Published online 2 August 2019

Nucleic Acids Research, 2019, Vol. 47, No. 18 ***e106***
doi: 10.1093/nar/gkz623

A 'poly-transfection' method for rapid, one-pot characterization and optimization of genetic systems

Jeremy J. Gam [1,2], Breanna DiAndreth [1,2], Ross D. Jones [1,2], Jin Huh [1,2] and Ron Weiss[1,2,*]

[1]Department of Biological Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA and [2]Synthetic Biology Center, MIT, Cambridge, MA 02139, USA

Received January 06, 2019; Revised June 24, 2019; Editorial Decision July 08, 2019; Accepted July 29, 2019

ABSTRACT

Biological research is relying on increasingly complex genetic systems and circuits to perform sophisticated operations in living cells. Performing these operations often requires simultaneous delivery of many genes, and optimizing the stoichiometry of these genes can yield drastic improvements in performance. However, sufficiently sampling the large design space of gene expression stoichiometries in mammalian cells using current methods is cumbersome, complex, or expensive. We present a 'poly-transfection' method as a simple yet high-throughput alternative that enables comprehensive evaluation of genetic systems in a single, readily-prepared transfection sample. Each cell in a poly-transfection represents an independent measurement at a distinct gene expression stoichiometry, fully leveraging the single-cell nature of transfection experiments. We first benchmark poly-transfection against co-transfection, showing that titration curves for commonly-used regulators agree between the two methods. We then use poly-transfections to efficiently generate new insights, for example in CRISPRa and synthetic miRNA systems. Finally, we use poly-transfection to rapidly engineer a difficult-to-optimize miRNA-based cell classifier for discriminating cancerous cells. One-pot evaluation enabled by poly-transfection accelerates and simplifies the design of genetic systems, providing a new high-information strategy for interrogating biology.

INTRODUCTION

Biologists frequently introduce exogenous DNA, RNA, or proteins into cells to modify cellular signaling or behavior. In many cases, multiple genetic elements (e.g. plasmids) must be delivered simultaneously. For instance, simultaneous expression of several transcription factors has been used to reprogram cell fate (1,2), and both guide RNA (gRNA) and Cas9 are needed for targeted nuclease activity (3). Often, the relative ratios of these elements are important for function; for example, specific stoichiometries of reprogramming transcription factors improve reprogramming efficiency by several fold (4,5). However, ratios are often chosen based on intuition, trial and error, or coarse optimization, which can hinder realization of the desired phenotype. For instance, popular plasmids for CRISPR/Cas9 (e.g. pX330 (3)) encode constitutive high expression of both gRNA and Cas9, but recent results show significant gains in genome editing efficiency after optimizing the ratio of gRNA to Cas9 (6,7). Moreover, the likelihood of using an unoptimal stoichiometry increases exponentially with each additional element in the system (8). Several methods have been prototyped to address this problem, but these involve complex and expensive pooled experiments (9–11), microfluidics (8,12), or time-consuming manual approaches (13). Pooled experiments link phenotype to genotype in a high throughput manner in mammalian cells, but include many steps that each require significant expertise, including: design of pooled DNA, assembly into a uniform library, delivery of the library into cells, and preparation and analysis of high throughput sequencing experiments (14,15). Microfluidic devices can combinatorially combine different elements in droplets, but require dedicated clean rooms, equipment, and expertise to fabricate, and are often difficult to reproduce (16). More commonly, coarse optimization is achieved by manually varying stoichiometry in many different independent co-transfection samples. This manual approach is often tedious for smaller systems and infeasible for larger systems such as those being developed in synthetic biology. Thus, there is a need for a simple and efficient method for rapidly optimizing the stiochiometry of different elements in genetic systems.

Here, we introduce a one-pot 'poly-transfection' method, which enables the evaluation of a given genetic system across a wide range of genetic element stoichiometries, all in a single sample. Notably, poly-transfection requires only a very simple change to a typical co-transfection proto-

*To whom correspondence should be addressed. Tel: +1 617 324 7555; Email: rweiss@mit.edu

col. In a co-transfection, plasmids are mixed before the transfection reagent is added, which results in highly correlated delivery of the plasmids to cells. Therefore, even though co-transfection results in single-cell expression that varies over several orders of magnitude (Figure 1A), only a single stoichiometry can be evaluated for the genetic system across that broad range (Figure 1B) (17,18). For poly-transfections, we instead mix each plasmid separately with transfection reagent before adding them to cells, resulting in decorrelated delivery of each plasmid (Figure 1C) (17,18). Also, rather than including a single transfection marker for an entire co-transfection sample,(19) we include a distinct transfection marker with each plasmid, such that each fluorescence color intensity serves as a proxy for the concentration of a specific plasmid. When combined with single-cell analysis methods such as flow cytometry, each poly-transfected cell provides an independent measurement of how the system behaves at a specific combination of plasmid concentrations. Therefore, poly-transfections take advantage of broad transfection distributions in order to widely sample various plasmid stoichiometries.

In addition to providing more informative data, poly-transfections simplify experimental planning and execution down to a single sample, resulting in significant savings in active experiment time (Supplementary Figure S1). We have demonstrated poly-transfections in systems containing up to four different plasmid-encoded elements, and depending on the experiment up to seven would be practical (Supplementary Figure S1), enabling efficient study and optimization of many current and future genetic systems (20). While in this study we focus on effects of stoichiometry of DNA plasmids, poly-transfections can likely also be applied to the delivery of other types of molecules like RNA or proteins.

MATERIALS AND METHODS

Detailed experimental methods and guidelines for designing poly-transfection experiments may be found on protocols.io under 'One-pot Optimization of Genetic Circuits using Poly-transfections' (dx.doi.org/10.17504/protocols.io.k98cz9w) and a quickstart guide is provided in Supplementary Table S1.

DNA assembly framework

The plasmids used in this study were assembled by a hierarchical Golden Gate-Gibson assembly method similar to previous hierarchical methods used in our lab (21). First, Golden Gate assembly was used to assemble plasmids for each transcription unit (termed plasmid level 1 or pL1 plasmids) from a number of basic input parts (termed plasmid level 0 or pL0) including insulators (pL0-I), promoters (pL0-P), 5' UTRs (pL0-5), gene coding sequences (pL0-G), 3' UTRs (pL0-3), and transcriptional terminators/poly-A sequences (pL0-T). Then plasmids containing multiple transcription units (termed pL2) were constructed from the assembled pL1 plasmids using Gibson assembly. For the Gibson step, plasmids were linearized using I-SceI endonuclease to generate overlapping regions, followed by assembly with Gibson Assembly Ultra Kit from SGI and electroporation into electro-competent cells. Typical Golden Gate and Gibson reaction protocols are given in Protocol Exchange. As a simplified alternative to hierarchical assembly, we have constructed plasmids that utilize a single Golden Gate step to obtain plasmids encoding two transcription units - one expressing a gene of interest and the other expressing a fluorescent protein marker (Supplementary Figure S2). Since the backbone vector already encodes fluorescent protein expression, Gibson assembly would no longer be required. DNA sequences comprising the simplified DNA assembly framework can be found in a Supplementary Data 1 (GenBank file) and are available to researchers via Addgene (#109150–109154). Full sequence information for other plasmids used in one-pot experiments can also be found in the Supplementary Data 1 (GenBank file).

Cell culture

The HEK293FT cell line was obtained from Thermo Fisher and HeLa cells were obtained from ATCC. Both cell lines were maintained in DMEM containing 4.5 g/l glucose, L-glutamine, and sodium pyruvate (Cellgro) which was supplemented with 10% FBS (Cellgro) and were grown at 37°C and 5% CO_2. All lines used in this study tested negative for mycoplasma.

