Biotechnology and Genetic Engineering Reviews - Vol. 24, 165-178 (2007)

Expanding Adeno-associated Viral Vector Capacity: A Tale of Two Vectors

ARKASUBHRA GHOSH AND DONGSHENG DUAN*

Department of Molecular Microbiology and Immunology, University of Missouri, School of Medicine, One Hospital Dr., Room M610G, MSB, Columbia, MO 65212, USA

Introduction

Adeno-associated Virus (AAV) is a human parvovirus originally isolated as a contaminant in adenoviral stocks (Atchison, *et al*., 1965). It is classified as a dependovirus, as opposed to autonomous parvovirus, because it needs a helper virus to complete a productive life cycle (Flotte and Berns, 2005). About 110 serotypes have been identified and they are classifed into seven clades (Gao, *et al*., 2004). Type-2 AAV (AAV-2) is the first serotype isolated. This is also the most extensively studied prototype AAV. Wild type AAV-2 consists of a 4681nt single-stranded DNA genome and is non-pathogenic (Srivastava, *et al*., 1983). The ends of the genome are palindromic repeat sequences termed the inverted terminal repeats (ITRs). The AAV genome consists of two open reading frames encoding structural/capsid proteins and replication/regulation proteins, respectively (*Figure 1*). AAV enters cell through receptor-mediated endocytosis followed by endosomal sorting and trafficking into the nucleus (Ding, *et al*., 2005). After uncoating, the AAV genome is converted to the double stranded form. In the presence of a helper virus, AAV is reproduced (Berns and Bohenzky, 1987) (*Figure 2*). Without helper virus, a latent infection cycle is established and AAV is propagated as an integrated provirus (Berns and Linden, 1995). In human cells, AAV-2 specifically integrates in the short arm of chromosome 19 (Kotin, *et al*., 1990). Recent advances in the field of gene therapy

*To whom correspondence may be addressed (duand@missouri.edu)

Abbreviations: AAV: adeno-associated virus; AP: alkaline phosphatase; BMD: Becker muscular dystrophy; CFTR: cystic fibrosis transmembrane conductance regulator; CMV: cytomegalovirus; DMD: Duchenne muscular dystrophy; Double-D: two D sequences; ITR: inverted terminal repeat; kb: kilobases; MW: molecular weight; nt: nucleotides; rAAV: recombinant adeno-associated viral vector; SA: splice acceptor; SD: splice donor; SV40: simian virus 40.

have led to a spurt in using recombinant AAV (rAAV) for gene delivery (Carter, 2004; Carter, 2005). In rAAV, the native AAV genes are replaced by the therapeutic gene(s) (*Figure 3*). After infection, rAAV mainly persists as a latent circular episomal genome in host cells (Duan, *et al*., 1998). These circular genomes are thought to be responsible for long term gene expression (Duan, *et al*., 1998; Duan, *et al*., 1999).

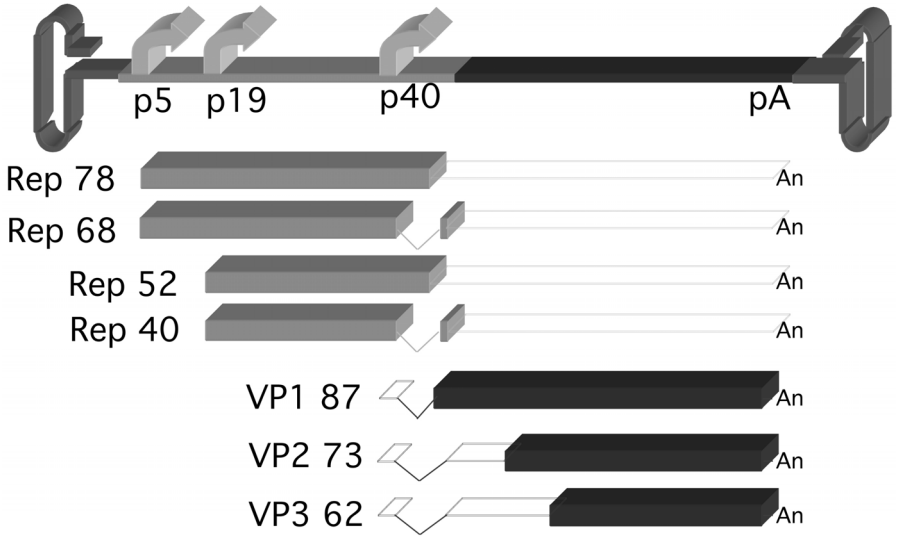

Figure 1. Organization of the wild-type AAV-2 genome. The wild type AAV genome has two open reading frames encoding four replication proteins (Rep78, Rep68, Rep52 and Rep40) and three structural proteins (VP1, VP2 and VP3; the numerical numbers indicate MW of each capsid protein). The replication proteins are expressed from the p5 and p19 promoters and undergo alternative splicing to generate different transcripts. The structural proteins are all expressed from the p40 promoter. They are generated by alternative splicing and/or different start codons.

