

# Yeast ORFan Gene Project: Module 7 Guide

## Gene Deletion Phenotypes

Your ORFan gene was most likely included in the yeast deletion collection and after creation of this collection some basic questions were asked of all of these strains. We will now mine this data to determine if there is any information that might aid in our determination of the function of your protein.

## ***Saccharomyces cerevisiae* Morphological Database**

The *Saccharomyces Cerevisiae* Morphological Database (SCMD) is a collection of micrographs of budding yeast mutants. Micrographs of mutants with altered cell morphology were taken by the Ohya Group, University of Tokyo, from a set of the haploid MATa deleted strains obtained from EUROSCARF. From the micrographs, disruptant cells were automatically extracted by a novel cell-image processing software developed by the Morishita Group, University of Tokyo. The abstract from the PNAS paper published by this group about their work: "One of the most powerful techniques for attributing functions to genes in uni- and multicellular organisms is comprehensive analysis of mutant traits. In this study, systematic and quantitative analyses of mutant traits are achieved in the budding yeast *Saccharomyces cerevisiae* by investigating morphological phenotypes. Analysis of fluorescent microscopic images of triple-stained cells makes it possible to treat morphological variations as quantitative traits. Deletion of nearly half of the yeast genes not essential for growth affects these morphological traits. Similar morphological phenotypes are caused by deletions of functionally related genes, enabling a functional assignment of a locus to a specific cellular pathway. The high-dimensional phenotypic analysis of defined yeast mutant strains provides another step toward attributing gene function to all of the genes in the yeast genome."\*

**Navigate to SCMD at** <http://yeast.gi.k.u-tokyo.ac.jp/datamine/> and enter your gene name into the Keywords search box and submit the search.

If strains with your gene mutated have been analyzed you will be presented with a page with your gene's name hyperlinked – click on the hyperlinked gene name to be directed to the data page.

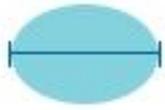
Once on the data page you will be presented with 3 types of data all centered on the average shapes of cells that are mutant for your gene of interest. Specifically, you will view data about the average shape of cell grouped by bud size, nucleus location and actin distribution. On the following pages there are diagrams that outline how each of the parameters in the data tables are calculated as well as a page of the average numbers for a wildtype cell for comparison of your data.

Copy the data tables for cells mutant for your gene into your **Module 7 Worksheet**, make comparisons to the wildtype data found below and determine if any parameters appear to be distinctly different. [Using the Snipping Tool/Grab to copy data tables as an image might be easiest]. Discuss any differences you observe in the Comments section of your **Module 7 Worksheet**.

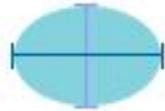
\* Information regarding the functionality of each software/algorithm and operational information was taken directly from the websites and included here for instructional purposes only.

# Yeast ORFan Gene Project: Module 7 Guide

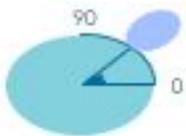
## Cell Shape Parameters



Long axis length



Roundness := long axis length / short axis length



Bud neck position angle (0-90)



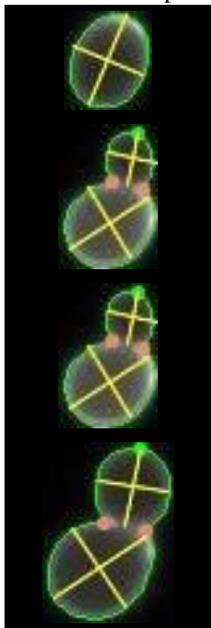
Bud growth direction angle (0-90)

In SCMD, the unit of displayed axis lengths is a *pixel* (= 0.129512 micrometer) on the images, and the unit of areas of cells is its square, pixel x pixel (= 0.016773358).

This value shows how well an ellipse fit to the outline of the cell. The smaller the better. Cell walls of some disruptants are more rectangular than elliptical. The *fitness* is useful to distinguish such shapes.

Fitness :

## Bud Size Groups

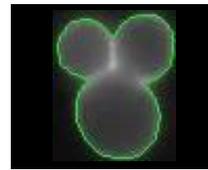


no No bud

small small bud cell  
( $0 < \text{cell size ratio}^{*1} < 0.5$ )

medium medium bud cell  
( $0.5 \leq \text{cell size ratio}^{*1} < 0.7$ )

large large bud cell  
( $0.7 \leq \text{cell size ratio}^{*1}$ )



complex complexed cells

## Nucleus Location Groups



A A single nucleus (unbudded cell)

A1 A single nucleus in mother cell (budded cell)

B A single nucleus at bud neck (budded cell). The yellow point shows the center of gravity of the elongated nucleus.