Transfections

Unless indicated otherwise, transfections were conducted using Lipofectamine 3000 (Thermo Fisher) and were performed according to the manufacturer's dilution protocols with modifications for poly-transfections. For poly-transfections each DNA-lipid mixture was prepared separately before being applied to cells. As an example, to perform an experiment with two plasmids in a single well in a 24-well plate, plasmid 1 was mixed in one tube with 0.5 ul of transfection reagent, plasmid 2 was mixed with 0.5 ul of transfection reagent in a separate tube, complexes were allowed to form, then both complex species were added simultaneously to the same well containing cells (Figure 1C). In contrast, for co-transfections, DNA was first mixed together at the specified ratios and only then added to the transfection reagent (Figure 1B). We used 100-300 ng of DNA per unique plasmid with 1 ul of P3000 reagent and 1 ul of Lipofectamine 3000, all prepared in 100 ul Opti-MEM per 24-well plate. In the 24-well plate format, 200 000 HEK293FT cells or 150 000 HeLa cells were transfected per well. The amount of DNA and transfection reagent per plasmid species was scaled down according to the total number of unique plasmids in order to maintain similar total amounts compared to traditional transfections. For a quick reference for setting up poly-transfections, see Supplementary Table S1. Transfection preparation information for each experiment is specified in Supplementary Data 2 (spreadsheet). We note several examples where we used poly-transfections to characterize genetic systems (Figures 2 and 3, Supplementary Figures S10–S17, Supplementary Notes 2 and 3).

Reliability of poly-transfections across different transfection reagents and protocols was assessed by repeating poly-transfections using the same amount of DNA and the corresponding amount of transfection reagent recommended

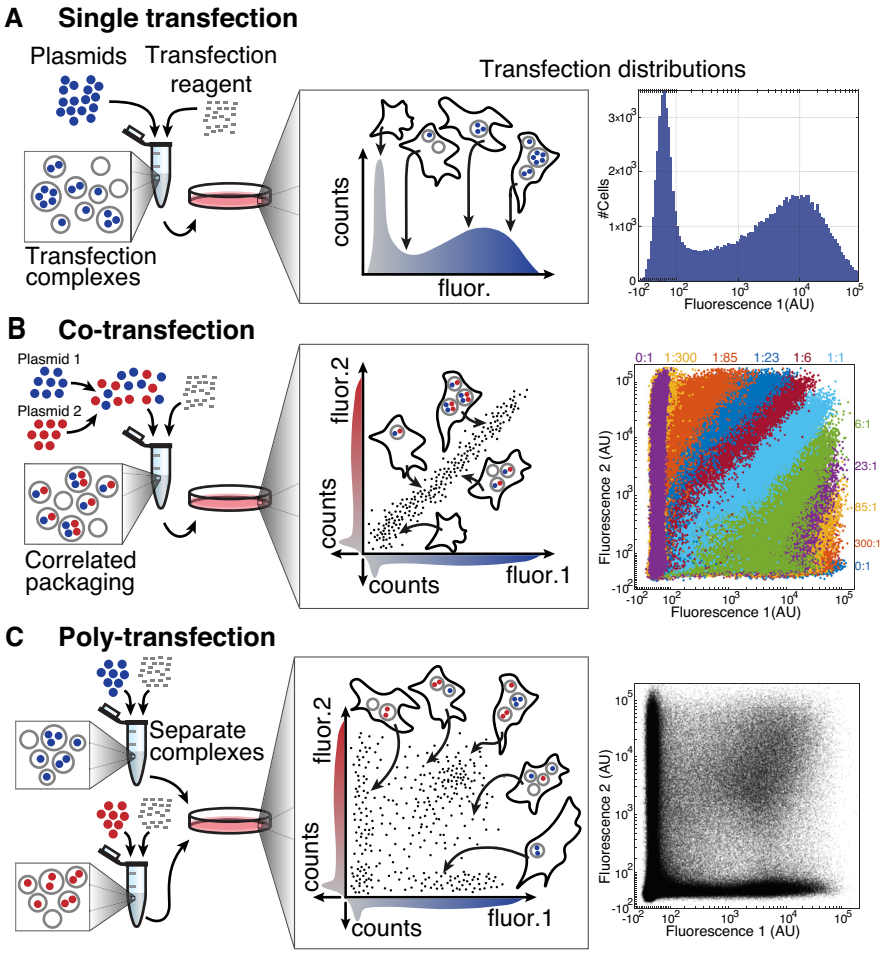

Figure 1. Overview and comparison of plasmid delivery with a single transfection, co-transfection of two plasmids, or poly-transfection of two plasmids. For each transfection method, the leftmost diagram shows formation of transfection complexes between negatively charged DNA and positively charged lipid. In these examples, each colored plasmid (blue and red) encodes expression of a different fluorescent protein. The center diagram shows examples of plasmid delivery to cells and also a schematic for the expected distributions in a histogram or scatter plot. Color intensity on the histogram corresponds to fluorescence from the corresponding plasmid color. The rightmost diagram shows real data from cells transfected using each given method. (**A**) In a transfection with a single plasmid species, plasmids are mixed with transfection reagent to form transfection complexes containing varying numbers of plasmids (left). In the schematic, transfected cells uptake varying numbers of complexes, with each complex having different numbers of plasmids, resulting in a broad distribution of transfected cells ranging from untransfected to very highly transfected (center). In actual single-plasmid transfection data, cells express a wide range of the encoded fluorescent protein. (**B**) In a co-transfection with two different plasmids, both plasmids are mixed together before adding to transfection reagent, resulting in highly correlated packaging of the two plasmid species. While the number of total plasmids per cell varies significantly, the fraction of plasmids that are either species is relatively constant with the major expected source of deviation being stochastic noise (left) (35). Transfected cells may uptake few or many complexes but since each complex has correlated amounts of each plasmid, the co-transfection only explores a small diagonal region of the concentration space between the two plasmids (center). In actual co-transfection data, cells exhibit correlated delivery of both plasmids. Colors indicate the different co-transfections and plasmid ratios used in each co-transfection (right). (**C**) In a poly-transfection, each plasmid species is mixed with transfection reagent separately, resulting in complexes that contain only a single plasmid species (left). Transfected cells uptake different combinations of the complexes, resulting in cells that contain neither, both, or a single plasmid species, and all at varying total plasmid amounts (center). In actual poly-transfection data, cells explore a wide range of the concentration space with many different plasmid stoichiometries explored simultaneously. A poly-transfection explores a similar range of fluorescence combinations compared to co-transfections (two dimensions ranging from -10^2 to 10^5 AU), but requires only the preparation of a single sample to obtain >90% coverage of concentration space compared to nine different samples with co-transfection (Supplementary Figure S1). With three plasmids, poly-transfections still require only a single sample, compared to as many as $11 \times 11 = 121$ co-transfection samples.

by the manufacturer (Supplementary Figures S3–S6). For miRNA classifier co-transfections, the 27 different ratios used are listed in Supplementary Table S2.

