



Published in final edited form as:

*Acta Biomater.* 2016 April 1; 34: 143–158. doi:10.1016/j.actbio.2015.12.036.

## High content analysis platform for optimization of lipid mediated CRISPR-Cas9 delivery strategies in human cells

Benjamin Steyer<sup>1</sup>, Jared Carlson-Stevermer<sup>1,2</sup>, Nicolas Angenent-Mari<sup>2</sup>, Andrew Khalil<sup>2</sup>, Ty Harkness<sup>1,2</sup>, and Krishanu Saha<sup>1,2,3,\*</sup>

<sup>1</sup>Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI, USA

<sup>2</sup>Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI, USA

<sup>3</sup>Department of Medical History and Bioethics, University of Wisconsin-Madison, Madison, WI, USA

### Abstract

Non-viral gene-editing of human cells using the CRISPR-Cas9 system requires optimized delivery of multiple components. Both the Cas9 endonuclease and a single guide RNA, that defines the genomic target, need to be present and co-localized within the nucleus for efficient gene-editing to occur. This work describes a new high-throughput screening platform for the optimization of CRISPR-Cas9 delivery strategies. By exploiting high content image analysis and microcontact printed plates, multi-parametric gene-editing outcome data from hundreds to thousands of isolated cell populations can be screened simultaneously. Employing this platform, we systematically screened four commercially available cationic lipid transfection materials with a range of RNAs encoding the CRISPR-Cas9 system. Analysis of Cas9 expression and editing of a fluorescent mCherry reporter transgene within human embryonic kidney cells was monitored over several days after transfection. Design of experiments analysis enabled rigorous evaluation of delivery materials and RNA concentration conditions. The results of this analysis indicated that the concentration and identity of transfection material have significantly greater effect on gene-editing than ratio or total amount of RNA. Cell subpopulation analysis on microcontact printed plates,

\*To whom correspondence should be addressed. Tel: 608-316-4313; ksaha@wisc.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### APPENDICES

- 7.1. Appendix A: Oligonucleotides employed in this study
- 7.2. Appendix B: Two-way ANOVA for comparison of image and flow cytometry
- 7.3. Appendix C: Conditions used in DOE 1
- 7.4. Appendix D: Conditions used in DOE 2
- 7.5. Appendix E: Additional univariate analysis from  $\mu$ Feature plate editing
- 7.6. Appendix F: Validation of PSLR model
- 7.7. Appendix G: Gating conditions for GFP intensity cell sorting

### AUTHOR CONTRIBUTIONS

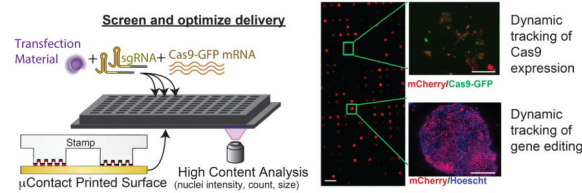
B.G.S. and K.S. conceived and supervised the whole project. B.G.S. performed the major research and wrote the manuscript with input from all authors. J.C.S., A.K., N.A., T.H. provided the technical support.

### DISCLOSURES

The authors declare no competing financial interests.

further revealed that low cell number and high Cas9 expression, 24 hours after CRISPR-Cas9 delivery, were strong predictors of gene-editing outcomes. These results suggest design principles for the development of materials and transfection strategies with lipid-based materials. This platform could be applied to rapidly optimize materials for gene-editing in a variety of cell/tissue types in order to advance genomic medicine, regenerative biology and drug discovery.

## Abstract



## Keywords

CRISPR-Cas9; gene-editing; cationic lipids; design of experiments; high content analysis; DNA/RNA; human embryonic cells

## 1. INTRODUCTION

Genomic medicine involving gene-editing to correct disease causing mutations or insert other genetic sequences into patients' cells is a growing area of biomedical research [1–6]. Recent gene-editing technologies utilize nucleases to generate a DNA double strand break (DSB), or “cut,” in genomic DNA at a desired location [7–11]. CRISPR-Cas9, an emerging gene-editing technology, exploits a modified bacterial immune defense mechanism termed CRISPR (clustered regularly interspaced short palindromic repeats) that cuts DNA at specific sequences.

The engineered CRISPR-Cas9 system encompasses two essential components: 1) an endonuclease, Cas9; and 2) a short, single-guide RNA (sgRNA) that forms a ribonucleoprotein (RNP) complex with Cas9 and targets endonuclease activity to a specific sequence in the genome [12,13]. Interaction of an approximately 20 nucleotide (nt) sgRNA sequence to the complementary genomic DNA increases the residence time of the Cas9-sgRNA complex at that specific genomic locus, enabling Cas9 nuclease to create targeted DSBs. However, the Cas9-sgRNA complex can associate with and cut the genome at other “off-target” sites [14,15]. Decreasing off-target mutagenesis has been achieved by titrating the quantity of Cas9 and sgRNA delivered [16,17] and by engineering the specificity of CRISPR-Cas9 components [18–20]. While RNP engineering efforts have increased efficiencies of editing and reduced off-target mutagenesis, a significant bottleneck in genomic medicine remains in effectively delivering these engineered components to human cells.

Employing CRISPR-Cas9 gene-editing requires cellular delivery and subsequent nuclear translocation of RNP complexes or plasmid DNA and RNAs that encode the two components of the system. Initial CRISPR-Cas9 experiments with human cells used

electroporation of plasmids encoding Cas9 and sgRNAs driven by constitutive promoters [9,10]. Viral delivery strategies have also been employed [21], although non-viral delivery strategies are typically preferred over viral delivery. This is because viruses can integrate into the genome causing insertional mutagenesis, which is problematic for many research and clinical applications.

A common method for delivering Cas9 components to cells makes use of synthetic biomaterials that form lipid nanoparticles containing DNA, RNA, or pre-formed RNP complexes [22–24]. These methods involve encapsulation or complexing of nucleic acid cargo through interactions between the negatively charged phosphate backbone of the cargo and the positively charged lipid head groups. In the case of RNPs, the highly negative charge of the sgRNA allows the RNP complex to be encapsulated by the cationic lipids [16]. Cellular uptake and subcellular trafficking can be mediated by endocytosis and macropinocytosis, although exact mechanisms are poorly understood and may vary widely across transfection reagents [25]. Direct comparison of liposomal based transfection of a Cas9-encoding plasmid to Cas9-encoding mRNA indicated that mRNA increased editing efficiencies and lowered off-target mutagenesis in many human cell lines [17]. However, each human cell line required different doses and formulations of lipid nanoparticles or complexes. Despite these challenges, lipid-based delivery of mRNA encoding Cas9 provides an attractive route to achieving precise editing in human cells.

Optimizing CRISPR-Cas9 gene-editing with RNA in human cells is a complex multiparametric space. For each cell line in research and each patient in the clinic, screening of many different formulations of the mRNA/sgRNA/material complex will likely be required for optimal editing [23,24]. Combinatorial synthesis and assessment of non-viral transfection agents for delivering plasmid DNA [26] and siRNA [27,28] have yielded successful strategies both *ex vivo* and *in vivo* [29]. To our knowledge these platforms have not been applied for optimizing CRISPR-Cas9 delivery. Indeed, there are limited platforms available to rapidly screen synthetic biomaterial transfection agents and optimize the amount, ratios, and chemical modifications to both components of the CRISPR-Cas9 system. Thus far, optimization has relied on low-throughput endpoints assays on samples homogenized from  $10^4$ - $10^6$  cells [17,16]. Such cell population averaged measurements of DSB formation utilize DNA sequencing or an endonuclease assay (e.g., SURVEYOR [30] or T7E1 [31]) of genomic DNA. These assays can be difficult to standardize [32,33] and do not monitor the cargo trafficking and editing processes *in situ*. Thus, the development of new materials with these low-throughput assays can be slow, and the delivery kinetics involved in optimized strategies are poorly defined.

Here we describe a high throughput method to optimize non-viral gene-editing strategies. Our platform exploits image cytometry and high content image analysis (HCA) in combination with a customized microcontact printed cell substrate [34,35] to simultaneously monitor non-viral delivery and editing in human cells. We use established image analysis methods [36] and design-of-experiments (DOE) based statistical techniques [37,38] to screen existing liposomal delivery materials and to optimize delivery inputs to maximize gene-editing outcomes. Dynamic tracking of Cas9 protein expression, subcellular localization and gene disruption within subpopulations of cells suggests that the cell number

and the level of Cas9 expression within 24 hours of delivery are important predictors of editing. This platform enables the rapid screening of material mediated CRISPR-Cas9 gene-editing strategies in human cells.

## 2. MATERIALS AND METHODS

### 2.1. H2B-mCherry Reporter Construction and Culture

A constitutively expressed histone 2B-mCherry reporter transgene was integrated into the genome of human embryonic kidney (HEK) 293T cells via CRISPR-Cas9 gene-editing. The H2B-mCherry plasmid was generated by cloning the H2B-mCherry sequence (Addgene #20972) into the EGFP sequence of the AAV-CAGGS-EGFP plasmid (Addgene #22212). This plasmid was electroporated with plasmids encoding human codon optimized Cas9 (Addgene #41815) and sgRNA targeting the *AAVS1* locus (Addgene #41818) into HEK 293T cells. Populations containing the transgene were purified through puromycin selection and clonal isolation followed by sequencing and fluorescent imaging to confirm expression.

HEK H2B-mCherry cells were maintained at 37°C and 5% CO<sub>2</sub> in growth media composed of DMEM (Thermo Scientific), 10% v/v FBS (Thermo Scientific), 2 mM L-Glutamine (Thermo Scientific), and 50 U/mL Penicillin-Streptomycin (Thermo Scientific). Media was changed every 2 days. Cells were passaged 1:40 every 4-5 days once 80-90% confluent with 0.05% Trypsin-EDTA (Thermo Scientific) onto tissue culture polystyrene (TCPS) plates (Thermo Scientific) coated overnight with a sterile 0.1% (w/v) gelatin A (Sigma) in RO water solution. At least two passages prior to experimental treatment, cells were transitioned to imaging media containing FluoroBrite DMEM (Thermo Scientific) supplemented with 10% v/v FBS and 2 mM L-Glutamine. Specific culture volumes and media change regimens during experimentation are described in Section 2.9.

### 2.2. Single-guide RNA (sgRNA) Design, Synthesis and Characterization

Coding sequences for genomic targets were obtained from NCBI Gene ([www.ncbi.nlm.nih.gov/gene](http://www.ncbi.nlm.nih.gov/gene)) and imported into an online sequence management tool (Benchling, [www.benchling.com](http://www.benchling.com)). Potential sgRNA sites were identified using Benchling's online genome-editing design tools. For mCherry reporter targeting, sgRNAs were designed against the mCherry coding sequence. sgRNA sequences were selected based on high on-target [39] and low off-target [40] ranking. In general, three sgRNAs were screened for each locus (sequences available in Appendix, Table A.1). Initial delivery work that monitored the amount of genomic disruption by a T7E1 assay identified "mCherry C" as the most effective sgRNA screened (data not shown).

Once designed, sgRNAs were synthesized as described previously [34]. Briefly, to construct template DNA for *in vitro* transcription (IVT), a 60 nt sgRNA specific forward primer, containing a truncated T7 promoter and the 20 nt target sequence, as well as a universal reverse primer were used to amplify a synthetic double stranded DNA template (Integrated DNA Technologies) encoding the conserved trans-activating CRISPR RNA (tracr) sequence. PCR was performed using Phusion High-Fidelity Polymerase (New England Biolabs) according to manufacturer protocols. IVT was performed using the HiScribe T7

transcription kit (New England Biolabs, E2040S) and incubated at 37°C overnight. For quantification and characterization of sgRNAs, purification was performed with the MEGAclean Transcription Clean-Up Kit (Ambion, AM1908) and quantified on a Nanodrop 2000 (Thermo Scientific). All sgRNAs used in this study were purified, however, unpurified sgRNAs have been delivered to cells resulting in measurable gene-editing with minimal toxicity (data not shown). A 2100 Bioanalyzer (Agilent Technologies) was used according to manufacturer's protocols to confirm size and uniformity of the sgRNA.

