Recombinant Adeno-Associated Virus Vector Genomes Take the Form of Long-Lived, Transcriptionally Competent Episomes in Human Muscle

Bruce C. Schnepf,¹ Jeffrey D. Chulay,² Guo-Jie Ye,² Terence R. Flotte,³
Bruce C. Trapnell,⁴ and Philip R. Johnson^{1,*}

¹The Children's Hospital of Philadelphia and the University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; ²Applied Genetic Technologies Corporation, Alachua, Florida; ³University of Massachusetts Medical School, Worcester, Massachusetts; ⁴Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio.

Gene augmentation therapy as a strategy to treat alpha-1 antitrypsin (AAT) deficiency has reached phase 2 clinical testing in humans. Sustained serum levels of AAT have been observed beyond one year after intramuscular administration of a recombinant adeno-associated virus (rAAV) vector expressing the AAT gene. In this study, sequential muscle biopsies obtained at 3 and 12 months after vector injection were examined for the presence of rAAV vector genomes. Each biopsy sample contained readily detectable vector DNA, the majority of which existed as double-stranded supercoiled and open circular episomes. Episomes persisted through 12 months, although at slightly lower levels than observed at 3 months. There was a clear dose response when comparing the low- and mid-vector-dose groups to the high-dose group. The highest absolute copy numbers were found in a high-dose subject, and serum AAT levels at 12 months confirmed that the high-dose group also had the highest sustained serum AAT levels. Sequence analysis revealed that the vast majority of episomes contained double-D inverted terminal repeats ranging from fully intact to severely deleted. Molecular clones of vector genomes derived directly from the biopsies were transcriptionally active, potentially identifying them as the source of serum AAT in the trial subjects.

INTRODUCTION

ALPHA-1 ANTITRYPSIN (AAT) is a protease inhibitor that is synthesized in the liver and circulates in the blood to reach bodily tissues, where it inactivates resident proteases that left unchecked can cause tissue breakdown. Deficiency of AAT is caused by mutations in the *SERPINA1* gene on human chromosome 14 that result in decreased AAT secretion from the liver. Certain *SERPINA1* mutations lead to early onset panacinar emphysema and liver cirrhosis.¹ Individuals with pulmonary manifestations are treated with weekly intravenous infusions of AAT derived from donated human plasma.² Recent studies show that such protein augmentation slows loss of lung density and emphysema progression, but the requisite frequent infusions remain inconvenient and very expensive.^{3,4}

As an alternative to protein augmentation infusions, AAT gene transfer using a recombinant

adeno-associated virus (rAAV) vector has shown promise in animal models and humans.^{5–15} In a recent phase 2 clinical trial, sustained AAT expression in serum was observed 1 year after intramuscular injection of rAAV1-AAT, and there was a clear dose response in the levels of serum AAT achieved.^{14,15} Although AAT levels in subjects who received the highest dose fell below the therapeutic target, these data demonstrated that relatively high levels of sustained protein expression are achievable using rAAV vector gene delivery to muscle.

As part of the ongoing assessment of subjects in the trial, muscle biopsies were performed at 3 and 12 months after vector injection.^{14,15} These unique biopsy samples afforded us the opportunity to quantify and analyze persisting rAAV1-AAT vector genomes. Herein we show that vector genomes formed double-stranded, circular episomes that

*Correspondence: Dr. Philip R. Johnson, Abramson Research Center, The Children's Hospital of Philadelphia, Room 1216B, 3615 Civic Center Boulevard, Philadelphia, PA 19104. E-mail: johnsonphi@email.chop.edu

persisted at least 12 months after vector administration. Episomes were open circular and supercoiled monomers, as well higher-ordered multimeric forms. Circular genomes were cloned directly from multiple biopsies taken at both time points. Sequence analysis revealed that nearly all clones contained the full intact AAT expression cassette with double-D inverted terminal repeats (ITRs). Importantly, the rescued molecular clones were capable of AAT protein expression *in vitro*, suggesting that these transcriptionally competent episomes are the source of sustained AAT expression in the human subjects.

MATERIALS AND METHODS

Quantification of rAAV vector genomes

Muscle tissue was obtained from biopsies as described previously,¹⁴ and genomic DNA was isolated using EZ1 DNA Tissue Kit (Qiagen, Valencia CA). AAV genome copy number in tissue samples was quantified using real-time TaqMan PCR analysis (ABI 7500). The primers and probe set were directed against the SV40 polyadenylation sequence of the rAAV1-AAT vector (SV40For: 5'-AGCAATAGCATCACAAATTTTCACAA-3' [300 nM]; SV40Rev: 5'-CAGACATGATAAGATACATTGATGAGTT-3' [300 nM]; SV40Probe, [6-FAM]5'-AGCATT TTTTTCCTGCTTCTAGTTGTGGTTTGTGC-3' [TAMRA-6-FAM] [270 nM]). PCR conditions were 50°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min using 100 ng of genomic DNA in 1× TaqMan Universal PCR master mix (Life Technologies, Carlsbad, CA) in a 25 µl volume. A plasmid (pCB-AT-Zero) containing the rAAV1-AAT vector genome⁸ was used to generate a standard curve at a range from 2.5×10^7 to 2.5×10^1 . A circular plasmid was used instead of a linear plasmid to generate the standard curve to more accurately quantify circular rAAV genomes present in the muscle samples.

