

Washington University in St. Louis

Washington University Open Scholarship

Volume 13

Washington University
Undergraduate Research Digest

Spring 2018

gRNA Design Pipeline

Shrikar Iragavarapu

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/wuurd_vol13

Recommended Citation

Iragavarapu, Shrikar, "gRNA Design Pipeline" (2018). *Volume 13*. 88.
https://openscholarship.wustl.edu/wuurd_vol13/88

This Abstracts A-I is brought to you for free and open access by the Washington University Undergraduate Research Digest at Washington University Open Scholarship. It has been accepted for inclusion in Volume 13 by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

gRNA DESIGN PIPELINE

Shrikar Iragavarapu

Mentor: William Buchser

In recent years, CRISPR-Cas9 has made the headlines in both the general worldwide media for good reason. It is an invaluable tool for both *in vitro* and *in vivo* genome editing, used by scientists around the world to break, insert, or remove genes of interest. As a result, the technique allows scientists to study particular genes and create a variety of cell lines used today. The strands of RNA crucial to the functioning of CRISPR-Cas9, called single guide RNAs (gRNAs), can be designed in the lab to target particular genes. However, current methods of designing gRNAs that have high efficiency and specificity are still imperfect. My project is to create an algorithm that can predict the activity score/effectiveness of any gRNA sequence, based on an experimental dataset with over 3000 gRNA sequences and their corresponding activity scores. In order to do this, certain features of the gRNAs nucleotide sequence were analyzed for a possible relationship with their corresponding overall activity scores. Using R, a linear model was determined based on the dataset of 3000 gRNA sequences and corresponding activity scores, indicating the possibility of activity score prediction based solely on gRNA nucleotide sequence. Future direction includes incorporation of additional gRNA features into the existing algorithm and to consider indel pattern for each gRNA in order to increase the predictive accuracy of the algorithm.