An integration-defective lentivirus-based resource for site-specific targeting of an edited safe-harbour locus in the human genome

R Torres[1], A Garcia[1], M Jimenez[1], S Rodriguez[2] and JC Ramirez[1]

Optimized gene transfer into human cells are still challenging the promise of human stem and induced pluripotent stem cells as resources for disease models, diagnostic screens and personalized cell therapy. These potential applications require precise control of the spatio-temporal action of gene switches and the coordinated regulation of modulators, effectors and differentiation factors during pluripotency, differentiation and homeostasis. Most studies require identical transgene environments for comparable analysis; however, this cannot be achieved by standard methods for transgenesis in human cells because of unintended epigenetic modifications, genetic instability, dose-dependent effects, and disruption or activation of host genes. Although gene targeting can circumvent these problems, human cells have proved difficult to target, and there is therefore a need to develop tools for targeted transgenesis at efficiencies similar to those achieved in mice. We present a simple strategy, KASTRINA 2.0, for reliable transgenesis in human cells, based on targeted recombinase-mediated cassette exchange and the safe episomal status conferred by integrase-deficient lentivirus (IDLV). By driving limited *cre* recombinase expression, the IDLV yields single site-specific recombination of a selectable donor cassette (TRINA) at the 'safe-harbour' *AAVS1* locus previously edited by zinc-finger nuclease to contain an acceptor site (KAS2.0).

Keywords: non-integrative lentivirus; site-specific integration; RMCE

INTRODUCTION

Transgenesis procedures in which the position and dosage of the inserted transgene are uncontrolled can lead to unpredictable effects due to disruption of host gene function, compromising data interpretation and placing severe limits on the applicability of this technology. Current methodologies for transgenesis suffer from a strong experimental effect due to the uncontrolled position and number of inserted transgenes. This leads to unpredictable outcomes that compromise data interpretation in research contexts and severely limits the potential clinical application of these technologies. To overcome these problems, procedures for gene gain- or loss-of-function must avoid all genome effects other than the desired modification. The almost 30 years since the first gene therapy assays[1] have witnessed extraordinary advances in systems for gene delivery and the control of transgene expression. These advances have been joined by techniques for the introduction of foreign gene cassettes at specific genome sites, the control of gene expression by RNAi, conditional knockout and knockdown and the study of specific cell-fate decisions. Development of these techniques for induced pluripotent stem cell (iPSC) applications promises the ability to manipulate disease pathways, genetically repair disease, and generate reporter cell lines to facilitate isolation of specific cell types. However, realizing these applications will require reliable tools for safe genetic modification of human cells. Targeted transgenesis in human cells has proved time consuming and unreliable, with human cells refractory to methodologies that promote high-frequency homologous recombination in mouse cells. These problems can be partially overcome by procedures for single-copy insertion at a specific genomic locus and by the use of vectors designed to deliver the cargo in a cell-specific manner. An important development in this area is the use of genome-editing approaches that co-opt endogenous DNA damage sensor pathways for the repair of homologous sequences. For example, engineered zinc-finger nucleases have been used to drive targeted integration of selectable markers in cells of diverse origin,[2,3] and new resources have recently become available, including meganucleases,[4] TAL effector nucleases (TALENS)[5] and more recently bacterial clustered, regularly interspaced, short palindromic repeats and its associated protein (CRISPR/Cas) system.[6] Another important technology is the rearrangement of DNA segments by site-specific recombinase-mediated integration (known as recombinase-mediated cassette exchange or RMCE), a powerful tool that targets defined genetic elements while having minimal secondary effects on host gene expression.[7] However, few studies have examined the potential of such approaches for targeting sites in the human genome.

The ideal sites for transgene insertion are 'safe harbours', sites susceptible to disruption and that support transcriptional activity in distinct cell types with minimal interference in host gene function.[8] The *AAVS1* locus, mapping to chromosome 19 and encoding the *PPP1R12C* gene, is widely regarded as innocuous for targeting, a view supported by the preferential integration of adeno-associated virus (AAV) sequences at this locus during natural infection.[9] Disruption of *PPP1R12C* by zinc-finger nucleases is safe and does not alter the pluripotency of human induced pluripotent stem cells (hiPSCs) or human embryonic stem (hES) cells;[2] moreover, the locus is transcriptionally active in primary and transformed cell lines.[10,11] The *AAVS1* locus thus displays the two essential characteristics of a transgene safe-harbour, a feature that

[1]Viral Vector Facility, Fundacion CNIC, Melchor Fernandez Almagro 3, Madrid, Spain and [2]Molecular Cytogenetics Group, Fundacion CNIO, Melchor Fernandez Almagro 3, Madrid, Spain. Correspondence: Dr JC Ramirez, Viral Vector Facility, Fundacion CNIC, Melchor Fernandez Almagro 3, Madrid 28029, Spain.
E-mail: jcramirez@cnic.es

probably evolved through in-field trial and error during AAV-cell co-evolution to select locus combining cell integrity with viral genome surveillance.

We previously demonstrated the use of integrase-defective lentivirus (IDLV) for targeted RMCE-mediated transgenesis at an integrated genomic landing pad in human cell lines.[1,10] Here we refined the procedure by combining the precise integration by RMCE with the unambiguous targeting achieved by the zinc-finger nuclease to develop a procedure for safe and reliable transgenesis into the human genome. The IDLV (TRINA, for transfer into acceptor) achieves rapid and efficient *cre* recombinase-mediated transgenesis at an *AAVS1* locus previously modified to bear a single-acceptor cassette (KAS2.0, for knock-down/in acceptor site version 2).

The method is designed to allow several replacements by repetition of the exchange reaction using IDLVs, with no need for further zinc-finger nuclease targeting. The major advantage of the KAS2.0 cassette is that it allows promoter trapping and double-positive/negative selection at each step, thus enabling selection without additional genetic analysis to detect cells undergoing proper recombination, as selection of random lentiviral integrations is negligible. Here we present data demonstrating single interchange at the *AAVS1* locus, define the optimal conditions for efficient RMCE by modulating enzyme levels and suggest the most convenient manner for delivery of the donor cassette.

RESULTS

Zinc-finger nuclease targeting inserts a single copy of the KAS2.0 landing pad at the *AAVS1* locus

We recently described a system for efficient, MOI-dependent cre-mediated RMCE of the donor cassette TRINA (transfer into acceptor) into randomly integrated copies of the acceptor cassette KAS (knock-in/out acceptor site).[10] Here we have redesigned the KASTRINA system by introducing additional features to the acceptor cassette. The new cassette, KAS2.0, includes dual selectable markers in a polycistronic expression cassette (Figure 1a). The KAS genomic landing pad contains a PGK promoter that controls the expression of the bi-cistronic eGFP-*neo* cassette as a selection marker, spaced by the porcine teschovirus CHYSEL 2A sequence and floxed by heterospecific lox sites (loxP/lox2272). The floxed construct is preceded (5′) by the cytomegalovirus (CMV) early enhancer element and the chicken beta-actin promoter (CAG) followed by a strong transcription stop signal. At the 3′ terminus is a promoterless second selectable marker (*hph*, hygromycin B phophotransferase). To achieve site-specific integration by zinc-finger nuclease-driven homologous recombination, the construct is flanked by the left and right homology arms of the human integration site for adeno-associated virus (*AAVS1*). Human HEK293A cells were electroporated with mRNA encoding zinc-finger nuclease (ZFN) and co-transfected with the KAS2.0 plasmid (pKAS2.0). Quantification after 10 days of selective culture showed that 0.5–1% of G418R cells gave rise to clones (Supplementary Figure 1). Twenty-four individual clones were picked and expanded under G418 selection, and randomly selected twelve of these were further characterized in detail. We first assessed ZFN-driven integration of the KAS2.0 construct by conventional PCR of genomic DNA. We used three primer pairs to discriminate between random and site-specific integration (Figure 1b): AL/KL (*AAVS1* left + KAS left), KR/AR (KAS right + *AAVS1* right) and KMf/KMr (KAS middle forward + KAS middle reverse). Eleven out of twelve clones were positive for insertion of KAS2.0, as shown by detection of the 1.4 kb band (Figure 1b). In five clones (*S1K7, S1K8*, *S1K16*, *S1K20* and *S1K22*), integration in the *AAVS1* locus by homologous recombination was confirmed by the presence of 1.2 and 0.9 kb bands (left arm and right arm, respectively; see Figure 1b, (+) labels below the gel). The ~30% frequency of site-specific integration (5 out of 11 clones) is comparable to previous reported data.[4,8] Analysis of the remaining eleven clones revealed a similar frequency (3 out of 12 clones, data not shown). To accurately determine the copy number of integrated KAS sequences, we detected *eGFP* in the 12 clones by quantitative PCR using primers hybridizing at the eGFP region (Figure 1b, lower chart). Cell numbers were determined with primers to the single-copy gene *albumin*; this gene maps to chromosome 4 (4q13.3; NM_000477A), which FISH analysis showed to be diploid in the hypotriploid HEK293A cell line. The analysis revealed that clones *S1K8, S1K16, S1K20 and S1K22* contained a single-KAS copy per diploid DNA content.