AAV vectors carry several features that are advantageous for gene therapy. These include their ability to infect both mitotic and post mitotic cells, longterm gene expression in a wide range of tissues and very low immune response. Many preclinical and clinical gene therapy studies have been performed with rAAV over the last two decades (Carter, 2004; Carter, 2005). Although most of the work has used the prototype AAV-2, investigators have begun to appreciate the newly identified AAV serotypes as gene therapy vehicles over the past few years (Gao, *et al*., 2005; Wu, *et al*., 2006). These serotypes display unique transduction patterns in different tissues. One of the most striking findings is that some serotypes, in particular AAV-6, -8 and -9, can mediate efficient whole body gene transfer (Gregorevic, *et al*., 2004; Inagaki, *et al*., 2006; Pacak, *et al*., 2006; Wang, *et al*., 2005; Ghosh, *et al.*, 2007). It is interesting to note that AAV-2 ITRs have been used as packaging signals in these studies. The serotype-specific differences in transduction are most likely attributed to the differences in viral uptake and/or intracellular trafficking.

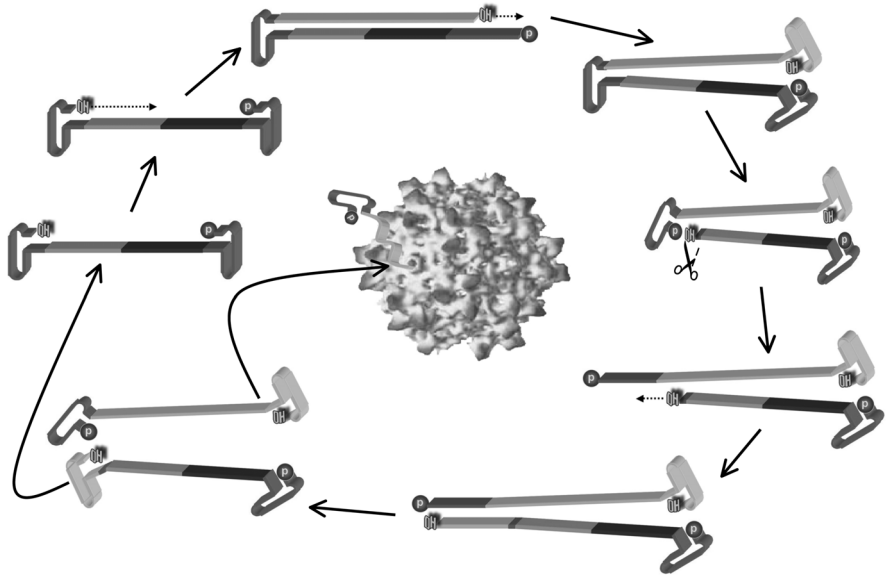

Figure 2. Replication and packaging of the AAV genome. AAV replication starts from a single-stranded template. Primed by the 3' hydroxy group at the end of the ITR, a second strand is synthesized. The second strand synthesis continues across the ITR on the other end, followed by resolution at the left-hand terminal resolution site (scissor sign). This allows completion of the left hand ITR replication. Finally a pair of positive and negative single-stranded AAV genomes is produced. These new genomes may now serve as templates for further rounds of amplification. Alternatively they can be packaged into pre-assembled empty capsid.

Figure 3. Organization of a typical recombinant AAV (rAAV) genome. In a rAAV vector, an expression cassette containing a reporter or a therapeutic gene of choice replaces all the viral genes of the wild-type virus. Only the ITRs are kept to serve as the essential replication/packaging signal.

Strategies for expanding AAV packaging capacity

Despite its promise as an excellent therapeutic vehicle, AAV has been excluded for the gene therapy of many diseases. This is primarily due to the small packaging capacity. Up to a 5 kb therapeutic expression cassette can be encapsidated in a rAAV virion, beyond which the packaging efficiency drops significantly (Dong, *et al*., 1996). This works perfectly for diseases involving small genes such as hemophilia B and

alpha 1-antitrypsin deficiency (Flotte, *et al*., 2004; Manno, *et al*., 2006). However, it cannot meet the need of a large therapeutic gene such as the dystrophin gene (the gene responsible for Duchenne and Becker muscular dystrophy, DMD and BMD) and the cystic fibrosis transmembrane conductance regulator (CFTR) gene (the gene responsible for cystic fibrosis).

