"D" matters in recombinant AAV DNA packaging

To the editor:

Unmethylated CpG dinucleotide-based motifs (CpGs) have been documented as pathogen-associated molecular patterns (PAMPs) that bind to Toll-like receptor (TLR)9 pattern recognition receptors and lead to adaptive effector functions that result in a loss of transduced cells.[1] Vector genome hypomethylation has been recognized as a deficiency of current methods of generation of adeno-associated viral (AAV) vectors.[2] Although wild-type parvovirus DNA has been shown to be generally hypomethylated[3] in its replication cycle, one additional source of hypomethylation in recombinant AAV (rAAV) genomes stems from the encapsidation of unreplicated vector genomes, which are often directly derived from bacterial plasmid DNA origin.[2] Since the mainstay method for clinical AAV production has been based on triple plasmid transfection, understanding the mechanism of the direct packaging of plasmid-derived DNA is necessary to find a solution to this problem.

Traditionally, we think of AAV vector production as involving DNA replication. AAV vector genomes are constructed by flanking the expression cassette with two copies of AAV inverted terminal repeat (ITR) as the essential *cis* elements to facilitate the replication of vector DNA and packaging.[4] A key element in the AAV ITR identified to control such a function is the D-element ("D"), which is a 20-nucleotide sequence that is highly conserved among various AAV serotypes. "D" is a unique region of AAV ITR that is not part of the palindromic sequences. In addition, the "D" is located near the terminal resolution site, which is the initial site for rescue, replication, and packaging.[5–7] AAV rep genes and capsid genes along with a set of helper virus genes are provided in *trans* to complement the functionality of the AAV ITR and D-element during vector production.

Nevertheless, it is challenging to study the role of D-element function because of the strong secondary structure in AAV ITRs and the necessity of replacing the two copies of D-elements in the AAV genomes simultaneously. We overcame this technical hurdle by utilizing a self-complimentary vector in our study, which allowed a random library of D-sequences to be analyzed in the rAAV packaging process. The analysis of the D-element requirement for rAAV vector encapsidation revealed an astonishing outcome: wild-type AAV (WTAAV) ITR D-sequences were recovered only ~25% of the time (Figure S1). This suggests that any sequence replacing the D-element (so called "pseudo-D" sequence) could be utilized by rAAV for packaging, indicating a lax requirement for D-sequences in rAAV encapsidation and providing an explanation for the existence of plasmid backbone in the rAAV preparations.

Textbooks state that the D-element is essential for WTAAV replication and packaging. However, the results from rAAV packaging and the above studies clearly show a difference between rAAV and WTAAV. We think that the DNA replication differences between the rAAV production system and WTAAV propagation are the primary cause for such a discrepancy. WTAAV propagation usually arises from initial low MOI infection, and successful productive lytic infection from a single virus infection or latent viral genome rescue would occur naturally. This means that viral DNA will have to be amplified extensively before encapsidation. As illustrated in Figure 1A, if a mutant "D" emerged because of replication error or for some other reason, even assuming that this mutant "D" had a modestly lower replication efficiency (say ~80%), DNA with a mutant "D" would only account for 0.1% of sequences and effectively be eliminated after 30 rounds of replication to reach a low titer of 1E10 genomes from a single genome. Therefore, such mutant D-elements, even with a slight disadvantage in replication, are less likely to survive during WTAAV replication and propagation, thus promoting predominance of the highly conservative D-sequences observed among many different AAV serotypes.

However, the packaging process of rAAV vectors is different. As illustrated in Figure 1B, AAV ITRs are mostly symmetric in the vector plasmid used for production. D-elements function to specifically promote the vector genome part of the plasmid to be replicated and packaged. Upon rescue of the vector genome from the plasmid molecule, owing to the symmetry of the ITRs, two DNA fragments are expected to be generated. One is the vector genome containing the D-sequences that would lead to the replication and packaging of the desired rAAV vectors. The other part is the plasmid backbone, which retains all AAV ITR palindromic sequences except the unique D-sequences. Therefore, we asked whether the 20 nucleotides immediately adjacent to the ITRs in the plasmid backbone side could act as alternative "D" sequences, or so called "pseudo-D" elements. Although such pseudo-D sequences have been documented to be deficient in replication, previous observations and our new data (see Figure S2) demonstrate that pseudo-"D"s are still compatible with AAV encapsidation. Indeed, in a typical AAV production system, viral particles with plasmid backbone may represent as much as one-tenth of all viral particles produced.[2,8]

