Self-cutting and integrating CRISPR plasmids (SCIPs) enable targeted genomic integration of large genetic payloads for rapid cell engineering

Darin Bloemberg[1], Daniela Sosa-Miranda[1], Tina Nguyen[1], Risini D. Weeratna[1], and Scott McComb[1,2]*

[1]Human Health Therapeutics Research Center, National Research Council Canada, Ottawa, Ontario, Canada

[2]Department of Biochemistry, Microbiology, and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Canada

Running Title: Genome editing using single CRISPR plasmids

Keywords: Cas9, CRISPR, plasmid, genetic engineering, genome editing, cell engineering

*Corresponding Author:
Dr. Scott McComb
National Research Council Canada
1200 Montreal Road, Building M54
Ottawa, ON, Canada K1A 0R6
scott.mccomb@nrc-cnrc.gc.ca
613-993-9387

Contribution to the Field

We present here the first report of a novel, straightforward, and highly time- and cost-efficient system for targeted delivery of large synthetic DNA payloads into almost any desired site within a mammalian genome. By including a self-cleaving site within a well-known CRISPR plasmid, this single-component platform enables genomic integration by simple plasmid transfection and is easily re-programmable. Introducing self-cleaving plasmids to human and mouse cell lines produces similarly efficient gene knockout as non-cleaving plasmids, but results in stable integration of whole-plasmid DNA sequences into targeted genomic sites within a practical proportion of cells (up to 6%) which can then be selected and expanded. We demonstrate how self-cleaving and integrating plasmids (SCIP) can be modified to deliver specific transgenes to targeted locations by recombining sgRNA sequences with more than one self-cleavage site. We believe that pSCIP represents a valuable cell engineering tool that is much simpler to use and modify than existing genome engineering techniques involving in vitro DNA synthesis, virus handling and production, or expensive synthesis of long single-stranded DNA repair templates.

Abstract

Since observations that CRISPR nucleases function in mammalian cells, many strategies have been devised to adapt them for genetic engineering. Here, we investigated self-cutting and integrating Cas9/CRISPR plasmids (SCIPs) as easy-to-use gene editing tools that insert themselves at CRISPR-guided locations. Leaky sgRNA/Cas9 expression in bacteria initially prevented SCIP assembly and production; this was ameliorated by inserting a mammalian intron into the Cas9 gene. SCIPs demonstrated similar expression kinetics and gene disruption efficiency in mouse (EL4) and human (Jurkat) cells, with stable integration in 3-6% of transfected cells. Clonal sequencing analysis indicated that integrants showed bi- or mono-allelic integration of entire CRISPR plasmids in predictable orientations and with limited indel formation. Interestingly, including longer homology arms (HAs) (500 bp) in varying orientations only modestly increased knock-in efficiency (~2-fold), indicating that SCIP integration favours homology-independent mechanisms. Using a SCIP-payload design (SCIPay) which liberates a promoter-less sequence flanked by HAs thereby requiring perfect homology-directed repair (HDR) for expression, longer HAs resulted in higher integration efficiency and precision of the payload but did not affect integration of the remaining plasmid sequence. As proofs-of-concept, we used SCIPay to 1) insert a gene fragment encoding tdTomato into the CD69 locus of Jurkat cells, thereby creating a NOVEL cell line that reports T cell activation, and 2) insert a chimeric antigen receptor (CAR) gene into the T cell receptor alpha constant (TRAC) locus. Here, we demonstrate that SCIPs function as simple, efficient, and programmable tools useful for generating gene knockout/knock-in cell lines and suggest future utility in knock-in site screening/optimization, unbiased off-target site identification, and multiplexed, iterative, and/or library-scale automated genome engineering.

Introduction

Many systems of stably introducing synthetic transgenes of arbitrary sequence homology into mammalian cells exist. The most common of these, lentiviruses, are mainstays of cell biology and are used for such diverse functions as: expressing novel protein-coding genes, stably introducing shRNA, delivering barcoded Cas/CRISPR components, expressing fluorescent markers or fluorescent-tagged proteins, and gene therapy. Despite their widespread use, lentiviruses involve several limitations: their routine use requires high level technical training and appropriate facilities; their production requires high-quality/quantity plasmid preparations and

handling of extra cell lines; production scale-up is complicated by toxicity of lentiviral proteins; their genomic integration site is relatively uncontrollable; their envelope proteins are immunogenic; and their payload size and cell specificity are limited (Broussau et al. 2008; Sakuma, Barry, and Ikeda 2012; Vannucci et al. 2013). Some of these limitations have contributed to a recent global lentiviral shortages for gene and cell therapy trials and treatments (Kolata 2017).

Since observations that Type II CRISPR endonucleases (i.e. Cas9) function in mammalian cells, many strategies have been devised that exploit their reprogrammable capacity for inducing double-strand breaks (DSBs) for performing gene editing. This includes delivering donor sequences via AdV/AAV, high-yield PCR product, plasmid, or single stranded DNA oligonucleotides, and Cas/CRISPR components via plasmid, RNA, RNP, or replication-incompetent viruses (Jin et al. 2019; Ehrke-Schulz et al. 2017; Mangeot et al. 2019; Liu et al. 2018; Gao et al. 2018; Moore et al. 2015; Tálas et al. 2017). In addition to targeted gene editing for fixing germ-line mutation diseases, the current "killer app" for inserting synthetic transgenes is CAR-T, whereby T cells are programmed to target cancer cells using an engineered chimeric antigen receptor (CAR) (Bloemberg, McComb, and Weeratna 2019). Of note, several non-lentiviral methods have shown practical use for CAR-T engineering. The Sadelain group inserted a CAR into the TRAC locus by electroporating T cells with Cas9/gRNA RNPs and delivering the CAR transgene using rAAV6 (Eyquem et al. 2017). The Chen group also recently generated CAR-T cells by electroporating Cpf1 mRNA while delivering gRNA and donor DNA sequences via AAV, a method which enabled genetic disruption of a second location (Dai et al. 2019). The Marson lab similarly delivered Cas9/gRNA RNP into T cells and used high-output PCR products as donor templates to correct a pathological *IL2RA* defect and replace endogenous TCR genes with a cancer antigen-specific transgenic TCR (Roth et al. 2018).

An alternative option is programmable transposases, particularly the Sleeping Beauty and piggyBac systems (Mátés et al. 2009; Ivics et al. 1997; Geurts et al. 2003; Wilson, Coates, and George 2007). In fact, not only have both been used to generate high-quality CAR-T cells (Singh et al. 2013; Manuri et al. 2010), they are currently being examined in clinical trials (Kebriaei et al. 2016), demonstrating that robust T cell manufacturing is possible without lentiviruses. Furthermore, because this strategy uses plasmid DNA directly, it could enable a rapid simplification and efficiency in CAR-T production (Ramanayake et al. 2015).

Regardless of the specific technology, efficient and scalable methods for mammalian gene manipulation are relatively lacking but would be useful for cell line engineering, serial/repetitive genome editing, and identifying optimal genomic loci for transgene knock-in. Inspired by lentiviral vectors, we developed a plasmid-based system for targeted integration of synthetic transgenes at specific genomic loci by engineering Cas9/CRISPR plasmids to simultaneously cut genomic sequences and themselves: a strategy we call self-cutting and integrating plasmids, SCIP. We demonstrate that SCIPs represent easy-to-use genome engineering tools for generating knockout cell lines, fluorescent reporter knock-in cell lines, and inserting novel genetic payload sequences at varying genomic sites.

Materials and Methods

Cell Culture

Jurkat, EL4, Raji (ATCC) and Nalm6 (generously provided by Dr. Beat Bornhauser, University Children's Hospital Zurich) cells were cultured in standard RPMI media supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin, 1 mM sodium pyruvate, 2 mM L-glutamine, and 55 μM β-mercatpoethanol at 37ºC and 5% CO_2. HEK293T, SKOV3, and MCF7

cells (ATCC) were cultured in similar DMEM-based media. Cell counting was performed using an automated cell counter (Cellometer; Nexcelcom, Lawrence, MA) which assesses live/dead counts using acridine orange/propidium iodide staining.

Plasmids

The Cas9-CRISPR plasmid pSpCas9(BB)-2A-GFP (px458) (Addgene #48138) (Ran et al. 2013) served as the source of pSCIP development. The pQC ("quick-change") backbone plasmid containing a second golden gate cloning site for CRISPR bait insertion was created by cloning synthetic DNA oligonucleotides (Alpha DNA, Montreal, Canada) encoding tandem oppositional BsaI sites between the XbaI and KpnI sites of px458 (Supplemental Figure 1A). Similar to the existing BpiI-based golden gate cloning cassette in px458 for gRNA DNA insertion, this enables parallel or serial bait DNA insertion via BsaI-based Golden Gate assembly. An intron derived from the mouse immunoglobulin precursor V-region (GenBank ID: AH002574.2) was then inserted via Gibson cloning (300 ng digested plasmid and 10 ng DNA fragments with 1x in-house assembly mixture) into the middle of the SpCas9 coding sequence, as previously demonstrated to limit leaky SpCas9 expression in bacteria (Petris et al. 2017). This plasmid is denoted pQCi (quick change with intron).

CRISPR guides were designed using CRISPOR (http://crispor.tefor.net/) and potential sequences were curated manually (targeting early exons, high predicted efficiency and specificity, low/non-coding off-targets) (Supplemental Table 3). Corresponding complimentary DNA oligonucleotides were then synthesized (Alpha DNA, Montreal, Canada) and cloned using a one-step restriction/ligation reaction involving 300 ng plasmid DNA, 1 µM annealed oligonucleotides, 0.25 µL/2.5 U BpiI (ThermoFisher ER1012), 0.5 uL/0.5 U T4 ligase (ThermoFisher 15224017), and T4 ligase buffer (ThermoFisher 46300018). To create self-cutting plasmids, corresponding bait DNA sequences including PAM sites were inserted using a similar protocol involving 0.25 µL/2.5 U BsaI (ThermoFisher ER0291) (Supplemental Figure 1A). These restriction/ligation reactions involving annealed oligonucleotides are extremely efficient and result in successful cloning in greater than 95% of antibiotic-resistant clones.

SCIPs targeting mouse B2M with varying HA lengths and orientations were created using a similar approach (Supplemental Figure 1B). For inverted HAs, synthetic DNA fragments of the mouse B2M gene (GenBank ID: 12010, assembly GRCm38.p6) were constructed (Twist Bioscience, San Francisco, CA) in defined lengths surrounding the CRISPR target site (Supplemental Table 3) with capping sequences allowing their seamless cloning via BsaI-based restriction/ligation. For forward HAs, synthetic DNA fragments were similarly constructed (Twist Bioscience) with BsaI cloning caps such that the B2M genome sequences (not their reverse complements) on either side of the CRISPR site were exchanged for each other, thereby providing "normal" homology upon plasmid linearization (see Figure 5A). These reactions consisted of 300 ng plasmid DNA, 10 ng DNA fragments, 0.25 µL/2.5 U BsaI (ThermoFisher, Waltham, MA; ER0291), 0.5 uL/0.5 U T4 ligase (ThermoFisher 15224017), and T4 ligase buffer (ThermoFisher 46300018). Note that forward homology plasmids contain 17 or 6 base pair non-homologous caps resulting from Cas9 cleavage of the bait site.

To create the pSCIPay-BFP-reporter backbone plasmid, a DNA sequence was constructed *in silico* containing a P2A-BFP transgene without BpiI, BsaI, and BsmBI restriction sites (subsequently validated to maintain functionality) containing terminal BsaI cloning cassettes with distinct cloning sticky ends (Supplemental Figure 1C). This DNA fragment (Twist Bioscience) was inserted into BsaI-digested pQCi using Gibson assembly. This modular plasmid allows

insertion of gRNA-coding DNA oligonucleotides using BpiI as outlined above and subsequent cloning of HA- and bait-coding DNA fragments engineered to be in-frame with the preceding genome sequence and following P2A-BFP sequence with BsaI using Gibson or restriction/ligation cloning.

SCIPay-human-CD69-tdTomato and SCIPay-human-TRAC-CAR were generated by first cloning gRNA-coding DNA oligonucleotides as described above and subsequently inserting DNA fragments (Twist Bioscience) designed *in silico* with gRNA-specific bait sequences into pQCi using Gibson assembly. EGFR-targeting CAR utilized a novel EGFR-targeting sequence under development at the National Research Council Canada.

Plasmid DNA was purified after overnight culture in standard lysogeny broth containing 100 µg/mL ampicillin using a miniprep plasmid DNA isolation kit (Sigma Aldrich, St Louis, MO; PLN350) and quantified using spectrophotometry (NanoDrop One, ThermoFisher). Plasmid construction was always validated to *in silico* sequences using Sanger sequencing.

Transformation and Colony Counts

Standard *Escherichia coli* DH5α chemically competent cells generated using an in-house protocol from an in-house stock were used for most experiments. Stbl3 (ThermoFisher) and BL21 (New England Biolabs, Ipswich, MA) cells were utilized when indicated. Briefly, 50 µL freshly thawed cells were incubated on ice with 10 ng plasmid DNA for 30 minutes, heat shocked at 42ºC for 35 seconds, and placed on ice for 2 minutes. Cells were recovered by adding 950 µL warmed SOC media and incubating at 37ºC while shaking vigorously for 1 hour. From this solution, 100 µL was plated on ampicillin-containing agar plates and incubated overnight at 37ºC. Colony counts were performed the following morning and reported values represent actual numbers of colonies, not normalized for dilutions.

Transfection

EL4 and Jurkat cells were transfected via electroporation according to a previously-outlined protocol (Chicaybam et al. 2013). After collection, $5x10^5$ – $1x10^6$ cells were suspended in 100 µL Buffer 1SM (5 mM KCl, 15 mM $MgCl_2$ 120 mM Na_2HPO_4/NaH_2PO_4, 25 mM sodium succinate, and 25 mM mannitol; pH 7.2) and briefly (<5 minutes) incubated with appropriate plasmid DNA (2-5 µg). This solution was transferred into 0.2 cm generic electroporation cuvettes (Biorad Gene Pulser; Bio-Rad laboratories, Hercules, CA) and immediately electroporated using a Lonza Nucleofector I (Lonza, Basel, Switzerland) and program X-05 (X-005 on newer Nucleofector models). Cells were cultured in pre-warmed recovery media (RPMI with 20% FBS, 2 mM L-glutamine and 1 mM sodium pyruvate) overnight before being transferred to normal culture media.

Flow Cytometry

Transfection, Knock-out, and Knock-in Detection

Transfection efficiencies were evaluated on cells 24 and 48 hours post-transfection. Cells were collected from culture, centrifuged for 5 minutes at 300g, re-suspended in flow cytometry staining buffer (FSB, PBS with 1% FBS, 10 mM HEPES, and 2 mM EDTA), and analyzed for GFP using the FITC channel.

