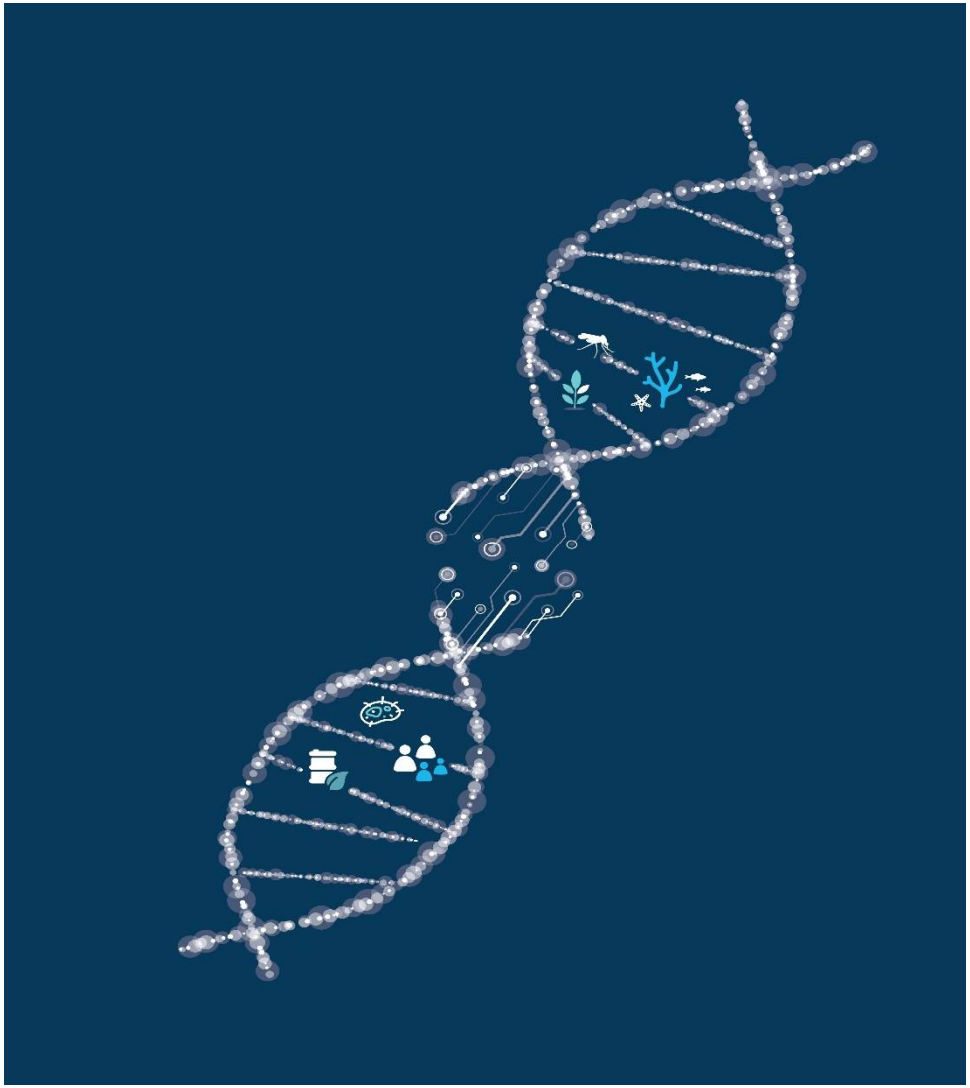
CSIRO

Machine Learning guided Design of Ribosome Binding Sites

Synthetic Biology FSP

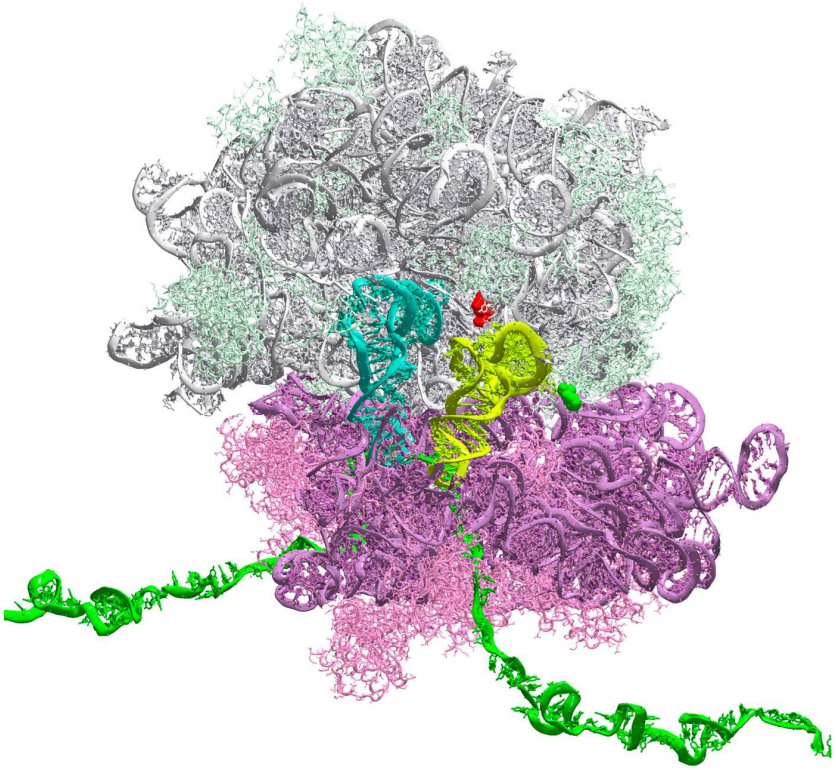
Maciej Holowko | 18/06/21



CSIRO

The RBS Problem

- Ribosome Binding Site (RBS) controls protein expression level
