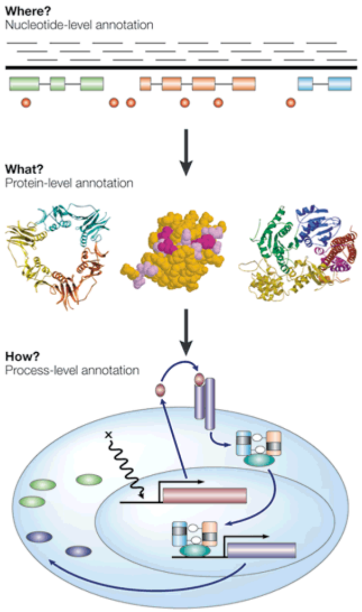
Dónde? Qué? Cómo?



Nature Reviews | Genetics

Proteomics studies

- Activities (function)
- Expression
- Post-translational modifications (PTM)
- Localization
- Complex formation

Protein Function

Two “broad”types of protein function

- “Catalytic”proteins, including enzymes, transporters, chaperons, etc.

- “Structural”proteins

Compared with Genomics

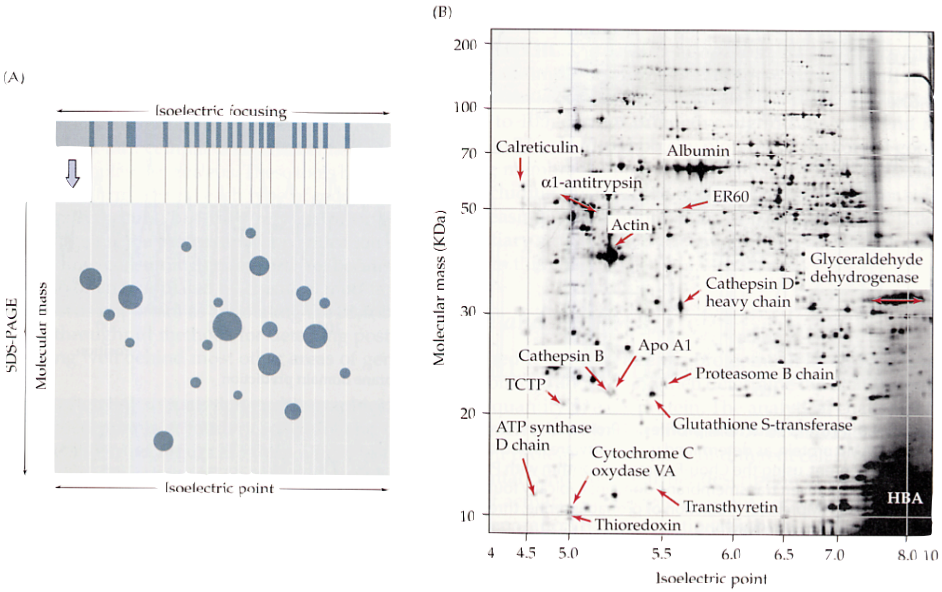
Similarities

- Static picture of dynamic processes
- High throughput analysis
- Technology-driven
- Computation intensive

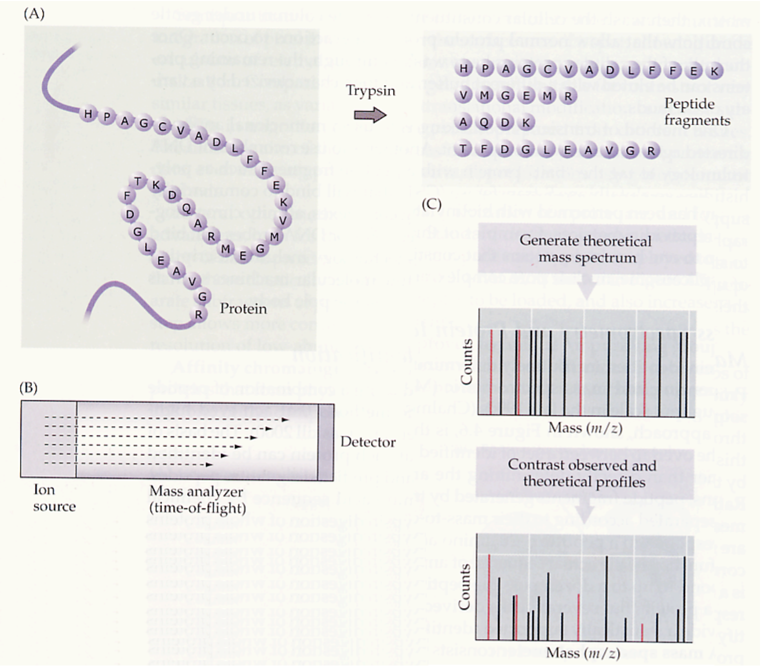
Differences

- Proteomics: “closer”to activity (function: PTM, location, turnover, protein complex, enzyme function)
- Protein dynamics and RNA dynamics do not always correlate

Two-Dimensional PAGE



Mass spectometry sequencing



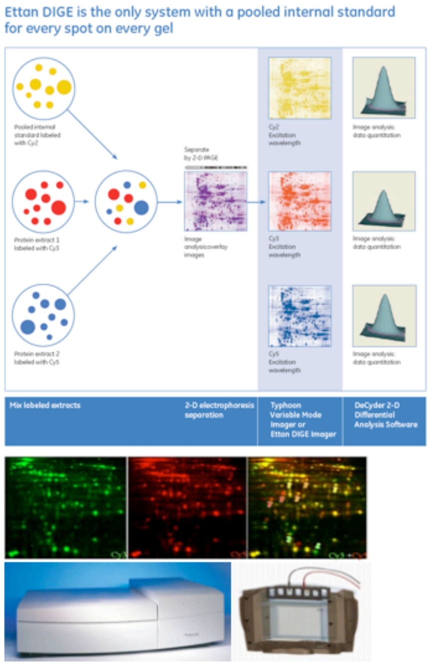
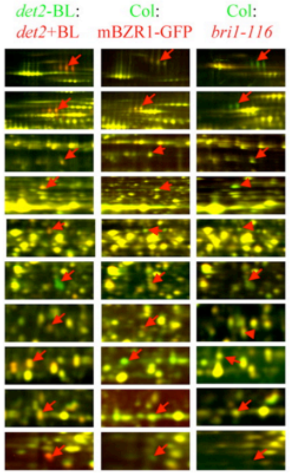


