\&NSmMut}$ replicon was delivered and co-expressed with RdRp and N in *N.*
295 *benthamiana* leaves, eGFP fluorescence was only detected in a single cells (Fig. 4*B*).
296 As expected, Western immunoblot analysis confirmed the production of eGFP protein
297 in leaves containing the $MR_{(-)eGFP\&NSmMut}$ replicon, and in significantly lower amounts
298 compared to the $MR_{(-)eGFP}$ replicon (Fig. 4*C*).

299

300 **Establishment of the systemic infection of $M_{(-)}$- and $S_{(+)}$-mini-replicon reporters**
301 **by co-expression of full-length antigenomic $L_{(+)}$ in *N. benthamiana***

302 With the establishment of S (g/ag)RNA-based mini-replicon systems, and a
303 movement-competent M gRNA-based mini-replicon, we set out to construct full
304 length genomic cDNA clones of $L_{(-)}$, $M_{(-)}$ and $S_{(-)}$, flanked by HH and HDV at 5'- and
305 3'-terminus, as a first step towards the rescue of TSWV entirely from cDNA clones.
306 At the same time, similar constructs were made for the anti-genomic $L_{(+)}$, $M_{(+)}$ and
307 $S_{(+)}$. However, attempts to recover infectious TSWV from *N. benthamiana* after
308 agrobacterium-mediated delivery of all binary expression constructs of $L_{(-)}$, $M_{(-)}$, $S_{(-)}$

309 together with N, RdRp and four VSRs, but also with the anti-genomic $L_{(+)}$, $M_{(+)}$ and
310 $S_{(+)}$ constructs, all failed (Table 1).

311 Since $MR_{(-)eGFP}$ was earlier shown to be movement competent, it was next
312 investigated whether the $M_{(-)}$- and $S_{(-)}$-minigenomes moved into the same plant cell in
313 the presence of both RdRp and N. Upon co-expression of $MR_{(-)mCherry}$, $SR_{(-)eGFP}$, RdRp
314 and N in *N. benthamiana* leaves, expression of both mCherry and eGFP from the
315 $MR_{(-)mCherry}$ and $SR_{(-)eGFP}$ mini-replicons, respectively, could be discerned. However,
316 the foci of mCherry fluorescence were separate from those showing eGFP
317 fluorescence (Fig. S4). Previously, ectopic expression of Tobacco crinkle virus (TCV)
318 RdRp was reported to cause superinfection exclusion, and prevented the entry of
319 progeny virus into the original cell expressing the RdRp (47). Ectopic expression of
320 TSWV RdRp and N would possibly cause superinfection exclusions and block
321 intercellular movement of $SR_{(-)eGFP}$ into cells containing $MR_{(-)mCherry}$, To avoid that, a
322 new strategy was employed in which RdRp and N were expressed from viral agRNAs.
323 To this end, a construct was made of the full-length L agRNA containing the
324 optimized RdRp and flanked by the HH and HDV ribozymes, denoted $L_{(+)opt}$ (Fig. 5*A*).
325 To test the expression and functionality of RdRp from this construct, $L_{(+)opt}$ was
326 co-expressed with the $SR_{(-)mCherry\&eGFP}$ mini-replicon, N and VSRs in *N. benthamiana*
327 leaves. The results showed clear eGFP and mCherry fluorescence and indicated that
328 $L_{(+)opt}$ was able to support $SR_{(-)mCherry\&eGFP}$ transcription and replication (Fig. S5*A*).
329 Furthermore, $L_{(+)opt}$ was also able to support the replication and transcription of the
330 $SR_{(+)eGFP}$ mini-replicon, without the additional ectopic expression of N (Fig. S5*B*),

and of the movement competent $MR_{(-)eGFP}$ (Fig. S5*C*).

In a next experiment $L_{(+)opt}$ was co-expressed with $MR_{(-)mCherry}$, $SR_{(+)eGFP}$ and four VSRs in *N. benthamiana* and plants analyzed for a systemic infection (Fig. 5*A*). At 6 dpi, mCherry and eGFP fluorescence were detected in the locally agroinfiltrated *N. benthamiana* leaves and in which some foci were found to express mCherry and eGFP together (Fig. S6*A*). At 15 dpi, necrotic symptoms became visual in systemic leaves of *N. benthamiana* (Fig. 5*B, a* and *c*). Using a handheld UV lamp, a clear eGFP fluorescence was also observed in those leaves (Fig. 5*B, b* and *d*). The eGFP signal was detected in 24 out of 30 agro-infiltrated *N. benthamiana* plants (Table 1). Both eGFP and mCherry fluorescence were detected in veins/stems and systemic leaves under a fluorescence microscope (Fig. 5*B, e*). The systemic infection of *N. benthamiana* with $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$ was further confirmed by RT-PCR analysis (Fig. S6*B*).

Recovery of infectious TSWV from the full-length cDNA clones

Based on the establishment of a systemic infection after *Agrobacterium*-mediated delivery of replicons $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$, we next generated a full-length construct for S agRNA without any reporter gene, designated as $S_{(+)}$ and co-expressed it with replicon constructs $L_{(+)opt}$, $M_{(-)}$ and four VSRs in *N. benthamiana* leaves. However, and surprisingly, no infectious TSWV was recovered from systemic leaves of *N. benthamiana* that were infiltrated with these constructs (Table 1). To find out whether this was due to failure of $S_{(+)}$, we next examined whether $S_{(+)}$ was able to

353 establish a systemic infection in combination with the functional $MR_{(-)eGFP}$ and $L_{(+)opt}$
354 constructs. When $L_{(+)opt}$, $MR_{(-)eGFP}$, $S_{(+)}$ and three VSRs (P19-HcPro-γb) were
355 co-expressed in *N. benthamiana* leaves, eGFP fluorescence was visible at 18 dpi in
356 systemic leaves of *N. benthamiana*, although not as efficient as in the case with the
357 $SR_{(+)eGFP}$ replicon, since only 7 out of 60 plants showed systemic infection (Fig. 5*D*
358 and Table 1). RT-PCR analysis confirmed the systemic infection with $S_{(+)}$, $MR_{(-)eGFP}$
359 and $L_{(+)opt}$ in those *N. benthamiana* (Fig. S6*C*). When $L_{(+)opt}$, $MR_{(-)eGFP}$ and $S_{(+)}$ were
360 co-expressed with four VSRs (P19-HcPro-γb and NSs) in *N. benthamiana* leaves,
361 intriguingly, no eGFP fluorescence was observed in the systemic leaves (Table 1)
362 suggesting that ectopic expression of NSs interfered with the rescue of virus from full
363 length $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$.

364 Next, we tested the rescuing of virus from full length $M_{(-)}$, $SR_{(+)eGFP}$ and $L_{(+)opt}$. To
365 this end, the constructs of $L_{(+)opt}$, $M_{(-)}$ and $SR_{(+)eGFP}$ were delivered and co-expressed
366 in *N. benthamiana* in the presence of either four (P19-HcPro-γb+NSs) or three
367 (P19-HcPro-γb) VSRs. The results showed no eGFP fluorescence in systemic leaves
368 of *N. benthamiana* at 15-50 dpi, indicating that $M_{(-)}$ was not able to complement and
369 rescue the $S_{(+)}$-mini-replicons into systemic leaves (Table 1). Northern blot analysis
370 showed that neither gRNAs nor agRNAs were detected for $M_{(-)}$ while, in contrast,
371 gRNAs and agRNAs were detected for $S_{(+)}$ (Fig. S7*A* and *B*). Earlier, the $MR_{(-)eGFP}$
372 mini-replicon was shown to replicate and transcribe (Fig. 4*D*). The only difference
373 between $M_{(-)}$ and $MR_{(-)eGFP}$ mini-replicon was the GP gene, which was exchanged for
374 eGFP in the second construct. Considering that primary $M_{(-)}$ transcripts were

375 produced in the nucleus by the 35S promoter, and putative splice sites were also
376 predicted in the GP sequence (*SI Appendix*, Table S2), it was likely that primary $M_{(-)}$
377 transcripts were prone to splicing before sufficient replication of the mini-replicon
378 and transcriptional-translational expression of the cell-to-cell movement protein gene
379 could take place. For this reason, codon optimization was performed on the GP gene
380 sequence in $M_{(-)}$, leading to a new construct designated as $M_{(-)opt}$ (Fig. 5*A*). Upon
381 co-expression of $L_{(+)opt}$, $M_{(-)opt}$ and the $SR_{(+)eGFP}$ mini-replicon in *N. benthmiana*
382 eGFP fluorescence was observed in systemic leaves (Fig. 5*C*). Fluorescence was
383 observed in 27 out of 30 agroinfiltrated plants, and demonstrated that $M_{(-)opt}$
384 produced a functional and stable M genomic RNA, able to replicate and support
385 systemic movement of S and L RNP molecules by its encoded NSm protein (Table 1).
386 RT-PCR analysis further confirmed a systemic infection of *N. benthamiana* with
387 $SR_{(+)eGFP}$, $M_{(-)opt}$ and $L_{(+)opt}$ (Fig. S6*D*).

388 In a final experiment, aiming to rescue "wild type" TSWV entirely from cDNA
389 clones, the binary constructs of $L_{(+)opt}$, $M_{(-)opt}$ and $S_{(+)}$ were agroinfiltrated into *N.*
390 *benthamiana* leaves together with three VSRs (P19-HcPro-γb). At 19 dpi, typical leaf
391 curling was observed in systemic leaves from *N. benthamiana* plants (Fig. 6*A* and
392 Table 1). Upon disease progression, plants started to exhibit a stunted phenotype
393 between 19-30 dpi. When the experiment was repeated with a large batch of plants, a
394 systemic infection was observed in 6 out of 60 plants (Table 1). Northern blot
395 analyses on samples collected from systemically infected leaves showed the presence
396 of gRNA and agRNA of S, M and L RNA segments (Fig. 6*B*), which was also

confirmed by RT-PCR (Fig. S8). Moreover, sequence analysis of the amplicons derived from the L and M RNA confirmed the presence of codon optimized RdRp and GP gene sequences (Fig. 6*C*). Immunoblot analysis on systemically infected leaf samples showed the presence of N, NSs, NSm, Gn and Gc proteins *N. benthamiana* (Fig. 6*D*), altogether indicating a successful systemic infection with rescued TSWV (rTSWV).

To demonstrate genuine virus particle rescue of rTSWV, samples were collected from newly infected systemic leaf tissues and subjected to transmission electron microscopy. As shown in Fig. 6*E*, typical enveloped and spherical virus particles were observed in rTSWV-infected tissue, altogether indicating that infectious TSWV (rTSWV) was successfully rescued from full-length cDNA clones of $L_{(+)opt}$, $M_{(-)opt}$ and $S_{(+)}$.

Discussion

The establishment of a reverse genetics system for a segmented NSV basically requires two steps. The first one involves the *in vivo* reconstitution of transcriptionally active RNPs, often managed by development of a mini-genome replication system. The second step involves virus rescue entirely from full-length "infectious" cDNA clones, based on tools developed and optimized with the mini-genome replication system. In this study, we first successfully reconstituted infectious RNPs based on TSWV $S_{(-)}$-gRNA and $S_{(+)}$-agRNA after having optimized the sequence of RdRp. Next, a movement competent mini-genome replication system was developed based

419 on $M_{(-)}$-gRNA, which was also able to complement and systemically rescue
420 reconstituted S RNPs. In a third step, full length constructs were made for
421 $S_{(+)eGFP}$-agRNA, $M_{(-)mCherry}$-gRNA and $L_{(+)}$-agRNA, to directly accommodate for
422 translation of (small amounts of) all three genomic (35S) transcripts into N, NSm and
423 RdRp proteins, respectively, and leave out the additional need of ectopically
424 expressed N and RdRp. *Agrobacterium*-mediated delivery of these constructs lead to a
425 successful systemic infection of *N. benthamiana* with rTSWV carrying eGFP
426 reporters. In a last step, the GP gene sequence of $M_{(-)}$ was optimized, that allowed the
427 final rescuing of infectious rTSWV particles entirely from full-length cDNA clones in
428 *N. benthamiana*.

429 The genomic RNAs of segmented NSVs possess neither a 5' cap-structure nor
430 3'-poly(A) tail (2, 48). Instead, their termini contain highly conserved sequences that
431 show inverted sequence complementary and fold into a panhandle structure with a
432 major role in RNA transcription and replication. Any additional nucleotide residues at
433 those termini in the past have been shown to disrupt/affect transcription-replication of
434 animal-infecting segmented NSVs (49). For this reason, the choice of plant promoter
435 to generate the first primary full-length genomic RNA templates (mimicking authentic
436 genomic RNA molecules) for initiating viral replication is one of the major and
437 critical factors for the construction of a reverse genetics system for TSWV. For
438 animal-infecting segmented NSVs, researchers in the past have been using various
439 systems. One of the first strategies employed bacteriophage T7 promoter constructs
440 co-expressed with a T7 RNA polymerase and later followed by the use RNA

polymerase I (Pol I) promoter constructs to generate the initial viral genome length RNA transcripts in mammalian cells (39, 40, 50-52). Unfortunately, attempts to establish the TSWV mini-replicon system based on the T7 promoter and T7 RNA polymerase strategy was unsuccessful (Fig. S1*A*, *C* and *D*). The activity of the Pol I promoter was shown to be species-dependent (53). Although a Pol I promoter has been reported from *Arabidopsis* (54, 55), while the transcription initiation +1 site is still not known. For *N. benthamiana* no Pol I promoter has been characterized yet. The 35S promoter, an RNA Pol II promoter, is well characterized and hence remains the only choice to establish a reverse genetics system for TSWV in plants. This is be in contrast to reverse genetics of segmented NSVs in animal cells, where all viruses have been reconstituted after T7/Pol I driven production of primary viral RNA templates for replication. The Pol II promoter has been used to produce the initial viral RNA transcripts of an animal-infecting nonsegmented NSV (56). Recently, the 35S/Pol II promoter was also successfully employed to produce primary viral RNA template of the first non-segmented plant NSV reconstituted, the SYNV rhabdovirus (42, 43). Here, we successfully deployed the 35S/Pol II promoter and two ribozymes at 5' and 3' ends of viral RNA sequences, to generate full length viral RNA transcripts that are recognized as initial/"authentic" RNA templates for TSWV replication and transcription by viral N and RdRp.

Besides the right promoter, the RdRp protein may present another bottleneck for the establishment of a reverse genetics system. Tospoviruses code for a single, unprocessed ~330 kDa RdRp from the 8.9 kb-sized L RNA (17, 18). The RdRp gene

463 sequence of TSWV was predicted to contain numerous intron splicing sites (*SI*
464 *Appendix*, Table S1). Since the first animal segmented negative strand RNA virus was
465 rescued in 1996 (39), numerous groups worldwide have attempted to construct a
466 reverse genetics system for a tospovirus in plants. Here, it is shown that codon
467 optimization and removal of potential intron splicing sites have been crucial for the
468 expression of a functional RdRp of tospovirus from 35S-driven constructs *in planta*
469 (Fig. 1*B*). While codon optimization may have contributed to increased protein
470 expression levels, removal of predicted potential intron splicing sites from the RdRp
471 genemay have helped to further stabilize and increase expression levels. After all,
472 TSWV is known to replicate in the cytoplasm (2, 48), and its RdRp gene may not
473 have been evolved to escape from the nuclear (pre-mRNA) splicing machinery.
474 However, after nuclear transcription of the RdRp gene by the 35S promoter, any
475 intron splicing site in the wild type RdRp transcript could thus be spliced and result in
476 a truncated, non-functional RdRp.

477 Not only for RdRp, but also an optimized GP gene sequence turned out to be
478 crucial to rescue a full length M RNA-based transcriptionally active RNP. Whereas
479 the $M_{(-)mCherry}$ mini-replicon was able to establish a systemic infection in *N.*
480 *benthamiana* when co-expressed with $S_{(+)eGFP}$ mini-replicon and $L_{(+)opt}$, the wild type
481 full length M segment did not. Like in the case with the RdRp gene sequence, the GP
482 gene sequence of TSWV was also predicted to contain numerous intron splicing sites
483 (*SI Appendix*, Table S2). The absence of antigenomic and genomic RNA strands from
484 the wild type full length M replicon on Northern blots (Fig. S7*B*) indicated the

485 possibility that primary transcripts could have been prone to splicing in the GP
486 sequence. This would not only lead to a loss of genome length RNA molecules, but
487 also inhibit the production of NSm protein (either from direct translation of the
488 primary M transcript, or after secondary transcription of NSm mRNA), needed for
489 cell-to-cell and systemic movement of viral RNPs.

490 Not only the wild type sequence of L and M RNA segments may be spliced in the
491 nucleus, also the S RNA segment generated by the 35S promoter could be prone to
492 splicing. This is supported by the experiments in which *N. benthamiana* where
493 infiltrated with the $S_{(+)eGFP}$ mini-replicon, $M_{(-)opt}$ and $L_{(+)opt}$ and resulted in 80% virus
494 recovery (Table 1), but when only the $S_{(+)eGFP}$ mini-replicon was exchanged for the
495 full length $S_{(+)}$ virus recovery dropped to 11.37 %. The very same reason may explain
496 the low infection rate (10 %) when full length $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ are co-expressed
497 (Table 1). Although this could be due to the splicing of S RNA, the (residual) levels of
498 full length S produced apparently have been sufficient to initiate viral replication.

499 Similar to other bunyaviruses (39, 52, 57), both the RdRp and N proteins are
500 required for reconstitution of infectious RNPs complexes for TSWV (Fig 1*C*).
501 However, high expression of either N or RdRp results in cell death and cause negative
502 effects on the replication and transcription of TSWV. Moreover, ectopic expression of
503 N and RdRp by the strong 35S promoter also seems to cause superinfection exclusion,
504 as earlier observed and reported with various viruses infecting humans, animals, and
505 plants (47, 58-60). During superinfection exclusion a preexisting infection of virus
506 prevents a secondary infection with the same or a highly similar virus. It is an active

507 virus-controlled process that may be determined by a specific viral protein. For
508 example, for potyvirus the coat protein and NIa protease have been identified to
509 control superinfection exclusion (60). For TCV, the p28, involved in replication
510 protein, was shown to confer superinfection exclusion as *a priori* expression of p28
511 blocked (re-)infection with TCV (47). Ectopic expression of N and RdRp may also
512 have blocked progeny L-, M- and S- RNAs from moving into neighboring plant cells.
513 However, the presence of L-, M- and S- RNA segments in the same cells is a
514 pre-requisite for the reconstitution of infectious TSWV and to systemic spread in *N.*
515 *benthamiana.* Fortunately, direct expression of N from $S_{(+)}$ and RdRp from $L_{(+)opt}$ have
516 helped to overcome superinfection exclusion and to recover infectious TSWV.
517 Whether this is due to the fact that N and RdRp are directly expressed from primary
518 viral genome transcripts and simultaneously associate to progeny L-, M- and S- RNA
519 segments in the same plant cells into infectious RNPs and/or involves a more
520 fine-tuned protein expression relative to RNA segment replication remains unclear.
521 Ectopic expression of NSs also inhibited the rescue of full length $S_{(+)}$ segment from
522 cDNA. Since TSWV NSs significantly enhanced the replication of S and M
523 mini-replicons lacking the NSs ORF, this indicated that the inhibition could relate to
524 (simultaneous / *a priori*) ectopic expression of NSs gene sequences with overlap to
525 the full length $S_{(+)}$ replicon. This could be tested by ectopic expression of an
526 untranslatable $NSs^{\Delta ATG}$ construct.

527 In summary, a series of issues has hampered the construction of a successful
528 TSWV reverse genetics system: the choice of promoter and construct design to

generate primary viral RNA transcripts in plants that mimick authentic viral RNA molecules, the expression of a very large viral RdRp, negative effects of ectopic expression of RdRp, N and NSs, and the absence of viral RNA synthesis of the wild type $M_{(-)}$ segment. In this study, we have been able to solve all these issues and successfully managed to establish a reverse genetics system for the tripartite RNA genome of TSWV. Using the S RNA mini-replicon system containing eGFP and mCherry reporter genes, the role of *cis*-and *trans*-acting elements for viral replication and transcription can be studied. Using the M-RNA mini-replicon system cell-to-cell movement of TSWV RNPs *in planta* can be studied. To track the virus during systemic infection of plants rTSWV can be generated containing fluorescent reporter genes at the genetic loci of either GP or NSs. The establishment of this RG system now provides us with a new and powerful platform to generate mutant viruses and study their disease pathology in a natural setting, including basic principles of all tospovirus life cycle and viral pathogenicity. As a personal communication, Jeanmarie Verchot's group has also recovered the Rose rosette emaravirus entirely from cDNA clones, a plant NSV with 7 RNA segments. The establishment of these RG systems presents the start of a new research era for the segmented plant NSVs.

Materials and Methods

Details of the methodology used are provided in *SI Appendix,* Materials and Methods, and include plasmid construction, plant material and growth conditions, agro-infiltration, immunoblot analysis, Northern blot analysis, RT-PCR, GFP imaging,

fluorescence microscopy and Electron microscopy. Primers used in this study are listed in *SI Appendix*, Table S3.

ACKNOWLEDGMENTS

We thank Dr. Yi Xu for critical review of this manuscript. This work was supported by the National Natural Science Foundation of China (31630062 and 31471746), the Fundamental Research Funds for the Central Universities (JCQY201904), Youth Science and Technology Innovation Program to XT and Postgraduate Research & Practice Innovation Program of Jiangsu Province to MF.

References

1. Fields BN, Knipe, DM, & Howley, PM (1996) *Fields Virology* (Lippincott-Raven, New York).
2. Elliott RM, Blakqori G (2011) Molecular biology of orthobunyaviruses. In: Plyusnin, A, Elliott, RM (Eds.), The Bunyaviridae: molecular and cellular biology. *Horizon Scientific Press,* Norwhich, UK.
3. German TL, Ullman DE, & Moyer JW (1992) Tospoviruses: diagnosis, molecular biology, phylogeny, and vector relationships. *Annu Rev Phytopathol* 30:315-348.
4. Kong L, Wu J, Lu L, Xu Y, & Zhou X (2014) Interaction between rice stripe virus disease-specific protein and host PsbP enhances virus symptoms. *Mol Plant* 7(4):691-708.
5. Lu G, *et al.* (2019) Tenuivirus utilizes its glycoprotein as a helper component to overcome insect midgut barriers for its circulative and propagative transmission. *PLoS Pathog* 15(3):e1007655.
6. Oliver JE & Whitfield AE (2016) The genus tospovirus: emerging bunyaviruses that threaten food security. *Annu Rev Virol* 3(1):101-124.
7. Prins M & Goldbach R (1998) The emerging problem of tospovirus infection and nonconventional methods of control. *Trends Microbiol* 6(1):31-35.
8. Turina M, Kormelink R, & Resende RO (2016) Resistance to tospoviruses in vegetable crops: epidemiological and molecular aspects. *Annu Rev Phytopathol* 54:347-371.
9. Whitfield AE, Ullman DE, & German TL (2005) Tospovirus-thrips interactions. *Annu Rev Phytopathol* 43:459-489.
10. Zhu M, van Grinsven IL, Kormelink R, & Tao X (2019) Paving the way to tospovirus infection: multilined interplays with plant innate immunity. *Annu Rev Phytopathol* 57:2.1-2.22.

