

Engineering an RNA-based tissue-specific platform for genetic editing through use of a miRNA-enabled Cas12a

Rasmus Møller, Kohei Oishi, and Benjamin R. tenOever

Icahn School of Medicine at Mount Sinai, New York, NY 10029

ABSTRACT

The capacity to edit genomes using CRISPR-Cas systems holds immense potential for countless genetic-based diseases. However, one significant impediment preventing broad therapeutic utilization is *in vivo* delivery. While genetic editing at a single cell level *in vitro* can be achieved with high efficiency, the capacity to utilize these same biologic tools in a desired tissue *in vivo* remains challenging. Non-integrating RNA virus-based vectors constitute efficient platforms for transgene expression and surpass several barriers to *in vivo* delivery. However, the broad tissue tropism of viral vectors raises the concern for off-target effects. Moreover, prolonged expression of the Cas proteins, regardless of delivery method, can accumulate aberrant RNAs leading to unwanted immunological responses. In an effort to circumvent these shortcomings, here we describe a versatile RNA virus-based technology that can achieve cell-specific activity and self-inactivation by combining host microRNA (miRNA) biology with the CRISPR-Cas12a RNA-guided nuclease. Exploiting the RNase activity of Cas12a, we generated a vector that self-inactivates upon delivery of Cas12a and an accompanying CRISPR RNA (crRNA). Furthermore, we show that maturation of the crRNA can be made dependent on cell-specific miRNAs, which confers cell-specificity. We demonstrate that this genetic editing circuit delivers diminished yet sufficient levels of Cas12a to achieve effective genome editing whilst inducing a minimal immunological response. It can also function in a cell-specific manner thereby facilitating *in vivo* editing and mitigating the risk of unwanted, off-target effects.

INTRODUCTION

The CRISPR-Cas systems of archaea and many bacteria are sequence-specific adaptive defense systems that have evolved to cleave foreign nucleic acid¹. This defense system is dependent on acquisition and integration of foreign DNA protospacers in a process generally referred to as adaptation². Once integrated, expression of the so-called protospacers generate a precursor CRISPR RNA (pre-crRNA) which is further processed and matured to produce crRNA – comprised of a 5' direct repeat and the spacer. Finally, crRNA is bound by a Cas nuclease to elicit interference on incoming DNA as defined by complementarity of its guide RNA. Moreover, as the protospacer DNA is inherited, adaptation of a single prokaryotic cell can result in Lamarckian evolution for its offspring³.

While most of the Cas-nucleases only possess RNA-guided DNase activity, Cas12a also has RNase function⁴. The RNase function is responsible for processing the pre-crRNA by cleaving direct repeat sequences that flank the protospacer⁴. The crRNA that is generated as a result of these processing events is sufficient for instilling specificity onto the DNase activity of Cas12a. Similar to Cas9, Cas12a has also been repurposed as a eukaryotic gene editor⁵⁻⁸. However, as Cas12a biology is still in its infancy, its optimization lags behind that of Cas9. Despite this, the ability of Cas12a to process its own crRNA enables one to use it to generate the crRNA from diverse types of RNA so long as it is flanked by direct repeats⁹. This activity not only allows one to generate multiple crRNAs for any number of targets, but it has also enabled the generation of mRNAs that both code for Cas12a and the desired guides on a single transcript¹⁰. This

is in contrast to the most commonly applied Cas9 editing tool which demands a separate DNA dependent RNA polymerase for production of Cas9 and the single guide RNA^{6,11}.

Despite the immense potential of both the Cas9 and Cas12a systems, one significant impediment remains delivery of these large proteins alongside the desired crRNA(s). This challenge is formidable, especially when one wishes to efficiently edit a large number of cells to repair a genetic defect *in vivo*. This problem is further confounded by the fact that maintaining Cas expression for longer periods of time can result in the generation of off-target effects, aberrant RNAs, chromosomal translocations, and/or removal of the Cas-expressing cells^{12,13}. Given these challenges, the most optimal genetic editors would be delivered with the efficiency of a virus in a manner that was free of any possibility of genomic integration and would function only in a desired cell type for the time required to achieve editing. To this end, we designed a self-inactivating RNA that expresses Cas12a and processes its own crRNA in a cell-specific manner that is amenable to RNA virus-based delivery.

Having some parallels with the CRISPR-Cas platform, host miRNAs are rooted in defense systems albeit optimized for the targeting of RNA, rather than DNA¹⁴. Generally described as RNA interference (RNAi), this defense system acquires double stranded RNA (dsRNA) fragments from incoming virus and processes them into small interfering RNAs (siRNAs) similar in length than that of crRNAs. Furthermore, like crRNAs, siRNAs provide specificity to an otherwise non-specific nuclease. In the case

of RNAi, siRNAs are bound by a family of Argonaute (AGO) proteins with the capacity to cleave target mRNA in a sequence-specific manner¹⁵. The emergence of this system was so evolutionarily successful, that the genes involved in this biology duplicated and generated a similar small RNA based system commonly referred to as microRNAs (miRNAs)¹⁶. Unlike the small RNA-based antiviral systems, miRNAs derive from the host genome but are ultimately processed and utilized in a similar function, even sharing some overlap with the processing machinery of the Type II CRISPR-Cas system¹⁷. The general repurposing of this biology is believed to have happened independently in plants and animals but, in both examples, the biogenesis of miRNA-mediated regulation enabled another level of transcriptional control which coincided with multicellularity¹⁶.

