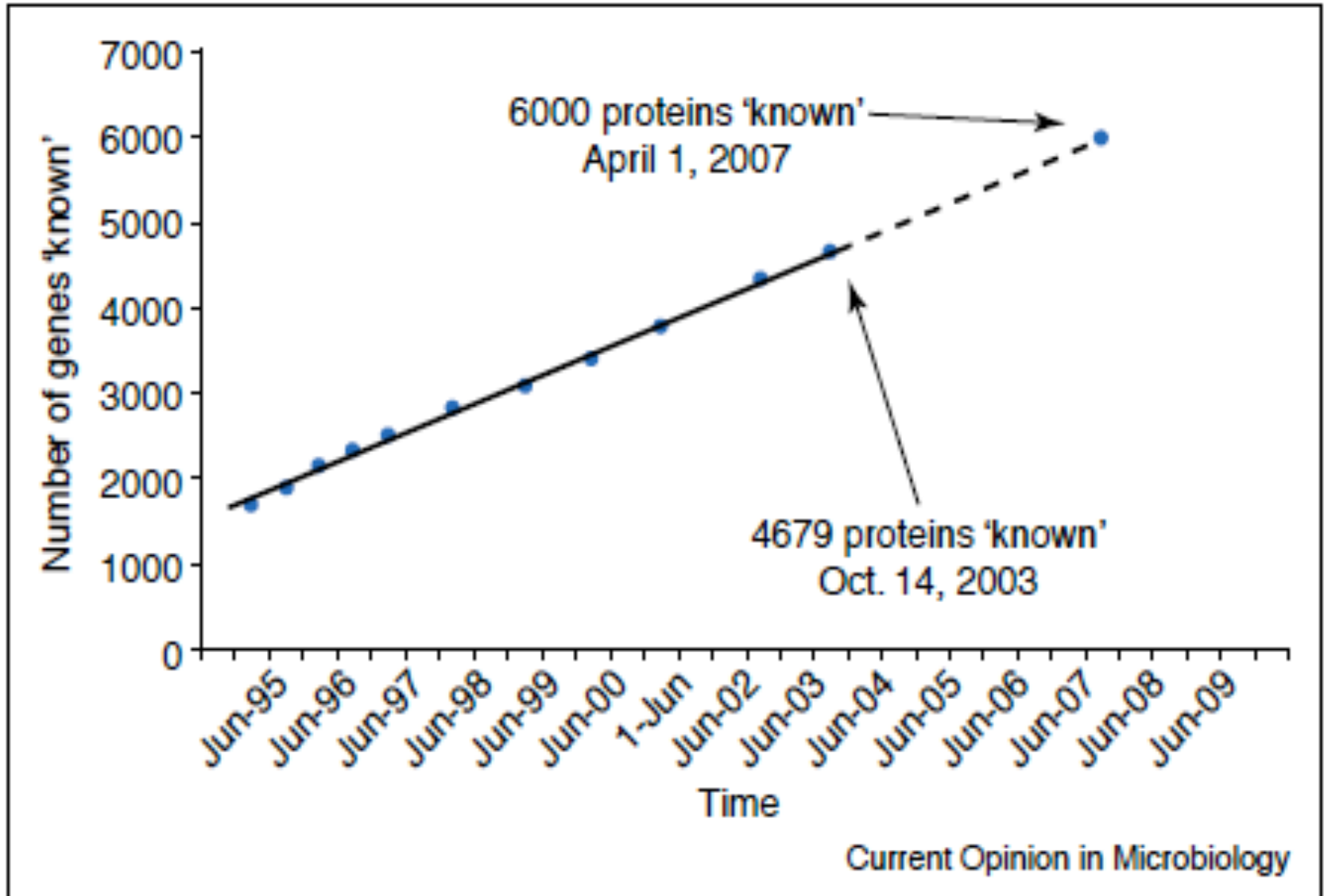# When will we know the functions of all genes in an organism?

Mark Johnston's joke answer from 2003

Saccharomyces Genome Database

SGD^{TM} 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

- **Expanding the View of the Functional Budding Yeast Genome** *- February 25, 2011*

    In an effort to provide a comprehensive view of all of the sequence-based functional elements in the Saccharomyces cerevisiae genome, SGD will collect sequence-based functional feature annotations from published datasets, aimed at charting the yeast transcriptome (e.g. protein-coding genes, non-coding RNAs), chromatin landscapes (e.g. nucleosome phasing, histone modifications and variations), as well as cataloging regulatory sequence elements (e.g. transcription factor binding sites, splicing signals) and more.

