ARTICLE

Characterization of integration frequency and insertion sites of adenovirus DNA into mouse liver genomic DNA following intravenous injection

Zhibin Wang [1✉], Philip J. Troilo[1], Thomas G. GriffithsII[1], Laural B. Harper[1], Amy B. Barnum[1], Stephen J. Pacchione[1], Cindy J. Pauley[1], Jose A. Lebron[1], Jayanthi Wolf [1] and Brian J. Ledwith[1]

While generally referred to as "non-integrating" vectors, adenovirus vectors have the potential to integrate into host DNA via random, illegitimate (nonhomologous) recombination. The present study provides a quantitative assessment of the potential integration frequency of adenovirus 5 (Ad5)-based vectors following intravenous injection in mice, a common route of administration in gene therapy applications particularly for transgene expression in liver. We examined the uptake level and persistence in liver of first generation (FG) and helper-dependent (HD) Ad5 vectors containing the mouse leptin transgene. As expected, the persistence of the HD vector was markedly higher than that of the FG vector. For both vectors, the majority of the vector DNA remained extrachromosomal and predominantly in the form of episomal monomers. However, using a quantitative gel-purification-based integration assay, a portion of the detectable vector was found to be associated with high molecular weight (HMW) genomic DNA, indicating potential integration with a frequency of up to ~44 and 7000 integration events per µg cellular genomic DNA (or ~0.0003 and 0.05 integrations per cell, respectively) for the FG and HD Ad5 vectors, respectively, following intravenous injection of 1×10^{11} virus particles. To confirm integration occurred (versus residual episomal vector DNA co-purifying with genomic DNA), we characterized nine independent integration events using Repeat-Anchored Integration Capture (RAIC) PCR. Sequencing of the insertion sites suggests that both of the vectors integrate randomly, but within short segments of homology between the vector breakpoint and the insertion site. Eight of the nine integrations were in intergenic DNA and one was within an intron. These findings represent the first quantitative assessment and characterization of Ad5 vector integration following intravenous administration in vivo in wild-type mice.

INTRODUCTION

Replication-defective adenovirus vectors are commonly employed as experimental gene delivery vehicles for gene therapy and genetic vaccination approaches because they can infect both proliferating and quiescent cells, and they can be easily prepared to high titers [1–3]. As of a review in 2017, 547 clinical trials (20.5%) among a total of 2670 trials worldwide had used adenovirus vector [4]. The FG Ad5 vector remains the most popular adenovirus vector currently used in clinical protocols. FG Ad5 vectors have a deletion in the E1 region that is required both for adenovirus replication and for downstream viral gene expression, and hence must be propagated in cell lines that provide the E1-proteins in trans. However, residual or "leaky" adenoviral gene expression from the FG vectors following administration in vivo can elicit a cytotoxic-T-lymphocyte (CTL)-mediated immune response, resulting in the rapid loss of transgene expression [2]. To overcome these limitations, HD Ad5 vectors, also known as "gutted" vectors, have been developed. HD Ad5 vectors, in which all of the adenoviral genes have been deleted, must be propagated by co-culture with a helper

adenovirus that provides all of the viral proteins necessary for virus replication and assembly in trans. Several studies have demonstrated that HD Ad5 vectors can achieve long-term transgene expression in vivo [3, 5, 6].

One of the primary safety concerns in using gene transfer vectors is the risk of insertional mutagenesis caused by integration of the foreign DNA into the host genomic DNA. Adenovirus vectors are generally considered to be "non-integrating" vectors because they lack native integration machinery [7, 8]. However, several studies have shown that replication-incompetent adenovirus vectors are capable of random integration into host chromosomes. In cell culture in vitro, adenovirus vectors have been shown to integrate at a rate of $\sim 10^{-2}$–10^{-5} per infected cell, with the absolute frequency depending on cell type, transgene, and multiplicity of infection [8–12]. Moreover, integration of adenovirus type 12 (Ad12) DNA has been demonstrated in tumors induced by subcutaneous injection of Ad12 in newborn hamsters [13]. Adenovirus vector integration was also observed in a FAH-deficient mice model by intravenous injection [14]. However, there has been no quantitative assessment of adenovirus vector

integration in vivo or in the absence of selective pressure for cells containing the integrated vector, which is essential for determining the risk of insertional mutagenesis in many of the clinical applications of Ad5 vectors.

In the present study, we assessed the potential integration frequency of FG and HD Ad5 vectors containing the mouse leptin gene (FG-leptin and HD-leptin, respectively [5, 6]) in liver DNA following intravenous injection in mice, a common route of administration particularly for transgene expression in liver. A gel-purification quantitative PCR (Q-PCR) assay analogous to that previously used to assess plasmid integration [15–18] was used to quantify an integration frequency. In this assay, HMW genomic DNA is purified away from extrachromosomal Ad5 vector DNA by multiple rounds of pulsed-field gel electrophoresis (PFGE) and assayed for potentially integrated vector using highly sensitive Q-PCR, capable of detecting 10 copies of vector per µg cellular genomic DNA. Following gel purification, the level of vector DNA remaining associated with HMW genomic DNA represents a worst-case integration frequency, since it is possible that some of the residual vector is episomal but co-purifies with the genomic DNA. Therefore, to confirm integration of the Ad5 vectors and to characterize some of the chromosomal insertion sites, we used RAIC PCR, which permits direct amplification and sequencing of vector-to-genomic DNA junctions and is capable of detecting low frequency integration events in nonclonal populations, and has previously been used to characterize integration events following intramuscular injection and electroporation of plasmid DNA in mice [18].

MATERIALS AND METHODS
Adenovirus vectors
The structure of the two vectors used in the RAIC PCR assay is shown in Fig. 1. The leptin expression cassette contains the human cytomegalovirus (HCMV) promoter, the human leptin cDNA, and the bovine growth hormone poly(A) sequence. The first generation adenovirus vector (FG-leptin) was constructed by replacing the E1 domain of the Ad5 genome with the leptin expression cassette [5, 6]. Construction of the helper-dependent adenovirus vectors (HD-leptin and HD-leptin-D) have been described in detail in a previous study [5, 6].

Mouse treatment and isolation of liver DNA
All mouse housing and experiments were conducted in strict accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) of Merck & Co., Inc., Kenilworth, NJ, USA. The FG-leptin, HD-leptin

and HD-leptin-D were injected into the tail vein of C57BL/J6 female mice, respectively. The dose was 10^{11} virus particles, measured by plaque-forming units [5, 6], per mouse, and control mice received vehicle. Liver was harvested at 2, 4, 6.5, and 8 weeks after dosing, frozen in liquid nitrogen, and stored at −70 °C. Cohort tissues (from three mice) were homogenized, pooled, and incubated in DNA lysis solution containing SDS, proteinase K and RNase. Total DNA for the standard DNA preparation was extracted with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol, and precipitated with isopropanol. DNA samples were quantitated by spectrophotometry and aliquots were run on analytical agarose gels to ensure that each was of HMW.

Filter DNA preparation
This procedure maintains the genomic DNA at a very high molecular weight such that more efficient PFGE separation parameters can be used for gel purification. The method has also been shown to be capable of removing some of the extrachromosomal vector that is present in cell lysates. Homogenized or minced tissue was placed on a 2 µm filter and gently lysed in a solution containing SDS and proteinase K. Following digestion, the solution and five rinses of TE buffer was pulled through the filter using a peristaltic pump. The high molecular weight DNA remaining on top of the filter was collected and quantitated by spectrophotometry.

Gel purification
For the quantitative integration assay and for DNA to be used as template for the RAIC PCR, HMW genomic DNA was separated from free vector using PFGE. All PFGE gel purifications were done using the BioRad CHEF (Clamped Homogeneous Electric Field) Mapper System. The electrophoresis parameters were determined by the instrument using auto-algorithms for resolution of 10–55 Kb (for the standard prep), or 20–300 Kb (for the filter prep) DNA fragments. Approximately 5 µg of DNA was loaded per well in gels cast using a BioRad 14-well comb. For multiple rounds of purification, the excised band of HMW genomic DNA from the first gel was directly subjected to successive rounds of gel purification without elution from the gel. After the final round of gel purification, DNA was electroeluted from the gel slices and then concentrated by centrifugation in Centricon-30 filtration units (Millipore, Inc.). In some cases, for the quantitative integration assay, gel-purified DNA was digested with Asc I or Srf I restriction enzyme prior to a final round of gel purification to remove any potential Ad5 concatemers that may have comigrated with the HMW genomic DNA. Gel-purified DNA for the RAIC PCR was not digested with a restriction enzyme.