One significant difference between antiviral RNAi and host miRNA biology is the complementarity between the small RNA and its target(s). In contrast to antiviral RNAi, miRNAs do not generally bind with perfect complementarity to their targets and therefore do not induce cleavage, owing to the fact that they are derived from host genes¹⁶. As a result, miRNAs are often considered ‘fine-tuners’ of host biology and are thought to act over the course of days or weeks¹⁸. In contrast, virus-derived siRNAs can engage their target with complete complementarity, as they derive from the pathogen itself, and thus enable both cleavage and the recycling of the small RNA – achieving silencing within hours¹⁶. However, as this difference in biology is almost completely defined by complementarity, if one introduces a perfect binding site for a given miRNA into a transcript, it will undergo potent silencing more akin to the normal

activity of antiviral RNAi¹⁹. This exploitation of miRNA biology therefore enables one to generate viruses or virus-based vectors which can be ubiquitously targeted or function in a species-, tissue-, cell-, or even signaling-specific manner based on endogenous miRNA expression²⁰⁻²⁶.

In an effort to generate an RNA-based DNA editor that functions in a cell-specific manner that would be amenable for *in vivo* use, we combined CRISPR-Cas and miRNA biology. In brief, we utilize the fact that Cas12a processes its own pre-crRNA to make a vector that delivers both Cas12a and crRNA and in doing so, inactivates itself. To this end, we encode crRNAs in the 3'-UTR of Cas12a and show that it leads to self-cleavage of its own transcript. Moreover, we demonstrate that delivery of this self-inactivating construct is sufficient to achieve efficient gene editing. And lastly, we show that processing of the pre-crRNA can be made to be dependent on miRNA-expression thereby conferring cell-type specificity on our editing platform.

RESULTS

In an effort to build a single mRNA transcript that could both self-inactivate and function in a cell-specific manner, we designed a construct encoding an enhanced green fluorescent protein (GFP) and an HA epitope-tagged Cas12a separated by a P2A peptide site²⁷ with its targeting instructions in the 3' UTR (Fig. 1a). To achieve self-inactivation, we incorporated a crRNA in the 3' untranslated region (UTR) of the GFP-Cas12a mRNA that consisted of a direct repeat and a spacer (DR). Accurate processing of this crRNA relies on a second motif found 3' of the spacer comprising a

perfect miRNA target (miR-T) site. In the presence of the cognate miRNA, Ago2 as part of the RNA induced silencing complex (RISC), will be recruited and result in 3' cleavage of the crRNA¹⁹. As miRNAs can be cell-specific, this synthetic construct would inactivate itself ubiquitously while only generating functional crRNA in a desired cell type where the cognate miRNA is present (Fig. 1a).

To initially characterize self-inactivation, we experimented with variations on the DR design to ascertain the relationship between Cas12a engagement and loss of GFP expression (Fig. 1b). To this end, we utilized either canonical direct repeats, direct repeats that would be poorly or unable to be cleaved by Cas12a (A18G and UUAA, respectively), or one in which the direct repeats were disrupted altogether (scrambled; scrbl)^{9,28}. To determine how these constructs would function, they were introduced into fibroblasts and monitored for GFP expression by both fluorescence microscopy and western blot (Fig. 1b-c). These data demonstrated that the GFP expression from the construct containing canonical direct repeats showed only low levels of fluorescence or expression by western blot which could also be correlated with HA-Cas12a expression. When the direct repeats were comprised of the A18G sites, fluorescence increased as compared to canonical sites. This enhanced expression could also be further corroborated by western blot analysis of both GFP and HA-Cas12a suggesting self-inactivation was diminished with the A18G sites. When the direct repeats were made to be uncleavable by Cas12a (UUAA), GFP expression was comparable to a construct lacking any direct repeats (scrbl). These data could be further validated by small RNA northern blot which indicated that wild type direct repeats, and to a lesser extent A18G,

was capable of generating a visible crRNA, inversely correlating to the GFP and HA-Cas12a protein levels (Fig. 1c-d).

We next sought to make correct processing of the crRNA dependent on miRNA-mediated cleavage to enable cell-specificity. To this end, we replaced the 3' direct repeat with two target sites for an endogenously expressed miRNA, miR-106a or an irrelevant control target sequence (ctrl-T) while keeping a wild-type direct repeat on the 5' end to mediate self-inactivation (Fig. 1b). GFP and Cas12a protein expression from these two constructs were comparable to that containing two canonical direct repeats indicating that a single direct repeat is sufficient for self-inactivation (Fig. 1b-c). In contrast, the crRNA is no longer processed when the 3' direct repeat is replaced with the control miRNA target sequence (ctrl-T) indicating a lack of cleavage (Fig. 1d). However, when ctrl-T is replaced with target sites corresponding to miR-106a, which is abundantly expressed in fibroblasts, we observe a distinct product corresponding to Ago2-based cleavage (Fig. 1d). As Ago2 cleaves its target site between bases 10-11¹⁶, the resulting crRNA contains an extended 3' terminus which should not impede Cas12a targeting or specificity²⁹.

