

A Toolkit for Rapid Modular Construction of Biological Circuits in Mammalian Cells

João Pedro Fonseca,^{†,¶} Alain R. Bonny,^{†,¶} G. Renuka Kumar,[‡] Andrew H. Ng,[§] Jason Town,[†] Qiu Chang Wu,^{||} Elham Aslankoochi,[†] Susan Y. Chen,[†] Galen Dods,[†] Patrick Harrigan,^{†,⊥} Lindsey C. Osimiri,^{†,¶} Amy L. Kistler,^{*,‡} and Hana El-Samad^{*,†,¶}

[†]Department of Biochemistry and Biophysics, California Institute for Quantitative Biosciences, University of California, San Francisco, San Francisco, California 94158, United States

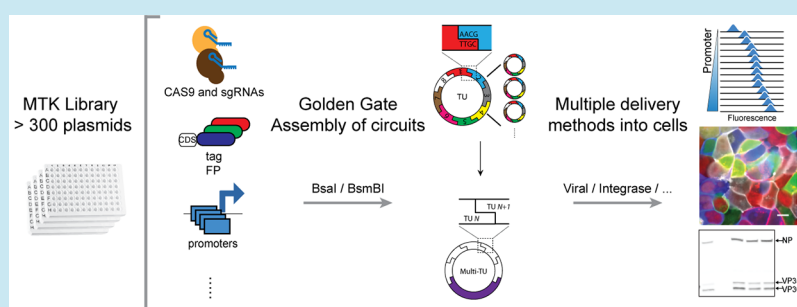
[‡]Chan Zuckerberg Biohub, San Francisco, California 94158, United States

[§]Cell Design Initiative, University of California, San Francisco, San Francisco, California 94158, United States

[¶]The UC Berkeley–UCSF Graduate Program in Bioengineering, University of California, San Francisco, San Francisco, California 94132, United States

^{||}Harvard Systems Biology Graduate Program, Cambridge, Massachusetts 02138, United States

Supporting Information



ABSTRACT: The ability to rapidly assemble and prototype cellular circuits is vital for biological research and its applications in biotechnology and medicine. Current methods for the assembly of mammalian DNA circuits are laborious, slow, and expensive. Here we present the Mammalian ToolKit (MTK), a Golden Gate-based cloning toolkit for fast, reproducible, and versatile assembly of large DNA vectors and their implementation in mammalian models. The MTK consists of a curated library of characterized, modular parts that can be assembled into transcriptional units and further weaved into complex circuits. We showcase the capabilities of the MTK by using it to generate single-integration landing pads, create and deliver libraries of protein variants and sgRNAs, and iterate through dCas9-based prototype circuits. As a biological proof of concept, we demonstrate how the MTK can speed the generation of noninfectious viral circuits to enable rapid testing of pharmacological inhibitors of emerging viruses that pose a major threat to human health.

KEYWORDS: modular cloning, virology, library generation

Molecular cloning is the cornerstone of modern biological research, enabling the generation of DNA vectors that encode a wide variety of molecules, which can further be organized into genetic circuits with applications ranging from basic discovery to medical platforms. To transform molecular cloning into a fast and reproducible engineering discipline, circuit parts need to be modular, vetted for their function, and easy to share. Furthermore, these circuits need to be rapidly assembled, amenable to fast prototyping, and easily deliverable. Conventional approaches to produce DNA vectors involve PCR amplification of DNA fragments, which are ligated following digestion with restriction enzymes. This procedure is time-consuming, laborious, error-prone, and requires sequencing control steps. Commercial DNA synthesis provides a potential avenue to reliably build libraries of DNA vectors.

However, current pricing, size limitation, and turnaround time remain a significant bottleneck.

While in yeast and plant systems toolkits for rapid assembly of genetic circuits from libraries of well described modular parts exist,^{1–6} no similar comprehensive resource exists for mammalian systems. Existing resources such as the Gibson modular assembly platform (GMAP),⁷ and more recently, the Mammalian Modular Cloning (mMoClo)⁸ and Extensible Mammalian Modular Assembly kit (EMMA)⁹ provide a starting point for developing this functionality. However, while these methods use sets of modular parts that can be assembled directionally into transcription units (TUs), they

Received: August 3, 2019

Published: November 5, 2019

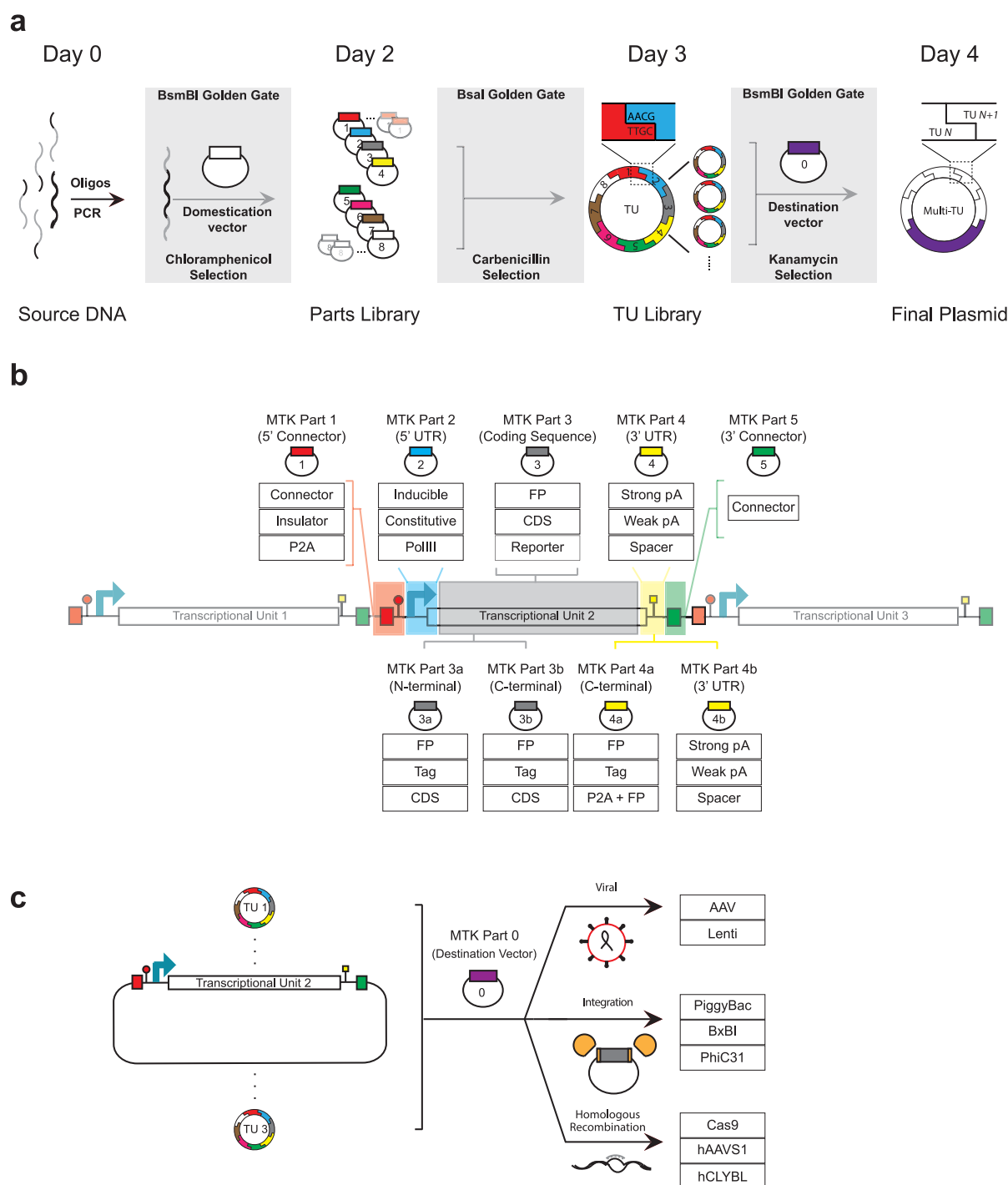


Figure 1. Schematic and definition of parts of the Mammalian Toolkit (MTK). (a) Workflow of the MTK starting with a BsmBI part domestication, BsaI transcriptional unit assembly, and a final BsmBI assembly of a multitranscriptional unit plasmid. Top text indicates approximate time from initial PCR with primers to receiving final plasmid. (b) Schematic of a standard transcriptional unit (TU). Tables below part definition summarize variations of that part that are present in the MTK. [Supporting Table S1](#) contains a comprehensive list of current parts. (c) Library of TUs can be reused with no further modification and delivered to cells in three methods already present in the MTK using a dedicated destination vector (MTK part 0).

have important remaining shortcomings. GMAP requires sequencing in between all cloning steps, which for large DNA elements (such as a 7 kb viral polymerase) makes the method slow and expensive since PCR reactions of such elements are difficult and prone to mutations, and require, in the viral polymerase example, at least 7 sequencing reactions for every tested clone. EMMA and mMoClo provide a method

to hierarchically assemble DNA vectors that encode large circuits, without the need of additional PCR and sequencing reactions. However, they both lack several characteristics that would make them useful for general use. mMoClo is not available in Addgene or upon request, making its adoption unrealizable. EMMA, while available to users, lacks a library of tested parts for facile use and integration of sgRNA cloning.

EMMA also provides good flexibility in terms of part building, but it has not been shown to be able to build circuits with more than 3 transcriptional units. Finally, because EMMA necessitates assembly of 25 parts to build functional plasmids, a substantial energy barrier for its adoption in the community at large seems to exist.

In this work, we present a toolkit that capitalizes on the strengths of previous efforts,^{3,10} but substantially extends their capabilities for mammalian molecular cloning, providing a first platform that can be immediately adopted by the community. The Mammalian ToolKit (MTK) is a library of over 300 parts including vetted promoters, 3' UTRs, fluorescent proteins, insulator and P2A elements that can be rapidly combined to build complex genetic circuits. To facilitate the MTK's ease of use, users can consolidate parts into ready-made, customizable destination vectors to bypass the need to reassemble multiple plasmids and often-used elements such as selection markers. These vectors can be delivered to a wide range of cell types through viral, recombinase, and CRISPR/Cas9 methods. Importantly, the MTK requires the assembly of only 8 unique parts for a functional transcriptional unit, which in turn can be immediately used to build plasmids carrying up to 9 transcriptional units that can encode the expression of a wide variety of coding DNA sequences as well as of sgRNAs for targeting of Cas9 proteins of two species. As a proof of concept, we built a hAAVS1 landing pad for single integration of genetic circuits, created and delivered libraries of protein variants and sgRNAs, and rapidly compared multispecies Cas9-based genetic circuits. Finally, using Ebola virus (EBOV) as an example, we demonstrate how the MTK can simplify and speed the *de novo* generation of noninfectious biosafety level 2 (BSL2) stable cell line expression systems commonly deployed to facilitate analyses of highly pathogenic emerging viruses that require biosafety level 4 (BSL4) handling.

