Plans for Today

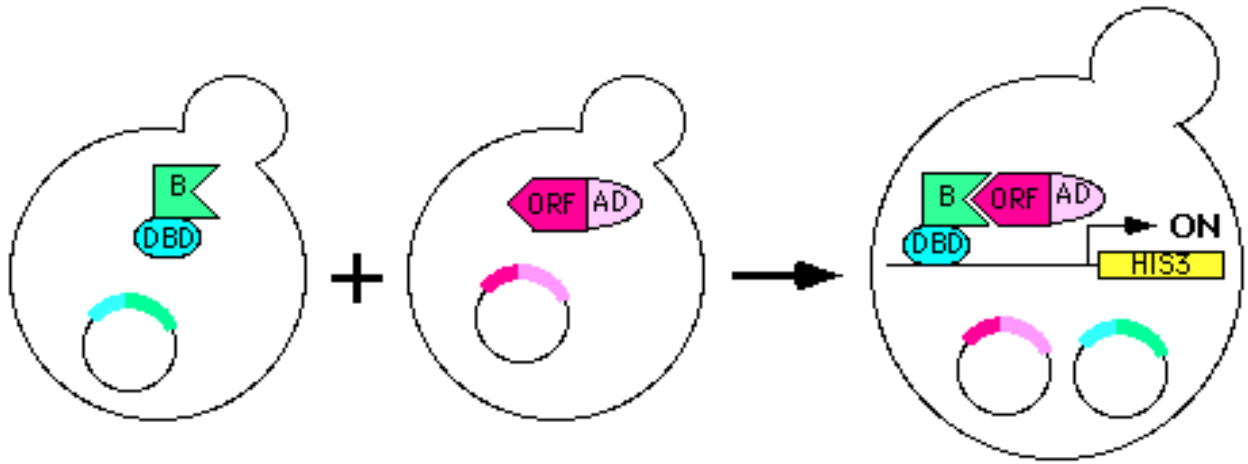
- Protein-Protein interaction
  - Physical interaction
    - Immunoprecipitation (IP)
    - Yeast two-hybrid (Y2H)
- Proteome-wide high throughput strategy
  - Binary complex
    - Yeast two-hybrid (Nature 2000)
    - Complementation system (LCI, PNAS 2004)
  - Mass Spectrometry (multiple complex)
    - TAP strategies (Nature 2002)
    - Immunoprecipitation (Nature 2002)
  - Biochemical Assay
    - Pooling strategy (Science 1999)
  - Protein array
    - Protein microarray (Science 2000)

Protein-protein interaction

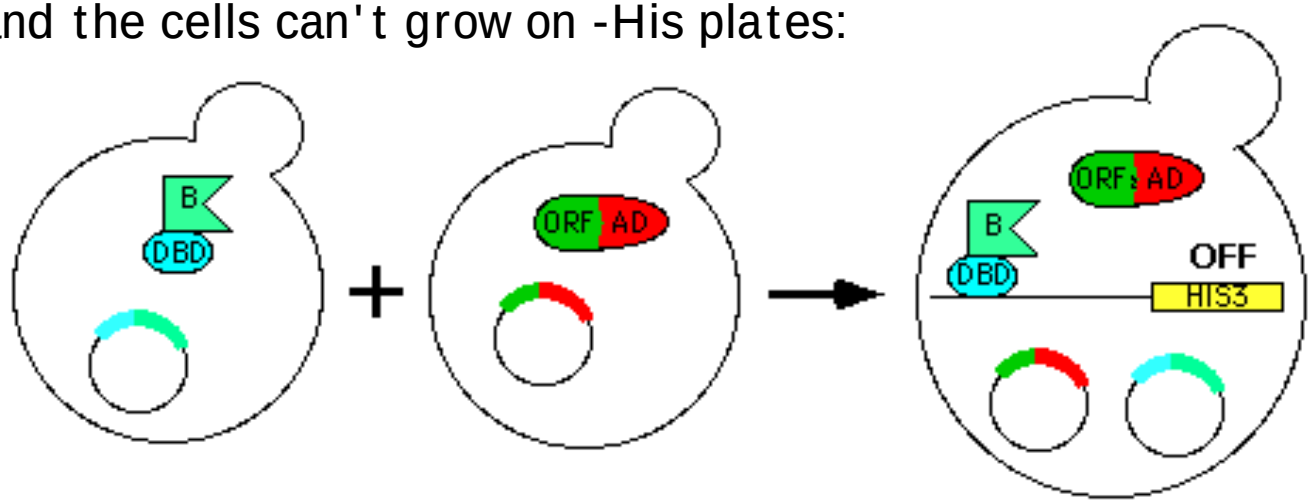
- Guilt-by-association
  - Prediction from sequence
    - In *silico* analysis
      - Protein A from species A: domain 1 and 2
      - Orthologous domains of species B
        - Protein 1' and protein 2'
      - Protein 1' interact with protein 2' in species B and function as Protein A of species A
      - Recognition ≠ sequence homology ?
  - Physical interactions
    - Yeast two-hybrid screen of whole genome
      - Binary interaction
    - Protein Chip
      - Complementary interaction
    - Tagged protein
      - Multi-protein complex
        - Tandem affinity purification (TAP) + MS
        - Immunoprecipitation + MS

Y2H: basic design

- If the two proteins interact (B + ORF), the reporter gene (here: HIS3) is switched on and the diploids can grow on -His plates:



- If the two proteins don't interact, the reporter gene remains inactive and the cells can't grow on -His plates:



Discovery Q

- Y2H protein interaction occurs inside nucleus of yeast cell. Is it OK?
- What is the proper control?
- Is it restricted to yeast proteins only?

Large-scale Y2H

- Yeast genome
  - Methods
    - Array screening
      - Prey: 6000 yeast strains each expressing a different molecule
      - Bait: 192 yeast strains (orthologues)
      - Much more time- and labor- intensive
      - Can rapidly identify the locations producing false-positive interaction
      - More positive identification (48 interactions/12 bait vs. 14/12)
    - Library screening
      - Prey: a pool of 6000 yeast strains (activation domain library)
      - Bait: 6000 yeast strains (64 x 96-well plate = 6144)
      - Hybrid cells are selected then screened for positive interactions
      - Reasonable time and effort
  - Results
    - Bioinfomatics platform for data analysis
      - http://portal.curagen.com
    - Database of Interacting Proteins (DIP)
      - http://dip.doe-mbi.ucla.edu/

Uetz, et al., *Nature* 2000 403, 623-627.
Schwikowski et al., 2000 *Nature Biotech*. 18, 1257-1261.

Uetz, et al., *Nature* 2000 403, 623-627.

Y2H Strategy Comparison

- Array vs. Library screen

Table 1 Summary of experimental results

45 % (a) Protein array — More stringent screen

Description	Total
Yeast ORFs screened	192
Yeast ORF yielding reproducible interactions	87
Total discrete interacting protein pairs (reproduced in a second screen)	281

Non-random choice ←

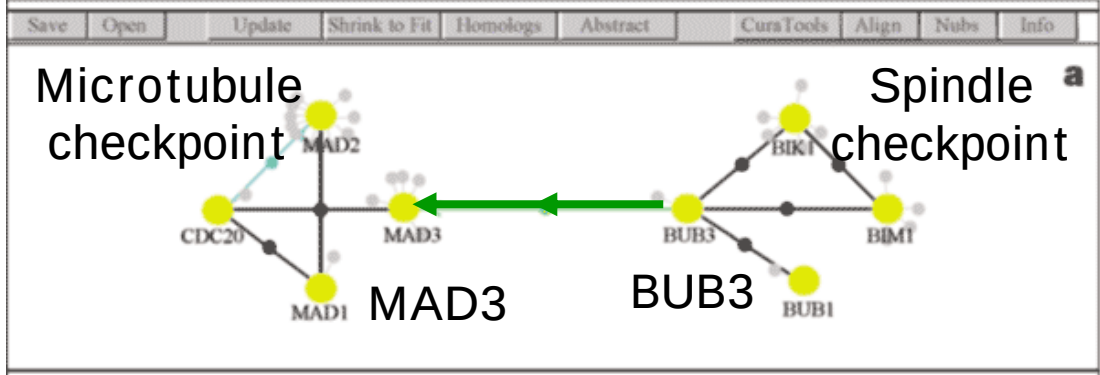
8 % (b) High-throughput screens — Less stringent screen

