

## Adeno-associated virus Rep-mediated targeting of integrase-defective retroviral vector DNA circles into human chromosome 19

Huang, Shuohao

Graduate School of Systems Life Sciences, Kyushu University

Kawabe, Yoshinori

Department of Chemical Engineering, Faculty of Engineering, Kyushu University

Ito, Akira

Department of Chemical Engineering, Faculty of Engineering, Kyushu University

Kamihira, Masamichi

Graduate School of Systems Life Sciences, Kyushu University | Department of Chemical Engineering, Faculty of Engineering, Kyushu University

<https://hdl.handle.net/2324/27290>

---

出版情報 : Biochemical and Biophysical Research Communications. 417 (1), pp.78-83, 2012-01-06.  
Elsevier

バージョン :

権利関係 : (C) 2011 Elsevier Inc.



# **Adeno-associated virus Rep-mediated targeting of integrase-defective retroviral vector DNA circles into human chromosome 19**

Shuohao Huang<sup>a</sup>, Yoshinori Kawabe<sup>b</sup>, Akira Ito<sup>b</sup>, Masamichi Kamihira<sup>a,b,\*</sup>

<sup>a</sup>*Graduate School of Systems Life Sciences, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan*

<sup>b</sup>*Department of Chemical Engineering, Faculty of Engineering, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan*

\*Corresponding author. Address: Department of Chemical Engineering, Faculty of Engineering, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.  
Fax: +81 (0)92 802 2793.

*E-mail address:* kamihira@chem-eng.kyushu-u.ac.jp (M. Kamihira).

**Abbreviations:** adeno-associated virus, AAV; AAV integration site, AAVS1; enhanced green fluorescent protein, EGFP; fetal bovine serum, FBS; integrase-defective retroviral vector, IDRV; long terminal repeat, LTR; p5 integration efficiency element, p5IEE; Rep-binding sites, RBS; AAVS1 terminal resolution site, TRS.

## **ABSTRACT**

Retroviral vectors have been employed in clinical trials for gene therapy owing to their relative large packaging capacity, alterable cell tropism, and chromosomal integration for stable transgene expression. However, uncontrollable integrations of transgenes are likely to cause safety issues, such as insertional mutagenesis. A targeted transgene integration system for retroviral vectors, therefore, is a straightforward way to address the insertional mutagenesis issue. Adeno-associated virus (AAV) is the only known virus capable of targeted integration in human cells. In the presence of AAV Rep proteins, plasmids possessing the p5 integration efficiency element (p5IEE) can be integrated into the AAV integration site (AAVS1) in the human genome. In this report, we describe a system that can target the circular DNA derived from non-integrating retroviral vectors to the AAVS1 site by utilizing the Rep/p5IEE integration mechanism. Our results showed that after G418 selection 30% of collected clones had retroviral DNA targeted at the AAVS1 site.

**Keywords:** integrase-defective retroviral vectors; adeno-associated virus (AAV); Rep; AAVS1; targeted integration

## 1. Introduction

Gammaretrovirus vectors effectively insert foreign genes into the genome of dividing cells. This integration is mediated by the retroviral integrase enzyme encoded by the viral *pol* gene. Integrase has three clearly identifiable domains: (i) the  $\text{Zn}^{2+}$  binding N-terminus containing an HHCC motif; (ii) the central catalytic domain with a DDE motif; and (iii) the host DNA binding C-terminus [1]. Type 1 mutations of integrase can disrupt the integration process without severely disturbing other steps of retroviral replication [1]. Integrase-defective retroviral vectors (IDRV) from type 1 mutation, for example, the D184A mutation in Moloney murine leukemia virus, deliver transgenes efficiently, and the linear retroviral cDNA can be turned into circular forms in the cell nucleus either by non-homologous joining or homologous recombination [2]. IDRV circles have been proven useful for many applications, especially for transient gene expression [2].

Adeno-associated virus (AAV) is a small single-strand DNA virus that infects human cells without causing a severe immune response. More importantly, AAV is the only known virus that can target its genome to a specific human chromosome locus designated AAVS1 (19q13.3-qter) [3,4]. Both of these characters help AAV attract much attention in the gene therapy field. Although the underlying mechanism of targeted integration remains poorly understood, several key components have been identified. Rep proteins (Rep68/78) are responsible for binding of the AAV genome to the AAVS1 sequence, nicking AAVS1 terminal resolution site (TRS), and inducing the replication-associated integration. Another component is the Rep-binding sites (RBS) located in both the AAV genome and AAVS1 [5,6].

