YMTHE, Volume *31*

Supplemental Information

A *trans*-amplifying RNA simplified to essential elements is highly replicative and robustly immunogenic in mice

Mario Perkovic, Stefanie Gawletta, Tina Hempel, Silke Brill, Evelin Nett, Ugur Sahin, and Tim Beissert

Supplemental figures and legends.

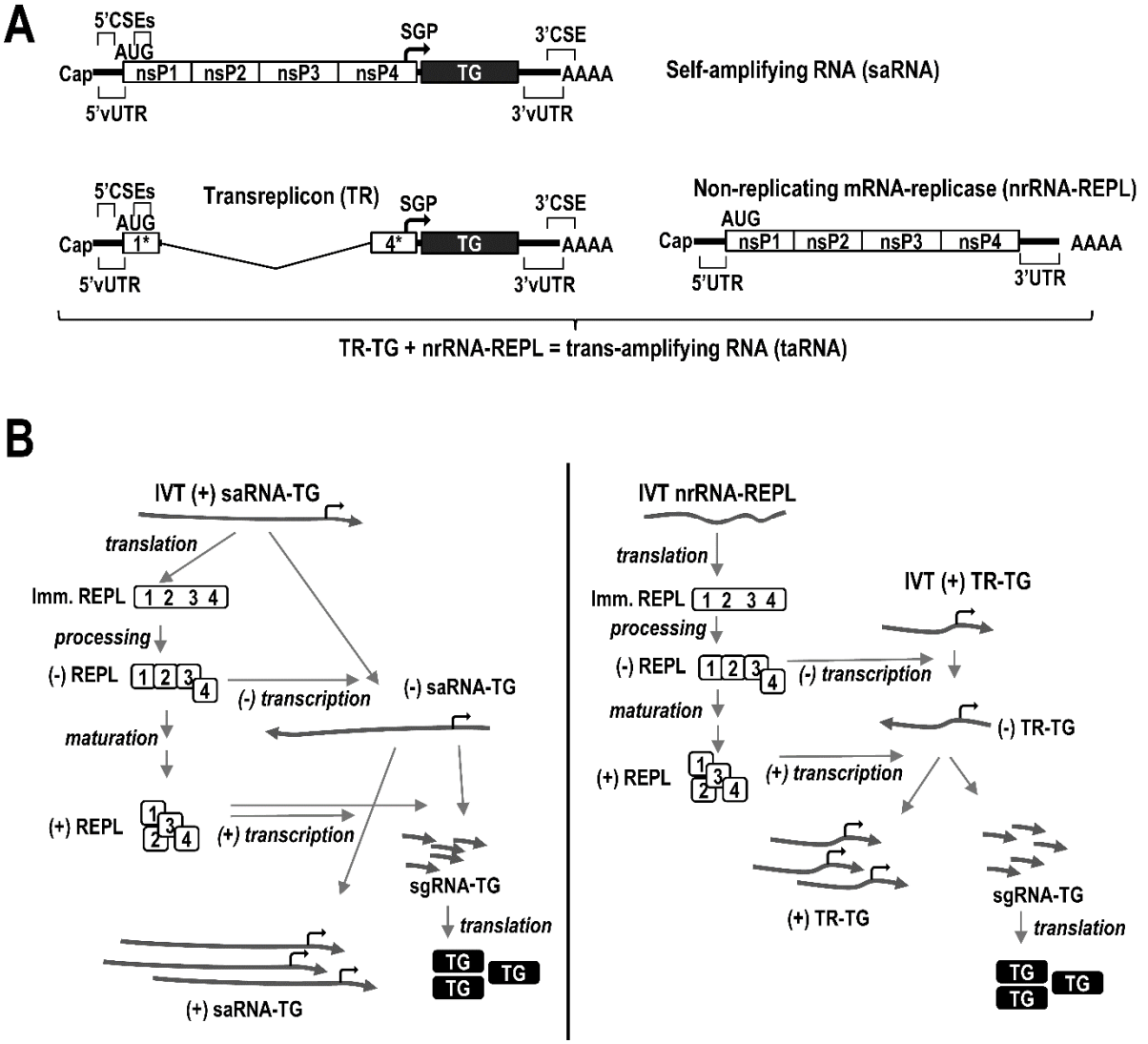

Figure S1: Self- and trans-amplifying RNA compared. (A) Structure of saRNA and taRNA. Self-amplifying RNA (saRNA) contains the following elements: a 5'cap, a 5'viral untranslated region (5'vUTR), the replicase gene composed of non-structural proteins 1 to 4 (nsP1 – 4), a subgenomic promoter (SGP), a transgene (TG), a 3'vUTR and a poly-A tail. Conserved sequence elements (CSEs) serving as replicase recognition signals are located within the 5'vUTR, at the beginning of nsP1, at the end of nsP4 (equivalent to SGP), and at the end of the 3'vUTR extending into the poly-A tail. Trans-amplifying RNA (taRNA) comprises 2 separate RNA molecules. The transreplicon (TR) is engineered from saRNA by removing largely the replicase gene except for the 5'end of the nsP1 gene containing the CSE (1*), and the 3'end of the nsP4 gene comprising the SGP (4*). The full replicase gene is encoded by a non-replicating mRNA (nrRNA-REPL) and flanked by non-viral UTRs that do not interact with replicase and thus do not support replication. **(B) saRNA and taRNA replication.