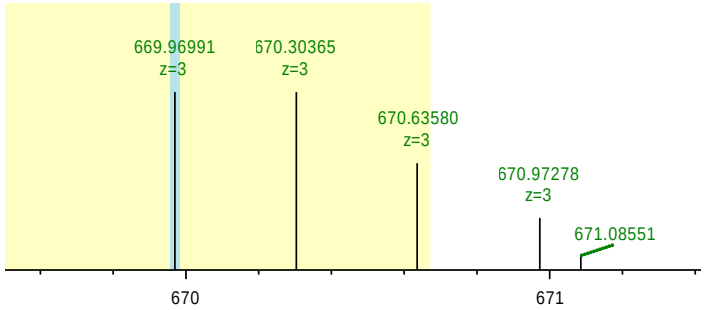
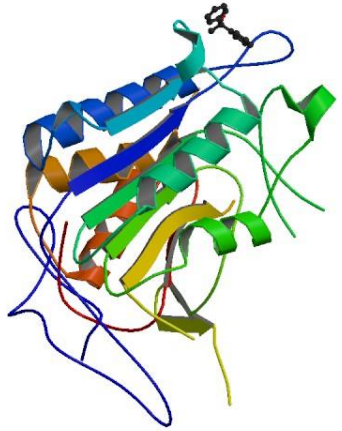
Production of omics data: Proteomics

Thibaut Léger, PhD

LERES, EHESP





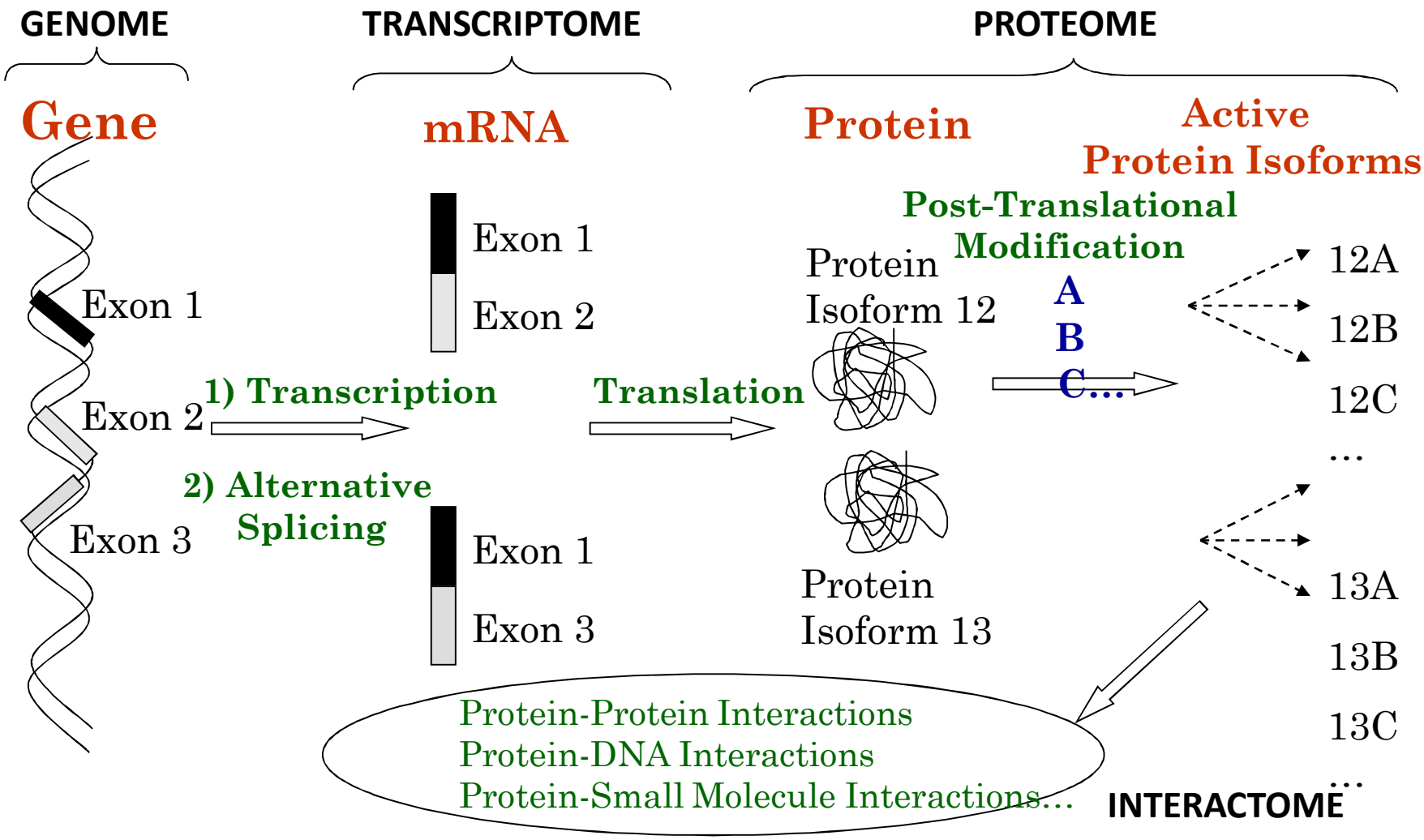
DUBii

10th march 2021

What is behind these data?

1													Normalized abundance							
2													Sain						Patient	
3	Accession	Peptide cou	Unique pep	Confidence	Anova (p)	q Value	Max fold ch	Power	Highest me	Lowest mea	Mass	Description	1845007-F1	1845007-F3	1845007-F5	1845007-F7	1845007-F9	1845007-F1	1845007-F2	1845007-F4
4	P40197	12	12	545.23	2.48E-07	4.74E-05	32.042418	1	Patient	Sain	60.921	Platelet glycoprotein V OS	1447.8857	898.47877	2114.249	3517.982	2506.9091	1186.7152	33178.605	73995.78
5	P02776	2	2	134.77	2.63E-07	4.74E-05	78.817355	1	Patient	Sain	10.838	Platelet factor 4 OS=Homc	3812.4369	3755.8358	1044.3911	2939.4867	3862.0883	1110.5549	333829.66	68071.51
6	Q13201	6	6	221.4	5.33E-07	6.40E-05	72.663027	1	Patient	Sain	138.023	Multimerin-1 OS=Homo sa	160.48528	939.93933	155.7663	732.57482	572.39757	196.60688	75519.303	23299.12
7	P04114	336	334	28302.18	7.99E-06	0.0006593	2.1384481	1	Patient	Sain	515.283	Apolipoprotein B-100 OS=	14117253	10601958	8472023.6	9982572.2	10056625	11898407	24415094	1744463
8	P07996	42	42	2678.71	9.16E-06	0.0006593	143.08362	1	Patient	Sain	129.3	Thrombospondin-1 OS=Hc	11604.367	10346.75	4059.3393	5652.5533	2121.1343	2367.1543	2303007.2	83175.08
9	Q15485	7	7	289.57	1.42E-05	0.0007087	24.427918	0.9999997	Patient	Sain	33.98	Ficolin-2 OS=Homo sapier	8778.176	6132.5206	10104.107	8194.6312	8690.2929	1737.9715	185819.2	105024.9
10	P10720	3	3	208.63	1.49E-05	0.0007087	36.552808	0.9999997	Patient	Sain	11.545	Platelet factor 4 variant O	32881.787	11758.269	1451.6395	5448.2604	4187.5362	4196.5766	546497.96	239272.0
11	Q12884	2	2	70.47	1.58E-05	0.0007087	Infinity	0.9999996	Patient	Sain	87.657	Prolyl endopeptidase FAP	0	0	0	0	0	0	961.55592	106.477
12	Q15061	1	1	3.67	3.71E-05	0.0014832	23.811179	0.9999905	Sain	Patient	74.843	WD repeat-containing prc	567513.12	751475.75	852125.74	1189728.8	625225.18	78135.037	46681.126	36585.06
13	P04075	8	8	371.85	4.58E-05	0.0016488	4.2937733	0.9999814	Patient	Sain	39.395	Fructose-bisphosphate al	33219.801	26234.171	28730.699	26696.908	33816.665	32024.353	128602.25	168571.3
14	P09486	10	9	463.23	6.25E-05	0.0020451	6.1746973	0.999953	Patient	Sain	34.61	SPARC OS=Homo sapiens (	9522.5101	37758.131	36721.019	20388.641	15117.423	28228.93	299004.24	97804.24
15	P10124	1	1	70.16	7.11E-05	0.002134	272.44302	0.9999321	Patient	Sain	17.641	Serglycin OS=Homo sapier	67.740323	0	958.26875	53.15204	160.31231	238.84011	105180.09	42246.24
16	P02775	6	6	388.98	8.21E-05	0.0022557	28.012451	0.9998996	Patient	Sain	13.885	Platelet basic protein OS=	61607.9	109352.83	105910.89	67216.894	60865.635	22752.188	3524260.5	1692399.5
17	Q9H1K0	2	2	12.94	8.77E-05	0.0022557	1.8919044	0.9998799	Patient	Sain	88.815	Rabenosyn-5 OS=Homo sa	871721.79	1194024.5	1103071.3	1121288.3	1020557.4	1361492.9	1506333.1	1757466.5
18	P35542	4	4	338.41	0.0001236	0.0029659	16.656374	0.9997122	Patient	Sain	14.737	Serum amyloid A-4 protein	192222.76	102064.74	9416.0169	8023.3014	46690.651	107524.48	1543185.2	1741696.5
19	P80188	6	6	223.58	0.0001847	0.0041561	9.1194461	0.9992657	Patient	Sain	22.574	Neutrophil gelatinase-as	2265.4484	1473.9176	6203.72	3676.0015	7827.3304	4603.3796	66633.817	26485.80
20	P02144	4	4	209.23	0.0002392	0.0050651	6.5032553	0.9987201	Patient	Sain	17.173	Myoglobin OS=Homo sapi	2754.3333	2223.1201	1035.6898	1813.5104	1462.164	5472.3528	17977.389	20813.25
21	P05067	5	5	179.76	0.000297	0.0059397	18.947351	0.9980148	Patient	Sain	86.888	Amyloid beta A4 protein C	408.20926	53.510036	338.118	494.28658	258.47482	516.31147	8561.0101	2161.775
22	Q9NPH3	7	7	223.36	0.0003417	0.0064734	2.9534965	0.9973949	Patient	Sain	65.377	Interleukin-1 receptor acc	19269.326	23550.392	17918.705	19071.463	29165.24	20448.453	74326.28	82743.30
23	P22352	7	7	322.73	0.0005632	0.0101356	2.3655577	0.9936247	Patient	Sain	25.537	Glutathione peroxidase 3	356078.09	500176.63	411199.92	354267.54	608490.2	620677.21	1139291.1	1214566.8
24	A0A075B6I0	1	1	30.88	0.0005981	0.0102519	15.884721	0.9929515	Sain	Patient	12.806	Immunoglobulin lambda	14017.761	25223.891	6493.2483	70640.84	15600.774	36585.067	4641.3232	1487.298
25	P05155	33	33	2513.9	0.0007389	0.0120898	5.4023431	0.9900925	Patient	Sain	55.119	Plasma protease C1 inhib	741421.19	592478.75	377555.54	465885.04	2331779.9	7791080.5	11109177	10511766
26	A0A0C4DH2	2	2	100.38	0.0013344	0.0208833	3.8865369	0.9766266	Sain	Patient	12.999	Immunoglobulin heavy va	50500.146	64330.439	38938.111	121804.39	27881.723	32180.678	24057.435	17302.863
27	P14780	6	6	207.75	0.0014988	0.0221418	62.066206	0.972746	Patient	Sain	78.408	Matrix metalloproteinase	794.97907	136.39171	262.24899	313.84328	0	555.64619	58928.664	9636.858
28	P02649	27	27	1941.9	0.0015378	0.0221418	2.4710215	0.971824	Patient	Sain	36.132	Apolipoprotein E OS=Hom	2552396.1	1536331.8	853676.88	1079240.3	1188651.7	1327575.2	2857655.7	2177433.1
29	Q3C1V8	1	1	14.2	0.0019728	0.0268284	5.551993	0.9615339	Patient	Sain	25.917	Brain-specific homeobox	119.76184	56.50442	16.821931	8.5075688	3888.8817	85.190432	6077.0842	3199.583
30	P02652	11	11	995.34	0.0020124	0.0268284	2.4014067	0.9606031	Patient	Sain	11.168	Apolipoprotein A-II OS=Hc	34056410	26771849	14632439	15663194	42268641	26890744	89512485	43236635
31	P08571	12	12	937.27	0.0025308	0.0320335	1.6971067	0.9485492	Sain	Patient	40.051	Monocyte differentiation	797874.17	434080.26	499432.28	477652.53	733837.28	418725.29	352767.33	318079.37
32	P26927	31	31	1222.76	0.0025808	0.0320335	3.3123536	0.9474046	Patient	Sain	80.268	Hepatocyte growth factc								380083.05
33	P02655	6	6	819.16	0.0030504	0.0361789	4.7206762	0.9368237	Patient	Sain	11.277	Apolipoprotein C-II OS=I								137623
34	P02671	8	8	369.69	0.0031158	0.0361789	191.8384	0.9353762	Patient	Sain	94.914	Fibrinogen alpha chain								107899
35	P02763	17	17	1378.36	0.0036234	0.0407586	2.8533872	0.924364	Sain	Patient	23.497	Alpha-1-acid glycoprote								2274357
36	Q14831	1	1	23.81	0.0038979	0.0425178	2.9788784	0.9185746	Sain	Patient	102.185	Metabotropic glutamate								31759.88
37	P02647	32	32	2367.65	0.0040205	0.0425648	2.2688435	0.9160271	Patient	Sain	30.759	Apolipoprotein A-I OS=H								27802701
38	P01876	21	21	1557.26	0.0044316	0.0455765	3.5420267	0.9076475	Sain	Patient	37.631	Ig alpha-1 chain C regio								5212718.5
39	Q9UNW1	8	8	227	0.005259	0.0525841	1.7900311	0.8915219	Sain	Patient	55.016	Multiple inositol polypl								30358.52
40	Q16853	2	2	81.52	0.0054293	0.0528194	6.1436713	0.8883207	Sain	Patient	84.568	Membrane primary ami								194.7495
41	P01833	8	8	219.66	0.006944	0.0656448	9.1555604	0.8614401	Sain	Patient	83.232	Polymeric immunoglobu								2483.330
42	P27169	28	27	2064.1	0.0071604	0.0656448	1.8486226	0.8578184	Patient	Sain	39.706	Serum paraoxonase/ary								6280114.8
43	P60174	1	1	28.75	0.0074	0.0656448	2.9166777	0.8538651	Patient	Sain	30.772	Triosephosphate isome								4675.589
44	P00736	41	39	3067.37	0.0074771	0.0656448	1.4866701	0.8526056	Sain	Patient	80.067	Complement C1r subcon								2638367.5
45	P09871	37	37	2588.95	0.0085558	0.0726349	1.93603	0.835606	Sain	Patient	76.635	Complement C1s subcor								2302798.8
46	A0A0B4J1V0	3	3	104.52	0.0086769	0.0726349	2.7965659	0.833766	Sain	Patient	12.917	Immunoglobulin heavy va	92353.063	20397.701	46790.32	87259.628	45135.867	81618.911	38454.812	17117.34
47	P01619	4	4	262.21	0.0095577	0.0781901	2.30903	0.8207663	Sain	Patient	12.549	Immunoglobulin kappa va	409287.43	332471.99	228180.64	675002.86	220569.52	187876.28	139130.56	148976.3
48	Q08380	21	21	1485.84	0.0115602	0.0900161	2.528023	0.7934679	Sain	Patient	65.289	Galectin-3-binding protei	938408.99	614190.09	1091308.6	2622296.8	1676638.8	2590031.7	640186.94	906404.5
49	P06732	9	9	314.24	0.0117011	0.0900161	5.8938011	0.7916528	Patient	Sain	43.074	Creatine kinase M-type OS	7288.867	9691.0748	1004.0051	6516.3248	4494.7392	6004.4549	19154.379	89665.94

CONCEPT


GENOME
TRANSCRIPTOME
PROTEOME
Gene
mRNA
Protein
Active Protein Isoforms
Post-Translational Modification
Exon 1
Exon 2
Exon 3
1) Transcription
2) Alternative Splicing
Translation
Exon 1
Exon 2
Exon 1
Exon 3
Protein Isoform 12
Protein Isoform 13
A
B
C…
12A
12B
12C
…
13A
13B
13C
…
Protein-Protein Interactions
Protein-DNA Interactions
Protein-Small Molecule Interactions…
INTERACTOME

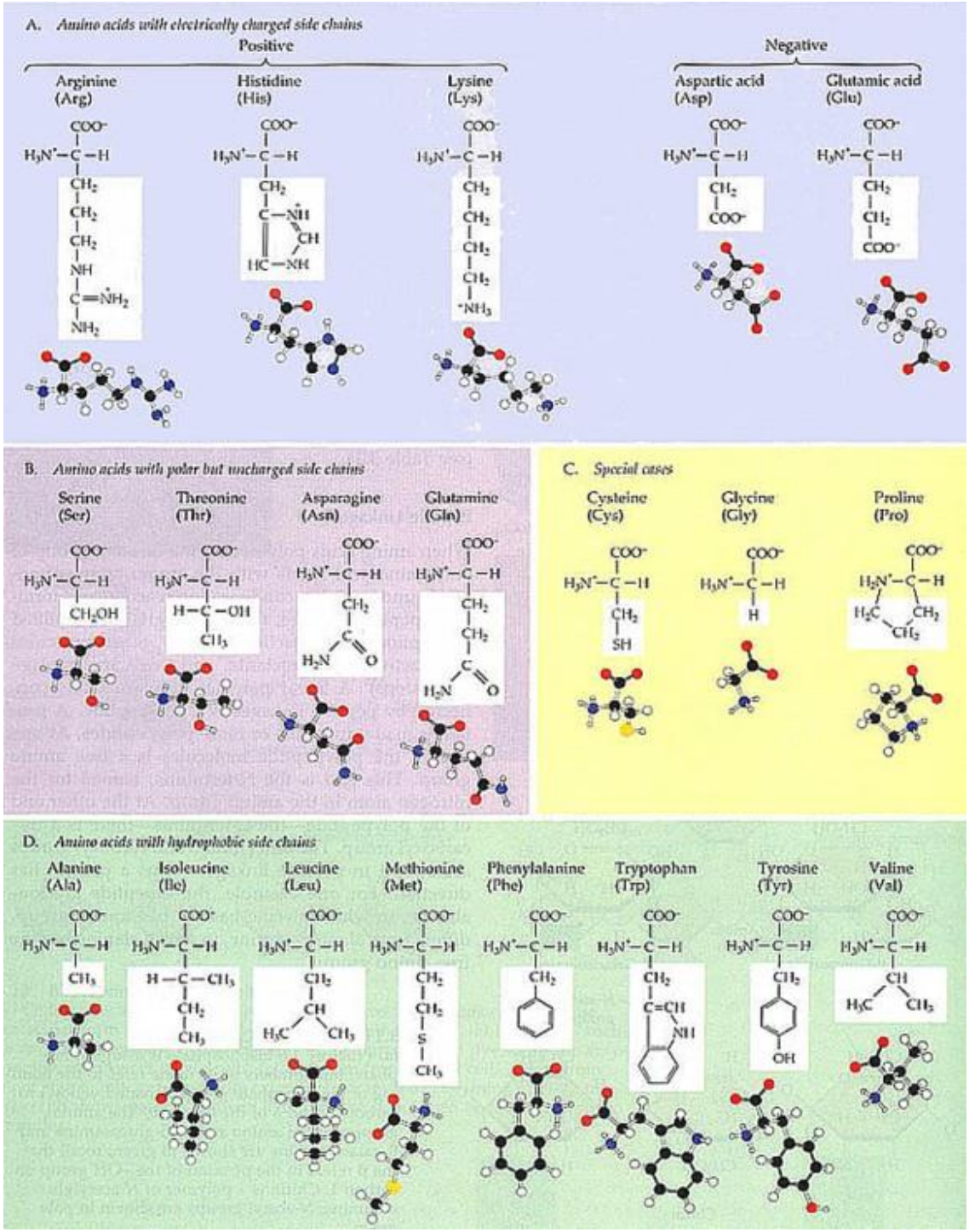
GENOMICS vs PROTEOMICS

Genome (DNA)

- Static (no change with time)
- Can be amplified (PCR)
- Little sample complexity
- *(4 base pairs, very similar, same order of concentration)*
- Good solubility

Proteome (proteins)

- Dynamic
- *(highly variable with time; many proteomes for one genome)*
- Cannot be amplified
- High sample complexity (wide variety of physical and chemical properties; concentrations can differ by 9 orders of magnitude)
- Various solubility; some proteins are insoluble in water



Alanine	A, Ala	71.079
Arginine	R, Arg	156.188
Asparagine	N, Asn	114.104
Aspartic acid	D, Asp	115.089
Cysteine	C, Cys	103.145
Glutamine	Q, Gln	128.131
Glutamic acid	E, Glu	129.116
Glycine	G, Gly	57.052
Histidine	H, His	137.141
Isoleucine	I, Ile	113.160
Leucine	L, Leu	113.160
Lysine	K, Lys	128.17
Methionine	M, Met	131.199
Phenylalanine	F, Phe	147.177
Proline	P, Pro	97.117
Serine	S, Ser	87.078
Threonine	T, Thr	101.105
Tryptophan	W, Trp	186.213
Tyrosine	Y, Tyr	163.176
Valine	V, Val	99.133

Proteomics?