The high level of contamination with plasmid backbone provides more evidence for limited replication of AAV vector DNA in these production systems. Since WT D is not the dominating or decisive factor to exclude packaging pseudo-D DNA into AAV capsid, the ratio of WT-D DNA and pseudo-D plasmid backbone DNA is therefore dictating the composition of DNA in the packaged capsid, as illustrated in Figure 1B. After rescue from the plasmid DNA, the ratio of pseudo-D DNA and vector DNA is 1:1, without any DNA replication. In this scenario, if pseudo-D DNA were packaged as efficiently as WT-D DNA, the amount of vector with plasmid backbone would be the same as vector DNA with D-sequences. However, if the vector DNA with D-element were to be replicated twice, then the ratio of plasmid backbone to vector DNA in the vector would be 1:3. Such a balance would vary according to the property of the "pseudo-D" that happens to be in the

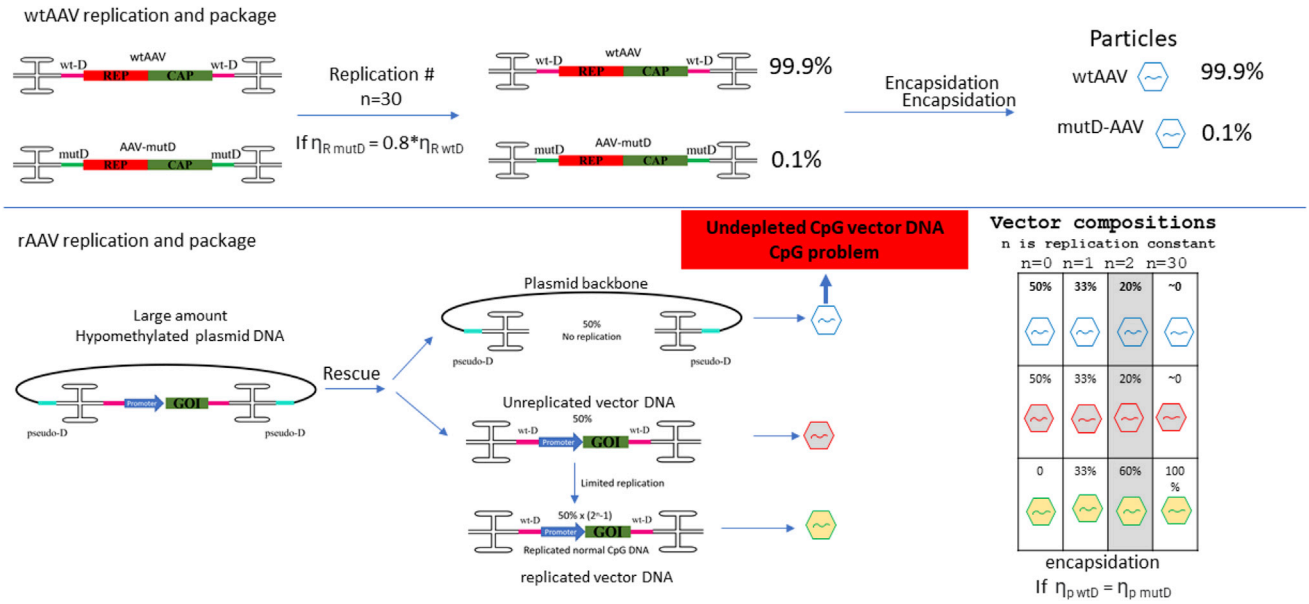


Figure 1. The role of D sequences in replication and packaging leads to the difference in composition between WTAAV and rAAV composition

(A) Illustration of WTAAV population composition in a hypothetical situation when a mutant D-element replication efficiency ($\eta_{R\ mutD}$) is only about 80% of WT D ($\eta_{R\ mutD}$). $\eta_{p\ wtTD}$ and $\eta_{p\ muD}$ are packaging efficiency of WTD and mutD, which is assumed to be similar in this case. The replication difference allows the WTAAV to become dominant. (B) Illustration of recombinant AAV composition when a large amount of vector plasmid DNA is introduced to the host cells, followed by rescue and limited replication of rescued vector genomes. The table shows the hypothetical composition based on the level of replicated vector genomes. η_R stands for replication efficiency and η_p stand for packaging efficiency.