** (left) A positive-sensed *in vitro* transcribed (IVT) saRNA is transfected into cells. First, the replicase gene is translated and forms the immature replicase (imm. REPL) which is rapidly cleaved by the nsP2-encoded protease at the nsP3/4 boundary. A complex formed by nsP4 and uncleaved nsP123 acts as negative-strand specific replicase ((-)REPL), which transcribes a negative-stranded copy of the transfected IVT saRNA. Upon further proteolytic cleavage the fully cleaved replicase eventually acts as positive-strand specific replicase ((+)REPL). The (+)REPL generates not only new full-length copies of the saRNA (novel genomic RNA), but also recognizes the SGP on the negative-stranded saRNA and generates a subgenomic RNA encoding the TG (sgRNA-TG)(for reviews of alphaviral replication please refer to literature [1; 2]). (right) The replication of taRNA differs from the replication of saRNA with respect to the RNA usage. The IVT nrRNA-REPL is only a template for replicase translation, but itself does not replicate, whereas the TR is the template for transcription by the replicase, and later for the translation of the TG. The maturation of replicase is identical to saRNA, and as for saRNA, the replicase first generates a negative-stranded full-length copy of the TR, followed by novel positive-stranded full-length copies (genomic-equivalents of the TR), as well as subgenomic transcripts. As for saRNA, the TG is translated from the sgRNA-TG.