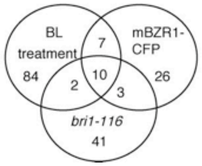
Fig 1. A. Principle of 2D-DIGE technology - (Ettan DIGE - GE web site)
B Fluorescent Images scanned with Typhoon (GE) and Ettan Dalt system (GE)

A

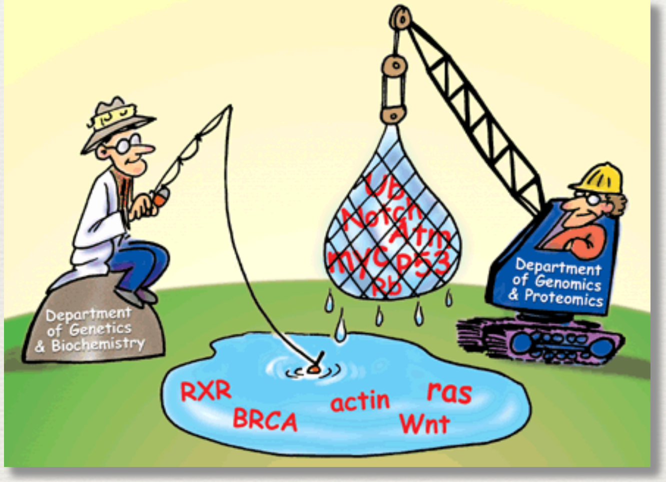


Spot Number	Gene Locus	Protein Identity	*bzr1-1D*/Col ratio	*bzr1-1D*/Col p value	*bri1-116*/Col ratio	*bri1-116*/Col p value
317	At1g22530	PATL2	1.49	5.1E-05	-2.59	8.5E-06
745	At1g72150	PATL1	1.51	2.0E-03	-2.03	2.0E-03
953	At3g46970	PHS2	1.12	3.9E-02	-1.29	1.4E-05
1433	At5g11670	NADP-ME2	-1.32	5.0E-02	-1.46	3.5E-05
1941	At3g09940	MDAR3	-1.44	1.9E-02	-1.27	3.1E-04
2898	At5g54770	THI1	1.25	4.1E-02	-1.48	1.7E-02
3034	At1g10370	GSTU17	1.31	4.4E-03	1.10	2.8E-01
3279	At4g02520	GSTF2	-1.82	9.2E-03	-1.05	6.0E-01
3286	At2g30860	GSTF9	-1.28	2.3E-03	1.05	5.6E-01
3299	At2g41090	CABP-22	-1.33	2.3E-02	1.01	6.9E-01

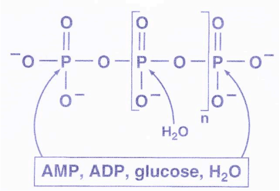
B



Systems Biology
From genes to function and back



¿Qué perturbamos?

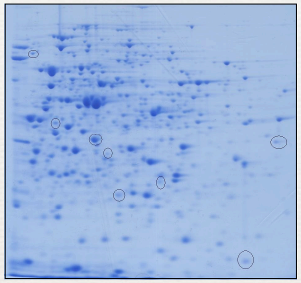


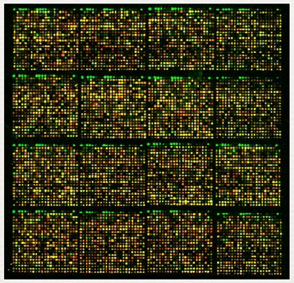
$$\begin{tikzcd}
& \mathrm{GTP} \arrow[d, shift right] & \\
\mathrm{ATP} \arrow[r, shift left, "\mathrm{PPK1}"] & \mathrm{PoliP} \arrow[l, shift left, "\text{Polifosfato kinasa}"] \arrow[u, shift right, "\mathrm{PPK2}"'] \arrow[r, "\mathrm{PPX}", "\text{Exopolifosfatasa}"'] & \mathrm{Pi}
\end{tikzcd}$$

- Reserva de fosfato
- Sustituto del ATP
- Homeostasis a metales
- Regulación pH intracelular
- Patogenicidad en diferentes bacterias como *H. Pylori*, *P. aeruginosa*, *V. Cholera*, *Salmonella* y *Shiguella*.
- Comunicación bacteriana, motilidad y formación de biopelículas.

¿Qué monitoreamos?

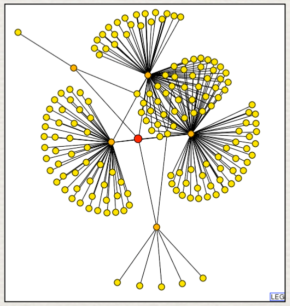
Proteómica

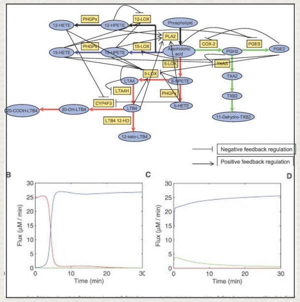




Genómica funcional

Bioinformática





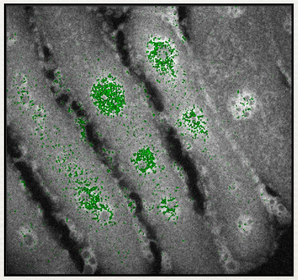
Modelamiento matemático

¿Cómo integramos?

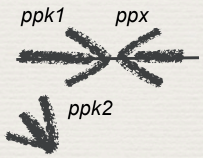
¿A quién perturbamos?

