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.

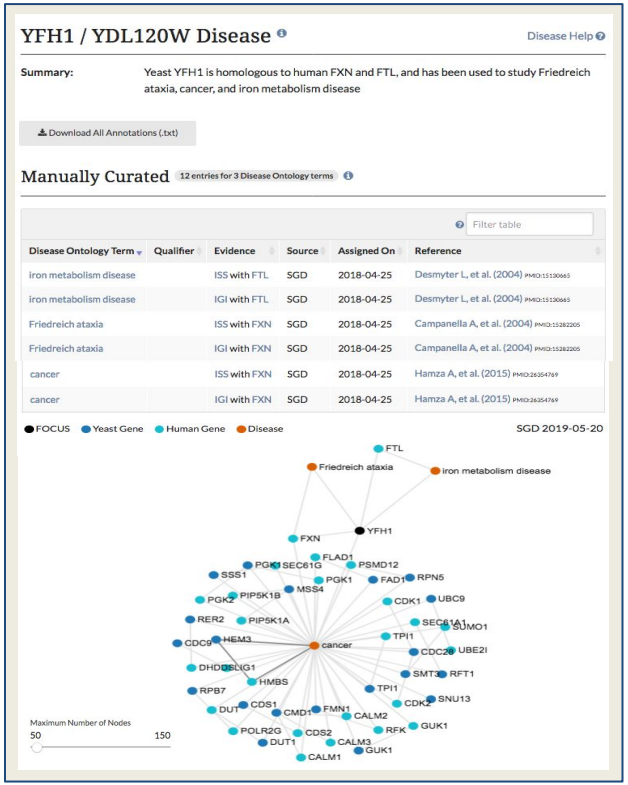
- Identified 639 associations for 209 genes, using three evidence codes: IGI (genetic interaction), ISS (sequence similarity) and/or IMP (mutant phenotype).

Human disease-related genes

Disease Summary

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

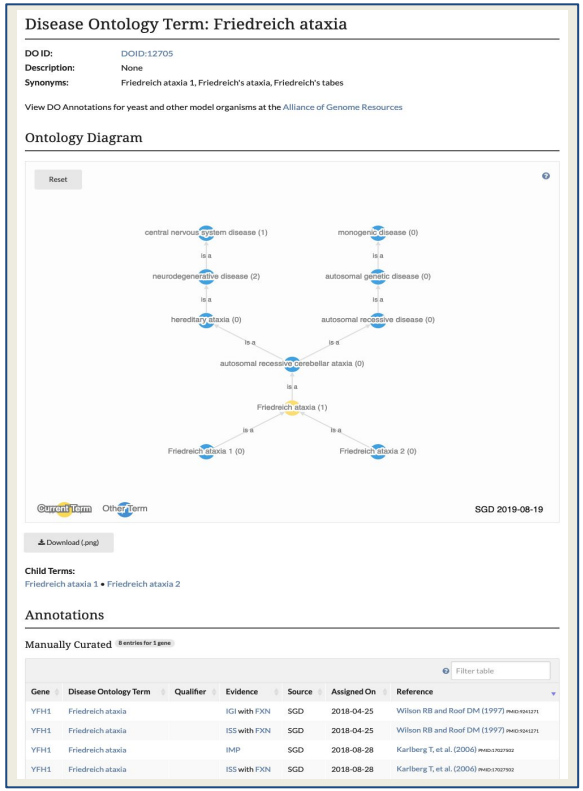
Network Diagram:
- ❑ Connect genes, orthologs and shared DO terms
- ❑ Move slider to adjust network (# of nodes)



