PNAS
www.pnas.org

Supplementary Information for

Rescue of Tomato spotted wilt virus 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]

Corresponding Author: Xiaorong Tao

Email: taoxiaorong@njau.edu.cn

This PDF file includes:

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 *N. benthamiana* leaves infected with TSWV lettuce isolate, then inserted into binary vector pCambia2300 or pCXSN to generate p2300-RdRp$_{wt}$, p2300-N, and pCXSN-NSs downstream of a double 35S promoter (2×35S). The pCB301-P19-HcPro-γb construct expressing three VSRs P19, HcPro and γb simultaneously was kindly provided by Dr. Xianbing Wang in College of Biological Sciences of China Agricultural University (1). The codon usage and intron-splicing sites of RdRp were optimized (*SI Appendix*, Fig. S9) and *de novo* synthesized by GenScript Biotech Corp (Nanjing, China), then inserted into binary vector pCambia2300 to generate p2300-RdRp$_{opt}$ downstream of the 2×35S promoter.

Construction of full-length TSWV genomic S$_{(-)}$, M$_{(-)}$, L$_{(-)}$, and antigenomic S$_{(+)}$, M$_{(+)}$ and L$_{(+)}$ cDNA clones. To generate constructs to express full-length TSWV genomic and antigenomic RNA of 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 self-cleaving hammerhead (HH) ribozyme (2) and inserted into binary expression vector pCB301-2×35S-RZ-NOS linearized by two restriction endonucleases StuI and SmaI (3). cDNA clones of 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 [$L_{(+)}$] were then generated. The full-length TSWV genomic RNA of $S_{(-)}$, $M_{(-)}$ and $L_{(-)}$, and antigenomic RNA of $S_{(+)}$, $M_{(+)}$ and $L_{(+)}$ were expressed downstream of the double 35S promoter (2×35S) and flanked with a self-cleaving hammerhead (HH) ribozyme at the 5′-terminus and a hepatitis delta virus (HDV) ribozyme at the 3′-terminus.

Construction of TSWV $SR_{(-)eGFP}$, $SR_{(-)mCherry\&eGFP}$ and $SR_{(+)eGFP}$ mini-replicons. For generating $SR_{(-)eGFP}$-genomic RNA mini-replicon, the eGFP ORF was amplified and used to replace the N gene in the pCB301-2×35S-HH-$S_{(-)}$-RZ-NOS by *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.

For generating $SR_{(-)mCherry\&eGFP}$ 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 the In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-$SR_{(-)mCherry\&eGFP}$-RZ-NOS [35S:$SR_{(-))mCherry\&eGFP}$] was generated. The T7:$SR_{(-)mCherry\&eGFP}$ mini-replicon (pCB301-T7-HH-$SR_{(-)mCherry\&eGFP}$-RZ-NOS) controlled by T7 promoter was constructed by the same strategy as 35S:$SR_{(-)mCherry\&eGFP}$.

For generating antigenomic $S_{(+)eGFP}$-mini-replicon, the eGFP ORF was amplified and used to replace the NSs gene in the pCB301-2×35S-HH-$S_{(-)}$-RZ-NOS by recombination using the 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 TSWV $MR_{(-)eGFP}$, $MR_{(-)mCherry}$ and $MR_{(-)eGFP\&NSmMut}$ mini-replicons. For generating $MR_{(-)eGFP}$ and $MR_{(-)mCherry}$ mini-replicons, 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 the 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 [$MR_{(-)mCherry}$] were generated.

For generating $MR_{(-)eGFP\&NSmMut}$ 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 the In-Fusion Cloning mixture (Clontech). The construct pCB301-2×35S-HH-$MR_{(-)eGFP\&NSmMut}$-RZ-NOS [$MR_{(-)eGFP\&NSmMut}$] was generated. All primers used above are listed in *SI Appendix*, Table S3.

Construction of full-length $L_{(+)opt}$ and $M_{(-)opt}$ cDNA clones. For generating the full-length $L_{(+)opt}$ cDNA clone**,** RdRp with the sequence codon and intron-splicing sites optimized 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 construct pCB301-2×35S-HH-$L_{(-)opt}$-RZ-NOS [$L_{(-)opt}$] was generated.

To generate the full-length $M_{(-)opt}$ cDNA clone**,** the GP gene with codon and intron-splicing sites optimized 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 *in vitro* recombination using the In-Fusion

Cloning mixture (Clontech). The construct 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 8-week-old plants of *Nicotiana benthamiana*, were used in all agroinfiltration assays. The TSWV isolate from asparagus lettuce (TSWV-LE) was used in this study (GenBank accession KU976396 for S, JN664253 for M and KU976394 for L) (4). The TSWV-LE isolate was maintained in *N. benthamiana*. For long-term storage, infected new leaves of *N. benthamiana* were kept in a -80 °C refrigerator. The *N. benthamiana* inoculated with TSWV-LE were grown in a growth chamber at 25 °C and 16 h light/8 h dark (5).

***Agrobacterium* infiltration**

Agrobacterium tumefaciens strain GV3101 was transfected with recombinant plasmids using electroporation and used for agroinfiltration of *N. benthaminana* leaves essentially as described before (6, 7). *A. tumefaciens* cells were resuspended in agroinfiltration buffer (10 mM MgCl2, 10 mM MES pH 5.6, and 100 μM acetosyringone) adjusted to an optical density OD$_{600}$ of 1.0 and incubated for 2 to 3 h in the dark at room temperature. Equal volumes of *Agrobacterium* cultures (final concentration OD$_{600}$ = 0.2) harboring the p2300-N, p2300-RdRp, pCB301-derived reporter or full-length infectious clone vector(s), were mixed with one volume of bacterial mixture (final concentration OD$_{600}$ = 0.05) containing the NSs and P19-

HcPro-γb. Fully expanded leaves of *N. benthamiana* plants at the 6-7-leaf stage were infiltrated with *Agrobacterium* cultures using 1 mL needleless syringes. The *N. benthamiana* plants infiltrated with *Agrobacterium* cells were grown in a growth chamber at 25 °C and 16 h light/8 h dark (5).

Western blot analysis

Total protein was extracted from 1 g *Agrobacterium*-infiltrated leaf patches, healthy or TSWV-systemically infected leaves of *N. benthamiana* in 1 mL extraction buffer (10 % v/v glycerol, 25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10 mM dithiothreitol, 2 % w/v polyvinylpolypyrrolidone, 0.5 % v/v Triton X-100 and 1× protease inhibitors cocktail) (8). Protein samples were separated by SDS-PAGE, transferred to PVDF membranes (GE Healthcare, UK), blocked with 5 % w/v skim milk solution and incubated with a polyclonal antiserum specific to TSWV N, NSm, NSs, Gn, Gc, GFP, mCherry or T7 RNA pol at room temperature for 1 h or overnight at 4 °C. After incubation in a secondary antibody containing HRP-conjugated goat antirabbit (1:10000) for 1 h, the blots were developed using the ECL Substrate Kit (Thermo Scientific, Hudson, NH, USA). The signal was visualized using the ChemiDoc Touch Imaging System (Bio-Rad). Protein loading was evaluated by staining the blots with Ponceau S. Rabbit polyclonal antibodies against RdRp, Gn, Gc, NSm, N and NSs were produced in our laboratory. Secondary antibody HRP-conjugated goat antirabbit IgG was purchased from Sigma-Aldrich (St. Louis, MO, USA).

Northern blot analysis

For Northern blot analysis of TSWV gRNAs, agRNAs or viral mRNA transcripts, total RNAs were extracted from *Agrobacterium*-infiltrated leaf patches, healthy or TSWV-systemically infected leaves using an RNAprep Pure Plant Kit (Tiangen Biotech, Beijing, China), respectively. DIG-labeled probes specific for sense or antisense eGFP, NSs, NSm, and L-5′UTR was synthesized using a DIG High Prime RNA labeling kit (Roche, Basel, Switzerland). Total RNAs were separated in 1 % formaldehyde agarose gels and transferred to Hybond-N^+ membranes (GE Healthcare, UK) (9). The membrane blots were hybridized with the specific DIG-labeled probe, then processed using a DIG-High Prime Detection Starter Kit II (Roche) and the manufacturer's protocol.

RT-PCR and sequencing analysis

For detecting the virus in leaves of *N. benthamina* systemically 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 leaf areas with systemic symptoms. First-strand cDNAs were synthesized using M-MLV Reverse Transcriptase (Promega, Madison, WI, USA). RT-PCRs were performed to detect the $SR_{(+)eGFP}$, $MR_{(-)mCherry}$, $MR_{(-)eGFP}$, $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ minigenomes 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

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

Transmission electron microscopy and immunogold labelling

Samples (1 mm × 4 mm) were excised from leaves of *N. benthamiana* plants infected with rTSWV rescued from the full-length infectious clones. The sample tissues were fixed in 2.5 % v/v glutaraldehyde and 1 % w/v osmium tetroxide in 100 mM phosphate buffer (PB; pH 7.0) as described by Li *et al.* (6, 10) and then embedded in Epon 812 resin as instructed by the manufacturer (SPI-EM, Division of Structure Probe, West Chester, PA, USA). Ultrathin sections (70 nm) were mounted on formvar-coated grids and then stained with uranyl acetate for 10 min, then with lead citrate for 10 min. The stained sections were examined with a transmission electron microscope (TEM; H-7650, Hitachi, Japan).