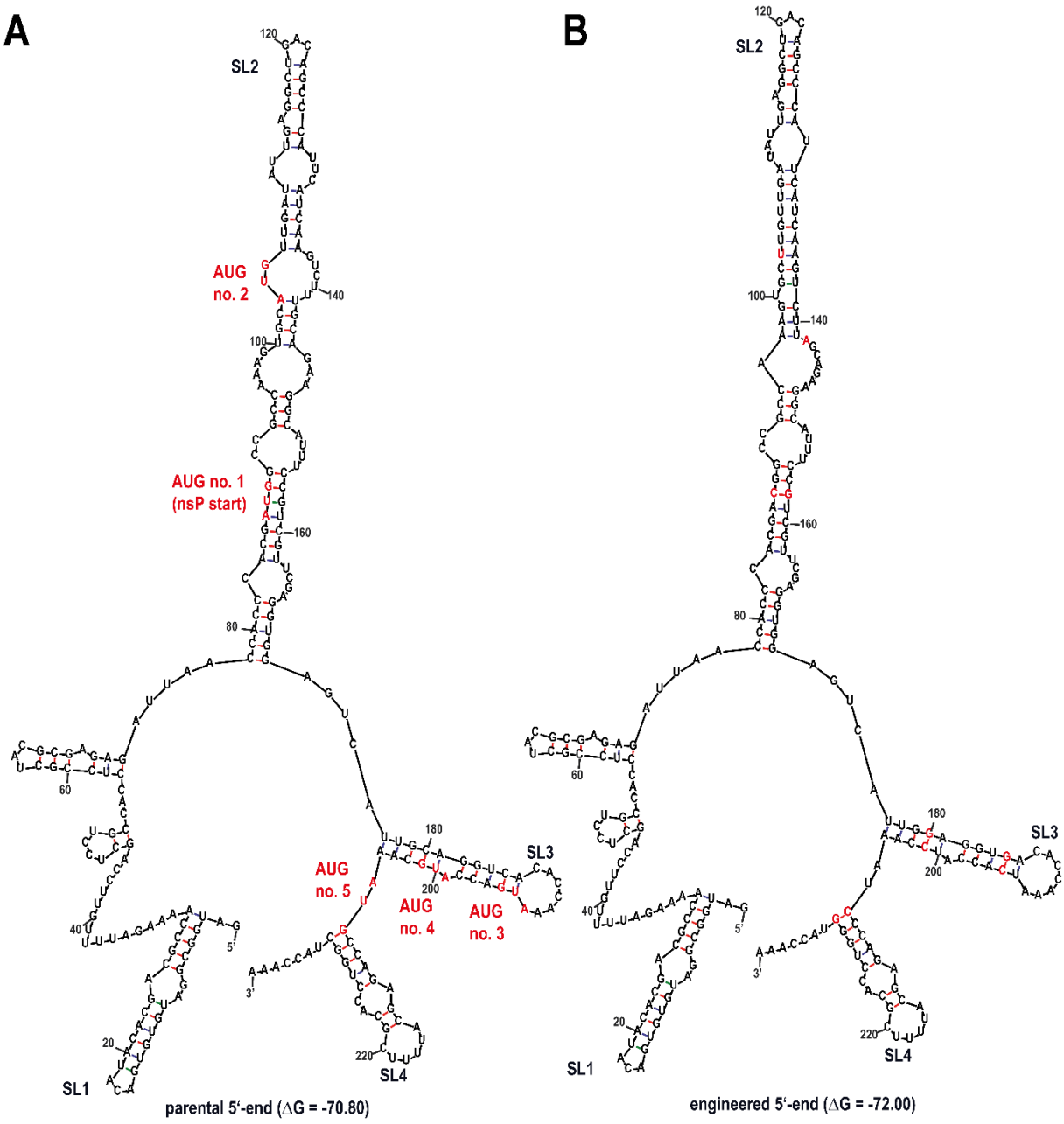

Figure S2. Redesign of the 5'UTR and 5'CSE region in STRs. (A) The predicted structure of the 5'terminal 236 nucleotides of the original SFV saRNA 5'end, also included in the first version of TRs. AUG triplets are highlighted in red. The numbering of stem loop structures (SL) were taken from literature.[3] SL3 and SL4 together are also known as "51 nucleotides CSE" which is critical for RNA replication and are located within the nsP1 coding region. **(B)** After removing the SGP we redesigned the 5'end of the RNA. AUG triplets were mutated in order to prevent translation initiation at the nsP1 start or any other AUG upstream of inserted transgenes, and a few compensating mutations were introduced to conserve the overall folding of critical RNA regions. The nucleotide changes are highlighted red within the predicted structure of the engineered 236 nucleotides region. We used mFOLD version 2.3 applying the default settings, except for setting the maximum distance between paired bases to 100 to avoid 5'-3' end pairing (http://www.unafold.org/mfold/applications/rna-folding-form-v2.php, accessed January 26th, 2022).