11. Kormelink R, Garcia ML, Goodin M, Sasaya T, & Haenni AL (2011) Negative-strand RNA viruses: the plant-infecting counterparts. *Virus Res* 162(1-2):184-202.
12. Scholthof KB, *et al.* (2011) Top 10 plant viruses in molecular plant pathology. *Mol Plant Pathol* 12(9):938-954.
13. Adams MJ, *et al.* (2017) Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses. *Arch Virol* 162(8):2505-2538.
14. Pappu HR, Jones RA, & Jain RK (2009) Global status of tospovirus epidemics in diverse cropping systems: successes achieved and challenges ahead. *Virus Res* 141(2):219-236.
15. Hogenhout SA, Ammar el D, Whitfield AE, & Redinbaugh MG (2008) Insect vector interactions with persistently transmitted viruses. *Annu Rev Phytopathol* 46:327-359.
16. Gilbertson RL, Batuman O, Webster CG, & Adkins S (2015) Role of the insect supervectors Bemisia tabaci and Frankliniella occidentalis in the emergence and global spread of plant viruses. *Annu Rev Virol* 2(1):67-93.
17. Adkins S, Quadt R, Choi TJ, Ahlquist P, & German T (1995) An RNA-dependent RNA polymerase activity associated with virions of tomato spotted wilt virus, a plant- and insect-infecting bunyavirus. *Virology* 207(1):308-311.
18. de Haan P, *et al.* (1991) Tomato spotted wilt virus L RNA encodes a putative RNA polymerase. *J Gen Virol* 72 (Pt 9):2207-2216.
19. Kikkert M, *et al.* (1999) Tomato spotted wilt virus particle morphogenesis in plant cells. *J Virol* 73(3):2288-2297.
20. Ribeiro D, *et al.* (2008) Tomato spotted wilt virus glycoproteins induce the formation of endoplasmic reticulum- and Golgi-derived pleomorphic membrane structures in plant cells. *J Gen Virol* 89(Pt 8):1811-1818.
21. Sin SH, McNulty BC, Kennedy GG, & Moyer JW (2005) Viral genetic determinants for thrips transmission of tomato spotted wilt virus. *Proc Natl Acad Sci U S A* 102(14):5168-5173.
22. Feng Z, *et al.* (2016) The ER-membrane transport system is critical for intercellular trafficking of the NSm movement protein and tomato spotted wilt tospovirus. *PLoS Pathog* 12(2):e1005443.
23. Kormelink R, Storms M, Van Lent J, Peters D, & Goldbach R (1994) Expression and subcellular location of the NSm protein of tomato spotted wilt virus (TSWV), a putative viral movement protein. *Virology* 200(1):56-65.
24. Soellick T, Uhrig JF, Bucher GL, Kellmann JW, & Schreier PH (2000) The movement protein NSm of tomato spotted wilt tospovirus (TSWV): RNA binding, interaction with the TSWV N protein, and identification of interacting plant proteins. *Proc Natl Acad Sci U S A* 97(5):2373-2378.
25. Storms MM, Kormelink R, Peters D, Van Lent JW, & Goldbach RW (1995) The nonstructural NSm protein of tomato spotted wilt virus induces tubular structures in plant and insect cells. *Virology* 214(2):485-493.
26. Storms MM, *et al.* (1998) A comparison of two methods of microinjection for assessing altered plasmodesmal gating in tissues expressing viral movement proteins. *Plant J* 13(1):131-140.
27. Li J, *et al.* (2015) Structure and function analysis of nucleocapsid protein of tomato spotted wilt virus interacting with RNA using homology modeling. *J Biol Chem* 290(7):3950-3961.

628 28. Komoda K, Narita M, Yamashita K, Tanaka I, & Yao M (2017) Asymmetric trimeric ring
629 structure of the nucleocapsid protein of tospovirus. *J Virol* 91(20):e01002-17.
630 29. Guo Y, *et al.* (2017) Distinct mechanism for the formation of the ribonucleoprotein complex
631 of tomato spotted wilt virus. *J Virol* 91(23):e00892-17.
632 30. Feng Z, *et al.* (2013) Nucleocapsid of tomato spotted wilt tospovirus forms mobile particles
633 that traffic on an actin/endoplasmic reticulum network driven by myosin XI-K. *New Phytol*
634 200(4):1212-1224.
635 31. Ribeiro D, *et al.* (2013) The cytosolic nucleoprotein of the plant-infecting bunyavirus tomato
636 spotted wilt recruits endoplasmic reticulum-resident proteins to endoplasmic reticulum export
637 sites. *Plant Cell* 25(9):3602-3614.
638 32. Bucher E, Sijen T, De Haan P, Goldbach R, & Prins M (2003) Negative-strand tospoviruses
639 and tenuiviruses carry a gene for a suppressor of gene silencing at analogous genomic
640 positions. *J Virol* 77(2):1329-1336.
641 33. Schnettler E, *et al.* (2010) Diverging affinity of tospovirus RNA silencing suppressor proteins,
642 NSs, for various RNA duplex molecules. *J Virol* 84(21):11542-11554.
643 34. Takeda A, *et al.* (2002) Identification of a novel RNA silencing suppressor, NSs protein of
644 tomato spotted wilt virus. *FEBS Lett* 532(1-2):75-79.
645 35. Hoang NH, Yang HB, & Kang BC (2013) Identification and inheritance of a new source of
646 resistance against tomato spotted wilt virus (TSWV) in Capsicum. *Sci Hortic-Amsterdam*
647 161:8-14.
648 36. de Ronde D, *et al.* (2014) Analysis of tomato spotted wilt virus NSs protein indicates the
649 importance of the N-terminal domain for avirulence and RNA silencing suppression. *Mol*
650 *Plant Pathol* 15(2):185-195.
651 37. Kim SB, *et al.* (2017) Divergent evolution of multiple virus-resistance genes from a
652 progenitor in Capsicum spp. *New Phytol* 213(2):886-899.
653 38. Brittlebank CC (1919) Tomato diseases. *J Agri Victoria* 27:231-235.
654 39. Bridgen A & Elliott RM (1996) Rescue of a segmented negative-strand RNA virus entirely
655 from cloned complementary DNAs. *Proc Natl Acad Sci U S A* 93(26):15400-15404.
656 40. Neumann G, *et al.* (1999) Generation of influenza A viruses entirely from cloned cDNAs.
657 *Proc Natl Acad Sci U S A* 96(16):9345-9350.
658 41. Flatz L, Bergthaler A, de la Torre JC, & Pinschewer DD (2006) Recovery of an arenavirus
659 entirely from RNA polymerase I/II-driven cDNA. *Proc Natl Acad Sci U S A*
660 103(12):4663-4668.
661 42. Ganesan U, *et al.* (2013) Construction of a sonchus yellow net virus mini-replicon: a step
662 toward reverse genetic analysis of plant negative-strand RNA viruses. *J Virol*
663 87(19):10598-10611.
664 43. Wang Q, *et al.* (2015) Rescue of a plant negative-strand RNA virus from cloned cDNA:
665 insights into enveloped plant virus movement and morphogenesis. *PLoS Pathog*
666 11(10):e1005223.
667 44. Ishibashi K, Matsumoto-Yokoyama E, & Ishikawa M (2017) A tomato spotted wilt virus S
668 RNA-based replicon system in yeast. *Sci Rep* 7(1):12647.
669 45. van Knippenberg I, Goldbach R, & Kormelink R (2005) Tomato spotted wilt virus S-segment
670 mRNAs have overlapping 3'-ends containing a predicted stem-loop structure and conserved
671 sequence motif. *Virus Res* 110(1-2):125-131.

672 46. de Haan P, Wagemakers L, Peters D, & Goldbach R (1990) The S RNA segment of tomato
673 spotted wilt virus has an ambisense character. *J Gen Virol* 71 (Pt 5):1001-1007.
674 47. Zhang XF, *et al.* (2017) A self-perpetuating repressive state of a viral replication protein
675 blocks superinfection by the same virus. *PLoS Pathog* 13(3):e1006253.
676 48. Elliott RM (2014) Orthobunyaviruses: recent genetic and structural insights. *Nat Rev*
677 *Microbiol* 12(10):673-685.
678 49. Ferron F, Weber F, de la Torre JC, & Reguera J (2017) Transcription and replication
679 mechanisms of bunyaviridae and arenaviridae L proteins. *Virus Res* 234:118-134.
680 50. Blakqori G & Weber F (2005) Efficient cDNA-based rescue of La Crosse bunyaviruses
681 expressing or lacking the nonstructural protein NSs. *J Virol* 79(16):10420-10428.
682 51. Ikegami T, Won S, Peters CJ, & Makino S (2006) Rescue of infectious rift valley fever virus
683 entirely from cDNA, analysis of virus lacking the NSs gene, and expression of a foreign gene.
684 *J Virol* 80(6):2933-2940.
685 52. Flick R & Pettersson RF (2001) Reverse genetics system for uukuniemi virus (Bunyaviridae):
686 RNA polymerase I-catalyzed expression of chimeric viral RNAs. *J Virol* 75(4):1643-1655.
687 53. Hempel WM, Cavanaugh AH, Hannan RD, Taylor L, & Rothblum LI (1996) The
688 spezhucies-specific RNA polymerase I transcription factor SL-1 binds to upstream binding
689 factor. *Mol Cell Biol* 16(2):557-563.
690 54. Doelling JH, Gaudino RJ, & Pikaard CS (1993) Functional analysis of Arabidopsis thaliana
691 rRNA gene and spacer promoters in vivo and by transient expression. *Proc Natl Acad Sci U S*
692 *A* 90(16):7528-7532.
693 55. SaezVasquez J & Pikaard CS (1997) Extensive purification of a putative RNA polymerase I
694 holoenzyme from plants that accurately initiates rRNA gene transcription in vitro. *P Natl Acad*
695 *Sci USA* 94(22):11869-11874.
696 56. Martin A, Staeheli P, & Schneider U (2006) RNA polymerase II-controlled expression of
697 antigenomic RNA enhances the rescue efficacies of two different members of the
698 Mononegavirales independently of the site of viral genome replication. *J Virol*
699 80(12):5708-5715.
700 57. Flick R, Flick K, Feldmann H, & Elgh F (2003) Reverse genetics for crimean-congo
701 hemorrhagic fever virus. *J Virol* 77(10):5997-6006.
702 58. Laliberte JP & Moss B (2014) A novel mode of poxvirus superinfection exclusion that
703 prevents fusion of the lipid bilayers of viral and cellular membranes. *J Virol*
704 88(17):9751-9768.
705 59. Webster B, Ott M, & Greene WC (2013) Evasion of superinfection exclusion and elimination
706 of primary viral RNA by an adapted strain of hepatitis C virus. *J Virol* 87(24):13354-13369.
707 60. Tatineni S & French R (2016) The coat protein and NIa protease of two potyviridae family
708 members independently confer superinfection exclusion. *J Virol* 90(23):10886-10905.

709

710

711

712

713

714 **Table 1.** Systemic infection rate of recombinant TSWV rescued in *N. benthamiana* in

715 the presence of viral suppressors of RNA silencing (VSRs).

Antigenome and genome derivatives	**VSRs**	**Systemic infection (No. of infected/inoculated plants)**
$S_{(-)}$+$M_{(-)}$ +$L_{(-)}$	N+RdRp+NSs+P19-HcPro-γb	0 % (0/30)
$S_{(+)}$+$M_{(+)}$+$L_{(+)}$	N+RdRp+NSs+P19-HcPro-γb	0 % (0/30)
$SM_{(+)eGFP}$+$MR_{(-)mCherry}$+$L_{(+)opt}$	NSs+P19-HcPro-γb	80 % (24/30)
$S_{(+)}$+$MR_{(-)eGFP}$+$L_{(+)opt}$	P19-HcPro-γb	11.37 % (7/60)
$S_{(+)}$+$MR_{(-)eGFP}$+$L_{(+)opt}$	NSs+P19-HcPro-γb	0 % (0/60)
$SR_{(+)eGFP}$+$M_{(-)}$+$L_{(+)opt}$	P19-HcPro-γb	0 % (0/60)
$SR_{(+)eGFP}$+$M_{(-)}$+$L_{(+)opt}$	NSs+P19-HcPro-γb	0 % (0/60)
$SR_{(+)eGFP}$+$M_{(-)opt}$+$L_{(+)opt}$	NSs+P19-HcPro-γb	90 % (27/30)
$S_{(+)}$+$M_{(-)opt}$+$L_{(+)opt}$	P19-HcPro-γb	10 % (6/60)

716

717 Mixture of *Agrobacterium* cultures harboring the plasmids encoding each of the S, M, L and

718 derivatives (final concentration OD_{600}=0.2), RdRp (OD_{600}=0.2), N (OD_{600}=0.2) and VSRs

719 (OD_{600}=0.05) were infiltrated into *N. benthamiana* leaves. Systemic infection was scored

720 15-30 dpi.

721

722

723

724

725

726

727

728 **Figure Legends:**

729 **Fig. 1** Construction of a TSWV $S_{(-)}$ RNA-based mini-replicon system in *N.*
730 *benthamiana*. (*A*) Schematic representation of binary constructs to express TSWV $S_{(-)}$
731 mini-replicon, TSWV N, RdRp and four RNA silencing suppressors (VSRs: NSs, P19,
732 HcPro and γb) proteins by agroinfiltration into *N. benthamiana*. The $S_{(-)}$-gRNA of
733 TSWV is shown on the top. $SR_{(-)mCherry\&eGFP}$: the NSs and N of $S_{(-)}$-gRNA were
734 replaced by mCherry and eGFP, respectively. (-) refers to the negative
735 (genomic)-strand of S RNA; 2×35S: a double 35S promoter; HH: hammerhead
736 ribozyme; RZ: Hepatitis Delta virus (HDV) ribozyme; NOS: nopaline synthase
737 terminator. (*B*) Foci of eGFP and mCherry fluorescence in *N. benthamiana* leaves
738 co-expressing $SR_{(-)mCherry\&eGFP}$, RdRp, N and four VSRs at 5 days post infiltration (dpi)
739 under a fluorescence microscope. The bar represents 400 μm. (*C*) Analysis of RdRp
740 and N requirement for $SR_{(-)mCherry\&eGFP}$ mini-genome replication in *N. benthamiana*
741 leaves. $SR_{(-)mCherry\&eGFP}$ was coexpressed with pCB301 empty vector (Vec), N, RdRp
742 or both in *N. benthamiana* leaves by agroinfiltration. Agro-infiltrated leaves were
743 examined and photographed at 5 dpi under a fluorescence microscope. Signal shown
744 reflects a merge of mCherry and eGFP fluorescence from both reporter genes. Bar
745 represents 400 μm. (*D*) Immunoblot analysis on the expression of N and eGFP
746 proteins in the leaves shown in panel (*C*) using specific antibodies against N and GFP,
747 respectively. Ponceau S staining of rubisco large subunit is shown for protein loading
748 control. (*E*) Northern blot analysis of $S_{(-)}$-mini-replicon replication and transcription
749 in the presence of N, RdRp or both in *N. benthamiana*. The S RNA genomic,

750 anti-genomic and subgenomic transcripts (eGFP mRNA) were detected by
751 DIG-labeled sense eGFP or anti-sense eGFP probes. The red and blue arrows indicate
752 the anti-genomic and genomic RNAs of $SR_{(-)mCherry\&eGFP}$, respectively. The green
753 arrow indicates the eGFP mRNA transcript. Ethidium bromide staining of ribosomal
754 RNA (rRNA) was used as RNA loading control.
755
756 **Fig. 2** Optimization of the $SR_{(-)mCherry\&eGFP}$ mini-replicon system. (*A*) Optimizing the
757 concentration of N and RdRp proteins for replication and transcription of
758 $SR_{(-)mCherry\&eGFP}$ in *N. benthamiana* leaves. Increasing amounts of *Agrobacterium,*
759 from OD_{600}= 0.2 to 0.8 and containing the binary expression constructs for N (upper
760 panels) or RdRp (bottom panels), were mixed with fixed amounts of *Agrobacterium*
761 containing the RdRp or N construct (OD_{600} 0.2), respectively, and their effects on
762 eGFP reporter expression were visualized under a fluorescence microscope at 5 dpi.
763 Bars represent 400 μm. (*B*) and (*C*) Western immunoblot detection of the N and eGFP
764 proteins expressed in the leaves shown in panel (*A*) using specific antibodies against
765 N and GFP, respectively. (*D*) Optimization of RNA silencing suppressors (VSRs) on
766 $SR_{(-)mCherry\&eGFP}$ mini-reporter replication and transcriptions as measured by eGFP and
767 mCherry expression. The $SR_{(-)mCherry\&eGFP}$, N and RdRp proteins were co-expressed
768 with pCB301 empty vector (Vec), NSs, P19-HcPro-γb or all four VSRs in *N.*
769 *benthamiana* leaves. Foci expressing eGFP and mCherry in agroinfiltrated leaves
770 were visualized under a fluorescence microscope at 5 dpi. Bars represent 400 μm. (*E*)
771 Western immunoblot detection of N and eGFP protein synthesis in the leaves shown

772 in panel (*D*) using N and GFP-specific antibodies, respectively. Ponceau S staining
773 was used as protein loading control.
774
775 **Fig. 3** Development of the $S_{(+)}$-agRNA mini-replicon system in *N. benthamiana*. (*A*)
776 Schematic representation of the TSWV $SR_{(+)eGFP}$ mini-replicon. The NSs gene of
777 TSWV S agRNA was replaced by eGFP. Anti-genomic RNA strands of the $SR_{(+)eGFP}$
778 mini-replicon are transcribed from a double 35S promoter (2×35S), and flanked by a
779 HH ribozyme and HDV ribozyme (RZ) sequence. (+) refers to the positive
780 (antigenomic)-strand of S RNA. (*B*) Foci of eGFP fluorescence in *N. benthamiana*
781 leaves co-expressing TSWV $SR_{(+)eGFP}$ with pCB301 empty vector (Vec), N, RdRp or
782 N+RdRp by agroinfiltration. The agroinfiltrated leaves were photographed at 3 dpi
783 under a fluorescence microscope. Bars represent 400 μm. (*C*) Western immunoblot
784 detection of N and eGFP protein synthesis in the leaves shown in panel (*B*) using N-
785 and GFP-specific antibodies, respectively. Ponceau S staining was used as a protein
786 loading control. (D) Northern blot analysis of the replication and transcription of
787 $SR_{(+)eGFP}$ mini-replicon in *N. benthamiana* co-expressed with empty vector (Vec), N,
788 RdRp or both. The anti-genomic RNAs (red arrow), genomic RNAs (blue arrow) and
789 eGFP mRNA transcripts (green arrow) were detected with DIG-labeled sense and
790 anti-sense eGFP probes, respectively. Ethidium bromide staining was used as RNA
791 loading control.
792
793 **Fig. 4** Establishment of a TSWV $M_{(-)}$-gRNA based mini-replicon system with

794 cell-to-cell movement competency in *N. benthamiana*. (*A*) Schematic representation
795 of the TSWV $MR_{(-)eGFP}$ mini-replicon and its mutant derivative $MR_{(-)eGFP\&NSmMut}$. The
796 GP gene of TSWV $M_{(-)}$ gRNA was substituted by eGFP. The genomic RNA of the
797 mini-replicon is transcribed from a double 35S promoter (2×35S) and flanked by a
798 Hammerhead (HH) ribozyme and HDV ribozyme (RZ). For $MR_{(-)eGFP\&NSmMut}$, a stop
799 codon was introduced immediately after the start codon of the NSm gene in the
800 $MR_{(-)eGFP}$ mini-replicon. (*B*) Foci of eGFP fluorescence expressed from the $MR_{(-)eGFP}$
801 or $MR_{(-)eGFP\&NSmMut}$ mini-replicon in *N. benthamiana* leaves co-expressed with the
802 empty vector (Vec), N, RdRp or N+RdRp by agroinfiltration. Agroinfiltrated leaves
803 were photographed at 4 dpi under a fluorescence microscope. Bars represent 400 μm.
804 (*C*) Western immunoblot detection of N and eGFP protein synthesis in the leaves
805 shown in panel (*B*) with specific antibodies against N and GFP, respectively. Ponceau
806 S staining was used as protein loading control. (*D*) Northern blot analysis of the
807 replication and transcription of the $MR_{(-)eGFP}$ mini-replicon co-expressed with Vec, N,
808 RdRp or N+RdRp in *N. benthamiana* leaves. The anti-genomic RNAs (red arrow),
809 genomic RNAs (blue arrow) and eGFP mRNA transcripts (green arrow) were detected
810 with DIG labeled sense and anti-sense eGFP probes, respectively. Ethidium bromide
811 staining was used as RNA loading control.
812
813 **Fig. 5** Establishment of a systemic infection in *N. benthamiana* with replicons $S_{(+)}$ and
814 $M_{(-)}$ co-expressed with full length antigenomic $L_{(+)opt}$. (*A*) Schematic representation of
815 constructs expressing TSWV full length antigenomic $L_{(+)opt}$ with optimized RdRp, full

length genomic $M_{(-)opt}$ with optimized GP, $MR_{(-)eGFP}$, $MR_{(-)mCherry}$, full length antigenomic $S_{(+)}$ and $SR_{(+)eGFP}$. Primary viral RNA transcripts are transcribed from a double 35S promoter (2×35S) and flanked by a HH and HDV ribozyme (RZ). (*B*) eGFP and mCherry fluorescence in *N. benthamiana* resulting from systemic infection of agroinfiltrated $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$ constructs. The systemic infected plants (*a* and *b*) and leaves (*c* and *d*) were photographed at 21 dpi under white light and (hand-held) ultraviolet (UV) light. Foci of eGFP and mCherry fluorescence in leaves shown in panel *d* as visualized under a fluorescence microscope. Bar represents 400 μm. (*C*) eGFP fluorescence in *N. benthamiana* resulting from systemic infection of agroinfiltrated $SR_{(+)eGFP}$, $M_{(-)opt}$ and $L_{(+)opt}$ constructs. Infected plants (*a* and *b*) and leaves (*c* and *d*) were photographed at 18 dpi under white light and (hand-held) UV light, respectively. (*D*) eGFP fluorescence in *N. benthamiana* resulting from systemic infection of agroinfiltrated $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$ constructs. Infected plants (*a* and *b*) and leaves (*c* and *d*) were photographed at 50 dpi under white light and (hand-held) UV light, respectively.

Fig. 6 Rescue of infectious TSWV from full-length cDNA clones in *N. benthamiana*. (*A*) Systemic infection of *N. benthamiana* plants with rescued TSWV (rTSWV) resulting from agoinfiltration of $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ and three VSRs (P19, HcPro and γb). The plant agroinfiltrated with pCB301 empty vector was used as a mock control. Images were taken at 19 dpi. Boxed areas of the plants that show stunting, mosaic and leaf curling are shown enlarged in the right panels. (*B*) Sequence

838 confirmation of codon optimized sequences of the GP gene (from the $M_{(-)opt}$ RNA
839 segment) and the RdRp gene (from the $L_{(+)opt}$ RNA segment) on RT-PCR fragments
840 obtained from systemic leaves of *N. benthamiana* infected with rTSWV. The
841 optimized sequence of GP from rTSWV is underlined by a blue dashed line, and wild
842 type GP sequence underlined in blue. The 3'-untranslated region (UTR) sequence of
843 the M genomic RNA is marked with a yellow line. The optimized sequence of RdRp
844 from rTSWV is underlined by a red dashed line, and wild type RdRp sequence
845 underlined in red. The 3'-UTR sequence of the L genomic RNA is marked with a
846 purple line. The stars indicate the codon optimization sites of GP and RdRp gene
847 sequences. (*C*) Northern blot detection of viral RNA from the S, M and L RNA
848 segment, respectively, in systemic leaves of *N. benthamiana* infected with rTSWV.
849 The anti-genomic RNAs (red arrow) and genomic RNAs (blue arrow) were detected
850 with DIG labeled sense and anti-sense NSs-, NSm- and L-5'UTR probes, respectively.
851 Lane 1 and 2 refer to two independent replicates. Ethidium bromide staining was used
852 as RNA loading control. (*D*) Western immunoblot detection of the N, NSm, NSs, Gc,
853 and Gn proteins from leaves systemically infected with rTSWV, using specific
854 antibodies against N, NSm, NSs, Gc, and Gn, respectively. Leaves infected with
855 wild-type TSWV were used as a positive control. Ponceau S staining was used as
856 protein loading control. (*E*) Electron micrographs of thin sections of *N. benthamiana*
857 plants infected with rTSWV. Boxed regions in the left panels and showing the
858 presence of virions, are shown enlarged in the right panels. Spherical enveloped virus
859 particles are indicated (white arrow head). Bars represent 0.2 μm.

