# Integration of Epigenetics Data Into CRISPR Off-Target Energy Model

Freja Kelleris
RTH, University of Copenhagen
35th TBI Winter Seminar, Bled

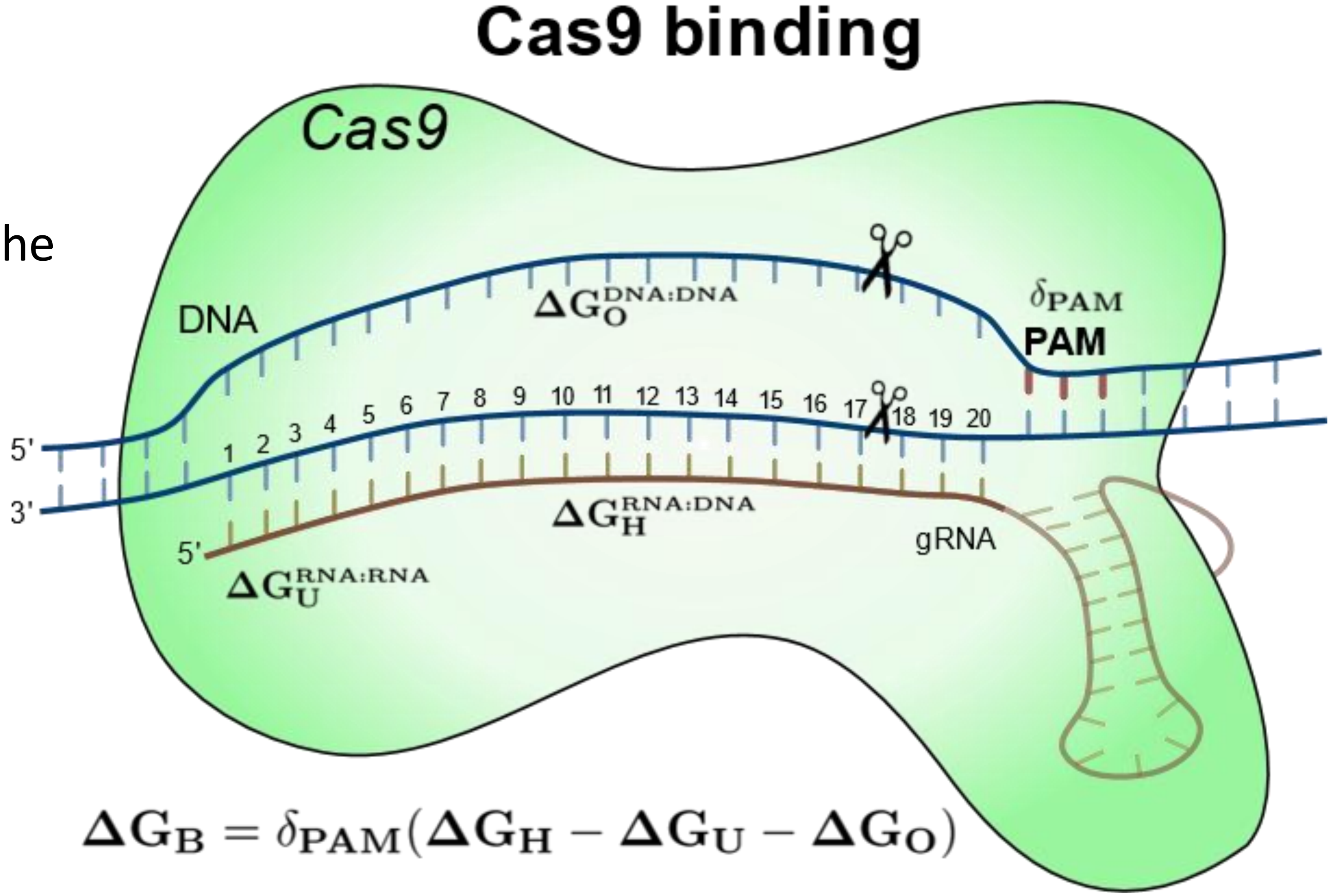
Feb 10, 2020

# Project Motivation

- CRISPR technology paving the way in biology, medicine, and biotechnology
- Being able to better predict the risks of potential off-targets
- Improving CRISPR technology and making it application safer

# CRISPR-Cas9 Energy model

CRISPRoff score:
The estimated free-energy
Contribution from binding to the
off-target site



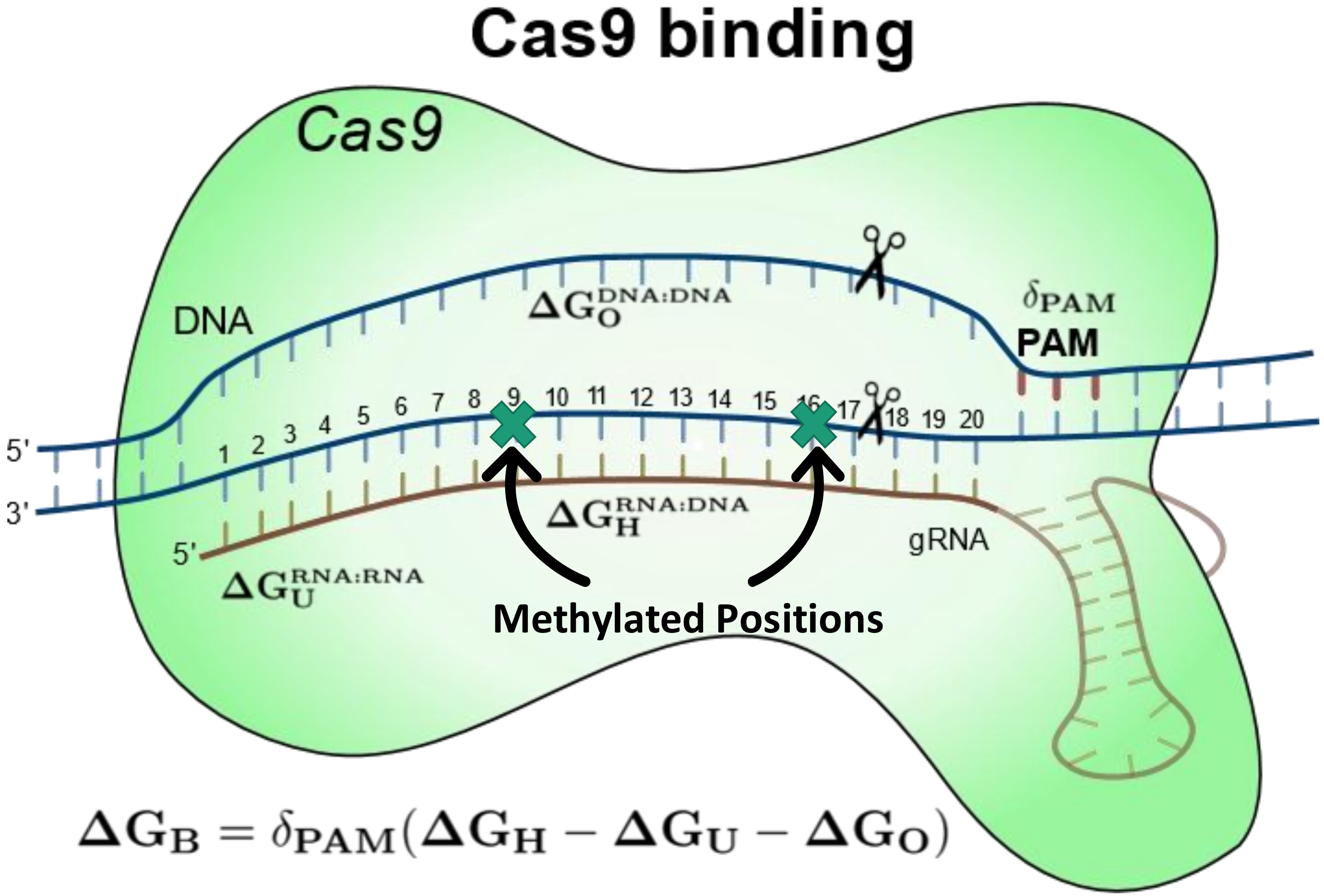
Cas9 binding

$$\Delta G_{B} = \delta_{PAM}(\Delta G_{H} - \Delta G_{U} - \Delta G_{O})$$

Alkan, Ferhat, et al. *Genome biology* 19.1 (2018)

# CRISPR-Cas9 Energy model

- Methylations at or near such positions might siginficantly impact the binding at that given site.



