



BME 335 Final Project

By Alan Fletcher, Kishan Bhakta, & Nabihah
Momin



Context

- CRISPR - Cas9 is one of the more well known approaches of editing the genome using CRISPR
- One of the most accurate, efficient, fast, and cheap ways to modify DNA
- Based on a naturally occurring system in which bacteria store a segment of viral DNA to disable the virus
- Goal is to prevent and treat human diseases
- SpCas9-BME335 was created to be better than the standard wild type SpCas9 method



Main Hypotheses

H_0 : The SpCas9-BME335 variant will NOT improve specific on-target DNA contacts when compared to the wild type SpCas9

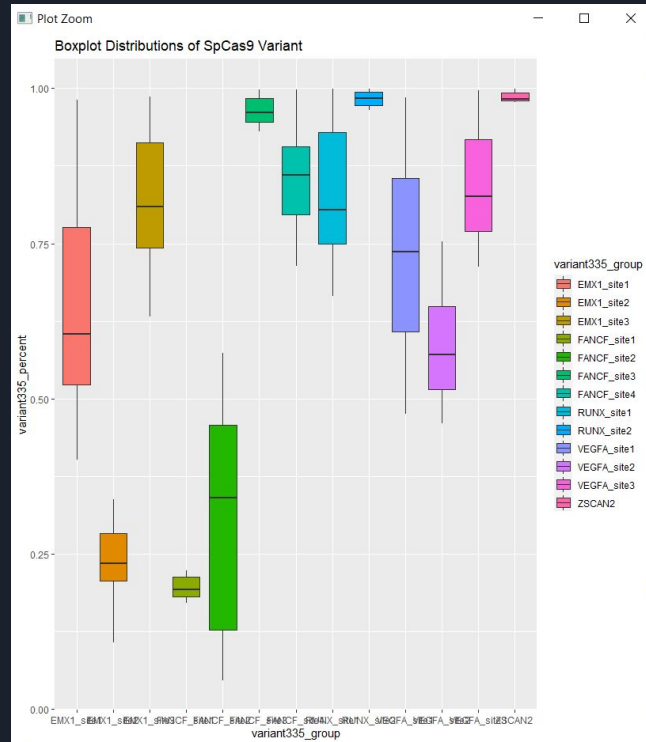
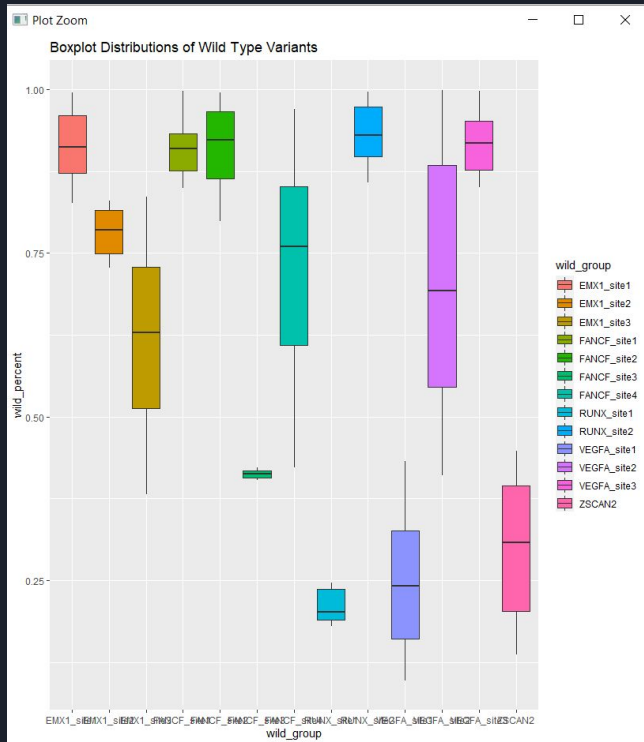
H_A : The SpCas9-BME335 variant WILL improve specific on-target DNA contacts when compared to the wild type SpCas9



Study 1 - ANOVA Test

- Background
 - Percent Modified Measurements of Functionality With On-Target Sites for SpCas9-BME335 and wild-type SpCas9
- Methods
 - Study Design
 - Observational
 - Study Population
 - 13 Different sgRNAs Targeted Sites in the Endogenous Human Genes (EMX1 Site1-3, FANCF Site 1-4, RUNX Site 1-2, ZSCAN 2, VEGFA Site 1-3)
 - Study Sample
 - On-target events observed for sgRNAs at each site for both variants
- Variables
 - Explanatory
 - Nuclease (SpCas9-BME335 or the wild type variant SpCas9) and sgRNA site Respectively
 - Response
 - Percent Modified Measurements of On-Target Sites for Each Nuclease

