Does size of AAV vector influence the transduction efficiency?

Jingyi Wang

March 14 2022

The Question

Does size of AAV vector influence the transduction efficiency?

Disclaimer: This project analyzes simulated data. The questions and hypotheses are real, but the results and conclusions are not.

Sub-Field of Biology:

Genetics

Rationale and Background:

Gene therapy is the process of altering genes inside the body to treat disease. Three main types of gene therapy are gene augmentation therapy, gene inhibition therapy, and the killing of specific cells. There are potential risks that might be targeting the wrong cells and infection caused by the virus. I think gene therapy is interesting because I am really interested in the field of genetics, and gene therapy is a relatively new topic. I heard about gene therapy's application to cancer but did not have the opportunity to do research to know more about it. It requires knowledge from different fields and the researchers are still determining when and how they should use it as a treatment. It still requires a high cost which made it limited for people. Therefore, more research and time are needed in order to extend the range of utilization of gene therapy on different diseases other than cancer. There is controversial whether or not people are allowed to apply gene therapy to basic traits. And people are concerned that gene modification will lead to intolerance for people with gene damage. These ethical concerns also require discussion. I am mostly interested about AAV vectors because it is the most commonly used vector platform. AAV stands for Adeno-associated virus. They do not cause disease but can cause a very mild immune response. They have properties which make them candidates for gene therapy.

Examples of relevant literature

Review article title:

Adeno-Associated Virus (AAV) as a Vector for Gene Therapy

URL: https://link.springer.com/article/10.1007/s40259-017-0234-5#citea

Abstract: There has been a resurgence in gene therapy efforts that is partly fueled by the identification and understanding of new gene delivery vectors. Adeno-associated virus (AAV) is a non-enveloped virus that can be engineered to deliver DNA to target cells, and has attracted a significant amount of attention in the field, especially in clinical-stage experimental therapeutic strategies. The ability to generate recombinant AAV particles lacking any viral genes and containing DNA sequences of interest for various therapeutic

applications has thus far proven to be one of the safest strategies for gene therapies. This review will provide an overview of some important factors to consider in the use of AAV as a vector for gene therapy.

Relevant original research article title:

Construction and analysis of compact muscle-specific promoters for AAV vectors

URL: https://www.nature.com/articles/gt2008104

Abstract: Adeno-associated viral (AAV) vectors have been broadly used for gene transfer in vivo for various applications. However, AAV precludes the use of most of the original large-sized tissue-specific promoters for expression of transgenes. Efforts are made to develop highly compact, active and yet tissue-specific promoters for use in AAV vectors. In this study, we further abbreviated the muscle creatine kinase (MCK) promoter by ligating a double or triple tandem of MCK enhancer (206-bp) to its 87-bp basal promoter, generating the dMCK (509-bp) and tMCK (720-bp) promoters. The dMCK promoter is shorter but stronger than some previously developed MCK-based promoters such as the enh358MCK (584-bp) and CK6 (589-bp) in vitro in C2C12 myotubes and in vivo in skeletal muscles. The tMCK promoter is the strongest that we tested here, more active than the promiscuous cytomegalovirus (CMV) promoter. Furthermore, both the dMCK and tMCK promoters are essentially inactive in nonmuscle cell lines as well as in the mouse liver (>200-fold weaker than the CMV promoter). The dMCK promoter was further tested in a few lines of transgenic mice. Expression of LacZ or minidystrophin gene was detected in skeletal muscles throughout the body, but was weak in the diaphragm, and undetectable in the heart and other tissues. Similar to other miniature MCK promoters, the dMCK promoter also shows preference for fast-twitch myofibers. As a result, we further examined a short, synthetic muscle promoter C5-12 (312-bp). It is active in both skeletal and cardiac muscles but lacks apparent preference on myofiber types. Combination of a MCK enhancer to promoter C5-12 has increased its strength in muscle by two- to threefold. The above-mentioned compact muscle-specific promoters are well suited for AAV vectors in muscle-directed gene therapy studies.

Hypotheses

The Biological Hypothesis:

AAV vector is efficient due to its small size.

The Biological Prediction:

The size of AAV vector is negatively correlated with transduction efficiency.

A Statistical Alternative Hypothesis:

The size of AAV vector is positively correlated with transduction efficiency.

A Statistical Null Hypothesis:

The size of AAV vecto does not correlate with transduction efficiency.

