2022

Predicting Gene Function of Unknown Yeast ORFs Through Phylogenetic Comparative Analysis

Lewis Barr
lewis.barr@bobcats.gcsu.edu

Predicting gene function of unknown yeast ORFs through phylogenetic comparative analysis

Lewis Barr
Department of Biological and Environmental Science
Georgia College, Milledgeville, GA

Introduction

Yeast (*Saccharomyces cerevisiae*) has been an instrumental model system for a diverse array of research applications for over a century now[4]. In this study, the open reading frames (ORFs) YOR280C, YGR066C, and YLR359W were chosen for analysis. These open reading frames code for the genes FSH3, GID10, and ADE13, respectively. FSH3 is a serine hydrolase and is a known orthologous tumor suppressor gene (OVCA2) in humans[3]. GID10 is the recognition component of the Pro/N-degron pathway[2]. It is orthologous to GID4 in humans. Lastly, ADE13 is an adenylosuccinate lyase involved in purine biosynthesis[1]. ADE13 is orthologous to the human gene ADSL. Though each of the gene functions has been described in literature, little is known about the cellular localization. This study offers a preliminary analysis of the localization by comparing the amino acid sequences of each protein with sequences from species with orthologous proteins. The known localization of these orthologs will be used to predict the localization of the proteins in *S. cerevisiae.*

Objective

- To predict the localization of gene products by comparing sequence similarity of orthologous genes in other species prior to future experiments.

Materials and Methods

- ORFs identified via the Saccharomyces Genome Database (SGD).
- Amino acid sequences of all species collected via NCBI.
- Sequences aligned using Aliview.
- Phylogenetic models were created using EMBL-EBI and iTOL software.
- Known localization of each gene product were compared with unknown ORFs to predict potential localization.

Acknowledgements

I would like to thank Drs. Ellen France, Rich Adams, and Ashok Hegde for their guidance and support throughout this study.

Results

- Collected phylogenic data matches expected mutations across species in both FSH3 and ADE13.
- Large evolutionary distances between species in GID10 comparison led to inconclusive results regarding sequence similarity and potential localization.

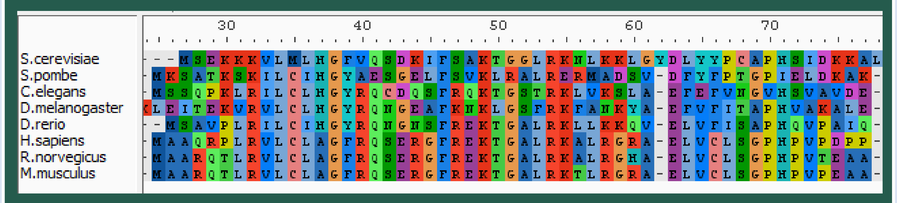

Figure 1. Snapshot of amino acid sequence (FSH3) aligned via Aliview. Sequences data collected from NCBI.

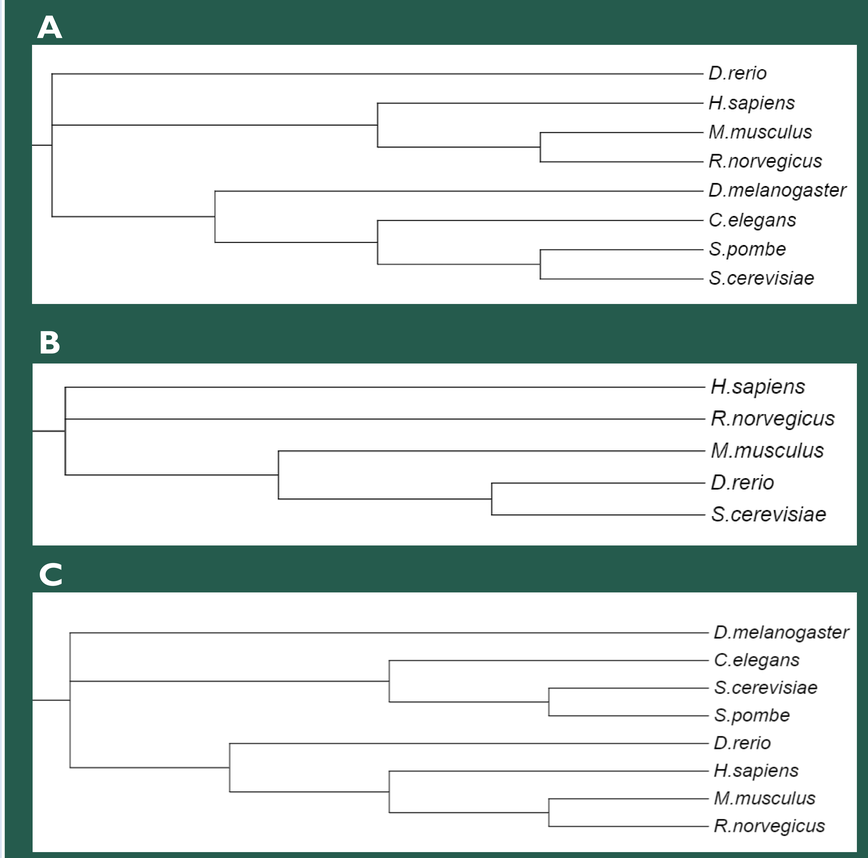

Figure 2. Phylogenies created using EMBL-EBI and iTOL browser software. (A) FSH3 amino acid sequence phylogeny. (B) GID10 amino acids sequence phylogeny. (C) ADE13 amino acid sequence similarity.

Significance

- *S. cerevisiae* has long been an important model system for research. Each gene chosen is orthologous to humans. In humans, deletions or mutation in each gene result defect or ailments.
- Mutations in ADSL result in adenylosuccinate lyase deficiency. OVCA2 is a biomarker of ovarian cancer. Deletion of GID4 is linked to Smith-Magenis syndrome in humans. Identification of protein localization may provide the framework for new experiments.

Conclusions

- Sequence data suggests that FSH3 will be localized in the cytoplasm and nucleus.
- ADE13 is predicted to localize in the cytosol.
- Due to high sequence difference, predictions can not be made on GID10 localization.
- Conclusions from this study should guide future work in identifying the localization of each protein.
- In the future, GFP tagging will be used to identify the localization of FSH3, GID10, and ADE13 products. GFP fluorescence will be compared via colocalization to confirm cellular localization. This study allows us to identify cellular location to more precisely target for colocalization analysis.

References

1. Dorfman BZ (1969) The isolation of adenylosuccinate synthetase mutants in yeast by selection for constitutive behavior in pigmented strains. Genetics 61(2):377-89
2. Melnykov A, et al. (2019) Gid10 as an alternative N-recognin of the Pro/N-degron pathway. Proc Natl Acad Sci U S A 116(32):15914-15923
3. Baxter SM, et al. (2004) Synergistic computational and experimental proteomics approaches for more accurate detection of active serine hydrolases in yeast. Mol Cell Proteomics 3(3):209-25
4. Jiang, Q., L. Lin, and T. Wang. (2008.) A new model for apoptosis research: Yeast. Progress in Biochemistry and Biophysics 35(261):4.