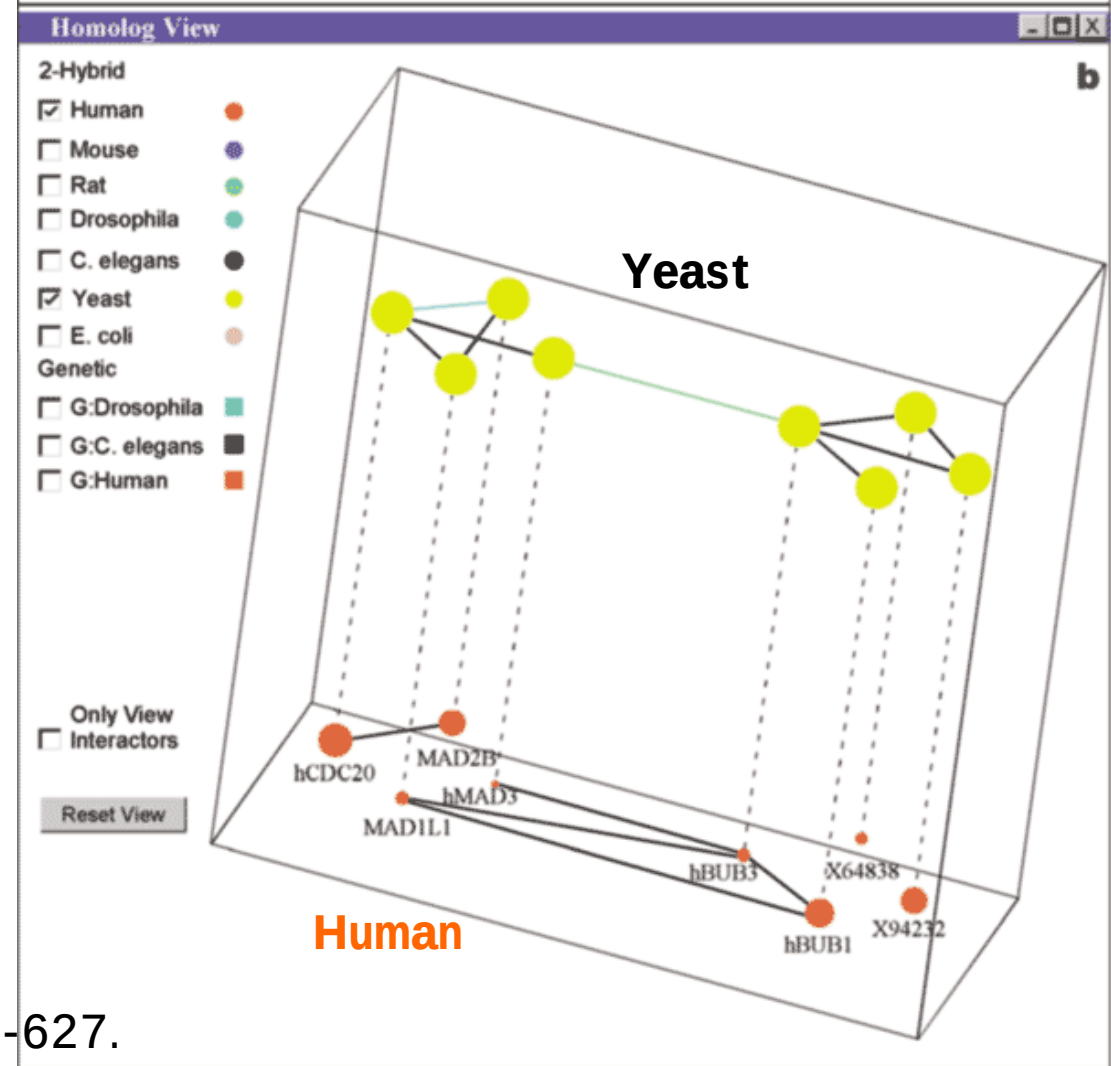
Description	Total
Yeast ORF PCR products	6,144
Yeast ORFs cloned*	5,345
ORFs pooled to generate the activation domain library	5,341†
ORF activating transcription on their own	680
Yeast ORFs identified to have interactions‡	817
Total discrete interacting protein pairs	692
Interactions identified in independent experiments§	286
Interactions identified multiple times in a single experiment‖	186
Interactions identified only once	220

Select against poor growth and mating ability

Results

- Systematic Y2H
  - Place functionally unclassified proteins into a biological context
  - Offer insight into novel interactions between proteins involved in the same biological function
  - Identify novel interactions that connect biological functions into larger cellular processes

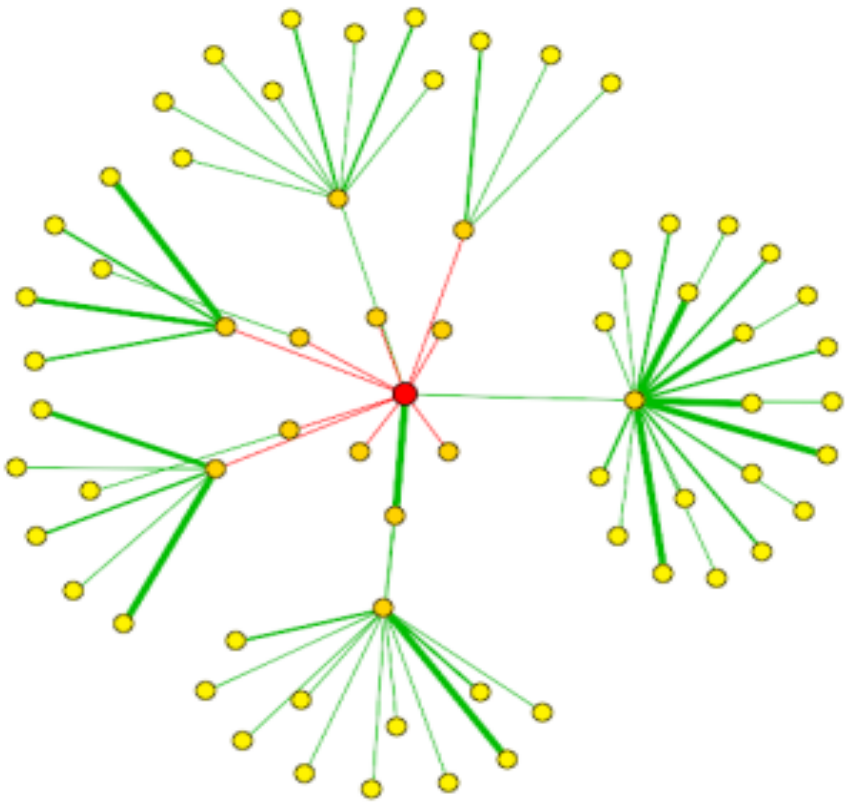


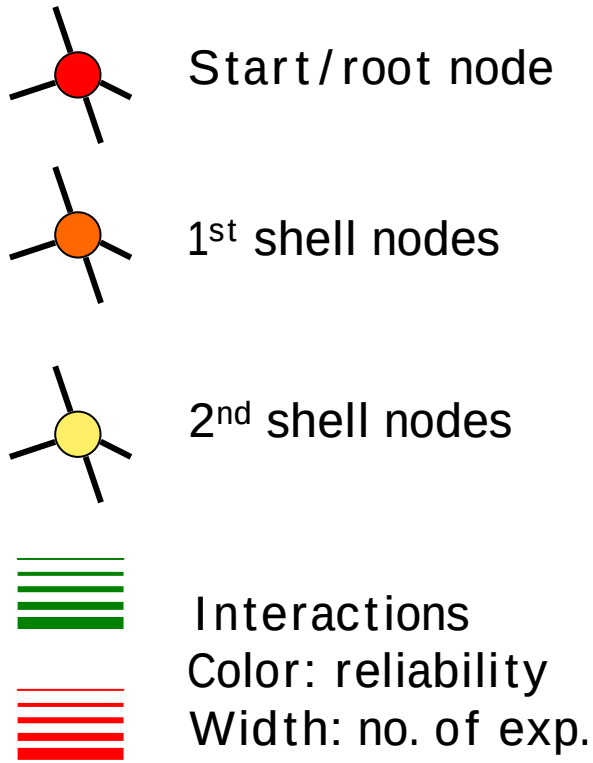


Uetz, et al., *Nature* 2000 403, 623-627.