vector DNA plasmid. Since the plasmid backbone is commonly observed at 10% of all genomes in AAV vectors, limited vector DNA replication must have taken place prior to rAAV packaging (if pseudo-D is not as efficient as WT D in packaging, vector DNA actually replicates even less). If, on the other hand, the vector DNA were to be replicated 30 times, vector plasmid DNA would be diluted out and therefore not affect the vector composition.

Thus, the lax requirement of D-element in packaging is a major factor contributing to vector composition, as reflected by the considerable level of contamination of plasmid backbone,[8] which is typically found in vector preparations using plasmids as template. Despite extensive DNase clearing, viral particles containing the bacterial backbone are still observed in even the most highly purified preparations.[9,10] Two aspects of current production systems further compound the problem. First, vast amounts of input plasmids are used for AAV production and highly efficient transfection methods are employed,[11] such that replicating forms of AAV genomes are not required to be the predominant species in rAAV production system for high-yield rAAV production, which greatly increases the chances of packaging plasmid backbone into virions. Second, due to the necessity of downregulating rep78 during packaging,[12] expression of the latter is repressed.[13,14] Lower levels of rep78 protein in combination with other helper and cellular factors is presumably another contributing factor, such that newly replicating DNA is not the dominant substrate for the packaging machinery. Because current production systems are prone to packing non-vector plasmid sequences, the field should consider extending CpG depletion beyond the AAV vector genome into the plasmid backbone in order to further decrease immunogenicity of the vectors.

Lax requirement of the D-element in rAAV packaging and inefficient replication of vector DNA are substantially affecting the composition of rAAV vectors used for human gene therapy, resulting in a high abundance of plasmid backbone contamination and increased presence of hypomethylated CpG sequences in rAAV genomes.[2] Increasing the size of the plasmid backbone was reported to significantly reduce, but not eliminate, reverse packaging.[15] Considering the less selective nature of the D-element in AAV encapsidation, it might be advantageous to adopt plasmid backboneless vector DNA for rAAV packaging. For instance, "doggie bone" DNA has been utilized for production of other biologicals such as lentiviral vectors. Finally, it has been proposed to develop production systems that incorporate methylation of CpG sequences, which might help drive high-level transgene expression.[2] Such an effort would certainly require an improved vector production system that eliminates direct packaging of unmethylated input vector sequences. If input DNA were fully methylated, limited DNA replication regulated by the D-element, might become an advantage since AAV DNA by itself is hypomethylated in the productive lytic infection. In conclusion, although D does not matter that much in rAAV packaging, it does matter immensely in vector performance and safety.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ymthe.2021.05.002.

Junping Zhang,[1] Ping Guo,[2] Yinxia Xu,[2] Patrick L. Mulcrone,[1] R. Jude Samulski,[3] and Weidong Xiao[1]

[1]Herman B Wells Center for Pediatric Research, Indiana University, Indianapolis, IN 46202, USA; [2]Lewis Katz School of Medicine, Temple University, Philadelphia, PA 19140, USA; [3]Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

https://doi.org/10.1016/j.ymthe.2021.05.002

Correspondence: R. Jude Samulski, Gene Therapy Center and Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
E-mail: rjs@med.unc.edu

Correspondence: Weidong Xiao, Herman B Wells Center for Pediatric Research, Indiana University, Indianapolis, IN 46202, USA.
E-mail: xiaow@iu.edu