- Translation Initiation Rate (TIR) of is hard to predict and is expensive to label
- RBS calculators are based on thermodynamic data



CSIRO

Machine Learning Algorithms

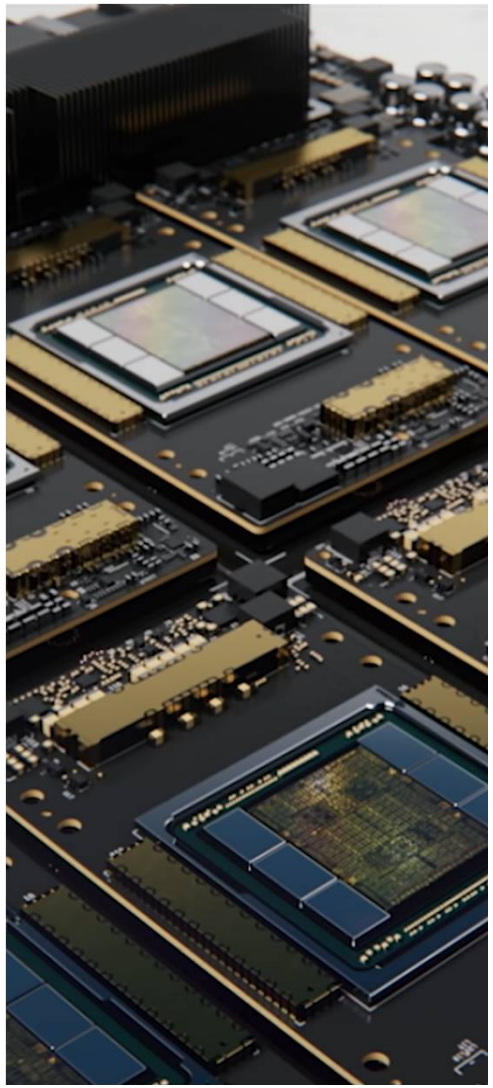
A (Bayesian) regression algorithm which **predicts** both

