



Comparison of the editing patterns and editing efficiencies of TALEN and CRISPR-Cas9 when targeting the human CCR5 gene

Arildo Nerys-Junior¹, Luciene P. Braga-Dias², Paula Pezzuto¹, Vinícius Cotta-de-Almeida³ and Amilcar Tanuri¹

¹Laboratório de Virologia Molecular, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil.

²Laboratório do Serviço de Biotecnologia e Desenvolvimento Animal, Instituto de Ciência e Tecnologia em Biomodelos, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil.

³Laboratório de Pesquisas Sobre o Timo, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil.

Abstract

The human C-C chemokine receptor type-5 (CCR5) is the major transmembrane co-receptor that mediates HIV-1 entry into target CD4+ cells. Gene therapy to knock-out the CCR5 gene has shown encouraging results in providing a functional cure for HIV-1 infection. In gene therapy strategies, the initial region of the CCR5 gene is a hotspot for producing functional gene knock-out. Such target gene editing can be done using programmable endonucleases such as transcription activator-like effector nucleases (TALEN) or clustered regularly interspaced short palindromic repeats (CRISPR-Cas9). These two gene editing approaches are the most modern and effective tools for precise gene modification. However, little is known of potential differences in the efficiencies of TALEN and CRISPR-Cas9 for editing the beginning of the CCR5 gene. To examine which of these two methods is best for gene therapy, we compared the patterns and amount of editing at the beginning of the CCR5 gene using TALEN and CRISPR-Cas9 followed by DNA sequencing. This comparison revealed that CRISPR-Cas9 mediated the sorting of cells that contained 4.8 times more gene editing than TALEN+ transfected cells.

Keywords: CCR5, CRISPR-Cas9, efficiency, gene editing, TALEN.

Received: March 15, 2017; Accepted: June 7, 2017.

Introduction

HIV-1 entry into target CD4+ cells requires the C-C chemokine receptor type 5 (CCR5) that acts as a co-receptor for the V3 loop of the gp120 viral adhesion protein (Hill *et al.*, 1997). In addition, mutant strains raised in advanced stages of the infection can use the C-X-C chemokine receptor type 4 (CXCR4) as a co-receptor to mediate viral entry. However, CXCR4 has not been defined as a preferential anti-HIV target since it is also the transmembrane protein that guides CD4+ cells to inflammatory sites and is thus regarded as highly relevant for immune activity (McGowan and Shah, 2010). On the other hand, a CCR5 natural 32-bp deletion (defined as the CCR5 Δ 32 allele) is an effective restriction condition against HIV-1 infection. As this mutant allele produces a truncated protein that is not expressed on the cell surface, individuals homozygous for CCR5 Δ 32

cannot be infected by the usual CCR5-tropic-only strains of HIV-1 (Grotto and Pardini, 2006). This mutation thus confers resistance to HIV-1 infection in the homozygous state and partial resistance to infection with a slower rate of progression to AIDS in the heterozygous state (Reiche *et al.*, 2008; Silva-Carvalho *et al.*, 2016). In addition, this mutant allele has been associated with increased susceptibility to systemic lupus erythematosus (Baltus *et al.*, 2016) and juvenile idiopathic arthritis (Scheibel *et al.*, 2008), as well as decreased susceptibility to pre-eclampsia (Telini *et al.*, 2014), osteomyelitis (Souza *et al.*, 2015) and rheumatoid arthritis (Pokorny *et al.*, 2005).

In a landmark heterologous transplant in 2009, an HIV-1-positive patient received a bone marrow transplant from a compatible HIV-1-negative CCR5 Δ 32 homozygote donor as treatment for his acute myeloid leukemia (Hütter *et al.*, 2009). After transplantation, antiretroviral therapy was discontinued, resulting in a rapid decrease in viral load followed by long-term viral absence. This was considered the first functional cure for HIV-1 infection, with the patient remaining functionally healed.

Send correspondence to Amilcar Tanuri. Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro (UFRJ), Avenida Carlos Chagas Filho, 373, Centro de Ciências da Saúde (CCS), Bloco A, Sala 121, Cidade Universitária, 21941-902, Rio de Janeiro, RJ, Brazil. E-mail: atanuri1@gmail.com

Despite the very encouraging results of heterologous transplantation, the large-scale application of this approach is not a trivial matter, mainly because of the low frequency of the CCR5 Δ 32 allele in the general population (10-20% in northern and northeastern Europe, which have the highest frequencies of CCR5 Δ 32 in the world) and the low frequency of compatible individuals (Gonzales *et al.*, 2001; Hütter *et al.*, 2009; Silva-Carvalho *et al.*, 2016).

To overcome these problems, gene therapy strategies in autologous transplantation have been proposed to treat HIV-1 infection (Hütter *et al.*, 2009; Cannon and June, 2011). In addition, the beginning of the CCR5 gene, defined as the 3' nucleotides immediately downstream from the ATG start codon, is a key region for planned targeting since it mediates properly CCR5 gene knock-out. Gene therapy for HIV-1 infection initially requires the identification and choice of a suitable genetic tool for editing the target gene. Currently, the two most modern and effective programmable endonucleases that mediate precise gene targeting are the transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) (Nemudryi *et al.*, 2014).

Chromosomal position is directly related to chromatin structure, and transcriptional rate (Narlikar *et al.*, 2002), as well as promoter and genetic position are directly related to epigenetic modifications such as DNA methylation (Moarii *et al.*, 2015). TALEN binds to methylated cytosines (Valton *et al.*, 2012a,b), whereas CRISPR-Cas9 does not (Vojta *et al.*, 2016). However, additional studies are needed to determine the patterns of sensitivity for TALEN and CRISPR-Cas9 in chromatin, epigenetics, histones, nuclear localization and different transcriptional landscapes. To determine the relevance of these conditions in targeting the genetic site prior to gene editing it is very desirable to understand the patterns and efficiency of gene editing using TALEN and CRISPR-Cas9 for each genetic region of interest in the target cell type (Arvey *et al.*, 2012; Lee *et al.*, 2012; Weber *et al.*, 2015).

TALEN and CRISPR-Cas9 are endonucleases that can be programmed to target DNA cleavage. TALEN recognizes thymine on its conserved n-terminal portion. This is the position "0" (zero) of the genomic target. Subsequent target genomic recognition process is performed by the assembled sequence of specific repetitive variable diresidue (RVD), where: "NI" RVD type recognizes adenine, "HD" RVD type recognizes cytosine, "NN" RVD type recognizes guanine and "NG" RVD type recognizes thymine; other forms allow additional nucleotide recognition with lower efficiency and double-strand breaks can be produced by dimerization of the *FokI* catalytic site from both TALEN arms (Cermak *et al.*, 2011a). In contrast, CRISPR recognizes the target through a short RNA sequence known as single guided RNA (sgRNA) and double-strand breaks are produced by the Cas9 protein from adaptive bacterial im-

munity Type II of *Streptococcus* spp., e.g., *Streptococcus pyogenes*, and *Archaeae* (Gasiunas *et al.*, 2012; Jinek *et al.*, 2012; Wiedenheft *et al.*, 2012; Riordan *et al.*, 2015).

