nature biomedical engineering

Article

https://doi.org/10.1038/s41551-024-01245-z

Robust genome and cell engineering via in vitro and in situ circularized RNAs

Received: 29 January 2023

Accepted: 24 July 2024

Published online: 26 August 2024

Check for updates

Michael Tong[1], Nathan Palmer[2], Amir Dailamy[1], Aditya Kumar[1], Hammza Khaliq[1], Sangwoo Han[1], Emma Finburgh[1], Madeleine Wing[3], Camilla Hong[1], Yichen Xiang[1], Katelyn Miyasaki[1], Andrew Portell[1], Joseph Rainaldi[4], Amanda Suhardjo[1], Sami Nourreddine[1], Wei Leong Chew[5], Ester J. Kwon[1] & Prashant Mali[1]

Circularization can improve RNA persistence, yet simple and scalable approaches to achieve this are lacking. Here we report two methods that facilitate the pursuit of circular RNAs (cRNAs): cRNAs developed via in vitro circularization using group II introns, and cRNAs developed via in-cell circularization by the ubiquitously expressed RtcB protein. We also report simple purification protocols that enable high cRNA yields (40–75%) while maintaining low immune responses. These methods and protocols facilitate a broad range of applications in stem cell engineering as well as robust genome and epigenome targeting via zinc finger proteins and CRISPR–Cas9. Notably, cRNAs bearing the encephalomyocarditis internal ribosome entry enabled robust expression and persistence compared with linear capped RNAs in cardiomyocytes and neurons, which highlights the utility of cRNAs in these non-dividing cells. We also describe genome targeting via deimmunized Cas9 delivered as cRNA and a long-range multiplexed protein engineering methodology for the combinatorial screening of deimmunized protein variants that enables compatibility between persistence of expression and immunogenicity in cRNA-delivered proteins. The cRNA toolset will aid research and the development of therapeutics.

RNAs have emerged as a powerful therapeutic class. However, their typically short half-life impacts their activity both as an interacting moiety (such as short interfering RNAs) and a template (such as messenger RNAs). Towards this, RNA stability has been modulated using a host of approaches, including engineering untranslated regions (UTRs), modulating secondary structures, incorporating cap analogues, modifying nucleosides and optimizing codons[1–5]. More recently, circularization strategies that remove free ends necessary for exonuclease-mediated degradation thereby rendering RNAs resistant to most mechanisms of turnover have emerged as a particularly promising methodology[6–15].

Simple and scalable approaches to achieve efficient production and purification of circular RNAs (cRNAs) are however lacking, thus limiting their broader application. Towards this, we developed two methods: '**o**utside developed' **c**RNA (ocRNA) via in vitro circularization using group II introns, and in situ or '**i**nside developed' **c**RNA (icRNA) via in-cell circularization using the ubiquitously expressed RtcB protein. As cRNA production commonly involves RNaseR enrichment and high performance liquid chromatography (HPLC) purification to reduce immune responses prompted by double-stranded RNA produced during in vitro transcription (IVT), resulting RNA yields are low (typically less than 1% of the input RNA[6]). We thus optimized

[1]Department of Bioengineering, University of California San Diego, La Jolla, CA, USA. [2]Biological Sciences Graduate Program, University of California San Diego, La Jolla, CA, USA. [3]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA, USA. [4]Biomedical Sciences Graduate Program, University of California San Diego, La Jolla, CA, USA. [5]Genome Institute of Singapore (GIS), Agency for Science, Technology and Research (A*STAR), Singapore, Republic of Singapore. e-mail: pmali@ucsd.edu

https://doi.org/10.1038/s41551-024-01245-z

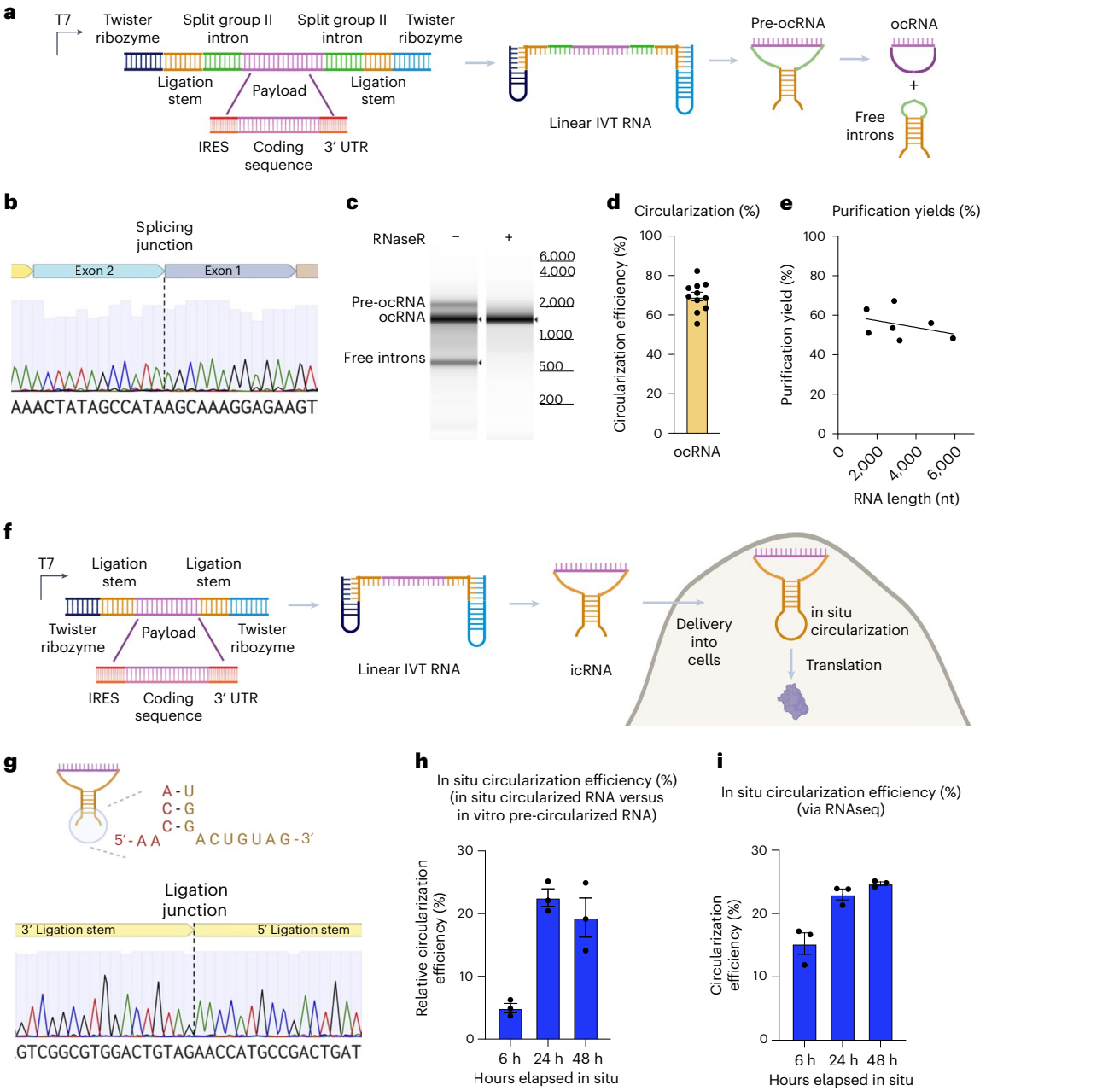


Fig. 1 | Engineering ocRNAs and icRNAs. a, Schematic describing the production of ocRNAs. These are generated via IVT of linear RNAs that bear twister ribozyme + permuted group II intron–exon sequence flanked IRES coupled to an mRNA of interest. Once transcribed, the flanking twister ribozymes rapidly self-cleave, enabling hybridization of the complementary ligation stems to one another. Then, autocatalytic intron splicing occurs, releasing the introns and ligating the spliced ends together. **b**, Sanger sequencing trace mapping the junction site formed upon ligation. **c**, Tapestation of ocRNA before and after RNaseR treatment. **d**, Circularization efficiencies quantified by tapestation analysis. Values represented as mean ± s.e.m. ($n = 11$). **e**, Yields from cellulose dsRNA purification across differing length RNA constructs. **f**, Schematic describing the production of icRNAs. These are generated via IVT of linear RNAs that bear a twister ribozyme flanked IRES coupled to an mRNA of interest. Once transcribed, the flanking twister ribozymes rapidly self-cleave, enabling hybridization of the complementary ligation stems to one another, and upon delivery into cells, these linear RNAs are then circularized in situ by the ubiquitous RNA ligase RtcB. **g**, Sanger sequencing trace mapping the junction site formed upon ligation. **h**, HEK293Ts were transfected with icRNA or in vitro pre-circularized icRNA, and RNA was isolated at 6 h, 24 h and 48 h. RT-PCR was performed and the ratio of the icRNA band to pre-circularized icRNA band was plotted to evaluate in situ circularization efficiencies. Values represented as mean ± s.e.m. ($n = 3$). **i**, HEK293Ts were transfected with icRNA and RNA was isolated at 6 h, 24 h and 48 h and RNAseq performed. cRNA counts relative to total RNA are shown. Values represented as mean ± s.e.m. ($n = 3$).

protocols that incorporated alternatives and our resultant methods use an HPLC-free purification process while maintaining low immunogenicity, persistence and robust expression. Importantly, they enabled both high yields and cost-effectiveness, which we leveraged to demonstrate a range of applications from stem cell engineering to robust genome and epigenome targeting via zinc finger (ZF) proteins and CRISPR–Cas9 systems.

Common to all these applications enabled via cRNA delivery is the critical consideration of their immune system interactions. Although for some applications, such as vaccines, robust immune responses to a delivered transgene are desirable, for other applications, such as genome and epigenome targeting, immune responses can instead inhibit therapeutic effects[16,17]. Inducing immune responses through transgene expression enabled by RNA delivery has been extensively

researched in vaccine development and proven through the success of COVID vaccines based on this technology[18–21]. However, despite substantial engineering efforts, deimmunization remains a tougher problem to crack[22]. Thus, to facilitate compatibility between persistence of expression and immunogenicity, especially when delivering non-human payloads via cRNAs, we also concurrently developed a **lo**ng-**ra**nge multiple**x**ed (LORAX) protein engineering methodology based on high-throughput screening of combinatorially deimmunized protein variants. We applied it to identify a Cas9 variant with seven key human leukocyte antigen (HLA) restricted epitopes simultaneously immunosilenced after a single round of screening, and showed that cRNA-mediated delivery of the same enabled genome targeting.

Results

Engineering ocRNAs and icRNAs

To engineer ocRNAs, we adopted the group II intron from *Clostridium tetani*[23–25], where domains I, II and III were split from domains V and VI at domain IV, to preserve the catalytic and structurally relevant properties for intron splicing while permitting circularization of our RNA construct. We permuted the domains of the intron to have domains V and VI at the 5′ end of our construct and domains I to III at the 3′ end of our construct. Since the *C. tetani* group II intron is seen to have alternative splicing with four groups of DV/DVI domains (A, B, C and D), we kept a minor leader sequence to domain DV/DVI group A as it was likely to house the crucial exon binding site (EBS) required for the domain I intron sequence to bind for selective splicing. Importantly, the *C. tetani* intron is ORF-less, suggesting that splicing action is not dependent on an intron-encoded protein and thus compatible with IVT applications. Finally, to improve long-range interactions, we added flanking twister ribozymes that rapidly self-cleave during IVT (removing the immunogenic 5′ phosphate) and enable hybridization of downstream complementary ligation stems to one another. The above circularization mechanics are adjacent to an internal ribosome entry site (IRES) (25, 26) coupled to an mRNA of interest and a 3′ UTR (Fig. 1a). We analysed the junction via Sanger sequencing of reverse transcription PCR (RT-PCR) product from IVT and found a small 26 nt splicing scar split between proposed EBS of each permuted domain group, confirming circularization (Fig. 1b). In further characterization, we observed three species on tapestation that correspond with the lengths of each expected component: the intron arms, circularized product and precursor (Fig. 1c). Evaluating IVT circularization efficiency of the group II intron, we found yields to be consistently around 70% (Fig. 1d). In downstream dsRNA purification, we also noted that the ocRNA construct retained an approximate 60% yield (Fig. 1e).

To engineer icRNAs, we took a simplified approach that did not require engineering beyond flanking the IRES and payload with twister ribozymes (Fig. 1f). Specifically, to engineer icRNAs, we generated in vitro transcribed linear RNAs that bear a twister ribozyme flanked IRES[26,27] coupled to an mRNA of interest and a 3′ UTR. Once transcribed, the flanking twister ribozymes rapidly self-cleave, enabling hybridization of the complementary ligation stems to one another. Upon delivery into cells, these linear RNAs are then circularized in situ via ligation of the proximal 5′ and 3′ ends by the ubiquitous cellular RNA ligase RtcB. Thus, contrary to the ocRNA, this construct completely relies on the endogenous RtcB action to circularize. As a first check of circularization, we confirmed that the ligation junction was mapped correctly to its predicted location via Sanger sequencing of the RT-PCR product from total RNA (Fig. 1g). To assess in situ circularization efficiency, HEK293T cells were transfected with icRNAs encoding for green fluorescent protein (GFP) or with the same icRNAs pre-circularized in vitro using RTCB ligation in a test tube followed by RNaseR treatment to enrich circularized RNA species. RT-PCR was performed using outward-facing primers and normalized to GAPDH expression. Across independent icRNA treatments, we observed efficient (~20%) in situ circularization levels for icRNAs compared with pre-circularized RNA (Fig. 1h). However, since RNA degrades with time in situ, we wanted to observe the proportions of icRNA as best we could with cell proliferation. With RNA sequencing (RNAseq) data across the same three time points of 6 h, 24 h and 48 h, we observed a net increase of cRNA proportion, consistent with circularization increasing over time and depletion of linear RNA through endonucleolytic attack (Fig. 1i). Notably, the largest change in proportion occurs within the first 24 h, probably owing to the short half-life of linear RNA. As another side-by-side comparison, we also engineered linear **i**n situ **c**ircularization **d**efective RNAs (icdRNAs) by utilizing catalytically inactive mutants of the twister ribozymes. Specifically, HEK293Ts were transfected with circular GFP icRNA or linear icdRNA and RNA was isolated at 6 h, 1 day, 2 days and 3 days after transfection. We observed similar amounts of GFP RNA at 6 h (Supplementary Fig. 1a, left), confirming that approximately equal quantities of icRNA and icdRNA were delivered to cells. However, GFP RNA with functional circularization was much higher at days 1, 2 and 3 than icdRNA, indicating improved RNA persistence via circularization. This improved RNA persistence also correlated with increased GFP translation at 3 days (Supplementary Fig. 1a, middle). To confirm that icRNAs were covalently circularized in cells upon delivery in vitro, we performed RT-PCR by designing outward-facing primers that selectively amplified only the circularized RNA molecules. Indeed, we only observed a PCR product for icRNAs, confirming successful circularization (Supplementary Fig. 1a, right).

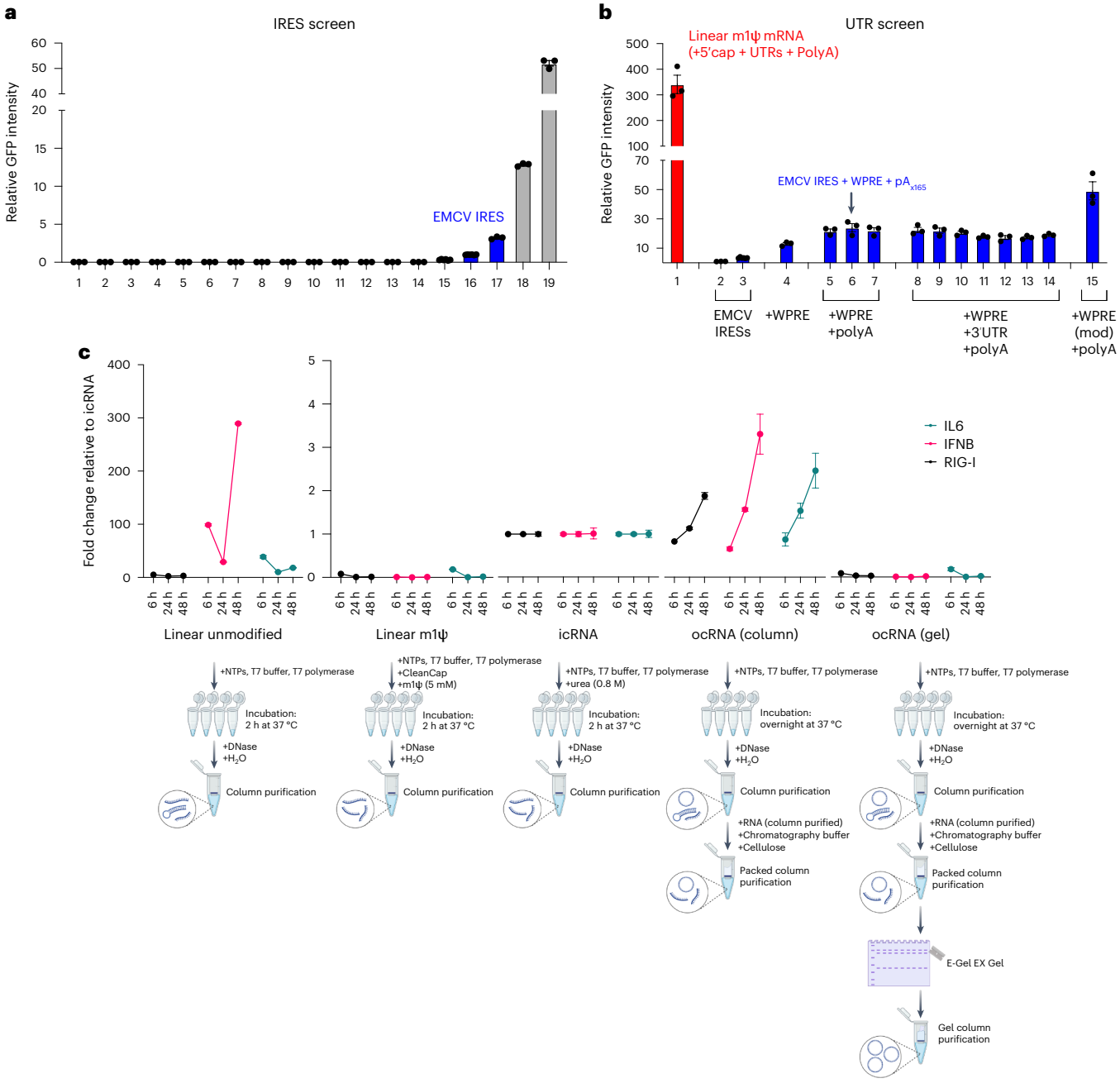
Engineering improved translation from cRNAs

To improve protein translation from cRNAs, we next screened a panel of 19 naturally occurring and synthetic IRES sequences[28–32] (Supplementary Table 1). We found the 6A form of the encephalomyocarditis virus IRES (EMCV; Fig. 2a, blue bar), the enterovirus 71 IRES (Fig. 2a, #18) and the Coxsackievirus B3 IRES (Fig. 2a, #19)[6] to be the best at enabling cap-independent protein translation. To further improve protein translation, we also screened a panel of 3′ UTRs (Supplementary

Fig. 2 | Optimization and characterization of ocRNAs and icRNAs. **a**, HEK293Ts were transfected with icRNAs containing various IRES sequences and GFP intensity was quantified by flow cytometry. cRNAs containing the EMCV IRES (blue bar) were selected for further optimization. Values represented as mean ± s.e.m. ($n = 3$). **b**, HEK293Ts were transfected with icRNAs containing the EMCV IRES coupled with various 3′ UTRs and poly(A) stretches (blue bars) and GFP intensity was quantified by flow cytometry. Addition of a WPRE and a poly(A) stretch substantially improved protein translation, and icRNAs bearing the EMCV IRES were used for all subsequent studies. These designs were also compared with capped linear N^1-methylpseudouridine-5′-triphosphate (m1Ψ) RNA (red bar). Values represented as mean ± s.e.m. ($n = 3$). **c**, A549s were transfected with various constructs encoding GFP and RNA isolated at 6 h, 24 h and 48 h. Immune markers RIG-I, IFNB and IL6 were quantified by RT-qPCR for each sample relative to GAPDH. Values represented as mean ± s.e.m. ($n = 3$). Shown below are detailed methods of synthesis and purification for each construct. **d**, A549s were transfected with the same constructs encoding GFP, and RNA was isolated at 6 h, 24 h and 48 h. Cell viability was quantified via CCK-8. Absorbance at 450 nm was measured for samples on day 0, day 1 and day 2. Data were normalized within each sample to day 0 values. Values represented as mean ± s.e.m. ($n = 3$). **e**, HEK293T cells were transfected with circular GFP icRNA containing encephalomyocarditis IRES with WPRE and 50 nt poly(A) stretch or 5′ capped linear RNA and the GFP mRNA amount was measured over time (left axis). Values were normalized to the amount at the 6 h time point for each respective group ($n = 3$, $P = 0.0072$ for day 1, $P = 0.0015$ for day 2, $P = 0.00086$ for day 3, $P = 0.0037$ for day 4 and $P = 0.000531$ for day 5; t-test, two-tailed). The ratio of icRNA GFP mRNA compared with linear RNA for each day is plotted (right axis). The increase in value over time illustrates improved persistence of icRNA.