Southern blot
Five micrograms of liver DNA, which was isolated by the standard procedure, was run on PFGE using an auto-algorithm for resolution of 5–150 Kb DNA fragments. After denaturing, neutralizing, and transferring to a ZETABIND® membrane (CUNO, Inc.), DNA was hybridized with ^{32}P-labeled probes based on the intact FG-leptin vector DNA for the FG-leptin-treated liver DNA sample and a PCR-generated fragment (from the mouse leptin coding region) for the HD-leptin-D-treated liver DNA sample.

TaqMan® PCR
Samples (0.5 µg of DNA per reaction) were assayed in quadruplicate by "TaqMan®" real-time quantitative fluorescence PCR using the ABI Prism® 7700 Sequence Detection System (Applied Biosystems) for the following leptin segment:

Amplicon length = 94 bp
Forward primer, LPI-F: 5′-GAC CAT TGT CAC CAG GAT CAA
Reverse primer, LPI-R: 5′-GTG AAG CCC AGG AAT GAA GTC
Fluorescent probe, LPI-P: 6FAM-TCC GCC AAG CAG AGG GTC ACT G-TAMRA

The leptin PCR segment is targeted to the leptin cDNA from nucleotides 969–1062 within FG-leptin; from nucleotides 8935–9028 within HD-leptin; and from nucleotides 6048–6141 and 26,841–26,934 for the two leptin inserts within HD-leptin-D. Prior to PCR, DNA samples were heat denatured at ~95 °C for ~15 min. TaqMan® PCR reactions were carried out in a final volume of 50 µL containing 1X TaqMan® Universal Master Mix (Applied Biosystems), 100 nM probe, 180 nM of each primer, and 0.5 µg of sample DNA. In each TaqMan® PCR experiment, a positive control titration curve consisting of 0, 10, 10^2, 10^3, 10^4, 10^5, and 10^6 copies of each vector per µg of control DNA was assayed along with the test samples. The level of
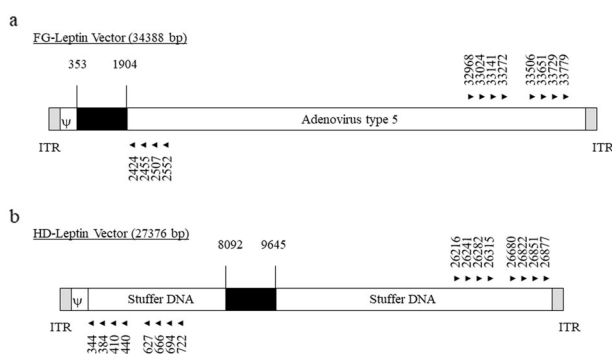
Fig. 1 FG-leptin and HD-leptin structures and vector primers for RAIC PCR. For each vector, the ITR, the packaging signal (ψ), and the leptin expression cassette (in black) is indicated. **a** For the FG-leptin vector, three sets of primers were designed. **b** For the HD-leptin vector, four sets of primers were designed. Each primer set included four nested vector primers (staggered triangles) that allowed up to three rounds of nested PCR followed by sequencing. The sequences of the primer sets used to characterize nine integration events are shown in Table 1.

vector DNA in a sample was determined by comparison of the sample's Ct value with those of the titration curve.

RAIC PCR

The RAIC PCR was performed essentially as previously described [18]. The templates used were PFGE gel-purified DNA from the liver of mice that were treated with FG-leptin or HD-leptin. Primers used for the RAIC PCR assay are listed in Table 1 and were purchased from Operon Technologies. The mouse B1 primer was based on the mouse B1 sequence (Genebank accession number, J00631). For the first round of PCR, the mouse B1 primer and GT primer contained a 5' Tag sequence and dUTPs were substituted for dTTPs. The primers for the FG-leptin sample were based on sequence of the FG-leptin adenovirus vector. One set of primers directed to 5' end and two sets of primers to 3' end were designed to amplify potential vector-genomic junctions (Fig. 1a). The primers for the HD-leptin sample were based on sequence of the HD-leptin adenovirus vector. Two sets of primers directed to each vector end were designed to amplify potential vector-genomic junctions (Fig. 1b). All primer sets for both vectors included four primers that allowed up to three rounds of nested PCR followed by sequencing. The 5'ends of all round 1 and 2 vector primers were labeled with biotin. The PCR products from the third round were purified using Microcon-PCR centrifugal units (Millipore, Inc.). A nested vector-specific primer and a Tag-nested primer were used to sequence both ends of the fragment. When direct sequencing failed to generate acceptable sequence, the specific PCR products were gel-purified, cloned into TOPO pCR2.1 vector using the TOPO TA Cloning Kit (Invitrogen), and each end of the insert was sequenced using forward and reverse M13 primers. All sequencing was performed on an ABI 377 DNA sequencer with the ABI Prism BigDye Terminator v3.1 (Applied Biosystems). When cloning was used to facilitate DNA sequencing, the resulting sequence was confirmed by reamplification of the uncloned PCR products using primers derived from the newly identified sequences, thus ruling out the artificial joining of vector and genomic DNA via the cloning procedure.

RESULTS

Uptake and persistence of FG-leptin and HD-leptin in liver

Two studies were performed in which FG-leptin or HD-leptin constructs were injected into the tail vein of C57BL/J6 female mice at a dose of 10^{11} virus particles per mouse. In the first study, mice were injected with FG-leptin or HD-leptin-D (in the latter construct, the leptin gene was duplicated as an inverted repeat; as described in ref. [5]) and liver was harvested at 2, 4, and 8 weeks after dosing. In the second study, mice were injected with HD-leptin (a second HD construct, without the duplication [5]) and liver was harvested at 6.5 weeks after dosing. Total liver DNA was isolated using a standard procedure. The level of vector DNA in liver was examined using real-time fluorescence Q-PCR (TaqMan® PCR) and the results are shown in Table 2. As expected, the level of vector persisting in the liver was much greater for the gutted vector. At 2 weeks after injection, the level of HD-leptin-D in liver was ~11-fold greater than the level of FG-leptin. Moreover, the FG-leptin level decreased five-fold between the 2 weeks and 8 weeks timepoints, whereas the level of HD-leptin-D remained unchanged.

Conformation of FG-leptin and HD-leptin in vivo

To investigate the conformation of the adenovirus vector in vivo, Southern blots were performed on liver DNA samples from both the FG-leptin and the HD-leptin-D-treated mice. As shown in

Table 1. Primers for the RAIC PCR assay[a].

