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

Arildo Nerys-Junior¹, Lendel C. Costa¹, Luciene P. Braga-Dias², Márcia Oliveira³, Átila D. Rossi¹, Rodrigo Delvecchio da Cunha¹, Gabriel S. Gonçalves¹ and Amilcar Tanuri¹

¹Laboratório de Virologia Molecular, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. 
²Plataforma de Animais Transgênicos, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil. 
³Laboratório de Pesquisas Sobre o Timo, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil.

Abstract

Engineered nucleases such as zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) are one of the most promising tools for modifying genomes. These site-specific enzymes cause double-strand breaks that allow gene disruption or gene insertion, thereby facilitating genetic manipulation. The major problem associated with this approach is the labor-intensive procedures required to screen and confirm the cellular modification by nucleases. In this work, we produced a TALEN that targets the human CCR5 gene and developed a heteroduplex mobility assay for HEK 293T cells to select positive colonies for sequencing. This approach provides a useful tool for the quick detection and easy assessment of nuclease activity.

Key words: CCR5, genome editing, gene knockout, TALEN, HMA.

Received: June 14, 2013; Accepted: October 1, 2013.

Introduction

Programmable nucleases such as zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) are useful tools for producing genomic modifications (Osborn et al., 2011; Moore et al., 2012). These enzymes consist of two polyprotein subunits that have a nuclear localization signal (NLS), as well as a specific programmable sequence that recognizes a desired DNA sequence, followed by a I restriction enzyme which works only when dimerized with the monomer from the other subunit (Cermak et al., 2011a; Mussolino et al., 2011). These nucleases produce double-strand breaks (DSB) by cleaving at specific sites in the chromosome. The resulting DSB are subsequently processed by the non-homologous end-joining (NHEJ) DNA repair mechanism that may give introduce random mutations, deletions and/or insertions that can modify the open reading frame (ORF) of the targeted protein (Pan et al., 2012; Xiao et al., 2013).

The use of engineered nucleases can also facilitate the introduction of knock-in genes. In this case, a given gene or desired construction with flanking homologous sequences containing 50-500 base pairs (bp) can be introduced at the cleaved site by the DNA repair machinery through homologous recombination (HR) (Orlando et al., 2010). Identifying the desired genomic modifications by sequencing the cell progeny after nuclease treatment is one of the biggest obstacles to the widespread use of nucleases, primarily because of the need to extract the genome from TALEN-treated cells, amplify targeted regions by the polymerase chain reaction (PCR), clone the amplicons into a cloning vector and, finally, sequence all of the colonies obtained.

One solution to this problem is to use a simple heteroduplex mobility assay (HMA) (Delwart et al., 2012) that can detect colonies carrying the TALEN-edited sequences and is easy to perform. In this work, we designed two TALENs that target the CCR5 (C-C chemokine receptor type 5) gene, the main co-receptor used by human immunodeficiency virus-1 (HIV-1) to enter host cells (Djin-Ye Oh et al., 2008; Holt et al., 2010; Ramalingam et al., 2012; Yao et al., 2012). Sanger sequencing was used to identify colonies carrying the modified gene.

Materials and Methods

TALEN assembly, plasmid construction and strategies

Two strategies were used to produce the human CCR5 gene knockout and were based on constructing
TALENs that recognized different sites at the start of the first CCR5 gene exon (Figure 1). In the first strategy, the TALEN recognition site started 135 bp downstream from the first CCR5 codon (Mussolino et al., 2011), and in the second strategy it started 155 bp from the first codon (Miller et al., 2011). TALEN plasmids were assembled utilizing a Golden Gate TALEN assembly kit (AddGene, Cambridge, MA, USA) according to the manufacturer’s instructions (Cermak et al., 2011b). To demonstrate intracellular TALEN activity, we constructed a reporter plasmid carrying the TALEN target CCR5 sequence, based on the pRGS-CR vector (Kim et al., 2011). The sequence was inserted between a red fluorescent protein (RFP) and green fluorescent protein (GFP). This insertion placed the GFP gene out of frame so that the transfected cells expressed only RFP. TALEN expression in transfected HEK 293T cells resulted in digestion of the target sequence in pRGS-CR and subsequent NHEJ-mediated repair of the plasmid restored the GFP reading frame and allowed pRGS-CR to express both fluorescent proteins (Figure 2).

