AAV BASED SITE-SPECIFIC INTEGRATION *IN VIVO*

by

Zhongya Wang

A THESIS

Presented to the Department of Cell & Developmental Biology
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of
Master of Science

December 2007

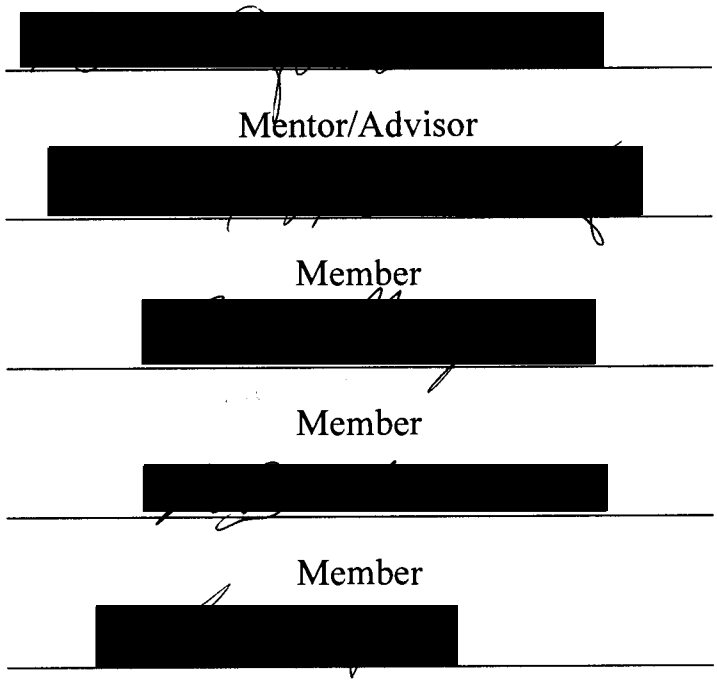
Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is certify that Master's thesis of

Zhongya Wang

has been approved



Mentor/Advisor

Member

Member

Member

Member

Abstract

Recombinant adeno-associated virus (AAV) vectors are mostly episomal and rarely integrate into the host genome. This feature of random integration poses significant safety concerns regarding the use of AAV in clinical gene therapy. When integration does occur, AAV can integrate randomly throughout the genome. For the purpose of gene therapy for genetic diseases, an integrating vector capable of site-specific integration would be ideal. In our lab, we have constructed an AAV vector in which a human fumarylacetoacetate hydrolase (Fah) expression cassette is flanked by ~ 2 kb of homologous rDNA in the region of the I-PpoI site. We hypothesized that the AAV vector could be targeted to this location by homologous recombination facilitated by the presence of a double strand DNA break created by I-ppoI. To prove our hypothesis, *Fah*$^{-/-}$ mice were injected with high dose of viral particles ($3x10^{11}$ vg) and phenotypic correction was measured by weight gain after NTBC withdrawal. Adult *Fah*$^{-/-}$ mice injected with a dose of $3x10^{11}$ particles (high dose) gained weight after NTBC withdrawal while control mice died after 4-6 weeks. $1x10^{9}$ AAV8-rDNA-Fah (low dose) rescued *Fah*$^{-/-}$ mice with initial weight loss, followed by weight gain. Next, the hepatocytes of weight-stabilized mice were serially transplanted into secondary *Fah*$^{-/-}$ recipients. All secondary recipients displayed weight gain after transplantation indicating the presence of a stably integrated Fah expression cassette. Site-specific junction PCR was done to confirm the presence of our hFah expression cassette within the rDNA. In addition, junction PCR would enable us to differentiate homologous recombination events from non-homologous end joining and thus was applied to all the mice including primary mice, transplanted mice, and neonatal mice injected with our AAV-rDNA-Fah. Site-specific junction

fragments were detected in all the mice, including neonatal mice. The sequence results of the junction fragments generated by site-specific PCR exactly matched the predicted sequence. Finally, a dose dependent comparison between AAV2-rDNA-Fah and AAV2-Fah showed that AAV2-rDNA-Fah can rescue the *Fah*$^{-/-}$ mice at significantly a lower dose (about 1/30-1/10) compared to the dose required for AAV2-Fah. AAV8-rDNA-Fah is able to rescue *Fah*$^{-/-}$ mice at a dose of $1x10^9$. Our strategy was also tested in a hemophilia B mouse model. AAV-rDNA-hFIX was constructed and injected into wild type mice. One month after injection, mice underwent partial hepatectomy (PHx) to remove most of the episomal AAV genome. Surprisingly, the post-PHx and pre-PHx ratio of hFIX level was significantly higher using AAV8-rDNA-hFIX and AAV2-rDNA-hFIX (40-50% and 70-80% respectively), compared to only 5-10% after PHx with conventional AAV-hFIX. The results are promising and may be of potential use in human gene therapy trials for hemophilia given the stable integration achieved with high efficiency. Surprisingly, we also detected evidence of site-specific integration systemically in other tissues such as the lung, heart, muscle and kidney in addition to liver as early as 3 days post injection. This highly suggests that AAV8-rDNA is capable of integrating systemically *in vivo*. This is the first report of site-specific integration in an organ system other than liver. This finding is significant for gene therapy of certain diseases that systemically affect patients. It is possible that with an appropriate tissue-specific promoter or with different AAV serotypes, AAV-rDNA can be specifically used in other tissues. Thereby, our AAV-rDNA vector is superior to currently existing AAV vectors due to its ability to integrate site-specifically within ribosomal DNA with high efficiency at a low rescue dose. Our novel vector opens the possibility for a new strategy for gene delivery for use in clinical gene therapy trials.

Part I Introduction

The adeno-associated virus (AAV) belongs to the dependovirus genus of the parvovirus family. The wildtype form of AAV is a non-enveloped virus with a 22 nm capsid, which contains a 4.7 kb single strand DNA genome with either plus or minus strand orientation. The life cycle of AAV has two phases: a productive phase, in which, AAV requires facilitation by a helper virus, like adenovirus, in order to replicate and infect new host cells; and a latent phase, in which, AAV can integrate into host genome (Berns *et al.*1996). The AAV genome is composed of two open reading frames (ORFs) and two inverted terminal repeats (ITRs). The left ORF encodes for four replication proteins: Rep 48, Rep 52, Rep 68, and Rep 78, which are responsible for latent integration and regulate the replication of the AAV genome. The right ORF encodes for structural proteins that form the capsid of the AAV particle. The 147 bps inverted terminal repeats (ITRs) form a T-shaped secondary structure with a free 3' hydroxyl group, which is necessary for the initiation of AAV genome replication (Berns. 1990).

The successful cloning of the AAV virus has made it possible to engineer recombinant AAV vectors capable of carrying foreign DNA that is expressed in mammalian cells (Hermonat *et al.* 1984). The unique ability of AAV to exist either in a productive phase (episomal) or latent

phase (integrated into host chromosome), as well as its non-pathogenicity are the advantages of AAV for its use as a gene therapy vector (McCarty *et al.* 2004).

A vector capable of permanently introducing exogenous DNA into the host genome would be superior to episomal delivery of target genes. Use of AAV as a gene therapy vector was driven largely by its ability to integrate into host chromosomes and allow stable expression of the transgene and allow transfer to daughter cells during cell division (McCarty *et al.* 2004). However recombinant AAV (rAAV) integration is inefficient with an integration frequency of only $0.2\text{-}1.0\text{x}10^{-3}$ in Rep-independent manner. In fact wild type AAV also integrates into chromosome 19 inefficiently at a rate of only 0.1-0.5% in Rep dependent manner (Flotte *et al.* 1994; Mccarty *et al.* 2004). For therapeutic purposes, integration efficiency is too low to provide clinical benefit. In addition, random integration randomly spread throughout the host genome with preferential targeting of actively transcribed genes and chromosome breakage points raises concerns about the safety of AAV and the potential for malignant transformation as a result of random integration into a promoter or tumor suppressor gene (Nakai *et al.* 2003; Miller *et al.* 2004). It has been reported that AAV carrying homologous sequences is able to correct point mutations within the host genome via homologous recombination. In addition, Miller *et al* have reported that

increased efficiency of correction of point mutations in the presence of double strand breakage sites in vitro in cultured cells (Miller *et al.* 2002; and Miller *et al.* 2003). Further support for correction of gene mutations by homologous recombination also come from Miller *et al* and Worthon *et al* who report successful targeted correction *in vivo* in 3 different mouse models including RosaAnZΔ4, MPS VII and HT1. They noted a correction efficiency of 10^{-4}-10^{-5} that was significantly below what is required for therapeutic correction and noted that the efficiency of random integration was higher than that of targeted integration (Miller *et al.* 2006; Worthon *et al.* unpublished data). This low efficiency of integration currently limits the application of rAAV based gene correction for use in clinical gene therapy trials. An alternative strategy in which a novel rAAV capable of site specific integration with high efficiency is greatly needed and would be of value in current human gene therapy trials.