- Mean
- Uncertainty

→ Gaussian Process Regression

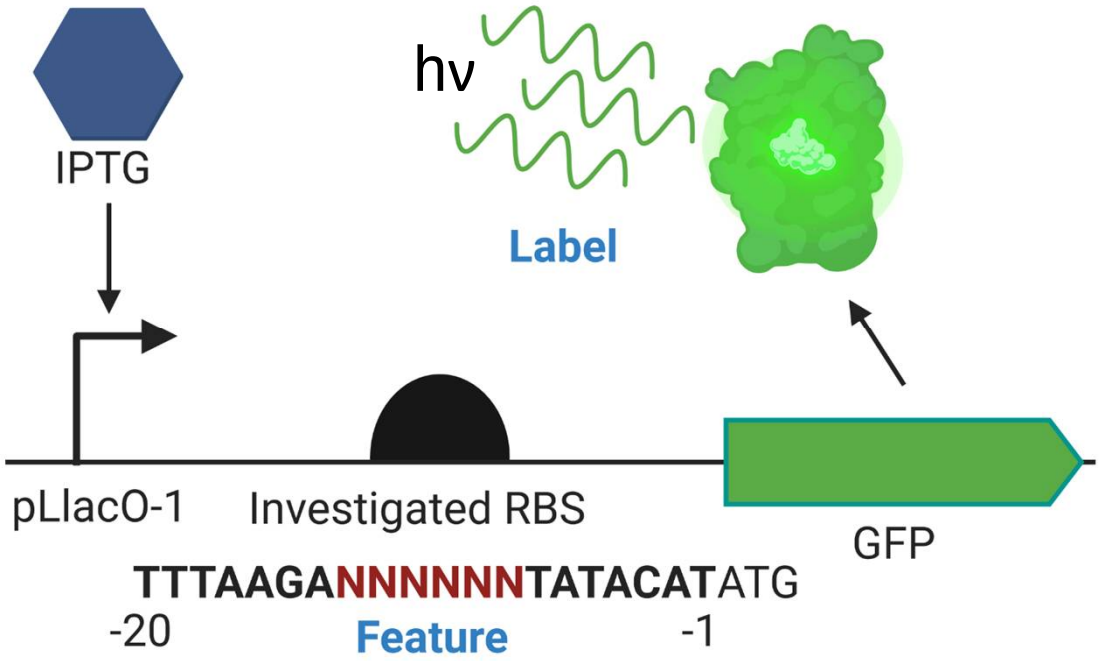
An online/batch algorithm which **recommends** sequences to test