Two clones (*S1K8* and *S1K16*) were selected by the aforementioned criteria of single-band amplicon corresponding to both HR arms as well as on copy per cell by qPCR (Figure 1b) and were further analysed to confirm that integration occurred exclusively at the *AAVS1* locus. The KAS2.0 integration site was studied by FISH analysis with a BAC-derived probe specific for the *AAVS1* site (chromosome 19) and another probe specific for the KAS2.0 insert. Consistent with the hypotriploid karyotype of HEK293A (CRL-1573) cells, the BAC-probe (red) labelled three copies of chromosome 19p13.3q (Figure 1c left, arrowhead). The KAS2.0 probe (green) labelled only one of the three chromosome 19 copies (Figure 1c left, arrow), indicating specific targeting by ZFN and confirming that clones *S1K8* and *S1K16* are composed exclusively of cells containing a single, site-specific integrated copy of the KAS2.0 cassette. To study the insertion on a molecular level, we performed a Southern blot analysis of the AAVS1 locus in these two selected clones. Genomic DNA was digested with *Eco*RV and probed with either an AAVS1 or an eGFP-specific probes (Figure 1c, right panel). As showed, both clones have a single-GFP insertion (Figure 1c right, upper 10.9Kbp band) that maps at the AAVS1 locus (Figure 1c right, lower 5.4 Kbp band).

A central issue in gene targeting is the stability of phenotype conferred by the recombination event, particularly under conditions of non-selection pressure. Culture of the five positive clones under non-selection conditions confirmed stability of the eGFP + phenotype over more than 30 population doublings (> 2 months), and clonogenic assay in the presence of G418 verified that cells remained G418R (Figure 1d and Supplementary Figure 2).

The *AAVS1*-integrated KAS2.0 genomic landing pad supports lentivirus-mediated RMCE

We next examined the ability of integration-deficient lentivirus encoding the TRINA cassette (Figure 2a) to promote RMCE in clones *S1K8* and *S1K16*. This vector is designed to unidirectional insert cherry fluorescent protein (mCherry) and the EF1α promoter between the KAS loxP sites upon expression of the cre recombinase (Figure 2a). Legitimate recombination leads to replacement of the selectable eGFP-2A-*neo* cassette flanked by heterospecific loxP sites in KAS with a dual-expression cassette encoding mCherry and hygromycin B phosphotransferase (*hph*) (Figure 2a). The TRINA lentiviral construct has no promoter in the LTR region, and is thus unlikely to be expressed when aberrantly integrated the genome, so that expression of the mCherry-*hph* cassette will occur only after legitimate recombination.[5,10] Thus, in a single step, this promoter-trapping strategy creates the capacity for both positive selection (gain of mCherry + hygro^R) and negative selection (loss of G418^R and eGFP +). Clones *S1K8* and *S1K16* were transduced with TRINA at high viral load (MOI = 3 TU per cell) to ensure transduction of a large number of cells. At 72 h post-transduction, cultures were split and maintained under hygromycin B selection for 10 days. Three subclones each were then picked from cultures of *S1K8* (*S1K8T1* to *3*) and *S1K16* (*S1K16T1* to *3*) and amplified for 1 month under non-selection conditions. Subsequent culture in selective media confirmed the inability of the subclones to grow in the presence of G418 and their capacity for growth in medium containing hygromycin B