Supplemental Figure Legends

Fig. S1 Functional analysis of wild type RdRp and the use of T7 promoter in a mini-genome replication assay. (*A*) Schematic diagram of TSWV 35S:SR$_{(-)mCherry\&eGFP}$ and T7:SR$_{(-)mCherry\&eGFP}$ mini-replicon reporters. (*B*) The wild type RdRp (RdRp$_{wt}$) or the empty vector (Vec) was co-expressed with 35S:SR$_{(-)mCherry\&eGFP}$, N, VSRs (VSRs: NSs, P19, HcPro and γb) in *N. benthamiana* leaves. The expression of eGFP was examined by a fluorescence microscope. (*C*) Constructs coding for T7:SR$_{(-)mCherry\&eGFP}$, T7 RNA polymerase (pol), N and VSRs were co-expressed with RdRp$_{wt}$ or RdRp$_{opt}$ in *N. benthamiana* leaves. Replication of T7:SR$_{(-)mCherry\&eGFP}$ was examined by monitoring for eGFP fluorescence with a fluorescence microscope. Empty vector (Vec) pCB301 was used as a negative control. Bars represent 200 μm. (*D*) Western immunoblot detection of T7 RNA pol using a T7 RNA pol specific antibody. Ponceau S staining was used as protein loading control. Lane 1: sample from leaves co-expressing T7:SR$_{(-)mCherry\&eGFP}$, T7 RNA Pol, N, VSRs and RdRp$_{wt}$; lane 2: sample from leaves co-expressing T7:SR$_{(-)mCherry\&eGFP}$, T7 RNA Pol, N, VSRs and RdRp$_{opt}$.

Fig. S2 Time course analysis on gene expression from the SR$_{(-)mCherry\&eGFP}$ mini-replicon in *N. benthamiana* leaves. (*A*) Foci of eGFP and mCherry fluorescence expressed from SR$_{(-)mCherry\&eGFP}$ in *N. benthamiana* leaves co-expressing N, RdRp and the VSRs at 3, 6, 9 and 12 dpi, respectively. Fluorescence of eGFP and mCherry were photographed under a fluorescence microscope using GFP and RFP filters,

882 respectively. Bars represent 400 μm. (*B*) Western immunoblot detection of the N,
883 eGFP and mCherry proteins in leaves shown in panel A, using specific antibodies
884 against N, GFP and mCherry, respectively. The empty vector (Vec) was used as a
885 negative control. Ponceau S staining was used as protein loading control.
886
887 **Fig. S3** The role of 5'-UTR, 3'-UTR and IGR on viral RNA synthesis from the
888 $SR_{(-)mCherry\&eGFP}$ mini-replicon. (*A*) Schematic representation of TSWV
889 $SR_{(-)mCherry\&eGFP}$ and derivatives with deletions of the 5'UTR, IGR or 3'UTR. (*B*) eGFP
890 and mCherry fluorescence expressed from TSWV $SR_{(-)mCherry\&eGFP}$ and mutant
891 derivatives in *N. benthamiana*. The $SR_{(-)mCherry\&eGFP}$ or its mutants were coexpressed
892 with N, RdRp and the four VSRs in *N. benthamiana* leaves. The agroinfiltrated leaves
893 were examined and photographed at 5 dpi under a fluorescence microscope using
894 GFP and RFP filters, respectively. Bars represent 200 μm. (*C*) Western immunoblot
895 detection of the N and eGFP proteins expressed from the $SR_{(-)mCherry\&eGFP}$
896 mini-replicon and mutant derivatives, using specific antibodies against N and GFP,
897 respectively. Ponceau S staining was used as protein loading control. (*D*) Northern
898 blot analysis of viral RNA synthesis from $SR_{(-)mCherry\&eGFP}$ and mutant derivatives. The
899 anti-genomic RNAs (red arrow), genomic RNAs (blue arrow) and eGFP mRNA
900 transcripts (green arrow) were detected with DIG–labeled sense eGFP or anti-sense
901 eGFP probes. Ethidium bromide staining was used as RNA loading control.
902
903 **Fig. S4** Cell-to-cell movement analysis of $SR_{(-)GFP}$ and $MR_{(-)mCherry}$ in *N. benthamiana*

co-expressing RdRp, N and four VSRs in *N. benthamiana*. Agroinfiltrated leaves were examined and photographed at 5 dpi under a fluorescence microscope. Bars represent 400 μm.

Fig. S5 Functional analysis of RdRp expressed from TSWV $L_{(+)opt}$, NSm from $MR_{(-)eGFP}$ and N from $SR_{(+)eGFP}$ using the mini-genome replication system in *N. benthamiana*. (*A*) Functional analysis of RdRp expressed from TSWV $L_{(+)opt}$ using the S RNA mini-replicon system in *N. benthamiana*. The $L_{(+)opt}$, RdRp, or pCB301 empty vector (Vec) was co-expressed with N, $SR_{(-)mCherry\&eGFP}$ and the four VSRs in *N. benthamiana* leaves. (*B*) Functional analysis of N expressed from $SR_{(+)eGFP}$ in *N. benthamiana*. $SR_{(+)eGFP}$ was co-expressed with the empty vector (Vec), N, RdRp or $N+L_{(+)opt}$ in *N. benthamiana* leaves in the presence of four VSRs. (*C*) Functional analysis of NSm expressed from $MR_{(-)eGFP}$ in *N. benthamiana*. $MR_{(-)eGFP}$ was co-expressed with the empty vector (Vec), N, RdRp or $N+L_{(+)opt}$ in *N. benthamiana* leaves in the presence of four VSRs. Foci showing mCherry and eGFP fluorescence in agroinfiltrated *N. benthamiana* leaves were examined at 3 dpi by a fluorescence microscope. Bars represent 400 μm.

Fig. S6 Analysis of *N. benthamiana* leaves agroinfiltrated with constructs of $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$, or $SR_{(+)eGFP}$, $MR_{(-)opt}$ and $L_{(+)opt}$ or $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$. (*A*) Local infection analysis of cell-to-cell movement of $SR_{(+)eGFP}$ and $MR_{(-)mCherry}$ co-expressing with $L_{(+)opt}$ and four VSRs in *N. benthamiana* by agroinfiltration. The

926 agro-infiltrated leaves were examined and photographed at 5 dpi under a fluorescence
927 microscope. Bars represent 400 μm. (*B*) RT-PCR analysis on systemically infected
928 leaves from *N. benthamiana* plants agroinfiltrated with $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and
929 $L_{(+)opt}$. (*C*) RT-PCR analysis on systemically infected leaves from *N. benthamiana*
930 plants agroinfiltrated with $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$ (*D*) RT-PCR analysis on
931 systemically infected leaves from *N. benthamiana* plants agroinfiltrated with $SR_{(+)eGFP}$,
932 $M_{(-)opt}$ and $L_{(+)opt}$. All agroinfiltrations were performed in the additional presence four
933 VSRs constructs. RT PCR was performed on total RNA purified from systemic leaves
934 for detection of S, M or L segments using segment-specific primers. Amplicons were
935 resolved by electrophoresis in a 1 % agarose gel. Lanes 1-2 represent two biological
936 replicates of systemic infected leaf samples; As positive controls (CK^+) for proper
937 fragment size, PCR was performed on plasmids carrying S, M , $L_{(+)opt}$ or derivatives.
938 As negative control (CK^-), RT-PCR was performed in the absence of nucleic acids.
939 DNA size markers are shown on the left hand side.
940
941 **Fig. S7** Northern blot detection of viral RNA synthesis produced from full length $S_{(+)}$
942 and wild type $M_{(-)}$ replicons. (*A* and *B*) The full length $S_{(+)}$ or wild type $M_{(-)}$ was
943 co-expressed with the empty vector (Vec), N, $RdRp_{opt}$ or $N+RdRp_{opt}$ in *N.*
944 *benthamiana* leaves in the presence of four VSRs. The genomic RNAs (blue arrow),
945 anti-genomic RNAs (red arrow) of S (*A*) and M (*B*) were detected with DIG–labeled
946 sense or anti-sense NSs and NSm probes, respectively. Ethidium bromide staining
947 was used as RNA loading control.

948

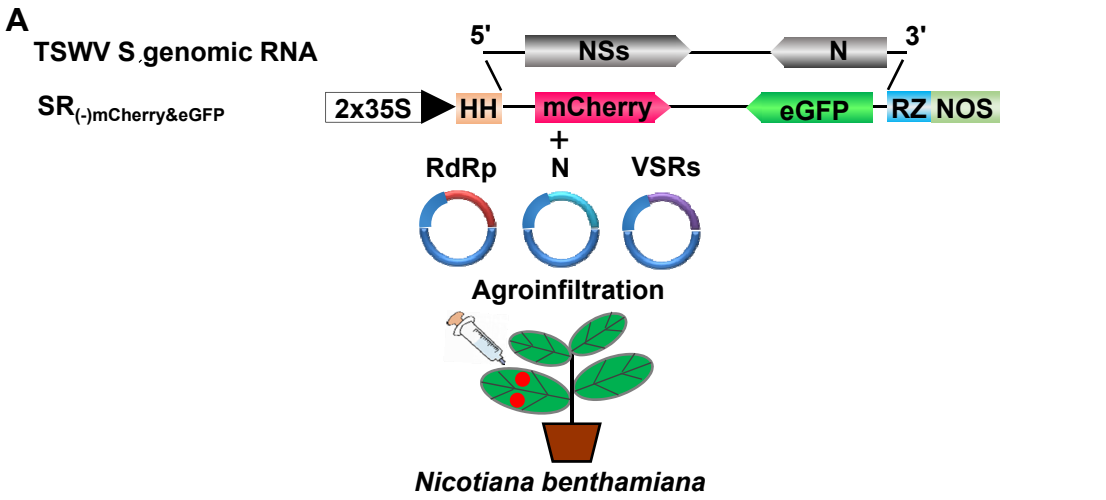
949 **Fig. S8** RT-PCR detection of $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ genomic RNA in systemic leaves
950 of *N. benthamiana* infected by rTSWV. The $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ and the four VSRs
951 were co-expressed in *N. benthamiana* leaves by agroinfiltration. Total RNA was
952 purified from systemic leaves of agroinfiltrated plants and the presence of $S_{(+)}$, $M_{(-)opt}$
953 and $L_{(+)opt}$ were detected by RT-PCR using segment-specific primers. RT-PCR
954 products were resolved by electrophoresis in a 1% agarose gel. Lanes 1-2, two
955 biological replicates of systemic infected leaf samples; RT-PCR on plasmids carrying
956 S, M and L as DNA template were used as positive controls (CK^+). RT-PCR without
957 adding the DNA template was used as negative controls (CK^-). DNA size markers are
958 shown on the left hand side.

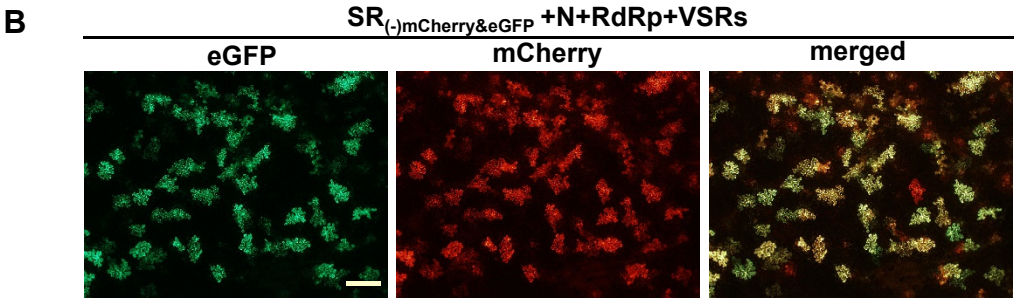
959

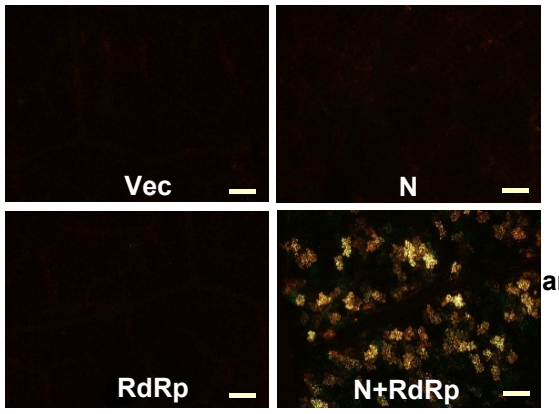
960 **Fig. S9** Optimized RdRp gene sequence used in the study.

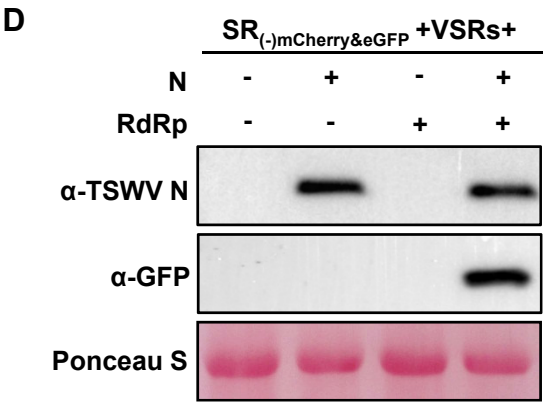
961

962 **Fig. S10** Optimized GP gene sequence used in the study.