Plasmid-Safe DNase analysis

Genomic DNA (1.5 µg) was digested in 22.5 µl for 3 hr with *AvrII*, a restriction enzyme that does not cut within the rAAV1-AAT genome, but cuts approximately every 5 kb within the human genome. *AvrII*-digested DNA was further diluted in 33 mM Tris (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, and 2 mM ATP in a 35 µl volume, and then divided into two separate tubes for (+) and (−) Plasmid-Safe DNase (PS-DNase) treatment. The (+) tube received 15 U PS-DNase, while the (−) tube received an equal volume of water. Both tubes were incubated for 16 hr at

37°C followed by nuclease inactivation for 30 min at 70°C. The amount of PS-DNase-resistant rAAV vector DNA (2.5 µl of PS-DNase-treated material, equivalent to 100 ng) was quantified by real-time TaqMan PCR analysis using the SV40 pA-specific primer–probe. The amount of residual genomic DNA present after PS-DNase treatment was quantified using 2.5 µl of PS-DNase-treated material with the TaqMan human β -actin control reagent kit (Life Technologies).

DNA hybridization analysis

For Southern blot hybridization, 5 µg of genomic DNA was first digested in a 60 µl volume for 3 hr with *AvrII*. *AvrII*-digested DNA was further diluted in 33 mM Tris (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, and 2 mM ATP in an 80 µl volume, and then divided into two separate tubes for (+) and (−) PS-DNase treatment. The (+) tube received 25 U PS-DNase, whereas the (−) tube received an equal volume of water. Both tubes were incubated for 16 hr at 37°C followed by nuclease inactivation for 30 min at 70°C. The DNA was fractionated on a 0.8% agarose gel, denatured for 30 min in 1.5 M NaCl and 0.5 M NaOH, followed by neutralization for 2×30 min in 1.5 M NaCl and 1 M Tris (pH 7.4), and transferred to a nylon membrane by capillary transfer. DNA hybridization conditions were 65°C for 16 hr in a buffer containing 6× SSC, 1× Denhardt's reagent, 0.5% SDS, and 100 µg/ml sonicated herring sperm DNA with a radiolabeled probe corresponding to the CMV enhancer region of the rAAV1-AAT vector genome. After hybridization, membranes were washed twice at 65°C in 2× SSC and 0.1% SDS for 15 min, and then twice at 65°C in 0.1× SSC and 0.1% SDS for 30 min, with a final brief wash in 0.1× SSC at room temperature. The hybridized membrane was exposed to X-ray film for visualization.

Isolation and characterization of rAAV1-AAT molecular clones

Genomic DNA was digested with *AvrII* and PS-DNase and treated as described for vector genome quantification. The resulting DNA template (2.5 µl, equivalent to 100 ng of input) was mixed in a final volume of 15 µl of sample buffer (TempliPhi; GE Healthcare, Piscataway, NJ) also containing 300 pmol each of two rAAV1-AAT vector-specific primers: CMVtemp1, 5'-GTATGTTCCCATAGTAC-3'; CMVtemp2, 5'-TGACGTCAATGGAAAGTC-3'. Each primer (one annealing to each strand) contained a phosphorothioate linkage between the last two 3' bases to increase primer and product stability. The reaction was heated to 95°C for

3 min, and then cooled slowly to 4°C to allow primer annealing, and mixed with 15 μ l of phi29 DNA phage polymerase reaction buffer containing phi29 polymerase, and incubated for 18 hr at 30°C. Amplified products were heat inactivated at 65°C for 10 min before digestion with *Hind*III (single cutter) and fractionation on a 0.8% agarose gel. A gel slice from a range of 3–4 kb was excised and the DNA was extracted using QIAquick gel extraction kit (Qiagen), followed by ligation into the *Hind*III site in pBlueScript and transformation into One Shot TOP10 bacteria (Life Technologies). Bacteria plates and liquid cultures were grown at 32°C to minimize ITR deletions and rearrangements. Positive clones were identified by a standard bacteria colony PCR screening assay using primers AGTC-FOR, 5'-CAACCTGGCTGAGTTCGCCTTCAG-3', and AGTCREV, 5'-TTAAACATGCCTAAACGCTTCATC-3', and sequenced by the NAPCore facility at The Children's Hospital of Philadelphia (Philadelphia, PA). To examine the molecular clones for protein expression, they were digested with *Hind*III to release the rAAV1-CB-hAAT insert, self-ligated to recreate an open reading frame, and transfected into HeLa cells using Superfect (Qiagen) transfection reagent. At 48 hr posttransfection, the equivalent of 0.5 μ l of culture medium was fractionated on a 4–12% SDS-polyacrylamide gel for transfer to PVDF membrane and Western blot analysis using a goat anti-alpha 1 antitrypsin-HRP-conjugated antibody (ab7635; Abcam, Cambridge, MA) at 1:5000 dilution with detection using ECL plus chemiluminescent reagent (Thermo Fisher Scientific, Rockford, IL).

RESULTS

Quantification of rAAV1-AAT vector genomes in muscle biopsies

Primary results of the phase 2 clinical trial evaluating the safety and efficacy of rAAV1-AAT are described elsewhere.^{14,15} In this trial, 3 cohorts (of 3 subjects each) received rAAV1-AAT by intramuscular injection at doses of 6×10^{11} (low), 1.9×10^{12} (mid), or 6×10^{12} (high) vector genomes (vg)/kg body weight. As part of the trial protocol, sequential muscle biopsies at the injection site were performed at 3 and 12 months after vector injections. Of the 9 subjects enrolled in the trial, sequential biopsies were available for all subjects except 305 (mid-dose group). Also, subject 303 had only a single biopsy specimen at 3 months. For all other subjects, two separate biopsies were taken at each time point (Table 1). Histologic and immunologic analyses of the biopsies have been reported.¹⁵

Table 1. rAAV1-AAT vector genome copies in muscle biopsies

Vector dose/kg	Subject	Time of biopsy, months	Sample designation ^a	Copies/ μ g ($\times 10^6$) ^b
Low (6×10^{11})	301	3	301-A	6.6
			301-B	6.9
		12	301-1	0.3
			301-2	0.5
	302	3	302-A	1.3
			302-B	3.7
		12	302-1	1.2
			302-2	1.2
	303	3	303-A	4.2
			303-B	NA
		12	303-1	5.4
			303-2	4.5
Mid (1.9×10^{12})	304	3	304-A	3.5
			304-B	3.1
		12	304-1	1.4
			304-2	1.5
	401	3	401-A	0.8
			401-B	0.8
		12	401-1	1.3
			401-2	1.2
	306	3	306-A	5.1
			306-B	7.5
		12	306-1	2.2
			306-2	6.4
	307	3	307-A	16.0
			307-B	27.0
		12	307-1	3.5
			307-2	8.5
High (6×10^{12})	308	3	308-A	6.6
			308-B	2.9
		12	308-1	8.1
			308-2	1.3

AAT, alpha-1 antitrypsin; NA, biopsy sample not available; rAAV, recombinant adeno-associated virus.