Experimental Design

Sampling Design:

The overall experiment work on determining if structural difference between each type of AAV vectors gave uniquness regarding the process of genetherapy. The independent variable in the experiment is the size of

AAV vectors in kb. The dependent variable is infection time in hours. Other factors that might influence the speed will be controled partially like temperature and kinds of vectors. However, there will always be some factors that might influence the result of the experiment like the life cycles that cannot be controlled because it is directly connected with the sizes of AAV vectors. Therefore, correlation study is chosen. Moreover, the labotary environemnt is in vivo, since the transduction efficiency of AAV will be measured on living organism like mice.The overall structure of the experiment does not include experimentation on human. Although the experiment was done on mice, the process of the experiment should be designed to eliminate maximum harm to mice. However, in general, if there is test of efficiency of AAV vector on human. Then there are numbers of ethical considerations. Firstly, the consent from patients and their families should be given prior to experiment. Determing if a patient should use gene therapy should be made.

Explanatory and Response Variables:

The explanatory variable is the size of the AAV vector and response variable is the speed for infection.

Sample size:

sample size is 9 including AAV vectors from 1 to 9.According to the research article that I found for my research "Transduction optimization of AAV vectors for human gene therapy of glaucoma and their reversed cell entry characteristics" which research about the transduction of 7 AAV vectors, including scAAV1,2,2.5,5,6,8,9. It is better this experiment to include all kinds of aav vectors to maximize the sample sizes and power of the test. Therefore the sample size is 9.

Alpha:

The prediction of the experiment is that the with the increase of the size of AAV vectors, the speed of infection decrease. Therefore the predicted r (correlation coefficient) is -0.9. The value for alpha for this experiment should be around 0.05.

Data Analysis Plan

Correlation

I chose correlation instead of regression because correlation examines the relationship between two varibales while regression examine if one variable affects another. In the experiment, while the relationship of sizes and transduction efficiency is explored, there are other factors that might influence the transduction efficiency that cannot be controlled. The factors are connected with the sizes of AAV vectors. Therefore, correlation is chosen.

Assumptions and Exploratory Data Analysis (EDA)

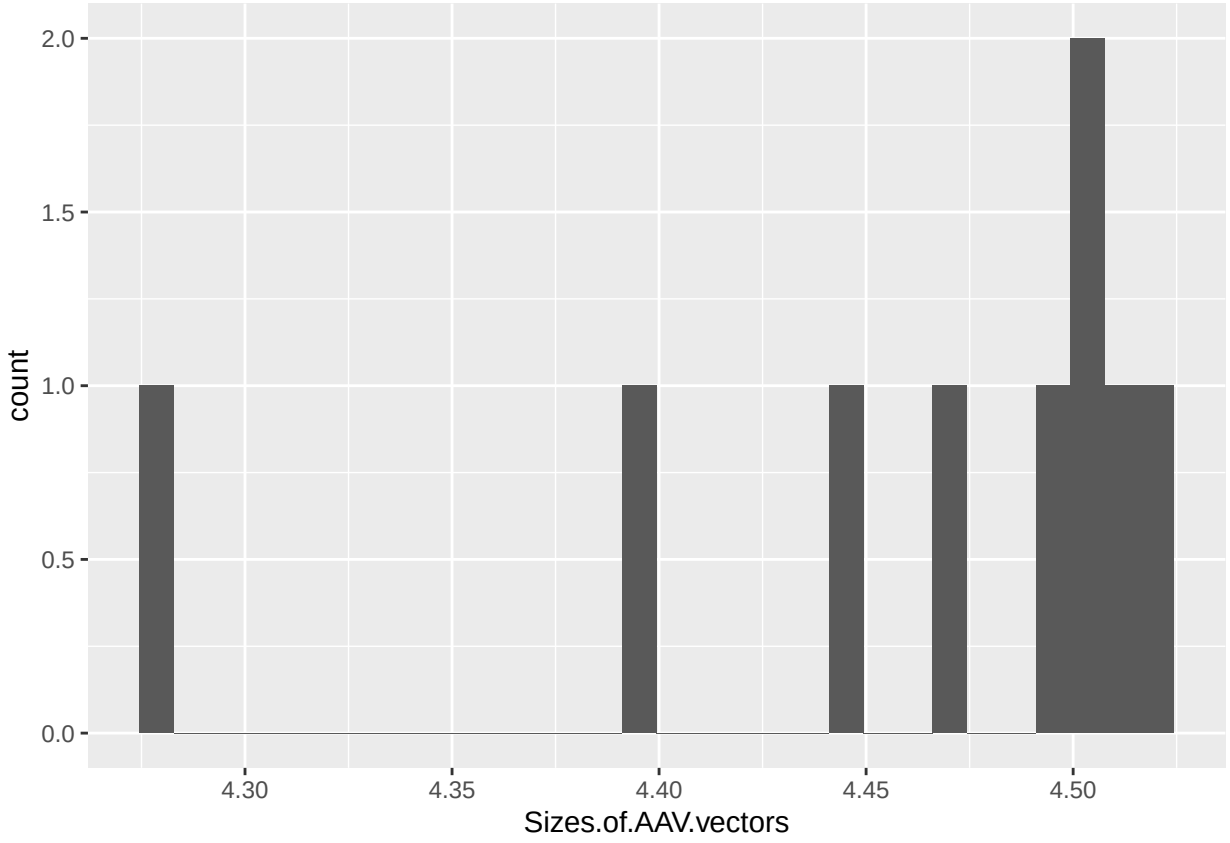
The data are normally distributed. The variables are independent.