The CRISPR-Cas9 target requires a protospacer adjacent motif (PAM) sequence (5'-NGG-3') and sgRNA has an anchorage sequence that anchors it to Cas9. This anchorage sequence is followed by the recognition sequence (without a PAM complementary sequence) that is complementary to the target. A double-strand break occurs in the base pairs after the PAM sequence (three base pairs after the beginning of the recognition sequence) where the HNH domain of Cas9 cleaves the strand that is paired with sgRNA and the RuvC domain cleaves the other DNA strand at the same position (Jinek *et al.*, 2012; Wiedenheft *et al.*, 2012; Riordan *et al.*, 2015). Whereas TALENs are assembled by a sequence of cloning and subcloning steps (Cermak *et al.*, 2011b), CRISPR-Cas9 is easily assembled in a single cloning step (Ran *et al.*, 2013a). This difference makes TALEN more time consuming to assemble compared to CRISPR-Cas9.

TALEN requires a reporter plasmid to indicate its action within the cell (Kim *et al.*, 2011). This plasmid has an operon modulated by a CMV promoter that regulates the expression of a red fluorescent protein (RFP) which, in turn, indicates successful transfection of the plasmid. This RFP sequence is immediately followed (without a stop codon) by the TALEN recognition sequence and an out-of-frame GFP sequence. Whenever TALEN is expressed and cleaves the plasmidial target, the non-homologous end-joining cellular repair mechanism inserts an InDel mutation that reestablishes the GFP frame in some cases (Kim *et al.*, 2011). For this to occur, the TALEN target must not contain a stop codon in the frame containing GFP. In some cases, TALEN cleaves the target but does not reestablish the GFP open reading frame. Conversely, several CRISPR-Cas9-coding plasmids already have a GFP reporter gene after the Cas9 gene where it is separated from the Cas9 protein by a T2A self-cleaving peptide (Ran *et al.*, 2013a). TALEN transfections require the co-transfection of three plasmids at the same time (right arm plasmid, left arm plasmid and reporter plasmid), but the reporter plasmid is optional in some cases (Kim *et al.*, 2011). On the other hand, CRISPR-Cas9 transfections require the transfection of a single plasmid that contains not only all the CRISPR-Cas9 molecular requirements to cleave the desired target, but also a puromycin resistance gene for drug-based cell selection, or a GFP reporter system that indicates Cas9 production by itself (Ran *et al.*, 2013a).

Cell sorting can be done in both TALEN and CRISPR transfection experiments, although in TALEN transfections, the use of a reporter plasmid to sort RFP⁺/GFP⁺ cells (produced by TALEN when acting in the nucleus) is required, (Kim *et al.*, 2011). In contrast, CRISPR-Cas9 transfections may allow the sorting of GFP⁺ cells by themselves, where GFP⁺ cells are produced by Cas9-T2A-GFP Open

Reading Frame (ORF) before Cas9 anchoring in the sgRNA produced by the same plasmid (Ran *et al.*, 2013a).

TALEN has low toxicity and is very efficient, very specific and rarely shows off-target effects (Mussolino *et al.*, 2011). In contrast, CRISPR-Cas9, despite being more efficient than TALEN, may generate higher off-target effects (Shen *et al.*, 2014; Tsai and Joung, 2016). These off-target effects of CRISPR-Cas9 can be easily and dramatically reduced by using truncated single-guided RNAs (sgRNAs < 20 nucleotides in length) (Fu *et al.*, 2014).

CRISPR-Cas 9 is highly efficient at inducing mutagenesis in certain human somatic cells and this characteristic can be used to mediate hematopoietic cell-based therapy (Mandal *et al.*, 2014). However, the differences in efficiency between CRISPR-Cas9 and TALEN should be tested for each target locus in each target cell type to assess the usefulness of these tools for each objective. This applies to cell cultures such as HEK293T cells and genes, such as the human CCR5 gene, that are targets for gene therapy.

The therapeutic potential of CRISPR-Cas9 has been known and explored for some time (Gaj *et al.*, 2013). Although CRISPR-Cas9 is better than TALEN for gene editing in HEK293FT cells (He *et al.*, 2016), it is unclear to what extent CRISPR-Cas9 is better than TALEN at editing the human CCR5 gene, including in HEK293T cells.

Gene therapy for the human CCR5 gene is ever closer to becoming a reality, and CRISPR-Cas9 has a prominent role in this process. The ablation of the CCR5 gene in NOD/Prkdc^{scid}/IL-2R γ ^{null} mice was found to confer long-term resistance to HIV-1 infection *in vivo* (Xu *et al.*, 2017). This finding has renewed interest in gene-therapy-based alternatives for curing HIV-1 infections using a hematopoietic stem cell procedure. The *in vivo* excision of HIV-1 provirus with a multiplex CRISPR-Cas9 system has been used in other animal models, such as transgenic mice, and may provide an alternative approach for gene therapy through CRISPR-Cas9 (Yin *et al.*, 2017).

Despite various advances in the use of gene therapy to treat HIV-1 infection, and the fact that TALEN and CRISPR-Cas9 have been used in most studies, including targeting of the CCR5 gene, nothing is known about possible variations in the patterns and efficiencies of TALEN and CRISPR-Cas9 in editing the beginning of the CCR5 gene. Such knowledge is important for corroborating the choice of a given tool in basic research and gene therapy targeting of the CCR5 gene.

To address our poor understanding of this matter, and to determine whether TALEN or CRISPR-Cas9 is the better of these two approaches for editing the beginning of the CCR5 gene, we compared the patterns and quantity of editions in the CCR5 gene using TALEN and CRISPR-Cas9. DNA sequencing was used to show that a version of CRISPR-Cas9 that carries a GFP reporter gene mediates the sorting of cells that contain five times more gene editing than the sorted TALEN⁺ transfected cells.

Material and Methods

TALEN assembly and reporter plasmid construction

To target the beginning of the CCR5 gene, the TALEN strategy proposed by Miller *et al.* (2011) was assembled through the Golden Gate TALEN assembly kit (AddGene, Cambridge, MA, USA) (Cermak *et al.*, 2011b). For gene editing, both right and left assembled TALEN plasmids were transfected together, with the recognition site starting 154 bp downstream from the CCR5 start codon (ATG) relative to the first standard thymine (T) of the recognition site. To demonstrate TALEN activity within cells (Kim *et al.*, 2011), a reporter plasmid containing the TALEN target was constructed using the pRGS vector (red-green system plasmid), referred to as the pRGS-CR reporter plasmid (pRGS to CCR5 Miller TALEN target) (Figure 1).

CRISPR-Cas9 assembly

The beginning of the CCR5 gene was analyzed using the online software CRISPR design tool from the Massachusetts Institute of Technology (available at <http://crispr.mit.edu/>) (Hsu *et al.*, 2013). The nearest possible CRISPR target site from the CCR5 start codon was chosen for testing (Figure 1). The CRISPR-Cas9 plasmid to target the first possible CRISPR site in the CCR5 gene was assembled in the pX458 vector (also referred to as pSpCas9(BB)-2A-GFP; Addgene plasmid #48138) using the standard assembly protocol (Ran *et al.*, 2013a). The necessary annealing primers that produce the insertion sequence encoding the target-complementary sequence of the sgRNA in the assembled CRISPR plasmid were designed manually (Figure 2). The assembled CRISPR-Cas9 mediates a no-overhang (blunt) double strand break between the 24th and 25th nucleotides downstream from the CCR5 start codon (Figure 1).