To overcome this size limitation, efforts have been made to delete the less important regions in a therapeutic gene and to generate minimized genes that can fit into a single AAV capsid. However, such mini-genes may be less functional than the full-length gene (Harper, *et al*., 2002; Lai, *et al*., 2005). Alternative strategies are needed to deliver a large gene with AAV vector. It has been demonstrated that AAV genomes undergo intermolecular recombination at their ITRs to form larger concatameres (Duan, *et al*., 1999; Duan, *et al*., 1998; Duan, *et al*., 2003; Song, *et al*., 2001; Song, *et al*., 2004; Yue and Duan, 2003). At the junction of the viral genomes is a unique structure called the double-D ITR (Duan, *et al*., 1999). It is further observed that the majority of the recombined genome exists in a head-to-tail orientation (Duan, *et al*., 1998; Duan, *et al*., 1999). This feature of the AAV genome metabolism offers a great opportunity to double the AAV packaging capacity with two vectors (Duan, *et al*., 2000; Sun, *et al*., 2000; Yan, *et al*., 2000). Several distinctive dual vector approaches have being developed since then. These include cis-activation and trans-splicing (*Figure 4*) (Duan, *et al*., 2002; Duan, *et al*., 2000; Yan, *et al*., 2000). In the cis-activation approach, the therapeutic expression cassette is split into two parts one containing the regulatory elements such as the promoter and the enhancer while the other carrying the protein coding sequence (*Figure 4A*). These two parts are packaged into independent AAV vectors. Co-administration will reconstitute the expression cassette. In the trans-splicing approach, the gene itself is split into two AAV vectors (*Figure 4B*). The split gene fragments are engineered with splicing signals. After co-infection, a functional open reading frame is reconstituted by cellular splicing machinery from the recombined vector genome.

Besides viral ITR-mediated recombination, intermolecular recombination between AAV genomes also occurs through homologous recombination (Senapathy, *et al*., 1984; Wang, *et al*., 1995). This yields another option to expand AAV packaging capacity called the overlapping approach (*Figure 4C*) (Duan, *et al*., 2001; Halbert, *et al*., 2002). Essentially, a large gene is split into two overlapping fragments and packaged into two AAV vectors. When these vectors enter the same cell, the full-length gene is reconstituted through homologous recombination at the overlapping region.

CIS-ACTIVATING A "PROMOTER-LESS" AAV

The size of many therapeutic genes (such as the CFTR cDNA) is in the range of ~4.6 kb. These genes will fit into a single AAV virion themselves but there is no space for a strong promoter/enhance to drive gene expression. The AAV ITR has weak promoter activity and it can drive minimal level expression when a strong exogenous promoter is omitted in an AAV vector (Flotte, *et al*., 1993). This is apparently not sufficient for gene therapy. One way to overcome this problem is to introduce strong enhancers through a second AAV vector to strengthen transcription. Duan *et al* tested this approach by constructing a "super-enhancer" AAV vector (Duan, *et al*., 2000). This