Flow cytometry

Flow cytometery was performed using a BD LSRFortessa equipped with a 355 nm laser with 379/28 nm filter at 410 PMT voltage for measuring fluorescence from Sirius, 405 nm laser with 450/50 nm filter at 200 V for measuring TagBFP, 488 laser with 530/30 filter at 180 V for measuring mNeonGreen, 561 nm laser with 582/15 nm filter at 210 V for measuring mKO2, and 640 laser with 710/50 nm filter at 320 V for measuring iRFP720. For co-transfections >50 000 cells were collected per sample, while for poly-transfections >200 000 cells were collected for 2-plasmid

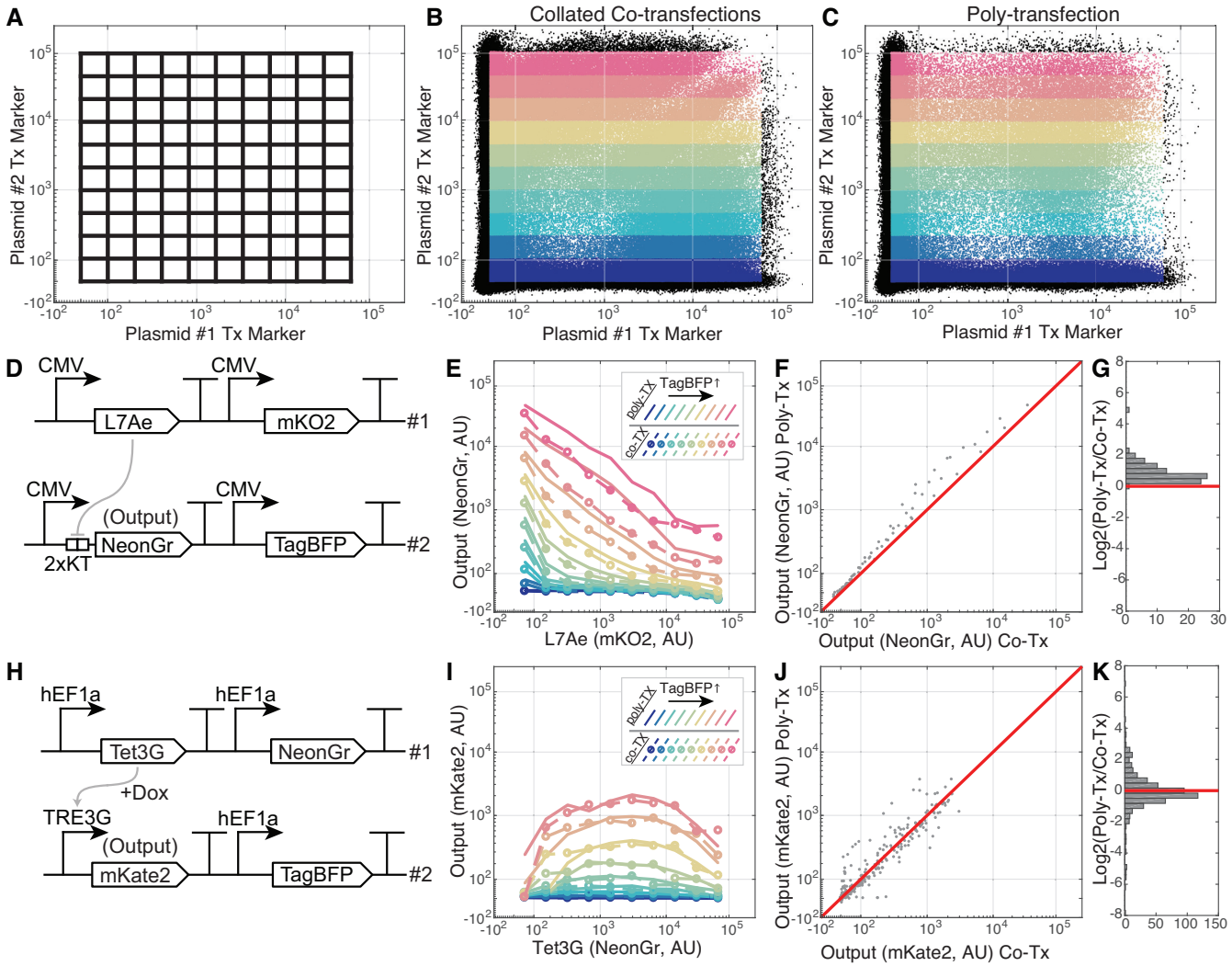

Figure 2. Benchmarking poly-transfection against co-transfection. (**A–C**) Binning workflow for bench-marking analysis. Ten biexponentially-spaced bins were assigned for each transfection marker dimension which approximate the levels of each of the two plasmids (A). Binning was performed on both a collated set of 11 co-TX samples spanning various plasmid ratios (B) and data from a single poly-TX (C). (B) and (C) each show ∼500 000 cells, representing the cumulative data from all 11 co-TX samples in (B), and all cells collected from the single poly-TX sample in (c). Colors correspond to sets of bins defined by gene 2 (TagBFP) levels. (D–K) Median output fluorescence was evaluated for the cells in each bin and compared between methods for each system: L7Ae translational repression (D–G) and TetG transcriptional activation (H–K). (**D**) Constructs for measuring L7Ae activity. mKO2 fluorescence serves as an estimate for delivery of L7Ae (gene #1), while TagBFP fluorescence serves as an estimate for delivery of the regulated mNeonGreen output (gene #2). (**E**) Multi-dimensional titration curves for L7Ae. Each line represents the set of bins at one level of TagBFP as denoted in (B) and (C). Solid lines denote poly-TX data while dashed lines denote co-TX data. (**F**) Visual comparison between co-TX and poly-TX for L7Ae system. Each point represents the measured output in corresponding bins for poly-TX and co-TX measurements, where more equivalent values are closer to the red 1:1 line. (**G**) Histogram for discrepancy between poly-TX and co-TX for the L7Ae system. The values are the $\log_2$ fold-changes from the co-TX to poly-TX for each point in (F). Mean = 0.87, standard deviation = 0.65. A mean close to zero implies high accuracy while a standard deviation close to zero implies high precision. The general underestimation of output in poly-TX compared to co-TX is discussed in Supplementary Note 1. (**H**) Constructs for measuring Tet3G activity. mNeonGreen fluorescence serves as an estimate for the delivery of Tet3G (gene #1), while TagBFP fluorescence serves as an estimate for delivery of the mKate2 output (gene #2). (**I**) Multi-dimensional titration curves for Tet3G at the highest level of doxycycline (Dox = 2.3 μM). The highest bin level of TagBFP was excluded due to low cell counts in the co-TX samples. (**J**) Visual comparison between poly-TX and co-TX for Tet3G system (inclusive of all doxycycline levels). (**K**) Histogram for discrepancy between poly-TX and co-TX for Tet3g system, calculated as in (G). Mean = –0.09, standard deviation = 1.31. Overall, the differences observed between poly-TX and co-TX-derived are low which contribute to the good correlation observed in (F) and (J) and provide confidence in the reliability of the poly-TX method.

