MATERIALS AND METHODS

AAV vectors: The vector genomes used in this study are shown in Figure 1. The FLAG-tagged beta-galactosidase gene (beta-galF) was produced by PCR with the reverse primer possessing a FLAG-tag sequence (5′-GATTACAAGGATG ACGACGATAAG) and previously produced pAAVbeta-gal (12) as the template. The beta-galF was inserted between the cytomegalovirus immediate early enhancer/promoter and SV40 polyA signal to produce the genome of AAV2(beta-galF). The fusion gene of EGFP and human alpha-tubulin (EGFPtub) was obtained from pEGFP-Tub (BD Bioscience Clontech, Palo Alto, Calif., USA) by cleavage with *Nhe*I and *Bam*HI. The EGFPtub was inserted between the human elongation factor 1 alpha (EF1alpha) promoter and SV40 polyA signal to produce the genome of AAV2(EGFPtub). By using PCR G at nucleotide 157 (nt157) (A at the first ATG of the EGFP gene is designated as nt1) was changed to T to produce a novel *Hin*dIII cleavage site to obtain EGFPatub. Similarly, G at nt 129 was changed to T to produce a *Hin*dIII cleavage site to obtain EGFPbtub. The amino acid sequences of EGFPatub and EGFPbtub were not affected by the nt substitutions. Each vector genome was flanked with the ITR sequence of AAV2.

beta-galF or EGFPtub was packaged into the AAV2 capsid in human 293 cells as described previously (12). Similarly, EGFPatub and EGFPbtub were packaged into the AAV10 and AAV11 capsids, respectively. The AAV2 vector stocks used in Experiments I, II, and III (Table 1) were purified by heparin affinity column chromatography (13), and the vector stocks used in Experiment IV were purified by CsCl equilibrium centrifugation (12). The infectivity of the AAV2 vector purified by CsCl centrifugation was comparable to that of the AAV2 vector purified by heparin column chromatography. Extract from 293 cells that had not been transfected with plasmids for vector production was similarly processed by heparin affinity column chromatography and was used as a mock inoculant.

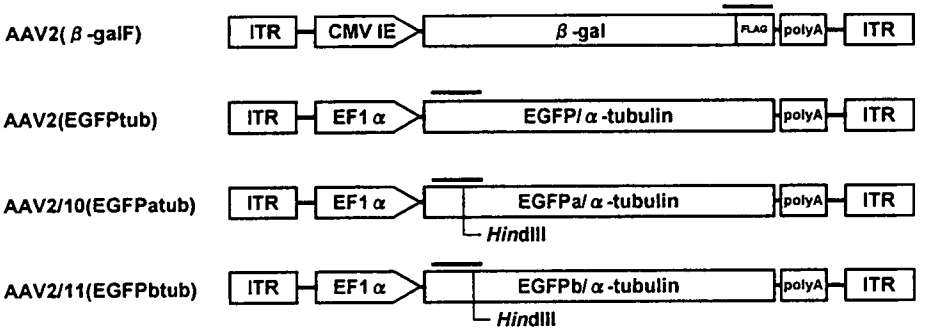

Fig. 1. Schematic representation of the vector genomes. AAV2(beta-galF) is the type 2 vector which has the FLAG-tagged beta-galactosidase gene driven by the cytomegalovirus immediate early promoter (CMV IE). AAV2(EGFPtub) is the type 2 vector which has the fusion gene of enhanced green fluorescent protein (EGFP) and alpha-tubulin driven by the human elongation factor 1 alpha promotor (EF1 alpha). AAV2/10(EGFPatub) is the type 10 pseudotyped vector which has the fusion gene of EGFPa and alpha-tubulin driven by EF1 alpha. EGFPa has a cleavage site of *Hin*dIII at the nucleotide (nt) 157 (A at the first ATG of EGFP cording region is numbered as nt 1). AAV2/11(EGFPbtub) is the type 11 pseudotyped vector which has the fusion gene of EGFPb and alpha-tubulin driven by EF1 alpha. EGFPb has a cleavage site of *Hin*dIII at the nt 124. The horizontal bars represent PCR amplicons used for detection of the vector genomes. ITR, inverted terminal repeat of AAV2; polyA, SV40 poly adenylation signal.

Table 1. Study design for vector administration in cynomolgus monkeys

Experiment	Monkey No.	Gender	AAV vector	Dose (genome copies/animal)	Time of sample collection
I	#1	FM	Mock	–	2 days
	#2	FM	AAV2(β-galF)	2.5×10^{10}	2 days
	#3	FM	AAV2(β-galF)	2.5×10^{10} [1)]	3 months
	#4	M	AAV2(β-galF)	2.5×10^{10}	3 months
	#5	FM	AAV2(β-galF)	2.5×10^{10}	3 months
II	#6	M	AAV2(EGFPtub)	2.5×10^{11}	3 months
	#7	FM	AAV2(EGFPtub)	2.5×10^{11}	3 months
	#8	M	AAV2(EGFPtub)	2.5×10^{11}	3 months
	#9	FM	AAV2(EGFPtub)	2.5×10^{11}	3 months
III	#10	FM	AAV2(EGFPtub)	1.0×10^{11}	5 months
	#11	FM	AAV2(EGFPtub)	1.0×10^{11}	5 months
	#12	M	AAV2(EGFPtub)	1.0×10^{11}	5 months
	#13	M	AAV2(EGFPtub)	1.0×10^{11}	5 months
IV	#14	FM	Mock	–	3 months
	#15	FM	AAV2, 2/10, 2/11 [2)]	1.0×10^{10} each	3 months
	#16	M	AAV2, 2/10, 2/11	1.0×10^{10} each	3 months
	#17	FM	AAV2, 2/10, 2/11	1.0×10^{10} each	7 months

[1)]: #3 was received second injection of AAV2(beta-galF) (5×10^{10} gc) at 60 days after the first injection.
[2)]: Mixture of AAV2(EGFPtub), AAV2/10(EGFPatub), and AAV2/11(EGFPbtub).

The DNase-resistant vector DNA in the vector stock was measured by Real-Time PCR (Applied Biosystems, Foster City, Calif., USA) with TaqMan probes (CCCAACGAGAAG CGCGATCACA) hybridized to EGFP DNA.

Animal experiments: Cynomolgus monkeys (4 to 5 years of age and weighing 3 to 5 kg) were obtained from the Tsukuba Primate Research Center of the National Institute of Biomedical Innovation (Ibaraki, Japan). The monkeys were sedated during all procedures by the administration of ketamine (10 mg/kg). The AAV vectors or the mock inoculant in 5 ml of physiological saline were administered intravenously into the femoral vein of the monkeys. The dose of the AAV vectors and the time of necropsy are indicated in Table 1. The monkeys were bled every 2 weeks until they were sacrificed. All animal studies were performed in accordance with the guidelines for animal experiments in National Institute of Infectious Diseases, Tokyo, Japan.

Extraction of DNA from tissues: Monkey tissues were harvested at necropsy and stored at −80°C until use. Before necropsy, blood was extensively drawn to avoid contamination of the tissues by blood. DNA was extracted from each frozen tissue type (approximately 25 mg) by using QIAamp DNA extraction kit (Qiagen GmbH, Hilden, Germany).

Detection of AAV vector DNA in tissue DNA: PCR was designed to amplify the 285-bp region of beta-galF (forward primer: 5′-GCGACTTCCAGTTCAACATC, reverse primer: 5′-TTACGCGAAATACGGGCAGA) and 323-bp region of EGFP (forward primer: 5′-ACAAGTTCAGCGTGTCCGGC, reverse primer: 5′-CCTCCTTGAAGTCGATGCCC) in the vector genomes. PCR consisted of an initial heating step at 94°C for 5 min, 37 cycles of incubation at 94°C for 30 s and at 68°C for 1 min, and incubation at 68°C for 5 min. The vector DNA fragment in the tissue DNA sample was amplified by PCR. The DNA sample contained 0.5 μg DNA (equivalent to 10^5 cells). For comparison, the DNA fragment in standard DNA solution, which contained a known amount of plasmid DNA having beta-galF or EGFPtub genes (10^2 to 10^5 copies) and DNA extracted from the liver of monkey (#1) that had received the mock inoculant were amplified by PCR in a similar manner. To verify the quality and quantity of DNA in the samples, a portion of G3PDH gene was amplified with the previously described primers (12). The PCR products were analyzed by electrophoresis on a 2% agarose gel followed by ethidium bromide staining. When the DNA sample contained more than 100 gc of the vector DNA, the amplified DNA fragment was clearly detected (Fig. 2).

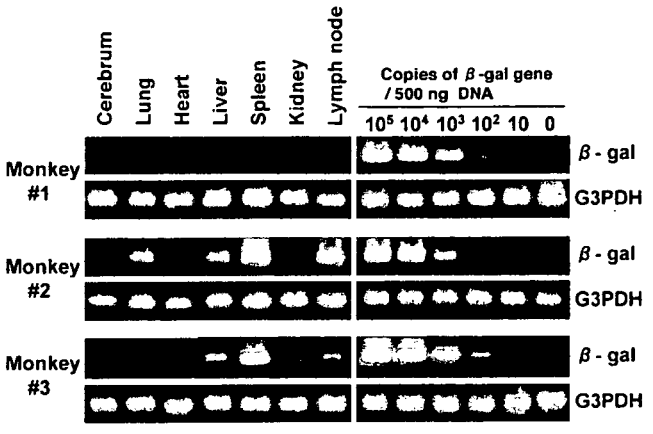

Fig. 2. Detection of AAV vector DNA in monkey tissues. DNA samples were extracted from the various tissues of monkey #1, injected with mock inoculate, and of monkeys #2 and 3, injected with AAV2(beta-galF). The segment of the vector genome in the DNA samples was amplified by PCR. G3PDH gene was amplified for references. Representative results of the agarose gel electrophoresis of the PCR products were presented.