For immunogold labelling, the TEM sections were incubated in 1 % BSA buffer for 10 min, then for 1 h in 1:200 v/v diluted rabbit polyclonal antibody against Gn. The

sections were rinsed with PB, then incubated in 1:30 v/v diluted goat antirabbit IgG conjugated with gold particles for 1 h. After several rinses with PB, the sections were examined with a transmission electron microscope.

Imaging GFP in infected plant using hand-held UV lamp

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

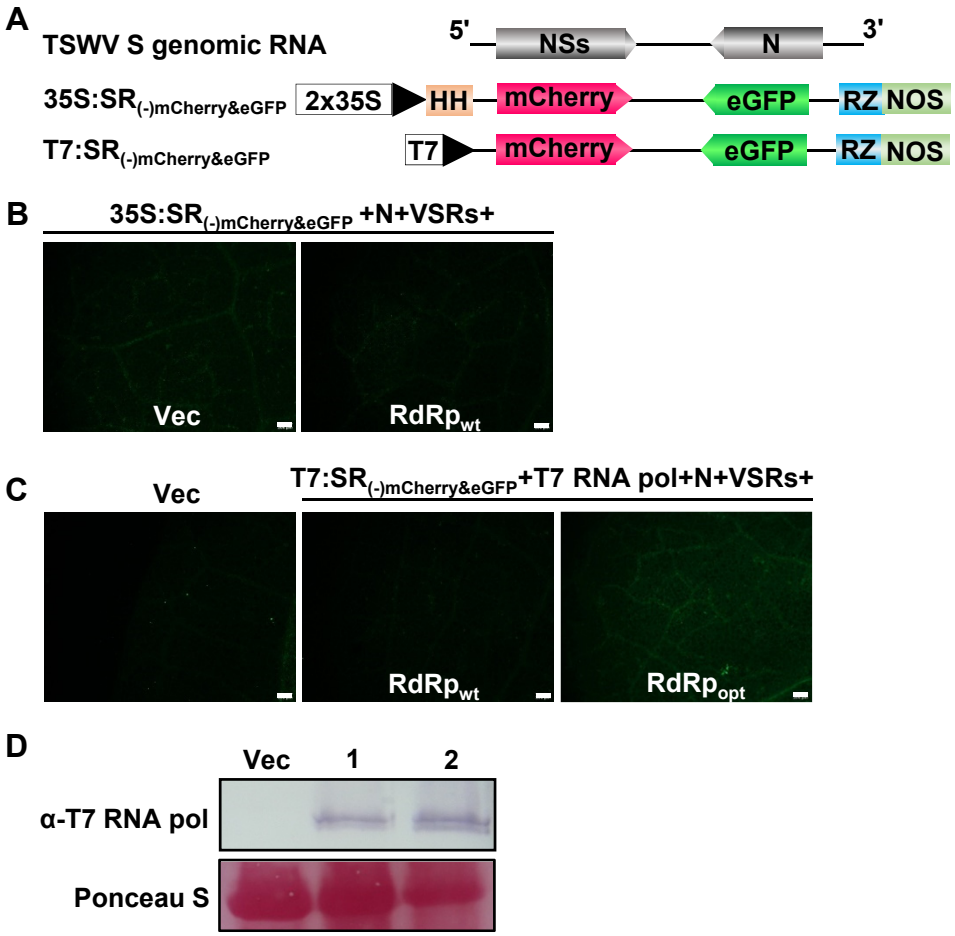

Fig. S1 Functional analysis of wild-type RdRp and use of T7 promoter in a mini-genome replication assay. (*A*) Schematic diagram of TSWV 35S:SR$_{(-)mCherry\&eGFP}$ and T7:SR$_{(-)mCherry\&eGFP}$ mini-replicon reporters. (*B*) Wild-type RdRp (RdRp$_{wt}$) or the empty vector (Vec) was co-expressed with 35S:SR$_{(-)mCherry\&eGFP}$, N, VSRs (VSRs: NSs, P19, HcPro and γb) in *N. benthamiana* leaves. Expression of eGFP was detected with 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 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 polymerase (pol) using a T7 RNA pol-specific antibody. Ponceau S staining was used as protein loading control. Lane 1: sample from leaves coexpressing T7:SR$_{(-)mCherry\&eGFP}$, T7 RNA Pol, N, VSRs and RdRp$_{wt}$; lane 2: sample from leaves coexpressing T7:SR$_{(-)mCherry\&eGFP}$, T7 RNA Pol, N, VSRs and RdRp$_{opt}$.

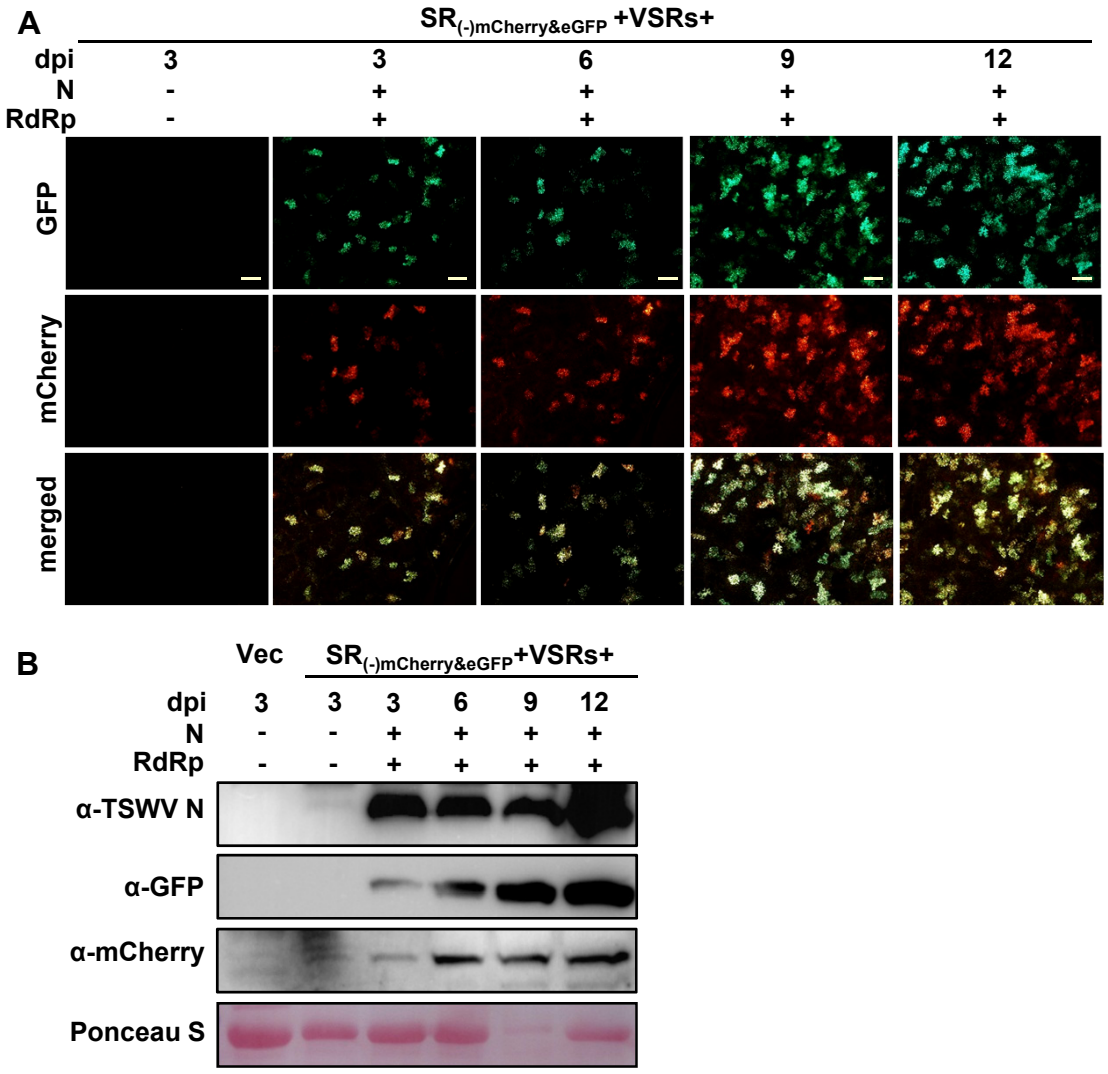

Fig. S2 Time-course analysis of gene expression from the $SR_{(-)mCherry\&eGFP}$ mini-replicon in *N. benthamiana* leaves. (*A*) Foci of eGFP and mCherry fluorescence expressed from $SR_{(-)mCherry\&eGFP}$ in *N. benthamiana* leaves coexpressing N, RdRp and the VSRs at 3, 6, 9 and 12 dpi, respectively. Fluorescence of eGFP and mCherry was photographed with a fluorescence microscope using GFP and RFP filters, respectively. Bars represent 400 μm. (*B*) Western immunoblot detection of N, eGFP and mCherry proteins in leaves shown in panel A, using specific antibodies against N, GFP and mCherry, respectively. The empty vector (Vec) was used as a negative control. Ponceau S staining was used as protein loading control.

