Comparison of different methods to overexpress large genes

Aline Lisie Ramos,[1] Fernanda Soares Niemann,[1] Adriana Silva Santos Duarte,[1] Karla Priscila Ferro,[1] Irene Santos,[1] Carolina Louzão Bigarella,[1] Antonio Filareto,[2,3] Sara Teresinha Olalla Saad[1]

[1]Department of Internal Medicine, School of Medical Science, Hematology and Transfusion Medicine, University of Campinas/Hemocentro-UNICAMP, Campinas, São Paulo, Brazil; [2]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA; [3]Veterans Affairs Palo Alto Health Care System, Palo Alto, California, USA

Abstract

Gain-of-function of very large transgene constructs can lead to genetic perturbations, providing researchers with the alternative of a powerful tool to identify pathway components which remain undetected when using traditional loss-of-function analysis. To promote longer-term expression, various systems for transgene integration have been developed, however large cDNA sequences are often difficult to clone into size-limited expression vectors. We attempted to overexpress *ARHGAP21*, a 5.874 kb gene, using different methodologies as plasmid, lentiviral and Sleeping Beauty (SB) transposon based gene transfer. Using lentiviral based transduction; an enormous amount of lentiviral supernatant was produced to obtain a satisfactory titration after double ultracentrifugation. However, U937 transduced cells showed only 50% of gene expression increase, which vanished after 5 days. SB transposon system application to overexpress *ARHGAP21* was a complete success. Nucleofecting SB-based vector plus SB100x transposase vector resulted in an expressive increase of gene and protein expression. Furthermore, the overexpression was maintained even after freezing and thawing processes. In conclusion, our work shows that the SB transposon system is the best choice for those seeking a stable and high gene expression. Once the overexpression is achieved, freezing cells and using them for a long time becomes possible.

Introduction

The discovery of restriction enzymes[1] and the development of recombinant DNA (rDNA) technology[2] has allowed researchers to manipulate DNA molecules to study the possible role of genes in biological systems and use them to develop new therapies and novel medicines.[3,4] Specific gene modification is referred to as gene targeting, in this sense, by using a host DNA repair pathway, via homologous recombination, an endogenous genomic locus can be replaced by an exogenous sequence when supplemented with a targeting vector. Gene targeting enables scientists to have control over the cellular genome through gene silencing, shut down, or overexpression.[5] Genome editing technologies enable the modification of any part of the human genome with extreme precision.[6,7] Rescue or gene corrections have been widely used as an important tool for understanding the mechanisms of a wide variety of diseases such as acute myeloid leukemia[8] and Duchene muscular dystrophy.[9] Overexpression of genes of interest (GOIs) helps to reveal their roles in the control of complex biological processes, providing a powerful tool to identify pathway components which may remain undetected using traditional loss of function analysis. Given the diversity of cell lines available for gene expression studies, taking transgenesis methods that permit the simple introduction of complex transgene systems in a consolidated manner into consideration, is quite a pertinent procedure.

Bacterial artificial chromosome (BAC) or simply plasmids are routinely used as expression vectors in non-viral gene therapy studies owing to their ease of construction and amplification. Moreover, plasmids are episomal and non-integrating, which reduces the risk of insertional mutagenesis compared with viral vectors. The choice of enhancer-promoter combination has a great impact on both level and duration of transgene expression. Viral enhancers and promoters derived from cytomegalovirus (CMV), respiratory syncytial virus and simian virus 40 are frequently used to achieve high levels of expression in mammalian cells and tissue types, this expression however, is usually transient.[10] Nevertheless, several high-profile studies of cell signaling have made use almost exclusively of transient overexpression.[11] The need for high levels of protein expression has prompted the search for new strategies, including technologies to obtain cells with a high number of copies of a given gene.[12] Therefore, the increased number of attempts to achieve *the best combination promoter sequence-50UTR* is not surprising.[13-15] For instance, experiments with the human cytomegalovirus major immediate-early gene virus promoter (pCMV) are part of expression vectors widely used for recombinant gene expression.[15,16] However some disadvantages render the use of plasmids imperfect, such as genetic instability, structural instability and metabolic burden.[17,18] In addition, resulting plasmids which are too large to be transformed into competent cells becomes an intractable problem.[19]

Viral-vectors based gene delivery can achieve higher transduction efficiency and long-term gene expression.[16,20,21] However, some unwanted effects may be associated, such as high costs, immunogenicity, carcinogenicity, poor target cell specificity, limitations in the size of the carried genes,[22] inability to transfer large genes,[23] limited DNA packaging capacity,[24-26] difficulty of vector production[22,27] and requirement for specialized biohazard containment in preparation or application.

Transposons are a family of mobile DNA elements that have been adapted as experimental tools for stable genomic integration of transgenes. In an effort to promote longer-term expression, transposition systems based on the recombinases, such as phiC31,[28] PiggyBac57 and Sleeping Beauty (SB),[29] have been developed. DNA transposons are discrete pieces of DNA with the ability to change their positions within the genome via a cut-and-paste mechanism called transposition. These mobile genetic elements can be harnessed as gene delivery vector systems that can be used as tools for versatile applications.[30] The paradigmatic use of any transposon-based vector system relies on transient expression of a transposase enzyme that enables genomic insertion of a GOI flanked by terminal inverted repeats (TIRs) of the transposon.[30] Moreover, as transposition proceeds through a cut-and-paste mechanism that only involves DNA, transposon vectors do not incorporate mutations by reverse transcription, as happens in retroviral stocks at reasonable frequencies, and due to their large capacity can tolerate larger and more complex transgenes.[31] Transposons have been shown to be applicable for development of cloned embryos[32] and iPS cell generation, through repeated expression of the transposase in reprogrammed cells, the chromosomally integrated vector can be excised from the genome, thereby resulting in genetically *clean* iPS cells.[33-35] However, there are applications where the use of one or another type of transposon systems could prove advantageous.

The SB transposon system[36] has several advantages over other transposon systems, including PB. In the first place, transposition efficiency has been greatly enhanced by the novel hyperactive SB100X transposase[37] that yields stable gene transfer efficiencies higher than that of PB[38] and comparable to those of integrating viral vectors, enabling highly efficient transgene integration and expression. Secondly, SB-based vectors are probably superior in terms of safety as SB shows a random genomic insertion profile with no overt preference for integrating into genes and their transcriptional regulatory regions, whereas PB shows preferential integration into transcription units[38,39] and in comparison to PB transposons[40] the TIRs of SB vectors have negligible enhancer/promoter activity.[41,42] Finally, as opposed to PB,[43] there are no SB-related sequences in mammalian genomes, thereby preventing potential cross-mobilization between endogenous and exogenously introduced transposons.

