Molecular Therapy
Methods & Clinical Development
Original Article

Analysis of vector genome integrations in multicentric lymphoma after AAV gene therapy in a severe hemophilia A dog

Lucas Van Gorder,[1,2] Bhavya S. Doshi,[1,2] Elinor Willis,[3] Timothy C. Nichols,[4] Emma Cook,[5] John K. Everett,[5] Elizabeth P. Merricks,[4] Valder R. Arruda,[1,2] Frederic D. Bushman,[5] Mary Beth Callan,[6] and Benjamin J. Samelson-Jones[1,2]

[1]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [2]Division of Hematology, Raymond G. Perelman Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [3]Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [4]Department of Pathology and Laboratory Medicine and the UNC Blood Research Center, University of North Carolina, Chapel Hill, NC 27599, USA; [5]Department of Microbiology Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [6]Department of Clinical Sciences and Advanced Medicine, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Adeno-associated viral (AAV) vectors have traditionally been viewed as predominantly nonintegrating, with limited concerns for oncogenesis. However, accumulating preclinical data have shown that AAV vectors integrate more often than previously appreciated, with the potential for genotoxicity. To understand the consequences of AAV vector integration, vigilance for rare genotoxic events after vector administration is essential. Here, we investigate the development of multicentric lymphoma in a privately owned dog, PC9, with severe hemophilia A that was treated with an AAV8 vector encapsidating a B domain–deleted canine coagulation *F8* gene. PC9 developed an aggressive B cell lineage multicentric lymphoma 3.5 years after AAV treatment. Postmortem analysis of the liver, spleen, and lymph nodes showed the expected biodistribution of the AAV genome. Integration events were found both in PC9 and a second privately owned hemophilia A dog treated similarly with canine *F8* gene transfer, which died of a bleeding event without evidence of malignancy. However, we found no evidence of expanded clones harboring a single integration event, indicating that AAV genome integrations were unlikely to have contributed to PC9's cancer. These findings suggest AAV integrations occur but are mostly not genotoxic and support the safety profile of AAV gene therapy.

INTRODUCTION

Although it has been traditionally accepted that adeno-associated viral (AAV) vectors are nonintegrating with limited concerns for triggering oncogenesis,[1,2] more recent studies have challenged the previously presumed safety profile of AAV vectors.[3,4] Neonatal mice treated with AAV vectors have a significantly increased incidence of hepatocellular carcinoma due to integrations of vector genomes.[5–7] In addition, 2 adult hemophilia A (HA) research colony dogs treated with AAV vectors showed an increase in transgene factor VIII (FVIII) levels over time.[8] Postmortem analysis found that these HA dogs had clonally expanded hepatocytes that harbored AAV integrations, although there was no evidence of oncogenesis.[8] In addition, benign integration events have been observed in both nonhuman primates and human patients treated with AAV vectors.[9] These findings highlight the need for surveillance to detect rare potentially genotoxic AAV vector integration events.

Four cancers have been reported to date in AAV recipients with hemophilia: 2 in HA recipients and 2 in hemophilia B (HB) recipients (Table 1).[10,11] Investigations into the cancers in the HB recipients did not identify a link between AAV vector treatment and oncogenesis.[10,11] Although integration analyses of the tumors in HA recipients have not yet been published, results reported at meetings suggest that there is also no evidence that AAV vector integrations played a role in oncogenesis. A recent report described the development of an epithelioid neoplasm of the spinal cord 14 months after the intravenous administration of onasemnogene abeparvovec; although AAV vector genomes were present in most tumor cells, integration site analysis was inconclusive because of limited sample availability.[12] These findings are overall reassuring; however, the possible time between exposure to AAV vector and formation of an AAV-related cancer is unknown.

Here, we investigate the development of multicentric lymphoma in PC9, a privately owned dog with severe HA, who received an AAV serotype 8 (AAV8) vector. The development of PC9's cancer is of particular concern due to the time after vector administration and the type of cancer. Multicentric lymphoma is the most common malignancy in dogs, with 25 cases reported per 100,000 dogs per year at

Received 16 June 2023; accepted 12 November 2023;
https://doi.org/10.1016/j.omtm.2023.101159.

Correspondence: Benjamin J. Samelson-Jones, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.
E-mail: samelsonjonesb@chop.edu

Table 1. Cancers reported in hemophilia AAV gene therapy clinical trial subjects

Disease	Vector dose (vg/kg)	AAV capsid	Cancer	Cancer VCN	Subject predispositions
Hemophilia A	6×10^{13}	AAV5	Parotid acinic cell carcinoma	No difference from nonneoplastic tissues	Not identified (European Hemophilia Consortium, Press Release, 2022)
	6×10^{13}	AAV5	B cell acute lymphoblastic leukemia	No difference from nonneoplastic tissues	Philadelphia Chromosome (Biomarin Pharmacetical Inc., SEC Filing, 2022)
Hemophilia B	1×10^{12}	AAV8	Squamous cell carcinoma of the tonsil	0	Not identified[10]
	2×10^{13}	AAV5	Hepatocellular carcinoma	3.21	Infection with hepatitis C and B viruses[11]

an average age of diagnosis of 6–9 years.[13,14] Although canine multicentric lymphoma can be of T or B cell origin, it is estimated that 70%–80% of cases are of B cell origin.[15] Studies in mice and nonhuman primates have shown that AAV8 transduces lymphoid tissues and lymphocytes,[16–18] suggesting that the vector that PC9 received likely was present in his lymphocytes. Likewise, data from AAV gene therapy clinical trials have shown that peripheral blood mononuclear cells clear AAV vector slower than other tissues;[19] this prolonged presence of vector in hematolymphoid tissue potentially allows more time for the AAV vector genome to integrate into lymphocytes and cause deleterious downstream effects. Furthermore, exposures to mutagens often have a latent period before the carcinogenic effects can be seen. The incubation time for blood cancers after Epstein-Barr virus infection is roughly 4 years.[20] These considerations raised the question about the role of AAV integration in the development of cancer in PC9.

RESULTS

Control dog PC1 and diagnosis of multicentric lymphoma in PC9

PC1 was a privately owned shepherd mix diagnosed with severe HA caused by a threonine to methionine point mutation at amino acid 62 in the canine *F8* gene. As previously reported,[21] PC1 was treated with 5×10^{13} total vector genomes per kilogram (vg/kg) of an AAV8 B domain–deleted (BDD) canine (c) FVIII dual vector system driven by a thyroxine-binding globulin gene promoter/enhancer (Figure 1, top) at 13 months of age.[22–24] He died from a bleeding event 8.8 years after treatment (B.S.D., B.S.J., M.B.C., unpublished data). Gross and histopathological clinical autopsy of PC1 did not demonstrate evidence of cancer. Postmortem liver samples were saved for further research and were used as controls to compare against PC9.