Figure 1 - TALEN genomic recognition sites and corresponding RVD sequences. (A) The underlined sequences represent the RVD recognition sites of the right and left subunits of CCR5 TALEN 1 proposed by Mussolino et al. (2011). (B) The underlined sequences represent the RVD recognition sites of the right and left subunits of CCR5 TALEN 2 proposed by Miller et al. (2011) (A and B). The nucleotides in bold indicate the “spacer site” for cleavage by FokI nuclease. The thymine highlighted in dark gray before the RVD recognition sites represents the first fixed nucleotide (thymine) targeted by any TALEN. The light gray highlighted nucleotide near the “spacer site” represents the nucleotide recognized by the last repeat RVD. (C) The nucleotides in bold represent the sequence inserted in the pRGS-CR reporter plasmid (Kim et al., 2011) between RFP and GFP out of frame (see Figure 2 for more details). (D) Amino acid sequence of RVD peptides and their respective targeted nucleotides. (E) RVD sequences and their corresponding target sites in the human genome.
Transfection of HEK 293T cells

Plasmids containing the TALENs (Miller et al., 2011) and the pRGS-CR reporter plasmid were co-transfected at a ratio of 1:1:2, in a total of 2 ug of DNA. Lipofectamine 2000™ was used for transfection, according to the manufacturer’s instructions. The HEK 293T cells were transfected at confluence (4x10⁵ cells) in six-well plates.

Flow cytometry and cell sorting

Fluorescent activated cell sorting (FACS) was done three days after transfection of HEK 293T cells. Control cells were transfected only with the pRGS-CR reporter plasmid whereas under evaluation transfected cells received TALEN-coding plasmids (one plasmid for the TALEN right arm and other for the TALEN left arm) and the pRGS-CR reporter plasmid. Cell sorting was done with a MoFlo™ flow cytometer (Dako Cytomation, Beckman Coulter, Brea, California) to segregate cells with the desired phenotype (RFP+/GFP+). The resulting data were used to calculate the proportion of cells with the desired genomic modifications.

Heteroduplex mobility assay (HMA)

Three days after transfection, cell genomic DNA was extracted 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 that included 200 bp up and downstream from the TALEN target site was cloned using a TA cloning kit (Invitrogen Life Technology, Carlsbad, CA, USA) and transformed in E. coli JM109 (Promega, Madison, WI, USA) competent cells. White positive colonies were screened using an LB (Luria Broth) 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).

The HMA was done using 150 ng of the CCR5 wild-type amplicon, 150 ng of the amplicon generated from the selected colony and 10% of the reaction volume of Neb Buffer 2 10X (New England Biolabs, Ipswich, MA, USA) or heteroduplex annealing buffer (Delwart et al., 2012). The HMA reactions was heated to 99.9 °C for 3 min and immediately placed on ice for at least 5 min to form the heteroduplexes. HMA products were analyzed by electrophoresis in 10% acrylamide gels and colonies with heteroduplex patterns were selected for sequencing.

T7 endonuclease assay

For the T7 endonuclease (New England Biolabs) digestion 50% of the HMA reaction volume and 1 U of T7 endonuclease were mixed in a separate tube and incubated at 37 °C for 1 h. In this case, Neb Buffer 2 replaced the heteroduplex annealing buffer. Controls without T7 enzyme were run simultaneously. Electrophoresis was done in 10% acrylamide gels in TBE 1X to confirm the positive samples. The gels are run at 120 V for 4 h when nondigested reactions are analyzed and for 2 h when digested reactions are loaded. The gels were subsequently stained with ethidium bromide and visualized in a UV transilluminator.