Our research group has identified a human gene composed by 25 exons and located on chromosome 10, initially termed *ARHGAP10*[44] and currently referred to *ARHGAP21*, that encodes for a 1957-aminoacid Rho-GAP, containing a PDZ, a PH, and a Rho-GAP domain. The cDNA is 7118 bp long and has an open reading frame (ORF) of 5874 bp. Back then, the authors suggested that this gene could play a role in hematopoiesis and could also be important for cell differentiation.[44] Almost fifteen years later, not much has been clarified regarding the functions of this gene or its encoded protein ARHGAP21, in the hematopoietic field. ARHGAP21 is a member of the RhoGAP family and has been described in many cell types as a controller of actin cytoskeleton dynamics. Moreover, as PDZ domains are commonly conserved protein-protein interaction domains[45] and as the modular structures of GAPs are important for their interaction with other proteins,[46] ARHGAP21 has been associated with several proteins, such as FAK,[47] PKCζ,[48] α catenin,[49] β arrestin,[50] Cdc42[47,48,51-53] and ARF1.[51,54,55] Furthermore, ARHGAP21 has been found to exert important functions such as mechanical stress,[48] cytoskeleton organization and cell migration[47,56,57] formation of cell-cell adherent junctions,[49,56] modulation of vesicles trafficking[53,58] differentiation,[44] cell proliferation, and gene expression.[57]

In this article we report the challenges of overexpressing large genes, particularly *ARHGAP21*, by using plasmid transfection, lentiviral transduction and SB transposon systems, in order to compare the efficiency of these methods and elect an efficient tool for further investigation of the role of this protein in hematopoiesis.

Materials and Methods

Cell culture

U937 and HL60 cell lines were used as model for human myeloid leukemia and HS-5 cell line as model for bone marrow stroma. These cell lines were obtained from ATCC (Philadelphia, PA, USA) and used in specific assays as appropriate. Cells were cultivated with the appropriate media (RPMI for U937, IMDM for HL-60 and DMEM for HS-5 cells), containing 10% (U937 and HS-5) or 20% (HL-60) of fetal bovine serum with penicillin/streptomycin and maintained at 37°C, 5% CO_2. HEK 293FT and HT1080 cell lines were obtained from ATCC (Philadelphia, PA, USA), and were used for lentivirus production and lentivirus supernatant stock titration, respectively. HEK 293FT cells were cultured in DMEM containing 0.1mM MEM Non-Essential Amino Acids, 6 mM L-glutamine, 1 mM MEM Sodium Pyruvate and 10% fetal bovine serum and HT1080 cells were cultured in MEM containing 10% fetal bovine serum with penicillin/streptomycin, both cell lines were maintained at 37°C, 5% CO_2.

ARHGAP21 plasmid electroporation based overexpression

The pCMVplasmid containing the cDNA encoding full-length human *ARHGAP21* (ARH) pCMV-ARH, and the empty vector (CTL) pCMV-CTL were purchased from OriGene Technologies (Rockville, MD, USA). Plasmid transfer was performed with the appropriate amount of vectors and using Gene Pulser II Electroporation System (Bio Rad), according to the manufacturer's instructions. After electroporation, HL60 cells were maintained for 48 hs in the appropriate medium at 37°C, 5% CO2. Geneticin was added to the culture medium for 10 days selection. After selection, cells were collected and ARHGAP21 expression analyzed by Real Time-PCR (RT-PCR) and Western Blotting (WB).

ARHGAP21 lentivirus based overexpression

The full-length protein sequence was obtained from a pEGFP-N2 (626) vector (kindly provided by Dr. Philip Chevrie - Curie Institute, Paris), by recombinant PCR using the *primers* A A A A A G G A T C C G G A G A A T G A A A T G A T G G C - CACGCGTCGG (forward) and AAAAAGAATTCGAAA-GACAGGGATGAAACTCTGC (reverse), and high fidelity DNA polymerase (PhusionTaq). The amplicon with the expected size (5.9 kb) was verified by agarose gel and extracted with QIAquick® PCR Purification kit (Qiagen) according to the manufacturer's instructions. The amplicon was digested with BamHI and EcoRI restriction enzymes and then cloned into a pENTR1A vector using T4 DNA ligase enzyme using Gateway® pENTR™ Dual Selection Vector kit (Invitrogen™), according to the manufacturer's instructions. Constructs were confirmed by DNA sequence analysis. Thereafter, we recombine pENTR-ARH and pLenti6.2/V5-DEST using pLenti6.2/V5-DEST Gateway® Vector Kit (Invitrogen™) according to the manufacturer's instructions. After confirmation of the integrity of the construction and correct orientation, lentiviral production was initiated. ViraPower™ Lentiviral Expression System (Invitrogen™) was used for lentivirus production, following the manufacturer's instructions. Lentiviral stock titration and U937 cell line transduction were performed according to kit recommendations. The minimum of three productions of lentiviral supernatant were necessary, which were ultra-centrifuged at 146 g per 2 hours at 4ºC, in order to obtain a satisfactory titration.

ARHGAP21 Sleeping Beauty system – transposon mediated overexpression

The transposon vector pKt2-iresGFP containing the full-length cDNA encoding ARHGAP21 and the empty vector were cloned by GenOne Biotechnologies (Rio de Janeiro, Rio de Janeiro, Brazil). *ARHGAP21* sequence was obtained from the pLenti6.2/V5-ARH construct, also used for lentivirus production, through XhoI and NotI digestion and was then cloned into a pKt2-iresGFP vector[9] kindly provided by Rita Perlingeiro (Lillehey Heart Institute, University of Minessota, USA). Constructs were confirmed by DNA sequence analysis. For transposase we used pCMV (CAT)T7-SB100, purchased from Addgene (Plasmid #34879), which yields high levels of transposon integration.[37]

The transfection was performed with the appropriate amount of plasmids and using Amaxa™ 4D-Nucleofector™, SF Cell Line 4D-Nucleofector® X Kit S (32 RCT) and adjusted Amaxa™ 4D-Nucleofector™ Protocol, according to the manufacturer's instructions. After nucleofection, HS-5 cells were maintained in the appropriate medium at 37°C, 5% CO2 until 75% confluence per well of a 6-well plate was reached. HS-5 GFP⁺ cell populations, HS-5 pKt2-ARH-iresGFP (ARH - ARHGAP21 overexpressing cells) and HS-5 pKt2-iresGFP (EPY- control cells), were then purified using cell sorting (BD FACSAria II). Serial sortings were performed until GFP⁺ populations reached 95-100%. Further, samples were collected and ARHGAP21 gene and protein expression analyzed by RT-PCR and WB. After the confirmation of ARHGAP21 overexpression, cells were expanded and frozen. Cells were then thawed, cultured and ARHGAP21 expression was analyzed again by RT-PCR and WB.

HL-60 pCMV-*ARHGAP21* xenograft model

10 days after thawing and Geneticin (G418) selection, HL-60 pCMV-ARH or pCMV-CTL cells were collected and ARHGAP21 gene and protein expression were analyzed by RT-PCR and WB. To verify the effect of ARHGAP21 overexpression on tumor development *in vivo*, male (NOD.CB17-Prkdcscid/J lineage) 6- to 8-week-old animals, from The Jackson Laboratory, bred at the Animal Facility Centre at the University of Campinas, under specific pathogen-free conditions, were matched for bodyweight before use. Experimental protocol was approved by the Institutional Committee of Animal Care. Mice were inoculated, s.c., in the dorsal region, on day 0 with 0.1 mL of HL-60 pCVM-ARH or pCMV-CTL ($1×10^7$ cells/mice). Every 7 days tumor volumes were evaluated according to the formula: tumor volume (mm3) ¼ (length × width2)/2. Mice were sacrificed after 21 days; tumors were then removed, minced, and homogenized in RNA extraction buffer.

