

PREDICTING YEAST TRANSLATION EFFICIENCY FROM RNA MODIFICATION AND DEGRADATION FEATURES: A MACHINE LEARNING APPROACH

Orlane Rossini ¹, Alice Cleyen ^{1,2}, Nikolay Shirokikh ² and Shafi Mahmud ²

¹IMAG, Univ Montpellier, CNRS, Montpellier, France

²John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia

November 25th 2024



UNIVERSITÉ DE
MONTPELLIER

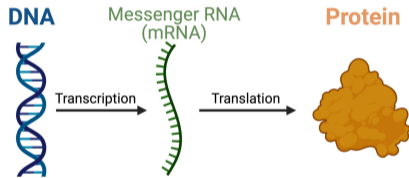


Australian
National
University

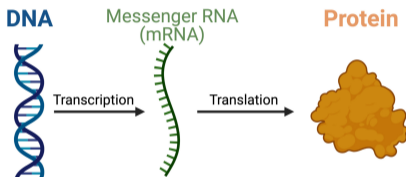
IMAG
INSTITUT MONTPELLIERAIN
ALEXANDER GROTHENDIECK



anr[®]



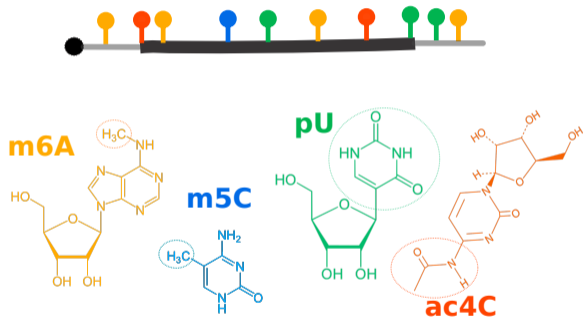
How can translation efficiency be predicted from RNA features?



How can translation efficiency be predicted from RNA features?

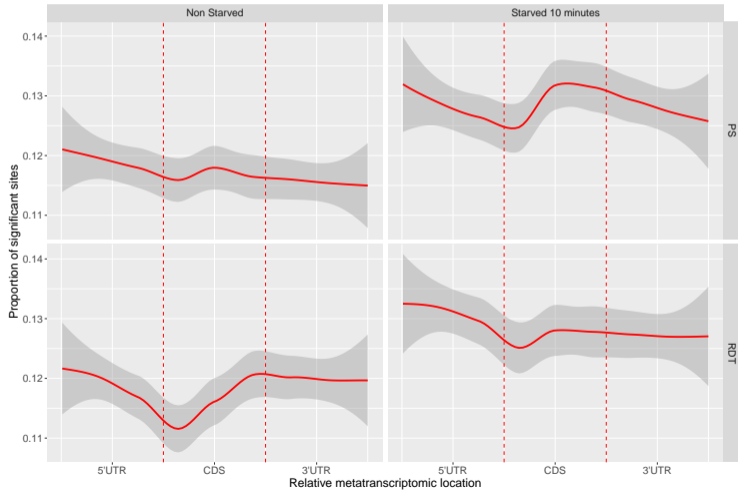
STE (*Stochastic Translation Efficiency*) = a normalised measure of translation efficiency constructed through artificial intelligence analysis.

