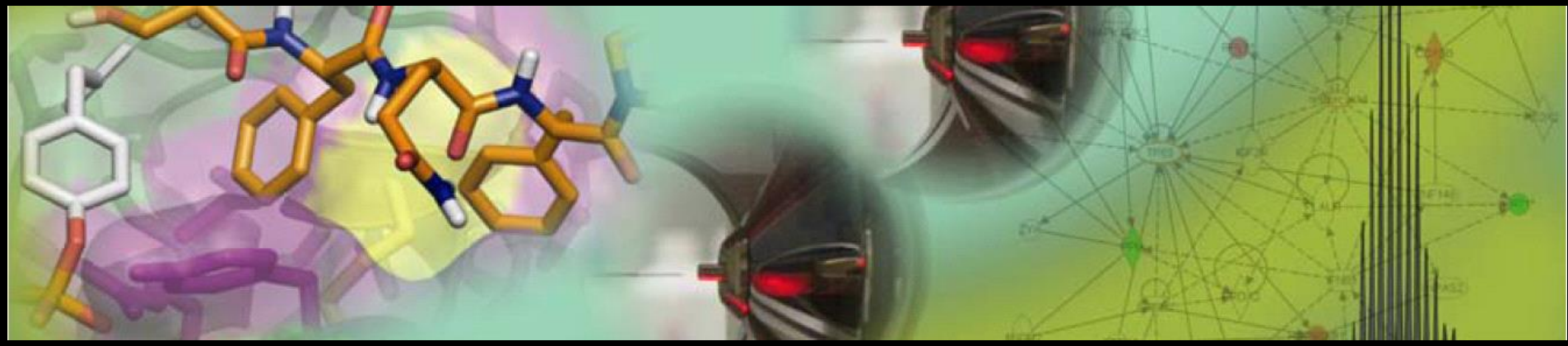


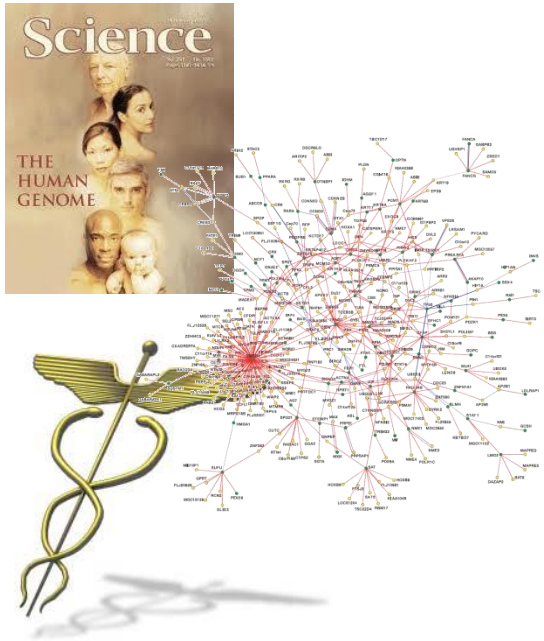
Proteins: Proteomics & Protein-Protein Interactions Part I

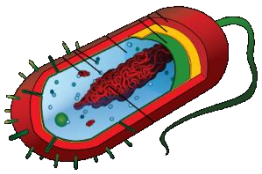
Jesse Rinehart, PhD

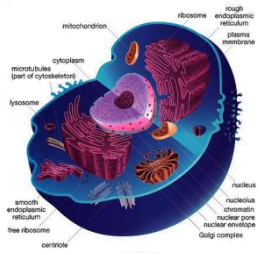
Yale Systems Biology Institute

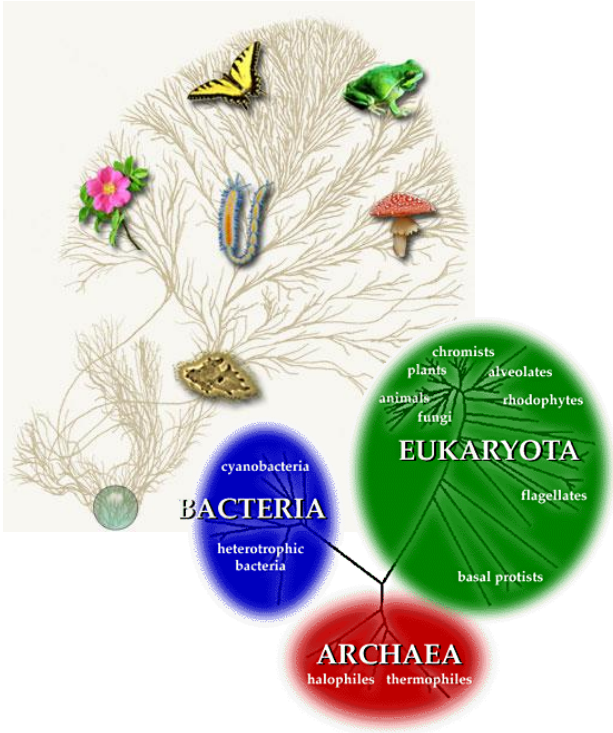
Cellular & Molecular Physiology
Yale University School of Medicine

DNA → RNA → PROTEIN

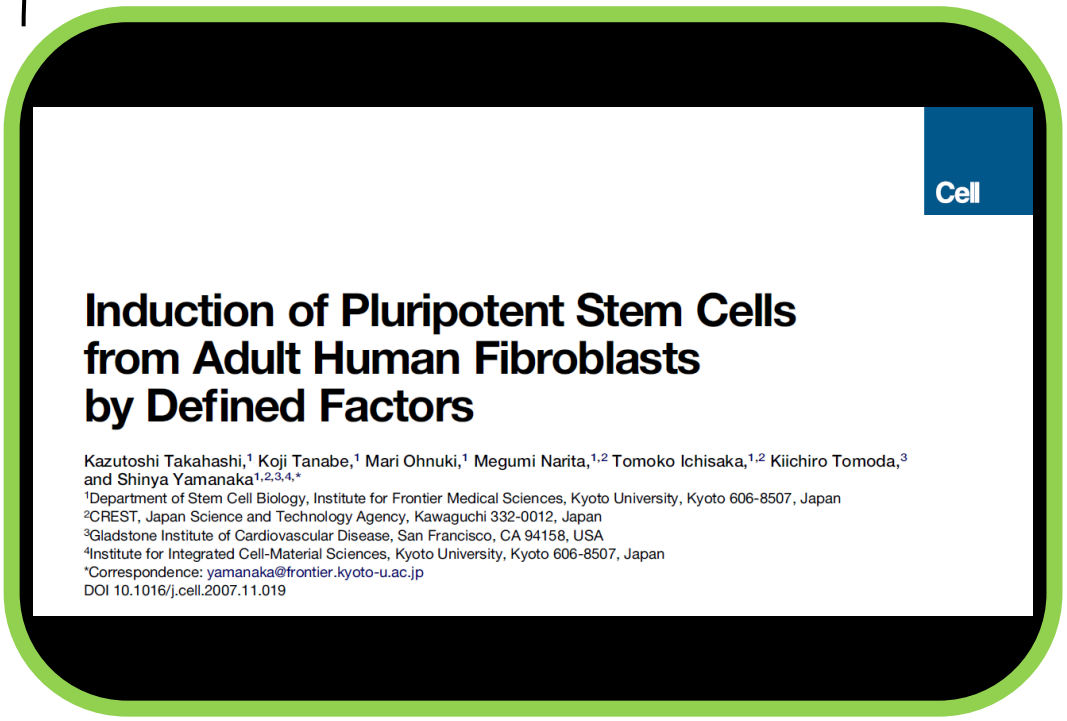








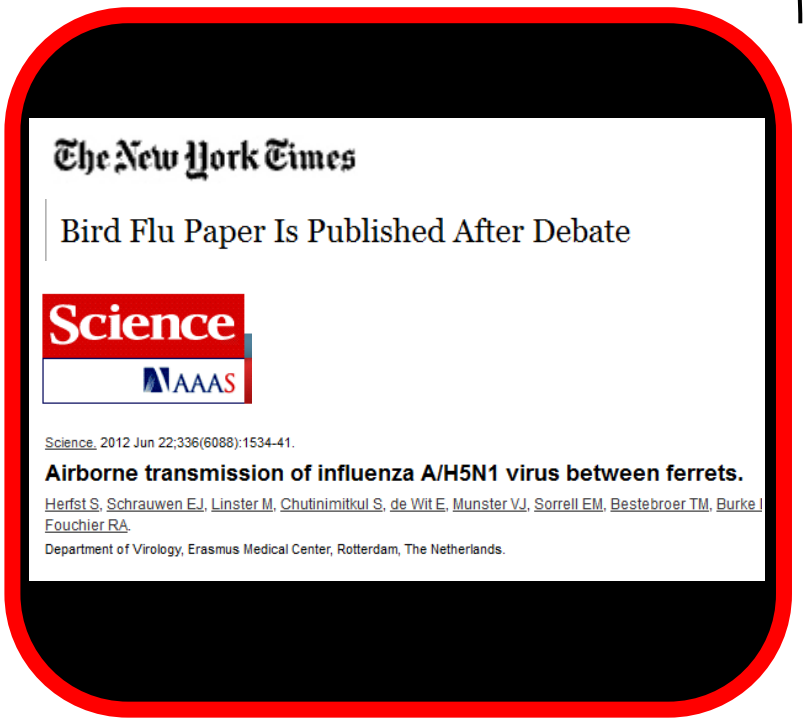
DNA → RNA → PROTEIN



Cell

Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors

Kazutoshi Takahashi,[1] Koji Tanabe,[1] Mari Ohnuki,[1] Megumi Narita,[1,2] Tomoko Ichisaka,[1,2] Kiichiro Tomoda,[3] and Shinya Yamanaka[1,2,3,4,*]
[1]Department of Stem Cell Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan
[2]CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan
[3]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA
[4]Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto 606-8507, Japan
*Correspondence: yamanaka@frontier.kyoto-u.ac.jp
DOI 10.1016/j.cell.2007.11.019



The New York Times

Bird Flu Paper Is Published After Debate

Science
AAAS

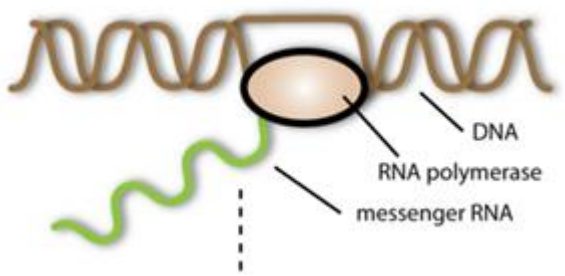
Science. 2012 Jun 22;336(6088):1534-41.

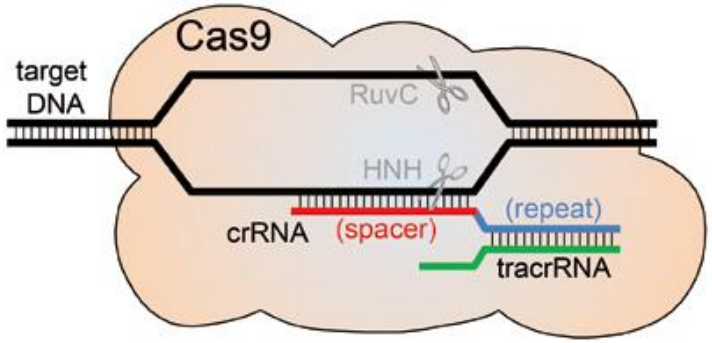
Airborne transmission of influenza A/H5N1 virus between ferrets.

Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke I
Fouchier RA.
Department of Virology, Erasmus Medical Center, Rotterdam, The Netherlands.

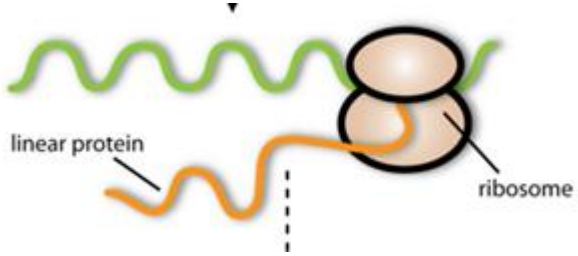
DNA → RNA → PROTEIN

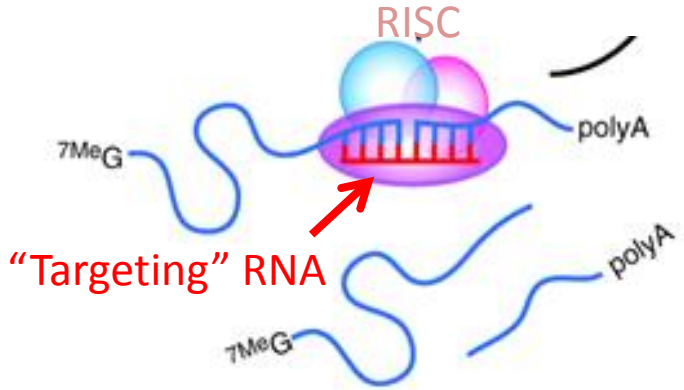
CRISPR





RNAi





Proteins: Proteomics & Protein-Protein Interactions

Overview

- **Techniques & Technologies**
  - Mass Spectrometry
  - Protein-Protein Interactions
    - Genetic & Biochemical Strategies
    - Protein Purification
  - Quantitative Proteomics
- **Applications**
  - Representative Studies
- **Putting it all together….**
  - Databases & Pathways

Proteins: Proteomics & Protein-Protein Interactions

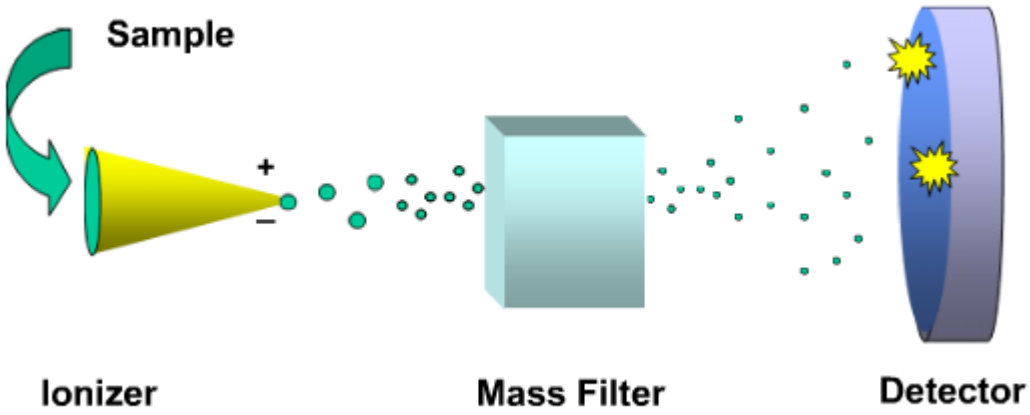
Overview

- **Techniques & Technologies**
  - **Mass Spectrometry**
  - Protein-Protein Interactions
    - Genetic & Biochemical Strategies
    - Protein Purification
  - Quantitative Proteomics
- **Applications**
  - Representative Studies
- **Putting it all together….**
  - Databases & Pathways

Principles of Mass Spectrometry (MS)

- In a mass spectrum we measure m/z (mass-to-charge)
- For proteins we measure peptide m/z
- A sample has to be ionizable in order to be analyzed

Basic Components of a Mass Spectrometer



Two major ionization techniques enabled the success of mass spectrometry in the life sciences.

- Electrospray Ionization (ESI)

  Fenn, J.B. et al, Science, 1989, 246, 64.

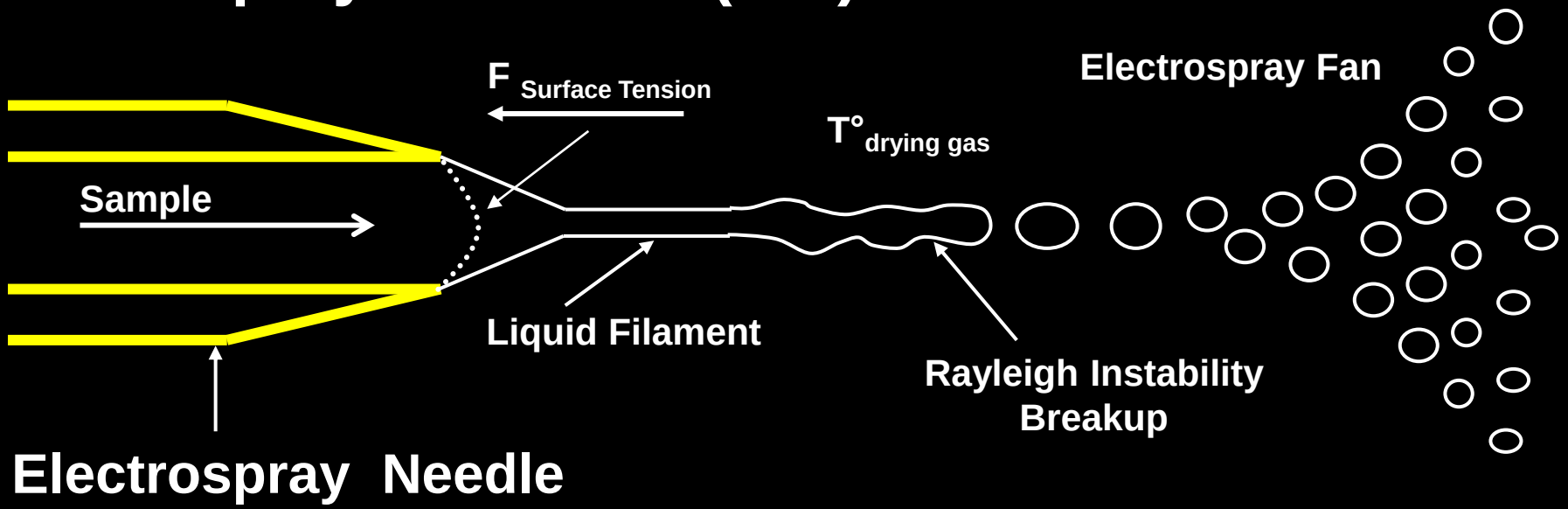
- Matrix Assisted Laser Desorption Ionization (MALDI)

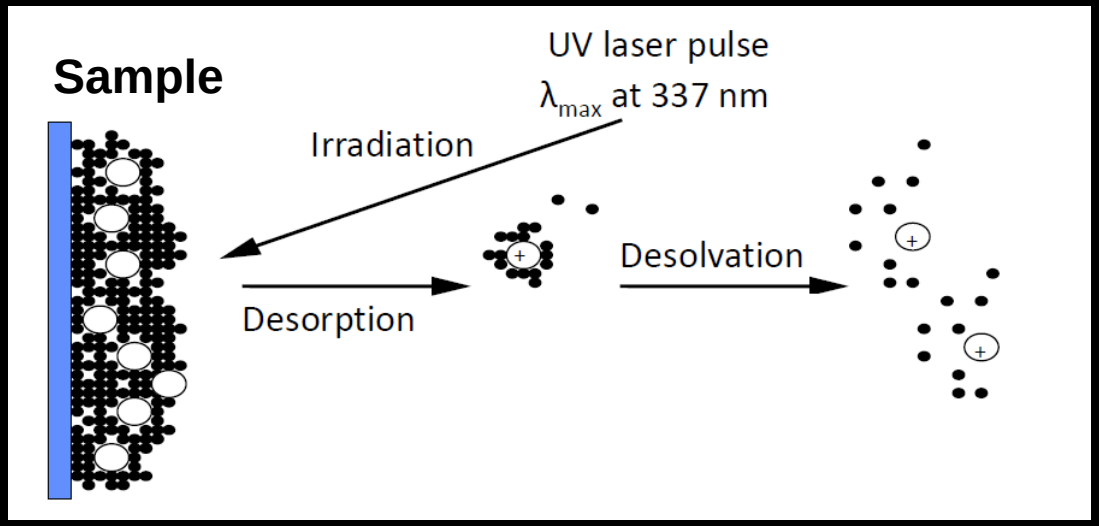
  Karas, M.; Hillenkamp, F., Anal. Chem., 1995, 60, 2299

MS based Proteomics is born:

- MS to measure weight of large intact proteins
- Non-covalently bonded protein complexes can also be measured (ESI only)
- **Intact peptides measured and "sequenced"**

Electrospray Ionization (ESI)





Matrix Assisted Laser Desorption Ionization (MALDI)

Mass Spectrometry takes the 2002 Nobel Prize in Chemistry

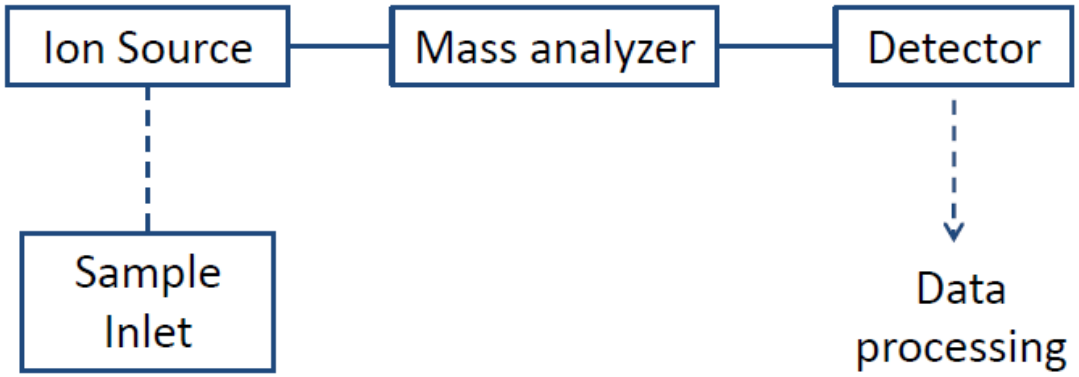
Awarded to John B. Fenn & Koichi Tanaka

***Fenn -** Discovered Electrospray Ionization (ESI)

***Pioneering work at Yale University in the Department of Chemical Engineering**

Tanaka - Discovered Matrix Assisted Laser Desorption Ionization (MALDI)

