Expanding AAV Packaging Capacity with *Trans*-splicing or Overlapping Vectors: A Quantitative Comparison

Dongsheng Duan,[1,3,*] Yongping Yue,[1,3] and John F. Engelhardt[1,2,3]

[1]*Department of Anatomy & Cell Biology,* [2]*Department of Internal Medicine, and*
[3]*Center for Gene Therapy of Cystic Fibrosis and Other Genetic Diseases, College of Medicine, The University of Iowa, Iowa City, Iowa 52242, USA*

**To whom correspondence and reprint requests should be addressed. Fax: (319) 335-7198. E-mail: dongsheng-duan@uiowa.edu.*

Recombinant adeno-associated (rAAV) viral vectors hold great therapeutic potential for human diseases. However, a relatively small packaging capacity (less than 5 kb) has limited the application of rAAV for certain diseases such as cystic fibrosis and Duchenne muscular dystrophy. Here we compared two mechanistically distinct approaches to overcome packaging restraints with rAAV vectors. The *trans*-splicing approach reconstitutes gene expression from two independent rAAV vectors, each encoding unique, nonoverlapping halves of a transgene. This process requires intermolecular concatamerization and subsequent splicing between independent vectors. A distinct overlapping vector approach uses homologous recombination between overlapping regions in two independent vectors. Using the β-galactosidase gene as template, *trans*-splicing approaches were threefold (in primary fibroblasts) and 12-fold (in muscle tissue) more effective in generating full-length transgene products than the overlapping vector approach. Nevertheless, the efficiency of *trans*-splicing remained moderate at approximately 4.3% (for muscle) and 7% (for fibroblasts) of that seen with a single vector encoding the full-length transgene. The efficiency of *trans*-splicing was augmented 185-fold by adenoviral E4, but not E2a, gene products. This augmentation was much less pronounced with the overlapping vectoring approach (12-fold). *Trans*-splicing and overlapping vector approaches are two viable alternatives to expand rAAV packaging capacity.

Key words: recombinant adeno-associated virus, rAAV, homologous recombination, circular intermediates, viral packaging, gene therapy, muscle, splicing

INTRODUCTION

The single-stranded adeno-associated virus (AAV) has evolved into one of the most important gene delivery vehicles in the field of molecular therapy [1]. Application of recombinant AAV (rAAV) vectors in multiple organ systems has also demonstrated therapeutic efficacy for diseases ranging from hemophilia to blindness [2,3]. Despite these successes, packaging limitation remains a fundamental problem for delivering genes larger than 5 kb. Based on the fact that independent AAV genomes can form large, circular concatamers through intermolecular recombination [4,5], breakthrough progress in expanding rAAV packaging capacity was recently demonstrated by several independent laboratories [6–9]. In these studies, either components of the transcription regulatory unit or the transgene itself is split into two parts and separately packaged in two AAV viruses. Co-infection with both viruses leads to the reconstruction of intact expression cassettes through inverted terminal repeat (ITR) mediated intermolecular concatamerization. Through this mechanism, full-length transgene expression (*trans*-splicing) or enhanced transgene expression (*cis* activation) can be achieved by heterodimerization.

In addition to viral genome circularization and intermolecular recombination, recent studies determining the basic biology of rAAV have also revealed other unique biological properties that have expanded the usefulness of this vector system. rAAV can mediate fairly efficient site-specific correction at sites within chromosomal DNA that are homologous to recombinant vectors [10]. The mechanism of this process seems to predominately involve mismatch repair, but may also involve homologous recombination. Homologous recombination between mutant proviral AAV plasmids has also been shown to yield infectious virus in a transfection assay [11], but little is known regarding homologous recombination between single-stranded AAV viral genomes. Homologous recombination involves homologous pairing, strand breaking and single-strand invasion, and formation and resolution of a Holliday-type structure. Single-stranded DNA initiates

1525-0016/01 $35.00

doi:10.1006/mthe.2001.0456, available online at http://www.idealibrary.com on IDEAL