PC9 was a privately owned mixed-breed dog diagnosed early in life with severe HA caused by a 10-bp deletion in exon 14 of the *F8* gene. PC9 was treated at 9 months of age with 6×10^{12} vg/kg of an AAV8 vector containing a codon optimized BDD-cFVIII variant gene (ΔF + V3) driven by a human α1 antitrypsin (hAAT) promoter (Figure 1, bottom).[25–29] PC9 tolerated treatment well, with transgene cFVIII antigen and activity in the moderate HA range (B.S.D., B.S.J., M.B.C., unpublished data).

Three and one-half years after gene therapy, at 4 years of age, PC9 was diagnosed with multicentric lymphoma. This cancer was aggressive, with the rapid onset of neurological symptoms leading to euthanasia on the same day as the diagnosis. The early age of onset coupled with the aggressive presentation warranted further study to determine whether the AAV vector was involved with oncogenesis. Gross and histopathological clinical autopsy determined that the cancer had infiltrated peripheral and visceral lymph nodes, the cranial mediastinum, the spleen, the liver, and the lungs. Postmortem samples of the liver, spleen, and lymph nodes were saved for further analysis.

Examination of H&E-stained sections demonstrated that neoplastic cells were present in the lymph nodes, spleen, and liver of PC9 (Figure 2). The architecture of the lymph node was entirely replaced by neoplastic lymphocytes, which extend through the capsule and into the perinodal adipose tissue (Figures 2A and 2B). The spleen had significant cancer infiltration, with sheets of neoplastic lymphocytes markedly expanding the white pulp (Figures 2C and 2D). The liver had modest tissue infiltration, with neoplastic lymphocytes expanding portal tracts (Figures 2E and 2F). The neoplastic lymphocytes are intermediate to large (Figures 2B, 2D, and 2F). Sections of liver from PC1, who was unaffected by cancer, showed normal morphology, with no evidence of neoplastic cells (Figure 2G). The samples from PC9 demonstrated that the spleen and lymph nodes were effaced by neoplastic cells, whereas the liver was only modestly affected.

Sections of the spleen were immunostained for CD3 (T cell marker),[14,30] CD79b (B cell marker),[14,31] and Pax5 (a mature B cell marker)[32] to determine the cell of origin. Neoplastic cells were negative for CD3 (Figure 3A). Approximately 90% of neoplastic cells were positive for CD79b (Figure 3B) and 75% were positive for PAX5 (Figure 3C). Combined, these findings indicate that the lymphoma is of B cell origin, as are a majority of canine multicentric lymphoma cases.

Vector copy number (VCN) determination

To determine the number of vector genomes present in each tissue, vector-specific TaqMan primers were designed to target the promoters (Figure 1; Table S1). Splenic and lymph node samples were confirmed by both gross and histopathological examination to be diffusely infiltrated by cancer, with neoplastic cells throughout each section. Genomic DNA (gDNA) isolated from 8 sections of each frozen tissue was analyzed with these TaqMan primers, and they demonstrated VCN per diploid cell of 0.05 in the liver and 0.02 in the spleen. The signal from amplification of the hAAT promoter in

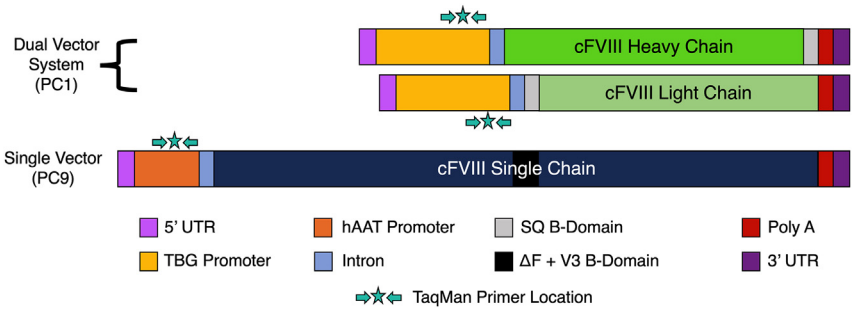

Figure 1. Schematics of vectors

Illustration of the dual vector system and single-chain vector system used to treat HA dogs PC1 and PC9, respectively. (Top) In the dual vector system shown, 2 separate AAV vectors delivered either the 2.5-kb cFVIII heavy chain or the 2.4-kb cFVIII light chain driven by the same liver-specific thyroxine binding globulin promoter/enhancer. These constructs used an embedded intron and SV40 polyadenylation signal to promote expression. The site of the TaqMan primers and probe are shown on the schematic near the 3′ end of the thyroxine binding globulin promoter. (Bottom) The single-chain vector contains a codon-optimized cFVIII cDNA sequence with the ΔF + V3 B domain linker (SFSQNPPVSKATNVSNNSNTSNDSNVS) driven by a liver-specific hAAT promoter. This construct also used an embedded intron and SV40 polyadenylation signal to drive expression. The site of the TaqMan primers and probe are shown on the schematic near the 3′ end of the hAAT promoter. SQ, XX.

the lymph nodes was below the level of detection (estimated at a VCN of <0.0006 genomes per diploid cell) (Figure 4).

Because the vector genome is not consistently present in all of the tissues infiltrated by lymphoma, these data are inconsistent with an AAV vector genome integration contributing to the development of lymphoma in PC9. If insertional mutagenesis by AAV contributed to transformation, then samples with a significant cancer burden such as the spleen and lymph nodes would be expected to have a VCN approaching 1 genome per diploid cell. The value could be higher if vector genomes were present in untransformed cells or if the neoplastic cells had more than 1 integrated AAV genome. However, this was not seen. The lower VCN values seen in these samples replicate the expected biodistribution of an AAV8 vector.

In comparison, the PC1 liver gDNA samples were analyzed for VCN, with TaqMan primers designed to target the thyroxine binding globulin (TBG) promoter. PC1 was determined to have a VCN of 29.07 genome copies per diploid cell. The higher VCN in PC1 compared to PC9 is likely due to the higher vector dose he received.

Vector genome integration analysis

To understand the vector composition of the dog samples in more detail, we also carried out an analysis of integration site distributions. The TaqMan qPCR showed that an intact AAV vector genome integration was not responsible for cancer transformation, but the integration of partial genomes may still have caused insertional mutagenesis. We thus undertook deep sequencing of tissue samples to evaluate for AAV vector integrations using an integration site recovery protocol that identifies AAV vector sequence integrations at the edges of the AAV vector genome that include AAV inverted terminal repeats (ITRs). Clonally expanded integration sites or integration sites near or within oncogenes were of highest concern.

gDNA from 3 disparately spaced samples of the liver, spleen, and lymph nodes from PC9 (9 samples total) were analyzed for AAV vector integrations using ligation mediated PCR which targets ITR sequences.[8] Three gDNA samples from the liver of PC1 were also analyzed as controls.

In the 9 samples from PC9, 41 distinct integration events were detected (Table S2). Of these, 30 integrations were detected in hepatic samples, 10 integrations were detected in splenic samples, and only 1 integration was detected in lymph node samples (Figure 5A). We were reassured that no single integration event was repeated between any gDNA samples. Based on the diffuse cancer infiltration of the spleen and lymph nodes, all of the samples contained gDNA from neoplastic cells. The absence of repeated integration events across neoplastic cells indicates that an AAV vector genome integration did not expand with the lymphoma cells and is inconsistent with an AAV vector genome integration driving oncogenesis. Furthermore, the distribution of these integration events correlated with the relative distribution of persisting vector genomes seen in the TaqMan qPCR assay, further supporting the conclusion that no integration is expanded within the lymphoma.