    We have upgraded our GBrowse genome viewer to allow users to quickly and easily browse this information-rich view of the yeast genome. In addition, 17 new data tracks have been added, including recent surveys of the budding yeast transcriptome (Nagalakshmi et al. 2008; Xu et al. 2009; Yassour et al. 2009) and catalogs of regions favored for recombination and replication events (Mancera et al. 2008; Buhler et al. 2007; Xu et al. 2006; Eaton et al. 2010). We invite authors to work with us to integrate their data into our GBrowse viewer pre- and/or post-publication as we move forward. Watch for the regular addition of new tracks to SGD's GBrowse in the future! Please contact us if you are interested in participating or have questions and comments.

DFG10/YIL049W Summary

Help

DFG10 BASIC INFORMATION

Standard Name	*DFG10* [1]
Systematic Name	YIL049W
Feature Type	ORF, Verified
Description	Probable polyprenol reductase that catalyzes conversion of polyprenol to dolichol, the precursor for N-glycosylation; involved in filamentous growth; mutations in human ortholog SRD5A3 confer CDG (Congenital Disorders of Glycosylation) ([1], [2])
Name Description	Defective for Filamentous Growth [1]

GO Annotations

All DFG10 GO evidence and references

View Computational GO annotations for DFG10

Molecular Function
Manually curated
- 3-oxo-5-alpha-steroid 4-dehydrogenase activity (ISS)

Biological Process
Manually curated
- dolichol biosynthetic process (IMP)
- pseudohyphal growth (IMP)

Cellular Component
Manually curated
- cellular_component unknown (ND)

Mutant Phenotype

All DFG10 Phenotype details and references

Classical genetics
null
- carboxypeptidase Y (Prc1p) modification: decreased

unspecified
- pseudohyphal growth: absent

Large-scale survey
null
- competitive fitness: decreased
- endocytosis: decreased
- resistance to cycloheximide: decreased
- viable

Interactions

DFG10 All interactions details and references

View additional details at BioGRID

10 total interaction(s) for 10 unique genes/features.

Physical Interactions
- Affinity Capture-RNA: 1

Genetic Interactions
- Negative Genetic: 1
- Phenotypic Suppression: 3
- Positive Genetic: 1
- Synthetic Growth Defect: 2
- Synthetic Haploinsufficiency: 1
- Synthetic Lethality: 1

DFG10 RESOURCES

Click on map for expanded view

SGD ORF map

259000 to 264000
PCL7 DFG10 NEO1
5'   5'
IX

GBrowse

chrIX:259158..2619
260k
YIL050W   YIL048
YIL049W
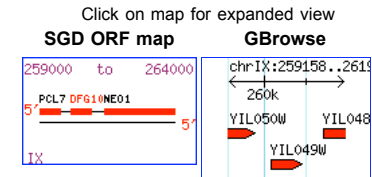
- **Literature**
  Literature Guide   View

- **Retrieve Sequences**
  Genomic DNA   View

- **Sequence Analysis Tools**
  BLASTP   View

- **Protein Info & Structure**
  Protein Info
  View

- **Localization Resources**
  YeastGFP DB (UCSF) at SGD   View

- **Interactions**
  BioGRID (Toronto)   View

- **Phenotype Resources**
  PROPHECY   View

- **Maps & Displays**
  Chromosomal Features Map   View

- **Comparison Resources**
  Fungal Alignment   View

- **Functional Analysis**
  Expression Summary   View

Click on histogram for expression summary
Expression Summary

Number of Experiments vs. Log₂ Ratios
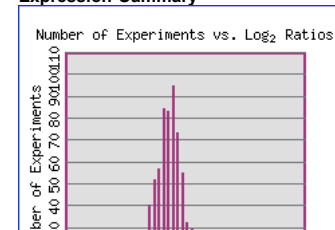
Community Info   Submit Data   BLAST   Primers   PatMatch   Gene/Seq Resources   Advanced Search   Community Wiki

DFG10/YIL049W Gene Ontology Annotations   Help

Summary   Locus History   Literature   Gene Ontology   Phenotype   Interactions   Expression   Protein   Wiki

This page displays GO annotations in different sections according to the annotation method used to add that annotation to SGD.

DFG10 Manually curated*:   Jump to: top | High-throughput | Computational
Last Reviewed on: 2011-03-22   Molecular Function | Biological Process | Cellular Component

Manually curated Molecular Function

Annotation(s)	Evidence	Reference(s)	Assigned By
3-oxo-5-alpha-steroid 4-dehydrogenase activity	ISS: Inferred from Sequence or structural Similarity with EBI:Q9H8P0 *Assigned on 2011-03-22*	**Cantagrel V, *et al.* (2010)** SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* 142(2):203-17 SGD Curated Paper PubMed Access Full Text	SGD

Manually curated Biological Process

Annotation(s)	Evidence	Reference(s)	Assigned By
dolichol biosynthetic process	IMP: Inferred from Mutant Phenotype *Assigned on 2011-02-01*	**Cantagrel V, *et al.* (2010)** SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* 142(2):203-17 SGD Curated Paper PubMed Access Full Text	SGD
pseudohyphal growth	IMP: Inferred from Mutant Phenotype *Assigned on 2002-09-27*	**Mosch HU and Fink GR (1997)** Dissection of filamentous growth by transposon mutagenesis in Saccharomyces cerevisiae. *Genetics* 145(3):671-84 SGD Curated Paper PubMed Access Full Text	SGD