One of the known hotspots of rAAV integration is within the ribosomal DNA locus and could serve as a potential target for directing site-specific integration. This locus is a region that is actively transcribed with more than 200 copies present in a mouse diploid genome (Nakai *et al.* 2003). This makes ribosomal DNA ideal as a site-specific integration target. In addition there is an endonuclease, I-PpoI, which specifically recognizes a 15bp sequence inside the rDNA gene, cleaves ribosomal DNA to generate a double stand break, and targets

site- specific integration of the rAAV via homologous recombination (Lin *et al.* 1998). To test our hypothesis, we generated a novel AAV construct containing an Fah expression cassette flanked by homologous rDNA sequences (AAV-rDNA-Fah) in which there was an I-PpoI recognition site. The I-PpoI recognition site within the AAV-rDNA-Fah was disrupted by insertion of the hFah expression cassette so that the rDNA-hFah was resistant to cleavage by the I-PpoI nuclease. Our vector was also tested in second disease model, the hemophilia B mouse model with AAV-rDNA-hFIX to confirm that site specific integration was not being directed by an intrinsic property of the Fah expression cassette. Here we describe a novel recombinant AAV capable of site-specific integration in the liver with improved efficiency compared to conventional AAV vectors and enhanced integration in the presence of double strand DNA breaks. Site integration was observed systemically as well in other tissues such as the lung, heart, kidneys, and muscle, in addition to the liver. Our new vector provides a safe and effective method for delivering therapeutic levels of target genes for the genetic correction of many disease states and is ideal for use in human gene therapy clinical trials.

Part II Materials and methods

Plasmids and rAAV vectors

The construction of pAAV-rDNA-Fah was based on a pAAV vector, a gift from Mark Kay's lab at Stanford University, Palo Alto, CA. This vector encodes a 2038 bp rDNA fragment cloned from HT1 mice liver genomic DNA, a respiratory syncytial virus (RSV) promoter, a human fumarylacetoacetate hydrolase (*Fah*) cDNA, and a poly(A) signal derived from the human β-globin gene (βglpA) between two ITRs, shown in figure 1. The bacterial β-lactamase (ampicillin resistance or Ampr) gene, the ColE1 plasmid origin of replication (ori), and a 1.6-kb of filler DNA from the bacterial *lacZ* gene were placed outside the two AAV-ITRs within the plasmid backbone. To construct pAAV-rDNA-hFIX, a fragment containing the RSV-hFAH-poly(A) expression cassette was replaced with human Factor IX gene (*FIX*) cDNA from pNEB193-SynEnh-TTR-hFIX10-90-spa. pAAV-I-PpoI was based on pAAV vector and CMV-NLS-I-PpoI expression cassette which was subcloned from pCNPpo6 plasmid, a gift from Ray Monnat at the University of Washington, Seattle, WA. The rAAVs were produced in HEK293 cells by the triple transfection method, purified and titer counts obtained at Stanford University.

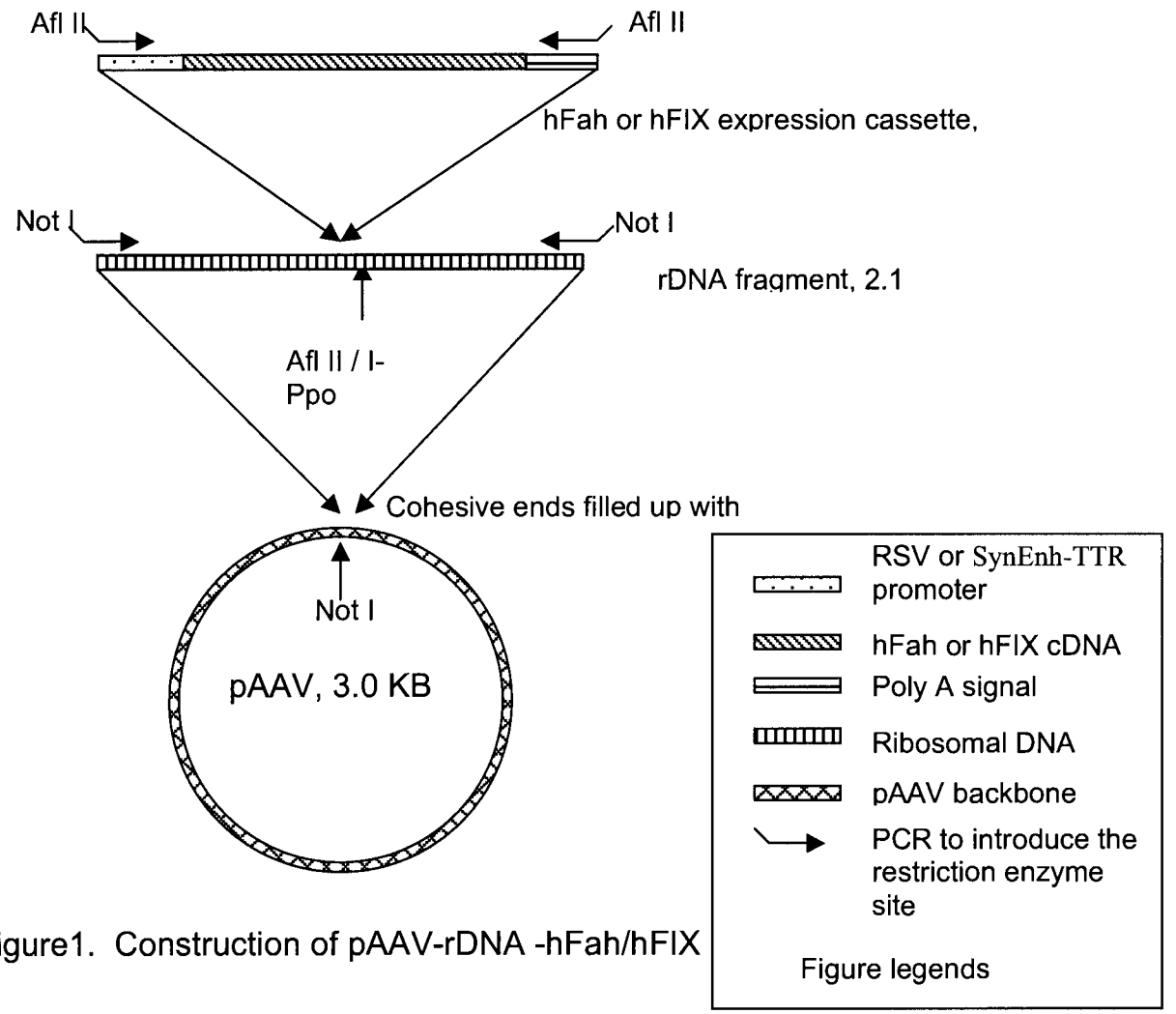


Figure1. Construction of pAAV-rDNA -hFah/hFIX

Mouse strains and animal husbandry

Fah mutant (Grompe *et al.* 1993), wild type mice were inbred on a 129S4 background. Fah null mice were maintained on drinking water containing NTBC (8 mg/L) (Grompe *et al.* 1995). All the experiments

were performed under the OHSU Institutional Animal Care and Use Committee Guide.

Partial hepatectomy

Liver resection of the left lateral and median lobes was performed after a midventral laparotomy. Surgery was performed under general anesthesia (Higgins *et al.* 1931). The mortality rate following PH was <5%. Following PH, liver tissue was rinsed with PBS to eliminate blood cells, then minced, and divided into aliquots for further processing.

Hepatocyte transplantation

After AAV infection, *Fah*$^{-/-}$ mice were withdrawn from NTBC for 12 weeks to select for AAV-rDNA-Fah integrated hepatocytes. To further select the AAV-rDNA-Fah integrated hepatocytes, serial transplantation was also performed. For serial transplantation, hepatocytes were isolated by a standard two-step collagenase perfusion (Berry *et al.* 1969) and enriched to >95% purity by differential centrifugation (Overturf *et al.* 1999). A total of 300,000 viable hepatocytes were injected intrasplenically into *Fah*$^{-/-}$ recipients as described (Overturf *et al.* 1996) and NTBC was stopped immediately after transplantation to allow for selection of Fah positive cells.

Retro orbital vein injection, *in vivo* selection

AAV was injected into 129S4 *Fah*$^{-/-}$ or wild type mice via retro orbital vein injection. We injected adult male HTI mice on NTBC with $3.0X10^{11}$ particles of AAV-rDNA-Fah with or without $5x10^{10}$ AAV8-I-PpoI via the retro orbital vein ($n = 8$). After injection, we withdrew NTBC from *Fah*$^{-/-}$ mice to select for hepatocytes containing integrated vector genomes and remove hepatocytes containing only non-integrated episoma AAV. After selection for 16 weeks *in vivo*, the liver was harvested. For the FIX experiment, after AAV injection, wild type mice were euthanized on the date as indicated; tissues were harvested, and frozen at -80°C.

Junction PCR and quantitative real time PCR

The following primers were used for the hFah junction PCR:

Junction PCR 34:

Forward 5'-CCGCTTTTCGCCTAAACACAC-3'

Reverse 5'-GCCGTATCGTTCCGCCTG-3'

Junction PCR 45:

Forward 5'-AACACACCCTAGTCCCCTCAGATAC-3'

Reverse 5'-CACCCCGTTTCCCAAGACG-3'

The following primers were used for the hFIX junction PCR:

Junction1: Forward 5'-GAGTATCGGAACACTCGCTCTACG-3'

Reverse 5'-CTGGGCGGGATTCTGACTTAG-3'

Junction2: Forward 5'-CGGAACACTCGCTCTACGAAATG-3'

Reverse 5'-TTCCCAAGACGAACGGCTCTCC-3'

Junction3: Forward 5'-CACTCGCTCTACGAAATGTGCAG-5'

Reverse 5'-CCCCGTTTCCCAAGACGAAC-3'

Junction4: Forward 5'-TCTGCAAGGGTCATCAGTAGTTTT-3'

Reverse 5'-TGTCGAGGGCTGACTTTCAATAG-3'

Junction5: Forward 5'-CCATCTTACTCAACATCCTCCCAG-3'

Reverse 5'-TGTCGAGGGCTGACTTTCAATAG-3'

(shared with junction 4)

Junction PCR was run under the following conditions: 94°C for 5min for hot start; 94°C for 1min for denaturation, 65°C >55 °C (touch down) for 1min for 10 cycles, 72 °C for 3 min; 94°C for 1min, 55 °C for 1min, 72 °C for 3 min for extension and run for 30 cycles with a final extension step at 72 °C for 10 min.