■ RESULTS

An Expansive, Modular Cloning Toolkit for Rapid Prototyping in Mammalian Cells. The basis for the MTK is a library of parts that are “domesticated” from source DNA using a Golden Gate (GG) reaction¹⁰ into a standard vector. This requires designing oligos that anneal to the source DNA, appending restriction sites that enable the resulting PCR product to be digested and ligated into the base parts vector with chloramphenicol resistance (MTK0_027). The resulting plasmid is then sequenced once to ensure fidelity to the source sequence, and can later be reused as a validated modular part in a variety of genetic constructs. Similar to previously published works,^{9,8,3,10} we utilize the type-IIS restriction enzymes BsaI and BsmBI with “reach over” endonuclease activity that leaves four arbitrary base overhangs adjacent to the recognition site. The defining feature for a part vector is a unique four base overhang that categorizes it to similar parts, and ensures an ordered 5' to 3' ligation of parts into a transcriptional unit (TU) of multiple parts. We used the same overhangs as defined in the Yeast Toolkit (YTK).³ These overhangs have been shown to enable highly efficient assemblies and, because the YTK has reached a wide community of users, parts that have been built for the YTK can be reused in the MTK. With a library of sequenced-verified part vectors, only diagnostic restriction digests are necessary to verify correct assembly in subsequent GG reactions. We employ BsmBI and BsaI GG reactions in an alternating manner to assemble a library of part vectors into libraries of TU

plasmids, and ultimately into multitranscriptional unit (multi-TU) plasmids. Like part vectors, TUs have unique overhangs that define their position in a multi-TU plasmid. The time from initial source DNA to final plasmid product is 4 days, assuming the use of fast growing *E. coli* and 18 hours for sequencing turnaround (Figure 1a). However, once MTK plasmid parts are constructed, they can be reused to assemble new configurations in only 2 days. This is a substantial gain for a general, nonspecialist, user who can achieve many complex circuits already from the large library of parts that the MTK provides.

The MTK encompasses part vector categories 1–8 that are sufficient to build and deliver a vast combinatorial library of genetic constructs to cells (Figure 1b). Parts 2, 3, and 4 form the core of a TU, specifying the 5' UTR, coding, and 3' UTR sequences, respectively. Part 2 corresponds to promoter sequences that can be specified to drive constitutive or inducible expression or recruit diverse polymerases. Part 3 vectors are canonical coding sequences that are typically proteins of interest. Part 4 corresponds to 3' UTR sequences that can encode polyadenylation (pA) sequences that terminate transcription or spacer sequences that couple transcription to the downstream TU. Parts 1 and 5 correspond to connector sequences that enable the sequential ordering of TUs into a multi-TU plasmid, with many versions provided in the MTK. For example, Part 1 implements connectors with insulator sequences¹¹ to minimize polymerase read-through between TUs or connectors with cis-acting P2A elements to enable strongly correlated expression from one promoter driving up to five downstream TUs.^{12,13} Overall, the connectors included in the MTK allow the construction of multi-TU plasmids encoding 9 TUs, and hence large genetic circuits. Further nested subdivision can be achieved—for example, Part 3 can be replaced with Part 3a and 3b, and still connect to a Part 4, allowing for combinations that implement tethering of localization tags, protein domains, or any desired coding sequence both N- and C-terminally with an innocuous linker sequence in between (Figure 1b). Additionally, coupled Part 234 vectors accommodate rapid cloning of small guide RNA (sgRNA) expression by oligo annealing for CRISPR/Cas9-related genetic constructs. Parts 6, 7, and 8 generally flank a typical TU and can encode the method of delivery to cells, such as homology arms for a locus of integration. Lastly, the MTK contains kanamycin resistant Part 0 destination vectors that allow delivery *via* viral transduction, transposase transfection,^{14,15} and homologous recombination from the same, recyclable collection of TUs (Figure 1c). Only Part 0 in the final GG reaction needs to be changed in order to accommodate different delivery methods into cells.

Overall, the MTK combines intuitive organization with an expansive library (345 parts listed in Supporting Table S1 and publicly available on Addgene) for rapid facile construction of genetic constructs that can be integrated into cells without redesign or resequencing, irrespective of the delivery method.

MTK Enables Facile Construction of Independent or Multicistronic TUs with Different Levels of Expression.

The MTK contains 17 characterized constitutive promoters derived from a mix of human, mouse, and viral¹⁶ origin, and two inducible promoters whose use is illustrated in Figure 5. This set provides a range of expression levels, and a balance of native and transgene promoters for the two conventional cell lines tested. To enable matching expression levels of different

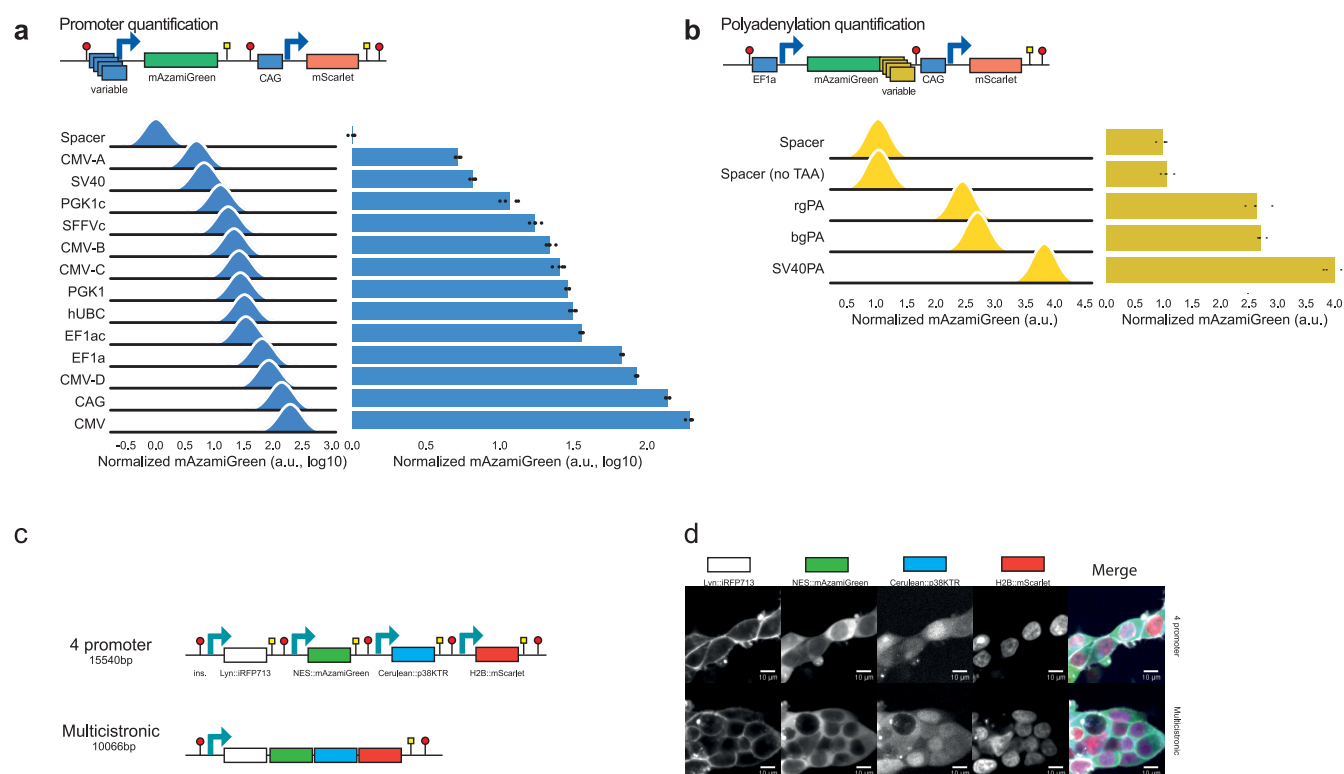


Figure 2. Characterization of constitutive promoters, 3' UTR and multicistronic connectors provided in the MTK. (a) Top: Schematic of a TU used to quantify promoter expression. Bottom: Ranked expression of mAzamiGreen (normalized to a constitutive mScarlet) driven by different promoters. Distributions show one of four biological replicates, and bar plots represent the mean of all four biological replicates. The mean of each replicate is also shown as a black dot. The letter “c” following the name of a promoter (e.g., PGK1 versus PGK1c) is used to designate a “crippled” promoter where the Kozak sequence is disrupted. (b) Top: Schematic of TU used to quantify the effect of the 3' UTR on mAzamiGreen expression driven by the EF1a promoter. Data representation is as described above. (c) Schematic of expression strategies of 4 fluorescent proteins targeted to 4 different cellular locations. In the first strategy, a plasmid contains 4 TUs, each containing a CAG promoter. In the first TU, CAG drives expression of membrane-targeted iRFP713 (Lyn::iRFP713), in the second expression of cytoplasmic mAzamiGreen (NES::mAzamiGreen), in the third expression of p38 kinase translocation reporter fused to mCerulean (mCerulean::p38KTR), and in the fourth histone H2B fused to mScarlet (H2B::mScarlet). In the second strategy, a multicistronic plasmid encodes the same proteins, but all are produced from a single transcript driven by a CAG promoter with 3 P2A sequences separating the 4 peptides. The size in bp of each plasmid is also indicated. (d) Confocal images of HEK293T cells expressing the 4 promoter (top) or multicistronic (bottom) plasmid. Merge panel shows Lyn::iRFP713 in white, NES::mAzamiGreen in green, Cerulean::p38KTR in blue, and H2B::mScarlet in red.

proteins without reusing the same promoter, this panel also includes select promoters that closely match each other.

To characterize the relative strengths of these promoters we assembled a panel of 14 TUs with varied promoter parts driving mAzamiGreen expression, flanked by the bovine growth hormone (Bgh) pA 3' UTR. An insulated downstream TU, where a CAG promoter expresses mScarlet with the rabbit beta-globin (Rgl) pA 3' UTR (Figure 2a), was used for normalization. Transient transfection in HEK293T cells demonstrated that this suite of promoters spanned a smooth continuum of 2 orders of magnitude over background, with the strongest promoter, CMV, more than 300-fold greater than a promoter-less mAzamiGreen (Figure 2a). The relative strengths of these promoters was largely maintained across human and mouse cell lines (Supporting Figure S1a,b), suggesting that they are portable across commonly used cell lines. Additionally, the rank of promoter expression was consistent between transient and stably integrated expression (Supporting Figure S1c,d).