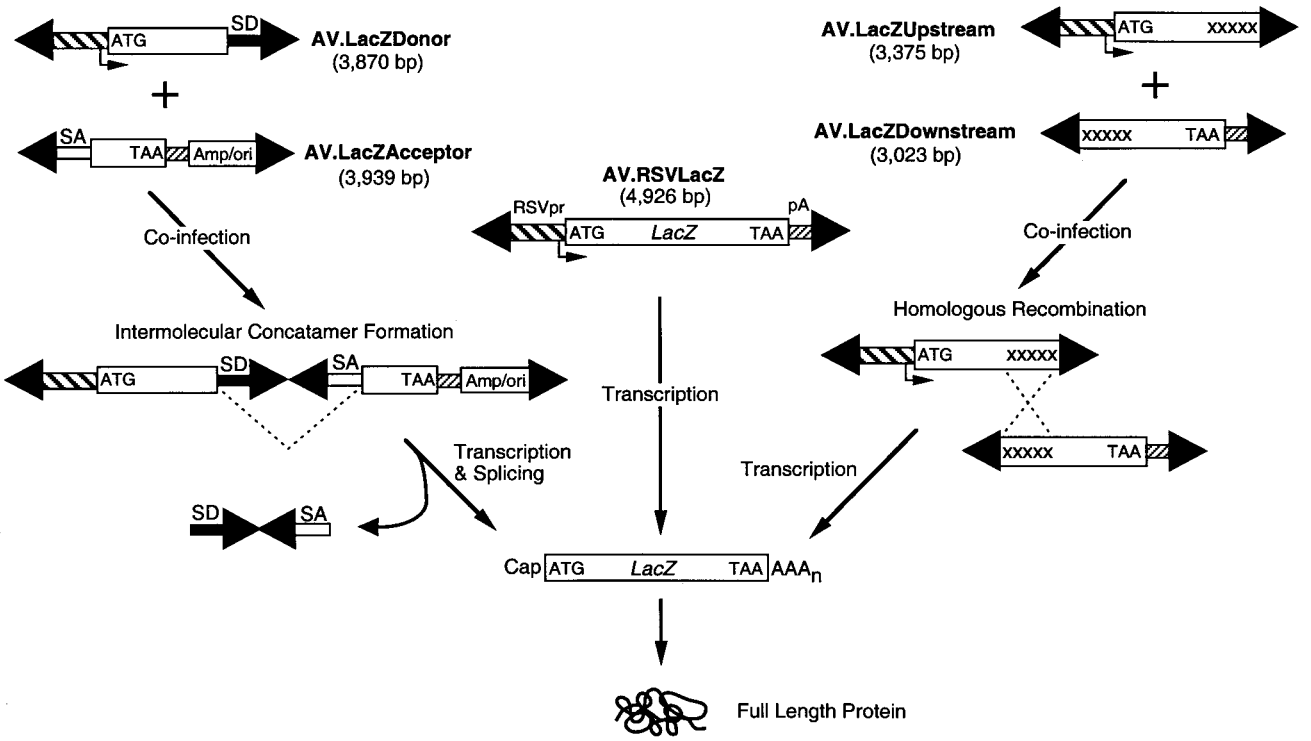

FIG. 1. *Trans*-splicing and overlapping dual rAAV vector approaches reconstitute β-galactosidase gene expression. The β-galactosidase gene was expressed from rAAV vectors using three distinct methods. A full-length β-galactosidase gene encoded in a single vector (AV.RSVLacZ) was used as the standard for the maximal achievable level of expression. Alternatively, the β-galactosidase gene was split into two viral vectors to evaluate *trans*-splicing (AV.LacZDonor/AV.LacZAcceptor) and overlapping (AV.LacZUpstream/AV.LacZDownstream) vector approaches. The 5′ portion and an artificial splice donor sequence (SD) were cloned after the RSV promoter in AV.LacZDonor. The splice acceptor sequence (SA), 3′ portion of the β-galactosidase gene, and a poly(A) signal (pA) were included in a second vector (AV.LacZAcceptor). Co-infection with both donor and acceptor vectors led to heterodimer formation and reconstitution of full-length *lacZ* mRNA and protein following splicing. A third method for reconstituting functional β-galactosidase gene expression from two vectors used homologous recombination between an overlapping set of vectors. The RSV promoter and the first two-thirds of the β-galactosidase gene were cloned in AV.LacZUpstream. The second two-thirds of the β-galactosidase gene and a poly(A) signal were cloned in a second vector, AV.LacZDownstream. The overlapping sequence of these two vectors is denoted by "XXXX." Homologous recombination between the overlapping region of homology in AV.LacZUpstream and AV.LacZDownstream reconstituted a full-length LacZ expression cassette.

recombination in *Escherichia coli* [12]. Hence, it is plausible that the single-stranded nature of AAV and subsequent double-strand conversion might facilitate genetic exchange through homologous recombination. To some extent, this process of homologous recombination between ITRs of independent viral genomes has already been implied through the documentation of double-D circular concatamers [4,5]. However, if homologous recombination also occurs between AAV genomes through ITR independent mechanisms, it might be possible to deliver large genes through an "overlapping" vectoring approach (Fig. 1). In this strategy, two distinct but overlapping halves of a transgene are encoded in two independent vectors that are delivered simultaneously to target cells. Here we compared the efficiency of "*trans*-splicing" and "overlapping" dual vectoring approaches to reconstitute β-galactosidase gene expression in primary fibroblasts and skeletal muscle.

The success of the *trans*-splicing and/or overlapping techniques in expanding the packaging capacity of rAAV in the clinical setting will depend on the ultimate efficiency that can be achieved in a given target organ. Using the human erythropoietin (EPO) genomic DNA as template, we previously demonstrated therapeutic levels of EPO production with *trans*-splicing rAAV vectors [7]. Nakai, *et al.*, have also demonstrated an efficiency of up to 70% of that from a single vector in hepatocytes *in vivo* [9]. However, in the latter case the experimental gene was actually not split in half. Therefore, this optimistic efficiency was a combined effect of *trans*-splicing and *cis* activation [13]. A systematic search for exogenous factors that might be capable of enhancing *trans*-splicing and/or overlapping vectoring approaches holds practical significance. As an initial step in this screening process, we examined the feasibility of increasing dual vector efficiencies with adenoviral proteins. Our results suggested that adenoviral

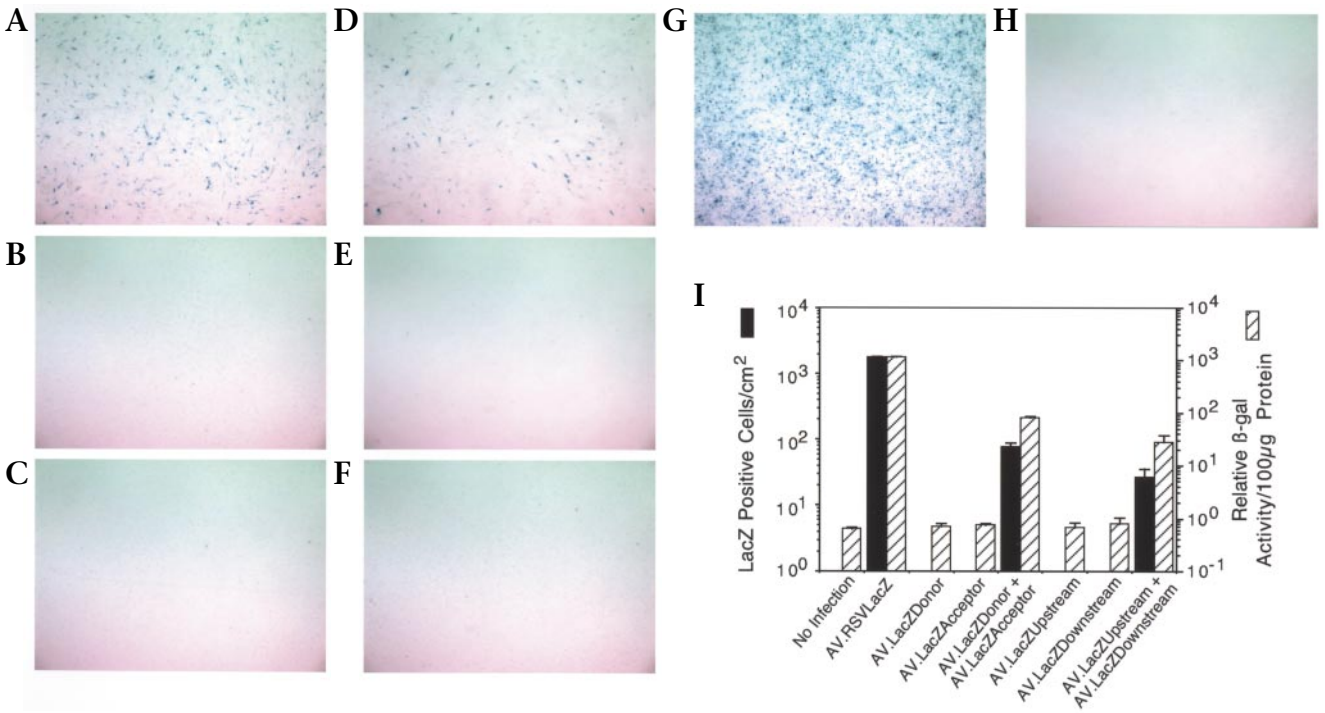

FIG. 2. *In vitro* analysis of *trans*-splicing and overlapping dual rAAV vector approaches in primary fetal fibroblasts. Primary fetal fibroblasts were infected at an MOI of 10,000 DNA particles/cell in serum-free medium with AV.LacZDonor and AV.LacZAcceptor (A), AV.LacZDonor alone (B), AV.LacZAcceptor alone (C), AV.LacZUpstream and AV.LacZDownstream (D), AV.LacZUpstream alone (E), AV.LacZDownstream alone (F), or AV.RSVLacZ (G), or were mock treated (H). The β-galactosidase activity was evaluated at 3 days postinfection by X-gal histochemical staining (A–H) or by quantifying X-gal positive cells and enzymatic activity using a solution-based assay (I). Data in (I) represent the mean (± S.E.M.) of three independent experiments.