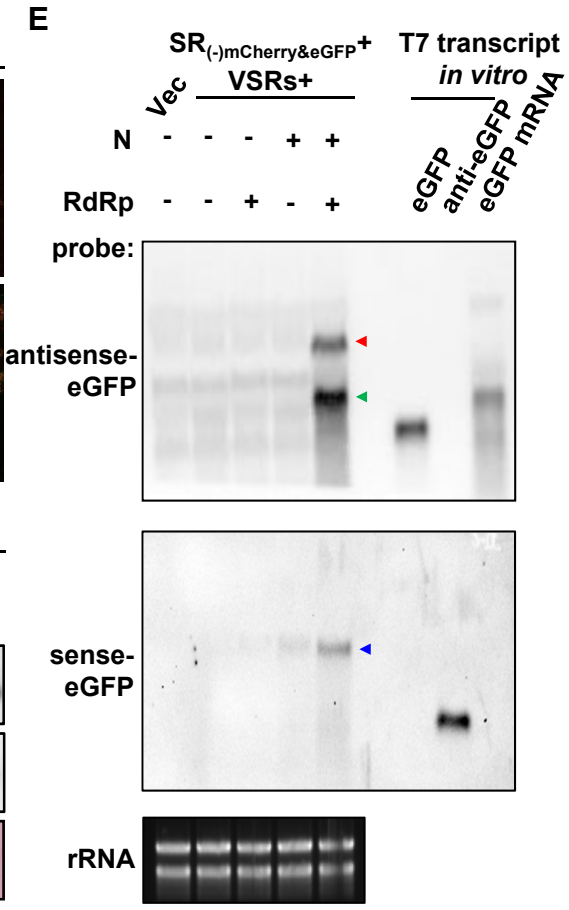


Fig. 1

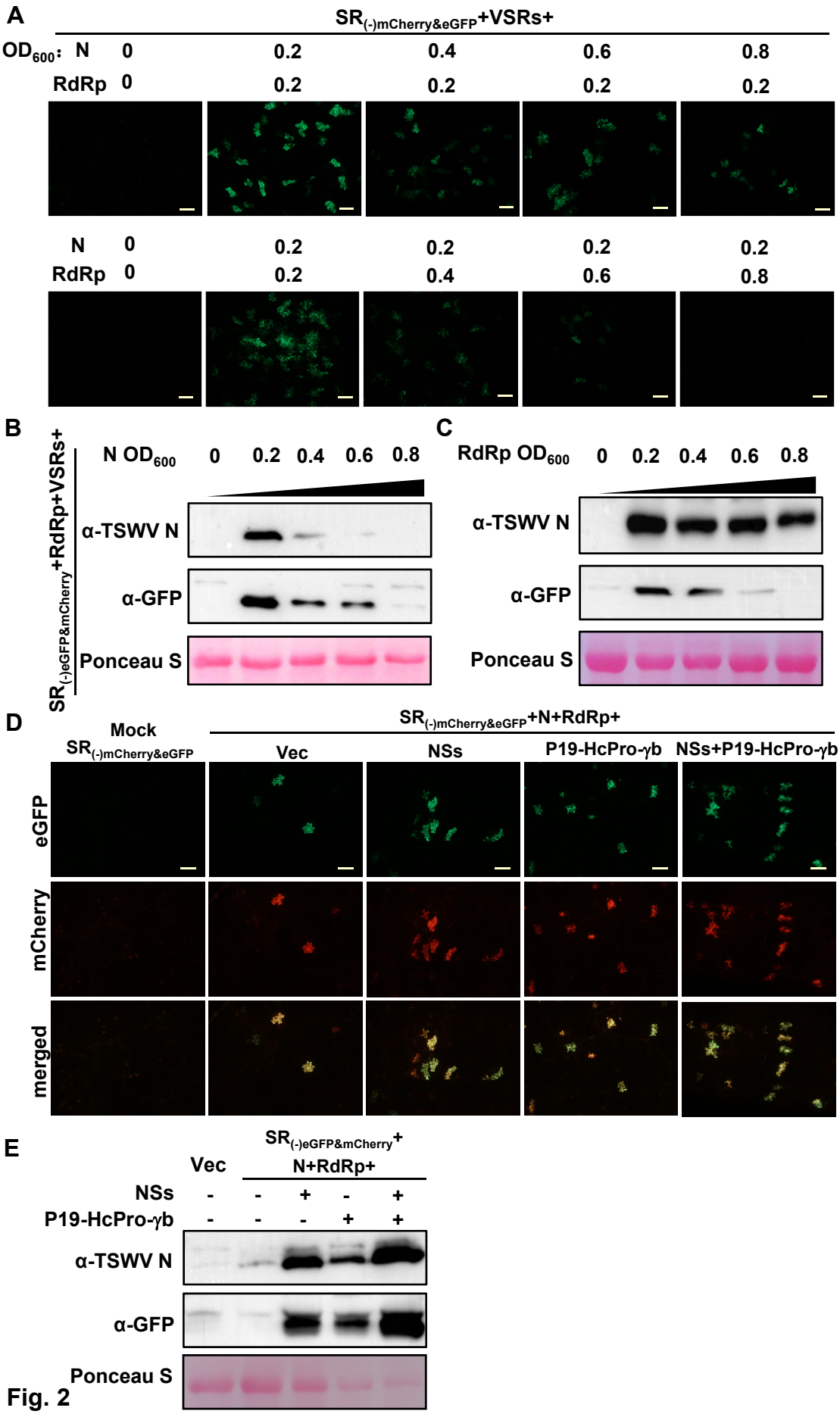


Fig. 2

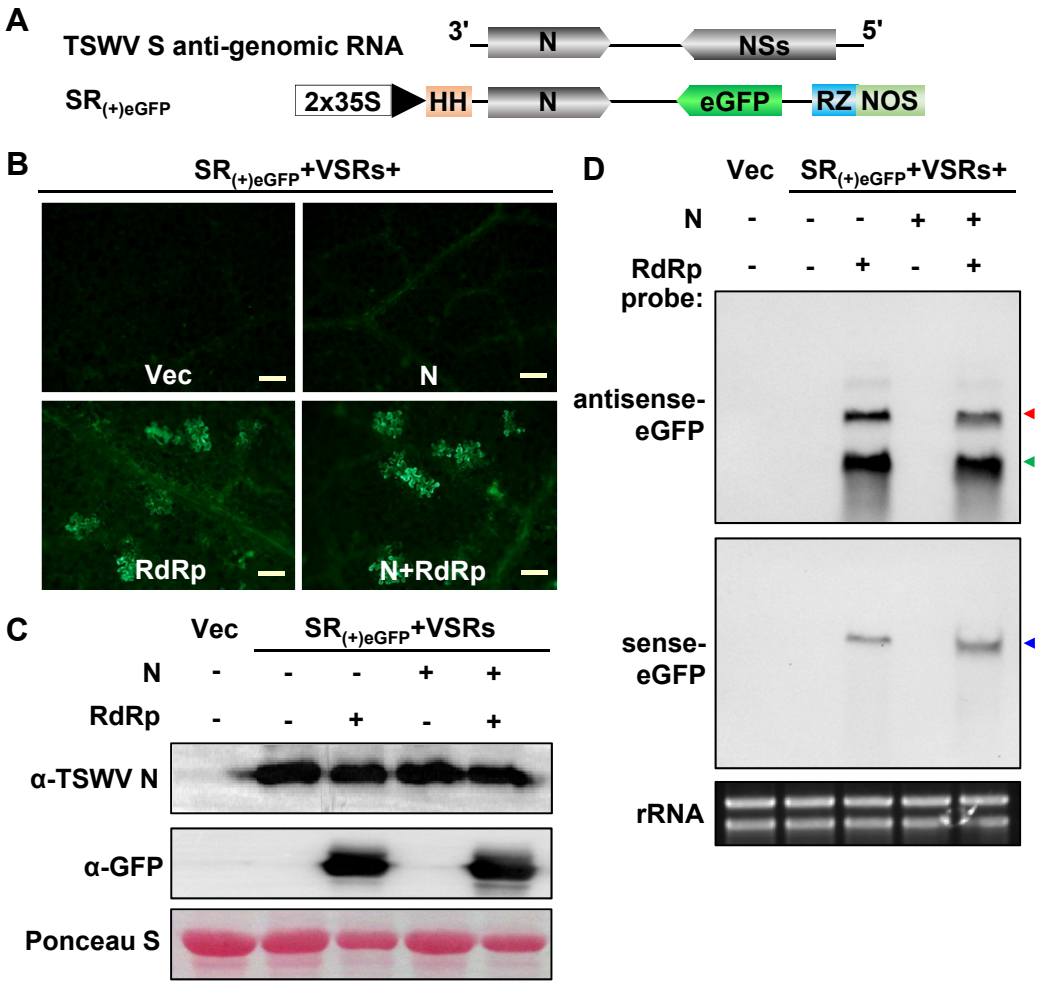


Fig. 3

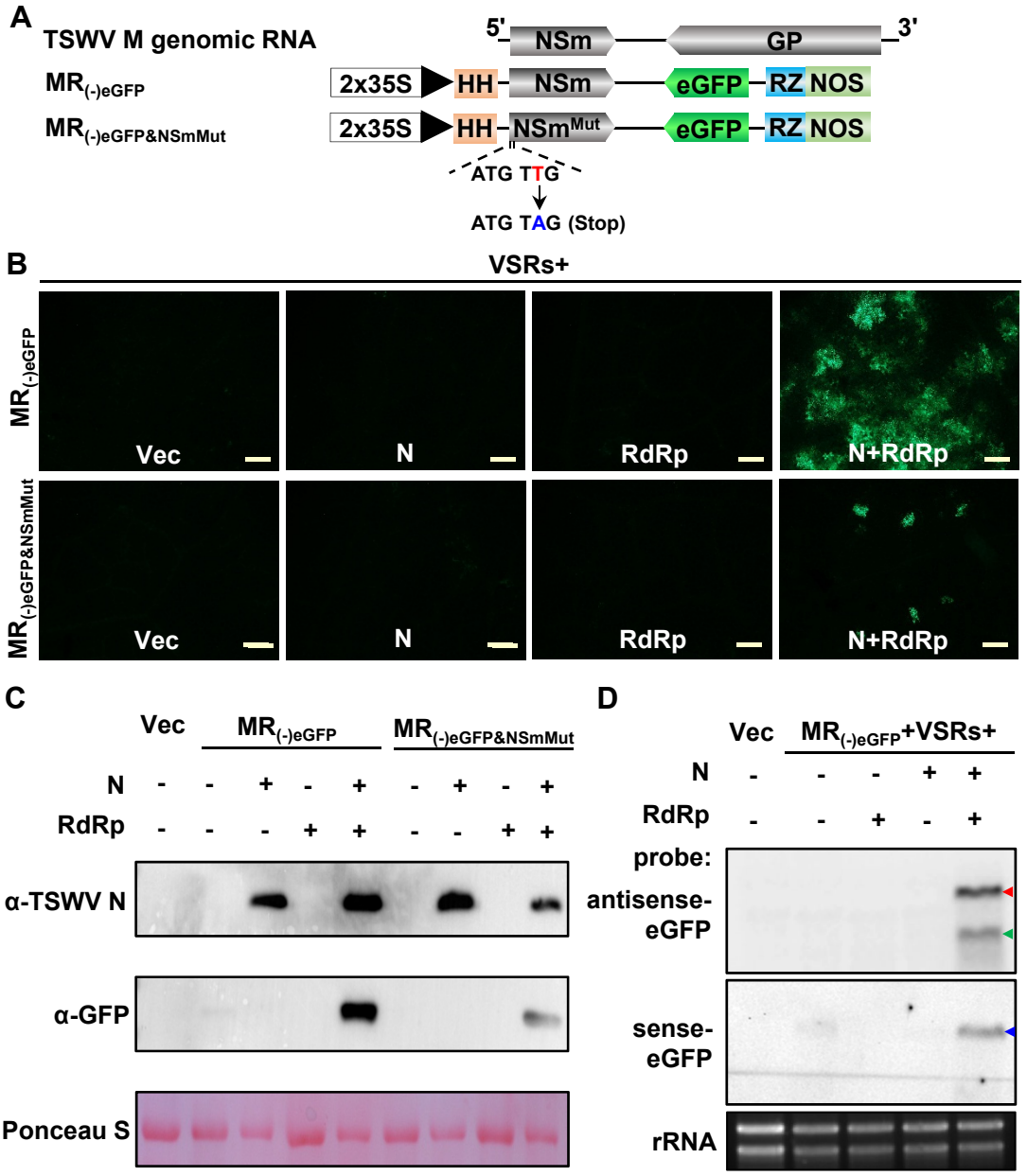


Fig. 4

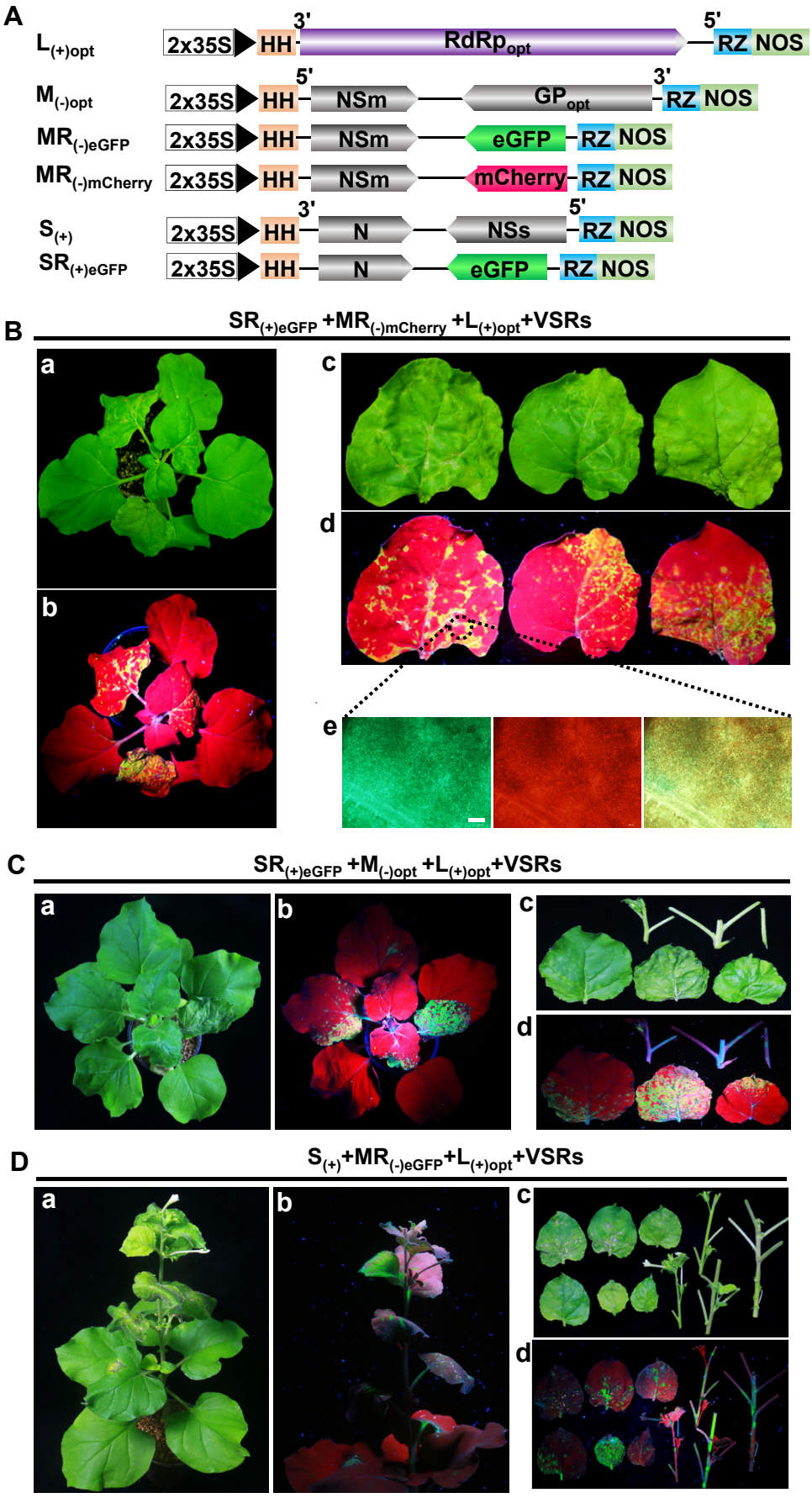


Fig. 5

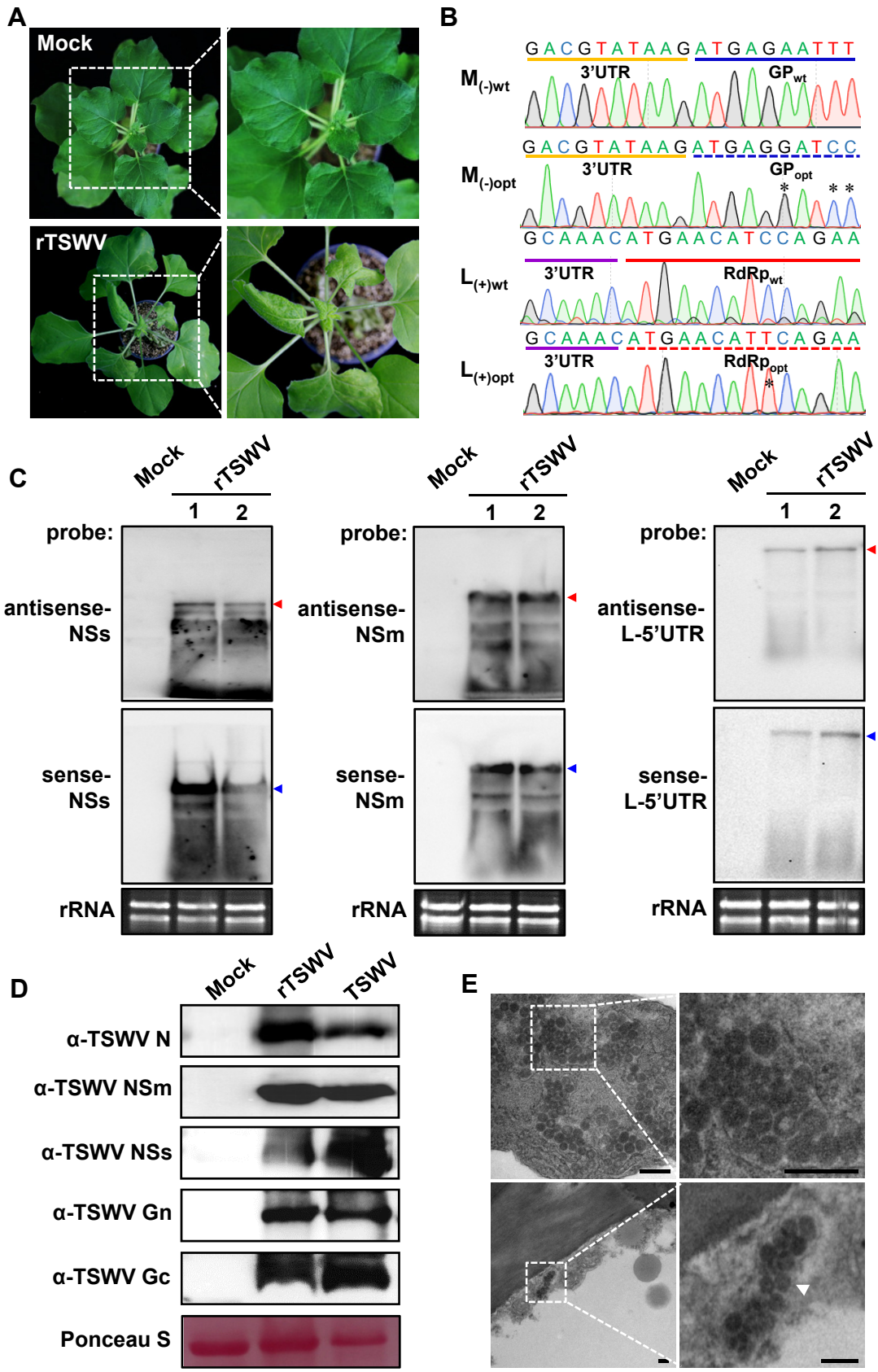


Fig. 6

Supplementary Materials and Methods

Plasmid construction

Construction of RdRp, RdRp$_{opt}$, N, NSs and VSRs. The cDNA of the RdRp, N, and NSs genes was amplified from the total RNA of TSWV-lettuce isolate infected tissues and inserted into a binary vector pCambia2300 or pCXSN to generate p2300-RdRp$_{wt}$, p2300-N, and pCXSN-NSs downstream a double 35S promoter (2×35S). The P19-HcPro-γb carrying three VSRs simultaneously in the pCB301 vector was kindly provided by Dr. Xianbing Wang in College of Biological Sciences of China Agricultural University. The codon usage and intron-splicing sites optimized sequence of RdRp (*SI Appendix,* Fig. S9) was *de novo* synthesized by GenScript Biotech Corp (Nanjing, China) and was inserted into a binary vector pCambia2300 to generate p2300S-RdRp$_{opt}$ downstream a 2×35S promoter.

Construction of full length TSWV genomic S$_{(-)}$, M$_{(-)}$, L$_{(-)}$, and anti-genomic S$_{(+)}$, M$_{(+)}$ and L$_{(+)}$ cDNA clones. To generate constructs to express full length TSWV genomic RNA and antigenomic RNA S, M and L segments, total RNA extracted from TSWV-lettuce infected leaves of *N. benthamiana* plants was reverse transcribed into cDNA, followed by PCR amplification with specific primers (*SI Appendix,* Table S3) using Phanta Super-Fidelity DNA Polymerase (Vazyme Biotech, Nanjing, China). The PCR products were fused with a self-cleaving hammerhead (HH) ribozyme (1) and inserted into a binary expression vector pCB301-2×35S-RZ-NOS linearized by two restriction endonucleases *Stu* I and *Sma* I (2). The pCB301-2×35S-HH-S$_{(-)}$-RZ-NOS [S$_{(-)}$], pCB301-2×35S-HH-M$_{(-)}$-RZ-NOS [M$_{(-)}$], pCB301-2×35S-HH-L$_{(-)}$-RZ-NOS [L$_{(-)}$], pCB301- pCB301-2×35S-HH-S$_{(+)}$-RZ-NOS [S$_{(+)}$], 2×35S-HH-M$_{(+)}$-RZ-NOS [M$_{(+)}$] and pCB301-2×35S-HH-L$_{(+)}$-RZ-NOS [L$_{(+)}$]

cDNA clones were generated. The full length TSWV genomic $S_{(-)}$, $M_{(-)}$ and $L_{(-)}$, and anti-genomic $S_{(+)}$, $M_{(+)}$ and $L_{(+)}$ were expressed downstream a double 35S promoter (2×35S) and franked with a self-cleaving hammerhead (HH) ribozyme at 5'-terminus and a Hepatitis delta virus (HDV) ribozyme at 3'-terminus.

Construction of TSWV $SR_{(-)eGFP}$, $SR_{(-)mCherry\&eGFP}$ and $SR_{(+)eGFP}$ minireplicons. To generate $SR_{(-)eGFP}$-genomic RNA minireplicon, the eGFP ORF was amplified and used to replace the N gene in the pCB301-2×35S-HH-$S_{(-)}$-RZ-NOS by the *in vitro* recombination using the In-Fusion Cloning mixture (Clontech, Japan). The construct pCB301-2×35S-HH-$SR_{(-)eGFP}$-RZ-NOS [$SR_{(-)eGFP}$] was generated.

To generate $SR_{(-)mCherry\&eGFP}$ in which the NSs and N genes in S gRNA were replaced with mCherry and eGFP, respectively, the mCherry ORF was amplified and used to exchange the NSs gene in the pCB301-2×35S-HH-$SR_{(-)eGFP}$-RZ-NOS by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-$SR_{(-)mCherry\&eGFP}$-RZ-NOS [35S:$SR_{(-))mCherry\&eGFP}$] was generated. The T7:$SR_{(-)mCherry\&eGFP}$ minireplicon (pCB301-T7-HH-$SR_{(-)mCherry\&eGFP}$-RZ-NOS) controlled by T7 promoter was constructed by the same strategy as 35S:$SR_{(-)mCherry\&eGFP}$.

To generate antigenomic $S_{(+)eGFP}$-minireplicon, the eGFP ORF was amplified and used to replace the NSs gene in the pCB301-2×35S-HH-$S_{(-)}$-RZ-NOS by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-$SR_{(+)eGFP}$-RZ-NOS [$SR_{(+)eGFP}$] was generated. The primers used above are listed in *SI Appendix,* Table S3.

Construction of $SR_{(-)mCherry\&eGFP\Delta5'UTR}$, $SR_{(-)mCherry\&eGFP\Delta IGR}$ and $SR_{(-)mCherry\&eGFP\Delta3'UTR}$ mutants. To generate $SR_{(-)mCherry\&eGFP\Delta5'UTR}$, the DNA copy of $SR_{(-)mCherry\&eGFP}$ without 5'-

UTR (88 nt) was amplified from pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS and used to recombine with backbone vector of pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP\Delta5'UTR}$-RZ-NOS [SR$_{(-)mCherry\&eGFP\Delta5'UTR}$] was generated.

To generate SR$_{(-)mCherry\&eGFP\Delta IGR}$, the DNA copy of SR$_{(-)mCherry\&eGFP}$ without IGR (550 nt) was amplified from pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS and used to recombine with the vector backbone from pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP\Delta IGR}$-RZ-NOS [SR$_{(-)mCherry\&eGFP\Delta IGR}$] was generated.

To generate SR$_{(-)mCherry\&eGFP\Delta3'UTR}$, the DNA copy of SR$_{(-)mCherry\&eGFP}$ without 3'-UTR (151 nt) was amplified from pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS and used to recombine with the vector backbone from pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP}$-RZ-NOS using the primer pair FMF48/P3382 by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-SR$_{(-)mCherry\&eGFP\Delta3'UTR}$-RZ-NOS [SR$_{(-)mCherry\&eGFP\Delta3'UTR}$] was generated. The primers used above are listed in *SI Appendix,* Table S3.

Construction of TSWV MR$_{(-)eGFP}$, MR$_{(-)mCherry}$ and MR$_{(-)eGFPNSmMut}$ minireplicons. To generate MR$_{(-)eGFP}$ and MR$_{(-)mCherry}$ minireplicons, the eGFP and mCherry ORFs were amplified and used to replace the GP gene in pCB3012×35S-HH-M$_{(-)}$-RZ-NOS, respectively, by recombination using In-Fusion Cloning mixture (Clontech). The constructs pCB301-2×35S-HH-MR$_{(-)eGFP}$-RZ-NOS [MR$_{(-)eGFP}$] and pCB301-2×35S-HH-MR$_{(-)mCherry}$-RZ-NOS [MR$_{(-)mCherry}$] was generated.

To generate MR$_{(-)eGFPNSmMut}$ in which a stop codon was introduced immediately after

the start codon of NSm, the NSm^{Mut} was amplified and used to replace the wild type NSm sequence in pCB301-2×35S-HH-$MR_{(-)eGFP}$-RZ-NOS by recombination using In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-$MR_{(-)eGFPNSmMut}$-RZ-NOS [$MR_{(-)eGFPNSmMut}$] was generated. All primers used above are listed in *SI Appendix,* Table S3.

Construction of full length $L_{(+)opt}$ and $M_{(-)opt}$ cDNA clones. To generate full length $L_{(+)opt}$ cDNA clone**,** the sequence codon and intron-splicing sites optimized RdRp was amplified and used to replace the wild type RdRp sequence in pCB301-2×35S-HH-$L_{(+)}$-RZ-NOS by recombination using the In-Fusion Cloning mixture (Clontech). The pCB301-2×35S-HH-$L_{(-)opt}$-RZ-NOS [$L_{(-)opt}$] was generated.

To generate full length $M_{(-)opt}$ cDNA clone**,** the codon and intron-splicing sites optimized GP gene was de novo synthesized by GenScript Biotech Corp (Nanjing, China) (*SI Appendix,* Fig. S10) and used to replace the wild type GP sequence in pCB301-2×35S-HH-$M_{(-)}$-RZ-NOS by the *in vitro* recombination using In-Fusion Cloning mixture (Clontech). The pCB301-2×35S-HH-$M_{(-)opt}$-RZ-NOS [$M_{(-)opt}$] was generated. The primers used above are listed in *SI Appendix,* Table S3.

Plant material and virus source

Six to eight weeks of *Nicotiana benthamiana* was used in all agroinfiltration assay. *N. benthamiana* plants were grown in a growth chamber setting at 25 °C, a 16 h light and 8 h dark photoperiod (3). The TSWV isolate from asparagus lettuce (TSWV-LE) was used in this study (GenBank accession number: KU976396 for S, JN664253 for M and KU976394 for L) (4). The TSWV-LE isolate was maintained on *N. benthamiana*. For long-term storage, the infected new leaves of *N. benthamiana* were kept in an 80 °C refrigerator.

***Agrobacterium* infiltration**

94 Recombinant plasmids were electroporated into *Agrobacterium tumefaciens* strain
95 GV3101 and agroinfiltrations were performed essentially as described (5, 6). *A.*
96 *tumefaciens* cells were resuspended by agroinfiltration buffer [10 mM MgCl2, 10 mM
97 MES (pH 5.6) and 100 μM acetosyringone] adjusted to an optical density OD_{600} of 1.0 and
98 incubated for 2 to 3 h in dark at room temperature. Equal volumes of *Agrobacterium*
99 cultures (final concentration OD_{600}=0.2) harboring the p2300-N, p2300-RdRp, pCB301-
100 derived reporter or full-length infectious clone vector(s), were mixed with one volume of
101 bacterial mixture (final concentration OD_{600}=0.05) containing the NSs and P19-HcPro-γb.
102 The *Agrobacterium* cultures were infiltrated into fully expanded leaves of 6-7 leaf stage *N.*
103 *benthaminan* plants using 1 mL needleless syringes.

104 **Immunoblot analysis**

105 Total protein was extracted from 1 g *Agrobacterium*-infiltrated leaf patches, healthy or
106 TSWV-infected *N. benthamiana* systemic leaves in a 1 mL extraction buffer [10 % (v/v)
107 glycerol, 25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10 mM dithiothreitol, 2 %
108 (w/v) polyvinylpolypyrrolidone, 0.5 % (v/v) Triton X-100 and 1× protease inhibitors
109 cocktail] (7). Protein samples were separated by SDS-PAGE gels, transferred to PVDF
110 membranes (GE Healthcare, UK), blocked with 5 % skim milk solution and incubated with
111 a polyclonal antiserum specific to the TSWV N, NSm, NSs, Gn, Gc, GFP, mCherry or T7
112 RNA pol at room temperature for 1 h or overnight at 4 ℃ and washed three times. After
113 incubation in a secondary antibody containing HRP-conjugated goat anti-rabbit (1:10000)
114 for 1 h, the blots were detected using the ECL Substrate Kit (Thermo Scientific, Hudson,
115 NH, USA). To evaluate protein loading, the blots were stained with Ponceau S.

116 **Northern blot analysis**

117 For Northern blot analysis of TSWV gRNAs, agRNAs or viral mRNA transcripts, total

RNAs were extracted from *Agrobacterium*-infiltrated leaf patches, healthy or TSWV-infected systemic leaves using an RNAprep Pure Plant Kit (Tiangen Biotech, Beijing, China), respectively. DIG-labeled specific probes for sense or antisense GFP, NSs, NSm, L-5'UTR was synthesized by DIG High Prime RNA labeling kit (Roche, Basel, Switzerland). The total RNAs were separated on 1 % formaldehyde agarose gels and transferred to Hybond-N+ membranes (GE Healthcare, UK) (8). The membrane blots were hybridized with a DIG-labeled specific probe and detected using a DIG-High Prime Detection Starter Kit II (Roche), following the manufacturer's protocol.

RT-PCR and sequencing analysis

To detect the virus in systemic leaves of *N. benthamina* infected with $SR_{(+)eGFP}$+$MR_{(-)mCherry}$+$L_{(+)opt}$, $S_{(+)}$+$MR_{(-)eGFP}$+$L_{(+)opt}$, $SR_{(+)eGFP}$+$M_{(-)opt}$+$L_{(+)opt}$ or rTSWV recovered from the full-length cDNA clones, total RNAs were extracted from systemic symptoms plant leaves. First-strand cDNAs were synthesized using M-MLV Reverse Transcriptase (Promega, USA). RT-PCRs were performed to detect the $SR_{(+)eGFP}$, $MR_{(-)mCherry}$, $MR_{(-)eGFP}$, $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ minigenome and genomic RNA using their specific-primers. The PCR products were inserted into a pMD19-T vector (Takara, Dalian, China) and sequenced by Sanger dideoxy-mediated chain-termination DNA sequencing method at Sangon Biotech (Shanghai, China). The primers used in this study are listed in *SI Appendix,* Table S3.

Fluorescence microscopy

The agro-infiltrated *N. benthamiana* leaves were examined for fluorescence expression using an OLYMPUS IX71-F22FL/DIC Inverted Fluorescence Microscope (OLYMPUS, Tokyo, Japan) with a green or red barrier filter. The leaf sample was fixed in water on a

141 microslider under a coverslip to detect the eGFP and mCherry fluorescence, respectively.
142 Fluorescence images were processed using ImagePro (OLYMPUS, Tokyo, Japan) and
143 Adobe (San Jose, CA, USA) Photoshop programs.

144 **Electron microscopy**

145 Small tissues (1 mm × 4 mm) were excised from leaves of the *N. benthamiana* with infected
146 rTSWV rescued from the full-length infectious clones. The sample tissues were fixed in
147 2.5 % glutaraldehyde and 1 % osmium tetroxide dissolving into 100 mM phosphate buffer
148 (pH 7.0) as described by Li *et al* (5, 9) and then embedded in Epon 812 resin as instructed
149 by the manufacture (SPI-EM, Division of Structure Probe, Inc., West Chester, USA).
150 Ultrathin sections (70 nm) were mounted on formvar-coate grids and then stained with
151 uranyl acetate for 10 min followed by lead citrate for 10 min. The stained sections were
152 examined under a transmission electron microscope (TEM; H-7650, Hitachi, Japan).

153 **Imaging GFP in infected plant by hand-held UV lamp**

154 GFP fluorescence in leaves was monitored with a hand-held 100 W, long-wave UV lamp
155 (UV Products, Upland, CA, USA) and the leaves were photographed using a Canon EOS
156 70D digital camera (Canon, Japan) with a 58 mm UV filter.

157

158 **References**

159 1. Herold J & Andino R (2000) Poliovirus requires a precise 5' end for efficient positive-strand RNA
160 synthesis. *J Virol* 74(14):6394-6400.
161 2. Shen Y*, et al.* (2014) A versatile complementation assay for cell-to-cell and long distance
162 movements by cucumber mosaic virus based agro-infiltration. *Virus Res* 190:25-33.
163 3. Voinnet O & Baulcombe DC (1997) Systemic signalling in gene silencing. *Nature* 389(6651):553.
164 4. Jiang L*, et al.* (2017) Occurrence and diversity of Tomato spotted wilt virus isolates breaking the
165 Tsw resistance gene of Capsicum chinense in Yunnan, southwest China. *Plant Pathol* 66(6):980-
166 989.
167 5. Wang Q*, et al.* (2015) Rescue of a plant negative-strand RNA virus from cloned cDNA: insights
168 into enveloped plant virus movement and morphogenesis. *PLoS Pathog* 11(10):e1005223.
169 6. Ganesan U*, et al.* (2013) Construction of a Sonchus yellow net virus minireplicon: a step toward

reverse genetic analysis of plant negative-strand RNA viruses. *J Virol* 87(19):10598-10611.