→ Multiarmed Bandits Algorithms: Upper Confidence Bound

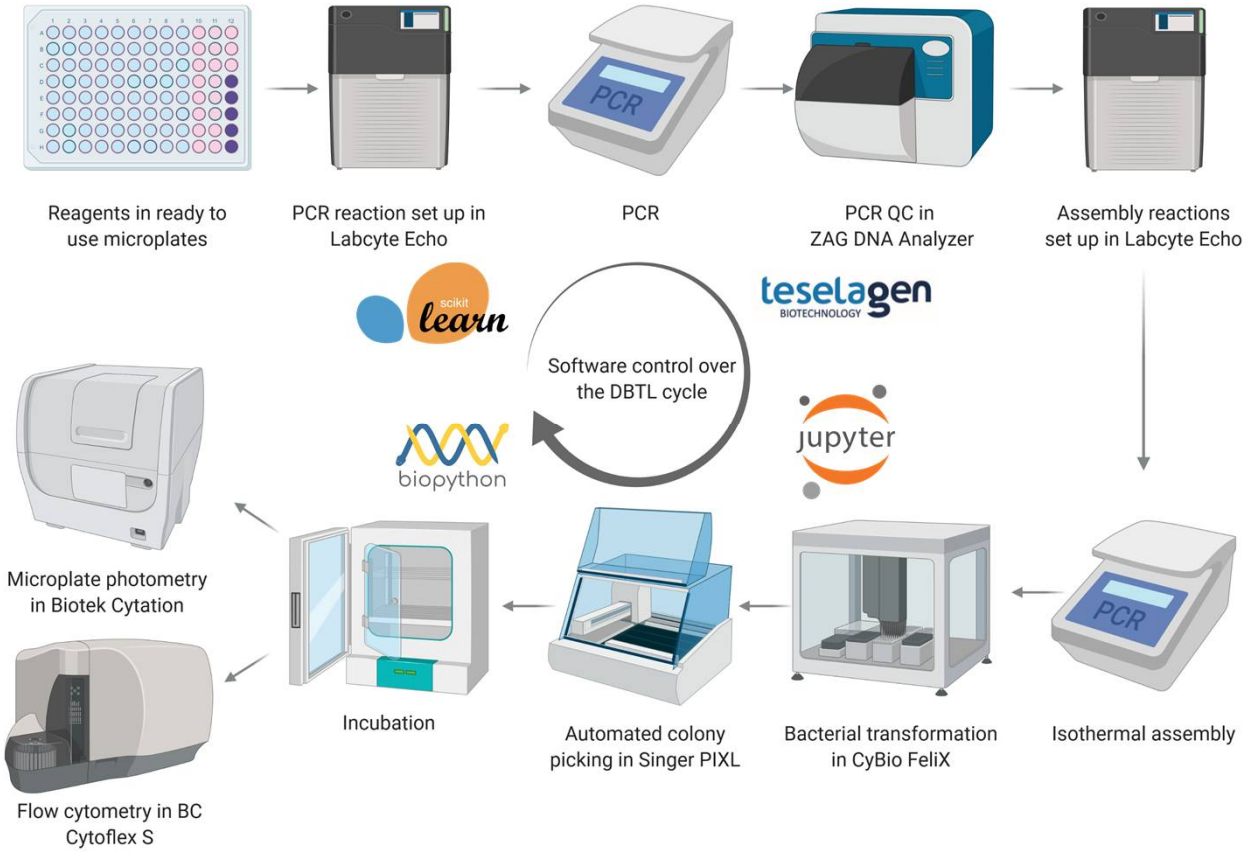


CSIRO

Genetic construct



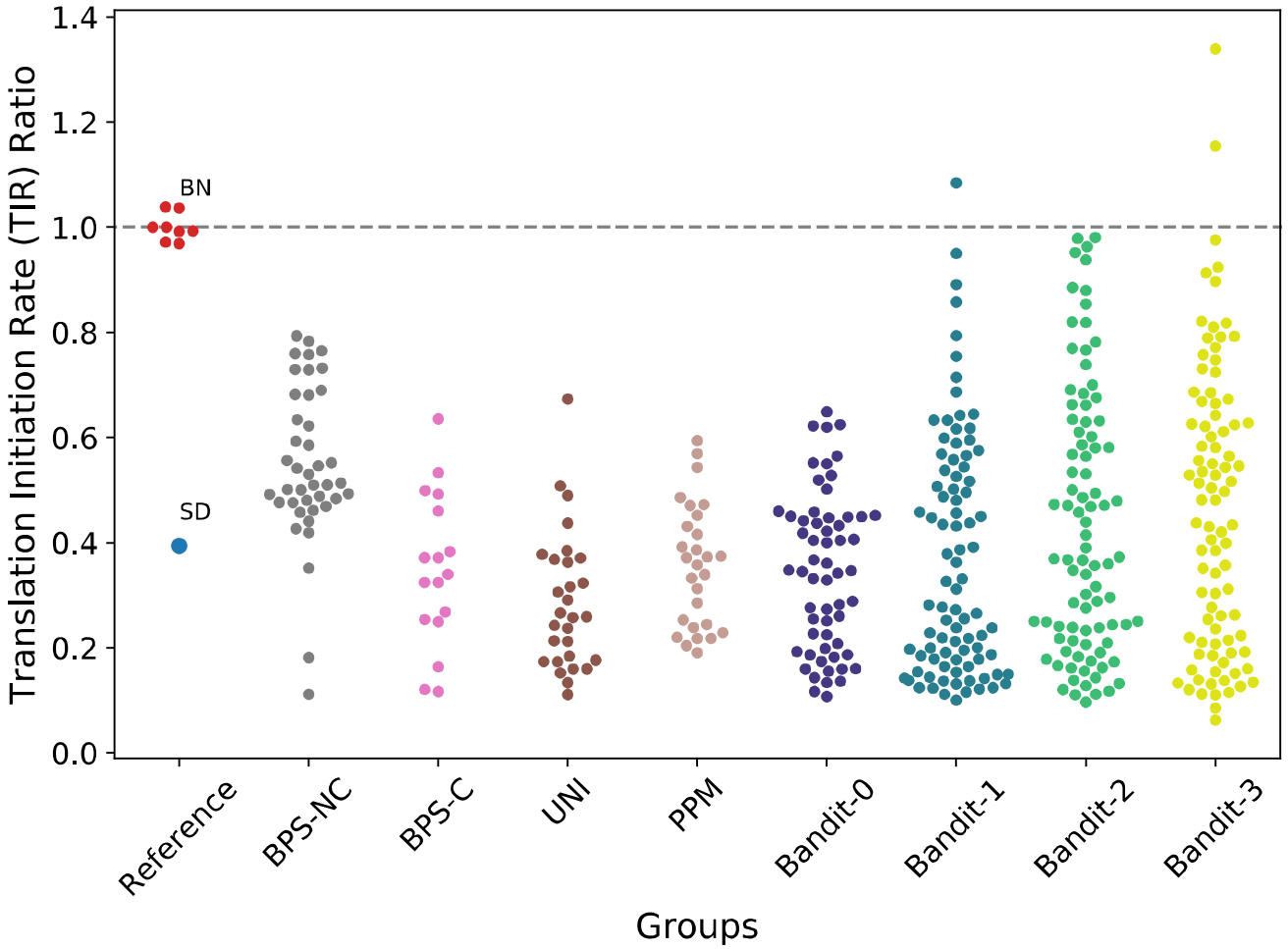
