BME 237 Problem Set 5: RNAi and CRISPR sgRNA scoring

My lab is interested in understanding the functional consequences of alternative

- aedaivds@ucsc.edu • 5/12/21

splicing. For example, we have identified alternative splicing changes associated with mutations in the splicing factor, U2AF1, which is frequently mutated in lung cancer and certain leukemias. Through RNA-Seq analysis, we can identify alternative splicing changes, transcriptome-wide. However, we do not understand if or how these splicing changes might change the function of these target genes. You can use this paper as a reference, but it is not necessary for this problem set: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0087361 One of the genes affected by U2AF1 mutation is CTNNB1 (Beta-catenin), which is a

well-characterized cancer driver. U2AF1 mutations cause a change in alternative 3' splice site usage in the 3' UTR of CTNNB1, which results in a longer 3' UTR. We

hypothesize that this longer 3' UTR will change the regulation of CTNNB1 due to additional miRNA binding sites that are present. We will start to test this hypothesis by using miRNA target prediction to find possible miRNAs that bind in this longer 3' UTR. Another gene affected by U2AF1 mutation is PTBP1 (which is, itself, a splicing factor). U2AF1 mutations cause skipping of an exon in PTBP1. We would like to experimentally test the downstream consequences of causing exon skipping in PTBP1 by designing sgRNAs that would direct Cas9 to cut close to either splice site of the

exon. Repairing the cut through non-homologous end joining may create mutations at the splice site, which would then cause exon skipping. **q1** To find possible miRNAs that specifically target the CTNNB1 3'UTR isoform associated with U2AF1 mutation, we will use TargetScan (Links to an external site.).

that are used in their database:

ENST00000396183.3) ENST00000396183.3 (more highly expressed in wild-type conditions. Downregulated with the U2AF1 mutation. Shorter 3' UTR than ENST00000396185.3)

ENST00000396185.3 (upregulated with U2AF1 mutation. Longer 3'UTR than

TargetScan has three different 3'UTRs deriving from annotated transcripts of CTNNB1

Use TargetScan to identify at least one miRNA that would target the 3'UTR of ENST00000396185.3, but not ENST00000396183.3. These are potential miRNA target sites that we can test to see if they affect CTNNB1 gene regulation in follow-up

experiments.

ENST00000349496.5 (expressed at a low level. Can disregard)

The miRNA is case-sensitive. For example, please give the name as miR-XX-XXX. If

the miRNA has an additional version number, please include, e.g. miR-XX-XXX.2

on the transcripts. This might be helpful. Look for a link that says "[View human genome browser (hg19)]"

TargetScan provides links to a UCSC Genome Browser view of the miRNA target sites

q1 Answer miR-139-5p and miR-142-3p.2 target scan

🔥 Music1 - Google D... 🖊 BME Graduate Stu... 📅 Bioinformatics Fre... 🙇 Discord

Release 7.2: March 2018

Gene expression a...

Agarwal et al., 2015

[Go to TargetScanMouse] [Go to TargetScanWorm]

[Go to TargetScanFly] [Go to TargetScanFish]

step 1 Target scan home page. Enter the gene we are

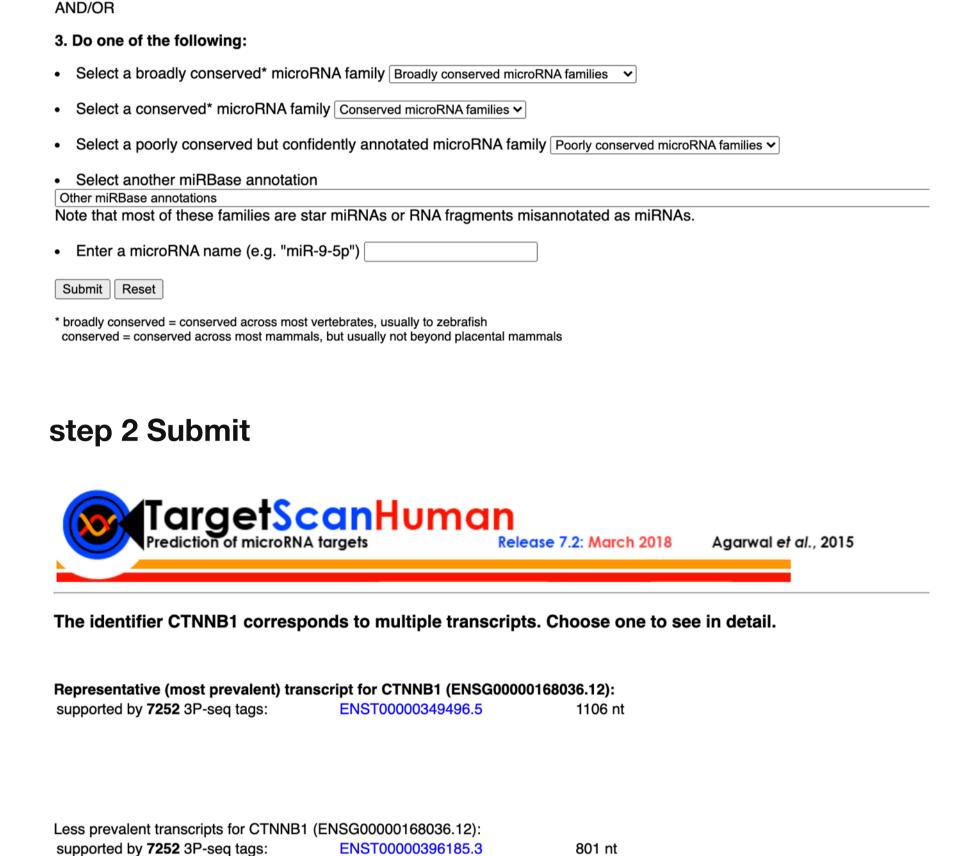
Search for predicted microRNA targets in mammals

