

New Data and Collaborations at the *Saccharomyces* Genome Database: Updated reference genome, alleles, and the Alliance of Genome Resources

Stacia R. Engel, Edith D. Wong, Robert S. Nash, Suzi Aleksander, Micheal Alexander, Eric Douglass, Kalpana Karra, Stuart R. Miyasato, Matt Simison, Marek S. Skrzypek, Shuai Weng, and J. Michael Cherry

Department of Genetics, Stanford University, Stanford, CA 94305

ABSTRACT

Saccharomyces cerevisiae is used to provide fundamental understanding of eukaryotic genetics, gene product function, and cellular biological processes. *Saccharomyces* Genome Database (SGD) has been supporting the yeast research community since 1993, serving as its de facto hub. Over the years, SGD has maintained the genetic nomenclature, chromosome maps, and functional annotation, and developed various tools and methods for analysis and curation of a variety of emerging data types. More recently, SGD and six other model organism focused knowledgebases have come together to create the Alliance of Genome Resources to develop sustainable genome information resources that promote and support the use of various model organisms to understand the genetic and genomic bases of human biology and disease. Here we describe recent activities at SGD, including the latest reference genome annotation update, the development of a curation system for mutant alleles, and new pages addressing homology across model organisms as well as the use of yeast to study human disease.

INTRODUCTION

Saccharomyces cerevisiae is used to provide fundamental understanding of eukaryotic genetics, gene product function, and cellular biological processes. The published scientific literature from the yeast research community is integrated into the biomedical knowledgebase *Saccharomyces* Genome Database (SGD; www.yeastgenome.org). Biocurators with expertise in genetics, cell biology, and molecular biology have collected information from more than 100,000 published papers and combined the results from diverse experiments into a comprehensive resource for researchers, educators, and students.

The *S. cerevisiae* nomenclature has been maintained by SGD since 1993 (Cherry et al. 1997). Soon thereafter, SGD began providing the genetic and physical maps for the 16 yeast nuclear chromosomes (Cherry et al. 1997), the catalog of all known yeast proteins (Chervitz et al. 1999; Weng et al. 2003), biological process and molecular function annotations using the Gene Ontology (Dwight et al. 2001), as well as gene expression data and tools for analysis (Ball et al. 2001). SGD has maintained the reference genome from strain S288C, which was the first completely sequenced eukaryotic genome (Goffeau et al. 1996), and its annotation since 1998 (Cherry et al. 1998), along with sequence analysis and retrieval tools for studying that reference genome (Christie et al. 2004; Balakrishnan et al. 2005; Hirschman et al. 2006), and later broadened the reference panel by adding genomes of 11 additional highly studied strains to more fully support the work of the yeast research community (Engel and Cherry, 2013). In parallel to these activities, SGD also developed principles and practices for the extraction and curation of various types of biological data (Ball et al. 2000; Dwight et al. 2004; Hong et al. 2008; Engel et al. 2010; Costanzo et al. 2011; Park et al. 2012; Balakrishnan et al. 2013; Skrzypek and Nash 2015). SGD then turned to further enhancing existing tools and curation practices, including the development of an automated pipeline for pan-genome analysis (Song et al. 2015), release of the Variant Viewer for analysis of the reference genome panel (Sheppard et al. 2016), curation of the complete set of

yeast metabolic pathways (Cherry 2015), and updated curation methods and data models for the capture of post-translational modifications (Hellerstedt et al. 2017), regulatory relationships (Engel et al. 2018), macromolecular complexes (Wong et al. 2019), and the yeast transcriptome (Ng et al. 2019).

As the de facto hub of the yeast research community, SGD also engages in a wide range of outreach and communication activities to disseminate published results, promote collaboration, facilitate scientific discovery, and inform users about new tools, data, or other database developments (MacPherson et al. 2017). These activities include participating in conferences and hosting workshops, direct contact with authors of yeast research papers, the posting of online help resources, and involvement in social media, including the production of video tutorials and webinars, all hosted on SGD's YouTube channel (<https://www.youtube.com/SaccharomycesGenomeDatabase>). To increase readership and reach a broad audience, content posted on one outreach platform is often publicized or announced on other outreach platforms. We also collaborate with the *Genetics Society of America* to annotate online journal articles published in GENETICS and G3: Genes | Genomes | Genetics to associate yeast genes listed in these articles to their respective gene pages at SGD.

More recently, SGD and six other model organism-focused knowledgebases - Mouse Genome Database (MGD; <http://www.informatics.jax.org>, Bult et al. 2019), Rat Genome Database (RGD; <http://rgd.mcw.edu>, Lauderkind et al. 2018), Zebrafish Information Network (ZFIN; <http://zfin.org>, Ruzicka et al. 2019), WormBase (<http://wormbase.org>, Lee et al. 2018), FlyBase (<http://flybase.org>, Thurmond et al. 2019), and the Gene Ontology (GO) Consortium (<http://www.geneontology.org>, The Gene Ontology 2019) - have created the Alliance of Genome Resources ("the Alliance"; <https://www.alliancegenome.org>; Alliance 2019; Alliance 2022). Together, we are working to create an online resource that integrates, develops, and provides harmonized data for all member projects (Alliance 2020). The aim is to more broadly facilitate the use of model organisms to understand the genetic bases of human biology and disease. These efforts build on the well-established collaborations between these groups, who have long worked together to enhance data consolidation, dissemination, visualization, and the application of shared standards (i.e., "data harmonization"). Working within the Alliance has influenced SGD's curation of mutant alleles and human diseases, as well as our use of homology throughout the SGD website to highlight the greater biological context of key findings in yeast research.

GENOME VERSION R64.3.1

Annotation updates and additions

The *S. cerevisiae* strain S288C reference genome annotation was updated for the first time since 2014 (Table 1). The new genome annotation is release R64.3.1, dated 2021-04-21. Resequencing of S288C in 2014 indicated that genomic sequence variation was less than expected between individual laboratory copies of this strain, illustrating that the underlying genome sequence is stable and complete. As such, while SGD updated the annotation of the genomic sequence, the fundamental sequence itself remains unchanged (Cherry et al. 1998; Engel et al. 2014). All new and updated annotations are sourced from the primary literature. The R64.3.1 update included the addition of seven open reading frames (ORFs), five noncoding RNAs (ncRNAs), two upstream ORFs (uORFs), and one long terminal repeat (LTR). Three ORFs had their translation starts shifted to a different methionine, and one ORF received a newly annotated intron and also had its translation stop shifted. We also changed the Sequence Ontology (SO; <http://www.sequenceontology.org/>; Eilbeck et al. 2005) term used to describe the non-transcribed spacers in the ribosomal DNA (rDNA) array.