EXAMPLE OF FEATURE



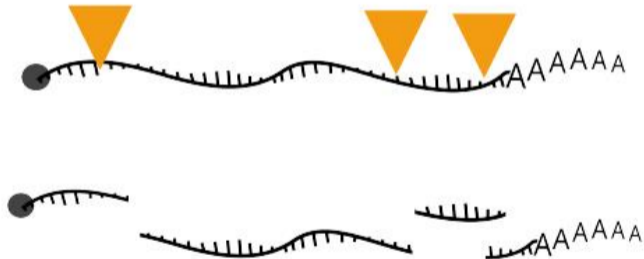
EXAMPLE OF FEATURE

m5C MODIFICATION RATE AS A FUNCTION OF GENE POSITION



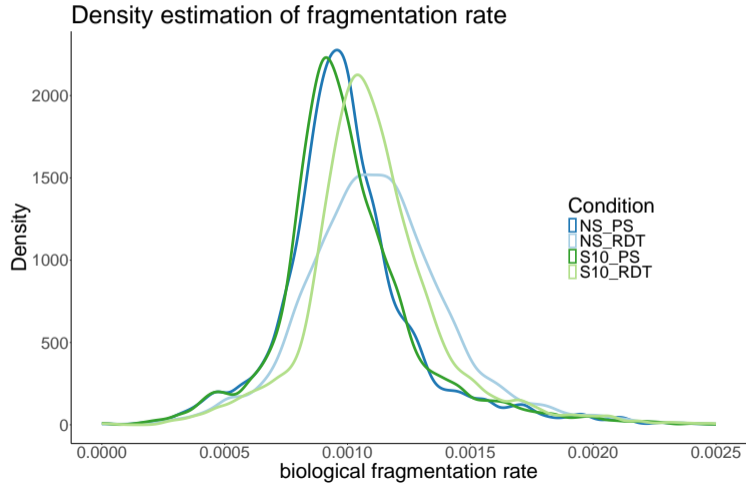
EXAMPLE OF FEATURE

DEGRADATION RATES

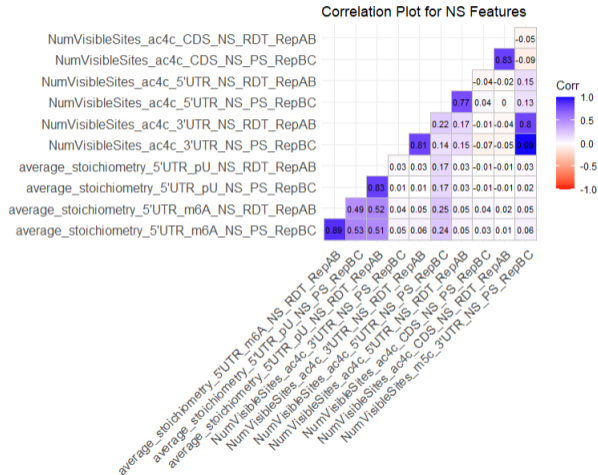


EXAMPLE OF FEATURE

DEGRADATION RATES



ISSUES AND CHALLENGES

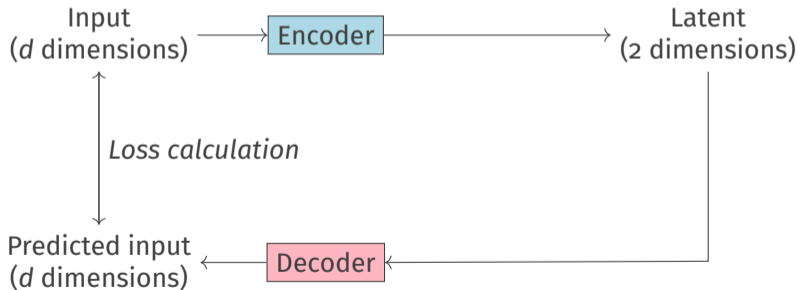


- ~ 450 features
- ~ 6500 genes
- correlated and redundant features

Hypothesis : Not all genes are equally influenced by RNA features in translation efficiency prediction

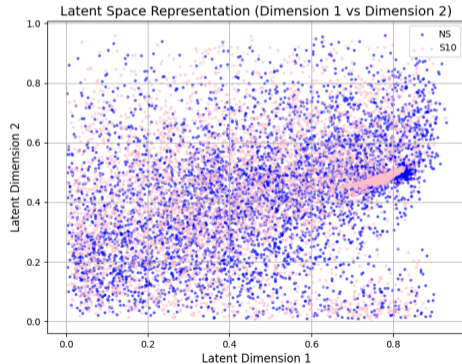
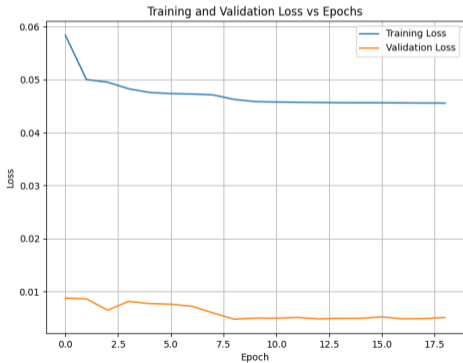
METHODS - DIMENSION REDUCTION

