



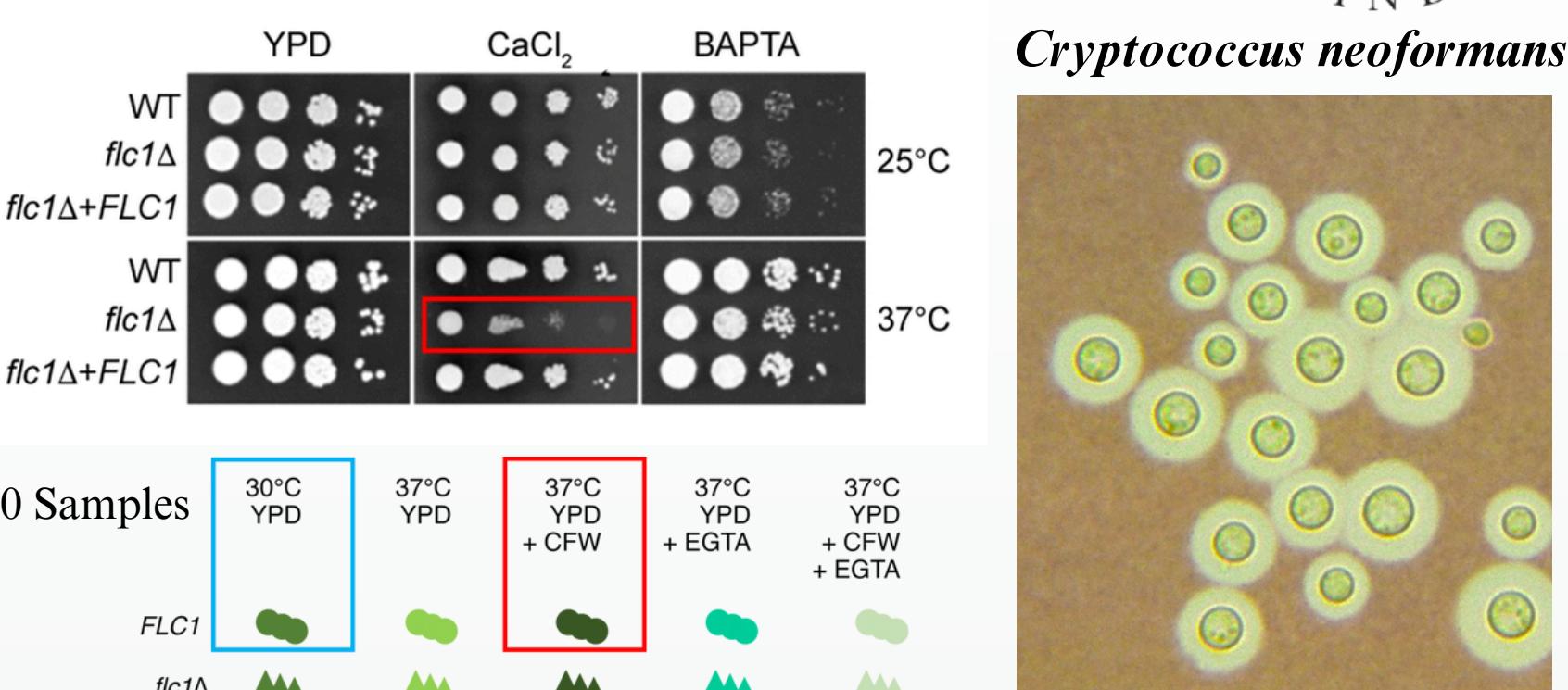
Integrative Transcriptomic and Motif Analysis Reveals Stress-Specific Regulatory Responses in *Cryptococcus neoformans* upon *FLC1* Deletion

Author: Xu Guo Instructors: Tim Cannings and Serveh Sharifi Far
University of Edinburgh, King's Buildings, South Layton Terrace, Edinburgh, EH9 3JL

Introduction

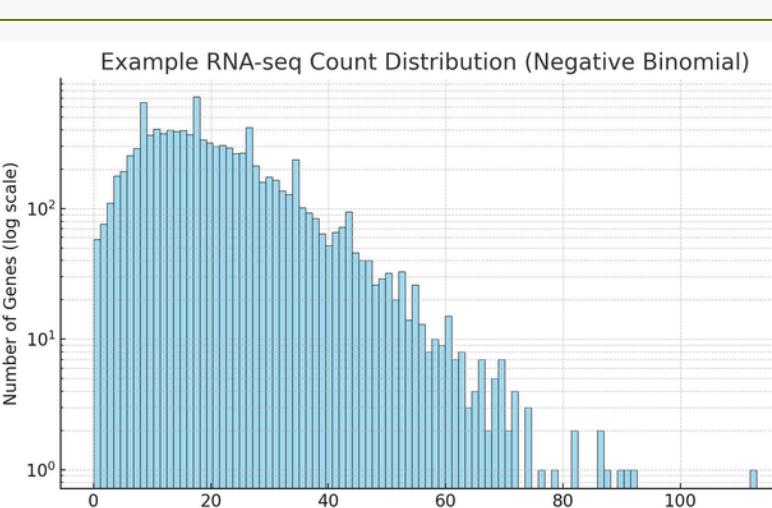
Cryptococcus neoformans is a pathogenic yeast that adapts to environmental stress, especially in the human host. The *FLC1* gene is crucial for calcium stability and cell wall integrity, and its deletion causes growth defects under **high calcium at 37°C**. Researchers use chemical stressors like Calcofluor White (**CFW**) and calcium chelators to study gene expression changes and regulatory mechanisms under stress.