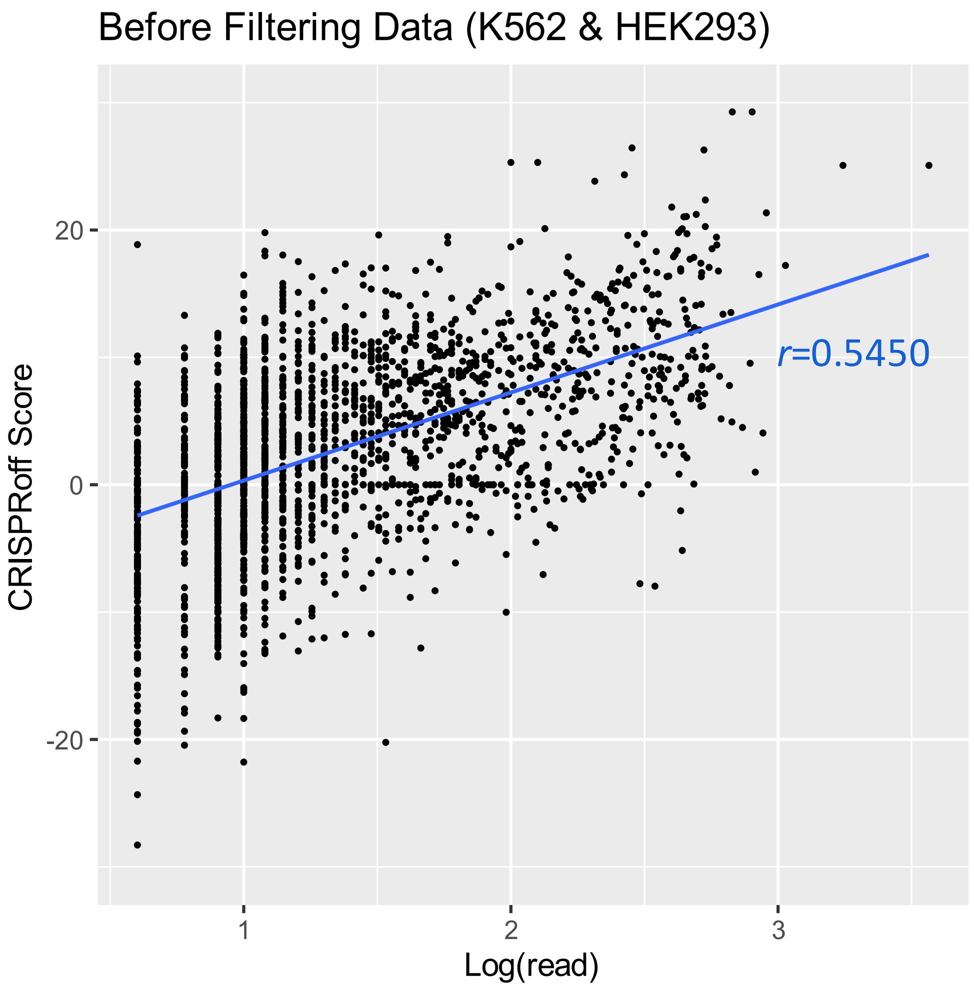
$$\Delta G_{B} = \delta_{PAM}(\Delta G_{H} - \Delta G_{U} - \Delta G_{O})$$

Alkan, Ferhat, et al. *Genome biology* 19.1 (2018)

# Testing Naive Off-target Correction Models

- How Can we improve the correlation between the CRISPR-off Target score and the off-target cleavage efficiency?
- Try to incorporate Epigenetics Data:
  - DNA Methylation Data
  - Histone Modification Data

- Log(read): measured off-target activity
- Using CIRCLE-Seq[1] data as basis



1. *Tsai et al.* *Nature Methods* volume 14, 607–614(2017)

# Incorporation of DNA Methylation Data

- ENCODE project and is available as tracks in the UCSC browser[1]
- Reduced Representation Bisulfite Sequencing (RRBS)[2]
- Multiple Cell lines are available
  - Includes K562 and HEK293

*1.ENCODE Project Consortium. (2011). A user's guide to the encyclopedia of DNA elements(ENCODE). PLoS biology,9(4), e1001046.*

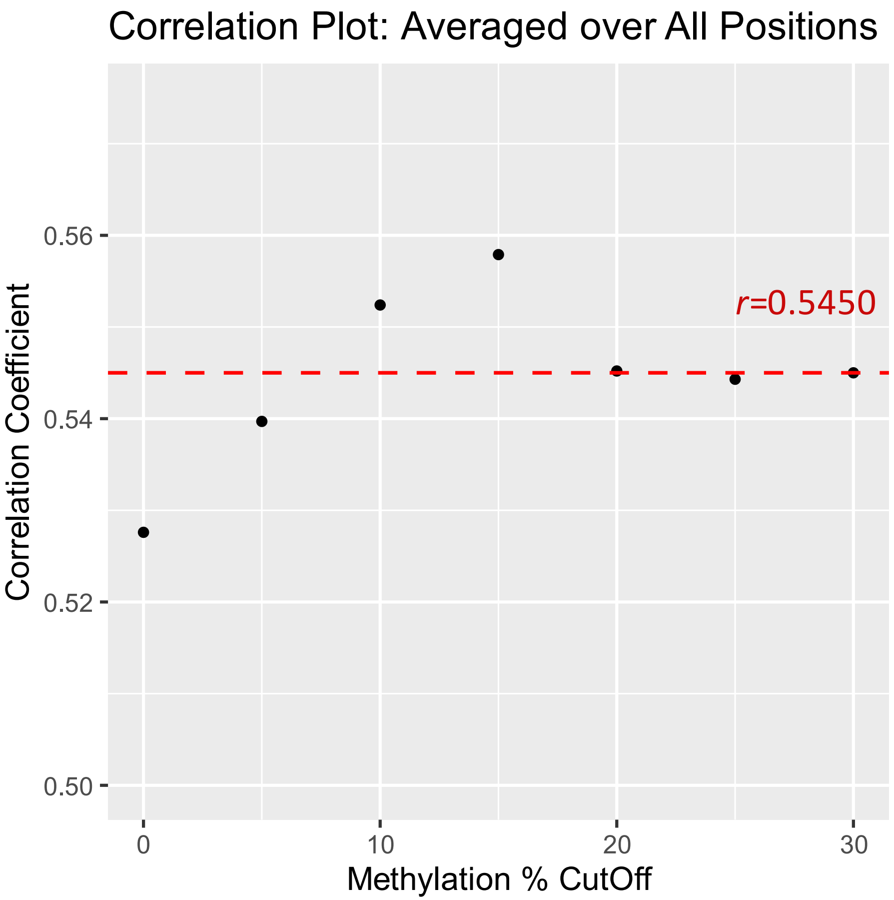
*2. Meissner et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008 ;454(7205):766-70.*

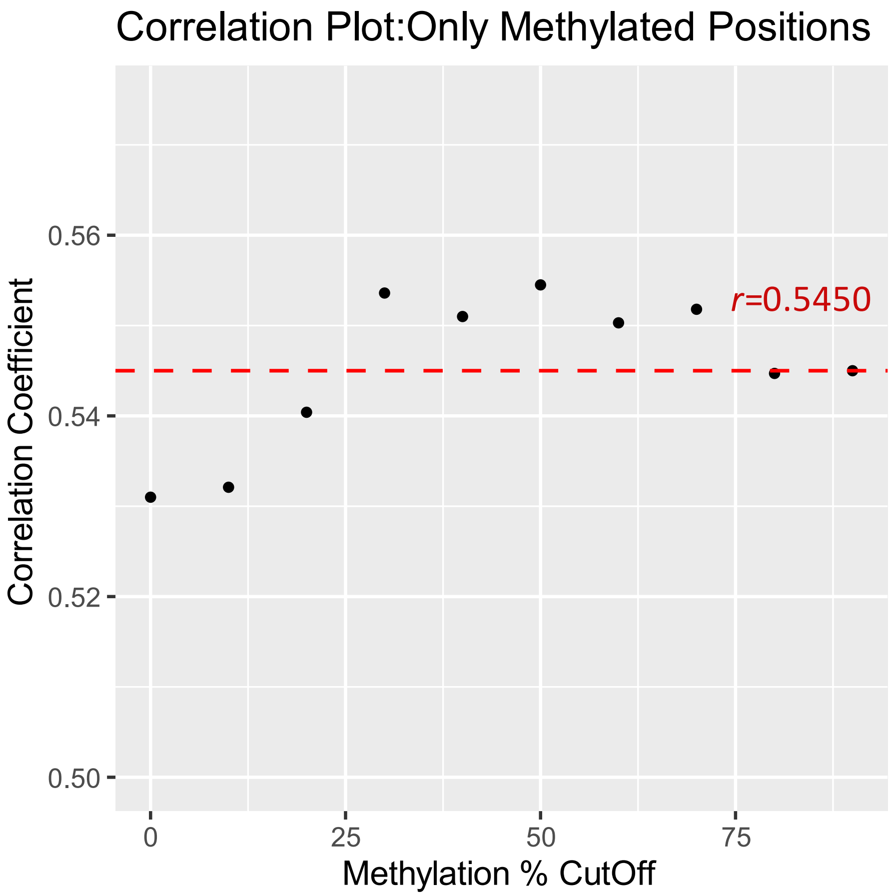
# DNA Methylation Data - Analysis setup

Three types of approaches was attempted:

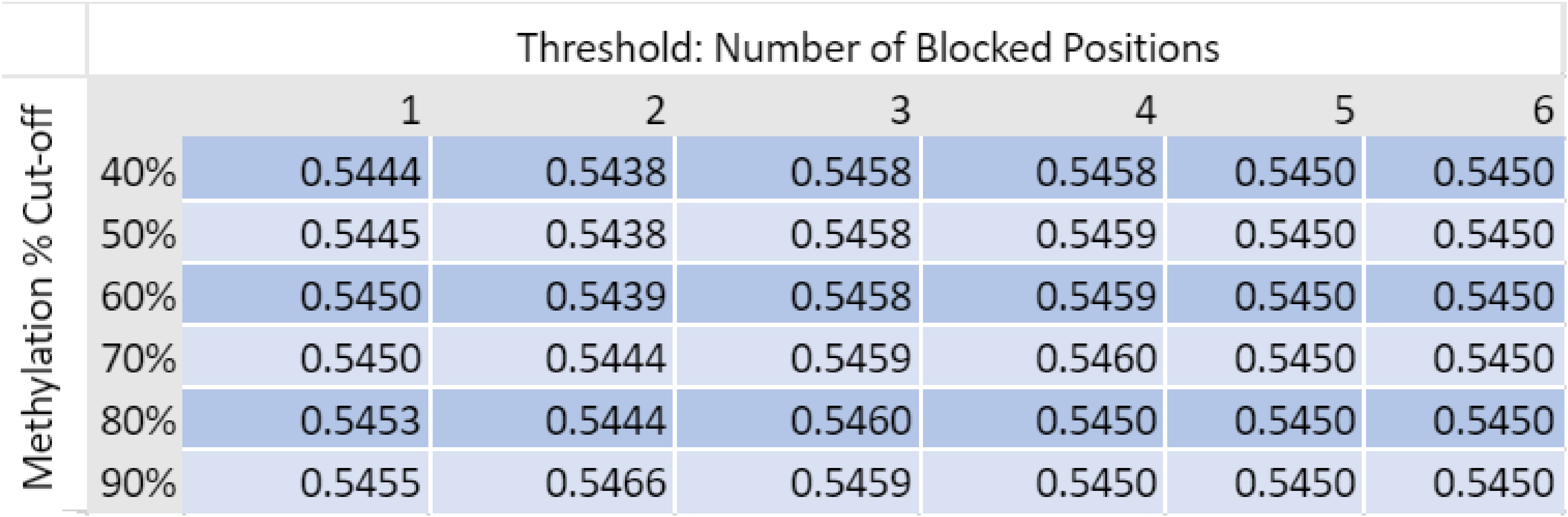
1. Average methylation signal percentage over all positions .
2. Average methylation signal percentage over only positions that have a reported methylation signal.
3. Number of blocked positions.
    - Positions blocked when the associated methylation signal excess a given threshold.