DNA sequencing

The colonies with heteroduplex patterns were sequenced using a ABI BigDye Terminator sequencing kit (Applied Biosystems, Carlsbad, CA, US) on an Applied Biosystems 3130 Genetic Analyzer. Some heteroduplex-negative colonies were also sequenced as a negative control for HMA selection. For this, genomic DNA extracted from non-transfected HEK 293T cells was used as a wild-type reference. All sequences were aligned using SeqMan software v8.1.2 (DNAStar, Madison, WI, US).
Results

To control the activity of human CCR5 gene-targeting TALENs, HEK 293T cells were co-transfected with a reporter plasmid pRGS-CR and vectors expressing the two TALEN-specific subunits. We tested two pairs of TALENs that target different specific recognition sites in the first exon of the CCR5 gene (see Figures 1 and 2 for details). Flow cytometry showed that both of the engineered nucleases had almost the same efficiency in producing double-labeled RFP+/GFP+ cells for vector pRGS-CR (Figure 3).

After FACS, genomic DNA was extracted from double-labeled cells and compared with genomic DNA from cells not subjected to FACS. For this, the target region was amplified and cloned. Thirty positive colonies for each condition (sorted and unsorted cells) were amplified using the same PCR reaction and the DNA products were further analyzed with the HMA. The frequency of HMA-positive colonies in sorted and unsorted cells was 20% and 3%, respectively (Figure 4), and the size of the heteroduplex band (assessed electrophoretically) increased proportionally to the size of the amplicon deletion; the heteroduplex band

![Flow cytometry](image)

**Figure 3** - Flow cytometry after transfection of TALEN-containing plasmids that target two distinct CCR5 first exon positions. Mock transfection was done solely with Lipofectamine 2000 (Invitrogen). The reporter transfection was done with Lipofectamine 2000 and the pRGS-CR reporter plasmid. Strategy 1 transfection was done with Lipofectamine 2000, the pRGS-CR reporter plasmid and the plasmid that encoded the CCR5 TALEN that targeted position 135th base pair of the first exon of CCR5 (Mussolino et al., 2011). Strategy 2 transfection was the same as Strategy 1 except that the plasmid contained a TALEN that targeted position 155 of the first exon of CCR5 (Miller et al., 2011). The percentages depicted represent the percentage of cells gated in each quadrant of the flow cytometry analysis.
was always clearly distinguishable from the homoduplex band and distant from the top of the gel. Digestion with T7 endonuclease assay to confirm the existence of heteroduplex strands showed that third of the colonies was positive and two thirds negative (Figure 5).

The sequencing of five colonies from unsorted (NS-C1 to NS-C5) and sorted (S-C1 to S-C5) cells shown in Figure 4 indicated that all HMA-positive colonies carried information on the genomic deletions induced by TALEN and that all HMA-negative colonies were wild-type colonies, as expected (Figure 6). Careful analysis of the HMA-positive colonies in Figure 4 (S-C2 and S-C5) showed that the larger the deletion the slower the migration, i.e., greater mass, of the heteroduplex band (compare the migration of S-C2, which has an 18 bp deletion, with that of S-C5, which has a 10 bp deletion).

**Discussion**

ZFPs and TALENs are powerful tools for genetic modification (Cermak et al., 2011a), but screening and detecting such events is difficult and laborious so that new approaches are required to solve these problems (Kim et al., 2011). The results described here agree with those of Kim et al. (2011), who observed that cells sorted with FACS had a greater number of genomically-edited cells than did unsorted cells. The use of red and green fluorescent proteins greatly facilitated the assessment of transfection efficiency and TALEN activity, respectively, in the pRGS-CR reporter plasmid. By careful selection, the number of edited cells in the pool of sorted cells was increased by eight-fold when compared with unsorted cells (~3.34% of TALEN-edited colonies before sorting vs. 26.7% after cell sorting).