The MTK provides options for further fine-tuning of TU expression using a collection of five different 3' UTR sequences. Similar to the promoter comparison, we generated a circuit with the same constitutive promoter (EF1a) and

varied the 3' UTR sequence to compare normalized mAzamiGreen expression (Figure 2b). We observed a range of 4-fold change in expression of mAzamiGreen among the three conventional 3' UTR sequences. While the Bgh pA and Rgl pA signals have the same effect on gene expression, the simian virus 40 (SV40) pA signal had a nearly 1.5-fold greater effect (Figure 2c). Moreover, replacing a canonical 3' UTR sequence with either a spacer sequence with a stop codon (used for lentiviral delivery) or without a stop codon (used for multicistronic read-through) diminished expression of its upstream coding sequence. While we explored here only a few combinations related to one promoter (EF1a), libraries of fine-tuned expression of a protein of interest can be easily generated using this platform by the interested user.

Finally, to enable concomitant expression of up to 6 TUs in a multicistronic vector, we incorporated ribosome skipping P2A elements^{12,13} into the MTK part 1. We used this part to build a vector with one CAG promoter driving membrane-tethered iRFP713, cytoplasmic mAzamiGreen, mCerulean-tagged p38KTR,¹⁷ and histone H2B fused to mScarlet separated by P2A elements, and compared expression from this construct to a parallel multi-TU construct (Figure 2c). The multicistronic construct conferred a 35% reduction in

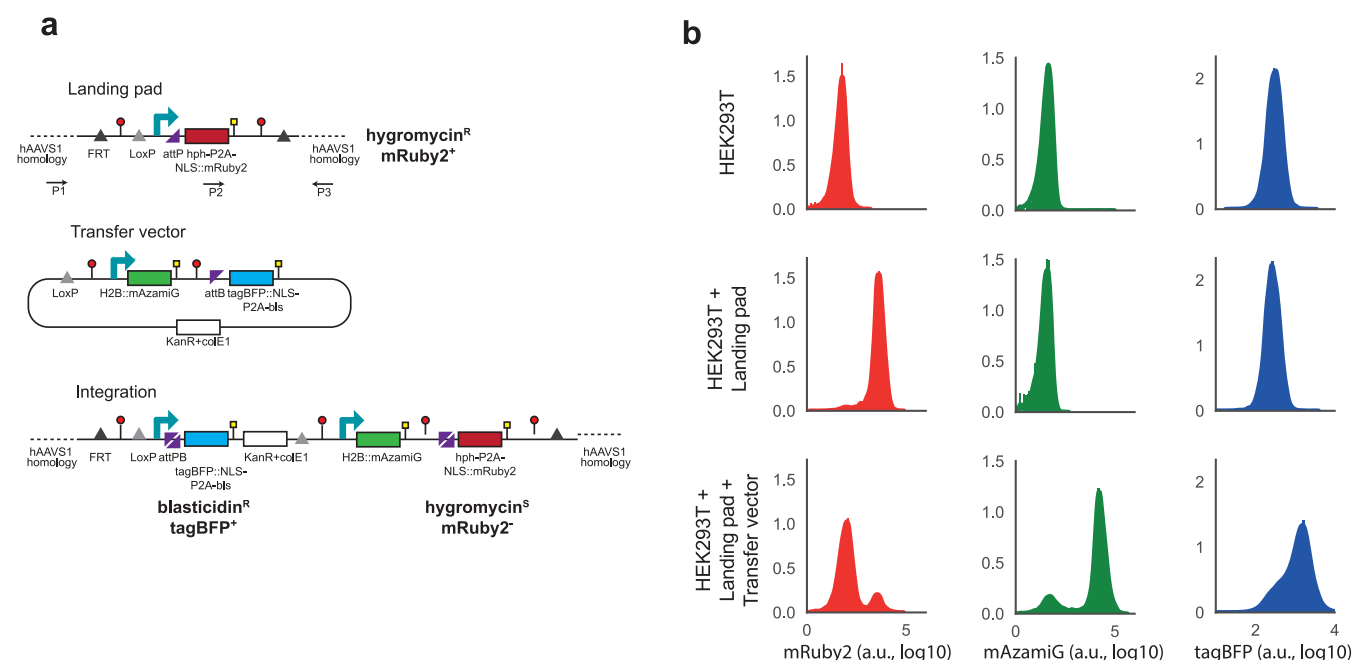


Figure 3. Generation and testing of landing pads for HEK293T cells using the MTK. (a) Schematic of landing pad, transfer vector and integrated vector. Also shown are genotyping primers. FRT, flippase recognition target; ins., Insulator; LoxP, LoxP site; attP, BxBI phage attachment site; hph, hygromycin resistance gene; bls, blasticidin resistance gene; NLS, nuclear localization signal; KanR, kanamycin resistance gene; colE1, colE1 origin of replication; attB, BxBI bacterial attachment site; mAzamiG, mAzamiGreen. (b) mRuby2, mAzamiGreen and tagBFP expression in populations of parental, Landing pad, and Landing pad with Transfer vector HEK293T cells. mRuby2 expression indicates presence of hAAVS1 landing pad. mAzamiGreen and tagBFP expression indicates precise integration of transfer vector in hAAVS1 landing pad.

vector size (15 540 to 10 020 bp) compared to the multi-TU construct yet maintained the correct localization of each of the proteins in transiently transfected HEK293T cells (Figure 2d).

These data demonstrate the opportunity to deploy P2A elements to generate multicistronic expression constructs to circumvent the size limitations of conventional viral delivery vectors.

MTK Contains a Landing Pad System and Accompanying Destination Vector. When delivering synthetic genetic circuits, it can be essential to have site-specific and reliable single copy integration. Therefore, the MTK includes a BxBI-dependent landing pad (LP) system for integrating synthetic circuits in a locus of choice. This system is divided into two parts: an LP cell line and an LP transfer vector. The LP transfer vector can be used with MTK generated LP cell lines, as well as cell lines carrying the SBI pinpoint system landing pad, or that were generated using other methods and that contain a BxBI attB site.⁸ We used a CRISPR/Cas9 approach to build a HEK293T cell line with a BxBI landing site in the well-characterized adeno-associated virus integration site 1 locus^{18,19} (hAAVS1 LP) and tested its ability for site-specific recombination. The vector used encoded a multicistronic cassette with hygromycin resistance and nuclear-localized mRuby2. The BxBI attP site was located between the promoter and the first gene of this cassette (Figure 3a). We verified the correct integration of the LP cassette by PCR in 8 clones (Supporting Figure S3) and continued its characterization in clone #8 with a monoallelic LP. We further confirmed the presence of the LP in the cell line by mRuby2 expression (Figure 3b).

In order to verify that the transfer vector was correctly integrated in the LP, we positioned a promoter-less multicistronic cassette encoding blasticidin resistance and nuclear

localized tagBFP downstream of the BxBI attB site. When site-specific integration is accomplished, the cell line switches fluorophore (from mRuby2 to tagBFP) and resistance (hygromycin to blasticidin).

To test the landing pad, we integrated a transcriptional unit that expresses H2B mAzamiGreen-fused histone H2B from the CAG promoter (Figure 3, Supporting Figure S3b). Upon integration of the transfer vector into the hAAVS1 LP, we noted the decrease in mRuby2 expression together with expression of mAzamiGreen and tagBFP in most cells (Figure 3b, Supporting Figure S3b), with a small fraction of cells showing no or incorrect integration of the transfer vector. Finally, while we chose to integrate the landing pad in one genomic locus, others can be easily targeted by replacing the homology arms in the CRISPR/Cas9 vector.

MTK Allows Rapid, One Step Combinatorial Construction of Libraries. The inherent modularity of the MTK workflow facilitates parallelization of large libraries of vectors in one-step combinatorial reactions. Moreover, such libraries can be reused and delivered repeatedly and in multiple ways. These features are particularly relevant for optimization of a large number of protein or circuit variants or CRISPR/Cas9 screens where efficient throughput and variant representation are critical. Here, we provide two examples of MTK combinatorial library construction that illustrate how it can be applied to streamline the generation of large libraries of TU variants. First, a library of fluorescent proteins that vary in their localization, and which can be integrated as a single-copy into the BxBI landing pad site; second, a library of viral-delivered sgRNAs that target Cas9 to two fluorescent proteins leading to their disruption.

In the first library, we combined one of three localization tags (NLS (nucleus), NES (cytoplasm), and Lyn (plasma