### 2.3. Transcription and Preparation of mRNA

Transcription template was prepared from plasmid vector pMJ920 (Addgene# 42234). The plasmid was linearized via overnight digestion with NsiI and XbaI (New England Biolabs). Following overnight incubation at 37°C, the reaction was precipitated using Ammonium-Acetate/EDTA/Ethanol. For each 200 µL IVT reaction, 20 µL of 3 M Ammonium Acetate, 37.5 µL of 0.5 M EDTA, and 400 µL of 200 proof Ethanol were added. This was incubated overnight at -20°C, then centrifuged at 15,000 g for 15 minutes at 4°C in a tabletop microcentrifuge. The supernatant was aspirated, and pellet was air-dried for 5-10 minutes at room temperature. The pellet was then resuspended in 25 µL of nuclease-free water, and concentration was measured using absorbance (NanoDrop 2000, Thermo Scientific) before diluting to 120 ng/µL. Aliquots were stored at -80°C.

Cas9-GFP mRNA was transcribed using the HiScribe T7 ARCA mRNA kit (with tailing) (New England Biolabs, E2060S) in 60 µL reactions as per the manufacturer's instructions. For each reaction, 3 µg of linearized pMJ920 template was added to 30 µL of 2× ARCA/NTP mix, 21 µL of nuclease-free water, and 6 µL of T7 RNA polymerase mix was incubated at 37°C for 30 minutes. After transcription, 6 µL of DNase I mix was added to the reaction and another 37°C incubation was carried out for 15 minutes. Finally, 60 µL of nuclease-free water, 15 µL of 10× poly(A) polymerase reaction buffer, and 15 µL of poly(A) polymerase were added to the reaction and incubated at 37°C for 30 minutes.

To purify the mRNA, 75 µL of LiCl solution (provided with IVT kit) was added to the reaction and then incubated overnight at -20°C. Precipitated mRNA was centrifuged for 15 minutes at 4°C in a tabletop microcentrifuge at 15,000 g. The supernatant was then carefully aspirated. The pellet washed with 500 µL cold 70% ethanol, and centrifuged at 15,000 g at 4°C for 10 minutes. Ethanol was carefully aspirated and the pellet was air dried for ~1-2 minutes at room temperature. The pellet was resuspended in TE buffer and solubilized by heating to 65°C for 10 minutes. mRNA concentration was determined using absorbance (NanoDrop 2000). The typical reaction yielded ~50 µg of product. mRNA was diluted to 500 ng/µL in TE buffer and 20 µL aliquots were stored at -80°C.

Cas9-GFP mRNA was labeled after IVT using Label IT Tracker Cy5 (Mirrus, #7021) according to manufacturer's protocols. After labeling, mRNA was again purified using the LiCl method described above.

### 2.4. T7 Endonuclease 1 Assay

Genomic DNA was harvested using QuickExtract (Epicentre) and the desired target locus was PCR amplified (primers for each locus available in the Table A.2) in a 100 µL reaction

using AccuPrime Taq polymerase (Thermo Scientific) in 1× AccuPrime Buffer II for 35 cycles. The PCR products were column purified and eluted in 30 µL nuclease free water and the concentration of each sample was measured via absorbance (NanoDrop 2000). Following concentration determination, 200 ng of PCR product was run for 45 minutes at 120V on 2% agarose/1× TAE gel with 1× SYBR Safe stain (Thermo Scientific). The intensity of each band was evaluated and relative DNA concentration of each sample was adjusted to ensure uniform loading between samples. Generally only several samples required adjustment where the absorbance readings greatly overestimated the DNA concentration. Using the adjusted concentrations, 400 ng of purified product was brought up to a total volume of 19 µL in 1× NEBuffer 2 (New England Biolabs). Each 19 µL sample was then run through a hybridization protocol in a Nexus Thermocycler (Eppendorf) which was as follows: 10 min at 95°C, a 2°C/s ramp down to 85°C, 1 min at 85°C, and a 0.1°C/s ramp down to 25°C. Following hybridization, 1 µL of T7 endonuclease I (New England Biolabs) was added and samples were incubated for 15 minutes at 37°C. Immediately following incubation, 1 µL of 0.5 M EDTA was added to inhibit enzyme activity. The entire volume of each sample was mixed with 5 µL of 50% (v/v) glycerol in nuclease free water loading buffer and run on a 2% agarose/1× TAE gel with 1× SYBR Safe for 30-45 minutes at 120 V. The intensity of the resulting cut and uncut bands were evaluated using Image Lab Software (BioRad). Percent gene modification was determined using the following formula (where a is the intensity of the uncut band, and b and c are the intensities of the two cut bands): % Gene Modification =  $100 * [1 - (1 - (b+c)/(a+b+c))^{1/2}]$ .

## 2.5. Flow Cytometry

Flow cytometry was conducted on samples fixed with 4% paraformaldehyde using a BD Canto cytometer (BD Biosciences). Cytometer was fitted with a high throughput sampler system and 8-color (4YG-2R-2B) optics. FSC-A vs SSC-A was used to gate for desired cell population and FSC-A vs FSC-H was used to gate for single cells. Single cell mCherry fluorescence was detected using a 561nm excitation laser and a 610/20 emission filter. GFP fluorescence was detected using 488nm excitation and 530/30nm emission. FCS 3.0 files were exported and uploaded to FlowJoV10 for analysis. Size and intensity gates for each sample were established using untransfected control H2B-mCherry cells fixed at each timepoint.

## 2.6. Microfeature Well-Plate Fabrication

Microcontact printed well plates were fabricated as previously described [35]. Briefly, gold coated glass plates were stamped with an alkanethiol ethanol initiator solution. Poly Ethylene Glycol (PEG) brushes were grown on areas stamped with the initiator solution as described previously [41]. This reaction forms distinct 600 µm diameter circular features (i.e., termed as "µFeatures") surrounded by PEG brushes that are resistant to cell adhesion. Microcontact printed glass plates were attached to bottomless 6 well plates using double sided tape. Plates were sterilized by submerging in 70% ethanol and allowed to dry for 2-4 hours in a sterile environment. Before cell seeding, patterned plates were incubated with 2 mL per well of a sterile 0.1% w/v gelatin A (Sigma) in water solution for 12-24 hours at 37°C to improve cell adhesion to µFeatures.



## 2.7. High Content Image Acquisition and Processing

Automated image acquisition was performed using a Nikon Eclipse TI epifluorescence microscope and NIS Elements Advanced Research (V4.30) software. The ND acquisition module was used to establish a XY coordinate grid to automatically acquire stitched images of each well (10× magnification, minimum of 4 field by 4 field image stitch) or  $\mu$ Feature (20× magnification, 2×2 stitch). Nikon Perfect Focus was used for experiments using glass culture substrates (Fig. 3-6) to ensure that all images were in the same Z-plane and in focus across varying XY coordinates. For experiments performed on TCPS plates (Fig. 2, 3), focus was obtained manually at the center of each well immediately prior to acquisition. Images were passed to CellProfiler for analysis (see Section 2.8). Processing was performed in a parallel manner using the Center for Throughput Computing (University of Wisconsin-Madison) and results were written to a local 16 MySQL database. MySQL Workbench 6.1 CE was used to retrieve data and join tables from distinct time points together on the basis of well position. Once joined, tables were exported as CSV files and imported into FlowJoV10 for analysis.

## 2.8. High Content Image Analysis and Image Cytometry

Image analysis was performed using CellProfiler [36] (Broad Institute). CellProfiler analysis pipelines were created for each individual experiment and pipeline files are available upon request. In general, to analyze gene-editing at the mCherry locus, cells were first stained in imaging medium with 5 $\mu$ g/mL Hoechst 33342 (Thermo Scientific) for 15 minutes at 37°C and nuclei were identified and defined as individual 2D objects within CellProfiler through a blue channel (Hoechst) image. These nuclei objects (7-20 pixels in size) were uniformly shrunk in area by 1 pixel to account for inclusion of any non-nuclei regions within the outline and to account for slight XY coordinate shifts between different channels. Shrunk nuclei objects were then super-imposed upon the red channel (mCherry) of the same image. Mean nuclear mCherry intensity was measured and tabulated for the area within each shrunk nuclei object. Nuclear GFP analysis was performed on Hoechst stained cells using similar methodology described for mCherry.

Nuclear GFP analysis on non-Hoechst stained cells could be performed up to 48 hours after transfection while mCherry expression was still observed in all nuclei. In this case, the mCherry image was enhanced to identify nuclei objects, which were shrunk by 1 pixel and superimposed onto the associated GFP channel. Cy5 labeled Cas9 mRNA localization was determined using a similar pipeline to above except that shrunk nuclei were superimposed onto the Cy5 image for intensity analysis.

Analysis of mean GFP intensity at the cell versus nuclei level for  $\mu$ Feature subpopulations was performed using a slightly modified pipeline. All individual nuclei objects within each  $\mu$ Feature were identified using enhanced mCherry images and corresponding cell objects were identified via CellProfiler's propagation algorithm, using a contrast enhanced GFP image. Once nuclei and cell outlines were determined for each image, the individual integrated GFP intensity for all nuclei and cell objects within a single  $\mu$ Feature was collected within the pipeline and used to determine the mean GFP intensity for each  $\mu$ Feature on a per cell or nuclei basis.

For all analysis, intensity per object data was exported directly to a comma spaced variable (csv) file or a csv file was created from the storage database. Data files were imported into FlowJoV10 to generate intensity histograms. Gating information was obtained using untransfected control cells from the same time point (unless otherwise noted) and subjected to the same image analysis.

## 2.9. Liposomal Delivery of RNA to Human Cells

For all experiments, HEK H2B-mCherry cells were plated 16-24 hours prior to transfection. In time course experiments (Fig. 2), HEK H2B-mCherry cells were plated at  $t = 0$  using 15,000-20,000 cells per well with 250  $\mu$ L of imaging media in gelatin A coated 48 well TCPS plates (Thermo Scientific). At  $t = 0$  hrs, cells were transfected with transfection solution prepared as follows: 500 ng Cas9-GFP mRNA and 120 ng mCherry C sgRNA were added to 12.5  $\mu$ L of OptiMEM (Thermo Scientific) and mixed prior to adding to a solution containing 12.5  $\mu$ L OptiMEM and 1  $\mu$ L of MessengerMax (Thermo Scientific). Transfection solution was incubated for 10 minutes at room temperature before adding to culture media. Media was not changed prior to transfection. At  $t = 24$  hrs an additional 250  $\mu$ L of imaging media was added to each well in culture. At  $t = 48$  hours and  $t = 96$  hours a half media change was performed (removed 250  $\mu$ L culture media, added 250  $\mu$ L fresh imaging media).

In DOE experiments (Figs. 3-4), HEK H2B-mCherry cells were plated at 7,500-10,000 cells per well with 125  $\mu$ L of imaging media in gelatin A coated 96 well glass plates (In Vitro Scientific, P96-1.5H-N). Cas9-GFP mRNA and sgRNA were delivered at specified levels in 10  $\mu$ L total of OptiMEM using specified volumes of Lipofectamine 2000, RNAiMax, MessengerMax (all Thermo Scientific), or Transit-mRNA (Mirus). Of note, Transit-mRNA is a two-component reagent and each component was added at the total volume specified in the experimental design. At  $t = 24$  hrs an additional 125  $\mu$ L of imaging media was added to each well in culture. At  $t = 48$  hours and  $t = 96$  hours a half media change was performed.

For experiments using micropatterned 6 well plates, cells were seeded at 25,000 cells per well in 2 mL of imaging media. Transfections were performed with varying amounts of Cas9-GFP mRNA and 600 ng of sgRNA with 5  $\mu$ L of MessengerMax in 250  $\mu$ L total of OptiMEM. Cas9-GFP mRNA was delivered to five wells at 1, 3, 5, 7, or 9  $\mu$ g with 600 ng of mCherry C sgRNA. To the remaining well, 600 ng of hROSA26 sgRNA was added with 5  $\mu$ g of Cas9-GFP mRNA. At  $t = 24$  hours, 2 mL of additional imaging media was added to each well. At  $t = 48$  hours a half media change was performed.

For sorting experiments cells were plated at 100,000-200,000 cells per well in a 6 well plate. Cas9-GFP mRNA was added to cells at 3 or 7  $\mu$ g with 600 ng of mCherry C or VEGF sgRNA and 5  $\mu$ L of MessengerMax in 250  $\mu$ L total of OptiMEM. The lipid/RNA formulation was added to cells in 2 mL total of imaging media. Cells were sorted at 24 hours after transfection and plated at 10,000 cells per well in a 24 well TCPS plate with 3-4 replicates per GFP intensity level (dependent on total sorted cell count). At 96 hours after transfection, cells were Hoechst stained, imaged, then singularized and fixed with 4% PFA for 15 min in preparation for flow cytometry.