RNA extraction and quantitative RT-PCR analysis

Total RNA from HL60-pCMV tumors, HS-5 pKT-ARH-iresGFP (ARH – ARHGAP21) and pKT2-iresGFP (EPY - Empty) cells was extracted using RNAspin Mini kit (GE Healthcare®, Cat. 25-0500-72), according to the manufacturer's instructions. RNAs integrity was analyzed with NanoVue Plus (GE Healthcare). Reverse transcription reaction was performed using RevertAid™ First Strand cDNA Synthesis Kit, according to the manufacturer's instructions (MBI Fermentas, St. Leon-Rot, Germany). Real-time detection of *ARHGAP21* amplification was performed in Realplex4 Mastercycler (Eppendorf) using Power SybrGreen PCR Master Mix (Applied Biosystems) and specific primers: forward 5'-AGGCAAACTTTGCTTGGTGCTA-3' and reverse 5'-ACT-GAGAAGTTTCCTTTCCGACTC-3'. HPRT was used as the housekeeping gene and the sequence of primers used was: forward 5'- AGGCAAACTTTGCTTGGTGCTA-3' and reverse 5'-CTCAGCCTTTCCTTTGAAGAGTCA-3'. Relative levels of gene expression were calculated using the equation: $2^{-\Delta\Delta CT}$[26]. A negative 'No Template Control' was included for each primer pair. Three replicas were run on the same plate for each sample.

Western blotting

Cell lysates were prepared using ice-cold 1x RIPA buffer (150 mM NaCl, 50 mM Tris-Cl (pH 7.5), 1% NP 40 (Igepal), 0.5% Sodium Deoxycholate, 0.1% SDS and 0.1% Sodium azide) in combination with phosphatase and protease inhibitors (10 mM Na_3VO_4, 10mM $Na_4P_2O_7$, 25mM PMSF and 0.1 mg/mL aprotinin). Equal amounts of cell lysates were subjected to SDS-PAGE (sodium sulfate polyacrylamide gels under reducing conditions), and transferred to nitrocellulose membranes, which were incubated overnight at 4°C with primary antibodies: rabbit anti-human ARHGAP21 1:500 (Bethyl*), and goat anti-human b-actin (1:2,000) (sc:1616, Santa Cruz). After incubation with HRP-conjugated secondary antibodies anti-rabbit (1:10,000), and anti-goat (1:4,000), membranes were incubated with Westar ECL-Sun

(Cyanagen, Cat.XLS06,0050) and exposed in Alliance 2.7 (UVITEC - Cambridge). Quantitative analyses of the optical intensities of protein bands were carried out with UVI BAND MAX – version 15.oba-Dongle (UVITEC - Cambridge) and normalized by actin for protein expression.

*Polyclonal antibody for ARHGAP21 was generated against a synthetic peptide (KSDSGSLGDAKNEKE), corresponding to residues 1856-1870 of the human protein and affinity purified by Bethyl Laboratories, Inc.

Statistical analysis

Statistical analysis was performed using GraphPad Prism7 software (GraphPad Software, Inc., San. Diego, CA, USA). Data were expressed as the mean ± SD. For comparisons, an appropriate Student's T-test or One-Way ANOVA test was used. The level of significance was set at $P<0.05$. All experiments were repeated a minimum of three independent times.

RESULTS

ARHGAP21 plasmid electroporation based overexpression

Initially, we attempted to overexpress *ARHGAP21* by transfecting a pCMV based plasmid containing this protein full-length sequence into HL60 cells. We choose HL60 cells because ARHGAP21 expression on this cell line is extremely low. After 10 days of transfection and Geneticin (G418) selection, cells were collected and ARHGAP21 expression analyzed by RT-PCR and WB. The transfection was very efficient and we obtained a gene expression 300 to 500 times higher than those in control cells. The increase in the expression was also observed at protein levels, which were almost undetectable before (Figure 1A). In addition to

in vitro assays, we designed *in vivo* experiments. For this purpose, the ARHGAP21 overexpressing cells had to be frozen to be transplanted later on. After thawing, the elevated expression initially obtained was not maintained, however there was still a satisfactory increase (Figure 1B). We then performed the subcutaneous xenograft transplantation of those cells. Unfortunately, we found no difference regarding tumor growth (data not shown). Furthermore, ARHGAP21 expressions in tumor cells after tumor harvest were almost undetectable (Figure 1C).

ARHGAP21 lentivirus based overexpression

In order to achieve a stable overexpression of ARHGAP21, we elected lentivirus (LV) based gene delivery as a method. The production of LV was not as efficient as desired. HEK 293FT cells had to be transfected many times to obtain a large volume of LV supernatant, and then ultra-centrifuged a minimum of two times to obtain a satisfactory titration (Figure 2A). For these experiments, we chose the U937 cells as this cell line is usually easily transduced. After 10 days of transduction and Blasticidin selection, we obtained an increase of approximately 50% in *ARHGAP21* gene expression (Figure 2B). However, once again the expression was completely lost after five days (Figure 2B).

Sleeping Beauty system – transposon mediated *ARHGAP21* overexpression

As our previous attempts to stably overexpress ARHGAP21 were not efficient, we then elected the SB Transposon System. Here we choose the HS-5 human bone marrow stromal cell line as the cells are adherent and we believed more easily nucleofected. We performed several nucleofections, under different conditions, until we were able to obtain a reasonable plasmid transfer efficiency in which we were able to generate a suitable population from which we started to purify GFP+ populations containing pKT2-iresGFP (EPY) and pKT2-ARH-iresGFP (ARH). The percentage
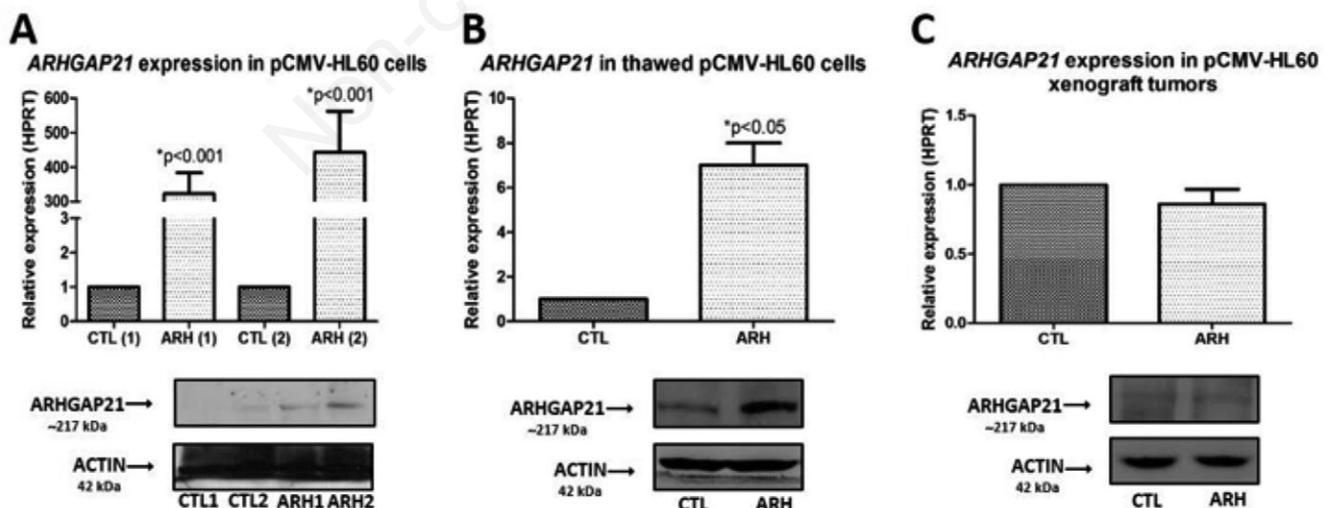