New ncRNA systematic nomenclature

Also included in the R64.3.1 update is the establishment of a new systematic nomenclature for yeast ncRNAs (Table S1). For many years, a widely adopted systematic nomenclature has existed for yeast protein-coding genes, or ORFs, as many yeast researchers call them (Cherry et al. 1998). With this most recent genome version, we have established a systematic nomenclature for yeast ncRNAs that is similar to, but distinct from, that used for ORFs. All annotated *S. cerevisiae* ncRNAs are designated by a symbol consisting of four uppercase letters, a four-digit number, and another letter, as follows: Y for “Yeast”, NC for “noncoding”, A-Q for the chromosome on which the ncRNA gene resides (where “A” is chromosome I, “B” is chromosome II, etc., up to “P” for chromosome XVI, and lastly “Q” for the mitochondrial chromosome), a four-digit number corresponding to the sequential order of the ncRNA gene on the chromosome (starting from the left telomere and counting toward the right telomere), and W or C indicating whether the ncRNA gene is encoded on the “Watson” or “Crick” strand (where “Watson” runs 5’ to 3’ from left telomere to right telomere, and “Crick” runs 3’ to 5’). For example, [YNCP0002W](#) is the second ncRNA gene from the left end of chromosome XVI and is encoded on the Watson strand. Going forward, when evidence is published pointing to new ncRNA genes, they will be added to the annotation using the next sequential number available for the specific chromosome on which the ncRNA gene resides. In cases in which more than one ncRNA gene is added to any particular chromosome during the same annotation update (i.e., same genome revision), they will be named using the next sequential number starting with the leftmost ncRNA gene and proceeding to the right of the chromosome.

Nomenclature updates - “legacy” gene names

SGD has long been the keeper of the official *Saccharomyces cerevisiae* gene nomenclature (Cherry 1998). Robert Mortimer handed over this responsibility to SGD in 1993 after maintaining the yeast genetic map and gene nomenclature for 30 years (Hawthorne and Mortimer 1960; Mortimer and Schild 1980). The accepted format for gene names in *S. cerevisiae* comprises three uppercase letters followed by a number. The letters typically signify a phrase (referred to as the “Name Description” in SGD) that provides information about a function, mutant phenotype, or process related to that gene, for example “ADE” for “ADENine biosynthesis” or “CDC” for “Cell Division Cycle”. Gene names for many types of chromosomal features follow this basic format regardless of the type of feature named, whether an ORF, a tRNA, another type of non-coding RNA, an ARS, or a genetic locus. Some *S. cerevisiae* gene names that pre-date the current nomenclature standards do not conform to this format, such as [MRLP38](#), [RPL1A](#), and [OM45](#). A few historical gene names predate both the nomenclature standards and the database, and are less computer-friendly than more recent gene names, due to the presence of punctuation. SGD recently updated these gene names to be consistent with current standards and to be more software-friendly by removing punctuation (Table 2).

Although non-standard historical names are maintained in SGD, any new names for yeast genes must conform to the standard format. The SGD Gene Registry (https://www.yeastgenome.org/reserved_name/new) is the agreed-upon system used by the *S. cerevisiae* community and SGD to reserve standard names for newly characterized genes. We have developed a set of Gene Naming guidelines and procedures in order to assist researchers in gene naming. We urge researchers to contact SGD prior to the publication of a new gene name (even if they have previously reserved it) to ensure that the gene name they wish to use is still appropriate. To date, 1363 currently annotated *S. cerevisiae* genes still have no standard gene name and are identified solely via their systematic names (ex. [YFL019C](#)). All *S. cerevisiae* ORFs are categorized into one of three groups: “verified” ORFs are those for which there is clear experimental evidence for the presence of expression of a protein-coding gene; “uncharacterized” ORFs are likely, but not yet established, to encode a protein; and

“dubious” ORFs are unlikely to encode a protein (Fisk et al. 2006). Of the 1363 currently annotated *S. cerevisiae* ORFs without a standard gene name, 131 are verified, 571 are uncharacterized, and the remaining 661 are classified as dubious.

ALLELES

To serve the needs of the yeast community more completely, and influenced by our active participation in the Alliance, SGD has recently revamped the way we curate mutant alleles of genes and has released new allele web pages. For over a decade, SGD has been recording mutant allele information as descriptive properties of phenotype annotations (Costanzo et al. 2009). We have recently elevated alleles in the database from phenotype properties to primary data objects, *i.e.*, standalone entities to which other annotations, such as phenotypes or genetic interactions, can be attached, and that have their own detailed web pages.

Previously, as part of phenotype curation, only allele names, mutant types, and free-text descriptions were collected. We now have expanded the information gathered and displayed for alleles to include affected gene, alias names, description in free text including relevant nucleotide and/or amino acid changes and correspondence to human alleles when known, and allele type. Allele type, which is captured using the *structural_variant* (SO:0001537) branch of the Sequence Ontology, describes the change to the relevant sequence feature with terms like *missense_variant* (SO:0001583), *stop_gained* (SO:0001587), and *frameshift_variant* (SO:0001589). To further describe the functional effect of the structural change within the context of phenotype annotations, SGD uses the *functional_effect_variant* (SO:0001536) branch of the SO, which contains terms like *null_mutation* (SO:0002055), *loss_of_function_variant* (SO:0002054), and *dominant_negative_variant* (SO:0002052).

New allele pages

The new allele pages at SGD contain different types of information divided into several sections: Overview, Phenotypes, Genetic Interactions, Shared Alleles network, external Resources, and relevant Literature (fig. 1). The Allele Overview provides general information about the allele, including its name, the affected gene, the type of allele (*e.g.*, *missense_variant*, as described above), and a description of sequence change and/or domain mutated. The Phenotype Annotations for an allele include an observable feature (*e.g.*, “cell shape”), a qualifier (*e.g.*, “abnormal”), a mutant type (*e.g.*, null), strain background, and a reference. In addition, annotations are classified as classical genetics or high-throughput (*e.g.*, large scale survey, systematic mutation set). The Genetic Interactions for an allele are defined as experimentally observed genetic interactions between that allele and another of a different gene. All interactions listed in SGD are curated by BioGRID (Oughtred et al. 2021). The Shared Alleles network displays positive genetic interactions, negative genetic interactions, and phenotypes that are shared between the given allele and other alleles (fig. 1). The Resources section provides links to allele-related information, such as mutant strain repositories, external phenotype and interaction databases, and information about the Yeast Phenotype Ontology. Additionally, all literature associated with an allele can be found on its Literature page (fig. 1).

Updates to Interactions data/pages

To further accommodate the enhanced focus on alleles and for improved clarity, SGD has made recent updates to the way we present Interaction data. Previously, genetic and physical interaction annotations were combined in one table. These annotations are now displayed in separate annotation tables on Interactions

pages (ex. <https://www.yeastgenome.org/locus/S000003424/interaction>), Reference pages (ex. <https://www.yeastgenome.org/reference/S000305076>), and in YeastMine (<https://yeastmine.yeastgenome.org/yeastmine/templates.do>) to allow for the listing of allele designations, Synthetic Genetic Analysis (SGA) Scores, and P-values from the global genetic interactions paper by Costanzo et al. (2016).

