The VEGAS platform is not suitable for mammalian directed evolution

Christopher E. Denes[1,4], Alexander J. Cole[2,4], Minh Thuan Nguyen Tran[3], Mohd Khairul Nizam Mohd Khalid[3], Alex W. Hewitt[3], Daniel Hesselson[2,5], G. Gregory Neely[1,5]

[1]The Dr. John and Anne Chong Lab for Functional Genomics, Charles Perkins Centre and School of Life & Environmental Sciences, The University of Sydney, Sydney, NSW, 2006, Australia

[2]Centenary Institute and Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, 2006, Australia

[3]Menzies Institute for Medical Research, School of Medicine, University of Tasmania, Tasmania, 7000, Australia

[4]These authors contributed equally

[5]These authors contributed equally

*Correspondence: d.hesselson@centenary.org.au (D.H.), greg.neely@sydney.edu.au (G.G.N.)

ORCIDs:

C.E.D.: 0000-0002-8669-8088

A.J.C.: 0000-0001-9019-6884

M.T.N.T.: 0000-0002-9554-6449

M.K.N.M.K.: 0000-0001-7200-3102

A.W.H.: 0000-0002-5123-5999

D.H.: 0000-0002-8675-9426

G.G.N.: 0000-0002-1957-9732

Keywords: VEGAS, Sindbis virus, directed evolution, replicability, reproducibility

Abstract

Directed evolution uses cycles of gene diversification and selection to generate proteins with novel properties. While traditionally directed evolution is performed in prokaryotic systems, recently a mammalian directed evolution system (viral evolution of genetically actuating sequences, or "VEGAS") has been described. Here we report that this platform has major technical issues precluding its use for directed evolution. These issues include a rapid loss of VEGAS system integrity, an inability to propagate VEGAS Sindbis particles across rounds of transduction, and widespread prevalence of circuit "cheaters". Similar results have been obtained in independent labs. It may be possible to use Sindbis virus for mammalian directed evolution in the future, however, in its reported form, VEGAS is not suitable for use as a mammalian directed evolution platform.

INTRODUCTION

A new generation of mammalian directed evolution systems have been reported, which exploit viral life cycles to couple mutagenesis of a virally packaged transgene-of-interest with selection for higher fitness variants (Berman et al. 2018; English et al. 2019). In these systems, essential structural components are removed from the viral genome and placed under the control of synthetic gene expression circuits within the host cell and regulated by the transgene-of-interest. Selection operates on natural mutations arising in the transgene during error-prone viral replication that then impact the expression of the transgene-regulated viral structural genes. Variants that more efficiently drive the synthetic circuit are then overrepresented in the packaged viral particles used for further rounds of evolution (Berman et al. 2018; English et al. 2019). The English, et al. (2019) viral evolution of genetically actuating sequences (VEGAS) system offered a simplified platform for simultaneous mutagenesis and selection using Sindbis virus (SINV). The VEGAS system was designed to link the activity of a transgene to expression of the Sindbis Structural Genome (SSG; comprising Capsid, E3, E2, 6K and E1 genes) (English et al. 2019). Thus, to apply the selective pressure required for directed evolution, SINV replication must be dependent on SSG-expression in the mammalian host cell.

Here, we find that the VEGAS system does not result in viral propagation across rounds of replication when performed as described. Investigating why the VEGAS system failed, we found a rapid loss of system integrity during rounds of replication. While SINV efficiently packaged a transgene and transduced cells as previously reported (Shapiro et al. 2010; Fayzulin et al. 2005), we found that propagation beyond this initial transduction event was hijacked by "cheaters" that packaged SSG components, thereby short-circuiting transgene-dependent evolution. We conclude that contrary to its initial description, the VEGAS system is not suitable for mammalian directed evolution.

RESULTS

In our efforts to implement the VEGAS system, we first attempted to reproduce SSG-dependent replication of SINV (Figure 1B from (English et al. 2019)). We electroporated six independent BHK-21 cultures with pSinHelper, pSinCapsid and pTSin-EGFP seed mRNA (**Figure S1A)** to package SSG-deficient SINV-EGFP particles. After 24 h, we detected 4.72 x $10^9 \pm 5.16 \times 10^8$ (SEM) genome copies (gc)/mL in the supernatant (Round (R)0; **Figure 1A**) and observed bright EGFP fluorescence in most packaging cells (R0; **Figure 1B**). Fresh BHK-

21 cultures were transfected with a CMV-SSG plasmid (+SSG) or control DNA (-SSG) and SINV-EGFP particles were added at a calculated multiplicity of infection (MOI) of 1 gc/cell (R1; **Figure 1A**). After 24h, viral titers did not show SSG-dependence (-SSG; 7.35 x 10^6 ± 2.30 x 10^5 gc/mL; +SSG, 7.33 ± 4.77 x 10^5 gc/mL; **Figure 1A**) and GFP fluorescence was rarely observed in the transduced cells (**Figure 1B**). Furthermore, these low titers precluded adding R1 SINV-EGFP particles to R2 cells at an apparent MOI of 1 (see below). Together, these data show that SINV particles produced using the VEGAS protocol are not generated at quantities required to perform further rounds of transduction.

One possible cause for the rapid loss of viral titers observed between R0 and R1 is that transfected host cells failed to express the Sindbis structural genes provided via DNA plasmid. Thus, we next examined host cell expression of the CMV-SSG plasmid used in R1. To avoid the possibility of transfected plasmids contributing to high background signal during qPCR, we developed a barcoding method to discriminate between plasmid DNA (pDNA) and expressed mRNA (**Figure S2A**). Using this approach, we confirmed that the CMV-SSG plasmid was highly expressed (**Figure S2B,C**). We next addressed possible explanations for the low expression of the SINV-EGFP transgene when transduced at an MOI of 1. We suspected that residual seed mRNA in the supernatant following R0 transfection could have been carried over in the R0 output and inflated titers, impeding accurate calculations of MOIs. To test this, we treated packaged R0 virus samples with RNase A to degrade seed mRNA that was not protected by encapsidation. Indeed, RNase A treatment of R0 samples decreased the measured virus titer by 93%, whereas R1 samples, which were not directly exposed to seed mRNA, were much less RNase-sensitive (**Figure S3**). This differential sensitivity indicated that residual packaging RNA was present in the output from R0. To determine whether the inaccurate SINV-EGFP R0 titers were responsible for the lack of SSG-dependent replication, we added neat (undiluted) packaged SINV-EGFP to CMV-SSG-expressing BHK-21 cells. This approach produced 5-10% GFP-positive cells at R1 (arrowheads, **Figure 1D**), demonstrating initial Sindbis packaging was successful, and generated sufficiently high viral titers to test SSG-dependence at R2 (**Figure 1C**). However, at R2 we still did not observe GFP-positive cells (**Figure 1D**), consistent with a dramatic drop in viral titers and lack of SSG-dependence (**Figure 1C**). Overall, initial R0 titer calculations are inflated by contaminating seed mRNA making R1 MOI calculations unreliable. Regardless of this issue, efficient transduction of neat

SINV preparations still leads to unproductive replication, indicating that directed evolution campaigns can not be performed with the VEGAS system in its reported form.

