

Using SGD GO Slim Mapper and Interaction Data to Predict Gene Function

The Gene Ontology (GO) is structured in a hierarchy, such that granular terms (“perinuclear space”) are connected and further down the hierarchy than their related broader terms (“nucleus”). However, for many purposes, such as reporting the upregulated cellular functions of a transcriptomics experiment, is very useful to focus on the broad, high-level part of the GO. For example, if you were interested in which of your upregulated genes are involved in DNA replication, it would be useful to map genes that have been annotated to specific terms (e.g. “synthesis of RNA primer involved in nuclear cell cycle DNA replication”) to more general terms (e.g. “DNA replication”).

The **Gene Ontology (GO) Slim Mapper** at SGD maps granular GO annotations of a group of genes to more general terms and/or bins them into broad categories, ie. “**GO Slim**” terms. Using GO Slim Mapper, predict what biological processes an uncharacterized gene may be involved in based on its genetic interactions.

- From the SGD home page (www.yeastgenome.org), go to the Locus Summary page for the uncharacterized gene **YLR287C**.
- Select **Interactions** tab. Here, we are interested in finding genes that have a genetic interaction with YLR287C, as the function of these genes may provide hints about the function of YLR287C.
- Scroll to the **Annotations** table and search for “synthetic”. This will filter the table for genes that, when knocked in combination with YLR287C, elicit some sort of synthetic growth defect, haploinsufficiency, lethality, etc. These harsh phenotypes may suggest clues about related functions in YLR287C.

Annotations ⓘ 213 entries for 202 genes							
<input type="text" value="synthetic"/>							
Interactor	Type	Assay	Annotation	Action	Modification	Phenotype	Reference
BUB3	Genetic	Synthetic Lethality	high-throughput	Hit		inviable Mutant Type: unspecified	Daniel JA, et al. (2006) PMID:16157669
CCS1	Genetic	Synthetic Growth Defect	high-throughput	Hit		vegetative growth: decreased Mutant Type: unspecified	Pan X, et al. (2006) PMID:16487579
CSM3	Genetic	Synthetic Growth Defect	high-throughput	Hit		vegetative growth: decreased Mutant Type: unspecified	Pan X, et al. (2006) PMID:16487579
CTF4	Genetic	Synthetic Growth Defect	high-throughput	Hit		vegetative growth: decreased Mutant Type: unspecified	Pan X, et al. (2006) PMID:16487579

- Find and click on the **Analyze** button at the bottom of the Annotation table. This will import the table you filtered to a page where you can send the genes to other SGD tools.
- On the next page that lists the YLR287C interactors, select **GO Slim Mapper**.

Tools	
GO Term Finder Find common GO annotations between genes.	GO Slim Mapper Sort genes into broad categories.
SPELL View expression data.	YeastMine Conduct advanced analysis.

Genes	
<input type="text" value="Filter table"/>	
Gene Name	Description
BUB3	Kinetochore checkpoint WD40 repeat protein; localizes to kinetochores during prophase and metaphase, delays anaphase in the presence of unattached kinetochores; forms complexes with Mad1p-Bub1p and with Cdc20p, binds Mad2p and Mad3p; functions at kinetochore to activate APC/C-Cdc20p for normal mitotic progression

- The GO Slim Mapper has three steps (plus one optional step) in which you can specify your query. In Step 1, the input box is preloaded with the list of genes you imported.
- In Step 2, choose a **GO Set** by selecting **Yeast GO-Slim: Process** from the pull-down.
- In Step 3, highlight **SELECT ALL Terms from Yeast GO-Slim: Process**.

Step 1: Choose Gene/ORF names	
Either Enter Gene/ORF names (separated by a return or a space) YOR026W YMR038C YMR048W YPR135W YOR080W YLR233C YGL086W YJL030W YDR432W YOR209C YDR510W YNL273W	OR Upload a file of Gene/ORF names: (.txt or .tab format) <input type="button" value="Choose File"/> No file chosen
Step 2: Choose GO SLIM Terms(s) by choosing a GO Set Terms from the selected GO Set will be automatically entered in the box in Step 3 Yeast GO-Slim: Process	
Step 3: Refine your list of GO Slim Terms <div> <div> SELECT ALL Terms from Yeast GO-Slim: Process DNA recombination DNA repair DNA replication </div> <div> Choose at least one term from the list • Select or unselect multiple options for GO terms by pressing the Control (PC) or Command (Mac) key while clicking • For information about a particular GO Term and its definition, type the GO Term in the Search box at the top of the page </div> </div>	
Search Reset This will map annotations made to your input list of genes from the Manually curated and High-throughput annotation methods. Go to Step 4 below for filtering options.	
Optional Step 4: Select Annotation Method(s) Default maps Manually curated and High-throughput Annotation Methods	
• Manually curated: <input checked="" type="radio"/> yes <input type="radio"/> no • High-throughput: <input checked="" type="radio"/> yes <input type="radio"/> no	• Manually curated - includes annotations based on published experiments or analyses or curatorial statements that are assigned by SGD curators. • High-throughput - includes annotations made from published experiments performed on a high-throughput or genome-wide basis.
Search Reset	

- Click the **Search** button to use the default settings or go further down to customize your query.
 - Optional Step 4 allows excluding manually curated or high-throughput annotations; leave the **yes** options checked for both.

Optional Step 4: Select Annotation Method(s) Default maps Manually curated and High-throughput Annotation Methods	
<ul style="list-style-type: none"> • Manually curated: <input checked="" type="radio"/> yes <input type="radio"/> no • High-throughput: <input checked="" type="radio"/> yes <input type="radio"/> no 	<ul style="list-style-type: none"> • Manually curated - includes annotations based on published experiments or analyses or curatorial statements that are assigned by SGD curators. • High-throughput - includes annotations made from published experiments performed on a high-throughput or genome-wide basis.
<input type="button" value="Search"/> <input type="button" value="Reset"/>	

- Results appear in a table with four columns:
 - a. GO Slim terms picked by GO Slim Mapper
 - b. Cluster frequency, the number and percentage of genes in your list annotated to each term
 - c. Genome frequency, the number and percentage of all genes in the genome annotated to each term
 - d. Genes from your list that are annotated to that term, hyperlinked to their Locus Summary pages. You can also download the results in a tab-delimited file.
- Based on the results, what biological processes might YLR287C be involved in?