^aTwo separate biopsies were obtained at each time point.

^bTaqMan primer-probe specific to SV40 polyadenylation signal in the vector. Copies are per microgram of genomic DNA.

For the purposes of this study, genomic DNA was extracted from tissue samples and vector genome copy number was determined by quantitative PCR (see Materials and Methods).

In samples obtained at 3 months, values ranged from 7.9×10^5 copies/ μ g of genomic DNA (sample 401-A) up to 2.7×10^7 copies/ μ g (307-B) (Table 1). Assuming 6 pg genomic DNA per diploid nucleus, this meant that there were approximately 160 copies per nucleus in sample 307-B. In 12-month samples, values ranged from 2.9×10^5 copies/ μ g (301-1) up to 8.5×10^6 copies/ μ g (307-2). Some subjects showed a modest decrease in copy number from 3 to 12 months (e.g., subjects 301 and 307), but many maintained relatively stable levels over the interval (e.g., subjects 303, 401, and 308). The highest copy numbers were found in the high-dose group (subject 307, samples 307-A and 307-B), and overall, the high-dose group harbored significantly more genome copies over the year than either the low-dose

Figure 1. Schematic of PS-DNase analysis of rAAV1-AAT vector genomes. Vector DNA in muscle tissue can exist as integrated or unintegrated linear, and as circular forms in the nucleus. To enrich for circular vector genomes, total cellular DNA is digested with an enzyme to linearize host genomic DNA, but does not cut within the rAAV vector genome (in this case, *AvrII*). Digested DNA is then treated with PS-DNase that hydrolyzes linear DNA (including integrated vector DNA), but leaves most circular rAAV vector DNA intact. The remaining circular genomes can be characterized by Southern blot, qPCR, and cloning. AAT, alpha-1 antitrypsin; PS-DNase, Plasmid-Safe DNase; rAAV, recombinant adeno-associated virus.

($7.9 \pm 7.2 \times 10^6$ vs. $3.3 \pm 2.5 \times 10^6$ copies/ μ g; $p=0.04$, paired t -test) or the mid-dose ($7.9 \pm 7.2 \times 10^6$ vs. $1.7 \pm 1.0 \times 10^6$ copies/ μ g; $p=0.04$, paired t -test) groups. This finding likely accounts for higher levels of serum AAT observed in this group out to one year.¹⁵

Exonuclease digestion of genomic DNA

In earlier work, we exploited the unique properties of a novel exonuclease (Plasmid-Safe DNase, PS-DNase) to analyze various *in vivo* forms of AAV genomes.^{16–18} PS-DNase selectively degrades linear DNA (single- and double-stranded), as well as single-stranded circular DNA, but does not digest double-stranded circular DNA molecules. Thus, in the context of this work, PS-DNase would degrade single-stranded and replicating rAAV vector genomes, but not circular episomal vector genomes.

The first step in using PS-DNase involves digestion of genomic DNA with a restriction enzyme that does not cut within the rAAV vector genome, but reduces the complexity of the genomic DNA that allows for more efficient hydrolysis of linear DNA (Fig. 1). The digested DNA is then treated with PS-DNase that degrades linear DNA (including free linear and integrated rAAV vector genomes) but leaves double-stranded, circular rAAV vector DNA intact. The PS-DNase-resistant rAAV vector DNA can then be subjected to (1) direct visualization of PS-DNase-resistant rAAV vector genomes by Southern blot; (2) quantification

of rAAV vector genomes resistant to PS-DNase treatment; and (3) cloning of PS-DNase-resistant rAAV vector genomes for sequencing and expression analyses.

Southern blot analysis of vector DNA

Given the high vector copy number in some of the biopsy samples, we believed that Southern blot analysis of unamplified vector DNA was feasible. Genomic DNA was digested with *AvrII* to reduce the sample complexity (*AvrII* did not cut the vector genome), and then half the sample was treated with PS-DNase, whereas the remaining half was left untreated. All samples were fractionated on an agarose gel for Southern blot analysis (Fig. 2). Because of the small physical size of the biopsy samples, we confined this analysis to samples with the most available tissue that contained higher vector genome copies. In one instance, DNA from the 3- and 12-month time points (samples 306-B and 306-2, respectively) from the same subject (306) was available and suitable for blot analysis.

As expected, genomic DNA treated with *AvrII* appeared on ethidium bromide-stained gels as a smear (Fig. 2A, C). When these same samples were then treated with PS-DNase, the visibly stained DNA was fully degraded (compare lanes labeled – and +). Gels were hybridized with a radiolabeled probe corresponding to the CMV enhancer region to visualize the vector genomes present before and after PS-DNase treatment (Fig. 2B, D). The vector

Figure 2. Southern blot analysis of rAAV-hAAT vector genomes in muscle. Genomic DNA from subject biopsies taken at 3 months (**A, B**) and 12 months (**C, D**) after vector administration was digested with an enzyme that does not cut within the vector genome (*AvrII*). Digested DNA was then treated with PS-DNase (lanes labeled "+"), or not treated (lanes labeled "-"). The resulting products were fractionated on agarose gels and stained with ethidium bromide (**A, C**) and then blotted and hybridized with a vector-specific probe (**B, D**). PS-DNase completely degraded linearized genomic DNA in the (+) lanes (**A, C**). Southern blots revealed multiple vector genome forms, including supercoiled monomer (sc), relaxed open circular monomer (rlx), and higher-order multimers. The molecular weight markers shown are a linear ladder. Supercoiled DNA markers were also run in the same gel (not shown). The injected rAAV vector genome was 3845 bp.