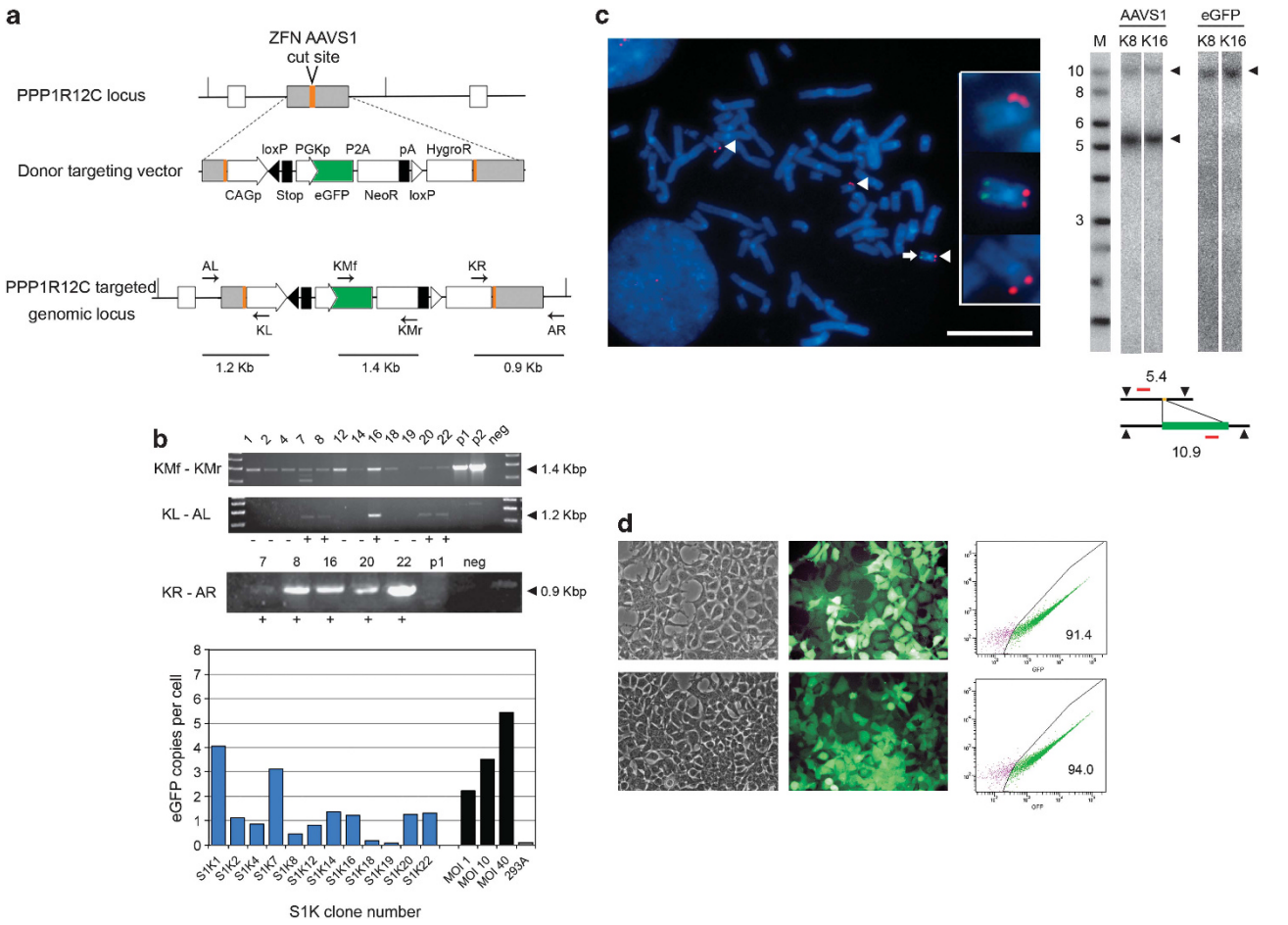

Figure 1. The genomic landing pad KAS2.0 targets the *AAVS1* site in human HEK293A cells through the action of zinc-finger nuclease. (**a**) Editing strategy for the generation of HEK293 cells bearing the KAS cassette in the *AAVS1* locus. The scheme shows the main elements in the KAS2.0 genomic landing pad in the donor targeting vector (*top*), and the anticipated genomic structure of the edited *AAVS1* locus (*bottom*). Promoters are represented by arrowed boxes, loxP sites by triangles and ORFs by open boxes. The target locus and homologous vector sequences are shaded grey, with the zinc-finger nuclease cut site in orange. Arrows show positions and orientations of primers used for PCR, and predicted amplimer sizes are shown underneath (A, *AAVS1*; K, KAS; L, M, R = left, middle and right; f, *r* = forward and reverse). (**b**) (*Upper panel*) PCR genotyping of KAS2.0 integration in 12 of 24 G418^R (neo^R) clones isolated. Genomic DNA of neo^R clones was amplified with the indicated primer pairs to detect integration of the KAS2.0 cassette (KMf/KMr product; 11 out of 12). Integration in the *AAVS1* site was detected in five clones (indicated by (+) label) by analysis of the left and right junctions (KL-AL and KR-AR products). Numbers identify the isolated neo^R clones. Amplimer sizes are indicated to the right. (p1, p2 plasmid DNA as positive control; neg, DNA from untransduced cells). (*Bottom panel*) Copy number of KAS integrated in twelve *S1K* clones (blue bars). *eGFP* was detected by qPCR of genomic DNA and quantified relative to the expression of *albumin*, which is diploid in HEK293A cells (see text for details). Of the five selected positive clones from the PCR analysis, only S1K8, S1K16, S1K20 and S1K22 have just one copy per diploid DNA content. The numbers of cells analysed for each clone were between 50 000 and 80 000. (MOI indicates values obtained with the parental transduced 293A cells at indicated values; 293 is for non-transduced parental cells). (**c**) (*Left*) FISH analysis of clone *S1K8*, showing labelling of the three copies of chromosome 19 (red probe) in the HEK293A cell line. Enlarged views are shown in the inset. One copy is positive for the specific KAS cassette signal (green) in the long arm of chromosome 19 (19q13), excluding the presence of copies of KAS at other loci. (*Right*) Southern blot analysis of the S1K8 and S1K16 clones. Genomic DNA was digested with *Eco*RV and blotted. Membranes were probed against AAVS1 (left lanes) or eGFP (right lanes). Positive hybridized bands are marked with an arrowhead. The scheme at the bottom represents the wild-type allele and the targeted allele, respectively, showing the relative position of the probes and the expected size of the fragments. Arrowheads, *Eco*RV site. M, molecular weight marker in Kbp. K denotes *S1*K clone (**d**) The KAS2.0 cassette is stable in targeted clones. Clone *S1K8* was cultured through 30 population doublings in the absence of selection and then cultured to confluence in the presence of G418. Resistance and eGFP expression were maintained throughout the 15 weeks of culture, as determined by a clonogenic assay and FACS analysis of cells maintained in non-selective medium (upper lane) or selective medium (lower lane). Almost all cells were neo^R and eGFP$^+$.

(Figure 2b left). As anticipated, all the *S1K8T* and *S1K16T* hygro^R clones lost eGFP expression and became mCherry$^+$ as analyzed by microscopy (Figure 2b right). To test the fidelity of the RMCE reaction, we first checked whether integration of TRINA occurred at the KAS2.0 site in three *S1K8T* and three *S1K16T* clones. Genomic DNA was amplified by conventional PCR using different pairs of primers specific for the insertion sites (KAS left + TRINA left and TRINA right + KAS right; Figure 2c). PCR products were compared with the signal obtained with internal primers targeting the EF1α promoter in the TRINA sequence. In all six clones, the TRINA sequence was integrated exclusively in the KAS genomic landing pad.

To check that there were no sequence alterations at the join sites, the PCR products from *S1K8T1*, *S1K8T2* and *S1K8T3* cells were TA-cloned into pGEMT, and for each clone 10 independent *E. coli DH5alpha* colonies were sequenced with universal primers.

