Molecular and Genetic Medicine

Ling et al., J Mol Genet Med 2015, 9:3
http://dx.doi.org/10.4172/1747-0862.1000175

Commentary Open Access

# The Adeno-Associated Virus Genome Packaging Puzzle

**Chen Ling[1,2], Yuan Wang[1-3], Yuan Lu[4], Lina Wang[1-3], Giridhara R Jayandharan[5], George V Aslanidi[1,2], Baozheng Li[1,2], Binbin Cheng[1,3], Wenqin Ma[1,2], Thomas Lentz[6], Changquan Ling[3], Xiao Xiao[6,7], R Jude Samulski[6], Nicholas Muzyczka[2,8,9] and Arun Srivastava[1,2,8-10*]**

[1]*Division of Cellular and Molecular Therapy, Department of Pediatrics, University of Florida College of Medicine, Gainesville, FL, USA*

[2]*Powell Gene Therapy Center, University of Florida College of Medicine, Gainesville, FL, USA*

[3]*Department of Traditional Chinese Medicine, Changhai Hospital, Second Military Medical University, Shanghai, China*

[4]*Department of Orthopedics & Rehabilitative Medicine, University of Florida College of Medicine, Gainesville, FL, USA*

[5]*Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur, India*

[6]*Gene Therapy Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*

[7]*Division of Molecular Pharmaceutics, University of North Carolina School of Pharmacy, Chapel Hill, NC, USA*

[8]*Department of Molecular Genetics & Microbiology, University of Florida College of Medicine, Gainesville, FL, USA*

[9]*Genetics Institute, University of Florida College of Medicine, Gainesville, FL, USA*

[10]*Shands Cancer Center, University of Florida College of Medicine, Gainesville, FL, USA*

[*]**Corresponding author:** Arun Srivastava, Division of Cellular & Molecular Therapy, Cancer and Genetics Research Complex, 2033 Mowry Road, Room 492-A, Gainesville, FL 32611-3633, USA, Tel: (352) 273-8259; Fax: (352) 273-8342; E-mail: aruns@peds.ufl.edu

**Received date:** May 21, 2015; **Accepted date:** July 08, 2015; **Published date:** July 15, 2015

**Copyright:** © 2015 Ling C, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Commentary

Adeno-associated virus (AAV) is a single-stranded (ss) DNA-containing non-pathogenic human parvovirus, and recombinant AAV (rAAV) vectors are currently in use in a number of gene therapy clinical trials [1,2]. Various steps in the life cycle of the wild-type (wt) as well as the rAAV vectors have been studied extensively [3]. Although the single-stranded AAV genomes of both [+] and [-] polarities are encapsidated into separate mature virions with equal frequency [4], the precise underlying mechanism of viral genome encapsidation into viral capsids remains somewhat unclear.