In the hepatic gDNA samples analyzed from the control dog PC1, there were 212 distinct integration events detected (Figure 5A; Table S3). Despite the large number of integrations, none of the integration sites found in PC9 were also found in PC1. The larger number of integration events in PC1 is likely due to the increased vector dose received. Because the probe used to find the AAV integrations is specific to the ITR sequences, there is no way to distinguish which vector (AAV-cFVIII-heavy chain or AAV-cFVIII-light chain) is integrated at any given location.

Estimates of the number of integration events per cell or sample are challenging.[33] Lower-bound approximations of integrations using the Chao1 estimator[34] suggest there are at least 460 events in PC1's liver samples and 33 events in PC9's liver samples. This is estimated to be between 1 and 3 integrations per cell for PC1 and between 0.1 and 0.2 integrations per cell for PC9. The Chao1 estimated integration events for PC9 are similar to the observed events because of the overall low number of events.

The distribution of identified integration sites between PC9 and PC1 was similar. Neither dog had evidence of an integration that occurred in the exon of an oncogene. Of all of the integrations found in PC9, 0% were in exons, 2.4% were in oncogene transcriptional units,

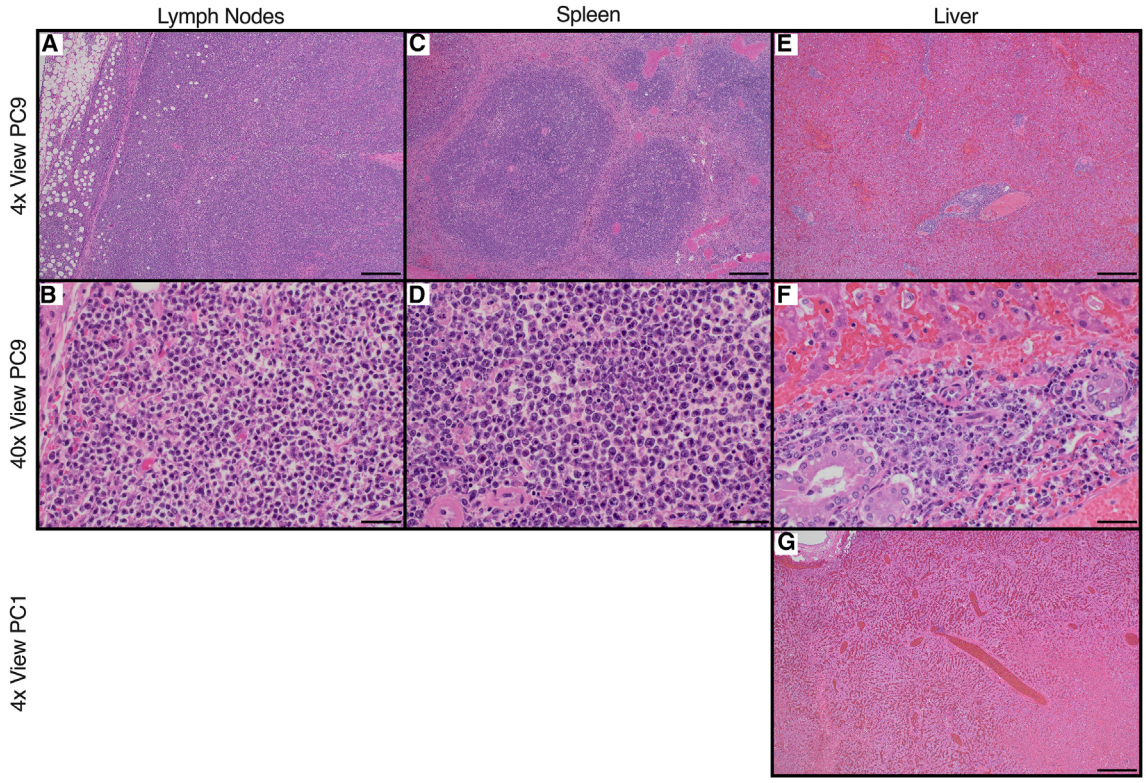

Figure 2. Lymph node, spleen, and liver histology of PC9 and PC1

Formalin-fixed and paraffin-embedded postmortem tissue samples from PC9's lymph node (A and B), spleen (C and D), and liver (E and F) were stained with H&E and viewed at either 4× (A, C, and E) or 40× (B, D, and F) magnification to assess cancer burden. Densely packed neoplastic cells efface the normal tissue architecture of the spleen and lymph node (A–D) and moderately infiltrate PC9's hepatic portal tracts (E and F). As a comparison, a postmortem liver sample from PC1 stained similarly (G) demonstrates normal liver morphology with no neoplastic cells. 4× scale bars, 200 μm; 40× scale bars, 20 μm.

39% were in the transcriptional units of nononcogenes, and 59% were in intergenic regions (Figure 5B). In all of the integration events found in PC1, 3.8% were in nononcogene exons, 6.6% were in oncogene transcriptional units, 37% were in the transcriptional units of nononcogenes, and 53% were in intergenic regions (Figure 5C). Although the observed frequency of identified integrations in PC1 and PC9 into transcriptional units is higher than would be randomly expected, the observed frequency into oncogene transcriptional units is lower than would be randomly expected (Figure S1). These results suggest that although the total number of integration events increases with higher vector doses, the distribution remains roughly the same. A total of 4.9% and 9.8% of PC9's integration events and 13% and 16% of PC1's integrations events were within 25 or 50 kb of oncogene transcriptional units, respectively.

The single integration event in an oncogenic transcriptional unit in PC9 occurred in the gene *STARD13*, which acts as a tumor suppressor in hepatocellular carcinomas. However, this integration event was not clonally expanded and is therefore not causative of PC9's cancer. The 8 exonic integrations in PC1 were also not clonally expanded. However, 3 separate integration events occurred within exons of *ZCCHC10*, which codes for a zinc finger nuclease thought to be important in p53 regulation and a cancer suppressor.[35] There is no sequence homology between the vector genome and *ZCCHC10*.

Six integration events were identified as a potential clonal expansion (present in ≥ 5 cells) in the gDNA samples from PC9 by the sonic abundance method (Table 2).[36] Three intergenic integrations were potentially clonally expanded in the spleen or liver. The most abundant was an intergenic integration found upstream of *TNFRSF19* in a hepatic gDNA sample; however, it was expanded to only 11 cells. Three transcriptional unit integrations were also clonally expanded in the liver or spleen, although only expanded to 7 cells each. The only integration found in the lymph node samples was identified in a single cell. None of these integration events are expanded sufficiently to suggest that they were consistently present in neoplastic cells.

