Toward a Recombinant Adeno-Associated Virus Origin of Replication

F. Curtis Hewitt

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Genetics and Molecular Biology

Chapel Hill
2009

Approved By:

Advisor: Richard J. Samulski

Reader: Christina Burch

Reader: Blossom Damania

Reader: Tal Kafri

Reader: Dale Ramsden

©2009
F. Curtis Hewitt
ALL RIGHTS RESERVED

Abstract
F. Curtis Hewitt
Toward a Recombinant Adeno-Associated Virus Origin of Replication
(Under the direction of Richard Jude Samulski)

Adeno-associated virus (AAV) is widespread throughout the human population. Many of the traits responsible for this ubiquity also make AAV an exceptional gene delivery vehicle. However, the widespread nature of AAV creates the risk that recombinant AAV vectors could be mobilized into unintended tissues or the general population by wild type (wt) AAV replication machinery. Alarmingly, the majority of characterized AAV serotypes are capable of rescuing and replicating the AAV2-based vectors currently used in AAV clinical trials.

There are a number of potential methods to prevent AAV vector mobilization. The origin of replication from a less prevalent human AAV serotype or a non-human serotype would decrease, but not eliminate, the risk of mobilization. Ultimately, AAV vectors must utilize an origin of replication incompatible with the Replication (Rep) proteins of any naturally occurring AAV serotype. Unfortunately, the nature of the AAV origin, the inverted terminal repeat (ITR), has been a barrier to understanding the mechanisms of replicative specificity necessary to synthesize a novel origin.

By generating a panel of chimeric and mutant ITRs and Rep proteins between two serotypes of AAV, we have mapped two independent DNA-protein interfaces required to generate replicative specificity. In vivo replication assays demonstrated that for AAV2, three residues in the Rep active site are required to make specific contacts with the nicking site of ITR2. AAV5 has a unique interaction between a 12 nucleotide extended Rep binding element

and a 49 amino acid region of Rep5 containing two DNA binding interfaces. Structural models display significant differences between serotypes in these regions. Understanding the separation of these elements led to the creation of a recombinant origin of replication with properties independent of either parent serotype. This novel origin stands to prevent AAV vector mobilization and expand our understanding of the basic mechanisms of AAV replication.

Acknowledgements

Without the support of family, the lab, and my committee, this work would not have been possible. I would like to thank Jude Samulski for pushing me to be a better scientist and for his optimistic outlook on both my positive and negative results. I would also like to thank my committee, Tal Kafri, Dale Ramsden, Blossom Damania, and Christina Burch who helped direct this work to a degree they may not recognize. Every member of the Samulski lab played a role in this project and never failed to assist me with any issue. I would like to thank Matt Hirsch in particular for our many scientific discussions, his critical reading, and the guidance he offered throughout my graduate career. I owe my work ethic, education, academic mindset, and so much more to my parents and the sacrifices they made. Finally, I thank my wife. She saw me through the highs and lows of graduate school, and our life and future together make all the hard work well worth it.

Table of Contents

List of tables....x

List of figures....xii

List of abbreviations.... xiiii

Chapter 1: Introduction....1

1A. AAV Virology....2

1B. AAV Replication....7

1C. Evolutionary Relationships of AAV Serotypes....10

1D. AAV as a Gene Therapy Vector....20

1E. AAV Vector Mobilization....23

Chapter 2: Creating a Cell Culture Assay for AAV Vector Mobilization....28

2A. Introduction....29

2B. Materials and Methods....30

B1. Plasmid Construction.....30

B2. Cell culture.....31

B3. Production of rAAV.....32

B4. Hirt DNA Purification and Southern Blot Analysis.....32

B5. *In Vitro* Transduction Assay.....33

B6. Quantitative PCR for Virus Titer.....33

B6. Dot Blot for Virus Titer.....34

B7. DNA Alignment and Generation of Phylogenetic Trees.....34

B8. Mobilization of ITR2 and ITR5 Vector Genomes by an AAV2 Provirus. 35

B9. Creation and Mobilization of AAV Vector Genomes from Stable GFP Expressing Cell Lines. 35

2C. Results 36

C1. The AAV5 Origin of Replication is Unique Among Human Serotypes. 36

C2. Creation and Characterization of Rep5 Helper Constructs for Cap1-5. 39

C3. ITR2 but not ITR5 Vectors can be Mobilized by a wt AAV2 Provirus. 43

C4. Persisting AAV Vector Genomes can be Mobilized. 46

2D. Discussion 49

Chapter 3: Mapping Protein/DNA Specificity of the AAV Origin of Replication 55

3A. Introduction 56

3B. Materials and Methods 59

B1. ITR Cloning. 59

B2. Rep Cloning. 59

B3. Western Blot Analysis 62

B4. Densitometry 62

B5. Molecular Modeling 63

B6. Plasmid Construction 63

3C. Results 63

C1. Characterization of Chimeric ITRs 63

C2. Nicking Stem Sequence and Height are Critical for ITR5 specificity 68

C3. Spacer Length is Critical for ITR2, not ITR5 72

C4. The ITR5 Spacer Acts as a RBE for Rep5. 75

C5. Identification of Regions of Rep Responsible for ITR Specificity. 79

C6. Characterization of Rep Regions Involved in ITR Specificity. 84

C7. Comparison of Rep Region 1 across Serotypes 89

C8. Structure-Function Model of Rep-ITR Specificity. 94

3D. Discussion 97

Chapter 4: Methods to Prevent Vector Mobilization 101

4A. Introduction 102

4B. Materials and Methods 103

B1. Plasmid Construction. 103

B2. Transduction Assay. 104

B3. qPCR Titering for Mobilization. 104

4C. Results 104

C1. Non-Homologous ITR Vectors. 104

C2. Snake AAV 108

C3. Novel Chimeric ITR as a rAAV Vector 112

C4. Prevention of AAV Vector Mobilization by a Novel Chimeric ITR. 115

4D. Discussion 117

Chapter 5: Future Directions 121

Appendix 1: Chimeric and Mutant Rep Proteins Created 126

Appendix 2: Chimeric and Mutant ITRs Created 127

References.....138

List of Tables

Table 1. Prevalence of Human AAV Serotypes 12

Table 2. Homology Between AAV Serotypes.....17

Table 3 Oligonucleotides Utilized for Chimeric Rep Cloning 61

List of Figures

Figure 1. Organization of the AAV genome. 3

Figure 2. Model of AAV genomic replication. 8

Figure 3. Phylogenetic comparison between AAV serotypes. 15

Figure 4. Comparison between ITR2 and ITR5. 19

Figure 5. AAV as a gene delivery vector. 21

Figure 6. Illustration of AAV Vector Mobilization. 25

Figure 7. The Rep5-ITR5 interaction is unique among the fully characterized AAV serotypes. 38

Figure 8. Creation and characterization of Rep5 helper constructs for Cap1-5. 41

Figure 9. TR2 but not TR5 vectors are mobilized by latent wt AAV2 and Ad coinfection. 45

Figure 10. AAV genomes conferring long-term transgene expression in cultured cells can be rescued, replicated, and packaged. 48

Figure 11. AAV Vector Mobilization Assay. 52

Figure 12. Diagram of ITR synthesis. 58

Figure 13. Cloning and Characterization of Chimeric ITRs. 67

Figure 14. Relation of Nicking Stem Height and Sequence to Rep-ITR Specificity 70

Figure 15. Effect of RBE-Nicking Stem Spacing on Rep-ITR Specificity. 74

Figure 16. The ITR5 Spacer Acts as a RBE for Rep5. 78

Figure 17. Cloning and Characterization of Chimeric Reps. 82

Figure 18. Characterization of Rep Regions Critical for ITR Specificity. 86

Figure 19. Alignment of Rep Region 1. 91

Figure 20. Chimeric Rep8 proteins are not compatible with ITR2 or ITR5 93

Figure 21. Model of Rep-ITR specificity. 96

Figure 22. Replication of non-homologous ITR vectors. 106

Figure 23. Snake AAV possesses a unique origin of replication with respect to characterized human/primate AAV serotypes. 111

Figure 24. Vector yields and transduction potential of a novel ITR. 114

Figure 25. Specificity of the novel recombinant origin of replication extends to vector mobilization. 116

List of Abbreviations

AAV	Adeno-Associated Virus
Ad	Adenovirus
bp	Base Pairs
Cap	Capsid Protein
CMV	Cytomegalovirus
CNS	Central Nervous System
d	Dimer
DMEM	Dulbecco's Modified Eagles Medium
ds	Double Strand
GFP	Green Fluorescent Protein
HEK	Human Embryonic Kidney
ITR	Inverted Terminal Repeat
kb	Kilobases
kDa	Kilodalton
m	Monomer
moi	Multiplicity of infection
nt	Nucleotides
orf	Open Reading Frame
pA	Poly Adenylation
PCR	Polymerase Chain Reaction
Pol	Polymerase
qPCR	Quantitative PCR

rAAV	Recombinant Adeno-Associated Virus
Rep	Replication Protein
RPE	Retinal Pigmented Epithelium
SAAV	Snake AAV
sc	Self Complimentary
sITR	Snake ITR
sRep	Snake Rep
ss	Single Strand
TK	Thimidine Kinase
vg	Vector Genome
wt	Wild Type

Chapter 1: Introduction

1A. AAV Virology

Adeno-associated virus (AAV) was discovered as a contaminating factor in adenoviral (Ad) preparations over 40 years ago (2). The potential of AAV for gene delivery has been the driving force behind research of the virus since then, due in large part to its lack of pathogenicity. While the requirement of a helper virus to replicate AAV once led to the label of "defective virus," today the elegant and complex nature of the virus is better understood. Indeed, AAV has succeeded evolutionarily to such a great degree that it is now ubiquitous in many vertebrate species including our own.

The 4.7kb single strand (ss) AAV genome is relatively simple (Figure 1). The genome is flanked by T-shaped inverted terminal repeats (ITRs) approximately 150 base pairs (bp) in length (4). Roughly half of the genome encodes four non-structural replication (Rep) genes from a single open reading frame (orf) via alternative splicing and an alternative transcriptional start site (58, 88). The p5 and p19 promoters drive transcription of the Rep gene products and harbor Rep binding sequences which allow Rep to auto-regulate its expression. The other half of the genome encodes three capsid (Cap) proteins from the same orf by virtue of alternative splicing. The Cap genes are transcribed from the p40 promoter which is trans-activated by the Rep protein (69).

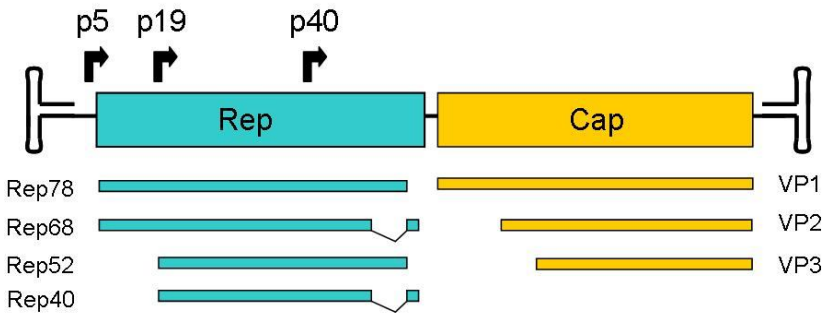


Figure 1. Organization of the AAV genome.
Two T-shaped ITRs flank the AAV genome. Four different Rep proteins (Rep78, Rep68, Rep52, and Rep40) are produced from the p5 and p19 promoters and alternative splicing. The p40 promoter drives expression of the three Cap proteins (VP1, VP2, and VP3) which are produced by alternative splicing.

The viral ITRs are the only cis-elements required for replication and encapsidation of the genome. Sequences within the ITR are bound site-specifically by Rep. The ATP-dependent Rep helicase unwinds the ITR structure allowing the formation of a secondary hairpin, or "nicking stem" which is nicked by Rep in order to catalyze AAV replication (9). The complex secondary structure and high G-C content of the ITR make it highly recombinogenic. These features are necessary for the latent phase of the virus as well as for efficient correction or repair of mutated ITRs upon entry into the lytic phase.

The AAV Rep proteins have multiple functions critical to the AAV replicative cycle. The N-terminus contains two separate DNA binding interfaces which bind specific sequences within the ITR. The Rep DNA binding domain also binds specifically to the three AAV promoter sequences as well as specific Ad promoters. This domain also contains the active site of the molecule. The active site interacts with the nicking stem of the ITR and contains the nucleophilic tyrosine molecule responsible for creating a site-specific nick (86). The C-terminus of Rep contains an oligomerization domain that participates in Rep multimerization on the ITR. The C-terminus also possesses residues which bind specifically to the viral capsid in an interaction required for encapsidation of the genome. Additionally, the C-terminus harbors an ATP-dependent SF3 DNA helicase required for both replication and encapsidation (20). It is believed that the Rep helicase "motor" bound to the pre-constructed capsid acts to thread the viral genome through a pore at the 5-fold axis of symmetry (43). The N-terminal DNA binding/nicking domain is contained exclusively by the large Rep proteins, Rep78 and Rep68 (Figure 1). These large Reps are necessary to catalyze replication due to their exclusive interaction with the ITR. Two small Rep proteins, Rep52 and Rep40,

transcribed from the p19 promoter, lack this domain. While these small Rep proteins are not necessary for replication, they are required for efficient encapsidation (43).

Sixty of the AAV Cap proteins come together to form a 22 nm diameter particle with T=1 icosahedral symmetry. Three different Cap proteins are produced: VP1, VP2, and VP3 (Figure 1). An assembled capsid consists of these proteins in a 5:5:50 ratio, respectively. Encapsidation of the genome is believed to occur through a pore in the center of one of the twelve pentamers which make up the capsid (8). This pore is also thought to allow the extrusion of the unique N-terminus of the VP1 and VP2 subunits. The N-terminus of VP1 contains a phospholipase domain required for transduction and is only externalized after the virion has been internalized into the endosome (26, 110). The N-terminus of VP1 and VP2 harbor nuclear localization sequences necessary to target the virion to the nucleus for delivery of the genome (87). VP3 is the major structural protein and contains the receptor binding motifs required for cell binding and entry. Changes in receptor binding are responsible for the wide range of tissue tropisms of different AAV serotypes. Heparan sulfate and N- or O-linked sialic acid are the major cell receptors for the majority of characterized AAV serotypes (99). Many of the specific regions and residues involved in receptor binding have been identified and can be altered via rational mutagenesis or directed evolution techniques to specifically target new cell types (54).

AAV replication requires co-infection by a helper virus or cellular stress (14, 50, 64). Ad typically serves as the wt helper virus for AAV, though Herpes Simplex Virus (HSV) can also complement AAV replication. AAV has evolved intricate mechanisms of viral latency due to its dependence on a helper virus for replication. The large Rep proteins auto-regulate both the p5 and p19 promoters (Figure 1) (47, 69). In the absence of helper, this auto-

regulation prevents Rep transcription and keeps the virus in a latent state. Upon superinfection by Ad, four early genes are required for AAV replication: E1A, E1B, E2A, and E4. E1A and E4 are required for transactivation of the AAV promoters. E1B and E2A are involved in aiding mRNA transport as well as stalling the cell cycle to allow uninterrupted viral production (11, 12, 13, 48, 51, 72, 78). Additionally, Ad virus-associated (VA) RNA is required to prevent the interferon host cell shutoff mechanism of translation. Ultimately, the presence of these helper genes (or cytotoxic stress) leads to the alteration of auto-regulatory Rep complexes, transforming them into trans-activators of all three AAV promoters (69). This leads to Rep-catalyzed replication and encapsidation of the genome.

Because not every cell infected with AAV will also be infected with helper virus, AAV is highly efficient at ensuring long-term persistence of its genome. Post infection, genomes typically circularize or form higher-order concatemers (18, 83, 101). In non-dividing tissues, these genomes have been shown to persist in an episomal state indefinitely (23). However, non-replicating episomes in dividing tissue are diluted and eventually lost in a manner similar to plasmid DNA. For that reason, long term persistence of AAV in dividing cells requires integration of the viral genome into the host chromosome. The recombinogenic ITRs contribute to a low level of random integration (75). Additionally, AAV is unique among animal viruses in the ability to direct its genome to integrate site-specifically into the host chromosome (14, 45, 80). The AAVS1 site on human chromosome 19 contains a Rep binding element and nicking site enabling Rep2 to direct viral genomes into that site (97). Due to the likelihood of losing all or part of the ITR during recombination, AAV can efficiently “gene correct” its ITRs, repairing a deleted or mutated ITR using the opposite ITR

as a template (79). This ability to self-repair the ITR is critical for the rescue of an AAV provirus entering the lytic phase.

1B. AAV Replication

The AAV ITRs are critical for nearly every aspect of the viral life-cycle. As they are the only *cis*-element required for replication and encapsidation, a recombinant AAV (rAAV) vector can be efficiently produced as long as the Rep and Cap genes are provided in trans (102). As the ss genome enters the nucleus, the secondary structure of the ITR provides a free 3' hydroxyl to act as a primer for synthesis of the second strand. This results in a ds genome capable of undergoing transcription (Figure 2A-C; 33). This replicative intermediate is known as a monomer (m). In the absence of Rep, the ITR with a free 3' hydroxyl can prime synthesis of the entire molecule again. This results in a dimmer (d) molecule (Figure 2I). This process can continue in the absence of Rep, allowing higher-order intermediates to form until Rep acts to resolve the ITRs.

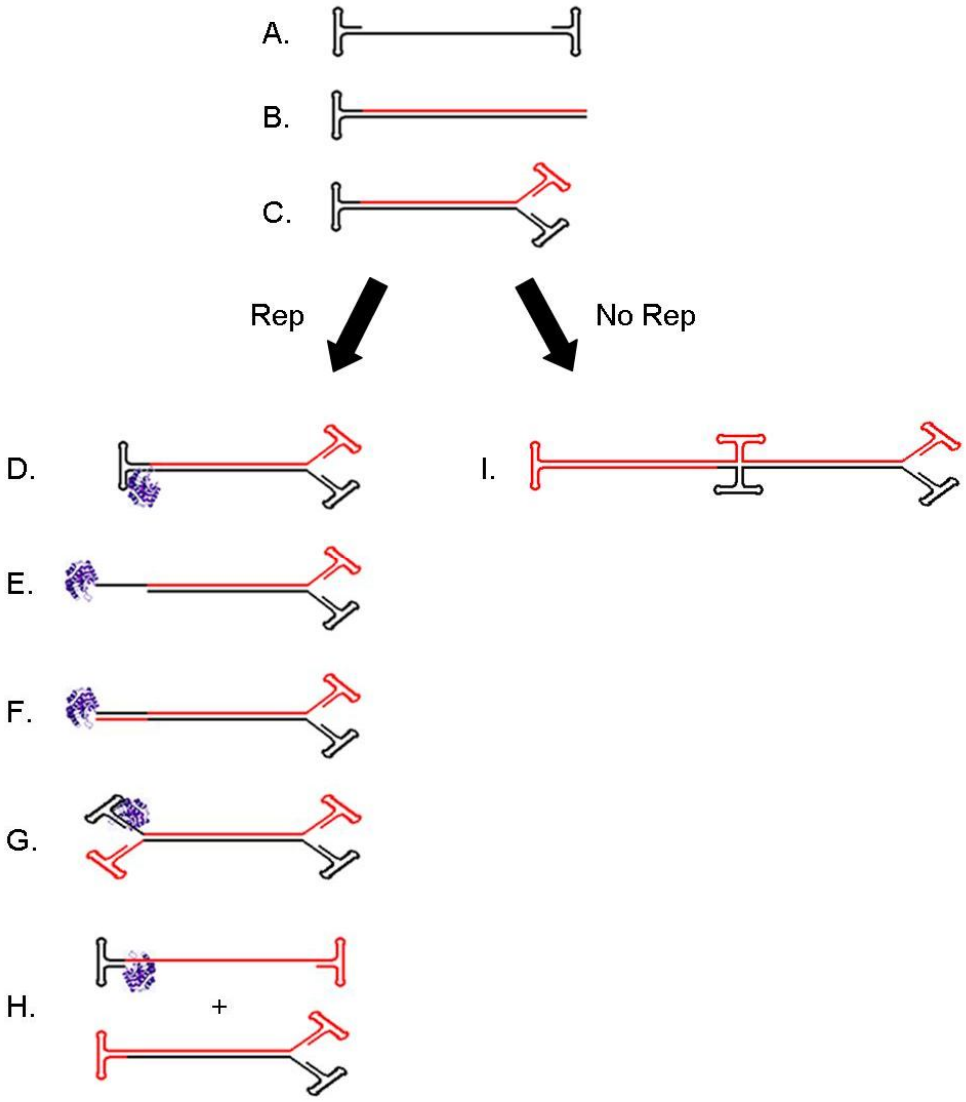


Figure 2. Model of AAV genomic replication.
The AAV genome enters the nucleus as a single strand molecule with the exception of the ITRs (A). The ITR acts to prime second strand synthesis (newly synthesized strand shown in red; B). This leads to duplication of the opposite ITR and transcription of viral genes (C). Rep binds (D) and nicks (E) the closed ITR allowing synthesis of the remainder of the genome (F). The fully replicated genome (G) can be displaced by the Rep helicase or by the subsequent synthesis of a new second strand (H). Alternately, in the absence of Rep, a dimeric genome can be synthesized using the 3' hydroxyl of the open ITR as a primer (I).

In the presence of Rep, the large Rep proteins (78 and 68 kDa) are responsible for binding site-specifically to the ITR on both the tetrad GAGY repeat known as the Rep binding element (RBE) and to the tip of one of the hairpin stems (RBE', Figure 2D; 10, 76). Rep molecules multimerize on the ITR, binding every four nucleotides along the RBE. Rep is thought to bind as two hexamers in a head-to-head orientation, although there is evidence that it may bind as two octamers or may only require dimerization for function (55, 60, 84, 85, 93). Upon ITR binding, the ATP-dependent SF3 helicase in the C-terminus of Rep unwinds the DNA, leading to the formation of an internal hairpin (Figure 4). This nicking stem contains the terminal resolution site (trs) and specific nicking site of the ITR (9). After nicking, Rep remains covalently attached to the 5' end of the ITR. The liberated 3' hydroxyl primes synthesis through the ITR to complete synthesis of the genome (Figure 2E-G; 70). Fully replicated genomes can undergo subsequent rounds of replication or can be encapsidated (Figure 2H). Unlike many autonomous parvoviruses, the (+) and (-) polarity strands are packaged interchangeably during AAV encapsidation (5).

Polymerase (Pol) δ has been identified as the host polymerase responsible for AAV second strand synthesis. As with normal Pol δ function, PCNA and RFC are also critical for second strand synthesis. Once the second strand has been synthesized, the genome typically circularizes or concatemerizes in a process which also utilizes host cell machinery. Cellular repair machinery recognizes the ITR as a double strand break repair intermediate. This results in the recruitment of Rec Q helicases, the MRN complex, as well as ATM and other DNA repair machinery which are necessary for circularization or contatemerization of the AAV genome (17).

As an alternative to second strand synthesis, because both (+) and (-) strands are encapsidated, it is possible that AAV genomes could anneal after infection. While wt AAV likely does not infect cells at high enough titer for genomes to anneal efficiently, this pathway might occur in assays utilizing AAV as a vector when as many as 100,000 vector genomes (vg)/cell are administered. Because synthesis of the second strand is the major rate-limiting step for wt AAV and rAAV infections, vector genomes have been designed to bypass this step. These vectors (termed self-complimentary) utilize one wt ITR and one ITR with a deletion preventing Rep nicking (62). Because of the mutant ITR, self-complimentary (sc) vectors package a ds genome with the closed, mutant ITR acting as a hinge to allow self-annealing and transcription of the genome without second strand synthesis. As a result, scAAV vectors have a quicker onset of transcription and subsequent expression of the transgene. Due to the packaging constraints of AAV, however, this requires sc vectors be no larger than 2.4 kb in length. Thereby, scAAV vectors can not be used for vector cassettes requiring the entire 4.7 kb capacity of AAV.

1C. Evolutionary Relationships of AAV Serotypes

AAV belongs to the family Parvoviridae. Historically this family has been split into three genera: Parvoviruses, Densoviruses, and Dependoviruses. Both Parvoviruses and Densoviruses are autonomous. Parvoviruses are typically vertebrate viruses while Densoviruses infect insects. AAV occupies the genus Dependovirus, a group of Parvoviruses dependent on a helper virus for their life cycle. Due to the general lack of pathogenicity displayed by AAV and other parvoviruses, discovery of new isolates and serotypes has taken place largely through PCR-mediated screening for related sequences.

The majority of AAV serotypes have been isolated from primate and human samples. Individual viral serotypes, by definition, can not be neutralized by sera which contains neutralizing antibodies to any other serotype. Because new AAV isolates are identified frequently and because of the challenge of cross-reacting potential new serotypes with antibodies against each of the others, the true number of AAV serotypes is in constant flux. Although AAV has been isolated from a wide range of species including snakes, birds, and swine, serotypes isolated from humans and primates have been explored in greater depth due to their potential for human gene therapy applications.

While humans and primates harbor many closely related AAV serotypes, four specific serotypes have been reported to make up the majority of AAV hosted by humans. By testing human sera for neutralizing antibodies to different AAV serotypes, AAV serotype 2 (AAV2) has been estimated to infect between 30-80% of the population (Table 1). AAV3 infects approximately the same percentage. AAV5 infects substantially fewer humans, with only 10-20% estimated to be seropositive. AAV6 is harbored at a slightly higher rate, around 30% (33, 40, 59, 91).

	Estimated % of Population Seropositive	Tissue Tropism	Rep Protein Replicates Vectors Flanked by:	
			ITR2	ITR5
AAV2	30-80%	Kidney	Yes	No
AAV3	30-80%	Cochlear Inner Hair Cells	Yes	No
AAV5	10-20%	CNS, RPE	No	Yes
AAV6	30%	Skeletal Muscle	Yes	No

Table 1. Prevalence of Human AAV Serotypes

AAV serotypes 1-4 and 6 were isolated as contaminants of Ad preparations. AAV5 was isolated from a human penile condylomatous wart (3). Most other AAV serotypes have been isolated from non-human tissue using PCR techniques to amplify DNA closely related to AAV genomic sequence. The driving force behind the search for new serotypes has been the promise of novel capsid variants which possess unique tropism and lead to improved rAAV vectors. Unfortunately, while the Rep and Cap genes can be easily isolated by PCR, the structure and G-C rich nature of the ITR has prevented the sequencing of these elements from most newly isolated serotypes. For this reason, the evolutionary relationships between the Rep and Cap portions of the known AAV serotypes are clearer than those of the ITR.

Most of these serotypes display markedly different tissue tropism in vivo. AAV2, largely due to its utilization of heparan sulfate as a receptor, is highly efficient in transduction of most tissue culture cells and is the major serotype used for cell culture studies. In vivo, AAV2 has demonstrated the ability to transduce kidney tissue (90). AAV3 has been used to transduce cochlear inner ear cells, though poor transduction efficiency of tissue culture cells has made determination of the cell receptor a greater challenge (56). AAV5 uses alpha-2,3-N-linked sialic acid as a receptor and has been utilized to transduce retinal cells (57). And AAV6, binding heparan sulfate, is able to efficiently transduce human skeletal muscle (7). Several non-human serotypes are also widely utilized for in vivo studies. AAV8 and AAV9 efficiently transduce liver tissue. AAV8 also transduces the heart and pancreas. And AAV1 is commonly used to achieve muscle transduction (100).

Comparing the amino acid sequence of Rep and Cap, as well as the nucleotide sequence of the ITRs between these serotypes by phylogenetic analysis demonstrates the highly conserved nature of these viruses (Figure 3). All trees were generated using AAV

serotypes 1-9, except the ITR map due to the lack of characterized ITR sequences for several serotypes. The autonomous parvovirus B19 was included as an outgroup. Also included were the bovine, goat, and snake AAV sequences.

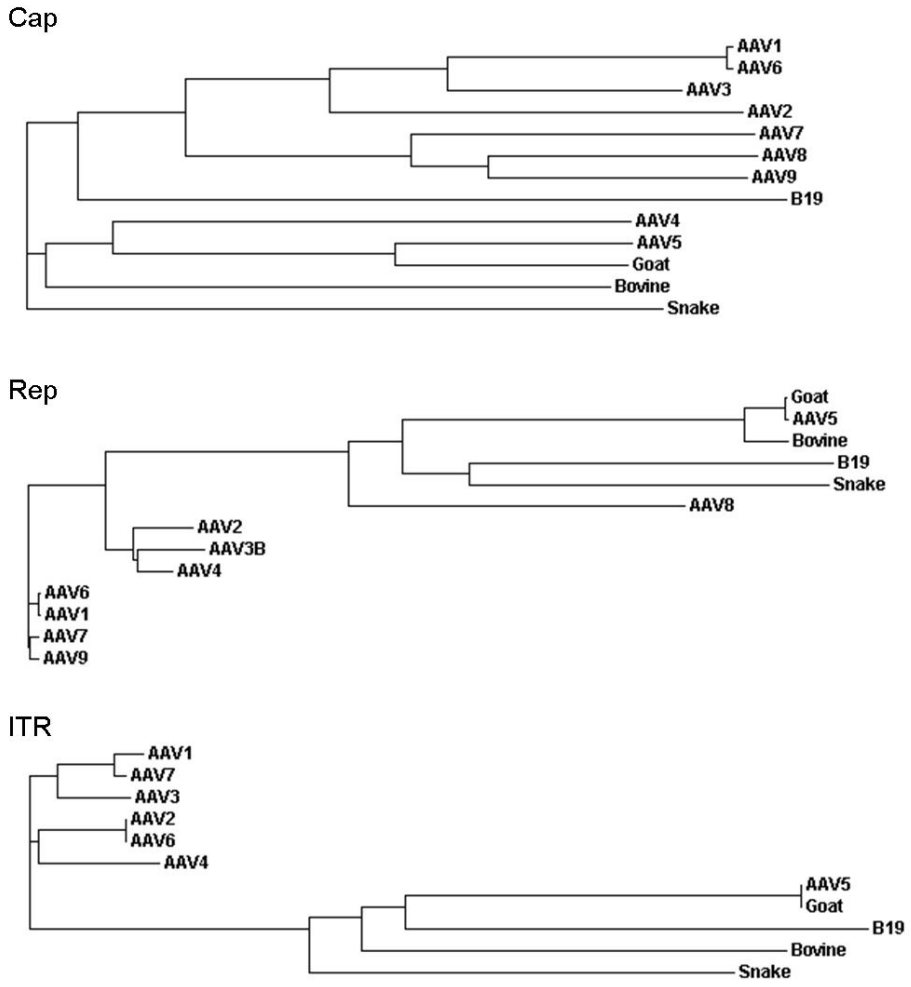


Figure 3. Phylogenetic comparison between AAV serotypes.
Distance from tree root represents level of divergence. Trees were generated using the amino acid sequence of Rep or Cap of the indicated serotypes and the nucleotide sequence of the ITR.

The AAV capsid proteins are all highly conserved. AAV4 is more closely related to AAV5 and the other non-human/primate AAV serotypes in its Cap sequence. However, Rep4 and ITR4 are not closely related to AAV5, suggesting a possible recombination event between an AAV2 and an AAV5-like serotype. Interestingly, receptor binding does not always correlate with serotype homology as AAV1 and AAV6 utilize different receptors despite extremely high homology.

The Rep proteins are also tightly conserved. Only two of the human/primate serotypes, AAV8 and AAV5, deviate in any significant manner. AAV5 Rep is almost 100% conserved with respect to goat AAV and is also highly related to bovine AAV. AAV5 is so poorly conserved with respect to the other human/primate serotypes that its Rep protein (Rep5) is unable to function on the ITRs of those serotypes (16). Due to their high degree of conservation, AAV 1-4, and 6 have demonstrated complete cross-compatibility between their Rep proteins and ITR sequences with respect to replication (30, 31). The high degree of conservation of AAV7 and 9-11 suggests this inter-compatibility should also extend to them.