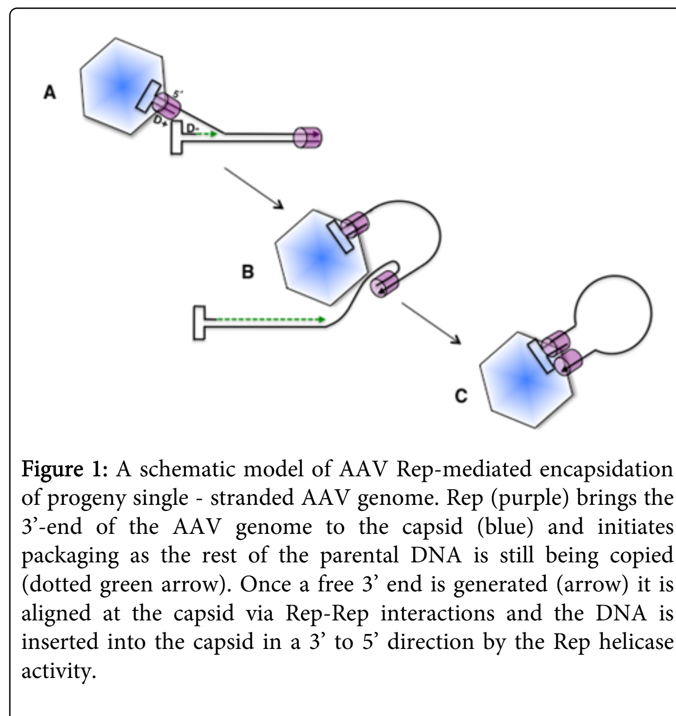


**Figure 1:** A schematic model of AAV Rep-mediated encapsidation of progeny single - stranded AAV genome. Rep (purple) brings the 3'-end of the AAV genome to the capsid (blue) and initiates packaging as the rest of the parental DNA is still being copied (dotted green arrow). Once a free 3' end is generated (arrow) it is aligned at the capsid via Rep-Rep interactions and the DNA is inserted into the capsid in a 3' to 5' direction by the Rep helicase activity.

One model of AAV genome encapsidation suggests that a Rep protein covalently attached to the 5' end of the displaced strand binds to a preformed capsid as the rest of the parental strand is still being copied (Figure 1). When the single strand is completely displaced, the 3'-end, which is also bound by Rep is then positioned on the capsid surface through Rep-Rep interactions to initiate packaging of the single strand in a 3' to 5' direction. This has generally been referred to as 'Rep-mediated encapsidation' [3]. Covalent attachment of Rep to the 5'- end occurs during terminal resolution (Figure 2). During this step of DNA replication, Rep binds to two elements within the terminal repeat, the Rep-binding element (RBE) in the stem structure, and to that involving a small palindrome comprising a single tip of an internal hairpin within the ITR (RBE'). The Rep helicase is activated to unwind the region containing the terminal resolution site (trs), which in turn is stabilized in a single-stranded configuration by a secondary structure (Figure 2A). Rep then nicks the single - stranded trs via a transesterification reaction to produce a new 3' hydroxyl end and a 5' covalently bound Rep via a tyrosine–$PO_4$ linkage (Figure 2B) [5-9].

The AAV genome is flanked by inverted terminal repeats (ITRs) of 145 nucleotides (nts), the terminal 125 of which form palindromic double-stranded T-shaped hairpin structures [10]. The remaining 20 nts in each ITR, termed the D-sequence, remain single-stranded. Although the terminal hairpin structure should be sufficient for initiating replication of the viral genome, what role, if any, the single-stranded D-sequence plays, remained elusive.

Subsequent studies by us and others documented that the D-sequence indeed plays a pivotal role in the life cycle of AAV, such as genome rescue, replication, and integration [11-13]. However, the role of the D-sequence in AAV DNA encapsidation and transgene expression from rAAV vectors remained unclear (Figure 3). We also identified two distinct host cell proteins which were shown to interact with these sequences, as the two D-sequences in the AAV genome are complementary to each other since the ssD[-]-sequence is always present at the 3'-end, and the ssD[+]-sequence is invariably present at the 5'-end of the viral genome [14-16].

**Citation:** Ling C, Wang Y, Lu Y, Wang L, Jayandharan GR, et al. (2015) The Adeno-Associated Virus Genome Packaging Puzzle. J Mol Genet Med 9: 175. doi:10.4172/1747-0862.1000175