To determine the serotype of the vector in Experiment IV, PCR products were digested with *Hin*dIII, and the size of the resultant DNA fragment was analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, Calif., USA). The *Hin*dIII digestion produced 324-bp, 243-bp, and 276-bp fragments from the genomes of AAV2(EGFPtub), AAV2/10(EGFPatub), and AAV2/11(EGFPbtub), respectively.

Immunoblot detection of EGFP/alpha-tubulin fusion protein in the tissue extracts: Monkey tissue (50 mg) was homogenized in a lysis buffer (0.1M MES, 2 mM EGTA, 1 mM $MgCl_2$, pH 6.8) and the homogenate was centrifuged at 100,000 g for 30 min at 4°C. The supernatant was mixed with Taxol (Paclitaxel; Sigma-Aldrich, St. Louis, Mo., USA) (20 μM) and GTP (Sigma-Aldrich) (0.5 mM), and was warmed to 37°C for 30 min in order to allow the tubulin to assemble into microtubules. Then, the supernatant was chilled on ice for 10 min and was centrifuged at 10,000 × g for 30 min at 4°C to precipitate the microtubules. The pellet was resuspended in PBS and electrophoresed in SDS-polyacrylamide gel. The proteins in the gel were transferred to a Hybond-P nylon membrane (Amersham Biosciences Corp., Piscataway, N.J., USA). After blocking the membrane with 5% skim milk, the EGFP/tubulin fusion protein on the membrane was allowed to bind with anti-EGFP rabbit polyclonal antibody (#632376; BD Bioscience Clontech, Palo Alto, Calif., USA) and anti-alpha-tubulin mouse monoclonal antibody (T-9026; Sigma-Aldrich). Horseradish peroxidase conjugated anti-rabbit and anti-mouse IgG goat antibodies (SC-2030 and SC-2031, respectively; Santa Cruz Biotechnology, Inc., Santa Cruz, Calif., USA) and an ECL Western Blotting Detection System (Amersham Biosciences) were used to detect rabbit and mouse IgGs on the membrane. Fluorescence was detected using a Storm Phosphor Imager (Amersham Biosciences).

A HeLa cell clone expressing EGFPtub was newly produced and used as a model for immunoblot detection of endogenous tubulin and EGFPtub. HeLa cells, cultured in growth medium (DMEM supplemented with 10% fetal bovine serum) in a 10-cm dish, were transfected with 2 μg of pEGFP-Tub (BD Bioscience Clontech) with the transfection reagent Effectene (Qiagen). After incubation of the cells for 48 h, they were passaged at a split ratio of 1 to 10 and cultured in growth medium containing a selective drug, G418 (500 μg/ml). Drug-resistant cell clones were obtained by two successive single colony isolations. One clone (HeLa/EGFPtub) that stably expressed EGFPtub was selected. Endogenous tubulin and EGFPtub were extracted from HeLa/EGFPtub and were used as markers for tubulin and EGFPtub.

Neutralization of AAV vectors with serum antibody: The neutralizing activities of the monkey serum were examined by testing the inhibition of the transduction of COS-1 cells by AAV vectors expressing beta-gal. COS-1 cells (2×10^4 cells/well) were seeded in 96-well plates at 6 h before inoculation. Fifteen microliters of vector solution containing 4×10^3 transducing unit/ml was mixed with 18 μl of serum that had been serially diluted from 1:10 to 1:6,250 with PBS and incubated for 1 h at 37°C. The number of vector genome copies required for 4×10^3 transducing unit with COS-1 were 10^6 gc, 10^8 gc, and 10^8 gc for the AAV2,

AAV2/10, and AAV2/11 vectors, respectively. Then, the sample was mixed with 30 μl of DMEM containing 10% FBS and was used to inoculate the cells in two wells (30 μl/well). After incubation of the cells with occasional rocking for 2 h at 37°C, 70 μl of fresh DMEM containing 10% FBS was added to each well. Two days later, the cells were fixed and stained using an In Situ beta-Galactosidase Staining Kit (Stratagene, La Jolla, Calif., USA). The cells that showed positive staining for beta-gal were counted under a microscope. The neutralizing titer of the antiserum was expressed as the reciprocal of the highest dilution that repressed the number of beta-gal-positive cells to half of the number obtained with the samples mixed with similarly diluted serum from a non-immunized mouse.

RESULTS

Distribution of AAV2(beta-galF) in cynomolgus monkeys: In Experiment I (Tables 1 and 2), one (#1) and four monkeys (#2, 3, 4, and 5) were intravenously injected into the femoral vein with a mock inoculate and 2.5×10^{10} gc of AAV2(beta-galF) (Fig. 1), respectively. No clinical symptoms were observed among these monkeys, indicating that the AAV2 vector particle did not induce any acute toxic effects, as has also been reported previously (4,14). In urine collected between 0 and 24 h post-inoculation (pi), a very low level of the vector DNA was found in monkey #2 sample (10^2 gc/30 μl), but this was not observed in the samples from the other monkeys. The vector DNA was not detected in the urine collected between 24 and 48 h pi from any of the five monkeys. The results indicated that excretion of the injected AAV2 vector was very limited. In whole blood samples obtained at 1 day pi, a very low level of the vector DNA was found in samples from monkeys #2, 3, 4, and 5 (10^2 gc/10 μl). Since the body weight of these monkeys was approximately 3 kg, the total blood of each monkey could be estimated as 250 ml. Therefore, only 1/10,000 of the vectors initially administered is thought to have remained in the circulating blood at 1 day pi. In whole blood samples from monkeys #3, 4, and 5 obtained at 1 week pi, no vector DNA was detected (data not shown).

Monkeys #1 and 2 were sacrificed at 2 days pi. The vector DNA was found in none of the monkey #1 samples, and it was found in various tissues of monkey #2 (Fig. 2 and Table 2). Relatively high levels of the vector DNA were observed in the spleen, tonsil, and axillary lymph node. Low levels of the vector DNA were detected in the brain, ovary, and uterus. Histological examination of formalin-fixed specimens of tissues positive for the vector DNA did not show any abnormalities or inflammatory reactions.

At 2 months pi, monkey #3 received a second intravenous injection with 5×10^{10} gc of AAV2(beta-galF). No clinical symptoms were observed after the injection, indicating that the second injection did not induce any strong acute abnormal immunological reactions.

Monkeys #3, 4, and 5 were sacrificed at 3 months pi. During the 3-month period, the hematological profiles were examined periodically, and no abnormalities were found in the blood samples of monkeys #3, 4, and 5 (data not shown). The vector DNA was present in various tissues (Fig. 2 and Table 2). The highest vector DNA level was detected in the spleen of monkey #3, which had been injected twice with the vector. Among monkeys #3, 4, and 5, the vector distribution pattern to tissues appeared to vary from monkey to monkey. For example, the vector DNA was not detected in the spleen of monkey #5. It was noteworthy that, the vector DNA was found in the ovaries of monkeys #3 and 5, albeit at low levels. The detection of the vector DNA in the spleen, lymph node, and ovary sections by in situ hybridization with PCR was not successful, because it was difficult to find sections positive for the vector DNA. Thus, no vector DNA-positive cell species within the tissues were identified.

Because the vector DNA was readily detected in lymphoid tissues such as the spleen and lymph nodes, we examined the susceptibility of peripheral blood mononuclear cells (PBMCs) to the vector. PBMCs were collected from a cynomolgus monkey that had received no AAV vectors before and were incubated with an AAV2 vector. PBMCs (10^5) or COS-1 cells (5×10^4) in the wells of a 48-well plate were inoculated with 10^9 genome copies of AAV2(EGFP-tub). Three days later, the cells were examined for EGFP expression under a fluorescence microscope. Whereas almost all of the COS-1 cells were positive for EGFP expression, none of the PBMCs were positive for EGFP. It is possible that following the low-dose intravenous injection to the monkeys, the vector did not infect the lymphocytes in the spleen and lymph node.

Table 2. Experiment I: Detection of the vector genomes in the various tissues. Monkeys injected with AAV2(beta-galF) (2.5×10^{10} gc) were sacrificed at 2 days or 3 months after the injection

Tissue	Monkey				
	#1 (Mock)	#2 (2d)	#4 (3m)	#5 (3m)	#3 (3m)[1)]
Cerebrum	–	+	+	–	–
Cerebellum	–	+	+	+	+
Bone marrow	–	++	+	–	–
Retina	–	–	–	–	ND
Skin	–	–	+	–	+
Muscle	–	–	–	–	–
Trachea	–	+	–	+	+
Lung	–	++	–	–	+
Heart	–	–	+	–	–
Liver	–	++	–	–	+
Gallbladder	–	++	+	–	+
Pancreas	–	+	+	+	–
Spleen	–	++++	++	–	+++
Esophagus	–	+	–	–	–
Stomach	–	+	–	–	+
Jejunum	–	ND	–	+	+
Ileum	–	+	+	+	–
Colon	–	+	+	+	ND
Kidney	–	–	+	–	+
Adrenal gland	–	–	–	–	–
Bladder	–	+	–	ND	+
Tonsil	–	+++	+	–	+
Thymus	–	+	–	–	–
Parotid gland	–	+	–	–	–
Submandibular gland	–	+	–	–	–
Thyroid gland	–	++	+	–	+
Axillary lymph node	–	+++	+	–	–
Hilar lymph node	–	+	ND	–	+
Mesenteric lymph node	–	++	–	–	+
Iliac lymph node	–	++	ND	ND	ND
Inguinal lymph node	–	+	+	–	+
Testis/Ovary	–	+	–	+	+
Epididymis/Uterus	–	+	ND	+	–

[1)]: #3 was received second injection of AAV2(beta-galF) (5×10^{10} gc) at 60 days after the first injection. (–), $<10^2$ gc/0.5 μgDNA; (+), 10^2-10^3 gc/0.5 μgDNA; (++), 10^3-10^4 gc/0.5 μgDNA; (+++), 10^4-10^5 gc/0.5 μgDNA; (++++), $>10^5$ gc/0.5 μgDNA; (ND), Not done.