We previously demonstrated that IDRV circles could be used in site-specific genomic modification via the Cre/*loxP* reaction [7]. A prerequisite step in that strategy is to build founder cells possessing *loxP* sites. In this study, without founder cells, we attempted to target IDRV to the AAVS1 site by employing the Rep/p5IEE system. The AAV p5IEE cis-element has been shown not only necessary but also sufficient for Rep mediated AAVS1 targeting [8]. Here, we show successful targeting of IDRV circles to the AAVS1 site, which extends the usage of retroviral vectors.

## 2. Materials and methods

### 2.1. Plasmid construction

The *rep* gene fragment was amplified from the plasmid, pBluescript SKII+AV1-235-4491 (Riken BioResource Center, Tsukuba, Japan) using the primers, RepFW, 5'-ATG CCG GGG TTT TAC GAG AT-3', and RepRV, 5'-TTG AGC TTC CAC CAC TGT CTT ATT-3'. The amplified PCR product was inserted into the *EcoRV* site of pBluescriptII KS(-) (Toyobo, Osaka, Japan) to generate the plasmid pBlue/Rep. The *rep* fragment from pBlue/Rep after digestion with *SmaI* and *ClaI* was ligated into a murine stem cell virus (MSCV)-based retroviral vector plasmid, pQMSCV/TAANeoGreen- [7] after digestion with *EcoRI* (blunt-ended) and *ClaI* to generate pQMSCV/Rep.

A neomycin resistance gene (*Neo*) tagged with an ATG-*loxP* sequence was amplified from the plasmid, pIRES2-DsRed-Express (Clontech, Palo Alto, CA) using the primers, 5'-CCC GCG GCC GCA TGA TAA CTT CGT ATA GCA TAC ATT ATA CGA AGT TAT TGA TTG AAC AAG ATG GAT TGC AC-3' and 5'-CGG GAT CCT CAG AAG AAC TCG TCA AGA AGG-3', in which *NotI* and *BamHI* sites (underlined) were appended onto both ends of the PCR product (wild-type *loxP* sequence is double

underlined). The PCR product was digested with *NotI* and *BamHI*, and ligated into *NotI*- and *BamHI*-digested pBlue/TAANeoEGFP [7] to generate pBlue/NIG. The *SacI*-*Neo*-IRES-EGFP-*XhoI* fragment from pBlue/NIG was ligated into the blunt-ended pQMSCV/NΔAβ [7] after digestion with *XhoI* to generate pQMSCV/NIG. An HSV-derived thymidine kinase (*TK*) fragment was amplified from the plasmid, OSDupDel.Hprt (Open Biosystems, Huntsville, AL) using the primers, TKFW, 5'-CCG CGG ATG GCT TCG TAC CCC TGC-3', and TKRV, 5'-CCG CGG TCA GTT AGC CTC CCC CAT CT-3' to append *SacII* sites (underlined) onto both ends. The PCR product was ligated into *SacII*-digested pQMSCV/NIG to generate pQMSCV/TNIG. A p5IEE sequence was amplified from the plasmid, pBluescript SKII+AV1-235-4491 using the primers, 5'-GTC CTG TAT TAG AGG TCA CGT GAG TGT TTT GCG ACA TTT TGC GAC ACC ATG TGG TCA CGC T-3' and 5'-CAA AAT GGA GAC CCT GCG-3', and ligated into the *EcoRV*-digested pBluescript (Toyobo) to generate pBlue/p5IEE. The *EcoRI*-p5IEE-*XhoI* fragment from pBlue/p5IEE was blunt-ended, and then ligated into the blunt-ended pQMSCV/TNIG after digestion with *EcoRI* to generate pQMSCV/PTNIG.

## 2.2. Cell culture, retroviral vector production and titration

As previously described [7], 293FT cells (Invitrogen, Carlsbad, CA) were cultured. HeLa cells were cultured at 37 °C in a 5% (v/v) CO<sub>2</sub> incubator and maintained in Dulbecco's modified Eagle's medium (Sigma, St. Louis, MO) supplemented with sodium bicarbonate (3.7 mg/ml), 10% FBS (Biowest, Nuaille, France), 100 U/ml penicillin G potassium, and 100 µg/ml streptomycin sulfate (Wako Pure Chemical

Industries, Osaka, Japan). To select cells with transgene integration, G418 (Sigma) was added to the medium at a final concentration of 1 mg/ml.