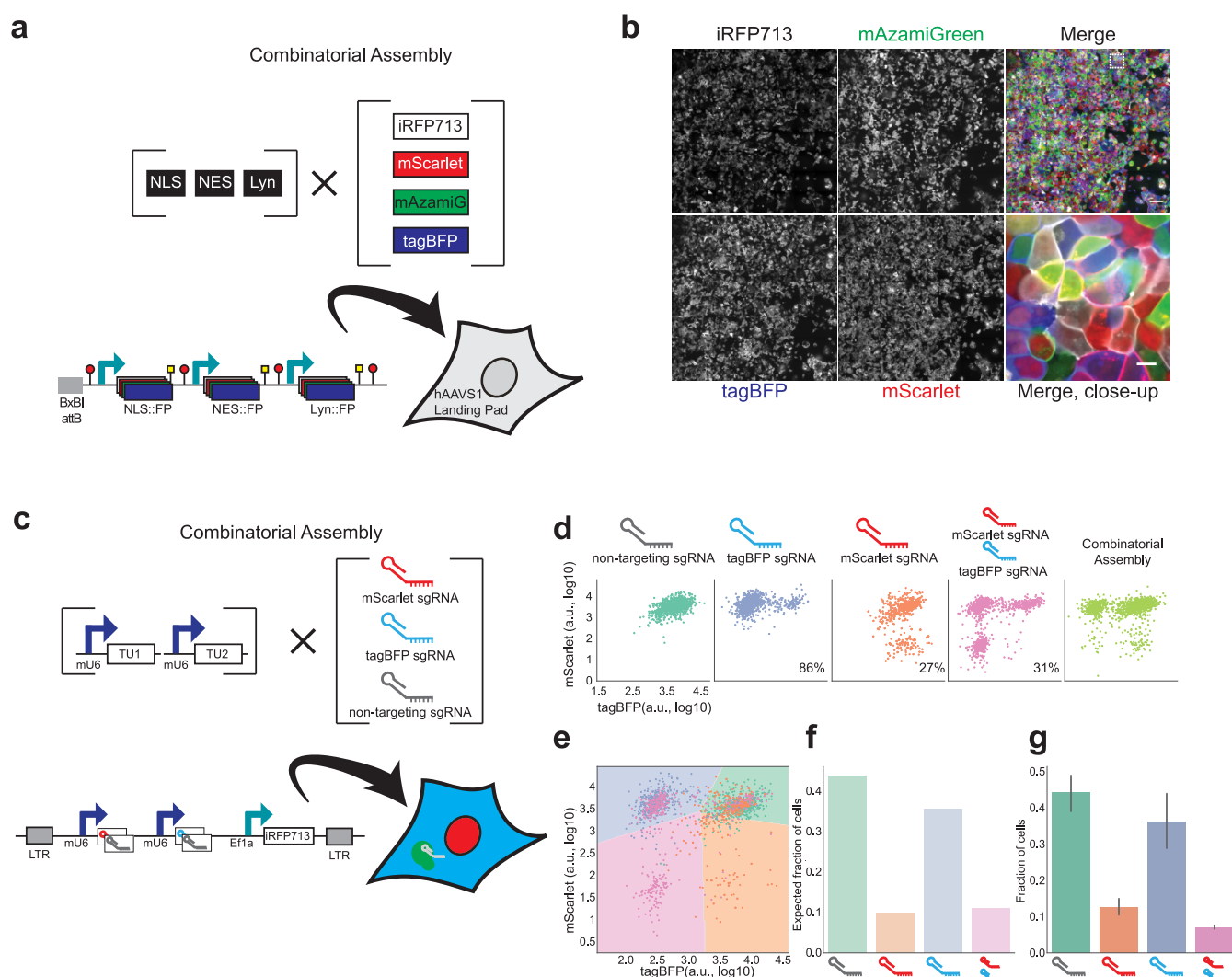


Figure 4. Combinatorial assembly of complex libraries using the MTK. (a) Overview of the MTK strategy for a combinatorial, one pot, generation of a library containing 12 combinations of 4 different fluorescent proteins (FP) targeted to three cellular locations. The final multi-TU construct is composed of 3 TUs: EF1a driving the nuclear localization of an FP; CMV expressing the cytoplasmic localization of an FP; and CAG producing a membrane-tethered FP. Each FP is one of 4 variable fluorophores (tagBFP, mAzamiGreen, mScarlet, and iRFP713), giving a total of 64 possible variants. Constructs are delivered to the landing pad in HEK293T cells. (b) Confocal image of HEK293T+LP, transfected with pooled library in panel (a). iRFP713 shown in white, mScarlet shown in red, mAzamiGreen shown in green, and tagBFP shown in blue. Merge and close-up show cells expressing 3 FPs in three subcellular locations displaying the variety of expected combinations. (c) Overview of the MTK strategy for a combinatorial, one pot library of lentiviral vectors carrying sgRNAs targeting tagBFP, mScarlet, both, or nontargeting. The final multi-TU construct is composed of three TUs: mU6 promoter driving the expression of mScarlet or nontargeting sgRNA; mU6 promoter driving the expression of tagBFP or nontargeting sgRNA; EF1a driving the expression of iRFP713 for identification of cells that have integrated the construct. Library was produced and transduced to HEK293T cells expressing tagBFP, mScarlet, and Cas9 fused to mAzamiGreen (3C cells). (d) Scatter plots of tagBFP and mScarlet fluorescence in populations of 3C cells where sgRNAs were individually expressed to target tagBFP, mScarlet, tagBFP and mScarlet, or with nontargeting sgRNA. Last panel shows the scatter plots of tagBFP and mScarlet fluorescence in populations transfected with combinatorially assembled library of sgRNAs. (e) Scatter plot of tagBFP and mScarlet fluorescence in populations of 3C cells, where an equal number of cells expresses sgRNAs that target tagBFP, mScarlet, tagBFP and mScarlet, or with nontargeting sgRNA. Shaded areas correspond to 4 classes identified by a linear classifier (green, nontargeting; blue, tagBFP; orange, mScarlet; pink, mScarlet and tagBFP). (f) Expected fraction of cells expressing different guide combinations as identified by linear classifier in panel (e) where 4 sgRNA combinations have equal ratios. (g) Measured fraction of cells expressing different guide combinations in combinatorial assembly is similar to (f). Bar plot and error bars represents mean and 95% CI of four biological repeats. NLS, Nuclear localization signal; NES, nuclear export signal; Lyn, plasma membrane tag; mAzamiG, mAzamiGreen.

membrane)) with four fluorescent proteins (tagBFP, mAzamiGreen, mScarlet, and iRFP713) (Figure 4a). Performed in parallel, this reaction generated 4 fluorescent variants per localization tag. These libraries of variants were further pooled in equimolar amounts with a LP destination vector. This single reaction was predicted to create a library of 64 distinct variants, where each vector in the library encodes three random

fluorescent proteins that are localized in the nucleus, cytoplasm, and plasma membrane. The final library was delivered to the hAAVS1-LP HEK293T cell line described before (Figure 3). Visualization of expression and localization of the fluorescent proteins after selection with blasticidin showed a qualitative assortment of random, single fluorescent

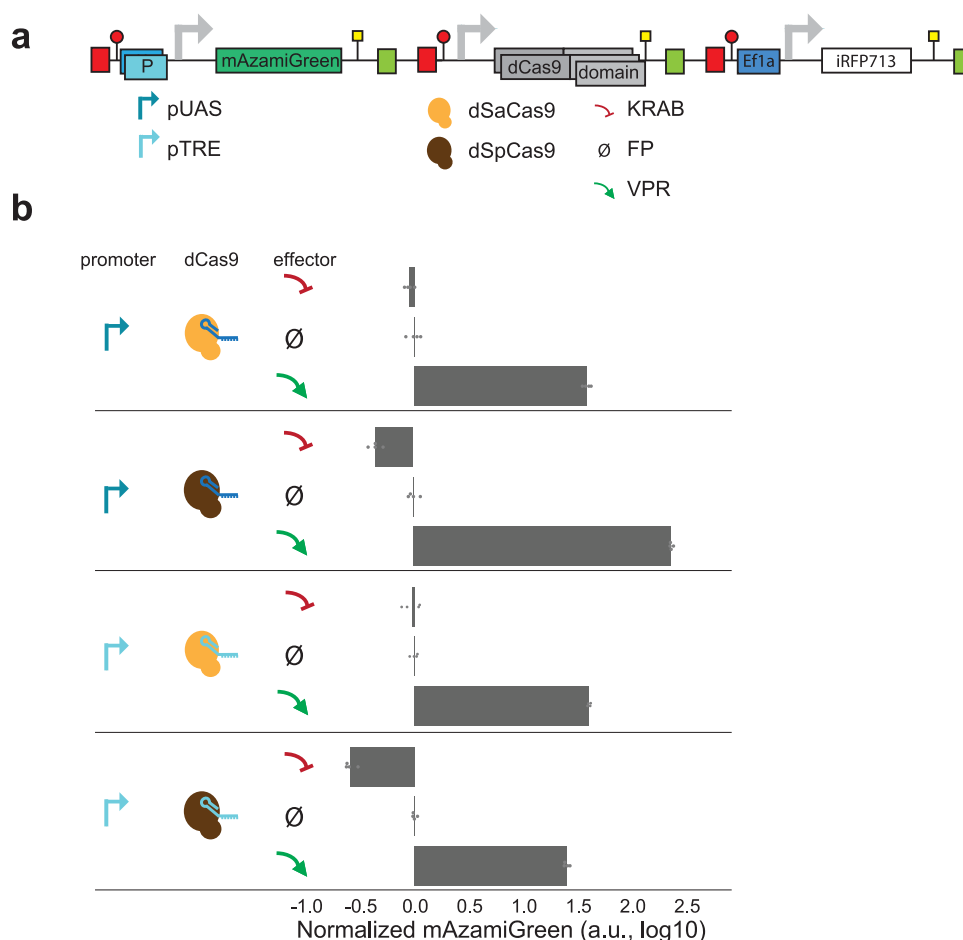


Figure 5. Parallelization of dCas9 circuit prototyping using the MTK. (a) Schematic of 12 dCas9 circuit variations assembled in parallel. Each variation is a multi-TU plasmid with either a TRE or UAS inducible promoter controlling the expression of mAzamiGreen, followed by Ef1a driving either dSpCas9 or dSaCas9 fused to a KRAB domain, fluorophore, or VPR domain. UAS, Upstream Activating Sequence. TRE, Tetracycline Responsive Element. dSpCas9 and dSaCas9, deactivated *S. pyogenes* and *S. aureus* Cas9, respectively. KRAB, Krueppel-associated box. FP, fluorescent protein. VPR, VP64-p65-Rta. (b) mAzamiGreen expression (normalized by iRFP713) for different configurations of the circuit. Every row is a different configuration, corresponding to a variation of the promoter, dCas9 used, and effector used. Configurations are grouped by their promoter-dCas9 pairing (different groups are separated by solid lines). Bar plot represents mean of 4 biological replicates, the mean of each is shown as a black dot.

markers located in the three subcellular compartments (Figure 4b), as predicted from the cloning strategy.

While other Golden Gate-based toolkits have the ability to generate combinatorial libraries of proteins, they cannot generate combinatorial libraries of sgRNAs. Such libraries are useful for targeting a combination of genes, either with Cas9 or with a wide variety of modified dCas9 variants. To address this need, we showcased the ability of the MTK to generate sgRNA combinatorial libraries by building a library of sgRNAs that target two fluorescent proteins. We first assembled TUs that contain sgRNAs that target tagBFP, mScarlet, or that are nontargeting.²⁰ We combined those in a multi-TU unit so that a final Lentiviral delivery plasmid contained 2 guide RNAs, ensuring that we had any one of four possible outcomes (knockout of tagBFP, mScarlet, tagBFP and mScarlet, or no knockout) in the final library (Figure 4c). Each multi-TU encoded iRFP713 to facilitate the identification of cells that express sgRNAs.

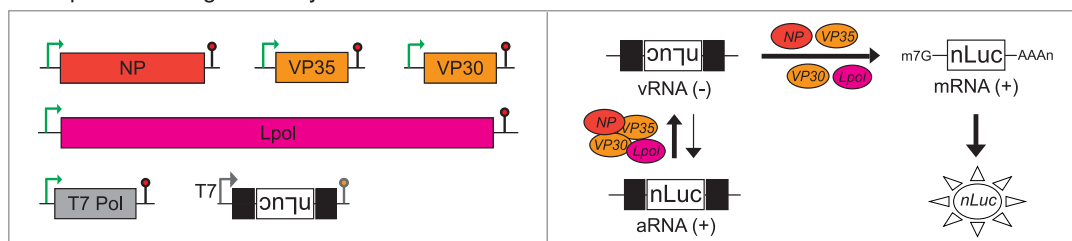
Since sgRNA targeting has variable efficiency,²¹ we concurrently generated four individual lentiviral plasmids for the four outcomes of the library as a control to compare the results of the combinatorial assembly. We independently

transduced the four control viruses in HEK293T cells expressing tagBFP, mScarlet and spCas9 (3C cell line) and quantified tagBFP and mScarlet expression 2 weeks after transduction. Cells that contained the nontargeting sgRNAs expressed high levels of both tagBFP and mScarlet (Figure 4d, first panel). Cells expressing the tagBFP-targeting sgRNA showed reduced expression of tagBFP in 86% of the population (Figure 4d, second panel), while only 27% of cells showed reduced expression of mScarlet when its guide was expressed (Figure 4d, third panel). Accordingly, when both guides were coexpressed, only 31% of cells show reduced expression for both proteins, while most cells had reduced tagBFP (Figure 4d, fourth panel).