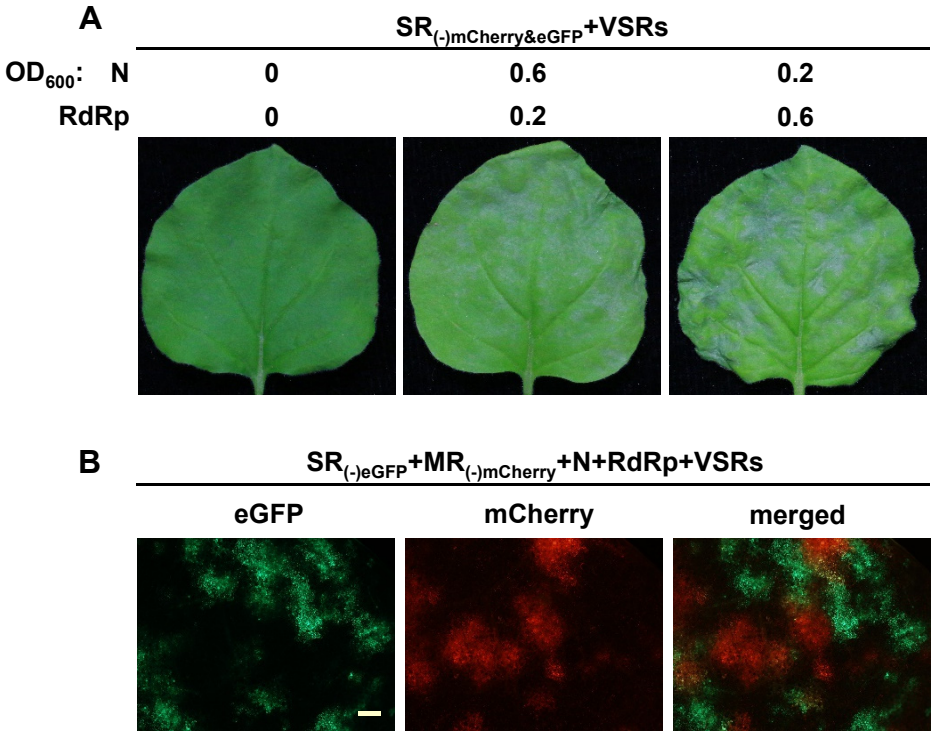

Fig. S3 Effects of ectopic expression of TSWV RdRp and N on host cells or intercellular movement of $SR_{(-)GFP}$ and $MR_{(-)mCherry}$ mini-replicons in *Nicotiana benthamiana*. (*A*) Expression of N or RdRp at high concentrations (OD_{600} > 0.6) of *Agrobacterium* induced cell death in the infiltrated leaves of *N. benthamiana*. *Agrobacterium* at OD_{600} = 0.6 containing the binary expression constructs for N or RdRp were mixed with fixed amounts of *Agrobacterium* containing the RdRp or N construct (OD_{600} 0.2), respectively, for infiltrating *N. benthamiana* leaves with *Agrobacterium* containing $SR_{(-)mCherry\&GFP}$ and four VSRs; cell death phenotype was photographed at 5 dpi. (*B*) Cell-to-cell movement analysis of $SR_{(-)GFP}$ and $MR_{(-)mCherry}$ in *N. benthamiana* coexpressing RdRp, N and four VSRs. Agroinfiltrated leaves were examined and photographed at 5 dpi with a fluorescence microscope. Bars represent 400 μm.

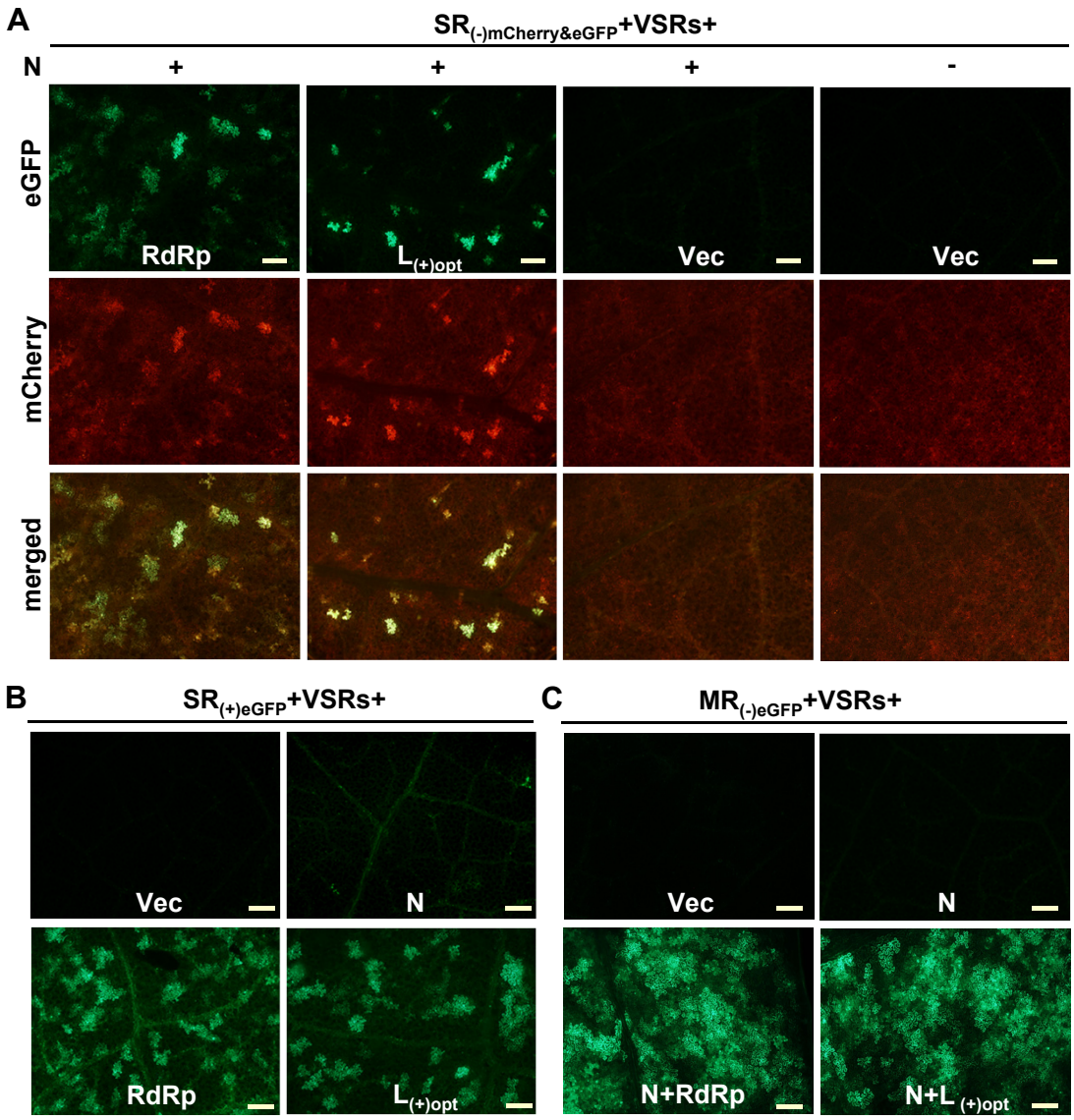

Fig. S4 Functional analysis of RdRp expressed from TSWV $L_{(+)opt}$, NSm from $MR_{(-)eGFP}$ and N from $SR_{(+)eGFP}$ using the mini-genome replication system in *N. benthamiana*. (*A*) Functional analysis of RdRp expressed from TSWV $L_{(+)opt}$ using the S RNA mini-replicon system in *N. benthamiana*. The $L_{(+)opt}$, RdRp, or pCB301 empty vector (Vec) was coexpressed with N, $SR_{(-)mCherry\&eGFP}$ and the four VSRs in leaves. (*B*) Functional analysis of N expressed from $SR_{(+)eGFP}$ in *N. benthamiana.* $SR_{(+)eGFP}$ was coexpressed with the empty vector (Vec), N, RdRp or $L_{(+)opt}$ in leaves in the presence of four VSRs. (*C*) Functional analysis of NSm expressed from $MR_{(-)eGFP}$ in *N. benthamiana.* $MR_{(-)eGFP}$ was coexpressed with the empty vector (Vec), N, N+RdRp or N+$L_{(+)opt}$ in leaves in the presence of four VSRs. Agroinfiltrated leaves were examined with a fluorescence microscope at 3 dpi for foci showing mCherry and eGFP fluorescence. Bars represent 400 μm.