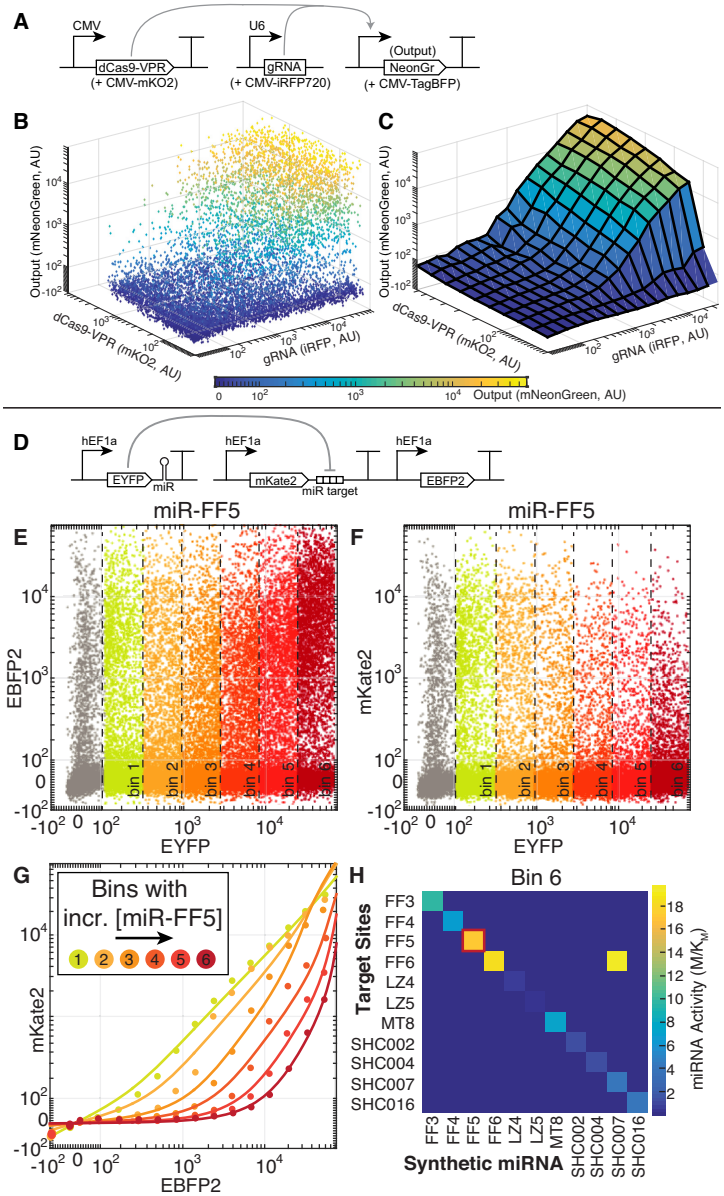

Figure 3. Novel characterizations enabled by poly-transfection. (**A–C**) Characterization of transcriptional activation of a fluorescent reporter by dCas9-VPR. Contributions to activation by dCas9-VPR and gRNA can be measured simultaneously with our one-pot method. Shown are the circuit diagram (A), subsampled scatter (B), and a surface plot (C) of output as a function of gRNA and dCas9-VPR at intermediate TagBFP transfection marker levels (10^3 to 10^4 AU). Surface plot indicates the medians of mNeonGreen within each bin and data at all TagBFP levels can be found in Supplementary Figure S24. (**D**) Genetic circuit for testing miRNA behavior. The circuit consists of two plasmids where the first expresses a synthetic miRNA based on a miR-155 expression platform and also EYFP marker serving as an indicator of miRNA expression. The second plasmid encodes a miRNA low sensor where mKate2 fluorescent protein contains miRNA target sites in the 3' UTR which mediate repression of mKate2 in the presence of the complementary miRNA. EBFP2 transfection marker indicates how much of this sensor plasmid is delivered to each cell. (**E**) Scatter plot of poly-transfection input parameter space for miR-FF5. A plot of miRNA concentration (∼EYFP) and sensor concentration (∼EBFP2) shows that most of the parameter space is covered by the poly-transfection. Data are colored according to EYFP fluorescence, with yellow to red marking the lowest to highest EYFP bins respectively. (**F**) Scatter plot of output as a function of miR-FF5 concentration. Relative to EBFP2 (E), mKate2 decreases when miR-FF5 concentration (∼EYFP) increases, demonstrating that miR-FF5 is able to repress the FF5 sensor. (**G**) Sensor performance at different miR-FF5 concentrations. Since sensors for endogenous miRNAs are often plotted with output as a function of transfection marker, we binned data in each of the EYFP bins further using EBFP2 and calculated the mKate2 medians (circles). Lines indicate fits to a miRNA repression model (29,30). As miRNA concentration increases, its ability to repress the sensor increases as indicated by further decreases in mKate relative to EBFP2. (**H**) Orthogonality of synthetic miRNAs and sensors. We generated orthogonality matrices for the six EYFP bins, where rows indicate which target sites were present on the sensor and columns indicate the synthetic miRNA introduced. Shown is the sixth bin, which corresponds to the highest synthetic miRNA expression. Colors indicate the miRNA activity as fit using a miRNA repression model. This allows us to compress the EBFP2 and mKate2 dimensions into a single measure for simplified visualization. The square corresponding to the miR-FF5 data displayed in (E–G) is outlined. Almost all miRNA/sensor pairs are highly orthogonal, as illustrated by miRNA activity along the diagonal. 11*11=121 poly-transfections were needed to generate the complete five-dimensional dataset. The entire set of six orthogonality matrices showing miRNA activity and orthogonality across a wide range of miRNA expression levels is shown in Supplementary Figure S10. For comparison, with co-transfection, only a single orthogonality matrix could be obtained for the same number of samples, and activity measurements would be biased since miRNA delivery would not be fixed with respect to sensor delivery since they are highly co-delivered.

experiments. For >2-plasmid experiments, roughly 10 times more cells were collected per plasmid species.

To determine an orthogonal set of fluorescent proteins, we separately transfected HEK293FT cells with 22 constitutively expressed fluorescent proteins and measured signal from all available channels on the flow cytometer (Supplementary Figure S7). Fluorescence score was calculated as log10(99th percentile channel fluorescence from sample)-log10(channel background from untransfected). We then determined that Sirius, TagBFP, mNeonGreen, mKO2 and iRFP720 formed a relatively orthogonal set.

Data analysis

We have developed a MATLAB pipeline to facilitate analysis of multi-dimensional poly-transfection data. Code may be found on Github at <https://github.com/Weiss-Lab/MATLAB_Flow_Analysis>. Generally, data was binned across fluorescence corresponding to relevant input parts and statistics (e.g. median) for the output calculated for each bin (Supplementary Figure S8). In some cases, data was 'sliced' to only analyze data with a certain level of transfection marker prior to binning (Supplementary Figure S8). Example analysis and associated code from this study may be found within the Github repository under <https://github.com/Weiss-Lab/Gam_OPO_Paper>.

For subsampling of poly-transfection data, a series of points sampling a ratiometric trajectory was determined and the distance between each cell and the trajectory determined. To subsample, a number of cells were selected for inclusion based on the distance between the cell and the trajectory according to a log-normal distribution. For our analysis, each length between points on the trajectory corresponded to selection of approximately 310, 160, 100, 50 and 20 cells at a distances of 0.1, 0.2, 0.3, 0.4, 0.5 (logicle transformed units). Also, uniform subsampling may be used to normalize poly-transfection data to reduce the effect of higher density of singly- and doubly-transfected cells (Supplementary Figure S9).

To visualize the stoichiometric effects for each plasmid species, data were binned according to fluorescence of each transfection marker and plotted as medians in 2D, 3D and 4D heatmaps (Figures 2 and 3, Supplementary Figure S10–S15). We also compared performance of 3-plasmid and 4-plasmid high sensors by using an ROC-like analysis, where each point within the ROC curve corresponds to the sensitivity and specificity calculated within a fluorescent bin (Supplementary Figure S16).

Statistical analysis

Statistical analysis for this study is based on log-normal distributions for production of fluorescent proteins in cells. Data collected in this study supported the assumption of geometric standard deviation of approximately 2.8. For co-transfections, >50 000 cells were collected such that at least 100 data points could be analyzed per transfection bin. For poly-transfections, 10 times more cells were collected per plasmid species to minimize sample collection time while still allowing at least 99% and 83% coverage of concentration space in 4- and 5-color experiments respectively.

Data availability

All data from this study are available from FlowRepository under repository IDs FR-FCM-ZYU3 through FR-FCM-ZYU7. DNA sequences and transfection information can be found in the supplementary, detailed experimental protocols are available from Protocol Exchange #6667, and general MATLAB scripts for poly-transfection analysis are accessible via github at https://github.com/Weiss-Lab/MATLAB_Flow_Analysis.