To ascertain whether the product of 5' direct repeat and a 3' miRNA cleavage site remains functional, we next expressed variants of our RNA construct that encoded a crRNA targeting beta 2 microglobulin (B2M). In comparing transcripts lacking direct repeats (scrbl), having both direct repeats, or containing a 5' direct repeat with either a control 3' target sequence (ctrl) or miR-106a 3' sites we observe loss of MHC Class I, a

proxy for B2M targeting, only in conditions in which the 3' end of the spacer contained a wild type direct repeat or the miR-106a target sites (Fig. 1d). These data demonstrate a ~14% reduction of MHC1 with the canonical Cas12a targeting system which increases to greater than 30% targeting in the presence of miR-106a despite the extended crRNA (Fig. 1d-e). Together, these data suggest that miRNA biology can be exploited in conjunction with Cas12a-based processing to generate a single RNA capable of both self-inactivation and cell-specific targeting.

Given the capacity of a single transcript to both yield a functioning Cas12a editing platform and undergo self-inactivation, we next assessed whether this biologic circuit could be applied to a viral modality that would be most amenable to *in vivo* delivery. For safety reasons³⁰, this work was done with an RNA virus incapable of spread and containing a scrambled crRNA with no complementarity to a genomic sequence. Utilizing only the RNA dependent RNA polymerase (RdRp) of Nodamura virus and the 5' and 3' uncoding material required for RdRp recognition as described elsewhere³¹, we generated a self-replicating RNA (herein referred to as a replicon) to express GFP, Cas12a, and a 3' crRNA-containing UTR (Fig. 2a). As observed from our original plasmid design, launching of this replicon showed self-inactivation was also achieved using a viral-based delivery system (Fig. 2b). Comparable to what was observed using DNA, the RNA-based circuit equivalent showed self-targeting with canonical direct repeats (either one or two) with an intermediate phenotype observed for the A18G variant.

To determine whether the transcriptional response to our self-inactivating circuit would be amenable to *in vivo* use, we next performed bulk RNA sequencing to ascertain the transcriptional response to Cas12a expression and/or crRNA processing. To this end, we first compared the expression of Cas12a that was capable of self-inactivation and compared it to one incapable thereof. Surprisingly, this sequencing data set revealed that in contrast to sustained expression of Cas12a alone, the self-inactivating plasmid resulted in a significant number of differentially expressed genes (DEGs) (Fig. 3a and Supplementary Table 1). All upregulated genes with a log2fold change greater than 1 and an adjusted p-value less than 0.01 were annotated as belonging to the interferon response (Supplementary Table 1). These data would indicated that Cas12a processing of its own RNA results in a significant accumulation of aberrant RNA capable of inducing the host antiviral defenses. In contrast, this same comparison using the replicon-based platform yielded no DEGs (Fig. 3b). To determine if the lack of an interferon signature in response to the replicon-based platform was simply the result of having it generated in both conditions as a result of RdRp activity, we next compared the plasmid-based Cas12a system with processable crRNA to the equivalent replicon platform (Fig. 3c). This comparison yielded a larger number of DEGs but the interferon signature remained limited to plasmid-based delivery of Cas12a and crRNA demonstrating that the replicon self-inactivation is potent enough to prevent a cellular antiviral response (Supplementary Table 2). This was further corroborated by replicon read numbers which show that self-inactivation prevents any accumulation of either positive or negative sense transcripts that might otherwise serve as pathogen associated molecular patterns (Fig. 3d).

DISCUSSION

Here we present data demonstrating that RNA-based platforms can be designed to support safe and effective genetic editing. Based on the dual RNase and DNase properties of Cas12a, we show that RNA constructs can be engineered to be self-targeting. This attribute ensures that Cas12a and crRNA expression is temporal, thereby minimizing off-target editing and the accumulation of aberrant and potentially inflammatory RNA. In addition, self-inactivation also allows one to adopt a RNA virus-based delivery system, as self-cleavage minimizes cytotoxicity and ensures no persistence or genomic integration.

Having shown that an RNA-based replicon can deliver a self-inactivated Cas12a and crRNA in the absence of a transcriptional response, we next sought to engineer this construct so that it would only function in a desired cell type. In general, nucleic acid-based therapeutics and gene therapy vectors rely on promoter elements that are uniquely specific to a desired cell type. While this strategy has achieved some noteworthy successes, use of DNA as a vector introduces other unwanted issues including the need for entry into the nucleus and the possibility of genomic integration. RNA-based vectors mitigate this risk by having no DNA phase and performing all of their function in the cytoplasm³². Given these attributes, we chose to adapt miRNA-based targeting as a means of instilling cell-specific activity. Here we show that the addition of a perfect complementary miRNA target site can replace the 3' direct repeat needed to liberate a desired crRNA and thus render its biology cell specific. Together

with the knowledge that every tissue or cell-type has a unique miRNA profile, these data suggest that one can engineer an RNA-based vector to efficiently enter the cytoplasm and then function only in those cells where editing is desired.