PD1 and B2M expression on EL4 cells were assessed with fluorescent antibody staining using an APC-conjugated PD1 antibody (BioLegend, San Diego, CA: 135210) or AlexaFluor 647-conjugated H-$2K^b$ antibody (BD Biosciences, San Jose, CA; 562832). Briefly, cells were collected

from culture, centrifuged at 300g for 5 minutes, re-suspended in antibody diluted in FSB (2.5 µL per million cells), incubated at room temperature for 30 minutes, centrifuged at 300g for 5 minutes, re-suspended in flow buffer, fixed with 0.5% formalin in PBS, and immediately analyzed. TCR expression on Jurkat cells was similarly assessed using a BV510-conjugated TCRαβ antibody (BD Bioscience 563625) or APC-conjugated CD3 antibody (BD Bioscience 561810).

As SCIP integration results in CMV-driven Cas9-GFP expression, knock-in events were detectable by simultaneous assessment of FITC fluorescence. Similarly, SCIPay-BFP-reporter and SCIPay-TRAC-CAR knock-in was assessed using the BV421 filter set and SCIPay-CD69-tDtomato was assessed using the PE filter set.

Jurkat Functional Testing

Jurkat activation during anti-CD3 administration or co-culture was assessed using flow cytometry detection of CD69 with an APC-conjugated anti-CD69 antibody (BD Bioscience 555533). Cells were not washed prior to or after staining: 0.2 µL antibody was added to 50 µL PBS per well of a 96-well plate, this diluted antibody solution was added to wells, cells were incubated for 30 min at 37ºC, formaldehyde was added to a final concentration of 0.5%, and the plate was immediately analyzed. For TRAC-CAR knock-in, Jurkat cells were distinguished from target cells via BFP (BV421).

Jurkat-CD69-tdTomato (J69) cell activation was assessed by detecting PE fluorescence and for J69-CAR, Jurkat cells were distinguished from target cells via tdTomato (PE) and GFP (FITC).

All flow cytometry analyses were performed using a BD LSR Fortessa configured with a high-throughput plate reader and the following lasers: 355 nm, 405 nm, 488 nm, 561 nm, and 640 nm.

Cell Line Generation Using SCIPs

Stable cell lines (clonal EL4-B2M-SCIP, J69, Jurkat-TRAC-CAR) were generated using high-speed fluorescence-based cell sorting using various staining and gating strategies as described in the text (MoFlo Astrios, Beckman Coulter, Mississauga, ON).

Genomic DNA Extraction, PCR, and Sequencing

Crude genomic DNA extracts were generated from clonal EL4 cells to confirm B2M-SCIP integration. Briefly, cell clones expanded to 50-100,000 cells were centrifuged, re-suspended in 100 µL lysis buffer (50 mM NaOH, 0.2 mM EDTA), incubated at 98ºC for 15 min, cooled to 4ºC, and 10 µL neutralization buffer (1 M Tris-HCl, pH 7.0) was added to each sample. DNA was then frozen at -80ºC for short-term storage.

PCR was performed using DreamTaq Green (ThermoFisher K1082) and combinations of the primers indicated in Table 1 to detect SCIP integration orientation and efficiency. Products from select reactions were then Sanger sequenced using either U6-F, CMVe-R, or B2M-R primers (Table S2).

Lentivirus Production and Jurkat Transduction

CD19-CAR lentivirus was generated using HEK293T cells. Briefly, 400,000 HEK cells were seeded in 6-well plates and the next day transfected with 1 µg psPAX2 (Addgene #12260), 3 µg pCMV-VSV.G (Addgene, Watertown, MA; #8454), and 4 µg CD19-CAR-GFP transfer plasmid (Bloemberg et al. 2020) with 24 µg linear PEI (Polysciences Inc., Warrington, PA) for 6 hours. Crude culture media supernatants containing virus particles were collected 72 hours later,

centrifuged at 500g for 10 min at 4°C to remove cell debris, and filtered through a 0.45 µm PVDF filter (Millipore/Sigma).

To make CAR-expressing Jurkat-CD69-reporter cells, 250,000 cells were incubated with a 1:1 mixture of this filtered viral supernatant and normal culture media containing 10 µg/mL Dextran for 48 hours. Cells were then washed and expanded in normal culture media. As the CAR transgene is followed by a P2A-GFP construct in the lentiviral transfer plasmid, all CAR-positive cells express GFP. Therefore, resting J69-CAR19 cells are CD69-negative, faintly tdTomato/PE-positive, and GFP-positive. Upon activation by co-culture with CD19 positive cells; these J69-CAR cells exhibit increased expression of tdTomato/PE.

Continuous Live-Cell Imaging

Continuous live-cell imaging (IncuCyte® S3, Sartorius) was used to assess J69 activation. After adding anti-CD3 antibody or CAR target cells, images were acquired every 30-60 minutes in light/phase and red (ex. 565-605 nm; em. 625-705 nm) fluorescence channels.

Statistics

Quantified results presented are means +/- standard error (SEM) and typically represent 3-5 independently run experiments; specific experiment descriptions are contained in respective figure captions.

Statistical comparisons were performed in appropriate situations where group size and replication numbers were not predetermined for power effects. T tests were calculated using Microsoft Excel and ANOVA analyses were performed using GraphPad Prism. Tukey's multiple comparison post hoc was used to identify locations of significance for ANOVAs. A p-value of 0.05 was considered statistically significant in all situations. Analysis specifics and descriptions are included in corresponding figure captions.

Results

Self-cutting and integrating plasmids (SCIP) as a non-viral strategy for mammalian genome editing

We hypothesized that including a "bait" site for the corresponding gRNA in a Cas9/CRISPR plasmid could generate an easy-to-modify and -manufacture single-component gene editing tool (Figure 1A). Upon transfection into mammalian cells, this plasmid would produce double-strand breaks (DSBs) and self-linearize, possibly leading to its integration (Figure 1A). This ability to stably integrate user-defined DNA payloads would mimic the function of lentiviral transgene delivery; however, we hypothesized that SCIPs could integrate at CRISPR-guided locations (Figure 1A).

We created a SCIP cloning plasmid by adding a second "Golden Gate" site to the popular GFP-expressing Cas9/CRISPR plasmid px458 (see Methods for details) to be used for inserting bait sequences (pQC; Supplemental Figure 1A). However, regardless of the specific guide+bait sequences used (Supplemental Table 3), SCIPs proved initially very inefficient to assemble and produce, as successful ampicillin-resistant transformant colony numbers were reduced ($p<0.05$) by 95% (Figure 1B) and colonies which did survive produced 90% less ($p<0.05$) plasmid DNA (Figure 1C).

Inserting an intron into the Cas9 sequence significantly improves self-cutting plasmid stability and production in bacteria

We theorized this instability was due to leaky Cas9 and gRNA production in the DH5α cells used for molecular cloning and plasmid production, which would cause plasmid degradation (Figure 2A). To prevent processing of these transcripts in bacteria cells, we inserted a mammalian intron into the Cas9 plasmid genetic sequence, a strategy previously used to increase the production of cleavable CRISPR components (Petris et al. 2017) (pQCi; Figure 2B and Supplemental Figure 1B). Plasmids targeting three different sites of the human TRAC gene were created with guide sequences, bait sequences, or both, in pQC and pQCi backbones (generating the 8 plasmid formats in Supplemental Figure 1B). Not only were transformant colony counts (Figure 2C), plasmid DNA yield (Figure 2D), and qualitative plasmid DNA quality (Figure 2E) unaffected ($p>0.05$) by the intron, but guide+bait plasmid instability was completely reversed in pQCi (Figure 2C & 2D). This improved stability of guide+bait-containing plasmids was also observed in Stbl3 and BL21 cell lines (Figure 2F & 2G), indicating that the phenomenon of leaky bacterial Cas9/sgRNA expression is not restricted to DH5α.

Self-cutting plasmids display similar expression and gene disruption efficiency as Cas9/CRISPR plasmids

The function and stability of self-cutting plasmids within mammalian cells was then tested by electroporating mouse EL4 cells with various plasmid formats. Here, GFP expression was similar ($p>0.05$) in pQC and pQCi empty, guide, and bait plasmids, demonstrating that transcript processing was unaffected by the intron (Figure 3A – 3C). Importantly, GFP expression produced by intron-containing pQCi plasmids with guide+bait sequences was significantly higher ($p<0.05$) compared to pQC plasmids and was similar to empty, guide, and bait counterparts (Figure 3A – 3C). When evaluating gene disruption efficiency for plasmids targeting beta-2 microglobulin (B2M) or programmed death receptor-1 (PD1), px458-guide, pQCi-guide, and pQCi-guide+bait produced similar percentages ($p>0.05$) of knockout cells (Figure 3D & 3E), again suggesting that the intron does not interfere with Cas9 and that self-cutting plasmids remain functional. The restored stability of self-cutting plasmids in bacteria cells and their apparently normal function in mammalian cells allowed us to further investigate SCIPs.

SCIPs integrate into CRISPR-guided locations

SCIP integration should result in stable expression of all plasmid ORFs. Here, this includes CMV-driven Cas9-eGFP, enabling simple knock-in detection using flow cytometry (Figure 4A). Using the same plasmids constructed targeting mouse B2M and PD1, the integration efficiency (the percentage of knockout cells that are GFP positive at 7 days after electroporation) increased ($p<0.05$) from <0.5% to 3% for B2M and from 2% to 6% for PD1 when EL4 cells were electroporated with SCIPs as compared to guide-only px458 and pQCi (Figure 4A & 4B). Similarly, the integration precision (the percentage of GFP positive cells that are knockouts) increased ($p<0.05$) from 50% to 80% for B2M and from 45% to 60% for PD1 with SCIPs compared to guide-only plasmids (Figure 4A & 4B). Importantly, these knock-out/knock-in events were stable and unchanging, as profiles 7 days post-transfection remained after 7 additional days (Supplemental Figure 3). Therefore, self-linearization enabled a 5-fold increase in DNA donation and 40% increase in the targeted precision of these editing events.

Non-homologous end joining (NHEJ) mediated integration of SCIPs should favour a specific integration orientation to prevent the formation of sgRNA target sites (Supplemental Figure 2B). Thus, to validate the presence and directionality of SCIP integration, EL4 cells were electroporated with B2M-targeted SCIPs and those displaying both B2M knockout and stable GFP

expression one week later were single-cell sorted. Genomic DNA extracted from these clonal cultures was then assessed using the PCR strategy outlined in Figure 4C and Supplemental Table 1. In addition to the predicted direction, 5 other possibilities causing simultaneous knockout and knock-in are also illustrated (Figure 4D). Of 36 measurable clones, 14 (39%, examples #13, 14, 16, 24 in Figure 4E) showed double-allele knock-in and 12 (33%, examples #17, 18, 21, 23 in Figure 4E) showed single-allele knock-in with indels likely present in the other allele (Table 1), representing the majority of clones. An additional 5 clones (14%, example #20 in Figure 4E) displayed wild-type bands and no B2M-specific integration and were therefore categorized as double-allele indels with assumed knock-in at a random location.

The PCR products from the 14 clones showing double-allele knock-in were additionally examined via Sanger sequencing to confirm the integration mechanism. Here, the side with 17 bp homology displayed more consistent ligation than the side with 6 bp homology (4 versus 6 clones with indels). However, the inverted homologous regions remained in all clones, providing important insight regarding the mechanism of these repair events. Overall, this shows that SCIPs integrate in a predictable orientation and with limited indel formation, thereby enabling targeted genomic integration of large genetic payloads (a whole plasmid of 9400 bp in this case) at efficiencies feasible for cell engineering.

Homology arm length and orientation minimally affect SCIP integration efficiency and precision

As the presence of short, homologous DNA sequences appeared to determine SCIP integration, we then constructed plasmids with increasing homology arm (HA) lengths surrounding the relatively efficient mouse B2M CRISPR site, expecting this to improve knock-in efficiency and precision. Here, instead of differently-sized micro-HAs (17 and 6 bp), HA length was created equal on each side of the plasmid bait site (Figure 5A, Supplemental Figure 1C). These plasmids were also generated in two formats: 1) inverted homology, in a similar orientation to that present in bait-containing plasmids (imagine the HAs in Supplemental Figure 2Biii getting longer), and 2) forward homology, in an overlapping orientation similar to typical DNA payloads designed for HDR (Figure 5A).

While there was a general trend for increasing B2M knockout efficiency in mouse EL4 cells with increasing HA length, there was no significant ($p>0.05$) difference compared to pQCi-guide and SCIP and no apparent effect of HA orientation (Figure 5B). Similarly, HA length and orientation did not affect knock-in precision, suggesting that 80% is the highest precision possible with this specific CRISPR gRNA (Figure 5C). With respect to knock-in efficiency, there was also a trend for increasing efficiency as forward HA length increased, although we noted that this did not dramatically increase knock-in efficiency compared to minimal bait-only SCIP, increasing only from ~1.5 to 3% with 500bp homology arms (Figure 5D). Overall, these data indicate that SCIP integration is likely primarily driven by a NHEJ-mediated blunt-end ligation of linearized plasmid into the genomic target site and is largely homology independent.

SCIPay design permits CRISPR-guided integration of specific donor DNA sequences via HDR

Our ultimate interest is enabling easier and higher-throughput directed gene editing, whereby user-defined DNA payloads can be inserted at specific locations or used to replace endogenous sequences. Therefore, we modified SCIP to contain a second bait site: by flanking these two bait sites around a DNA payload with HAs, we hypothesized that this plasmid should be able to deliver specific DNA sequences via homology directed repair (HDR), without integrating the remaining plasmid (Figure 6A, Supplemental Figure 1D). To test this strategy, we first created

a fluorescent reporter SCIPay plasmid containing a promoter-less P2A-BFP transgene, thereby allowing HDR-mediated editing to be detected by BFP production (Supplemental Figure 1D). This construct was given appropriate restriction sites enabling simultaneous insertion of both HA and bait sequences during a single cloning reaction. Continuing with the high-efficiency mouse B2M gRNA used above, SCIPays were created with different HA length (100, 200, 300, or 500 bp) in-frame with the mouse B2M genome flanked by bait sites that will yield short 6 bp scars facing the transgene payload (Figure 6A). SCIPay plasmids were equally ($p>0.05$) efficient as gRNA-only containing plasmids (QCi-guide) at producing B2M knockout cells (Figure 6B & 6H). While integration of the Cas9-eGFP plasmid portion occurred in up to 2% of cells transfected with bait-containing SCIPay, the efficiency and precision of these insertions remained consistent ($p>0.05$) over increasing HA length (Figure 6C & 6D). In contrast, the percentage of BFP-positive cells increased with longer HAs, leading to elevated ($p<0.05$) efficiency and precision scores (Figure 6E & 6F). The ratio of BFP- to GFP-positive cells with B2M knockout approached 1.5 with 500 bp HA, indicating that HDR-mediated editing involving the payload sequence occurred 50% more often than any NHEJ editing event for the remaining plasmid, wherein GFP expression is drivent by a constitutive promoter (Figure 6G). These results show that SCIPay can be used to insert specific gene sequences at specific genomic locations, but appropriate design and selection criteria are required to isolate desired genotypes.