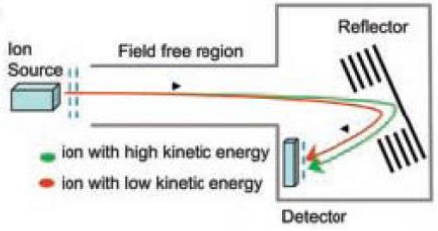
Basic Mass Spectrometer



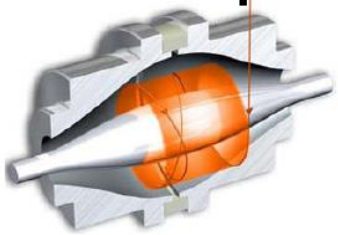
Typical LC-MS Setup

Mass Analyzers

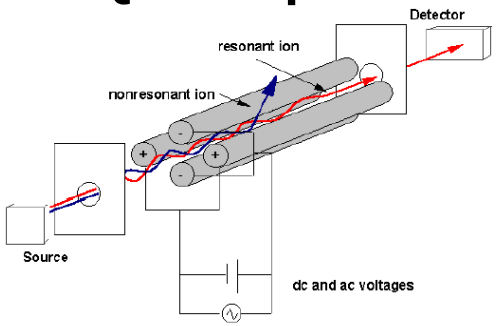
Time of Flight (TOF)



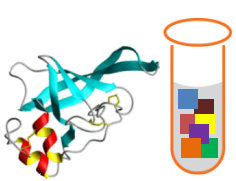
Orbitrap



Quadrupole



Typical work flow for LC-MS "shotgun proteomics"



Protein mixture

Enz. Digest

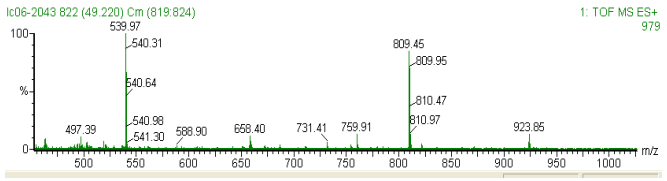
Peptide Pool

LC

n-UPLC

LTQ-Orbitrap MS

MS



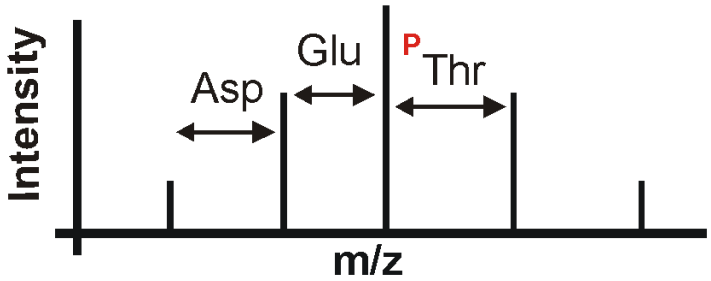
peptide

peptide

isolate & fragment

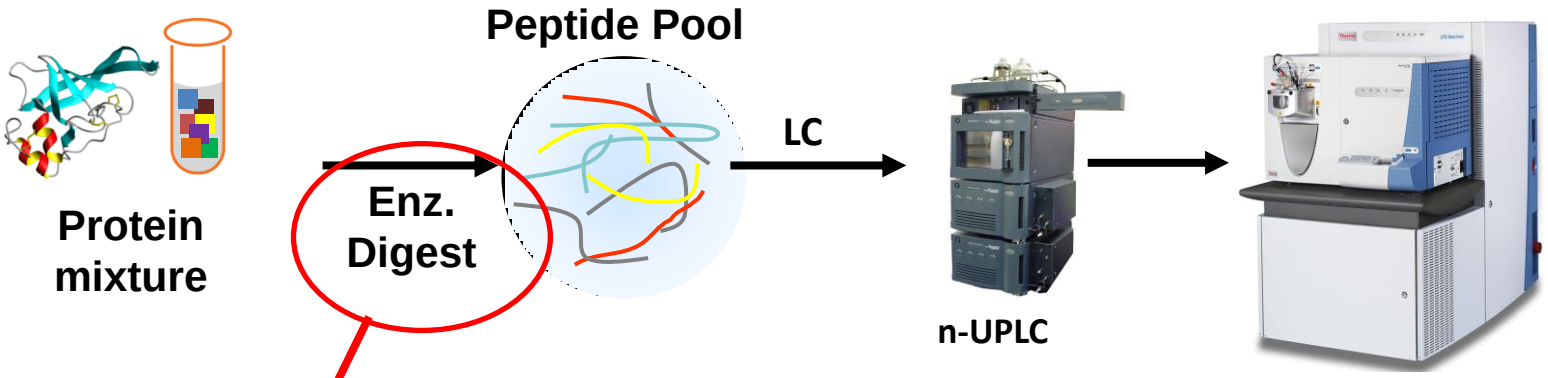
peptide

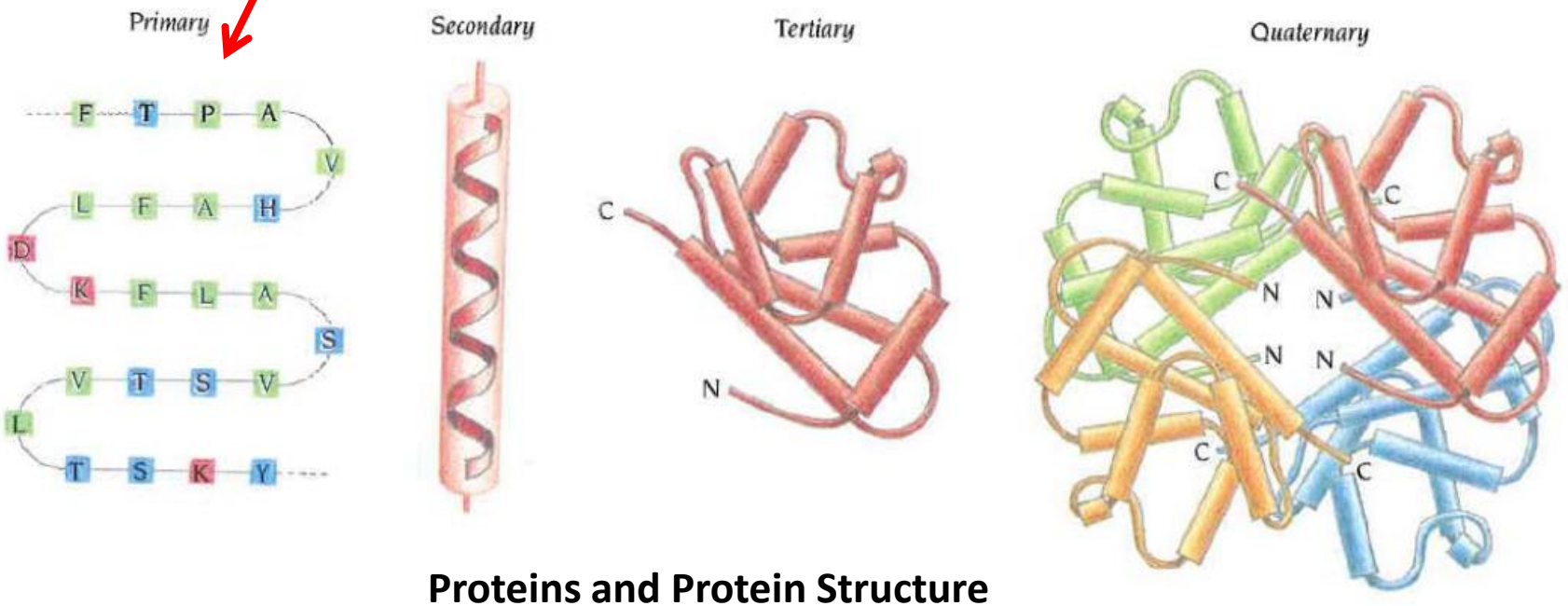
MS/MS



peptide fragments

Typical work flow for LC-MS "shotgun proteomics"





Proteins and Protein Structure

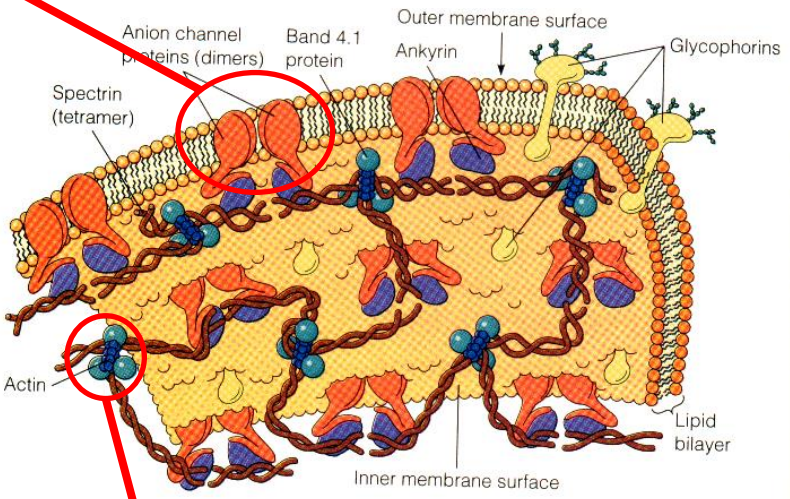
(Branden, C. and Tooze, J. ***Introduction to Protein Structure***)

Trypsin digest followed by LC-MS: Examples of "***Sequence Coverage***"

Band 3 Anion Transporter

Matched peptides shown in **Bold Red**

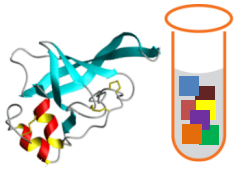
```
   1 MEELQDDYED MMEENLEQEE YEDPDIPESQ MEEPAAHDTE ATATDYHTTS
  51 HPGTHKVYVE LQELVMDEKN QELRWMEAAR WVQLEENLGE NGAWGRPHLS
 101 HLTFWSLLEL RRVFTKGTVL LDLQETSLAG VANQLLDRFI FEDQIRPQDR
 151 EELLRALLLK HSHAGELEAL GGVKPAVLTR SGDPSQPLLP QHSSLETQLF
 201 CEQGDGGTEG HSPSGILEKI PPDSEATLVL VGRADFLEQP VLGFVRLQEA
 251 AELEAVELPV PIRFLFVLLG PEAPHIDYTQ LGRAAATLMS ERVFRIDAYM
 301 AQSRGELLHS LEGFLDCSLV LPPTDAPSEQ ALLSLVPVQR ELLRRRYQSS
 351 PAKPDSSFYK GLDLNGGPDD PLQQTGQLFG GLVRDIRRRY PYYLSDITDA
 401 FSPQVLAAVI FIYFAALSPA ITFGGLLGEK TRNQMGVSEL LISTAVQGIL
 451 FALLGAQPLL VVGFSGPLLV FEEAFFSFCE TNGLEYIVGR VWIGFWLILL
 501 VVLVVAFEGS FLVRFISRYT QEIFSFLISL IFIYETFSKL IKIFQDHPLQ
 551 KTYNYNVLMV PKPQGPLPNT ALLSLVLMAG TFFFAMMLRK FKNSSYFPGK
 601 LRRVIGDFGV PISILIMVLV DFFIQDTYTQ KLSVPDGFKV SNSSARGWVI
 651 HPLGLRSEFP IWMMFASALP ALLVFILIFL ESQITTLIVS KPERKMVKGS
 701 GFHLDLLLVV GMGGVAALFG MPWLSATTVR SVTHANALTV MGKASTPGAA
 751 AQIQEVKEQR ISGLLVAVLV GLSILMEPIL SRIPLAVLFG IFLYMGVTSL
 801 SGIQLFDRIL LLFKPPKYHP DVPYVKRVKT WRMHLFTGIQ IICLAVLWVV
 851 KSTPASLALP FVLILTVPLR RVLLPLIFRN VELQCLDADD AKATFDEEEG
 901 RDEYDEVAMP V
```



β-actin

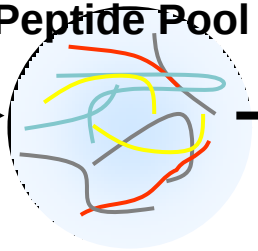
Matched peptides shown in **Bold Red**

```
   1 MDDDIAALVV DNGSGMCKAG FAGDDAPRAV FPSIVGRPRH QGVMVGMGQK
  51 DSYVGDEAQS KRGILTLKYP IEHGIVTNWD DMEKIWHHTF YNELRVAPEE
 101 HPVLLTEAPL NPKANREKMT QIMFETFNTP AMYVAIQAVL SLYASGRTTG
 151 IVMDSGDGVT HTVPIYEGYA LPHAILRLDL AGRDLTDYLM KILTERGYSF
 201 TTTAEREIVR DIKEKLCYVA LDFEQEMATA ASSSSLEKSY ELPDGQVITI
 251 GNERFRCPEA LFQPSFLGME SCGIHETTFN SIMKCDVDIR KDLYANTVLS
 301 GGTTMYPGIA DRMQKEITAL APSTMKIKII APPERKYSVW IGGSILASLS
 351 TFQQMWISKQ EYDESGPSIV HRKCF
```



Protein mixture

Enz. Digest

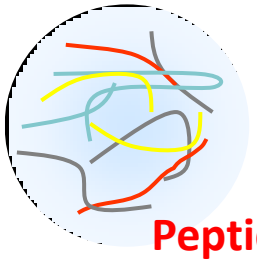


Peptide Pool

LC

n-UPLC

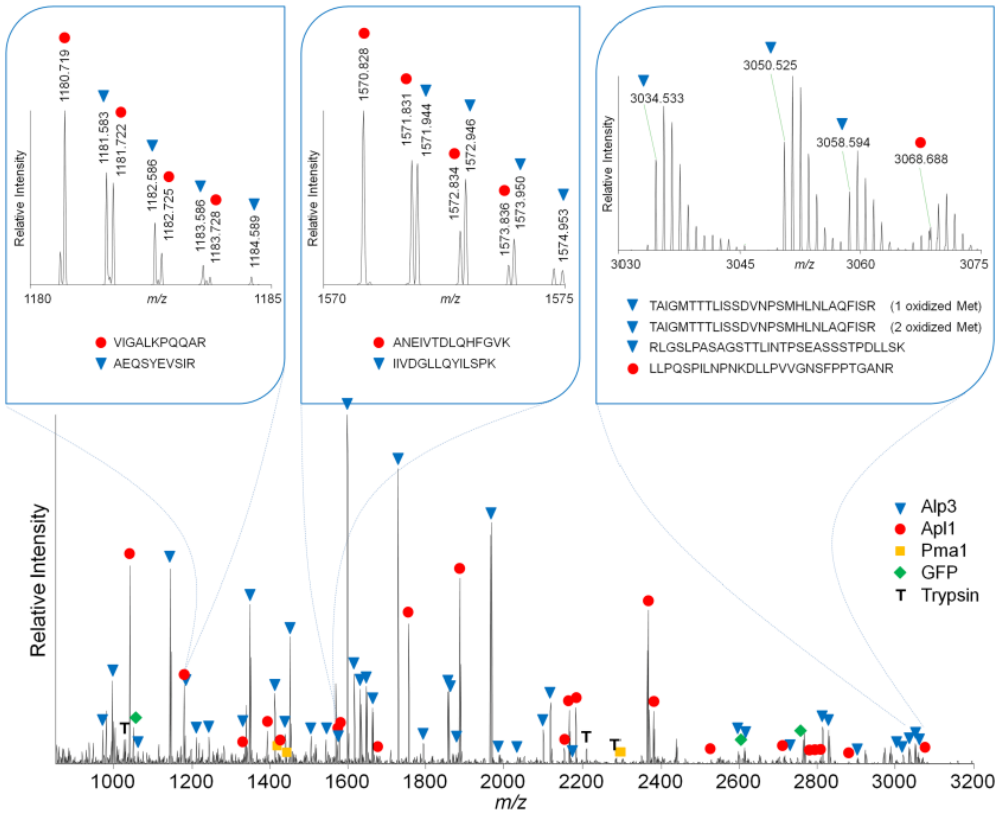
LTQ-Orbitrap MS

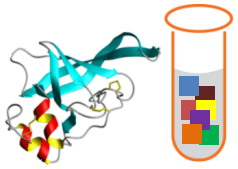


Peptide ions have a mass (m) and a charge (z).

100 Da peptide:
+1 = 100 m/z
+2 = 50 m/z
+3 = 33.3 m/z

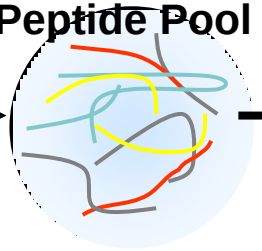
Mass Spectrum





Protein mixture

Enz. Digest

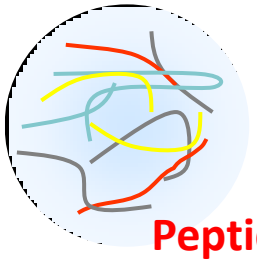


Peptide Pool

LC

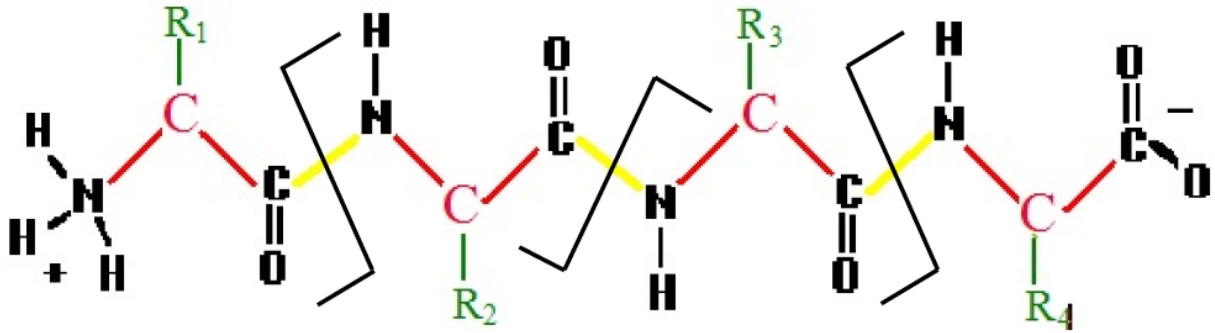
n-UPLC

LTQ-Orbitrap MS

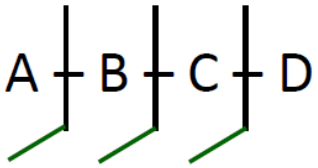


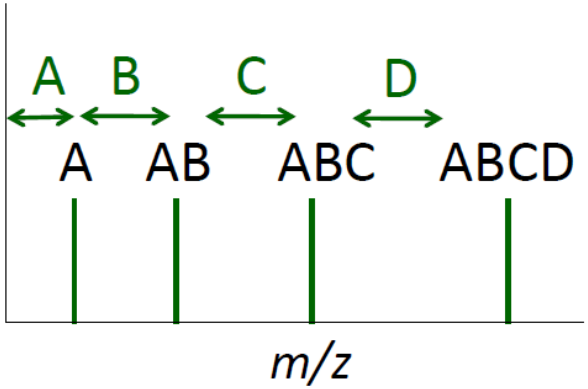
Peptide ions are isolated and "sequenced"

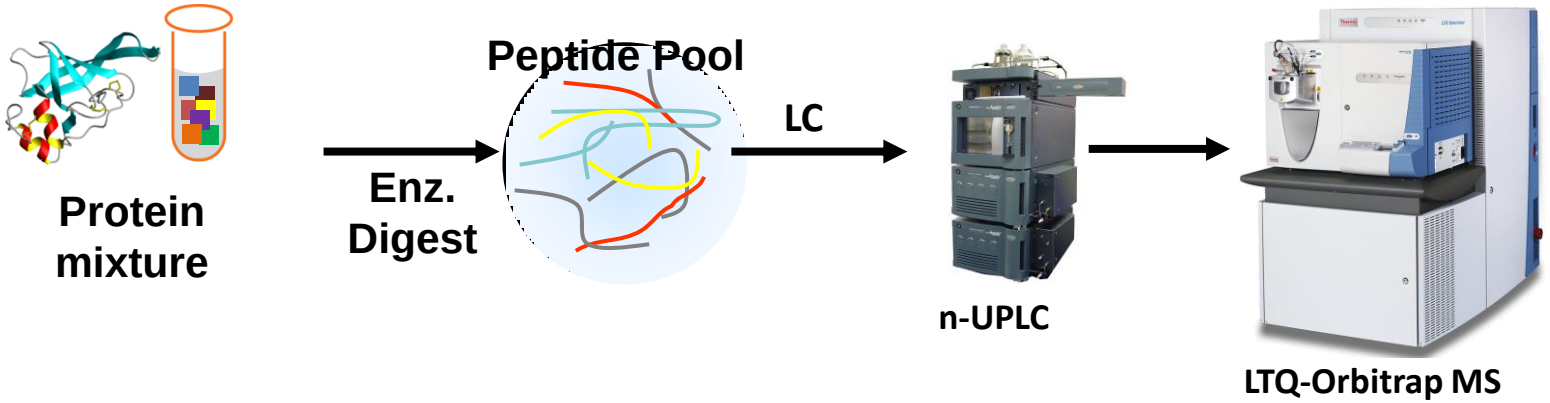
Peptide sequencing



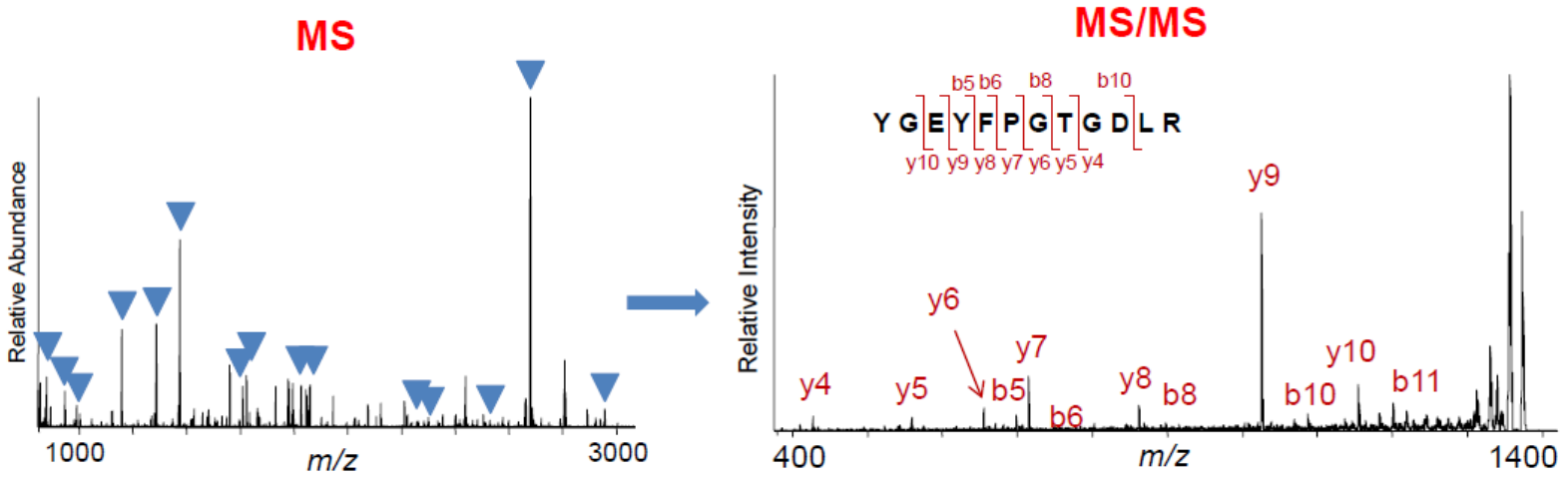
Simplified concept of peptide fragmentation







Database searching – at MS or MS/MS level



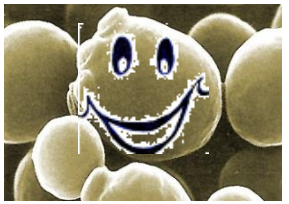
Computational Steps: massive amounts of MS data are read & interpreted. Databases searched to match peptide sequences.