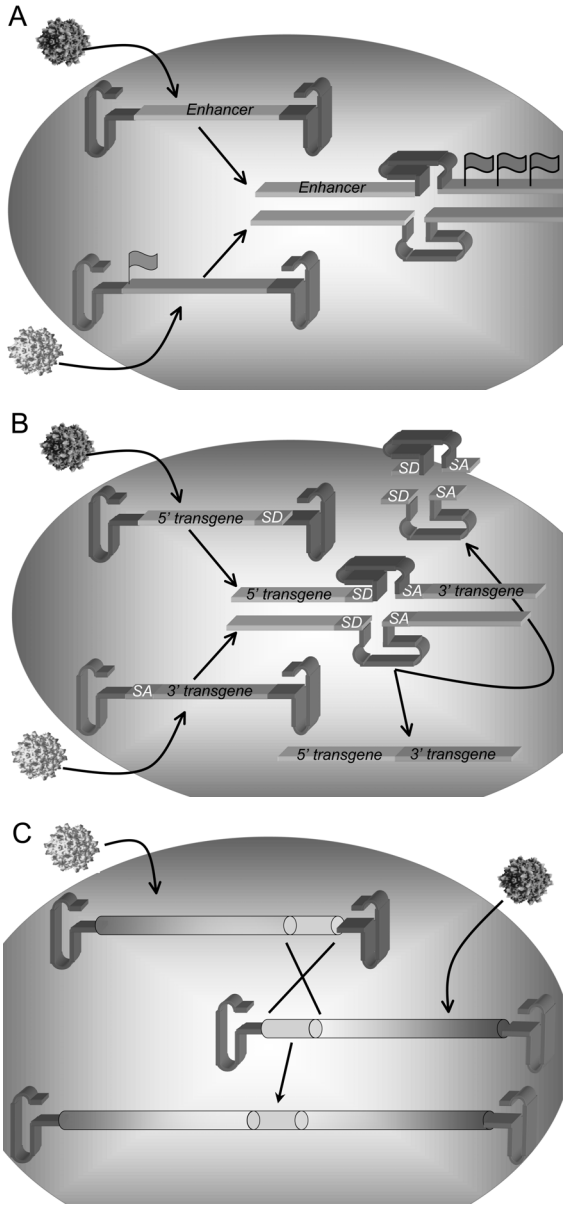

Figure 4. Schematic illustration of the different dual vector approaches to expand the AAV packaging capacity. **A,** Cis-activation. In this case, a vector carrying a promoterless transgene (or a transgene with a weak promoter) is co-infected with a vector containing the "super-enhancer". The rAAV genomes undergo ITR mediated recombination to form heterodimeric molecules. The transgene expression is highly enhanced by the presence of the "super-enhancer" elements in the concatamer. **B,** Trans-splicing. In this approach, the transgene is split into two halves in the middle of an intron. The AV.Donor molecule contains the 5' half of the expression cassette with the splice donor (SD) and the AV.Acceptor contains the 3' half of the expression cassette containing the splice acceptor (SA). Upon co-infection with both vectors, ITR mediated heterodimers are formed. The transgene is transcribed from the recombined vector genome. The splice signals from the intron will then splice out the ITR junction and generate a full length mature mRNA transcript. **C,** Overlapping. This method requires the transgene to be split in such a way that the two split parts share a sequence of homology between them. Upon co-infection, the two split parts will recombine at their homologous sequence. This recombination event will thus lead to the generation of an intact expression cassette.

vector contains two enhancer elements from the simian virus 40 (SV40) and cytomegalovirus (CMV) genomes, respectively. They also made a "promoter-less" luciferase AAV vector. Co-administration of the "super-enhancer" AAV vector resulted in 200-fold increase in transgene expression from the "promoter-less" vector. Molecular analysis of the retrieved vector genomes revealed intermolecular recombination between the two incoming vector genomes and the formation of a strong expression cassette containing the ITR promoter, the luciferase gene and the super-enhancer. The cis-activation method has several advantages for certain gene therapy applications. First, it does not entail the splitting of the gene within the coding sequence. Second, it works irrespective of the orientation of the recombined vector genome. However, there are also apparent limitations. One of the biggest limitations is that it cannot work for the genes where the coding sequence is beyond the packaging capacity of a single AAV. Safety is another important concern. Random integration of the super-enhancer elements may lead to oncogenic expression of the neighboring genes.

EXPANDING PACKAGING CAPACITY WITH THE TRANS-SPLICING VECTORS

Eukaryotic genes consist of coding regions called exons interspersed with non-coding regions called introns. At the junctions of these introns and exons are the splicing signals. These splicing signals direct the formation of the spliceosome complex at the appropriate locations. Subsequently, the introns are removed and mature mRNA is generated to make protein. The trans-splicing approach takes advantage of the cellular splicing process. Basically, the transgene is divided into two parts either at an endogenous intron or an engineered synthetic "intron". The vector carrying the splicing donor signal is called the AV.Donor. This vector also carries the promoter and the 5' half of the transgene. The vector carrying the splicing acceptor signal is called the AV.Acceptor. This vector also carries the 3' half of the transgene and the polyadenylation signals. Upon co-infection, two vector genomes form the head-to-tail heterodimer. The RNA transcript from such a heterodimer contains both the 5' and the 3' halves of the transgene as well as splicing signals and the double-D ITR at the junction. The junction region is then spliced out from pre-mRNA to form mature mRNA for protein expression.

A number of therapeutic and/or reporter transgenes have been tested in the trans-splicing vectors (Chao, *et al*., 2002; Duan, *et al*., 2001; Lai, *et al*., 2005; Nakai, *et al*., 2000; Reich, *et al*., 2003; Sun, *et al*., 2000; Yan, *et al*., 2000). Yan *et al* split the human erythopoitin gene at its 3rd intron and packaged them into the trans-splicing vectors (Yan, *et al*., 2000). Co-infection in mouse skeletal muscle resulted in therapeutic level of erythropoietin in the serum. Single vector administration did not lead to any erythropoietin expression. Chao *et al* generated the factor 8 gene trans-splicing vectors (Chao, *et al*., 2002). In immune deficient mice, intra-portal vein co-administration produced 2% of normal level of factor VIII in blood. Several groups also split the LacZ reporter gene using the synthetic splicing signals (Duan, *et al*., 2001; Nakai, *et al*., 2000; Reich, *et al*., 2003; Sun, *et al*., 2000). In these proof of principle studies, the trans-splicing vectors consistently regenerated the split gene and lead to the expression of functional protein. However, the actual transduction efficiency is much lower than that of a single intact AAV vector (Duan, *et al*., 2001).

There are a number of potential rate limiting factors in the trans-splicing approach (Lai, *et al*., 2005; Xu, *et al*., 2004; Yan, *et al*., 2005). These include co-infection efficiency, formation of the head-to-tail heterodimer genome, transcription and the stability of the double-D ITR containing pre-mRNA and splicing across the double-D ITR structure (*Figure 5*). It is now clear that co-infection is not a barrier. Even in a diseased tissue, co-infection efficiency can reach 90% (Xu, *et al*., 2004). To test whether unidirectional heterodimer formation is a rate-limiting step, Yan *et al* generated trans-splicing AAV vectors with AAV-2 ITR and AAV-5 ITR at the opposite ends of the genomes (Yan, *et al*., 2005). Recombination between AAV-2 ITR and AAV-5 ITR is very inefficient. This leads to preferential unidirectional recombination of the ITRs from the same serotype. Comparing with the traditional trans-splicing vectors which carry the same ITR at both ends, the hybrid ITR vectors resulted in 6~10-fold enhancement in gene expression. This study confirms that head-to-tail intermolecular recombination is a rate-limiting barrier.