- Proteomics is the large-scale study of proteomes, it means all proteins from a cell, an organelle, a tissue, an organ or from an organism at a one point, under specific conditions.

- Proteomics is at the crossroads of biochemistry, analytical chemistry and bioinformatics.

⇒Proteins can be modified by different biological or chemical processes; The different variants of proteins are called now:

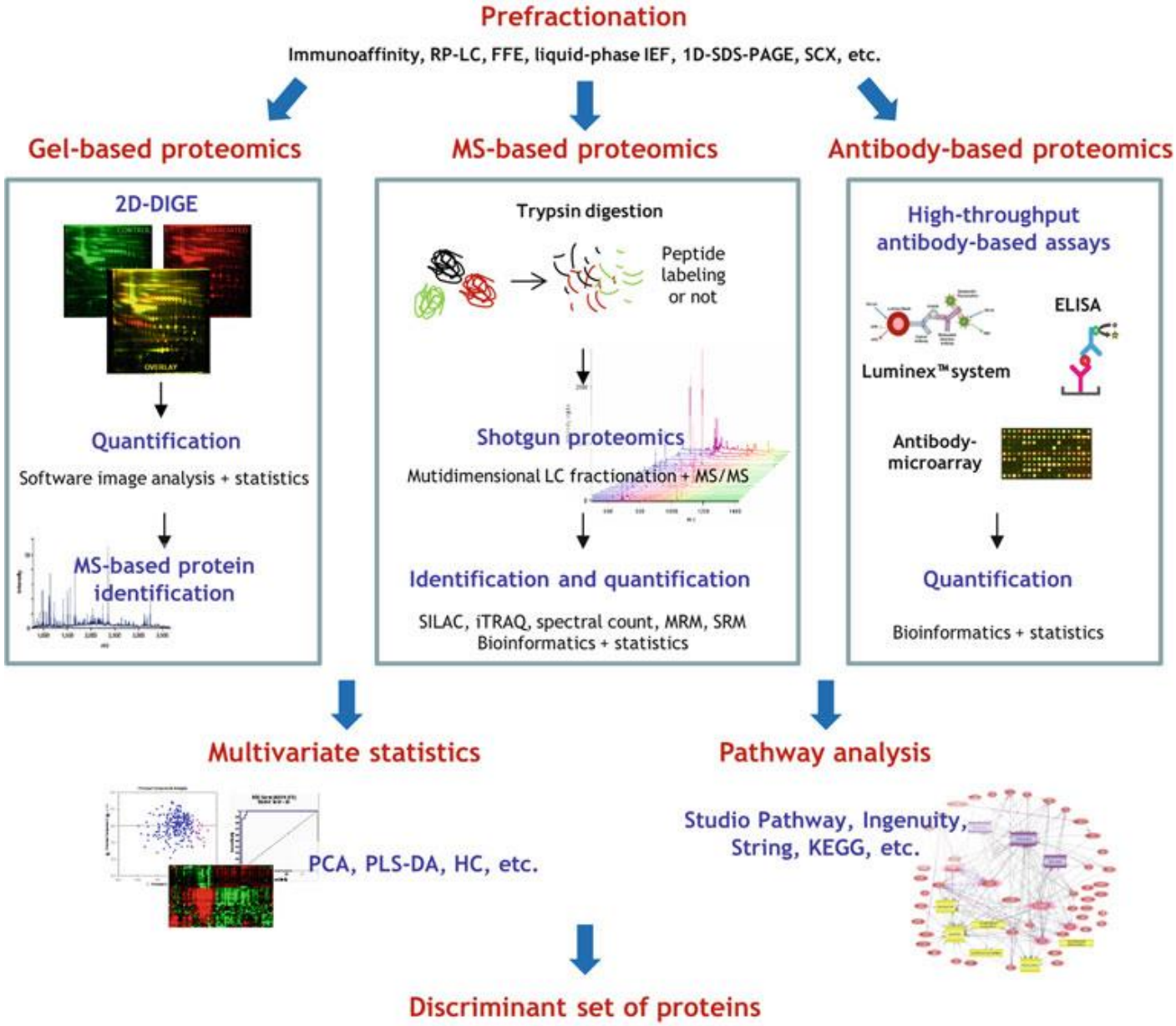
Proteoforms

Nat Methods. 2013 Mar;10(3):186-7. doi: 10.1038/nmeth.2369.
Proteoform: a single term describing protein complexity.
Smith LM, Kelleher NL; Consortium for Top Down Proteomics

PROTEOMICS GOALS

- Identification of all proteins in a proteome
- Search for new, hypothetical or predicted proteins
- Analysis of differential expression between 2,3,... different conditions (protein up- or downregulation)
- Identification of post-translational modifications
- Characterization of proteins by function, pathway, cellular location, etc.
- Study of protein-protein interactions

Proteomics techniques



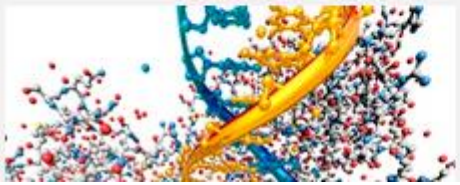
HUMAN PROTEOME MAP

Statistics

Organs/cell types	30
Genes identified	17,294
Proteins identified	30,057
Peptide sequences	293,700
N-terminal peptides	4,297
Splice junctional peptides	66,947
Samples	85
Adult tissues	17
Fetal tissues	7
Cell types	6

Welcome to ProteomicsDB!

ProteomicsDB is a joint effort of the Technische Universität München (
human proteome and its use across the scientific community.



Browse proteins

Explore the human proteome protein by protein.

Status

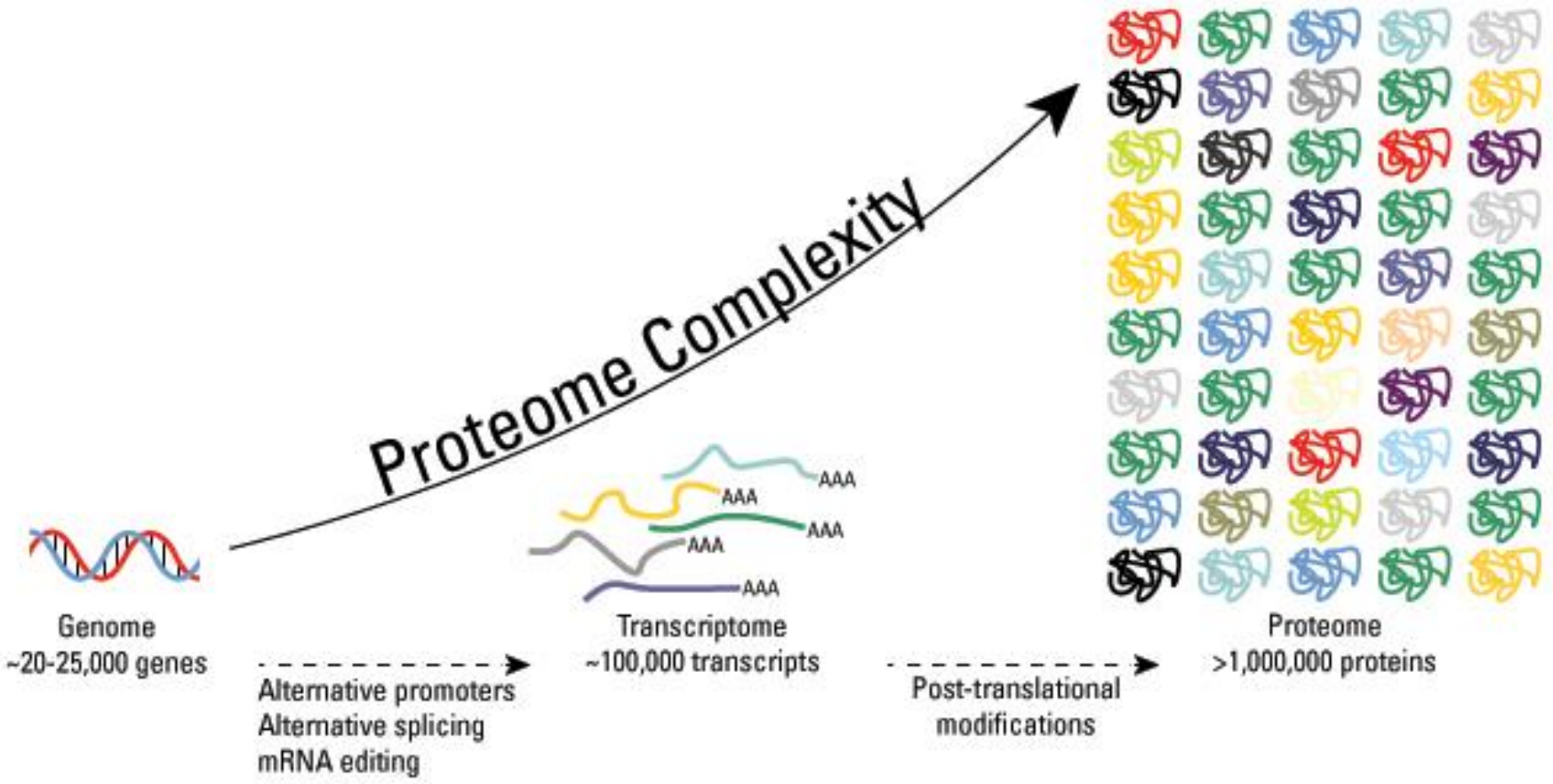
Human Proteome

Coverage:	80%
Proteins:	15721 of 19629
Isoforms:	11353 of 86771
Unique Peptides (Isoform):	113944
Unique Peptides (Gene):	455289
Spectra:	43237800

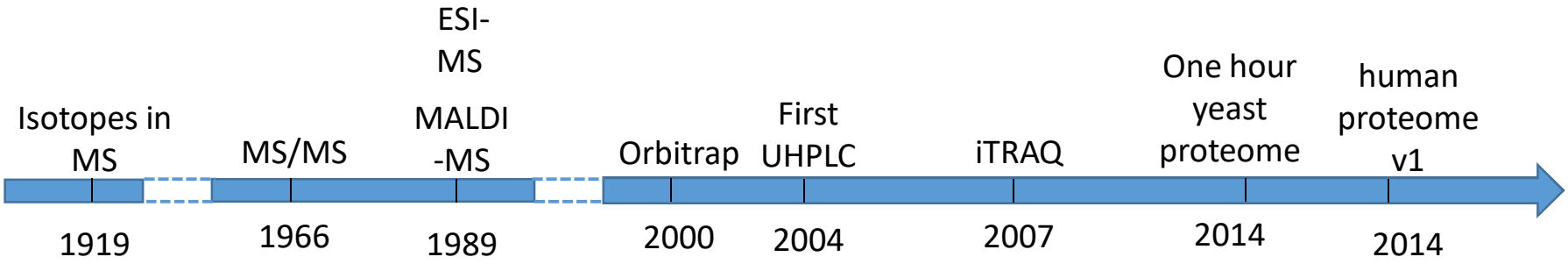
Repository

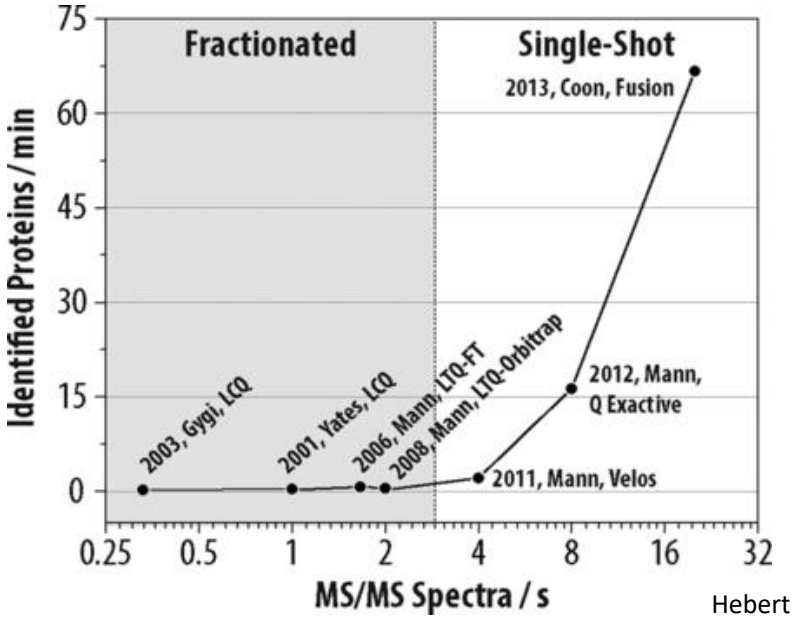
Registered Users:	533
Projects:	75
Experiments:	397
Files:	19459
Data Volume:	7.84 TB

