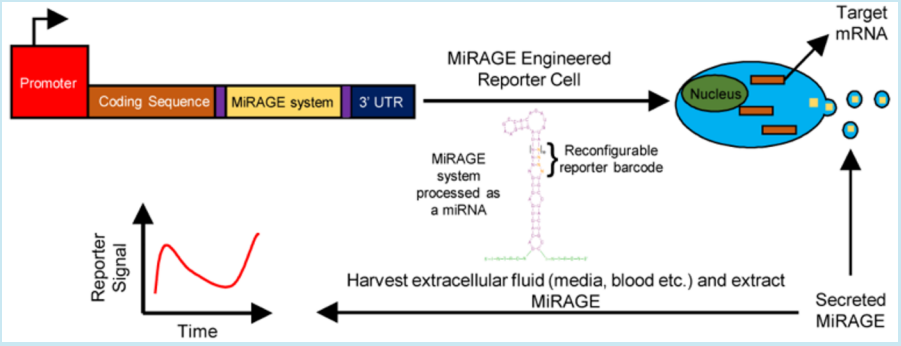
ACS Synthetic Biology
Cite This: ACS Synth. Biol. 2019, 8, 25−33
Letter
pubs.acs.org/synthbio

S-MiRAGE: A Quantitative, Secreted RNA-Based Reporter of Gene Expression and Cell Persistence

Kinshuk Mitra,[†,#] William N. Feist,[‡,#] Simone Anfossi,[§] Enrique Fuentes-Mattei,[§] Maria Ines Almeida,[§,▽] Jean J. Kim,[∥] George A. Calin,[§,⊥] and Aryeh Warmflash*[,†,‡]

[†]Department of Bioengineering, Rice University, Houston, Texas 77005, United States
[‡]Department of Biosciences, Rice University, Houston, Texas 77005, United States
[§]Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, United States
[∥]Department of Molecular and Cellular Biology, Stem Cells and Regenerative Medicine Center, Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, Texas 77030, United States
[⊥]The Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, United States

Supporting Information



ABSTRACT: Nondestructive measurements of cell persistence and gene expression are crucial for longitudinal research studies and for prognostic assessment of cell therapies. Here we describe S-MiRAGE, a platform that utilizes small secreted RNA molecules as sensitive and quantitatively accurate reporters of cellular processes. S-MiRAGE allows cellular numbers or gene expression to be measured from culture media or from biofluids. We show that multiple S-MiRAGE reporters can be multiplexed, and demonstrate the utility of S-MiRAGE by monitoring the differentiation status of human embryonic stem cells *in vitro* and tumor growth in a mouse model *in vivo*.

A number of proteins are frequently used as genetically encoded reporters including fluorescent proteins,[1] luciferase,[2] and magnetic resonance imaging (MRI) or positron emission tomography (PET)-detectable proteins,[3] with the choice of protein depending on the needs of the application. All of these systems suffer from limited multiplexability and poor signal-to-noise in challenging applications,[4] for example, when used as a reporter of weakly expressed endogenous genes. Motivated by recent successes in adopting circulating nucleic acids (cNAs) for clinical use,[5] we built a platform in which small expressed RNA molecules are used as reporters.

A previous attempt to use secreted synthetic miRNAs demonstrated the feasibility of this approach;[6] however, this study suffered from several limitations. Nonspecific detection of the synthetic miRNA prevented discrimination with certainty between animals with and without cells expressing synthetic miRNAs injected. Furthermore, no quantitative correlations between existing reporters and secreted miRNAs were reported, and to date, no studies have shown the ability to measure endogenous gene expression with these reporters. Overcoming these limitations could allow for the development of synthetic cNA-based reporters for a variety of research applications, and, in the future, for use as biomarkers[6,7] for medical interventions lacking quantitative metrics of success such as autologous therapies.

In this study, we report the development of S-MiRAGE (Synthetic-MicroRNA-like Reporters Assaying Gene Expression, hereafter referred to as MiRAGE). We show that choosing miRNA sequences from distantly related species as secreted reporters allows measurement from mammalian culture media or biofluids with very low background, thus circumventing the problem that has prevented implementation of this strategy to date. When these sequences are placed within the genome of