https://doi.org/10.1038/s41551-024-01245-z

Table 2) in the context of the EMCV 6A IRES. Addition of a Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) and a poly(A) stretch improved relative protein translation by over five-fold, but additional 3′ UTRs[5] did not improve translation efficiency (Fig. 2b, blue bars). This is consistent with similar improvements in linear RNA, as poly(A) binding proteins (PABPs) are known to interact with the translation complex from the open 3′ end of the RNA structure in a sandwich conformation[33,34]. Therefore, we hypothesized that a



https://doi.org/10.1038/s41551-024-01245-z

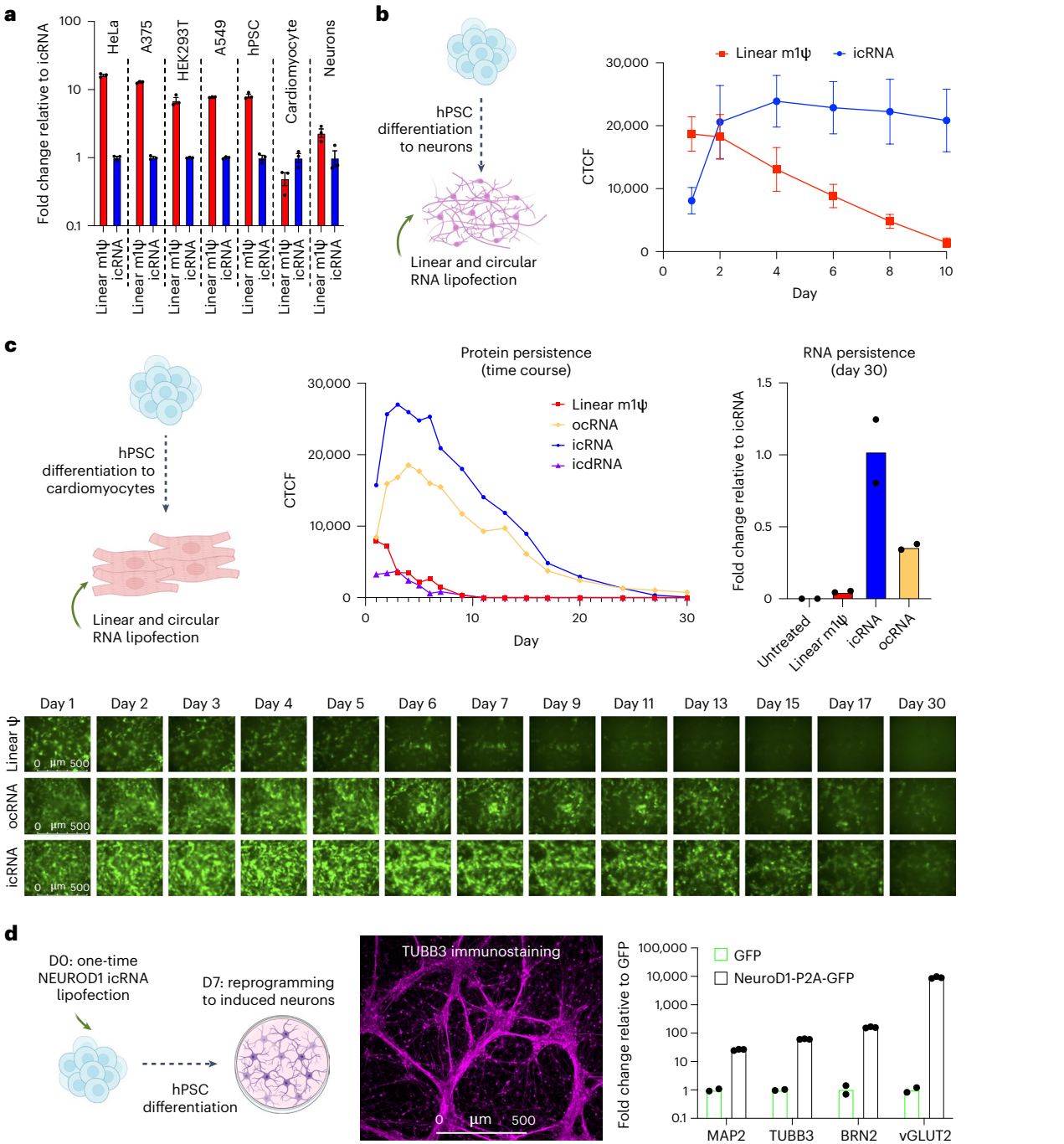


Fig. 3 | Assessing persistence and activity of ocRNAs and icRNAs. a, Indicated cell types were transfected with linear m1Ψ and icRNA. GFP intensity was quantified on day 1 by flow cytometry relative to icRNA. (Cardiomyocytes and neurons are represented by calculated total cell fluorescence (CTCF) image data.) Values represented as mean ± s.e.m. ($n = 3$). **b**, Left: post differentiation of stem cells into neurons, icRNAs or linear m1Ψ RNAs were transfected into cells and images were taken over 10 days. CTCF mCherry expression over time was plotted for icRNA and linear m1Ψ RNA. **c**, Left: post differentiation of stem cells into cardiomyocytes, ocRNAs, icRNAs or linear m1Ψ RNAs were transfected into cells and images were taken over 30 days. Middle: CTCF GFP expression over time was plotted for icRNA and linear m1Ψ RNA. Right: relative GFP RNA quantified with RT-qPCR after 30 days. Values represented as mean ($n = 2$). Bottom: representative images are shown illustrating icRNA and ocRNA persistence. **d**, Left: schematic of a one-time transfection of 500 μg icRNA encoding NeuroD1-P2A-GFP onto stem cells (H1). Middle: day 7 TUBB3 immunostaining of differentiated neurons. Right: neural markers MAP2, TUBB3, BRN2 and vGLUT2 were quantified by RT-qPCR for each sample relative to GAPDH. Values for GFP are represented as mean ($n = 2$) and values for NeuroD1-P2A-GFP are represented as mean ± s.e.m. ($n = 3$).

poly(A) tail in this conformation assists in recruiting PABPs in cRNAs. For all subsequent studies, we thus used cRNAs based on the EMCV IRES coupled to a modified WPRE and 165 bp poly(A) stretch (Fig. 2b, #15). Here a crucial 'Y-stem' secondary structure noted for recruiting the core translation complex eIF4G was found to have high stacking probability with a short sequence of the WPRE via RNAfold, and thus the sequence was removed from the WPRE, recovering an approximate twofold improvement in expression[35,36]. Notably, these designs showed robust activity when benchmarked with contemporary optimized cRNA constructs[15] (Supplementary Fig. 1b, green bar).

Developing facile purification protocols for cRNAs

Conventional purification strategies rely on HPLC or enzyme-based techniques, which can be complex to implement and/or costly and/or yield limiting and variable results (as seen by incomplete RNaseR reactions). Therefore, we sought to utilize simpler and more economical methods. Specifically, for icRNAs, we included urea (0.8 M) as a denaturing agent during the IVT process to reduce the highly immunogenic and spurious dsRNA production inherent to transcription via T7 polymerases[37]. For ocRNA constructs, optimal circularization yields were obtained via overnight IVT reactions but in the absence of urea (as crucial secondary structures for splicing are abrogated upon addition of denaturing agents) (Supplementary Fig. 1c). For ocRNA purification, we thus instead relied on cellulose chromatography[38] that can selectively remove dsRNA above 30 bp in length (which is above that of the crucial IRES secondary structures within cRNAs). Finally, resulting RNA was treated with phosphatase. Next, we evaluated the immune response of thus purified icRNAs and ocRNAs versus linear RNA benchmarks in A549 cells. As reported in previous literature, linear unmodified RNA demonstrated a large immune response, but linear RNA with N^1-methylpseudouridine-5′-triphosphate (m1Ψ) modification demonstrated negligible responses from the retinoic acid-inducible gene I (RIG-I), interleukin-6 (IL6) and interferon beta (IFNB) immune markers[9,14] (Fig. 2c, top). Notably, icRNA and ocRNA constructs demonstrated mild immune responses, with the ocRNA (gel) demonstrating almost undetectable immune responses (similar to the linear capped and modified RNAs). Accordingly, all the circular constructs also demonstrated robust cell viability (Fig. 2d). While HPLC purification will be the method of choice of clinical translation, the cRNA purification approaches above provide a viable alternative for most in vitro and in vivo applications.

Functional characterization of cRNAs

We next sought to validate cRNA persistence across relevant in vitro and in vivo settings. We first confirmed that cRNAs were indeed improved in persistence compared with their linear counterparts in HEK293Ts (Fig. 2e). However, consistent with previous observations comparing IRES-based cRNAs versus 5′ capped linear RNA[14,15], we observed that the latter (Fig. 2b, red bar) enabled higher instantaneous protein translation 24 h post delivery (2–3-fold). Hypothesizing that IRES translation may be cell type specific[39], we thus screened a variety of cells to determine the optimal cellular contexts to deploy our cRNAs (Fig. 3a). Specifically, since the IRES that we chose was the EMCV IRES, we rationalized that a test in neurons and cardiomyocytes could demonstrate higher efficiencies as it represents a closer cellular environment to what is natural for the parent cardiovirus type. To that end, we transfected icRNAs containing the EMCV, modified WPRE, and a 165 poly(A) stretch or a linear capped RNA with m1Ψ nucleotide substitution and optimized UTRs into stem cell-derived neurons (Fig. 3b, left). Interestingly, we found that the icRNAs provide extended expression and increased translation after a short delay (Fig. 3b, right). Next, we transfected the same panel with the additional conditions of ocRNAs and defective icdRNAs into stem cell-derived cardiomyocytes (Fig. 3c, left). Interestingly, our observations confirmed that both EMCV-driven circularized RNA constructs exhibited higher or equivalent expression levels on day 1 compared with linear RNA. Notably, they again rapidly surpassed linear RNA in expression by day 2 and far past their own day 1 expression (Fig. 3c, middle, bottom). Moreover, linear and circularization defective icdRNAs demonstrated rapid and similar decreases over the course of a week. To ensure that trends were reproducible, the experiment was repeated and image analysis performed in triplicate at a lower exposure to optimize representation of GFP intensity, resulting in a nearly identical profile (Supplementary Fig. 1d). On day 30, cells were collected, and qPCR analysis of GFP RNA was performed (Fig. 3c, right). Results indicated a considerably lower RNA level of linear RNAs compared with icRNAs and ocRNAs further confirming persistence of circular species. Overall, these results suggest a potential cellular context dependence on the activity of IRESs and/or a preference for non-dividing cells, which compared with dividing cells do not as rapidly dilute RNA levels post transfection. On that note, we hypothesized that the differentiation of stem cells into non-dividing cells might similarly better leverage the heightened expression and apparent protein accumulation provided by persistent icRNA. Towards this, we were indeed able to robustly differentiate H1 human pluripotent stem cells (hPSCs) to neurons within 1 week via a single transfection of circular NEUROD1 RNAs (Fig. 3d).

Finally, we sought to extend these results in vivo. To ensure that icRNAs were also capable of circularization in vivo, we generated lipid nanoparticles (LNPs) containing either circular icRNA or icdRNA and retro-orbitally injected them into mice. LNPs for each condition were successfully generated of similar size (Supplementary Fig. 2a, left). Livers were isolated 3 and 7 days after injection and quantification and visualization by RT-qPCR confirmed circularization in vivo persisting for at least 7 days (Supplementary Fig. 2a, middle, right). To benchmark against true linear RNA, we synthesized LNPs[40,41] bearing either icRNA or linear RNAs (Supplementary Fig. 2b, left). RNA/LNPs were retro-orbitally injected into C57BL/6 mice and liver transcriptomes assayed 7 days after injection. While linear RNA was largely degraded by day 7, RT-qPCR showed that icRNAs persisted (Supplementary Fig. 2b, middle). Furthermore, RT-PCR also again confirmed successful in situ circularization of icRNAs in vivo (Supplementary Fig. 2b, right).

Genome and epigenome engineering via ZF proteins delivered as cRNAs

Spurred by the above results, we hypothesized that this increased persistence of cRNAs, in addition to enabling applications entailing sustained transgene expression, could also facilitate efficient genome and especially epigenome targeting. Towards this, we first explored icRNA utility in the context of ZF proteins, as being a solely protein-based genome engineering toolset we anticipated that ZFs would be particularly suited for this mode of delivery. Indeed, we observed more efficient genome editing via zinc finger nuclease (ZFN) icRNAs compared with corresponding icdRNAs targeting the GFP and CCR5 genes[42,43] (Fig. 4a). Building on this observation, we next investigated delivery of ZF-KRAB proteins to programmably repress genomic targets. We focused on PCSK9, a gene that encodes an enzyme that regulates low-density lipoprotein receptor degradation. Loss-of-function mutations in PCSK9 are associated with reduced risk of cardiovascular disease with no documented adverse side effects[44–46]. Antibodies, antisense oligonucleotides and CRISPRs have all been utilized to target PCSK9[47–53], and here we explored if a transient pulse of ZF epigenome regulators could enable repression of PCSK9. First, we screened in HeLa cells a panel of 67 ZF-KRAB proteins delivered as icRNAs (Fig. 4b, top). These ZFs tiled across 0–1,200 bp of the transcription start site (Fig. 4b, middle) of hPCSK9, and notably, 16 of these enabled robust gene repression (90–50%) (Fig. 4b, bottom). Next, to achieve inheritable repression, we fused a 3A3L DNA methylator with the ZF-KRAB modules[54,55]. Notably, we observed sustained repression of PCSK9 over a 12-day period (Fig. 4c). Taken together, these results establish that a transient pulse of ZF epigenome regulators delivered as cRNAs can enable robust gene repression, enabling a genomically scarless and safe approach for modulating therapeutic gene expression.

Genome and epigenome engineering via deimmunized CRISPR–Cas delivered as cRNAs

Building on these results with ZF proteins, we next explored if cRNA persistence could similarly enhance the activity of CRISPR–Cas9 systems. However, unlike for ZFs, which are built on a human protein chassis, this feature of persistence may aggravate immune responses in therapeutic settings for CRISPR systems as those are derived from prokaryotes, including some residing in the human gut microbiome[56–60]. Thus, to enable compatibility between persistence of expression and

https://doi.org/10.1038/s41551-024-01245-z

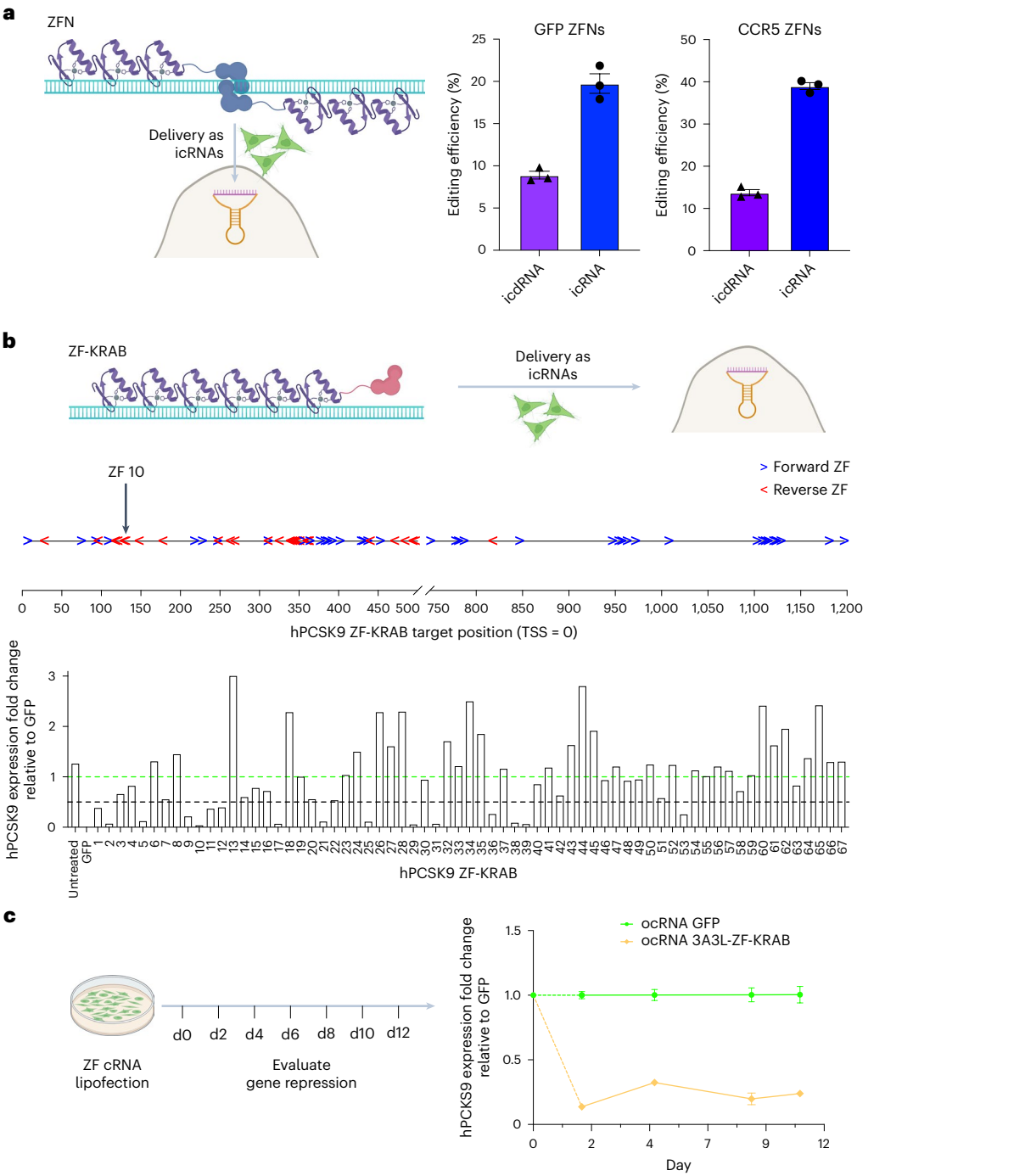


Fig. 4 | Application of icRNAs and ocRNAs to ZF-mediated genome and epigenome targeting. a, Editing efficiency of circular icRNA or circularization defective icdRNA ZFNs targeting a stably integrated GFP gene or the endogenous CCR5 gene in HEK293T cells is plotted. Values represented as mean ± s.e.m. ($n = 3$). **b**, Target locations of ZFs are indicated alongside forward (sense) and reverse (antisense) strand binding. Repression efficiency of ZF-KRAB proteins produced by circular icRNA in HeLa cells is plotted on the left as hPCSK9 expression fold change relative to the circular GFP icRNA quantified with RT-qPCR after 48 h. The black dotted line indicates the measure for successful repression of hPCSK9 by a human ZF-KRAB protein. The green dotted line indicates hPCSK9 levels in the GFP control. **c**, Transient hPCKS9 repression efficiency produced by a 3A3L and KRAB fusion with ZF10 delivered as ocRNA in HeLa cells. hPCSK9 levels are quantified with RT-qPCR at each time point. Values represented as mean ± s.e.m. ($n = 3$).

immunogenicity, we sought first to develop a methodology to screen progressively deimmunized Cas9 proteins by combinatorially mutating particularly immunogenic epitopes[61].

While variant library screening has proven to be an effective approach to protein engineering[62–69], applying it to deimmunization faces three important technical challenges. One, the need to mutate multiple sites simultaneously across the full length of the protein; two, reading out the associated combinatorial mutations scattered across large (>1 kb) regions of the protein via typical short-read sequencing platforms; and three, engineering fully degenerate combinatorial libraries, which can very quickly balloon to unmanageable numbers of variants[22,70]. To overcome these challenges, we developed several methodological innovations, which, taken together, comprise a LORAX protein engineering method capable of screening millions of

https://doi.org/10.1038/s41551-024-01245-z

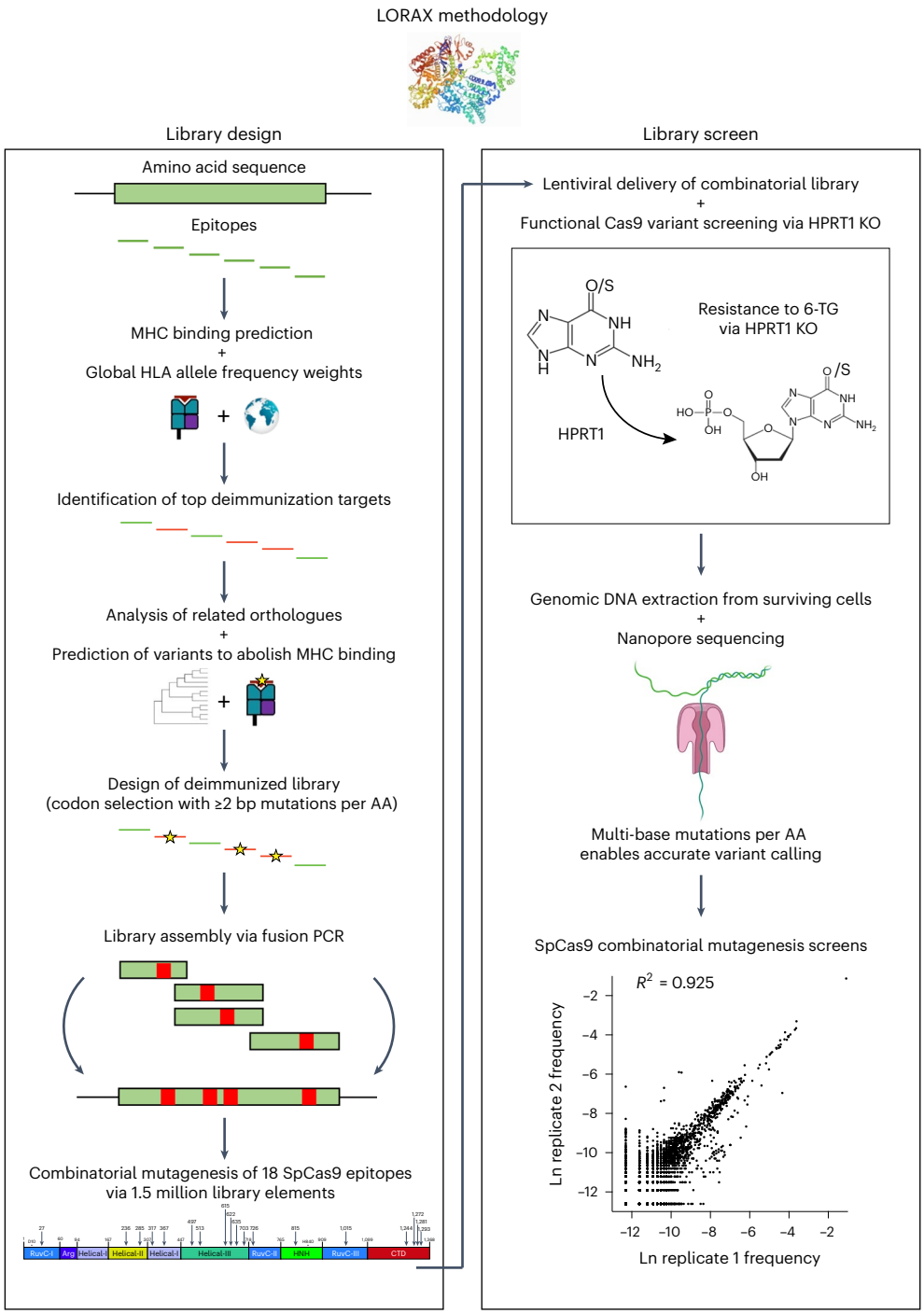


Fig. 5 | LORAX protein engineering methodology to screen progressively deimmunized Cas9 variants. Left: library design. Low-frequency SNPs that have a limited effect on Cas9 function were identified and immunogenicity was evaluated in silico using the netMHC epitope prediction software to identify candidate mutations. This analysis was performed for many Cas9 orthologues. Mutations were generated such that 2 bp was changed to account for nanopore sequencing accuracy. A library was then generated by fusion PCR of blocks containing WT and mutations at specific epitopes. Location of epitopes in SpCas9 that were combinatorially mutated and screened is shown. Right: library screen. The screen was performed by transducing HeLa cells with a lentiviral library containing the Cas9 variants and a guide that cuts the HPRT1 gene. HPRT1 knockout produces resistance to 6-TG. After 2 weeks, DNA is extracted from surviving cells and Cas9 variant sequences are PCR amplified from the genomic DNA and nanopore sequenced. High accuracy of variant identification is possible owing to the use of 2 bp mutations for each amino acid change. Post-screen library element frequencies across two independent replicates are shown. Replicate correlation was calculated excluding the over-represented WT sequence.

combinatorial variants simultaneously with mutations spread across the full length of arbitrarily large proteins (Fig. 5).

Towards library design, to narrow down the vast mutational space associated with combinatorial libraries, we utilize an approach guided by evolution and natural variation[71,72]. As deimmunizing protein engineering seeks to alter the amino acid sequence of a protein without disrupting functionality, it is extremely useful to narrow down mutations to those less likely to result in non-functional variants. To identify these mutants, we generated large alignments of Cas9 orthologues from publicly available data to identify low-frequency single

nucleotide polymorphisms (SNPs) that have been observed in natural environments. Such variants are likely to have a limited effect on protein function, as highly deleterious alleles would tend to be quickly selected out of natural populations (if Cas9 activity is under purifying selection) and therefore not appear in sequencing data[73]. To further subset these candidate mutations, we evaluated for immunogenicity in silico using the netMHC epitope prediction software[74,75], to determine to what degree the candidate mutations are likely to result in the deimmunization of the most immunogenic epitopes in which they appear. This is a critical step as many mutations may have little effect on overall immunogenicity[76–78]. Screening for decreased peptide–major histocompatibility complex (MHC) class I binding filters out amino acid substitutions, which are likely immune neutral, substantially increasing the likelihood of functional hits with enough epitope variation to evade immune induction[78,79].