Graphical representation

- Graphical representation of massive information
  - e.g. Database of Interacting Proteins (DIP)

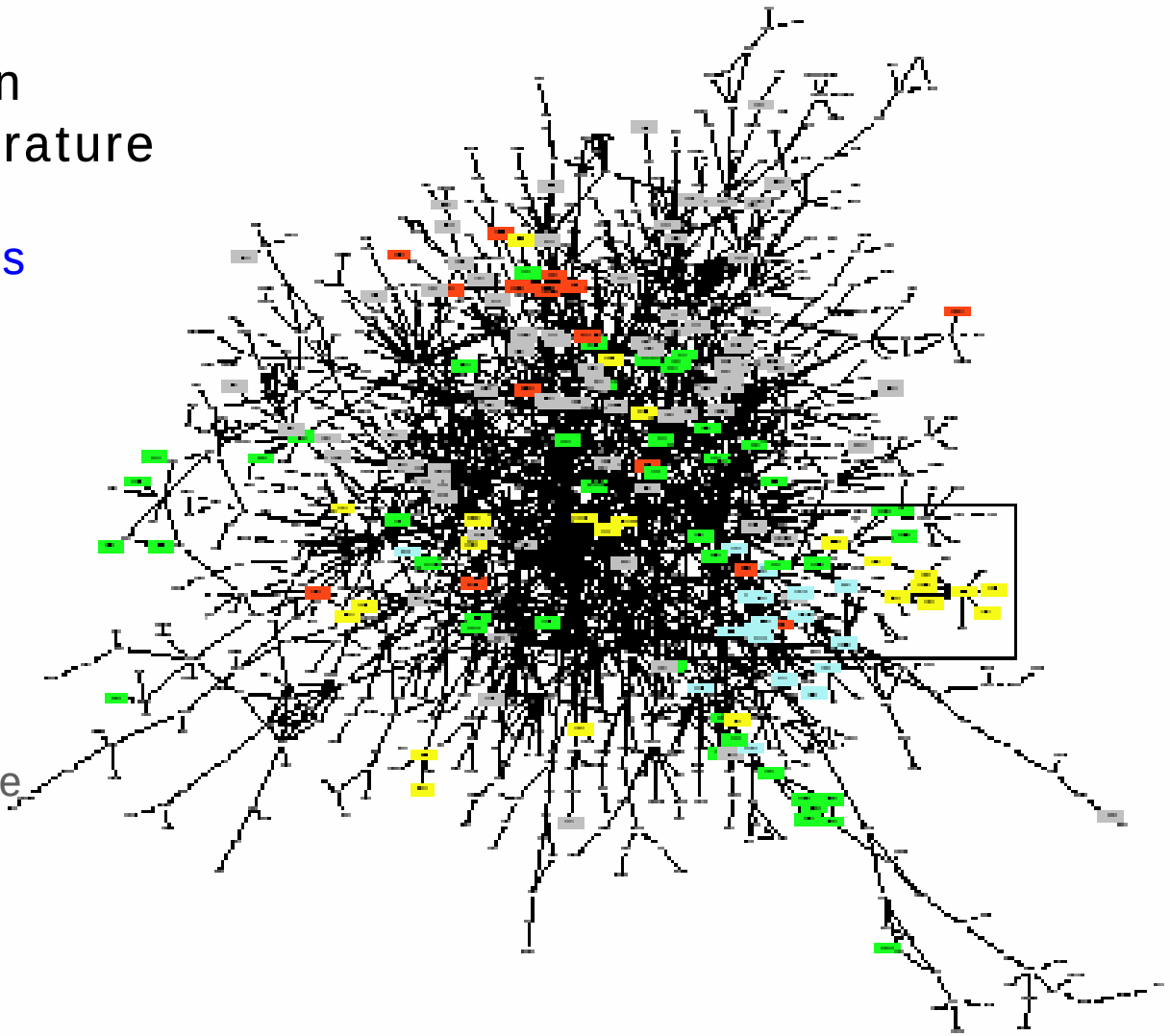




http://dip.doe-mbi.ucla.edu/

Interaction map of yeast proteome

- Built by association
- Compiled from literature
  - 1548 proteins
  - 2358 interactions

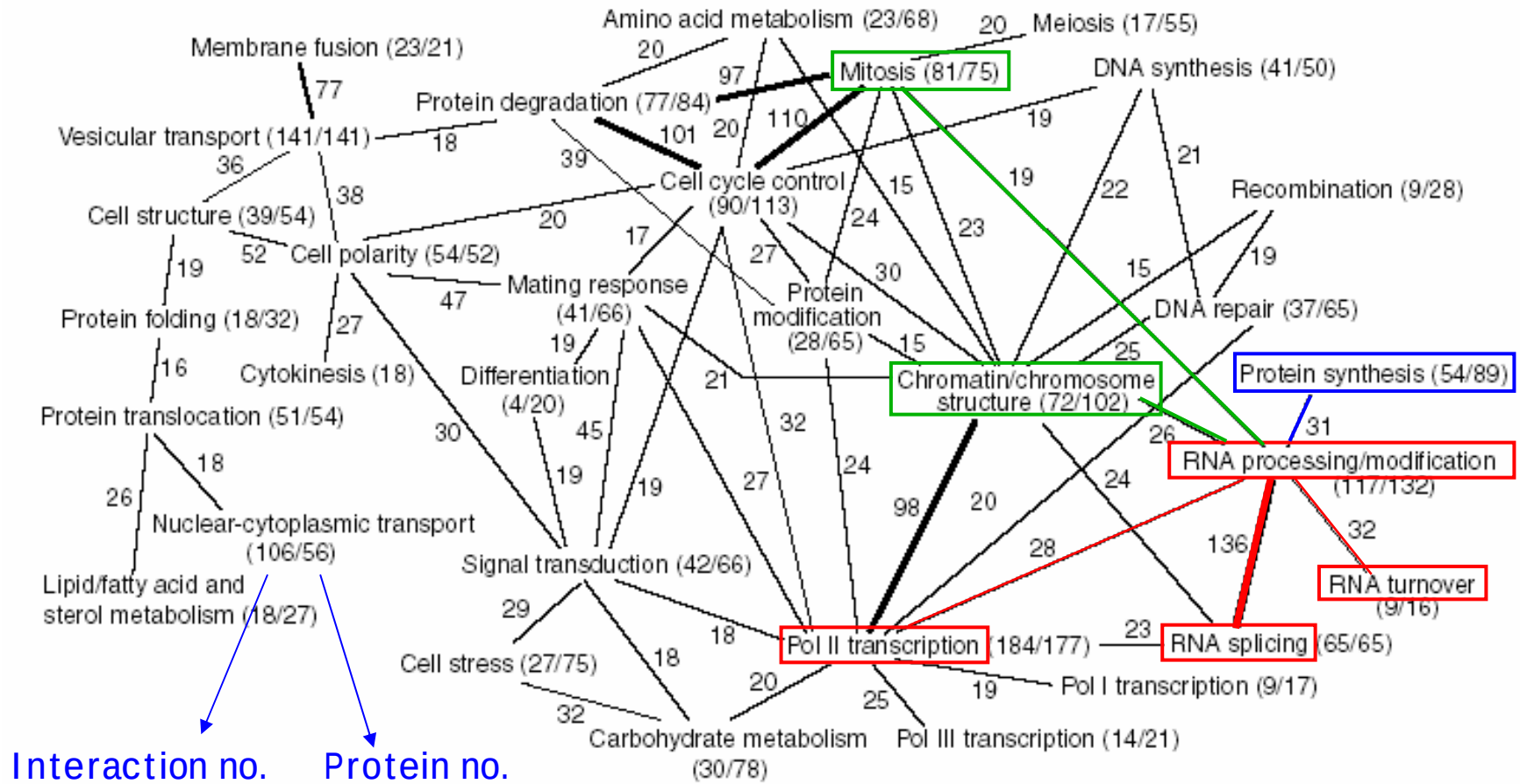


Blue: membrane fusion
Gray: chromatin structure
Green: cell structure
Yellow: lipid metabolism
Red: cytokinesis

Schwikowski et al., 2000 *Nature Biotech*. 18, 1257-1261.

