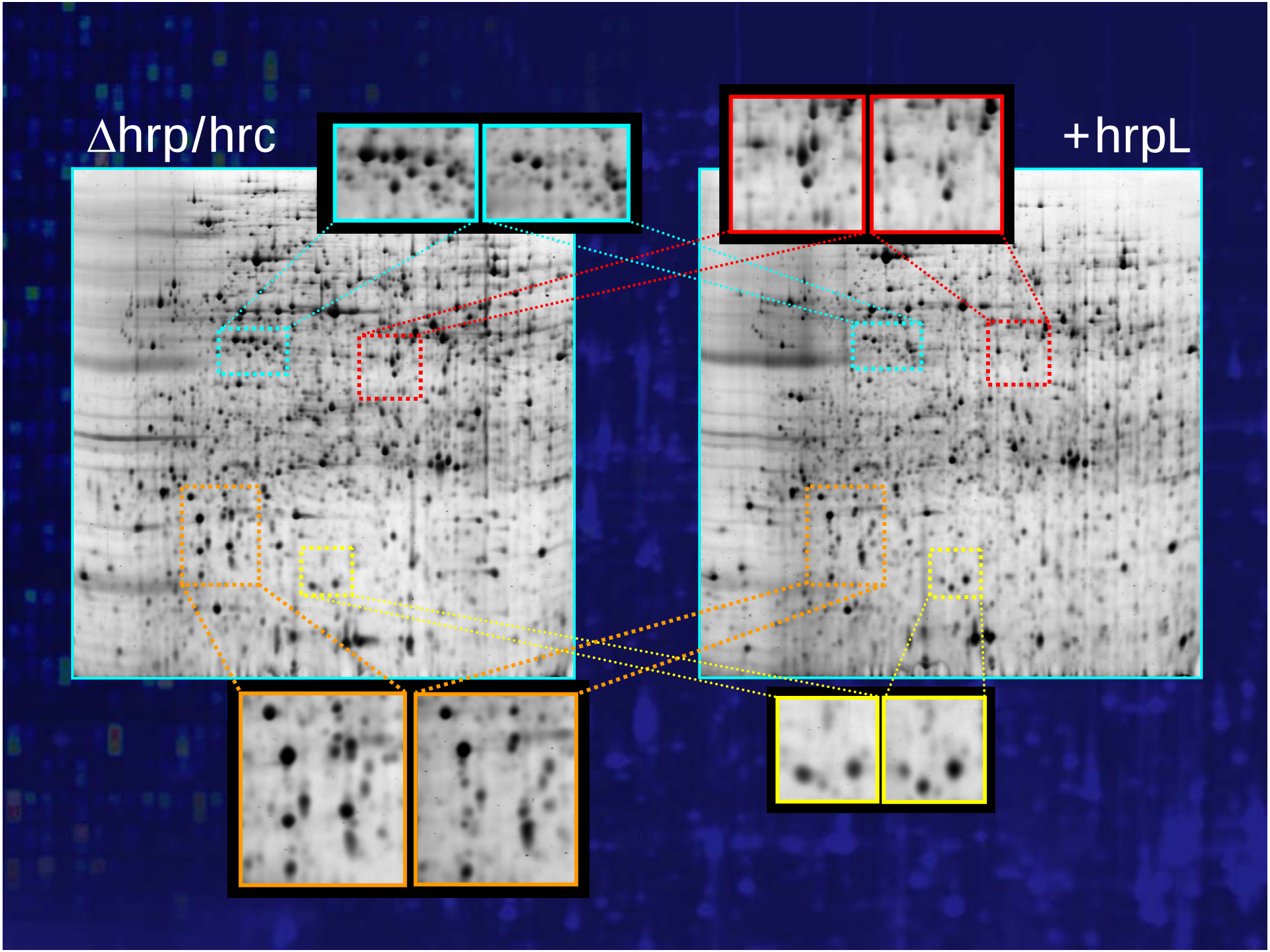
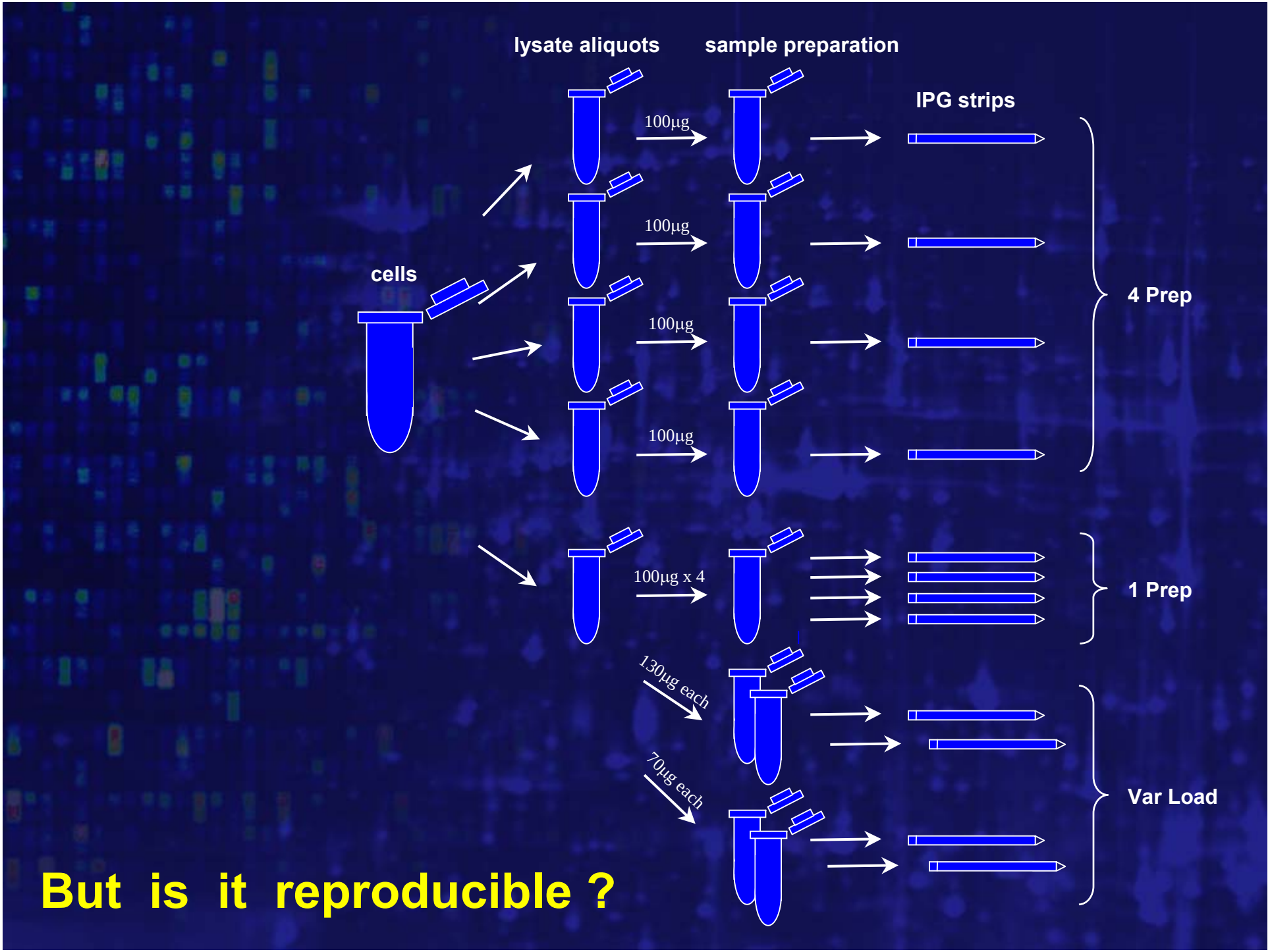
Proteomics and the Antemortem Diagnosis of Neurodegenerative Disease (and a vignette)

Kelvin H. Lee

Chemical and Biomolecular Engineering
Cornell Proteomics Program
Cornell University


Δhrp/hrc
+hrpL



But is it reproducible ?

Melanie 3 Analysis

Spots Matched in 2 out of 4 gels

Expt	#spots	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1 Prep	815	76%	88%	94%	97%	99%	99%	100%	100%
4 Prep	842	65%	79%	90%	96%	98%	99%	100%	100%
Var Load	919	72%	85%	92%	97%	98%	99%	100%	100%

Spots Matched in 3 out of 4 gels

Expt	#spots	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1 Prep	793	76%	88%	94%	97%	99%	99%	99%	100%
4 Prep	757	65%	79%	90%	96%	99%	99%	100%	100%
Var Load	879	73%	85%	93%	97%	99%	99%	100%	100%

Spots Matched in 4 out of 4 gels

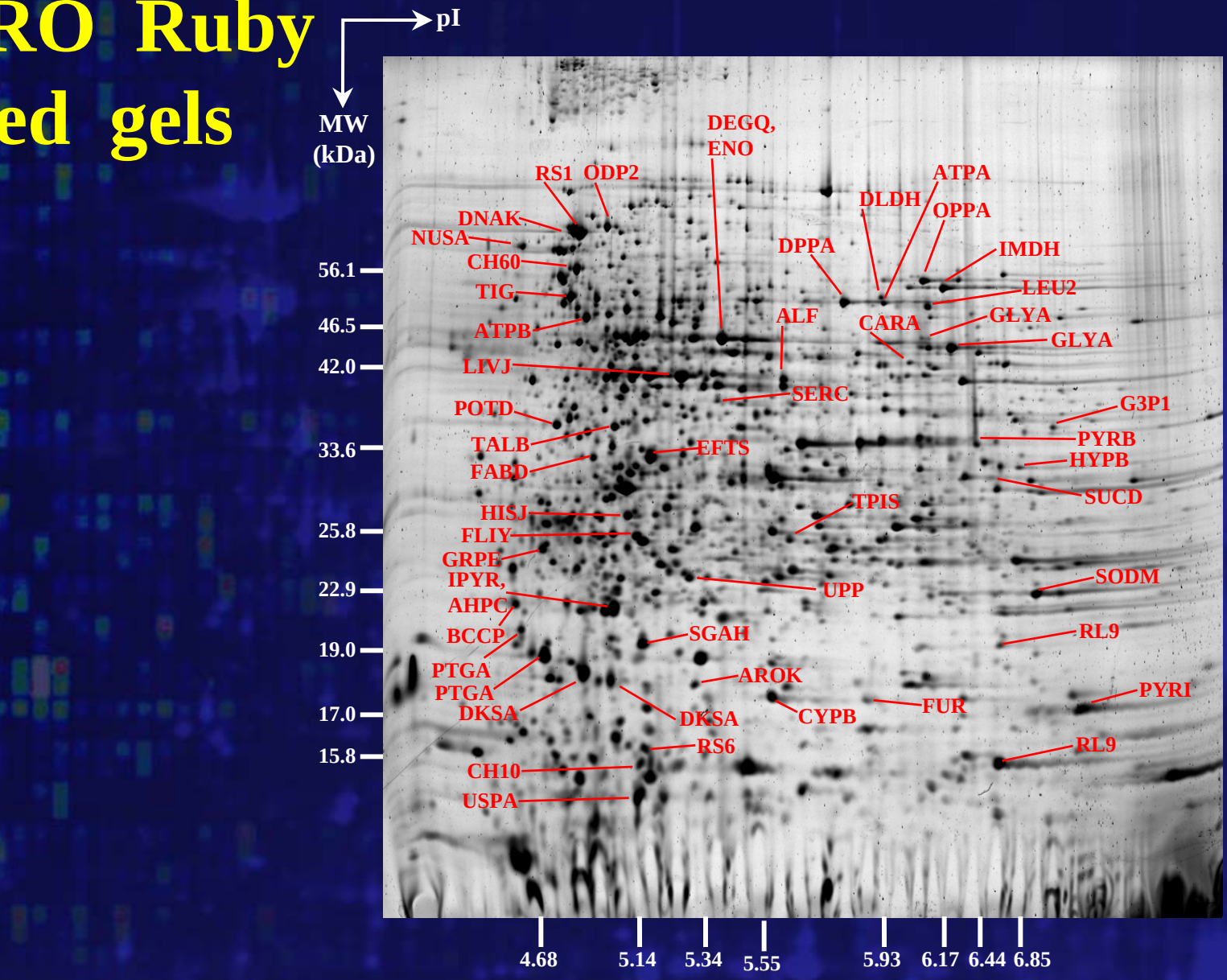
Expt	#spots	CV=0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1 Prep	771	76%	88%	94%	97%	99%	99%	99%	100%
4 Prep	641	65%	79%	90%	96%	99%	100%	100%	100%
Var Load	824	74%	86%	93%	97%	99%	99%	100%	100%

Expt	Qual CV
1 Prep	0.03
4 Prep	0.14
Var Load	0.05

Bottom line (with various caveats) :

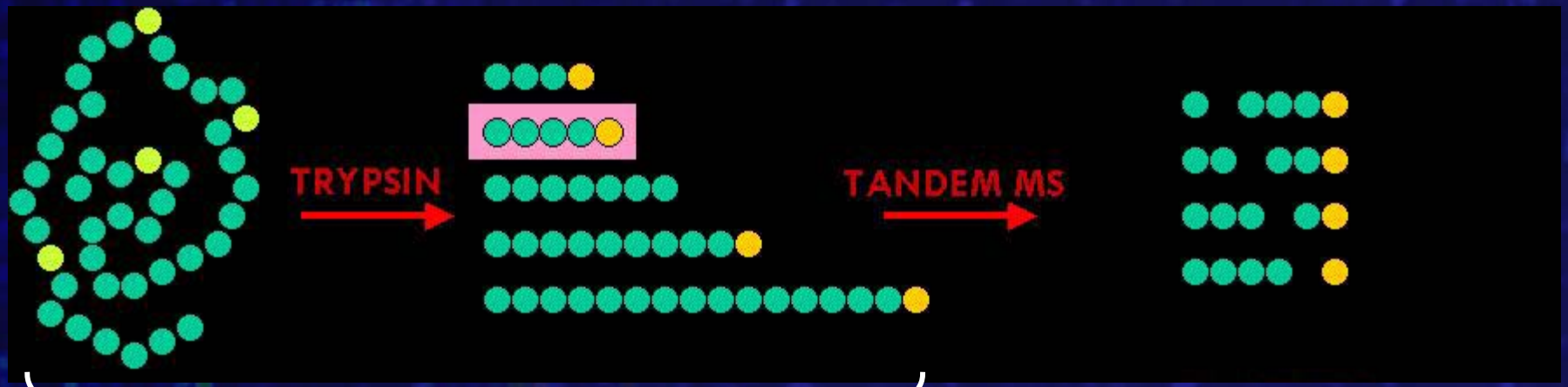
95% of the spots exhibit a CV < 0.52.

SYPRO Ruby stained gels



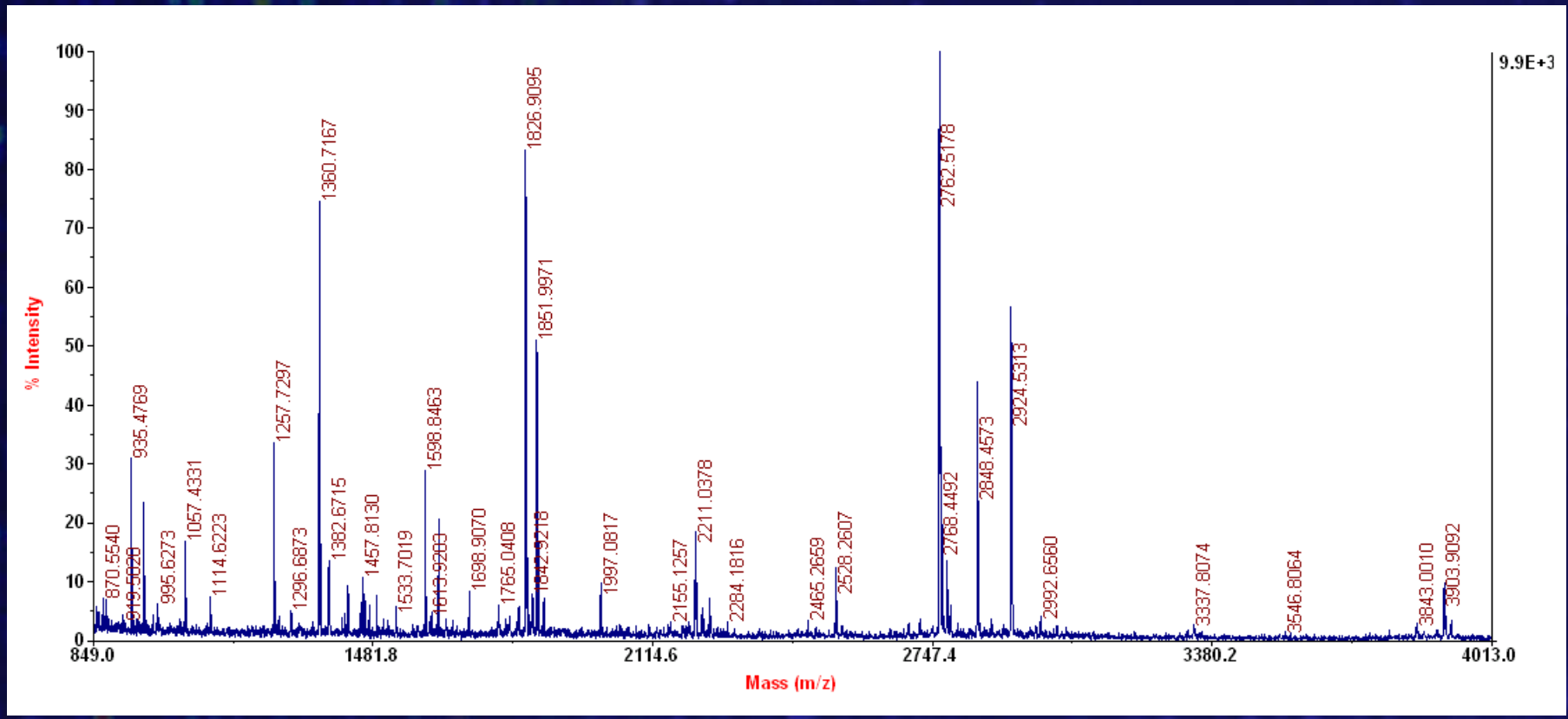
Characterization

Mass spectrometer measures mass of biomolecules with high accuracy.