Escherichia coli
K12

Pseudomonas aeruginosa
PAO1

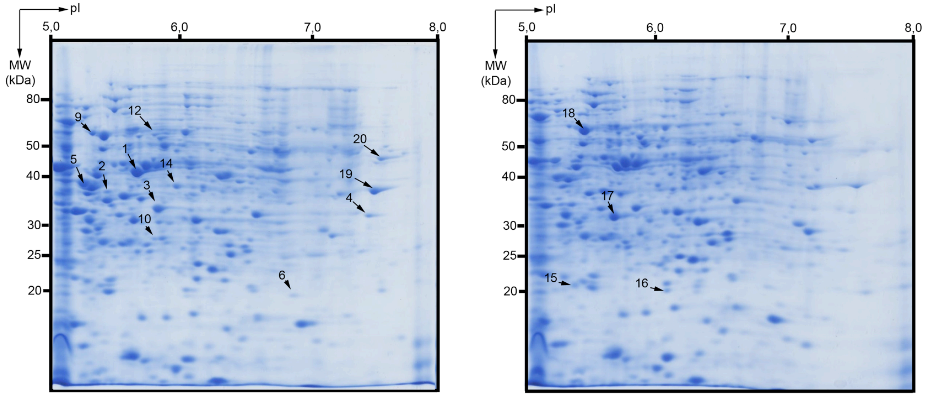


Pseudomonas sp.
B4



?

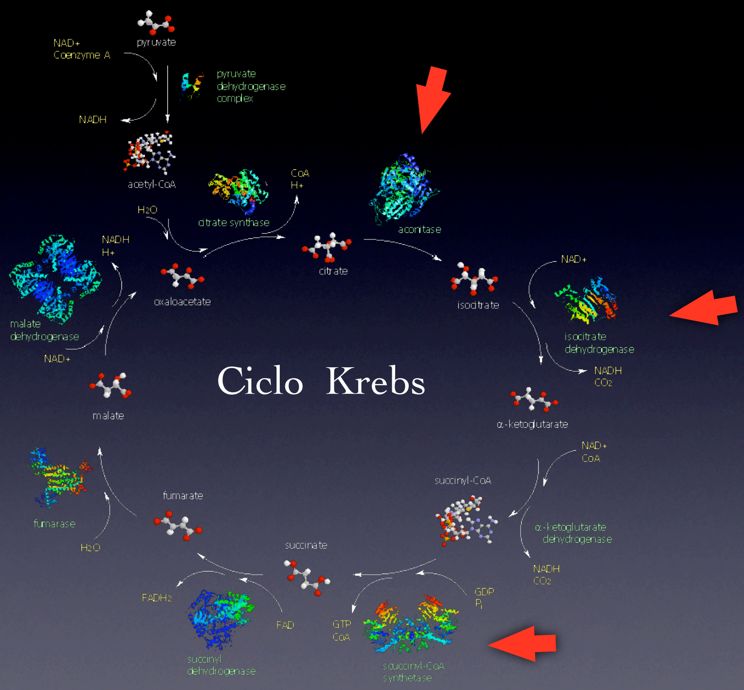
Classic Proteomics
Pseudomonas sp. B4

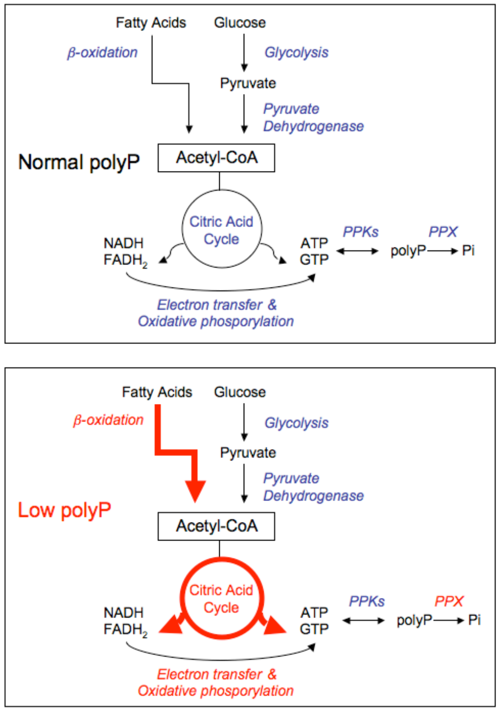


Ps. B4 pMLS7

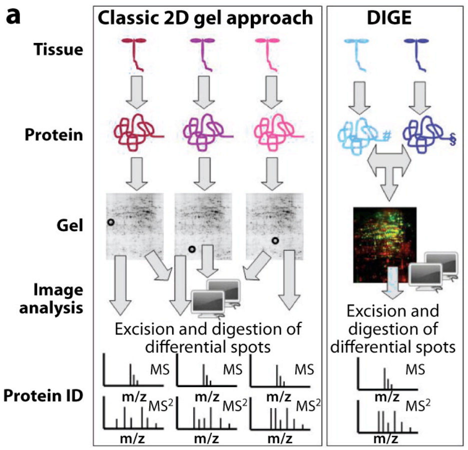
Ps. B4 pS7PPX1

Ciclo Krebs


pyruvate
NAD+
Coenzyme A
pyruvate
dehydrogenase
complex
NADH
acetyl-CoA
citrate synthase
CoA
H+
H2O
aconitase
citrate
oxaloacetate
NADH
H+
malate
dehydrogenase
NAD+
isocitrate
NAD+
isocitrate
dehydrogenase
NADH
CO2
α-ketoglutarate
malate
NAD+
CoA
succinyl-CoA
α-ketoglutarate
dehydrogenase
NADH
CO2
fumarase
H2O
fumarate
succinate
GDP
Pi
FADH2
FAD
GTP
CoA
succinyl
dehydrogenase
succinyl-CoA
synthetase