Next, to enable readout, we applied long-read nanopore sequencing to measure the results of the screens of our combinatorial libraries. This circumvents the limit of short target regions and obviates the need for barcodes altogether by single-molecule sequencing of the entire target gene, enabling library design strategies that can explore any region of the protein in combination with any other region without any complicated cloning procedures required to facilitate barcoding[80]. So far, the adoption of nanopore sequencing has been limited by its high error rate, around 95% accuracy per DNA base[81], compared with established short-read techniques, which are multiple orders of magnitude more accurate. To address this challenge, we designed our libraries such that each variant that we engineered would have multiple nucleotide changes for each single target amino acid change, effectively increasing the sensitivity of nanopore-based readouts with increasing numbers of nucleotide changes per library member. The large majority of amino acid substitutions are amenable to a library design paradigm in which each substitution is encoded by two, rather than one, nucleotide changes, owing to the degeneracy of the genetic code and the highly permissive third 'wobble' position of codons.

The scale of engineering that would be required to generate an effectively deimmunized Cas9 is not fully understood, as combinatorial deimmunization efforts at the scale of proteins thousands of amino acids long have not yet been possible. Therefore, to roughly estimate these parameters, we developed an immunogenicity scoring metric that takes into account all epitopes across a protein and the known diversity of MHC variants in a species weighted by population frequency to generate a single combined score representing the average immunogenicity of a full-length protein as a function of each of its immunogenic epitopes[82]. Formally, this score is calculated as

$$I_x = \frac{\Sigma_i^m \Sigma_j^n w_j \left(1 - \log\left(k_{ij} \times \hat{j}\right)\right)}{y}$$

where I_x is the immunogenicity score of protein x, i is the epitopes, j is the HLA alleles, $\hat{j}$ is the allele specific standardization coefficient, w_j is the HLA allele weights, k_{ij} is the predicted binding affinity of epitope i to allele j, and y is the protein specific scaling factor. We then predicted the overall effect of mutating the top epitopes in several Cas9 orthologues (Supplementary Fig. 3a). As might be expected, this analysis suggests that single-epitope strategies are woefully inadequate to deimmunize a whole protein for multiple HLA types, and also that there are diminishing returns as more and more epitopes are deimmunized. Our analysis suggests that it may require on the order of tens of deimmunized epitopes to make a notable impact on overall, population-wide protein immunogenicity. The scale of engineering demanded by these immunological facts has previously been intractable, but by applying LORAX, we conjectured that one could now make substantial steps, several mutations at a time, through the mutational landscape of the Cas9 protein.

Specifically, applying the procedure above, we designed a library of Cas9 variants based on the SpCas9 backbone containing 23 different mutations across 18 immunogenic epitopes (Fig. 5). Combining these in all possible combinations yields a library of 1,492,992 unique elements. With this design, we then constructed the library in a stepwise process. First, the full-length gene was broken up into short blocks of no more than 1,000 bp, which overlap by 30 bp on each end. Each block is designed such that it contains no more than four target epitopes to mutagenize. With few epitopes per block and few variant mutations per epitope, it becomes feasible to chemically synthesize each combination of mutations for each block. Each of these combinations was then synthesized and mixed at equal ratios to make a degenerate block mix. This was repeated for each of the blocks necessary to complete the full-length protein sequence via fusion PCR.

To identify functional variants still capable of editing DNA, we next designed and carried out a positive selection screen targeting the hypoxanthine phosphoribosyltransferase 1 (HPRT1) gene[83]. In the context of the screen, HPRT1 converts 6-thioguanine (6-TG), an analogue of the DNA base guanine, into 6-TG nucleotides that are cytotoxic to cells via incorporation into the DNA during S-phase[84]. Thus, only cells containing functional Cas9 variants capable of disrupting the HPRT1 gene can survive in 6-TG-containing cell culture media. To first identify the optimal 6-TG concentration, HeLa cells were transduced with lentivirus particles containing wild-type (WT) Cas9 and either an HPRT1-targeting guide RNA (gRNA) or a non-targeting guide. After selection with puromycin, cells were treated with 6-TG concentrations ranging from 0 μg ml^{-1} to 14 μg ml^{-1} for 1 week. Cells were stained with crystal violet at the end of the experiment and imaged. About 6 μg ml^{-1} was selected as all cells containing a non-targeting guide had died while cells containing the HPRT1 guide remained viable (Supplementary Fig. 3b).

To perform the screen, replicate populations of HeLa cells were transduced with lentiviral particles containing the variant SpCas9 library along with the HPRT1-targeting gRNA at 0.3 multiplicity of infection (MOI) and at greater than 75-fold coverage of the library elements. Cells were selected using puromycin after 2 days and 6-TG was added once cells reached 75% confluency. After 2 weeks, genomic DNA was extracted from remaining cells and full-length Cas9 amplicons were nanopore sequenced on the Oxford Nanopore (ONT) MinION platform.

MinION sequencing confirmed that the majority of the pre-screened library consists of Cas9 sequences with many mutations, with most falling into a broad peak between 6 and 14 mutations per sequence, each of which knocking out a key immunogenic epitope (Supplementary Fig. 3c). Interestingly, the post-screening library was significantly shifted in the mutation density distribution, suggesting that the majority of the library with large (>4) numbers of mutations resulted in non-functional proteins that were unable to survive the screen. Meanwhile, WT, single and double mutants were generally enriched as these proteins proved more likely to retain functionality and pass through the screen (Supplementary Fig. 3c). In addition, the two independent replicates of the screen showed strong correlation ($R^2 = 0.925$) providing further evidence of robustness (Fig. 5). We also analysed the change in overall frequency of mutations in the pre- and post-screen libraries to see if a pattern of mutation effects could be inferred. Although the WT allele was enriched at every site in the post-screen sequences, nearly every site retained a reasonable fraction of mutated alleles, suggesting that the mutations, at least individually, are fairly well tolerated and do not disrupt Cas9 functionality (Supplementary Fig. 3d).

To select hits for downstream validation and analysis, we devised a method for differentiating high-support hits likely to be real from noise-driven false-positive hits. To do this, we hypothesized that the fitness landscape of the screen mutants is likely to be smooth; that is, variants that contain similar mutations are more likely to have similar fitness in terms of editing efficiency compared with randomly selected pairs[85]. We confirmed this by computing a predicted screen score for

https://doi.org/10.1038/s41551-024-01245-z

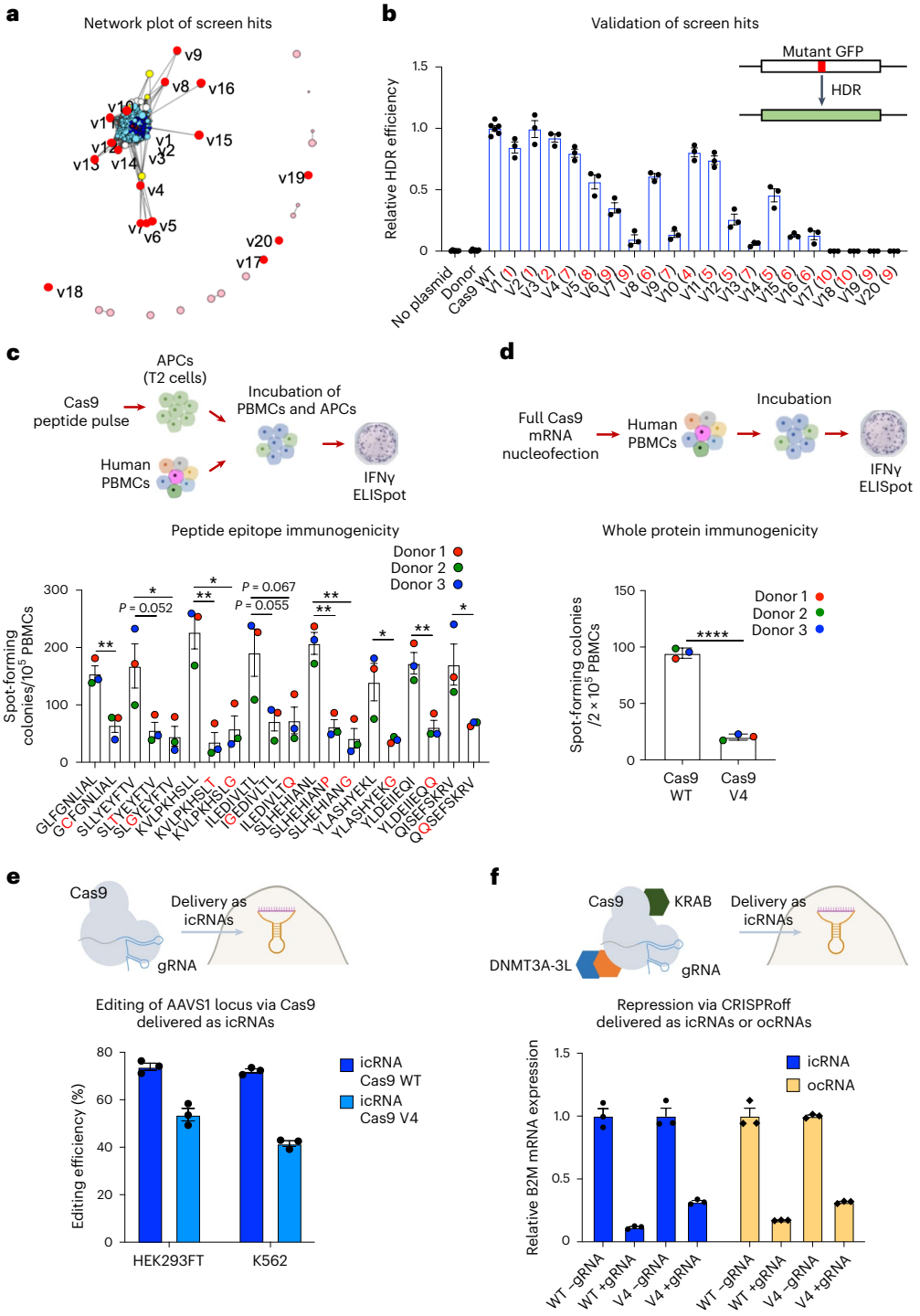


Fig. 6 | Validation of LORAX screen identified Cas9 variants for deimmunization, and genome and epigenome targeting via delivery as icRNAs and ocRNAs. a, Network reconstruction connecting Cas9 variants with similar mutational patterns. Node colours indicate the number of deimmunized epitopes (dark blue < 3, light blue = 3, white = 4, yellow = 5, pink > 5). Circles in red represent tested variants and labelled with their respective names. **b**, HEK293T bearing a GFP coding sequence disrupted by the insertion of a stop codon and a 68 bp genomic fragment of the AAVS1 locus were used as a reporter line. WT or Cas9 variants, an sgRNA targeting the AAVS1 locus and a donor plasmid capable of restoring GFP function via HDR were transfected into these cells and flow cytometry was performed on day 3. Relative quantification of GFP expression restoration by HDR is plotted. The number in parentheses represents the number of mutations in the variant. Values represented as mean ± s.e.m. ($n = 3$). **c**, T2 cells were pulsed with WT and variant peptides, cultured with PBMCs, and an ELISpot assay was performed to assess PBMC IFNγ secretion to WT and variant peptides. The number of spot-forming colonies for each peptide is plotted ($n = 3$, mean ± s.e.m., $^*P < 0.05$, $^{**}P < 0.01$, unpaired t-test, two-tailed). Red letters in the peptide sequences represent the mutated amino acid. **d**, RNA encoding for Cas9 WT or variant V4 was electroporated into PBMCs to assess the whole protein immunogenicity. ELISpot assay was performed to assess PBMC IFNγ secretion to WT and variant protein. The number of spot-forming colonies for each peptide is plotted ($n = 3$, mean ± s.e.m., $^{****}P < 0.0001$, unpaired t-test, two-tailed). **e**, Circular icRNA for Cas9 WT or variant V4, along with an sgRNA targeting the AAVS1 locus, was introduced into HEK293T and K562 cells. Editing efficiency at the AAVS1 locus in the two cell lines is plotted. Values represented as mean ± s.e.m. ($n = 3$). **f**, Circular icRNA and ocRNA for CRISPRoff WT or variant V4, along with an sgRNA targeting the B2M gene, were introduced into HEK293T cells. B2M gene repression of CRISPRoff constructs in the presence or absence of sgRNA is plotted. Values represented as mean ± s.e.m. ($n = 3$).

each variant based on a weighted regression of its nearest neighbours in the screen. This metric correlates well with the actual screen scores and approaches the screen scores even more closely as read coverage increases. This provides good evidence that the fitness landscape is indeed somewhat smooth (Supplementary Fig. 4a). Next, we reasoned that because the fitness landscape is smooth, real hits should reside in broad fitness peaks, which include many neighbours that also show high screen scores, whereas hits that are less supported by near neighbours are more likely to be spurious as they represent non-smooth fitness peaks. Formalizing this logic, we performed a network analysis to differentiate noise-driven hits from bona fide hits by looking at the degree of connectivity with other hits (Fig. 6a).

Applying these analyses to the screen output led us to select and construct 20 variants (V1–20) for validation and characterization. We applied two independent methods to quantify editing of the deimmunized Cas9 variants. First, we performed a gene-rescue experiment using low-frequency homology-directed repair (HDR) to repair a genetically encoded broken GFP gene[86] (Fig. 6b). Second, we quantified NHEJ-mediated editing by genomic DNA extraction and Illumina next-generation sequencing (NGS) using the CRISPResso2 package[87] (Supplementary Fig. 4b). Variants highly connected to neighbours were capable of editing, whereas those not connected were non-functional, validating the network-based approach that we used to select hits as enriching for truly functional sequences. Among the screen hits was the L616G mutation first identified in ref. 61 as a functional Cas9 variant with a critical immunodominant epitope deimmunized (V1). This concordance with previous work provided further confidence in our screening method. Interestingly, we discovered another deimmunizing mutation within the same epitope, L623Q (V2), which similarly retains Cas9 functionality, but appears to be more epistatically permissive, as many of our multi-mutation hits combine this mutation with other deimmunized epitopes. From these multi-mutation hits, we chose V4, which demonstrated high editing capability while still bearing simultaneous mutations across seven distinct epitopes, as well as family members V3, a variant bearing two mutations, and V5, a variant bearing the seven changes from V4 plus one additional mutation.

To confirm that mutation of these epitopes indeed elicited deimmunization, we assessed T cell response to WT and variant peptides by measuring IFNγ secretion in the ELISpot assay[17,61]. We chose to use peripheral blood mononuclear cells (PBMCs) from three separate donors that carried the HLA-A*0201 allele as peptides were presented to cells using the TAP-deficient cell line T2 (HLA-A*0201 positive)[88]. Correspondingly, we synthesized peptides for epitopes 2, 7, 8, 9, 12, 15 and 16 as our predictions suggested that these epitopes would induce a reduction in immune response for the HLA-A*0201 allele (Supplementary Fig. 5a). Importantly, since SpCas9v4 carries four of these mutations, this assay would also provide confirmation of deimmunization for this variant. We found that mutant peptides for all epitopes tested indeed resulted in fewer spot-forming colonies for all three donors compared with WT peptides (Fig. 6c and Supplementary Fig. 5b), thereby confirming our predictions. To assess whether the full-length protein is deimmunized, we next generated mRNA encoding for SpCas9WT and SpCas9V4 and electroporated this into the PBMC populations. As PBMCs have a mixture of both antigen-presenting cells (APCs) and T cells[89,90], we are able to introduce the RNA to the APCs and measure T cell response via the ELISpot assay. Excitingly, we observed significantly fewer spot forming colonies for full-length SpCas9v4 compared with SpCas9WT and similarly when compared with SpCas9 L616G (Fig. 6d and Supplementary Fig. 5c,d).

On the basis of this, we then further evaluated the efficacy of these mutants side by side with WT SpCas9 across a panel of genes and cell types, and assessed V4 activity across both targeted genome editing and epigenome regulation experiments[91] (Supplementary Fig. 6a–c). Together, these results confirmed that leveraging our unique combinatorial library design and screening strategy, we were able to produce Cas9 variants with multiple top immunogenic epitopes simultaneously mutated while still retaining genome targeting functionality. Spurred by this, we next evaluated delivery of SpCas9WT and SpCas9v4 and CRISPRoff versions of the same as icRNAs. CRISPRoff represents one of the newest additions to the CRISPR toolbox with the exciting capability to permanently silence gene expression upon transient expression[55].

We conjectured that Cas9 and CRISPRoff would represent exciting applications of cRNAs for hit-and-run genome and epigenome targeting, as the prolonged persistence could enable robust targeting, while the use of partially deimmunized Cas9 proteins would enable greater safety in therapeutic contexts. Towards this, icRNAs for WT SpCas9 or SpCas9v4, along with single guide RNAs (sgRNAs) targeting the AAVS1 locus, were transfected into HEK293Ts. Excitingly, we observed AAVS1 genome editing and approximately 50% relative circularization rates for icRNAs as quantified by RNAseq. In addition, icRNAs and ocRNAs for WT dSpCas9 or dSpCas9v4 CRISPRoff along with sgRNAs targeting the B2M gene were transfected into HEK293Ts[54], and we confirmed robust B2M gene repression via both cRNA formats (Fig. 6e–f). Importantly, the in vitro circularized ocRNAs were found to have a circularization efficiency of approximately 40% and 20% for SpCas9 and CRISPRoff inserts, respectively, as quantified via tapestation analyses. Lastly, to assess the specificity of SpCas9v4 targeting, we performed RNAseq on WT and SpCas9v4 CRISPRoff samples with and without the B2M guide. As expected, B2M was heavily downregulated in SpCas9WT and SpCas9v4 samples containing the sgRNA compared with samples with no guide (Supplementary Fig. 6d, red dot). Importantly, all differentially expressed genes (DEGs) for V4 were also DEGs for WT, suggesting that SpCas9v4 and SpCas9WT are comparably specific in this assay (Supplementary Fig. 6d, purple dots).

Discussion

To facilitate the pursuit of cRNAs, we developed two methods: 'outside developed' cRNAs (ocRNAs) via in vitro circularization using group II introns (which, to the best of our knowledge, is the first demonstration of gene of interest expression from cRNAs prepared by a group II intron), and in situ or 'inside developed' cRNAs (icRNAs) via in-cell circularization using the ubiquitously expressed RtcB protein[92,93]. Furthermore, we developed HPLC-free purification protocols for these enabling high yields while maintaining low immune responses. The resulting simplicity and scalability of production facilitated a range of applications from stem cell engineering to robust genome and epigenome targeting via ZF proteins and CRISPR–Cas9 systems. In particular, icRNAs and ocRNAs bearing the EMCV IRES demonstrated robust expression (compared with linear capped and modified RNAs) in cardiomyocytes and neurons, and also prolonged RNA persistence highlighting substantial promise and utility in non-dividing cells. In addition, our circularization strategies allowed for efficient generation and delivery of large constructs, such as Cas9 and CRISPRoff, which would be otherwise cumbersome to deploy via lentiviruses and adeno-associated viruses owing to packaging limits[55].

Concurrently, to enable compatibility between persistence of expression and immunogenicity, we also developed the LORAX protein engineering method that can be applied iteratively to tackle particularly challenging multiplexed protein engineering tasks by exploring huge swaths of combinatorial mutation space unapproachable using previous techniques. We demonstrated the power of this technique by creating a Cas9 variant with seven simultaneously deimmunized epitopes, which still retains functionality in a single round of screening. This opens up the application of gene editing to long-persistence therapeutic modalities such as AAV or icRNA delivery. Furthermore, whereas this methodology is particularly suited to the unique challenges of protein deimmunization, it is also applicable to any potential protein engineering goal, so long as there exists an appropriate screening procedure to select for the desired functionality.

While ocRNAs and icRNAs are a versatile system with broad applications, we anticipate multiple avenues for further engineering their efficacy: one, the upper limits of payload circularization for both icRNAs and ocRNAs require further exploring; two, despite comparable cell viability results, the full extent of immunogenicity presented by icRNAs and ocRNAs has yet to be explored in depth, and may serve to provide useful insights into future targeted improvements[9,94–96]; three, we observed persistence of cRNAs in non-dividing cells, but approaches to improve their persistence in dividing cells (where they are otherwise rapidly diluted in every cell division) could further broaden their utility; and four, as typically the 5′ cap was observed to be more efficient than IRESs in recruiting the translation machinery, comprehensive screens of tissue specificity of IRESs[97–101] will be crucial to deploy corresponding cRNAs in their optimal setting.

Similarly, the versatility of the LORAX method comes with a set of limitations and trade-offs that must be managed to leverage its utility. Naturally, library design is of critical importance. Here we have leveraged several features such as Cas9 evolutionary diversity, MHC-binding predictions, HLA allele frequencies and calculated immunogenicity scores to generate a useful library of variants to test. Other approaches may bring in more sources of information from places such as protein structure[102], coevolutionary epistatic constraints[103], amino acid signalling motifs[104] or T/B cell receptor binding repertoires[105], among other possibilities. Another critical factor is careful selection of hits downstream of screening, especially given the sparse coverage owing to nanopore sequencing. Here we have developed a network-based method for differentiating spurious from bona fide hits leveraging known aspects of protein epistasis and fitness landscapes. Similar customizations and tweaks relevant to the specific biology of a given problem may yield substantial returns in applying LORAX or other large-scale combinatorial screening methods to various protein engineering challenges.

Looking ahead, in addition to its core utility in applications entailing transgene delivery, we anticipate that cRNAs will be particularly useful in scenarios where a longer duration pulse of protein production is required. These include, for instance, epigenome engineering and cellular reprogramming, as well as transient healing and rejuvenation applications. Taken together, we anticipate that the highly simple and scalable ocRNA and icRNA methodologies could have broad utility in basic science and therapeutic applications.

Methods

Cell culture

HEK293T, A549, A375 and HeLa cells were cultured in DMEM supplemented with 10% FBS and 1% antibiotic–antimycotic (Thermo Fisher). K562 cells were cultured in RPMI supplemented with 10% FBS and 1% antibiotic–antimycotic (Thermo Fisher).

H1 human embryonic stem cells (hESCs) were maintained under feeder-free conditions in mTeSR1 medium (StemCell Technologies). Before passaging, tissue-culture plates were coated with growth factor-reduced Matrigel (Corning) diluted in DMEM/F-12 medium (Thermo Fisher Scientific). Cells were dissociated and passaged using the dissociation reagent Versene (Thermo Fisher Scientific) and passaged at a 1:4 ratio.

All cells were cultured in an incubator at 37 °C and 5% CO_2.

DNA transfections were performed by seeding HEK293T cells in 12-well plates at 25% confluency and adding 1 µg of each DNA construct and 4 µl of Lipofectamine 2000 (Thermo Fisher). RNA transfections were performed by adding a given pmol of each RNA construct and Lipofectamine MessengerMax (Thermo Fisher) (refer Supplementary Table 3 for specific details). Electroporations were performed in K562 cells using the SF Cell Line 4D-Nucleofector X Kit S (Lonza) per the manufacturer's protocol.

IVT. DNA templates for generating desired RNA products were created by PCR amplification from plasmids or gBlock gene fragments (IDT) and purified using a PCR purification kit (Qiagen). Plasmids were then generated with these templates containing a T7 promoter followed by 5′ ribozyme sequence, a 5′ ligation sequence, an IRES sequence linked to the product of interest, a 3′ UTR sequence, a 3′ ligation sequence, a 3′ ribozyme sequence and lastly a poly-T stretch to terminate transcription. Linearized plasmids were used as templates and RNA products were then produced using the HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB E2040) per the manufacturer's protocol. In icRNA conditions, urea was added to a final concentration of 0.8 M in a 20 µl reaction through a freshly prepared 6 M urea solution. Linear mRNA was produced using the HiScribe T7 mRNA Kit with CleanCap Reagent AG (NEB E2080) unless otherwise mentioned. All UTP was replaced with N^1-methylpseudouridine-5′-triphosphate (Trilink Biotechnologies, N-1081) for m1Ψ conditions.