Distribution of AAV2(EGFPtub) in cynomolgus monkeys and expression of EGFPtub: In Experiment II, 2.5×10^{11} gc of AAV2(EGFPtub) (Fig. 1) was administered into four monkeys (#6, 7, 8, 9) intravenously into the femoral vein, and the monkeys were sacrificed at 3 months pi (Tables 1 and 3). The vector DNA in various tissues was examined by similar procedures to those used in Experiment I. With some variation from monkey to monkey, high levels of the vector DNA (>10^3 gc/0.5 μg DNA) tended to be detected in the spleen, liver, gallbladder, tonsils, and lymph nodes, and low levels of the vector DNA (<10^3 gc/0.5 μg DNA) were present in the cerebrum, bone marrow, muscle, trachea, lung, heart, pancreas, esophagus, colon, kidney, adrenal gland, bladder, and parotid gland. At 3 months after injection, the vector DNA was not detected in the blood samples. Therefore, we concluded that the vector DNA in these specimens was not derived from contamination with blood.

EGFPtub expression was detected in the axillary lymph nodes of monkeys #7, 8, and 9 (Fig. 3). Endogenous tubulin and EGFPtub proteins were extracted from the liver, spleen, tonsils, and axillary lymph nodes of the monkeys. The two proteins were readily co-purified and concentrated by a previously described method (15). The tubulin/EGFPtub complex extracted from approximately 10^7 cells was analysed by SDS-polyacrylamide gel electrophoresis followed by immunoblotting with a mixture of anti-EGFP and anti-alpha-tubulin antibodies. The tubulin/EGFPtub complex extracted from a newly produced HeLa cell clone expressing EGFPtub was used as a size-marker for endogenous tubulin and EGFPtub protein. Comparison with commercially available purified tubulin (Sigma-Aldrich) enabled us to estimate the approximate amounts of EGFPtub (Fig. 3A). Samples for immunoblotting (50 mg) contained approximately 10 to 20 ng of EGFPtub. Under these conditions, transgene expression was detected in 6 out of 23 axillary lymph nodes tested. No transgene expression was detected in the other tissues including the spleen and liver, in which the injected vector DNA was readily found, most likely because the analyzable volume of a single piece of tissue is limited. Therefore, it remains unclear whether the transgene was not expressed in most of the tissues or the cells expressing the transgene were escaped from the sampling.

Table 3. Experiment II: Detection of the vector genomes in the various tissues. Monkeys injected with AAV2(EGFPtub) (2.5×10^{11} gc) were sacrificed at 3 months after the injection

Tissue	Monkey			
	#6 (3m)	#7 (3m)	#8 (3m)	#9 (3m)
Cerebrum	+	–	–	–
Cerebellum	–	–	–	–
Spinal cord	–	–	–	–
Bone marrow	++	++	–	–
Skin	–	–	–	–
Muscle	++	++	–	–
Trachea	+	++	+	–
Lung	–	–	–	+
Heart	+	+	–	+
Liver	+++	+++	–	+
Gallbladder	+	+	–	+++
Pancreas	+	+	–	–
Spleen	++++	++++	++	+++
Esophagus	+	–	–	–
Stomach	–	–	–	–
Jejunum	–	–	–	–
Ileum	–	–	–	–
Colon	+	–	–	–
Kidney	+	–	+	+
Adrenal gland	+	–	–	++
Bladder	+	–	–	–
Tonsil	++	+	+++	++
Thymus	–	–	–	–
Parotid gland	–	–	–	++
Submandibular gland	–	–	–	–
Thyroid gland	–	–	–	–
Axillary lymph node	++	++	+	++
Hilar lymph node	++	++++	++	–
Mesenteric lymph node	++	+	+	–
Iliac lymph node	++	+++	++	–
Inguinal lymph node	+	++	+++	–
Testis/Ovary	–	–	–	–
Epididymis/Uterus	–	–	–	–

(–), <10^2 gc/0.5 μgDNA; (+), 10^2-10^3 gc/0.5 μgDNA; (++), 10^3-10^4 gc/0.5 μgDNA; (+++), 10^4-10^5 gc/0.5 μgDNA; (++++), >10^5 gc/0.5 μgDNA.

In Experiment III 1.0×10^{11} gc of AAV2(EGFPtub) (Fig. 1) was administered to four monkeys (#10, 11, 12, 13), as was done in Experiment II, and the vector DNA in various tissues was examined at 5 months pi (Table 4). The vector distribution was similar to that observed at 3 months pi, indicating the vector DNA was stably maintained in various tissues. The levels of vector DNA in the tissues were similar to those observed at 3 months pi (Experiment II).

Distribution of AAV2/10(EGFPatub) and AAV2/11 (EGFPbtub) in cynomolgus monkeys: Since the data in the several recent studies indicated that the AAVs of different serotypes have different tissue tropism, in Experiment IV, the AAV2/10 and AAV2/11 pseudotype vectors (Fig. 1) and the AAV2 vector were compared as to the vector DNA distribution pattern in monkeys (Table 5). The entire coding regions of

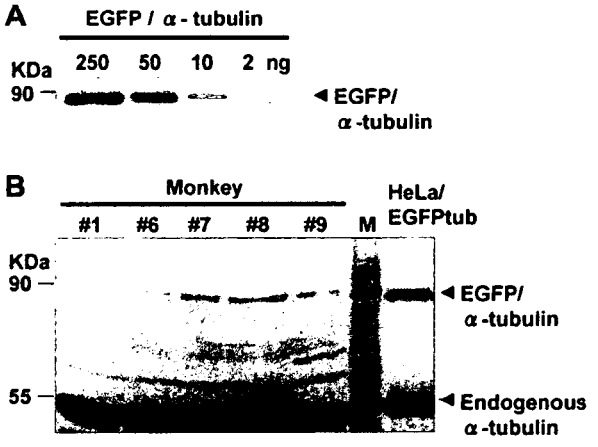

Fig. 3. Detection of transgene expression in monkey tissues. (A) Sensitivity of immunoblotting for fusion protein of EGFP and alpha-tubulin (EGFP/alpha-tubulin). The complex of endogenous tubulin and EGFP/alpha-tubulin was extracted from HeLa cells expressing EGFP/alpha-tubulin. The extracts were electrophoresed in a SDS-polyacrylamide gel and transferred to a nylon membrane. The EGFP/alpha-tubulin was detected with a mixture of anti-EGFP and anti-alpha-tubulin antibodies. The amount of EGFP/alpha-tubulin was estimated by the comparison with known amount of commercially available purified tubulin. (B) Expression of EGFP/alpha-tubulin in the lymph node of monkeys. The complex of endogenous tubulin and EGFP/alpha-tubulin was extracted from the axillary lymph node of monkeys, injected with mock inoculate (monkey #1) or AAV2(EGFPtub) (monkeys #6, 7, 8, and 9). The EGFP/alpha-tubulin was detected as described above.

AAV10 and AAV11 were recently isolated from cynomolgus monkeys, and the pseudotyped vectors were produced as previously described (12). To compare the distribution of the three vectors in the same monkey, AAV2(EGFPtub), AAV2/10(EGFPatub), and AAV2/11(EGFPbtub) were mixed together, and a mixture containing 1.0×10^{10} gc of each vector was administered intravenously to three monkeys (#15, 16, 17). One monkey (#14) received saline as a negative control. The PCR DNA derived from three vector genomes are distinguishable by measuring the sizes of DNA fragments produced by the digestion of the PCR DNA with *Hin*dIII, because the EGFP genes of AAV2/10(EGFPatub) and AAV2/11(EGFPbtub) have the *Hin*dIII site at different positions (Fig. 1). The vector inoculations did not induce any clinical symptoms, thus indicating that the AAV10 and AAV11 capsids exerted no acute toxicity. Serum samples were collected 1 week before and 1, 5, 9, and 11 weeks pi. Three monkeys (#14, 15, 16) were sacrificed at 3 months pi, and one monkey (#17) was sacrificed at 7 months pi.

The AAV2/10(EGFPatub) DNA was found primarily in lymphoid tissues such as the spleen and lymph nodes (Table 5). The levels of the AAV2/10(EGFPatub) DNA in the lymphoid tissues at 3 months pi (monkeys #15 and 16) were similar to those at 7 months pi (monkey #17), suggesting that the vector genome was stably maintained. The AAV2/11 (EGFPbtub) DNA was similarly found in the lymphoid tissues of monkeys #15 and 17, but was not found in any tissues from monkey #16 (Table 5). Consistent with our previous observation that the vector DNA was not found in the liver of mice intravenously injected with the AAV2/11 pseudotyped vector (12), no AAV2/11(EGFPbtub) DNA was found in the livers of monkeys #15 and 17.