Problem of proteome complexity



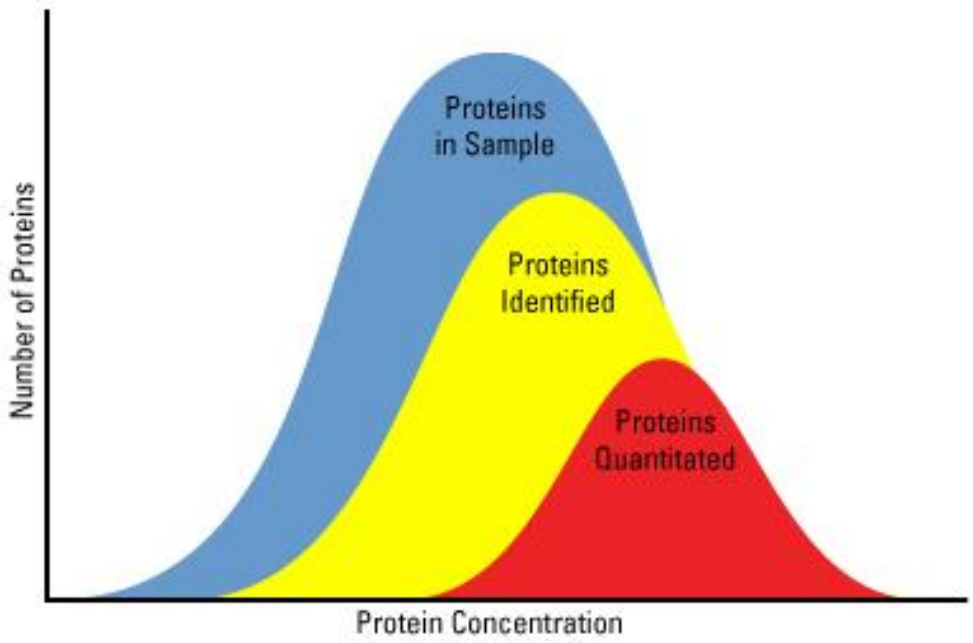
Evolution of proteomics performances

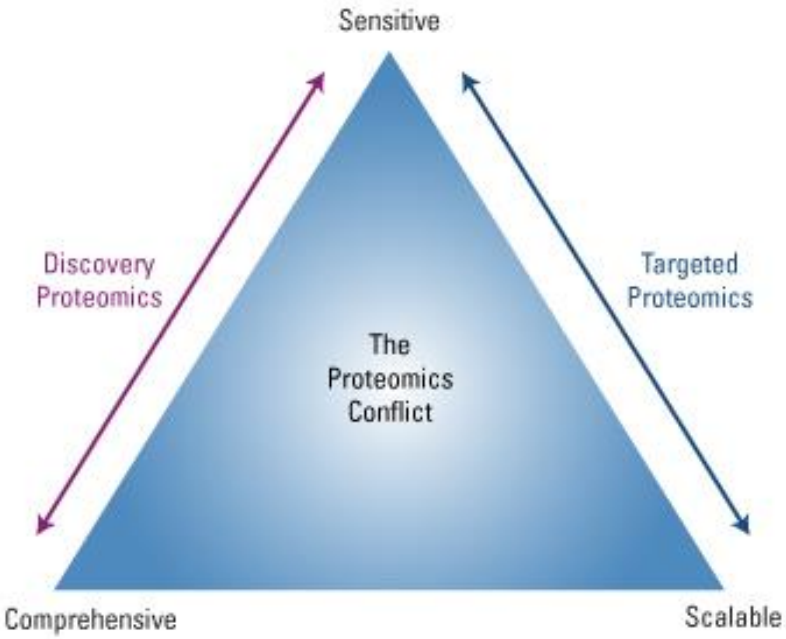




Hebert *et al.* 2009

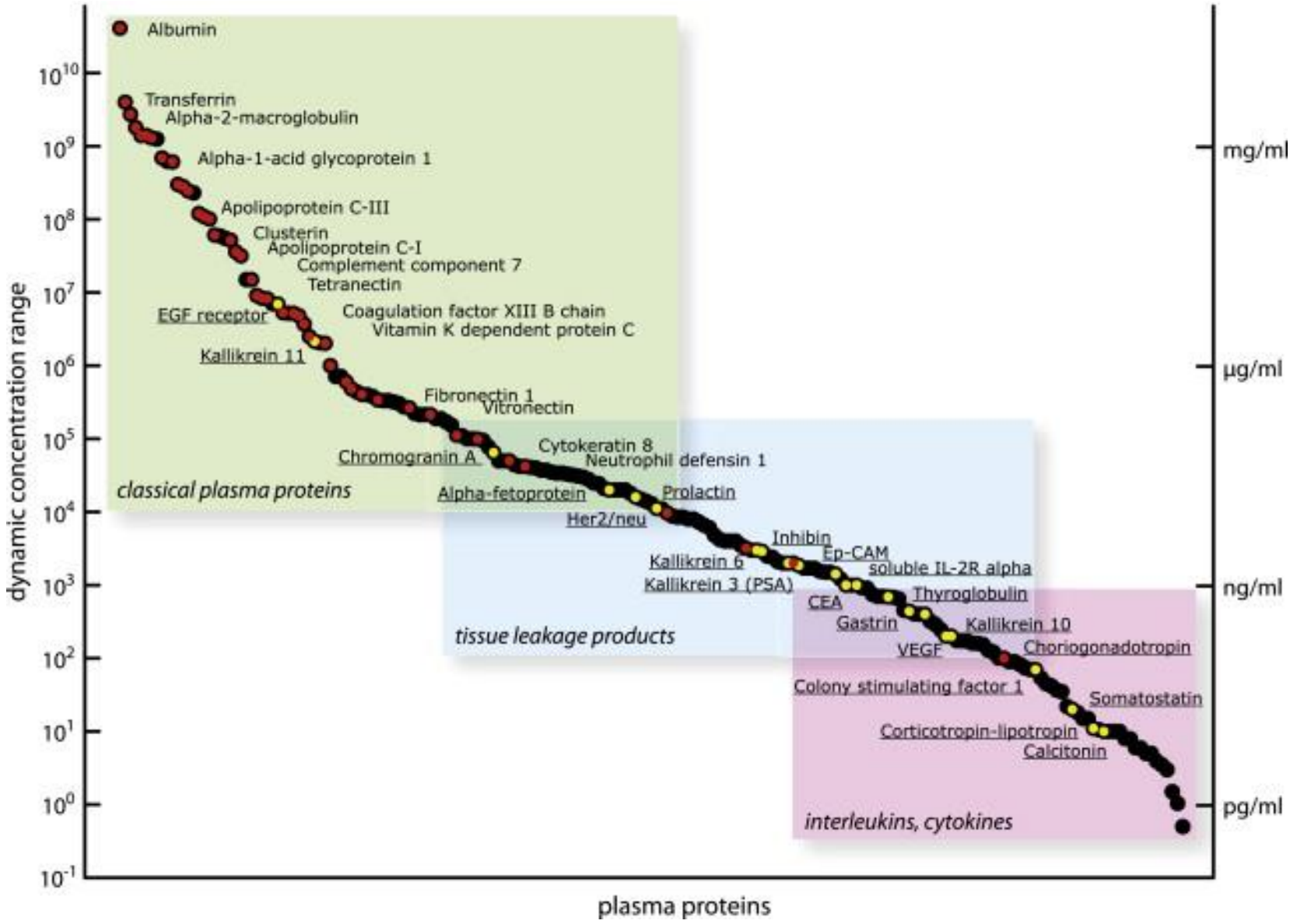
Inherent dilemma linked to proteomics



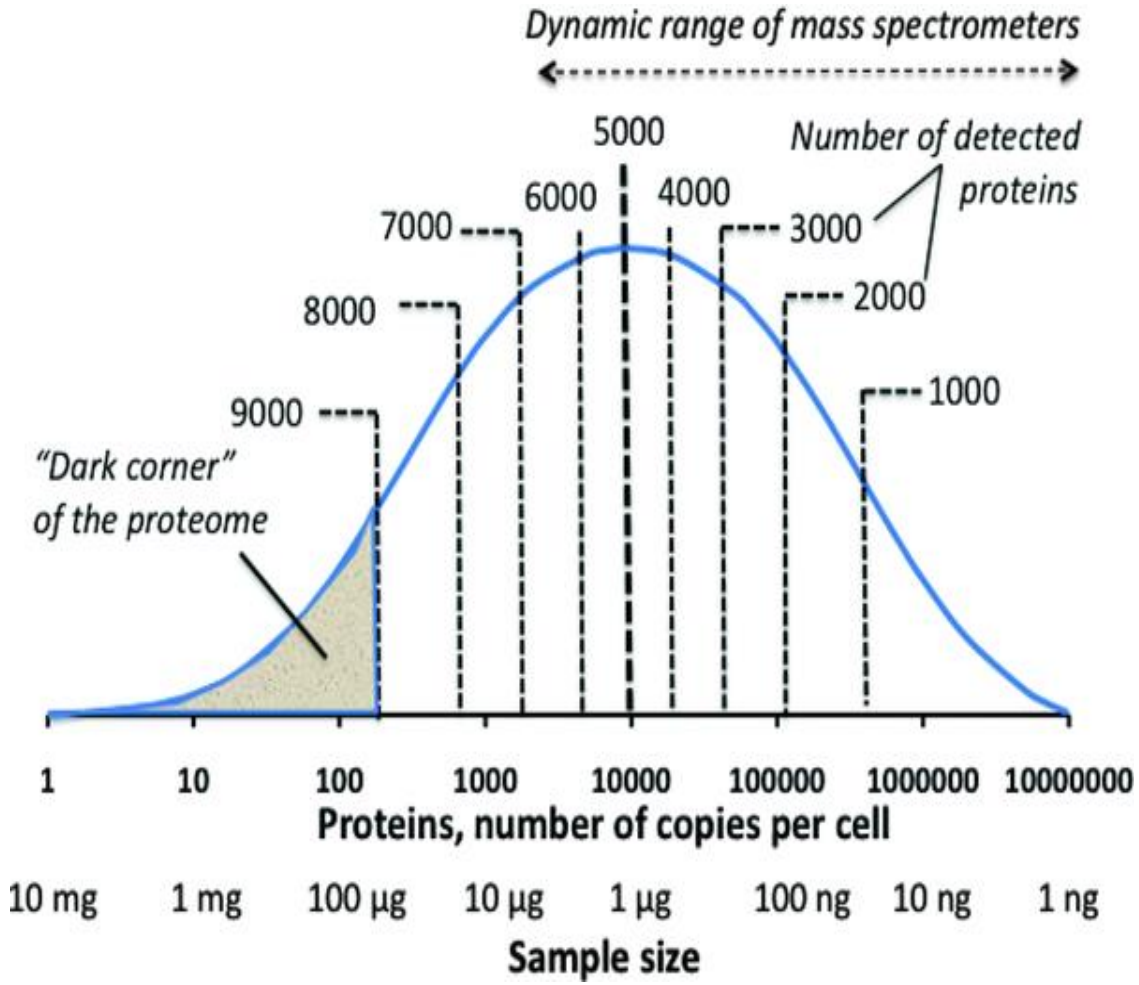


http://www.piercenet.com/method/quantitative-proteomics

Dynamic range in human plasma



Proteomics and proteome coverage



Zubarev *et al.* 2013

Dynamic range of proteomes

Key questions in proteomics

- What is the protein content of my biological sample?

  => problem of identification

- What is the abundance of my protein of interest?

  => quantification

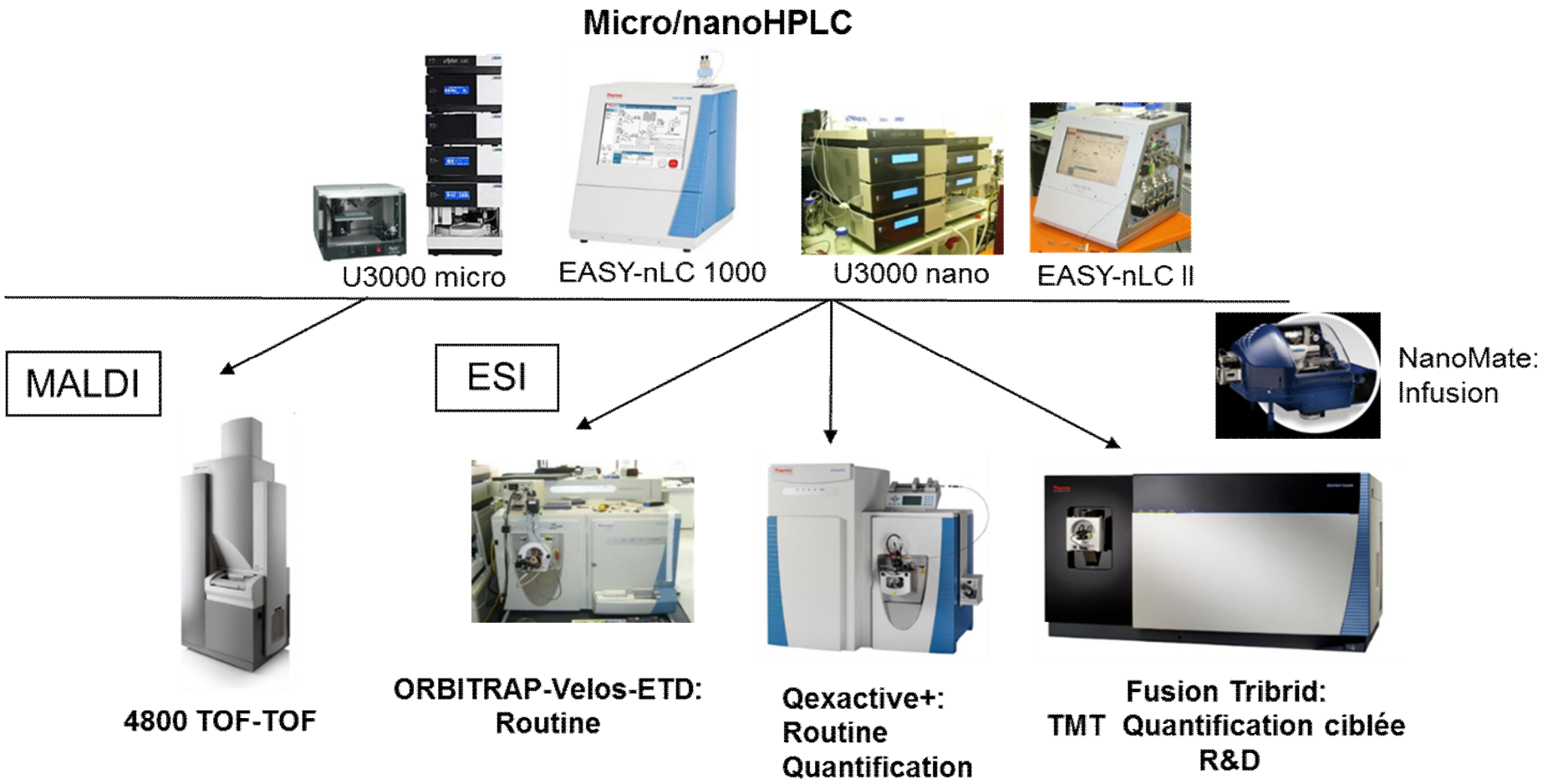
- Relative question: What are the protein abundance variations of the proteomes studied?

- What are the partners of my protein of interest?

- Are there any signature proteins related to a particular biological process?

  => biomarkers identifications and quantifications

Instrumentations



Proteomics workflows

Top ï Down

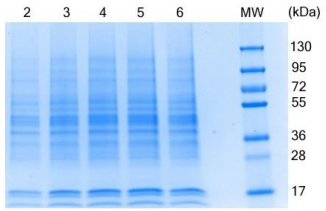
(< 50000 Da in particular conditions)

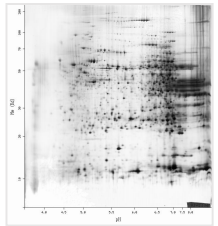
Proteins

Peptides

HPLC

PROTEASES



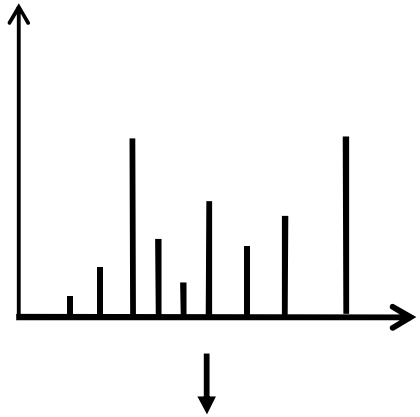


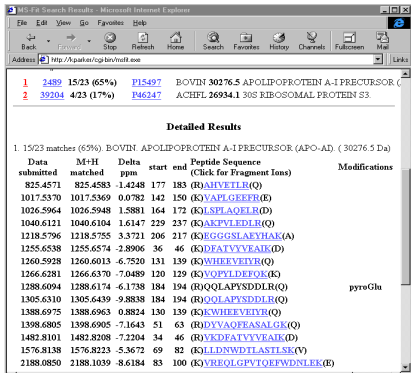
Bottom ï Up

(< 4000 Da)

MS acquisition

- MALDI TOF/TOF
- Orbitrap





Identification
Quantification

The strategy is dependent of the sample complexity

Search engine (Mascot, Sequest, Peaks, Maxquant, OMSSA, Prosight, Byonik)

BOTTOM-UP PROTEOMICS: PRO'S AND CON'S

Advantages

- Less sophisticated instrumentation and expertise
- High throughput
- More info about proteins with "extreme" phys.-chem. properties (hydrophobic, Hi/Low MW, acidic/basic)

Disadvantages

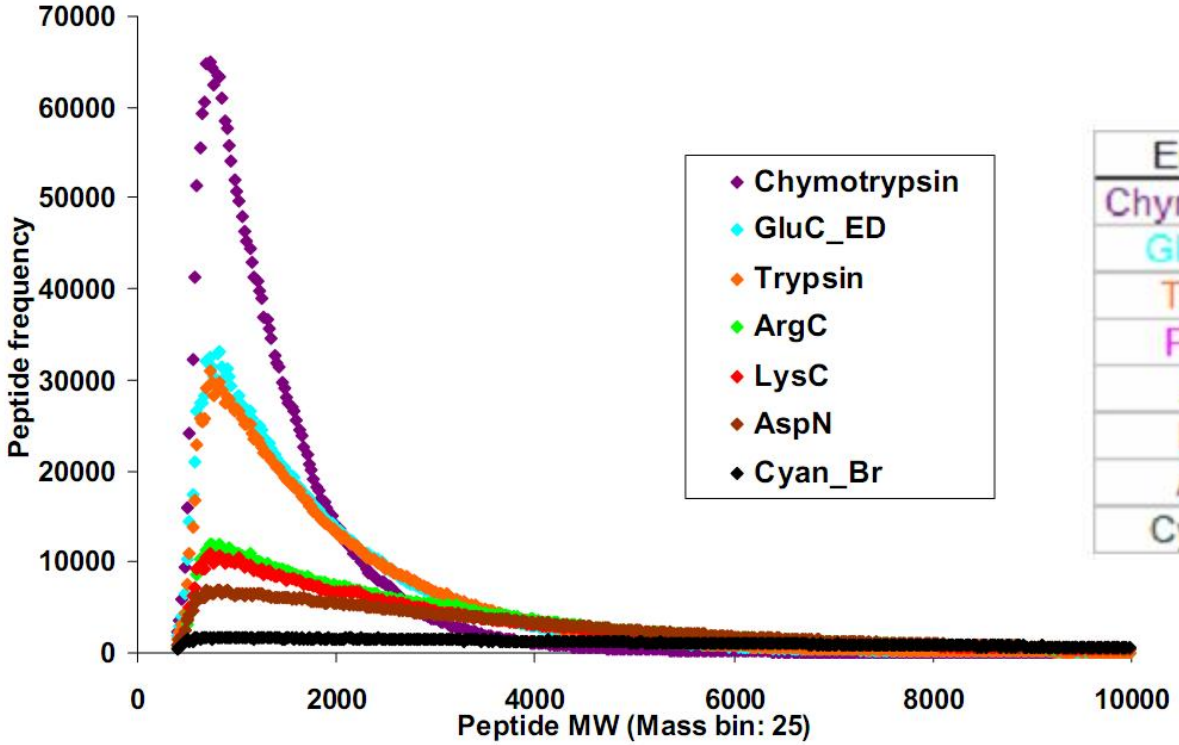
- Confidence in protein ID strongly depends on restriction criteria (subjective; potential bias)
- Since protein ID is often done by 1-2 peptides, PTM and isoform information is often lost

Cleavage rules of proteases

Enzyme or Reagent	Cleaves where?	Exceptions
Trypsin	C-terminal side of K or R	if P is C-term to K or R
Trypsin	(C-term to K/R, even before P)	C-terminal side of K or R
Trypsin (higher specificity)	C-terminal side of K or R	if P is C-term to K or R; after K in CKY, DKD, CKH, CKD, KKR; after R in RRH, RRR, CRK, DRD, RRF, KRR
Lys C	C-terminal side of K	
CNBr	C-terminal side of M	
Arg C	C-terminal side of R	if P is C-term to R
Asp N	N-terminal side of D	
Asp N + N-terminal Glu	N-terminal side of D or E	
Glu C (bicarbonate)	C-terminal side of E	if P is C-term to E, or if E is C-term to E
Glu C (phosphate)	C-terminal side of D or E	if P is C-term to D or E, or if E is C-term to D or E
Chymotrypsin	(C-term to F/Y/W/M/L, not before P, not after Y if P is C-term to Y)	
	C-terminal side of F, L, M, W, Y	if P is C-term to F, L, M, W, Y, if P is N-term to Y
Chymotrypsin (C-term to F/Y/W/, not before P, not after Y if P is C-term to Y)		
	C-terminal side of F, Y, W	if P is C-term to F, Y, W, if P is N-term to Y
Trypsin/Chymotrypsin (C-term to K/R/F/Y/W, not before P, not after Y if P is C-term to Y)		
	C-terminal side of K, R, F, Y, W	if P is C-term to K, R, F, Y, W, if P is N-term to Y
Pepsin (pH 1.3)	C-terminal side of F, L	
Pepsin (pH > 2)	C-terminal side of F, L, W, Y, A, E, Q	
Proteinase K	C-terminal side of A, C, G, M, F, S, Y, W	