REFERENCES

1. Faust, S.M., Bell, P., Cutler, B.J., Ashley, S.N., Zhu, Y., Rabinowitz, J.E., and Wilson, J.M. (2013). CpG-depleted adeno-associated virus vectors evade immune detection. J. Clin. Invest. *123*, 2994–3001.
2. Wright, J.F. (2020). Codon Modification and PAMPs in Clinical AAV Vectors: The Tortoise or the Hare? Mol. Ther. *28*, 701–703.
3. Tóth, R., Mészáros, I., Hüser, D., Forró, B., Marton, S., Olasz, F., Bányai, K., Heilbronn, R., and Zádori, Z. (2019). Methylation Status of the Adeno-Associated Virus Type 2 (AAV2). Viruses *11*, 11.
4. Xiao, X., Xiao, W., Li, J., and Samulski, R.J. (1997). A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. J. Virol. *71*, 941–948.
5. Snyder, R.O., Samulski, R.J., and Muzyczka, N. (1990). In vitro resolution of covalently joined AAV chromosome ends. Cell *60*, 105–113.
6. Wang, X.S., Ponnazhagan, S., and Srivastava, A. (1995). Rescue and replication signals of the adeno-associated virus 2 genome. J. Mol. Biol. *250*, 573–580.
7. Wang, X.S., 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.
8. Wright, J.F. (2014). Product-Related Impurities in Clinical-Grade Recombinant AAV Vectors: Characterization and Risk Assessment. Biomedicines *2*, 80–97.
9. Chadeuf, G., Ciron, C., Moullier, P., and 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.
10. Salvetti, A., Orève, S., Chadeuf, G., Favre, D., Cherel, Y., Champion-Arnaud, P., David-Ameline, J., and Moullier, P. (1998). Factors influencing recombinant adeno-associated virus production. Hum. Gene Ther. *9*, 695–706.
11. Grieger, J.C., Soltys, S.M., and Samulski, R.J. (2016). Production of Recombinant Adeno-associated Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector. Mol. Ther. *24*, 287–297.
12. Mouw, M.B., and Pintel, D.J. (2000). Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. J. Virol. *74*, 9878–9888.
13. Li, J., Samulski, R.J., and Xiao, X. (1997). Role for highly regulated rep gene expression in adeno-associated virus vector production. J. Virol. *71*, 5236–5243.
14. Xiao, W., Chirmule, N., Berta, S.C., McCullough, B., Gao, G., and Wilson, J.M. (1999). Gene therapy vectors based on adeno-associated virus type 1. J. Virol. *73*, 3994–4003.
15. Hauck, B., Murphy, S.L., Smith, P.H., Qu, G., Liu, X., Zelenaia, O., Mingozzi, F., Sommer, J.M., High, K.A., and Wright, J.F. (2009). Undetectable transcription of cap in a clinical AAV vector: implications for preformed capsid in immune responses. Mol. Ther. *17*, 144–152.

YMTHE, Volume *29*

Supplemental Information

"D" matters in recombinant AAV DNA packaging

Junping Zhang, Ping Guo, Yinxia Xu, Patrick L. Mulcrone, R. Jude Samulski, and Weidong Xiao

Supplementary results

Utilizing the scheme depicted in supplementary Figure 1, we used self-complimentary vectors and systematically analyzed the D-element requirement for rAAV vector encapsidation. Since there are two copies of ITR in a typical AAV vector along with two identical D-elements at each end of the AAV genomes, it is a challenge to replace the two D-sequences in the AAV genomes simultaneously. As shown in Figure 1B, native AAV ITR in scAAV is replicated and duplicated in replication and packaging process. This allowed us to make a plasmid vector library with randomized D -element. After such vector DNA libraries were introduced into AAV production cells, for those molecules that were competent for AAV packaging, the ITRs with mutation were duplicated and restored to the normal dual ITR configuration during the rescue and replication process. Therefore, replacing the single D sequence in scAAV vector had the same biological effects as replacing the two ITRs in a regular AAV vector. The obtained vectors were used to infect the test cell line GM16095 cells and the recovered vector DNA in the test cells were sequenced and the newly identified nucleotides in the D-sequence region of ITR were compared to wild type D sequences. The results shown in Figure 1C clearly demonstrated wtAAV ITR D sequences were only recovered at around 25% or so, which shows the lax requirement for D sequences in rAAV packaging.

To answer the issue of plasmid backbone packaging in rAAV production, we tested the compatibility of the junction sequence next to ITR in the common AAV vector for supporting AAV packaging. The results are shown in the supplementary figure 2. Although wt-D is 10 times more efficient than those non-canonical D-sequences, this backbone “pseudo-D” still allows approximately 10% packaging level as that of the wt-D element.

