

ORIGINAL ARTICLE

Preferable sites and orientations of transgene inserted in the adenovirus vector genome: The E3 site may be unfavorable for transgene position

M Suzuki, S Kondo, Z Pei¹, A Maekawa, I Saito and Y Kanegae

The adenovirus vector (AdV) can carry two transgenes in its genome, the therapeutic gene and a reporter gene, for example. The E3 insertion site has often been used for the expression of the second transgene. A transgene can be inserted at six different sites/orientations: E1, E3 and E4 sites, and right and left orientations. However, the best combination of the insertion sites and orientations as for the titers and the expression levels has not sufficiently been studied. We attempted to construct 18 AdVs producing GFP or LacZ gene driven by the EF1 α promoter and Cre gene driven by the α -fetoprotein promoter. The AdV containing GFP gene at E3 in the rightward orientation (GFP-E3R) was not available. The LacZ-E3R AdV showed 20-fold lower titer and 50-fold lower level of fiber mRNA than the control E1L AdV. Notably, we found four aberrantly spliced mRNAs in the LacZ-E3L/R AdVs, probably explaining their very low titers. Although the transgene expression levels in the E4R AdVs were about threefold lower than those in the E1L AdVs, their titers are comparable with that of E1L AdVs. We concluded that E1L and E4R sites/orientations are preferable for expressing the main target gene and a second gene, respectively.

Gene Therapy (2015) 22, 421–429; doi:10.1038/gt.2014.124; published online 15 January 2015

INTRODUCTION

First-generation (E1 deleted) adenovirus vectors (FG AdVs), which lack the E1 and E3 regions, are popularly used in basic studies to elucidate gene functions, and have been employed for gene therapy.^{1–4} Because the DNA fragments of up to about 7 kilobases (kb) in total can be inserted into the AdV genome, the AdVs are frequently used to produce two proteins simultaneously from two independent transgenes expressing both the target gene and the reporter gene, for example. In the studies using the cultured cells and in the animal experiments, the GFP and luciferase are used as the reporters. Recently, positron emission tomography has clinically been used in patients for diagnoses and in experimental animal models. Therefore, the AdVs containing both the therapeutic gene and the positron emission tomography reporter gene would be valuable in the gene therapy fields, because the therapeutic effects, the vector duration and distribution can simultaneously be monitored.^{5–8} Probably one would wish for high-titer AdVs with the highest expression for the therapeutic gene and with the second highest for the reporter gene not causing any trouble, if the insertion sites and orientations in the AdV genome can be chosen. However, the titers and the expression levels of the AdVs may considerably be influenced by the sites and orientations of the transgenes. Such information may be very valuable for construction of the best vector, especially in the vector containing both the therapeutic gene and the reporter gene.

The simultaneous expression of two genes could be achieved by inserting the two genes into the E1 site under the control of a single promoter using the internal ribosomal entry sites or using porcine teschovirus-1 2A.^{9,10} In the former approach, the

expression of the second gene might be influenced by the sequences between internal ribosomal entry sites and its initiation codon, and in the latter, the manipulation is necessary to remove the stop codon of the first gene and to adjust the frames of the two genes. When two genes driven by the independent promoters are inserted into the E1 site, they might interfere with each other. However, when two independent expression units are inserted in different sites in the AdV genome, no interference occurs. Moreover, the advantage of this approach is that the main target gene can easily be changed using the AdV cassette that already contains the reporter gene.

There are three insertion sites and two orientations: a transgene can be inserted into the AdV genome by substitution of the E1 or E3 gene and by simple insertion at a position upstream of the E4 gene. Therefore, there are six different possible sites/orientations for any given transgene. Moreover, not only the potent promoters such as EF1 α but also tissue-specific promoters such as α -fetoprotein (AFP) can also be employed. Although the studies examining which sites/orientations are superior to others are practically important, they have been very limited^{11,12} and systematic analyses have not been reported so far.

As it is known that the expression level of a transgene varies considerably depending on the site in the cell chromosome of the human genome, the phenomenon is called the 'position effect'.^{13,14} Although CG-methylation in the cell chromosome is clearly one reason, it is not observed in the AdV genome. Therefore, it would be of interest to examine whether the 'position effect' might also be observed similarly in the AdV genome for the potent promoter and for the tissue-specific promoter.

FG AdVs retain almost all viral genes. They are normally not expressed in the target cells, because E1A protein, the essential transactivator for expression of all other viral genes, is not present. However, there is one report of splicing of aberrant mRNAs from the inserted foreign genes to a viral gene.¹⁵ In this case, the aberrant mRNAs are transcribed by strong foreign promoters and produce transgene-viral gene fusion proteins, which elicit strong immune responses. However, it is not known whether the production of the aberrant gene product between the inserted transgene and viral gene is rare or not.

In this study, we examined the AdV titers and expression levels of an identical transgene inserted at the E1, E3 and E4 sites. We used three transgenes, namely, GFP, LacZ and Cre, and two promoters, namely, the potent EF1 α promoter and the cancer-specific AFP promoter, and attempted to construct AdVs using all combinations, that is, 18 AdVs, and succeeded in constructing 17 of them. We found that insertion at the E1 and E4 sites yielded mostly high titers, whereas the one at the E3 yielded variable titers. Surprisingly, four aberrantly spliced mRNAs between the transgenes and viral genes were found in the vector obtained by insertion at the E3 site, which was probably the reason for the very low titers. As for the expression levels, clear differences were observed among the vectors obtained with insertion at the E1, E3 and E4 sites despite using the identical transgene, indicating that the position effect was certainly present for the AdV genome and that aberrant splicing may, at least in part, explain this effect. We also propose a strategy to avoid generation of the aberrantly spliced mRNAs.

RESULTS

The vector titers were significantly influenced by the insertion sites and orientations of the transgene

We first examined whether the vector titers were influenced by the site/orientations of the transgenes containing a potent EF1 α promoter. Towards this end, we attempted to construct six GFP-expressing (EF-GFP) and six LacZ-expressing (EF-LacZ) vectors in all possible combinations, that is, the E1, E3 and E4 insertion sites and the two orientations (Figure 1), and measured the vector titers (Figure 2a) (hereinafter, the vectors will be designated as per the following; the vectors containing the GFP gene and LacZ gene at the E1 insertion site and in the left orientation shall be denoted as G-E1L and Z-E1L vectors, respectively). Among the GFP-expressing vectors, high titers were obtained for G-E1L, G-E3L, G-E4L and G-E4R vectors (Figure 2a, bars 1, 3, 5 and 6), while the titer for the G-E1R vector was lower (bar 2). Notably, the G-E3R vector, that is, vector with the GFP transgene inserted in the E3 site in the rightward orientation, could not be obtained despite three independent attempts (bar 4, denote 'x'). Therefore, although exactly the same EF1 α -GFP expression unit was inserted in these vectors, the sites and orientations exerted considerable influence on the vector titers and even determined whether the vector was available or not. Similar results were obtained for vectors expressing LacZ: the titers of the Z-E1L, Z-E4L and Z-E4R vectors (bars 7, 11 and 12) were high, and that of the Z-E1R vector was also low (bar 8). However, the results of insertion at the E3 site differed for GFP and LacZ. The titer ratio of Z-E3L was significantly lower than that of G-E3L (compare bars 3 and 9, described later), and the Z-E3R vector was available, although its titer was extremely low (bar 10). Therefore, the GFP gene and LacZ gene themselves influenced the vector titers.