Normal polyP
Fatty Acids
Glucose
β-oxidation
Glycolysis
Pyruvate
Pyruvate Dehydrogenase
Acetyl-CoA
Citric Acid Cycle
NADH
FADH2
ATP
GTP
PPKs
PPX
polyP
Pi
Electron transfer & Oxidative phosporylation
Low polyP
Fatty Acids
Glucose
β-oxidation
Glycolysis
Pyruvate
Pyruvate Dehydrogenase
Acetyl-CoA
Citric Acid Cycle
NADH
FADH2
ATP
GTP
PPKs
PPX
polyP
Pi
Electron transfer & Oxidative phosporylation

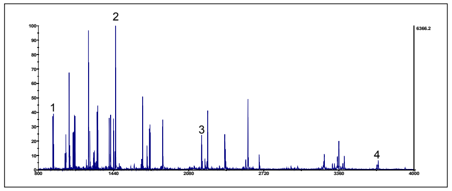
Classical vs Q- proteomics



Schulze WX, Usadel B. 2010.
Annu. Rev. Plant. Biol. 61:491–516

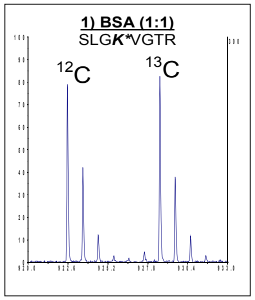
ICPL- Isotope-Coded Protein labelling

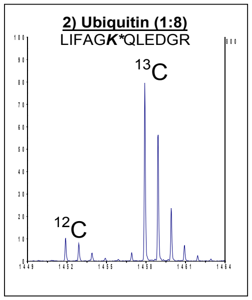
MS-spectrum of the tryptic digest

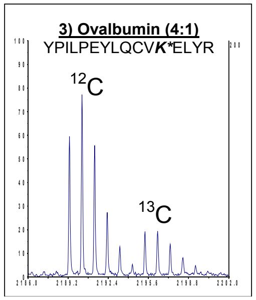


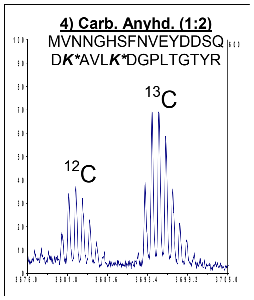
Protein	Protein-Mix A	Protein-Mix B	Relative Amount
BSA (bovine)	13.5 µg	13.5 µg	**1:1**
Ovalbumin (chicken)	24 µg	6 µg	**4:1**
Carbonic Anhydrase II (bovine)	12 µg	24 µg	**1:2**
Ubiquitin (bovine)	0.5 µg	4 µg	**1:8**

Four peptide pairs as an example for an expanded view:









Summary of peptide analysis Q-proteomics

wt v/s *ppk-, ppx-*

- Total number of peptides: 3.113
- Total number of quantified peptides:1.794
- Total number of proteins: 313
- Total number of differencied proteins: 53

wt v/s *ppx-*

- Total number of peptides: 2.457
- Total number of quantified peptides: 1.558
- Total number of proteins: 270
- Total number of differencied proteins: 59

wt v/s *ppk1-*

- Total number of peptides: ????
- Total number of quantified peptides: ????
- Total number of proteins: ????
- Total number of differencied proteins: ????

Q-Proteomics (wt vs *ppx*-)

Up Regulated proteins

Protein Name	Score	SC [%]	# Pept.	# (L/H)	Avg.	log2(Avg-median)
1-pyrroline dehydrogenase	78,54	6,33	3	2	3,36	1,748461233
Glutamate decarboxylase	783,33	24,89	29	18	13,34	3,737686761
Ribosomal protein L32	56,20	33,33	2	2	5,38	2,427606173
Selenium metabolism protein SsnA	51,39	3,62	2	2	4,63	2,211012193
Heat shock protein HslVU, ATPase subunit HslU	39,62	9,93	4	2	3,77	1,914564523

Down Regulated proteins:

Protein Name	Score	SC [%]	# Pept.	# (L/H)	Avg.	log2(Avg-median)
Malate synthase G	475,16	19,23	18	13	0,2	-4,070389328
Delta-1-pyrroline-5-carboxylate dehydrogenase	770,25	20,00	37	27	0,47	-2,837728571
M20/DapE family protein YgeY	88,20	7,69	5	4	0,39	-3,106915204
Aconitate hydratase 2	1013,54	30,87	44	24	0,4	-3,070389328
Transcriptional regulator, AsnC family	77,07	18,90	3	3	0,76	-2,144389909

Q-Proteomics (wt vs *polyP*-)

Up Regulated proteins:

Protein	Score	SC [%]	# Pept.	# (L/H)	Avg.	log2(WT/*polyP*-) median
(P)ppGpp synthetase I, SpoT/RelA	42,80	2,42	4	3	3,39	1,913288367
ATP-dependent protease La	120,56	8,16	6	4	1,83	1,023846742
Phosphoenolpyruvate synthase	133,15	8,46	6	4	1,78	0,983880335
ATP synthase subunit alpha	334,22	23,59	19	12	1,48	0,717600269

Down Regulated proteins:

Protein	Score	SC [%]	# Pept.	# (L/H)	Avg.	log2(WT/*polyP*-) median
Arginine decarboxylase	94,0833333	6,38297872	6	5	0,52	-0,791413378
Malate dehydrogenase (Oxaloacetate-decarboxylating) (NADP(+))., Phosphate acetyltransferase	57,84	3,55731225	4	3	0,52	-0,791413378
6-phosphogluconolactonase	49,41	11,48	5	4	0,22	-2,032421478