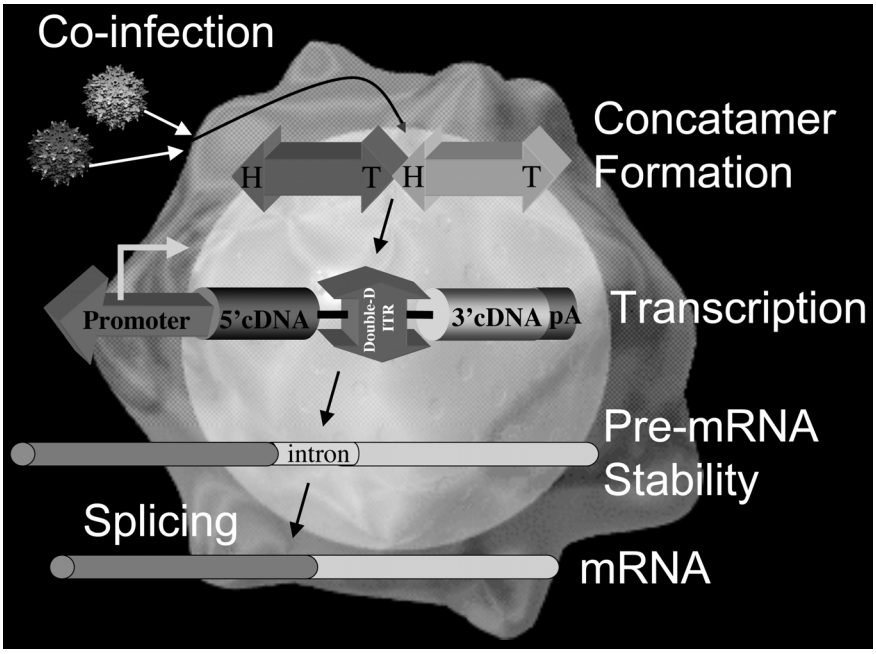

Figure 5. Rate-limiting steps in the trans-splicing AAV vectors. There are several potential rate-limiting barriers in the trans-splicing approach. These include co-infection efficiency, heterodimer formation, transcription across the double-D ITR, pre-mRNA stability and splicing.

To test whether the double-D ITR represents a critical rate-limiting factor, Xu *et al* generated a mini-cassette that resembles the recombined vector genome (Xu, *et al*., 2004). Essentially, a double-D ITR structure is inserted in the intron of a synthetic LacZ gene. Comparing these with the original LacZ construct and an intron-containing LacZ construct, the total mRNA transcript level was significantly reduced in the presence of the double-D ITR structure. Pre-mRNA stability was also reduced by the double-D ITR structure. Ultimately, gene expression was reduced by 50% at the

protein level. These results demonstrated that mRNA accumulation is also a rate-limiting factor.

An often neglected issue is the gene splitting site. Depending on the sequence specificity, different exon/intron/exon junctions are spliced at different efficiencies. Furthermore, surrounding exonic splicing enhancer elements may also modulate the overall level of mRNA production. To test whether the gene splitting site influences transduction efficiency of the trans-splicing vectors, Lai *et al* systematically screened a number of different exon/intron/exon junctions in a 6 kb mini-dystrophin gene (Lai, *et al*., 2005). At first, they evaluated mRNA accumulation in synthetic mini-cassettes that contain the double-D ITR structure. Two sets of the trans-splicing vectors were then generated based on the sites that yielded the highest levels of mRNA. In vivo examination in a mouse model of DMD revealed wide-spread mini-dystrophin expression from the vector set that performed the best in mRNA accumulation screening. The other vector set was developed on a different site that had ~50% mRNA level of the best site. This second vector set failed to yield therapeutic level transduction.

Taken together, tremendous progress has been made in the last few years in optimizing the trans-splicing vectors. Identification of the potential rate-limiting factors has resulted in novel approaches to bypass these barriers. Therapeutic levels of both secretary and structural proteins have been achieved. The development of the trans-splicing vector has finally moved from the proof-of-principle stage to the gene therapy application stage.

DEVELOPMENT OF THE OVERLAPPING VECTOR APPROACH

The overlapping vector approach provides an alternative to the trans-splicing approach where a homologous recombination dependent strategy is employed to reconstitute the intact transgene. Homologous recombination has been shown to regenerate the deleted regions in the ITR (Senapathy, *et al*., 1984; Wang, *et al*., 1995). Recent evidence suggests that homologous recombination may also occur between rAAV genome and genomic DNA (Hendrie and Russell, 2005; Liu, *et al*., 2004; Miller, *et al*., 2006; Russell and Hirata, 1998). It is therefore expected that AAV vectors would be capable of homologous recombination between each other should they share a region of homology.