Then, we constructed six vectors containing the AFP promoter and Cre gene (AFP-Cre) and measured their titers (Figure 2b). Although these vectors contained the AFP promoter and Cre gene, this transgene unit served as a nonfunctional DNA, because the AFP promoter, which is hepatocarcinoma-cell-specific, is not active in the 293 cells. The titers of the all six vectors were very

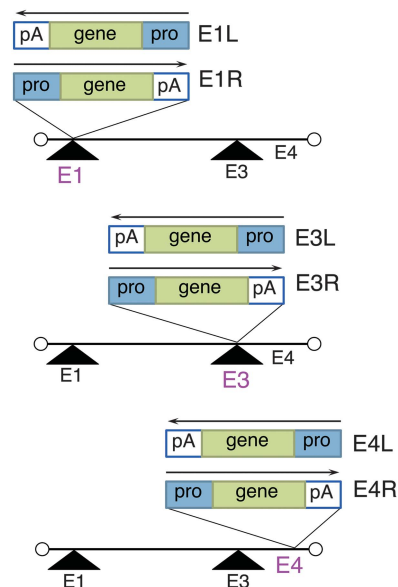


Figure 1. The FG AdV structures of six different site/orientations in all possible combinations. The box containing 'pro,' 'gene' and 'pA' represents the expression unit and the arrows show the orientation of transcription. 'pro,' EF1 α and AFP promoter; gene, GFP, LacZ and Cre; pA, rabbit β -globin polyadenylation signal. For example, the vector containing the transgene at the E1 insertion site and in the left orientation is denoted as 'E1L.'

similar (Figure 2b). Thus, the site/orientation does not always influence the vector titers, and it appeared that there may be some specific reasons why the titers were low for vectors containing the EF1 α promoter expressing the GFP and LacZ genes.

Aberrant chimera mRNAs were produced in the vectors containing the expression unit at the E3 site

The E3 transgene is present within the large intron from the major late promoter (MLP) to the fiber gene (Figure 3a, except the first). We previously reported an aberrant splicing from a cryptic donor site present in the LacZ gene to the viral pIX acceptor site, which produces a LacZ-pIX fusion protein.¹⁵ Therefore, we speculated that similar aberrant splicing might occur for the LacZ gene inserted at the E3 site.

Total RNA was prepared from the 293 cells infected with the E3R vector and reverse-transcribed to detect such aberrantly spliced mRNA spanning from the LacZ cryptic donor site to the possible fiber acceptor site, which is the only acceptor site present downstream of the LacZ donor site. In fact, we identified an aberrant mRNA spliced from this LacZ donor to the fiber acceptor (Figure 3a, second; Figure 3b, 0.7-kb band). The splicing donor site in the LacZ gene was identical to that of the reported LacZ gene inserted at E1 site to the viral pIX acceptor site, and the fiber acceptor site was the same as that normally spliced from the MLP donor site (Supplementary Table S1). This is quite abnormal because, in general, splicing occurs between only specific donor and acceptor sites, suggesting that an inserted transgene could disturb normal splicing.

We also examined whether any other aberrantly spliced mRNA upstream of the transgene was present or not. Surprisingly, we also detected an abnormal mRNA spliced from the donor site of the third exon of the viral MLP to the acceptor site of the second exon of the EF1 α promoter (Figure 3a, third; Figure 3b, 1.2-kb band; the junction sequence is shown in Supplementary Table S1). These results mean that the normal splicing from the MLP donor to the fiber acceptor are doubly competed with aberrant splicing

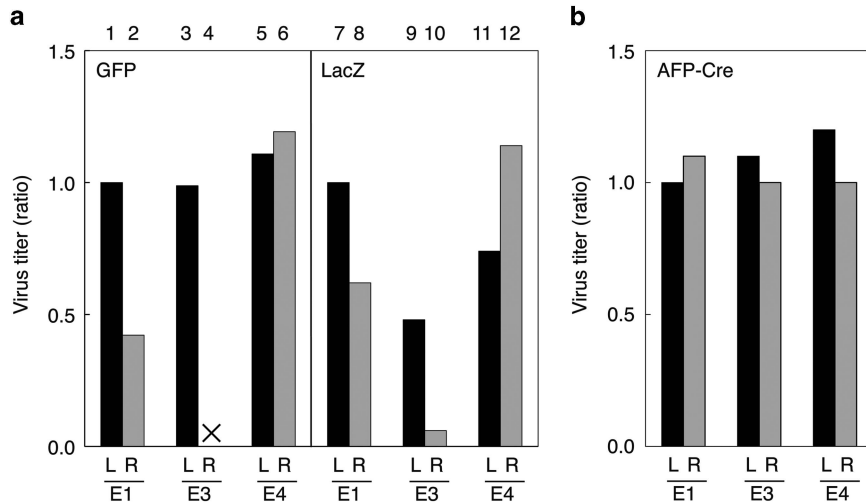


Figure 2. Titers of the virus vectors containing identical expression units. **(a)** Virus titers of the AdVs containing the EF1 α promoter. The AdV genomes transduced into the HuH-7 cells were measured 3 days post infection. The virus titers were calculated relative to the copy numbers of the AdVs.¹⁶ The titer of the E1L vector was set as 1; G-E1L, 8.3×10^8 relative virus titer (rVT)/ml, L-E1L, 5.0×10^9 rVT/ml. 'x' indicates that G-E3R could not be obtained. **(b)** The titers of the virus vector containing Cre gene driven by the AFP promoter. E1L vector was used as the control. * $P < 0.05$, ** $P < 0.01$.