Figure 1. Efficacy of plasmid based gene transfer. (A) Gene (upper panel) and Protein (lower panel) expression of ARHGAP21 in two different pools of HL60 cells transfected with pCMV-ARHGAP21 or pCMV-Control and selected for 10 days with Geneticin; **(B)** Gene (upper panel) and Protein (lower panel) expression of ARHGAP21 in HL60 cells transfected with pCMV-ARHGAP21 or pCMV-Control after thawing; **(C)** Gene (upper panel) and Protein (lower panel) expression of ARHGAP21 in xenograft tumors generated by thawed HL60 cells transfected with pCMV-ARHGAP21 or pCMV-Control. HPRT (hypoxanthine phosphoribosyl transferase)=internal control gene for quantitative real-time-PCR expression analysis.

of GFP+ cells on the first FACS sorting of the HS-5 nucleofected cells were extremely low for both sequences, approximately 8.000 HS-5-pKT2-ARH-iresGFP and 24.000 HS-5-pKT2-iresGFP cells (Data not shown). At the second FACS sorting, cells with pKT2-iresGFP sequence yielded 50% of GFP+ cells, and from there on achieved more than 99%. The purification of the HS-5 carrying pKT2-ARH-iresGFP sequence was a bit more laborious (Figure 3). This was probably due to the differences in sequence size, pKT2-iresGFP has 5.3kb whereas pKT2-ARH-iresGFP has 11,2kb which facilitates a high number of copy insertions and also favors genome replication.[3] Despite the low number of cells obtained after the first separation, the cells survived and proliferated very well. After 6 weeks pure populations with a great number of cells were acquired. ARHGAP21 overexpression was highly efficient, even after over a month in culture a 23 times higher gene expression was verified by RT-PCR and a 27 time increase in protein levels was verified by WB (Figure 3B and C).

Sleeping Beauty Transposon System not only allowed us to finally stably overexpress ARHGAP21 in a cell line, but also promoted long-term expression. After the purification of the cells and confirmation of the protein overexpression, several vials of cells were frozen. Furthermore, ARHGAP21 protein overexpression was verified to remain present even after thawing once (Figure 4C and D) or twice (Figure 4E and F).

Finally, we were able to design an efficient protocol to clone and stably express a large cDNA sequence (ARHGAP21 5,87Kb) into human cell line.

Discussion

The ideal overexpression method should have high efficiency, low cell toxicity, minimal effects on normal physiology, and should be both easy to use and reproducible, providing a powerful tool to identify pathway components which may remain undetected using traditional loss of function analysis.

Transfection of a pCMV based plasmid containing ARHGAP21 protein full-length sequence into human cell line, initially seemed to be a very good option, simple, easy and efficient. For short time studies this might be a good choice, considering the increase of 300-500x in the levels of gene expression we obtained. However, when cells need to be expanded for longer than a month to reach the enough quantity to be transplanted in animal models, as in our study, definitely plasmid-based expression is not a good option. The longer the transfected cells remain in the culture, the more they will lose the transgene expression. Freezing the cells in the attempt to conserve the expression, in our hands, did not work either, considering that after thawing cells needed to be maintained in culture, for at least 10 days; in order for us to carry out the antibiotic selection and then expand them before transplantation. After this entire process the cells continued to demonstrate loss of expression. Even using a promoter derived from CMV, usually used to achieve high levels of expression in mammalian cells,[10] the ARHGAP21 expression was transient. Short-term expression of transgenes is one of the problems frequently associated with non-viral gene transfer.[59] The silencing of transgene expression from plasmid DNA might be due to complex intra-cellular events in which multiple factors are involved. Empirical studies have clearly demonstrated that the type of promoter, the bacterial backbone[60] and other epigenetic modifications,[61,62] explain, at least in part, the silencing of transgene expression from plasmid DNA. Finally, in the case plasmid vectors were inserted into the genome, the expression might be controlled differently from the expression from episomal plasmid vectors. Considering our extremely high levels of ARHGAP21 expression right after transfection, the complete loss of expression after tumor growth came as a major disappointment. Thus, we proceeded to a more stable transfection using lentivirus system.

Virus-mediated transduction is known to be highly efficient and easy to achieve sustainable transgene expression *in vivo* owing to the viral nature of integration into the host genome.[63] However, this method did not work for our purpose either. The
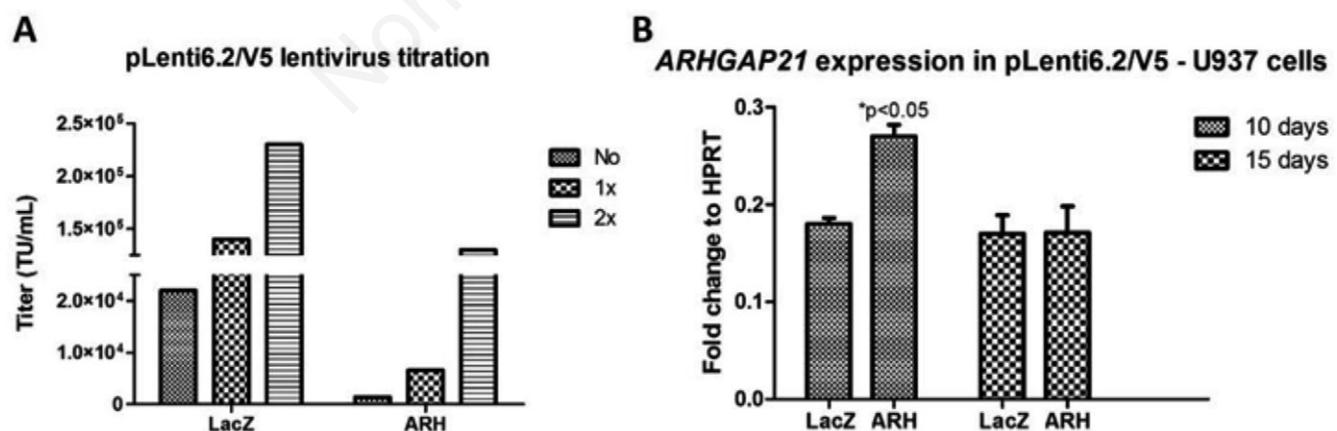

Figure 2. Efficacy of lentiviral based gene transfer. (A) Titration of lentivirus supernatant. The titration was performed in HT1080 cells with supernatant collected from 293FT cells transfected with pLenti6.2/V5-ARH or pLenti6.2/V5-LacZ and submitted to none, one or double ultracentrifugation; (B) ARHGAP21 gene expression in U937 cells transduced with lentivirus supernatant and Blasticidine selected over 10 or 15 days. ARH=sequence which encodes ARHGAP21 protein, co-transfected along with lentivirus packaging vectors to make a recombinant ARHGAP21 lentivirus; LacZ=sequence which encodes the enzyme β-galactosidase, co-transfected along with entivirus packaging vectors to make a recombinant control lentivirus. HPRT (hypoxanthine phosphoribosyl transferase)=internal control gene for quantitative real-time-PCR expression analysis.