# DNA Methylation Data - Results



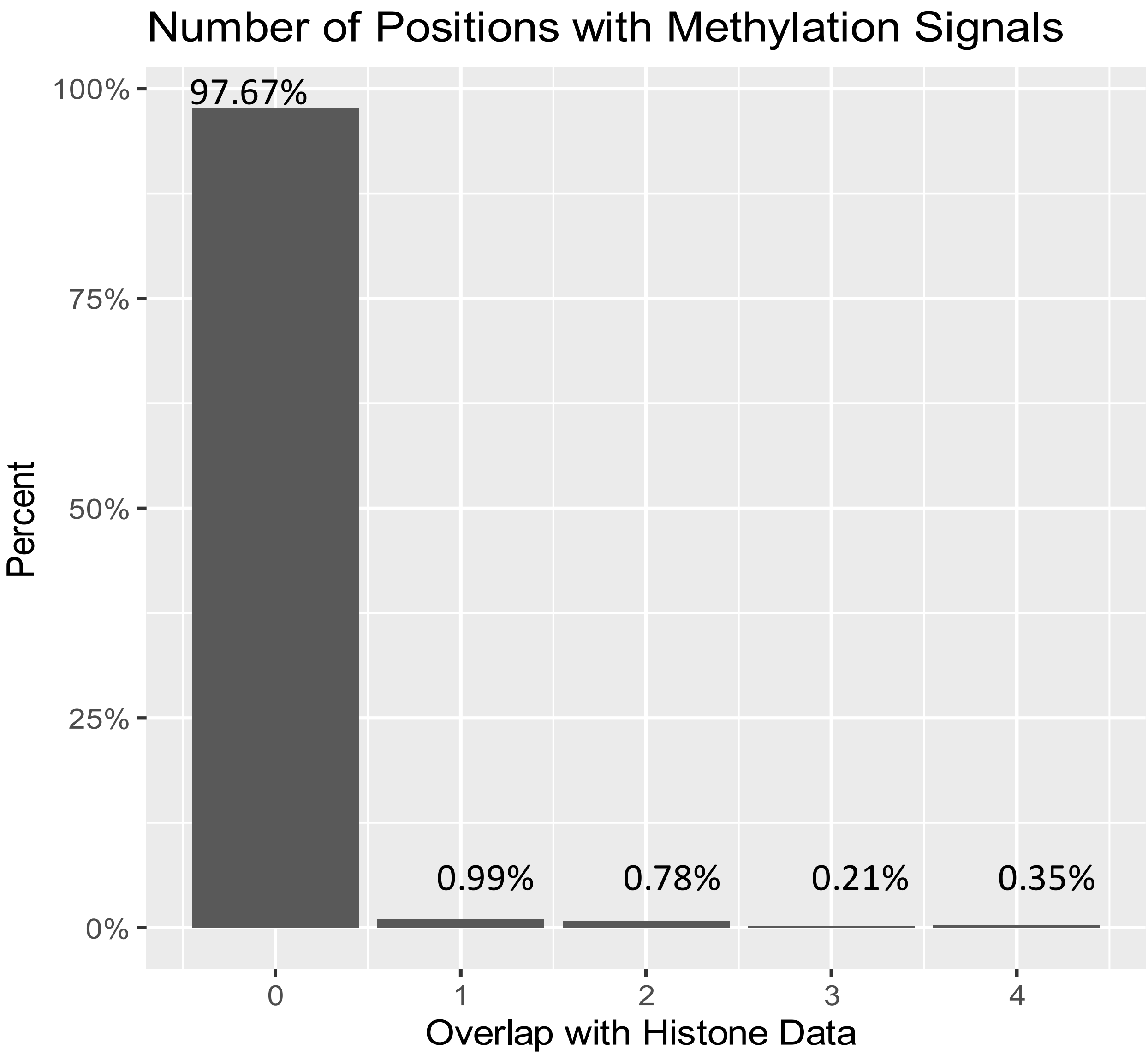


# DNA Methylation Data- Results

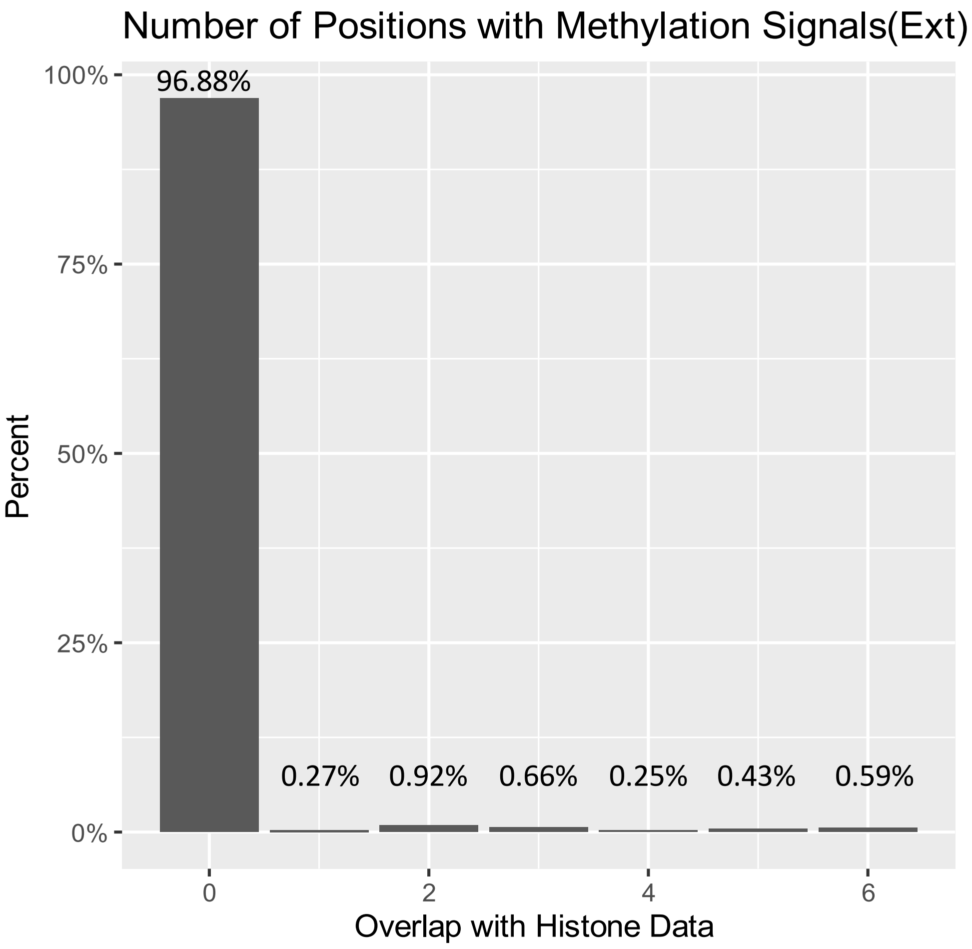


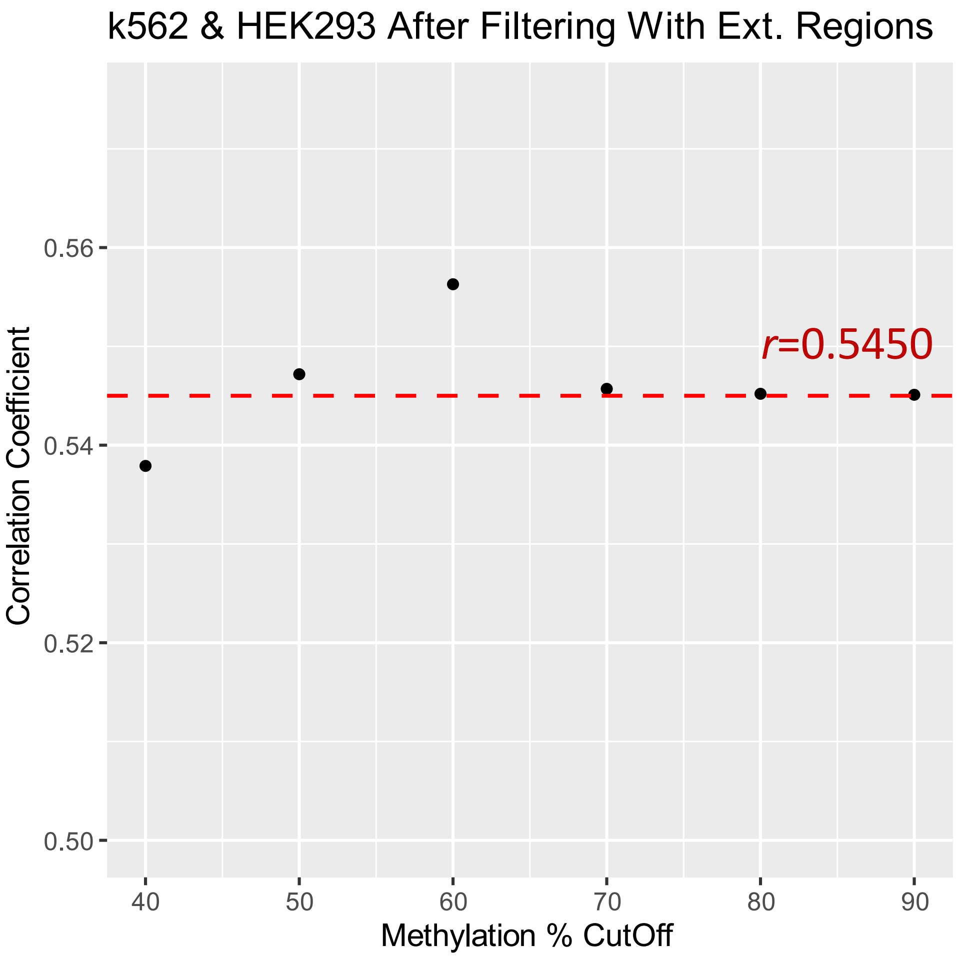
| | | Threshold: Number of Blocked Positions | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| Methylation % Cut-off | 40% | 0.5444 | 0.5438 | 0.5458 | 0.5458 | 0.5450 | 0.5450 |
| | 50% | 0.5445 | 0.5438 | 0.5458 | 0.5459 | 0.5450 | 0.5450 |
| | 60% | 0.5450 | 0.5439 | 0.5458 | 0.5459 | 0.5450 | 0.5450 |
| | 70% | 0.5450 | 0.5444 | 0.5459 | 0.5460 | 0.5450 | 0.5450 |
| | 80% | 0.5453 | 0.5444 | 0.5460 | 0.5450 | 0.5450 | 0.5450 |
| | 90% | 0.5455 | 0.5466 | 0.5459 | 0.5450 | 0.5450 | 0.5450 |