Experiment Material and methods:

Cell lines

HEK 293 cells and GM16095 cells (a human fibroblast cell line purchased from the Coriell Institute) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 100 μg of penicillin/ml, and 100 units/ml of streptomycin (Invitrogen, Carlsbad, CA). All cells were maintained in a humidified 37 °C incubator with 5% CO_2.

rAAV production and purification

The transgene plasmids packaged in the AAV vectors were described previously[1]. AAV vectors were generated using the triple plasmid co-transfection method[1,2]. Briefly, pAAV-Rep-Cap (serotypes 2 and 8), pAd helper, and the transgene plasmids were co-transfected into HEK293 cells and cultured in roller bottles at a ratio of 1:1:1. The vectors from transfected cells and medium were harvested 72 hours post transfection and purified by two rounds of CsCl gradient ultracentrifugation and extensively dialyzed against PBS containing 5% D-sorbitol. Viral genomes (vg) were quantified by real-time PCR and vector titers are expressed as vg/ml.

D-sequence library and D-variant construction

The plasmid pdsAAV-CB-GFP containing one 145-nt ITR and 88-nt mutant ITR was used as backbone. The 145-nt ITR was replaced with 131-nt wild type ITR by PvuII and Mscl restriction enzymes. New constructed pdsAAV-CB-GFP plasmid was digested with XbaI and Bbsl restriction and ligated with 8-nt, 18-nt, 24-nt random D sequence library respectively, as shown in Figure 1. Thus, Core sequences of the terminal resolution sites (trs) site were replaced with randomized 8, 18, and 24 oligonucleotides. D sequence library was referred to as pseudo “D” sequence lib.

The pseudo “D” lib-CB-GPF plasmid was co-transfected with pFΔ6 and ph22 into HEK293 cells at the ratio of 1:1:1. Cell culture medium was collected at 48 hours after transfection and was used to infect GM16095 cells for rAAV vector evaluation. Cells were lysed, and DNA was extracted at 12 hours following infection. Amplification of pseudo “D” lib was performed using 5 μl of the extracted DNA as input to a 50 μl AccuTaq LA (Sigma) reaction mixture, according to the manufacturer’s instructions, using 1 μl Forward primer (5’-CTTCTAGTTGCCAGCCATC-3’) and 1 μl Reverse primer(5′-GGCCTCAGTGAGCGAGCGAG-3′) with PCR cycling conditions of 98°C for 30 s, 30 cycles of 94°C for 15 s, 50°C for 20 s, and 68°C for 2 min, followed by 68°C for 10 min. TA cloning of PCR products were performed using pCR™2.1 cloning vector (The TA Cloning® Kit with pCR™2.1 vector, Thermos fisher Scientific, US), according to the manufacturer’s instructions. Positive colonies were verified by restriction digestion and sequenced by Sanger sequencing. These verified D-sequences were separately inserted into pdsAAV-CB-GFP plasmid containing wild type ITR by XbaI and Bbsl restriction enzymes, thus specific D-variant vector was constructed.