Purification protocols. IVT RNA reactions were cleaned with the Monarch RNA Cleanup Kit (500 µg) (T2050) according to manufacturer's instructions. For icRNA and ocRNA, CIP treatment is preferred to ensure the removal of any remaining triphosphates from IVT, including cleaved twister ribozyme product. For ocRNA, cellulose chromatography was performed twice according to ref. 38 per 100 µg of RNA. Resulting RNA was again cleaned with the Monarch RNA Cleanup Kit (500 µg) (T2050) according to manufacturer's instructions. For gel-extracted samples, RNA was further separated on precast 2% E-Gel EX agarose gels and then gel extraction performed with the Monarch RNA Cleanup Kit (50 µg) (T2040) according to manufacturer's instructions. For RNaseR (Lucigen RNR07250)-treated samples, up to 89 µl of water was added to 20 µg of RNA, then heated to 70 °C for 3 min and briefly put on ice. RNaseR reaction buffer and 20 U of RNaseR were added and heated at 37 °C for 15 min, with an additional 10 U of RNaseR halfway. In the case of ref. 15, cRNA, no additional RNaseR was added and the reaction was heated at 37 °C for an hour. After RNaseR digestion, reactions were cleaned with the Monarch RNA Cleanup Kit (50 µg) (T2040) according to manufacturer's instructions.

In vitro immune and cell viability experiments. To assess immune response, A549s were transfected with uncapped and unmodified linear RNA, linear m1Ψ RNA, icRNA, ocRNA (column) and ocRNA (gel), processed via previously mentioned methods additionally outlined in the associated figure schematic, and then RNA isolated at 6 h, 24 h and 48 h. RT-qPCR was performed to quantify IFNB, RIG-I and IL6 mRNA levels relative to GAPDH. To assess cell viability, A549s were seeded and transfected with the same panel of RNA previously mentioned. A cell counting kit-8 (CCK-8) assay was performed before transfection on day 0, and then again on days 1 and 2, with media replacement before each measurement. For each well, 5 µl of CCK-8 reagent (Dojindo, CK04) was used per 100 µl of media and incubated at cell culture conditions for 1 h. Absorbance was measured at 450 nm on a plate reader.

In vitro persistence experiments. To assess persistence of circular icRNA, HEK293T cells were transfected with circular icRNA GFP or linear icdRNA and RNA was isolated 6 h, 1 day, 2 days and 3 days after transfection. RT-qPCR was performed to assess the amount of GFP RNA and RT-PCR was performed to confirm cRNA persistence in cells receiving icRNA.

Persistence of circular icRNA containing EMCV IRES, WPRE and a 50 adenosine poly(A) stretch compared with commercially sourced RNA (Trilink Biotechnologies, L-7601) with a 5′ cap and a poly(A) tail was similarly performed, with additional time points at days 4 and 5. RNA was isolated from cells and RT-qPCR was performed to assess the amount of GFP RNA.

For cardiomyocyte experiments, H1 hESCs were differentiated into cardiomyocytes using established protocols[106,107]. In brief, stem cells were dissociated using Accutase and seeded into 12-well Matrigel-coated plates. Cells were maintained in mTeSR1 (StemCell Technologies) for 3–4 days until cells reached about 95% confluence.

Media was changed to RPMI containing B27 supplement and 10 μM CHIR99021. After 24 h, media was changed to RPMI containing B27 supplement without insulin. Two days later, media was changed such that half of the cultured media was mixed with fresh RPMI containing B27 supplement without insulin and 5 μM IWP2. After 2 days, media was changed to RPMI containing B27 supplement without insulin. Media was then changed to RPMI containing B27 supplement every 2 days. After 6 days, media was replaced with cardiac metabolic enrichment media[108]. After 2 days, media was again changed to RPMI containing B27 supplement every 2 days. Cardiomyocytes were transfected with icRNA, ocRNA and icdRNA containing EMCV IRES, modified WPRE, and a 165 adenosine poly(A) stretch or linear m1Ψ mRNA 2 weeks after CHIR99021 induction. Five images were taken for each biological replicate of each condition at each time point and GFP intensity was quantified for representative images using FIJI (NIH).

For neuron experiments, H1 hPSC clonal lines overexpressing NEUROD1 were seeded and neural differentiation media consisting of 1:1 DMEM/F12–Neurobasal media (Thermo Fisher Scientific) + 100x Glutamax, 10 ng ml^{-1} BDNF (Peprotech), 10 ng ml^{-1} NT3 (Peprotech), 0.75 μg ml^{-1} puromycin, 1 μg ml^{-1} doxycycline (Sigma-Aldrich), 50x B27 supplement (Thermo Fisher Scientific) and 100x N2 supplement (Thermo Fisher Scientific) was added. Media was changed every 2 days. Three images were taken for each biological replicate of each condition at each time point and GFP intensity was quantified using FIJI (NIH).

In vitro differentiation experiments. H1 hPSCs were seeded on day −1 in a Matrigel-coated 24-well plate in mTeSR. On day 0, mTeSR media was replaced and cells were transfected with RNA. On day 3, 2 μM Ara-C was added to mTeSR media to eliminate dividing hPSCs. mTeSR media was replaced every day until RNA collection or staining.

Immunostaining. To prepare for staining, medium was removed and cells were washed with 1 ml of PBS and 500 μl of 4% paraformaldehyde solution subsequently added for fixation at room temperature for 1 h, shielded from light. Post-fixation, cells were washed once with 500 μl of PBS. Afterwards, 500 μl of blocking buffer (3% FBS, 1% BSA, 0.5% Triton-X and 0.5% Tween) was added and incubated for 1 h at room temperature. Following blocking, primary antibody (diluted 1:1,000 in blocking buffer) was added and incubated overnight at 4 °C, shielded from light. After overnight incubation, cells were washed three times for 5 min each before the application of secondary antibody (1:2,000 dilution in blocking buffer) and incubated for 1 h at room temperature, shielded from light. Following secondary staining, cells were washed three times for 5 min each with PBS. Antibodies used include anti-tubulin β 3 (TUBB3) (BioLegend, 801201) and anti-mouse IgG (H + L) Alexa Fluor Plus 647 (Invitrogen, A32728).

Flow cytometry experiments. To assess in vitro protein translation efficiencies, equimolar amounts of icRNA or linear mRNA were transfected into HEK293Ts, A549s, HeLas, hPSCs or A375s and GFP intensity was quantified 24 h later. GFP intensity, defined as the mean intensity of the cell population, was quantified after transfection using a BD LSRFortessa cell analyser.

Quantifying circular efficiency. To assess circular efficiency, icRNA containing EMCV IRES, WPRE and a 165 adenosine poly(A) stretch was generated. RNA was then either frozen or pre-circularized using the RTCB ligase (NEB M0458S) per manufacturer's instructions. To remove any linear RNA, pre-circularized RNA was treated with RNaseR (Lucigen RNR07250) per manufacturer's instructions. icRNA or pre-circularized icRNA was then transfected into HEK293Ts and RNA was isolated from cells at 6 h, 24 h and 48 h. RT-PCR was performed and the intensity of the circular band for icRNA compared with pre-circularized icRNA was defined as the circular efficiency. All circular intensity values were normalized to respective GAPDH band intensity.

To assess circularization efficiency of ocRNA, samples frozen overnight were run on an Agilent Tapestation to qualify ocRNA profiles. Then, gel analysis was performed on tapestation results to quantify circularization efficiency on ocRNA samples with band intensity normalization by nucleotide.

RNAseq. RNAseq was performed on HEK293T cells 6 h, 24 h and 48 h after transfection. Three biological replicates were sequenced. Total RNA was isolated from cells via an RNeasy kit (Qiagen) with on-column DNase I treatment. For RNA capture, 1 μg per sample of total RNA was used with NEBNext Poly(A) mRNA Magnetic Isolation Module (E7490S). Subsequently, an NEBNext Ultra RNA Library Prep Kit (E7530S) was used to generate Illumina-compatible RNAseq libraries. Sequencing was performed on an Illumina NovaSeq 6000, with paired end 100 bp reads. Reads were aligned using BWA-MEM2 to custom genomes derived from the GRCh38 Human Reference. Custom references were created using Cell Ranger to include GFP and Cas9 (720 bp) and ligation region (198 bp) for alignment. Featurecounts was used to quantify gene counts on BAM files output from BWA-MEM2. RPKM normalization was then performed and normalized counts evaluated to determine % circularization of transfected RNA.

LNP formulations. (6Z,9Z,28Z,31Z)-Heptatriaconta-6,9,28,31-tetraen-19-yl-4-(dimethylamino)butanoate (DLin-MC3-DMA) was purchased from BioFine International. 1,2-Distearoyl-*sn*-glycero-3-phosphocholine (DSPC) and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG-2000) were purchased from Avanti Polar Lipids. Cholesterol was purchased from Sigma-Aldrich. mRNA LNPs were formulated with DLin-MC3-DMA:cholesterol:DSPC:DMG-PEG at a mole ratio of 50:38.5:10:1.5 and a N/P ratio of 5.4. To prepare LNPs, lipids in ethanol and mRNA in 25 mM acetate buffer (pH 4.0) were combined at a flow rate of 1:3 in a PDMS staggered herringbone mixer[109,110]. The dimensions of the mixer channels were 200 by 100 μm, with herringbone structures 30 μm high and 50 μm wide. Immediately after formulation, 3 volumes of PBS was added and LNPs were purified in 100 kDa molecular weight cut-off (MWCO) centrifugal filters by exchanging the volume three times. Final formulations were passed through a 0.2 μm filter. LNPs were stored at 4 °C for up to 4 days before use. LNP hydrodynamic diameter and polydispersity index were measured by dynamic light scattering (Malvern NanoZS Zetasizer). The mRNA content and percent encapsulation were measured with a Quant-iT RiboGreen RNA Assay (Invitrogen) with and without Triton X-100 according to the manufacturer's protocol.

Animal experiments. All animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of the University of California, San Diego. All mice were acquired from Jackson Labs.

To confirm circularization of icRNA constructs in vivo, 10 μg of circular GFP icRNA or linear GFP icdRNA LNPs was injected retro-orbitally into C57BL/6J mice. After 3 and 7 days, livers were isolated and placed in RNAlater (Sigma-Aldrich). RNA was later isolated using QIAzol Lysis Reagent and purified using an RNeasy mini kit (Qiagen) according to the manufacturer's protocol. The amount of circularized RNA was assessed by RT-qPCR.

To assess icRNA persistence in vivo, equal concentration of icRNA containing EMCV, WPRE, and a 165 adenosine poly(A) stretch (15 μg LNPs for EMCV) or linear RNA was injected retro-orbitally into C57BL/6J mice. On day 7, livers were isolated and RNA was extracted. RT-qPCR was performed to assess mRNA expression among the conditions and RT-PCR was performed to ensure circularization for icRNA conditions.

Cas9 alignment and mutation selection. Naturally occurring variation in Cas9 sequence space was explored by aligning BLAST hits of the SpCas9 amino acid sequence. This set was then pruned by

removing truncated, duplicated or engineered sequences, and those sequences whose origin could not be determined. At specified immunogenic epitopes and key anchor residues, top alternative amino acids were obtained using frequency in the alignment weighted by overall sequence identity to the WT SpCas9 sequence, such that commonly occurring amino acid substitutions appearing in sequences highly similar to the WT were prioritized for further analysis and potential inclusion in the LORAX library.

HLA frequency estimation and binding predictions. HLA-binding predictions were carried out using netMHC4.1 or netMHCpan3.1. Global HLA allele frequencies were estimated from data at allelefrequencies.net as follows. Data were divided into 11 geographical regions. Allele frequencies for each of those regions were estimated from all available data from populations therein. These regional frequencies were then averaged weighted by global population contribution. Alleles with greater than 0.001% frequency in the global population, or those with greater than 0.01% in any region, were included for further analysis and predictions.

Immunogenicity scores. The vector of predicted nM affinities output by netMHC was first normalized across alleles to account for the fact that some alleles have higher affinity across all peptides and to allow for the relatively equivalent contribution of all alleles. These values were then transformed using the 1-log(affinity) transformation also borrowed from netMHC such that lower nM affinities will result in larger resulting values. These transformed, normalized affinities are then weighted by population allele frequency and summed across all alleles and epitopes. Finally, the scores are standardized across proteins to facilitate comparison.

Identification of HPRT1 guide. The lentiCRISPR-v2 plasmid (Addgene #52961) was first digested with Esp3I and a guide targeting the HPRT1 gene was cloned in via Gibson assembly. After lentivirus production, HeLa cells were seeded at 25% confluency in 96-well plates and transduced with virus (lentiCRISPR-v2 with or without HPRT1 guide) and 8 μg ml^{-1} polybrene (Millipore). Virus was removed the next day and 2.5 μg ml^{-1} puromycin was added to remove cells that did not receive virus 2 days later. After 2 days of puromycin selection, 0–14 μg ml^{-1} 6-TG was added. After 5 days, cells were stained with crystal violet, solubilized using 1% sodium dodecyl sulfate, and absorbance was measured at 595 nm on a plate reader. Owing to the lack of cells in the negative control, 6 μg ml^{-1} was chosen.

Generation of variant Cas9 library. Cas9 variant sequences were generated by separating the full-length gene sequence into small sections, where each section contained WT or variant Cas9 sequences. Degenerate pools of these gBlocks were PCR amplified and annealed together, yielding a final library size of 1,492,992 elements. Specifically, the Cas9 blocks used as input to the fusion PCR were synthesized as linear DNA. The nucleotide numbers that define the limits of each block are given in Supplementary Table 4. Note the 30 bp overlaps to enable fusion.

The upper row of pie charts (Supplementary Fig. 3) corresponds to the composition of each individually synthesized block, with each section corresponding to a singular block sequence. In the lower row, each pie corresponds to a single mutation site (note that one block may contain up to three mutation sites depending on the block). Once the fusion PCR was performed, the full-length library elements were purified by size selection using AmpureXP magnetic beads formulated to a very low 0.4× concentration to enable selection of only high-molecular-weight DNA greater than ~3 kb. This insert was cloned into lentiCRISPRv2 (containing the HPRT1 guide) using Gibson assembly, and the plasmid product purified by 30 min dialysis before electroporation. Electroporation was done using NEB Stbl3 cells made electrocompetent as follows. A 20 ml overnight liquid culture was inoculated from a single clone picked from an antibiotic-free plate streaked with Stbl3s. This culture was used to inoculate two 1 l flasks, which were grown to a target OD_{600} of 0.4. These cultures were centrifuged, washed with pure, cold water, and concentrated 10× to 2× 100 ml. This was repeated twice more to yield a final volume of 2 × 1 ml electrocompetent cells. These cells were electroporated in 200 μl aliquots, each with 100 ng of purified assembled library DNA. Dilutions of each of these electroporations were plated to estimate transformation efficiency before pooling into the final library. Plasmid DNA was isolated using the Qiagen Plasmid Maxi Kit and this DNA was then used to create lentivirus containing the variant Cas9 library.

Cas9 screen. HeLa cells were seeded in 15 15 cm plates at a density of 10 million cells per plate and transduced with virus containing the variant Cas9 library and 8 μg ml^{-1} polybrene the next day at an MOI of 0.3. Media was changed the next day and 2.5 μg ml^{-1} puromycin was added to remove cells that did not receive virus 2 days later. Once cells reached 90% confluency, 6 μg ml^{-1} 6-TG was added to media. Media was changed every other day for 10 days to allow for selection of cells containing functional Cas9 variants. After 10 days, cells were lifted from the plates and DNA was isolated using the DNeasy Blood and Tissue Kit per the manufacturer's protocol.

Nanopore sequencing. Pre-screen analysis of the Cas9 variant library elements was performed by amplifying the sequence from the plasmid. About 1 μg of the variant Cas9 sequences was used for library preparation using the Ligation Sequencing Kit (Oxford Nanopore Technologies, SQK-LSK109) per manufacturer's instructions. DNA was then loaded into a MinION flow cell (Oxford Nanopore Technologies, R9.4.1). Post-screen analysis of library elements was performed by amplifying the Cas9 sequences from 75 μg of genomic DNA. About 1 μg of the variant Cas9 sequences was similarly prepared using the Ligation Sequencing Kit and sequenced on a MinION flow cell.

Base calling and genotyping. Raw reads coming off the MinION flow cell were base-called using Guppy 3.6.0 and aligned to an SpCas9 reference sequence containing non-informative NNN bases at library mutation positions, so as not to bias calling towards WT or mutant library members, using Minimap2's map-ont presets. Reads covering the full length of the Cas9 gene with high mapping quality were genotyped at each individual mutation site and tabulated to the corresponding library member. Reads with ambiguous sites were excluded from further analysis.

Cluster analysis. Network analysis was performed by first thresholding genotypes to include only those identified as hits from the screen. These were genotypes appearing in the pre-screen plasmid library, both post-screen replicates, and having a fold change enrichment larger than the WT sequence (4.5-fold enrichment). These hits were used to create a graph with nodes corresponding to genotypes and node sizes corresponding to fold change enrichment. Edges were placed between nodes at most 4 mutations distant from each other, and edge weights were defined by $1/d$, where d is the distance between genotypes. Network analysis was done using Python bindings of igraph. Plots were generated using the Fruchterman–Reingold force-directed layout algorithm.

HDR validation. Lentivirus was produced from a plasmid containing a GFP sequence with a stop codon and 68 bp AAVS1 fragment. HEK293T cells were treated with 8 μg ml^{-1} polybrene and lentivirus. After puromycin selection to create a stable line, cells were transfected with plasmids containing variant Cas9 sequences, a guide targeting the AAVS locus and a GFP repair donor plasmid. After 3 days, FACS was performed and percent GFP-positive cells were quantified.

Genome engineering experiments. To validate variant Cas9 functional cutting, variant Cas9 and guides were transfected into HEK293T cells. After 2 days, genomic DNA was isolated. Genomic DNA was also isolated after 2 days from K562 cells after electroporation. To assess activity of CCR5 ZFNs delivered as icRNAs, HEK293Ts were transfected with circular icRNA or linear icdRNA and genomic DNA was isolated after 3 days. Assessment of GFP ZFN was performed by transfecting HEK293Ts stably expressing a broken GFP with circular icRNA or linear icdRNA and isolating genomic DNA after 3 days. To assess activity of Cas9 delivered as icRNAs, HEK293Ts and K562 were transfected or nucleofected with Cas9 WT or Cas9 v4 along with a gRNA (synthesized via Synthego) and genomic DNA was isolated after 3 days.

ZF experiments were performed by transfecting HEK293T cells with 0.5 µg of left and right arms of each ZF as either icRNA or icdRNA. After 3 days, genomic DNA was isolated.

Epigenome engineering experiments. ZF-KRAB icRNA experiments were performed by transfecting HeLa cells with icRNA encoding various in-house designed ZF sequences targeting the hPSCK9 gene. ZFs consisted of six variable DNA contacting regions as shown in Supplementary Table 5, inserted into the following backbone.

MAPKKKRKVGIHGVPAAMAERPFQCRICMRNFS(**ZF1**)HIRTHTGEK PFACDICGRKFA(**ZF2**)HTKIHTGSQKPFQCRICMRNFS(**ZF3**)HIRTHTGE KPFACDICGRKFA(**ZF4**)HTKIHTGSQKPFQCRICMRNFS(**ZF5**)HIRTHTG EKPFACDICGRKFA(**ZF6**)HTKIHLRQKDAARGS

RNA was isolated 2 days later and repression of hPCSK9 was assessed by RT-qPCR. Similarly, 3A3L-ZF-KRAB ocRNA experiments were performed by transfecting HeLa cells with ocRNA encoding 3A3L-ZF-KRAB (specifically ZF10) targeting the hPSCK9 gene. Cells were passaged at 25% every 2 days or at 80–90% confluency, and the remaining 75% of cells at each passage collected for RNA isolation and repression of hPCSK9 was assessed by RT-qPCR. dCas9-VPR experiments were performed by transfecting HEK293T cells with dCas9wt-VPR or dCas9v4-VPR with or without a gRNA targeting the ASCL1 gene. Likewise, KRAB-dCas9 experiments were performed by transfecting cells with KRAB-dCas9wt or KRAB-dCas9v4 with or without a gRNA targeting the CXCR4 gene. CRISPRoff experiments were performed by transfecting HEK293T cells with icRNA CRISPRoffwt or CRISPRoffv4 with or without a gRNA targeting the B2M gene (Synthego). RNA was isolated 3 days later and repression or activation of genes was assessed by RT-qPCR.

Quantification of editing using NGS. After extraction of genomic DNA, PCR was performed to amplify the target site. Amplicons were then indexed using the NEBNext Multiplex Oligos for the Illumina kit (NEB). Amplicons were then pooled and sequenced using a Miseq Nano with paired end 150 bp reads. Editing efficiency was quantified using CRISPResso2.

Cas9 specificity. RNA isolated from the CRISPRoff experiment was used to assess specificity. RNAseq libraries were generated from 300 ng of RNA using the NEBNext Poly(A) mRNA magnetic isolation module and NEBNext Ultra II Directional RNA Library Prep kit for Illumina and sequenced on the Illumina NovaSeq 6000 with paired end 100 bp reads. Fastq files were mapped to the reference human genome hg38 using STAR aligner. Differential gene expression was analysed using the Bioconductor package DESeq2 with the cut-off of $\log_2$(fold change) greater than 0.5 or less than −0.5 and a P value less than 10^{-3}. To identify DEGs, CRISPRoff WT and V4 samples containing the B2M guide were compared with samples not receiving the guide.

ELISpot assay. TAP-deficient T2 cells were a generous gift from Stephen Schoenberger lab. PBMCs were purchased from StemCell Technologies. All donors contained the HLA-A*0201 allele. Both cell lines were maintained in RPMI1640 media supplemented with 10% FBS, 1% penicillin–streptomycin, 10 mM HEPES and 1 mM sodium pyruvate. On the first day, PBMCs were thawed and rested overnight at a density of 10^6 cells per ml. T2 cells were pulsed with peptides at 10 µg ml^{-1} overnight. Peptides were produced from Genscript's Custom Peptide Synthesis service at crude purity. Lastly, 96-well plates (Immobilon-P, Millipore) were coated with 10 µg ml^{-1} anti-IFNγ monoclonal antibody (1-D1K, Mabtech) overnight at 4 C. The next day, T2 cells were washed two times and 50,000 T2 cells and 100,000 PBMCs were added to each well. Four replicates were used per condition. After 22 h, cells were removed from the plate and 2 µg ml^{-1} biotinylated anti-IFNγ secondary antibody (7-B6-1, Mabtech) was added for 2 h. Plates were washed and 1:1,000 streptavidin–ALP (3310-10-1000, Mabtech) was added for 45 min. Plates were washed and colour was developed by adding BCIP/NBT-plus substrate (3650-10, Mabtech) for 10 min. Plates were thoroughly washed with water and dried at room temperature, and spots were automatically counted using an ELISpot plate reader.

To assess the immunogenicity of the full-length Cas9 WT and variant protein, in vitro transcribed RNA encoding for WT, L616G or V4 was electroporated into PBMCs as previously described[89,90]. As PBMCs contain both APCs and T cells, it is possible to electroporate RNA directly into these APCs and assess T cell response via the ELISpot. Electroporation was performed using the P3 Primary Cell 4D-Nucleofector X Kit (Lonza V4XP). In brief, PBMCs were first thawed and rested overnight at a density of 10^6 cells per ml. The next day, 1×10^6 PBMCs were resuspended in 20 µl of Lonza P3 nucleofector solution and mixed with 1 µg RNA. In some cases, electroporation was performed using the Lonza P3 Primary Cell 4D-Nucleofector X Kit L, with 3×10^6 PBMCs resuspended in 100 µl of Lonza P3 nucleofector solution and mixed with 6 µg RNA. After electroporation, 2×10^5 cells were added to each well of an ELISpot plate already coated with anti-IFNγ monoclonal antibody as described above. After 28 h, cells were removed from the plate and the ELISpot assay and analysis was performed as described.