Table 4. Experiment III: Detection of the vector genomes in the various tissues. Monkeys injected with AAV2(EGFPtub) (1.0×10^{11} gc) were sacrificed at 5 months after the injection

Tissue	Monkey			
	#10 (5m)	#11 (5m)	#12 (5m)	#13 (5m)
Cerebrum	–	–	–	–
Cerebellum	–	–	–	–
Bone marrow	–	–	–	+
Skin	–	–	–	–
Muscle	–	+	–	+
Trachea	–	–	–	–
Lung	+	+	–	+
Heart	+	+	–	–
Liver	++	+++	–	+++
Gallbladder	–	++	–	–
Pancreas	++	–	+	–
Spleen	++	++++	+++	+++
Esophagus	–	–	–	–
Stomach	–	–	–	–
Jejunum	–	–	–	–
Ileum	–	+	–	–
Colon	–	+	–	+
Kidney	–	+	+	+
Adrenal gland	–	+	–	+
Bladder	–	+	–	–
Tonsil	+	+	++	–
Thymus	–	+	–	+
Parotid gland	–	–	–	–
Submandibular gland	–	–	+	–
Thyroid gland	–	+++	–	–
Axillary lymph node	–	++	++++	+
Mesenteric lymph node	–	+	+	–
Iliac lymph node	–	+++	–	–
Inguinal lymph node	+	+++	++	–
Testis/Ovary	–	–	–	–
Epididymis/Uterus	–	+	–	–

(–), $<10^2$ gc/0.5 μgDNA; (+), 10^2-10^3 gc/0.5 μgDNA; (++), 10^3-10^4 gc/0.5 μgDNA; (+++), 10^4-10^5 gc/0.5 μgDNA; (++++), $>10^5$ gc/0.5 μgDNA.

Immune responses to the vectors varied from monkey to monkey. Since mouse anti-AAV2, AAV10, and AAV11 VP2 sera neutralized the AAV2, 10, and 11 vectors in a type-specific manner (12), it is very likely that the neutralizing activity of monkey antibody against AAV2, 10, and 11 is type-specific. Since it is possible that the monkeys had been infected with AAVs immunologically cross-reactive to AAV2, 10, and 11, the pre-administration serum obtained 1 week before the inoculation was considered as the baseline antibody titer (Fig. 4). Monkey #15 developed anti-AAV11 and low-level anti-AAV10 neutralizing antibodies, but did not develop anti-AAV2 antibody (Fig. 4), although similar levels of AAV2 (EGFPtub), AAV2/10(EGFPatub), and AAV2/11(EGFPbtub) DNAs were found in various lymphoid tissues (Table 5). The low level of anti-AAV11 neutralizing antibody that was present in the pre-administration serum was not found to inhibit the distribution of AAV2/11(EGFPbtub) (Table 5).

Monkey #16 developed anti-AAV2 and anti-AAV10 neutralizing antibodies, but did not develop anti-AAV11 antibody (Fig. 4). As was the case with monkey #15, the low level neutralizing antibody against AAV10 in the pre-administration serum (Fig. 4) did not inhibit the distribution of AAV2/10(EGFPatub) (Table 5). Since no anti-AAV11 neutralizing antibody was detected in the pre-administration serum of monkey #16, the absence of AAV2/11(EGFPbtub) DNA in monkey #16 (Table 5) could not be explained by the neutralization of the vector.

Monkey #17 responded to the three vector capsids and developed antibodies against AAV2, 10, and 11. It is likely that low doses of the vectors induced various immune responses of the monkeys against the vectors.

DISCUSSION

In this study, we intravenously administered AAV vectors (AAV2, AAV2/10, and AAV2/11) at a dose of 2×10^9 gc to 5×10^{10} gc/kg weight to cynomolgus monkeys, and we examined the behavior of the vectors and the responses of the monkeys. We assumed that intravenous injections at a relatively low dose of vector would mimic systemic conditions caused by targeted high-dose injections. We found that vector DNA persists for a long period of time, probably without replication, in various tissues, and in particular in the lymphatic tissues. We also demonstrated that the transgene is expressed in the axillary lymph nodes. These data are pertinent for the assessment of vector safety, although the status of the persisting vector DNA remains unclear, and the factors that regulate the transgene expression must still be investigated.

Following AAV2 vector inoculation, rapid excretion of the vector in the urine was very limited. By 1 week pi, the AAV2 vector DNA became undetectable in the circulating blood. At 2 days pi, the vector DNA was readily detected in the lymphoid tissues, especially in the spleen (Table 2). The vector DNA was detected at lower levels in the brain, lung,

Table 5. Experiment IV: Serotype-specifc detection of the vector genomes in the various tissues. Monkeys injected with the mixture of AAV2(EGFPtub), AAV2/10(EGFPatub) and AAV2/11(EGFPbtub) (1.0×10^{10} gc each) were sacrificed at 5 or 7 months after the injection

Tissue	Monkey								
	#15 (3m)			#16 (3m)			#17 (3m)		
	2	2/10	2/11	2	2/10	2/11	2	2/10	2/11
Cerebrum	–	–	–	–	–	–	–	–	–
Cerebellum	–	–	–	–	–	–	–	–	–
Bone marrow	–	–	–	–	–	–	–	–	–
Skin	–	–	–	–	–	–	–	+	–
Retina	–	–	–	–	–	–	+	–	–
Muscle	–	–	–	–	–	–	–	–	–
Trachea	–	–	–	–	–	–	–	–	–
Lung	–	–	–	–	–	–	–	–	–
Heart	–	–	–	–	–	–	–	–	–
Liver	+	+	–	–	–	–	–	–	–
Gallbladder	–	–	–	–	–	–	–	–	–
Pancreas	–	–	–	–	–	–	–	–	–
Spleen	++	++	++++	++	++++	–	++	+++	++
Esophagus	–	–	–	–	–	–	–	–	–
Stomach	–	–	–	–	–	–	–	–	–
Jejunum	–	–	–	–	–	–	–	–	–
Ileum	–	–	–	+	+	–	–	–	–
Colon	+	–	–	+	+	–	+	–	–
Kidney	–	–	–	–	–	–	–	–	–
Adrenal gland	–	–	–	–	–	–	–	–	–
Bladder	++	++	+++	–	–	–	–	–	–
Tonsil	+	++	++	+	++	–	–	–	–
Thymus	–	–	–	–	–	–	–	–	–
Parotid gland	–	–	–	–	+	–	–	+	–
Submandibular gland	–	–	–	+	+++	–	–	–	–
Thyroid gland	–	–	–	–	–	–	–	–	–
Axillary lymph node	+	++	++	+	+++	–	+	++	+
Hilar lymph node	–	++	+	+	++	–	+	++	–
Mesenteric lymph node	+	+	+	–	+	–	++	++	–
Iliac lymph node	–	–	–	+	++	–	++	++	+
Inguinal lymph node	+	++	++	+	++	–	–	+	–
Testis/Ovary	–	–	–	–	–	–	–	–	–
Epididymis/Uterus	–	–	–	–	–	–	–	–	–

(–), <10^2 gc/0.5 μgDNA; (+), 10^2-10^3 gc/0.5 μgDNA; (++), 10^3-10^4 gc/0.5 μgDNA; (+++), 10^4-10^5 gc/0.5 μgDNA; (++++), >10^5 gc/0.5 μgDNA.

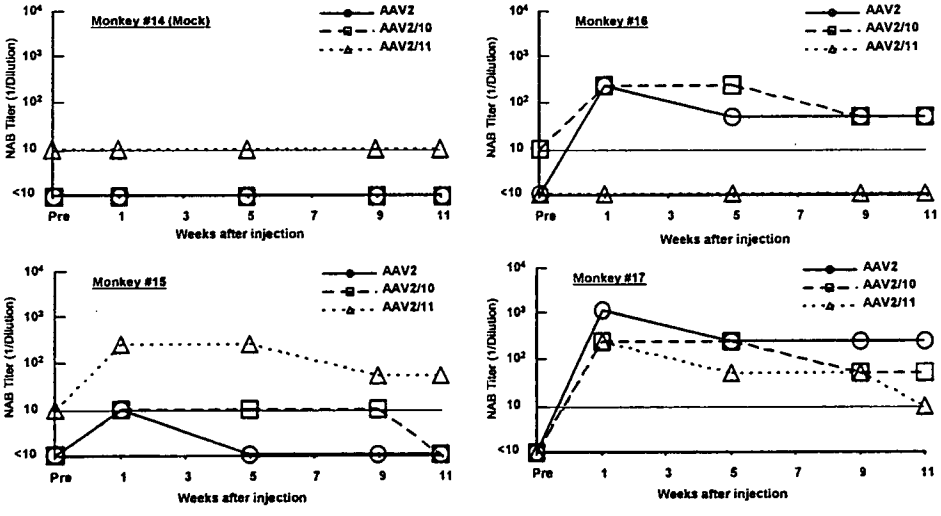

Fig. 4. Induction of anti-AAV neutralizing antibodies in monkeys. The serum from the monkey injected with saline (monkey #14) and the sera from those infected with the mixture of AAV2(EGFPtub), AAV2/10(EGFPatub), and AAV2/11(EGFPbtub) (monkeys #15, 16, and 17) were examined for their neutralizing activities by inhibition of transduction of COS-1 cells by the AAV vectors expressing beta-gal. Neutralizing titer (NAB Titer) of the antiserum was expressed as the reciprocal of the highest dilution that repressed the number of beta-gal positive cells to half of the number obtained with the samples mixed with the similarly diluted serum from a non-immunized mouse.

liver, gallbladder, pancreas, colon, ovary, uterus, and other organs (Table 2). The transfer of AAV2 vector to the brain via the blood-brain barrier agreed with the results of previous studies; when AAV2 vector has been injected into the liver or lung of non-human primates, vector DNA has been detected in the brain (6,8). No histological abnormalities or inflammatory reactions were observed in the tissues positive for the vector DNA. These results indicated that the vector in the blood was not readily excreted in the urine and rapidly attached to (or trapped by) various tissues, most readily the lymphoid tissues, without inducing any clinical symptoms. The presence of the vector DNA in the lymphoid tissues was consistent with previous findings; when an AAV2 vector containing the EGFP gene was injected into the liver of rhesus monkey fetuses, 0.01 - 0.05% of the monocytes in the spleen and lymph nodes of the infants were EGFP-positive (8). However, such EGFP-positive cells have not yet been characterized in detail.

