Content Update:

➢ Disease-related Genes
➢ Protein Abundance

ICYGMB; Aug 21, 2019
Disease associations - background

- Collected human-yeast cross-species functional complementation data. 1014 total annotations from P-POD and SGD representing 569 genes available in YeastMine.

- A subset of papers refer to associations between human genes and disease. Connections between the yeast gene, the human homolog and associated diseases were collected using Disease Ontology (DO) terms.

- Identified 639 associations for 209 genes, using three evidence codes: IGI (genetic interaction), ISS (sequence similarity) and/or IMP (mutant phenotype).
Human disease-related genes

Disease Summary

Association details:
- DO term
- Evidence
- Source
- Date assigned
- Reference

Network Diagram:
- Connect genes, orthologs and shared DO terms
- Move slider to adjust network (# of nodes)

Tabbed page disease content

Interactive DO diagram with relationships

DO term annotation table

Disease Ontology term page
Data in YeastMine*

* https://yeastmine.yeastgenome.org/yeastmine/begin.do
Future directions

● Continue curating disease-associated genes from the literature while collecting human-yeast cross-species complementation data.

● Create a computational section of disease-associations using orthologous gene calls from the Alliance using the ortholog prediction tool, DIOPT, and the IEA (Inferred from Electronic Annotation) evidence code.
Brandon Ho, Anastasia Baryshnikova and Grant Brown reanalyzed abundance data from 21 quantitative proteomic analyses (Ho et al. 2018; PMID: 29361465).

- Mode-shift normalization and scaling were used to convert all data to the same abundance unit (molecules per cell).
- Asked to incorporate this data into SGD.

We used to display abundance data from this GFP-tagging study.

Collected data from many other proteome-wide studies, but reported in various units (arbitrary units, molecules/cell).
We added normalized values and metadata (media, visualization and strain background) from both untreated cells and cells treated with various environmental stressors.

- Used several ontologies for the metadata:
  - Experimental Factor Ontology (EFO) - growth media
  - Evidence & Conclusion Ontology (ECO) - visualization
  - Chemical Entities of Biological Interest (ChEBI) and GO - treatments

- Fold change was included when abundance in stressed cells was more than two standard deviations from the untreated average.
Display data and metadata in a table within the experimental data section of the protein tab page.
Authors requested that we include the median value for untreated cells.

- They calculated the median value and median absolute deviation.
- Displayed these on Locus Summary Pages
Data in YeastMine

Templates

Templates are predefined queries, each has a simple form and a description. Click on a template to run it, you can search for templates by keyword and filter them by category.

Filter: abundance  Filter: Proteins  Reset

Actions: Export selected  Options: Show descriptions  Show Tags

You are not logged in. Log in to mark items as favourites 🌟.

- Gene → Protein Abundance
  Retrieve protein abundance for the protein(s) encoded by the specified gene(s).

- Gene → Median Protein Abundance
  Retrieve median protein abundance and median absolute deviation for the protein(s) encoded by the specified gene(s).

▲top
Thank you!

SGD Website
yeastgenome.org

Questions/Comments
sgd-helpdesk@lists.stanford.edu

YouTube Channel
youtube.com/SaccharomycesGenomeDatabase

Download this presentation
bit.ly/ICYGMB-2019