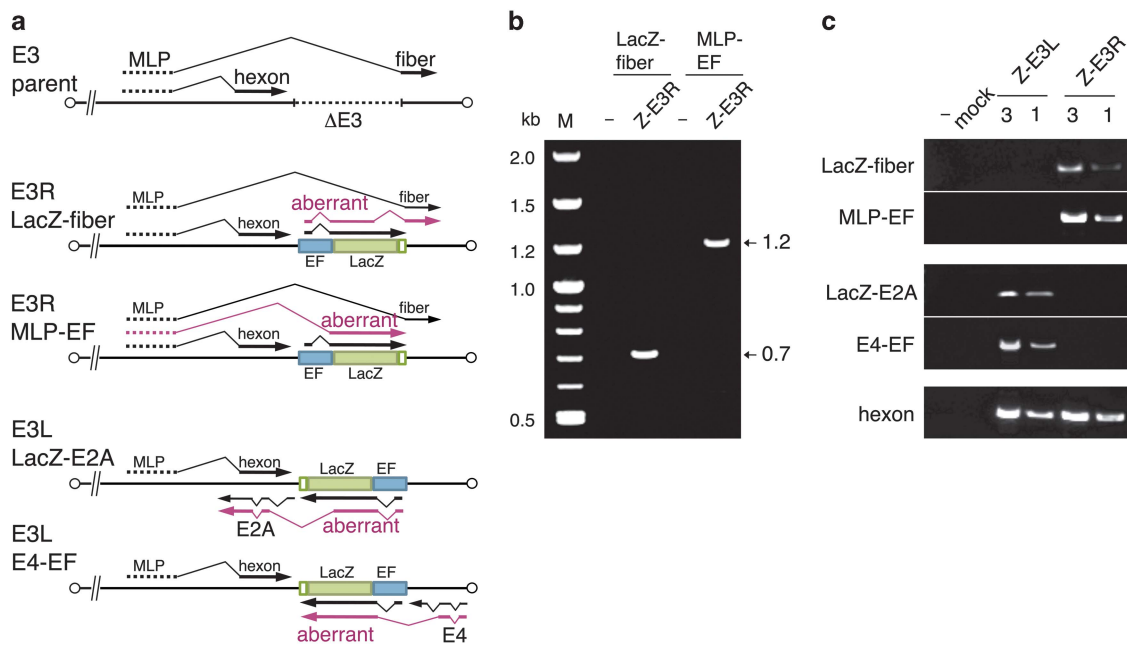


Figure 3. Structures of aberrant chimeric mRNAs. **(a)** Schematic representation of the aberrantly spliced mRNA and the expression unit in the E3 region. The LacZ expression units in the E3 region are shown. Aberrant mRNAs are shown in red. The bold lines and thin polygonal lines represent the exon and intron of the transcription, respectively. Arrow, orientation of the transcription; EF, EF1 α promoter; LacZ, LacZ DNA. The 'parent' denotes the vectors before the insertion into the E3. 'LacZ-fiber,' 'MLP-EF,' 'LacZ-E2A' and 'E4-EF' indicate combinations of primers for detection of the chimeric mRNAs. **(b)** Detection of aberrant splicing by PCR. The 293 cells were infected with the Z-E3L and Z-E3R vectors, as indicated. The bands are generated from the chimera-specific mRNA between the viral gene and the inserted transgene. The primer sequences are shown in Supplementary Table S2. M, size maker; -, no DNA. **(c)** Specificity of the aberrant splicing. The 293 cells were infected with either Z-E3L or Z-E3R. The splicing from the MLP to hexon was used as the control. The bands of LacZ-fiber and MLP-EF in Z-E3R are the same as those described in **(b)** (0.7 and 1.2 kb, respectively). These bands were not detected in Z-E3L (lanes 3 and 1); threefold more DNA was loaded in lane 3 than in lane 1 to clearly show the semi-quantitative difference in the amount of cDNA. mock, mock infection of the 293 cells.

from the MLP donor to the EF1 α acceptor and from LacZ donor to the fiber acceptor. We confirmed that these LacZ-fiber and MLP-EF aberrant mRNAs observed for the Z-E3R vector were not detected for the Z-E3L (Figure 3c, first and second rows).

We further examined whether such abnormal chimera mRNA was present for the Z-E3L vector of the opposite orientation. Actually, we detected two chimera mRNAs using the same PCR analysis; a viral E4 donor was spliced to the EF1 α acceptor, and the

cryptic LacZ donor was spliced to a viral E2A acceptor (Figure 3a, fourth and fifth; Figure 3c, third and fourth rows). They were not detected for Z-E3R). The EF1 α acceptor and LacZ donor were the same as those found in the E3L vector, and the sequences of the viral donor and the acceptor were identical to those found in the wild-type adenovirus, although the combinations were abnormal (Supplementary Table S1).

Then, we measured the amounts of fiber mRNA for the Z-E3R vector (very low titer) and compared them with those for the Z-E1L vector (high titer) and Z-E3L vector (medium titer) by conventional PCR and quantitative PCR (qPCR). These PCR and qPCR primers were designed to detect specifically the normal MLP-fiber mRNA, but not the aberrant mRNA, because they are prepared at the sequence junction: the forward and reverse primers are located in the MLP and fiber, respectively. The cDNAs from the 293 cells infected with Z-E1L, Z-E3L or Z-E3R were diluted from 10^0 to 10^{-3} before PCR (shown as '0' to '-3' in Figure 4a) for semi-quantitative detection (Figure 4a). The level of normal fiber mRNA of Z-E3L (middle titer) was lower and that of Z-E3R (very low titer) was much lower than that of Z-E1L (high titer), that is, the fiber mRNA levels and titers were well correlated. Notably, the amount of fiber mRNA of the Z-E3R vector was only 2% of that for the Z-E1L vector (Figure 4b, E3R). This may probably explain why the titer of Z-E3R was very low.

The expression of the E1 transgenes was higher than that of the E3 or E4 transgenes

The titers show amounts of infections virus particles produced by the vectors growing in the 293 cells, while expression levels are also important for the vector. The amounts of the produced gene products being influenced by the position of an identical expression unit on the cell chromosomes is referred to as the 'position effect' of gene expression.^{13,14} To examine whether

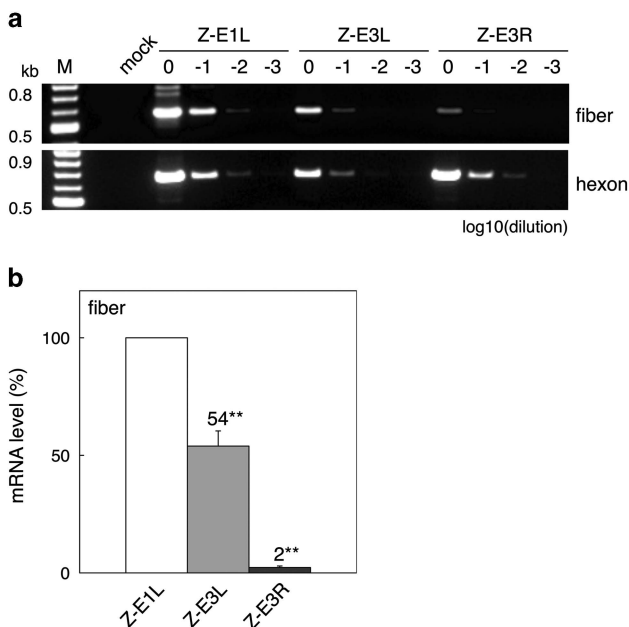


Figure 4. The relative levels of fiber mRNAs. The 293 cells were infected with the Z-E1L, Z-E3L and Z-E3R at MOI 5. (a) PCR detection of the fiber and hexon mRNAs. '0, -1, -2 and -3' mean ' 10^0 , 10^{-1} , 10^{-2} and 10^{-3} dilution of the cDNAs', respectively. M, size marker; mock, mock infection of 293 cells. (b) The quantities of the fiber and hexon mRNAs determined by qPCR. Amount of fiber mRNA relative to those of each hexon and Z-E1L fiber are regarded as 100%. $n=3$; means \pm s.d. ** $P < 0.01$.

identical transgenes inserted into vector genomes are influenced or not by the 'position effect,' we infected the HuH-7 cells with the same numbers of active vector particles¹⁶ expressing GFP under the control of the EF1 α promoter at multiplicity of infection (MOI) 3 and 10 and measured the amounts of GFP mRNAs by qPCR (Figure 5a).

