Journal of Visualized Experiments

CRISPR/Cas9-based genome engineering to generate Jurkat reporter models for HIV-1 infection with selected proviral integration sites

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE58572R2
Full Title:	CRISPR/Cas9-based genome engineering to generate Jurkat reporter models for HIV-1 infection with selected proviral integration sites
Keywords:	HIV; HIV latency; HIV integration sites; cell culture model; genome engineering; CRISPR/Cas9
Corresponding Author:	Ulrike Lange Heinrich-Pette-Institut Leibniz-Institut fur Experimentelle Virologie Hamburg, Hamburg GERMANY
Corresponding Author's Institution:	Heinrich-Pette-Institut Leibniz-Institut fur Experimentelle Virologie
Corresponding Author E-Mail:	ulrike.lange@leibniz-hpi.de
Order of Authors:	Julia K Bialek
	Thomas Walther
	Joachim Hauber
	Ulrike C Lange
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Heinrich-Pette-Institute, Leibniz Institute for Experimental Virology, Department of Antiviral Strategies, Martinistr 52, 20251 Hamburg, Germany

TITLE:

CRISPR-Cas9-Based Genome Engineering to Generate Jurkat Reporter Models for HIV-1 Infection with Selected Proviral Integration Sites

AUTHORS AND AFFILIATIONS:

Julia K Bialek[1,2,*], Thomas Walther[1,*], Joachim Hauber[1,3], Ulrike C Lange[1,2,3]

[1]Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Hamburg, Germany
[2]Department of Anesthesiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[3]German Center for Infection Research (DZIF), Hamburg Partner Site, Germany
*Authors contributed equally

Corresponding Author:

Ulrike C Lange
ulrike.lange@leibniz-hpi.de
Tel: +49-40-48051-347

Email Addresses of Co-authors:

Julia K Bialek (bialek.julia@mh-hannover.de)
Thomas Walther (t.walther@uni-luebeck.de)
Joachim Hauber (joachim.hauber@leibniz-hpi.de)

KEYWORDS:

HIV, HIV latency, HIV integration sites, cell culture model, genome engineering, CRISPR-Cas9

SUMMARY:

We present a genome engineering workflow for the generation of new *in vitro* models for HIV-1 infection that recapitulate proviral integration at selected genomic sites. Targeting of HIV-derived reporters is facilitated by CRISPR-Cas9-mediated, site-specific genome manipulation. Detailed protocols for single-cell clone generation, screening, and correct targeting verification are provided.

ABSTRACT:

Human immunodeficiency virus (HIV) integrates its proviral DNA non-randomly into the host cell genome at recurrent sites and genomic hotspots. Here we present a detailed protocol for the generation of novel *in vitro* models for HIV infection with chosen genomic integration sites using CRISPR-Cas9-based genome engineering technology. With this method, a reporter sequence of choice can be integrated into a targeted, chosen genomic locus, reflecting clinically relevant integration sites.

In the protocol, the design of an HIV-derived reporter and choosing of a target site and gRNA sequence are described. A targeting vector with homology arms is constructed and transfected into Jurkat T cells. The reporter sequence is targeted to the selected genomic site by homologous

recombination facilitated by a Cas9-mediated double-strand break at the target site. Single-cell clones are generated and screened for targeting events by flow cytometry and PCR. Selected clones are then expanded, and correct targeting is verified by PCR, sequencing, and Southern blotting. Potential off-target events of CRISPR-Cas9-mediated genome engineering are analyzed.

By using this protocol, novel cell culture systems that model HIV infection at clinically relevant integration sites can be generated. Although the generation of single-cell clones and verification of correct reporter sequence integration is time-consuming, the resulting clonal lines are powerful tools to functionally analyze proviral integration site choice.

INTRODUCTION:

Integration of proviral DNA into the host genome upon infection is a critical step in the life cycle of human immunodeficiency virus (HIV). Following integration, HIV persists by establishing latency in long-lived CD4+ T cell subsets such as memory CD4+ T cells. HIV integration appears to be non-random[1,2]. A number of genomic hotspots with recurrently integrated proviral DNA has been detected in several studies through the sequencing of integration sites in acutely and chronically infected individuals[2-8]. Interestingly, at some of these integration sites, the same locus was detected in a large fraction of infected cells, leading to the idea that integration at recurrent sites might positively affect clonal expansion[1].

To advance our understanding of the significance of recurrent integration sites, proviral integration site choice must be explored. However, several technical aspects hamper studying HIV integration site choice and the consequences. Broadly used cell culture models for HIV latency like JLat cell lines do not reflect clinically relevant recurrent integration sites[9]. Studies on primary patient-derived cells, on the one hand, enable description of integration site landscape by sequencing but do not allow for functional analyses. To our knowledge, no adequate experimental model is available to functionally analyze selected clinically relevant integration sites.

Here we present a detailed workflow to generate novel models for HIV infection using CRISPR-Cas9-based genome engineering technology. The workflow described herein can be used to generate T cell-derived reporter cell lines that model HIV infection, carrying a genomically integrated proviral reporter at a chosen integration site. They are thus serving as new tools to explore how the proviral integration site can impact HIV biology and how the provirus responds to different treatment strategies (*e.g.,* inducibility by latency reversing agents). Our method uses the advantages of CRISPR-Cas9-based genome engineering, in which integration of the reporter sequence by homologous recombination is facilitated by a Cas9 nuclease-induced double-strand break at the target site. Target sites for integration are chosen according to proximity to the described recurrent integration sites from studies on HIV-infected individuals and the presence of suitable PAM motifs for Cas9-mediated genome engineering.

In our exemplary results, we have focused on the BACH2 gene locus, which codes for the BTB And CNC Homology transcriptional regulator 2. In chronically HIV-infected individuals on antiretroviral therapy, BACH2 is one of the loci showing enrichment of integrated HIV-1

sequences[3,6-8,10]. We have chosen a minimal HIV-derived reporter consisting of HIV-1-derived long terminal repeat (LTR), tdTomato coding sequence, and bovine growth hormone (BGH) polyadenylation signal (PA), which we have targeted to two specific sites in BACH2 intron 5. The presented protocol is optimized for Jurkat cells, a human CD4+ T cell-derived suspension cell line, but other cell lines may be used and the protocol adapted accordingly. We present a detailed workflow for selection of target site, construction of target vector with homology arms, CRISPR-Cas9-mediated targeting of the reporter into the chosen genomic site, generation and selection of clonal lines, and comprehensive verification of newly generated, targeted reporter cell lines.

PROTOCOL:

1. Targeting Strategy for Genome Engineering and Targeting Vector (tv) Design

Note: The first step of genome engineering involves selection and generation of the necessary tools for CRISPR-Cas9-mediated targeting. Selection of a genomic integration site locus, choice of cell type for targeting, and design of an HIV-derived reporter for integration should precede this step. This protocol describes targeting of an HIV-LTR_tdTomato_BGH-PA minimal reporter into Jurkat target cells. A flow chart of the workflow for CRISPR-Cas9-based targeting, generation, screening and verification of clonal lines is depicted in **Figure 1**. The described targeting strategy uses the *S. pyogenes* Cas9 (SpCas9) to generate gRNA-directed dsDNA breaks at a selected integration site. The reporter is then targeted into the chosen genomic locus through homologous recombination by providing a non-linearized targeting vector (tv) that contains the reporter sequence flanked by so-called 5' and 3' homology arms (HA)[11].

1.1. Choice of targeted locus, gRNA, and targeting vector design

1.1.1. Choose thegenomic locus to be targeted based on the individual scientific question. Use published lists of recurrent integration sites of HIV found in patients in different studies[2-8] as a guideline. *In silico* extract the genomic sequence of the desired genomic locus to be targeted (sequence of the complete gene or at least 5 kb of genomic sequence) using UCSC genome browser (http:// genome.edu.ucsc.edu).

1.1.2. Choose guide RNAs (gRNAs) of 20 nt for targeting of the chosen genomic locus using the E-CRISP webtool (http://www.e-crisp.org).

1.1.2.1. Select "Homo sapiens GRCh38" as the organism. Input 2000 bp of the genomic sequence covering the desired genomic locus extracted in step 1.1.1.

1.1.2.2. Start a gRNA search using medium application settings (any PAM, any 5' base, off-targets tolerate mismatches, and introns/CPG islands are excluded). A list with possible gRNA designs will appear, ranking from highest to lower scores for specificity and efficiency.

1.1.2.3. Select a gRNA that preferably shows a high score for specificity and efficiency and is as close as possible to the desired genomic locus to be targeted.

Note: A compromise between proximity to the desired genomic locus and design of specific and efficient gRNA has to be found.

1.1.3. Blast the chosen gRNA sequence against the reference genome using the NCBI blast browser (https://blast.ncbi.nlm.nih.gov) to check for uniqueness of the gRNA binding site.

1.1.3.1. Select "human" as the genome. Input the gRNA sequence as the query sequence. Select "highly similar sequences" (megablast) as the program. Ensure that the gRNA sequence is unique. If not, chose a different gRNA from step 1.1.2.3 and blast again.

1.1.4. Once gRNA sequence is chosen, select *in silico* 1000 bp upstream and downstream of gRNA sequence from genomic sequence extracted in step 1.1.1 as 5' and 3' HA accordingly.

Note: The gRNAs should be homologous to the chosen genomic integration site locus and located adjacent to a protospacer adjacent motif (PAM; *e.g.,* NGG for SpCas9) (**Figure 2a**). The tv contains the reporter sequence that is 5' and 3' flanked by HAs. HAs cover 1000 bp upstream and downstream of the gRNA sequence[11]. The full gRNA sequence should not be included in the HA. An overlap of up to 5 nt is acceptable (**Figure 2a**).

1.2. Generation of gRNA and targeting vectors

Note: For vector schemes, see **Figure 2b**.

1.2.1. To generate a vector for expression of SpCas9 and gRNA, use the pX330-U6-Chimeric_BB-cBh-hSpCas9 as the backbone from which both SpCas9 and the single guide RNA (sgRNA) can be simultaneously expressed. To clone the gRNA sequence into the backbone, use the BbsI restriction sites[12].

1.2.2. To generate the tv, choose a high-copy plasmid as backbone (*e.g.,* pMK or cDNA3.1).

1.2.2.1. First, assemble the reporter (in this protocol: LTR_tdTomato_BGH-PA) into the construct backbone by Gibson Assembly clong[13] using a commercial assembly cloning kit, and introduce 5' and 3' flanking restriction sites (*e.g.,* 5' PacI and 3' SmaI) for subsequent restriction digestion cloning of the HAs.

1.2.2.2. Amplify 1000 bp of the HA fragments chosen in step 1.1.4 from genomic DNA (gDNA) of the cell type to be targeted (in this protocol: Jurkat cells) using a DNA polymerase with proofreading activity (see **Tables 1** and **2** for PCR ingredients and cycling conditions). Then, introduce reporter flanking restriction sites on the 5' and 3' ends of each HA (PacI on 5' HA on both ends, SmaI on 3' HA on both ends).

