# Yeast Biochemical Pathways Tool

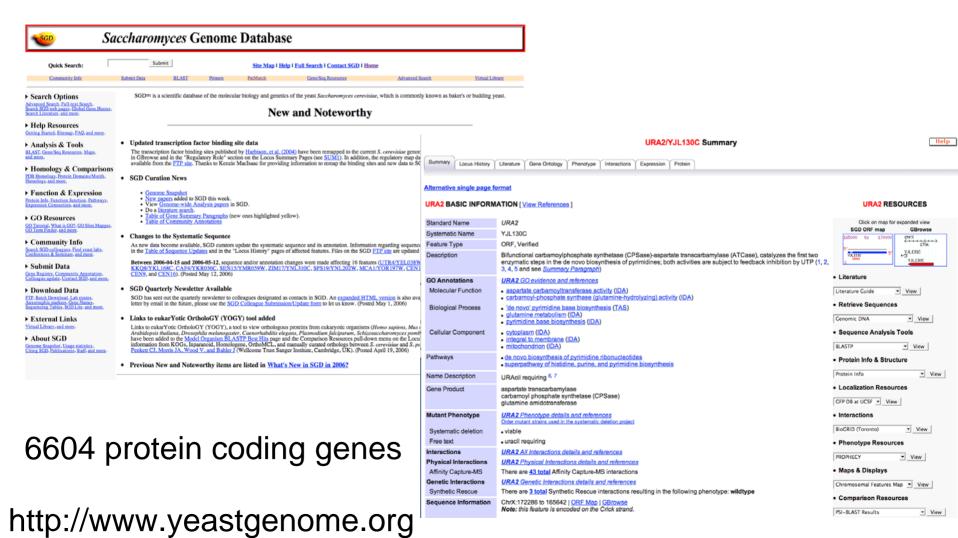
Rama Balakrishnan SGD Stanford University

# About S. cerevisiae

- Simplest eukaryotic unicellular organism
- Its power:
  - Model organism to study genetics, cellular processes
  - Several industrial applications

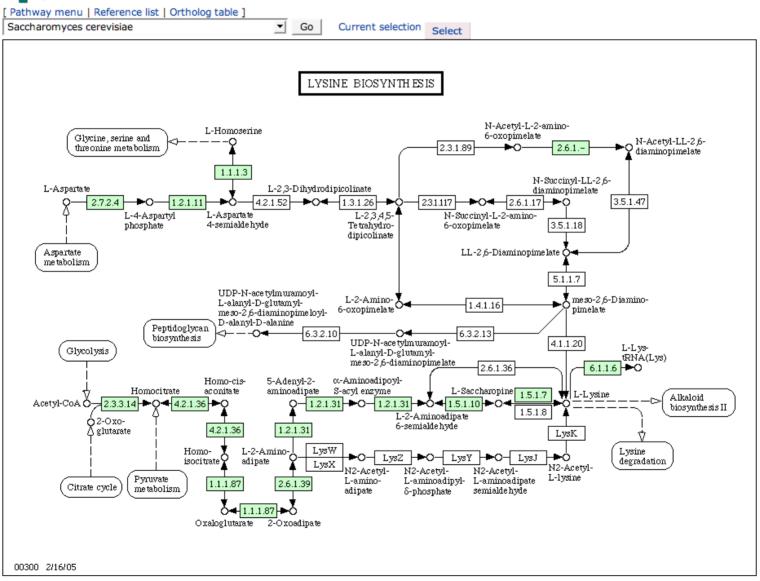
### **About SGD**

Online database for the genomic sequence, genetics and molecular biology of *Saccharomyces cerevisiae*, commonly known as baker's or budding yeast.



# Lysine Biosynthesis from KEGG





# Building Yeast Biochemical Pathways

- Used GO Function ontology annotations
  - EC2GO mapping
- Gene product or Description field information
  - Matches 'ase', manual mapping
- Initial build done with 731 genes
- We have 6604 protein coding genes in the database

# Summary

Total number of genes: 731

EC # matches: 635

Function name matches: 17

(no E.C. available)

Failed matches: 79

(many turned out to be protein kinases, tRNA modification, or egosterol biosynthesis genes)

Pathways Predicted: 125

Reactions mapped: 652

# I nitial Cleanup after Build

- Resolve ambiguous EC numbers
- Fill in missing reactions
- Delete pathways that don't occur in yeast
- Add pathways unique to yeast
- Contribute new pathways to MetaCyc

### Curation

- Types of pathway information captured:
  - References
  - Summary paragraph
  - Evidence code
  - Reactions; EC numbers
  - Genes and gene product references
- Compound IDs from KEGG, CheBI, PubChem

