

Señora

Directora de la Dirección de Nuevas Creaciones

Superintendencia de Industria y Comercio

Ministerio de Comercio, Industria y Turismo

E. S. D.

Re.: Expediente No. **NC2017/0011344**

Solicitud de Patente PCT entrada en Fase Nacional, para:

“MÉTODO PARA PRODUCIR VECTORES ADENO-ASOCIADOS DE GRAN TAMAÑO EMPLEANDO LÍNEAS CELULARES PRODUCTORAS” (nuevo título)

Solicitante: **GENZYME CORPORATION**

RESPUESTA A EXAMEN DE FONDO

Yo, **MARGARITA CASTELLANOS ABONDANO**, mayor de edad, identificada y domiciliada como aparece al pie de mi firma, obrando en mi calidad de apoderada de la sociedad solicitante del asunto citado en la referencia, en atención a su oficio No. **14563** del 24 de noviembre de 2020, atentamente me permito dar respuesta al mismo de la siguiente manera:

1. Modificación al título de la solicitud

Siguiendo la amable sugerencia del examinador, se ha modificado el título de la solicitud para señalar:

“MÉTODO PARA PRODUCIR VECTORES ADENO-ASOCIADOS DE GRAN TAMAÑO EMPLEANDO LÍNEAS CELULARES PRODUCTORAS”

De esta manera, se refleja con mayor claridad el objeto reivindicado.

2. Modificación a las Reivindicaciones

De acuerdo con lo establecido en el Artículo 34 de la Decisión 486 y en vista de las observaciones realizadas por el examinador, se presenta en esta oportunidad un nuevo pliego reivindicatorio conformado por 38 reivindicaciones, en donde se han realizado las siguientes modificaciones:

- Atendiendo la amable sugerencia del examinador, se ha modificado el contenido de las reivindicaciones 1 y 26, a fin de indicar como característica

esencial que el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño, se integran de forma estable en el genoma de la línea celular productora.

- Se ha corregido el error respecto a la presencia de la letra d en la reivindicación 2.
- Del mismo modo, siguiendo la amable sugerencia del examinador, se ha incorporado el proceso de obtención de la composición en la reivindicación 15.

No se ha adicionado materia nueva, por medio de estas reivindicaciones modificadas.

3. Recibo por concepto de modificación de solicitud en trámite

Teniendo en cuenta lo establecido en la Resolución No. 59669 del 2020, adjunto se presenta recibo por concepto de modificación de solicitud en trámite.

4. Claridad del pliego reivindicatorio

De conformidad con las modificaciones introducidas a las reivindicaciones, puede establecerse que las nuevas reivindicaciones son claras y concisas. En particular:

- El examinador ha objetado el contenido de las reivindicaciones 1 y 26, indicando que comprenden como característica opcional, una característica esencial: *“el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora”*.

Según se indicó arriba, las nuevas reivindicaciones 1 y 26 han sido modificadas siguiendo la amable sugerencia del examinador, presentando ahora la característica señalada como esencial.

- El examinador señala que la reivindicación 2 no es clara al contener una letra d en su contenido.

Teniendo en cuenta las modificaciones introducidas, la nueva reivindicación 2 ya no contiene dicha letra.

- El examinador considera que la reivindicación 15 es clara al no incluir el proceso de obtención de la composición.

La nueva reivindicación 15 incorpora el proceso de obtención de la composición, siguiendo la amable sugerencia del examinador.

De esta manera, se superan las objeciones formuladas y puede establecerse que la presente solicitud cumple con los requerimientos del Artículo 30 de la Decisión 486.

5. Nivel inventivo

El examinador considera que la combinación de las enseñanzas de los documentos D2: WO2015/038625 y D3: "Complete Correction of Hemophilia A with Adeno-associated Viral Vectors Containing a Full-Size Expression Cassette" hace obvio el objeto de las anteriores reivindicaciones 1 a 38.

A este respecto, se desea manifestar nuestro desacuerdo con la opinión del examinador, por cuanto el objeto de la presente solicitud no resulta obvio a la luz de la combinación de D2 y D3, por las siguientes razones:

Como se señaló anteriormente, las reivindicaciones 1 y 26 han sido modificadas para recitar, en parte, *"en donde el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora"* (énfasis adicionado). Igualmente, la reivindicación 15 ha sido modificada para incluir el método de obtención de la composición que comprende partículas de rAAV. Es importante entonces indicar que la característica según la cual, *"el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora"*, **no** es opcional en las nuevas reivindicaciones 1, 15 y 26.

Teniendo en cuenta lo anterior, las reivindicaciones modificadas son inventivas sobre D2 y D3 porque D3 describe la producción de partículas de rAAV por transfección triple y señala el fracaso de dicho método para producir partículas de rAAV con genomas de gran tamaño. Además, se sabía en la técnica que la triple transfección no permitía obtener, de forma reproducible, partículas de rAAV con genomas de gran tamaño. Por su parte, D2 no permite superar las deficiencias de D3, y la combinación de D2 y D3 no proporciona enseñanza o sugerencia alguna sobre un método, como se reivindica, el cual permite encapsidar con éxito partículas de rAAV con genomas de gran tamaño.

5.1. D3 describe la producción de partículas de rAAV por triple transfección y la incapacidad de encapsidar partículas de rAAV con genomas de gran tamaño

En primer lugar, se observa que los experimentos descritos en D3 implican producir partículas de rAAV mediante triple transfección. La triple transfección es un método en el que tres plásmidos **separados** se transfectan en una línea celular, que incluye (1) un plásmido que codifica los genes rep y cap de AAV, (2) un plásmido auxiliar con genes de un virus auxiliar y (3) el vector rAAV para ser empaquetado. Específicamente, la sección Materiales y Métodos de D3 establece:

"Preparación del vector AAV"

El vector de AAV recombinante con el gen del factor VIII se produjo mediante el método de cotransfección de plásmido triple como se describió anteriormente (Hauck et al., 2003)" (ver D3, página 649).

El método de triple transfección de D3 no es lo mismo que un método que implica el cultivo de una línea celular productora de AAV en condiciones para generar partículas de rAAV, en donde el ácido nucleico que codifica los genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño **se integran de forma estable en el genoma de la línea celular productora**, como se reivindica.

El examinador afirma que D3 enseña un método en donde el vector AAV tiene un tamaño de 5,7 kb (ver tabla de comparación entre las características técnicas esenciales de la invención reivindicada, D2 y D3). Sin embargo, no estamos de acuerdo. Si bien D3 describe un **intento** de empaquetar dicho vector, lo que es más importante, D3 señala el fracaso del método de triple transfección para producir realmente partículas de rAAV con genomas de gran tamaño. Por ejemplo, D3 señala:

*“Para demostrar el tamaño de los genomas de AAV empaquetados en los viriones de AAV, realizamos un análisis en gel de agarosa alcalina. Como se muestra en la Fig. 2D, similar al AAV-TTRFVIII de tamaño estándar, la mayoría del ADN monocatenario empaquetado a partir de pAAV-CB-FVIII era de aproximadamente 5 kb. Se detectaron pocos ADN de AAV en el rango de 5,75 kb. **Estos resultados sugirieron que no todas las secuencias de vector estaban completamente empaquetadas en cápsides de AAV.** La amplia gama de frotis para AAV-CB-FVIII también confirmó que una mezcla heterogénea de genomas con diferentes longitudes fue empaquetada en cápsides de AAV”* (ver D3, página 652, énfasis adicionado).

Los autores de D3 señalan además que probablemente hubo delecciones en el genoma del vector durante el proceso de encapsidación, y especulan que la complementación entre las partículas de AAV con diferentes porciones del vector encapsulado puede explicar el nivel de expresión transgénica que observaron:

*“Como se muestra en la Fig. 2, **la secuencia de ADN en el virión parecía estar en el rango de 5 kb, muy por debajo del supuesto genoma del vector de 5,75 kb. Esto sugirió que hubo delecciones de secuencias de vectores durante el proceso de encapsidación.** La amplia gama de densidades de los vectores también sugirió una mezcla heterogénea de genomas de varias longitudes, lo que fue confirmado por análisis de gel alcalino. Es interesante observar que los ADN de vector con delecciones pudieron complementarse entre sí y regenerar casetes de expresión de longitud completa, lo que condujo a un nivel significativo de expresión transgénica (Fig. 3)”*. (ver D3, páginas 652-653, énfasis adicionado).

Por último, los autores de D3 concluyen que *“puede ser necesario un desarrollo futuro para mejorar el rendimiento de empaquetamiento de vectores de AAV con viriones grandes que contienen genoma”* (ver D3, página 653).

Tomado en conjunto, D3 no enseña un método o línea celular para producir una partícula de AAV que comprenda un genoma de AAV recombinante de gran tamaño, según afirma el examinador. Además, como se describe en detalle a continuación, D3

no es el único que demuestra el fracaso de los métodos de triple transfección para producir partículas de AAV con genomas de gran tamaño.

5.2. La triple transfección no permite obtener, de forma reproducible, partículas de rAAV con genomas encapsidados y de gran tamaño

El uso de un método de triple transfección para intentar producir partículas de rAAV con genomas de gran tamaño, como se describe en D3, no hace que la invención reivindicada sea obvia porque en el momento en que se allegó la presente solicitud, una persona del oficio normalmente versada en la materia habría entendido que los intentos de producir partículas de rAAV con genomas relativamente grandes fueron generalmente infructuosos. Esto se indica en los “Antecedentes de la invención” de la presente solicitud, en donde se establece:

“Los primeros estudios definen el límite de empaquetamiento en 4,7 a 4,8 kb (Dong, J-Y et al. (1996) Human Gene Therapy 7:2101-2112). Estudios más recientes han confirmado un límite de genomas de vector empaquetados aproximadamente en 5,0 a 5,2 kb de tamaño para cápsides de AAV2, 30 AAV5 o AAV8. En estos estudios, los genomas de gran tamaño (o “fragmentados”) de ambas polaridades fueron suprimidos normalmente en el extremo 5' y los genomas más empaquetados no excedieron de ~5,2 kb (Lu, H. et al. (2008) Human Gene Therapy 19:648-654; Wu, Z. et al. (2010) Molecular Therapy 18:80-86; Grose, W.E. et al. (2012) PLoS One 7:e39233).” (ver párrafo [0004] de la Descripción).

De hecho, antes de que se allegara la presente solicitud, varios estudios científicos habían demostrado los desafíos de producir partículas de rAAV con genomas de gran tamaño. Primero, Dong et al., en “*Characterization of Genome Integrity for Oversized Recombinant AAV Vector*” (Mol Ther. 2010 Ene; 18 (1): 87-92), (adjunto al presente memorial) intentaron empaquetar vectores AAV de hasta 7,2 kb de tamaño. Dong et al. concluyeron que **“nuestros resultados y otros estudios respaldaron que los genomas de AAV empaquetados eran <5 kb. Sin embargo, a través de otro mecanismo molecular, muy probablemente, la conversión del genoma de AAV a partir de una secuencia superpuesta, los vectores de AAV producidos a partir de genomas de AAV de gran tamaño pudieron dar lugar a una expresión transgénica significativa”** (ver Dong et al., Página 90, último párrafo de Discusión, énfasis adicionado).

Además, Wu et al. en “*Effect of genome size on AAV vector packaging*” (Mol Ther. 2010 Ene; 18 (1): 80-6) (adjunto al presente memorial), intentaron empaquetar vectores de AAV con una longitud de 4,7 a 8,7 kb, utilizando proteínas de la cápside de los serotipos de AAV 2, 5 y 8. Wu indica: **“descubrimos que el tamaño de los ADN del genoma del vector empaquetado nunca excedía los 5,2 kb, independientemente del tamaño del vector codificado por el plásmido de producción. Esto fue cierto para todos los tipos de cápside y vectores probados. El ADN del genoma del vector empaquetado de los vectores codificados por plásmidos que superan los 5,2 kb constaba de una población heterogénea de genomas con una distribución de tamaño amplia y una longitud máxima de ~ 5 kb. La heterogeneidad se generó por la**

eliminación en el extremo 5' (ver Wu et al., Página 80, último párrafo de la introducción, énfasis adicionado). Wu et al. señalaron: “tampoco está claro por qué nuestros datos no concuerdan con informes anteriores de que los genomas de vectores de hasta 8,9 kb pueden empaquetarse y son funcionales para la transducción. Para controlar las diferencias en los plásmidos de empaquetamiento y las técnicas de purificación de vectores, repetimos nuestros experimentos con otros dos conjuntos de plásmidos de empaquetamiento y otro protocolo de purificación sin cambios en el resultado (datos no mostrados)” (ver Wu et al., Página 85, segundo al último párrafo de la columna izquierda).

De manera similar, Lai et al., en “Evidence for the Failure of Adeno-associated Virus Serotype 5 to Package a Viral Genome ≥ 8.2 kb” (Mol Ther. 2010 Ene; 18 (1): 75-9), probaron específicamente la idea que las cápsides de AAV5 tienen una capacidad elevada para empaquetar genomas de gran tamaño, como se informó en Allocca et al. (Revista de investigación clínica 2008 118 (5): 1955-64). Lai et al. señalaron “nuestros resultados sugieren que AAV-5 no puede empaquetar un genoma $\geq 8,2$ kb. De manera similar a AAV-2, la partícula infecciosa de AAV-5 ITR única se genera cuando el genoma viral excede la capacidad máxima de empaquetamiento (Figura 5). Sin embargo, la producción de dicho virión ITR único es sustancialmente menos eficaz. Sospechamos que un mecanismo similar también puede explicar (al menos parcialmente) los hallazgos de Allocca et al. **Especulamos que el virus AAV-5 generado por Allocca et al. puede contener genomas virales parcialmente empaquetados.** Una posible superposición de secuencias entre los virus que portan el genoma del extremo 5' y el extremo 3' puede reconstituir la expresión mediante recombinación homóloga (Figura 5b). Debido a que la reconstitución basada en superposición depende de la secuencia del gen diana, la falta de expresión de minidistrofina probablemente refleje una pobre eficiencia de recombinación del minigen $\Delta H2-R15$ (Figura 4)” (ver Lai, página 77, columna de la derecha, énfasis adicionado). Lai et al. concluyeron, “Junto con los resultados de Dong et al. y Wu et al. creemos que AAV-5 puede no encapsidar un genoma viral $\geq 8,2$ kb. La capacidad máxima de carga de una sola partícula de AAV sigue siendo limitada” (ver Lai, et al., Página 77, columna derecha, énfasis adicionado).

Se tiene entonces que, al igual que D3, tanto Dong et al., como Wu et al. y Lai et al. describen la producción de partículas de rAAV mediante triple transfección, como se describió anteriormente. En conjunto, dados los hallazgos de Dong et al., Wu et al., Y Lai et al., una persona del oficio normalmente versada en la materia no habría sido conducida a la invención reivindicada, simplemente siguiendo las enseñanzas de D2 y/o D3. Por el contrario, considerando la falta de reproducibilidad en la técnica y los desafíos persistentes de producir partículas de rAAV con genomas de gran tamaño, **no** habría sido obvio, para una persona del oficio normalmente versada en la materia, llegar al método de la reivindicación 1, la composición de la reivindicación 15, o la línea celular de la reivindicación 26.

5.3. D2 no supera las deficiencias de D3

El documento D2 no supera las deficiencias de D3. El examinador afirma que la integración estable de los genes rep y cap de AAV en el genoma de las células

productoras, ya se describe en el documento D2 y argumenta que la persona del oficio normalmente versada en la materia estaría motivada para integrar los genes rep y cap en el genoma de la célula productora, como se sugiere en el documento D2, dentro del método descrito en D3. No estamos de acuerdo con esta opinión.

Primero, se observa que D2 no contiene información alguna con respecto a la integración estable del ácido nucleico que codifica el genoma de rAAV de gran tamaño en el genoma de una línea celular productora. En consecuencia, D2 no enseña ni sugiere un método, composición o línea celular en donde *“el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora”*.

En segundo lugar, el examinador afirma que en una modalidad particular de D2: *“El vector comprende el constructo DC172 2xHCR ATG FVIII que tiene un tamaño de 6962 pares de bases (6.9 kb)”*. Sin embargo, D2 no proporciona ejemplos prácticos sobre cómo utilizar realmente el vector de 6,9 kb, y mucho menos producir una partícula de AAV que comprende un genoma de rAAV de gran tamaño utilizando el vector de 6,9 kb. Por tanto, D2 no enseña ni sugiere un método, composición o línea celular que implique una línea celular productora de AAV que comprenda un genoma de rAAV de gran tamaño, en donde el genoma de rAAV de gran tamaño sea de 5,4 kb - 6,7 kb, según se reivindica en la presente solicitud.

Por último, no habría sido obvio para una persona del oficio normalmente versada en la materia, modificar las enseñanzas de D2 a fin de producir partículas de rAAV con genomas de gran tamaño, ni habría sido obvio combinar las enseñanzas de D2 y D3 para llegar a la invención reivindicada. Esto es especialmente cierto, dado que los desafíos persistentes respecto al intento de producir partículas de rAAV con genomas de gran tamaño eran bien conocidos en la técnica, como se indicó anteriormente.

5.4. La presente solicitud resuelve el problema técnico de producir partículas de rAAV con genomas de gran tamaño mediante el uso de una línea celular productora

Por el contrario, la presente solicitud describe el uso **de una línea de células productoras (“PCL”, por sus siglas en inglés) específicamente para resolver el problema de producir partículas de rAAV con genomas de gran tamaño**. Los inventores descubrieron que tanto el tamaño como la integridad de los genomas de gran tamaño de las partículas de rAAV producidas mediante una PCL eran superiores a los de las partículas comparables producidas por triple transfección. En particular, los genomas de AAV están encapsulados en la dirección 3' a 5'. Por lo tanto, evaluar si un genoma de AAV estaba encapsulado requiere detectar la presencia o ausencia del extremo 5' del genoma. Los inventores de la presente solicitud descubrieron que los vectores rAAV producidos por PCL eran de hecho más propensos a estar sobredimensionados (ver Figura 8D) y más propensos a poseer extremos 5' encapsulados (ver Figuras 8C y 9C). Estas propiedades no pueden haber sido fácilmente deducidas o logradas por un experto en la materia basándose simplemente en D2 y/o D3.