Searching for alleles

All alleles for a specific gene can be accessed via the Alleles section on the Locus Summary Page or downloaded from YeastMine using the Genes -> Alleles template (https://yeastmine.yeastgenome.org/yeastmine/template.do?name=Gene_Alleles&scope=all). The alleles have also been added to SGD's Elasticsearch, with 'facets' for publications, allele types, affected genes, and phenotypes, which allow for the browsing, partitioning, and viewing of the approximately 14,000 yeast alleles annotated so far in SGD.

HOMOLOGY AND THE ALLIANCE OF GENOME RESOURCES

SGD and six other model organism focused knowledgebases - MGD, RGD, ZFIN, WormBase, FlyBase, and the GO Consortium - have recently created a new knowledgebase for model organisms. These major community-driven model organism projects have taken up the mission of harmonizing common data types and creating an integrated web resource. The Alliance of Genome Resources has brought together biocurators and software engineers from all the MODs to build a new central resource. Teams have been created to define data models for the major data types provided by the MODs. These teams also define how the information should be presented. Software engineers from the MODs work together to create the cloud computational environment, component tools, and web pages that define such an important resource. Gene products, proteins, ncRNAs, and pseudogenes are connected by their homology, molecular functions, biological processes, cellular component location, anatomical expression, and association with disease. Human gene details are provided by RGD.

One of the first work products to come out of the Alliance was a consolidated set of orthologs, using data from several different computational and manually curated sources (Howe et al. 2018). Many aspects of data integration presented at the Alliance require a common set of orthology relationships among genes for the organisms represented, including human. The Alliance provides the results of all methods that have been benchmarked by the Quest for Orthologs Consortium (QfO; <https://questfororthologs.org/>, Linard et al. 2021). The homolog inferences from the different methods have been integrated using the DRSC Integrative Ortholog Prediction Tool (DIOPT; Hu et al. 2011), which integrates a number of existing methods including those used by the Alliance: Ensembl Compara (Vilella et al. 2009), HUGO Gene Nomenclature Committee (HGNC; Povey et al. 2001), Hieranoid (Kaduk and Sonnhammer 2017), InParanoid (O'Brien et al. 2005), the Orthologous MATrix project (OMA; Schneider et al. 2007), OrthoMCL (Li et al. 2003), OrthoFinder (Emms and Kelly 2015), OrthoInspector (Linard et al. 2011), PANTHER (Mi et al. 2018), PhylomeDB (Huerta-Cepas et al. 2008), Roundup (Deluca et al. 2006), TreeFam (Li et al. 2006), and ZFIN (Westerfield et al. 1997). This set of assertions (the "orthology set") is key because it allows the inference of function from one species to another, and is especially helpful because work using a highly studied and experimentally tractable organism such as yeast can be informative regarding the biology of other organisms in which targeted experiments are not possible, such as human (Dolinski and Botstein 2007). Currently included in the Alliance and in the orthology set are five other leading model organisms in addition to *S. cerevisiae* (yeast): *Caenorhabditis elegans* (worm),

Drosophila melanogaster (fruit fly), *Danio rerio* (zebrafish), *Rattus norvegicus* (rat), and *Mus musculus* (mouse).

SGD takes advantage of the homology data from the Alliance to provide easy access to information about homologous genes in just one click. At SGD, we have used the orthology set to provide links between SGD gene pages and those for orthologous genes at the Alliance (fig. 2). On SGD gene pages, users will find hexagonal icons representing each model organism (human, mouse, rat, zebrafish, fly, worm, and yeast) for which there is homologous gene information at the Alliance. Clicking on the icon immediately directs the user to the gene page for the selected model organism on the Alliance website. The Alliance gene pages present a variety of data types for all of the various organisms, including functional annotations using the Gene Ontology, phenotypes, disease associations, alleles, variants, sequence features, expression information, and both physical and genetic interactions. Data for individual genes can be downloaded directly from Alliance gene pages, and bulk downloads are available from the Alliance Downloads site (<https://www.alliancegenome.org/downloads>).

Homology pages

We have also used the Alliance orthology set to establish new Homology pages at SGD for protein-coding genes. The information displayed on the Homology Pages is divided into several sections: Homologs, Functional Complementation, Fungal Homologs, and External Identifiers (fig. 3). The Homologs section lists information about genes in the Alliance orthology set, including species, gene ID, and gene name, with links to the corresponding gene page at the Alliance. The Functional Complementation section presents data about cross-species complementation between yeast and other species, as curated by SGD and including legacy data from the Princeton Protein Orthology Database (P-POD; Heinicke et al. 2007). These curated Functional Complementation data are also displayed on SGD Reference pages. The Fungal Homologs section shows curated information including species, gene ID, and database source for orthologs in 24 additional species of fungi, such as those in *Candida*, *Neurospora*, and *Aspergillus*, among others. The External Identifiers section lists identifiers for the protein from various database sources.

Disease pages

To promote the use of yeast as a catalyst for biomedical research, SGD utilizes the Disease Ontology (DO; Schriml et al. 2018) to describe human diseases that are associated with yeast homologs. Disease Ontology annotations to yeast genes are now available through SGD's new Disease pages, each of which corresponds to a Disease Ontology term, such as *amyotrophic lateral sclerosis* (DOID:332), and lists out all yeast genes annotated to the term by SGD. Yeast genes with one or more human disease associations also have a new Disease tab accessible from the genes' respective locus pages (fig. 4). The Disease tab shows all manually curated, high-throughput, and computational disease annotations for the yeast gene. Additionally, these pages feature a network diagram that depicts shared disease annotations for other yeast genes and their human homologs.

While the Alliance will provide much of the information researchers will want to explore, there will continue to be data that cannot be harmonized across organisms because of its unique characteristics, such as an aspect of biology only found in one of the species or the nuance of an experimental assay that is unique to one community. In time, these unique bits of information will be provided directly from the Alliance site. However, until then, the main knowledgebases are essential. All of the MODs look forward to the common user interface and ease of discovering common associations between genes in other organisms. This is an important step to

allow researchers, educators, and students to have access to the gold standard of expertly curated information on each of these foundations of biological science.

OTHER UPDATES

Supplemental data and published datasets on reference pages

SGD procures and displays supplemental materials for references stored in our database. We are hosting data from past, present, and future papers on our literature pages. To access these data, simply search SGD with the PubMed ID and then look for the “Downloadable Files.” Additionally, published large-scale datasets from Gene Expression Omnibus (GEO) are displayed in the “Published Datasets” section. The dataset title is linked to a dataset-specific page and a controlled vocabulary of terms is used to bin similar datasets into broad categories.

Submit data form

Authors can submit data and information about their publications, including novel results, datasets (including accession IDs), or other important information, using SGD’s simple “Submit Data” form (<https://www.yeastgenome.org/submitData>).