The positive colony shown in Figure 5 indicated that the T7 assay is a good way to screen for and confirm HMA-positive samples (Kim et al., 2011; Mussolino et al., 2011). Two colonies showing heteroduplexes and three
with no sign of heteroduplex formation were selected from sorted cells and sequenced in both strands. As a control, we sequenced five random colonies from unsorted cells that showed no HMA formation. The sequences from colonies showing heteroduplexes in the HMA had evidence of TALEN-mediated mutagenesis, in contrast to HMA-negative colonies that had only the wild-type sequence (Figure 6). Interestingly, all TALEN-edited sequences in the genomes of sorted cells consisted of deletions 10-18 bp in size.

The efficiency of TALEN editing observed here was comparable to that reported by Miller et al. (2011) and Mussolino et al. (2011). The sensitivity and specificity of the test were 100% in both cases, with CI 95% of 49-100% and 89-100%, respectively. These results indicate that the HMA is a suitable and reliable test for detecting TALEN-induced modifications. In other fields, the HMA is also a reliable method for identifying mutations, insertions and deletions in genes of interest. Indeed, this assay has been successfully used in epidemiological assays and in searches for linkage in infections, as well as in studies of polymorphism, genotyping, HIV-1 subtyping, phylogenetic analysis, natural HIV-1 recombination, viral genetic heterogeneity and diversity (Heyndrickx et al., 2000; White et al., 2000; Agwale et al., 2001; Mehta et al., 2010; Huang et al., 2011; Mendonça et al., 2011; Manigart et al., 2012).

These findings, together with recent work in zebrafish (Ota et al., 2013), show that the HMA is a reliable assay for detecting DNA edited by TALEN. The assay is sufficiently sensitive and specific to detect colonies carrying mutations, insertions and deletions of interest, as shown by the fact that DNA sequencing confirmed the mutations in all colonies with a heteroduplex band in the HMA. Digestion with T7 endonuclease confirmed the nature of the heteroduplex structure. In addition, FACS efficiently concentrated cells with the genomic alterations of interest when combined with a suitable pRGS-CR reporter plasmid (Kim et al., 2011). FACS can also be used to enrich the pool of genetically modified cells during cell sorting; this could be useful for gene therapy involving autotransplantation since it would involve the transfer of a lower number of unedited cells.

In conclusion, HMA combined with FACS is a reliable and accurate means of detecting TALEN-induced mutations. This combination of techniques could be particularly useful for low resource laboratories that wish to analyze TALEN experiments.

Acknowledgments

We thank Drs. Adam J. Bogdavone and Clarice Schimidt of the Department of Plant Pathology, Iowa State University, USA, for important discussions during this work. This research was sponsored by CNPq, FAPERJ and CAPES.

References

Agwale SM, Robbins KE, Odama L, Saekhou A, Zeh C, Edubio A, Njoku OM, Sani-Gwarzo N, Gboun MF, Gao F, et al. (2001) Development of an env gp41-based heteroduplex
mobility assay for rapid human immunodeficiency virus type 1 subtyping. J Clin Microbiol 39:2110-2114.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Sonia 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.

Djin-Ye Oh, Jessen H, Kücherer C, Neumann K, Oh N, Poggensee G, Bartmeyer B, Jessen A, Pruss A, Schumann RR, et al. (2008) CCR5Δ32 genotypes in a German HIV-1 seroconverter cohort and report of HIV-1 infection in a CCR5Δ32 homozygous individual. PLoS ONE 3:e2747.

Heyndrickx L, Janssens W, Zekeng L, Musonda R, Anagonou S, Van der Auwera G, Coppens S, Vereecken K, De Witte K, Van Rampelbergh R, et al. (2000) Simplified strategy for detection of recombinant human immunodeficiency virus type 1 group M isolates by gag/env heteroduplex mobility assay. Study Group on Heterogeneity of HIV Epidemics in African Cities. J Virol 74:363-370.

Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, Crooks GM, Kohn DB, Gregory PD, Holmes MC, et al. (2010) Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol 28:839-847.