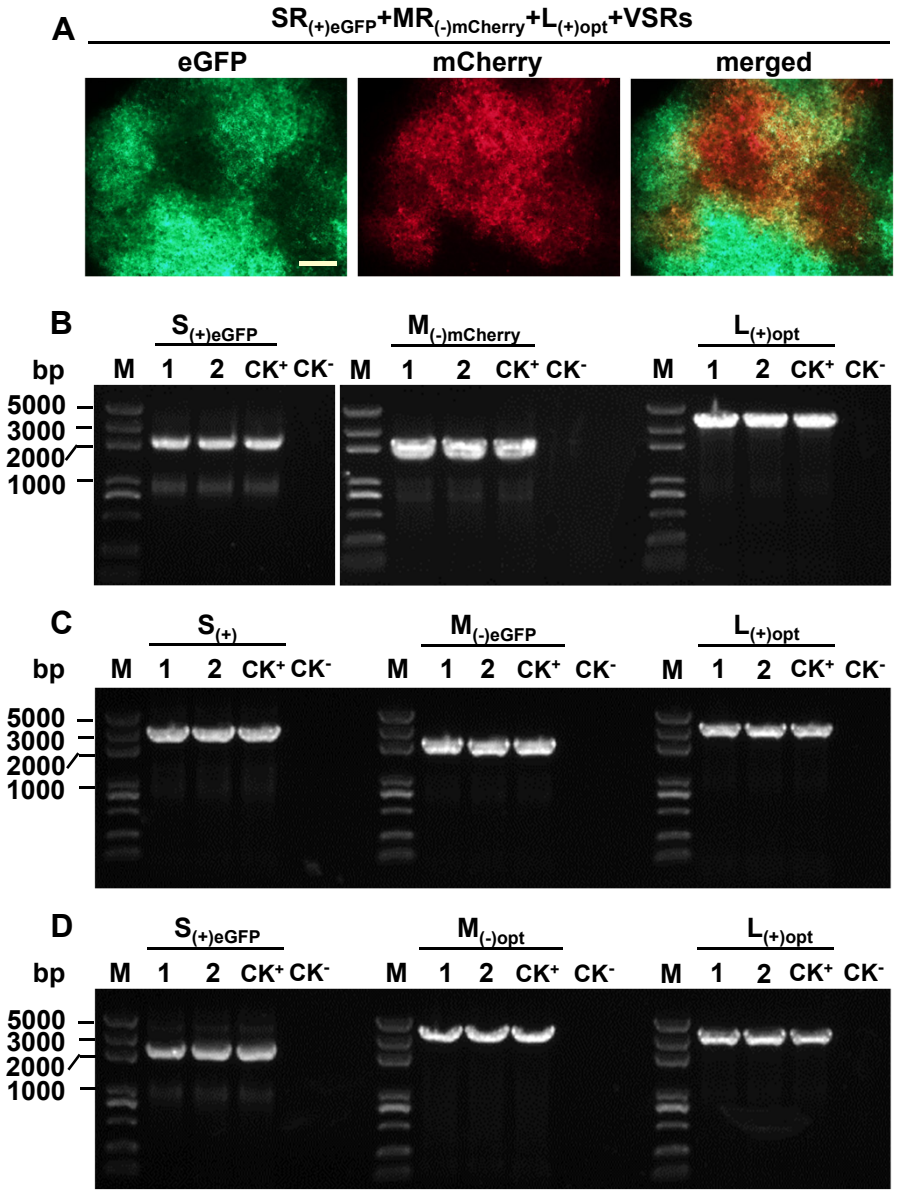

Fig. S5 Analysis of *N. benthamiana* leaves agroinfiltrated with constructs of $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$, or $SR_{(+)eGFP}$, $M_{(-)opt}$ and $L_{(+)opt}$, or $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$. (*A*) Local infection analysis of cell-to-cell movement of $SR_{(+)eGFP}$ and $MR_{(-)mCherry}$ coexpressing with $L_{(+)opt}$ and four VSRs in leaves after agroinfiltration. Agroinfiltrated leaves were examined and photographed at 5 dpi using a fluorescence microscope. Bars represent 400 μm. (*B*) RT-PCR analysis of systemically infected leaves from *N. benthamiana* plants after agroinfiltration with $SR_{(+)eGFP}$, $MR_{(-)mCherry}$ and $L_{(+)opt}$ at 15 dpi. (*C*) RT-PCR analysis of systemically infected leaves from *N. benthamiana* plants agroinfiltrated with $S_{(+)}$, $MR_{(-)eGFP}$ and $L_{(+)opt}$ at 15 dpi. (*D*) RT-PCR analysis of systemically infected leaves from *N.*

benthamiana plants agroinfiltrated with $SR_{(+)eGFP}$, $M_{(-)opt}$ and $L_{(+)opt}$ at 15 dpi. All agroinfiltrations included four VSRs. For RT-PCR, total RNA was purified from systemically infected leaves to detect S, M or L segments using segment-specific primers. Amplicons were resolved by electrophoresis in 1 % agarose. Lanes 1 and 2 represent two biological replicates of systemically infected leaf samples. For positive controls (CK^+) of proper fragment size, PCR was performed on plasmids carrying $S_{(+)}$, $M_{(-)opt}$, $L_{(+)opt}$ or derivatives. As the negative control (CK^-), RT-PCR was performed in the absence of nucleic acids. DNA size markers are shown on the left.

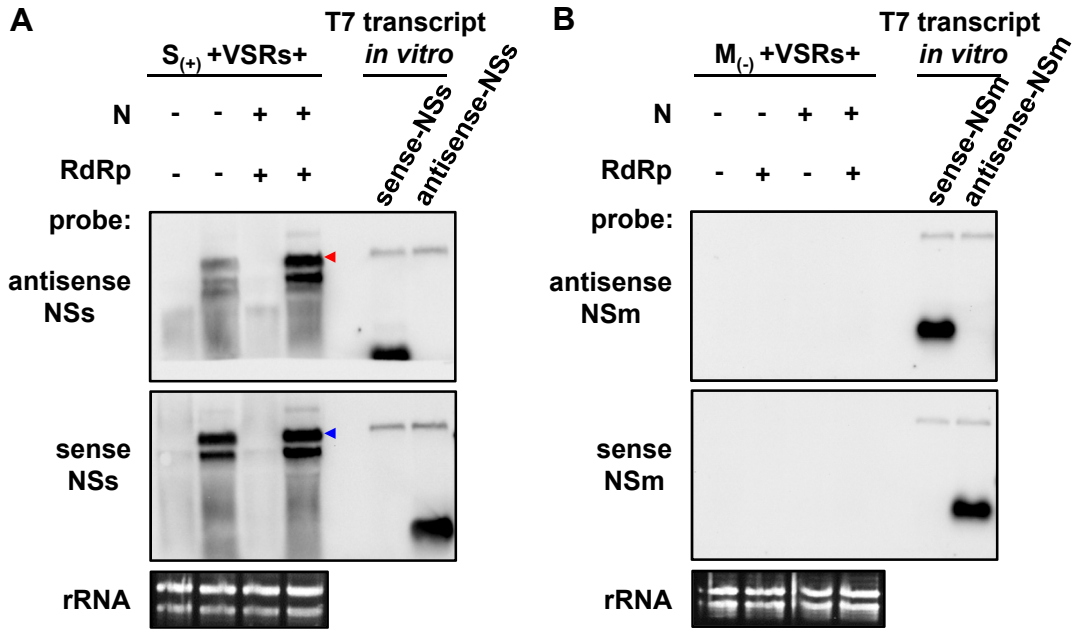

Fig. S6 Northern blot detection of viral RNA synthesis produced from full-length $S_{(+)}$ and wild-type $M_{(-)}$ replicons. (*A* and *B*) The full-length $S_{(+)}$ or wild-type $M_{(-)}$ was coexpressed with the empty vector (Vec), N, $RdRp_{opt}$ or $N+RdRp_{opt}$ in *N. benthamiana* leaves in the presence of three (P19, HcPro and γb) VSRs for $S_{(+)}$ or four VSRs (P19, HcPro, γb and NSs) for $M_{(-)}$. Genomic RNAs (blue arrow), antigenomic RNAs (red arrow) of S (*A*) were detected with DIG-labeled sense or antisense NSs and NSm probes, respectively. Ethidium bromide staining was used as RNA loading control.