Interaction between groups

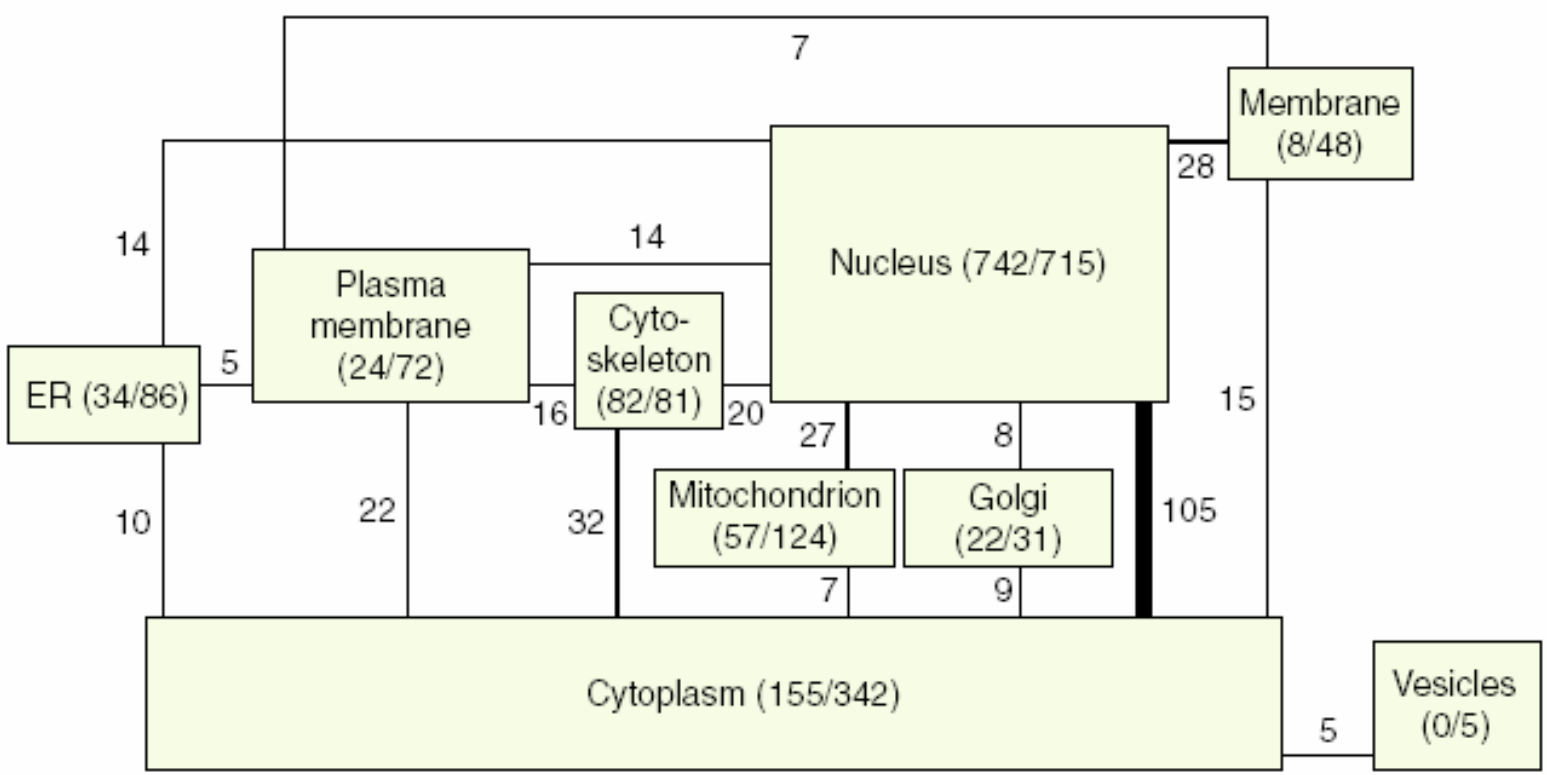
- Crosstalk between and within functional groups



Schwikowski et al., 2000 *Nature Biotech*. 18, 1257-1261.

By location

- Crosstalk between and within subcellular compartments



Schwikowski et al., 2000 *Nature Biotech*. 18, 1257-1261.

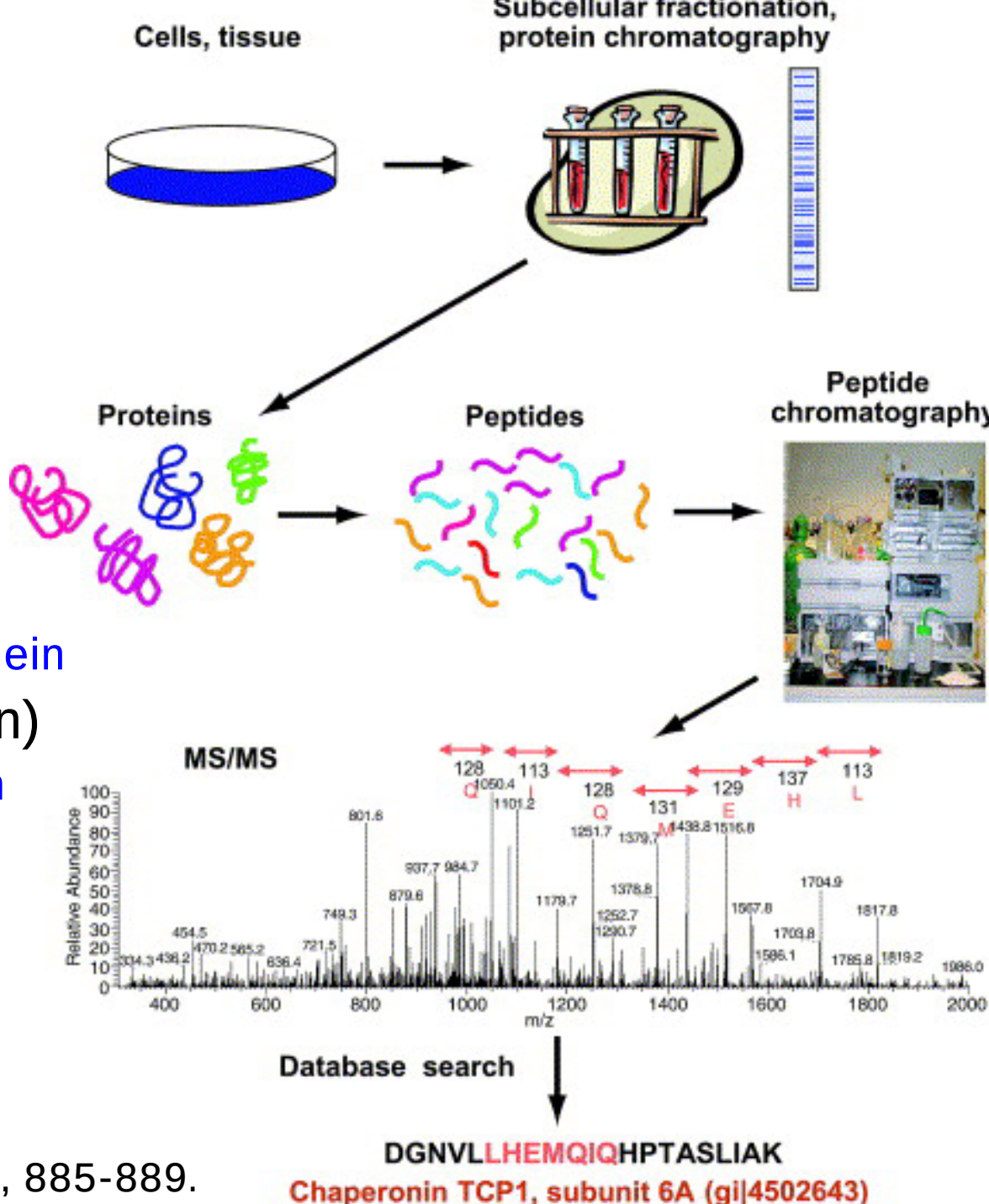
Two-hybrid System

- Pros
  - Function of the unknown protein can be inferred through interactions with proteins with known function.
  - Identify previous unrecognized interactions between proteins involved in the same biological process.
  - Provided clues for seeing how individual biological events are integrated into larger cellular process
- Cons
  - Lacks context-dependency
    - Identify potential protein interactions but not in their biological contexts

Schwikowski et al., 2000 *Nature Biotech*. 18, 1257-1261.