Primers for GT and Mouse B1 (B1 consensus sequence is underlined)	
B1 Forward	
CAGUUCAAGUGUCUGCUGACGCCUACGCAGAUCAUGC<u>CAGGCGGAUUUCUGAGUUCGAGG</u>	
B1 Reverse	
CAGUUCAAGUGUCUGCUGACGCCUACGCAGAUCAUGC<u>UGCCUCUGCCUCCCAAGUGCU</u>	
GT	
CAGUUCAAGUGUCUGCUGACGCCUACGCAGAUCAUGCGUGUGUGUGUGUGUGUGUGUGUGUG	
Tag	
CAAGUGUCUGCUGACGCCUACG	
Tag-nested	
UGACGCCUACGCAGAUCAUGC	
Primers used to detect integration HD49, HD72, and HD220	
1722 (1st round PCR)	[bioTEG] TGTAGGACTGAACGTCTTGCTCGAG
1694 (2nd round PCR)	[bioTEG] TGATGAAGGAGATGGGAGGCCATC
1666 (3rd round PCR)	TGTAGCCCTCTGTGTGCTCAAGG
1627 (Sequencing)	TGCTGACCTGCTGGATTACATC
Primers used to detect integration FG22, FG117, FG388, FG1043, and FG1756	
2552 (1st round PCR)	[bioTEG] TCCAACAGCTGCTGAGAAACGAC
2507 (2nd round PCR)	[bioTEG] TCCAATTGTGCCAAAAGAGCCGTC
2455 (3rd round PCR)	AAGCTGCACTGCTTGCAAGCG
2424 (Sequencing)	AGCAAAGTCAGTCACAATCCC
Primers used to detect integration FG34388	
33506 (1st round PCR)	[bioTEG] AGGTAAGCTCCGGAACCACCAC
33651 (2nd round PCR)	[bioTEG] AGCATAAGACGGACTACGGCCATG
33729 (3rd round PCR)	TCCTCGGTCATGTCCGGAGTC
33779 (Sequencing)	TTGATTCATCGGTCAGTGC

[a]The sequence of the primers used for the mouse B1 short interspersed nucleotide element (B1 SINE), GT dinucleotide repeat, Tag sequence, and the nested primer sets for the FG Ad5 or HD Ad5 vectors used to characterize nine integration events are shown.

**Table 2.** Uptake and persistence of adenovirus vector in mouse liver.

Vector	Copies of vector per µg of liver DNA (Copies of vector per cell[a])			
	(at various timepoints following intravenous injection)			
	2 week	4 week	6.5 week[c]	8 week
FG-leptin	61,000 (0.41)	28,000 (0.19)	n.d.	12,000 (0.08)
HD-leptin-D[b]	700,000 (4.67)	900,000 (6)	n.d.	700,000 (4.67)
HD-leptin	n.d.	n.d.	600,000 (4)	n.d.

n.d. not determined, *FG* first generation Ad5 vector, *HD* helper-dependent Ad5 vector.
[a]Divided copies of vector per µg DNA by 150,000 cells/µg DNA.
[b]HD-leptin-D contained two copies of the leptin gene per viral particle due to a duplication within the vector. Therefore, the leptin gene segment was present at twice the vector level indicated in this table.
[c]Separate study examining vector levels in the liver at 6.5-weeks post-administration.
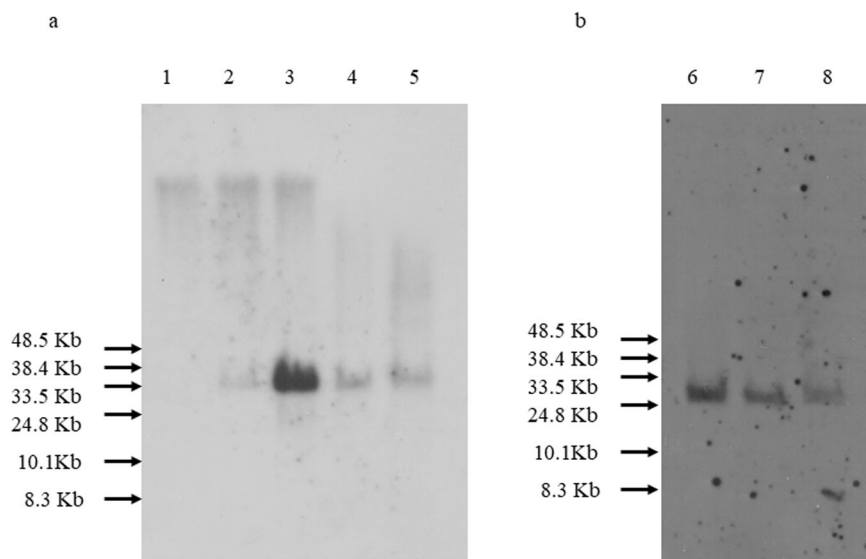
Fig. 2 Southern analysis to examine the in vivo state of adenoviral DNA. Molecular weight marker was shown on the left. **a** First generation adenoviral DNA (FG-leptin) was detected. Lane 1, 5 µg of negative control mouse genomic DNA; lane 2, 10^5 copies of FG-leptin spiked in 5 µg of the control DNA; lane 3, 10^6 copies of FG-leptin spiked in 5 µg of the mouse genomic DNA; lane 4, 5 µg of the FG-leptin-treated mouse liver DNA at 2 weeks after dosing; lane 5, 5 µg of the FG-leptin-treated mouse liver DNA at 4 weeks after dosing. **b** Helper-dependent adenoviral DNA (HD-leptin-D) was detected. lane 6, 5 µg of the HD-leptin-D-treated mouse liver DNA at 2 weeks after dosing; lane 7, 5 µg of the HD-leptin-D-treated mouse liver DNA at 4 weeks after dosing; lane 8, 5 µg of the HD-leptin-D-treated mouse liver DNA at 8 weeks after dosing.

Fig. 2a and b, the vector that persisted in the liver was found to migrate as a distinct band identical to that observed in the vector control lane, indicating that most of the vector persisted as linear monomers identical in conformation to what had been injected. The HMW signal observed in the treated-mouse samples was similar to that observed with control mouse DNA (indicating it was due to nonspecific hybridization). These results indicate that the majority of the adenovirus vector remained extrachromosomal and predominantly persisted as an episomal monomer in liver at least 8 weeks post injection. However, the Southern blot analysis, which is relatively insensitive, does not rule out low levels of other forms such as concatemers or integrated vector (as assessed below).

Gel-purification based assay for integration into host genomic DNA
To quantitatively assess the level of potentially integrated adenoviral vector, mouse genomic DNA was first purified away from free vector DNA by preparative PFGE, and then assayed for integrated vector DNA by TaqMan® PCR. The original PFGE parameters were set to resolve 10–55 Kb DNA, where HMW genomic DNA migrates as a compact band, well separated from the adenovirus vector monomer (Fig. 3a). Using this procedure, the level of FG-leptin in liver DNA samples (from 8-weeks post-dose) was decreased from 16,000 copies per µg DNA prior to gel purification to 2000 copies per µg DNA following gel purification (Table 3). Likewise, the level of the HD-leptin-D in liver DNA samples (from 8-weeks post-dose) was decreased from 700,000 copies per µg DNA prior to gel purification to 75,000 copies per µg DNA following gel purification, and the level of the HD-leptin in liver DNA samples (from 6.5-weeks post-dose) was decreased from 600,000 copies per µg DNA prior to gel purification to 100,000 copies per µg DNA following gel purification (Table 3). These findings demonstrate that at least 80–90% of the vector persisting in liver at 6–8-weeks post-injection was extrachromosomal. Note, under the conditions used for these assays (no restriction-enzyme digestion), the gel-purified HMW genomic DNA is representative of the whole genomic DNA since the genomic DNA is fragmented by random, mechanically shearing [17].

To better remove extrachromosomal vector monomers and concatemers (which could comigrate with the HMW genomic DNA), two enhancements to the gel-purification integration assay
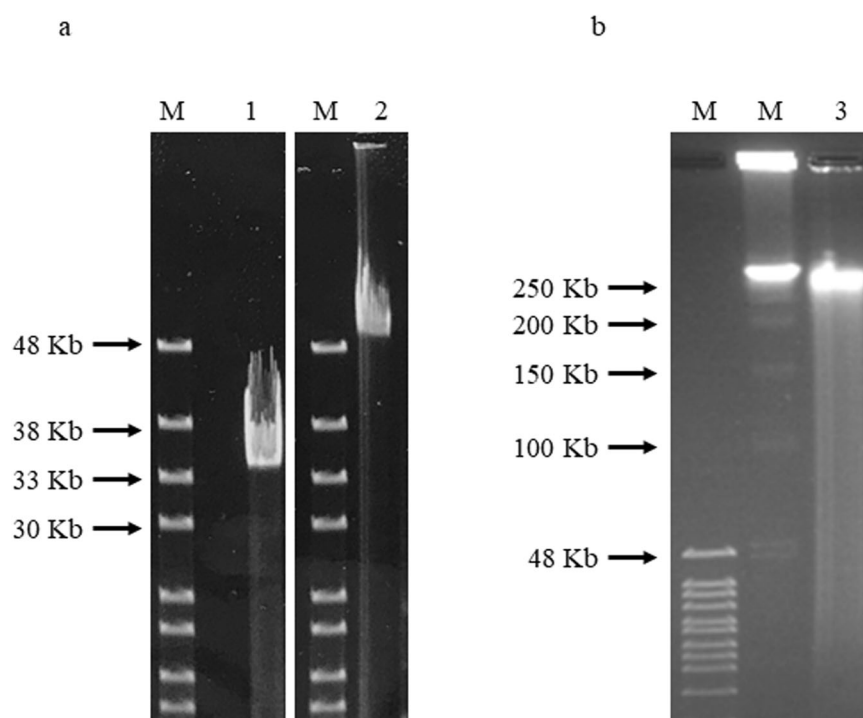
Fig. 3 Pulsed-field gel to remove free adenovirus vector DNA from HMW genomic DNA. The two types of pulsed-field run parameters: **a** 10–55 Kb PFGE and **b** 10–150 Kb PFGE. Lanes M, molecular weight standards; lane 1, FG-Leptin vector DNA; lane 2, standard preparation DNA; and lane 3, filter prep DNA.