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.3     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
SpeedData<-read.csv(file="speed.csv")
head(SpeedData)
```

```
##   Sizes.of.AAV.vectors time.for.infection
## 1             4.472050           150.0755
## 2             4.514326           100.3242
## 3             4.503529           128.4917
## 4             4.518446           129.4309
## 5             4.506560           135.9029
## 6             4.398183           167.2312
```

```r
#USE THIS BLOCK TO INPUT NECESSARY CODE.
Histogram1<-ggplot(data=SpeedData, mapping=aes(x=Sizes.of.AAV.vectors))+geom_histogram(bins=30)
Histogram1
```

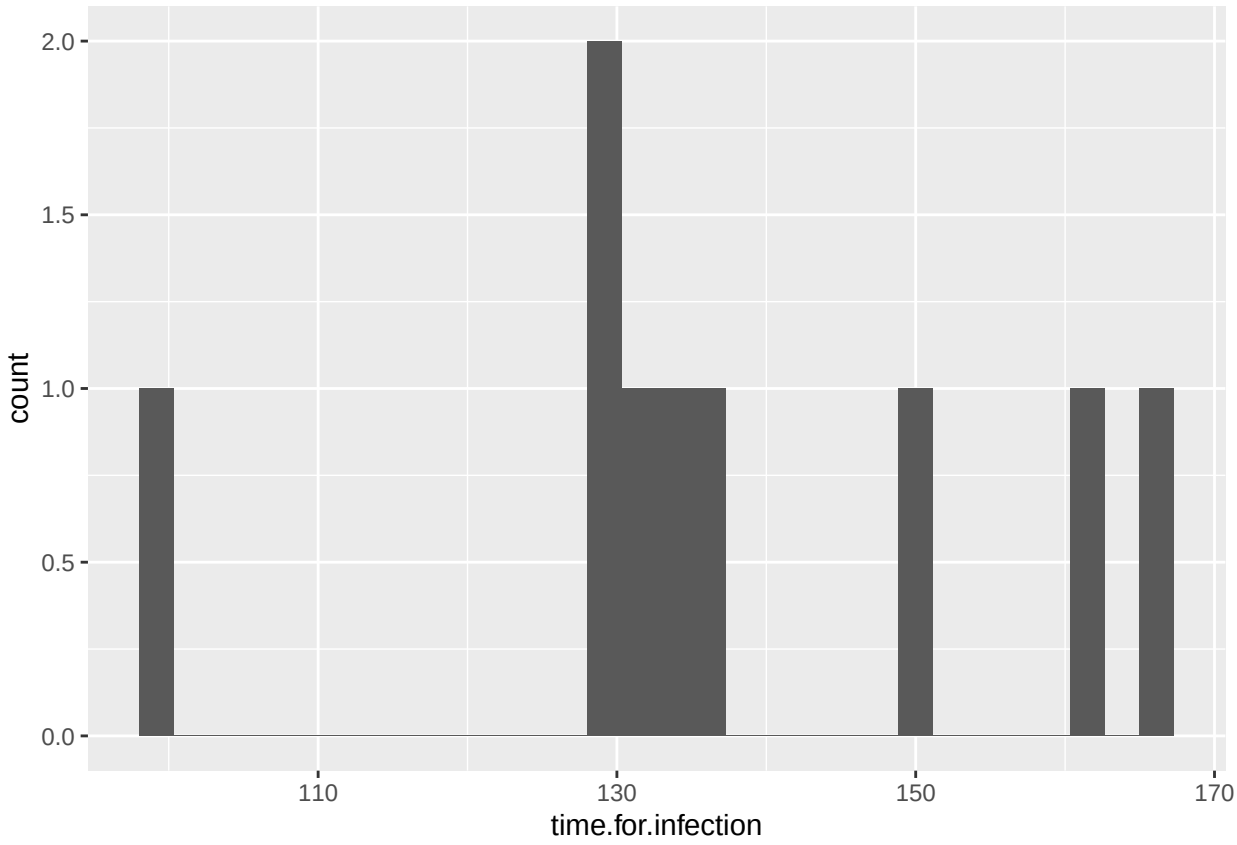