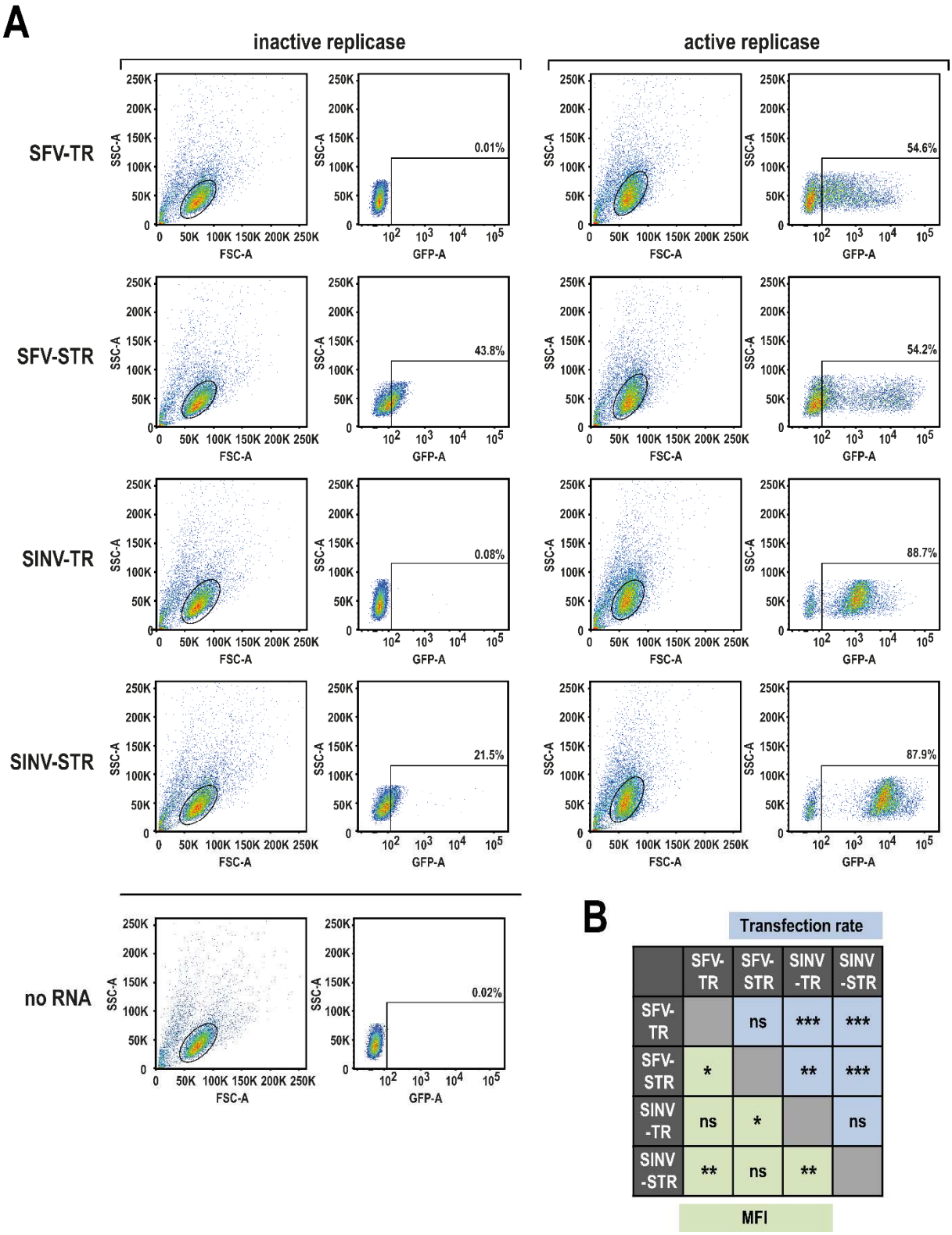

Figure S3. STRs improve taRNA performance compared to TRs. (A) Representative dot plots. Representative GFP expression of cells from one of the 3 experiments summarized in Figure 1C is shown. K562 cells were co-electroporated with 1 µg nrRNA encoding either inactive or active SFV-replicase, and 0.2 µg TRs or STRs of SFV or SINV. **(B) Statistical analysis.** Unpaired Student's t-test was applied to determine if the differences in transfection rates (shaded blue) and MFI (shaded green) of the taRNA transfected groups with active replicase are statistically significant. (ns: not significant; * $p<0.05$; **$p<0.01$; ***$p<0.001$)