The completion of the human genome project ushered in hope that treating genetic diseases would soon become possible³³. This dream was further fueled by the discovery and repurposing of the bacterial and archaeal CRISPR-Cas immune systems which provided unprecedented genome editing capabilities¹. Indeed, recent efforts have suggested that this promise is closer to reality than ever before as demonstrated through the *ex vivo* editing of human T cells for therapeutic purposes¹³. While this latter accomplishment may provide countless genetic treatments, the full potential of CRISPR-based therapeutic still will require a vehicle for *in vivo* editing. Here we provide a platform to enable this next advancement. By exploiting the biology of RNA viruses, Cas12a, and miRNAs, here we demonstrate that one can design a single RNA that could be systemically delivered with high efficiency but only function for a set amount of time and only within a desired cell lineage. The basic principle underlying this biology could be used with the replicon based strategy outlined here or with existing RNA virus based genetic editors³². Taken together, with the rapid advancements of synthetic biology, new CRISPR-Cas systems, and our understanding of virus-host interactions, our progress towards *in vivo* editing may allow us to view genetic-based errors as something that can be seamlessly overwritten.

METHODS

Cells

All fibroblasts used in this paper were HEK293T cells which were maintained in Dulbecco's Modified Eagle Medium (Gibco®) supplemented with 1x penicillin-streptavidin solution (Corning®) and 10% fetal bovine serum (Corning®).

Plasmids

All constructs were synthesized by Twist Biosciences. The sequence will be deposited on NCBI and the plasmid will be made available on AddGene following publication.

Western blot

Whole cell extract was prepared from live cells lysed in 1% NP-40 lysis buffer. 2µg (pEGFP plasmid-based experiments) and 70µg (nodamuravirus-based experiments) were analyzed by SDS-PAGE on a 4-15% acrylamide gradient gel (BIO-RAD®) and gels were subsequently blotted onto a 0.45µm nitrocellulose membrane (BIO-RAD®) and blocked in 5% milk in TBS for 1h at room temperature. Blots were probed with the following primary antibodies in 5% milk in TBS-T overnight at 4°C: anti-HA (clone HA-7, MilliporeSigma®), anti-GFP (ab290, Abcam®), anti-actin (clone Ab-5, Thermo Scientific®) and anti-GAPDH (G9545, MilliporeSigma®). After 4 x 5min washes in 1x TBS-T, blots were probed with HRP-linked secondary antibody for 1h at room temperature (anti-mouse, NA931V or anti-rabbit, NA934V, GE Healthcare®) and developed using the Immobilon Western HRP Substrate Kit (MilliporeSigma®).

Small RNA Northern blot

Total RNA was extracted from live cells using TRIzol (Invitrogen). Northern blot was performed as described in with 20µg total RNA per sample³⁴. Probes included the following: B2M-crRNA

(5'-GCTGGATAGCCTCCAGGCCA-3') miR-106a (5'-CTACCTGCACTGTAAGCACTTTT-3') and U6 (5'-GCCATGCTAATCTTCTCTGTATC-3'). Probes were labeled with ATP-P32 using T4 polynucleotide kinase (NEB®) and blot was exposed to a phosphor screen (GE®) and developed on a Typhoon TRIO Storage Phosphorimager (GE®).

Flow cytometry

Roughly 7.5×10^5 cells/well were plated on 6-well plates. After attaching overnight, cells were transfected using lipofectamine 2000 (Invitrogen®) and were passaged 1:5 when they reached ~80% confluency for up to ten days. For flow cytometry analysis, cells were trypsinized and washed 2 x stain buffer (FBS) (BD Biosciences®) and stained using the BD Cytofix/Cytoperm Fixation/Permeabilization Kit (BD Biosciences®). The following antibodies and dyes were used: anti-human HLA-A,B,C Pacific Blue (clone W6/32, BioLegend®), anti-HA Alexa Fluor 647 (clone HA.11, BioLegend) and LIVE/DEAD stain Aqua (ThermoFisher®). Fixed cells were analyzed on a 2019 Attune NxT Flow Cytometer. Data processing was done with FlowJo v. 10.6.2.

RNA sequencing

Total RNA was extracted using TRIzol (Invitrogen®). 1µg of total RNA per sample was used as starting material for building Illumina libraries for deep sequencing. Libraries were built following Illumina protocols using the TruSeq Stranded mRNA library prep kit (Illumina®). Libraries were enriched for polyadenylated transcripts and sequenced on an Illumina NextSeq 500. Alignment of reads was done using STAR alignment to the human reference genome (hg38) and differential gene expression and statistics were determined by the DEseq2 protocol. Reads mapping to the Nodamuravirus replicon were aligned using Bowtie2.