1.2.2.3. Sequentially clone HAs into construct backbone already containing the reporter (generated in step 1.2.2.1) by restriction enzyme cloning[14,15]. First, clone in 5' HA using PacI restriction sites, then clone in 3' HA using SmaI restriction sites.

Note: If tv backbone contains an additional fluorescent reporter, unwanted backbone integration can be assessed by flow cytometry (see steps 3.2.2 and 3.2.3). If tv backbone contains no fluorescent reporter, backbone integration must be assessed using PCR (see step 3.2.8).

[Place **Figure 1** here] [Place **Figure 2** here]

2. CRISPR-Cas9-Based Targeting of Jurkat Cells

2.1. Transfection of Jurkat cells

2.1.1. 24 h prior to transfection, plate 1.25 x 10^6 Jurkat T cells in 2.5 mL of RPMI 1640 supplemented with 10% (v/v) fetal calf serum (FCS) and 4 mM L-glutamine [referred to as "RPMI w.o. antibiotics (AB)"] per well of a 6-well cell culture plate. For a single targeting experiment, prepare one complete 6-well plate (*i.e.,* 6 wells each with 2.5 mL of cell suspension).

2.1.2. On the following day, co-transfect the cells with circular tv and pX330-U6-Chimeric_BB-cBh-hSpCas9/gRNA using a transfection reagent specific for Jurkat cells.

2.1.2.1. Add 2 µg of circular tv and 2 µg of pX330-U6-Chimeric_BB-cBh-hSpCas9/gRNA per well to 250 µL of commercial RPMI medium with reduced serum concentration optimized for transfection (RPMI with 50% reduction in serum) in a reaction tube and mix well.

2.1.2.2. Add 12 µL of transfection reagent slowly to the DNA/medium without touching the wall of the tube and swirl. Let the mixture incubate for 15 min and add dropwise to one well of cells. Incubate the cells at 37 °C and 5% CO_2.

Note: The preparation of transfection reaction can be scaled up. No medium change is required after transfection.

2.2. Enrichment of transfected cells by fluorescence-activated cell sorting (FACS)

2.2.1. 72 h post-transfection, pool the transfected cells, count them, and prepare for enrichment by FACS. Collect the cells in a 50 mL conical tube, centrifuge at 300 x g and room temperature (RT) for 4 min, wash the cells once in PBS, centrifuge again, suspend the pellet in an appropriate amount of FACS buffer (PBS supplemented with 1% FCS + 1 mM EDTA) at 1 x 10^7 cells/mL, and finally transfer into a FACS tube.

2.2.2. Subject the cells to FACS and sort those that express the fluorescent reporter of the tv (*e.g.,* tdTomato in this protocol). Collect the cells in RPMI 1640 supplemented with 10% FCS, 4 mM L-glutamine, and 50 U/mL penicillin and streptomycin (referred to as "RPMI w/ AB").

2.2.3. After FACS sorting, wash the cells once by adding 20 mL of RPMI w/ AB to sorted cells and centrifuge at 300 x g for 4 min at RT. Resuspend the cell pellet in an appropriate amount of RPMI w/ AB and plate the cells in one well of a cell culture plate with the appropriate volume according to cell number post-FACS.

Note: Culture sort the cells in a small volume (*e.g.,* 24-well), as considerable levels of cell death have been observed in the first week post-targeting (up to 80-90%).

2.2.4. Expand the mixed targeted cell population up to a density of 1 x 10^6 cells/mL in a 75 cm^2 cell culture flask. This will take around 10-14 days post-FACS sorting.

2.3. Confirmation of targeting events by flow cytometry in the mixed targeted cell population

2.3.1. After 10-14 days of expansion (when cells have reached a density of 1 x 10^6 cells/mL in a 75 cm^2 cell culture flask), plate two wells with 1 x 10^6 cells of the mixed targeted cell population in 1 mL of RPMI w/ AB in a 12-well cell culture plate.

2.3.2. Induce the HIV LTR of the reporter (generation of tv is described in step 1.2.2.1.) in one of the wells by adding 50 ng/mL phorbol 12-myristate 13-acetate (PMA) and 1 µM Ionomycin (referred to as PMA-Iono). Use cells in the second well as the non-induced control. Culture both the induced and the non-induced cells for 24 h.

2.3.3. Take 0.5 mL of cell suspension of non-induced and induced cells (each), wash them once in PBS, and suspend each in 200 µL of FACS buffer.

2.3.4. Analyze 100,000 cells by flow cytometry. Gate the viable single-cells based on size in forward and sideward scatter, and analyze fluorescent reporter gene expression.
Note: At this step (10-14 days post-FACS sorting), transient expression of fluorescent reporter by transfection should no longer be detectable. Fluorescent expression at this time point indicates genomic integration of reporter sequence.

2.4. **Detection of targeting events by PCR on genomic DNA of the mixed targeted cell population**
Note: To detect targeting events via PCR, design two primer pairs specific for the 5' integration (int.) junction and 3' int. junction. For the 5' int. junction PCR, the forward primer should bind upstream of the 5' HA and the reverse primer in the LTR of the reporter (primers P1 and P2 in **Figure 3a**). The primer pair for the 3' int. junction PCR should span from the PA of the reporter to 100-200 bp downstream of the 3' HA (primers P3 and P4 in **Figure 3a**). Primers P1 and P4 will also serve for amplification of the non-targeted allele in the mixed targeted population. For a schematic, see **Figure 3a**.

2.4.1. Prepare gDNA from 2 mL of cell suspension of the mixed targeted cell population from step 2.2.4. Use a gDNA extraction kit according to the manufacturer's protocol. Then prepare the gDNA of non-targeted cells as a control.

2.4.2. Perform int. junction PCRs (primers P1/P2 and P3/P4 for 5' and 3' int. junction, respectively) and non-targeted allele PCR (primers P1/P4) using a high-fidelity DNA polymerase (see **Tables 3** and **4** for PCR ingredients and cycling conditions). Analyze 5 µL of PCR products on a 1.5% agarose/TAE gel.
Note: If the mixed targeted population contains cells that have undergone genome engineering, a specific PCR product should be observed that is not detectable in gDNA of non-targeted cells (negative control). For the non-targeted allele PCR, one should observe a product of the same size for both the targeted and non-targeted cells (positive control for genomic P1 and P4 primers). If no bands are observed, consider altering PCR cycling conditions by increasing the number of cycles or altering the PCR buffer (for example through addition of DMSO or increased amounts of Mg^{2+}), or by changing the polymerase.

3. Generation of Clonal Lines and Screening for Correct Targeting

Note: After confirmation of the targeting events in the mixed targeted cell population by flow cytometry and PCR (sections 2.2-2.4), generate single-cell clones (duration: 28 to 35 days) and screen for correct integration of the reporter sequence.

3.1. Generation of single-cell clones through dilution plating

3.1.1. Prepare Jurkat-conditioned medium in advance: take off RPMI w/ AB medium from healthy, untreated Jurkat T cells grown to 1 x 10^6 cells/mL, centrifuge for 5 min at 300 x g, and filter the supernatant using a 0.22 µm syringe filter unit.

Note: Keep the conditioned medium at 4 °C for short-term storage or at -20 °C for storage longer than 1 week. Prepare 20 to 30 mL of conditioned medium prior to dilution plating.

3.1.2. Count the targeted cells from step 2.2.4 after 10-14 days of expansion and dilute them in RPMI w/ AB to a concentration of 1 x 10^5 cells/mL. Take 100 µL of 1 x 10^5 cells/mL solution and dilute with 9.9 mL of medium to achieve a concentration of 1000 cells/mL. Take 1 mL of 1000 cells/mL solution and dilute with 9 mL of medium to achieve a concentration of 100 cells/mL.

3.1.3. Plate 96-well plates containing 1 cell per well and 2 cells per well. For 1 cell per well, take 1 mL of 100 cells/mL solution and mix gently with 5 mL of conditioned medium and 4 mL of fresh medium in a sterile reagent reservoir.

3.1.4. For 2 cells per well, take 2 mL of 100 cells/mL solution and mix gently with 5 mL of conditioned medium and 3 mL of fresh medium. Plate 96-well round-bottom plates with 100 µL of the respective cell dilution per well using a multichannel pipet.

Note: 5 to 10 96-well plates per targeting construct are sufficient to obtain clones for screening.

3.1.5. Stack the 96-well plates, cover each stack with a 6-well plate containing 3 mL of PBS in each well, and incubate the plates at 37 °C in a humidified cell culture incubator with 5% CO_2 for 3 weeks.

Note: Do not change the cell culture medium during this time. Do not open the incubator more than once or twice a week. The best results are observed in incubators with open water reservoir.

3.1.6. After 3 weeks of incubation, visually confirm the presence of grown colonies using light microscopy (4X magnification) and mark the wells with grown colonies so they are visible as points on the bottom of the wells.

3.1.7. Prepare one 96-well round-bottom plate with 100 µL of RPMI w/ AB per well. Gently resuspend cells of a marked well by pipetting. Transfer 100 µL of cell suspension into one well of the new 96-well plate already containing 100 µl of RPMI w/ AB, then mix gently by pipetting. Transfer 100 µL of this cell suspension into a second empty 96-well round-bottom plate to duplicate the plate.

3.1.8. Continue with all marked wells with grown colonies. Fill all the blank wells with 200 µL of RPMI w/ AB medium. Incubate the plates at 37° C and 5% CO_2.

Note: One of these plates will serve for expansion of single-cell clones ("stock plate") and the other as a "duplicate plate" for screening.

3.2. Screening of single-cell clones by flow cytometry and PCR

Note: While the single-cell clones are expanding, use the duplicate plate from step 3.1.8 to screen single-cell clones for presence of reporter sequence by PCR (steps 3.2.4-3.2.12) and expression of fluorescent reporter by flow cytometry (steps 3.2.2-3.2.3) (**Figure 3c**).

3.2.1. Let the duplicate plate incubate for 24 to 48 h and duplicate the plate again. To do this, add 100 µL of RPMI w/ AB to every well, mix gently by pipetting, and transfer 100 µL to a new 96-well round-bottom plate using a multichannel pipet. Use one plate for flow cytometry screening and the other for PCR-based screening.

3.2.2. For flow cytometry screening, stimulate the cells with PMA-Iono. Prepare a mastermix of 0.1 µL of Ionomocin (1 mM stock), 0.1 µL of PMA (50 µg/µL stock), and 4.8 µL of RPMI w/ AB per number of wells, then add 5 µL of mastermix per well.

Note: Induction is necessary to successfully identify clones, where the LTR might be transcriptionally silent and therefore the fluorescent reporter is not expressed.