Manually curated Cellular Component

Annotation(s)	Evidence	Reference(s)	Assigned By
cellular_component unknown	ND: No Biological Data Available *Assigned on 2002-09-27*	**SGD (2002)** Use of the ND evidence code for Gene Ontology (GO) terms in SGD () SGD Paper	SGD

* Manually curated GO annotations reflect our best understanding of the basic molecular function, biological process, and cellular component for this gene product. Manually curated annotations are assigned by SGD curators based on published papers when available, or by curatorial statements if necessary. Curators periodically review all Manually curated GO annotations for accuracy and completeness. The "Last Reviewed on:" date at the top of this section indicates when these annotations were last reviewed.

DFG10 High-throughput**:   Jump to: top | Manually curated | Computational

There are no High-throughput annotations for DFG10

** GO annotations from High-throughput experiments are made based on a variety of large scale high-throughput experiments, including genome-wide experiments. Many of these annotations are made based on GO annotations (or mappings to GO annotations) assigned by the authors, rather than SGD curators. While SGD curators read these publications and often work closely with authors to incorporate the information, each individual annotation may not necessarily be reviewed by a curator. GO Annotations from high-throughput experiments will be assigned only when this type of data is available, and thus may not be assigned in all three aspects of the Gene Ontologies.

DFG10 Computational***:   Jump to: top | Manually curated | High-throughput
Molecular Function | Biological Process | Cellular Component

Computational Molecular Function

Table 1

197 protein-coding genes whose existence is supported by expression and/or conservation over evolution, but which are completely uncharacterized on Saccharomyces Genome Database and which are not present in any of the major yeast functional genomics data sets analyzed here.

YAL016C-B	YGL218W	YLR154W-F	YOL166W-A
YAL037C-A	YGL258W-A	YLR156C-A	YOR011W-A
YAL037W	YGL262W	YLR157C-C	YOR012W
YAL063C-A	YGR035W-A	YLR157W-A	YOR020W-A
YAL064C-A	YGR121W-A	YLR157W-C	YOR032W-A
YAL067W-A	YGR127W	YLR159C-A	YOR034C-A
YAR035C-A	YGR146C-A	YLR159W	YOR072W-B
YAR068W	YGR169C-A	YLR161W	YOR161C-C
YBL008W-A	YGR174W-A	YLR162W	YOR192C-C
YBL039W-A	YGR204C-A	YLR162W-A	YOR293C-A
YBL071C-B	YGR240C-A	YLR264C-A	YOR316C-A
YBL101W-C	YHL015W-A	YLR285C-A	YOR338W
YBL108C-A	YHL048C-A	YLR307C-A	YOR376W-A
YBL112C	YHR007C-A	YLR312C-B	YOR381W-A
YBR056W-A	YHR022C-A	YLR342W-A	YOR394C-A
YBR072C-A	YHR050W-A	YLR361C-A	YPL038W-A
YBR182C-A	YHR086W-A	YLR406C-A	YPL039W
YBR196C-A	YHR175W-A	YLR412C-A	YPL119C-A
YBR196C-B	YHR199C-A	YLR466C-B	YPL152W-A
YBR200W-A	YHR212W-A	YML003W	YPL189C-A
YBR221W-A	YHR213W-A	YML054C-A	YPR089W
YBR296C-A	YHR213W-B	YML100W-A	YPR108W-A
YBR298C-A	YHR214C-D	YMR001C-A	YPR159C-A
YCL001W-A	YHR214C-E	YMR013W-A	
YCL012C	YIL002W-A	YMR030W-A	
YCL047C	YIL014C-A	YMR105W-A	