E1L and E1R vectors expressed much more GFP mRNA than the other three vectors, that is, the E3L, E4L and E4R vectors, both at MOI 3 and MOI 10 (bars 2 and 3 to 4, 6 and 7; bars 8 and 9 to 10, 12 and 13). In regard to the expression levels of these vectors, the mRNA amounts of the E4 vectors were about one-third of those of the E1 vectors. Similar results were obtained for the vectors expressing LacZ: E1L/R vectors expressed much more LacZ mRNA than all the E3 and E4 vectors, both at MOI 3 and MOI 10 (Figure 5b, bars 2 and 3 to 4–7; bars 8 and 9 to 10–13). Therefore, similar effects were observed using two different genes. These results might suggest that the position effect observed here was not dependent on the inserted transgene. It should be noted that about as much as a 50-fold more virus stock solution of Z-E3R vector could be needed than that of the Z-E1L vector to obtain the same expression level, because the titer of Z-E3R vector was about one-tenth and the expression level obtained was about one-fifth when the same volume of the virus stock solution is used for infection (Figure 2a, bars 7 and 10; Figure 5b, bars 2–5 and 8–11). We also measured the expressed protein levels of GFP and LacZ using fluorometry and β -galactosidase assay, respectively (Figure 6a and b). The G-E1 vectors produced significantly more GFP than the G-E4 vectors (Figure 6a, bars 2 and 3 to 6 and 7; bars 8 and 9 to 12 and 13), although the G-E3L vector expressed a similar level to that of the G-E1 vectors (bars 4 and 10). Also, the Z-E1 vectors expressed more LacZ than the Z-E3 and Z-E4 vectors (Figure 6b, bars 2 and 3 to 4–7; bars 8 and 9 to 10–13). These results in respect of the protein level confirm the mRNA expression levels measured by qPCR shown in Figure 5a and b.

To examine whether the position effects may also be observed for a tissue-specific promoter and for genes other than GFP and LacZ, the HuH-7 cells were infected with the vector expressing Cre under the control of the AFP promoter and the Cre expression levels were measured. The AFP promoter is specifically active in the HuH-7 cells derived from hepatocarcinoma in contrast to the case in the 293 cells. Because tissue-specific promoters, including the AFP promoter, are generally weak, the expressed Cre mRNA level was too low to measure quantitatively. Therefore, we used the method of 'excisional expression,' where the Cre enzyme driven by the AFP promoter switched on the potent EF1 α promoter and specifically enhanced the expression level of dsRed by about 50-fold¹⁷ (the strategy is shown in Supplementary Figure S1). The results were again very similar to those obtained using the EF1 α promoter (Figure 6a and b): the AFP-E3 and E4 vectors expressed only about a half to one-fifth of dsRed mRNA than the AFP-E1 vectors (Figure 5c, bars 3 and 4 to 5–8, 9 and 10 to 11–14). Therefore, although the vector titers obtained using the AFP promoter were not influenced by their insertion sites (Figure 2b), the position effects at the E1, E3 and E4 sites showed very similar patterns to those of the EF1 α promoter. Altogether, the E3L/R and E4L/R vectors expressed about two to fivefold less transgene products than the E1L/R vectors, not only when the potent EF1 α promoter was used, but also when the tissue-specific AFP promoter was used, suggesting there may be a mechanism common to these promoters.

To examine whether the position effect of expression observed in the HuH-7 cells may also be observed in other cells, the HeLa cells were infected with the GFP-expressing vectors at MOI 3 and at MOI 10 (Figure 7a and b). The G-E1L vector expressed a significantly greater amount of mRNA than the G-E3 and G-E4 vectors (Figure 7a, bars 2 to 4, 6 and 7; bars 8 to 10, 12 and 13). However, the mRNA level of the E1R vector was not significantly different from those of the E3 and E4 vectors, because the

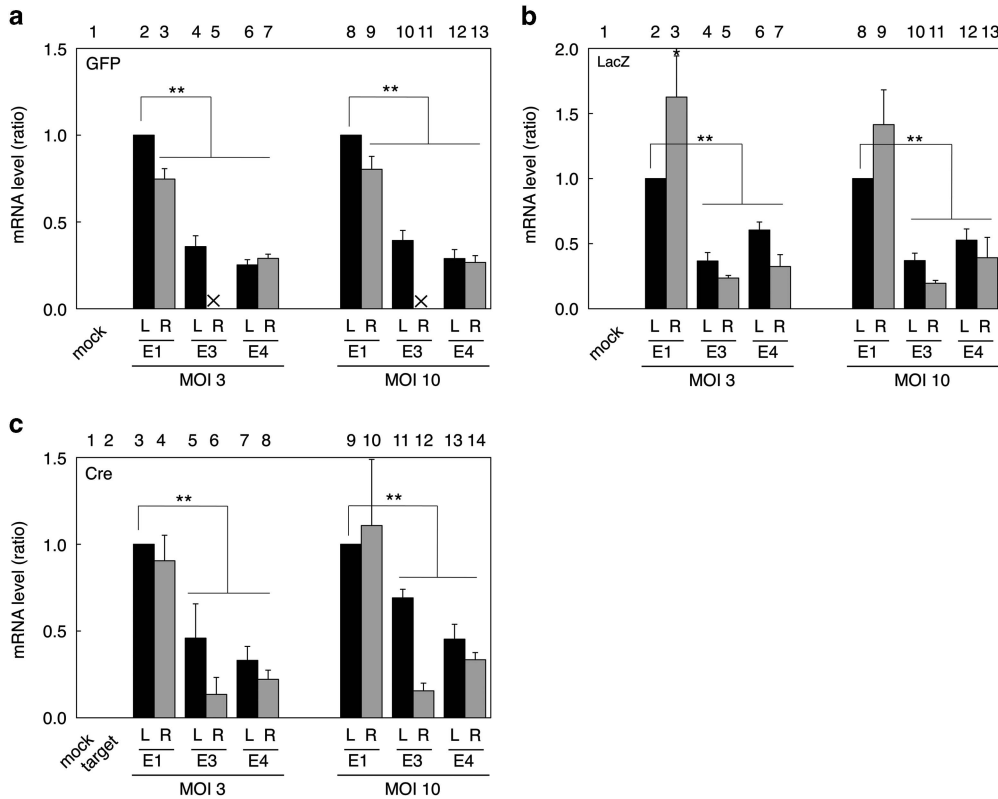


Figure 5. Expression levels of the transgene mRNA in the infected HuH-7 cells. The mRNA levels are shown relative to the mRNA level of the transgene in E1L-infected cells set as 1. **(a)** GFP mRNA levels. The cells were infected with the EF-GFP AdVs at the indicated MOIs. The GFP mRNA levels were quantified by qPCR, $n=3$. **(b)** LacZ mRNA levels. The cells were infected with the EF-LacZ AdVs. LacZ mRNA levels are shown in the same manner as that described in **(a)**, $n=4$. **(c)** Cre mRNA levels. The cells were co-infected with the AFP-Cre AdVs (switch vectors) and a target vector which expressed dsRed by Cre-mediated recombination (excisional expression). The dsRed mRNA levels were quantified in the same manner as that described in **(a)**, $n=6$. Error bars indicate mean \pm s.d.; mock, HuH-7 cells without infection; target, HuH-7 cells infected with the target vector only; * $P < 0.05$, ** $P < 0.01$. The other representations are the same as those in Figure 1.

