M2D7: Examine features in gRNA-targeted genomic sequences

- Prelab discussion
- Perform computational analysis of MG1655 regulatory elements potentially affected by gRNA binding
- Work on research article with extra time

M2D7: Examine features in gRNA-targeted genomic sequences

- Prelab discussion
- Perform computational analysis of MG1655 regulatory elements potentially affected by gRNA binding
- Work on research article with extra time

Quiz ~~Today~~ Friday

Office Hours (T/R): T- Becky 16-317

R- Noreen 16-319

Either- Jamie 16-469

Mod2 Overview

Research goal: Increase the yield of commercially valuable byproducts in *E.coli* using CRIPSRi technology to target genes involved in mixed-acid fermentation pathway.

Last Lab:

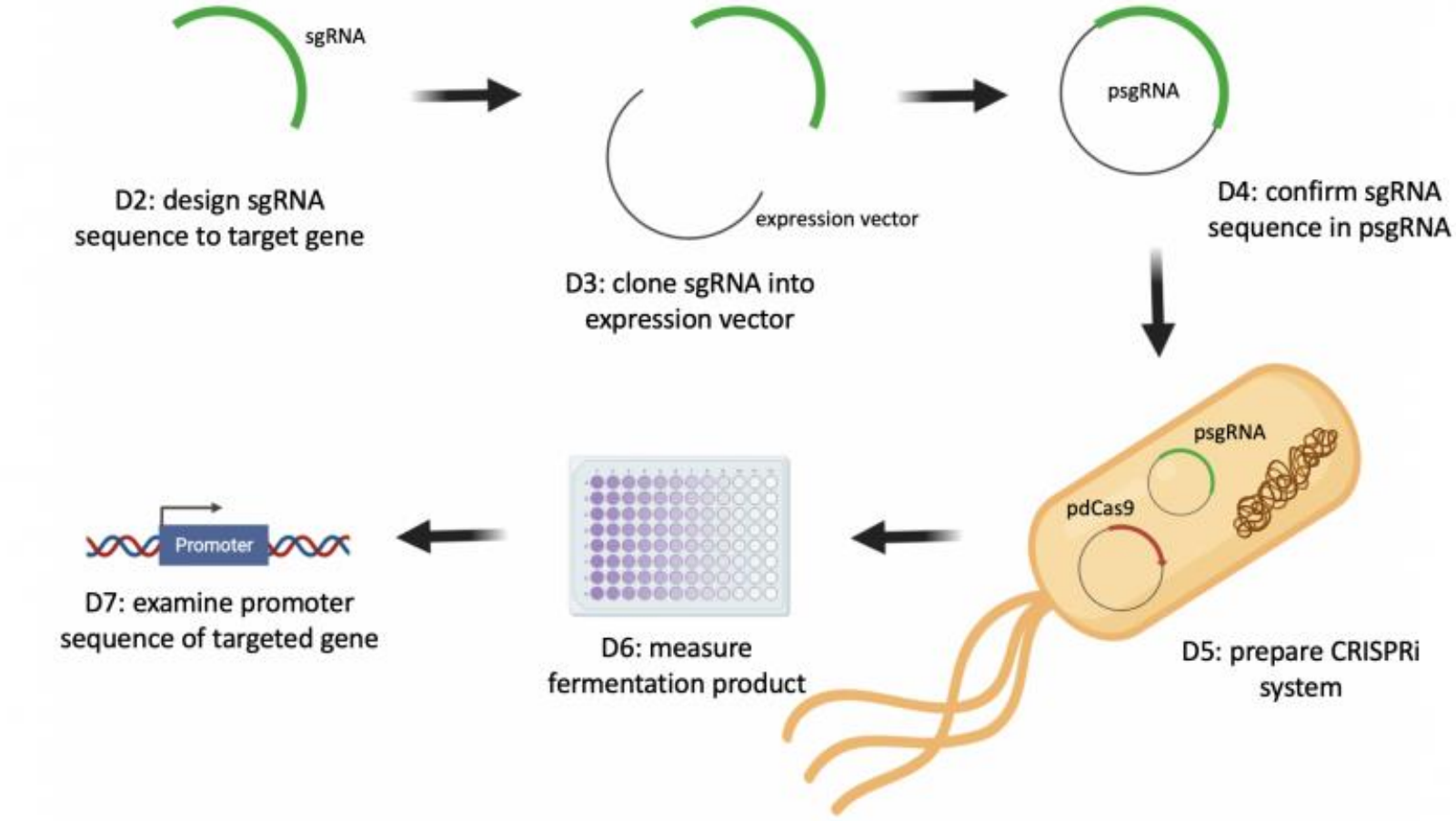
Measure bacteria O.D. and fermentation products