E4 gene products can significantly augment *trans*-splicing and, to a lesser extent, overlapping dual vector gene delivery. Furthermore, this enhancement is distinct from the reported effects of E4 ORF3 on promoter activation as determined by proviral plasmid transfection assays [14,15]. These results support new avenues of research for further optimization of rAAV dual vector mediated gene delivery through heterodimerization.

RESULTS

Construction of *Trans*-splicing and Overlapping AAV Vector Sets

Previous studies have demonstrated that a transgene can be split into two rAAV viruses and a full-length transgene transcript can be reconstituted following intermolecular concatamerization of co-infected viral genomes. These approaches have used either endogenous [7] or heterologous intronic sequences [8] to facilitate splicing. In the strategy described by Sun, *et al*. [8], *trans*-splicing vectors were generated by first cloning an intron into the transgene. The intron-containing transgene was then divided into two parts within the intronic region, thereby separating the donor and acceptor sides of the intron/transgene cassette for cloning into donor and acceptor rAAV proviral plasmids. To explore alternative strategies for constructing *trans*-splicing AAV vectors, we took a different approach. We began by dividing the transgene into two rAAV proviral plasmids using restriction sites near a candidate splice site. Next, splicing signals were introduced to these two vectors by a PCR-mediated process (Fig. 1). This new approach allows for more flexible cloning of *trans*-splicing vectors for any given transgene [16]. In addition, this approach is more amenable for testing different combinations of splicing signals within a given context.

An essential question in designing overlapping dual vector sets is the minimal and/or optimal length of a redundant DNA sequence required for efficient homologous recombination. These parameters have not been determined for intermolecular recombination between rAAV genomes. However, based on previous studies of homologous recombination between AAV proviral plasmids [11], it is expected that larger regions of overlap will provide higher levels of recombination. However, it should be pointed out that increasing the length of overlapping sequences may eventually jeopardize the usefulness of this approach for expanding rAAV packaging capacity. For this proof-of-principle study, we have chosen a relatively large region of overlap encompassing approximately one-third of the *lacZ* DNA. The 3′ end of AV.LacZUpstream virus and

doi:10.1006/mthe.2001.0456, available online at http://www.idealibrary.com on IDEAL

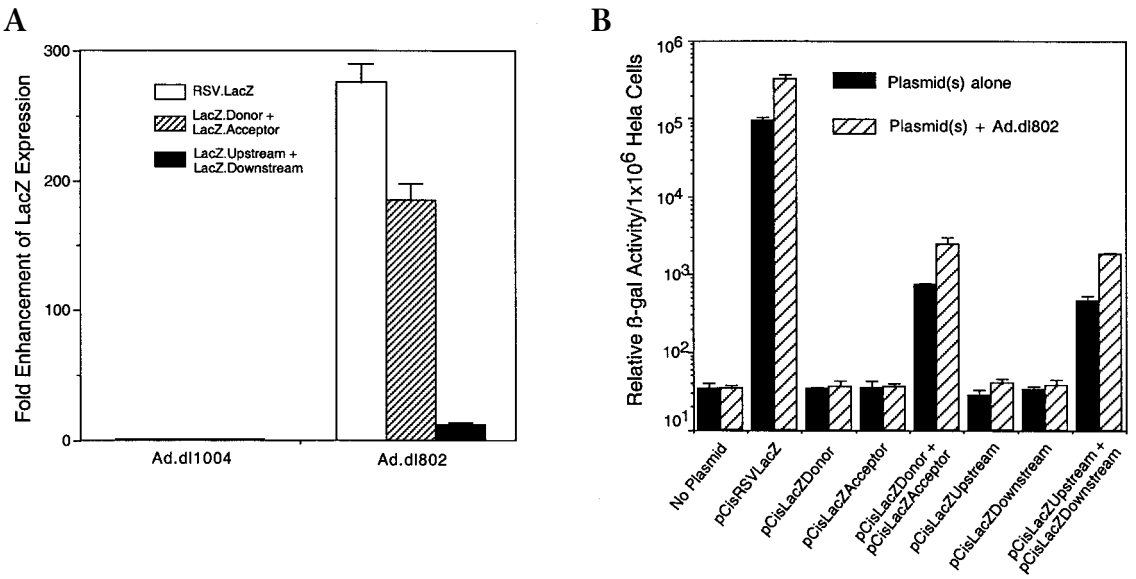

FIG. 3. Effect of adenoviral proteins on the efficiency of *trans*-splicing and overlapping dual rAAV vector approaches. To evaluate the effect of adenoviral proteins on rAAV *trans*-splicing and overlapping vector approaches, fetal fibroblasts were co-infected with dual vectors in the presence of Ad.dl802 or Ad.dl1004 mutants. β-Galactosidase activity was quantified at 2 days postinfection. (A) The fold enhancement of LacZ expression by the two different adenoviruses. Significant induction was achieved following infection with Ad.dl802 with the following fold augmentation: AV.RSVLacZ (276 ± 15-fold), *trans*-splicing dual vector set (185 ± 13-fold), and overlapping dual vector set (11.8 ± 1.8-fold). To understand the potential inductive contribution of Ad.dl802 at the transcriptional level, Hela cells were transfected with *cis* plasmids in the presence or absence of Ad.dl802. β-Galactosidase expression was quantified at 18 h post-transfection. (B) A less than fourfold increase in transgene expression by Ad.dl802 infection from pCisRSVLacZ alone, from pCisLacZDonor and pCisLacZAcceptor cotransfection, and from pCisLacZUpstream and pCisLacZDownstream cotransfection. This result indicated that promoter activation by adenoviral E4 proteins had a minor role in Ad.dl802 facilitated augmentation of *trans*-splicing and overlapping efficiency.

the 5′ end of AV.LacZDownstream virus shared this 1-kb nucleotide overlapping sequence (Fig. 1).