Cellular transfections

The plasmids encoding the TALEN right and left arms were simultaneously co-transfected with the pRGS-CR reporter plasmid at a ratio of 1:1:2 in a total of 2 μ g of DNA and the CRISPR-Cas9-encoding plasmid was transfected in a total of 3 μ g of DNA. For DNA transfection in both cases, Lipofectamine 2000TM was used according to the manufacturer's instructions and HEK293T cells were co-transfected at confluence (4x10⁵ cells/well) in six-well plates (BD Falcon, Corning).

Flow cytometry (FC) and fluorescence activated cell sorting (FACS)

Flow cytometry was done using a BD Accuri C6 flow cytometer 24, 48 and 72 h after transfection to determine the highest proportion of GFP⁺ transfected HEK293T cells in the CRISPR-Cas9 transfections. This same approach was previously used to show that the in-

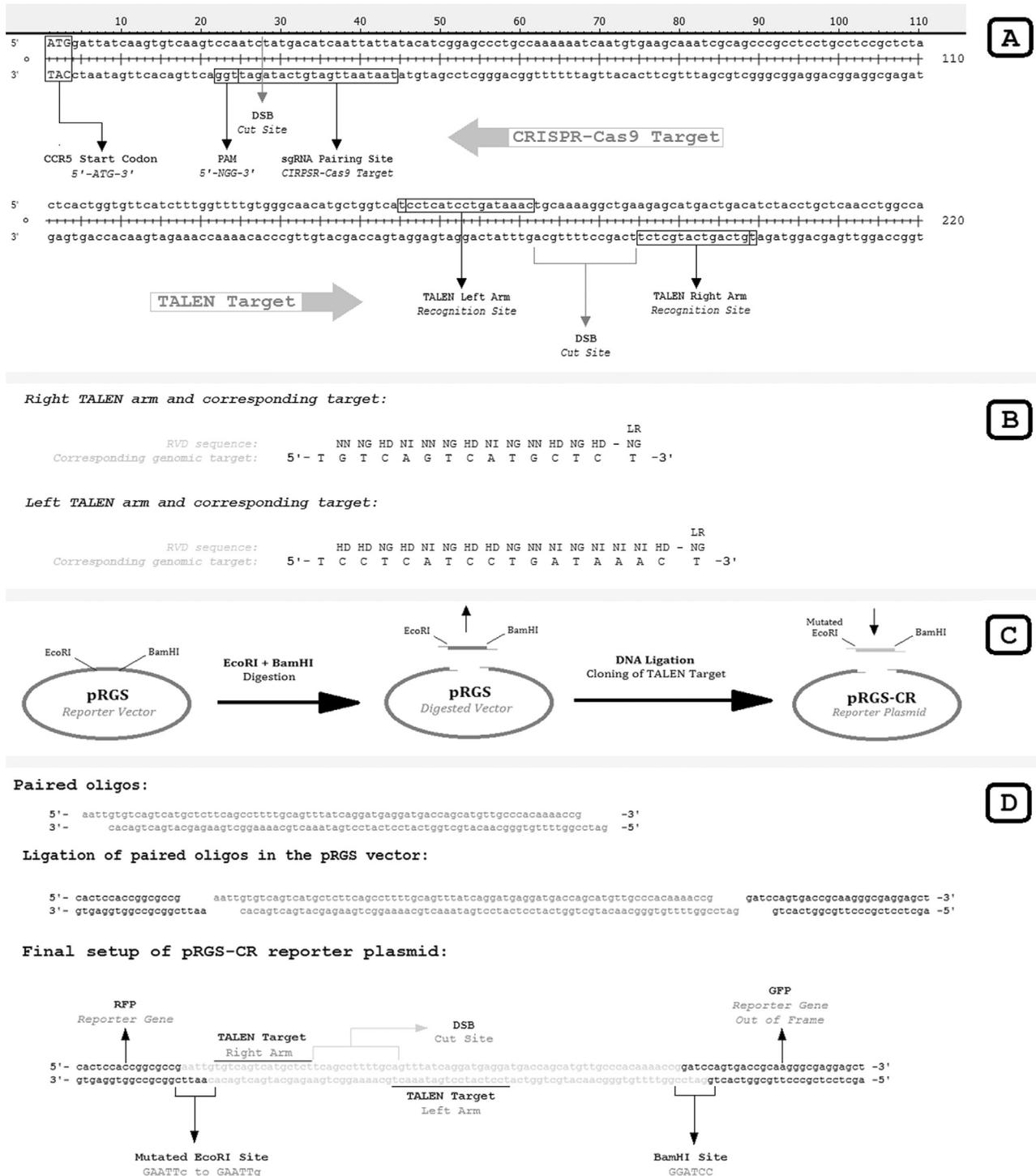


Figure 1 - CRISPR-Cas9 and TALEN recognition sites with assembly description and pRGS-CR reporter plasmid construction. **A.** The beginning of the CCR5 gene with TALEN and CRISPR-Cas9 recognition sites. Whereas the assembled CRISPR-Cas9 mediates a blunt double-strand break between the 24th and 25th nucleotides from the CCR5 start codon (ATG), TALEN mediates an overhanging double-strand break (DSB) between the 168th and 181th nucleotides from the CCR5 start codon. **B.** Repetitive variable diresidue (RVD) sequence of right and left TALEN arms with corresponding genomic recognition site. LR indicates the last repeat RVD. **C.** Steps of pRGS-CR reporter plasmid assembly. The pRGS vector (Plasmidial Red and Green system) is co-digested with *EcoRI* and *BamHI* to release a short DNA strand and expose *EcoRI* and *BamHI* overhangs. Annealed oligos are inserted in the digested pRGS vector through corresponding overhangs. The annealed oligos contain a mutated *EcoRI* site 5'-GAATTC-3' (right) to 5'-GAATTg-3' (mutated) in the *EcoRI* overhang to allow digestion before transformation, thereby avoiding the transformation of unwanted assembled constructs. **D.** Details of the assembled pRGS-CR reporter plasmid. Annealed oligos with corresponding overhangs are indicated, as are the recognition sites for each TALEN arm, the *BamHI* splicing site and the mutated *EcoRI* sequence of the annealed oligos.