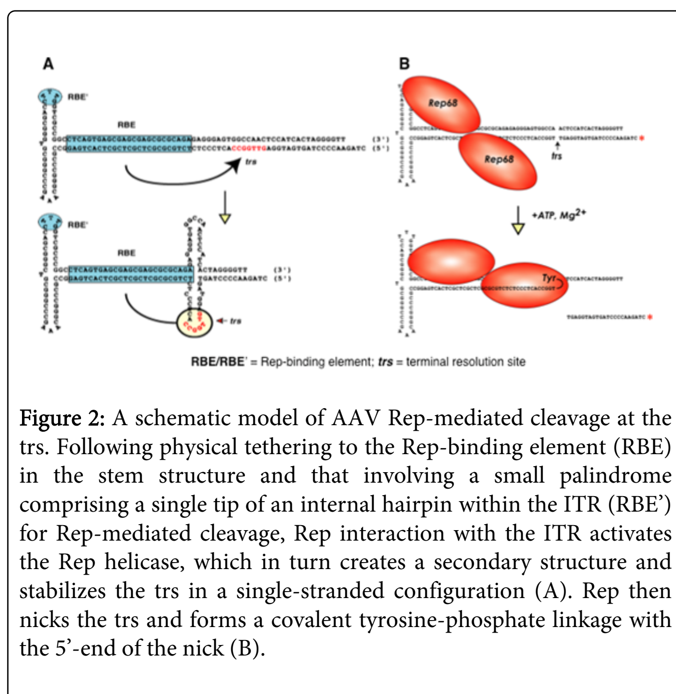


**Figure 2:** A schematic model of AAV Rep-mediated cleavage at the trs. Following physical tethering to the Rep-binding element (RBE) in the stem structure and that involving a small palindrome comprising a single tip of an internal hairpin within the ITR (RBE') for Rep-mediated cleavage, Rep interaction with the ITR activates the Rep helicase, which in turn creates a secondary structure and stabilizes the trs in a single-stranded configuration (A). Rep then nicks the trs and forms a covalent tyrosine-phosphate linkage with the 5'-end of the nick (B).

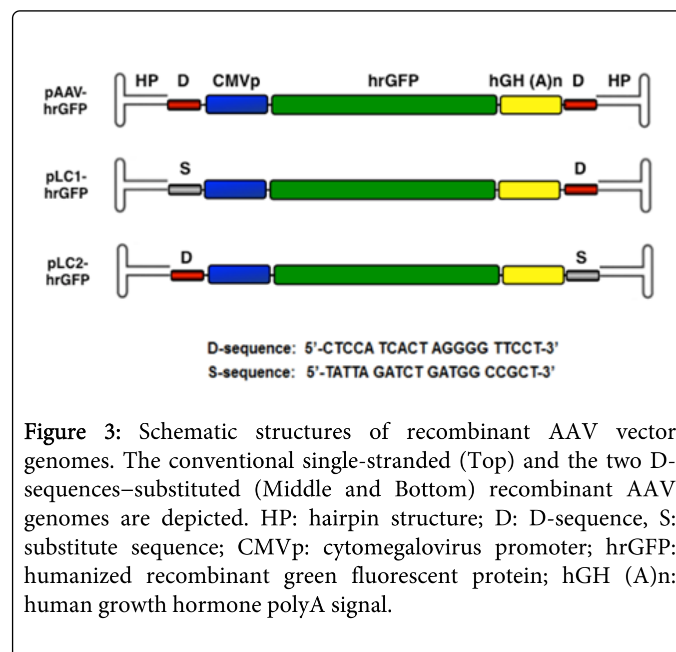


**Figure 3:** Schematic structures of recombinant AAV vector genomes. The conventional single-stranded (Top) and the two D-sequences–substituted (Middle and Bottom) recombinant AAV genomes are depicted. HP: hairpin structure; D: D-sequence, S: substitute sequence; CMVp: cytomegalovirus promoter; hrGFP: humanized recombinant green fluorescent protein; hGH (A)n: human growth hormone polyA signal.