Explore button

SGD introduced a new “Explore SGD” button on our homepage, which allows users to peruse SGD data and pages without an initial search query. After selecting the “Explore SGD” button, users are redirected to a search results page where they can browse all of the information in SGD. The tool is designed for both new and veteran users alike, as new users are provided a glimpse into the warehouse of information SGD contains, while seasoned users may discover something new. After clicking on the “Explore SGD” button, use the categories on the left to navigate through the various pages and examine areas of interest. An “Explore” link has also been added to the selection of links available in the black bar at the top of every SGD webpage, giving users the ability to access the search results page from anywhere on the SGD website.

microPublication Biology

SGD has partnered with *microPublication Biology* (<https://www.micropublication.org>) to promote the growing model of rapid publication of research results. *microPublication* further expands this rapid-release model by coupling publication with curation and deposition into community-directed authoritative databases such as SGD and the other Alliance member databases mentioned above. Each publication is peer-reviewed, assigned a digital object identifier (DOI), published online, indexed in PubMed (<https://pubmed.ncbi.nlm.nih.gov>), and then integrated into the relevant databases, speeding information dissemination and scientific discovery. SGD encourages authors to utilize *microPublication* for new research findings or reagents, experimental results that do not fit into a larger existing or future narrative, negative results, successful replications of recent work, cautionary findings regarding unsuccessful attempts at replication of recent work, “data not shown” from other publications, and data from student projects. All published yeast articles are immediately available in SGD (<https://www.yeastgenome.org/search?category=reference&journal=microPublication.%20Biology>).

ACKNOWLEDGEMENTS

Thanks to Jodi Lew-Smith for help preparing this manuscript.

FUNDING

Research reported in this publication was supported by the National Human Genome Research Institute of the National Institutes of Health under Award Number U24HG001315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- Alliance of Genome Resources Consortium, 2019 The Alliance of Genome Resources: Building a Modern Data Ecosystem for Model Organism Databases. *Genetics*. 213: 1189-1196. doi: 10.1534/genetics.119.302523.
- Alliance of Genome Resources Consortium, 2020 Alliance of Genome Resources Portal: unified model organism research platform. *Nucleic Acids Res*. 48(Database Issue): D650-D658. doi: 10.1093/nar/gkz813
- Alliance of Genome Resources, 2022 *Genetics* the same issue as this paper
- Awasthi, A., V. Nain, C.V. Srikanth, and R. Puria, 2020 A regulatory circuit between lncRNA and TOR directs amino acid uptake in yeast. *Biochim Biophys Acta Mol Cell Res*. 1867: 118680. doi: 10.1016/j.bbamcr.2020.118680.
- Balakrishnan, R., K.R. Christie, M.C. Costanzo, K. Dolinski, S.S. Dwight *et al.*, 2005 Fungal BLAST and Model Organism BLASTP Best Hits: new comparison resources at the Saccharomyces Genome Database (SGD). *Nucleic Acids Res*. 33(Database Issue): D374-D377.
- Balakrishnan, R., M.A. Harris, R. Huntley, K. Van Auken, and J.M. Cherry, 2013 A guide to best practices for Gene Ontology (GO) manual annotation. *Database (Oxford)* doi: 10.1093/database/bat054.
- Ball, C.A., K. Dolinski, S.S. Dwight, M.A. Harris, L. Issel-Tarver, *et al.*, 2000 Integrating functional genomic information into the Saccharomyces genome database. *Nucleic Acids Res* 28: 77-80.
- Ball, C.A., H. Jin, G. Sherlock, S. Weng, J.C. Matese, *et al.*, 2001 Saccharomyces Genome Database provides tools to survey gene expression and functional analysis data. *Nucleic Acids Res* Jan 1;29(1):80-1.
- Bult, C.J., J.A. Blake, C.L. Smith, J.A. Kadin, J.E. Richardson, *et al.*, 2019 Mouse genome database (MGD) 2019. *Nucleic Acids Res*. 47: D801–D806. doi: 10.1093/nar/gky1056
- Bunina, D., M. Štefl, F. Huber, A. Khmelinskii, M. Meurer, *et al.*, 2017 Upregulation of SPS100 gene expression by an antisense RNA via a switch of mRNA isoforms with different stabilities. *Nucleic Acids Res*. 45: 11144-11158. doi: 10.1093/nar/gkx737.

- 377 Camblong, J., N. Iglesias, C. Fickentscher, G. Dieppois, and F. Stutz. Antisense RNA stabilization induces
378 transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell*. 131:706-717. doi:
379 10.1016/j.cell.2007.09.014.
- 380
- 381 Cherry, J.M., 1998 Genetic nomenclature guide. *Saccharomyces cerevisiae*. *Trends Genet* 14: S.10-S.11
- 382
- 383 Cherry, J.M., 2015 The *Saccharomyces* Genome Database: Exploring Biochemical Pathways and Mutant
384 Phenotypes. *Cold Spring Harb Protoc*. doi: 10.1101/pdb.prot088898.
- 385
- 386 Cherry, J.M., C. Ball, S. Weng, G. Juvik, R. Schmidt, *et al.*, 1997 Genetic and physical maps of
387 *Saccharomyces cerevisiae*. *Nature* 387 (6632 Suppl), 67-73.
- 388
- 389 Cherry, J.M., C. Adler, C.A. Ball, S.A. Chervitz, S.S. Dwight, *et al.*, 1998 SGD: *Saccharomyces* Genome
390 Database. *Nucleic Acids Res* 26: 73-80.
- 391
- 392 Chervitz, S.A., E.T. Hester, C.A. Ball, K. Dolinski, S.S. Dwight, *et al.*, 1999 Using the *Saccharomyces* Genome
393 Database (SGD) for analysis of protein similarities and structure. *Nucleic Acids Res* 27: 74-78.
- 394
- 395 Christie, K.R., S. Weng, R. Balakrishnan, M.C. Costanzo, K. Dolinski, *et al.* 2004 *Saccharomyces* Genome
396 Database (SGD) provides tools to identify and analyze sequences from *Saccharomyces cerevisiae* and related
397 sequences from other organisms. *Nucleic Acids Res*. 32(Database issue): D311-D314.
- 398
- 399 Costanzo, M., B. VanderSluis, E.H. Koch, A. Baryshnikova, C. Pons, *et al.* 2016 A global genetic interaction
400 network maps a wiring diagram of cellular function. *Science* 353(6306):aaf1420. doi: 10.1126/science.aaf1420.
- 401
- 402 Costanzo, M.C., M.S. Skrzypek, R. Nash, E. Wong, G. Binkley, *et al.* 2009 New mutant phenotype data
403 curation system in the *Saccharomyces* Genome Database. *Database*; doi: 10.1093/database/bap001.
- 404
- 405 Costanzo, M.C., J. Park, R. Balakrishnan, J.M. Cherry, and E.L. Hong, 2011 Using computational predictions
406 to improve literature-based Gene Ontology annotations: a feasibility study. *Database (Oxford)* doi:
407 10.1093/database/bar004.
- 408
- 409 Deluca, T.F., I.H. Wu, J. Pu, T. Monaghan, L. Peshkin, *et al.*, 2006 Roundup: a multi-genome repository of
410 orthologs and evolutionary distances. *Bioinformatics* 22: 2044–2046.
- 411
- 412 Dolinski, K., and D. Botstein, 2007 Orthology and functional conservation in eukaryotes. *Annu Rev Genet*. 41:
413 465-507. doi: 10.1146/annurev.genet.40.110405.090439.
- 414
- 415 Dwight, S.S., M.A. Harris, K. Dolinski, C.A. Ball, G. Binkley, *et al.*, 2002 *Saccharomyces* Genome Database
416 (SGD) provides secondary gene annotation using the Gene Ontology (GO). *Nucleic Acids Res*. 30: 69-72.
- 417
- 418 Dwight, S.S., R. Balakrishnan, K.R. Christie, M.C. Costanzo, K. Dolinski, *et al.*, 2004 *Saccharomyces* genome
419 database: underlying principles and organisation. *Brief Bioinform*. 5: 9-22.
- 420
- 421 Eilbeck, K., S.E. Lewis, C.J. Mungall, M. Yandell, L. Stein, *et al.*, 2005 The Sequence Ontology: A tool for the
422 unification of genome annotations. *Genome Biology* 6:R44 doi: 10.1186/gb-2005-6-5-r44
- 423

