QBio Extra Credit Assignment Report

High-throughput binding assays:

In Vitro experiments:

SELEX–seq (Systematic Evolution of Ligands by Exponential Enrichment Sequencing)

SELEX-seq is a method used to identify DNA or RNA sequences that have a high affinity for specific proteins. It identifies high-affinity nucleic acid sequences that bind specific proteins using iterative selection and amplification, followed by sequencing.

Advantages – identifies high affinity binding sites.

Disadvantages – in vitro settings, potential PCR biases

PBM (Protein Binding Microarray)

This assesses protein DNA interactions by incubating a microarray with thousands of DNA sequences with a labelled protein and analysing binding via fluorescence.

Advantages – identifies high-affinity binding sites.

Disadvantages – in vitro, limited to arrayed sequences

In Vivo experiments:

ChIP – seq (Chromatin Immunoprecipitation – sequencing)

ChIP-seq determines in vivo protein DNA interactions by capturing DNA bound to specific protein in cells, then sequencing the enriched DNA fragments.

Advantages – in vivo insights, genomic context

Disadvantages – requires specific antibody, limited resolution.

A graph of different colored bars

Description automatically generated with medium confidence

From these results, it appears that both models have very similar average R-squared results for the three data sets. Since a higher R-squared result indicated a better fir of the model to the data, for both mad and max datasets the 1-mer+shape model is a better fit and for myc, 1-mer is a better fit. Since the results are quite consistent across the three data sets, it suggests that the models are robust to dataset-specific variations.