3.2.3 Let cells incubate for 24 h and prepare cells for flow cytometry as described in step 2.3.3. Gate any viable single-cells based on size in forward and sideward scatter and analyze fluorescent reporter gene expression by flow cytometry (for example results, see **Figure 3c**). If tv backbone contains a second fluorescent reporter with promoter (*e.g.,* GFP), screen any clones for backbone reporter expression also (see step 1.2.2 and the following note for explanation).

Note: Backbone reporter expression indicates unwanted integration of backbone sequences.

3.2.4. Once the clones in the second duplicate plate have grown sufficiently (usually 24 to 48 h after duplication of the 96-well plate), prepare cell lysates containing gDNA for PCR screening. Centrifuge the plate for 10 min at 300 x g at RT. Carefully take off the supernatant without disturbing the cell pellet.

Note: All steps for the preparation of lysates and PCR reactions can be performed with multichannel pipettes.

3.2.5. Wash cells with 100 µL of PBS by gentle pipetting and centrifuging the plate for 5 min at 300 x g at RT. Take off the PBS and add 200 µL of lysis buffer [200 mN NaCl, 100 mM Tris-HCl pH 8-8.5, 5 mM EDTA, 0.1% SDS; then add 250-1000 µg/mL of proteinase K (lyophilized powder, weigh in freshly)] per well. Mix gently by pipetting, and transfer the suspension to a new PCR plate.

3.2.6. Seal the plate with paraffin film and incubate for 1 h at 55 °C in a thermocycler. Centrifuge at maximum speed for 10 min (3000 x g) to spin down cell debris, and transfer the supernatant to a new PCR plate.

Note: Cell lysates in plates can be stored at this stage at 4 °C until further use.

3.2.8. Prepare a 96-well PCR plate with 110 µL of dH_2O and add 10 µL of cell lysate (1:12 dilution). Cell lysates might be viscous and difficult to pipet. Use at least 20-µL pipet tips.

3.2.9. Inactivate proteinase K by incubation for 10 min at 99 °C in a thermocycler. Subsequently use the inactivated and diluted cell lysates for PCR screening.

3.2.10. Design primers for screening PCR (P5 and P6) based on the chosen reporter sequence to amplify 500-800 bp of the reporter sequence. For positive control PCR, use primers P7 and P8 that amplify 630 bp of a wild-type, non-targeted genomic locus (*NUP188* gene) (**Figure 3c** and **Table 5**). Design a third primer pair that amplifies 500-600 bp of the tv backbone as a control for unspecific integration of tv backbone sequences (backbone PCR).

3.2.11. For screening, control, and backbone PCR, use a commercial PCR mastermix (see **Tables 6** and **7** for PCR ingredients and cycling conditions). Use 2 µL of the diluted and inactivated lysate prepared in step 3.2.8 as a template and run PCR for 38 to 40 cycles of PCR amplification in 96-well format.

3.2.12. Analyze 5 µL of PCR products on a 1.5% agarose/TAE gel.

Note: For control PCR, a specific band of 630 bp should be observed for every sample, confirming that the quality of cell lysates is adequate for PCR. A specific band in screening PCR (500-800 bp depending on primer design) indicates integration of the reporter sequence. A specific band for backbone PCR (500-600 bp, depending on primer design) indicates unwanted integration of tv backbone sequences (for example results, see **Figure 3c**).

3.2.13. Combine the results of flow cytometry (step 3.2.3) and PCR-based screening (step 3.2.12). Select clones which show correct sizes of PCR products in screening PCR and positive control PCR and expression of fluorescent reporter after induction with PMA-Iono in flow cytometry. Exclude clones that show any PCR product in backbone PCR or expression of tv backbone-encoded fluorescent protein, indicating unspecific integration of tv backbone sequence.

3.2.14. Gradually expand selected clones from the 96-well stock plate to bigger well formats (48/24/12/6-well) until achieving a T75 cell culture flask format by adding fresh medium every 2 to 3 days. Maintain a cell density between 1 x 10^5 and 1 x 10^6 cells/mL.

3.2.15 Make sure to prepare cell stocks of clones during expansion: count the cells, centrifuge at 300 x g for 5 min at RT, discard the supernatant, and suspend the cells gently in FCS + 10% DMSO at 5 x 10^6 cells/mL. Aliquot in cryogenic vials and use a cryo-freezing container to freeze the cells to 80 °C at 1 °C/min. For long term storage, transfer them to liquid N_2.

Note: It is advisable to retain a T75 cell culture flask (*i.e.,* around 1 x 10^7 cells) during the expansion in preparation of gDNA for verification of targeting by Southern blotting (see section 3.4).

3.3. Verification of integration sites by PCR/sequencing in selected clones

Note: 5' and 3' int. junctions of the selected clones are PCR amplified and submitted to Sanger sequencing to verify correct targeting at the DNA sequence level.

3.3.1. Prepare gDNA of the selected clones and Jurkat wild-type cells using a commercial gDNA extraction kit.

3.3.2. Use primer pairs binding the 5' end of the reporter and upstream of 5' HA for 5' int. junction (primers P1 and P2) and the 3' end of the reporter and downstream of 3' HA for 3' int. junction (primers P3 and P4) as described in step 2.4. Use primers P1 and P4 to amplify the targeted integration site on the allele without reporter integrant (**Figure 4a**).

3.3.3. Prepare PCR reactions with 100-200 ng of gDNA as a template and perform PCR using a polymerase with proofreading activity (see **Tables 1** and **2** for PCR ingredients and cycling conditions).

Note: If no bands are observed, consider altering the PCR cycling conditions by increasing the number of cycles or altering the PCR buffer (for example, by adding DMSO or increased amounts of Mg^{2+}), or by changing the polymerase.

3.3.4. Analyze 5 µL of PCR products on a 1.5% agarose/TAE gel. If correct band sizes are observed, purify the remaining PCR product using a commercial kit and subject it to Sanger sequencing. Verify sequences of 5' int. junction, 3' int. junction, and targeted site of the allele without reporter integrant by aligning them with expected sequences.

Note: The homologous allele where the reporter has not integrated will likely show Cas9-mediated changes at the target site, such as nucleotide insertions or deletions (**Figure 4a**).

3.3.5. For clones that show correct int. junction sequences after alignment, perform a PCR amplifying the whole targeted reporter and subject it to Sanger sequencing to verify the correct sequence of the integrant.

3.4. Southern blot analysis for verification of targeting in selected clones

Note: Southern blot analysis of selected clones is required to verify correct targeting and exclude Cas9-mediated recombination events that may have occurred at the targeted integration site.

3.4.1. Develop a strategy for appropriate gDNA digestion and probe the design prior to starting the experiment.

3.4.1.1. Select a restriction enzyme for gDNA restriction that generates appropriate fragments of 2 to 10 kb length at the targeted site. Certain restriction enzymes, such as *Asp*718, *Bam*HI, *Bgl*I, *Bgl*II, *Eco*RV, *Hind*III, *Nco*I, *Pst*I, *Pvu*II, *Sca*I, *Stu*I, and *Sst*I have been successfully used for digesting high molecular weight gDNA.

3.4.1.2. Design two different Southern probes: a reporter-specific probe and genomic probe. The reporter-specific probe hybridizes to a sequence within the reporter (*i.e.,* tdTomato-specific probe). The genomic probe hybridizes to a genomic region close to (but not overlapping) one HA.

3.4.1.3. Choose the genomic probe so that the fragment generated by gDNA digestion that will be detected by genomic probe binding differs in length (more than 2 kb) from the targeted and non-targeted alleles (**Figure 4b**). A probe length of 400 to 1000 bp is recommended.

3.4.1.4. Design PCR primers to amplify the two required probes. Amplify the reporter-specific probe from tv template using a high-fidelity DNA polymerase (see **Tables 3** and **4** for PCR ingredients and cycling conditions).

3.4.1.5. Amplify the genomic probe from wild-type Jurkat gDNA prepared with a commercial gDNA extraction kit using a DNA polymerase with proofreading activity (see **Tables 1** and **2** for PCR ingredients and cycling conditions). Purify the PCR products on an agarose/TAE gel and extract the fragments using a commercial gel extraction kit according to the manufacturer's instructions.

3.4.2. Extract high molecular weight gDNA from 1 x 10^7 cells of the wild-type Jurkat cells and selected clones from step 3.2.14.

3.4.2.1. Pellet the cells by centrifugation at 300 x g for 5 min at RT, wash it once with PBS, and suspend the pellet in 4 mL of lysis buffer [200 mM NaCL, 100 mM Tris-HCl pH 8, 5 mM EDTA, 0.1% SDS; then add 250-1000 µg/mL proteinase K (lyophilized powder, weigh in freshly)]. Incubate o/n at 55 °C, shaking at 350 rpm in a tabletop thermomixer.

3.4.2.2. Add 4 mL of isopropanol and mix by inversion 10 to 20 times. The gDNA should become visible as white precipitate. Spool the precipitated gDNA onto the fine tip of a glass pipette, wash by emerging in 750 µL of 70% EtOH, and let dry at RT (5 to 10 min).

3.4.2.3. Shed the precipitate into a 1.5 mL reaction tube containing 500 µL of 1x TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and leave to dissolve o/n at 4 °C, shaking at 350 rpm. Any pipetting of gDNA from this stage should be done with wide-bore tips to avoid shearing.
Note: Preparation of high molecular weight gDNA is essential for Southern blot analysis, and commercially available gDNA preparation kits are not suitable.

3.4.3. Digest (two times) 15 µg of gDNA of the selected clones and wild-type Jurkat cells with selected restriction enzyme (see step 3.4.1.1) in a 60 µL reaction with 6 µL of enzyme (20 Units/µL): first, add DNA, digestion buffer, and ddH_2O, incubate o/n at 37 °C, then add enzyme and incubate o/n at enzyme-specific digestion temperature. 15 ug of digested gDNA is required per Southern probe.

3.4.4. Use 7 µL of the 60 µL restriction digest for analytical gel electrophoresis on a 1% agarose/TAE gel. A smear indicates complete digestion and good DNA quality for Southern blot analysis.

3.4.5. Precipitate the remaining restriction digest by adding 1:10 3 M sodium acetate and 2 volumes 100% EtOH, then incubate for 1 h at -80 °C and centrifuge for 30 min at 15600 x g at 4 °C.

3.4.6. Discard the supernatant and wash the pellet with 70% EtOH. Centrifuge for 15 min at 15600 x g at 4 °C, discard the supernatant, let the pellet dry briefly at RT, and dissolve in 20 µL of ddH_2O.