To investigate the integrity of the VEGAS system over successive rounds of transduction we performed long-read nanopore sequencing of viral RNA (**Figure 1E**). Using the approach for transgene recovery as described (English et al. 2019), all reads should map to the transgene subjected to directed evolution. Surprisingly, we found only 53.4% of reads mapped to the EGFP transgene in the initial packaged virus (R0; **Figure 1E**) and transgene inclusion was reduced to 30.3% by R2. The remaining reads mapped to envelope and capsid sequences derived from the pSinHelper and pSinCapsid plasmids. We note that the *in vitro*-transcribed VEGAS pSinHelper and pSinCapsid mRNAs retain the SINV NSP1 packaging signal (Weiss, Geigenmüller-Gnirke, and Schlesinger 1994) (**Figure S4**), which could allow seed mRNA to compete with the SINV-EGFP transgene for packaging. These data show that on application, the VEGAS system rapidly loses integrity and preferentially packages SINV structural elements (cheater particles), further impacting the utility of the VEGAS system as an efficient tool for mammalian directed evolution.

Although replication-competent SINV can retain a GFP transgene for at least six viral passages at low MOIs (0.1 PFU/cell) (Thomas et al. 2003), we next considered the possibility that the VEGAS system requires selective pressure within a circuit to produce sufficient viral titers to propagate past R2 and maintain system integrity. To test this, we first validated a simple circuit whereby a serum response factor DNA-binding domain (SRF)-NLS-VP64 fusion protein activates a serum response element (SRE)-regulated firefly luciferase reporter (SRE_LUC) (**Figure 2A**). In our hands, SRF-NLS-VP64 activated SRE_LUC similarly to the maximal activation of SRE-Luciferase previously reported (Figure S3B from (English et al. 2019)), confirming the integrity of this circuit. Next, we subcloned SRF-NLS-VP64 into the VEGAS pTSin vector and packaged it in a 1:1 ratio with pTSin-EGFP to provide an evolutionarily neutral competitor. Under selection, the EGFP transgene should drop out while the higher fitness SRF-NLS-VP64-containing viruses, capable of driving expression of the Sindbis structural genome, should rapidly dominate the culture. As predicted, the percentage of GFP-positive cells decreased with serial passage (**Figure 2B),** however by R2 we still did not detect SSG-dependent replication (**Figure 2C**). Thus, even this simple VEGAS circuit is not capable

of driving the viral replication required for a directed evolution campaign. Importantly, the circuit-inducing SRF-NLS-VP64 transgene did not outcompete the control EGFP transgene (3.2% SRF-NLS-VP64 vs 21.7% GFP by R2), and both were rapidly outcompeted by "cheater" sequences containing the Sindbis structural elements (75.1%, **Figure 2D**). Together, these data show that even under selective pressure, simple VEGAS feedback circuit components do not show a selective advantage nor lead to the productive rounds of replication required for directed evolution campaigns. Instead, these strategies rapidly become "short-circuited" by cheater particles containing Sindbis structural genes.

DISCUSSION

While it is possible that undocumented methodological subtleties could explain why VEGAS Sindbis particles fail to propagate across rounds of transduction, we have thus far not identified any that could account for the loss of VEGAS system integrity. To generate the Sindbis particles used to initiate VEGAS campaigns, English *et al.* used the NEON electroporator, whereas we used the Amaxa electroporator that has previously been used successfully to deliver RNA to the cytosol (Ekstrom and Dean 2011) and efficiently generate functional Sindbis particles (Shapiro et al. 2010). It is unlikely that using a different electroporator would explain our inability to run VEGAS directed evolution campaigns, since both machines produced Sindbis particles that transduced R1 target cells with equivalent efficiency (~5-10%). Instead, the VEGAS system is non-functional due to technical limitations at R2 and beyond.

Furthermore, our experiments involved VEGAS components generated and made publicly available by English *et al.* when provided (Addgene plasmids #127692–127695). Of note, we could not obtain the CMV-SSG plasmid from English *et al.* and were required to generate this component ourselves, which we validated and used in Figure 1. Moreover, English *et al.* similarly did not provide the original VEGAS circuits required to reproduce their original work, and thus we generated our own serum response factor circuit which we used in Figure 2. Again, while we confirmed this circuit is functional and could drive expression from the SRE as in English *et al.*, we found this circuit could not support viral propagation past R2. A similar inability to use the VEGAS system as described has been observed independently across our separate laboratories over the 3 years since the platform was first described (English et al. 2019). We have consulted with English et al. but did not receive advice that resolved these

172 critical technical issues. Together, we conclude that the VEGAS system in its published form
173 is not suitable for directed evolution campaigns.

174

175 Importantly, as Sindbis virus is pathogenic to humans and has the potential to cause meningitis,
176 extreme caution should be used when trying to implement the published VEGAS system in a
177 BSL2 environment since a single recombination event between the transgenic SINV genome
178 and the host cell SSG expression plasmid (which share homologous sequences) could produce
179 a replication-competent virus. In our hands we tested for and did not detect competent virus,
180 however the VEGAS system is almost immediately hijacked by "cheaters" that packaged SSG
181 components, making it currently unsuitable for directed evolution. While we believe there is a
182 future for Sindbis-based mammalian directed evolution systems, this will require substantial
183 efforts to address critical design flaws that limit VEGAS utility. Optimally, any repairs or
184 upgrades should be confirmed by independent groups prior to publication.