Serotype	Cap Homology (aa) Compared to AAV2	Rep Homology (aa) Compared to AAV2	ITR Homology (nt) Compared to AAV2
AAV1	83%	87%	83%
AAV3b	87%	89%	85%
AAV4	60%	90%	83%
AAV5	58%	58%	49%
AAV6	83%	87%	100%
AAV7	82%	88%	83%
AAV8	83%	86%	
AAV9	84%	87%	
Goat	57%	58%	49%
Bovine	59%	58%	60%
Snake	52%	37%	63%
B19	22%	22%	11%

Table 2. Homology Between AAV Serotypes

Thus, the phylogenetic relationships of the origin of replication between fully characterized human/primate AAV serotypes breaks down into two groups: The AAV5 origin of replication and the remaining serotypes with AAV2-like origins of replication. An understanding of the factors which drive Rep-ITR specificity would then likely extend to all serotypes with the possible exception of AAV8.

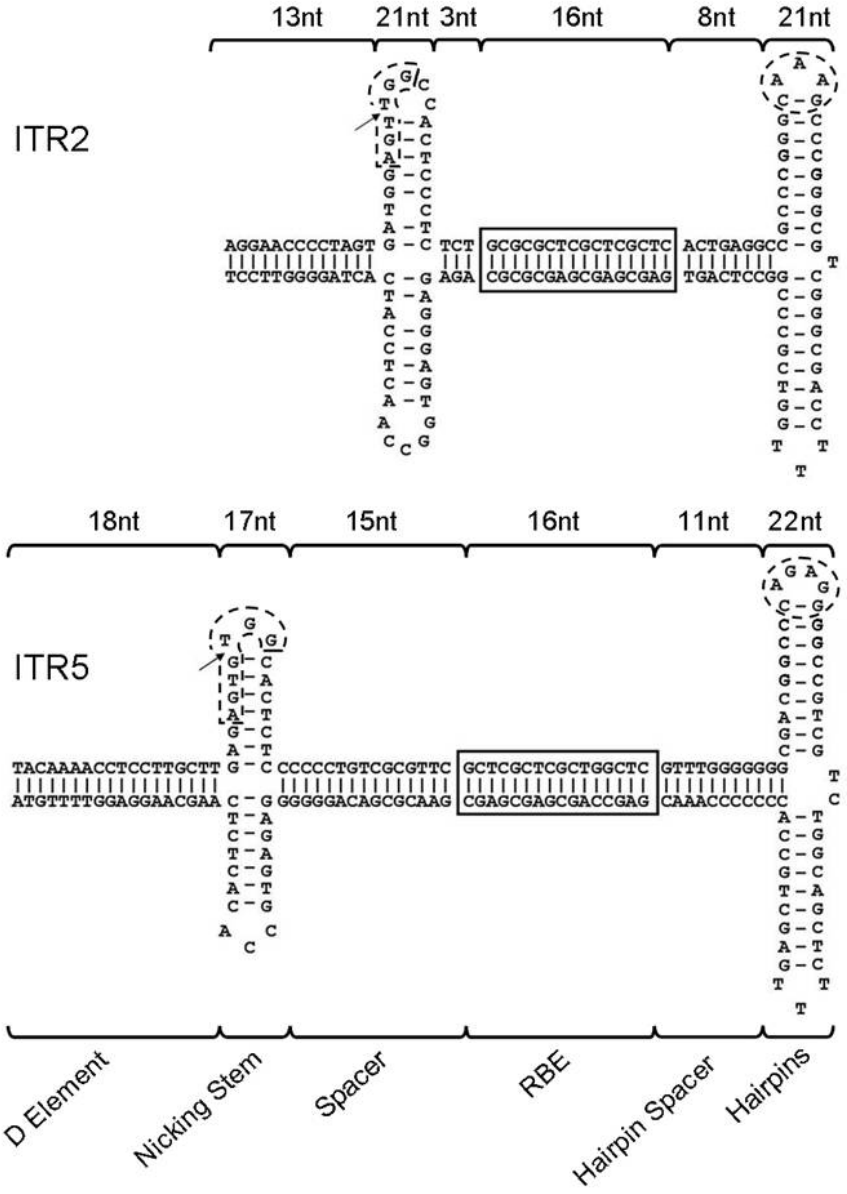


Figure 4. Comparison between ITR2 and ITR5.
Number of nt in each section is indicated. The Rep Binding Element (RBE) is boxed. The Rep Binding Element' (RBE') is circled. The terminal resolution site (trs) sequence is shown (hatched) and the nicking site is indicated by an arrow.

Replicative specificity between AAV2 and AAV5 does not exist at the level of binding, as Rep2 and Rep5 can bind interchangeably to ITR2 or ITR5 (16). Instead, specificity is created by the inability of Rep to cleave the ITR of the opposite serotype. This occurs despite high conservation between the ITR2 and ITR5 sequence, secondary structure, and location of elements required for Rep interaction (RBE, RBE', trs, nicking stem; Figure 4).

Identification of the elements involved in Rep-ITR specificity stands to increase the understanding of viral and cellular DNA binding and endonucleolytic proteins. It is likely that similar interactions take place in a wide range of viral and cellular replication and repair pathways. Localization of these elements may also facilitate the identification of other unique Parvovirus origins of replication. Here, we demonstrate two unique mechanisms at the DNA and protein level to achieve Rep-ITR specificity and utilize these factors to create a novel AAV origin of replication.

1D. AAV as a Gene Therapy Vector

Current clinical trials are using AAV to deliver genes to treat diseases such as cystic fibrosis, muscular dystrophy, hemophilia B, Parkinson's, Alzheimer's, and Canavan's. In the laboratory setting, AAV is also utilized as a gene delivery tool for cell culture and in vivo studies. While the danger of adding exogenous DNA to cells does apply to rAAV vectors, the benefits of this virus with respect to other viral gene delivery methods have powered the rapid expansion of rAAV in the field.

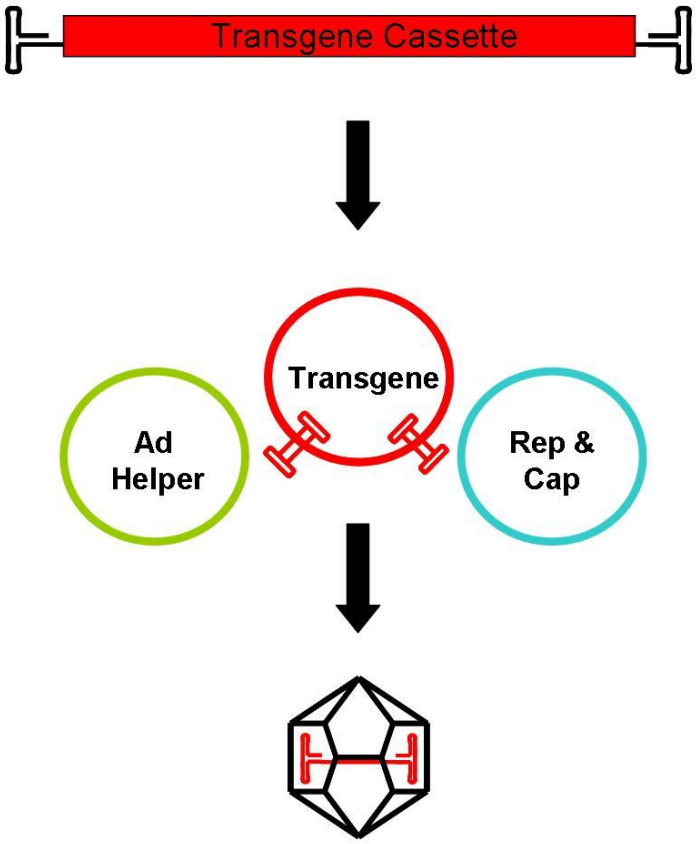


Figure 5. AAV as a gene delivery vector.
As the viral ITRs are the only cis-requirement for replication and encapsidation of the genome, the remainder of the viral genome can be replaced by exogenous sequence. A triple-transfection of the ITR-flanked construct along with the Adenovirus helper genes required for Rep and Cap expression leads to production of rAAV harboring the genetic payload.

The ability to utilize AAV as a gene delivery vector is due to the ITRs being the only cis-elements required for replication and encapsidation of the genome (Figure 5). In this manner, any DNA sequence can be assembled into a rAAV vector by placing it between two ITRs. To produce rAAV, a stable plasmid carrying the vector is triple-transfected into cells (typically HEK 293 cells). This requires an Ad helper plasmid (pXX680) which encodes the four Ad proteins required for AAV replication. It also requires an AAV helper plasmid encoding the viral Rep and Cap genes. A wide array of AAV helper plasmids have been created. These allow encapsidation into the Cap of any serotype and possess the Rep necessary to replicate the ITRs (36, 71). Upon transfection, the Rep proteins are expressed and act to rescue the genome from the exogenous plasmid backbone allowing canonical AAV replication to occur. Forty-eight to 72 hours after transfection, the cells are harvested, lysed, and the resulting vector is purified through a variety of methods (28). Viral titers are usually measured via Southern blotting or quantitative PCR (qPCR), and transduction efficiency can be determined if the viral payload contains a reporter construct.

There are a number of advantages to the use of AAV as a vector. AAV is largely non-immunogenic in relation to other viral gene delivery methods (111). The natural tissue tropisms of different serotypes allow for the targeting of specific tissues, overcoming the concern that a vector may transduce unintended organs. This can also allow for systemic administration yielding transduction of tissues in a predictable manner. In addition, pseudotyping or using different serotypes can allow re-administration of a vector to avoid immune response (46).

Another significant advantage is vector persistence. In non-dividing cells essentially all rAAV vector genomes circularize and persist in an episomal state, preventing the risk of

insertional mutagenesis (83). Vectors have been shown to persist for years in such an episomal state (101). Even though Rep2 has the ability to direct targeted integration of an ITR2-flanked genome, this is rarely a concern as Rep is not included in most rAAV vectors. These vectors can be delivered without concern for cytotoxic stress due to the general lack of pathogenicity of AAV.

However, there are drawbacks to the use of rAAV vectors. The size of the AAV genome is often too small to accommodate the transgenic cassette for most therapeutic purposes. The vector can only be expanded to about 5.2 kb in order to achieve adequate vector yield (27). This limitation has led to the creation of truncated gene products, truncated promoters, and split vectors which exploit the natural tendency of ITR-flanked genomes to concatemerize (96, 104). AAV also faces the greatest challenge to gene therapy; the risk that the addition of exogenous DNA can lead to insertional mutagenesis. While AAV persists almost exclusively in an episomal state, integration into the host chromosome does occur at a low rate, especially in dividing cells where double strand breaks are most likely to occur (66). Additionally, the ITR appears to drive an increase in ectopic chromosomal integration over plasmid-derived therapeutic methods. Finally, the ubiquity of wt AAV in the human population creates the possibility of innate immunity to certain serotypes. It also leads to the risk that co-infection of a cell with a rAAV vector with wt AAV would allow the rAAV vector to be replicated and mobilized into new cells or the population at large (1, 36).

1E. AAV Vector Mobilization

All current AAV vectors in clinical trials utilize ITR2s. This is due to the first rAAV vectors being developed from AAV2. However, using ITR2s for therapeutic purposes creates a safety risk due to the ubiquity of AAV2 in the human population as well as AAV3 and

AAV6 (whose Rep proteins can replicate ITR2s),. The Rep proteins of any of these serotypes (AAV1-4, 6) can replicate and encapsidate an ITR2 flanked vector as they cannot discriminate between the ITRs of their wt genome and those of the vector. In this manner, rAAV vectors have the potential to be “mobilized” out of the target tissue into different tissues of the body or into other individuals in the population (Figure 6; 36).

For nearly thirty years it has been known that a chromosomally integrated AAV genome could be rescued and replicated to allow AAV to enter the lytic phase of its life cycle (14). Ever since AAV was first cloned into a recombinant plasmid it has been known that the viral or vector genome could be rescued and replicated from exogenous DNA (77). Over ten years ago the first evidence for in vivo mobilization of an AAV vector transgene was presented, describing a potential scheme of vector mobilization between individuals (1). Yet, to date, only one paper has been published on preventing AAV vector mobilization (based on Chapter 2 of this work), and the ever increasing quantity of AAV vectors used for clinical trials have failed to adapt to address this concern.

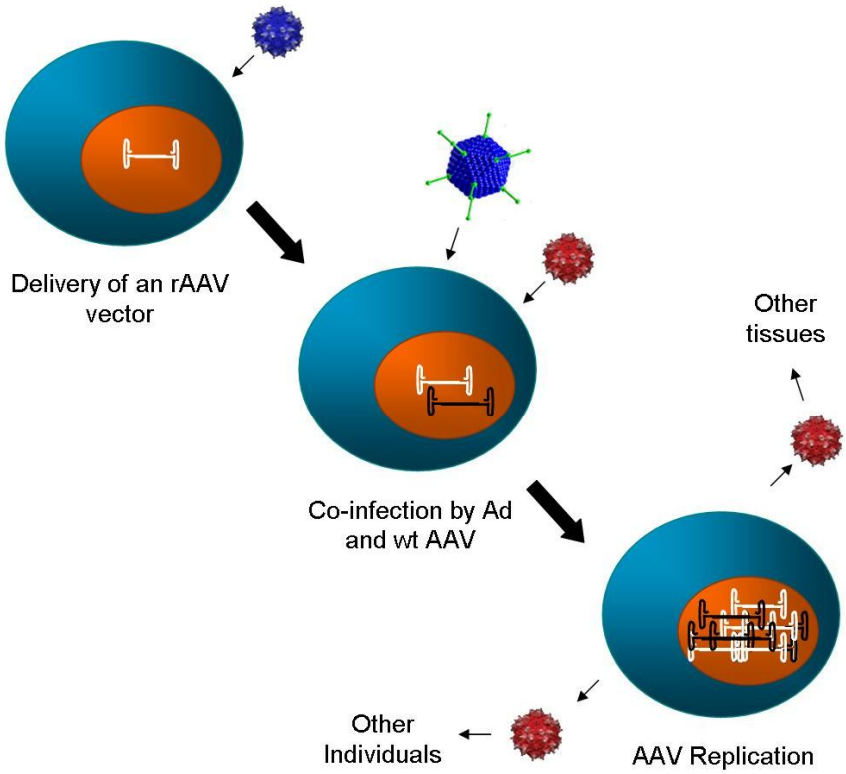


Figure 6. Illustration of AAV Vector Mobilization.
Administration of a rAAV vector to a cell (blue) results in delivery of the rAAV transgene (blue capsid and white vector) to the nucleus (orange). Co-infection by Ad and wt AAV (red capsid and black genome) allows AAV replication. Wt Rep proteins replicate all compatible ITR flanked vectors and encapsidate them into the wt capsid. These mobilized vectors could potentially infect new tissue within the body or other individuals.

There are four possible solutions to vector mobilization. The simplest method to decrease the risk of vector mobilization would be to use the ITRs of a serotype which is not as prevalent as AAV2. Rep2, Rep3, and Rep6 can all replicate the current ITR2 rAAV vectors (30, 31). Using ITR5 rAAV vectors would allow only wt AAV5 to mobilize the vector. Due to the significantly lower prevalence of AAV5 in the population, ITR5 vectors should be considerably safer than ITR2 vectors.

A second method builds upon the use of ITR5s. AAV vector mobilization requires a cell to be infected by a rAAV vector, wt AAV, and a helper virus such as Ad. By creating a vector carrying one ITR2 and one ITR5, a cell would have to be infected by the vector, two individual AAV serotypes as well as Ad for mobilization to occur. These non-homologous ITR vectors have already been used to direct concatemerization of split AAV vectors (104). However, such vectors would need to show no replication in the presence of a single Rep from either serotype, and would also need to be packaged efficiently such that useful titers for therapeutic purposes could be achieved.

While each of these methods would decrease the likelihood of AAV vector mobilization, they still use ITRs from AAV serotypes which infect the human population. In order to completely prevent any possibility of vector mobilization, an ITR from a non-human AAV serotype could be used, provided that it could not be replicated by the Rep proteins of any human serotype. Because AAV has been found in a large range of species, ITRs may exist which can not be replicated by Rep2 (or its compatible serotypes) or by Rep5. However, such divergent vectors may require host-specific co-factors in order to replicate, or a host-specific helper virus. The possibility also exists that AAV may jump from these species into our own, nullifying any advantage from using a rAAV vector flanked by these ITRs.

The ideal solution to vector mobilization would be the creation of a novel Rep-ITR interaction. Such a vector could not be mobilized by any of the wt AAV serotypes which infect humans, nor the non-human serotypes which could switch to human hosts. However, rational design or random mutagenic approaches to create a novel AAV origin of replication are difficult as both the ITR and Rep would need to co-evolve to become unique from their wt counterparts. A novel ITR would be useless without a corresponding Rep to replicate and encapsidate it. Additionally, not enough is known about Rep-ITR specificity to target any specific portion of the protein or ITR. Due to these limitations, the first step in creating a novel AAV origin of replication must be to determine why replicative specificity exists between currently characterized AAV serotypes.

Chapter 2: Creating a Cell Culture Assay for AAV Vector Mobilization

2A. Introduction

The potential risk of AAV vector mobilization has been thoroughly characterized. Integrated wt AAV or rAAV genomes can be rescued from exogenous plasmid DNA, chromosomal DNA, and episomally persisting forms by the expression of Rep and Cap (14, 77). Further, in vivo rAAV mobilization has been observed in primates by administration of wt AAV either before or after rAAV administration (1).

The potential danger of AAV vector mobilization should make the development of a non-mobilizable vector a priority for those looking to take AAV vectors into the clinic. While potential solutions are available, such as the use of non-homologous vectors, ITRs from non-human AAV serotypes, or the creation of entirely novel Rep-ITR interactions, the first step must be the creation of mobilization assays to test such novel vectors.

As in vivo methods would be inefficient for high-throughput screening of potential mobilization resistant vectors, the creation of cell culture assays to screen for mobilization is the best option. As two different mobilization scenarios exist, two mobilization assays must be created. The first is that of the presence of an AAV2 chromosomally integrated provirus existing prior to rAAV administration. This is due to the ability of AAV2 to integrate site-specifically into human chromosome 19 (80). This assay would entail the co-administration of a rAAV vector flanked by any ITR and Ad helper into a cell line harboring an AAV2 provirus. Assays measuring rAAV vector replication and mobilization by the wt AAV2 genome can then be performed. However, this assay will be limited due to the lack of proviral cell lines for other AAV serotypes.

A second assay must be created to account for the possibility of a co-infection of wt AAV and helper virus mobilizing a long-term persisting rAAV vector. In a cell culture

environment with dividing cells, the persisting rAAV vector would almost certainly be integrated into the host chromosome. A population of dividing cells harboring a persisting rAAV vector would approximate tissue treated with rAAV. Such cells could then be exposed to Rep and Cap from any serotype as well as helper virus in order to screen for mobilization of the rAAV vector.

Whether the Rep-ITR specificity displayed by AAV2 and AAV5 extend to mobilization is unknown. It is possible that integrated AAV genomes can act as mobile genetic elements and recombine to excise themselves. It is also possible that Rep does not need to be able to replicate the ITR in order for it to catalyze excision from exogenous DNA. However, in order for Rep to truly mobilize a vector, it must be able to replicate and encapsidate the vector. Because the AAV2 and AAV5 origins of replication are exclusive with respect to their cognate Rep protein, AAV2 should be incapable of mobilizing an ITR5 flanked vector, just as AAV5 should be unable to mobilize an ITR2 flanked vector (15). If true, the relative prevalence of these serotypes in the human population (up to 80% for AAV2, up to 20% for AAV5) should make ITR5 flanked vectors significantly safer (33, 40, 59, 91).

2B. Materials and Methods

B1. Plasmid Construction. The Rep5 fragment was amplified from the plasmid AAV5-2 (gift from R.M Kotin, hereafter referred to as pRep5Cap5) via polymerase chain reaction (PCR) using the forward primer rep-5 F1 5'-CGAGCTCGGCGCGTATGAGTTCTCGC-3' and the reverse primer rep-5 R1 5'-GACTACTCGCTTTATTTACTGTTC-3', which added an upstream *SacI* restriction site. The Cap2 fragment was amplified from the plasmid pXR2 (71) using forward primer 5'-

ATGGCTGCCGATGGTTATCTTC-3' and reverse primer 5'-TGGTGATGACTCTGTCGCCC-3', which included a *NotI* restriction site downstream from Cap2. The Rep5 and Cap2 fragments were ligated back into the pXR2 plasmid via the *SacI* and *NotI* restriction sites, in a triple-fragment ligation, to make pRep5Cap2. The pRep5Cap2 plasmid was used as a template for construction of the remaining plasmids. The pXR plasmid series contain Rep2 or chimeric replication genes with the cap genes from AAV serotypes that correspond to the plasmid number (71). The forward primers were pxr1/3 F (5'-ATGGCTCGGCATCCTTATCTTC-3') and pxr4F (5'- ATGGCTGCTGACGGTTACC-3'), and the reverse primers were pxr1/3 R (5'- CTATGACCATGATTACGCCAAGC-3') and pxr4 R (5'-CAGCTATGACCATGATTACGC-3'). The amplified region included a *NotI* restriction site downstream from the capsid genes. The Rep5 and Cap fragments were ligated into the pRep5Cap2 backbone via *PpuMI* and *NotI* restriction sites. Rep5Cap1-4 plasmids were confirmed by restriction site analysis and DNA sequencing. Vector plasmids included the transgene eGFP preceded by the CMV promoter and followed by the SV40 polyadenylation (p(A)) signal, flanked by AAV type 2 ITR (TR2-eGFP) or AAV type 5 ITR (TR5-eGFP) (Figure 7A). A neomycin cassette lies downstream driven by a TK promoter and the bovine growth hormone p(A) signal. Note that the ITR5-eGFP plasmid has an additional 500 bp insert upstream of the 3' ITR in order to distinguish it from ITR2-eGFP on a gel. The pRep6Cap6 plasmid was obtained from David Russell (75).

B2. Cell culture. HEK 293 and Cos1 cell lines were originally obtained from the American Type Culture Collection (ATCC) (Rockville, Md.). Detroit 5 (D5) and Detroit 6 (D6) human bone marrow cells were a gift from K. Berns. The D5 cell line was originally cultured as a clone of wt AAV2 latently-infected D6 cells (14) and contained wt AAV2 DNA

integrated into the chromosomal DNA. All cells were maintained at 37°C with 5% CO_2 saturation in media supplemented with 10% fetal bovine serum (Sigma) and 20 units/ml penicillin/streptomycin. HEK 293, Cos1, D5 and D6 cells were cultured in Dulbecco modified Eagle medium, while CHO pgsD-677 cells were grown in Ham F12 medium. CHO pgsD-677 cells were also obtained from the ATCC (89).

B3. Production of rAAV. Approximately 2x107 HEK 293 cells were triple-transfected with 10 µg AAV helper plasmid (encoding Rep2 or Rep5 and Cap2), 10 µg pXX680 (Ad helper plasmid), and 10 µg either ITR5-eGFP or ITR2-eGFP plasmid transgene vector. These constructs were mixed with 500µl DMSO and 100µl PEI prior to 5 min incubation at 25oC and dropwise addition to HEK 293 cells. At 48 hours post-transfection, cells were harvested for collection of cell lysate, Hirt DNA, and protein. Virus particles were collected from cell lysates produced from triple freezing and thawing of cells and centrifugation to remove cell debris.

B4. Hirt DNA Purification and Southern Blot Analysis. Hirt DNA purification was performed as described (41). Briefly, cells were harvested 48-72 hours post-transfection and washed in Phosphate Buffered Saline (PBS) and resuspended in 370ul Hirt Solution (0.01M Tris-HCl pH 7.5 and 0.1M EDTA) prior to addition of 25ul 10% SDS and 165ul 5M NaCl. Samples were then stored at 4°C overnight prior to centrifugation. Supernatant was purified by the addition of UltraPure Phenol:Chloroform:Isoamyl Alcohol (25:24:1) (Invitrogen) and DNA was precipitated and isolated by the addition of an equal volume of isopropanol. Samples were resuspended in 50ul sterile ddH_2O. 5ul of each sample was digested with 4U *DpnI* (NEB) 2-4 hours at 37°C prior to gel electrophoresis and Southern blot analysis as described (19). The nylon membrane (Hybond-XL) (GE Healthcare Life Sciences) was

hybridized to a probe corresponding to the GFP open reading frame labeled with the Random Primed DNA Labeling Kit (Roche) and d-CTP P^{32}. Blots were visualized after exposure to a phosphorimager screen (GE Healthcare Life Sciences).

B5. *In Vitro* Transduction Assay. Lysate from each transfection was tested for infectivity to ascertain virus production. 1 x 10^5 HEK 293 or Cos1 cells were seeded in each well of 48 well plates. Cos1 and CHO pgsD cells were used to assay AAV4 and AAV6 respectively since they transduce HEK 293 cells at a low efficiency. Original cell lysates harvested from transfections were diluted to a range from 1x 10^{-1} to 1 x 10^{-8}. 100 µl of each serial dilution in addition to Ad (m.o.i. of 5) were added to cells. Each infection was done in duplicate. At 24 hours post-infection GFP positive cells were visualized. Under 200x magnification, the number of GFP positive cells per field was counted. Ten fields per well were counted and the average number of GFP positive cells/field was determined. This number was multiplied by the number of fields per well and divided by the amount of lysate added to each well (as given by the dilution factor) to determine the number of transducing units (TU) per µl of cell lysate.

B6. Quantitative PCR for Virus Titer. Virus particles/µl was determined for each virus preparation. 10 µl of partially-purified cell lysate (see "Production of rAAV" in Methods) was added to 90 µl DNaseI Solution (10 µg DNaseI, 10 mM tris pH 7.5, 10 mM $MgCl_2$, 2 mM $CaCl_2$) and incubated for 1 hour at 37°C and stopped with the addition of 6 µl 0.5M EDTA. Then 120 µl of Proteinase K solution (1M NaCl, 1% Sarkosyl, 20 µg Proteinase K) was added and each sample was incubated for 2 hours at 55°C. Each sample was heated at 95°C for 10 minutes, diluted 1:500 or 1:5000, and used directly as template for qPCR.

For quantitiation of DNase-resistant viral genomes in the sample, 2 µl were used as template in a 10 µl real-time PCR reaction. The LightCycler FastStart DNA master SYBR Green I kit (Roche # 12239264001) was used following the manufacturer's instructions, using 3 mM $MgCl_2$ and 500 nM of each primer. All reactions were carried out on a Roche diagnostic real-time PCR LightCycler 2.0. For each reaction, the cycling parameters were as follows: 10 minutes at 95°; 5 cycles at 95° for 15 sec, 64° for 5 sec, and 72° for 15 sec; 5 cycles at 95° for 15 sec, 62° for 5 sec, and 72° for 15 sec; 40 cycles at 95° for 15 sec, 60° for 5 sec, and 72° for 15 sec. At the end of each run, a melting curve analysis was performed in which the PCR products were annealed at 72° and the temperature was gradually raised to 99°. In all cases, the PCR products melted in a narrow temperature range, indicating a pure PCR product without detectable non-specific amplification. The following primers were used: QGFP1-F 5'-AGCAGCACGACTTCTTCAACTCC-3' and QGFP1-R 5'-TGTAGTTGTACTCCAGCTTGTGCC-3'. A plasmid containing GFP, pTR2-GFP, was used to generate the standard curve.

B6. Dot Blot for Virus Titer. Dot blots were performed using 10µl per sample of purified cell lysate used in the various transduction assays, as described (21). The samples were blotted to a nylon membrane using a dot blot manifold and UV crosslinked at 60mJoules (UV Stratalinker 1800, Stratagene). The Southern blot was hybridized to a radiolabeled GFP probe (see "Hirt DNA Analysis" above) and exposed to film or visualized using a Storm Phosphoimager (Molecular Dynamics, Sunnyvale, CA, USA).

B7. DNA Alignment and Generation of Phylogenetic Trees. ClustalW2 was used for all DNA alignments and to generate phylogenetic trees (49). The amino terminal 200 amino acids of the large Rep proteins from AAV serotypes 1-12 were used to generate a

phylogenetic tree by the PHYLIP method. This program also calculated the percent homology between the Rep proteins of different serotypes.

B8. Mobilization of ITR2 and ITR5 Vector Genomes by an AAV2 Provirus. D6 and D5 cell lines were infected under four different conditions. A 10 cm plate of each cell line approximately 75% confluent was incubated overnight at 37oC in 5% CO2 saturation with 50µl AAV2 ITR2-eGFP or AAV2 ITR5-eGFP cell lysate. Cells were then washed four times with 10 ml DMEM and co-infected with Ad (m.o.i. of 10). Control plates were infected with only AAV2 ITR2-eGFP or only Ad. At 96 hours post-infection cell lysate and Hirt DNA from each infection were harvested as described above. Hirt DNA was separated by electrophoresis on a 0.8% agarose gel and transferred to a nylon membrane. The blot was probed for GFP as described above. The membrane was stripped by boiling with 1% SDS. The membrane was re-hybridized with a probe specific for Rep2 and Cap2. Additionally, cell lysates from each infection and Ad (m.o.i. 5) were used to infect 1 x 105 fresh HEK 293 cells. Cells positive for GFP were visualized 24 hours post-infection.

B9. Creation and Mobilization of AAV Vector Genomes from Stable GFP Expressing Cell Lines. 1 x 105 HEK 293 cells were infected with 10,000 vector genomes/cell of ITR2- or ITR5-eGFP rAAV and were then cultured for 18 days. Cells (approximately 1% GFP positive) were then trypsinized and subjected to flow sorting with a Beckman-Coulter Dako Moflo to isolate GFP positive cells. 1500 GFP positive cells were pooled into one well of a 96-well plate and allowed to propagate. GFP expression was monitored to insure a homogeneous GFP positive population in which the transgene was maintained. AAV helper plasmids were transfected into these cells in 10-cm dishes and incubated for 48 hours prior to Hirt DNA extraction and preparation of crude lysate.

2C. Results

C1. The AAV5 Origin of Replication is Unique Among Human Serotypes. As previously reported, Rep1-4 and 6 are unable to catalyze replication of ITR5 flanked genomes, while Rep5 is unable to catalyze the replication with vectors flanked by ITR1-4, and 6 (30, 31). Due to high sequence homology, it is likely that AAV serotypes 7-12 are also compatible with ITR2s and not ITR5s (Figure 3 and Table 1). In order to demonstrate the specificity of the Rep5-ITR5 interaction, two GFP vectors were utilized (Figure 7A). These constructs, ITR2-eGFP and ITR5-eGFP, are flanked by either ITR2s or ITR5s and expressing GFP from a CMV promoter. The rAAV vectors were transfected into 293 cells along with Ad helper plasmid (pXX680), and either Rep2Cap2, Rep5Cap2, or Rep6Cap6. Rep6 was included to confirm the cross-compatibility of AAV2 and AAV6 replication, as well as to underscore the ability of other naturally occurring serotypes to replicate ITR2 vectors. After 48 hours, cells were harvested for Hirt DNA and crude lysate. Hirt DNA (41) was analyzed by Southern blot with a probe for the GFP orf. A *DpnI* digestion was performed to remove transfected methylated plasmid DNA, but not unmethylated genomes which had been replicated in the cell. The results validated Rep-ITR specificity for the serotypes used, with Rep2 and Rep6 driving replication of only ITR2 vectors, and Rep5 driving replication of only the ITR5 vector. (Figure 7B). Crude lysate from these cells was used to transduce HEK 293 cells (highly transducible by Cap2) and CHO pgsD cells (transducible by Cap6) (Figure 7C). Specificity in the production of rAAV vectors followed the same pattern as replication, with Rep2 and Rep6 each able to produce infectious ITR2 rAAV particles and Rep5 able to produce only infectious ITR5 rAAV particles.