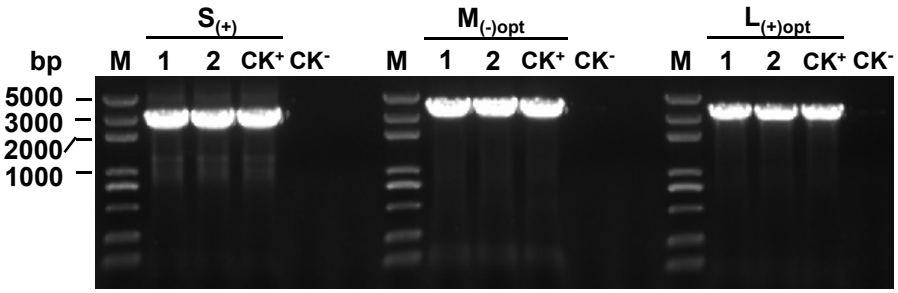

Fig. S7 RT-PCR detection of $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ genomic RNA in leaves of *N. benthamiana* systemically infected with rTSWV. $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ and the four VSRs were coexpressed in leaves via agroinfiltration. Total RNA was purified from systemically leaves of agroinfiltrated plants, and $S_{(+)}$, $M_{(-)opt}$ and $L_{(+)opt}$ were detected by RT-PCR using segment-specific primers. RT-PCR products were resolved by electrophoresis in 1 % agarose. Lanes 1 and 2, two biological replicates of systemic infected leaf samples; RT-PCR on plasmids carrying S, M and L as DNA template were used as positive controls (CK^+). RT-PCR running without the DNA template were used as negative controls (CK^-). DNA size markers are shown on the left.

Fig. S8. Wild-type and optimized RdRp gene sequence used in the study.

Wild-type RdRp gene sequence

ATGAACATCCAGAAAATACAAAAATTAATAGAGAATGGAACTACTTTATTGTTGTCTATTGAGGATTGTGTAGGTTCTAATT
ACGACCTAGCTTTGGATTTACATAAGAGAAATAGTGATGAGATCCCAGAAGATGTGATTATTAATAATAATGCAAAAAATTA
TGAGACGATGAGAGAGTTAATTGTCAAAATCACTGCTGATGGAGAAGGACTAAACAAAGGGATGGCAACTGTAGACGTC
AAGAAGTTAAGTGAGATGGTTTCTCTATTTGAGCAAAAATACCTAGAAACAGAGTTAGCAAGGCATGATATTTTTGGAGAG
CTGATCTCCAGGCACCTGAGAATAAAGCCCAAACAAAGAAGTGAAGTGGAGATAGAGCATGCGCTCAGAGAATATCTGGA
TGAACTTAACAAAAAGTCCTGCATTAATAAGCTCTCTGATGATGAGTTTGAGAGAATAAATAAAGAATATGTGGCAACTAA
CGCCACACCTGACAACTATGTGATATATAAAGAATCAAAAAACAGTGAGCTTTGTTTAATCATTTATGATTGGAAAATATCT
GTCGATGCCAGGACAGAAACTAAAACAATGGAGAAATACTACAAAAATATCTGGAAATCTTTCAAAGATATAAAAGTGAAT
GGAAAGCCATTCCTGGAAGATCACCCTGTTTTCGTTTCTATAGTTATATTGAAACCTATTGCTGGGATGCCAATCACTGTTA
CTAGCAGCAGGGTTTTGGAGAAATTTGAAGATTCTCCATCAGCATTGCATGGAGAAAGAATAAAGCATGCTAGAAATGCC
AAATTGCTAAATATTTCTCATGTTGGGCAAATAGTTGGAACCACACCCACAGTGGTGAGAAACTATTATGCAAACACTCAA
AAGATCAAATCTGAAGTCAGAGGAATACTAGGTGATGATTTTGGATCTAAAGATGTGTTTTTTAGTCACTGGACCAGCAAA
TACAAAGAAAGAAATCCTACTGAAATAGCCTATTCTGAAGATATTGAAAGAATAATTGATTCACTTGTTACAGATGAAATAA
CTAAAGAGGAAATAATACATTTTTTGTTTGGAAATTTCTGTTTCCACATTGAAACAATGAATGACCAGCATATCGCCGACA
AATTTAAAGGGTACCAAAGCTCTTGTATCAATTTAAAAATAGAGCCAAAAGTTGATTTAGCTGATTTGAAGGACCACTTAA
TCCAGAAGCAGCAAATATGGGAATCTCTGTATGGGAAACATCTTGAGAAAATCATGCTTAGAATTAGAGAAAAAAAGAAA
AAAGAAAAAGAAATACCTGACATAACCACAGCTTTTAACCAGAATGCTGCTGAATATGAAGAAAAGTACCCAAACTGTTTT
ACAAATGATCTCTCTGAAACTAAAACTAACTTTTCTATGACTTGGTCCCCAAGTTTTGAAAAAATAGAATTGAGCTCAGAG
GTAGATTACAACAACGCAATTATAAACAAGTTTCGGGAAAGCTTCAAAAGTTCTTCAAGGGTTATTTATAATAGCCCATATA
GTTGCATAAATAACCAAACAAATAAAGCAAGAGATATAACAAACTTAGTTAGACTGTGTTTAACAGAGCTAAGTTGTGATA
CAACAAAAATGGAAAAGCAGGAGCTTGAGGATGAAATAGATATAAACACTGGGAGTATTAAAGTTGAGAGAACAAAAAAA
TCTAAAGAATGGAATAAGCAGGGTTCGTGTTTAACCAGGAACAAAAATGAATTTTGCATGAAAGAGACAGGCAGGGAGAA
CAAAACTATCTATTTTAAAGGCTTAGCAGTAATGAATATAGGAATGAGTTCTAAAAAAAGAATTCTAAAAAAAGAAGAAATA
AAAGAGAGGATTTCTAAAGGCTTGGAATATGATACCTCTGAAAGGCAGGCTGACCCAAATGATGATTATTCAAGTATAGAC
ATGTCATCTCTGACTCATATGAAAAAACTGATAAGACATGACAACGAGGATAGCTTAAGCTGGTGTGAAAGGATTAAGGAC
TCTTTGTTTGTTCTTCATAATGGTGATATAAGAGAGGAAGGCAAGATCACATCTGTTTACAATAATTATGCTAAAAATCCTG
AATGCTTGTACATTCAAGATTCAGTACTGAAAACTGAATTAGAGACTTGCAAAAAGATAAACAAATTATGCAATGACCTAG
CCATTTACCATTACTCTGAAGACATGATGCAATTCTCCAAAGGTTTAATGGTGGCTGACAGGTACATGACCAAAGAAAGTT
TCAAAATACTAACCACAGCAAATACGAGCATGATGCTATTGGCATTCAAAGGAGATGGAATGAACACCGGAGGATCAGGA
GTTCCTTACATAGCATTGCATATAGTGGATGAAGACATGTCAGATCAATTTAACATATGTTATACTAAAGAAATTTATAGCTA
TTTCCGAAATGGTAGTAATTACATTTATATAATGAGACCACAGAGACTCAATCAGGTGAGGCTGCTAAGCCTTTTCAAAAC
GCCTAGTAAAGTTCCTGTATGTTTTGCACAATTTTCAAAGAAAGCTAACGAAATGGAAAAATGGCTGAAAAACAAAGATAT
AGAAAAAGTAAATGTCTTTTCTATGACAATGACTGTAAAACAGATATTAATAAATATTGTGTTTTCATCTGTCATGATAGGA
ACTGTGACAAAGCTCAGCAGGATGGGAATTTTTGACTTCATGCGGTATGCAGGTTTTTTGCCACTGTCTGATTATTCTAAT
ATAAAAGAGTACATTAGAGACAAATTTGATCCTGATATAACCAATGTGGCAGATATCTATTTCGTTAATGGAATCAAAAAGC
TATTGTTCAGAATGGAAGATCTCAATTTAAGCACAAATGCCAAGCCTGTTGTTGTAGATCACGAAAATGATATTATAGGAG
GGATAACAGACTTGAATATAAAATGTCCAATAACAGGATCAACTCTACTGACTCTTGAGGATTTGTATAATAATGTCTATTT
GGCTATTTACATGATGCCTAAATCATTGCACAATCACGTTCACAATCTAACAAGCTTGTTAAATGTCCCTGCTGAGTGGGA
GCTAAAGTTCAGAAAAGAGTTAGGTTTCAACATATTTGAGGACATATACCCTAAGAAAGCAATGTTTGATGACAAAGACCT
ATTCTCTATAAATGGAGCTTTGAACGTGAAAGCATTATCTGATTACTATCTAGGAAATATAGAAAATGTGGGTTTGATGAGA
TCAGAAATAGAAAATAAAGAAGATTTTCTAAGCCCTTGTTATAAAATATCTACTTTAAAATCTTCAAAAAAATGCTCGCAGT
CAAACATTATAAGTACTGATGAGATAATAGAGTGTCTCCAGAATGCAAAGATCCAAGATATAGAAAATTGGAAAGGAAATA
ACCTGGCCATTATAAAAGGGCTTATAAGAACCTACAATGAGGAGAAAAACAGATTGGTGGAATTTTTTGAAGATAATTGTG
TCAATTCACTGTATCTTGTAGAAAAGCTTAAAGAGATAATTAATAGTGGATCGATAACTGTAGGGAAATCCGTAACATCTAA
ATTTATAAGAAATAACCATCCTTTAACAGTAGAAACATATCTCAAAACAAAACTATATTATAGAAATAATGTGACTGTTTTAA
AGTCTAAAAAAGTATCAGAGGAGCTTTATGACCTTGTAAAACAGTTTCATAACATGATGGAAATAGACTTAGATTCTGTTAT