The Yeast Deletion Collection: 6000 Strains
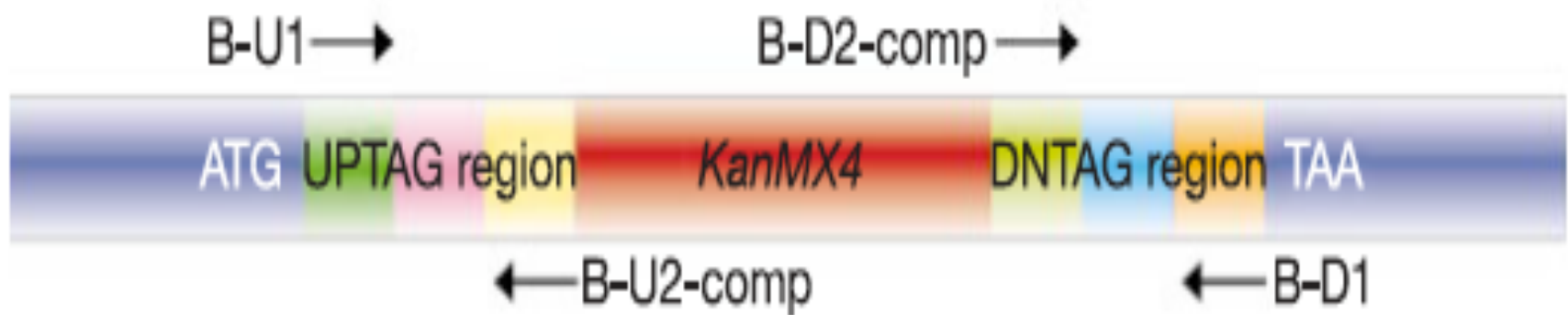
Figure 1 The *KanMX* deletion cassette module. The biotin-labelled, deletion-specific primers (B-U1, B-U2-comp, B-D1 and B-D2-comp; see Methods for structure) are used to amplify the unique UPTAG and DNTAG sequences from genomic preparations generated in the fitness-profiling studies.

How can we learn about functions of unknown or poorly characterized genes/proteins?
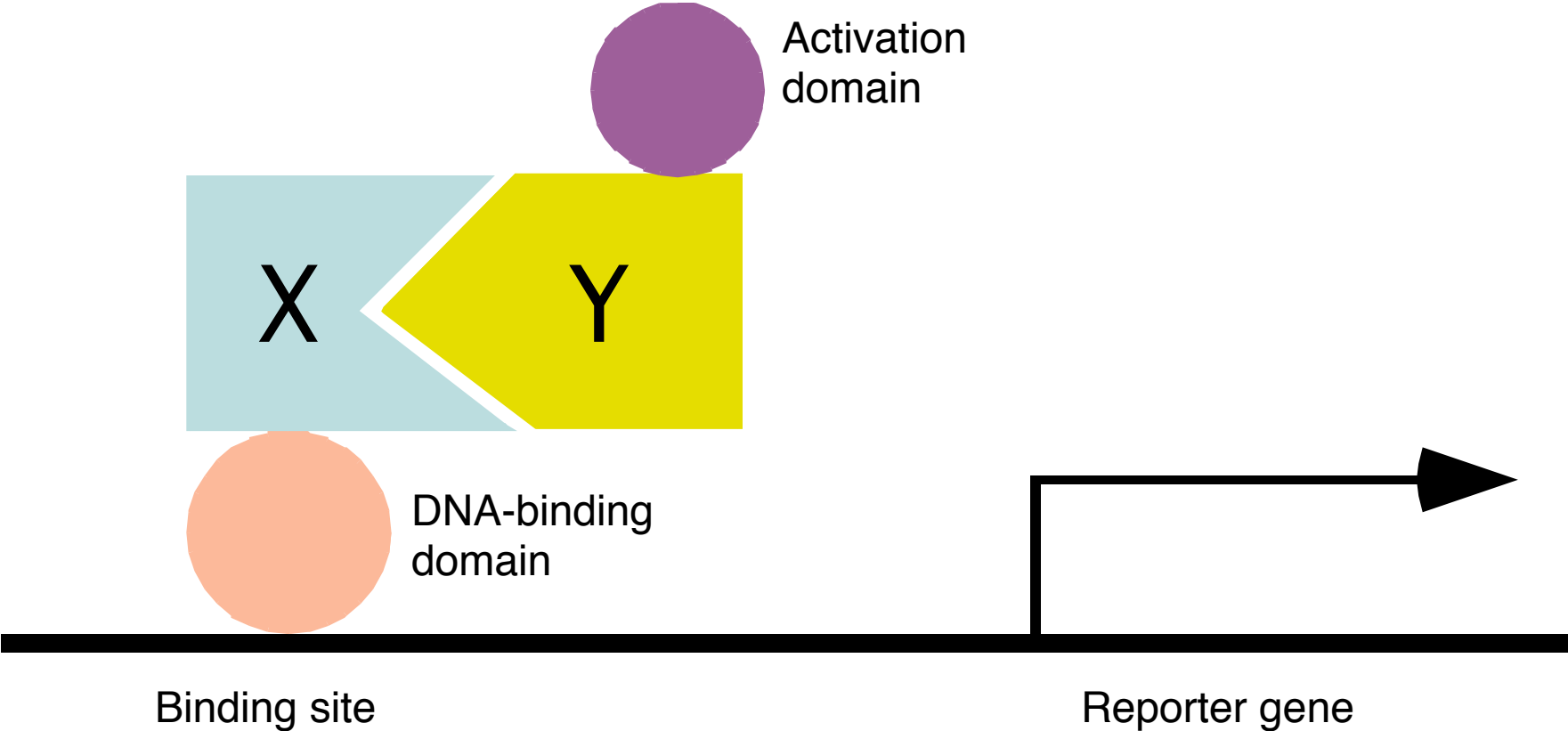
Protein interaction approaches

2-hybrid screens

Mass spectrometry

Genetic interaction approaches

Synthetic lethality (or other synthetic interactions)

Competitive fitness

Two-Hybrid System to Detect Protein Interactions

Activation
domain

X Y

DNA-binding
domain

Binding site

Reporter gene

2-hybrid tests

Create plasmid-borne version of protein of interest (bait) fused to DNA binding domain

Screen library of plasmids in which random bits of genomic DNA have been fused to a transcription activation domain. Transform the library into the strain harboring the bait plasmid and look for expression of the reporter.