DNA forms detected by hybridization were very consistent across the samples tested, regardless of copy number or time of biopsy. Similar forms were also observed in a single subject (306) over time (Fig. 2B, D). Multiple PS-DNase-resistant vector genome species were observed before and after PS-

DNase digestion, indicating the presence of double-stranded vector DNA. The most abundant forms were relaxed and supercoiled monomers (rlx mono and sc mono). Also present were various-sized multimers migrating at sizes consistent with dimers, trimers, and higher-ordered forms, probably

the result of intermolecular recombination. A smear of hybridized signal ranging from the gel origin down to approximately 1 kb was also observed. A plausible interpretation is that the formation of the double-stranded circular vector genomes was not a precise process, and a range of intermediate forms were generated. Positive and negative strands of single-stranded genomes that might anneal to form double-stranded linear species would also have been digested by PS-DNase and would not be detected.

Quantification of episomal circular vector DNA

Once we showed that vector episomes were present in biopsy samples, we used quantitative PCR to determine the amount of PS-DNase-resistant rAAV vector DNA in each sample. A ratio of vector DNA present before and after PS-DNase digestion was expressed as percent resistant DNA. The average level of PS-DNase-resistant vector DNA in the 3- and 12-month subject samples tested was $61.0 \pm 16.4\%$ and $58.6 \pm 13.5\%$, respectively. To control for PS-DNase activities, we tested two known templates. First, a TaqMan primer–probe set corresponding to the human β -actin gene was used to quantify the amount of host cell DNA present after PS-DNase treatment. Over 99% of the genomic DNA (human β -actin target sequence) was degraded, thereby confirming the desired activity of the exonuclease. Second, we tested PS-DNase activity on a plasmid target that was greater than 99% circular. Plasmid pCB-AT-Zero, which was used to generate the rAAV1-AAT vector used in the clinical study,¹³ was spiked into naïve genomic DNA and treated with PS-DNase identically to the

biopsy samples. An average of multiple experimental replicates ($n = 30$) showed that $84.5 \pm 25.3\%$ of circular plasmid DNA was resistant. Thus, under these experimental conditions, there is a modest background of hydrolysis of double-stranded circular DNA.

Cloning of rAAV1-AAT genomes from subject samples

The presence of extrachromosomal DNA forms allowed us to clone and sequence vector DNA directly from biopsy samples. Methods previously developed to clone and characterize infectious molecular clones of wild-type AAV from human tissues¹⁸ were used to amplify circular vector genomes. After digestion with a noncutting restriction enzyme, genomic DNA was enriched for circular genomes by treatment with PS-DNase (Fig. 1). Circular genomes were then annealed to a mixture of random hexamer primers as well as vector-specific primers (to the CMV enhancer region), and subjected to isothermal rolling circle amplification via strand displacement. We found that the addition of gene-specific primers to the random hexamers increased the robustness of the amplification compared with random hexamers alone. We consistently amplified as few as 200 vector copies/ μ g of genomic DNA (equivalent to 1.2×10^{-3} vector copies/nucleus).

The result of rolling circle amplification with Phi29 polymerase is long linear concatemers of vector genomes that can be digested with a restriction enzyme that cuts once within the vector genome (in this case, *HindIII*) to create head-to-tail, head-to-head, or tail-to-tail unit length vector genome fragments (Fig. 3). *HindIII*-digested DNA

Figure 3. Cloning of rAAV1-AAT vector episomes. Circular episomal vector genomes (a head-to-tail configuration is shown) were amplified using rolling circle amplification with Phi29 polymerase to produce linear concatemers (see Materials and Methods). Concatemers were digested with a single-cut enzyme (*HindIII*) and then cloned into a standard plasmid (see Materials and Methods). Individual clones were subsequently sequenced (Table 2) and analyzed for AAT expression (Fig. 4). Triangles represent inverted terminal repeats (ITRs) at the ends of the vector genome.

was fractionated on an agarose gel and a gel slice representing fragments from 3 to 4 kb was excised, and the DNA was extracted and cloned into the *Hind*III site of a standard plasmid cloning vector. The head-to-tail, head-to-head, and tail-to-tail fragments were predicted to be 3845, 3933, and 3507 bp in size, respectively, and so there was the opportunity of identifying all possible vector genome orientations. For the 3- and 12-month samples, we isolated 21 clones from 11 different subject samples (3 months: 302-A, 304-A, 304-B, 306-A, 307-A; 12 month: 302-1, 302-2, 303-1, 304-2, 306-2, 307-2) and sequenced their ITR junctions.