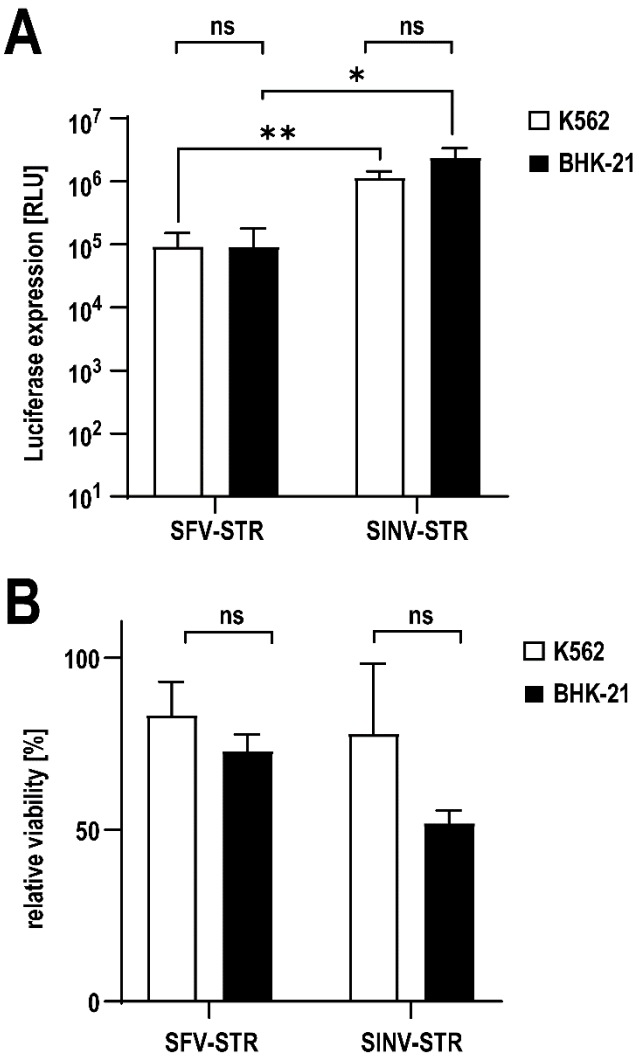

Figure S4: taRNA performance in K562 and BHK21 cells is comparable. K562 and BHK21 cells were co-electroporated with 4 µg capped SFV-replicase nrRNA and 1.8 µg uncapped SFV- or SINV-STR encoding firefly luciferase. (A) Luciferase expression and (B) viability were determined 24h after electroporation. Data shown as mean + SD of 3 independent experiments; unpaired Student's t-test was performed (ns: not significant; * $p<0.05$; **$p<0.01$; ***$p<0.001$). Background luciferase signals detected with untransfected cells range from 10 – 20 RLU.

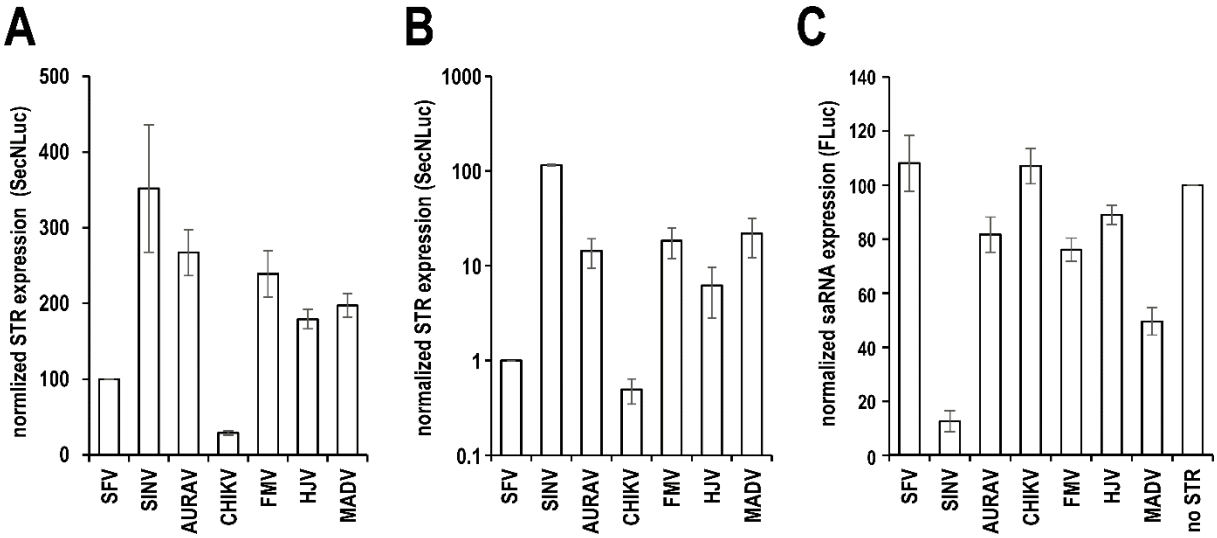

Figure S5: SFV replicase amplifies many heterologous STRs more efficiently than the homologous SFV-STR. (A) Template cross-utilization by SFV-replicase. K562 cells were co-lipofected in 96-well plates with 80 ng capped SFV-replicase nrRNA and 20 ng of the indicated alphaviral GFP-SecNLuc-coding uncapped STRs per well. SecNLuc expression was measured the next day and normalized to the expression of the homologous SFV-derived STR. **(B, C) STR and saRNA co-amplification.** K562 cells were co-lipofected with the collection of uncapped STRs encoding SecNLuc and capped SFV-saRNA expressing firefly luciferase. SecNLuc expression of the STRs **(B)** and firefly luciferase expression of the SFV-saRNA **(C)** was measured the next day. To balance inter-experimental variation, saRNA expression was normalized to the control sample without STR (no STR, C). Data are shown as mean and SD of 3 independent experiments (abbreviations: SFV: Semliki forest virus; SINV: Sindbis virus; AURAV: Aura virus; CHIKV: Chikungunya virus; FMV: Fort Morgan virus; HJV: Highlands J virus; MADV: Madariaga virus).