### 3. RESULTS

#### 3.1. Image cytometry enables faithful, dynamic observation of population level gene-editing outcomes

**3.1.1. HEK H2B-mCherry reporter cells facilitate tracking of Cas9-sgRNA mediated editing**—A cell reporter based system to track single cell gene-editing outcomes is illustrated in Figure 1. CRISPR-Cas9 mediated gene-edits can be quantified by tracking loss of nuclear mCherry expression in transgenic HEK 293T cells that constitutively express a histone 2B-mCherry fusion protein. Two components of the CRISPR-Cas9 system (sgRNA and Cas9-GFP mRNA) were delivered to the cells using cationic lipid based materials (Fig. 1A). After transit across the cell membrane, Cas9-GFP mRNA is translated by ribosomes in the cytoplasm into a functional Cas9 GFP fusion protein. The Cas9-GFP fusion protein complexes with sgRNA enabling recognition (Fig. 1B) of a specific 19-20 bp genomic DNA sequence specified by the sgRNA. Cas9 DNA endonuclease activity forms DSBs in the genomic DNA (Fig. 1C) at a location 3 bases upstream of the 3' end of the sgRNA target sequence. Non-homologous end joining (NHEJ) or other error prone DNA repair (Fig. 1D) leads to insertion or deletion mutations (gene-edits) in the coding region of the targeted loci. Edits at the mCherry reporter locus can result in frameshift, nonsense or missense mutations that lead primarily to loss of mCherry expression (Fig. 1F), whereas unedited DNA retains mCherry reporter expression (Fig. 1E).

**3.1.2. *In vitro* transcription (IVT) and purification of Cas9-GFP mRNA and mCherry sgRNA**—Representative 2100 Bioanalyzer (Agilent Technologies) electropherograms of transcribed and purified Cas9-GFP RNA are shown in Figure 1G. The size and purity of the Cas9-GFP mRNA (~4.3 Kb) are similar to that of commercially prepared non-GFP Cas9 (~4.0 Kb) (Thermo Scientific, A25640). Peaks at 25 nt are from standard size markers in the Bioanalyzer running buffer. LiCl precipitation was necessary to remove template DNA and excess ribonucleotides from the reaction prior to concentration determination and delivery.

Electropherograms of purified IVT mCherry C sgRNA are displayed in Figure 1H. mCherry C sgRNA was selected based on initial screening of three sgRNAs targeting the mCherry locus as described in Section 2.2. Size (~100 nt) and purity of IVT sgRNA conformed to predictions based on the size of the template and were consistent with prior work [34]. Column purification of sgRNAs was necessary to remove unincorporated ribonucleotides prior to determining concentration.

**3.1.3. Timecourse analysis of delivery and gene-editing outcomes with standard and image cytometry based methods**—Standard techniques for quantifying population level gene-editing outcomes include T7 endonuclease 1 (T7E1) digestion and flow cytometry. T7E1 analysis is time consuming and requires cell lysis and isolation of genomic DNA. This is followed by PCR amplification of the targeted loci, denaturing and rehybridizing amplified DNA, and subsequent digestion with the T7E1 enzyme to cleave edited DNA for size separation analysis. Flow cytometry can also be used to detect population level gene-editing outcomes. Editing can be tracked in live cells, if the

locus of interest has a reporter phenotype or the outcome can be assessed with a live cell stain. Editing can also be quantified in fixed and immunostained cells, using flow cytometry, if the edit results in loss of a functional protein or if an antibody is specific to mutation at a specific domain. We hypothesized that quantitative image cytometry (method described in section 2.8) could be used to faithfully track gene-editing outcomes, in H2B-mCherry reporter cells, comparable to T7E1 and flow cytometry methods and at increased workflow efficiency.

To evaluate this capability, we set up a gene-editing timecourse to track gene-editing outcomes for 120 hours after delivery of Cas9-GFPmRNA/sgRNA/material complexes. We quantified editing at the mCherry locus using both standard and image cytometry methods (Fig. 2A). This assay was performed in 48 well standard tissue culture polystyrene plates (TCPS) to obtain sufficient numbers of cells for flow analysis and sufficient genomic DNA for the T7E1 assay. For each method, gene-editing was quantified at specific time points from 3 physical replicate wells. The same amounts of Cas9-GFP mRNA (500 ng), mCherry sgRNA (120 ng) and material (MessengerMax, 1.0  $\mu$ L) were used across all wells. Editing, measured as percent DNA modification, was first detected using T7E1 assay beginning at 24 hours and increased up to a maximum of 47% at 48 hours (Fig. 2B, C). Flow cytometry analysis of per cell GFP intensity (to indicate Cas9 expression) and mCherry intensity (to indicate gene-editing) (Fig. 2D-F) was performed on cell populations at each time interval. From the GFP intensity histograms (Fig. 2D) the maximum percentage of GFP positive cells peaked at 24 hours (33%) and decreased rapidly after 48 hours. It was evident from GFP intensity histograms (Fig. 2D) that Cas9-GFP expression varied across each well population and could not be separated into distinct GFP positive and GFP negative populations.

Flow cytometry analysis of mCherry expression indicated that 67% of cells lost expression of mCherry between 72 – 96 hours. Maximal loss of mCherry expression measured with flow cytometry was observed 24-48 hours after maximal editing was seen in T7E1 assay. This is expected, as after mutation of the genomic loci, existing H2B-mCherry protein must be slowly degraded (*in vivo* half-life of mCherry ~24 hours [42]) or diluted by cell division before it is no longer detectable within the nucleus. Analysis of gene-editing with flow cytometry reported higher maximum editing efficiency ( $67 \pm 6.3\%$ ) compared to identically prepared T7E1 samples ( $47 \pm 0.2\%$ ). This result may reflect the poor efficiency at which T7E1 detects single base indels [32]. There may also be a single indel that is highly abundant within the population leading to perfect self-annealing in the hybridization step of the T7E1 assay.

Image cytometry analysis was performed on images taken of cells in TCPS culture wells (Fig. 2G) before those same wells were prepared for flow cytometry analysis. Mean red channel (mCherry) intensity of each individual nucleus within the image provides single cell data comparable to that obtained from flow cytometry (Fig. 2 H, I). Maximum mCherry loss was again observed between 72-96 hours ( $78 \pm 10\%$ ) (Fig. 2H) and no significant difference was observed between gene-editing outcomes detected with image cytometry versus flow cytometry (Appendix, Fig. B.1 A, B). Image cytometry histograms from the untransfected control samples (Fig. 2H) display a general increase in red intensity over time (observed as a shift of the average intensity of the mCherry positive population towards the right). This

could be due to overlapping of cells as the cell density within the culture well increases, leading to contribution of fluorescent signal from adjacent cells. Importantly, GFP intensity data could not be reliably quantified on the TCPS plates used in this experiment due to high background fluorescence at well edges in the GFP channel. Subsequent image studies were performed on glass bottom plates and/or plates with black well walls, which addressed image background illumination inconsistencies.

While GFP image analysis was limited in 48 well TCPS plates, we were able to employ image cytometry to track delivery of Cy5 labeled Cas9 mRNA to cells (Figure 3). Labeled Cas9-GFP was labeled and purified following IVT. A total of 500ng labeled mRNA was delivered to cells along with mCherry sgRNA (120 ng) and material (MessengerMax, 1.0  $\mu$ L). Cells were seeded 24 hours prior to transfection and each well was imaged over 12 hours at 1 hour intervals (Fig. 3A). Image analysis data was passed to FloJo to create intensity histograms (Fig 3A, right panel). Analysis of n=3 physical replicates determined that ~90% of cells contained detectable Cy5 labeled mRNA (Fig. 3B). This analysis indicates that a majority of cells receive Cas9-GFP mRNA by 12 h after transfection, while only a third have detectable levels of Cas9-GFP protein later at peak expression at 24 h.

These experiments illustrate a potential advantage of image cytometry where *in situ* analysis alleviates the need for singularization and washing steps, which increases throughput and yields greater numbers of cells for analysis (greater total cell counts in Fig. 2H versus Fig. 2D) at early timepoints ( $t = 0 - 24$  hours) with low cell density.

### 3.2. DOE based screening of material and delivery parameters for primary effectors of gene-editing efficiency

Based on comparable analysis of gene-editing using image cytometry versus standard methods (Fig. 2), we sought to focus our image acquisition and processing platform to perform high throughput optimization of cationic lipid mediated delivery of CRISPR-Cas9 RNA to human cells. Our focus in these experiments was to maximize loss of mCherry expression as a proxy of gene-editing efficiency. Even with a single output, the number of factors affecting gene-editing and the range of possible levels for each of these factors is vast. Importantly, many of these factors are hypothesized to interact so comparison of delivery materials, for example, must be performed across different concentrations of the material and ratios of the Cas9 and sgRNA cargo. We hypothesized that our image cytometry method in combination with DOE analysis could be used to rapidly and efficiently screen for the most significant effectors of gene-editing efficiency and to quantify their leverage on the output. DOE is regularly applied in engineering to optimize complex processes and has been applied previously in biological systems to optimize cell culture conditions [37,38]. Use of DOE in this application is particularly prescient as identification of factors that primarily influence gene-editing efficiency can inform future experiments that are designed to optimize only the most important inputs over a much greater range of levels.

An initial DOE analysis was performed to screen four factors (delivery material, material amount, Cas9 mRNA amount, and sgRNA amount) for their effect on gene-editing efficiency. Three commercially available liposomal delivery materials (Lipofectamine 2000, RNAiMax, and MessengerMax) with prior use in Cas9 delivery [16] were screened, while

material amount, Cas9 mRNA amount and sgRNA amount were delivered at two levels in a full factorial design ( $3 \times 2^3 = 24$  conditions). In order to increase throughput and minimize background signal, DOE analysis was performed in 96 well glass bottom plates. Levels for each factor (Fig. 4A) were selected based on levels used in prior 48 well format time course experiments (Fig. 2) and were volume adjusted to take into account the lower culture volumes used in the 96 well plate. Utilizing two levels for material, Cas9 mRNA and sgRNA amount enables replicates and lowers setup complexity, but requires assumption of a linear relationship between output and changing factor levels. To evaluate if a linear assumption is appropriate, a center point condition for each material was added to the design ("0", Fig. 4A). Final full factorial design with 27 conditions was conducted with two additional replicates for a total of  $n=3$  replicates (Full design is available in Table C.1).

Analysis of gene-editing outcomes at the mCherry locus after 120 hours of culture resulted in editing efficiency between 0-10% (Fig 4B). Scaling down the assay to 96 well glass bottom plates increased the throughput of the screen, but lowered the amount of editing compared to timecourse experiments conducted in 48 well format (Fig. 2). Lower editing efficiency in 96 well plates may be due to lower overall cell attachment on glass versus TCPS culture surfaces. Maximum gene-editing efficiency was observed using the MessengerMax delivery material at high levels (Fig. 3B). Outputs at center point conditions do not lie outside the output range of high and low conditions indicating that the linearity assumption is not violated across the selected level range (center point conditions are indicated by "0" levels, Fig. 4B). Ignoring factor interactions, a least squares linear regression model identified delivery material and material amount as significant factors (Fig. 4C). According to the regression model, mRNA and sgRNA amount were not significant factors across the range of levels studied (Fig. 4C). Parameter estimates for the regression model indicate use of MessengerMax, as the delivery material, has the greatest effect on gene-editing output, followed by use of higher amounts of delivery material (Fig. C.1 B). When accounting for interactions between factors, the least squares linear regression model identified material type\*material amount and Cas9-mRNA amount\*sgRNA amount as additional significant factors (Fig. C.1 C). Identification of a significant interaction between the amount of Cas9-mRNA and sgRNA suggests that while neither factor appears to significantly effect gene-editing outcome in isolation, there may be a specific Cas9 mRNA to sgRNA ratio that optimizes gene-editing efficiency.