We had documented that replacement of both D-sequences in the AAV-ITRs with a non-AAV substitute sequence (S-sequence) significantly reduced the efficiency of rescue and replication of the AAV genome from recombinant plasmids [11,13], and that no encapsidation of the viral genome into AAV capsids occurred [12]. In our recently published studies [17], we examined the consequences of substitution of only one D-sequence in either ITR on the efficiency of rescue, replication, and encapsidation of the AAV genome. To this end, a plasmid, designated pLC-1, was constructed which contained the human cytomegalovirus (CMV) promoter-driven humanized recombinant green fluorescent protein (hrGFP) gene flanked by a wt AAV2 right ITR and a left ITR in which the D-sequence was replaced with the S-sequence. In a second plasmid, designated pLC-2, the D-sequence in the right ITR was replaced with the S-sequence. A recombinant plasmid, designated pAAV-hrGFP, containing the identical expression cassette flanked by two wt AAV2-ITRs were used as an appropriate control (Figure 3). AAV genome rescue and replication were nearly the same as that from pAAV-hrGFP suggesting that the presence of one D-sequence in either of the two ITRs is necessary and sufficient for the efficient rescue and replication of the AAV genome. Similarly, the presence of one D-sequence in one of the two ITRs was necessary and sufficient for efficient packaging of the AAV genome. Interestingly, however, further analyses revealed that in contrast to plasmid pAAV-hrGFP from which ssAAV vectors of both polarities were generated, as expected, the use of both mutant plasmids pLC-1 and pLC-2 led to the generation of predominantly only single-polarity vectors, as determined by using either [+] or [-] polarity hrGFP-specific oligonucleotide probes. Because the two vector genomes are complementary, our recently published data suggested that only the AAV genomes containing the 5' D[+]-sequence –substitution and a 3' wt D[-]-sequence undergo successful encapsidation. We refer to this as 'D-sequence–mediated encapsidation'.

Most of the published data support the proposed model, shown schematically in Figure 1, in which the displaced strand covalently linked to Rep is targeted for packaging. For example, Nony et al. demonstrated that plasmids that did not have ITRs, but only had an internal Rep-binding element/terminal resolution site (RBE/trs) in the p5 promoter, were also packaged into AAV, albeit at a lower level [18]. They performed deletion mapping and discovered that this was due to the Rep binding element and the cryptic trs, i.e., when either was deleted, they no longer observed packaging (or replication of the ITR-negative AAV genomes). The p5 RBE/trs do not have a D-sequence as seen in the ITR. They also showed that vector (pBR) sequences were packaged, and this was also likely due to the presence of ITRs on the pBR plasmid DNA [18-22]. Huser et al. showed that chromosome 19 sequences, involved in site-specific integration of the AAV genome, were also packaged during normal lytic AAV replication in the presence of co-infection with adenovirus [23]. Chromosome 19 sequences have an RBE and a putative trs that was identified by Giraud et al., and Linden et al. [24,25], but they also do not have the D-sequences that are present in the normal ITR. In spite of that, they can both replicate and package during the AAV replication cycle [23]. Linden et al. [25] also showed that the sequences between the RBE and trs in the chromosome 19 were important for integration, and Weitzman et al., showed that like the normal AAV ITR, the sequences flanking the chromosome 19 trs form a cruciform, which appears to be necessary to stabilize the trs site in a single-stranded conformation [26]. They also showed that this secondary structure was necessary for nicking [26], just as we showed, it was necessary for nicking in the normal AAV ITR [7]. This was true for chromosome 19, even though the sequences flanking the trs were different from the regular ITR [8]. Our D-sequence-deletions/substitutions would not be able to do this, and therefore, they would not replicate, and they would not form covalent ITR/Rep complexes [7-9]. More recently, it was shown by McAlister and Owens [27] that if the AAV RBE and trs are substituted with the chromosome 19 RBE and trs on both ends of AAV, the extent of AAV replication/packaging is reduced by ~2-fold, which would argue that the D-sequences per se are not required, but that a

secondary structure that can stabilize the trs is needed so that Rep can nick and form a covalent linkage, as shown schematically in Figure 2.

However, it is not readily apparent whether the strand-specificity of encapsidation that was observed in our studies is consistent with this model. Thus, even if a secondary structure at the trs cannot be formed, nicking at this site could still occur, albeit at a lower level, which might provide a reasonable explanation as to why exclusive packaging of only one strand was not seen with our D-sequence–substituted mutants [17,28-30].