RESULTS

Benchmarking poly-transfection against co-transfection

We first determined whether a single poly-transfection can accurately recapitulate results from co-transfections. For this purpose we tested two regulatory modules: translational repression by RNA binding protein L7Ae and transcriptional activation by Tet3G (a variant of rtTA). Each of these systems has two constituent plasmid elements: one encoding the regulatory protein (L7Ae or Tet3G), and a second plasmid encoding a fluorescent reporter that is controlled by the regulatory protein (Figure 2D, H). We also encoded expression of a transfection marker for each plasmid using a different constitutive fluorescent protein on each of the two plasmids in order to track their concentrations. For a given two-plasmid poly-transfection, we found that 11 co-transfections were required to cover the two-dimensional concentration space, achieved by varying the ratio of these two plasmids in each co-transfection (Figure 2B and C). Next, data from these sets of co-transfections were collated to form a single dataset that could be compared to poly-transfection data. We then binned both the poly-transfection data and collated co-transfection data according to the two transfection markers and evaluated output reporter fluorescence in each bin (Figure 2E, I). Output fluorescence proved to be quite similar (Figure 2F, G, J, K) with 66% of all L7Ae values and 87% of all Tet3G values achieved by poly-transfection falling within 2-fold change of their associated co-transfection-derived values. For an explanation of binning and other analysis, see Supplementary Figure S8. Of note is a bias toward singly- and doubly-transfected cells in poly-transfections, which we discuss in Supplementary Note 1. Despite this bias, we obtained sufficient numbers of cells spanning the concentration space. We also note that expression of the transfection markers does not result in cellular burden with our flow cytometry-based measurements (Supplementary Note 1 and Supplementary Figure S25). Overall, the concordance between co- and poly-transfections supports the notion that a cell expressing a stoichiometry of transfection markers has been transfected with the corresponding stoichiometry of plasmids, regardless of whether the delivery was through a co- or poly-transfection.

New characterization and insights via poly-transfection

In addition to benchmarking poly-transfections and comparing to traditional co-transfections, we used poly-transfection to generate new insights and characterization for biological parts. We first determined the activation

strength for a dCas9-VPR/gRNA transcriptional activation (CRISPRa) system. CRISPRa systems have been used for various applications, though quantitative information about how activation strength depends on the stoichiometry of dCas9-activator and gRNA has not been explored in depth (22,23). Using a single poly-transfection, we were able to quantify how relative concentrations of a gRNA-expressing plasmid and a dCas9-VPR-expressing plasmid contribute to target gene activation. Interestingly, we found that increasing gRNA concentration contributes to significant increases in activation, whereas dCas9-VPR concentration does not greatly affect the degree of activation, as long as it is above a relatively low threshold (Figure 3A–C and Supplementary Figure S24). These results are consistent with our co-transfection experiments (Supplementary Figure S18) and CRISPR editing studies where efficiency of Cas9 editing depended more on gRNA concentration than Cas9 mRNA concentration (6,7). Therefore, activation efforts with CRISPRa and editing should benefit more from improved delivery of DNA coding for gRNA rather than Cas9.

We also used poly-transfection to examine the potency of synthetic miRNA sequences (24) in repressing a gene at the post-transcriptional level (25,26). Synthetic miRNA sequences have been recently leveraged to reduce leaky gene expression or generate more complex systems like feed-forward loops (27,28). We sought to enable these applications by quantifying both the potency of synthetic miRNAs as well as the orthogonality between miRNAs and the targeted genes, which would allow for more informed design choices. Plasmids expressing miRNA and target fluorescent proteins were each associated with a transfection marker (Figure 3D). As with other poly-transfections, a broad combinatorial space of combinations between the miRNA and miRNA sensors was explored (Figure 3E). As miRNA concentration increased, reporter fluorescence decreased more significantly compared to the transfection marker in a threshold-like manner as previously observed from similar sensor constructs (Figure 3F, G) (29,30). We then demonstrated that detailed information about miRNA orthogonality could be obtained from poly-transfection data, at a reduced number of transfections compared to co-transfections. We conducted poly-transfections using all possible combinations of eleven different miRNA and sensor sequences, resulting in $11 \times 11 = 121$ separate poly-transfections which were completed in less than two hours (Supplementary Figure S10). To obtain similar data using co-transfections would have required approximately $10 \times 11 \times 11 = 1210$ transfections, since around 10 co-transfections are required to adequately cover a 2D concentration space (Figure 1B). Analysis showed that synthetic miRNAs FF5 and FF6 have the strongest potency and all tested synthetic miRNAs have minimal cross-talk except for the miR-FF6 / miR-SHC007 pair, since both are derived from the same luciferase sequence. We were also able to easily extract key miRNA activity metrics for each synthetic miRNA (Figure 3H) (30). These examples show how CRISPRa and synthetic miRNA systems—and by extension, future yet uncharacterized systems - can be easily characterized using poly-transfection.

Optimizing a complete genetic system with poly-transfection

Next, we used poly-transfections to rapidly engineer a difficult-to-optimize genetic system for discriminating cancerous from noncancerous cells (26). Cancer cells often express the biomarker miRNA-21-5p highly; (31) thus, we built and optimized a cell classifier which produces a genetic output only in the presence of miR-21-5p (i.e. a single high miRNA classifier). Our classifier is composed of three DNA components and responds to miR-21-5p in the following way: (i) at high miR-21-5p activity, the miRNA degrades the transcript encoding the BM3R1 transcriptional repressor, (ii) degradation of BM3R1 allows transcription from the promoter driving mKO2 reporter, and (iii) a Gal4-VP16 transcriptional activator is needed for output expression (Figure 4A). Constraining this system to accurately produce output only in miR-21-5p expressing cells requires balancing the ratios of plasmid concentrations for BM3R1, Gal4-VP16, and reporter.

We optimized transfection-based classification of cells via the miR-21-5p classifier by using poly-transfection to identify an optimal DNA ratio of the three classifier plasmids that would specifically distinguish HeLa from HEK293FT cells, which have high and low miR-21-5p activity respectively. We evaluated classification accuracy at different plasmid ratios by computationally subsampling the 3D poly-transfection data, designating stoichiometric 'trajectories' through the concentration space (Figure 4B), and including only cells that were in close proximity to the trajectory (Supplementary Figure S19). Subsampling in this manner provided data similar to traditional co-transfections from which classification metrics including specificity, sensitivity, and accuracy could be calculated. Using an optimization method (Supplementary Note 3), we identified a well-performing ratio of 10.9:1.5:1 for Gal4-VP16:output:BM3R1 plasmids at which HeLa and HEK293FT cells were distinguished with 91% specificity, 62% sensitivity, and 77% accuracy (Figure 4C, Supplementary Figure S20). We verified this prediction by performing a co-transfection at the corresponding ratio of plasmids and found that the specificity, sensitivity, and accuracy metrics agreed with the values computed from the subsampled poly-transfection (99%, 68%, 84% respectively; Figure 4d). Further comparisons between poly- and co-transfections at 27 different DNA ratios representing titrations of each unique plasmid showed good correlation for specificity, sensitivity, and accuracy (Supplementary Figure S20). Analysis showed that high classification accuracy was possible with Gal4-VP16 expressed at a wide range of high levels, and with BM3R1 and output within a narrow band of lower concentrations (Figure 4E). Overall, these data indicate that subsampled data from single poly-transfection samples can be used to quickly evaluate system performance, reducing or eliminating the need to iterate through different physical designs (Supplementary Figure S20).