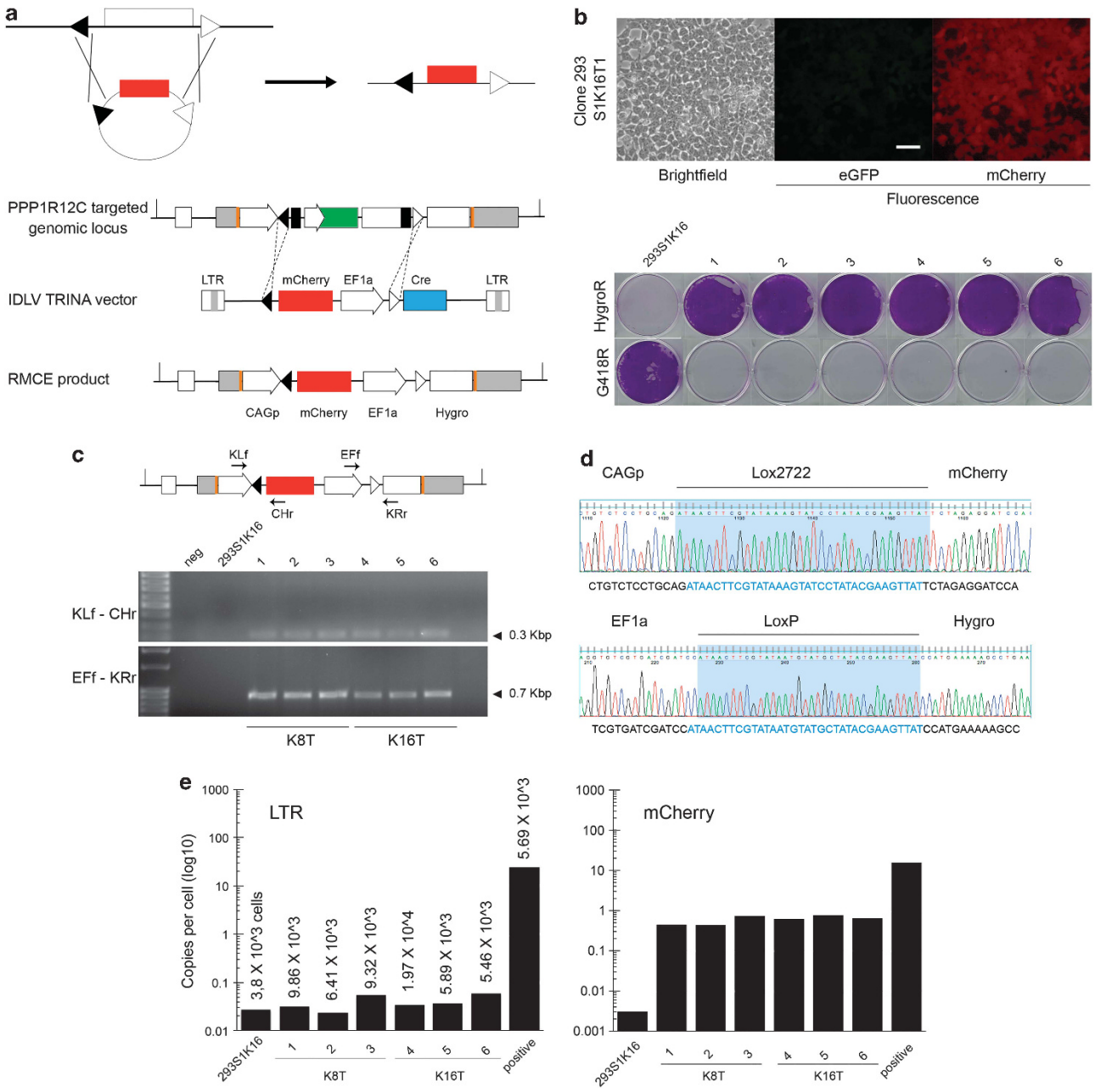

Figure 2. IDLV-TRINA exclusively targets the single KAS-edited *AAVS1* locus, with no detectable non-specific provirus integration. (**a**) Anticipated RMCE reaction in HEK293 cells bearing the KAS cassette in the *AAVS1* locus upon transduction with the IDLV-TRINA vector. The upper scheme is an overview of the anticipated RMCE reaction detailed in the lower scheme. The red and blue boxes in IDLV-TRINA depict the mCherry and cre expression components of the promoterless cassette. The scheme of the RMCE product highlights the promoter trapping strategy, in which the integration of TRINA components places mCherry under the control of the PGK promoter in the KAS2.0 cassette (displacing eGFP and Neo^R) and introduces the EF1α promoter to activate hygromycin B resistance. The scheme of the PPP1R2C-targeted locus is represented as in Figure 1 (**b**) Cultures of the parental clones *S1K-8* and *-16* or their TRINA-derived clones (*K8T* and *K16T*) were selected during 10 days in medium containing hygromycin or G418. Viability of TRINA-derived clones in hygromicin B (crystal violet staining) reveals cre-driven promoter trapping and legitimate RMCE between KAS and TRINA cassettes. The micrographs show fluorescence microscopy analysis of clone *K16T1* subsequently grown without selection for 2 weeks, confirming expression of mCherry instead of eGFP. (**c**) PCR genotyping of clones *K8T* and *K16T* using the primer pairs KLf plus CHf or EFf plus KRr (indicated on diagram) to confirm insertion in the *AAVS1* site by replacement of the KAS cassette. (**d**) Confirmation of correct gene insertion by sequencing the genomic locus at the join ends. The PCR products of clones *K8T1–3* indicated in (**c**) were cloned and sequenced. The figure shows the sequence of clone *K8T3* at the join ends around the lox sites. (**e**) Copy number of TRINA integrated in three *S1K8T* and three *S1K16T* clones. *mCherry* was detected by qPCR of genomic DNA and quantified relative to the expression of *albumin*, as before. All six *S1KT* clones contained a single TRINA copy per diploid DNA content (*left*). Analysis of random insertion with primers for the LTR sequence, which is not integrated in the RMCE reaction, detected <1 integration per cell. The numbers of cells analyzed for each clone are indicated above the bars. *293S1K16* indicates values obtained with the parental non-TRINA-transduced *S1K16* clone; *positive* indicates results obtained with the parental clone transduced with integrative lentivirus at high MOI.

The sequence of a 500 bp region flanking the RMCE product was identical for all clones and corresponded to the predicted region spanning the cre cut-and-join reaction site, thus confirming correct targeting of the *AAVS1* locus (Figure 2d). Given that cells in all the clones each contain just one KAS genomic landing pad (Figure 1d), these results confirm the high fidelity of the KAS2.0-TRINA RMCE reaction. To accurately determine the copy number of integrated TRINA sequences, we detected *mCherry* by qPCR as aforementioned (Figure 1b). All six *S1KT* clones contained a single TRINA copy per diploid DNA content (Figure 2e, left), compared with ~20 integrated copies per cell achieved with an integrative lentivirus.

IDLV-TRINA is a circular entity containing either one or two LTRs. This super helical-circular state is a major requirement for efficient RMCE and serves to prevent random integration. To determine the level of random insertion, we detected lentiviral sequences using LTR-specific primers, as the LTR sequence is not integrated in the RMCE reaction (Figure 2a). In cells cultured under hygromycin selection for 2 weeks, the number of lentiviral integrations was below one (Figure 2e, right), indicative of the absence of both random integrated IDLV-TRINA and RMCE off-products containing LTR sequences. This is similar to the number detected in control non-transduced parental *S1K16* cells, indicating that no IDLV DNA was present in any of the six clones after hygromycin selection.

These results confirm that integration of TRINA occurs only at the target locus and that the *cre* recombinase delivered by the TRINA IDLV catalyzes 100% efficient cassette exchange in hygro^R plus mCherry cell clones at the *AAVS1* site.

The cre expression level determines RMCE efficiency and can be optimized by separate delivery

The RCME reaction in KASTRINA 2.0 is auto-limiting because it renders the cre-encoding sequence in retro-transcribed TRINA promoterless assuming that RMCE leads to the loss of the EF1α promoter (Figure 2a). Donor vector has commonly to be applied at an excess to drive RMCE to be efficient, at least in case of Flp recombinase.[12] To test whether the cre levels produced limit the efficiency of RMCE beause of reduced turnover or enzymatic properties, as reported for other models,[6,12,13] we assessed independent delivery of the cre enzyme and the TRINA vector (Figure 3a). We constructed a TRINA vector lacking the cre cassette (IDLV-TRINA.Δcre) and a separate IDLV vector encoding only cre (IDLV-cre) under the control of the CMV promoter, a stronger promoter than the EF1α promoter used to drive cre expression in TRINA. RMCE efficiency was measured in clones *S1K8* and *S1K16* by counting the number of hygro^R colonies raised after transduction with IDLV-TRINA.Δcre alone or in combination with increasing amounts of IDLV-cre, and the results compared with transduction with IDLV-TRINA (Figure 3b, Supplementary Table 1). As expected, no hygro^R clones were isolated from cells transduced with IDLV-TRINA.Δcre or IDLV-cre alone (Figure 3b insert), demonstrating the requirement of both the enzyme and the target sequence for the RMCE to occur. The highest frequency of hygro^R colonies (4×10^{-3}) was obtained not at the highest IDLV-cre:IDLV-TRINA.Δcre ratio tested (10:2), but at a ratio of 2:10. This frequency is nearly equal to that obtained when cre is delivered in the TRINA backbone (Figure 3b insert, IDLV.TRINA). Interestingly, MOIs of IDLV-cre above 2 TU per cell decreased the frequency of hygro^R colonies by 5- to 10-fold. At higher MOIs, the lower hygro^R colony frequencies, small colony sizes and cell appearance (Supplementary Figure 3) are indicative of toxicity associated with high expression of cre enzyme.[7,14]

The IDLV-cre:IDLV-TRINA.Δcre ratio is also limited by the need to ensure efficient co-transduction, a fact that relies on the MOI employed of each virus, which is difficult to achieve at lower ratios (2:5 and 2:2).[8,15] The correspondence between lentivirus dose (IDLV-cre or TRINA) and the level of cre enzyme was tested by western blot at 48 h post-transduction. IDLV-cre at MOI = 2 achieved levels of cre expression comparable to those obtained

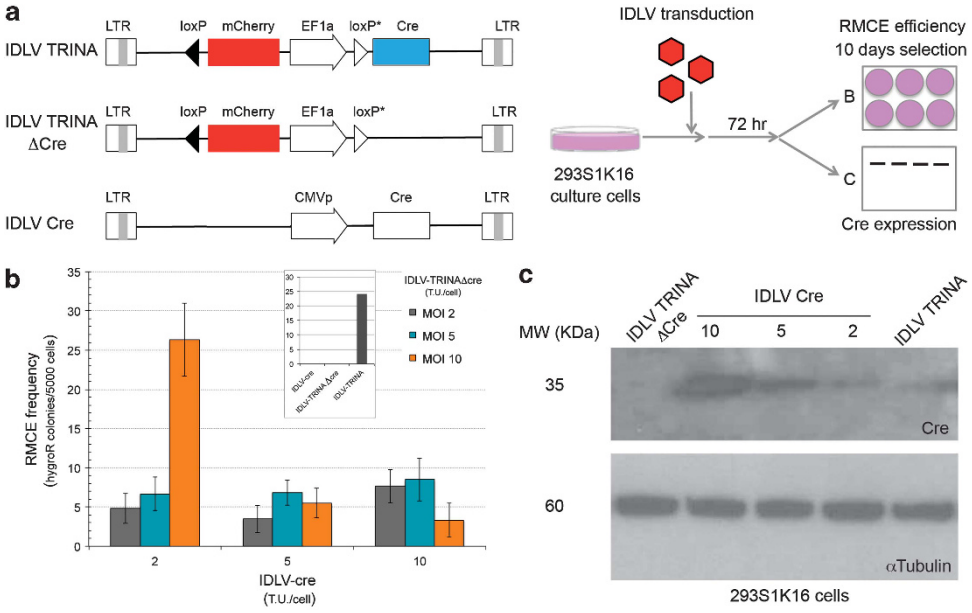

Figure 3. Optimization of RMCE in *AAVS1*-targeted cells by independent lentiviral delivery of *cre* recombinase. (**a**) IDLV TRINA and alternative lentiviral vectors generated to test the effect of independent delivery of TRINA and *cre* recombinase. The scheme shows the experimental strategy followed for the analysis in (**b**) and (**c**). (**b**) Quantitative representation of the number of hygro^R clones rescued from *S1K16* cultures co-transduced with IDLV-cre and IDLV-TRINA.Δcre at variable MOI(2 to 10 TU per cell). Average number of two independent experiments is represented according to the legend. For comparison is represented the frequency of hygro^R clones obtainedfor single transduction (MOI = 3 TU per cell) with IDLV-cre, IDLV-TRINA.Δcre or IDLV-TRINA (insert graph). Hygro^R clones obtained 14 days after transduction with the corresponding virus were counted under a magnifying glass on triplicate cultures. Average ± s.d. are represented. (**c**) Immunoblot showing expression of cre protein in cultures transduced with IDLV-cre at the indicated MOIs. Results obtained with IDLV-TRINA.Δcre and IDLV-TRINA (MOI = 3) are shown for comparison. α-Tubulin was detected as a loading control. MW, molecular weight marker.