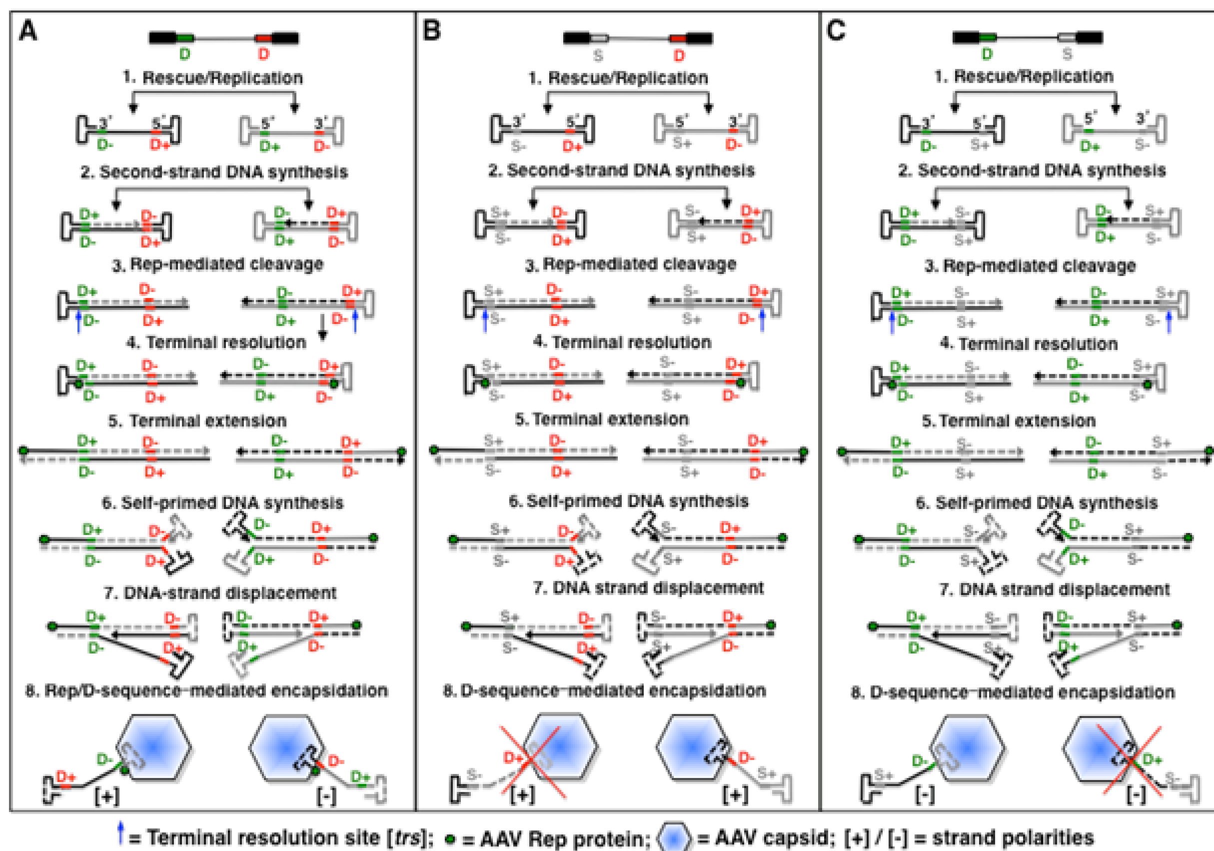


**Figure 4:** Schematic models for replication and encapsidation of AAV genomes from recombinant plasmids containing the wild-type ITRs (A), or the D – sequence – substitutions (grey S) in the left (B), or the right (C) ITR in the viral genome. The conventional AAV genome is shown containing the two D-sequences in the left and the right ITRs as green and red rectangles. Arrows indicate the 3' - end; solid circles indicate covalently bound Rep molecules. See text for details of each step.