Por tanto, los inventores resolvieron el problema técnico no resuelto anteriormente de producir partículas de rAAV con genomas de gran tamaño.

Además, el método, la composición y la línea celular reivindicados implican la producción de una partícula de rAAV con un genoma de gran tamaño a través de una célula productora, en donde el ácido nucleico que codifica los genes rep y cap de AAV y el genoma de rAAV de gran tamaño se integran de manera estable en el genoma de la línea celular productora, proporcionando así un efecto técnico inesperado y superior sobre los métodos que utilizan la triple transfección, ya que los métodos de triple transfección no generan de forma reproducible partículas de AAV con genomas de gran tamaño, como se describe anteriormente.

En consecuencia, las reivindicaciones modificadas poseen nivel inventivo sobre D2 y D3.

5.5. Respuesta a los argumentos del examinador sobre el nivel inventivo de la solicitud

Agradecemos al examinador señalar que D3 no menciona explícitamente que el ácido nucleico que codifica los genes rep y cap y/o el genoma se mantienen de manera estable en la línea celular productora, se integran de manera estable en el genoma de la línea celular productora, o se mantienen de manera estable en la línea celular productora y se integran de manera estable en el genoma de la línea celular productora. No obstante, el examinador ha argumentado que la materia reivindicada carece de nivel inventivo en vista de los documentos D3 y D2, por cuanto este último sugiere la integración estable del ácido nucleico que codifica los genes rep y cap de AAV en el genoma de la célula productora. No nos encontramos de acuerdo con esta opinión. Como se describe anteriormente, las reivindicaciones modificadas son inventivas sobre D3 y D2. En particular, se observa que D2 no enseña ni sugiere la integración estable del ácido nucleico que codifica el genoma de rAAV de gran tamaño en el genoma de una célula productora.

En cuanto a los argumentos relacionados con el efecto técnico del método reivindicado, el examinador afirma que éste solo presentó un efecto diferencial con respecto al método de triple transfección cuando tanto el ácido nucleico que codifica el genoma de gran tamaño como el ácido nucleico que codifica los genes rep y cap se integraron de manera estable en el genoma de la línea celular productora, y no hay evidencia de un efecto mejorado cuando los genes rep y cap o el genoma de gran tamaño se insertan solos. A este respecto y según lo indicado anteriormente, las reivindicaciones han sido modificadas para recitar, en parte, que *“el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora”*. Tal como ha reconocido el examinador, los ejemplos de la presente solicitud demuestran el efecto técnico de producir una partícula de rAAV con un genoma de gran tamaño a través de una célula productora, en donde el ácido nucleico que codifica los genes rep y cap de AAV y el genoma de rAAV de gran tamaño **están integrados de forma estable en el genoma**

de la línea celular productora. En consecuencia, las reivindicaciones modificadas poseen el efecto técnico de producir partículas de rAAV con genomas de gran tamaño.

Finalmente, con respecto a lo que se describe en el documento D3, el examinador afirma que dicho documento sí revela que el genoma de AAV es mayor de 5,4 kb, ya que, aunque la mayoría de los vectores producidos se acercaron a aproximadamente 5,0 kb, fue posible producir un número suficiente de vectores de 5,75 kb para pruebas *in vivo* (página 652, columna derecha, párrafo 3). No nos encontramos de acuerdo con esta opinión. Como se señaló anteriormente, los autores de D3 especulan que la razón por la que se logró suficiente expresión transgénica para sus pruebas *in vivo* fue que pudo haber ocurrido la complementación **entre** partículas de AAV con diferentes porciones del vector encapsidado. Por consiguiente, no estamos de acuerdo en que D3 demuestre que era posible producir un número suficiente de vectores de 5,75 kb para pruebas *in vivo*.

Por lo tanto, la presente solicitud cumple con los requerimientos del Artículo 18 de la Decisión 486.

6. Conclusión

Teniendo en cuenta las modificaciones introducidas en esta oportunidad y los argumentos que se presentan, es posible establecer que las reivindicaciones del nuevo pliego reivindicatorio cumplen con los requerimientos de la Decisión 486.

7. Anexos

- Nuevo pliego reivindicatorio conformado por treinta y ocho (38) reivindicaciones.
- Recibo por concepto de modificación de las reivindicaciones en trámite.
- Dong *et al.*, "Characterization of Genome Integrity for Oversized Recombinant AAV Vector" (*Mol Ther.* 2010 Jan;18(1):87-92)
- Wu *et al.* in "Effect of genome size on AAV vector packaging" (*Mol Ther.* 2010 Jan;18(1):80-6)
- Lai *et al.*, in "Evidence for the Failure of Adeno-associated Virus Serotype 5 to Package a Viral Genome ≥ 8.2 kb" (*Mol Ther.* 2010 Jan;18(1):75-9)

Señora Directora,



Margarita Castellanos Abondano
C.C. No. **39.779.654** de Bogotá
T.P.A. No. **70.819** del C.S.J.
Calle 117 No. 5A-35 - Bogotá, D.C.
Teléfono: 7560770

REIVINDICACIONES

1. Un método para producir una partícula de virus adenoasociado (AAV) que comprende un genoma de AAV recombinante (rAAV) de gran tamaño, comprendiendo el método
 - a) cultivar una línea celular productora de AAV en condiciones para generar partículas rAAV, en donde la línea celular productora de AAV comprende
 - i) ácido nucleico que codifica genes rep y cap de AAV, y
 - ii) un genoma de rAAV de gran tamaño, en donde el genoma de rAAV es 5,4kb – 6,7 kb;
 - b) proporcionar funciones auxiliares de AAV por adenovirus, HSV o baculovirus; y
 - c) recolectar las partículas de rAAV que comprenden genomas de rAAV de gran tamaño,

en donde el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora.
2. El método de la reivindicación 1, en donde el genoma de rAAV es 5,4 kb.
3. El método de cualquiera de las reivindicaciones 1-2, en donde el genoma de rAAV de gran tamaño comprende un transgen heterólogo y una o más de:
 - (i) una o dos repeticiones terminales invertidas (ITRs) de AAV;
 - (ii) un intrón; o
 - (iii) una señal de poliadenilación.
4. El método de la reivindicación 3, en donde la una o dos ITRs son ITRs de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12, AAV2R471A, AAV DJ, AAV de cabra, AAV bovino, o AAV de ratón.
5. El método de la reivindicación 3 o 4, en donde el transgén heterólogo codifica un péptido terapéutico o ácido nucleico terapéutico.
6. El método de cualquiera de las reivindicaciones 3-5, en donde el transgén heterólogo codifica el Factor VIII, la distrofina, la disferlina, o el regulador de conductancia transmembrana de la fibrosis quística (CFTR).
7. El método de cualquiera de las reivindicaciones 3-6, en donde el transgén heterólogo es un transgén humano, está unido operativamente a un promotor, o es un transgén humano y está unido operativamente a un promotor.
8. El método de cualquiera de las reivindicaciones 1-7, en donde la partícula de rAAV comprende una cápside de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7,

AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12, AAV2R471A, AAV2/2-7m8, AAV DJ, AAV2 N587A, AAV2 E548A, AAV2 N708A, AAV V708K, un AAV de cabra, AAV1/AAV2 quimérica, AAV bovina o cápside de AAV de ratón rAAV2/HBoV1.

9. El método de cualquiera de las reivindicaciones 3-8, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan del mismo serotipo de AAV.
10. El método de cualquiera de las reivindicaciones 3-9, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de AAV2.
11. El método de cualquiera de las reivindicaciones 3-8, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de diferentes serotipos de AAV.
12. El método de cualquiera de las reivindicaciones 3-8 o la reivindicación 11, en donde la una o dos ITRs son ITRs de AAV2 y la cápside es una cápside AAVrh8R o cápside AAV8.
13. El método de cualquiera de las reivindicaciones 1-12, en donde la línea celular productora se deriva de células de primate, células HeLa, células 293, células A549 o células Perc.6.
14. El método de cualquiera de las reivindicaciones 1-13, que comprende además purificación de las partículas de rAAV.
15. Una composición que comprende partículas de rAAV en donde al menos 45% de la partícula de rAAV encapsida un genoma de rAAV de gran tamaño que es de 5,4 kb - 6,7kb,

en donde las partículas de AAV se producen mediante un método para producir una partícula de rAAV que comprende un genoma de rAAV de gran tamaño, comprendiendo el método

- a) cultivar una línea celular productora de AAV en condiciones para generar partículas de rAAV, en donde la línea celular productora de AAV comprende
 - i) ácido nucleico que codifica los genes rep y cap de AAV, y
 - ii) un genoma de rAAV de gran tamaño, en donde el genoma de rAAV de gran tamaño es de 5,4 kb - 6,7 kb;
- b) proporcionar funciones auxiliares de AAV mediante adenovirus, HSV o baculovirus; y
- c) recolectar las partículas de rAAV que comprenden genomas de rAAV de gran tamaño,

en donde el ácido nucleico que codifica los genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora.

16. La composición de la reivindicación 15, en donde el genoma de rAAV de gran tamaño comprende un transgén heterólogo y uno o más de:
 - (i) una o dos repeticiones terminales invertidas (ITRs) de AAV;
 - (ii) un intrón;
 - (iii) una señal de poliadenilación.
17. La composición de la reivindicación 16, en donde las ITRs son ITRs de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12, AAV2R471A, AAV DJ, AAV de cabra, AAV bovino o AAV de ratón.
18. La composición de la reivindicación 16 o 17, en donde el transgén heterólogo codifica un péptido terapéutico o ácido nucleico terapéutico.
19. La composición de cualquiera de las reivindicaciones 16-18, en donde el transgén heterólogo codifica el Factor VIII, la distrofina, la disferlina o el regulador de conductancia transmembrana de la fibrosis quística (CFTR).
20. La composición de cualquiera de las reivindicaciones 16-18, en donde el transgén heterólogo es un transgén humano, está unido operativamente a un promotor, o es un transgén humano y está unido operativamente a un promotor.
21. La composición de cualquiera de las reivindicaciones 15-20, en donde la partícula de rAAV comprende una cápside de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12, AAV2R471A, AAV2/2-7m8, AAV DJ, AAV2 N587A, AAV2 E548A, AAV2 N708A, AAV V708K, un AAV de cabra, AAV1/AAV2 quimérico, AAV bovino o cápside de AAV de ratón rAAV2/HBoV1.
22. La composición de cualquiera de las reivindicaciones 15-21, en donde la ITR y la cápside de la partícula de rAAV se derivan del mismo serotipo de AAV.
23. La composición de cualquiera de las reivindicaciones 15-22, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de AAV2.
24. La composición de cualquiera de las reivindicaciones 15-21, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de diferentes serotipos de AAV.
25. La composición de cualquiera de las reivindicaciones 15-21 o la reivindicación 24, en donde la una o dos ITRs son ITRs de AAV2 y la cápside es una cápside AAVrh8R o cápside AAV8.

26. Una línea celular para producir una partícula de virus adenoasociado (AAV) que comprende un genoma de AAV recombinante de gran tamaño, comprendiendo la línea celular

- a) ácido nucleico que codifica genes rep y cap de AAV, y
- b) un genoma de rAAV de gran tamaño, en donde el genoma de rAAV es 5,4 kb – 6,7kb,

en donde el ácido nucleico que codifica genes rep y cap de AAV y el ácido nucleico que codifica el genoma de rAAV de gran tamaño se integran de forma estable en el genoma de la línea celular productora,

y en donde la línea celular productora se deriva de células de primates, células HeLa, células 293, células A549 o células Perc.6.

27. La línea celular de la reivindicación 26, en donde el genoma de rAAV es 5,4 kb.

28. La línea celular de cualquiera de las reivindicaciones 26-27, en donde la línea celular está adaptada para el crecimiento en suspensión.

29. La línea celular de cualquiera de las reivindicaciones 26-28, en donde el genoma de rAAV de gran tamaño comprende un transgen heterólogo y uno o más de:

- (i) una o dos repeticiones terminales invertidas (ITRs) de AAV;
- (ii) un intrón; o
- (iii) una señal de poliadenilación.

30. La línea celular de la reivindicación 29, en donde las ITRs son ITRs de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12, AAV2R471A, AAV DJ, AAV de cabra, AAV bovino, o AAV de ratón.

31. La línea celular de la reivindicación 29 o 30, en donde el transgén heterólogo codifica un péptido terapéutico o ácido nucleico terapéutico.

32. La línea celular de una cualquiera de las reivindicaciones 29-31, en donde el transgén heterólogo codifica el Factor VIII, la distrofina, la disferlina o el regulador de conductancia transmembrana de la fibrosis quística (CFTR).

33. La línea celular de una cualquiera de las reivindicaciones 29-32, en donde el transgén heterólogo es un transgén humano, está unido operativamente a un promotor, o es un transgén humano y está unido operativamente a un promotor.

34. La línea celular de cualquiera de las reivindicaciones 26-33, en donde la partícula de rAAV comprende una cápside de serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh8R, AAV9, AAV10, AAVrh10, AAV11, AAV12,

AAV2R471A, AAV2/2-7m8, AAV DJ, AAV2 N587A, AAV2 E548A, AAV2 N708A, AAV V708K, un AAV de cabra, AAV1/AAV2 quimérico, AAV bovino o cápside de AAV de ratón rAAV2/HBoV1.

35. La línea celular de cualquiera de las reivindicaciones 29-34, en la donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan del mismo serotipo de AAV.
36. La línea celular de cualquiera de las reivindicaciones 29-35, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de AAV2.
37. La línea celular de cualquiera de las reivindicaciones 29-34, en donde la una o dos ITRs y la cápside de la partícula de rAAV se derivan de diferentes serotipos de AAV.
38. La línea celular de cualquiera de las reivindicaciones 29-34 o la reivindicación 37, en donde la una o dos ITRs son ITRs de AAV2 y la cápside es una cápside AAVrh8R o cápside AAV8.



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT: 800.176.089-2

COMPROBANTE DE RECAUDO SISTEMA SIPI

Radicado: NC2017/0011344
Recibo: 21-13493
Fecha: 12/02/2021

MARGARITA CASTELLANOS ABONDANO
CC: 39.779.654
Pago PSE No: 500607
CUS: 891455956

Patente PCT
PRODUCCIÓN DE VECTORES ADENO-ASOCIADOS DE GRAN TAMAÑO

Detalle del pago

Id.	Rentístico	Concepto y Servicio	Cant.	Val. Unit.	Subtotal
1	50005-01-03	Modificaciones / Correcciones al registro Patente PCT (DESCUENTO INTERNET)	1	\$ 171,500	\$ 171,500
TOTAL					\$ 171,500

See page 6

Characterization of Genome Integrity for Oversized Recombinant AAV Vector

Biao Dong¹, Hiroyuki Nakai² and Weidong Xiao^{1,3}

¹The Sol Sherry Thrombosis Research Center, Temple University School of Medicine, Philadelphia, Pennsylvania, USA; ²Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Philadelphia, Pennsylvania, USA; ³Department of Microbiology and Immunology, Temple University School of Medicine, Philadelphia, Pennsylvania, USA

Application of recombinant adeno-associated virus (rAAV) in gene therapy has been limited by its packaging capacity. Recent studies suggested that rAAV could achieve persistent transgene expression beyond 4.7-kb packaging limit. To clarify the mechanism leading to transgene expression from oversized rAAV vector, we constructed a series of rAAV vectors with genomes ranging from 2.9 to 7.2 kb. A plasmid replication origin and an ampicillin-resistant marker were included in the vector to facilitate the recovery of circularized, post-transduction AAV genome. Southern dot-blot analysis and silver staining confirmed that rAAVs could be produced at varying vector size. However, the vector yields decreased approximately tenfold for oversized vectors as compared to regular ones. Alkaline Southern blot hybridization suggested that the packaged genomes for oversized vectors were truncated. In the cells transduced by the above vectors, circularized rAAV monomers could be rescued at 24 hours after infection. Few recovered AAV genomes were >5 kb regardless of the initial vector size. In mice receiving the above vectors, larger circularized rAAV genomes could be recovered for oversized vectors at day 21 after vector administration. Our studies suggested that the partially packaged rAAV sequences may complement each other to restore full expression cassette.

Received 25 August 2009; accepted 13 October 2009; published online 10 November 2009. doi:10.1038/mt.2009.258

INTRODUCTION

Adeno-associated virus (AAV) is a nonpathogenic and replication-defective human parvovirus. In addition to two open reading frames encoding the *rep* gene and *cap* gene, the 4.7-kb single-stranded DNA genome is flanked by two inverted terminal repeats (ITRs). There are 145 nucleotides in each ITR and the palindromic sequences within the ITR can form a T-shape structure, which facilitates AAV DNA replication in a self-priming manner.¹ These two ITRs are the only *cis* elements in recombinant AAV (rAAV) genome essential for efficient packaging.