Prior work with Cas9 protein delivery indicates gene-editing efficiency is dependent on amount of Cas9 and sgRNA delivered [16] so it was unexpected that the amount of Cas9 mRNA or sgRNA delivered had no significant effect on gene-editing outcome in our model. We reasoned that the low levels ("–", in Fig. 4A) of Cas9 mRNA, as well as sgRNA, used in our initial DOE were not limiting to overall gene-editing outcome. To test this hypothesis, a second DOE analysis (Fig. 5) was designed to examine gene-editing at lower levels of Cas9 mRNA, sgRNA and liposomal material. Although increased material amounts in DOE 1 yielded increased gene-editing, we chose to lower the amount of liposomal delivery material to capture its delivery performance over a wider range. Additionally, we screened another material (Transit-mRNA) along with the leading material identified in the previous DOE analysis (MessengerMax). Further, we simultaneously evaluated gene-editing outcomes with Cas9-GFP mRNA prepared in our lab to a commercially available Cas9 mRNA. For

continuous variables (mRNA, sgRNA, material amount), high factor levels (“+” in Fig. 5A) were set similar to low levels specified in the first DOE analysis, and low levels were selected to cover levels used in recently published work with Cas9 mRNA delivery [17]. The final DOE design included 5 factors each at 2 levels ( $2^5 = 32$  conditions, with  $n=2$  replicates) and is available in Table D.1.

In the second DOE experiment, MessengerMax performed superiorly to Transit-mRNA across all delivery conditions (Fig. 5B). Again, higher levels of delivery material appeared to increase gene-editing efficiency. However maximum gene-editing efficiencies using MessengerMax in this DOE were higher (10% versus 8%) than gene-editing efficiencies in DOE 1, despite use of lower amounts of material. Also of interest, was that Cas9-GFP mRNA had similar gene-editing efficiency compared to commercial Cas9 using MessengerMax as the delivery agent, whereas commercial Cas9 resulted in higher editing efficiency than Cas9-GFP when Transit-mRNA was used as the delivery agent (Fig. 5B). This data supports a potential interaction between the type of mRNA cargo and the delivery material. Again, ignoring factor interactions, delivery material and material amount were the only significant factors identified (Fig. 5C). Parameter estimates for the regression model indicate that material choice is the most important factor influencing gene-editing outcome (Fig. D.1B). When accounting for interactions between factors, material\*mRNA type was identified as a significant factor (Fig. D.1C). This supports earlier observations (Fig. 5B) of differential editing efficiency with mRNA type based on delivery material used. Again, this model did not indicate that Cas9 mRNA and sgRNA amount had a significant effect on gene-editing outcomes (Fig. D.1D), even when delivered at concentrations up to 3-5 fold below levels reported in prior studies [17].

### 3.3. Longitudinal high content analysis of gene-editing within subpopulations

We considered the possibility that variable Cas9 mRNA trafficking and processing within cells in an individual well of the previous DOE experiments (Figs. 5-6) led to non-significant differences in gene-editing efficiency at varying Cas9-mRNA delivery levels. Single cell analysis using flow and image cytometry indicated significant variation in nuclear Cas9-GFP intensity across gene-edited cell populations in 48 well (Fig. 2D) and 96 well plate format (data not shown).

To control for variable Cas9-GFP expression within each delivery condition we sought to track editing outcomes in subpopulations of cells with known Cas9-GFP expression. To this end, we employed microcontact printed cell culture surfaces to spatially separate ~450 subpopulations of cells into circular microfeatures ( $\mu$ Features, 600  $\mu$ m diameter) within each well of a 6 well plate. The polyethylene-glycol (PEG)-based surface chemistry of the platform spatially separates distinct populations of cells by preventing cell adhesion outside of the  $\mu$ Features [35] and enables tracking of distinct subpopulations over several days of culture. Furthermore, these features are printed on 0.2 mm thickness optically clear glass culture substrates, which enable high-resolution image acquisition for tracking subcellular fluorescence expression *in situ*. One day prior to CRISPR-Cas9 delivery, cells were seeded onto the microcontact printed plates at overall well density of 1,500 - 2,500 cells per  $\text{cm}^2$ . After one day of culture, Cas9-GFP mRNA at concentrations ranging from 0.3 - 3 nM (500 -

4500 ng/mL) was added to the culture media along with sgRNAs targeting either mCherry or human ROSA26 (hROSA26) loci [43].

High content imaging prior to transfection was utilized to obtain an initial cell number for each population ( $t = 0$  h; representative images in Fig. 6A). At 24 h post transfection, wells were imaged to determine both cellular and nuclear Cas9-GFP intensity and cell number ( $t = 24$  h; representative images in Fig. 6A). Mean intensity of cellular and nuclear Cas9-GFP was calculated on a per area basis for each cell/nuclei and averaged over all cells/nuclei within each  $\mu$ Feature to capture subpopulation specific data. The average of subpopulation Cas9-GFP expression within each well generally increased between low and high concentration mRNA formulations however, as before; expression was highly variable within individual wells (Fig. 6B).

### 3.4. Predicting gene-editing outcomes from longitudinal analysis

We first performed univariate analysis - plotting single factors against one another - in order to determine if any single factors had a strong correlation with GFP intensity (at  $t = 24$  h) or with % gene-editing ( $t = 96$  h). It is important to note that these comparisons are subject to confounding factors from related variables not included in the analysis. Quantification of mean nuclear and mean cell Cas9-GFP expression at  $t = 24$  h indicated that the two parameters were indistinguishable at the subpopulation level (Fig. E.1A). Nuclear Cas9-GFP expression at  $t = 24$  h was weakly negatively correlated with the initial number of cells ( $t = 0$  h) present within each  $\mu$ Feature ( $R^2 = 0.30$ ) (Fig. E.1B). We also determined the change in cell number over the 24 h after transfection and found that slower dividing subpopulations were associated with higher mean Cas9-GFP expression (Fig. 6C;  $R^2 = 0.37$ ).

Upon an additional three days of culture ( $t = 96$  h), nuclei were stained using a cell permeable Hoechst dye and imaged along with the mCherry channel in order to determine the percentage of cells within each  $\mu$ Feature that were mCherry negative (representative images in Fig. 6A). Association of % mCherry negative cells within each  $\mu$ Feature to original mRNA dose (Fig. 6D) indicated highly variable gene-editing outcomes amongst subpopulations within the same well. This result supports the importance of monitoring gene-editing outcomes at the subpopulation level. Furthermore, variable gene-editing outcomes within each well align closely to variation in Cas9-GFP expression observed within each well at  $t = 24$  h (Fig. 6B). Univariate analysis of cell growth from  $t = 24$  to  $t = 96$  hours versus editing indicates a weak positive correlation between cell growth and editing percentage (Fig. E.1C). Preliminary analysis of mean cellular Cas9-GFP intensity at  $t = 24$  h against loss of mCherry expression at  $t = 96$  h indicated strong positive correlation (Fig. 6C,  $R^2 = 0.6623$ ). This suggests that high mean cellular Cas9-GFP expression may be a useful independent predictor of downstream editing outcomes.

To more stringently test the relationship between Cas9 expression and gene-editing, we conducted multivariate analysis, including principal components analysis (PCA) and partial least squares regression (PLSR) to account for closely related variables (Figure 7). PCA reduces the dimensionality of data, while maximally preserving the original information, enabling interpretation of complex relationships. When individual data points for each subpopulation were superimposed on the PCA score plot, the subpopulations with highest



resulting gene-editing efficiency at  $t = 96$  hours (high mCherry loss, blue dots in Fig. 7A) clustered near higher mean nuclear and cellular Cas9-GFP intensity measured at  $t = 24$  hours. Conversely, subpopulations with low levels of gene-editing (low mCherry loss, red dots in Fig. 7A) clustered around high measures of cell number.

To quantitatively compare the individual contribution of all measured factors on gene-editing efficiency, we conducted a PLSR analysis. PLSR is useful for quantifying relative effects of independent factors when they are highly correlated. PLSR was implemented in JMP Pro and yielded the model depicted in Figure 7B. A model containing three factors was selected based on cross validation (Fig. F.1A). This model explains 70% of the variation in % mCherry gene-editing (Fig. F.1B; percent variation in Y). Model coefficients for each factor confirmed that mean nuclei and cell GFP intensity at  $t = 24$  hrs had the strongest weights in predicting gene-editing (Fig. 7C). This analysis also confirmed that more rapid cell division in the 24 hour period after transfection, as well as from 24 hours to 96 hours, is negatively correlated with gene-editing outcomes, as indicated by the negative values for the Delta(cells) coefficients (Fig. 7C). When accounting for other factors using PSLR, a greater number of cells at  $t = 0$  h was found to have a small positive contribution to gene-editing efficiency, contradicting the earlier univariate correlation observation (Fig. E.1B). Greater cell growth from  $t = 24$  to  $t = 96$  h had a smaller negative weight, which contradicts the univariate correlation observation made in Fig. E.1C. Taking the PCA and PLSR results together, longitudinal analysis within small subpopulations was able to identify features of delivery that may be useful for predicting gene-editing efficiency.

To reinforce our conclusions from PCA and PLSR analysis, we conducted a cell sorting experiment diagramed in Fig. 8A. Cells were transfected at  $t = 0$  with lipid/RNA formulations containing 1 or 2 nM Cas9-GFP mRNA and either mCherry or VEGFA sgRNA. At  $t = 24$  hours cells were sorted on the basis of GFP expression into low, medium and high bins (Fig. G.1) and re-plated at 10,000 cells per well in 24 well plate format. After 3 days of culture ( $t = 96$  h) cells were analyzed for gene-editing and cell count. Cell count at  $t = 96$  h, a measure of cell growth between  $t = 24 - 96$  h, was not significantly different between GFP expression bins (Fig. 8B). Percent gene-editing was significantly higher in groups with greater GFP intensity at  $t = 24$  hrs, as determined by image and flow cytometry (Fig 8C). Editing reached a maximum of >90% mCherry negative in the high GFP group (Fig. 8C). Furthermore, VEGFA sgRNA treated cells sorted into low and high GFP groups also demonstrated higher % of gene modification in the high GFP group (Fig. 8D) These data support that relative GFP intensity 24 hours after transfection is reliable predictor of gene editing efficiency using lipid/CRISPR-Cas9 RNA formulations.

## 4. STATISTICS

### 4.1. Confidence Intervals and Paired t-Tests

Unless stated otherwise, all error bars correspond to 95% confidence intervals of the mean determined from at least  $n=3$  samples. Confidence intervals were calculated using the CONFIDENCE.NORM function in Excel (Microsoft) with  $\alpha = 0.05$ . Single tailed paired t-tests were also conducted on a minimum of  $n=3$  samples using the T.Test function in Excel.

## 4.2. ANOVA

A two-way ANOVA to compare image and flow cytometry methods was implemented using GraphPad Prism. P values < 0.05 were considered significant.

## 4.3. Design of Experiments (DOE)

Full Factorial DOE designs with added center points were designed using JMP Pro 11 software. Regression model fits were also performed using JMP Pro 11.

## 4.4. Principal Components Analysis (PCA) and Partial Least Squares Regression (PLSR)

PCA and PLSR were conducted using JMP Pro 11 software. PCA was conducted on correlations using the default estimation method. PLSR was performed using the SIMPLS algorithm. Validation, to determine the proper number of factors in the PLSR model, was performed using the leave-one-out method.

# 5. DISCUSSION

This work examined the delivery of two RNA-encoded components of the Cas9 gene-editing system with four lipid based biomaterials. Image based longitudinal tracking of many subpopulations enabled quantitative *in situ* analysis of key processes involved in non-viral RNA mediated gene editing, including: protein expression, subcellular localization and functional gene disruption. Nearly all reported work optimizing CRISPR-Cas9 system delivery has monitored these processes using endpoint assays with population-averaged measurements. This work provides additional insight with in-parallel collection of multiparametric data from live cells at single cell to subcellular level resolution.

Comparing our image-based method to standard methods of quantifying gene-editing confirmed the reliability of image based cytometry in tracking gene-editing outcomes (Fig. 2). In comparison to flow cytometry, *in situ* analysis enabled by image based methods allowed for reliable quantification of gene-editing in subpopulations with low cell counts not suitable for flow cytometry (Fig. 5,6). Furthermore, tracking gene-editing in cell subpopulations over time with high content imaging enabled reliable tracking of cell number as well as subcellular localization of Cas9-GFP expression.