Only 1 integration in PC1, in an intron of *PEBP4,* was found to be potentially expanded to 5 cells (Table 3). *PEBP4* codes for a secreted phosphatidylethanolamine binding protein with many biological functions.[37,38] It has been suggested to be a regulator of signaling pathways in many cancer types, including gastric and non-small cell lung cancers. Of note, an AAV vector genome integration in

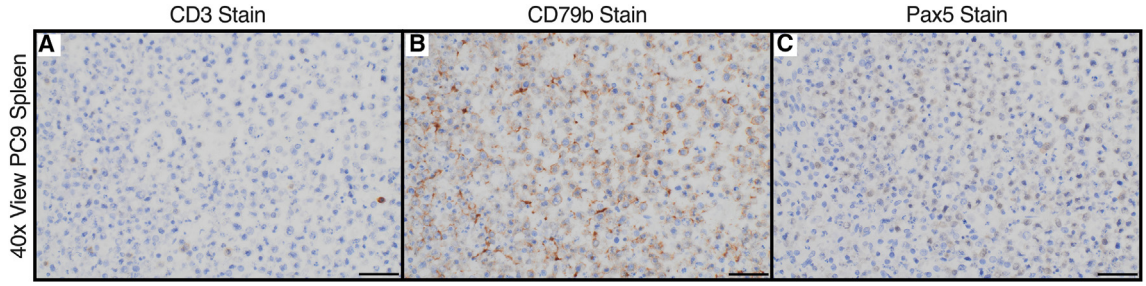

Figure 3. Immunohistochemistry staining shows neoplastic cells are of B cell origin

To determine the cell of origin of the lymphoma, immunohistochemistry for T cell (CD3) and B cell (CD79b and Pax5) markers was performed on formalin-fixed paraffin-embedded sections of spleen. (A) Neoplastic cells are negative for CD3. (B) 90% of neoplastic cells exhibit cytoplasmic staining for CD79b, whereas (C) roughly 75% exhibit nuclear staining for Pax5. Scale bars, 20 μm.

PEBP4 leading to a clonal expansion has also been previously reported in an AAV-treated dog.[8] This is the only integration identified that has been previously observed after AAV vector administration.

These results support a model in which oncogenesis in PC9 was independent of AAV vector administration; the putatively clonally expanded integrations in PC9 are neither sufficiently expanded nor consistently found across multiple samples. PC1 also had many unique integration events, including a putatively clonally expanded integration, but with no evidence of oncogenesis. Combined, these data indicate that AAV vector genome integrations occur but mostly without significant genotoxic effects.

DISCUSSION

PC9 was diagnosed with an aggressive lymphoma at 4 years of age, well before the typical 6- to 9-year-old age range of diagnosis. Postmortem examination showed multicentric B cell lymphoma involving lymph nodes, spleen, liver, lungs, and cranial mediastinum. Because of the uniqueness of this model of HA (privately owned pet dog living in real-world conditions) and the growing preclinical data suggesting that AAV vector genome integrations can cause clonal expansions, which may increase the incidence of cancer, further investigation into the cause of the cancer was undertaken. A single research colony HA dog has previously been reported to have developed a localized solid tumor after AAV vector administration, with the tumor having no detectable vector.[39]

Despite the early age of diagnosis and aggressive nature of the lymphoma, methods evaluating ITR and promoter integrations provide no evidence that PC9's cancer was caused by the integration of the recombinant AAV8-cFVIII vector. A copy of the promoter would be expected to be present in every neoplastic cell if a full AAV vector genome integration contributed to the lymphoma; this would have resulted in VCNs closer to 1 genome per diploid cancer cell in cancer-infiltrated tissues. In contrast, we found the expected biodistribution of vector genomes after AAV8 gene therapy: a high concentration of AAV genomes in the liver, lower levels in the spleen, and few to no AAV vector genomes present in the lymph nodes. However, this experiment only assayed for the presence of the vector-specific promoter due to the sequence-specific limitations of TaqMan-based PCR and the high oncogenic potential of a promoter integration. Partial AAV vector genome integrations that do not include the promoter can occur and would not be captured in this assay.

Integration analysis also did not support a role for partial ITR-containing AAV genome integrations in cancer transformation in PC9. Although many ITR-containing integration events were found, none of these integrations was notably expanded. Furthermore, no ITR-containing integration event was found in more than one sample, suggesting that there was no driving ITR-containing integration event that contributed to oncogenesis and expanded with the cancer.

This integration analysis did not link ITR-containing AAV genome integrations to oncogenesis, but it did provide other insights. PC9 had 6 integrations that were putative clonal expansions. The clinical significance of these integrations and putative expansions is yet unknown, although their presence stresses the importance of continued monitoring for genotoxicity after AAV vector administration in both preclinical and clinical studies.

Integration analysis in PC1, who received a vector dose of 5×10^{13} vg/kg and had 29.07 vector copies per diploid genome in his liver, showed a significantly increased burden of AAV vector integration events with 212 integration events across 3 hepatic samples. In contrast, PC9, who received a lower vector dose of 4×10^{12} vg/kg and had 0.05 vector copies per diploid genome in his liver, had only 30 integration events in 3 hepatic samples. These data support an association between the dose of AAV vector administered and the number of integration events, with larger doses leading to an increased incidence of AAV genome integrations. The relationship between AAV vector dose and number of vector integration events in PC1 and PC9 is similar to the association observed in 6 research colony HA dogs.[8] In these 6 dogs treated with AAV doses between 1.2×10^{13} and 4×10^{13} vg/kg, integration analysis identified between 161 and 764 vector genome integrations across 3 liver samples. Combined, these results suggest that larger vector doses likely increase the total number of integration events, but do not change the integration targeting.

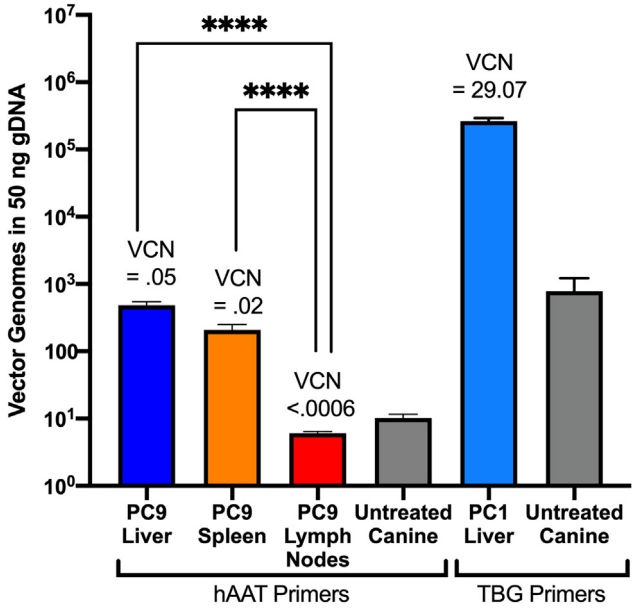

Figure 4. VCN analysis in PC9 and PC1

TaqMan qPCR targeted to the hAAT and TBG promoters (see Figure 1) were used to determine the VCN in 50 ng of gDNA from PC9's liver, spleen, and lymph nodes, and PC1's liver. Vector copies were quantified using a 12-point standard of 1:10 dilutions from 1×10^{11} copies per reaction down to 1 copy per reaction, with the expression cassette containing plasmid. VCNs were determined using 5.5 pg gDNA per diploid canine cell. Data shown as an average of 2 experiments of 8 gDNA samples run in duplicate with SD. ****$p < 0.0001$.