3.4.7. Run 1% agarose/TAE blotting gel, loading 20 uL of digested gDNA per lane. Run the gel for 2 h at 60 V, 400 mA.

Note: The percentage of agarose gel and running time/voltage may be adjusted according to expected fragment size for Southern blot detection calculated in step 3.4.1.1. The following steps of Southern blot analysis are described in detail in a supplementary protocol (steps 1 to 18). These steps comprise of: washing of the blotting gel, blotting onto a nylon membrane, radioactive probe generation, probe hybridization, and development of autoradiograph film. Compare the obtained banding pattern after autoradiograph development in step 1 (supplementary protocol) with the expected pattern according to Southern strategy (for example results, see **Figure 4b**).

3.5. Analysis of off-target events

Note: Since CRISPR-Cas9-mediated genome engineering can generate off-target effects, PCR-amplify the ten highest-ranked *in silico*-predicted off-target sites in selected clones and subject them to Sanger sequencing.

3.5.1. Use CCTop[16] (http://crispr.cos.uni-heidelberg.de) to generate a list of the ten highest-ranked *in silico* predicted off-target sequences.

3.5.1.1. Input the gRNA sequence including PAM as used for targeting as the query sequence. Select “NGG” as PAM and “Human genome” as the reference for off-target prediction.

3.5.1.2. Set maximal total mismatches to “4” and target site length to the length of the gRNA without PAM. The output file will provide a ranked list of genomic off-target sites for respective gRNA.

3.5.2. *In silico* extract the genomic sequence 500 bp upstream and downstream of each of the ten highest-ranked off-target hits using UCSC Genome Browser (http://genome.edu.ucsc.edu) and the position of the off-target hit from CCTop results list.

3.5.3. For each off the target sites to be analysed, design a PCR primer pair that amplifies a fragment of 600 to 700 bp in length including the predicted off-target site.

3.5.4. Extract gDNA from the selected clones and Jurkat wild-type cells using a commercial gDNA extraction kit. For every off-target site, perform a PCR using a DNA polymerase with proofreading activity (see **Tables 1** and **2** for PCR ingredients and cycling conditions) on wild-type and the respective clone-derived gDNA.

3.5.5. Analyze 5 µL of PCR products on a 1.5% agarose/TAE gel. If correct band sizes are observed, purify the remaining PCR product using a commercial PCR purification kit and subject it to Sanger sequencing. Compare sequences of the off-target sites in Jurkat cells and the targeted clones.

REPRESENTATIVE RESULTS:

In this representative experiment we have chosen to target a minimal HIV-1-derived reporter consisting of a LTR, tdTomato-coding sequence, and polyA-signal sequence to two loci in intron 5 of the BACH2 gene[17]. The loci for targeting were chosen according to proximity to published recurrent integration sites found in different studies on primary T cells from HIV-infected patients [2,4-8] and the presence of a PAM motif NGG necessary for Cas9-mediated induction of double-strand breaks. Target vectors were constructed and transfected according to the described protocol. The workflow for transfection, checking for targeting events, generation of single-cell clones, and screening and selection of clones is schematically shown in **Figure 1**.

Two weeks after FACS-enrichment of transfected cells, we were able to detect reporter gene expression after PMA-Iono induction by flow cytometry in 4 to 12% of cells, depending on integration site (data not shown) and PCR products on genomic DNA spanning the whole 5' and 3' integration junction from upstream of 5' HA into reporter sequence and downstream of 3' HA into reporter sequence, respectively (**Figures 3a** and **3b**). Having confirmed that targeting events occurred, we went on to generate single-cell clones by limiting dilution plating. In our hands, plating 5 to 10 96-well plates per targeting construct was sufficient to obtain enough clones for a successful screen. In the protocol (section 3.2), it is described how to perform a screening of single-cell clones in duplicated 96-wells. In this regard, only positive clones must be expanded, which saves both time and effort. In **Figure 3c,** example data of FACS-screening after PMA-Iono induction and PCR screening are shown. We observed a number of clones with high, low, and no fluorescent reporter gene expression (**Figure 3c**). For PCR-screening, it is important to include a control PCR which amplifies a genomic locus of choice to determine whether the quality of cell lysates is adequate for PCR. In our case, we have chosen a 630 bp sequence in the NUP188 gene locus for control PCR (for primer sequences, see **Table 5**). Primers for screening PCR were then designed to amplify a shorter sequence located in the reporter. Additionally, PCR for a sequence on the target vector backbone was performed to exclude any clones which had unspecifically integrated target vector backbone sequences (backbone PCR, data not shown).

Pre-selected clones were then expanded and further analyzed for correct targeting by Southern blot and PCR and sequencing. Southern blot was performed using a reporter-specific probe and a probe specific for the genomic locus binding outside the reporter but not within the homology arms of the targeting vector. Interestingly, all the clones that were tested by Southern blot were positive in screening PCR beforehand, but only a portion showed correct band sizes in Southern blot analysis and had heterozygously integrated the reporter (**Figure 4b**). It is therefore necessary

to not only rely on screening PCR results but also verify correct targeting by Southern blotting. To verify correct targeting at the DNA sequence level, integration junctions were amplified by PCR and products were subjected to Sanger sequencing (**Figure 4a**). Notably, sequencing of the target site homologous allele, where the reporter had not integrated, revealed Cas9-mediated changes. In **Figure 4a,** examples for a Cas9-mediated deletion is shown. To test for Cas9-mediated off-target effects, a list of the highest-ranked off-target sites was generated as described in section 3.5. Ten highest-ranked off-target sites were PCR-amplified from genomic DNA of the clones, and the products subjected to Sanger sequencing. In the targeted single-cell clones we generated, no variations from Jurkat wild-type sequences were observed at the ten highest-ranked off-target sites.

[Place **Figure 3** here]

[Place **Figure 4** here]

FIGURE AND TABLE LEGENDS:

Figure 1: Workflow for CRISPR-Cas9-mediated targeting, generation, and selection of clonal reporter lines with defined integration site. (**A**) Generate the target vector and transduce Jurkat T cells with the target vector and Cas9/gRNA expression plasmid. (**B**) Enrich the transfected cells 72 h post transfection by FACS. (**C**) Let the cells grow for 10 to 14 days and confirm the occurrence of targeting events by PCR and flow cytometry. (**D**) Generate single-cell clones by limiting dilution and let clones grow for 3 weeks. (**E**) Screen the clones for correct targeting by PCR and flow cytometry in 96-well format. Expand selected clones. (**F**) Verify correct targeting in selected clones by Southern blot, PCR and sequencing, and analysis of off-target events of Cas9 endonuclease activity.

Figure 2: Targeting strategy and vector design. (**a**) gRNA and choice of homology arms. 20 nt gRNA is homologous to the chosen genomic target site and situated adjacent to a PAM. Homology arms are complementary to 1000 bp up- and downstream of the gRNA and should not include the gRNA sequence. (**b**) Schematics of targeting vector and gRNA/Cas9 vector. The targeting vector consists of the chosen reporter sequence that is 5' and 3' flanked by the homology arms. The gRNA/Cas9 vector is based on the pX330-U6-Chimeric_BB-cBh-hSpCas9 backbone. (**c**) Schematic of targeting by homologous recombination. Target vector and guideRNA/Cas9 vector are transfected into Jurkat cells. Cas9 mediates a double strand break at genomic target site (indicated by *) and facilitates homologous recombination and integration of reporter sequence into the genomic target locus.

Figure 3: Screening of single-cell clones for targeting events. (**a**) Schematic of primer design for detection of targeting events by PCR in mixed targeted cell population. Primer pairs for detection of 5' integration junction (P1, P2) and 3' integration junction (P3, P4) are indicated as arrows. Primer P1 and P4 also serve to amplify targeted locus of the allele without reporter integrant (**b**) Genomic DNA of mixed targeted cell population was prepared 10 days after FACS enrichment of transfected cells and analyzed for 5' integration junction (P1, P2), 3' integration junction (P3, P4;

data not shown) and wild-type allele of the targeted locus (P1, P4). Wild-type Jurkat cells (wt) served as control. (**c**) First screen of single-cell clones in 96-well plates. 96-well plates with single-cell clones were screened for correct targeting by flow cytometry after PMA-Iono induction and PCR analysis. Results for example clones are shown. PCR was performed on cell lysates with reporter-specific primers (screening PCR; P5, P6) and primers amplifying a genomic locus (control PCR, here: NUP188 locus; P7, P8). Cell lysates of wild-type Jurkat cells (wt) and genomic DNA from HIVisB2 clone (B2)[18] prepared with commercially available kit served as a negative control for screening PCR and positive control for control PCR, respectively.

Figure 4: Analysis of selected single-cell clones for correct reporter integration. (**a**) Sequence analysis of integration junctions and targeted locus on the allele without reporter integrant. Primer pairs for detection of 5' integration junction (P1, P2), 3' integration junction (P3, P4) and targeted locus of allele without reporter integrant (P1, P4) were used. PCR products were amplified from genomic DNA of single-cell clones and subjected to Sanger sequencing. Sequencing results were aligned to expected sequences. Shown are example data of one clone. Sequence chromatograms from genome/HA junction and HA/reporter junction are shown for both 5' and 3' integration junction. Matches are indicated as dots. Binding site of gRNA is marked with a box. Cas9-induced mutations are highlighted in red. Schematic of primer design is shown below. Arrows indicate primer positions used for amplification. (**b**) Southern blot analysis of selected targeted single-cell clones and Jurkat wild-type cells. Southern blot analysis was carried out on genomic DNA with a reporter-specific probe and a probe recognizing a genomic sequence in both targeted and wild-type allele (genomic probe). Data of 7 example clones is shown. Clones with correct band sizes are marked with boxes. Diluted target vector plasmid served as positive control (+). Schematic for Southern blot analysis is shown below.

Table 1: Recipe for PCR using a polymerase with proofreading activity. The amount of each reagent to be added per PCR reaction is indicated. PCR is intended for amplification using genomic DNA as template. The recipe is used in step 1.2.2.2 (amplification of homology arms), step 3.3.3 (verification of integration sites), step 3.4.1.5 (generation of genomic probe for Southern blot), and step 3.5.4 (analysis of off-target events).

Table 2: Cycling conditions for PCR using a polymerase with proofreading activity. Cycling conditions correspond to PCR recipe listed in **Table 1**.

Table 3: Recipe for PCR using a high-fidelity polymerase. The amount of each reagent to be added per PCR reaction is indicated. PCR is intended for robust amplification of DNA templates. The recipe is used in step 2.4.2 (detection of targeting events) and step 3.4.1.4 (generation of reporter-specific probe for Southern blot).

Table 4: Cycling conditions for PCR using a high-fidelity polymerase. Cycling conditions correspond to PCR recipe listed in **Table 3**.

Table 5: Oligonucleotide sequences for control PCR. Forward and reverse primers for the amplification of a 630 bp fragment of NUP188 locus are indicated.

Table 6: Recipe for screening-PCR. The amount of each reagent to be added per PCR reaction is indicated. PCR recipe is intended for screening PCR in step 3.2.11.