High-throughput RBS engineering



CSIRO

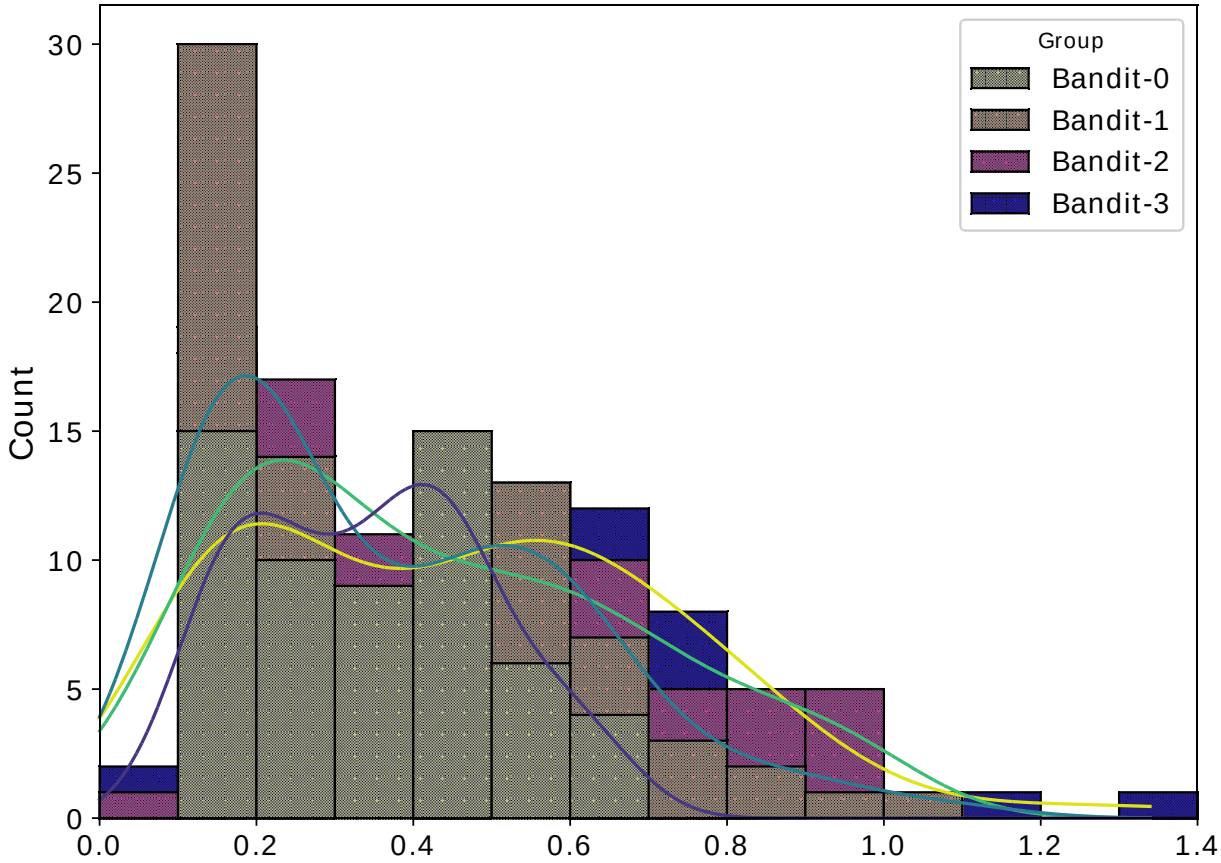
CSIRO

Results



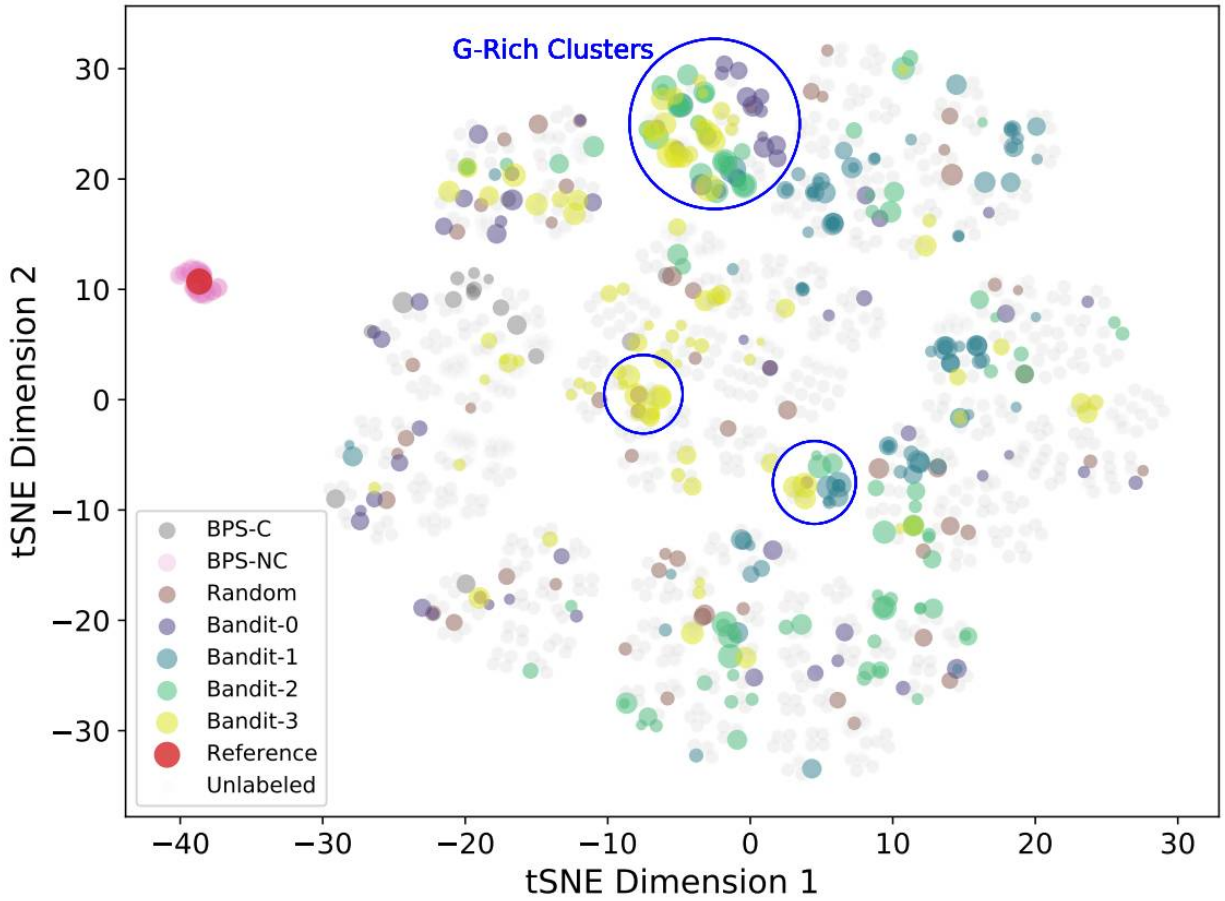
CSIRO