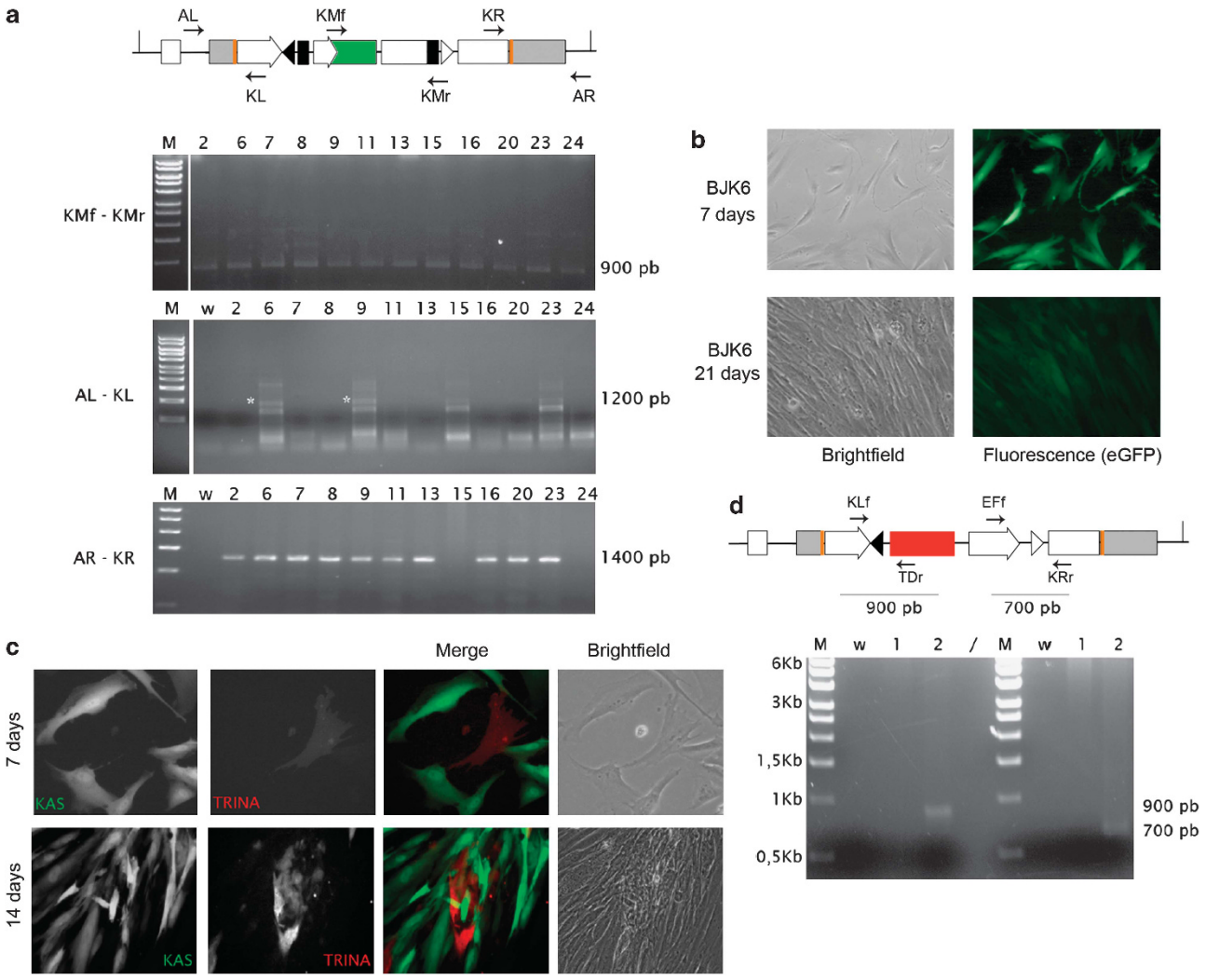

Figure 4. Integration of a modified TRINA cassette (TRINA2) into the KAS-edited *AAVS1* locus of primary human fibroblasts using an IDLV RMCE strategy. (**a**) PCR genotyping of KAS integration in twelve isolated G418^R (neo^R) BJ fibroblast clones. Amplification of genomic DNA from the indicated primer pairs follows the same strategy used in Figure 1b to determine correct left and right junction insertion in *AAVS1*. Primer positions in relation to the KAS-edited *AAVS1* locus are indicated in the figure. Numbers identify individual neo^R clones. PCR product sizes are indicated to the right. (*) Indicates the expected positive band (**b**) Genetic stability of the integrated KAS2.0 cassette. Clone BJ-K6 was maintained for 3 weeks (six population doublings) without selection and then cultured to confluence in the presence of G418. Both G418 resistance (*left*) and eGFP expression (*right*) were maintained throughout this period. (**c**) Transduction of BJ-K6 fibroblasts with IDLV TRINA2. The legitimate RMCE in the *AAVS1* locus generated the exchange and shift from green (eGFP/KAS) to mtdTomato (TRINA2), highlighting the promoter trapping strategy and the integration of the cassette. The left panel shows the generation under pressure conditions and the right one without them. (**d**) PCR genotyping of BJ-K6Tpool using the primer pairs (depicted on the diagram) confirms insertion in the *AAVS1* site by replacement of the KAS cassette in the mix population (lane 2), whereas this product is not present in the parental one (lane 1).

with TRINA at MOI = 5, reflecting the relative strength of the CMV promoter (Figure 3c). This result was confirmed by immunofluorescence with anti-cre antibody, which also revealed that cre subcellular localization was independent of the delivery method (Supplementary Figure 4).

KASTRINA 2.0 creates a safe harbour for human primary cells

To test the potential of KASTRINA 2.0, we assessed the ability of a tailored TRINA vector to target human primary BJ fibroblasts bearing the KAS2.0 cassette (Figure 4a). BJ cells were nucleofected with the *AAVS1*-targeting KAS2.0 cassette and electroporated with ZFN mRNA. Cells were subsequently cultured under selection conditions for 2 weeks to isolate G418^R BJ-KAS2.0 (BJ-K) clones. Although primary cells are very sensitive to G418, by titrating the responsiveness of BJ cells to the drug we obtained a low frequency (5×10^{-4}) of G418^R BJ-K clones showing a healthy appearance (Figure 4b). Growth of these clones was difficult and required culture on feeder layers of irradiated non-transduced BJ cells. Twelve G418^R, eGFP$^+$ clones were expanded, and integration of the KAS2.0 cassette in the *AAVS1* locus was confirmed by standard PCR in 2 out of 12 of the clones (Figure 4a). Further analysis of one of this clones, clone BJ-K6, by Southern blot confirmed correct integration of a single copy (data not shown). The karyotype of clone BJ-K6 is identical to that of parental BJ cells (Supplementary Figure 5), and eGFP expression and G418 resistance were maintained over more than six population doublings during 3 weeks' non-selective culture (Figure 4b). These results confirm that BJ-K6 is a stable clone of primary human cells bearing a site-specific-integrated single copy of the genomic landing pad KAS2.0 in the 'safe-harbour' *AAVS1* locus.