PEPTIDE LENGTH AND NUMBER OF PEPTIDES GENERATED DEPENDING ON ENZYME USED FOR DIGESTION

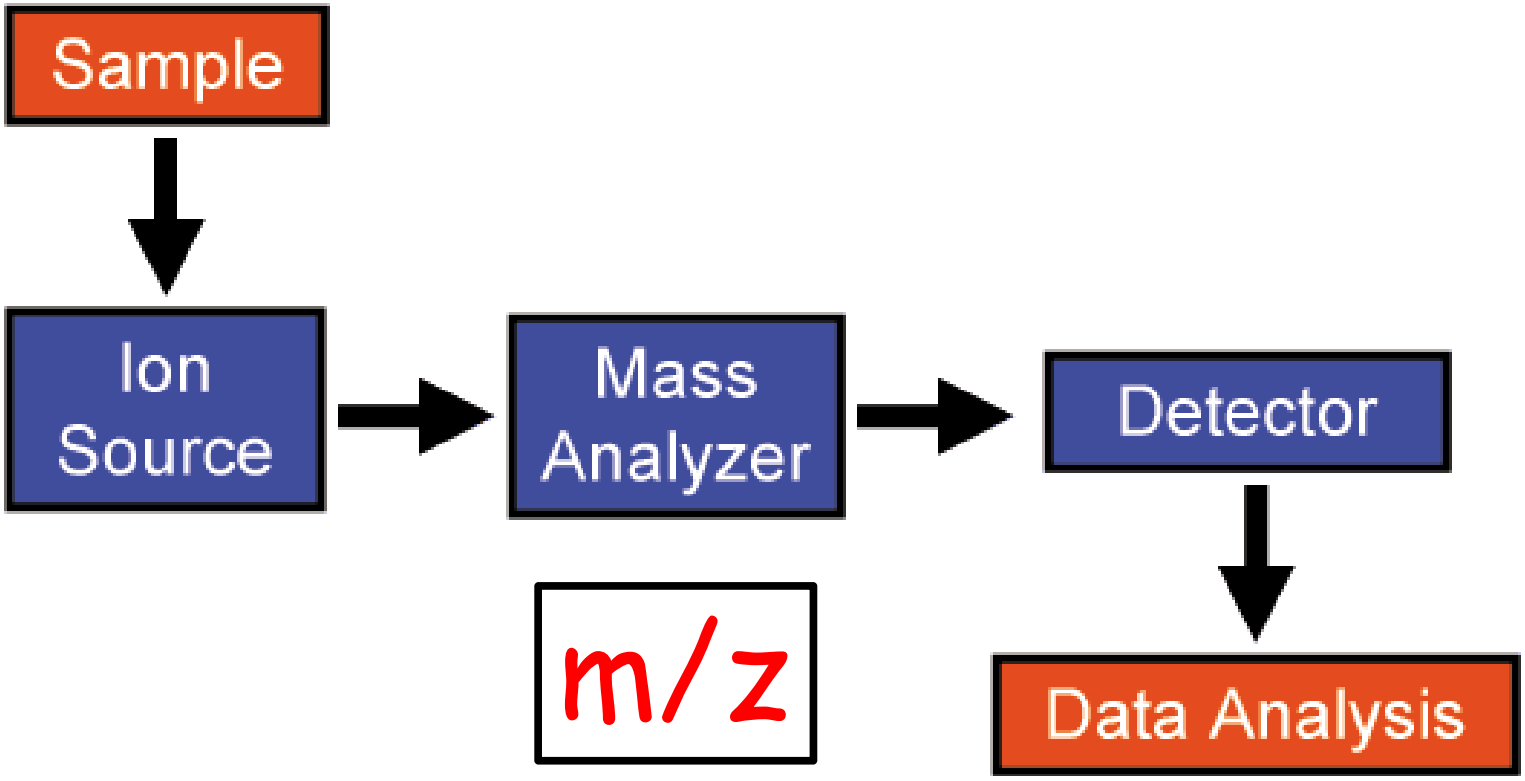
Other enzymes with more or less specific cleavage:



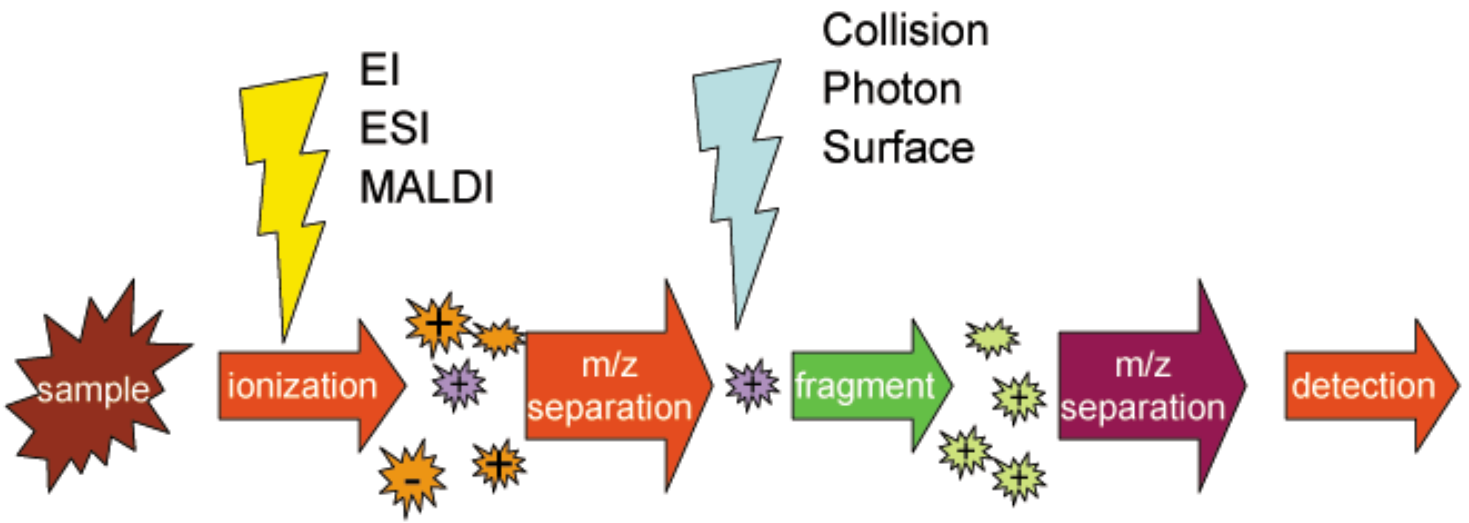
Enzyme	Cleaves at:
Chymotrypsin	FWYL
GluC_ED	ED
Trypsin	KR
Pepsin	FLIWY
ArgC	R
LysC	K
AspN	D
Cyan_Br	M

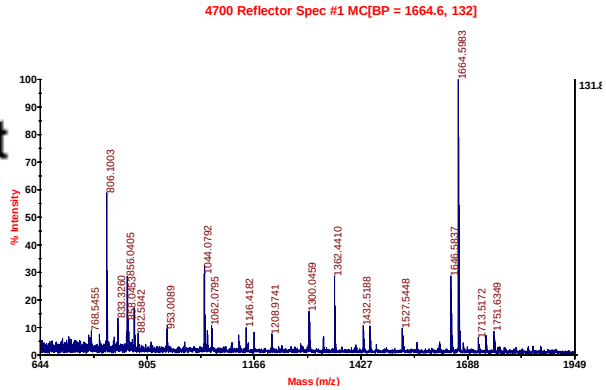
Advantages of a new proteomic approach that uses accurate mass measurements, LC retention time, isoelectric point and dual enzymatic digestion. Petritis K. et. al., Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352; ASMS'2007 poster presentation http://www.chem.agilent.com/Library/posters/Public/Petritis_ASMS_2007.pdf

What is MS?

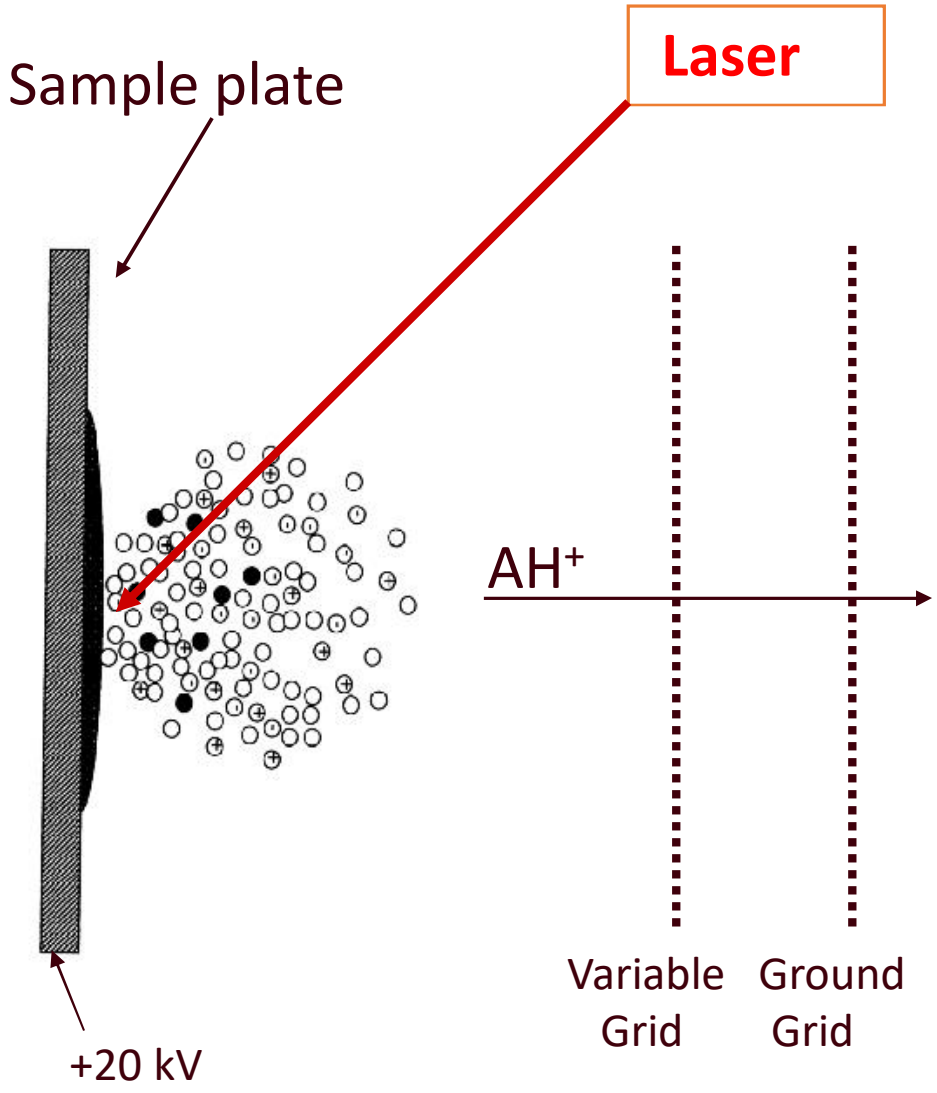


MS and MS/MS


EI
ESI
MALDI
Collision
Photon
Surface
sample
ionization
m/z
separation
fragment
m/z
separation
detection
MS1
MS2
Precursor
ion
Product
ion


4700 Reflector Spec #1 MC[BP = 1664.6, 132]
% Intensity
Mass (m/z)
644
905
1166
1427
1688
1949
768.5455
806.1003
833.3260
856.0405
882.5842
953.0089
1044.0792
1062.0795
1146.4182
1208.9741
1300.0459
1382.4410
1432.5188
1527.5448
1664.5983
1713.5172
1751.6349

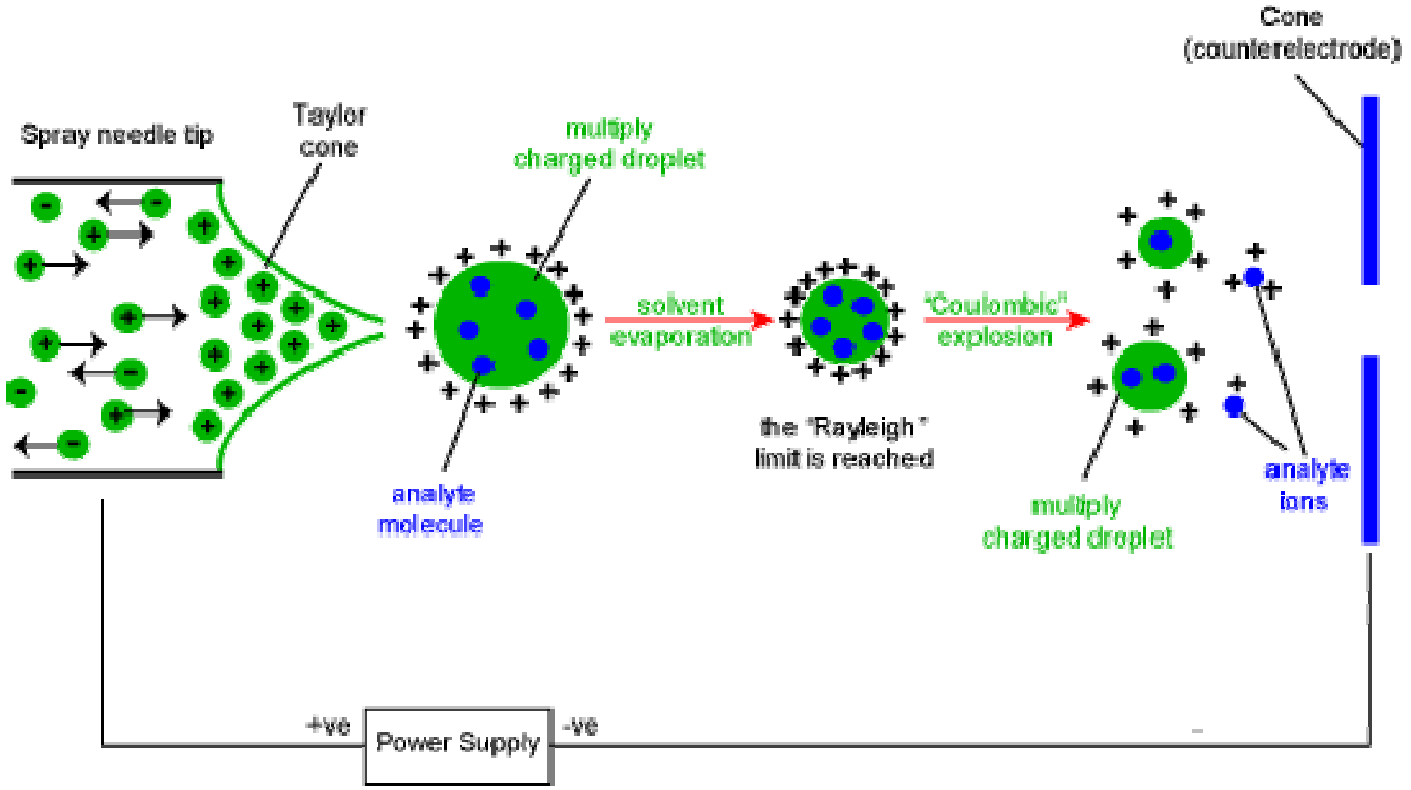
MALDI ionization (Matrix Assisted Laser Desorption Ionization)



1. L'échantillon (A) est mélangé avec un excès de matrice (M) et séché sur la plaque MALDI
2. Le flash Laser ionise les molécules de matrice
3. Les molécules d'échantillon sont ionisées par transfert de protons de la matrice:

   $MH^+ + A \rightarrow M + AH^+$.