Which genes inhibit the growth of the fungi under condition of 37°C+Δ*FLC1*+CFW?



RNAseq is a high-throughput technique used to measure genome-wide gene expression by sequencing RNA transcripts.

Show a right-skewed, overdispersed distribution with many zeros, often modeled using **Negative Binomial (NB) distribution**.



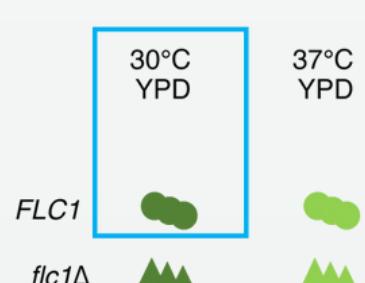
Objective

Discover co-expression profiles across conditions
Use Principal Component Analysis (PCA) to examine differences of global gene expression patterns

Identify differentially expressed genes (DEGs)
Use DESeq2 under a negative binomial framework

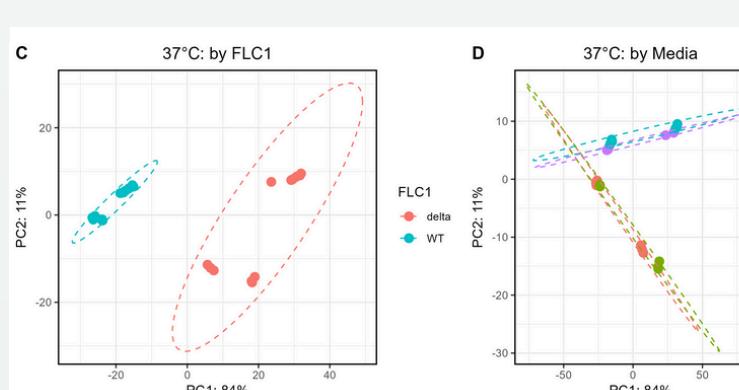
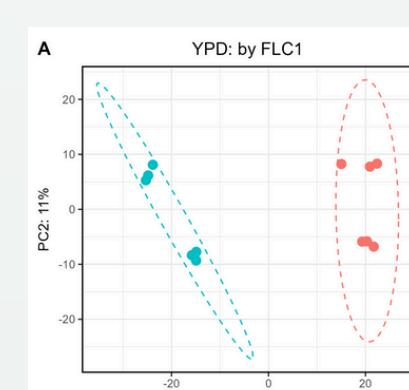
PCA

Design 1: control media
A: Δ*FLC1*
B: Temperature
E: Δ*FLC1*:Temperature



Design 2: control media

C: Δ*FLC1*
D: Media
F: Δ*FLC1*:Media



NB Regression Model

$$K_{ij} \sim NB(\mu_{ij}, \alpha_i), \quad \mu_{ij} = s_j \cdot q_{ij}, \quad \log(q_{ij}) = X_j \beta_i$$

Model 1:

~Temperature + Δ*FLC1* + Temperature:Δ*FLC1*

Model 2:

~Media + Δ*FLC1* + Media:Δ*FLC1*

K: the raw read count

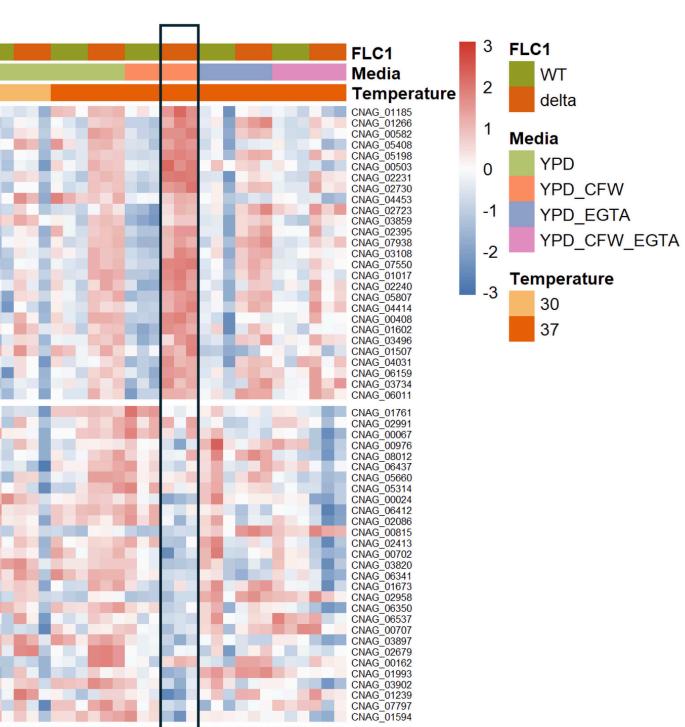
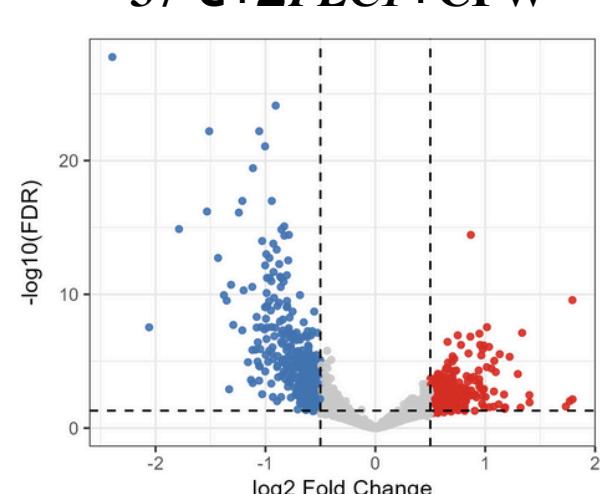
μ: the expected count

α: a gene-specific dispersion

S: the sample-specific size factor

q: the true gene expression

56 genes express differently under the following conditions, specifically: 37°C+Δ*FLC1*+CFW



Motif is a short and recurring DNA sequence which is likely to be a binding site of transcript factors (TFs).

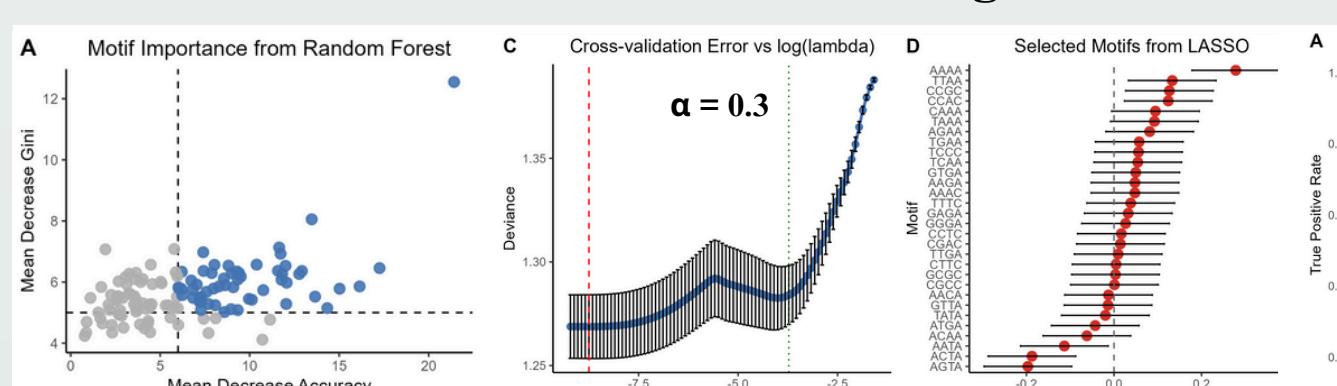
Objective

Determine whether DNA sequence motifs explain stress-induced gene expression

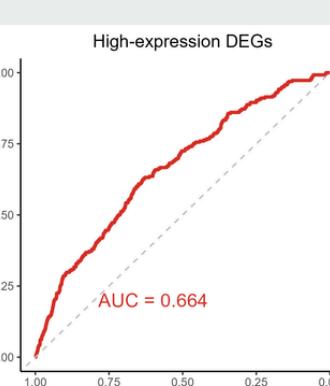
Use Random Forest and Lasso Regression to screen predictors from 4 mer motif-gene matrix

Random Forest

High



ROC



Limitation

Incomplete control experiment (30°C+media). Lack of accuracy for high-expression predictors. Lack of model validation.

Overlook genes involved in combinatorial regulation.

Conclusion

In this study, we found that *FLC1* deletion causes distinct gene responses under **temperature** and **media conditions**, with a unique signature at 37 °C + CFW consistent with the observed growth inhibition. Differential expression and motif analysis further revealed that **specific motifs** are linked to stress-induced regulation, providing new insights into fungal adaptation and pathogenesis.

Low