FIGURES

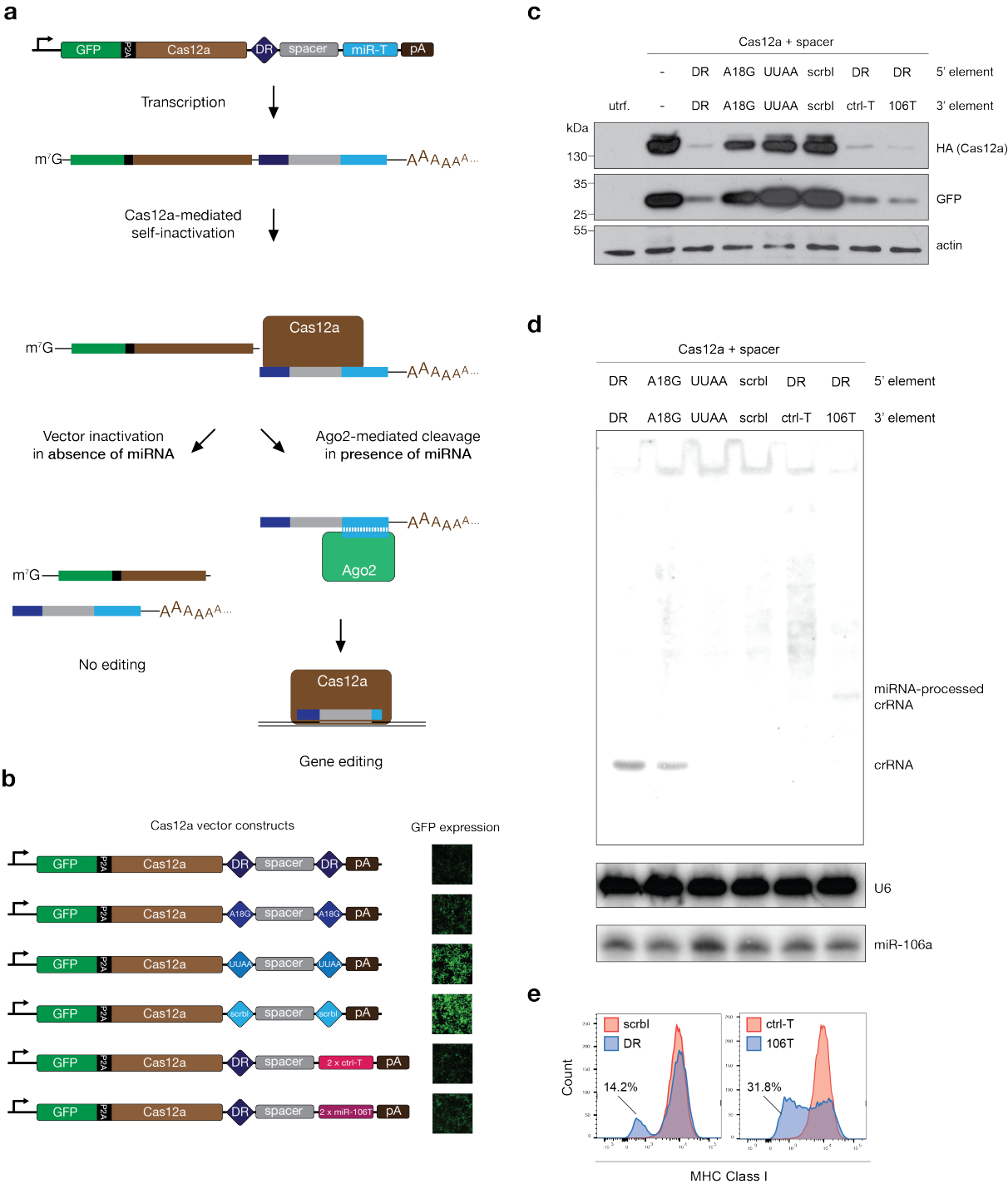
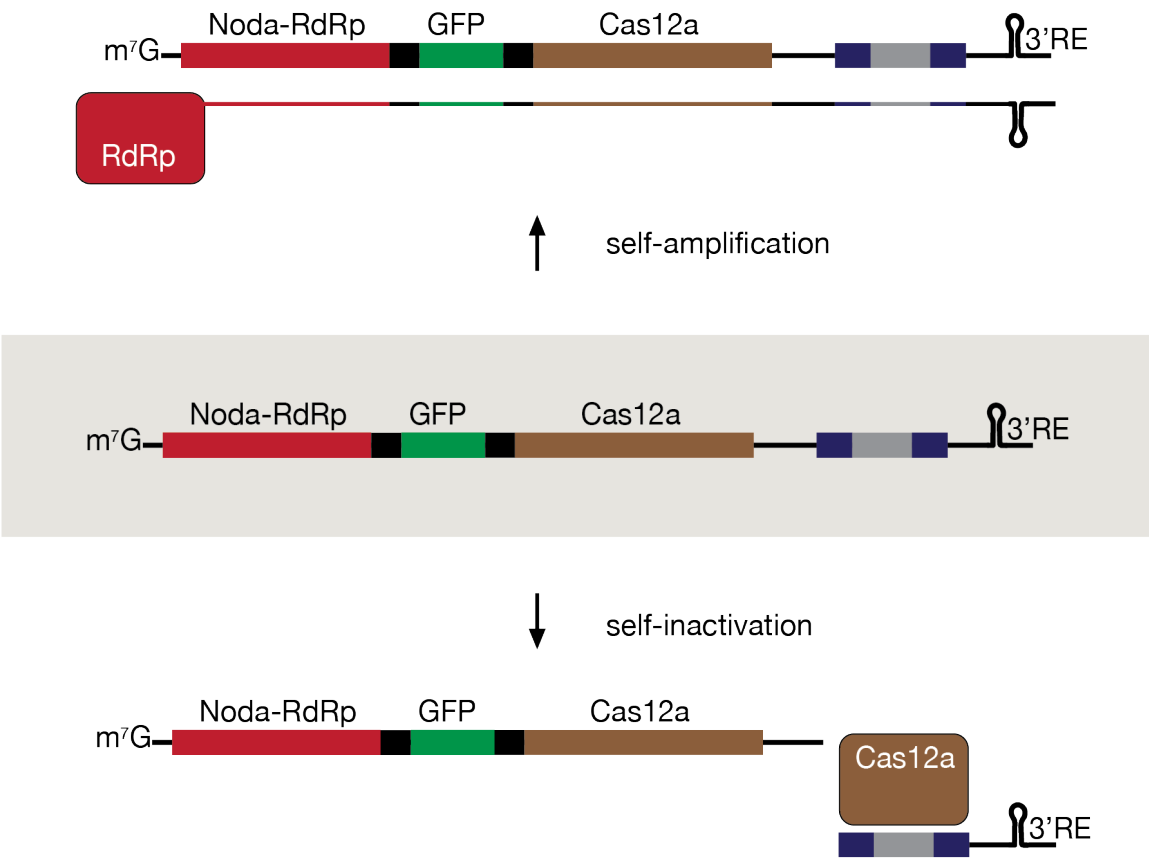