7. Li J*, et al.* (2019) A plant immune receptor adopts a two-step recognition mechanism to enhance viral effector perception. *Mol Plant* 12(2):248-262.

8. Feng MF*, et al.* (2018) Identification of Strawberry vein banding virus encoded P6 as an RNA silencing suppressor. *Virology* 520:103-110.

9. Kong L, Wu J, Lu L, Xu Y, & Zhou X (2014) Interaction between Rice stripe virus disease-specific protein and host PsbP enhances virus symptoms. *Mol Plant* 7(4):691-708.

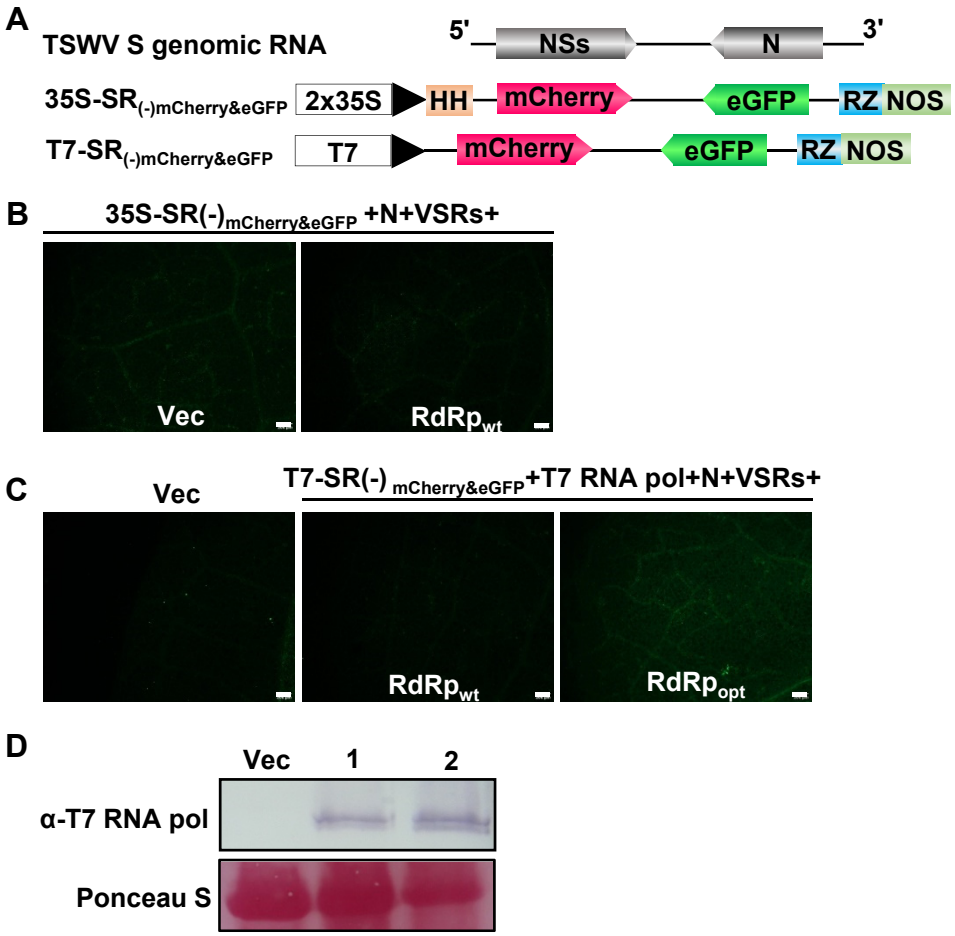
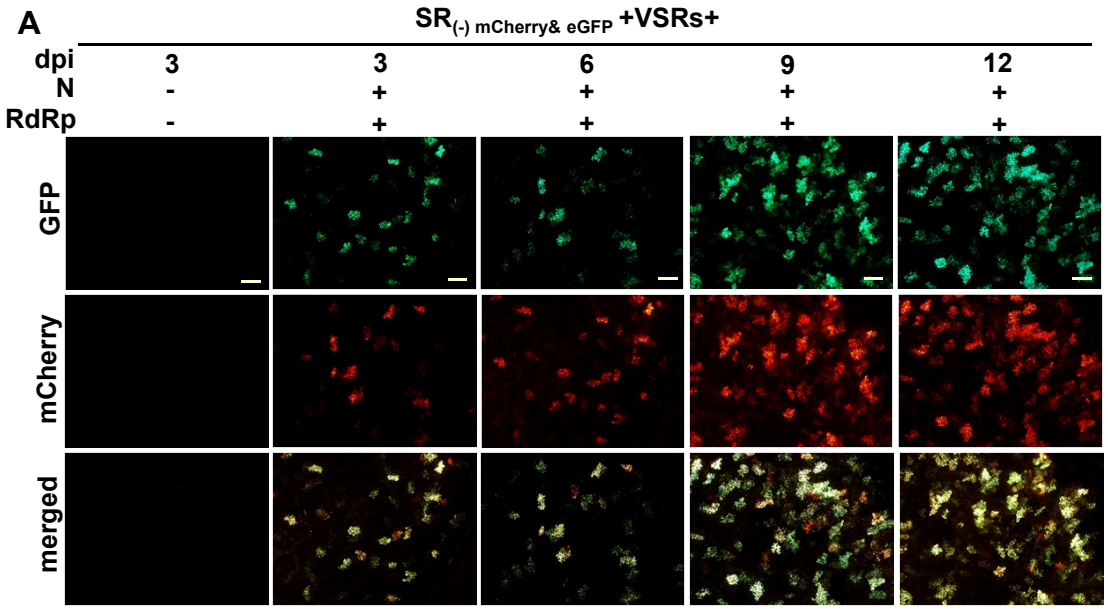


Fig. S1



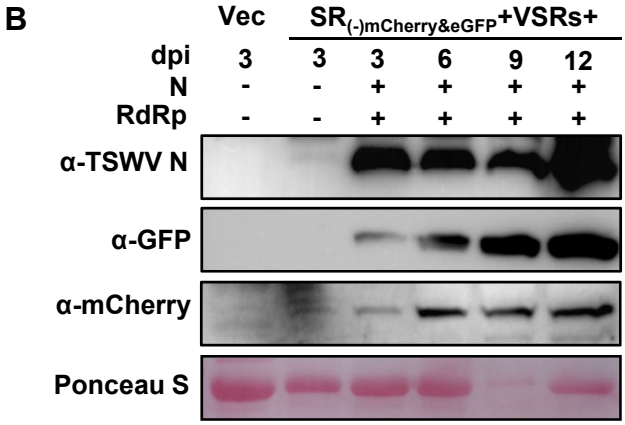


Fig. S2

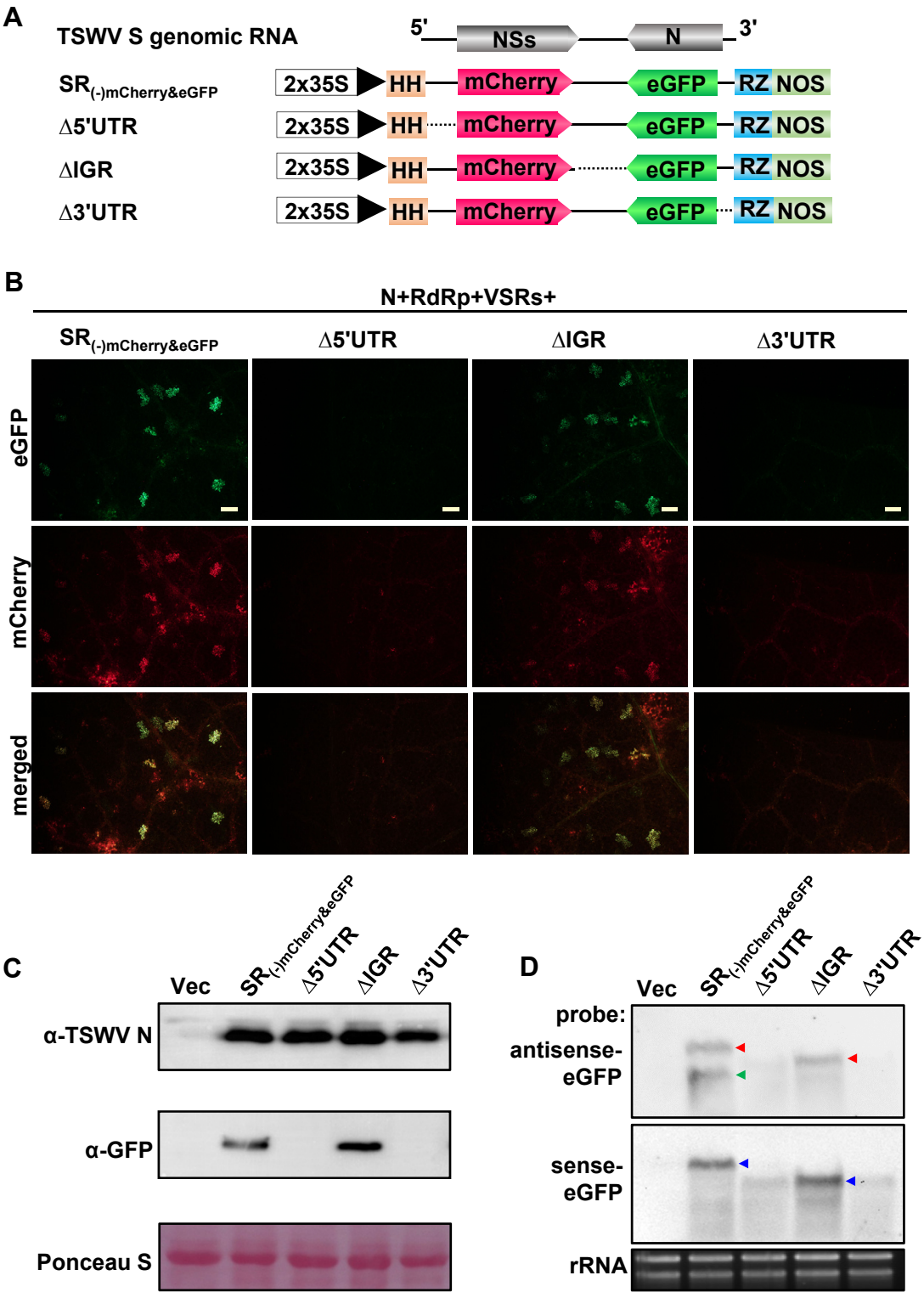


Fig. S3

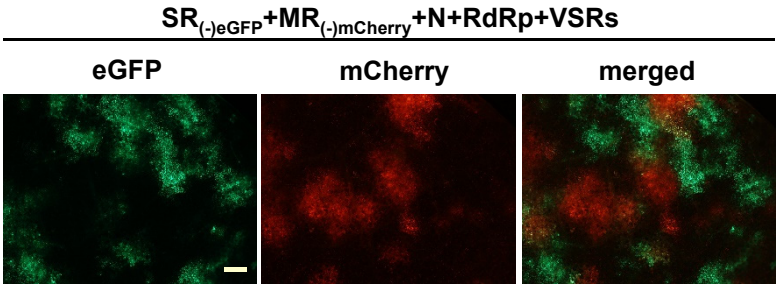


Fig. S4

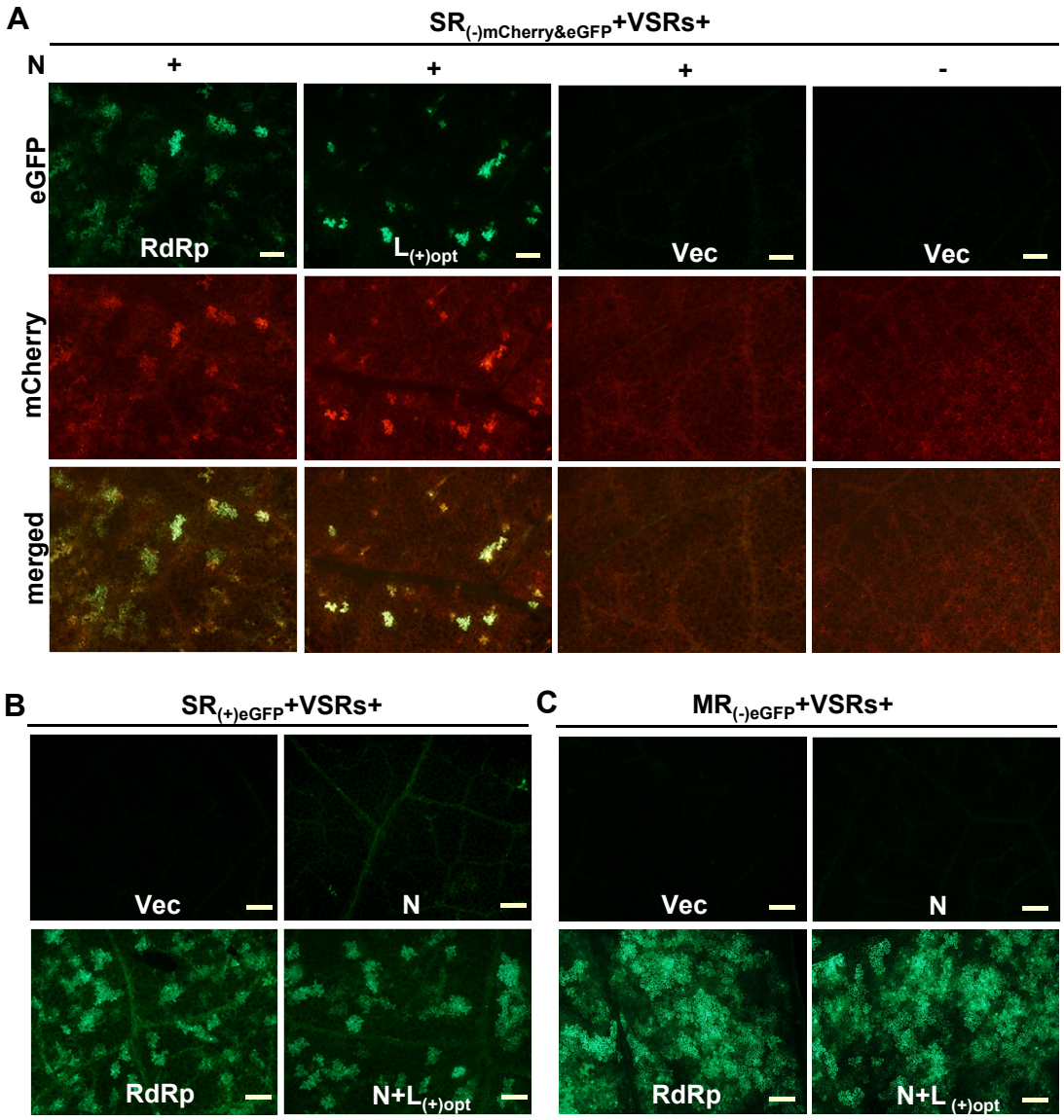


Fig. S5

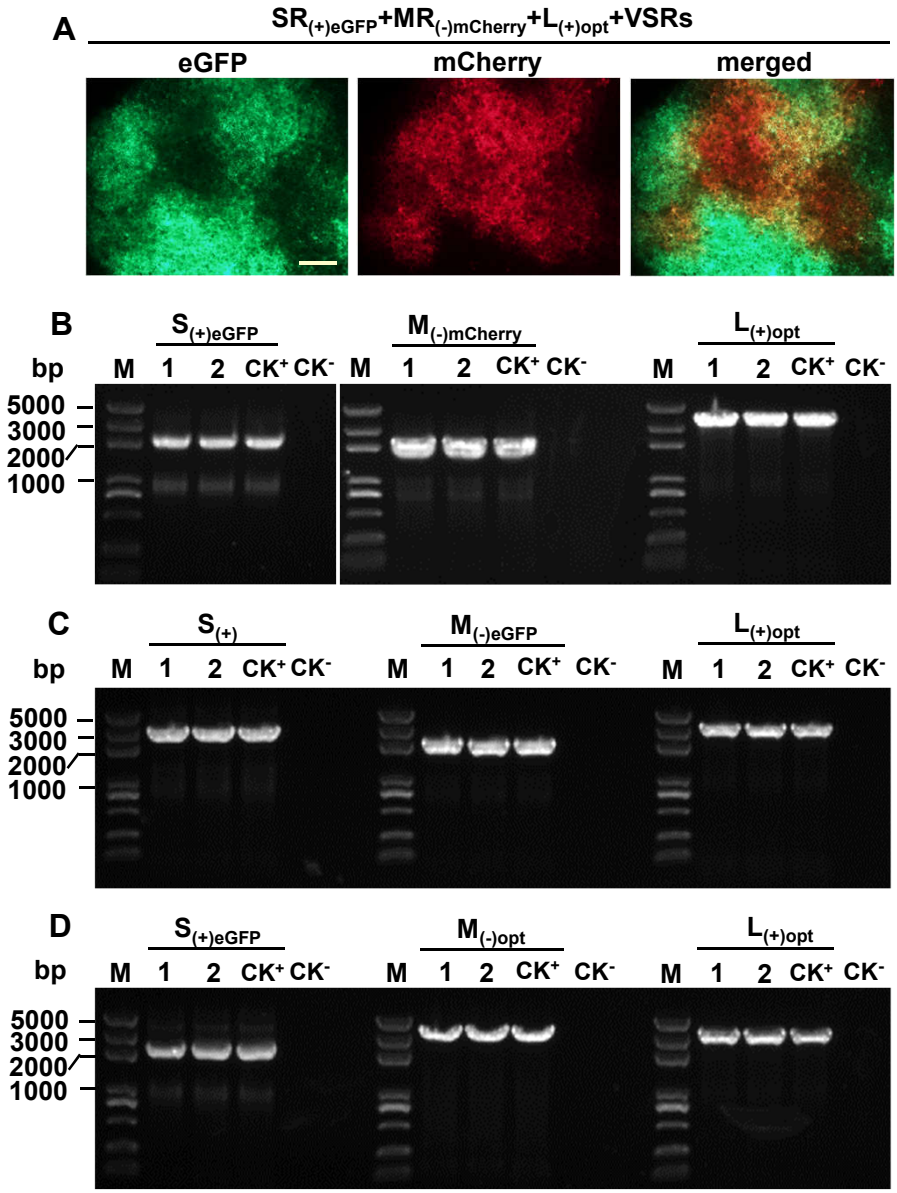


Fig. S6

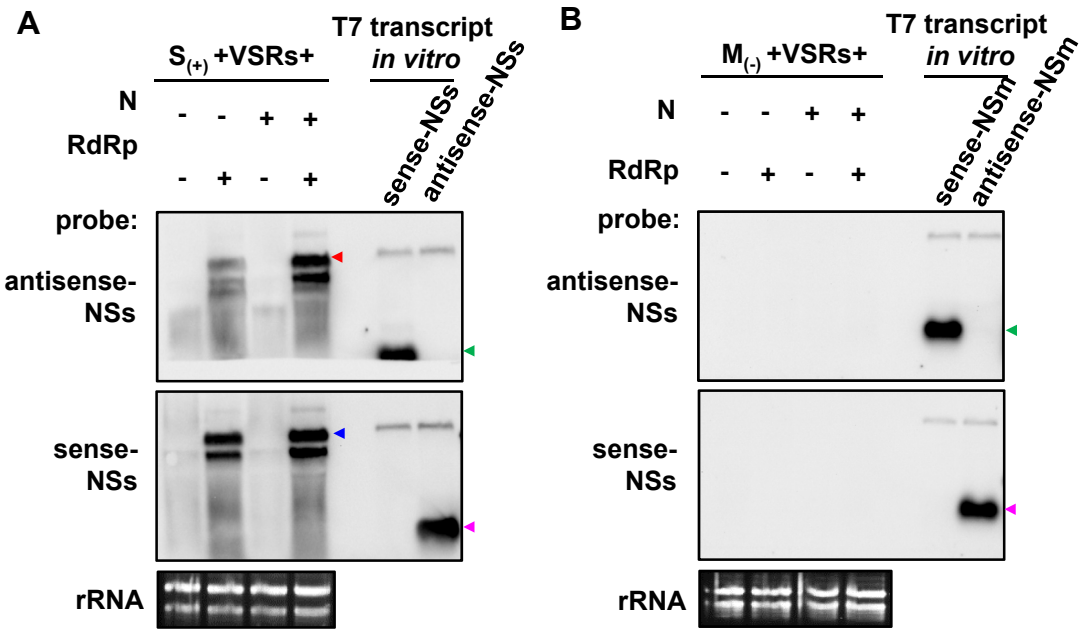


Fig. S7

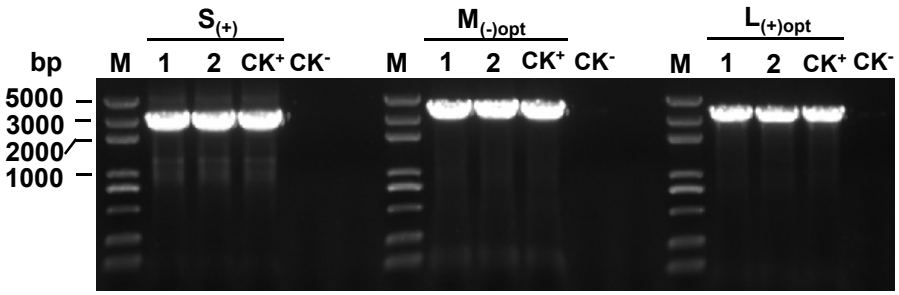


Fig. S8

Fig. S9. Optimized RdRp gene sequence used in the study.