Table 7: Cycling conditions for screening-PCR. Cycling conditions correspond to PCR recipe listed in **Table 6.**

DISCUSSION:

Here, we describe a protocol to generate HIV-1-derived Jurkat reporter models with chosen proviral integration sites applying CRISPR-Cas9-based genome engineering.

Several points of the protocol require careful attention during the planning stage. First, the locus to be targeted should be chosen carefully, as some loci might be easier to target than others (*e.g.,* depending on the chromatin status of the region and the target sequence itself). Repetitive sequences are hard to clone into the targeting vector and are often not unique within the genome. Regions of repressive chromatin are harder to target with the CRISPR-Cas9 system[19,20].

Second, the choice of gRNA is crucial for CRISPR-Cas9-mediated targeting. For the purpose of generating models for HIV infection with representative integration sites, one would want the gRNA to bind as close as possible to an integration hotspot. However, this site might not be ideal for Cas9 recruitment; therefore, a compromise must be made between proximity of the gRNA sequence to the integration site of choice and gRNA quality. We have found the E-CRISP webtool reliable in predicting functional gRNAs. It is also possible to carry out a gRNA pre-test by transiently expressing several gRNAs together with Cas9 in the cell type to be targeted, followed by a screening for mutations. A suitable gRNA will direct Cas9 efficiently to the target site and the gRNA complementary site will show mutations. The length of HAs (1000 bp upstream and downstream of the integration site) was chosen according to previously published studies[11]. Generally, reducing the length of homology arms will result in reduced targeting frequency. A length of 1000 bp of homology arm presents a good compromise between sufficient targeting frequency and ease of target vector construction.

Third, enough time must be spent on a good design of the reporter construct. The minimal reporter used in this protocol, which contains an HIV-1 NL4-3 derived LTR and a tdTomato coding sequence, was designed based on the following principle: single LTRs have been described as solely remaining proviral fragments in several cases of clonal expansion in chronically HIV-1-infected individuals[21]. It is expected that the LTR is strongly influences the chromatin status at integration sites and organizes the recruitment of cellular complexes. We have chosen a minimal HIV reporter focusing on the HIV LTR as main regulatory genetic element, then introduced tdTomato as a fluorescent LTR activity marker instead of further HIV-1 genes, as genome engineering frequency was reported to be higher with smaller targeting inserts[11]. Considering the time-consuming steps of tv cloning, clonal selection, screening, and verification of the clones, it is advised to carefully consider the design of the HIV reporter in the context of functional studies that will eventually be carried out on targeted cell lines. One might, for example, consider

the inclusion of gag leader sequence, 3' HIV LTR, and/or other viral gene sequences in the reporter. The protocol can be readily adapted to such different reporter constructs.

Generally, it is important to consider the choice of cell type to be targeted as single-cell clone generation can be difficult with certain cell lines. We found that Jurkat cells are not very efficient in single clone generation; however, we decided to choose this cell type for its previously known use in latent HIV infection models[9]. We obtained the best results in Jurkat clonal dilution plating with 50% conditioned medium, when plates were left undisturbed in an incubator that was opened no more than 3 times a week. If using a different cell line is desired, it is advised to pre-test the possibility of generating single-cell clones by carrying out a dilution plating experiment. Another point to keep in mind is the variation of transfection rates of different cell lines. If transfection is not feasible for the chosen cell line, electroporating the cells for targeting may be necessary. Note that the choice of cell line used for targeting may not be determined by technical aspects only, such as efficiency for transfection or single clone generation. Functional aspects may also be considered, such as transcriptional activity or inducibility of the targeted gene locus. This might require screening of different cell lines prior to the targeting workflow.

It should be emphasized that the described protocol is time-intensive. It should be expected to take 3 to 6 months from the first step to final clonal cell lines. The Southern blot analysis, which is used to check for site-specific single integration events, may seem cumbersome, but in our experience it is highly important - as only a subset of single-cell clones that showed the expected integration junctions per PCR showed a correct patterning in the Southern blot. Ideally, experimenters should generate a number of clones with the same insert targeted to the same integration site to control for clonal effects in any subsequent experiments that make use of the clones. It is possible to do heterozygous as well as homozygous targeting. In heterozygous reporter cell lines, the allele without an integrant is likely to show modifications at the Cas9 targeting site, which can be screened by PCR (**Figure 4a**). For homozygous targeting, we suggest that both alleles be targeted consecutively.

Taking into account these considerations, this workflow provides a means to generate powerful cellular models that can be used to increase the understanding of chronic HIV infection. For example, proviral activity in response to latency-reversing agents can be tested in the context of recurrent integration sites. This may be of particular interest to researchers in the field, since position effects have been postulated to impact HIV latency and reversal[22].

ACKNOWLEDGMENTS:
We thank Britta Weseloh and Bettina Abel for technical assistance. We also thank Arne Düsedau and Jana Hennesen (flow cytometry technology platform, Heinrich Pette Institut) for technical support.

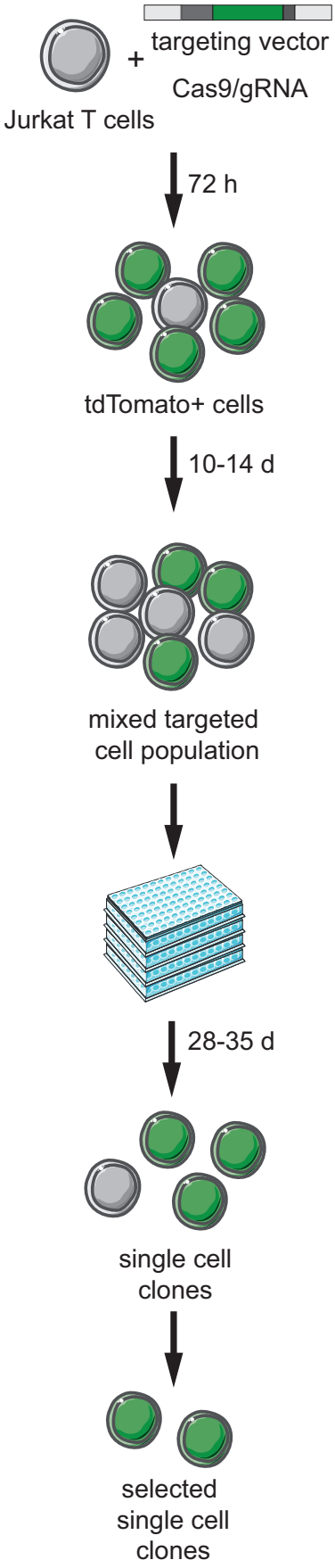
DISCLOSURES:
The authors have nothing to disclose.

REFERENCES:

1. Hughes, S. H., Coffin, J. M. What Integration Sites Tell Us about HIV Persistence. *Cell Host and Microbe.* **19**, (5), 588–598 (2016).
2. Marini, B., Kertesz-Farkas, A., *et al.* Nuclear architecture dictates HIV-1 integration site selection. *Nature.* **521**, (7551), 227–231 (2015).
3. Cesana, D., Santoni de Sio, F. R., *et al.* HIV-1-mediated insertional activation of STAT5B and BACH2 trigger viral reservoir in T regulatory cells. *Nature Communications.* **8**, (1), 498 (2017).
4. Cohn, L. B., Silva, I. T., *et al.* HIV-1 Integration Landscape during Latent and Active Infection. *Cell.* **160**, (3), 420–432 (2015).
5. Han, Y., Lassen, K., *et al.* Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *Journal of Virology.* **78**, (12), 6122–33 (2004).
6. Ikeda, T., Shibata, J., Yoshimura, K., Koito, A., Matsushita, S. Recurrent HIV-1 integration at the BACH2 locus in resting CD4+ T cell populations during effective highly active antiretroviral therapy. *The Journal of Infectious Diseases.* **195**, (5), 716–725 (2007).
7. Wagner, T. A., Mclaughlin, S., *et al.* Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science.* **345**, (6196), 570–573 (2014).
8. Maldarelli, F., Wu, X., *et al.* Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science.* **345**, (6193), 179–183 (2014).
9. Jordan, A., Bisgrove, D., Verdin, E. HIV reproducibly establishes a latent infection after acute infection of T cells *in vitro*. *The EMBO Journal.* **22**, (8), 1868–1877 (2003).
10. Mack, K. D., Jin, X., *et al.* HIV insertions within and proximal to host cell genes are a common finding in tissues containing high levels of HIV DNA and macrophage-associated p24 antigen expression. *Journal of Acquired Immune Deficiency Syndromes.* **33**, (3), 308–320 (2003).
11. Byrne, S. M., Ortiz, L., Mali, P., Aach, J., Church, G. M. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Research.* **43**, (3), 1–12 (2014).
12. ZhangLab CRISPR Genome Engineering Toolbox: Target Sequence Cloning Protocol. *Addgene website*. at <https://www.addgene.org/static/cms/filer_public/e6/5a/e65a9ef8-c8ac-4f88-98da-3b7d7960394c/zhang-lab-general-cloning-protocol.pdf> (2013).
13. Addgene, Moser, F. Gibson Assembly Protocol. *Addgene website*. at <https://www.addgene.org/protocols/gibson-assembly/> (2009).
14. Addgene Plasmid Cloning by PCR. *Addgene website*. at <https://www.addgene.org/protocols/pcr-cloning/> (2014).
15. Addgene Plasmid Cloning by Restriction Enzyme Digest (aka Subcloning). *Addgene website*. at <https://www.addgene.org/protocols/subcloning/> (2013).
16. Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., Mateo, J. L. CCTop: An intuitive, flexible and reliable CRISPR-Cas9 target prediction tool. *Public Library of Science (PLoS) ONE.* **10**, (4), (2015).
17. Lange, U. C., Bialek, J. K., Walther, T., Hauber, J. Pinpointing recurrent proviral integration sites in new models for latent HIV-1 infection. *Virus Research.* **249**, (March), (2018).
18. Bialek, J. K., Dunay, G. A., *et al.* Targeted HIV-1 Latency Reversal Using CRISPR-Cas9-Derived Transcriptional Activator Systems. *PloS ONE.* **11**, (6), e0158294 (2016).
19. Lee, C. M., Davis, T. H., Bao, G. Examination of CRISPR-Cas9 design tools and the effect of target site accessibility on Cas9 activity. *Experimental Physiology.* **103**, (4), 456–460 (2018).

20. Jensen, K. T., Fløe, L., *et al.* Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Letters.* **591**, (13), 1892–1901 (2017).
21. Simonetti, F. R., Sobolewski, M. D., *et al.* Clonally expanded CD4 + T cells can produce infectious HIV-1 *in vivo*. *Proceedings of the National Academy of Sciences.* **113**, (7), 1883–1888 (2016).
22. Chen, H. C., Martinez, J. P., Zorita, E., Meyerhans, A., Filion, G. J. Position effects influence HIV latency reversal. *Nature Structural and Molecular Biology.* **24**, (1), 47–54 (2017).