Study 1-Preliminary Visualization





Study 1: Hypotheses

- Null Hypothesis (H_0)
 - The Mean Percent Modified Measurements of On-Target Sites for Each Nuclease is the Same
- Alternative Hypothesis (H_A)
 - The Mean Percent Modified Measurements of On-Target Sites for Each Nuclease is the Significantly Different

Study 1 - ANOVA Test Results

```
> SpCas9AOV <- aov(variant335_group ~ variant335_percent, data = my_data_variant335)
> summary(SpCas9AOV)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
variant335_percent	1	1229	1229.0	108.3	<2e-16 ***
Residuals	453	5141	11.3		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> SpCas9AOV_wildType <- aov(wild_group ~ wild_percent, data = my_data_wild)
> summary(SpCas9AOV_wildType)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wild_percent	1	466	466.2	35.93	4.35e-09 ***
Residuals	427	5540	13.0		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Study 2 - Sign Test

Background: The purpose of Study 2 was to test if the variant, SpCas9-BME335, significantly reduces the number of off-target events in cells at each site when compared to the wild-type SpCas9 variant

Methods:

- Study Design: Experimental
- Study Population: Eight different sgRNAs targeted sites in the endogenous human genes (EMX1 Site1, EMX1 Site 2, FANCF Site 1, FANCF Site 2, FANCF Site 3, FANCF Site 4, RUNX1 Site1, and ZSCAN2)
- Study Sample: Off-target events that was observed for these sgRNAs at each site for the SpCas9-BME335 and the wild type variant SpCas9

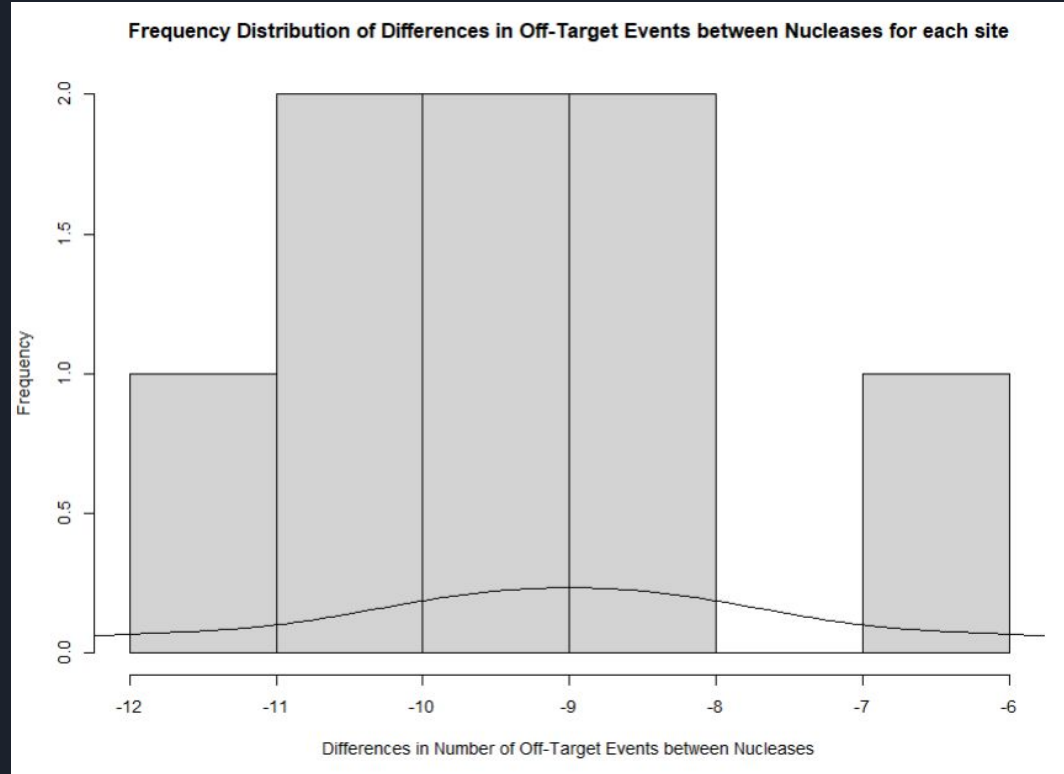
Variables:

- Explanatory: Nuclease (SpCas9-BME335 or the wild type variant SpCas9) and sgRNA site
- Response: Number of off-target events

Study 2 - Preliminary Data

Violates the assumptions of normality needed to conduct a paired t-test

Differences in the Number of Off-Target Events were calculated by taking the difference between the SpCas9-BME335 and the wild type variant SpCas9





Study 2 - Sign Test Hypotheses and Descriptive Statistics

H_0 : The median value of the difference between the SpCas9-BME335 and the wild type variant SpCas9 is equal to ZERO

H_A : The median value of the difference between the SpCas9-BME335 and the wild type variant SpCas9 is NOT equal to zero

<u>Descriptive Statistics for Study 2</u>		
	SpCas9-BME335	Wild Type SpCas9
Mean	1.25	10.25
Median	1.5	10.5
Standard Deviation	0.89	2.38
Variation	0.79	5.64



Study 2 - Results and Conclusion

Since the p value is less than the significance value of 0.05 we can reject the null hypothesis. This is further supported by the fact that the median values for the SpCas9-BME335 variant is way lower than the corresponding wild type SpCas9. Therefore, we can conclude SpCas9-BME335 significantly reduces the number of off-target events in cells at each site

