

# FUN-PROSE: A Deep Learning Approach to Predict Condition-Specific Gene Expression in Fungi

## Background/Objective

Information on the mRNA levels of all the genes in a genome is critical to defining the overall state of a cell in a given environmental condition. Reconstructing such condition-specific gene expression in fungal genomes is useful because it measures the organism's response to different environmental perturbations, which informs how one can metabolically engineer this system to produce desired chemicals. Most existing deep learning approaches predict the average expression levels of a gene based on its promoter sequence, ignoring its variation across different conditions. Here, we developed FUN-PROSE (FUNgal PRomoter to cONdition-Specific Expression), a deep learning model trained to predict the differential expression of individual genes across various conditions using their promoter sequences and the expression levels of all transcription factors (TFs).

## Approach

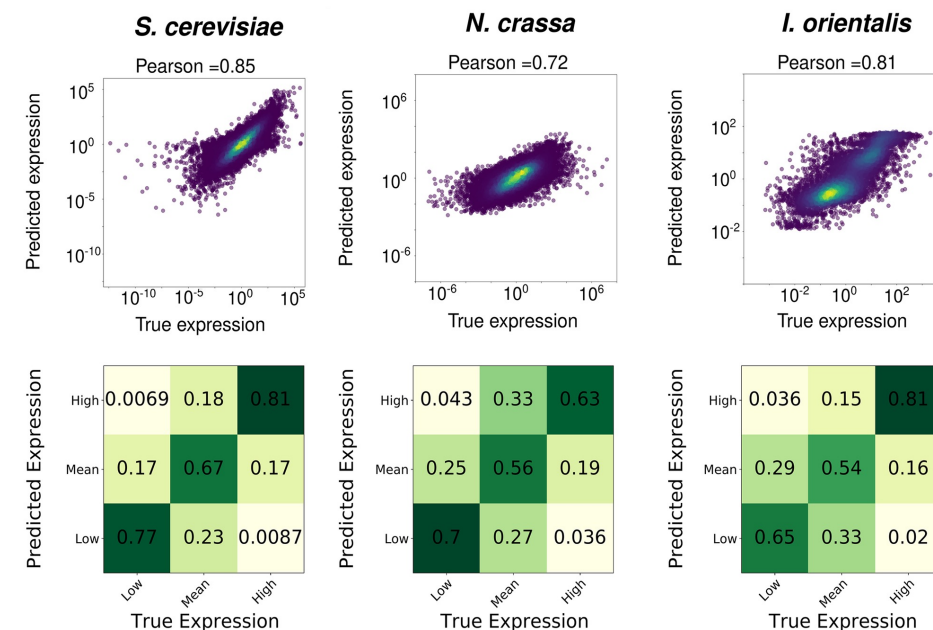
The model was trained and tested using data from three fungal species, *Saccharomyces cerevisiae*, *Neurospora crassa*, and *Issatchenkia orientalis*. The correlation between the predicted and observed condition-specific gene expression was optimized to be as high as 0.85. To understand how the model works, promoter sequence motifs responsible for variable expression of individual genes were extracted. The model was also used to connect individual TFs to their gene targets.

## Results

A sizeable fraction of both sequence motifs and TF-gene interactions learned by the model agreed with previously known biological information, and the rest corresponded to either novel biological facts or indirect correlations.

## Significance/Impacts

FUN-PROSE is a publicly available tool ([github.com/maslov-group/FUN-PROSE](https://github.com/maslov-group/FUN-PROSE)) that can predict metabolic engineering outcomes and gene expression in novel environmental conditions for an organism, an important part of optimizing their use in the production of desired chemicals.



**FUN-PROSE model predicts condition-specific gene expression in fungi.**