Received: September 26, 2018
Published: December 14, 2018

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

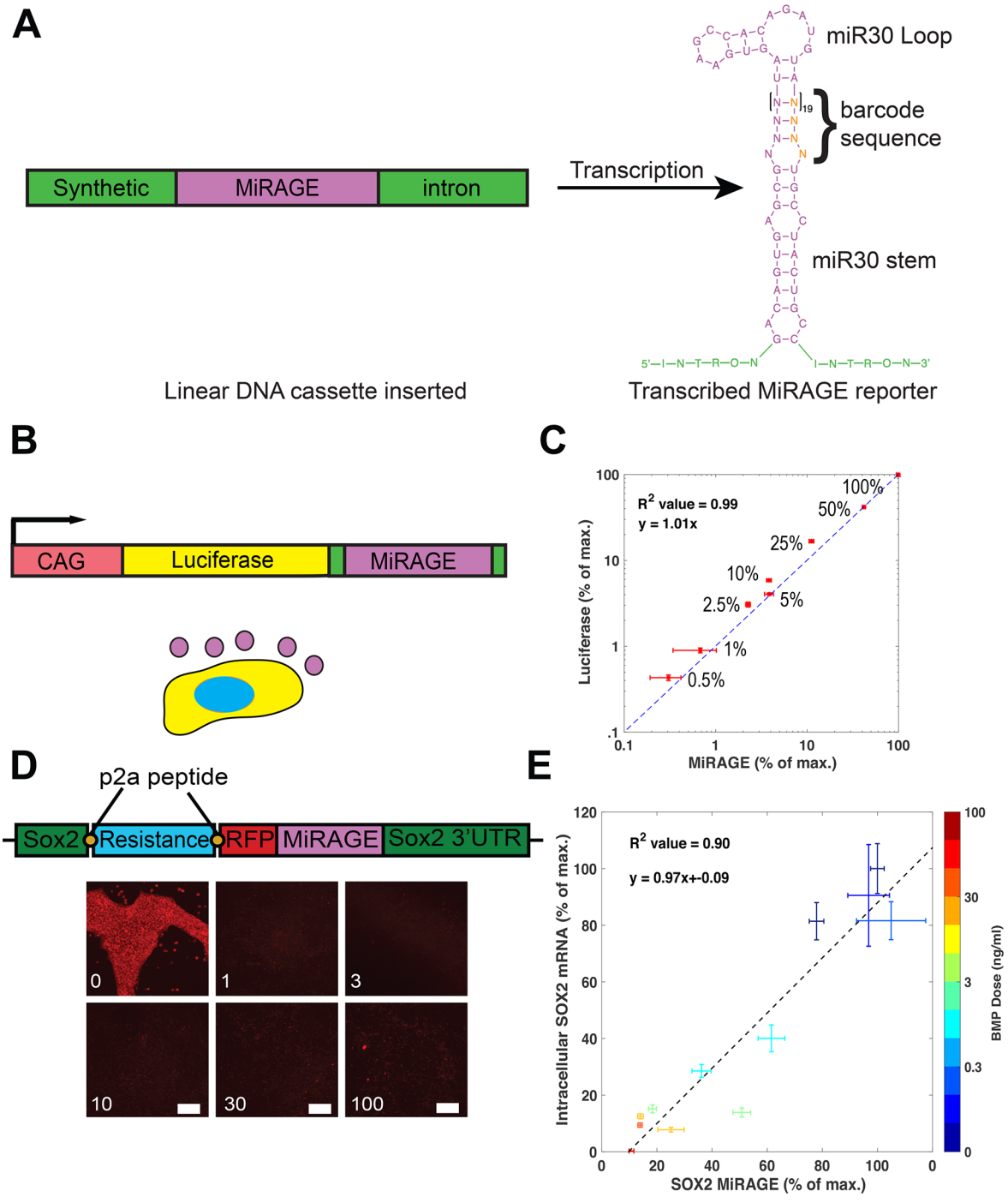


Figure 1. The MiRAGE system reports on cellular presence and endogenous gene expression. (A) Schematic of the MiRAGE system based on a modified hsa-miR-30 sequence flanked by synthetic introns (green). The custom 22 or 23 nucleotide barcode sequence is labeled in orange. (B) Schematic of a luciferase and MiRAGE expressing cell. MiRAGE is secreted by the cell, while luciferase is retained intracellularly. (C) Comparison of change in luciferase and MiRAGE reporter levels in media. 4T1-MiRAGE-luciferase cells were mixed with unmodified cells in varying proportions. Each data point is labeled with the ratio of reporter to wild-type cells; null hypothesis of slope of zero rejected at $p < 0.001$. (D) Schematic of modified SOX2 allele in the ESI017-SOX2-RFP-MiRAGE reporter cell line (top) and images of the RFP reporter upon treatment with the indicated concentrations of BMP4 (in ng/ml) for 4 days (bottom); scale bar 100 μm. (E) Comparison of intracellular *SOX2* mRNA with the MiRAGE reporter assayed from culture media. Both the *SOX2* mRNA levels and the miR levels were determined by qPCR. mRNA is normalized to GAPDH, while the MiRAGE reporter is normalized to miR-17. Data points are colored by the dose of BMP4 (in ng/mL); null hypothesis of slope of zero rejected at $p < 0.001$. Fold changes are as compared to a BMP4 dose of 0 ng/mL. In (C) and (E), each data point represents averages and standard deviations across three biological replicates. Note that in (C) and (E) each set of data was normalized to its maximum so that the fold changes from each reporter could be compared directly. The slope of the line is an indicator of the ratio of the fold changes as measured by the two reporters (*i.e.*, luciferase and MiRAGE or SOX2 qPCR and MiRAGE). The slope is 1 to within the error in all cases, indicating perfect agreement between the methods.

mammalian cells, they can be used to report nondestructively on the presence of those cells, or, when placed downstream of an endogenous gene using CRISPR-Cas9, on the expression of that gene. We show quantitative agreement between MiRAGE