Although the distribution patterns of the AAV2 vector at 3, 5, and 7 months pi varied from monkey to monkey (Tables 3, 4, and 5), the vector DNA was detected in various tissues of all of the monkeys sacrificed in this study. There was a tendency that from 2 days to 3 months pi, the vector DNA level was lowered to 1/100 in the spleen and to 1/2 in the other tissues. From 3 months pi to 5 months pi, the decrease in the vector DNA in the tissues, including that in the spleen, was marginal. The data strongly suggest that the AAV2 vector DNA would be maintained in various tissues for a long period of time, probably for years, without inducing any clinical symptoms.

The expression of the transgene was observed, at least in the lymph nodes, at 3 months pi. Although this remains to be determined, it is possible that the transgene could be expressed in other tissues harboring the vector DNA at a lower level. Then, the safety of the in vivo administration of AAV vectors would be affected by the properties of the transgene product.

AAV2/10(EGFPatub), AAV2/11(EGFPbtub), and AAV2 (EGFPtub) showed a similar pattern of vector distribution throughout the monkey tissues (Table 5). Previously, AAV2/10 and 2/11 vector DNAs were detected in the muscle tissues of BALB/c mice injected with AAV2/10 and 2/11 vectors (2×10^{11} gc/kg weight) via the tail vein at 6 weeks pi, which suggested that these vectors preferentially enter muscle cells (12). However, no AAV2/10(EGFPatub) and AAV2/11 (EGFPbtub) vector DNAs were detected in the muscle tissues of monkeys #15, 16, and 17 (Table 5). This apparent discrepancy may be ascribed to the difficulty of detecting vector DNA in the muscles of monkeys administered with a low vector dose (2×10^{9} gc/kg weight) such as that used in this study.

The vector distribution obtained in this study is consistent with those reported in previous studies, in which AAV vectors have been instilled in the bronchial epithelium of rhesus monkeys (6) and have been injected into the liver of rhesus fetuses (8) and injected into the muscle of rhesus and cynomolgus monkeys (7). That is, AAV vectors that enter the circulating bloodstream appear to be distributed to various tissues, primarily the lymphoid tissues, and are maintained for a long period of time. Since a small portion of the tissues was subjected to examination for the presence of vector DNA and transgene products, it cannot be ruled out that the vector genome and the transgene products were present in other portions of tissues. Furthermore, extensive histological studies of the tissues and organs will be necessary in order to define the vector DNA-positive cell-species (e.g., liver cells, lung cells, or lymphocytes) and to accurately identify those cells targeted by the vector. In addition, the long-term consequences of vector persistence remain unknown.

The distribution and persistence of the AAV vectors varied from monkey to monkey. This variation is thought to be associated, at least in part, with the genetic heterogeneity of the monkeys. More data with other monkeys should be gathered in order to evaluate safety issues related to gene therapy using AAV vectors.

ACKNOWLEDGMENTS

We thank Dr. Kunito Yoshiike for his critical reading of the manuscript.

This work was supported by a grant-in-aid from the Ministry of Health, Labour and Welfare for the Research on Human Genome, Tissue Engineering.

REFERENCES

1. Muzyczka, N. and Berns, K.I. (2001): Parvoviridae: The viruses and their replication. p. 2327-2359. *In* D.M. Knipe and P.M. Howley (ed.), Fields Virology. vol. 2. Lippincott Williams & Wilkins, Philadelphia.
2. Zolotukhin, S. (2005): Production of recombinant adeno-associated virus vectors. Hum. Gene Ther., 16, 551-557.
3. Grimm, D. and Kay, M.A. (2003): From virus evolution to vector revolution: use of naturally occurring serotypes of adeno-associated virus (AAV) as novel vectors for human gene therapy. Curr. Gene Ther., 3, 281-304.
4. Tenenbaum, L., Lehtonen, E. and Monahan, P.E. (2003): Evaluation of risks related to the use of adeno-associated virus-based vectors. Curr. Gene Ther., 3, 545-565.
5. McCarty, D.M., Young, S.M., Jr. and Samulski, R.J. (2004): Integration of adeno-associated virus (AAV) and recombinant AAV vectors. Annu. Rev. Genet., 38, 819-845.
6. Conrad, C.K., Allen, S.S., Afione, S.A., Reynolds, T.C., Beck, S.E., Fee-Maki, M., Barrazza-Ortiz, X., Adams, R., Askin, F.B., Carter, B.J., Guggino, W.B. and Flotte, T.R. (1996): Safety of single-dose administration of an adeno-associated virus (AAV)-CFTR vector in the primate lung. Gene Ther., 3, 658-668.
7. Favre, D., Provost, N., Blouin, V., Blancho, G., Cherel, Y., Salvetti, A. and Moullier, P. (2001): Immediate and long-term safety of recombinant adeno-associated virus injection into the nonhuman primate muscle. Mol. Ther., 4, 559-566.
8. Lai, L., Davison, B.B., Veazey, R.S., Fisher, K.J. and Baskin, G.B. (2002): A preliminary evaluation of recombinant adeno-associated virus biodistribution in rhesus monkeys after intrahepatic inoculation in utero. Hum. Gene Ther., 13, 2027-2039.
9. Nathwani, A.C., Davidoff, A.M., Hanawa, H., Hu, Y., Hoffer, F.A., Nikanorov, A., Slaughter, C., Ng, C.Y., Zhou, J., Lozier, J.N., Mandrell, T.D., Vanin, E.F. and Nienhuis, A.W. (2002): Sustained high-level expression of human factor IX (hFIX) after liver-targeted delivery of recombinant adeno-associated virus encoding the hFIX gene in rhesus macaques. Blood, 100, 1662-1669.
10. Davidoff, A.M., Gray, J.T., Ng, C.Y., Zhang, Y., Zhou, J., Spence, Y., Bakar, Y. and Nathwani, A.C. (2005): Comparison of the ability of adeno-associated viral

vectors pseudotyped with serotype 2, 5, and 8 capsid proteins to mediate efficient transduction of the liver in murine and nonhuman primate models. Mol. Ther., 11, 875-888.
11. Provost, N., Le Meur, G., Weber, M., Mendes-Madeira, A., Podevin, G., Cherel, Y., Colle, M.A., Deschamps, J.Y., Moullier, P. and Rolling, F. (2005): Biodistribution of rAAV vectors following intraocular administration: evidence for the presence and persistence of vector DNA in the optic nerve and in the brain. Mol. Ther., 11, 275-283.
12. Mori, S., Wang, L., Takeuchi, T. and Kanda, T. (2004): Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. Virology, 330, 375-383.
13. Auricchio, A., Hildinger, M., O'Connor, E., Gao, G.P. and Wilson, J.M. (2001): Isolation of highly infectious and pure adeno-associated virus type 2 vectors with a single-step gravity-flow column. Hum. Gene Ther., 12, 71-76.
14. Zaiss, A.K. and Muruve, D.A. (2005): Immune responses to adeno-associated virus vectors. Curr. Gene Ther., 5, 323-331.
15. Vallee, R.B. and Bloom, G.S. (1983): Isolation of sea urchin egg microtubules with taxol and identification of mitotic spindle microtubule-associated proteins with monoclonal antibodies. Proc. Natl. Acad. Sci. USA, 80, 6259-6263.

Chapter 2

Targeted Insertion of Transgene into a Specific Site on Chromosome 19 by Using Adeno-Associated Virus Integration Machinery

Masashi Urabe*, Yoko Obara, Takayuki Ito, Hiroaki Mizukami, Akihiro Kume and Keiya Ozawa*

Site-specific integration of the therapeutic transgene is favorable for gene therapy applications since it minimizes the risk of insertional mutagenesis and thereby prevents target cells from developing tumors. Adeno-associated virus (AAV), a member of parvovirus, is unique in that it integrates its genome into a specific site termed the AAVS1 locus (19q13.4) in the human genome. A non-structural replication initiator protein of AAV, Rep78 or Rep68, binds the inverted terminal repeat (ITR) sequence at either end of the AAV genome via tandem repeats of the GAGC/GCTC motif. A homologous sequence exists at the AAVS1 site. The Rep protein recognizes it and drives the integration of the AAV genome into AAVS1. The ITR is a *cis* element sufficient for AAVS1-specific integration. The incorporation of the ITR sequence into plasmid DNA is thus discussed in terms of Rep-mediated site-specific integration and of AAVS1 as a hazard-free target for transgene integration. Therefore, the use of the AAV integration machinery should allow us to develop a safer gene delivery system.

Keywords: Site-specific transgene integration; adeno-associated virus; Rep protein; AAVS1 locus.

*Corresponding authors. murabe@jichi.ac.jp, kozawa@ms2.jichi.ac.jp
Division of Genetic Therapeutics, Center for Molecular Medicine,
Jichi Medical University, Tochigi 329-0498, Japan

1. Introduction

Stem cells such as hematopoietic stem cells, embryonic stem (ES) cells and mesenchymal stem cells (MSCs) are attractive targets for gene therapy since they replicate themselves and differentiate into various cell lineages. To manipulate genes in these cells, it is especially important to utilize a system to introduce therapeutic DNA with a minimal risk of insertional mutagenesis. Insertion of the gene of interest into a defined site in the human chromosome is desirable. However, current strategies that achieve the integration of transgene into host chromosomal DNA insert it randomly, which is an insertional mutagenic/oncogenic hazard as shown with retroviral vectors. It is undoubtedly a milestone in gene therapy that 10 patients with X-linked severe combined immune deficiency, a lethal inherited disease characterized by an early block in T and natural killer lymphocyte differentiation due to mutations of the gene encoding the γc cytokine receptor subunit, were treated by infusion of autologous $CD34^+$ hematopoietic stem cells transduced with a murine retrovirus vector encoding the common γ chain (Cavazzana-Calvo *et al.*, 2000; Hacein-Bey-Abina *et al.*, 2002). Unfortunately T-cell leukemia developed in four patients approximately three years after gene therapy (Hacein-Bey-Abina *et al.*, 2003; Baum, 2007). The vector sequence was integrated into the upstream region or the first intron of the LMO2 gene, which is known to be a T-cell proto-oncogene. The activation of the LMO2 gene by chromosomal translocation has been reported in patients with T-cell acute leukemia (Nam and Rabbitts, 2006). The integrated vector promoter, the long terminal repeat (LTR) enhanced the activity of the LMO2 promoter and resulted in aberrant LMO2 expression and premalignant cell proliferation (Hacein-Bey-Abina *et al.*, 2003). To prevent such an adverse event, it is absolutely necessary to employ a strategy that introduces foreign DNA specifically into a predefined safe region of chromosomal DNA.