Using SCIPay to generate a cell line that reports T cell activation

This HDR efficiency is sufficiently high to perform genetic cellular engineering with SCIPay in a manner typically performed using lentivirus transduction and antibiotic selection. To demonstrate its utility in this regard and generate a useful tool for our ongoing CAR screening and development efforts (Bloemberg et al. 2020), we designed a SCIPay construct containing a P2A-tdTomato transgene targeted to a site immediately following the human CD69 promoter (Figure 7A, Supplemental Figure 1E).

The Jurkat human T cell line was electroporated with this plasmid and then sorted, stimulated with anti-CD3 antibody, and re-sorted according to the flow diagram in Figure 7B, producing a polyclonal population (P1 or J69) with medium-low PE/tdTomato expression in 100% of cells. When incubated with anti-CD3 antibody to induce T cell activation, normal Jurkat cells displayed a dose-response increase in surface CD69 expression as assessed using flow cytometry, while neither the intermediate (G1) nor final (P1) reporter cells showed CD69 expression above baseline (Figure 7C). Conversely, P1 cells displayed a dose-response increase in PE/tdTomato with anti-CD3 administration that mimicked the CD69 response observed in normal Jurkat cells (Figure 7D & 7E), demonstrating the ability to properly report T cell activation.

While the polyclonal J69 cells produced consistent data, we attempted to derive a clonal J69 line hypothesizing it would further reduce variability. Here, 24 clones were isolated from the polyclonal J69 population and assessed for responsiveness (Supplemental Figure 4). While most clones (20 of 24) functioned similar to the polyclonal population (CD69 negative and increased PE upon anti-CD3/CD28 stimulation), 3 clones also displayed GFP-dependent responses and raw signal magnitudes varied significantly (Supplemental Figure 4). These findings highlight the need for proper screening strategies and emphasize the promiscuity and randomness of using Cas9 for genome editing.

The utility of the J69 reporter cell line was further investigated by transducing cells with a CD19-targeted CAR lentivirus (Bloemberg et al. 2020) and assessing their ability to report CAR-mediated T cell activation (Figure 7B). Here, J69-CAR19 cells increased PE/tdTomato expression

at increasing target cell ratios when cultured alongside CD19-positive Nalm6 (Figure 7F) or Raji (Figure 7G & 7H) leukemia cells, responses that were not observed in non-transduced J69. Overall, these results clearly demonstrate that SCIPay can be an effective and efficient tool for cell line engineering.

Using SCIPay to insert a CAR into the T cell receptor alpha constant (TRAC) gene

As a final proof-of-concept, a SCIPay construct was designed containing an anti-EGFR CAR followed by a P2A-BFP transgene targeted to the human TRAC locus (Figure 8A). Here, delivering a single plasmid causes TCR-knockout and stable CAR integration into a specific genetic site in HDR-edited cells which are indicated by BFP production. Similar to before, this plasmid's function was demonstrated by electroporating and sorting Jurkat cells (Figure 8B).

After sorting, a population (B1) was isolated that was 40% BFP- (BV421) positive and 90% TRAC/TCRβ-negative (Figure 8C). Notably, we consistently observe that WT Jurkat cells routinely cycle between 50-80% TCR positivity (see middle panel of Figure 8C) and here, the CAR-BFP is driven by the TRAC promoter and not a synthetic/constitutive promoter. Therefore, 40% is likely under-representative of the total CAR-positive cells in B1, as BFP-positive cells has been isolated from G1 (Figure 8A). Regardless, when co-cultured alongside two cancer cell lines with different EGFR expression (SKOV3, EGFR++; MCF7, EGFR+), B1 increased surface CD69 expression in response to increasing target cell ratios, while normal Jurkat cells did not (Figure 8D). This response also stratified by EGFR expression on target cells, as SKOV3 caused greater CD69 induction than MCF7 (Figure 8D).

Discussion

Numerous strategies for genome editing have been conceived, with several showing promise as scalable technologies with therapeutic intentions. Although many demonstrate robust function, these protocols typically employ relatively complex technical processes, including RNA, protein, and/or virus manufacturing and handling. Here, we tested two self-cleaving plasmid formats for their ability to perform gene editing, which enabled the delivery of all necessary components using a single DNA molecule without requiring virus/RNA/protein production.

Speaking to the excitement and interest in optimizing genome editing, several SCIP-like technologies have appeared in the literature during the execution of the work presented here. In particular, a series of experiments from Anja Ehrhardt's team at Witten/Herdecke University employ AdV/AAV for delivering HDR-compatible DNA editing machinery (Ehrke-Schulz et al. 2017; Gao et al. 2018; Bergmann et al. 2018; Schiwon et al. 2018). Using a canine hemophilia genetic model, they transfected a single plasmid (or transduced a single AdV) containing a Tet-ON Cas9 cassette and a donor DNA template flanked by gRNA bait sites into cultured cells to insert this donor repair sequence at the *cFIX* mutation site, thereby generating functional protein (Gao et al. 2018). Here, the repair sequence used alternate codons to avoid Cas9/sgRNA cleavage of itself, while the bait sites in the donor plasmid match the target sequence from the genome and the HA are simply the 200bp surrounding the target (Gao et al. 2018). Coincidentally, their HDR success rates (2-5%) match ours using SCIPay (Figure 6E). Notably, using a Tet-ON doxycycline-inducible promoter for Cas9 likely prevented plasmid degradation in bacteria cells as found here, which we reversed by inserting a mammalian intron. Although used for different purposes, the Gao paper essentially uses a vector similar to our SCIPay system that requires the targeted genetic mutation to reside within a Cas9 gRNA target site.

Other relatively straightforward AdV (Jin et al. 2019; Palmer, Turner, and Ng 2020) and other viral (Mangeot et al. 2019; Petris et al. 2017) methods for gene knock-in have also been described. A single-component AdV system, where sgRNA are copied from an existing plasmid and inserted into a Cas9-GFP AdV-Shuttle vector using Gateway, enabled efficient gene knockout (70% of cells with indels) after two rounds of transduction (Jin et al. 2019). A similar all-in-one helper-dependent AdV Cas9 vector allowed efficient HDR-mediated repair of mutated *Cftr* via an included donor DNA sequence (3500 successfully edited colonies per 2 million cells versus 30 without all-in-one Cas9) in iPSCs, although the vast majority of progeny possessed single-allele insertions with indels or large deletions (Palmer, Turner, and Ng 2020). Here, similar to the strategy described above, self-cleavage was prevented by altering the target sequences' PAM in the donor DNA payload (Palmer, Turner, and Ng 2020). In another example using novel virus-like particles exploiting a MLV-derived Gag:Cas9 fusion protein separated by a MLV protease site, indel formation in up to 70% of cells was achieved in numerous cell types, including iPSCs and mouse bone marrow cells (Mangeot et al. 2019). While particle production required transfecting HEK cells with four separate plasmids, including a fit-for-purpose sgRNA plasmid, particles for HDR-mediated knock-in were generated by complexing them with a synthetic ssDNA oligomer, thereby resulting in successful genome editing in HEK cells and iPSCs (Mangeot et al. 2019). Finally, intending to regulate Cas9 levels and function, Petris *et al* described a Cas9 expression plasmid containing a Cas9-targeting sgRNA with reduced off-target editing effects upon HEK transfection (Petris et al. 2017). After transferring this expression construct into a Tet-ON lentiviral vector with the same Cas9-intron which we employed in this study to allow plasmid production, the resulting lentiviral particles limited Cas9 expression levels and enabled equally efficient indel formation as normal Cas9 lentiviruses but with reduced off-target DNA cleavage (Petris et al. 2017).

Other examples of integrating plasmid systems also exist (Liu et al. 2018; Moore et al. 2015; Tálas et al. 2017). First, a two-plasmid system was utilized to evaluate the cutting efficiency of gRNA candidates involving: 1) a cleaved GFP plasmid containing a unique/non-genomic gRNA and corresponding bait site along with 2) a Cas9 plasmid containing the gRNA sequence of interest (Tálas et al. 2017). By detecting the NHEJ-mediated integration percentage of the linearized GFP-expressing DNA sequence (which occurred relatively frequently: up to 30% of cells), gRNA efficiency was evaluated (Tálas et al. 2017). Using a similar approach, self-targeting plasmids were shown to limit Cas9 delivery/stability within target cells, thereby limiting off-target effects, although direct integration was not examined using this system (Moore et al. 2015). Lastly, Liu *et al*. used two CRISPR/Cas9 or ZFN plasmids to insert donor DNA sequences from bait-containing plasmids at guide-targeted genomic sites, very similar to what we performed using a single plasmid (Liu et al. 2018). Here, similar to our own observations, integration events occurred with preferred directionality and genomic sequences resulting from CRISPR/Cas9 editing maintained the short regions of homology representing the bait sites (Liu et al. 2018). The use of a single plasmid for CRISPR editing and donor delivery as we have reported here should be significantly more amenable to high-throughput production.

Clearly, the diversity and creativity in exploiting programmable nucleases for conducting mammalian genome editing is impressive. While the strategies just described appear robust and function as intended, they all involve more than one component, with most requiring RNA, protein, and/or viruses. These entities are each far more difficult to generate and handle than plasmid DNA, in addition to the inherent additive effects with their production (ie. precursor products). Instead, SCIPs incorporate all the elements required for target-specific genome engineering without

requiring helper plasmids, RNA, protein, or synthetic DNA manufacturing, additional cell lines, virus production, or any other helper component. For simple knockout cell line generation, target-specific SCIPs are created by two sequential restriction/ligation reactions with short and simple annealed oligonucleotides, meaning unique SCIPs can be generated in 5 days with little hands-on time. These reactions are so efficient that we typically avoid Sanger sequencing confirmation and have begun performing them simultaneously. Once produced, SCIPs can be immediately transfected into cells which are then assessed for functional effects. In this scenario, using a SCIP-puromycin (instead of SCIP-GFP), the data in Figure 4 shows that 80% of the puromycin-resistant cells would be knockouts. While constructing SCIPays for genomic knock-in of user-defined payload sequences is slightly more complicated and costly, as this involves higher-level genetic design and larger synthetic oligonucleotide procurement, the process is similarly straightforward. Furthermore, pursuing a lentiviral strategy to achieve such an outcome still involves designing and constructing a fit-for-purpose transfer plasmid, while also requiring helper plasmid production, helper cell line culturing, and additional time and resources to generate lentiviral particles.

While the overall HDR-mediated genome editing success rates with SCIPay were relatively modest (3% with 500 bp HAs), this was enough to engineer useful cell lines, and further refinement to the genetic design and delivery protocol specifics are ongoing to optimize these for specific purposes. Regardless, in comparison to lentiviral integration, SCIPay can result in simultaneous knock-in and knock-out, such as through inserting a CAR into TRAC as we performed here, if such an outcome is desired. Another potential application for SCIPs is the ability to edit iteratively, as this process is non-viral and there is less chance for cells to build up immunity/resistance to subsequent engineering efforts. Additionally, while SCIP integration results in permanent Cas9 expression, this could be considered a feature instead of a flaw for specific applications, as subsequent editing events could be performed by delivering smaller plasmids or other components (ie. other sgRNA and donor DNA). In fact, making SCIP-Cas9 integration reversible could be achieved with a recombination system such as Cre-Lox. Lastly, the robustness and efficiency of SCIP cloning permits the creation of plasmid arrays and/or libraries that would function like typical CRISPR libraries delivered lentivirally. However, SCIP's unique function enables new opportunities in functional genomics: one potential application we suggest is using SCIP libraries to identify novel safe harbor sites.

Lastly, alongside SCIP's benefits as a modular DNA editing tool are potential safety considerations. While they are easier to create and use than lentiviruses, we treated SCIPs as BSL2 biohazards during their cloning and purification. Although genetic transfer via plasmid DNA should be significantly less efficient than via lentiviruses, SCIPs are essentially synthetic gene drives where their integration results in their production, thereby potentially leading to spreading to every available sgRNA site. Despite this, we do not believe that SCIPs pose safety risks beyond other forms of gene editing tools.

Conclusion

While DNA base editors (Gaudelli et al. 2017) and their "prime editing" counterparts (Anzalone et al. 2019) likely represent the immediate future for performing therapeutic genome editing in settings where very few nucleotides require replacement, there remains a need for simple and expandable tools with which to conduct genome engineering research. In addition to the uses described here (ie. making knock-out cell lines and inserting DNA payloads at targeted sites), SCIPs show promise in more advanced and unique gene-editing situations, such as identifying unbiased off-target and safe harbor sites and performing library-scale multiplexed and/or iterative

automated genome engineering. Therefore, we believe that SCIPs represent innovative technologies and intend that their presentation here will stimulate their use by the wider research community.

Acknowledgements

The authors wish to thank the National Research Council Canada's genomics core facility, including Dr Qing Yan Liu and Joy Lei for technical advice and performing Sanger sequencing runs.

Authorship Confirmation Statements

SCIP conception (SM, DB), study design and planning (all), conducting experiments and collecting data (DB, DSM, SM, TN), data analysis and figure preparation (DB, DSM, SM), and manuscript writing (DB, SM). All authors approved of the final version of this manuscript and figures prior to its submission.

This manuscript and the data contained herein are not published or currently under peer-review for publication elsewhere.

Author Disclosure Statements

All authors declare that they do not have a competing financial interest.

References

Anzalone, Andrew V., Peyton B. Randolph, Jessie R. Davis, Alexander A. Sousa, Luke W. Koblan, Jonathan M. Levy, Peter J. Chen, et al. 2019. "Search-and-Replace Genome Editing without Double-Strand Breaks or Donor DNA." *Nature* 576 (7785): 149–57. https://doi.org/10.1038/s41586-019-1711-4.

Bergmann, Thorsten, Eric Ehrke-Schulz, Jian Gao, Maren Schiwon, Verena Schildgen, Stephan David, Oliver Schildgen, and Anja Ehrhardt. 2018. "Designer Nuclease-Mediated Gene Correction via Homology-Directed Repair in an in Vitro Model of Canine Hemophilia B." *The Journal of Gene Medicine* 20 (5): e3020. https://doi.org/10.1002/jgm.3020.