Alternatively, mate the strain carrying the bait plasmid to an ordered array of all 6000 yeast genes, each fused to a transcription activation domain, and select for diploids carrying both plasmids. Again, look for expression of the reporter.
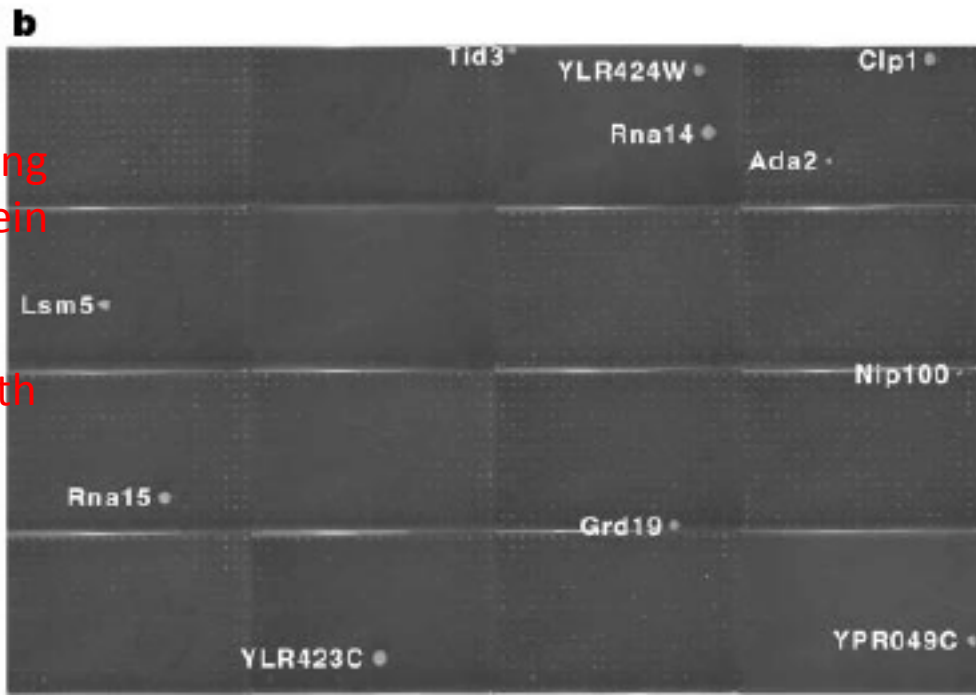
Desk-top Robot