# DNA Methylation – Overlap issue



# DNA Methylation Data-Extending Genomic Ranges



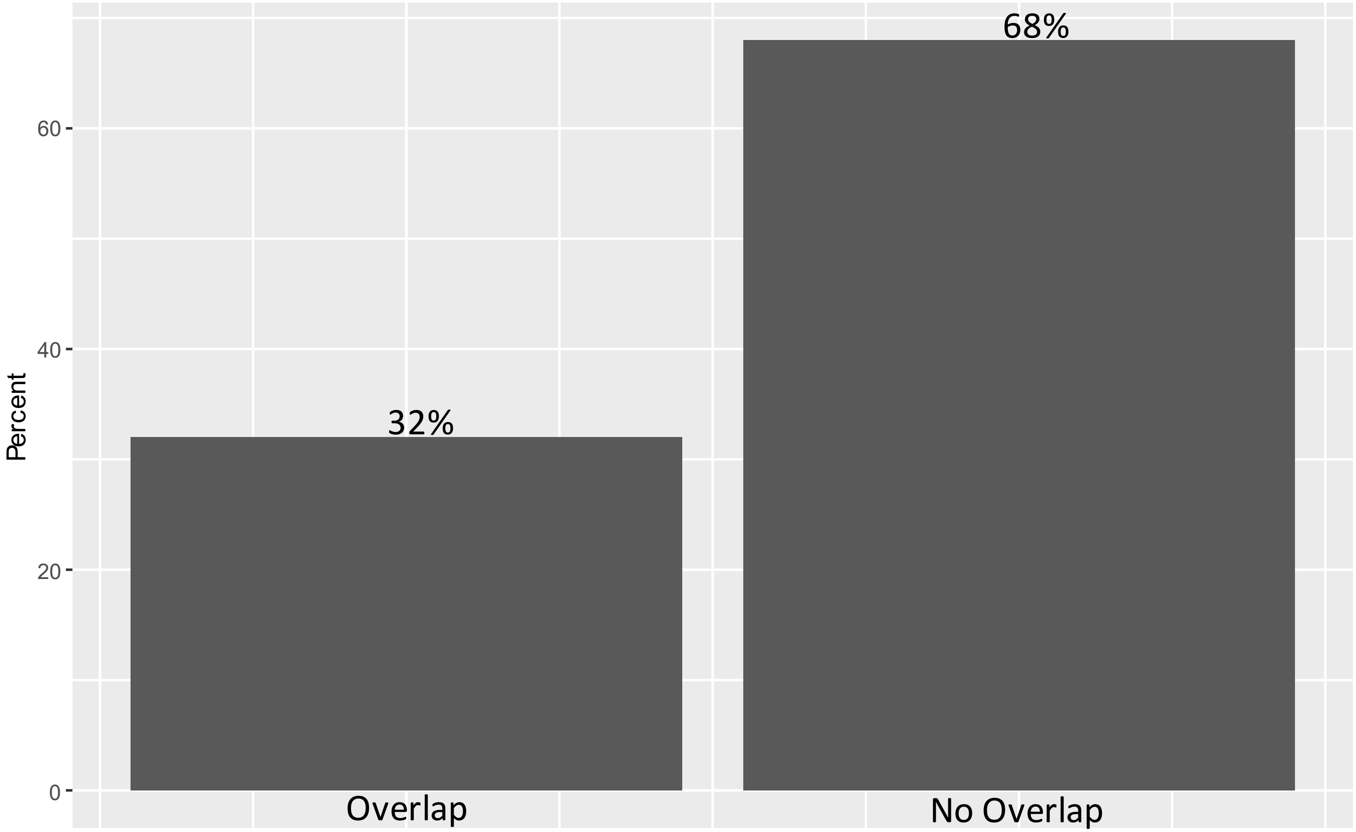


# Incorporation of Histone Modification Data

**Data:** Histone Modifications by ChIP-seq from ENCODE/Broad Institute[1,2]

**Antibody:** H3k27me3

**Cell line:** K562

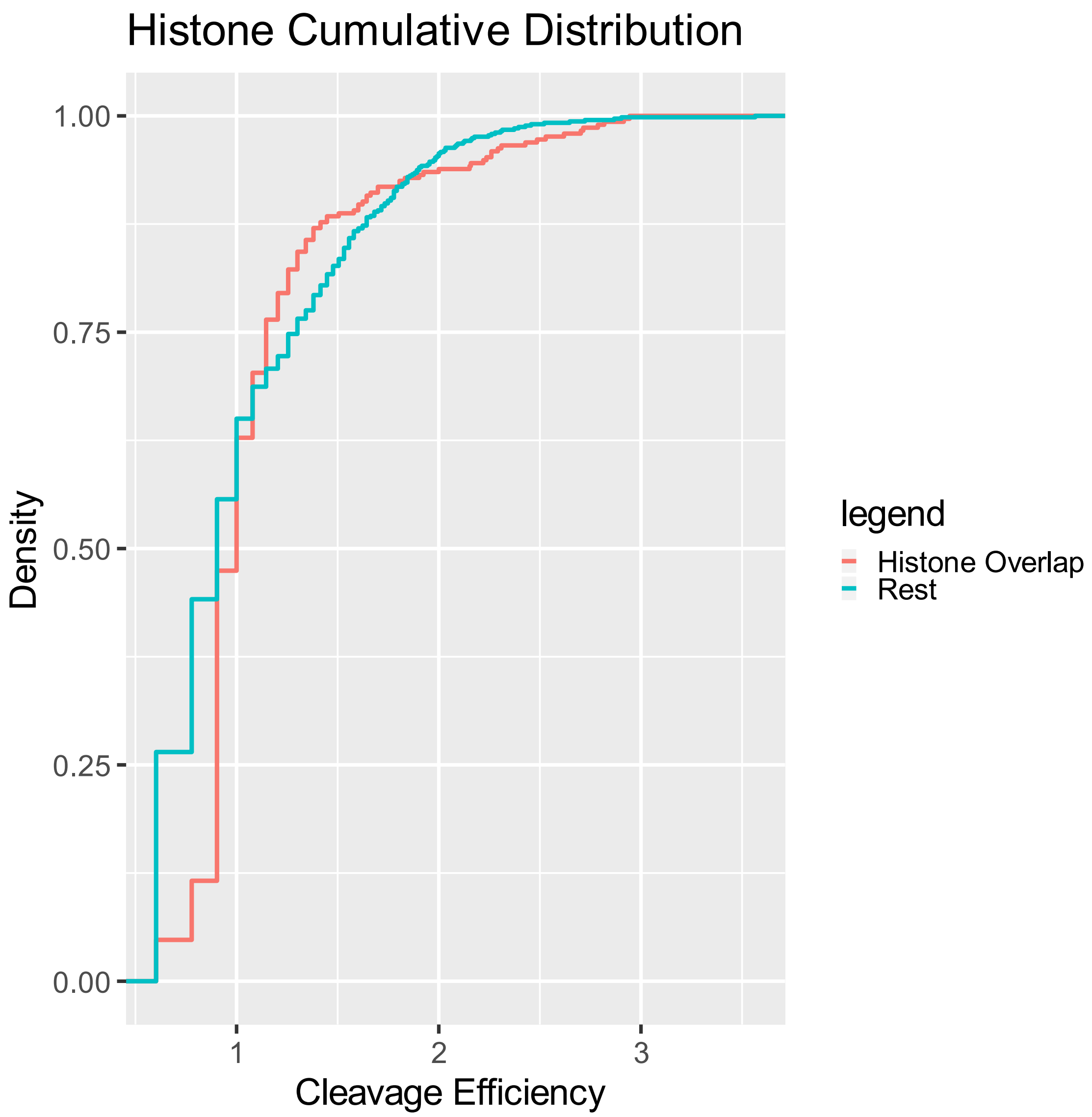


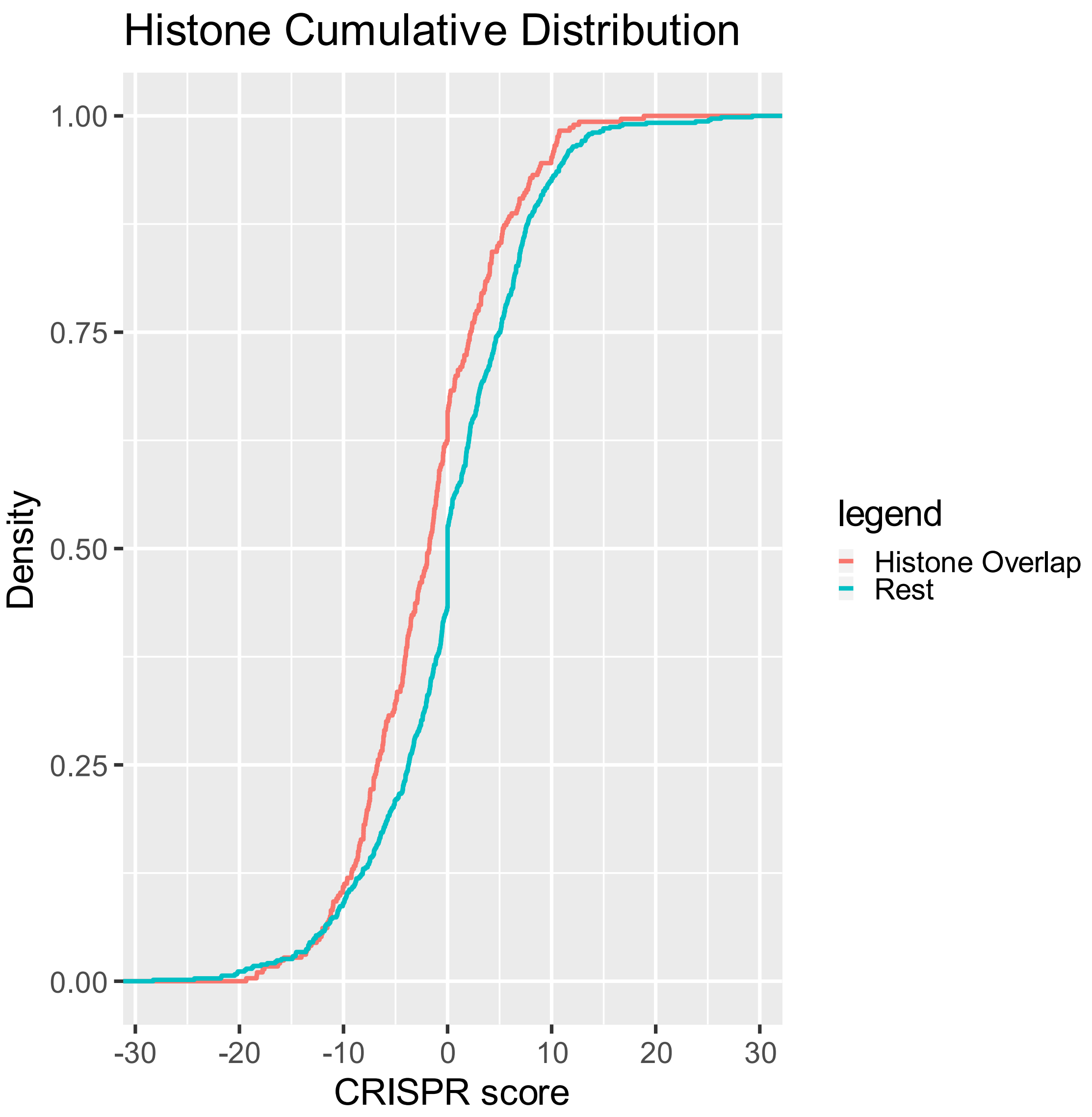
---

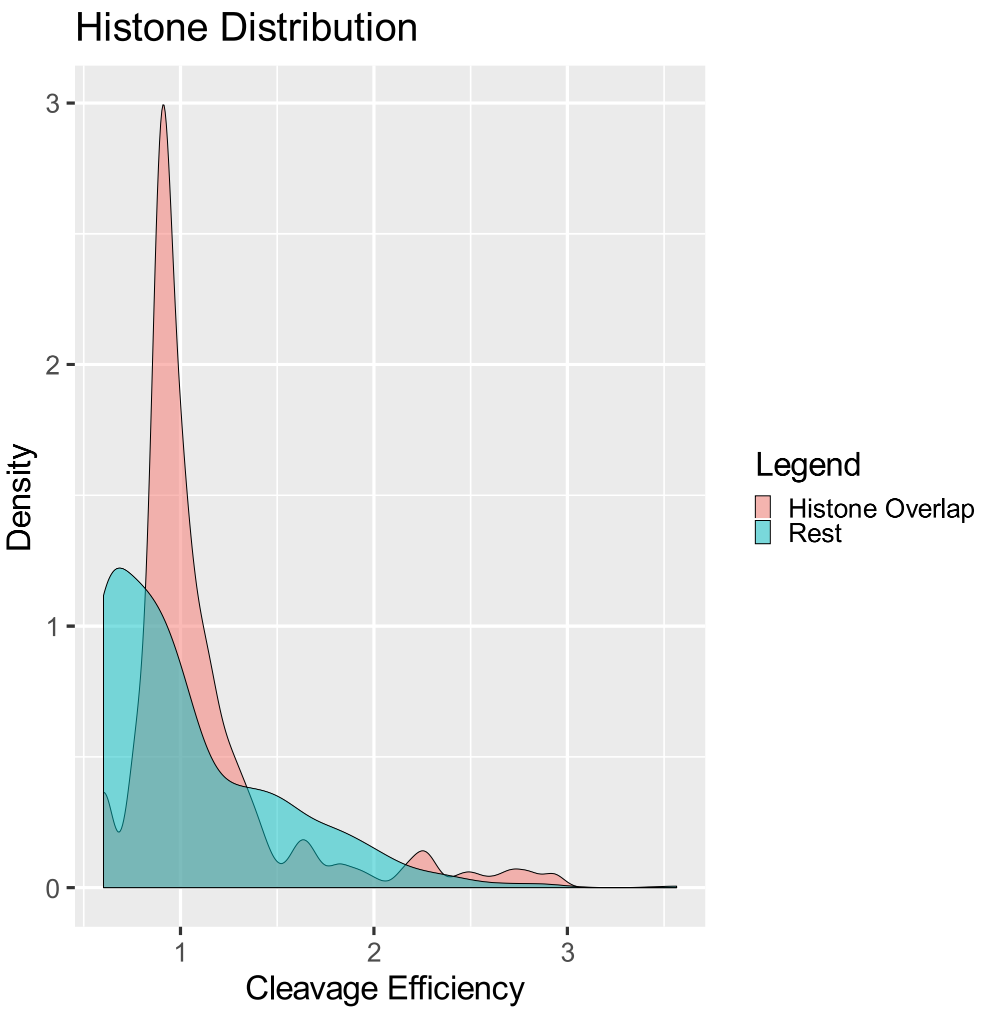