MS-based Proteomics

- Top-down
  - Analyze intact protein
- Bottom-up (shutgun)
  - Analyze peptides in proteolytic digest

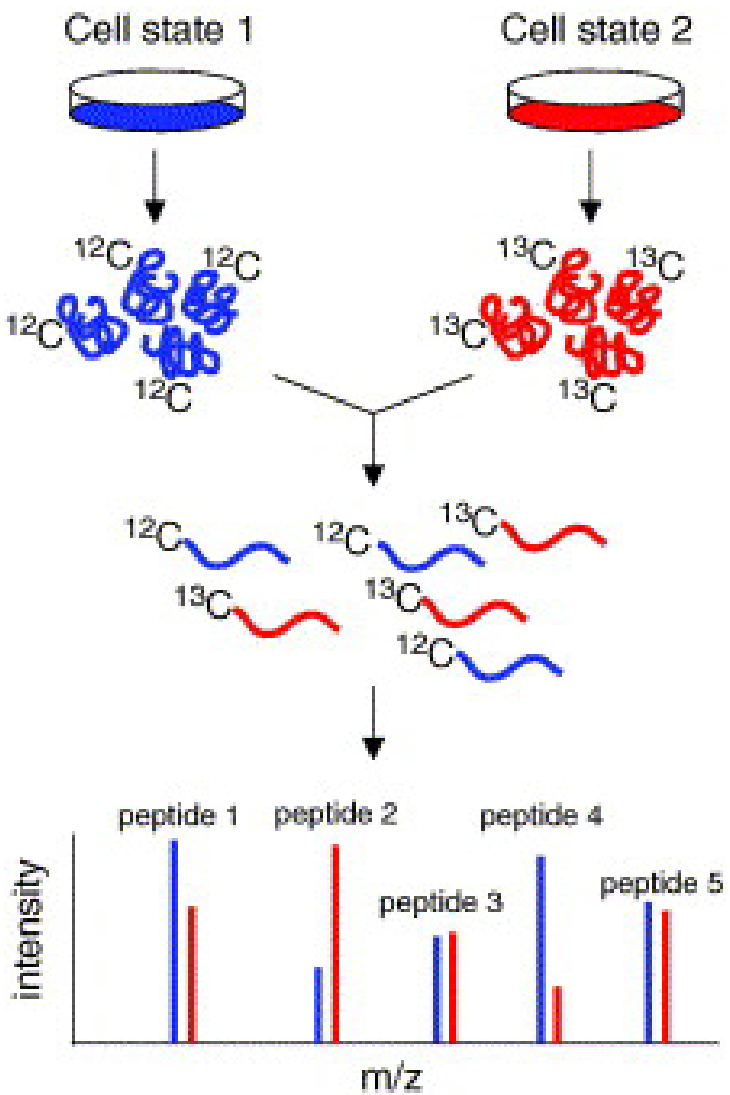


Resing and Ahn, 2005 FEBS 579, 885-889.

Resing and Ahn, 2005 FEBS 579, 885-889.

Quantification

- Relative abundance between samples



Stable isotope labeling (Chemically or metabolically

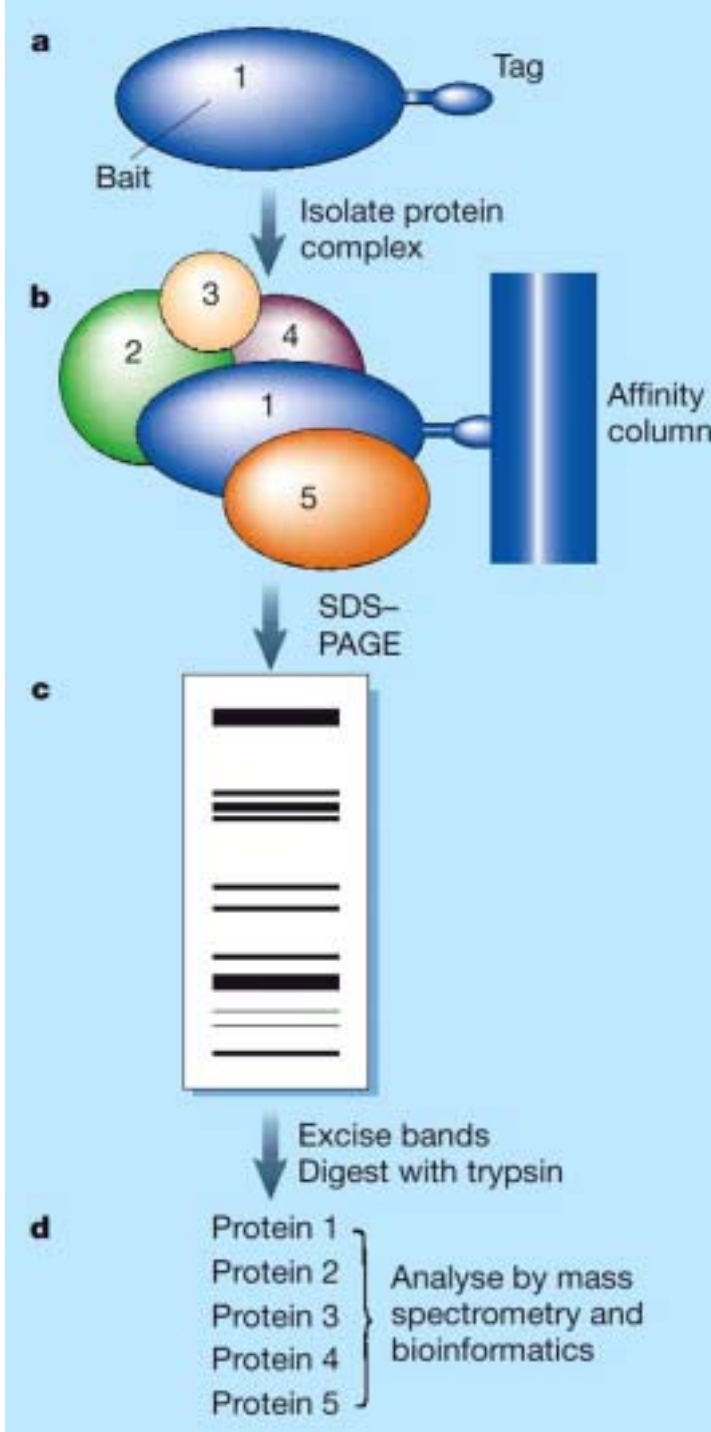
Labeled proteins Combine, fractionate proteolyze

MS analysis of isotopically labeled peptides

Quantify relative changes in protein abundance by measuring MS peak ratios

Tag-protein + MS

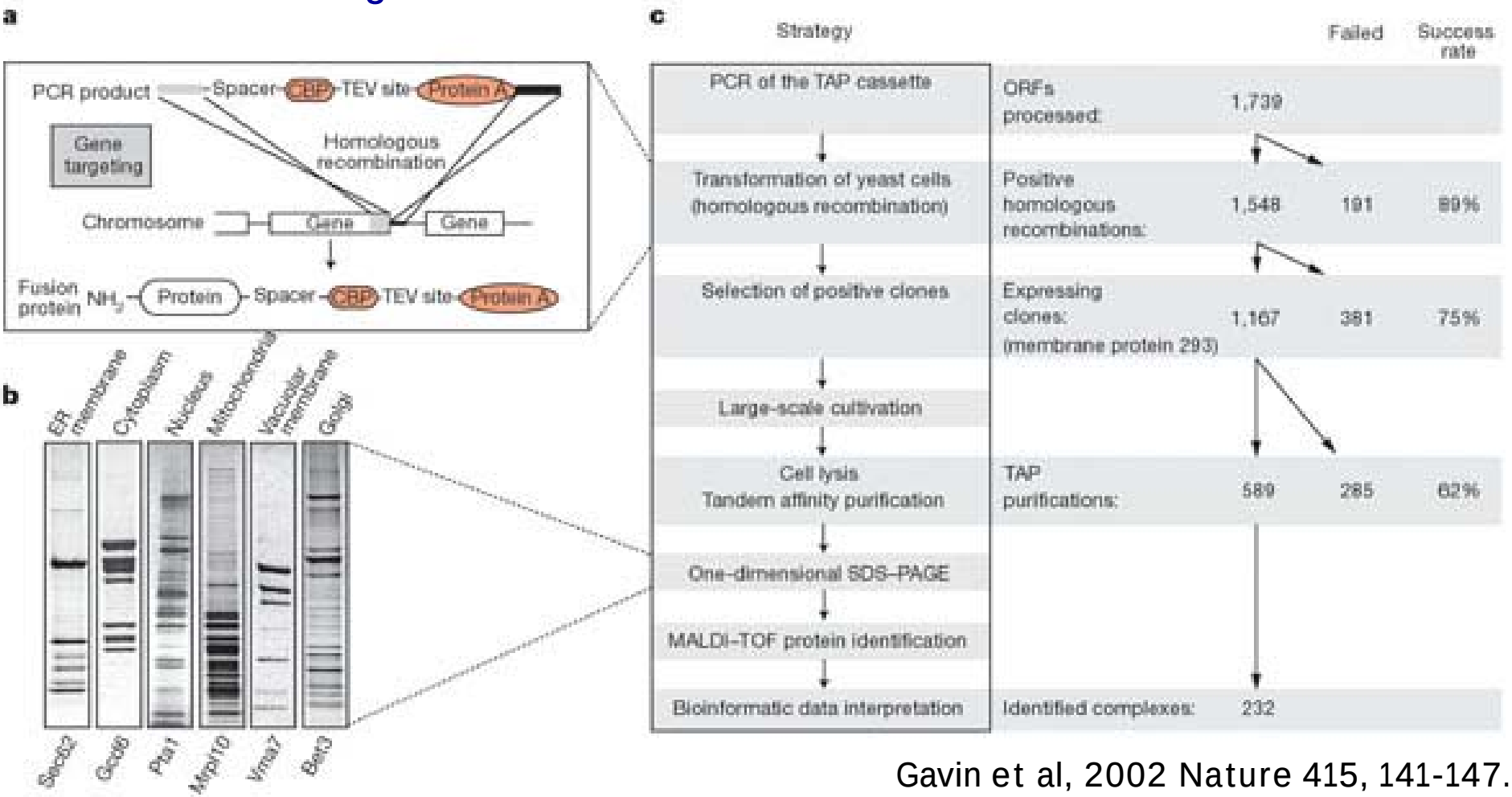
- Construct tagged “bait” protein
- Isolate protein complex
  - Tandem affinity purification (TAP)
    - 1st high-affinity purification
      - Elute with a site-specific protease
    - 2nd affinity purification
      - Protein complex
        - » Membrane associated
        - » Non-membrane associated
  - Immunoprecipitation
- Separation of protein complex
  - SDS-PAGE
  - Trypsin-digestion
  - Extraction from gel
- Identification
  - MALDI-TOF Mass Spectrometry
  - Database search algorithm (bioinformatics)