Twenty of 21 clones contained head-to-tail junctions. All clones contained at least some ITR sequences with the exception of clone 306A-16. This clone contained no ITR sequences, but the promoter sequence was intact. In addition, most of the SV40 polyadenylation sequence was deleted in this clone, with only the first 11 bp of the 222 bp SV40 pA remaining. Eight of the 21 (38%) clones had complete double-D ITR (DD-ITR) structures, with both flip and flop orientations present (Table 2). This value is consistent with our previous findings that 50% of wild-type AAV genomes contain intact DD-ITRs.¹⁸ The complete DD-ITR junctions contained a single AAV ITR flanked on either end with a D region. Clone 302-1-1 was interesting because of the presence of an illegitimate DD-ITR. The ITR junction in this clone contained two do-

main B hairpins instead of the typical B and C hairpins. We called this clone a double-D and double-B ITR (DD-BB-ITR). Although all of the individual B domain sequences were complete in this DD-BB-ITR clone, they were separated by nucleotides "ACGCC" instead of the expected "A" nucleotide only. The majority of the remaining clones contained D regions with complete or partially deleted A regions, with very little, if any, intervening C or B sequences. Clone 302A-8 was unique in that it was not a simple head-to-tail junction and contained recombination among 3 rAAV1-AAT vector genomes, with a head-to-head orientation at one end and a tail-to-tail orientation at the other. The length of the entire 302A-8 genome was 3.5 kb, which explains why it was isolated along with intact, unit-length genomes. None of the three individual vector genomes in clone 302A-8 contained an intact open reading frame. We were unable to obtain reliable sequence from clone 302-2-1 isolated from the 12-month subject samples because of the strong secondary structure created from the ITR hairpin in this particular clone. However, based on the restriction digestion pattern, it was predicted to contain a partially deleted ITR similar to clone 303-1-3.

rAAV1-AAT clones express AAT

The abundance of episomal genomes in the biopsy samples led us to wonder whether these genomes were transcriptionally competent. All 12 molecular clones isolated from the 3-month biopsies were released from the cloning vector by digestion with *Hind*III and then self-annealed with DNA ligase to restore the AAT open reading frame. The ligated products were transfected into HeLa cells, and the culture medium was analyzed by Western blot. All of the clones, with the exception of clone 302A-8, expressed AAT at the expected size of approximately 51 kb (Fig. 4). A slightly smaller species was also seen that most likely corresponded to posttranslational proteolytic processing that results in an N-terminal truncation.¹⁹ The protein products from the rescued clones appeared to be identical to the protein produced from plasmid pCB-AT Zero that was used to generate the rAAV1-AAT vector.¹³ Clone 302A-8 was predicted not to express protein based on sequence analysis, revealing the lack of an open reading frame because of complex rearrangements and recombination. Interestingly, clone 306A-16, which is missing almost the entire SV40 polyadenylation sequence, was able to express AAT although at a relatively lower level (Fig. 4). Given that the vector genomes were circular, it is possible that another poly-

Table 2. ITR junctions of rAAV1-AAT molecular clones

Time, months	Clone	ITR junction sequence ^a
3	302-A-8	Complex rearrangements with no open reading frame
	302-A-12	pA-D-ΔA-C-A'-D'-CB
	304-A-16	pA-ΔA-D'-CB
	304-B-23	pA-D-A-B-B'-C-C'-A'-D'-CB (complete DD-ITR—"flip")
	304-B-41	pA-D-A-C-C'-B-B'-A'-D'-CB (complete DD-ITR—"flop")
	304-B-68	pA-D-ΔA-ΔA'-D'-CB
	306-A-5	pA-ΔD-ΔD'-CB
	306-A-16	ΔpA-CB
	307-A-3	pA-D-ΔA-ΔA'-D'-CB
	307-A-4	pA-D-A-C-C'-B-B'-A'-D'-CB (complete DD-ITR—"flop")
	307-A-9	pA-D-A-ΔA'-D'-CB
	307-A-23	pA-D-A-C-C'-B-B'-A'-D'-CB (complete DD-ITR—"flop")
12	302-1-1	pA-D-A-B-B'-(accgc)-B-B'-A'-D'-CB (illegitimate DD-ITR)
	302-1-2	pA-D-A-C-C'-B-B'-A'-D'-CB (complete DD-ITR—"flop")
	302-2-1	pA-D—unreadable—D'-CB
	303-1-1	pA-D-A-B-B'-C-C'-A'-D'-CB (complete DD-ITR—"flip")
	303-1-2	pA-D-A-B-B'-C-C'-A'-D'-CB (complete DD-ITR—"flip")
	303-1-3	pA-D-ΔA-ΔA'-D'-CB
	304-2-1	pA-D-A-C-C'-B-B'-A'-D'-CB (complete DD-ITR—"flop")
	306-2-1	pA-D-ΔA-ΔA'-D'-CB
	307-2-1	pA-ΔA-D'-CB

^aITR junction is oriented with the SV40 polyadenylation signal (pA) on the left and the CMV enhancer CB-actin promoter (CB) on the right.

Deleted sequences/regions are designated by "Δ." DD, double-D; ITR, inverted terminal repeat.

Figure 4. Transgene (AAT) expression from cloned vector genomes. Clones were excised from the plasmid vector with *HindIII* (single cut within the vector genome), self-ligated to recreate an open reading frame, and transfected into HeLa cells (see Materials and Methods). Medium from transfected cells was analyzed for AAT expression by Western blot. All clones expressed protein of the expected size except for 302A-8, which, based on sequence analysis, was not predicted to express. Plasmid pCB-AAT Zero, which was used to generate the rAAV1-AAT vector used in the clinical trial, was included as a positive control. A GFP plasmid was a negative control.

adenylation signal was utilized further downstream from the missing SV40 pA. In fact, the sequence CATAAA was present 190 bp from the translational stop codon and could function as a surrogate polyadenylation signal.²⁰ Taken together, these data suggest that the AAV ITRs do not need to be intact, or even present, within the circular episomal vector genomes for protein expression. Transgene expression is possible as long as the correct regulatory elements such as the promoter and polyadenylation signal are present.