Table S1. Sequences.

Construct	5'untranslated region including 5'CSE (*; **) preceded by T7 promoter.	3'untranslated region incl. 3'CSE***
SFV-TR	TAATACGACTCACTATA**G**ATGGCGGATGTGTGACATACACGACGCCAAAA GATTTTGTTCCAGCTCCTGCCACCTCCGCTACGCGAGAGATTAACCACCC ACGATGGCCGCCAAAGTGCATGTTGATATTGAGGCTGACAGCCCATTCAT CAAGTCTTTGCAGAAGGCATTTCCGTCGTTCGAGGTGGAGTCATTGCAGG TCACACCAAATGACCATGCAAATGCCAGAGCATTTTCGCACCTGGCTACC AAATTGATCGAGCAGGAGACTGACAAAGACACACTCATCTTGGATATCGG CAGTGCGCCTTCCAGGAGAATGATGTCGACATGAAACGAGATGTCAAAGT CACTCCAGGGACGAAACACACAGAGGAAAGACCCAAAGTCCAGGTAATT CAAGCAGCGGAGCCATTGGCGACCGCTTACCTGTGCGGCATCCACAGGG AATTAGTAAGGAGACTAAATGCTGTGTTACGCCCTAACGTGCACACATTG TTTGATATGTCGGCCGAAGACTTTGACGCGATCATCGCCTCTCACTTCCA CCCAGGAGACCCGGTTCTAGAGACGGACATTGCATCATTCGACAAAAGC CAGGACGACTCCTTGGCTCTTACAGGTTTAATGATCCTCGAAGATCTAGG GGTGGATCAGTACCTGCTGGACTTGATCGAGGCAGCCTTTGGGGAAATA TCCAGCTGTCACCTACCAACTGGCACGCGCTTCAAGTTCGGAGCTATGAT GAAATCGGGCATGTTTCTGACTTTGTTTATTAACACTGTTTTGAACATCAC CATAGCAAGCAGGGTACTGGAGCAGAGACTCACTGACTCCGCCTGTGCG GCCTTCATCGGCGACGACAACATCGTTCACGGAGTGATCTCCGACAAGC TGATGGCGGAGAGGTGCGCGTCGTGGGTCAACATGGAGGTGAAGATCAT TGACGCTGTCATGGGCGAAAAACCCCCATATTTTTGTGGGGGATTCATAG TTTTTGACAGCGTCACACAGACCGCCTGCCGTGTTTCAGACCCACTTAAG CGCCTGTTCAAGTTGGGTAAGCCGCTAACAGCTGAAGACAAGCAGGACG AAGACAGGCGACGAGCACTGAGTGACGAGGTTAGCAAGTGGTTCCGGAC AGGCTTGGGGGCCGAACTGGAGGTGGCACTAACATCTAGGTATGAGGTA GAGGGCTGCAAAAGTATCCTCATAGCCATGGCCACCTTGGCGAGGGACA TTAAGGCGTTTAAGAAATTGAGAGGACCTGTTATACACCTCTACGGCGGT CCTAGATTGGTGCGTTAATACACAGAATTCTGATTATAGCGCACTATTATA GCACCACTAGTACGTTTAAACCCGGGTAAGCTTCCACC**ATG**	**TAA**GGGATGGATCCCTCG AGGCGGCCGCAGGAGCT TAATTCGACGAATAATTG GATTTTTATTTTATTTTGC AATTGGTTTTTAATATTTC CA(n)