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

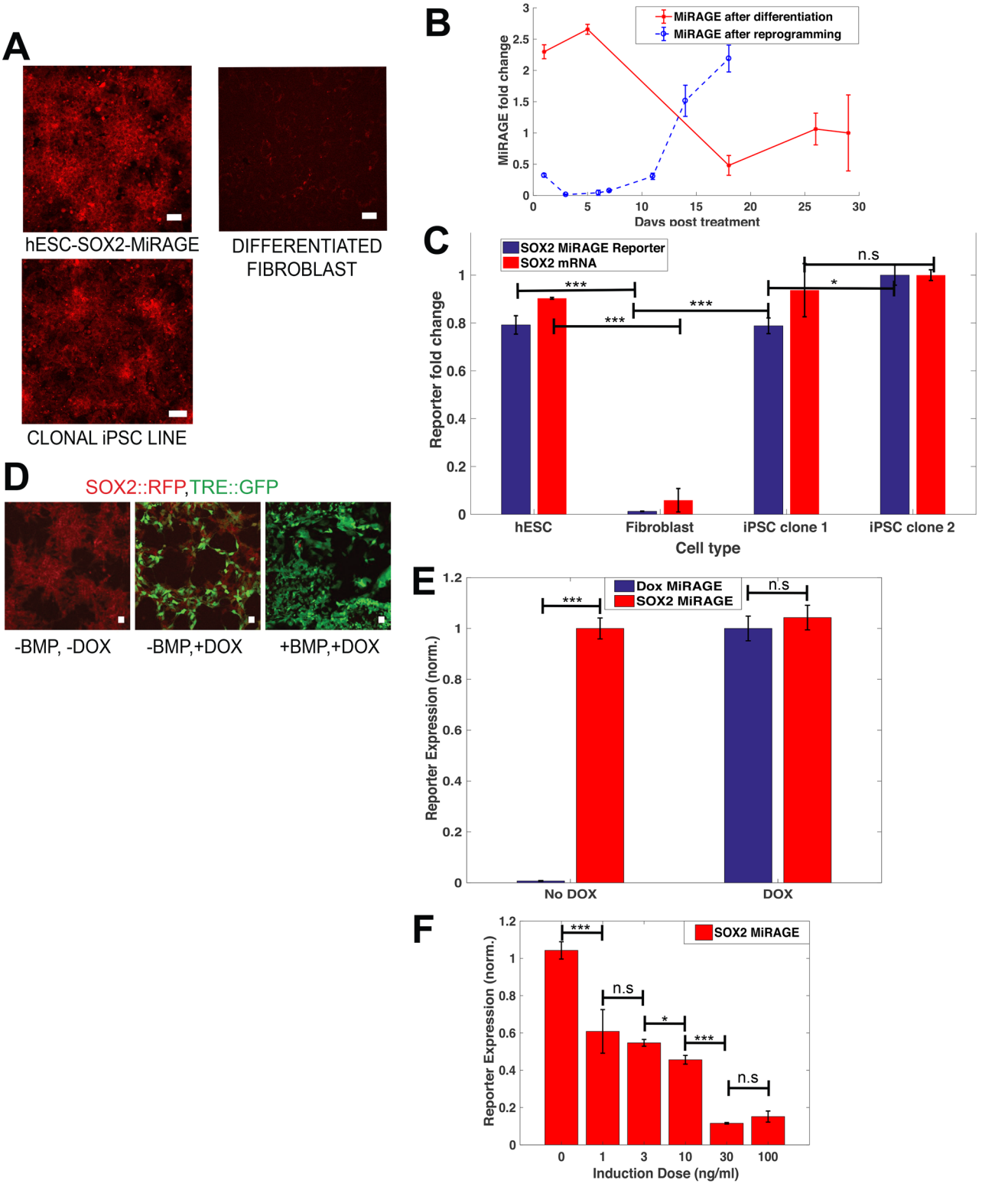


Figure 2. MiRAGE reports on reprogramming and allows for multiplexed monitoring of gene expression. (A) Images of ESI017-SOX2-RFP-MiRAGE cells at the indicated points in the reprogramming and differentiation process; scale bar 50 μm. (B) Expression level of the MiRAGE reporter in time during differentiation to fibroblasts and reprogramming assayed from the culture media by qPCR. MiRAGE levels were measured after beginning differentiation of hESCs into fibroblasts (day 0, differentiation) or after the removal of Sendai virus from cell culture during reprogramming (day 0, reprogramming). Fold changes are relative to the MiRAGE reading at the 28th day post differentiation. (C) Comparison of qPCR measurements of *SOX2* mRNA levels to MiRAGE levels in undifferentiated hESCs, differentiated fibroblasts, and 2 reprogrammed iPSC clones. Data set is normalized to readings obtained from iPSC clone 2. (D) Expression of fluorescent reporters of SOX2 and doxycycline with addition of either BMP4 or doxycycline; scale bar 20 μm. (E) Multiplexed detection of MiRAGE reporters for doxycycline and *SOX2* mRNA from the same cells in the presence and absence of Dox. Each reporter is normalized to readings taken under Dox induction. (F) SOX2MiRAGE expression levels as a function of BMP4 dose in the presence of Dox. Fold changes are relative to the value at 0 ng/ml BMP4 with Dox. Dox was used at a concentration of 100 ng/mL.

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

measurements and established techniques such as luciferase and quantitative PCR (qPCR), and demonstrate the capabilities of MiRAGE by monitoring the differentiation and reprogramming status of human embryonic and induced pluripotent stem cells, and by tracking cancer progression *in vivo* in a mouse model.

RESULTS AND DISCUSSION

Identification of Orthogonal Probes with Minimal Background. MiRAGE consists of a nontargeting miR placed downstream of a gene of interest (Figure 1a). MicroRNAs encapsulated in extracellular vesicles (microvesicles and exosomes) or bound to protein (Ago2)[8] and lipoprotein (HDL) are actively secreted by cells, and can be detected in the extracellular fluid.[9] The use of nucleic acids allows for multiplexing by simply changing the sequence, and sensitive detection *via* amplification by quantitative reverse transcription polymerase chain reaction (qRT-PCR). We placed the miR-30 backbone,[10] whose stem strands can be modified, within a synthetic intron[11] to reduce disruption of gene expression. We found that the MiRAGE reporter could be readily detected from the media of cultured cells, and that its expression level measured from media was the same whether placed at the 5′ or 3′ terminus of the protein-coding sequence of RFP (Supplementary Figure S1A,B). To minimize disruption of translation, all further experiments were performed with the MiRAGE cassette placed at the 3′ terminus of the protein coding sequence. We also found that probes against various miRs including a randomly generated sequence containing a previously described motif for exosomal sorting (ExoMotif)[12] yielded some signal both in media from unmodified cells and mouse serum (Supplementary Figure S1C). To reduce this background, we used qRT-PCR to test whether existing Taqman probes for zebrafish and *C. elegans* microRNAs produce background signals from mammalian biofluids (mouse serum and IP fluid). We selected two probes (cel-miR-2 and dre-miR-458, hereafter miR-2 and miR-458, respectively), which are not evolutionary conserved in mammals, for use in further studies. miRNAs with fewer targets are known to sort more efficiently to exosomes,[13] and therefore use of nonmammalian miRNAs may aid in maximizing exosomal secretion. We found that each probe specifically detected its targets and exhibited no cross talk with the second probe or background signal in mammalian biofluids.

MiRAGE Quantitatively Reports on Cell Presence and Gene Expression. Luciferase is a commonly used reporter due to its sensitivity and wide dynamic range.[14] We placed a MiRAGE cassette within a synthetic intron at the 3′ end of firefly luciferase under the control of a constitutive promoter and used the ePiggyBac system to stably integrate this construct into mouse 4T1 cells (Figure 1B). We mixed these cells with unlabeled 4T1 cells at varying ratios and observed a perfect linear relationship between the luciferase signal and MiRAGE levels detected by qRT-PCR on samples extracted from the culture media, both of which were directly proportional to the number of MiRAGE cells in the culture (Figure 1C, $R^2 = 0.99$). This linear relationship had a slope of one showing that the fold changes measured by MiRAGE and luciferase were identical over a large dynamic range ($>10^3$). To confirm that the MiRAGE reporters were secreted through exosomes or other extracellular vesicles (EVs), we purified EVs from the media and compared MiRAGE expression from these exosomes with the luciferase measurements. We found a similar linear relationship between MiRAGE and luciferase (Supplementary Figure S2A); however, detection was not as robust at very low levels, possibly due to loss of material during purification.