Accurate mass measurement (ppm) of resulting tryptic peptides can be compared to *in silico* tryptic digests of sequence databases - Peptide Mass Fingerprinting.

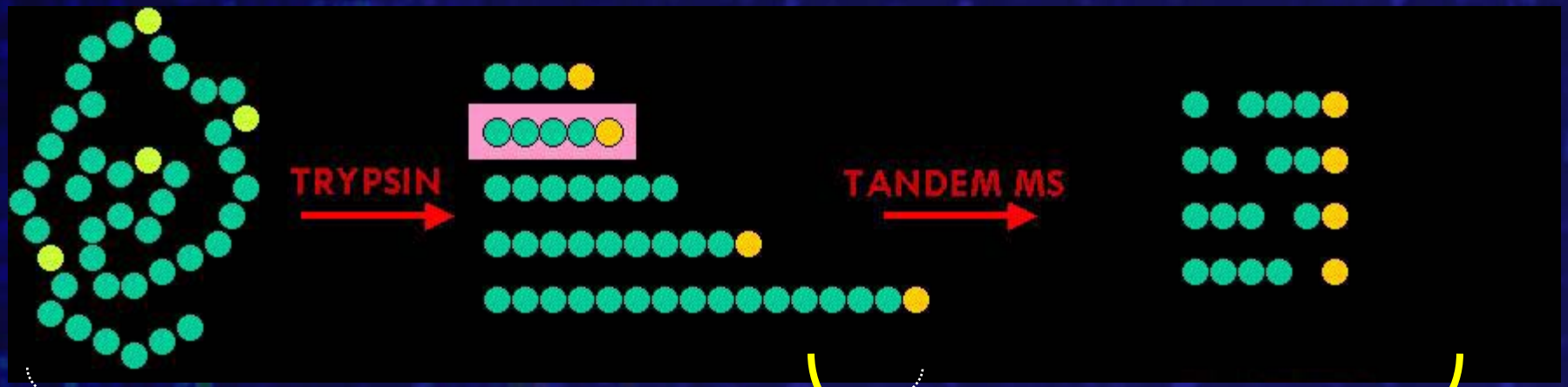
Peptide Mass Fingerprinting Enabled by Availability of DNA Sequence Information



Works well for prokaryotes. Less so for eukaryotes.

Characterization

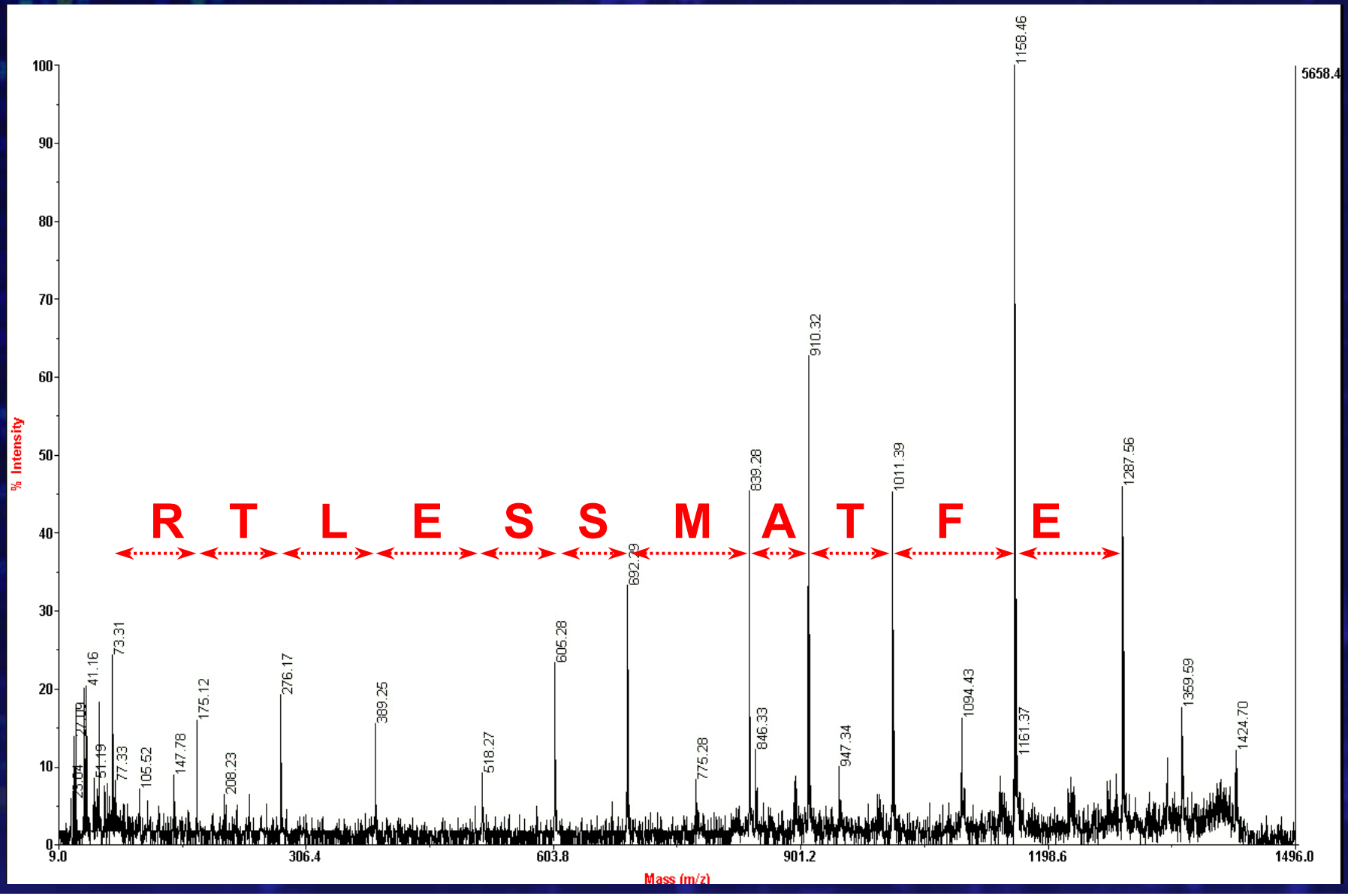
Mass spectrometer measures mass of biomolecules with high accuracy.



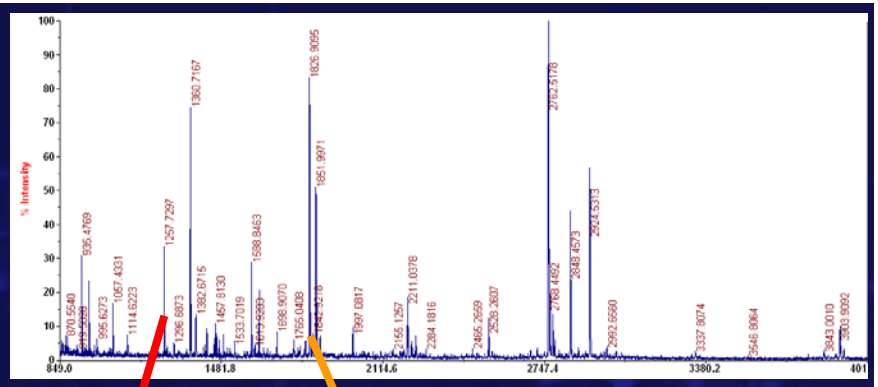
Peptide Mass Fingerprinting.

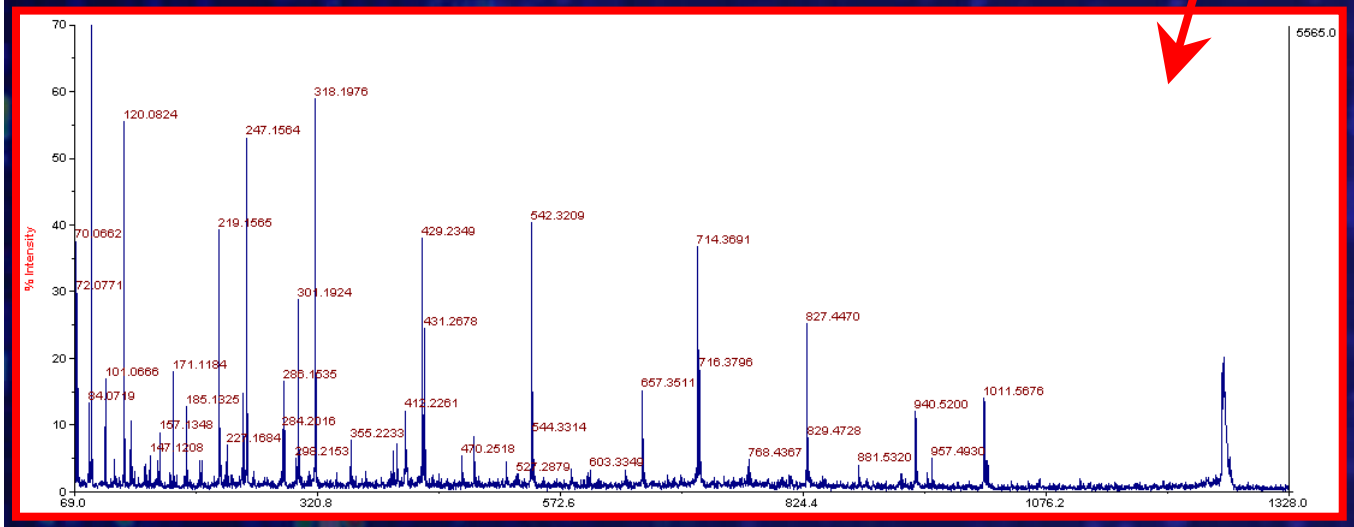
Individual peptides can be selected automatically by certain mass spectrometers for a tandem MS experiment.

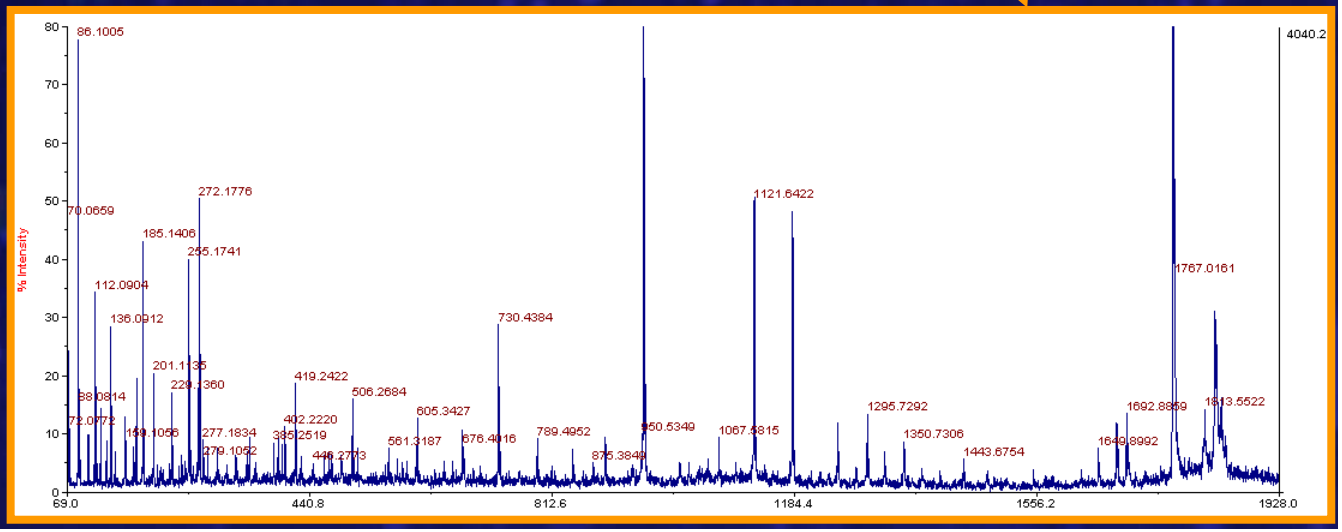
MS/MS Analysis of HO_2C-R-T-L-E-S-S-M-A-T-F-E-NH_2



Tandem MS enables sequencing & high confidence characterization. MALDI facilitates simplicity.







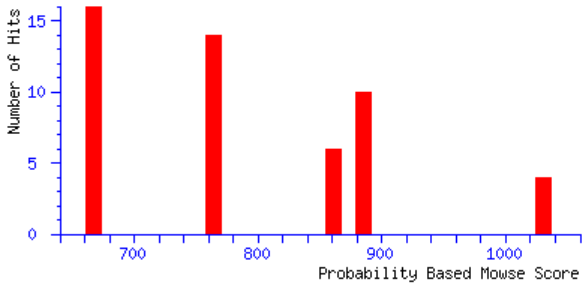
Mascot Search Results

Pseudomonas syringae TOF-TOF Example

Search title : SampleSetID: 2022, AnalysisID: 4593, MaldiWellID: 58508, S
Database : syringae6frame new (1836856 sequences; 298367028 residues)
Taxonomy : Bacteria (Eubacteria) (1836856 sequences)
Timestamp : 3 Oct 2003 at 17:57:19 GMT
Top Score : 1030 for 16853.00116675, 4887162..4888607

Probability Based Mowse Score

Score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 75 are significant (p<0.05).



Match to: 16853.00116675; Score: 1030
4887162..4888607

Nominal mass (M_r): 51882; Calculated pI value: 6.35
NCBI BLAST search of 16853.00116675 against nr

Variable modifications: Carbamidomethyl (C),Oxidation (M),Propionamide (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 82%

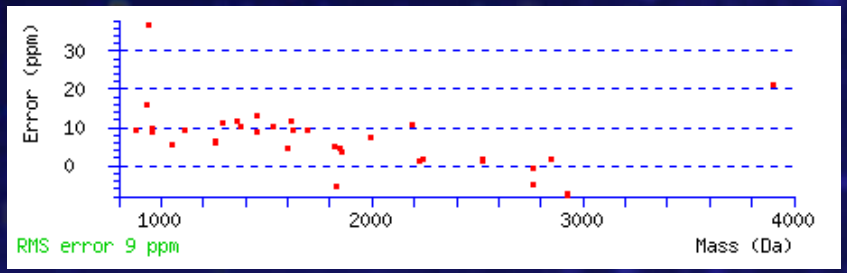
Matched peptides shown in Bold Red

```
    1 TVRRTKIVAT LGPASNSPEV IEQLILSGLD VARLNFSHGT PDEHKARAKL
   51 IREIAARHGR FVALLGDLQG PKIRIAKFTD KRIELKVGDK FTFSTAHPLT
  101 SGNQQIVGID YPDLVKDCGV GDELLLDDGR VVMRVDTQTA HELHCTVLIG
  151 GPLSDHKGIN RRGGGLTAPA LTEKDKQDIK LAAEMDLDYL AVSFPRDAAD
  201 MEYARQLRDE SGGTAWLVAK IERAEAVAND EVLDALIRAS DAVMVARGDL
  251 GVEIGDAELV GVQKRIILHA RRHNKAVIVA TQMMESMISS PMPTRAEVSD
  301 VANAVLDYTD AVMLSAESAA GSYPVEAVQA MARICVGAEK HPTGKTSSHR
  351 IGHSFTRCDE SIALAAMYTA NHFPGVKAII ALTESGYTPL IMSRIRSSVP
  401 IYAFSPHRGT QARAAMFRGV YTVPFDPGAL PPGQVSQAAV DELLKRGLVE
  451 QGDWVILTKG DSYHTIGGTN GMKILHVGDP LV
```