We now propose possible models for replication and packaging of rAAV genomes containing either two D-sequences, or those containing the D-sequence deletions in each of the ITRs, which are depicted schematically in Figure 4. The left and the right ITRs are denoted as green and red rectangles, respectively. Upon successful rescue and replication (Step 1), the two single-stranded AAV genomes are generated that are complementary to each other (black and grey lines, respectively). The 3'-end of each of the ssAAV genomes then serves as a primer for the host cell DNA polymerase to catalyze the viral second-strand DNA synthesis to generate the double-stranded DNA monomeric replicative intermediates (Step 2). Following unfolding of the hairpin structures and terminal extension, the two replicative DNA intermediates from both genomes undergo terminal resolution by the AAV Rep protein (Step 3). Following terminal resolution (Step 4), terminal extension (Step 5), self-primed DNA synthesis (Step 6), and DNA-strand displacement (Step 7), single-stranded progeny genomes containing the Rep protein covalently attached at the 5'-ends are encapsidated into AAV capsids (Step 8). The net result is that whereas both [+] and [-] polarity strands of the wt AAV genomes are packaged in separate mature virions, by 'Rep/D-sequence–mediated encapsidation' (Figure 4A), largely single-polarity genomes are generated from AAV genomes containing D-sequence –substitutions, and as a consequence, either the [+] polarity (Figure 4B), or the [-] polarity (Figure 4C) ssD[-]-sequence–containing genome are packaged by 'D-sequence–mediated encapsidation'. This implies that the strand containing a covalently bound Rep protein is not targeted for packaging as suggested by the model in Figure 1. Rather, once terminal resolution has occurred, the opposite end of the genome initiates strand displacement to produce a ssDNA containing a 3' D[-] end and a free D[+] (or S[+]) end with no covalent Rep attached, and this is the genome that is packaged. How the decision to initiate strand displacement synthesis from the end with no covalently bound Rep is made (i.e., the end that had not undergone terminal resolution) is not clear. Conceivably, a covalently bound Rep could

sterically hinder assembly of a cellular DNA replication complex. In addition, even the substituted S[-] end could bind Rep non-covalently (because it has both RBE and RBE') and yet is preferentially not packaged. This suggests that the model in which Rep is the primary signal for packaging is not entirely correct.

In summary, the precise mechanism of AAV genome encapsidation still remains a puzzle, but it is readily apparent that both Rep and the D-sequence play a critical role.

## Acknowledgments

We thank Drs. Kenneth I. Berns and Sergei Zolotukhin for a critical review of this manuscript. This research was supported in part by grants from Bankhead-Coley Cancer Research Program (to CL); and Public Health Service grants P01 DK-058327 (Project 1), R01 HL-097088, and R21 EB-015684 from the National Institutes of Health (to AS), and R01 GM-109524 (to NM). LW was supported by a Fellowship Award from the Government of China, and GRJ was supported in part by an 'Overseas Associate Fellowship-2006' from the Department of Biotechnology, Government of India.