interested in 'CTNNB1'

▲ Not Secure | targetscan.org/vert_72/

1. Select a species Human 2. Enter a human gene symbol (e.g. "Hmga2") CTNNB1

or an Ensembl gene (ENSG00000149948) or transcript (ENST00000403681) ID



ENST00000396183.3 ENST00000349496.5

ENST00000396185.3

supported by **7252** 3P-seq tags:

Notice the arrow points to a link to the UCSC Genome browser TargetScanHuman
Prediction of microRNA targets Re Release 7.2: March 2018 Agarwal et al., 2015

step 3select track ENST00000396185.3

Human CTNNB1 ENST00000396185.3 3' UTR length: 801

Conserved sites for miRNA families broadly conserved among vertebrates

[Show sites for other miRBase annotations, most of which are miRNA* sequences or RNA fragments

miR-150-5p

100

miR-142-3p.2

Show poorly conserved sites for miRNA families conserved among vertebrates] [Show conserved sites for miRNA families conserved only among mammals] Show poorly conserved sites for miRNA families conserved among mamma

[Show sites for poorly conserved but confidently annotated miRNA families]

miR-139-5p

isannotated as miRNAs]

[Show all species]

[Download SVG image of miRNA sites] [View table of miRNA sites]

[View human genome browser (hg19)]

ENST00000396183.3

642 nt

miR-142-3p.2

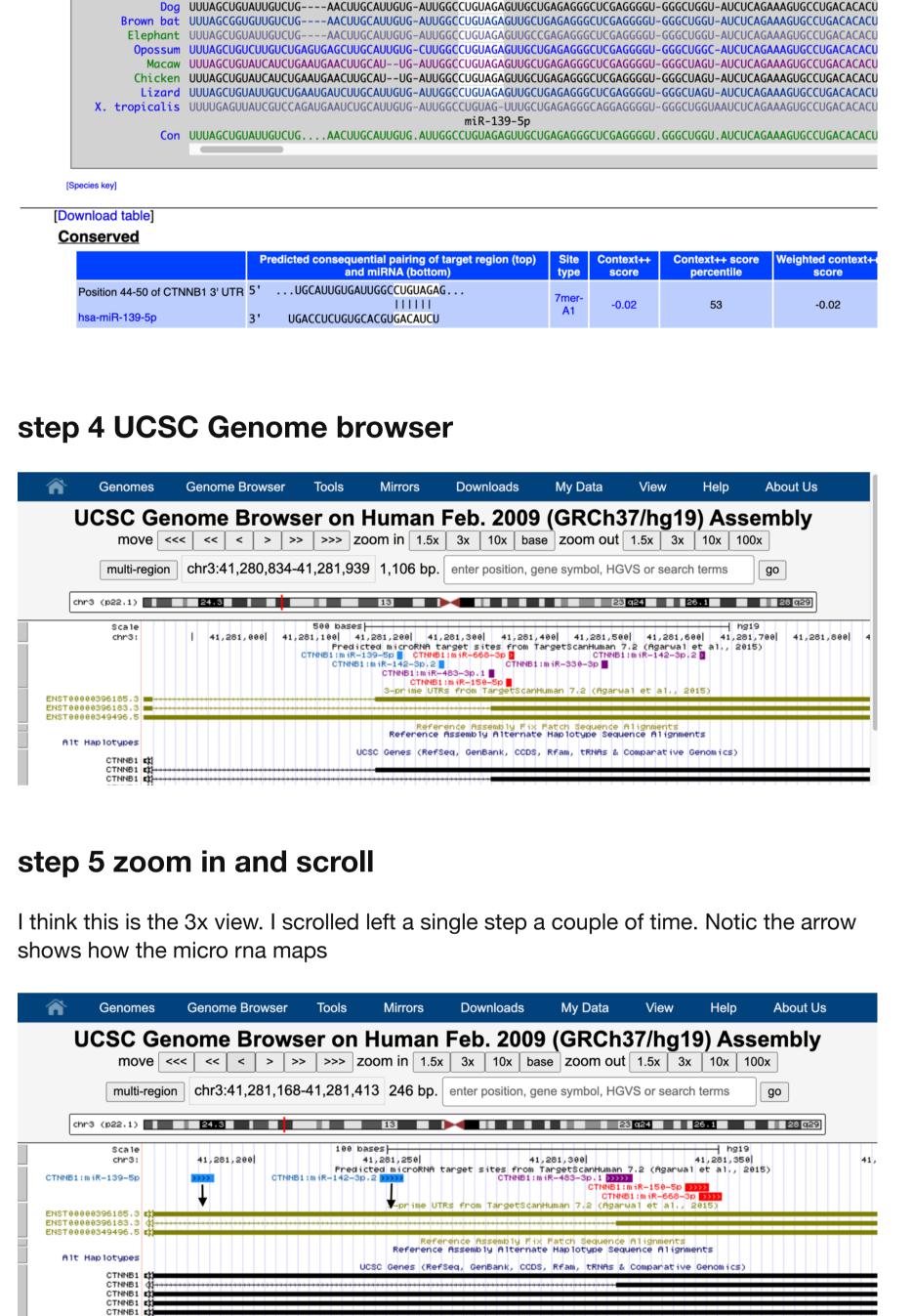
Sites with higher probability of preferential conservation