With a vector plasmid carrying the transgene cassette between two ITRs, rAAV can be produced by supplying necessary AAV

trans elements, and helper virus functions, such as adenovirus or herpes virus.² Extensive studies have documented rAAV as one of the most promising vector for human gene therapy.^{3–5} Many new AAV serotypes have been isolated or generated through directed evolution, which broadened the cell types and tissues that rAAV can efficiently transduce and direct long-term transgene expression.^{6–10} AAV vectors are currently being evaluated in human clinical trials for a variety of diseases and the preliminary results are very encouraging.^{11–17}

Wild-type AAV virus is able to integrate site-specifically into a specific locus of human chromosome 19 at a high frequency in a latent infection.^{18,19} However, cellular junctions between rAAV vectors and host chromosomes are rare to detect using various approaches.^{20,21} On the other hand, episomal AAV genomes, linear or circular, can be isolated readily in both tissue culture cells and transduced target tissues. As demonstrated by Duan *et al.* the circularized rAAV genomes can be formed right after rAAV vectors enter the host cells.^{22–24} The major transgene expression by rAAV transduction most likely resulted from episomal AAV genomes.²⁵

The packaging capacity for rAAV is a critical issue that may determine how rAAV vectors can be applied for human gene therapy, as many target genes and regulatory elements are large in size. It was reported that an rAAV genome <4.9 kb could be completely packaged into a virus.²⁶ By step-increasing the wild-type AAV genome, it was suggested that 5.3 kb of nucleotides should be the maximum for efficiently packaging AAV.²⁷ Using the same constructs as Dong *et al.*,²⁶ it was found that the full length of 6-kb rAAV genome could be packaged into virus, with or without VP2 protein.²⁸ Another report showed that a 3.3 kb of double-stranded DNA could be packaged into rAAV, which suggested the rAAV packaging capacity could be enlarged to 6.6 kb (ref. 29). Recently, it was reported that large vectors up to 8.9 kb could be packaged into rAAVs by selection of suitable AAV serotypes.³⁰ In our effort to package large vectors, it was shown that a 5.8-kb vector for the expression of factor VIII was only partially packaged, although this vector could functionally correct the phenotype of hemophilia A.³¹

Shuttle vectors have been used as a tool to study rAAV transduction mechanisms for conventional vectors with rAAV genomes <4.7 kb (refs. 22,32,33). In an rAAV shuttle vector, the ampicillin-resistant gene and plasmid replication region are presented in

Correspondence: Weidong Xiao, Department of Microbiology and Immunology, Sol Sherry Thrombosis Research Center and Section of Hematology, Temple University, 3400 N. Broad Street, OMS 531, Philadelphia, PA 19140, USA. E-mail: wxiao@temple.edu

the rAAV genomes. Upon transduction, duplex rAAV genome may become a circular molecule via the two ITRs. The circularized rAAV genome could be rescued by transforming competent *Escherichia coli*. Here, we took advantage of AAV genome circularization and characterized the genome integrity of rAAV vectors produced from oversized rAAV genome vectors. The rescued rAAV genomes were analyzed by restriction enzyme digestion mapping and DNA sequencing. Our detailed analysis of encapsidated AAV genomes demonstrated that the packaging of oversized vectors was inefficient and only partial AAV genomes can be packaged. Complementation of incomplete AAV genomes may have led to the regeneration of full AAV gene expression cassette.

RESULTS

Analysis of the effects of vector genome size on rAAV packaging

In order to characterize the vector genome integrity, we choose to employ a rescue strategy as previously published.^{22,32,33} The key elements in this approach are the plasmid replication origin (*ori*) and ampicillin resistant (*amp*) flanked by the two regular AAV ITRs. The detailed illustration for these plasmids is shown in **Figure 1**. The sequences flanked by AAV ITRs based on these vector plasmids ranged from 2.9 to 7.2 kb. The stuffer sequences for these vector plasmids were nonfunctional human factor VIII codons.

The packaging efficacy using these vector plasmids were then analyzed following our standard, triple plasmid transfection, AAV production protocol.³⁴ Under cesium chloride–gradient ultracentrifugation, we noticed that the usual distinctive AAV vector bands could no longer be observed from preparations that used the oversized vector plasmids DB6.4 and DB7.2 for rAAV production. Individual gradient fractions from these vector preparations were then collected and analyzed by both Southern dot-blot and silver staining. The packaged vector DNA as revealed by Southern blot (**Figure 2a**) suggested that vector size >4.7 kb (DB5.4, DB6.4, and DB7.2) caused a considerable decrease in vector yields. Despite that there were some variations in the yields of packaged of 7.2, 6.4, and 5.4 kb vectors, the vector yields from these vector plasmids were comparable to each other based on our quantitative analysis. Overall vector yields were approximately tenfold less than their smaller counter parts (DB2.9 and DB 3.9). On the other hand, AAV capsids revealed by silver staining showed that the ratio

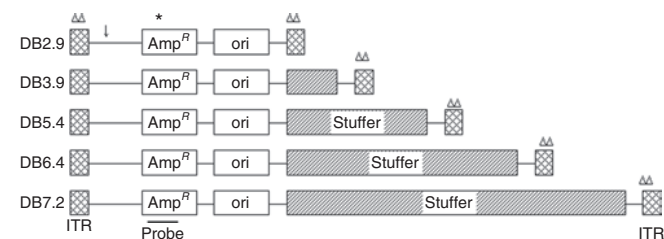


Figure 1 Illustration of AAV vector genomes. The stuffer fragments were partial sequences of factor VIII gene. The vectors were named according to their rAAV genome size. The DNA fragment used for the probe for Southern blot hybridization was identified, which corresponds to the sequences in the ampicillin-resistant gene. AAV, adeno-associated virus; *amp^R*, an ampicillin-resistant gene, β -lactamase; ITR, inverted terminal repeat sequence; *ori*, a plasmid replication region of pUC18; $\downarrow$, *AleI* site; *, *Bam*HI site; Δ , *SmaI* site.

of vector capsid proteins VP1, VP2, and VP3 were similar among all vector preparations, which is independent of vector genome size. Therefore, capsid formation, either empty particles or full particles, was not affected by the size of AAV vector genomes.

The size of vector DNA packaged into AAV capsids was further analyzed by alkaline Southern blot analysis. The peak fraction of AAV vectors from the above preparations were digested by proteinase K, separated in agarose gel, and probed with DNA sequence specific to the *amp^R* region. This result is presented in **Figure 3**. The major vector size for DB2.9 and DB 3.9 were close to 2.9 and 3.9 kb, which was nearly identical to their original size. However, the vector DNA from preparation from vector plasmids DB5.4, DB6.4, and DB7.2 were not >4.7 kb. It suggested that only partial DNA sequences can be encapsidated if vector DNA is >4.7 kb.

Rescue of rAAV DNA *in vitro*

As AAV genomes in circular configuration can be formed in the host cells after rAAV transduction, the inclusion of *amp* and *ori* in the AAV vector DNA could therefore facilitate the rescue of circularized AAV genomes. We transduced 293 cells with equal genome titers of AAV vectors (multiplicity of infection 10,000) made from vector plasmids that varied in their genome size. Small-molecular weight DNA was then extracted from these cells at 24 hours postinfection. After electro-transformation into competent cells, the number of *amp*-resistant colonies are counted and graphed in **Figure 4**. On average, AAV vector with small/regular genomes

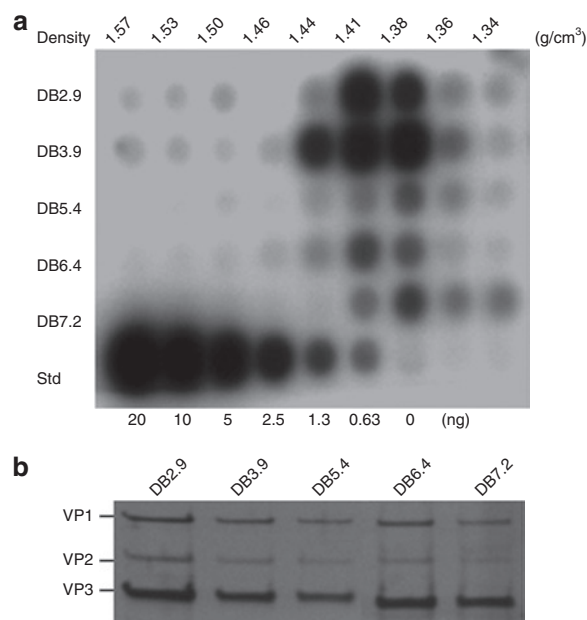


Figure 2 Oversized AAV vectors are inefficient in packaging. rAAV vectors were packaged using AAV5 capsids using triple plasmid transfection method as described in the Materials and Methods section. The resulting rAAV vectors were purified by two rounds of cesium chloride–gradient ultracentrifugation. **(a)** rAAV vectors in each gradient fraction were quantified by use of Southern dot-blotting. Each dot represented 2 μ l of purified vectors from each gradient fraction. The membrane was hybridized with a probe to a region that is identical in all vectors. **(b)** Detection of capsid proteins by silver staining. A volume of 1×10^{10} viral particles for each vector were used for the silver staining. rAAV, recombinant adeno-associated virus; Std, the standard for quantification of the purified viral genomes; VP1, VP2, VP3, are AAV capsid proteins.

such as DB2.9 and DB3.9 gave rise to 4,000–6,000 colonies. In contrast, vectors with oversized genomes, DB5.4, DB6.4, and DB7.2, resulted in substantial fewer colonies, which were in the range of 500–1,000.

Analysis of AAV genome regeneration from large AAV vectors *in vivo* and *in vitro*

The circularized AAV genomes recovered from transduced 293 cells by DB7.2 were then mapped by both restriction digestions and DNA sequencing. The missing AAV fragments are summarized in Figure 5. Of all clones recovered, 80% of clones had deletions >5kb, 7% had deletions between 3 and 4 kb, and only 13% had deletions between 2 and 3 kb. There were no clones with deletions <1 kb.

Similar experiments were carried out *in vivo*. Vectors made from DB7.2 were injected into the muscle of C57BL6 mice. The small-molecular weight DNA was extracted from the muscle tissue at the injection site on day 21. Upon transformation in competent *E. coli*, the clones containing AAV genomes are shown in

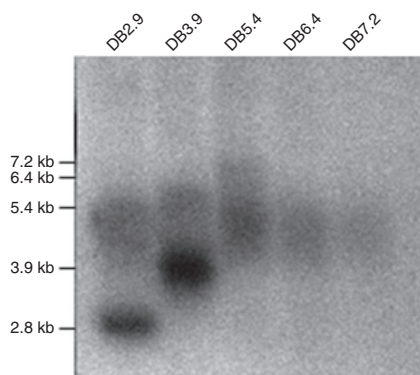


Figure 3 Analysis of the rAAV genome integrity. The size of the single-stranded DNAs packaged in rAAV capsids were analyzed using alkaline agarose gel electrophoresis. After extensive DNaseI treatment, the packaged DNAs from different vectors were resolved on 1% alkaline agarose gel. The membrane was hybridized with a probe specific to the ampicillin-resistant gene. rAAV, recombinant adeno-associated virus.

the bottom panel of Figure 5. In contrast to the *in vitro* results, the clones with large deletions (>4 kb) decreased to 17%. Clones with deletion in the range of 3, 2, and 1 kb are 26, 35, and 15%, respectively. Seven percent of the clones had deletions <1 kb. It suggested that AAV vector genomes undergo significant changes *in vivo*.

DISCUSSION

Traditionally, it has been documented that a virus can generally package no more than 105% nucleotides of normal virus genome. Analysis of the wild-type AAV virus packaging revealed that increasing AAV genome size drastically reduces virus encapsidation.^{26,35} An oversized wild-type AAV virus has also been explored as a helper for rAAV production due to its deficiency in packaging itself.³⁶ Despite the outstanding performance of rAAV vectors in transduction, the packaging capacity of AAV vectors has been a limiting factor, preventing broader application of rAAV vectors for human gene therapy. Recently, several studies explored oversized AAV vectors in a variety of settings.^{26–31} All these studies suggested that oversized AAV vectors may be useful for human

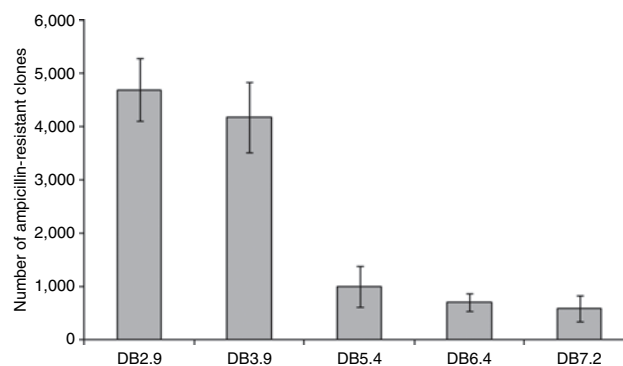


Figure 4 Rescued rAAV genomes from HeLa cells. HeLa cells were transduced with similar amounts of rAAVs at a multiplicity of infection of 10,000. After 24-hour infection, low-molecular-weight DNAs were extracted and transformed into competent cells. The ampicillin-resistant colonies growing on Luria–Bertani plates were counted for each vector. The experiment was repeated twice. rAAV, recombinant adeno-associated virus.

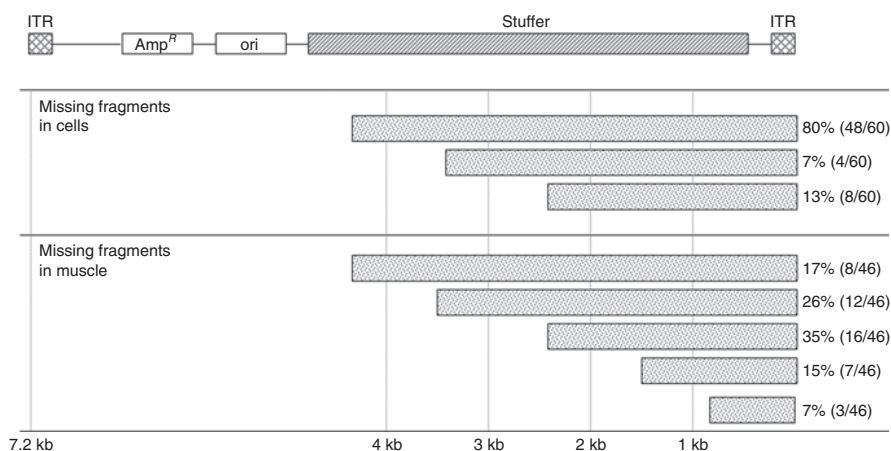


Figure 5 Summary of the deleted regions in the rescued rAAV genomes for DB7.2. All plasmids rescued from HeLa cells and mouse muscle tissues were analyzed by restriction enzyme digestion and DNA sequencing. The percentage was calculated as dividing the total number of rescued plasmids by the number of classified plasmids. The semi-filled box illustrates missing region from recovered rAAV genome. *amp^R*, an ampicillin-resistant gene, β -lactamase; ITR, inverted terminal repeat sequence; *ori*, a plasmid replication region of pUC18; rAAV, recombinant adeno-associated virus.

gene therapy. On the other hand, these results also raised a critical question of how AAV vectors could package more than the normal size of AAV genomes.

In our studies, we generated a series of AAV vectors with genomes ranging from 2.9 to 7.2 kb. In contrast to previous studies which focused on transgene expression, we included both an *amp* and *ori* sequence in the rAAV sequence to specifically facilitate the recovery of circularized AAV molecules. It is known that AAV vectors below the wild-type size generally performed similarly to wild-type-genome-sized vectors. Vectors DNA <4.7 kb (DB2.9 and DB 3.9) packaged and behaved similar to that standard-sized 4.7-kb vector. The packaging of oversized AAV genomes appeared to be inefficient with at least a tenfold decrease in packaging efficiency. This result is consistent with many previous studies²⁷ showing that the optimal packaging requires the size of AAV vector, including AAV ITRs, not to exceed 4.7–5.3 kb. In addition, we also observed that vector >5.3 kb gave rise to similar low yield, suggesting a common packaging problem for the large vectors. The debatable issue is how much DNA AAV can encapsidate into AAV genomes when AAV vector genome size increases. In several recently published studies, significant transgene expression could be observed *in vivo* using an AAV vector with large genomes ranging from 5 to 9 kb (refs. 26–31). However, the reported size of packaged DNA varied. Our previous results²⁷ suggested that only 3.9 kb could be packaged by alkaline gel analysis of AAV genomes, which confirms again that DNA inside the AAV capsid is <5 kb regardless of the size of the AAV genome (Figure 3). Analysis of different density fractions (higher up to 1.46–1.48 g/ml and lower to 1.35–1.38 g/ml) did not show any size difference as well (data not shown). Thus, we conclude that through physical limitation an AAV capsid cannot accommodate >5,000 nucleotides.

The conclusion that only limited nucleotides can be encapsidated was also supported by our data from rescuing AAV genomes *in vitro* and *in vivo*. It is clear that the larger the vector size, the fewer circularized AAV genomes that could be recovered (Figure 4). This result suggested that there were serious defects in encapsidated AAV genomes when rAAV genomes were over the normal range. The deletion of an ITR from the larger genomes during the truncation/deletion process of the larger genomes may have resulted in lower circularization efficiencies. In addition, deletion in elements controlling formation, *amp* gene, and plasmid *ori* would have reduced the clones, which could be recovered. Further analysis of recovered, circularized AAV genomes provided direct evidence supporting this hypothesis. The circular molecules recovered from AAV transduction of tissue cultures, at only 24 hours post-transduction, revealed that >80% of them contained large deletions. Of those with a 7.2-kb genome >80% had deletions >4 kb. There is nothing >6 kb, let alone full-size genome. This result is consistent with our Southern blot analysis of AAV genomes, that >5-kb genomes could not be packaged (Figure 3). We sequenced ~200 clones that were recovered (data not shown), which suggested that AAV ITR and non-AAV ITR sequences were joined together. Thus, we speculate that one ITR and non-ITR sequence junction could be formed by direct intramolecular ends joining or nonhomologous recombination. On the other hand, most of these clones were 2–3 kb, much <4.7 kb. The truncated genomes were probably being further processed in 293 cells after

uncoating occurred to achieve circularization. Missing one ITR may also increase rAAV genome degradation/processing. In contrast, the circularized AAV molecules recovered from the muscle tissue that has been transduced by rAAV vectors exhibited a very different pattern (Figure 5). Not only did the number of clones with large deletions decrease dramatically, but also clones close to full AAV genomes could be recovered. These phenomena were observed with all other vectors >5.4 kb. Therefore, the above results suggested that the AAV vectors underwent significant changes *in vivo*. Longer time in the target tissue probably facilitated AAV genome conversion, which allowed larger AAV genomes containing the full-length transgene to be regenerated.