To generate BJ-TRINA cells from clone BJ-K6, we constructed a modified TRINA vector (TRINA2) containing a promoterless cassette encoding for the membrane tdTomato fluorescent

protein instead of mCherry, because of its traceability and its lower toxicity,[9,16] together with the standard TRINA readout (Figure 4d). On the basis of the results of the cre titration experiments (Figure 3b), we transduced BJ-K6 cells with LV-TRINA2 at an MOI ratio of 5 (TU/cell). Cells were cultured under selective pressure for 2 weeks (hygro^R), but no clones could be rescued after such a long period, although cells were co-cultured on feeder cells. Because of the difficulties in obtaining lasting BJ cells capable to continue clonally proliferative, we decided to ascertain just whether the recombination procedure was taking place in the transduce culture. To determine whether the RMCE reaction is feasible in these primary cells, we transduced BJ-K6 clone with TRINA2 and assessed if replacement of the KAS cassette and the expression of the TRINA2 from that genomic location were occurring. On the basis of aforementioned difficulties in culturing BJ cells, we pooled the cultures upon transduction with TRINA2. The pooled culture (BJ-K6Tpool) contained a high proportion of cells arose by proper recombination events as demonstrated by the solely presence of the membrane-associated red fluorescent marker, but not the green fluorescence expression driven by the KAS cassette at 7 and 14 days post-transduction (Figure 4c). In spite we cultured cells as before in the presence of hygromicin B, we unsuccessfully isolated TRINA2 BJ cells as we just managed to isolate fibroblasts with no clonogenic capacity (not shown). Finally we characterized genotypically the BJ-K6Tpool of cells to assess whether integration occurred at the AAVS1-pre-targeted locus. Following the strategy as before (Figure 2), genomic DNA was amplified by PCR using primers that anneal at the expected junction sites in both orientations (KLf + Tdr and EFf + KRr; Figure 4d). In cells of the BJ-K6Tpool sequence proper recombination was detected by the presence of 0.9-Kb and 0.7-Kb in length-specific bands (Figure 4d lanes number as 2) that were not present in the parental cell clone BJ-K6 (Figure 4d lanes number as 1).

Hence, we can conclude that the KAS/Trina system operates either in modifying the genome of a human primary fibroblast as well as in established cell line, suggesting a wide applicability of the resource.

DISCUSSION

We have investigated the ability of a genetic resource based on lentivirus and recombinases to promote human gene editing in a 'safe-harbour' locus. The resource is based on our previous work that demonstrated that randomly integrated copies of an acceptor cassette could be targeted at high frequency by an integration-deficient lentivirus delivering donor cassette and recombinase in a recombinase-mediated cassette exchange reaction. Here we have advanced the procedure to yield frequencies above 0.1% of cell clones with single site-specific insertion in a safe harbour, yielding resource for safe site-specific single insertion.

The KASTRINA 2.0 procedure takes advantage of the specificity of zinc-finger nucleases and evidence indicating that the human PPP1R2C locus is a safe genetic harbour. This locus, also known as *AAVS1*, is the preferred target for insertion of the adeno-associated virus (AAV) genome during non-productive infections, at frequencies estimated at around 1%. Several reports demonstrate the safety of this locus both in hES cells[2,17] and in hiPS cells, as well as in hHSC.[18] PPP1R2C is moreover transcribed in all primary human cells studied and in commonly used transformed cell lines, thus satisfying two of the accepted criteria for a transgene 'safe harbour'.[8] Zinc-finger nucleases can be used in a wide spectrum of cell types of human origin,[8] while the *AAVS1* locus supports stable levels of expression of both promoterless (reliant on the native *PPP1R12C* promoter) and promoter-containing inserted constructs. We reasoned that combination of ZFN-mediated transgenesis with the site-specific advantages of the RMCE reaction would provide an efficient, simple and inexpensive methodology for replacing cassettes in the *AAVS1* locus.

KASTRINA 2.0 is based on integrase-deficient lentivirus-mediated replacement of a gene at a predetermined locus, and is in principle widely applicable irrespective of locus and cell type or origin. As proof of principle, we tested the KASTRINA 2.0 system in the human HEK293A cell line and in human primary BJ fibroblasts. The procedure includes several key components devised to synergize with the aforementioned advantages of the ZFN/AAVS1 technology. First, the DNA partners to be exchanged contain genetic elements designed to ensure directionality and irreversibility. Second, the strength of the promoter-trap strategy ensures that only clones that underwent legitimate recombinase-mediated integration are isolated. Third, the limited expression of the recombinase avoids the genotoxicity associated with long-term expression. Our results show that transduction with TRINA lentivirus promotes RMCE at the targeted *AAVS1* locus, with no measurable insertion at other locations, thus yielding a single copy of the lentiviral construct in the desired orientation in a safe harbour. Expression is stable, and the frequency achieved prevents the selection of cells in which the lentivirus, despite being integration-deficient, is randomly inserted. Both the donor and the acceptor cassettes are provided with selectable markers, but a key feature of the donor cassette is that expression is limited to copies correctly inserted in a promoter-containing region: the absence of U3 promoters in the lentivirus and the reverse orientation of the cassette in relation to the viral LTRs preclude expression from the delivery vector. Although the TRINA lentivirus is integration deficient, random insertion of the retrotranscriptase intermediate is in principle possible and has been reported at frequencies ranging from 10^{-2} to 10^{-4} depending on the strategies followed and the cell systems used for measurement.[19] The use of a promoterless donor cassette ensures that transgene expression after such events is negligible. Our data show integration of the IDLV-driven cassette only in the AAVS1-integrated KAS2.0 genomic landing: highly sensitive qPCR detected no copies at other loci. As anticipated by the features of the KASTRINA 2.0 system, all hygro^R clones lost eGFP expression and gained expression of mCherry (red). Thus, expression of the selectable markers encoded in TRINA relies on the proper performance of the RMCE, and cells undergoing random integration are not rescued.[10]

The exchange of only a single copy of the *AAVS1*-integrated KAS2.0 cassette for the incoming cDNA generated upon lentiviral retrotranscription is both strength and a weakness. The sensitivity of the process is attested by the low multiplicity of infection used (3 TU per cell) and the low number of TRINA cDNA copies generated per retrotranscription cycle (one per molecule of RNA genome). The system is moreover highly efficient, as recombinase expression is limited to the number of cre mRNA copies transcribed, the half-life of the enzyme and the enzymatic properties related to kinetic of RMCE. However, these features inevitably limit the RMCE reaction and consequently the frequency of rescue of the recombination-driven phenotype. This weakness can be overcome, however, by increasing the MOI, and indeed we previously showed the effect of MOI on the efficiency of RMCE with randomly integrated copies of the acceptor cassette.[10]

KASTRINA 2.0 proved to be almost equally efficient in primary human BJ fibroblasts as in the human HEK293A cell line. Differences in the rescue of KAS-bearing BJ fibroblast clones are related to the specific difficulties of clonal growth of these cells, and we did not detect differences in RMCE-driven replacement or in the frequency of non-desired integration of the vector cDNA in random loci. We could not manage to obtain lasting proliferative BJ clones after a number of serial passages required for the resource to be ready, suggesting that probably cells were entering into a senescence programme.[20] Alternatively, cell types that either go quickly into senescence or with limited proliferative capabilities to proliferate might be reversibly manipulated

previously to ensure proliferation before stable clones were generated. In this regard, other authors used regularly immortalized BJ-derived fibroblasts[11,21,22] in order to use stable clones as *in vitro* models. Indeed we have demonstrated the *bona fide* activity of our approach in non-modified BJ fibroblast, thus representing an improvement in the generation and use of isogenic cells and its comparison. However, our results cannot be extrapolated to other cell types or cells of different origin. A particular concern is the well-established difference between the homologous recombination frequencies of human and mouse stem cells.[23] If the KASTRINA 2.0 resource works equally well in stem cells, it might help to overcome this difficulty and become an attractive approach for unravelling the processes controlling stemness maintenance and specific differentiation.

Lentiviral vectors currently yield higher reprogramming frequencies than other methods[24] like adenovirus, excisable transposons, minicircles to mention few of significance. Future development of this approach should include safety features such as capacity for excision of the reprogramming construct once the encoded activity is no longer required, and beyond this for successive replacement of the insert to yield different desired cell specifications. There is a clear need for this technology, in order to avoid undesired effects associated with random insertion of lentiviral inserts, as highlighted in a recent report describing a powerful strategy for tracing the contribution of haematopoietic stem cells.[25]

The KASTRINA 2.0 resource has the potential to be developed along these lines and will find application in transgenesis systems where outcomes need to avoid interference from biological effects driven by the insertion itself. In addition, the resource is built with lentiviral vectors, an issue that provides main advantages to target stem and reprogrammed human cells as such vectors are the current gold standards in delivering genes to them. The KASTRINA2.0 represents a platform that partially answered some key questions that are being claimed recently by several authors[18,23] the need to develop effective and safe have claimed the need of developing strategies

MATERIALS AND METHODS

Mammalian cell culture, transfection and electroporation

The human embryonic kidney cell line HEK293A (CRL-1573, ATCC) and the Bj fibroblasts (CRL-2522, ATCC) were cultured under standard conditions in DMEM (Lonza, Barcelona, Spain) supplemented with 1% Glutamax (Invitrogen, Barcelona, Spain), 10 mg ml^{-1} antibiotics (penicillin and streptomycin) and 10% foetal bovine serum (Gibco, Invitrogen). Cells were transfected with endotoxin-free DNA (Qiagen, Madrid, Spain) and transfections were carried out in six-well plates (BD Biosciences, Madrid, Spain) by the calcium-phosphate method unless otherwise stated.