For this study we focused on DOE based screening of several factors– mRNA type, mRNA amount, sgRNA amount, lipid amount and lipid type –and determined that material identity and material concentration had the greatest effect on gene-editing efficiency (Fig. 3,4). Interestingly, despite examining molar concentrations of sgRNA and Cas9 mRNA over a 3-5 fold change above and below those reported previously [17], we did not identify these factors as significant effectors of editing efficiency (Fig 3,4). Prior work with lipid-based delivery of preformed Cas9-sgRNA RNPs indicated that editing outcomes were dependent on the concentration of Cas9 and sgRNA delivered [16]. Conversely, our experiments with mRNA indicate that, over the concentrations used, delivery of RNAs encoding the CRISPR-Cas9 system is primarily limited by material and/or cell mediated factors rather than amounts of RNA added to the lipid formulation.

The observation that >90% of cells have detectable levels of labeled Cas9 mRNA 12 h after transfection (Fig. 3B), while in a separate experiment, using the same material and RNA concentrations, only 30-35% of cells achieve detectable Cas9-GFP expression (Fig. 2E) suggests that mRNA trafficking or degradation within the endosome may be a major limiting factor governing expression of CRISPR-Cas9 components. Utilizing the imaging approach described here with labeled mRNA and compatible live cell endosome dyes could facilitate high throughput screening of material libraries to identify compounds that improve endosomal escape and prevent mRNA degradation.

Another important observation from our DOE analysis was a significant interaction between material type and mRNA type (Fig. 4B). Lower editing outcomes with Cas9-GFP mRNA versus commercial Cas9 mRNA only when using Transit-mRNA, highlight the need to consider cargo-material interactions during delivery screening. Size and charge topology characteristics of the cargo have been shown to impact complex formation, influence endosomal escape and impact cellular trafficking [44,45]. Further examination of physical or electrostatic interactions of CRISPR-Cas9 components with non-viral delivery materials with respect to both RNA delivery and subsequent gene editing efficiency is warranted and could be rapidly evaluated with DOE analysis using the imaging method described here.

Further, high content analysis of cell subpopulations using  $\mu$ Feature plates has distinct advantages that identified additional factors impacting RNA delivery and gene-editing efficiency. Specifically, our HCA platform enabled collection of subcellular localization and tracking of delivery and editing outcomes in distinct subpopulations of cells over several days in culture. We did not detect a difference between total cellular Cas9-GFP or nuclear Cas9-GFP and gene editing outcome at the subpopulation level (Fig. 7A,C) indicating that nuclear import is not a rate limiting step for these formulations. Another key result was that greater changes in cell number 24 hours after CRISPR-Cas9 delivery were associated with lower gene-editing efficiency (Fig. 7C). Higher cell proliferation may dilute the effects of any delivered RNA in daughter cells. Further,  $\mu$ Features with lower numbers of cells may have more quiescent cells. DNA repair is regulated by the cell cycle, and prior work in synchronizing cells resulted in higher gene-editing [46]. Further work is warranted, but our results indicate that optimization of lipid based transfection approaches should not merely lower cellular toxicity or increase endosomal escape, but rather concomitantly effects on the cell cycle or proliferation are likely important.

Our work exploited mRNA encoding a Cas9-GFP fusion protein and demonstrated that GFP fluorescence may be a useful tool for gene-editing optimization, because of the strong correlation between early Cas9-GFP expression and editing in both transgenes (Fig. 6) and endogenous loci (Fig. 8D). This relationship is corroborated by prior work with Cas9-2A-GFP plasmid delivery [47] and is a useful tool in cases where biomaterials need to be optimized for editing a specific locus or a cell line that does not contain a fluorescent reporter. Gene-editing seemed variable in our assays, ranging from 5-80% with identical cells, RNA and materials. These results have been observed with electroporation [17]. While there could be methods to reduce variability such as synchronizing cell cycle, other methods of delivery – such as nanoneedles [48] and mechanical squeezing [49] – may have less variability for applications of CRISPR-Cas9 where noise is an issue.

The potential design space for materials to improve gene-editing outcomes is vast, as is the need for systems to rapidly optimize new formulations. In terms of targets, gene-editing of fibroblasts, iPSCs, and CD34+ cells, all notable clinically relevant cell types, could be optimized using this platform. Many permutations of delivery cargo can also be screened. For example, chemical modifications to sgRNAs have been shown to increase the stability of sgRNA, resulting in increased editing [50]. mRNA can also be chemically modified to increase stability and reduce immunogenicity *in vivo* [51]. Further, there will likely be a need to increase the delivery of DNA repair templates [52,53] for precise insertion of therapeutic sequences at the Cas9 cut site. Given the flexible nature of the platform (i.e. capability of monitoring a wide range of morphological or fluorescent based parameters), we anticipate the ability to improve gene-editing efficacy by aiding in the development and optimization of novel materials-mediated gene-editing strategies.

## 6. CONCLUSIONS

Optimizing delivery of all components of the CRISPR-Cas9 system to cells is a current challenge in several areas of biomedical research. This work used a new screening platform to systematically evaluate gene-editing outcomes with four different lipid transfection reagents over a range of RNAs encoding the Cas9 system. The screening platform exploits microcontact printing to generate isolated cell-adhesive areas on a transparent cell culture surface. Image based analysis of Cas9 expression and editing of a fluorescent transgene reporter gene within human embryonic kidney cells was monitored over several days after transfection. We found editing efficiency to be highly dependent on transfection material identity. Amount of mRNA and sgRNA, as well as nuclear import of Cas9 protein did not affect gene-editing in this study. Additionally, we found Cas9 mRNA was delivered to a high percentage of cells, while Cas9-GFP expression and subsequent gene editing was detected in a fraction of the population. This indicates that intracellular trafficking and translation of mRNA may be a major limiting factor in gene editing efficiency using lipid based CRISPR-Cas9 RNA delivery strategies.

Higher Cas9 expression levels and lower cell growth within 24 hours after transfection resulted in higher gene-editing. Taken together, these results suggest that new transfection strategies with lipid based materials should seek to engineer delivery strategies that can rapidly deliver mRNA to translational machinery, while simultaneously influencing cell cycle and growth parameters that optimize editing. This high content imaging based approach to simultaneously monitor non-viral delivery parameters and gene-editing outcomes will likely be useful in optimizing non-viral delivery and editing in a variety of human cells, ultimately advancing and informing genomic medicine, regenerative biology and drug discovery.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

We acknowledge generous financial support from the Wisconsin Institute for Discovery (K.S.), NSF (CBET-1350178; K.S.), the Society in Science Foundation (K.S.), NIH (5T32HG002760 and T32GM008692-16; B.S.), the VitreoRetinal Surgery Foundation (B.S.) and the UW Stem Cell and Regenerative Medicine Center (J.C.S.). We thank John Zuris and David Liu for sharing Cas9 plasmids and plasmid depositors to Addgene. Additionally, we thank the UW Carbone Cancer Center Flow Cytometry Laboratory and the Thomson lab for their aid with flow sorting and cytometry experiments.

## REFERENCES

- [1]. Doudna JA. Genomic engineering and the future of medicine. *JAMA*. 2015; 313:791–792. doi: 10.1001/jama.2015.287. [PubMed: 25710652]
- [2]. Lander ES. Brave New Genome. *N. Engl. J. Med.* 2015; 373:5–8. doi:10.1056/NEJMp1506446. [PubMed: 26039524]
- [3]. Zhang F. CRISPR/Cas9: Prospects and Challenges. *Hum. Gene Ther.* 2015; 26:409–410. doi: 10.1089/hum.2015.29002.fzh. [PubMed: 26176430]
- [4]. Cox DBT, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* 2015; 21:121–131. doi:10.1038/nm.3793. [PubMed: 25654603]
- [5]. Baltimore D, Berg P, Botchan M, Carroll D, Charo RA, Church G, et al. A prudent path forward for genomic engineering and germline gene modification. *Science*. 2015; 348:36–38. doi: 10.1126/science.aab1028. [PubMed: 25791083]
- [6]. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 2013; 31:397–405. doi:10.1016/j.tibtech.2013.04.004. [PubMed: 23664777]
- [7]. Charpentier E, Doudna JA. Biotechnology: Rewriting a genome. *Nature*. 2013; 495:50–51. doi: 10.1038/495050a. [PubMed: 23467164]
- [8]. Carroll D. Genome Engineering with Targetable Nucleases. *Annu. Rev. Biochem.* 2014; 83:409–439. doi:10.1146/annurev-biochem-060713-035418. [PubMed: 24606144]
- [9]. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339:823–826. doi:10.1126/science.1232033. [PubMed: 23287722]
- [10]. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 2013; 339:819–823. doi:10.1126/science.1231143. [PubMed: 23287718]
- [11]. Jinek M, East Alexandra, Cheng Aaron, Lin Steven, Ma Enbo, Doudna Jennifer. RNA-programmed genome editing in human cells. *ELife Sci.* 2013; 2 doi:10.7554/eLife.00471.
- [12]. O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*. 2014; 516:263–266. doi:10.1038/nature13769. [PubMed: 25274302]
- [13]. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014; 507:62–67. doi:10.1038/nature13011. [PubMed: 24476820]
- [14]. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 2013; 31:822–826. doi:10.1038/nbt.2623. [PubMed: 23792628]
- [15]. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 2015; 33:187–197. doi:10.1038/nbt.3117. [PubMed: 25513782]
- [16]. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 2015; 33:73–80. doi:10.1038/nbt.3081. [PubMed: 25357182]
- [17]. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotechnol.* 2015; 208:44–53. doi: 10.1016/j.jbiotec.2015.04.024. [PubMed: 26003884]