Histogram

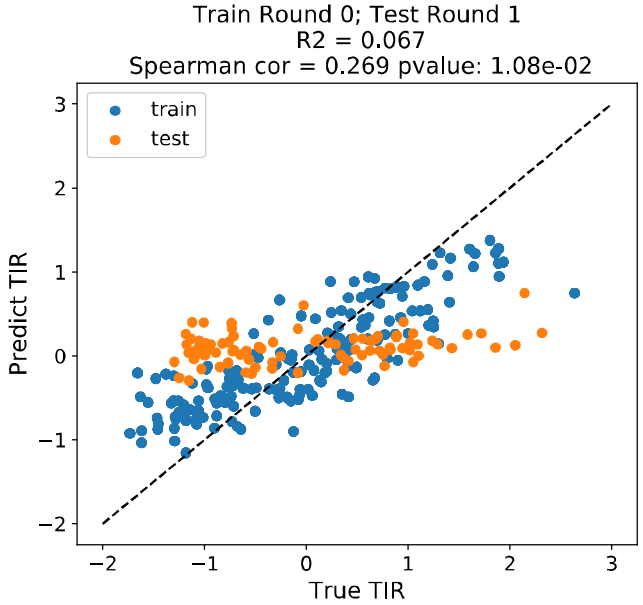


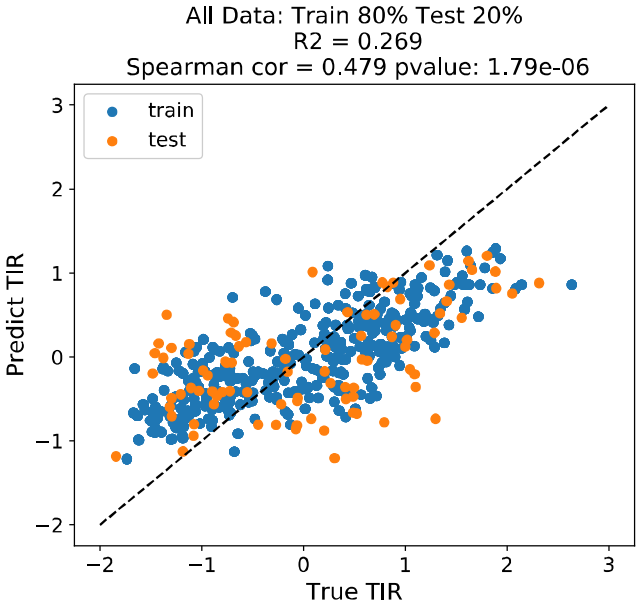
CSIRO

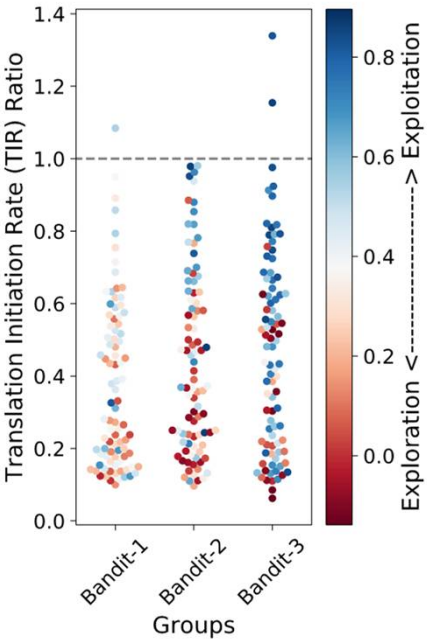
t-SNE


G-Rich Clusters
BPS-C
BPS-NC
Random
Bandit-0
Bandit-1
Bandit-2
Bandit-3
Reference
Unlabeled
tSNE Dimension 1
tSNE Dimension 2
30
20
10
0
−10
−20
−30
−40
−30
−20
−10
0
10
20
30