We wish to resolve the issue on AAV encapsidation of large genomes so that oversized AAV vectors can have broad applications in human gene therapy. Our results and other studies supported that AAV genomes packaged were <5 kb. However, through other molecular mechanism, most likely, AAV genome conversion from overlapping sequence, AAV vectors produced from oversized AAV genomes were able to give rise to meaningful transgene expression.³⁷ The successful use of this approach will probably rest on the further improvement of AAV vector-packaging technology.

MATERIALS AND METHODS

Cell lines. HEK 293 cells and HeLa cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 µg of penicillin/ml, 100 U of streptomycin/ml as recommended by the manufacturer (Invitrogen, Carlsbad, CA). All cells were maintained in a humidified 37°C incubator with 5% CO₂.

Plasmid construction. The plasmid pAAV-IsceI.AO3 (accession number EU022316) has been described previously.³³ Plasmids containing various rAAV genomes were made in the following procedures. First, pAAV-IsceI.AO3 was digested by *NheI* and *EcoRV* and subsequently filled-in by T₄ DNA polymerase. The 2,855-bp fragment was then self-ligated, or dephosphorylated by calf intestinal alkaline phosphatase (NEB, Ipswich, MA) and used as the backbone for the construction of the other vectors. The self-ligated plasmid has around 2.9 kb of rAAV genome, so it was named as DB2.9. The above plasmid designation was also used for the other constructs DB3.9, DB5.4, DB6.4, and DB7.2, respectively. These plasmids were constructed by blunt-end ligation of the backbone and the filled-in DNA fragments digested from the plasmid pAAV-TTR-hF8. Specially, the filled-in DNA fragments were 1,046-bp fragment by *ClaI* and *AccI*, 2,564-bp fragment by *KpnI* and *XhoI*, 3,522-bp fragment by *ClaI* and *ApaI*, and 4,389-bp fragment by *ClaI* and *XhoI*. The constructs were confirmed by DNA sequencing.

rAAV production. A triple plasmid co-transfection method was used to produce rAAVs as described previously.³⁸ One vector plasmid, one AAV5 helper plasmid, and one mini adenovirus function helper plasmid pFΔ6, were co-transfected into HEK293 cells cultured in roller bottles at a ratio of 1:2:1. The transfected cells were harvested 3 days later. rAAVs were then purified by two rounds of cesium chloride–gradient ultracentrifuge. After extensive buffer exchange against phosphate-buffered saline with 5% D-sorbitol, the peak fractions of purified virus were pooled and stored at –80°C before administration.

rAAV titer determination. Genome titers were determined by dot-blot hybridization according to the standard protocol. Briefly, 2 µl of purified virus from each gradient fraction was used to release the genome in a reaction mix containing 50 mmol/l Tris–HCl, pH 8.0, 1 mmol/l EDTA, 0.5%

sodium dodecyl sulfate (SDS), and 0.5 mg/ml proteinase K. The mixtures were extracted once using phenol:chloroform:isoamyl alcohol (25:24:1). After denaturing, the mixtures were dotted on a Hybond-XL nylon membrane (GE Healthcare, Little Chalfont, UK). After the membrane was rinsed with 2× sodium chloride and sodium citrate, it was dried at 65°C. For hybridization, the probe was a DNA fragment located in the ampicillin-resistant gene that was shared by all vectors. This fragment was PCR amplified using a pair of primers 5'-GCCTTCCTGTTTTGCTCAC-3' and 5'-TTGCCGGAAGCTAGAGTAA-3'. The purified PCR products were labeled by [α -³²P]dCTP using Prime-It II Random Primer Labeling Kit (Stratagene, La Jolla, CA).

Silver staining. Purified virus (~1 × 10¹⁰ particles) was mixed with protein-loading dye and NuPAGE sample-reducing reagent (Invitrogen), the mixture was heated at 95°C for 5 minutes. The denatured samples were then separated on 10% Tris-glycine gel (Invitrogen). After electrophoresis, the SDS-poly acrylamide gel was stained by SilverXpress silver staining kit (Invitrogen) according to the manufacturer's instruction.

rAAV DNA analysis. Alkaline agarose gel electrophoresis was used to analyze the size of the single-stranded DNA packaged in rAAV capsids as described previously.³¹ Briefly, viral genomes were treated with DNaseI and resolved on 1% alkaline agarose gel. After gel blotting, the membrane was hybridized with a ³²P-labeled probe that was used for the titer determination. The membrane was washed once under low-stringent condition (2× salt-sodium citrate + 0.1% SDS, at room temperature), followed by washing for three times under high-stringent condition (0.2× salt-sodium citrate + 0.1% SDS, at 65°C). The membrane was then analyzed by X-ray autoradiography.

Animal procedures. The 4- to 5-week-old C57BL/6 mice were housed in a pathogen-free environment with a normal diet. All surgical procedures involving mice were in accordance with institutional guidelines under approved protocols at the Children's Hospital of Philadelphia, PA. For each vector, three hindlimb muscles were injected independently and each injection contained about 5 × 10¹⁰ viral particles.

rAAV genome rescue. HeLa cells were grown to 95% confluence in a 6-well plate and transduced with rAAVs at an multiplicity of infection of 10,000. After 24-hour infection, cells were rinsed once with 1× phosphate-buffered saline, scraped into centrifuge tubes, and pelleted by centrifugation. To rescue rAAV genomes from HeLa cells, low-molecular-weight DNA was extracted using High Pure plasmid isolation kit (Roche Diagnostics, Mannheim, Germany) and eluted using 50 µl of elution solution (10 mmol/l Tris, pH 8.0).

To rescue rAAV genome from muscle tissue, transduced muscles were harvested on 21 day. Muscle tissue (100 mg) was chopped into small pieces and subsequently homogenized in 1 ml of buffer containing 10 mmol/l Tris (pH 8.0), 10 mmol/l EDTA, 1% SDS, and 100 µg/ml RNase (DNase free; Sigma, St Louis, MO). After incubation at 37°C overnight, NaCl was added to a final concentration of 1.1 mol/l. The samples were further incubated at 4°C overnight. The processed samples were then extracted twice with phenol-chloroform and once with chloroform, followed by ethanol precipitation. After washing with 70% ethanol, DNA pellet was resuspended in 20 µl of Tris-EDTA buffer.

For each vector, 3 µl of low-molecular-weight DNA prepared above was used to transform 40 µl of SURE electroporation-competent *E. coli* (Stratagene). Bacterial clones growing on Luria-Bertani plates containing 100 µg/ml ampicillin were randomly selected and cultured for plasmid extraction. At least 20 clones for each vector were analyzed by restriction enzyme digestion and DNA sequencing. The total colony number for each vector was also calculated according to the portion of DNA used for transformation.

Restriction enzyme digestion. To characterize the rescued rAAV genome, DNA fingerprinting was generated by *AleI* and *BamHI* double digestion. In a 20-µl reaction system, 300 ng of plasmid was digested with 2.5 units of *AleI* and 10 units of *BamHI* at 37°C for 3 hours. To estimate the ITR integrity, 10 units of *SmaI* was used in a 20 µl reaction system that was incubated at 25°C for 3 hours.

DNA sequencing. To assess rAAV genome accurately, a serial of primers were designed for DNA sequencing as following: DB7.2-2398F, 5'-ACC TCT GAC TTG AGC GTC GA-3'; DB7.2-2560F, 5'-AGA GCT CTC CAC CTG CTT CT-3'; DB7.2-3055F, 5'-TGA CCC ACT GTG CCT TAC CT-3'; DB7.2-3548F, 5'-TCC CAC CA ACA TGA TGG CAT-3'; DB7.2-4391F, 5'-CTC CCC AAT CCA GCT GGA GT-3'; DB7.2-4529F, 5'-AGC ATT GGA GCA CAG ACT GA-3'; DB7.2-5085F, 5'-ACA GGG CTC AGA GTG GCA GT-3'; DB7.2-5485F, 5'-ACC CCT TCT GGT CTG CCA CA-3'; DB7.2-6138F, 5'-TCA AGG TGG ATC TGT TGG CA-3'; DB7.2-6613F, 5'-GCT GCA AGT GGA CTT CCA GA-3'; DB7.2-6929F, 5'-AGC TTG GCA ATG CAT G-3'; DB7.2-321R, 5'-GTA GTT GTA CTC CAG CTT GT-3'; DB7.2-444R, 5'-GTG TTC TGC TGG TAG TGG T-3'; DB7.2-940R, 5'-CCA GTT CGA TGT AAC CCA CT-3'.

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health (HL080789 and HL084381 to W.X.).

REFERENCES

- Muzyczka, N and Berns, KI (2002). Parvoviridae: the viruses and their replication. In: Knipe, DM and Howley, PM (eds). *Fields Virology*, 4th edn. Lippincott Williams & Wilkins: Philadelphia, pp 2327–2360.
- Ferrari, FK, Xiao, X, McCarty, D and Samulski, RJ (1997). New developments in the generation of Ad-free, high-titer rAAV gene therapy vectors. *Nat Med* **3**: 1295–1297.
- Samulski, RJ (2003). AAV vectors, the future workhorse of human gene therapy. *Ernst Schering Res Found Workshop*: 25–40.
- Herzog, RW, Yang, EY, Couto, LB, Hagstrom, JN, Elwell, D, Fields, PA *et al.* (1999). Long-term correction of canine hemophilia B by gene transfer of blood coagulation factor IX mediated by adeno-associated viral vector. *Nat Med* **5**: 56–63.
- Fisher, KJ, Jooss, K, Alston, J, Yang, Y, Haeccker, SE, High, K *et al.* (1997). Recombinant adeno-associated virus for muscle directed gene therapy. *Nat Med* **3**: 306–312.
- Li W, Zhang L, Johnson JS, Zhijian W, Grieger JC, Ping-jie X *et al.* (2009). Generation of novel AAV variants by directed evolution for improved CFTR delivery to human ciliated airway epithelium. *Mol Ther* (pub ahead of print).
- Asokan, A and Samulski, RJ (2006). AAV does the shuffle. *Nat Biotechnol* **24**: 158–160.
- Grimm, D and Kay, MA (2003). From virus evolution to vector revolution: use of naturally occurring serotypes of adeno-associated virus (AAV) as novel vectors for human gene therapy. *Curr Gene Ther* **3**: 281–304.
- Gao, G, Vandenbergh, LH, Alvira, MR, Lu, Y, Calcedo, R, Zhou, X *et al.* (2004). Clades of adeno-associated viruses are widely disseminated in human tissues. *J Virol* **78**: 6381–6388.
- Gao, GP, Alvira, MR, Wang, L, Calcedo, R, Johnston, J and Wilson, JM (2002). Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci USA* **99**: 11854–11859.
- Maguire, AM, Simonelli, F, Pierce, EA, Pugh, EN Jr, Mingozzi, F, Bencicelli, J *et al.* (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* **358**: 2240–2248.
- Hasbrouck, NC and High, KA (2008). AAV-mediated gene transfer for the treatment of hemophilia B: problems and prospects. *Gene Ther* **15**: 870–875.
- High, KA (2007). Update on progress and hurdles in novel genetic therapies for hemophilia. *Hematology Am Soc Hematol Educ Program*: 466–472.
- Mingozzi, F, Meulenberg, JJ, Hui, DJ, Basner-Tschakarjan, E, Hasbrouck, NC, Edmonson, SA *et al.* (2009). AAV-1-mediated gene transfer to skeletal muscle in humans results in dose-dependent activation of capsid-specific T cells. *Blood* **114**: 2077–2086.
- Rodino-Klapac, LR, Lee, JS, Mulligan, RC, Clark, KR and Mendell, JR (2008). Lack of toxicity of α -sarcoglycan overexpression supports clinical gene transfer trial in LGMD2D. *Neurology* **71**: 240–247.
- Hajjar, RJ, Zsebo, K, Deckelbaum, L, Thompson, C, Rudy, J, Yaroshinsky, A *et al.* (2008). Design of a phase 1/2 trial of intracoronary administration of AAV1/SERCA2a in patients with heart failure. *J Card Fail* **14**: 355–367.
- Eberling, JL, Jagust, WJ, Christine, CW, Starr, P, Larson, P, Bankiewicz, KS *et al.* (2008). Results from a phase I safety trial of hAADC gene therapy for Parkinson disease. *Neurology* **70**: 1980–1983.
- Kotin, RM, Siniscalco, M, Samulski, RJ, Zhu, XD, Hunter, L, Laughlin, CA *et al.* (1990). Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci USA* **87**: 2211–2215.
- Yang, CC, Xiao, X, Zhu, X, Ansardi, DC, Epstein, ND, Frey, MR *et al.* (1997). Cellular recombination pathways and viral terminal repeat hairpin structures are sufficient for adeno-associated virus integration *in vivo* and *in vitro*. *J Virol* **71**: 9231–9247.
- Nakai, H, Iwaki, Y, Kay, MA and Couto, LB (1999). Isolation of recombinant adeno-associated virus vector-cellular DNA junctions from mouse liver. *J Virol* **73**: 5438–5447.

21. Nakai, H, Wu, X, Fuess, S, Storm, TA, Munroe, D, Montini, E *et al.* (2005). Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. *J Virol* **79**: 3606–3614.
22. Duan, D, Sharma, P, Yang, J, Yue, Y, Dudus, L, Zhang, Y *et al.* (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.
23. Duan, D, Sharma, P, Dudus, L, Zhang, Y, Sanlioglu, S, Yan, Z *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.
24. Duan, D, Yan, Z, Yue, Y and Engelhardt, JF (1999). Structural analysis of adeno-associated virus transduction circular intermediates. *Virology* **261**: 8–14.
25. Nakai, H, Yant, SR, Storm, TA, Fuess, S, Meuse, L and Kay, MA (2001). Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction *in vivo*. *J Virol* **75**: 6969–6976.
26. Dong, JY, Fan, PD and Frizzell, RA (1996). Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum Gene Ther* **7**: 2101–2112.
27. Hermonat, PL, Quirk, JG, Bishop, BM and Han, L (1997). The packaging capacity of adeno-associated virus (AAV) and the potential for wild-type-plus AAV gene therapy vectors. *FEBS Lett* **407**: 78–84.
28. Grieger, JC and Samulski, RJ (2005). Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol* **79**: 9933–9944.
29. Wu, J, Zhao, W, Zhong, L, Han, Z, Li, B, Ma, W *et al.* (2007). Self-complementary recombinant adeno-associated viral vectors: packaging capacity and the role of rep proteins in vector purity. *Hum Gene Ther* **18**: 171–182.
30. Allocca, M, Doria, M, Petrillo, M, Colella, P, Garcia-Hoyos, M, Gibbs, D *et al.* (2008). Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* **118**: 1955–1964.
31. Lu, H, Chen, L, Wang, J, Huack, B, Sarkar, R, Zhou, S *et al.* (2008). Complete correction of hemophilia A with adeno-associated viral vectors containing a full-size expression cassette. *Hum Gene Ther* **19**: 648–654.
32. Nakai, H, Storm, TA, Fuess, S and Kay, MA (2003). Pathways of removal of free DNA vector ends in normal and DNA-PKcs-deficient SCID mouse hepatocytes transduced with rAAV vectors. *Hum Gene Ther* **14**: 871–881.
33. Inagaki, K, Ma, C, Storm, TA, Kay, MA and Nakai, H (2007). The role of DNA-PKcs and artemis in opening viral DNA hairpin termini in various tissues in mice. *J Virol* **81**: 11304–11321.
34. Chen, L, Zhu, F, Li, J, Lu, H, Jiang, H, Sarkar, R *et al.* (2007). The enhancing effects of the light chain on heavy chain secretion in split delivery of factor VIII gene. *Mol Ther* **15**: 1856–1862.
35. Cao, L, Liu, Y, During, MJ and Xiao, W (2000). High-titer, wild-type free recombinant adeno-associated virus vector production using intron-containing helper plasmids. *J Virol* **74**: 11456–11463.
36. Cao, L, During, M and Xiao, W (2002). Replication competent helper functions for recombinant AAV vector generation. *Gene Ther* **9**: 1199–1206.
37. Samulski, RJ, Srivastava, A, Berns, KI and Muzyczka, N (1983). Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. *Cell* **33**: 135–143.
38. Chen, L, Lu, H, Wang, J, Sarkar, R, Yang, X, Wang, H *et al.* (2009). Enhanced factor VIII heavy chain for gene therapy of hemophilia A. *Mol Ther* **17**: 417–424.

See page 6

Effect of Genome Size on AAV Vector Packaging

Zhijian Wu¹, Hongyan Yang¹ and Peter Colosi¹¹Ocular Gene Therapy Laboratory, Neurobiology-Neurodegeneration and Repair Laboratory, National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA

Adeno-associated virus (AAV) vector genomes have been limited to 5 kilobases (kb) in length because their packaging limit was thought to be similar to the size of the parent AAV genome. Recent reports claim that significantly larger vector genomes can be packaged intact. We examined the packaged vector genomes from plasmid-encoded AAV vectors that ranged from 4.7 to 8.7 kb in length, using AAV types 2, 5, and 8 capsids. Southern blot analysis indicated that packaged AAV vector genomes never exceeded 5.2 kb in length irrespective of the size of the plasmid-encoded vector or the capsid type. This result was confirmed by vector genome probing with strand-specific oligonucleotides. The packaged vector genomes derived from plasmid-encoded vectors exceeding 5 kb were heterogeneous in length and truncated on the 5' end. Despite their truncated genomes, vector preparations produced from plasmid-encoded vectors exceeding 5.2 kb mediated reporter gene expression *in vitro* at high multiplicity of infection (MOI). The efficiency of expression was substantially lower than that of reporter vectors with genomes <5 kb in length. We propose that transcriptionally functional, intact vector genomes are generated in cells transduced at high MOI from the fragmentary genomes of these larger vectors, probably by recombination.