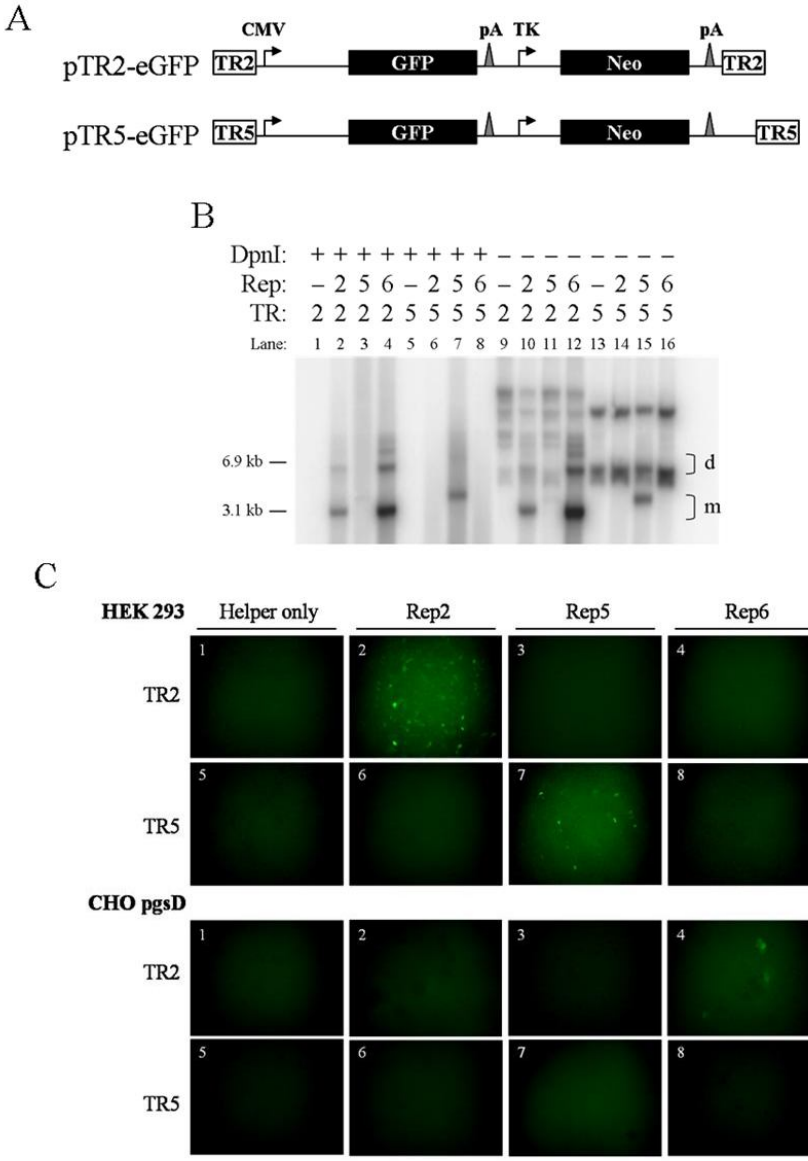

A
CMV
pA
TK
pA
pTR2-eGFP
TR2
GFP
Neo
TR2
pTR5-eGFP
TR5
GFP
Neo
TR5
B
DpnI: + + + + + + + + – – – – – – – –
Rep: – 2 5 6 – 2 5 6 – 2 5 6 – 2 5 6
TR: 2 2 2 2 5 5 5 5 2 2 2 2 5 5 5 5
Lane: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
6.9 kb —
3.1 kb —
d
m
C
HEK 293
Helper only
Rep2
Rep5
Rep6
TR2
TR5
CHO pgsD
TR2
TR5
1
2
3
4
5
6
7
8

Figure 7. The Rep5-ITR5 interaction is unique among the fully characterized AAV serotypes.
(A) The vector constructs used in this study. GFP expression was driven by a CMV promoter and SV40 poly(A) element. The Neomycin cassette included the thymidine kinase promoter and the bovine growth hormone poly(A) element. TR2s or TR5s flanked the vectors. pTR5-eGFP contained an additional 500bp ahead of the 3' TR.
(B) Southern blot of Hirt DNA comparing the ability of Rep2Cap2, Rep5Cap2 and Rep6Cap6 to replicate TR2s or TR5 flanked vector genomes. Hirt DNA was isolated 48 hours after transfection. DpnI cuts only the input plasmid, not the newly replicated AAV genomes. The two major replicative forms of AAV are indicated (m-double stranded monomer, d-double stranded dimer). Higher-order replicative forms are also visible.
(C) Transduction of HEK 293 cells (transducible by Cap2) or CHO pgsD cells (transducible by Cap6 or Cap2) with crude lysate from cells harvested 48 hours after transfection of Ad-helper plasmid only or triple-transfection of Ad-helper plasmid, TR2 or TR5 GFP, and either Rep2Cap2, Rep5Cap2, or Rep6Cap6. The numbers shown correspond to the lane of the gel in figure 2B.

C2. Creation and Characterization of Rep5 Helper Constructs for Cap1-5. As the prevalence of AAV5 in the human population is lower than AAV2 or AAV6, and considering the unique specificity between Rep5 and ITR5, we decided to create a rAAV ITR5 production system for transcapsidation into Cap1-5 similar to previous described system for type 2 (71). In order to confirm the efficacy of ITR5 vectors with respect to existing ITR2 vectors, Rep5 helper constructs were created (Figure 8A). This new system for producing virus vectors utilizes triple transfection with AAV helper plasmids containing the AAV5 Rep gene and one of the AAV serotype 1-5 Cap genes (pRep5Cap1-5), a reporter transgene plasmid with GFP flanked by AAV5 ITRs, and an Ad helper plasmid (XX680). HEK 293 cells were transfected with pRep5Cap1-5, Ad helper plasmid, and ITR2 or ITR5 eGFP. Analyses of Hirt DNA extracted from these cells showed that Rep5 functioned properly, generating the expected *DpnI*-resistant AAV monomer and dimer replication intermediates when delivered with the ITR5-eGFP but not the ITR2-eGFP vector (Figure 8B). Additionally, cell lysate harvested from each transfection was tested for infectivity to ascertain the system's ability to produce functional recombinant virus. HEK 293 or Cos1 cells were exposed to lysate from the ITR5 transfections and assayed for GFP expression at 24 hours post-infection. Lysate carrying capsid-specific sequences (types 1-5) all produced GFP-positive cells when ITR5 was complemented with Rep5 expression plasmids during vector production (Figure 8C, panels 1-5). Cells exposed to lysate from ITR2 transfections in the presence of Rep5 proteins were negative for transgene expression (Figure 3C, panel 6).

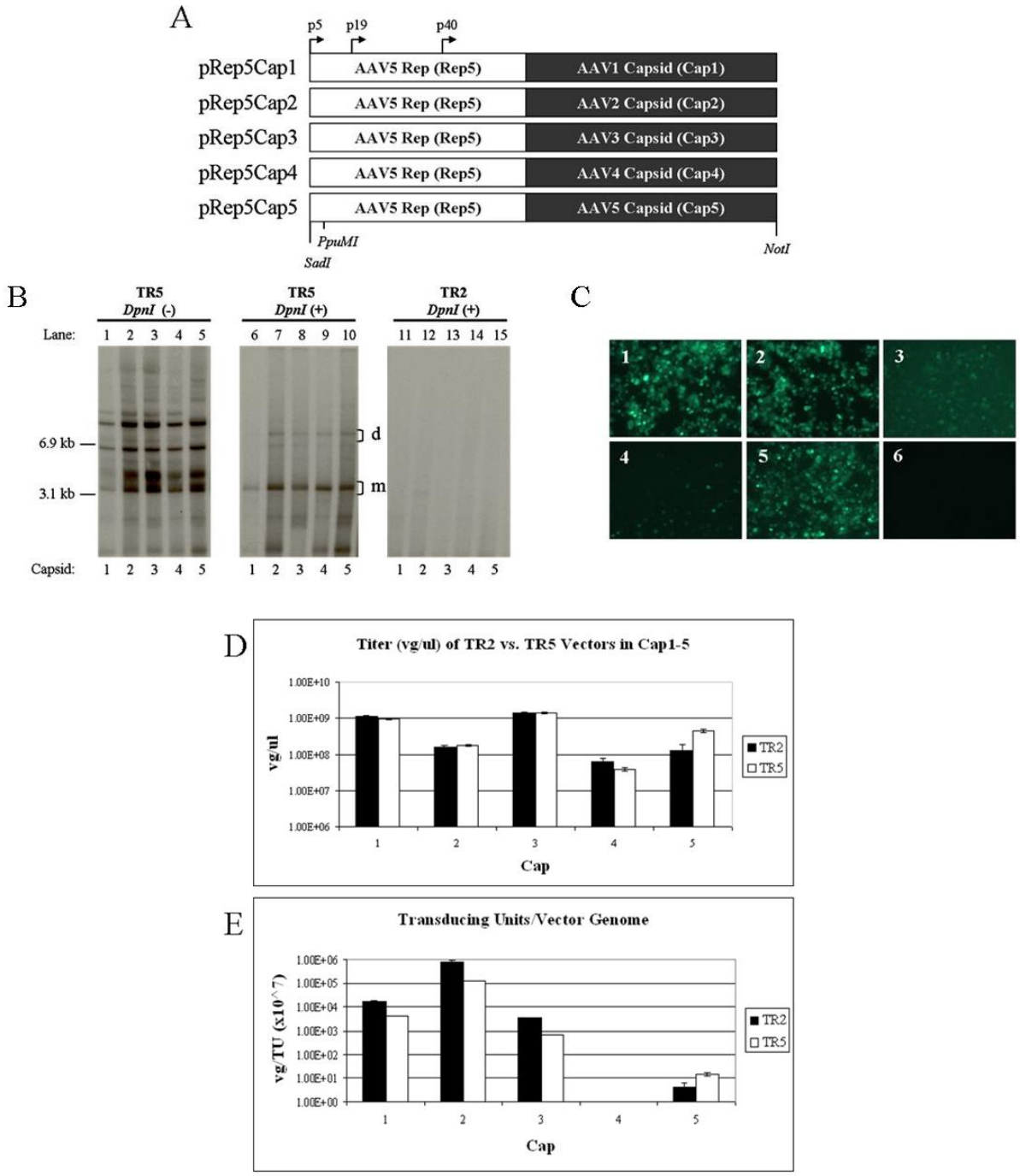

A
p5
p19
p40
pRep5Cap1
AAV5 Rep (Rep5)
AAV1 Capsid (Cap1)
pRep5Cap2
AAV5 Rep (Rep5)
AAV2 Capsid (Cap2)
pRep5Cap3
AAV5 Rep (Rep5)
AAV3 Capsid (Cap3)
pRep5Cap4
AAV5 Rep (Rep5)
AAV4 Capsid (Cap4)
pRep5Cap5
AAV5 Rep (Rep5)
AAV5 Capsid (Cap5)
PpuMI
SadI
NotI
B
TR5
DpnI (-)
TR5
DpnI (+)
TR2
DpnI (+)
Lane: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
6.9 kb
3.1 kb
d
m
Capsid: 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
C
1
2
3
4
5
6
D
Titer (vg/ul) of TR2 vs. TR5 Vectors in Cap1-5
vg/ul
1.00E+10
1.00E+09
1.00E+08
1.00E+07
1.00E+06
1 2 3 4 5
Cap
TR2
TR5
E
Transducing Units/Vector Genome
vg/TU (x10^7)
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
1 2 3 4 5
Cap
TR2
TR5

Figure 8. Creation and characterization of Rep5 helper constructs for Cap1-5.
(A) The AAV5 Rep and AAV1-4 Cap genes were subcloned into pXR2, a non-ITR-containing plasmid (see Methods for details). These new helper plasmids were used to package TR5 vectors into Cap1-5.
(B) Southern blot using a GFP specific probe of Hirt DNA extracted from HEK 293 cells transfected with pRep5Cap1-5, Ad helper plasmid, and TR5-GFP or TR2-GFP. The two replicative forms of the vector are indicated.
(C) Cell lysate from triple transfections described above were used to infect naïve HEK 293 cells. Cells infected with lysate from TR5-eGFP vector transfection of capsid serotypes 1-5 were positive for GFP (panels 1-5 corresponding to AAV1-5). HEK 293 cells infected with lysate from TR2-eGFP vector transfection of capsid serotypes 1-5 did not express GFP (panel 6, AAV1; representative of AAV2-5).
(D) Graph comparing the relative titers achieved in the production of TR2 vs. TR5 rAAV. Samples were titered in duplicate by Q-PCR. Standard error is indicated.
(E) Graph comparing the transducing units per vector genome of rAAV TR2 vs. TR5 vectors. Note that values on the y-axis are multiplied by 1x10-7. Virus was serially diluted and used to infect cells. GFP positive cells were quantitated and transducing units per microliter was calculated before conversion to transducing units per vector genome. Samples were measured in duplicate and standard error is indicated.

Having developed a functional AAV capsid production system using ITR5s, we sought to compare it to current Rep2-ITR2 production yields. The protocol designed by our lab (71) for production of transcapsidated rAAV2 (i.e., using helper plasmids that contain AAV Rep2 and serotypes 1-5 capsid genes, a reporter transgene (GFP) in an ITR2 vector cassette, and Ad helper plasmid) was used as a comparison to evaluate the new Rep5-ITR5 system. Cell lysate was harvested from each transfection and assayed for virus production by Q-PCR titering to determine virus particle number per unit volume of lysate (Figure 8D). The measurements obtained from the dot blots are comparable for ITR2 and ITR5 vectors when packaged in AAV serotype 1-4 capsids. Interestingly, ITR2 vector titers were noticeably reduced when packaged in an AAV5 capsid, possibly due to the evolutionary divergence of AAV5 with respect to the other serotypes (Figure 3 and Table 1). Both ITR5 and ITR2 vector production systems were also assayed for infectivity as measured by the number of transducing viral units per vector genome (Figure 8E). 1 x 10^5 HEK 293 cells were exposed to serial dilutions of cell lysate from each transfection, and the number of resulting transgene-positive cells was used to calculate the transducing units per microliter (TU/µl) of cell lysate. This was then divided by the viral titer (vg/ul) to yield the transducing units per vector genome (TU/vg). ITR5 vectors displayed a minor drop in transduction efficiency compared to ITR2 vectors, ranging from 5-10 fold in Cap1-3, while ITR5 vectors performed better in Cap5. This data demonstrates the transduction potential of each individual vector genome, again suggesting that due to evolutionary divergence, ITR5 vectors may have slightly better transduction potential when encapsidated in an AAV5 capsid. Transduction of the HEK 293 cells by Cap4 was below the detection threshold of the assay. The results indicate that while the yields of ITR5 rAAV vector production is equivalent to

the widely utilized ITR2 system, there may be a capsid specific effect on the ability of these vectors to transduce cells efficiently.

C3. ITR2 but not ITR5 Vectors can be Mobilized by a wt AAV2 Provirus. AAV2 has been consistently demonstrated to be the most prevalent natural AAV serotype in the human population (33, 40, 59, 91); thus there is a strong likelihood that a large percentage of human individuals harbor a latent AAV2 infection (65). For that reason, we obtained Detroit 5 (D5) cells demonstrated to contain a latent wt AAV2 infection in order to model the potential for ITR2 or ITR5 vector mobilization upon rAAV and Ad infection. The D5 cell line contains the wt AAV2 genome which is stably integrated at chromosome 19 and is rescuable upon infection by helper virus (14, 80). The parental line, Detroit 6 (D6), is negative for wt AAV and was used as a control. Cell lysate containing ITR2 or ITR5-flanked rAAV GFP genomes encapsidated into Cap2 were harvested from triple-plasmid transfections of HEK 293 cells. Each type of lysate was used to infect D5 and D6 cells. After 24 hour incubation, cells were washed and co-infected with Ad helper virus. Control plates were exposed to lysate containing either AAV2 ITR2-eGFP only or Ad only. Transgene GFP expression was observed in all cells receiving the original cell lysate, confirming the infectivity of AAV2 ITR2-eGFP and AAV2 ITR5-eGFP in D5 and D6 cells (data not shown). Hirt DNA analysis from infected cells revealed rescue of latent AAV2 genes, in the form of AAV2 replication intermediates, in D5 cells infected with Ad (Figure 9A). As expected, D6 cells without latent AAV or D5 cells without Ad did not show AAV2 replication intermediates (Figure 9A). In addition, rescued latent wt AAV2 genomes were able to complement rAAV vector genomes when assayed by Southern blot analysis. For example, replication intermediates were observed in D5 cells exposed to the ITR2 vector

while no vector replication was observed in the cells exposed to the ITR5 vector. (Figure 9B). Longer exposure revealed minor ITR5 vector signal in D5 cells, comparable to background levels of ITR5 and ITR2 vector signal found in D6 Hirt DNA, indicating a lack of replication in the presence of Ad.

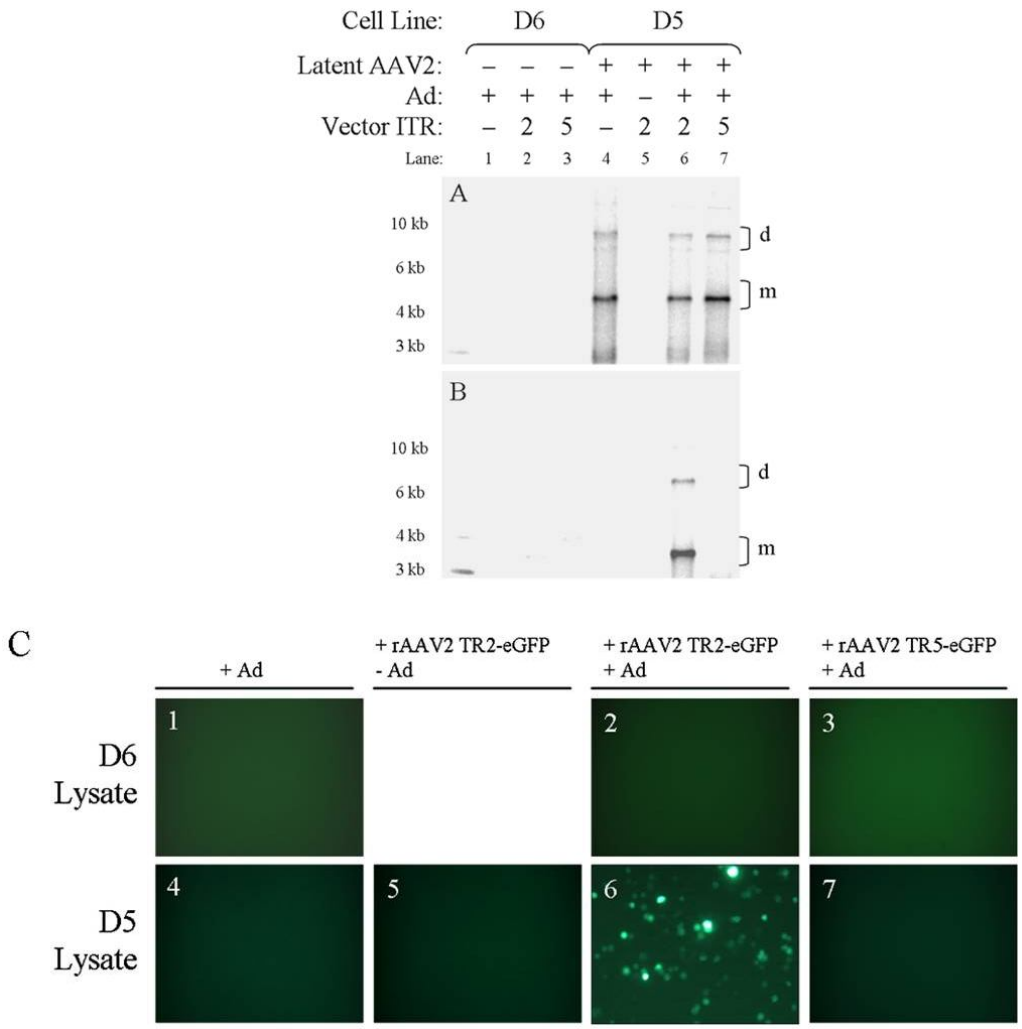


Figure 9. TR2 but not TR5 vectors are mobilized by latent wt AAV2 and Ad coinfection. Detroit 6 (D6, Latent AAV2 (-)) or Detroit 5 (D5, latent AAV2 (+)) cell lines were infected with rAAV2 from crude lysate containing TR2 or TR5-eGFP in the presence or absence of adenovirus. Hirt DNA was isolated and analyzed by Southern blot 48 hours post-infection using either an AAV2 (A) or GFP (B) probe. The two major replicated forms of AAV DNA are indicated. Larger replicative forms are also visible. Size marker is denoted. (C) GFP expression was visualized in HEK 293 cells after crude lysate was added from either D5 or D6 cells infected with the vectors shown above the panels. The numbers refer to the gel lane in figure 5B.

Lysate taken from ITR2 or ITR5 eGFP vector-infected D5 and D6 cells was used to infect naïve HEK 293 cells (Figure 9C). HEK 293 cells given lysate taken from D5 cells exposed to the AAV2/TR2-eGFP vector and Ad were positive for GFP (Figure 9C, panel 6), demonstrating that latent wt AAV2 was able to provide Rep and Cap in *trans* to mobilize the ITR2-flanked GFP vector. In contrast, HEK 293 cells given lysate from D5 cells infected with the ITR5-eGFP vector did not express GFP (Figure 9C, panel 7), demonstrating that the latent AAV2 was not able to mobilize the ITR5-flanked GFP vector. As expected, control lysate from D6 cells did not produce infectious GFP vectors using either ITR, and infection of D5 cells without Ad helper or without ITR2-eGFP vector did not produce infectious GFP vectors (Figure 9C, panels 1-5).

C4. Persisting AAV Vector Genomes can be Mobilized. During infection, AAV genomes not degraded have two fates: episomal formation or chromosomal integration (17, 63). While wt AAV2 has been shown in tissue culture cells to integrate into the human chromosome in a site specific fashion (14, 80), it has been demonstrated that Rep is required for this form of latency (4). Ideally, rAAV vectors should be delivered in the absence of Rep, wherein numerous studies have determined that rAAV genomes remain episomal, typically circularizing or forming into concatemers as a mechanism of vector persistence. (18, 83, 102) Thus, the infrequent event of integration by rAAV genomes is not site specific (63). Regardless of the method of molecular persistence, AAV genomes are able to excise themselves from the chromosome or episome upon Ad superinfection to enter the lytic phase of the AAV lifecycle, suggesting wt AAV persists in a conservative manner, keeping at least one ITR sequence intact. (80)

To test whether persisting rAAV vectors could be rescued after infection in the absence of site-specific integration, we infected HEK 293 cells with 10,000 vector genomes/cell of either ITR2 or ITR5 GFP virus as determined by dot blot. After 18 days, GFP cells were sorted and pooled to approximate a population of cells infected by the vector. While we did not confirm these vectors had integrated into the host chromosome, we did confirm that after sorting, GFP persistence remained in 100% of the cells for greater than two months. Rep2Cap2, Rep5Cap2, or Rep6Cap6 as well as Ad helper plasmid were transfected into the mock, ITR2, or ITR5 containing cell lines and both Hirt DNA and crude lysate were isolated. Figure 10A reveals both ITR2 and ITR5 genomes were capable of being rescued and replicated (lanes 6, 8, 11). Specificity remained consistent for these vectors, with both Rep2 and Rep6 able to rescue and replicate ITR2s and only Rep5 able to replicate ITR5s. The pTR5-eGFP panel was exposed longer than the ITR2 or mock due to the small amount of replicated vector DNA isolated from these cells.

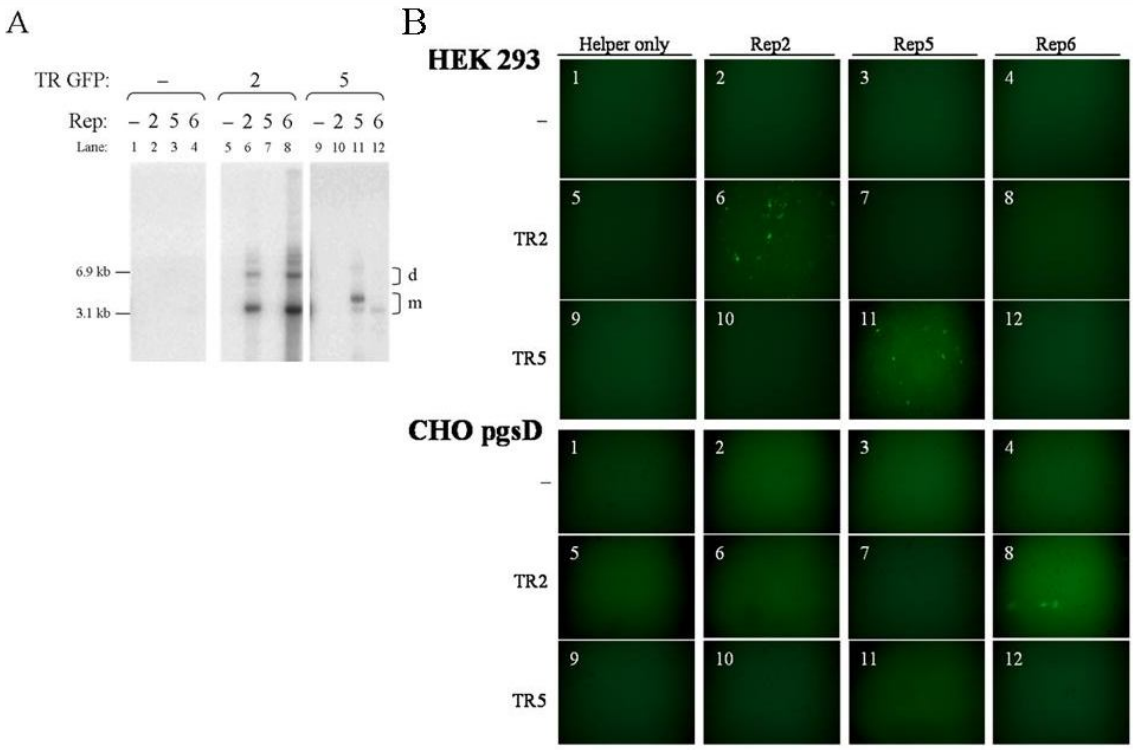


Figure 10. AAV genomes conferring long-term transgene expression in cultured cells can be rescued, replicated, and packaged.
(A) 293 cells were infected with either ITR2 or ITR5 eGFP vectors and passaged 18 days before cells still expressing GFP were sorted and pooled. AAV helper plasmids (Rep2Cap2, Rep5Cap2, Rep6Cap6) were then added to assay for the ability of persistent AAV genomes to be rescued and undergo replication in the presence or absence of Ad helper plasmid. Hirt DNA was isolated and assayed via Southern blot with a probe for the GFP ORF. The two major replicative forms of the vector genomes are indicated. Larger replicative forms are also visible. The TR5 (+) helper plasmid panel of the blot was subjected to longer exposure in order to visualize the replicating vector genomes.
(B) Mobilized genomes were assayed for infectivity by transducing HEK 293 or CHO pgsD cells with crude lysate from the cells described in figure 10A (control, TR2, or TR5 persisting vector genomes transfected with the helper plasmids described) 48 hours after addition of helper plasmids (All transfections for lysate used in figure 10B included Ad helper plasmid).

To determine if these rescued, replicating genomes could be encapsidated and mobilized to naïve cells, lysate was added to HEK 293 or CHO pgsD cells. Figure 10B shows that rescued genomes were encapsidated and that persisting rAAV genomes can be mobilized into previously non-transduced cells (HEK 293 panels 6 and 11 and CHO pgsD panel 8). Predictably, the transduction profile of the mobilized ITR2 or ITR5 vector genomes was dependent on the capsid into which they were packaged (Cap2 or Cap6), highlighting the potential danger of ITR2 vector mobilization being driven by a range of wt AAV serotypes.

2D. Discussion

This aim suggests that AAV5 based vectors are significantly less likely to be mobilized after administration than the AAV2 based vectors currently used in clinical trials. The two most prevalent human AAV serotypes (AAV2 and AAV6) both have the ability to replicate the ITR2 flanked vectors currently used in AAV clinical trials (30). A less widespread AAV serotype, AAV5, has a unique Rep-ITR interaction making it the only human serotype able to replicate ITR5 flanked vectors (16, 30, 52). This replicative specificity, as well as the relative abundance of these serotypes in the population, (AAV5 over four-fold less abundant that AAV2 (33, 40, 59, 91) led us to hypothesize that ITR5 flanked vector genomes have a significantly reduced risk of vector mobilization.

To test this hypothesis, we created an ITR5 based vector production system similar to the ITR2 based system currently used to produce rAAV (Figure 8A). This system worked well, exclusively packaging ITR5 flanked vectors into Cap1-5 while yielding comparable viral titers and transduction efficiency with respect to the current ITR2 vector production system. While our ITR5 vectors may have shown a minor inherent decrease in transduction efficiency (potentially in a capsid specific manner) optimization of our system may eliminate

this disparity. A similar comparison between ITR2 and ITR5 vectors *in vivo* showed no such bias, (30) suggesting the differential may be due to the sensitivity of our *in vitro* system or otherwise restricted to our assay.

In order to confirm that Rep-ITR replicative specificity extended to vector mobilization we adopted two cell culture assays. While the transformed cells used for these assays had the potential to behave differently from the primary cell types AAV vectors would encounter *in vivo*, we reasoned that Rep-ITR specificity would remain consistent regardless of cell type. That said, demonstrating vector mobilization in primary cells remains an important step in establishing the potential danger to future gene therapy candidates.

First, we showed that in cells with latent wt AAV2 infection, introduction of an ITR2 vector and subsequent superinfection by Ad resulted in replication of the wt genome and the ITR2-flanked transgene, (Figure 9A and 9B) and led to the production of infectious rAAV particles (Figure 9C). These results demonstrated that latent wt AAV2 plus Ad reconstituted the replication-deficient ITR2 vector system, allowing for mobilization of transgene vectors. Once again, AAV2 was unable to replicate or mobilize an ITR5 flanked genome (Figure 9B, lane 7), underlining the potential of our ITR5 based system to decrease AAV vector mobilization due to relative AAV2, AAV5, and AAV6 prevalence in the population. While inclusion of a cell line harboring a latent AAV5 genome would have been ideal for this study, there are no reports of AAV5 integrating site-specifically into the human chromosome. Thus, any cell line harboring an AA5 genome should be recapitulated by our mobilization system in figure 4 where a persisting ITR5 flanked genome is rescued and replicated.

Next, cell lines were first created containing stably persisting ITR2 or ITR5 flanked vector genomes. While we did not determine whether these genomes were integrated into the

host chromosome, the persistence of GFP signal in 100% of these cells two months after sorting suggests chromosomal integration. However, the possibility that they are persisting in some other manner only lends credence to the mobilization assay we have developed, as such genomes may recapitulate any number of modes of persistence *in vivo*. We next demonstrated that these persisting rAAV genomes could be rescued and replicated upon the transfection of AAV helper plasmids (Figure 10A). These genomes were also encapsidated and able to transduce naïve cells (Figure 10B). Predictably, the cell/tissue tropism of these mobilized genomes was dependent on the capsid into which they were mobilized.