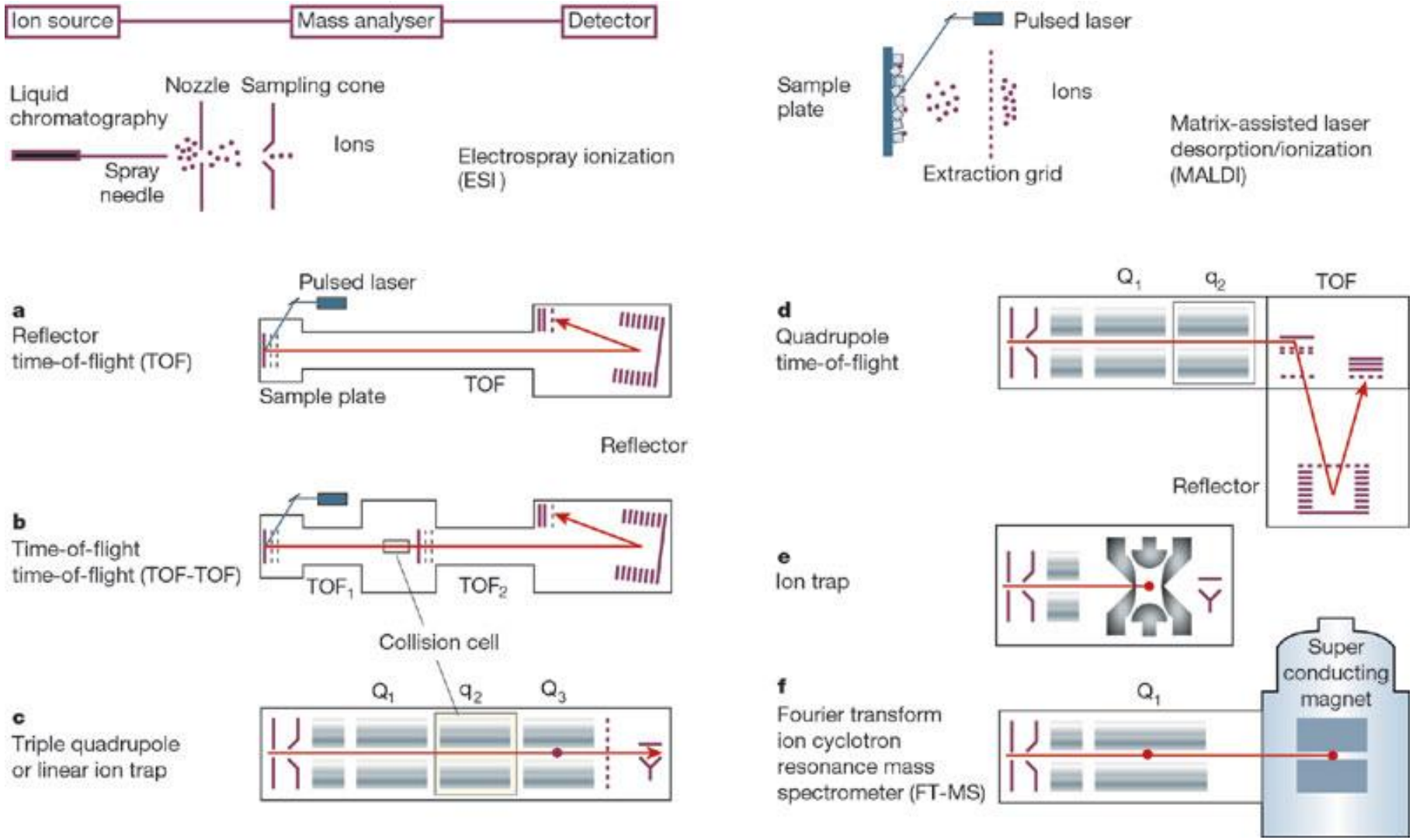
Ionization by electrospray



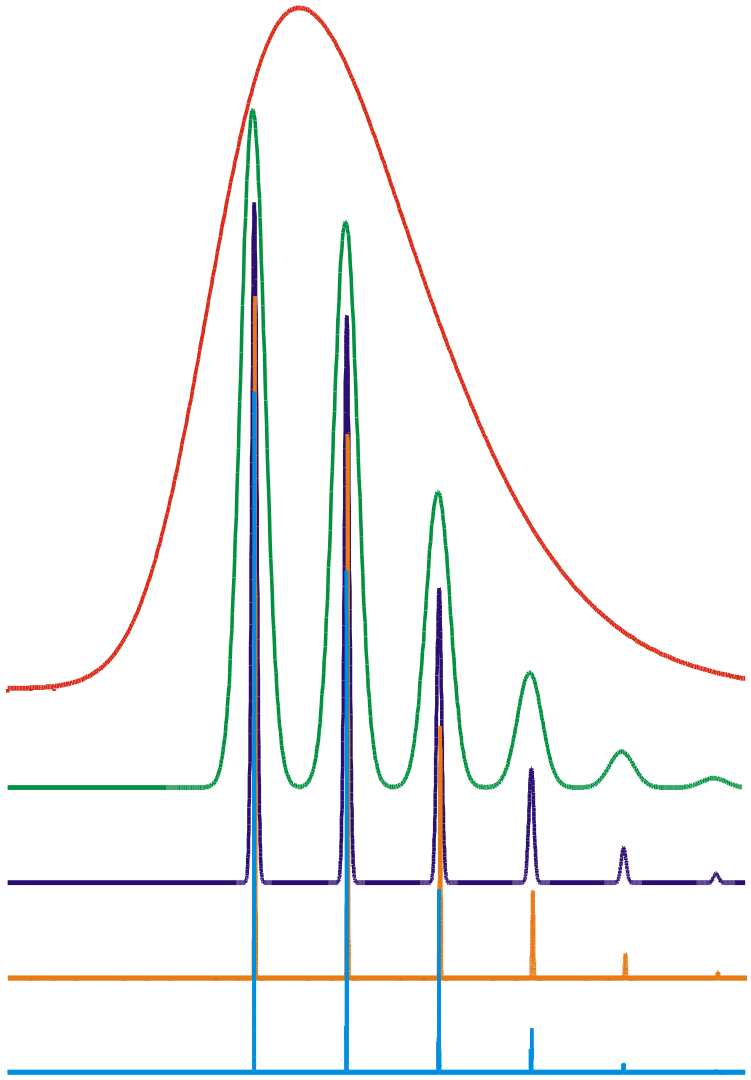
Electrospray

nanospray

Different instrumental design



Importance of spectral resolution

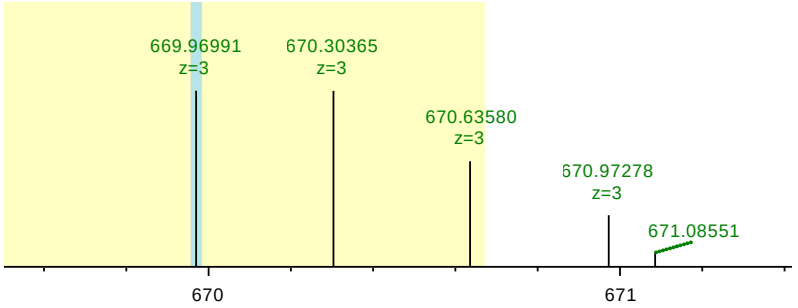


Resolution

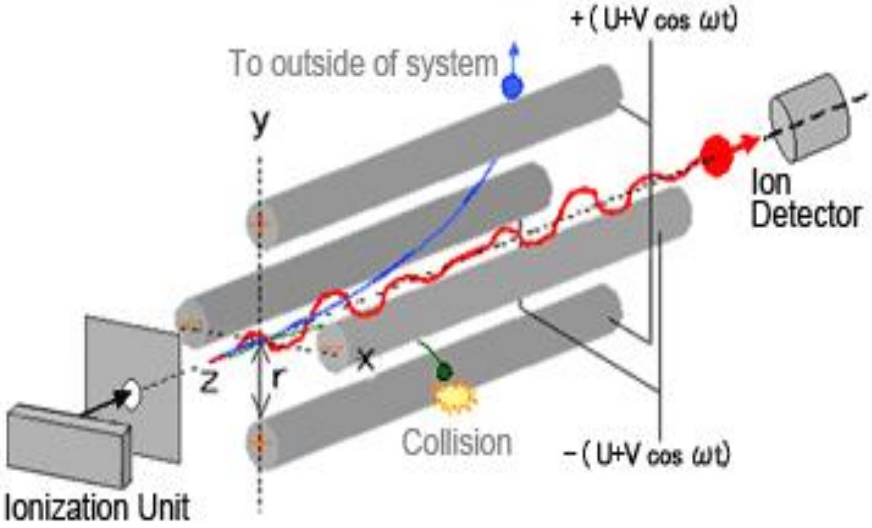
1.000	linear TOF w/o DE
5.000	reflector TOF w/o DE
25.000	reflector TOF with DE
125.000	FTMS wideband mode
1.000.000	FTMS high-res mode

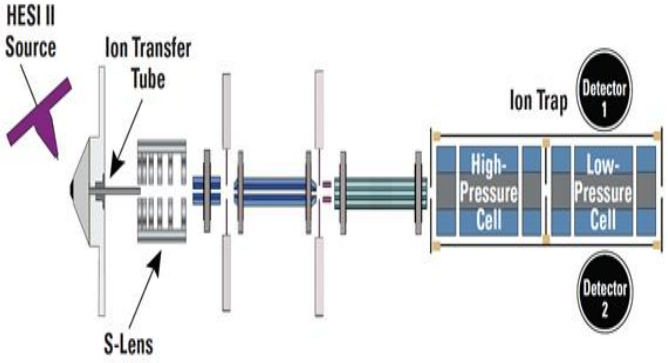
Natural abundance of atoms isotopes in proteins

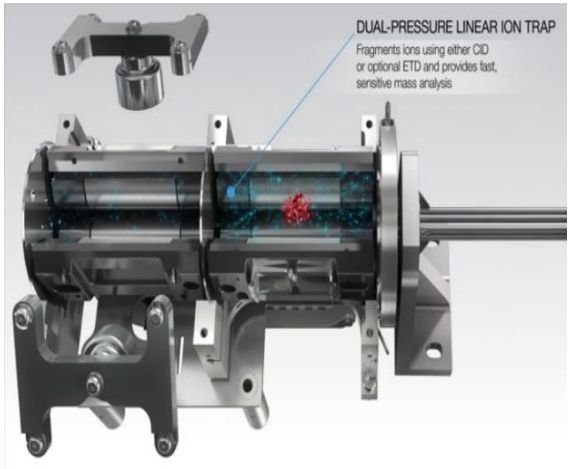
Name	Symbol	Mass (Da)	Abundance (%)
Hydrogen	H	1.007825	99.9885
Deuterium	H	2.014102	0.0115
Carbon	C	12.000000	98.9300
	C	13.003355	1.0700
Nitrogen	N	14.003074	99.6320
	N	15.000109	0.3680
Oxygen	O	15.994915	99.7570
	O	16.999132	0.0380
	O	17.999160	0.2050
Phosphorus	P	30.973762	100.0000
Sulfur	S	31.973762	94.9300
	S	32.971458	0.7600
	S	33.967867	4.2900
	S	35.967081	0.0200



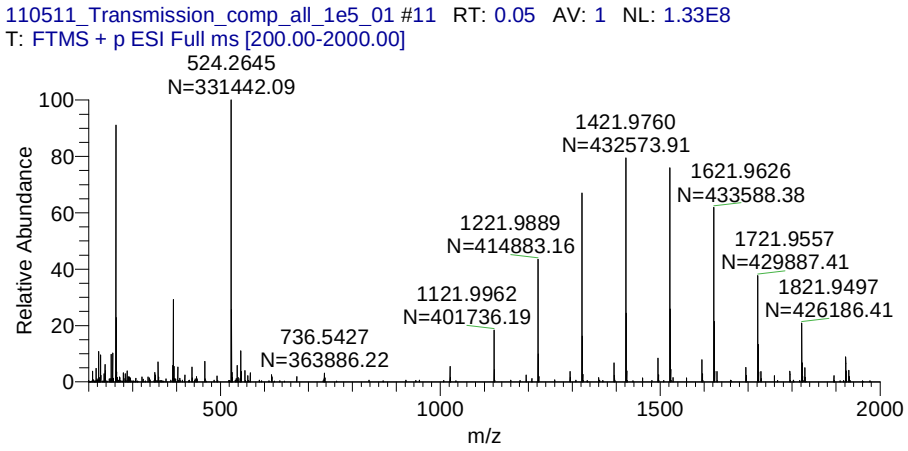
Different instrumental design



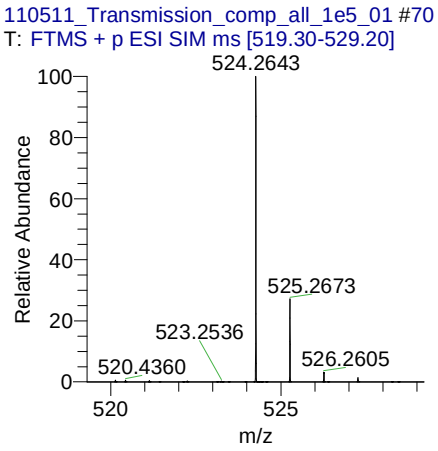




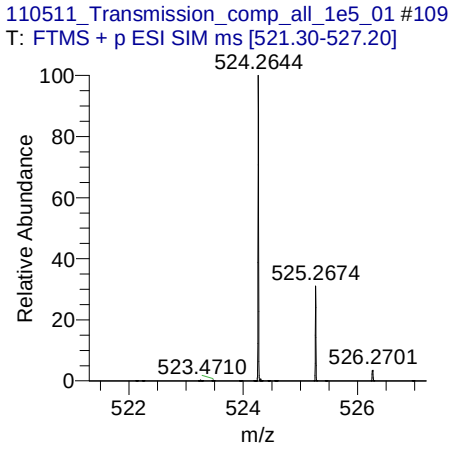
Isolation Width: Full MS to 1 amu (MRFA)



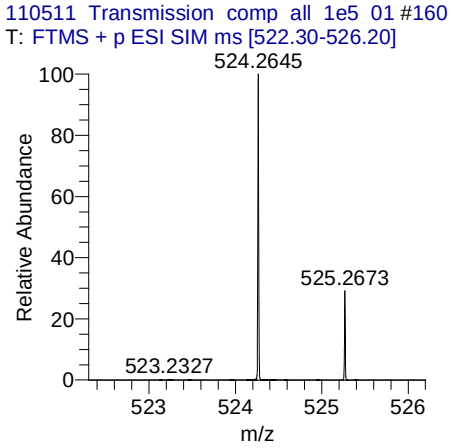
IsoW= 1800



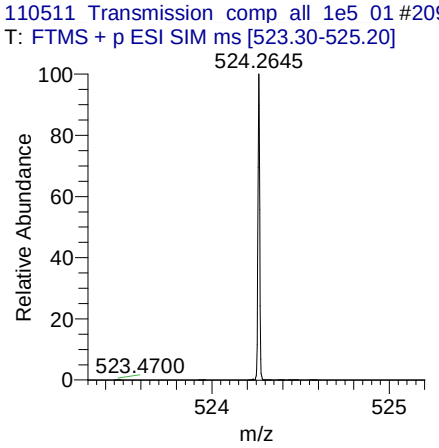
IsoW= 10



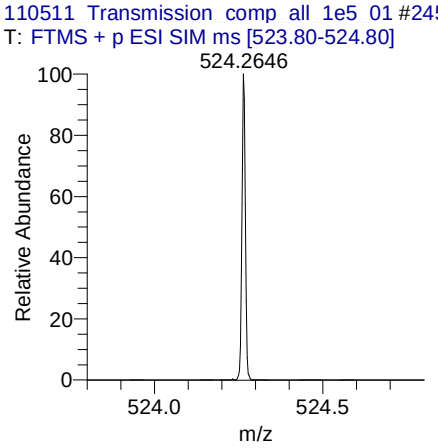
IsoW= 6



IsoW= 4

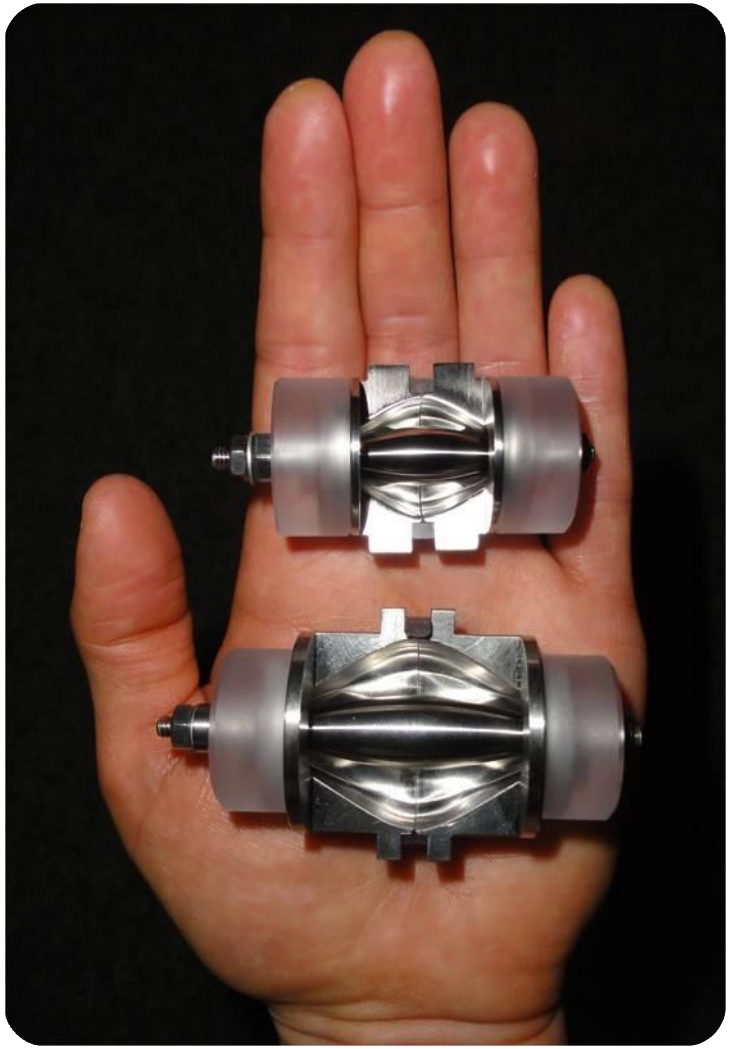


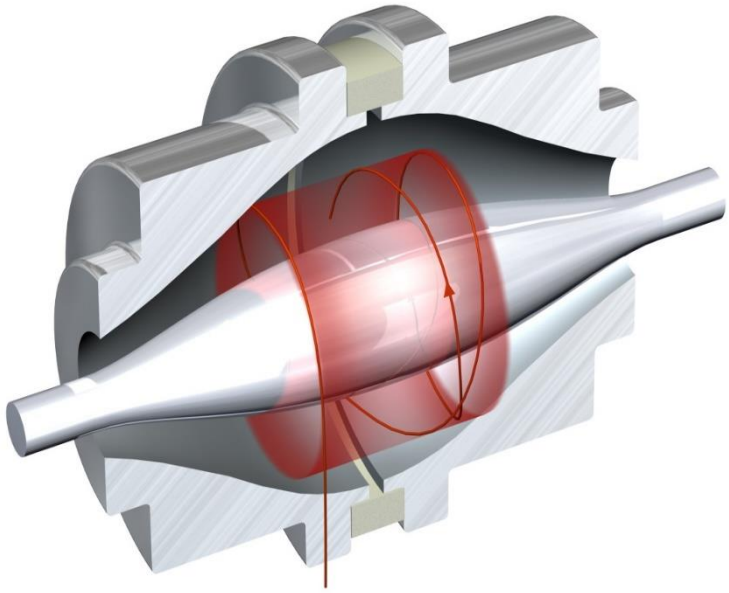
IsoW= 2



IsoW= 1

the orbitrap cell





$$\omega_z = \sqrt{\frac{k}{m/z}}$$

Developpement of the orbitrap family

2007 LTQ Orbitrap XL and Discovery

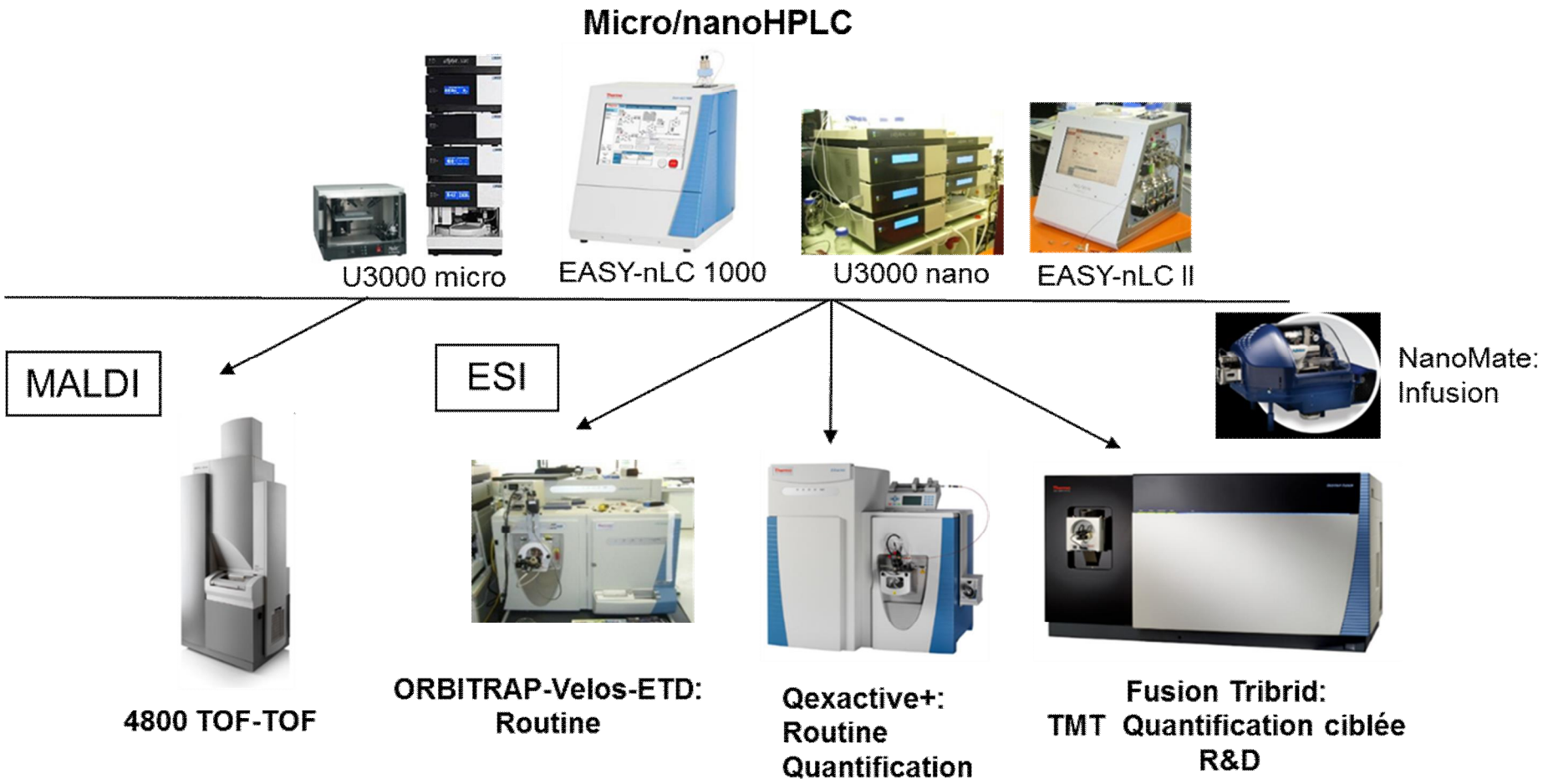
2008 LTQ Orbitrap XL ETD

2009 LTQ Orbitrap Velos

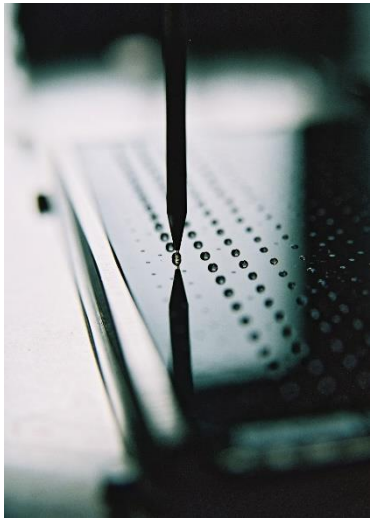
2011 LTQ Orbitrap Velos Pro
Orbitrap Elite