Received 25 August 2009; accepted 13 October 2009; published online 10 November 2009. doi:10.1038/mt.2009.255

INTRODUCTION

The adeno-associated viruses (AAVs) are single-stranded DNA viruses with genomes that are ~4.7 kilobases (kb) in length. They comprise the Dependovirus branch of the parvovirus family and require co-infection with a helper virus, typically adenovirus or herpes virus, to replicate. Recombinant vectors derived from AAVs are attractive for the development of human therapeutics because they mediate long-term gene expression and the parent AAVs produce no known pathology in humans. AAV vectors were recently used to successfully treat patients with Leber congenital amaurosis, a severe form of genetically acquired childhood blindness.^{1–3} A substantial limitation of AAV vectors is their small packaging capacity that is generally considered to be <5 kb. This limitation makes vector design for diseases involving larger genes challenging or impossible. Larger genes are relatively common and

are involved in diseases such as cystic fibrosis, Duchenne muscular dystrophy, hemophilia A, and some forms of Leber congenital amaurosis and retinitis pigmentosa. To address this problem, strategies based on trans-splicing^{4–6} and homologous recombination⁷ were developed. All of these methods require multiple vectors, and the efficiency of some of these methods is inherently limited by the efficiencies of intron processing or homologous recombination. Two recent reports suggest that AAV capsids are capable of packaging substantially larger vector genomes in an intact form. Grieger and Samulski⁸ reported that AAV types 1–5 can package up to 6 kb DNA. Most recently, Allocca *et al.*⁹ reported that AAV type 5 can package up to 8.9 kb DNA and mediate efficient transgene expression both *in vitro* and *in vivo*.

In this study, we evaluated the packaged genome size and *in vitro* expression characteristics of AAV vectors ranging from 4.7 to 8.7 kb in length. The vectors examined were a mouse CEP290 vector and a set of lacZ reporter vectors of increasing length. These were packaged into AAV types 2, 5, and 8 capsids using the three-plasmid transfection method. We found that the size of packaged vector genome DNAs never exceeded 5.2 kb no matter what size vector was encoded by the production plasmid. This was true for all capsid types and vectors tested. The packaged vector genome DNA from plasmid-encoded vectors exceeding 5.2 kb consisted of a heterogeneous population of genomes with a broad size distribution and a maximum length of ~5 kb. The heterogeneity was generated by deletion on the 5' end. Vector preparations consisting of heterogeneous fragmentary vector genomes were capable of gene expression *in vitro* when used at high multiplicity of infection (MOI). The expression efficiency of these vector preparations was substantially lower than vector preparations with genomes that were <5 kb in length.

RESULTS

AAV 2, 5, and 8 capsids do not efficiently package an intact 8.6 kb AAV mCEP290 vector genome

Mutations in the human *CEP290* gene account for ~15% of patients with Leber congenital amaurosis,¹⁰ a severe form of inherited blindness. CEP290 is a ubiquitously expressed, 2,480 amino-acid protein that plays roles in ciliary transport and centrosomal function.¹¹ The *rd16* mouse carries an in-frame deletion in the *CEP290* gene and exhibits an early-onset retinal degeneration that is similar to the human disease.¹¹ In an attempt to produce a vector that could be used to complement the CEP290 deficiency in the *rd16* mouse, we constructed an AAV vector plasmid that encoded a

Correspondence: Peter Colosi, Ocular Gene Therapy Laboratory, Neurobiology-Neurodegeneration and Repair Laboratory, National Eye Institute, National Institutes of Health, 6 Center Dr., Room 331B, MSC 0610, Bethesda, Maryland 20892, USA. E-mail: colosip@nei.nih.gov

mouse CEP290 expression cassette driven by the cytomegalovirus (CMV) promoter (Figure 1a). The total length of this AAV vector genome was 8.6 kb including the two inverted terminal repeats (ITRs). We attempted to package this vector into AAV types 2, 5, and 8 capsids using a conventional three-plasmid transfection method.¹² Vector yields, as measured by quantitative PCR, were

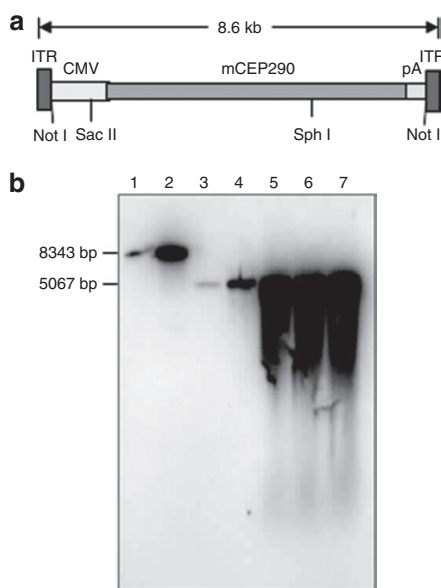


Figure 1 Electrophoresis of AAV-mCEP290 vectors. (a) Structure of the AAV mouse CEP290 vector. (b) Alkaline agarose electrophoresis and Southern blot analysis of isolated AAV mCEP290 vector genomes. Lanes 1 and 2: 8,343 bp vector plasmid DNA marker, 0.1 and 1 ng, respectively; lanes 3 and 4: 5,067 bp vector plasmid DNA marker, 0.1 and 1 ng, respectively; lanes 5–7: DNA from 1×10^{10} vector genomes isolated from AAV2 mCEP290, AAV5 mCEP290, and AAV8 mCEP290 vector preparations, respectively. AAV, adeno-associated virus; bp, base pair; CMV, cytomegalovirus; ITR, inverted terminal repeat; kb, kilobase.

not markedly different between the three AAV serotypes but were approximately tenfold lower than yields from vectors with lengths <4,850 base pairs (bp). Vector genome DNA was purified from the vector preparations, and subjected to alkaline electrophoresis and Southern analysis to assess the size and integrity of the packaged vector genomes (Figure 1b). In contrast to previous studies,^{8,9,13} our results showed that the packaged vector genomes consisted of a heterogeneous population of genomes with a broad size distribution and a maximum length of ~5 kb. No intact 8.6 kb vector genomes were detected. Fragmentary genome packaging was observed with all three serotypes. AAV5 displayed no particular ability to package large vector genomes.

The packaged AAV 5 mCEP290 vector genomes are truncated on the 5' end

To further characterize the packaged AAV5 mCEP290 vector genome population, strand-specific oligonucleotides corresponding to positions located throughout the mCEP290 vector were used to probe dilutions of dot-blotted vector genomes and control plasmid DNA that encodes the complete vector (Figure 2a). The strength of the hybridization signal from the packaged genome was compared to the signal from a linearized vector plasmid control (Figure 2b). The hybridization signals for control plasmid DNA were strong for all probes irrespective of hybridization position in the vector sequence. For the isolated vector genome DNA, the strongest hybridization signals were observed for probes that hybridize near the 3' end of the vector genome. Almost no vector sequences occurring >5.2 kb from the 3' end of the vector were detected. This was true for probes specific to either DNA strand. This result indicated that the majority of the vector genomes in the population have intact 3' ends but truncated 5' ends. The truncations occurred at multiple positions located <5.2 kb from the 3' end of the vector genome. Our study did not determine whether the 5' ends of the packaged genomes contain an ITR.

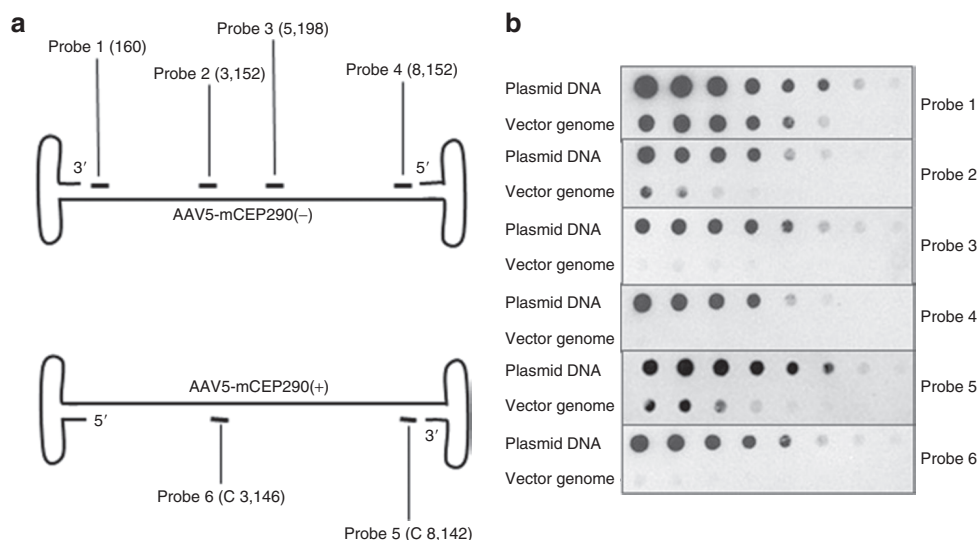


Figure 2 Genome integrity of AAV5-mCEP290 vector. (a) Hybridization positions of single-stranded oligonucleotides on the vector genome. (b) Strand-specific dot-blot hybridization of vector genome DNA. Twofold serial dilutions of vector plasmid DNA (top row) and isolated AAV5 mCEP290 vector genome DNA (bottom row). The masses of the vector plasmid DNAs are 50, 25, 12.5, 6.25, 3.1, 1.6, 0.78, and 0.39 ng (wells 1–8, top row). The calculated masses of the isolated vector DNAs are 7, 3.5, 1.7, 0.86, 0.43, 0.21, 0.11, and 0.05 ng (wells 1–8, bottom row). The DNA in the first well on the bottom row was isolated from 2.3×10^9 vector genomes.

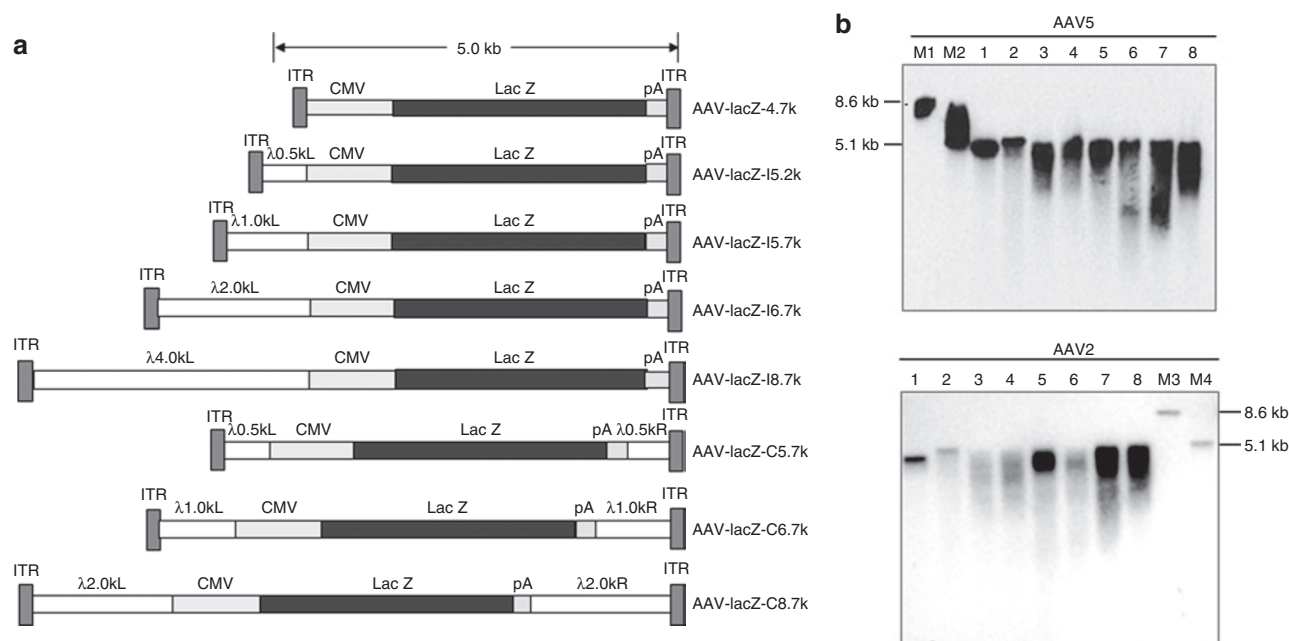


Figure 3 Packaging of AAV2 and AAV5 lacZ vectors. **(a)** Structures of plasmid-encoded vectors ranging from 4.7 to 8.7 kb. All of these vectors contained a 4.4 kb CMV lacZ expression cassette, located adjacent to the ITR (I designation in name) or centrally (C designation in name), and phage lambda DNA was used to increase the total length. **(b)** Southern blot analysis of isolated vector genome DNA separated on an alkaline agarose gel. 1×10^{10} vector genomes were analyzed for the AAV5 vectors (upper panel), and 1×10^9 vector genomes were analyzed for the AAV2 vectors (lower panel). Vector plasmid DNA markers in lanes M1: 8.6 kb (1 ng), M2: 5.1 kb (1 ng), M3: 8.6 kb (0.1 ng), and M4: 5.1 kb (0.1 ng). The DNA in M2 is an unintended partial digestion. Lanes 1–8 in both panels: genomes isolated from AAV-lacZ-4.7k, AAV-lacZ-5.2k, AAV-lacZ-5.7k, AAV-lacZ-6.7k, AAV-lacZ-8.7k, AAV-lacZ-C5.7k, AAV-lacZ-C6.7k, and AAV-lacZ-C8.7k vector preparations, respectively. CMV, cytomegalovirus; ITR, inverted terminal repeat; kb, kilobase.

Estimation of the size limit for efficient packaging of intact vector genomes

To estimate the packaging limit for AAV 2 and 5 capsids, a set of AAV lacZ vectors of increasing length were constructed (Figure 3a). The size of the plasmid-encoded vectors ranged from 4.7 to 8.7 kb. All of these vectors contained a 4.4 kb CMV lacZ expression cassette, located adjacent to the ITR or centrally, and phage lambda DNA was used to increase the total length of the vectors. After packaging the vectors, alkaline gel electrophoresis and Southern analysis were carried out on purified vector genome DNAs to assess their size (Figure 3b). The results were similar for vectors packaged in AAV2 and AAV5 capsids. Both the 4.7 and 5.2 kb vectors showed clear bands at their respective sizes. Vector genome DNAs derived from plasmid-encoded vectors larger than 5.2 kb consisted of a heterogeneous population of genomes with a broad size distribution and a maximum length of ~5 kb. The upper packaging limit for both AAV2 and AAV5 capsids appears to be ~5.2 kb.

Effect of vector size on transgene expression efficiency *in vitro*

The AAV type 2 vector preparations made from plasmid-encoded vectors ranging from 4.7 to 8.7 kb in length (Figure 3a) were assessed for expression *in vitro*. Subconfluent 293 cells were infected with these preparations at MOIs of 2×10^3 and 1×10^4 vector genomes per cell. After 24 hours, the cultures were fixed, stained with X-gal, and counted (Figure 4a). The expression efficiency of the vectors fell into three groups.

The vector preparation made from the 4.7 kb plasmid-encoded vector (pAAV-lacZ-4.7k) had the highest expression at both MOIs and alone comprised the first group (Figure 4b). The second group was made up of vector preparations made from plasmid-encoded vectors that were longer than 5 kb in length but that had intact lacZ expression cassettes and one ITR located entirely within 5,050 bp of continuous DNA. This group was 4- to 15-fold less efficient than AAV2-lacZ-4.7k at both MOIs and included AAV2-lacZ-5.2k, AAV2-lacZ-5.7k, AAV2-lacZ-6.7k, AAV2-lacZ-8.7k, and AAV2-lacZ-C5.7k. The third group was made up of vector preparations made from plasmid-encoded vectors that were longer than 5 kb in length but did not have intact lacZ expression cassettes and one ITR located entirely within 5,050 bp of continuous DNA. This group had expression efficiencies ranging from 77- to 1,600-fold lower than AAV2-lacZ-4.7k depending on the MOI and included AAV2-lacZ-C6.7k and AAV2-lacZ-C8.7k. Most of the vector preparations showed a two- to fourfold increase in transduction efficiency when the MOI was increased fivefold. The expression efficiency of AAV2-lacZ-C8.7k preparation was extremely sensitive to the fivefold increase in MOI and increased 18-fold.

Interestingly, all of the vector preparations mediated some degree of expression even though Southern analysis could not detect packaged genomes >5.2 kb in length. It appears that vectors that can package at least one ITR and the entire lacZ expression cassette in ≤5,050 bp of continuous DNA have greater expression efficiencies than those that cannot.