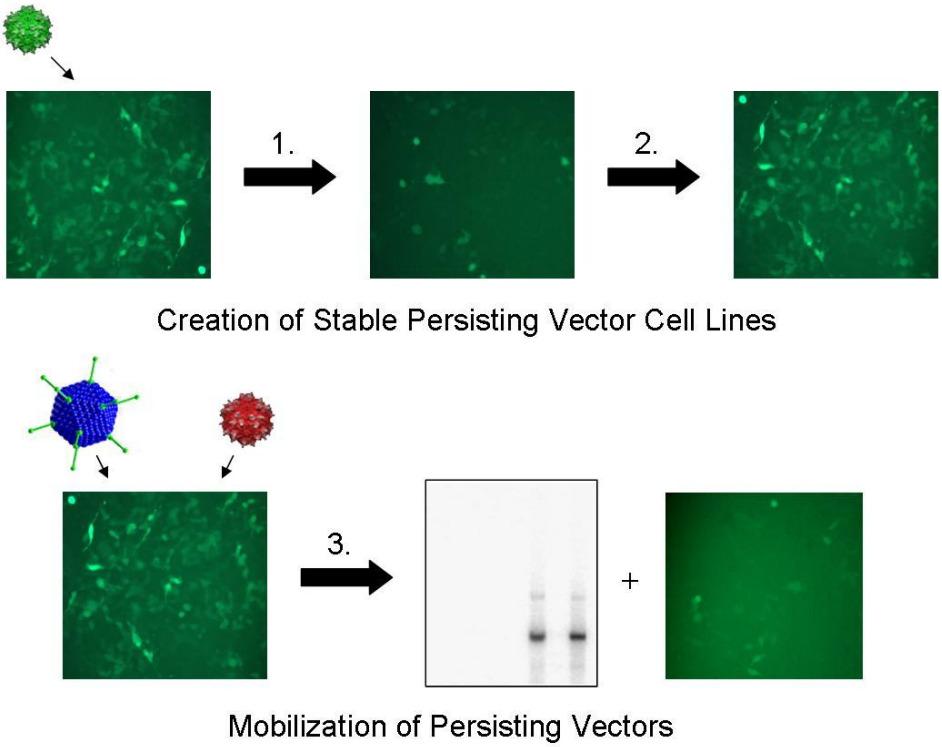


Figure 11. AAV Vector Mobilization Assay.
Cells infected with a rAAV GFP reporter flanked by the ITR of any serotype are passaged for 18-20 days, typically resulting in 1-4% of cells GFP positive (1). GFP positive cells are sorted and pooled (2). Cells are infected or transfected with Ad and any AAV serotype in order to mobilize genomes (3). Mobilization is determined by Southern blot of replicating genomes as well as infection of naïve cells with lysate.

Slightly different levels of replication were detected for pTR2-eGFP vectors in the presence of Rep2 or Rep6. (Figure 7B and 10A). While this may suggest Rep6 replicates ITR2s with higher fidelity than Rep2, it is more likely that more Rep6 protein was produced by the plasmid constructs. Western blots were not performed due to the lack of a suitable Rep6 antibody. Interestingly, our ITR5 vector was rescued with lower fidelity from 293 cells than our ITR2 vector (Figure 10A). We have confirmed that our Rep2 and Rep5 constructs produce equivalent amounts of protein by western blot (data not shown) and that they drive comparable amounts of vector genome replication (Figure 7B). As such, the decreased rescue of ITR5 genomes is most likely due to the inability of a subset of the ITR5-flanked GFP vectors in this population to be rescued due to deletions of the integrated or concatamerized ITRs as seen with AAV 2 latent genomes (106). It is possible, however, that persisting ITR5 genomes may be somewhat refractory to rescue and further experiments may be required to definitively answer this question.

It is impossible to quantify the degree of safety ITR5 based vectors would add to AAV clinical applications. Based on the exclusivity of the Rep5-ITR5 interaction, as well as the small degree of AAV5 in the population compared to AAV2 and AAV6, we can only postulate that ITR5 vectors possess a significantly lower risk of spreading after rAAV administration. While ITR5 based vectors may be markedly safer, they are not a solution.

More importantly, this work provides an assay to test the ability of any AAV vector to be mobilized by any wt AAV serotype. While the ability to test for mobilization of a vector by an AAV2 provirus is critical, such an approach would require the creation of new cell lines containing a provirus of every AAV serotype so that each ITR vector could be screened for mobilization. Instead, the ability to infect with a vector harboring any ITR and

then mobilize with the Rep and Cap of any serotype will allow efficient screening of any novel or non-human ITR.

Chapter 3: Mapping Protein/DNA Specificity of the AAV Origin of Replication

3A. Introduction

Having developed assays to screen vectors for mobilization against an AAV2 provirus as well as to screen any vector against mobilization by any AAV serotype, the next step towards preventing vector mobilization was the creation of a novel Rep-ITR interaction. To do so, however, required an understanding of the elements of the AAV origin of replication that govern replicative specificity. Identifying these elements was necessary to aid the rational design that would be required to engineer a new and unique Rep-ITR interaction.

Initial mapping of the Rep protein has revealed that only the unique N-terminus of the large Rep proteins possesses the ability to bind site-specifically and to nick the ITR, specifically the N-terminal 208aa (107). As such, chimeric Rep proteins have been created which carry the N-terminal 200 residues of one serotype and the C-terminus of the other. These chimeric Reps specifically replicate the ITR corresponding to its N-terminus (107).

Previously, AAV replicative specificity was postulated to be driven by the trs sequence (16). Rep2 can nick the ITR2 trs (AGT/TGG) and the AAVS1 trs of human chromosome 19 (GGT/TGG; 98). Rep5 nicks only the ITR5 trs (AGTG/TGG). However, alignment of the ITR2 and ITR5 sequences revealed several significant sequence and structural differences outside the trs sequence (Figure 4). The spacing between the putative RBE and the nicking stem was significantly different; three nt for ITR2 and 15 nt for ITR5. Additionally, while the trs sequence is not tightly conserved between ITR2 and ITR5, neither is the height or overall length of the putative nicking stem.

In order to address these concerns the ITR was synthesized and amplified in halves (Figure 12). Assembly of the halves required the inclusion of a *SfiI* site in one of the hairpin arms of the ITR. *SfiI* allowed the conservation of the RBE' sequence (10). Cloning the ITR in

a DD format required only one ITR per plasmid for replication (103). The three core Rep functions necessary for AAV replication (Rep binding, helicase, and nicking) were analyzed by the presence or absence of intracellular replication of the plasmid. This assay provided the ability to quantitate Rep-ITR function in a physiological setting, removing the concern that highly purified Rep protein might take on aberrant function *in vitro*. This system also avoided concerns that previous *in vitro* assays used only a fragment of the ITR or that oligos used to recapitulate the ITR might not fold correctly.

Identification of the elements involved in Rep-ITR specificity stands to increase the understanding of viral and cellular DNA binding and endonucleolytic proteins. It is likely that similar interactions take place in a wide range of viral and cellular replication and repair pathways. Localization of these elements may also facilitate the identification of other unique Parvovirus origins of replication. Here, we demonstrate two unique mechanisms at the DNA and protein level to achieve Rep-ITR specificity and utilize these factors to create a novel AAV origin of replication.

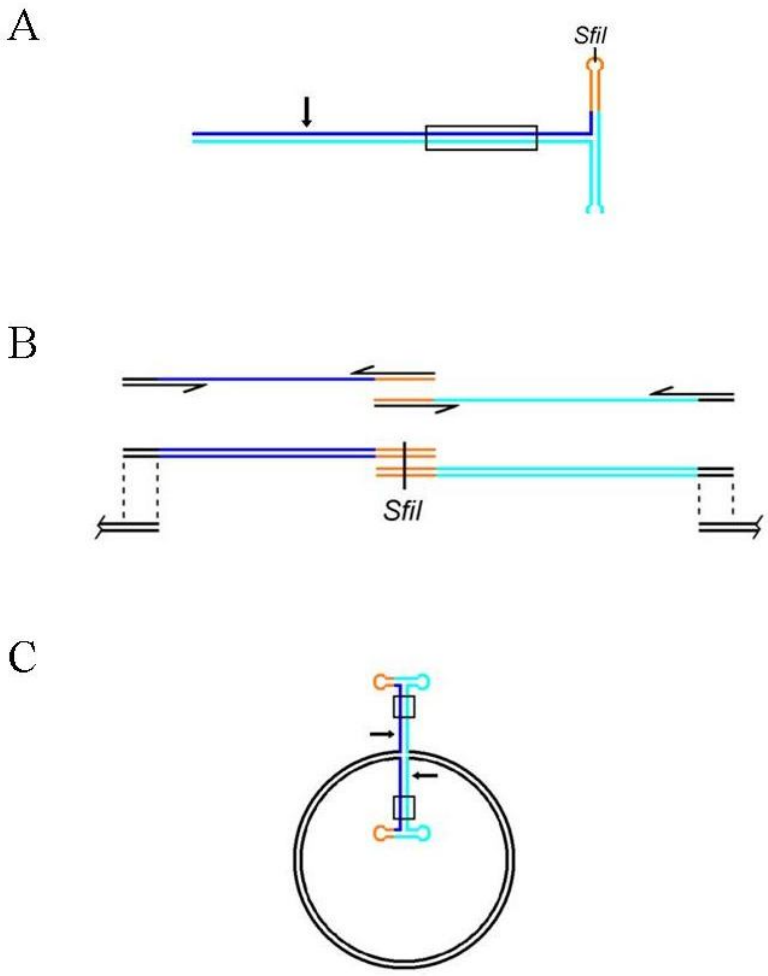


Figure 12. Diagram of ITR synthesis.
(A) The ITR was synthesized as oligonucleotides in two pieces (dark blue and light blue) overlapping across one hairpin stem containing the *SfiI* site (orange).
(B) Each half was amplified via PCR prior to digestion and cloning.
(C) Proper triple-ligation with pUC18-CMV GFP produced an ITR in DD format.

3B. Materials and Methods

B1. ITR Cloning. ITRs were cloned into a pUC-18 plasmid with a GFP cassette (CMV promoter, SV40 polyA) cloned between the *KpnI* and *EcoRI* restriction sites. The ITRs were synthesized in two halves as 4nmol Ultramer DNA oligos (Integrated DNA Technologies). *SfiI* restriction sites were incorporated into one hairpin arm the ITR for cloning (Figure 12). Due to inconsistencies of the reported sequence at the tip of the ITR5 hairpins between Chiorini et al., 1999, the published genbank sequence (genbank accession number NC_006152), and restriction mapping (data not shown), an ITR2 hairpin was utilized for the ITR5 construct (Figure 12). 200pg of each oligo was amplified in a PCR reaction using the ITR primers listed in supplemental table 1. 2.5U of PfuTurbo DNA Polymerase (Stratagene) was used to amplify each half of the ITR as follows: 1 cycle at 94^{o}C for 4 minutes, 35 cycles of 45 seconds at 94^{o}C, 30 seconds at 50^{o}C, and 30 seconds at 72^{o}C, 1 cycle of 10 minutes at 72^{o}C. PCR reactions were purified and subject to digestion by *KpnI* and *SfiI* or *HindIII* and *SfiI* (NEB). A triple ligation with the pUC-18 GFP plasmid and each half of the ITR was performed with T4 DNA Ligase (Invitrogen) for 1.5 hours at room temperature. All constructs were verified by DNA sequencing at the UNC-CH Genome Analysis Facility after linearization of the plasmid and ablation of the ITR secondary structure by *SfiI* digestion.

B2. Rep Cloning. pXR2 (Rep2Cap2) and pRep5Cap2 AAV helper plasmids served as templates for Rep cloning. The primer sequences used are indicated in Table 3. Two cloning strategies were used. Existing restriction sites were incorporated into primers for PCR (PCR-RD in Table 3) utilizing either pXR out fw or pXR out rev primers. PfuTurbo DNA Polymerase (Stratagene) was used at the manufacturer's recommendations for all PCR

reactions. PCR-RD products were digested with the enzymes indicated in Table 3 (NEB) prior to ligation with T4 DNA Ligase (Invitrogen) according to manufacturer's instructions. Alternately, an overlap-extension mediated PCR (OE-PCR) approach was used to produce Rep chimeras (39). The Rep2 and Rep5 junction was incorporated into forward and reverse primers which were used in separate PCR reactions with the pXR out fw and rev primers (Table 3, only fw oligos indicated, rev oligos complimentary to fw). These overlapping PCR products were combined into a single PCR reaction and cycled as follows: 1 cycle at 94^{o}C for 30 seconds, 18 cycles of 30 seconds at 94^{o}C, 30 seconds at 65^{o}C, and 4 minutes at 72^{o}C, 1 cycle of 10 minutes at 72^{o}C. 1ul of this reaction was used as template for a nested PCR with the pXR in fw and rev primers. Chimeras with the N-terminus of Rep2 and C-terminus of Rep5 were cloned into the Rep25aa166 construct between the *PpuMI* and *MfeI* sites. Chimeras with the N-terminus of Rep5 and C-terminus of Rep2 were cloned into the 52aa160 construct between the *PpuMI* and *BstBI* sites. All constructs were verified by DNA sequencing at the UNC-CH Genome Analysis Facility.

Clone/ Primer	Cloning Method	Orienta-tion	Sequence
pXR out fw		Forward	5'CGAAAAGTGCCACCTGACGTCTAAGAAACC
pXR in fw		Forward	5'TCGAATTCGACGGCCAGTGAATTGTAATACGACTC
pXR out rev		Reverse	5'CCATGATTACGCCAAGCTCGGAATTAACCGCATGCG A
pXR in rev		Reverse	5'CCATGGCCGGGCCCGGATTCACC
Rep52aa84	PCR-RD *AleI*	Reverse	5'TTCACCCCGGTGGTTTCCACGAGCACGTGCATGTGG AAGTAGCTCTCTCCCTTTTCAAACTGCACAAAG
Rep52aa110	PCR-RD *EagI*	Forward	5'CCTCGGCCGCTACGTGAGTCAGATTCGCGAAAAACT GATTCAGAG
Rep52aa126	OE PCR	Forward	5'GTGGTCTTCCAGGGAATTGAACCCACTTTGCCAAAC TGGTTCGCGGTC
Rep52aa138	OE PCR	Forward	5'CTGGGTCGCCATCACCAAGGTAAAGAAGGGAGGCG GGAACAAGGTGGTGGATGAG
Rep52aa146	OE PCR	Forward	5'GCGGAGCCAATAAGGTGGTGGATGAGTGCTACATCC CCAATTACTTGCTC
Rep52aa160	PCR-RD *Bpu10I*	Reverse	5'ACTGGAGCTCAGGTTGGACCTTCGGCAGCAGGTAG
Rep52aa175	OE PCR	Forward	5'CGTGGACAAACCTGGACGAGTATAAATTGGCCTGTT TGAATCTCACGGAGCGTAAAC
Rep52aa187	OE PCR	Forward	5'CTGAATCTGGAGGAGCGCAAACGGTTGGTGGCGCAG CATCTGACGCAC
Rep52aa207	PCR-RD *SgrAI*	Reverse	5'GATCACCGGCGCATCCGAGAACTCACGCTGCGAAGC
Rep25aa77	OE PCR	Forward	5'TAAGGCCCCGGAGGCCCTTTTCTTTGTGCAGTTTGAA AAGGGATCTG
Rep25aa97	OE PCR	Forward	5'CCACATGCACGTGCTCGTGGAAACCTCCGGCATCTC TTCCATGGTCCTCG
Rep25aa116	PCR-RD *NruI*	Forward	5'TCAGATTCGCGAAAAACTGGTGAAAGTGGTCTTCCA GG
Rep25aa125	OE PCR	Forward	5'GAATTTACCGCGGGATCGAGCCGCAGATCAACGACT GGGTCGCCATC
Rep25aa141	OE PCR	Forward	5'GGTCACAAAGACCAGAAATGGCGCCGGCGGAGCCA ATAAGGTGGTGGATTCTGG
Rep25aa149	OE PCR	Forward	5'GAGGCGGGAACAAGGTGGTGGATTCTGGGTATATTC CCGCCTACCTGC
Rep25aa166	PCR-RD *Bpu10I*	Forward	5'CCAGCCTGAGCTCCAGTGGGCGTGGACAAACCTG
Rep25aa187	OE PCR	Forward	5'GTTTGAATCTCACGGAGCGTAAACGGCTCGTCGCGC AGTTTCTGGCAG
Rep25aa216	PCR-RD *SgrAI*	Forward	5'ATGCGCCGGTGATCAAAAGCAAGACTTCCCAGAAAT ACATGG
ITR2 Half1 Kpn		Forward	5'ATTATAGGTACCAGGAACCCCTAGTGATG
ITR2 Half 1 Sfi		Reverse	5'TAATAGGGCCCAAAGGGCCGGG
ITR2 Half2 Sfi		Forward	5'TTAATAGGCCCTTTGGGCCGGG
ITR2 Half2 Hind		Reverse	5' TATAATAAGCTTAGGAACCCCTAGTGATGGAG
ITR5 Half1 Kpn		Forward	5' ATTATAGGTACCTACAAAACCTCCTTGCTTGAG
ITR5 Half1 Sfi		Reverse	5'TTAATAGGCCCTTTGGGCCGTCGC
ITR5 Half2 Sfi		Forward	5'TTAATAGGCCCAAAGGGCCGTCGTC
ITR5 Half2 Hind		Reverse	5'TATAATAAGCTTTACAAAACCTCCTTGCTTGAGAG

Table 3. Oligonucleotides Utilized for Chimeric Rep Cloning. OE PCR indicates the oligo was used for the overlap extension method of cloning. PCR-RD indicates that PCR was performed and a restriction digest of the product was used for cloning.

B3. Western Blot Analysis. Samples for Western blot analysis were harvested 48-72 hours after transfection of Ad-helper plasmid and the appropriate AAV helper construct. Cells were washed and resuspended in 100ul PBS prior to addition of 100ul 2x Laemmli Sample Buffer (100mM Tris pH 6.8, 4% SDS, 200mM DTT, 20% glycerol, 0.1% Bromophenol Blue). Samples were briefly sonicated and boiled for 10 minutes. Samples were run on NUPAGE 4-12% Bis-Tris gels (Invitrogen) at 160 volts for 90 minutes. Protein was transferred to a Nitrocellulose membrane (Invitrogen) via a wet transfer for 60 minutes at 30 volts. Gels were blocked overnight in 10% nonfat dry milk in 1x PBS/Tween (0.05%). Detection of both Rep2 and Rep5 proteins (all four sizes) was achieved with a monoclonal Anti-Adeno-Associated Virus Rep Protein antibody (clone 259.5, American Research Products) at a 1:20 dilution in PBS/Tween for 60 minutes at room temperature. After washing, a secondary HRP anti-mouse antibody was added at a 1:5,000 dilution in PBS/Tween for one hour at room temperature. After washing, SuperSignal West Femto Maximum Sensitivity Substrate (Pierce) was added and blots were exposed to X-ray film (Kodak).

B4. Densitometry. Densitometry was performed using the public domain NIH Image program (developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/). Agarose gels stained with Ethidium Bromide were imaged and a *DpnI*-resistant cellular DNA was used as a loading control. Values from densitometric analysis vector replication of Southern blots were divided by values obtained from the loading control. The lowest value (absence of any vector replication) was then subtracted from all values as background. In order to gauge relative replication efficiency, values for replicating TR2 vectors were divided by the value obtained from the Rep2-ITR2 control. Replicating TR5 vectors were similarly compared to the Rep5-ITR5 control.

B5. Molecular Modeling. Molecular models were generated using Swiss-Model (http://swissmodel.expasy.org). The published crystal structure of the amino-terminus of Rep5 complexed with the RBE (PDB accession #1rz9) was used as a template for all models. Visualization of protein structure rendering of images were performed with PyMOL (http://pymol.sourceforge.net).

B6. Plasmid Construction. The N-terminus of Rep8 was synthesized by Gene Art AG (Regensburg, Germany). A *PpuMI* site was incorporated into the Rep8 sequence 31 bp from the transcriptional start site corresponding to the same site in Rep2. Synthesis ended at the *BamHI* site at residue 244 common to both serotypes. Rep8 sequence was cloned into the pXR2 (Rep2Cap2) backbone used for AAV helper plasmids to create the plasmid pRep82aa244. This plasmid was then used as a template for a PCR reaction with pXR out fw (5' CGAAAAGTGCCACCTGACGTCTAAGAAACC) and Rep282 Bpu10I rev (5' ACTGGAGCTCAGGCTGAGTCTTGGGCAGGAGGTAG). The PCR product was then digested with SfiI (common to Rep8 and Rep2) as well as Bpu10I (incorporated onto the Rep8 sequence using the primer) and cloned into the pXR2 backbone to create the clone Rep282aa73-163.

3C. Results

C1. Characterization of Chimeric ITRs. An alignment of ITR2 and ITR5 revealed several divergent elements which might infer Rep specificity (Figure 4). The RBE and hairpins seemed unlikely to impact specificity as Rep2 and Rep5 have been reported to bind ITR2 and ITR5 interchangeably (16). Additionally, no evidence has been presented for Rep interactions with the portion of the D-element outside the nicking stem. Therefore, the spacer and nicking stem elements appeared to be the most likely candidates for unique interactions

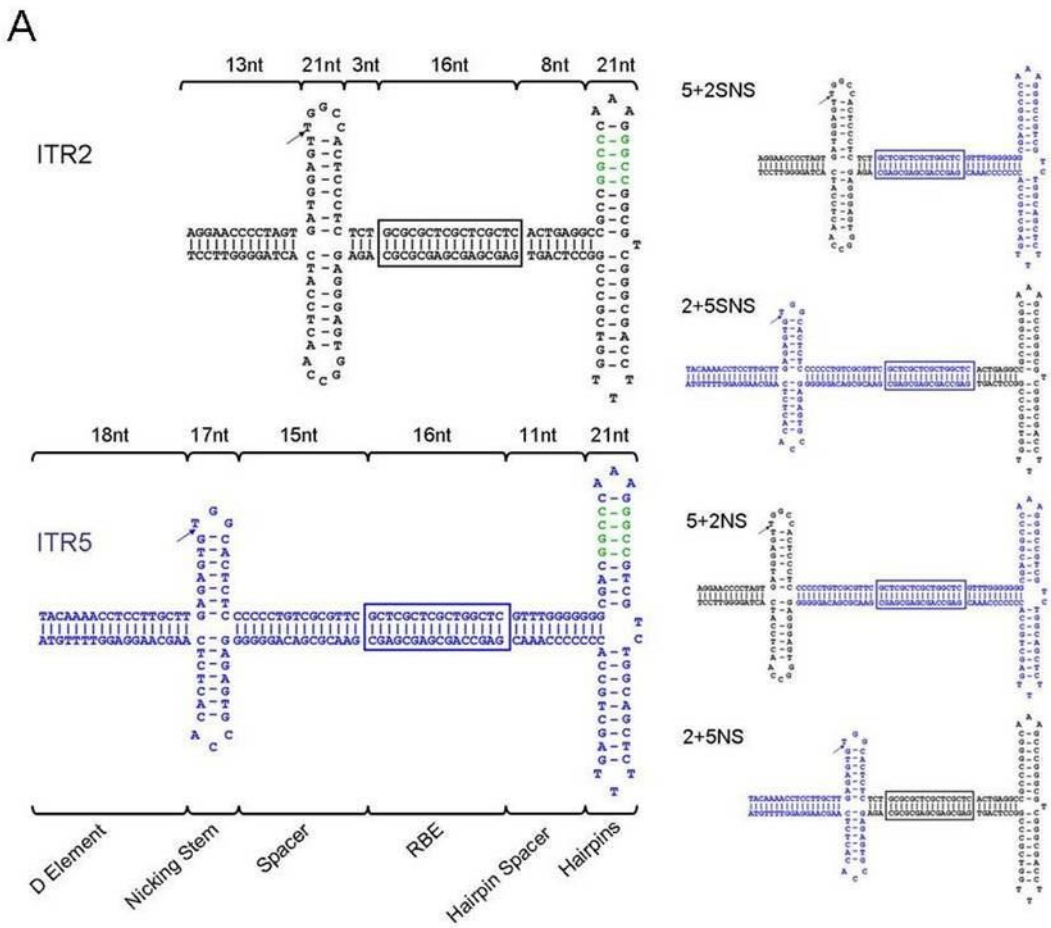
with their cognate Rep protein. This hypothesis was supported by low homology of these elements between AAV2 and AAV5.

Wt ITRs containing the *SfiI* site functioned as expected with Rep2 specific to ITR2 and Rep5 specific to ITR5 (Figure 13B). Rep2-ITR2 replicated approximately 2-fold more efficiently than Rep5-ITR5. This is potentially due to the lower folding energy of ITR5 resulting in reduced plasmid stability prior to replication. Due to this minor difference in replicative fidelity, all ITRs replicated with Rep2 were normalized to Rep2-ITR2, while ITRs replicated with Rep5 were normalized to Rep5-ITR5 (Figure 13B).

In order to confirm that the RBE and hairpin arms played no role in Rep specificity, we generated a chimeric ITR with ITR5 binding elements and an ITR2 spacer and nicking stem (ITR5+2SNS). Only Rep2 replicated this ITR, confirming the determinants of replicative specificity lie in the spacer/nicking stem elements (Figure 13B). While ITR5+2SNS replication was not as efficient as ITR2-Rep2, it was replicated at ITR5-Rep5 levels. Conversely, Rep5 specifically replicated an ITR comprised of ITR2 hairpins and hairpin spacer and the ITR5 spacer and nicking stem (ITR2+5SNS, Figure 13B). Rep5 replicated this ITR at wt levels. This data demonstrated that Rep-ITR specificity lies outside of the ITR binding regions.

Next, we explored whether the nicking stem or the spacing between the RBE and nicking stem harbored unique interactions with the Rep protein by creating chimeric ITRs which divided these individual elements. An ITR with the ITR5 binding elements and spacer and the ITR2 nicking stem could not be replicated by either Rep2 or Rep5 (ITR5+2NS, Figure 13B). The corresponding chimeric ITR (ITR2 binding elements and spacer with an ITR5 nicking stem) was replicated by both Rep2 and Rep5 (ITR2+5NS, Figure 13B). This

disparity suggested that the spacer and nicking stem play different roles in Rep-ITR specificity between AAV2 and AAV5.


A
13nt
21nt
3nt
16nt
8nt
21nt
ITR2
18nt
17nt
15nt
16nt
11nt
21nt
ITR5
D Element
Nicking Stem
Spacer
RBE
Hairpin Spacer
Hairpins
5+2SNS
2+5SNS
5+2NS
2+5NS

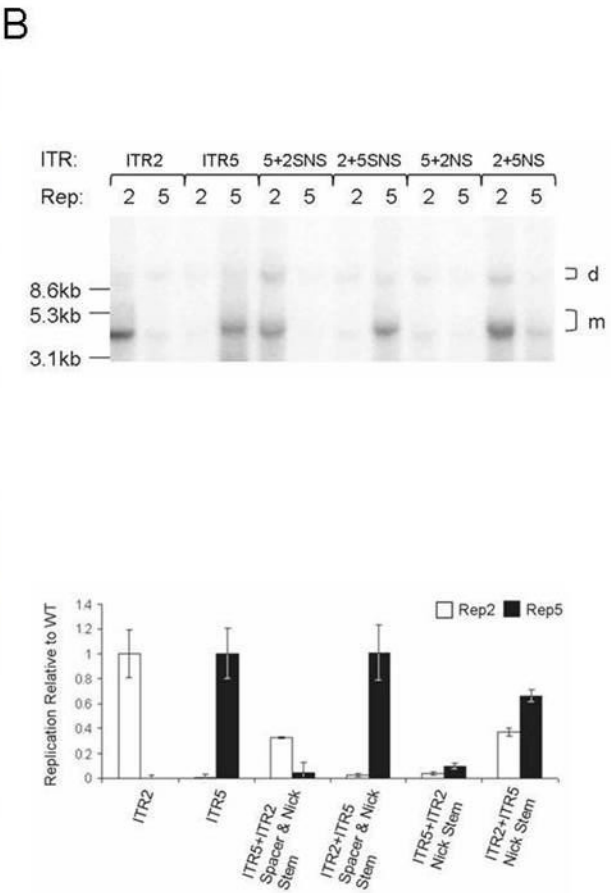

B
ITR:
ITR2
ITR5
5+2SNS
2+5SNS
5+2NS
2+5NS
Rep:
2 5 2 5 2 5 2 5 2 5 2 5
8.6kb
5.3kb
3.1kb
d
m
Replication Relative to WT
Rep2
Rep5
ITR2
ITR5
ITR5+ITR2 Spacer & Nick Stem
ITR2+ITR5 Spacer & Nick Stem
ITR5+ITR2 Nick Stem
ITR2+ITR5 Nick Stem

Figure 13. Cloning and Characterization of Chimeric ITRs.
(A) Sequence and structure of ITR2 (black) and ITR5 (blue) shown with incorporation of *SfiI* sites for cloning (green). Length in nt of ITR elements indicated above brackets. RBE is boxed. RBE' is indicated by a hatched circle. Nicking stem is extruded with arrow indicating the nicking site and hatched box indicating the trs. The four initial chimeric ITRs generated are shown (right).
(B) Replication assay and quantitation of chimeric Reps. Replication products from the indicated ITR and either Rep2 or Rep5 were analyzed by Southern blot. Monomeric (m) and dimeric (d) replicating species are indicated. The level of replication of each sample was measured by densitometric analysis and compared to wt replication.

C2. Nicking Stem Sequence and Height are Critical for ITR5 specificity.

ITR2+5NS established that Rep2 is capable of nicking an ITR with an ITR5 nicking stem and that Rep-ITR specificity is not driven exclusively by the trs sequence (Figure 13B). In order to determine the flexibility of Rep2 toward mutant nicking stems, we generated ITR2s containing altered forms of the hairpin (Figure 14A). Rep2 is able to replicate an ITR with an ITR5 nicking stem even though the ITR5 nicking stem contains a different trs sequence, is one bp shorter, and has two fewer unpaired nucleotides at its tip (Figure 14A). The substitution of the ITR5 nicking stem into ITR2 also allowed replication by Rep5.

To determine which element of the ITR2 nicking stem prevented Rep5 activity, we altered specific portions of the ITR2 stem. First, one bp at the top of the putative ITR2 nicking stem was removed to lower the height to that of ITR5 (ITR2-TA). Removing the T-A bp also resulted in a trs resembling ITR5, nicking between G/T opposed to T/T. Rep2 continued to function on this ITR as did Rep5, demonstrating that Rep5 can tolerate five unpaired nucleotides at the tip of the stem as long as the stem height and nt sequence are correct. A similar deletion from the base of the ITR2 nicking stem reduced the height to that of ITR5 while retaining the ITR2 nicking site (ITR2-GC). Rep2 continued to function efficiently on this ITR while Rep5 activity was ablated. This data suggested that the inability of Rep5 to function on ITR2 is primarily the sequence of the trs, specifically the requirement for a nick to be generated between G/T.