ATGAACATTCAGAAGATCCAAAAGCTGATCGAGAACGGCACCACCTTGCTGCTTTCTATTGAGGATTGCGTGGGCAGCAA

CTACGATCTTGCTCTTGATCTGCACAAGCGGAACTCCGATGAGATTCCTGAGGACGTGATCATCAACAACAACGCCAAGA

ACTACGAGACTATGCGTGAGCTGATCGTGAAGATCACTGCTGATGGTGAGGGTCTGAACAAAGGTATGGCTACCGTGGAC

GTGAAGAAACTGTCTGAGATGGTGTCTCTGTTCGAGCAGAAGTACCTTGAGACTGAGCTTGCTAGGCACGATATCTTCGG

CGAGCTTATTTCTAGGCACCTGAGGATCAAGCCGAAGCAGAGATCTGAGGTTGAGATTGAGCACGCTCTGAGAGAGTAC

CTGGACGAGCTTAACAAGAAGTCCTGCATCAACAAGCTGAGCGACGATGAGTTCGAGCGGATCAACAAAGAATATGTGG

CCACCAACGCTACCCCGGACAATTACGTTATCTACAAAGAGAGCAAGAACAGCGAGCTGTGCCTGATCATCTACGACTGG

AAGATCTCTGTGGATGCCAGGACTGAGACTAAGACCATGGAAAAGTACTACAAGAACATCTGGAAGTCCTTCAAGGACAT

CAAGGTGAACGGCAAGCCTTTCTTGGAGGATCATCCTGTGTTCGTGAGCATCGTGATCCTGAAGCCTATTGCTGGCATGC

CTATTACCGTGACCTCTAGCAGAGTGCTCGAGAAGTTTGAGGACTCTCCTTCTGCACTTCACGGCGAGAGGATTAAGCAC

GCTAGGAACGCTAAGCTGCTGAACATTTCTCACGTGGGTCAGATCGTGGGTACTACTCCTACTGTGGTGAGGAACTACTA

CGCTAACACCCAGAAGATCAAGTCCGAGGTTAGGGGTATCCTGGGTGATGATTTCGGCAGCAAGGACGTGTTCTTCTCTC

ACTGGACCTCCAAGTACAAAGAGCGGAACCCTACCGAGATCGCTTACTCTGAGGATATCGAGCGTATCATCGACAGCCTT

GTGACCGACGAGATCACCAAAGAAGAGATTATCCACTTCCTGTTCGGCAACTTCTGCTTCCACATCGAGACAATGAACGA

CCAGCACATTGCCGACAAGTTCAAGGGCTATCAGTCCTCTTGCATCAACCTGAAGATCGAGCCTAAGGTGGACCTGGCTG

ATCTTAAGGATCACCTGATTCAGAAGCAGCAGATTTGGGAGAGCCTGTACGGTAAGCACCTCGAGAAGATTATGCTGCGG

ATCCGTGAAAAGAAGAAGAAAGAGAAAGAGATCCCGGACATCACCACCGCCTTCAATCAAAATGCTGCCGAGTACGAAG

AGAAGTACCCGAACTGCTTCACCAACGACCTGTCAGAGACAAAGACCAACTTCTCTATGACCTGGTCACCGAGCTTCGA

GAAAATCGAGCTGTCTAGCGAGGTGGACTACAACAATGCCATTATCAACAAGTTCCGTGAGAGCTTCAAGTCCAGCAGCA

GGGTGATCTACAACAGCCCTTACAGCTGCATTAACAACCAGACCAACAAGGCCCGGGATATCACTAACCTTGTGAGGCTT

TGCCTGACCGAGCTTTCTTGCGATACCACCAAGATGGAAAAGCAAGAGCTTGAGGACGAGATCGACATCAACACCGGTA

GCATTAAGGTCGAGCGGACCAAGAAAAGCAAAGAGTGGAACAAGCAGGGCTCCTGCCTTACTAGGAACAAGAACGAGTT

CTGCATGAAGGAAACCGGCCGAGAGAACAAGACCATCTACTTCAAGGGTCTTGCCGTGATGAACATCGGCATGTCCTCTA

AGAAGCGGATTCTGAAAAAAGAGGAAATCAAAGAGAGGATCAGCAAGGGCCTCGAGTACGATACTTCTGAGAGACAGGC

TGATCCGAACGACGACTACAGCAGCATCGATATGTCATCACTGACCCATATGAAGAAGCTCATCCGGCACGATAACGAGG

ACTCACTTTCTTGGTGCGAGCGAATCAAGGACAGCTTGTTCGTGCTTCACAACGGCGATATCAGAGAAGAGGGCAAGAT

CACCAGCGTGTACAACAATTACGCTAAGAACCCCGAGTGCCTGTACATCCAGGATTCTGTGCTTAAGACCGAGCTGGAAA

CCTGCAAGAAAATCAACAAACTGTGCAACGATCTGGCCATCTACCACTACAGCGAGGACATGATGCAGTTCTCCAAGGGT

TTGATGGTGGCCGATCGGTACATGACCAAAGAGTCTTTCAAGATCCTGACCACCGCCAATACCAGCATGATGCTGCTTGC

TTTCAAAGGCGACGGTATGAACACTGGTGGTTCTGGTGTTCCTTACATTGCCCTGCACATCGTGGATGAGGATATGTCCG

ATCAGTTCAACATCTGCTACACCAAAGAGATCTACAGCTACTTCCGGAACGGCTCCAACTACATCTACATCATGAGGCCGC

AGAGGCTTAACCAGGTGAGGCTTCTTTCACTGTTCAAGACCCCTTCTAAGGTGCCAGTTTGCTTCGCCCAGTTCAGCAAA

AAGGCCAACGAAATGGAAAAATGGCTGAAGAACAAGGACATTGAGAAGGTCAACGTGTTCAGCATGACCATGACCGTGA

AGCAGATCCTGATCAACATCGTGTTCTCCAGCGTGATGATTGGCACCGTGACTAAGCTTTCTCGGATGGGCATCTTCGAC

TTCATGAGGTACGCTGGTTTCTTGCCGCTGTCCGACTACTCCAACATCAAAGAGTACATCCGGGACAAGTTCGACCCCGA

TATTACCAATGTGGCCGATATCTACTTCGTGAACGGGATCAAGAAACTGCTGTTCCGGATGGAAGATTTGAACCTGAGCA

CCAACGCAAAGCCTGTTGTGGTGGATCACGAGAACGATATCATCGGTGGCATCACCGACCTGAACATCAAGTGCCCTATT

ACTGGTTCTACCCTGCTGACCCTTGAGGACCTGTATAACAATGTGTACCTCGCCATCTACATGATGCCGAAGTCTCTGCAT

AACCACGTGCACAACCTTACCAGCCTGCTTAATGTTCCTGCTGAGTGGGAGCTGAAGTTCCGGAAAGAGCTTGGCTTCA

ACATTTTCGAGGACATCTACCCGAAGAAAGCCATGTTCGATGACAAGGACCTCTTCAGCATTAACGGCGCTCTTAACGTG

AAGGCCCTGAGCGATTACTACCTGGGTAACATCGAGAATGTGGGCCTGATGAGGTCCGAGATTGAGAACAAAGAGGACT

TCCTGTCTCCGTGCTACAAGATCTCTACCCTGAAGTCCAGTAAGAAGTGCAGCCAGAGCAACATCATCAGCACTGATGAG

ATCATTGAGTGCCTGCAGAACGCAAAGATTCAGGACATCGAAAACTGGAAGGGCAACAACCTGGCTATTATCAAGGGCCT

GATCCGGACCTACAACGAGGAAAAGAATCGGCTGGTTGAGTTCTTCGAGGATAACTGCGTGAACAGCCTGTACCTGGTC

GAGAAGCTTAAAGAGATCATTAACAGCGGCAGCATCACCGTGGGAAAGTCTGTGACTAGCAAGTTCATCCGTAACAATCA

CCCGCTGACCGTGGAAACCTACCTCAAGACTAAGCTGTACTATCGGAACAACGTGACCGTGCTGAAGTCTAAGAAGGTG

AGCGAGGAACTGTACGACCTCGTTAAGCAGTTCCACAACATGATGGAAATCGACCTGGACTCTGTGATGAACCTTGGTAA

GGGTACTGAGGGGAAGAAGCACACCTTCTTGCAGATGCTTGAGTTCGTGATGAGCAAGGCCAAGAATGTGACCGGTTCT

GTGGATTTCCTCGTGAGCGTGTTCGAAAAGATGCAGAGGACCAAGACCGACCGTGAGATCTACCTGATGAGCATGAAGG

TGAAGATGATGCTGTACTTCATCGAGCATACCTTCAAGCACGTGGCCCAGTCTGATCCTTCTGAGGCTATTAGCATCAGCG

GCGACAACAAGATTAGGGCTCTGTCTACCCTTAGCCTGGACACCATTACCAGCTACAACGACATCCTCAACAAGAATAGC

AAGAAGTCTCGGCTGGCTTTCCTGAGCGCTGATCAATCTAAGTGGTCCGCTTCTGACCTGACCTACAAGTACGTGCTGGC

CATCATTCTGAACCCGATTCTTACTACTGGCGAGGCCTCTCTTATGATCGAGTGCATTCTGATGTACGTGAAGCTGAAGAA

AGTGTGCATCCCGACCGACATTTTCTTGAACCTTAGGAAGGCTCAGGGCACCTTCGGTCAAAACGAAACCGCTATTGGTC

TGCTGACCAAGGGACTTACCACCAACACTTACCCGGTGTCTATGAACTGGCTGCAGGGTAACCTGAACTACCTCTCTTCT

GTGTACCACTCCTGCGCTATGAAGGCTTACCACAAGACTCTCGAGTGCTATAAGGACTGCGACTTTCAGACCCGGTGGAT

CGTTCACTCTGATGATAACGCAACCAGCCTGATCGCTTCAGGTGAGGTTGACAAGATGCTGACCGACTTCTCCTCTTCAT

CTCTGCCTGAGATGCTGTTCAGATCCATCGAGGCTCACTTCAAGTCTTTCTGCATCACACTCAACCCCAAGAAGTCATAC

GCCAGCTCAAGCGAGGTCGAGTTCATTTCTGAGAGGATTGTGAACGGCGCTATCATCCCACTTTACTGCAGGCATCTTGC

TAACTGCTGCACCGAGTCCTCTCACATCTCCTACTTCGATGACCTGATGTCCCTGTCTATCCACGTGACCATGTTGCTGAG

AAAGGGCTGCCCTAATGAGGTGATCCCTTTTGCTTACGGTGCTGTGCAAGTGCAGGCTCTCTCTATCTACTCTATGCTGC

CTGGTGAGGTGAACGACAGCATCCGGATCTTTAAGAAGCTGGGCGTCAGCCTCAAGTCCAACGAGATTCCTACTAACATG

GGCGGTTGGCTGACCTCTCCTATTGAGCCTTTGTCTATTCTGGGCCCCAGCAGCAACGACCAGATTATCTACTACAACGT

GATCCGGGATTTCCTGAACAAGAAATCCCTGGAAGAGGTGAAGGACTCCGTGTCCTCATCTTCTTACCTGCAGATGAGGT

TCAGGGAACTCAAGGGCAAGTACGAGAAGGGTACTCTGGAAGAGAAGGACAAAAAGATGATCTTCCTTATCAACCTCTTC

GAGAAGGCCAGCGTGTCCGAGGATAGTGATGTGCTTACCATCGGGATGAAGTTCCAGACCATGCTGACCCAGATCATCAA

GCTGCCCAACTTCATCAACGAGAACGCCTTGAACAAGATGAGCAGCTACAAGGACTTCTCCAAGCTGTACCCCAACCTCA

AGAAGAACGAGGATCTGTACAAGTCCACCAAGAACCTCAAAATCGACGAGGACGCTATCCTCGAGGGTGATGAGCTTTAT

GAGAAGATCGCCAGCAGCCTCGAGATGGAATCTGTGCACGACATCATGATCAAGAACCCGGAAACCATTCTGATCGCTCC

GCTGAACGATAGGGACTTTCTGCTGTCTCAGCTGTTCATGTACACCTCTCCGTCTAAGAGGAACCAGCTGTCTAATCAGT

CTACCGAGAAGCTGGCTCTTGATAGGGTGTTGAGATCTAAGGCTAGGACCTTCGTGGACATCTCCTCTACCGTTAAGATG

ACCTATGAAGAGAACATGGAAAAAAAGATCTTGGAGATGCTCAAGTTCGATCTCGACAGCTACTGCAGCTTCAAGACTTG

CGTGAACCTGGTGATCAAGGATGTGAACTTCTCCATGCTTATCCCGATCCTCGACTCTGCTTATCCTTGCGAGTCTAGGAA

GCGGGACAACTACAACTTCCGTTGGTTCCAGACTGAGAAGTGGATCCCTGTTGTTGAGGGTTCTCCTGGACTTGTGGTG

ATGCATGCTGTGTACGGGTCTAACTATATCGAGAACCTCGGCCTTAAGAACATCCCGCTTACCGACGACTCTATCAACGTG

CTGACTTCTACCTTCGGCACCGGTCTGATTATGGAAGATGTCAAGAGCCTGGTGAAGGGCAAAGACTCATTTGAGACAGA

GGCCTTCAGCAACTCTAACGAGTGCCAGAGACTTGTGAAGGCCTGCAACTACATGATCGCTGCTCAGAACAGGCTCCTG

GCTATCAATACCTGCTTTACCCGGAAGTCTTTCCCGTTCTACTCCAAGTTCAATCTTGGCAGGGGCTTCATCAGCAACACC

CTTGCTTTGCTGTCCACCATCTACAGCAAAGAGGAATCCTACCACTTCGTGTCCACCGCCTCTTACAAGCTGGATAAGAC

TATCCGGACCGTGATCTCTGCACAGCAGGATATGAACCTGGAAAAGATCCTGGATACCGCCGTGTACATCAGCGACAAGC

TTCAGTCTTTGTTCCCGACCATCACCAGAGAGGACATCGTTCTCATTCTCCAGAACGTGTGCCTGGACTCCAAGCCTATT

TGGCAGTCTTTGGAGGACAAGATGAAGAAGATTAACAACAGCACCGCCAGCGGCTTCACCGTGTCTAATGTGATTCTGAG

CCACAACTCCGAGCTGAACACCATCCAGAAACAGATCGTGTGGATGTGGAACATGGGCCTTTGCTCTCATAGGACCCTGG

ATTTCGTGATCCGTTACATCAGGCGGTCCGATGTGAGATACGTCAAGACTGAGGAACAGGACGAGAGCGGGAACTACATT

AGCGGAACCATGTACAAGATTGGCATCATGACCCGGTCTTGCTACGTGCAGTTGATCGCATCAGATCAGGATGTGGCTGT

GTCTCTTAGGACCCCTTTCGAGATTCTCAACGAGAGGGATTACCTGTTCGACACCTACCGTGAGTCCATCGAAAAGCTGC

TGCAAAAGTTCATGTTCGACAAGGTCAACATTATCAAGTCTAAGCAGCCGCAGATTGTGTTCCTCGAGCCAGGTGATGCT

TGCATTAGGATGACCACCGACAACAAAATGATCGTCAAGGTTAACGCCACCAGCAGGCAGATTAGGCTTGAGAACGTTAA

GCTGGTGGTGAAGATTAAGTATGAGAACGTGAACTCCGACGTGTGGGACATCATTGAGAGCCAGAAGTCTCTCGTTCTC

CGGCTTCCTGAAGTTGGTGAGTGCTTCAGCGACATGTATAAGACCGCTGACAGCGAGACTGAGACAATCAAGACTGTGA

AGAACCGGCTGATGACCTCTCTGACCTTCATTGAGGCTTTCGGCAACCTCAGCCAGCAGATCAAAGAAATCGTGGACGAT

GACATCAGGGAAACCATGGACGAGTTCCTCATGAACATCCGTGATACCTGCCTTGAGGGTCTTGAGAATTGCAAGAGCGT

GGAAGAGTACGACTCCTACCTTGATGAGAACGGGTTCAACGATACCGTCGAGCTGTTCGAGAATCTTCTGAGGACCCAC

GACAACTTCGAGAACGAGTACTCTCCACTGTTCAGCGAGATCGTGGATAAGGCTAAGCAGTACACCAGGGATCTCGAGG

GCTTTAAAGAAATCCTGCTGATGCTGAAGTACTCTCTGATCAACGACGCCTCCGGGTTCAAGTCATATAGAGCTACTGGTA

TGCACGCCGTTGAGCTGATGGCTAAGAAACACATTGAGATCGGCGAGTTCAACCTGCTGGGTATGATTCAGCTTATCAAG

GCTTGCGAGACTTGCCACAACAACGACTCCATTCTCAACCTGGCCAGCCTTAGGAATGTGCTGTCTAGGACTTACGCTAC

TTTCGGCAGGCGGATCAGGTTGAATCACGATCTGGATCTGCAGAACAACCTTATGGAAAAGAGCTACGACTTCAAGACCT

TGGTGCTCCCCGAGATTAAGCTCTCTGAGCTGAGCAGGGAAATCCTGAAAGAGAACGGCTTCGTTATCTCCGGCGAGAA

CCTTAAGATGGACAGGTCCGATGAAGAGTTCGAGGGGCTCGCTTCATTCAATGTGCTTAGGTTGGACGAGGAAGAGATG

TACGAGGGGCTGATCAAAGAAATGAAGATCAAGCGGAAGAAGAAGGGCTCTCTGTTCCCTGCTAATACCCTGCTTCTGTC

CGAGTTGATCAAGTTCCTGATCGGCGGTATTAAGGGCACCAGCTTCGATATTGAGACTCTGCTCCGGAACTCATTCAGGC

CGGATATCTTCTCTACCGACAGGCTTGGTAGGCTCTCTTCTTCTGTGCCTGCTCTTAAGGTGTACGCCACTGTGTACATGG

AGTACAAGAACGTCAACTGCCCTCTGAACGAGATCGCCGATTCTCTTGAGGGTTACCTGAAGCTCACCAAGTCCAAGAG

CAAAGAACACTTCTTGAGCGGCAGGGTGAAGAAGGCTCTTATTCAGTTGAGGGACGAGCAGTCCAGGACCAAAAAGCTC

GAGGTCTACAAGGATATCGCCAACTTCCTTAGCAGGCACCCTCTTTGCCTGTCTGAAAAGACACTGTACGGCCGGTACAC

CTACAGCGATATCAACGATTACATCATGCAGACCCGGGAAATCATCCTCTCTAAGATCTCCGAGTTGGATGAGGTGGTGG

AAACTGACGAGGACAACTTCCTGCTCTCTTACCTTAGGGGCGAAGAGGATGCTTTCGATGAGGACGATTCTGATGAGGA

AGAGGACACCGATTAG

Fig. S10. Optimized GP gene sequence used in the study.

ATGAGGATCCTGAAGCTTCTTGAGCTGGTGGTGAAGGTGAGCCTGTTTACCATTGCTCTGTCCTCTGTGCTTCTGGCCTT

CCTTATTTTCAGGGCTACCGACGCTAAGGTCGAGATCATTAGAGGCGATCATCCTGAGGTGTACGACGACTCTGCTGAGA

ATGAAGTTCCTACCGCTGCTAGCATTCAGCGGAAGGCTATTCTTGAGACTCTGACCTCTCTGATGCTCGAGTCTCAAACT

CCTGGAACCAGGCAGATTCGTGAGGAAGAGTCTACCATTCCTATCTTCGCTGGCAGCACTACCCAGAAGATTATCTCCGT

TAGCGACCTGCCTAACAACTGCCTGAACGCTTCATCTCTGAAGTGCGAGATCAAGGGCATCAGCACCTACAACGTTTACT

ACCAGGTCGAGAACAACGGCGTGATCTACTCTTGCGTGTCAGATTCTGCTGAGGGCCTTGAGAAGTGCGACAACTCTCT

TAACCTGCCGAAGCGGTTCTCTAAGGTGCCAGTGATTCCTATCACCAAGCTGGACAACAAGCGGCACTTCTCTGTGGGCA

CCAAGTTCTTCATTAGCGAGTCTCTGACCCAGGACAACTACCCGATTACCTACAACAGCTACCCTACCAACGGAACCGTG

TCTCTTCAGACCGTTAAGCTGTCTGGCGATTGCAAGATCACCAAGAGCAACTTCGCTAACCCGTATACCGTGAGCATTAC

CTCTCCAGAGAAGATCATGGGCTACCTGATCAAGAAGCCTGGCGAAAACGCTGAGCACAAGGTGATCTCATTCTCCGGCT

CTGCTTCCATTACCTTCACCGAAGAGATGCTGGATGGTGAGCACAACCTTCTCTGCGGTGATAAGTCTGCTAAGATCCCT

AAGACCAACAAGCGTGTGAGGGACTGCATCATCAAGTACAGCAAGAGCATCTACAAGCAGACCGCCTGCATCAACTTCTC

TTGGATTAGGCTGATCCTGATCGCCCTGCTGATCTACTTCCCTATTAGATGGCTGGTGAACAAGACCACCAAGCCGCTTTT

CTTGTGGTACGATCTGATCGGCCTGATTACTTACCCAATCCTGCTGCTGATTAACTGCCTGTGGAAGTACTTTCCGTTCAA

GTGCAGCAACTGCGGCAACCTGTGCATTATTACTCACGAGTGCACCAAGATCTGCATCTGCAACAAGAGCAAGGCCAGC

AAAGAACACTCTAGCGAGTGCCCGATCCTGAGCAAAGAAACCGATCACGACTACAACAAGCACAAGTGGACCAGCATGG

AATGGTTCCACCTTATCGTGAACACCAAGCTCAGCTTCAGCCTGCTTAAGTTCGTGACCGAGATTCTGATCGGTCTGGTG

ATCCTTAGCCAGATGCCTATGTCTATGGCTCAGACTACCCAGTGCCTTAGCGGTTGCTTTTATGTTCCTGGTTGCCCTGTG

CTGGTGACCTCTAAGTTTGAAAAGTGCCCTGAGAAGGACCAGTGCTACTGCAACGTGAAAGAGGACAAGATCATCGAGT

CCATCTTCGGCACCAACATCGTGATTGAGGGTCCTAACGACTGCATCGAGAACCAGAATTGTGTGGCTCACCCGAGCATC

GACAACCTGATTAAGTGTAGACTGGGCTGCGAGTACCTGGACCTGTTTAGAAACAAGCCTCTGTACAACGGCTTCAGCGA

CTACACCGGTTCTTCACTTGGTCTTACCTCTGTGGGACTGTACGAGGCTAAGAGGCTTAGGAACGGCATCATCGACTCTT

ACAACCGGACCGATAAGATCAGCGGTATGATCGCTGGTGACAGCCTGGATAAGAACGAGACTTCTATCCCCGAGAACATC

CTGCCAAGGCAGTCTCTGATTTTCGACTCTGTGGTGGATGGCAAGTACCGGTATATGGTTGAGCAGAGCCTTCTTGGAGG

TGGTGGAACTGTGTTCATGCTGAACGATAAGACCTCTGAGAAGGCCAAGAAGTTCGTCATCTACATCAAGAGCGTGGGC

ATCCACTACGAGGTGTCAGAGAAGTATACCACCGCTCCTATTCAGTCTACCCACACCGATTTCTACTCTACCTGCACCGGT

AATTGCGATACCTGCAGAAAGAACCAGGCTCTGACAGGCTTTCAGGACTTCTGCATTACCCCTACCTCTTACTGGGGTTG

TGAAGAGGCTTGGTGCTTCGCTATTAACGAGGGTGCTACTTGCGGCTTCTGCCGGAATATCTACGACATGGACAAGAGCT

ACCGGATCTACAGCGTGCTGAAGTCTACTATCGTTGCCGACGTGTGCATCTCCGGTATTCTTGGTGGACAGTGCTCTAGG

ATTACCGAAGAGGTTCCATACGAGAACGCTCTGTTCCAGGCTGATATTCAGGCTGATCTGCACAACGATGGCATCACCAT

TGGTGAGCTTATTGCTCACGGACCGGACAGCCATATCTACTCCGGTAACATTGCTAACCTGAACGACCCGGTGAAGATGT

TCGGTCATCCTCAGCTTACTCATGACGGCGTGCCAATCTTCACCAAGAAAACCCTTGAGGGCGACGACATGTCTTGGGAT

TGTGCTGCTATCGGCAAGAAGTCCATCACCATCAAGACCTGCGGTTACGACACTTACAGGTTCAGGTCTGGTCTTGAGCA

GATCTCTGACATCCCCATCAGCTTCAAGGACTTCAGCTCATTCTTCCTGGAAAAGAGCTTCAGCTTGGGGAAGCTGAAGA

TCGTGGTGGATCTGCCTAGCGATCTTTTCAAGGTTGCACCTAAGCGGCCGTCTATCACTTCTACTAGGCTTAACTGCAAC

GGCTGCCTTCTTTGTGGTCAGGGCCTTTCTTGCATCCTCGAGTTCTTCTCTGATCTGACCTTCAGCACCGCCATCTCTATC

GATGCTTGCTCTTTGTCTACCTACCAGCTGGCTGTGAAGAAGGGCAGCAACAAGTACAACATCACCATGTTCTGCAGCGC

AAACCCGGACAAGAAAAAGATGACTCTGTACCCTGAGGGCAACCCCGATATTTCTGTTGAGGTGCTCGTGAACAACGTG

ATCGTTGAGGAACCTGAGAACATCATCGATCAGAACGACGAGTACGCTCACGAGGAACAGCAGTACAACAGCGATTCATC

TGCTTGGGGATTCTGGGATTACATCAAGTCCCCGTTCAACTTCATTGCCAGCTACTTCGGCAGCTTCTTCGATACAATCAG

GGTGATCCTGCTTATCGCCTTCATCTTCCTGGTTATCTACTTCTGCAGCATCCTGACCACCATCTGCAAGGGTTACGTGAA

GAACAAGTCCTACAAGAGCCGGTCCAAGATCGAGGATGATGACGACTCAGAGATTAAGGCCCCGATGCTGATGAAGGAC

ACTATGACTAGAAGGCGGCCTCCGATGGATTTCTCTCATCTTGTTTAG

Table S1. The predicted intron splicing sites of wild-type RdRp gene.