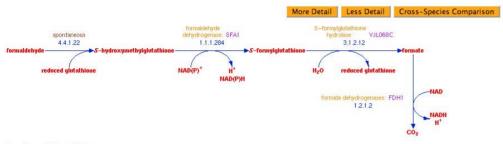
Current Statistics: 154 pathways, 460 genes

Available for download from our FTP site

# **Curated Pathway**

S. cerevisiae Pathway: formaldehyde oxidation II (glutathione-dependent)





Locations of Mapped Genes:



Synonyms: formaldehyde oxidation II (GSH-dependent)

Superclasses: Degradation/Utilization/Assimilation -> C1 Compounds

#### Comment:

Formaldehyde is formed by oxidative demethylation reactions in many plants and methylotrophic organisms, but Saccharomyces cerevisiae is a nonmethylotrophic yeast and cannot metabolize methanol to formaldehyde. However, S. cerevisiae is exposed to exogenous formaldehyde from plant material or in polluted air and water. Concentrations of formaldehyde of 1mM or higher are cytostatic or cytotoxic to haploid wild-type cells.

Any free formaldehyde in vivo spontaneously reacts with glutathione to form S-hydroxymethylglutathione [ Achkor03 ], [ Grey96 ], [ Degrassi99 ]. The level of enzymes involved in the degradation of formaldehyde, such as Sfa1p and Yjl068p, determine the level of formaldehyde toxicity, and cells overproducing Sfa1p are resistant to formaldehyde and null mutants in either sfa1 or yjl068c are hypersensitive to formaldehyde. Sfa1p is induced in response to chemical stresses [ Wehner93 ], [ Grey96 ], [ Degrassi99 ], [ Wehner93a ], [ GompelKlei89 ], [ Jelinsky99 ], [ Schaus01 ].

Formate dehydrogenase is encoded by FDH1/YOR388C and FDH2. In some strain backgrounds of S. cerevisiae, FDH2 is encoded by a continuous open reading frame comprised of YPL275W and YPL276W. However, in the systematic sequence of S288C, FDH2 is represented by these two separate open reading frames due to an in frame stop codon [ Overkamp02 ] .

#### References

Achkor 3: Achkor H, Diaz M, Fernandez MR, Biosca JA, Pares X, Martinez MC (2003). "Enhanced formaldehyde detoxification by overexpression of glutathione-dependent formaldehyde dehydrogenase from Arabidopsis." Plant Physiol 132(4):2248-55. PMID: 12913179

Degrassi99: Degrassi G, Uotila L, Klima R, Venturi V (1999). "Purification and properties of an esterase from the yeast Saccharomyces cerevisiae and identification of the encoding gene." Appl Environ Microbiol 65(8):3470-2. PMID: 10427036

GompelKlei89: Gompel-Klein P, Mack M, Brendel M (1989). "Molecular characterization of the two genes SNQ and SFA that confer hyperresistance to 4-nitroquinoline-N-oxide and formaldehyde in Saccharomyces cerevisiae." Curr Genet 16(2):65-74. PMID: 2557161

Grey96: Grey M, Schmidt M, Brendel M (1996), "Overexpression of ADH1 confers hyper-resistance to formaldehyde in Saccharomyces cerevisiae." Curr Genet 29(5);437-40. PMID: 8625422

Jelinsky99: Jelinsky SA, Samson LD (1999). "Global response of Saccharomyces cerevisiae to an alkylating agent." Proc Natl Acad Sci U S A 96(4):1486-91. PMID: 9990050

Overkamp02: Overkamp KM, Kotter P, van der Hoek R, Schoondermark-Stolk S, Luttik MA, van Dijken JP, Pronk JT (2002). "Functional analysis of structural genes for NAD(+)-dependent formate dehydrogenase in Saccharomyces cerevisiae." Yeast 19(6):509-20. PMID: 11921099

# How did we integrate Pathways with other resources?

From the home page, from Quick Search box, from individual locus pages



#### Search Options

Advanced Search, Full-text Search, Search SGD web pages, Global Gene Hunter, Search Literature, and more.

#### Help Resources

Getting Started, Sitemap, FAQ, and more

#### ► Analysis & Tools

BLAST, Gene/Seq Resources, Maps, and more.

#### ► Homology & Comparisons

PDB Homologs, Protein Domains/Motifs, Homologs, and more.

#### ► Function & Expression

Protein Info, Function Junction, Pathways, Expression Connection, and more

#### GO Resources

GO Tutorial, What is GO?, GO Slim Mapper, GO Term Finder, and more.

