

Supplementary Table 1. for MacMullen, W. John (2006). Quantifying literature citations, index terms, and Gene Ontology annotations in the Saccharomyces Genome Database to assess results-set clustering utility. In Proceedings of the 69th Annual Meeting of the American Society for Information Science & Technology (ASIS&T), Vol. 43.

Supplementary Table 1. Title, abstract, and index terms from 10 randomly-selected citation/gene pairs

SGDID, PMID 1st author, year	Journal J. Descriptors	MeSH terms[1]	GO Terms[2]	SGD Topics	High-freq. title / abstract terms[3]
1. S000001855 11959868 Musib, 2002	<i>J. Biol. Chem.</i> Biochemistry	Acrylamides / chemistry; Actins / *chemistry / metabolism; Calcium / metabolism; Cysteine / chemistry; Kinetics; Leucine / chemistry; Models, Molecular; Mutagenesis, Site-Directed; Myosins / metabolism; Oligonucleotides / chemistry; Plasmids / metabolism; Protein Binding; Protein Structure, Tertiary; Saccharomyces cerevisiae / *metabolism; Spectrometry, Fluorescence; Time Factors; Valine / chemistry	MF: structural constituent of cytoskeleton BP: cell wall organization and biogenesis* CC: actin cortical patch*	Mutants / Phenotypes; Protein / Nucleic Acid Structure; Strains / Constructs; Protein-protein Interactions	probe (6), pyrene (5), F-actin (5), fluorescence (3), hydrophobic (3)
2. S000001149 7823940 Konopka, 1995	<i>Mol. Cell. Biol.</i> Cytology; Molecular Biology	Blotting, Western; Cell Cycle Proteins /genetics; Fungal Proteins / biosynthesis / *genetics; Gene Expression; *Genes, Fungal; Genotype; Morphogenesis / genetics; Mutagenesis; Peptides / genetics / physiology; Phenotype; Pheromones / physiology; Profilins; Saccharomyces cerevisiae / cytology / *genetics / metabolism; *Saccharomyces cerevisiae Proteins	MF: cell wall organization and biogenesis* BP: GTPase activity* CC: cytoplasm*	Fungal Related Genes / Proteins; Protein-protein Interactions; Strains / Constructs; Mutants / Phenotypes	AFR1 (11), bud (5), filament (3), receptor (3), morphogenesis (3)
3. S000005589 7017711 Fried, 1981	<i>PNAS</i> Science	Anisomycin / pharmacology; *Cloning, Molecular; Cycloheximide / pharmacology; DNA, Fungal / genetics; Drug Resistance, Microbial; Plasmids; Pyrroles / pharmacology; Ribosomal Proteins / *genetics / metabolism; Saccharomyces cerevisiae / drug effects / *genetics; Sesquiterpenes / *pharmacology; Transformation, Genetic; Trichodermin / *pharmacology	MF: structural constituent of ribosome BP: protein biosynthesis* CC: cytosolic large ribosomal subunit (sensu Eukaryota)	Alias; Strains / Constructs; Mutants / Phenotypes; Function / Process	protein (7), trichodermin (6), gene (5), yeast (5), ribosomal (4)
4. S000005952 7957107 Hirst, 1994	<i>EMBO J.</i> Molecular Biology; Biotechnology	Acid Phosphatase / *biosynthesis / genetics; Alkaline Phosphatase / biosynthesis / genetics; Amino Acid Sequence; Base Sequence; Cell Cycle / *physiology; Cyclin-Dependent Kinases / *physiology; *Cyclins; DNA, Fungal / genetics / metabolism; *DNA-Binding Proteins; Enzyme Induction / drug effects; Fungal Proteins / *physiology; Gene Expression Regulation, Fungal /	MF: cyclin-dependent protein kinase activity BP: protein amino acid phosphorylation* CC: nucleus	Archived Literature	Pho4 (7), phosphate (5), transcription (4), Pho5 (3), Pho80-Pho85 (3)