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
7 - 33	2762.52	2761.51	2761.52	-0.01	0	IVATLGPASNSPEVIEQLILSGLDVAR (Ions score 136)
7 - 33	2762.52	2761.51	2761.52	-0.01	0	IVATLGPASNSPEVIEQLILSGLDVAR (No match)
34 - 45	1381.66	1380.66	1380.64	0.01	0	LNFSHGTPDEHK (No match)
50 - 57	941.62	940.62	940.58	0.03	1	LIREIAAR (No match)
61 - 72	1257.73	1256.72	1256.71	0.01	0	FVALLGDLQGPK (Ions score 96)
61 - 72	1257.73	1256.72	1256.71	0.01	0	FVALLGDLQGPK (No match)
91 - 116	2848.46	2847.45	2847.44	0.00	0	FTFSTAHPLTSGNQQIVGIDYPDLVK (Ions score 71)
91 - 116	2848.46	2847.45	2847.44	0.00	0	FTFSTAHPLTSGNQQIVGIDYPDLVK (No match)
117 - 130	1533.70	1532.69	1532.68	0.02	0	DCGVGDELLLDDGR Carbamidomethyl (C) (No match)
135 - 157	2528.26	2527.25	2527.25	0.00	0	VDTQTAHELHCTVLIGGPLSDHK Carbamidomethyl (C) (Ions score 45)
163 - 174	1114.62	1113.61	1113.60	0.01	0	GGGLTAPALTEK (No match)
181 - 196	1826.91	1825.90	1825.89	0.01	0	LAAEMDLDYLAVSFPR Oxidation (M) (Ions score 79)
181 - 196	1826.91	1825.90	1825.89	0.01	0	LAAEMDLDYLAVSFPR Oxidation (M) (No match)
197 - 205	1057.43	1056.42	1056.42	0.01	0	DAADMEYAR Oxidation (M) (No match)
197 - 205	1057.43	1056.42	1056.42	0.01	0	DAADMEYAR Oxidation (M) (Ions score 5)
221 - 238	1997.08	1996.07	1996.06	0.01	1	IERAEAVANDEVLDALIR (No match)
221 - 238	1997.08	1996.07	1996.06	0.01	1	IERAEAVANDEVLDALIR (Ions score 10)
224 - 238	1598.85	1597.84	1597.83	0.01	0	AEAVANDEVLDALIR (Ions score 36)
224 - 238	1598.85	1597.84	1597.83	0.01	0	AEAVANDEVLDALIR (No match)
239 - 247	935.48	934.47	934.45	0.01	0	ASDAVMVAR Oxidation (M) (No match)
248 - 264	1698.91	1697.90	1697.88	0.02	0	GDLGVEIGDAELVGVQK (Ions score 35)
248 - 264	1698.91	1697.90	1697.88	0.02	0	GDLGVEIGDAELVGVQK (No match)
248 - 265	1855.00	1853.99	1853.98	0.01	1	GDLGVEIGDAELVGVQKR (No match)
266 - 272	878.58	877.57	877.56	0.01	1	IILHARR (No match)
276 - 295	2228.05	2227.04	2227.04	0.00	0	AVIVATQMMESMISSPMPTR 3 Oxidation (M) (No match)
276 - 295	2244.04	2243.03	2243.03	0.00	0	AVIVATQMMESMISSPMPTR 4 Oxidation (M) (No match)
296 - 333	3903.91	3902.90	3902.82	0.08	0	AEVSDVANAVLDYTDAVMLSAESAAGSYPVEAVQAMAR 2 Oxidation (M) (No r
334 - 345	1296.69	1295.68	1295.67	0.01	1	ICVGAEKHPTGK Carbamidomethyl (C) (No match)
358 - 377	2195.05	2194.04	2194.02	0.02	0	CDESIALAAMYTANHFPGVK Carbamidomethyl (C) (No match)
378 - 394	1835.98	1834.98	1834.99	-0.01	0	AIIALTESGYTPLIMSR (No match)
378 - 394	1852.00	1850.99	1850.98	0.01	0	AIIALTESGYTPLIMSR Oxidation (M) (No match)
378 - 394	1852.00	1850.99	1850.98	0.01	0	AIIALTESGYTPLIMSR Oxidation (M) (Ions score 18)
395 - 408	1629.90	1628.89	1628.88	0.02	1	IRSSVPIYAFSPHR (No match)
395 - 408	1629.90	1628.89	1628.88	0.02	1	IRSSVPIYAFSPHR (Ions score 16)
397 - 408	1360.72	1359.71	1359.69	0.02	0	SSVPIYAFSPHR (Ions score 59)
397 - 408	1360.72	1359.71	1359.69	0.02	0	SSVPIYAFSPHR (No match)
419 - 445	2768.45	2767.44	2767.44	-0.00	0	GVYTVPFDPGALPPGQVSQAAVDELLK (No match)
419 - 446	2924.53	2923.52	2923.54	-0.02	1	GVYTVPFDPGALPPGQVSQAAVDELLKR (Ions score 53)
419 - 446	2924.53	2923.52	2923.54	-0.02	1	GVYTVPFDPGALPPGQVSQAAVDELLKR (Ions score 53)
419 - 446	2924.53	2923.52	2923.54	-0.02	1	GVYTVPFDPGALPPGQVSQAAVDELLKR (Ions score 108)
446 - 459	1613.92	1612.91	1612.89	0.02	1	RGLVEQGDWVILTK (No match)
447 - 459	1457.81	1456.81	1456.79	0.01	0	GLVEQGDWVILTK (No match)
460 - 473	1453.66	1452.65	1452.63	0.02	0	GDSYHTIGGTNGMK Oxidation (M) (No match)
474 - 482	962.58	961.57	961.56	0.01	0	ILHVGDPLV (No match)
474 - 482	962.58	961.57	961.56	0.01	0	ILHVGDPLV (Ions score 42)

MOWSE scores represent probability of random match

1030 ⇒ P = 1.85×10^{-43}



Cerebrospinal Spinal Fluid (CSF) Proteins as Central Nervous System Disease Markers

Lumbar CSF

Analysis

Normal

Different

- disease specific
- disease stage (early/late)
- symptom specific
- pathology specific
- severity
- response to therapy
- sample specific

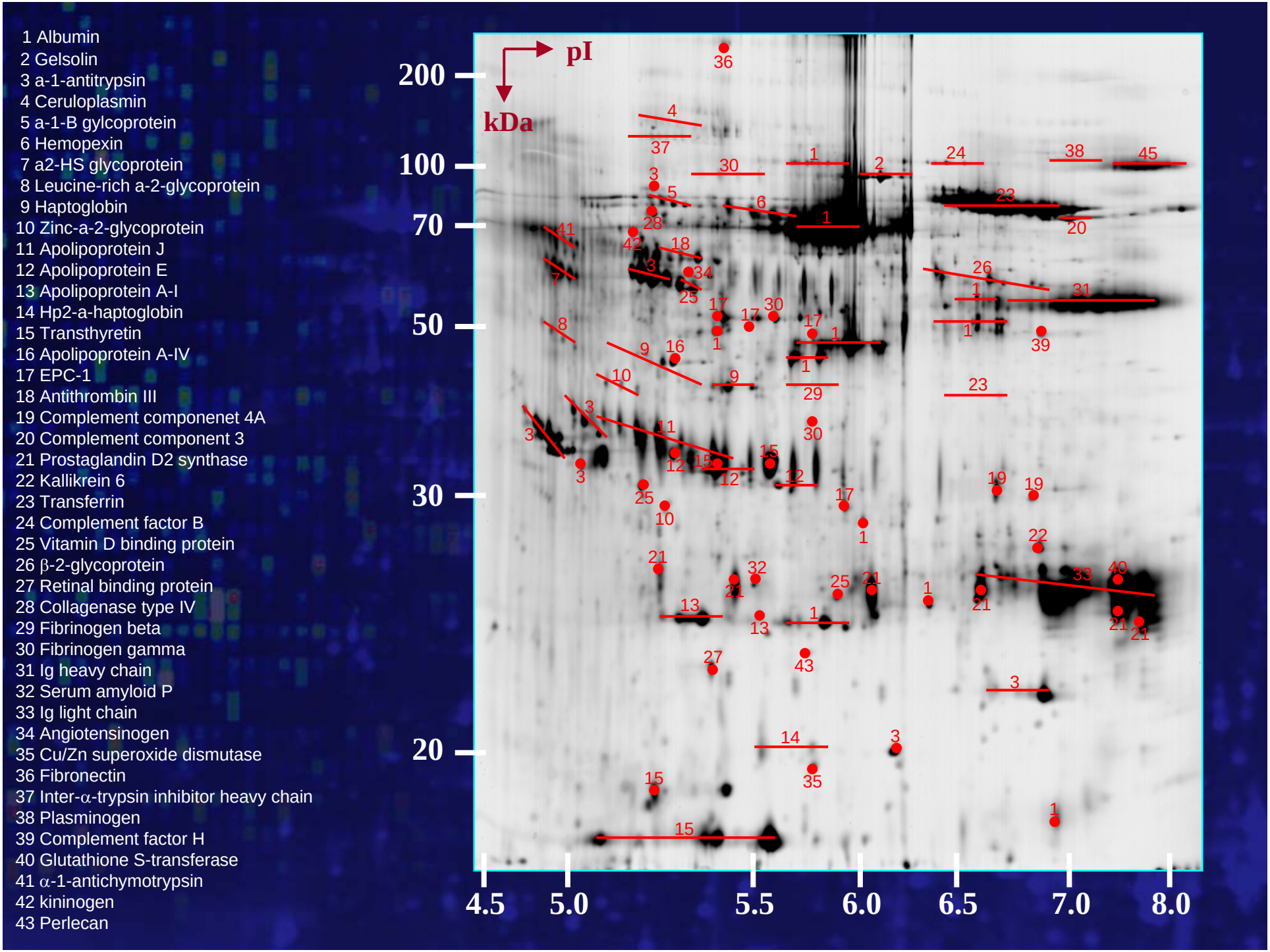
Are there any dementia biomarkers in CSF ?

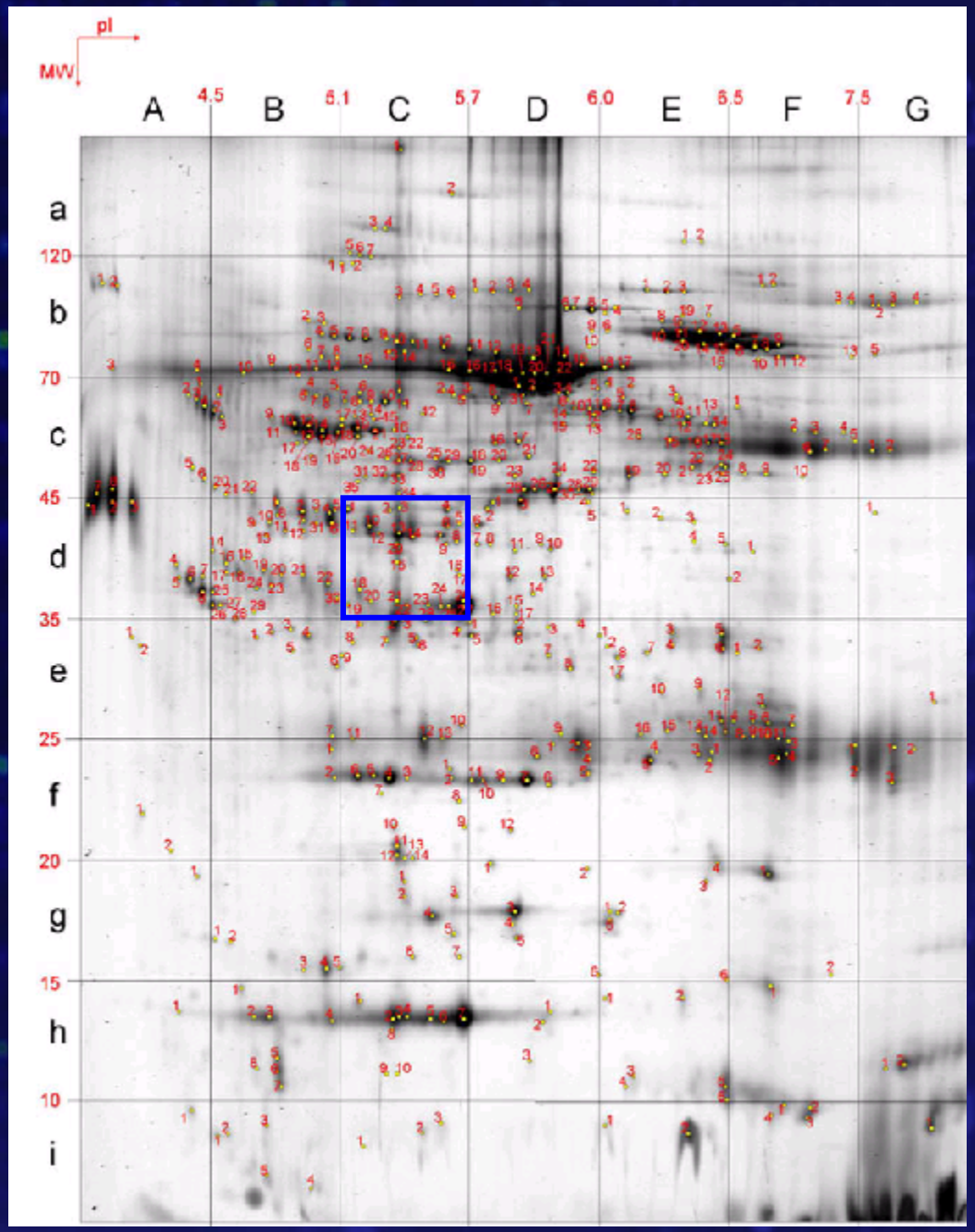
Vascular dementia - none
Dementia with Lewy bodies - none
Frontotemporal dementia - none
Alzheimer's disease - $A\beta_{1-42}$, Tau, AD7C-NTP

$A\beta_{1-42}$	83% sens, 82% spec
Tau	88% sens, 92% spec
$A\beta_{1-42}$ + Tau	85% sens, 87% spec
AD7C-NTP	70% sens, 87% spec

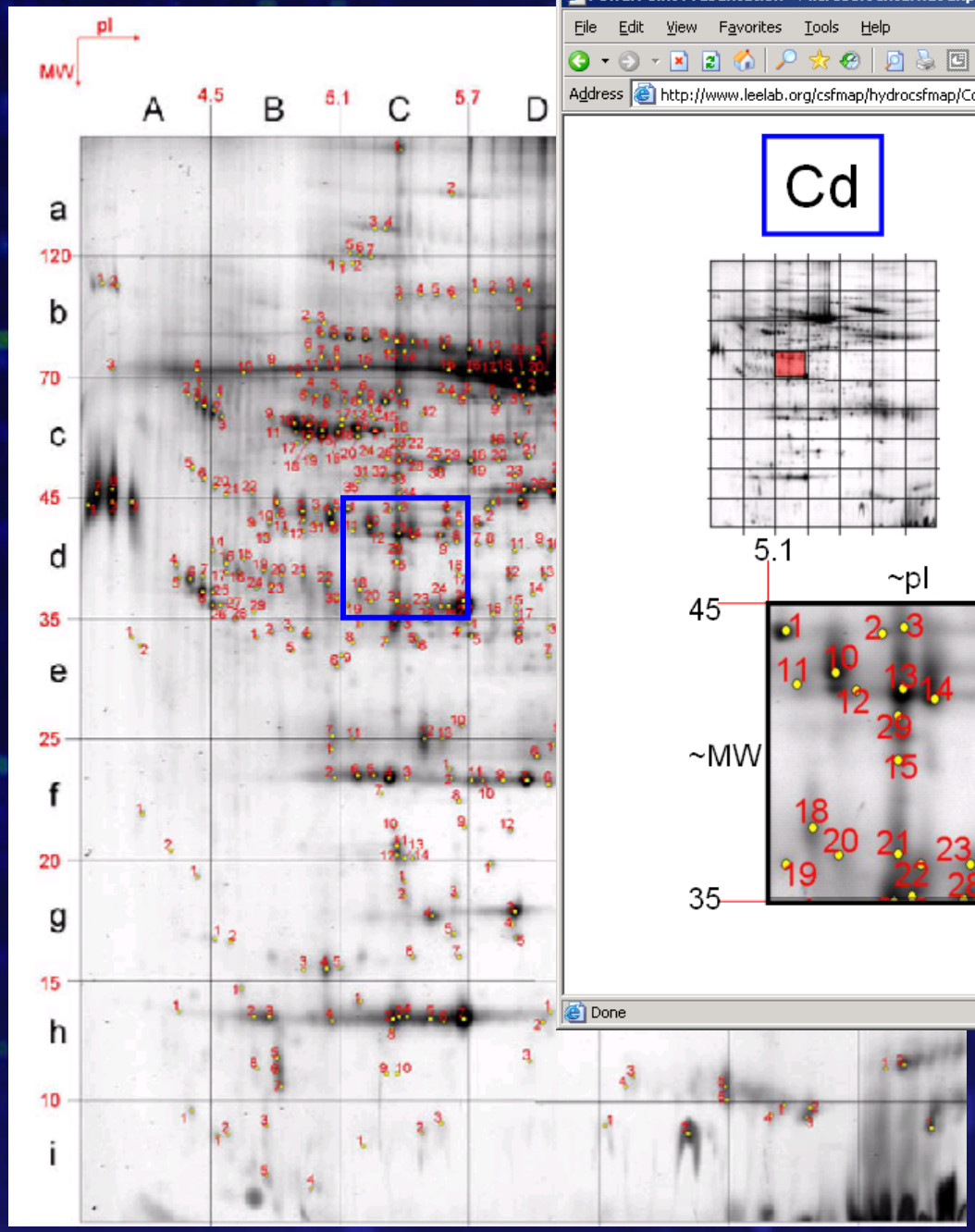
Sensitivity - true positive
Specificity - true negative

Neurology (2001) v56: 1143 Practice Parameter: Diagnosis of Dementia.