Figure 1



A: transfection

B: FACS enrichment

C: check for targeting events
- **flow cytometry**
- **PCR**

D: generation of single cell clones

E: first screen of single cell clones
- **flow cytometry**
- **PCR**

F: verification of targeting
- **Southern blot**
- **PCR/sequencing**
- **analysis of off-target events**

Figure 2

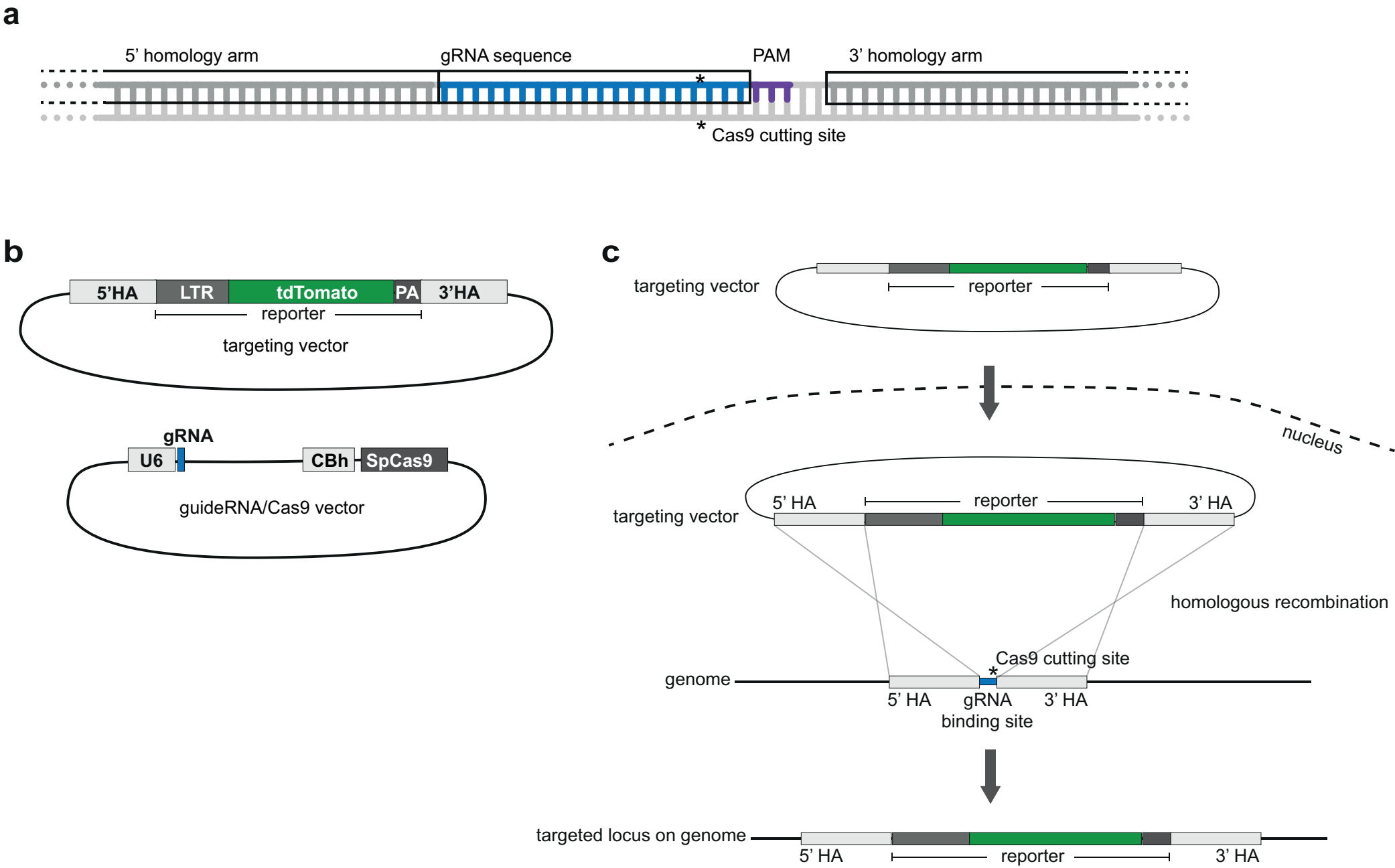


Figure 3

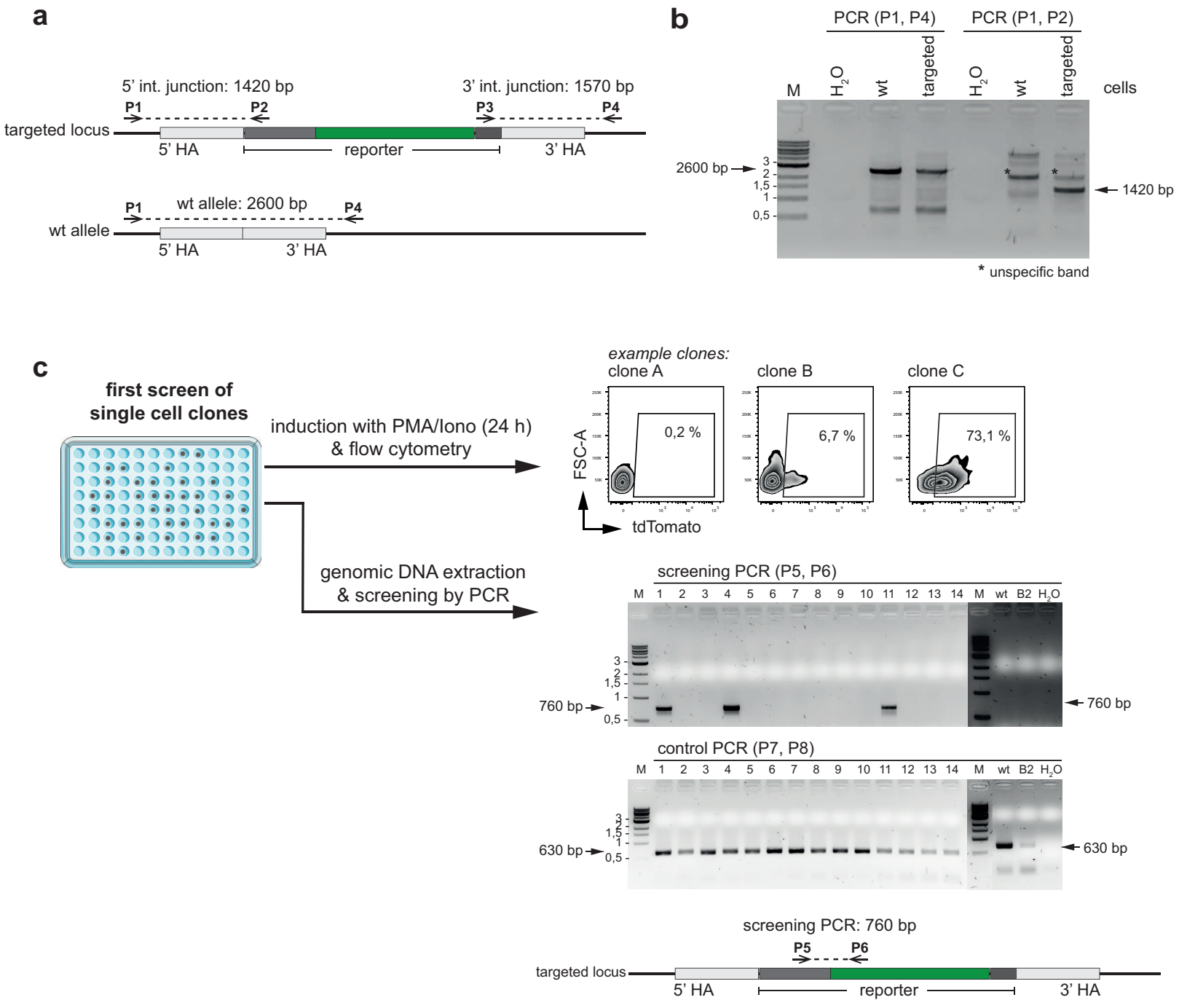


Figure 4

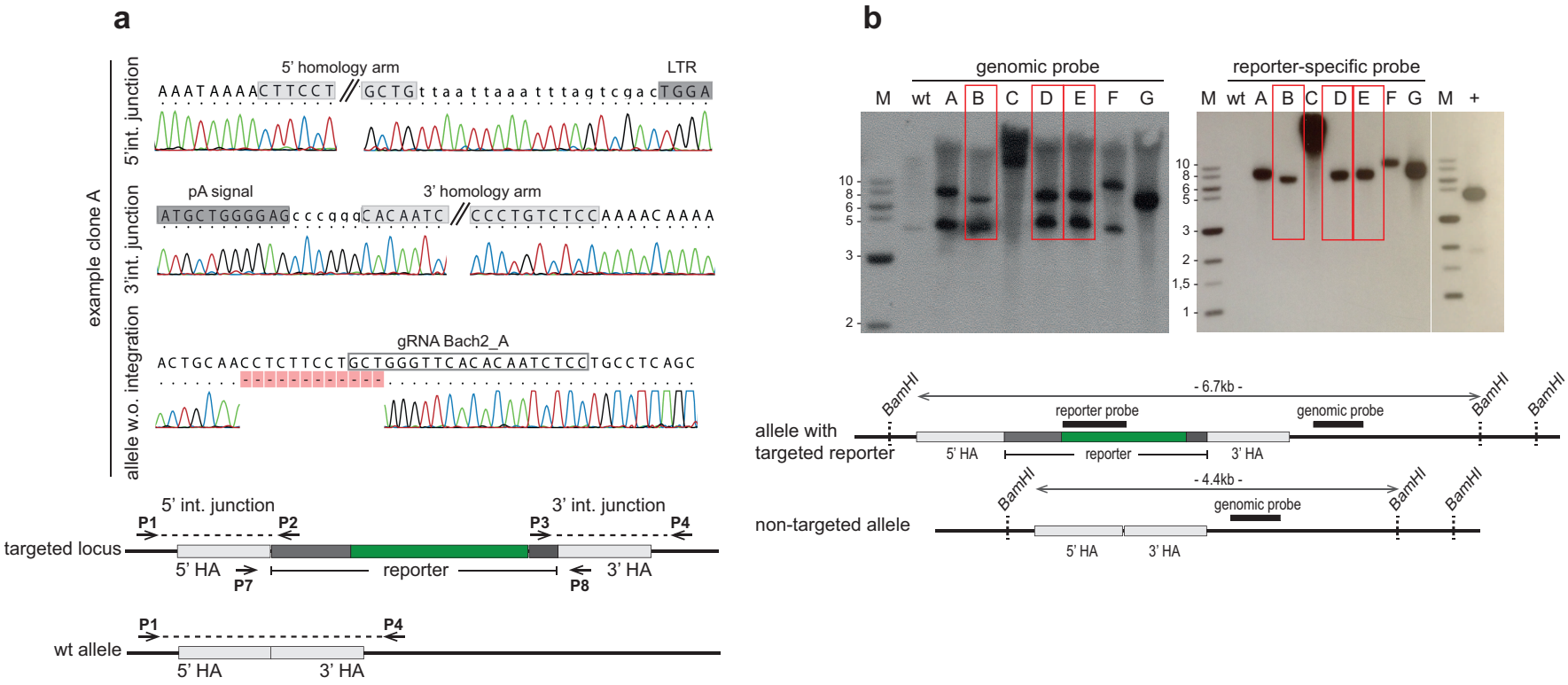


Table 1

Reagent	**Add per reaction**
Forward Primer (20 µM)	1 µL
Reverse Primer (20 µM)	1 µL
10x Taq buffer	5 µL
1 µl dNTPs (2.5 mM each)	1 µL
MgCl (50 mM)	1 µL
DMSO	1.5 µL
gDNA (50 - 100 ng/µL)	2 µL
Nuclease-free water	Fill up to 49.5 µL
Taq DNA polymerase (5 U/mL)	0.5 µL
reaction volume	**50 µL**