Preparation and quantitative PCR-based titration of conventional retroviral vectors and IDRV were performed as previously described [7]. The infectious titers for MSCV/PTNIG and IDRV/PTNIG were determined by scoring green-fluorescing cells using flow cytometry (BD FACSCalibur™ Flow Cytometer; BD Bioscience, San Jose, CA).

### *2.3. Targeting plasmid and IDRV to the AAVS1 site*

For plasmid targeting, 400 ng of pQMSCV/PTNIG with or without 25 ng of pQMSCV/Rep was transfected into 90%-confluent HeLa cells in the wells of 24-well plates using a lipofection reagent (Lipofectamine 2000; Invitrogen, Carlsbad, CA). Cells were then re-seeded 24 h post-transfection at a density of  $5 \times 10^4$  cells/well into six-well plates and then screened using medium with the addition of G418.

For IDRV targeting, IDRV/PTNIG ( $1.6 \times 10^{10}$  copies by quantitative-PCR) with or without IDRV/Rep ( $0.7 \times 10^{10}$  copies) was used to infect HeLa cells at a density of  $5 \times 10^4$  cells/well in 24-well plates. Two days after infection, G418 screening was initiated by re-seeding cells at a density of  $2 \times 10^5$  cells/well into six-well plates.

### *2.4. Southern blotting, genomic PCR and sequencing*

Southern blotting, genomic PCR, and genetic sequencing were performed, as previously described [7]. Briefly, AAVS1 and Neo probes were prepared by PCR using the primers, 5'-GAA CTC TGC CCT CTA ACG CTG C-3' and 5'-CAC CAG ATA AGG AAT CTG CC-3' for AAVS1 and 5'-CCG GAT CCA TGA TTG AAC AAG ATG GAT

TGC-3' and 5'-CCG GAT CCT CAG AAG AAC TCG TCA AGA AGG-3' for Neo. Genomic DNA (10 µg) digested with *EcoRI* was electrophoresed on a 0.7% (w/v) agarose gel. Genomic PCR covering possible junctions was performed using nested primer pairs (J1, 5'-GCA GAA GCC AGT AGA GCT CA-3' and J2, 5'-GCG CAA AGT GAC AAT GGC C-3' in the AAVS1 region; P1, 5'-GGT TCT GGT AGG AGA CGA GA-3' and P2, 5'-CCC CTT GAA CCT CCT CGT T-3' in the packaging single region; N, 5'-GCA TCA GAG CAG CCG ATT GT-3' and TKRV, in the transgene region). Long terminal repeat (LTR) structures were determined using primers: E, 5'-AAC GAG AAG CGC GAT CAC-3' and PRV, 5'-CTT CTT GAC ATC TAC CGA CTG G-3'. *TK* was detected using primers TKFW and TKRV. IDR/Rep integration was detected using primers RepFW and RepRV. *GAPDH* amplification was performed using primers: Fw, 5'-AGC CAC ATC GCT CAG ACA C-3' and Rv, 5'-ATC GCC CCA CTT GAT TTT G-3'.

### 3. Results

#### 3.1. Construct design and targeted integration of plasmid

In the previous studies, an ATG-complementation strategy was introduced into the viral vector constructs to facilitate the collection of cells undergoing expected recombination [7,9]. The strategy is based on the insertion of an ATG-deficient selection marker gene fused with a recombinase target site (*loxP* or *FRT*) into the viral vector, whose functional expression unit can be reconstituted after the recombinase (Cre or Flp)-mediated site-directed insertion of viral genome into the target site harboring the ATG start codon. A problem inherent in such a system is the inability to analyze non-targeted random integration, since clones derived from auto-integration usually have no



selection marker gene expression to survive the screening process using drugs. However, the ATG-complementation strategy eased the finding of expected clones and saved the labor required for the genomic characterization of transgene. In addition and most importantly, it proved the feasibility of using non-integrating gammaretroviral or lentiviral vectors as substrates for the recombinase-mediated targeting. With the same intention, we constructed a vector MSCV/PTNIG (Fig. 1A), where an out-of-frame *TK* gene with its stopping codon was inserted between p5IEE and *Neo* to inhibit the proper translation of *Neo*. Similar strategies have been used in Cre/*loxP* recombination mediated gene activation, where a floxed gene cassette was placed upstream of a second gene to inhibit the expression of the second gene [10]. We also constructed a Rep protein expression vector, MSCV/Rep (Fig. 1A). To test Rep-mediated targeting to AAVS1, we first transfected pQMSCV/PTNIG into HeLa cells, with or without pQMSCV/Rep. After G418 selection, no colonies formed in the absence of pQMSCV/Rep (Fig. 1B). We went on to analyze nine randomly picked clones in the co-transfection group using Southern blotting for the co-migration of AAVS1 and Neo probes, an indicator of targeted integration at AAVS1 [11], and identified five clones with targeted integration (data not shown). These results suggested that pQMSCV/PTNIG construct could eliminate cells with Rep-independent integration. The addition of Rep promoted *Neo* expression in HeLa cells and enabled them to obtain G418 resistance. We reasoned from these results that the construct of MSCV/PTNIG should be useful in the IDRV targeting experiments.