```r
#Test for normality of the second language scores in this code box.

ks.test(x=SpeedData$Sizes.of.AAV.vectors,y="pnorm",mean(SpeedData$Sizes.of.AAV.vectors),sd(SpeedData$Si
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  SpeedData$Sizes.of.AAV.vectors
## D = 0.24697, p-value = 0.5621
## alternative hypothesis: two-sided
```

```r
#visualize the gray matter density data here.
Histogram2<-ggplot(data=SpeedData, mapping=aes(x=time.for.infection))+geom_histogram(bins=30)
Histogram2
```




```r
#Test for normality of the second language scores in this code box.

ks.test(x=SpeedData$time.for.infection,y="pnorm",mean(SpeedData$time.for.infection),sd(SpeedData$time.f
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  SpeedData$time.for.infection
## D = 0.21203, p-value = 0.7389
## alternative hypothesis: two-sided
```

Interpretation of EDA:

If the statsitical significance level is 0.05, p-value is equal to 0.5621 and 0.7389 respectively. They are greater than 0.05. Therefore, we cannot reject null hypothesis which is that the sizes of AAV vector data and speed of infection are normally distributed. We can conclude that the sizes of AAV vector data and speed of infection are normally distributed at 0.05. Histograms do not look very normal distributed, but it is because its sample size which is equal to 9 which is pretty small.

Primary Statistical Analysis

Primary Statistical Analysis is the correlation test.

```
#USE THIS BLOCK TO INPUT NECESSARY CODE.

cor.test(SpeedData$Sizes.of.AAV.vectors,SpeedData$time.for.infection)
```

```
##
##  Pearson's product-moment correlation
##
## data:  SpeedData$Sizes.of.AAV.vectors and SpeedData$time.for.infection
## t = -2.7326, df = 7, p-value = 0.02923
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.9359759 -0.1038735
## sample estimates:
##       cor
## -0.7184343
```

Summary Statistics

```
#USE THIS BLOCK TO INPUT NECESSARY CODE.
note<-"From the test above, it can be concluded that the t value is equal to -2.7326. p-value is equal
print(note)
```