ARHGAP21 lentivirus production was neither easy nor efficient as desired. ARHGAP21 gene has a ~5.9Kb ORF, in this way, the limited LV production yields obtained could be attributed to the difficult of packaging the ARHGAP21 transgene. The large size of full-length coding sequence of ARHGAP21 not only limited LV production, but also appears to impair gene transfer, and more importantly, disabled long-term gene expression. After transduction and antibiotic selection of the target cells we obtained an increase of only 50% in ARHGAP21 gene expression, which was completely lost after five days in culture. The virus-mediated method has been reported to present some drawbacks,[22-27] transduction disadvantages such as a virus package limited space for a foreign gene to keep infectivity,[22,64] inability to transfer large genes[23] and difficulty of vector production.[22,27] For these reasons, much effort has been put into developing non-viral transfection methods even though virus mediated transfection is highly effective and easy to use.[64]

In this scenario of failed attempts, the possibility of uses transposon-based transgene transfer came to our attention. Transposons present a number of benefits over classic viral or bacterial plasmid transgenesis. Stable transgenesis of human cells by transposons is ease of application, efficient and simple to implement.

Transposons are deployed in cells as plasmid DNA, and as such, the preparation of vectors requires basic techniques of molecular biology, thereby providing simplicity and safety to the user. Moreover, transposon vectors can be stored and distributed easily. Integration into the genome occurs in a large proportion of transfected cells, where transposon copy numbers may be scaled through control of the amount of transposon plasmid.[65] Importantly, the cargo capacity of transposons is massive compared to virus; BAC-sized elements have been shown to undergo full-length integration,[66] a major advantage over the unpredictable end resection or fragmentation seen with random BAC integration. Therefore, finally we were able to obtain a reasonable plasmid transfer efficiency in which we could generate a suitable population from which we purified GFP+ populations containing pKT2-iresGFP (EPY) and pKT2-ARH-iresGFP (ARH). Despite the low number of cells obtained after the first separation, the cells survived and proliferated very well without loss of ARHGAP21 expression. After 6 weeks pure populations with a great number of cells were acquired and ARHGAP21 overexpression was highly efficient. In addition to stably overexpressing ARHGAP21 in a human cell line, and promote long-term expression, SB transposon-based gene transfer conveniently enabled the expansion and
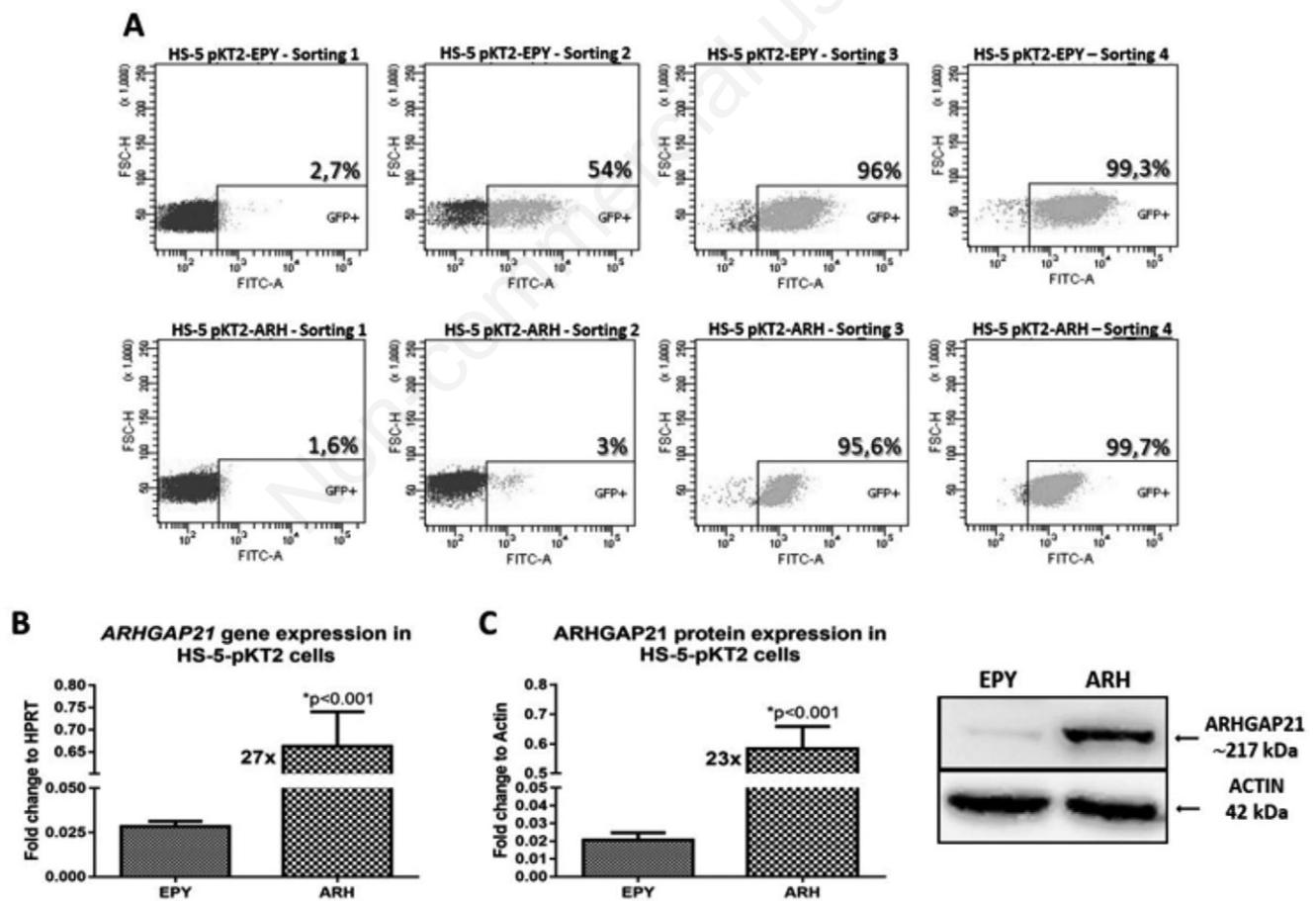

Figure 3. Efficacy of Sleeping Beauty transposon based gene transfer (A) Flow Cytometry showing a gradual increase of GFP+ cell populations during purification of HS-5 cells nucleofected with pKT2-iresGFP (pKT2-EPY=empty/control transposon) and pKT2-ARHGAP21-iresGFP (pKT2-ARH=ARHGAP21 expression transposon); (B) ARHGAP21 gene expression; (C) ARHGAP21 protein levels in GFP+ populations, after purification. ARH=ARHGAP21; EPY=empty; HPRT (hypoxanthine phosphoribosyl transferase)=internal control gene for quantitative real-time-PCR expression analysis; Actin=internal control for protein western blotting expression analysis.

freezing of the overexpressing cells. Thus, wherever we need to start an experiment we simply have to thaw the cells and expand them in culture.

Therefore, we finally were able to design an efficient protocol to clone and stably express ARHGAP21 cDNA sequence (a very large transgene 5,9Kb) into a human cell line, which is an efficient tool for further investigation of the role of this protein in hematopoiesis.