### 3.2. Targeted integration of circular IDRV DNA

IDRV/PTNIG and IDRV/Rep were produced and used to infect HeLa cells. We measured viral titers for conventional (MSCV/PTNIG) and mutant (IDRV/PTNIG) vectors. Due to the IRES-*EGFP* fragment, virus titer could be deduced from the number of EGFP-positive cells on day 2 post-infection, which were  $2 \times 10^4$  and  $4 \times 10^3$  IU/ml for the conventional and mutant vectors, respectively. By quantifying the retroviral RNA based on the standard curve calculated from the RNA control template in the titration kit, virus RNA numbers were determined as  $1.9 \times 10^{10}$  and  $3.2 \times 10^{10}$  copies/ml for the conventional and the mutant, respectively.

After IDRV infection and screening, a 10-fold increase of colony number was observed for the co-infection group (Fig. 1C). After three rounds of experiments, 37 total clones (representing one third of all the G418 resistant clones obtained in the combined three experiments) were analyzed by Southern blot, eleven of which exhibited targeted integration (Fig. 2A). Nested PCR was performed on the eleven clones to understand the IDRV circle-AAVS1 junctions and six clones gave amplified products (Fig. 3). Sequencing results suggested that the breakpoints on IDRV were scattered within p5IEE, and the breakpoints on the AAVS1 locus were near the TRS-RBS site. Two breakpoints on AAVS1 appeared 170 bp and 394 bp downstream of the RBS (Fig. 3A, B). For clones 3-6, we observed the insertion of unknown DNA of variable lengths (Fig. 3C). We also checked the LTR and *TK* structures for the six clones and found that clones 1, 3, 5, and 6 possessed only 2-LTR circles. Clone 4 possessed two types of circles. For clone 2, no circular DNA was detected (Fig. 2B). We also went on to check the auto-integration of IDRV/Rep by genomic PCR, and found no IDRV/Rep integration for all the six clones (Fig. 2B). Therefore, LTR structures were solely derived from IDRV/PTNIG. Five of the six integrants were derived from 2-LTR

circular DNA. Yu *et al.* observed the prevalence of 2-LTR circles over 1-LTR circles for gammaretroviral vectors [12], while 1-LTR forms are more prevalent for lentiviral vectors [2,13]. The formation of 1-LTR and 2-LTR circles depends on different cell repair pathways, and is cell cycle related [13]. Since we only analyzed six transformants, this result might not reflect the true prevalence of LTR-circles. As for *TK*, clones 2, 4, 5, and 6 showed the presence of an intact *TK* gene (Fig. 2B).

Clones 1 and 3 had single copy 2-LTR-circle integration at AAVS1 with disrupted *TK*. Clones 2, 4, and 6 had two copies of transgene integration, as indicated by two Neo bands in the Southern blot. Clone 2 and 6 had one of the IDRVs integrated at AAVS1, while for clone 4 the two transgenes were both targeted at AAVS1. PCR did not detect LTR circles for clone 2, which could be explained as Rep-mediated linear IDRV integration or a disruption at IDRV LTR regions. PCR suggested that for clone 4 the two IDRVs at AAVS1 assumed two types of LTR formations. For clone 6, the IDRV at AAVS1 could either be 2-LTR circular or linear. Genomic PCR could only prove the existence of 2-LTR circle in the genome, and inserted linear IDRV gave no amplified product. Therefore, the two IDRVs had equal chances of taking the 2-LTR circular formation, regardless of the genomic position. For clones 2, 4 and 6, at least one of the integrated IDRVs contained an intact *TK*. Clone 5 had single copy 2-LTR-IDRV integration at AAVS1, and this IDRV possessed intact *TK*.