Bloemberg, Darin, Scott McComb, and Risini Weeratna. 2019. "Building a Better CAR: Emerging High-Throughput in Vitro Tools for CAR Selection and Optimization." *Cell & Gene Therapy Insights* 5 (7): 681–92. https://doi.org/10.18609/cgti.2019.078.

Bloemberg, Darin, Tina Nguyen, Susanne MacLean, Ahmed Zafer, Christine Gadoury, Komal Gurnani, Anindita Chattopadhyay, et al. 2020. "A High-Throughput Method for Characterizing Novel Chimeric Antigen Receptors in Jurkat Cells." *Molecular Therapy - Methods & Clinical Development* 16 (March): 238–54. https://doi.org/10.1016/j.omtm.2020.01.012.

Broussau, Sophie, Nadine Jabbour, Guillaume Lachapelle, Yves Durocher, Rosanne Tom, Julia Transfiguracion, Rénald Gilbert, and Bernard Massie. 2008. "Inducible Packaging Cells for Large-Scale Production of Lentiviral Vectors in Serum-Free Suspension Culture." *Molecular Therapy: The Journal of the American Society of Gene Therapy* 16 (3): 500–507. https://doi.org/10.1038/sj.mt.6300383.

Chicaybam, Leonardo, Andressa Laino Sodre, Bianca Azevedo Curzio, and Martin Hernan Bonamino. 2013. "An Efficient Low Cost Method for Gene Transfer to T Lymphocytes." *PLoS ONE* 8 (3). https://doi.org/10.1371/journal.pone.0060298.

Dai, Xiaoyun, Jonathan J. Park, Yaying Du, Hyunu R. Kim, Guangchuan Wang, Youssef Errami, and Sidi Chen. 2019. “One-Step Generation of Modular CAR-T with AAV-Cpf1.” *Nature Methods* 16 (3): 247–54. https://doi.org/10.1038/s41592-019-0329-7.

Ehrke-Schulz, Eric, Maren Schiwon, Theo Leitner, Stephan Dávid, Thorsten Bergmann, Jing Liu, and Anja Ehrhardt. 2017. “CRISPR/Cas9 Delivery with One Single Adenoviral Vector Devoid of All Viral Genes.” *Scientific Reports* 7 (December). https://doi.org/10.1038/s41598-017-17180-w.

Eyquem, Justin, Jorge Mansilla-Soto, Theodoros Giavridis, Sjoukje J. C. van der Stegen, Mohamad Hamieh, Kristen M. Cunanan, Ashlesha Odak, Mithat Gönen, and Michel Sadelain. 2017. “Targeting a CAR to the TRAC Locus with CRISPR/Cas9 Enhances Tumour Rejection.” *Nature* 543 (7643): 113–17. https://doi.org/10.1038/nature21405.

Gao, Jian, Thorsten Bergmann, Wenli Zhang, Maren Schiwon, Eric Ehrke-Schulz, and Anja Ehrhardt. 2018. “Viral Vector-Based Delivery of CRISPR/Cas9 and Donor DNA for Homology-Directed Repair in an In Vitro Model for Canine Hemophilia B.” *Molecular Therapy. Nucleic Acids* 14 (December): 364–76. https://doi.org/10.1016/j.omtn.2018.12.008.

Gaudelli, Nicole M., Alexis C. Komor, Holly A. Rees, Michael S. Packer, Ahmed H. Badran, David I. Bryson, and David R. Liu. 2017. “Programmable Base Editing of A•T to G•C in Genomic DNA without DNA Cleavage.” *Nature* 551 (7681): 464–71. https://doi.org/10.1038/nature24644.

Geurts, Aron M., Ying Yang, Karl J. Clark, Geyi Liu, Zongbin Cui, Adam J. Dupuy, Jason B. Bell, David A. Largaespada, and Perry B. Hackett. 2003. “Gene Transfer into Genomes of Human Cells by the Sleeping Beauty Transposon System.” *Molecular Therapy: The Journal of the American Society of Gene Therapy* 8 (1): 108–17. https://doi.org/10.1016/s1525-0016(03)00099-6.

Ivics, Z., P. B. Hackett, R. H. Plasterk, and Z. Izsvák. 1997. “Molecular Reconstruction of Sleeping Beauty, a Tc1-like Transposon from Fish, and Its Transposition in Human Cells.” *Cell* 91 (4): 501–10. https://doi.org/10.1016/s0092-8674(00)80436-5.

Jin, Yuan-Hu, Hyunjeong Joo, Kwangjun Lee, Hyeongseok Kim, Ruth Didier, Young Yang, Heungsop Shin, and Choogon Lee. 2019. “Streamlined Procedure for Gene Knockouts Using All-in-One Adenoviral CRISPR-Cas9.” *Scientific Reports* 9 (1): 1–13. https://doi.org/10.1038/s41598-018-36736-y.

Kebriaei, Partow, Harjeet Singh, M. Helen Huls, Matthew J. Figliola, Roland Bassett, Simon Olivares, Bipulendu Jena, et al. 2016. “Phase I Trials Using *Sleeping Beauty* to Generate CD19-Specific CAR T Cells.” *The Journal of Clinical Investigation* 126 (9): 3363–76. https://doi.org/10.1172/JCI86721.

Kolata, Gina. 2017. “Gene Therapy Hits a Peculiar Roadblock: A Virus Shortage.” *The New York Times*, November 27, 2017, sec. Health. https://www.nytimes.com/2017/11/27/health/gene-therapy-virus-shortage.html.

Liu, Xiaofeng, Min Wang, Yufeng Qin, Xuan Shi, Peiqing Cong, Yaosheng Chen, and Zuyong He. 2018. “Targeted Integration in Human Cells through Single Crossover Mediated by ZFN or CRISPR/Cas9.” *BMC Biotechnology* 18 (October). https://doi.org/10.1186/s12896-018-0474-6.

Mangeot, Philippe E., Valérie Risson, Floriane Fusil, Aline Marnef, Emilie Laurent, Juliana Blin, Virginie Mournetas, et al. 2019. “Genome Editing in Primary Cells and in Vivo Using

Viral-Derived Nanoblades Loaded with Cas9-SgRNA Ribonucleoproteins." *Nature Communications* 10 (1): 1–15. https://doi.org/10.1038/s41467-018-07845-z.

Manuri, Pallavi V. Raja, Matthew H. Wilson, Sourindra N. Maiti, Tiejuan Mi, Harjeet Singh, Simon Olivares, Margaret J. Dawson, et al. 2010. "PiggyBac Transposon/Transposase System to Generate CD19-Specific T Cells for the Treatment of B-Lineage Malignancies." *Human Gene Therapy* 21 (4): 427–37. https://doi.org/10.1089/hum.2009.114.

Mátés, Lajos, Marinee K. L. Chuah, Eyayu Belay, Boris Jerchow, Namitha Manoj, Abel Acosta-Sanchez, Dawid P. Grzela, et al. 2009. "Molecular Evolution of a Novel Hyperactive Sleeping Beauty Transposase Enables Robust Stable Gene Transfer in Vertebrates." *Nature Genetics* 41 (6): 753–61. https://doi.org/10.1038/ng.343.

Moore, Richard, Alec Spinhirne, Michael J. Lai, Samantha Preisser, Yi Li, Taek Kang, and Leonidas Bleris. 2015. "CRISPR-Based Self-Cleaving Mechanism for Controllable Gene Delivery in Human Cells." *Nucleic Acids Research* 43 (2): 1297–1303. https://doi.org/10.1093/nar/gku1326.

Palmer, Donna J., Dustin L. Turner, and Philip Ng. 2020. "A Single 'All-in-One' Helper-Dependent Adenovirus to Deliver Donor DNA and CRISPR/Cas9 for Efficient Homology-Directed Repair." *Molecular Therapy - Methods & Clinical Development* 0 (0). https://doi.org/10.1016/j.omtm.2020.01.014.

Petris, Gianluca, Antonio Casini, Claudia Montagna, Francesca Lorenzin, Davide Prandi, Alessandro Romanel, Jacopo Zasso, Luciano Conti, Francesca Demichelis, and Anna Cereseto. 2017. "Hit and Go CAS9 Delivered through a Lentiviral Based Self-Limiting Circuit." *Nature Communications* 8 (May). https://doi.org/10.1038/ncomms15334.

Ramanayake, Saumya, Ian Bilmon, David Bishop, Ming-Celine Dubosq, Emily Blyth, Leighton Clancy, David Gottlieb, and Kenneth Micklethwaite. 2015. "Low-Cost Generation of Good Manufacturing Practice-Grade CD19-Specific Chimeric Antigen Receptor-Expressing T Cells Using PiggyBac Gene Transfer and Patient-Derived Materials." *Cytotherapy* 17 (9): 1251–67. https://doi.org/10.1016/j.jcyt.2015.05.013.

Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. 2013. "Genome Engineering Using the CRISPR-Cas9 System." *Nature Protocols* 8 (11): 2281–2308. https://doi.org/10.1038/nprot.2013.143.

Rees, Holly A., Wei-Hsi Yeh, and David R. Liu. 2019. "Development of HRad51–Cas9 Nickase Fusions That Mediate HDR without Double-Stranded Breaks." *Nature Communications* 10 (1): 1–12. https://doi.org/10.1038/s41467-019-09983-4.

Roth, Theodore L., Cristina Puig-Saus, Ruby Yu, Eric Shifrut, Julia Carnevale, P. Jonathan Li, Joseph Hiatt, et al. 2018. "Reprogramming Human T Cell Function and Specificity with Non-Viral Genome Targeting." *Nature* 559 (7714): 405–9. https://doi.org/10.1038/s41586-018-0326-5.

Sakuma, Toshie, Michael A. Barry, and Yasuhiro Ikeda. 2012. "Lentiviral Vectors: Basic to Translational." *Biochemical Journal* 443 (3): 603–18. https://doi.org/10.1042/BJ20120146.

Schiwon, Maren, Eric Ehrke-Schulz, Andreas Oswald, Thorsten Bergmann, Thomas Michler, Ulrike Protzer, and Anja Ehrhardt. 2018. "One-Vector System for Multiplexed CRISPR/Cas9 against Hepatitis B Virus CccDNA Utilizing High-Capacity Adenoviral Vectors." *Molecular Therapy - Nucleic Acids* 12 (September): 242–53. https://doi.org/10.1016/j.omtn.2018.05.006.

Singh, Harjeet, Matthew J. Figliola, Margaret J. Dawson, Simon Olivares, Ling Zhang, Ge Yang, Sourindra Maiti, et al. 2013. "Manufacture of Clinical-Grade CD19-Specific T Cells Stably Expressing Chimeric Antigen Receptor Using Sleeping Beauty System and Artificial Antigen Presenting Cells." *PLOS ONE* 8 (5): e64138. https://doi.org/10.1371/journal.pone.0064138.

Tálas, András, Péter István Kulcsár, Nóra Weinhardt, Adrienn Borsy, Eszter Tóth, Kornélia Szebényi, Sarah Laura Krausz, et al. 2017. "A Convenient Method to Pre-Screen Candidate Guide RNAs for CRISPR/Cas9 Gene Editing by NHEJ-Mediated Integration of a 'Self-Cleaving' GFP-Expression Plasmid." *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes* 24 (6): 609–21. https://doi.org/10.1093/dnares/dsx029.

Vannucci, Laura, Michele Lai, Flavia Chiuppesi, Luca Ceccherini-Nelli, and Mauro Pistello. 2013. "Viral Vectors: A Look Back and Ahead on Gene Transfer Technology." *The New Microbiologica* 36 (1): 1–22.

Wilson, Matthew H., Craig J. Coates, and Alfred L. George. 2007. "PiggyBac Transposon-Mediated Gene Transfer in Human Cells." *Molecular Therapy: The Journal of the American Society of Gene Therapy* 15 (1): 139–45. https://doi.org/10.1038/sj.mt.6300028.

Wolfs, Jason M., Thomas A. Hamilton, Jeremy T. Lant, Marcon Laforet, Jenny Zhang, Louisa M. Salemi, Gregory B. Gloor, Caroline Schild-Poulter, and David R. Edgell. 2016. "Biasing Genome-Editing Events toward Precise Length Deletions with an RNA-Guided TevCas9 Dual Nuclease." *Proceedings of the National Academy of Sciences of the United States of America* 113 (52): 14988–93. https://doi.org/10.1073/pnas.1616343114.

Figure Captions

<u>Figure 1. Self-cutting and integrating plasmid (SCIP) function and lack of bacterial stability.</u>
(A) Overview of theoretical SCIP function. Identification and cleavage of the bait site by the corresponding sgRNA-Cas9 complexes produced within mammalian cells linearizes the plasmid while performing a genomic DSB. In this way, a single, modular plasmid allows stable integration of user-defined DNA payloads, in a manner similar to lentiviruses. (B & C) Colony forming units (B) and mini-prep DNA yield (C) in SCIP compared to guide-containing CRISPR plasmids. Results in (B) and (C) represent means +/- SEM of 3 independent experiments. Asterisk (*) represents significant difference, as calculated using T-tests ($p<0.05$).

<u>Figure 2. Inserting an intron into the Cas9 sequence significantly improves SCIP stability and production in bacteria.</u>
(A) Proposed mechanism contributing to reduced SCIP stability in bacteria cells. Leaky CMV and hU6 promoters lead to Cas9/sgRNA complex production and subsequent plasmid cleavage. (B) Overview of SCIP backbone (pQC), plasmid, and intron structures. (C) Number of colonies formed from cloning reactions when creating various normal/Cas9 and intron-containing plasmids. (D) Plasmid DNA production from various normal/Cas9 and intron-containing plasmids. (E) Plasmid DNA quality as demonstrated in agarose gel. (F & G) Colony counts (F) and plasmid DNA yield (G) when using the indicated competent cell strains to generate guide+bait plasmids. Results in (C), (D), (E), and (F) represent means +/- SEM of 4-5 independent experiments. Asterisks (*) represent significant difference between Cas9 and $Cas9^{intron}$, as calculated using T-tests ($p<0.05$).

Figure 3. SCIPs display similar expression and gene disruption efficiency as CRISPR/Cas9 plasmids.

(A – C) The indicated plasmids were electroporated into mouse EL4 cells and assessed for GFP production using flow cytometry 24- (A & C) and 48-hours (B) after transfection. (D & E) Cells were similarly electroporated with the indicated plasmids containing gRNA targeting B2M or PD1 and analyzed for expression of the respective protein 7 days later. Results in (A), (B), and (D) represent means +/- SEM of 3 independent experiments. Asterisks (*) represent significant difference between Cas9 and $Cas9^{intron}$, as calculated using T-tests ($p<0.05$).

Figure 4. SCIPs integrate into CRISPR-guided locations.