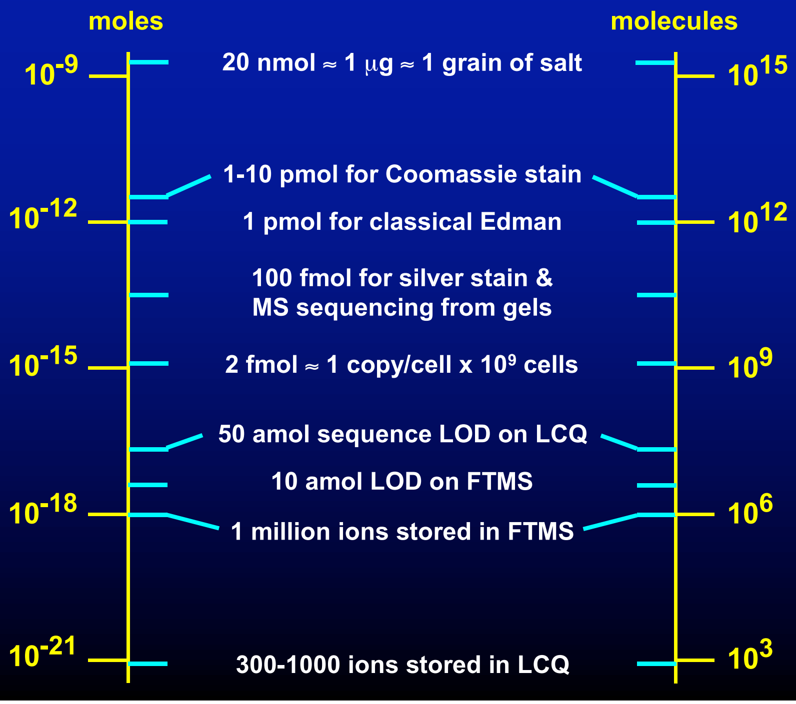
Q-Proteomics vs microarrays (ppk1-,*ppx*-)

Protein (wt vs ppx-)	**Q-Proteomics (Change)**	**Gene**	**Microarray (Change)**	
Malate synthase G	-3,03552	glcB	-1.896.972	-
Citrate synthase	-2,40325	gltA	-2.279.975	-
PTS system, galactitol-specific IIC subunit	-2,40325	treB	-3.890.282	-
Succinate dehydrogenase and fumarate reductase iron-sulfur protein	-2,07204	frdB	-1.987.034	-
FAD dependent oxidoreductase	-2,07204	ydiS	-1.656.145	-
Hydro-lyase, Fe-S type, tartrate/ fumarate subfamily, beta subunit	-1,93118	adhE	-2.208.965	-
2-oxoglutarate dehydrogenase, E2 subunit, dihydrolipoamide succinyltransferase	-1,60256	aceF	-2.754.167	-
Phosphoenolpyruvate carboxykinase (ATP)	-1,40325	pckA	-1.589.001	-
Periplasmic binding protein/LacI transcriptional regulator	-0,99989	ycjW	-1.888.503	-
Formate dehydrogenase, alpha subunit	-0,81829	hyfd	-1.735.920	-
Endoribonuclease L-PSP	2,25036	yhaR	5.158.855	+
Ribosomal protein L32	2,42761	rpmF	1.914.481	+

Q-Proteomics vs microarrays (*ppx*-)

Protein	**log2(WT/ polip(-))-median**	**Gene**	**Zscore**	**Regulation**
Arginine decarboxylase	-0,79141	adiA	-1,59218	**-**
Malate dehydrogenase (Oxaloacetate-decarboxylating) (NADP(+))., Phosphate acetyltransferase	-0,79141	mdh	-2,37323	**-**
Efflux transporter, RND family, MFP subunit	-1,13430	yhiI	-2,40360	**-**
6-phosphogluconolactonase	-2,03242	yieK	-3,29833	**-**
(P)ppGpp synthetase I, SpoT/ RelA	1,91329	relA	1,63127	**+**
ATP-dependent protease La	1,02385	lon	2,42645	**+**
3-oxoacyl-(Acyl-carrier-protein) reductase	0,71760	fabG	2,27737	**+**
L-asparaginase, type II	0,66802	ansB	2,42421	**+**

Numbers to Remember (1 mole ≈ 10^{24} molecules)



moles	molecules
10^{-9}	10^{15}
10^{-12}	10^{12}
10^{-15}	10^{9}
10^{-18}	10^{6}
10^{-21}	10^{3}

20 nmol ≈ 1 μg ≈ 1 grain of salt

1-10 pmol for Coomassie stain

1 pmol for classical Edman

100 fmol for silver stain &
MS sequencing from gels

2 fmol ≈ 1 copy/cell x 10^9 cells

50 amol sequence LOD on LCQ

10 amol LOD on FTMS

1 million ions stored in FTMS

300-1000 ions stored in LCQ

Gene Insertion

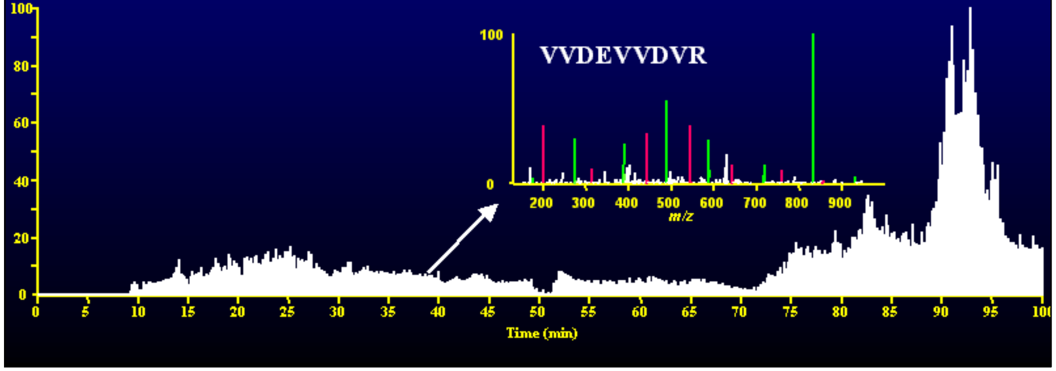
2 mL E. coli ≈ 2e9 cells
- 1300 proteins/cell