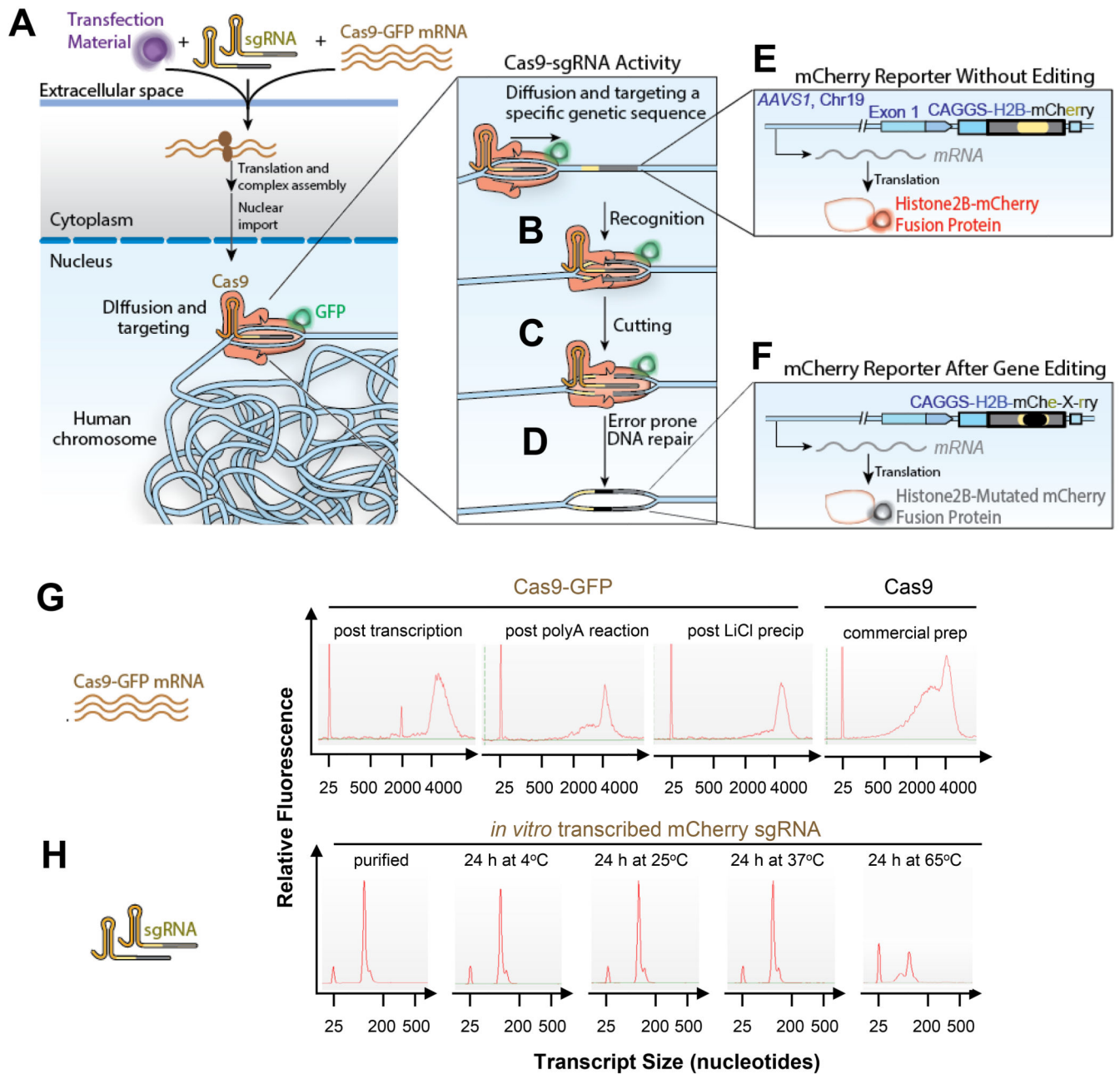
- [18]. Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A. Optical Control of CRISPR/Cas9 Gene Editing. *J. Am. Chem. Soc.* 2015; 137:5642–5645. doi:10.1021/ja512664v. [PubMed: 25905628]
- [19]. Nihongaki Y, Kawano F, Nakajima T, Sato M. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 2015; 33:755–760. doi:10.1038/nbt.3245. [PubMed: 26076431]
- [20]. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* 2015; 523:481–485. doi:10.1038/nature14592. [PubMed: 26098369]
- [21]. Kabadi AM, Ousterout DG, Hilton IB, Gersbach CA. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* 2014; 42:e147. doi:10.1093/nar/gku749. [PubMed: 25122746]
- [22]. Yin H, Kanasty RL, Eltoukhy AA, Vegas AJ, Dorkin JR, Anderson DG. Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* 2014; 15:541–555. doi:10.1038/nrg3763. [PubMed: 25022906]
- [23]. Li L, He Z-Y, Wei X-W, Gao G-P, Wei Y-Q. Challenges in CRISPR/CAS9 Delivery: Potential Roles of Nonviral Vectors. *Hum. Gene Ther.* 2015; 26:452–462. doi:10.1089/hum.2015.069. [PubMed: 26176432]
- [24]. Gori JL, Hsu PD, Maeder ML, Shen S, Welstead GG, Bumcrot D. Delivery and Specificity of CRISPR/Cas9 Genome Editing Technologies for Human Gene Therapy. *Hum. Gene Ther.* 2015; 26:443–451. doi:10.1089/hum.2015.074. [PubMed: 26068008]
- [25]. Xiang S, Tong H, Shi Q, Fernandes JC, Jin T, Dai K, et al. Uptake mechanisms of non-viral gene delivery. *J. Controlled Release.* 2012; 158:371–378. doi:10.1016/j.jconrel.2011.09.093.
- [26]. Green JJ, Zhou BY, Mitalipova MM, Beard C, Langer R, Jaenisch R, et al. Nanoparticles for Gene Transfer to Human Embryonic Stem Cell Colonies. *Nano Lett.* 2008; 8:3126–3130. doi: 10.1021/nl8012665. [PubMed: 18754690]
- [27]. Dahlman JE, Barnes C, Khan OF, Thiriot A, Jhunjhunwala S, Shaw TE, et al. In vivo endothelial siRNA delivery using polymeric nanoparticles with low molecular weight. *Nat. Nanotechnol.* 2014; 9:648–655. doi:10.1038/nnano.2014.84. [PubMed: 24813696]
- [28]. Sahay G, Querbes W, Alabi C, Eltoukhy A, Sarkar S, Zurenko C, et al. Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling. *Nat. Biotechnol.* 2013; 31:653–658. doi:10.1038/nbt.2614. [PubMed: 23792629]
- [29]. Mahiny AJ, Dewerth A, Mays LE, Alkhaled M, Mothes B, Malaeksefat E, et al. In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency. *Nat. Biotechnol.* 2015; 33:584–586. doi:10.1038/nbt.3241. [PubMed: 25985262]
- [30]. Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor nuclease. *BioTechniques.* 2004; 36:702–707. [PubMed: 15088388]
- [31]. Kim HJ, Lee HJ, Kim H, Cho SW, Kim J-S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* 2009; 19:1279–1288. doi: 10.1101/gr.089417.108. [PubMed: 19470664]
- [32]. Vouillot L, Th  lie A, Pollet N. Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases. *G3 GenesGenomesGenetics.* 2015; 5:407–415. doi:10.1534/g3.114.015834.
- [33]. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 2014; 42:e168. doi:10.1093/nar/gku936. [PubMed: 25300484]
- [34]. Carlson-Stevermer J, Goedland M, Steyer B, Prestil R, Angenent-Mari N, Saha K. High content analysis of gene-edited human embryonic stem cells using CRISPR-Cas9. *Stem Cell Rep.* (in revision).
- [35]. Harkness T, McNulty JD, Prestil R, Seymour SK, Klann T, Murrell M, et al. High-content imaging with micropatterned multiwell plates reveals influence of cell geometry and cytoskeleton on chromatin dynamics. *Biotechnol. J.* 2015 doi:10.1002/biot.201400756.
- [36]. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006; 7:R100. doi:10.1186/gb-2006-7-10-r100. [PubMed: 17076895]



- [37]. Marinho PA, Chailangkarn T, Muotri AR. Systematic optimization of human pluripotent stem cells media using Design of Experiments. *Sci. Rep.* 2015; 5 doi:10.1038/srep09834.
- [38]. Fan X, Liu T, Liu Y, Ma X, Cui Z. Optimization of primary culture condition for mesenchymal stem cells derived from umbilical cord blood with factorial design. *Biotechnol. Prog.* 2009; 25:499–507. doi:10.1002/btpr.68. [PubMed: 19319963]
- [39]. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 2014; 32:1262–1267. doi:10.1038/nbt.3026. [PubMed: 25184501]
- [40]. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 2013; 31:827–832. doi:10.1038/nbt.2647. [PubMed: 23873081]
- [41]. Sha J, Lippmann ES, McNulty J, Ma Y, Ashton RS. Sequential nucleophilic substitutions permit orthogonal click functionalization of multicomponent PEG brushes. *Biomacromolecules.* 2013; 14:3294–303. doi:10.1021/bm400900r. [PubMed: 23937610]
- [42]. Maye P, Fu Y, Butler DL, Chokalingam K, Liu Y, Floret J, et al. Generation and characterization of Col10a1-mcherry reporter mice. *Genesis.* 2011; 49:410–418. doi:10.1002/dvg.20733. [PubMed: 21328521]
- [43]. Irion S, Luche H, Gadue P, Fehling HJ, Kennedy M, Keller G. Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat. Biotechnol.* 2007; 25:1477–1482. doi:10.1038/nbt1362. [PubMed: 18037879]
- [44]. Hsu CYM, Uluda H. Effects of size and topology of DNA molecules on intracellular delivery with non-viral gene carriers. *BMC Biotechnol.* 2008; 8:23. doi:10.1186/1472-6750-8-23. [PubMed: 18312664]
- [45]. Munye MM, Ravi J, Tagalakis AD, McCarthy D, Ryadnov MG, Hart SL. Role of liposome and peptide in the synergistic enhancement of transfection with a lipopolyplex vector. *Sci. Rep.* 2015; 5:9292. doi:10.1038/srep09292. [PubMed: 25786833]
- [46]. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife.* 2014; 3:e04766. doi:10.7554/eLife.04766. [PubMed: 25497837]
- [47]. Duda K, Lonowski LA, Kofoed-Nielsen M, Ibarra A, Delay CM, Kang Q, et al. High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. *Nucleic Acids Res.* 2014; 42:e84. doi:10.1093/nar/gku251. [PubMed: 24753413]
- [48]. Chiappini C, De Rosa E, Martinez JO, Liu X, Steele J, Stevens MM, et al. Biodegradable silicon nanoneedles delivering nucleic acids intracellularly induce localized in vivo neovascularization. *Nat. Mater.* 2015; 14:532–539. doi:10.1038/nmat4249. [PubMed: 25822693]
- [49]. Sharei A, Zoldan J, Adamo A, Sim WY, Cho N, Jackson E, et al. A vector-free microfluidic platform for intracellular delivery. *Proc. Natl. Acad. Sci.* 2013; 110:2082–2087. doi:10.1073/pnas.1218705110. [PubMed: 23341631]
- [50]. Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* 2015; 33:985–989. doi:10.1038/nbt.3290. [PubMed: 26121415]
- [51]. Uchida S, Kataoka K, Itaka K. Screening of mRNA Chemical Modification to Maximize Protein Expression with Reduced Immunogenicity. *Pharmaceutics.* 2015; 7:137–151. doi:10.3390/pharmaceutics7030137. [PubMed: 26213960]
- [52]. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, et al. Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat. Methods.* 2014; 11:291–293. doi:10.1038/nmeth.2840. [PubMed: 24509632]
- [53]. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced Efficiency of Human Pluripotent Stem Cell Genome Editing through Replacing TALENs with CRISPRs. *Cell Stem Cell.* 2013; 12:393–394. doi:10.1016/j.stem.2013.03.006. [PubMed: 23561441]

**STATEMENT OF SIGNIFICANCE**

CRISPR-Cas9 is a new gene-editing technology for “genome surgery” that is anticipated to treat genetic diseases. This technology uses multiple components of the Cas9 system to cut out disease-causing mutations in the human genome and precisely suture in therapeutic sequences. Biomaterials based delivery strategies could help transition these technologies to the clinic. The design space for materials based delivery strategies is vast and optimization is essential to ensuring the safety and efficacy of these treatments. Therefore, new methods are required to rapidly and systematically screen gene-editing efficacy in human cells. This work utilizes an innovative platform to generate and screen many formulations of synthetic biomaterials and components of the CRISPR-Cas9 system in parallel. On this platform, we watch genome surgery in action using high content image analysis. These capabilities enabled us to identify formulation parameters for Cas9-material complexes that can optimize gene-editing in a specific human cell type.



**Figure 1. Diagram of CRISPR-Cas9 gene-editing and the transgene reporter used in this study** Cas9-GFP mRNA and sgRNA are delivered to the cytoplasm (**A**) where Cas9-GFP mRNA is translated into a Cas9-GFP fusion protein. Cas9-GFP complexes with sgRNA enabling recognition (**B**) of a specific 19-20 bp genomic target. Cas9 DNA endonuclease activity forms DNA double strand breaks (DSBs) at the target site (**C**). Non-homologous end joining (NHEJ) or other error prone DNA repair (**D**) leads to insertion or deletion mutations (gene-edits) at the target loci. Edits at the mCherry reporter can result in mutations that lead to loss of mCherry fluorescence (**F**), whereas unedited DNA retains mCherry reporter expression (**E**). Bioanalyzer electropherograms of transcribed and purified Cas9-GFP RNA (**G**) and