Kumar and Snyder, 2002 Nature 415, 123-124

TAP strategy

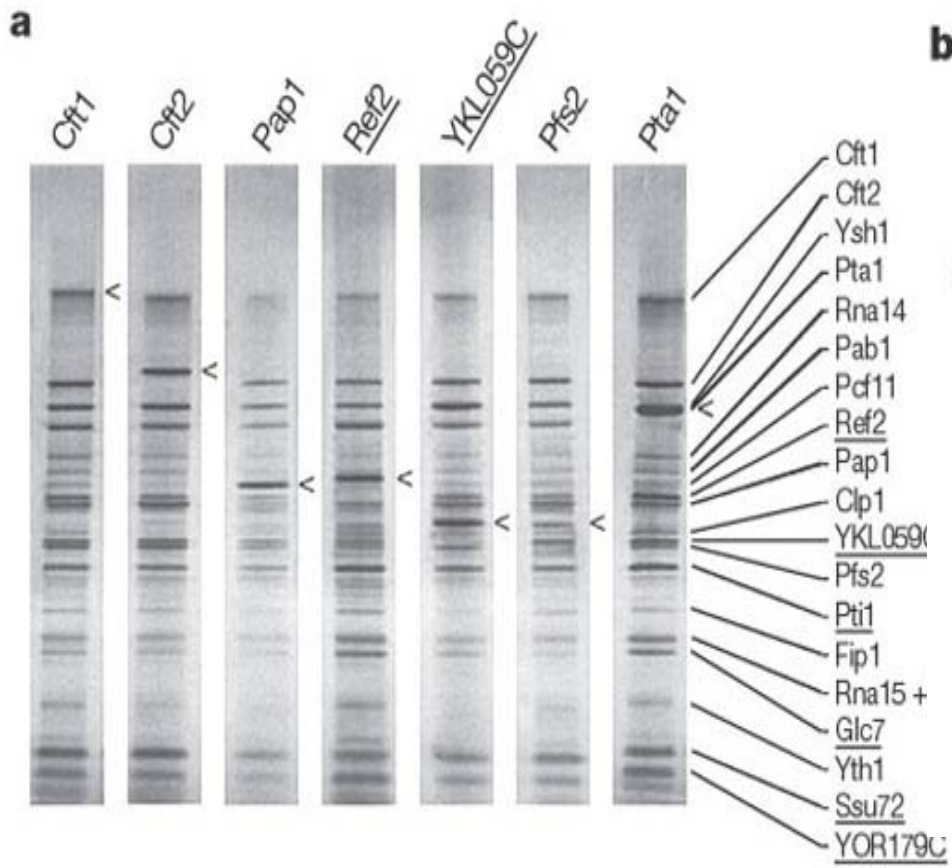
- Tandem affinity purification
  - Physiological expression level
  - In biological context

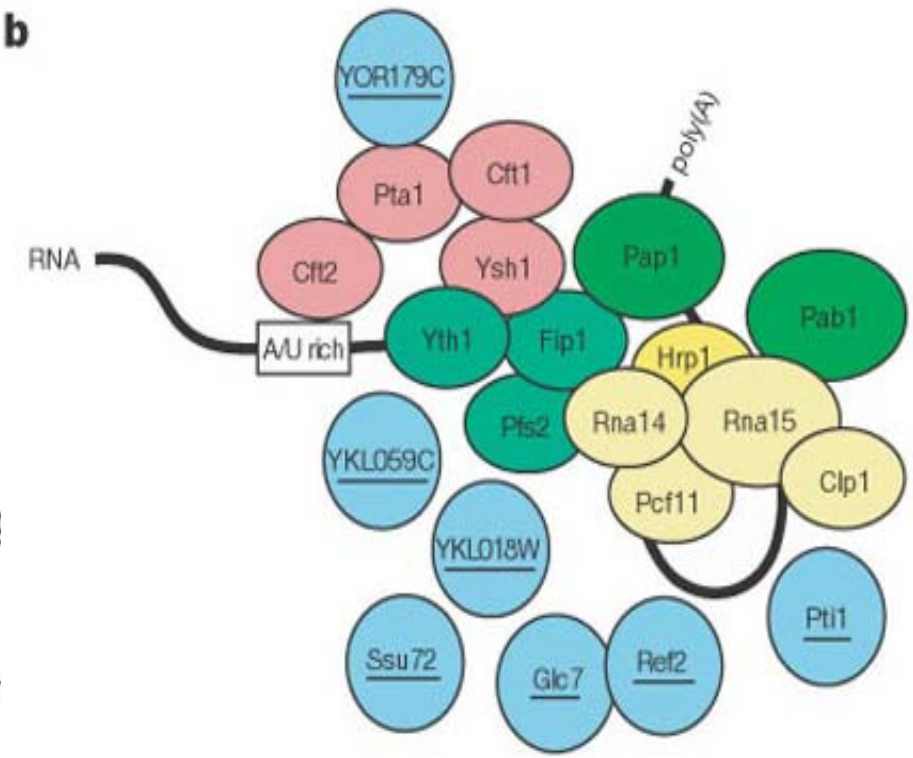


Gavin et al, 2002 Nature 415, 141-147.