#### Community Info

Search SGD colleagues, Find yeast labs, Conferences & Seminars, and more.

#### ► Submit Data

Gene Registry, Community Annotation, Colleague update, Contact SGD, and more.

#### ▶ Download Data

FTP, Batch Download, Lab strains, Auxotrophic markers, Gene Names, Sequencing Tables, SGD Lite, and more

#### External Links

Virtual Library, and more.

#### About SGD

Genome Snapshot, Usage statistics, Citing SGD, Publications, Staff, and more SGDIM is a scientific database of the molecular biology and genetics of the yeast Saccharomyces cerevisiae, which is commonly known as baker's or budding yeast.

#### **New and Noteworthy**

#### · Updated transcription factor binding site data

The transcription factor binding sites published by <u>Harbison, et al.</u> (2004) have been remapped to the current *S. cerevisiae* genome sequence. These data can be viewed in GBrowse and in the "Regulatory Role" section on the Locus Summary Pages (see <u>SUMI</u>). In addition, the regulatory map data from <u>MacIsaae</u>, et al. (2006) are available from the <u>FTP</u> site. Thanks to Kenzie MacIsaae for providing information to remap the binding sites and new data to SGD. (Posted June 2, 2006)

#### SGD Curation News

- Genome Snapshot
- New papers added to SGD this week.
- · View Genome-wide Analysis papers in SGD.
- Do a literature search.
- Table of Gene Summary Paragraphs (new ones highlighted yellow).
- Table of Community Annotations

#### · Changes to the Systematic Sequence

As new data become available, SGD curators update the systematic sequence and its annotation. Information regarding sequence and annotation changes can be found in the Table of Sequence Updates and in the "Locus History" pages of affected features. Files on the SGD FTP site are updated weekly.

Between 2006-04-15 and 2006-05-12, sequence and/or annotation changes were made affecting 16 features (UTR4/YEL038W, YJR003C, YKL207W, KKO8/YKL168C, CAF4/YKR036C, SEN15/YMR059W, ZIM17/YNL310C, SPS19/YNL202W, MCA1/YOR197W, CEN1, CEN4, CEN5, CEN6, CEN7, CEN9, and CEN16). (Posted May 12, 2006)

#### • SGD Quarterly Newsletter Available

SGD has sent out the quarterly newsletter to colleagues designated as contacts in SGD. An <u>expanded HTML version</u> is also available. If you would like to receive this letter by email in the future, please use the SGD Colleague Submission/Update form to let us know. (Posted May 1, 2006)

#### · Links to eukarYotic OrtholoGY (YOGY) tool added

Links to eukarYotic OrtholoGY (YOGY), a tool to view orthologous proteins from eukaryotic organisms (Homo sapiens, Mus musculus, Rattus norvegicus, Arabidopsis thaliana, Drosophila melanogaster, Caenorhabditis elegans, Plasmodium falciparum, Schizosaccharomyces pombe, and Saccharomyces cerevisiae), have been added to the Model Organism BLASTP Best Hits page and the Comparison Resources pull-down menu on the Locus Summary Pages. This tool provides information from KOGs, Inparanoid, Homologene, OrthoMCL, and manually curated orthologs between S. cerevisiae and S. pombe. YOGY was developed by Penkett CJ, Morris JA, Wood V, and Bahler J (Wellcome Trust Sanger Institute, Cambridge, UK). (Posted April 19, 2006)

• Previous New and Noteworthy items are listed in What's New in SGD in 2006?

# Quick search results



#### SGD Quick Search Result

Go to Advanced Search Page

Below are the search results for your query, ergosterol biosynthesis. If you would like to broaden your search, you may use one or more wildcard characters (\*) to indicate the location(s) where any text will be tolerated in your search term.

#### Search Results for: ergosterol biosynthesis

- 0 Gene names (gene name/alias/ORF name)
- 0 Gene products
- 1 Gene Ontology terms (GO terms, synonyms)
- 0 Colleagues (by last name)
- . 0 Authors (by last name, first initial)
- 1 Biochemical pathways
- 15 Descriptions
- 0 PubMed ID
- 0 Gene Ontology ID

# Yeast Biochemical Pathways



#### **Yeast Biochemical Pathways**

This page offers tools for visualizing biochemical pathways of Saccharomyces cerevisiae at SGD.

Help

- Main Query Page for Yeast Biochemical Pathways
- Main query page for searching Pathways, Reactions, Enzymes by name or EC number, Compounds and more.
- Metabolic Map

A 'bird's eve' view of S.cerevisiae metabolism. (This page could take a moment to load).