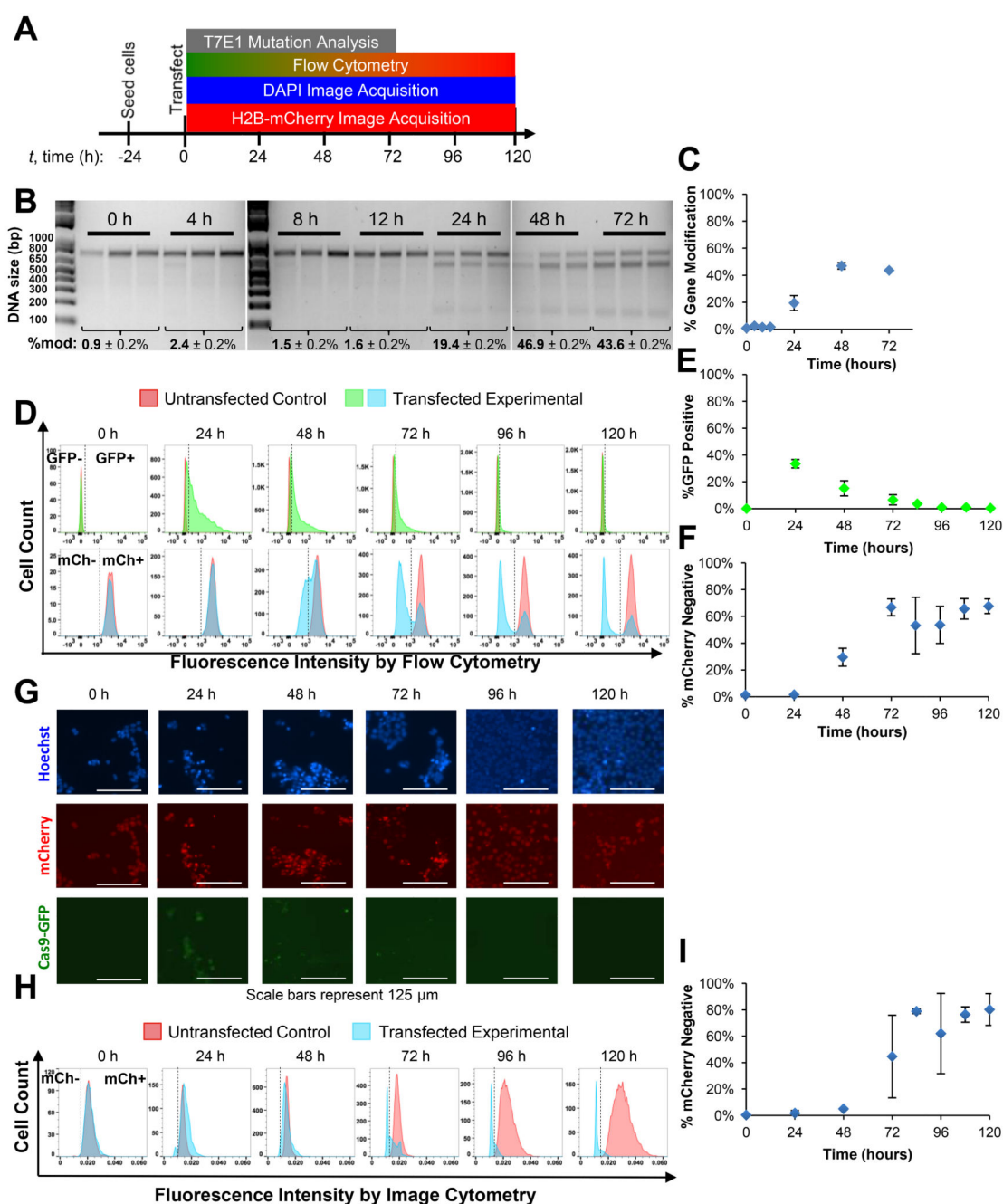
purified in vitro transcribed (IVT) mCherry sgRNA (**H**). Commercial prep is Thermo Scientific's Gene-Art product.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2. Gene-editing outcomes tracked using standard and image cytometry methods**  
Schematic of analysis during timecourse after delivery of Cas9 materials (A). Agarose gel result from T7E1 assay (n=3) (B). % Modification is provided +/- 95% confidence intervals of the mean. Quantification of percent gene modification over time from T7E1 assay (C). Representative flow cytometry histograms of GFP and mCherry expression in experimental treatments versus untransfected controls (D). Quantification of % mCherry positive cells (E) and % GFP negative cells (F) from flow cytometry assays. Representative images of nuclei (blue), H2B-mCherry (red) and Cas9-GFP (green) acquired for HCA (G). Image cytometry

histograms from CellProfiler output (**H**). Quantification of image cytometry data (**I**). n=3 samples at each time point for flow cytometry and image cytometry analysis. All error bars describe 95% confidence intervals of the mean.