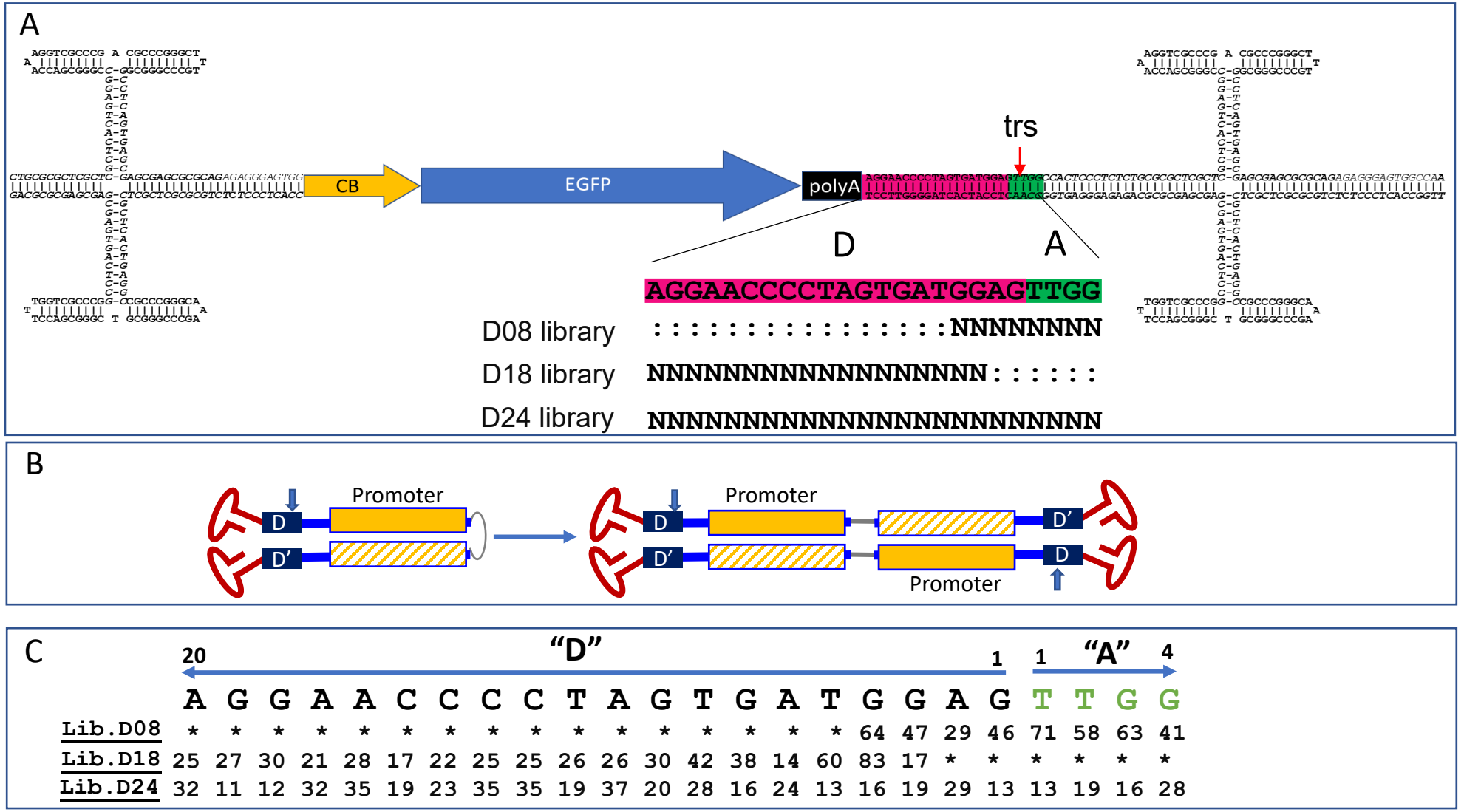
Figure Legend:

Supplementary Figure 1. A. Schematic illustration of D element library construction and screening diagram. The ITR related sequences are shown in the scAAV vector with D-element and terminal resolution region highlighted. The region randomized in the D08, D18 and D24 library are shown as N in the alignment with wild type ITR sequences. B. Illustration of the duplication of D sequence in scAAV replication. C. The summary of the percentage of recovered wild type AAV nucleotide in D sequences.