Proteomics

The study of the expression, location, interaction, function, and structure of all the proteins in a given cell, organelle, tissue, organ, or whole organism.

[Study of post-translational modifications (protein phosphorylation, acetylation, glycosylation …) via MS has grown in recent years to dramatically expand the field of Proteomics]

A tour of proteomics: Studies with the budding yeast *Saccharomyces cerevisiae*



2000 & 2001

Uetz et al, A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* .
& Ito et al, A comprehensive two-hybrid analysis to explore the yeast protein interactome . *PNAS.*

➲ **Large scale yeast two hybrid screens to map proteome wide interactions.**

2001

Washburn, et al. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol*.

➲ **Established the 'shotgun' technology by showing that many proteins in a yeast-cell lysate could be identified in a single experiment.**

2002

Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*.
& Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* .

➲ **Protein–protein interaction maps can be obtained by MS; the yeast cell is organized into protein complexes.**

2003

Ghaemmaghami, S. et al. Global analysis of protein expression in yeast. *Nature*. **&** Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature*.

➲ **TAP-Tag and expression studies & GFP-Tag and localization studies**

2006

Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*.

➲ **TAP-Tag and Protein-Protein Interaction**

2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

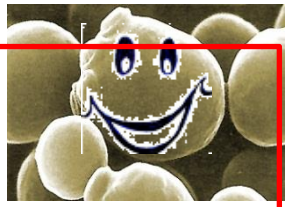
➲ **SILAC based quantitation of an entire proteome.**

2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell.*

➲ **Towards proteome wide targeted proteomics.**

A tour of proteomics: Studies with the budding yeast *Saccharomyces cerevisiae*



2000 & 2001

Uetz et al, A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* .
& Ito et al, A comprehensive two-hybrid analysis to explore the yeast protein interactome . *PNAS.*

➲ **Large scale yeast two hybrid screens to map proteome wide interactions.**

2001

Washburn, et al. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol*.

➲ **Established the 'shotgun' technology by showing that many proteins in a yeast-cell lysate could be identified in a single experiment.**

2002

Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*.
& Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* .

➲ **Protein–protein interaction maps can be obtained by MS; the yeast cell is organized into protein complexes.**

2003

Ghaemmaghami, S. et al. Global analysis of protein expression in yeast. *Nature*. **&** Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature*.

➲ **TAP-Tag and expression studies & GFP-Tag and localization studies**

2006

Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*.

➲ **TAP-Tag and Protein-Protein Interaction**

2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

➲ **SILAC based quantitation of an entire proteome.**

2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell.*

➲ **Towards proteome wide targeted proteomics.**

The *pace of proteomics is set by a combination of techniques and technological advances.

*orders of magnitude behind genome technologies (*sequencing*)

Yeast proteome reported in **Washburn et al. *Nature Biotech* 2001**:
~82 hours* = 1,484 proteins

*estimates from paper: 3 fractions @ 15 X 110 minute "runs" for each fraction

Yeast proteome by **Hunter, Colangelo, Rinehart, et al *unpublished 2010:***
One 60 minute run = 1,286 proteins AB SCIEX TripleTOF 5600

E. coli* Proteome** by **Rinehart, et al *unpublished 2012

Instrument	(samples) time	Peptides	Proteins	% Proteome (ref. Krug, 2013)
Velos (Thermo)	(1) 200 min	5,731	823	31%
Velos (Thermo)	**(35) 3,150 min**	**16,757**	**1,920**	**74%**
Q-Ex. (Thermo)	**(1) 240 min**	**15,669**	**1,750**	**67%**

A Next Generation Mass Spectrometer (Q-Exactive)
Provides Deeper and *Faster* Quantitative Proteomics

Proteins: Proteomics & Protein-Protein Interactions

Overview

- **Techniques & Technologies**
  - **Mass Spectrometry**

  - **Protein-Protein Interactions**
    - **Genetic & Biochemical Strategies**
    - **Protein Purification**

  - Quantitative Proteomics

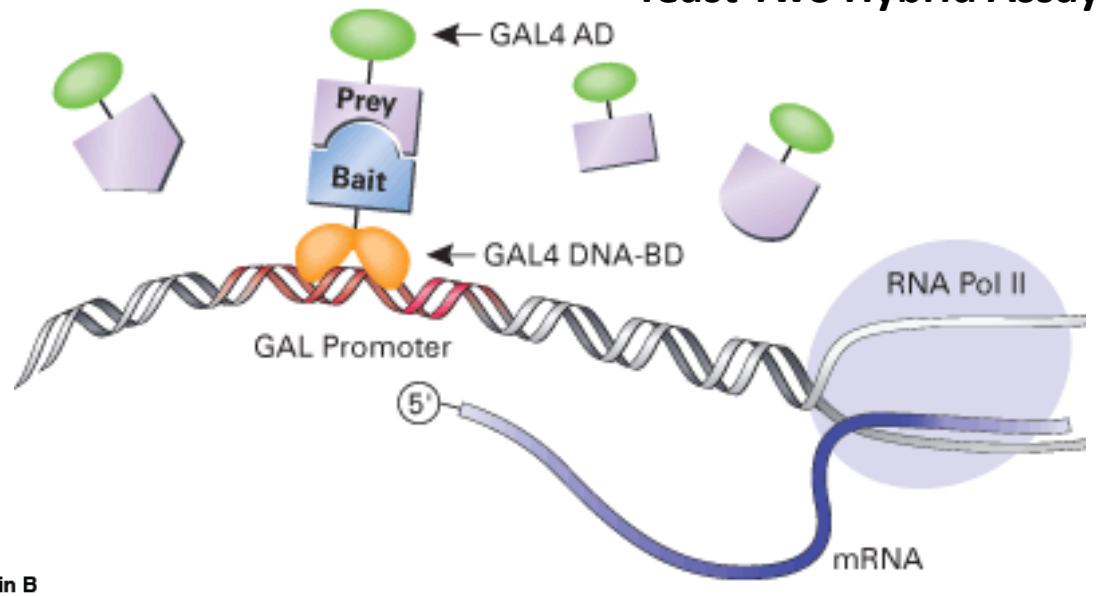
- **Applications**
  - Representative Studies

- **Putting it all together….**
  - Databases & Pathways

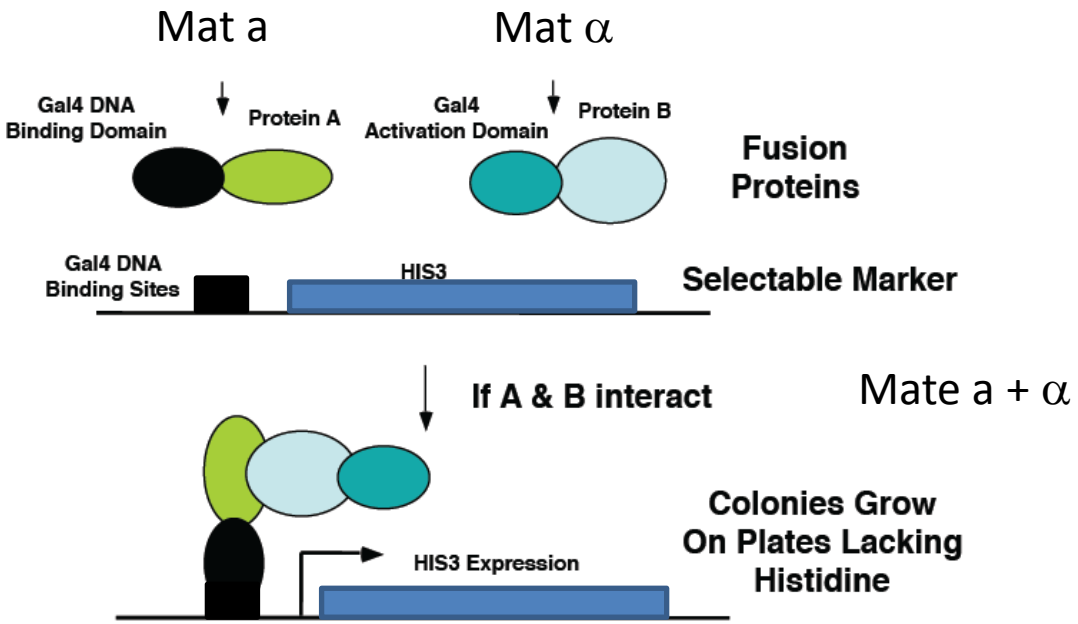
A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.

Uetz et al, Nature 2000
Ito et al, PNAS 2001

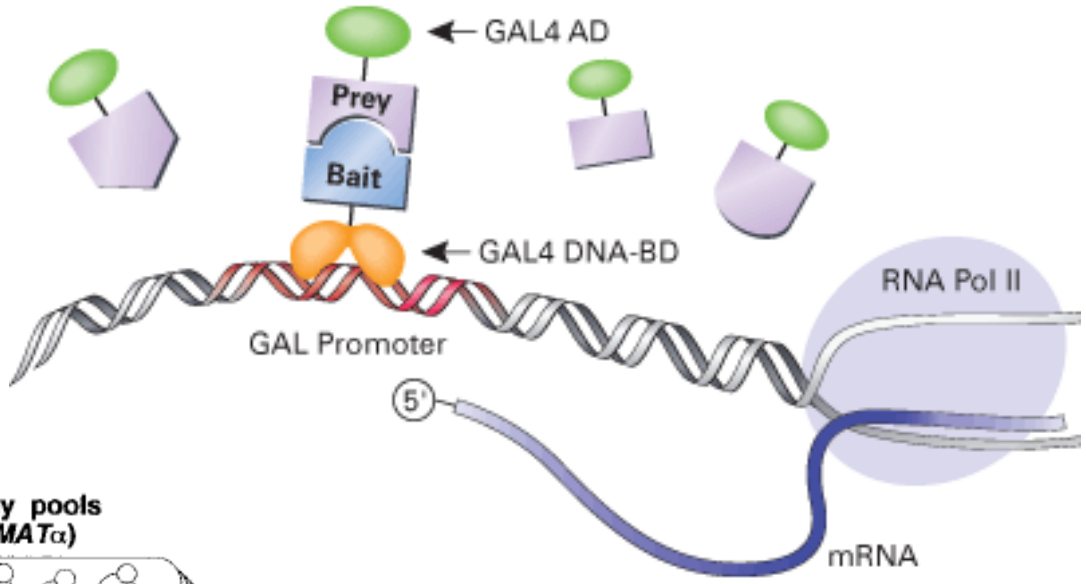
Yeast Two Hybrid Assay

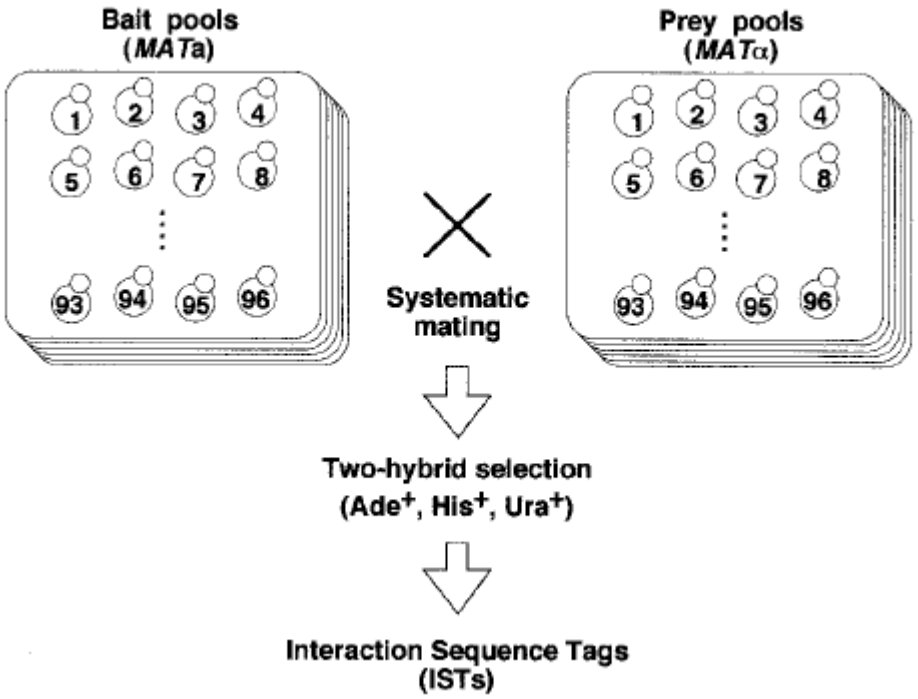


Clone bait and prey constructs and place in separate strains.



Uetz et al, Nature 2000
Ito et al, PNAS 2001





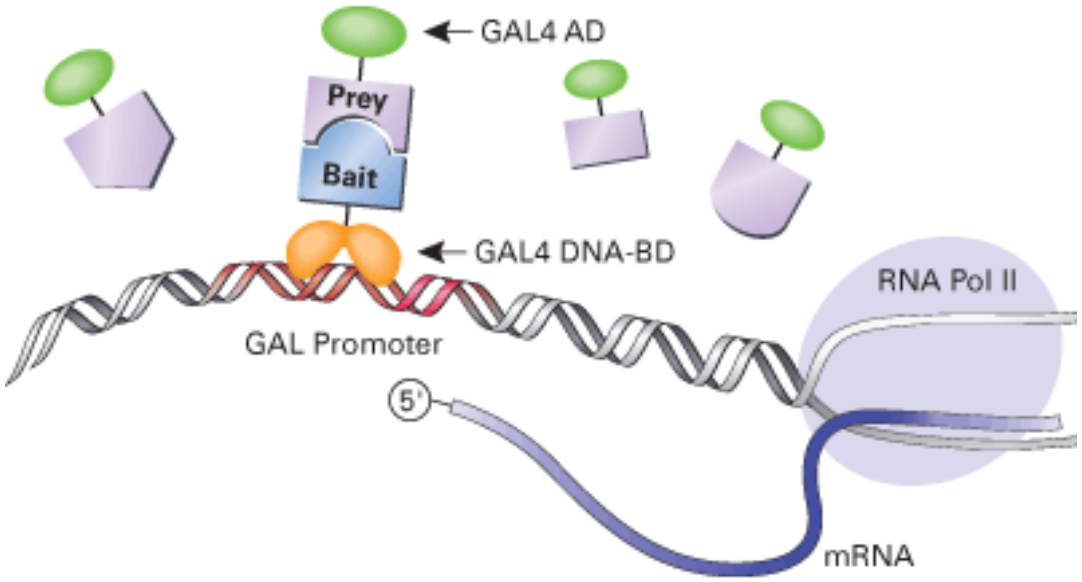
Results of Two Studies

1) 4,549 Interactions Among 3,278 Proteins (Ito et al.)

2) 957 Interactions 1004 proteins (Utez et al.)

A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.

Uetz et al, Nature 2000
Ito et al, PNAS 2001



Yeast Two Hybrid Assay

Advantages:

- ***In vivo* assay**
- **Simple**

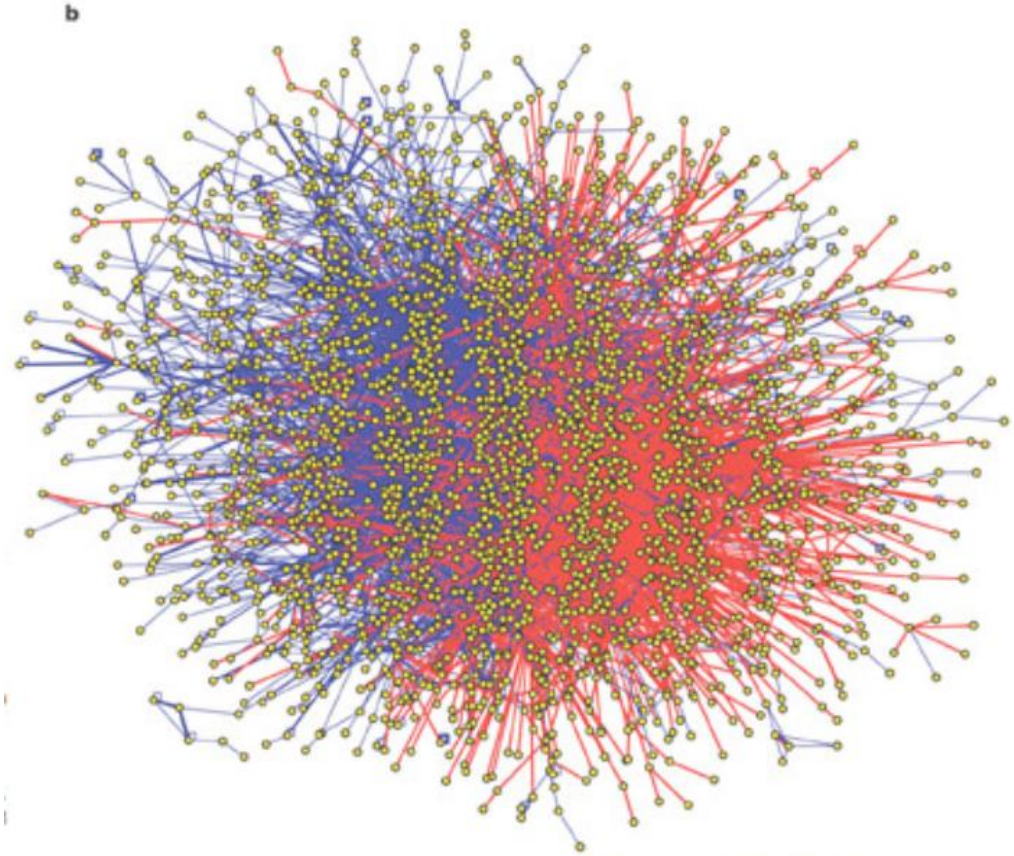
Some Disadvantages

- **Hard to execute on large scale**
- **False positives: a real interaction or "possible" interaction**
- **Interaction in nucleus (required for GAL system)**
- **Clones are fusion proteins and sometimes "partial" proteins**
- **Multiple protein complexes not "captured"**

Human Two Hybrid Map

8,100 ORFs (~7,200 genes)

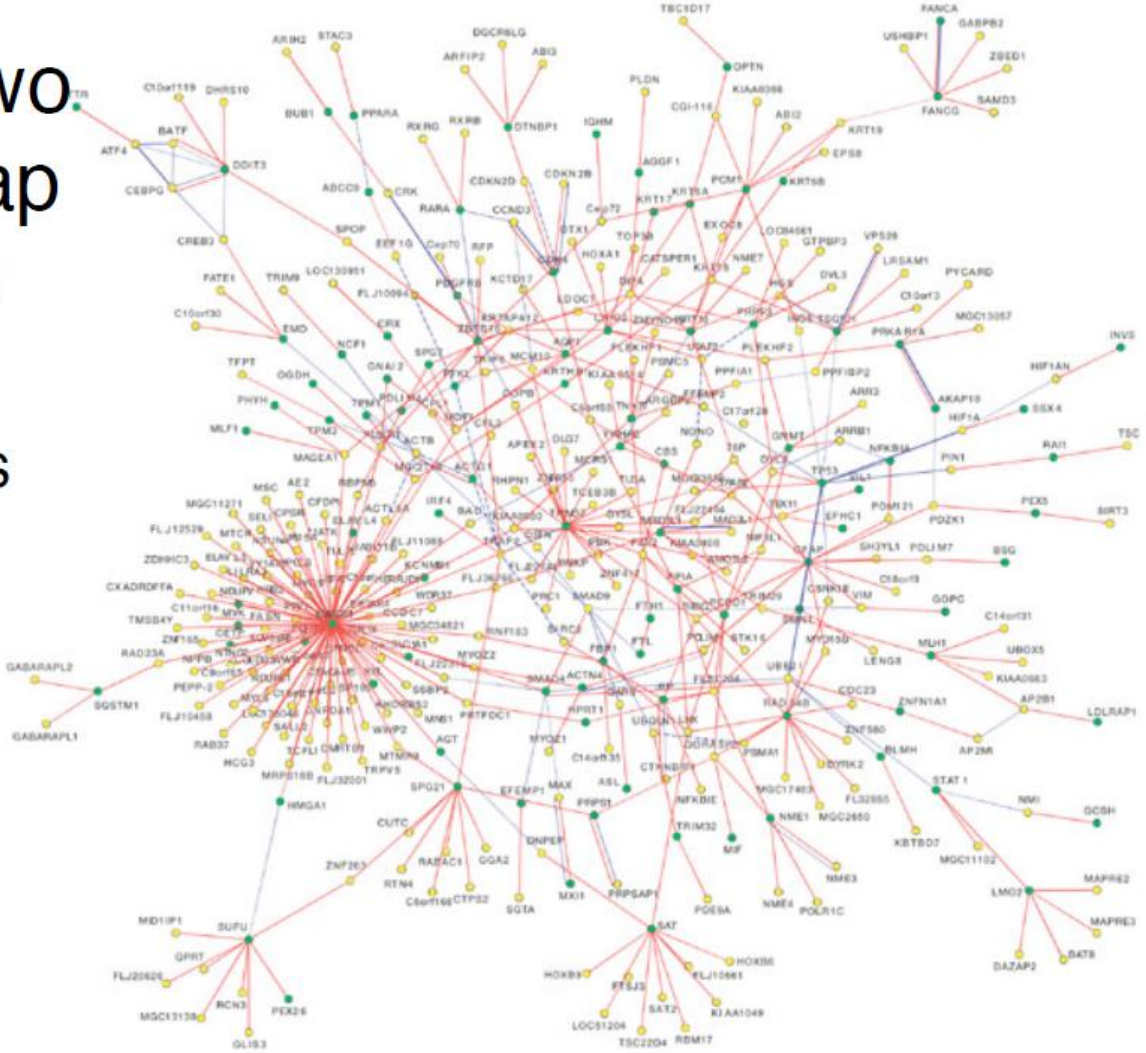
10,597 interactions



Rual et al. Nature 2005

Human Two Hybrid Map Disease Genes

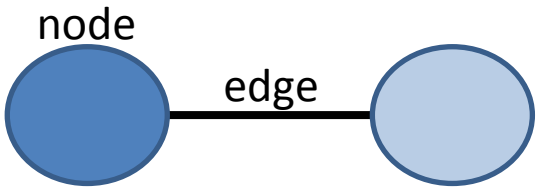
(121 genes (green))