Footprint about the size of 2 waffle irons.
Capable of pinning in 96, 384, and 1536 formats.
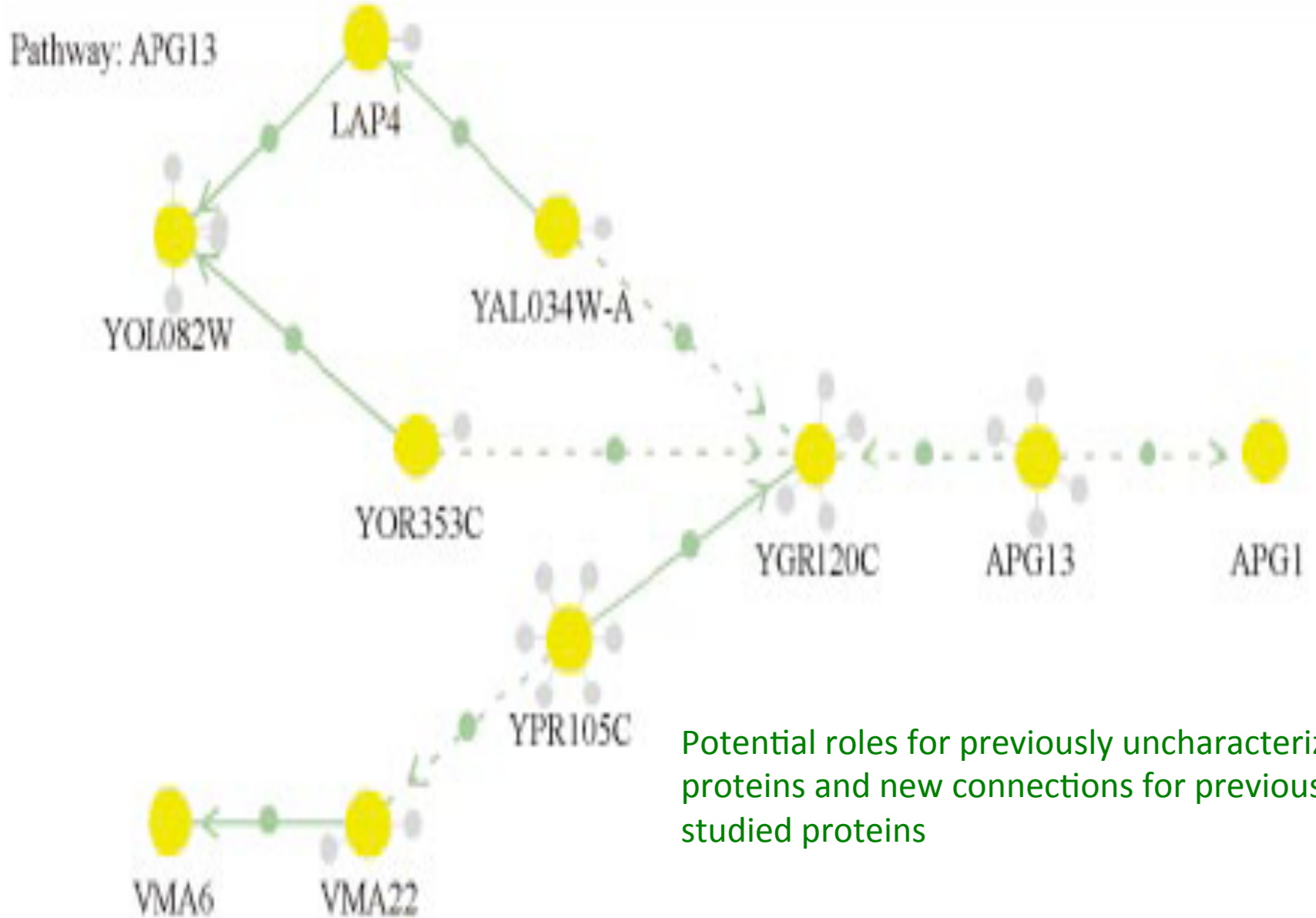
Sample 2-hybrid test using ordered array
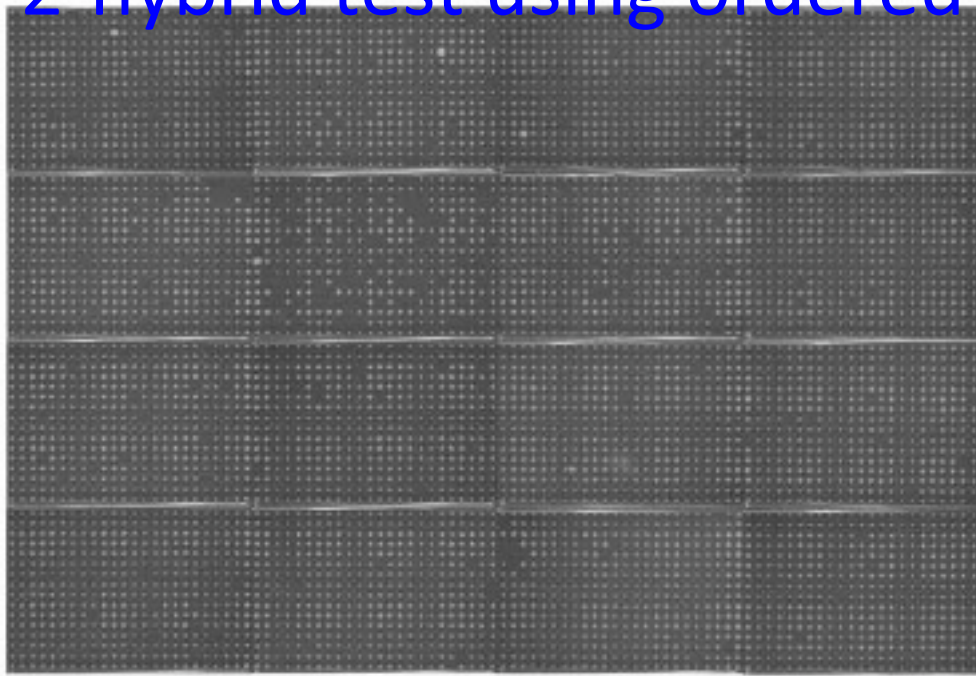
6000 transformants, each expressing a different ORF fused to the GAL4-AD

Mate to strain expressing GAL4-BD-Pcf11, a protein involved in pre-mRNA cleavage and polyadenylation. Growth indicates presumptive interaction.