(A & B) EL4 cells were electroporated with the indicated plasmids and assessed for gene knock-out and SCIP knock-in 7 days later. (A) FITC versus APC-[marker] flow cytometry plots allow evaluation of knock-out and knock-in events, quantified as efficiency and precision percentages (B). (C) EL4 cells electroporated with B2M-targeted SCIP were single-cell cloned and analyzed for knock-in configurations using the indicated PCR primers. (D) Illustration of possible SCIP integration configurations. (E) PCR DNA bands resulting from the indicated primers from clones #13-24. (F) Sanger sequencing results from measurable PCR reactions (B2M-F + CMVe-R or U6-F + B2M-R) produced from clones showing configuration #1 in (D). Results in (B) represent means +/- SEM of 4 independent experiments. Pound signs (#) represent significant difference between the indicated group and all other groups, as calculated using 1-way ANOVA ($p<0.05$).

Figure 5. Homology arm length and orientation minimally affect SCIP integration efficiency and precision.

(A) Overview of structural design for SCIPs with longer HAs. (B – D) EL4 cells were electroporated with SCIPs targeting B2M and assessed for knock-out/knock-in 7 days later. (B) Percentage of B2M knock-out cells and corresponding precision (C) and efficiency (D) scores. Results in (B), (C), and (D) represent means +/- SEM of 4 independent experiments.

Figure 6. SCIPay design permits CRISPR-guided integration of specific donor DNA sequences via homology directed repair.

(A) Overview of SCIPay design. Here, HA present in the payload sequence lead to its genomic integration via HDR. (B – G) EL4 cells were electroporated with SCIPay containing a BFP transgene targeting B2M with various HA lengths. (B) Percentage of knock-out cells. (C) Efficiency and (D) precision scores for GFP, denoting integration of the non-payload plasmid sequence. (E) Efficiency and (F) precision scores for BFP, denoting HDR integration of BFP. (G) Ratio of BFP- to GFP-positive cells. (H) Representative flow cytometry pltos. Results in (B – G) represent means +/- SEM of 4 independent experiments. Data was analyzed using 1-way ANOVA and significant differences are indicated with lower case letters, where groups with different letters are significantly different than each other ($p<0.05$).

Figure 7. Using SCIPay to generate a cell line that reports T cell activation.

(A) Overview of SCIPay designed to insert tdTomato into the human CD69 gene. (B) Outline of cell engineering strategy. After electroporation, successfully transfected cells were sorted, activated, and sorted again to enrich PE-responding cells. (C – E) The indicated cell populations were incubated with increasing anti-CD3 and assessed for surface CD69 (C) and tdTomato (D)

expression using flow cytometry. (E) Microscopy images of similar cells. (F – H) P1/J69 was transduced with an anti-CD19 CAR and cultured alongside CD19-positive cells. PE response to Nalm6 (F) and Raji (G & H) co-cultures. Results in (C – H) represent means +/- SEM of 3 independent experiments.

Figure 8. Using SCIPay to insert a chimeric antigen receptor (CAR) into the T cell receptor alpha subunit (TRAC) gene.
(A) Overview of SCIPay designed to insert an anti-EGFR CAR with BFP into the human TRAC gene. (B) Outline of cell engineering strategy. After electroporation, successfully transfected cells were sorted twice to enrich BFP-positive cells. (C) BFP and TCR/TRAC expression on double-sorted population (B1). (D) Surface CD69 expression during co-culture between the indicated cell population and SKOV3 or MCF7 cells. Results in (D) represent means +/- SEM of 3 independent experiments.

Tables

Table 1: EL4-B2M-SCIP Genomic PCR Results

	No bands	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**
Clone #	1,8,9,22	2,7,13,14,16,24,31, 32,33,39, 40,42,47, 48		15	3,10,11, 17,18,21, 23,26,27, 28,34,35		20,25,29, 38,41
Count	4	14	0	1	12	0	5

Made 48 clones; 36 grew

Supplemental Information

Figure Captions

Supplemental Figure 1. Overview and nomenclature of relevant plasmids.
(A) The widely used Cas9-GFP plasmid px458 served as the backbone for SCIP development, which began by inserting a second "golden-gate" restriction site (BsaI) for inserting bait sequences alongside the sgRNA cloning site (BpiI). (B) Combinations of guide- and/or bait-containing plasmids without (left side) or with (right side) the mammalian intron as used in Figures 2-4. Final "pSCIP" structure is pQCi-guide+bait: pQC with the intron, with the sgRNA, and with the bait sequence. (C) Structural depictions of pSCIPs with varying homology arm lengths and orientations, as used in Figure 5. (D) Structure of pSCIPay-BFP backbone/cloning plasmid, along with mouse B2M-targeting version, as used in Figure 6. (E) Structure of SCIPays used for cellular engineering in Figures 7 and 8.

Supplemental Figure 2. SCIP genomic integration mechanism possibilities.
(A) Upon Cas9/sgRNA production and RNP formation, DSBs are made in genomic (bottom) and plasmid (top) DNA sequences, producing the indicated genetic structures. This example illustrates the mouse B2M site/sgRNA used throughout the manuscript. (B) Potential integration mechanisms. (Bi) NHEJ orientation 1 re-creates the Cas9/sgRNA site and should be selected against. (Bii) NHEJ in the opposite orientation leads to SCIP integration and gene interruption.

(Biii) Due to CRISPR guide and bait structures, short sequences of inverted homology exist for NHEJ orientation 2, potentially contributing to site- and direction-specific integration. Figure 4 and Table 1 demonstrate this is in fact true.

Supplemental Figure 3. SCIP integration 14 days post-transfection.

This data supplements Figure 4B, where EL4 cells were electroporated with the indicated plasmids and assessed for gene knock-out and SCIP knock-in 14 days later. (A) Efficiency and (B) precision scores, similar to those presented in Figure 4B. Results represent means +/- SEM of 4 independent experiments. Pound signs (#) represent significant difference between the indicated group and all other groups, as calculated using 1-way ANOVA ($p<0.05$).

Supplemental Figure 4. Generating clonal J69 cell lines

To derive clonal J69 cells, the polyclonal population was single-cell sorted and 24 surviving clones were assessed for T cell responses. (A) FITC-GFP, PE-tdTomato, and CD69-APC responses after incubating cells with anti-CD3/CD28 beads. Data are expressed relative to non-stimulated cells. (B) Representative flow cytometry plots from select clones. (C) Raw green (FITC-GFP), red (PE-tdTomato), and CD69 (APC) mean fluorescence intensities.

Tables

Supplemental Table 1: EL4-B2M-SCIP Genomic PCR Possibilities

Primer Pair	**Size (bp)**	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**
B2M-F + B2M-R	345 (WT)	No	No	No	Maybe/ probably	Maybe/ probably	Maybe/ probably
U6-F + B2M-R (predicted)	558	Yes	No	Yes	Yes	No	No
U6-F + B2M-F (backwards)	541	No	Yes	Yes	No	Yes	No
CMVe-R + B2M-F (predicted)	410	Yes	No	Yes	Yes	No	No
CMVe-R + B2M-R (backwards)	428	No	Yes	Yes	No	Yes	No

Supplemental Table 2: PCR Primers

Name	**Sequence**
U6-F	CAAAGATATTAGTACAAAATACGTGACGTA
CMVe-R	GGTCATGTACTGGGCACAATG
B2M-F	ACTTTGGGGGAAGCAGATCAC
B2M-R	CAGTCTCAGTGGGGGTGAAT

Supplemental Table 3: CRISPR gRNA Target Sequences

Name	**Sequence**
Mouse B2M 1	CTGGTGCTTGTCTCACTGAC*CGG*
Mouse B2M 2	TCGGCTTCCCATTCTCCGGT*GGG*
Mouse B2M 3	CCGAGCCCAAGACCGTCTAC*TGG*

Mouse PD1	AGCGGGCATCCTGGACGGGT*TGG*
Human CD69 1	AGCTCTTTGCATCCGGAGAG*TGG*
Human CD69 2	TTACTTTCTTGTCCACTCTC*CGG*
Human CD69 3	AGAGATACCAGAGAGTAAAT*AGG*
Human TRAC 1	GGGAGTCAAAGTCGGTGAAC*AGG*
Human TRAC 2	GTTAAAAGATCCTCGGTCTC*AGG*
Human TRAC 3	CTCGACCAGCTTGACATCAC*AGG*

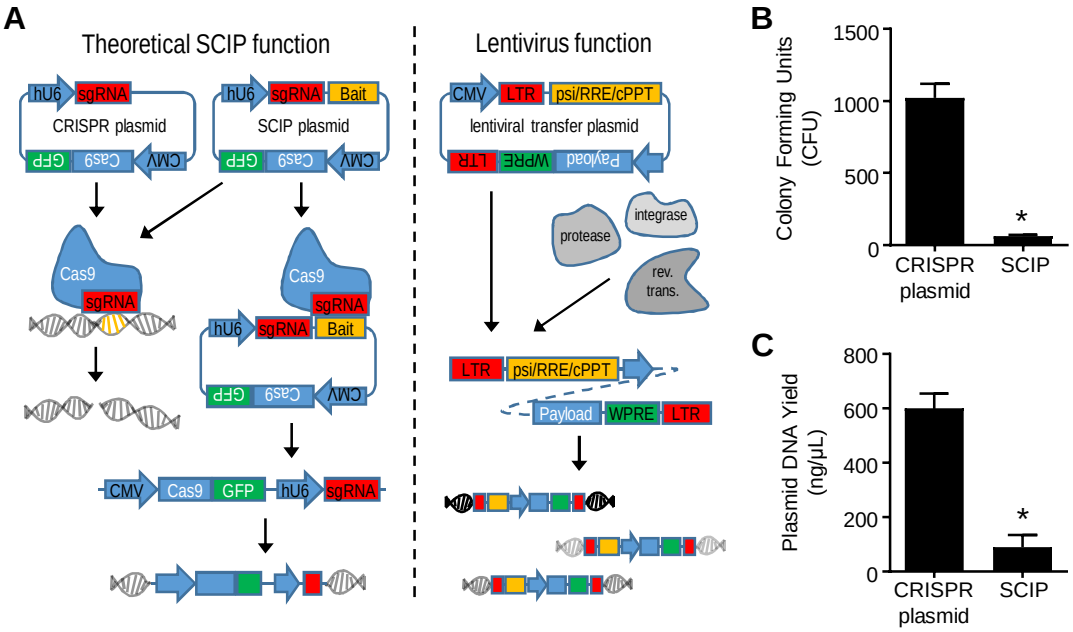

A
Theoretical SCIP function
hU6
sgRNA
CRISPR plasmid
CMV
Cas9
GFP
hU6
sgRNA
Bait
SCIP plasmid
Cas9
sgRNA
Cas9
sgRNA
hU6
sgRNA
Bait
CMV
Cas9
GFP
CMV
Cas9
GFP
hU6
sgRNA
Lentivirus function
CMV
LTR
psi/RRE/cPPT
lentiviral transfer plasmid
LTR
WPRE
Payload
protease
integrase
rev.
trans.
LTR
psi/RRE/cPPT
Payload
WPRE
LTR
B
Colony Forming Units
(CFU)
1500
1000
500
0
*
CRISPR
plasmid
SCIP
C
Plasmid DNA Yield
(ng/µL)
800
600
400
200
0
*
CRISPR
plasmid
SCIP

Figure 1

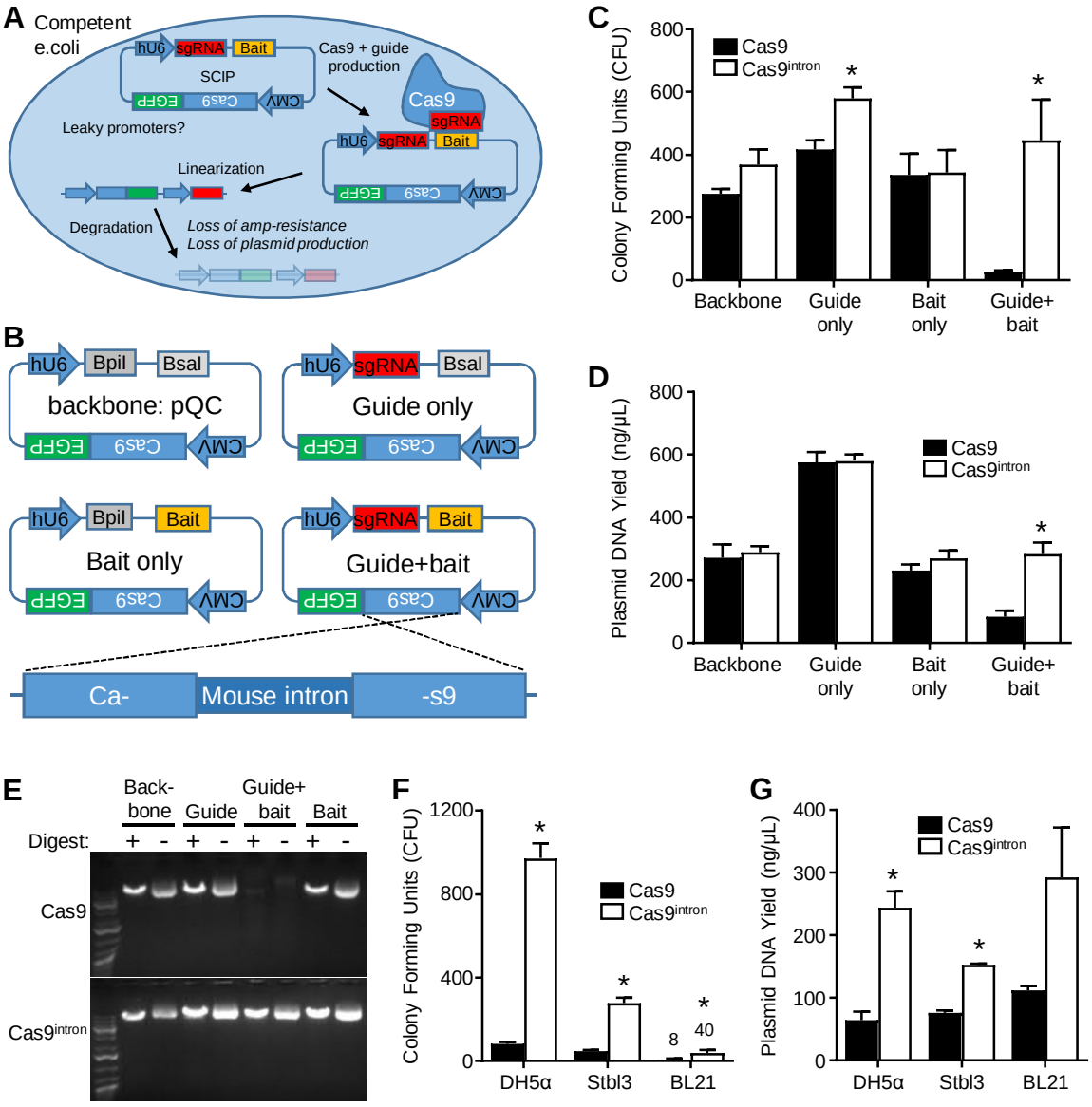

A
Competent e.coli
hU6
sgRNA
Bait
SCIP
Cas9
EGFP
CMV
Cas9 + guide production
Leaky promoters?
Linearization
Degradation
Loss of amp-resistance
Loss of plasmid production
B
hU6
BpiI
BsaI
backbone: pQC
Guide only
Bait only
Guide+bait
Ca-
Mouse intron
-s9
C
Colony Forming Units (CFU)
Cas9
Cas9intron
Backbone
Guide only
Bait only
Guide+ bait
D
Plasmid DNA Yield (ng/µL)
E
Back-bone
Guide
Guide+ bait
Bait
Digest:
+ - + - + - + -
Cas9
Cas9intron
F
Colony Forming Units (CFU)
DH5α
Stbl3
BL21
8
40
G
Plasmid DNA Yield (ng/µL)