SINV-TR	TAATACGACTCACTATA**G**GATTGACGGCGTAGTACACACTATTGAATCAAA CAGCCGACCAATTGCACTACCATCACAATGGAGAAGCCAGTAGTAAACGT AGACGTAGACCCCCAGAGTCCGTTTGTCGTGCAACTGCAAAAAAGCTTCC CGCAATTTGAGGTAGTAGCACAGCAGGTCACTCCAAATGACCATGCTAAT GCCAGAGCATTTTCGCATCTGGCCAGTAAACTAATCGAGCTGGAGGTTCC TACCACAGCGACGATCTTGGACATAGGCAGCGCACCGGCTCGTAGAATG ATATCGTCGACAGATCTCAAGACGACGCTATGGCGTTAACCGGTCTGATG ATCTTGGAGGACCTGGGTGTGGATCAACCACTACTCGACTTGATCGAGTG CGCCTTTGGAGAAATATCATCCACCCATCTACCTACGGGTACTCGTTTTAA ATTCGGGGCGATGATGAAATCCGGAATGTTCCTCACACTTTTTGTCAACA CAGTTTTGAATGTCGTTATCGCCAGCAGAGTACTAGATGAGCGGCTTAAA ACGTCCAGATGTGCAGCGTTCATTGGCGACGACAACATCATACATGGAGT AGTATCTGACAAAGAAATGGCTGAGAGGTGCGCCACCTGGCTCAACATG GAGGTTAAGATCATCGACGCAGTCATCGGTGAGAGACCACCTTACTTCTG CGGCGGATTTATCTTGCAAGATTCGGTTACTTCCACAGCGTGCCGCGTGG CGGAACCCCTGAAAAGGCTGTTTAAGTTGGGTAAACCGCTCCCAGCCGA CGACGAGCAAGACGAAGACAGAAGACGCGCTCTGCTAGATGAAACAAAG GCGTGGTTTAGAGTAGGTATAACAGGCACTTTAGCAGTGGCCGTGACGA CCCGGTATGAGGTAGACAATATTACACCTGTCCTACTGGCATTGAGAACT TTTGCCCAGAGCAAAAGAGCATTCCAAGCCATCAGAGGGGAAATAAAGCA TCTCTACGGTGGTCCTAAATAGTCAGCATAGTACATTTCATCTGACTAATA CTACAACACCACCACCGAATTCTGATTATAGCGCACTATTATAGCACCACT AGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC CACGCAGCGTCTGCATAA CTTTTATTATTTCTTTTATT AATCAACAAAATTTTGTTT TTAACATTTCA(n)
SFV-STR	TAATACGACTCACTATA**G**GATGGCGGATGTGTGACATACACGACGCCAAA AGATTTTGTTCCAGCTCCTGCCACCTCCGCTACGCGAGAGATTAACCACC CACGACGGCCGCCAAAGTGCTTGTTGATATTGAGGCTGACAGCCCATTCA TCAAGTCTTAGCAGTAATACGACTCACTATA**G**AAGGCATTTCCGTCGTTCG AGGTGGAGTCATTGGAGGTGACACCAAATCACCATCCAAATCCCAGAGCA	**TAG**CTCGAGGCGGCCGC AGGAGCTTAATTCGACGA ATAATTGGATTTTTATTTT ATTTTGCAATTGGTTTTTA ATATTTCCA(n)