Fig. 1: Self-inactivation and tissue specificity of Cas12a vector.

a, A schematic of how self-inactivation and miRNA-based tissue specificity is incorporated into one single Cas12a vector. GFP and Cas12a are encoded in one open reading frame separated by a P2A site. The 3'-UTR of the same transcript consists of a direct repeat (DR), a spacer with complementarity to a genomic target and a sequence with perfect complementarity to a miRNA referred to as a miRNA-target site (miR-T) and lastly an SV40 polyadenylation signal (pA). **b**, Variations on the construct depicted in **a** with a mutation of nucleotide 18 in the direct repeat from an A to a G (A18G), inversion of nucleotides 16-19 from AAUU to UUAA (UUAA) or a scrambled sequence of the entire direct repeat (scrbl). The last two constructs have two target sites to either a control-miRNA (ctrl-miRT) or miR-106a (miR-106T). Fluorescence images are representative of GFP expression 48 hours after transfection with the constructs indicated. **c**, Western blot of cells transfected for 48 hours with the constructs depicted in **b**, -, no crRNA in 3'-UTR, utr., untransfected cells. **d**, Small RNA Northern blot of total RNA from cells as in **c**, U6, U6 snRNA (loading control). **e**, MHC Class I cell surface expression measured ten days by flow cytometry ten days post transfection. Data from cells transfected with the constructs overlaid as indicated.