Table 3. Potential integration of adenovirus vector revealed by gel purification.

Vector	Timepoint	DNA prep method[a]	Copies of vector per µg of mouse liver DNA (Copies of vector per cell[b])		
			Pre-Gel	Post-Gel	Post RE and Gel[c]
FG-leptin	8 week	Standard	16,000 (0.11)	2000 (0.01)	22 (0.00015)[d]
HD-leptin-D	8 week	Standard	700,000 (4.7)	75,000 (0.5)	n.d.
HD-leptin	6.5 week	Standard	600,000 (4)	100,000 (0.67)	n.d.
		Filter	67,000 (0.45)	62,000 (0.41)	3500 (0.023)[d]

[a]Standard DNA preparation is total cellular DNA, and filter DNA preparation is cellular HMW DNA that remains after PK digestion and filtering through a 2 µm filter.
[b]Divided copies of vector per µg DNA by 150,000 cells/µg DNA.
[c]Vector levels after removal of potential concatemers by restriction-enzyme (RE) digestion, followed by gel purification (see text).
[d]Under the conditions used, RE digestion could result in ~50% of the integrations being lost during gel purification [17], hence integration level could be approximately twice that of the vector level shown: 44 and 7000 integrations per µg of mouse DNA (0.0003 and 0.05 integrations per cell) for FG-leptin and HD-leptin, respectively.

were employed. First, a "Filter Prep" method for isolating HMW liver DNA was used. DNA was isolated from tissue directly on a 2 µm filter, resulting in the removal of some free vector and yielding DNA of very high molecular weight, which could then be gel-purified using more stringent PFGE parameters (Fig. 3b). As above, the HMW DNA isolated using the Filter Prep can be considered representative of the whole genomic DNA since the genomic DNA is fragmented by random, mechanically shearing. Second, restriction-enzyme (RE) digestion was performed prior to gel purification, using a rare-cutting RE that is known to cleave within the vector DNA such that the vector DNA is digested while the HMW genomic DNA remains relatively intact. The advantage of the RE digestion is that it facilitates the separation of free vector (particularly concatemers) from HMW genomic DNA. Clearly, if the vector were integrated into genomic DNA as monomers or concatemers, these monomers or concatemers would also be digested. However, the PCR target segments were

selected to be outside of the RE site and thus, while internal copies of concatemers would be liberated, at least one copy of the PCR target would usually remain associated with HMW genomic DNA at each integration site (see Discussion).

For the liver DNA sample from FG-leptin treated mice, after RE (Asc I) digestion and an additional round of gel purification, the level of the FG-leptin associated with HMW genomic DNA was reduced to 22 copies per µg DNA (Table 3). For the liver DNA sample from HD-leptin treated mice, using a filter prep as well as RE (Srf I) digestion prior to gel purification, the level of the HD-leptin associated with HMW genomic DNA was reduced to 3500 copies/µg DNA. Since under the conditions used, RE digestion could result in ~50% of the integrations being lost during gel purification [17], the integration levels could be approximately twice the levels of vector detected, or ~44 and 7000 integrations per µg DNA for FG-leptin and HD-leptin, respectively. Thus, even after the most stringent attempts to remove extrachromosomal

vector, substantial levels of the vectors remained associated with genomic DNA, suggesting a portion of the residual copies were integrated into the mouse genome.

Confirmation and characterization of Ad5 vector integration

To confirm Ad5 vector integration and identify the insertion sites, we utilized the RAIC PCR assay which is able to detect rare, single-copy, integration events in a complex mixture in vivo [18]. A schematic of the RAIC PCR process is shown in Fig. 4. In this study, in addition to the B1 SINE primer used previously [18], a GT-repeat primer was also used as a genomic anchor. There are ~184,500 copies of GT repetitive elements interspersed in the 2.5-gigabase haploid mouse genome, and each occurs, on average, every 13.6 Kb [19]. Therefore, a random integration event would fall, on average, within ~6.8 Kb of a GT repetitive element. Thus, as is the case with a B1 genomic anchor, a subset of integration junctions should be able to be amplified using a vector-specific primer and a GT primer, long-PCR conditions, steps to remove 'repeat:repeat' products (UDG digestion and biotin capture of vector-containing products), and reamplification using nested PCR.

In the RAIC assay for integrated FG-leptin vector, liver DNA from 4-weeks post-dose was purified by two rounds of preparative PFGE, using instrument parameters optimized to resolve 10–55 Kb
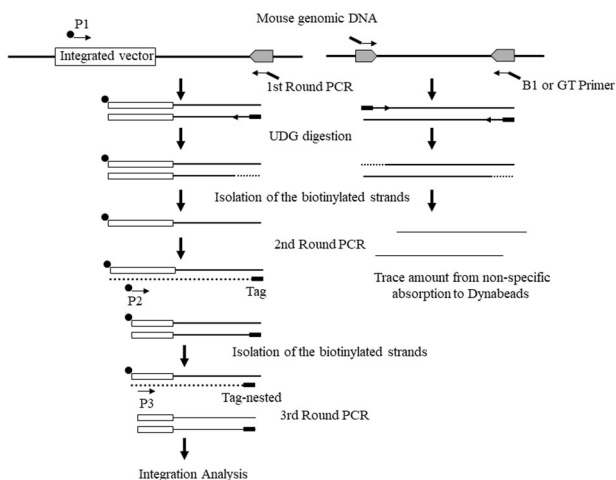

Fig. 4 Strategy of the RAIC PCR assay. The left panel represents the vector-to-B1 or GT amplification, and the right panel shows the B1-to-B1 or GT-to-GT amplification. The open box and line represent the integrated vector sequence and unknown mouse genomic sequence, respectively. The shadow arrow boxes represent B1 or GT elements on the mouse genome in different orientations. The arrow with a bold bar shows the B1 or GT primer in which dTTPs were substituted with dUTPs, and where the bold bar indicates the Tag sequence for the successive rounds of PCR. The closed circles represent biotin coupled to the 5′ end of the vector primer. The short-dotted lines indicate the B1 or GT primer where dUTP was digested by uracil DNA glycosylase (UDG). The long-dotted line represents the initial product from Tag or Tag-nested primer extension. P1 is the vector-specific primer in the first round PCR, and P2, P3 is the nested vector-specific primer in the second and third round PCR, respectively. The assay consists of the following procedures. The first round of PCR is carried out between a biotinylated vector-specific primer and the B1 or GT primer in which dTTPs are replaced with dUTPs. PCR products are treated with UDG. Then, the biotinylated DNA fragments that originate from the vector primer are isolated and purified using streptavidin-coated beads. The second round of PCR is performed using the purified biotinylated strands as templates and the nested biotinylated vector primer together with a Tag primer. The third round of PCR is performed using the purified biotinylated strands from the second round PCR as templates and the nested vector primer together with a Tag-nested primer. Finally, the third round PCR products are sequenced and analyzed.