No integrations at the same site occurred between PC1 and PC9. This suggests that there is not a common locus for insertion, but rather that genomes may be inserted into transcriptionally active parts of the genome. However, 2 integration sites stand out. First, PC1 had multiple unique integrations in the gene *ZCCHC10*. *ZCCHC10* is known to be expressed in the liver and may be an integration target due to being in an open chromatin state.[40] Second, the single integration site that was potentially clonally expanded in PC1, *PEBP4*, has been previously reported to be clonally expanded in a research colony HA dog (J60) that received a similar dual AAV vector gene therapy with BDD-cFVIII, albeit with a different serotype vector (AAV9) than PC1.[8] Postmortem liver samples of J60 demonstrated that vector integrations into *PEBP4* were clonally expanded to >100 cells; this represented the largest clonal expansion reported in the study.[8] Integration in *PEBP4* leading to putative clonal expansions in >1 dog warrants further study.

In both PC1 and PC9, AAV vector genome integrations were identified within and close to oncogenes. These integrations are of the highest concern because they have the potential to initiate or contribute to oncogenesis. However, in this study of long-term samples 8.9 and 3.5 years after vector administration, the integrations in transcriptional units of oncogenes were not observed to be clonally expanded. These data show that although these high-concern integrations do occur, there is no evidence to support the fact that they have genotoxic effects. Furthermore, these results suggest that integrations in the transcriptional units of oncogenes are a rare event rather than a locus for AAV vector genome integrations.

The analyses used in this study are limited in their abilities to detect partial vector genome integrations. The TaqMan-based VCN analysis only identified the presence of the vector-specific promoters, whereas the vector genome integration analysis probed only the 5′ and 3′ ITR sequences. However, these assays are the emerging standard to analyze gDNA samples for AAV genome integration events. The lack of ITR-including integrations and low promoter-based VCN in the lymph nodes suggest that AAV vector genomes were likely not present in neoplastic tissue in large quantities. Nonetheless, it is important to note that the possibility of partial AAV vector genome integrations causing insertional mutagenesis cannot be wholly eliminated in PC9 because of these experimental limitations.

Another consideration for oncogenic potential is the transgene being delivered by the AAV vector. A recent study reported that mice transiently expressing BDD-FVIII have a higher incidence of hepatocellular carcinoma than do mice expressing full-length FVIII, hypothesized to be due to cellular stress induced by BDD-FVIII protein.[41] However, in our study, the transgene is driven by a liver-specific promoter. Since the cancer is of B cell origin, transgene BDD-FVIII protein was unlikely to contribute to its development. Furthermore, the VCN analysis showed that even if this promoter were leaky and driving ectopic expression in lymphocytes, the vector is not consistently present in the lymphoma, further supporting the conclusion that transgene BDD-FVIII did not contribute to oncogenesis.

These data augment the safety profile of AAV gene therapy emerging from clinical trials. Similar to the recent investigation of the development of hepatocellular carcinoma in a hemophilia B AAV gene therapy subject,[11] here, we found no evidence that the AAV vector played a role in the formation of PC9's cancer. In the subject with hepatocellular carcinoma, 95 total AAV genome integration sites were identified with no pathogenically expanded or dominant integration site.[11] In our study, both PC9 and PC1 had a similar order of magnitude of AAV vector genome integration events, with no dominant integration sites or large putative clonal expansions. These results reinforce the low risk of AAV vector genotoxicity, but they also emphasize the importance of continued surveillance for rare oncogenic events given the number of vector integration events observed, including at previously reported sites.

MATERIALS AND METHODS

Privately owned dogs

The study was approved by the University of Pennsylvania Institutional Animal Care and Use Committee (protocol no. 804598), the Matthew J. Ryan Veterinary Hospital of the University of Pennsylvania Privately Owned Animal Protocol Committee (protocol no. 366), and the Children's Hospital of Philadelphia Institutional Animal Care and Use Committee (protocol no. 001441).

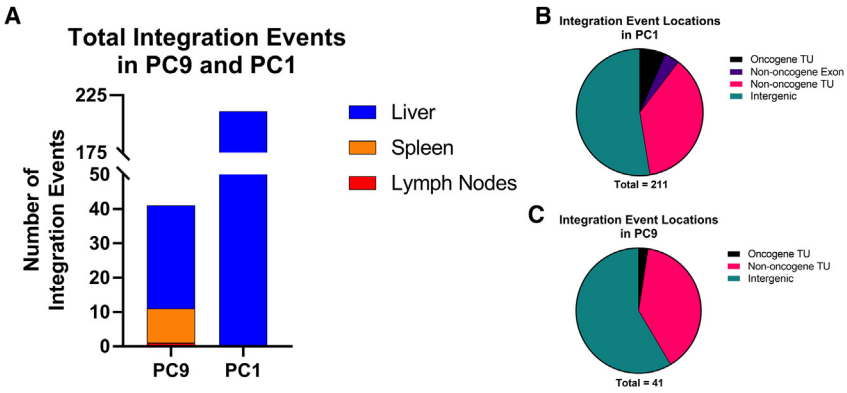

Figure 5. AAV genome integration events are similarly distributed in PC9 and PC1

Sites of AAV genome integration locations were determined in 9 (3 from each tissue) gDNA samples from PC9 and 3 liver gDNA samples from PC1 by ligation-mediated PCR, sequencing, and alignment to a reference canine genome. (A) The distribution of the 41 integration events in PC9 matches the VCN distribution seen in Figure 4, with 30 integration events in the liver, 10 in the spleen, and only 1 in the lymph nodes. PC1 has 212 integration events in the liver alone. The spleen and lymph node of PC1 could not be analyzed because samples were not saved. (B) The genomic location of all 41 integrations in PC9. One integration was in an oncogene transcriptional unit (TU), 16 integrations were in nononcogene TUs, and 24 integrations were intergenic. (C) The genomic locations of all 212 integrations in PC1. Eight integrations occurred in the nononcogene exons, 14 occurred in oncogene TUs, 78 occurred in nononcogene TUs, and 111 were intergenic.

Immunohistochemistry on paraffin-embedded sections

Tissues were fixed in 10% neutral buffered formalin. Trimmed tissues were routinely processed for paraffin embedding, sectioned to 5 μm thickness, and mounted on ProbeOn slides (Thermo Fisher Scientific). The immunostaining procedure was performed using a Leica BOND RX^m automated platform combined with the Bond Polymer Refine Detection kit (Leica DS9800).