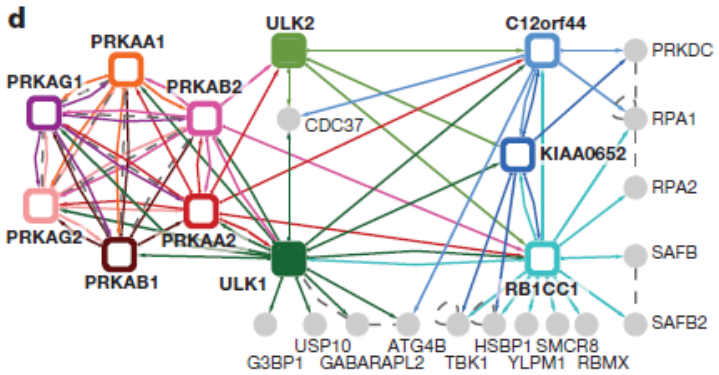


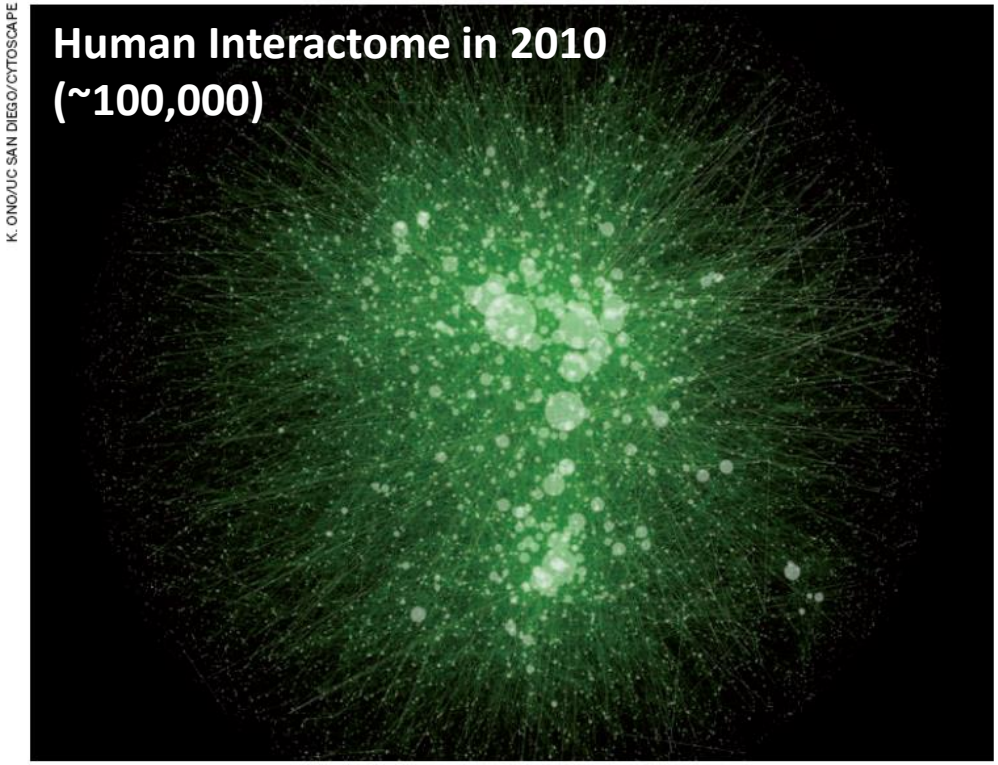
Rual et al. Nature 2005 Vol 437

Protein-Protein interaction maps:

Proteins are represented by **nodes** and interactions are represented by **edges** between nodes.

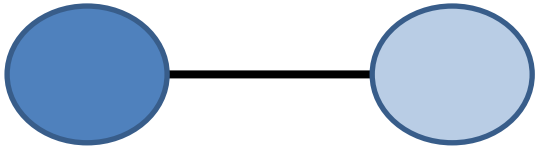






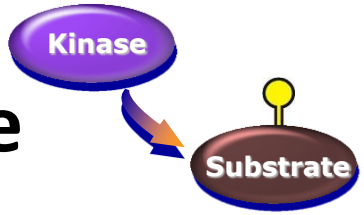
Bonetta, *Nature* 2010

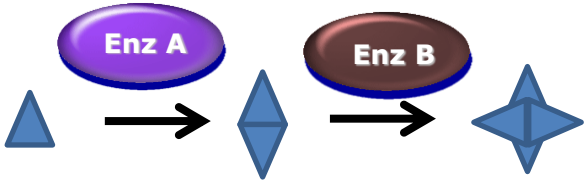
Protein-Protein interactions:



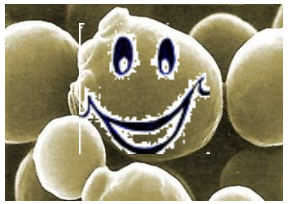
Some examples:

- Physical and direct
- Physical and indirect
  - Multi-protein complexes
  - Scaffolds
- Transient
  - Kinase & substrate
- Metabolic





A tour of proteomics: Studies with the budding yeast *Saccharomyces cerevisiae*



2000 & 2001

Uetz et al, A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* .
& Ito et al, A comprehensive two-hybrid analysis to explore the yeast protein interactome . *PNAS.*

➲ **Large scale yeast two hybrid screens to map proteome wide interactions.**

2001

Washburn, et al. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol*.

➲ **Established the 'shotgun' technology by showing that many proteins in a yeast-cell lysate could be identified in a single experiment.**

2002

Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*.
& Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* .

➲ **Protein–protein interaction maps can be obtained by MS; the yeast cell is organized into protein complexes.**

2003

Ghaemmaghami, S. et al. Global analysis of **protein expression** in yeast. *Nature*. **&** Huh, W. K. et al. Global analysis of **protein localization** in budding yeast. *Nature*.

➲ **TAP-Tag and expression studies & GFP-Tag and localization studies**

2006

Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*.

➲ **TAP-Tag and Protein-Protein Interaction**

2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

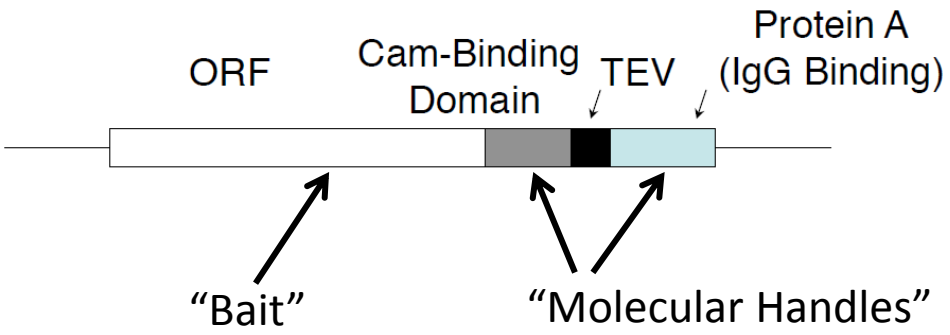
➲ **SILAC based quantitation of an entire proteome.**

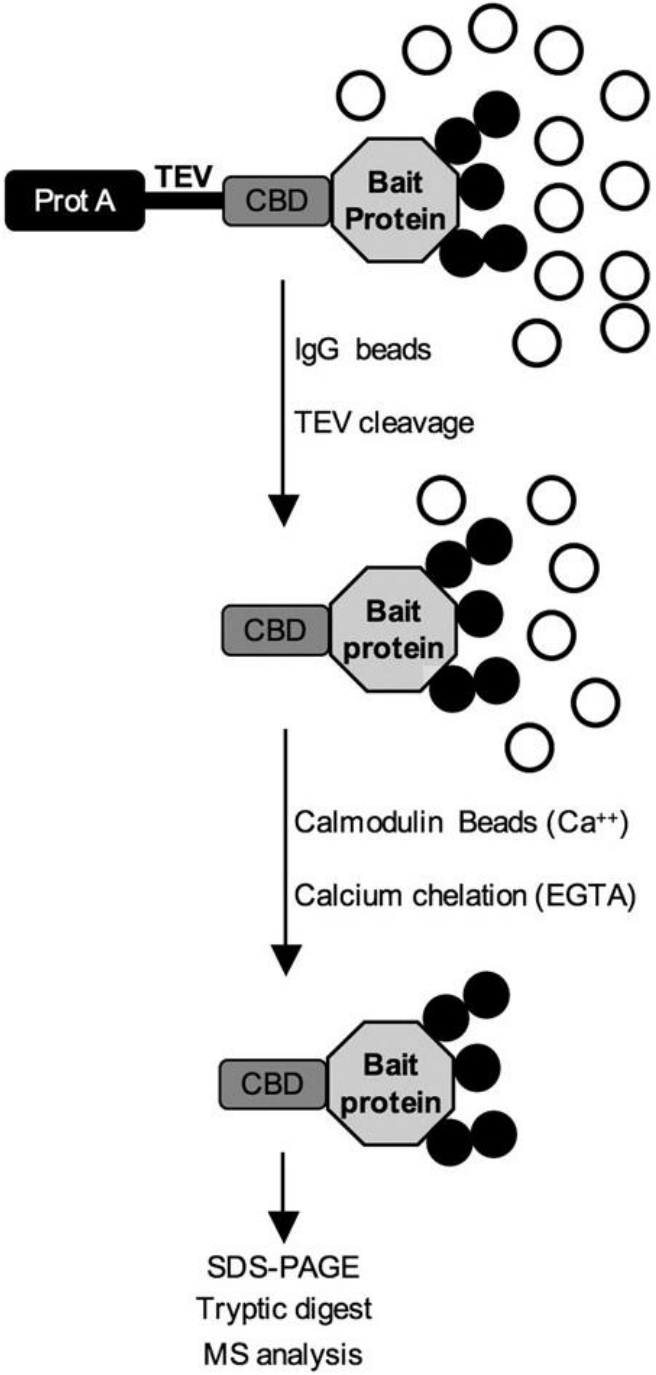
2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell.*

➲ **Towards proteome wide targeted proteomics.**

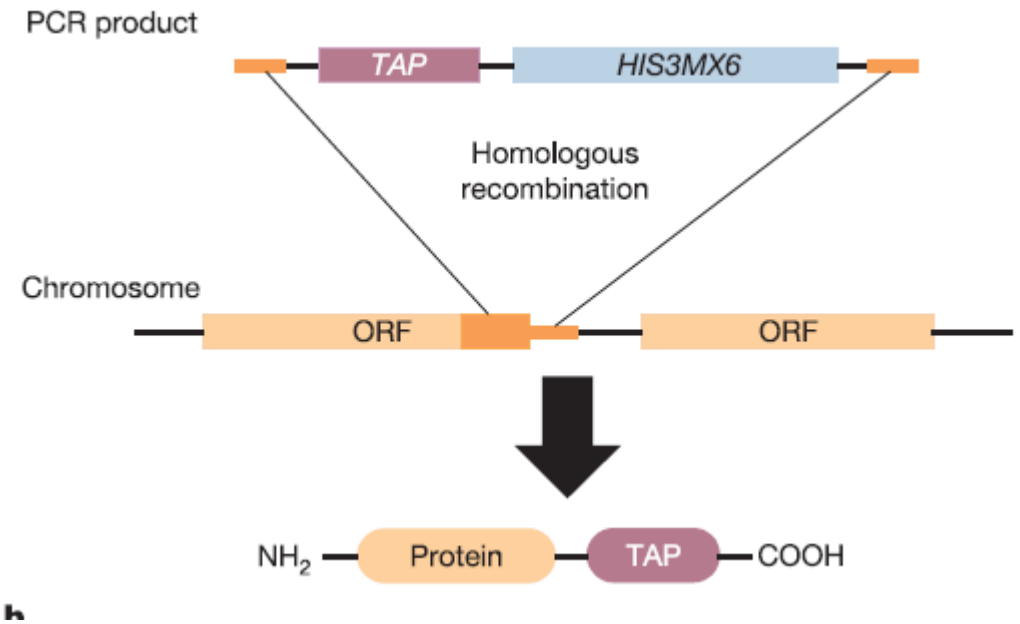
Tandem Affinity Purification (TAP) Tagging





Cannavo E et al. J. Biol. Chem. 2007

Global TAP Tagging in yeast



2003

Ghaemmaghami, S. et al. Global analysis of protein expression in yeast. *Nature*. **&** Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature*.

➲ **TAP-Tag and expression studies & GFP-Tag and localization studies**

2002

Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*.
& Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* .

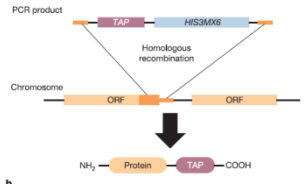
➲ Protein–protein interaction maps can be obtained by MS; the yeast cell is organized into protein complexes.

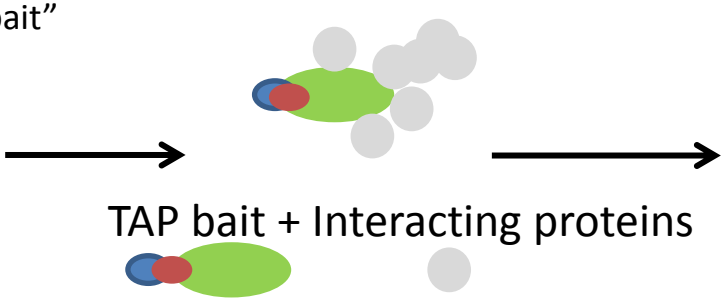
2006

Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*.

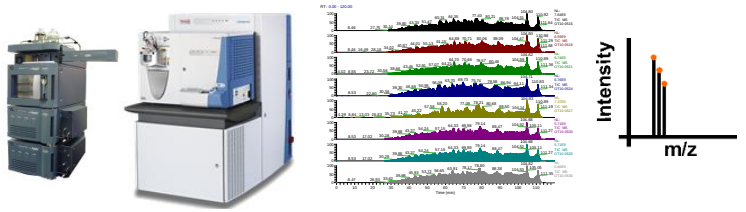
➲ TAP-Tag and Protein-Protein Interaction

Collection of tagged "bait" expression strains





Multiple runs of "shot gun" MS & SDS-PAGE with MS on individual proteins

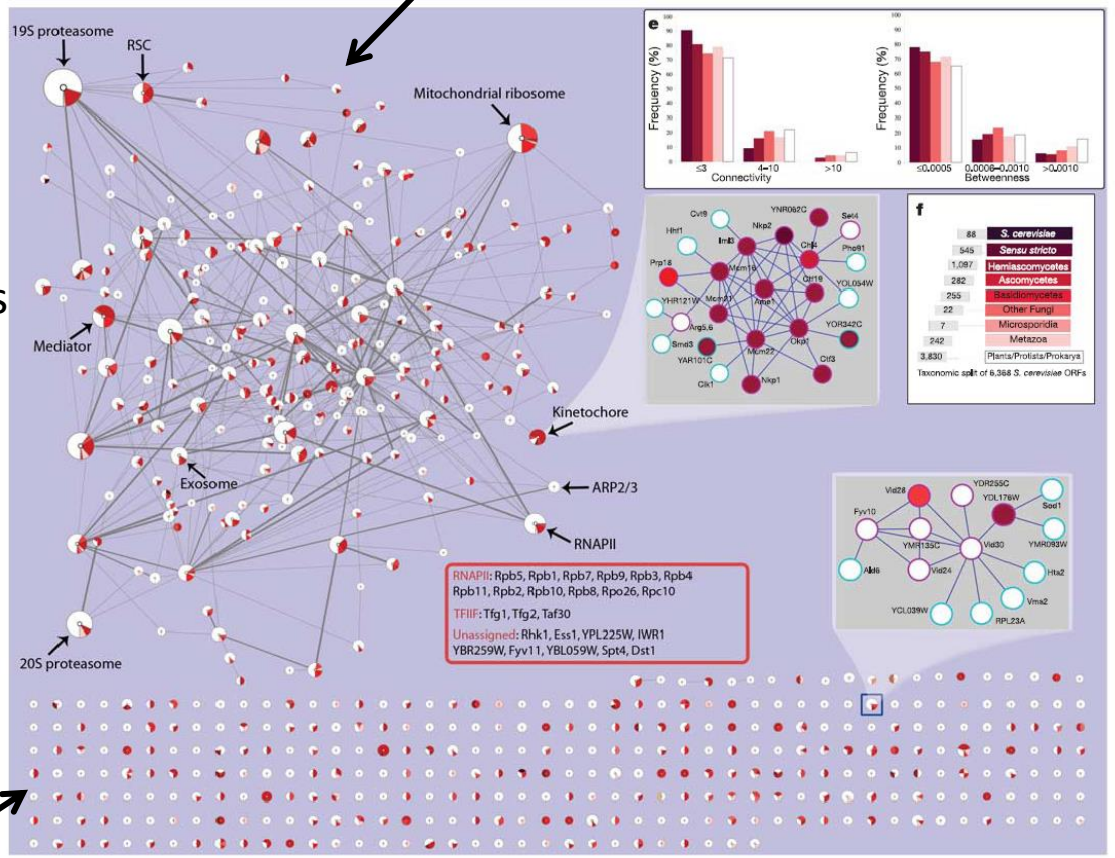


Krogan et al. observed 7,123 protein–protein interactions:

<u>Important aspects:</u>

- **Tagged the native genes and did not overexpress the fusion proteins**
- **Could immediately validate partners (reciprocal purification in data set)**
- **Complementary MS techniques, deeper coverage of complexes**
- **Authors state, "…rigorous computational procedures to assign confidence values to our predictions…"**

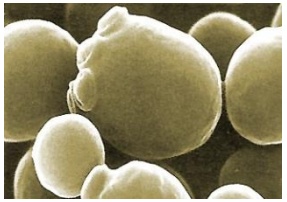
- 4,562 tagged proteins
- 2,357 successful purifications
- Identified 4,087 interacting proteins ~72 % proteome
- Majority of the yeast proteome is organized into complexes
- Many complexes are conserved in other species



Complexes with little or no interconnectivity

Krogan NJ, et al. *Nature.* 2006

How do we learn more about the organization of the human proteome?