Reporter is a nutritional gene required for growth

New participants in autophagy revealed by 2-hybrid screens

Pathway: APG13

LAP4
YAL034W-A
YOL082W
YOR353C
YGR120C
APG13
APG1
YPR105C
VMA6
VMA22

Potential roles for previously uncharacterized proteins and new connections for previously studied proteins

Mass Spectrometry

Take purified proteins, digest with trypsin, and perform mass spec. Because the sequence of the genome is known, the mass of all tryptic peptides can be predicted and compared to what is observed.
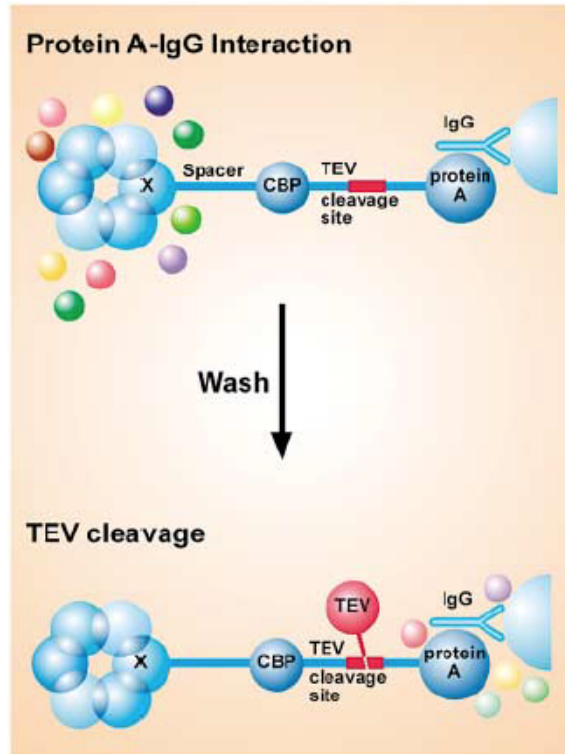
What would you do next, once candidate interacting proteins are ID'd by mass spec?
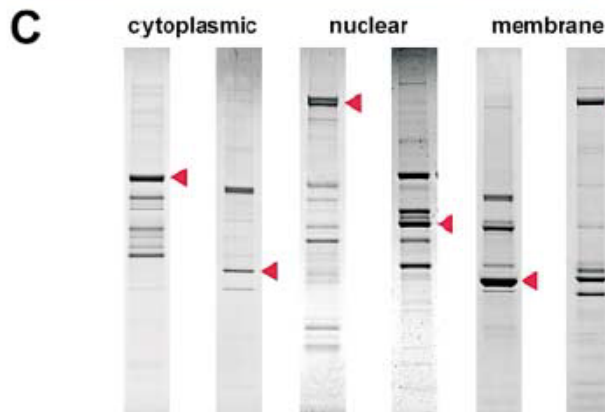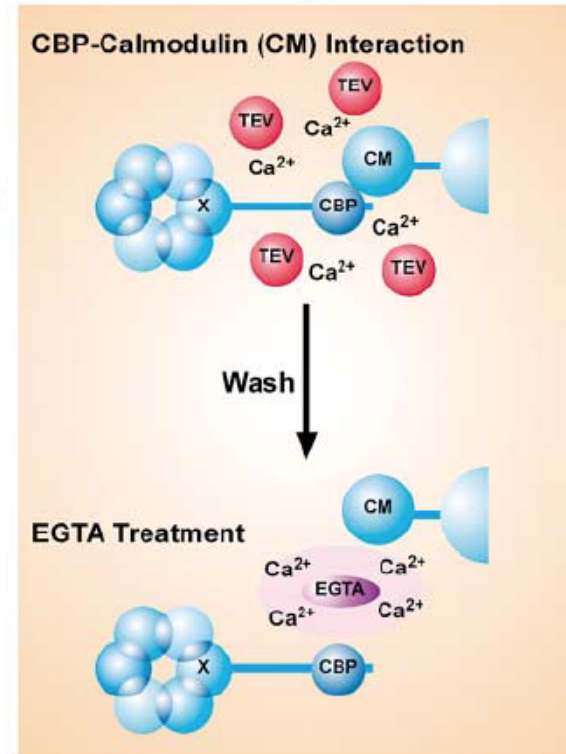
TAP Tag Mass Spec