					Activations**		
Position (bp)	**Putative splice site**	**Sequence**	**Score***	**Intron GC***	**Alt./Cryptic**	**Constitutive**	**Confidence****
69	Alt. isoform/cryptic donor	TGAGGATTGCgtgggcagca	4.835	0.529	0.954	0.035	0.963
175	Alt. isoform/cryptic donor	GAGACTATGCgtgagctgat	4.841	0.514	0.933	0.051	0.946
205	Constitutive donor	ACTGCTGATGgtgagggtct	12.212	0.500	0.370	0.546	0.322
223	Alt. isoform/cryptic donor	CTGAACAAAGgtatggctac	10.886	0.500	0.802	0.142	0.823
261	Alt. isoform/cryptic donor	GTCTGAGATGgtgtctctgt	5.341	0.486	0.882	0.086	0.903
280	Alt. isoform/cryptic acceptor	gttcgagcagAAGTACCTTG	5.730	0.514	0.943	0.055	0.942
281	Alt. isoform/cryptic donor	TCGAGCAGAAgtaccttgag	5.847	0.500	0.816	0.135	0.834
336	Constitutive acceptor	ttatttctagGCACCTGAGG	4.932	0.486	0.212	0.770	0.725
579	Alt. isoform/cryptic acceptor	tggatgccagGACTGAGACT	4.087	0.500	0.696	0.295	0.576
602	Alt. isoform/cryptic donor	CCATGGAAAAgtactacaag	6.059	0.457	0.939	0.044	0.953
642	Alt. isoform/cryptic donor	GGACATCAAGgtgaacggca	7.873	0.514	0.821	0.135	0.835
681	Alt. isoform/cryptic donor	TCCTGTGTTCgtgagcatcg	6.292	0.529	0.839	0.123	0.853
738	Alt. isoform/cryptic acceptor	tgacctctagCAGAGTGCTC	3.060	0.500	0.836	0.157	0.812
741	Alt. isoform/cryptic acceptor	cctctagcagAGTGCTCGAG	3.393	0.514	0.762	0.227	0.703
831	Alt. isoform/cryptic donor	CATTTCTCACgtgggtcaga	5.665	0.500	0.943	0.042	0.955
835	Alt. isoform/cryptic donor	TCTCACGTGGgtcagatcgt	6.026	0.500	0.936	0.048	0.949
843	Alt. isoform/cryptic donor	GGGTCAGATCgtgggtacta	5.056	0.514	0.903	0.074	0.918
864	Alt. isoform/cryptic donor	TCCTACTGTGgtgaggaact	11.320	0.500	0.901	0.072	0.920
906	Alt. isoform/cryptic donor	CAAGTCCGAGgttaggggta	10.127	0.514	0.709	0.226	0.682
913	Alt. isoform/cryptic donor	GAGGTTAGGGgtatcctggg	4.945	0.514	0.953	0.032	0.967
971	Alt. isoform/cryptic donor	GGACCTCCAAgtacaaagag	5.665	0.514	0.934	0.049	0.948
1185	Alt. isoform/cryptic donor	CGAGCCTAAGgtggacctgg	4.643	0.514	0.716	0.215	0.700
1207	Alt. isoform/cryptic acceptor	tgatcttaagGATCACCTGA	2.743	0.500	0.596	0.388	0.349
1249	Alt. isoform/cryptic donor	AGCCTGTACGgtaagcacct	11.990	0.457	0.579	0.329	0.431
1361	Alt. isoform/cryptic donor	ACGAAGAGAAgtacccgaac	5.393	0.529	0.904	0.069	0.923
1388	Alt. isoform/cryptic donor	CCAACGACCTgtcagagaca	5.572	0.500	0.910	0.068	0.925
1421	Alt. isoform/cryptic donor	CTATGACCTGgtcaccgagc	5.097	0.471	0.952	0.036	0.962
1492	Alt. isoform/cryptic donor	AACAAGTTCCgtgagagctt	6.085	0.500	0.565	0.376	0.334
1515	Alt. isoform/cryptic acceptor	agtccagcagCAGGGTGATC	2.527	0.486	0.940	0.057	0.939
1542	Alt. isoform/cryptic acceptor	gcccttacagCTGCATTAAC	3.273	0.486	0.526	0.454	0.136

1587	Alt. isoform/cryptic donor	CACTAACCTTgtgaggcttt	7.336	0.514	0.944	0.040	0.958
1672	Constitutive donor	ATCAACACCGgtagcattaa	5.999	0.500	0.395	0.516	0.233
1740	Constitutive acceptor	gccttactagGAACAAGAAC	7.612	0.529	0.146	0.841	0.827
1801	Alt. isoform/cryptic acceptor	ctacttcaagGGTCTTGCCG	3.996	0.471	0.717	0.270	0.624
1889	Alt. isoform/cryptic donor	AGGGCCTCGAgtacgatact	5.278	0.500	0.956	0.032	0.966
2006	Alt. isoform/cryptic donor	CACTTTCTTGgtgcgagcga	6.544	0.529	0.958	0.031	0.968
2125	Alt. isoform/cryptic acceptor	gtacatccagGATTCTGTGC	6.304	0.514	0.555	0.432	0.222
2205	Alt. isoform/cryptic acceptor	accactacagCGAGGACATG	4.878	0.457	0.847	0.147	0.827
2230	Alt. isoform/cryptic acceptor	gttctccaagGGTTTGATGG	2.346	0.471	0.692	0.290	0.580
2321	Constitutive acceptor	gctttcaaagGCGACGGTAT	6.287	0.457	0.385	0.587	0.344
2326	Alt. isoform/cryptic donor	AAAGGCGACGgtatgaacac	7.959	0.486	0.760	0.181	0.761
2487	Constitutive donor	GCTTAACCAGgtgaggcttc	14.109	0.500	0.246	0.676	0.637
2512	Alt. isoform/cryptic acceptor	actgttcaagACCCCTTCTA	4.626	0.514	0.933	0.062	0.933
2523	Alt. isoform/cryptic donor	CCCTTCTAAGgtgccagttt	10.939	0.471	0.614	0.306	0.501
2545	Alt. isoform/cryptic acceptor	cttcgcccagTTCAGCAAAA	6.490	0.529	0.692	0.300	0.566
2550	Alt. isoform/cryptic acceptor	cccagttcagCAAAAAGGCC	4.578	0.500	0.800	0.194	0.758
2601	Alt. isoform/cryptic donor	CATTGAGAAGgtcaacgtgt	6.294	0.500	0.944	0.041	0.956
2664	Constitutive acceptor	tgttctccagCGTGATGATT	6.507	0.486	0.462	0.524	0.117
2720	Alt. isoform/cryptic donor	ACTTCATGAGgtacgctggt	9.451	0.543	0.706	0.226	0.680
2765	Alt. isoform/cryptic donor	ACATCAAAGAgtacatccgg	4.583	0.500	0.944	0.041	0.956
2942	Constitutive donor	TGAACATCAAgtgccctatt	6.540	0.500	0.357	0.573	0.378
3217	Alt. isoform/cryptic donor	TACTACCTGGgtaacatcga	6.561	0.500	0.885	0.083	0.906
3292	Alt. isoform/cryptic acceptor	gtgctacaagATCTCTACCC	3.251	0.514	0.960	0.038	0.961
3307	Alt. isoform/cryptic acceptor	taccctgaagTCCAGTAAGA	3.209	0.500	0.930	0.067	0.928
3310	Alt. isoform/cryptic donor	CTGAAGTCCAgtaagaagtg	9.173	0.471	0.712	0.218	0.694
3370	Constitutive acceptor	gtgcctgcagAACGCAAAGA	3.539	0.500	0.477	0.500	0.047
3557	Alt. isoform/cryptic donor	CCGTGGGAAAgtctgtgact	6.027	0.500	0.919	0.059	0.935
3642	Alt. isoform/cryptic donor	TCGGAACAACgtgaccgtgc	5.533	0.500	0.919	0.061	0.934
3666	unclassified donor	GTCTAAGAAGgtgagcgagg	13.918	0.500	0.418	0.490	0.000
3745	Alt. isoform/cryptic donor	ATGAACCTTGgtaagggtac	11.753	0.500	0.609	0.303	0.502
3751	Alt. isoform/cryptic donor	CTTGGTAAGGgtactgaggg	6.065	0.500	0.952	0.033	0.966
3784	Constitutive acceptor	cttcttgcagATGCTTGAGT	7.888	0.514	0.201	0.782	0.742
3816	Alt. isoform/cryptic donor	GGCCAAGAATgtgaccggtt	6.207	0.543	0.936	0.047	0.950

3840	Alt. isoform/cryptic donor	GGATTTCCTCgtgagcgtgt	5.915	0.514	0.899	0.079	0.912
3880	Alt. isoform/cryptic donor	AAGACCGACCgtgagatcta	5.588	0.457	0.542	0.380	0.299
3906	Alt. isoform/cryptic donor	GAGCATGAAGgtgaagatga	8.763	0.500	0.560	0.358	0.361
4023	Alt. isoform/cryptic acceptor	ctacccttagCCTGGACACC	6.193	0.514	0.626	0.365	0.416
4136	Alt. isoform/cryptic donor	TGACCTACAAgtacgtgctg	9.678	0.500	0.596	0.323	0.459
4211	Alt. isoform/cryptic donor	GCATTCTGATgtacgtgaag	7.123	0.514	0.847	0.112	0.868
4230	Alt. isoform/cryptic donor	GCTGAAGAAAgtgtgcatcc	6.203	0.529	0.934	0.048	0.949
4266	Alt. isoform/cryptic acceptor	tgaaccttagGAAGGCTCAG	7.185	0.457	0.849	0.138	0.837
4306	Alt. isoform/cryptic donor	ACCGCTATTGgtctgctgac	5.056	0.529	0.865	0.097	0.888
4344	Alt. isoform/cryptic donor	CACTTACCCGgtgtctatga	6.061	0.514	0.900	0.077	0.914
4366	Alt. isoform/cryptic donor	TGGCTGCAGGgtaacctgaa	6.041	0.500	0.750	0.183	0.755
4414	Alt. isoform/cryptic acceptor	cgctatgaagGCTTACCACA	3.152	0.514	0.945	0.052	0.945
4459	Alt. isoform/cryptic acceptor	cgactttcagACCCGGTGGA	3.313	0.500	0.787	0.200	0.746
4510	Constitutive donor	ATCGCTTCAGgtgaggttga	11.745	0.500	0.262	0.665	0.607
4511	Constitutive acceptor	atcgcttcagGTGAGGTTGA	6.293	0.529	0.441	0.541	0.185
4701	Constitutive acceptor	tttactgcagGCATCTTGCT	8.159	0.514	0.411	0.576	0.288
4773	Alt. isoform/cryptic donor	GTCTATCCACgtgaccatgt	6.070	0.529	0.520	0.403	0.224
4809	Alt. isoform/cryptic donor	CCCTAATGAGgtgatccctt	6.780	0.529	0.856	0.106	0.877
4873	Alt. isoform/cryptic donor	ATGCTGCCTGgtgaggtgaa	9.597	0.529	0.762	0.188	0.753
4878	Alt. isoform/cryptic donor	GCCTGGTGAGgtgaacgaca	5.312	0.500	0.940	0.044	0.953
5001	Alt. isoform/cryptic acceptor	tgggccccagCAGCAACGAC	4.531	0.529	0.765	0.227	0.703
5004	Alt. isoform/cryptic acceptor	gccccagcagCAACGACCAG	3.016	0.529	0.837	0.157	0.813
5067	Alt. isoform/cryptic donor	CCTGGAAGAGgtgaaggact	8.199	0.514	0.764	0.180	0.764
5104	Alt. isoform/cryptic acceptor	ttacctgcagATGAGGTTCA	11.438	0.514	0.591	0.398	0.325
5140	Alt. isoform/cryptic donor	TACGAGAAGGgtactctgga	5.097	0.471	0.916	0.059	0.936
5263	Alt. isoform/cryptic acceptor	gctgacccagATCATCAAGC	2.350	0.557	0.593	0.388	0.345
5416	Alt. isoform/cryptic donor	ATCCTCGAGGgtgatgagct	5.258	0.471	0.829	0.126	0.848
5542	Alt. isoform/cryptic acceptor	gctgtctcagCTGTTCATGT	8.742	0.500	0.683	0.304	0.555
5549	Alt. isoform/cryptic donor	AGCTGTTCATgtacacctct	5.135	0.514	0.868	0.098	0.887
5569	Alt. isoform/cryptic acceptor	tccgtctaagAGGAACCAGC	3.653	0.514	0.927	0.070	0.924
5616	Alt. isoform/cryptic acceptor	ctcttgatagGGTGTTGAGA	3.796	0.486	0.802	0.185	0.769
5616	Alt. isoform/cryptic donor	TCTTGATAGGgtgttgagat	4.789	0.457	0.911	0.062	0.931
5668	Alt. isoform/cryptic acceptor	taccgttaagATGACCTATG	4.339	0.500	0.796	0.196	0.754

5730	Alt. isoform/cryptic acceptor	atctcgacagCTACTGCAGC	3.975	0.414	0.845	0.148	0.825
5829	Alt. isoform/cryptic acceptor	gcgagtctagGAAGCGGGAC	4.419	0.486	0.590	0.389	0.340
5863	Alt. isoform/cryptic acceptor	ttggttccagACTGAGAAGT	7.317	0.543	0.600	0.390	0.350
5921	Alt. isoform/cryptic donor	TGCATGCTGTgtacgggtct	4.883	0.500	0.957	0.031	0.968
6139	Alt. isoform/cryptic acceptor	cgctgctcagAACAGGCTCC	2.442	0.529	0.816	0.176	0.784
6199	Alt. isoform/cryptic acceptor	ctactccaagTTCAATCTTG	4.578	0.529	0.930	0.066	0.929
6213	Alt. isoform/cryptic acceptor	atcttggcagGGGCTTCATC	4.671	0.500	0.678	0.310	0.543
6258	Alt. isoform/cryptic acceptor	ccatctacagCAAAGAGGAA	4.973	0.500	0.822	0.171	0.793
6337	Alt. isoform/cryptic acceptor	ctctgcacagCAGGATATGA	3.577	0.529	0.610	0.372	0.390
6340	Alt. isoform/cryptic acceptor	tgcacagcagGATATGAACC	5.307	0.543	0.575	0.395	0.313
6420	Alt. isoform/cryptic acceptor	ccatcaccagAGAGGACATC	2.479	0.514	0.962	0.036	0.963
6445	Constitutive acceptor	cattctccagAACGTGTGCC	9.539	0.500	0.209	0.784	0.733
6624	Constitutive acceptor	gctctcatagGACCCTGGAT	10.048	0.500	0.176	0.817	0.784
6654	Alt. isoform/cryptic acceptor	gttacatcagGCGGTCCGAT	3.376	0.500	0.765	0.221	0.712
6663	Alt. isoform/cryptic donor	GCGGTCCGATgtgagatacg	6.227	0.486	0.936	0.047	0.950
6722	Alt. isoform/cryptic donor	GCGGAACCATgtacaagatt	5.221	0.514	0.960	0.028	0.971
6763	Alt. isoform/cryptic acceptor	ctacgtgcagTTGATCGCAT	5.764	0.529	0.844	0.148	0.824
6801	Constitutive acceptor	tgtctcttagGACCCCTTTC	5.629	0.514	0.255	0.731	0.652
6850	Alt. isoform/cryptic donor	GACACCTACCgtgagtccat	8.138	0.443	0.627	0.307	0.511
6894	Alt. isoform/cryptic donor	GTTCGACAAGgtcaacatta	6.897	0.471	0.855	0.108	0.874
6895	Alt. isoform/cryptic acceptor	gttcgacaagGTCAACATTA	2.380	0.486	0.870	0.126	0.855
6943	Alt. isoform/cryptic donor	CTCGAGCCAGgtgatgcttg	8.185	0.486	0.833	0.125	0.850
6944	Alt. isoform/cryptic acceptor	ctcgagccagGTGATGCTTG	5.597	0.457	0.761	0.232	0.696
6960	Constitutive acceptor	cttgcattagGATGACCACC	2.622	0.471	0.452	0.510	0.115
6990	unclassified donor	GATCGTCAAGgttaacgcca	6.349	0.471	0.500	0.414	0.000
7038	Alt. isoform/cryptic donor	TAAGCTGGTGgtgaagatta	5.500	0.471	0.942	0.041	0.957
7071	Alt. isoform/cryptic donor	GAACTCCGACgtgtgggaca	4.732	0.543	0.918	0.059	0.936
7124	Alt. isoform/cryptic acceptor	cttcctgaagTTGGTGAGTG	3.877	0.543	0.743	0.247	0.667
7126	Alt. isoform/cryptic donor	CCTGAAGTTGgtgagtgctt	12.153	0.500	0.727	0.210	0.711
7242	Alt. isoform/cryptic acceptor	gcaacctcagCCAGCAGATC	3.540	0.500	0.601	0.379	0.370
7355	Alt. isoform/cryptic donor	GCGTGGAAGAgtacgactcc	4.790	0.514	0.943	0.040	0.957
7464	Alt. isoform/cryptic acceptor	cactgttcagCGAGATCGTG	6.043	0.529	0.637	0.349	0.453
7535	Alt. isoform/cryptic donor	TGATGCTGAAgtactctctg	4.763	0.529	0.907	0.069	0.924

7585	Alt. isoform/cryptic donor	AGAGCTACTGgtatgcacgc	9.638	0.500	0.752	0.188	0.750
7648	Alt. isoform/cryptic donor	AACCTGCTGGgtatgattca	7.568	0.500	0.698	0.229	0.672
7722	Alt. isoform/cryptic acceptor	ccagccttagGAATGTGCTG	3.055	0.500	0.807	0.185	0.771
7758	Alt. isoform/cryptic acceptor	ctttcggcagGCGGATCAGG	6.039	0.514	0.607	0.379	0.375
7792	Alt. isoform/cryptic acceptor	ggatctgcagAACAACCTTA	3.876	0.514	0.696	0.289	0.586
7922	Alt. isoform/cryptic donor	AGATGGACAGgtccgatgaa	6.684	0.514	0.941	0.042	0.955
7968	Constitutive acceptor	atgtgcttagGTTGGACGAG	2.272	0.529	0.343	0.637	0.461
8169	Alt. isoform/cryptic acceptor	ctaccgacagGCTTGGTAGG	8.327	0.529	0.960	0.038	0.961
8173	Alt. isoform/cryptic donor	GACAGGCTTGgtaggctctc	8.257	0.486	0.774	0.169	0.782
8205	Alt. isoform/cryptic donor	TGCTCTTAAGgtgtacgcca	6.332	0.514	0.932	0.049	0.948
8206	Alt. isoform/cryptic acceptor	tgctcttaagGTGTACGCCA	5.395	0.514	0.524	0.455	0.133
8331	Alt. isoform/cryptic acceptor	tgagcggcagGGTGAAGAAG	5.008	0.514	0.856	0.139	0.838
8353	Alt. isoform/cryptic acceptor	tcttattcagTTGAGGGACG	2.472	0.500	0.856	0.133	0.845
8418	Alt. isoform/cryptic acceptor	acttccttagCAGGCACCCT	2.679	0.500	0.944	0.052	0.945
8421	Alt. isoform/cryptic acceptor	tccttagcagGCACCCTCTT	6.406	0.514	0.782	0.208	0.734
8459	Alt. isoform/cryptic donor	TGTACGGCCGgtacacctac	4.997	0.471	0.903	0.074	0.918
8472	Alt. isoform/cryptic acceptor	acacctacagCGATATCAAC	3.154	0.543	0.614	0.363	0.409
8497	Alt. isoform/cryptic acceptor	catcatgcagACCCGGGAAA	2.271	0.471	0.944	0.053	0.944
8521	Constitutive acceptor	cctctctaagATCTCCGAGT	5.017	0.500	0.424	0.562	0.245
8538	Alt. isoform/cryptic donor	GTTGGATGAGgtggtggaaa	4.678	0.529	0.945	0.039	0.958
8583	Constitutive acceptor	cttaccttagGGGCGAAGAG	8.039	0.486	0.196	0.792	0.753

The putative intron splicing sites of wild-type RdRp gene sequence was predicted by Alternative Splice Site Predictor (ASSP) (http://wangcomputing.com/assp/).

* Scores of the preprocessing models reflecting splice site strength, i.e. a PSSM for putative acceptor sites, and an MDD model for putative donor sites. Intron GC values correspond to 70 nt of the neighboring intron.

** Activations are output values of the backpropagation networks used for classification. High values for one class with low values of the other class imply a good classification. Confidence is a simple measure expressing the differences between output activations. Confidence ranges between zero (undecided) to one (perfect classification).

Table S2. The predicted intron splicing sites of wild-type GP gene.