SGDID, PMID 1st author, year	Journal J. Descriptors	MeSH terms[1]	GO Terms[2]	SGD Topics	High-freq. title / abstract terms[3]
		*drug effects; Helix-Loop-Helix Motifs; *Homeodomain Proteins; Macromolecular Substances; Molecular Sequence Data; Phosphates / pharmacology; Protein Binding; Recombinant Fusion Proteins / metabolism; *Repressor Proteins; Saccharomyces cerevisiae / genetics / *physiology; *Saccharomyces cerevisiae Proteins; *Signal Transduction; Trans-Activators / *physiology; *Transcription Factors; Transcription, Genetic / drug effects			
5. S000001217 7502579 Carmen, 1995	<i>Yeast</i> Yeasts; Yeasts – genetics; Microbiology	Cytochrome c Group / biosynthesis / genetics; *Cytochromes c; DNA, Recombinant; Gene Expression Regulation, Fungal / genetics; Genes, Fungal / *genetics; Glucose; Phosphopyruvate Hydratase / *genetics; Regulatory Sequences, Nucleic Acid / *genetics; Saccharomyces cerevisiae / *genetics / growth & development; *Saccharomyces cerevisiae Proteins; Sequence Deletion; Transcription, Genetic / *genetics	MF: phosphopyruvate hydratase activity BP: glycolysis* CC: soluble fraction*	DNA/RNA Sequence Features; Transcription; Mutants / Phenotypes	ENO1 (10), URS element (7), ENO2 (4), gene (4), repression (3)
6. S000000479 7896088 Liu, 1994	<i>Genetics</i> Genetics	Alleles; Amino Acid Sequence; Base Sequence; Comparative Study; DNA Mutational Analysis; DNA, Fungal / *genetics; DNA-Binding Proteins / physiology; Fungal Proteins / physiology; GTP- Binding Proteins / chemistry / *genetics / physiology; *Gene Expression Regulation, Fungal; Molecular Sequence Data; Mutagenesis, Site- Directed; Point Mutation; *Protein Structure, Tertiary; Repressor Proteins / physiology; Saccharomyces cerevisiae / *genetics; *Saccharomyces cerevisiae Proteins; Sequence Alignment; Sequence Deletion; Sequence Homology, Amino Acid; *Silent Information Regulator Proteins, Saccharomyces cerevisiae; Telomere / *physiology; *Telomere-Binding Proteins; Trans-Activators / physiology; rap GTP- Binding Proteins	MF: protein binding* BP: chromatin silencing at telomere* CC: nuclear telomere cap complex	Function / Process; Mutants / Phenotypes; Strains / Constructs	c-terminal (10), alleles (8), silencing (8), telomeric (8), tail (7)
7. S000000316 9461451	<i>Nucleic Acids Res.</i> Nucleic Acids; Biochemistry	Base Sequence; Cyclins / genetics; DNA, Fungal / genetics; Gene Expression Regulation, Fungal / *drug effects; Genes, Fungal; Genes, Regulator;	MF: transcription coactivator activity* BP: chromatin	Techniques and Reagents	expression (5), repressor (3), Ssn6 (2),