```
> sort(diff)
[1] -12 -10 -10 -9 -9 -8 -8 -6
> hypothesized_median <- 0
>
> min_val <- min(sum(diff = hypothesized_median), sum(diff != hypothesized_median))
>
> binom.test(min_val,length(diff),p=0.5, alternative = "two.sided")

Exact binomial test

data:  min_val and length(diff)
number of successes = 0, number of trials = 8, p-value = 0.007812
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.0000000 0.3694166
sample estimates:
probability of success
0
```



Study 4: Set up

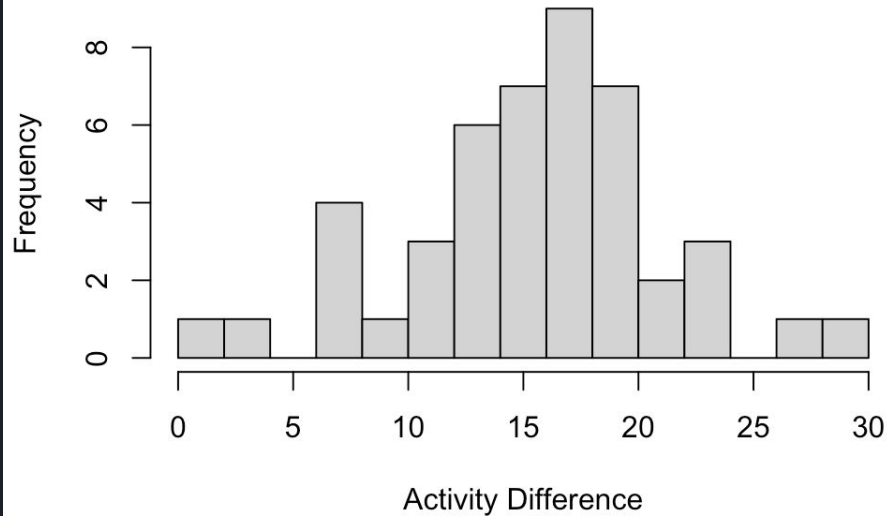
- Compares the off target activity of the newly modified SpCas9-BME335-2 and the wild-type SpCas9 from a previous study.
- Study Design: Experimental
- Study Population: Sites altered by SpCas9-BME335-2 and wild type SpCas9
- Study Sample: Off target activity observed on each sgRNA site for both the BME335-2 variant and the wild type method

Variables:

- Explanatory: Nuclease (SpCas9-BME335-2 or wild type SpCas9) and sgRNA site
- Response: Off target activity

Study 4: Preliminary Visualization

Activity Difference Distribution



Descriptive Statistics

diff_mean	15.2839239446674
diff_standard_de...	5.54633102024879
median	15.8701608714982
n	46
SE	0.817761895108068



Study 4: Paired Test

- A paired test was then performed on the data
- H_0 : The means of the SpCas9-BME335-2 and the wild-type SpCas9 methods are the same.
- H_A : There is a significant difference between the means of the SpCas9-BME335-2 and the wild-type SpCas9 methods.

Results:

$$t = 16.884, p\text{-value} < 2.2e-16$$



Study 4: Conclusion

- Since the p value < 0.05 , we reject the null hypothesis and conclude that the means of the two groups is significantly different
- Based on the data collected, we can also conclude that the SpCas9-BME335-2 had higher activity than the wild-type SpCas9
- That means that the wild-type SpCas9 protein had less off target activity



Main Conclusion

- Study 1: The mean off target activity for the variant is the same as the off target activity of the wild type. — Not an improvement
- Study 2: The median value of off target events for the variant were lower than the median of the off target events of the wild type. — An improvement
- Study 4: The off target activity of the variant is larger than the off target activity of the wild type. — Not an improvement

Based on results from the statistical tests, we FAIL reject our main null hypothesis. The BME335 variant did not improve specific on target contacts.