We used these results to build a linear classifier, trained on the individual targeted variants of sgRNA to determine the likely sgRNAs that each cell received in a population transduced with the full sgRNA library. Due to the inefficient targeting of the mScarlet guide, the linear classifier had an average precision of 0.62 and an average recall of 0.58 (Figure 4e, Supporting Figure S4) and predicted higher than expected proportions of nontargeting and tagBFP sgRNA-expressing

a Zaire Ebolavirus (ZEBOV) (-)ssRNA genome

ZEBOV 6 plasmid minigenome system



ZEBOV-4cis multicistronic construct

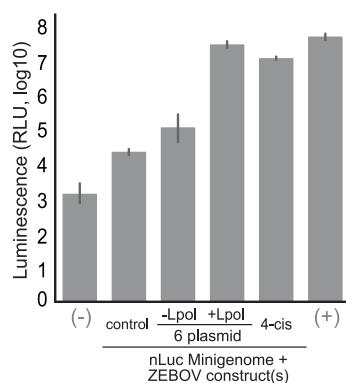
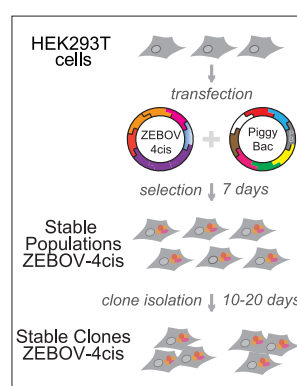
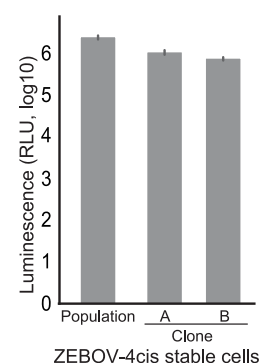
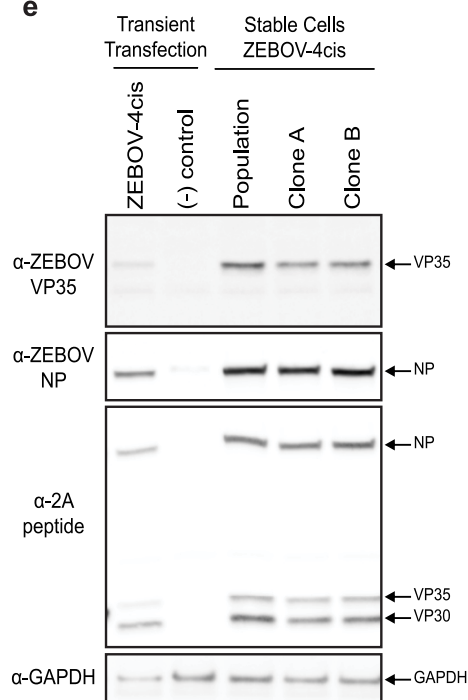
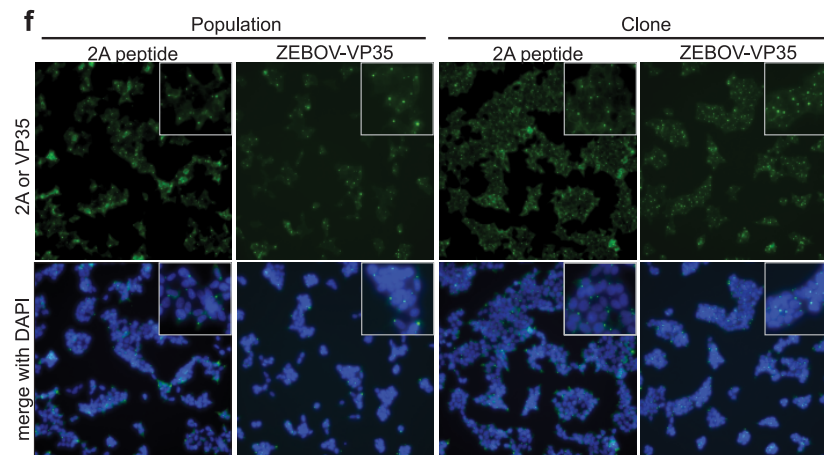
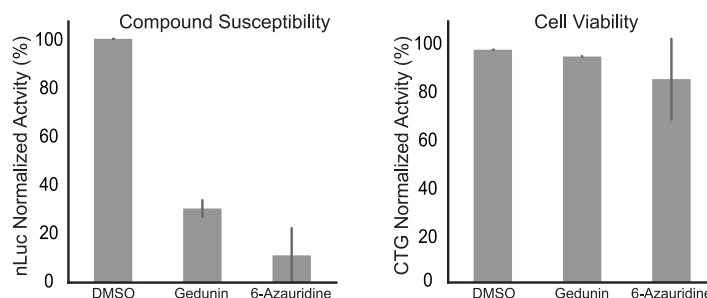
**b****c****d****e****f****g**

Figure 6. Generating multicistronic constructs for Zaire ebolavirus ribonucleoproteins in a mammalian cell host using the MTK. (a) Schematic of the different constructs. Panel 1: Schematic of the Zaire ebolavirus (ZEBOV) negative sense single stranded RNA (-ssRNA) genome. Panel 2: Functional minigenome assay for ZEBOV using the 6 plasmid system that includes expression constructs for NP, VP35, VP30, and Lpol viral

Figure 6. continued

proteins as well as the T7 polymerase along with a T7-driven minigenome reporter construct flanked by viral noncoding 3' and 5' UTRs, the leader and trailer, respectively. Transfection of the 6 plasmids into mammalian cells results in expression of NP, VP35, VP30, Lpol, and T7 and transcription of the minigenome viral RNA (vRNA(−)). The vRNA(−) is replicated to generate antigenomic RNA (aRNA(+)) and transcribed into mRNA by the viral proteins to yield reporter gene expression. Panel 3: Schematic of the Zaire ebolavirus multicistronic construct (ZEBOV-4cis) with NP, VP35, VP30, and Lpol separated by P2A ribosomal skipping site elements. (b) Luminescence measurements of ZEBOV minigenome activity. HEK293T cells were transfected with ZEBOV nano luciferase (nLuc) minigenome in combination with pCAGGs empty control plasmid (control), the ZEBOV 6 plasmids (with or without Lpol: +Lpol, −Lpol), or with ZEBOV-4cis in PiggyBac part 0 vector. Positive (+) and negative (−) controls for nLuc expression levels include transfection of only pCAGGS-nLuc plasmid or pCAGGs empty plasmid, respectively. Nano luciferase activity was measured 2 days post transfection. Bar plots represent the mean of biological replicates ($n = 2$). (c) Schematic of ZEBOV-4cis stable cell line generation using PiggyBac transposon. HEK293T cells were co-transfected with MTK043-ZEBOV-4cis and a PiggyBac expression construct for 3 days and selected with hygromycin for 7 days to generate stable cell populations expressing ZEBOV RNP complex proteins. Clones were isolated from these populations via limited dilution plating in additional 10–20 days. (d) Luminescence measurements of minigenome activity in ZEBOV-4cis stable population. ZEBOV-4cis stable cells and clones were transfected with a T7-driven ZEBOV minigenome construct encoding an nLuc reporter along with T7 polymerase for 2 days followed by nano luciferase assay. Bar plots represent mean of technical replicates ($n = 10$). (e) Western blot confirmation of protein expression. ZEBOV-4cis stable population and clones as well as HEK293T cells transfected with MTK0–43-ZEBOV-4cis were processed for Western blot analysis with mouse anti-ZEBOV VP35, rabbit anti-ZEBOV-NP, mouse anti-2A peptide, and rabbit anti-GAPDH antibodies. (f) Immunofluorescence analysis of viral protein localization. ZEBOV-4cis population and clone were stained with anti-2A peptide and anti-ZEBOV-VP35 primary antibodies and Alexa-488 secondary antibody. DAPI stained nuclei are shown in the merged images. Insets represent 2× magnified fields. (g) Effect of chemical compound inhibitors on minigenome activity in ZEBOV-4cis stable cells. ZEBOV-4cis stable population was transfected with a ZEBOV minigenome construct encoding a nano luciferase reporter and treated with DMSO (1%), Gedunin (5 μ M), or 6-Azauridine (5 μ M). After 2 days cells were processed for nano luciferase assay for functional minigenome activity and cell titer glo assays for cell viability. Levels of signal relative to DMSO control is plotted. Bar plots represent mean of biological replicates ($n = 2$).

cells when presented with an equal proportion of each variant of the library (Figure 4f).

To test the combinatorial assembly approach, we transduced the 3C cells with 4 biological replicates of the sgRNA library. After 14 days of selection, we measured tagBFP and mScarlet expression and identified cells belonging to the four possible outcomes of the viral library, suggesting that all combinatorial possibilities are achievable through this method (Figure 4g). Importantly, the fraction of each outcome was in accordance with the predictions of the classifier, indicating that the MTK is able to rapidly generate libraries for combinatorial exploration of proteins variants or sgRNAs that maintain the correct ratio of variants.

MTK Facilitates Optimization of Combinatorial Gene Circuits for Synthetic Biology Applications. To capitalize on the potential of the MTK for parallel testing of combinatorial circuits, we constitutively expressed deactivated *Streptococcus pyogenes* and *Streptococcus aureus* CRISPR Cas9 (dSpCas9 and dSaCas9, respectively) C-terminally fused to either a fluorescent protein (FP), repressor domain (Krüppel associated box, KRAB), or an activator domain (VP64-p65-Rta, VPR).^{22–26} These protein-effector combinations were targeted to two inducible promoters included in the MTK: the GAL4 Upstream Activating Sequence (UAS) or the Tet Responsive Element (TRE) (Figure 5a). Each gene circuit was assembled upstream of a constant TU control with constitutive EF1a expression of iRFP713. In total, we transiently transfected 12 different circuits into HEK293T cells and normalized changes in mAzamiGreen expression by iRFP713. The plasmids encoding these circuits have an average size of 16 kbp. While our parallel building protocol required only 2 test restriction enzyme digestions to verify the correct assembly of the circuits, methods based on Gibson assembly would have required about 200 sequencing reactions (12 circuits, sequencing 14 kbp of each one) for sequence verification of the 12 circuits assembled from 9 PCR products.

