The Mouse Genome Database (MGD)

Eppig J.T., et al. (2005). The Mouse Genome Database (MGD): From genes to mice: A community resource for mouse biology. Nucleic Acids Research 33(1), D471-D47

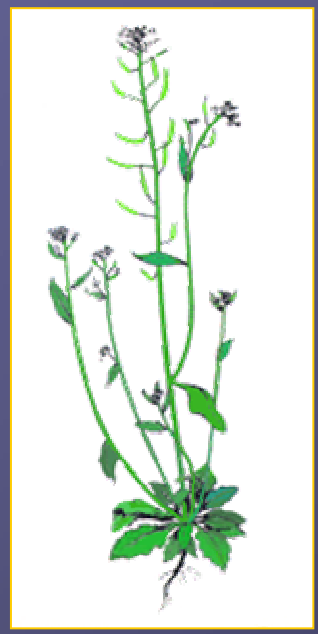
SILS Biomedical Informatics Journal Club

http://ils.unc.edu/bioinfo/

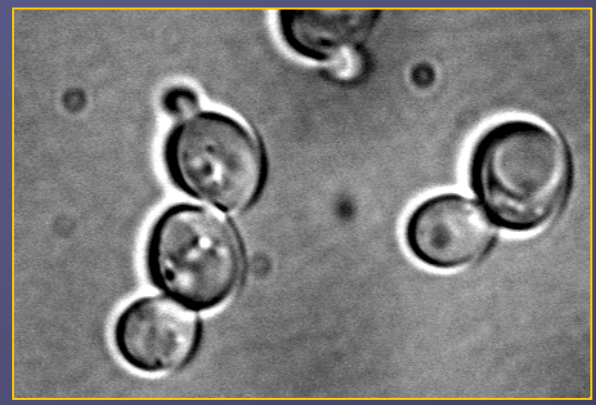
2005-09-13

Model organisms

Mus musculus [src]

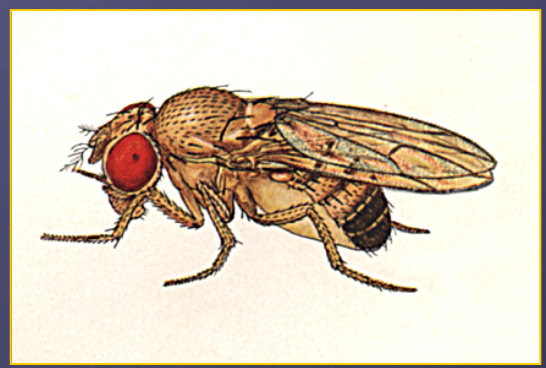


Arabidopsis thaliana [src]



Saccharomyces cerevisiae [src]

Dictyostelium discoideum [src]



Drosophila melanogaster [src]

Caenorhabditis elegans [src]

What's in a MOD?

- Evolution from simple repositories of gene and genome sequences to comprehensive, richly annotated resources
- Integration of sequence data with structure and function data
- Genotype and phenotype data
- Linkages to underlying experimental literature
- Linkages to orthologous genes in other organisms

MGD content

Table 1. Snapshot of data content in MGD: October 7, 2004

MGD data statistics	October 7, 2004
Number of genes with sequence data	28 287
Number of genes (including unmapped mutants)	33 207
Number of markers (including genes)	57 521
Number of markers mapped	53 082
Number of genes with links to Swiss-Prot	7769
Number of genes with GO annotations	15 309
Number of mouse/human curated orthologies	14 893
Number of mouse/rat curated orthologies	12 679
Number of genes with one or more phenotypic alleles	4996
Number of cataloged phenotypic alleles	10 949
Number of references	87 527
Number of mouse nucleotide sequences integrated into the MGI system (includes ESTs)	>7 600 000

‘evidence-based’ curation

- “MGD is updated on a daily basis by biologists on our curatorial staff who scan the current scientific literature, extract relevant data, and enter it in MGD.” [src]
- “All assertions of orthology are supported by a statement of evidence and a citation” [D473].