Functional β-Galactosidase Expression through *Trans*-splicing and Overlapping Vector Approaches in Primary Cells *in Vitro*

To compare the efficiency of *trans*-splicing and overlapping vector approaches, primary (passage 3) fetal fibroblasts were infected with the various vectors for a given set either alone or in combination at a multiplicity of infection (MOI) of 10,000 DNA particles per cell for each virus. Transgene expression was evaluated 3 days later using histochemical X-gal staining (Figs. 2A–2H) and a more quantitative chemiluminescent assay (Fig. 2I). Consistent with previous reports on *trans*-splicing, LacZ-positive cells were detected in cultures co-infected with both AV.LacZDonor and AV.LacZAcceptor vectors (Fig. 2A), but not in cells infected with either vector alone (Figs. 2B and 2C). Similarly, co-infection of cells with AV.LacZUpstream and AV.LacZDownstream produced LacZ-expressing cells (Fig. 2D) that were not evident in cells infected with each vector individually (Figs. 2E and 2F). To compare the efficiency of these two approaches, we also quantified the number of LacZ-positive cells and β-galactosidase activity in whole cell lysates. Both assays demonstrated that the *trans*-splicing method was approximately threefold more efficient than the overlapping approach in generating the full-length transgene product in these cells (Fig. 2I). To fully evaluate the practical feasibility of these two approaches, we also compared the level of transgene expression obtained using dual vectors with that of a single vector. In 293 cells, the level of protein expression from a *trans*-spliced *lacZ* was 15.8% that achieved from a single, intact, gene-encoding vector [8]. Our results evaluating efficiencies of β-galactosidase expression in reference to a full-length single vector encoding identical regulatory elements (AV.RSVLacZ) demonstrated only 7% and 2.4% for *trans*-splicing and overlapping vector approaches, respectively. The slightly lower level of *trans*-splicing efficiency in our system, compared with that previous report [8], may be due to different splicing sequences and/or target cell type specificity.

Adenoviral E4 Proteins Differentially Enhance the Efficiency of Dual Vector Gene Delivery

Although the current achievable level of transgene expression might be of therapeutic value for certain secreted proteins such as EPO [7], it may be unsatisfactory for structural proteins such as dystrophin or channels such as the cystic fibrosis transmembrane conductance regulator (CFTR). Hence, efforts to better understand the molecular events that control heterodimerization and homologous recombination may aid in developing strategies to augment the efficiency of these processes and enhance the

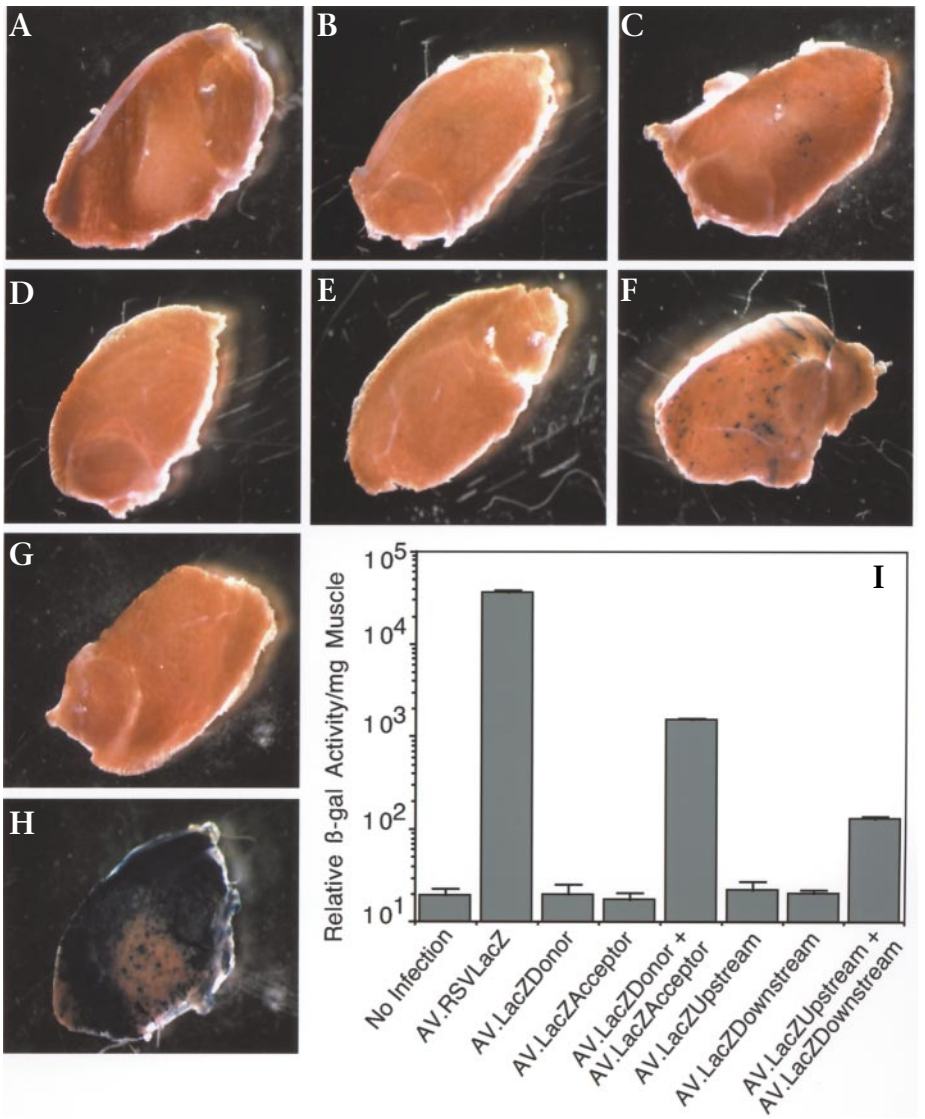

FIG. 4. *Trans*-splicing is more efficient than overlapping vectors in mouse skeletal muscle *in vivo*. To compare the efficiency of dual vector approaches *in vivo*, the anterior tibialis muscles of 7-week-old C57/BL6 mice were either injected with 5×10^{10} DNA particles of single viruses alone or co-infected with the same dose of *trans*-splicing or overlapping vector sets. Muscle tissues were harvested at 7 weeks postinfection for *in situ* X-gal staining (A–H) and lysate β-galactosidase activity assays (I). Muscles were infected with AV.LacZUpstream alone (A), AV.LacZDownstream alone (B), co-infected with both AV.LacZUpstream and AV.Lac.Downstream (C), AV.LacZDonor alone (D), AV.LacZAcceptor alone (E), or both AV.LacZDonor and AV.LacZAcceptor (F). (G) A buffer-injected control muscle. The maximal achievable level of LacZ expression is shown in an AV.RSVLacZ infected muscle in (H). LacZ-positive fibers were detected only following dual vector co-infection or in AV.RSVLacZ infected muscles. Results from enzymatic assays are shown in (I) and present more accurate quantifications from whole muscle lysates. Data in (I) represent the mean (± S.E.M.) of four independent muscle samples. β-Galactosidase expression from the *trans*-splicing dual vector approach is approximately 11-fold higher than that from the overlapping dual vector approach.

usefulness of dual vector approaches. To this end, we have evaluated whether E2a and/or E4 adenoviral proteins affect the efficiencies of gene delivery using these two approaches.