## References

1. Daya S, Berns KI (2008) Gene therapy using adeno-associated virus vectors. Clin Microbiol Rev 21: 583-593.
2. Hastie E, Samulski RJ (2015) Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success--a personal perspective. Hum Gene Ther 26: 257-265.
3. Samulski RJ, Muzyczka N (2014) AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annual Review of Virology 1: 427-451.
4. Berns KI, Rose JA (1970) Evidence for a single-stranded adenovirus-associated virus genome: isolation and separation of complementary single strands. J Virol 5: 693-699.
5. Im DS, Muzyczka N (1989) Factors that bind to adeno-associated virus terminal repeats. J Virol 63: 3095-3104.
6. Ashktorab H, Srivastava A (1989) Identification of nuclear proteins that specifically interact with adeno-associated virus type 2 inverted terminal repeat hairpin DNA. J Virol 63: 3034-3039.
7. Brister JR, Muzyczka N (2000) Mechanism of Rep-mediated adeno-associated virus origin nicking. J Virol 74: 7762-7771.
8. Jang MY, Yarborough OH 3rd, Conyers GB, McPhie P, Owens RA (2005) Stable secondary structure near the nicking site for adeno-associated virus type 2 Rep proteins on human chromosome 19. J Virol 79: 3544-3556.
9. Brister JR, Muzyczka N (1999) Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. J Virol 73: 9325-9336.
10. Srivastava A, Lusby EW, Berns KI (1983) Nucleotide sequence and organization of the adeno-associated virus 2 genome. J Virol 45: 555-564.
11. Wang XS, Ponnazhagan S, Srivastava A (1995) Rescue and replication signals of the adeno-associated virus 2 genome. J Mol Biol 250: 573-580.
12. Wang XS, Ponnazhagan S, 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.
13. Wang XS, Qing K, Ponnazhagan S, 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.
14. Qing K, Wang XS, Kube DM, Ponnazhagan S, Bajpai A, et al. (1997) Role of tyrosine phosphorylation of a cellular protein in adeno-associated virus 2-mediated transgene expression. Proc Natl Acad Sci U S A 94: 10879-10884.
15. Qing K, Hansen J, Weigel-Kelley KA, Tan M, Zhou S, et al. (2001) Adeno-associated virus type 2-mediated gene transfer: role of cellular FKBP52 protein in transgene expression. J Virol 75: 8968-8976.
16. Jayandharan GR, Aslanidi G, Martino AT, Jahn SC, Perrin GQ, et al. (2011) Activation of the NF-kappaB pathway by adeno-associated virus (AAV) vectors and its implications in immune response and gene therapy. Proc Natl Acad Sci U S A 108: 3743-3748.
17. Ling C, Wang Y, Lu Y, Wang L, Jayandharan GR, et al. (2015) Enhanced transgene expression from recombinant single-stranded D-sequence-substituted adeno-associated virus vectors in human cell lines in vitro and in murine hepatocytes in vivo. J Virol 89: 952-961.
18. Nony P, Tessier J, Chadeuf G, Ward P, Giraud A, et al. (2001) Novel cis-acting replication element in the adeno-associated virus type 2 genome is involved in amplification of integrated rep-cap sequences. J Virol 75: 9991-9994.
19. Tessier J, Chadeuf G, Nony P, Avet-Loiseau H, Moullier P, et al. (2001) Characterization of adenovirus-induced inverted terminal repeat-independent amplification of integrated adeno-associated virus rep-cap sequences. J Virol 75: 375-383.
20. Nony P, Chadeuf G, Tessier J, Moullier P, Salvetti A (2003) Evidence for packaging of rep-cap sequences into adeno-associated virus (AAV) type 2 capsids in the absence of inverted terminal repeats: a model for generation of rep-positive AAV particles. J Virol 77: 776-781.
21. Chadeuf G, Ciron C, Moullier P, Salvetti A (2005) Evidence for encapsidation of prokaryotic sequences during recombinant adeno-associated virus production and their in vivo persistence after vector delivery. Mol Ther 12: 744-753.
22. François A, Guilbaud M, Awedikian R, Chadeuf G, Moullier P, et al. (2005) The cellular TATA binding protein is required for rep-dependent replication of a minimal adeno-associated virus type 2 p5 element. J Virol 79: 11082-11094.
23. Hüser D, Weger S, Heilbronn R (2003) Packaging of human chromosome 19-specific adeno-associated virus (AAV) integration sites in AAV virions during AAV wild-type and recombinant AAV vector production. J Virol 77: 4881-4887.
24. Giraud C, Winocour E, Berns KI (1994) Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. Proc Natl Acad Sci U S A 91: 10039-10043.
25. Linden RM, Winocour E, Berns KI (1996) The recombination signals for adeno-associated virus site-specific integration. Proc Natl Acad Sci U S A 93: 7966-7972.
26. Weitzman MD, Kyöstiö SR, Kotin RM, Owens RA (1994) Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc Natl Acad Sci U S A 91: 5808-5812.
27. McAlister VJ, Owens RA (2010) Substitution of adeno-associated virus Rep protein binding and nicking sites with human chromosome 19 sequences. Virol J 7: 218.
28. Xiao X, Xiao W, Li J, Samulski RJ (1997) A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. J Virol 71: 941-948.
29. Zhou X, Zeng X, Fan Z, Li C, McCown T, et al. (2008) Adeno-associated virus of a single-polarity DNA genome is capable of transduction in vivo. Mol Ther 16: 494-499.
30. Zhong L, Zhou X, Li Y, Qing K, Xiao X, et al. (2008) Single-polarity recombinant adeno-associated virus 2 vector-mediated transgene expression in vitro and in vivo: mechanism of transduction. Mol Ther 16: 290-295.