Tabbed page disease content

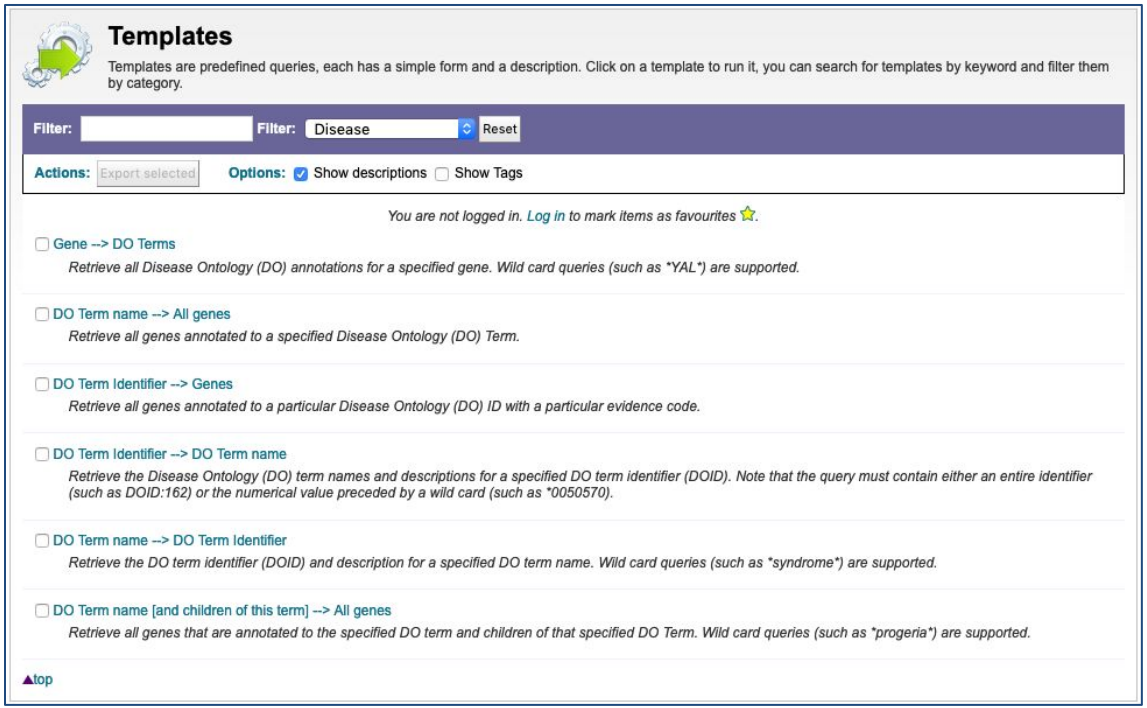
Interactive DO diagram with relationships

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.

Protein abundance - background

Global analysis of protein expression in yeast

Sina Ghaemmaghami[1,2], Won-Ki Huh[1,3], Kiowa Bower[1,2], Russell W. Howson[1,3], Archana Belle[1,3], Noah Dephoure[1,3], Erin K. O'Shea[1,3] & Jonathan S. Weissman[1,2]

[1]Howard Hughes Medical Institute, [2]Departments of Cellular & Molecular Pharmacology and [3]Biochemistry & Biophysics, University of California–San Francisco, San Francisco, California 94143-2240, USA

- 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)

CDC28/YBR160W Protein Information

Summary | Sequence | Locus History | Literature | Gene Ontology | Phenotype | Interactions | Expression | Regulation | Protein | Wiki

Protein Info | Domains/Motifs | Physico-chemical Props

Standard Name	Cdc28p 1, 2
Systematic Name	Ybr160wp
Alias	Cdk1p , Hsl5p , Srm5p
ORF Classification	Verified
Description	Cyclin-dependent kinase (CDK) catalytic subunit; master regulator of mitotic and meiotic cell cycles; alternately associates with G1, S, G2/M phase cyclins, which provide substrate specificity; regulates metabolism, basal transcription, chromosome dynamics, growth and morphogenesis; transcript induction in osmostress involves antisense RNA; human homologs CDK1, CDK2, CDK3 can complement yeast conditional cdc28 mutants; human CDK1, CDK2 can complement yeast cdc28 null mutant (10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 3, 4, 5, 6, 7, 8, 9)
Name Description	Cell Division Cycle 20
Experimental Data	
Molecules/cell	6670 21

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.

Unification of Protein Abundance Datasets Yields a Quantitative *Saccharomyces cerevisiae* Proteome

Brandon Ho,[1] Anastasia Baryshnikova,[2,3] and Grant W. Brown[1,4,*]
[1]Department of Biochemistry and Donnelly Center, University of Toronto, Toronto, ON M5S 1A8, Canada
[2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
[3]Present address: Calico Life Sciences, South San Francisco, CA 94080, USA
[4]Lead Contact

Abundance data integration

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.