- Eisenberg-Bord, M., M. Mari, U. Weill, E. Rosenfeld-Gur, O. Moldavski, *et al.*, 2017 Identification of seipin-linked factors that act as determinants of a lipid droplet subpopulation. *J Cell Biol.* 217:269-282. doi: 10.1083/jcb.201704122.
- Emms, D.M. and S. Kelly, 2015 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Gen Biol* 16: 157 doi: 10.1186/s13059-015-0721-2
- Engel, S.R. and J.M. Cherry, 2013 The new modern era of yeast genomics: community sequencing and the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the Saccharomyces Genome Database. Database (Oxford) doi: 10.1093/database/bat012.
- Engel, S.R., R. Balakrishnan, G. Binkley, K.R. Christie, M.C. Costanzo, *et al.*, 2010 Saccharomyces Genome Database provides mutant phenotype data. *Nucleic Acids Res.* 38(Database issue): D433-D436.
- Engel, S.R., F.S. Dietrich, D.G. Fisk, G. Binkley, R. Balakrishnan, *et al.*, 2014 The Reference Genome Sequence of *Saccharomyces cerevisiae*: Then and Now. *G3 (Bethesda)*. 4: 389-398. doi: 10.1534/g3.113.008995.
- Engel, S.R., M.S. Skrzypek, S.T. Hellerstedt, E.D. Wong, R.S. Nash, *et al.*, 2018 Updated regulation curation model at the Saccharomyces Genome Database. Database (Oxford). doi: 10.1093/database/bay007.
- Fisk, D.G., C.A. Ball, K. Dolinski, S.R. Engel, E.L. Hong, *et al.*, 2006 *Saccharomyces cerevisiae* S288C genome annotation: a working hypothesis. *Yeast* 23: 857-865.
- Geisler, S., L. Lojek, A.M. Khalil, K.E. Baker, and J. Collier, 2012 Decapping of long noncoding RNAs regulates inducible genes. *Mol Cell.* 45:279-291. doi: 10.1016/j.molcel.2011.11.025.
- Goffeau, A., B.G. Barrell, H. Bussy, R.W. Davis, B. Dujon, *et al.*, 1996 Life with 6000 genes. *Science* 274(5287):546, 563-7. doi: 10.1126/science.274.5287.546.
- Hawthorne, D.C. and R.K. Mortimer, 1960 Chromosome Mapping in *Saccharomyces*: centromere-linked genes. *Genetics* 45: 1085-1110
- He, C., C. Jia, Y. Zhang, and P. Xu, 2018 Enrichment-Based Proteogenomics Identifies Microproteins, Missing Proteins, and Novel smORFs in *Saccharomyces cerevisiae*. *J Proteome Res.* 17:2335-2344. doi: 10.1021/acs.jproteome.8b00032.
- Heinicke, S., M.S. Livstone, C. Lu, R. Oughtred, F. Kang, *et al.*, 2007 The Princeton Protein Orthology Database (P-POD): A Comparative Genomics Analysis Tool for Biologists. *PLOS One*. doi: 10.1371/journal.pone.0000766.
- Hellerstedt, S.T., R.S. Nash, S. Weng, K.M. Paskov, E.D. Wong, *et al.*, 2017 Curated protein information in the Saccharomyces Genome Database. doi: 10.1093/database/bax011.
- Hirschman, J.E., R. Balakrishnan, K.R. Christie, M.C. Costanzo, S.S. Dwight, *et al.*, 2006 Genome Snapshot: a new resource at the Saccharomyces Genome Database (SGD) presenting an overview of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res* 34(Database issue): D442-D445.

- Hong, E.L., R. Balakrishnan, Q. Dong, K.R. Christie, J. Park, *et al.*, 2008 Gene Ontology annotations at SGD: new data sources and annotation methods. *Nucleic Acids Res.* 36(Database issue): D577-D581.
- Houseley, J., L. Rubbi, M. Grunstein, D. Tollervey, and M. Vogelauer, 2008 A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol Cell.* 32:685-695. doi: 10.1016/j.molcel.2008.09.027.
- Howe, D.G., J.A. Blake, Y.M. Bradford, C.J. Bult, B.R. Calvi, *et al.*, 2018 Model organism data evolving in support of translational medicine. *Lab Anim (NY)*. doi: 10.1038/s41684-018-0150-4.
- Hu, Y., I. Flockhart, A. Vinayagam, C. Bergwitz, B. Berger, *et al.*, 2011 An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*. doi: 10.1186/1471-2105-12-357.
- Huber, F., D. Bunina, I. Gupta, A. Khmelinskii, M. Meurer, *et al.*, 2016 Protein Abundance Control by Non-coding Antisense Transcription. *Cell Rep.* 15:2625-2636. doi: 10.1016/j.celrep.2016.05.043.
- Huerta-Cepas, J., A. Bueno, J. Dopazo and T. Gabaldon, 2008 PhylomeDB: a database for genome-wide collections of gene phylogenies. *Nucleic Acids Res* 36(Database Issue): D491–D496.
- Jeon, S. and J. Kim, 2010 Upstream open reading frames regulate the cell cycle-dependent expression of the RNA helicase Rok1 in *Saccharomyces cerevisiae*. *FEBS Lett.* 584: 4593-4598. doi: 10.1016/j.febslet.2010.10.019.
- Kaduk, M. and E. Sonnhammer, 2017 Improved orthology inference with Hieranoid 2, *Bioinformatics* 33: 1154-1159
- Laulederkind, S.J.F., G.T. Hayman, S.J. Wang, J.R. Smith, V. Petri, *et al.*, 2018 A primer for the rat genome database (RGD). *Methods Mol. Biol.* 1757: 163–209. doi: 10.1007/978-1-4939-7737-6_8
- Lee, R.Y.N., K.L. Howe, T.W. Harris, V. Arnaboldi, S. Cain, *et al.*, 2018 WormBase 2017: molting into a new stage. *Nucleic Acids Res.* 46: D869–D874. doi: 10.1093/nar/gkx998
- Li, H., A. Coghlan, J. Ruan, L.J. Coin, J.K. Heriche, *et al.*, 2006. TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* 34(Database Issue): D572–D580.
- Li, L., C.J. Stoeckert Jr, and D.S. Roos, 2003 OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13: 2178–2189. doi: 10.1101/gr.1224503
- Linard, B., J.D. Thompson, O. Poch, and O. Lecompte, 2011 OrthoInspector: comprehensive orthology analysis and visual exploration. *BMC Bioinformatics* 12: 11 doi: 10.1186/1471-2105-12-11
- Linard, B., I. Ebersberger, S. McGlynn, N. Glover, T. Mochizuki, *et al.*, 2021 Ten years of collaborative progress in the Quest for Orthologs. *Molecular Biology and Evolution*, 38: 3033–3045, doi: 10.1093/molbev/msab098.