Lentivirus production. HEK293FT cells were seeded in 1 15 cm plate and transfected with 36 µl Lipofectamine 2000, 3 µg pMD2.G (Addgene #12259), 12 µg pCMV delta R8.2 (Addgene #12263) and 9 µg of the lentiCRISPR-v2 plasmid. The supernatant containing viral particles was collected after 48 h and 72 h, filtered with 0.45 µm Steriflip filters (Millipore), concentrated to a final volume of 1 ml using an Amicon Ultra-15 centrifugal filter unit with a 100,000 nominal molecular weight limit (NMWL) cut-off (Millipore) and frozen at −80 C.

RT-qPCR. cDNA was synthesized from RNA using the Protoscript II First Strand cDNA Synthesis Kit (NEB). qPCR was performed using a CFX Connect Real Time PCR Detection System (Bio-Rad). All samples were run in triplicates and results were normalized against GAPDH expression. Primers for qPCR are listed in Supplementary Table 6.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All key reagents will be made available via Addgene. Source data are provided with this paper.

Code availability

The code is available at https://github.com/natepalmer/lorax.

References

1. Karikó, K., Muramatsu, H., Ludwig, J. & Weissman, D. Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Res.* **39**, e142 (2011).

2. Presnyak, V. et al. Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111–1124 (2015).
3. Kuhn, A. N. et al. Phosphorothioate cap analogs increase stability and translational efficiency of RNA vaccines in immature dendritic cells and induce superior immune responses in vivo. *Gene Ther.* **17**, 961–971 (2010).
4. Holtkamp, S. et al. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood* **108**, 4009–4017 (2006).
5. Orlandini von Niessen, A. G. et al. Improving mRNA-based therapeutic gene delivery by expression-augmenting 3′ UTRs identified by cellular library screening. *Mol. Ther.* **27**, 824–836 (2019).
6. Wesselhoeft, R. A., Kowalski, P. S. & Anderson, D. G. Engineering circular RNA for potent and stable translation in eukaryotic cells. *Nat. Commun.* **9**, 2629 (2018).
7. Petkovic, S. & Müller, S. RNA circularization strategies in vivo and in vitro. *Nucleic Acids Res.* **43**, 2454–2465 (2015).
8. Müller, S. & Appel, B. In vitro circularization of RNA. *RNA Biol.* **14**, 1018–1027 (2017).
9. Wesselhoeft, R. A. et al. RNA circularization diminishes immunogenicity and can extend translation duration in vivo. *Mol. Cell* **74**, 508–520.e4 (2019).
10. Abe, N. et al. Rolling circle translation of circular RNA in living human cells. *Sci. Rep.* **5**, 16435 (2015).
11. Fan, X. et al. Pervasive translation of circular RNAs driven by short IRES-like elements. *Nat. Commun.* **13**, 3751 (2022).
12. Hansen, T. B. et al. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
13. Jeck, W. R. & Sharpless, N. E. Detecting and characterizing circular RNAs. *Nat. Biotechnol.* **32**, 453–461 (2014).
14. Kameda, S., Ohno, H. & Saito, H. Synthetic circular RNA switches and circuits that control protein expression in mammalian cells. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkac1252 (2023).
15. Chen, R. et al. Engineering circular RNA for enhanced protein production. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01393-0 (2022).
16. Li, A. et al. AAV-CRISPR gene editing is negated by pre-existing immunity to Cas9. *Mol. Ther.* **28**, 1432–1441 (2020).
17. Charlesworth, C. T. et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
18. Chaudhary, N., Weissman, D. & Whitehead, K. A. mRNA vaccines for infectious diseases: principles, delivery and clinical translation. *Nat. Rev. Drug Discov.* **20**, 817–838 (2021).
19. Corbett, K. S. et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **586**, 567–571 (2020).
20. Saunders, K. O. et al. Neutralizing antibody vaccine for pandemic and pre-emergent coronaviruses. *Nature* **594**, 553–559 (2021).
21. Thomas, S. J. et al. Efficacy and safety of the BNT162b2 mRNA COVID-19 vaccine in participants with a history of cancer: subgroup analysis of a global phase 3 randomized clinical trial. *Vaccine* https://doi.org/10.1016/j.vaccine.2021.12.046 (2021).
22. Zinsli, L. V., Stierlin, N., Loessner, M. J. & Schmelcher, M. Deimmunization of protein therapeutics—recent advances in experimental and computational epitope prediction and deletion. *Comput. Struct. Biotechnol. J.* **19**, 315–329 (2021).
23. McNeil, B. A., Simon, D. M. & Zimmerly, S. Alternative splicing of a group II intron in a surface layer protein gene in *Clostridium tetani*. *Nucleic Acids Res.* **42**, 1959–1969 (2013).
24. Pyle, A. M. Group II intron self-splicing. *Annu. Rev. Biophys.* **45**, 183–205 (2016).
25. Zimmerly, S. & Semper, C. Evolution of group II introns. *Mob. DNA* **6**, 7 (2015).
26. Chen, C. Y. & Sarnow, P. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268**, 415–417 (1995).
27. Jang, S. K. et al. A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* **62**, 2636–2643 (1988).
28. Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **19**, 568–576 (2012).
29. Alkemar, G. & Nygård, O. Secondary structure of two regions in expansion segments ES3 and ES6 with the potential of forming a tertiary interaction in eukaryotic 40S ribosomal subunits. *RNA* **10**, 403–411 (2004).
30. Bhat, P. et al. The beta hairpin structure within ribosomal protein S5 mediates interplay between domains II and IV and regulates HCV IRES function. *Nucleic Acids Res.* **43**, 2888–2901 (2015).
31. Chen, J. et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140–1146 (2020).
32. Hershey, J. W. B., Sonenberg, N. & Mathews, M. B. Principles of translational control: an overview. *Cold Spring Harb. Perspect. Biol.* **4**, a011528 (2012).
33. Bradrick, S. S., Dobrikova, E. Y., Kaiser, C., Shveygert, M. & Gromeier, M. Poly(A)-binding protein is differentially required for translation mediated by viral internal ribosome entry sites. *RNA* **13**, 1582–1593 (2007).
34. Machida, K. et al. Dynamic interaction of poly(A)-binding protein with the ribosome. *Sci. Rep.* **8**, 17435 (2018).
35. Mailliot, J. & Martin, F. Viral internal ribosomal entry sites: four classes for one goal. *Wiley Interdiscip. Rev.* **9**, e1458 (2018).
36. Imai, S., Kumar, P., Hellen, C. U. T., D'Souza, V. M. & Wagner, G. An accurately preorganized IRES RNA structure enables eIF4G capture for initiation of viral translation. *Nat. Struct. Mol. Biol.* **23**, 859–864 (2016).
37. Piao, X. et al. Double-stranded RNA reduction by chaotropic agents during in vitro transcription of messenger RNA. *Mol. Ther. Nucleic Acids* **29**, 618–624 (2022).
38. Baiersdörfer, M. et al. A facile method for the removal of dsRNA contaminant from in vitro-transcribed mRNA. *Mol. Ther. Nucleic Acids* **15**, 26–35 (2019).
39. Plank, T.-D. M., Whitehurst, J. T. & Kieft, J. S. Cell type specificity and structural determinants of IRES activity from the 5′ leaders of different HIV-1 transcripts. *Nucleic Acids Res.* **41**, 6698–6714 (2013).
40. Jayaraman, M. et al. Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo. *Angew. Chem. Int. Ed. Engl.* **51**, 8529–8533 (2012).
41. Sabnis, S. et al. A novel amino lipid series for mRNA delivery: improved endosomal escape and sustained pharmacology and safety in non-human primates. *Mol. Ther.* **26**, 1509–1519 (2018).
42. Lombardo, A. et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* **25**, 1298–1306 (2007).
43. Zou, J. et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* **5**, 97–110 (2009).
44. Abifadel, M. et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat. Genet.* **34**, 154–156 (2003).
45. Maxwell, K. N. & Breslow, J. L. Adenoviral-mediated expression of Pcsk9 in mice results in a low-density lipoprotein receptor knockout phenotype. *Proc. Natl Acad. Sci. USA* **101**, 7100–7105 (2004).
46. Cohen, J. C., Boerwinkle, E., Mosley, T. H. Jr & Hobbs, H. H. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006).

47. Thakore, P. I. et al. RNA-guided transcriptional silencing in vivo with *S. aureus* CRISPR-Cas9 repressors. *Nat. Commun.* **9**, 1674 (2018).
48. Ran, F. A. et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191 (2015).
49. He, N.-Y. et al. Lowering serum lipids via PCSK9-targeting drugs: current advances and future perspectives. *Acta Pharmacol. Sin.* **38**, 301–311 (2017).
50. Ridker, P. M. et al. Cardiovascular efficacy and safety of bococizumab in high-risk patients. *N. Engl. J. Med.* **376**, 1527–1539 (2017).
51. Sabatine, M. S. et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N. Engl. J. Med.* **376**, 1713–1722 (2017).
52. Fitzgerald, K. et al. A highly durable RNAi therapeutic inhibitor of PCSK9. *N. Engl. J. Med.* **376**, 41–51 (2017).
53. Ding, Q. et al. Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* **115**, 488–492 (2014).
54. Amabile, A. et al. Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell* **167**, 219–232.e14 (2016).
55. Nuñez, J. K. et al. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503–2519.e17 (2021).
56. Moreno, A. M. et al. Author correction: immune-orthogonal orthologues of AAV capsids and of Cas9 circumvent the immune response to the administration of gene therapy. *Nat. Biomed. Eng.* **3**, 842 (2019).
57. Chew, W. L. et al. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016).
58. Jawa, V. et al. T-cell dependent immunogenicity of protein therapeutics pre-clinical assessment and mitigation–updated consensus and review 2020. *Front. Immunol.* **11**, 1301 (2020).
59. Moghadam, F. et al. Synthetic immunomodulation with a CRISPR super-repressor in vivo. *Nat. Cell Biol.* **22**, 1143–1154 (2020).
60. Hakim, C. H. et al. Cas9-specific immune responses compromise local and systemic AAV CRISPR therapy in multiple dystrophic canine models. *Nat. Commun.* **12**, 6769 (2021).
61. Ferdosi, S. R. et al. Multifunctional CRISPR-Cas9 with engineered immunosilenced human T cell epitopes. *Nat. Commun.* **10**, 1842 (2019).
62. Allen, B. D., Nisthal, A. & Mayo, S. L. Experimental library screening demonstrates the successful application of computational protein design to large structural ensembles. *Proc. Natl Acad. Sci. USA* **107**, 19838–19843 (2010).
63. Sun, M. G. F., Seo, M.-H., Nim, S., Corbi-Verge, C. & Kim, P. M. Protein engineering by highly parallel screening of computationally designed variants. *Sci. Adv.* **2**, e1600692 (2016).
64. Cao, J. et al. High-throughput 5′ UTR engineering for enhanced protein production in non-viral gene therapies. *Nat. Commun.* **12**, 4138 (2021).
65. Hu, J. H. et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
66. Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296 (2020).
67. Kleinstiver, B. P. et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
68. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
69. Charles, E. J. et al. Engineering improved Cas13 effectors for targeted post-transcriptional regulation of gene expression. Preprint at *bioRxiv* https://doi.org/10.1101/2021.05.26.445687 (2021).
70. Griswold, K. E. & Bailey-Kellogg, C. Design and engineering of deimmunized biotherapeutics. *Curr. Opin. Struct. Biol.* **39**, 79–88 (2016).
71. Doud, M. B., Lee, J. M. & Bloom, J. D. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat. Commun.* **9**, 1386 (2018).
72. Gasiunas, G. et al. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* **11**, 5512 (2020).
73. Takeuchi, N., Wolf, Y. I., Makarova, K. S. & Koonin, E. V. Nature and intensity of selection pressure on CRISPR-associated genes. *J. Bacteriol.* **194**, 1216–1225 (2012).
74. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32**, 511–517 (2016).
75. Nielsen, M. et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017 (2003).
76. Osipovitch, D. C. et al. Design and analysis of immune-evading enzymes for ADEPT therapy. *Protein Eng. Des. Sel.* **25**, 613–623 (2012).
77. Choi, Y., Verma, D., Griswold, K. E. & Bailey-Kellogg, C. in *Computational Protein Design* (ed. Samish, I.) 375–398 (Springer New York, 2017).
78. King, C. et al. Removing T-cell epitopes with computational protein design. *Proc. Natl Acad. Sci. USA* **111**, 8577–8582 (2014).
79. Mazor, R. et al. Elimination of murine and human T-cell epitopes in recombinant immunotoxin eliminates neutralizing and anti-drug antibodies in vivo. *Cell. Mol. Immunol.* **14**, 432–442 (2017).
80. Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* **39**, 1348–1365 (2021).
81. Rang, F. J., Kloosterman, W. P. & de Ridder, J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* **19**, 90 (2018).
82. Schubert, B. et al. Population-specific design of de-immunized protein biotherapeutics. *PLoS Comput. Biol.* **14**, e1005983 (2018).
83. Liao, S., Tammaro, M. & Yan, H. Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. *Nucleic Acids Res.* **43**, e134 (2015).
84. Yang, F. et al. HPRT1 activity loss is associated with resistance to thiopurine in ALL. *Oncotarget* **9**, 2268–2278 (2018).
85. Meini, M.-R., Tomatis, P. E., Weinreich, D. M. & Vila, A. J. Quantitative description of a protein fitness landscape based on molecular features. *Mol. Biol. Evol.* **32**, 1774–1787 (2015).
86. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
87. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
88. Ninkovic, T. et al. Identification of *O*-glycosylated decapeptides within the MUC1 repeat domain as potential MHC class I (A2) binding epitopes. *Mol. Immunol.* **47**, 131–140 (2009).
89. Etschel, J. K. et al. HIV-1 mRNA electroporation of PBMC: a simple and efficient method to monitor T-cell responses against autologous HIV-1 in HIV-1-infected patients. *J. Immunol. Methods* **380**, 40–55 (2012).
90. Van Camp, K. et al. Efficient mRNA electroporation of peripheral blood mononuclear cells to detect memory T cell responses for immunomonitoring purposes. *J. Immunol. Methods* **354**, 1–10 (2010).
91. Moreno, A. M. et al. In situ gene therapy via AAV-CRISPR-Cas9-mediated targeted gene regulation. *Mol. Ther.* **28**, 1931 (2020).

92. Litke, J. L. & Jaffrey, S. R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat. Biotechnol.* **37**, 667–675 (2019).
93. Katrekar, D. et al. Efficient in vitro and in vivo RNA editing via recruitment of endogenous ADARs using circular guide RNAs. *Nat. Biotechnol.* **40**, 938–945 (2022).
94. Chen, Y. G. et al. Sensing self and foreign circular RNAs by intron identity. *Mol. Cell* **67**, 228–238.e5 (2017).
95. Chen, Y. G. et al. N6-Methyladenosine modification controls circular RNA immunity. *Mol. Cell* **76**, 96–109.e9 (2019).
96. Abe, B. T. et al. Circular RNA migration in agarose gel electrophoresis. *Mol. Cell* **82**, 1768–1777 (2022).
97. Chen, C.-K. et al. Structured elements drive extensive circular RNA translation. *Mol. Cell* **81**, 4300–4318.e13 (2021).
98. Yang, Y. et al. Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res.* **27**, 626–641 (2017).
99. Meyer, K. D. et al. 5′ UTR m6A promotes cap-independent translation. *Cell* **163**, 999–1010 (2015).
100. Weingarten-Gabbay, S. et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* **351**, aad4939 (2016).
101. Sample, P. J. et al. Human 5′ UTR design and variant effect prediction from a massively parallel translation assay. *Nat. Biotechnol.* **37**, 803–809 (2019).
102. Stiffler, M. A. et al. Protein structure from experimental evolution. *Cell Syst.* **10**, 15–24.e5 (2020).
103. Green, A. G. et al. Large-scale discovery of protein interactions at residue resolution using co-evolution calculated from genomic sequences. *Nat. Commun.* **12**, 1396 (2021).
104. Saylor, K., Gillam, F., Lohneis, T. & Zhang, C. Designs of antigen structure and composition for improved protein-based vaccine efficacy. *Front. Immunol.* **11**, 283 (2020).
105. Joglekar, A. V. et al. T cell antigen discovery via signaling and antigen-presenting bifunctional receptors. *Nat. Methods* **16**, 191–198 (2019).
106. Lian, X. et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat. Protoc.* **8**, 162–175 (2013).
107. Kumar, A. et al. Mechanical activation of noncoding-RNA-mediated regulation of disease-associated phenotypes in human cardiomyocytes. *Nat. Biomed. Eng.* **3**, 137–146 (2019).
108. Tohyama, S. et al. Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. *Cell Stem Cell* **12**, 127–137 (2013).
109. Chen, D. et al. Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation. *J. Am. Chem. Soc.* **134**, 6948–6951 (2012).
110. Belliveau, N. M. et al. Microfluidic synthesis of highly potent limit-size lipid nanoparticles for in vivo delivery of siRNA. *Mol. Ther. Nucleic Acids* **1**, e37 (2012).

Acknowledgements

We thank members of the Mali lab for discussions, advice and help with experiments. We also thank T. Long, S. Brightman and A. Sutherland for their advice on performing the ELISpot assay. This work was generously supported by UCSD Institutional Funds, NIH grants (R01HG012351, OT2OD032742, R01NS131560, U54CA274502 and DP2NS111507), Department of Defense Grant (W81XWH-22-1-0401), a Longevity Impetus Grant from Norn Group, a UC San Diego Gene Therapy Initiative Grant (GTI-2024-018), and an American Heart Association Postdoctoral Fellowship (AHA 916973). This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (S10 OD026929). This work was performed in part at the San Diego Nanotechnology Infrastructure (SDNI) of UCSD, a member of the National Nanotechnology Coordinated Infrastructure, which is supported by the National Science Foundation (grant ECCS-2025752). Some schematics were created using BioRender.

Author contributions

Conceptualization: M.T., N.P., A.K. and P.M. Experiments: M.T., N.P., A.K., A.D., H.K., S.H., E.F., M.W., C.H., Y.X., K.M., A.P., J.R., A.S., S.N. and P.M. Computational analyses: M.T. and N.P. Design: M.T., N.P., W.L.C., E.J.K. and P.M. Writing: M.T., N.P., A.K. and P.M. with input from all authors.

Competing interests

The authors have filed patents based on this work. P.M. is a scientific co-founder of Shape Therapeutics, Navega Therapeutics, Pi Bio, Boundless Biosciences and Engine Biosciences. The terms of these arrangements have been reviewed and approved by the University of California, San Diego, in accordance with its conflict-of-interest policies. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41551-024-01245-z.

Correspondence and requests for materials should be addressed to Prashant Mali.

Peer review information *Nature Biomedical Engineering* thanks the anonymous reviewers for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Prashant Mali

Last updated by author(s): May 17, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	GraphPad prism version 9.0.0 was used for plotting figures and computing the associated p values. FIJI ImageJ was used for cardiomyocyte image Calculated Total Cell Fluorescence (CTCF) analysis. CRISPResso2.1.0 was used for the analysis of amplicon sequencing. BWA-MEM2 and Featurecounts packages were used for RNA-Seq analysis. netMHC4.1 and netMHCpan3.1 were used to carry out HLA-binding predictions. Custom code is available at https://github.com/natepalmer/lorax.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All raw data will be made available via GEO, and all key reagents via Addgene. Source data for the figures are provided with this paper.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	The study did not involve human research participants.
Population characteristics	–
Recruitment	–
Ethics oversight	–

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	At least 2 independent samples per tested condition were evaluated for most experiments. No statistical methods were used to predetermine sample size. Cardiomyocyte CTCF values were obtained from a single representative space, and thus image, of each condition sampled repeatedly over 30 days.
Data exclusions	No data were excluded.
Replication	Two or more independent biological replicates were evaluated for both in vitro studies in cultures cells and in vivo studies using mice. The findings were consistent in all replicates.
Randomization	Mice were randomly assigned into groups prior to injection. For cell-culture experiments, transfection wells were randomly assigned.
Blinding	No blinding was carried out.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: anti-Tubulin β 3 (TUBB3) (BioLegend), anti-Mouse IgG (H+L) Alexa FluorTM Plus 647 (Invitrogen), anti-IFNγ monoclonal antibody (Mabtech), biotinylated anti-IFNγ secondary antibody (Mabtech).

Validation: Validated by the vendors.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): HEK293FT (ATCC), HeLa (ATCC), A549 (ATCC), A375 (ATCC) and H1 (WiCell).

Authentication: STR, by the vendor.

Mycoplasma contamination: Tested by the vendor; no mycoplasma contamination was found.

Commonly misidentified lines (See ICLAC register): HEK293FT cells were used for the cell-culture experiments, as per established procedures.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals: The laboratory mice used in this study were obtained from the Jackson Laboratory. C57BL/6J mice, male, 6–8 weeks of age.

Wild animals: The study did not involve wild animals.

Reporting on sex: Sex was not considered.

Field-collected samples: The study did not involve samples collected from the field.

Ethics oversight: Institutional Animal Care and Use Committee of the University of California, San Diego.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: HEK, HeLa, A549, A375 and H1 cells were cultured and dissociated using standard procedures.

Instrument: BD FACS LSRFortessa

Software: FlowJo

Cell population abundance	At least 10k gated cells were analysed per sample.
Gating strategy	Gates were drawn to exclude cell debris and clumps.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

nature biomedical engineering

Supplementary information

https://doi.org/10.1038/s41551-024-01245-z

Robust genome and cell engineering via in vitro and in situ circularized RNAs

In the format provided by the authors and unedited

CONTENTS

Supplementary Table 1 | IRES Constructs screened.