1 Albumin
2 Gelsolin
3 a-1-antitrypsin
4 Ceruloplasmin
5 a-1-B gylcoprotein
6 Hemopexin
7 a2-HS glycoprotein
8 Leucine-rich a-2-glycoprotein
9 Haptoglobin
10 Zinc-a-2-glycoprotein
11 Apolipoprotein J
12 Apolipoprotein E
13 Apolipoprotein A-I
14 Hp2-a-haptoglobin
15 Transthyretin
16 Apolipoprotein A-IV
17 EPC-1
18 Antithrombin III
19 Complement componenet 4A
20 Complement component 3
21 Prostaglandin D2 synthase
22 Kallikrein 6
23 Transferrin
24 Complement factor B
25 Vitamin D binding protein
26 β-2-glycoprotein
27 Retinal binding protein
28 Collagenase type IV
29 Fibrinogen beta
30 Fibrinogen gamma
31 Ig heavy chain
32 Serum amyloid P
33 Ig light chain
34 Angiotensinogen
35 Cu/Zn superoxide dismutase
36 Fibronectin
37 Inter-α-trypsin inhibitor heavy chain
38 Plasminogen
39 Complement factor H
40 Glutathione S-transferase
41 α-1-antichymotrypsin
42 kininogen
43 Perlecan
pI
kDa
200
100
70
50
30
20
4.5
5.0
5.5
6.0
6.5
7.0
8.0



www.leelab.org/csfmap



PowerPoint Presentation - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.leelab.org/csfmap/hydrocsfmap/Cd_files/slide0001.htm

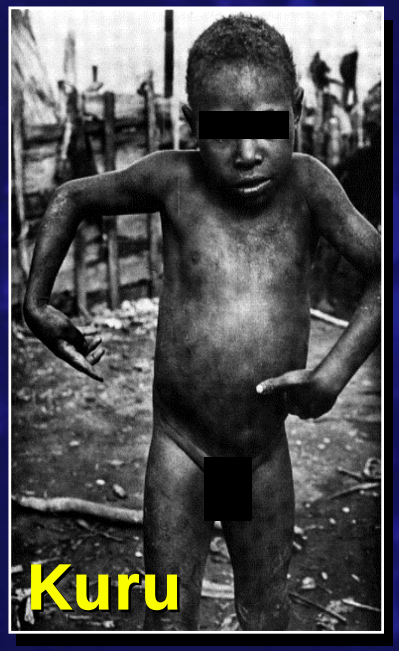
Spot	Identification	gi#	score	CI	Link
Cd1	Apolipoprotein AIV	178779	199	100	
Cd2	β-actin	14250401	167	100	
Cd3	Serum albumin	113576	63	99	
Cd4	Serum albumin	113576	67	99.9	
Cd5	Serum albumin	113576	73	99.9	
Cd6	Serum albumin	113576	152	100	
Cd7	Serum albumin	113576	141	100	
Cd8	Haptoglobin 1	123507	99	100	
Cd8	Serum albumin	113576	51	96	
Cd9	Serum albumin	113576	60	99.6	
Cd11	α-1-antitrypsin	1942629	92	100	
Cd12	Cathepsin L	4503155	25i	99	
Cd13	Haptoglobin 1	123507	154	100	
Cd13	Serum albumin	113576	87	99.9	
Cd14	Haptoglobin 1	123507	80	99.9	
Cd14	Serum albumin	113576	114	100	
Cd15	Haptoglobin 1	123507	22i	97	
Cd16	Haptoglobin 1	123507	27i	99.7	
Cd17	Apolipoprotein J	178855	20i	95	
Cd18	Apolipoprotein J	178855	149	100	
Cd19	Apolipoprotein E	178563	113	100	
Cd20	Apolipoprotein E	178563	56	98	
Cd21	Apolipoprotein J	178855	172	100	
Cd22	Apolipoprotein E	178563	104	100	
Cd23	Apolipoprotein E	178563	111	100	
Cd24	Apolipoprotein E	178563	119	100	
Cd25	Apolipoprotein J	178855	61	99.6	
Cd25	Transthyretin	339685	154	100	
Cd26	Transthyretin	339685	273	100	
Cd27	Transthyretin	339685	212	100	
Cd28	Apolipoprotein E	178563	232	100	
Cd29	Serum albumin	113576	70	99.9	

Done

Transmissible Spongiform Encephalopathy (TSE)

In humans:

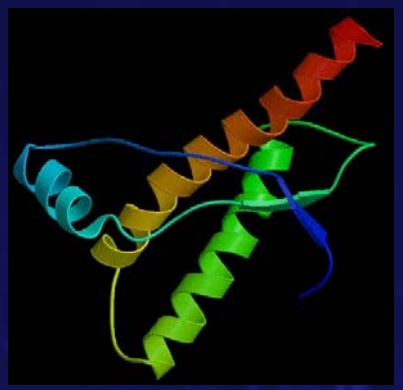
- Creutzfeldt–Jakob disease (CJD)
- New variant CJD
- Kuru
- others


Kuru

In animals:

- Scrapie (sheep)
- Bovine spongiform encephalopathy (BSE)
- Transmissible mink encephalopathy (TME)
- others

Molecular Pathology - An Enigma



- Normal prion protein (PrP) has unknown function.
- In TSEs, PrP is "misfolded" ($PrP \rightarrow PrP^{res}$) and amplified.
- BUT, one cannot reconstitute the disease with this pure PrP^{res} and different strains of PrP^{res} appear to exist.
- Heat, denaturants and proteases do not eliminate transmissibility.

Unifying Characteristics of TSEs

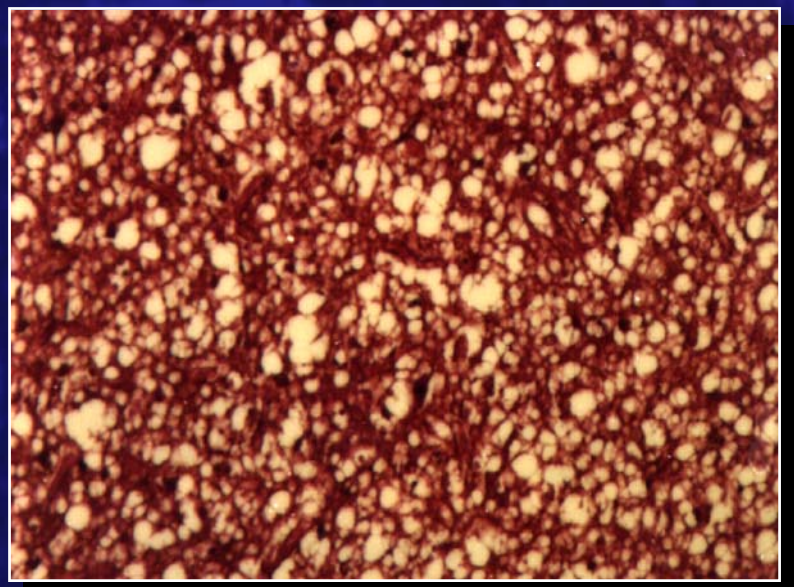
Transmissible (inter and intraspecies) - agent is very resistant to decontamination

Prion - no inflammatory response

Exposure, long incubation, onset of symptoms

Fatal

Spongiform pathology



Transmissible Spongiform Encephalopathy (TSE)

In humans:

Sporadic CJD — many years

New variant CJD — since 1996

others

In animals:

Scrapie (sheep) — many years

"Mad Cow disease" (BSE) — since 1986

others

Slaughter of all UK cattle > 30 months old has a significant economic impact.

Commercially - Available, Postmortem Screening Tools for BSE in Cattle

Several antibody-based tests with 100% sensitivity

Tests not validated for use on humans

A need to classify vCJD vs spCJD

No antemortem tests available

CJD in UK (since 1993)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
spCJD	38	51	35	40	59	63	62	49	55	73	57
iatrogenic CJD	4	1	4	4	6	3	6	1	3	0	4
other CJD	4	7	5	6	5	5	2	3	4	5	4
prob vCJD (alive)	0	0	0	0	0	0	0	0	0	0	0
vCJD (pending PM)	0	0	0	0	0	0	0	0	0	0	0
vCJD (confirmed)	0	0	3	10	10	18	15	28	20	17	18
Total	46	59	47	60	80	89	84	81	82	95	83

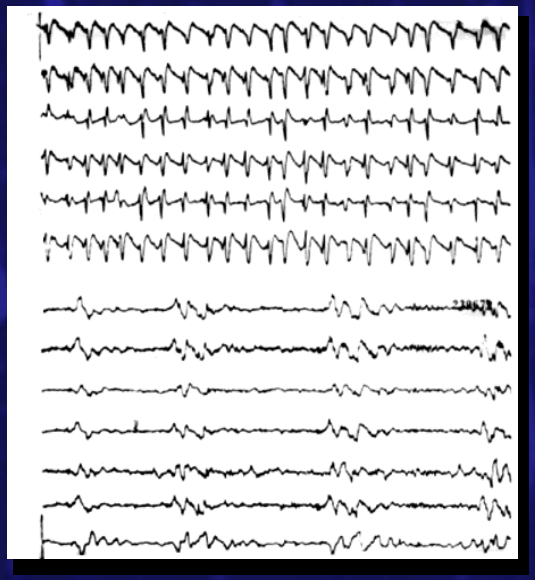
Total def & pr vCJD = 139 in UK

Differential Diagnosis: Alzheimer's, AIDS dementia, multi–infarct dementia, etc.

Antemortem Diagnosis - Humans

<u>Clinical Diagnosis</u>

- Atypical, subacute dementia
- Startle myoclonus
- "At risk" populations
- Brain biopsy
- Characteristic EEG

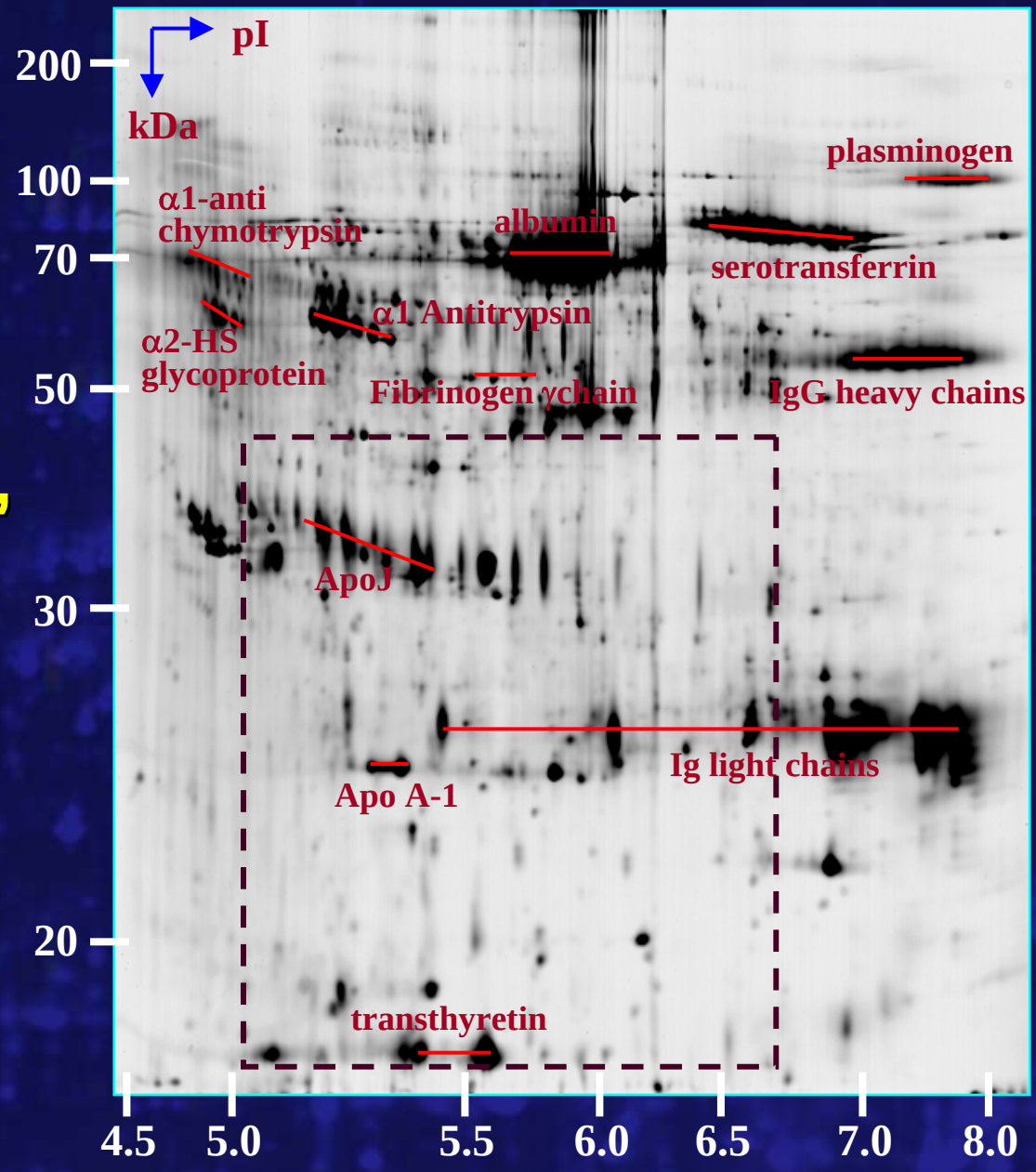


Iatrogenic transmission and may miss site of pathology

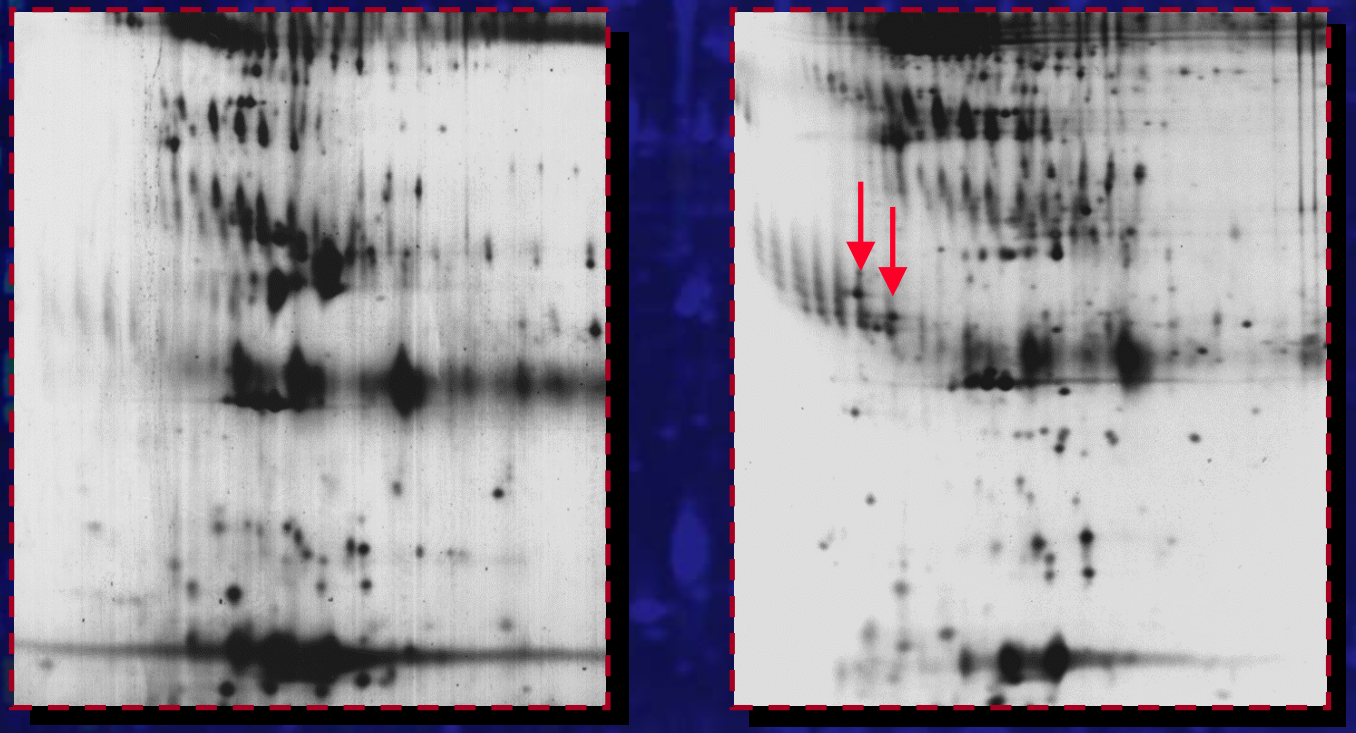
Not for atypical cases like vCJD

(Misdiagnosis of BSE and of animal TSEs)

Silver Stain 2DE–Separated Proteins in Cerebrospinal Fluid from a "Normal" Volunteer



Proteomics Identifies Two Creutzfeldt–Jakob Disease Associated Proteins



Results with the 130/131 Assay in the Differential Diagnosis of Demented Patients

69 of 70 (>98%)

CJD patients were positive

297 of 298 (>99%)*

control patients were negative

*primary CNS lymphoma

Proteomics as a Discovery Tool in Assay Development

2D Gels	Immunoassay
Complicated technology	Simple technology
Not automated	Automated
Too slow for the clinic	Fast
Lower throughput	High throughput

How can we purify enough material for sequencing ?

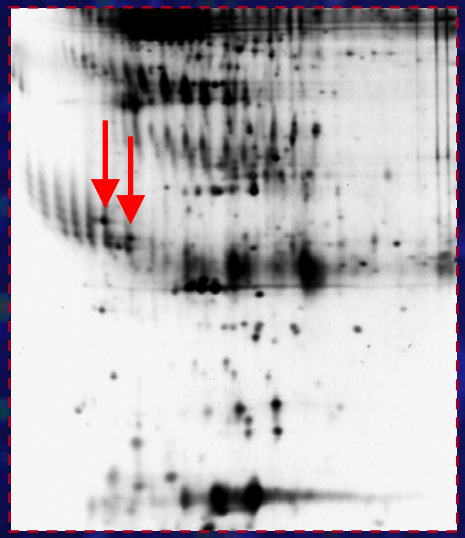
Hypothesis:

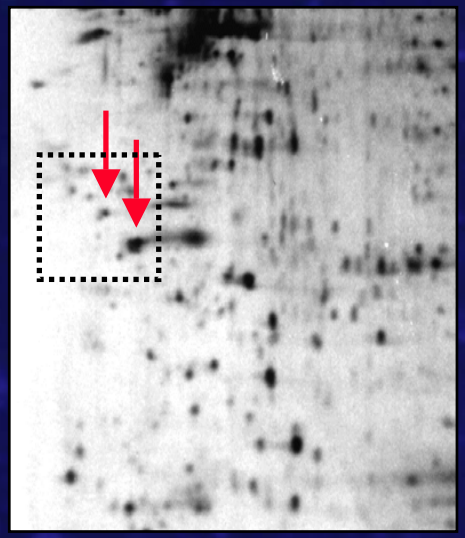
The appearance of 130/131 in CSF of CJD patients is a result of spongiform pathology. As neurons are destroyed their protein content leaks into the CSF.

Can we find 130/131 in brain extract from normal individuals?