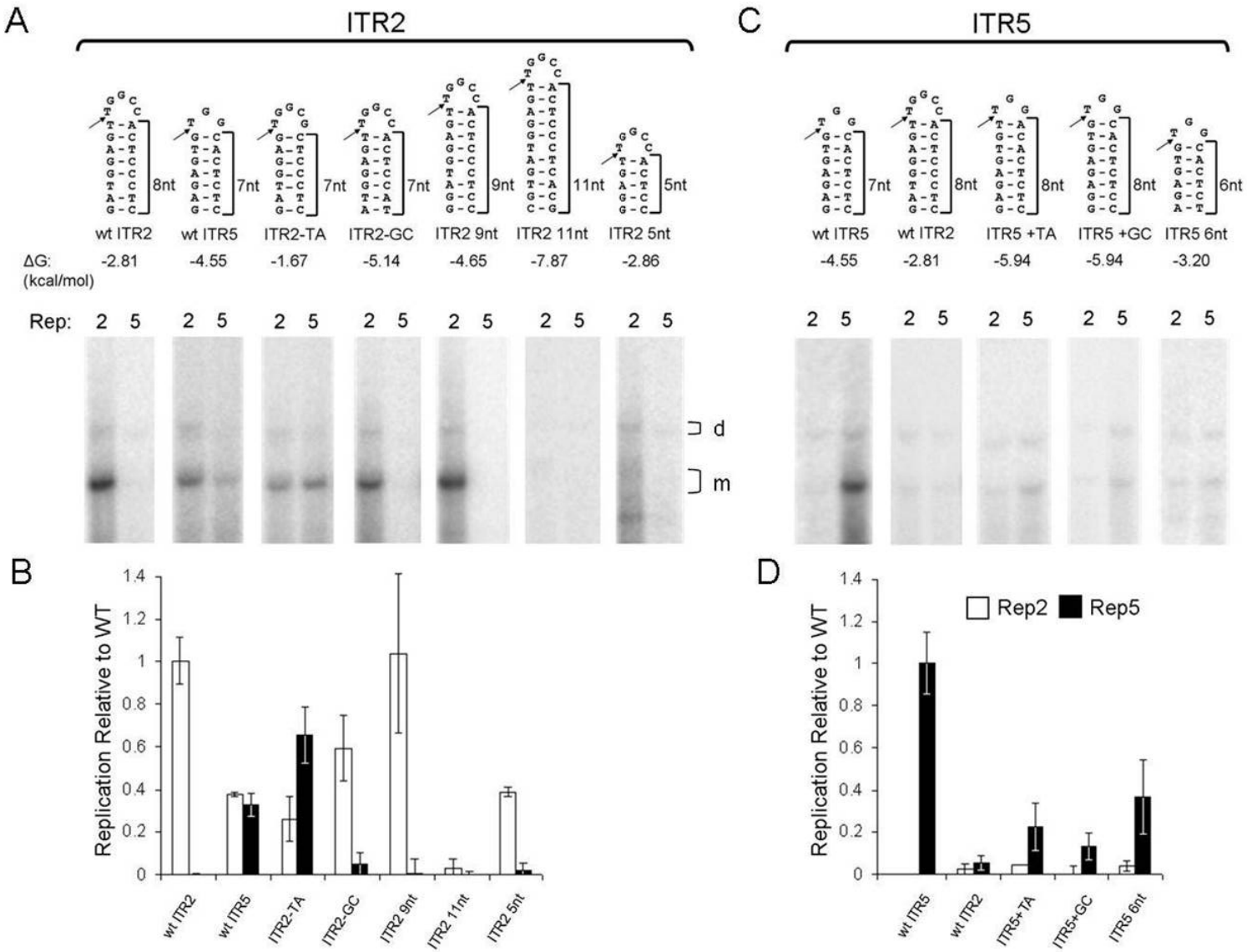

A
ITR2
8nt
7nt
7nt
7nt
9nt
11nt
5nt
wt ITR2
wt ITR5
ITR2-TA
ITR2-GC
ITR2 9nt
ITR2 11nt
ITR2 5nt
ΔG: (kcal/mol)
-2.81
-4.55
-1.67
-5.14
-4.65
-7.87
-2.86
Rep: 2 5 2 5 2 5 2 5 2 5 2 5 2 5
d
m
C
ITR5
7nt
8nt
8nt
8nt
6nt
wt ITR5
wt ITR2
ITR5 +TA
ITR5 +GC
ITR5 6nt
-4.55
-2.81
-5.94
-5.94
-3.20
2 5 2 5 2 5 2 5 2 5
B
Replication Relative to WT
1.4
1.2
1
0.8
0.6
0.4
0.2
0
wt ITR2
wt ITR5
ITR2-TA
ITR2-GC
ITR2 9nt
ITR2 11nt
ITR2 5nt
D
Rep2
Rep5
Replication Relative to WT
wt ITR5
wt ITR2
ITR5+TA
ITR5+GC
ITR5 6nt

Figure 14. Relation of Nicking Stem Height and Sequence to Rep-ITR Specificity.
(A) Sequence of nicking stem in an otherwise ITR2 context. Arrow indicates trs site. Brackets indicate height of putative stems in nt from the base of the stem to the putative nicking site. Predicted ΔG values for the hairpins are below. Southern blot analysis of the ITRs replicated by Rep2 or Rep5 is shown below.
(B) Quantitation of the Southern blots relative to wt replication from (A).
(C) Same as (A), except nicking stems indicated were used in an ITR5 context.
(D) Quantitation of the Southern blots relative to wt replication from (C).

To determine the extent of Rep2 flexibility for different nicking stems, we created three additional ITR2 mutants. Extending the nicking stem by one bp at the base had no effect on replication by Rep2 (ITR2 9 nt). However, a three bp extension was sufficient to ablate Rep2 function on the ITR (ITR2 11 nt). Surprisingly, Rep2 was able to tolerate a three bp deletion from the base of the stem, underlining the flexibility of Rep2 with respect to nicking stem substrates (ITR2 5 nt).

In order to explore the level of flexibility Rep5 possessed toward non-wt nicking stems, we created a panel of mutant ITR5s harboring altered nicking stems (Figure 14C). Curiously, Rep2 replicated none of these ITRs, suggesting an element outside the ITR5 nicking stem is responsible for preventing Rep2 function. As in Figure 13B, replacement of the ITR5 nicking stem with that of ITR2 resulted in the ablation of replication by Rep5, attributable to the incompatible trs sequence. The addition of one bp at the top of the ITR5 nicking stem severely decreased the ability of Rep5 to replicate the ITR (ITR5 +TA, Figure 14D). This insertion disrupted the ITR5 trs sequence and increased the size of the stem one bp. However, the low level of replication by Rep5 on ITR5 +TA suggests that the entire trs site of ITR2 is necessary to confer Rep2 specificity, not just the presence of a T/T nick site.

The addition of one bp to the base of the ITR5 nicking stem, preserving the ITR5 trs at the tip, nearly eliminated replication by Rep5 (ITR5 +GC). Likewise, the removal of one bp from the base of the ITR5 nicking stem strongly decreased replication by Rep5 (ITR5 6nt, Figure 14D). This data suggests that Rep5 is sensitive both to the height of the nicking stem as well as to the sequence of the trs. Thus, Rep5 is unable to replicate ITR2 because the ITR2 nicking stem is one bp too tall and has an incompatible trs sequence.

C3. Spacer Length is Critical for ITR2, not ITR5. While Rep2 can replicate a vector with an ITR5 nicking stem, it can not replicate wt ITR5 (Figure 13B). The only difference between ITR5+2SNS (which Rep2 can replicate) and ITR5+2NS (which Rep2 can not) is the ITR5 spacer (Figure 13B). The wt Rep2 spacer is three nt long while the wt Rep5 spacer is 15 nt long. Thus, we hypothesized that Rep2 may be sensitive to spacer length. Previous *in vitro* data supported this conclusion as insertions into the ITR2 spacer prevented nicking by Rep2 (10).

To explore the effect of spacer length on ITR2 and ITR5, we generated a series of mutant ITR2s and ITR5s with differing spacer lengths (Figure 15A and 15C). An insertion extending the ITR2 spacer to 10 nt ablated replication by Rep2 (ITR2 10nt, Figure 3B). Similarly, substitution of the ITR2 spacer with the 15 nt spacer of ITR5 also ablated replication by Rep2 (ITR2 15nt, Figure 15B). Rep5 was unable to replicate any of these vectors due to the presence of the ITR2 stem loop.

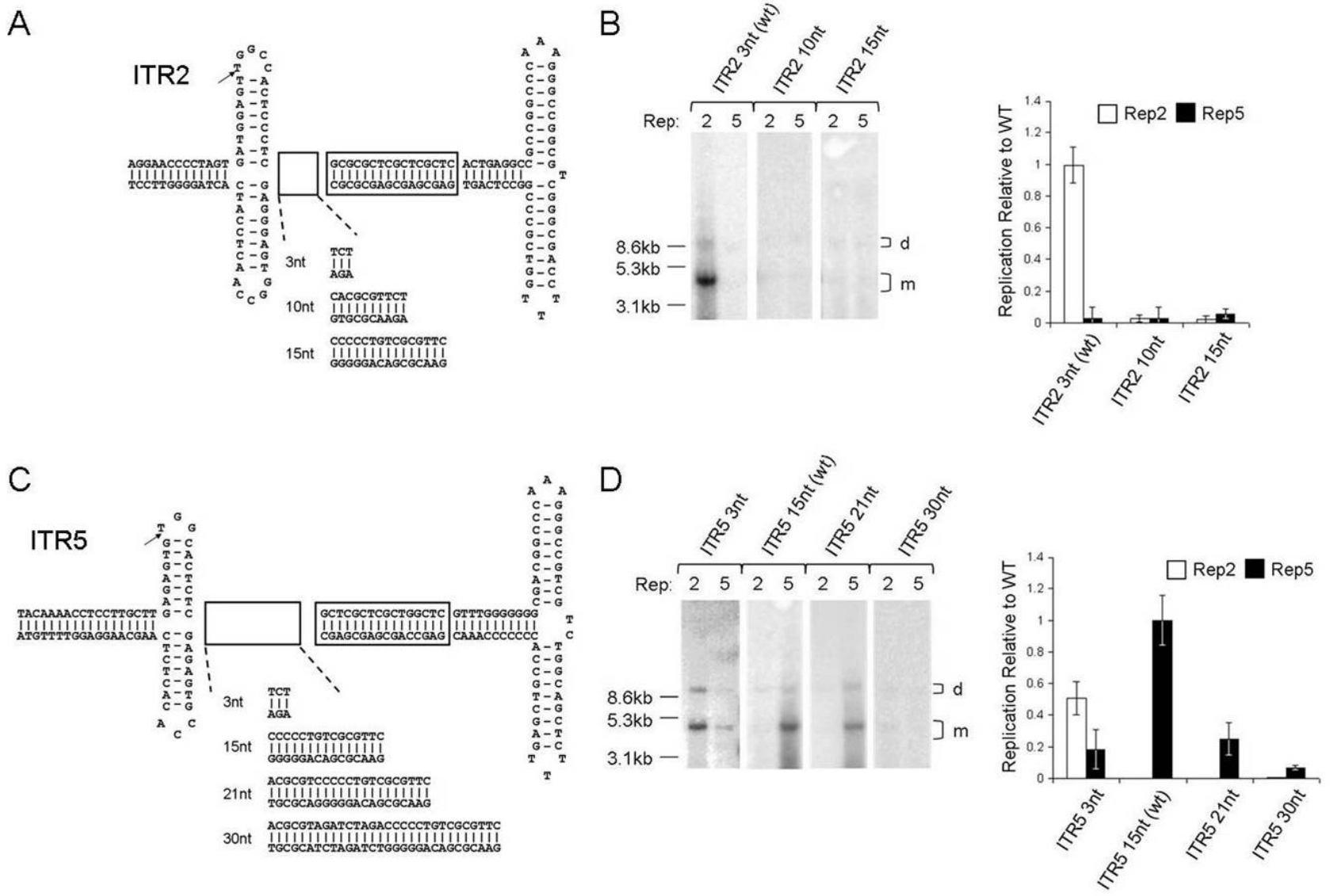

A
ITR2
3nt
10nt
15nt
B
ITR2 3nt (wt)
ITR2 10nt
ITR2 15nt
Rep: 2 5 2 5 2 5
8.6kb
5.3kb
3.1kb
d
m
Replication Relative to WT
Rep2
Rep5
C
ITR5
3nt
15nt
21nt
30nt
D
ITR5 3nt
ITR5 15nt (wt)
ITR5 21nt
ITR5 30nt

Figure 15. Effect of RBE-Nicking Stem Spacing on Rep-ITR Specificity.
(A) ITR2 mutants were synthesized with the indicated spacing between the RBE and nicking stem.
(B) Southern blot analysis of the ITRs depicted in (A) replicated by either Rep2 or Rep5 (Left). Quantitation of Southern blots relative to wt replication (Right).
(C) ITR5 mutants synthesized as in (A).
(D) Southern blot analysis and quantitation of (C).

Rep5 displayed greater flexibility toward spacer elements of differing lengths. Replacing the 15 nt ITR5 spacer with that of ITR2 resulted in an ITR which Rep5 retained the ability to replicate at a reduced level (ITR5 3nt, Figure 15D). Additionally, the presence of the three nt spacer allowed Rep2 to function on this ITR. The addition of six nt to the ITR5 spacer (for a total spacer length of 21nt) resulted in an ITR capable of being replicated by Rep5 at an efficient level (ITR5 21nt, Figure 15D). Replication by Rep5 was effectively abolished only after the insertion of 15 nt into the spacer (ITR5 30nt, Figure 15D). This panel of mutant ITR5s demonstrates the requirement for a three nt spacer element for Rep2 function.

This data confirmed that the length of the ITR5 spacer was critical to block Rep2 function. Even small insertions into the ITR2 spacer were not tolerated by Rep2. Meanwhile, Rep5 is flexible in regard to spacer length, demonstrating the ability to function on ITRs with spacers from 3-21 nt.

C4. The ITR5 Spacer Acts as a RBE for Rep5. The inability of Rep2 to function on ITRs with spacers longer than three nt led to the question of why Rep5 was so flexible in this regard. We hypothesized that Rep5 might specifically bind the ITR5 spacer just as it binds the RBE. The inability of Rep2 to bind this sequence would preclude its function on ITR5. Supporting this hypothesis was a moderately conserved GAGY Rep binding motif extending throughout the ITR5 spacer (Figure 16A). Additionally, as Rep monomers bind every four nt, the binding of three Rep5 monomers to the 15 nt spacer element would result in a three nt spacer, similar to that of ITR2 (38).

If Rep5 does bind the loosely conserved GAGY motif, the removal of that motif from the spacer should abolish Rep5 function. Indeed, the ITR5 No GAGY mutant could not be

replicated by Rep2 or Rep5 (Figure 16B). This suggested that the specific sequence of the ITR5 spacer plays an active role in the Rep5-ITR5 interaction. Conversely, a spacer with a pure GAGY repeat should not disrupt the ability of Rep5 to function on the ITR. Indeed, Rep5 was able to replicate this ITR at wt levels (ITR5 GAGY, Figure 16B). Rep2 was also able to replicate this ITR efficiently, suggesting the poorly conserved nature of the GAGY repeat within the ITR5 spacer prevents a critical DNA-protein interaction with Rep2 necessary for replication.

To explore how the ITR5 spacer functioned as an RBE, we removed three GAGY repeats from the hairpin side of the RBE (ITR5 Spacer RBE, Figure 16A). This essentially shifted the 16 nt RBE 12 nt closer to the nicking stem. Rep5 replicated this ITR efficiently, confirming the ITR5 spacer acts as a RBE (ITR5 Spacer RBE, Figure 16B). The slight reduction in replication fidelity of this ITR with respect to wt ITR5 may signal the inability Rep to properly interact with the RBE' (10). Rep2 was again unable to replicate ITR5 Spacer RBE due to its inability to interact with the ITR5 spacer.

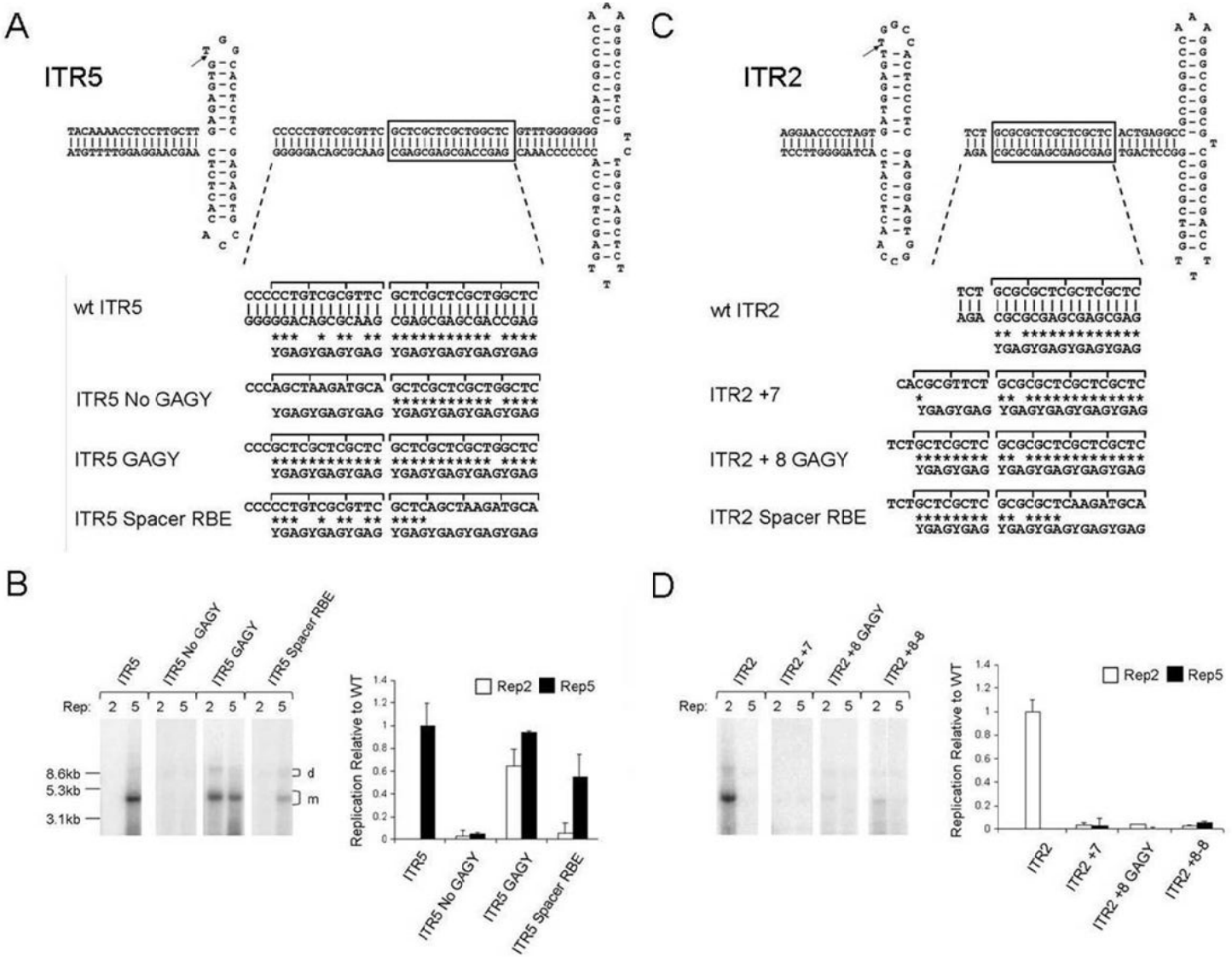

A
ITR5
wt ITR5
ITR5 No GAGY
ITR5 GAGY
ITR5 Spacer RBE
B
ITR5
ITR5 No GAGY
ITR5 GAGY
ITR5 Spacer RBE
Rep: 2 5 2 5 2 5 2 5
8.6kb
5.3kb
3.1kb
d
m
Rep2
Rep5
Replication Relative to WT
C
ITR2
wt ITR2
ITR2 +7
ITR2 + 8 GAGY
ITR2 Spacer RBE
D
ITR2
ITR2 +7
ITR2 +8 GAGY
ITR2 +8-8

Figure 16. The ITR5 Spacer Acts as a RBE for Rep5.
(A) ITR5 mutants were synthesized with the indicated RBE and spacer sequence. Brackets indicate individual tetranucleotide repeats bound by Rep monomers. Both strands of the wt ITR5 sequence are shown to illustrate conservation with the GAGY motif (indicated by *). Only one strand shown on others.
(B) Southern blot analysis of the ITRs depicted in (A) replicated by either Rep2 or Rep5 (Left). Quantitation of Southern blots relative to wt replication (Right).
(C) ITR2 mutants were generated with the RBE and spacer sequences indicated.
(D) Southern blot analysis and quantitation for (C).

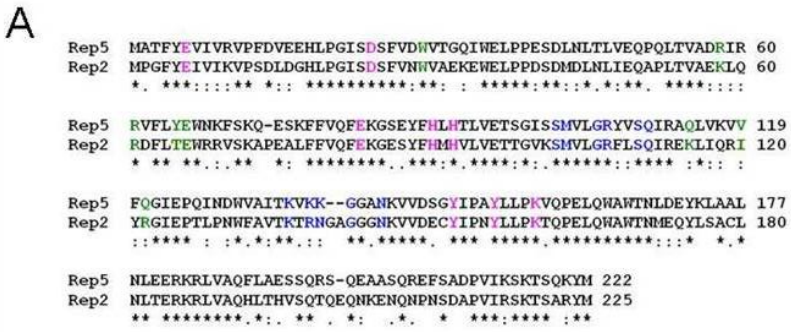
Next, we sought to extend the ITR2 spacer element to function as an extended RBE (Figure 16C). The seven nt insertion attempted in Figure 15A possessed essentially no GAGY homology (ITR2 +7, Figure 16C). As a result, Rep2 could not replicate this ITR (Figure 16D). Eight nt (two four nt GAGY repeats) inserted into the ITR2 spacer between the RBE and the existing spacer prevented replication by Rep2, demonstrating that the ITR2 RBE can not be extended.

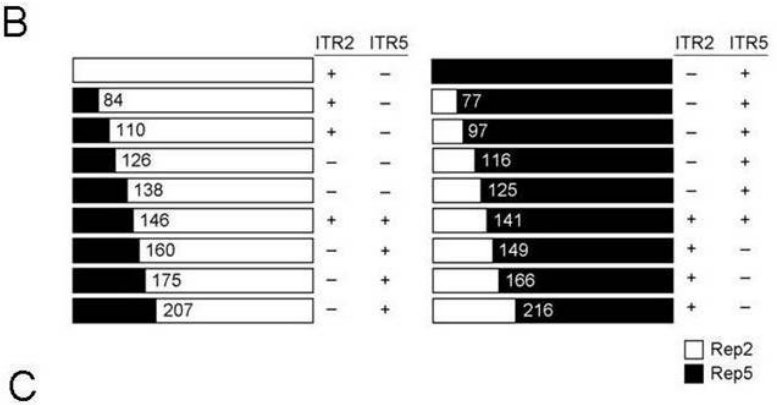
Similar to ITR5 Spacer RBE, we retained the eight nt GAGY insertion into ITR2 while removing eight nt of GAGY from the hairpin side of the RBE (ITR2 +8 -8, Figure 16C). This shifted the RBE eight nt closer to the nicking stem. Rep2 replicated this ITR very inefficiently at a level below the detection threshold of densitometric analysis (Figure 16D, Southern).

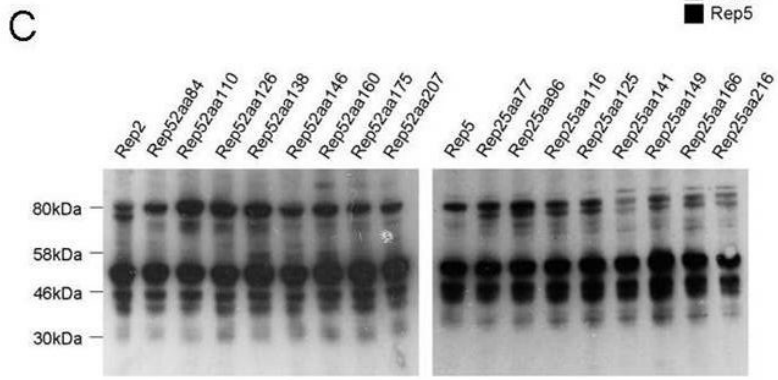
C5. Identification of Regions of Rep Responsible for ITR Specificity. Identifying the two elements of the ITR responsible for Rep specificity allowed us to map the regions of Rep2 and Rep5 involved in ITR specificity. We focused exclusively on the N-terminal 208 aa of the large Rep proteins as this region encompasses the DNA binding and endonucleolytic activity of the protein (107). This region displays approximately 60% sequence conservation evenly distributed across the protein sequence (Figure 17A). Residues involved in the active site of the protein are 100% conserved between Rep2 and Rep5 (37). Residues implicated in binding the RBE' are highly conserved (38). Residues which bind the RBE display nearly perfect conservation except for two conservative substitutions near aa 140.

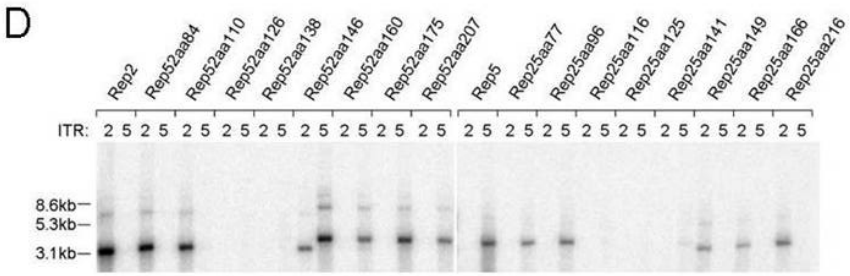
In order to map the regions of Rep involved in ITR specificity, we generated a panel of chimeric Reps derived from Rep2 and Rep5 (Figure 17B). The ability of each chimeric

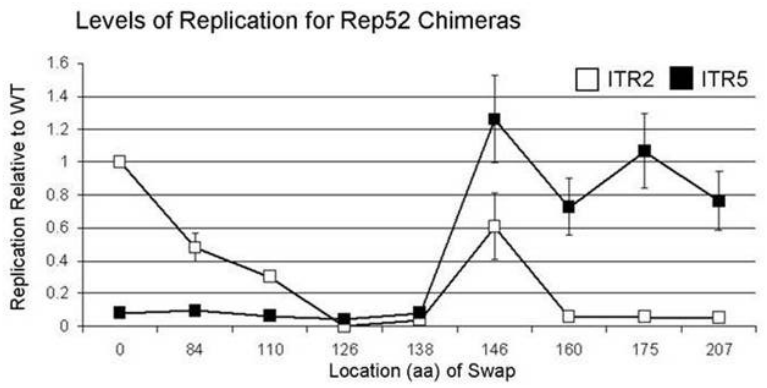
Rep to replicate an ITR2- or ITR5-flanked vector in HEK 293 cells was determined by Southern blot (Figure 17B and 17D). Each Rep in the panel was verified by DNA sequencing and Western blot analysis (Figure 17C). Every chimeric Rep showed similar protein expression profiles compared to wt. Densitometric analysis provided a comparison of the replication efficiency of each chimeric Rep with that of wt Rep2 or Rep5 (Figure 17E). Chimeric Reps were named according to the aa location of the swap between serotypes; for instance, Rep25aa77 possesses the N-terminal 76 aa of Rep2 and the C-terminus of Rep5.


A
Rep5 MATFYEVIVRVPFDVEEHLPGISDSFVDWVTGQIWELPPESDLNLTLVEQPQLTVADRIR 60
Rep2 MPGFYEIVIKVPSDLDGHLPGISDSFVNWVAEKEWELPPDSDMDLNLIEQAPLTVAEKLQ 60
Rep5 RVFLYEWNKFSKQ-ESKFFVQFEKGSEYFHLHTLVETSGISSMVLGRYVSQIRAQLVKVV 119
Rep2 RDFLTEWRRVSKAPEALFFVQFEKGESYFHMHVLVETTGVKSMVLGRFLSQIREKLIQRI 120
Rep5 FQGIEPQINDWVAITKVKK--GGANKVVDSGYIPAYLLPKVQPELQWAWTNLDEYKLAAL 177
Rep2 YRGIEPTLPNWFAVTKTRNGAGGGNKVVDECYIPNYLLPKTQPELQWAWTNMEQYLSACL 180
Rep5 NLEERKRLVAQFLAESSQRS-QEAASQREFSADPVIKSKTSQKYM 222
Rep2 NLTERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRSKTSARYM 225


B
ITR2 ITR5
84
110
126
138
146
160
175
207
77
97
116
125
141
149
166
216
Rep2
Rep5


C
Rep2
Rep52aa84
Rep52aa110
Rep52aa126
Rep52aa138
Rep52aa146
Rep52aa160
Rep52aa175
Rep52aa207
Rep5
Rep25aa77
Rep25aa96
Rep25aa116
Rep25aa125
Rep25aa141
Rep25aa149
Rep25aa166
Rep25aa216
80kDa
58kDa
46kDa
30kDa


D
ITR: 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5
8.6kb
5.3kb
3.1kb


E
Levels of Replication for Rep52 Chimeras
ITR2
ITR5
Replication Relative to WT
Location (aa) of Swap
0 84 110 126 138 146 160 175 207

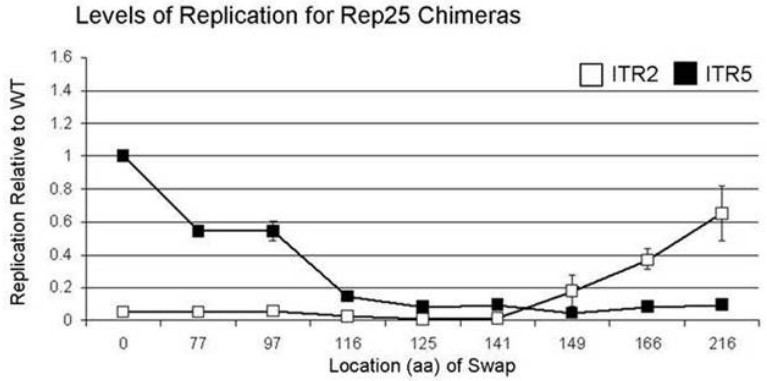

Levels of Replication for Rep25 Chimeras
ITR2
ITR5
Replication Relative to WT
Location (aa) of Swap
0 77 97 116 125 141 149 166 216

Figure 17. Cloning and Characterization of Chimeric Reps.
(A) An alignment of the N-termini of Rep2 and Rep5. (*) represents conserved amino acids. (: and .) indicates conservative substitutions. Blue indicates residues implicated in RBE binding interactions. Pink indicates residues which participate in the endonucleolytic active site. Green indicates residues implicated in RBE' binding.
(B) Chimeric Reps created and their ability to replicate ITR2 or ITR5 flanked vectors. Numbers indicate the aa position of the switch from one Rep to the other. (+) indicates the presence of replication, (–) indicates the absence.
(C) Western blot for expression of the chimeric Reps.
(D) Southern blot demonstrating replication of an ITR2 or an ITR5 vector by the chimeric Reps. Note that the ITR5 vector is 500bp larger than the ITR2 vector.
(E) Level of replication of the chimeric Reps relative to wt Rep2 or Rep5.

In the case of Rep5, replacement of the N-terminal 77 or 97 aa with Rep2 had no effect on ITR specificity nor a significant impact on replicative fidelity (Figure 17D and 17E). Larger pieces of Rep2 substituted onto the N-terminus of Rep5 were sufficient to prevent efficient replication of ITR5s (Rep25aa116, Rep25aa125, and Rep25aa141). This suggested that these chimeras possessed interruptions of a critical region of Rep5 for ITR5 specificity.

Rep2-based chimeras were unable to replicate ITR5s without the inclusion of the N-terminal 146 aa of Rep5 (Rep52aa146, Figure 17D). Rep52aa146 replicated ITR5 at wt levels, as did the three chimeras with larger portions of Rep5 on the N-terminus (Rep52aa160, Rep52aa175, Rep52aa207). This mapping reveals that the critical region for ITR specificity in Rep5 lies between aa 97-146. Surprisingly, the Rep52aa146 clone also functioned efficiently on ITR2, constituting a Rep capable of replicating ITR2 and ITR5. This suggested that ITR specificity existed in two different regions of Rep.