Figure 2

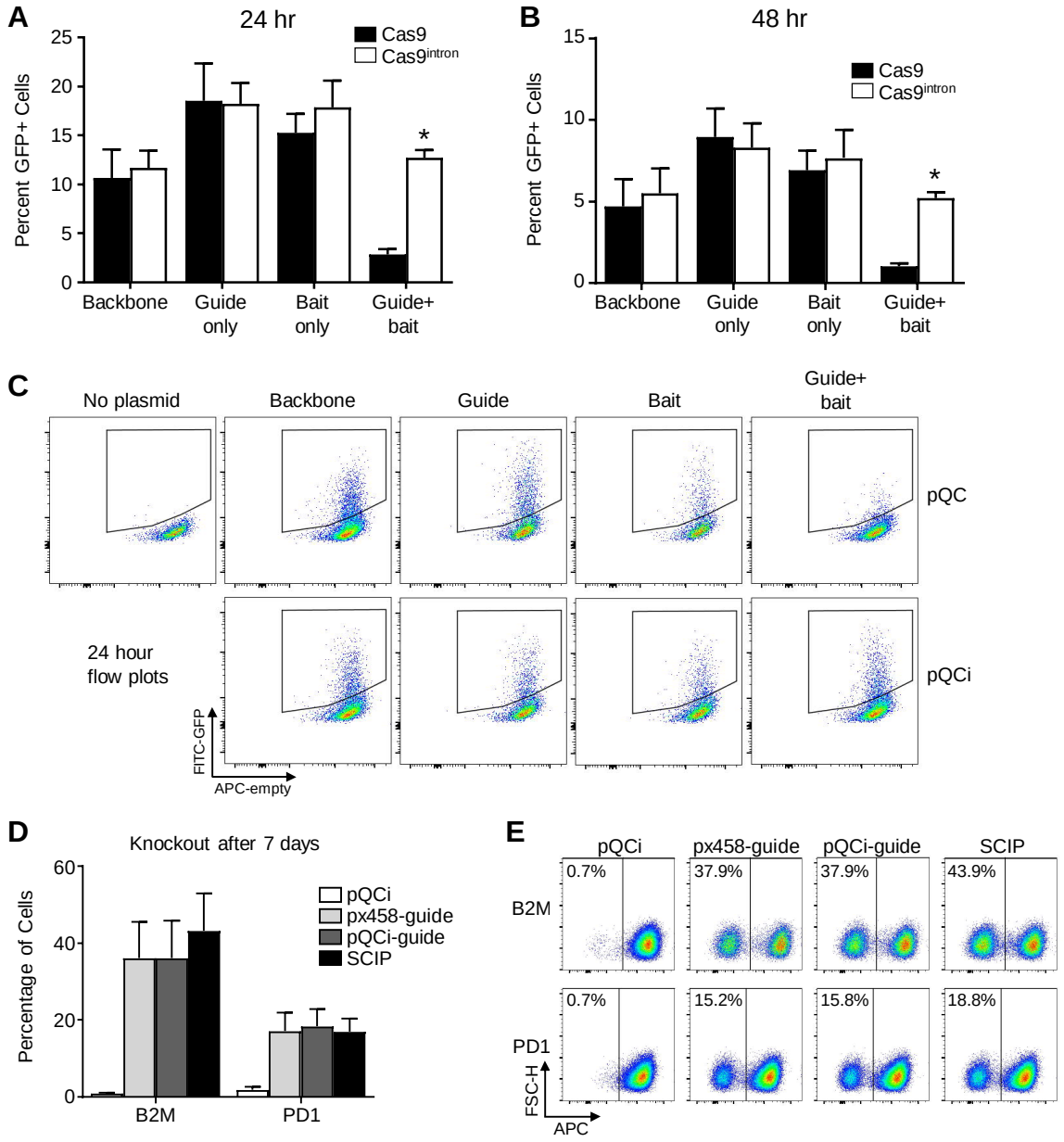

A
24 hr
Cas9
Cas9intron
Percent GFP+ Cells
25
20
15
10
5
0
*
Backbone
Guide only
Bait only
Guide+ bait
B
48 hr
Cas9
Cas9intron
Percent GFP+ Cells
15
10
5
0
*
Backbone
Guide only
Bait only
Guide+ bait
C
No plasmid
Backbone
Guide
Bait
Guide+ bait
pQC
24 hour flow plots
pQCi
FITC-GFP
APC-empty
D
Knockout after 7 days
Percentage of Cells
60
40
20
0
pQCi
px458-guide
pQCi-guide
SCIP
B2M
PD1
E
pQCi
px458-guide
pQCi-guide
SCIP
B2M
0.7%
37.9%
37.9%
43.9%
PD1
0.7%
15.2%
15.8%
18.8%
FSC-H
APC

Figure 3

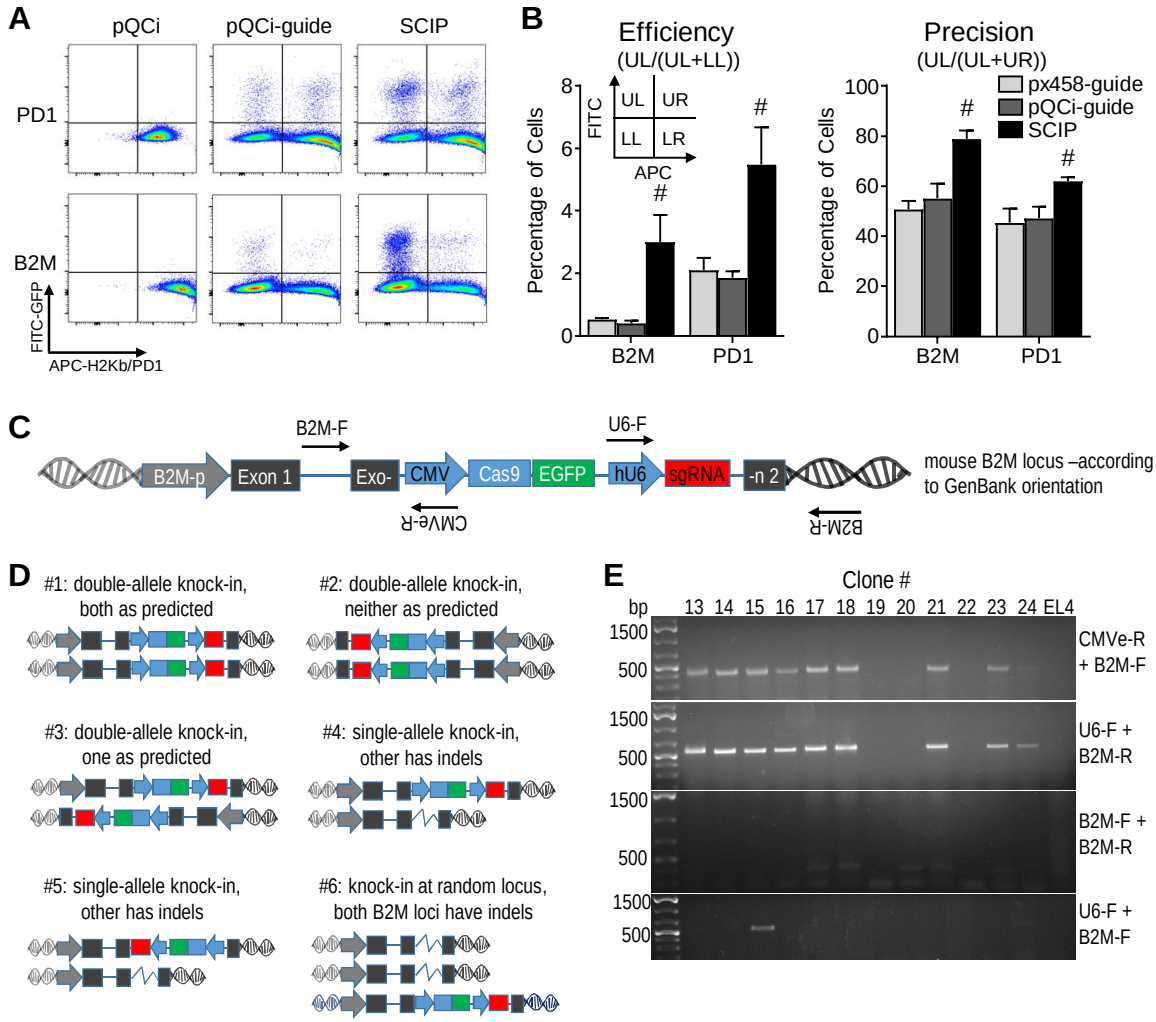

A
pQCi
pQCi-guide
SCIP
PD1
B2M
FITC-GFP
APC-H2Kb/PD1
B
Efficiency
(UL/(UL+LL))
Percentage of Cells
FITC
APC
UL
UR
LL
LR
B2M
PD1
Precision
(UL/(UL+UR))
px458-guide
pQCi-guide
SCIP
C
B2M-F
U6-F
B2M-p
Exon 1
Exo-
CMV
Cas9
EGFP
hU6
sgRNA
-n 2
CMVe-R
B2M-R
mouse B2M locus –according to GenBank orientation
D
#1: double-allele knock-in, both as predicted
#2: double-allele knock-in, neither as predicted
#3: double-allele knock-in, one as predicted
#4: single-allele knock-in, other has indels
#5: single-allele knock-in, other has indels
#6: knock-in at random locus, both B2M loci have indels
E
Clone #
bp
13 14 15 16 17 18 19 20 21 22 23 24 EL4
1500
500
CMVe-R + B2M-F
U6-F + B2M-R
B2M-F + B2M-R
U6-F + B2M-F
F
Mouse B2M gene exon 1
Genomic-scar (6nt)
Bait-scar (6nt)
Predicted>cccctcaaattcaagtatactcacgccacccACC-G---G-TGGGgctagcggtacccgttacataacttacggtaaatggcccgcctggctgaccgccc>3284
Clone 2<CCCCTCAAATTCAAGTATACTCACCCCACCCCCCGG---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<164
Clone 3<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<163
Clone 7<CCCATAAAACTCAAGTATAGGCACGCCACCCACC-G---GGTGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<163
Clone 10<CCCCTAAAATTAAAGTATACTCACCCCCCCCCCC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<161
Clone 11<CCCGTCAAATTCAAGTATACTCACGCCACCCACC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<161
Clone 13<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-G---GGTGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<162
Clone 14<CCCTTAAAATTCAAGTATCTTCACCCCCCCCCCC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<161
Clone 15<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<163
Clone 16<CCCCTAAAATTCAAGTACCCCCACGCCACCCGCC-G-----------AGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<161
Clone 17<CCCCTTCAATTCAAGTATACTCACGCCACCCCCC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<163
Clone 18<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-G---GGTGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACTTCCC<166
Clone 21<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-GACGG-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<162
Clone 23<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-GCG-G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<160
Clone 24<CCCCTCAAATTCAAGTATACTCACGCCACCCACC-G---G-TGGGGCTAGCGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC<164
5' insertion site
indels 6/14 clones
Whole plasmid sequence
Bait-scar (17nt)
Genomic-scar (17nt)
Mouse B2M gene exon 1
Predicted>atggctctagaacgcgtAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAacatactgaactgctacgtaacacagttccacccgcctcaca>12697
Clone 2>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>406
Clone 3>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>405
Clone 7>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATT----GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTACCACAGTTCCACCCGCCTCACA>404
Clone 10>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>406
Clone 11>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>414
Clone 13>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCAA------GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>401
Clone 14>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-----------AAGCCAAGCATACTGAACTGCTCCGTACCACAGTTCCACCCGCCTCACA>397
Clone 15>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>408
Clone 16>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>405
Clone 17>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>408
Clone 18>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTGGGGGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>411
Clone 21>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>406
Clone 23>ATGGCTCTAGAACGCGTAGGCCTTCGGCTTCCCATTCTCC-GGAGAATGGGAAGCCGAACATACTGAACTGCTACGTAACACAGTTCCACCCGCCTCACA>408
3' insertion site
Indels 4/13 colonies