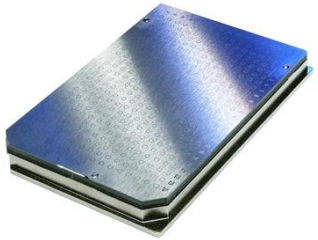
2013 ***Orbitrap Fusion Tribrid***

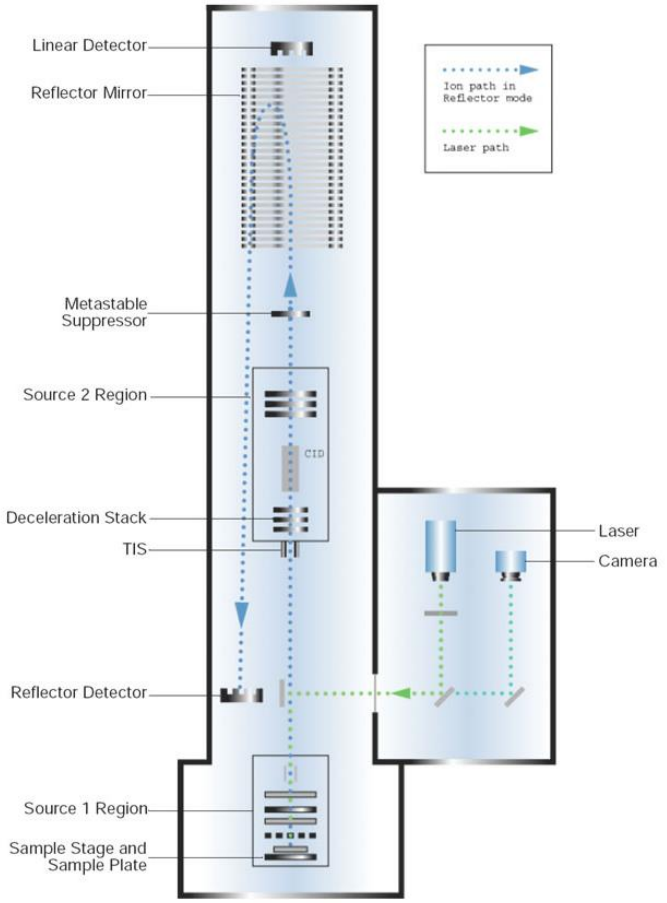
Instrumentations



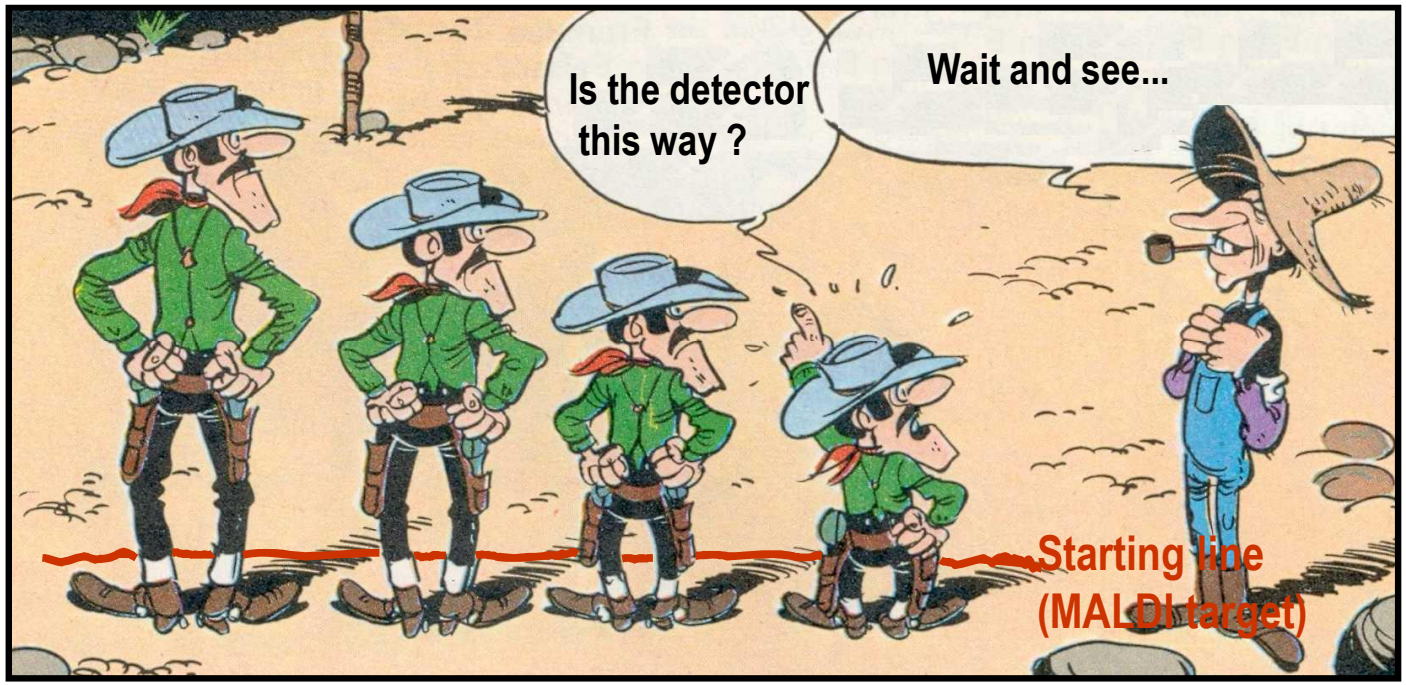
MALDI-TOF/TOF







Time of flight – principles (TOF)

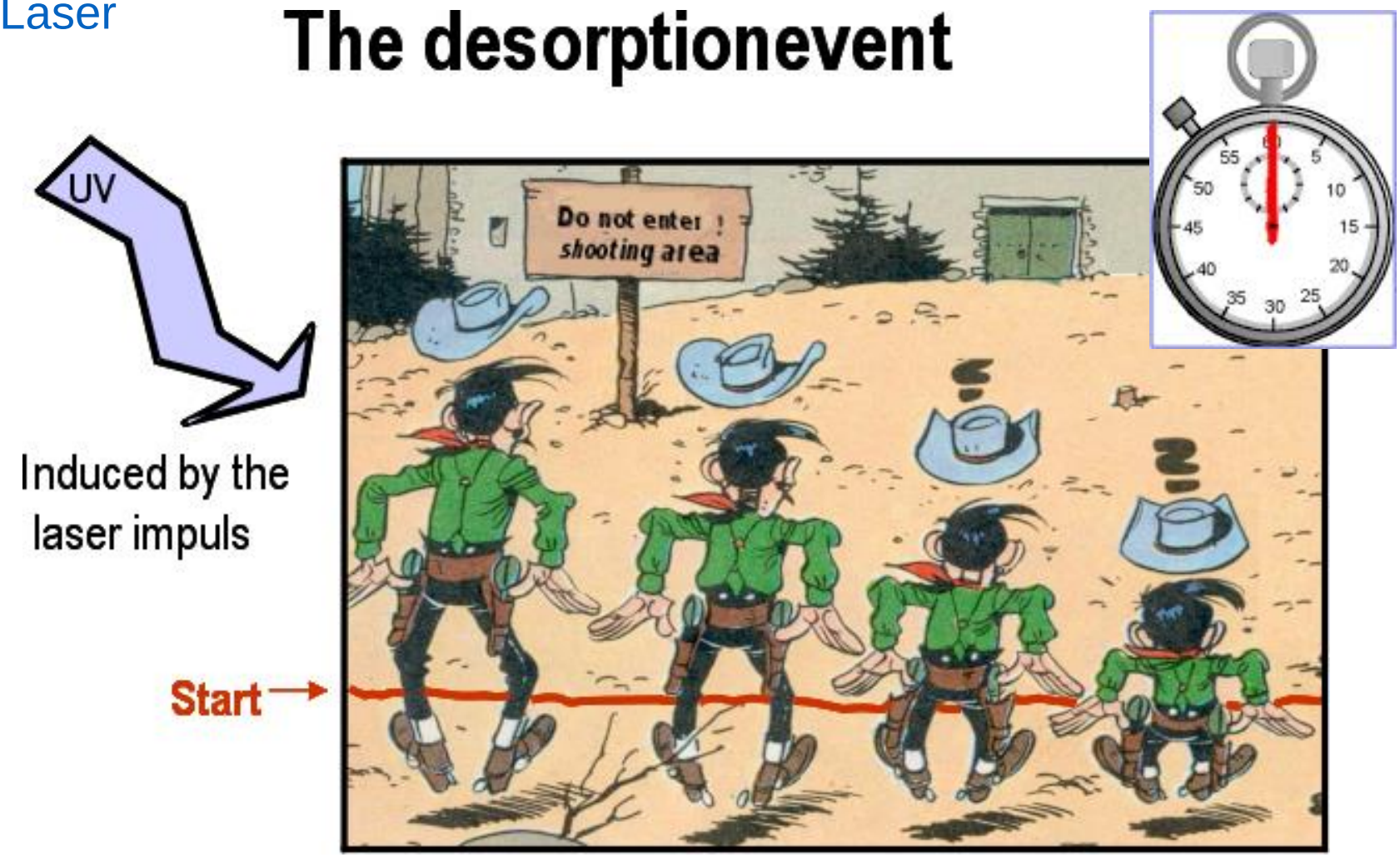


Remember : Mass of an ion is measured in the Dalton units !

Start !

Laser

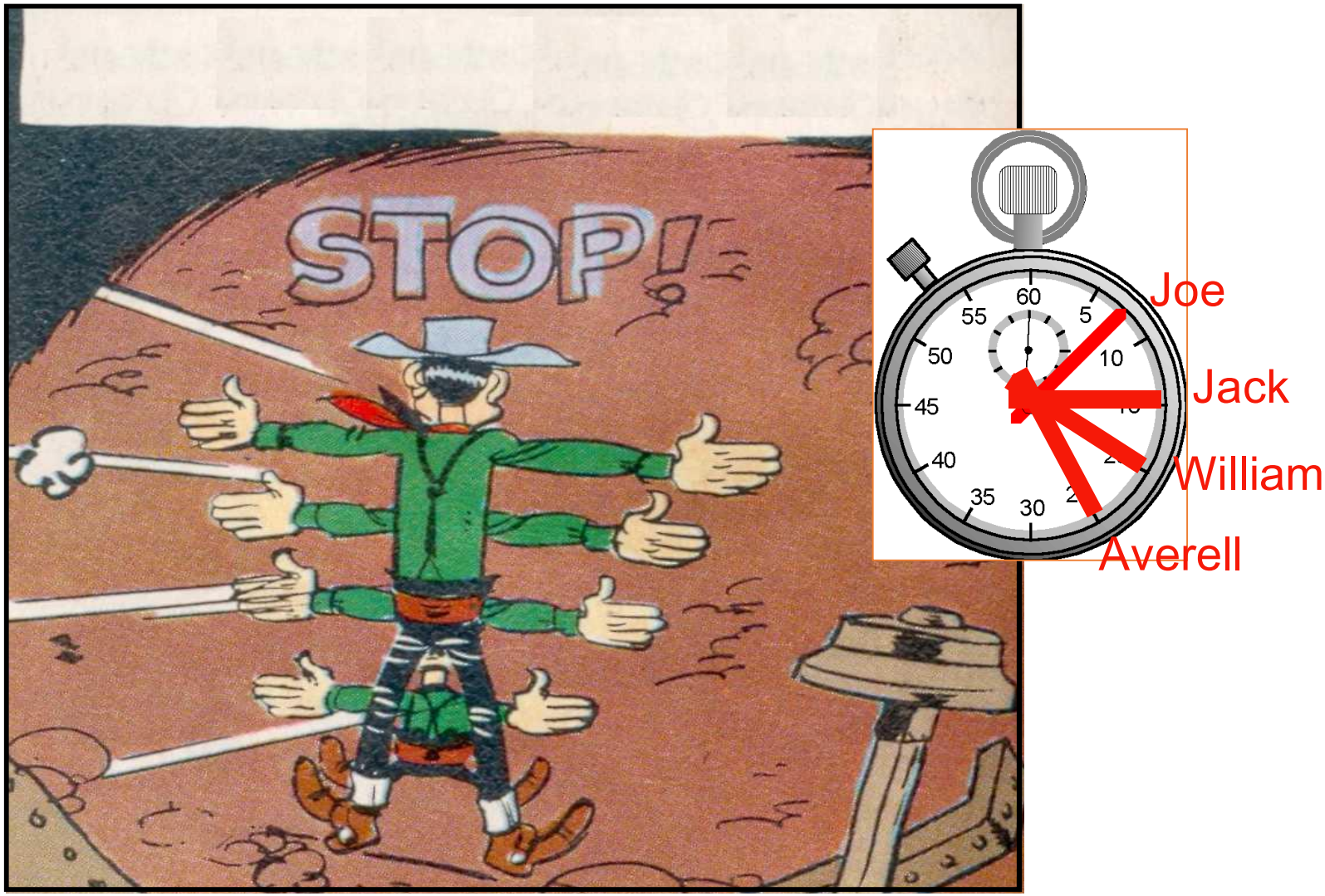
The desorptionevent



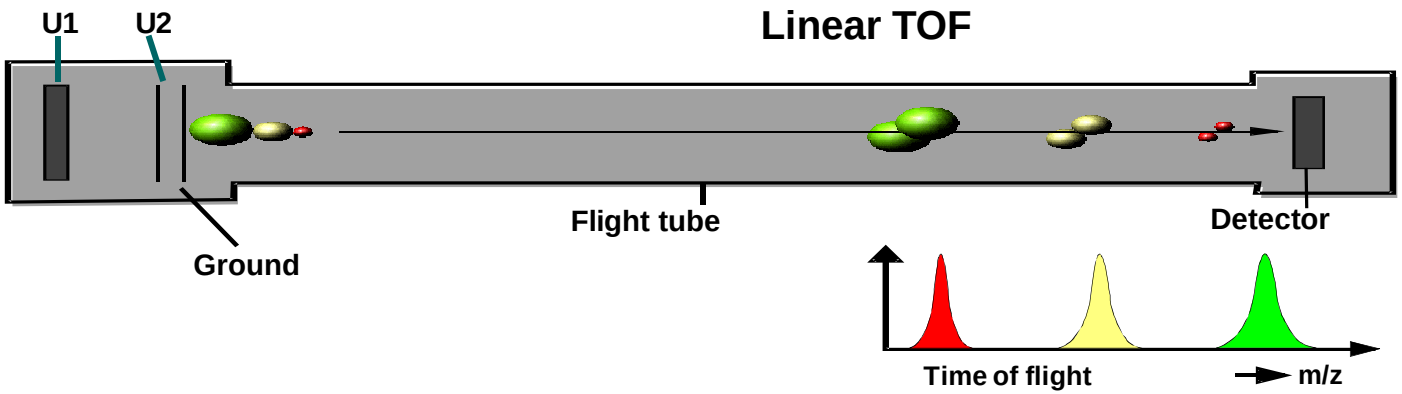
Ions in the time of flight (TOF)

Increasing MW

Ions in the detector



Ions analysis in linear mode



Electric field : $Ec = qU = 1/2\ mV^2$ Identical for all ions

(V=L/t L : tube lenght)

$\Rightarrow$ Simple relation $t^2 = mL^2/2qU$ = Constante x m/z

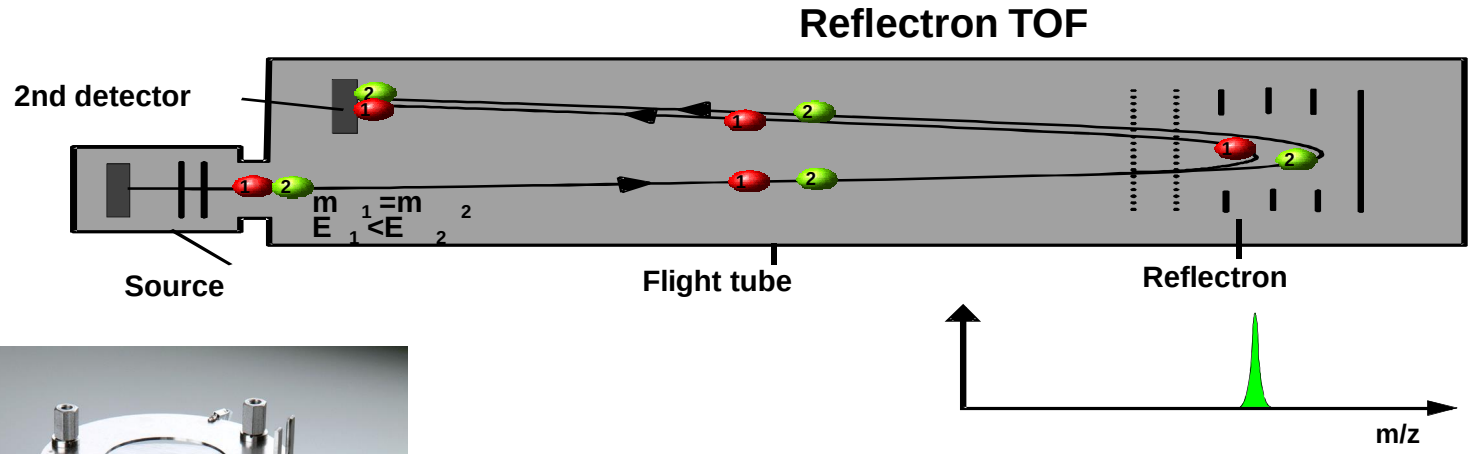
Light Corrections :

$t^2 = Am^2 + Bm + C$ (A : initial desorption Ec

C : Extraction Delay)

=> Simple Quadratic equation

Ions analysis in reflectron mode



Reflectron : 2 effects on resolution :

- Increasing flight path (better separation of particles of different masses) (equivalent 3m flight tube)
- Focusing effet for particles with same mass

Key questions in proteomics

- What is the protein content of my biological sample?

  => problem of identification

- What is the abundance of my protein of interest?