185

186 In summary, although we confirm that packaging and transduction of SINV particles efficiently
187 delivers transgenes to fresh host cells as previously documented (English et al. 2019; Fayzulin
188 et al. 2005; Shapiro et al. 2010), we conclude that due to fundamental biological limitations
189 (lack of virus propagation across rounds, loss of VEGAS system integrity, and widespread
190 prevalence of circuit "cheaters"), the VEGAS system as published is not suitable for use as a
191 mammalian directed evolution platform.

192

METHODS

Cell Culture

BHK-21 [C-13] cells were purchased from the American Type Culture Collection (#CCL-10). Cells were grown in a humidified 37°C (5% CO_2) atmosphere in MEM α (ThermoFisher, #32571101) supplemented with 5% HyClone fetal bovine serum (FBS) (Cytiva Life Sciences, #SH30084.03, AU origin) and 10% tryptose phosphate broth (TPB) (ThermoFisher, #CM0283B), referred to as BHK-21 Growth Medium. During transduction and recovery, cells were maintained in serum-free MEM α with 10% TPB, referred to as BHK-21 Recovery Medium (Serum-Free).

Molecular Biology and Plasmid Construction

All plasmids were designed in SnapGene® (version 5.3.2). Plasmids were generated by PCR amplification of sequences of interest with Velocity DNA Polymerase (Bioline, #BIO-21099) using primers synthesized by IDT or restriction enzyme digestion with NEB High-Fidelity enzymes. Assembly of amplicons was performed using the NEBuilder HiFi DNA Assembly Master Mix (NEB, #E2621). Assembled products were transformed into NEB® 10-beta Competent *E. coli* (NEB, #C3019) and selected on LB agar plates (ThermoFisher, #22700025) supplemented with 100 µg/ml ampicillin (Sigma-Aldrich, #A9518). Individual colonies grown overnight in liquid LB broth (ThermoFisher, #12795-084) supplemented with 100 µg/ml ampicillin were processed with either the ISOLATE II Plasmid Mini Kit (Bioline, #BIO-52057) for sequence verification, or the PureYield™ Plasmid Maxiprep System (Promega, #A2393) for transfection and *in vitro* transcription applications. Plasmid constructs were verified by restriction digestion and Sanger sequencing at the Australian Genome Research Facility (AGRF). A list of plasmids used and generated in this study is available in Supplementary File 1 – Plasmid List.

mRNA synthesis

mRNA for electroporation and transfection was produced using the mMESSAGE mMACHINE™ SP6 Transcription Kit (ThermoFisher, #AM1340) from XbaI-linearized SP6-driven plasmids as outlined previously (English et al. 2019). mRNA concentrations were calculated using the Qubit™ RNA BR Assay Kit (ThermoFisher, #Q10210). mRNA integrity

was assessed by gel electrophoresis (**Figure S1**). mRNA was frozen at -80°C immediately after transcription as per (English et al. 2019)**.**

Packaging of SINV Particles (referred to as Round 0 (R0))

1 x 10^6 BHK-21 cells were electroporated with a total of 7.8 µg of mRNA (1:1:1 of pSinHelper, pSinCapsid and pTSin-EGFP/pTSin-SRF-NLS-VP64) using Amaxa 2B (Lonza) as per the manufacturer's instructions for BHK-21 cells. The Amaxa electroporator efficiently delivers SINV RNA into the cell cytosol (Ekstrom and Dean 2011), allowing for the packaging of mature SINV particles (Shapiro et al. 2010). Electroporated cells were plated in BHK-21 Recovery Medium (Serum-Free). Virus-containing supernatants were collected 24 hours after mRNA delivery and centrifuged at 500 *g* for 5 minutes to pellet cellular debris. Clarified supernatants were collected for titration and subsequent transduction experiments.

Viral Titration

Viral supernatants were titrated as outlined by English and colleagues (English et al. 2019). Following collection and clarification, undiluted SINV-containing supernatants were combined with the TaqMan™ Fast Virus 1-Step Master Mix (ThermoFisher, #4444434) and the appropriate primer-probe sets in technical triplicates. Plates were run on a QuantStudio™ 7 Flex Real-Time PCR System (ThermoFisher). As described previously, serially diluted *in vitro*-transcribed pTSin mRNA was used as the standard curve (ranging between 10^3-10^7 genome copies (gc) per mL) (English et al. 2019). Standard curves were used to determine viral titers in gc/mL. Primer sequences are listed in Supplementary File 2 – Primer List.

Transduction of Viral Particles (Round 1 (R1)-onwards)

BHK-21 cells were seeded in 6-well plates at 6.5 x 10^4 cells/well, incubated for 24 hours, and transfected with a total of 2.5 µg of SSG-encoding DNA/well using *Trans*IT-2020 Transfection Reagent (Mirus Bio, #MIR5400) following the manufacturer's recommendations. 6 hours post-transfection, cells were rinsed twice with DPBS before viral inoculum (either undiluted or diluted to a calculated MOI of 1 gc/cell) was applied in a 380 µl volume of BHK-21 Recovery Medium (Serum-Free). Cells were incubated with virus for 1 hour and rinsed twice with DPBS before 1.5 mL BHK-21 Recovery Medium (Serum-Free) was added for a further 23 hours of incubation. Virus-containing supernatants were collected and processed for titration and transduction experiments as described in *Packaging of SINV Particles*.

We note that recombination between the SINV genome and homologous sequences in the SSG 3'UTR (**Figure S4)** could produce competent viral populations. While this has not been observed in our independent labs, we urge responsible research practices and approvals for working in appropriate Biosafety Level 2 facilities, and the application of appropriate testing protocols for determining replication competency. Assaying viral titer amplification on cells ± SSG elements in parallel can be used to quickly ascertain structural element dependency status.

RNase Digestion

To test if inflated R0 viral titers were a result of residual packaging mRNA, pooled viral supernatants were digested with 0 or 400 ng RNase A (Macherey-Nagel, #740505.50) for 4 hours at 37°C. Samples were then incubated for 20 minutes at room temperature with 80 units of RNaseOUT (ThermoFisher, #10777019). Samples were titrated as described in *Viral Titration*. RNase A treatment was not applied to any samples used for subsequent transduction assays.