Adenovirus is the most commonly used helper virus for AAV production. Several adenoviral proteins, including E1 and E4 proteins, facilitate rAAV genome strand conversion through lytic phase replication mechanisms and enhance AAV transduction [17,18]. These adenoviral proteins also affect the formation of circular AAV transduction intermediates [19,20]. To test whether adenoviral proteins can also modulate *trans*-splicing and overlapping vector mediated gene delivery, we co-infected cells with different adenoviral mutants at the time of rAAV dual vector infection. Ad.dl802 contains a deletion in the E2a region that leads to a complete loss of the multifunctional adenoviral DNA-binding protein (DBP) [21]. In contrast, Ad.dl1004 has most of the E4 open reading frames (ORF) deleted (except of E4 ORF1) [22]. Consistent with previous reports, co-administration of Ad.dl802 increased transduction by 276 ± 15-fold in AV.RSV.LacZ infected cells (Fig. 3) [17,20]. Ad.dl802 co-infection also increased transgene expression by 185 ± 13- and 11.8 ± 1.8-fold with *trans*-splicing and overlapping dual vector sets, respectively. Controls co-infected with Ad.dl802 and the individual vectors from each set gave no transgene expression (data not shown). Additionally, no significant induction was obtained in any rAAV vector group following Ad.dl1004 co-infection (Fig. 3A).

It has been suggested that adenoviral E4 ORF3 enhances transgene expression from heterologous viral promoters at the transcriptional level [14,15]. To distinguish between E4-mediated effects leading to enhanced viral concatamerization/splicing from those involving transcriptional activation, we carried out control

doi:10.1006/mthe.2001.0456, available online at http://www.idealibrary.com on IDEAL

experiments in Hela cells using proviral plasmid transfection. Infection with Ad.dl802 resulted in a 3.4-fold enhancement of β-galactosidase expression from the proviral plasmid pCisRSV.LacZ (Fig. 3B). However, this effect was significantly lower than the 276-fold enhancement observed following infection with a single virus encoding the full-length β-galactosidase gene (Fig. 3A). Consistent with this observation, transgene expression was also increased by Ad.dl802 when *cis* plasmids for *trans*-splicing (3.25-fold increase) and overlapping (3.99-fold increase) vector sets were evaluated. These data indicate that a less than fourfold augmentation can be attributed to the increased activity of the transcription unit in our vectors. Furthermore, these findings demonstrate that a low level of recombination between proviral plasmids can also facilitate transgene expression from *trans*-splicing and overlapping vector sets in the absence of Ad.dl802 infection. Given the fourfold background of transcriptional activation in the presence of Ad.dl802, we conclude that *trans*-splicing mechanisms are more efficiently augmented by E4 proteins (181-fold) compared with homologous recombination mechanisms in overlapping vectors (eightfold). These data suggest that adenoviral E4 proteins differentially affect processes involved in intermolecular concatamerization/splicing and homologous recombination.

Comparing *Trans*-splicing and Overlapping Dual Vector Approaches in Mouse Skeletal Muscle *in Vivo*

To compare the efficiency of rAAV dual vector *trans*-splicing and overlapping approaches for gene delivery *in vivo*, we injected the anterior tibialis muscle of C57/BL6 mice with 5×10^{10} particles of various recombinant viruses either individually or in combination. Previous studies evaluating rAAV *trans*-splicing in mouse skeletal muscle have demonstrated protein expression from spliced EPO and *lacZ* genes [7,8], but the relative efficiency of *trans*-splicing in comparison with a single vector was not determined. To gain more information on the efficiency of *trans*-splicing *in vivo*, muscles were harvested at 7 weeks postinfection and evaluated by whole-mount X-gal histochemical staining of tissue sections and lysate enzymatic assays. In agreement with our *in vitro* results from fetal fibroblasts, maximal levels of β-galactosidase expression were achieved in AV.RSVLacZ infected tissues (Figs. 4H and 4I). LacZ-positive fibers were also detected in tissues co-infected with *trans*-splicing and, to a lesser extent, in overlapping vector sets (Fig. 4). Transgene expression was not detected in any muscles infected with individual vectors that encoded part of *lacZ*. Evaluation of β-galactosidase enzyme activity demonstrated that *trans*-splicing vectors gave rise to 4.28% of the activity seen following a single AV.RSVLacZ infection. In contrast, co-infection with overlapping rAAV vectors demonstrated a significantly reduced efficiency, at 0.37% that seen in muscles infected with AV.RSVLacZ alone. These results demonstrate that dual vector approaches with the *trans*-splicing design are approximately 12-fold more effective at reconstituting an intact gene product than overlapping vectors.

DISCUSSION

Recent studies on rAAV transduction biology have significantly enriched our understanding of AAV entry, intracellular processing, and gene conversion events that influence the efficiency of this virus as a gene delivery vector. Cellular receptors and/or co-receptors for type-2, type-4, and type-5 AAV have been identified [23–26]. Strategies are also being developed to overcome rate-limiting steps in rAAV entry and intracellular processing [27,28]. Despite these advances, the packaging limitation remains a major impediment to the wide application of this vector system. The development of dual vector *trans*-splicing and overlapping technologies have broadened the range of diseases to which rAAV could be applied. Important to the utility of these approaches are several unanswered questions. First, the relative efficiency of such dual vector approaches will clearly affect their utility for disease intervention. Here we have quantitatively compared the relative efficiency of dual and single vector approaches. Second, the mechanisms that control intermolecular concatamerization and homologous recombination are largely unknown. Information in this regard may ultimately lead to optimization of these approaches and improve efficiency of gene delivery.

Our results comparing *trans*-splicing and overlapping vectors have demonstrated the feasibility of both approaches. Both *in vitro* and *in vivo* studies have demonstrated that *trans*-splicing mediates higher levels of transgene reconstitution than overlapping vectors with 1 kb of overlapping sequence. Several relevant models of rAAV genome conversion differentially influence the manner in which dual vector technologies are perceived. The oldest model involves self-priming through the ITRs and subsequent elongation through a mechanism similar to wild-type AAV genome replication. In the context of rAAV, this mechanism of genome conversion has been implicated as the predominant pathway for augmentation of rAAV-mediated transduction by adenoviral proteins [17,18]. A second, more recent, model involving annealing of plus (+) and minus (–) single-stranded rAAV genomes has been suggested as the molecular basis for assembling transcriptional active duplex forms in hepatocytes [29]. A third model proposes that circular double-stranded and/or single-stranded genomes may be an early conversion product of rAAV [30] and that time-dependent concatamerization may result from ITR-mediated intermolecular recombination [5]. Regardless of how duplex-form rAAV genomes are generated, it is commonly accepted that both circular and linear episomal forms can exist within cells as large head-to-tail concatamers [30,31]. Whether the preintegration intermediate is linear or circular remains to be conclusively determined. These studies have set the

foundation for the development of dual vector *trans*-splicing technology.