fragments. The vector primers that were used are illustrated in Fig. 1a. Six independent integration events were identified. For integration FG388 (an arbitrary name based on vector breakpoint), a 2.6 Kb product was detected after three rounds of PCR (Fig. 5d, lane 2). DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 2455–388) covalently linked to mouse genomic DNA (containing a B1 element) localized to an intergenic region on mouse chromosome 11 (Fig. 5a). For integration FG1043, a 1.8 Kb fragment was obtained after three rounds of PCR (Fig. 5e, lane 5), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 2455–1043) covalently linked to mouse DNA (containing a B1 element) localized to an intergenic region on mouse chromosome 6 (Fig. 5b). For integration FG22, a 5 Kb fragment was obtained after three rounds of PCR (Fig. 5e, lane 6), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 2455–22) covalently linked to mouse DNA (containing a B1 element) localized to an intergenic region on mouse chromosome 5 (Fig. 5c). This integration featured a rearrangement within the mouse genomic sequence: an ~2.5 Kb genomic segment was deleted at the insertion site, the genomic ends shared 6 bp of homology at the junction, and an additional 11 bp segment at the vector-mouse junction was present that was potentially generated from homologous recombination between vector ITRs (Fig. 6). It is not known if these genomic rearrangements occurred before, during, or after the vector integration. For integration FG1756, a 3 Kb product was detected after three rounds of PCR (Fig. 7d, lane 2). DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 2455–1756) covalently linked to mouse genomic DNA (containing a GT repetitive element) localized to an intergenic region on mouse chromosome 4 (Fig. 7a). For integration FG117, a 2.6 Kb fragment was obtained after three rounds of PCR (Fig. 7d, lane 2), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 2455–117) covalently linked to mouse DNA (containing a GT repetitive element) localized to an intergenic region on mouse chromosome 2 (Fig. 7b). For integration FG34388, a 1.1 Kb fragment was obtained after three rounds of PCR (Fig. 7e, lane 6), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 33,729–34,388) covalently linked to mouse DNA (containing a B1 element) localized to an intron 4 of Nsun7 gene on mouse chromosome 5 (Fig. 7c).

In the RAIC assay for the integrated HD-leptin vector, liver DNA from 6.5-weeks post-dose was purified by filter preparation, followed by four rounds of preparative PFGE, using instrument parameters optimized to resolve 20–300 Kb fragments. The vector primers that were used are illustrated in Fig. 1b. Three independent integration events were identified. For integration HD49, a 1.8 Kb product was detected after three rounds of PCR (Fig. 8d, lane 3). DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 1666–49) covalently linked to mouse genomic DNA (containing a B1 element) localized to an intergenic region on mouse chromosome 7 (Fig. 8a). For integration HD72, a 2 Kb fragment was obtained after three rounds of PCR (Fig. 8d, lane 3; same reaction as HD49), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 1666–72) covalently linked to mouse DNA (containing a B1 element) localized to an intergenic region on mouse chromosome 17 (Fig. 8b). For integration HD220, a 2.2 Kb fragment was obtained after three rounds of PCR (Fig. 8e, lane 6), and DNA sequencing confirmed that the PCR product contained vector sequences (nucleotides 1666–220) covalently linked to mouse DNA (containing a GT repetitive element) localized to an intergenic region on mouse chromosome 8 (Fig. 8c).

Sequence analyses of FG-leptin and HD-leptin insertion sites
Sequencing of the insertion sites revealed a short segment of homology (1–5 bp) between the vector and genomic DNA
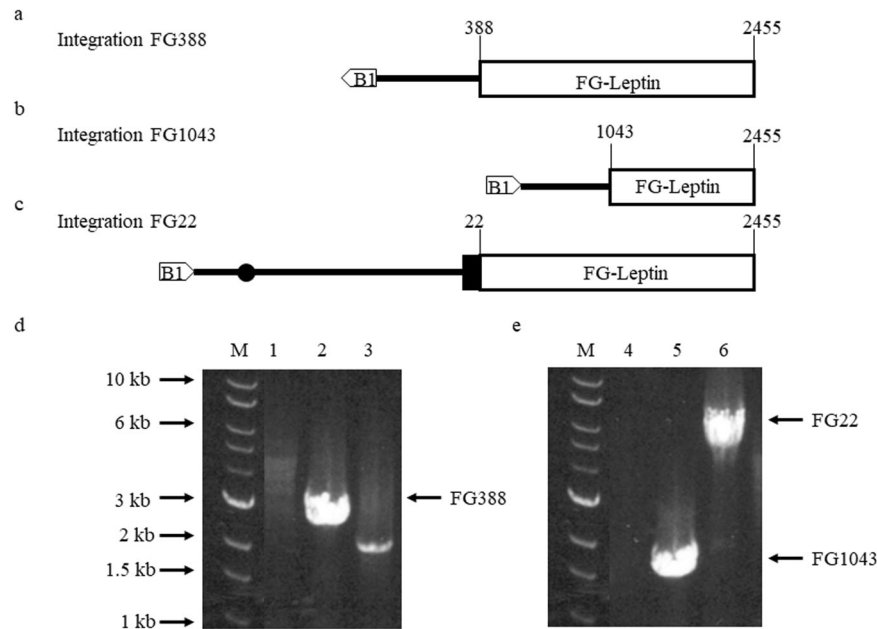
Fig. 5 Integrations in the FG-leptin-treated mouse liver DNA identified by RAIC PCR. a Structure of the junction fragment in FG388. The open box indicates vector sequence from 388 to 2455; the bold line represents genomic sequence terminating with the mouse B1 element. **b** Integration FG1043, with vector sequence from 1043 to 2455 and the genomic with B1 element. **c** Integration FG22, with vector sequence from 22 to 2455 and the genomic with B1 element. The filled circle represents 6 bp of homology at a rearrangement site within the mouse genomic DNA. The filled box represents an 11 nucleotide segment recombined from the ITR of the vector (see Fig. 6). **d, e** Third round integration junction products identified by RAIC PCR: FG388, FG1043, and FG22. Lanes M, molecular weight standards; lanes 1 and 4, 2000 copies of FG-leptin vector spiked in 0.4 μg of mouse DNA were used as a nonintegrated assay controls; lanes 2, 3, 5, and 6, 0.4 μg of the FG-leptin-treated mouse liver DNA as template. The sequence of the product in lane 3 was derived exclusively from vector.
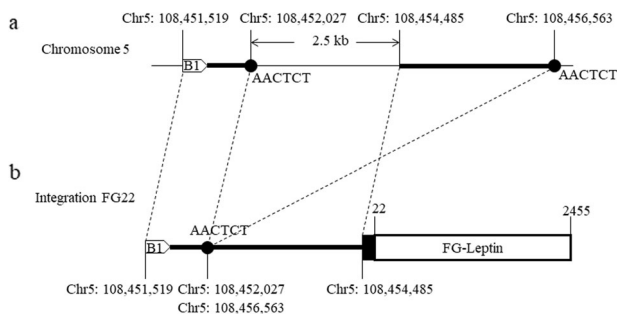
Fig. 6 Integration FG22 contained a rearrangement of both the mouse genomic and the vector DNA. a Mouse chromosome 5. the two segments of mouse sequence shown in bold line are 2.5 Kb apart, and are found in the FG22 junction fragment. The filled circle symbols 6 bp (AACTCT) of homology at the rearrangement site of the mouse genomic DNA. **b** Structure of the junction fragment in integration FG22. The open box indicates vector sequence from 22 to 2455. The bold line represents mouse sequences from the DNA segment at chromosome 5. The arrow box is the mouse B1 element. The filled box represents an 11 nucleotide segment from the vector ITR. Within the sequenced junction fragment, 2.5 Kb of genomic DNA is deleted.

sequences at the breakpoints in all integrations except integration FG22, which involved an insertion of an additional 11 bp segment at the junction site (Table 4 and Fig. 6). As described above, the chromosomal localization of each insertion indicated that all 9 integration events detected in this study occurred at different genomic sites, and all but FG22 and FG34388 were localized to different mouse chromosomes (Table 5). None of the integrations were generated by typical homologous recombination, and only one in nine was located within a gene. By viewing a 100 Kb window flanking each genomic integration site, it was determined that six of the nine integrations were within 50 Kb

of the nearest intron/exon and three of the nine integrations were farther than 50 Kb to the nearest gene (Table 5). Finally, AT-richness of the insertion sites (i.e., a 300-bp window flanking each insertion site) was examined and the average frequency of AT base pairs was 59% (Table 5), similar to the average for the entire mouse genome [19].