DISCUSSION

rAAV vectors have an established record of high-efficiency gene transfer in a variety of model systems and are now being tested as therapeutic modalities in a wide range of human diseases, including AAT deficiency.^{21–26} Muscle biopsies taken as part of a phase 2 trial for AAT deficiency^{14,15} afforded us (1) a unique opportunity to study the long-term fate of rAAV vector genomes in transduced human muscle, and (2) the potential to elucidate *in vivo* mechanisms of protein expression. Given the doses of vector injected and the resulting AAT expression in serum, we expected to detect vector genomes at the 3-month time point, and did so in all samples analyzed. Likewise, vector genomes were also found in all

follow-up biopsies at 12 months. Any general inferences regarding vector copy number and persistence must be considered in light of the small number of subjects studied and potential sampling bias when obtaining small biopsies. Nonetheless, three conclusions about copy number seem likely. First, although there was a slight trend downward in copy number from 3 to 12 months, overall, vector genome numbers were very stable. Second, there was a dose response when comparing the low- and mid-dose groups to the high-dose group. The highest absolute copy numbers were found in a high-dose subject (307), and serum AAT levels at 12 months confirmed that the high-dose group also had the highest sustained serum AAT levels.¹⁵ Finally, vector copy numbers in our human subjects were very similar to copy numbers in nonhuman primates given a comparable dose of an rAAV vector carrying a different transgene.²⁷ In the monkey study, animals received a dose of 5×10^{12} vg/kg and muscle biopsies were obtained 14–34 months later. In the 9 monkeys studied, vector copies/ μ g were comparable to the human biopsies ($3.6 \pm 2.9 \times 10^6$ vs. $7.9 \pm 7.2 \times 10^6$, respectively).

Earlier studies in humans also reported persistence of rAAV vector genomes in muscle. An rAAV2 vector carrying the human factor IX gene was given by multiple intramuscular injections, with individual doses of 1.5×10^{12} vg/site.²⁸ Analysis of vector genome copies from biopsies of the injection site revealed an average copy number of 0.5 vector copies per diploid genome that persisted for at least 4 years.^{28,29} Subjects with lipoprotein lipase (LPL) deficiency were injected with vector (rAAV1-LPL) at multiple sites at doses of 1.6 – 4.2×10^{11} vg/site.³⁰ Vector genome copy numbers ranged from undetectable to 14.4 (average 3.5) copies per diploid genomes in a 6–9-month window of analysis.³¹ Although direct comparisons with our study are not possible, when considered together, these data suggest that rAAV transduction of human muscle leads to stable, long-lived transduction.

Studies in animals have repeatedly shown that rAAV vector genomes persist *in vivo* predominantly as circular episomes.^{16,27,32–35} We also showed that wild-type AAV can persist as circular episomes in humans, and that infectious molecular clones could be isolated directly from human tissues.^{17,18,36} Building on that experience with wild-type AAV in humans, we used a similar approach to characterize the forms of rAAV vector DNA in the muscle biopsies. Episomes were readily detected in biopsies taken at 3 and 12 months, and their relative abundance did not change over the study period. From Southern blots (Fig. 2), the most abundant

forms were relaxed (open circular) monomeric double-stranded episomes, followed by supercoiled monomers. Lower levels of high-molecular-weight concatemers were also seen, indicating intermolecular recombination among clones.³⁷ This pattern is essentially identical to previous analyses of rAAV vector genomes in mice and nonhuman primate tissues.^{16,32,34,35,38} Interestingly, rAAV episomes in nonhuman primate tissues were shown to assimilate into chromatin-like structures similar to the host cellular chromosomal DNA nucleosome pattern,³² suggesting a potential mechanism for long-term persistence.

rAAV episomes probably arise through the initial formation of a double-strand intermediate from second-strand synthesis.^{39,40} This is followed by circularization through ITR recombination to form monomeric or higher-order forms (concatemers) with a double-D ITR (DD-ITR) structure.^{16,27,33,41,42} This process is often imprecise, giving rise to incomplete DD-ITRs that contain deletions and rearrangements.^{16,27,41,42} Using PS-DNase followed by rolling circle amplification, we previously isolated wild-type AAV episomes from human tissues that also contained DD-ITRs.^{17,18} All of the wild-type AAV clones except one were in a head-to-tail conformation, which has been shown to be the predominant configuration of rAAV and wild-type AAV genomes *in vitro* and *in vivo*.^{16–18,27,32,38,41,42} For wild-type AAV episomes, although a variety of ITR deletions were observed, approximately 50% of the DD-ITR junctions in the tissues contained intact DD-ITRs.¹⁸

In the current study, analysis of the DD-ITR junctions of rAAV1-AAT clones revealed similar results in that 38% (8/21) of the recovered clones contained an intact DD-ITR (Table 2). There were an equal number of clones from both the 3- and 12-month samples that possessed a complete DD-ITR, which suggested that these structures are stable *in vivo*. The majority of the remaining clones had completely absent B and C regions with other deletions in the A regions. Similar extensive internal deletions were also observed for rAAV ITR junctions *in vitro*⁴² and *in vivo* from nonhuman primates.²⁷ Interestingly, a DD-ITR of any length did not appear to be a necessity for episomal persistence. Clone 306-A-16 from the 3-month time-point contained no ITR sequence of any kind, and had additional deletions of the polyadenylation signal. We could not rule out that this deletion occurred during the amplification and cloning process, but, nonetheless, this clone was able to direct expression of AAT *in vitro*. Moreover, extensive ITR deletions did not affect transgene expression since

the majority of episomal clones tested were capable of AAT protein expression (Fig. 4).