Barcoded qPCR

RNA was extracted using the FavorPrep™ Blood/Cultured Cell Total RNA Mini Kit (Fisher Biotec, #FABRK 001-1) and 400 ng of cDNA was synthesized using the iScript™ Select cDNA Synthesis Kit (Bio-Rad Laboratories, #1708897) on a Mastercycler® nexus PCR Thermal Cycler (Eppendorf): 42°C for 60 min, 85°C for 5 min. cDNA was produced using two primers: a specific primer for hamster GAPDH, and a SINV structural genome primer containing a barcode. To remove residual primers, cDNA was bound to AMPure XP magnetic beads (Beckman Coulter, #A63881) in 0.75 M LiCl, 20% PEG 8000 buffer at a 1:0.9 DNA:buffer volume ratio, and washed twice with 70% ethanol. qPCR was performed using SYBR™ Select Master Mix (ThermoFisher, #4472908) following manufacturer's recommendations in triplicate on a QuantStudio™ 7 Flex Real-Time PCR System (ThermoFisher). Primer sequences are listed in Supplementary File 2 – Primer List.

Luciferase Assay

BHK-21 cells were seeded in 96-well plates at 2200 cells/well. Cells were transfected with a total 90 ng of plasmid constructs using *Trans*IT-2020 Transfection Reagent following the manufacturer's recommendations. After 6 hours, cells were washed twice with DPBS and switched to BHK-21 Recovery Medium. At 24 hours post-transfection, luciferase activity was

assessed using the Dual-Glo Luciferase Assay System (Promega, #E2940) following manufacturer's recommendations in black-bottom plates. Both Renilla and firefly luminescence were measured on an Infinite M1000 PRO microplate reader (Tecan). Raw firefly luminescence values were normalized to Renilla luminescence and log2-transformed.

Transgene Isolation

Sextuplicate viral supernatants from each round of viral replication were pooled in equal ratios. Viral RNA from 400 µl aliquots was isolated with the MagMAX™ Viral RNA Isolation Kit (ThermoFisher, #AM1939) as per the manufacturer's recommendations. Transgene sequences, situated between nsP4 and the viral 3' untranslated region (UTR) were reverse-transcribed and PCR-amplified using the SuperScript™ IV One-Step RT-PCR System (ThermoFisher, #12594025) and the 26S-F primer and pooled SinRev primer. Primer sequences are listed in Supplementary File 2 – Primer List. Amplicons of the appropriate size range for the transgene of interest were then gel extracted for nanopore sequencing using the ISOLATE II PCR and Gel Kit (Bioline, #BIO-52060) (**Figure S5**).

Nanopore Sequencing

Sample Processing

Isolated transgene DNA was processed to generate libraries for full-length transgene sequencing using Oxford Nanopore Technologies (ONT) Flongle flow cells (ONT, FLO-FLG001). Samples were prepared for sequencing according to the ONT protocol 'Amplicons by Ligation' (version ACDE_9064_V109_REVP_14AUG2019). 200 fmol of DNA (determined following quantification with the Qubit™ dsDNA HS Assay Kit) was processed using the NEBNext Companion Module for ONT Ligation Sequencing (NEB, #E7180). Sequencing adapters were added using the Ligation Sequencing Kit (ONT, #SQK-LSK109).

Sequencing and Basecalling

Up to 40 fmol of DNA library was loaded onto ONT Flongle flow cells (R9.4.1) in a MinION Mk1B Sequencer fitted with a Flongle Adapter (ONT). Sequencing was performed using MinKNOW (ONT, version 4.2.8) under default parameters and the following specified inputs: kit used, SQK-LSK109; 0.5 hours between MUX scans; basecalling, disabled. A minimum of 120,000 raw reads were obtained for each sample. Raw FAST5 files were processed using

Guppy (version 4.5.2) with the minimum q-score set to 7.0. Raw FAST5 and basecalled FASTQ reads were deposited at the European Nucleotide Archive (PRJEB47639).

Alignment

Quality-filtered base-called reads (in FASTQ format) were processed using EPI2ME Desktop Agent (ONT, version 3.3.0.1031). Reference files, consisting of complete transgene-coding sequences or helper- and capsid-coding sequences, were uploaded using the Fasta Reference Upload workflow (v2021.07.15). For each sample, 100,000 reads were aligned to reference files using the Fastq Custom Alignment workflow (v2021.03.25) using default parameters.

Microscopy

Phase contrast and EGFP fluorescence images were captured at 5X magnification on an Axio Vert.A1 FL (Zeiss) fitted with an AxioCam ICM1 camera (Zeiss 60N-C 2/3" 0.63X adapter). EGFP images were obtained with a BP475/40 excitation and BP530/50 emission filter (FT500 beam splitter). Images were collected using Zen 2 Blue Edition (Zeiss, version 2.0.0.0).

Statistics

Statistical analyses were performed in GraphPad Prism 9.2.0 for Mac, GraphPad Software, San Diego, California USA, www.graphpad.com. Fold-changes in luciferase activity were statistically analysed with a Brown-Forsythe and Welch ANOVA test, assuming Gaussian distribution and unequal standard deviations. Means were compared to the control baseline mean. qPCR data was analysed using multiple student t-test. p-values < 0.05 were considered significant. All data were plotted as mean ± SEM.

ACKNOWLEDGMENTS

A.W.H. was supported by National Health and Medical Research Council (NHMRC) Practitioner Fellowship GNT1150144. D.H. is funded by NHMRC grant GNT1185002 and the Centenary Institute. G.G.N. is funded by NHMRC grants GNT1107514, GNT1158164, GNT1158165, GNT1185002, the NSW Ministry of Health, and a kind donation from Dr. John and Anne Chong. Figure schematics were created with BioRender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, CED, AJC, DH, GGN; Methodology, CED, AJC, MTNT, DH; Investigation, CED, AJC; Resources, AWH, GGN; Data analysis, CED, AJC, MKNMK, MTNT, DH; Writing - original draft, CED, AJC, MTNT, DH; Writing - review & editing, CED, AJC, MTNT, MKNMK, AWH, DH, GGN; Supervision and project administration, AWH, DH, GGN; Funding acquisition, AWH, DH, GGN.

DECLARATION OF INTEREST

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community.

FIGURE LEGENDS

Figure 1. VEGAS generates unproductive transductions.