a



b

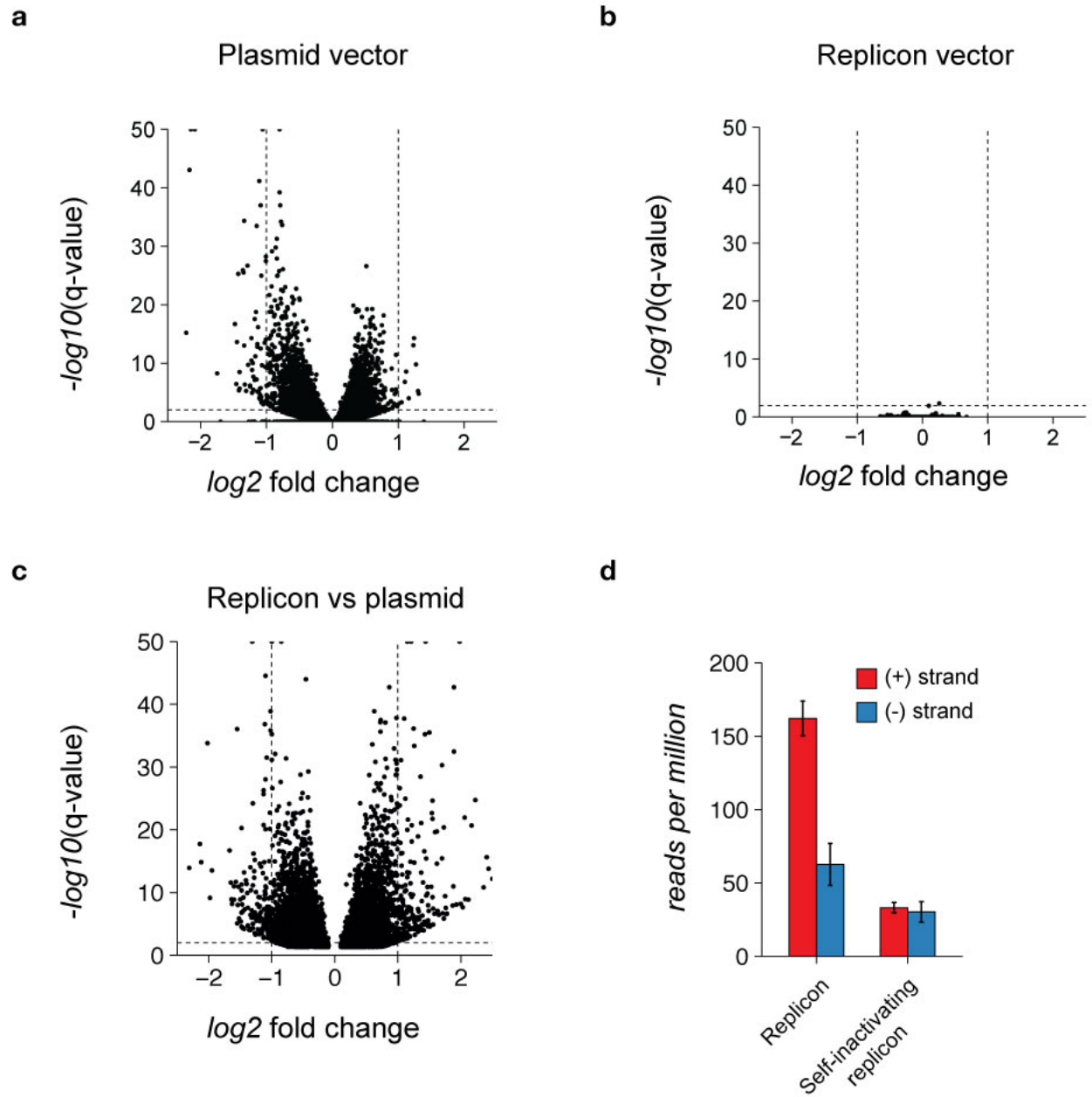


334

335 **Fig. 2: Delivery of Cas12a from a self-inactivating replicon.**

336 **a**, Schematic of a self-inactivating replicon construct. Nodamuravirus RNA dependent RNA
 337 polymerase (Noda-RdRp), GFP and Cas12a are encoded in one open reading frame separated
 338 by P2A sites. Downstream of the open reading frame, a spacer (grey) flanked by two direct
 339 repeats (dark blue) are inserted. **b**, Western blot of cells transfected with the replicon constructs
 340 indicated (as in Fig. 1c).

341



342

343

344

345

346

347

348

Fig. 3: Transcriptional response to self-inactivating Cas12a vectors.

a, Plot depicting differential gene expression of host genes in cells transfected with a plasmid-based Cas12a construct containing direct repeats in the 3'-UTR compared to cells transfected with a comparable construct without direct repeats. Each dot represents a gene plotted by its log₂ fold change between the two conditions and -log₁₀ of the adjusted p-value (q) determined based on triplicate samples. Horizontal line marks a q-value = 0.01 and Vertical lines mark a log₂ fold change of -1 and 1. **b**, same as **a**, but comparing replicon-based Cas12a construct containing direct repeats against a comparable construct without direct repeats. **c**, same as **a**, but comparing plasmid-based Cas12a with direct repeats to replicon-based Cas12a with direct repeats. **d**, Stranded read numbers aligning to the replicon as number of reads per million of total reads. Error bars represent standard deviation from three replicates.

SUPPLEMENTARY MATERIAL

Table 1. Differentially expressed genes in response to self-inactivating vs. non-processable plasmid-based Cas12a vector. The control group are fibroblasts expressing GFP, Cas12a, and a crRNA lacking direct repeats performed in triplicate. The experimental group are fibroblasts expressing GFP, Cas12a, and a crRNA flanked with direct repeats, also performed in triplicate. Total reads per sample are given. Shown are gene names, mean count per gene, log2fold change in comparing triplicate samples in deSeq2, Standard error and adjusted p-value (q value). Raw data to be deposited on NCBI GEO.

Table 2. Differentially expressed genes in response to self-inactivating plasmid-based Cas12a versus the corresponding replicon-based construct. The control group are fibroblasts expressing GFP, Cas12a, and a crRNA flanked with direct repeats performed in triplicate. The experimental group are fibroblasts expressing GFP, Cas12a, and a crRNA flanked with direct repeats as delivered by RNA replicon, also performed in triplicate. Total reads per sample are given. Shown are gene names, mean count per gene, log2fold change in comparing triplicate samples in deSeq2, Standard error and adjusted p-value (q value). Raw data to be deposited on NCBI GEO.

REFERENCES

- 1 Marraffini, L. A. & Sontheimer, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**, 181-190, doi:10.1038/nrg2749 (2010).
- 2 Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-736, doi:10.1038/nrmicro3569 (2015).
- 3 van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M. & Brouns, S. J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* **34**, 401-407, doi:10.1016/j.tibs.2009.05.002 (2009).
- 4 Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A. & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517-521, doi:10.1038/nature17945 (2016).
- 5 Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232, doi:10.1038/nbt.2507 (2013).
- 6 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).
- 7 Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**, 227-229, doi:10.1038/nbt.2501 (2013).
- 8 Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826, doi:10.1126/science.1232033 (2013).
- 9 Zetsche, B. *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol* **35**, 31-34, doi:10.1038/nbt.3737 (2017).
- 10 Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat Methods* **16**, 887-893, doi:10.1038/s41592-019-0508-6 (2019).