Protein tab page

Display data and metadata in a table within the experimental data section of the protein tab page

Experimental Data

Protein Half Life

Filter table

Experiment	Result	Reference
half-life	10.4 hr	Christiano R, et al. (2014)

Download (.txt)

Protein Abundance

Filter table

Abundance (molecules/cell)	Media	Treatment	Treatment time	Fold Change	Visualization	Strain background	Original Reference	Reference
4319	SD	untreated			confocal microscopy evidence	S288C	Breker M, et al. (2013)	Ho B, et al. (2018)
5182	SD	2 mM 1,4-dithiothreitol	2 hr		confocal microscopy evidence	S288C	Breker M, et al. (2013)	Ho B, et al. (2018)
4540	SD	1 mM hydrogen peroxide	1 hr		confocal microscopy evidence	S288C	Breker M, et al. (2013)	Ho B, et al. (2018)
4489	SD minus nitrogen	cellular response to nitrogen starvation	15 hr		confocal microscopy evidence	S288C	Breker M, et al. (2013)	Ho B, et al. (2018)
5533	SD	untreated			confocal microscopy evidence	S288C	Chong YT, et al. (2015)	Ho B, et al. (2018)

Showing 1 to 5 of 30 entries 10 records per page

« 1 2 3 4 5 6 »

Locus Summary 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

CDC28 / YBR160W Overview

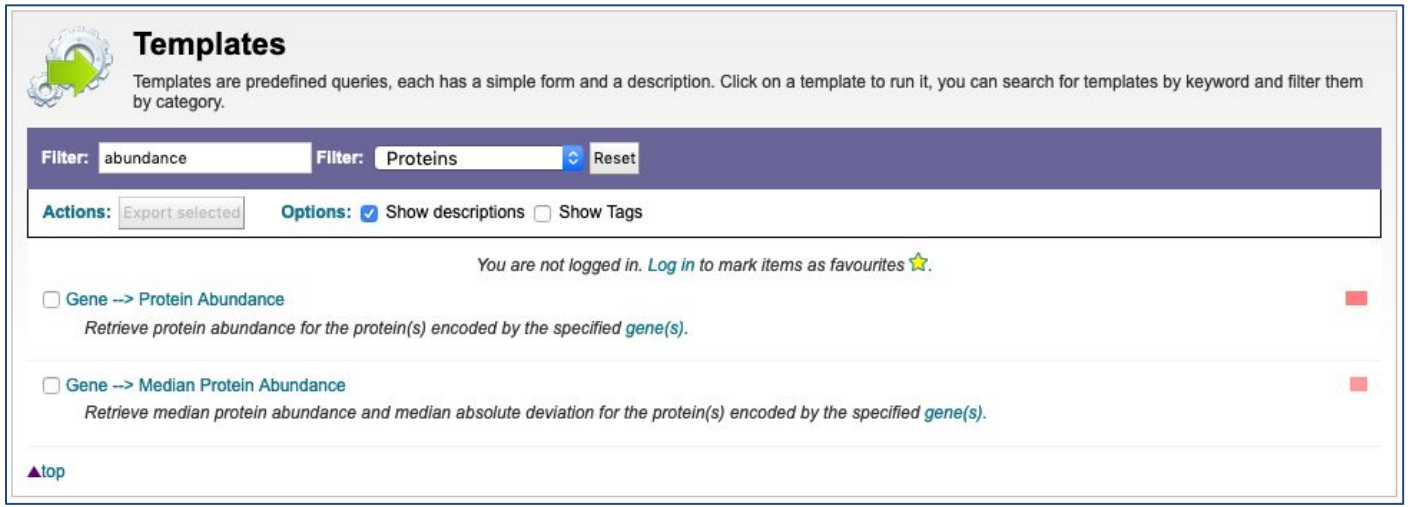
Standard Name:	CDC28 [1 2]
Systematic Name:	YBR160W
SGD ID:	SGD:S000000364
Aliases:	CDK1 , HSL5 , SRM5 [63]
Feature Type:	ORF , Verified
Description:	Cyclin-dependent kinase (CDK) catalytic subunit; master regulator of mitotic and meiotic cell cycles; alternately associates with G1, S, G2/M phase cyclins, which provide substrate specificity; regulates metabolism, basal transcription, chromosome dynamics, growth and morphogenesis; transcript induction in osmostress involves antisense RNA; human homologs CDK1, CDK2, CDK3 can complement yeast conditional cdc28 mutants; human CDK1, CDK2 can complement yeast cdc28 null mutant [3 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20]
Name Description:	Cell Division Cycle [4]
Comparative Info:	Integrated model organism details available at the Alliance of Genome Resources website

Protein

Protein Details

Length (a.a.):	298
Mol. Weight (Da):	34064.8
Isoelectric Point:	8.55
Median Abundance (molecules/cell):	6670 +/- 1539

Data in YeastMine



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