In addition to the abundant circular episomes, other forms of vector DNA were likely present in the biopsy specimens. As a minimum estimate, episomes comprised at least 60% of the total genomes (i.e., 60% were resistant to PS-DNase treatment). This number clearly underestimated the relative abundance of episomes for two reasons. First, we showed that the exonuclease (PS-DNase) degraded approximately 25% of control plasmid DNA, portending that vector genome episomes were also degraded. Second, a portion of the episomal vector genomes in the muscle tissue were probably damaged (e.g., become linearized) during extraction and handling processes, rendering them susceptible to exonuclease hydrolysis. In addition, at lower levels, single-stranded vector genomes might still be present (either free or encapsidated) that would be degraded by PS-DNase.¹⁶

This study was not designed to evaluate integration of vector DNA into host cell chromosomes. We have shown previously using a genome-wide amplification strategy that more than 99.5% of rAAV vector genomes were not integrated in mouse skeletal muscle.¹⁶ We also used linear-amplification-mediated (LAM)-PCR to examine wild-type AAV genomes in human tissue¹⁷ and showed that integration is a rare event for wild-type AAV during natural infection. More recently, Nowrouzi et al.²⁷ extensively examined the extent of integrated rAAV vector genomes in nonhuman primate muscle and liver. The authors coupled LAM-PCR with deep sequencing to estimate the frequency of rAAV vector integration. In that study, rAAV vector integration occurred at a low frequency of 10^{-4} to 10^{-5} in both nonhuman primate liver and muscle. When integration events were detected, there was no preference for specific genomic loci, and integrations were three times more likely in liver than in muscle. The authors concluded that rAAV integration is a rare event *in vivo*, and that concatemeric episomes were the predominant genome form out to 34 months after vector administration. These and other studies^{16,18,27,32,43–45} allow us to conclude that the overwhelming majority of vector genomes in the muscle biopsies persisted as episomal concatemers.

In summary, rAAV vectors expressing AAT persisted in human muscle predominantly as double-stranded circular episomes for at least 12 months after injection. Episomes isolated directly from biopsy specimens were transcriptionally competent, potentially identifying them as the source of serum AAT in the trial subjects. Our observations are remarkably consistent with data

derived from small- and large-animal models, and further validate the use of such models in developing gene transfer strategies for humans. Finally, these data reinforce the biologic similarities of rAAV vectors and wild-type AAV. Both forms of AAV genomes persist via circular episomes, and neither appears to form integrated proviruses as part of its life cycle.

ACKNOWLEDGMENTS

This study was supported by Applied Genetic Technologies Corporation. The authors thank the

Children's Hospital of Philadelphia NAPCORE for their expertise in sequencing of the rAAV1-AAT vector genomes.

AUTHOR DISCLOSURE

J.D.C. and G.-J.Y. are employees and own shares of Applied Genetic Technologies Corporation, and have a conflict of interest to the extent that this work potentially increases personal financial interests. None of the other authors have competing interests.

REFERENCES

- Silverman EK, Sandhaus RA. Clinical practice. Alpha1-antitrypsin deficiency. *N Engl J Med* 2009;360:2749–2757.
- Stoller JK, Aboussouan LS. A review of alpha1-antitrypsin deficiency. *Am J Respir Crit Care Med* 2012;185:246–259.
- Strange C, Beiko T. Treatment of alpha-1 antitrypsin deficiency. *Semin Respir Crit Care Med* 2015;36:470–477.
- Traclet J, Delaval P, Terrioux P, et al. Augmentation therapy of alpha-1 antitrypsin deficiency associated emphysema. *Rev Mal Respir* 2015;32:435–446.
- Song S, Morgan M, Ellis T, et al. Sustained secretion of human alpha-1-antitrypsin from murine muscle transduced with adeno-associated virus vectors. *Proc Natl Acad Sci USA* 1998;95:14384–14388.
- Song S, Scott-Jorgensen M, Wang J, et al. Intramuscular administration of recombinant adeno-associated virus 2 alpha-1 antitrypsin (rAAV-SERPINA1) vectors in a nonhuman primate model: Safety and immunologic aspects. *Mol Ther* 2002;6:329–335.
- Poirier AE, Combee LA, Marton AT, et al. Toxicology and biodistribution studies of a recombinant adeno-associated virus 2 (rAAV2) alpha-1 antitrypsin (AAT) vector. *Mol Ther* 2004;9:S40.
- Brantly ML, Chulay JD, Wang L, et al. Sustained transgene expression despite T lymphocyte responses in a clinical trial of rAAV1-AAT gene therapy. *Proc Natl Acad Sci USA* 2009;106:16363–16368.
- Brantly ML, Spencer LT, Humphries M, et al. Phase I trial of intramuscular injection of a recombinant adeno-associated virus serotype 2 alpha1-antitrypsin (AAT) vector in AAT-deficient adults. *Hum Gene Ther* 2006;17:1177–1186.
- De BP, Heguy A, Hackett NR, et al. High levels of persistent expression of alpha1-antitrypsin mediated by the nonhuman primate serotype rh.10 adeno-associated virus despite preexisting immunity to common human adeno-associated viruses. *Mol Ther* 2006;13:67–76.
- Liquan Wang R, McLaughlin T, Cossette T, et al. Recombinant AAV serotype and capsid mutant comparison for pulmonary gene transfer of alpha-1-antitrypsin using invasive and noninvasive delivery. *Mol Ther* 2009;17:81–87.
- Halbert CL, Madtes DK, Vaughan AE, et al. Expression of human alpha1-antitrypsin in mice and dogs following AAV6 vector-mediated gene transfer to the lungs. *Mol Ther* 2010;18:1165–1172.
- Chulay JD, Ye GJ, Thomas DL, et al. Preclinical evaluation of a recombinant adeno-associated virus vector expressing human alpha-1 antitrypsin made using a recombinant herpes simplex virus production method. *Hum Gene Ther* 2011;22:155–165.
- Flotte TR, Trapnell BC, Humphries M, et al. Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing alpha1-antitrypsin: Interim results. *Hum Gene Ther* 2011;22:1239–1247.
- Mueller C, Chulay JD, Trapnell BC, et al. Human Treg responses allow sustained recombinant adeno-associated virus-mediated transgene expression. *J Clin Invest* 2013;123:5310–5318.
- Schnepp BC, Clark KR, Klemanski DL, et al. Genetic fate of recombinant adeno-associated virus vector genomes in muscle. *J Virol* 2003;77:3495–3504.
- Schnepp BC, Jensen RL, Chen CL, et al. Characterization of adeno-associated virus genomes isolated from human tissues. *J Virol* 2005;79:14793–14803.
- Schnepp BC, Jensen RL, Clark KR, et al. Infectious molecular clones of adeno-associated virus isolated directly from human tissues. *J Virol* 2009;83:1456–1464.
- Umekawa T, Kohri K, Amasaki N, et al. Sequencing of a urinary stone protein, identical to alpha-one antitrypsin, which lacks 22 amino acids. *Biochem Biophys Res Commun* 1993;193:1049–1053.
- Beaudoing E, Freier S, Wyatt JR, et al. Patterns of variant polyadenylation signal usage in human genes. *Genome Res* 2000;10:1001–1010.
- Coura Rdos S, Nardi NB. The state of the art of adeno-associated virus-based vectors in gene therapy. *Virol J* 2007;4:99.
- Daya S, Berns KI. Gene therapy using adeno-associated virus vectors. *Clin Microbiol Rev* 2008;21:583–593.
- Aalbers CJ, Tak PP, Vervoordeldonk MJ. Advancements in adeno-associated viral gene therapy approaches: Exploring a new horizon. *F1000 Med Rep* 2011;3:17.
- Loring HS, Flotte TR. Current status of gene therapy for alpha-1 antitrypsin deficiency. *Expert Opin Biol Ther* 2015;15:329–336.
- Schnepp BC, Johnson PR. Vector-mediated antibody gene transfer for infectious diseases. *Adv Exp Med Biol* 2015;848:149–167.
- Hastie E, Samulski RJ. Adeno-associated virus at 50: A golden anniversary of discovery, research, and gene therapy success—a personal perspective. *Hum Gene Ther* 2015;26:257–265.
- Nowrouzi A, Penaud-Budloo M, Kaeppl C, et al. Integration frequency and intermolecular recombination of rAAV vectors in non-human primate skeletal muscle and liver. *Mol Ther* 2012;20:1177–1186.
- Manno CS, Chew AJ, Hutchison S, et al. AAV-mediated factor IX gene transfer to skeletal muscle in patients with severe hemophilia B. *Blood* 2003;101:2963–2972.
- Jiang H, Pierce GF, Ozelo MC, et al. Evidence of multiyear factor IX expression by AAV-mediated gene transfer to skeletal muscle in an individual with severe hemophilia B. *Mol Ther* 2006;14:452–455.
- Stroes ES, Nierman MC, Meulenberg JJ, et al. Intramuscular administration of AAV1-lipoprotein lipase S447X lowers triglycerides in lipoprotein lipase-deficient patients. *Arterioscler Thromb Vasc Biol* 2008;28:2303–2304.
- Mingozzi F, Meulenberg JJ, Hui DJ, et al. AAV-1-mediated gene transfer to skeletal muscle in humans results in dose-dependent activation of capsid-specific T cells. *Blood* 2009;114:2077–2086.