Prokaryotic site-specific recombinases such as Cre (Sauer and Henderson, 1988) or Flt (Andrews *et al.*, 1985) are widely used to introduce DNA into a site that their respective recognition sequence in eukaryotic cells as well as prokaryotic cells. However, the recognition sequence must be inserted into target chromosomal DNA in advance. Phage phiC31 integrase recognizing phage *attP* and *attB* sites has been shown to mediate site-specific DNA integration in human genome at native "pseudo" *attP* sites (Thyagarajan *et al.*, 2001).

AAV integrates its genome into a particular site in human chromosome 19, termed AAVS1 (19q13.4) (Kotin *et al.*, 1992; Samulski *et al.*, 1991), through the activity of a specific replicase/integrase protein, Rep. Taking advantage

of the AAV integration machinery, systems for AAVS1-specific integration of therapeutic DNA have been developed (Balague *et al.*, 1997; Surosky *et al.*, 1997). The systems are particularly valuable for *ex vivo* gene therapy applications for stem cells. Here the authors will review the current trend of the development of Rep-mediated AAVS1-targeted integration as well as basic biology of the site-specific integration of AAV.

2. Overview of AAV

The adeno-associated virus (AAV) belongs to the family *Parvoviridae* and is classified into the genus *Dependovirus*. A number of AAV serotypes have been reported so far. The AAV serotype 2 was sequenced more than 20 years ago and has been most extensively studied (for general review, see: Muzyczka, 2001; Smith, 2002). The type 2 is referred here as AAV. The AAV depends for its replication and propagation on a helper virus such as adenovirus, herpes virus, papilloma virus and vaccinia virus. When the AAV infects cells alone, it enters a latent infection phase and integrates into the human genome preferentially into the AAVS1 locus on chromosome 19 (19q14.2) (Kotin *et al.*, 1992; Samulski *et al.*, 1991). The integrated genome can be activated and rescued by subsequent superinfection by a helper virus (Fig. 1). There is no disease reported associated with AAV infection. More than 80% of adults are seropositive for antibodies against the AAV. The AAV is used as a gene-transfer vector particularly for long-term gene expression in livers, neurons and muscles. However, AAV vectors that are devoid of the *rep* gene fail to integrate into AAVS1 (see below for details).

The wild-type AAV has been reported to integrate into the AAVS1 site in immortalized cell lines at a frequency of 68% (Kotin *et al.*, 1990) or 94% (Kearns *et al.*, 1996). The insertion of AAV DNA into the AAVS1 site was also reported in human testis tissue (Mehrle *et al.*, 2004). The AAV vector genome has also been shown to persist extrachromosomally in cells (Afione *et al.*, 1996; Duan *et al.*, 1998; Nakai *et al.*, 2001). A recent study reported that the AAV genome in humans appears to persist as episomal forms, not as integrated forms; an attempt to detect a junction DNA sequence between the AAV genome and host chromosome by PCR from clinical tonsil-adenoid samples obtained from children was made. The junction sequence, however, could not be amplified, suggesting that the AAV genome existed in the respiratory tissues as an episomal form (Schnepp *et al.*, 2005). Thus, the mode of persistence of the AAV genome may be different among cell types.

The AAV genome is a single-stranded DNA of 4.8 kb in size. Either end of the genome shows a unique T-shaped hairpin configuration, which is

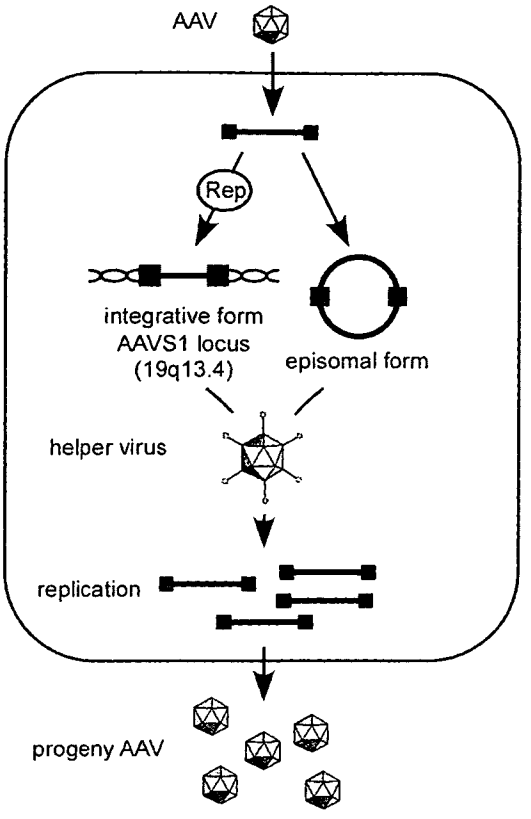

Fig. 1. Life cycle of AAV. Following infection, the AAV genome persists as an integrative or episomal form. The AAV genome preferentially integrates into the AAVS1 site (19q13.4) in the presence of the Rep protein. When AAV and a helper virus (e.g., adenovirus) coinfect cells or a helper virus superinfects the cells harboring the AAV genome, burst replication of AAV occurs.

called an inverted terminal repeat (ITR). The ITR serves as an origin of AAV genome replication. Between the ITRs are two open reading frames corresponding to *rep* and *cap* [Fig. 2(A)]. The *rep* gene encodes four overlapping nonstructural proteins, Rep78, Rep68, Rep52 and Rep40. The *cap* gene codes for structural proteins, VP1, VP2 and VP3. On the genome are three promoters, p5, p19, and p40, designated according to their map positions. The unspliced and spliced transcripts from the p5 promoter encode Rep78 and Rep 68, respectively, while Rep52 and Rep40 are translated from p19-unspliced and -spliced transcripts. Rep78 or Rep68 is a replication initiator of the AAV genome, which possesses site-specific, strand-specific endonuclease activity, ATP-dependent helicase activity (Im and Muzyczka, 1990), and ligase activity (Smith and Kotin, 2000). The Rep78 and Rep68 proteins bind

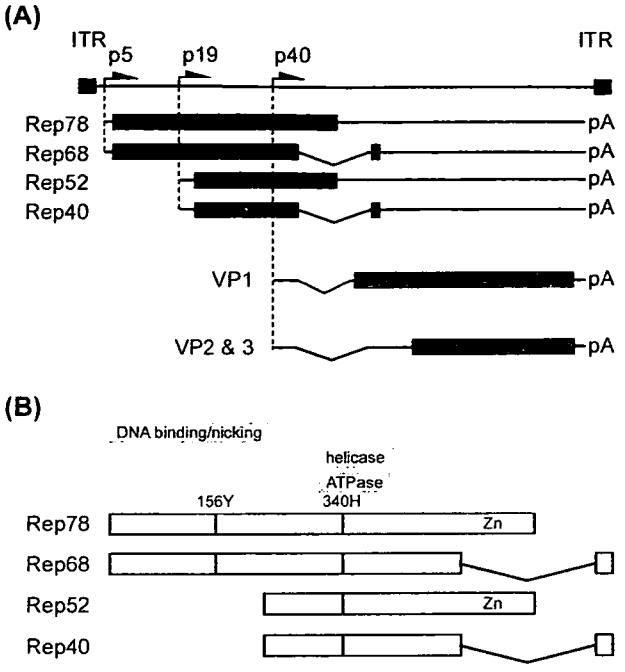

Fig. 2. AAV genome and Rep proteins. **(A)** Map of the AAV genome. Three promoters, p5, p19, and p40 drive transcription of Rep78/68, Rep52/40, and VP proteins, respectively. Spliced transcript from p5 and p19 encode Rep68 and Rep40 proteins, respectively. **(B)** Map of Rep proteins. The domains responsible for DNA binding/nicking, helicase and ATPase domains are indicated. The C-terminal portion of Rep 78 or Rep52 comprises a Zn-finger like motif. The tyrosine and histidine residues at position 156 and 340 are indicated, respectively.

the ITR via the Rep binding site (RBS) consisting of five tandem repeats of the GAGC/GCTC tetranucleotide, self-associate to form a hexameric complex (Hickman *et al.*, 2004; Smith *et al.*, 1997), extrude the terminal resolution site (*trs*), thereby forming a stem-loop structure via their helicase activity (Brister and Muzyczka, 1999) and nicking at the *trs* between the thymidine residues. The protrusion of the *trs* is a prerequisite for nicking at this site by the Rep protein. The site-specific nicking event is followed by unwinding of the terminal hairpin. The smaller Rep52 and Rep40 proteins also have a helicase activity (Collaco *et al.*, 2003; Smith and Kotin, 1998) and are involved in the packaging of the AAV genome into viral capsids (King *et al.*, 2001).

A number of mutational studies and the determination of the three-dimensional structure of the Rep protein (Hickman *et al.*, 2002 and 2004) revealed that the N-terminal half of the large Rep polypeptide is responsible for DNA binding and endonuclease activity. The central portion is essential for helicase activity, ATPase activity and multimerization of Rep proteins.

The lysine residue at position 340 associates with ATP. The tyrosine residue at position 156 covalently links to the 5′-end of single stranded DNA, a product derived from the nicking reaction mediated by the Rep protein [Fig. 2(B)]. Charged amino acids that are important for the site-specific integration were identified in the N-terminal half of Rep78 (Urabe *et al.*, 1999).