For Rep2, the N-terminal 83 or 109 aa of Rep5 could be substituted with no effect on ITR specificity or major influence on replicative fidelity (Rep52aa84 and Rep52aa110, Figure 17D and 17E). Chimeras including slightly larger portions of Rep5 were unable to replicate either ITR, again suggesting the interruption of a domain critical for ITR specificity (Rep52aa126 and Rep52aa138). Rep5-based chimeras were unable to replicate ITR2s without the inclusion of the N-terminal 149 aa of Rep2. However, ITR2 replication was inefficient (Figure 52aa149, Figure 17D and 17E). The inclusion of larger portions of Rep2 allowed replication of ITR2s to increase to wt levels (Rep25aa166, Rep25aa216). This data maps the Rep2 region involved in ITR specificity to aa 110-149. However, unlike Rep5, this was not the only region which played a role in ITR specificity. The ability of the Rep52aa146 chimera to replicate ITR2 and ITR5 vectors demonstrated a second region of Rep2 between

aa 138-160 sufficient to allow replication of ITR2s even when the other critical region (aa 110-149) was Rep5. The isolation of two different Rep regions involved in ITR specificity was consistent with the discovery of two independent elements governing specificity within the ITR.

C6. Characterization of Rep Regions Involved in ITR Specificity. To characterize the Rep domains identified in Figure 5, we created chimeric Rep proteins which specifically exchanged the regions implicated in ITR specificity (Figure 18A). Region 1 existed in Rep2 from aa 110-149 and in Rep5 from aa97-146. Region 2 lay within Rep2 from aa 149-187 and Rep5 from aa 146-187. As in Figure 5, all chimeras were verified by DNA sequencing and Western blot analysis (Figure 18B). Chimeras were then assayed for the ability to replicate ITR2- or ITR5-flanked vectors (Figure 18C).

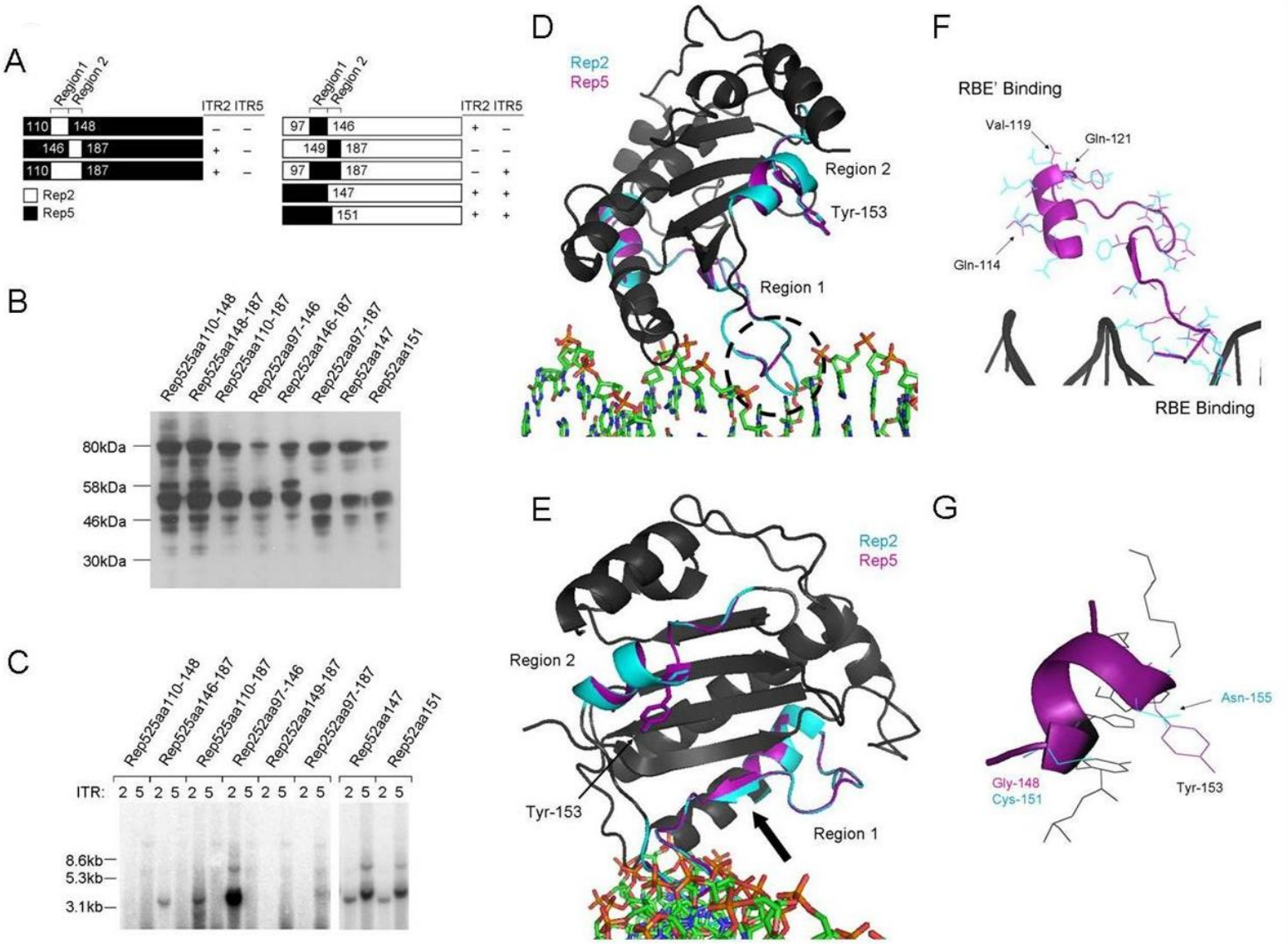

A
Region1
Region 2
ITR2 ITR5
110 148 – –
146 187 + –
110 187 + –
Rep2
Rep5
97 146 + –
149 187 – –
97 187 – +
147 + +
151 + +
B
Rep525aa110-148
Rep525aa148-187
Rep525aa110-187
Rep252aa97-146
Rep252aa146-187
Rep252aa97-187
Rep52aa147
Rep52aa151
80kDa
58kDa
46kDa
30kDa
C
Rep525aa110-148
Rep525aa146-187
Rep525aa110-187
Rep252aa97-146
Rep252aa149-187
Rep252aa97-187
Rep52aa147
Rep52aa151
ITR: 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5
8.6kb
5.3kb
3.1kb
D
Rep2
Rep5
Region 2
Tyr-153
Region 1
E
Rep2
Rep5
Region 2
Tyr-153
Region 1
F
RBE' Binding
Val-119
Gln-121
Gln-114
RBE Binding
G
Asn-155
Gly-148
Cys-151
Tyr-153

Figure 18. Characterization of Rep Regions Critical for ITR Specificity.
(A) Chimeric Reps and their ability to replicate ITR2 or ITR5 flanked vectors. Numbers indicate the aa position of the switch from one Rep to the other. (+) indicates the presence of replication, (–) indicates the absence. Region 1 and 2 involved in Rep-ITR specificity are indicated.
(B) Western blot for expression of chimeric Reps.
(C) Southern blot demonstrating replication of an ITR2 or ITR5 vector by the chimeric Reps. Note that the ITR5 vector is 500bp larger than the ITR2 vector.
(D) Structural model illustrating the two Rep regions. Rep2 structure is blue, Rep5 is purple. The nucleophilic tyrosine is indicated. Black hatched circle indicates the predicted structural difference of region 1 in the major groove of the ITR.
(E) Structural model as in (D). The nucleophilic tyrosine is indicated.
(F) Detailed structural view of region 1. The side-chains of non-conserved residues from Rep5 (purple) and Rep2 (blue) are shown. Three Rep5 residues implicated in RBE' binding are indicated.
(G) Detailed structural view of region 2. Side chains of active site residues are shown in black. Side chains of non-conserved residues in this region are shown for Rep2 (blue) and Rep5 (purple). The nucleophilic tyrosine is indicated, as is the adjacent Rep2 Asn-155.

Replacing Rep5 region 1 with Rep2 yielded a clone unable to replicate either vector, suggesting the chimera lacked the ability to bind the ITR5 spacer or nick the ITR2 nicking stem (Rep525aa110-148, Figure 18C). Replacing Rep5 region 2 with that of Rep2 allowed this chimera to replicate an ITR2 vector, suggesting region 2 of Rep2 was critical to nick the ITR2 nicking stem (Rep525aa146-187). The inability of this chimera to recognize ITR5 is harder to explain as Rep52aa146 could replicate ITR2 and ITR5 efficiently (Figure 17B). This result suggests that Rep2 region 2 makes specific contacts within Rep2 aa 188-208 which are necessary in order to function on the ITR5 nicking stem. Replacing regions 1 and 2 of Rep5 with Rep2 resulted in a Rep chimera which replicated only ITR2s (Rep525aa110-187).

Replacing Rep2 region 1 with Rep5 resulted in replication of only ITR2s, again demonstrating a connection between Rep2 region 2 and the ITR2 nicking stem (Rep252aa97-146). The lack of ITR5 replication by Rep252aa97-146 is difficult to explain based on the Rep52aa146 chimera which replicates ITR2s and ITR5s efficiently (Figure 17B). This result suggests that Rep5 region 1 makes specific contacts within the preceding 96 aa of Rep5 which are necessary in order to replicate ITR5. Replacing Rep2 region 2 with Rep5 resulted in a chimera unable to replicate either ITR (Rep252aa149-187). This chimeric Rep possesses neither Rep2 region 2 (required to nick the ITR2 nicking stem) nor Rep5 region 1 which appears necessary to interact with the ITR5 spacer. Finally, replacing both Rep2 regions 1 and 2 with Rep5 resulted in a chimera capable of replicating only ITR5 vectors (Rep252aa97-187).

The crystal structure of the N-terminal 193 aa of Rep5 complexed to the RBE allowed the location of these two critical regions to be modeled (38). The structure of the N-terminus

of Rep2 was modeled with Swiss-Model software using Rep5 as a template. The location of region 1 supports its involvement with the spacer/RBE (Figure 18D). This region interacts with the major groove of the ITR where one of the most apparent structural differences between Rep2 and Rep5 is predicted (Figure 18D, hatched circle). Rep2 contains a two aa insertion in this loop with respect to Rep5. This insertion and other non-conservative substitutions are likely responsible for the inability of Rep2 to interact with the ITR5 spacer.

Viewing Rep along the length of the ITR illustrates that region 1 constitutes much of the base of the protein (Figure 18E). Both Reps are predicted to participate in a β-sheet motif in the center of this region, while areas of reduced homology exist toward either side (the loop interacting with the major groove of the ITR on one side, RBE' interactions on the other). A more detailed look at region 1 reveals the greatest disparity between Rep2 and Rep5 occurs at the RBE binding interface in the major groove of the ITR (Figure 18F).

There is very little predicted structural difference between region 2 of Rep2 and Rep5 (Figure 18D and 18E). In an effort to dissect this region, we created two additional clones: Rep52aa147 and Rep52aa151 (Figure 18A). Like Rep52aa146, both of these Reps were able to replicate ITR2 and ITR5 vectors (Figure 18C). Rep52aa146 and Rep52 aa147 replicated ITR2 and ITR5 vectors with equivalent efficiency, suggesting E147 of Rep2 is not involved in ITR specificity. Rep52aa151 did display a modest reduction in ITR2 replication compared to Rep52aa146, suggesting that C151 of Rep2 plays a role in ITR2 specificity. Because Rep52aa160 can not replicate ITR2, this leaves only two other non-conserved residues between Rep2 and Rep5 in this region (N155 and T161). Both of these residues lie near the active site and are likely to interact with the nicking stem or active site. N155 lies directly

adjacent to Y156, the nucleophilic tyrosine, and may play a major role in ITR2 specificity (Figure 18G).

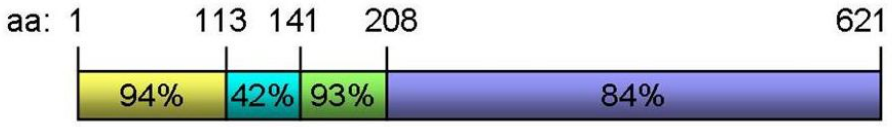
C7. Comparison of Rep Region 1 across Serotypes. Finding two relatively small regions of the Rep protein important for ITR specificity allowed those regions in the other human/primate AAV serotypes to be analyzed in silico. An alignment of the two Rep regions in AAV serotypes 1-11 was performed by ClustalW analysis. Only Rep5 possessed a significant lack of conservation in Rep region 2 with respect to the other serotypes (data not shown). This suggests that all other AAV serotypes may have flexibility similar to AAV2 toward the size and sequence of the nicking stem of the ITR. For region 1, however, there were three distinct groups (Figure 19A). Every serotype except AAV8 and AAV5 was highly conserved with AAV2. This supports previous reports of the inter-compatibility of serotypes 1-4 and 6 in their Rep-ITR interaction (30, 31). Rep5 was significantly divergent from Rep2, accounting for the inability of Rep2 to bind the extended ITR5 RBE. Surprisingly, AAV8 was poorly conserved in relation to both AAV2 and AAV5. Both the RBE' and RBE binding interfaces of region 1 of Rep8 were poorly conserved, despite extremely high homology between the remainder of the N-terminus of Rep8 with Rep2 (Figure 19B). The C-terminus of the protein was also well conserved, suggesting that the differences in region 1 may contribute to a unique origin of replication for AAV8. Additionally, the co-crystal structure of Rep5 bound to its RBE allowed us to model the N-terminus of Rep8. These models suggested that Rep8 possessed unique structural confirmations in both the RBE and RBE' interacting domains of Rep region 1 with respect to Rep2 and Rep5. This altered confirmation likely plays a critical role in ITR specificity (Figure 19C and D).

A

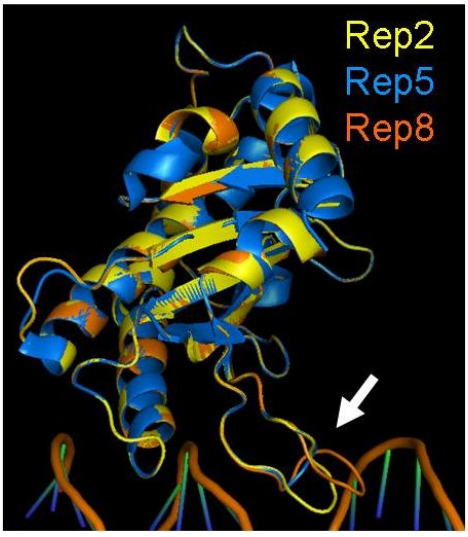
```
AAV2      EKL-IQRIYRGIEPTLPNWFAVTK--TRNGA

AAV1      D···V·T·························
AAV3b     ····VT·····V··Q·················
AAV4      ····VT·····V··Q·················
AAV5      AQ··VKVVFQ····QIND·V·I····VKK--
AAV6      D···V·T·························
AAV7      ····V·T····V····················
AAV8      ···GPDHLPA·SS··········DAVMAP·
AAV9      ····V·T·························
AAV10     DR··V·T····V····················
AAV11     DR··V·T····V····················
```

B



C



D

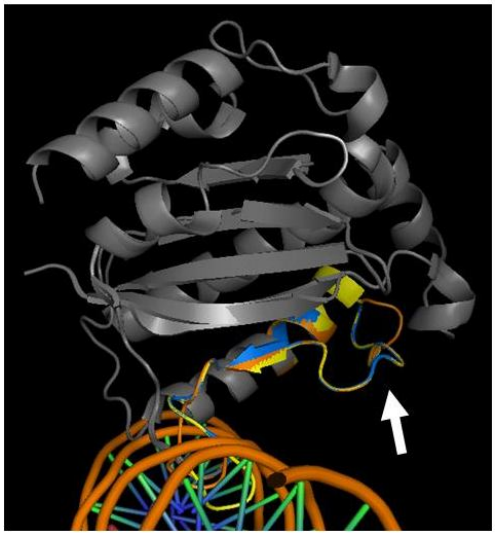


Figure 19. Alignment of Rep Region 1.
(A) The majority of characterized AAV serotypes are highly conserved with AAV2 in Rep region 1. However, Rep5 and Rep8 display poor conservation to all other human/primate AAV serotypes in this region.
(B) Homology between Rep2 and Rep8 along the full length of the Rep78 protein. Region 1 of Rep8 (aa 113-141) is poorly conserved, sharing only 42% homology to Rep2 (cyan). The remainder of the N-terminus is highly conserved (yellow and green). The C-terminus is also well conserved (purple).
(C and D) Structural models of the N-terminus of Rep2, Rep5, and Rep8. Rep8 (orange) is predicted to possess a significantly altered structure at the RBE binding interface of region 1 (white arrow in C) as well as the RBE' binding interface (white arrow in D).

Although the Rep8 sequence has been published, the ITR8 sequence has not. As we were unable to obtain the Rep8 construct, Gene Art AG (Regensburg, Germany) synthesized the N-terminal 244 aa. Rep chimeras were created with the N-terminal 244 aa of Rep8 and the C-terminus of Rep2 as well as a 50 aa section of Rep8 replacing Rep2 region 1 (Figure 20A). The N-terminal 244 aa of Rep8 was also used with the C-terminus of Rep5 (data not shown). None of these chimeric Reps were able to replicate an ITR2 or ITR5 vector, suggesting that this region of Rep8 may confer specificity unique specificity to its cognate origin of replication with respect to all other AAV serotypes (Figure 20B).

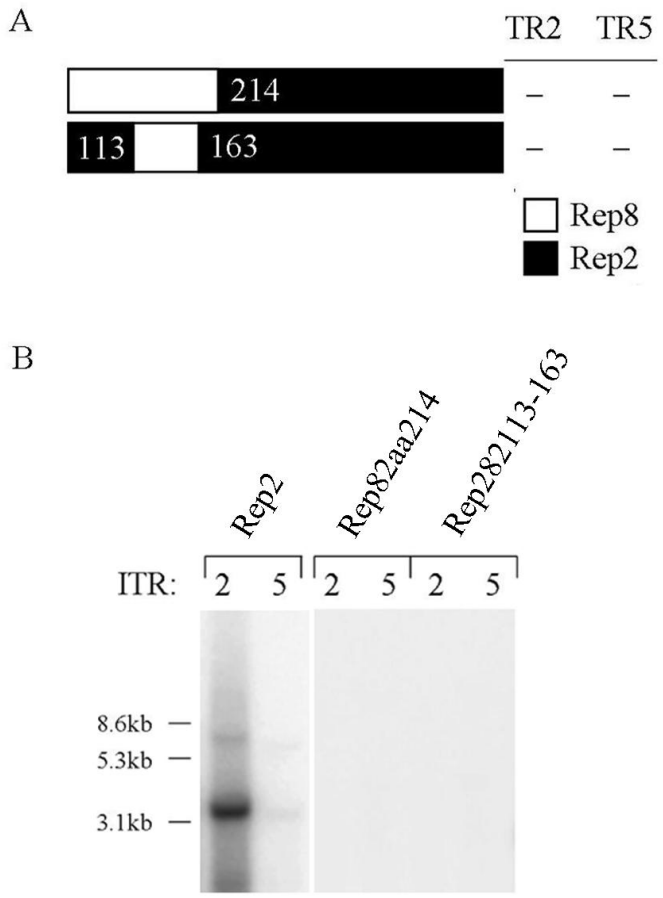


Figure 20. Chimeric Rep8 proteins are not compatible with ITR2 or ITR5.
(A) The Rep8-Rep2 chimeric proteins generated. (-) indicates the absence of replication on the indicated ITR. Numbers indicate the aa boundary of the chimera.
(B) Southern blot showing replication of ITR2 and ITR5 vectors by Rep2 and the two Rep8 chimeric constructs from (A).

C8. Structure-Function Model of Rep-ITR Specificity. In order to unify the ITR and Rep elements involved in specificity into a single model, we utilized the chimeric Reps separating region 1 and region 2 along with the chimeric ITRs separating the nicking stem and spacer. Rep2, Rep5, Rep52aa146 (which divides region 1 and 2 of Rep and can replicate ITR2 and ITR5), and Rep25aa149 (essentially no ITR2 or ITR5 replication) were selected. These Reps were tested for their ability to replicate ITR2, ITR5, ITR2+5NS (which is replicated by both Rep2 and Rep5), and ITR5+2NS (which is replicated by neither Rep2 or Rep5).

Only Rep2 and Rep52aa146 efficiently replicated ITR2 (Figure 21A and 21B). Only Rep5 and Rep52aa146 replicated ITR5. As in Figure 1, Rep2 and Rep5 replicated ITR2+5NS. Additionally, Rep25aa149 and Rep52aa146 replicated ITR2+5NS. This ITR is universally replicated by every Rep due to the exclusion of DNA elements involved in protein specificity. The three nt ITR2 spacer functions with the DNA binding region 1 of Rep2 or Rep5. The seven bp tall ITR5 nicking stem functions with region 2 of Rep2 or Rep5. Thus, any combination of these regions constitutes a Rep protein capable of replicating ITR2+5NS.

Finally, neither Rep2 nor Rep5 replicated ITR5+2NS. Rep2 is unable to interact properly with the 15 nt ITR5. Rep5 is unable to function on the ITR2 nicking stem. For these reasons, Rep25aa149 was also unable to catalyze replication. However, Rep52aa149 was able to replicate this ITR due to the proper combination of Rep regions (Figure 21C). Rep52aa149 possesses Rep5 region 1 which is necessary to interact with the 15 nt ITR5 spacer. This chimera also possesses Rep2 region 2, essential for function on the ITR2 nicking stem. This recombinant DNA-protein interaction is unique from either AAV2 or AAV5 and constitutes a novel Parvovirus origin of replication.

A
ITR:
2
5
2+5NS
5+2NS
Rep2
Rep5
Rep25aa149
Rep52aa146
8.6kb
5.3kb
3.1kb
d
m
B
Rep2
Rep5
25aa149
52aa146
ITR2
ITR5
ITR2+5NS
ITR5+2NS
C
Extended RBE
Putative RBE
ITR2
ITR5

Figure 21. Model of Rep-ITR specificity.
(A) Southern blot of Hirt DNA demonstrating replication of the indicated ITR vector by the indicated Rep.
(B) Table indicating the presence (+) or absence (–) of replication of the gel from (A).
(C) Model of a novel AAV origin of replication. The chimeric ITR can be replicated only by a chimeric Rep protein. Rep5 sequence in region 1 (blue) is required for the extended RBE of ITR5 (purple). Rep2 sequence in region 2 (yellow) is required to function on an ITR2 nicking stem (cyan).

3D. Discussion

Taken as a whole, this work illustrates two specific mechanisms of DNA-protein specificity at the Parvovirus origin of replication. Chimeric ITRs narrowed the DNA elements involved in specificity to the spacer and nicking stem sequences (Figure 13B). These results contradicted previous assertions that Rep-ITR specificity were driven solely by the nicking sequence as Rep2 efficiently nicked an ITR harboring the ITR5 nicking stem (16). Rep2 is highly flexible in the sequence and height of its nicking stem while Rep5 is highly specific to its cognate stem (Figure 14).

Three residues of Rep2 are necessary to cleave the ITR2 nicking stem (Figure 17 and 18). Residues C151, N155, and T161 all lie in the active site of the protein in a predicted alpha helix along with the nucleophilic tyrosine Y156. How these residues (termed Rep region 2) grant Rep2 flexibility toward mutant nicking stems remains unclear. The corresponding Rep5 residues (G148, A152, and V158) may participate in highly specific interactions which require specific height and sequence considerations for the ITR5 nicking stem.

AAV5 Rep-ITR specificity is mediated by the ITR5 spacer. Replacement of the three nt ITR2 spacer with the 15 nt ITR5 spacer ablated replication by Rep2 (Figure 14B). A poorly conserved Rep binding element allows Rep5 to interact with the elongated ITR5 spacer (Figure 16B). Mutating the spacer to include a strong Rep binding element allowed Rep2 and Rep5 to replicate the ITR. However, insertion of a Rep binding element into the ITR2 spacer still largely decreased Rep2 function. While this data might suggest that additional Rep5 molecules bind to ITR5, previous *in vitro* experiments have not come to this

conclusion, although those studies were performed in the absence of hairpins on the ITRs (16).

A 49 aa region of Rep5 interacts with the ITR5 spacer (aa 97-146, Figure 17 and 18). The crystal structure of the N-terminus of Rep5 reveals that this region (region 1) possesses residues which specifically bind to the RBE and RBE' of the ITR. Major structural differences in the Rep5 loop which binds the major groove of the RBE likely account for the majority of ITR5 spacer specificity. While Figure 1B predicts RBE' binding should not play a role in Rep-ITR specificity, it is possible that RBE' contacts alter the secondary structure of region 1 as it interacts with the RBE.

Because the regions of Rep critical for ITR specificity were separate (region 1 of Rep5 from aa97-146 and region 2 of Rep2 from aa151-161), a chimeric Rep possessing both regions was able to efficiently replicate ITR2 and ITR5. An ITR which could be replicated by any wt or chimeric Rep was constructed by excluding the DNA elements required for specificity; the ITR5 spacer and the ITR2 nicking stem. Most significantly, a novel origin of replication was generated. This ITR contained both of the critical elements for Rep specificity; the ITR5 spacer and the ITR2 nicking stem. As a result, only a chimeric Rep protein made up of Rep5 region 1 and Rep2 region 2 was able to replicate the ITR. The creation of a unique origin of replication highlights the power of studying the DNA-protein interactions of a viral origin of replication.

The creation of a unique DNA-protein interaction was possible because of the separation of the specific Rep-ITR interactions in AAV2 and AAV5. How and why these two different DNA-protein interactions evolved is unclear. It is likely due to evolutionary divergence in the ITR sequence which may have occurred in different hosts (AAV2 is related

to other primate AAVs, AAV5 is related to non-primate AAVs such as goat and bovine). This model of replicative specificity can likely be extended to other parvoviruses such as snake AAV which has a highly conserved T-shaped ITR structure but different spacer and nicking stem lengths (24). Similar DNA-protein interactions likely occur in distantly related viruses such as the autonomous human Parvovirus B19 which can have ITRs as long as 400 bp (22). Less conserved Parvovirus origins of replication might also employ additional DNA-protein interactions outside of the nicking stem and spacer sequences.

Additionally, this work may provide further insight into why AAV2 is the only known animal virus capable of integrating site-specifically into the human chromosome (80). Integration occurs due to the specific cleavage of the AAVS1 site on chromosome 19 by Rep2. Rep2 is highly flexible in its nicking substrates, functioning on nicking stems from five bp to nine bp in height and on poorly conserved trs sequences. Thus, the only requirement for Rep2 to nick the human chromosome would be a functional nicking stem within three nt of a consensus RBE. As there are an estimated $2x10^5$ consensus RBEs in the human chromosome, the likelihood of such an occurrence is high (108). This may also explain why an integration locus for AAV5 has not been identified. Rep5 is highly specific to both the height and sequence of the ITR5 nicking stem. There is likely no ITR5 nicking stem homolog in the human chromosome within range of a consensus RBE to allow nicking by Rep5. It is possible that other hosts infected by AAV5-related serotypes might possess chromosomal integration sites for AAV5.

These results also stand to improve the safety of future AAV therapeutic vectors. The danger of AAV vector mobilization by wt AAV could be averted if therapeutic vectors harbored ITRs which no wt Rep could replicate (36). The mechanisms of Rep-ITR specificity

described here might extend to cellular elements. For instance, the C. elegans mobile element Tc1 contains terminal repeats which are specifically bound and endonucleolytically cleaved by its transposase, Tc1A (95). Biology at the related SV40 T antigen and papillomavirus E1 origins of replication likely possess conserved interactions (37). Bacteriophage φX174 and plant geminiviruses, as well as other viruses which employ rolling circle mechanisms of replication also possess homology to the AAV origins of replication (43). In this way, dissection of specificity at the AAV origin of replication provides a broad platform to investigate other DNA-protein interactions.

Chapter 4: Methods to Prevent Vector Mobilization

4A. Introduction

The creation of a novel AAV origin of replication resulted in three different vectors to test for anti-mobilization capabilities. While ITR5 flanked vectors should be at decreased risk of vector mobilization, they do not eliminate the underlying problem (36). The three remaining possibilities for mobilization resistant vectors are: Non-homologous ITR vectors, ITRs from non-human AAV serotypes, and a novel, lab-derived AAV origin of replication.

While not complete in the prevention of vector mobilization, mixed ITR or non-homologous ITR vectors should effectively remove all risk. Mobilization of a standard vector requires a cell to be triple-infected by a rAAV vector, a corresponding wt AAV serotype, and a helper virus. The inclusion of one ITR from a different serotype should require quadruple infection by the rAAV vector, two independent AAV serotypes, and a helper virus. Decreasing the risk further would be the natural tropism of the wt AAV serotypes (in this case, AAV2 and AAV5) which is unlikely to infect the same tissue as AAV2 and AAV5 were isolated from completely separate cell types and use different cell receptors. If such a vector behaved as hypothesized, it would decrease the minor risk of vector mobilization to irrelevant levels.

Better still is the promise of using unique AAV origins of replication from serotypes which do not infect humans. If such an ITR could be used to produce vector with their cognate Rep in human cells and package vector into the capsids from human/primate AAV serotypes, the risk of mobilization would be effectively eliminated. The only risk would be the unlikely occurrence of the non-human serotype jumping into and spreading through the human population.

While Chapter 3 provided evidence that the AAV8 origin of replication may be unique within the human/primate AAV serotypes, the ITRs are not readily available for the creation of new vector. Additionally, the isolation of AAV8 from primates would suggest the virus could jump into humans relatively efficiently (25). Therefore, we looked toward distantly related AAV serotypes, discovering that the complete genomic sequence of SAAV had been characterized (24). The snake ITR (sITR) varied substantially from the ITR of AAV2 and AAV5, as did the N-terminus of the snake Rep (sRep) protein. Thus, we chose the SAAV origin of replication for study.

Finally, the findings in Chapter 3 resulted in the creation of a novel chimeric ITR which could be replicated only by a chimeric Rep protein made up of Rep2 and Rep5. Mobilization of this vector would require the quadruple infection of a cell with a rAAV vector, a helper virus, wt AAV2, wt AAV5, and a recombination event between AAV2 and AAV5 in a 42 nt region of the Rep protein. Therefore, if this novel origin can be adapted to an efficient vector production system, it will effectively negate any risk of vector mobilization by wt AAV.

4B. Materials and Methods

B1. Plasmid Construction. Non-homologous ITR plasmids were constructed from the pTR2 or pTR5-eGFP constructs from Chapter 2. Each plasmid was double digested with *AgeI* and *PciI* which resulted in two fragments, each carrying a single ITR. The corresponding fragments were then ligated to create the plasmids pTR2-5 or pTR5-2 eGFP depending whether the 5' ITR with respect to the GFP transgene was ITR5 or ITR2. The SAAV construct was a gift of Dr. Peter Tijssen. The sRepCap2 vector was created by overlap extension PCR. sRep was amplified using the primers 5' ATATATGATGCGAGCAATGACGTC

AGCGGACATGTC and CTGGAAGATAACCATCGGCAGCCATATCTGTTATCAGTAAGTTTATTGTTC CTTGTCCAT. Cap2 was amplified with the primers 5' ATGGACAAGGAACAATAAACTTACTGA TAACAGATATGGCTGCCGATGGTTATCTTCCAG and 5' CCATGATTACGCCAAGCTCGGAATTA ACCGCATGCGA. After overlap extension PCR, the sRepCap2 fragment was amplified by the nested primers 5' ATATATCATATGAGCTCCAGC GGACATGTCTGGACATGTCTTTG and 5' CCATGGCCGGGCCCGGATTCACC and inserted into pXR2 utilizing *SacI* and *SacII*.

B2. Transduction Assay. rAAV harboring ITR2, ITR5, or ITR5+2NS genomes were purified by a single CsCl gradient were titered by qPCR as in Chapter 2B. $5x10^4$ HEK 293 cells were infected by each virus at 10,000, 1,000, and 100 vg/cell. 48 hours post-infection cells were assayed visually to determine the extent of transduction.