We then applied our findings to construct optimized single-plasmid versions of the miR-21-5p classifier (Figure 5A), which better approximates delivery conditions for therapeutics and other classifier applications like cell purification. Given that relative DNA copy numbers are fixed in a single-plasmid system, we instead tuned expression ratios of

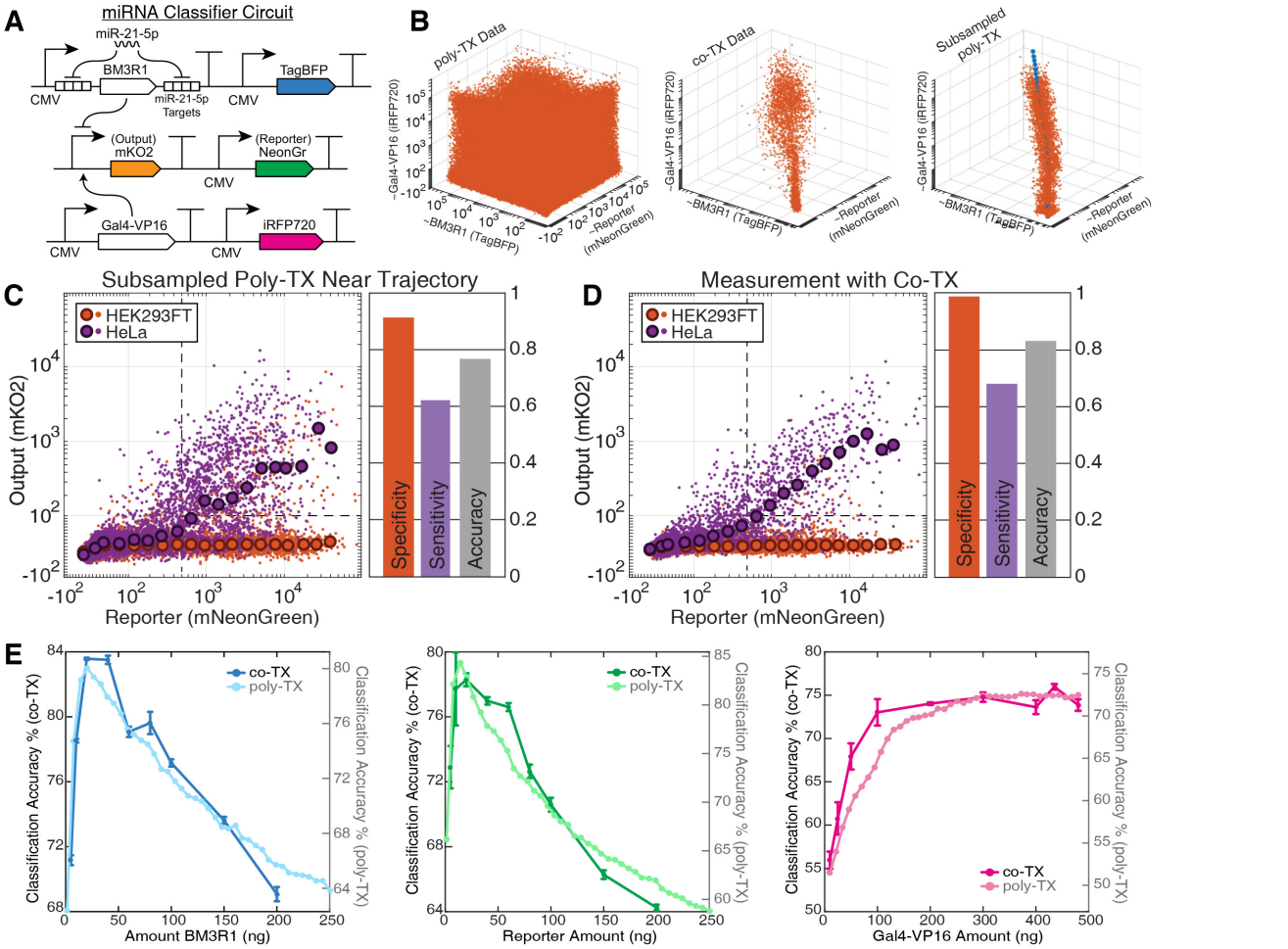

Figure 4. Optimization of a miRNA classifier circuit using poly-transfection. (**A**) miRNA classifier to be optimized. In cell lines where miR-21-5p activity is high (e.g. HeLa), BM3R1 is repressed by the miRNA, derepressing output mKO2 in the presence of Gal4-VP16. (**B**) Poly-transfection (poly-TX), co-transfection (co-TX), and subsampled poly-TX data in three dimensions representing concentration of each gene. Poly-TX data shows sufficient coverage in three input dimensions such that it can be subsampled according to a ratio trajectory (blue) to yield data similar to co-TX. For poly-transfection, at least 1.5 million cells were assayed. (**C**) Classification in HEK293FT and HeLa cells based on the subsampled poly-TX. Scatter diagram of mKO2 output is plotted as a function of reporter marker fluorescence (mNeonGreen). Vertical dotted lines indicate the threshold for cells designated as transfected (mNeonGreen = 5×10^2 AU) and horizontal lines indicate the threshold for determining if cells express high or low output (mKO2 = 10^2 AU). Poly-TX data was subsampled according to a trajectory corresponding to Gal4-VP16 = 435 ng of DNA, reporter = 60 ng, and BM3R1 = 40 ng. For HEK293FT cells (red), output remained low in most cells across all reporter levels, while in HeLa cells (purple) a majority of transfected cells expressed high output (specificity = 91%, sensitivity = 62%, accuracy = 77%). Larger outlined circles indicate mKO2 median data binned according to delivery of the mNeonGreen reporter plasmid. (**D**) Cell classification in a co-TX experiment at same stoichiometric ratio as in (C). Similar to the subsampled poly-TX, output was low in HEK293FT cells and high in HeLa cells (specificity = 99%, sensitivity = 68%, accuracy = 84%). (**E**) Classification accuracy as a function of plasmid amounts. For either co-TX or subsampled poly-TX data, accuracy showed similar dependency on BM3R1, reporter, and Gal4-VP16 amounts. DNA amounts correspond to those in (C) with one plasmid varied at a time.

each gene to the optimized levels identified above by inserting a repressive upstream open reading frame (uORF) into the 5'UTR of the mKO2 reporter and truncating the CMV promoter driving BM3R1 (Figure 5C, Supplementary Figures S21 and S22) (32,33). The miR-21-5p classifier optimized in this way showed higher classification accuracy in single-plasmid transfections compared to both an unoptimized classifier (high expression of all genes) and a poorly optimized classifier (low Gal4-VP16 expression via truncated CMV) (Figure 5C). Additionally, when we replaced the fluorescent output with the apoptosis regulator Bax, the optimized design showed greater ability to selectively induce apoptosis in HeLa cells compared to the other classifier variants (Figure 5C), and demonstrated similar selectivity compared to other more complex miRNA classifiers (26,34).

DISCUSSION

Poly-transfections bridge the gap between simple but low information co-transfections and complex high-information pooled approaches (Supplementary Figure S23). By taking advantage of innate variations in transfection efficiency, poly-transfections yield high-information data similar to modern single-cell analysis techniques, while also being substantially simpler. Specifically, the increased cost and complexity of FlowSeq methods (9) and