We next evaluated whether our system can monitor endogenous gene expression. We integrated a MiRAGE reporter, RFP, and a blasticidin resistance gene at the C-terminus of the *SOX2* gene, a marker for pluripotency, in the human embryonic stem cell (hESC) line ESI017 using CRISPR-Cas9[15] (Figure 1D). We determined that modifications to the wild-type *SOX2* allele occurred in less than 5% of cells heterozygous for the reporter (Supplementary Figure S3A,B). We treated these reporter cells with varying concentrations of Bone Morphogenic Protein 4 (BMP4), causing differentiation to a trophectoderm-like fate, and loss of SOX2 expression.[16,17] Fluorescent signal from the integrated RFP gene was completely lost at doses as low as 1 ng/mL (Figure 1d), which mirrored the loss of SOX2 protein as measured by immunofluorescence (Supplementary Figure S4). We compared the level of the MiRAGE reporter measured from either the media or the cells with the amount of *SOX2* mRNA in the same cells and found a nearly perfect correlation (Figure 1e, $R^2 = 0.98$ and Supplementary Figure S5). The best correlation between cellular *SOX2* mRNA and MiRAGE was observed when the expression of the latter was normalized to that of a constitutively expressed microRNA, miR-17,[18] rather than an mRNA (GAPDH), potentially revealing variability in the total secretion of miRNA which can be accounted for by proper normalization (Supplementary Figure S5). Thus, the MiRAGE reporter was a quantitative, nondestructive readout for the levels of SOX2 mRNA that varied more gradually with BMP4 levels than either the SOX2 protein or the RFP reporter. Finally, we tested whether the ability to measure gene expression was specific to MiRAGE reporter by assaying for SOX2 expression directly from the media. We found that although SOX2 mRNA could be detected in the media, its levels did not correlate with those in the cells, indicating that mRNA in the media is a poor reporter for gene expression (Supplementary Figure S2B).

MiRAGE Can Monitor Reprogramming to Pluripotency. We next used the MiRAGE-*SOX2* reporter to monitor reprogramming to pluripotency directly from the culture media. We differentiated ESI017-MiRAGE-*SOX2* cells into fibroblasts using an established protocol,[19] and confirmed differentiation *via* immunostaining that showed loss of the pluripotency marker SOX2 and upregulation of the fibroblast marker Vimentin (Supplementary Figure S6). We then initiated reprogramming by overexpressing canonical reprogramming factors: OCT4, KLF4, SOX2, and c-MYC using a Sendai virus platform.[20] Imaging of the cells and analysis of MiRAGE levels in cell-culture media indicated that the expression of both reporter types (RFP and MiRAGE) fell below the threshold of detection during differentiation. Following the initiation of reprogramming, the reporters reappeared and gradually rose to the level in undifferentiated cells (Figure 2A,B). We compared the MiRAGE reporter with *SOX2* mRNA at critical points in the differentiation and reprogramming process, and again found nearly perfect agreement (Figure 2C).

MiRAGE Reporters Are Multiplexable. Current fluorescent and luminescent reporters also suffer from a limited capacity to multiplex due to spectral overlap. In contrast, we expected MiRAGE reporters with different stem sequences to be orthogonal. As proof of principle, ESI017-MiRAGE-SOX2 cells were engineered to express miR-2 and GFP in response to the addition of doxycycline (Figure 2D). Adding doxycycline strongly induced the miR-2 MiRAGE reporter without affecting