  => quantification problem

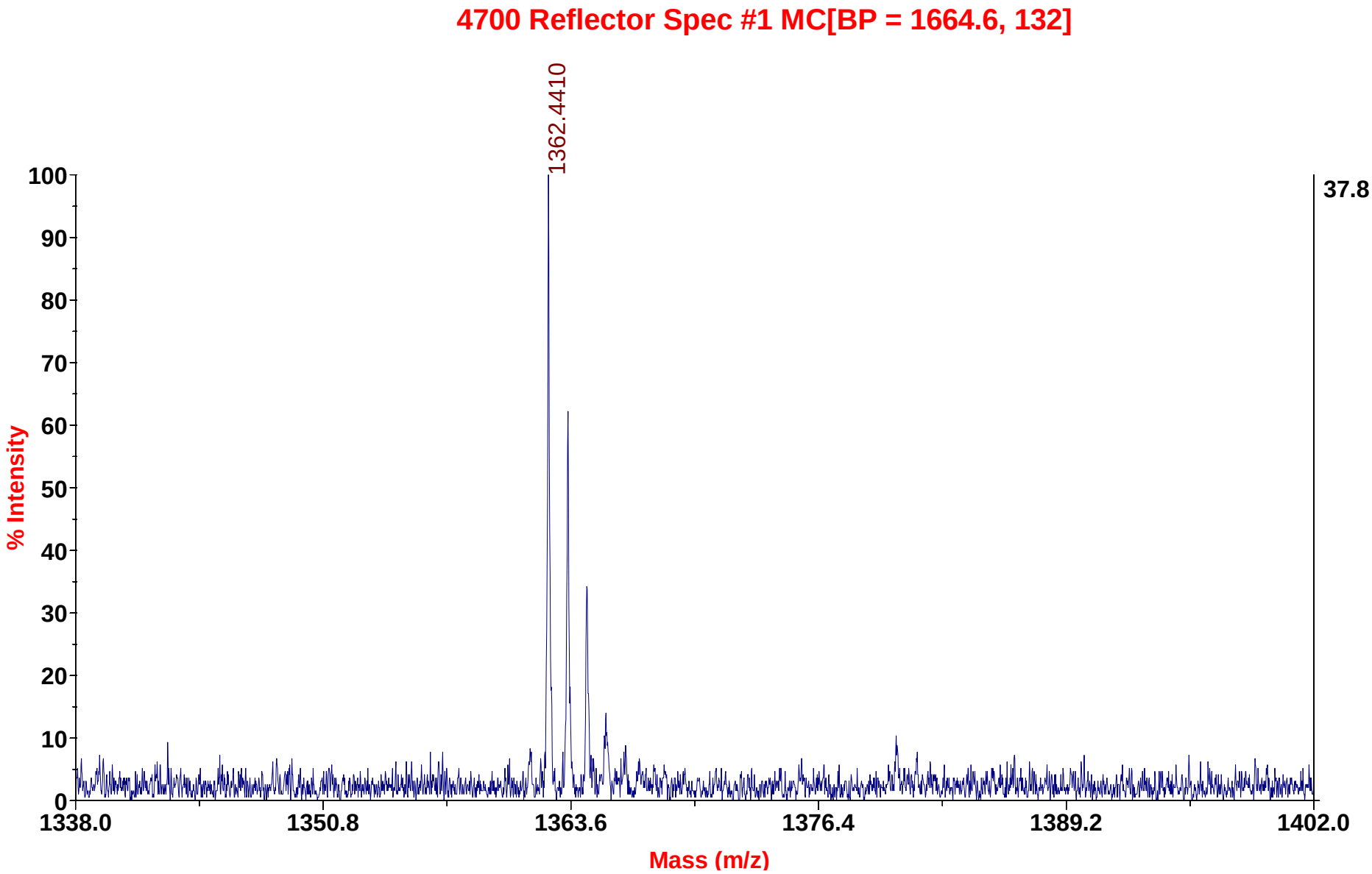
- Relative question: What are the protein abundance variations of the proteomes studied?

- What are the partners of my protein of interest?

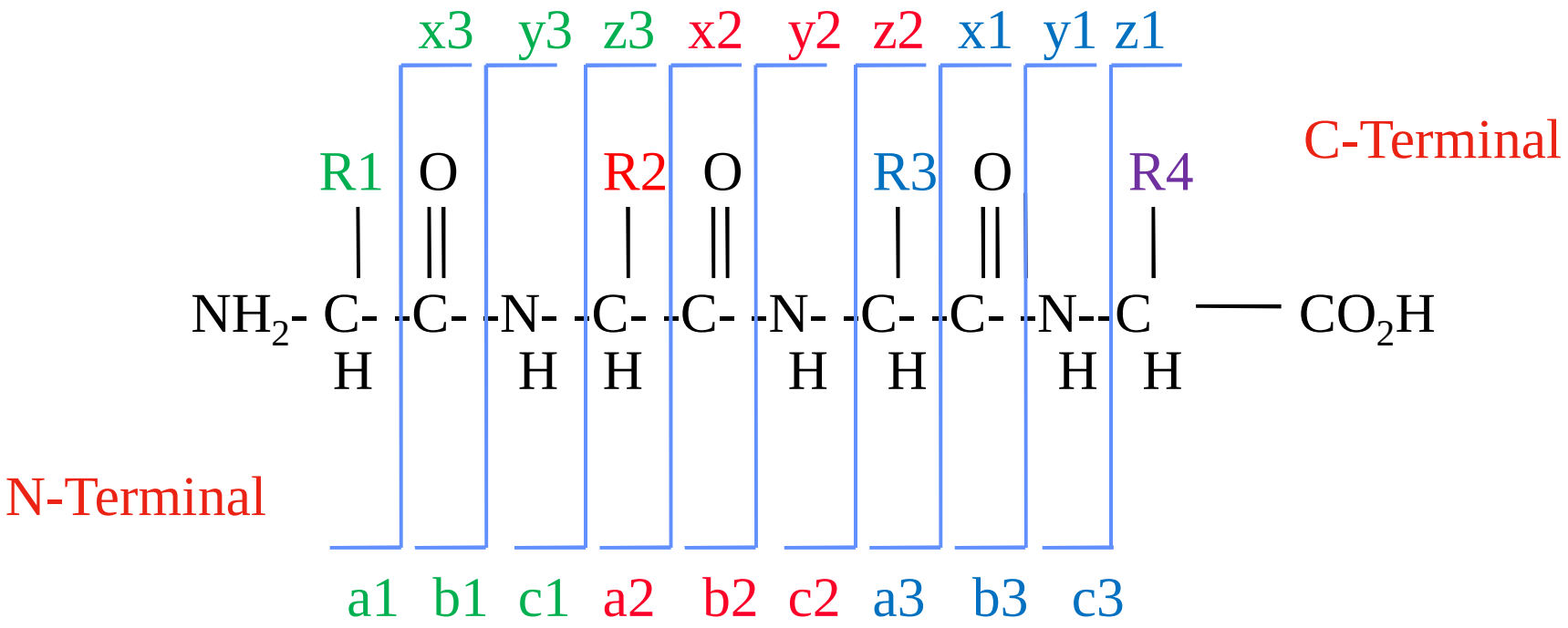
- Are there any signature proteins related to a particular biological process?

  => biomarkers identifications and quantifications

Ion precursor selection



MS/MS fragmentation for peptides



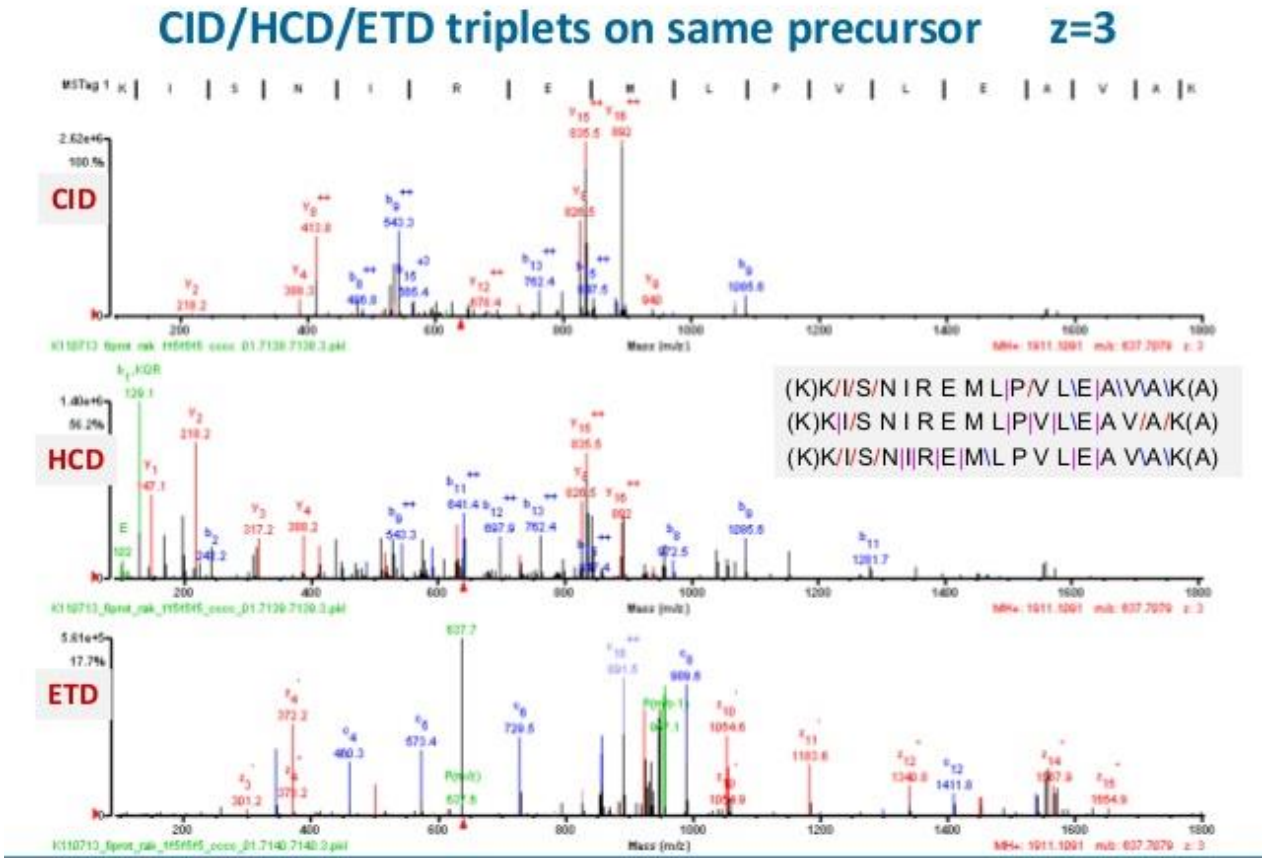
a $H — (NH\text{-}CHR\text{-}CO)\text{-}\overset{+}{N}H{=}CHR$

b $H — (NH\text{-}CHR\text{-}CO)\text{-}NH{=}CHR\text{-}C{=}\overset{+}{O}$

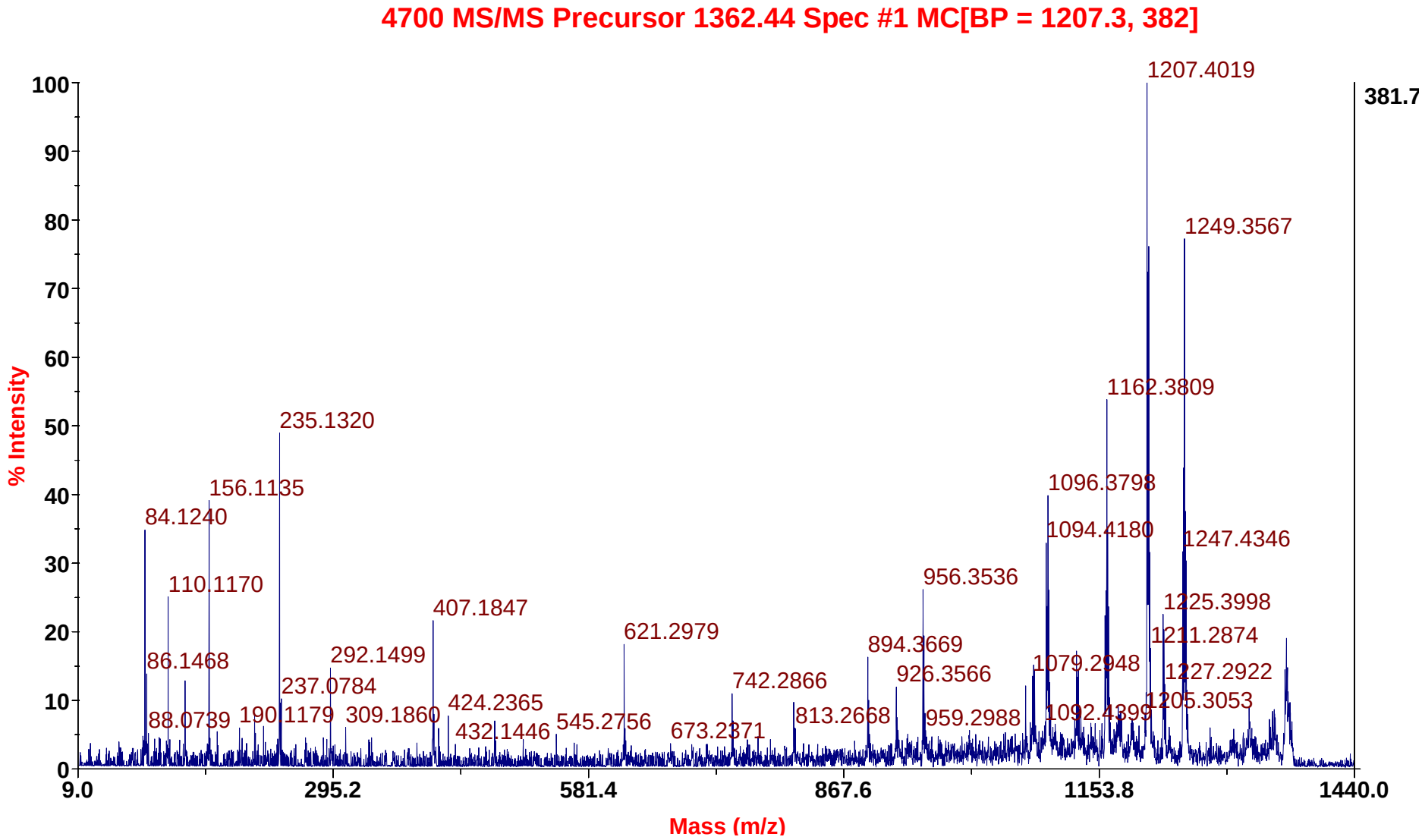
y $H — (\underbrace{NH\text{-}CHR\text{-}CO}_{+H})\text{-}OH$

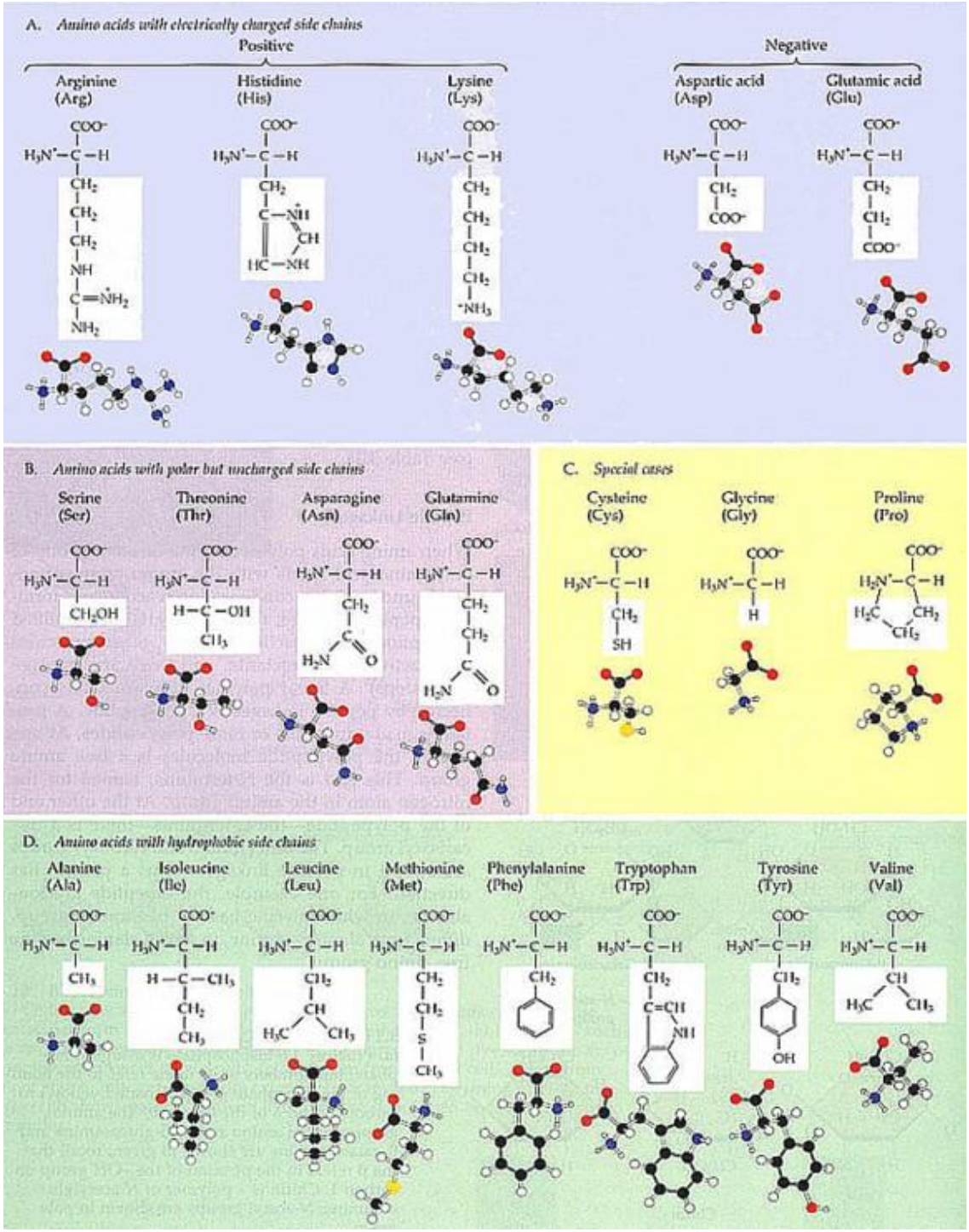
MS/MS fragmentation for peptides

DISSOCIATION INDUITE PAR COLLISION (CID)
HIGHER ENERGY COLLISIONAL DISSOCIATION (HCD)
ELECTRON TRANSFER DISSOCIATION (ETD)



MS/MS spectrum of the precursor 1362.44 m/z





Alanine	A, Ala	71.079
Arginine	R, Arg	156.188
Asparagine	N, Asn	114.104
Aspartic acid	D, Asp	115.089
Cysteine	C, Cys	103.145
Glutamine	Q, Gln	128.131
Glutamic acid	E, Glu	129.116
Glycine	G, Gly	57.052
Histidine	H, His	137.141
Isoleucine	I, Ile	113.160
Leucine	L, Leu	113.160
Lysine	K, Lys	128.17
Methionine	M, Met	131.199
Phenylalanine	F, Phe	147.177
Proline	P, Pro	97.117
Serine	S, Ser	87.078
Threonine	T, Thr	101.105
Tryptophan	W, Trp	186.213
Tyrosine	Y, Tyr	163.176
Valine	V, Val	99.133

MS/MS spectra interpretation

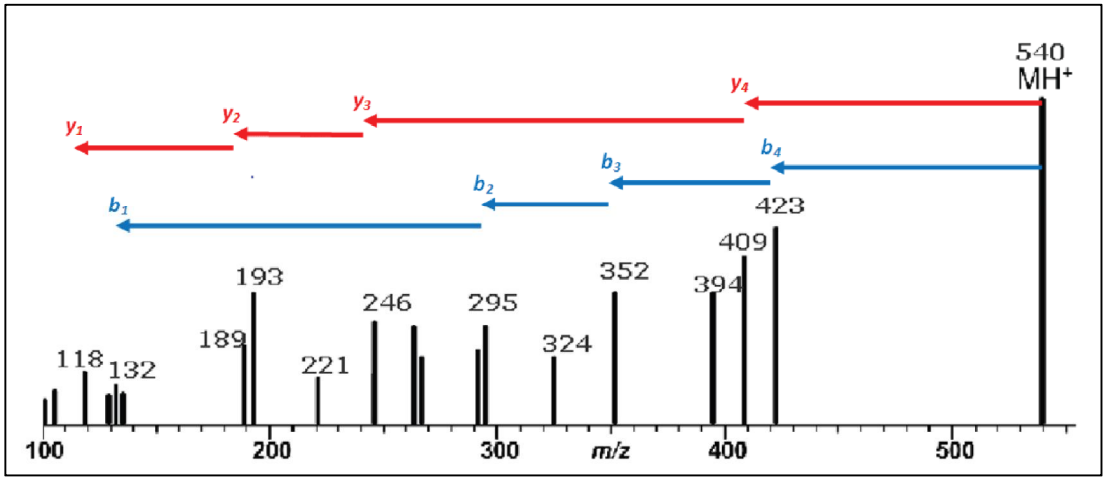


Table 1

Ion	*m/z*	Neutral loss (from previous ion in the series)	Amino Acid Residue
Precursor $[M+H]^+$	540		
y_4	409	131	M
y_3	246	163	Y
y_2	189	57	G
y_1	118	71	A
b_4	423	117 (99+18)	V
b_3	352	71	A
b_2	295	57	G
b_1	132	163	Y
a_4	395?		
a_3	324		
a_2	267		
a_1	104		