- Nene, R.V., C.D. Putnam, B.Z. Li, K.G. Nguyen, A. Srivatsan, *et al.*, 2018 Cdc73 suppresses genome instability by mediating telomere homeostasis. PLoS Genet. 14:e1007170. doi: 10.1371/journal.pgen.1007170.
- MacPherson, K.A., B. Starr, E.D. Wong, K.S. Dalusag, S.T. Hellerstedt, *et al.*, 2017 Outreach and online training services at the Saccharomyces Genome Database. doi: 10.1093/database/bax002.
- Makanae, K., R. Kintaka, K. Ishikawa, and H. Moriya, 2015 Small toxic protein encoded on chromosome VII of Saccharomyces cerevisiae. PLoS One. 10: e0120678. doi: 10.1371/journal.pone.0120678.
- Mi, H., A. Muruganujan, D. Ebert, X. Huang, and P.D. Thomas, 2018 PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res 47(Database Issue): D419-D426
- Mortimer, R.K. and D. Schild, D., 1980 Genetic map of Saccharomyces cerevisiae. Microbiol. Rev. 44: 519-571
- Ng, P.C., E.D. Wong, K.A. MacPherson, S. Aleksander, J. Argasinska, *et al.*, 2019 Transcriptome visualization and data availability at the Saccharomyces Genome Database. Nucleic Acids Res. 48(Database Issue) doi: 10.1093/nar/gkz892.
- O'Brien, K.P., M. Remm, and E.L. Sonnhammer, 2005 Inparanoid: a comprehensive database of eukaryotic orthologs. Nucleic Acids Res 33(Database Issue): D476–D480.
- Oughtred, R., J. Rust, C. Chang, B.-J. Breitkreutz, C. Stark, *et al.*, 2021 The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci. 30: 187-200
- Park, J., M.C. Costanzo, R. Balakrishnan, J.M. Cherry, and E.L. Hong, 2012 CvManGO, a method for leveraging computational predictions to improve literature-based Gene Ontology annotations. Database (Oxford) doi: 10.1093/database/bas001.
- Paul, V.D., U. Mühlenhoff, M. Stümpfig, J. Seebacher, K.G. Kugler, *et al.*, 2015 The deca-GX3 proteins Yae1-Lto1 function as adaptors recruiting the ABC protein Rli1 for iron-sulfur cluster insertion. Elife. 4: e08231. doi: 10.7554/eLife.08231.
- Povey, S., R. Lovering, E. Bruford, M. Wright, M. Lush, *et al.*, 2001 The HUGO Gene Nomenclature Committee (HGNC). Hum Genet 109: 678-680
- Ruzicka L, Howe DG, Ramachandran S, Toro S, Van Slyke CE, et al., 2019 The Zebrafish Information Network: new support for non-coding genes, richer Gene Ontology annotations and the Alliance of Genome Resources. Nucleic Acids Res. Jan 8;47(D1):D867-D873. doi: 10.1093/nar/gky1090
- Sadhu, M.J., J.S. Bloom, L. Day, J.J. Siegel, S. Kosuri, *et al.*, 2018 Highly parallel genome variant engineering with CRISPR-Cas9. Nat Genet. 50:510-514. doi: 10.1038/s41588-018-0087-y.

- Sampath, V., B. Liu, S. Tafrov, M. Srinivasan, R. Rieger, *et al.*, 2013 Biochemical characterization of Hpa2 and Hpa3, two small closely related acetyltransferases from *Saccharomyces cerevisiae*. *J Biol Chem.* 288:21506-21513. doi: 10.1074/jbc.M113.486274.
- Sheppard, T.K., B.C. Hitz, S.R. Engel, G. Song, R. Balakrishnan, *et al.*, 2016 The *Saccharomyces* Genome Database Variant Viewer. *Nucleic Acids Res.* 44(Database Issue): D698-D702. doi: 10.1093/nar/gkv1250.
- Schriml L.M., Mitarka E., Munro J., Tauber B., Schor M., *et al.*, 2018 Human Disease Ontology 2018 update: classification, content and workflow expansion. *Nucleic acids research*, 47(Database Issues): D955–D962. doi: doi:10.1093/nar/gky1032
- Schneider, A., C. Dessimoz, and G.H. Gonnet, 2007 OMA Browser--exploring orthologous relations across 352 complete genomes. *Bioinformatics* 23: 2180–2182. doi: 10.1093/bioinformatics/btm295
- Skrzypek, M.S. and R.S. Nash, 2015 Biocuration at the *Saccharomyces* Genome Database. *Genesis.* 53: 450-457. doi: 10.1002/dvg.22862.
- Song, G., B.J. Dickins, J. Demeter, S. Engel, B. Dunn, *et al.*, 2015 AGAPE (Automated Genome Analysis Pipeline) for Pan-Genome Analysis of *Saccharomyces cerevisiae*. *PLoS One.* 10: e0120671. doi: 10.1371/journal.pone.0120671.
- The Gene Ontology Consortium, 2019 The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47: D330–D338. doi: 10.1093/nar/gky1055
- Thurmond, J., J.L. Goodman, V.B. Strelets, H. Attrill, L.S. Gramates, *et al.*, 2019 FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47: D759–D765. doi: 10.1093/nar/gky1003.
- Vilella, A.J., J. Severin, A. Ureta-Vidal, L. Heng, R. Durbin, *et al.*, 2009 EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res* 19: 327–335.
- Weng, S., Q. Dong, R. Balakrishnan, K. Christie, M. Costanzo, *et al.*, 2003 *Saccharomyces* Genome Database (SGD) provides biochemical and structural information for budding yeast proteins. *Nucleic Acids Res.* 31: 216-218.
- Westerfield, M., E. Doerry, A.E. Kirkpatrick, W. Driever, and S.A. Douglas, 1997 An on-line database for zebrafish development and genetics research. *Semin. Cell Dev. Biol.* 8: 477–488. doi: 10.1006/scdb.1997.0173.
- Wong, E.D., M.S. Skrzypek, S. Weng, G. Binkley, B.H.M. Meldal, *et al.*, 2019 Integration of macromolecular complex data into the *Saccharomyces* Genome Database. *Database (Oxford)*. doi: 10.1093/database/baz008.
- Wu, X. and J.E. Haber, 1996 A 700 bp cis-acting region controls mating-type dependent recombination along the entire left arm of yeast chromosome III. *Cell.* 87: 277-85. doi: 10.1016/s0092-8674(00)81345-8.
- Xu, Z., W. Wei, J. Gagneur, F. Perocchi, S. Clauder-Münster, *et al.*, 2009 Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457: 1033-1037. doi: 10.1038/nature07728.