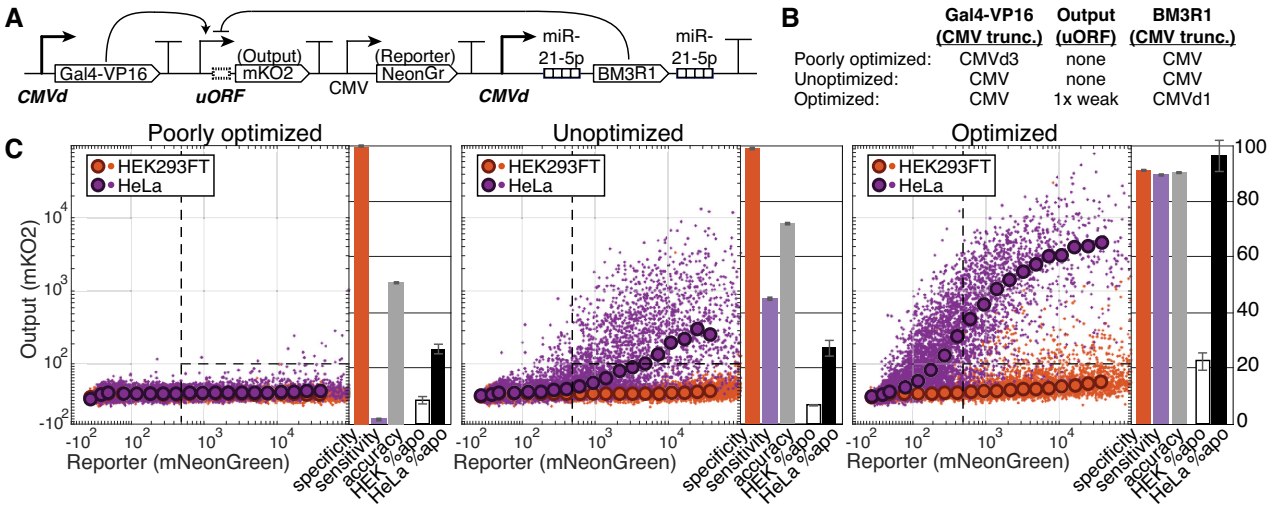

Figure 5. Evaluation of single-plasmid miRNA classifiers *in vitro*. (**A**) Circuit diagram for single-plasmid miRNA high sensors. Tunable parts include either upstream open reading frames (uORFs) or CMV truncations (CMVd), both of which tune down expression to a defined degree (Supplementary Figs. S21, S22) (32,33). (**B**) Table of CMV truncations and uORFs used in single-plasmid constructs. The optimized circuit encodes high expression of Gal4-VP16 and reduced levels of output and BM3R1, the unoptimized circuit encodes high expression of all three, while the poorly optimized circuit encodes low expression of Gal4-VP16 and high expression of output and BM3R1. (**C**) Performance of an optimized single-plasmid high sensor compared to unoptimized and poorly optimized variants. Scatter plots of binned and raw output (mKO2) as a function of transfection marker (mNeonGreen) in the poorly optimized high sensor (left) show low output in both HEK293FT and HeLa cells, resulting in high specificity (100%) but inferior sensitivity (2%) and classification accuracy (51%). The unoptimized high sensor (middle) shows low output in HEK293FT (specificity = 99%), but only high output in a fraction of HeLa cells (sensitivity = 45%), and an overall lowered ability to classify cells (accuracy = 72%). For the optimized high sensor (right), output remained low in most transfected HEK293FT cells (specificity = 91%) and high in most transfected HeLa cells (sensitivity = 90%), resulting in a generally high classification accuracy (accuracy = 90%). Results from the optimized classifier with Bax as output show much higher degree of killing in on-target HeLa cells (black) with low killing in off-target HEK293FT cells (white). Apoptosis percentages are relative to positive controls with constitutively expressed Bax. Error bars for bar charts indicate standard deviations for technical triplicates.

droplet microfluidics (8) makes them unsuitable for small- to medium-scale systems (2–4 genes) commonly used in biology. Additionally, poly-transfections are simpler to perform and scale drastically better than co-transfections, which require exponentially more samples for each unique plasmid within the system (Supplementary Figure S1). Furthermore, our poly-transfection method can be readily extended to newer single-cell analysis technologies such as spectral analyzers and mass cytometers (e.g. CyTOF) that enable even higher dimensional datasets than are possible with the 5-color flow cytometers used in this study.

In immortalized cell lines such as HEK293 and HeLa cells, both genetic and non-genetic variability contribute to noise in flow cytometry and other measurements. In general, noise causes cells receiving the same plasmid dosages to show divergent responses. For poly-transfections specifically, intrinsic noise could result in deviations in transfection marker fluorescence from that expected at a given plasmid dosage (i.e. fluctuation in plasmid delivery measurements). Our measurements of intrinsic noise for two fluorescent reporters on the same plasmid or co-transfected plasmids (Supplementary Figure S5) show transfection marker noise that is several orders of magnitude less than the variation caused by differential plasmid uptake during transfection. Noise also affects distributions of output fluorescence and is often measured to determine system robustness. Poly-transfections, like other single cell methods, allow for noise in the output to be measured as a function of the concentration space landscape (Figure 3B). It is also important to note that existing methods including co-transfection, microfluidics, or pooled experiments face the same factors for cell-to-cell variability.

Overall our poly-transfection method represents a convenient and powerful one-pot approach for evaluating effects of DNA stoichiometries for diverse experimental systems. In addition to Cas9 and classifier systems shown here, we anticipate that future insights gained from poly-transfections will help biologists extract deeper information from genomic perturbations, design reporters with higher signal to noise ratio, and build tools that more robustly modulate gene expression, among other applications. Therefore, in addition to poly-transfections being immediately applicable to fields such as synthetic biology where optimization of complex genetic systems is already essential, we envision that the ease of our method will make such one-pot strategies more attractive to a wide range of other fields within biology.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to acknowledge Brian Teague, Katherine Kiwimagi, Scott Olesen, Douglas Lauffenburger and Phillip Sharp for discussion; Brian Teague for manuscript editing; Patricia Rogers for spectral analyzer help; Olga Parkin and Cameron Haase-Pettingell for administrative support; and

Aishwarya Jagtap for technical support. Sirius/pcDNA3 was a gift from Takeharu Nagai (Addgene plasmid #51957) and mKO2-N1 was a gift from Michael Davidson (Addgene plasmid #54625).

FUNDING

National Institutes of Health [R01CA173712, R01CA207029, P50GM098792]; National Science Foundation [1745645]; Cancer Center Support (core) Grant [P30CCA14051, in part] from the NCI. Funding for open access charge: National Institutes of Health [P50GM098792].
Conflict of interest statement. None declared.