This Lab:

Examine DNA regulatory elements that may impact the efficacy of your CRISPRi system

Next Lab:

Organize figures and outline text for the research article



Design an experiment

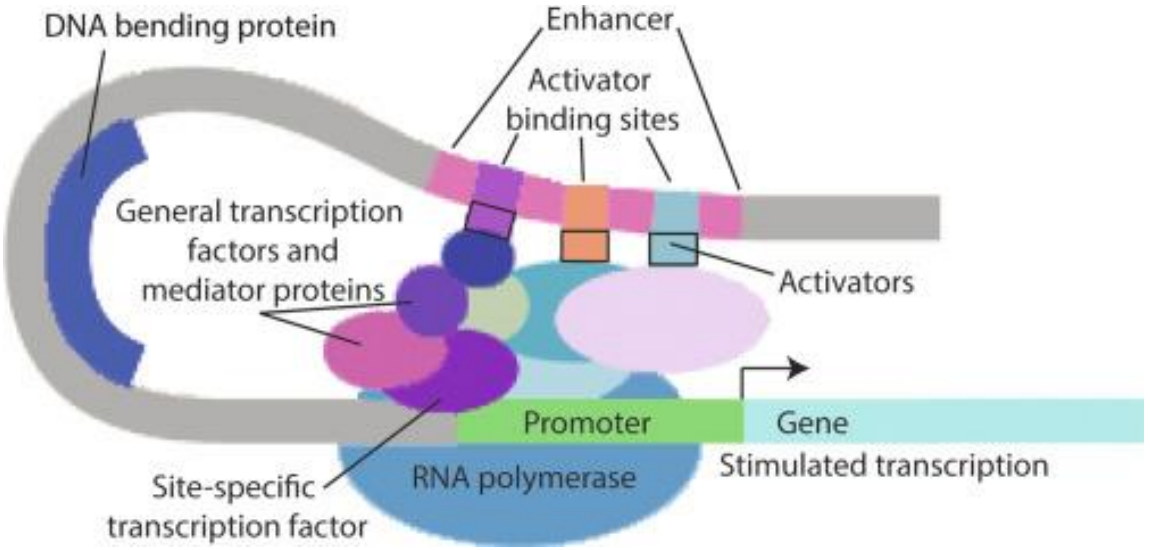
Collect data

Interpret data

Reflect

Why would you care about transcription factor binding if you didn't target a TFBS?

- We have been focused on using our CRISPRi system to block RNAP
- Other regulatory elements, including transcription factors, can be affected by the binding of your CRISPRi complex
- Your project involved engineering a gene regulation complex
  - When considering the efficacy of your designed system, you also need to consider any potential unexpected consequences