A. Viral titers after packaging and transduction of pTSin-EGFP into control or SSG-expressing BHK-21 cells (N = 6). Horizontal gray bars indicate the batch-specific thresholds for viral detection.
B. Phase contrast and EGFP fluorescence images of BHK-21 cells used for packaging (R0) and transduction (R1 [MOI 1]). Scale bars represent 200 μm.
C. Viral titers after transduction with undiluted R0 virus into SSG-expressing BHK-21 cells. At R2, viruses were transduced at MOI 1 into control or SSG-expressing BHK-21 cells (N = 6).
D. Phase contrast and EGFP fluorescence images of BHK-21 cells transduced with undiluted R0 virus (R1 [Neat]) and MOI 1 at R2. Scale bars represent 200 μm. White arrowheads indicate EGFP-positive cells. Green boxes denote a magnified section of the image with enhanced brightness and contrast to highlight GFP-positive cells.
E. Nanopore sequencing of viral RNA isolated from pooled samples in Panels C and D (N = 6). Reads (~100,000 per sample) were aligned to viral reference sequences.

Figure 2. Selective pressure does not rescue system integrity.

A. Activation of an SRE-regulated luciferase reporter (SRE_LUC) by SRF-NLS-VP64. Error bars represent mean ± SEM (N = 3, with 3 technical replicates). A Brown-Forsythe and Welch ANOVA test was used to determine statistical significance. **** $p < 0.0001$.
B. Phase contrast and EGFP fluorescence images of BHK-21 cells at R0-R2 for an SRF-NLS-VP64/SRE_SSG circuit. Scale bars represent 200 μm.
C. Viral titers of an SRF-NLS-VP64 transgene-carrying virus packaged in a 1:1 ratio with an EGFP-carrying virus (N = 6).
D. Nanopore sequencing of viral RNA isolated from pooled samples (N = 6). Reads (~100,000 per sample) were aligned to viral reference sequences.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Integrity of *in vitro*-transcribed mRNA.

A. Gel electrophoresis of VEGAS *in vitro*-transcribed mRNA.
B. Gel electrophoresis of *in vitro*-transcribed mRNA used for SRF-NLS-VP64 selection circuit.

Figure S2. Expression of the SSG in transfected BHK-21 cells.

A. Schematic of the barcoded qPCR method used to discriminate mRNA from pDNA during transient transfection. A barcoded SSG-specific primer was used for reverse transcription. Unincorporated barcoded primers were depleted by bead clean-up (indicated by gray circles). qPCR was performed using an SSG-specific forward primer and a barcode-specific reverse primer to discriminate between SSG cDNA and residual pDNA.
B. Raw Ct values from barcoded qPCR for SSG expression of the CMV-SSG plasmid in BHK-21 cells, ± reverse transcription (RT). t-tests were used to determine statistical significance. ***$p < 0.001$, ns = non-significant.
C. SSG expression was calculated relative to GAPDH. Error bars represent mean ± SEM (N = 3). t-tests were used to determine statistical significance. *$p < 0.05$.

Figure S3. RNase A-digested R0 and R1 viruses.

Pooled SINV from R0 and R1 (N = 6, produced on +SSG BHK-21 cells) were treated with RNAse A and titrated.

Figure S4. Packaging signal alignment in seed mRNA sequences.

Schematic showing the position of the SINV packaging signal within VEGAS seed mRNAs (asterisks). Genes and truncations drawn to scale.

Figure S5. Isolated transgene RT-PCR amplicons for nanopore sequencing.

A. Transgenes from electroporated VEGAS viruses (from Figure 1) were amplified by RT-PCR (boxed in red) and gel extracted for nanopore sequencing.
B. Transgenes from the SRF-NLS-VP64/SRE_SSG circuit (from Figure 3C) were amplified by RT-PCR (boxed in red) and gel extracted for nanopore sequencing.