#### **4. Discussion**

The AAV integration mechanism has been employed in the site-specific integration of plasmids [14], Herpes simplex virus vector [15], and hybrid adenovirus/adeno-associated virus vector [16]. In this study, we evaluated if the same mechanism could

direct IDR V circles to the AAVS1 site. Transiently expressed Rep can recognize the p5IEE on IDR V circles and perform the integration reaction. IDR V circles can be inserted into a pre-determined site via recombinase [7,9]. However, recombinase-mediated targeting requires target sites. IDR Vs can also be directly targeted to a genomic site by homologous recombination (HR) [17]. However, the HR-based strategy demands beforehand knowledge of the target site sequences and the incorporation of homologous sequences into the viral vector. The system described here demands neither founder cells nor understanding of target site genomic DNA sequences.

In this study, we constructed a transgene-delivering vector MSCV/PTNIG, where *Neo* expression was expected to be prohibited by *TK*. Both the plasmid transfection and IDR V infection experiments suggested that the sole presence of transgene vector in infected cells could not effectively provide them with G418 resistance (Fig. 1B and C). Abnormal proviral structures have been found after residual integration of integrase-mutated MLV retrovirus [18], and modification of vector sequences was also found for integrase defective lentiviral vectors [13]. None of these modifications seemed very effective to induce *Neo* expression in MSCV/PTNIG. Although it remains to be determined as to the exact nature of Rep-induced turning on of *Neo*, the accessing of Rep to the p5IEE in IDR V led to efficient IDR V modification. After IDR V insertion, the existing promoter in IDR V or a genomic promoter initiated *Neo* transcription. Regardless of the different possible pathways that promote *Neo* expression, the participation of Rep played a positive role in obtaining cell clones with Rep-mediated IDR V integration.

We also found that the remaining 70% of analyzed clones were not targeted at AAVS1. In previous studies, using the same p5IEE and positive selection, 30% [19] to

48.4% [20] clones had AAVS1 targeting. Howden *et al.* found the remaining clones having random integration [19], while Yue *et al.* suspected Rep-induced integration at non-AAVS1 sites. Recently, novel hot spots such as AAVS2 and AAVS3 were identified for wild type AAV-2 [21]. We suspect that some of the off-AAVS1 transformants in this study might be Rep-mediated insertion at non-AAVS1 sites.

Our ultimate goal is to develop a safer gammaretroviral gene delivery system. Here, we created a new system where IDRV gene delivery and Rep/p5IEE gene integration were combined. Data produced by our laboratory describing junction sequences suggested that IDRV circle insertion via Rep shared some similarities with plasmid [11], although the lower targeted integration efficiency (Table 1) and the appearance of viable colonies following infection in the absence of Rep (Fig. 1C) suggested Rep-mediated IDRV integration was more complicated than plasmid targeting. At this stage, the detailed mechanism of Rep mediated targeting at AAVS1 still remains elusive [3]. This fact makes it even more difficult for us to have a clear view about the mechanism behind Rep-induced IDRV modification and integration. Nevertheless, Southern blot analysis demonstrated that after IDRV targeting by Rep, one allele of AAVS1 remained intact (Fig. 2A). Targeting IDRV to the AAVS1, a “safe harbor” for gene integration [22], should ameliorate safety concerns for gammaretroviral vectors. This study expands the application of IDRV and paves the way for a better retroviral vector-based gene delivery system.

## **Acknowledgments**

S. Huang is a recipient of the Japanese Government Scholarship from the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

## References

- [1] O. Delelis, K. Carayon, A. Saib, E. Deprez, J.F. Mouscadet, Integrase and integration: biochemical activities of HIV-1 integrase, *Retrovirology* 5 (2008) 114.
- [2] M.B. Banasik, P.B. McCray, Jr., Integrase-defective lentiviral vectors: progress and applications, *Gene Ther.* 17 (2010) 150-157.
- [3] E. Henckaerts, R.M. Linden, Adeno-associated virus: a key to the human genome?, *Future Virol.* 5 (2010) 555-574.
- [4] R.M. Kotin, M. Siniscalco, R.J. Samulski, X.D. Zhu, L. Hunter, C.A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, K.I. Berns, Site-specific integration by adeno-associated virus, *Proc. Natl. Acad. Sci. U. S. A.* 87 (1990) 2211-2215.
- [5] R.H. Smith, Adeno-associated virus integration: virus versus vector, *Gene Ther.* 15 (2008) 817-822.
- [6] R.M. Linden, E. Winocour, K.I. Berns, The recombination signals for adeno-associated virus site-specific integration, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 7966-7972.
- [7] S. Huang, Y. Kawabe, A. Ito, M. Kamihira, Cre recombinase-mediated site-specific modification of a cellular genome using an integrase-defective retroviral vector, *Biotechnol. Bioeng.* 107 (2010) 717-729.
- [8] N.J. Philpott, J. Gomos, K.I. Berns, E. Falck-Pedersen, A p5 integration efficiency element mediates Rep-dependent integration into AAVS1 at chromosome 19, *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 12381-12385.
- [9] B. Moldt, N.H. Staunstrup, M. Jakobsen, R.J. Yanez-Munoz, J.G. Mikkelsen, Genomic insertion of lentiviral DNA circles directed by the yeast FLP recombinase, *BMC Biotechnol.* 8 (2008) 60.
- [10] S.J. Kaczmarczyk, J.E. Green, A single vector containing modified cre recombinase and LOX recombination sequences for inducible tissue-specific