As expected, in each of the circuit iterations, expression of mAzamiGreen relative to the iRFP713 control was below basal

with the repressor domain, or above basal with the activator domain (Figure 5b). dSpCas9 achieved approximately 3-fold reduction in fluorescence through the KRAB effector and 30-fold induction in both the UAS and TRE promoters. In contrast, the dSaCas9 fused to VPR was an effective activator, increasing expression 30-fold, but when fused to a KRAB domain, dSaCas9-mediated repression was comparable to background FP expression levels. The TRE promoter exhibited higher basal activity due to its overall greater repression, and lower activation when compared to the UAS promoter. This information expedites the rational design of dCas9-based gene circuits and can be easily scaled to screen and customize various combinations of genetic circuits.

MTK Streamlines the Generation of Endogenous Viral Circuits. Viruses represent a class of naturally occurring genetic circuits that are particularly amenable to MTK construction due to their modular genome organization. As such, the MTK provides virologists the opportunity to rapidly “boot up” the genes, replicons, or the complete life cycles of emerging viral agents as soon as their sequences are available. A key example corresponds to the filoviruses, a family of emerging highly pathogenic (BSL4) RNA viruses that includes the Ebola viruses.²⁷ Global efforts to contain filovirus outbreaks have resulted in several candidate vaccines and antivirals; however, there is currently no approved preventive or therapeutic treatment.

A challenge to advancing our understanding of EBOV was its highly pathogenic nature and the need for BSL4 containment. In the late 1990s a BSL2 system was developed for Zaire Ebola virus (ZEBOV) to facilitate research on its transcription and genome replication sublife cycle and antiviral discovery efforts.^{28,29} However, this system requires transient transfection of 6 plasmids (four viral proteins (NP, VP35, VP30, and Lpol), T7 polymerase, and a T7-driven minigenome reporter construct²⁸) (Figure 6a). The complexity and variability of this system has limited feasibility to perform industrial scale compound screening campaigns to identify effective inhibitors of EBOV. More recently, a robust stable

ZEBOV RNP cell line system was developed *via* sequential integration of RNP viral proteins;³⁰ however, this approach entailed a time-intensive, multistep approach that took ~1 month. This time scale is undesirable if the goal is to enable rapid investigation of emerging viruses or variants of the same virus.

Given the urgent need and time sensitive nature that outbreaks of highly pathogenic emerging viral agents pose to human health, we explored how the MTK workflow can streamline and improve the *de novo* generation of BSL2 tools like the EBOV RNP minigenome replicon system. We used the MTK system to design a single 4-cistronic construct of Zaire ebolavirus (ZEBOV-4cis, Figure 6a) directly and simultaneously into five different Part 0 destination vectors (Supporting Figure S5a). In transient transfections, all variants of the ZEBOV-4cis constructs displayed levels of minigenome activity similar to the 6-plasmid system, with 40–60-fold higher activity than cells lacking the viral polymerase Lpol. All five of these systems showed robust cell viability (Figure 6b, Supporting Figure S5b). Stable cell lines harboring the ZEBOV-4cis construct were generated using PiggyBac transposon-mediated integration in 10 days followed by isolation of clonal cells (Figure 6c, Supporting Figure S5c).

Comparable viability and minigenome reporter activity were observed in two independent ZEBOV-4cis stable populations and clonal cell lines (Figure 6d). The P2A ribosome skipping sites encoded in the ZEBOV-4cis construct confer 2A peptide tags at the C-terminus of each of the NP, VP35, and VP30 proteins, enabling the simultaneous detection of all 3 proteins with a 2A peptide antibody (Figure 6e). Larger multi-ORF fusion protein products were not detected with the 2A peptide antibodies or NP and VP35 antibodies, indicating efficient “self-cleavage” occurs at each of the P2A sites. As expected, VP35 showed diminished expression, probably due to the known effect of P2A elements on downstream peptides. This is a trait that may be desirable in certain systems such as viruses that regulate expression levels of downstream genes.³¹ Parallel immunofluorescence analysis showed a punctate cytoplasmic localization pattern (Figure 6f) consistent with previously described ZEBOV inclusion bodies that correspond to sites of viral replication.^{32–34} A lack of available specific antibodies or 2A tag encoded in the Lpol gene made parallel protein expression analyses infeasible for Lpol; however, the Lpol dependence of minigenome activity in this system (Figure 6b and 6c) provides confirmation of functional Lpol expression.

We also examined the susceptibility of the ZEBOV-4cis stable cells to previously described small molecule inhibitors: Gedunin, an inhibitor of heat shock protein 90, and 6-Azauridine, a nucleoside analogue.^{35,36} Treatment of ZEBOV-4cis stable cells with 5 μ M of each compound revealed >50% inhibition of minigenome activity, with minimal impact on cell viability (Figure 6g). These data indicate that the ZEBOV-4cis stable cell lines generated here are similarly susceptible to known inhibitors identified in transient RNP minigenome systems or recombinant virus systems.^{35,36} Taken together, these data demonstrate that the MTK workflow facilitates rapid generation of BSL2 systems to study filoviruses and other viral agents as they emerge, providing a critical opportunity to increase the throughput and the rate at which we can screen for and identify candidate antiviral compounds.

DISCUSSION

A key element of cellular engineering is the ability to quickly build, test, and iterate on designs. This is currently infeasible in mammalian cells using conventional methods in molecular cloning. Here, we presented the MTK, a platform that takes an important step to remove bottlenecks in mammalian cellular engineering.

A major asset of the MTK is a large, characterized suite of modular parts to build versatile TUs, which we showcased by driving different levels of gene expression for single TUs and creating multicistronic constructs. Additionally, we included lentiviral, recombinase, and Cas9 delivery vectors to maximize the number of contexts where circuits can be implemented. Due to the modularity of the MTK, it is straightforward to generate new variants of delivery vectors, which we demonstrated by creating BxBI Landing Pad for the hAAVS1 locus.

The set of validated, interchangeable parts of MTK constitute fundamental tools for the rapid and facile assembly of genetic circuits. We illustrated this point in two ways: first, we assembled a combinatorial library of transcriptional units that encode different proteins or guide RNAs; second, we combinatorially assembled new circuits that use different MTK parts. Such libraries are generated through the MTK in a straightforward way, reproducing all variants included in the combinatorial assembly process and maintaining their ratios through all the cloning steps. While these studies tackled a small set of applications, many new combinatorial libraries can be built with any parts in the MTK as it currently exists, or as users add to its parts. For example, introducing a panel of C-terminal fluorescent protein tags into the MTK would only require a one-time PCR and sequence verification for each tag to “domesticate” the series as 4a Parts. Additionally, the facile assembly of sgRNAs expedites the production of expression vectors that are easily amenable to CRISPR screens of all scales. With the many reagents encoded in the MTK library, diverse TUs with multiple functions can be rapidly assembled, screened, and repurposed to generate a wide array of final circuit vectors.

The capabilities provided in the MTK are particularly relevant in the context of infectious disease outbreaks where time is of the essence. Responding to and containing such outbreaks is an important public health challenge that demands rapid production and iteration of viral circuits in mammalian cells to enable discovery of inhibitors and analysis of their basic biology. Using ZEBOV as an example, we demonstrated how the MTK can reduce the lead time needed to generate functional reagents and respond to outbreaks of highly pathogenic viruses. We generated functional cell lines stably expressing the ZEBOV replication complex components in days, while maintaining a library of these components for diverse future applications. Furthermore, given that discovery or availability of a virus sequence does not necessarily correlate with successful culturing of the virus in a lab,³⁷ the MTK system provides a powerful and parallelizable method to rapidly test multiple viral strains and primary clinical isolates, as well as engineered variants of each of these that are optimized for expression in cell culture. While this presents a scaling challenge for conventional cloning approaches, the MTK workflow enables faster iteration and identification of optimal viral circuits for downstream analyses. Thus, the MTK presents a tool for virologists to begin to scale functional

experimental studies apace with the recent explosive growth in viral genome sequences.^{38,39}

Finally, while not explored in this work, the MTK constitutes a launching platform for additional exciting capabilities, including the incorporation of barcoding capabilities for bulk and single-cell sequencing technologies,^{40–42} enhanced flexibility in genetic circuit design,⁴³ and automation to streamline assembly. These advances are poised to position the MTK as a major catalyst for biological research and biotechnology.

METHODS

Bacterial Cell Culture. Commercial MachI and XL10 strains (QB3MacroLab) were used to transform plasmid vectors. A typical transformation mixture consists of 2 μ L of the Golden Gate reaction product and 48 μ L bacteria incubated on ice for 30 min, heat shocked at 42 °C for 1 min, and recovered on ice for 5 min, and the reaction mixture is plated onto selective agar and incubated overnight at 37 °C. In the case of multi-TU transformations, cells are recovered in LB media for 30 min after heat shock at 37 °C before plating reaction onto kanamycin selective agar plates. Cells were cultured in antibiotic concentrations of 100 μ g/mL chloramphenicol, 25 μ g/mL carbenicillin, and 100 μ g/mL kanamycin.

Golden Gate Reactions. The general reaction mixture follows 0.5 μ L per PCR product, annealed oligos, geneblock, or plasmid (50 fmol/ μ L⁻¹); 1 μ L T4 DNA Ligase Buffer (10 $\times$) (NEB) with 0.25% PEG; 0.5 μ L T4 DNA Ligase (NEB) diluted with water to a total volume of 9.5 μ L. For BsaI GG reactions, we added 0.5 μ L BsaI-HFv2 (NEB # R3733). For BsmBI GG reactions we used 0.5 μ L of either BsmBI (NEB) or FastDigest Esp3I (Thermo Scientific FD0454 NEB) (both 10 000 U/mL).

Thermocycler Protocols. The GG protocol is primarily used for assembly reactions. The reaction temperature is initially held at 45 °C for 2 min to digest the plasmids followed by 20 °C for 4 min to anneal constituent parts together. After repeating these first two steps 24 times, the temperature is increased to 60 °C for 10 min to digest the remaining recognition sites and inactivate the ligase. Then the temperature is held at 80 °C for 10 min to inactivate the enzyme. Lastly, the reaction is held at 12 °C indefinitely. The “GG End-On” protocol is used when BsaI or BsmBI sites need to be retained in the final product. The temperature is initially held at 45 °C for 2 min to digest the plasmid followed by 20 °C for 5 min to anneal and ligate the resulting plasmid. These steps are cycled 24 times and then held at 16 °C indefinitely.