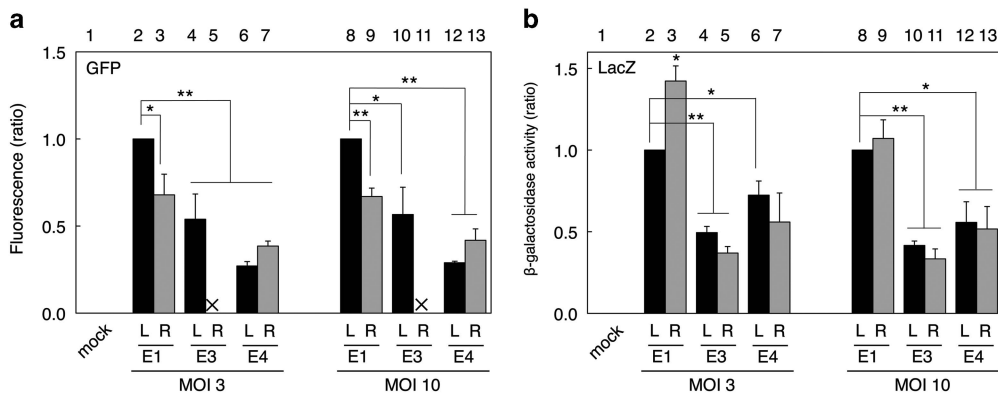


Figure 6. Protein expression levels in the infected HuH-7 cells. **(a)** Fluorescence of GFP. The cells were infected with the EF-GFP AdVs at the indicated MOIs. The fluorescence of GFP was quantified by Ascent fluorometry. The fluorescences are shown relative to the fluorescence level in E1L-infected cells set as 1, $n=4$. **(b)** Activities of β -galactosidase. The cells were infected with the EF-LacZ AdVs. β -galactosidase activities were evaluated by the β -gal assay, $n=3$. Error bars represent \pm s.d.; mock, mock infection of HuH-7 cells; * $P < 0.05$, ** $P < 0.01$. The other representations are the same as those in Figure 1.

expressed mRNA level of the E1R was lower than that of the E1L (bars 3 to 4, 6 and 7; bars 9 to 10, 12 and 13). Similar results were obtained using GFP fluorometry: the G-E1L vector exhibited significantly more fluorescence than the other G-E3 and G-E4 vectors (Figure 7b, bars 2 to 4, 6 and 7), whereas the E1R vector

expression was not statistically significant (bars 3 to 4, 6 and 7). These results were confirmed by fluorescence microscopy (Supplementary Figure S2a). Moreover, the same results were obtained using the CV-1 cell line derived from monkey fibroblasts (Supplementary Figure S2b). Therefore, very similar position effect

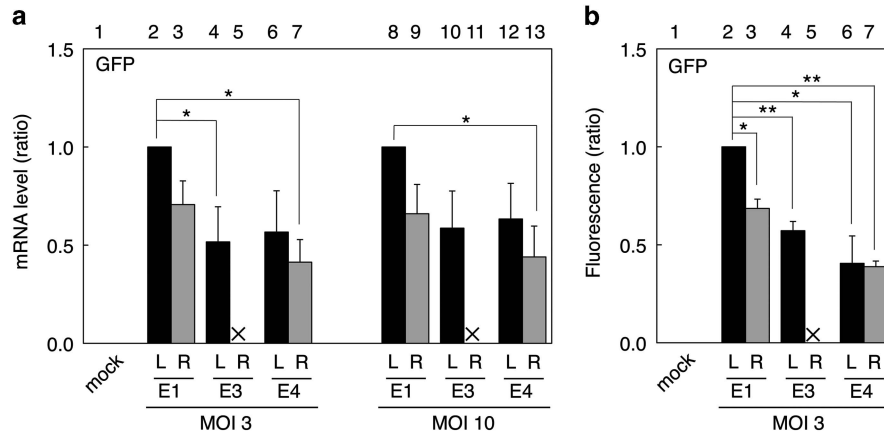


Figure 7. Expression levels of the transgene mRNA and protein in the infected HeLa cells. **(a)** GFP mRNA levels. The HeLa cells were infected with the EF-GFP AdVs at the indicated MOIs. The other representations are the same as those in Figure 4a, $n = 3$. **(b)** Fluorescence of GFP. The HeLa cells were infected with EF-GFP AdVs at MOI 3. The other representations are the same as those in Figure 5a, $n = 4$. Error bars indicate mean \pm s.d.; mock, mock infection of HeLa cells; * $P < 0.05$, ** $P < 0.01$.

among the E1, E3 and E4 insertion sites are obtained at least for the G-E1L vector.

DISCUSSION

We demonstrated in this study that the inserted sites and orientations for a given transgene greatly influenced the vector titers and expression levels. Especially, when the transgene was inserted in E3R, the GFP-expressing vector could not be obtained and the LacZ-expressing vector titer was extremely low. Also, the titer of E3L AdV was lower than that of E1L AdV. Because the aberrantly spliced mRNAs from the transgenes to a viral gene have been reported for the E1R site/orientation¹⁵ (described below), and similar aberrant splicing might have occurred for E4L site/orientation in the same mechanism. Therefore, considering the titers, aberrant splicing and expression levels, E1L and E4R sites/orientations were preferable for the main target gene and the second gene, respectively, in the simultaneous expression. As for the titer, the information might be useful not only for FG AdVs but also for the replication-competent AdVs containing E1A gene under the control of a cancer-specific promoter, because they both are prepared using the 293 cells.

We have demonstrated that the vectors containing transgene at E1L/R showed higher titers and expression levels than other vectors. E1L/R is the most frequently used, probably because the E1 site is required or convenient for the major methods of AdV construction which are now commonly used.^{18–20} For example, for the method established by the Graham's group, the use of the E1 site is essential because it exploits the viral packaging sequences partially overlapping with the E1 region. Consequently, we think that the E1 site, found to be the best in this work, might have been chosen in the currently popular methods. However, there seems to be one concern with E1R: aberrant splicing has been reported to occur to viral pIX gene from the cryptic donors present not only in the LacZ gene but also in the herpes thymidine-kinase (TK) gene, which is used for positron emission tomography as a reporter gene and for suicide gene therapy. Consequently, TK-pIX and LacZ-pIX fusion proteins were produced, and the pIX protein evokes strong immune responses.¹⁵ For this reason, we always adopt the leftward orientation for the E1 site.

Currently, in the simultaneous expression of the target gene and the reporter gene, the E3 sites are mostly employed for the reporter gene.^{21–25} However, as described here, when the LacZ-expressing transgene driven by the EF1 α promoter was inserted at E3R, the vector titer and expression level were very low, probably because of the aberrant splicing, and the G-E3R AdV could not be

obtained. In our experiences, the E3R AdV containing GFP gene driven by SR α promoter²⁶ could also not be obtained. In contrast, however, the E3R AdV containing GFP driven by CMV promoter could be obtained. The reason of such difference is unclear, but it might be related to the fact that both EF1 α and SR α promoters contain the splicing unit including the splicing acceptor site in their promoters, which might produce the aberrant mRNA spliced from the MLP donor, but the CMV promoter contain no splicing unit. Because aberrant splicing was detected even at E3L, both E3R and E3L were problematic and to be avoided, if possible.