	TTTTCGCACCTGGGTACCAAATTGATCGAGCAGGAGACTGACAAAGACAC ACTCATCTTGGATCTAGTGCCACC**ATG**	
SINV-STR	TAATACGACTCACTATA**G**GATTGACGGCGTAGTACACACTATTGAATCAAA CAGCCGACCAATTGCACTACCATCACAACGGAGAAGCCAGTAGTAAACGT AGACGTAGACCCCCAGAGTCCGTTTGTCGTGCAACTGCAAAAAAGCTTCC CGCAATTTGAGGTAGTAGCACAGGAGGTGACTCCAAATCACCATCCTAAT CCCAGAGCATTTTCGCATCTGGGCAGTAAACTAATCGAGCTGGAGGTTCC TACCACAGCGACGATCTTGGACATAGGCAGCGCACCGGCTCGTAGAACG ATAAACCCCTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC CACGCAGCGTCTGCATAA CTTTTATTATTTCTTTTATT AATCAACAAAATTTTGTTT TTAACATTTCA(n)
AURAV-STR	TAATACGACTCACTATA**G**GATAGCGGACGGACTAGTACTTGTACTACAGA ATTAACTGCCGTGTGCCGCCCGCTAAACTAGCCCCAATCATCGAAAATCG AGAAACCGACAGTGCACGTTGACGTAGACCCCCAAAGTCCGTTTGTGCTA CAACTGCAGAAGAGTTTCCGACAATTCGAGATTGTGGCTCAGGAGGTGAC TCCGAATCACCATCCTAATCCCAGAGCTTTTTCGCATCTGGGTAGTAAACT GATCGAACATCAGATCCCACTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC TACTACGGTAGTTGAAAA TAACTTTATAGAATTTTAA AATTTTCTTTATTAAAATC TTTTGTTTTTCTTTTATTAT TTCAAAATTTTGTTTTTAA TATTTCA(n)
CHIKV-STR	TAATACGACTCACTATA**G**GATGGCTGCGTGAGACACACGTAGCCTACCAG TTTCTTACTGCTCTACTCTGCAAAGCAAGAGATTAAGAACCCATCTTGGAT CCTGTGTACGTGGACATAGACGCTGACAGCGCCTTTTTGAAGGCCCTGC AACGTGCGTACCCCAAGTTTGAGGTGGAACCTAGGGAGGTGACACCGAA TCACCATCCTAATCCTAGAGCGTTCTCGCATCTAGGTATAAAACTAATAGA GCAGGAAATTGATCCCGACACTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC GCCGAACTCACTTTGAGA TGTAGGCATAGCATACCG AACTCTTCCACGATTCTC CGAACCCACAGGGACGT AGGAGATGTTATTTTGTTT TTAATATTTCA(n)
FMV-STR	TAATACGACTCACTATA**G**GATAGGGTATCGTTTAGAGGCGCGTACCCTAC TTAACCGATCCAAACATCGAGAAACTGCATCTAGACTTAGACGTTGATAG CCCGTTCGTCAAGTCTTTGGAGAAGAGGTTTCCAGAGTTTGAGATCGAAG CAAAGGAGGTGACTGACAATCACCATCCTAATCCCAGAGCGTTCTCGCAT CTGGGTACTAAGCTTATAGAAAGCGAAGTCGACCGGGAACAAGTGATATT GGACATTGGTAGCGCACCGACTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC ACTTAGTTTCAATTTATTT TTCTTACATTTAACTTTAA ACCTTTTATTCTTTATCCT TATTTTATTTAGTCTACTA GATTAGTTTTGTTTTTAAT ATTTCA(n)
HJV-STR	GTAATACGACTCACTATA**G**ATAGGGCATCGTATAGAGGCACGTACCCTAC AAACCGATCCAAACAGGGAAAAAGTTCACGTTGACTTAGACGCCGACAGC CCGTTCGTCAAGTCACTGCAACGGAGCTTTCCCCAGTTTGAGATCGAAGC CAAGGAGGTGACAGACAATCACCATCCTAATCCCAGAGCGTTTTCGCAAG TGGGTACAAAGCTCATAGAGAGCGAAGTCGACCGGGAGCAGATTATACT GGACATAGGAAGTGCACCTGACTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC ATTCAAACACACTACACA CTTATAACAACTTTTAAAA TTTTGATTAGATTACTATA TTTTTCTTTTCTTTATTTTT CTTTTATTTTGTTTTTAAAA TTTCA(n)
MADV-STR	TAATACGACTCACTATA**G**GATAGGGTATCGTGTAGAGGCAACGACCCGAC CTATCCTATCCAAAATCGAGAAAGTACACGTGGACTTAGACGCTGACAGC CCTTACGTCAAGTCACTGCAGAAGTGCTTTCCGGATTTTGAGATAGAAGC CACGGAGGTGACTGACAATCACCATCCTAATCCTAGAGCGTTTTCGCATC TAGGCACCAAACTCATCGAAAGCGAAGTGGACCAAGACCAGGTTATCCTG GATATTGGAAGCGCGCCTGACTAGTGCCACC**ATG**	**TAG**CTCGAGGCGGCCGC TTACAACACTACTGGCGG CGCATAATGCTGCCTTTT ATAAAATCTTTAAAATTCA TATACAATTTTTTCTTTTAT GTTTTTATTTTGTTTTTAAT ATTTCA(n)

* First transcribed nucleotide and start codon of the transgene in bold

** TR sequences translatable to a truncated NsP1-peptide are underlined

*** Stop codon of the transgene in bold

Supplemental References

1 Pietilä, M. K., Hellström, K., and Ahola, T. (2017). Alphavirus polymerase and RNA replication. Virus research 234, 44–57.

2 Ahola, T., McInerney, G., and Merits, A. (2021). Alphavirus RNA replication in vertebrate cells. Advances in virus research 111, 111–156.

3 Frolov, I., Hardy, R., and Rice, C. M. (2001). Cis-acting RNA elements at the 5'end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA (New York, N.Y.) 7, 1638–1651.