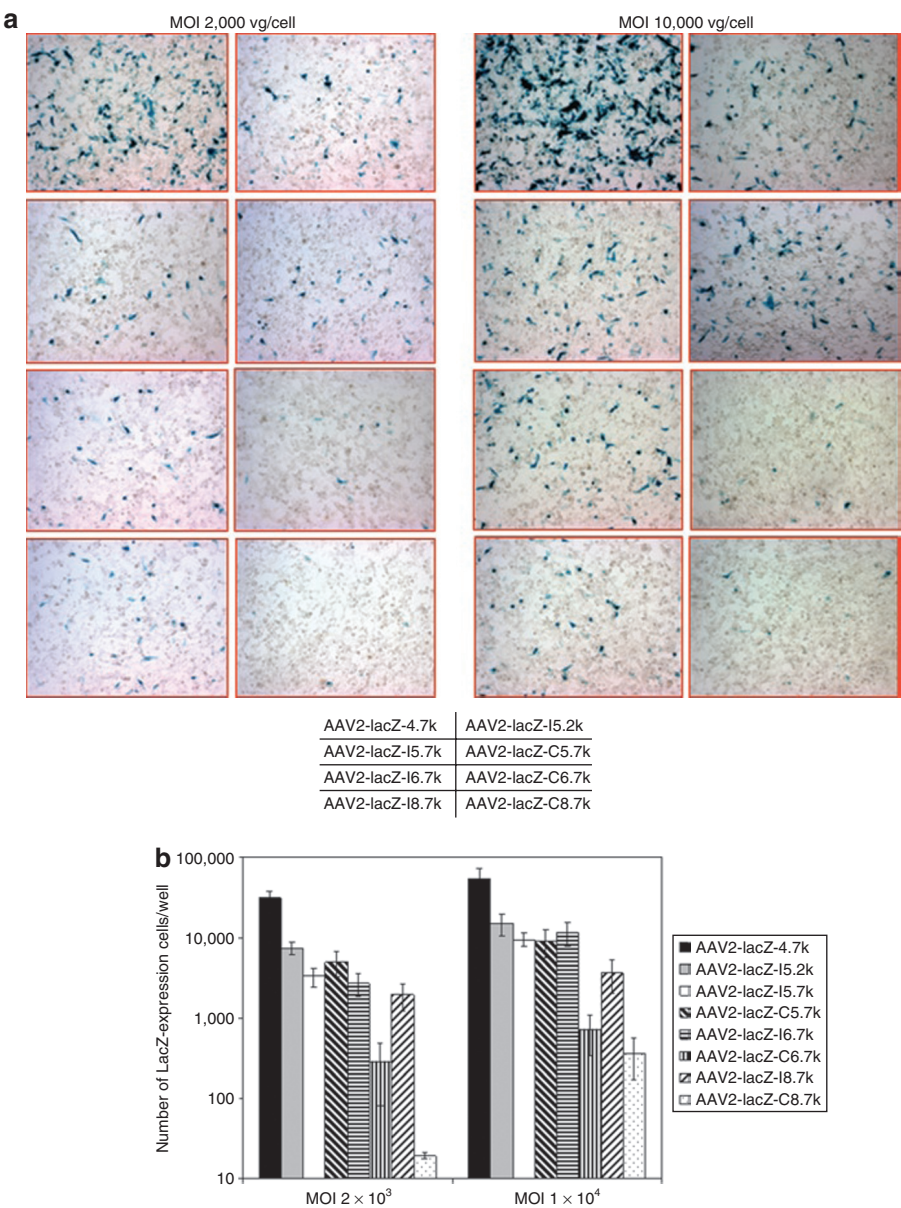


Figure 4 LacZ expression in 293 cells infected with AAV2-lacZ vectors. **(a)** LacZ expression in 293 cells infected with AAV2-lacZ vectors at MOI 2,000 (left panels) or 10,000 (right panels). **(b)** Quantification of transduction of the AAV2 lacZ vector set. MOI, multiplicity of infection.

DISCUSSION

Several groups have examined AAV packaging and have tried to assess the maximum packaging capacity of AAV vectors. Dong *et al.*¹⁴ were the first to systematically address the subject and reported that AAV type 2 vectors made from plasmid-encoded vectors exceeding 4.9kb showed substantially reduced expression when used to transduce cells *in vitro*. Subsequently, Grieger and Samulski⁸ reported that genomes up to 6.0kb could be packaged in AAV types 1–5 capsids, but that vectors exceeding 5.2kb were largely degraded in infected cells by the proteasome. Two years later, Wu *et al.*¹³ reported that a self-complementary vector with an intact 6.6kb genome could be produced using the AAV type 2 capsid. Most recently, Allocca *et al.*⁹ reported that AAV type 5 was able to package intact AAV genomes up to 8.9kb in length and that these vector preparations mediated gene expression *in vitro* and *in vivo*.

In this study, we investigated ability of AAV types 2, 5, and 8 capsids to package nine different plasmid-encoded AAV vectors. The vectors examined were an 8.6kb mouse CEP290 vector and a set of lacZ reporter vectors that ranged from 4.7 to 8.7kb in length. In contrast to some of the previous studies, we found that maximum size of the packaged vector DNA never exceeded ~5.2kb, no matter what size plasmid-encoded vector or AAV serotype was used. Given the limit of sensitivity of our Southern analysis (0.33% of the applied vector genome DNA), if intact 8.6kb CEP290 or 8.7kb lacZ genomes were packaged, they comprised <1 part in 300 (by mass) of the total packaged DNA. Additionally, we found that vector genomes isolated from vector preparations produced with plasmid-encoded vectors exceeding 5.2kb were heterogeneous with a broad size distribution and had a maximum length of ~5kb. Strand-specific probing of packaged

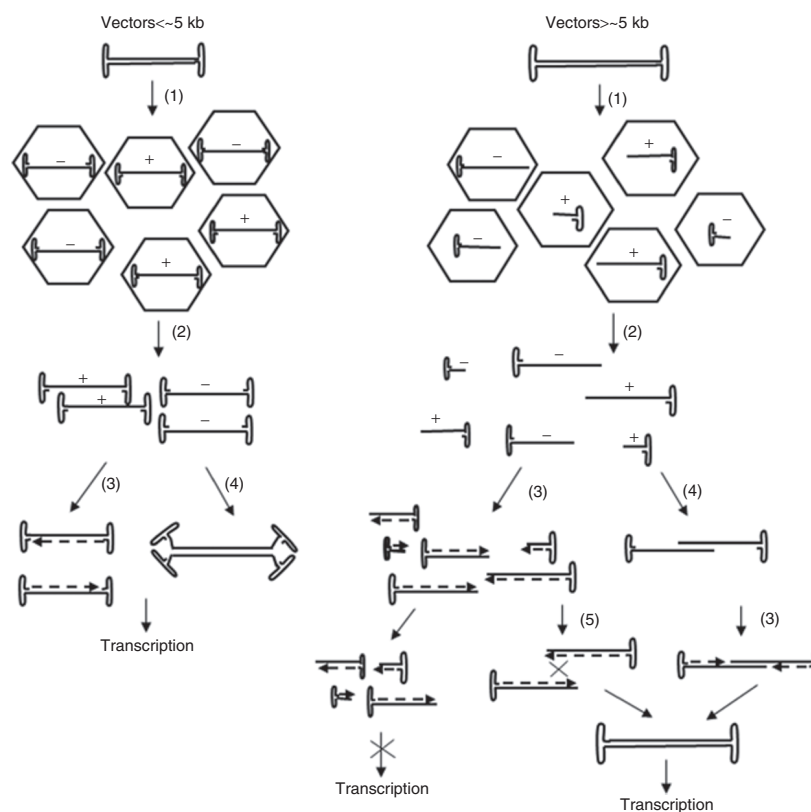


Figure 5 Proposed model for packaging and transduction of AAV2 lacZ vectors. (1) Packaging, (2) uncoating, (3) second-strand DNA synthesis, (4) annealing, and (5) homologous recombination.

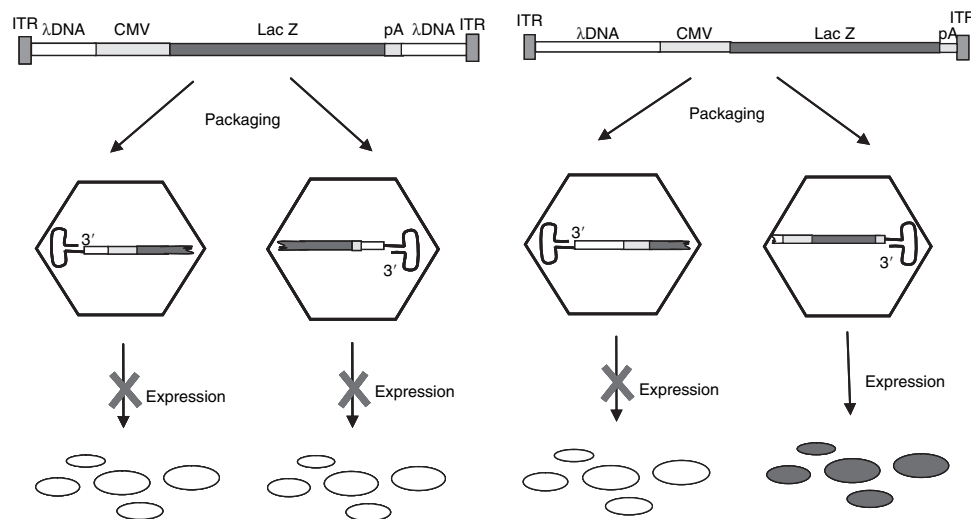


Figure 6 Proposed model for fragmentary genome transduction. Right panel: one ITR and intact lacZ cassette contained in DNA fragment <5,050bp in length. Left panel: intact lacZ cassette and either ITR contained in a DNA fragment >5,050bp in length. The higher transduction efficiency of vectors with ITR-proximal lacZ expression cassettes may indicate that these vectors package the first 5,050bp from the 3' end, which encodes one ITR and an intact lacZ expression, and that this fragmentary genome is capable of mediating gene expression without having to anneal to a second vector. bp, base pair; CMV, cytomegalovirus; ITR, inverted terminal repeat.

CEP290 vector genomes indicated they were exclusively derived from the 5 kb region located at the 3' end of the vector genomes (of both polarities). Apparently, the genomes were truncated at the 5' end. This may be due to the fact that AAV vectors package genomes unidirectionally into preformed capsids starting with the 3' end.¹⁵ With larger vectors, the packaging process may stall after

the capsid is full. The remaining unpackaged 5' end of the genome may then be removed by cellular nucleases.

We then asked whether these fragmentary vector genomes, packaged in AAV type 2 capsids, were capable of mediating gene expression in cells. Somewhat surprisingly, all of the AAV2 lacZ vector preparations mediated some degree of reporter gene

expression when used to transduce 293 cells at MOIs exceeding 1. The reason that vector preparations with fragmentary genomes mediate gene expression is not clear. A likely possibility is that overlapping vector genomes of opposite polarities anneal and regenerate full-length vector genomes in cells that are infected by multiple vector particles (Figure 5). Generation of full-length vector genomes may also occur by homologous recombination. Multiple infections are probably common at the high MOIs used in our experiment. Intracellular vector recombination of AAV vectors has been previously described.⁷

The location of the 4.4kb CMV lacZ cassette in the vector genome had a large effect on transduction activity of the vectors. Vector preparations produced from 5.2 to 8.7kb plasmid-encoded vectors with intact lacZ expression cassettes and one ITR located entirely within 5,050bp of continuous DNA-mediated expression are about tenfold more efficiently than vector preparations derived from plasmid-encoded vectors that did not have intact lacZ expression cassettes and one ITR located entirely within 5,050bp of continuous DNA. The higher efficiency of vectors with ITR-proximal lacZ expression cassettes may indicate that these vectors simply package the first 5,050bp from the 3' end, which encodes one ITR and an intact lacZ expression, and that this partial genome is capable of mediating gene expression without having to anneal to a second vector. This idea is illustrated in Figure 6.

The transduction efficiency of the AAV2-lacZ-C8.7k vector preparation alone demonstrated a very high sensitivity to MOI, jumping 18-fold, versus two- to fourfold for the other vectors, in response to a fivefold increase in MOI. This vector and the AAV2-lacZ-C6.7k vector do not have intact lacZ expression cassettes within 5.2kb of either end of their genomes and would likely require intracellular recombination to regenerate an intact lacZ expression cassette. Assuming a maximum packaged genome size of 5.2kb, the AAV2-lacZ-C8.7k vector would have the smallest overlap between genomes of opposite polarities (1.7kb versus 3.7kb for the AAV2-lacZ-C6.7k vector). This smaller overlap may have made the expression efficiency of this vector more MOI sensitive. The genome populations in these vector preparations are heterogeneous with respect to size. Certainly, the smaller overlap would reduce the fraction of genomes capable of annealing within the cell.

It is also unclear why our data do not agree with previous reports that vector genomes as large as 8.9kb can be packaged and are functional for transduction. To control for differences in packaging plasmids and vector purification techniques, we repeated our experiments with two other sets of packaging plasmids^{9,16} and one other purification protocol¹⁷ with no change in outcome (data not shown).

In conclusion, we found no evidence that AAV capsids efficiently package vector genomes that are substantially larger than the wild-type viral genomes. The limitations imposed by the packaging limits of AAV capsids are major hindrances to the development of many otherwise promising and clinically relevant applications. If AAV vector-based gene delivery is to be used for these applications, strategies need to be developed to accommodate the genome size limitation. The recent reports of large genome packaging usually come with steep reductions in packaging or

transduction efficiency. Considering our data, or the data of Dong *et al.*,¹⁴ this is not surprising. More work certainly needs to be done to see whether capsid or genome modifications can be found that will allow the efficient packaging of larger genomes. But we believe that the current trend of simply trying to package larger genomes into the commonly used AAV capsids is akin to fitting a small shoe onto a large foot.

MATERIALS AND METHODS

AAV vector constructs and production. AAV2 ITRs were used in all AAV vector constructs. The mouse CEP290 complementary DNA was provided by Anand Swaroop (National Eye Institute/National Institutes of Health). To clone the mCEP290 complementary DNA into an AAV vector plasmid, PCR was carried out using the following primers: 5'-AGT CAT CGA TGC CAC CAT GCC ACC TAA TAT AAA GTG GAA AGA ATT AA-3' and 5'-AGT CAG ATC TCT AAT AAA TAG GGA AAC TAT-3'. The PCR product was digested with Cla I and Bgl II, and then inserted into an AAV vector plasmid with an expression cassette consisting of a CMV promoter and a human β -globin polyadenylation site. The parent plasmid was described previously and was used for enhanced green fluorescent protein expression.¹⁸ The resulting plasmid was designated pV4.6c mCEP290. The plasmid encoding the 4.7kb, CMV lacZ vector was previously described.¹⁹ Phage lambda DNA of different lengths was used to create the larger lacZ vectors. PCR was used to create lambda DNA fragments of the appropriate sizes. The PCR products were inserted with either 3' or 5' of the CMV lacZ expression cassette. The designations of the resulting vectors are shown in Figure 3.

AAV production and purification was carried out as previously described.¹² Real-time PCR was employed to titer the vectors. The following fluorescence-labeled probe (5'-TCC AAA ATG TCG TAA CAA CT-3') and primers (5'-TGG GAG TTT GTT TTG CAC CAA-3' and 5'-CGC CTA CCG CCC ATT TG-3') were used for all vectors.

Alkaline gel electrophoresis and Southern blot analysis of AAV vector DNA. DNA was extracted from 1×10^9 or 1×10^{10} full vector particles as previously described.⁹ Before extraction, the vector particles were treated with DNase I to remove unpackaged DNA. Alkaline agarose gel electrophoresis was performed as previously described.²⁰ DIG-labeled (Roche, Indianapolis, IN) or ³²P-labeled probes were employed for DNA hybridization. To analyze AAV mCEP290 vectors, a 408bp Eco RI fragment on pV4.6c mCEP290 was used as a probe. The 5.1 and 8.3kb markers were produced by digestion of the pV4.6c mCEP290 with Sac II and Sph I, or Not I alone. To analyze AAV LacZ vectors, a 657bp Spe I-Age I fragment from the vector plasmid was used as a probe. The 5.1 and 8.6kb markers were generated by Sma I digestion of pAAV-lacZ-15.2k and pAAV-lacZ-C8.7k plasmids, respectively.

Table 1 Probes for strand-specific hybridization of AAV5-mCEP290 genome

Probe	Location at vector genome	Sequence
1	160–199	5'-ATT AAT AGT AAT CAA TTA CGG GGT CAT TAG TTC ATA GCC C-3'
2	3,152–3,174	5'-ACA TCT TTT GCA GGA ACT AGA CA-3'
3	5,198–5,222	5'-TCT TAA AAC ACA AGC AAC TTG CAA G-3'
4	8,152–8,173	5'-TTG GAT TTG AAC TGC CTA GTC C-3'
5	C8,161–C8,142	5'-TCA AAT CCA AAC TGT TCC GA-3'
6	C3,173–C3,146	5'-GTC TAG TTC CTG CAA AAG ATG TAT TAA G-3'

Strand-specific dot-blot hybridization of AAV5 mCEP290 vector. AAV5 mCEP290 vector DNA was extracted as previously described.⁹ Extracted vector DNA from 2.3×10^9 vector genomes was applied to the first well of the manifold. Twofold serial dilutions of the extracted vector DNA were applied to the remaining seven wells of the manifold. The corresponding amounts of vector DNA applied to the eight wells of the manifold were 7, 3.5, 1.7, 0.86, 0.43, 0.21, 0.11, and 0.05 ng. Standard DNA generated by Sma I digestion of pV4.6c mCEP290 was diluted in the same fashion. The amounts of plasmid DNA dotted onto the eight wells of the manifold were 50, 25, 12.5, 6.25, 3.1, 1.6, 0.78, and 0.39 ng. Both vector and plasmid DNA were denatured by 0.25 mol/l NaOH before loading to nylon membrane. Single-strand oligonucleotides corresponding to several positions in the vector genome (Table 1) were designed, synthesized, and labeled using DIG Oligonucleotide 3'-End Labeling Kit (Roche). Hybridization was carried out overnight at 10°C below the melting temperature of each probe. After probing, the membrane was stripped of DIG-labeled probe by washing twice with 0.2 mol/l NaOH/0.1% sodium dodecyl sulfate for 15 minutes at 37°C and rinsing in 2× saline-sodium citrate for 5 minutes.

Infection and quantification of transduced cells. HEK 293 cells were seeded on 24-well plates at a concentration of 1×10^5 cells per well 40 hours before infection. The cells were transfected with pAdeno5, plasmid encoding the E2A, VA RNA, and E4 regions of adenovirus type 2 (ref. 19). The cells were then infected with the AAV2 LacZ vectors at MOIs of 2×10^3 or 1×10^4 . Etoposide was added to each well to produce a final concentration of 20 nmol/l. After 24 hours, the cells were fixed with phosphate-buffered saline containing 2% formaldehyde and 0.2% glutaraldehyde for 5 minutes at room temperature. After fixation, lacZ expression was evaluated by histochemical staining with a solution of 1 mg/ml X-gal in phosphate-buffered saline containing 5 mmol/l $K_3Fe(CN)_6$, 5 mmol/l $K_4Fe(CN)_6 \cdot 3H_2O$, and 2 mmol/l $MgCl_2$. The cells were stained for 4 hours at 37°C. Eight fields per well were counted, and the average number of blue cells per well was determined.