+

100 fmol prnA
- 30 copies/cell
- 61 kDa, 481 AA

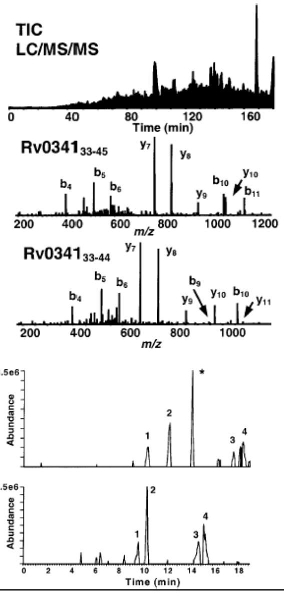
load 1/30 sample
100 min DD LCQ run
3000 MS/MS spectra

Automatic SEQUEST
analysis of spectra
5 hits for prnA target
11% coverage by AA

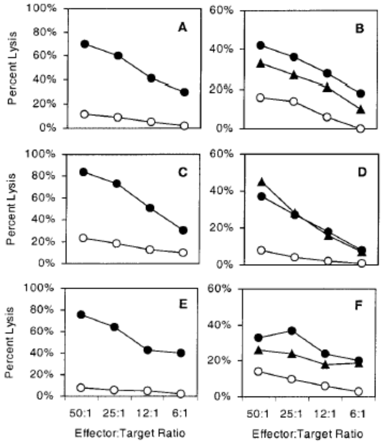


INFECTION AND IMMUNITY, June 2002, p. 2926–2932 Vol. 70, No. 6
0019-9567/02/$04.00+0 DOI: 10.1128/IAI.70.6.2926–2932.2002

Identification by Mass Spectrometry of $CD8^+$-T-Cell *Mycobacterium tuberculosis* Epitopes within the Rv0341 Gene Product

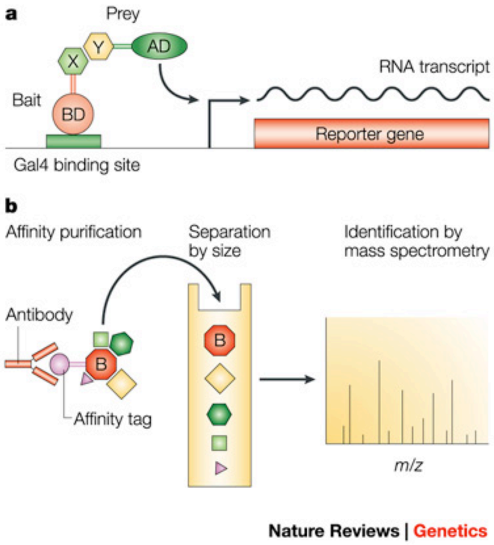


Rv0341 peptides found in purified MHC Class I

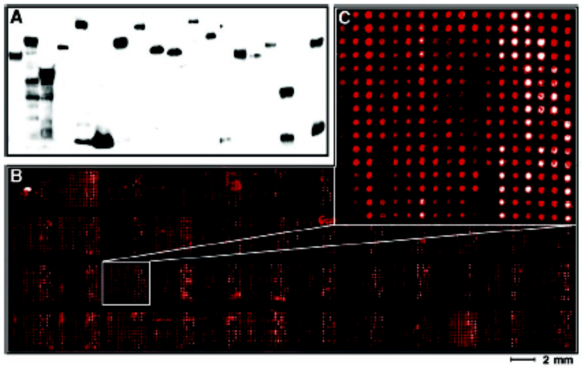


Rv0341 peptides recognize and lyses *M. tuberculosis*-infected dendritic cells.

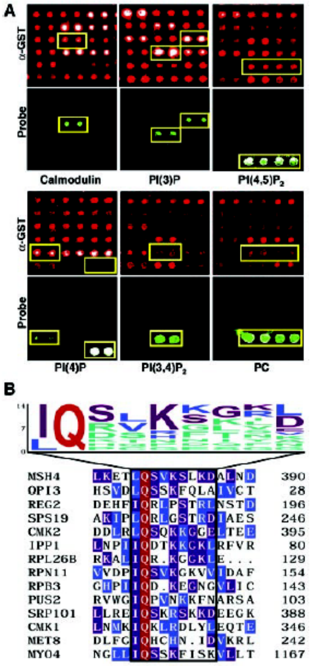
Systematic Proteomics



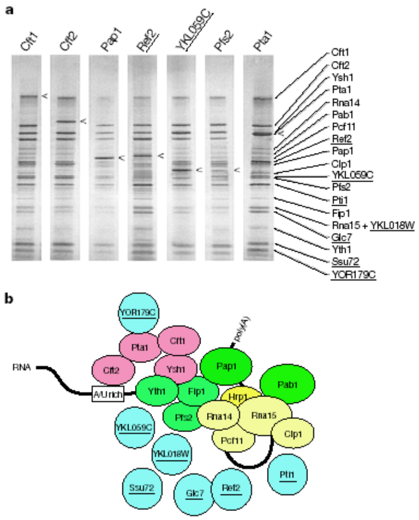
Yeast Protein Chip
(*In vitro*)



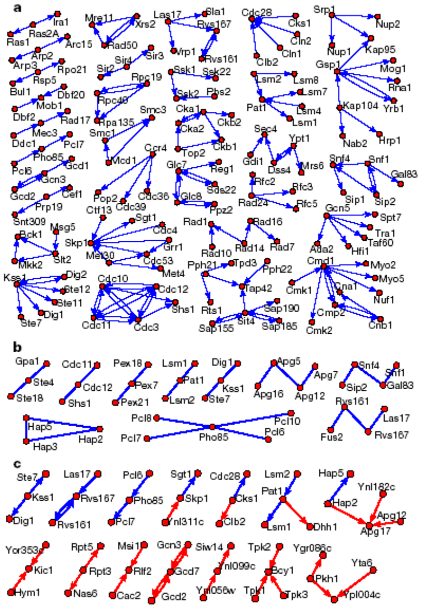
www.sciencemag.org SCIENCE VOL 293 14 SEPTEMBER 2001