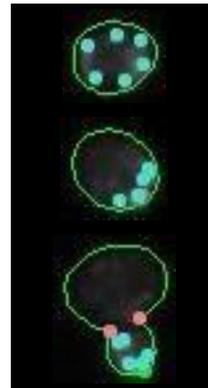
C Nuclei in both mother and daughter cell (budded cell)

D Multi-nucleated mother cell

E No nucleus is detected

F A single nucleus in daughter cell (budded cell)

## Actin Localization Groups



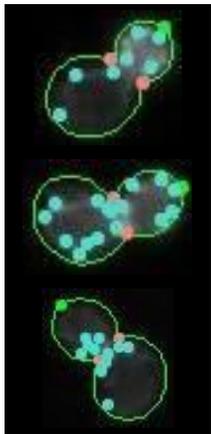
A Delocalized (unbudded cell)

B Localized at bud size (unbudded cell)

api Hyperlocalized at bud tip (budded cell)

\* Information regarding the functionality of each software/algorithm and operational information was taken directly from the websites and included here for instructional purposes only.

# Yeast ORFan Gene Project: Module 7 Guide



- iso Localized at bud (budded cell)
- E Delocalized (budded cell)
- F Localized at bud neck (budded cell)

## WILDTYPE CHARACTERISTIC SHAPES

	no	small	medium	large
<b>Bud Size</b>				
Percentage (%)	30.8	19.6	23.6	26.0
Long Axis	34.366	37.483	35.440	34.132
Roundness	1.174	1.373	1.217	1.166
Bud Neck Position		36.027	31.100	36.455
Bud Growth Direction		62.341	30.458	46.863
Daughter / Mother	0.000	0.131	0.407	0.636
[group by sheet]	[datasheet]	[datasheet]	[datasheet]	[datasheet]
	[teardrop view]	[teardrop view]	[teardrop view]	[teardrop view]
<b>Search Similar Shape</b>	[search]	[search]	[search]	[search]

	A	A1	B	C
<b>Nucleus Location</b>				
Percentage (%)	30.6	41.3	6.4	21.2
Long Axis	34.366	35.440	34.015	34.132
Roundness	1.174	1.217	1.307	1.166
Bud Neck Position		31.100	28.301	36.455
Bud Growth Direction		30.458	57.995	46.863
Daughter / Mother	0.000	0.407	0.484	0.636
[group by sheet]	[datasheet]	[datasheet]	[datasheet]	[datasheet]
	[teardrop view]	[teardrop view]	[teardrop view]	[teardrop view]
<b>Search Similar Shape</b>	[search]	[search]	[search]	[search]

### Average Shapes Grouped by Nucleus Location

	D	E	F
<b>Nucleus Location</b>			
Percentage (%)	0.3	0.3	0.0
Long Axis	37.000	37.696	
Roundness	1.121	1.201	
Bud Neck Position		55.115	
Bud Growth Direction		74.107	
Daughter / Mother	0.000	0.612	
[group by sheet]	[datasheet]	[datasheet]	[datasheet]
	[teardrop view]	[teardrop view]	[teardrop view]
<b>Search Similar Shape</b>	[search]	[search]	[search]

### Average Shapes Grouped by (peculiar) Nucleus Location

	A	B	apl	iso	E	F
<b>Actin Distribution</b>						
Percentage (%)	9.1	21.5	23.1	26.6	9.4	10.2
Long Axis	27.203	34.366	37.483	35.440	35.903	34.132
Roundness	1.193	1.174	1.373	1.217	1.130	1.166
Bud Neck Position			36.027	31.100	8.952	36.455
Bud Growth Direction			62.341	30.458	0.346	46.863
Daughter / Mother	0.000	0.000	0.131	0.407	0.284	0.636
[group by sheet]	[datasheet]	[datasheet]	[datasheet]	[datasheet]	[datasheet]	[datasheet]
	[teardrop view]					
<b>Search Similar Shape</b>	[search]	[search]	[search]	[search]	[search]	[search]

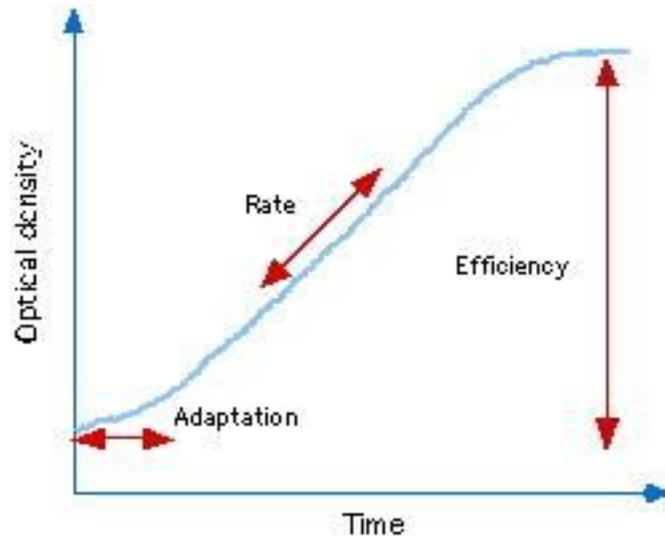
### Average Shapes Grouped by Actin Distribution

\* Information regarding the functionality of each software/algorithm and operational information was taken directly from the websites and included here for instructional purposes only.