GAACCTTGGGAAAGGTACAGAAGGGAAAAAACACACATTCTTGCAGATGCTTGAATTTGTCATGTCCAAAGCTAAAAATG
TCACCGGGTCTGTAGATTTTCTAGTTTCTGTTTTTGAAAAAATGCAGAGAACCAAAACAGACAGAGAAATATACTTGATGA
GTATGAAAGTGAAAATGATGCTTTATTTTATAGAGCACACATTCAAACATGTTGCGCAGAGTGATCCATCAGAAGCCATAT
CTATAAGTGGAGACAATAAAATAAGGGCACTTTCTACATTATCTTTGGACACAATCACGTCTTACAACGACATTTTAAACAA
AAATTCAAAGAAGTCAAGATTGGCTTTCCTGTCTGCTGATCAATCGAAATGGTCGGCATCAGACCTTACCTATAAATATGT
TTTAGCTATCATATTAAATCCAATTTTAACTACTGGTGAAGCTAGTTTAATGATAGAATGTATTTTAATGTATGTTAAATTGAA
GAAGGTTTGTATACCAACAGATATTTTTTTGAATCTAAGAAAAGCTCAAGGAACTTTCGGGCAAAATGAAACTGCCATAGG
ACTTTTGACTAAAGGTTTGACGACAAACACATACCCTGTTAGCATGAACTGGTTACAAGGCAATTTAAATTATCTGTCTTC
TGTTTATCACTCTTGTGCAATGAAAGCTTATCACAAGACTTTAGAATGCTACAAAGACTGTGATTTCCAAACTAGATGGAT
TGTGCACTCTGATGATAATGCGACATCATTGATAGCCAGTGGAGAGGTCGATAAAATGCTAACAGACTTTTCAAGCTCATC
TCTGCCAGAAATGTTGTTTAGAAGCATTGAAGCTCATTTCAAAAGCTTTTGCATAACTTTGAACCCAAAAAAGAGTTATGC
TTCTTCATCAGAAGTAGAGTTTATATCTGAAAGAATTGTAAATGGAGCAATTATTCCTCTCTATTGCAGACATTTAGCAAAT
TGTTGCACAGAATCTTCACATATAAGTTATTTTGATGATCTAATGTCACTAAGTATACATGTTACGATGCTTCTGAGAAAAG
GCTGTCCTAATGAAGTTATACCTTTTGCTTATGGGGCTGTACAGGTACAAGCATTAAGCATCTATTCAATGCTTCCTGGTG
AAGTGAATGATAGCATCAGAATTTTTAAGAAGCTTGGAGTAAGTTTAAAGTCAAACGAGATTCCCACAAACATGGGGGGC
TGGTTGACTTCTCCTATAGAGCCATTGTCTATATTAGGTCCATCATCGAATGATCAAATCATCTATTACAATGTGATAAGAG
ATTTTTTGAACAAAAAAAAGTTTAGAGGAAGTAAAAGATAGCGTCTCTTCTTCCAGCTATCTACAGATGAGATTCAGAGAGC
TAAAAGGAAAGTATGAAAAAGGAACTCTGGAAGAAAAAGATAAAAAGATGATATTTCTTATCAATCTGTTTGAGAAAGCAT
CAGTGTCTGAAGATTCAGACGTTCTAACAATCGGGATGAAATTTCAAACTATGTTAACTCAGATTATAAAATTGCCCAATTT
TATAAATGAGAATGCTTTAAACAAGATGTCAAGTTATAAAGATTTTTCAAAACTTTACCCTAATTTGAAAAAGAATGAAGAT
TTATATAAAAGTACTAAGAACTTAAAGATAGACGAGGATGCTATTTTAGAGGGAGATGAGTTATATGAGAAGATTGCATCTA
GCTTAGAAATGGAATCTGTTCATGACATAATGATAAAAAATCCTGAAACAATTCTGATAGCACCATTGAATGATAGAGATTT
TTTACTTAGTCAGCTGTTCATGTACACAAGCCCTTCTAAAAGGAATCAGTTATCTAACCAATCTACAGAGAAACTTGCTTT
GGATAGAGTATTAAGGTCAAAAGCTAGAACATTTGTAGACATTTCTTCCACTGTGAAGATGACTTATGAAGAAAACATGGA
AAAGAAAATCTTGGAAATGCTAAAATTTGATTTAGATTCATATTGTTCATTTAAAACATGTGTAAATCTAGTGATCAAGGAT
GTTAATTTCAGCATGCTGATCCCAATATTGGATTCTGCATACCCTTGTGAATCTAGGAAAAGAGATAACTACAATTTCAGGT
GGTTTCAGACTGAGAAGTGGATACCTGTAGTTGAAGGCTCTCCGGGATTAGTGGTAATGCATGCTGTCTATGGATCAAAT
TATATAGAAAATTTAGGTTTAAAAAACATCCCTCTAACAGACGATAGCATTAATGTTTTAACAAGCACGTTTGGAACAGGTT
TAATCATGGAGGATGTAAAATCCTTAGTTAAAGGCAAAGACAGCTTCGAAACAGAGGCTTTCAGCAATTCTAACGAATGTC
AAAGATTGGTGAAAGCATGCAATTATATGATAGCAGCACAAAACAGGCTTTTAGCAATTAACACATGCTTTACTAGGAAAA
GCTTCCCCTTCTATTCTAAATTCAACCTAGGGAGAGGGTTTATCTCAAACACATTAGCTCTCCTATCCACCATCTACAGTAA
AGAAGAATCCTATCATTTTGTTTCTACAGCTAGTTATAAATTAGACAAAACTATTAGAACTGTGATAAGTGCTCAGCAAGAT
ATGAACTTAGAGAAAATACTGGACACTGCTGTATACATATCAGATAAATTGCAGTCACTTTTCCCAACAATTACAAGAGAG
GATATAGTTTTGATATTGCAAAATGTGTGCCTAGACAGCAAACCTATATGGCAGAGTCTAGAAGACAAAATGAAAAAGATT
AACAATTCAACAGCAAGTGGCTTCACAGTGTCAAATGTGATTCTGTCACATAACAGTGAATTGAACACAATCCAGAAACA
AATTGTCTGGATGTGGAACATGGGTTTATGTTCTCATAGAACATTAGATTTTGTTATCAGGTATATTAGGAGAAGTGATGTA
AGATATGTGAAAACTGAAGAACAAGACGAATCAGGAAATTATATCTCTGGAACTATGTACAAAATCGGGATCATGACAAGA
AGCTGTTATGTTCAATTGATAGCATCTGATCAAGATGTAGCAGTTTCTTTGAGGACACCATTTGAGATATTGAATGAAAGA
GATTATCTTTTTGACACATACAGAGAAAGTATAGAGAAATTATTGCAGAAATTTATGTTTGATAAAGTGAACATAATAAAAT
CAAAACAACCACAGATTGTTTTCCTAGAACCAGGAGATGCCTGCATCAGAATGACCACAGACAACAAGATGATTGTAAAG
GTTAATGCCACATCAAGACAAATAAGACTAGAGAATGTGAAATTAGTTGTGAAGATAAAATATGAAAACGTGAACTCCGAT
GTGTGGGATATTATAGAAAGCCAAAAATCTCTAGTCTTAAGGCTCCCTGAAGTAGGGGAATGTTTCTCTGATATGTACAAA
ACTGCAGACTCTGAAACTGAAACAATCAAAACCGTAAAAAACAGGCTTATGACTTCTTTAACTTTCATAGAAGCCTTTGGA
AACTTATCACAGCAGATCAAAGAGATTGTGGATGATGATATCAGAGAAACGATGGATGAATTCTTAATGAACATCCGGGAT
ACCTGCTTAGAAGGTTTGGAAAACTGCAAAAGTGTGGAAGAATATGATAGCTATCTTGATGAAAATGGGTTTAATGATACA
GTAGAACTATTCGAAAACTTGTTAAGAACACATGACAACTTTGAAAATGAGTATAGTCCTCTGTTTTCAGAGATTGTTGAC
AAAGCAAAACAGTATACTAGAGATTTAGAAGGTTTCAAAGAAATACTGCTCATGCTTAAATATTCTTTGATAAATGATGCAT
CAGGATTTAAAAGCTATAGAGCCACTGGAATGCATGCTGTCGAACTAATGGCAAAAAAGCACATAGAGATAGGGGAATTC
AACTTGCTAGGAATGATCCAACTGATTAAAGCTTGCGAAACATGCCACAACAACGACTCTATATTAAACTTAGCAAGTTTG
AGGAATGTTCTTAGCAGGACATATGCCACATTTGGGAGAAGAATAAGATTGAATCATGATCTGGATTTGCAAAATAACTTG