Huang TP, Yeh Y and Tzeng DD (2011) Barcode-like heteroduplex DNA pattern as an aid for rapid identification of anthracnose fungi. Nat Biotechnol 28:72-78.

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

Manigart O, Boers D, Karita E, Hawkins PA, Vwalika C, Bakker J, Mulenga J, Derdeyn CA, Allen S and Hunter E (2012) A gp41-based heteroduplex mobility assay provides rapid and accurate assessment of intrasubtype epidemiological linkage in HIV type 1 heterosexual transmission pairs. AIDS Res Hum Retroviruses 28:1745-1755.

Mehta PR, Nema S, Paramjpe S, Ingole N, Wanjare S and Nataraj G (2010) Study of HIV-1 subtypes in serodiscordant couples attending an integrated counselling and testing centre in Mumbai using heteroduplex mobility analysis, protocol version 5 from the Laboratory of James I. Mullins. NIH AIDS Research & Reference Reagent Program. http://www.addgene.org/static/cms/files/GoldenGateTALENassembly2011.pdf (January 15, 2013).

Mendonaça MC, Ferreira AM, Santos MG, Oviedo EC, Bello MS, Siqueira MM, Maceira JM, von Hubinger MG and Couto CM (2011) Genotyping of human parvovirus B19 in clinical samples from Brazil and Paraguay using heteroduplex mobility assay, single-stranded conformation polymorphism and nucleotide sequencing. Mem Inst Oswaldo Cruz 106:502-504.

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.

Moore FE, Reyon D, Sander JD, Martinez SA, Blackburn JS, Khayter C, Ramirez CL, Joung J and Langenau DM (2012) Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS One 7:e37877.

Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye T and Cathomen T (2011) A novel TALEN nuclelease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39:9283-9293.

Ota S, Hisano Y, Muraki M, Hoshijima K, Dahlem TJ, Grunwald DJ, Okada Y and Kawasawa A (2013) Efficient identification of TALEN-mediated genome modifications using heteroduplex mobility assays. Genes Cells 18:450-459.

Pan Y, Xiao L, Li AS, Zhang X, Sirosi P, Zhang J and Li K (2012) Biological and biomedical applications of engineered nucleases. Mol Biotechnol 55:54-62.

Ramalingam S, London V, Kandavelou K, Cebotaru L, Guggino W, Civin C and Chandrasegaran S (2012) Generation and Genetic Engineering of Human Induced Pluripotent Stem Cells Using Designed Zinc Finger Nucleases. Stem Cells Dev 22:595-610.

White PA, Zhai X, Carter I, Zhao Y and Rawlinson WD (2000). Simplified hepatitis C virus genotyping by heteroduplex mobility analysis. J Clin Microbiol 38:477-482.

Xiao A, Wu Y, Yang Z, Hu Y, Wang W, Zhang Y, Kong L, Gao G, Zhu Z, Lin S, et al. (2013). EENdb: A database and knowledge base of ZFNs and TALENs for endonuclease engineering. Nucleic Acids Res 41:D415-D422.

Yao Y, Nashun B, Zhou T, Qin L, Zhao S, Xu J, Esteban MA and Chen X (2012) Generation of CD34+ cells from CCR5-disrupted human embryonic and induced pluripotent stem cells. Hum Gene Ther 23:238-242.

Internet Resources

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Sonia NV, Bogdanove AJ and Voytas DF (2011b) Golden Gate TALEN Assembly. Addgene, 2011. http://www.addgene.org/static/cms/files/GoldenGateTALENassembly2011.pdf (January 15, 2013).

Delwart EL, Hanley EW, Herring B, Lear M, Lyagoba F, Rodriguez AG and Shankarappa R (2012) Heteroduplex mobility analysis, protocol version 5 from the Laboratory of James I. Mullins. NIH AIDS Research & Reference Reagent Program. http://ubik.microbiol.washington.edu/HMA/index.html; http://ubik.microbiol.washington.edu/HMA/HMA_Manual.pdf (January 15, 2013).

Associate Editor: Emmanuel Dias Neto