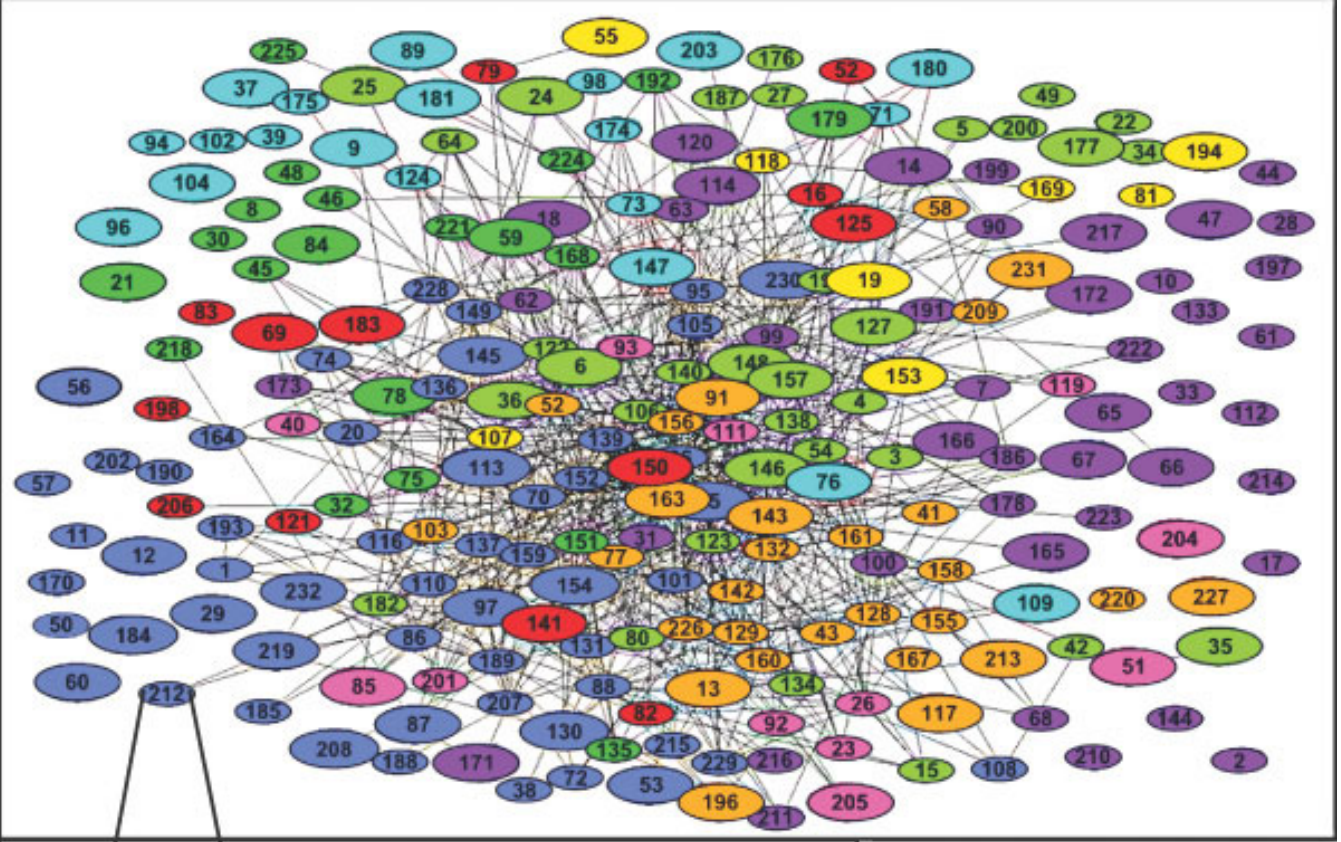
Complex composition

- Primary validation by 'reverse' purification
  - Use different "bait" protein as entry points
  - Identify indirect interacting components



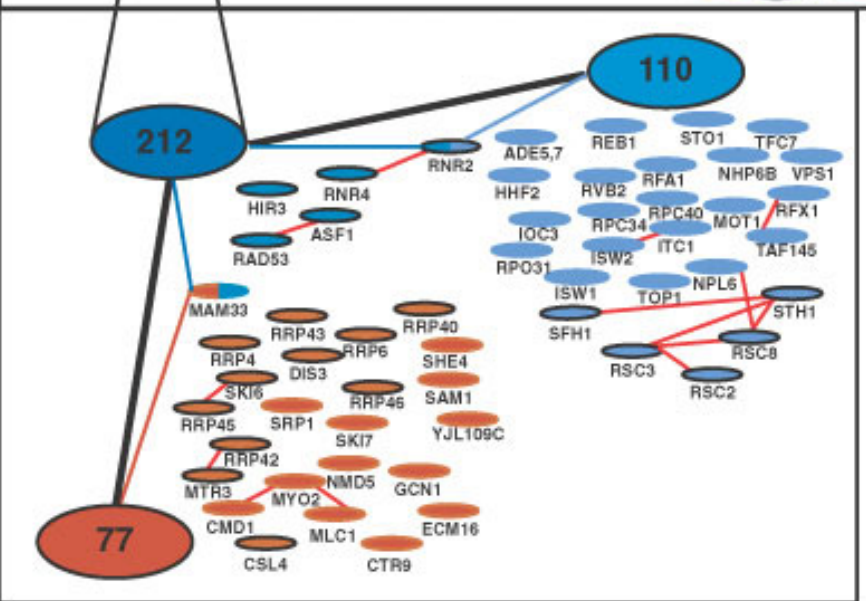


Gavin et al, 2002 Nature 415, 141-147.



Grouping

- Position (center)
- Size (large)
- Color (role)

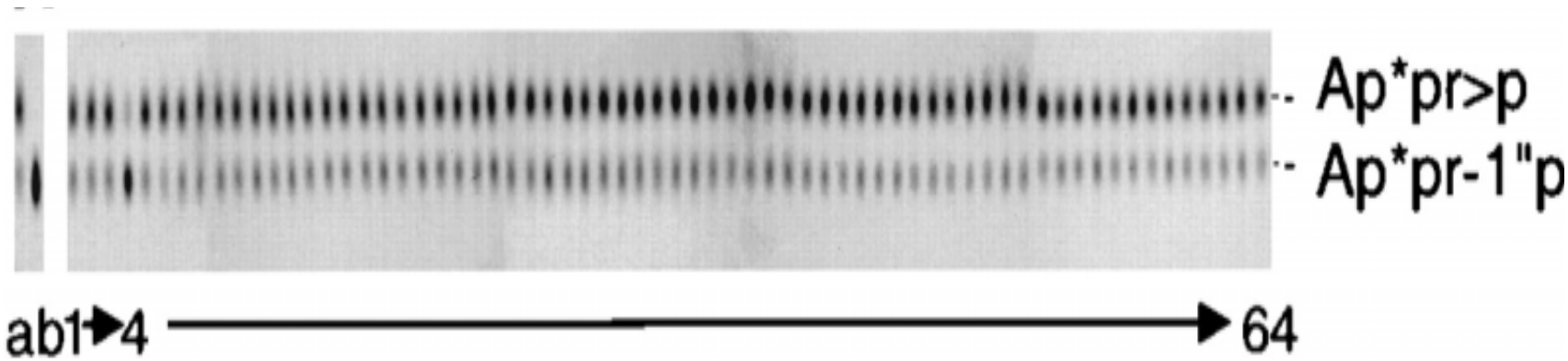


Protein complex network

Gavin et al, 2002 Nature 415, 141-147.

Pooling functional assay

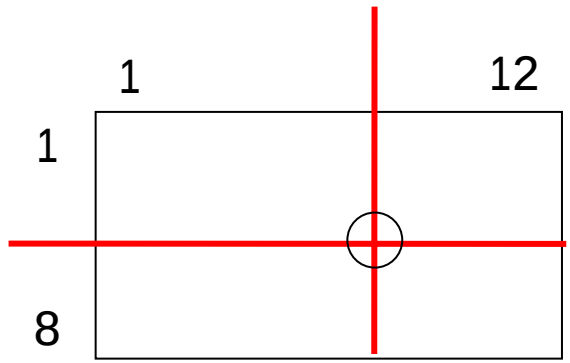
- Biochemical assay for activity
  - 6144 GST-ORF strains
    - 1st round assay: 64 pools of 96 fusions/plate each
  - Deconvolution of the positive pool
    - 2nd round assay: pools of 12 columns and 8 columns
  - Example: cyclic phosphodiesterase (Ap*pr>p → Ap*pr-1”p)
    - a: substrate only (Ap*pr>p)
    - b: substrate + enzyme → product (Ap*pr-1”p)
    - Separated on thin-layer plates

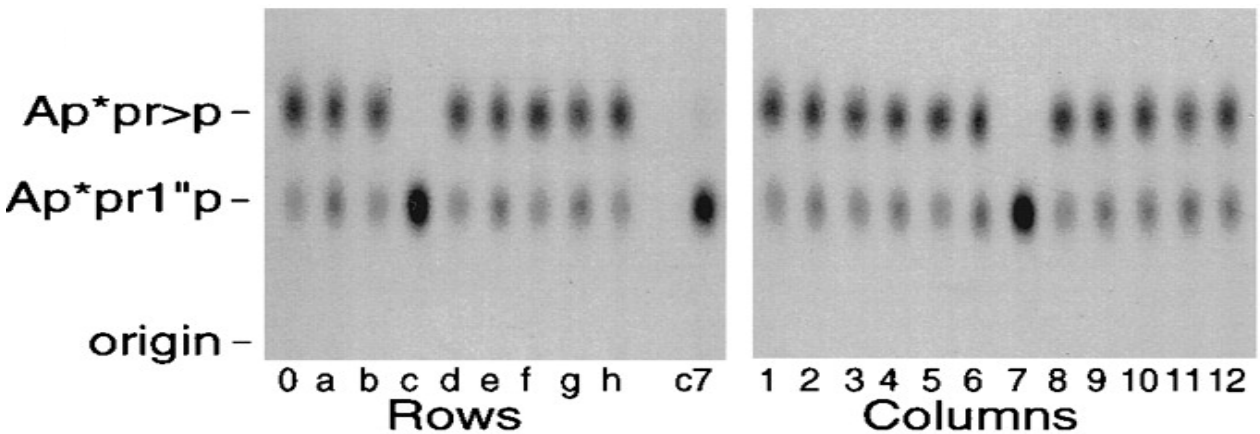


Martzen et al, 1999 Science 286, 1153-1155.

Deconvolution

- Example
  - Cyclic phosphodiesterase (CPDase)
  - Assay 64 pools to identify positive pool(s) → pool 4
  - Deconvolute the positive pool(s) in rows and columns → C7 (YGR247w)
- Priori
  - The GST-ORF is functional
  - Soluble after extraction
  - Remain functional
    - Retains other required components when purified
- Potentials
  - Fast and sensitive
  - Determine the range of the substrate proteins
  - Identifying gene leads to the binding of particular molecule, ligand, or drug.