Conclusions

In conclusion, our study shows that the SB transposon system is the best choice for those who seek a stable and high gene expression. Our protocol is advantageous, as in addition to obtaining the gene overexpression, the freezing of cells for using them over a long period of time is possible.
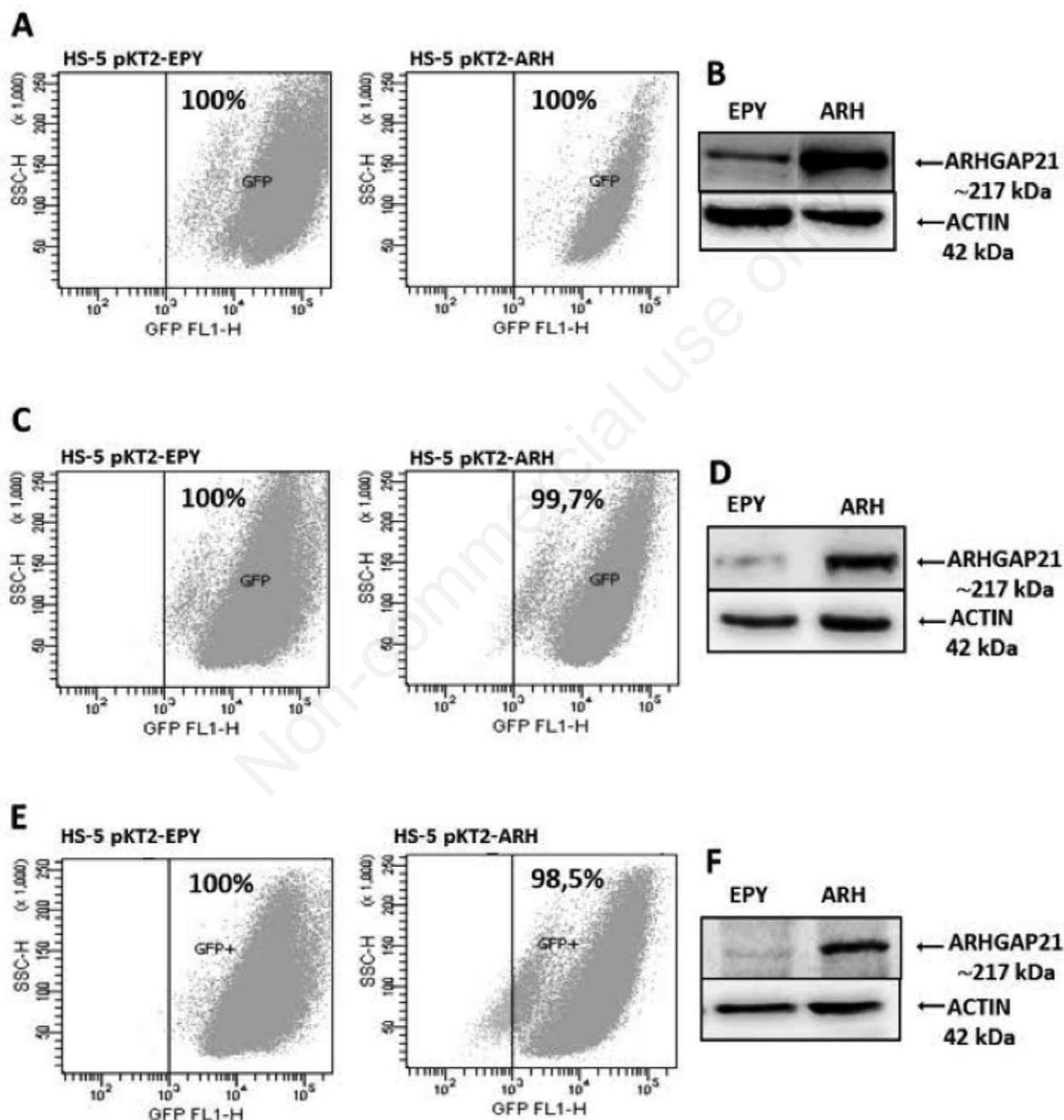

Figure 4. ARHGAP21/iresGFP expression in HS-5 cells nucleofected with pKT2-iresGFP (pKT2-EPY=empty/control transposon) and pKT2-ARHGAP21-iresGFP (pKT2-ARH=ARHGAP21 expression transposon), after 60 days in culture (A and B), after freezing once (C and D) and twice (E and F). The protein expression was verified by Flow Cytometry through GFP tag (A, C and E), and by Western Blotting (B, D and F). ARH=ARHGAP21; EPY=empty; Actin=internal control for protein western blotting expression analysis.
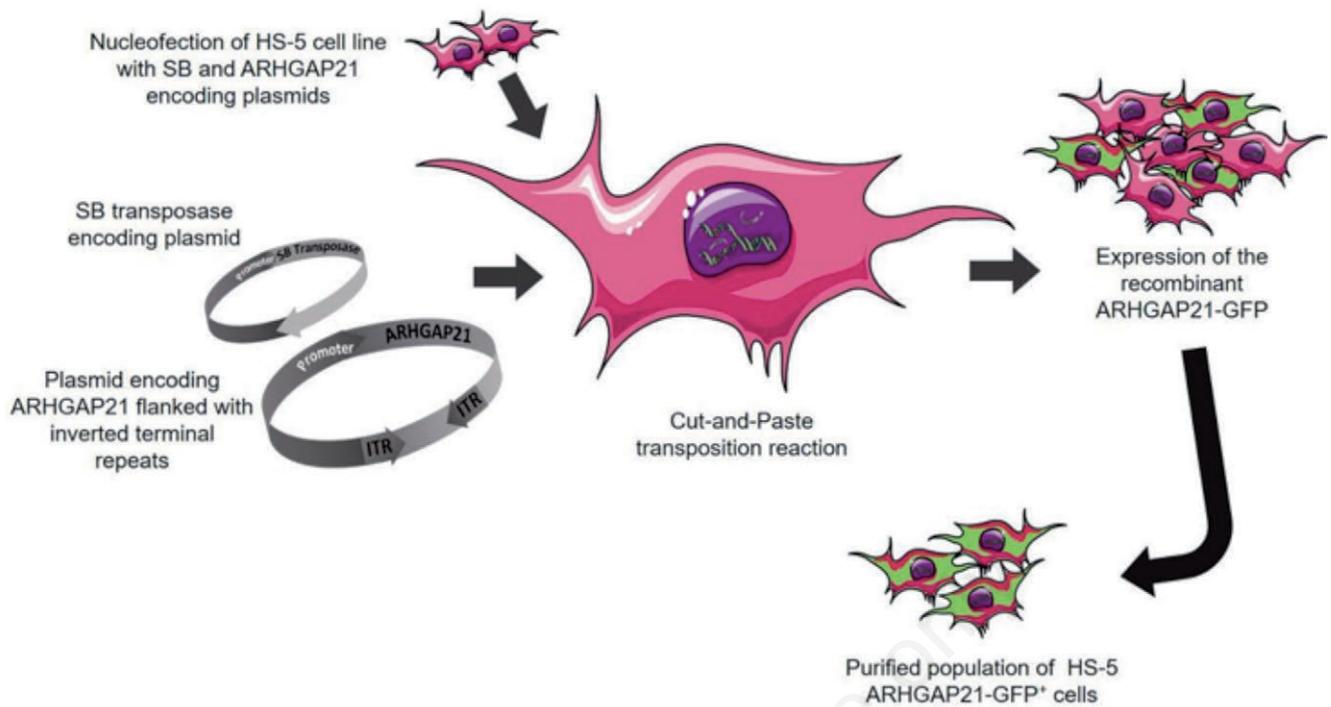
Figure 5. Schematic figure illustrating ARHGAP21 overexpression in HS-5 cell line protocol. Figure was produced using Servier Medical Art, http://www.servier.com/Smart/ImageBank.aspx?id=729