Yagoub, D., A.P. Tay, Z. Chen, J.J. Hamey, C. Cai, *et al.*, 2015 Proteogenomic Discovery of a Small, Novel Protein in Yeast Reveals a Strategy for the Detection of Unannotated Short Open Reading Frames. J Proteome Res. 14:5038-5047. doi: 10.1021/acs.jproteome.5b00734.

Table 1. The *S. cerevisiae* strain S288C reference genome annotation was updated. The new genome annotation is release R64.3.1, dated 2021-04-21.

Chromosome	Feature	Description of change	Reference
II	YNCB0008W aka GAL10-ncRNA	New ncRNA antisense to GAL10: coordinates 276805..280645	Houseley et al. 2008; Pinskaya et al. 2009; Geisler et al. 2012
II	YNCB0014W aka TBRT/XUT_2F-154	New ncRNA antisense to TAT1: coordinates 376610..378633	Awasthi et al. 2020
III	RE/RE301	New recombination enhancer: coordinates 29108..29809	Wu and Haber 1996
V	YELWdelta27	New Ty1 LTR: coordinates 449274..449626	Nene et al. 2018
V	HPA3/YEL066W	Moved translation start to Met19: old coordinates 26667..27206 new coordinates 26721..27206	Sampath et al. 2013
VII	OTO1/YGR227C-A	New ORF: coordinates 949052..949225 Crick	Makanae et al. 2015
VII	ROK1/YGL171W	Two new uORFs: coordinates uORF1 182286..182407 coordinates uORF2 182291..182329	Jeon and Kim 2010
VIII	YHR052C-B	New ORF: coordinates 212519..212692 Crick	He et al. 2018
VIII	YHR054C-B	New ORF: coordinates 214517..214690 Crick	He et al. 2018
VIII	SUT169/YNCH0011W	New ncRNA: coordinates 378254..379237	Xu et al. 2009; Geisler et al. 2012; Huber et al. 2016; Bunina et al. 2017
X	YJR012C	Moved start to Met76: old coordinates 459795..460418 Crick new coordinates 459795..460193 Crick	Sadhu et al. 2018
X	YJR107C-A	New ORF coordinates 628457..628693 Crick	Yagoub et al. 2015; He et al. 2018
XI	YKL104W-A	New ORF: coordinates 245032..245286	He et al. 2018
XII	YLR379W-A	New ORF: coordinates 877444..877716	Internal reanalysis of results from Song et al. 2015 to find and annotate missing S288C ORFs

XII	NTS1-2 , NTS2-1 , NTS2-2	Change feature_type/SO_term from SO:0001637 rRNA_gene to SO:0000183 non_transcribed_region	
XIII	LDO45/YMR147W	Shift stop to be same as LDO16/YMR148W, add intron: old coordinates 559199..559870 new coordinates 559199..559780, 560156..560812	Eisenberg-Bord et al. 2018
XIII	YMR008C-A	New ORF: coordinates 283081..283548 Crick	Internal reanalysis of results from Song et al. 2015 to find and annotate missing S288C ORFs
XIII	YNCM0001W aka PHO84 lncRNA	New ncRNA: coordinates 23564..26578	Camblong et al. 2007
XIV	LTO1/YNL260C	Move start to Met37: old coordinates 156859..157455 Crick new coordinates 156859..157347 Crick	Paul et al. 2015
XVI	YNCP0002W aka GAL4 lncRNA	New ncRNA: coordinates 79562..82648	Geisler et al. 2012

618
619

Table 2. “Legacy” gene names which predate the database have been updated to be more software-friendly by removing unnecessary punctuation.

ORF	Old gene name	New gene name
YGL234W	ADE5,7	ADE57
YER069W	ARG5,6	ARG56
YBR208C	DUR1,2	DUR12
YIL154C	IMP2'	IMP21

Fig. 1. SGD allele pages include (A) Overview with name, affected gene, type of allele, and a description of sequence change and/or domain mutated; phenotype annotations; (B) genetic interaction annotations; and (C) Shared Alleles network diagram depicting shared phenotypes and interactions with other alleles.

Fig. 2. (A) SGD provides links between SGD gene pages and those for orthologous genes at the Alliance of Genome Resources. Hexagonal icons represent each organism (human, mouse, rat, zebrafish, fly, worm, and yeast) for which there is gene information at the Alliance. (B) Curated information from SGD is presented on yeast gene pages at the Alliance of Genome Resources. The various types of yeast data that can be found at the Alliance can be viewed using the menu on the left side of the gene page, and include orthologs, functional annotations using the Gene Ontology, cellular pathways, phenotypes, disease associations, alleles, variants, sequence features, expression, and interactions. Data can be downloaded in bulk from <https://www.alliancegenome.org/downloads>.

Fig. 3. SGD homology pages (A) include an Overview of general information about the yeast gene, with links to homologous gene pages at the Alliance; (B) information about known homologs including species corresponding Gene ID, and name of the homolog; and (C) data about cross-species functional complementation between yeast and other species. Also included on the page, but not shown here, are Fungal Homologs, gene identifiers in other databases, and links to external resources.

Fig. 4. SGD disease pages use the Disease Ontology to describe human diseases that are associated with yeast homologs, and include (A) disease annotation summary and individual annotations; and (B) shared annotation network that depicts shared disease annotations for other yeast genes and their human homologs.