Several possibilities exist for the mechanism underlying homologous recombination between independent rAAV genomes. For example, homologous DNA sequence exchange could occur between circular and/or linear double-stranded viral genomes. Alternatively, single-stranded viral DNA genomes could serve as a template by which double-stranded genomes are modified by a recombinational event. It is unlikely that the partial annealing between + and – strands of two overlapping vectors could reconstitute a full-length functional transgene in the absence of some mechanism for precisely removing the intervening ITR sequences.

Both viral and nonviral stimuli have been evaluated for their ability to enhance the efficiency of rAAV transduction. Here we found that adenoviral E4 gene products also enhanced dual vector transduction efficiency, most notably in the *trans*-splicing vector approach. The underlying mechanism for this finding is not readily apparent because E4 proteins have been associated with increased head-to-head and tail-to-tail R_f (replication form) concatamer formation. These R_f structures should not reconstitute a *trans*-spliced message. Furthermore, transcriptional enhancement of the promoters in our vector constructs by E4 ORF3 seems to have only a minor role in this induction. Previous studies have demonstrated that E4 proteins have little effect on the formation of circular intermediates, hence this pathway is not likely responsible for the augmentation [19]. Given the present findings, we conclude that E4 proteins influence alternative functions that augment gene expression from *trans*-splicing vectors. In support of this notion, it has been shown that E2a mutant Ad.dl802 also enhances intranuclear recruitment of the rAAV genome [32]. Additionally, adenoviral E4 proteins might also help facilitate the splicing process [33,34].

Despite the lack of clearly defined mechanisms for intermolecular concatamer formation and homologous recombination, both *trans*-splicing and overlapping vector approaches have their respective advantages. The cloning of overlapping vectors is relatively simple and straightforward. However, the overall packaging capacity is smaller than that of the *trans*-splicing approach. For example, in a vector containing 1 kb of regulatory elements (promoter and poly(A)) and a 2 kb of overlapping transgene sequence (~ 1/3 transgene size), the maximal transgene size accommodated by this approach would be approximately 6.4 kb. In contrast, in *trans*-splicing vectors containing 1.5 kb of regulatory elements (promoter, poly(A), and introns), the maximal transgene size would be approximately 7.9 kb. On the other hand, major obstacles in the use of *trans*-splicing vectors are the difficulty of cloning and variability in splicing efficiency within the context of a given cDNA.

At present, the efficiencies of the dual vector systems reported here are much lower than can be achieved with a single vector. Although dual vector efficiencies will never match that of the single vector, refinements of these approaches may ultimately enhance efficiencies to their theoretical maximums. For example, defining the optimal length of overlapping sequence will likely increase the efficiency of intermolecular recombination approaches. Additionally, as both dual vector approaches require coinfection of the same cell with two viral genomes, improvements in infectivity using alternative serotypes and higher titer virus will likely also increase the efficiencies of these approaches. Finally, several stimulatory molecules have been described for enhancing rAAV transduction efficiency. These include proteasome [27] and tyrosine kinase [35] inhibitors. It remains to be determined whether these exogenous agents can also enhance the efficacy of dual vector approaches.

MATERIALS AND METHODS

Proviral plasmids. Three sets of proviral plasmids were used to generate type-2 rAAV stocks. The pCisRSVLacZ plasmid contains the full-length prokaryotic *lacZ* gene under the transcriptional control of the RSV promoter and the SV40 poly(A) [27]. The AV.LacZ virus generated from this proviral plasmid was used to determine the maximal level of transgene expression from a single vector under our experimental conditions.

The second set of proviral plasmids used to produce LacZ *trans*-splicing viruses included donor (pCisRSVLacZDonor) and acceptor (pCisRSVLacZAcceptor) vectors. To generate these plasmids, the β-galactosidase gene was divided at a putative conserved splicing junction (CAG/G) between nucleotides 2742 and 2743 (the nucleotide numbering accords with the β-galactosidase gene sequence in pCMVβ from Clontech) and a PCR-mediated swapping procedure was used during plasmid construction. For pCisRSVLacZDonor, a 780-bp PCR product was obtained from the pCisRSVLacZ template using a forward primer (EL752, 5′-GTCATAGCGATAACGAGCTCCTGCAC-3′) and a reverse primer (EL753, 5′-GCGCGTCGACTATTGGTCTCCTTAAACCTGTCTTGTAACCTTGATACTTACCTGCGCCAGCTGGCAGTTCAGGCCAATCCGCGCCGG-3′). The reverse primer EL753 also contained the intronic donor sequences (underlined) from the first intron of the human β-globulin gene. This 780-bp PCR product was then used to replace an 1820-bp fragment in pCisRSVLacZ to make the donor proviral plasmid (pCisRSVLacZDonor). The final pCisRSVLacZ.Donor construct contains the RSV promoter, the 5′ portion of *lacZ* (nt 1 to 2742), and the intron splicing donor signal. To generate pCisLacZAcceptor, a 1-kb PCR product was amplified from the pCisRSVLacZ template with a forward primer (EL751, 5′-GCGCCTGCAGCTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGC-3′) and a reverse primer (EL688, 5′-GCGCCTGCAGCATACCACATTTGTAGAGGTTTTAC-3′). The upstream primer EL751 also contained intronic acceptor sequences (underlined) from the intron of the heavy chain of the human immunoglobulin gene. Finally, this PCR product was used to swap out a 1765-bp fragment in pCisAVGFP3ori to make the LacZ acceptor proviral plasmid (pCisLacZAcceptor). The complete pCisLacZAcceptor plasmid contained the 3′ portion of *lacZ* (from nt 2743 to 3144), SV40 poly(A) sequence, the ampicillin resistance gene, and a bacterial replication origin. To reduce the possibility of PCR-induced mutations, all PCR reactions were carried out with the Expand High Fidelity PCR system (Roche Diagnostic Corp., Indianapolis, IN). Furthermore, all the PCR fragments and cloning junctions in both pCisRSVLacZDonor and pCisLacZAcceptor were confirmed by DNA sequencing.