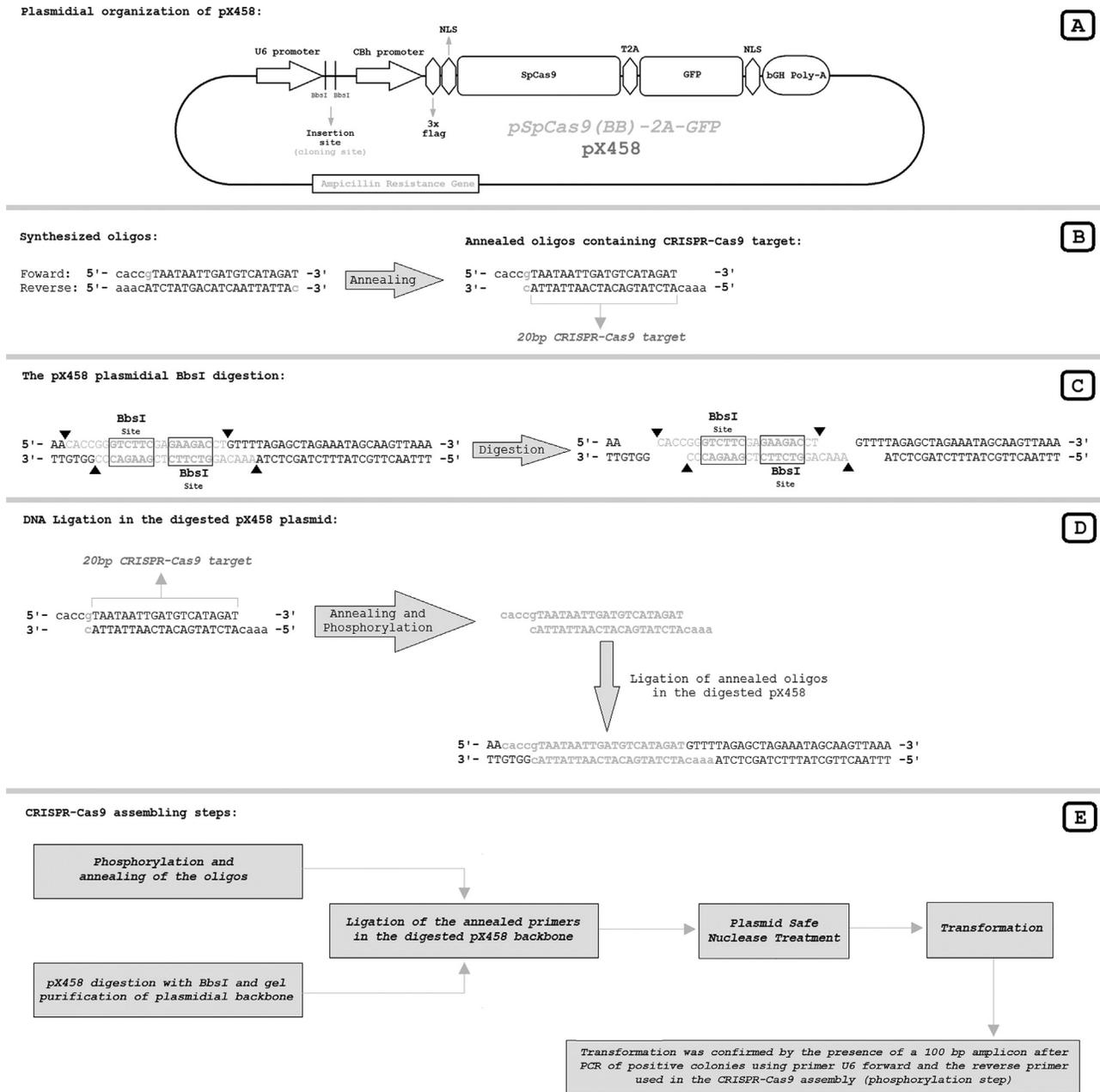


Figure 2 - Sequential steps for assembling the CRISPR-Cas9 plasmid using the pX458 vector. **A.** The pX458 vector (also known as pSpCas9(BB)-2A-GFP; Addgene Plasmid #48138) map. The U6 promoter encodes short RNAs and regulates the synthesis of sgRNA. The CBh promoter modulates Cas9 expression that is followed by a T2A self-cleaving peptide that releases a GFP reporter protein. The bovine growth hormone polyadenylation signal (bGH-PolyA) is used after the GFP gene. **B.** Oligos used to assemble the CRISPR-Cas9. Annealed oligos have overhangs complementary to the *BbsI* digested pX458 vector. The 20-bp genomic CRISPR-Cas9 target that encodes the genomic recognition part of sgRNA is indicated. The gray highlighted guanine (g) that is paired with the gray highlighted cytosine (c) are not part of the recognition site of sgRNA, but are requirements of the U6 promoter for proper sgRNA production. **C.** Details of *BbsI* digestion of the pX458 vector. Both *BbsI* sites are released in the digestion since they are located within the DNA sequence that is replaced by the annealed oligos in the ligation. **D.** Detailed ligation of annealed oligos in the *BbsI*-digested pX458 vector. **E.** General steps for CRISPR-Cas9 assembly.

interval with the highest proportion of RFP⁺/GFP⁺ cells in TALEN+pRGS-CR transfections was 72 h after transfection (Kim *et al.*, 2011; Nerys-Junior *et al.*, 2014) and this interval was used in cell sorting of TALEN+pRGS-CR transfections. For CRISPR-Cas9-transfected HEK293T cells, the best interval (highest

proportion of GFP⁺ cells) was 48 h post-transfection. Cell sorting was done with a MoFlo™ flow cytometer (Dako Cytomation, Beckman Coulter, Brea, CA, USA) to isolate HEK293T cells with the desired phenotype (RFP⁺/GFP⁺ for TALEN+pRGS-CR transfections and GFP⁺ for CRISPR-Cas9 transfections).

Genomic extraction and DNA cloning

HEK293T cells were grown for five days after flow cytometric analysis and cell sorting. Genomic DNA was extracted from HEK293T cells using a QIAamp DNA blood mini kit (Qiagen, Hilden, Germany) and the CCR5 gene was amplified using the primers 5'-TGGAGGGCAACTAAATACATTCTAGG-3' (forward) and 5'-CAGGTACCTATCGATTGTCAGGAGGA-3' (reverse) with the following cycle conditions: 95 °C for 5 min in the pre-PCR phase, followed by 38 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min, with a final extension (post-PCR phase) of 72 °C for 7 min. The 445 bp amplicon, which included 200 bp up and downstream from the TALEN target site, and also 100 bp upstream and 300 bp downstream from the CRISPR-Cas9 site, was cloned using a pGEM[®]-T Easy Vector (Promega, Madison, WI, USA) and transformed in *E. coli* JM109 (Promega) competent cells. White positive colonies were screened using a Luria broth (LB) medium plate containing ampicillin (50 µg/mL), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and isopropyl β-D-1-thiogalactopyranoside (IPTG). Subsequent comparison allowed distinction between sorted and unsorted white colonies.

DNA sequencing

White colonies were sequenced using an ABI BigDye Terminator sequencing kit (Applied Biosystems, Carlsbad, CA, US) on an Applied Biosystems 3130 Genetic Analyzer. Genomic DNA extracted from non-transfected HEK293T cells was used as a wild-type reference that was validated based on the wild-type CCR5 genetic sequence from GeneBank (National Center for Biotechnology Information - NCBI, Bethesda MD, USA). All sequences were aligned using SeqMan software v8.1.2 (DNASar, Madison, WI, USA).

Results

Fluorescence measurements by flow cytometry showed that GFP⁺ cells were most abundant in CRISPR-Cas9-transfected cells 48 h after transfection (Figure 3). Whereas Miller's TALEN transfections resulted in ~10% of RFP⁺/GFP⁺ cells (Nerys-Junior *et al.*, 2014), CRISPR-Cas9 transfections resulted in 57.2% of GFP⁺ cells (Figure 3). The transfections were repeated 10 times in the same conditions and in all cases the proportion of gated cells showed no more than 2% variability.

Although CRISPR-Cas9 transfections were ~47% more efficient than TALEN transfections in generating GFP⁺ cells, as indicated by flow cytometric analysis, locus modifications still need to be evaluated by sequencing to show direct nucleotide In-Del alterations. For this, RFP⁺/GFP⁺ cells were sorted 72 h after TALEN transfections and GFP⁺ cells were sorted 48 h after CRISPR-Cas9 transfections. The sorted cells were cultured for five

days in one well each of a 6-well plate, at which point they reached 80% confluence. The genomic DNA of both groups of cells was subsequently extracted for PCR. The resulting 445 bp amplicon containing TALEN and CRISPR-Cas9 sites was cloned into a pGEM[®]-T Easy Vector (Promega) and transformed in *E. coli* JM109 (Promega) competent cells that were then plated on an ampicillin/X-Gal/IPTG plate. A portion of cells was separated before cell sorting for subsequent extraction of genomic DNA and the amplicon containing TALEN and CRISPR-Cas9 target cloned into a pGEM[®]-T Easy Vector for subsequent comparison between sorted and unsorted cells.