A

Summary

Literature

Allele: **cdc28-13**

Overview

Phenotype Annotations

Interaction Annotations

Shared Alleles

References

Resources

Allele: **cdc28-13**

Allele Name:

cdc28-13¹

SGD ID:

SGD:S000277751

Affected Gene:

CDC28

Allele Type:

missense variant²

Description:

R283Q²

Phenotype Annotations

15 entries for 11 phenotypes

Filter table

Gene	Phenotype	Experiment Type	Mutant Information	Strain Background	Chemical	Details	Reference
CDC28	actin cytoskeleton morphology: abnormal	classical genetics	conditional Allele: cdc28-13	Other		Temperature: elevated temperature, 37 °C Details: actin cable length increased from 4.3 μm to 8.2 μm at the restrictive temperature, as cable length scales with cell length	McInally SG, et al. (2021) PMID:34114567
CDC28	cell cycle progression in G1 phase: arrested	classical genetics	conditional Allele: cdc28-13	W303		Temperature: elevated temperature, 37 °C Details: 96% of cells unbudded and 95% viable	Buscemi G, et al. (2000) PMID:10704371

B

Allele: **cdc28-13**

Overview

Phenotype Annotations

Interaction Annotations

Shared Alleles

References

Resources

Genetic Interaction Annotations

178 entries for 147 genes

Filter table

	Interactor	Interactor	Allele	Assay	Annotation	Action	Phenotype	SGA score	P-value	Reference
	CDC28	ACT1	cdc28-13, act1-159	Negative Genetic	high-throughput	Bait-Hit		-0.1657	0.02865	Costanzo M, et al. (2016) PMID:27708008
	CDC28	ADK1	adk1-Δ, cdc28-13	Negative Genetic	high-throughput	Hit-Bait		-0.1209	0.003639	Costanzo M, et al. (2016) PMID:27708008
	CDC28	ALG2	cdc28-13, alg2-1	Negative Genetic	high-throughput	Bait-Hit		-0.1603	0.000007189	Costanzo M, et al. (2016) PMID:27708008
	CDC28	ALT1	cdc28-13, alt1-Δ	Negative Genetic	high-throughput	Bait-Hit		-0.1801	0.009543	Costanzo M, et al. (2016) PMID:27708008
	CDC28	APC11	cdc28-13, apc11-13	Positive Genetic	high-throughput	Bait-Hit		0.161	0.02863	Costanzo M, et al. (2016)

bioRxiv preprint doi: <https://doi.org/10.1101/2021.09.16.460706>; this version posted September 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

C

Allele: **cdc28-13**

Overview

Phenotype Annotations

Interaction Annotations

Shared Alleles

References

Resources

Shared Alleles

SGD 2021-09-15

● FOCUS ● ALLELE ● POSITIVE INTERACTION ● NEGATIVE INTERACTION ● PHENOTYPE

All Positive Interactions Negative Interactions Phenotypes

Download (.png)

B

bioRxiv preprint doi: <https://doi.org/10.1101/2021.09.16.460706>; this version posted September 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Drosophila Sod1 76 Screenshot

A

Summary
Sequence
Protein
Gene Ontology
Phenotype
Disease
Interactions
Regulation
Expression
Literature
Homology

CDC28 / YBR160W Homology

Standard Name: CDC28^{1 2}

Systematic Name: YBR160W

SGD ID: SGD:S000000364


Aliases: CDK1, HSL5, SRM5⁶³

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 Cycle⁴

Comparative Info:



B

Homologs 12 entries for 12 homologs

Species	Gene ID	Gene name	Source
<i>Caenorhabditis elegans</i>	WB:WBGene00000405	cdk-1	Alliance
<i>Danio rerio</i>	ZFIN:ZDB-GENE-010320-1	cdk1	Alliance
<i>Danio rerio</i>	ZFIN:ZDB-GENE-040426-2741	cdk2	Alliance
<i>Drosophila melanogaster</i>	FB:FBgn0004106	Cdk1	Alliance
<i>Drosophila melanogaster</i>	FB:FBgn0004107	Cdk2	Alliance
<i>Homo sapiens</i>	HGNC:1722	CDK1	Alliance
<i>Homo sapiens</i>	HGNC:1771	CDK2	Alliance
<i>Homo sapiens</i>	HGNC:1772	CDK3	Alliance
<i>Mus musculus</i>	MGI:88351	Cdk1	Alliance
<i>Mus musculus</i>	MGI:104772	Cdk2	Alliance

C


Functional Complementation 12 entries for 12 complements

Species	Gene ID	Strain background	Direction	Details	Source	Reference
<i>Homo sapiens</i>	HGNC:1722	S288C	other complements yeast	Human CDK1 complements the thermosensitivity of the yeast cdc28-4 mutant as well as the inviability of the cdc28 null mutant.	SGD	Elledge SJ and Spottswood MR (1991) PMID:1714386
<i>Homo sapiens</i>	HGNC:1722	S288C	other complements yeast	Human CDK1 complements the thermosensitivity of the yeast cdc28-13 mutant as well as rescuing the inviability of the null mutant.	SGD	Ninomiya-Tsuji J, et al. (1991) PMID:1717994
<i>Homo sapiens</i>	HGNC:1722	S288C	other complements yeast	Human CDK1 complements the inviability of the <i>S. cerevisiae</i> cdc28 null mutant.	SGD	Mayi T, et al. (2015) PMID:25541464
<i>Homo sapiens</i>	HGNC:1722	S288C	other complements yeast	Human CDK1 complements the inviability of the <i>S. cerevisiae</i> cdc28 null mutant.	SGD	Hamza A, et al. (2015) PMID:26354769

A

Summary:

Yeast STE24 is homologous to human ZMPSTE24, and has been used to study type 2 diabetes, non-alcoholic fatty liver disease, the premature aging disease Hutchinson-Gilford progeria syndrome (HGPS), and the related progeroid disorders mandibuloacral dysplasia type B (MAD-B) and restrictive dermopathy (RD)

 Download All Annotations (.txt)

Manually Curated 10 entries for 4 Disease Ontology terms ?

? Filter table

Disease Ontology Term	Qualifier	Evidence	Source	Assigned On	Reference
progeria		ISS with ZMPSTE24	SGD	2018-04-25	Barrowman J, et al. (2012) PMID:22718200
progeria		IGI with ZMPSTE24	SGD	2018-04-25	Barrowman J, et al. (2012) PMID:22718200
progeria		ISS with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150
progeria		IGI with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150
type 2 diabetes mellitus		ISS with ZMPSTE24	SGD	2021-08-13	Kayatekin C, et al. (2018) PMID:29526462
type 2 diabetes mellitus		IGI with ZMPSTE24	SGD	2021-08-13	Kayatekin C, et al. (2018) PMID:29526462
lethal restrictive dermopathy		ISS with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150
lethal restrictive dermopathy		IGI with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150
non-alcoholic fatty liver disease		ISS with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150
non-alcoholic fatty liver disease		IGI with ZMPSTE24	SGD	2021-08-31	Spear ED, et al. (2018) PMID:29794150

bioRxiv preprint doi: <https://doi.org/10.1101/2021.09.16.460706>; this version posted September 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

B

Shared Annotations ?