The day before transfection, 500 000 cells were seeded in each well, next day and under a confluence of around 80–90% the transfection was carried out. A total amount of 6 mg of plasmid DNA was prepared, adding 30 μl of $CaCl_2$ 2.5 M in a final amount of 250 μl of milliQ water; in parallel 250 μl of a HBS 2 × solution was prepared and both tubes were mixed. Finally, the transfection mix was added in a drop wise manner to the cells, let them stayed in the incubator for 4–6 h and lastly the medium was changed. After 24–48 h the cells were processed properly for the subsequent analysis.

In the case of the Bj fibroblasts, it was necessary to transfect them by electroporation methods. The culture dishes were seeded as stated before to reach a culture density of around 90% the next day. The day of the experiment, the cells were trypsinized and washed with DPBS (Invitrogen, Life Technologies, Barcelona, Spain), they were counted and resuspended in R solution from the Neon transfection System (Invitrogen, Life Technologies). They were electroporated with 10 μl tips using between 50 000 and 100 000 cells using a programme with a single pulse of 1650 V and 20 ms. Next, cells were seeded in a 12-well plate. The next day the media was replaced with normal media and the cells were processed as designed for the next assays.

Lentivirus generation, titration and transduction

Viruses were produced by transient plasmid transfection into 293T cells by the calcium phosphate method, as previously described.[10] In brief, cells were seeded at 1.1×10^7 cells/dish in 15-cm dishes the day before transfection. Cells were transfected by the calcium phosphate method with 3 μg pRSV-Rev, 3.75 μg pMD.2G (VSV-G), 13 μg pMDLg/pRRE (or 13 μg pMDLg/pD64VRRE for production of non-integrative lentiviral vectors) and 35 μg of transfer plasmid (pTRINA, pTRINA dCre or pCre). The medium was collected after 48 h, cleared by low-speed centrifugation and filtered through 0.45-mm-pore-size PVDF filters. Viral stocks were concentrated by ultracentrifugation in SW28 Beckman rotor at 90 000 g (26 000 r.p.m.) for 2 h at 4 °C. Pellets containing lentivirus were air dried and resuspended O/N at 4 °C in 400–600 ml of media. Viral titres were calculated by FACS analysis on transduced HEK293T cells when vectors expressed fluorescent proteins (transduction units per ml) and particles quantitated by qPCR on supernatants (particles per ml). Values were around 10^7–10^8 TU per ml and in a 1:100 TU per particles ratio. Cells were transduced at different MOI in media containing polybrene (8 μg ml^{-1}, final concentration). Cells were incubated at 37 °C for 6 h. After that viral supernatant was replaced with cell medium and returned to 37 °C for 72 h to further analysis.

Clonal expansion

Once the cells were transfected or transduced, the process carried out was the same. In brief, 72 h after genetic modification, media was changed and it was replaced by the culture media in which selective agent was added. In the case of KAS cell line generation, it was G418 (Geneticin, Invitrogen, Life Technologies) at a 450 μg ml^{-1} and 150 μg ml^{-1}, depending on whether it was applied to HEK cells or to BJ fibroblasts, respectively. Regarding the RMCE experiments *per se*, after transduction cells were analysed by FACS or by starting hygro selection (100 μg ml^{-1} or 40 μg ml^{-1}, for HEK293 or BJ fibroblasts, respectively).

Standard PCR analysis and quantification of transgene copy numbers by qPCR analysis

Standard PCR was performed using the following conditions: template denaturation at 95 °C for 1 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 61 °C for 30 s, and extension at 72 °C for 1.5 min and a final extension of 5 min.

The primers used were for pKAS integration detection
AL Fw: 5′-GGCCCTGGCCATTGTCACTT-3′
KL Rv: 5′-CATAAGGTCATGTACTGGGC-3′
Product: 1200 bp
KMf Fw: 5′-ACCTTGATGCCGTTCTTC-3′
KMr Rv: 5′-GCGTGCAATCCATCTTG-3′
Product: 1400 bp
KR Fw: 5′-CTGTGTAGAAGTACTCGCCG-3′
AR Rv: 5′-GGAACGGGGCTCAGTCTG-3′
Product: 900 bp
and for RMCE product detection
KLf: 5′-TTCGGCTTCTGGCGTGTG-3′
CHr: 5′-GAACTCCTTGATGATGGCC-3′
Product: 300 bp
EFf: 5′-TTGACCTCAGCGTCGTAGTG-3′
KRr: 5′-GATGTTGGCGACCTCGTATT-3′
Product: 700 bp.

The genomic DNA was extracted following standard procedure and serial dilutions quantitated using NanoDrop ND 1000 Spectrophotometer (NanoDrop Technologies, Bonsai Tecnologies Group SL, Madrid, Spain). Real-time PCR was performed using the SYBR Green methodology. The number of integrated copies of pTRINA plasmid in the 293AKAS cells was measured as previously described,[10] with some modifications. PCR efficiency was examined with five dilutions of genomic DNA, and the specificity of individual gene primers was validated by the melting curve at the end of each PCR assay. Standard curves were obtained with diluted amounts of the pTRINA plasmid from 0.01 pg to 100 ng, corresponding 10^2–10^9 copies. Ct values obtained upon amplification of LTR and ChFP regions contained in pTRINA using the specific primers listed below were interpolated and the relative number of copies calculated. Cell equivalents were calculated using Ct values of the single copy gene Albumin on diluted samples of the genomic DNA. The qPCR reaction was carried out under the following conditions: template denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 15 s, annealing at 60 °C for 20 s and extension at 72 °C for 30 s. The primers used were:

qLTR Fw: 5′-TGTGTGCCCGTCTGTTGTGT-3′
qLTR Rv: 5′-GAGTCCTGCGTCGAGAGAGC-3′
Product: 95 bp
qChFP Fw: CCCCGTAATGCAGAAGAAGA-3′
qChFP Rv: 5′-TTGACCTCAGCGTCGTAGTG-3′
Product: 102 bp
qhAlb Fw: 5′-GCTGTCATCTCTTGTGGGCTGT-3′
qhAlb Rv: 5′-ACTCATGGGAGCTGCTGGTTC-3′
Product: 124 bp
qeGFP Fw: 5′-CCACATGAAGCAGCACGAC-3′
qeGFP Rv: 5′-GTGCTCAGGTAGTGGTTG-3′
Product: 290 bp.

Southern blot

Genomic DNA digested with *Eco*RV enzyme were loaded on a 0.8% agarose gel in 1 × TAE and separated by running at 70–100 V constant voltage until the dye front is 1–3 cm from the end of the gel. Control, quantitated plasmids containing the sequences to be detected were loaded to provide a reference for size and quantity (20 pg, 200 pg and 2 ng). The gel was photographed under 300 nm ultraviolet light. The gel was washed in distilled water and then equilibrated in 2 × SSC for 15 min at RT. The gel was blotted by capillarity on nylon membranes Hybond-N plus (GE Healthcare, Alcobendas, Spain) with NaOH 0.5 N overnight at RT. The membrane was fixed by irradiation in a Stratalinker (Stratagene, Basel, Switzerland) and used for further hybridization.