The third set of *cis* proviral plasmids was generated to test an overlapping vector strategy using intermolecular homologous recombination. In this case, the first two-thirds of *lacZ* was cloned into an upstream

doi:10.1006/mthe.2001.0456, available online at http://www.idealibrary.com on IDEAL

vector and the last two-thirds of *lacZ* was cloned into a downstream vector. To generate the pCisLacZUpstream proviral plasmid, pCisRSVLacZDonor (described above) was double digested with *Bst*XI and *Sal*I. The 7025-bp larger fragment was then re-ligated by blunt end ligation. This construct contains the RSV promoter and the first two-thirds of *lacZ* (from nt 1 to 2304). To generate the pCisLacZDownstream proviral plasmid, we first constructed an intermediate vector (pDD323) by inactivating one of the two *Sal*I sites (at nt 8375) in pCisRSVLacZ. This pDD323 plasmid was then double digested with *Sal*I (at nt 3817) and *Eco*RV and the resulting 6655-bp larger fragment was blunted and re-ligated. The pCisLacZDownstream proviral construct contains the second two-thirds of *lacZ* from nt 1207 to 3144 and the SV40 poly(A) signal. Taken together, there is 1093 bp of overlapping sequence in *lacZ* (from nt 1027 to 2299) between pCisLacZUpstream and pCisLacZDownstream.

Recombinant AAV-2 production and purification. Type-2 recombinant AAV was produced with a three-plasmid transfection method as described [36]. Briefly, 60% confluent 293 cells were co-transfected with a *cis* proviral plasmid, AAV-2 helper plasmid (pXX-2), and adenoviral helper plasmid (pXX6-80) in a ratio of 1:1:3. The viral lysate was harvested at 50 h post-transfection with the aid of a cell scraper. After centrifugation for 15 min ($\times$ 500*g*), the cell pellet was resuspended in 10 mM Tris, pH 8.0. Viral particles were then released by physical stresses (including three rounds of freeze-thaw and five passages through a 25G needle) and chemical treatments (including DNaseI digestion, trypsin digestion, and deoxycholate incubation) as described [16]. Finally, the viral lysate was diluted in 20 mM Tris-Cl, pH 8.0, 150 mM NaCl, and 0.5% sodium deoxycholate and loaded on a Poros heparin column (PerSeptive, Applied Biosystems) in a Beckman Biosys 2000 HPLC Workstation. Recombinant AAV-2 was eluted with a linear NaCl gradient and viral fractions were pooled and dialyzed against HEPES buffer (20 mM Hepes, 150 mM NaCl, pH 7.8). Contamination by wild-type AAV-2 was determined as described to be less then one functional particle per 1×10^{10} rAAV particles [7].

Viral infection of fetal fibroblasts. To analyze the efficiency of *trans*-splicing and overlapping vectors *in vitro*, we used primary cultured fetal fibroblasts as described [20]. Briefly, cells (passage 3) were cultured in high glucose DMEM (Dulbecco's modified Eagle medium) supplemented with 10% FBS (fetal bovine serum), penicillin (100 U/ml), and streptomycin (100 μg/ml). Cells were plated at a density of 5×10^5 cells per well in 6-well plates the day before infection. Immediately before the infection, cells were washed with serum-free medium. AAV infection was also performed in serum-free medium at a MOI of 10,000 DNA particles/cell for each virus. At 1 h postinfection, the final concentration of FBS was increased to 10% by the addition of serum. Cell cultures were collected for analysis of transgene expression at 72 h postinfection. In experiments in which adenoviral gene products were evaluated, the mutant adenovirus Ad.dl802 [21] or Ad.dl1004 [22] was applied to cells at a MOI of 1000 particles/cell at the time of rAAV infection.

β-Galactosidase assays. We used two different approaches to evaluate β-galactosidase expression in virally transduced cells. The histochemical localization of β-galactosidase activity was evaluated using an X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) staining protocol as described [27]. Briefly, cells were fixed in 0.5% glutaraldehyde and then stained in X-gal staining solution (1 mg/ml X-gal, 5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$ and 1 mM $MgCl_2$ in phosphate-buffered saline (PBS)) for 2 h at 37°C. To decrease the background staining from endogenous β-galactosidase, 5 mM EGTA was also added during fixation [37]. The quantitative evaluation of β-galactosidase expression in cell lysates was performed with a Galacto-Light Plus system (Applied Biosystems, Bedford, MA) according to the manufacturer's recommendation. Briefly, cells cultured in 6-well plates were rinsed with PBS twice, then lysed in 200 μl Tropix lysis buffer (Applied Biosystems, Bedford, MA). Following a 2-min spin at 12,000*g* in a bench-top microcentrifuge, supernatants were collected as cell lysates. The endogenous eukaryotic β-galactosidase activity was then inactivated by heating cell lysates at 48°C for 60 min. Finally the β-galactosidase activity was measured with a Luminometer (TD-20/20; Turner Designs Instrument, Sunnyvale, CA) at a sensitivity of 60% (for fetal fibroblasts) or 70% (for Hela cells).

Animal experiments. All animal experiments were carried out in accordance with NIH and institutional guidelines of the University of Iowa. C57/BL6 mice were purchased from The Jackson Laboratory. The *in vivo* transduction efficiency was examined in the anterior tibialis muscle as described [30]. To minimize the effect of the injection volume on transgene expression, all the viruses (single or co-infection) were delivered in 40 μl Hepes-buffered saline (HBS). To fully evaluate the efficiency of different approaches, four groups of mice (four mice per group) were infected with different viral constructs. In group 1, the left anterior tibialis was infected with 5×10^{10} particles of control virus AV.RSVLacZ. The β-galactosidase expression from this group reflected the maximal achievable level. The right anterior tibialis muscles in this group were injected with an equal volume of HBS to serve as a negative control for the background β-galactosidase activity. In group 2, the left anterior tibialis was infected with AV.LacZDonor and the right anterior tibialis was infected with AV.LacZAcceptor. In group 3, the left muscle of each mouse was infected by AV.LacZUpstream and the right side of the same mouse was infected with AV.LacZDownstream. In groups 2 and 3, 5×10^{10} particles of each indicated virus were infused in each individual muscle. The group 4 mice were co-infected with two vectors in each muscle. In particular, the left muscles were co-infected with 5×10^{10} particles of AV.LacZDonor and 5×10^{10} particles of AV.LacZAcceptor. The right muscles were co-infected with 5×10^{10} particles of AV.LacZUpstream and 5×10^{10} particles of AV.LacZDownstream. Results from this group indicated the relative efficiency for either of the two approaches. All the viral infected mice were sacrificed at 7 weeks postinfection and β-galactosidase expression in the anterior tibialis muscles was analyzed as described below. *In situ* X-gal staining of the fresh muscle tissue was performed *en bloc* with 1-mm thick *trans*-sectional tissue specimens obtained from the center of the anterior tibialis muscle. After 30 min fixation in 0.5% gluteraldehyde, tissue samples were incubated in 1 mM $MgCl_2$/PBS for an additional 30 min. Finally, the muscle tissues were stained in X-gal solution for 15 h at room temperature. To quantify β-galactosidase activity, the entire anterior tibialis muscle was frozen in liquid nitrogen and pulverized by hand grinding with an ice-cold porcelain mortar and pestle. The muscle powder was then homogenized in 600 μl of 1× Reporter lysis buffer (Promega) by additional grinding. The crude lysates were spun for 2 min at 16,000*g* and the supernatants were used for β-galactosidase assay as described above.