32. Pénau-Budloo M, Le Guiner C, Nowrouzi A, et al. Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. *J Virol* 2008;82:7875–7885.
33. Duan D, Sharma P, Yang J, et al. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J Virol* 1998;72:8568–8577.
34. Afione SA, Conrad CK, Kearns WG, et al. *In vivo* model of adeno-associated virus vector persistence and rescue. *J Virol* 1996;70:3235–3241.
35. Song S, Laipis PJ, Berns KI, et al. Effect of DNA-dependent protein kinase on the molecular fate of the rAAV2 genome in skeletal muscle. *Proc Natl Acad Sci USA* 2001;98:4084–4088.
36. Flotte TR, Schwiebert EM, Zeitlin PL, et al. Correlation between DNA transfer and cystic fibrosis airway epithelial cell correction after recombinant adeno-associated virus serotype 2 gene therapy. *Hum Gene Ther* 2005;16:921–928.
37. Yang J, Zhou W, Zhang Y, et al. Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination. *J Virol* 1999;73:9468–9477.
38. Sun X, Lu Y, Bish LT, et al. Molecular analysis of vector genome structures after liver transduction by conventional and self-complementary adeno-associated viral serotype vectors in murine and nonhuman primate models. *Hum Gene Ther* 2010;21:750–761.
39. McCarty DM, Fu H, Monahan PE, et al. Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction *in vivo*. *Gene Ther* 2003;10:2112–2118.
40. Choi VW, Samulski RJ, McCarty DM. Effects of adeno-associated virus DNA hairpin structure on recombination. *J Virol* 2005;79:6801–6807.
41. Duan D, Sharma P, Dudas L, et al. Formation of adeno-associated virus circular genomes is differentially regulated by adenovirus E4 ORF6 and E2a gene expression. *J Virol* 1999;73:161–169.
42. Yan Z, Zak R, Zhang Y, et al. Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. *J Virol* 2005;79:364–379.
43. Chenuaud P, Larcher T, Rabinowitz JE, et al. Optimal design of a single recombinant adeno-associated virus derived from serotypes 1 and 2 to achieve more tightly regulated transgene expression from nonhuman primate muscle. *Mol Ther* 2004;9:410–418.
44. Rivera VM, Gao GP, Grant RL, et al. Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer. *Blood* 2005;105:1424–1430.
45. Toromanoff A, Cherel Y, Guilbaud M, et al. Safety and efficacy of regional intravenous (r.i.) versus intramuscular (i.m.) delivery of rAAV1 and rAAV8 to nonhuman primate skeletal muscle. *Mol Ther* 2008;16:1291–1299.

Received for publication October 2, 2015;
accepted after revision November 22, 2015.

Published online: December 2, 2015.