Quantitative real time PCR:

The following primers were used for real time PCR:

MAAI: Forward 5'GTACCATTGAGGTGGGCCTA-3'

Reverse 5'GCTGGTTCGTCCACTCTTTC-3'

Rosa (beta-Gal): Forward 5'-GTGCGGATTGAAAATGGTCT-3'

Reverse 5'-GACCTGACCATGCAGAGGAT-3'

hFah: Forward 5'-CATAACAGGTACTGCCAGG-3'

Reverse 5'-GAGAAAATCTCATGGCAGG-3'

PCR conditions were as follows:

Hot start at 95.0 °C for 03:00 followed by 45 cycles of denaturation, annealing, and extension under conditions below.

Cycle 2: (45X)

Step 1:	95.0 °C	for 00:15.
Step 2:	64.0 °C	for 00:20.
Step 3:	72.0 °C	for 00:20.

Data collection and real-time analysis enabled.

Cycle 3: (101X)

Step 1:	45.0 °C-95.0 °C	for 00:10.

Increase set point temperature after cycle 2 by 0.5 °C

Melt curve data collected and analysis enabled.

The quantitative PCR results were compared to a standard curve generated using serial dilutions of Rosa26 genomic DNA. AAV8-rDNA-Fah treated mice with southern blotting demonstrated about 1 copy/ genome and samples were then quantified based on standard curve data.

Histology and immunohistology

Liver tissues were fixed in 10% phosphate-buffered formalin, pH 7.4, dehydrated in 80% ethanol, and embedded in paraffin wax at 58°C. Four-micrometer sections were rehydrated and stained with a polyclonal chicken antibody against human Fah. The antibody was diluted in PBS, pH 7.4, and applied at concentrations of 1:300,000 at 37°C for 30 min.

Endogenous peroxidase activity was blocked with 3% H_2O_2 and methanol. Avidin and biotin pretreatment was used to prevent endogenous background staining. The secondary antibody was biotinylated goat anti-chicken IgG used at 1:250 dilution (BA-1000; Vector Laboratories, Burlingame, CA, USA).

Liver function tests

Animals were anesthetized with mouse cocktail and blood was collected in Microtainer plasma separator tubes containing lithium heparin (Becton–Dickinson Vacutainer Systems, Franklin Lakes, NJ, USA). After a brief centrifugation, the plasma was frozen at -80°C. Twenty microliters of plasma were mixed with 80 microliters of a 7% (w/v) bovine serum albumin solution and assayed for Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), bilirubin, and conjugated bilirubin levels using a Kodak Ektachem 700 chemistry analyzer (Eastman Kodak, Rochester, NY, USA).

Measurement of human FIX in serum

Enzyme-linked immunosorbent assay specific for human Factor IX was performed to measure human Factor IX in mouse serum.

Part III Results

1. Both the AAV2-rDNA-Fah and AAV8 can rescue *Fah*$^{-/-}$ mice

Fah$^{-/-}$ mice can only survive in the presence of exogenously supplemented Fah enzyme activity or by rescue by the chemical 2-(2-nitro-4-trifluoro-methylbenzyol)-1,3-cyclohexanedione (NTBC). In hepatocytes, NTBC blocks an upstream step in the tyrosine metabolism pathway resulting in accumulation of fumarylacetoacetate (FAA), a toxic metabolite in hepatocytes. Without other intervention, *Fah*$^{-/-}$ mice lose weight and die within 3 weeks of NTBC withdrawal. We hypothesized that our vector would integrate site specifically into the genome of *Fah*$^{-/-}$ hepatocytes and result in phenotypic rescue and weight gain in mice off of NTBC. To test the activity of AAV-rDNA-Fah, $3x10^{11}$ vg AAV2 and AAV8-rDNA-Fah were injected into 5 and 6 *Fah*$^{-/-}$ mice, respectively, via the retro-orbital route. The mice that received the highest dose ($3x10^{11}$ vg) of AAV showed normal behavior, normal urea (data not show) and weight gain for 6 months post- injection, whereas control mice injected with placebo died within 4 weeks (data not shown). This phenotypic correction suggests the Fah expression cassette in the rDNA fragments expressed sufficient functional Fah enzyme and

restored up to 50% Fah enzyme activity and liver function compared to wild type liver (see figure 2A and 2B and data not shown).

2. Injection of AAV-rNDA-Fah leads to integration into the mouse genome

The majority of AAV vectors injected into mice hepatocytes usually remain episomal, mostly as concatamers, although the mechanism remains partly unclear (McCarty *et al.* 2004). Only a very small portion of the AAV vectors will actually integrate into the host genome. Although the integration is randomly spread through the genome, there a a certain number of known integration hot spots, including one in a rDNA locus (Nakai *et al.* 2003).

In contrast to integrated AAV, episomal AAV genome, (irrespective of concatermer status) will be diluted and lost during host cell division and proliferation (Nakai *et al.* 2001; McCarty *et al.* 2004). Only AAV genome that has integrated into host chromosome can be replicated and passed to the daughter cells during hepatocyte division and proliferation. One unique feature of the liver is its capacity for self-regeneration following hepatocyte injury in the form of partial hepatectomy, or chemically induced cell toxicity. A widely used method to stimulate hepatocyte proliferation and liver regeneration *in vivo* is partial hepatectomy (PHx). To confirm AAV-rDNA-Fah integration rather

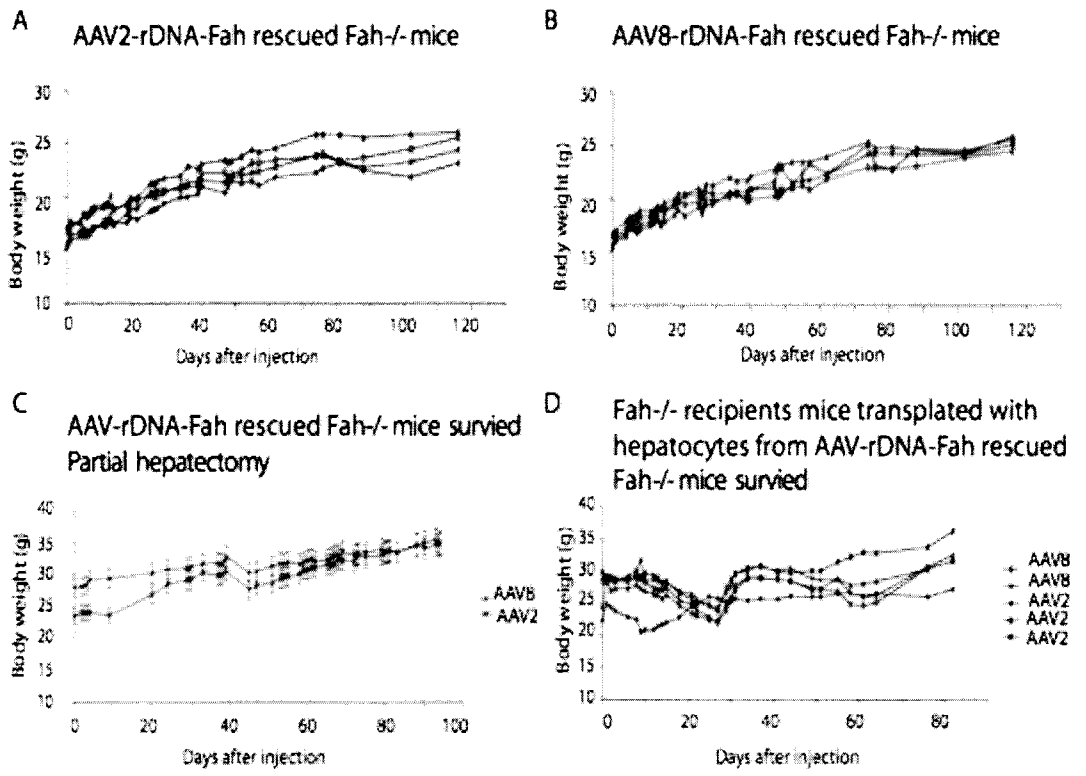


Figure 2: AAV-rDNA-Fah rescue Fah deficient mice with integration

than simply episomal vector in hepatocytes, *Fah*$^{-/-}$ mice injected with AAV2-rDNA-Fah and AAV8-rDNA-Fah at $3x10^{11}$ vg dose received partial hepatectomy. We hypothesized that *Fah*$^{-/-}$ mice can survive partial hepatectomy with restoration of normal liver function only if AAV-rDNA-Fah has been integrated into the genome with expression of hFah protein at sufficient levels for normal Fah enzyme activity. As figure 2C shows, both AAV2 and AAV8-rDNA-Fah treated mice survived the PHx without weight loss and strongly supports the presence of stably integrated AAV-rDNA-Fah within the hepatocytes.