Figure 4

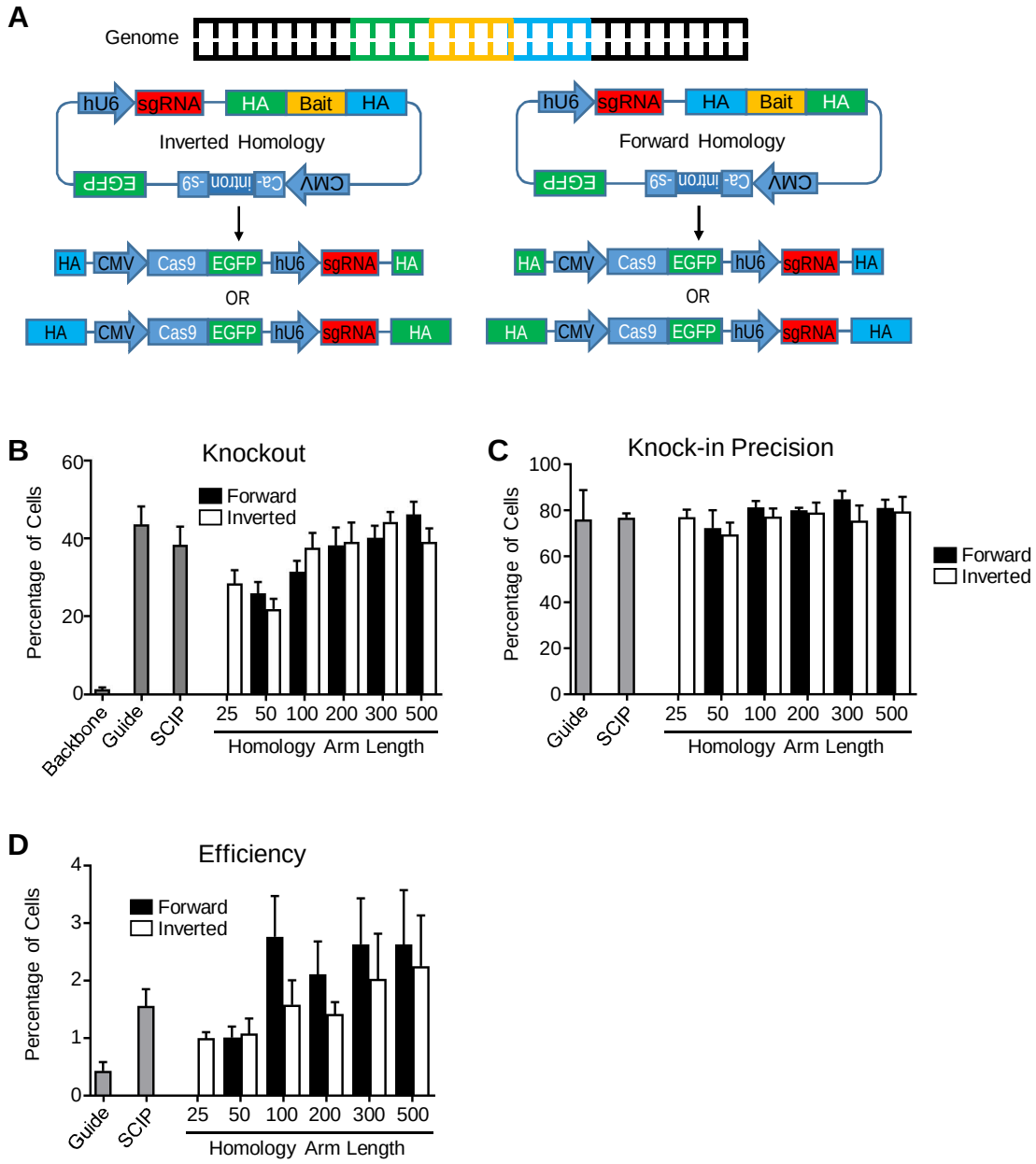

A
Genome
hU6
sgRNA
HA
Bait
HA
Inverted Homology
Forward Homology
CMV
Ca-intron
-S9
EGFP
HA
CMV
Cas9
EGFP
hU6
sgRNA
HA
OR
B
Knockout
Forward
Inverted
Percentage of Cells
Backbone
Guide
SCIP
25 50 100 200 300 500
Homology Arm Length
C
Knock-in Precision
Percentage of Cells
Guide
SCIP
25 50 100 200 300 500
Homology Arm Length
D
Efficiency
Forward
Inverted
Percentage of Cells
Guide
SCIP
25 50 100 200 300 500
Homology Arm Length

Figure 5

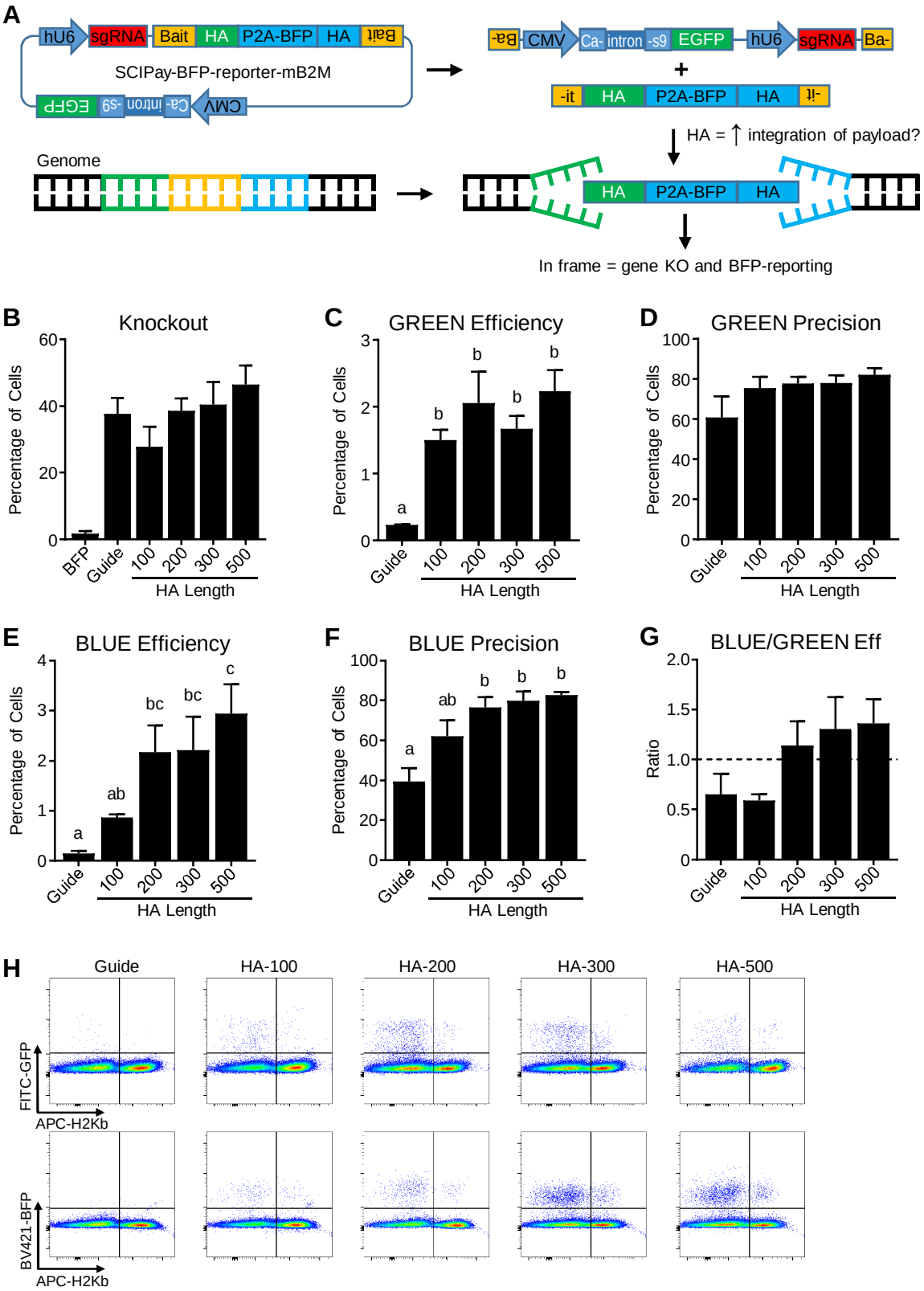

A
hU6
sgRNA
Bait
HA
P2A-BFP
HA
Bait
SCIPay-BFP-reporter-mB2M
CMV
Ca-
intron
-s9
EGFP
-Ba-
CMV
Ca-
intron
-s9
EGFP
hU6
sgRNA
-Ba-
+
-it
HA
P2A-BFP
HA
-it
HA = ↑ integration of payload?
Genome
HA
P2A-BFP
HA
In frame = gene KO and BFP-reporting
B
Knockout
Percentage of Cells
BFP
Guide
100
200
300
500
HA Length
C
GREEN Efficiency
Percentage of Cells
a
b
b
b
b
Guide
100
200
300
500
HA Length
D
GREEN Precision
Percentage of Cells
Guide
100
200
300
500
HA Length
E
BLUE Efficiency
Percentage of Cells
a
ab
bc
bc
c
Guide
100
200
300
500
HA Length
F
BLUE Precision
Percentage of Cells
a
ab
b
b
b
Guide
100
200
300
500
HA Length
G
BLUE/GREEN Eff
Ratio
Guide
100
200
300
500
HA Length
H
Guide
HA-100
HA-200
HA-300
HA-500
FITC-GFP
APC-H2Kb
BV421-BFP
APC-H2Kb

Figure 6

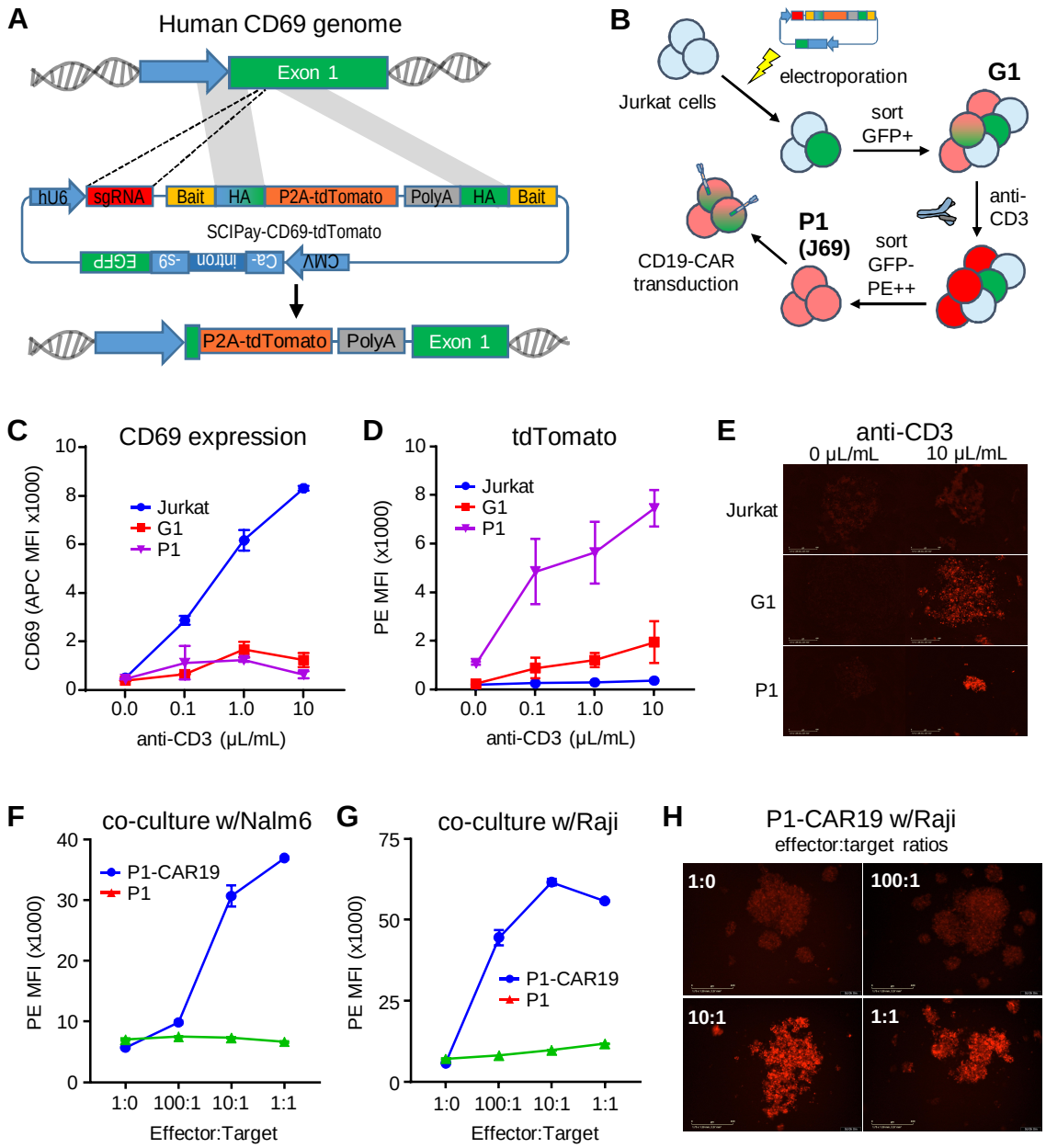

A
Human CD69 genome
Exon 1
hU6
sgRNA
Bait
HA
P2A-tdTomato
PolyA
HA
Bait
SCIPay-CD69-tdTomato
EGFP
-s9
intron
Ca-
CMV
P2A-tdTomato
PolyA
Exon 1
B
Jurkat cells
electroporation
G1
sort
GFP+
anti-
CD3
P1
(J69)
sort
GFP-
PE++
CD19-CAR
transduction
C
CD69 expression
CD69 (APC MFI x1000)
Jurkat
G1
P1
anti-CD3 (μL/mL)
D
tdTomato
PE MFI (x1000)
Jurkat
G1
P1
anti-CD3 (μL/mL)
E
anti-CD3
0 μL/mL
10 μL/mL
Jurkat
G1
P1
F
co-culture w/Nalm6
PE MFI (x1000)
P1-CAR19
P1
Effector:Target
G
co-culture w/Raji
PE MFI (x1000)
P1-CAR19
P1
Effector:Target
H
P1-CAR19 w/Raji
effector:target ratios
1:0
100:1
10:1
1:1