ATGGAAAAAAGTTATGATTTCAAGACACTGGTTTTACCAGAAATAAAATTATCAGAACTATCAAGGGAGATACTAAAAGAG
AATGGGTTTGTTATATCTGGAGAGAATCTAAAAATGGATAGGTCTGATGAAGAATTTGAAGGTCTTGCCAGTTTTAATGTG
TTGAGGCTAGATGAGGAAGAAATGTATGAAGGTTTGATCAAAGAAATGAAGATTAAAAGAAAAAAGAAAGGGTCTTTATT
CCCGGCAAATACACTTTTACTAAGTGAGTTGATAAAGTTCTTGATTGGAGGGATAAAGGGAACCAGTTTTGATATAGAGAC
ATTGTTGCGGAATAGTTTTAGACCAGACATATTTTCAACTGACAGACTGGGCAGATTAAGTTCCAGTGTACCTGCACTCAA
AGTTTATGCAACTGTTTATATGGAATATAAGAATGTCAATTGTCCTTTAAATGAGATAGCTGACAGCTTAGAAGGTTATCTA
AAACTGACAAAAAGCAAGTCTAAGGAGCATTTCTTATCTGGAAGAGTTAAGAAAGCTTTGATACAATTAAGAGATGAACA
ATCGCGAACTAAAAAACTAGAGGTCTATAAAGATATCGCAAATTTCCTTTCTAGGCACCCATTATGCTTATCAGAAAAAACA
TTGTATGGAAGATACACTTACTCTGATATCAATGATTATATCATGCAAACAAGAGAGATTATTTTGAGTAAAATAAGTGAGT
TGGATGAGGTTGTTGAAACAGATGAAGACAATTTCTTGCTTAGTTATCTAAGAGGGGAAGAAGATGCCTTTGACGAAGAT
GATTCTGATGAAGAAGAAGACACGGATTAA

Opimized RdRp gene sequence

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. S9. Wild-type and optimized GP gene sequence used in the study.

Wild type GP gene sequence

ATGAGAATTTTAAAACTACTAGAACTAGTGGTAAAAGTGAGTCTTTTCACAATTGCCCTGAGTTCTGTCTTGTTGGCATTC
TTGATCTTCAGAGCCACAGATGCCAAAGTAGAAATAATTCGTGGAGATCATCCTGAGGTTTATGATGATTCTGCTGAGAAT
GAAGTACCCACTGCTGCATCGATTCAACGCAAGGCTATCTTGGAGACTTTAACTAGTCTAATGCTAGAATCTCAGACTCCT
GGAACCCGTCAGATACGAGAAGAAGAATCAACCATCCCTATTTTTGCTGGGTCAACTACGCAAAAAATAATCTCTGTCTCG
GATCTTCCTAACAACTGCTTGAATGCTTCTTCATTAAAATGCGAGATAAAAGGGATATCCACTTATAATGTTTATTATCAAG
TGGAAAATAATGGTGTCATATATTCCTGTGTTTCTGATTCAGCAGAAGGGTTAGAAAAATGTGACAATTCTTTAAATTTGCC
AAAGAGATTCTCCAAAGTCCCGGTTATTCCCATTACCAAACTTGACAACAAAAGGCATTTTTCAGTTGGAACAAAATTCTT
CATTTCAGAAAGCCTGACACAAGATAATTATCCTATAACTTATAACTCGTACCCTACTAATGGAACAGTATCATTACAAACC
GTAAAGTTATCTGGAGACTGCAAAATAACTAAATCAAACTTTGCCAATCCCTATACTGTTAGCATCACTAGCCCCGAGAAG
ATCATGGGTTATTTGATAAAAAAACCTGGTGAAAATGCGGAACACAAGGTCATATCTTTTTCTGGATCAGCAAGCATCACT
TTTACTGAAGAAATGTTGGATGGTGAGCACAATCTCTTGTGCGGTGACAAATCAGCTAAAATACCAAAAACAAACAAAAG
AGTTAGAGATTGCATAATCAAATATTCAAAGAGCATTTATAAGCAAACAGCCTGCATCAATTTCTCTTGGATAAGATTAATA
TTGATAGCTTTGTTGATCTATTTCCCTATCCGGTGGTTAGTAAACAAGACAACTAAACCTCTCTTTCTCTGGTATGATCTTA
TAGGCTTGATCACATACCCTATTTTGTTGCTCATAAATTGCTTATGGAAATATTTTCCATTCAAATGTTCTAACTGTGGCAAT
TTGTGCATAATCACACATGAGTGTACTAAAATTTGCATCTGCAACAAAAGCAAAGCCTCAAAAGAACACTCTTCAGAGTGT
CCCATACTGTCCAAAGAAACAGATCATGACTACAACAAACATAAGTGGACTAGCATGGAATGGTTCCATCTAATAGTGAAC
ACTAAGCTAAGCTTTAGTTTGCTGAAATTTGTGACTGAAATCTTAATAGGTTTGGTCATTTTGTCTCAAATGCCCATGTCTA
TGGCTCAAACTACTCAATGTTTGAGTGGATGTTTTTATGTTCCAGGCTGTCCAGTTTTGGTTACAAGCAAATTTGAAAAAT
GCCCTGAAAAAGATCAATGTTACTGTAATGTAAAAGAAGATAAGATTATAGAAAGTATCTTTGGCACTAACATTGTTATAGA
AGGTCCTAATGATTGCATAGAGAACCAGAATTGTGTTGCACACCCATCTATTGATAATCTTATAAAATGTAGATTAGGTTGC
GAATACTTAGATTTATTTCGCAACAAACCTTTGTATAATGGGTTTTCAGATTATACAGGAAGCTCTTTGGGGTTAACATCGG
TTGGTCTGTATGAGGCTAAGAGATTGAGGAATGGTATAATAGATTCCTATAACCGTACAGACAAGATTTCCGGAATGATTG
CAGGAGATTCTCTAGACAAAAATGAAACAAGCATACCAGAGAACATTCTGCCTAGACAATCATTGATCTTTGATTCTGTTG
TGGATGGGAAATATAGATATATGGTAGAACAATCTCTTTTAGGAGGAGGAGGGACTGTATTCATGTTAAATGATAAGACCT
CAGAAAAAGCAAAAAAATTCGTGATTTATATCAAAAGTGTGGGAATTCATTATGAAGTGTCTGAAAAATACACAACAGCTC
CTATCCAAAGCACTCACACAGATTTTTATTCCACTTGTACAGGAAACTGTGACACTTGCAGAAAAAATCAAGCTTTAACAG
GTTTCCAAGATTTTTGTATAACACCAACTTCTTATTGGGGATGTGAAGAAGCTTGGTGTTTTGCAATTAATGAAGGTGCTA
CATGCGGGTTCTGTCGAAATATTTATGACATGGATAAATCATATAGAATTTATTCAGTGCTAAAATCAACTATAGTGGCAGA
TGTTTGCATTTCTGGTATTTTAGGAGGTCAATGCTCAAGGATTACTGAAGAGGTTCCTTATGAAAATGCATTGTTTCAAGC
TGATATACAAGCAGATCTGCACAATGATGGTATCACTATAGGTGAACTGATAGCTCATGGACCTGACAGCCATATTTATTCT
GGAAATATTGCAAACTTGAATGATCCTGTCAAAATGTTTGGTCATCCACAATTGACTCATGATGGAGTGCCTATTTTTACTA
AGAAAACTCTAGAAGGAGATGACATGTCTTGGGATTGTGCAGCAATAGGGAAAAAATCAATCACTATTAAAACATGCGGA
TATGACACATACAGATTTAGATCTGGTTTAGAGCAAATATCAGATATTCCCATTAGTTTCAAAGATTTCTCTAGTTTTTTCTT
GGAAAAATCTTTTAGTTTAGGGAAACTCAAGATTGTCGTTGATCTTCCATCTGATCTTTTTAAAGTTGCTCCTAAGAGACC
TTCCATAACTTCGACAAGATTGAATTGCAATGGCTGTCTTCTATGCGGTCAAGGTTTATCTTGCATTTTGGAATTTTTCTCA
GATTTGACATTTTCTACTGCAATTTCTATAGATGCTTGCTCTCTATCTACTTATCAGCTGGCTGTTAAAAAAGGGTCTAACA
AATACAACATAACAATGTTTTGTTCTGCTAATCCGGATAAGAAGAAAATGACATTGTATCCAGAAGGCAATCCGGATATTTC
TGTGGAAGTTCTGGTCAATAATGTTATTGTAGAAGAACCGGAGAACATAATAGATCAAAATGATGAGTATGCTCATGAAGA
ACAACAATATAATTCTGATTCCTCAGCATGGGGCTTCTGGGATTATATTAAAAGCCCATTCAATTTCATTGCAAGTTACTTT
GGCTCATTTTTTGATACTATCAGAGTGATATTGCTTATTGCATTCATTTTCCTTGTGATTTATTTCTGTTCTATTCTAACAAC
AATTTGTAAAGGATATGTCAAGAATAAATCTTATAAATCTAGATCCAAGATAGAGGATGATGATGATTCTGAGATCAAAGCC
CCTATGTTAATGAAAGATACAATGACAAGACGAAGGCCACCTATGGATTTCTCTCACCTTGTCTGA