The E4 site has not frequently been used.^{17,19,27–29} The E4 insertion position, *Sna*BI site, is located 162 nucleotides (nt) from the right end of the AdV genome. Vectors containing a transgene at the E4 site showed high titers, although their expression levels were lower than those of the E1 vectors. The titers and expression levels do not significantly differ between E4L and E4R vectors. However, the E4L might produce aberrantly spliced mRNAs as observed for the E3L/R. The E4R site/orientation was successfully used as the position of the second transgene¹⁷ where Cre was expressed under the control of the AFP promoter. The expressed Cre turned on the potent promoter present at the E1L and the high-level expression of the target gene was obtained, while maintaining strict specificity. Interestingly, it was also recently reported that the E4L, not the E4R, is better than the E1 site for short hairpin RNA expression.²⁷ The difference may be related to the use of the RNA polymerase III promoter for the short hairpin RNA production, whereas the polymerase II promoter is used for the protein production. If so, the E4L may be advantageous not only for the production of short hairpin RNA, but also of guide RNAs used for the CRISPR/Cas9 system.^{30,31} The reported results might not contradict with the results described here, because the RNA polymerase III expression is not involved with splicing.

Altogether, therefore, the E1L and E4R sites/orientations appear to be the best for use in AdVs for the simultaneous expression of the target gene and reporter gene, respectively. Importantly, the occurrence of aberrant splicing sometimes yields a viral-transgene fusion protein, which may induce strong immune responses caused by the viral encoding region.¹⁵ The probability is one-third to coincide the coding frame of the transgene with that of the viral gene. The coding regions of the four aberrant mRNAs described in this report were, by chance, connected out-of-frame, yielding no transgene-fusion protein, but only a truncated LacZ protein composed of several amino acids. However, in a previous study,¹⁵ LacZ-pIX and TK-pIX fusion proteins were produced under the control of potent promoters. Thus, the production of such fusion proteins by aberrant splicing is not a rare event.

In general, use of the same promoter may cause troubles by the homologous recombination in the simultaneous recombination. We have experienced that, when the identical promoters in the E1L and E4R sites/orientations were applied, a rearrangement through the homologous recombination occurred, whereas the AdVs in the E1L and E4L was stable. Therefore, use of different promoters would be preferable.

The harmful aberrant splicing can be avoided using the technique described herein. The aberrant splicing can be identified by PCR analysis as reported here. The preferred sequences of the splicing donor site are AGGT/AAGT (slash denotes the exon-intron junction), where the underlined GT dinucleotides are definitely required for splicing. Nucleotide mutations of these sequences, especially the GT nucleotides, can be introduced without changing the amino acid encoded by the therapeutic gene, and successful disruption of the aberrant splicing can be easily confirmed by a subsequent PCR using the same primer set. As an example, the change from T to C, G or T (Gly) in the aberrant splicing present in the TK gene removes the splicing while keeping the amino acid sequence intact, and this mutated TK gene is expected to be safer than the original TK gene for suicide gene therapy and for positron emission tomography analyses using TK gene as a reporter.

The results described here demonstrated that position effects were evident for the expression of transgenes present on the AdV genome. The mechanisms of the position effect in the AdV genome are unknown. We think that viral enhancers, silencers or other *cis*-acting sequences near to the foreign transgene promoter might influence its expression. It might be related to the fact that similar results were obtained for a strong EF1 α promoter and a cancer-specific AFP promoter. In contrast to the expression levels, the vector titers were probably not influenced by the position of the transgene, because similar and high titers were obtained using a tissue-specific promoter, which does not produce the transgene mRNA in the 293 cells (Figure 2b). The results might suggest that the vector titers were mainly influenced by the combination of a strong promoter and aberrant splicing between the transgene and viral genes (Figure 2a). These results are probably valuable for efficient and safe gene therapy using FG AdVs.

In this work, the range of the E3 deletion is between two *Xba*I sites in the E3 region (nt position 28,592–30,470). Because the lengths of the E3 deletions are slightly different among the commercially available AdV construction kits, the differences might influence the results obtained here. We surmise that these differences might not influence the conclusion described here as far as the same transgene is used, because the same splicing event is essentially expected to occur by the same mechanism. However, the possibility cannot be ruled out.

The cassette plasmids (also used as cosmids) are available on request in a collaboration basis, which bear the full-length AdV genome containing the unique *Swa*I site at the E1 site and the unique *Clal* site at the E4 site (pA_{xw4cit2}) and that inversely containing the *Clal* site at the E1 site and *Swa*I at the E4 site (pA_{xc4wit2}). The method to insert a transgene into these sites is described in the Materials and Methods section.

MATERIALS AND METHODS

Cells and virus titration

Human 293,³² HeLa and HuH-7 cell lines are derived from the human embryonic kidney, human cervical carcinoma and human hepatocellular carcinoma, respectively. The CV-1 cell line is derived from African green monkey kidney. The cells were cultured in Dulbecco's Modified Eagles Medium supplemented with 10% fetal calf serum. The 293 cells constitutively express adenoviral E1 genes and support the replication of E1-substituted AdVs. After infection with AdVs, the cells were maintained in Dulbecco's Modified Eagles Medium supplemented with 5% fetal calf serum without geneticin. FG AdVs were titrated using the method

described by Pei *et al.*¹⁶ Briefly, the copy numbers of a viral genome that was successfully transduced into the infected target cells were measured by real-time PCR (relative virus titer). HuH-7 cells were used as the target cells. The titer of the standard virus was determined using the copy number of serially diluted plasmid DNA. When FG AdVs are used, the relative virus titer (copies per ml) normally corresponds to about one-fifth of the Tissue Culture Infectious Dose 50 titer, when the gene product is not deleterious; the probable reason for this difference is that the transduction efficiency of the 293 cells is exceptionally high as compared with that of the other cells. The sequences of the TaqMan probes for the titration are derived from adenovirus-5 (Ad5) pIX gene: forward primer, 5'-TGTG ATGGGCTCCAGCATT-3'; probe, 5'-ATGGTCGCCCGCTCCTGCC-3'; reverse primer, 5'-TCGTAGGTCAAGGTAGTAGAGTTTGC-3'.

Vector construction

All the AdVs described here were constructed using the cosmid cassette pA_{xcwit2} containing the full-length AdV genome.^{19,33} The GFP/LacZ-expression unit were under the control of the EF1 α promoter³⁴ and the Cre-expression unit was driven by the AFP promoter; the AFP promoter used here was the (AB) 256 AFP promoter.³⁵ All the expression units were inserted into the *Swa*I cloning site at the E1 substitution region as E1L or R vectors. All the E3L and E3R vectors possessed the same expression units at the *Xba*I site in the E3 region of pA_{xcwit2}; the *Xba*I site is originally generated by deletion between the two *Xba*I sites in the Ad5 genome (nt position 28,592–30,470). The E4 cloning site represented by the *Sna*BI site (nt position 35770) located in the E4 region at 165-nt downstream from the right end of the Ad5 genome.¹⁷

Conventional PCR

The PCR experiments were essentially performed by the standard method.³⁶ Typically, the 293 cells in the 6-well plate were infected at MOI 5 and 16 h post infection, total RNAs were prepared and reverse-transcribed using oligo(dT) primer; the resultant cDNAs were amplified by PCR with Tks Gflex DNA polymerase (Takara Bio, Shiga, Japan) and a PCR system (ProFlex PCR system, Applied Biosystems, Foster City, CA, USA). The PCR cycling conditions were in accordance with the manufacturer's protocol (Takara Bio): 94 °C for 1 min, followed by 30 cycles at 98 °C for 10 s, 60 °C for 15 s, and 68 °C for 30 s. The primer sets are described in Supplementary Table S2. The PCR products containing each splice junction were subjected to agarose gel electrophoresis.