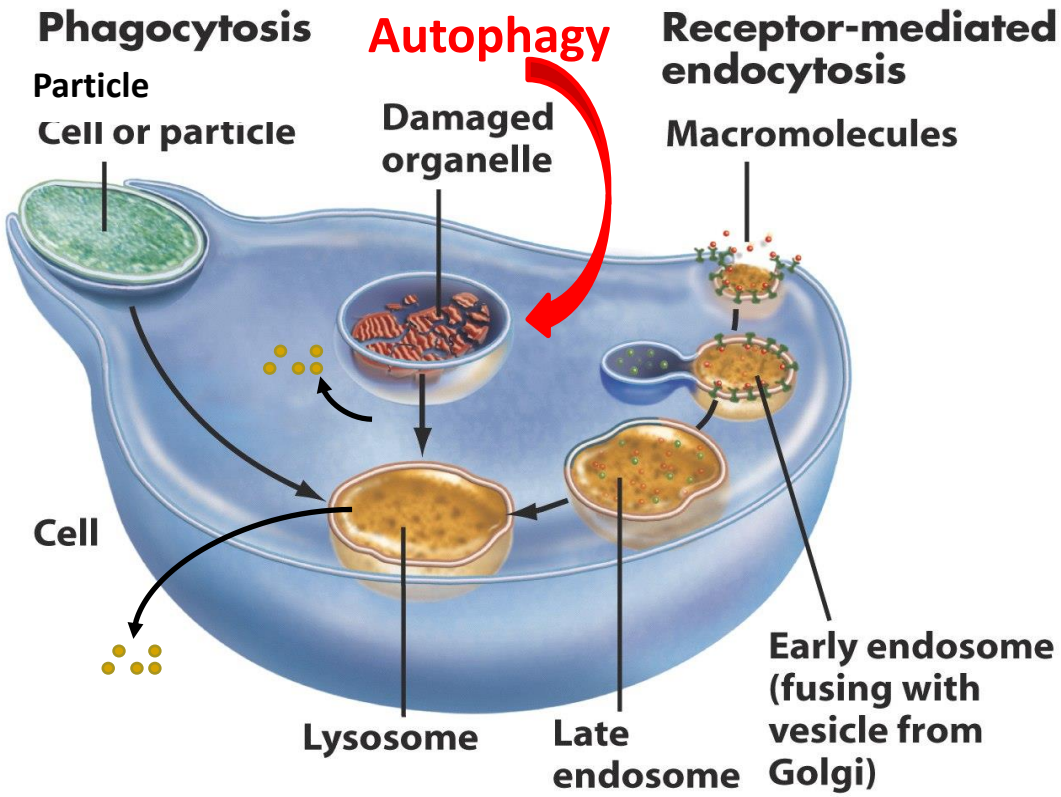


nature Vol 466|1 July 2010|doi:10.1038/nature09204

ARTICLES

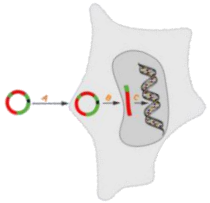
Network organization of the human autophagy system

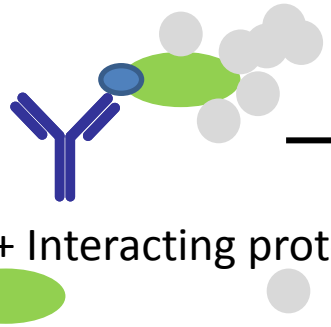
Christian Behrends[1], Mathew E. Sowa[1], Steven P. Gygi[2] & J. Wade Harper[1]



Pearson Prentice Hall, Inc. 2005
www.stolaf.edu/people/giannini/cell/lys.htm

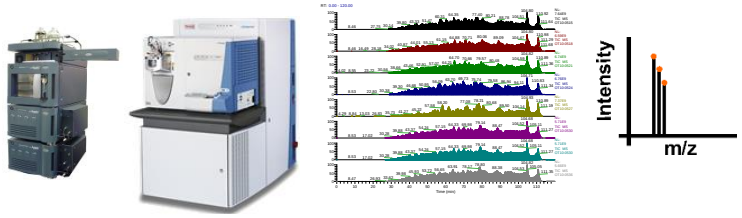
Transfect tagged "bait"





IP Bait + Interacting proteins

Multiple runs of "shot gun" LC-MS/MS



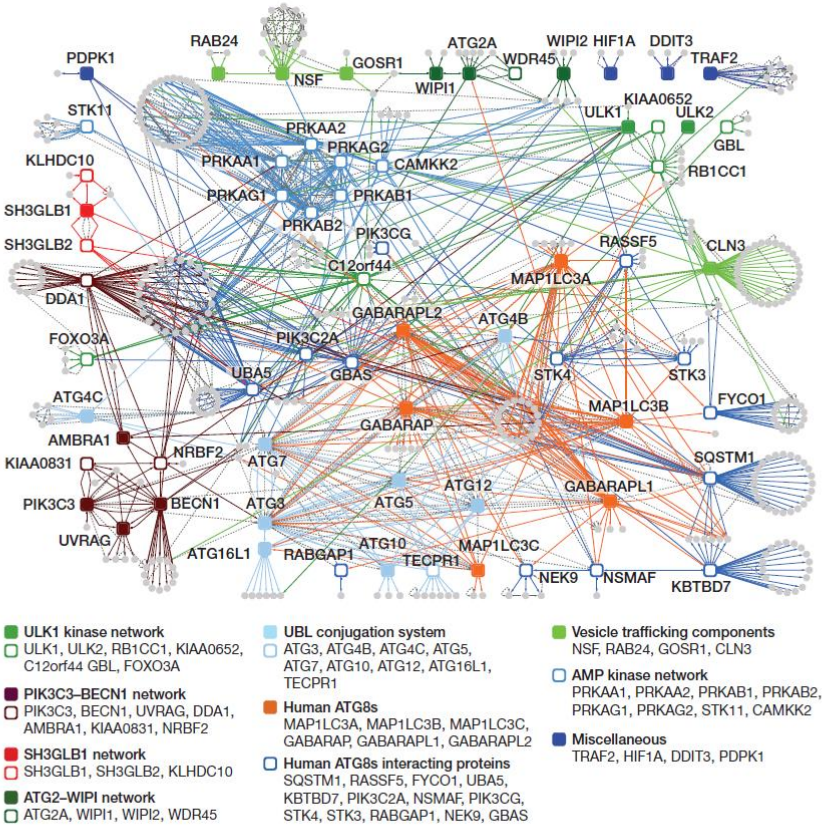
~65 bait proteins
LC-MS/MS identifies
2553 proteins

Data analysis to sort out real interaction from background

Authors use CompPASS to identify High-Confidence Interacting Proteins (HCIP)

763 HCIPs identified that compose The Autophagy Interaction Network

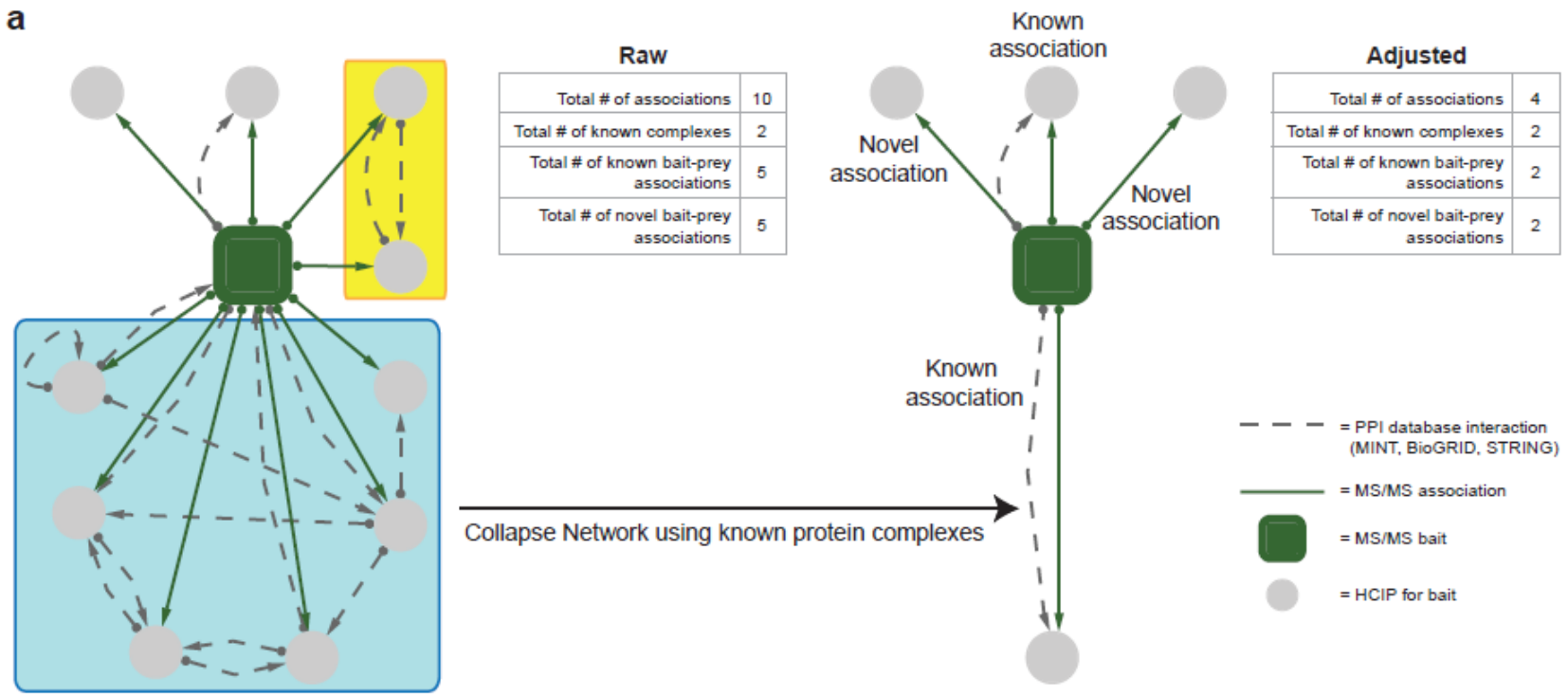
Autophagy Interaction Network



Behreands et al, Nature 2010

Figure 1 | Overview of the autophagy interaction network (AIN). HCIPs within the autophagy network are shown for 32 primary baits (filled squares) and 33 secondary baits (open squares). Subnetworks are colour-coded. Interacting proteins are indicated by grey circles.

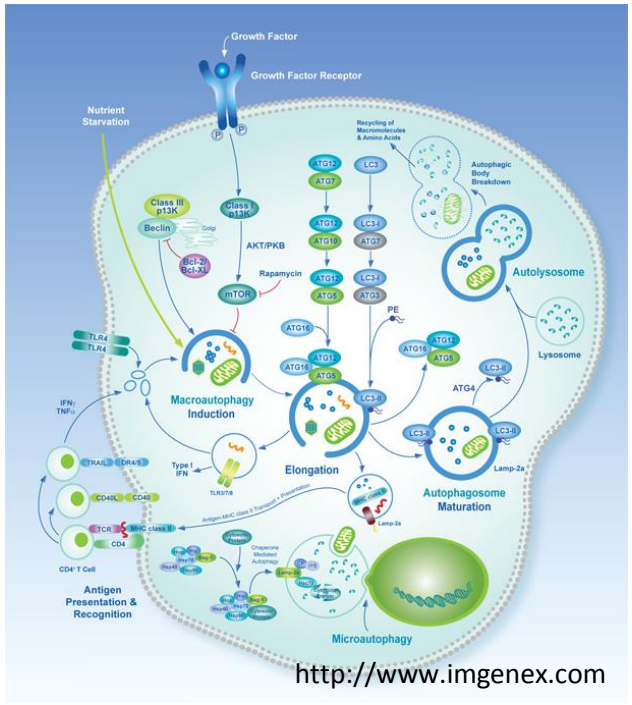
Concatenating Data: Supplementing observed protein complexes with the Protein-Protein Interaction Databases: MINT, BioGRID, STRING



HCIPs	751	Raw Network
Known Protein Complexes (KPC)	84	
Known Bait-KPC Interactions	40	
Total Associations	497	Collapsed Network
Known Bait-HCIP/KPC Associations	68	
Novel Bait-HCIP Associations	429	

Behreands et al, Nature 2010

A Functional Organization of the Network



http://www.imgenex.com

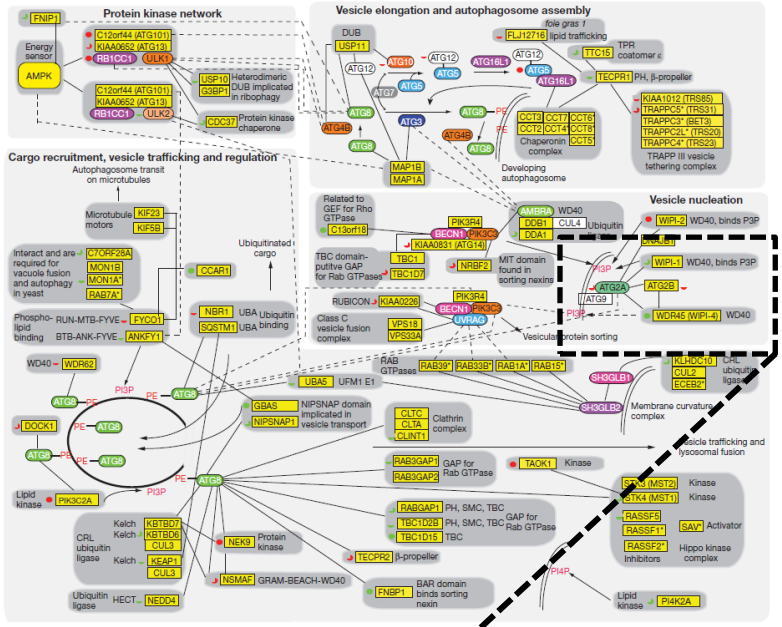
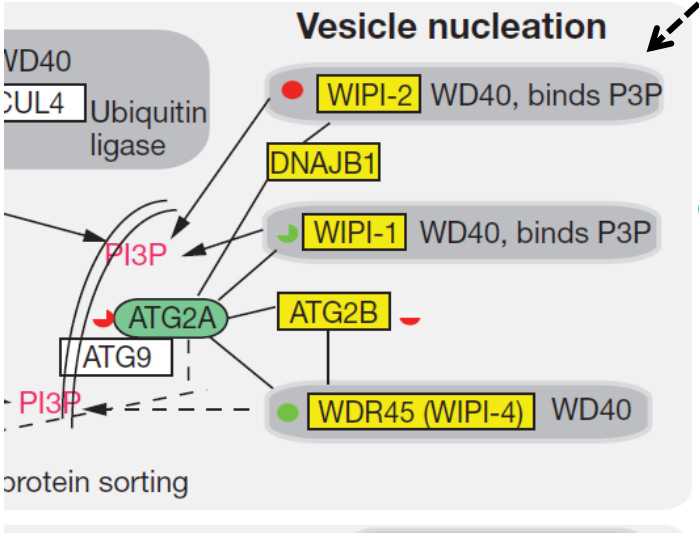


Figure 6 | Functional integration of the autophagy interaction network. Proteins in yellow boxes are HCIPs and proteins labelled with an asterisk were subthreshold (WD^N-score <1.0) for HCIP identification. Dotted lines, cross-module interaction; solid lines, this study; dotted arrows, potential functional interactions. Red or green full circle, three-quarter circle and half-circle represent a reduction or increase in autophagosomes with 4, 3 and 2 siRNAs, respectively, in GFP–MAP1LC3B-expressing U2OS cells without rapamycin. Proteins in white boxes were not found by proteomics. The six ATG8 family members are represented by ATG8.



Phenotype after siRNA

- Less autophagosomes
- More autophagosomes

Behreands et al, Nature 2010

The Hippo Signaling Pathway Interactome

Young Kwon,[1] Arunachalam Vinayagam,[1]* Xiaoyun Sun,[3]* Noah Dephoure,[4] Steven P. Gygi,[4] Pengyu Hong,[3] Norbert Perrimon[1,2]†

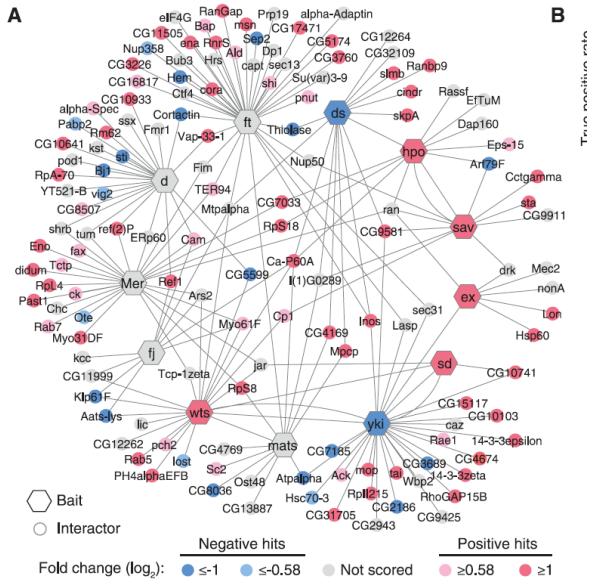
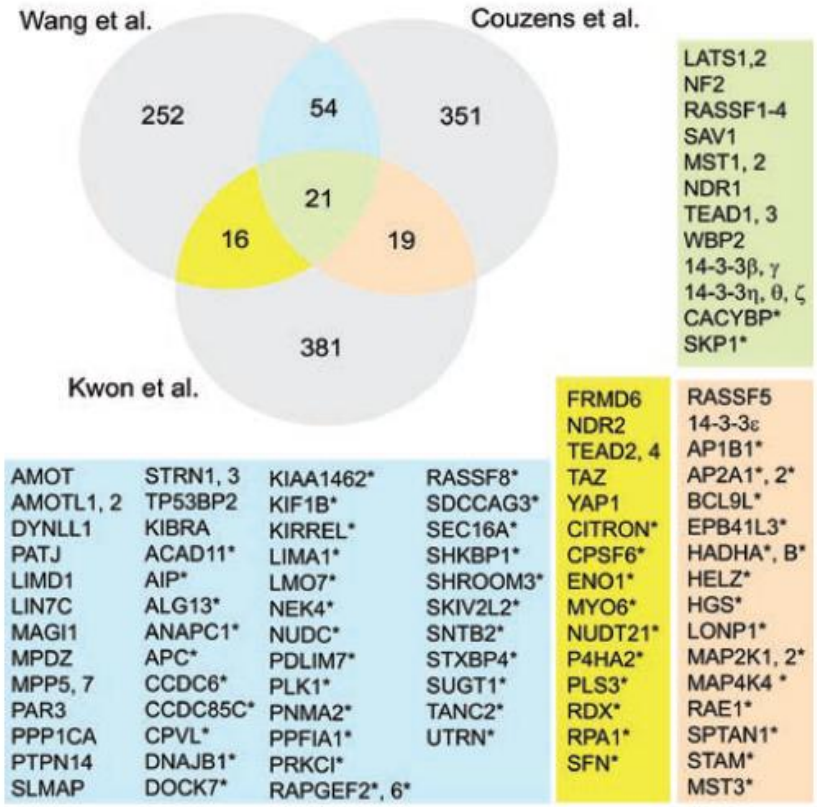


Fig. 2. Validation of Hippo-PPIN with functional RNAi screen and co-IP. (**A**) Distribution of Yki-reporter values for individual double-stranded RNAs (dsRNAs) in our focused RNAi screen. About 70% of genes are covered by two dsRNAs. (**B**) Recovery of Hippo pathway components from RNAi screen [fold-change (log_2) cutoff ± 1]. (**C**) The positive



Cell Research (2014) 24:137-138.

www.nature.com/cr

RESEARCH HIGHLIGHT

Discovering the Hippo pathway protein-protein interactome

Cell Research (2014) **24**:137-138. doi:10.1038/cr.2014.6; published online 14 January 2014

Kwon Y, *et al*. *Science* 2013; Couzens AL, *et al*. *Sci Signal* 2013; Wang W, *et al*. *Mol Cell Proteomics* 2014

Protein-Protein Interaction Databases

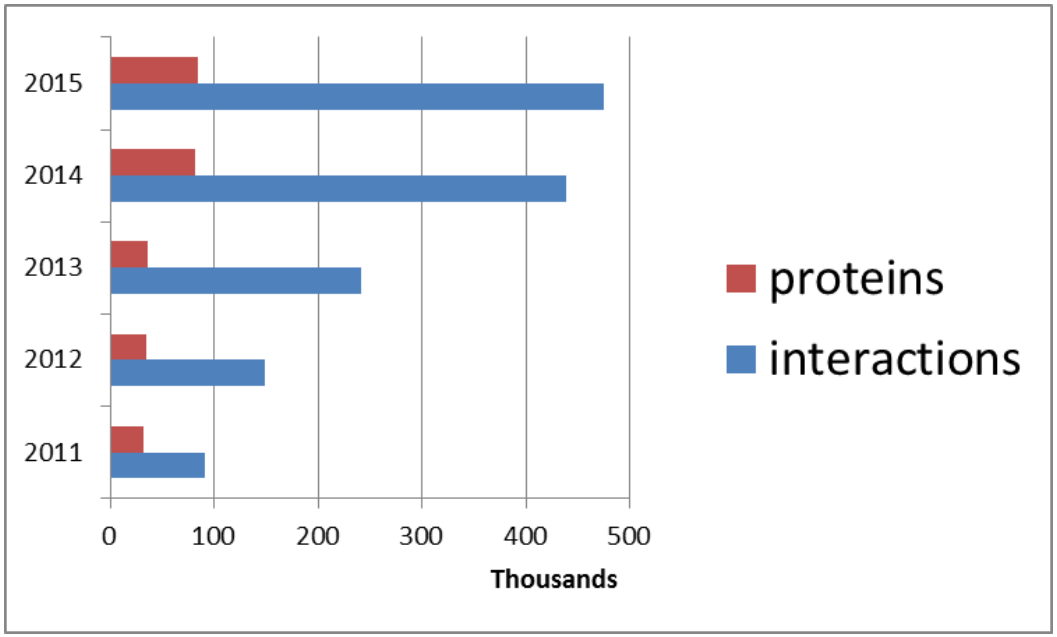
http://www.ebi.ac.uk/intact/

EMBL-EBI

IntAct

Data Content

- Publications: **13297**
- Interactions: **475733**
- Interactors: **84952**



2015 IntAct

+ 234,275 interactions

+ 49,520 proteins

2013 MINT

Statistics:

241458 interactions

35432 proteins

5510 pmids

Protein interaction networks:

Some of the many important aspects:

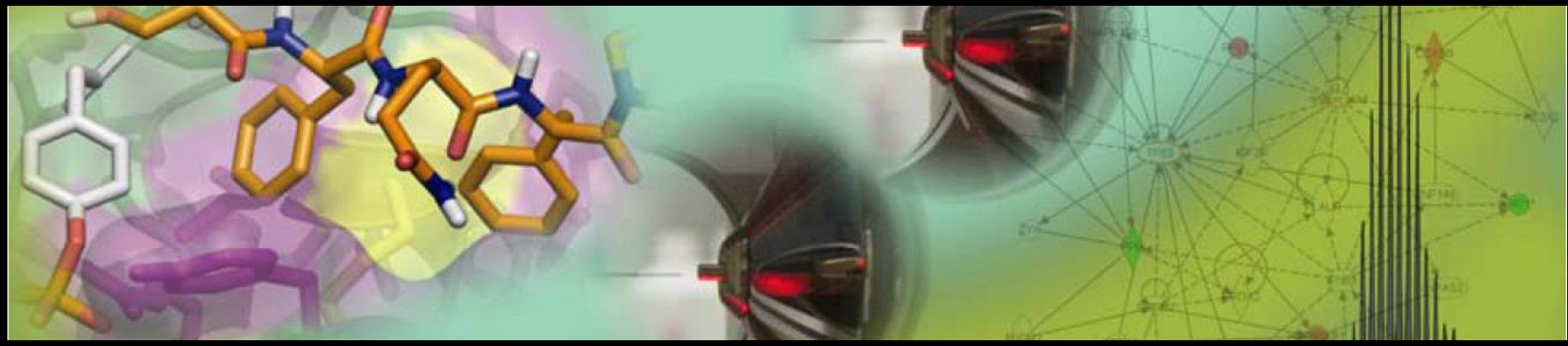
- Parts List
- Organization and assembly
- Biological function can be inferred

However:

- Interaction data is largely static

Next Step:

- How do protein interaction networks change over time?