In order to confirm AAV-rDNA-Fah integration into the mouse genome, mice were injected with AAV-rDNA-Fah (both AAV2 and AAV8 serotypes) at a dose of $3x10^{11}$ vg and perfused. Hepatocytes were harvested and then transplanted into secondary *Fah*$^{-/-}$ recipients. In the *Fah*$^{-/-}$ mouse liver, we and others have shown that Fah positive hepatocytes have a selective growth advantage over Fah negative hepatcytes. After transplantation, *Fah*$^{-/-}$ recipients were withdrawn from NTBC, which also protects the *Fah*$^{-/-}$ hepatocytes from genomic toxicity. In addition, AAV-rDNA-Fah transduced hepatocytes exhibit a selective growth advantage and can proliferate and go on to repopulate livers of the *Fah*$^{-/-}$ recipients. After successful gene correction and transplantation, Fah positive hepatocytes will go on to replicate for tens to hundreds of times to repopulate the *Fah*$^{-/-}$ liver (Overturf *et al.* 1996). Only hepatocytes containing an AAV-rDNA-Fah expression cassette that has stably integrated into the host chromosome are capable of surviving the repopulation. In other words, survival of recipients transplanted with hepatocytes from AAV-rDNA-Fah is indicative of and confirms integration of AAV-rDNA-Fah into host genome. As shown in figure 2D, all the *Fah*$^{-/-}$ recipients survived with NTBC support for only 5 days followed by withdrawal. This strongly suggests the existence of AAV-rDNA-Fah genome integration.

3. AAV-rDNA-Fah integrates into the host genome in a site-specific manner

Once we confirmed the integration of AAV-rDNA-Fah into host genome, questions regarding the mechanism arose. First, where exactly does the AAV-rDNA-Fah integrate and does it do so in a site-specific as expected or randomly throughout the genome? And second, what is the mechanism of integration, and is it primarily via homologous recombination (HR) only or is there some contribution via non-homologous end joining (NHEJ)? To address our first question we used junction PCR to screen for the presence of an integrated hFAH expression cassette in the rDNA locus. Primers were designed to specifically target the Fah expression cassette and the rDNA locus sequence in a region not contained within our AAV-rDNA-Fah vector. With this strategy, we expect the PCR to result in a specific band only if AAV-rDNA-Fah had integrated into designed rDNA locus, generating a new rDNA sequence containing the Fah expression cassette. Four primer pairs were designed to target the PCR junction fragment. First we determined the specificity of our 4 primer pairs. Junction PCR of AAV-rDNA-Fah infected liver were run out on agarose gels and a band of the expected size was seen but absent in control $Fah^{-/-}$ genomic DNA plus non-integrated pAAV-rDNA-Fah plasmid, shown in figure 3A. $Fah^{-/-}$ genomic DNA plus pAAV-rDNA-Fah plasmid were used as the negative

control and represent host genome and non-integrated AAV genome, and allowed us to rule out the possibility that a positive signal is the result of template switching. We then tested for the presence of site-specific integration in all mice that were injected. Junction PCRs were applied to all samples from AAV2-rDNA-Fah treated adult *Fah*$^{-/-}$ mice, AAV8-rDNA-Fah treated adult *Fah*$^{-/-}$ mice, *Fah*$^{-/-}$ recipient mice transplanted with hepatocytes from AAV2 or AV8-rDNA-Fah treated *Fah*$^{-/-}$ mice, and AAV8-rDNA-Fah treated neonatal *Fah*$^{-/-}$ mice, shown in figure 3B. PCR results suggest that all the samples from AAV2 or AAV8 treated mice had a positive signal with amplification of the desired junction fragment. In contrast, our negative control (genomic DNA plus pAAV-rDNA-Fah) did not have amplification of a junction fragment. The PCR results confirm site-specific integration of our AAV-rDNA mediated Fah expression cassette into rDNA of hepatocytes at the expected locus. To gain further insight into the integration events, we cloned the junction fragments from individual mice to compare integration patterns. As shown in figure 3C, EcoR I digestion gave the same digestion pattern in each mouse suggesting that the integration events into rDNA locus were identical or at least similar from mouse to mouse. We sequenced 4 cloned junction fragments in total and exact alignment of

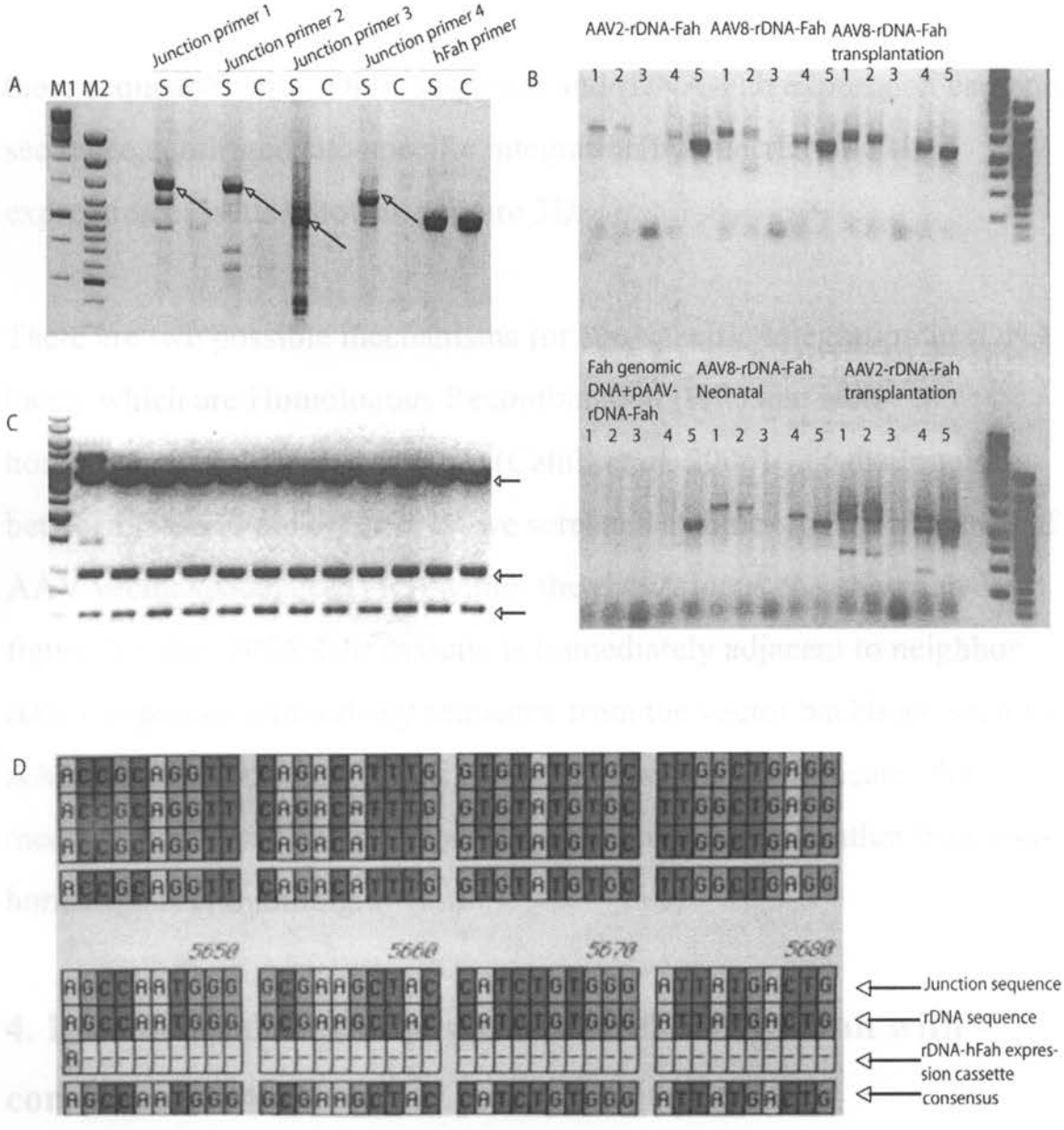


Figure 3 Junction PCR and sequence suggest site specific integration. (A) test junction PCR primers. All the jucntion PCR primers gave specific band (arrowhead pointed)in AAV8-rDNA-Fah injected Fah deficient mouse (S) but not in the mixture of Fah mutant mouse genomic DNA and pAAV-rDNA-Fah (C). (B) Fah deficient mice received AAV2-rDNA-Fah, AAV8-rDNA-Fah, and recipient Fah mutant mice serial transplanted from hepatocytes from AAV2-rDNA-Fah or AAV8-rDNA-Fah injected mice, and AAV8-rDNA-Fah injected Fah deficient neonatal mice have been detected the junction fragments. 1, 2, 3, 4 different junction PCR primer pairs; 5 hFah PCR primer pair (C)EcoR I Digestion of pCR-blunt II -juction PCR fragment from individual mouse gave the same digestion pattern. (D) Alignment of Junction fragment sequence and rDNA sequence, and rDNA-Fah expression cassette sequence prove the site specific integration.

these sequences with rDNA sequence and rDNA-Fah expression cassette sequence confirmed site-specific integration by our rDNA-Fah expression cassette, shown in figure 3D.

There are two possible mechanisms for site-specific integration in rDNA locus, which are Homologous Recombination (HR) and Non-homologous End Joining (NHEJ) (Cahill *et al.* 2006). To distinguish between these two mechanisms, we screened samples for the presence of AAV vector sequence (ITRs) within the rDNA locus. As shown in figure 3D, the rDNA-Fah cassette is immediately adjacent to neighbor rDNA sequence without any sequence from the vector backbone such as AAV inverted terminal repeats present. This strongly indicates the mechanism of integration is homologous recombination rather than non-homologous end joining.