Quantitative real-time PCR

The sequences of the GFP primers and the dsRed primers have been described previously.^{17,37} The sequences of the LacZ primers were as follows: forward primer, 5'-ATCAGGATATGTGGCGGATGA-3'; probe, 5'-CGG CATTTCGGTGACGTCT-3'; reverse primer, 5'-TGATTTGTGTAGTCGGTTTA TGCA-3'. The primer sequences of the normal splicing junctions of MLP-fiber and MLP-hexon were as follows: for MLP-fiber detection forward primer in MLP third exon, 5'-AAAGGCGTCTAACCAGTCACAGT-3'; probe in the fiber gene, 5'-AGCGCGAAGACCGTCTGAAGATAACC-3'; reverse in the fiber gene, 5'-CCGCTTTCGGTGCATATGG-3'; and for MLP-hexon detection forward primer in MLP third exon, 5'-TCTAACCAGTCACAGTCGCAAGA-3'; probe in the hexon gene, 5'-CGCGCCCGCTTCCAAGATG-3'; reverse in the hexon gene, 5'-CACTGCGGCATCATCGAA-3'.

The mRNA levels were calculated as described by Maekawa *et al.*³⁷ Briefly, the total RNA of the infected cells was extracted, and the amounts of the expressed target RNAs and 18S-rRNA (correction standard) were quantified using reverse transcription (TaqMan Reverse Transcription Reagents, Roche, Basel, Switzerland) and real-time PCR (Applied Biosystems Prism 7000); the ratio of the target RNA to 18S-rRNA was then calculated. To quantify the AdV genome, the infected total cell DNA was prepared from the cells using a previously described method^{29,38} or a DNA preparation kit (Macherey-Nagel, through Takara Bio). Quantitative PCR was performed to detect the AdV genome using a probe for the pIX gene, as described above.¹⁷ The amount of chromosomal DNA was simultaneously measured to correct the Ct values of the viral genome per cell, and the corrected Ct is shown throughout. The probes were derived from the sequence of the human β -actin gene for the HeLa and HuH7 cell lines. The qPCR reaction was performed using the following cycling conditions according to the manufacturer's protocol (Applied Biosystems): 50 °C for 2 min and 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s and 60 °C for 1 min.

Measurement of the expressed GFP and LacZ

The HuH-7, HeLa and CV1 cells were infected at MOI 3 or MOI 10 of each vector in a 24-well plate in triplicated experiments. Three days after the infection, the infected cells were washed twice with phosphate-buffered saline. The cells in the three wells were then fixed with 4% paraformaldehyde to quantify the GFP fluorescence using Labsystems Fluoroskan Ascent FL (GMI, Ramsey, MN, USA) or by fluorescence microscopy. The cells infected with the LacZ vectors were harvested for the quantification of β -galactosidase (β -gal) (β -gal assay kit, Invitrogen, Carlsbad, CA, USA). To quantify the β -gal activity, the infected cells were disrupted by sonication and the lysate was subjected to a color reaction assay using o-Nitrophenyl β -D-galactopyranoside. The stained color standard was determined using 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal).

Insertion of a transgene into the cassette plasmid (cosmid) containing both the E1 and E4 sites

The method of inserting a given transgene into the *SwaI* site of pAxc4cit2 and pAxc4wit2 is followed by the protocol of Takara Bio (Adenovirus Dual Expression Kit). Briefly, the plasmid containing the transgene DNA fragment is treated with the appropriate restriction enzymes and treated with the DNA polymerase I Klenow fragment, followed by agarose gel electrophoresis. The isolated transgene fragment of about 50 ng is ligated with 1–2 μ g of the cassette cosmid at a volume of 15 μ l at 15°C for overnight. The ligated DNA is cleaved with *SwaI* to remove the self-ligated parent plasmid, and then transformation or lambda *in vitro* packaging is performed. The latter method is highly efficient and removes the deleted plasmids smaller than 38 kb. To insert a given transgene DNA to the *Clal* site, the DNA fragment is treated with the Klenow polymerase and ligated with a DNA linker of *BspT104I* (Takara Bio, reaction temperature is 37°C) or *BstBI* (New England Biolabs, Ipswich, MA, USA, 65°C), 5'-GGTTCGAACC-3' (the underline shows the recognition sequences), for example, and digested with either enzyme. Because the termini produced with this enzyme can be ligated with *Clal*-cleaved DNA and the ligated DNA cannot be cleaved with either enzyme. Therefore, after ligation of the transgene and *Clal*-cleaved cassette, the self-ligated parent plasmid can be removed by *Clal* digestion. Alternatively, both *SwaI* and *Clal* sites can be converted to *I-CeuI*, *I-SseI* or both by insertion of the cleavage-site oligonucleotides.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank N Goda for research assistance and Ms T Shino for her secretarial assistance. This study was supported in part by Grants-in-Aids from the Ministry of Education, Culture, Sports, Science and Technology to YK and SK; The Program for Intractable Disease Research utilizing Disease-specific iPS Cells from JST to YK; a grant for Practical Research on Hepatitis B (009) from the Ministry of Health, Labour and Welfare of Japan to IS.