MAIN FIGURES

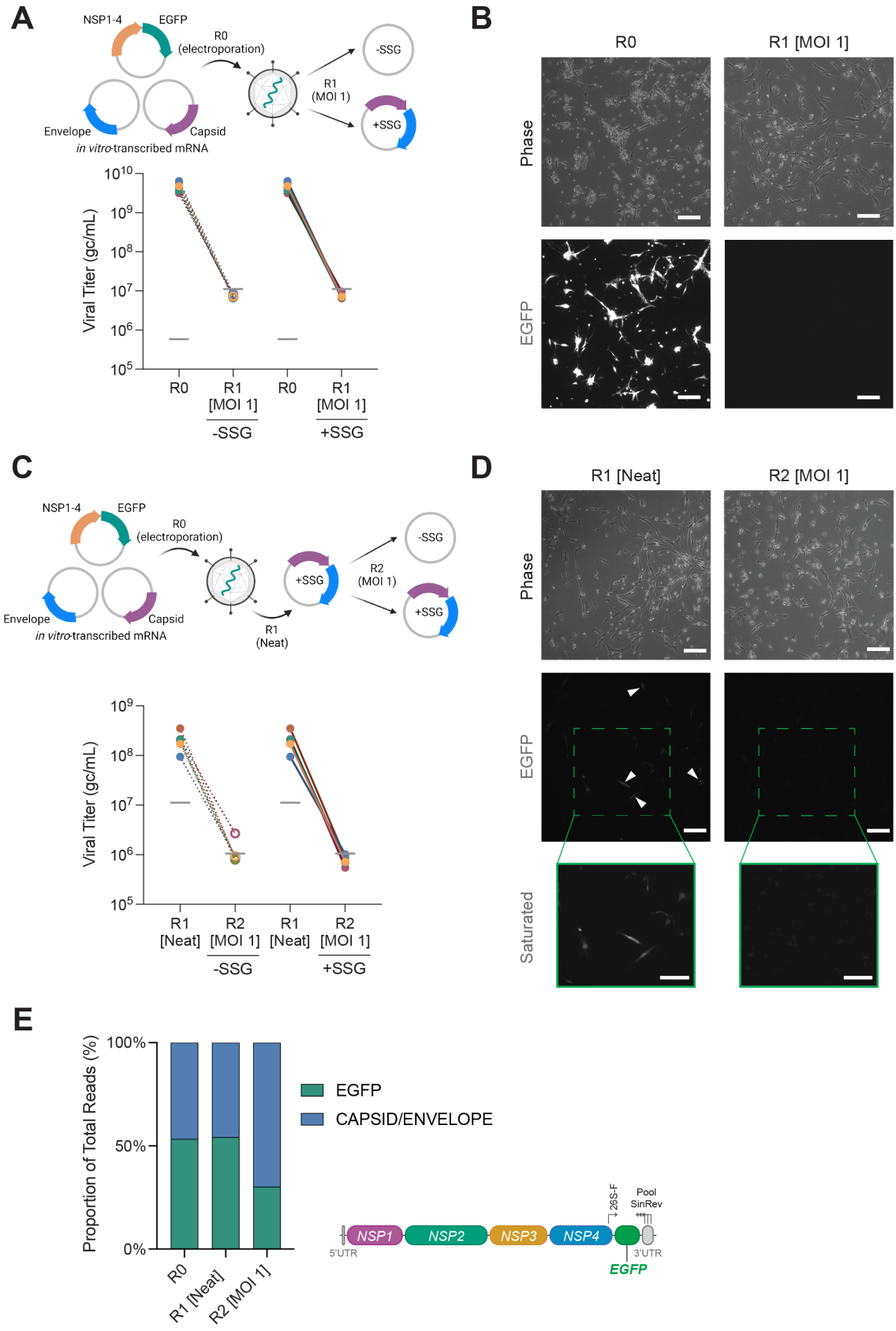


Figure 1

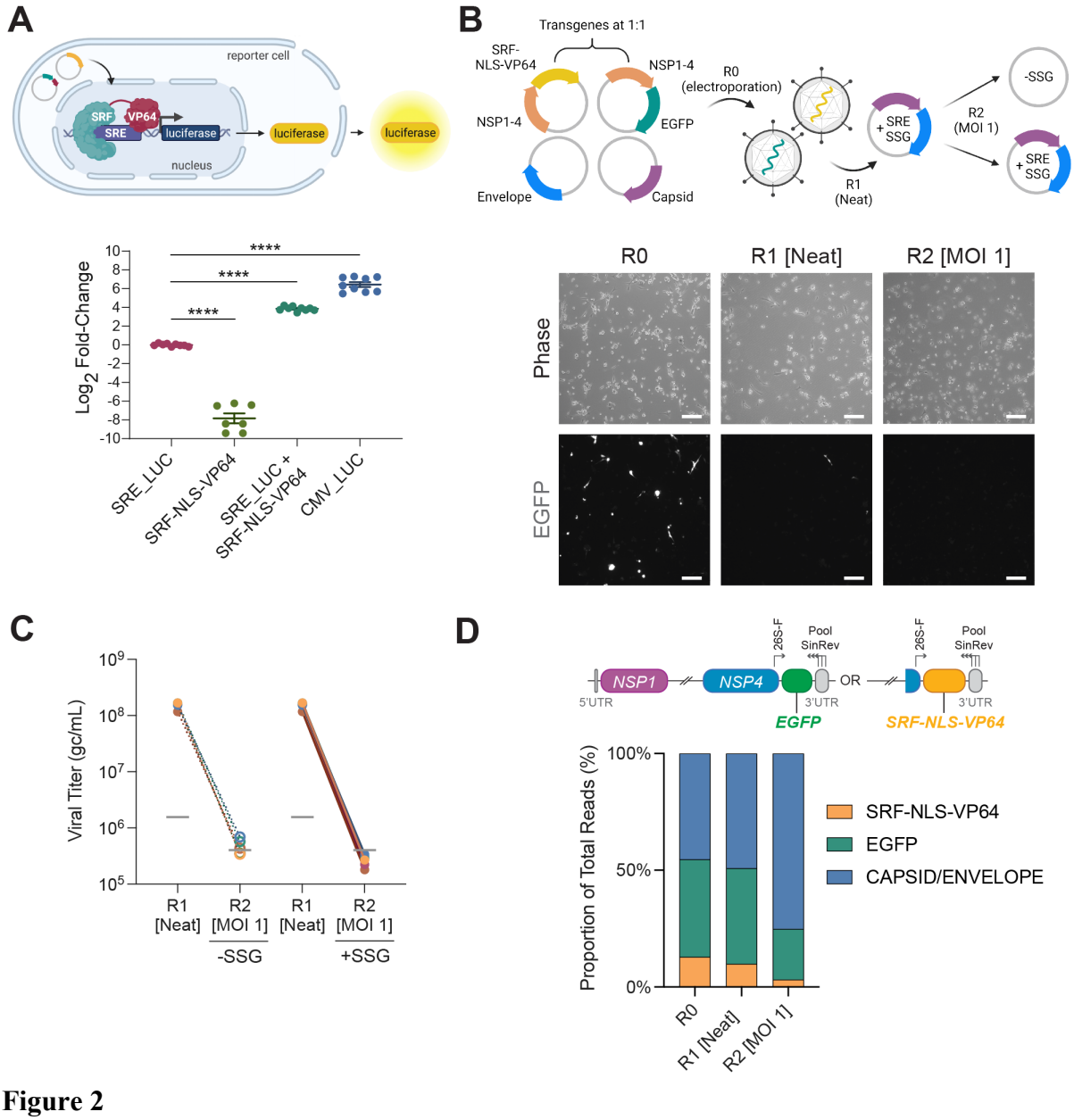
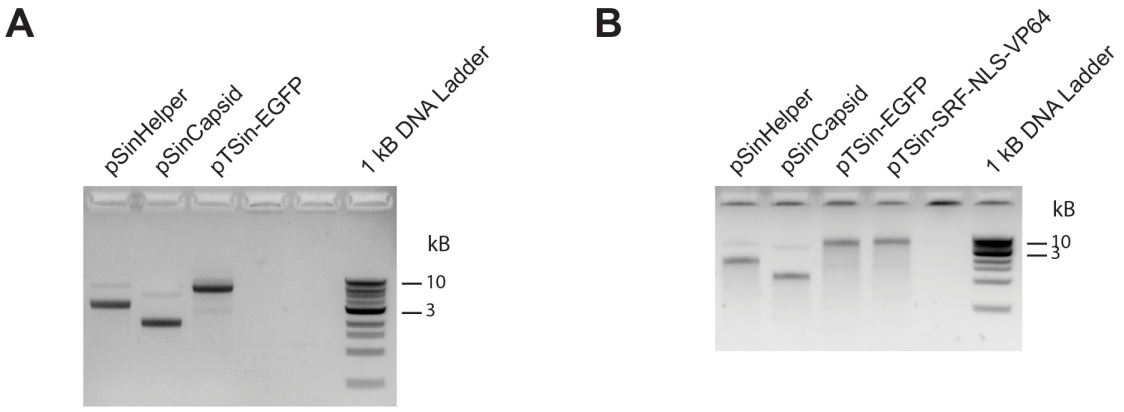
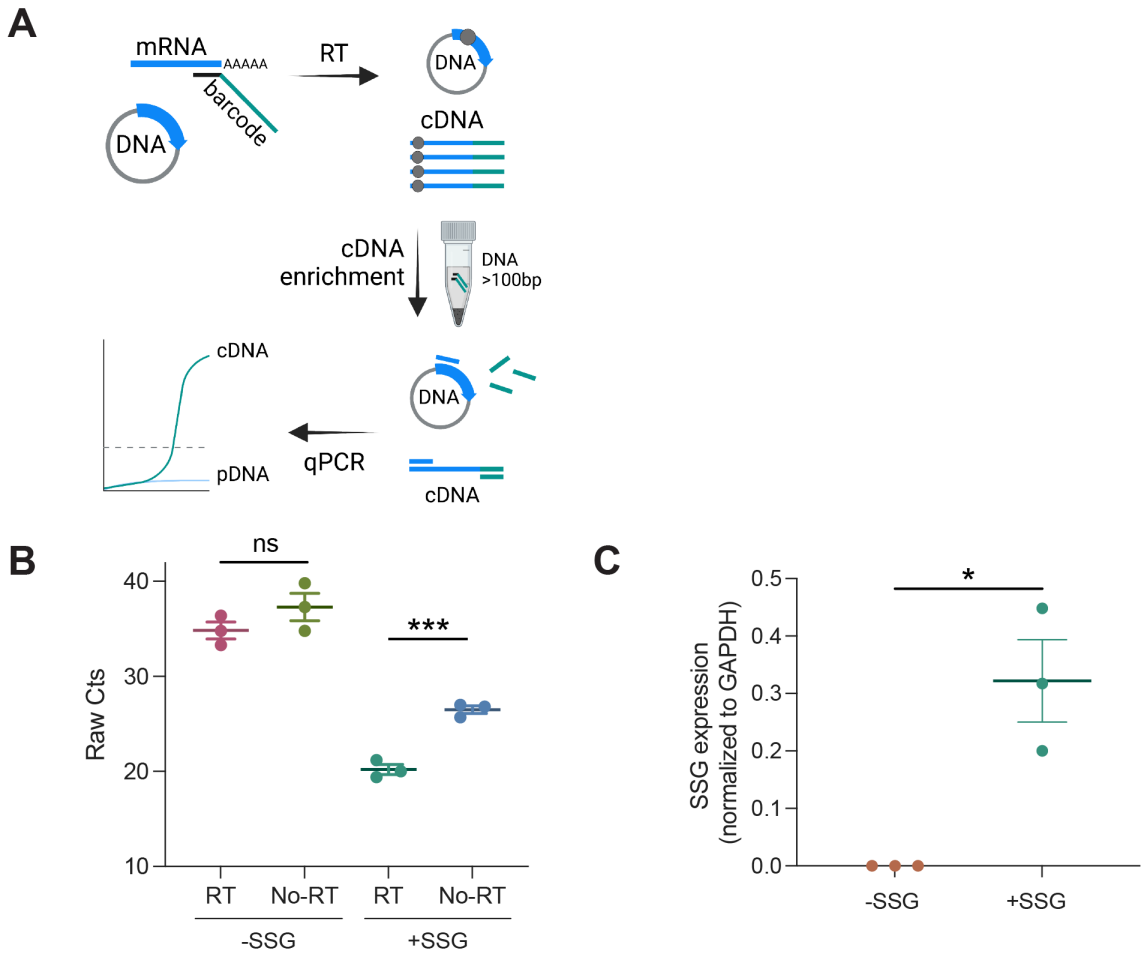