Identification of CSF Proteins 130/131 in Normal Brain Extract

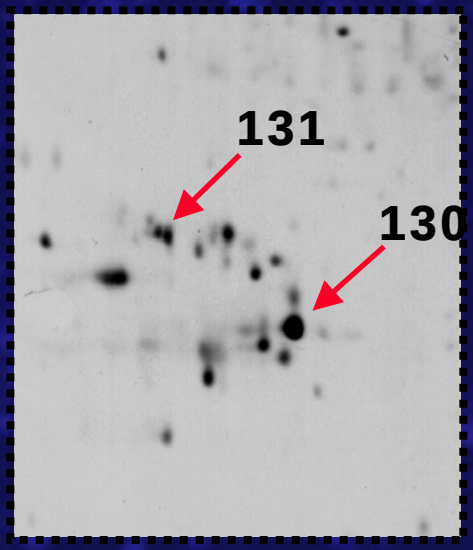
CJD CSF





Normal Brain

Normal Brain



Characterization of Spot 130

Brain spot corresponding to CSF 130 pooled from 10 blots
LysC digestion
4 sequences obtained (initial yield 4 ± 2 pmol):

VTELNEPLXNEDXNLLSVA

DYYXYLAEVATGEK

NVVXARRSSXRVISSIEQ

YSEAXEIS

human 14–3–3γ †

† Sequence information from A. Aitken

14–3–3 Family of Proteins

Seven isoforms

Highly conserved sequence

Variety of functions:

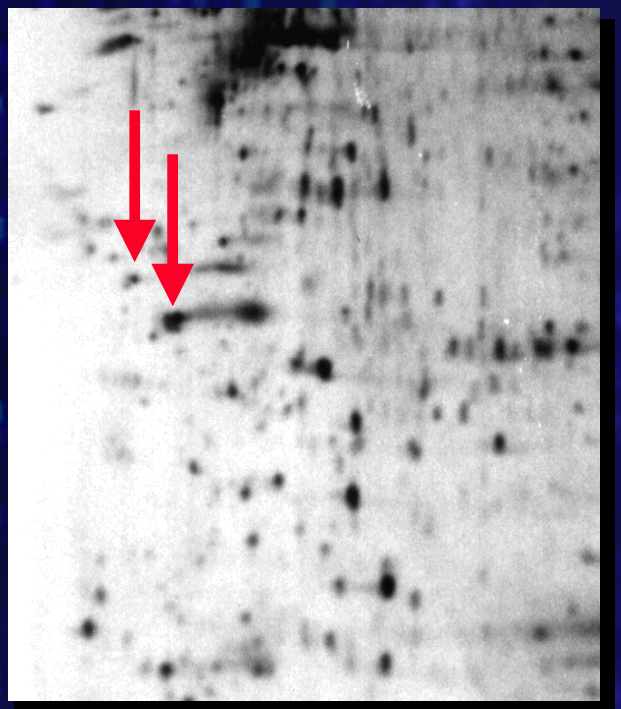
- signal transduction
- cell cycle control
- exocytosis
- melatonin biosynthesis

Identification of CSF 130/131 Proteins in Brain

sequence → antibody

brain

CSF



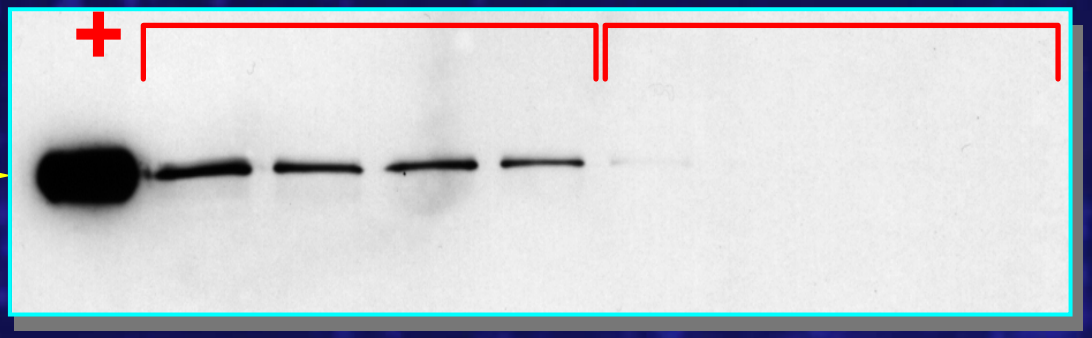
Anti–14–3–3 antibodies recognize both 130 and 131 on 2DE immunoblots of CJD CSF and brain extract.

1D Assay for TSEs including BSE

1: Alzheimer's
2 & 3: CJD
4: normal cow
5: experimental TME in cow
6: normal human serum
7: CJD serum
8: normal human brain
9: PrP–pur. brain extract – CJD

◄ 30 kD

BSE control



30 kD ►

ANIMALS	TOTAL SAMPLES	POSITIVE SAMPLES
Cattle: induced TME		
+ pathology	5	3
– pathology	1*	1*
Cattle: induced scrapie	4	3
Cattle: normal controls	15	0
Cattle: BSE	10	10
Cattle: BSE normal controls	6	1
Sheep: scrapie	6	5
Sheep: normal controls	1	0
Chimps: induced TSE	15	15
Chimps: normal controls	77	0
Total animals with + path	40	36
Total animals with – path	100	2

***This cow had clinical symptoms of TSE but normal histopathology**

Effectiveness of Assays for TSE

2DE - Human

Overall Sensitivity > 98 % (69/70)

Overall Specificity > 99 % (297/298)

1D - Human

Overall Sensitivity > 96 % (68/71)

Overall Specificity > 99 % (90/91)

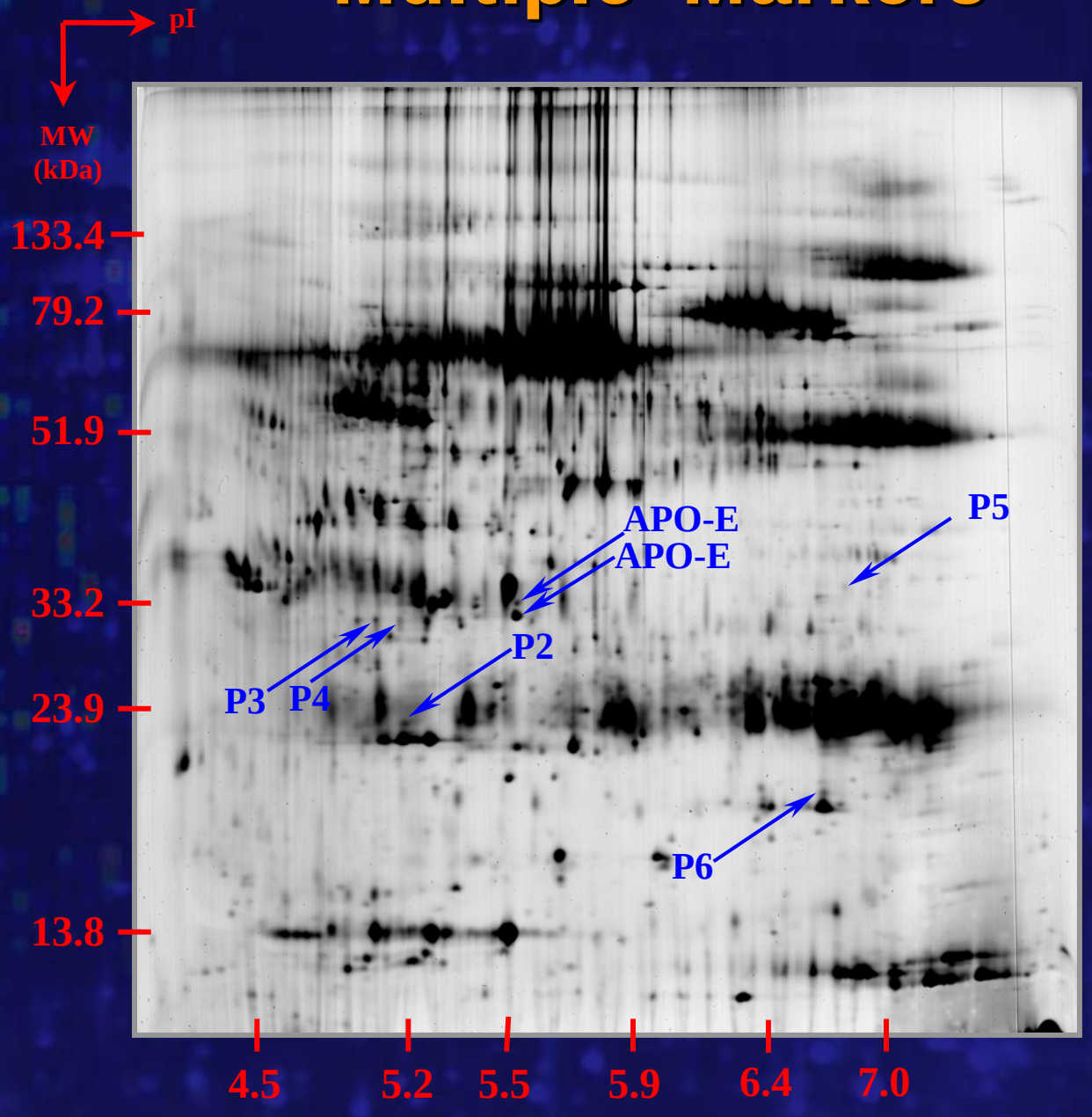
1D - Animal

Overall Sensitivity 90 % (36/40)

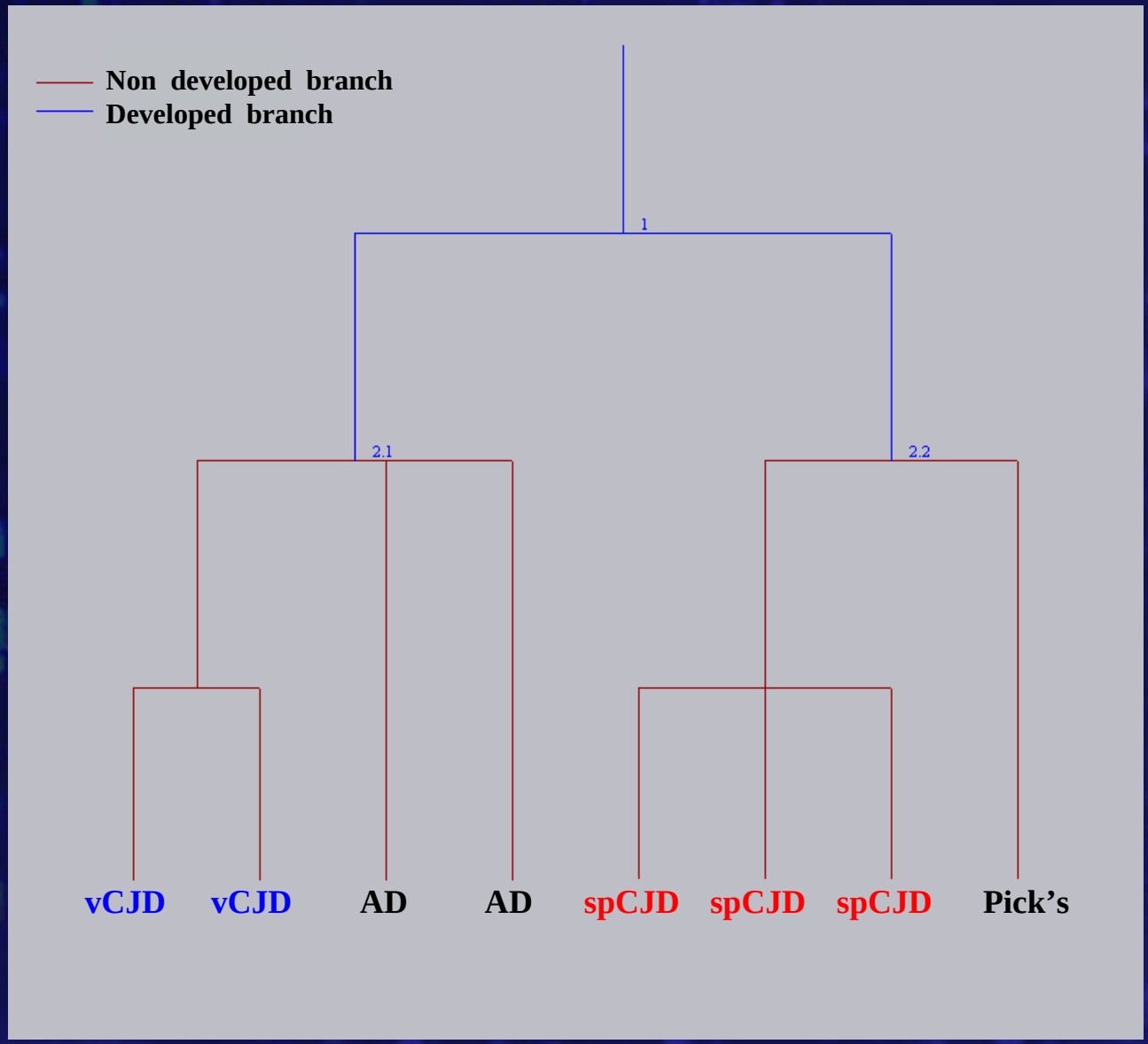
Overall Specificity 98 % (98/100)

Multiple Markers

vCJD
vs
spCJD
vs
other



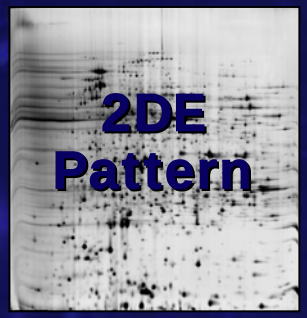
Heuristic Clustering – Class Level 2



Developing Disease Bar Codes

Biomarker Profiles

Multiple sclerosis
Schizophrenia
Parkinson's disease
Alzheimer's disease
CJD (BSE in cattle, scrapie in sheep)
Animal diseases



Screen

Quantitation
- Computer Algorithms
- CSF Database

Isolate Individual Biomarkers

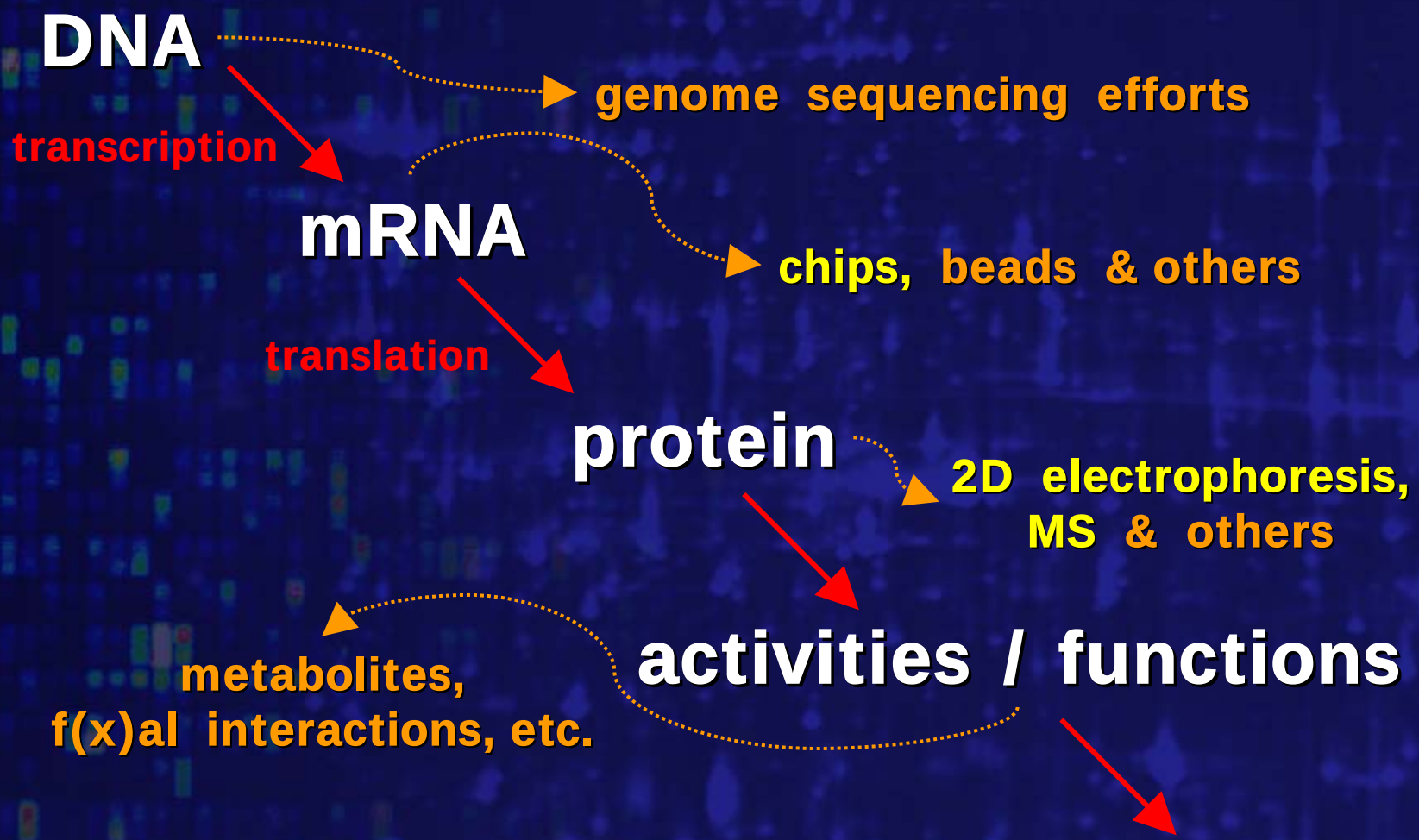
Purify / MS

Develop **antibodies** to biomarkers

Western–blotting

Immunoassay

The "Central Dogma" (?)