					Activations**		
Position (bp)	**Putative splice site**	**Sequence**	**Score***	**Intron GC***	**Alt./Cryptic**	**Constitutive**	**Confidence****
36	unclassified donor	GGTGGTGAAGgtgagcctgt	13.078	0.514	0.408	0.496	0.000
93	Constitutive acceptor	ttattttcagGGCTACCGAC	11.782	0.500	0.182	0.810	0.775
183	Alt. isoform/cryptic acceptor	ccgctgctagCATTCAGCGG	3.056	0.500	0.929	0.067	0.928
259	Alt. isoform/cryptic donor	AGGCAGATTCgtgaggaaga	4.761	0.500	0.860	0.108	0.875
297	Alt. isoform/cryptic acceptor	tcgctggcagCACTACCCAG	5.498	0.514	0.952	0.046	0.952
318	Alt. isoform/cryptic donor	GATTATCTCCgttagcgacc	4.792	0.529	0.928	0.056	0.940
324	Alt. isoform/cryptic acceptor	tctccgttagCGACCTGCCT	2.244	0.471	0.679	0.308	0.547
364	Alt. isoform/cryptic acceptor	atctctgaagTGCGAGATCA	3.458	0.500	0.845	0.145	0.828
405	Alt. isoform/cryptic donor	TTACTACCAGgtcgagaaca	5.199	0.529	0.954	0.034	0.965
406	Alt. isoform/cryptic acceptor	ttactaccagGTCGAGAACA	5.238	0.500	0.592	0.383	0.354
437	Alt. isoform/cryptic donor	ACTCTTGCGTgtcagattct	4.710	0.529	0.923	0.057	0.938
443	Alt. isoform/cryptic acceptor	tgcgtgtcagATTCTGCTGA	5.120	0.514	0.630	0.357	0.433
504	Alt. isoform/cryptic donor	GTTCTCTAAGgtgccagtga	9.430	0.500	0.831	0.124	0.851
529	Alt. isoform/cryptic acceptor	tatcaccaagCTGGACAACA	2.328	0.500	0.853	0.139	0.837
552	Alt. isoform/cryptic donor	GCACTTCTCTgtgggcacca	6.302	0.514	0.893	0.079	0.912
592	Constitutive acceptor	tctgacccagGACAACTACC	3.778	0.543	0.341	0.630	0.460
649	Alt. isoform/cryptic acceptor	gtctcttcagACCGTTAAGC	6.880	0.529	0.693	0.295	0.574
708	Alt. isoform/cryptic donor	CCCGTATACCgtgagcatta	8.383	0.514	0.655	0.278	0.576
728	Constitutive acceptor	acctctccagAGAAGATCAT	6.756	0.514	0.388	0.598	0.352
780	Alt. isoform/cryptic donor	TGAGCACAAGgtgatctcat	9.170	0.514	0.526	0.387	0.264
824	Alt. isoform/cryptic acceptor	ttcaccgaagAGATGCTGGA	3.783	0.529	0.805	0.185	0.770
835	Alt. isoform/cryptic donor	ATGCTGGATGgtgagcacaa	9.886	0.514	0.525	0.389	0.260
894	Alt. isoform/cryptic donor	CAACAAGCGTgtgagggact	8.781	0.500	0.825	0.132	0.840
914	Alt. isoform/cryptic donor	GCATCATCAAgtacagcaag	6.287	0.514	0.880	0.089	0.899
969	Constitutive acceptor	cttggattagGCTGATCCTG	5.033	0.486	0.190	0.798	0.761
1008	Constitutive acceptor	tccctattagATGGCTGGTG	7.158	0.500	0.176	0.808	0.782
1046	Alt. isoform/cryptic donor	TTTTCTTGTGgtacgatctg	8.598	0.486	0.789	0.158	0.800
1106	Alt. isoform/cryptic donor	GCCTGTGGAAgtactttccg	5.205	0.500	0.934	0.049	0.948
1123	Alt. isoform/cryptic acceptor	tccgttcaagTGCAGCAACT	4.432	0.471	0.799	0.192	0.760
1128	Alt. isoform/cryptic acceptor	tcaagtgcagCAACTGCGGC	3.238	0.486	0.858	0.136	0.842

1328	Alt. isoform/cryptic donor	GCCTGCTTAAgttcgtgacc	5.154	0.514	0.929	0.051	0.945
1356	Alt. isoform/cryptic donor	GATCGGTCTGgtgatcctta	4.853	0.486	0.642	0.284	0.558
1368	Alt. isoform/cryptic acceptor	tgatccttagCCAGATGCCT	2.891	0.500	0.795	0.198	0.751
1372	Alt. isoform/cryptic acceptor	ccttagccagATGCCTATGT	3.041	0.529	0.730	0.259	0.646
1443	Alt. isoform/cryptic donor	CCCTGTGCTGgtgacctcta	6.777	0.471	0.866	0.100	0.884
1463	Alt. isoform/cryptic donor	AGTTTGAAAAgtgccctgag	5.015	0.543	0.964	0.026	0.973
1631	Alt. isoform/cryptic donor	TGGGCTGCGAgtacctggac	4.582	0.500	0.967	0.024	0.975
1783	Alt. isoform/cryptic donor	AAGATCAGCGgtatgatcgc	7.260	0.529	0.671	0.262	0.610
1795	Alt. isoform/cryptic donor	ATGATCGCTGgtgacagcct	7.023	0.500	0.803	0.149	0.814
1883	Alt. isoform/cryptic donor	TGGATGGCAAgtaccggtat	5.025	0.500	0.926	0.053	0.943
1889	Alt. isoform/cryptic donor	GCAAGTACCGgtatatggtt	5.618	0.486	0.814	0.135	0.834
1992	Alt. isoform/cryptic donor	CATCAAGAGCgtgggcatcc	5.600	0.514	0.825	0.138	0.833
2010	Alt. isoform/cryptic donor	CCACTACGAGgtgtcagaga	7.706	0.500	0.765	0.178	0.767
2044	Alt. isoform/cryptic acceptor	tcctattcagTCTACCCACA	4.199	0.514	0.801	0.186	0.768
2077	Constitutive donor	ACCTGCACCGgtaattgcga	7.342	0.500	0.414	0.502	0.176
2097	Alt. isoform/cryptic acceptor	atacctgcagAAAGAACCAG	4.833	0.529	0.869	0.122	0.860
2125	Constitutive acceptor	aggctttcagGACTTCTGCA	3.662	0.500	0.451	0.529	0.148
2191	Alt. isoform/cryptic donor	ATTAACGAGGgtgctacttg	4.511	0.529	0.955	0.033	0.966
2253	Alt. isoform/cryptic acceptor	ggatctacagCGTGCTGAAG	3.191	0.514	0.910	0.086	0.906
2280	Alt. isoform/cryptic donor	CGTTGCCGACgtgtgcatct	6.880	0.514	0.904	0.073	0.919
2331	Alt. isoform/cryptic donor	TACCGAAGAGgttccatacg	4.991	0.500	0.959	0.029	0.970
2359	Constitutive acceptor	tctgttccagGCTGATATTC	5.600	0.514	0.373	0.611	0.389
2401	Alt. isoform/cryptic donor	ATCACCATTGgtgagcttat	11.534	0.529	0.522	0.395	0.242
2443	Alt. isoform/cryptic donor	ATCTACTCCGgtaacattgc	9.301	0.529	0.591	0.337	0.430
2469	Alt. isoform/cryptic donor	GAACGACCCGgtgaagatgt	5.862	0.500	0.788	0.167	0.788
2494	Alt. isoform/cryptic acceptor	tcatcctcagCTTACTCATG	6.466	0.514	0.705	0.287	0.594
2618	Alt. isoform/cryptic donor	ACACTTACAGgttcaggtct	7.293	0.500	0.904	0.068	0.925
2619	Alt. isoform/cryptic acceptor	acacttacagGTTCAGGTCT	5.949	0.500	0.773	0.215	0.722
2624	Alt. isoform/cryptic donor	ACAGGTTCAGgtctggtctt	8.772	0.486	0.925	0.054	0.942
2625	Alt. isoform/cryptic acceptor	acaggttcagGTCTGGTCTT	3.359	0.514	0.684	0.296	0.567
2641	Alt. isoform/cryptic acceptor	tcttgagcagATCTCTGACA	4.449	0.529	0.945	0.051	0.946
2661	Alt. isoform/cryptic acceptor	tccccatcagCTTCAAGGAC	3.473	0.514	0.769	0.222	0.711
2668	Alt. isoform/cryptic acceptor	cagcttcaagGACTTCAGCT	3.339	0.514	0.785	0.206	0.737

2676	Alt. isoform/cryptic acceptor	aggacttcagCTCATTCTTC	2.357	0.500	0.795	0.197	0.753
2751	Alt. isoform/cryptic donor	TCTTTTCAAGgttgcaccta	5.168	0.529	0.935	0.047	0.950
2752	Alt. isoform/cryptic acceptor	tcttttcaagGTTGCACCTA	3.558	0.500	0.699	0.281	0.598
2787	Constitutive acceptor	cttctactagGCTTAACTGC	5.710	0.500	0.455	0.528	0.138
2815	Alt. isoform/cryptic donor	CTTCTTTGTGgtcagggcct	9.649	0.529	0.757	0.184	0.757
2821	Alt. isoform/cryptic acceptor	ttgtggtcagGGCCTTTCTT	6.984	0.529	0.502	0.483	0.039
2865	Alt. isoform/cryptic acceptor	tgaccttcagCACCGCCATC	4.796	0.529	0.876	0.117	0.867
2908	Alt. isoform/cryptic acceptor	tacctaccagCTGGCTGTGA	6.831	0.500	0.696	0.293	0.579
2933	Alt. isoform/cryptic donor	GCAGCAACAAgtacaacatc	6.089	0.500	0.851	0.108	0.873
2958	Alt. isoform/cryptic acceptor	tgttctgcagCGCAAACCCG	4.988	0.500	0.635	0.350	0.448
3021	Alt. isoform/cryptic donor	TTCTGTTGAGgtgctcgtga	5.200	0.514	0.963	0.028	0.971
3146	Alt. isoform/cryptic donor	ATTACATCAAgtccccgttc	5.088	0.500	0.945	0.041	0.956
3171	Alt. isoform/cryptic acceptor	tcattgccagCTACTTCGGC	3.870	0.471	0.816	0.173	0.788
3183	Alt. isoform/cryptic acceptor	acttcggcagCTTCTTCGAT	3.003	0.486	0.868	0.128	0.853
3201	Alt. isoform/cryptic acceptor	atacaatcagGGTGATCCTG	2.559	0.471	0.787	0.200	0.745
3201	Alt. isoform/cryptic donor	TACAATCAGGgtgatcctgc	5.357	0.500	0.829	0.130	0.843
3249	Alt. isoform/cryptic acceptor	acttctgcagCATCCTGACC	7.746	0.471	0.516	0.472	0.086

The putative intron splicing sites of wild type GP gene sequence was predicted by ASSP).

* Scores of the preprocessing models reflecting splice site strength, i.e. a PSSM for putative acceptor sites, and an MDD model for putative donor sites. Intron GC values correspond to 70 nt of the neighboring intron.

** Activations are output values of the backpropagation networks used for classification. High values for one class with low values of the other class imply a good classification. Confidence is a simple measure expressing the differences between output activations. Confidence ranges between zero (undecided) to one (perfect classification).

Table S3. List of primers used in the study.

Construct	Abbreviation	Primer sequence (5' to 3')	Purpose
p2300-N	N	**F:** GGGGTACCATGTCTAAGGTTAAGCTCA **R:** ACGTCGACTTAAGCAAGTTCTGCAAGTTTTG	To amplify TSWV N and cloned into p2300S
p2300-RdRp^{wt}	RdRp^{wt}	**F:** CGGGATCCATGAACATCCAGAAAATAC **R:** GACGTCGACTTAATCCGTGTCTTCTTCTTC	To amplify TSWV wildtype RdRp and cloned into p2300S
p2300-RdRp^{opt}	RdRp^{opt}	**F:** CTCGGTACCATGAACATTCAGAAGATCCAAAAGC **R:** GACTCTAGACTAATCGGTGTCCTCTTCCTC	To amplify TSWV optimized RdRp and cloned into p2300S
pCXSN-NSs	NSs	**F:** CTCGGTACCATGTCTTCAAGTGTTTATGAG **R:** GACTCTAGATTATTTTGATCCTGAAGATATG	To amplify TSWV NSs and cloned into pCXSN
pCB301-HH-S$_{(-)}$-RZ-NOS	S$_{(-)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAGAGCAATTGTGTCATAATTTTATTC **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATTGTGTCAATTTTATTCAAAC	To amplify the TSWV genomic RNA sequence for construction of S$_{(-)}$
		F: GTTTGAATAAAATTGACACAATTGCTCTGGGTCGGCATGGCATCTCCACC **R:** GAATAAAATTATGACACAATTGCTCTGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of S$_{(-)}$
pCB301-HH-S$_{(+)}$-RZ-NOS	S$_{(+)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAGAGCAATTGTGTCAATTTTATTCAAAC **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATTGTGTCATAATTTTATTCTTA	To amplify the TSWV antigenomic RNA sequence for construction of S$_{(+)}$
		F: GAATAAAATTATGACACAATTGCTCTGGGTCGGCATGGCATCTC **R:** GTTTGAATAAAATTGACACAATTGCTCTGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of S$_{(+)}$
pCB301-HH-S$_{(-)eGFP}$-RZ-NOS	SR$_{(-)eGFP}$	**F:** GCTTTTTTATAATTTAACTTACAACTGCTTTTACTTGTACAGCTCGTCCATGCCGAGA **R:** GTCAAAGCATATAACAACTTCTACGATCATCATGGTGAGCAAGGGCGAGGAGCTGTTC	To amplify the eGFP for construction of SR$_{(-)eGFP}$
		F: GAACAGCTCCTCGCCCTTGCTCACC	To amplify the pCB301 backbone for

		ATGATGATCGTAGAAGTTGTTATATGCT TTGAC **R:** TCTCGGCATGGACGAGCTGTACAAG TAAAAGCAGTTGTAAGTTAAATTATAAA AAAGC	construction of $SR_{(-)eGFP}$
pCB301-HH-$S_{(-)mCherry\&eGFP}$-RZ-NOS	$SR_{(-)mCherry\&eGFP}$	**F:** CACAGTACCAATAACCATAATGGTGA GCAAGGGCGAGGAGGATAAC **R:** GAAAAGCTGGACACGGCAAGATTA AGATCTGTACAGCTCGTCCATGCCGC	To amplify the mCherry for construction of $SR_{(-)mCherry\&eGFP}$
		F: GCGGCATGGACGAGCTGTACAGATC TTAATCTTGCCGTGTCCAGCTTTTC **R:** GTTATCCTCCTCGCCCTTGCTCACCA TTATGGTTATTGGTACTGTG	To amplify the pCB301 backbone for construction of $SR_{(-)mCherry\&eGFP}$
pCB301-T7-$S_{(-)mCherry\&eGFP}$-RZ-NOS	T7:$SR_{(-)mCherry\&eGFP}$	**F:** GAAATTAATACGACTCACTATAGGAG AGCAATTGTGTCAATTTTATTCAAAC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATTGTGTCATAATTTTATTCTTA	To amplify the eGFP and mCherry expression cassette for construction of T7:$SR_{(-)mCherry\&eGFP}$
		F: GTTTGAATAAAATTGACACAATTGCT CTCCTATAGTGAGTCGTATTAATTTC **R:** GAATAAAATTATGACACAATTGCTCT GGGTCGGCATGGCATCTC	To amplify the pCB301 backbone for construction of T7:$SR_{(-)mCherry\&eGFP}$
pCB301-HH-$S_{(+)eGFP}$-RZ-NOS	$SR_{(+)eGFP}$	**F:** GGAAAAGCTGGACACGGCAAGATTA CTTGTACAGCTCGTCCATGCCGAG **R:** GAACACAGTACCAATAACCATAATG GTGAGCAAGGGCGAGGAGCTGTTC	To amplify the eGFP for construction of $SR_{(+)eGFP}$
		F: GAACAGCTCCTCGCCCTTGCTCACCA TTATGGTTATTGGTACTGTGTTC **R:** CTCGGCATGGACGAGCTGTACAAGT AATCTTGCCGTGTCCAGCTTTTCC	To amplify the pCB301 backbone for construction of $SR_{(+)eGFP}$
pCB301-HH-$S_{(-)mCherry\&eGFP}^{\Delta 5'UTR}$-RZ-NOS	$SR_{(-)mCherry\&eGFP}^{\Delta 5'UTR}$	**F:** CGAAAACCCGGTATCCCGGGTTCAT GGTGAGCAAGGGCGAGGAGGATAAC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATTGTGTCAATTTTATTCAAAC	To amplify the Δ5'UTR expression cassette for construction of $SR_{(-)mCherry\&eGFP}^{\Delta 5'UTR}$
		F: GTTTGAATAAAATTGACACAATTGCT CTGGGTCGGCATGGCATCTCCACC **R:** GTTATCCTCCTCGCCCTTGCTCACCA TGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $SR_{(-)mCherry\&eGFP}^{\Delta 5'UTR}$
pCB301-HH-$S_{(-)mCherry\&eGFP}^{\Delta IGR}$-RZ-NOS	$SR_{(-)mCherry\&eGFP}^{\Delta IGR}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATTGTGTCATAATTTTATTC **R:** GCATGGACGAGCTGTACAAGTAATT AAGATCTGTACAGCTCGTC	To amplify the ΔIGR expression cassette for construction of $SR_{(-)mCherry\&eGFP}^{\Delta IGR}$
		F: GACGAGCTGTACAGATCTTAATTAC TTGTACAGCTCGTCCATGCCGAGA	To amplify the pCB301 backbone for construction of $SR_{(-)mCherry\&eGFP}^{\Delta IGR}$

		R: GAATAAAATTATGACACAATTGCTCT GAACCCGGGATACCGGGTTTTCG	
pCB301-HH-$S_{(-)mCherry\&eGFP}^{\Delta 5'UTR}$-RZ-NOS	$SR_{(-)mCherry\&eGFP}^{\Delta 3'UTR}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATTGTGTCATAATTTTATTC **R:** GGTGGAGATGCCATGCCGACCCATG GTGAGCAAGGGCGAGGAGCTGTTC	To amplify the Δ3'UTR expression cassette for construction of $SR_{(-)mCherry\&eGFP}^{\Delta 3'UTR}$
		F: GAACAGCTCCTCGCCCTTGCTCACC ATGGGTCGGCATGGCATCTCCACC **R:** GAATAAAATTATGACACAATTGCTCT GAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $SR_{(-)mCherry\&eGFP}^{\Delta 3'UTR}$
pCB301-HH-$M_{(-)}$-RZ-NOS	$M_{(-)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATCAGTGCATCAGAAATATACC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATCAGTGCAAACAAAAAC	To amplify the TSWV genomic M-RNA sequence for construction of $M_{(-)}$
		F: GTTTTTGTTTGCACTGATTGCTCTGG GTCGGCATGGCATCTCCACC **R:** GGTATATTTCTGATGCACTGATTGCT CTGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $M_{(-)}$
pCB301-HH-$M_{(+)}$-RZ-NOS	$M_{(+)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATCAGTGCAAACAAAAACTC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATCAGTGCGTCAGAAATATAC	To amplify the TSWV antigenomic M-RNA sequence for construction of $M_{(+)}$
		F: GTATATTTCTGACGCACTGATTGCTC TGGGTCGGCATGGCATCTCCACC **R:** GAGTTTTTGTTTGCACTGATTGCTCT GAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $M_{(+)}$
pCB301-HH-$M_{(-)eGFP}$-RZ-NOS	$MR_{(-)eGFP}$	**F:** GAATCAAATTTAGCCTGTGACAAGC AGACTTACTTGTACAGCTCGTCCATGC **R:** CCATTATAATCTGAGCAGACGTATA AGATGGTGAGCAAGGGCGAGGAGCTG	To amplify the eGFP for construction of $MR_{(-)eGFP}$
		F: CAGCTCCTCGCCCTTGCTCACCATCT TATACGTCTGCTCAGATTATAATGG **R:** GCATGGACGAGCTGTACAAGTAAGT CTGCTTGTCACAGGCTAAATTTGATTC	To amplify the pCB301 backbone for construction of $MR_{(-)eGFP}$
pCB301-HH-$M_{(-)mCherry}$-RZ-NOS	$MR_{(-)mCherry}$	**F:** GAATCAAATTTAGCCTGTGACAAGC AGACTTAAGATCTGTACAGCTCGTCCAT GC **R:** CCATTATAATCTGAGCAGACGTATA AGATGGTGAGCAAGGGCGAGGAGGAT AAC	To amplify the mCherry for construction of $MR_{(-)mCherry}$
		F: GTTATCCTCCTCGCCCTTGCTCACCA TCTTATACGTCTGCTCAGATTATAATGG **R:** GCATGGACGAGCTGTACAGATCTTA	To amplify the pCB301 backbone for construction of $MR_{(-)mCherry}$

		AGTCTGCTTGTCACAGGCTAAATTTGATTC	
pCB301-HH-$M_{(-)eGFP\&NSmMut}$-RZ-NOS	$MR_{(-)eGFP\&NSmMut}$	**F:** CTCTACCTTAGGCTGTTGAACTCAAAATGTAGACTCTTTTCGGTAATAAGG **R:** GCATGGACGAGCTGTACAAGTAAGTCTGCTTGTCACAGGCTAAATTTGATTC	To amplify the NSm^{Mut} for construction of $MR_{(-)eGFP\&NSmMut}$
		F: GAATCAAATTTAGCCTGTGACAAGCAGACTTACTTGTACAGCTCGTCCATGC **R:** CCTTATTACCGAAAAGAGTCTACATTTTGAGTTCAACAGCCTAAGGTAGAG	To amplify the pCB301 backbone for construction of $MR_{(-)eGFP\&NSmMut}$
pCB301-HH-$L_{(-)}$-RZ-NOS	$L_{(-)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAGAGCAATCAGGTACAACTAAAAC **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATCAGGTAACAACGAT	To amplify the TSWV genomic L-RNA for construction of $L_{(-)}$
		F: ATCGTTGTTACCTGATTGCTCTGGGTCGGCATGGCATCTCCACC **R:** GTTTTAGTTGTACCTGATTGCTCTGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $L_{(-)}$
pCB301-HH-$L_{(+)}$-RZ-NOS	$L_{(+)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAGAGCAATCAGGTAACAACGAT **R:** GTGGAGATGCCATGCCGACCCAGAGCAATCAGGTACAACTAAAAC	To amplify the TSWV antigenomic L-RNA for construction of $L_{(+)}$
		F: GTTTTAGTTGTACCTGATTGCTCTGGGTCGGCATGGCATCTCCAC **R:** ATCGTTGTTACCTGATTGCTCTGAACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $L_{(+)}$
pCB301-HH-$L_{(+)opt}$-RZ-NOS	$L_{(+)opt}$	**F:** ATCAGGTAACAACGATTTTAAGCAAACATGAACATTCAGAAGATCCAAAAGCTG **R:** CATGCATTGTTAGGCATTACTTTTAATCTAATCGGTGTCCTCTTCCTCATCAG	To amplify the RdRp-optimized for construction of $L_{(+)opt}$
		F: CTGATGAGGAAGAGGACACCGATTAGATTAAAAGTAATGCCTAACAATGCATG **R:** CAGCTTTTGGATCTTCTGAATGTTCATGTTTGCTTAAAATCGTTGTTACCTGAT	To amplify the pCB301 backbone for construction of $L_{(+)opt}$
pCB301-HH-$M_{(-)opt}$-RZ-NOS	$M_{(-)opt}$	**F:** ACCATTATAATCTGAGCAGACGTATAAGATGAGGATCCTGAAGCTTCTTG **R:** GAATCAAATTTAGCCTGTGACAAGCAGACCTAAACAAGATGAGAGAAATC	To amplify the GP-optimized for construction of $M_{(-)opt}$
		F: GATTTCTCTCATCTTGTTTAGGTCTGCTTGTCACAGGCTAAATTTGATTC **R:** CAAGAAGCTTCAGGATCCTCATCTT	To amplify the pCB301 backbone for construction of $M_{(-)opt}$

		ATACGTCTGCTCAGATTATAATGGT	
pGEM-NSs	-	**F:** GTTAATACTAACGGAGTGAAAC **R:** GATTGAAATTTGGCTTGAAACAGTAC	To amplify the sense-NSs for construction of pGEM-NSs to generate the DIG-tabled probes of S vRNA in Northern blotting
pGEM-anti-NSs	-	**F:** GATTGAAATTTGGCTTGAAACAGTA **R:** GTTAATACTAACGGAGTGAAAC	To amplify the antisense-NSs for construction of pGEM-anti-NSs to generate the DIG-tabled probes of S cRNA in Northern blotting
pGEM-NSm	-	**F:** GCTTTGACTAAAGCTATGGATAC **R:** TCTTGTATTCTTGGCTGCACATC	To amplify the sense-NSm for construction of pGEM-NSm to generate the DIG-tabled probes of M vRNA in Northern blotting
pGEM-anti-NSm	-	**F:** TCTTGTATTCTTGGCTGCACATC **R:**GCTTTGACTAAAGCTATGGATAC	To amplify the antisense-NSm for construction of pGEM-anti-NSm to generate the DIG-tabled probes of M cRNA in Northern blotting
pGEM-L 3'UTR	-	**F:** AGAGCAATCAGGTACAACTAAAAC **R:** AAGTAATGCCTAACAATGCATGA	To amplify the L 3'UTR for construction of pGEM- L 3'UTR to generate the DIG-tabled probes of L vRNA in Northern blotting
pGEM-anti-L 3'UTR	-	**F:** AAGTAATGCCTAACAATGCATGA **R:** AGAGCAATCAGGTACAACTAAAAC	To amplify the antisense-L 3'UTR for construction of pGEM-anti-L 3'UTR to generate the DIG-tabled probes of L cRNA in Northern blotting
pGEM-eGFP	-	**F:** ATGGTGAGCAAGGGCGAGGAGCTGTTC **R:** ATGGTGAGCAAGGGCGAGGAGCTGTTC	To amplify the sense-eGFP for construction of pGEM-eGFP to generate the DIG-tabled probes of antisense-eGFP RNA in Northern blotting
pGEM-anti-eGFP	-	**F:** ATGGTGAGCAAGGGCGAGGAGCTGTTC **R:** ATGGTGAGCAAGGGCGAGGAGCTGTTC	To amplify the antisense-eGFP for construction of pGEM-anti-eGFP to generate the DIG-tabled probes of sense-eGFP RNA in Northern blotting
-	-	**F:** GGTGGAGATGCCATGCCGACCCAGAGCAATTGTGTCATAATTTTATTCTTA **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATTGTGTCAATTTTATTCAAAC	To amplify the of $S_{(+)eGFP}$ minigenome by RT-PCR
-	-	**F:** GTTCATTTCATTTGGAGAGGAGAGCATCAGTGCAAACAAAAAC **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATCAGTGCGTCAGAAATATAC	To amplify the of $M_{(-)mCherry}$ and $M_{(-)eGFP}$ minigenome by RT-PCR
-	-	**F:** GAATCAAATTTAGCCTGTGACAAGC	To amplify the of $M_{(-)opt}$ genome by RT-

		AGACCTAAACAAGATGAGAGAAATC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATCAGTGCAAACAAAAAC	PCR
-	-	**F:** GATCAAGGATGTTAATTTCAGCATGC TTATCCCGATCCTCGAC **R:** GAATCAAATTTAGCCTGTGACAAGC AGACCTAAACAAGATGAGAGAAATC	To amplify the of $L_{(+)opt}$ antigenome by RT-PCR