Either large Rep protein is capable of supporting the replication of the AAV genome (Holscher *et al.*, 1995). However, functional differences between the two have been reported. Rep68 shows a stronger nicking activity than its counterpart and is more efficient in processing dimers to monomer duplex DNA (Ni *et al.*, 1998 and 1994), which is an intermediate replicative form of the AAV genome. The helicase activity of Rep78 has been reported to be stronger (Wollscheid *et al.*, 1997). Rep78 suppresses CREB-dependent transcription by the interaction of a domain unique to Rep78 with protein kinase A (PKA) (Chiorini *et al.*, 1998; Di Pasquale and Stacey, 1998) which implies that Rep78 (or Rep52) indirectly inhibits adenoviral nuclear transport by PKA (Suomalainen *et al.*, 2001) and CREB-dependent adenovirus promoters E1A, E2 and E4 (Leza and Hearing, 1989). Rep68, but not Rep78, associates with 14-3-3 proteins through phosphorylated serine at position 535 and its interaction may affect the life cycle of AAV (Han *et al.*, 2004).

3. Mechanism of AAVS1-Targeted Integration of the AAV Genome

There are similar RBS and *trs* sequences in the AAVS1, which correspond to the first exon and intron of the *MBS85* gene (Fig. 3). The Rep protein binds the RBS, which is the first event during the replication and amplification of the AAV genome. The Rep protein has been shown to mediate complex formation between AAV DNA and AAVS1 DNA (Weitzman *et al.*, 1994). It has been reported that the Rep protein asymmetrically amplifies the AAVS1 sequence (Urcelay *et al.*, 1995), suggesting that a similar event observed on the ITR during the replication of the AAV genome takes place on the AAVS1 locus as well. A model of deletion-substitution mechanism has been proposed for the mechanism of the site-specific integration of the AAV genome (Linden *et al.*, 1996a).

The ITR sequence alone appears to enhance the integration into host chromosomal DNA (Lieber *et al.*, 1999; Philip *et al.*, 1994), albeit not site-specifically. The minimal *cis* element for AAVS1-specific integration is controversial. A series of analyses of AAV integration events on an EBV-based episomal vector harboring the human AAVS1 sequence in the HEK293 cell

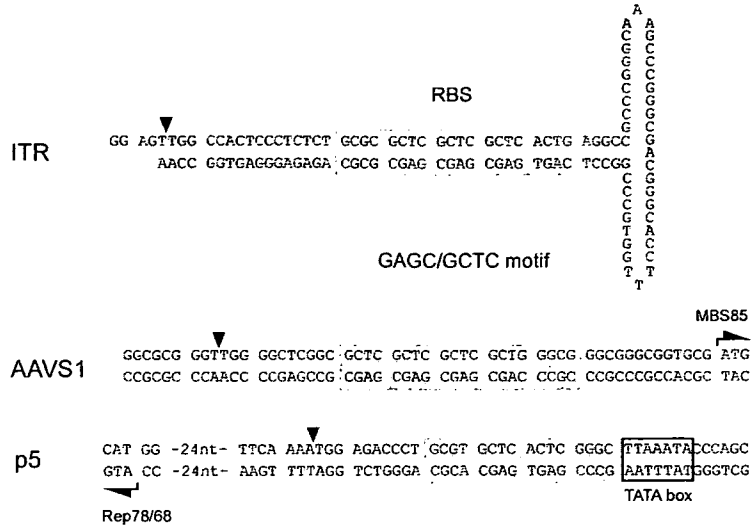

Fig. 3. Comparison of the ITR, AAVS1 and p5 promoter. Terminal resolution site (*trs*) or nicking site is indicated by an arrowhead. A motif consisting of five or four tandem repeats of the GAGC/GCTC tetranucleotide constitutes a Rep binding site (RBS). The initiation codon for myosin binding subunit 85 (MBS85) or for Rep78/68 is indicated by an arrow. The TATA box on the p5 promoter is indicated by a box.

line expressing EBNA-1, revealed that both the RBS and the *trs* sequences are required for AAVS1-targeted integration (Linden *et al.*, 1996a and 1996b). Removal of the sequence upstream of the *trs* appears to result in a decrease in the frequency of AAVS1 specific-integration (Linden *et al.*, 1996a and 1996b). The eight-base spacer sequence between the RBS and *trs* also appears important for site-specific integration, as partial replacement with unrelated nucleotides in the spacer sequence greatly reduces the integration frequency (Meneses *et al.*, 2000). On the other hand, a recent study reported that the RBS motif alone is sufficient for the Rep-mediated AAVS1-specific insertion of a GFP/Neo plasmid into HEK293 cells (Feng *et al.*, 2006). Seven out of 19 (37%) G418-resistant clones were considered to have the GFP/Neo transgene integrated at the AAVS1 site by Southern analysis. The RBS motif within the p5 promoter could direct the GFP/Neo gene into AAVS1 albeit with a lower frequency (2 of 13 clones).

Although the RBS in the ITR is the primary binding site for the Rep protein, a tip of the T-shaped hairpinned structure of the ITR is identified as another Rep association site (Ryan *et al.*, 1996). It has been shown that the affinity of the Rep protein for the RBS in the ITR is higher than for the RBS analog in the p5 promoter, as revealed by electrophoretic mobility shift assay

(Glauser *et al.*, 2005). The ITR thus is the perfect substrate for Rep binding and seems to mediate AAVS1-specific integration most efficiently. The Rep protein is capable of associating with transcription factor Sp1 (Hermonat *et al.*, 1996; Pereira and Muzyczka, 1997) and the TATA binding protein (TBP) (Francois *et al.*, 2005; Hermonat *et al.*, 1998; Su *et al.*, 2000), in both cases binding to the p5 promoter region. The association of Rep with Sp1 and TBP may stabilize the binding of Rep to the p5 element, compensate for Rep's lower affinity for the DNA substrate, and enhance the site-specific integration.

Besides driving transcription of Rep78 and Rep68 mRNA, the p5 promoter sequence also serves as an origin of replication that promotes the amplification of integrated AAV genome (Francois *et al.*, 2005; Nony *et al.*, 2001). The p5 sequence element also enhances the AAVS1-specific integration and is sufficient for the site-specific integration (Philpott *et al.*, 2002a and 2002b). The p5 promoter comprises a RBS and *trs* homolog like the ITR and the AAVS1 site. The Rep protein binds the p5-RBS (Kyostio *et al.*, 1995) and nicks at the *trs* homolog (Francois *et al.*, 2005). The analysis of junction sequences between Rep-mediated integrants and host AAVS1 site revealed that some break points occur at the p5 promoter (Tsunoda *et al.*, 2000), which corroborates the presence of a *cis* element for AAVS1-specific integration.

Cellular factors involved in Rep-mediated integration into AAVS1 have not been fully identified although the Rep protein is a major player. A protein capable of associating with the Rep protein may be involved in the targeted integration. The DNA-dependent protein kinase, which is involved in the repair of double-stranded DNA break and in V(D)J recombination by nonhomologous end-joining, is responsible for the circularization and concatemerization of the AAV genome in cells and inhibits the site-specific integration into host DNA (Song *et al.*, 2004). Recently the TAR RNA-binding protein of 185 kDa (TRP-185), which was identified as a protein binding to the TAR RNA loop of human immunodeficiency virus type 1 (Wu *et al.*, 1991), has been shown to bind to the RBS region. Inhibiting the integration at the RBS of the AAV genome, TRP-185 appears to promote the hexamerization of the Rep protein and to enhance the integration downstream of the RBS (Yamamoto *et al.*, 2007).

4. AAV Target Site: AAVS1 Locus (19q13.4)

The AAVS1 locus (19q13.4) overlaps with the first exon of the *MBS85* gene encoding myosin binding subunit 85 or protein phosphatase 1 regulatory

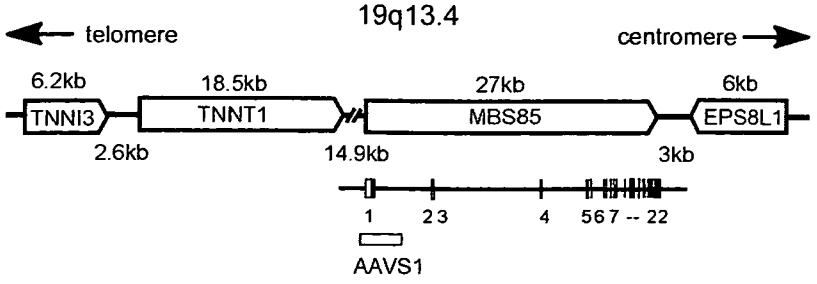

Fig. 4. Location of the AAVS1 site. The AAVS1 site of approximately 2 kb in length overlaps with the first exon and intron of *MBS85*, a hot spot of AAV integration. TNNI3, cardiac troponin I; TNNT1, slow skeletal troponin T1; MBS85, myosin binding subunit 85; EPS8L1, epidermal growth factor receptor pathway substrate 8-like protein 1. (Adapted from a figure by Dutheil *et al.* 2004.)

subunit 12C (Fig. 4). MBS85 is one of the effector molecules in Rho signaling pathway and is involved in the assembly of myosin chains (Tan *et al.*, 2001). Rho is a small molecule with GTPase activity (Bishop and Hall, 2000). RhoA associates with a kinase and the kinase phosphorylates MBS85. The phosphorylated MBS85 inhibits the myosin light chain phosphatase and a net increase in myosin light chain phosphorylation activates myosin and finally induces stress fiber formation. A region upstream of the initiation codon for MBS85 is homologous to the AAV origin of replication, at which are two elements essential for AAV integration: a RBS homolog and a *trs* homolog.