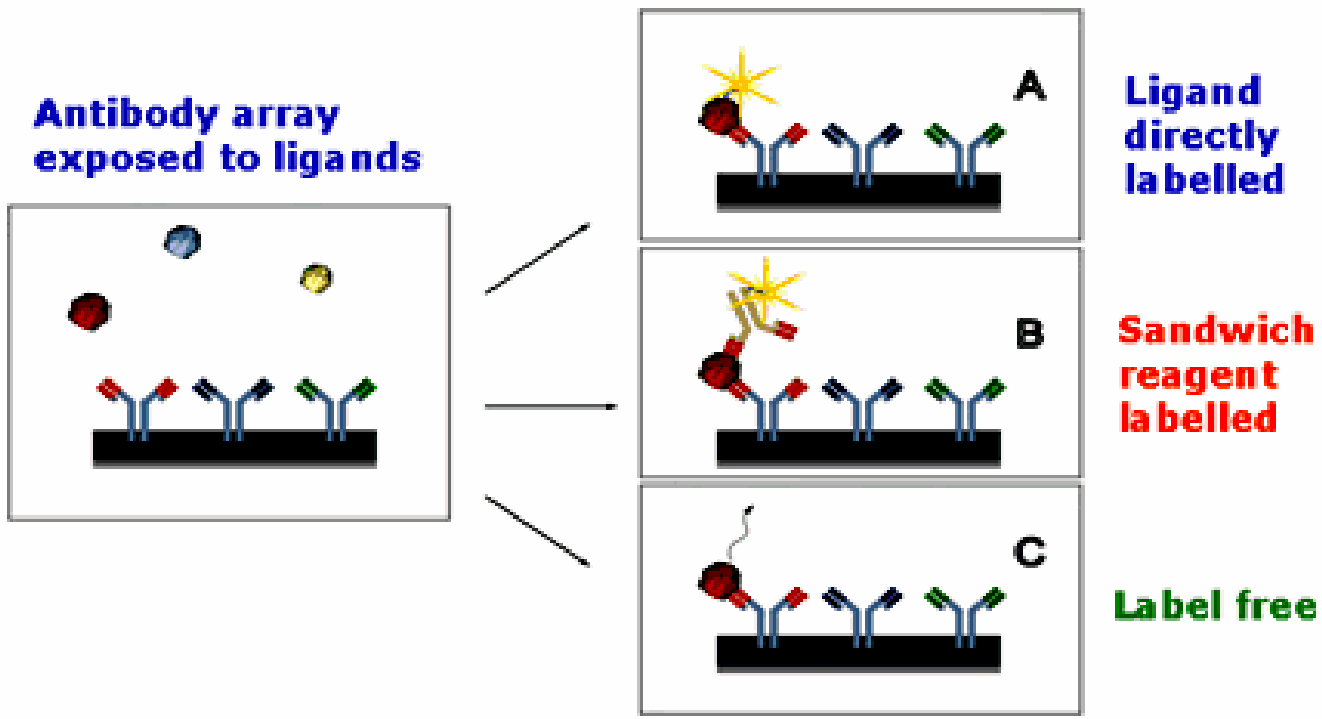
Martzen et al, 1999 Science 286, 1153-1155.

Protein Chip

- Miniaturized bait-and-capture assay
  - Protein-protein interaction
  - Protein-DNA interaction
  - Protein-RNA interaction
  - Protein-Ligand interaction
- Common elements
  - Immobilized of bait molecules on a substratum
  - Complex capture molecules
- Detection
  - Direct detection
  - Reverse-phase microarray

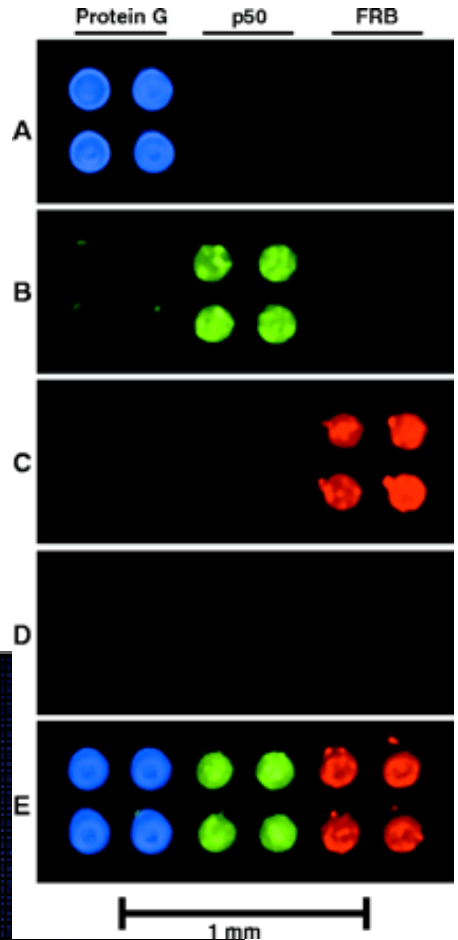
Detection strategy

- Label-free methods
  - Mass spectrometry
  - Surface plasmon resonance
  - Atomic force microscope
- Labeled probe methods
  - Direct
  - Indirect
  - Sandwich

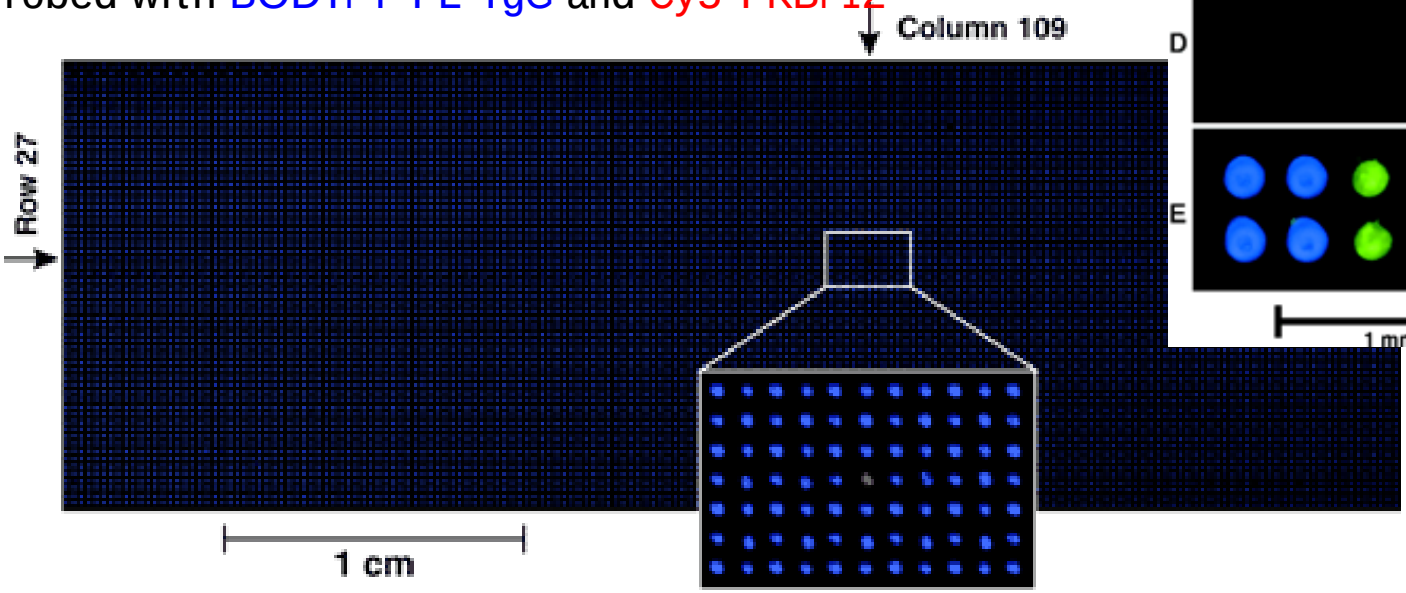


Protein Microarray

- Probe
  - BODIPY-FL-IgG
  - Cys3
  - Cy5



10,800 spots = 10,799 protein G + 1 FRB
Probed with BODIPY-FL-IgG and Cy5-FKBP12



MacBeath and Schreiber, 2000 Science 289, 1760-1763.

Protein Chip Challenges

- Protein vs. DNA
- Challenges
  - Low abundance protein in a mixture
  - Specificity and high-affinity antibody
  - Denatured vs. native state protein