After dewaxing (Leica AR9222) and rehydration in distilled H_2O, sections were pretreated with the epitope retrieval BOND ER2 high pH buffer (EDTA-based pH 9.0, Leica AR9640) for 20 min at 95°C. Endogenous peroxidase was inactivated with 3% H_2O_2 for 10 min at room temperature. Nonspecific protein-protein interactions were blocked with Leica PowerVision IHC/ISH Super Blocking solution (Leica PV6122) for 30 min at room temperature. The primary antibody (CD3ε, Bio-Rad, MCA1477T, rat monoclonal antibody; PAX-5, Cell Signaling Technology, no. 12709, D7H5X, rabbit monoclonal antibody; or CD79b, Cell Signaling Technology, no. 96024, rabbit monoclonal antibody) was diluted 1/600 using the same solution used for the blocking step and applied for 45 min at room temperature. A biotin-free polymeric IHC detection system (Leica DS9800) consisting of horseradish peroxidase–conjugated Gt anti-rat immunoglobulin G secondary antibody was applied for 25 min at room temperature. Immunoreactivity was revealed with the diaminobenzidine chromogen reaction. Slides were finally counterstained in hematoxylin, dehydrated in an ethanol series, cleared in xylene, and permanently mounted with a resinous mounting medium (Thermo Scientific ClearVue coverslipper). Tris-buffered saline (Leica AR9590) was used as washing solution in between steps.

H&E tissue staining

Tissues were fixed in 10% neutral buffered formalin. Trimmed tissues were routinely processed for paraffin embedding, sectioned to 5 μm thickness, and stained with H&E.

gDNA isolation

Frozen tissues samples were obtained from the clinical veterinarians that performed the autopsies. Pieces of tissue were excised, homogenized, and separated into 10-mg portions. gDNA was isolated via the Qiagen DNeasy Blood and Tissue kit (Qiagen no. 69506) standard protocol.

A total of 180 μL of buffer ATL and 20 μL Proteinase K was added to each tissue sample. This mixture was vortexed and incubated at 56°C until lysed. The sample was then vortexed, 200 μL buffer AL was added, and then vortexed again. The sample then had 200 μL 100% ethanol added and was thoroughly mixed by vortex. The sample was placed on a DNeasy Mini spin column and centrifuged at 10,000 × g for 1 min. The flow through was discarded and the column was washed with 500 μL AW1 by centrifugation at 10,000 × g for 1 min. The column was washed a second time with 500 μL AW2 by centrifugation at 20,000 × g for 3 min. The gDNA was then eluted in 200 μL buffer AE by the addition of the AE, incubation at room temperature for 1 min, and then centrifugation at 10,000 × g for 1 min. The concentration of the resulting gDNA was then determined by absorbance before being used in subsequent assays.

TaqMan qPCR

All TaqMan qPCR reactions were performed using the 2× Primetime Gene Expression Master Mix (Integrated DNA Technologies no. 0000796148) on a Bio-Rad CFX384 Real-Time System. Reactions were run in 384-well plates (Thermo Scientific AB3384B) with a total volume of 20 μL. The primers used (Table S1) were ordered from Integrated DNA Technologies and brought up to 100 μM in biological grade H_2O. The probe uses Integrated DNA Technologies' dual quencher system with a 5′ Cy5 fluorophore, TAO quencher in the middle of the oligo, and a 3′ Iowa Black RQ-Sp quencher. The final concentration in each well was 0.25 nM for the forward and reverse primers and 0.15 nM for the probe. The total amount of gDNA in each well was 100 ng. Each sample was run in duplicate. Cycling

Table 2. Top clonal expansions in PC9

Sample	Position ID	No. of cells	Nearest gene	Distance to nearest gene, bp	Location	Nearest oncogene	Distance to Nearest oncogene, bp	Oncogene location
Liver no. 8	chr25 + 15086085.1	11	TNFRSF19	10,386	Upstream	FGF9	1,544,240	Upstream
Spleen no. 6	chr16–55072312.1	9	CSMD1	68,403	Downstream	MCPH1	3,234,449	Upstream
Spleen no. 7	chr7+2395624.1	7	CAMSAP2	0	Intron	KIF14	152,627	Upstream
Spleen no. 6	chr30 + 9893210.1	7	UBR1	0	Intron	HSP90AA1	98,747	Downstream
Liver no. 7	chr19–33689377.1	7	DPP10	0	Intron	CKS1B	300,947	Downstream
Spleen no. 8	chr5+18591802.1	6	NXPE2	13,326	Upstream	NNMT	434,715	Upstream

conditions were as follows: 3 min at 95°C, 15 s at 95°C, 30 s at 60°C, read Cy5 fluorescence, and repeat steps 2–4 40 times. Vector copies were quantified by comparing Ct values against a standard of known copy numbers made from the plasmid used in the production of the AAV vector. The standard used 10-fold dilutions from 1×10^{11} down to 1 genome per well.

VCNs were determined by dividing the number of vector genomes by the number of diploid genomes present in the gDNA sample, which was estimated based on an average of 5.5 pg gDNA per diploid canine cell.[8]

Integration analysis

Integration analysis was performed as previously described using sonic abundance ligation–mediated PCR.[8,36] In brief, gDNA was sheared to a size of ~1,000 bp, purified, and ligated to barcoded linkers. Ligated gDNA then underwent 2 different PCR protocols that used linker-specific primers mixed with virus-specific primers to enrich integration sites. Nested PCR reactions with primers landing near the inner edge of ITR sequences (D region) selectively amplified genomic fragments containing AAV/genomic DNA juncture sequences (Table S1). Samples were purified post-PCR, pooled, and quantified with an Illumina MiSeq. AAV genome integrations were determined using the AAVenger software pipeline, had ITR sequencing remnants removed, and then were aligned to the canine genome to identify integration positions. Flip and flop ITR forms were observed when ITR remnants were not largely truncated or rearranged. The number of unique fragment lengths at each position were then used to determine clonal abundance. The list of considered oncogenes and their sources for the informatic analysis is available at http://www.bushmanlab.org/links/genelists. Vector integrations per cell were estimated similarly to VCNs.

DATA AND CODE AVAILABILITY

All of the original data are available from the authors upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2023.101159.

ACKNOWLEDGMENTS

The authors sincerely thank the owners of PC1 and PC9 for entrusting their pets to this research study. We also acknowledge research funding from NIH/National Heart, Lung, and Blood Institute (HL158781 to V.R.A. and B.J.S.-J. and N01_7N92019D00041 to T.C.N.). L.V.G. is supported by the NIH-sponsored Hematology Research Training Program at the University of Pennsylvania (HL07439). This study was also supported in part by the Companion Animal Research Fund from the Department of Clinical Sciences and Advanced Medicine, University of Pennsylvania School of Veterinary Medicine and the Institute for Translational Medicine and Therapeutics of the Perelman School of Medicine and the School of Veterinary Medicine at the University of Pennsylvania. Work in the Bushman laboratory was supported in part by P30AI045008, U01AI125051, R01CA241762, UM1AI164570, R01HL142791, R01HL160748, U19AI174998, and U19AI149680. We also thank the Penn Vet Comparative Pathology Core (CPC) for their help with tissue preparation and imaging. The Penn Vet CPC is supported by the Abramson