Sites with lower probability of preferential conservation 8mer 7mer-m8 7mer-A1 non-canonical

8mer 7mer-m8 7mer-A1 non-canonical

Key:

.10......20.......30.......40......50......60......70.....80......90......100. Human UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Chimp UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Rhesus UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Squirrel UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Mouse UUUAGCUGUAGUGUCUG----AACGUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGCUGGU-AUCUCAGAAAGUGCCUGACACACU Rat UUUAG-UGUAUUGUCUG----AACUUGCAUUGUG-ACUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Rabbit UUUAGCUAUAUUGUCUG----AACUUGCAAUGUG-AUUGGCCUGUAGAG-UGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Pig UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Cow UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU Cat UUUAGCUGUAUUGUCUG----AACUUGCAUUGUG-AUUGGCCUGUAGAGUUGCUGAGAGGGCUCGAGGGGU-GGGCUGGU-AUCUCAGAAAGUGCCUGACACACU



RefSeq genes from NCBI

About Us

refresh

My

ZO(

CRISPR Targets

Generan Genes

Predicted microRNA target sites from TargetSc

UCSC Genes (RefSeq, GenBank, CCDS, Rfam

A 305 S 306 P 307 Y 308 A 309 G 310 A 311 G

3-prime UTRs from TargetScanHuman 7 Reference Assembly Fix Patch Reference Assembly Alternate Haplot

300 Y 301 A 302 G 303 A 304 G

CRISPR/Cas9 -NGG Targets

RefSeq genes from

The UCSC Genome Browser contains a track called "CRISPR Targets", which is based

on the CRISPOR tool. We will use this to design our sgRNAs to target near either

What is one possible good sgRNA to select. These would be sequences colored

green, indicating "high predicted cleavage". Please provide the guide sequence.

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

enter position, gene symbol, HGVS or search terms

| 865,565| 865,516| 865,515| 865,526| 865,525| 865,536| 865,535| 865,546| 865,545| 865,556| 865,555| 865,556| G T A T A A T C T C A G C C T C T C C G T A T G C A G G A G C T G G T T T C C C T C C C A C C T T T G C C A T T C C T C A Predicted microRNA target sites from TargetScanHuman 7.2 (Agarwal et al., 2015)

move | << | < | > | >> | zoom in | 1.5x | 3x | 10x | base | zoom out | 1.5x | 3x

The PTBP1 exon coordinate is chr19:805,492-805,569 (hg19).

chr19:805,492-805,569 78 bp.

UCSC Genome Browser after entering coordinates **Genome Browser Tools**

GGAGAGGCTGAGATTATACC

chr19 (p13.3) 19p13.3

Alt Haplotype

DNase Cluster

refresh button

<u> AceView Genes</u>

hide

=x∩ninhv

q2 answer

splice site of the PTBP1 exon.