DISCUSSION

The enhanced persistence of HD-leptin versus FG-leptin is consistent with previous observations [3, 4, 6]. Following adenovirus-mediated gene transfer in vivo, the innate immune response to adenovirus vectors results in inflammation of transduced tissues and substantial loss of vector DNA within the first 2 days [20]. This early response occurs in response to both HD and FG vectors, as it is induced by the viral particle and is independent of viral gene transcription. The progressive clearance of adenovirus vectors is mainly attributed to the adaptive immune response that is directed against the residual expression of the viral genes that exist in FG but not HD vectors.

Southern blot analysis indicated that the FG-leptin and HD-leptin DNA persists in liver predominantly in the form of episomal monomers. To investigate the level of integrated adenovirus vector, HMW genomic DNA was purified away from free adenovirus DNA by PFGE and then assayed by Q-PCR for associated vector. Using conditions in which the HMW band of genomic DNA isolated from the gel can be considered representative of the whole genomic DNA, the level of FG-leptin and HD-leptin in liver DNA samples was reduced by 80–90%, indicating that at least 80–90% of vector persisting in liver is extrachromosomal. Some of the remaining 10–20% of the detectable vector could represent extrachromosomal vector that co-purifies with the genomic DNA (as purification procedures are rarely 100% effective). As an additional step in the analysis of integration, RE digestion was used to facilitate the further separation of free
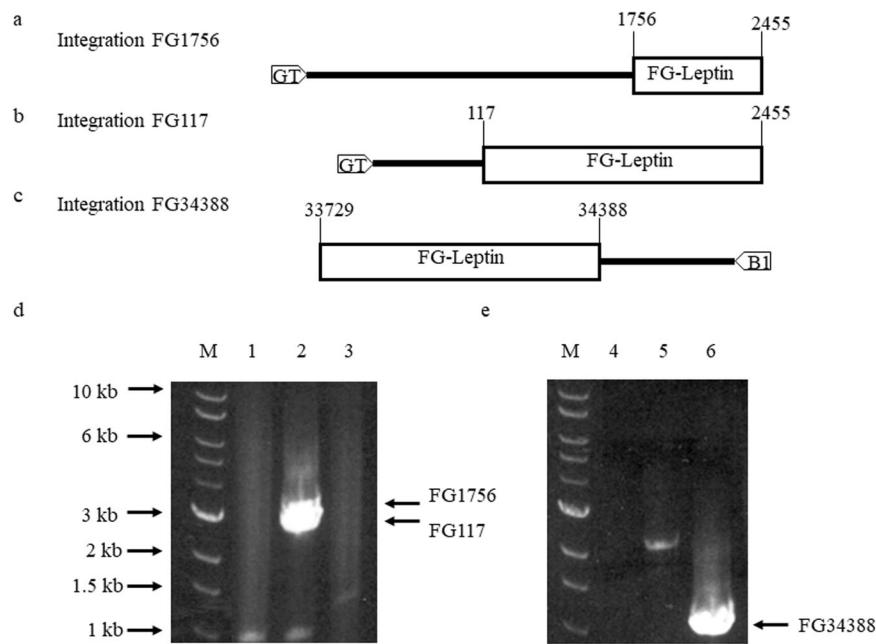
Fig. 7 Integrations in the FG-leptin-treated mouse liver DNA identified by RAIC PCR. a Structure of the junction fragment in FG1756. The open box indicates vector sequence from 1756 to 2455; the bold line represents genomic sequence terminating with the mouse GT repeat. **b** Integration FG117, with vector sequence from 117 to 2455 and the genomic with GT repeat. **c** Integration FG34388, with vector sequence from 33729 to 34388 and the genomic with B1 element. **d, e** Third round integration junction products identified by RAIC PCR: FG1756, FG117, and FG34388. Lanes M, molecular weight standards; lanes 1 and 4, 2000 copies of FG-leptin vector spiked in 0.4 µg of mouse DNA were used as a nonintegrated assay controls; lanes 2, 3, 5, and 6, 0.4 µg of the FG-leptin-treated mouse liver DNA as template. The sequence of the RAIC PCR products in lane 5 was derived exclusively from mouse genomic DNA.
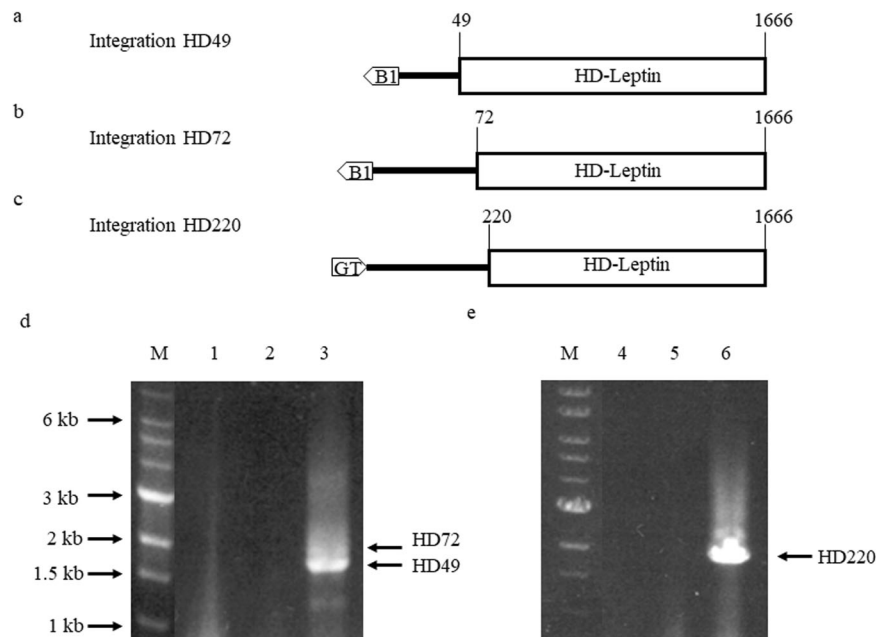
Fig. 8 Integrations in the HD-leptin-treated mouse liver DNA identified by RAIC PCR. a Structure of the junction fragment in HD49. The open box indicates vector sequence from 49 to 1666; the bold line represents genomic sequence terminating with the mouse B1 element. **b** Integration HD72, with vector sequence from 72 to 1666 and the genomic with B1 element. **c** Integration HD220, with vector sequence from 220 to 1666 and the genomic with GT repeat. **d, e** Third round integration junction products identified by RAIC PCR: HD49, HD72, and HD220. Lanes M, molecular weight standards; lanes 1 and 4, 0.1 µg of control mouse DNA as control template; lanes 2 and 5, 5000 copies of HD-leptin vector spiked in 0.1 µg of mouse DNA were used as a nonintegrated assay controls; lanes 3, and 6, 0.1 µg of the HD-leptin-treated mouse liver DNA as template.

vector from HMW genomic DNA. A rare-cutting RE that cleaves within the vector was used to digest potential vector concatemers (or trapped monomers) into shorter fragments while leaving HMW genomic DNA relatively intact, allowing for more efficient separation of HMW genomic DNA from free vector. Using RE digestion prior to gel purification, the level of FG-leptin and HD-leptin DNA was reduced to 22 and 3500 copies per µg DNA (or 150,000 diploid cells), respectively. Under the conditions used in

Table 4. DNA sequence of each junction site[a].