```
## [1] "From the test above, it can be concluded that the t value is equal to -2.7326. p-value is equal
```

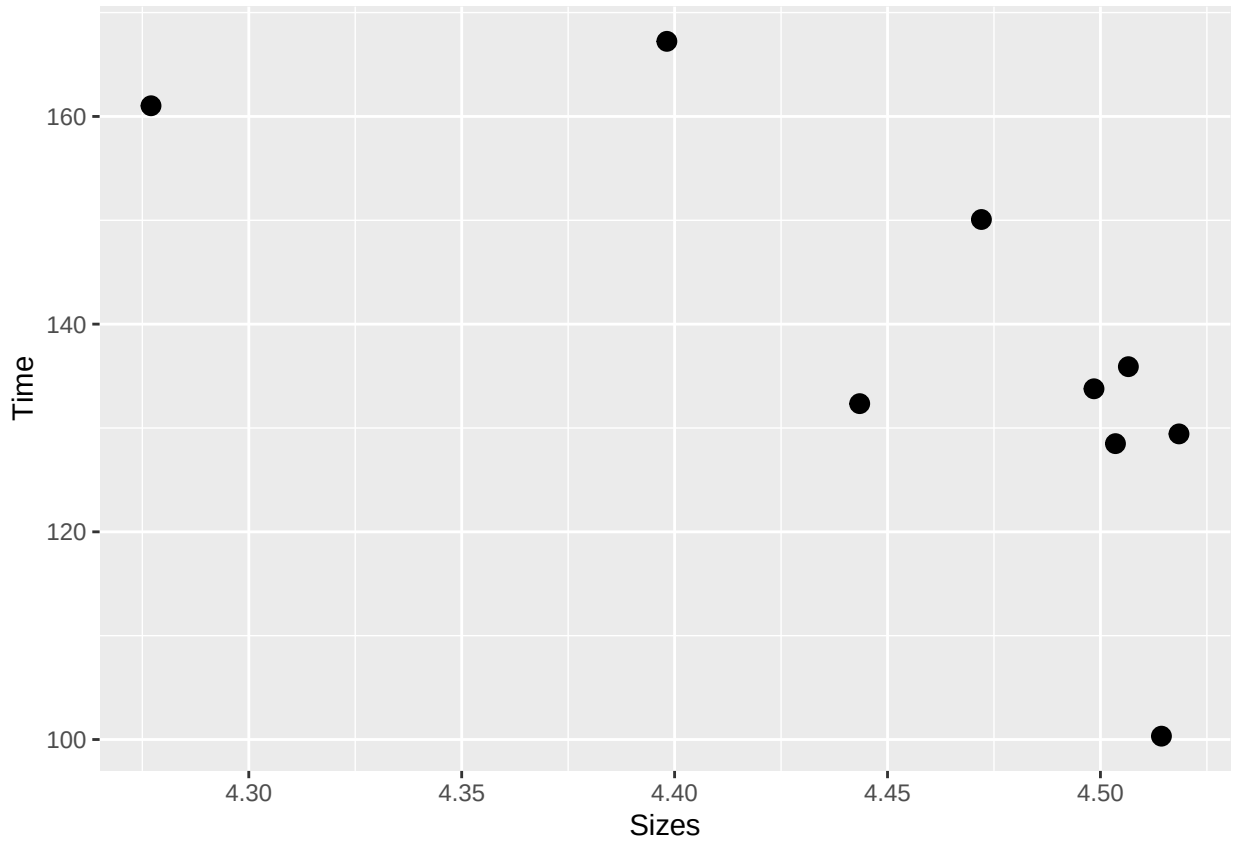
Data Visualization

scatterplot

```
#USE THIS BLOCK TO INPUT NECESSARY CODE.
scatterplot<-ggplot(data=SpeedData,aes(x=SpeedData$Sizes.of.AAV.vectors,y=SpeedData$time.for.infection)
scatterplot
```

```
## Warning: Use of `SpeedData$Sizes.of.AAV.vectors` is discouraged. Use
## `Sizes.of.AAV.vectors` instead.
```

```
## Warning: Use of `SpeedData$time.for.infection` is discouraged. Use
## `time.for.infection` instead.
```



Conclusions

With the statistical significance level is 0.05, the p-value of the estimated corrleation coefficient is 0.02923, which is less than 0.05. We can reject the null hypothesis that the correlation between two data is 0. We conclude that there is statistically significant correlation between the two data at significance level of 0.05. Since the estimated correlation coefficient is -0.7184 and the 95% CI is less than 0. There is a negative correlation between two data. Therefore, it can be conluded that with the increase of the size of AAV vectors, the speed of infection decreases.

Future Directions

For future directions, different factors' effect on the efficiency of transduction can be determined.

Citations

["Adeno-Associated Virus." Wikipedia, Wikimedia Foundation, 25 Jan. 2022, https://en.wikipedia.org/wiki/Adeno-associated_virus#Structure. "Gene Therapy." Mayo Clinic, Mayo Foundation for Medical Education and Research, 29 Dec. 2017, https://www.mayoclinic.org/tests-procedures/gene-therapy/about/pac-20384619. Wang, B., Li, J., Fu, F. H., Chen, C., Zhu, X., Zhou, L., Jiang, X., & Xiao, X. (2008, June 19). Construction

and analysis of compact muscle-specific promoters for Aav Vectors. Nature News. Retrieved March 14, 2022, from https://www.nature.com/articles/gt2008104 Naso, M. F., Tomkowicz, B., Perry, W. L., & Strohl, W. R. (2017, July 1). Adeno-associated virus (AAV) as a vector for gene therapy - biodrugs. SpringerLink. Retrieved March 14, 2022, from https://link.springer.com/article/10.1007/s40259-017-0234-5#citea]