The possibility of using homologous recombination-based approach to expand AAV packaging capacity was initially tested by Duan *et al* (Duan, *et al*., 2001). They generated two vectors including AV.Upstream and AV.Downstream, respectively. The AV.Upstream vector carries the promoter and the 5' two-thirds of the LacZ gene. The AAV.Downstream vector carries the 3' two-thirds of the LacZ gene followed by the polyadenylation signal. There is a 1 kb LacZ sequence (the middle one-third) shared by both vectors. After co-infection in muscle, LacZ positive fibers were detected. Despite the proof-of-principle, transduction efficiency of the overlapping vectors is several logs lower than that of a single intact vector (Duan, *et al*., 2001). Interestingly, when Halbert *et al* applied the similar strategy to the alkaline phosphatase (AP) gene in the lung, they observed an extraordinarily high level expression, a level that was comparable to a single intact AAV vector (Halbert, *et al*., 2002). To resolve the

difference between these two studies, Ghosh and coworkers compared the LacZ and the AP overlapping vectors in skeletal muscle (Ghosh, *et al*., 2006). Surprisingly, despite a poor performance of the LacZ overlapping vectors, the AP overlapping vectors were able to efficiently transduce muscle. Using a similar quantification method as used by Halbert *et al*, Ghosh *et al* demonstrated that transduction efficiency of the overlapping vectors could also reach that of a single intact vector in muscle in the context of the AP gene (Ghosh, *et al*., 2006). Taken together, these results suggest that the success of the overlapping vectors is largely dependent on the transgene itself. If a large gene contains a favorable recombinogenic region, one may expect to achieve therapeutic level expression from the overlapping AAV vectors.

Summary and future perspective

The demonstration that AAV can mediate safe and longterm gene expression has spawned a very deep interest in developing this virus for gene therapy. Being one of the smallest viruses, AAV faces a unique size-related challenge. It cannot deliver a large therapeutic expression cassette. This has essentially excluded AAV gene therapy for a number of common diseases such DMD/BMD, hemophilia A and cystic fibrosis. The dual vector strategies developed over the past few years have offered a real promise in overcoming the size limitation. With the increased packaging capacity, the territory of AAV gene therapy can now be expanded further. The next important step will be to move from the exciting results seen in rodents to large animal models of human diseases and eventually to human patients. Studies in several areas may likely accelerate this process. The first is to better understand the basic transduction biology of rAAV. A good example of this is the development of new vectors based on alternative AAV serotypes. These new vectors have not only enhanced transduction from a single intact AAV virion, but they have also played a critical role for the success of dual AAV vectors (Ghosh, *et al*., 2006; Halbert, *et al*., 2002; Reich, *et al*., 2003). Currently, only AAV-2 has been approved for patient use. Clinical application of these new serotype vectors will certainly benefit from a clear understanding of their transduction process such as their receptor(s) and co-receptor(s), intracellular trafficking and potential integration. Secondly, it is important to evaluate systemic gene delivery with dual AAV vectors. Many diseases affect multiple tissues/organs in the body. Local gene transfer may only yield limited benefit to these patients. From this standpoint, a body-wide gene therapy will be required to cure the disease. None of the dual vector strategies has been evaluated for whole body transduction. Current successes are all based on local gene transfer (Note, since the submission of this manuscript, Ghosh *et al* have achieved efficient whole body transduction in newborn mice with the trans-splicing AAV vectors). And finally, the dual vector design needs to be further streamlined and optimized. Although we have a collection of different dual vector strategies to choose for different applications, all these strategies have their inherent limitations. For example, in the trans-splicing approach, we may need to screen a large number of potential gene splitting sites before we can find one that works. On the other hand, the overlapping approach may only work for the selected genes. Creative means have to be developed for generic application of the dual vector approach in any gene in the future.

Acknowledgements

We thank helpful discussion from Duan lab members. This work is supported by grants from the National Institutes of Health (AR-49419, DD), the Muscular Dystrophy Association (DD) and the Cystic Fibrosis Foundation (DD).

References

ATCHISON, R. W., CASTO, B. C. and HAMMON, W. M. (1965). Adenovirus-associated defective virus particles. *Science*. **149**. 754-6

BERNS, K. I. and BOHENZKY, R. A. (1987). Adeno-associated viruses: an update. *Adv Virus Res*. **32**. 243-306

BERNS, K. I. and LINDEN, R. M. (1995). The cryptic life style of adeno-associated virus. *Bioessays*. **17**. 237-45.

CARTER, B. J. (2004). Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective. *Mol Ther*. **10**. 981-9

CARTER, B. J. (2005). Adeno-associated virus vectors in clinical trials. *Hum Gene Ther*. **16**. 541-50

CHAO, H., SUN, L., BRUCE, A., XIAO, X. and WALSH, C. E. (2002). Expression of human factor VIII by splicing between dimerized AAV vectors. *Mol Ther*. **5**. 716-22.