ARHGAP21 overexpression in HS-5 cell line protocol

As seen in Figure 5, for *each well* of a 16 well Nucleocuvette™ Strips [SF Cell Line 4D-Nucleofector® X Kit S (32 RCT)]:

Step 1: Solution+Plasmids

- 16,4ul Nucleofector Solution
- 3,6ul Supplement
- 0,5ug SB100x
- 2ug pKT2 (ARHGAP21 or EMPTY)
- Incubate 10-30 minutes at RT

Step 2: Cells

- Trypsinize and centrifuge 90g per 10 minutes
- Adjust to 3.2×10^5 cells
- Wash with PBS and centrifuge 90g per 10 minutes
- Keep pellets on ice

It is very important that the cellular confluence should not be above 80%, cells must be as healthy as possible before nucleofection.

Step 3: Nucleofection

- Ressuspend cell pellet with 20uL of Solutions+Plasmids
- Carefully pipet into each well
- Make sure that cells are very well homogenized, the volume should be exact 20uL and that no bubbles are present into the well
- At Amaxa™ 4D-Nucleofector™ select cell type program: HEK293 (CM 130)

Always run at least one well with each control: pmaxGFP™ Vector (0,4ug) and only cells to check the efficiency of nucleofection (Figure 6).

Step 4: Culture

- After nucleofection transfer cells into a 6-wells plate with 3 mL of pre-warmed DMEM+20% FBS per well (1 well for each cell group)
- Incubate for 24 h at 37 ºC and 5% CO_2
- After incubation aspirate media with dead cells
- Add 3 mL of fresh pre-warmed DMEM+20% FBS
- Culture cells until confluence reach 75%

Step 5: GFP+ cells purification

- Trypsinize cells and centrifuge 90g per 10 minutes
- Ressuspend cell pellet in 300 uL-800uL of PBS+0,5% BSA+2mM EDTA+0,25 ug/mL propidium iodide (PI)
- Cell debris and dead cells are excluded from the analysis based on scatter signals and PI fluorescence.
- The high GFP+ population must be sorted

Serial sortings must be performed until GFP+ population has reached 95-100%. Furthermore, samples must be collected and ARHGAP21 expression analyzed by RT-PCR and WB. After con-

Figure 6. Nucleofection microwells scheme.

firmation of ARHGAP21 overexpression, cells must be expanded and then frozen.

Step 6: Freezing cells

- Trypsinize cells and centrifuge 90g per 10 minutes
- Freeze $5\text{-}10\times10^6$ in 1 mL of DMEM+20% FBS+10% Dimethyl sulfoxide (DMSO)
- For short time preservation maintain cells at -80°C
- For long time preservation maintain cells in liquid nitrogen

References

1. Kelly TJ Jr, Smith HO. A restriction enzyme from Hemophilusinfluenzae. II. J Mol Biol 1970;51:393-409.
2. Jackson DA, Symons RH, Berg P. Biochemical methods for inserting new genetic information into DNA of Simian Virus 40, Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proc Natl AcadSci 1972;69:2904-9.
3. Prelich G. Gene overexpression, uses, mechanisms and interpretation. Genetics 2012;190:841-54.
4. Govindan G, Ramalingam S. Programmable site-specific nucleases for targeted genome engineering in higher eukaryotes. J Cell Physiol 2016;231:2380-92.
5. Hotta A, Yamanaka S. From genomics to gene therapy, induced pluripotent stem cells meet genome editing. Annu Rev Genet 2015;49:47-70.
6. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 2014;346:1258096.
7. Pellagatti A, Dolatshad H, Valletta S, Boultwood J. Application of CRISPR/Cas9 genome editing to the study and treatment of disease. Arch Toxicol 2015;89:1023-34.
8. Boudreaux SP, Ramirez-Herrick AM, Duren RP, Connely OM. Genome-wide profiling reveals transcriptional repression of MYC as a core component of NR4A tumor suppression in acute myeloid leukemia. Oncogenesis 2012;1:e19.
9. Filareto A, Parker S, Darabi R, et al. An ex vivo gene therapy approach to treat muscular dystrophy using inducible pluripotent stem cells. Nat Commun 2013;4:1549.
10. Gill DR, Pringle IA, Hyde SC. Progress and prospects, the design and production of plasmid vectors. Gene Ther 2009;16:165-71.
11. Gibson TJ, Seiler M, Veitia RA. The transience of transient overexpression. Nat Methods 2013;10:715-21.
12. Liao MML, Sunstrom NA. A transient expression vector for recombinant protein production in Chinese hamster ovary cells. J Chem Technol Biotechnol 2006;81:82-8.
13. Xu ZL, Mizuguchi H, Ishii-Watabe A, et al. Optimization of transcriptional regulatory elements for constructing plasmid vectors. Gene 2001;272:149-56.
14. Kim SY, Lee J, Shin H, et al. The human elongation factor 1 alpha (EF-1a) first intron highly enhances expression of foreign genes from the murine cytomegalovirus promoter. J Biotechnol 2002;93:183-7.
15. Mariati SCL, Ho MGS, Yap YY. Evaluating post-transcriptional regulatory elements for enhancing transient gene expression levels in CHO K1 and HEK293 cells. Protein Expr Purif 2010;69: 9-15.
16. Mufarrege EF, Antuña S, Etcheverrigaray M, et al. Development of lentiviral vectors for transient and stable protein overexpression in mammalian cells. A new strategy for recombinant human FVIII (rhFVIII) production. Protein Expr Purif 2014;95:50-6.
17. Friehs K. Plasmid copy number and plasmid stability. Adv Biochem Eng Biotechnol 2004;86:47-82.
18. Rozkov A, Avignone-Rossa CA, Ertl PF, et al. Characterization of the metabolic burden on Escherichia coli DH1 cells imposed by the presence of a plasmid containing a gene therapy sequence. Biotechnol Bioeng 2004;88:909-15.
19. Cleary JM, Rosen I, Harding NE, Cabanas DK: Recombinant DNA plasmid for xanthan gum synthesis. Merck 1987:EP0233019.
20. Wu Y, Hao GM, He J, et al. Lentivirus mediated over expression of CGRP inhibited oxidative stress in Schwann cell line. Neurosci Lett 2015;598:52-8.
21. Gómez-Mauricio G, Moscoso I, Martín-Cancho MF, et al. Combined administration of mesenchymal stem cells overexpressing IGF-1 and HGF enhances neovascularization but moderately improves cardiac regeneration in a porcine model. Stem Cell Res Ther 2016;7:94.
22. Braun S. Muscular gene transfer using nonviral vectors. Curr Gene Ther 2008;8:391-405.
23. Wang W, Li W, Steinhoff G. Non-viral gene delivery methods. Curr Pharm Biotechnol 2013;14:46-60.
24. Brown HC, Wright JF, Zhou S, et al. Bioengineered coagulation factor VIII enable long-term correction of murine hemophilia A following liver-direct adeno-associated viral vector delivery. Mol Ther Methods Clin Dev 2014;6:14036.
25. Wang N, Zhang H, Zhang BQ, et al. Adenovirus-mediated efficient gene transfer into cultured three-dimensional organoids. PLoS One 2014;9:e93608.
26. Sheridan C. Gene therapy finds its niche. Nat Biotechnol 2001;29:121-8.
27. Yin H, Kanasty RL, Eltoukhy AA, et al. Non-viral vectors for gene-based therapy. Nat Rev Genet 2014;15:541-55.
28. Ehrhardt A, Xu H, Huang Z, et al. A direct comparison of two nonviral gene therapy vectors for somatic integration, in vivo evaluation of the bacteriophage integrase phiC31 and the Sleeping Beauty transposase. Mol Ther 2005;11:695-706.
29. Aronovich EL, McIvor RS and Hackett PB. The Sleeping Beauty transposon system, a non-viral vector for gene therapy. Hum Mol Genet 2011;20:R14-R20.
30. Ivics Z, Li MA, Mátés L, et al. Transposon-mediated genome manipulation in vertebrates. Nat Methods 2009;6:415-22.
31. Zayed H, Izsvak Z, Walisko O, Ivics Z. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Mol Ther 2004;9:292-304.
32. Lee S, Moon JH, Song K, et al. Establishment of transgenic porcine fibroblasts expressing a human klotho gene and its effects on gene expression and preimplantation development of cloned embryos. DNA Cell Biol 2017;36:42-9.
33. Woltjen K, Michael IP, Mohseni P, et al. Piggybac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 2009;458:766-70.
34. Yusa K, Rad R, Takeda J, Bradley A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat Methods 2009;6:363-9.
35. Kim SI, Oceguera-Yanez F, Sakurai C, et al. Inducible Transgene Expression in Human iPS Cells Using Versatile All-in-One piggyBac Transposons. Methods Mol Biol 2016;1357:111-31.
36. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish,