ACKNOWLEDGMENTS

We thank Chun Y Gao of the National Eye Institute for assistance with cell imaging.

REFERENCES

- Maguire, AM, Simonelli, F, Pierce, EA, Pugh, EN Jr, Mingozzi, F, Bennicelli, J *et al.* (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* **358**: 2240–2248.
- Bainbridge, JW, Smith, AJ, Barker, SS, Robbie, S, Henderson, R, Balaggan, K *et al.* (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* **358**: 2231–2239.
- Cideciyan, AV, Aleman, TS, Boye, SL, Schwartz, SB, Kaushal, S, Roman, AJ *et al.* (2008). Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proc Natl Acad Sci USA* **105**: 15112–15117.
- Nakai, H, Storm, TA and Kay, MA (2000). Increasing the size of rAAV-mediated expression cassettes in vivo by intermolecular joining of two complementary vectors. *Nat Biotechnol* **18**: 527–532.
- Yan, Z, Zhang, Y, Duan, D and Engelhardt, JF (2000). Trans-splicing vectors expand the utility of adeno-associated virus for gene therapy. *Proc Natl Acad Sci USA* **97**: 6716–6721.
- Sun, L, Li, J and Xiao, X (2000). Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. *Nat Med* **6**: 599–602.
- Halbert, CL, Allen, JM and Miller, AD (2002). Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol* **20**: 697–701.
- Grieger, JC and Samulski, RJ (2005). Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol* **79**: 9933–9944.
- Allocca, M, Doria, M, Petrillo, M, Colella, P, Garcia-Hoyos, M, Gibbs, D *et al.* (2008). Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* **118**: 1955–1964.
- den Hollander, AI, Roepman, R, Koenekoop, RK and Cremers, FP (2008). Leber congenital amaurosis: genes, proteins and disease mechanisms. *Prog Retin Eye Res* **27**: 391–419.
- Chang, B, Khanna, H, Hawes, N, Jimeno, D, He, S, Lillo, C *et al.* (2006). In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse. *Hum Mol Genet* **15**: 1847–1857.
- Grimm, D, Zhou, S, Nakai, H, Thomas, CE, Storm, TA, Fuess, S *et al.* (2003). Preclinical in vivo evaluation of pseudotyped adeno-associated virus vectors for liver gene therapy. *Blood* **102**: 2412–2419.
- Wu, J, Zhao, W, Zhong, L, Han, Z, Li, B, Ma, W *et al.* (2007). Self-complementary recombinant adeno-associated viral vectors: packaging capacity and the role of rep proteins in vector purity. *Hum Gene Ther* **18**: 171–182.
- Dong, JY, Fan, PD and Frizzell, RA (1996). Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum Gene Ther* **7**: 2101–2112.
- King, JA, Dubielzig, R, Grimm, D and Kleinschmidt, JA (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20**: 3282–3291.
- Rabinowitz, JE, Rolling, F, Li, C, Conrath, H, Xiao, W, Xiao, X *et al.* (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.
- Wu, Z, Sun, J, Zhang, T, Yin, C, Yin, F, Van Dyke, T *et al.* (2008). Optimization of self-complementary AAV vectors for liver-directed expression results in sustained correction of hemophilia B at low vector dose. *Mol Ther* **16**: 280–289.
- Park, TK, Wu, Z, Kjellstrom, S, Zeng, Y, Bush, RA, Sieving, PA *et al.* (2009). Intravitreal delivery of AAV8 retinoschisin results in cell type-specific gene expression and retinal rescue in the Rsl-KO mouse. *Gene Ther* **16**: 916–926.
- Matsushita, T, Elliger, S, Elliger, C, Podsakoff, G, Villarreal, L, Kurtzman, GJ *et al.* (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther* **5**: 938–945.
- Sambrook, J and Russell, DW (eds) (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York. pp 999.

See page 6

Evidence for the Failure of Adeno-associated Virus Serotype 5 to Package a Viral Genome ≥ 8.2 kb

Yi Lai¹, Yongping Yue¹ and Dongsheng Duan¹¹Department of Molecular Microbiology and Immunology, School of Medicine, The University of Missouri, Columbia, Missouri, USA

Limited packaging capacity hinders adeno-associated virus (AAV) gene therapy. A recent study seems to have provided a solution to this problem. Allocca *et al.* reported that AAV-5 could package an 8.9 kb vector genome. Here we tested whether this approach can be used to deliver a large genome for Duchenne muscular dystrophy (DMD) gene therapy. We first evaluated AAV-5 packaging of an 8.2 kb genome. This vector carries two independent reporter gene cassettes, one for *alkaline phosphatase* (AP) and another for *LacZ*. Viral yield was log-fold lower than that of a regular AAV-5. Nevertheless, both AP and LacZ genes were detected in purified virus. Injection to dystrophic muscle resulted in both AP and LacZ expression. On electron microscopy, virion structure appeared normal. Surprisingly, we did not find the full-length single-stranded viral genome by alkaline gel electrophoresis. Neither did we see the full-length double-stranded replication forms in adenovirus coinfecting cells. We suspect that AP and LacZ expression may have come from partially packaged 5' or 3'-half of the genome. Additional studies revealed failure of AAV-5 to package and express an 8.7 kb *minidystrophin* gene cassette. In summary, our results do not support the extraordinary packaging capacity of AAV-5.

Received 25 August 2009; accepted 13 October 2009; published online 10 November 2009. doi:10.1038/mt.2009.256

INTRODUCTION

Adeno-associated virus (AAV) has become one of the most favorite gene delivery vehicles over the last two decades.^{1–4} Recent clinical success further raises the hope of treating inherited diseases with AAV gene therapy.^{5–7} Although a highly efficient and quite safe viral vector,⁸ AAV has a major limitation. It is generally believed that the maximal packaging capacity of an AAV vector is 5 kb. This presents a great hurdle for therapeutic expression cassettes that exceed this limit.

To overcome this obstacle, a number of single and dual vector approaches have been explored with various levels of success.^{9,10} Dual vector approaches often require complicated molecular engineering and transgene reconstitution may highly depend on the target gene sequence.^{11–13} On the other hand, several groups have reported packaging of a ~6 kb genome in a single

AAV virion.^{14–16} Though encouraging, a 6 kb packaging capacity remains insufficient for many therapeutic genes such as a 7 kb *minidystrophin* gene for Duchenne muscular dystrophy (DMD) gene therapy.¹⁷

Recently, Allocca *et al.* screened a series of AAV serotypes for their tolerance to large viral genome.¹⁸ Surprisingly, they noticed a transgene and purification method independent packaging of an up to 8.9 kb genome by AAV serotype 5 (AAV-5).¹⁸ Importantly, they achieved structural and functional amelioration of recessive Stargardt's disease in a mouse model by delivering an 8.9 kb *adenosine triphosphate-binding cassette transport family* gene, *ABCA4*, expression cassette to the retina.¹⁸ Encouraged by this observation, here we tested whether AAV-5 can be used to deliver a large genome to dystrophin-deficient mdx mice, a murine DMD model. We examined two vector genomes including (i) an 8.2 kb genome carrying two intact reporter gene expression cassettes for *alkaline phosphatase* (AP, at the 5' end) and β -galactosidase (*LacZ*, at the 3' end), and (ii) an 8.7 kb genome containing the 7 kb Δ H2-R15 *minidystrophin* gene. The viruses derived from these two genomes were referred to as AV.AP.LacZ and AV. Δ H2-R15.

In contrast to the results reported by Allocca *et al.*,¹⁸ we failed to detect incorporation of the full length of these oversized genomes in AAV-5 virus. Immunostaining with N-terminal and C-terminal specific antibodies also failed to reveal minidystrophin expression in AV. Δ H2-R15 infected mdx muscle. Interestingly, we did observe coexpression of AP and LacZ in AV.AP.LacZ infected mdx muscle. However, this expression appeared to have derived from a partial packaging of either the 5' or the 3' end of the genome. In summary, our results do not support serotype-specific packaging of an oversized genome by AAV-5.

RESULTS

Characterization of AAV-5 AV.AP.LacZ virus

To examine AAV-5 packaging of a large genome, we constructed an 8.2 kb template by flanking two independent reporter gene cassettes with AAV-2 inverted terminal repeats (ITRs) (Figure 1). Each expression cassette contains its own transcriptional regulatory elements including a Rous sarcoma virus (RSV) promoter and a SV40 pA signal. The 5'-end cassette expresses the AP gene and the 3'-end one expresses the *LacZ* gene (Figure 1).

After confirming expression, the template plasmid was used for AAV-5 production.¹⁹ After four rounds of isopycnic ultracentrifugation in cesium chloride gradient, we obtained a viral stock

Correspondence: Dongsheng Duan, Department of Molecular Microbiology and Immunology, One Hospital Drive, Columbia, Missouri, USA. E-mail: duand@missouri.edu

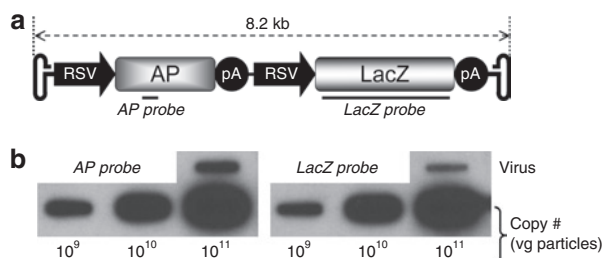


Figure 1 AAV-5 packaging of an 8.2 kb viral genome. **(a)** Schematic outline of the putative full-length viral genome. The 5'-half genome contains an AP expression cassette and the 3'-half genome contains a LacZ expression cassette. RSV, Rous sarcoma virus promoter; pA, SV40 polyadenylation signal; AP, alkaline phosphatase gene; LacZ, β -galactosidase gene. The locations of the AP and LacZ probes are marked. **(b)** A representative slot blot of the purified AV.AP.LacZ virus. vg, viral genome.

with a physical titer of $\sim 7 \times 10^8$ viral genome (vg) particles/ μ l. Both AP and LacZ sequences were detected by slot blot (**Figure 1b**). To compare the packaging efficiency with that of an ideal sized genome, we generated AAV-5 AV.EGFP virus using the same protocol. AV.EGFP has a 4.7 kb genome and it expresses the enhanced green fluorescence protein (EGFP).²⁰ Consistent with Allocca *et al.*,¹⁸ the yield of AV.EGFP was approximately tenfold higher than that of AV.AP.LacZ (data not shown).

Next, we delivered AV.AP.LacZ virus to the tibialis anterior muscle of adult mdx mice. One month later, we examined AP and LacZ expression by histochemical staining (**Figure 2**). Robust AP expression was observed whereas LacZ staining appeared weak. Importantly, the relative levels of AP expression closely mirrored that of LacZ expression in the corresponding cells in serial sections (**Figure 2**).

Wild-type AAV-5 virion is a 25 nm icosahedral particle.²¹ To determine whether a large genome has alternated virion structure, we compared AV.AP.LacZ and AV.EGFP by transmission electron microscopy (**Figure 3a**). In both cases, more than 98% viral particles were uniform fully packaged virions with an average size of approximately 25 nm. No gross change was detected in AV.AP.LacZ stock (**Figure 3a**).

To determine whether the full 8.2 kb genome was incorporated in viral particle, we examined purified virus by alkaline gel Southern blot (**Figure 3b**). Surprisingly, we did not see the 8.2 kb full-length band. Instead, AP and LacZ probes yielded distinctive sets of bands, all smaller than 5 kb. The AP probe revealed two prominent bands migrated at ~ 3.6 kb and ~ 3.9 kb. There also appeared a faint smear around 5 kb. Interestingly, the LacZ probe revealed a ~ 5 kb band and some barely visible low molecular weight smear (**Figure 3b**). It is currently not clear why the AP and LacZ probes yielded the viral genome of different lengths. Nevertheless, our result supports a model of segmental rather than full-length genome packaging.

To confirm this unexpected finding, we examined replication forms of the AV.AP.LacZ virus. HeLa cells were coinfecting with AV.AP.LacZ and adenovirus, Ad.dl802.²² Low molecular weight Hirt DNA was extracted 24 hours later. Southern blot was performed using either the AP or LacZ probe (**Figure 3c**).²² Consistent with our alkaline gel result, we failed to detect the 8.2 kb replication monomer. Using the AP probe, we observed a series of bands

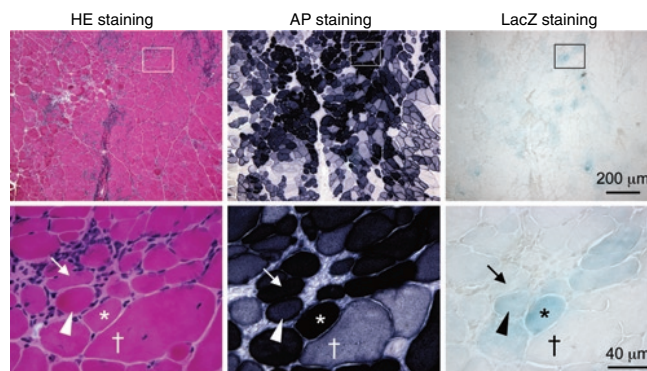


Figure 2 Characterization of AV.AP.LacZ infection in dystrophic muscle. Representative photomicrographs of HE, AP, and LacZ stained muscle sections. Transgene expression was detected at 1 month after AV.AP.LacZ infection ($N = 4$ muscles). Bottom panels are high magnification photomicrographs of the boxed areas in the respective low power images. Asterisk, a myofiber with high level AP expression also showed high LacZ expression; Cross, a myofiber with relatively weak AP expression also showed minimal LacZ expression; Arrow, a myofiber with robust AP but poor LacZ expression; Arrowhead, a myofiber with relatively high level LacZ but weak AP expression.

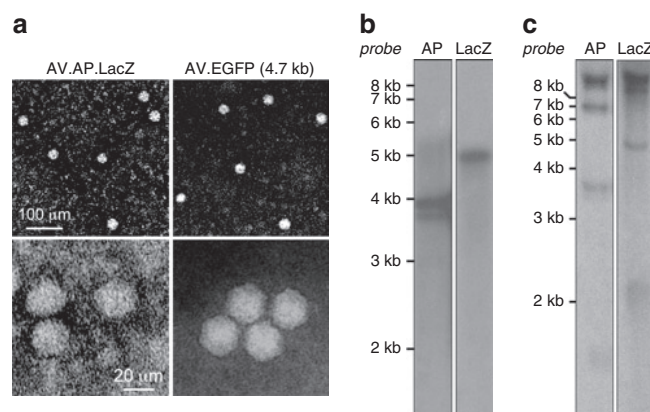


Figure 3 Characterization of AV.AP.LacZ viral particle and the vector genome. **(a)** Representative electron microscopic images of AV.AP.LacZ and AV.EGFP viral particles. AV.EGFP carries a 4.7 kb vector genome. **(b)** A representative alkaline gel Southern analysis of the AV.AP.LacZ viral genome. **(c)** A representative Southern blot examination of the replication forms of AV.AP.LacZ virus. The probes used are marked for each blot. Results in each panel were obtained from at least three independent experiments.

migrated at ~ 3.6 kb, ~ 7.2 Kb and ~ 11 Kb. The size of these bands appears to correlate with the replication form monomer, dimer and trimer of a ~ 3.6 kb viral genome (**Figure 3c**). A similar sets of replication form bands corresponding to a 5 kb viral genome were revealed by the LacZ probe (**Figure 3c**).

Characterization of AAV-5 AV. Δ H2-R15 virus

Our ultimate goal is to deliver a large-size genome for DMD gene therapy. We recently reported a highly functional 7 kb Δ H2-R15 human *minidystrophin* gene. Here we tested whether functional AAV-5 vectors can be generated to express this minigene. Together with the promoter, pA signal and the flanking viral ITRs, the estimated vector genome size is 8.7 kb (**Figure 4a**). This should fit within the 8.9 kb packaging capacity suggested by Allocca *et al.*¹⁸

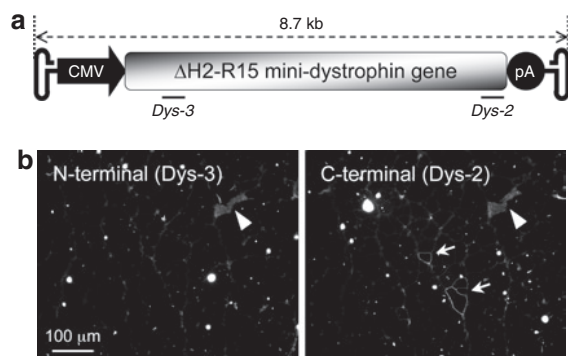


Figure 4 AV.ΔH2-R15 fails to express the minidystrophin gene in mdx muscle. **(a)** Schematic outline of the putative full-length AV.ΔH2-R15 genome. The locations of the Dys-3 and Dys-2 antibodies are marked. **(b)** Representative photomicrographs of serial muscle sections stained with a human dystrophin hinge-1 specific Dys-3 antibody or a dystrophin C-terminal domain specific antibody (Dys-2). Arrow, rare occurring revertant fibers detected by the Dys-3 antibody. Arrowhead, infiltration of immunoglobulin in necrotic myofiber.

AV.ΔH2-R15 viral stock was generated the same way as described above. Electron microscopy revealed normal sized particles at ~25 nm (data not shown). Slot blot with 5' and 3'-end minidystrophin gene probes suggested that the purified viruses contained sequences from both ends (data not shown). A total of 8×10^9 vg particles of AV.ΔH2-R15 were delivered to each tibialis anterior muscle in four adult mdx mice. Immunofluorescence staining was performed at 50 days postinfection. To allow for full-length minidystrophin detection, we used two independent antibodies. Dys-3 specifically reacts with the N-terminal end of human dystrophin whereas Dys-2 recognizes dystrophin C-terminal domain from any species. We were only able to find rare occurring revertant fibers with the Dys-2 antibody (**Figure 4**). Additional alkaline gel and adenovirus coinfection Southern blot experiments confirmed the absence of full-length genome packaging (data not shown).