A sequence homologous to human AAVS1 is also found in the simian (Amiss *et al.*, 2003) as well as rodent genome (Dutheil *et al.*, 2004). Integration of the AAV genome into the simian AAVS1 ortholog was identified in a cell line, COS7 (Amiss *et al.*, 2003). Rep68 that are derived from serotype 4 AAV, which was originally isolated from monkey, has a higher affinity for simian RBS than AAV2 Rep68. A survey in genomic banks revealed the presence of an AAVS1 ortholog in the mouse genome too (Dutheil *et al.*, 2004). *In vitro* studies showed that a recombinant type 2 Rep68 can bind to mouse RBS and nick at *trs*, suggesting that the AAV genome would target the mouse AAVS1 locus if the genome is delivered to the nucleus.

The integration of the AAV genome occurs mostly within approximately 2 kb downstream of the RBS (Linden *et al.*, 1996a). The integration of the AAV genome or a transgene into AAVS1 is accompanied by deletion of the target site. After nicking at the *trs*, the Rep protein unwinds target DNA progressively and the DNA with the ITR sequences replaces the existing host DNA, which is contrasted to the microdeletion accompanying the retrovirus-mediated integration event.

5. Multifunctional Rep Protein

The Rep proteins are "sticky" and have been reported to associate with a number of cellular proteins, including TBP (Hermonat *et al.*, 1998; Su *et al.*, 2000); Sp1 (Hermonat *et al.*, 1996; Pereira and Muzyczka, 1997); E2F-1 (Batchu *et al.*, 2001), p53 (Batchu *et al.*, 1999), pRB(Batchu *et al.*, 2002); a topoisomerase I binding protein (Topors) (Weger *et al.*, 2002); protein kinase A (PKA); a protein kinase (PrKX) (Chiorini *et al.*, 1998; Di Pasquale and Stacey, 1998); transcription-positive cofactor 4 (PC4) (Muramatsu *et al.*, 1998; Weger *et al.*, 1999); 14-3-3 proteins (association with Rep68 via phosphrylated serine at position 535) (Han *et al.*; 2004); UBC9 (E2 conjugating enzyme for the small ubiquitin-related polypeptide SUMO-1) (Weger *et al.*, 2004); TRP-185 (Yamamoto *et al.*, 2007); and high mobility group chromosomal protein 1 (HMG1) (Costello *et al.*, 1997).

The anti-proliferative action of the Rep protein is partially explained by the association of the Rep protein with TBP (Hermonat *et al.*, 1998; Su *et al.*, 2000), which is required for the assembly of the transcription initiation complex. The CREB protein (cAMP responsive element binding protein) is a transcription factor and is activated by PKA, which plays a central role in cell growth and development. Thus, the inhibiton of PKA by the Rep protein eventually suppresses the CREB-dependent transcriptional activation (Chiorini *et al.*, 1998; Di Pasquale and Stacey, 1998). PC4 is involved in the downregulation of the p5, p19, and p40 promoters in the absence of a helper virus (Weger *et al.*, 1999). Coinfection of adenovirus relieves the suppression of the AAV promoters (Weger *et al.*, 1999) and enhances the replication of the AAV genome (Muramatsu *et al.*, 1998). UBC9 is involved in the attachment of small ubiquitin-like modifier (SUMO-1), which protects a protein from ubiquitin-mediated degradation. A prolonged life span of the Rep protein (being protected by UBC9) may establish AAV latency in cells (Weger *et al.*, 2004). The association of adenovirus E1b with p53 induces ubiquitin-mediated degradation of the p53 tumor suppressor gene product and this disturbs the cell cycle pathway (Ciechanover *et al.*, 1994). AAV Rep78 associates with p53 and prevents it from being degradated (Batchu *et al.*, 1999). Topors appears to enhance the expression of Rep78, Rep52 and capsid proteins as a transcriptional regulator in the absence of a helper virus (Weger *et al.*, 2002). The 14-3-3 proteins are associated with a number of cellular proteins and implicated in their modification (Fu *et al.*, 2000). The interaction between Rep68 and 14-3-3 proteins results in reduced DNA binding activity of Rep68. However, the significance of the interaction has to be elucidated (Han *et al.*, 2004). A transcription factor E2F-1 is suppressed by its interaction

with the retinoblastoma protein (pRB). Rep78 associates with E2F-1 and stabilizes the E2F-1-pRB complex (Batchu *et al.*, 2002). The Rep protein has been reported to interact with Sp1 and to inhibit a promoter bearing an Sp1 binding motif (Hermonat *et al.*, 1996). The Rep protein also binds to the E2 transactivator of human papilloma virus (HPV) type 16, disrupts the binding of E2 to the cellular transcriptional coactivator p300, and inhibits the replication and transforming activities of HPV (Marcello *et al.*, 2000).

Although the consensus Rep binding site on the ITRs is a five tandem repeat of the GAGC/GCTC tetranucleotide, the Rep proteins can bind to imperfect RBS motifs (Chiorini *et al.*, 1995). Within the AAV p5, p19, and p40 promoter regions, incomplete RBS motifs were identified and the Rep protein has been shown to bind there (Kyostio *et al.*, 1995). The Rep protein regulates the AAV p5, p19, and p40 promoter activity. In latent infection, the large Rep protein suppresses the p5 promoter, which is suitable for persistence of the AAV genome. In the presence of adenovirus, the large Rep protein trans-activates the p19 and p40 promoters.

Systematic *in vitro* studies indicated that the Rep protein can binds to many cellular promoter regions, including the c-*sis* proto-oncogene, glucose transporter, and KIP2 (cyclin-dependent kinase inhibitor) (Wonderling and Owens, 1996 and 1997). The Rep polypeptides have been shown to down-regulate the promoters of a number of genes such as c-H-*ras* , c-*fos*, c-*myc* (Batchu *et al.*, 1994; Hermonat, 1991 and 1994); c-*sis* (Wonderling and Owens, 1996), E2F-1 (Batchu *et al.*, 2001); and the LTR promoter of the human immunodeficiency virus 1 (Batchu and Hermonat, 1995; Oelze *et al.*, 1994). A detailed study showed that Rep suppresses the HPV type 16 p97 promoter by inhibiting the binding of TBP to the p97 promoter (Su *et al.*, 2000) or by directly binding the p97 promoter region (Zhan *et al.*, 1999). While most promoter activities are down-regulated by Rep78 or Rep68, some promoters, such as the human cytomegalovirus (CMV) immediate early promoter (Wonderling *et al.*, 1997) and the c-*sis* promoter (Wonderling and Owens, 1996) are trans-activated. The Rep protein also suppresses translation (Takeuchi *et al.*, 2000).

AAV inhibits the replication of viruses (Bantel-Schaal and zur Hausen, 1988) and cellular transformation by SV40 or Ad E1a plus *ras* oncogene (Khleif *et al.*, 1991). Establishment of cellular transformants by plasmid transfection was also inhibited (Labow *et al.*, 1987). A cellular factor induced by AAV infection that suppresses cell proliferation has been reported (Bantel-Schaal, 2001). It is reported that AAV infection results in an increase in p21 level, a cyclin-dependent kinase inhibitor (Hermanns *et al.*, 1997). Rep68 and Rep78 induce cell cycle arrest in G1 and G2 phases and, in addition,

Rep78 inhibits S-phase progression by accumulating active hypophosphorylated pRb (Saudan *et al.*, 2000), a negative regulator of transcription factors (Weinberg, 1995). Infection with AAV suppresses indirectly the HPV18 promoter activity in transgenic mice (Walz *et al.*, 2002). The Rep protein has been reported to suppress the expression of c-*myc*, c-*myb*, and pRb whereas it up-regulates c-*fos* expression (Klein-Bauernschmitt *et al.*, 1992).

6. The Use of the AAV Integration Machinery to Achieve Site-Specific Integration

Since the elements required for AAVS1-specific integration of AAV are Rep78 or Rep68, and the ITR sequences, it is possible to insert any DNA sequence linked to an ITR sequence into AAVS1 if the Rep78 or Rep68 protein is expressed simultaneously. Early studies demonstrated that a reporter gene was efficiently inserted into the AAVS1 site in cultured cells such as HeLa cells and HEK293 cells, although the frequency of the site-specific integration differed (Balague *et al.*, 1997; Shelling and Smith, 1994; Surosky *et al.*, 1997). Table 1 summarizes the studies of AAVS1-targeted integration in cultured cells.

Shelling and Smith inserted a *Neo* gene downstream of the p40 promoter, the original promoter for capsid protein expression (Shelling and Smith, 1994). The plasmid construct harbored the coding sequence for the p5-, p19-*rep* genes and a p40-driven *Neo* gene flanked by the ITRs (Tratschin *et al.*, 1985). Since the open reading frames (ORFs) of *rep78* and *rep68* extend over the p40 promoter, the Rep proteins expressed from this plasmid were C-terminally truncated. After transfecting it into HeLa cells and HEK293 cells, they analyzed G418-resistant clones by Southern blot and showed that nine out of 12 clones had the *Neo* gene integrated in AAVS1. They also generated an AAV vector with the same *rep* and *Neo* genes. The AAVS1-specific integration efficiency was 82% (nine out of 11 clones).

Balague *et al.* adopted a similar strategy, but the difference is that a *rep78* gene was placed outside the ITR-flanked *GFP* transgene portion on the plasmid. One day after transfecting HEK293 cells, they sorted fluorescent cells and plated them at 1 cell per well, thereby isolating single-cell clones without selection for stable transgene expression analysis. Forty five percent (103 out of 227) of the clones were GFP-positive as compared to eight out of 167 with control *rep78*-free plasmid. Southern analysis showed that in five out of nine (55%) clones integration of the GFP gene was into AAVS1. The results indicated that Rep78 enhances the chromosomal integration of an ITR plasmid both site-specifically and randomly (Balague *et al.*, 1997).