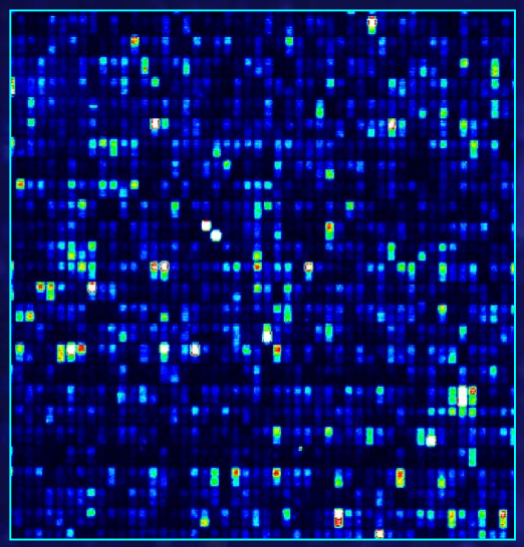
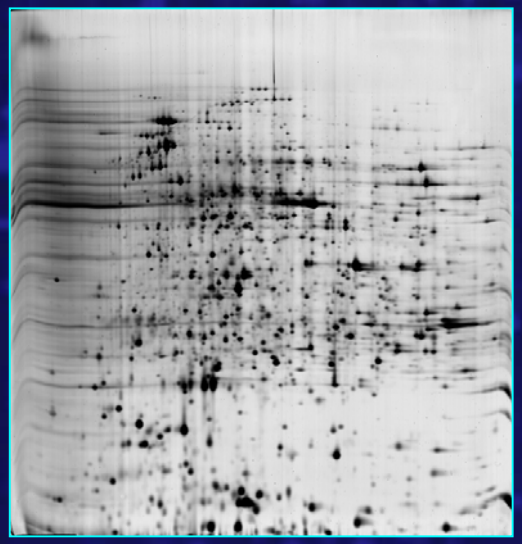


The machine does not isolate us from the great problems of nature but plunges us more deeply into them.

The Wind, Sand and Stars, Antoine de Saint-Exupéry, 1939

A second story

Chip to Gel Mapping

f  ? 

DNA

↓

mRNA

↓

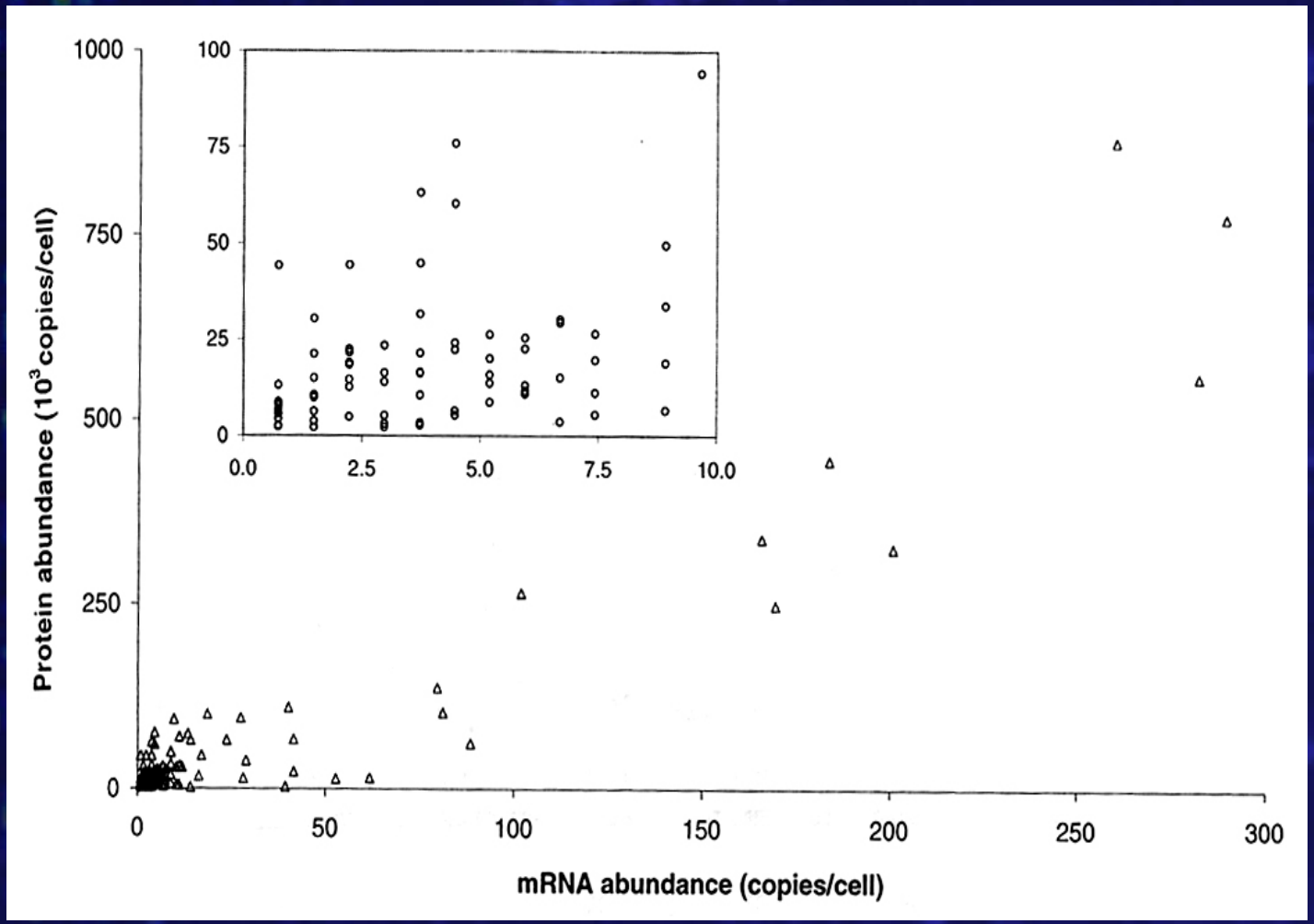
protein

↓

activities

↓

Yeast



Gygi *et al* 1999

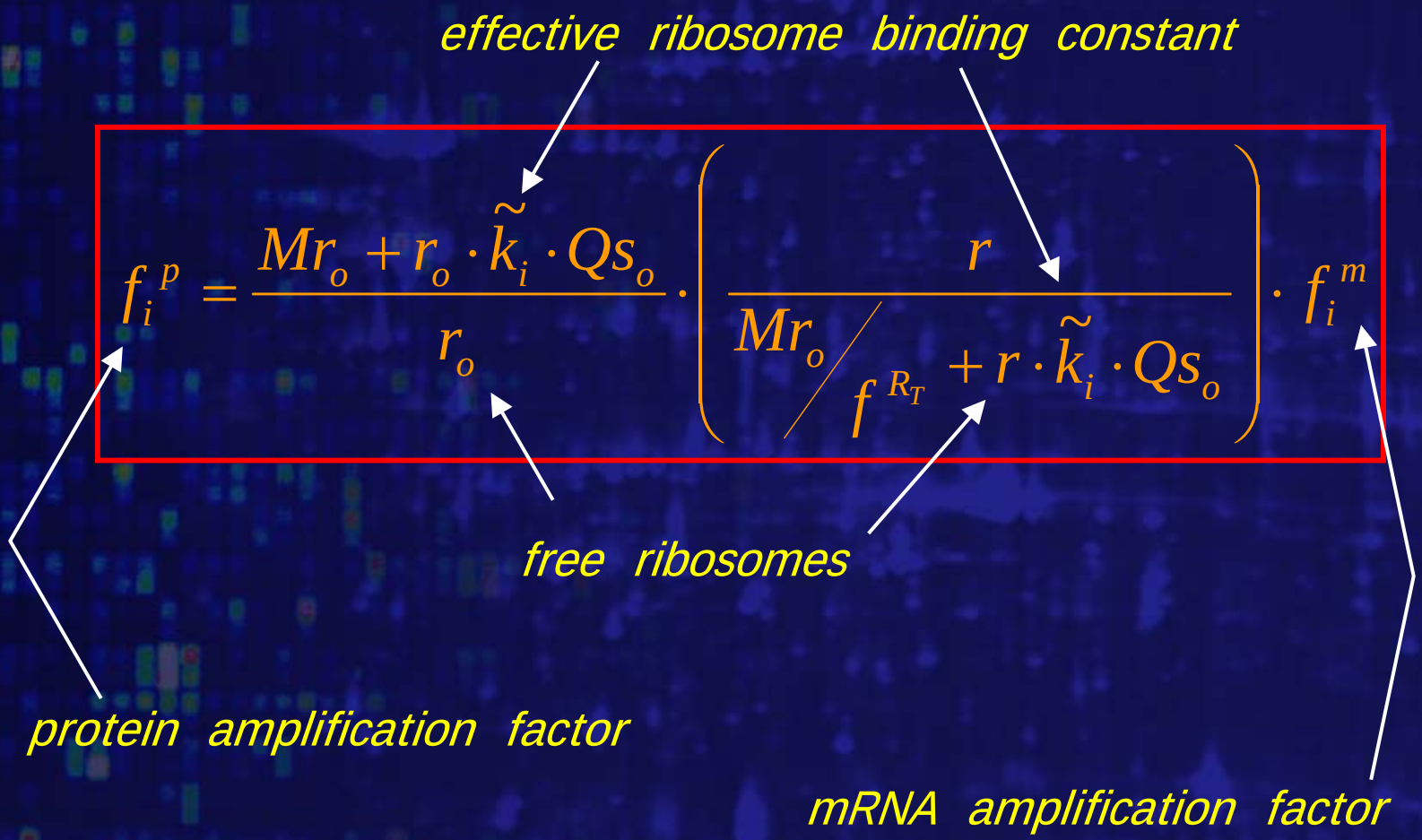
$$f_j^m \equiv \frac{M_{T,j}}{M_{T,j,o}} \overset{?}{\approx} f_j^p \equiv \frac{P_j}{P_{j,o}}$$

"mRNA amplification factor"

"protein amplification factor"

If mRNA for gene X increases 2-fold when the system is perturbed, how much does the protein for gene X increase by?

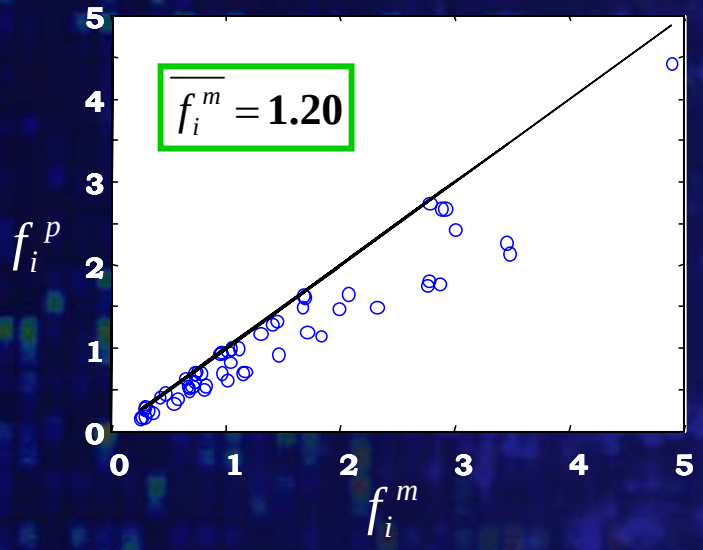
SOLVING THE SYSTEM : A KEY RESULT

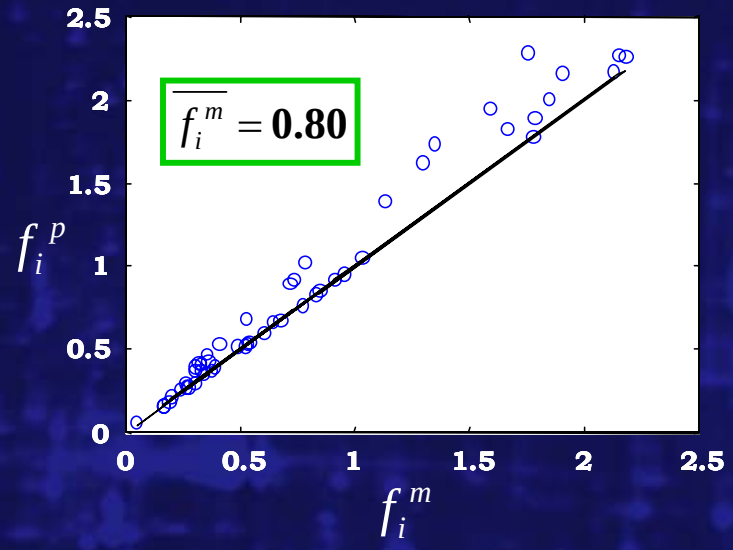


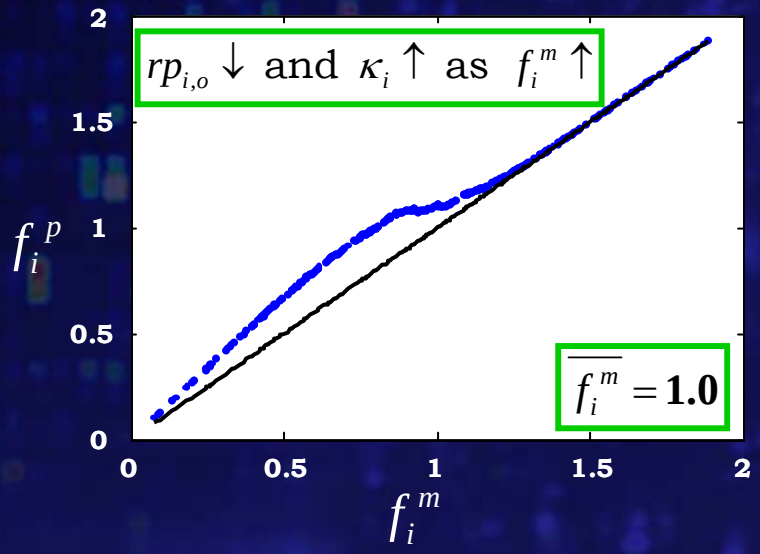
where $f^{\,R_T} \equiv \frac{R_T}{R_{T,o}}$ ratio of total # ribosomes

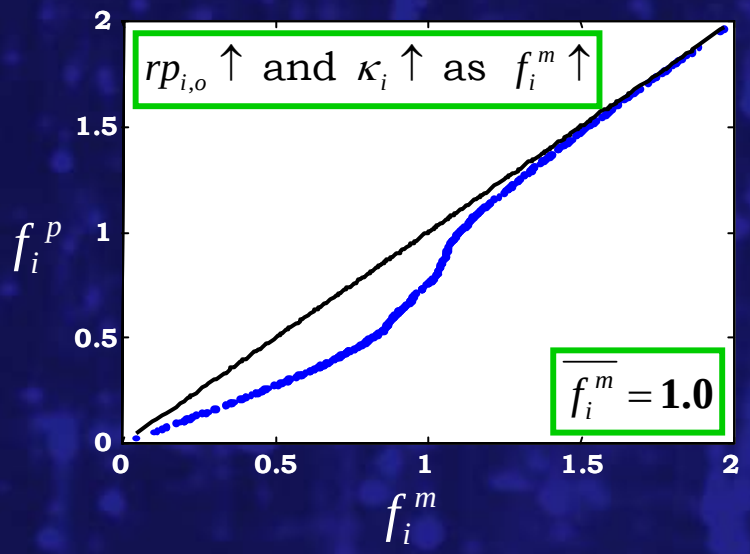
COMPUTATIONAL STUDIES

$\overline{f_i^m}$ total mRNA in the cell



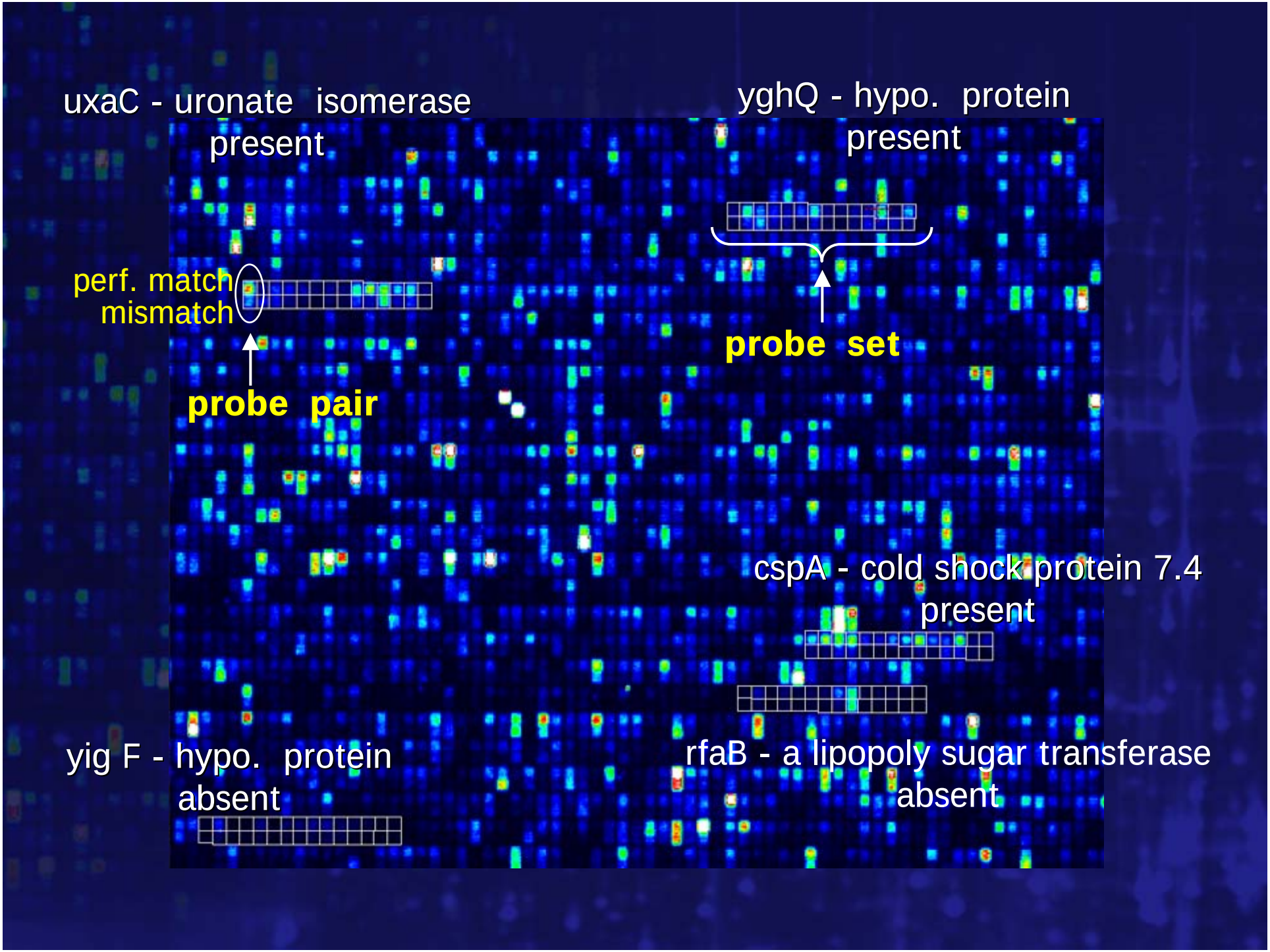




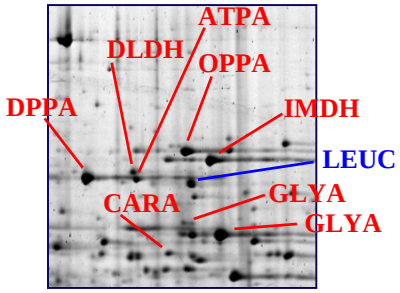


Experimental Results - *E. coli*

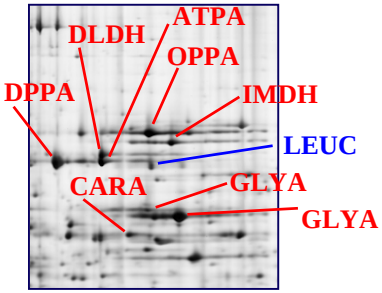
- Induction of parallel cultures with 0, 0.1, 1.0 mM IPTG.
- Comparison of Hly super-secreting mutant vs controls.
- Studies on the effect of Fis expression level.
- mRNA analysis by Affymetrix Genechips - 3 metrics.
- Protein measurements by 2DE/MS & ICAT - 4 metrics.
- Biological Replicates


uxaC - uronate isomerase
present
yghQ - hypo. protein
present
perf. match
mismatch
probe pair
probe set
cspA - cold shock protein 7.4
present
yig F - hypo. protein
absent
rfaB - a lipopoly sugar transferase
absent