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

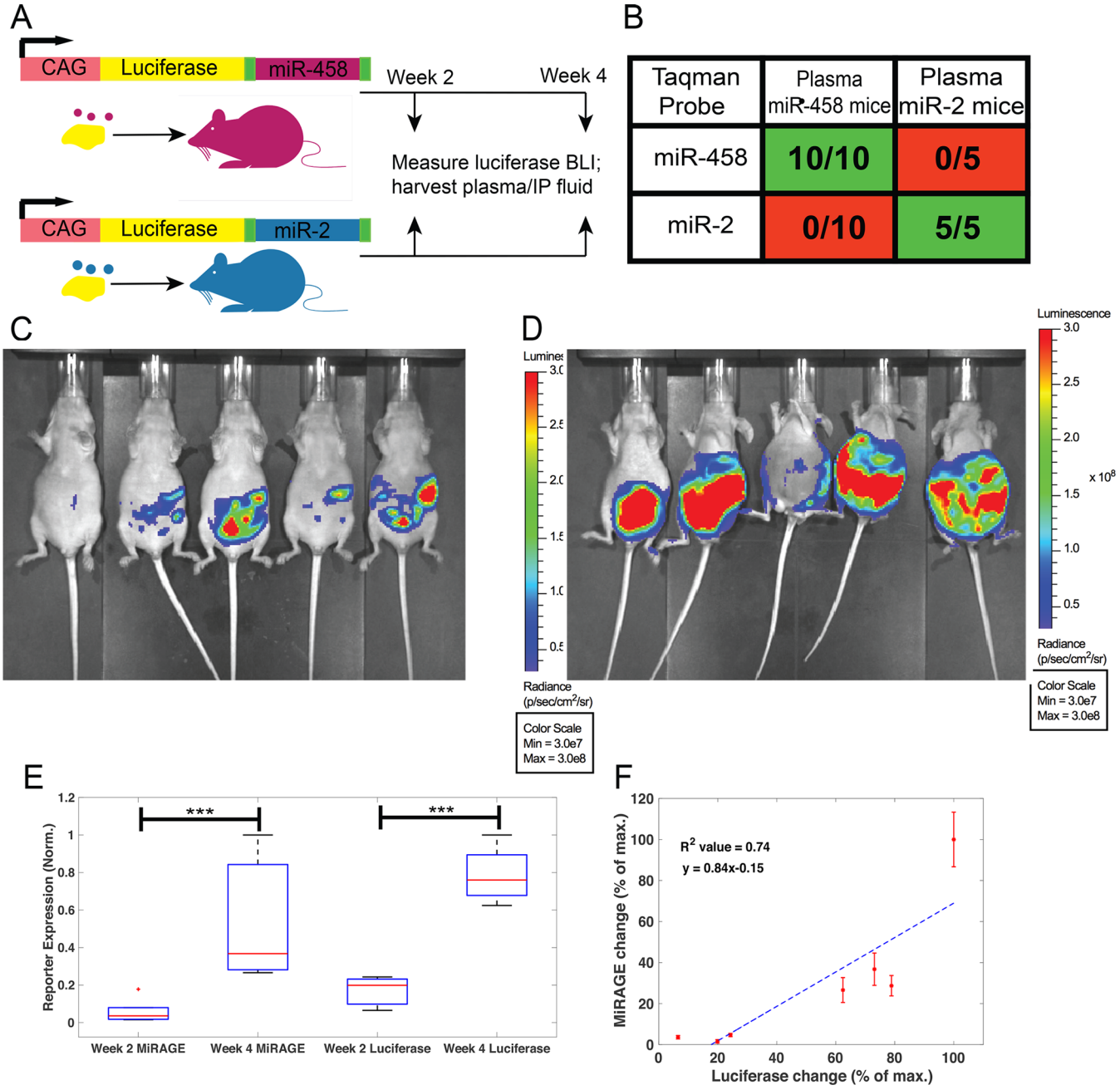


Figure 3. MiRAGE can be used to report on cell presence in vivo. (A) Schematic of the experiment. (B) Table showing the fraction of mice in which the indicated probe yielded positive results from the indicated blood plasma sample. All successful detections represent samples with CT values of less than 36 for all three replicates, while all negative samples either had no amplification or CT values greater than 36 for all three replicates. (C,D) Whole body bioluminescent images of miR-458 mice after week 2 (C) and week 4 (D) of tumor implantation. (E) Boxplot of MiRAGE and luciferase signal across week 2 and week 4 miR-458 mice cohorts; 5 mice per boxplot group. Data for each reporter is normalized to the mouse with the highest expression of that reporter. (F) Comparison of the level of circulating MiRAGE determined by qPCR from blood plasma with luciferase dosimeter signal detected for miR-458 mice. Each data point corresponds to one animal. We removed from consideration three mice whose IP fluid buildup interfered with the bioluminescent imaging; null hypothesis of slope of zero rejected at $p < 0.05$. Data sets for MiRAGE and luciferase are normalized to their maximum values.

the expression of the miR-458 reporter of *SOX2* expression (Figure 2E). Similarly, the downregulation of the miR-458 upon addition of BMP was not affected by the doxycycline-induced upregulation of miR-2 (Figure 2F).

MiRAGE Reports on Cell Presence *in Vivo*. A limited subset of reporter genes can be detected *in vivo* due to the restricted ability to detect signal through deep tissue. We reasoned that MiRAGE, detected by a blood test, could function as an efficient reporter for applications not requiring spatial information and negate the need for expensive instrumentation. To test this hypothesis, we injected nude female mice intraperitoneally with HeYA8 ovarian cancer cells engineered to express luciferase and either MiRAGE miR-458 or MiRAGE miR-2 (Figure 3A; 10 mice each group). Five of the miR-2 mice did not survive the experiment due to rapid cancer progression and were not included in the analysis, so that 15 mice were

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

analyzed in total. At either the second or fourth week after HeYA8 cell injection, blood and intraperitoneal (IP) fluid were harvested, the blood processed into plasma, and total small RNA extracted from both biofluids. We assayed all fluids for both miR-2 and miR-458 and found that the miR detected in both plasma and IP fluid (15/15 mice in both cases) matched that expressed in the cell line injected into the animal, while the miR expressed by the noninjected cell line was never detected in the blood plasma (0/15 mice) and only in a small number of cases in the IP fluid (3/15 mice) (Figure 3B, Supplementary Figure S7A). Thus, MiRAGE is a sensitive and specific means to detect the presence of engineered cancer cells in vivo, particularly from the blood plasma.

To determine whether the quantitative levels of the MiRAGE reporter can be used as readout of tumor growth, mice were also imaged for total whole-body bioluminescence after 2 and 4 weeks prior to biofluid collection (Figure 3C,D). We found that the fold increase in MiRAGE signal from plasma was identical to that determined by luminescence (Figure 3E). We compared MiRAGE signal from plasma to luciferase dosimeter values in individual mice from both week 2 and week 4 cohorts. We observed agreement between the two reporters with the MiRAGE signal increasing somewhat faster than the luminescence at high levels (Figure 3F, $R^2 = 0.74$). In contrast, whereas we did detect the MiRAGE reporter in the IP fluid, its levels did not correlate quantitatively with that of the luciferase, possible due to poor clearance from the IP compartment or dilution of MiRAGE by increased IP fluid accumulation (Supplementary Figure S6B,C).

Our results here demonstrate proof of principle that a secreted miRNA reporter can be used as a quantitative, multiplexable readout of cell numbers or gene expression both *in vitro* and *in vivo*. Leveraging recent advances in methods for RNA detection[21] will further improve the sensitivity of these measurements. Thus, MiRAGE will be a powerful tool for multiplexed nondestructive measurements and could potentially be used as a surrogate secreted biomarker when a natural one is absent.

MATERIALS AND METHODS

Cell Culture, Transfections, and Selection. Human ESI017 cells were cultured on Matrigel (BD Biosciences, 1:200 in DMEM-F12)-coated dishes in mTeSR1 media (Stem Cell Technologies, Vancouver, BC, Canada). Cells were passaged with Dispase when 60−80% confluent. Cells were regularly tested for mycoplasma and found negative. The ESI017-Sox2-RFP-MiRAGE line was created *via* CRISPR-Cas9 by nucleofecting ESI017 cells with 2.5 μg each of SOX2 Cloning Free Homology Donor and Guide RNA using a P3 Primary Cell Kit and 4D-Nucleofector Core Unit (Lonza). Nucleofected cells contained a resistance cassette and were selected after 2 days with 2 μg/mL blasticidin for 10 days before being sorted by flow cytometry for expression of RFP.

Human HEYA8 and Mouse 4T1 cells were cultured in RPMI 1640 medium (Corning) containing 10% fetal bovine serum. HEYA8 and 4T1 cell lines expressing MiRAGE were created by transfecting cells with the Lipofectamine 2000 kit (Invitrogen, mixed in Opti-MEM medium from Gibco) followed by 1 μg/mL puromycin selection for 4 days.

Stem Cell Differentiation and Reprogramming. *BMP4-Mediated Differentiation.* 1 $\times 10^5$ cells were seeded using accutase passaging in 35 mm μ-Dishes (Ibidi, Martinsried, Germany) and grown in mTeSR1 with ROCK-Inhibitor Y-27672 (10 μM; StemCell Technologies). After 2 days cells were treated with BMP4 (R&D systems) at predefined doses (1, 3, 5, 10, 30, or 100 ng/mL) in mTeSR1 for 4 days.