4. Dose dependent comparison of AAV-rDNA-Fah with conventional AAV

We compared the integration efficiency of our novel AAV-rDNA vector to that of the conventional AAV vector to determine if our vector could integrate at higher frequencies. To test our hypothesis, we injected *Fah*$^{-/-}$ mice with either our AAV2-rDNA-Fah or conventional AAV2-Fah at

several different doses of $3x10^{11}$, $1x10^{10}$, $1x10^{9}$ and placebo. After injection, all the mice were withdrawn from NTBC and weight was monitored daily. Weight data here correlates to restoration of Fah enzyme activity and liver function. As shown in figure 4A, AAV2-rDNA-Fah is able to rescue *Fah*$^{-/-}$ mice at a dose of only $1x10^{10}$ with initial weight loss followed by weight gain. The higher dose of $3x10^{11}$ dose is able to rescue *Fah*$^{-/-}$ without weight loss. The lowest dose of $1x10^{9}$ dose is not enough to rescue *Fah*$^{-/-}$ mice after NTBC withdrawal. For AAV2-Fah, only the highest dose of $3x10^{11}$ rescues the *Fah*$^{-/-}$ mice with initial weight loss followed by weight gain, shown in figure 4B. Of note, this pattern of initial weight loss followed by gain is similar to that observed after injection with our novel AAV2-rDNA-Fah at the lower dose of $1x10^{10}$. Mice injected with a lower dose of AAV2-Fah at $1x10^{10}$ were unable to survive and died at earlier time than those that received AAV2-rDNA-Fah at the lowest dose of $1x10^{9}$. Our functional rescue data suggested that AAV2-rDNA-Fah can rescue the *Fah*$^{-/-}$ mouse at a much lower dose (10-30 fold lower) and integrates into the genome at much higher efficiency than conventional AAV2-Fah. Because of slow release of AAV2 genome, it is not an ideal serotype to use for determining the minimal effective dose of AAV-rDNA-Fah. For that reason, we switched to serotype 8 and injected *Fah*$^{-/-}$ mice with AAV8-rDNA-Fah again at decreasing concentrations: $1x10^{10}$, $1x10^{9}$, $3x10^{8}$, $1x10^{8}$, and $3x10^{7}$ and measured dose response in terms of weight gain

after NTBC withdrawal. As shown in figure 4C, AAV8-rDNA-Fah is able to rescue *Fah*$^{-/-}$ mice at a dose of 1x10^{10} without weight loss and at the lower dose of 1x10^{9}, albeit with initial weight loss followed by weight gain after NTBC withdrawal.

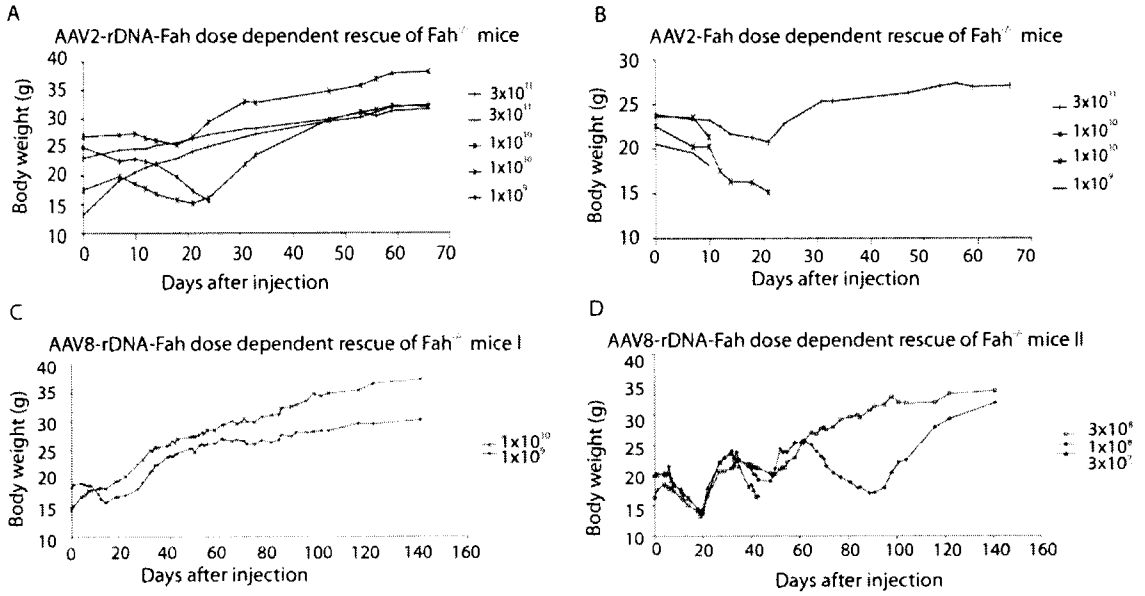


Figure 4. dose dependent comparasion study of AAV-rDNA-Fah.
(A) AAV2-rDNA-Fah is able to rescue Fah deficient mice at middle dose (1x10^{10})with weight lost firstly and weight gained finally while high dose (3x10^{11}) rescues Fah deficient mice without weight lost and low dose (1x10^{9}) is unable to rescue Fah deficient mice. (B) AAV2-Fah is able to rescue Fah deficient miceonly at high dose (3x10^{11}), while Fah deficient mice receiving middle dose (1x10^{10}) is unble to survive. (C) AAV8-rDNA-Fah is able to rescue Fah deficient mice at low dose (1x10^{9}) while middle dose (1x10^{10}) is able to rescue Fah deficient mice without weight lost. (D) Furtherly, 3x10^{9} AAV8-rDNA-Fah is able to rescue Fah deficient mice wtih 1 time NTBC on for 5 days while 1x10^{8} is able to rescue Fah deficient mice with 2 times NTBC on and 3x10^{7}is not able to rescue Fah deficient mice.

At the lowest dose of 3x10^{8}, mice required a brief period of exogenous NTBC rescue after viral injection to rescue the *Fah*$^{-/-}$ mouse and at the lower dose of 1x10^{8}, mice required two periods of NTBC supplementation for 5 days to rescue the *Fah*$^{-/-}$ mice. At the lowest dose of 3x10^{7}, mice were unable to survive as shown in figure 4D.

5. Absolute frequency of AAV-rDNA-Fah integration

As indicated before, the *Fah*$^{-/-}$ mouse is a selective mouse model in which Fah+ hepatocytes will have a selective growth advantage

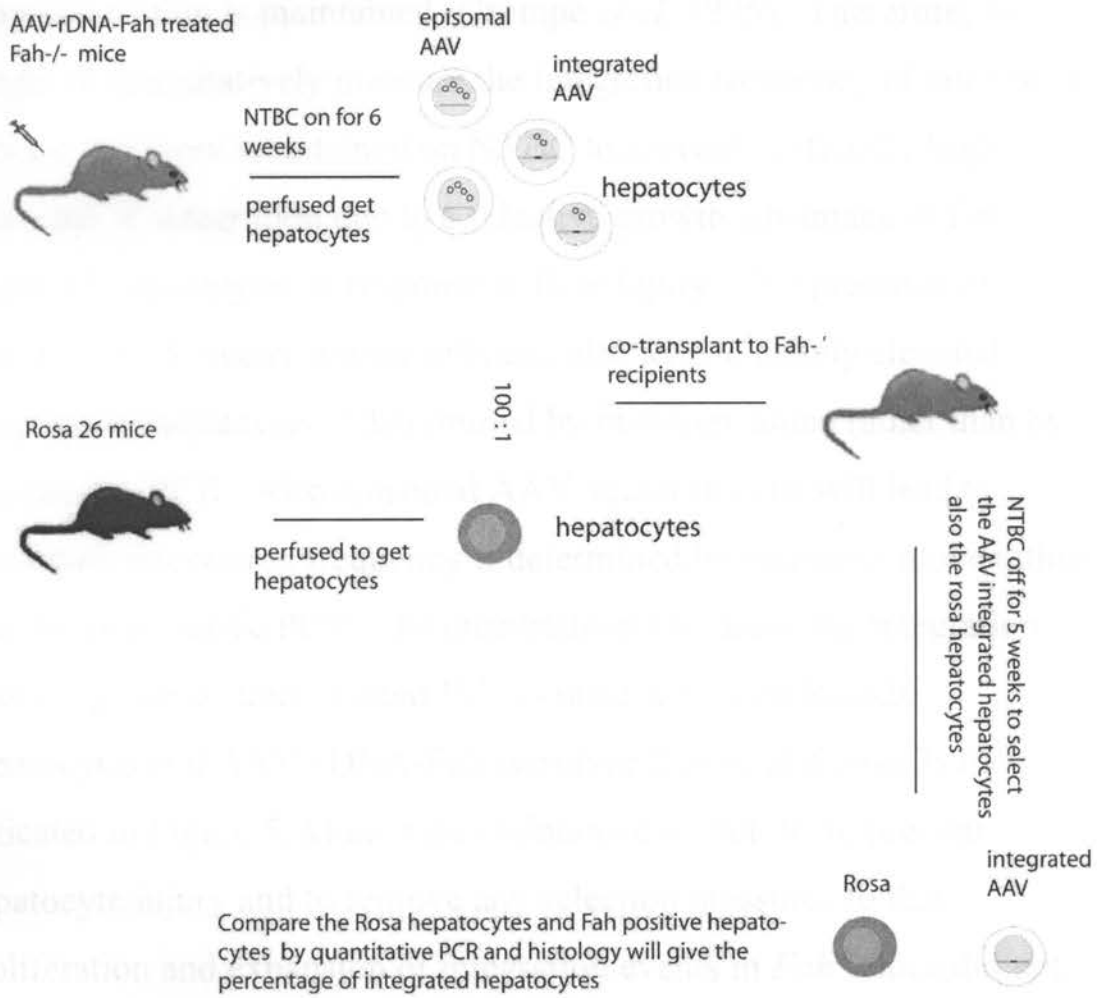
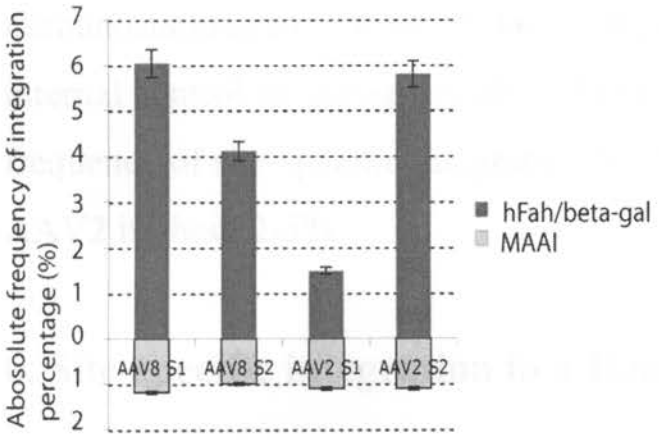