- amplification of gene expression, *Nucleic Acids Res.* 29 (2001) E56-56.
- [11] D. Feng, J. Chen, Y. Yue, H. Zhu, J. Xue, W.W. Jia, A 16bp Rep binding element is sufficient for mediating Rep-dependent integration into AAVS1, *J. Mol. Biol.* 358 (2006) 38-45.
- [12] S.S. Yu, K. Dan, H. Chono, E. Chatani, J. Mineno, I. Kato, Transient gene expression mediated by integrase-defective retroviral vectors, *Biochem. Biophys. Res. Commun.* 368 (2008) 942-947.
- [13] K. Wanisch, R.J. Yanez-Munoz, Integration-deficient lentiviral vectors: a slow coming of age, *Mol. Ther.* 17 (2009) 1316-1332.
- [14] N.J. Philpott, J. Gomos, E. Falck-Pedersen, Transgene expression after re-mediated site-specific integration into chromosome 19, *Hum. Gene Ther.* 15 (2004) 47-61.
- [15] T. Heister, I. Heid, M. Ackermann, C. Fraefel, Herpes simplex virus type 1/adeno-associated virus hybrid vectors mediate site-specific integration at the adeno-associated virus preintegration site, AAVS1, on human chromosome 19, *J. Virol.* 76 (2002) 7163-7173.
- [16] T. Ueno, H. Matsumura, K. Tanaka, T. Iwasaki, M. Ueno, K. Fujinaga, K. Asada, I. Kato, Site-specific integration of a transgene mediated by a hybrid adenovirus/adeno-associated virus vector using the Cre/loxP-expression-switching system, *Biochem. Biophys. Res. Commun.* 273 (2000) 473-478.
- [17] Y. Okada, Y. Ueshin, H. Hasuwa, K. Takumi, M. Okabe, M. Ikawa, Targeted gene modification in mouse ES cells using integrase-defective lentiviral vectors, *Genesis* 47 (2009) 217-223.
- [18] L.A. Donehower, H.E. Varmus, A mutant murine leukemia virus with a single missense codon in pol is defective in a function affecting integration, *Proc. Natl. Acad. Sci. U. S. A.* 81 (1984) 6461-6465.
- [19] S.E. Howden, L. Voullaire, J. Vadolas, The transient expression of mRNA coding

for Rep protein from AAV facilitates targeted plasmid integration, *J. Gene Med.* 10 (2008) 42-50.

- [20] Y.B. Yue, Y.Y. Xue, L. Tian, J.L. Xue, J.Z. Chen, W. Jia, Functional differentiation between Rep-mediated site-specific integration and transcriptional repression of the adeno-associated viral p5 promoter, *Hum. Gene Ther.* 21 (2010) 728-738.
- [21] D. Huser, A. Gogol-Doring, T. Lutter, S. Weger, K. Winter, E.M. Hammer, T. Cathomen, K. Reinert, R. Heilbronn, Integration preferences of wildtype AAV-2 for consensus rep-binding sites at numerous loci in the human genome, *PLoS Pathog.* 6 (2010) e1000985.
- [22] J. Zou, C.L. Sweeney, B.K. Chou, U. Choi, J. Pan, H. Wang, S.N. Dowey, L. Cheng, H.L. Malech, Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting, *Blood* 117 (2011) 5561-5572.