Protein-Protein Interaction (*In vivo*)

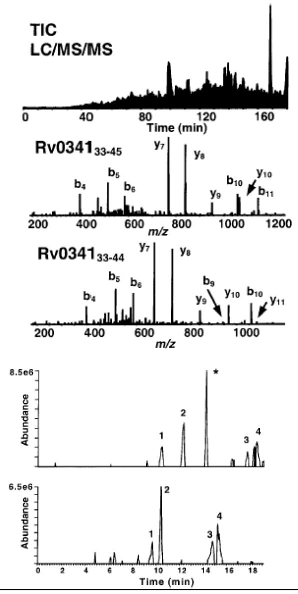


NATURE | VOL 415 | 10 JANUARY 2002 | www.nature.com

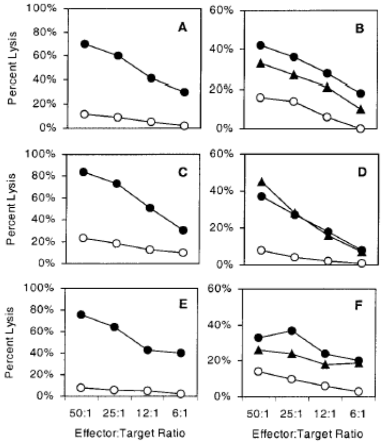


INFECTION AND IMMUNITY, June 2002, p. 2926–2932 Vol. 70, No. 6
0019-9567/02/$04.00+0 DOI: 10.1128/IAI.70.6.2926–2932.2002

Identification by Mass Spectrometry of CD8$^+$-T-Cell *Mycobacterium tuberculosis* Epitopes within the Rv0341 Gene Product

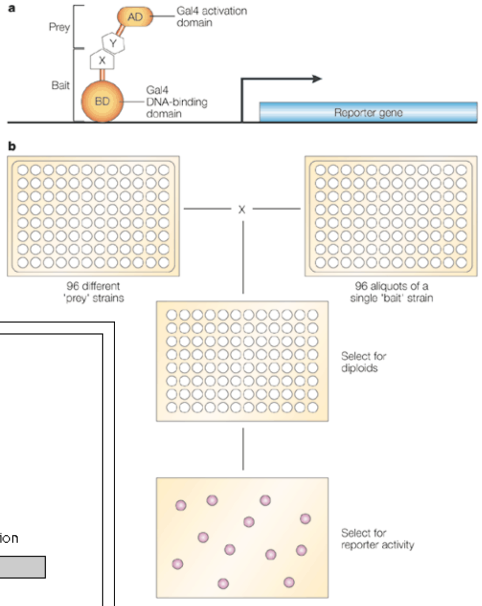


Rv0341 peptides found in purified MHC Class I

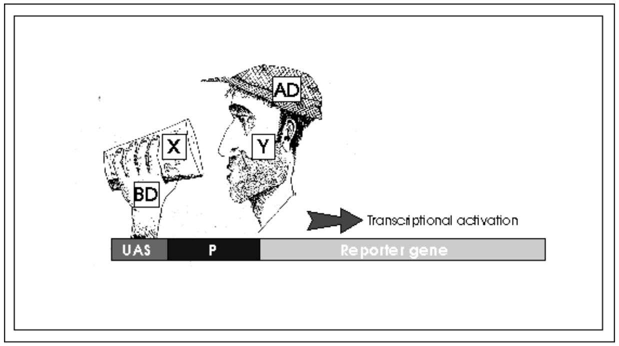


Rv0341 peptides recognize and lyses *M. tuberculosis*-infected dendritic cells.