Domestication of Parts. Forward and reverse primers ordered from IDT (www.idtdna.com) were manually designed to anneal to source DNA. See [Supporting Table S3](#) for part-specific design of domestication primers. In summary, internal BsaI and BsmBI sites were removed and tandem BsaI and BsmBI sites were appended to both the 5' and 3' ends of the sequence (see [Supporting Tables S3 and S4](#)). PCR was performed with the general reaction mixture of 10 μ L Q5 polymerase master mix (NEB MO492S) using 1 μ L each of the forward and reverse primers (10 μ M) and 0.5 μ L of template DNA. The desired PCR product was gel extracted (ThermoFisher Gel Extraction K0691) and 1 μ L of the final elution added to a BsmBI-mediated GG reaction (described above) with the MTK0_027 domestication vector. The resulting product was transformed into bacteria and grown in

selective LB overnight. Plasmid DNA was extracted (ThermoFisher MiniPrep K0503) and sequenced verified.

Removal of Internal BsaI and BsmBI Sites. To remove an internal BsaI or BsmBI site when domesticating a coding DNA sequence part, overlapping forward and reverse primers are designed at the restriction site with a synonymous mutation made to ablate the internal site while preserving the coded amino acid. For noncoding parts, substitutions were made to maintain CG content when possible. Flanking the site are complementary BsmBI overhangs such that upon digestion of the resulting PCR products, the components anneal together.

Oligo Annealing. Oligos were designed such that the desired sequence was at least 15 bp long, and that when complemented the oligos generated the overhangs associated with the part number. To anneal oligos into double stranded DNA, we prepared a reaction mixture as follows: 1 μ L of each oligo (100 μ M), 1 μ L T4 Ligase Buffer (NEB), 1 μ L T4 PNK (NEB), 6 μ L water. Mixture was incubated at 37 °C for 1 h and then diluted to a volume of 200 μ L. To anneal the oligos, a thermocycler protocol was prepared to hold the temperature of the reaction at 96 °C for 6 min, and ramp down 0.1 °C per second to 23 °C. The reaction is then held at 23 °C indefinitely. In the case of sgRNA design, this final mix is then added to a BsmBI-mediated GG reaction into its corresponding sgRNA destination vector. In the case of multioligo assembly, each oligo is added to the reaction mixture with MTK0_027 in a BsmBI-mediated GG reaction as described above.

Geneblocks. Gene fragments were ordered from IDT as either whole constructs or partial constructs with complementary overhangs to ensure proper domestication.

TU Assembly. Part 1–5 plasmids were pooled together with a recipient Part 678 plasmid following a 2:1 molar ratio in a BsaI-mediated GG reaction as described above.

MTU Assembly. Constituent transcriptional unit plasmids were pooled in a 2:1 molar ratio with the destination vector in a BsmBI-mediated GG reaction as described above.

Mammalian Cell Culture and Transfection. HEK293T and 3T3 cells were maintained in DMEM (Dulbecco's Modified Eagle Medium, Gibco) supplemented with 10% Fetal Calf Serum (SAFC) and passaged every ~3 days.

Clonal cell lines of 293T cells carrying BxbI landing pad were obtained by single cell sorting (FACS Aria2) of cells expressing mRuby2, following transfection of parental cells with plasmids carrying CAS9 and 3 guide RNAs for hAAVS1(JPF0432) and landing pad (MTK0_057).

Transfections for transient expression were done in quadruplicate, in 96 well (5 $\times$ 10⁴ cells, 300 ng total plasmid DNA) or 6 well plate (1 $\times$ 10⁵ cells, 4 μ g total plasmid DNA) format with Lipofectamine 2000 (Invitrogen), according to manufacturer's instructions. CAS9, PiggyBac, and BxbI transfections were performed in triplicate with Lipofectamine 2000, according to the manufacturer's instructions. In summary, transfections were performed with 1:1 ratio of CAS9/transposase/recombinase (JPF0432/pCMV-hyPBBase/pCAG-NLS-HA-Bxb1) to transfer vector, for a total of 800 ng of DNA per well in 24 well plates seeded with 5 $\times$ 10⁴ cells. The next day following transfection, media was replaced with normal growth media with the appropriate antibiotic selection. Hygromycin (100 μ g/mL, Invivogen) and Blasticidin (10 μ g/mL, Neta Scientific) selection was performed for at least 7 days, with media changes every 48 h. pCMV-hyPBBase was a gift from Wendell Lim and pCAG-NLS-HA-Bxb1 was a gift

from Pawel Pelczar (Addgene plasmid # 51271; <http://n2t.net/addgene:51271>; RRID: Addgene_51271).

Transfections for Zaire ebolavirus RNP assays were performed in duplicate or triplicate using standard calcium phosphate transfection methods. Briefly, complexes were generated in 100 μ L of 1 $\times$ Hepes Buffered Saline (HBS) (Thermo AAJ62623AK) and 12.5 μ L of 2 M calcium chloride (Fisher Scientific 50995817) per 1 μ g of DNA for 15 min at room temperature prior to transfection of cells.

Lentiviral Production. For lentiviral generation, 24 h before transfection, SE5 HEK293T cells were plated in a 6 well plate containing 2 mL of growth media. To prepare virus, transfections of 4 μ g total DNA of equimolar amounts of pCMV-dR8.91, pCMV-VSV-G, and transfer vector were performed using Lipofectamine 2000, and following the manufacturer's instructions. Media was changed after 16 h and virus were collected 48 h after transfection, by filtering the supernatant through a 0.45 μ m filter (Millipore SLHV033RS).

For transductions, 24 h before addition of virus target cells were seeded at a density of 1×10^5 in a 6 well plate and transduced with 1 mL, 100 or 10 μ L of viral supernatant supplemented with 4 μ g/mL of Polybrene (SCBT sc-134220).

pMD2.G was a gift from Didier Trono (Addgene plasmid # 12259; <http://n2t.net/addgene:12259>; RRID: Addgene_12259) and pCMV-dR8.91 was a gift from Wendell Lim.

Flow Cytometry and Data Analysis. For the analysis of promoter, 3' UTR expression and hAAVS1 LP expression, cells were collected in 96 well plates (Corning) and measured using an LSR2 flow cytometer (BD) with the four laser configuration (488, 635, 355, 405 nm). mAzamiGreen (excitation at 488 nm, emission at 530 nm), mRuby2 or mScarlet (excitation at 561 nm, emission between 610 and 620 nm) and tagBFP (excitation at 355 nm, emission at 450 nm) fluorescence levels were recorded for 10 000 events. Gating of single cells, normalization of fluorescence levels, and statistical analysis was performed with custom python scripts (<https://github.com/jpfon/MTK>).

Microscopy and Image Processing. For the imaging of combinatorial assembly of the fluorescent protein and localization tags library, SE4 HEK293T cells were plated in an 8-Well μ -Slide (Ibidi) that contains 200 μ L of growth media. After 24 h and before imaging, growth media was replaced with 200 μ L of Fluorobrite DMEM (Gibco). Imaging was performed in a temperature and atmosphere controlled chamber on a Zeiss microscope equipped with a Yokagawa CSUX1-A1N-E confocal spinning disk. Images were collected with a 40 $\times$ 1.1 NA water immersion objective and Photometrics Evolve 512 EMCCD camera. Images were stitched (ZEN, Zeiss) and gamma-corrected (FIJI) for perception enhancement.

For the imaging of multicistronic fluorescent proteins, 45E4 HEK293T cells were plated in 8-Well μ -Slide (Ibidi) containing 200 μ L of growth media. Imaging was performed in a temperature and atmosphere controlled chamber on a Nikon Ti Microscope equipped with a Andor Borealis CSU-W1 confocal spinning disk. Images were collected through a 20 $\times$ 0.75 NA air objective, using an Andor 4 Laser Launch for tagBFP (excitation at 405 nm, collection between 425 and 475 nm), mAzamiGreen (excitation at 488 nm, collection between 500 and 550 nm), mScarlet (excitation at 561 nm, collection between 590 and 650 nm) and iRFP713 (excitation at 640 nm, collection between 665 and 736 nm). An Andor Zyla 4.2

sCMOS was used to detect the images and pixel size was 325 nm.

Viral DNA Sequences and Plasmids. Synthetic cDNA sequences for viral NP, VP35, VP30, and Lpol from Zaire ebolavirus Mayinga 1976 isolate (Accession: NC_002549, H.sapiens-tc/COD/1976/Yambuku-Mayinga) were codon optimized (NP, VP35, VP30) and synthesized by IDT, followed by Gibson assembly into the pCAGGs vector backbone. These yielded pCAGGs-NP, pCAGGs-VP35, pCAGGs-VP30, and pCAGGs-Lpol expression plasmids. In order to generate the 4-cistronic (ZEBOV-4cis), viral proteins were PCR amplified from pCAGGs expression vectors and domesticated into MTK0_027 entry vector using the BsmBI site followed by verification with Sanger sequencing. The T7-driven viral minigenome construct p2.0_3ESE_eGFP was a gift from Elke Mühlberger (Addgene plasmid # 69359; <http://n2t.net/addgene:69359>; RRID: Addgene_69359), which was modified *via* NdeI and NotI restriction sites to generate p2.0–3ESE-nLuc. T7opt in pCAGGS was a gift from Benhur Lee (Addgene plasmid # 65974; <http://n2t.net/addgene:65974>; RRID: Addgene_65974).

Compounds. Gedunin (CAS 2753-30-2) (Fisher Scientific # 33871) and 6-Azauridine (Sigma # A1882-1G) were resuspended in DMSO to generate a 1 mM stock solution. Stock solutions were serially diluted in DMSO and media to treat cells with a final concentration of 5 μ M in 1% DMSO.

Generation of ZEBOV-4cis Stable Cell Populations and Clones. HEK293T cells were seeded in 6-well plates at a density of 0.4×10^6 cells/well for 24 h and then transfected with 2 μ g of total DNA (1 μ g of MTK043-ZEBOV-4cis construct and 1 μ g of PiggyBac transposon expression construct (pCMV-hyPBBase)) using calcium phosphate transfection. After 2 days cells were transferred to a 10 cm dish to begin selection with hygromycin B (0.4 mg/mL) (Fisher Scientific # MIR5930). Complete selection was observed in 7 days resulting in ZEBOV-4cis expressing stable populations from which clones were isolated using limited dilution plating. Briefly, cells were plated into 96-well plates at a density of 0.5 cells/well in the presence of hygromycin for 10 days and subsequently expanded to yield stable clones.