*1.* ENCODE Project Consortium. *Nature* 489.7414 (2012): 57-74

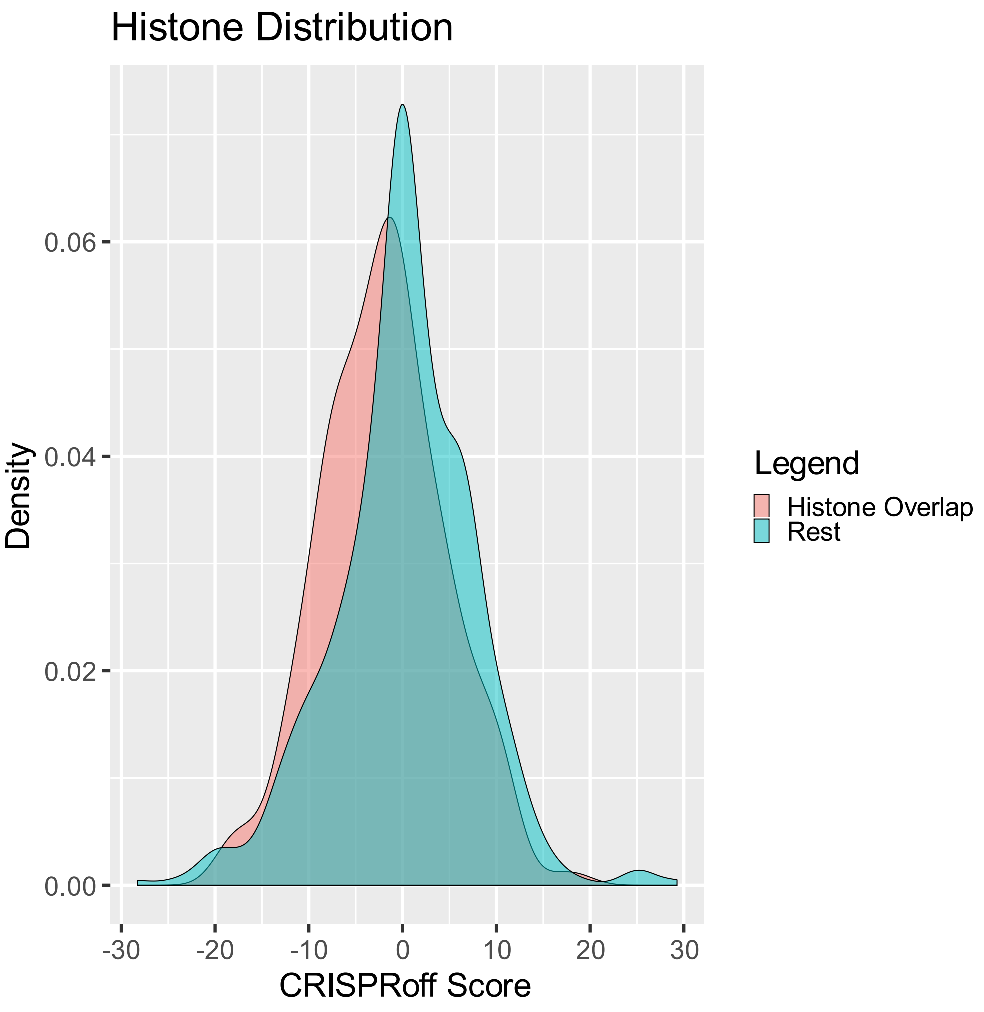
*2.* Meissner, Alexander, et al. *Nature* 454.7205 (2008): 766-770.

# Histone Distributions






Histone Distribution
Density
3
2
1
0
1
2
3
Cleavage Efficiency
Legend
Histone Overlap
Rest

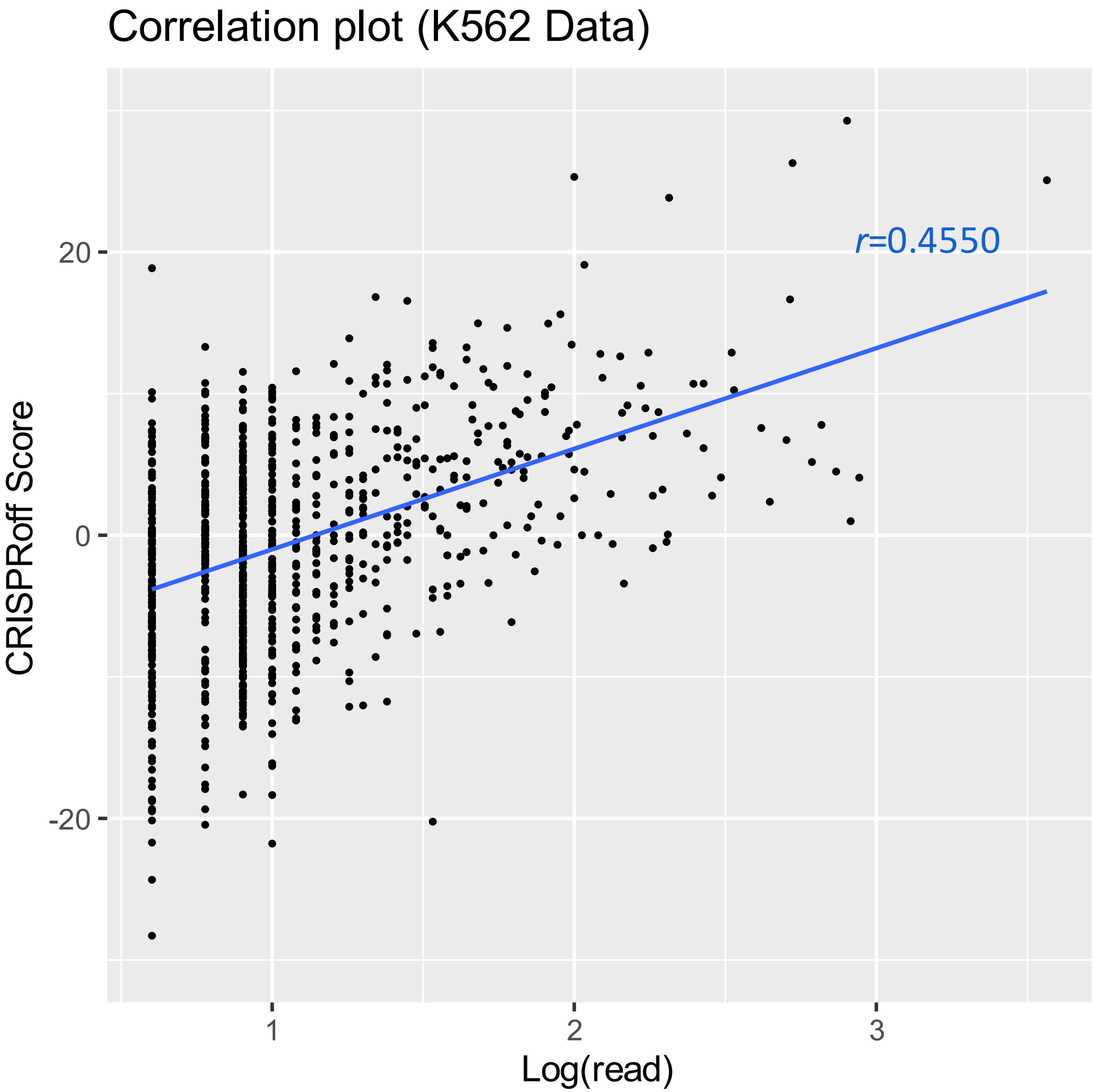

Histone Distribution
Density
0.06
0.04
0.02
0.00
-30
-20
-10
0
10
20
30
CRISPRoff Score
Legend
Histone Overlap
Rest