DISCUSSION

From many standpoints, nano-sized AAV appears to be an ideal gene delivery vehicle.^{1–4} However, a broad application of AAV gene therapy has been hindered by the small viral packaging capacity. A recent report challenged AAV packaging dogma.¹⁸ Allocca *et al.* examined AAV-1, 2, 3, 4, 5, 7, 8, and 9. Surprisingly they found that AAV-5 could tolerate a genome as large as 8.9 kb.¹⁸ It was recently shown that the assembly and the overall mechanical stiffness of the parvovirus are influenced by the interaction between the viral genome and capsid.²³ Because AAV-5 is phylogenetically the most divergent AAV serotype,²⁴ it seems plausible that AAV-5 capsid could be more adaptable to a larger genome.

To capitalize this intriguing finding for DMD gene therapy, we examined AAV-5 packaging and expression from two large viral genomes. We first tested an 8.2 kb genome containing both AP and LacZ expression cassettes. Consistent with Allocca *et al.*,¹⁸ we detected both AP and LacZ sequences in purified virus (**Figure 1**). Further, we observed AP and LacZ coexpression in mdx muscle (**Figure 2**). Despite the encouraging initial results, we were surprised that there was no apparent difference between AAV-5

virions carrying a larger or normal sized genome (**Figure 3a**). Additional molecular characterization suggests that the purified viral stock actually represents a mixture of two partially packaged viruses. One contains the AP expression cassette and the other the LacZ expression cassette (**Figures 3b,c** and **5a**). The observed coexpression may likely reflect efficient coinfection of the same cell by two simultaneously delivered viruses.²⁵

To further confirm this contradictory observation, we evaluated AAV-5 packaging of an 8.7 kb genome containing a single expression cassette for a 7 kb minidystrophin gene (**Figure 4**). On slot blot, electron microscopy, and alkaline gel and adenovirus coinfection Southern blots, we observed results similar to these of AV.AP.LacZ. The only difference was that we did not see minidystrophin expression whereas both AP and LacZ expression were detected in AV.AP.LacZ infected muscle (**Figures 2** and **4**).

Our results suggest that AAV-5 cannot package a genome ≥ 8.2 kb. Similar to AAV-2,^{26–28} infectious single ITR AAV-5 particle is generated when viral genome exceeds the maximal packaging capacity (**Figure 5**). However, the production of such single ITR virion is substantially less efficient.^{26,27} We suspect that a similar mechanism may also explain (at least partially) the findings from Allocca *et al.*¹⁸ We speculate that AAV-5 virus generated by Allocca *et al.* may contain partially packaged viral genomes. A potential sequence overlap between viruses carrying the 5'-end and 3'-end genome may reconstitute expression by homologous recombination (**Figure 5b**).^{29,30} Because overlapping-based reconstitution depends on the target gene sequence,¹² the lack of minidystrophin expression may likely reflect a poor recombination efficiency of the ΔH2-R15 minigene (**Figure 4**).

An interesting aspect of our model is whether the 5'- and the 3'-ITRs are equally efficient in generating single ITR viral particles. Although histochemical staining revealed much intense AP expression (**Figure 2**), we believe that this may not necessarily suggest a preferential packaging of the AP expression cassette. It is well established that the sensitivity of the AP staining method is much higher than that of the LacZ staining method.^{12,25,31} Future studies are needed to completely resolve this issue.

Together with the results from Dong *et al.*³² and Wu *et al.*³³ we believe that AAV-5 may not encapsidate a viral genome ≥ 8.2 kb. The maximal carrying capacity of a single AAV particle remains limited.

MATERIALS AND METHODS

Construction of the cis proviral plasmids. pcis.AP.LacZ (also called YL203) was generated by inserting an RSV-LacZ expression cassette downstream of the AP expression cassette in a previously reported AAV packaging plasmid, pcis.RSV.AP.³⁴ The RSV-LacZ expression cassette was obtained from an already described plasmid, pcis.RSV.LacZ.³⁵ pcis.CMV.ΔH2-R15 (also called YL198) was generated by PCR-mediated multistep cloning. Briefly, the C-terminal domain of the dystrophin gene was PCR amplified using pDysEHpA (a gift from Jeffery Chamberlain, University of Washington, Seattle, WA) as the template and the following pair of primers, forward primer 5'-tcgagactttgccaaggtac (DL1070); reverse primer 5'-g cgcGCATGCaagccatgaattcatatggatcacg (DL1071, the underlined capital nucleotides represent the Sph I site introduced during cloning. This site is located downstream the 3' end of the minidystrophin gene). The PCR product was digested with BstB I and Sph I and swapped into pcis.CMV.ΔR4/ΔC to obtain an intermediate plasmid, pYL197.³⁶ The Nsi I/BsB I fragment of pYL197 was then replaced by a 5,574 bp fragment from pHSA.

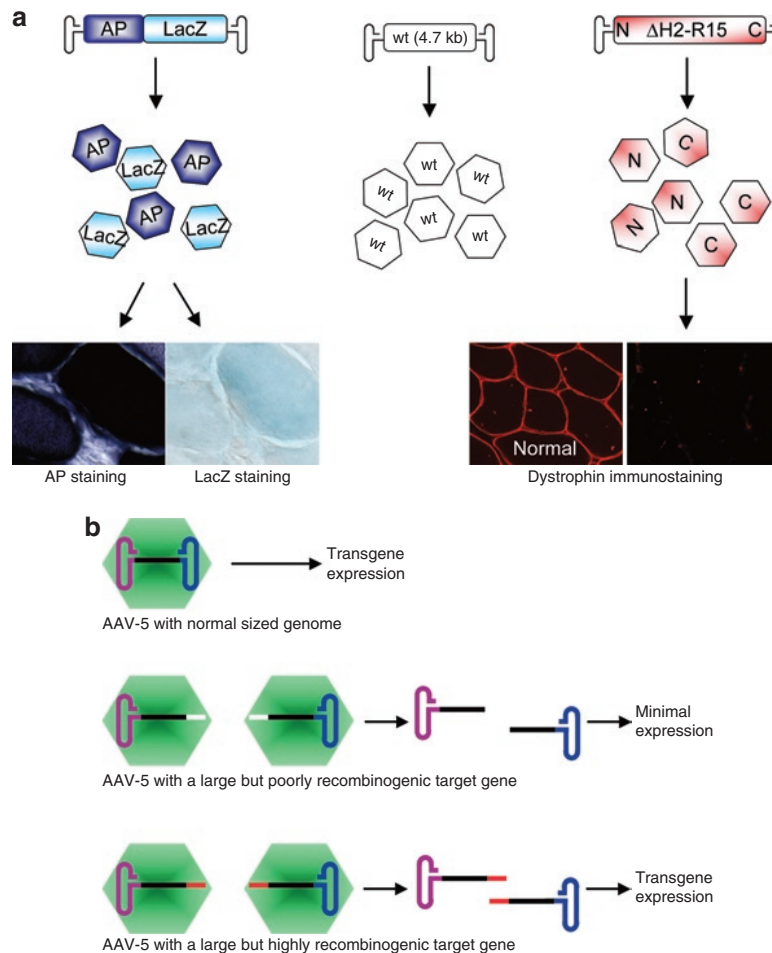


Figure 5 AAV-5 packaging of normal and large sized viral genomes. **(a)** Schematic presentation of AV.AP.LacZ (left panel) and AV.ΔH2-R15 (right panel) packaging. Middle panel depicts AAV-5 packaging of a wild-type (wt) sized genome. AV.AP.LacZ and AV.ΔH2-R15 are consisted of particles carrying either the 5'-half or the 3'-half of the vector genome. In AV.AP.LacZ, independent packaging of each expression cassette by individual viral particle resulted in AP and LacZ coexpression in the same muscle cell. In AV.ΔH2-R15, failure to express minidystrophin is most likely due to an incomplete packaging of the full-length cassette in a single virion. Also included are a representative histochemical and dystrophin immunostaining images from AV.AP.LacZ or AV.ΔH2-R15 infected mdx muscle. The immunostaining image from normal muscle reveals sarcolemmal dystrophin expression. N, the N-terminal end of minidystrophin; C, the C-terminal end of minidystrophin. **(b)** Transgene expression from AAV-5 carrying different size viral genomes. For virus with a normal sized genome, both the 5' and the 3'-end inverted terminal repeats (ITRs) are encapsidated in mature viral particle. For these with an oversized genome, we hypothesize that there is only one ITR in each mature virion. For a large but highly recombinogenic target gene, we speculate that transgene expression may have derived from homologous recombination between the overlapping regions of partially packaged viral genomes.

ΔH2-R15 (also called pYL66).¹⁷ The resulting plasmid was named pYL198. The region obtained by PCR and the regions containing cloning junctions were confirmed by DNA sequencing.

AAV-5 production. AAV-5 stock was generated according to a previously published protocol.¹⁹ Experimental virus was purified from crude viral lysate through four rounds of cesium chloride banding.

Animal studies. Animal experiments were performed in accordance with the NIH and institutional guidelines of the University of Missouri. Male mdx mice were purchased from the Jackson Laboratory (Bar Harbor, ME). Recombinant AAV-5 virus was delivered to the tibialis anterior muscle according to a previously described protocol (6×10^9 vg particles/muscle for AV.AP.LacZ and 8×10^9 vg particles/muscle for AV.ΔH2-R15).²⁰ LacZ and AP expression was analyzed by histochemical staining as described before.^{12,19} Minidystrophin expression was determined by immunofluorescence staining using a human dystrophin hinge-1 specific antibody (Dys-3; Novocastra, Newcastle, UK) and a dystrophin C-terminal

domain specific antibody (Dys-2, Novocastra) according to our published protocol.³⁷

Electron microscopy. Purified AAV-5 virus was placed on a 200 mesh carbon-coated copper grid for five minutes. After four to five rounds of gentle washing in ultra pure water, virus was stained with 1% uranyl acetate for 5 minutes. Viral particles were visualized using a JEOL JEM-1400 transmission electron microscope.

Alkaline gel Southern blot. Viral DNA was extracted from purified virus as reported before.³⁸ After electrophoresis in a 1% alkaline agarose gel, Southern blot was performed using either an AP or a LacZ probe. The AP probe consists of a 359 bp Sac I fragment from pcis.RSV.AP. The LacZ probe consists of a 3,474 bp Not I fragment from pcis.RSV.LacZ.

Examination of replication forms of AAV genome by Southern blot. Hela cells were coinfectd with AAV-5 and adenovirus dl802 as described before.²² Low molecular weight Hirt DNA was extracted 24 hours later and

resolved in a 1% agarose gel.²² Replication form viral genome was examined by AP and LacZ probes as described in alkaline gel Southern.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (AR-49419, DD) and the Muscular Dystrophy Association (DD). The authors also thank Robert J McDonald Jr., MD, for the generous support to Duchenne muscular dystrophy research in Duan lab.

REFERENCES

- Flotte, TR and Berns, KI (2005). Adeno-associated virus: a ubiquitous commensal of mammals. *Hum Gene Ther* **16**: 401–407.
- Carter, BJ (2004). Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective. *Mol Ther* **10**: 981–989.
- Carter, BJ (2005). Adeno-associated virus vectors in clinical trials. *Hum Gene Ther* **16**: 541–550.
- Srivastava, A (2008). Adeno-associated virus-mediated gene transfer. *J Cell Biochem* **105**: 17–24.
- Cideciyan, AV, Hauswirth, WW, Aleman, TS, Kaushal, S, Schwartz, SB, Boye, SL *et al.* (2009). Vision 1 year after gene therapy for Leber's congenital amaurosis. *N Engl J Med* **361**: 725–727.
- Bainbridge, JW, Smith, AJ, Barker, SS, Robbie, S, Henderson, R, Balaggan, K *et al.* (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* **358**: 2231–2239.
- Maguire, AM, Simonelli, F, Pierce, EA, Pugh, EN Jr, Mingozzi, F, Bennicelli, J *et al.* (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* **358**: 2240–2248.
- Frank, KM, Hogarth, DK, Miller, JL, Mandal, S, Mease, PJ, Samulski, RJ *et al.* (2009). Investigation of the cause of death in a gene-therapy trial. *N Engl J Med* **361**: 161–169.
- Duan, D, Yan, Z and Engelhardt, JF (2006). Expanding the capacity of AAV vectors. In: Bloom, ME, SF Cotmore, RM Linden, CR Parrish and JR Kerr (eds). *Parvoviruses*. Hodder Arnold; Distributed in the U.S.A. by Oxford University Press: London, New York. pp. 525–532.
- Ghosh, A and Duan, D (2007). Expanding adeno-associated viral vector capacity: a tale of two vectors. *Biotechnol Genet Eng Rev* **24**: 165–177.
- Lai, Y, Yue, Y, Liu, M, Ghosh, A, Engelhardt, JF, Chamberlain, JS *et al.* (2005). Efficient *in vivo* gene expression by trans-splicing adeno-associated viral vectors. *Nat Biotechnol* **23**: 1435–1439.
- 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, Lai, Y and Duan, D (2008). A hybrid vector system expands adeno-associated viral vector packaging capacity in a transgene-independent manner. *Mol Ther* **16**: 124–130.
- Grieger, JC and Samulski, RJ (2005). Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol* **79**: 9933–9944.
- Warrington, KH Jr, Gorbatyuk, OS, Harrison, JK, Opie, SR, Zolotukhin, S and Muzyczka, N (2004). Adeno-associated virus type 2 VP2 capsid protein is nonessential and can tolerate large peptide insertions at its N terminus. *J Virol* **78**: 6595–6609.
- Lu, H, Chen, L, Wang, J, Huack, B, Sarkar, R, Zhou, S *et al.* (2008). Complete correction of hemophilia A with adeno-associated viral vectors containing a full-size expression cassette. *Hum Gene Ther* **19**: 648–654.
- Lai, Y, Thomas, GD, Yue, Y, Yang, HT, Li, D, Long, C *et al.* (2009). Dystrophins carrying spectrin-like repeats 16 and 17 anchor nNOS to the sarcolemma and enhance exercise performance in a mouse model of muscular dystrophy. *J Clin Invest* **119**: 624–635.
- Allocca, M, Doria, M, Petrillo, M, Colella, P, Garcia-Hoyos, M, Gibbs, D *et al.* (2008). Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* **118**: 1955–1964.
- Duan, D, Yan, Z, Yue, Y, Ding, W and Engelhardt, JF (2001). Enhancement of muscle gene delivery with pseudotyped adeno-associated virus type 5 correlates with myoblast differentiation. *J Virol* **75**: 7662–7671.
- Duan, D, Sharma, P, Yang, J, Yue, Y, Dudus, L, Zhang, Y *et al.* (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.
- Walters, RW, Agbandje-McKenna, M, Bowman, VD, Moninger, TO, Olson, NH, Seiler, M *et al.* (2004). Structure of adeno-associated virus serotype 5. *J Virol* **78**: 3361–3371.
- Duan, D, Sharma, P, Dudus, L, Zhang, Y, Sanlioglu, S, Yan, Z *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.
- Carrasco, C, Castellanos, M, de Pablo, PJ and Mateu, MG (2008). Manipulation of the mechanical properties of a virus by protein engineering. *Proc Natl Acad Sci USA* **105**: 4150–4155.
- Gao, G, Vandenberghe, LH, Alvira, MR, Lu, Y, Calcedo, R, Zhou, X *et al.* (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol* **78**: 6381–6388.
- Xu, Z, Yue, Y, Lai, Y, Ye, C, Qiu, J, Pintel, DJ *et al.* (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.
- Wang, XS and Srivastava, A (1997). A novel terminal resolution-like site in the adeno-associated virus type 2 genome. *J Virol* **71**: 1140–1146.
- Wang, XS, Ponnazhagan, S and Srivastava, A (1996). Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions. *J Virol* **70**: 1668–1677.
- Wang, XS, Qing, K, Ponnazhagan, S and Srivastava, A (1997). Adeno-associated virus type 2 DNA replication *in vivo*: mutation analyses of the D sequence in viral inverted terminal repeats. *J Virol* **71**: 3077–3082.
- Halbert, CL, Allen, JM and Miller, AD (2002). Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol* **20**: 697–701.
- Duan, D, Yue, Y and Engelhardt, JF (2001). Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison. *Mol Ther* **4**: 383–391.
- Bell, P, Limberis, M, Gao, G, Wu, D, Bove, MS, Sanmiguel, JC *et al.* (2005). An optimized protocol for detection of *E. coli* beta-galactosidase in lung tissue following gene transfer. *Histochem Cell Biol* **124**: 77–85.
- Dong, B, Nakai, H, and Xiao, W. Characterization of genome integrity for oversized recombinant AAV vector. *Mol Ther* (in the press).
- Wu, Z, Yang, H and Colosi, P. Effect of genome size on AAV vector packaging. *Mol Ther* (in the press).
- Duan, D, Yue, Y, Yan, Z, Yang, J and Engelhardt, JF (2000). Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus. *J Clin Invest* **105**: 1573–1587.
- Yue, Y and Dongsheng, D (2002). Development of multiple cloning site cis-vectors for recombinant adeno-associated virus production. *BioTechniques* **33**: 672, 674, 676–672, 674, 678.
- Liu, M, Yue, Y, Harper, SQ, Grange, RW, Chamberlain, JS and Duan, D (2005). Adeno-associated virus-mediated microdystrophin expression protects young mdx muscle from contraction-induced injury. *Mol Ther* **11**: 245–256.
- Yue, Y, Li, Z, Harper, SQ, Davisson, RL, Chamberlain, JS and Duan, D (2003). Microdystrophin gene therapy of cardiomyopathy restores dystrophin-glycoprotein complex and improves sarcolemma integrity in the mdx mouse heart. *Circulation* **108**: 1626–1632.
- McCarty, DM, Monahan, PE and Samulski, RJ (2001). Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther* **8**: 1248–1254.