RefSeq Curated |

q2

move <<<

PTBP1 📢

PTBP1

PTBP1

Alt Haplotypes

RefSeq Curated

step 2 select CRISPR Targets

AUGUSTUS

Updated GENCODE

step 3 click on one of the green CRISPR track

hide

Genes and Gene Predictions

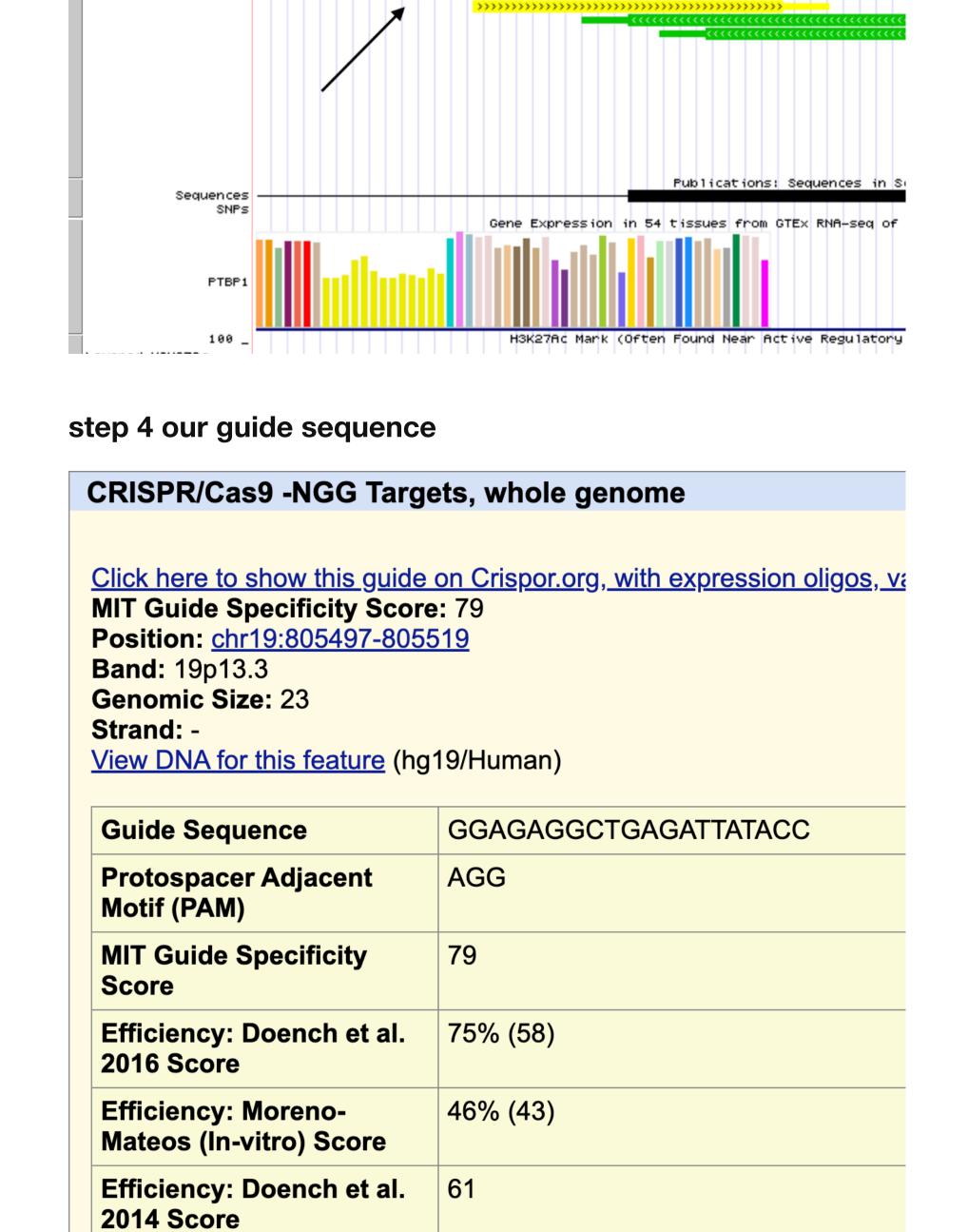
Genomes Genome Browser **Tools** Mirrors **Downloads** UCSC Genome Browser on Human Feb. 2009 zoom in 1.5x chr19:805,492-805,569 78 bp. multi-region enter position, gene symbol, chr19 (p13.3) 19p13.3 | 805,500| 805,505| 805,510| 805,515| 805,520| 805,525| 805,530| 805 | GTGCACCTGGTATAATCTCAGCCTCTCCGTATGCAGGAGCTG chr19:

scroll down or search. make sure the track is not hidden. Make sure you press the

CCDS

Geneid Genes

hide



55

Bae et al. Out-of-Frame

Score