The probes were digested products purified from the plasmids pEGFP-N1 (Clontech, Saint-Germain-en-Laye, France) for eGFP and a PCR product of the AAVS1 locus obtained from genomic DNA for AAVS1 detection. Labelling was performed using 100 μCi per reaction of ^{32}P-alpha-dCTP using the commercial kit RediPrime II (GE Healthcare) following the manufacturer instructions. Non-incorporated nucleotides were removed using a Micro Bio-Spin P-30 (BioRad, Madrid, Spain) column according to the manufacturer instructions and quantitated by scintillation counting. Filters were wetted in 2 × SSC and prehybridized and hybridized in 25 cm tubes with 10–20 ml of PerfectHyb Plus (Sigma Aldrich, Madrid, Spain). Hybridization solution contained 1–2 × 10^7 c.p.m. ml^{-1} and was left to proceed overnight at 66 °C. Filters when then washed with successive 20 min washes at 65 °C in 0.1 × SSC 0.5% SDS. Wet membranes were exposed to PhosphorImager screen and developed using a STORM scanner.

Western blot

Proteins were extracted following standard procedures in the presence of Complete Protease Inhibitor Cocktail Tablets (Roche Applied Science, Madrid, Spain). Western blots were carried out by standard methods on proteins transferred to PVDF using TransFi (Invitrogen). Membranes were probed for cre with monoclonal anti-Cre antibody (Novagen, EMD, Merck – España, Madrid, Spain) (1/2000) or anti-alpha tubulin (AbCam, Cambridge, UK) (1/5000) in PBS/0.1% Tween-20 (PBSTween). Secondary antibodies were HRP-conjugated goat antimouse IgG (Santa Cruz Biotechnology, Heidelberg, Germany), and blots were developed with ECL (GE Healthcare).

Cytometry analyses

Flow cytometry analysis was performed after 72 h post-transduction. Cells were trypsinized and collected, washed twice in culture-grade 1 × PBS and analyzed in a FACS DIVA (Becton Dickinson, Madrid, Spain) sorter with an appropriate laser for CherryFP excitation. In every case, 10 000 events were counted in triplicate.

Immunofluorescence

For localization of antigens by double label indirect immunofluorescence (IF), cells were seeded onto glass coverslips coated with poly-L-lysine (Sigma), transfected or transduced. 24–48 h later, cells were washed twice with PBS and fixed (5 min) in cold 4% paraformaldehyde, permeabilized with 0.5% TritonX-100, PBS and blocked with 4% normal goat serum. Thereafter, samples were incubated (45 min, 37 °C) with anti-Cre (1/200 dilution, Novagen EMD Millipore), diluted in PBS supplemented with 5% goat serum.

Secondary antibody Alexa-488-conjugated (Molecular Probes, Invitrogen, Life Technologies) was used at 1/500 at RT during 1 h. Finally, samples were dyed with DAPI (Invitrogen, Life Technologies) to stain DNA, air dried and mounted in Gelatin (Sigma Aldrich). Samples were examined with a Leica SPE confocal laser scanning microscope (Leica Microsystems Holdings, Wetzlar, Germany) fitted with two lasers giving excitation at 488 nm (secondary antibody detection), 543 nm (mCherry or tdTOMATO visualization) and 405 nm (DAPI, nuclear staining). Data were collected sequentially at a resolution of 1024 × 1024 pixels from 0.5 to 1.0 μm thick optical slices.

Karyotyping

The cell line was incubated at 37 °C in culture flasks in an atmosphere of 5% CO_2 in air. Metaphase cells were prepared by standard cytogenetic methods. Mitotic arrest with Colcemid (0.1 μg ml^{-1}, 1.5 h, 37 °C; GIBCO, Strachclyde, UK) was followed by hypotonic treatment (75 mm KCl, 15 min, 37 °C) and fixed with methanol/acetic acid (3:1) before spreading onto slides.

Fluorescence *in situ* hybridization analysis

Two sets of probes were used to localize virus integration sites. To detect the adeno-associated virus integration site, we used bacterial artificial chromosomes (BAC) that map at 19p13.3 (RP11-787B21) and 4q13.3 (RP11-49L9) cytobands as controls. TRINA integration site was detected using the DNA from the vector plasmid. The BACs were obtained from Human BAC Clone Library RPCI-11 (Children's Hospital Oakland Research Institute, Oakland, CA, USA). All probes were labelled directly by nick translation according to the manufacturer's specifications (Vysis, Dowers Grove, IL, USA) with SpectrumGreen– or Spectrum Orange–dUTPs (Vysis). The probes were blocked with Cot-1 DNA (Vysis) to suppress repetitive sequences. Metaphase spreads, obtained from the patient's cell line and a normal control, were hybridized overnight at 37 °C with labelled probes. After post-hybridization washes, the chromosomes were counterstained with DAPI in anti-fade solution. Cell images were captured using a cooled charge-coupled device (CCD) camera (Photometrics SenSys camera) connected to a computer running the Chromofluor image analysis system (Cytovision, Applied Imaging Ltd, Newcastle, UK).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grant of the Ministry of Economy and Compititiveness [INNPACTO IPT-010000-2010-40] to JCR; and National Centre for Cardiovascular Research (CNIC) institutional funding from the Fundacion Pro-CNIC to the Viral Vector Technical Unit.

DISCLAIMER

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1 Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, Clerici M *et al.* T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* 1995; **270**: 475–480.
2 Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 2009; **27**: 851–857.
3 Zou J, Maeder ML, Mali P, Pruett-Miller SM, Thibodeau-Beganny S, Chou B-K *et al.* Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 2009; **5**: 97–110.
4 Chapdelaine P, Pichavant C, Rousseau J, P Acirc Ques F, Tremblay JP. Meganucleases can restore the reading frame of a mutated dystrophin. *Gene Ther* 2010; **17**: 846–858.
5 Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF *et al.* A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 2011; **29**: 143–148.
6 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE *et al.* RNA-guided human genome engineering via Cas9. *Science* 2013; **339**: 823–826.
7 Turan S, Galla M, Ernst E, Qiao J, Voelkel C, Schiedlmeier B *et al.* Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges. *J Mol Biol* 2011; **407**: 193–221.

8 DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res* 2010; **20**: 1133–1142.
9 Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer* 2012; **12**: 51–58.
10 Torres R, García A, Payá M, Ramirez JC. Non-integrative lentivirus drives high-frequency cre-mediated cassette exchange in human cells. *PLoS One* 2011; **6**: e19794.
11 Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee Y-L, Kim KA *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* 2007; **25**: 1298–1306.
12 Kolb AF. Selection-marker-free modification of the murine beta-casein gene using a lox2272 [correction of lox2722] site. *Anal Biochem* 2001; **290**: 260–271.
13 Qiao J, Oumard A, Wegloehner W, Bode J. Novel tag-and-exchange (RMCE) strategies generate master cell clones with predictable and stable transgene expression properties. *J Mol Biol* 2009; **390**: 579–594.
14 Silver DP, Livingston DM. Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol Cell* 2001; **8**: 233–243.
15 Fehse B, Kustikova OS, Bubenheim M, Baum C. Pois(s)on-it's a question of dose. *Gene Ther* 2004; **11**: 879–881.
16 Muzumdar MD, Tasic B, Miyamichi K, Li L, Luo L. A global double-fluorescent Cre reporter mouse. *genesis* 2007; **45**: 593–605.
17 Smith JR, Maguire S, Davis LA, Alexander M, Yang F, Chandran S *et al.* Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 2008; **26**: 496–504.
18 van Rensburg R, Beyer I, Yao XY, Wang H, Denisenko O, Li Z-Y *et al.* Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. *Gene Ther* 2013; **20**: 201–214.
19 Wanisch K, Yáñez-Muñoz RJ. Integration-deficient lentiviral vectors: a slow coming of age. *Mol Ther* 2009; **17**: 1316–1332.
20 Brown JP, Wei W, Sedivy JM. Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science* 1997; **277**: 831–834.
21 Marión RM, Strati K, Li H, Murga M, Blanco R, Ortega S *et al.* A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 2009; **460**: 1149–1153.
22 Li H, Collado M, Villasante A, Strati K, Ortega S, Cañamero M *et al.* The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* 2009; **460**: 1136–1139.
23 Hockemeyer D, Jaenisch R. Gene targeting in human pluripotent cells. *Cold Spring Harbor Symposia on Quantitative Biology* 2011; **75**: 201–209.
24 González F, Boué S, Belmonte J. Methods for making induced pluripotent stem cells: reprogramming a la carte. *Nat Rev Genet* 2011; **12**: 231–242.
25 Neff NF, Quake SR, Lu R, Weissman IL. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat Biotechnol* 2011; **29**: 928–933.