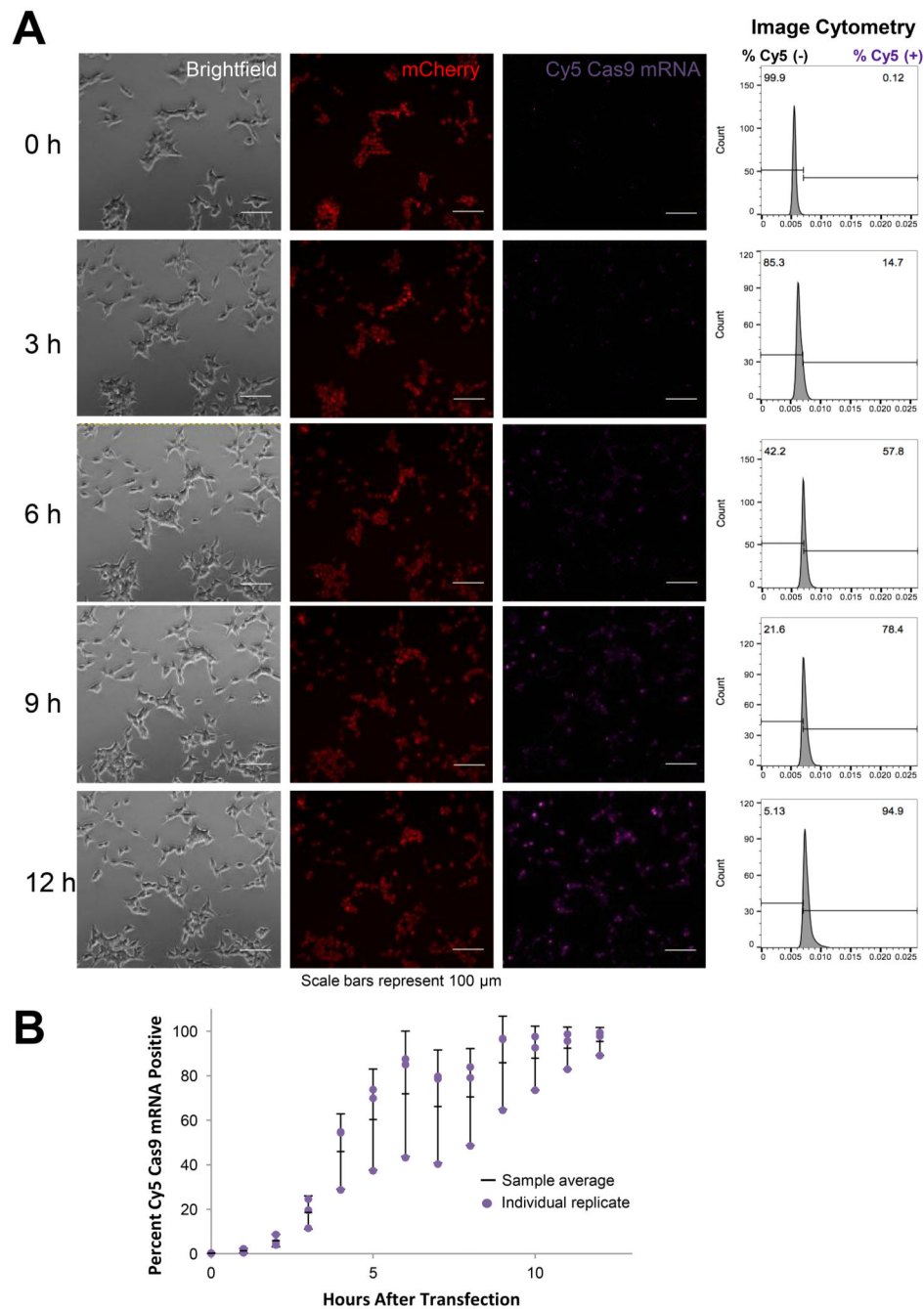
Author Manuscript

Author Manuscript

Author Manuscript

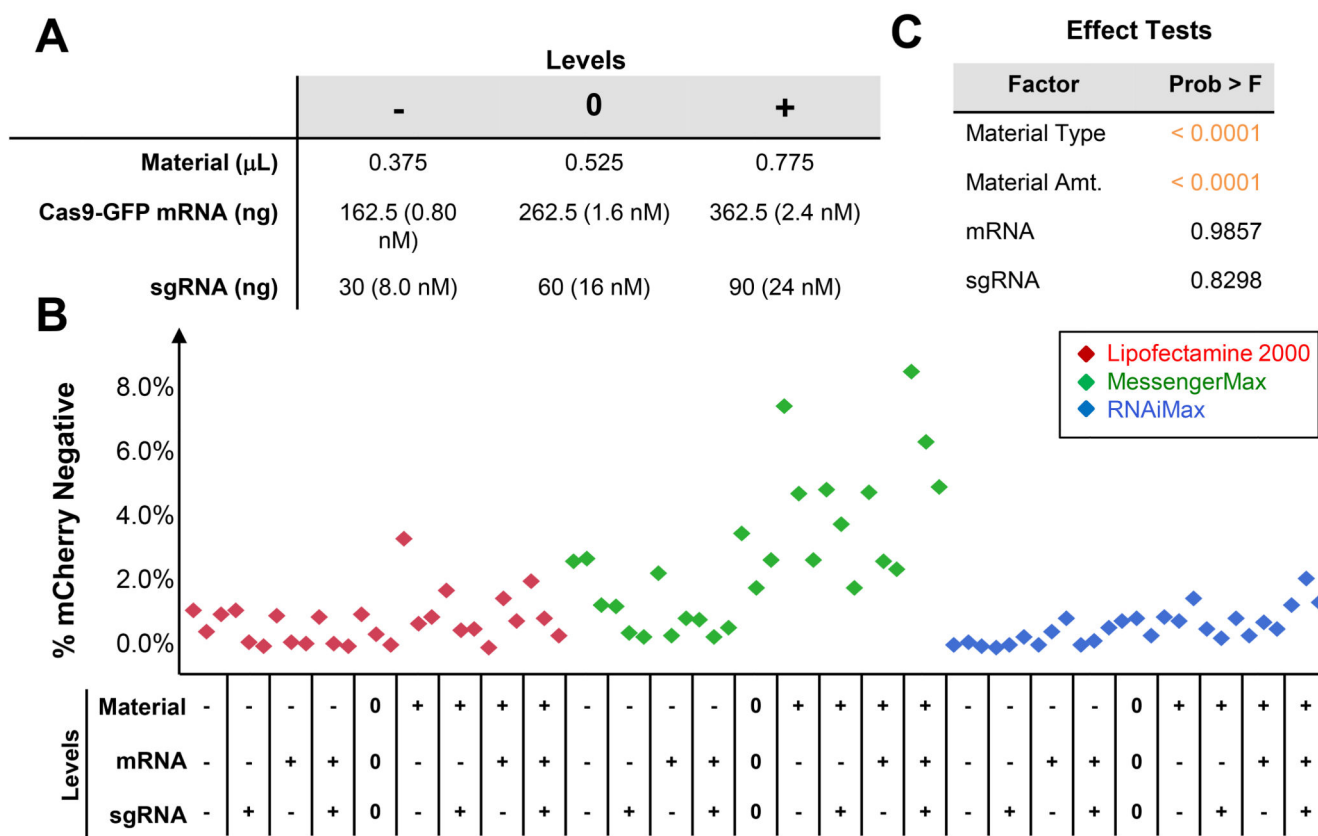
Author Manuscript





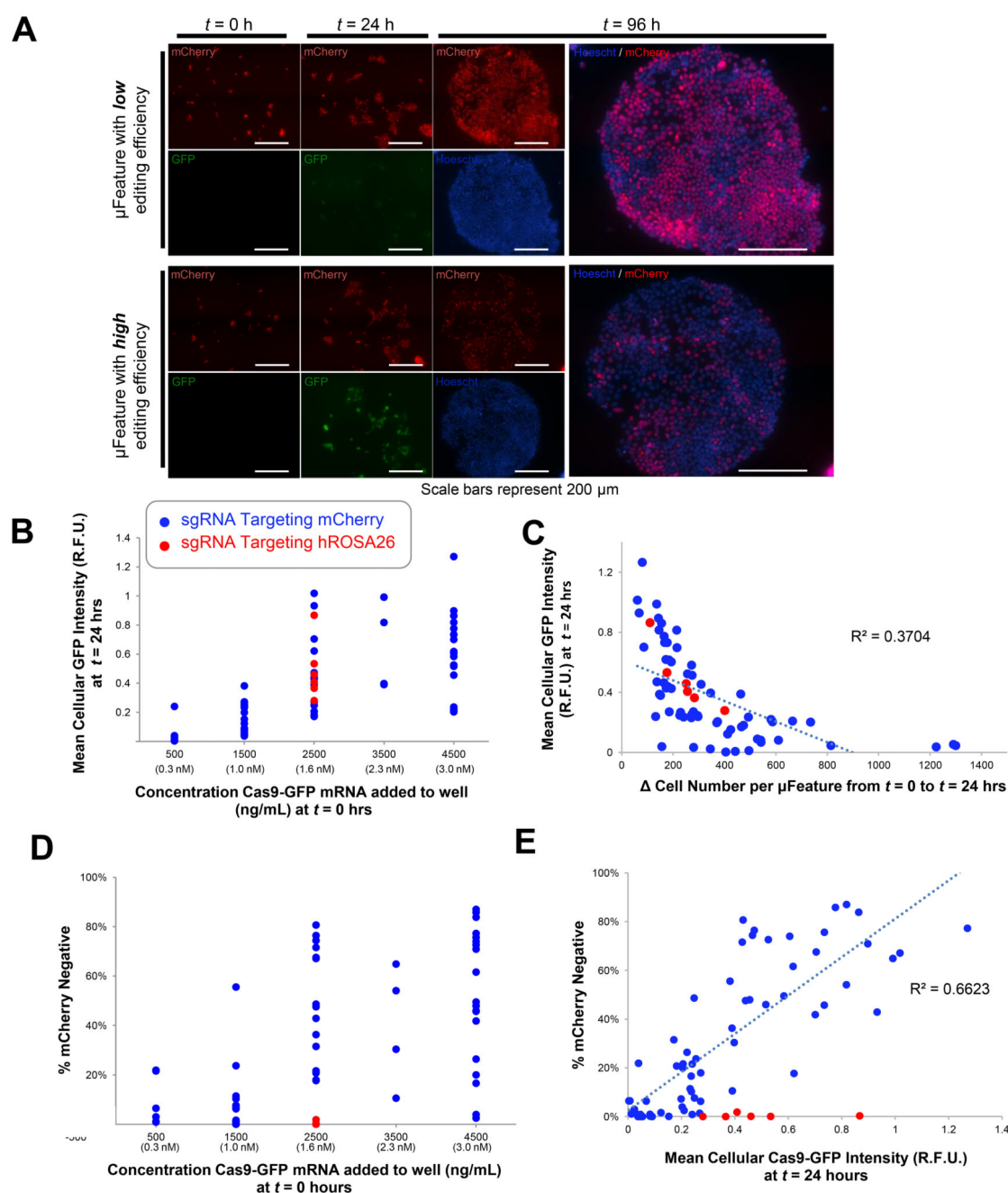
**Figure 3. Tracking fluorescently-labeled Cas9 mRNA delivery**

Automated time course to track Cy5 labeled Cas9 mRNA delivery to cells. Representative images and image cytometry histograms (A) and image cytometry quantification of percent Cy5 positive cells in n=3 physical replicates (B). Control gating was performed on samples at t=0 immediately following transfection. Error bars describe 95% confidence intervals around the mean.



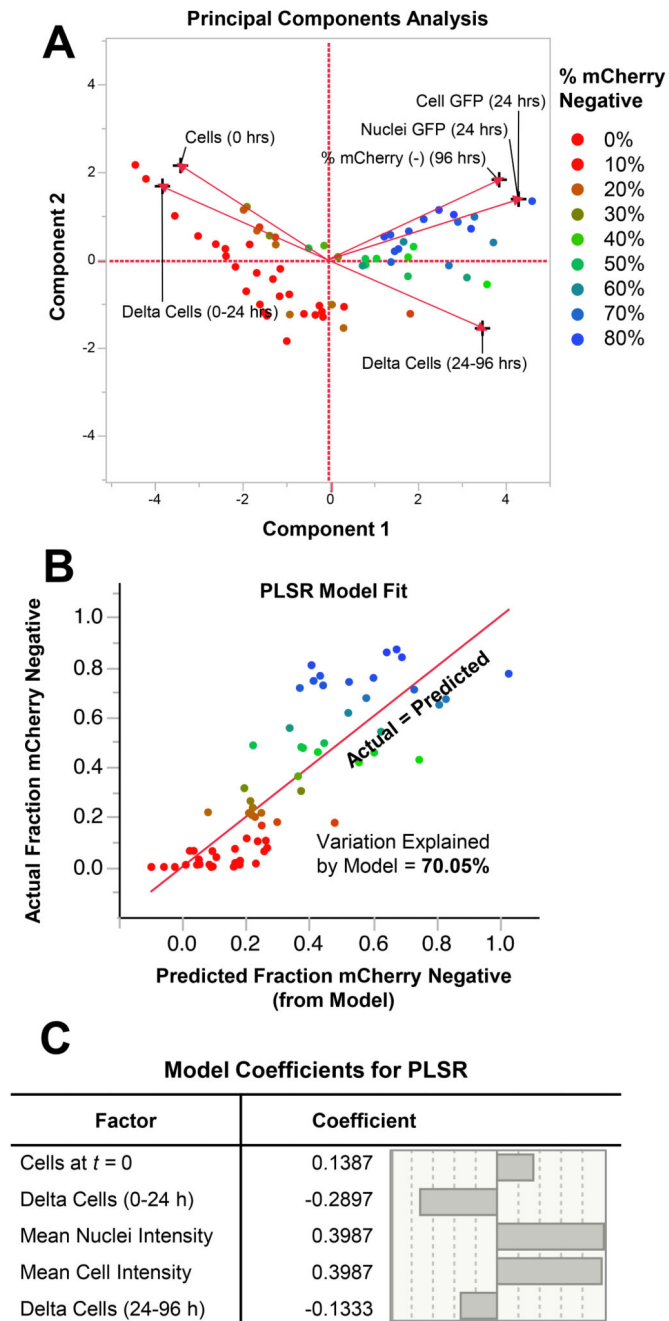
**Figure 4. Design of experiments screening analysis 1 for optimal gene-editing parameters**  
 Three different lipid based delivery materials (Lipofectamine 2000, MessengerMax, and RNAiMax) were screened at specified amounts across variable amounts of Cas9-GFP mRNA and mCherry C sgRNA (A). Gene-editing efficiencies of delivery materials across all specified conditions (n=3 replicates per condition) (B). Probability of an effect for each factor (full statistics described in Figure C.1A) (C). Orange text indicates a significant effect ( $p < 0.05$ ).





**Figure 6. High content analysis of CRISPR-Cas9 RNA delivery and editing using  $\mu$ Feature plates** Representative images of nuclei (blue), H2B-mCherry (red) and Cas9-GFP (green) of cells on  $\mu$ Features passed to CellProfiler for analysis (A). Enlarged merged image on right panel is provided to highlight cells without mCherry expression. Variable expression of mean Cas9-GFP for each subpopulation of cells within micropatterned wells at  $t = 24$  hours after transfection (B). Blue indicates delivery of mCherry C sgRNA, while red indicates delivery of a control sgRNA against hROSA26. Correlation of change in cell number per  $\mu$ Feature (cell number at  $t = 24$  hours – cell number at  $t = 0$  hours) to mean cellular GFP at  $t = 24$

hours (**C**). Dashed trendlines indicate linear regression. Variable outcomes of mCherry gene-editing within the same micropatterned well at  $t = 96$  hours after transfection (**D**). Correlation of subpopulation level percent mCherry loss at  $t = 96$  hours versus cellular Cas9-GFP intensity at  $t = 24$  hours (**E**).



**Figure 7. Multivariate analysis of gene-editing outcomes using  $\mu$ Feature plates**

Biplot of principal components analysis (A). Red circles on the biplot correspond to  $\mu$ Features with low editing (0-10%), while green circles correspond to medium editing (30-60%) and blue circles to high editing efficiency (70-80%). Proximity of factors (labeled points at ends of arrows) with  $\mu$ Features displaying low or high editing, indicates relationship of those factors with associated editing outcome. Actual versus predicted plot for final PLSR model (B). Proximity of data points to red line (actual = predicted) indicates goodness of fit for model. The final PLSR model explained 70% of the variability of percent



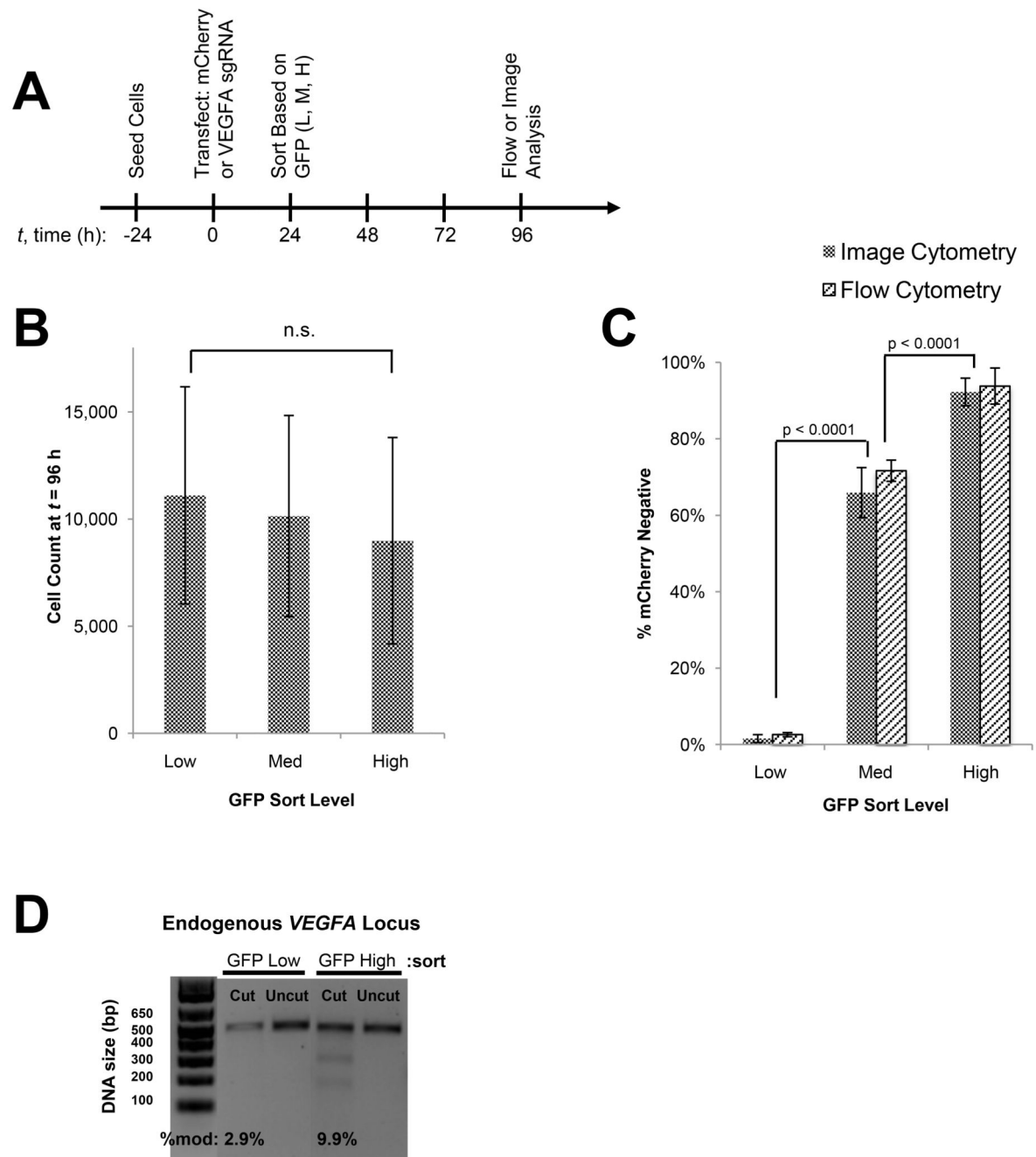
mCherry expression loss (see model validation in Appendix, Fig. F.1A,B). Coefficients from PLSR model for centered and scaled data (C). Positive coefficients indicate positive effect on percent mCherry gene-editing and negative coefficients indicate negative effect on percent mCherry gene-editing. Larger coefficients indicate greater overall effect on gene-editing outcome.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 8. Fluorescence assisted cell sorting (FACS) analysis of relationship between GFP intensity and editing outcomes**

Diagram of setup used in this experiment (**A**). No significant difference ( $p > 0.05$ ) in cell count at  $t = 96$  hours for cells sorted for low, medium and high relative GFP intensity ( $n = 3$ -4 physical replicates) (**B**). Example gating for FACS provided in Appendix (Figure G.1). Higher relative GFP intensity results in greater gene-editing at mCherry transgene (**C**). Higher relative GFP intensity results in increased % gene modification at endogenous *VEGFA* locus (**D**). Error bars represent 95% confidence intervals of the mean. Significant

difference determined using single tailed paired t-test where  $p < 0.05$  was considered significant.