## Figure legends

**Fig. 1.** Schematic representation of retroviral vector constructs and comparison of G418-resistant colony number after transfection or infection. (A) Structure of retroviral vector constructs. MSCV/PTNIG and its plasmid form, pQMSCV/PTNIG, were used for delivering the transgene into HeLa cells. MSCV/Rep and its plasmid form, pQMSCV/Rep, were used for expressing Rep. LTR, long terminal repeat derived from mouse stem cell virus (MSCV);  $\Psi$ +, virus packaging signal sequence from MSCV; p5IEE, p5 integration efficiency element from adeno-associated virus 2 (AAV); *TK*, thymidine kinase gene from herpes simplex virus; *Neo*, neomycin resistance gene; IRES, internal ribosomal entry site sequence from EMCV; *EGFP*, enhanced green fluorescent protein gene; *Rep 68/78*, DNA fragment expressing Rep 68 and 78 proteins from AAV. (B) Comparison of G418-resistant colony numbers after transfection with pQMSCV/PTNIG in the presence or absence of pQMSCV/Rep. (C) Comparison of G418-resistant colony numbers after infecting HeLa cells with IDR/PTNIG in the presence or absence of IDR/Rep. Values represent the averaged colony numbers from a single well of a six-well plate. All experiments were performed in triplicate and the data are presented as the mean value plus standard deviation.

**Fig. 2.** Genomic characterization of HeLa cells co-infected with IDR/PTNIG and IDR/Rep. (A) Southern blot analysis of six clones showing targeted integration at AAVS1. White arrows indicate bands co-binding to both Neo and AAVS1 probes after *EcoRI* digestion. *EcoRI* digestion does not cut IDR/ PTNIG, and releases an 8 kb AAVS1 fragment (wild-type) from the human genome. Black arrows indicate random

integration of IDRV. (B) PCR-based analysis for the six clones with targeted integration. LTR shows the LTR-structures of inserted IDRV circles. TK shows the presence or absence of intact *TK* in the cell genome. *Rep* amplification detects residual integration of IDRV/Rep. *GAPDH* amplification was performed as a control.

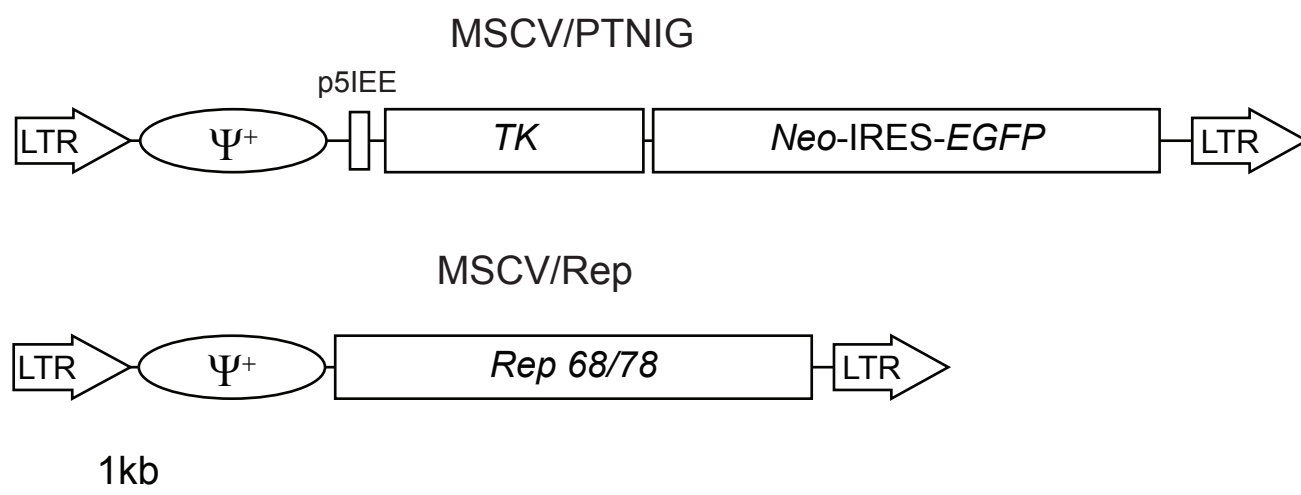
**Fig. 3.** IDRV-AAVS1 junctions in the six HeLa clones. (A) Schematic representation of IDRV/PTNIG with primer-binding sites (P1, P2, TKRV, and N) and integration sites (indicated by arrows). (B) Schematic representation of the AAVS1 locus. J1 and J2 indicate AAVS1-specific primers. Integration sites are indicated by arrows. (C) Sequences of integration junctions. Sequences from IDRV/PTNIG are italicized. Sequences from AAVS1 are bolded. Sequences inserted at the junctions between IDRV/PTNIG and AAVS1 are also given. Numbers indicating the breakpoints in AAVS1 are given according to the sequence from GenBank ID: S51329.1.