SGDID, PMID 1st author, year	Journal J. Descriptors	MeSH terms[1]	GO Terms[2]	SGD Topics	High-freq. title / abstract terms[3]
Belli, 1998		Genes, Reporter; Kinetics; Lac Operon; Plasmids / genetics; Repressor Proteins / genetics; Saccharomyces cerevisiae / *drug effects / *genetics; Tetracycline / *pharmacology; Tetracycline Resistance / genetics; Transformation, Genetic	remodeling* CC: nucleus		budding (2), activator/repressor (2)
8. S000003113 9506516 Patel, 1998	<i>Cell Cytology</i>	Adenosine Triphosphate / metabolism; *Carrier Proteins; Cell Cycle Proteins / genetics / metabolism / physiology; Endoplasmic Reticulum / *metabolism; Fungal Proteins / genetics / metabolism / *physiology; *Glycoproteins; Intracellular Membranes; Membrane Fusion / *physiology; Membrane Glycoproteins / physiology; Membrane Proteins / analysis / physiology; Microsomes / metabolism; Mutation; Peptide Fragments; Qa-SNARE Proteins; Qb-SNARE Proteins; Recombinant Fusion Proteins; SNARE Proteins; *Saccharomyces cerevisiae Proteins; *Vesicular Transport Proteins; Yeasts / cytology	MF: unknown BP: retrograde vesicle-mediated transport, Golgi to ER CC: endoplasmic reticulum	Mutants / Phenotypes; Protein-protein Interactions; Strains / Construct	membrane (8), fusion (7), organelle (4), t-SNARE (2), vesicular (2)
9. S000000065 11266464 Yoshida, 2001	<i>J. Cell Biol. Cytology</i>	4-Nitroquinoline-1-oxide / pharmacology; Active Transport, Cell Nucleus; Bleomycin / pharmacology; Carrier Proteins / genetics / *metabolism; Cell Nucleus / *metabolism; Cytoplasm / metabolism; DNA Repair; DNA-Binding Proteins / *metabolism; Gene Deletion; *Karyopherins; Microbial Sensitivity Tests; Protein Binding; Receptors, Cytoplasmic and Nuclear / genetics / *metabolism; Replication Factor A; Saccharomyces cerevisiae; *Saccharomyces cerevisiae Proteins; Transcription Factors / genetics / *metabolism; ran GTP-Binding Protein / metabolism	MF: single-stranded DNA binding* BP: DNA recombination* CC: cytoplasm*	Cellular Location; Regulation of; Substrates / Ligands / Cofactors; Protein-protein Interactions; Function / Process; Techniques and Reagents	Kap142p (6), nuclear (6), replication protein A (RPA) (6), protein (5), DNA (5)
10. S000001249 10394911 Charizanis, 1999	<i>Mol. Gen. Genet. Genetics; Molecular Biology</i>	Base Sequence; Cyclic AMP-Dependent Protein Kinases / *metabolism; DNA-Binding Proteins / *metabolism; Fungal Proteins / metabolism; Genes, Reporter; Genotype; Hydrogen Peroxide / pharmacology; Kinetics; Molecular Sequence Data; Mutagenesis, Insertional;	MF: transcription factor activity* BP: response to osmotic stress* CC: nucleus	Function / Process; Genetic Interactions; Regulatory Role; Regulation of; Mutants / Phenotypes; Strains /	Pos9/Skn7 (7), oxidative stress(6), stress (6), gene (6), transcription (5)

SGDID, PMID 1st author, year	Journal J. Descriptors	MeSH terms[1]	GO Terms[2]	SGD Topics	High-freq. title / abstract terms[3]
		Oligodeoxyribonucleotides; *Oxidative Stress; Plasmids; Polymerase Chain Reaction; Saccharomyces cerevisiae / drug effects / genetics / *physiology; *Saccharomyces cerevisiae Proteins; Signal Transduction; Transcription Factors / *metabolism; beta-Galactosidase / biosynthesis / genetics; ras Proteins / *metabolism		Constructs	

[1] MeSH terms with asterisks indicate MeSH Major Topics [MAJR]. Nine of the ten citations had either the MeSH terms ‘Research Support, U.S. Gov’t, P.H.S.’ or ‘Research Support, Non-U.S. Gov’t’ or both, which were removed to increase readability and to save space.

[2] GO terms with asterisks indicate that more than one GO term was annotated to this citation, but the SGD literature file contains only the most used term.

[3] High-frequency terms do not include words such as ‘a’, ‘and’, ‘the’, etc. These may be different terms from those identified in the term co-occurrence counts