424 11 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive
425 bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).

426 12 Koo, T., Lee, J. & Kim, J. S. Measuring and Reducing Off-Target Activities of
427 Programmable Nucleases Including CRISPR-Cas9. *Mol Cells* **38**, 475-481,
428 doi:10.14348/molcells.2015.0103 (2015).

429 13 Stadtmauer, E. A. *et al.* CRISPR-engineered T cells in patients with refractory cancer.
430 *Science* **367**, doi:10.1126/science.aba7365 (2020).

431 14 tenOever, B. R. The Evolution of Antiviral Defense Systems. *Cell Host Microbe* **19**, 142-
432 149, doi:10.1016/j.chom.2016.01.006 (2016).

433 15 Mello, C. C. & Conte, D., Jr. Revealing the world of RNA interference. *Nature* **431**, 338-
434 342, doi:10.1038/nature02872 (2004).

435 16 Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20-51, doi:10.1016/j.cell.2018.03.006
436 (2018).

437 17 Aguado, L. C. *et al.* RNase III nucleases from diverse kingdoms serve as antiviral
438 effectors. *Nature* **547**, 114-117, doi:10.1038/nature22990 (2017).

439 18 Aguado, L. C. *et al.* microRNA Function Is Limited to Cytokine Control in the Acute
440 Response to Virus Infection. *Cell Host Microbe* **18**, 714-722,
441 doi:10.1016/j.chom.2015.11.003 (2015).

442 19 tenOever, B. R. RNA viruses and the host microRNA machinery. *Nat Rev Microbiol* **11**,
443 169-180, doi:10.1038/nrmicro2971 (2013).

444 20 Brown, B. D., Venneri, M. A., Zingale, A., Sergi Sergi, L. & Naldini, L. Endogenous
445 microRNA regulation suppresses transgene expression in hematopoietic lineages and
446 enables stable gene transfer. *Nat Med* **12**, 585-591, doi:10.1038/nm1398 (2006).

447 21 Perez, J. T. *et al.* MicroRNA-mediated species-specific attenuation of influenza A virus.
448 *Nat Biotechnol* **27**, 572-576, doi:10.1038/nbt.1542 (2009).

- 22 Langlois, R. A. *et al.* MicroRNA-based strategy to mitigate the risk of gain-of-function influenza studies. *Nat Biotechnol* **31**, 844-847, doi:10.1038/nbt.2666 (2013).
- 23 Langlois, R. A., Varble, A., Chua, M. A., Garcia-Sastre, A. & tenOever, B. R. Hematopoietic-specific targeting of influenza A virus reveals replication requirements for induction of antiviral immune responses. *Proc Natl Acad Sci U S A* **109**, 12117-12122, doi:10.1073/pnas.1206039109 (2012).
- 24 Pham, A. M., Langlois, R. A. & TenOever, B. R. Replication in cells of hematopoietic origin is necessary for Dengue virus dissemination. *PLoS Pathog* **8**, e1002465, doi:10.1371/journal.ppat.1002465 (2012).
- 25 Moller, R. *et al.* miRNA-mediated targeting of human cytomegalovirus reveals biological host and viral targets of IE2. *Proc Natl Acad Sci U S A* **115**, 1069-1074, doi:10.1073/pnas.1719036115 (2018).
- 26 Aguado, L. C. *et al.* Homologous recombination is an intrinsic defense against antiviral RNA interference. *Proc Natl Acad Sci U S A* **115**, E9211-E9219, doi:10.1073/pnas.1810229115 (2018).
- 27 Sharma, P. *et al.* 2A peptides provide distinct solutions to driving stop-carry on translational recoding. *Nucleic Acids Res* **40**, 3143-3151, doi:10.1093/nar/gkr1176 (2012).
- 28 Zhong, G., Wang, H., Li, Y., Tran, M. H. & Farzan, M. Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. *Nat Chem Biol* **13**, 839-841, doi:10.1038/nchembio.2410 (2017).
- 29 Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J. & Russell, R. Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. *Mol Cell* **71**, 816-824 e813, doi:10.1016/j.molcel.2018.06.043 (2018).
- 30 Baltimore, D. *et al.* Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* **348**, 36-38, doi:10.1126/science.aab1028 (2015).

- 475 31 Gitlin, L., Hagai, T., LaBarbera, A., Solovey, M. & Andino, R. Rapid evolution of virus
476 sequences in intrinsically disordered protein regions. *PLoS Pathog* **10**, e1004529,
477 doi:10.1371/journal.ppat.1004529 (2014).
- 478 32 Park, A. *et al.* Sendai virus, an RNA virus with no risk of genomic integration, delivers
479 CRISPR/Cas9 for efficient gene editing. *Mol Ther Methods Clin Dev* **3**, 16057,
480 doi:10.1038/mtm.2016.57 (2016).
- 481 33 Collins, F. S., Morgan, M. & Patrinos, A. The Human Genome Project: lessons from
482 large-scale biology. *Science* **300**, 286-290, doi:10.1126/science.1084564 (2003).
- 483 34 Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of
484 small RNA. *Nat Protoc* **3**, 1077-1084, doi:10.1038/nprot.2008.67 (2008).

485