Table 2

Steps	Temperature	Time	Cycles
Initial Denaturation	95 °C	2 min	1
Denaturation	95 °C	40 s	30 -35
Annealing	58 °C	45 s	
Extension	72 °C	1 min/kb	
Final Extension	72 °C	10 min	1

Table 3

Reagent	**Add per reaction**
Forward Primer (20 µM)	1 µL
Reverse Primer (20 µM)	1 µL
5x high-fidelity buffer	5 µL
1 µl dNTPs (2.5 mM each)	1 µL
gDNA (50 - 100ng/µL) or plasmid DNA (50 ng/µL)	2 µL gDNA or 1 µL plasmid DNA
Nuclease-free water	Fill up to 49.5 µL
High-Fidelity DNA polymerase	0.5 µL
reaction volume	**50 µL**

Table 4

Steps	Temperature	Time	Cycles
Initial Denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	30 - 35
Annealing	58 °C	30 s	
Extension	72 °C	30 s/kb	
Final Extension	72 °C	10 min	1

Table 5

Primer	Sequence (5‘ – 3‘)
P7 control PCR F	CTTTGTTGGGTAAGCATGGAGGTC
P8 control PCR R	CAGTTACTCACCTTTGCACATAGG

Table 6

Click here to download Table Table 6.xlsx

Reagent	**Add per reaction**
ready-to-use PCR Master Mix	8 µL
Forward Primer (20 µM)	1 µL
Reverse Primer (20 µM)	1 µL
Nuclease-free water	8 µl
cell lysate (1:12 dilution)	2 µL
total reaction volume	**20 µL**

Table 7

Click here to download Table Table 7.xlsx

Steps		Temperature	Time	Cycles
Pre-PCR	Denaturation	94 °C	2 min	1
	Annealing	58 °C	1 min	
	Extension	72 °C	1 min	
Denaturation		94 °C	30 s	38
Annealing		58 °C	30 s	
Extension		72 °C	1 min	
Final Extension		72 °C	10 min	1

Table of Materials

Click here to download Table of Materials 58572_R2_JoVE_Materials.xls

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
pX330-U6-Chimeric_BB-cBh-hSpCas9	Addgene	42230	vector for expression of SpCas9 and gRNA
pMK	GeneArt		mammalian expression vector for cloning
cDNA3.1	Invitrogen	V79020	mammalian expression vector for cloning
BbsI	New England Biolabs	R0539S	restriction enzyme
NEBuilder Hifi DNA Assembly Cloning Kit	New England Biolabs	E5520S	Assembly cloning kit used for target vector generation
TaqPlus Precision PCR System	Agilent Technologies	600210	DNA polymerase with proofreading activity used for amplification of homology arms (step 1.2.2.2), verification of integration site and reporter sequence (step 3.3.3 and 3.3.5), generation of genomic probe for Southern blot (step 3.4.1.5) and analysis of off-target events (step 3.5.4)
96-well tissue culture plate (round-bottom)	TPP	92097	tissue culture plates for dilution plating
Phusion High-Fidelity DNA polymerase	New England Biolabs	M0530 L	DNA polymerase used for detection of targeting events (step 2.4.2) and generation ofreporter-specific probe for Southern blot (step 3.4.1.4)
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	D9170	dimethyl sulfoxide as PCR additive
Magnesium Chloride ($MgCl_2$) Solution	New England Biolabs	B9021S	$MgCl_2$ solution as PCR additive
Deoxynucleotide (dNTP) Solution Mix	New England Biolabs	N0447S	dNTP mixture with 10 mM of each nt for PCR reactions
5PRIME HotMasterMix	5PRIME	2200400	ready-to-use PCR mix used for screening PCR (step 3.2.11)
QIAamp DNA blood mini kit	Qiagen	51106	DNA isolation and purification kit
QIAquick PCR Purification Kit	Qiagen	28106	PCR Purification Kit
RPMI 1640 without glutamine	Lonza	BE12-167F	cell culture medium
Fetal Bovine Serum South Africa Charge	PAN Biotech	P123002	cell culture medium supplement
L-glutamine	Biochrom	K 0282	cell culture medium supplement
Penicillin/Streptomycin 10.000 U/ml / 10.000 µg/ml	Biochrom	A 2212	cell culture medium supplement
Gibco Opti-MEM Reduced Serum Media	Thermo Fisher Scientific	31985062	cell culture medium with reduced serum concentration optimized for transfection
TransIT-Jurkat	Mirus Bio	MIR2125	transfection reagent
phorbol 12-myristate 13-acetate	Sigma-Aldrich	P8139-1MG	cell culture reagent
Ionomycin	Sigma-Aldrich	I0634-1MG	cell culture reagent
Syringe-driven filter unit, PES membrane, 0,22 µm	Millex	SLGP033RB	filter unit for sterile filtration
Heracell 150i incubator	Thermo Fisher Scientific	51026280	tissue culture incubator
Amershan Hybond-N+	GE Healthcare	RPN1520B	positively charged nylon membrane for DNA and RNA blotting
Stratalinker 1800	Stratagene	400072	UV crosslinker
High Prime	Roche	11585592001	kit for labeling of DNA with radioactive dCTP using random oligonucleotides as primers
illustra ProbeQuant G-50 Micro Columns	GE Healthcare	28-9034-08	chromatography spin-columns for purification of labeled DNA

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: CRISPR/Cas9-based genome engineering to generate Jurkat reporter models for HIV-1 infection with selected proviral integration sites.

Author(s): Julia K. Bialek, Thomas Walther, Joachim Hauber, Ulrike C. Lange

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☒ Standard Access ☐ Open Access

Item 2: Please select one of the following items:

☒ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Ulrike C. Lange
Department:	DEPARTMENT OF ANESTHESIOLOGY
Institution:	UNIVERSITY MEDICAL CENTER HAMBURG-EPPENDORF
Title:	DR.
Signature:	U/L
Date:	1/06/2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Rebuttal letter

Ref: JoVE58572_R1
Title: CRISPR/Cas9-based genome engineering to generate Jurkat reporter models for HIV1 infection with selected proviral integration sites
Authors: Bialek JK, Walther T, Hauber J, Lange UC

We would like to thank the editors for the thorough review of our manuscript. We have addressed the comments as follows:

Comment [A1]: We enforce a 10-page limit on the protocol section. The current protocol section is close to 11.5 pages long. Can some portions be trimmed or moved to a supplementary file?

- We would suggest to move one part of step '3.4 Southern blot Analysis for verification of targeting in selected clones' to a supplementary file. In step 3.4 to 3.4.7 the purpose of Southern blot analysis, design of Southern probes and generation of Southern probes and preparation of gDNA from targeted single cell clones is explained. We would keep this part in the main protocol. The later steps (3.4.8 – 3.4.18) are more general, and we have moved them to a supplementary protocol.
- Please find the supplementary protocol with step 3.4.8 to 3.4.18 of the Southern blot protocol uploaded together with the main manuscript. We have added an explanatory note in line 541 to 548 of the revised manuscript which refers to the supplementary protocol.

Comment [A2]: Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. Please add more specific details (e.g. button clicks for software actions, numerical values for settings, etc) to your protocol steps. Please address the comments below.

- The protocol has been revised carefully and more details have been added, especially to step 1 and following. All comments have been addressed.

Comment [A3]: Some stes were adjusted for language and clarity. Steps that were not in the imperative voice were made into notes. Notes are generally not filmable and have been excluded from highlighting.

- We would like to thank the editors for the thorough review of our protocol.

Comment [A4]: Please provide additional details for inputs you enter in the „organim", „target reion", etc fields.

- Detailed steps for selection of gRNAs using E-CRISP webtool have been added in step 1.1.2.1 to 1.1.2.4.

Comment [A5]: Please mention all button clicks and field entries in order to esure filmability.

- Single actions for blasting of gRNA sequences have been added in step 1.1.3.1.

Comment [A6]: Please add this to the reference list and use superscripted citation numbering.

- Protocol for gRNA cloning in pX330 has been added as reference #12.

Comment [A7]: These were not included with the submission, please double check. Please add all materials used to the table of materials.

- All tables (table 1 - 7) have been uploaded together with the revised manuscript. Tables were missing in the original submission due to an uploading error.
- Materials used have been added to table of materials.

Comment [A8]: This note is redundant, Please see the line above and make steps crisper.

- Redundant parts has been removed. Cloning of targeting vector with homology arms has been divided into substeps (Steps 1.2.2.2 to 1.2.2.3) and steps have been clarified.

Comment [A9]: Reference for how to do this?

- Homology arms are cloned into targeting vector by restriction enzyme cloning. This information has been added to step 1.2.2.3 and a reference for restriction enzyme cloning has been added (reference 14 and 15).

Comment [A10]: Some parts are action steps and can be converted to steps and substeps in imperative tense.

- Action steps have been added to step 1.2.2.2. and 1.2.2.3.

Comment [A11]: To adhere to JoVE's poicy oon avoiding commercial biases, please replace this with a generic alternative. We have indicated all such instances in red font.

- Commercial language has been removed and replaced with generic terms.

Comment [A12]: Which medium? RPMI?

- Transfection reactions are set up with RPMI with reduced serum concentration optimized for transfection (Gibco Opti-MEM, see table of materials). This information has been added.

Comment [A13]: How and when are the cells counted?

- Cells are counted after pooling and before centrifugation and washing (line 221 of revised manuscript)

Comment [A14]: Wash how? Mention centrifuge speed and duration.

- After FACS sorting, wash cells once by adding 20 mL of RPMI w/ AB to sorted cells and centrifuge for at 300 x g for 4 min at RT. This has been added to step 2.2.3.