REFERENCES

1. Yu,J., Hu,K., Smuga-Otto,K., Tian,S., Stewart,R., Slukvin,I.I. and Thomson,J.A. (2009) Human induced pluripotent stem cells free of vector and transgene sequence. *Science*, **324**, 797–801.
2. Warren,L., Manos,P.D., Ahfeldt,T., Loh,Y.H., Li,H., Lau,F., Ebina,W., Mandal,P.K., Smith,Z.D., Meissner,A. *et al.* (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, **7**, 618–630.
3. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–824.
4. Papapetrou,E.P., Tomishima,M.J., Chambers,S.M., Mica,Y., Reed,E., Menon,J., Tabar,V., Mo,Q., Studer,L., Sadelain,M. *et al.* (2009) Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 12759–12764.
5. Nagamatsu,G., Saito,S., Kosaka,T., Takubo,K., Kinoshita,T., Oya,M., Horimoto,K. and Suda,T. (2012) Optimal ratio of transcription factors for somatic cell reprogramming. *J. Biol. Chem.*, **287**, 36273–36282.
6. Min,Y.-L., Li,H., Rodriguez-Caycedo,C., Mireault,A.A., Huang,J., Shelton,J.M., McAnally,J.R., Amoasii,L., Mammen,P.P.A., Bassel-Duby,R. *et al.* (2019) CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. *Sci. Adv.*, **5**, eaav4324.
7. Fujii,W., Kawasaki,K., Sugiura,K. and Naito,K. (2013) Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.*, **41**, e187.
8. Hori,Y., Kantak,C., Murray,R.M. and Abate,A.R. (2017) Cell-free extract based optimization of biomolecular circuits with droplet microfluidics. *Lab. Chip.*, **17**, 3037–3042.
9. Kosuri,S., Goodman,D.B., Cambray,G., Mutalik,V.K., Gao,Y., Arkin,A.P., Endy,D. and Church,G.M. (2013) Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 14024–14029.
10. Woodruff,L.B.A., Gorochowski,T.E., Roehner,N., Mikkelsen,T.S., Densmore,D., Gordon,D.B., Nicol,R. and Voigt,C.A. (2017) Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. *Nucleic Acids Res.*, **45**, 1553–1565.
11. Ghodasara,A. and Voigt,C.A. (2017) Balancing gene expression without library construction via a reusable sRNA pool. *Nucleic Acids Res.*, **45**, 8116–8127.
12. Woodruff,K. and Maerkl,S.J. (2017) A high-throughput microfluidic platform for mammalian cell transfection and culturing. *Sci. Rep.*, **6**, 23937.
13. Li,Y., Jiang,Y., Chen,H., Liao,W., Li,Z., Weiss,R. and Xie,Z. (2015) Modular construction of mammalian gene circuits using TALE transcriptional repressors. *Nat. Chem. Biol.*, **11**, 201–213.
14. Wong,A.S., Choi,G.C., Cheng,A.A., Purcell,O. and Lu,T.K. (2015) Massively parallel high-order combinatorial genetics in human cells. *Nat. Biotechnol.*, **33**, 952–961.
15. Wong,A.S.L., Choi,G.C., Cui,C.H., Pregernig,G., Milani,P., Adam,M., Perli,S.D., Kazer,S.W., Gaillard,A., Hermann,M. *et al.* (2016) Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 2544–2549.
16. Kong,D.S., Thorsen,T.A., Babb,J., Wick,S.T., Gam,J.J., Weiss,R. and Carr,P.A. (2017) Open-source, community-driven microfluidics with Metafluidics. *Nat. Biotechnol.*, **35**, 523–529.
17. Schwake,G., Youssef,S., Kuhr,J.T., Gude,S., David,M.P., Mendoza,E., Frey,E. and Rädler,J.O. (2010) Predictive modeling of non-viral gene transfer. *Biotechnol. Bioeng.*, **105**, 805–813.
18. Bhise,N.S., Shmueli,R.B., Gonzalez,J. and Green,J.J. (2012) A novel assay for quantifying the number of plasmids encapsulated by polymer nanoparticles. *Small*, **8**, 367–373.
19. Davidsohn,N., Beal,J., Kiani,S., Adler,A., Yaman,F., Li,Y., Xie,Z. and Weiss,R. (2015) Accurate predictions of genetic circuit behavior from part characterization and modular composition. *ACS Synth. Biol.*, **4**, 673–681.
20. Xie,M. and Fussenegger,M. (2018) Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nat. Rev. Mol. Cell Biol.*, **19**, 507–525.
21. Guye,P., Li,Y., Wroblewska,L., Duportet,X. and Weiss,R. (2013) Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.*, **41**, 300–330.
22. Dominguez,A.A., Lim,W.A. and Qi,L.S. (2016) Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation *Nat. Rev. Mol. Cell Biol.*, **17**, 5–15.
23. Chavez,A., Scheiman,J., Vora,S., Pruitt,B.W., Tuttle,M., P R Iyer,E., Lin,S., Kiani,S., Guzman,C.D., Wiegand,D.J. *et al.* (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods.*, **12**, 326–328.
24. Wang,T., Xie,Y., Tan,A., Li,S. and Xie,Z. (2015) Construction and characterization of a synthetic microRNA cluster for multiplex RNA interference in mammalian cells. *ACS Synth. Biol.*, **5**, 1193–1200.
25. Brown,B.D., Gentner,B., Cantore,A., Colleoni,S., Amendola,M., Zingale,A., Baccarini,A., Lazzari,G., Galli,C., Naldini,L. *et al.* (2007) Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. *Nat. Biotechnol.*, **25**, 1457–1467.
26. Xie,Z., Wroblewska,L., Prochazka,L., Weiss,R. and Benenson,Y. (2011) Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science*, **333**, 1307–1311.
27. Lillacci,G., Benenson,Y. and Khammash,M. (2018) Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.*, **46**, 9855–9863.
28. Shimizu,K., Sakurai,F., Tomita,K., Nagamoto,Y., Nakamura,S., Katayama,K., Tachibana,M., Kawabata,K. and Mizuguchi,H. (2014) Suppression of leaky expression of adenovirus genes by insertion of microRNA-targeted sequences in the replication-incompetent adenovirus vector genome. *Mol. Ther. Methods Clin. Dev.*, **1**, 14035.
29. Mukherji,S., Ebert,M.S., Zheng,G.X., Tsang,J.S., Sharp,P.A. and van Oudenaarden,A. (2011) MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.*, **43**, 854–859.
30. Gam,J.J., Babb,J. and Weiss,R. (2018) A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity. *Nat. Commun.*, **9**, 2430.
31. Bullock,M.D., Pickard,K.M., Nielsen,B.S., Sayan,A.E., Jenei,V., Mellone,M., Mitter,R., Primrose,J.N., Thomas,G.J. and Packham,G.K. (2013) Pleiotropic actions of miR-21 highlight the critical role of deregulated stromal microRNAs during colorectal cancer progression. *Cell Death Dis.*, **4**, 1–10.
32. Ferreira,J.P., Overton,K.W. and Wang,C.L. (2013) Tuning gene expression with synthetic upstream open reading frames. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 11284–11289.
33. Ferreira,J.P., Peacock,R.W.S., Lawhorn,I.E.B. and Wang,C.L. (2011) Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters. *Syst. Synth. Biol.*, **5**, 131–138.
34. Wroblewska,L., Kitada,T., Endo,K., Siciliano,V., Stillo,B., Saito,H., Weiss,R. *et al.* (2015) Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. *Nat. Biotechnol.*, **33**, 839–841.
35. Raser,J.M. and O'Shea,E.K. (2015) Control of Stochasticity in Eukaryotic Gene Expression. *Science*, **304**, 1811–1814.

APPENDIX

TERMINOLOGY

In this work, we use terminology as follows: A **genetic system or circuit** refers to a set of genes that together carry out a defined function. Popular examples include the CRISPR/Cas9 system, CRY2/CIB1 optogenetic system, and Gal4-VP16 activation of a gene. Each of the genes in a genetic system is referred to as an **element**, which is often encoded on a single plasmid (and therefore *plasmid* is sometimes used to describe each genetic element). **Concentration space** represents the set of all combinations of plasmid concentrations that can be delivered to cells. For instance, Figure 2c shows a two dimensional concentration space where the x-axis corresponds to the concentration of one plasmid delivered to a cell and the y-axis corresponds to that of another plasmid. Performance of the circuit in each cell represents an additional dimension that can be calculated as a function of the concentration space. Generally, the concentration space encompasses the same number of dimensions as the number of plasmids within the system. A **stoichiometric ratio** refers to the molar ratio at which DNA is present within a cell. For co-transfections, we have shown that the fluorescence of transfection markers within each cell corresponds to the ratio at which the DNA is mixed during transfection (Supplementary Figure S5). For example, in Figure 1b individual stoichiometries can be shown within concentration space as diagonal lines. The 1:1 stoichiometry is the 45° line passing through the origin. A **co-transfection (co-TX)** refers to the standard method for transfecting multiple plasmids, where DNA is premixed together before being added to the transfection reagent. A **poly-transfection (poly-TX)** refers to a transfection where each plasmid is mixed with transfection reagent separately, forming transfection complexes that contain only one plasmid species. Then, these different transfection complexes are added to cells such that they uptake a decorrelated amount of each unique plasmid. **Binning** refers to the flow cytometry analysis technique where cells are divided into bins according to the fluorescence of one or more transfection marker fluorescent proteins. With binning, a metric (e.g. median) for the output fluorescence can be calculated and plotted and/or modeled as a function of the transfection marker(s).