1	IRES_Cricket Paralysis Virus	GCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTAGTGC TATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCCCC ACAATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGAA ATTTACCTGCTACATTTCAA
2	IRES_Homo Sapiens IGF2	GCGACCGGGCATTGCCCCCAGTCTCCCCCAAATTTGGGCATTGTCCCCGGGTCTTCCA ACGGACTGGGCGTTGCTCCCGGACACTGAGGACTGGCCCCGGGGTCTCGCTCACCTT CAGCAG
3	IRES_Hepatovirus A	CAGGGTTCTTAAATCTGTTTCTCTATAAGAACACTCATTTTTCACGCTTTCTGTCTTCTTT CTTCCAGGGCTCTCCCCTTGCCCTAGGCTCTGGCCGTTGCGCCCGGCGGGGTCAACTC CATGATTAGCATGGAGCTGTAGGAGTCTAAATTGGGGACACAGATGTTTGGAACGTCAC CTTGCAGTGTTAACTTGGCTTTCATGAATCTCTTTGATCTTCCACAAGGGGTAGGCTACG GGTGAAACCTCTTAGGCTAATACTTCTATGAAGAGATGCCTTGGATAGGGTAACAGCGG CGGATATTGGTGAGTTGTTAAGACAAAAACCATTCAACGCCGGAGGACTGACTCTCATC CAGTGGATGCATTGAGTGGATTGACTGTCAGGGCTGTCTTTAGGCTTAATTCCAGACCT CTCTGTGCTTAGGGCAAACATCATTTGGCCTTAAATGGGATTCTGTGAGAGGGGATCCC TCCATTGACAGCTGGACTGTTCTTTGGGGCCTTATGTGGTGTTTGCCTCTGAGGTACTC AGGGGCATTTAGGTTTTTCCTCATTCTTAAATAATAATG
4	IRES_Hepatitis C Virus H77 Isolate	CACGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCC CCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG GACGACCGGGTCCTTTCTTGGATAAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCC CCGCAAGACTGCTAGCCGAGTAGTGTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGA TAGGGTGCTTGCGAGTGCCCCGGGAGG
5	IRES_Homo Sapiens FGF1	CGCTCCAGGGGAATCAGGGCATCGCCTCCTTTTCTGGGAGGACACTCCCTTCTGATGG TGAATGGGAACTCCCTTCCTCCTGCAGCAGCCTGCCTGCAGCTGTCCTGGTAGAACAGT GTGGACATTGCAGAAGCTGTCACTGCCCCAGAAAGAAAGCACCCCAGAGCC
6	IRES_Bovine Viral Diarrhea Virus 1	AGTAGGACTAGCATAATGAGGGGGGTAGCAACAGTGGTGAGTTCGTTGGATGGCTTAA GCCCTGAGTACAGGGTAGTCGTCAGTGGTTCGACGCCTTGGAATAAAGGTCTCGAGAT GCCACGTGGACGAGGGCATGCCCAAAGCACATCTTAACCTGAGCGGGGGTCGCCCAG GTAAAAGCAGTTTTAACCGACTGTTACGAATACAGCCTGATAGGGTGCTGCAGAGGCCC ACTGTATTGCTACTAAAAATCTCTGCTGTACATGGCACATGGAGTT
7	IRES_Human Rhinovirus A89	GTGAGAAGCCTAATTATTGACAAGGTGTGAAGAGCCGCGTGTGCTCAGTGTGCTTCCTC CGGCCCCTGAATGTGGCTAACCTTAACCCTGCAGCCGTTGCCCATAATCCAATGGGTTT GCGGTCGTAATGCGTAAGTGCGGGATGGGACCAACTACTTTGGGTGTCCGTGTTTCCT GTTTTTCTTTTGATTGCATTTTATGGTGACAATTTATAGTGTATAGATTGTCATCATGGGT GCACAAGTATCTAGACAAAATGTT
8	IRES_Pan Paniscus LIMA1	GCGCGCCAGTCCTCATCCTCCTGCTGAAGTGACAAGCCACGCTGCTTCTGGAGCCAAA GCTGACCAAGAAGAACAAATCCACCCCAGATCTAGACTCAGGTCACCTCCTGAAGCCCT CGTTCAGGGTCGATATCCCCACATCAAGGACGGTGAGGATCTTAAAGACCACTCAACAG AAAGTAAAAAACGACTCGGACCG
9	IRES_Human Adenovirus 2	GGCGCGCCAGTCCTAGTGAAAACGTTCCTGCTCTCACAGATCACGGGACGCTACCGCT GCGCAACAGCATCGGAGGAGTCCAGCGAGTGACCATTACTGACGCCAGACGCCGCAC CTGCCCCTACGTTTACAAGGCCCTGGGCATAGTCTCGCCGCGCGTCCTATCGAGCCGC ACTTTTTGAGCAAACCGACTCGGACCGATG
10	IRES_Montana Myotis Leukoencephelitis Virus	GGCGCGCCAGTCCTTGGTTGAATGAGATGCAAAAAATTAGGAAGGACATTGACCTGTGG AAACCTTCAGCAGGTCATTGCAACTGGGAGGAAGTAGAGTTCTGTTCCAACCACTTTCA CAAGCTTGTTATGAGAGATGGGAGAACCTTGGTTGTTCCTTGCAGACATGAAGTTGAGC TTGTGGGGCGTCGACTCGGACCG
11	IRES_Homo Sapiens RANBP3	GGCGCGCCAGTCCTAGACTCGGGAACTGCCTGAATGTGGTTTGGGACACGAGACCTCA TCATATTGATGAGCGAACAAACAAGAACATTTCCTCCCTCCCCTCCTTTGAATTGAAATG GCACATTAAGACTTGTCACGGCTTCTCACTGGGACTGGAGACCTCGTTCCTTCACCCCG CGTGTCGCCAGCGACTCGGACCGATG
12	IRES_Pestivirus Giraffe 1	GGCGCGCCAGTCCTAGTACAGGGCAGTCGTCAACAGTTCAACACGCAGAATAGGTTTG CGTCTTGATATGCTGTGTGGACGAGGGCATGCCCACGGTACATCTTAACCTATCCGGG GGTCGGATAGGCGAAAGTCCAGTATTGGACTGGGAGTACAGCCTGATAGGGTGTTGCA GAGACCCATCTGATCGACTCGGACCGATG
13	IRES_Homo Sapiens TGIF1	GGCGCGCCAGTCCTCCTAGTGAGAGGACTCGGGACAGGGAATTGGCCCTGGGAGAAA ACGCGCGGGGGGCGTCCGAGACGCCCCGTGAAAGCCGTGCCGACCCTTGGGAGGACT GACAGGTCTAGAGACACGTAGCCATCAGCAATGTGGGCCTCCCGGGAATAAGTGAGGG GCTCTGTGTTTCGAGGCGACTCGGACCGATG
14	IRES_Human Poliovirus 1 Mahoney	CATGGGACGCTAGTTGTGAACAAGGTGTGAAGAGCCTATTGAGCTACATAAGAATCCTC CGGCCCCTGAATGCGGCTAATCCCAACCTCGGAGCAGGTGGTCACAAACCAGTGATTG GCCTGTCGTAACGCGCAAGTCCGTGGCGGAACCGACTACTTTGGGTGTCCGTGTTTCC TTTTATTTTATTGTGGCTGCTTATGGTGACAATCACAGATTGTTATCATAAAGCGAATTGG ATTGGCCATCCGGTG

15	IRES_Foot-and-Mouth Disease Virus Type O	AAGCAGGTTTCCACAACTGATAAAACTCGTGCAACTTGAAACTCCGCCTGGTCTTTCCA GGTCTAGAGGGGTTACACTTTGTACTGTGCTCGACTCCACGCCCGGTCCACTGGCGGG TGTTAGTAGCAGCACTGTTGTTTCGTAGCGGAGCATGGTGGCCGTGGGAACTCCTCCTT GGTGACAAGGGCCCACGGGGCCGAAAGCCACGTCCAGACGGACCCACCATGTGTGCA ACCCCAGCACGGCAACTTTTACTGCGAACACCACCTTAAGGTGACACTGGTACTGGTAC TCGGTCACTGGTGACAGGCTAAGGATGCCCTTCAGGTACCCCGAGGTAACACGGGACA CTCGGGATCTGAGAAGGGGATTGGGACTTCTTTAAAAGTGCCCAGTTTAAAAAGCTTCT ACGCCTGAATAGGCGACCGGAGGCCGGCGCCTTTCCATTACCCACTACTAAATC
16	IRES_Encephalomyocarditis Virus_7A	CCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGG TGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCC CGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAA AGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAA GACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAG GTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCC CAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGT ATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGG GGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCC CCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACC
17	IRES_Encephalomyocarditis Virus_6A	CCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGG TGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCC CGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAA AGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAA GACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAG GTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCC CAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGT ATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGG GGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCC CGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACC
18	IRES_Enterovirus 71	TTAAAACAGCTGTGGGTTGTCACCCACCCACAGGGTCCACTGGGCGCTAGTACACTGG TATCTCGGTACCTTTGTACGCCTGTTTTATACCCCCTCCCTGATTTGCAACTTAGAAGCA ACGCAAACCAGATCAATAGTAGGTGTGACATACCAGTCGCATCTTGATCAAGCACTTCT GTATCCCCGGACCGAGTATCAATAGACTGTGCACACGGTTGAAGGAGAAAACGTCCGTT ACCCGGCTAACTACTTCGAGAAGCCTAGTAACGCCATTGAAGTTGCAGAGTGTTTCGCT CAGCACTCCCCCCGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCG TGGCGGTGGCTGCGTTGGCGGCCTGCCTATGGGGTAACCCATAGGACGCTCTAATACG GACATGGCGTGAAGAGTCTATTGAGCTAGTTAGTAGTCCTCCGGCCCCTGAATGCGGCT AATCCTAACTGCGGAGCACATACCCTTAATCCAAAGGGCAGTGTGTCGTAACGGGCAAC TCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCTTTTTATTCTTGTATTGGCTGC TTATGGTGACAATTAAAGAATTGTTACCATATAGCTATTGGATTGGCCATCCAGTGTCAA ACAGAGCTATTGTATATCTCTTTGTTGGATTCACACCTCTCACTCTTGAAACGTTACACAC CCTCAATTACATTATACTGCTGAACACGAAGCG
19	IRES_Coxsackievirus B3	TTAAAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTGGGCGCTAGCACTCTGGTA TCACGGTACCTTTGTGCGCCTGTTTTATACCCCCTCCCCCAACTGTAACTTAGAAGTAAC ACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTT ACCCCGGACTGAGTATCAATAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATC CGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTTCGCTCAG CACTACCCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGG CGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAAACCCATGGGACGCTCTAATACAGAC ATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCTGAATGCGGCTAAT CCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTC TGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTT ATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAAT AGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACAT TACAATTCATTGTTAAGTTGAATACAGCAAA

Supplementary Table 2 | 3' UTR Constructs Screened, icRNA, ocRNA, Chen et Al., circRNA.

1	Linear mRNA	TAATACGACTCACTATAGGGAGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAG AGAACCCACGCGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACCTGG TACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCC CCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCCACTCACCACCTCTGCT AGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCAC ACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCT ATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACACCCTGGAGCTAGCAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTTTT
2	IRES_Encephalomyocarditis Virus_7A	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACA CGATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTG GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCG GCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCA CCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGC AGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACA ACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTC AAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCC AGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAACTGCC ATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCCCGGATAAAAGT GGAGGGTACAGTCCACGCTTTTTT
3	IRES_Encephalomyocarditis Virus_6A	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA

		CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAAAACTGCCA TCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCCCGGATAAAAGTG GAGGGTACAGTCCACGCTTTTTT
4	IRES_Encephalomyocarditis Virus_6A+WPRE	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGTAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCC GGTCCCAAGCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
5	IRES_Encephalomyocarditis Virus_6A+WPRE+pA50 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT

		TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCCCGG ATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
6	IRES_Encephalomyocarditis Virus_6A+WPRE+pA120 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCC AAGCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
7	IRES_Encephalomyocarditis Virus_6A+WPRE+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA

		GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGC CATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCCCGGATAAAAG TGGAGGGTACAGTCCACGCTTTTTTT
8	IRES_Encephalomyocarditis Virus_6A+WPRE+mtRNR1-AES 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAACAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCC ACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAA CCCCAGGGTTGGTCAATTTCGTGCCAGCCACACCCTGGTACTGCATGCACGCAATGCT AGCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCCGACCTCGGGTCCCAGGTA TGCTCCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCGAAA CACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAA GCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT

9	IRES_Encephalomyocarditis Virus_6A+WPRE+mtRNR1-LSP1 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCCTCTCCCTCCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAACAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCC ACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAA CCCCAGGGTTGGTCAATTTCGTGCCAGCCACACCTTCCAGCCAGACACCCGCCCCCCG GCCCTGGCTAAGAAGTTGCTTCCTGTTGCCAGCATGACCTACCCTCGCCTCTTTGATGC CATCCGCTGCCACCTCCTTTTGCTCCTGGACCCTTTAGCCTCTCTGCCCTTCCACTCTCT GACCCCCGAAACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAA TGCCGGTCCCAAGCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
10	IRES_Encephalomyocarditis Virus_6A+WPRE+AES-mtRNR1 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCCTCTCCCTCCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT

		TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAACTGGTACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGGTA CCCCGAGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCCACT CACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCT TAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAG TTTAACTAAGCTATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACACCCGAAAC ACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAA GCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
11	IRES_Encephalomyocarditis Virus_6A+WPRE+AES-hBg 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAACTGGTACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGGTA CCCCGAGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCCACT CACCACCTCTGCTAGTTCCAGACACCTCCGAGAGCTCGCTTTCTTGCTGTCCAATTTCTA TTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCT TGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCTGCGTCCGAAACACA CACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCC CGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
12	IRES_Encephalomyocarditis Virus_6A+WPRE+FCGRT-hBg 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG

		CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAATGCCCGTCCTCACCAAGACTGACTGCCTGCTGCTTTGCTACTGCCCGG GCCCATGAGACTGACTTCCCACTGCTCTGCCTGCCTCTCCCCACTGCACTGGCACAGC CCCGCCTTGCCGCTGCTGATCCATTGCCGGTGTGACCGAGAGCTCGCTTTCTTGCTGT CCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTAT GAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCTGCGTC CGAAACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGT CCCAAGCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
13	IRES_Encephalomyocarditis Virus_6A+WPRE+2hBg 3' UTR +pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGAGAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCC CTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGC CTAATAAAAAACATTTATTTTCATTGCTGCGTCGAGAGCTCGCTTTCTTGCTGTCCAATTT CTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGG CCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCTGCGTCCGAAAC ACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAA

		GCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
14	IRES_Encephalomyocarditis Virus_6A+WPRE+HBA1 3' UTR+pA165 3'	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCGCCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGA AGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCG TCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAG GGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCG GAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACA CCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAG TCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACC CCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACAC GATGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG CGTCTTCGAAGCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCTTGGGCCTCCCCCCAG CCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGC GGCACGAAACACACACACACACACACACCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATG CCGGTCCCAAGCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
15	IRES_Encephalomyocarditis Virus_6A+Modded WPRE+pA165 3' (icRNA) Structure Code: T7 Promoter, Twister Ribozyme 5', Ligation Area 5', EMCV IRES, GFP, WPRE (Modified), PolyA (165nt), Ligation Area 3', Twister Ribozyme 3', Terminator Hairpin	TAATACGACTCACTATAGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAGG GGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGGCC CGCGGAATTCCGCCCCCCCCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAG CCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTC TTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGG GGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGT TCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGA ACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACC TGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTC AAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCC ATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGG TTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGA TGATAATATGGCCACAACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC

		CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACAATCCAGCGGA CCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGAAGTAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGCCATCAGTCGGCGTG GACTGTAGAACACTGCCAATGCCGGTCCCAAGCCCGGATAAAAGTGGAGGGTACAGTC CACGCTTTTTTT
-	Group II Intron-Exon 1 + IRES_Encephalomyocarditis Virus_6A+Modded WPRE+pA165 3' + Group II Intron-Exon 2 (ocRNA) Structure Code: T7 Promoter, Twister Ribozyme 5', Ligation Area 5', Group II Intron-Exon 1, EMCV IRES, GFP, WPRE (Modified), PolyA (165nt), Group II Intron-Exon 2, Ligation Area 3', Twister Ribozyme 3', Terminator Hairpin	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGCTATTATCGAGCGAACGCCT TATGCGATGAAAGTCGCACGTAGGGTGTAGACCAAGCGAAATCCTATGCATTTAGGATA GTGAGGTATAGCAAAGGAGAAGTCGACGGGCCCGCGGAATTCCGCCCCCCCCCCCTC TCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCG TTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAA CCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAAT GCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAA CAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCT CTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGC CACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAA CAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCT CGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAAC CACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACCATGGTGA GCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG GCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCAC CCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATG AAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCA TCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGA CACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATC CTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACA AGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCT GCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA TGGACGAGCTGTACAAGTAAGGATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGAT TGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGC CTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTG GTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGC ACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCT TTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGC CTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGT CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGC GGGACGTCCTTCTGCTACAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTC TGCGGCCTCTTCCGCGTCTTCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAACTATAGCCATACAATAAAAGTGCGAAACGTTATCCTATAAGTAAGAAAGTTTT AAAATTTTCTTACGAAAAGGATAGAACTTAAAAGTTCTAACTGTTCTACTAAAGTAATAAG TGAAAATCTTATTTAAAGCAAACAACCAAGTAGCTTTAAGTCTAAGTCCCCTACACAAGT TTTATACTACTATGCAAAACTTGTGAAGCTAGGTAAGGTCGTAATCCGTGAAAGTCGGAT GCGGGGCTCCTTAAAAGATTACTATGGTAAACATAAGCTAATCCATTAAGATGCGATTTA TATGTATTTTATACTGTTAAATATTTTTGTGCTTGTGGCTTGGTATAAAACAGTTAAGATG AAGTACTTAACTGGTTTTGGAATAATTGGTTGTTAAACTAAAACATTATAAATCGTTAGTG GATACCTAAGGTAATCAAAAATAGGGATAGGTAGAATGGAACGTTTGATGCTGTATATGA AGAGGTTTAGTAGAACCTAGGACACATATACGGGCTCAGCAGGTTCATAGTAGCTATGA TACTCAGCCGGAAGTCAAATTAATTTTGAAATACTTCTATGGTAACATAGGAGAAGGATA AAACTGAGTGAGCCAAGGAACCTAGTCGGTAATAGCTGCCATCAGTCGGCGTGGACTG TAGAACACTGCCAATGCCGGTCCCAAGCCCGGATAAAAGTGGAGGGTACAGTCCACGC TTTTTTT

-	IRES_Enterovirus 71 Virus_6A+Modded WPRE+pA165 3' (icRNA - EV71)	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCTTAAAACAGCTGTGGGTTGTCACCCACCCACAGGGTCCACTGGGC GCTAGTACACTGGTATCTCGGTACCTTTGTACGCCTGTTTTATACCCCCTCCCTGATTTG CAACTTAGAAGCAACGCAAACCAGATCAATAGTAGGTGTGACATACCAGTCGCATCTTG ATCAAGCACTTCTGTATCCCCGGACCGAGTATCAATAGACTGTGCACACGGTTGAAGGA GAAAACGTCCGTTACCCGGCTAACTACTTCGAGAAGCCTAGTAACGCCATTGAAGTTGC AGAGTGTTTCGCTCAGCACTCCCCCCGTGTAGATCAGGTCGATGAGTCACCGCATTCCC CACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCTATGGGGTAACCCATAGG ACGCTCTAATACGGACATGGCGTGAAGAGTCTATTGAGCTAGTTAGTAGTCCTCCGGCC CCTGAATGCGGCTAATCCTAACTGCGGAGCACATACCCTTAATCCAAAGGGCAGTGTGT CGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCTTTTTATT CTTGTATTGGCTGCTTATGGTGACAATTAAAGAATTGTTACCATATAGCTATTGGATTGG CCATCCAGTGTCAAACAGAGCTATTGTATATCTCTTTGTTGGATTCACACCTCTCACTCT TGAAACGTTACACACCCTCAATTACATTATACTGCTGAACACGAAGCGCATGGTGAGCA AGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACG TAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCT CGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCT GGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAG CAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCG TGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC TGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA GCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATG GACGAGCTGTACAAGTAAGGATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTG ACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTT TGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTT GCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACT GTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTC CGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTT GCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCG GGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGG GACGTCCTTCTGCTACAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTG CGGCCTCTTCCGCGTCTTCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAA GCCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
-	Group II Intron-Exon 1 + IRES_Enterovirus 71 Virus_6A+Modded WPRE+pA165 3' + Group II Intron-Exon 2 (ocRNA - EV71)	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGCTATTATCGAGCGAACGCCT TATGCGATGAAAGTCGCACGTAGGGTGTAGACCAAGCGAAATCCTATGCATTTAGGATA GTGAGGTATAGCAAAGGAGAAGTCGACGGGCCCGCGGAATTCCTTAAAACAGCTGTGG GTTGTCACCCACCCACAGGGTCCACTGGGCGCTAGTACACTGGTATCTCGGTACCTTTG TACGCCTGTTTTATACCCCCTCCCTGATTTGCAACTTAGAAGCAACGCAAACCAGATCAA TAGTAGGTGTGACATACCAGTCGCATCTTGATCAAGCACTTCTGTATCCCCGGACCGAG TATCAATAGACTGTGCACACGGTTGAAGGAGAAAACGTCCGTTACCCGGCTAACTACTT CGAGAAGCCTAGTAACGCCATTGAAGTTGCAGAGTGTTTCGCTCAGCACTCCCCCCGT GTAGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCTGCCTATGGGGTAACCCATAGGACGCTCTAATACGGACATGGCGTGAAG AGTCTATTGAGCTAGTTAGTAGTCCTCCGGCCCCTGAATGCGGCTAATCCTAACTGCGG AGCACATACCCTTAATCCAAAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAAC CGACTACTTTGGGTGTCCGTGTTTCTTTTTATTCTTGTATTGGCTGCTTATGGTGACAATT AAAGAATTGTTACCATATAGCTATTGGATTGGCCATCCAGTGTCAAACAGAGCTATTGTA TATCTCTTTGTTGGATTCACACCTCTCACTCTTGAAACGTTACACACCCTCAATTACATTA TACTGCTGAACACGAAGCGCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC

		CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACAATCCAGCGGA CCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTATAGCCATACAATAAAAGTG CGAAACGTTATCCTATAAGTAAGAAAGTTTTAAAATTTTCTTACGAAAAGGATAGAACTTA AAAGTTCTAACTGTTCTACTAAAGTAATAAGTGAAAATCTTATTTAAAGCAAACAACCAAG TAGCTTTAAGTCTAAGTCCCCTACACAAGTTTTATACTACTATGCAAAACTTGTGAAGCTA GGTAAGGTCGTAATCCGTGAAAGTCGGATGCGGGGCTCCTTAAAAGATTACTATGGTAA ACATAAGCTAATCCATTAAGATGCGATTTATATGTATTTTATACTGTTAAATATTTTTGTGC TTGTGGCTTGGTATAAAACAGTTAAGATGAAGTACTTAACTGGTTTTGGAATAATTGGTT GTTAAACTAAAACATTATAAATCGTTAGTGGATACCTAAGGTAATCAAAAATAGGGATAG GTAGAATGGAACGTTTGATGCTGTATATGAAGAGGTTTAGTAGAACCTAGGACACATATA CGGGCTCAGCAGGTTCATAGTAGCTATGATACTCAGCCGGAAGTCAAATTAATTTTGAA ATACTTCTATGGTAACATAGGAGAAGGATAAAACTGAGTGAGCCAAGGAACCTAGTCGG TAATAGCTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCC CGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
-	IRES_Coxsackievirus B3 Virus_6A+Modded WPRE+pA165 3' (icRNA - CVB3)	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGAGCGGCCGCGTCGACGGG CCCGCGGAATTCCTTAAAACAGCCTGTGGGTTGATCCCACCCACAGGCCCATTGGGCG CTAGCACTCTGGTATCACGGTACCTTTGTGCGCCTGTTTTATACCCCCTCCCCCAACTG TAACTTAGAAGTAACACACACCGATCAACAGTCAGCGTGGCACACCAGCCACGTTTTGA TCAAGCACTTCTGTTACCCCGGACTGAGTATCAATAGACTGCTCACGCGGTTGAAGGAG AAAGCGTTCGTTATCCGGCCAACTACTTCGAAAAACCTAGTAACACCGTGGAAGTTGCA GAGTGTTTCGCTCAGCACTACCCCAGTGTAGATCAGGTCGATGAGTCACCGCATTCCCC ACGGGCGACCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATGGGGAAACCCATGGGA CGCTCTAATACAGACATGGTGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCC CTGAATGCGGCTAATCCTAACTGCGGAGCACACACCCTCAAGCCAGAGGGCAGTGTGT CGTAACGGGCAACTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTGTTTCATTTTATT CCTATACTGGCTGCTTATGGTGACAATTGAGAGATCGTTACCATATAGCTATTGGATTGG CCATCCGGTGACTAATAGAGCTATTATATATCCCTTTGTTGGGTTTATACCACTTAGCTT GAAAGAGGTTAAAACATTACAATTCATTGTTAAGTTGAATACAGCAAAATGGTGAGCAAG GGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAG CTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCG TGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCA GCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGG GGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCA GAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTG CAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGC GCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA CGAGCTGTACAAGTAAGGATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGAC TGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTG TATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGC TGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGT GTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCC GGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTG CCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGG GGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGG ACGTCCTTCTGCTACAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGC GGCCTCTTCCGCGTCTTCGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAACTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAG CCCGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT

-	Group II Intron-Exon 1 + IRES_Coxsackievirus B3 Virus_6A+Modded WPRE+pA165 3' + Group II Intron-Exon 2 (ocRNA - CVB3)	TAATACGACTCACTATAGGGCCATCAGTCGCCGGTCCCAAGCCCGGATAAAATGGGAG GGGGCGGGAAACCGCCTAACCATGCCGACTGATGGCAGCTATTATCGAGCGAACGCCT TATGCGATGAAAGTCGCACGTAGGGTGTAGACCAAGCGAAATCCTATGCATTTAGGATA GTGAGGTATAGCAAAGGAGAAGTCGACGGGCCCGCGGAATTCCTTAAAACAGCCTGTG GGTTGATCCCACCCACAGGCCCATTGGGCGCTAGCACTCTGGTATCACGGTACCTTTGT GCGCCTGTTTTATACCCCCTCCCCCAACTGTAACTTAGAAGTAACACACACCGATCAAC AGTCAGCGTGGCACACCAGCCACGTTTTGATCAAGCACTTCTGTTACCCCGGACTGAGT ATCAATAGACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAACTACTTC GAAAAACCTAGTAACACCGTGGAAGTTGCAGAGTGTTTCGCTCAGCACTACCCCAGTGT AGATCAGGTCGATGAGTCACCGCATTCCCCACGGGCGACCGTGGCGGTGGCTGCGTT GGCGGCCTGCCCATGGGGAAACCCATGGGACGCTCTAATACAGACATGGTGCGAAGA GTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCTGAATGCGGCTAATCCTAACTGCGGA GCACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCGGAAC CGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTTATGGTGACAAT TGAGAGATCGTTACCATATAGCTATTGGATTGGCCATCCGGTGACTAATAGAGCTATTAT ATATCCCTTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGTTAAAACATTACAATTCATT GTTAAGTTGAATACAGCAAAATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGG CATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGC TTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGT TGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT TGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCT GTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC TCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACAATCCAGCGGA CCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTATAGCCATACAATAAAAGTG CGAAACGTTATCCTATAAGTAAGAAAGTTTTAAAATTTTCTTACGAAAAGGATAGAACTTA AAAGTTCTAACTGTTCTACTAAAGTAATAAGTGAAAATCTTATTTAAAGCAAACAACCAAG TAGCTTTAAGTCTAAGTCCCCTACACAAGTTTTATACTACTATGCAAAACTTGTGAAGCTA GGTAAGGTCGTAATCCGTGAAAGTCGGATGCGGGGCTCCTTAAAAGATTACTATGGTAA ACATAAGCTAATCCATTAAGATGCGATTTATATGTATTTTATACTGTTAAATATTTTTGTGC TTGTGGCTTGGTATAAAACAGTTAAGATGAAGTACTTAACTGGTTTTGGAATAATTGGTT GTTAAACTAAAACATTATAAATCGTTAGTGGATACCTAAGGTAATCAAAAATAGGGATAG GTAGAATGGAACGTTTGATGCTGTATATGAAGAGGTTTAGTAGAACCTAGGACACATATA CGGGCTCAGCAGGTTCATAGTAGCTATGATACTCAGCCGGAAGTCAAATTAATTTTGAA ATACTTCTATGGTAACATAGGAGAAGGATAAAACTGAGTGAGCCAAGGAACCTAGTCGG TAATAGCTGCCATCAGTCGGCGTGGACTGTAGAACACTGCCAATGCCGGTCCCAAGCC CGGATAAAAGTGGAGGGTACAGTCCACGCTTTTTTT
-	Chen et Al., CircRNA GFP	TAATACGACTCACTATAGGGGTTAACAAAAAAAAAAAAAAAAAAAAAAATCACCGACCG ATCTATCTGAACCGGTGGGAATTCTAGAGAAAATTTCGTCTGGATTAGTTACTTATCGTG TAAAATCTGATAAATGGAATTGGTTCTACATAAATGCCTAACGACTATCCCTTTGGGGAG TAGGGTCAAGTGACTCGAAACGATAGACAACTTGCTTTAACAAGTTGGAGATATAGTCT GCTCTGCATGGTGACATGCAGCTGGATATAATTCCGGGGTAAGATTAACGACCTTATCT GAACATAACGCTACCGTTTAATATTGCGTCATATTCGGCGACCATTTGTGTGGTAAAAAA AAAAAACCAAAAAAAAAAAAACAAAAAAAAAAAAATAATTGACTAAGATATCTTAAAACAGCG GATGGGTACCCCACCATCCGACCCACTGGGTGTAGTACTCTGGTACTTCGTACCTTTGT ACGCCTGTTCTTCCCATTGTACCCTTCCTGAACTTCCAACCCAAGTAACGTTAGAAGCTC AACATTTAGTACAACAGGAAGCACCACATCCAGTGGTGTTTAGTACAAGCACTTCTGTTT CCCCGGAGCGAGGTATAGGCTGTACCCACTGCCAAAAACCTTTAACCGTTATCCGCCAA CCAACTACGTAAAAGCTAGTAGTATTATGTTTTTAACTAGGCGTTCGATCAGGTGGATTT CCCCTCCACTAGTTTGGTCGATGAGGCTAGGAATTCCCCACGGGTGACCGTGTCCTAG CCTGCGTGGCGGCCAACCCAGCCCACTCACTATTTGTTTTCGCGCCCAGTTGCAAAAAG TGTCGGGGCTGGGACGCCTTTTTATAGACATGGTGTGAAGACTCGCATGTGCTTGGTTG TGATTCCTCCGGCCCCTGAATGCGGCTAACCTTAACCCTGGAGCCTTGTGTCACAAACC AGTGATGATAAGGTCGTAATGAGCAATTCCGGGACGGGACCGACTACTTTGGGTGTCC GTGTTTCTTATTTTTCTTATTATTGTCTTATGGTCACAGCATATATATAACATATACTGTGA TCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGC TGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG

		CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCC CTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCC GACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGG AGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCA TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGA GGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG CCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGAC CCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCA CTCTCGGCATGGACGAGCTGTACAAGTAATAACAATTGGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCC CGTGGTCTTTGAATAAAGTCTGAACCACACAAATGGTCGCCGACTCAGTAGATGTTTTCT TGGGTTAATTGAGGCCTGAGTATAAGGTGACTTATACTTGTAATCTATCTAAACGGGGAA CCTCTCTAGTAGACAATCCCGTGCTAAATTGTAGGACTGCCCTTTAATAAATACTTCTAT ATTTAAAGAGGTATTTATGAAAAGCGGAATTTATCAGATTAAAAATACTTTCTCTAGAGTC GACCTGCAGACGCGTTCAGATAGATCGGTCGGTGAAAAAAAAAAAAAAAAAAAAAAAA AGCGGCCGCCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

Supplementary Table 3 | RNA Transfection conditions.

Cell Line	Number of Cells Seeded	Lipofectamine MessengerMAX (uL/well)	Well Format	Transfected mRNA (pmol/well)	Figure
Hek293T	200k	3.5	12 well	3.75	Fig 1i
Hek293T	200k	3.5	12 well	1.5	Fig 1h, 2a, 2b, 2e, 4a, S1a
Hek293T	125k	3.5	12 well	0.65	Fig 6e, 6f
K562	-	-	12 well	0.65	Fig 6e
Cardiomyocytes	300k	3.5	12 well	1.5	Fig 3ba
HeLa	150k	3.5	12 well	2	Fig 4b, 4c
Hek293T	100k	1	24 well	0.6	Fig 3a
HeLa	100k	1	24 well	0.6	Fig 3a
A549	100k	1	24 well	0.6	Fig 2c, S1b, 3a
A375	100k	1	24 well	0.6	Fig 3a
hPSCs	50k	1	24 well	0.6	Fig 3d, 3a
Neurons	50k	1	24 well	1.2	Fig 3b
A549	7.5k	0.25	96 well	0.15	Fig 2d

Supplementary Table 4 | Cas9 block start and end positions.

Block	Start	End
1	0	680
2	650	1030
3	1000	1630
4	1600	1990
5	1960	2400
6	2370	3000
7	2970	3790
8	3760	4140

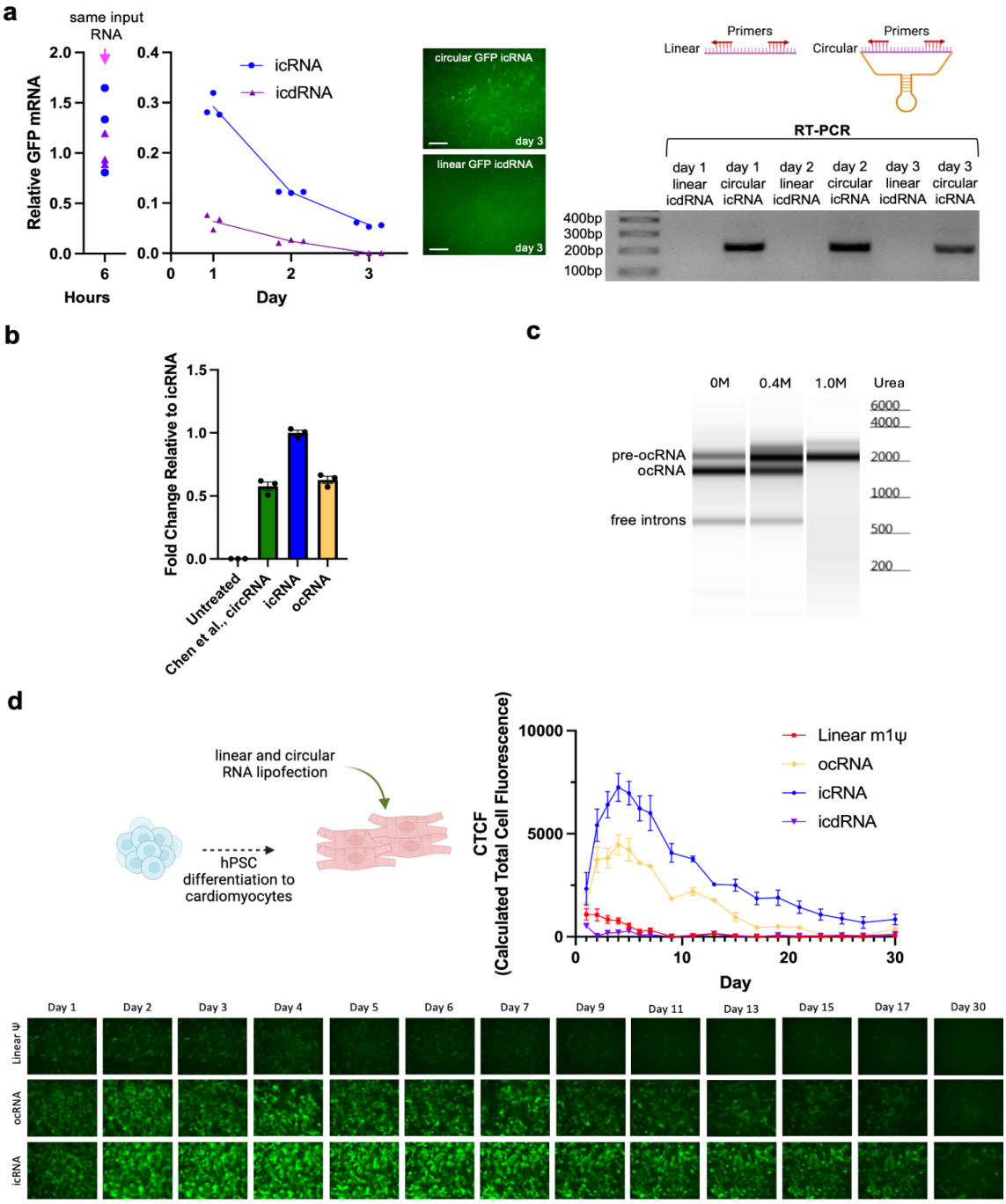
Supplementary Table 5 | Sequences of ZFs targeting hPCSK9.

Position	ZF1	ZF2	ZF3	ZF4	ZF5	ZF6
1	TSGSLSR	RKELLRS	DRSYRNT	RRSDLKR	RSDNLAR	DRSALAR
2	RSDDLTR	HRRSRDQ	RSADLTR	RSDDLTR	RSDVLST	WGRLRKL
3	QSSDLSR	RRDALLM	RSCCLHL	RNASRTR	YQGVLTR	RSDNLRE
4	RSADLTR	DRSNLTR	RSDHLSQ	DNSHRTR	RSADLTR	RSDNLRE
5	RSDSLTR	QSGDLTR	QSGDLTR	QSGDLTR	QSGDLTR	QSGDLTR
6	DRSNLSR	RSDALAR	DRSDLTR	DRSTLAR	DRSDLSR	DRTDLTR
7	QSSDLSR	HRSTRNR	ERGTLAR	QSGHLSR	DDWNLSQ	RSANLTR
8	RSDSLSR	RNASLAT	DRSDLSR	DRSDLSR	DRSALSR	QSGTLAR
9	QSSDLSR	QSGHLSR	RSDVLSE	RNQHRKT	YPKDLSK	QSGNLAR
10	QSSDLSR	QSGHLSR	RSDVLSE	RNQHRKT	YPKDLLK	QSGNLAR
11	HACCRNM	RSADLSR	DRSDLSR	RRTDLRR	RSDHLSE	QSATRTK
12	TSGSLSR	QSGNLAR	QSSDLSR	QSGHLSR	QSGDLTR	LRTSLSK
13	DRGTLSR	DRSHLTR	RSDTLSK	DNSTRIK	RSDHLSR	RSHDRTK
14	RCGHLYD	QNATRTK	QAHHLRT	DRSNLTR	RSDSLLR	DRSHLAR
15	QSGDLTR	RRADRAK	DNRDLST	DRSNLTR	RSDNLST	DNSYLKT
16	QNAHRKT	IRSTLRD	QSAHRKN	RHSHLTS	RSHSLLR	RLDWLPM
17	RSDTLSA	SNRARRR	QSSDLSR	YKWTLRN	RSDNLAR	QSGHLSR
18	QSAHLAR	QSGDLTR	RSGNLST	DRSNRIR	RSDNLAR	RSDHLSR
19	QSSDLSR	HRKSLSR	RSDSLLR	RLDWLPM	DRSALSR	RSDDLTR
20	QAHHLRT	DRSNLTR	RSDNLSE	SSRNLAS	DRSHLTR	NSRNLRN
21	RSDDLTR	QAATLSR	RSDSLSV	QSGDLTR	QSSDLSR	QWSTRKR
22	RSDDLSR	QSGDLTK	RSDSLTR	QSGDLTR	QSGDLTR	QSGDLTR
23	QSGDLTR	RSDALAR	QSGDLTR	QSGDLTR	QSGDLTR	QSGDLTR
24	RSSDLTE	RSADLSR	QSSDLSR	HRSTLSR	QSSDLSR	QSSTLSR
25	QSGSLTR	QSGDLTR	QSGDLTR	QSGDLTR	QSGDLTR	QSGDLTR
26	QSSDLRR	RSDDLSR	QSGDLTR	QSGDLAR	QSSALTR	QSSALTR
27	DRSNLSR	RSDNLTR	DRSNLTR	RSDNLTR	RSDNLSR	DSGHLSR
28	QSGSLTR	QSGDLTR	QSGDLTR	QSGHLAR	QSGDLSQ	DSHHLTR
29	HACCRNM	RSADLSR	DRSDLSR	RRTDLRR	RSDHLSE	QSATRKK
30	RSGHLSE	QNATRTK	QSGNLAR	HRSSLSR	RSDSLLQ	DSSTRIK
31	HRQRLEE	RNASRTR	DRSDLSR	RRTDLRR	RSDHLSE	RNRDRIT
32	DRSNLSR	RSSHLAR	DRSNLTR	RSDNRIR	DRSNLVR	RSDNLTR
33	DRSNLAR	DRSNLTR	TSGHLSR	DRSNLTR	RSDNLSR	DRSNLTR
34	RSDNLAR	RSDNLAR	DRSNLRT	DRSNLTR	DRSNLTR	DRSNLTR
35	ERGTLTR	RSDALAR	DRDSLTR	DRSNLAR	RSDNLAR	RSDNLTR
36	QKATRTT	RNASRTR	DDWNLSQ	RSANLTR	RSDNLSR	DRSNRTK
37	RSDALSV	DSSHRTR	DRSDLSR	RSDNLTR	RSDNLAR	DSSDRKK

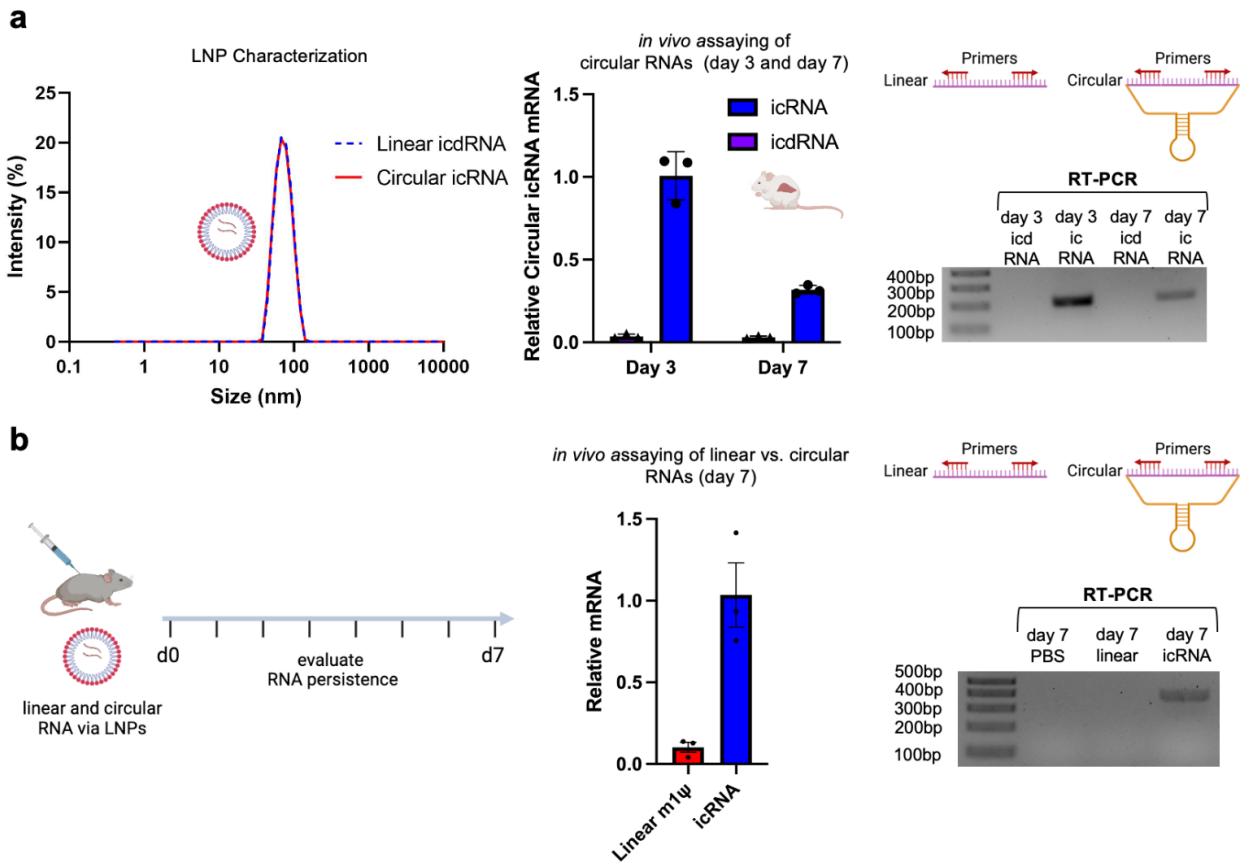
38	QSGSLTR	QSGNLAR	ERRDLSR	RSDALPM	DRSDLSR	DRSNLTR
39	QSGSLTR	QSGDLTR	QSGDLTR	QSGDLTR	QSGHLTR	QKGTLGE
40	DRSALSR	RSDALAR	QSSDLSR	RSDALAR	RSDNLAA	QAHHRIT
41	YPKDLVK	QSGHLSR	DQSNLRA	RSANLTR	DSSTRKT	QSATRTK
42	WRSCRSA	DSSHRTR	TSGHLSR	QSGNLAR	RSDALSR	RSDDLTR
43	QSGHLAR	QSGNLAR	RSDALSR	RSDDLTR	QSSDLSR	DRSHLTR
44	QSGNLSR	RSDALAR	RSSDLTR	ASSTRIK	DRSHLSR	RSDALTD
45	ERGTLAR	RSDDLTR	QSGDLTR	RSGHLAR	QSGDLTR	QSGDLTR
46	RSDHLSR	RSADLRR	DRSALTR	RSAHLTR	QSGHLTR	RSDALAR
47	RSDHLSR	QSGNLAR	RSDNLAR	DRSDLSR	RSDNLAR	RNQDRTN
48	DRSHLAR	DRSHLAR	QSSNLAR	RSDNLAR	DRGTLSR	RSDNLAR
49	RSDNLAR	QSGDLTR	QSGNLAR	DRSHLAR	QSGNLAR	RSDNLTR
50	QSGHLAR	QSGDLTR	RSDHLSE	RSSNRKK	QSGNLSR	RSDHLTT
51	QSSHLSR	RSDARTR	RSDTLSE	QNSNRTK	RSDTLSR	DRRDRKK
52	DRSHLVR	RSDALRT	RSDNRKT	QSGNLAR	RSDSLSQ	TSHHLSR
53	RSDNLSR	RSADLSR	QSGDLTR	RRDHLTT	QSDVLSR	QSATLSK
54	TSSHLSR	RSDNLAR	RSDTLSE	QSSTRIR	RSDHLSQ	QSATRKT
55	QSGALKK	QSGHLTR	TSGHLSR	RSDNLTR	RSDTLTR	QSGDLSR
56	RSDNLAR	QSGHLTR	RSGDLSV	RSDALAR	RSSHLAR	TSSHLTR
57	QSGHLAR	QSGNLAR	RSDHLSR	RSDHLSR	RSDHLSR	DSDHLSR
58	QSSNLSR	QSSDRIK	QSDSLST	QSGNLAR	RSDNLST	QSGHRIN
59	QSGSLTR	RSDHLSR	QSAHLRR	TSSDLRR	QSSDLSR	QAGNLSK
60	RSDHLSR	QSGDLTR	RSDHLSR	RNANRKR	QSSDLSR	QASNLSK
61	RSDDLAR	RSDHLSR	RSDDLTR	RSDHLSR	QSSDLTR	QSSDLRR
62	RSDHLSR	RNDNRKT	RSDNLAR	RSDHLSR	RSDTLSR	RSDDRTR
63	RSDDLTR	RSDHLSR	QNDNLSA	RSAHLTR	RSDHLSE	DSANRIK
64	RSDHLSE	TSSHLSR	RSDNLSA	RSDNLQR	RSDHLSR	QSGDRTK
65	QSDSLAR	RSDNLTR	RSDHLSE	RSDHLSR	RSDNLAR	QSGHLSR
66	RSDHLSE	NSASRTK	RSDSLSR	RLDNRKK	RSAHLAR	QSGDLTR
67	RSDNLAR	RSDDLTR	RSAHLAR	RSDARKR	QSSDLTR	TSGHLSR

Supplementary Table 6 | qPCR primers.