Differentiation of hESCs into Fibroblast-like Cells. On Differentiation (Diff) Day 0, ESI-SOX2M hESCs were pretreated with 10 μM Y-27632 in mTeSR1 medium (StemCell Technologies) for 30 min. Colonies were nonenzymatically detached as small clumps using ReLeSR Passaging Reagent (StemCell Technologies) according to manufacturer's instructions, and then resuspended in embryoid body (EB) medium (consisting of 20% fetal bovine serum (FBS) (Hyclone), 80% KnockOut Dulbecco's Modified Eagle's Medium (KO−DMEM) (Thermo Fisher Gibco), 1× GlutaMAX, 1× MEM Non-Essential Amino Acids (NEAA), 1× Pen Strep, and 55 μM β-mercaptoethanol) supplemented with 10 μM of Y-27632 on the first day only. The cell clumps were grown in suspension for 4 days in an ultralow attachment 6-well plate (Corning Costar) to form EBs. On Diff Day 4, EBs were plated onto 0.1% gelatin-coated 6-well plates (BD Falcon), and cultured for 10 more days with complete EB medium changes every 2−3 days. Cells of various morphologies grew out radially from the adherent EBs.

On Diff Day 14, the outgrowth culture was passaged (Passage 1) using Accutase (StemCell Technologies) and replated to gelatin-coated 6-well plates in Fibroblast (Fib) medium (consisting of 10% FBS (GE HyClone), 90% high glucose DMEM (Thermo Fisher Gibco), 1× MEM NEAA, and 1× Antibiotic-Antimycotic). The monolayer culture was passaged two more times (Diff Day 18 and 28) to expand fibroblast-like cells. Medium conditioned for at least 24 h was collected on Diff Days 0, 4, 17, 25, and 28. Cells from Passage 3 were subsequently replated at 250 000 cells per well onto a gelatin-coated 6-well plate and used for the Reprogramming (Rep) experiment (Diff Day 31 = Rep Day −1).

Reprogramming of hESC-derived fibroblasts into hiPSCs using Sendai virus. To initiate reprogramming (Rep Day 0), fibroblast-like cells differentiated from hESCs were infected with nonintegrating Sendai viruses expressing OCT4, SOX2, KLF4, and C-MYC (CytoTune-iPS 2.0, Thermo Fisher) according to manufacturer's recommendations. Four days after Sendai virus infection (Rep Day 4), cells were dissociated using Accutase, and replated in triplicate wells in Fib medium at 100 000 cells per well onto plates coated with hESC-Qualified Matrigel (Corning). The Fib medium was replaced with ReproTeSR medium (StemCell Technologies) the next day (Rep Day 5). Medium conditioned for at least 24 h was collected on Rep Day 1 (uninfected well only), Rep Day 3 (uninfected and infected wells), and Rep Days 6, 7, 11, 14, and 18 (triplicate from infected wells). On Rep Day 19, six human induced pluripotent cell (hiPSC) colonies were manually picked and grown in TeSR-E8 medium (StemCell Technologies). Two clonal cell lines were successfully expanded for further characterization.

Luciferase Detection *in Vitro*. Mouse 4T1 cells dually labeled for MiRAGE and luciferase expression were mixed with unlabeled cells at varying dilutions. After 24 h, cell lysate from seeded cells were collected and luciferase catalyzed using the Luciferase Assay System (Promega, E1500). Luciferase was detected in a 96-well white-walled transparent bottom plate (Corning) using a M1000 Plate Reader (Tecan).

Multiplexing Test. For testing crosstalk, 3×10^5 ESI-SOX2-RFP-MiRAGE(miR-458)/TRE-GFP-MiRAGE (miR-2) cells were seeded in a 24-well dish (Ibidi). Cells were cultured in the absence or presence of doxycycline (100 ng/mL) and with varying doses of BMP (0, 1, 3, 10, 30, 100 ng/mL). After 4 days,

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

cells were imaged, and RNA was collected from the media as described above.

RNA Collection. Cellular samples were first stored in RNA*later* Stabilization Solution (Ambion) for 1−2 days at 4 °C before processing for RNA using the *mir*Vana PARIS RNA and Native Protein Purification Kit (Ambion). Media samples were first centrifuged at 500*g* for 4 min to sediment cellular debris, and the supernatant was harvested. RNA was extracted from the supernatant media using the *mir*Vana PARIS kit, following manufacturers protocol. RNA from IP fluid and plasma was harvested using the same kit but with the initial centrifugation step modified to 3500*g* for IP fluid and plasma.

Extracellular Vesicle (EV) Purification. Media samples were centrifuged at 3000*g* for 15 min to remove cellular debris while preserving exosomes in the supernatant. EVs were isolated using ExoQuick TC EV Isolation Kit (SBI Biosciences, Palo Alto CA) and miRNAs were extracted from EVs using *mir*Vana kit (Thermo Fisher Scientific Carlsbad CA) according to the manufacturers' instructions.

qRT-PCR. For miRNA measurements, total RNA isolated from cell lines, supernatants, and plasma samples was reverse-transcribed to cDNA using TaqMan MicroRNA Reverse Transcription kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. In brief, 10 ng of total RNA were reverse-transcribed in a total volume reaction of 15 μL containing dNTPs, MultiScribe reverse transcriptase, reverse transcription buffer, RNase inhibitor, specific miRNA primer (TaqMan MicroRNA Assays, Applied Biosystems), and nuclease-free water. The reaction was performed using the DNA Engine Peltier Thermal Cycler (BIO-RAD) at 16 °C for 30 min, 42 °C for 30 min, and 85 °C for 5 min. For mRNA measurements, total RNA isolated from cell lines was reverse-transcribed to cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. In brief, 400 ng of total RNA were reverse-transcribed in a total volume reaction of 20 μL containing reverse transcription buffer 10×, dNTP 25×, MultiScribe reverse transcriptase 50 U/μL, RNA inhibitor 40 U/μL, random primer 10×. The reaction was performed using the DNA Engine Peltier Thermal Cycler (BIO-RAD) at 25 °C 10 min, 37 °C for 120 min, and 85 °C for 5 min.

The levels of miRNAs and mRNAs were measured by quantitative reverse transcription-polymerase chain reaction (qRT-PCR) using TaqMan MicroRNA Assays and TaqMan Gene Expression Assays (Applied Biosystems) according to the manufacturer's instructions. Reactions were performed using CFX384 Real-Time System (BIO-RAD) at 95 °C for 3 min and 40 cycles at 95 °C for 5 s and 60 °C for 30 s.