Autoencoder



METHODS - DIMENSION REDUCTION

WORK IN PROGRESS



METHODS - CLUSTERING

WORK IN PROGRESS

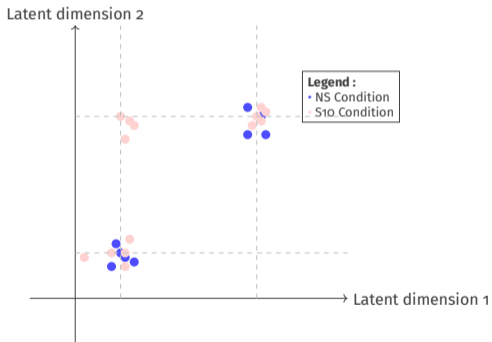


FIGURE – Example of what we want

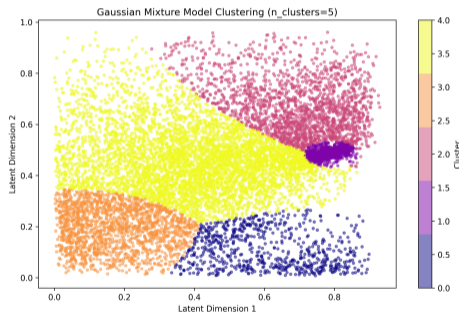
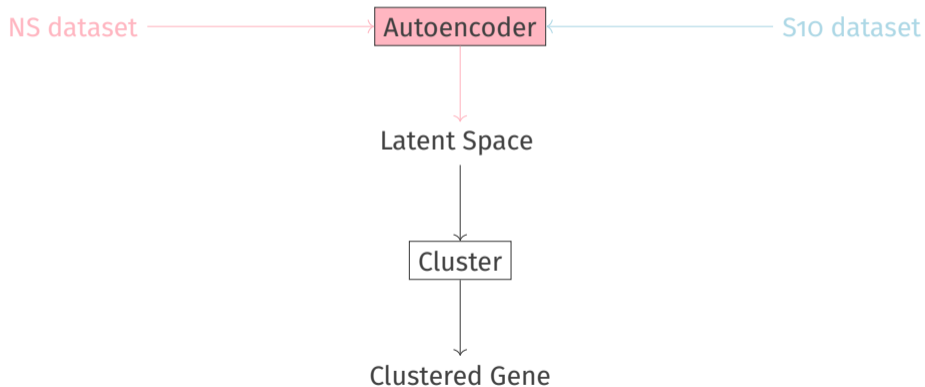
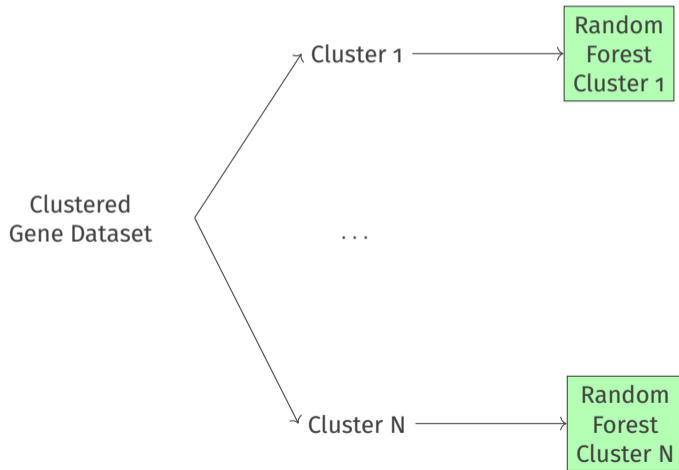


FIGURE – Example of what we get

SUMMARY OF DATA PREPROCESSING



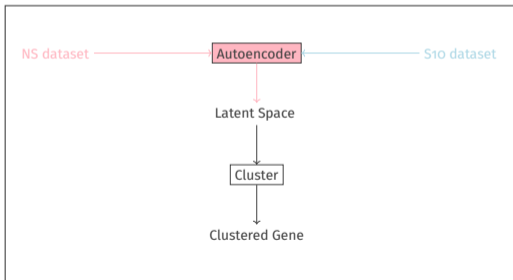
METHODS - RANDOM FOREST



CONCLUSION

WORK IN PROGRESS

1 - DATA PREPROCESSING



2 - ANALYSIS