Prediction Efficiency







CSIRO

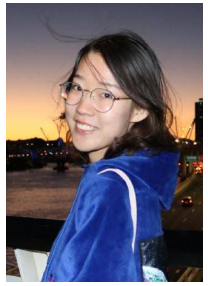
CSIRO

Summary

- Biologically important rules found
- Significant decrease in discovery time
- Reliable library with a number of very strong RBSs

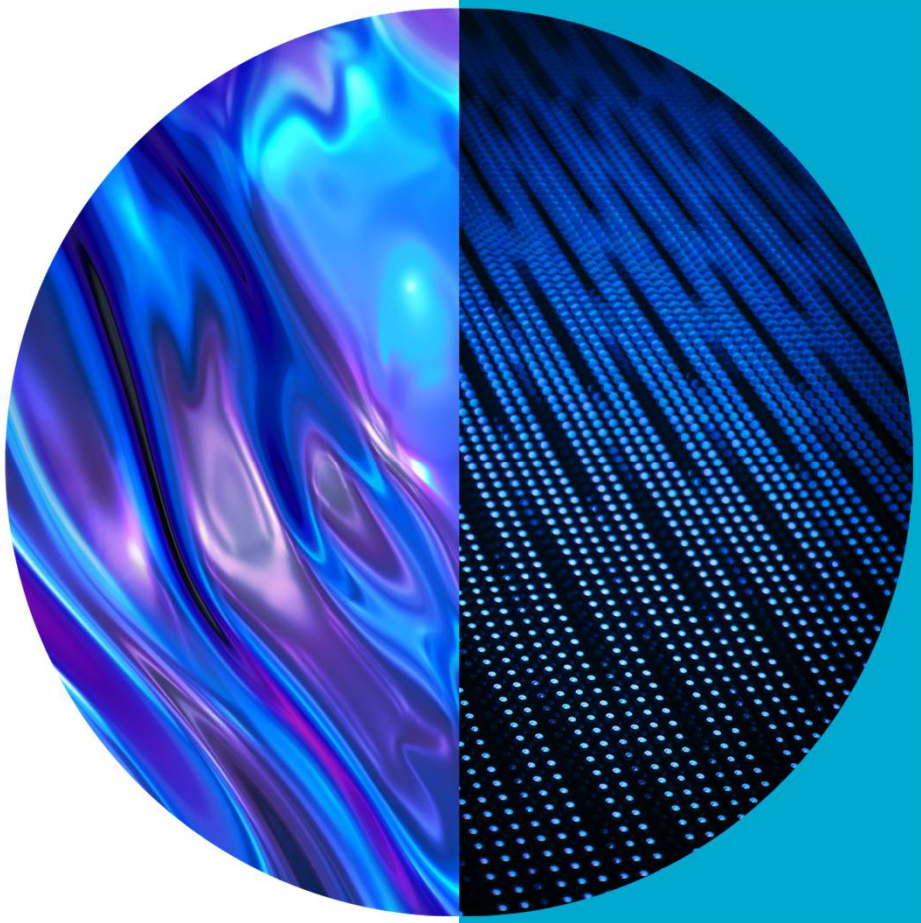
Top RBS Core	TIR Ratio
GGGGGC	1.34
GGGGGT	1.15
GGCTAT	1.08
AGGAGA	1
GGCGTT	0.98
GGGGGG	0.98
GGCGAC	0.98
CAGGAG	0.96
GGCGAG	0.95
AGGAGG	0.39

CSIRO



Mengyan Zhang

Huw Hayman Zumpe



Cheng Soon Ong

Claudia Vickers

Australia's Pre-eminent National Science Organization

CSIRO

Thank you

Land & Water/SynBio FSP
Maciej Holowko
Synbio Stack Developer

+61 2 94905764
maciej.holowko@csiro.au
https://research.csiro.au