Mice. All animal experiments were conducted in accordance with American Association for Laboratory Animal Science regulations and the approval of The University of Texas MD Anderson Cancer Center Institutional Animal Care and Use Committee. HEYA8 cells expressing GFP, luciferase, and either MiRAGE-MiR2MiRAGE MiR458 were selected by sorting for GFP expression using a FACS Aria II instrument (BD Bioscience). The cells were resuspended in phosphate-buffered saline (PBS) solution and 5 × 10^6 cells in 100 μL PBS were injected *via* IP into the peritoneum of Nu/Nu mice. The *in vivo* imaging of tumors was performed using a Xenogen IVIS 100 optical *in vivo* imaging system at week 2 and week 4 after injection of the cells.

Imaging. All images were acquired from cells in imaging quality plasticware (Ibidi) using a 20×, Na 0.75 objective on Olympus IX83 inverted confocal microscope. Images were organized using ImageJ[22] and analyzed using ilastik and custom software written in MATLAB (Mathworks).

Plasmid Building/Cloning. Plasmids used in this study were cloned using conventional restriction digest methods or Gibson assembly. The plasmids were integrated into the genomes of cells either through ePiggyBac-based transposable elements, or through CRISPR-Cas9 based site-specific integration. The plasmids used were as follows:

1. epb-tt-puro-synthetic intron-shOCT4-RFP (Plasmid bearing synthetic intron, a gift of Ali Brivanlou[11]).
2. epb-tt-puro-GFP-MiRAGE (to express MiRAGE and GFP in response to doxycycline) was created from the above plasmid by replacing the small hairpin OCT4-targetting shRNA with a MiRAGE reporter cassette. Finally, the RFP was replaced with a PCR amplified GFP gene fragment.
3. epb-ubc-GFP-cag-fluc barcode (to express GFP, firefly luciferase and MiRAGE) was created by replacing the tet-responsive promoter in the above-described plasmid with a CAG promoter. Subsequently, a gene segment coding for firefly luciferase and an N-terminus t2a peptide was inserted between the GFP and MiRAGE segments.
4. Sox2HD-rfp-blas-barcode (Plasmid containing insert for knock-in *via* homologous recombination to create SOX2 reporter cell line) was created by the Gibson assembly of RFP (from Plasmid 1), MiRAGE barcode (from plasmid 3), blasticidin and backbone (from plasmids previously described[16]). Assembly was performed using HiFi DNA Assembly master mix (NEB), following manufacturers protocol. The cloning free homology donor was created by using two rounds of PCR to introduce the homology donors into the inset DNA *via* primer overhangs (italicized).

   Forward primer: *CGGCACGGCCATTAACGG-CACACTGCCCCTCTCACACA*TGGAGGGCAG-AGGAAGTCTTCT, and reverse primer: *CTTTGA-AAATTTCTCCCCCCTCCAGTTCGCTGTCCG-GCCCTTA*GGTGGCGACCGGTGGAT.

   For the second round of PCR, forward primers: *TCACATGTCCCAGCACTACCAGAGCGGCCC-GGTGCCCGGCAC*CGGCACGGCCATTAACGG, and reverse primer: *AAAAACGAGGGAAATGGGAGGGGT-GCAAAAGA*TGGAGGGGGGAGAAATTTTCAAAG.
5. Px330 SOX2 gRNA (Guide RNA plasmid to create SOX2 reporter cell line) with the sequence CCGGAC-AGCGAACTGGAGGG was cloned by inserting and duplexing an oligo pair to create the double stranded guide RNA with staggered overhangs compatible with *BbsI* restriction cut site. The duplexed oligo was then inserted into the PX330 plasmid (Addgene Plasmid #42230) bearing Cas9 and scaffolding RNA.

Tide Analysis. Tide analysis was performed to compare SOX2 loci of ESI017 MiRAGE reporter cells (ESI-SOX2M) to unmodified ESI017 by collecting genomic DNA using the DNeasy Blood and Tissue Kit (Qiagen). Briefly, the unedited allele was PCR amplified with primers targeting the unmodified Sox2 gene on either end of the CRISPR-Cas9 cut site (Guide RNA: CCGGACAGCGAACTGGAGGG; Forward primer: GTGGTTACCTCTTCCTCCCACT; Reverse primer: AAGTTTTCTTGTCGGCATCG). For the WT sample, DNA Clean and Concentrator (Zymo Research) was used to

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

purify the PCR product. For the CRISPR sample, the PCR product representing the unmodified allele was gel extracted from a 0.7% agarose gel using the Zymoclean Gel DNA Recovery Kit (Zymo Research). The purified products were sequenced (LoneStar Laboratories; GGGACATGATCA-GCATGTATC) and analyzed using the free TIDE web tool.[23]

Immunostaining. 35 mm imaging dishes (Ibidi) were rinsed with PBS, fixed for 20 min using 4% PFA, rinsed twice with PBS, and blocked for 30 min at room temperature. The blocking solution contained 3% donkey serum and 0.1% Triton-X in 1× PBS. After blocking, the cells were incubated with primary antibodies at 4 °C overnight. The next day the cells were washed three times with PBST (1× PBS with 0.1% Tween20) and incubated with secondary antibodies (AlexaFluor488 (Abcam) cat#A21206 or AlexaFluor647 (Life Technologies) cat#A31573) and DAPI dye for 30 min at room temperature. After secondary antibody treatment, samples were washed twice in PBST at room temperature. Primary antibodies were used against Sox2 D6D9 XP (Cell Signaling Technologies 5024S; dilution 1:200), (BD Biosciences 611203, dilution 1:400), NANOG (R&D Systems AF1997, dilution 1:200), E-Cadherin (Cell Signaling Technologies 31955, dilution 1:200), and Vimentin (Cell Signaling Technologies 5741T, dilution 1:100).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00406.

Figures S1−S7 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: aryeh.warmflash@rice.edu.

ORCID

Aryeh Warmflash: 0000-0002-5679-2268

Present Address

[∇]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal.

Author Contributions

[#]KM and WNF contributed equally to this work.

Notes

The authors declare the following competing financial interest(s): KM and AW have submitted a patent application based on the technology described in this paper.