Proteins: Proteomics & Protein-Protein Interactions Part II

Jesse Rinehart, PhD

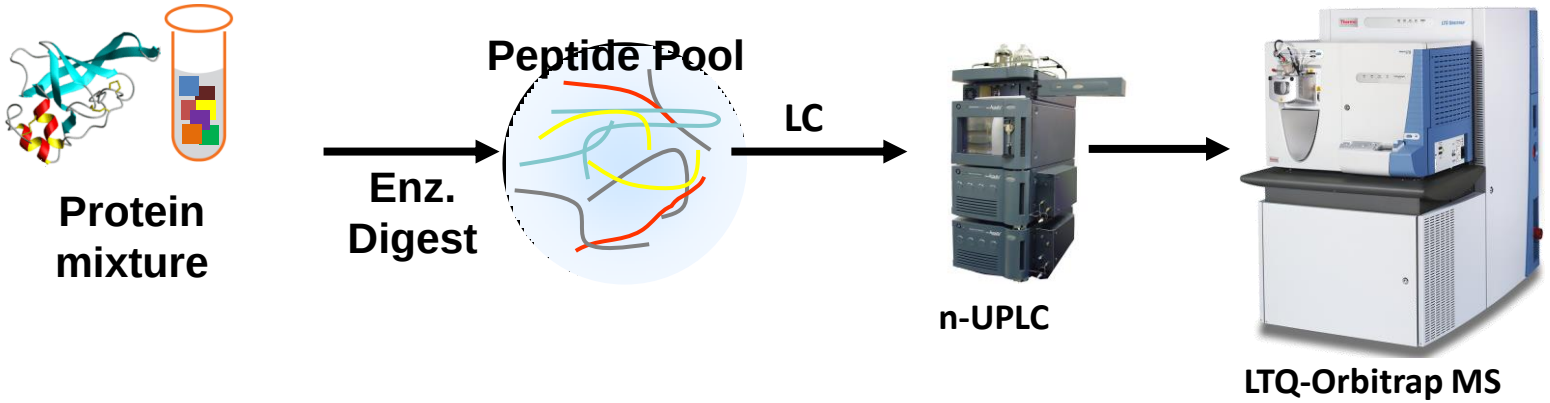
Yale Systems Biology Institute

Cellular & Molecular Physiology
Yale University School of Medicine

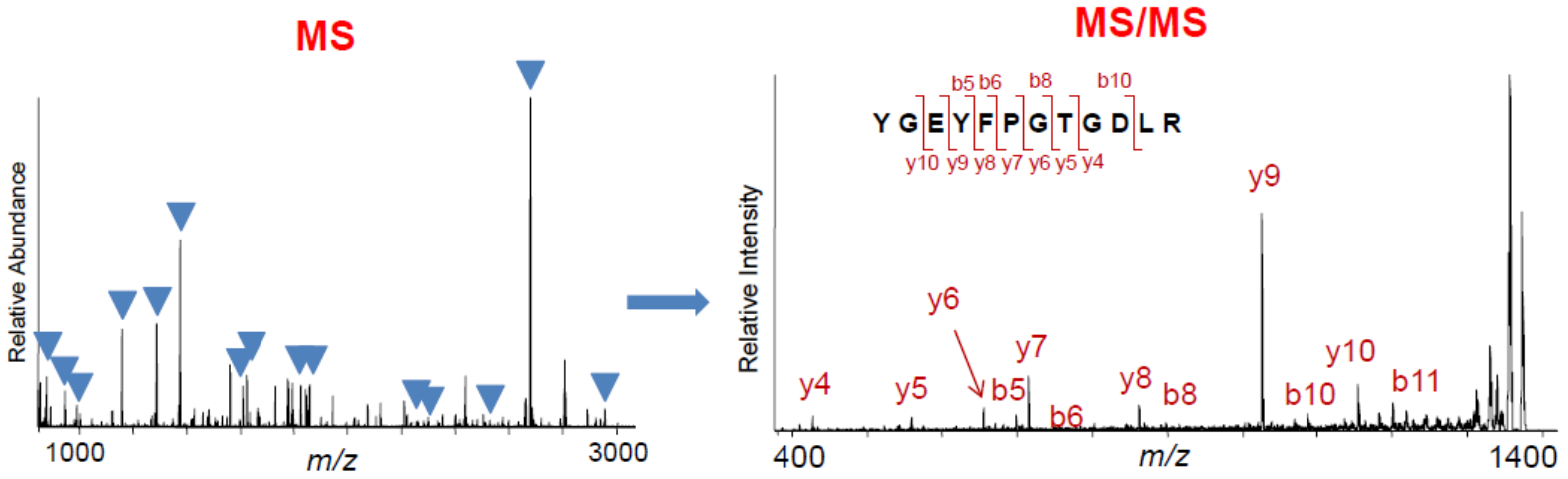
Proteins: Proteomics & Protein-Protein Interactions

Overview

- **Techniques & Technologies**
  - **Mass Spectrometry**
  - Protein-Protein Interactions
    - Genetic & Biochemical Strategies
    - Protein Purification
  - **Quantitative Proteomics**
- **Applications**
  - **Representative Studies**
- **Putting it all together….**
  - Databases & Pathways

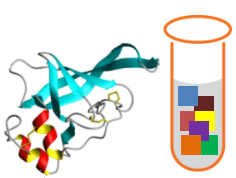


Database searching – at MS or MS/MS level



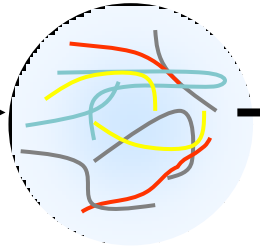
Computational Steps: massive amounts of MS data are read & interpreted. Databases searched to match peptide sequences.

Typical work flow for LC-MS "shotgun proteomics"



Protein mixture

Enz. Digest



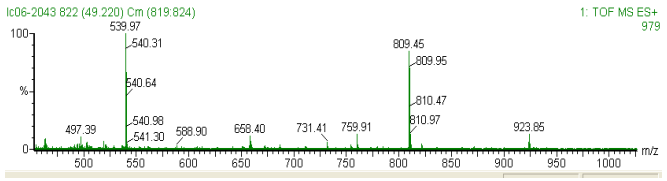
Peptide Pool

LC

n-UPLC

LTQ-Orbitrap MS

MS



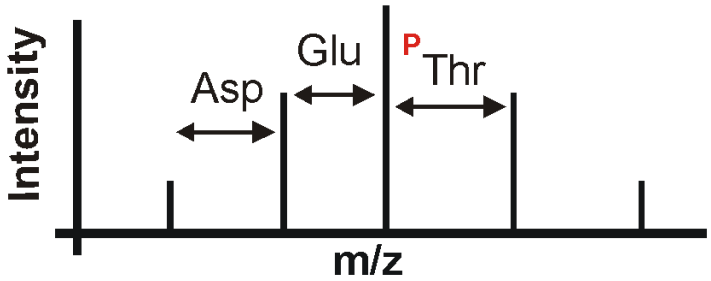
peptide

peptide

isolate & fragment

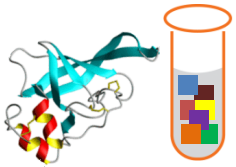
peptide

MS/MS



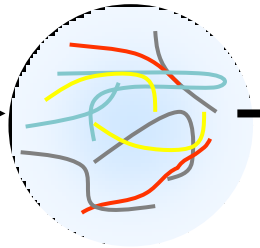
peptide fragments

Typical work flow for LC-MRM "targeted proteomics"



Protein mixture

Enz. Digest



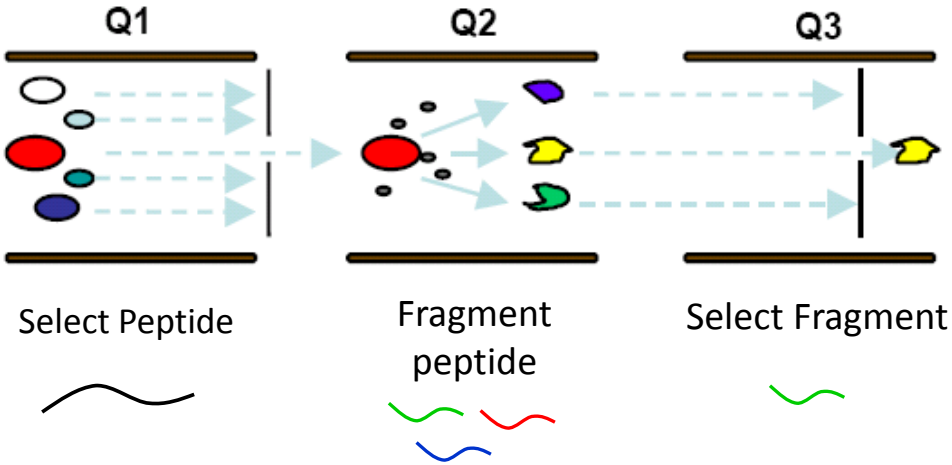
Peptide Pool

LC

n-UPLC

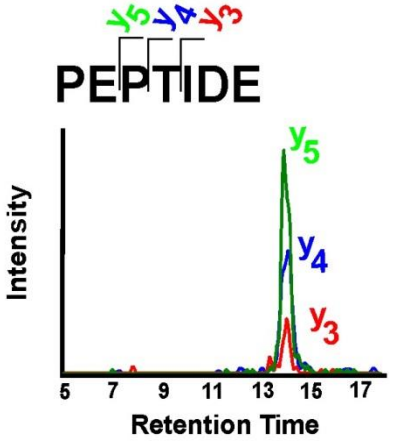
QQQ MS

Multiple Reaction Monitoring (MRM)



Detect Fragment

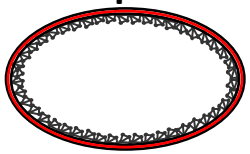
MRM Spectra

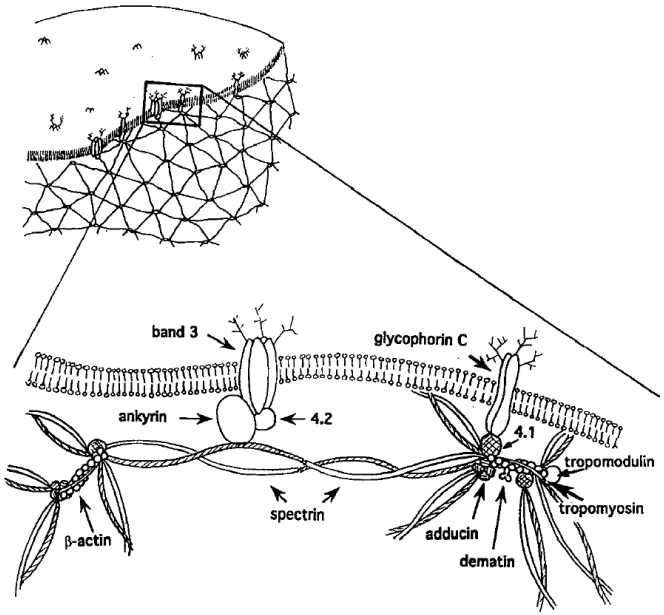


MS Data is not inherently quantitative, *but …*

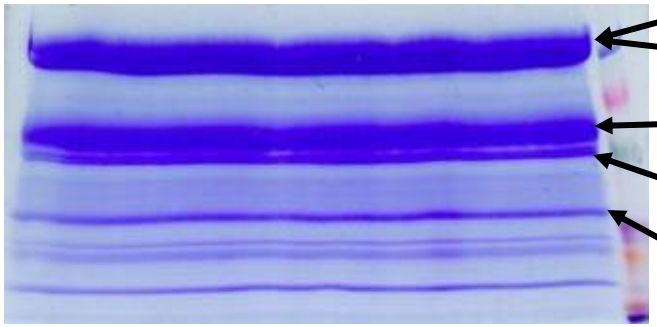
Red Blood Cell

RBC membrane:
a native multi-protein complex





RBC membrane proteome
Coomassie Stained
SDS-PAGE (250 ug Protein)
~16 bands



RBC membrane proteome
Shotgun Proteomics
1ug Peptides (242 Proteins)

# peptides (unique)	
352 (291)	Spectrin alpha chain, erythrocyte OS=Homo sapiens GN=SPTA1 PE=1 SV=5
291 (233)	Spectrin beta chain, erythrocyte OS=Homo sapiens GN=SPTB PE=1 SV=5
172 (134)	Ankyrin-1 OS=Homo sapiens GN=ANK1 PE=1 SV=3
57 (46)	Band 3 anion transport protein OS=Homo sapiens GN=SLC4A1 PE=1 SV=3
52 (39)	Erythrocyte membrane protein band 4.2 OS=Homo sapiens GN=EPB42 PE=1 SV=3
43 (34)	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1
30 (20)	Actin, alpha cardiac muscle 1 OS=Homo sapiens GN=ACTC1 PE=1 SV=1
22 (9)	Beta-actin-like protein 2 OS=Homo sapiens GN=ACTBL2 PE=1 SV=2
28 (6)	POTE ankyrin domain family member J OS=Homo sapiens GN=POTEJ PE=3 SV=1
68 (49)	Protein 4.1 OS=Homo sapiens GN=EPB41 PE=1 SV=4

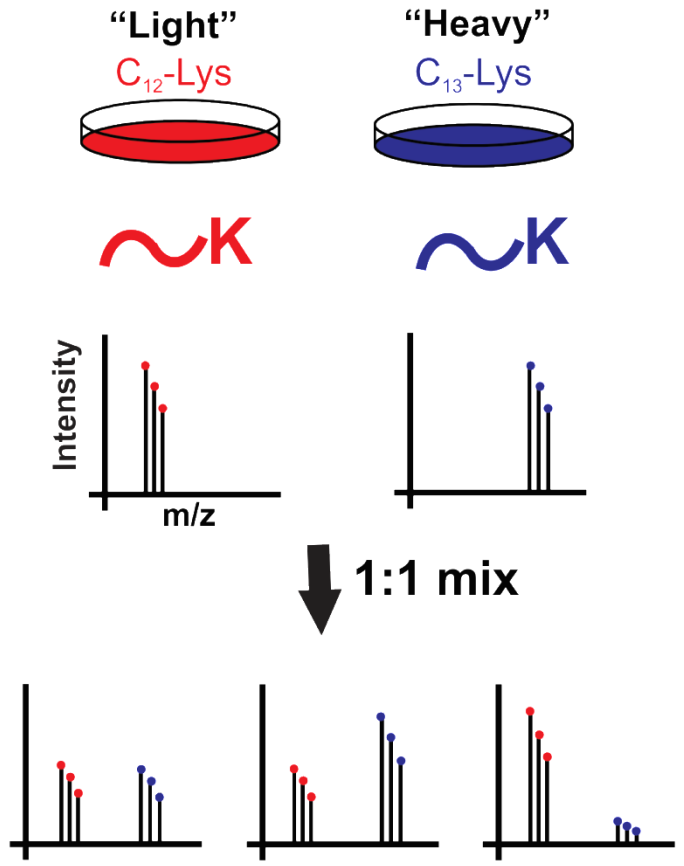
Rinehart et al., *unpublished*

Quantitative Proteomics

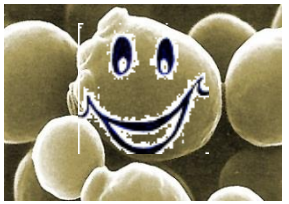
S.I.L.A.C. - Stable isotope labeling with amino acids in cell culture

-Ong S.E. et al. *Molecular & Cell Proteomics* 2002

- Stable isotopes are *not radioactive*, and they occur naturally in nature. For example, 99% of all carbon in the world is carbon-12 (^{12}C) and 1% is carbon-13 (^{13}C).
- SILAC reagents have enriched stable isotopes that have been placed into compounds in abundances much greater than their natural abundance.
- We can obtain labeled compounds with ~95-99% ^{13}C.
- Because a mass spectrometer separates ions by mass, we use mass spectrometry to distinguish isotopes in compounds by their mass.
- Simultaneous comparison in the same MS run is key



A tour of proteomics: Studies with the budding yeast *Saccharomyces cerevisiae*



2000 & 2001

Uetz et al, A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* .
& Ito et al, A comprehensive two-hybrid analysis to explore the yeast protein interactome . *PNAS.*

➲ **Large scale yeast two hybrid screens to map proteome wide interactions.**

2001

Washburn, et al. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol*.

➲ **Established the 'shotgun' technology by showing that many proteins in a yeast-cell lysate could be identified in a single experiment.**

2002

Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*.
& Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* .

➲ **Protein–protein interaction maps can be obtained by MS; the yeast cell is organized into protein complexes.**

2003

Ghaemmaghami, S. et al. Global analysis of protein expression in yeast. *Nature*. **&** Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature*.

➲ **TAP-Tag and expression studies & GFP-Tag and localization studies**

2006

Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*.

➲ **TAP-Tag and Protein-Protein Interaction**

2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

➲ **SILAC based quantitation of an entire proteome.**

2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell.*

➲ **Towards proteome wide targeted proteomics.**

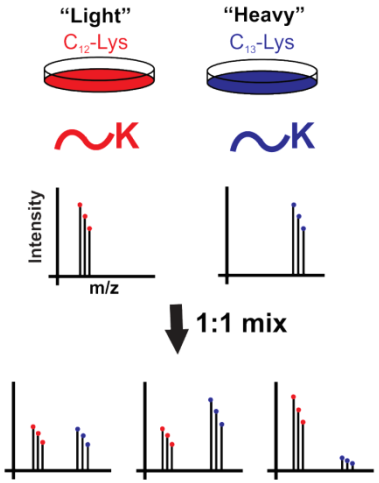
2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

➲ **SILAC based quantitation of an entire proteome.**

S.I.L.A.C. - Stable isotope labeling with amino acids in cell culture

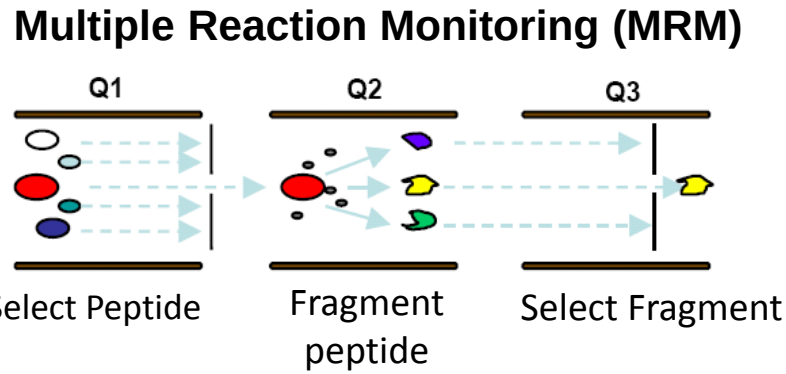
-Ong SE et al. *Molecular & Cell Proteomics* 2002.

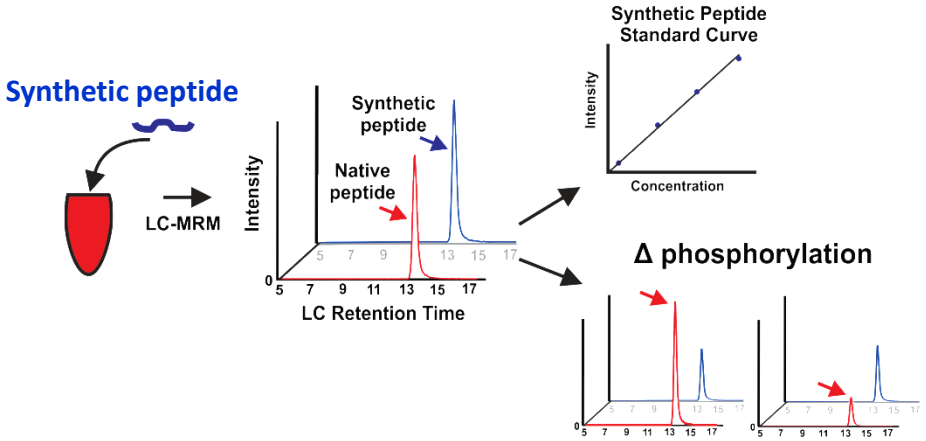


2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell*.

➲ **Towards proteome wide targeted proteomics.**





2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*. 30;455(7217):1251-4.

➲ **SILAC based quantitation of an entire proteome.**

Table 1 | Yeast ORFs identified by SILAC-based quantitative proteomics

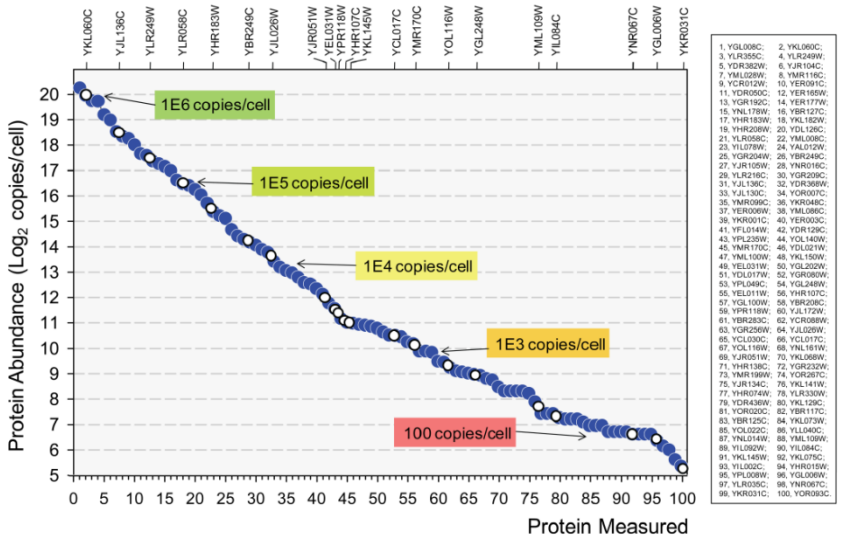
	Number of ORFs	TAP	GFP	nanoLC–MS
Total yeast ORFs	6,608	4,251	4,154	4,399
Characterized yeast ORFs	4,666	3,629	3,581	3,824
Uncharacterized yeast ORFs	1,128	581	539	572
Dubious yeast ORFs	814	26 (3%)	23 (3%)	3 (<1%)
Not present in ORF database		15	11	0

Comparative sequencing shows that 814 of the 6,608 yeast ORFs are never expressed (dubious ORFs, http://www.yeastgenome.org). Of these only six were identified in this experiment and three were validated by SILAC-assisted *de novo* sequencing of several peptides (Supplementary Table 5 and Supplementary Figs 2–4). Two of the three validated ones were reclassified as genuine yeast genes during writing of this manuscript (YGL041W-A and YPR170W-B). This leaves three potential false-positives (0.37% of 815) and suggests that our estimate of a false-positive identification rate of maximally 1% is conservative.

2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell*.

➲ **Towards proteome wide targeted proteomics.**



2008

de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*.

➲ **SILAC based quantitation of an entire proteome.**

Pheromone signaling is required for mating of haploid cells and is absent from diploid cells.