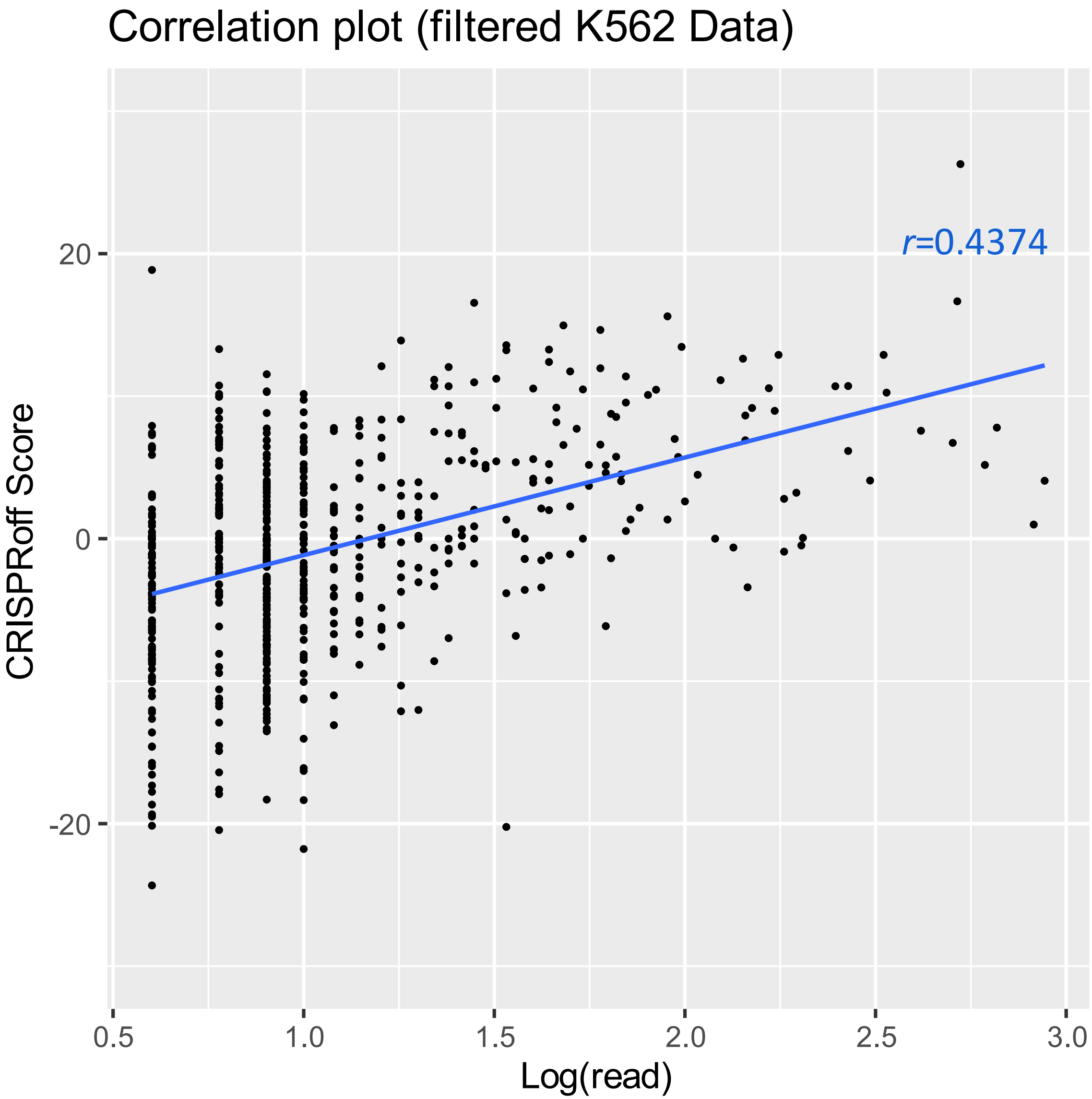
# Analysis of H3k27me3

Two approaches were attempted:

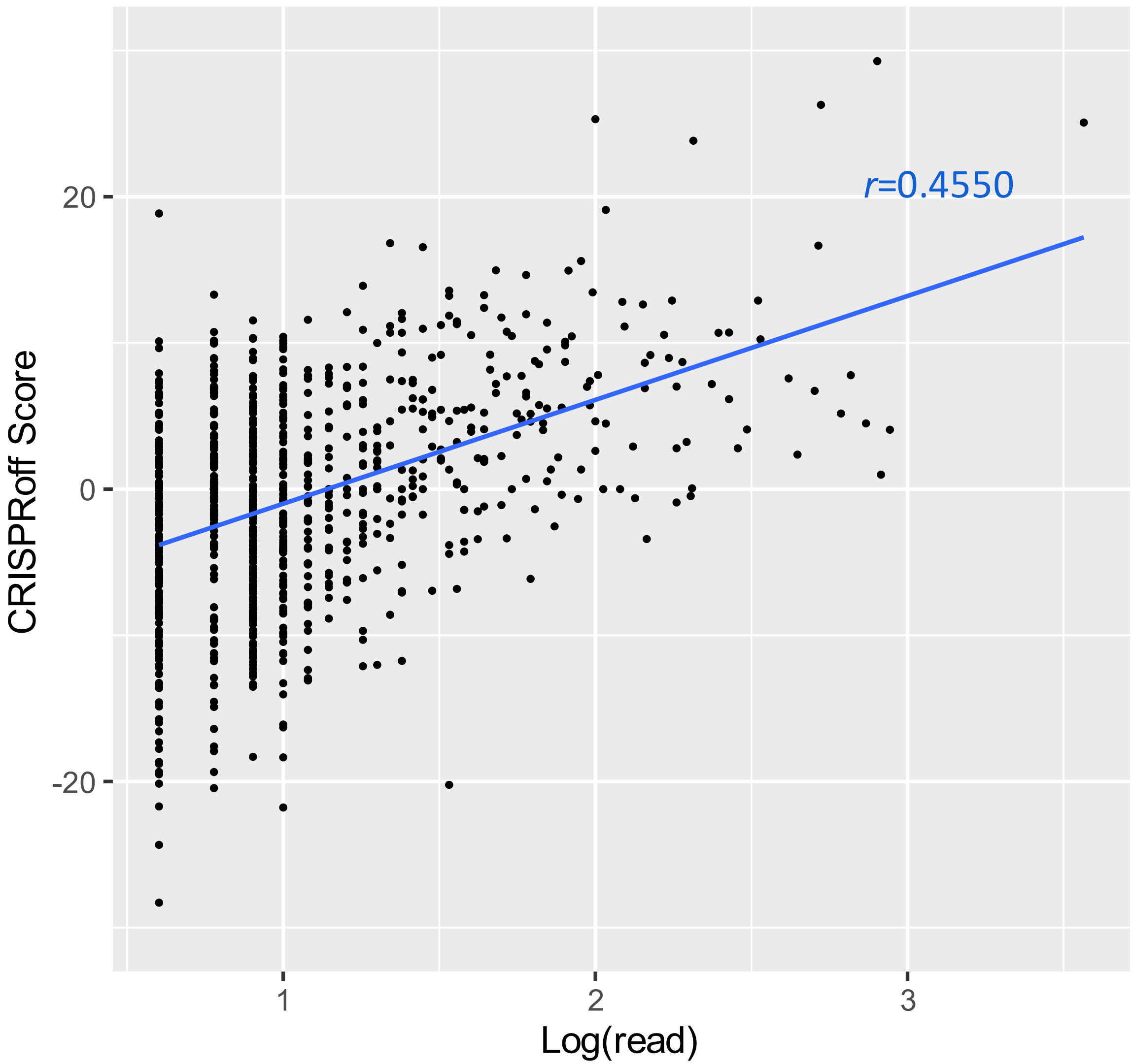
1. Removing any off-targets in a genomic region that has a reported Histone signal.
2. Removing off-targets in a genomic region that has a reported Histone signal that exceeds a given threshold (> 300, 400…).