Table 3. Top clonal expansions in PC1

Sample	Position ID	No. of cells	Nearest gene	Distance to nearest gene, bp	Location	Nearest oncogene	Distance to nearest oncogene, bp	Oncogene location
Liver #2	chr25 + 34466837.1	5	PEBP4	0	Intron	EGR3	124,451	Upstream
Liver #3	chr3+48611060.1	3	SV2B	78,585	Upstream	AKAP13	380,379	Upstream
Liver #3	chr1-19778565.1	2	LINC01416	18,111	Upstream	MBD2	1,457,646	Upstream
Liver #3	chr1-54737525.1	2	RPS6KA2	0	Intron	FGFR1OP	115,635	Upstream
Liver #3	chr1-88043170.1	2	FAM189A2	0	Intron	TJP2	82,973	Upstream
Liver #3	chr10–21814404.1	2	SULT4A1	0	Intron	PARVB	213,787	Downstream

Cancer Center Support Grant (P30 CA016520). The Aperio Versa 200 Scanner used for whole-slide imaging was acquired thanks to a NIH Shared Instrumentation Grant (S10 OD123465-01A1). The Leica BOND RXm automated instrument for IHC was acquired through Penn Vet's Institute for Infectious and Zoonotic Diseases Core Pilot Grant Opportunity 2022. The graphical abstract was created with BioRender.

AUTHOR CONTRIBUTIONS

L.V.G. designed and performed experiments, analyzed data, and wrote the manuscript. B.S.D., E.W., E.C., J.K.E., and E.P.M. designed and performed experiments, analyzed data, and revised the manuscript. T.C.N., F.D.B., and M.B.C. designed and supervised experiments, analyzed data, and revised the manuscript. V.R.A. designed and supervised experiments but passed away before the completion of the manuscript. B.J.S.-J. designed and supervised experiments, analyzed data, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Schnepp, B.C., Clark, K.R., Klemanski, D.L., Pacak, C.A., and Johnson, P.R. (2003). Genetic fate of recombinant adeno-associated virus vector genomes in muscle. J. Virol. *77*, 3495–3504.
2. Schnepp, B.C., Jensen, R.L., Chen, C.L., Johnson, P.R., and Clark, K.R. (2005). Characterization of adeno-associated virus genomes isolated from human tissues. J. Virol. *79*, 14793–14803.
3. Sabatino, D.E., Bushman, F.D., Chandler, R.J., Crystal, R.G., Davidson, B.L., Dolmetsch, R., Eggan, K.C., Gao, G., Gil-Farina, I., Kay, M.A., et al. (2022). Evaluating the state of the science for adeno-associated virus integration: An integrated perspective. Mol. Ther. *30*, 2646–2663.
4. Chandler, R.J., Sands, M.S., and Venditti, C.P. (2017). Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models. Hum. Gene Ther. *28*, 314–322.
5. Donsante, A., Vogler, C., Muzyczka, N., Crawford, J.M., Barker, J., Flotte, T., Campbell-Thompson, M., Daly, T., and Sands, M.S. (2001). Observed incidence of tumorigenesis in long-term rodent studies of rAAV vectors. Gene Ther. *8*, 1343–1346.
6. Donsante, A., Miller, D.G., Li, Y., Vogler, C., Brunt, E.M., Russell, D.W., and Sands, M.S. (2007). AAV vector integration sites in mouse hepatocellular carcinoma. Science *317*, 477.
7. Chandler, R.J., LaFave, M.C., Varshney, G.K., Trivedi, N.S., Carrillo-Carrasco, N., Senac, J.S., Wu, W., Hoffmann, V., Elkahloun, A.G., Burgess, S.M., and Venditti, C.P. (2015). Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. J. Clin. Invest. *125*, 870–880.
8. Nguyen, G.N., Everett, J.K., Kafle, S., Roche, A.M., Raymond, H.E., Leiby, J., Wood, C., Assenmacher, C.A., Merricks, E.P., Long, C.T., et al. (2021). A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells. Nat. Biotechnol. *39*, 47–55.
9. Gil-Farina, I., Fronza, R., Kaeppel, C., Lopez-Franco, E., Ferreira, V., D'Avola, D., Benito, A., Prieto, J., Petry, H., Gonzalez-Aseguinolaza, G., and Schmidt, M. (2016). Recombinant AAV Integration Is Not Associated With Hepatic Genotoxicity in Nonhuman Primates and Patients. Mol. Ther. *24*, 1100–1105.
10. Konkle, B.A., Walsh, C.E., Escobar, M.A., Josephson, N.C., Young, G., von Drygalski, A., McPhee, S.W.J., Samulski, R.J., Bilic, I., de la Rosa, M., et al. (2021). BAX 335 hemophilia B gene therapy clinical trial results: potential impact of CpG sequences on gene expression. Blood *137*, 763–774.
11. Schmidt, M., Foster, G.R., Coppens, M., Thomsen, H., Dolmetsch, R., Heijink, L., Monahan, P.E., and Pipe, S.W. (2023). Molecular Evaluation and Vector Integration Analysis of HCC Complicating AAV Gene Therapy for Hemophilia B. Blood Adv. *7*, 4966–4969.
12. Retson, L., Tiwari, N., Vaughn, J., Bernes, S., Adelson, P.D., Mansfield, K., Libertini, S., Kuzmiski, B., Alecu, I., and Gabriel, R. (2023). Epithelioid Neoplasm of the Spinal Cord in a Child with Spinal Muscular Atrophy Treated with Onasemnogene Abeparvovec. Mol. Ther. *31*, 2991–2998.
13. Edwards, D.S., Henley, W.E., Harding, E.F., Dobson, J.M., and Wood, J.L.N. (2003). Breed incidence of lymphoma in a UK population of insured dogs. Vet. Comp. Oncol. *1*, 200–206.
14. Atherton, M.J., and Mason, N.J. (2022). Bite-size introduction to canine hematologic malignancies. Blood Adv. *6*, 4073–4084.
15. Vail, D.M., and MacEwen, E.G. (2000). Spontaneously occurring tumors of companion animals as models for human cancer. Cancer Invest. *18*, 781–792.
16. Gao, G., Lu, Y., Calcedo, R., Grant, R.L., Bell, P., Wang, L., Figueredo, J., Lock, M., and Wilson, J.M. (2006). Biology of AAV serotype vectors in liver-directed gene transfer to nonhuman primates. Mol. Ther. *13*, 77–87.
17. Wang, L., Bell, P., Somanathan, S., Wang, Q., He, Z., Yu, H., McMenamin, D., Goode, T., Calcedo, R., and Wilson, J.M. (2015). Comparative Study of Liver Gene Transfer With AAV Vectors Based on Natural and Engineered AAV Capsids. Mol. Ther. *23*, 1877–1887.
18. Breuer, C.B., Hanlon, K.S., Natasan, J.S., Volak, A., Meliani, A., Mingozzi, F., Kleinstiver, B.P., Moon, J.J., and Maguire, C.A. (2020). In vivo engineering of lymphocytes after systemic exosome-associated AAV delivery. Sci. Rep. *10*, 4544.
19. George, L.A., Monahan, P.E., Eyster, M.E., Sullivan, S.K., Ragni, M.V., Croteau, S.E., Rasko, J.E.J., Recht, M., Samelson-Jones, B.J., MacDougall, A., et al. (2021). Multiyear Factor VIII Expression after AAV Gene Transfer for Hemophilia A. N. Engl. J. Med. *385*, 1961–1973.
20. Hjalgrim, H., Askling, J., Rostgaard, K., Hamilton-Dutoit, S., Frisch, M., Zhang, J.S., Madsen, M., Rosdahl, N., Konradsen, H.B., Storm, H.H., and Melbye, M. (2003). Characteristics of Hodgkin's lymphoma after infectious mononucleosis. N. Engl. J. Med. *349*, 1324–1332.
21. Callan, M.B., Haskins, M.E., Wang, P., Zhou, S., High, K.A., and Arruda, V.R. (2016). Successful Phenotype Improvement following Gene Therapy for Severe Hemophilia A in Privately Owned Dogs. PLoS One *11*, e0151800.
22. Sarkar, R., Mucci, M., Addya, S., Tetreault, R., Bellinger, D.A., Nichols, T.C., and Kazazian, H.H., Jr. (2006). Long-term efficacy of adeno-associated virus serotypes 8 and 9 in hemophilia a dogs and mice. Hum. Gene Ther. *17*, 427–439.
23. Finn, J.D., Ozelo, M.C., Sabatino, D.E., Franck, H.W.G., Merricks, E.P., Crudele, J.M., Zhou, S., Kazazian, H.H., Lillicrap, D., Nichols, T.C., and Arruda, V.R. (2010). Eradication of neutralizing antibodies to factor VIII in canine hemophilia A after liver gene therapy. Blood *116*, 5842–5848.
24. Sabatino, D.E., Lange, A.M., Altynova, E.S., Sarkar, R., Zhou, S., Merricks, E.P., Franck, H.G., Nichols, T.C., Arruda, V.R., and Kazazian, H.H., Jr. (2011). Efficacy and safety of long-term prophylaxis in severe hemophilia A dogs following liver gene therapy using AAV vectors. Mol. Ther. *19*, 442–449.
25. McIntosh, J., Lenting, P.J., Rosales, C., Lee, D., Rabbanian, S., Raj, D., Patel, N., Tuddenham, E.G.D., Christophe, O.D., McVey, J.H., et al. (2013). Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant. Blood *121*, 3335–3344.
26. Nguyen, G.N., George, L.A., Siner, J.I., Davidson, R.J., Zander, C.B., Zheng, X.L., Arruda, V.R., Camire, R.M., and Sabatino, D.E. (2017). Novel factor VIII variants with a modified furin cleavage site improve the efficacy of gene therapy for hemophilia A. J. Thromb. Haemostasis *15*, 110–121.
27. Siner, J.I., Iacobelli, N.P., Sabatino, D.E., Ivanciu, L., Zhou, S., Poncz, M., Camire, R.M., and Arruda, V.R. (2013). Minimal modification in the factor VIII B-domain sequence ameliorates the murine hemophilia A phenotype. Blood *121*, 4396–4403.
28. Siner, J.I., Samelson-Jones, B.J., Crudele, J.M., French, R.A., Lee, B.J., Zhou, S., Merricks, E., Raymer, R., Nichols, T.C., Camire, R.M., and Arruda, V.R. (2016). Circumventing furin enhances factor VIII biological activity and ameliorates bleeding phenotypes in hemophilia models. JCI Insight *1*, e89371.