B3. qPCR Titering for Mobilization. qPCR titering was performed as outlined in Chapter 2B. Cells harboring a persisting GFP encoding vector flanked by ITR2s, ITR5s, or ITR5+2NSs, were transfected with Ad-helper plasmid and either pXR2 (Rep2Cap2), pRep5Cap2, or pRep52aa149 (chimeric Rep specific to ITR5+2NS, Cap2). 48 hours later Hirt DNA was isolated and 10 ul was used in a qPCR reaction.

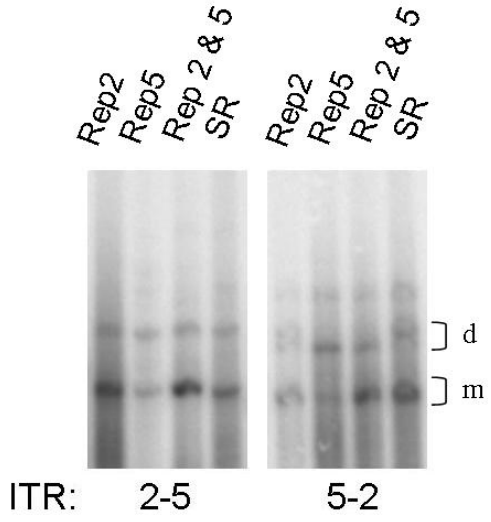
4C. Results

C1. Non-Homologous ITR Vectors. Vectors flanked by one ITR2 and one ITR5 should require co-infection of wt AAV2 and AAV5 to be mobilized. Such vectors have been used to aid directional concatemerization of rAAV genomes; however, no account of their replicative efficiency or mobilization potential have been reported (104). In order to remain consistent with the results from Chapter 2, we utilized pTR2- and pTR5-eGFP to create two non-homologous ITR vectors (Figure 22A). These constructs differed by serotype of the ITR at the 5' end of the GFP reporter gene.

A

pTR5-2-eGFP
TR5
GFP
Neo
TR2
pTR2-5-eGFP
TR2
GFP
Neo
TR5

B


Rep2
Rep5
Rep 2 & 5
SR
Rep2
Rep5
Rep 2 & 5
SR
d
m
ITR:
2-5
5-2

C

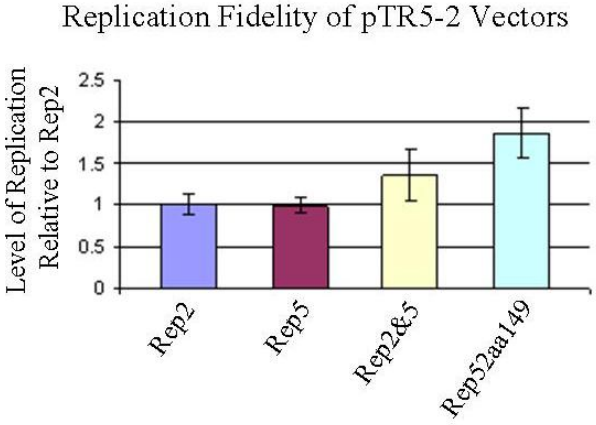

Replication Fidelity of pTR5-2 Vectors
Level of Replication Relative to Rep2
2.5
2
1.5
1
0.5
0
Rep2
Rep5
Rep2&5
Rep52aa149

Figure 22. Replication of non-homologous ITR vectors.
(A) Two non-homologous ITR vectors were created with either ITR2 or ITR5 at the 5' end of the GFP reporter gene.
(B) Southern blot analysis of replicating non-homologous ITR vectors. The indicated ITR construct was co-transfected with Ad helper plasmid and the AAV helper plasmids listed. Hirt DNA was isolated and *DpnI* digested prior to Southern blot probed with labeled fragment corresponding to the GFP orf. Monomer (m) and dimer (d) replication intermediates are indicated.
(C) Densitometery analysis of replication levels of TR5-2 vectors replicated by the Reps indicated. Replication was normalized to the Rep2-only replication of the vector. Analysis was performed in duplicate and standard error is indicated.

As Rep-ITR specificity appears consistent for replication and mobilization, the non-homologous ITR vectors were first assayed for replication in the presence of Rep2, Rep5, both Rep2 and Rep5, as well as the chimeric Rep52aa149 capable of replicating ITR2 and ITR5 (Figure 22B). Forty-eight hours after transfection, Hirt DNA was isolated and assayed by Southern blot. Replicating monomer genomes were quantitated by densitometry analysis as in Chapter 3 (Figure 22C). These results verified that Rep2 or Rep5 alone was able to catalyze replication of the genome, though at a lower level than Rep2 and Rep5 in combination. In fact, the inclusion of both Reps increased the number of replicating genomes less than 1.5-fold. Rep52aa149 increased the number of replicated genomes by nearly 2-fold. That this Rep was slightly more efficient than Rep2 and Rep5 in combination may suggest a slight inhibitory effect of one Rep on the other. This may also be due to the decreased transfection efficiency of four plasmids (Ad helper, vector, Rep2 helper, Rep5 helper) into a single cell compared to three (Ad helper, vector, chimeric Rep).

Despite the ability of a single Rep to efficiently replicate these vectors, we examined their utility as anti-mobilization tools. To do so, non-homologous ITR vector virus was used in the mobilization assay developed in Chapter 2. Twenty days after infection, GFP positive cells were sorted and pooled and persisting vectors were mobilized (data not shown). Not surprisingly, Rep2 or Rep5 alone were able to mobilize and replicate these genomes.

While this data has implications for several aspects of AAV biology, it removes non-homologous ITR vectors from consideration as anti-mobilization vectors. The ability of Rep2 or Rep5 to replicate these vectors demonstrates that they are no different for anti-mobilization purposes than the ITR2 or ITR5 flanked vectors tested in Chapter 2. Further, due to the incompatibility of the Rep-ITR interaction between AAV2 and AAV5, the ability

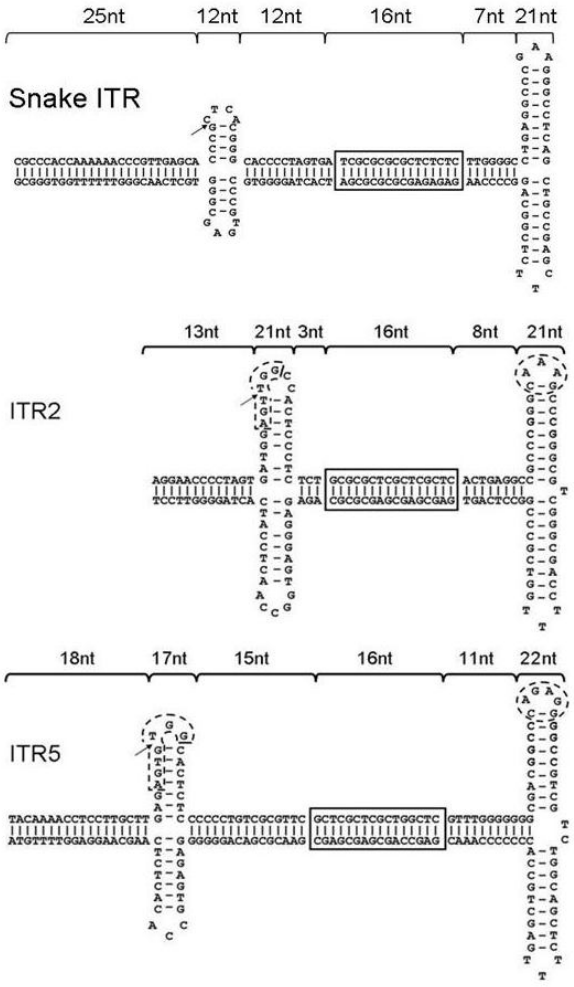
of these vectors to replicate efficiently in the presence of a single Rep suggests a gene correction event resulting in an ITR2 or ITR5 flanked genome.

C2. Snake AAV. Using the ITRs of non-human/primate AAV serotypes for AAV therapeutic vectors holds great promise in preventing AAV vector mobilization in humans. Due to almost perfect conservation with AAV5, the goat AAV origin of replication would be ineffective in the prevention of vector mobilization as Rep5 could likely rescue and replicate a goat-ITR vector. Alternately, the ITRs of the autonomous parvovirus B19 and several avian AAVs range from 400-600 bp in length and do not possess the T-shaped structure of human AAV (23). Such ITRs would further shrink the packaging capacity of AAV and may be too divergent to be encapsidated into human AAV capsids.

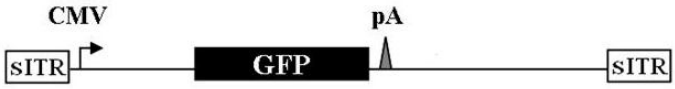
The snake AAV origin of replication overcame most of these limitations. SAAV, isolated from Python regius and Boa constrictor snakes, has been fully sequenced (24). The snake Rep proteins (sRep) are only 37% conserved with Rep2 and 34% conserved with Rep5. The snake ITR (sITR) has a highly conserved secondary structure to ITR2 and ITR5; however, there is significant divergence in the nt sequence and the size/length of critical elements for Rep specificity (Figure 23A). The RBE and RBE' sequences are reasonably conserved between serotypes, as is the hairpin spacer which separates them. This suggests that Rep2 and Rep5 may be able to bind sITR. However, there are major differences between all three serotypes in the ITR regions critical for specificity identified in Chapter 3. sITR contains an intermediate spacer length (12 nt) compared to ITR2 and ITR5. However, while Rep5 requires a partial GAGY consensus motif in its spacer, sITR has essentially no GAGY conservation in this element. This lack of a strong RBE in the spacer is likely to preclude Rep2 function on sITR. The sITR putative nicking stem is also significantly different from

that of ITR2 or ITR5, at only four nt in height and 12 nt in total length. This is likely to preclude Rep5 function which has specific requirements for both the height and sequence of the nicking stem.

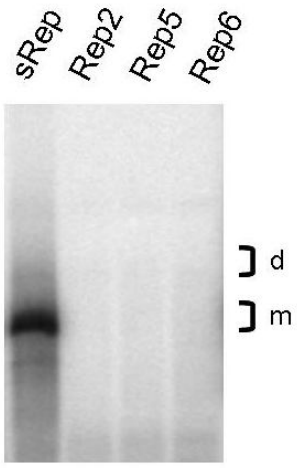
A



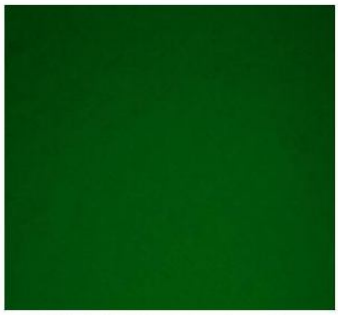
B



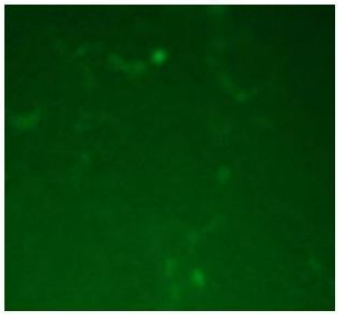
C



D



sITR-GFP+sRepCap2
+Ad Helper



sITR-GFP+sRepCap2
+Ad Helper
+Rep2Cap2

Figure 23. Snake AAV possesses a unique origin of replication with respect to characterized human/primate AAV serotypes.
(A) Comparison of sITR with ITR2 and ITR5. The putative RBE is boxed. The putative nicking stem is shown extruded, and the putative nicking site is indicated by an arrow. The length of each segment is indicated above.
(B) Schematic of sITR-GFP.
(C) Southern blot of replication of sITR-GFP by the sRepCap2, Rep2Cap2, Rep5Cap2, and Rep6Cap6 helper plasmids. Replication of a replication deficient snake vector. Monomer (m) and dimer (d) replication intermediates are indicated.
(D) sRepCap2 is unable to produce infectious vector as observed by a lack of transduction of HEK 293 cells (top). The addition of Rep2Cap2 rescues packaging of sITR-GFP producing vector which can efficiently transduce cells.

Thus, we created a GFP vector flanked by sITRs (Figure 23B). To determine whether sRep was able to encapsidate a sITR vector into a human capsid, we created a helper plasmid comprised of sRep and Cap2. Surprisingly, sRep was able to efficiently replicate the sITR-flanked vector in HEK 293 cells with the human Ad-helper plasmid (Figure 23C). As predicted, Rep2, Rep5, and Rep6 were all unable to replicate the sITR vector.

The inability of Rep2 and Rep5 to replicate sITRs confirmed their potential for preventing vector mobilization. However, efficient replication of the sITR vector by sRep was only the first step in vector production. The next step was to determine whether sRep could encapsidate the vector into Cap2. Infecting naïve 293 cells with lysate from cells transfected with sITR-GFP, sRepCap2, and Ad helper resulted in no GFP positive cells, suggesting vector was not produced (Figure 23D, top). Q-PCR results supported the conclusion that sRep failed to encapsidate the sITR vector into Cap2, as well as sCap (data not shown). A Western blot confirmed that sRepCap2 produced Cap2 as expected (data not shown). In an attempt to rescue encapsidation of the sITR vector, we added Rep2Cap2 along with sRepCap2. Lysate from this transfection contained infectious particles, effectively transducing naïve 293 cells. This results demonstrate that a chimeric SAAV-AAV2 sITR vector production system can be developed and should effectively prevent AAV vector mobilization.

C3. Novel Chimeric ITR as a rAAV Vector. The novel AAV origin of replication created in Chapter 3 has the potential to prevent AAV vector mobilization. Neither parent serotype (AAV2 and AAV5) can replicate the chimeric ITR. Thus, mobilization of this ITR would require a recombination even between Rep2 and Rep5 in order to reconstitute the chimeric Rep necessary to replicate the novel origin. First, the ability of the novel ITR

needed to be tested in a vector setting to assure functionality comparable to current vector production methods. For this purpose, this ITR5+2NS was compared to the DD forms of ITR2 and ITR5 to determine whether it could be used for vector systems.

The vector was prepared by the triple transfection of Ad-helper plasmid and either Rep2Cap2 with ITR2, Rep5Cap2 with ITR5, or Rep52aa149Cap2 with ITR5+2NS into five 15cm plates. A CsCl gradient was used to purify the vector and the titer was determined by qPCR (28). The ITR2 and ITR5 vectors were both produced with equivalent titers as previously reported (36). The ITR5+2NS vector was consistently one to two orders of magnitude lower in titer than the wt ITRs (Figure 24A). This was consistent with the data from Chapter 3 suggesting the origin may replicate less efficiently than wt.

In order to determine whether the transduction efficiency of the novel ITR was also lower than wt, we infected HEK 293 cells with equivalent viral titers of the three vectors. Cells were infected with 10,000 vg/cell, 1,000 vg/cell, and 100vg/cell of each and the relative transduction efficiency of each was assessed (Figure 24B). All three ITRs demonstrated equivalent transduction potential by this assay, verifying that ITR5+2NS vectors are as infectious as ITR2 and ITR5 vectors. This suggests that ITR5+2NS vectors are stable and their low replication efficiency is a result of the Rep-ITR interaction, not the ITR itself.

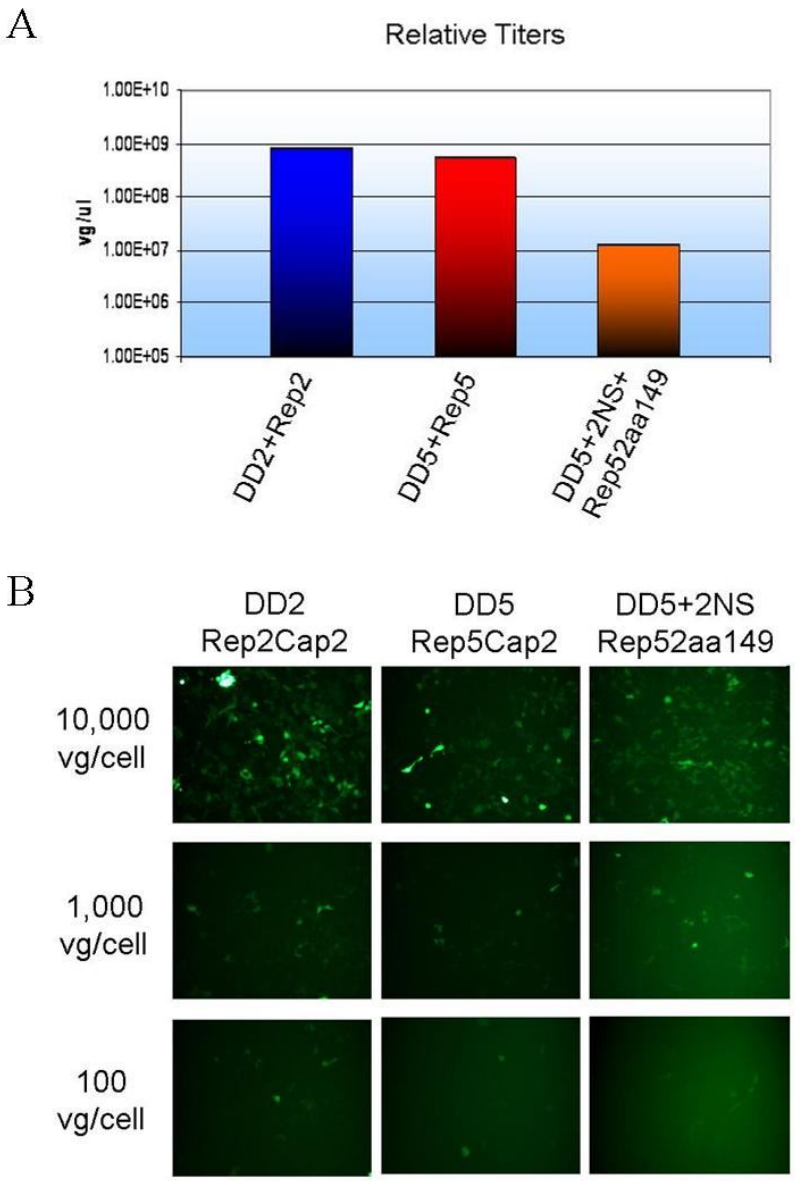


Figure 24. Vector yields and transduction potential of a novel ITR.
(A) Virus was purified via CsCl gradient and viral titer was measured by qPCR.
(B) Viral vector was diluted in order to infect HEK 293 cells with the titer indicated. The vectors were flanked by the ITR and replicated by the Rep indicated. All were encapsidated into Cap2.

C4. Prevention of AAV Vector Mobilization by a Novel Chimeric ITR. While the vector production efficiency of the novel chimeric ITR might preclude it from widespread therapeutic use, its anti-mobilization properties could make it critical to the future of gene therapy. Additionally, while the vector yields are lower, the transduction efficiency is equivalent to the current ITR2 vectors that are widely used. In order to determine whether vectors flanked by this ITR are truly at a reduced risk of vector mobilization, we used the mobilization assay developed in Chapter 2 (Figure 26).

HEK 293 cells were infected with 10,000 vg/cell of ITR2, ITR5, and ITR5+2NS flanked virus and passaged for 20 days. GFP positive cells were then sorted and pooled. The cell populations were grown out and remained 100% GFP positive. Each cell population was then transfected with Ad-helper plasmid and either no AAV helper, Rep2Cap2, Rep5Cap2, or Rep52aa149Cap2. Hirt DNA was isolated 48 hours later and 10 ul was used for qPCR with a primer set corresponding to the GFP coding sequence. Additionally, crude lysate was used to infect naïve 293 cells.

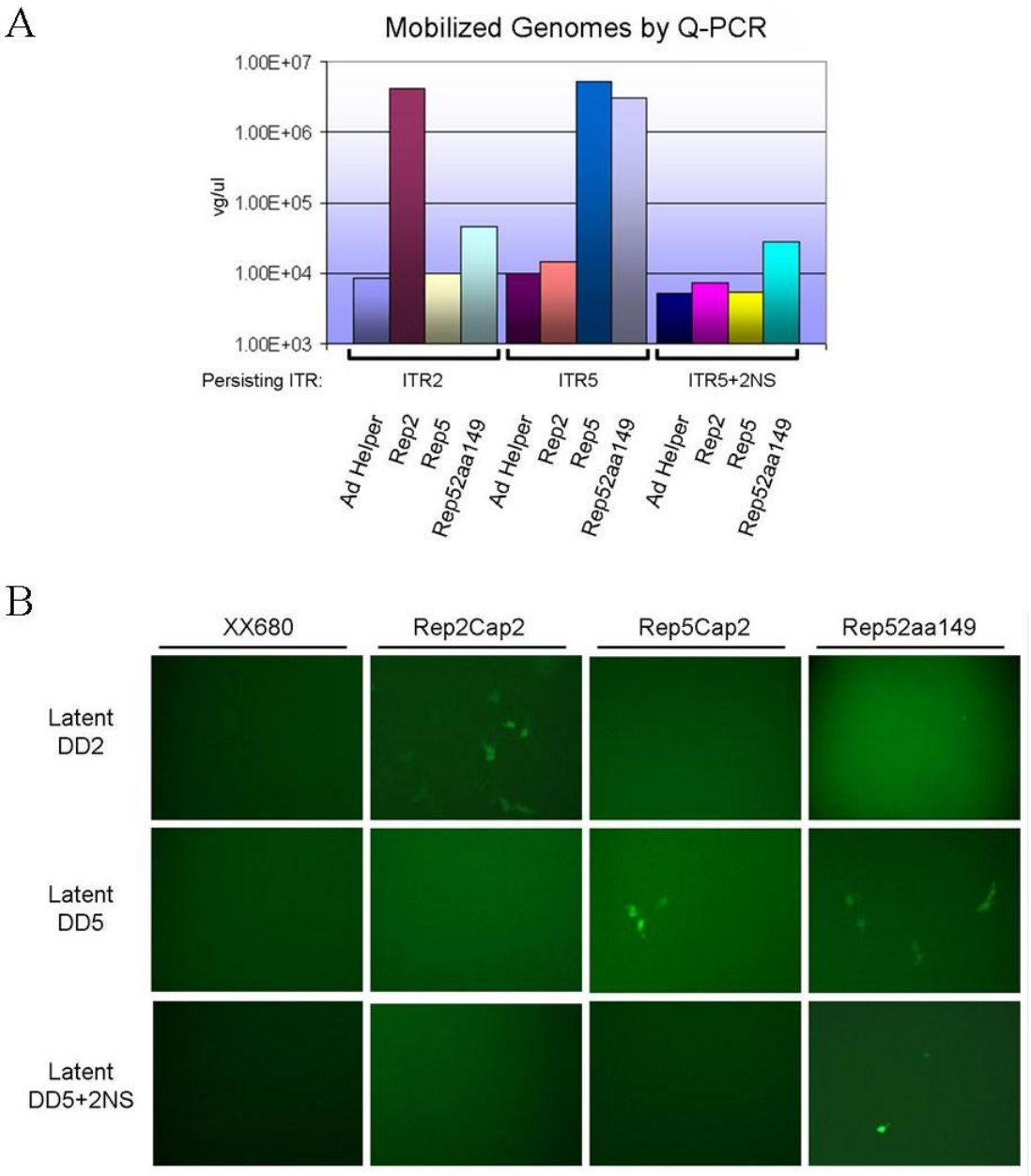


Figure 25. Specificity of the novel recombinant origin of replication extends to vector mobilization.
(A) qPCR analysis of mobilized ITR2, ITR5, and ITR5+2NS GFP genomes from Hirt DNA isolated from cells mobilized with Ad-helper plasmid and the AAV helper indicated.
(B) Lysate taken from cells containing mobilized genomes from (A) was used to infect naïve 293 cells. Pictures were taken 48 hours after infection.

The ITR2 flanked genomes were efficiently mobilized by Rep2 (Figure 25A). Rep5 resulted in no mobilization compared to background (Ad-helper only). Rep52aa149 gave modest mobilization of the ITR2 flanked transgene. The decrease in mobilization by Rep52aa149 with respect to Rep2 was supported by its ability to replicate an ITR2 flanked transgene with approximately 60% of Rep2 efficiency (Figure 17E). The ITR5 flanked genomes were efficiently mobilized only by Rep5 and Rep52aa149. The ITR5+2NS flanked genome could be mobilized only by Rep52aa149, supporting its possible utility as an anti-mobilization vector. The transgene was not mobilized efficiently, likely due to the inability of Rep52aa149 to replicate it with wt efficiency. However, this decrease in mobilization efficiency would be beneficial when creating a mobilization resistant vector. While ITR5+2NS serves as an anti-mobilization vector, Rep52aa149 serves as a universal mobilization protein as it was able to mobilize ITR2, ITR5, and ITR5+2NS genomes. However, even if this Rep was reconstituted in nature, it would be unable to mobilize ITR5+2NS flanked vector genomes as efficiently as ITR2 or ITR5 flanked vectors. Transduction of mobilized vectors into naïve 293 cells supported the qPCR data (Figure 25B). Again, only Rep52aa149 was able to mobilize ITR5+2NS vector genomes. Rep52aa149 was also able to mobilize ITR2 and ITR5 vectors, though mobilization of ITR2 and ITR5+2NS genomes was inefficient, resulting in very few visible GFP positive cells.

4D. Discussion

The assays developed in Chapter 2 yielded the tools to test novel vectors which might prevent AAV vector mobilization in humans. While conceptually well-founded, the use of non-homologous ITR vectors was obviated by the nature of the ITR itself. The inability of the Rep2 or Rep5 to work on the ITR of the opposite serotype has been well established

previously and in this work (16). That Rep2 or Rep5 alone can replicate these vectors efficiently suggests that they are actively being converted into homologous ITR vectors upon transfection. This is almost certainly due to the ability of one ITR to serve as a template to repair the opposite one (79). This is so efficient that the use of a chimeric Rep capable of replicating each ITR type only increased the amount of replicated genomes by 2-fold. This suggests that of the genomes replicated by the chimeric Rep, half have likely undergone ITR repair to become homologous ITR vectors.

Due to these considerations, the ability of a single Rep to mobilize these genomes from cells was not surprising. What is quite striking, however, is the rate at which one ITR can be used to repair/replace the other. This has likely arisen out of necessity due to the latent phase which AAV must enter in the absence of helper virus. Either mode of AAV genomic persistence (circularization/concatemerization and chromosomal integration) requires that the ITR undergo recombination. This often results in partial deletion of the ITR, especially with respect to chromosomal integration thought to occur using homologous recombination through microhomologies anywhere in the ITR (75). Thus, the ability of AAV to repair itself as it enters the lytic cycle is critical for the survival of the virus. It is possible that the ability to repair the ITR is the selective force which has led to the evolution of such a recombinogenic structure.

The use of non-human AAV ITRs to prevent vector mobilization posed a number of obstacles. First, the Rep protein needed to be efficiently expressed in a functional state in human HEK 293 cells. Additionally, due to the lack of a species-specific Ad-helper, the viral genes needed to be efficiently complemented by human Ad. The ITR needed to be divergent enough from the human/primate ITRs so that the human Reps could not catalyze its

replication; however, the ITR needed to be conserved enough to be packaged into human/primate Cap required for current rAAV vectors.

By these criteria, SAAV shows enormous promise in the prevention of vector mobilization. The virus can efficiently replicate in human-derived cells with human-derived Ad-helper genes, despite its reptilian lineage and poor homology to primate AAVs. Critically, the sITR can not be replicated by Rep2, Rep5, or Rep6. Unfortunately, sRep is unable to package an sITR vector into Cap2 or sCap. The inability of sRep to encapsidate the vector into its cognate Cap suggests either a mutation in the sRep or sCap construct or that SAAV encapsidation needs to occur in reptilian cells or in the presence of Snake Ad. However, the addition of Rep2 rescued sITR encapsidation. It is unclear whether Rep2 is interacting with the SAAV replication machinery or is binding and encapsidating replicated ss sITR-flanked genomes. In either case, we now possess a sITR vector production system ready for optimization and incorporation into rAAV applications.

Perhaps the ultimate method for preventing vector mobilization comes not from optimizing current AAV origins of replication but from creating entirely new Rep-ITR interactions. While the use of a non-human/primate AAV will always carry the risk of that same non-human serotype jumping into our species, a novel recombinant ITR would possibly be unable to be replicated by any naturally occurring AAV in existence. Chapter 3 describes the creation of such a novel origin of replication.

Unfortunately, while the novel origin is possesses a unique Rep-ITR interaction with respect to its parent serotypes, it also lacks the replicative efficacy of its parents. Vector preparations using this ITR were consistently one to two orders of magnitude lower in titer than ITR2 or ITR5 vectors. In a therapeutic setting where current ITR2 vectors are often

limited by the titer that can be achieved, 10-100 times lower titer would prevent the use of any vector. However, the recombinant ITR displays the same transduction efficiency as ITR2 or ITR5 vectors. This demonstrates that the rate limiting step for ITR5+2NS is vector production efficiency, and if that shortcoming be addressed, it could likely serve as the ITR in all future vector applications.

The same decreased interaction between the recombinant ITR and recombinant Rep also explains the low degree of vector mobilization of this ITR seen in Figure 25A. Only Rep52aa149 can mobilize ITR5+2NS at any level. Even then, rescue and replication of the vector is so weak that even if a recombination event in nature produced a Rep52aa149-like Rep, the chance of it leading to sustained mobilization of the vector is exceedingly low. Between the promise of this vector and the potential shown by non-human ITRs, the risk of vector mobilization can be eliminated by the time rAAV vectors are ready to progress beyond clinical trials.

Chapter 5: Future Directions

Future Directions

The number of individuals in the population carrying a rAAV vector stands to increase dramatically as AAV-mediated gene delivery progresses further into the clinic. Current clinical trials have seen rAAV vectors administered to nearly every major organ in the body. As these treatments become mainstream, the risk of these ITR2 flanked vectors being mobilized by an Ad/AAV infection will only grow.

By characterizing the AAV origin of replication, we have taken the first steps towards circumventing the risks posed by vector mobilization. Determining the mechanisms of DNA-protein specificity in the Rep-ITR interaction has also unlocked new questions about the AAV origin of replication. The only human/primate AAV serotype (apart from AAV5) with significant divergence in Rep homology from Rep2 is AAV8. Chimeric Rep8-Rep2 proteins failed to replicate ITR2 or ITR5, suggesting AAV8 may have a novel origin of replication with respect to the other AAV serotypes. The determinants of this specificity should be readily determined upon the publication of the ITR8 sequence. Rep8-ITR8 specificity will likely exist within the spacer/nicking stem regions of the ITR, just as it does in AAV2 and AAV5. However, the AAV8 origin may possess a unique Rep-ITR interaction outside this section of the ITR, potentially deriving specificity from the RBE'.

Further biochemical and structural studies are needed to define the individual determinants of Rep-ITR specificity. Several residues seem particularly critical in this regard. Such studies are necessary to identify the specific Rep interactions with both the nicking stem and spacer of the ITR. Additional studies are also needed to explore the function of Rep region 1. This region possesses interactions with both the RBE and RBE' of the ITR. Because a single Rep molecule can not be bound to RBE, the RBE', and interact with the

nicking stem, it is unclear why the RBE' portion of Rep region 1 impacts the RBE binding and nicking function of the protein. Nonetheless, in order for any Rep to function on any ITR, the RBE and RBE' interacting domains of region 1 must originate from the same serotype. As Rep is able to function on ITRs lacking the RBE' and hairpins entirely, this observation is likely due to proper folding of the Rep protein (16).

Additionally, this work has implications with respect to AAV site-specific integration. AAV2 can direct integration of its genomes into the AAVS1 site of human chromosome 19. This occurs because the chromosome contains a RBE and a trs on the chromosome which can be nicked by Rep2 (80, 97). Rep2 is believed to nick and unwind the chromosome at the trs, resulting in the recruitment of host repair machinery. AAV genomes are then inserted into the chromosome through mini-homologies within the ITR via the homologous recombination pathway. The ability of other AAV serotypes to integrate into the AAVS1 site has not been characterized. However, high conservation between AAV2 and AAV serotypes 1-4, 6, 7, and 9-11 suggests that these serotypes should be capable of targeted integration into the same site.