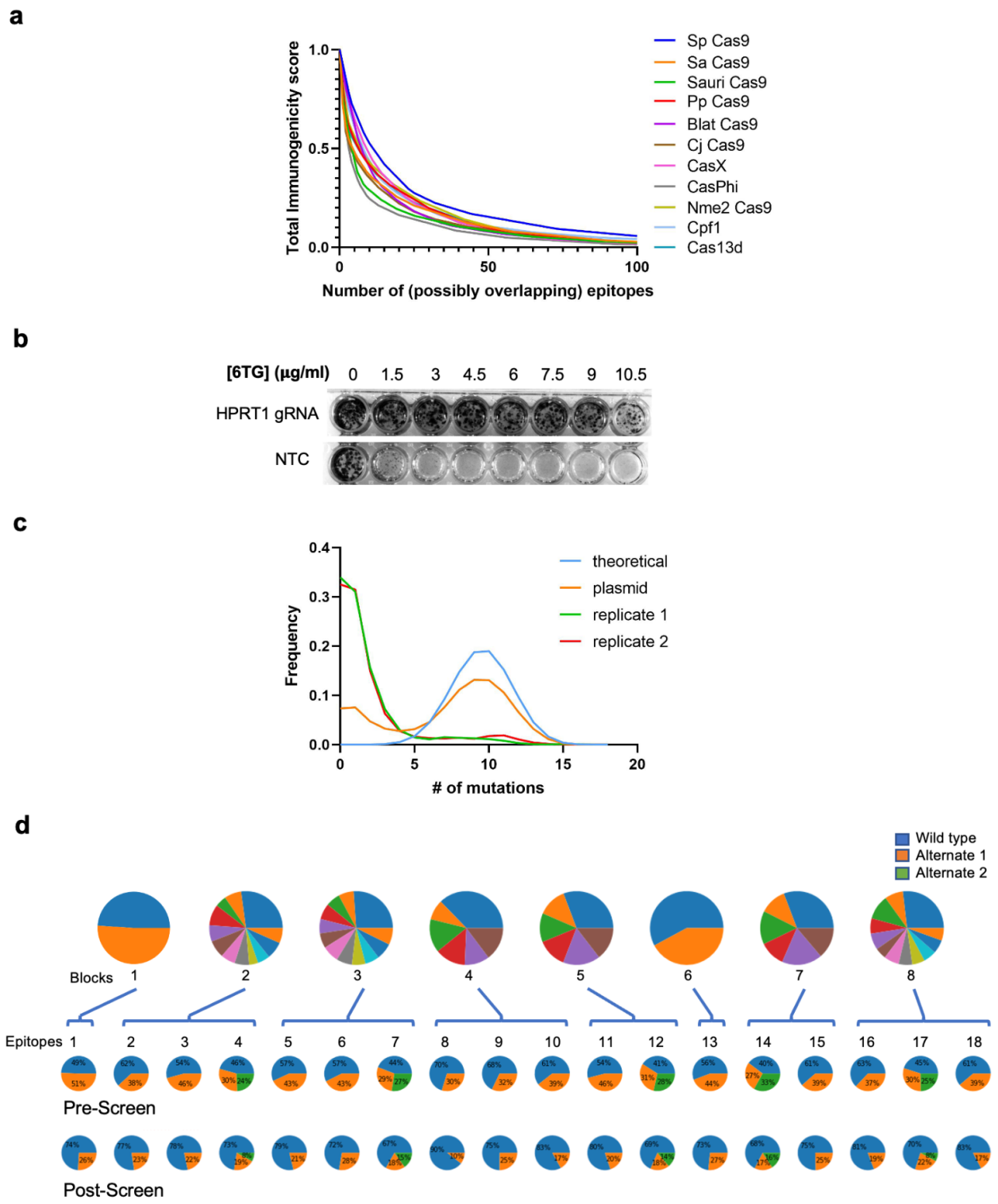
CXCR4_F	GAAGCTGTTGGCTGAAAAGG
CXCR4_R	CTCACTGACGTTGGCAAAGA
ASCL1_F	CGCGGCCAACAAGAAGATG
ASCL1_R	CGACGAGTAGGATGAGACCG
B2M_F	TATGCCTGCCGTGTGAACCATGT
B2M_R	GGCATCTTCAAACCTCCATGATGCT
hGAPDH_F	ACAGTCAGCCGCATCTTCTT
hGAPDH_R	ACGACCAAATCCGTTGACTC
hPCSK9_F	ATCCACGCTTCCTGCTGC
hPCSK9_R	CACGGTCACCTGCTCCTG
hEPO_F	CAGCGCCCTTGAGGACTAT
hEPO_R	TGTTCTACATGCCTCACCTGT
GFP_F	ACGTAAACGGCCACAAGTTC
GFP_R	AAGTCGTGCTGCTTCATGTG
circ_F	ACAACCACTACCTGAGCACC
circ_R	AATTCCGCGGGCCCGT
mGAPDH_F	TGGCCTTCCGTGTTCCTAC
mGAPDH_R	GAGTTGCTGTTGAAGTCGCA
mPCSK9_F	GCTTCTGCTCCAGAGGTCAT
mPCSK9_R	CTCCGATGATGTCCTTCCCG
hRIG-I_F	GTTGTCCCCATGCTGTTCTT
hRIG-I_R	GCAAGTCTTACATGGCAGCA
hIFN-B_F	CTCTCCTGTTGTGCTTCTCC
hIFN-B_R	GTCAAAGTTCATCCTGTCCTTG
hIL-6_F	AGCCACTCACCTCTTCAGAAC
hIL-6_R	GCCTCTTTGCTGCTTTCACAC
hMAP2_F	CTCAGCACCGCTAACAGAGG
hMAP2_R	CATTGGCGCTTCGGACAAG
hTUBB3_F	GGCCAAGGGTCACTACACG
hTUBB3_R	GCAGTCGCAGTTTTCACACTC
hVGLUT2_F	GGGAGACAATCGAGCTGACG
hVGLUT2_R	TGCAGCGGATACCGAAGGA
hBRN2_F	AAGCGGAAAAAGCGGACCT
hBRN2_R	GTGTGGTGGAGTGTCCCTAC



Supplementary Fig. 1 l Characterization of circular RNAs *in vitro*. a, Left, HEK293T cells were transfected with circular GFP icRNA and linear icdRNA and GFP mRNA amount was measured over time. The 6-hour time point was included to assess initial RNA input (left panel, n=3, p=.414; t-test, two-tailed). Data from days 1, 2, and 3 illustrate persistence of icRNA (middle panel). Values represented as mean +/- SEM (n=3, p=0.000143 for day 1, p<0.0001 for day 2, p<0.0001 for day 3 t-test, two-tailed). Values were normalized to the 6-hour time point. GFP protein expression was largely gone by day 3 in linear icdRNA transfected cells (right panel). Right, RT-PCR based confirmation of icRNA circularization in cells. **b**, A549 cells were transfected with icRNA, RnaseR treated ocRNA, and RnaseR treated circRNA with enhanced protein translation components by Chen et Al. GFP intensity was quantified by flow cytometry relative to icRNA. Values represented as mean +/- SEM (n=3) **c**, Tapestation of ocRNAs with increasing molar concentrations of urea added prior to IVT. ocRNA product band intensity is relatively decreased as molarity of the RNA denaturing agent increases. **d**, Left, Post differentiation of stem cells into cardiomyocytes, ocRNAs, icRNAs or linear m1Ψ RNAs were transfected into cells and images were taken over 30 days. Right, calculated total cell fluorescence (CTCF) GFP expression over time was plotted for icRNA and linear m1Ψ RNA based on a lower exposure to ensure optimal intensity representation. Values represented as mean +/- SEM (n=3). Bottom, representative images are shown illustrating icRNA and ocRNA persistence.

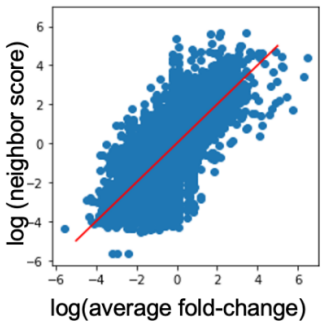


Supplementary Fig. 2 | Persistence of circular RNAs *in vivo*. **a**, Left, Characterization of lipid nanoparticles (LNPs) encapsulating icRNAs by dynamic light scattering. No differences in size were observed for LNPs containing circular icRNAs or linear circularization defective icdRNAs. Right, LNPs containing either circular icRNA or linear icdRNA were retro orbitally injected into C57BL/6J mice and RNA was isolated from livers on days 3 and 7. The ratio of circular RNA detected normalized to icRNA day 3 expression was plotted (n=3). RT-qPCR confirmed icRNA persistence up to day 7 *in vivo*. **b**, Left, Lipid nanoparticles containing icRNA or linear mRNA were retro orbitally injected into C57BL/6J mice. Middle, RNA expression in the liver after 7 days normalized to icRNA was quantified by RT-qPCR and plotted (n=3). Right, RT-qPCR confirmed *in vivo* circularization of icRNA constructs.

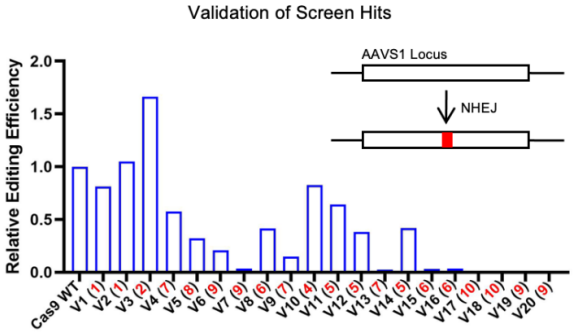


Supplementary Fig. 3 | LORAX screen design and results. a, Immunogenicity scores for Cas orthologs, demonstrating reduced immunogenicity (averaged across HLA types) as the number of no mutated epitopes increases. **b**, Presence of HPRT1 converts 6TG into a toxic nucleotide analog. HeLa cells transduced with wildtype Cas9 and either a HPRT1 targeting or nontargeting (NTC) guide. Only cells where the HPRT1 gene is disrupted are capable of living in various concentrations of 6TG. 6 μg/mL 6TG was used for the screen as this concentration was sufficient for complete killing of NTC-bearing cells. **c**, Variant Cas9 sequences were amplified from the plasmid library or genomic DNA post-screen. Long-read nanopore sequencing was performed and the mutational density distribution for the predicted library, the constructed Cas9 variant library, and the two replicates post-screen are plotted. **d**, Cas9 block composition and pre- and post-screen allele frequencies at each of the 18 mutational sites. Each block and site shows enrichment of the wild-type allele, but all sites retain a substantial fraction of mutant alleles.

a



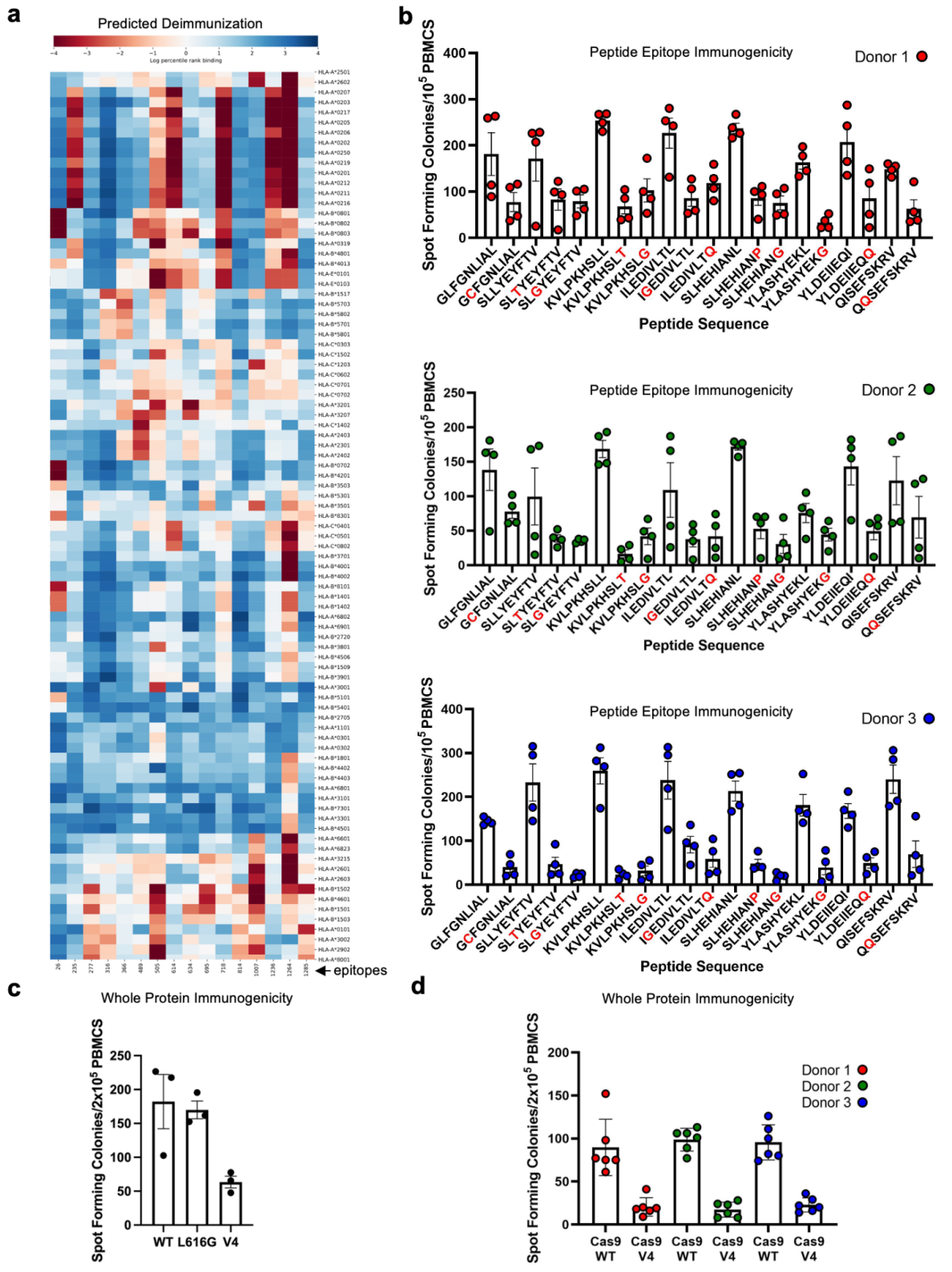
b



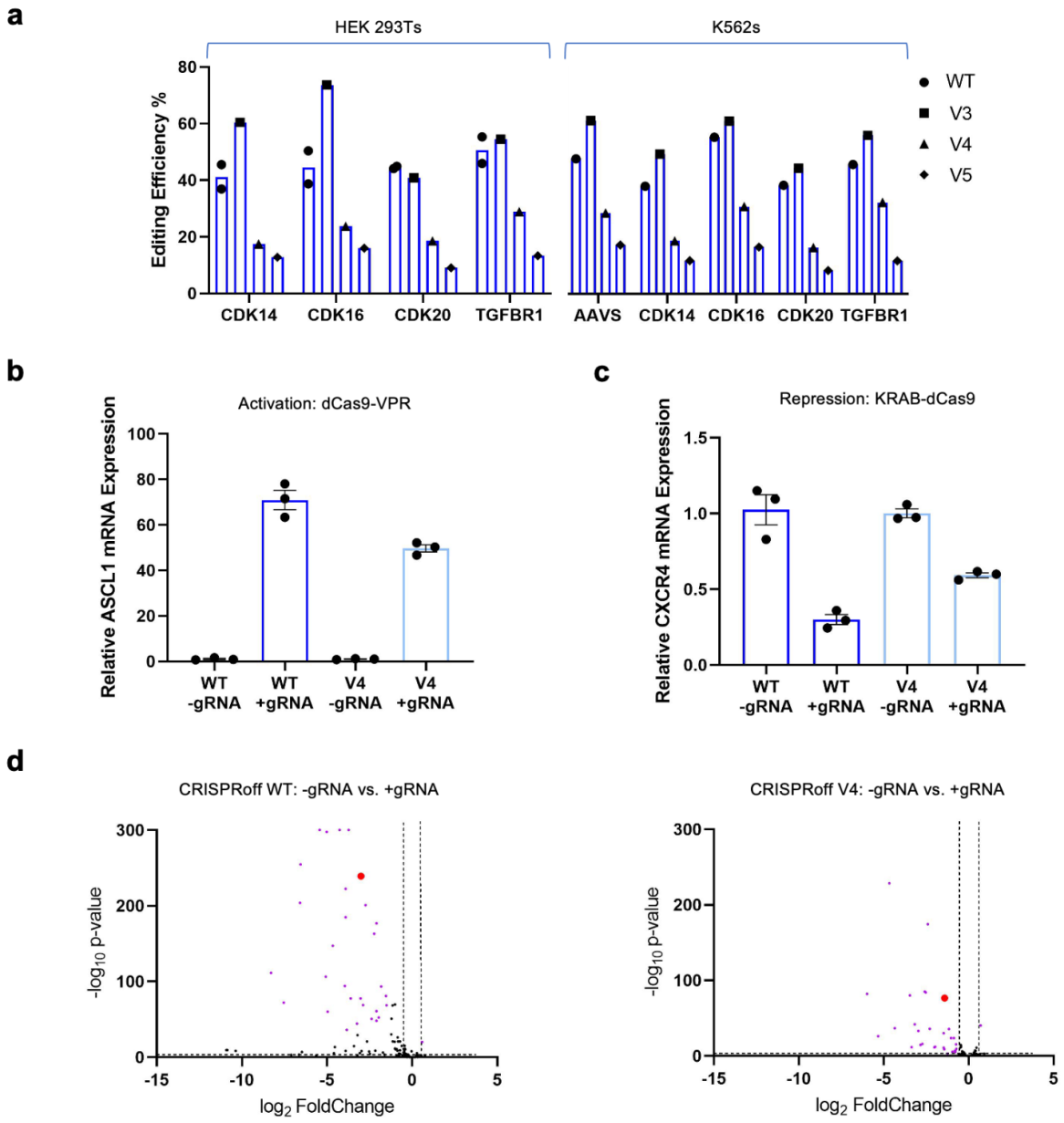
Variant Genotypes

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
V1	-	-	-	-	-	-	-	L616G	-	-	-	-	-	-	-	-	-	-
V2	-	-	-	-	-	-	-	-	L623Q	-	-	-	-	-	-	-	-	-
V3	-	-	-	-	-	-	L514T	-	L623Q	-	-	-	-	-	-	-	-	-
V4	-	-	Y286Q	-	-	-	L514T	-	L623Q	-	-	L727G	L816D	-	L1245G	-	L1282A	-
V5	-	-	Y286Q	-	S368C	-	L514T	-	L623Q	-	-	L727G	L816D	-	L1245G	-	L1282A	-
V6	P28L	-	Y286Q	-	S368C	-	L514T	-	L623Q	-	-	L727G	L816D	-	L1245G	-	L1282A	-
V7	-	-	Y286Q	S318H	S368C	-	L514T	-	L623Q	-	-	L727G	L816D	-	L1245G	-	L1282A	-
V8	-	L237C	-	-	S368C	F498T	-	L616G	-	-	-	-	-	Y1016G	-	-	-	Y1294Q
V9	-	L237C	-	-	S368C	F498T	-	L616G	-	-	-	-	-	Y1016G	-	I1273Q	-	Y1294Q
V10	-	-	-	S318H	-	F498T	L514G	-	L623Q	-	-	-	-	-	-	-	-	-
V11	-	-	-	-	-	F498T	-	-	L623Q	-	-	L727G	-	-	-	-	L1282A	Y1294Q
V12	-	-	-	-	-	-	-	-	L623Q	L636D	F704A	-	-	Y1016G	L1245G	-	-	-
V13	-	-	-	-	-	-	-	-	L623Q	L636D	F704A	-	L816D	-	L1245G	I1273Q	L1282E	-
V14	-	-	-	-	S368C	-	-	-	L623Q	-	-	L727P	-	Y1016G	-	I1273Q	-	-
V15	-	L237C	-	S318H	-	-	-	-	-	-	-	-	L816D	-	-	I1273Q	L1282E	Y1294Q
V16	-	-	-	S318C	-	-	L514T	-	-	L636D	F704A	-	-	Y1016G	-	-	-	Y1294Q
V17	-	-	-	S318H	S368C	-	-	-	-	-	F704A	L727G	L816D	Y1016G	L1245G	I1273Q	L1282A	Y1294Q
V18	P28L	-	-	S318H	S368C	F498T	-	L616G	-	-	F704A	-	-	Y1016K	-	I1273Q	L1282A	Y1294Q
V19	-	-	Y286Q	S318C	-	F498T	-	-	-	L636D	F704A	L727P	-	Y1016G	L1245G	-	L1282A	-
V20	-	-	-	S318C	-	-	-	-	L623Q	L636D	F704A	L727P	L816D	Y1016K	-	-	L1282A	Y1294Q

Supplementary Fig. 4 l Validations of LORAX screen identified Cas9 variants. a, Correlation between the fold change of a Cas9 variant and its predicted fold-change based on a k-nearest neighbors regression. Neighboring variants are those that share similar mutational patterns. The strong correlation suggests a smooth fitness landscape in which variants with similar mutation patterns will be more similar in fitness, on average, than those with divergent mutation patterns. **b**, Cas9 wildtype or variants V1-20 and sgRNA targeting the AAVS1 locus were introduced into HEK293T cells. NHEJ mediated editing at the AAVS1 locus was quantified via NGS for Cas9 WT and variants V1-20 is plotted. The number in parentheses represents the number of mutations in the variant. Variant genotypes are listed in the lower panel.



Supplementary Fig. 5 l Prediction and experimental confirmation of Cas9 epitope deimmunization. a, Predicted mutation-specific reduction in immunogenicity based on the epitope mutated and the HLA typing is depicted for each mutation included in the screen. **b**, Technical replicates of spot forming colonies in the ELISpot assay to assess peptide epitope immunogenicity are plotted for each donor (n=4). **c**, Technical replicates of spot forming colonies in the ELISpot assay to assess whole protein immunogenicity between WT Cas9, a sole L616G mutation, and V4 (n=3). **d**, Technical replicates of spot forming colonies in the ELISpot assay to assess whole protein immunogenicity are plotted for each donor (n=6).



Supplementary Fig. 6 | Characterization of Cas9 variants across genome and epigenome targeting assays. a, Cas9 wild-type or variants V3, V4, or V5, along with sgRNAs targeting the respective genes, were introduced into HEK293T and K562 cells. Editing efficiency of variants across 4 loci in HEK293Ts and 5 loci in K562s is plotted. **b**, ASCL1 mRNA expression in cells transfected with dCas9 WT-VPR or dCas9 V4-VPR and sgRNA or no sgRNA is shown. Values represented as mean +/- SEM (n=3). **c**, CXCR4 mRNA expression in cells transfected with dCas9 WT-KRAB or dCas9 V4-KRAB and sgRNA or no sgRNA is shown. Values represented as mean +/- SEM (n=3). **d**, Volcano plot demonstrating differentially expressed genes for CRISPRoff wildtype with or without the B2M guide and variant V4 with or without guide. Dotted lines indicate the cutoff for significance (log_2(fold change) greater than 0.5 or less than -0.5 and $-log_{10}$ p-value greater than 3). B2M downregulation is confirmed by the red dots. Differentially expressed genes found in both wildtype and V4 are labeled with purple dots.