After Sanger sequencing, the analysis of 41 white colonies obtained from unsorted cells revealed only one CRISPR-Cas9-edited colony that contained a 30 bp deletion; all the other 40 white colonies were wild-type. In contrast, of 41 white colonies obtained from sorted cells, 26 were found to be CRISPR-Cas9-edited colonies. Of these, 73.1% (19 colonies) involved 4-36 bp deletions and 26.9% (7 colonies) involved 1-53 bp insertions in the CRISPR-Cas9 cut site (Figure 4).

In unsorted CRISPR-Cas9 transfected cells, only 2.4% of the white colonies (0.73 colonies/30 colonies analyzed) showed editing compared to sorted CRISPR-Cas9 transfected cells in which 63.4% of the colonies were edited (19 colonies/30 colonies analyzed). Thus, target gene editing was ~26-fold greater in CRISPR-Cas9 sorted white colonies than in unsorted white colonies.

Miller's TALEN transfections resulted in one edited colony for every 30 colonies analyzed (3.3%, or 3.3 for every 100 analyzed) when no cell sorting was applied before genomic purification, and four edited colonies for every 30 analyzed (13.3%, 13.3 for every 100 analyzed) when genomic DNA extracted from sorted RFP⁺/GFP⁺ cells was analyzed five days after cell sorting (Nerys-Junior *et al.*, 2014). The sequencing of 32 white colonies in the sorted and unsorted groups yielded the same proportion as previously described. All the TALEN-edited colonies involved deletions ranging from 9 to 21 bp; no insertions were observed (Figure 5).

Table 1 summarizes the main findings of this investigation and shows that CRISPR-Cas9 was much better at editing the beginning of the CCR5 gene than the most efficient TALEN described for this same genetic site (Miller *et al.*, 2011; Nerys-Junior *et al.*, 2014).

Discussion

TALEN and CRISPR-Cas9 transfections involve different reporter systems. Whereas TALEN requires the co-transfection of three plasmids (one for each TALEN arm and one for the pRGS-CR reporter system), CRISPR-Cas9 assembled in the pX458 plasmid requires the transfection of only one plasmid that encodes the entire CRISPR-Cas9 system and the reporter system simultaneously.

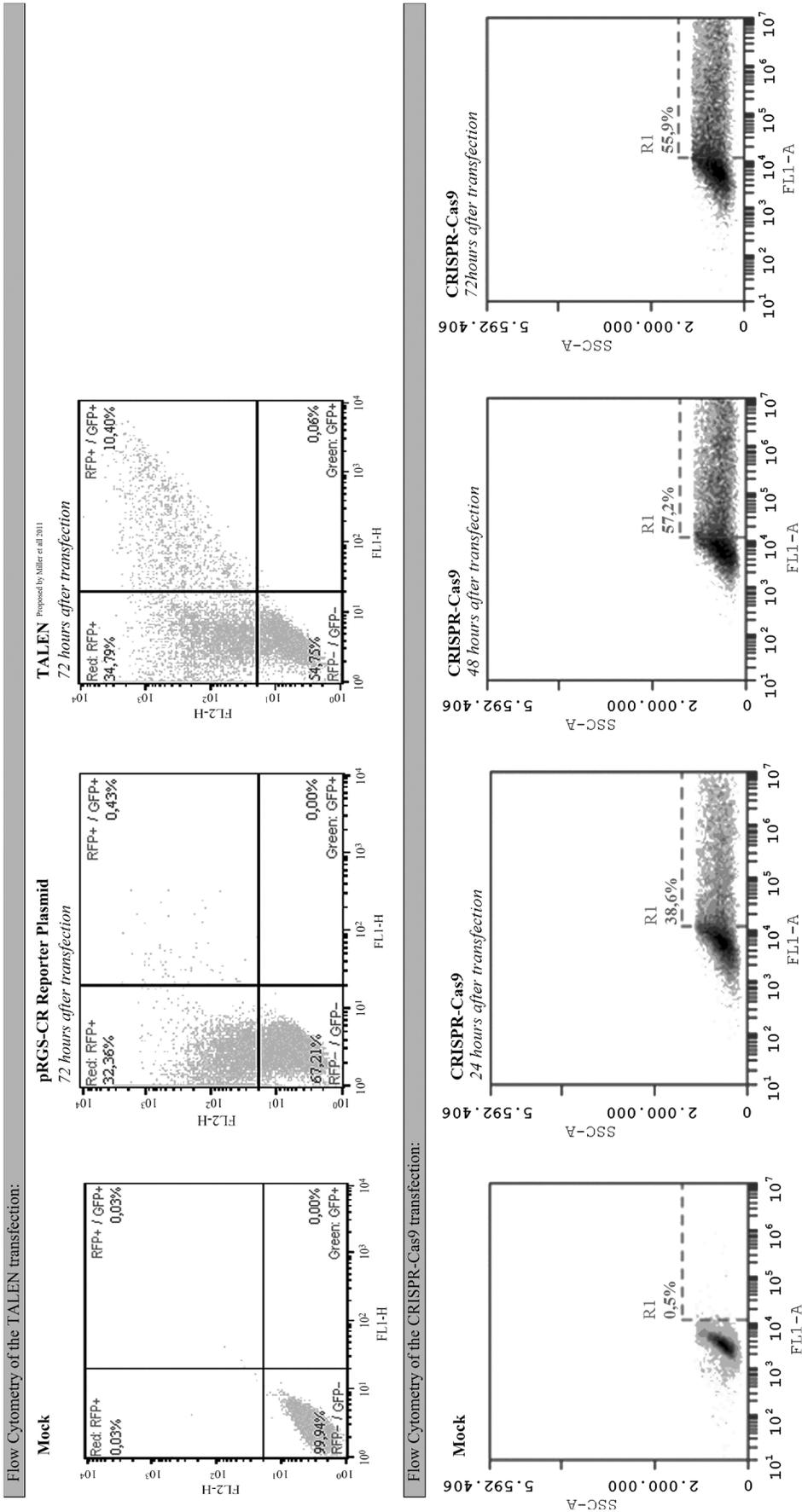
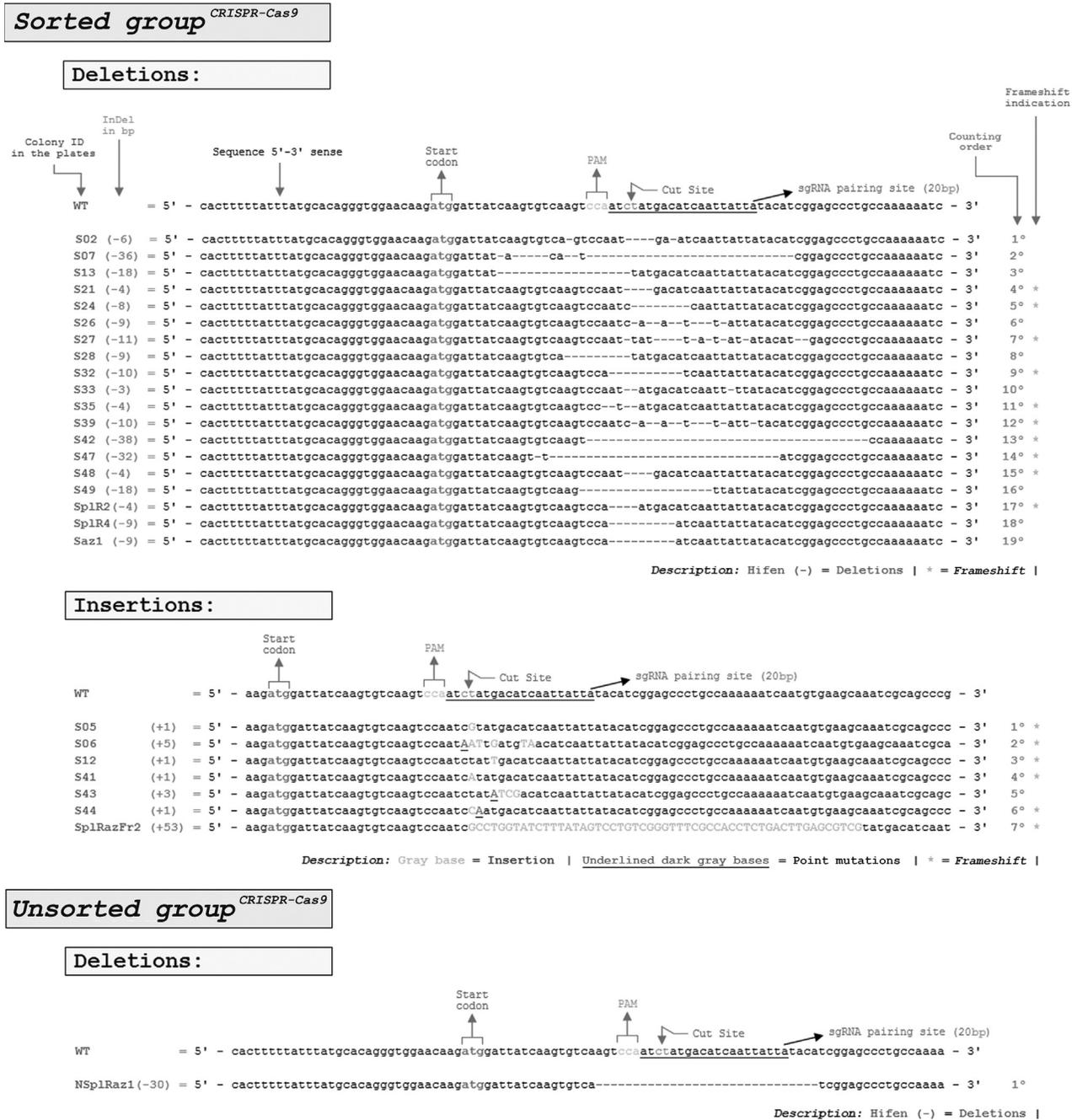