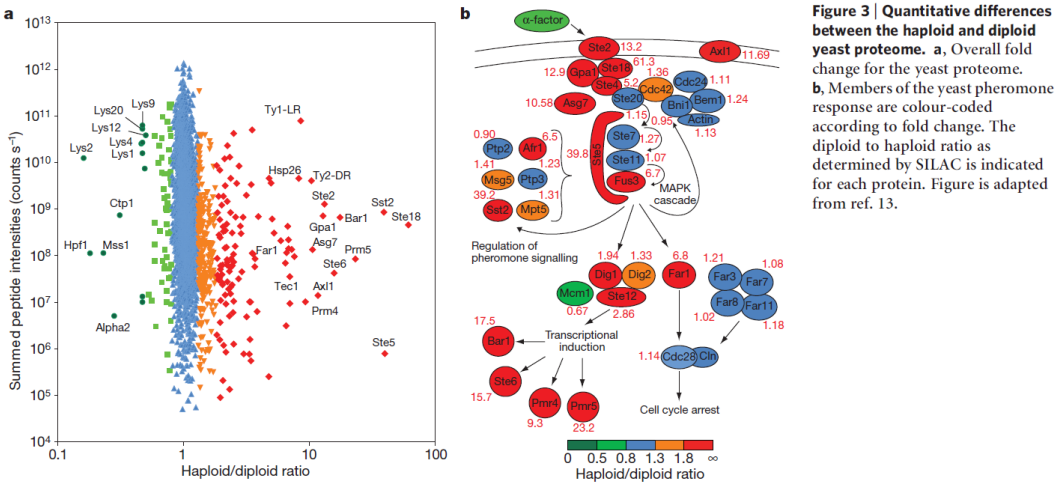


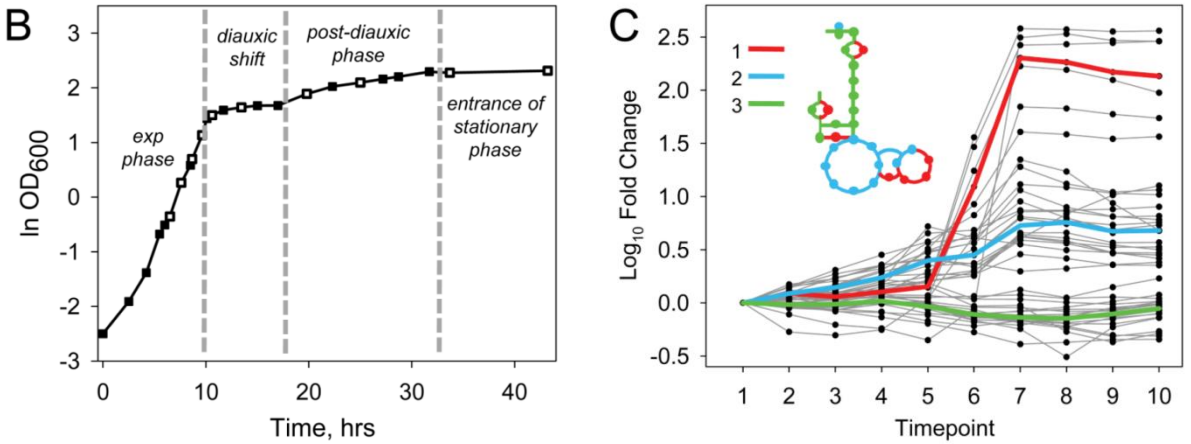
Figure 3 | Quantitative differences between the haploid and diploid yeast proteome. **a**, Overall fold change for the yeast proteome. **b**, Members of the yeast pheromone response are colour-coded according to fold change. The diploid to haploid ratio as determined by SILAC is indicated for each protein. Figure is adapted from ref. 13.

2009

Picotti P, et al. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell.

➲ **Towards proteome wide targeted proteomics.**

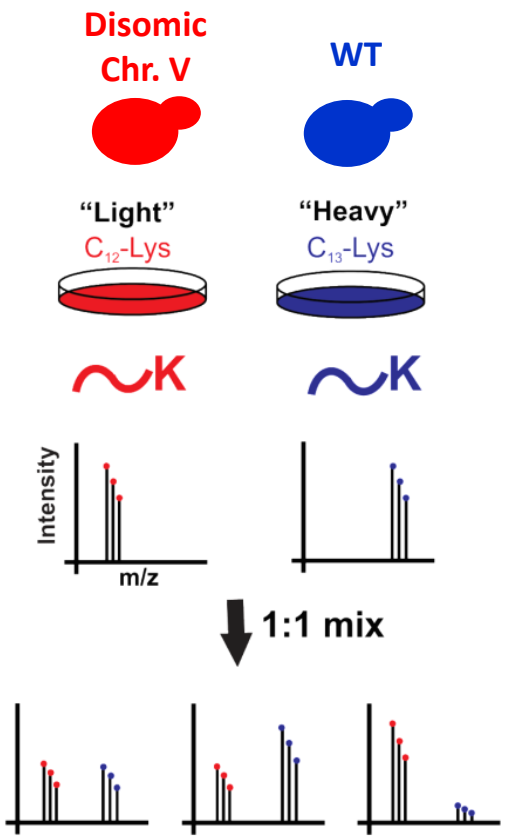
Network expression dynamics

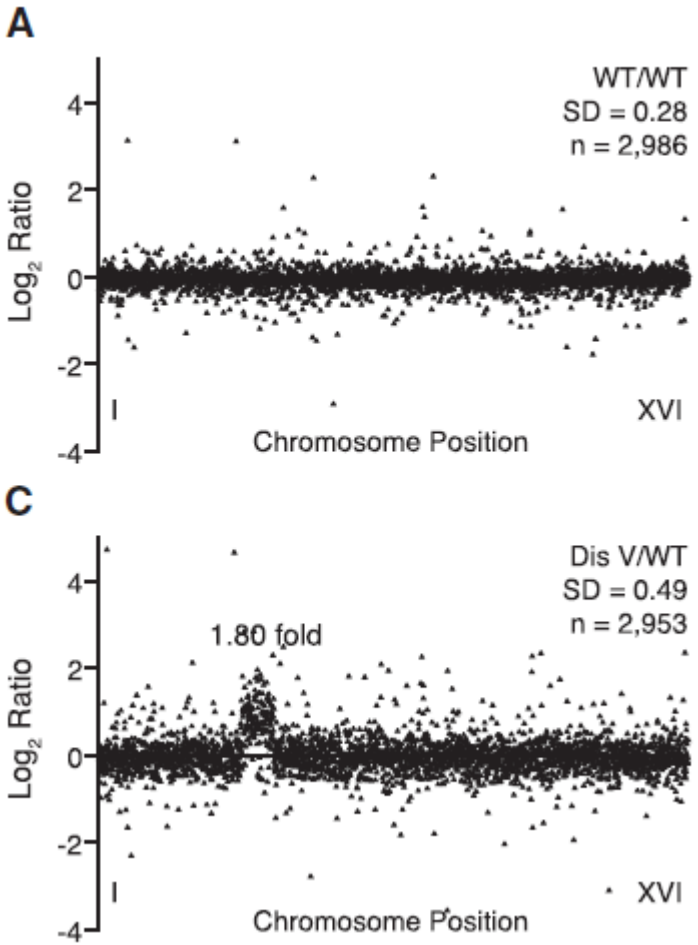


Identification of Aneuploidy-Tolerating Mutations

Cell *143*, 71–83, October 1, 2010

Eduardo M. Torres,[1,2] Noah Dephoure,[3] Amudha Panneerselvam,[1] Cheryl M. Tucker,[4] Charles A. Whittaker,[1] Steven P. Gygi,[3] Maitreya J. Dunham,[5] and Angelika Amon[1,2,*]





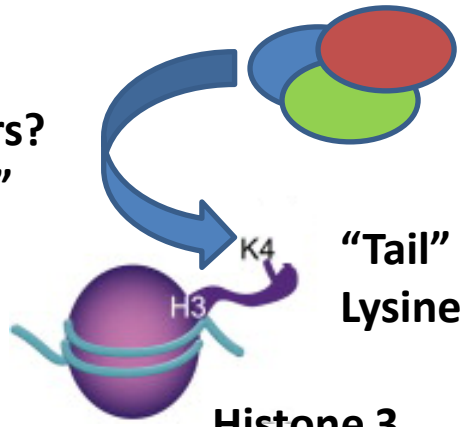
Resource

Cell

Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers

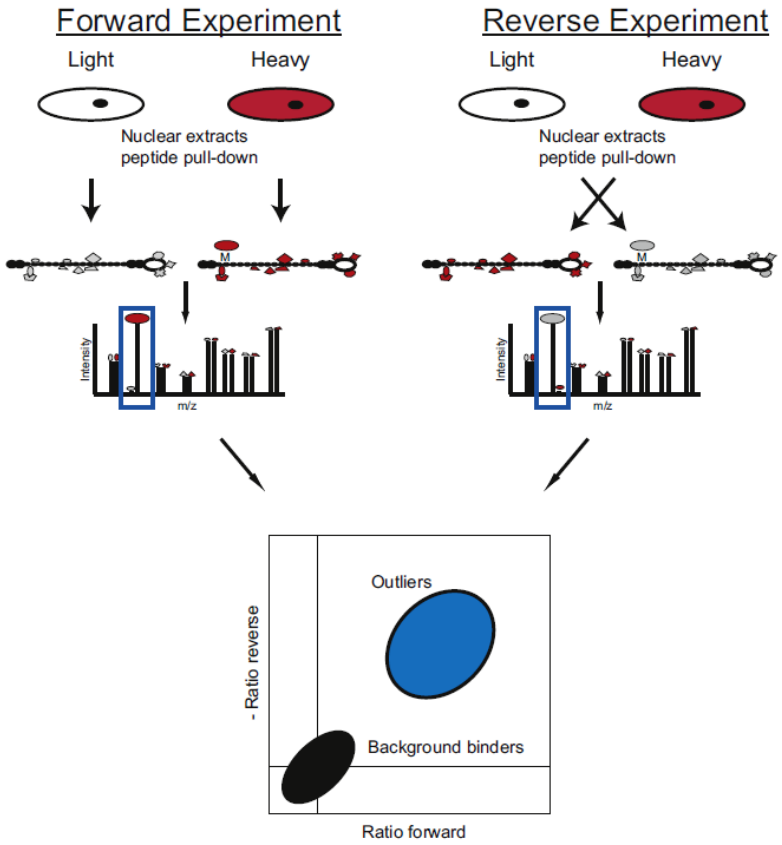
Michiel Vermeulen,[1,6,7,*] H. Christian Eberl,[1,6] Filomena Matarese,[2,6] Hendrik Marks,[2] Sergei Denissov,[2] Falk Butter,[1] Kenneth K. Lee,[3] Jesper V. Olsen,[1,5] Anthony A. Hyman,[4] Henk G. Stunnenberg,[2,*] and Matthias Mann[1,*]

Protein Regulators? "Readers"



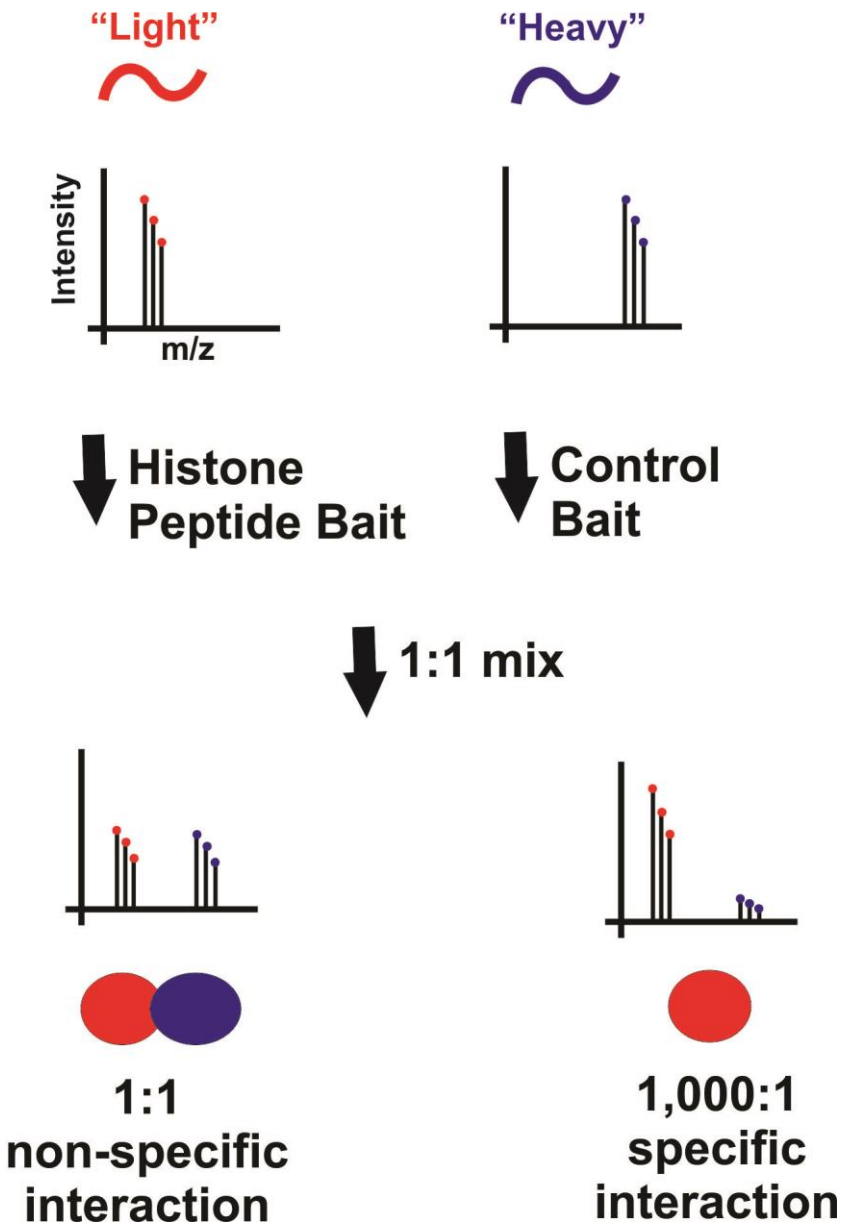
"Tail" Lysine 4

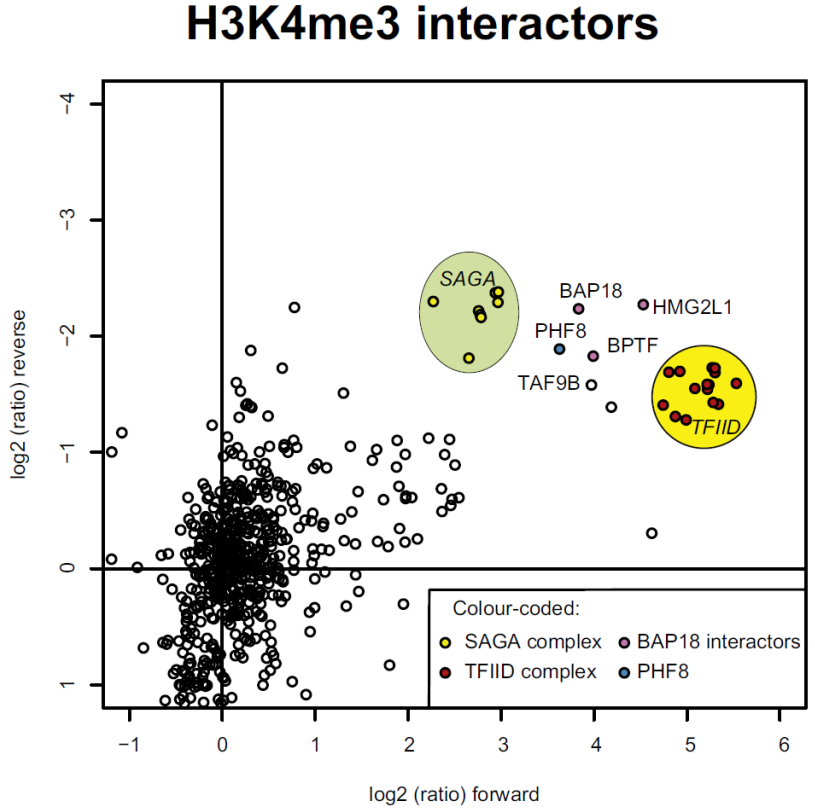
Histone 3

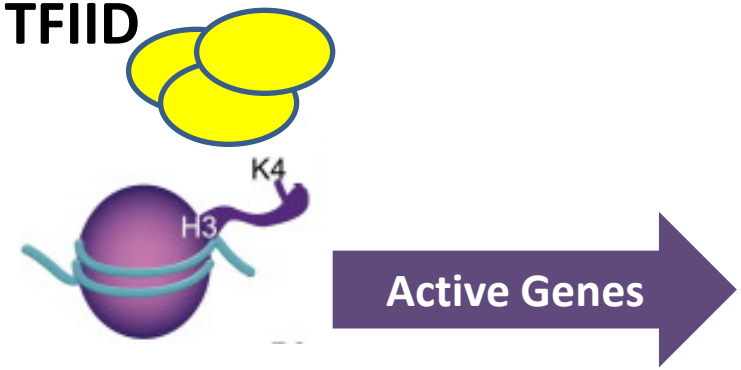


Vermeulen et al., **Cell** 2010

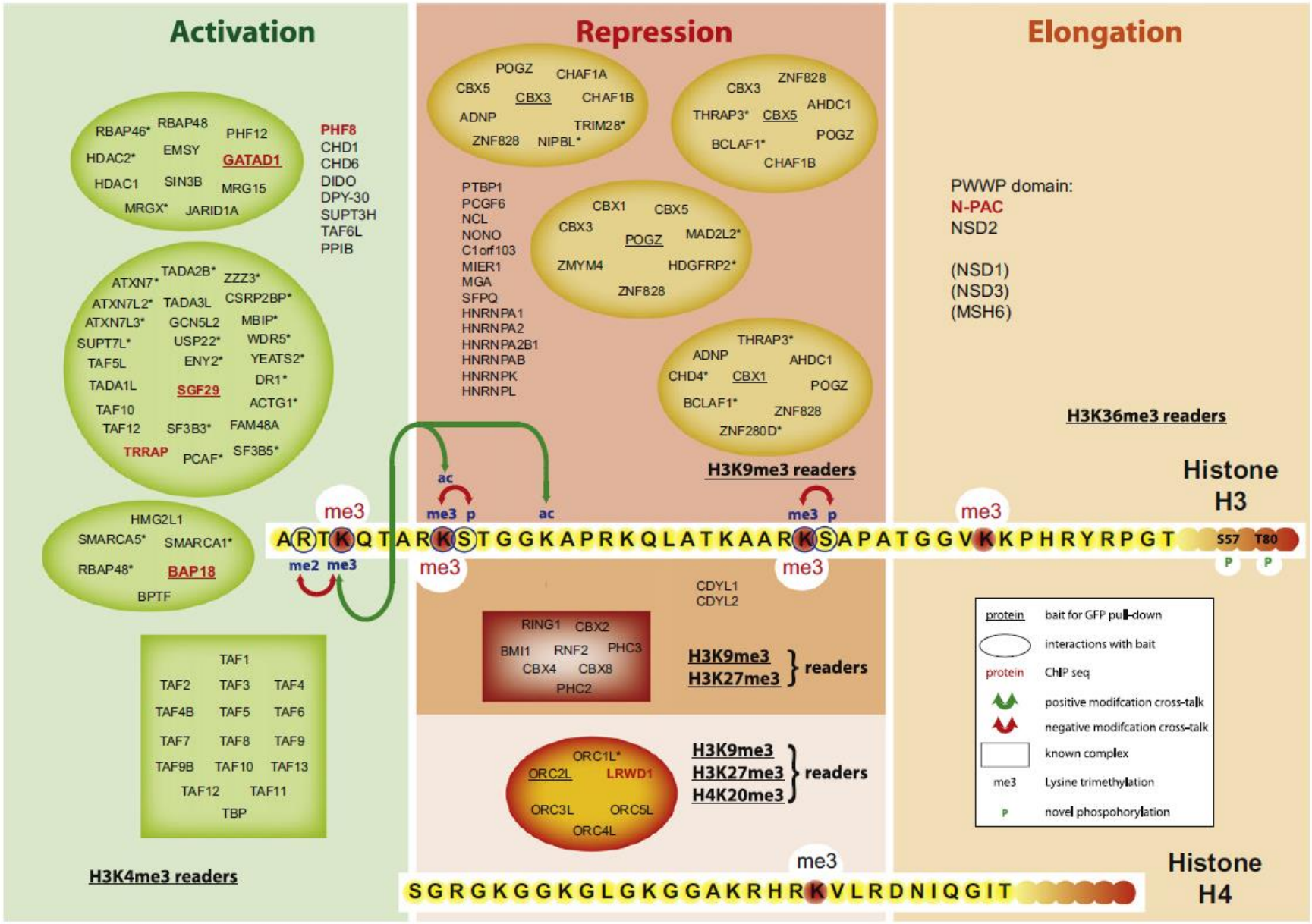
The major lysine methylation sites on the N terminus of histone H3 and histone H4 with a clearly defined biological function are H3K4me3, H3K9me3, H3K27me3, H3K36me3, and H4K20me3, which are associated with different functional states of chromatin. H3K4me3 is almost exclusively found on promoter regions of actively transcribed genes while H3K36me3 is linked to transcription elongation. H3K9me3, H3K27me3, and H4K20me3 are generally found on silent heterochromatic regions of the genome. Part of the functional distinction between these methylation sites relates to the proteins interacting with them. A number of these "chromatin readers" for various histone methyl lysine sites have already been identified and character-ized (Kouzarides, 2007; Shilatifard, 2006; Taverna et al., 2007),