Attempts to discover a human chromosomal integration locus for AAV5 have been unsuccessful. It is likely that the determinants of Rep5-ITR5 specificity identified in Chapter 3, specifically the sensitivity to changes in the nicking stem, are responsible for the absence of Rep5 nicking sites in the human chromosome. However, having identified these requirements, it should be possible to engineer an AAV5 targeted integration site into human cell lines. The ability to achieve differential Rep-mediated targeted integration of a vector into different sites dependent on the Rep supplied would impact therapeutic endeavors as well as biotechnological methods. It is also possible that the chromosomes of non-human

organisms which harbor closely related AAV serotypes (goat, bovine), possess AAV5 integration sites. While the likelihood of identifying such a site is unlikely through homology searches (AAVS1 is poorly conserved with ITR2), directed integration studies with goat or bovine cell lines may yield a novel site for site-specific AAV5 integration.

Initial data indicates that SAAV holds the potential to eliminate the risk of vector mobilization. The sRep-sITR interaction is unique with respect to human AAV serotypes, yet SAAV vectors can be replicated in human cells with human Ad-helper. It is unknown whether sRep can encapsidate sITR vectors directly into human/primate capsids, or whether a chimeric Rep possessing the sRep N-terminus and Rep2 C-terminus will be required. The relative vector yields of a SAAV production system will need to be comparable to those of the current AAV2 based vectors. Additionally, the relative integration efficiency of sITR compared to ITR2 or ITR5 must be determined to insure that SAAV vectors are safe. Also, because the SAAV origin of replication is unique with respect to AAV2 and AAV5, the existence of a possible SAAV site-specific integration locus in the human chromosome must be determined. Finally, the mobilization assays described in Chapter 2 must be repeated to demonstrate the increased safety potential of sITRs.

Perhaps the most significant result of this work was the creation of a novel AAV origin of replication. This recombinant ITR provides the greatest means to eliminate the risk of vector mobilization. However, several steps lie between it and the clinic. The depressed vector yields achieved with this ITR must be improved. Because the use therapeutic rAAV vectors is often limited by titer, a log-order decrease in the titer of any vector with respect to an ITR2 vector would make it impractical for clinical purposes. Directed evolution or

rational mutagenesis of this novel ITR is necessary in order to generate an origin of replication with wt AAV levels of replicative fecundity.

If the vector yields can be increased, this novel ITR must be fully characterized before it can enter the clinic. We have already demonstrated that the relative transduction efficiency, adjusted for titer, of this ITR is equivalent to ITR2 and ITR5. The random integration potential of this ITR must be explored and shown to be equivalent to its parent serotypes, AAV2 and AAV5. The ability of this ITR to undergo site-specific integration should also be explored. The chimeric Rep specific to this ITR also functions on ITR2 (albeit at a lower than wt efficiency) and should still recognize the AAVS1 site of chromosome 19. Thus, this novel origin may still undergo site-specific integration into AAVS1 at some level. Additionally, Rep52aa146 may be able to direct rAAV vectors into a separate site in the chromosome. This chimeric Rep possesses AAV2-like flexibility toward nicking stem substrates as well as the binding promiscuity of Rep5. Thus, it is entirely possible that a location for site-specific integration of ITR5+2NS already exists in the human chromosome.

The driving force behind these experiments was the clinical risk of vector mobilization. Of all the methods tested to prevent mobilization, the most successful came out of studying the basic biology of AAV. In that way, this work straddled both the clinical and virological aspects of AAV. Future experiments are likely to walk a similar path. Through its evolution, AAV has been faced with many of the same challenges which complicate the use of rAAV vectors today. Studying nature's solutions stands to have a greater impact on the future of AAV-mediated gene delivery than any rational engineering approach.

Appendix 1: Chimeric and Mutant Rep Proteins Created

Name	Description	Replicates ITR2	Replicates ITR5	Confirmed by Western
Rep52aa73		Yes	No	Yes
Rep52aa84	Described in Chapter 3	Yes	No	Yes
Rep52aa110	Described in Chapter 3	Yes	No	Yes
Rep52aa126	Described in Chapter 3	No	No	Yes
Rep52aa138	Described in Chapter 3	No	No	Yes
Rep52aa146	Described in Chapter 3	Yes	Yes	Yes
Rep52aa147	Described in Chapter 3	Yes	Yes	Yes
Rep52aa151	Described in Chapter 3	Yes	Yes	Yes
Rep52aa156		No	Yes	Yes
Rep52aa160	Described in Chapter 3	No	Yes	Yes
Rep52aa175	Described in Chapter 3	No	Yes	Yes
Rep52aa207	Described in Chapter 3	No	Yes	Yes
Rep25aa73		No	Yes	Yes
Rep25aa77	Described in Chapter 3	No	Yes	Yes
Rep25aa89		No	Yes	Yes
Rep25aa97	Described in Chapter 3	No	Yes	Yes
Rep25aa116	Described in Chapter 3	No	Yes (low)	Yes
Rep25aa125	Described in Chapter 3	No	Yes (low)	Yes
Rep25aa136		No	Yes (low)	No
Rep25aa141	Described in Chapter 3	Yes	Yes (low)	Yes
Rep25aa149	Described in Chapter 3	Yes	No	Yes
Rep25aa166	Described in Chapter 3	Yes	No	Yes
Rep25aa187		Yes	No	No
Rep25aa216	Described in Chapter 3	Yes	No	Yes
Rep25aa244		Yes	No	Yes
Rep25aa101		No	No	No
Rep25aa116K100S	K100S does not rescue Rep25aa116	No	Yes (low)	No
Rep2 K72A	Contradicts previous findings (93)	Yes	No	Yes
Rep5 K72A		No	Yes	Yes
Rep2 C179A	Determining importance of Rep2 C179	No	No	No
Rep52aa146 C179A	Determining importance of Rep2 C179	Yes	Yes	No
Rep525aa110-148	Described in Chapter 3	No	No	Yes
Rep525aa146-187	Described in Chapter 3	Yes	No	Yes
Rep525aa110-187	Described in Chapter 3	Yes	No	Yes
Rep525aa136-146		No	No	No
Rep252aa97-146	Described in Chapter 3	Yes	No	Yes
Rep252aa149-187	Described in Chapter 3	No	No	Yes
Rep252aa97-187	Described in Chapter 3	No	Yes	Yes
Rep252aa136-141		Yes	No	No
Rep82aa244	Described in Chapter 3	No	No	Yes
Rep282aa113-163	Described in Chapter 3	No	No	Yes
Rep85aa163		No	No	No

Appendix 2: Chimeric and Mutant ITRs Created

Name	Description	Replicated by Rep2	Replicated by Rep5
ITR2	ITR2 in DD format w/ SfiI site for cloning	Yes	No
ITR5	ITR5 in DD format w/ SfiI site for cloning	No	Yes
ITR5+2NS	ITR5 with ITR2 nicking stem - Replicated only by Rep52aa149	No	No
ITR2+5NS	ITR2 with ITR5 nicking stem	Yes	Yes
ITR2+7	ITR2 with 7bp spacer insertion	No	No
ITR5+6	ITR5 with 6bp spacer insertion	No	Yes
ITR5+15	ITR5 with 15 bp insertion in the spacer	No	No
ITR2-TA	ITR2 with T-A bp deleted from top of nicking stem	Yes	Yes
ITR5+2SNS	ITR5 with ITR2 spacer and nicking stem	No	No
ITR5+2S	ITR5 with ITR2 spacer	Yes	Yes
ITR5+TA	ITR5 with T-A bp insertion at top of nicking stem	No	Yes
ITR2 Crowfoot	ITR2 with extra RBE hairpin (for a total of three)	Yes	No
S1 ITR	ITR2 with the RBE, spacer, and nicking stem of the AAVS1 site on Chromosome 19	Yes	No
ITR5-GC	ITR5 with G-C bp deleted from base of nicking stem	No	Yes
ITR2+GC	ITR2 with G-C bp inserted into base of nicking stem	Yes	No
ITR5+2SNS-GGG	ITR5 with ITR2 nicking stem and spacer with a three nucleotide deletion from the hairpin spacer	Yes	No
ITR2 No Stem	ITR2 with the nicking stem deleted	No	No
ITR2 Reconstituted Stem	ITR2 with nicking stem but mutant trs sequence	No	No
ITR5 No Stem	ITR5 with the nicking stem deleted	No	No
ITR5 Reconstituted Stem	ITR5 with nicking stem but mutant trs sequence	No	Yes (low)
S1 ITR No hairpins	AAVS1 P1 element cloned with an SfiI site and no RBE'	Yes	No
ITR2-2ntRBE	ITR2 with spacer and first two nt of the RBE deleted which supports AAVS1 nicking stem	Yes	No
ITR2+5SNS	ITR2 with an ITR5 spacer and nicking stem	Yes	No
ITR2+5SNS+GGG	ITR2 with an ITR5 spacer and nicking stem and three nt (GGG) inserted into the hairpin spacer	No	No
ITR5 No GAGY	ITR5 with the GAGY from the spacer removed	No	No
ITR5 GAGY	ITR5 with perfect GAGY homology in the spacer	Yes	Yes
ITR5 Spacer RBE	ITR5 with GAGY consensus in the 12nt of the RBE nearest the hairpins removed	No	No
ITR2 +8 GAGY	ITR2 with 8nt spacer insertion with perfect GAGY homology	No	No
ITR2 +8-8 GAGY	ITR2 with 8nt spacer insertion with perfect GAGY homology and 8nt of GAGY homology removed from the RBE nearest the hairpins	Yes	No
ITR5 No Hairpins	ITR5 with no hairpins (for RBE')	No	No
ITR2-GC	ITR2 with one bp removed from the base of the stem	Yes	No
ITR2 5bp	ITR2 with 5bp tall nicking stem and conserved trs	Yes	No
ITR2 11bp	ITR2 with 11bp nicking stem and conserved trs	No	No
ITR5 5bp	ITR5 with 5bp tall nicking stem (deletions from nicking stem base)	No	Yes

ITR5 9bp	ITR5 with 9bp tall nicking stem (insertions at base of stem)	No	Yes
ITR5+GC	ITR5 with G-C bp inserted into the base of nicking stem	No	Yes
ITR5+2NS-GC	ITR5 with an ITR2 nicking stem with a GC bp deleted from the base of the nicking stem	No	No

References

1. **Afione, S. A., C. K. Conrad, W. G. Kearns, S. Chunduru, R. Adams, T. C. Reynolds, W. B. Guggino, G. R. Cutting, B. J. Carter, and T. R. Flotte.** 1996. In vivo model of adeno-associated virus vector persistence and rescue. J. Virol. **70:**3235-3241.

2. **Atchinson, R. W., Castro, B.C., Hammon, W.M.** 1965. Adenovirus-Associated Defective Virus Particles. Science. 13;149:754-6.

3. **Bantel-Schaal U, zur Hausen H.** 2004. Characterization of the DNA of a defective human parvovirus isolated from a genital site. Nat Genet. Jul;36(7):767-73.

4. **Berns, K. I. a. C. R. P.** 2007. Parvoviridae, p. 2437-2477. *In* D. M. K. a. P. M. Howley (ed.), Fields virology, 5th ed. Lippincott Williams & Wilkins, New York, NY.

5. **Berns, K.I., Adler, S.** 1972. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. J. Virol. *5*, 693-69

6. **Blacklow, N. R., M. D. Hoggan, and W. P. Rowe.** 1968. Serologic evidence for human infection with adenovirus-associated viruses. J Natl Cancer Inst **40:**319-27.

7. **Blankinship, M. J., et al.** 2004. Efficient transduction of skeletal muscle using vectors based on adeno-associated virus serotype 6. Mol. Ther. 10: 671–678.

8. **Bleker, S., M. Pawlita, and J. A. Kleinschmidt.** 2006. Impact of capsid conformation and Rep-capsid interactions on adeno-associated virus type 2 genome packaging. J. Virol. 80:810-820.

9. **Brister, J.R., Muzyczka, N.** 1999. Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. J. Virol. *73*, 9325-36.

10. **Brister, J.R., Muzyczka, N.** 2000. Mechanism of Rep-mediated adeno-associated virus origin nicking. J Virol. *74*, 7762-71.

11. **Chang, X. B., and J. H. Wilson.** 1986. Formation of deletions after initiation of simian virus 40 replication: influence of packaging limit of the capsid. J Virol 58:393-401.

12. **Chang, L. S., and T. Shenk.** 1990. The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. J Virol. 64:2103-9.

13. **Chang, L. S., Y. Shi, and T. Shenk.** 1989. Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. J Virol 63:3479-88.

14. **Cheung, A. K., M. D. Hoggan, W. W. Hauswirth, and K. I. Berns.** 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J Virol **33:**739-48.

15. **Chiorini, J. A., F. Kim, L. Yang, and R. M. Kotin.** 1999. Cloning and characterization of adeno-associated virus type 5. J. Virol. *73*, 1309-1319.

16. **Chiorini, J. A., S. Afione, and R. M. Kotin.** 1999. Adeno-associated virus (AAV) type 5 Rep protein cleaves a unique terminal resolution site compared with other AAV serotypes. J Virol **73:**4293-8.

17. **Choi, V. W., D. M. McCarty, and R. J. Samulski.** 2006. Host cell DNA repair pathways in adeno-associated viral genome processing. J Virol **80:**10346-56.

18. **Choi, V. W., R. J. Samulski, and D. M. McCarty.** 2005. Effects of adenoassociated virus DNA hairpin structure on recombination. J. Virol. **79:**6801-6807.

19. **Chomczynski, P.** 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. Anal. Biochem. *201,* 134.

20. **Collaco, R., V. Kalman-Maltese, A. D. Smith, J. D. Dignam, and J. P. Trempe**. 2003. A biochemical characterization of the adeno-associated virus Rep40 helicase. J. Biol. Chem. 278:34011-34017.

21. **Davis, M. D., J. Wu, and R. A. Owens.** 2000. Mutational analysis of adenoassociated virus type 2 Rep68 protein endonuclease activity on partially singlestranded substrates. J. Virol. **74:**2936-2942.

22. **Deiss, V., J. D. Tratschin, M. Weitz, and G. Siegl.** 1990. Cloning of the human parvovirus B19 genome and structural analysis of its palindromic termini. Virology 175:247-254.

23. **Duan, D., P. Sharma, J. Yang, Y. Yue, L. Dudus, Y. Zhang, K. J. Fisher, and J. F. Engelhardt.** 1998. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. J. Virol. 72:8568-8577.

24. **Farkas, S. L., Z. Zadori, M. Benko, S. Essbauer, B. Harrach, and P. Tijssen.** 2004. A parvovirus isolated from royal python (Python regius) is a member of the genus Dependovirus. J. Gen. Virol. *85*, 555-561.

25. **Gao, G. P., M. R. Alvira, L. Wang, R. Calcedo, J. Johnston, and J. M. Wilson.** 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. Proc. Natl. Acad. Sci. USA 99:11854-11859. Virology. 1984 Apr 15;134(1):52-63.

26. **Girod, A., C. E. Wobus, Z. Zadori, M. Ried, K. Leike, P. Tijssen, J. A. Kleinschmidt, and M. Hallek.** 2002. The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. J Gen Virol 83:973-8.

27. **Grieger, J. C., and R. J. Samulski.** 2005. Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. J. Virol. 79:9933-9944.

28. **Grieger, J. C., V. W. Choi, and R. J. Samulski.** 2006. Production and characterization of adeno-associated viral vectors. Nat Protoc **1:**1412-28.

29. **Grimm, D., et al.** 2003. Preclinical in vivo evaluation of pseudotyped adeno-associated virus vectors for liver gene therapy. Blood 102: 2412–2419.

30. **Grimm, D., K. Pandey, H. Nakai, T. A. Storm, and M. A. Kay.** 2006. Liver transduction with recombinant adeno-associated virus is primarily restricted by capsid serotype not vector genotype. J. Virol. **80:**426-439.

31. **Grimm, D., M. A. Kay, and J. A. Kleinschmidt.** 2003. Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. Mol. Ther. **7:**839–850.

32. **Flotte, T. R., S. A. Afione, R. Solow, M. L. Drumm, D. Markakis, W. B. Guggino, P. L. Zeitlin, and B. J. Carter.** 1993. Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. J. Biol. Chem. **268:**3781–3790.

33. **Halbert, C.L., Miller, A.D., Mcnamara, S., Emerson, J.,Gibson, R.L., Ramsey, B., and Aitken, M.L.** 2006. Prevalence of neutralizing antibodies against adeno-associated virus (AAV) types 2, 5, and 6 in cystic fibrosis and normal populations: Implications for gene therapy using AAV vectors. Hum. Gene Ther. **17,** 440–447.

34. **Hauswirth, W.W., and Berns, K.I.** 1977. Origin and termination of adeno-associated virus DNA replication. Virology. *78*, 488-499.

35. **Hermonat, P.L., Batchu, R.B.** 1997. The adeno-associated virus Rep78 major regulatory protein forms multimeric complexes and the domain for this activity is contained within the carboxy-half of the molecule. FEBS Lett. *20*, 180-4.

36. **Hewitt, F.C., Li, C., Gray, S.J., Cockrell, S., Washburn, M., Samulski, R.J.** 2009. Reducing the Risk of AAV Vector Mobilization with AAV5 Vectors. J. Virol. *83*, 3919-29

37. **Hickman, A.B., Ronning, D.R., Kotin, R.M., Dyda, F.** 2002**.** Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. Mol. Cell. *10*, 327-37.

38. **Hickman, A.B., Ronning, D.R., Perez, Z.N., Kotin, R.M., Dyda, F.** 2004**.** The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces. Mol. Cell. *13*, 403-14.

39. **Higuchi, R., Krummel, B., and R. K. Saiki.** 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. *16*, 7351-7367.

40. **Hildinger, M., A. Auricchio, G. Gao, L. Wang, N. Chirmule, and J. M. Wilson.** 2001. Hybrid vectors based on adeno-associated virus serotypes 2 and 5 for muscle-directed gene transfer. J Virol **75:**6199-203. 28

41. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J Mol Biol **26:**365-9.

42. **Ilyina, T.V., and Koonin, E.V.** 1992. Conserved sequence motifs in binding domain of the Epstein-Barr virus origin-binding protein, the initiator proteins for rolling circle DNA replication encoded by EBNA1, bound to DNA. Cell 84, 791–800.

43. **King, J.A., Dubielzig, R., Grimm, D., and J. A. Kleinschmidt.** 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. EMBO J. *20*, 3282-3291.

44. **Koerber, J.T., Jae-Hyung Jang, J., and D. V. Schaffer**. 2008. DNA Shuffling of Adeno-associated Virus Yields Functionally Diverse Viral Progeny. Mol Ther. 16 10, 1703–1709

45. **Kotin, R. M., and K. I. Berns.** 1989. Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. Virology 170:460-7.

46. **Kwon, I., and D. V. Schaffer.** 2007. Designer gene delivery vectors: molecular engineering and evolution of adeno-associated viral vectors for enhanced gene transfer. Pharm. Res.doi:10.1007/s11095-007-9431-0.

47. **Kyostio, S. R., R. S. Wonderling, and R. A. Owens.** 1995. Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. J. Virol. 69:6787-6796.

48. **Janik, J. E., M. M. Huston, and J. A. Rose.** 1981. Locations of adenovirus genes required for the replication of adenovirus-associated virus. Proc Natl Acad Sci U S A 78:1925-9.

49. **Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J. and Higgins D.G.** 2007. ClustalW and ClustalX version 2. Bioinformatics 2007 23(21): 2947-2948.

50. **Laughlin, C. A., C. B. Cardellichio, and H. C. Coon.** 1986. Latent infection of KB cells with adeno-associated virus type 2. J Virol 60:515-24.

51. **Laughlin, C. A., N. Jones, and B. J. Carter.** 1982. Effect of deletions in adenovirus early region 1 genes upon replication of adeno-associated virus. J Virol 41:868-76.

52. **Li, C. and R. J. Samulski**. 2005. Serotype-specific replicating AAV helper constructs increase recombinant AAV type 2 vector production. Virology. 25;335(1):10-21.

53. **Li, W, Asokan, A, Wu, Z, Van Dyke, T, Diprimio, N, J, SJ *et al.*** 2008. Engineering and selection of shuffled AAV genomes: a new strategy for producing targeted biological nanoparticles. Mol Ther **16**:1252–1260.

54. **Li, W., et al.** 2008. Engineering and selection of shuffled AAV genomes: A new strategy for producing targeted biological nanoparticles. Mol Ther.;16:1252–1260.

55. **Li, Z. L., R. Brister, D.-S. Im, and N. Muzyczka.** 2003. Characterization of the adenoassociated virus Rep protein complex formed on the viral origin of DNA replication. Virology 313:364-376.

56. **Liu Y, Okada T, Sheykholeslami K, Shimazaki K, Nomoto T, Muramatsu S, Kanazawa T, Takeuchi K, Ajalli R, Mizukami H, Kume A, Ichimura K, Ozawa K.** 2005. Specific and efficient transduction of Cochlear inner hair cells with recombinant adeno-associated virus type 3 vector. Mol Ther. Oct;12(4):725-33.

57. **Lotery AJ, Yang GS, Mullins RF, Russell SR, Schmidt M, Stone EM, Lindbloom JD, Chiorini JA, Kotin RM, Davidson BL.** 2003. Adeno-associated virus type 5: transduction efficiency and cell-type specificity in the primate retina. Hum Gene Ther. Nov 20;14(17):1663-71.

58. **Lusby, E., K. H. Fife, and K. I. Berns.** 1980. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. J Virol 34:402-9.

59. **Mayor, H. D**, **Drake, S**, **Stahmann, J**, **Mumford, D. M.** 1976. Antibodies to adeno-associated satellite virus and herpes simplex in sera from cancer patients

and normal adults. Am J Obstet Gynecol. 1;126(1):100-4.

60. **Mastrangelo IA, Hough PV, Wall JS, Dodson M, Dean FB, Hurwitz J.** 1989. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. Nature. Apr 20;338(6217):658-62.

61. **Matsushita T, Okada T, Inaba T, Mizukami H, Ozawa K, Colosi, P**. 2004. The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production. J Gen Virol, 85(Pt 8):2209-2214.

62. **McCarty DM, Fu H, Monahan PE, Toulson CE, Naik P, Samulski RJ.** 2003. Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. Gene Ther. 10(26):2112-8.

63. **McCarty, D. M., S. M. Young, Jr., and R. J. Samulski.** 2004. Integration of adeno-associated virus (AAV) and recombinant AAV vectors. Annu Rev Genet **38:**819-45.

64. **McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka.** 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. J Virol. 62:1963-73.

65. **Mehrle S., Rohde V., Schlehofer J.R.** 2004. Evidence of chromosomal integration of AAV DNA in human testis tissue. Virus Genes. Jan;28(1):61-9.

66. **Miller DG, Petek LM, Russell DW.** 2004. Adeno-associated virus vectors integrate at chromosome breakage sites. Nature Genetics. 36:767 - 773

67. **Ni, T.H., Zhou, X., McCarty, D.M., Zolotukhin, I., Muzyczka, N.** 1994. In vitro replication of adeno-associated virus DNA. J. Virol. *68*, 1128-38.

68. **Penaud-Budloo, M., Le Guiner, C., Nowrouzi, A., Toromanoff, A., Chérel Y., Chenuaud, P., Schmidt, M., von Kalle, C., Rolling, F., Moullier, P., Snyder, R.O.** 2008. Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. J Virol. Aug;82(16):7875-85.

69. **Pereira, D. J., D. M. McCarty, and N. Muzyczka.** 1997. The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. J. Virol. 71:1079-1088.

70. **Prasad, K.M., Trempe, J.P.** 1995. The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. Virology. *214*, 360-70.

71. **Rabinowitz, J. E., F. Rolling, C. Li, H. Conrath, W. Xiao, X. Xiao, and R. J. Samulski.** 2002. Cross-packaging of a single adeno-associated virus (AAV) type 2 vector genome into multiple AAV serotypes enables transduction with broad

specificity. J Virol **76:**791-801.

72. **Richardson, W. D., and H. Westphal.** 1984**.** Requirement for either early region 1a or early region 1b adenovirus gene products in the helper effect for adeno-associated virus. J Virol 51:404-10.

73. **Rose, J. A., M. D. Hoggan, and A. J. Shatkin.** 1966. Nucleic acid from an adeno-associated virus: chemical and physical studies. Proc. Natl. Acad. Sci. USA 56:86-92.

74. **Rutledge, E. A., and D. W. Russell.** 1997. Adeno-associated virus vector integration junctions. J. Virol. 71:8429-8436.

75. **Rutledge, E. A., C. L. Halbert, and D. W. Russell.** 1998. Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. J. Virol. **72:**309-319.

76. **Ryan J.H., Zolotukhin S., and Muzyczka, N.** 1996. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. J Virol. *70*, 1542-1553.

77. **Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka.** 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. USA **79:**2077-2081.

78. **Samulski, R. J., and T. Shenk.** 1988. Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. J Virol 62:206-10.

79. **Samulski, R.J., Srivastava, A., Berns, K.I. and Muzyczka, N.** 1983. Rescue of adeno-associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. Cell. **33**:135–143

80. **Samulski, R. J., X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter.** 1991. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. Embo J **10:**3941-50.29

81. **SantaLucia, J.,Jr.** 1998. A unified view of polymer, dumbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc. Natl Acad. Sci. *95*, 1460–1465.

82. **Schmidt, M., E. Grot, P. Cervenka, S. Wainer, C. Buck, and J. A. Chiorini.** 2006. Identification and characterization of novel adeno-associated virus isolates in ATCC virus stocks. J. Virol. **80:**5082-5085.

83. **Schnepp, B. C., K. R. Clark, D. L. Klemanski, C. A. Pacak, and P. R. Johnson.** 2003. Genetic fate of recombinant adeno-associated virus vector

genomes in muscle. J. Virol. **77:**3495-3504.32.

84. **Smelkova, N. V., and J. A. Borowiec.** 1997. Dimerization of simian virus 40 T-antigen hexamers activates T-antigen DNA helicase activity. J. Virol. 71:8766-8773.

85. **Smith, R. H., A. J. Spano, and R. M. Kotin.** 1997. The Rep78 gene product of adeno-associated virus (AAV) self-associates to form a hexameric complex in the presence of AAV ori sequences. J. Virol. 71:4461-4471.

86. **Smith, R. H., and R. M. Kotin.** 2000. An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. J. Virol. 74:3122-3129.

87. **Sonntag, F., S. Bleker, B. Leuchs, R. Fischer, and J. A. Kleinschmidt.** 2006. AAV2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. J. Virol. 80:11040-11054.

88. **Srivastava, A., E. W. Lusby, and K. I. Berns.** 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. J Virol 45:555-64.

89. **Summerford, C., and R. J. Samulski.** 1998. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. J. Virol. **72:**1438-1445.

90. **Takeda, S., et al.** 2004. Successful gene transfer using adeno-associated virus vectors into the kidney: comparison among adeno-associated virus serotype 1-5 vectors in vitro and in vivo. Nephron Exp. Nephrol. 96: e119–e126.

91. **Tobiasch E**, **Rabreau M**, **Geletneky K**, **Laruë-Charlus S**, **Severin F**, **Becker N**, **Schlehofer JR.** 1994. Detection of adeno-associated virus DNA in human genital tissue and in material from spontaneous abortion. J Med Virol. 44(2):215-22.

92. **Tratschin, J. D., I. L. Miller, M. G. Smith, and B. J. Carter.** 1985. Adenoassociated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol. Cell. Biol. **5:**3251-3260.

93. **Urabe, M., Y. Hasumi, A. Kume, R. T. Surosky, G. J. Kurtzman, K. Tobita, and K. Ozawa.** 1999. Charged-to-alanine scanning mutagenesis of the N-terminal half of adeno-associated virus type 2 Rep78 protein. J. Virol. 73:2682-2693.

94. **Valle, M., C. Gruss, L. Halmer, J. M. Carazo, and L. E. Donate.** 2000. Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. Mol. Cell. Biol. 20:34-41.

95. **Vos, J.C., van Luenen, H.G.A.M., and Plasterk, R.H.A.** 1993. Characterization of the Caenorhabditis elegans Tc1 transposase in vivo and in vitro. Genes Dev. 7, 1244–1253.

96. **Wang B, Li J, Fu FH, Xiao X.** 2009. Systemic human minidystrophin gene transfer improves functions and life span of dystrophin and dystrophin/utrophin-deficient mice.J Orthop Res. 27(4):421-6.

97. **Weitzman, M. D., S. R. M. Kyöstiö, R. M. Kotin, and R. A. Owens.** 1994. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc. Natl. Acad. Sci. USA 91:5808-5812. Virology. 2003 Sep 1;313(2):364-76.

98. **Wu, J., Davis, M.D., Owens, R.A.** 2001. A Rep recognition sequence is necessary but not sufficient for nicking of DNA by adeno-associated virus type-2 Rep proteins. Arch. Biochem. Biophys. *389*, 271-7.

99. **Wu, Z., Asokan, A., Samulski, R.J.** 2006 Adeno-associated virus serotypes: vector toolkit for human gene therapy. Mol Ther. 14, 316–327.

100. **Xiao, W., Chirmule, N., Berta, S. C., McCullough, B., Gao, G. and Wilson, J. M.** 1999. Gene therapy vectors based on adeno-associated virus type 1.J. Virol. 73: 3994–4003.

101. **Xiao, X., J. Li, and R. J. Samulski.** 1996. Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. J. Virol. **70:**8098-8108.

102. **Xiao, X., J. Li, and Samulski, R.J.** 1998**.** Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. J. Virol. *72*, 2224-2232.

103. **Xiao, X., Xiao, W., Li, J., Samulski, R.J.** 1997. A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. J. Virol. *71*, 941-8.

104. **Yan, Z., D. C. Lei-Butters, Y. Zhang, R. Zak, and J. F. Engelhardt.** 2007. Hybrid adeno-associated virus bearing nonhomologous inverted terminal repeats enhances dual-vector reconstruction of minigenes in vivo. Hum. Gene Ther. 18:81-87.

105. **Yan, Z., R. Zak, Y. Zhang, and J. F. Engelhardt.** 2005. Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. J. Virol. **79:**364-379.

106. **Yang, J., W. Zhou, Y. Zhang, T. Zidon, T. Ritchie, and J. F. Engelhardt.** 1999. Concatemerization of adeno-associated virus circular genomes

occurs through intermolecular recombination. J. Virol. **73:**9468-9477.

107. **Yoon, M., Smith, D.H., Ward, P., Medrano, F.J., Aggarwal, A.K., Linden, R.M.** 2001. Amino-terminal domain exchange redirects origin-specific interactions of adeno-associated virus rep78 in vitro. J. Virol. *75*, 3230-9.

108. **Young, S.M. Jr, McCarty, D.M., Degtyareva, N., Samulski, R.J.** 2000. Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination. J. Virol.*74*, 3953-66.

109. **Young, S.M. Jr, Xiao, W., and Samulski, R.J.** 2000. Site-specific targeting of DNA plasmids to chromosome 19 using AAV cis and trans sequences. Methods Mol. Biol. *133*, 111-26.

110. **Zadori, Z., J. Szelei, M. C. Lacoste, Y. Li, S. Gariepy, P. Raymond, M. Allaire, I..R. Nabi, and P. Tijssen.** 2001. A viral phospholipase A2 is required for parvovirus infectivity. Dev Cell 1:291-302.

111. **Zaiss, A. K., Q. Liu, G. P. Bowen, N. C. Wong, J. S. Bartlett, and D. A. Muruve.** 2002. Differential activation of innate immune responses by adenovirus and adeno-associated virus vectors. J. Virol. 76:4580-4590.

112. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406-3415.