Comment [A15]: This was note defined so far.

- "RPMI w/ AB" is defined in step 2.2.2.

Comment [A16]: Where is this from ? Did you generate it (provide the step number)?

- The reporter sequence is assembled (cloned) in step 1.2.2.1. In our example we used HIV-derived long terminal repeat sequence (LTR), followed by tdTomato fluorescent reporter sequence and BGH polyA-sequence. However, experimenters are free to choose components of reporter and targeting vector depending on research question. Assembly of reporter sequence is described in 1.2.2.1. Components of reporter sequence (LTR, tdTomato sequence, BGH-PolyA) and rationale of reporter design is introduced in introduction (line 90 – 92 of revised manuscript) and in step 1.

Comment [A17]: What do you analyze? Mention gating strategy if relevant.

- Cells are analyzed for fluorescent reporter gene expression and gated on viable single cells based on size in forward and sideward scatter. This information has been added to step 2.3.4.

Comment [A18]: Do you plate them after sorting? How? Do you perform FACS sorting again after 10-14 days?

- Cells are not sorted a second time. After sorting in step 2.2.3, cells are cultivated for 10 to 14 days (step 2.2.4) and analyzed by FACS and PCR after these 10-14 days of expansion. This has been clarified in step 2.3.1. and in the Note after step 2.3.4.

Comment [A19]: Please provide the step number for greater clarity.

- Cells from step 2.2.4 after expansion of 10 – 14 days are used for gDNA preparation. This has been added to step 2.4.1.

Comment [A20]: Need reference or details

- Design of primers for integration junction PCRs is described in the Note after Step 2.4. (line 266 – 272 of revised manuscript). Primers have to be designed by the experimenter depending on the chosen integration site and reporter sequence. PCR reactions and cycling conditions are provided in table 3 and 4.

Comment [A21]: Not provided.

- Tables 1 – 7 are provided with the revised manuscript. Tables were missing in the original submission due to an uploading error.

Comment [A22]: Which cells?

- Cells from step 2.2.4 after expansion are used for generation of single cell clones. This information has been added to step 3.1.2.

Comment [A23]: Mention how and when cells are counted.

- Cells from step 2.2.4.are counted 10 – 14 days after expansion and used for single cell plating subsequently.

Comment [A24]: Mention magnification

- Magnification (4x) is mentioned in step 3.1.6.

Comment [A25]: Already containing 100 uL RPMI w/AB?

- Cells are transferred to a 96- well plate containing 100 µL medium. 100 µl of this cell suspension is transferred to an empty plate, to duplicate the plate. This has been clarified in step 3.1.7.

Comment [A26]: Mention primers and reacton conditions.

- Primers and reaction conditions are mentioned in step 3.2.10 and following. The note was intended to explain the time frame of Step 3.2.: Single cell clones have been picked and transferred to new 96-well plate and 96-well plate has been duplicated (step 3.1.6 – 3.1.8). Single cell clones on ‘stock plate’ are now expanding, while in the meantime ‘duplicate plate’ is used to screen single cell clones. This is described in detail in the following steps (3.2.1 – 3.2.12)

- To clarify, we have referenced the corresponding parts of the protocol (screening PCR step 4.2.4 – 3.2.12 and flow cytometry step 3.2.2. – 3.2.3) in the note in line 348 – 350 of the revised manuscript.

Comment [A27]: Mention gating /analysis strategy.

- Preparation for flow cytometry and gating strategy is explained in step 3.2.3. Please see also answer to comment A26.

Comment [A28]: Mention gating approach.

- Viable single cells are gated based on size in forward and sideward scatter and fluorescent reporter gene expression is analyzed. This information has been added to step. 3.2.3.

Comment [A29]: Unclear what is done and what we would film

- The reporter sequence contains one fluorescent reporter gene (tdTomato). Expression of fluorescent reporter is analyzed in step 3.2.3. The backbone of the plasmid used for target vector cloning might also contain a second fluorescent reporter gene, for example GFP, based on the chosen plasmid (this is explained in Note following step 1.2.2.3). If fluorescent expression resulting from target vector backbone is detected by FACS in step 3.2.3, this indicates unwanted integration of backbone sequences and corresponding clones have to be excluded from further analysis (step 3.2.13).
  We unhighlighted this section for the script.

Comment [A30]: We cannot film designing the primers, but I am wary of unhighlighting this as it will lead to a discontinuity. Please edit the step to add filmable actions if possible.

- One could show a screen showing the annotated sequence of the reporter sequence and the position of the primer pairs and length of PCR products *in silico*. We use CLC workbench software for this purpose.

Comment [A31]: Not provided

- All tables have been provided with the revised manuscript.

Comment [A32]: Please provide all tables in the revised submission as individual excel files.

- All tables have been provided with the revised manuscript.

Comment [A33]: Specifcally which size?

- The design of the reporter to be targeted (choice of fluorescent reporter, promoter, additional elements like polyA-sequences) is up to the experimenter, depending on research question. Therefore primer pairs for screening and backbone PCR (step 3.2.10) have to be designed individually as well. A concrete size of PCR products cannot be given in the protocol for screening and backbone PCR. For a positive result, PCR products should show the corresponding specific size. This has been clarified in the note following step 3.2.12.
  For positive control PCR, primer sequences are given in table 5. PCR products should show a specific band of 630 bp. This information has been added to 'note' following step 3.2.12.

Comment [A34]: Specifcally which size?

- See answer to comment A33.

Comment [A35]: Not filmable so i have unhighlighted. We can sometimes include brief non-filmable steps at the conclusion of the protocol section.

- One could explain in one sentence in the end as a conclusion, that clones have been screened by PCR and flow cytometry, and those which show positive results in both screens can be selected for further analysis.

Comment [A36]: I split this up, please verify and edit accordingly.

- The edit has been verified and confirmed.

Comment [A37]: What is the size again here?

- For analysis, PCR product sequences are aligned to expected sequences after Sanger sequencing. This has been clarified in step 3.3.5.
- Sizes of PCR products in step 3.3.4 cannot be given, as PCR primers are designed individually by the experimenter based on targeted locus and reporter sequence. An example is given in representative results (Figure 4a).

Comment [A38]: Please replace the commercial names.

- Commercial names (HighPrime in step 3.2.13.2 and illustra ProbeQuant in step 3.4.13.3.) have been replaced.
- (this part of the protocol has been moved to supplementary protocol. Comment [A38] shows up as Comment [A1] in supplementary protocol)

Comment [A39]: These need to be referenced in-text somewhere. Please upload all tables with your submission.

- All tables have been uploaded with the revised manuscript. Table 6 and 7 are referenced in step 3.2.11 (screening PCR).
- Recipes for PCR reactions vary depending on the polymerase and supplied buffer systems used (Phusion polymerase, TaqPlus precision polymerase etc.). We have used three different commercial polymerases in this protocol, each of which was tested extensively and showed to work best for the corresponding part of the protocol (amplification from gDNA or plasmid DNA, amplification of long fragments, more robust or more sensitive polymerases). We would like to share our experience which polymerase to use for which application, as this took some time and effort for optimization. We have now deleted commercial language from table headlines (Phusion polymerase, TaqPlus precision polymerase), but have stated the name of the polymerase in the table itself to make this information available to the experimenter.

1. Subsequently submit blotting gel to following washes:

1.1. Depurination: Add 1 L 0.25 M HCl in ddH_2O per gel, incubate for 20 min at RT on a rocking platform at 40 rpm, remove the solution and rinse once with ddH_2O.

1.2. Denaturation: Add 1 L 0.5 M NaOH + 1.5 M NaCl in ddH_2O per gel, incubate for 30 min at RT on a rocking platform at 40 rpm, remove the solution and wash once with ddH_2O.

1.3. Neutralization: Add 1 L 0.5 M Tris pH 7.4 + 1.5 M NaCl in ddH_2O per gel, incubate for 30 min at RT on a rocking platform at 40 rpm, remove the solution and wash once with ddH_2O.

2. Blot the gel o/n at RT onto a positively charged nylon membrane through capillary transfer (see manufacturer's recommendations) using 20x saline-sodium citrate (SSC) buffer (0.3 M sodium citrate, 3 M NaCl).

3. After blotting, dry membrane for 30 min at 80 °C and crosslink using an UV crosslinker. At this stage, membranes can be stored at RT for several days.

Note: The following steps involving radioactive probe generation, hybridization and post-hybridization washes should be carried out in an appropriate facility for work with radioactive materials.

4. Prepare Church hybridization buffer. For this, first prepare pH buffer (720 mL 1 M Na_2HPO_4 + 280 mL 1 M NaH_2PO_4). To prepare Church buffer supplement 500 mL Ph buffer with 10 g BSA, 2 mL 0.5 M EDTA, 70 g SDS and fill up to 1 L with ddH_2O.

5. Transfer blot membrane into glass hybridization tube and add 15 mL Church buffer. Pre-hybridize for 1 h at 65 °C rotating. Use different hybridization tubes for each probe.

6. Prepare radioactively labelled probes.

6.1. Denature 25 ng of probe DNA prepared in step 3.4.2 in 20 µL ddH_2O for 10 min at 95 °C in screw cap tubes and transfer onto ice.

6.2. Add 4 µL of 5x concentrated random primer mix containing 1 U/mL Klenow polymerase, 0.125 mM dATP, dGTP and dTTP each in 50% (v/v) glycerol (commercial kit for random-primed labeling of DNA) and mix thoroughly by pipetting. Add 2 µL 10 mCi/mL [α-^{32}P]dCTP, mix and incubate for 10 min at 37 °C. Stop the reaction by adding 2 µL 0.2 M EDTA (pH 8).

6.3. Remove non-incorporated [α-^{32}P]dCTP using commercial chromatograpy spin-colums for purification of labeled probes according to the manufacturer’s recommendations.

6.4. Denature radioactive labelled probe by incubation for 10 min at 95 °C and transfer immediately onto ice.

7. Add denatured probe to the 15 mL Church buffer of the pre-hybridized membranes. Hybridize o/n at 65 °C rotating.

8. After hybridization, discard hybridization buffer and rinse membranes once in prewarmed low stringency buffer (2x SSC + 0.1% SDS). Subsequently wash membranes in 20 mL prewarmed low stringency buffer at 65 °C rotating for 20 min.

9. Discard first wash buffer and incubate in 20 mL prewarmed high stringency buffer (0.5x SSC + 0.1% SDS) for 10 to 20 min at 65 °C rotating. Discard second wash and rinse membranes once in high stringency buffer.

10. Drain excess liquid from washed membranes using filter paper, wrap membranes in clear plastic foil and expose to autoradiography film in appropriate cassette o/n at -80 °C.

11. Develop autoradiograph film. Compare obtained banding pattern with expected pattern according to Southern strategy (for example results see **Figure 4b**).