Sequence source	Integration HD49	Integration HD72	Integration HD220
Vector	aatatgata**ATGaggggg**tg	ttgtgacg**TGGcgcggggcg**	aca**ATTT**tcg**C**gcggtttta
Cellular	**gggattcacATG**gtgtctaa	**aagaaggaTGG**gttgaaggt	**ggaggcATTTC**ctcaactga
Junction	**gggattcacATGaggggg**tg	**aagaaggaTGGcgcggggcg**	**ggaggcATTTC**gcggtttta

Sequence source	Integration FG1043	Integration FG117	Integration FG1756
Vector	gaaatgtcat**TGatcctggt**	tgtggcggaa**GTGTgatgtt**	gggcgggggt**Cgttgggcgg**
Cellular	**ctgagaggaaTG**gcaggcaa	**gtgtttctctGTGT**tacccg	**tgaatcctgcC**aatccatga
Junction	**ctgagaggaaTGatcctggt**	**gtgtttctctGTGTgatgtt**	**tgaatcctgcCgttgggcgg**

Sequence source	Integration FG388	Integration FG34388
Vector	tagagccc**ACcgcatcccc**a	**Atatacctta**
Cellular	**ctcagtttACt**tggtaaatt	**ttactagtgaA**gttagaatt
Junction	**ctcagtttACtcgcatcccc**	**ttactagtgaAtataccttta**

Sequence source	Integration FG22
Vector	tcccttccagctctctgcccc**ttttggattgaagccaatatgataa**
Cellular	**acacacacacacacacacaca**cacacacagggcttactttatacag
Junction	**acacacacacacacacacaca**_tatataccttta_**ttttggattgaagccaatatgataa**

Bases shown in CAPITAL letters are homologous between the vector and cellular sequences, and retained in the junction sequence. Bases from the vector and cellular sequences that are shown in bold are found in the junction sequence. Additional 11 bp at junction site of integration FG22 was shown in italic letters.
[a]The table shows the vector, cellular, and the junction sequences of the nine integrations that were isolated and sequenced.

Table 5. The genomic site for integration[a].

Integrations	Chromosome localization	Distance to nearest gene	Frequency of AT[b]
FG22	Chr5: 108454485	5.6 kb to Mfsd7a	62.00%
FG117	Chr2: 94579612	>50 kb	57.00%
FG338	Chr11: 101662526	3 kb to Arl4d	55.70%
FG1043	Chr6: 22397341	41.3 kb to Fam3c	55.30%
FG1756	Chr4: 121598681	15.3 kb to Gm12887	66.30%
FG34388	Chr5: 66265313	Hit intron 4 of Nsun7	64.00%
HD49	Chr7: 30560941	0.9 kb to Psenen	53.70%
HD72	Chr17: 81940270	>50 kb	62.00%
HD220	Chr8: 30949802	>50 kb	54.30%
			58.9% (average)

[a]Mouse bioinformatics data obtained from UCSC Genome Bioinformatics Mus musculus GRCm38/mm10.
[b]Percent AT base pairs within a 300 bp window flanking the integration site.

the present study, RE digestion could cause in the loss of ~50% of the integrated vector during gel purification [17]. Thus, the frequency of integration of FG-leptin and HD-leptin could be as high as ~44 and 7000 integrations per μg DNA, respectively, assuming the residual vector copies associated with genomic DNA were integrated and not just a result of incomplete purification. Since 1 μg of cellular genomic DNA represents ~150,000 diploid genomes, this would correspond to ~0.0003 and 0.05 integrations per cell for FG-leptin and HD-leptin, respectively. Previously, an adenovirus vector expressing fumarylacetoacetate hydrolase (FAH) was found to be integrated in liver nodules which grew under selective pressure in the FAH$^{\Delta exon5}$ mouse model at a frequency of 6.72×10^{-5} per transduced hepatocyte; however, this occurred in proliferating nodules under selective pressure [14]. The integration frequency of FG-leptin is similar to that reported for rAAV vectors in vivo, which was observed to be approximately and 6×10^{-4} integrations per diploid genome in mice, although the efficiency of recovery of integration events in the method

used is not known [21]. Similarly, rAAV integration frequencies of 10^{-4}–10^{-5} integrations per diploid genome were observed in nonhuman primates [22] and humans [23].

To put the potential integration frequencies of FG-leptin and HD-leptin into perspective for risk assessment, we compare them with the spontaneous mutation frequency. The median somatic mutation frequency (single nucleotide variations) in humans is estimated to be approximately 1680 mutations per cell (2.8×10^{-7} mutations per base pair (bp) $\times 6 \times 10^9$ bp per diploid cell) [24]. With the worst-case assumption that all vector DNA associated with HMV genomic DNA is integrated, the mutation frequency due to integration of HD-leptin and FG-leptin would be 10^4- and 10^6-fold lower, respectively, than this spontaneous mutation frequency. Since integration of foreign DNA may be potentially more disruptive to gene expression than single nucleotide variation, we also make a more conservative, worst-case comparison to the spontaneous rate of gene-inactivating mutations, which in human lymphocytes is estimated to be in

the order of 10^{-6} mutations per gene copy (i.e., per one of the two copies of genes on autosomal chromosomes) [25]. Making the hypothetical worst-case assumption that every integration falls in or near a gene copy and affects its expression (which based on the characterization of nine integrations in this study is actually likely to be a rare event), and dividing integrations per cell by a conservative estimate of 40,000 gene copies per cell (2 copies × 20,000 genes per diploid cell) [26], the frequency of gene copy mutation would be ~10^{-6} or 10^{-8} for HD and FG-leptin, respectively. In this hypothetical worst-case scenario, the integration frequency observed with HD-leptin would be comparable to the spontaneous frequency of gene copy mutation, while the integration frequency observed with FG-leptin would be approximately 100-fold below the spontaneous frequency. Again, this is a worst-case comparison because many or most integrations would probably not fall within or affect a gene copy (and furthermore, affecting one copy of an autosomal gene may or may not have a phenotypic consequence since the other copy would unlikely be affected). These worst-case comparisons suggest a low risk of insertional mutagenesis due to integration of Ad5 vectors following intravenous administration.

To confirm integration and identify insertion sites, we used RAIC PCR, which was developed to detect rare, single-copy integration events in a complex mixture in vivo [18]. We identified nine independent and unique integration events in liver DNA from mice treated with FG-leptin and HD-leptin. Characterization of the insertion sites revealed several features that are consistent with illegitimate (nonhomologous) recombination [27]. Each of the nine integrations found in this study were located at different sites within the mouse genome. The junction sites identified in the present study contained microhomology between the vector breakpoint and the insertion site within the cellular DNA, resulting in deletions and/or additions at the junction sites, a feature typically observed in illegitimate integration [18, 27] including studies of Ad5 and Ad12 vectors in vitro [10, 11]. Plasmid and several viral vectors have been reported to preferentially integrate within or near transcriptionally active genes [18, 28–30]. An in vitro study has revealed that half of the adenovirus integrations took place within genes [12]. In the present study, although only nine integrations were characterized, one of the nine integrations characterized fell within a gene, which is not consistent with previously observed propensity of foreign DNA to preferentially integrate into genes.

The concern about insertional mutagenesis by DNA and viral vectors is the potential activation of oncogenes or inactivation of tumor suppressor genes, which could potentially lead to oncogenesis. Insertional oncogenesis has generally been considered a remote, theoretical concern due to the relatively low chance of hitting an oncogene or suppressor gene and the potential need for two or more "cooperative" hits based on the multi-step nature of carcinogenesis. Concern over insertional oncogenesis was originally heightened by the observation and characterization of two cases of leukemia [31], and later a third case [32], in a gene therapy trial for X-linked severe combined immunodeficiency (X-SCID). This trial, which involved ex vivo transduction of immature lymphocytes from infants with high doses of a retroviral vector containing the γc chain transgene. In the leukemia cases, it is thought that the γc chain transgene cooperated with the oncogene (the LMO2 gene in the first two cases), which had been activated by insertion of the retroviral vector LTR in a uniquely sensitive population of immature lymphocytes [33, 34]. Studies of recombinant AAV (rAAV) vectors have also raised concerns about insertional oncogenesis, as high doses of rAAV in newborn mice have been associated with the development of hepatocellular carcinoma [35]. More recently, clonal expansion of transduced liver cells with integrated vector was observed in some dogs in long-term follow-up after rAAV gene therapy for hemophilia [36]. With respect to the risk of insertional

oncogenesis by Ad5 vectors, a HD Ad5-vectored HIV vaccine was negative for tumorigenicity when tested in newborn hamsters and newborn rats at the maximum tolerated dose of 10^{10} and 10^{11} virus particles (subcutaneous), respectively [37].