Figure 2

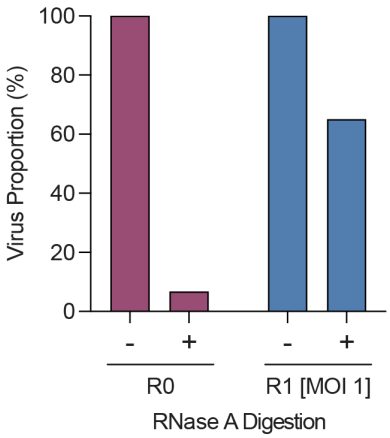
SUPPLEMENTAL FIGURES



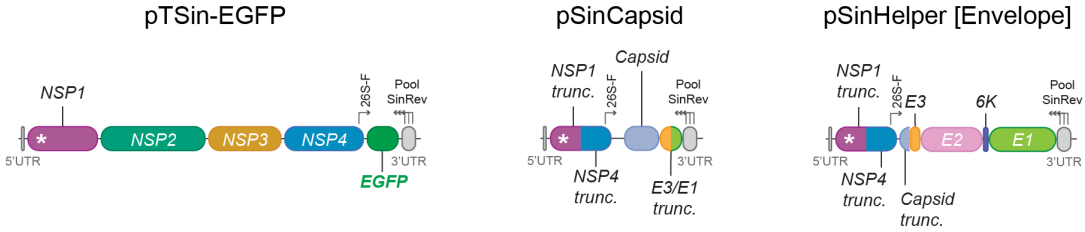
Supplementary Figure 1



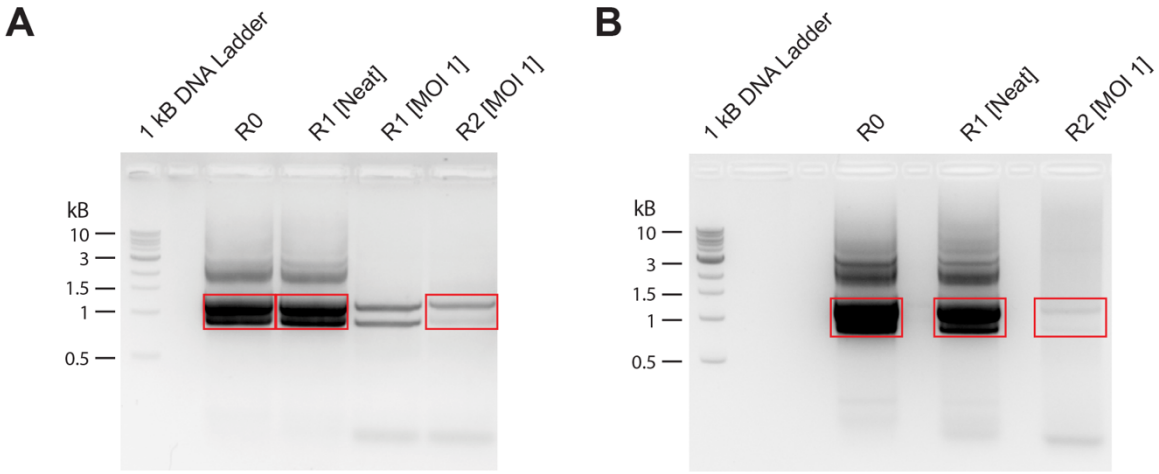
Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5