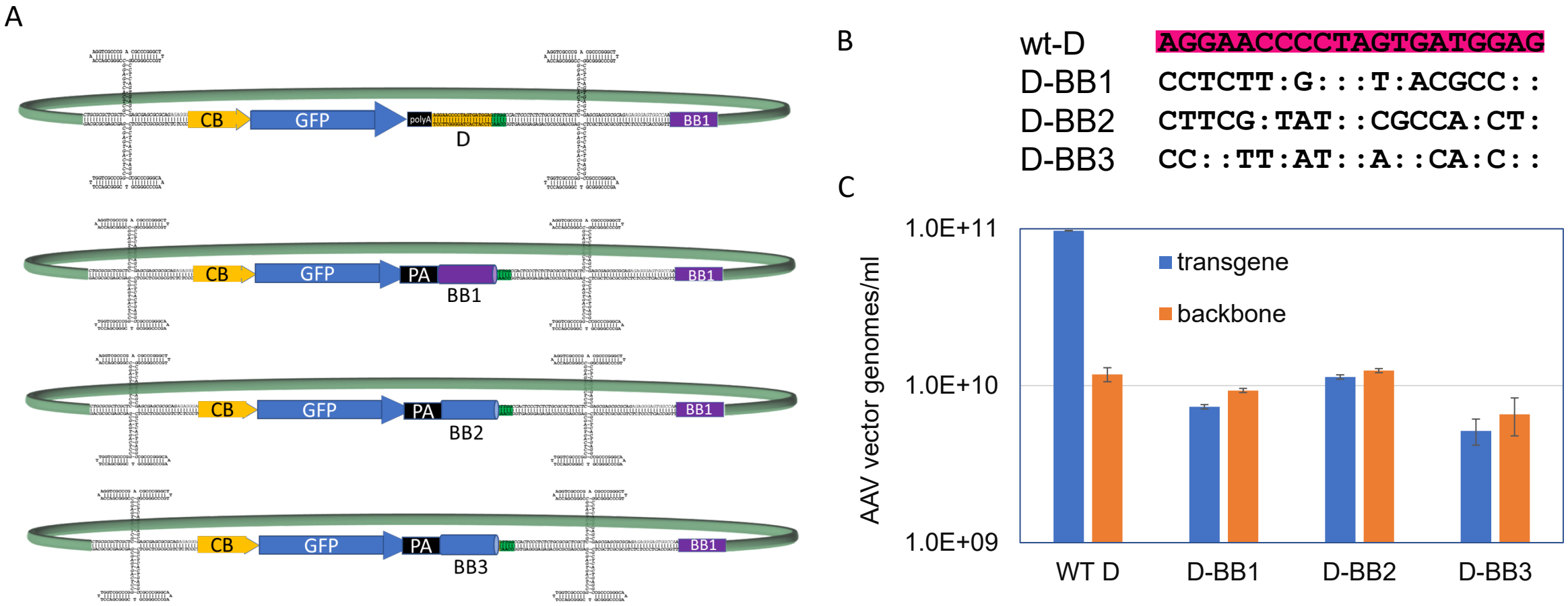
Supplementary Figure 2. Nucleotides adjacent to ITR in rAAV vector plasmid serve the functionality of D-element in rAAV packaging. A. Illustration of the source of plasmid backbone nucleotides tested in this study. The top construct is a self-complementary AAV vector with a wt D sequences. The lower three vector plasmids have the D-sequences replaced by the nucleotides next to ITR in the plasmid backbone moiety. BB1 was derived from pscAAV-CB-GFP, BB2 was derived from pssAAV-CB-lacZ, BB3 was derived from pssAAV-hFIX. B. Alignment of BB1, BB2 and BB3 to the wt-D element. C. The vectors plasmid shown in A. were used to produce rAAV vector. The results showed BB1, BB2 and BB3 element all supported rAAV production.

References:

1. Wang, Q.*, et al.* High-Density Recombinant Adeno-Associated Viral Particles are Competent Vectors for In Vivo Transduction. *Human gene therapy* **27**, 971-981 (2016).
2. Wang, Q.*, et al.* Efficient production of dual recombinant adeno-associated viral vectors for factor VIII delivery. *Human gene therapy methods* **25**, 261-268 (2014).



Supplementary Figure 1. A. Schematic illustration of D element library construction and screening diagram. The ITR related sequences are shown in the scAAV vector with D-element and terminal resolution region highlighted. The region randomized in the D08, D18 and D24 library are shown as N in the alignment with wild type ITR sequences. B. Illustration of the duplication of D sequence in scAAV replication. C. The summary of the percentage of recovered wild type AAV nucleotide in D sequences.



Supplementary Figure 2. Nucleotides adjacent to ITR in rAAV vector plasmid serve the functionality of D-element in rAAV packaging. A. Illustration of the source of plasmid backbone nucleotides tested in this study. The top construct is a self-complementary AAV vector with a wt D sequences. The lower three vector plasmids have the D-sequences replaced by the nucleotides next to ITR in the plasmid backbone moiety. BB1 was derived from pscAAV-CB-GFP, BB2 was derived from pssAAV-CB-lacZ, BB3 was derived from pssAAV-hFIX. B. Alignment of BB1, BB2 and BB3 to the wt-D element. C. The vectors plasmid shown in A. were used to produce rAAV vector. The results showed that BB1, BB2 and BB3 element all supported rAAV production.