DING, W., ZHANG, L., YAN, Z. and ENGELHARDT, J. F. (2005). Intracellular trafficking of adeno-associated viral vectors. *Gene Ther*. **12**. 873-80

DONG, J. Y., FAN, P. D. and FRIZZELL, R. A. (1996). Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum Gene Ther*. **7**. 2101-12

DUAN, D., SHARMA, P., DUDUS, L., ZHANG, Y., SANLIOGLU, S., YAN, Z., YUE, Y., YE, Y., LESTER, R., YANG, J., FISHER, K. J. and ENGELHARDT, J. F. (1999). Formation of adeno-associated virus circular genomes is differentially regulated by adenovirus E4 ORF6 and E2a gene expression. *J Virol*. **73**. 161-9.

DUAN, D., SHARMA, P., YANG, J., YUE, Y., DUDUS, L., ZHANG, Y., FISHER, K. J. and ENGELHARDT, J. F. (1998). Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long term episomal persistence in muscle. *J. Virol.* **72**. 8568-8577

DUAN, D., YAN, Z., YUE, Y. and ENGELHARDT, J. F. (1999). Structural analysis of adeno-associated virus transduction intermediates. *Virology*. **261**. 8-14.

DUAN, D., YUE, Y. and ENGELHARDT, J. F. (2003). Consequences of DNA-dependent protein kinase catalytic subunit deficiency on recombinant adeno-associated virus genome circularization and heterodimerization in muscle tissue. *J Virol*. **77**. 4751-9

DUAN, D., YUE, Y. and ENGELHARDT, J. F. (2001). Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison. *Mol Ther*. **4**. 383-91.

DUAN, D., YUE, Y., YAN, Z. and ENGELHARDT, J. F. (2000). A new dual-vector approach to enhance recombinant adeno-associated virus-mediated gene expression through intermolecular cis activation. *Nat. Med.* **6**. 595-8

FLOTTE, T. R., AFIONE, S. A., SOLOW, R., DRUMM, M. L., MARKAKIS, D., GUGGINO, W. B., ZEITLIN, P. L. and CARTER, B. J. (1993). Expression of the cystic fibrosis

transmembrane conductance regulator from a novel adeno-associated virus promoter. *J Biol Chem*. **268**. 3781-90

Flotte, T. R. and Berns, K. I. (2005). Adeno-associated virus: a ubiquitous commensal of mammals. *Hum Gene Ther*. **16**. 401-7

Flotte, T. R., Brantly, M. L., Spencer, L. T., Byrne, B. J., Spencer, C. T., Baker, D. J. and Humphries, M. (2004). Phase I trial of intramuscular injection of a recombinant adeno-associated virus alpha 1-antitrypsin (rAAV2-CB-hAAT) gene vector to AAT-deficient adults. *Hum Gene Ther*. **15**. 93-128

Gao, G., Vandenberghe, L. H., Alvira, M. R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J. M. (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol*. **78**. 6381-8

Gao, G., Vandenberghe, L. H. and Wilson, J. M. (2005). New recombinant serotypes of AAV vectors. *Curr Gene Ther*. **5**. 285-97

Ghosh, A., Yue, Y. and Duan, D. (2006). Viral serotype and the transgene sequence influence overlapping adeno-associated viral (AAV) vector-mediated gene transfer in skeletal muscle. *J. Gene Med.* **8**. 298-305

Ghosh, A., Yue, Y., Long, C., Bostick, B. and Duan, D. (2007) Efficient whole-body transduction with trans-splicing adeno-associated viral vectors. *Mol Ther.* **15.** 750-5.

Gregorevic, P., Blankinship, M. J., Allen, J. M., Crawford, R. W., Meuse, L., Miller, D. G., Russell, D. W. and Chamberlain, J. S. (2004). Systemic delivery of genes to striated muscles using adeno-associated viral vectors. *Nat Med*. **10**. 828-34

Halbert, C. L., Allen, J. M. and Miller, A. D. (2002). Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol*. **20**. 697-701.

Harper, S. Q., Hauser, M. A., DelloRusso, C., Duan, D., Crawford, R. W., Phelps, S. F., Harper, H. A., Robinson, A. S., Engelhardt, J. F., Brooks, S. V. and Chamberlain, J. S. (2002). Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nat Med*. **8**. 253-61.