ACKNOWLEDGMENTS

We thank Terry Ritchie (University of Iowa) for editorial assistance in the preparation of this manuscript. This work was supported by NIH grant HL58340 (J.F.E. and D.D.), the Center for Gene Therapy co-funded by NIH and the Cystic Fibrosis Foundation (P30 DK54759; J.F.E.), and research grants from the Muscular Dystrophy Association (D.D.) and Targeted Genetics Corp. (J.F.E.).

RECEIVED FOR PUBLICATION MAY 30; ACCEPTED JULY, 20 2001.

REFERENCES

1. Kay, M. A., Glorioso, J. C., and Naldini, L. (2001). Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat. Med.* **7:** 33–40.
2. Kay, M. A., *et al.* (2000). Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* **24:** 257–261.
3. Acland, G. M., *et al.* (2001). Gene therapy restores vision in a canine model of childhood blindness. *Nat. Genet.* **28:** 92–95.
4. Duan, D., Yan, Z., Yue, Y., and Engelhardt, J. F. (1999). Structural analysis of adeno-associated virus transduction intermediates. *Virology* **261:** 8–14.
5. Yang, J., *et al.* (1999). Concatamerization of adeno-associated viral circular genomes occurs through intermolecular recombination. *J. Virol.* **73:** 9468–9477.
6. 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:** 595–598.
7. Yan, Z., Zhang, Y., Duan, D., and Engelhardt, J. F. (2000). Trans-splicing vectors expand the utility of adeno-associated virus for gene therapy. *Proc. Natl. Acad. Sci. USA* **97:** 6716–6721.
8. Sun, L., Li, J., and Xiao, X. (2000). Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. *Nat. Med.* **6:** 599–602.
9. 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. *Nat. Biotechnol.* **18:** 527–532.

10. Russell, D. W., and Hirata, R. K. (1998). Human gene targeting by viral vectors. *Nat. Genet.* **18:** 325–330.
11. Senapathy, P., and Carter, B. J. (1984). Molecular cloning of adeno-associated virus variant genomes and generation of infectious virus by recombination in mammalian cells. *J. Biol. Chem.* **259:** 4661–4666.
12. Taylor, A. F. (1992). Movement and resolution of Holliday junctions by enzymes from E. coli. *Cell* **69:** 1063–1065.
13. Samulski, R. J. (2000). Expanding the AAV package. *Nat. Biotechnol.* **18:** 497–498.
14. Lusky, M., *et al.* (1999). Regulation of adenovirus-mediated transgene expression by the viral E4 gene products: requirement for E4 ORF3. *J. Virol.* **73:** 8308–8319.
15. Yew, N. S., *et al.* (1999). Increased duration of transgene expression in the lung with plasmid DNA vectors harboring adenovirus E4 open reading frame 3. *Hum. Gene Ther.* **10:** 1833–1843.
16. Duan, D., Yue, Y., Yan, Z., and Engelhardt, J. F. (2001). Trans-splicing vectors expand the packaging limits of adeno-associated virus for gene therapy applications. In *Methods in Molecular Medicine, Viral Vectors for Gene Therapy Methods and Protocols* (C. A. Machida, Ed.), Humana Press, Inc., Totowa, NJ, in press.
17. Fisher, K. J., *et al.* (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J. Virol.* **70:** 520–532.
18. Ferrari, F. K., Samulski, T., Shenk, T., and Samulski, R. J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.* **70:** 3227–3234.
19. Duan, D., *et al.* (1999). Formation of adeno-associated virus circular genomes is differentially regulated by adenovirus E4 ORF6 and E2a gene expression. *J. Virol.* **73:** 161–169.
20. Sanlioglu, S., Duan, D., and Engelhardt, J. F. (1999). Two independent molecular pathways for recombinant adeno-associated virus genome conversion occur after UV-C and E4orf6 augmentation of transduction. *Hum. Gene Ther.* **10:** 591–602.
21. Rice, S. A., and Klessig, D. F. (1985). Isolation and analysis of adenovirus type 5 mutants containing deletions in the gene encoding the DNA-binding protein. *J. Virol.* **56:** 767–778.
22. Bridge, E., and Ketner, G. (1989). Redundant control of adenovirus late gene expression by early region 4. *J. Virol.* **63:** 631–638.
23. Summerford, C., Bartlett, J. S., and Samulski, R. J. (1999). αVβ5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat. Med.* **5:** 78–82.
24. Qing, K., *et al.* (1999). Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat. Med.* **5:** 71–77.
25. Walters, R. W., *et al.* (2001). Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J. Biol. Chem.* **21:** 21.
26. Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J., and Chiorini, J. A. (2001). Adeno-associated virus serotype 4 (aav4) and aav5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J. Virol.* **75:** 6884–6893.
27. Duan, D., Yue, Y., Yan, Z., Yang, J., and Engelhardt, J. F. (2000). Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus. *J. Clin. Invest.* **105:** 1573–1587.
28. Hansen, J., Qing, K., and Srivastava, A. (2001). Adeno-associated virus type 2-mediated gene transfer: altered endocytic processing enhances transduction efficiency in murine fibroblasts. *J. Virol.* **75:** 4080–4090.
29. Nakai, H., Storm, T. A., and Kay, M. A. (2000). Recruitment of single-stranded recombinant adeno-associated virus vector genomes and intermolecular recombination are responsible for stable transduction of liver in vivo. *J. Virol.* **74:** 9451–9463.
30. Duan, D., *et al.* (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.
31. 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. USA* **98:** 4084–4088.
32. Weitzman, M. D., Fisher, K. J., and Wilson, J. M. (1996). Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. *J. Virol.* **70:** 1845–1854.
33. Nissim-Rafinia, M., Chiba-Falek, O., Sharon, G., Boss, A., and Kerem, B. (2000). Cellular and viral splicing factors can modify the splicing pattern of CFTR transcripts carrying splicing mutations. *Hum. Mol. Genet.* **9:** 1771–1778.
34. Estmer Nilsson, C., *et al.* (2001). The adenovirus E4-ORF4 splicing enhancer protein interacts with a subset of phosphorylated SR proteins. *EMBO J.* **20:** 864–871.
35. Mah, C., *et al.* (1998). Adeno-associated virus type 2-mediated gene transfer: role of epidermal growth factor receptor protein tyrosine kinase in transgene expression. *J. Virol.* **72:** 9835–9843.
36. Xiao, X., Li, J., 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.
37. Kaplitt, M. G., *et al.* (1994). Long-term gene expression and phenotypic correction using adeno-associated virus vectors in the mammalian brain. *Nat. Genet.* **8:** 148–154.