This report described the integration of FG and HD Ad5 vectors in mouse genomic DNA following intravenous injection in mice, in a manner consistent with random, illegitimate recombination. With a dose of 10^{11} virus particles in mice, the frequency of integration of FG-leptin and HD-leptin was orders of magnitude lower than the spontaneous somatic mutation frequency. While this study involved intravenous injection, we have carried out studies involving intramuscular injection of Ad5-vectored HIV vaccines and obtained similar results (Troilo et al., unpublished results; briefly summarized in a meeting report from the WHO [38]).

REFERENCES

1. Hitt MM, Addison CL, Graham FL. Human adenovirus vectors for gene transfer into mammalian cells. Adv Pharmacol. 1997;40:137–206.
2. Lai CM, Lai YK, Rakoczy PE. Adenovirus and adeno-associated virus vectors. DNA Cell Biol. 2002;21:895–913.
3. St George JA. Gene therapy progress and prospects: adenoviral vectors. Gene Ther. 2003;10:1135–41.
4. Ginn SL, Anais K, Amaya AK, Ian E, Alexander IE, Edelstein M, et al. Gene therapy clinical trials worldwide to 2017: an update. J Gene Med. 2018;20:e3015. https://doi.org/10.1002/jgm.3015.
5. Morsy MA, Gu M, Motzel S, Zhao J, Jing Lin J, Su Q, et al. An adenoviral vector deleted for all viral coding sequences results in enhanced safety and extended expression of a leptin transgene. Proc Natl Acad Sci USA. 1998;95:7866–71.
6. Morsy MA, Gu MC, Zhao JZ, Holder DJ, Rogers IT, Pouch WJ, et al. Leptin gene therapy and daily protein administration: a comparative study in the ob/ob mouse. Gene Ther. 1998;5:8–18.
7. European Medicines Agency. Expert Committee on Medicinal Products Gene Therapy. Report from the CPMP Gene Therapy Expert Group Meeting 26th-27th February 2004. EMEA/CPMP/1879/04/Final.
8. Mitani K, Kubo S. Adenovirus as an integrating vector. Curr Gene Ther. 2002;2: 135–44.
9. Harui A, Suzuki S, Kochanek S, Mitani K. Frequency and stability of chromosomal integration of adenovirus vectors. J Virol. 1999;73:6141–6.
10. Hillgenberg M, Tonnies H, Strauss M. Chromosomal integration pattern of a helper-dependent minimal adenovirus vector with a selectable marker inserted into a 27.4-kilobase genomic stuffer. J Virol. 2001;75:9896–908.
11. Doerfler W. A new concept in (adenoviral) oncogenesis: integration of foreign DNA and its consequences. Biochim Biophys Acta. 1996;1288:F79–F99.
12. Stephen SL, Sivanandam VG, Kochanek S. Homologous and heterologous recombination between adenovirus vector DNA and chromosomal DNA. J Gene Med. 2008;10:1176–89.
13. Hilger-Eversheim K, Doerfler W. Clonal origin of adenovirus type 12-induced hamster tumors: nonspecific chromosomal integration sites of viral DNA. Cancer Res. 1997;57:3001–9.
14. Stephen SL, Montini E, Sivanandam VG, Al-Dhalimy M, Kestler HA, Finegold M, et al. Chromosomal integration of adenoviral vector DNA in vivo. J Virol. 2010; 84:9987–94.
15. Nichols WW, Ledwith BJ, Manam SV, Troilo PJ. Potential DNA vaccine integration into host cell genome. Ann N Y Acad Sci. 1995;772:30–39.
16. Ledwith BJ, Manam S, Troilo PJ, Barnum AB, Pauley CJ, Griffiths TG II, et al. Plasmid DNA vaccines: assay for integration into host genomic DNA. Dev Biol (Basel). 2000; 104:33–43.
17. Ledwith BJ, Manam S, Troilo PJ, Barnum AB, Pauley CJ, Griffiths TG II, et al. Plasmid DNA vaccines: investigation of integration into host cellular DNA following intramuscular injection in mice. Intervirology. 2000;43:258–72.
18. Wang Z, Troilo PJ, Wang X, Griffiths TG II, Pacchione SJ, Barnum AB, et al. Detection of integration of plasmid DNA into host genomic DNA following intramuscular injection and electroporation. Gene Ther. 2004;11:711–21.
19. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002;420:520–62.
20. Liu Q, Muruve DA. Molecular basis of the inflammatory response to adenovirus vectors. Gene Ther. 2003;10:935–40.
21. Li H, Malani N, Hamilton SR, Schlachterman A, Bussadori G, Edmonson SE, et al. Assessing the potential for AAV vector genotoxicity in a murine model. Blood. 2011;117:3311–9.

22. Nowrouzi A, Penaud-Budloo M, Kaeppel C, Appelt U, Guiner CL, Moullier P, et al. Integration frequency and intermolecular recombination of rAAV vectors in non-human primate skeletal muscle and liver. Mol Ther. 2012;20:1177–86.

23. Kaeppel C, Beattie SG, Fronza R, Logtenstein RV, Salmon F, Schmidt S, et al. A largely random AAV integration profile after LPLD gene therapy. Nat Med. 2013;19:889.

24. Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg JJ. Differences between germline and somatic mutation rates in humans and mice. Nat Commun. 2017; 8:15183.

25. Cole J, Skopek TR. International Commission for Protection Against Environmental Mutagens and Carcinogens. Working paper no. 3. Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. Mutat Res. 1994;304:33–105.

26. Willyard C. New human gene tally reignites debate. Nature. 2018;558:354–5.

27. Wurtele H, Little KC, Chartrand P. Illegitimate DNA integration in mammalian cells. Gene Ther. 2003;10:1791–9.

28. Nakai H, Montini E, Fuess S, Storm TA, Grompe M, Kay MA. AAV serotype 2 vectors preferentially integrate into active genes in mice. Nat Genet. 2003;34:297–302.

29. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002;110:521–9.

30. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003;300:1749–51.

31. McCormack MP, Rabbitts TH. Activation of the T-cell oncogene LMO2 after gene therapy for X-linked severe combined immunodeficiency. N Engl J Med. 2004; 350:913–22.

32. Kaiser J. Gene therapy. Panel urges limits on X-SCID trials. Science. 2005;307: 1544–5.

33. Insertional mutagenesis and oncogenesis: update from non-clinical and clinical studies. Gene Therapy Expert Group of the Committee for Proprietary Medical Products (CPMP), European Agency for the Evaluation of Medical Products - June 2003 meeting. J Gene Med. 2004;6:127–9.

34. Cichutek K. Lessons learned from gene therapy concerning and the use of integrating vectors and the possible risk of insertional oncogenesis. Dev Biol (Basel). 2006;123:29–34.

35. Chandler RJ, Sands MS, Venditti CP. Recombinant adeno-associated viral integration and genotoxicity: insights from animal models. Human Gene Ther. 2017; 28:314–22.

36. Nguyen GN, Everett JK, Kafle S, Roche AM, Raymond HE, Leiby J. A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells. Nat Biotech. 2021;39:47–55.

37. Ledwith BJ, Lanning CL, Gumprecht LA, Anderson CA, Coleman JB, Gatto NT, et al. Tumorigenicity assessments of Per.C6 cells and of an Ad5-vectored HIV-1 vaccine produced on this continuous cell line. Dev Biol (Basel). 2006;123:251–63.

38. WHO. Meeting Report. WHO informal consultation on characterization and quality aspect of vaccines based on live viral vectors. Geneva: World Health Organization; 2003. pp. 9–10.

AUTHOR CONTRIBUTIONS

COMPETING INTERESTS

ADDITIONAL INFORMATION