Hendrie, P. C. and Russell, D. W. (2005). Gene targeting with viral vectors. *Mol Ther*. **12**. 9-17

Inagaki, K., Fuess, S., Storm, T. A., Gibson, G. A., McTiernan, C. F., Kay, M. A. and Nakai, H. (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol Ther*. **14**. 45-53

Kotin, R. M., Siniscalco, M., Samulski, R. J., Zhu, X. D., Hunter, L., Laughlin, C. A., McLaughlin, S., Muzyczka, N., Rocchi, M. and Berns, K. I. (1990). Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci U S A*. **87**. 2211-5

Lai, Y., Yue, Y., Liu, M., Ghosh, A., Engelhardt, J. F., Chamberlain, J. S. and Duan, D. (2005). Efficient in vivo gene expression by trans-splicing adeno-associated viral vectors. *Nat Biotechnol*. **23**. 1435-9

Liu, X., Yan, Z., Luo, M., Zak, R., Li, Z., Driskell, R. R., Huang, Y., Tran, N. and Engelhardt, J. F. (2004). Targeted correction of single-base-pair mutations with adeno-associated virus vectors under nonselective conditions. *J Virol*. **78**. 4165-75

Manno, C. S., Arruda, V. R., Pierce, G. F., Glader, B., Ragni, M., Rasko, J., Ozelo, M. C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Nakai, H., Chew, A., Leonard, D., Wright, J. F., Lessard, R. R., Sommer, J. M., Tigges, M., Sabatino, D., Luk, A., Jiang, H., Mingozzi, F., Couto, L., Ertl, H. C., High, K. A. and Kay, M. A. (2006). Successful transduction of liver in hemophilia

by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med*. **12**. 342-347

MILLER, D. G., WANG, P. R., PETEK, L. M., HIRATA, R. K., SANDS, M. S. and RUSSELL, D. W. (2006). Gene targeting in vivo by adeno-associated virus vectors. *Nat Biotechnol*. **24**. 1022-6

NAKAI, H., STORM, T. A. and KAY, M. A. (2000). Increasing the size of rAAV-mediated expression cassettes in vivo by intermolecular joining of two complementary vectors [see comments]. *Nat. Biotechnol.* **18**. 527-32

PACAK, C. A., MAH, C. S., THATTALIYATH, B. D., CONLON, T. J., LEWIS, M. A., CLOUTIER, D. E., ZOLOTUKHIN, I., TARANTAL, A. F. and BYRNE, B. J. (2006). Recombinant adeno-associated virus serotype 9 leads to preferential cardiac transduction in vivo. *Circ Res*.

REICH, S. J., AURICCHIO, A., HILDINGER, M., GLOVER, E., MAGUIRE, A. M., WILSON, J. M. and BENNETT, J. (2003). Efficient trans-splicing in the retina expands the utility of adeno-associated virus as a vector for gene therapy. *Hum Gene Ther*. **14**. 37-44

RUSSELL, D. W. and HIRATA, R. K. (1998). Human gene targeting by viral vectors. *Nat. Genet.* **18**. 325-30.

SENAPATHY, P., TRATSCHIN, J. D. and CARTER, B. J. (1984). Replication of adeno-associated virus DNA. Complementation of naturally occurring rep- mutants by a wild-type genome or an ori- mutant and correction of terminal palindrome deletions. *J Mol Biol*. **179**. 1-20.

SONG, S., LAIPIS, P. J., BERNS, K. I. and FLOTTE, T. R. (2001). Effect of DNA-dependent protein kinase on the molecular fate of the rAAV2 genome in skeletal muscle. *Proc. Natl. Acad. Sci. U S A*. **98**. 4084-8.

SONG, S., LU, Y., CHOI, Y. K., HAN, Y., TANG, Q., ZHAO, G., BERNS, K. I. and FLOTTE, T. R. (2004). DNA-dependent PK inhibits adeno-associated virus DNA integration. *Proc Natl Acad Sci U S A*. **101**. 2112-6

SRIVASTAVA, A., LUSBY, E. W. and BERNS, K. I. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J Virol*. **45**. 555-64

SUN, L., LI, J. and XIAO, X. (2000). Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. *Nat. Med.* **6**. 599-602

WANG, X. S., PONNAZHAGAN, S. and SRIVASTAVA, A. (1995). Rescue and replication signals of the adeno-associated virus 2 genome. *J Mol Biol*. **250**. 573-80

WANG, Z., ZHU, T., QIAO, C., ZHOU, L., WANG, B., ZHANG, J., CHEN, C., LI, J. and XIAO, X. (2005). Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat Biotechnol*. **23**. 321-8

WU, Z., ASOKAN, A. and SAMULSKI, R. J. (2006). Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol Ther*. **14**. 316-27

XU, Z., YUE, Y., LAI, Y., YE, C., QIU, J., PINTEL, D. J. and DUAN, D. (2004). Trans-splicing adeno-associated viral vector-mediated gene therapy is limited by the accumulation of spliced mRNA but not by dual vector coinfection efficiency. *Hum Gene Ther*. **15**. 896-905

YAN, Z., ZAK, R., ZHANG, Y. and ENGELHARDT, J. F. (2005). Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. *J Virol*. **79**. 364-79

YAN, Z., ZHANG, Y., DUAN, D. and ENGELHARDT, J. F. (2000). From the cover: trans-splicing vectors expand the utility of adeno-associated virus for gene therapy. *Proc.*

Natl. Acad. Sci. U S A. **97**. 6716-6721

YUE, Y. and DUAN, D. (2003). Double strand interaction is the predominant pathway for intermolecular recombination of adeno-associated viral genomes. *Virology*. **313**. 1-7