A X — CBP — TEV cleavage site — protein A

B

First affinity purification

Protein A-IgG Interaction

Spacer, CBP, TEV cleavage site, protein A, IgG

Wash

TEV cleavage

X, CBP, TEV cleavage site, TEV, protein A, IgG

Second affinity purification

CBP-Calmodulin (CM) Interaction

TEV, Ca²⁺, CM, CBP, Ca²⁺, TEV, Ca²⁺, TEV

Wash

EGTA Treatment
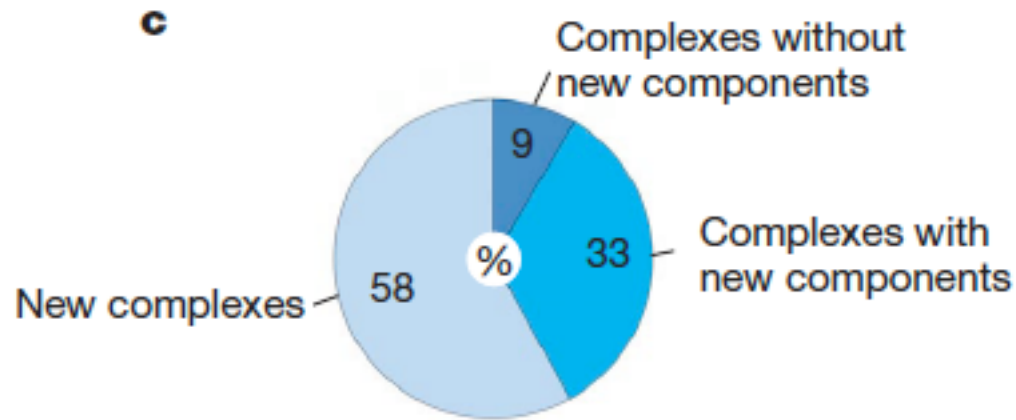
CM, Ca²⁺, Ca²⁺, EGTA, Ca²⁺, Ca²⁺, X, CBP
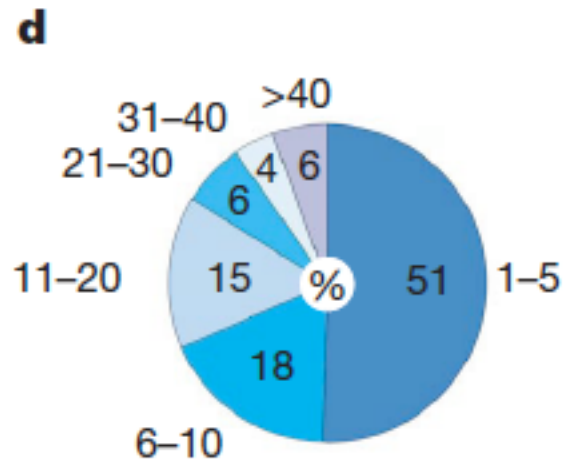
C cytoplasmic   nuclear   membrane

Some examples, unusually good examples.

Red arrows indicate baits.

Large scale mass spec experiment using 589 different baits

c

Complexes without new components — 9%

Complexes with new components — 33%

New complexes — 58%

Novelties in complexes

d

>40 — 6%

31–40 — 4%

21–30 — 6%

11–20 — 15%

6–10 — 18%

1–5 — 51%

Number of proteins per complex

e

Transcription/DNA maintenance/chromatin structure — 24%

Cell cycle — 6%

Cell polarity and structure — 3%

Intermediate and energy metabolism — 19%

Membrane biogenesis/turnover — 9%

Protein synthesis/turnover — 14%

Protein/RNA transport — 5%

RNA metabolism — 12%

Signalling — 9%
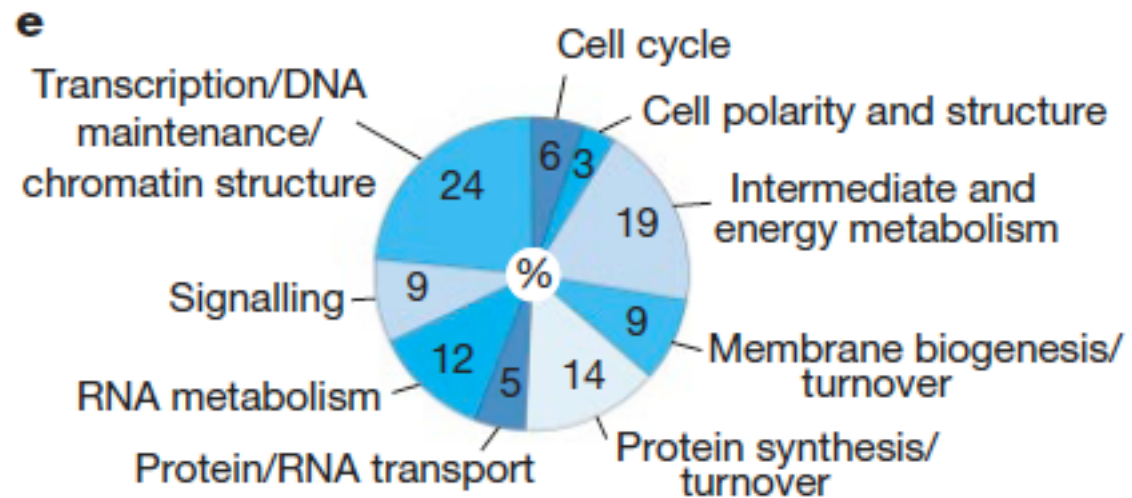
Distribution of complexes according to function