ScienceDirect

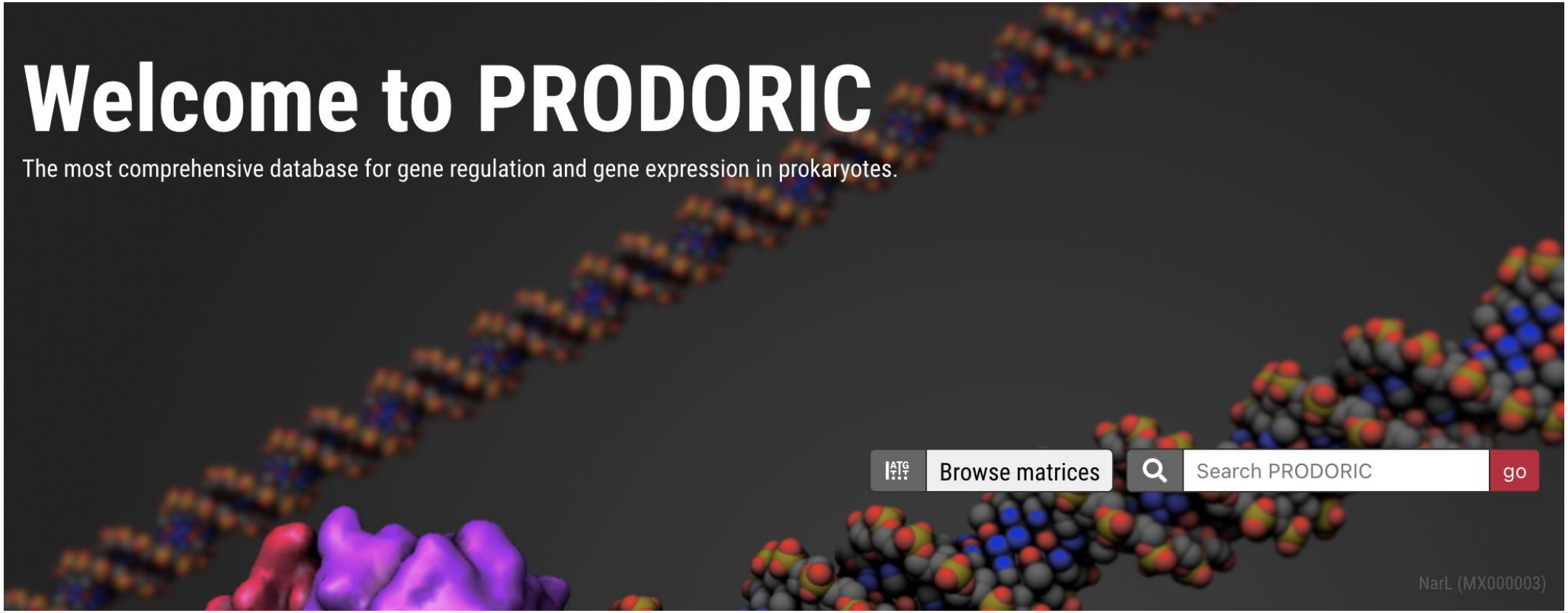
Basic overview of today's lab

Get your gRNA sequence + some bases up and down

Check if 25 transcription factors might bind in/around your gRNA target sequence

Transcription factors bind all sorts of different sites – use a **P**ositional **W**eight **M**atrix to check all of them

Determine the potential for your gRNA to inhibit transcription factor binding



From ChIP-Seq to PWM

Chromatin
Immuno**P**recipitation

Transcription factor sequence binding data

PFM f

base b					
A	5	0	1	0	0
C	0	2	2	4	0
G	0	3	1	0	4
T	0	0	1	1	1

Positional **Frequency** Matrix (PFM)

Easy conversion to Positional **Probability** Matrix (PPM)

~Math~

Log (probability of finding this base)

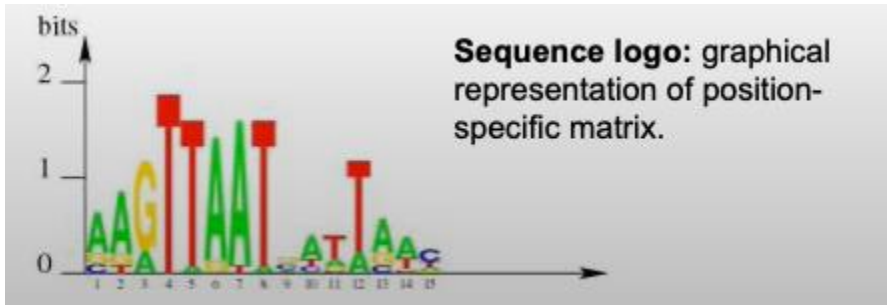
(background nucleobase frequency)

PWM

A	1.6	−1.7	−0.2	−1.7	−1.7
C	−1.7	0.5	0.5	1.3	−1.7
G	−1.7	1.0	−0.2	−1.7	1.3
T	−1.7	−1.7	−0.2	−0.2	−0.2

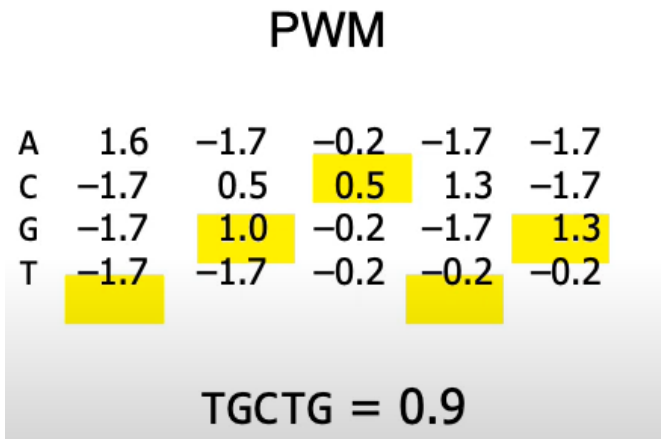
TGCTG = 0.9

Positional **Weight** Matrix



> 0 – More likely to be functional
< 0 – More likely to be random

Position Weight Matrix Takeaway



1) PWMs allow you to check for putative TF binding because TFs bind multiple sites

2) PWMs are a function of:
    1) How likely it is at each position for a certain base to show up
    2) The background frequency of that particular base

3) A score for a predicted sequence can be calculated by summing a base's score at every position

4) Larger total scores = more likely to be hits
   Smaller total scores = more likely to be random