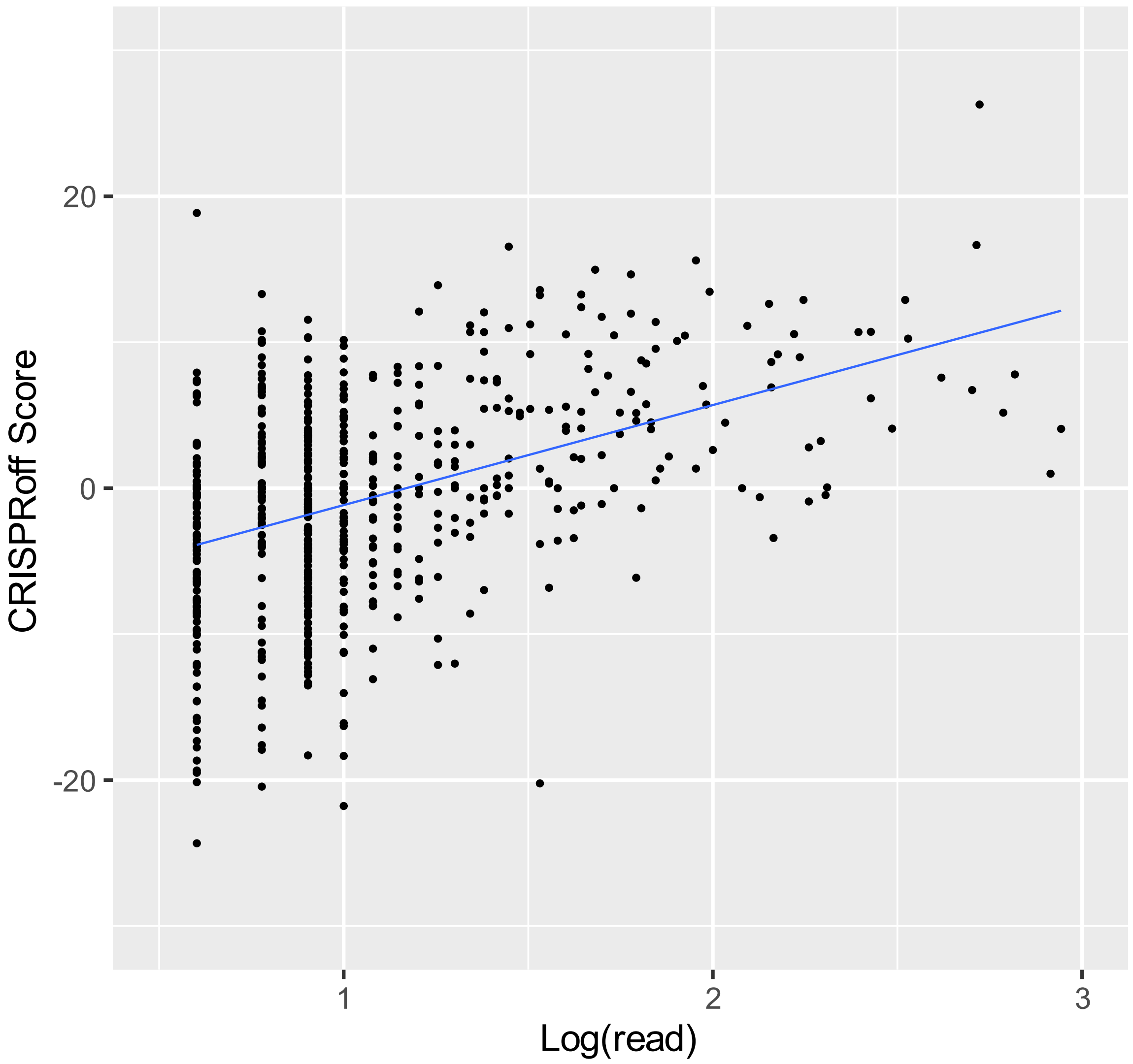
# Histone Modification Data  - Results





# Histone Modification Data - Results


Correlation plot (K562 Data)


Correlation plot(Filtered K562 Data)

# Exploration of Histone Modification Results

Randomly Removing Data points

1. Removing samples of same size(293)
2. Calculate Correlation Coefficient
3. Repeat 30 times
4. Compare to Histone filtered coefficient

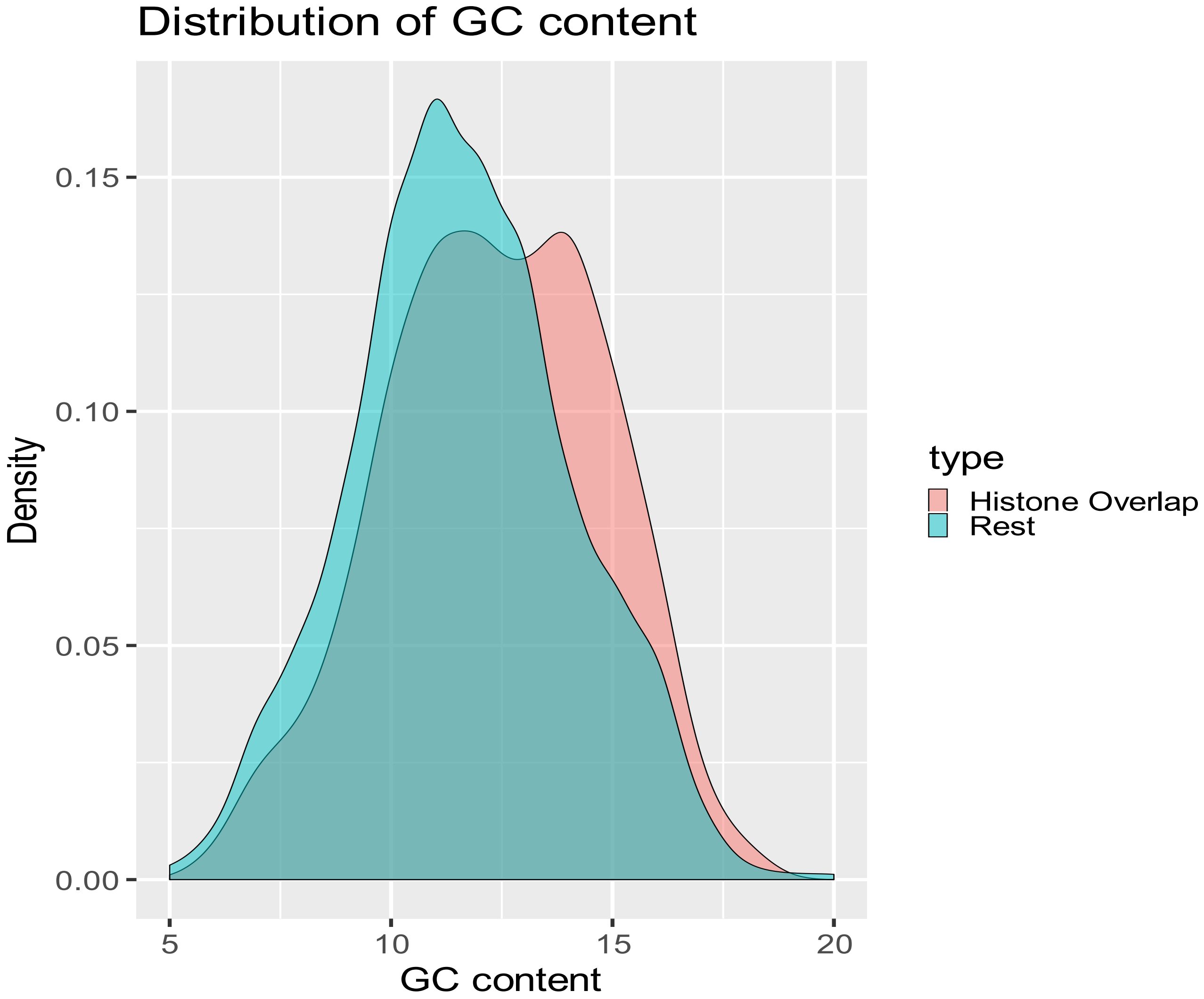
<u>Results</u>

Mean: 0.4436

SD: 0.019

<u>Result from Histone Exp.</u>

*r* = 0.4374



# Summary

- More suitable data is needed to proceed:
  - More significant overlap between Methylation and Off-target data.
  - Off-Target Data from an in Vivo study.

- Moving forward: We may need to generate our own data.

# Acknowledgements

**Special Thank You To:**

- Giulia Corsi
- Ferhat Alkan

  (Netherlands Cancer Institute)
- Christian Anthon
- Yonglun Luo

  (Aarhus University, Department of Biomedicine)
- Poul Hyttel

  (Copenhagen University, Pathobiological Sciences)
- Kristine Freude

  (Copenhagen University, Pathobiological Sciences)
- Jan Gorodkin

**Funding:**

Independent Research Fund Denmark, Technology and Production (9041-00317B)