Figure 5. Strategy for absolute integration frequency.

compared to Fah – hepatocytes after NTBC withdrawal. Thus, given this selective growth advantage, it would not be accurate to measure the

absolute frequency of AAV-rDNA-Fah integration in *Fah*$^{-/-}$ mice after NTBC withdrawal. However, previous work in our lab has shown that *Fah*$^{-/-}$ mice maintained on NTBC drinking water at a concentration of 8mg/L, do not suffer liver injury in *Fah*$^{-/-}$ mice, and thus the selection pressure for Fah corrected hepatocytes does not occur as long as NTBC supplementation is maintained (Grompe *et al.* 1995). Therefore, we sought to quantitatively measure the integration frequency of our vector in mice that were maintained on NTBC to prevent artificially high estimates of integration due to a selective growth advantage of Fah corrected hepatocytes in response to liver injury. The presence of episomal AAV vector within cells can also lead to falsely elevated integration frequencies if determined by histology alone rather than by quantitative PCR. Also episomal AAV vector in cells will lead to inaccurate integration frequency if determined by histology alone rather than by quantitative PCR. To quantitatively measure the integration frequency, we co-transplanted Fah -/- mice with with Rosa26 hepatocytes and AAV-rDNA-Fah (serotype 2 or 8) at a dose $3x10^{11}$, indicated in Figure 5. Mice were maintained on NTBC to prevent hepatocyte injury and to remove any selection pressure, so that proliferation and expansion of integration events in *Fah*$^{-/-}$ mice did not occur. Six weeks post-transplantation, the mice were sacrificed, and the AAV treated mouse livers were perfused and harvested. Hepatocytes



	hFah	beta-gal	MAAI
AAV8 S1	5.31	0.88	1.18
AAV8 S2	5.54	1.35	1
AAV2 S1	1.27	0.85	1.1
AAV2 S2	1.39	0.24	1.1

Figure 6. Absolute frequency of integration
The value of hFah/beta-gal represents integrated hepatocytes vs total tranplanted hepatocytes. Each value compares to the standard hFahVS beta-gal standard. MAAI serve as input control.

containing either episomal AAV and/or integrated AAV genome were collected and co-transplanted with Rosa 26 hepatocytes at a ratio of 100:1 into *Fah*$^{-/-}$ recipient mice. The transplant recipients were withdrawn from NTBC for 5 weeks to allow selection and proliferation of hepatocytes containing integrated hFah. NTBC withdrawal provides the selection pressure required to dilute and purge the liver of hepatocytes containing episomal AAV-genome. After 5 weeks of selective pressure, we hypothesized that only hepatocytes with AAV genome stably integrated into the host chromosome and passed onto daughter cells or Rosa 26 hepatocytes will be expanded to form Fah positive nodules in the *Fah*$^{-/-}$ liver. At that time, *Fah*$^{-/-}$ recipients were harvested, and the liver was analyzed both by quantitative PCR, PCR targeting the expression cassette, hFah and beta-gal, and by

immunostaining for Fah and beta-gal. Rosa 26 hepatocytes serve as an internal control for transplantation. As shown in figure 6, the absolute frequency of site-specific integration for AAV8 is about 3-6%, and for AAV2 is about 2-5%.

6. Site-specific integration in a Hemophilia B murine model

To rule out the possibility that site-specific integration at the rDNA locus was being driven by the hFah expression cassette, we tested the ability of our vector to integrate site specifically in a second disease model using a different expression cassette. We chose the hemophilia B mouse model. The hFah expression cassette with the RSV promoter in the AAV-rDNA-Fah vector was replaced by the hFIX expression cassette with a synthetic enhancer and human transthyretin (TTR) gene promoter. Wild type mice were injected with this vector and then screened for the presence of specific junction fragments, shown in figure 7. Junction fragments were specifically detected in AAV-rDNA-hFIX treated mice but not in placebo treated mice, nor in the mixture of pAAV-rDNA-hFIX and wildtype genomic DNA (negative control). It is unlikely that both hFIX and hFah expression cassettes with different promoters favorably integrate into the same locus. Instead, it is much more likely that site-specific integration is mediated by rDNA fragment and homologous recombination.

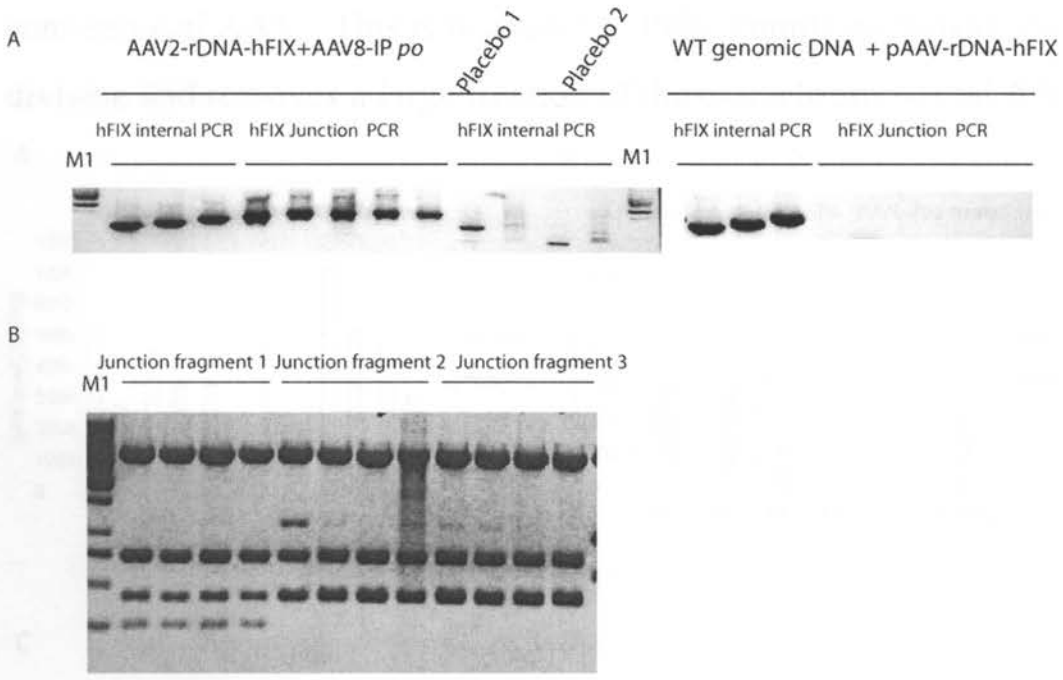


Figure 7. Site specific integration for hemophilia mice model.
(A) Junction fragments has benn detected in mice received AAV2-rDNA-Fah and AAV8-IP*po* but not the placebo treated mice and not in the mixture of pAAV-rDNA-hFIX and wild type geomic DNA. (B) EcoR I digestion of cloned junction fragment in pCR blunt II vector give the expected size band.

To confirm integration, we cloned the junction fragments into the pCR blunt II vector and digested these clones with EcoR I. The digestion gives the predicted two bands and vector backbone, as seen in figure 7B. Next to confirm stable integration and resultant expression of our target gene, AAV2-rDNA-hFIX or AAV8-rDNA-hFIX at a dose of $3x10^{11}$ (with or without AAV8-I-PpoI at a dose of $5x10^{10}$) was injected into wild type mice. One month following injection, the hFIX level was measured and 2/3 PHx was performed in these mice. The hFIX level dropped to about 5-10% of the pre-PHx level with the injection of

conventional AAV. This is because the PHx stimulates hepatocytes division and removes a large fraction of the extrachromosomal AAV