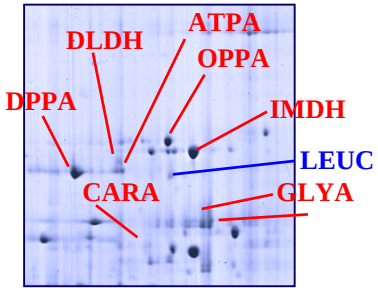
leuC (P30127; B0072)



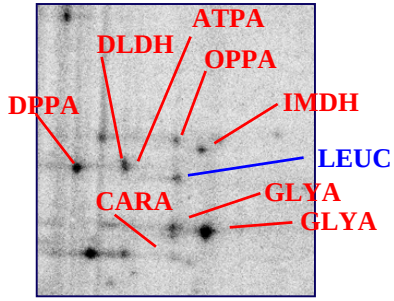
SYPRO Ruby



Ammonical Silver



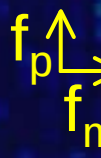
Colloidal Blue



^{35}S methionine

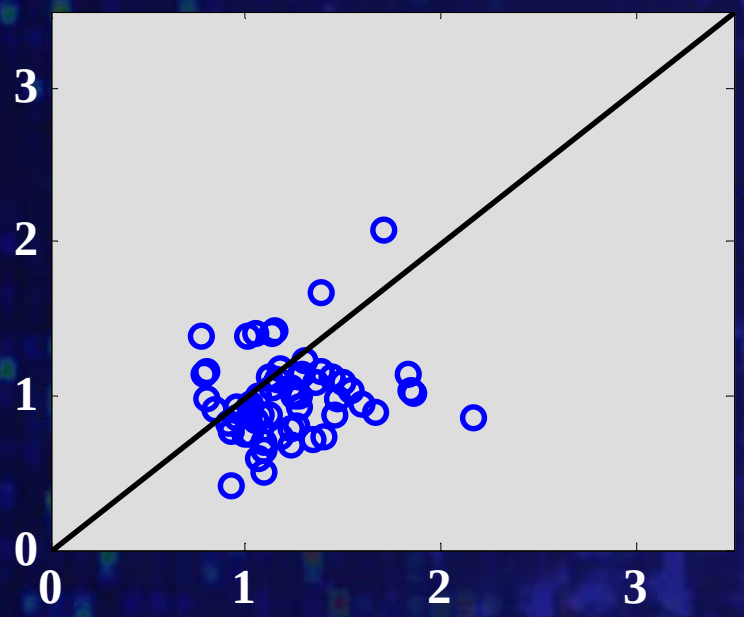
Fold change (f_p)	(low)	(high)
Ruby (%vol)	1.14	0.88
Silver (%vol)	1.99	1.75
Blue (%vol)	0.97	0.58
S35 (vol/stds)	0.76	0.35

leuC protein is downregulated (~25% for the culture) for shift from 0 mM IPTG to 0.1 mM IPTG to 1.0 mM IPTG.



expt

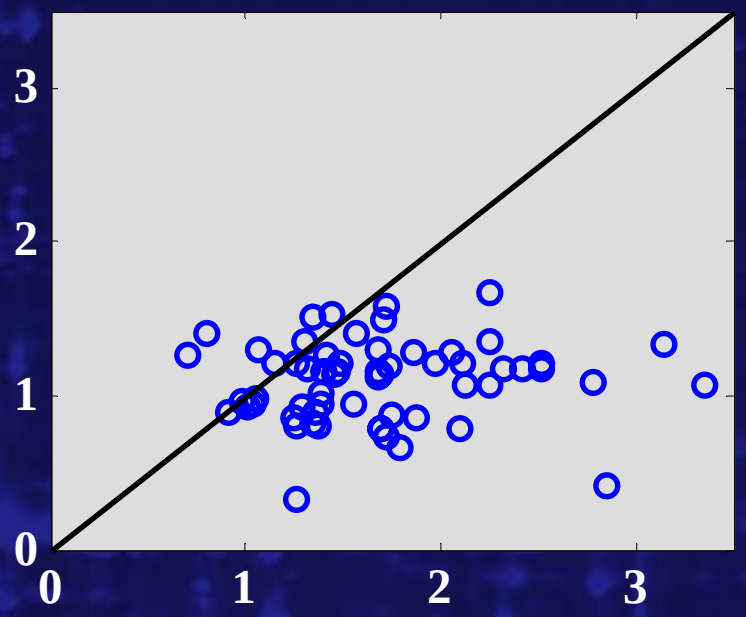
shift from 0 mM to 0.1 mM IPTG



X: $mRNA_i$ 0.1 / $mRNA_i$ 0.0

Y: $prot_i$ 0.1 / $prot_i$ 0.0

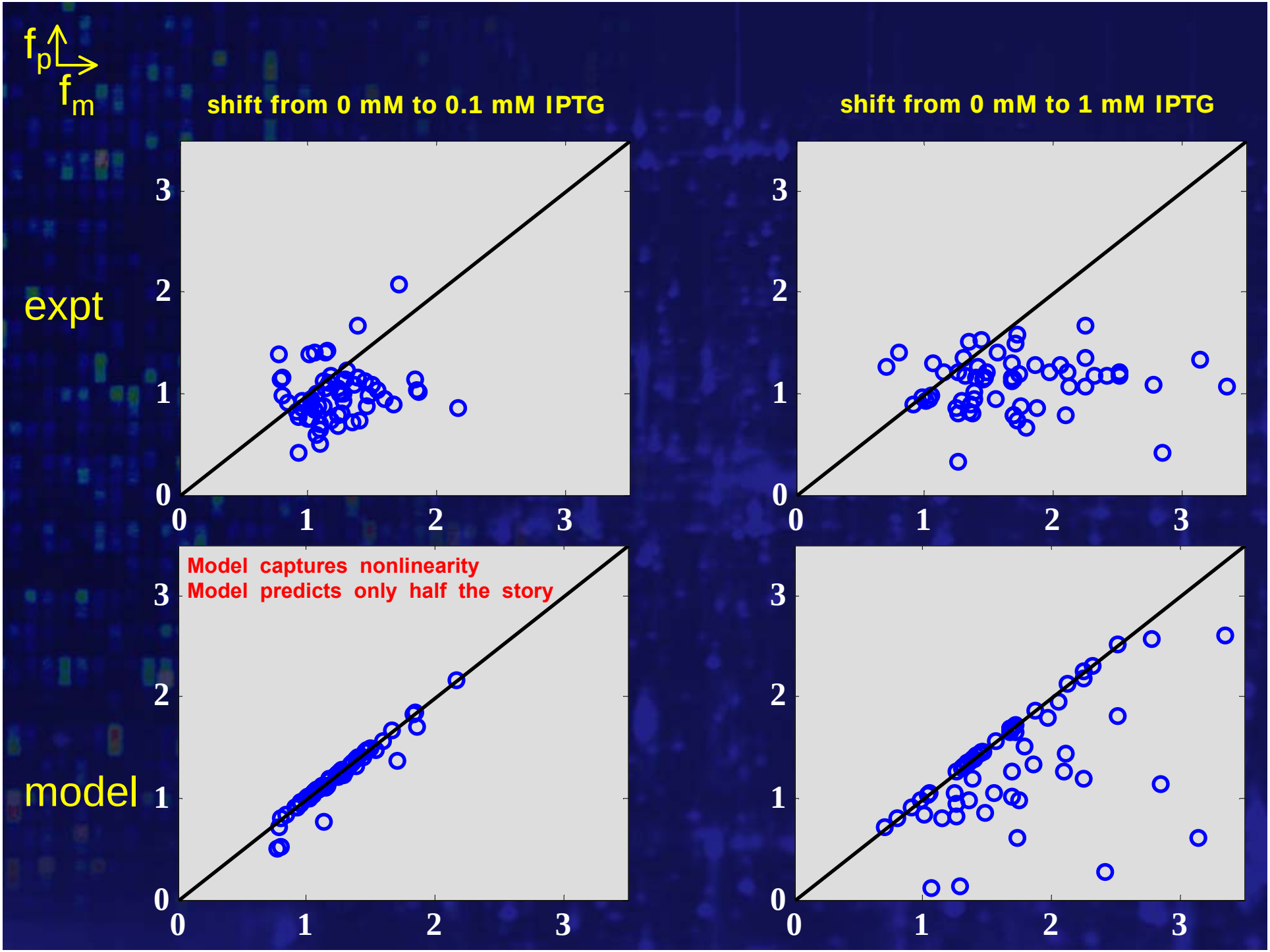
shift from 0 mM to 1 mM IPTG



X: $mRNA_i$ 1.0 / $mRNA_i$ 0.0

Y: $prot_i$ 1.0 / $prot_i$ 0.0

Relationship clearly nonlinear.


f_p
f_m
shift from 0 mM to 0.1 mM IPTG
shift from 0 mM to 1 mM IPTG
expt
model
Model captures nonlinearity
Model predicts only half the story
0
1
2
3

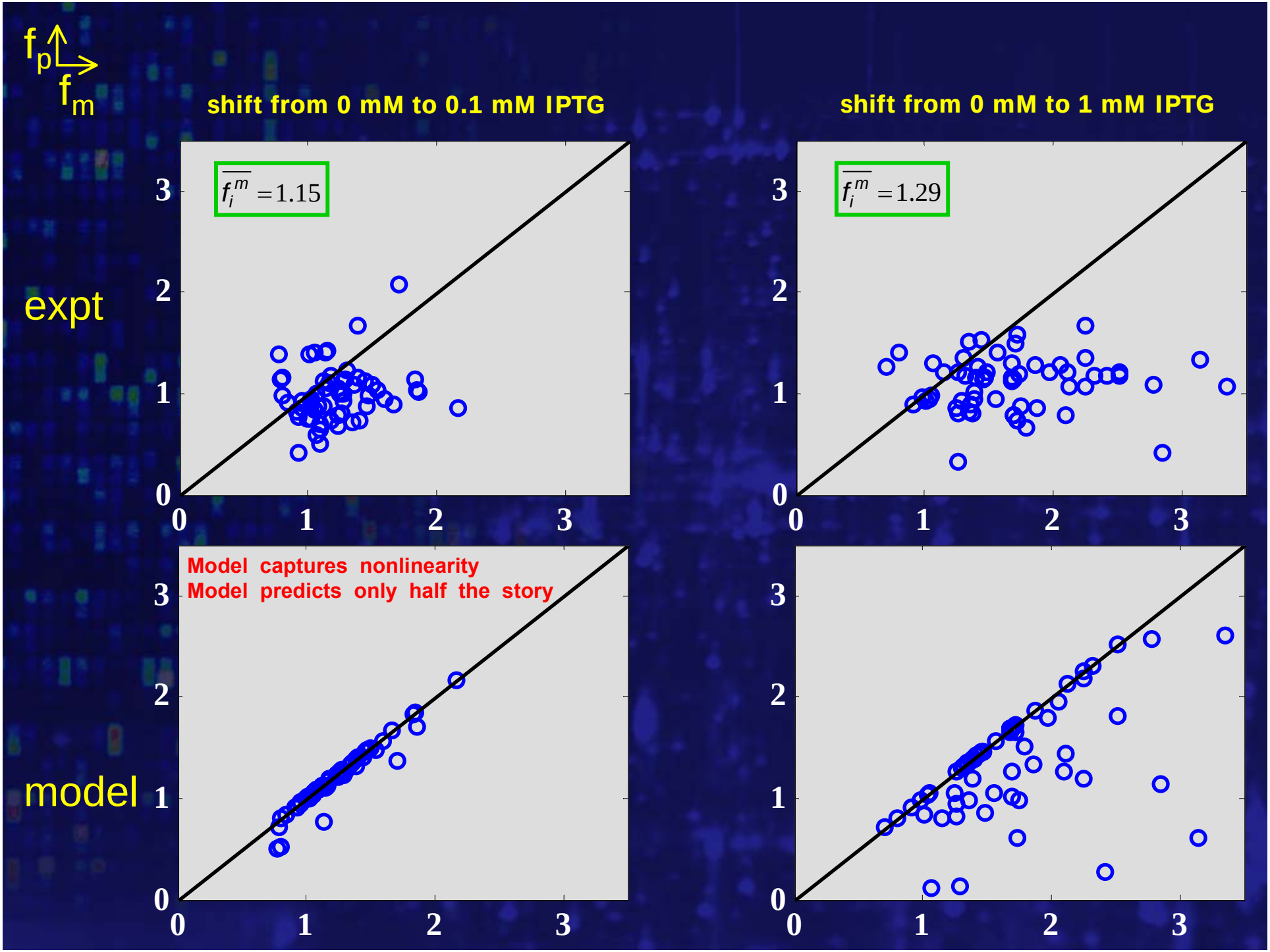
$$\bar{f}_i^m < 1 \Rightarrow r > r_o \Rightarrow f_i^p > f_i^m$$

$$\bar{f}_i^m > 1 \Rightarrow r < r_o \Rightarrow f_i^p < f_i^m$$

A perturbation (0.1 to 1) could lead to, e.g., increased $\bar{f}_i^m$ which model suggests by scaling arguments should result in a shift in the data further below the 1 to 1 line.

Scaling arguments suggest that if the total mRNA in the cell increases, then the ratio of any reasonable subset of observed data should reflect fewer proteins than mRNA - data migrates further below 1:1 line.

(Free ribosome availability becomes limiting).