29. Hafenrichter, D.G., Wu, X., Rettinger, S.D., Kennedy, S.C., Flye, M.W., and Ponder, K.P. (1994). Quantitative evaluation of liver-specific promoters from retroviral vectors after *in vivo* transduction of hepatocytes. Blood *84*, 3394–3404.

30. van Dongen, J.J., Krissansen, G.W., Wolvers-Tettero, I.L., Comans-Bitter, W.M., Adriaansen, H.J., Hooijkaas, H., van Wering, E.R., and Terhorst, C. (1988). Cytoplasmic expression of the CD3 antigen as a diagnostic marker for immature T-cell malignancies. Blood *71*, 603–612.

31. Riondato, F., and Comazzi, S. (2021). Flow Cytometry in the Diagnosis of Canine B-Cell Lymphoma. Front. Vet. Sci. *8*, 600986.

32. Torlakovic, E., Torlakovic, G., Nguyen, P.L., Brunning, R.D., and Delabie, J. (2002). The value of anti-pax-5 immunostaining in routinely fixed and paraffin-embedded sections: a novel pan pre-B and B-cell marker. Am. J. Surg. Pathol. *26*, 1343–1350.

33. Bushman, F.D., Cantu, A., Everett, J., Sabatino, D., and Berry, C. (2021). Challenges in estimating numbers of vectors integrated in gene-modified cells using DNA sequence information. Mol. Ther. *29*, 3328–3331.

34. Chao, A., Chazdon, R.L., Colwell, R.K., and Shen, T.J. (2006). Abundance-based similarity indices and their estimation when there are unseen species in samples. Biometrics *62*, 361–371.

35. Ning, Y., Hui, N., Qing, B., Zhuo, Y., Sun, W., Du, Y., Liu, S., Liu, K., and Zhou, J. (2019). ZCCHC10 suppresses lung cancer progression and cisplatin resistance by attenuating MDM2-mediated p53 ubiquitination and degradation. Cell Death Dis. *10*, 414.

36. Berry, C.C., Gillet, N.A., Melamed, A., Gormley, N., Bangham, C.R.M., and Bushman, F.D. (2012). Estimating abundances of retroviral insertion sites from DNA fragment length data. Bioinformatics *28*, 755–762.

37. Huang, R.Q., Wang, S.Q., Zhu, Q.B., Guo, S.C., Shi, D.L., Chen, F., Fang, Y.C., Chen, R., and Lu, Y.C. (2019). Knockdown of PEBP4 inhibits human glioma cell growth and invasive potential via ERK1/2 signaling pathway. Mol. Carcinog. *58*, 135–143.

38. Zhang, D., Dai, Y., Cai, Y., Suo, T., Liu, H., Wang, Y., Cheng, Z., and Liu, H. (2016). PEBP4 promoted the growth and migration of cancer cells in pancreatic ductal adenocarcinoma. Tumour Biol. *37*, 1699–1705.

39. Batty, P., Mo, A.M., Hurlbut, D., Ishida, J., Yates, B., Brown, C., Harpell, L., Hough, C., Pender, A., Rimmer, E.K., et al. (2022). Long-term follow-up of liver-directed, adeno-associated vector-mediated gene therapy in the canine model of hemophilia A. Blood *140*, 2672–2683.

40. Fagerberg, L., Hallström, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell. Proteomics *13*, 397–406.

41. Kapelanski-Lamoureux, A., Chen, Z., Gao, Z.H., Deng, R., Lazaris, A., Lebeaupin, C., Giles, L., Malhotra, J., Yong, J., Zou, C., et al. (2022). Ectopic clotting factor VIII expression and misfolding in hepatocytes as a cause for hepatocellular carcinoma. Mol. Ther. *30*, 3542–3551.