Western Blot and Immunofluorescence Analyses. Cell lysates were prepared in RIPA lysis and extraction buffer (Thermo # 89900) containing protease inhibitors (Sigma # P8340) for Western blotting. Lysates were resolved by SDS-polyacrylamide gel electrophoresis (PAGE), transferred to a polyvinylidene difluoride (PVDF) membrane, and subjected to Western blotting using primary antibodies: rabbit polyclonal Zaire ebolavirus NP (IBT Bioservices # 0301-045) (1:1000 dilution), mouse monoclonal Zaire ebolavirus VP35 (Kerafast # EMS702) (1:1000 dilution), mouse monoclonal 2A peptide (Novus Biologicals # NBP2-59627) (1:2000 dilution), Rabbit polyclonal GAPDH (Thermo Fisher Scientific # PA1-987) (1:5000 dilution), and secondary antibodies: goat antirabbit or goat antimouse polyclonal IRDye-800CW antibodies (VWR, 1:5000 dilution). For immunofluorescence analysis, cells seeded in 12-well tissue culture plates for 24 h were fixed in 4% formaldehyde for 20 min, incubated in permeabilization buffer (1% [vol/vol] Triton X-100 and 0.1% [wt/vol] sodium citrate in PBS) for 10 min and then in blocking buffer (1% [vol/vol] Triton X-100, 0.5% [vol/vol] Tween 20, and 3% bovine serum albumin in PBS) for 30 min. Cells were then incubated overnight with primary antibodies for Zaire ebolavirus VP35 or 2A peptide, secondary antibodies goat

antimouse or goat antirabbit Alexa-488 for 30 min, and stained with DAPI solution (GeneTex # GTX16206) for 10 min. Cells were imaged at 10× (Leica light microscope).

Minigenome Reporter Assays (Luciferase, GFP). ZEBOV minigenome reporter construct (p2.0–3E5E-nLuc, 250 ng) and T7-expression plasmid (pCAGGs-T7opt, 250 ng) were co-transfected with four ZEBOV expression plasmids (pCAGGs-ZEBOV-NP, pCAGGs-ZEBOV-VP35, pCAGGs-ZEBOV-VP30, pCAGGs-ZEBOV-Lpol; 250 ng each) or multicistronic constructs (ZEBOV-4cis, 1000 ng), and pCAGGs empty vector for a total of 1500 ng DNA per 0.5E6 HEK293T cells using calcium phosphate transfection (see method above). Suspension transfected cells were seeded at a density of 20 000 cells/well into 8 wells of 96-well plates for 2 days in duplicates. Nano luciferase assays for minigenome function and cell titer glo assays for cell viability were performed as per manufacturer's instructions (Promega). For minigenome assays in ZEBOV-4cis stable cells, 500 ng each of p2.0-T7–3E5E-nLuc or p2.0-T7–3E5E-eGFP and 500 ng of T7-expression plasmid were cotransfected. Nano luciferase levels were assayed as above and eGFP signal was captured *via* microscopy (Leica, 4× magnification).

Statistical Analysis. Bar plots show mean of biological replicates and, when shown, error bars denote 95% confidence intervals of mean. Statistical analysis was done in Python (<https://github.com/jpfon/MTK>).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acssynbio.9b00322](https://doi.org/10.1021/acssynbio.9b00322).

Supporting Figures S1–S5, Tables S1–S4, including library contents (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: amy.kistler@czbiohub.org.

*E-mail: hana.el-samad@ucsf.edu.

ORCID

João Pedro Fonseca: 0000-0002-4465-7165

Hana El-Samad: 0000-0001-6239-9916

Present Address

[†]Zymergen, Emeryville, California 94608, United States.

Author Contributions

[†]J.P.F. and A.R.B. contributed equally. J.P.F., A.R.B., and H.E.-S. conceived of the study. A.L.K. and G.R.K. conceived of the virology application. J.P.F., A.R.B., G.R.K., A.H.N., J.T., Q.C.W., E.A., S.Y.C., G.D., P.H., and L.C.O. constructed the parts library. J.P.F., A.R.B., and G.R.K. collected and processed the data. J.P.F., A.R.B., G.R.K., A.L.K., and H.E.-S. interpreted the results, wrote and edited the manuscript.

Notes

The authors declare no competing financial interest.

All supporting data for the findings and plasmids presented in this study are available on github (<https://github.com/jpfon/MTK>).

■ ACKNOWLEDGMENTS

The authors thank the members of the El-Samad lab and Joe DeRisi (UCSF) for helpful feedback, and John Dueber (UC

Berkeley) for the gift of constructs used in this study. This work was supported by the National Science Foundation Award # 1715108 to H.E.-S and the Defense Advanced Research Projects Agency [grant number HR0011-16-2-0045 to H.E.-S]. The content and information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. H.E.-S is an investigator in the Chan Zuckerberg Biohub. This work was also supported by the National Defense Science & Engineering Graduate (NDSEG) Fellowship awarded to A.R.B., A.H.N., and L.C.O. L.C.O. is also supported by the Paul and Daisy Soros Fellowship for New Americans and National Institute for General Medical Sciences (NIGMS) Initiative for Maximizing Student Development (IMSD) Fellowship. Research performed by G.R.K. and A.L.K. is funded by the Chan Zuckerberg Biohub.

■ REFERENCES

- (1) Engler, C., et al. (2014) A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843.
- (2) Hernanz-Koers, M., et al. (2018) FungalBraid: A GoldenBraid-based modular cloning platform for the assembly and exchange of DNA elements tailored to fungal synthetic biology. *Fungal Genet. Biol.* 116, 51–61.
- (3) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* 4, 975–986.
- (4) Pérez-González, A., et al. (2017) Adaptation of the GoldenBraid modular cloning system and creation of a toolkit for the expression of heterologous proteins in yeast mitochondria. *BMC Biotechnol.* 17, 80.
- (5) Pollak, B., et al. (2019) Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New Phytol.* 222, 628–640.
- (6) Halleran, A. D., Swaminathan, A., and Murray, R. M. (2018) Single Day Construction of Multigene Circuits with 3G Assembly. *ACS Synth. Biol.* 7, 1477–1480.
- (7) Akama-Garren, E. H., et al. (2016) A Modular Assembly Platform for Rapid Generation of DNA Constructs. *Sci. Rep.* 6, 16836.
- (8) Duportet, X., et al. (2014) A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res.* 42, 13440–13451.
- (9) Martella, A., Matjusaitis, M., Auxillos, J., Pollard, S. M., and Cai, Y. (2017) EMMA: An Extensible Mammalian Modular Assembly Toolkit for the Rapid Design and Production of Diverse Expression Vectors. *ACS Synth. Biol.* 6, 1380–1392.
- (10) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6, No. e16765.
- (11) Liu, M., et al. (2015) Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol.* 33, 198–203.
- (12) Szymczak, A. L., and Vignali, D. A. A. (2005) Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin. Biol. Ther.* 5, 627–638.
- (13) Liu, Z., et al. (2017) Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.*, DOI: [10.1038/s41598-017-02460-2](https://doi.org/10.1038/s41598-017-02460-2).
- (14) Hermann, M., et al. (2014) Binary recombinase systems for high-resolution conditional mutagenesis. *Nucleic Acids Res.* 42, 3894–3907.
- (15) Yusa, K., Zhou, L., Li, M. A., Bradley, A., and Craig, N. L. (2011) A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U. S. A.* 108, 1531–1536.
- (16) Nalaskowski, M. M., Ehm, P., Giehler, S., and Mayr, G. W. (2012) A toolkit for graded expression of green fluorescent protein fusion proteins in mammalian cells. *Anal. Biochem.* 428, 24–27.

- (17) Regot, S., Hughey, J. J., Bajar, B. T., Carrasco, S., and Covert, M. W. (2014) High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* 157, 1724–1734.
- (18) Smith, J. R., et al. (2008) Robust, Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration. *Stem Cells* 26, 496–504.
- (19) Sadelain, M., Papapetrou, E. P., and Bushman, F. D. (2012) Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* 12, 51–58.
- (20) Doench, J. G., et al. (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191.
- (21) Moreno-Mateos, M. A., et al. (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* 12, 982–988.
- (22) Yeo, N. C., et al. (2018) An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods* 15, 611–616.
- (23) Kiani, S., et al. (2015) Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* 12, 1051–1054.
- (24) Margolin, J. F., et al. (1994) Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A.* 91, 4509–4513.
- (25) Chavez, A., et al. (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328.
- (26) Qi, L. S., et al. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183.
- (27) Emanuel, J., Marzi, A., and Feldmann, H. (2018) Filoviruses: Ecology, Molecular Biology, and Evolution. *Adv. Virus Res.* 100, 189–221.
- (28) Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999) Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–2342.
- (29) Hoenen, T., and Feldmann, H. (2014) Reverse genetics systems as tools for the development of novel therapies against filoviruses. *Expert Rev. Anti-Infect. Ther.* 12, 1253–1263.
- (30) Tao, W., Gan, T., Guo, M., Xu, Y., and Zhong, J. (2017) Novel Stable Ebola Virus Minigenome Replicon Reveals Remarkable Stability of the Viral Genome. *J. Virol.*, DOI: 10.1128/JVI.01316-17.
- (31) Brauburger, K., et al. (2014) Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. *J. Virol.* 88, 12558–12571.
- (32) Trunschke, M., et al. (2013) The L-VP35 and L-L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals. *Virology* 441, 135–145.
- (33) Groseth, A., et al. (2009) The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction identified. *Virus Res.* 140, 8–14.
- (34) Hoenen, T., et al. (2012) Inclusion bodies are a site of ebolavirus replication. *J. Virol.* 86, 11779–11788.
- (35) Edwards, M. R., et al. (2015) High-Throughput Minigenome System for Identifying Small-Molecule Inhibitors of Ebola Virus Replication. *ACS Infect. Dis.* 1, 380–387.
- (36) Uebelhoefer, L. S., et al. (2014) High-throughput, luciferase-based reverse genetics systems for identifying inhibitors of Marburg and Ebola viruses. *Antiviral Res.* 106, 86–94.
- (37) Lohmann, V., and Bartenschlager, R. (2014) On the history of hepatitis C virus cell culture systems. *J. Med. Chem.* 57, 1627–1642.
- (38) Woolhouse, M. E. J., et al. (2008) Temporal trends in the discovery of human viruses. *Proc. R. Soc. London, Ser. B* 275, 2111–2115.
- (39) Woolhouse, M., Scott, F., Hudson, Z., Howey, R., and Chase-Topping, M. (2012) Human viruses: discovery and emergence. *Philos. Trans. R. Soc., B* 367, 2864–2871.
- (40) Adamson, B., et al. (2016) A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* 167, 1867–1882.
- (41) Datlinger, P., et al. (2017) Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* 14, 297–301.
- (42) Dixit, A., et al. (2016) Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* 167, 1853–1866.
- (43) Yeung, E., et al. (2017) Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Syst* 5, 11–24.