Optimized GP gene sequence

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 wild-type 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:** GACTCTAGATTATTTTGATCCTGAAGCATATG	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 $SR_{(+)eGFP}$
		R: GAATAAAATTATGACACAATTGCTCT GAACCCGGGATACCGGGTTTTCG	
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 AGTCTGCTTGTCACAGGCTAAATTTGAT TC	To amplify the pCB301 backbone for construction of $MR_{(-)mCherry}$
pCB301-HH-$M_{(-)eGFP\&NSmMut}$-RZ-NOS	$MR_{(-)eGFP\&NSmMut}$	**F:** CTCTACCTTAGGCTGTTGAACTCAA AATGTAGACTCTTTTCGGTAATAAGG **R:** GCATGGACGAGCTGTACAAGTAAGT CTGCTTGTCACAGGCTAAATTTGATTC	To amplify the NSm^{Mut} for construction of $MR_{(-)eGFP\&NSmMut}$
		F: GAATCAAATTTAGCCTGTGACAAGC AGACTTACTTGTACAGCTCGTCCATGC **R:** CCTTATTACCGAAAAGAGTCTACAT TTTGAGTTCAACAGCCTAAGGTAGAG	To amplify the pCB301 backbone for construction of $MR_{(-)eGFP\&NSmMut}$
pCB301-HH-$L_{(-)}$-RZ-NOS	$L_{(-)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATCAGGTACAACTAAAAC **R:** GGTGGAGATGCCATGCCGACCCAGA GCAATCAGGTAACAACGAT	To amplify the TSWV genomic L-RNA for construction of $L_{(-)}$
		F: ATCGTTGTTACCTGATTGCTCTGGGT CGGCATGGCATCTCCACC **R:** GTTTTAGTTGTACCTGATTGCTCTGA ACCCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $L_{(-)}$
pCB301-HH-$L_{(+)}$-RZ-NOS	$L_{(+)}$	**F:** CGAAAACCCGGTATCCCGGGTTCAG AGCAATCAGGTAACAACGAT **R:** GTGGAGATGCCATGCCGACCCAGAG CAATCAGGTACAACTAAAAC	To amplify the TSWV antigenomic L-RNA for construction of $L_{(+)}$

		F: GTTTTAGTTGTACCTGATTGCTCTGG GTCGGCATGGCATCTCCAC **R:** ATCGTTGTTACCTGATTGCTCTGAAC CCGGGATACCGGGTTTTCG	To amplify the pCB301 backbone for construction of $L_{(+)}$
pCB301-HH-$L_{(+)opt}$-RZ-NOS	$L_{(+)opt}$	**F:** ATCAGGTAACAACGATTTTAAGCAA ACATGAACATTCAGAAGATCCAAAAGC TG **R:** CATGCATTGTTAGGCATTACTTTTAA TCTAATCGGTGTCCTCTTCCTCATCAG	To amplify the RdRp-optimized for construction of $L_{(+)opt}$
		F: CTGATGAGGAAGAGGACACCGATT AGATTAAAAGTAATGCCTAACAATGCA TG **R:** CAGCTTTTGGATCTTCTGAATGTTCA TGTTTGCTTAAAATCGTTGTTACCTGAT	To amplify the pCB301 backbone for construction of $L_{(+)opt}$
pCB301-HH-$M_{(-)opt}$-RZ-NOS	$M_{(-)opt}$	**F:** ACCATTATAATCTGAGCAGACGTAT AAGATGAGGATCCTGAAGCTTCTTG **R:** GAATCAAATTTAGCCTGTGACAAGC AGACCTAAACAAGATGAGAGAAATC	To amplify the GP-optimized for construction of $M_{(-)opt}$
		F: GATTTCTCTCATCTTGTTTAGGTCTG CTTGTCACAGGCTAAATTTGATTC **R:** CAAGAAGCTTCAGGATCCTCATCTT ATACGTCTGCTCAGATTATAATGGT	To amplify the pCB301 backbone for construction of $M_{(-)opt}$
pGEM-NSs	-	**F:** GTTAATACTAACGGAGTGAAAC **R:** GATTGAAATTTGGCTTGAAACAGTA C	To amplify the sense-NSs for construction of pGEM-NSs to generate the DIG-labeled probes of S agRNA in Northern blot
pGEM-anti-NSs	-	**F:** GATTGAAATTTGGCTTGAAACAGTA C **R:** GTTAATACTAACGGAGTGAAAC	To amplify the antisense-NSs for construction of pGEM-anti-NSs to generate the DIG-labeled probes of S gRNA in Northern blot
pGEM-NSm	-	**F:** GCTTTGACTAAAGCTATGGATAC **R:** TCTTGTATTCTTGGCTGCACATC	To amplify the sense-NSm for construction of pGEM-NSm to generate the DIG-labeled probes of M agRNA in Northern blot
pGEM-anti-NSm	-	**F:** TCTTGTATTCTTGGCTGCACATC **R:** GCTTTGACTAAAGCTATGGATAC	To amplify the antisense-NSm for construction of pGEM-anti-NSm to generate the DIG-labeled probes of M gRNA in Northern blot
pGEM-L 5′ UTR	-	**F:** AGAGCAATCAGGTACAACTAAAAC **R:** AAGTAATGCCTAACAATGCATGA	To amplify the L 5′ UTR for construction of pGEM-L 5′ UTR to generate the DIG-labeled probes of L agRNA in Northern blot
pGEM-anti-L 5′ UTR	-	**F:** AAGTAATGCCTAACAATGCATGA	To amplify the antisense-L 5'UTR for

		R: AGAGCAATCAGGTACAACTAAAAC	construction of pGEM-anti-L 5'UTR to generate the DIG-labeled probes of L gRNA in Northern blot
pGEM-eGFP	-	**F:** ATGGTGAGCAAGGGCGAGGAGCTGTTC **R:** TTACTTGTACAGCTCGTCCATGCCGAGA	To amplify the sense-eGFP for construction of pGEM-eGFP to generate the DIG-labeled probes of antisense-eGFP RNA in Northern blot
pGEM-anti-eGFP	-	**F:** TTACTTGTACAGCTCGTCCATGCCGAGA **R:** ATGGTGAGCAAGGGCGAGGAGCTGTTC	To amplify the antisense-eGFP for construction of pGEM-anti-eGFP to generate the DIG-labeled probes of sense-eGFP RNA in Northern blot
-	-	**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:** GAATCAAATTTAGCCTGTGACAAGCAGACCTAAACAAGATGAGAGAAATC **R:** GGTGGAGATGCCATGCCGACCCAGAGCAATCAGTGCAAACAAAAAC	To amplify the of $M_{(-)opt}$ genomic RNA by RT-PCR
-	-	**F:** GATCAAGGATGTTAATTTCAGCATGCTTATCCCGATCCTCGAC **R:** GAATCAAATTTAGCCTGTGACAAGCAGACCTAAACAAGATGAGAGAAATC	To amplify the of $L_{(+)opt}$ antigenomic RNA by RT-PCR

References

1. Fang XD, *et al.* (2019) A cytorhabdovirus phosphoprotein forms mobile inclusions trafficking on the actin/ER network for viral RNA synthesis. *J Exp Bot* 70(15): 4049-4062.
2. Herold J & Andino R (2000) Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis. *J Virol* 74(14):6394-6400.
3. Shen Y, *et al.* (2014) A versatile complementation assay for cell-to-cell and long distance movements by cucumber mosaic virus based agro-infiltration. *Virus Res* 190:25-33.
4. Jiang L, *et al.* (2017) Occurrence and diversity of tomato spotted wilt virus isolates breaking the Tsw resistance gene of Capsicum chinense in Yunnan, southwest China. *Plant Pathol* 66(6):980-989.
5. Voinnet O & Baulcombe DC (1997) Systemic signalling in gene silencing. *Nature* 389(6651):553.
6. Wang Q, *et al.* (2015) Rescue of a plant negative-strand RNA virus from cloned cDNA: insights into enveloped plant virus movement and morphogenesis. *PLoS Pathog* 11(10):e1005223.
7. 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.
8. 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.
9. Feng MF, *et al.* (2018) Identification of strawberry vein banding virus encoded P6 as an RNA silencing suppressor. *Virology* 520:103-110.
10. 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.