5) A Core score is calculated by the max score of any 6 consecutive nucleotides

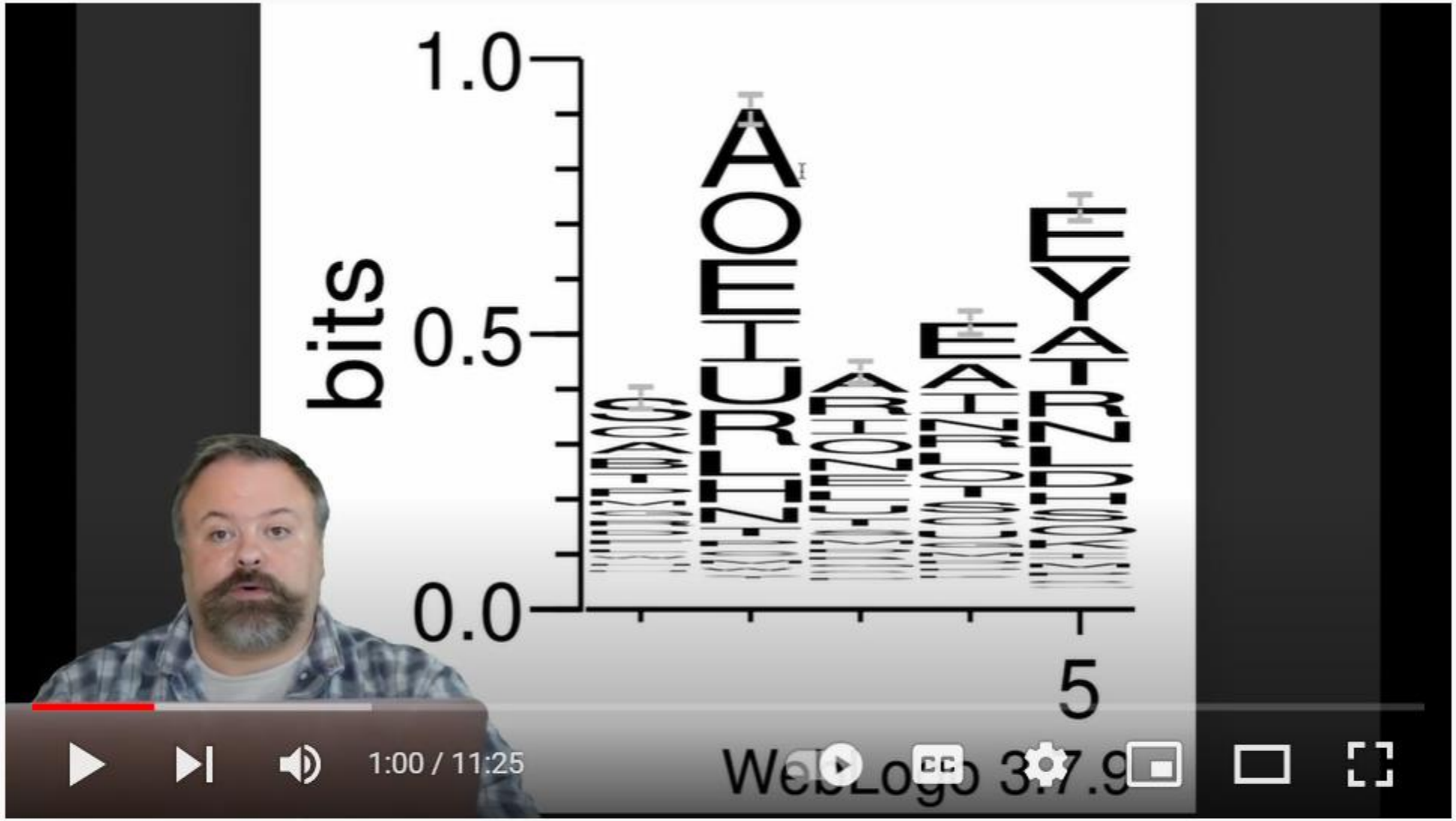
Extra cool math

$$R_i = \log_2(4) - (H_i + e_n)$$

4 bases

Entropy – high if outcome is uncertain

Low if outcome is certain



Creating a sequence logo of 5-letter words for Wordle

111 views...

4 DISLIKE SHARE DOWNLOAD SAVE ...

For Today

- Complete the wiki exercises on regulatory elements
- With extra time, work on:
  - Data analysis
  - Research article outline
  - Homework for M2D8

For M2D8…

- Outline the discussion section for the Research Article

Discussion section

- Start with ‘here we showed…’
  - Restate major results and broad implications
  - Follow same order as in Figures/Results
- Describe your conclusions from your data
  - If necessary, describe caveats of experiment and suggest improvements
- Identify unknowns and speculate (within reason)
  - Don’t make huge generalizations or overreach
- Propose future experiments, identify new questions that arise
- Come back to the big picture / impact statement topic introduced in background