and its transposition in human cells. Cell 1997;91:501-10.

37. Mates L, Chuah MK, Belay E, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet 2009;41:753-61.

38. Grabundzija I, Irgang M, Mátés L, et al. Comparative analysis of transposable element vector systems in human cells. Mol Ther 2010;18:1200-9.

39. Ammar I, Gogol-Döring A, Miskey C, et al. Retargeting transposon insertions by the adeno-associated virus Rep protein. Nucl Acids Res 2012;40:6693-712.

40. Cadinanos J, Bradley A. Generation of an inducible and optimized piggyBac transposon system. Nucleic Acids Res 2007;35:e87.

41. Moldt B, Yant SR, Andersen PR, et al. Cis- acting gene regulatory activities in the terminal regions of sleeping beauty DNA transposon-based vectors. Hum Gene Ther 2007;18:1193-204.

42. Walisko O, Schorn A, Rolfs F, et al. Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. Mol Ther 2008;16:359-69.

43. Newman JC, Bailey AD, Fan HY, et al. An abundant evolutionarily conserved CSB-PiggyBac fusion protein expressed in Cockayne syndrome. PLoS Genet 2008;4:e1000031.

44. Bassères DS, Tizzie EV, Duarte AA, et al. ARHGAP10, a novel human gene coding for a potentially cytoskeletal Rho-GTPase activating protein. BiochemBiophys Res Commun 2002;294:579-85.

45. Nourry C, Grant SG, Borg JP. PDZ domain proteins plug and play! Sci STKE 2003;179:RE7.

46. Kandpal RP. Rho GTPase activating proteins in cancer phenotypes. Curr Protein Pept Sci 2006;7:355-65.

47. Bigarella CL, Borges L, Costa FF, Saad ST. ARHGAP21 modulates FAK activity and impairs glioblastoma cell migration. Biochim Biophys Acta 2009;1793:806-16.

48. Borges L, Bigarella CL, Baratti MO, et al. ARHGAP21 associates with FAK and PKCzeta and is redistributed after cardiac pressure overload. Biochem Biophys Res Commun 2008;374:641-6.

49. Sousa S, Cabanes D, Archambaud C, et al. ARHGAP10 is necessary for alpha-catenin recruitment at adherens junctions and for Listeria invasion. Nat Cell Biol 2005;7:954-60.

50. Anthony DF, Sin YY, Vadrevu S, et al. β-Arrestin 1 inhibits the GTPase-activating protein function of ARHGAP21, promoting activation of RhoA following angiotensin II type 1A receptor stimulation. Mol Cell Biol 2011;31:1066-77.

51. Schultz ML, Tecedor L, Stein CS, et al. CLN3 deficient cells display defects in the ARF1-Cdc42 pathway and actin-dependent events. PLoS One 2014;9:e96647.

52. Wang S, Li H, Chen Y, et al. Transport of influenza virus neuraminidase (NA) to host cell surface is regulated by ARHGAP21 and Cdc42 proteins. J Biol Chem 2012;287:9804-16.

53. Hehnly H, Longhini KM, Chen JL, Stamnes M. Retrograde Shiga toxin trafficking is regulated by ARHGAP21 and Cdc42. Mol Biol Cell 2009;20:4303-12.

54. Mènètrey J, Perderiset M. Structural basis for ARF1-mediated recruitment of ARHGAP21 to Golgi membranes. EMBO J 2007;26:1953-62.

55. Dubois T, Chavrier P. ARHGAP10, a novel RhoGAP at the cross-road between ARF1 and Cdc42 pathways, regulates Arp2/3 complex and actin dynamics on Golgi membranes. Med Sci 2005;21:692-4.

56. Barcellos KS, Bigarella CL, Wagner MV, et al. ARHGAP21 protein, a new partner of α-tubukin involved in cell-cell adhesion formation and essential for epithelial-mesenchymal transition. J Biol Chem 2013;288:2179-89.

57. Lazarini M, Traina F, Machado-Neto JA, et al. ARHGAP21 is a RhoGAP for RhoA and RhoC with a role in proliferation and migration of prostate adenocarcinoma cells. Biochim Biophys Acta 2013;1832:365-75.

58. Ferreira SM, Santos GJ, Rezende LF, Gonçalves LM. ARHGAP21 prevents abnormal insulin release through actin rearrangement in pancreatic islets from neonatal mice. Life Sci 2015;127:53-8.

59. Zang L, Nishikawa M, Ando M, Takahashi Y, Takakura Y. Contribution of Epigenetic Modifications to the Decline in Transgene Expression from Plasmid DNA in Mouse Liver. Pharmaceutics 2015;7:199-212.

60. Chen ZY, He CY, Meuse L, Kay MA. Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. Gene Ther 2004;11:856-64.

61. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C and Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol 2004;6:73-7.

62. Okitsu CY, Hsieh CL. DNA methylation dictates histone H3K4 methylation. Mol Cell Biol 2007;27:2746-57.

63. Elsner C, Bohne J. The retroviral vector family: something for everyone. Virus Genes 2017;53:714-22.

64. Kyung Kim T, Eberwine JH. Mammalian cell transfection: the present and the future. Anal Bioanal Chem 2010;397:3173-8.

65. Wang W, Lin C, Lu D, et al. Chromosomal transposition of piggyBac in mouse embryonic stem cells. Proc Natl Acad Sci USA 2008;105:9290-5.

66. Rostovskaya M, Fu J, Obst M, et al. Transposon-mediated BAC transgenesis in human ES cells. Nucleic Acids 2012;40:e150.