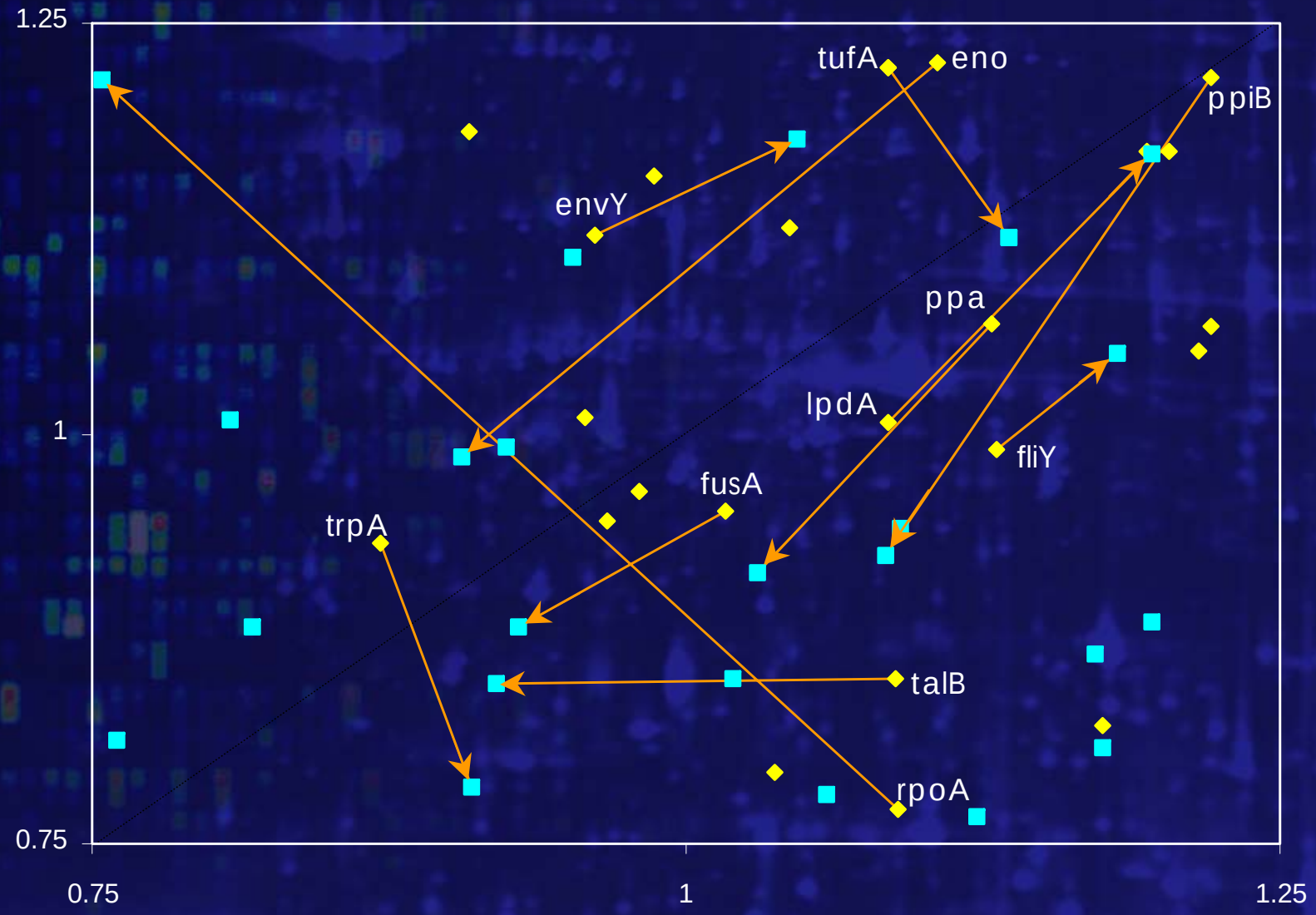

f_p
f_m
shift from 0 mM to 0.1 mM IPTG
shift from 0 mM to 1 mM IPTG
expt
model
$\overline{f_i^m} = 1.15$
$\overline{f_i^m} = 1.29$
Model captures nonlinearity
Model predicts only half the story
0
1
2
3

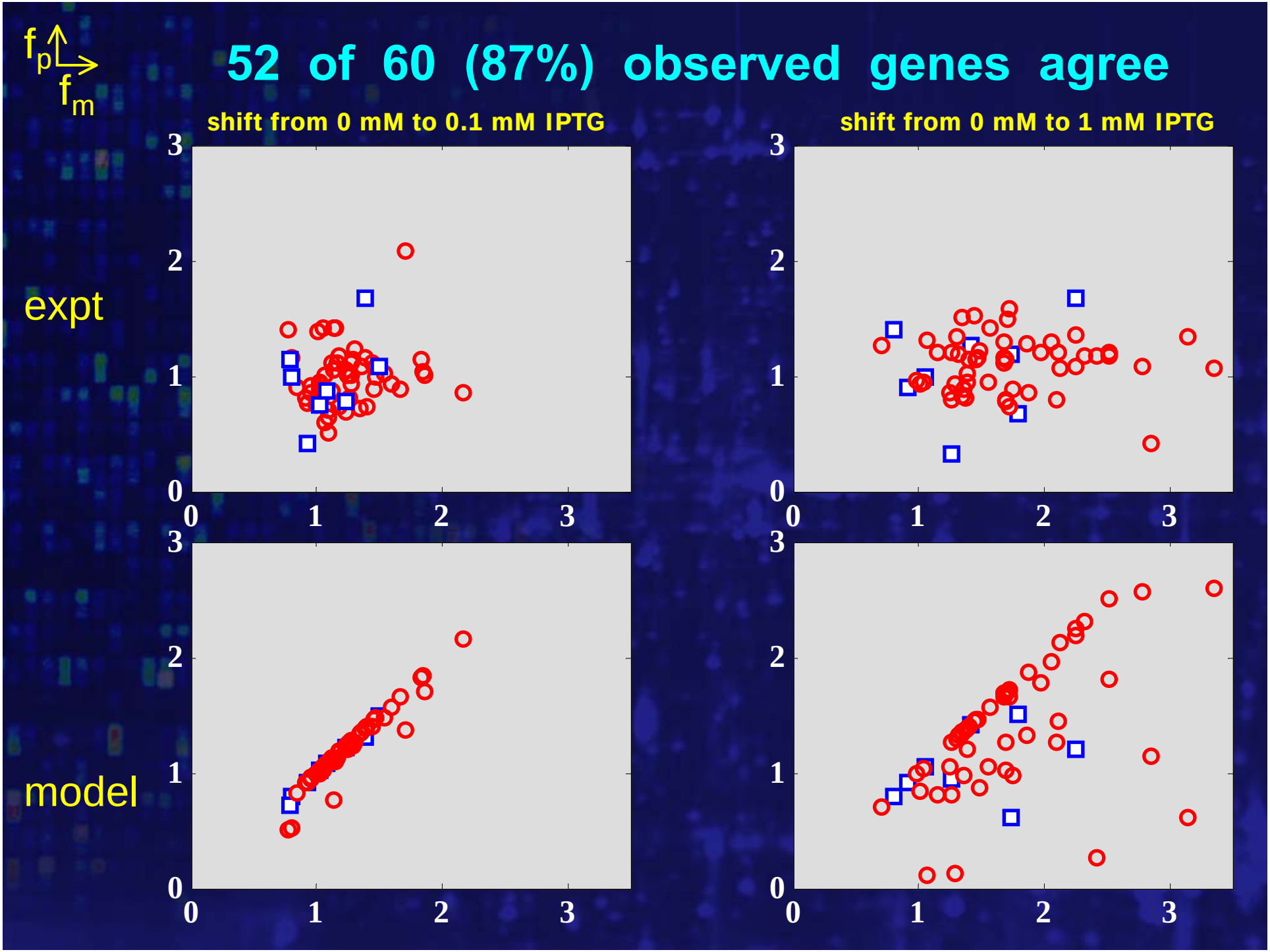
- The framework captures qualitatively the shift in the $f_p : f_m$ relationship when $\overline{f_i^m}$ increases.

- Can the framework be used to predict changes in gene expression for individual genes ?

Use the experimental data to identify the effective ribosome binding constants that capture the relative responses of the $f_p : f_m$ relationship in the experiments.

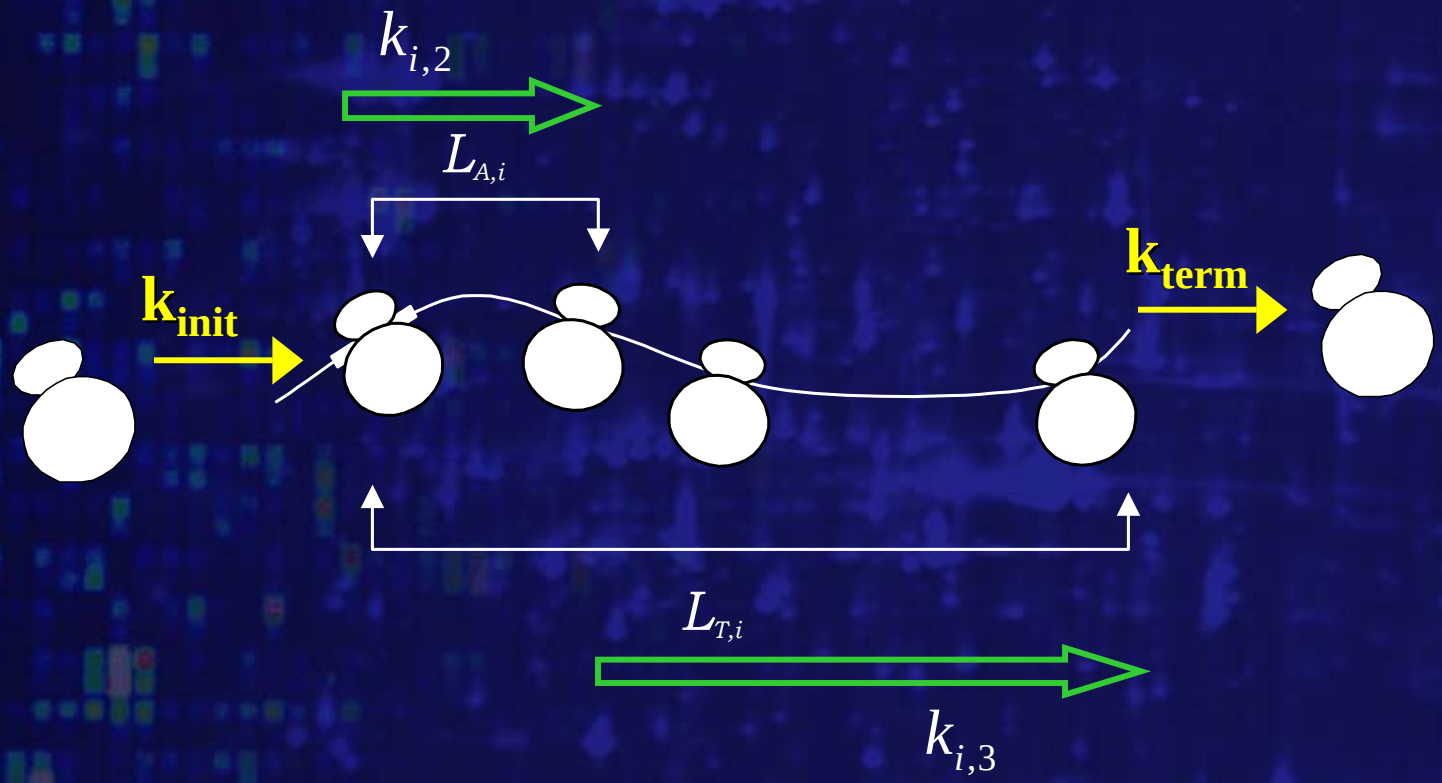
As culture conditions change (0.1mM/0mM to 1mM/0mM), the relative fp:fm ratio will shift.
Can the framework predict this shift for individual genes?




52 of 60 (87%) observed genes agree
shift from 0 mM to 0.1 mM IPTG
shift from 0 mM to 1 mM IPTG
expt
model
f_p
f_m
0
1
2
3

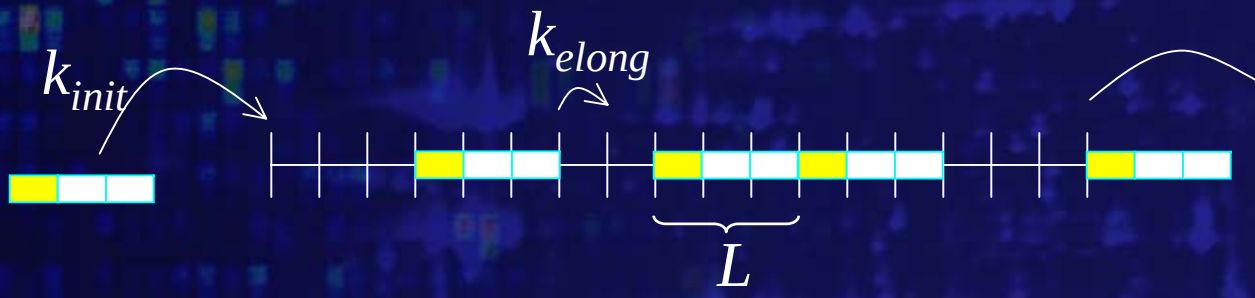
We just saw impact of ribosome competition and binding constants.

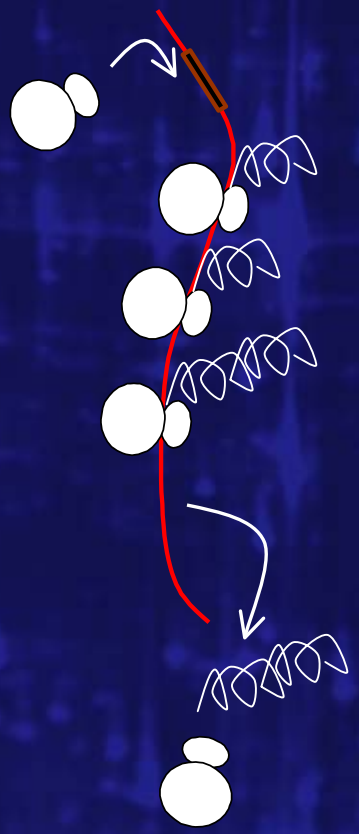
Translation Depends on the Codon Usage



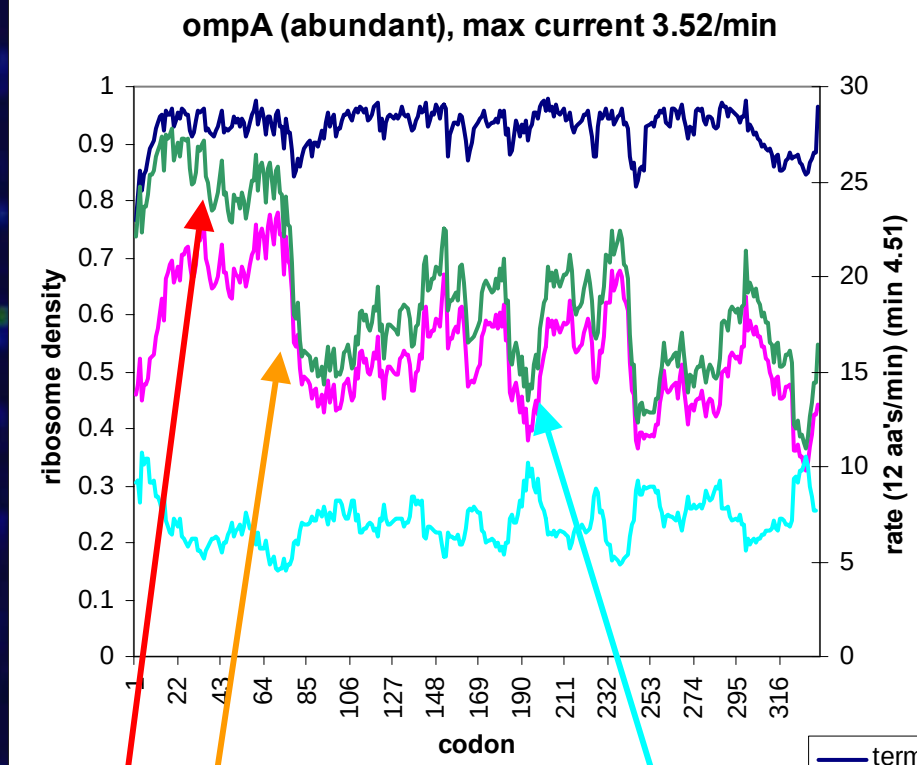
We know sequence / codon usage, [aa-tRNA].
Assume 100 copies compete for 4500 ribosomes.
Continuous time Monte Carlo simulations ...

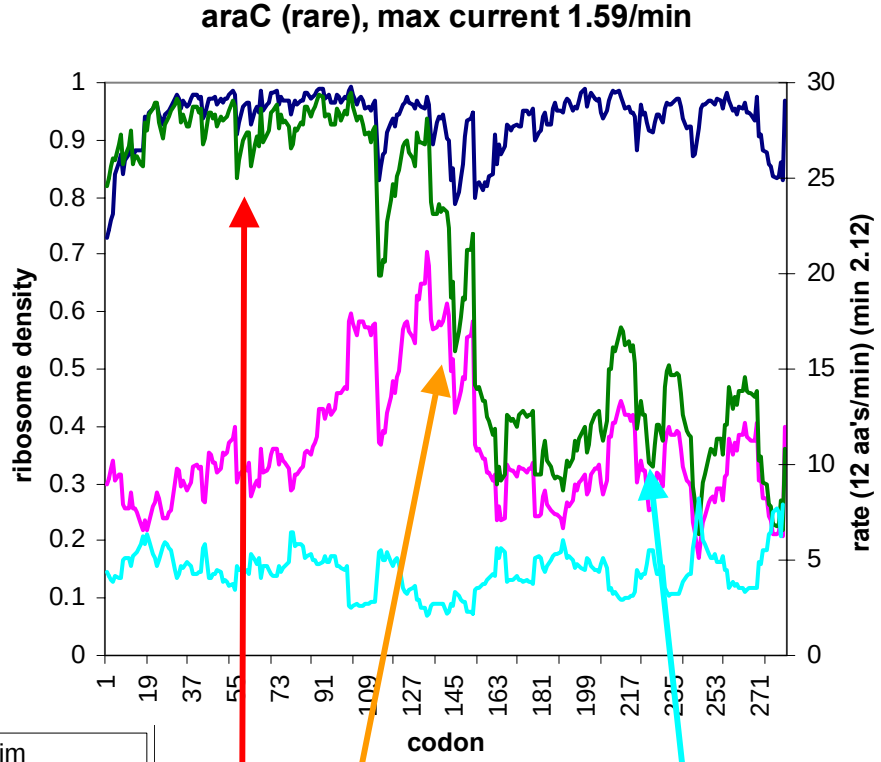
Single lattice model

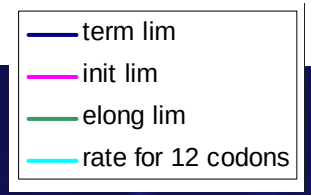




- mRNA modeled as a lattice
- ribosomes have steric hindrance
- parameters
  - ribosome length L (=12)
  - Initiation rate $k_{init}=k\times$[free ribosomes]
  - Elongation rates $k_{elong}=k\times$[free aa-tRNA]
  - termination rate k_{term}







Higher density of ribosomes

A slow site

Fewer ribosomes

Higher density

A slow site

Lower density

The mRNA - Protein Relationship Depends on the DNA sequence.

- Ribosome binding affinity
- Length of message
- Ribosome velocity
  - Codon usage and frequency
  - mRNA secondary structure