REFERENCES

- Crystal RG. Adenovirus: the first effective in vivo gene delivery vector. *Hum Gene Ther* 2014; **25**: 3–11.
- Watanabe M, Nasu Y, Kumon H. Adenovirus-mediated REIC/Dkk-3 gene therapy: Development of an autologous cancer vaccination therapy (Review). *Oncol Lett* 2014; **7**: 595–601.
- Wold WS, Toth K. Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. *Curr Gene Ther* 2013; **13**: 421–433.
- Danthinne X, Imperiale MJ. Production of first generation adenovirus vectors: a review. *Gene Ther* 2000; **7**: 1707–1714.
- Gil JS, Machado HB, Campbell DO, McCracken M, Radu C, Witte ON *et al*. Application of a rapid, simple, and accurate adenovirus-based method to compare PET reporter gene/PET reporter probe systems. *Mol Imaging Biol* 2013; **15**: 273–281.
- Qin C, Lan X, He J, Xia X, Tian Y, Pei Z *et al*. An in vitro and in vivo evaluation of a reporter gene/probe system hEFL(18)F-FES. *PLoS One* 2013; **8**: e61911.
- Zhang G, Lan X, Yen TC, Chen Q, Pei Z, Qin C *et al*. Therapeutic gene expression in transduced mesenchymal stem cells can be monitored using a reporter gene. *Nucl Med Biol* 2012; **39**: 1243–1250.
- Gambhir SS, Barrio JR, Phelps ME, Iyer M, Namavari M, Satyamurthy N *et al*. Imaging adenoviral-directed reporter gene expression in living animals with positron emission tomography. *Proc Natl Acad Sci USA* 1999; **96**: 2333–2338.
- Chan HY, V S, Xing X, Kraus P, Yap SP, Ng P *et al*. Comparison of IRES and F2A-based locus-specific multicistronic expression in stable mouse lines. *PLoS One* 2011; **6**: e28885.
- Kim JH, Lee SR, Li LH, Park HJ, Park JH, Lee KY *et al*. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 2011; **6**: e18556.
- Small JC, Kurupati RK, Zhou X, Bian A, Chi E, Li Y *et al*. Construction and characterization of E1- and E3-deleted adenovirus vectors expressing two antigens from two separate expression cassettes. *Hum Gene Ther* 2014; **25**: 328–338.
- Pham L, Nakamura T, Gabriela Rosales A, Carlson SK, Bailey KR, Peng KW *et al*. Concordant activity of transgene expression cassettes inserted into E1, E3 and E4 cloning sites in the adenovirus genome. *J Gene Med* 2009; **11**: 197–206.
- Yankulov K. Dynamics and stability: epigenetic conversions in position effect variegation. *Biochem Cell Biol* 2013; **91**: 6–13.
- Wilson C, Bellen HJ, Gehring WJ. Position effects on eukaryotic gene expression. *Annu Rev Cell Biol* 1990; **6**: 679–714.
- Nakai M, Komiya K, Murata M, Kimura T, Kanaoka M, Kanegae Y *et al*. Expression of pIX gene induced by transgene promoter: possible cause of host immune response in first-generation adenoviral vectors. *Hum Gene Ther* 2007; **18**: 925–936.
- Pei Z, Kondo S, Kanegae Y, Saito I. Copy number of adenoviral vector genome transduced into target cells can be measured using quantitative PCR: Application to vector titration. *Biochem Biophys Res Commun* 2012; **417**: 945–950.
- Kanegae Y, Terashima M, Kondo S, Fukuda H, Maekawa A, Pei Z *et al*. High-level expression by tissue/cancer-specific promoter with strict specificity using a single-adenoviral vector. *Nucleic Acids Res* 2011; **39**: e7.
- Bett AJ, Haddara W, Prevec L, Graham FL. An efficient and flexible system for construction of adenovirus vectors with insertions or deletions in early regions 1 and 3. *Proc Natl Acad Sci USA* 1994; **91**: 8802–8806.
- Miyake S, Makimura M, Kanegae Y, Harada S, Sato Y, Takamori K *et al*. Efficient generation of recombinant adenoviruses using adenovirus DNA-terminal protein complex and a cosmid bearing the full-length virus genome. *Proc Natl Acad Sci USA* 1996; **93**: 1320–1324.
- Mizuguchi H, Kay MA. Efficient construction of a recombinant adenovirus vector by an improved in vitro ligation method. *Hum Gene Ther* 1998; **9**: 2577–2583.
- Shin SP, Seo HH, Shin JH, Park HB, Lim DP, Eom HS *et al*. Adenovirus Expressing Both Thymidine Kinase and Soluble PD1 Enhances Antitumor Immunity by Strengthening CD8 T-cell Response. *Mol Ther* 2013; **21**: 688–695.
- Suzuki T, Sasaki T, Yano K, Sakurai F, Kawabata K, Kondoh M *et al*. Development of a recombinant adenovirus vector production system free of replication-competent adenovirus by utilizing a packaging size limit of the viral genome. *Virus Res* 2011; **158**: 154–160.
- Trujillo MA, Oneal MJ, McDonough S, Qin R, Morris JC. A probasin promoter, conditional for example replicating adenovirus that expresses the sodium iodide symporter (NIS) for radiotherapy of prostate cancer. *Gene Ther* 2010; **17**: 1325–1332.
- Mailly L, Boulade-Ladame C, Orfanoudakis G, Deryckere F. A novel adenovirus vector for easy cloning in the E3 region downstream of the CMV promoter. *Viral J* 2008; **5**: 73.
- Mizuguchi H, Kay MA, Hayakawa T. In vitro ligation-based cloning of foreign DNAs into the E3 and E1 deletion regions for generation of recombinant adenovirus vectors. *Biotechniques* 2001; **30**: 1112–1114; 1116.
- Takebe Y, Seiki M, Fujisawa J, Hoy P, Yokota K, Arai K *et al*. SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cellular Biol* 1988; **8**: 466–472.
- Pei Z, Shi G, Kondo S, Ito M, Maekawa A, Suzuki M *et al*. Adenovirus vectors lacking virus-associated RNA expression enhance shRNA activity to suppress hepatitis C virus replication. *Sci Rep* 2013; **3**: 3575.
- Mizuguchi H, Xu ZL, Sakurai F, Yamuyi T, Hayakawa T. Tight positive regulation of transgene expression by a single adenovirus vector containing the rTA and tTS expression cassettes in separate genome regions. *Hum Gene Ther* 2003; **14**: 1265–1277.
- Saito I, Oya Y, Yamamoto K, Yuasa T, Shimojo H. Construction of nondefective adenovirus type 5 bearing a 2.8-kilobase hepatitis B virus DNA near the right end of its genome. *J Virol* 1985; **54**: 711–719.
- Cong L, Ran FA, Cox D, Lin SL, Barretto R, Habib N *et al*. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013; **339**: 819–823.
- Mali P, Yang LH, Esvelt KM, Aach J, Guell M, DiCarlo JE *et al*. RNA-guided human genome engineering via Cas9. *Science* 2013; **339**: 823–826.

- 32 Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* 1977; **36**: 59–74.
- 33 Fukuda H, Terashima M, Koshikawa M, Kanegae Y, Saito I. Possible mechanism of adenovirus generation from a cloned viral genome tagged with nucleotides at its ends. *Microbiol Immunol* 2006; **50**: 643–654.
- 34 Kim DW, Uetsuki T, Kaziro Y, Yamaguchi N, Sugano S. Use of the human elongation factor-1-alpha promoter as a versatile and efficient expression system. *Gene* 1990; **91**: 217–223.
- 35 Sato Y, Tanaka K, Lee G, Kanegae Y, Sakai Y, Kaneko S *et al*. Enhanced and specific gene expression via tissue-specific production of Cre recombinase using adenovirus vector. *Biochem Biophys Res Commun* 1998; **244**: 455–462.
- 36 Sambrook J, Russell DW. *Molecular Cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2001.
- 37 Maekawa A, Pei Z, Suzuki M, Fukuda H, Ono Y, Kondo S *et al*. Efficient production of adenovirus vector lacking genes of virus-associated RNAs that disturb cellular RNAi machinery. *Sci Rep* 2013; **3**: 1136.
- 38 Nakano M, Odaka K, Takahashi Y, Ishimura M, Saito I, Kanegae Y. Production of viral vectors using recombinase-mediated cassette exchange. *Nucleic Acids Res* 2005; **33**: e76.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Supplementary Information accompanies this paper on Gene Therapy website (<http://www.nature.com/gt>)