■ ACKNOWLEDGMENTS

The authors thank Ali Brivanlou for the plasmid ePB-tta-puro-synthetic intron-shOCT4-RFP with the synthetic intron, and Alessandro Rosa for advice on using the synthetic intron. We also thank the flow cytometry core at MD Anderson, the Human Stem Cell Core at Baylor College of Medicine, and the members of the AW lab for helpful discussions and feedback. This work was funded by grants to AW from the Cancer Prevention and Research Institute of Texas (CPRIT grant number RR140073) and the National Science Foundation (NSF grant MCB-1553228). Work in Dr. Calin's laboratory is supported by National Institutes of Health (NIH/NCATS) grant UH3TR00943-01 through the NIH Common Fund, Office of Strategic Coordination (OSC), the NIH/NCI grant 1 R01 CA182905-01, a U54 grant, UPR/MDACC Partnership for Excellence in Cancer Research 2016 Pilot Project, a Team DOD (CA160445P1) grant, a Ladies Leukemia League grant, a CLL Moonshot Flagship project, a SINF 2017 grant, and the Estate of C. G. Johnson, Jr.

■ REFERENCES

(1) Davis, R. P., Ng, E. S., Costa, M., Mossman, A. K., Sourris, K., Elefanty, A. G., and Stanley, E. G. (2008) Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak−like cells and enables isolation of primitive hematopoietic precursors. *Blood 111*, 1876−1884.

(2) Wurdinger, T., Badr, C., Pike, L., De Kleine, R., Weissleder, R., Breakefield, X. O., and Tannous, B. A. (2008) A secreted luciferase for ex vivo monitoring of in vivo processes. *Nat. Methods 5*, 171−173.

(3) Keu, K. V., Witney, T. H., Yaghoubi, S., Rosenberg, J., Kurien, A., Magnusson, R., Williams, J., Habte, F., Wagner, J. R., and Forman, S. (2017) Reporter gene imaging of targeted T cell immunotherapy in recurrent glioma. *Sci. Transl. Med. 9*, eaag2196.

(4) Sun, N., Lee, A., and Wu, J. C. (2009) Long term non-invasive imaging of embryonic stem cells using reporter genes. *Nat. Protoc. 4*, 1192.

(5) Wan, J. C., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., Pacey, S., Baird, R., and Rosenfeld, N. (2017) Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat. Rev. Cancer 17*, 223.

(6) Ronald, J. A., D'Souza, A. L., Chuang, H.-Y., and Gambhir, S. S. (2016) Artificial MicroRNAs as Novel Secreted Reporters for Cell Monitoring in Living Subjects. *PLoS One 11*, e0159369.

(7) Warren, A. D., Kwong, G. A., Wood, D. K., Lin, K. Y., and Bhatia, S. N. (2014) Point-of-care diagnostics for noncommunicable diseases using synthetic urinary biomarkers and paper microfluidics. *Proc. Natl. Acad. Sci. U. S. A. 111*, 3671−3676.

(8) Arroyo, J. D., Chevillet, J. R., Kroh, E. M., Ruf, I. K., Pritchard, C. C., Gibson, D. F., Mitchell, P. S., Bennett, C. F., Pogosova-Agadjanyan, E. L., and Stirewalt, D. L. (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. U. S. A. 108*, 5003−5008.

(9) Kosaka, N., Iguchi, H., Yoshioka, Y., Takeshita, F., Matsuki, Y., and Ochiya, T. (2010) Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J. Biol. Chem. 285*, 17442−17452.

(10) Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D. Y., Barbosa, I. A., Kwon, J. S., and Guan, Y. (2013) An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep. 5*, 1704−1713.

(11) Arduini, B. L., and Brivanlou, A. H. (2012) Modulation of FOXD3 activity in human embryonic stem cells directs pluripotency and paraxial mesoderm fates. *Stem Cells 30*, 2188−2198.

(12) Villarroya-Beltri, C., Gutiérrez-Vázquez, C., Sánchez-Cabo, F., Pérez-Hernández, D., Vázquez, J., Martin-Cofreces, N., Martinez-Herrera, D. J., Pascual-Montano, A., Mittelbrunn, M., and Sánchez-Madrid, F. (2013) Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nat. Commun. 4*, 2980.

(13) Squadrito, M. L., Baer, C., Burdet, F., Maderna, C., Gilfillan, G. D., Lyle, R., Ibberson, M., and De Palma, M. (2014) Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells. *Cell Rep. 8*, 1432−1446.

(14) Badr, C. E., and Tannous, B. A. (2011) Bioluminescence imaging: progress and applications. *Trends Biotechnol. 29*, 624−633.

(15) Arbab, M., Srinivasan, S., Hashimoto, T., Geijsen, N., and Sherwood, R. I. (2015) Cloning-free CRISPR. *Stem Cell Rep. 5*, 908−917.

(16) Nemashkalo, A., Ruzo, A., Heemskerk, I., and Warmflash, A. (2017) Morphogen and community effects determine cell fates in response to BMP4 signaling in human embryonic stem cells. *Development 144*, 3042−3053.

(17) Horii, M., Li, Y., Wakeland, A. K., Pizzo, D. P., Nelson, K. K., Sabatini, K., Laurent, L. C., Liu, Y., and Parast, M. M. (2016) Human pluripotent stem cells as a model of trophoblast differentiation in both

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33

normal development and disease. *Proc. Natl. Acad. Sci. U. S. A. 113*, E3882−E3891.
(18) Mestdagh, P., Van Vlierberghe, P., De Weer, A., Muth, D., Westermann, F., Speleman, F., and Vandesompele, J. (2009) A novel and universal method for microRNA RT-qPCR data normalization. *Genome biology 10*, R64.
(19) Araki, R., Uda, M., Hoki, Y., Sunayama, M., Nakamura, M., Ando, S., Sugiura, M., Ideno, H., Shimada, A., and Nifuji, A. (2013) Negligible immunogenicity of terminally differentiated cells derived from induced pluripotent or embryonic stem cells. *Nature 494*, 100.
(20) Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009) Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn. Acad., Ser. B 85*, 348−362.
(21) Jin, J., Vaud, S., Zhelkovsky, A. M., Posfai, J., and McReynolds, L. A. (2016) Sensitive and specific miRNA detection method using SplintR Ligase. *Nucleic Acids Res. 44*, e116−e116.
(22) Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., and Schmid, B. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods 9*, 676.
(23) Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res. 42*, e168−e168.

DOI: 10.1021/acssynbio.8b00406
ACS Synth. Biol. 2019, 8, 25−33