Figure 3 - Flow cytometry 24, 48, and 72 h after CRISPR-Cas9 transfection and 72 h after TALEN transfection. Our previous work and other reports showed that the highest proportion of RFP⁺/GFP⁺ cells was obtained 72 h after TALEN transfection. The frequency of GFP⁺ cells in CRISPR-Cas9 transfections was compared with a mock transfection without DNA. This same mock transfection was used to measure the frequency of RFP⁺/GFP⁺ cells in TALEN transfections. The highest proportion of GFP⁺ cells (57.2%) was obtained 48 h after CRISPR-Cas9 transfection.



Insertions:

Unsorted group CRISPR-Cas9

Deletions:

WT = 5' - cactttttattatgacacaggtggaacaagatggattatcaagtgccaagtccatctatgacatcaattattatcacatcggagccctgccaaaa - 3'

NSplRaz1 (-30) = 5' - cactttttattatgacacaggtggaacaagatggattatcaagtgcca-----tcggagccctgccaaaa - 3' 1°

Description: Hifen (-) = Deletions |

Figure 4 - Genomic editions identified in the CRISPR-Cas9 transfections. For both sorted and unsorted groups 41 *E. coli* JM109 white colonies were sequenced by Sanger sequencing. Whereas only one colony was edited in the unsorted group, 26 colonies were edited in the sorted group, indicating 26-fold more gene editions in the sorted group compared to the unsorted group. While the only identified colony in the unsorted group was a 30-bp deletion, in the sorted group 73.1% of the genomic editing (19 colonies) consisted of deletions and 26.9% (7 colonies) consisted of insertions. In the sorted group, approximately two-thirds of the editing generated a frameshift (16 of 26 editions), indicating random mutations.

Despite the disadvantage of TALEN compared to CRISPR-Cas9, when the need for co-transfections in TALEN experiments is required, in sorted RFP⁺/GFP⁺ TALEN-transfected cells TALEN production and action in the cell nucleus can be ensured, even though editing of the pRGS-CR reporter plasmid does not necessarily imply ed-

iting of the target gene (although this association is generally valid). On the other hand, not all editing of the pRGS-CR reporter plasmid will restore the GFP open reading frame and, in some cases, there may be target genome editing without internalization of the pRGS-CR reporter plasmid, or there may be target genome editing without ed-

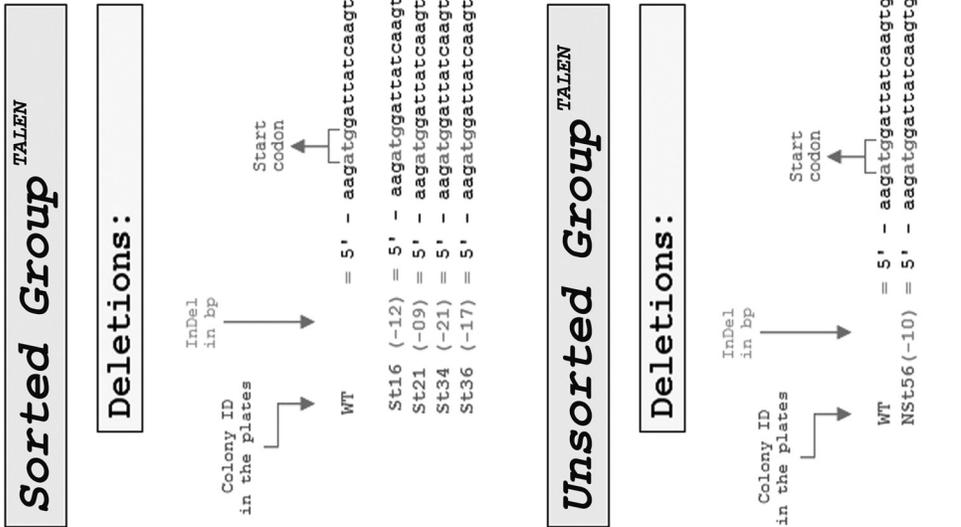


Figure 5 - Genomic editing identified in TALEN transfections. For both sorted and unsorted groups 32 *E. coli* JM109 white colonies were sequenced by Sanger sequencing. Whereas only one colony was edited in the unsorted group, four colonies were edited in the sorted group, indicating four-fold more gene editing in the sorted group compared with the unsorted group. The only edited colony in the unsorted group contained a 10-bp deletion, while in the sorted group 100% of the genomic editing (four colonies) consisted of deletions and no insertions. Frameshift editing appeared to be random since around two-thirds of the deletions generated a frameshift.