• Expression Viewer

Overlay expression data on the Metabolic map.

Yeast Biochemical Pathways are created using the Pathway tools software developed by Peter Karp and his colleagues at <u>SRI</u> International. The current datasets of pathways at SGD were generated using the Pathway Tool's PathoLogic module, which generates an initial set of pathways by comparing SGD annotations to a reference database (MetaCyc). Although PathoLogic creates a genome database, this feature is not curated by SGD.

These automatically generated pathways are then manually curated and corrected, based on published *S.cerevisiae* literature. When necessary, yeast-specific biochemical pathways are added. Since the functions of many of the yeast genes are not yet known, many of the pathways could be incomplete or may even contain errors. Manual curation of pathways is an ongoing process at SGD and we welcome feedback from the research community. If you notice any problems or errors, please send a message to SGD curators.

For more information about searching and browsing the Yeast Biochemical Pathways, please read the help document.

#### NOTE

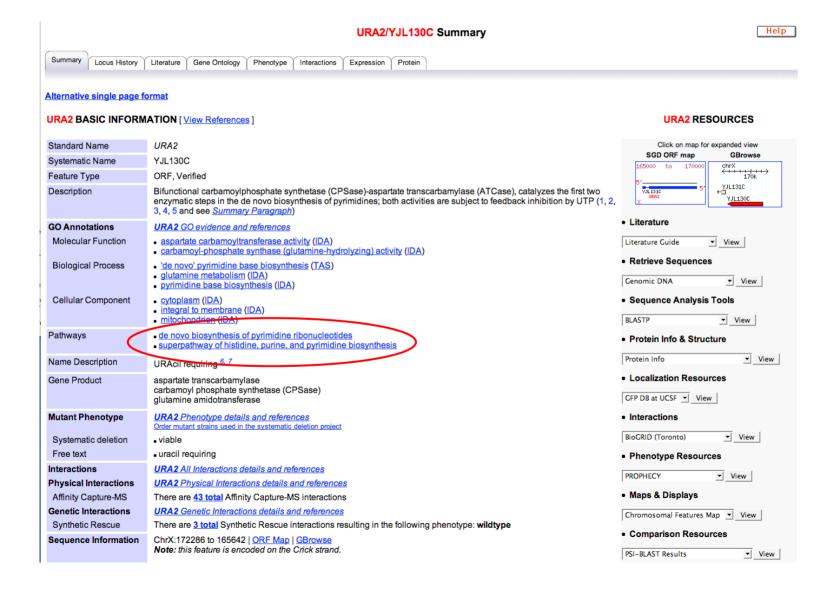
Pathway tools uses a non-standard port (:8555). If you are working behind a firewall, you will not be able to access these pathways at SGD. If this is the case, request your network administrator to open the :8555 port on the firewall.

#### DOWNLOAD

Yeast Biochemical Pathways datasets can be downloaded from our ftp site.

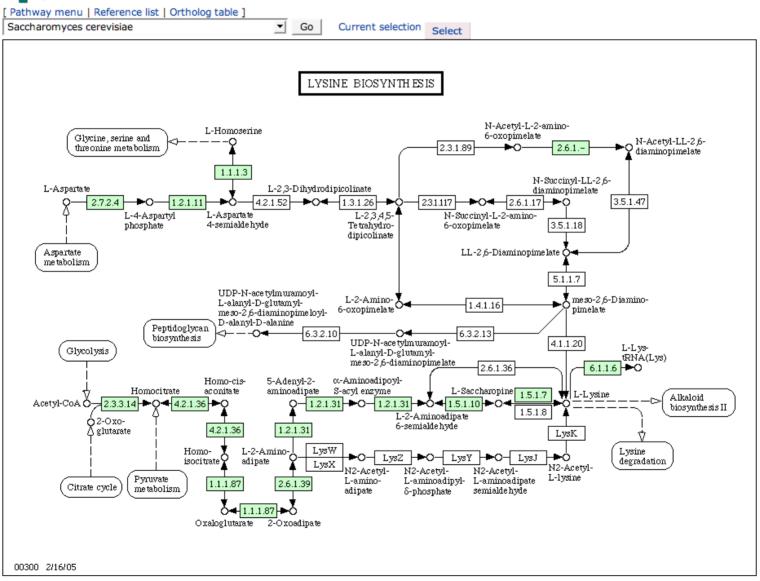


# From the Locus Summary Pages



# Lysine Biosynthesis from KEGG





# Lysine Biosynthesis - Detailed View

S. cerevisiae Pathway: lysine biosynthesis