**Table 1.** Efficiency of targeted integration

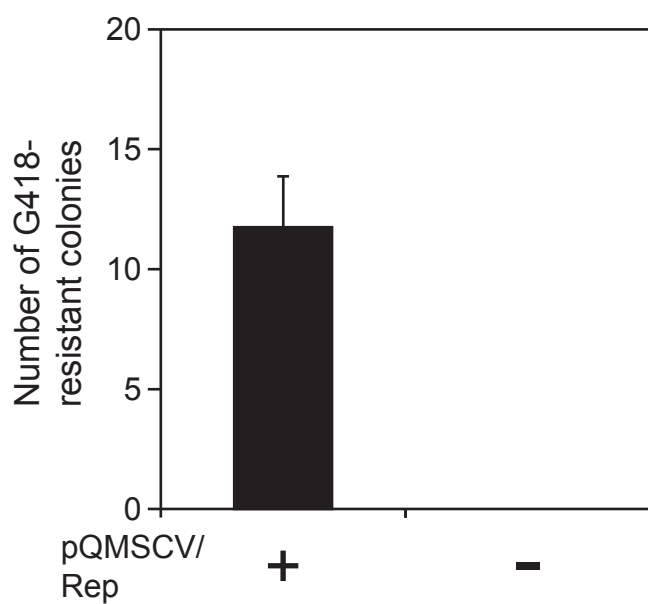
| Transgene delivery method                                      | No. of clones analyzed | AAVS1 disruption <sup>a</sup> | Targeted integration <sup>a</sup> |
|--|------------------------|-------------------------------|-----------------------------------|
| Co-transfection (transgene and Rep gene delivered as plasmids) | 9                      | 6/9 (66%)                     | 5/9 (55%)                         |
| Co-infection (transgene and Rep gene delivered as IDRVs)       | 37                     | 18/37 (49%)                   | 11/37 (30%)                       |

<sup>a</sup>AAVS1 disruption and targeted integration were analyzed by Southern blotting.

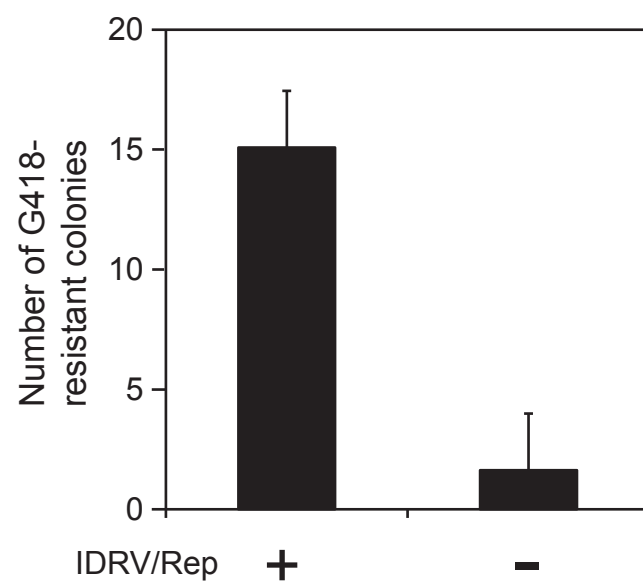
A



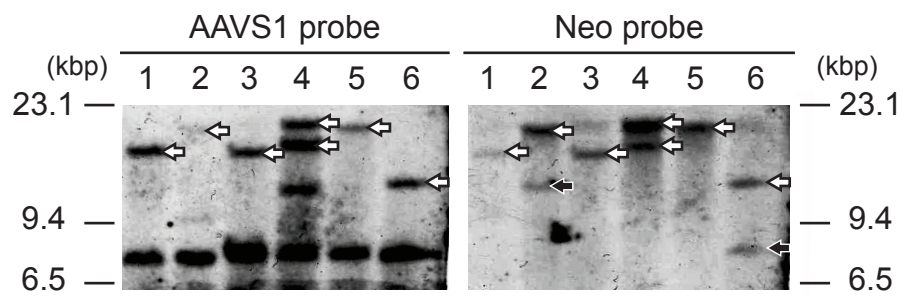
B



C



A



B