Table 1 - Summary of the differences between TALEN and CRISPR-Cas9 transfections. The percentage of gene knockouts from the total number of edited colonies (insertions and deletions) in CRISPR-Cas9 sorted cells was 61.5%.

	Cell sorting	Number of sequenced colonies	Number of edited colonies	% of edited colonies	Edited colonies/30 sequenced	Improvement in efficiency between sorted and unsorted cases	Difference in efficiency between sorted TALEN and sorted CRISPR-Cas9	Deletions (%)	Insertions (%)	% of gene knock-out from the total number of edited colonies (deletions)	% of gene knock-out from the total number of edited colonies (insertions)	Deletion range (bp)	Insertion range (bp)
TALEN	Unsorted	32	1	3.1	0.9	~4 times	~5 times	100	0	100	0	10	None
	Sorted	32	4	12.5	3.8			100	0	25	0	9-21	None
CRISPR-Cas9	Unsorted	41	1	2.4	0.7	~26 times		100	0	0	0	30	None
	Sorted	41	26	63.4	19			73.1	26.9	52.6	85.7	4-36	1-53
								(n=19)	(n=7)				

iting of the internalized pRGS-CR reporter plasmid. Despite these limitations and uncertainties, for to assure TALEN entry and action within the target cells, the usage of the pRGS-CR reporter plasmid is currently the best option.

In experiments with CRISPR-Cas9 assembled in the pX458 vector, the on-board GFP reporter system allows easier transfections and ensures Cas9 production within the GFP⁺ cell, but does not indicate the anchorage of the sgRNA in Cas9 for the correct production, assembly and action of the CRISPR-Cas9 system within the cell nucleus. Nevertheless, the system mediated by the pX458 vector allows an extreme potential correlation between GFP⁺ cells and target genome editing.

The sorting of RFP⁺/GFP⁺ cells in TALEN transfections appears to be the only effective approach for ensuring TALEN production and action within the sorted cell nucleus. In the case of CRISPR-Cas9 assembly, the sorting of GFP⁺ cells in CRISPR-Cas9 transfections is a highly efficient procedure that ensures CRISPR-Cas9 system production within the sorted cell, therefore strongly indicating a potential target genome editing.

Our previous work (Nerys-Junior *et al.*, 2014) showed the same efficiency for TALEN transfection compared to that reported by Miller *et al.* (2011), and the efficiency observed here for CRISPR-Cas9 transfection was comparable to that of previous studies (Ran *et al.*, 2013a,b). However, direct comparison of the efficiencies of TALEN and CRISPR-Cas9 for the same genetic portion of the CCR5 gene under the same conditions is a new important finding that has a direct bearing on the development of CCR5 gene editing studies and new gene therapies in the CCR5 gene. In this context, the use of a shorter genetic site to evaluate the efficiency of gene editing by TALEN and CRISPR-Cas9 is an important consideration, especially because chromatin structure, transcription rate and DNA methylation of the chromosomal position of the target gene influence TALEN and CRISPR-Cas9 action equally within the cell nucleus. Thus, differences in efficiencies are directly related to the efficiencies of TALEN and CRISPR-Cas9 themselves and are locus-specific, i.e., they are not necessarily applicable to other genetic loci.

Off-target effects apparently did not affect our analysis of the efficiency of TALEN and CRISPR-Cas9 in editing the CCR5 gene. The off-target activity of the CRISPR-Cas9 system can be easily overcome by using a shorter (< 20 nucleotides) recognition portion of sgRNA that does not affect on-target CRISPR efficiency (Fu *et al.*, 2014).

In our experimental conditions, unsorted and sorted TALEN transfections generated 3.3 and 13.3 edited colonies, respectively, for every 100 colonies analyzed. Thus, cell sorting in TALEN transfections using the pRGS-CR reporter plasmid generates four times more editions than in the unsorted TALEN group. In contrast, unsorted and

sorted CRISPR-Cas9 transfections generated 2.4 and 63.4 edited colonies, respectively, for every 100 colonies analyzed, indicating that cell sorting in CRISPR-Cas9 transfections using the pSpCas9(BB)-2A-GFP plasmid (*pX458*) as backbone generated 26 times more editions than in the unsorted CRISPR-Cas9 group. Together, these findings indicate that CRISPR-Cas9 was 4.8 fold more efficient than TALEN in editing the beginning of the CCR5 gene (13.3 edited colonies/100 colonies for sorted TALEN transfections versus 64.4 edited colonies/100 colonies for sorted CRISPR-Cas9 transfections). Our results also show that it is only possible to detect differences in efficiency when untransfected cells are separated from the correctly transfected cells.

In conclusion, CRISPR-Cas9 was better than TALEN for editing the beginning of the CCR5 gene, especially when greater editing efficiency and a higher proportion of edited cells are required.

Acknowledgments

This work is part of a thesis by Arildo Nerys-Junior, in partial fulfillment of the requirements for a PhD in the Post-Graduate Program in Genetics at the Federal University of Rio de Janeiro (UFRJ). The authors thank Rodrigo Delvecchio da Cunha, Gabriel dos Santos Gonçalves, Átila Duque Rossi, Ana Luiza Chaves Valadão, Cynthia Chester Cardoso, Renato Santana de Aguiar and Rodrigo de Moraes Brindeiro for critical discussions during the experiments, and Márcia Oliveira and Yolanda Alcântara for help and technical support during the experiments. We also thank members of the Laboratório de Virologia Molecular Animal (LAVIMOAN), the Laboratório de Pesquisas sobre o Timo (LPT, Instituto Oswaldo Cruz - Fiocruz) and the Serviço de Biotecnologia e Desenvolvimento Animal (SBDA, Instituto de Ciência e Tecnologia em Biomodelos – ICTB, Fundação Oswaldo Cruz - Fiocruz) for help during this study. This work was supported by the Programa de Pós-Graduação em Genética (PgGen) of the Instituto de Biologia (IB), Federal University of Rio de Janeiro (UFRJ), and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) and Fundação Universitária José Bonifácio (FUJS).

References

Arvey A, Agius P, Noble WS and Leslie C (2012) Sequence and chromatin determinants of cell-type-specific transcription factor binding. *Genome Res* 22:1723-1734.

Baltus TH, Kallaur AP, Lozovoy MA, Morimoto HK, Delongui F, Alfieri DF, Iriyoda TM, Dichi I, Simão AN and Reiche EM (2016) CCR5Δ32 (rs333) polymorphism is associated with the susceptibility to systemic lupus erythematosus in female Brazilian patients. *Rheumatol Int* 36:7-15.

Cannon P and June C (2011) Chemokine receptor 5 knockout strategies. *Curr Opin HIV AIDS* 6:74-79.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ and Voytas DF (2011a) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39:e82.

Fu Y, Sander JD, Reyon D, Cascio VM and Joung JK (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* 32:279-284.

Gaj T, Gersbach CA and Barbas 3rd CF (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31:397-405.

Gasiunas G, Barrangou R, Horvath P and Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* 109:E2579-2586.

Gonzales MJ, Machezano RN and Shafer RW (2001) Human immunodeficiency virus type 1 reverse-transcriptase and protease subtypes: classification, amino acid mutation patterns, and prevalence in a northern California clinic-based population. *J Infect Dis* 184:998-1006.

Grotto RMT and Pardini MIMC (2006) Molecular biology of the HIV-1 and genetics of human resistance to AIDS. *Arq Ciênc Saúde* 13:61-64.