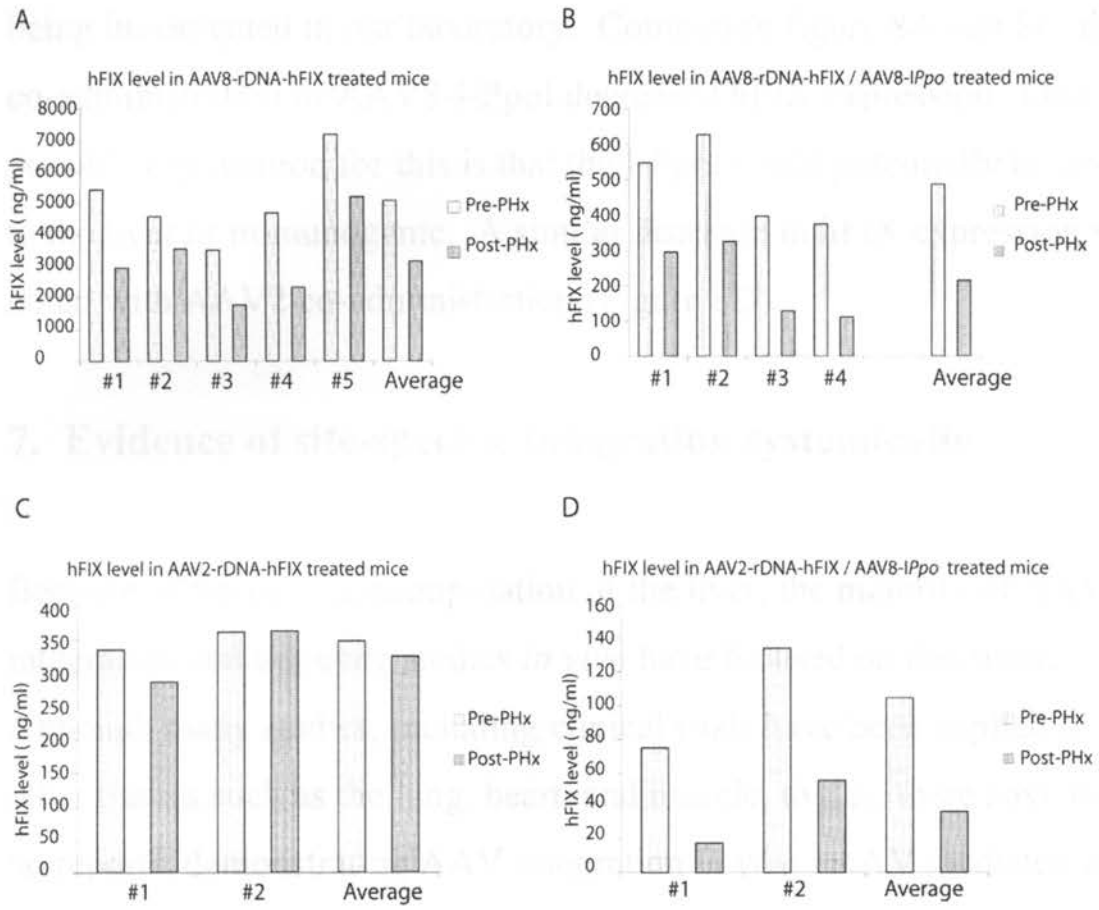


Figure 8. Majority of hFIX level persists after partial hepatectomy (PHx).

genome (Nakai *et al.* 2001). One month after PHx surgery, hFIX level was measured again. In contrast to conventional AAV, hFIX level in AAV8-rDNA-hFIX treated mice still remains about 40- 60% after PHx, which suggests that the majority of hFIX expression comes from integrated AAV-rDNA-hFIX, shown in Figure 8. With AAV2, about 70-90% of hFIX from AAV2-rDNA-hFIX treated mice persists after

PHx, see figure 9C. Note however that the absolute hFIX expression for AAV2-rDNA-hFIX is low for unknown reasons, and this is currently being investigated in our laboratory. Comparing figure 8A and 8B, the co-administration of AAV8-I-PpoI decreased hFIX expression. One possible explanation for this is that the I-PpoI could potentially be toxic to the liver or immunogenic. A similar decrease in hFIX expression was noted with AAV2 co-administration (Figure 8D).

7. Evidence of site-specific integration systemically

Because of the ease of manipulation of the liver, the majority of AAV integration and targeting studies *in vivo* have focused on this tissue. Although many studies, including clinical trials have been applied to other tissues such as the lung, heart, and muscle, to date there have been no reports, demonstrating AAV integration *in vivo*. AAV mediated gene targeting *in vivo* has not shown correction in any other tissues other than liver (Miller *et al.* 2006). Given the high integration efficiency of our AAV-rDNA vector, we tested for the presence of site-specific integration in tissues other than the liver. To test our hypothesis, AAV8-rDNA-Fah at a dose of $3x10^{11}$ (with or without AAV8-I PpoI at a dose of $5x10^{10}$) was injected into wild type mice. Liver, kidney, lung, heart, and muscle were harvested at day 3, day 7, and day 21. As described earlier, immuohistology cannot distinguish episomal AAV

mediated expression from either randomly integrated or site-specifically integrated AAV mediated expression. Therefore, junction fragment PCR was performed on this genomic DNA from AAV-rDNA-Fah treated tissue to confirm the presence of site-specific integration.

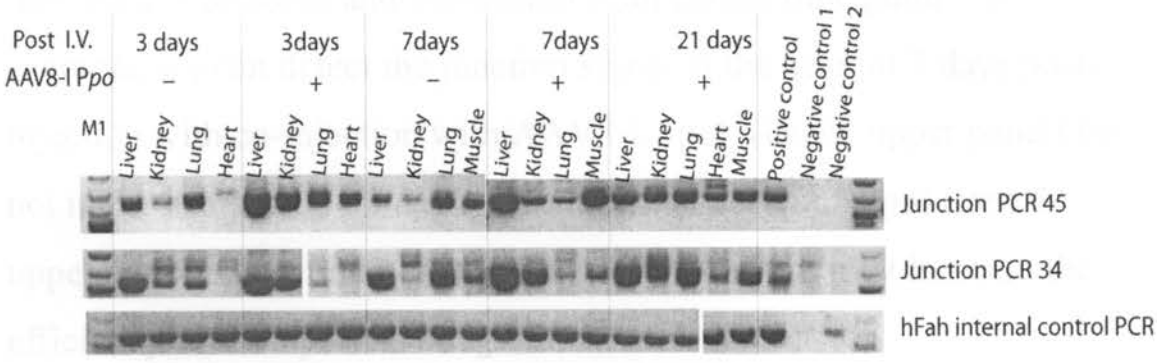


Figure 9. Site specific integration junction fragments has been detected systematicall. Wild type mice received AAV-rDNA-Fah at high dose ($3x10^{11}$)with or without AAV8-I P*po* at dose ($5x10^{10}$) have been sacrificed at the time point indicated. Liver, kidney, lung, muscle, heart has been taken for genomic DNA extraction and PCR analysis. Positive signal has been dected in virtually every tissues taken from each sample except heart from mice after 3 days without AAV8-I P*po*. Positive control is liver genomice DNA from $Fah^{-/-}$ serially trans-panted with hepatocytes from AAV8-rDNA-Fah treated mice. Negtive control1 is mixture of pAAV-rDNA-Fah and WT genomic DNA. Negative control 2 is pure pAAV-rDNA-Fah.

As shown in Figure 9, the site-specific junction fragment was detected in the liver, kidney, and lung with or without I-PpoI and detected in heart tissue only in the presence of I-PpoI as early as 3 days post-injection. In addition, site-specific integration was detected in muscle at 7 and 21 days post-injection, with or without I-PpoI. These results confirm that site-specific integration is occurring systemically. In addition, there is a tissue specific response based on serotype such that in some tissues, for

example the kidney, the AAV8 serotype occurs with low efficiency. If we compare the PCR signal in lane 2 vs lane 6, and lane 7 vs lane 11 in the upper two panels from liver with and without I-PpoI, we find that co-injection with AAV8-I-PpoI resulted in a stronger signal indicating increased expression and expression at an earlier time point. For example, we can detect the junction signal in the heart at 3 days post-injection with co-injection with AAV8-I-PpoI (lane9, upper panel) but not in the heart within 3 days post-injection without I-PpoI (lane 5, upper panel). Our findings here suggest that I-PpoI may increase the efficiency of site-specific integration.

Part IV Discussion

We have developed a novel strategy for gene delivery using an AAV gene therapy vector capable of site-specific integration at high efficiency *in vivo* with clinical efficacy and phenotype rescue in HT1 and hemophilia mice.

AAV is currently considered a safe gene therapy vector due to its non-pathogenicity and quiescent nature in the absence of a helper virus such as adenovirus (Mccarty *et al.* 2004). An important feature of AAV is its ability to integrate into the host genome and makes AAV an attractive gene therapy vector due to the ability of transduced cells to pass on integrated target genes to daughter cells during replication and proliferation. Stable transmission of the target gene to daughter cells would obviate the need for repeat dosing of the AAV vector in gene therapy trials. However, despite these advantages, AAV integration is currently very inefficient and with current vectors, the integration is random (Mccarty *et al.* 2004). Random integration carries with it the risk of potential tumorgenesis. In a neonatal mouse liver, random integration could potentially cause hepatocelluar carcinoma (Dansante *et al.* 2007). Another AAV gene therapy strategy, gene targeting, has been successfully applied in vivo to correct mutations in two genes, a *Lac*Z and *GusB* gene. However, the correction frequency is only about 10^{-4}-

10^{-5}, approximately ten to one hundred fold lower than the therapeutic level required for disease correction (Miller *et al.* 2006). With the increasing number of AAV clinical trials currently being approved, and as the population treated with AAV vectors increases, the risk of insertional mutagenesis remains a concern. A safer AAV gene therapy strategy without the risk of insertional mutagenesis from random insertion and a higher efficiency of gene targeting would be a significant improvement over currently available vectors. Our results suggest that our novel AAV-rDNA vector is a suitable candidate because of its ability to integrate into the host chromosome site-specifically at a high integration frequency. Although here we cannot completely exclude the possibility of a random integration event using this vector, all the mice including neonatal mice (n>30) injected with our AAV-rDNA vector have not gone on to develop any tumors. Theoretically, AAV-rDNA is much more likely to target to rDNA than any other locus within the host chromosome due to the presence of more than 200 copies of the homologous region that is actively transcribed. This makes the target locus an intrinsic hot spot for integration even without rDNA fragment in AAV vecor (Nakai *et al.* 2003). In addition, more than 40% percent of the AAV vector sequence is composed of rDNA sequence (2 kb ribosomal DNA sequence of the 4.3 kb AAV genome), which drives homologous recombination of AAV-rDNA with preferential attachment to the rDNA locus. And finally, the AAV vector or genome has been

reported to localize to nucleolus around tandem repeats of rDNA. This will provide the physical accessibility of the AAV genome to rDNA locus (Johnson *et al.* 2007). In addition, safety regarding the use of AAV vectors resurfaced within the public's consciousness this summer with the untimely death of a patient receiving treatment in an AAV clinical gene therapy trial (Wadman. 2007). In addition, there has been insertional tumorgenesis reported in AAV treated neonatal mice (Dansante *et al.* 2007). Our vector proposes a novel and safe way to correct target genes without the risk of random integration events at significantly higher integration efficiencies than seen previously.