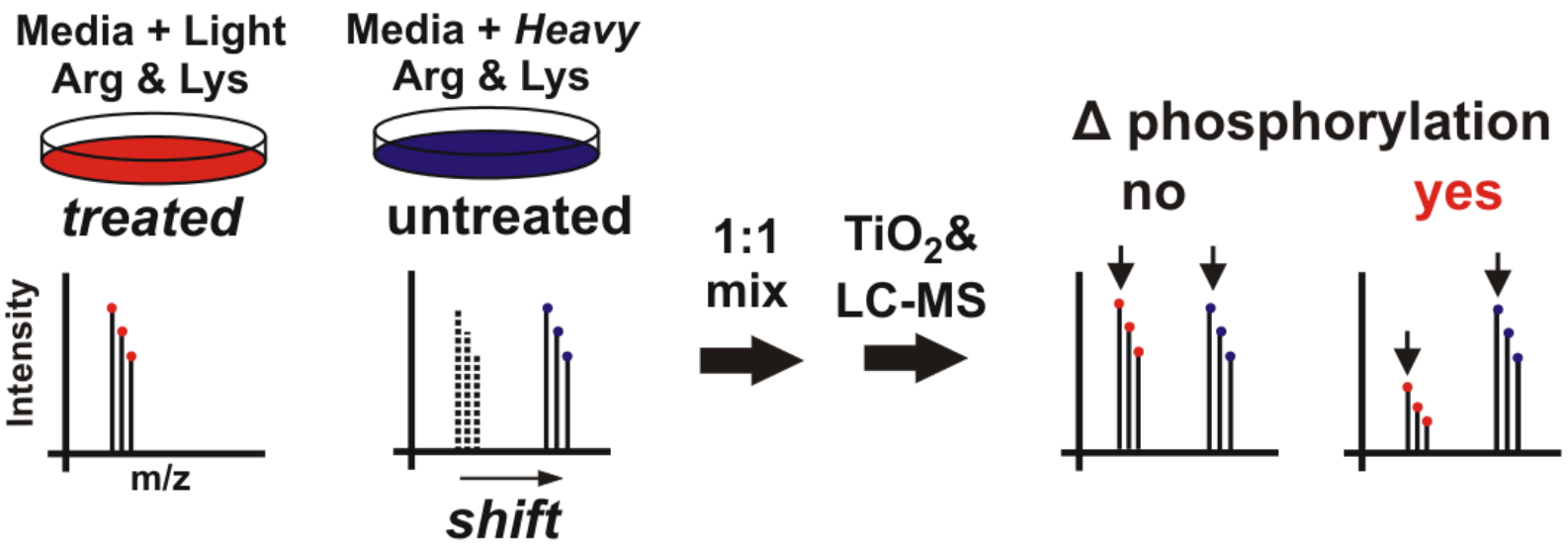


Vermeulen et al., **Cell** 2010

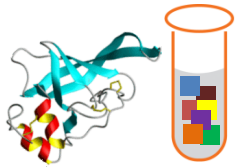


Vermeulen et al., **Cell** 2010

A SILAC approach to study protein phosphorylation dynamics



Major technological advances in mass spectrometers and phosphopeptide enrichment



Protein mixture

Digest

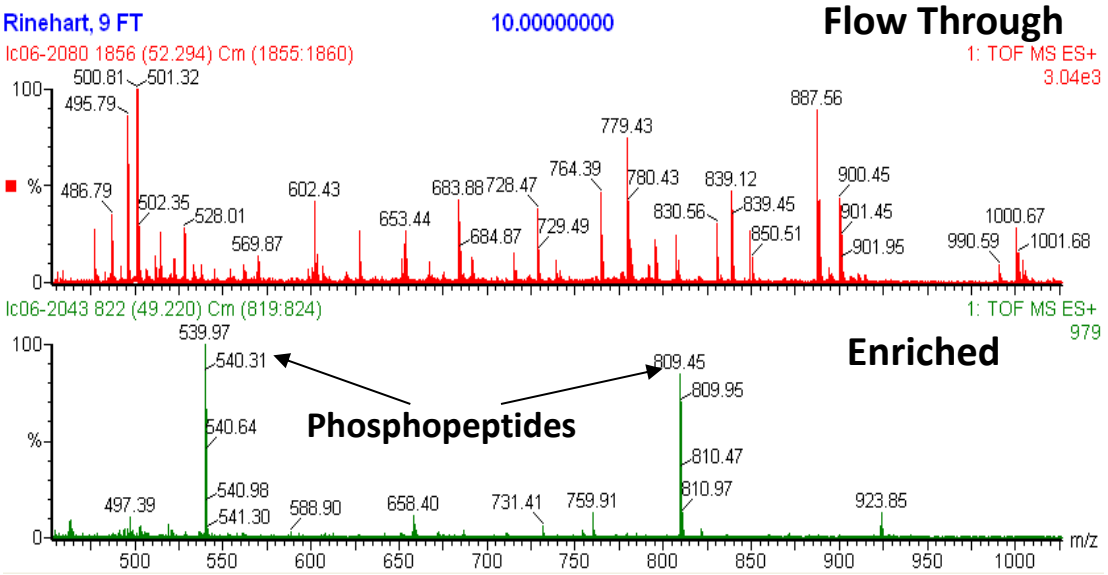
Peptides

TiO_2 Enrich

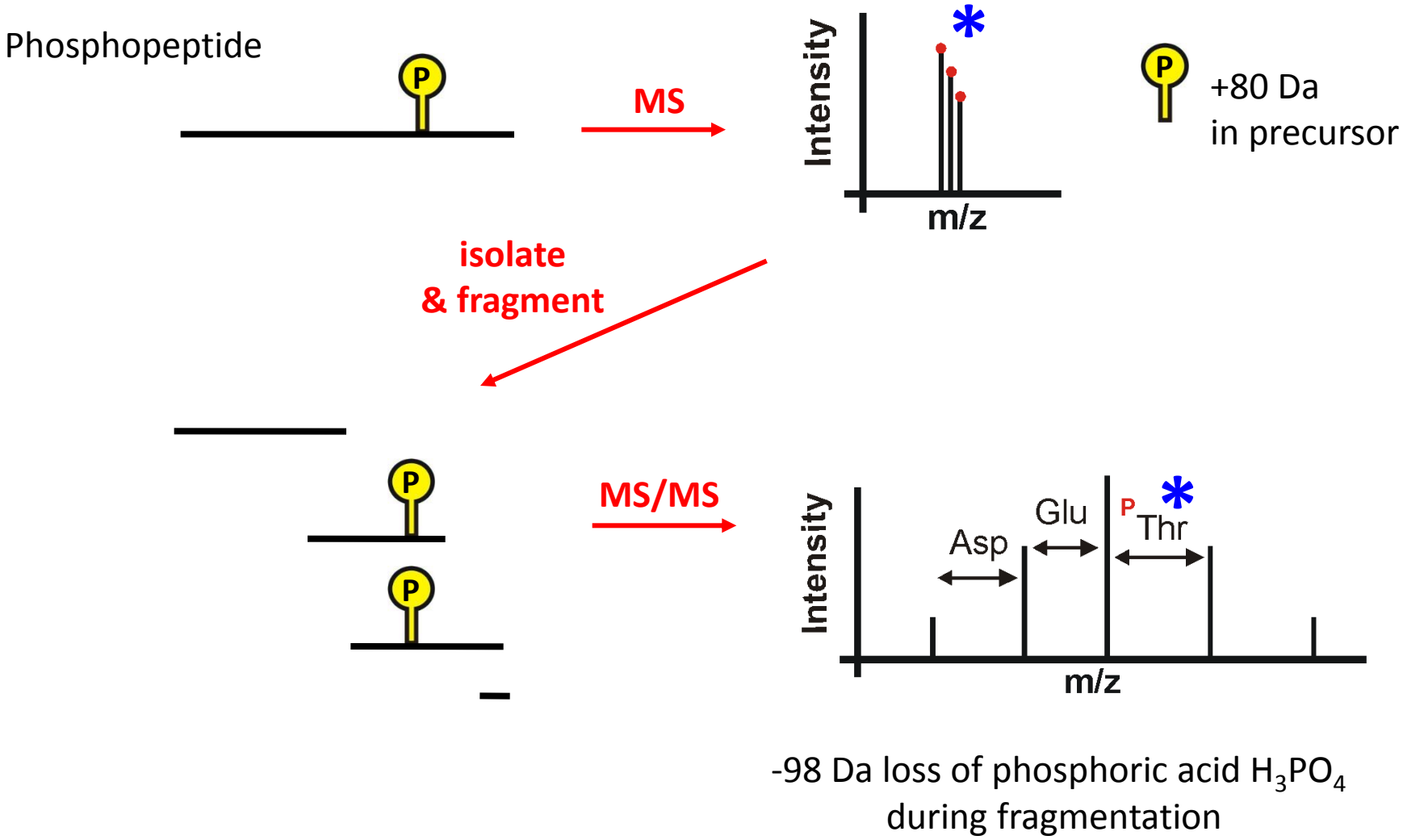
Phosphopeptides

MS

TiO_2 Enrichment

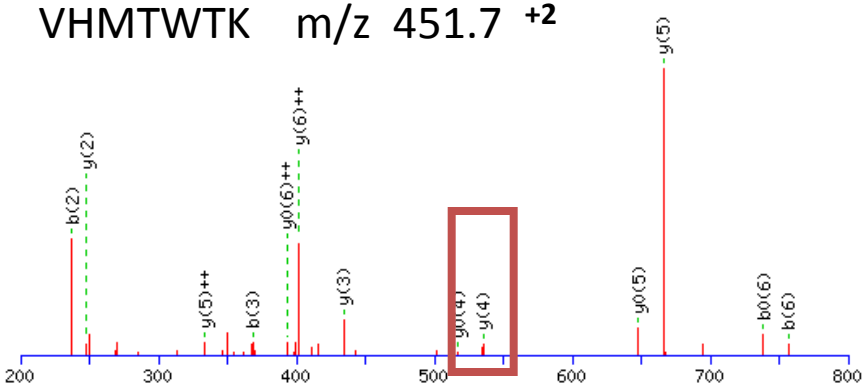


*Phosphopeptide signatures in MS



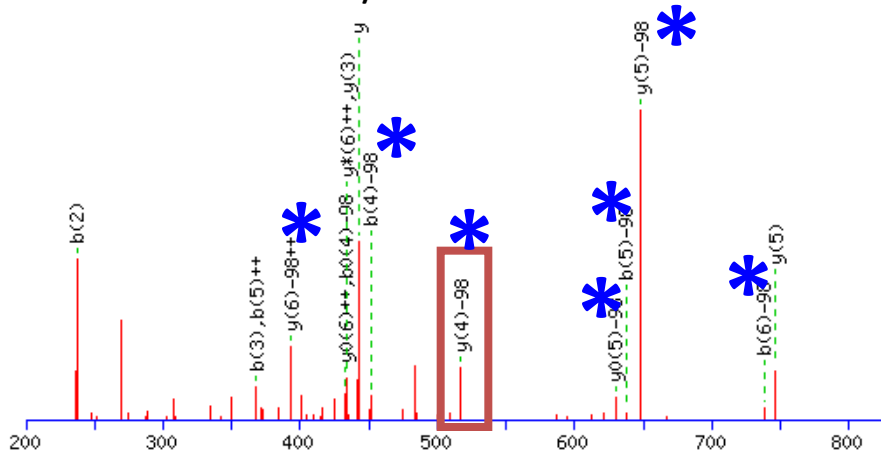
V	
H	803.3869
M	666.3280
T	535.2875
W	434.2398
T	248.1605
K	147.1128

VHMTWTK m/z 451.7 +2



V	
H	785.3763
M	648.3174
-18 T	517.2769
W	434.2398
T	248.1605
K	147.1128

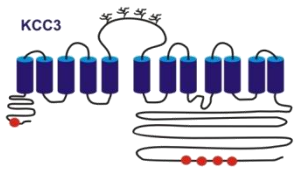
VHMT^{P}WTK m/z 491.7 +2



P +80 Da in precursor

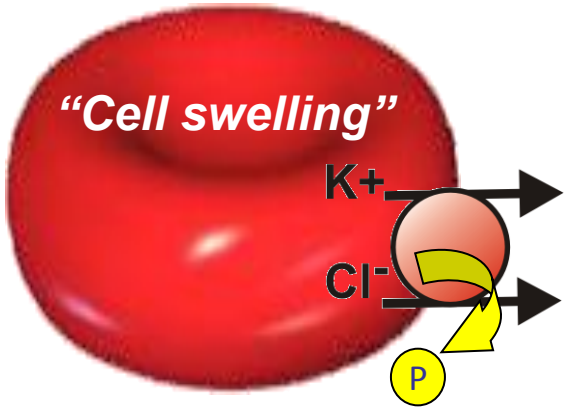
* - H_3PO_4, 98 Da

(Threonine changes to 2-aminodehydrobutyric acid, -18 Da)



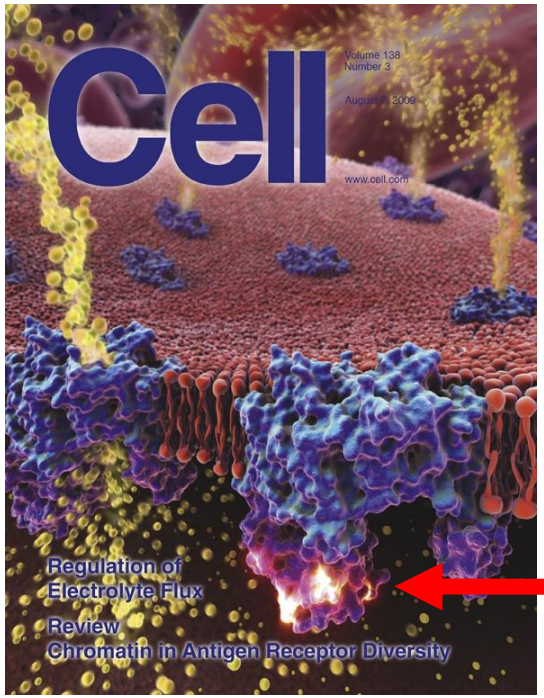
Sites of Regulated Phosphorylation that Control K-Cl Cotransporter Activity

Jesse Rinehart,[1,5] Yelena D. Maksimova,[2] Jessica E. Tanis,[3] Kathryn L. Stone,[5,6] Caleb A. Hodson,[1] Junhui Zhang,[1] Mary Risinger,[7] Weijun Pan,[4] Dianqing Wu,[4] Christopher M. Colangelo,[5,6] Biff Forbush,[3] Clinton H. Joiner,[7] Erol E. Gulcicek,[5,6] Patrick G. Gallagher,[2] and Richard P. Lifton[1,5,*]

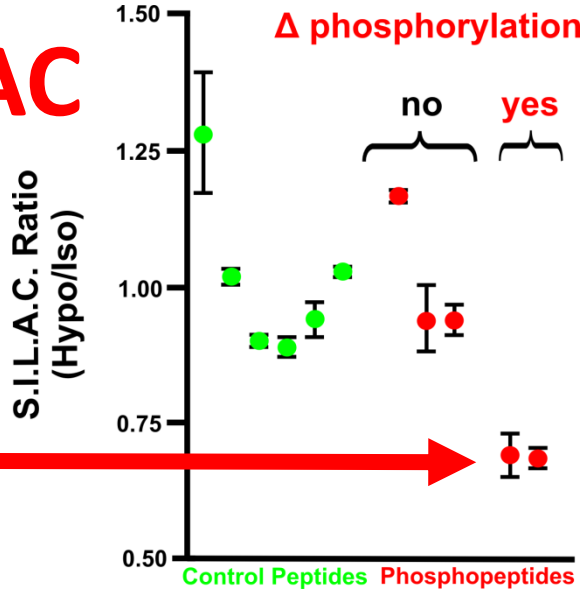


ON!

De-phosphorylation *activates* KCC3



SILAC

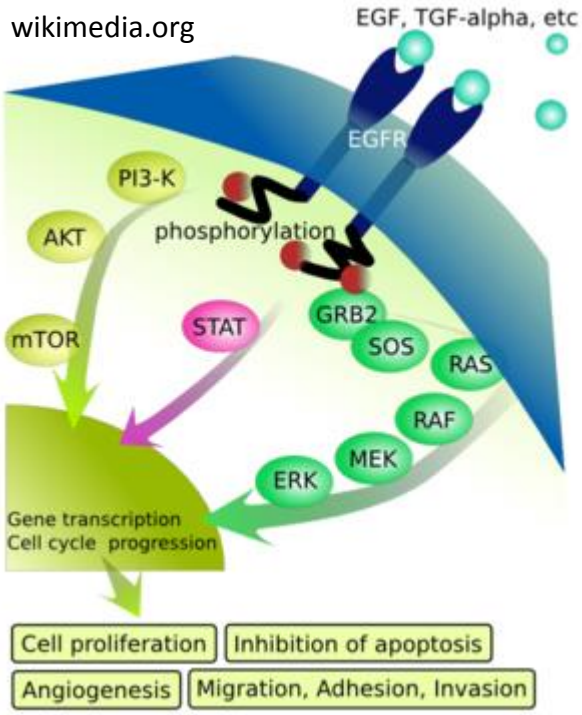


Adapted from Rinehart, et al. *Cell,* 2009

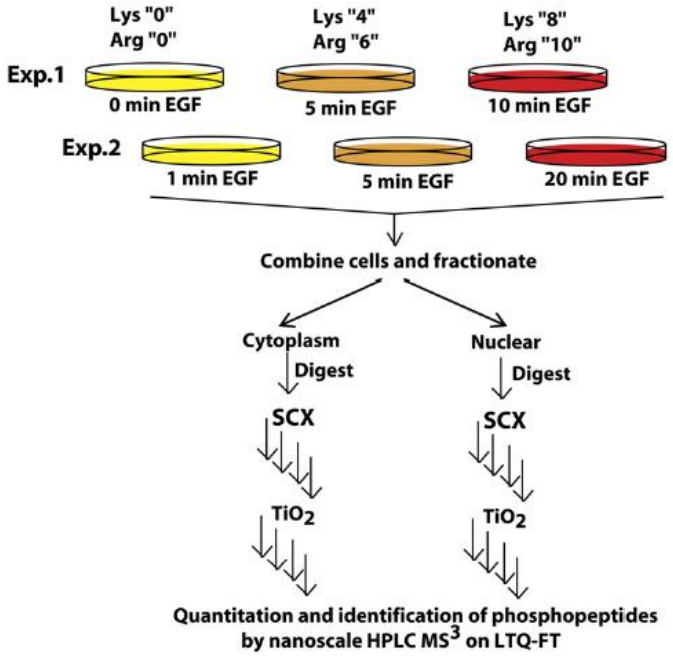
Quantitative Proteomics Reveals Dynamics in Signaling Networks

Phosphorylation dynamics
after EGF stimulation

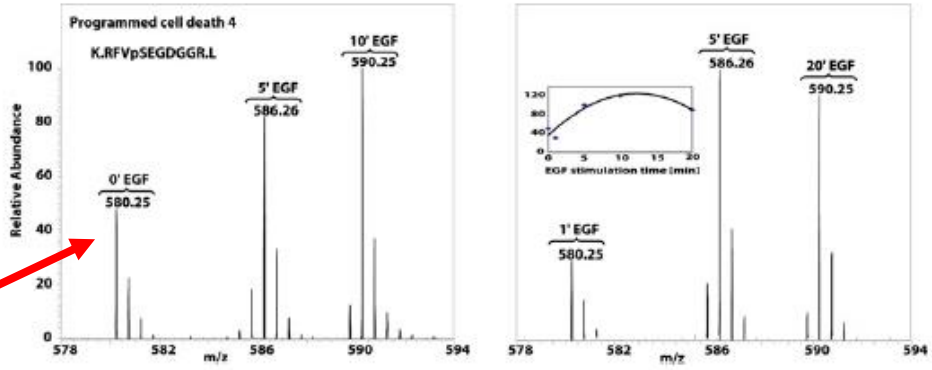
wikimedia.org



SILAC approach enables dynamic analysis



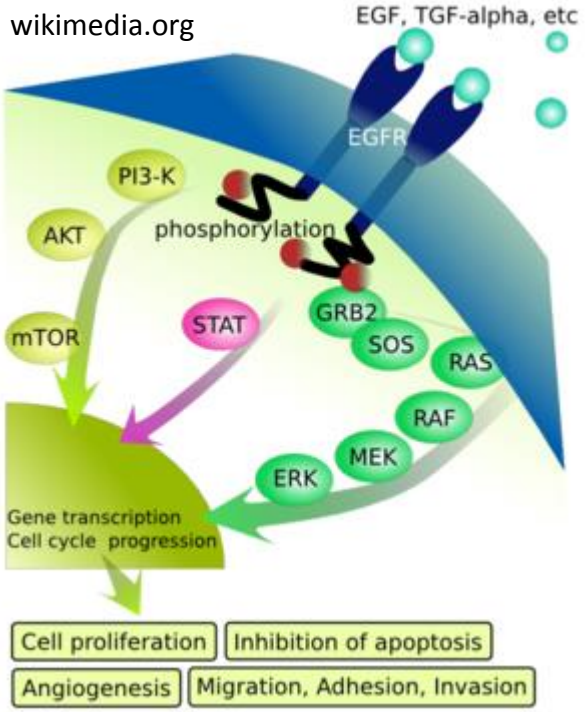
MS spectra triplets

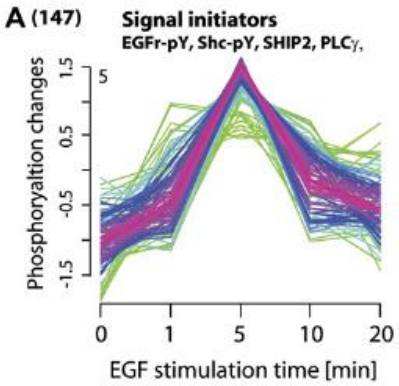


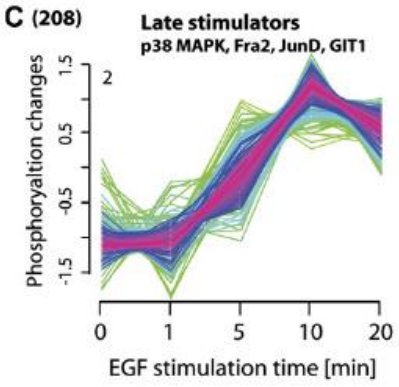
Olsen, et al. *Cell*, 2006

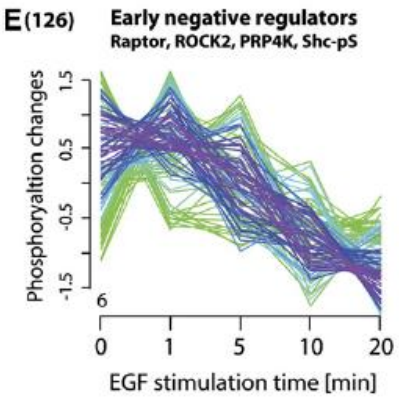
Phosphorylation dynamics after EGF stimulation

wikimedia.org









Olsen, et al. *Cell*, 2006

Proteins: Proteomics & Protein-Protein Interactions

Overview

- **Techniques & Technologies**
  - Mass Spectrometry
  - Protein-Protein Interactions
    - Genetic & Biochemical Strategies
    - Protein Purification
  - Quantitative Proteomics
- **Applications**
  - Representative Studies
- **Putting it all together….**
  - Databases & Pathways