He Z, Proudfoot C, Whitelaw CB and Lillico SG (2016) Comparison of CRISPR/Cas9 and TALENs on editing an integrated EGFP gene in the genome of HEK293FT cells. *SpringerPlus* 5:814.

Hill CM, Deng H, Unutmaz D, Kewalramani VN, Bastiani L, Gorny MK, Zolla-Pazner S and Littman DR (1997) Envelope glycoproteins from human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus can use human CCR5 as a coreceptor for viral entry and make direct CD4-dependent interactions with this chemokine receptor. *J Virol* 71:6296-6304.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31:827-832.

Hütter G, Nowak D, Mossner M, Ganepola S, Müssig A, Allers K, Schneider T, Hofmann J, Kücherer C, Blau O, *et al.* (2009) Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* 360:692-698.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier EA (2012) Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816-821.

Kim H, Um E, Cho SR, Jung C, Kim H and Kim JS (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* 8:941-943.

Lee BK, Bhingee AA, Battenhouse A, McDaniel RM, Liu Z, Song L, Ni Y, Birney E, Lieb JD, Furey TS, *et al.* (2012) Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome Res* 22:9-24.

Mandal PK, Ferreira LM, Collins R, Meissner TB, Boutwell CL, Friesen M, Vrbanc V, Garrison BS, Stortchevoi A, Bryder D, *et al.* (2014) Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* 15:643-652.

- DNA endonuclease in adaptive bacterial immunity. *Science* 337:816-821.
- Kim H, Um E, Cho SR, Jung C, Kim H and Kim JS (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* 8:941-943.
- Lee BK, Bhinghe AA, Battenhouse A, McDaniell RM, Liu Z, Song L, Ni Y, Birney E, Lieb JD, Furey TS, *et al.* (2012) Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome Res* 22:9-24.
- Mandal PK, Ferreira LM, Collins R, Meissner TB, Boutwell CL, Friesen M, Vrbanac V, Garrison BS, Stortchevoi A, Bryder D, *et al.* (2014) Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* 15:643-652.
- McGowan JP and Shah S (2010) Understanding HIV Tropism. Physicians' Research Network (PRM), New York, Vol. 15, p. 1002-1010.
- Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29:143-148.
- Moarii M, Boeva V, Vert JP and Reyal F (2015) Changes in correlation between promoter methylation and gene expression in cancer. *BMC Genomics* 16:873.
- Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye T and Cathomen T (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res* 39:9283-9293.
- Narlikar GJ, Fan HY and Kingston RE (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108:475-487.
- Nemudryi AA, Valetdinova KR, Medvedev SP and Zakian SM (2014) TALEN and CRISPR/Cas genome editing systems: tools of discovery. *Acta Naturae* 6:19-40.
- Nerys-Junior A, Costa LC, Braga-Dias LP, Oliveira M, Rossi AD, da Cunha RD, Gonçalves GS and Tanuri A (2014) Use of the heteroduplex mobility assay and cell sorting to select genome sequences of the CCR5 gene in HEK 293T cells edited by transcription activator-like effector nucleases. *Genet Mol Biol* 37:120-126.
- Pokorny V, McQueen F, Yeoman S, Merriman M, Merriman A, Harrison A, Highton J and McLean L (2005) Evidence for negative association of the chemokine receptor CCR5 d32 polymorphism with rheumatoid arthritis. *Ann Rheum Dis* 64:487-490.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA and Zhang F (2013a) Genome engineering using the CRISPR-Cas9 system. *Nat Protocols* 8:2281-2308.
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, *et al.* (2013b) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 155:479-580.
- Reiche EM, Ehara Watanabe MA, Bonametti AM, Morimoto HK, Akira Morimoto A, Wiechmann SL, Matsuo T, Carvalho De Oliveira J and Vissoci Reiche F (2008) Frequency of CCR5-Δ32 deletion in human immunodeficiency virus type 1 (HIV-1) in healthy blood donors, HIV-1-exposed seronegative and HIV-1-seropositive individuals of southern Brazilian population. *Int J Mol Med* 22:669-675.
- Riordan SM, Heruth DP, Zhang LQ and Ye SQ (2015) Application of CRISPR/Cas9 for biomedical discoveries. *Cell Biosci* 5:33.
- Scheibel I, Veit T, Neves AG, Souza L, Prezzi S, Machado S, Kohem C, Icarelli M, Xavier R, Brenol JC, *et al.* (2008) Differential CCR5Δ32 allelic frequencies in juvenile idiopathic arthritis subtypes: evidence for different regulatory roles of CCR5 in rheumatological diseases. *Scand J Rheumatol* 37:13-17.
- Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, *et al.* (2014) Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods* 11:399-404.
- Silva-Carvalho WH, de Moura RR, Coelho AV, Crovella S and Guimarães RL (2016) Frequency of the CCR5-Δ32 allele in Brazilian populations: A systematic literature review and meta-analysis. *Infect Genet Evol* 43:101-107.
- Souza MS, Souza CA, Cunha LM, Souza AQ, Morais MS and Rabenhorst SH (2015) A new look at osteomyelitis development – focus on CCR5Δ32. Study in patients from northeast Brazil. *Infect Genet Evol* 31:61-63.
- Telini B, Veit TD, Chies JA and Vianna P (2014) The CCR5Δ32 polymorphism as a pre-eclampsia susceptibility marker: an evaluation in Brazilian women. *Arch Gynecol Obstet* 290:1-3.
- Tsai SQ and Joung JK (2016) Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* 17:300-312.
- Valton J, Daboussi F, Leduc S, Molina R, Redondo P, Macmaster R, Montoya G and Duchateau P (2012a) 5'-Cytosine-phosphoguanine (CpG) methylation impacts the activity of natural and engineered meganucleases. *J Biol Chem* 287:30139-30150.
- Valton J, Dupuy A, Daboussi F, Thomas S, Maréchal A, Macmaster R, Melliand K, Juillerat A and Duchateau P (2012b) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J Biol Chem* 287:38427-38432.
- Vojta A, Dobrinic P, Tadic V, Bockor L, Korac P, Julg B, Klasic M and Zoldos V (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res* 44:5615-5628.
- Weber D, Heisig J, Kneitz S, Wolf E, Eilers M and Gessler M (2015) Mechanisms of epigenetic and cell-type specific regulation of Hey target genes in ES cells and cardiomyocytes. *J Mol Cell Cardiol* 79:79-88.
- Wiedenheft B, Sternberg SH and Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482:331-338.
- Xu L, Yang H, Gao Y, Chen Z, Xie L, Liu Y, Liu Y, Wang X, Li H, Lai W, *et al.* (2017) CRISPR/Cas9 mediated CCR5 ablation in human hematopoietic stem/progenitor cells confers HIV-1 resistance *in vivo*. *Mol Ther* 25:1782-1789.
- Yin C, Zhang T, Qu X, Zhang Y, Putatunda R, Xiao X, Li F, Xiao W, Zhao H, Dai S, *et al.* (2017) *In vivo* excision of HIV-1 provirus by saCas9 and multiplex single-guide RNAs in animal models. *Mol Ther* 25:1168-1186.

Internet Resources

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ and Voytas DF (2011b) Golden Gate TALEN assembly. Addgene.

<http://www.addgene.org/static/cms/files/GoldenGateTAL>

Assembly2011.pdf (January 15, 2013 and October 31, 2014).

Associate Editor: Carlos F.M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.