Low integration efficiencies have been a limitation of AAV vectors. Our data on comparing the dose response for AAV-rDNA-Fah to AAV-Fah (Figure 4) show that AAV-rDNA-Fah has about 10-30 higher efficiency of integration than AAV-Fah. Lower doses are required to achieve rescue in our hereditary tyrosinaemia type I mouse model. This most likely is the result of the higher integration frequency of our novel AAV vector.

In our experiments using AAV-rDNA-hFIX, the hFIX measurement revealed more than 50% correction of hFIX levels after injection with AAV8-rDNA-hFIX and PHx and more than 70% correction of the hFIX level with injection with AAV2-rDNA-Fah followed by PHx. These

results are reasonable due to the following facts: Firstly, given that hFIX is a secreted protein, expressed hFIX protein will be released into the blood regardless of where it is synthesized. Also since AAV-rDNA vector can integrate systemically, hFIX could be expressed in other organ tissues and contribute to the systemic correction of the FIX level. Finally, injection with the current AAV2-hFIX vector followed by PHx results in 5-10% correction of hFIX levels. If our AAV-rDNA is 10-30 times more efficient at integrating into the host genome than regular AAV, this would account for a 6-7 folds correction in hFIX levels. In addition the rDNA locus is a transcriptionally active hot spot within the nucleus and thus an expression cassette that integrates into these regions may undergo increased levels of transcription. In any event, given the remarkable level of factor correction, our vector results may be of benefit in gene therapy clinical trials for hemophilia B. It is possible that a durable phenotypic correction could be obtained with a single administration of AAV-rDNA-hFIX.

In terms of the mechanism for AAV-rDNA mediated site-specific integration, it has been reported that double strand DNA breaks introduced by the I-SceI endonuclease can increase AAV mediated recombination efficiency (Miller *et al.* 2003). In addition, it has been reported that disruption of homologous recombination pathway but non-homologous end-joining pathway by RNAi has a significant effect on

the frequency of gene targeting (Vasileva *et al.* 2006). However, it is questionable whether homologous recombination is possible with an exogenous gene expression cassette up to 2 kbs (~100% of the homologous arm). The sequenced junction fragments prove the absence of any viral vector sequence (the ITRs) within the junction. This absence of vector sequence within the junction fragments strongly confirms that homologous recombination rather than non-homologous end-joining is responsible for AAV-rDNA mediated site-specific integration. This suggests that AAV is purely a delivery vector for the expression cassette, rDNA-hFah (hFIX), and actual viral sequence does not integrate into the host genome.

Taken together, we have designed a novel strategy for site-specific integration. We have tested our vector in two separate disease models, HT1 and hemophilia and document phenotypic correction. The rescue of function data for both HT1 and hemophilia suggest our AAV-rDNA vector is a safer, more efficient vector delivery system than current conventional AAV vectors. One drawback of our vector is the size limitation of the expression cassette, which is only 2.5 kb. However, given the increased safety profile, and efficiency of our vector, it is likely to be of benefit in gene therapy trials where the target gene is less than 2.5 kb.

Part V Conclusion

Our novel AAV-rDNA vector mediates site-specific integration in two separate diseases models (Hereditary tyrosinemia I and Hemophilia B) and results in therapeutic correction.

HT1 mouse model

The hepatocytes of weight-stabilized mice were serially transplanted into secondary *Fah*$^{-/-}$ recipients. All secondary recipients displayed weight gain after transplantation confirming the presence of a stably integrated Fah expression cassette. Next, site-specific junction PCR was used to prove the presence of hFah expression cassette within the rDNA locus and without the presence of vector sequence. Junction PCR was performed on all the mice including primary injected mice, transplanted mice, and neonatal mice injected with AAV-rDNA-Fah. Site-specific junction fragments were detected in all mice. The sequence of the junction fragment generated by site-specific PCR exactly aligns with the predicted junction sequence. Similar results were obtained in neonatal mice. A dose dependent comparison between our AAV2-rDNA-Fah and conventional AAV2-Fah was done. AAV2-rDNA-Fah rescued the *Fah*$^{-/-}$ mice at a significantly lower dose (10-30 fold lower) than the dose AAV2-Fah required to cause phenotypic correction. In addition, AAV8-rDNA-Fah was able to resuce *Fah*$^{-/-}$ mice at dose of only $1x10^9$.

Hemophilia B mouse model

AAV-rDNA-hFIX was made and injected into wild type mice. One month after injection, partial hepatectomy(PHx) was performed to drive proliferation in

response to liver injury. This would result in loss of episomal AAV genome. Surprisingly, hFIX levels were significantly higher using AAV8-rDNA-hFIX and AAV2-rDNA-hFIX (40-50% and 70-80% respectively), compared to only 5-10% of hFIX after PHx with conventional AAV-hFIX. These results are promising and our vector delivery system may be of potential benefit in human gene therapy trials for hemophilia B.

Systemic Integration

We detected evidence of site-specific integration systemically in the lung, heart, muscle and kidney in addition to the liver as early as 3 days post injection. This highly suggests that AAV8-rDNA is capable of integrating systemically *in vivo* and this is the first report of site-specific integration in an organ system other than liver. This finding may be relevant to gene therapy of systemic diseases. In addition, it is possible that with the appropriate tissue-specific promoter and the appropriate AAV serotype, AAV-rDNA can be integrate into many host tissues.

Overall

Therefore, our novel AAV-rDNA vector is clearly superior to existing AAV vectors with its capacity for site-specific integration with high efficiency at a low rescue dose. Thus, our vector provides a new strategy for target gene delivery for use in clinical gene therapy trials using a systemically integrating, site-specific vector with high efficiency of integration.

Part VI Reference

1 Berns KI, Giraud C. *Curr Top Microbiol Immunol.* 1996; 218: 1-23.

2 Berns KI. *Microbiol Rev.* 1990 Sep; 54(3): 316-29.

3 Berry MN, Friend DS. *J Cell Biol.* 1969 Dec; 43(3): 506-20.

4 Cahill D, Connor B, Carney JP. *Front Biosci.* 2006 May 1; 11: 1958-76.

5 Donsante A, Miller DG, Li Y, Vogler C, Brunt EM, Russell DW, Sands MS. *Science.* 2007 Jul 27; 317(5837): 477.

6 Flotte TR, Afione SA, Zeitlin PL. Am J Respir *Cell Mol Biol.* 1994 Nov; 11(5): 517-21.

7 Grompe M, al-Dhalimy M, Finegold M, Ou CN, Burlingame T, Kennaway NG, Soriano P. *Genes Dev.* 1993 Dec; 7(12A): 2298-307.

8 Grompe M, Lindstedt S, al-Dhalimy M, Kennaway NG, Papaconstantinou J, Torres-Ramos CA, Ou CN, Finegold M *Nat Genet.* 1995 Aug; 10(4): 453-60.

9 Higgins GM, Anderson RM *Exp Pathol* 1931, 12:186-202.

10 Johnson JS, Samulski RJ, 10th ASGT meeting, 2007.

10 Lin J, Vogt VM. *Mol Cell Biol.* 1998 Oct; 18(10): 5809-17.

11 Miller DG, Petek LM, Russell DW. *Mol Cell Biol.* 2003 May; 23(10): 3550-7.

12 Miller DG, Wang PR, Petek LM, Hirata RK, Sands MS, Russell DW.

Nat Biotechnol. 2006 Aug; 24(8): 1022-6.

13 Miller DG, Petek LM, Russell DW. *Nat Genet.* 2004 Jul; 36(7): 767-73.

14 Miller DG, Petek LM, Russell DW. *Mol Cell Biol.* 2003 May; 23(10): 3550-7.

15 Miller DG, Rutledge EA, Russell DW *Nat Genet.* 2002 Feb; 30(2): 147-8.

16 McCarty DM, Young SM Jr, Samulski RJ. *Annu Rev Genet.* 2004; 38: 819-45.

17 Nakai H, Montini E, Fuess S, Storm TA, Grompe M, Kay MA. *Nat Genet.* 2003 Jul; 34(3): 297-302.

18 Nakai H, Yant SR, Storm TA, Fuess S, Meuse L, Kay MA. *J Virol.* 2001 Aug; 75(15): 6969-76.

19 Overturf K, Al-Dhalimy M, Finegold M, Grompe M *Am J Pathol.* 1999 Dec; 155(6): 2135-43.

20 Overturf K, Al-Dhalimy M, Tanguay R, Brantly M, Ou CN, Finegold M, Grompe M. *Nat Genet.* 1996 Mar; 12(3): 266-73.

21 Sambrook J and Rusell DW, *Molecular Cloning: A Laboratory Manual* 2000, December.

22 Vasileva A, Linden RM, Jessberger R. *Nucleic Acids Res.* 2006 Jul 5; 34(11): 3345-60.

23 Wadman, M, *Nature* 449, 270 (20 September 2007)

Acknowledgement

I want to thank my advisor Dr. Markus Grompe for his training and patience. I want to thank my thesis advisory committee member, Dr. Bruce Magun, Dr. Mihail Iordanov, Dr. Cary Harding, Dr. David Koeller for their patience and advice for my thesis work. I want o thank the Grompe lab for supporting me. And I want to thank my parents in China and wife, Xi, for their support.