Figure 7

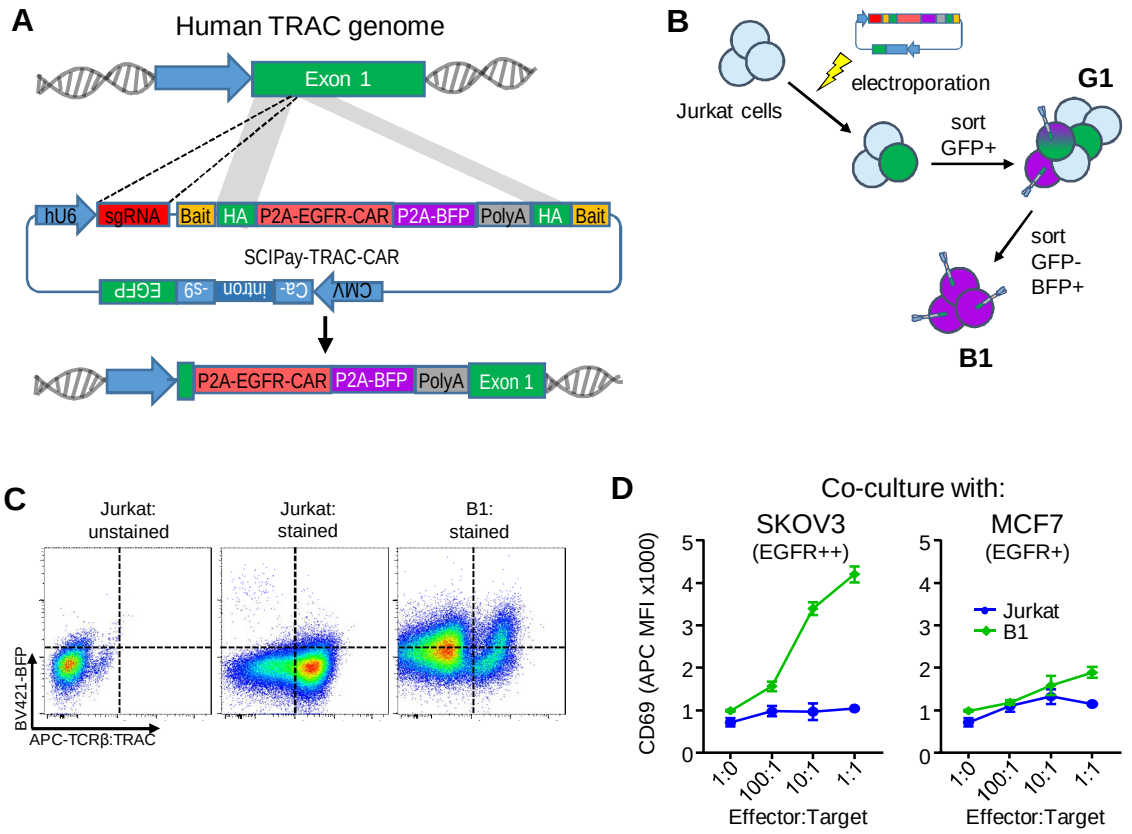
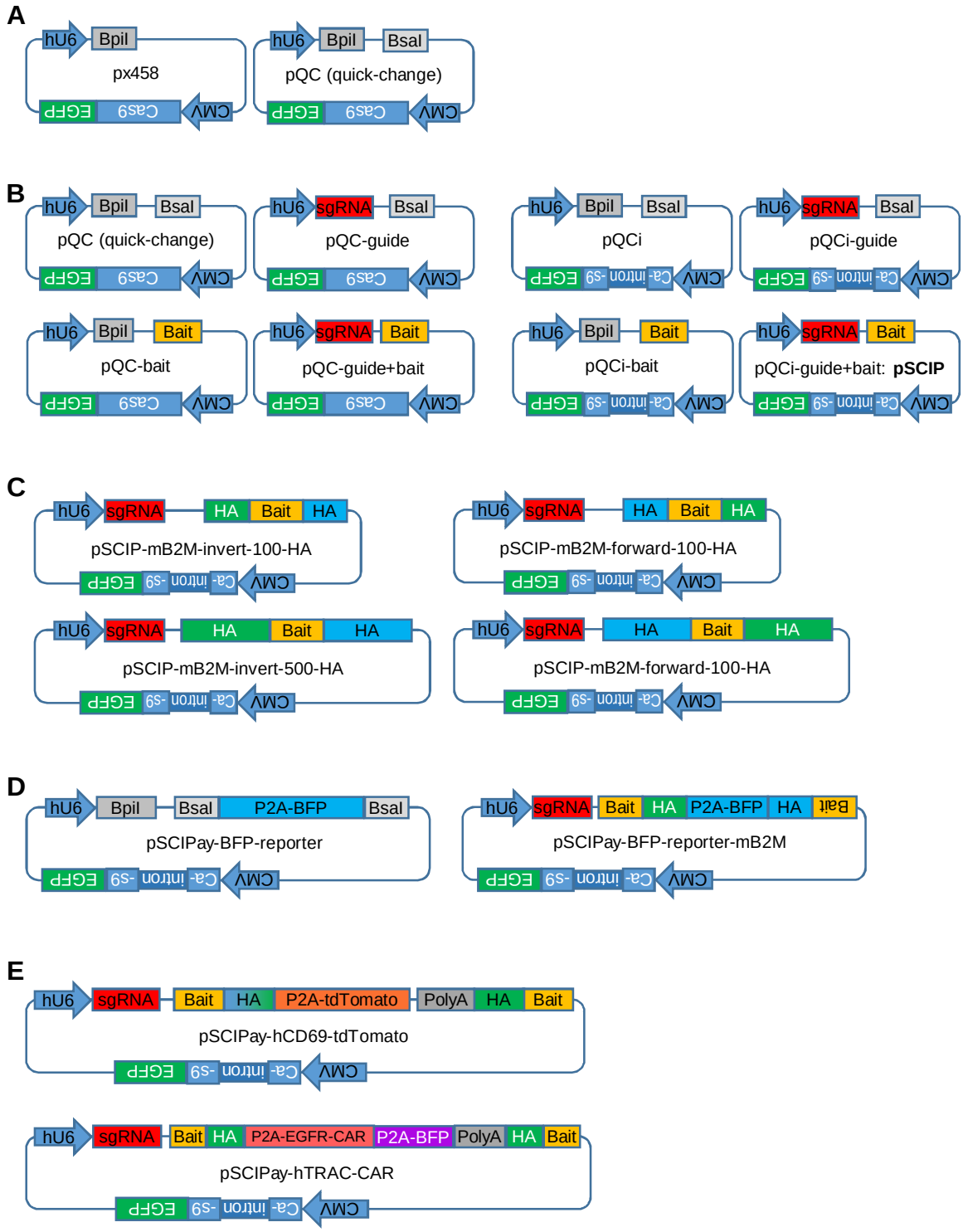
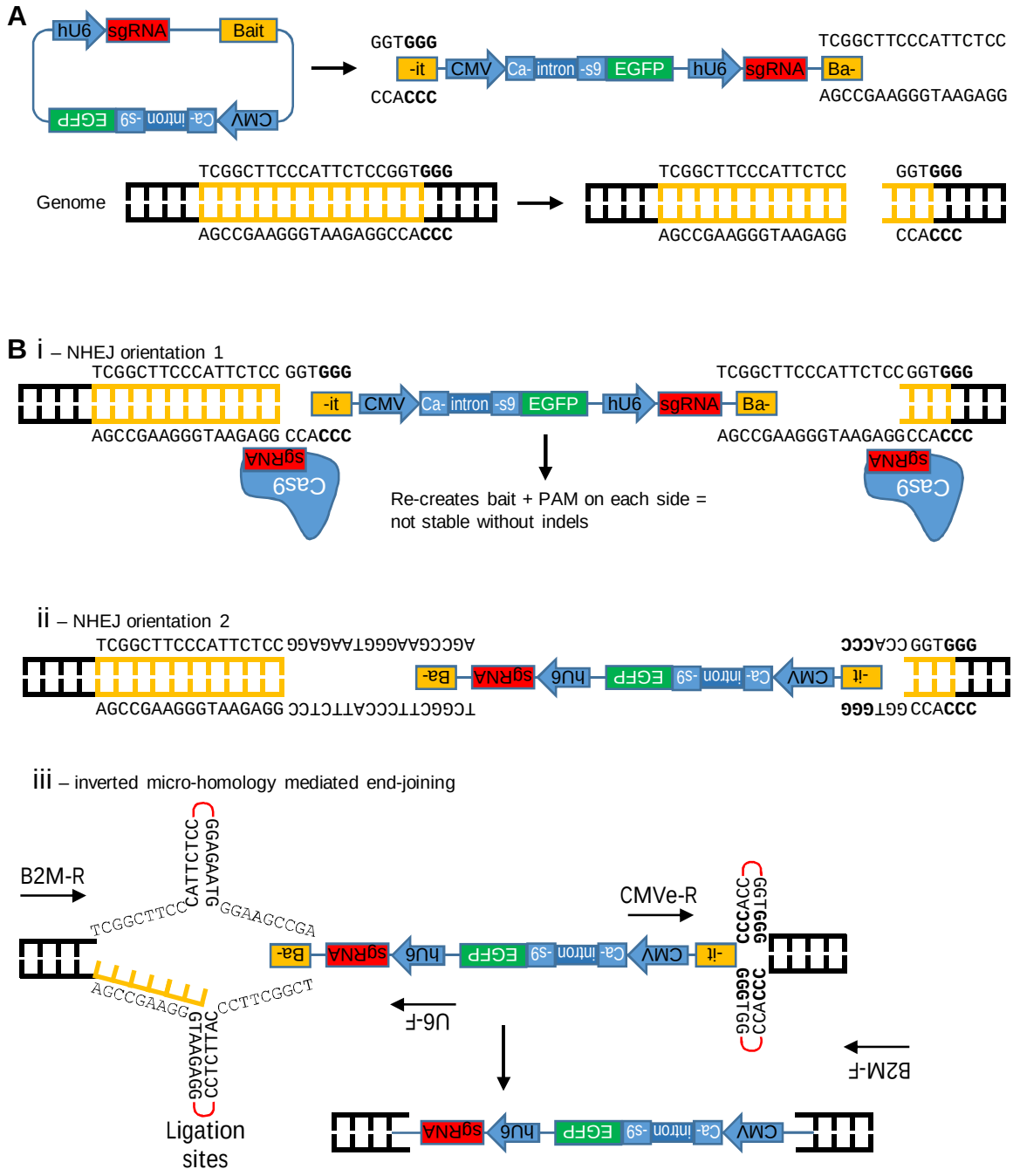

A
Human TRAC genome
Exon 1
hU6
sgRNA
Bait
HA
P2A-EGFR-CAR
P2A-BFP
PolyA
HA
Bait
SCIPay-TRAC-CAR
EGFP
-S9
intron
Ca-
CMV
P2A-EGFR-CAR
P2A-BFP
PolyA
Exon 1
B
Jurkat cells
electroporation
G1
sort
GFP+
sort
GFP-
BFP+
B1
C
Jurkat:
unstained
Jurkat:
stained
B1:
stained
BV421-BFP
APC-TCRβ:TRAC
D
Co-culture with:
SKOV3
(EGFR++)
MCF7
(EGFR+)
CD69 (APC MFI x1000)
Jurkat
B1
1:0
100:1
10:1
1:1
Effector:Target

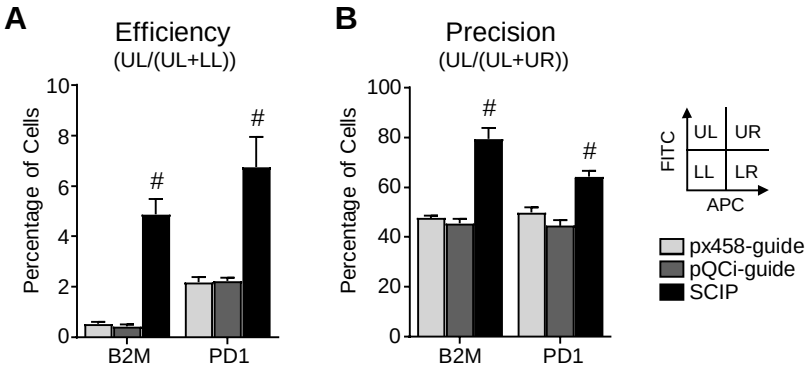
Figure 8



Supplemental Figure 1

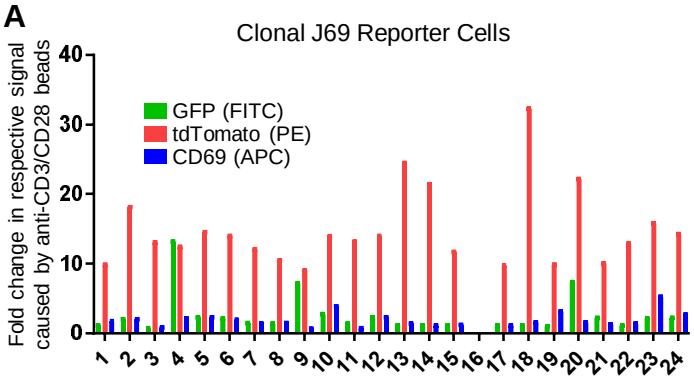


Supplemental Figure 2

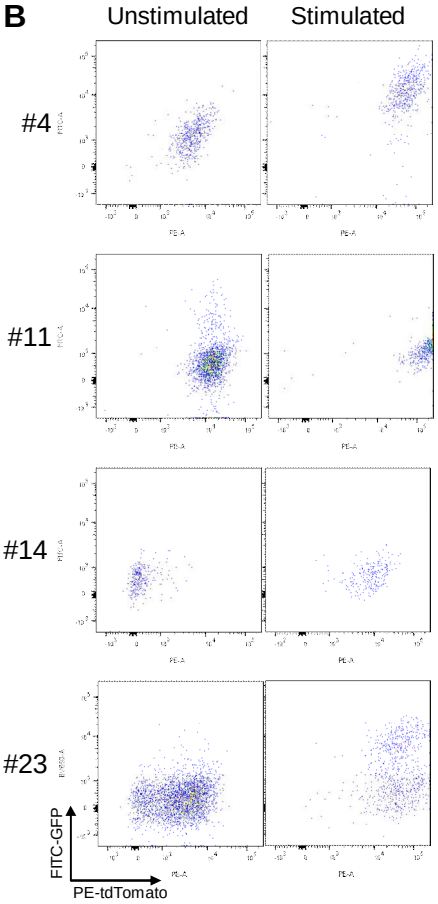


Supplemental Figure 3

A



B



C

Clone	Raw signal: unstimulated Green	Red	CD69	Raw signal: stimulated Green	Red	CD69	Fold change: unstim/stim Green	Red	CD69
1	682	3359	285	971	33894	493	1.4	10.1	1.7
2	579	2671	283	1282	48860	787	2.2	18.3	2.8
3	851	3351	482	892	44797	445	1.0	13.4	0.9
4	1511	5521	455	18831	67700	1202	12.5	12.3	2.6
5	575	9837	231	1486	143611	843	2.6	14.6	3.6
6	736	12569	300	1681	167641	888	2.3	13.3	3.0
7	845	13274	437	1641	165525	916	1.9	12.5	2.1
8	565	4268	293	921	45051	461	1.6	10.6	1.6
9	3137	8641	451	23532	80470	489	7.5	9.3	1.1
10	610	12762	230	1869	175776	1073	3.1	13.8	4.7
11	965	15706	676	1711	202766	806	1.8	12.9	1.2
12	637	13116	279	1677	184817	711	2.6	14.1	2.5
13	2412	2116	659	3636	54440	2363	1.5	25.7	3.6
14	719	2583	327	1030	57562	545	1.4	22.3	1.7
15	527	2116	239	735	24970	400	1.4	11.8	1.7
16	168	406	97.7	43.2	4689	350	0.3	11.5	3.6
17	590	4122	224	821	41468	377	1.4	10.1	1.7
18	572	314	228	729	10136	532	1.3	32.3	2.3
19	700	2414	312	950	25316	1771	1.4	10.5	5.7
20	944	2090	362	6491	45202	991	6.9	21.6	2.7
21	872	2644	343	2070	26756	1061	2.4	10.1	3.1
22	923	3644	420	1290	47624	528	1.4	13.1	1.3
23	722	3643	361	1639	58395	3916	2.3	16.0	10.8
24	1262	3545	737	1987	51407	1965	1.6	14.5	2.7

Supplemental Figure 4