449 **KEY RESOURCES TABLE:**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
NEB® 10-beta Competent *E. coli* (High Efficiency)	NEB	C3019
Chemicals, Peptides, and Recombinant Proteins		
Ampicillin sodium salt	Sigma-Aldrich	A9518
RNase A	Macherey-Nagel	740505.50
Critical Commercial Assays		
Dual-Glo® Luciferase Assay Kit	Promega	E2940
FavorPrep™ Blood/Cultured Cell Total RNA Mini Kit	Fisher Biotec	FABRK 001-1
iScript™ Select cDNA Synthesis Kit	Bio-Rad Laboratories	1708897
ISOLATE II PCR and Gel Kit	Bioline	BIO-52060
ISOLATE II Plasmid Mini Kit	Bioline	BIO-52057
Ligation Sequencing Kit	Oxford Nanopore Technologies	SQK-LSK109
MagMAX™ Viral RNA Isolation Kit	ThermoFisher	AM1939
mMESSAGE mMACHINE™ SP6 Transcription Kit	ThermoFisher	AM1340
NEBNext® Companion Module for ONT Ligation Sequencing	NEB	E7180
NEBuilder® HiFi DNA Assembly Master Mix	NEB	E2621
PureYield™ Plasmid Maxiprep System	Promega	A2393

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
Qubit™ dsDNA HS Assay Kit	ThermoFisher	Q32851
Qubit™ RNA BR Assay Kit	ThermoFisher	Q10210
SuperScript™ IV One-Step RT-PCR System	ThermoFisher	12594025
SYBR™ Select Master Mix	ThermoFisher	4472908
TaqMan™ Fast Virus 1-Step Master Mix	ThermoFisher	4444434
*Trans*IT®-2020 Transfection Reagent	Mirus Bio	MIR5400
Velocity DNA Polymerase	Bioline	BIO-21099
Deposited Data		
European Nucleotide Archive	This paper	PRJEB47639
Experimental Models: Cell Lines		
Hamster: BHK-21 [C-13]	ATCC	CCL-10
Software and Algorithms		
EPI2ME Desktop Agent (Version 3.3.0.1031)	Oxford Nanopore Technologies	https://www.nanoporetech.com
GraphPad Prism (Version 9.2.0)	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Guppy (Version 4.5.2)	Oxford Nanopore Technologies	https://www.nanoporetech.com
MinKNOW (Version 4.2.8)	Oxford Nanopore Technologies	https://www.nanoporetech.com
SnapGene® (Version 5.3.2)	Insightful Science	https://www.snapgene.com

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
UMIC-seq	GitHub	https://github.com/fhlab/UMIC-seq
Zen 2 Blue Edition (Version 2.0.0.0)	Zeiss	https://www.zeiss.com
Recombinant DNA		
See Supplementary File 1 - Plasmid List. Plasmid maps are available upon request.	N/A	N/A
Oligonucleotides		
See Supplementary File 2 - Primer List.	N/A	N/A
Other		
0.45 μm filter	Merck Millipore	SLHV033RS
6-well plate	Corning	3516
96-well plate	Corning	3596
96-well plate, black-bottom	Interpath	655209
AMPure XP Magnetic Beads	Beckman Coulter	A63881
DMEM, high glucose	ThermoFisher	11965118
DPBS	Sigma-Aldrich	D8537
Fetal Bovine Serum (HyClone), AU origin	Cytiva Life Sciences	SH30084.03
Flongle Flow Cell (R9.4.1)	Oxford Nanopore Technologies	FLO-FLG001
LB Agar	ThermoFisher	22700025
LB Broth	ThermoFisher	12795-084

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
MEM α with nucleosides	ThermoFisher	32571101
Opti-MEM Reduced Serum Medium (GlutaMAX Supplement)	ThermoFisher	51985034
RNaseOUT Recombinant Ribonuclease Inhibitor	ThermoFisher	10777019
Tryptose Phosphate Broth	ThermoFisher	CM0283B
XbaI	NEB	R0145

450

REFERENCES

Berman, Chet M., Louis J. Papa 3rd, Samuel J. Hendel, Christopher L. Moore, Patreece H. Suen, Alexander F. Weickhardt, Ngoc-Duc Doan, et al. 2018. "An Adaptable Platform for Directed Evolution in Human Cells." *Journal of the American Chemical Society* 140 (51): 18093–103.

Ekstrom, Marta H., and David A. Dean. 2011. "The Amaxa Nucleofector System and Standard Electroporation: A Comparison of the Efficiency of DNA Nuclear Localization." *Molecular Therapy: The Journal of the American Society of Gene Therapy* 19 (May): S269–70.

English, Justin G., Reid H. J. Olsen, Katherine Lansu, Michael Patel, Karoline White, Adam S. Cockrell, Darshan Singh, Ryan T. Strachan, Daniel Wacker, and Bryan L. Roth. 2019. "VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells." *Cell* 178 (3): 748–61.e17.

Fayzulin, Rafik, Rodion Gorchakov, Olga Petrakova, Evgenia Volkova, and Ilya Frolov. 2005. "Sindbis Virus with a Tricomponent Genome." *Journal of Virology* 79 (1): 637–43.

Shapiro, Jillian S., Andrew Varble, Alissa M. Pham, and Benjamin R. Tenoever. 2010. "Noncanonical Cytoplasmic Processing of Viral microRNAs." *RNA* 16 (11): 2068–74.

Thomas, John M., William B. Klimstra, Kate D. Ryman, and Hans W. Heidner. 2003. "Sindbis Virus Vectors Designed to Express a Foreign Protein as a Cleavable Component of the Viral Structural Polyprotein." *Journal of Virology* 77 (10): 5598–5606.

Weiss, B., U. Geigenmüller-Gnirke, and S. Schlesinger. 1994. "Interactions between Sindbis Virus RNAs and a 68 Amino Acid Derivative of the Viral Capsid Protein Further Defines the Capsid Binding Site." *Nucleic Acids Research* 22 (5): 780–86.