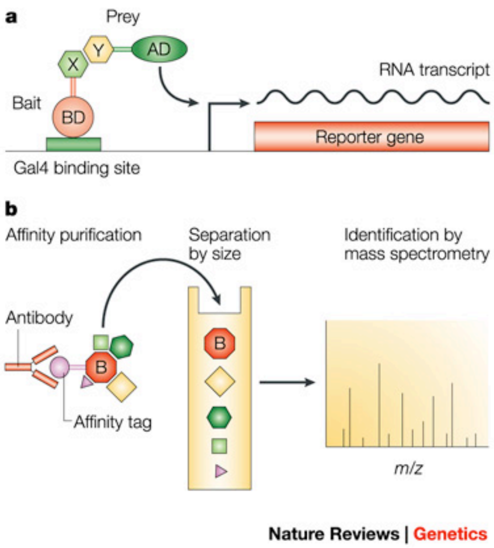
Two-hybrid Interaction Microarray



Nature Reviews | Genetics



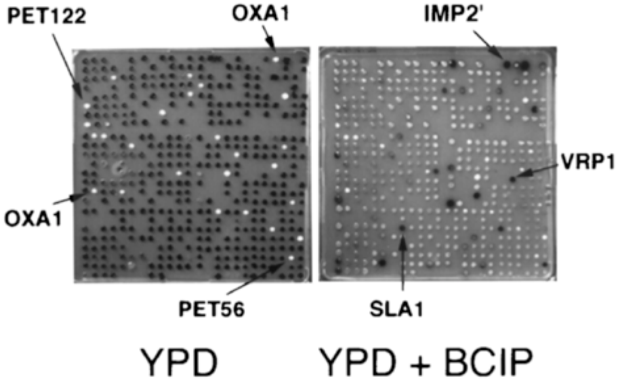
Systematic Proteomics




nature

Global yeast proteomics

Global Yeast Two-Hybrid



6000 Yeast transformants expressing an ORF as a fusion protein to an activation domain

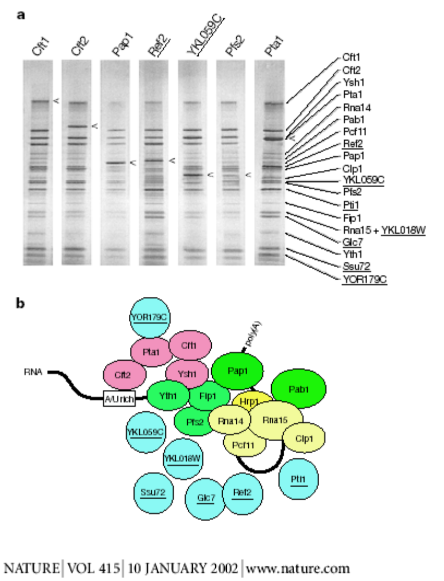
mate with

192 Yeast DNA binding domains hibrids

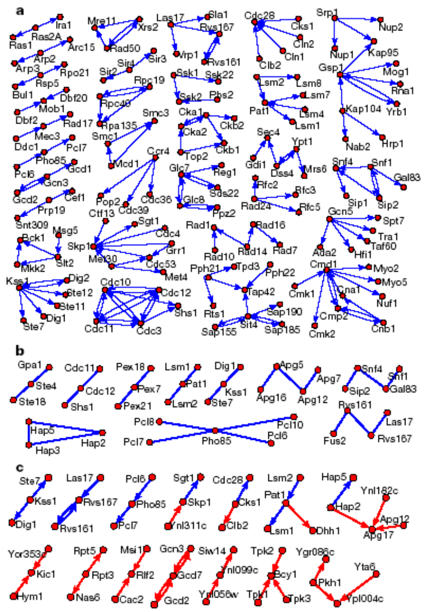
Diploid selection with markers carried on the two hibrid plasmids

Nature **403**, 623 - 627 (2000)

Protein-Protein Interaction (*In vivo*)



NATURE | VOL 415 | 10 JANUARY 2002 | www.nature.com



Molecular Cell, Vol. 12, 1353–1365, December, 2003, Copyright ©2003 by Cell Press

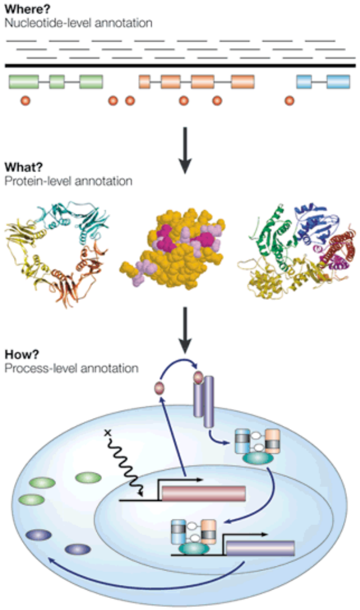
Assigning Function to Yeast Proteins by Integration of Technologies

Table 1. Assignment of GO Terms Based on the Experimental and Computational Evidence from the Four Technologies

	GO Terms Assignment			Experimental and Computational Evidence			
ORF	Biological process	Component	Molecular function	Process predicted by copurification	Process predicted by two-hybrid	Localization	Remote homology search and Protein structure prediction
YJR072C	Aerobic respiration	Cytoplasm	Signal peptide binding	None	Aerobic respiration	Cytoplasm	Signal peptide-binding domain superfamily; GTPase domain of the signal sequence recognition protein Ffh
YJR013W	Amino acid transport	Nuclear envelope-ER	Mannosyl transferase	None	Amino acid transport	Nuclear envelope-ER	Mannosyl transferase (PSI)
YFR003C	Cell cycle	Cytoplasm, nucleus[a]	Protein phosphatase inhibitor	Cell cycle	Morphogenesis*	No signal	Protein phosphatase type 1 inhibitor (PSI)
YIR010W (DSN1)	Chromosome segregation	Kinetochore	Unknown	Chromosome segregation	None*	Kinetochore	Myosin motor domain
YPL233W (NSL1)	Chromosome segregation	Kinetochore	Unknown	Chromosome segregation	None	Kinetochore	Spectrin repeat
YKL088W	Coenzyme A biosynthesis	Cytoplasm	Phosphopanto-thenoylcysteine decarboxylase	Coenzyme A biosynthesis	Cellular process	Cytoplasm	Phosphopantothenoylcysteine decarboxylase; Pantothenate metabolism flavoprotein superfamily
YDR531W	Coenzyme A biosynthesis[b]	Cytoplasm, nucleus	Pantothenate kinase	None	None	Cytoplasm, nucleus	Fumble, similar to pantothenate kinases
YIL083C	Coenzyme A biosynthesis[b]	Cytoplasm, nucleus	Phosphopanto-thenoylcysteine ligase	None	No positives	Cytoplasm, nucleus	Phosphopantothenoylcysteine synthetase/decarboxylase (PSI); Pantothenate metabolism flavoprotein superfamily
YKR079C	DNA catabolism; RNA catabolism	Nucleus, Mitochondrion	Nuclease	DNA or RNA catabolism	No positives	Nucleus, Mitochondrion	tRNA endoribonuclease (PSI) Metallohydrolase/oxidoreductase
YDR489W (SLD5)	DNA repair	Nucleus[a]	Unknown	Unknown complex 1	Unknown complex 1*	No data	Interferon-induced guanylate-binding protein
YOL146W (PSF3)	DNA repair	Nucleus[a]	Unknown	Unknown complex 1	None	No data	Ribosome recycling factor
YDR013W (PSF1)	DNA repair	Nucleus[a]	Unknown	Unknown complex 1	DNA repair*	No signal	None
YJL072C (PSF2)	DNA repair	Nucleus	Unknown	Unknown complex 1	Unknown complex 1*	Nucleus	ARM repeat, Cytokines superfamily
YML023c	DNA repair	Nucleus	Unknown	DNA repair	Cell growth and/or maintenance*	Nucleus	None
YDR288W	DNA repair	Nucleus	DNA binding	DNA repair	Response to DNA damage stimulus*	Nucleus	Adenylation domain of NAD+-dependent DNA ligase; Winged helix DNA-binding domain

(*continued*)

Where? What? How?



Nature Reviews | Genetics