# Yeast ORFan Gene Project: Module 7 Guide

## PROPHECY : PROFiling of PHENotypic Characteristics in Yeast

“PROPHECY provides quantitative information about phenotypes for the complete collection of deletion strains in yeast (*Saccharomyces cerevisiae*).” \*



“PROPHECY evaluates the phenotype of a deletion strain on the basis of growth behavior during micro-cultivation. PROPHECY quantifies growth aberrations by estimating the **rate of growth**, the **efficiency of growth** and the **adaptation time**.”\*

Navigate to PROPHECY at <http://prophecy.lundberg.gu.se/Main.aspx> and enter your standard gene name in the Quick Search box and submit the Search. (Note – for some reason with this site you may need to click the Search button twice to get the website to proceed.)

The data for cells harboring a deletion mutation of your gene will appear on a new page with graphical information about the growth of these strains in environments of high salt (0.85M NaCl), oxidative stress [cells ability to tolerate production of destructive reactive oxygen species] (Diamide 1.4mM) and (Paraquat 200 µg/ml), reductive stress [ability of cells to maintain redox homeostasis] (DTT 1.6mM)\*. As well as an overall Gene-by-environment phenotype summary box which highlights any significant changes. Click on each individual graph to open the data in a larger window and analyze how cells mutant for your gene, grow relative to their wildtype counterparts.

Copy the large version of each graph and incorporate them into the given space in your **Module 7 Worksheet**. Also report the overall summary from the Gene-by-environment phenotype summary box on the first page and comment on any changes you noticed in the graphs that contributed to the summary provided.

## Yeast Fitness Database

\* Information regarding the functionality of each software/algorithm and operational information was taken directly from the websites and included here for instructional purposes only.

## Yeast ORFan Gene Project: Module 7 Guide

This site provides a graphical interface to the supplementary data presented in two publications: "The Chemical Genomic Portrait of Yeast: Uncovering a phenotype for all genes" (Hillenmeyer et al., Science 2008) and "Chemical genomic profiles predict gene function and mechanism of drug action" (Hillenmeyer et al., Genome Biol. 2010). This data should therefore provide us with information about the phenotype of cells with your gene mutated when exposed to various drugs and then based on these data the program will report what other gene mutations show similar sensitivities to the same drugs to which your mutated strains is sensitive.

**Navigate to the Yeast Fitness Database at <http://chemogenomics.med.utoronto.ca/fitdb/fitdb2.cgi>** and enter your standard gene name in the Search box and submit the Search.

Sensitivity data can be informative based on how and what the drug is known to target. Scroll down to the Homozygous knockout section, which shows the most impactful drug effects on cellular fitness when the cell lacks all copies of your gene of interest. Identify the top 3 drugs with significant p-values (<0.05) and perform an Internet search to determine how these drug work in a cell. (You need to separately look up the mechanism of action for the drugs showing significant fitness defects because the list of tested drugs is too long to provide this information in this document). In your **Module 7 Worksheet** record the drug, the Fitness Defect (z-score), the p-value and the mechanism of action of the drug (include the citation of where you found this information).

In Section 2: Co-fitness interactions of the Yeast Fitness Database an algorithm is used to identify if cells with other mutated genes show similar sensitivities to the drugs that cells mutated for your gene of interest do, as this can imply a similarity in function. Look up the functions of up to all of the 10 given genes and record that information in your **Module 7 Worksheet**, including taking the time to correlate the information and generate a hypothesis. [To look up the genes and their known functions navigate to SGD and type the gene names into the search box, when the gene page loads copy the Description portion into your worksheet.]

Not only are the names of these genes provided for you to investigate their known functions through SGD – but they have also been classified by Gene Ontology groupings (to the right of the homozygous co-fitness interaction diagram). [The Gene Ontology (GO) project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. The project provides **a controlled vocabulary of terms** for describing gene product characteristics and gene product annotation data from GO Consortium members, as well as tools to access and process this data.\*] All proteins with known functions are grouped into all of the categories for which their function applies with respect to Molecular Function, Biological Process and Cellular Component. (For instance, if you wanted to classify information about a Strawberry you could put it into groups of Color: Red, Type: Berry, Seeds: Outside – etc. and then look for other types of fruit that are red, or berries or have seeds on the outside as an indicator of potential relatedness.) Once again this data can be useful because it may indicate that your gene/protein will have a similar function and could be assigned to the groups/ontologies that are identified. Record any Enriched GO terms of interactors that are indicated on your Yeast Fitness Database page in your **Module 7 Worksheet**. If you do not know what any of the terms in the GO annotation mean look them up individually and also together in their full GO term to decipher their functional implication.

\* Information regarding the functionality of each software/algorithm and operational information was taken directly from the websites and included here for instructional purposes only.