MYGAV
User AA Formula 1: C2 H3 N1 O1
Elemental Composition: C24 H38 N5 O7 S1
MH+1(av) MH+1(mono)
540.6627 540.2486
[–] Main Sequence Ions

b				y
---	1	M	5	---
295.1111	2	Y	4	409.2082
352.1326	3	G	3	246.1448
423.1697	4	A	2	189.1234
---	5	V	1	118.0863

http://www.sepscience.com/Information/Archive/MS-Solutions/246-

Current post-translational modifications (PTMs)

Acids & amides (E/D/Q/N)	Pyroglutamic acid (Q)	-17.0306	Deamidation (Q/N)	+0.9847
	Carboxylation (E/D)	+44.0098		

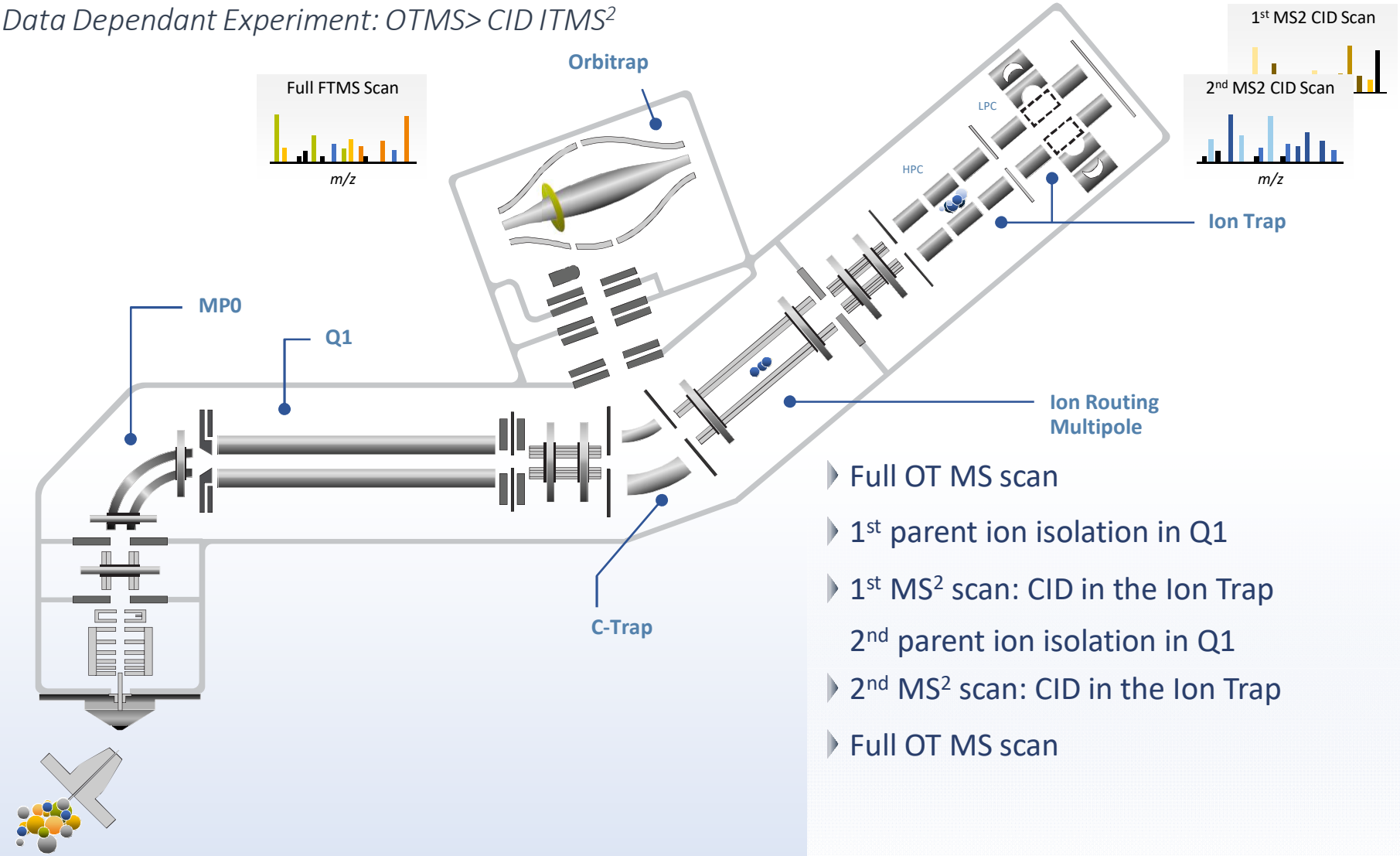
Hydroxyl groups (S/T/Y)	Phosphorylation	+79.9799	Sulphation	+80.0642

Carbohydrates (S/T/N)	Pentoses	+132.1161	Deoxyhexoses	+146.1430
	Hexosamines	+161.1577	Hexoses	+162.1424
	N-acetylhexosamines	+203.1950	Sialic acid	+291.2579

Sulphydryls (C)	Disulphide bond	-2.0159	Oxidation	+15.9994
	Cysteinylation	+119.1442	Glutathionylation	+305.3117

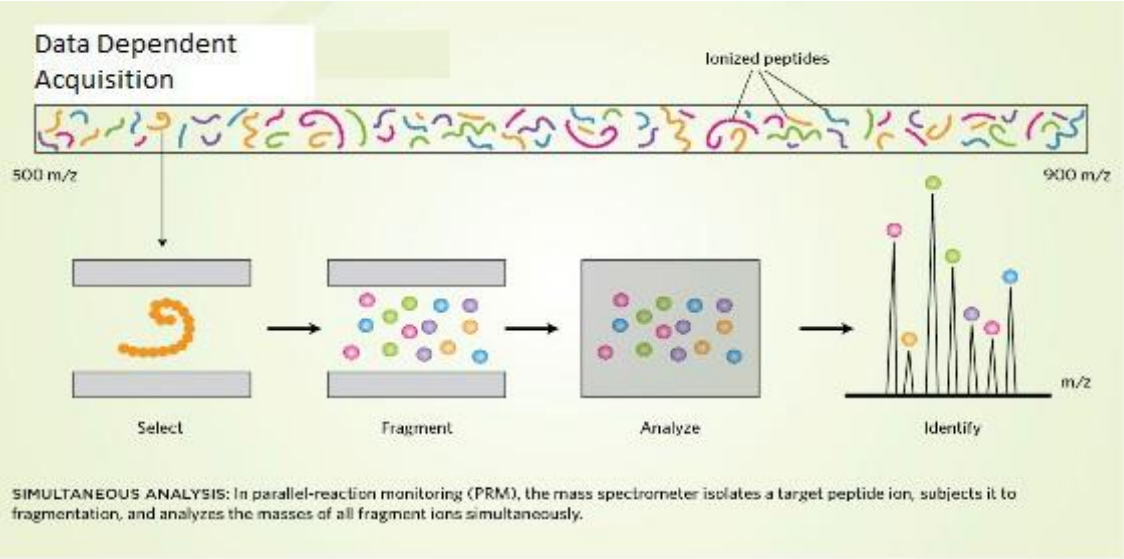
MS and MS/MS spectra generation

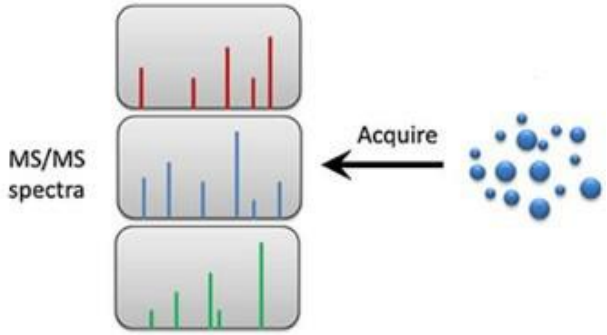
Data Dependant Experiment: OTMS> CID ITMS²

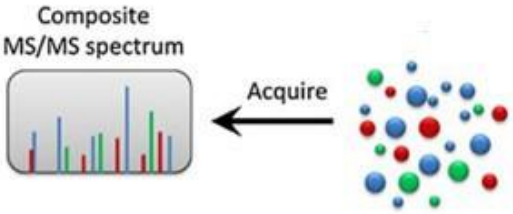


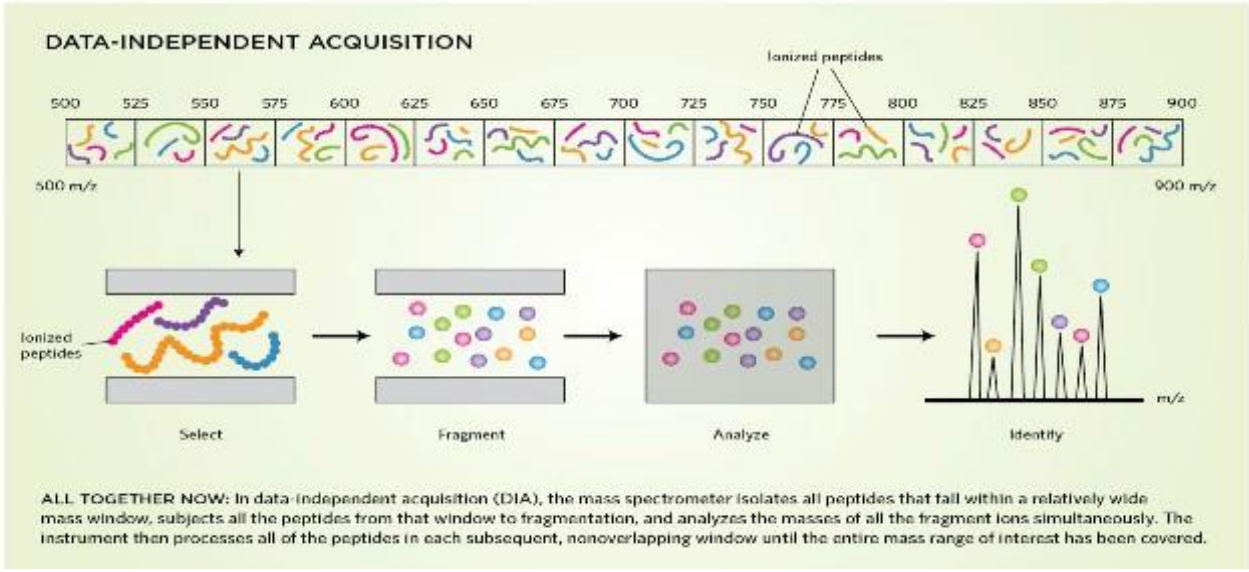
- Full OT MS scan
- 1st parent ion isolation in Q1
- 1st MS² scan: CID in the Ion Trap
- 2nd parent ion isolation in Q1
- 2nd MS² scan: CID in the Ion Trap
- Full OT MS scan

DDA versus DIA

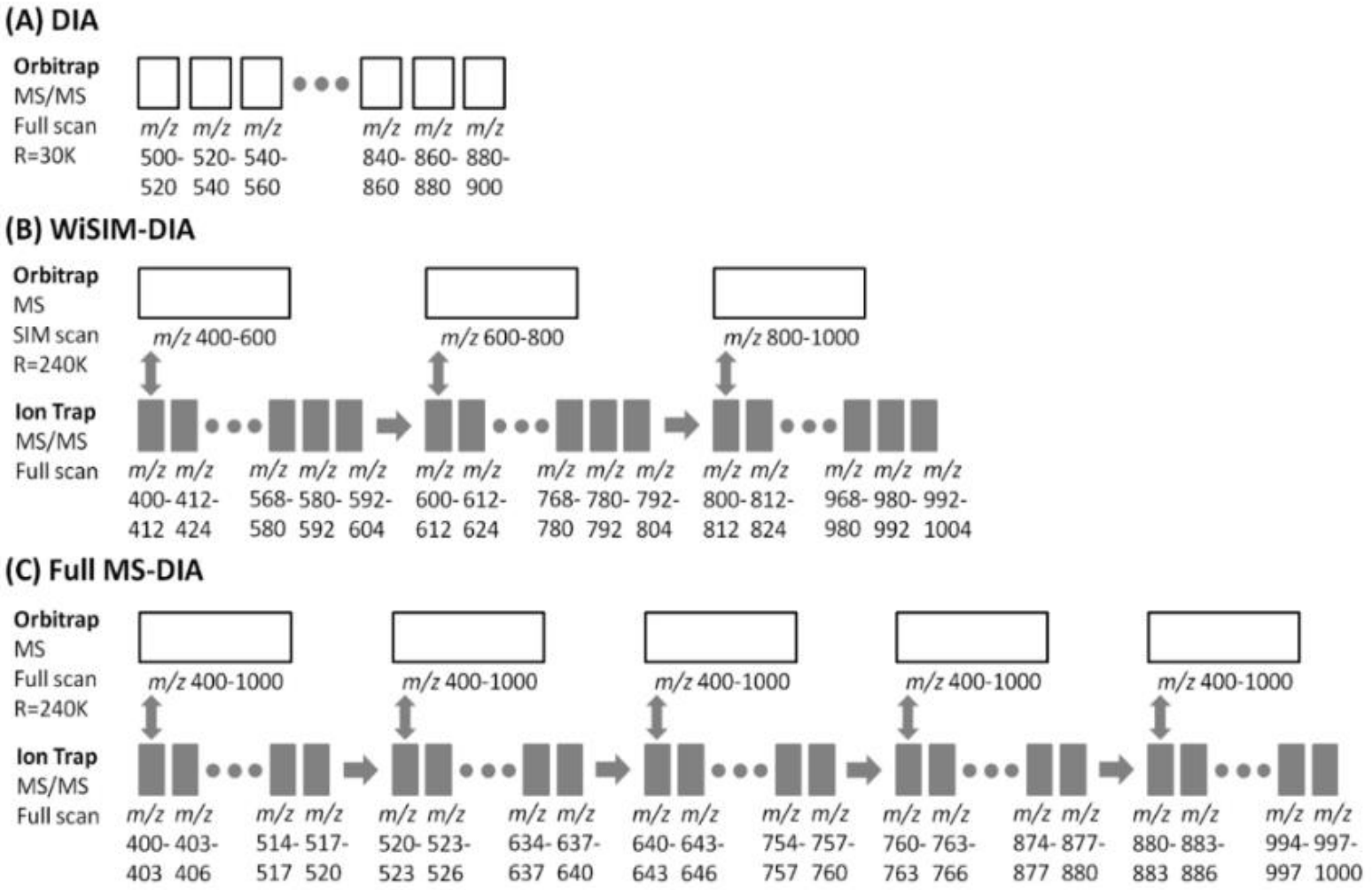




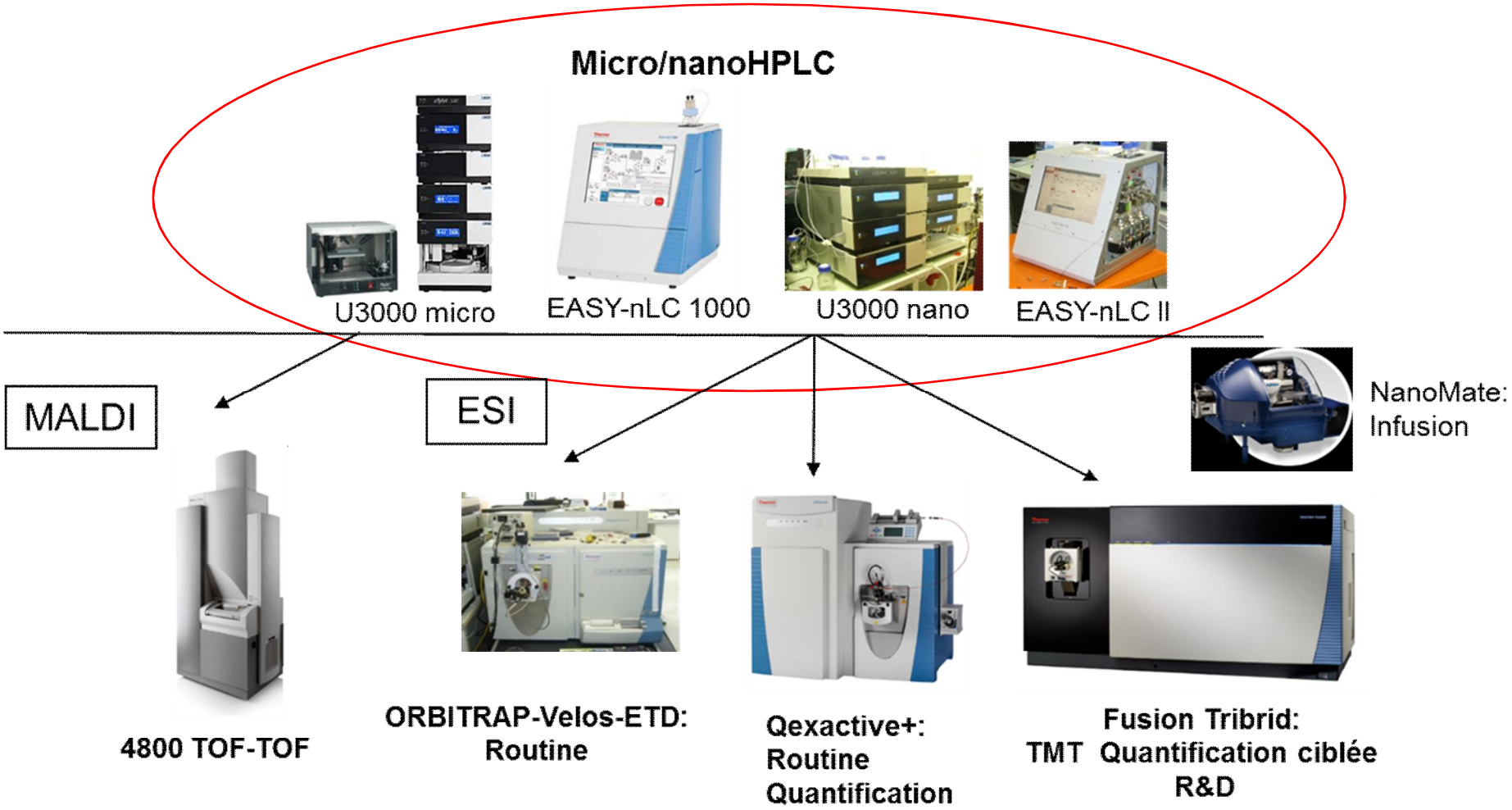




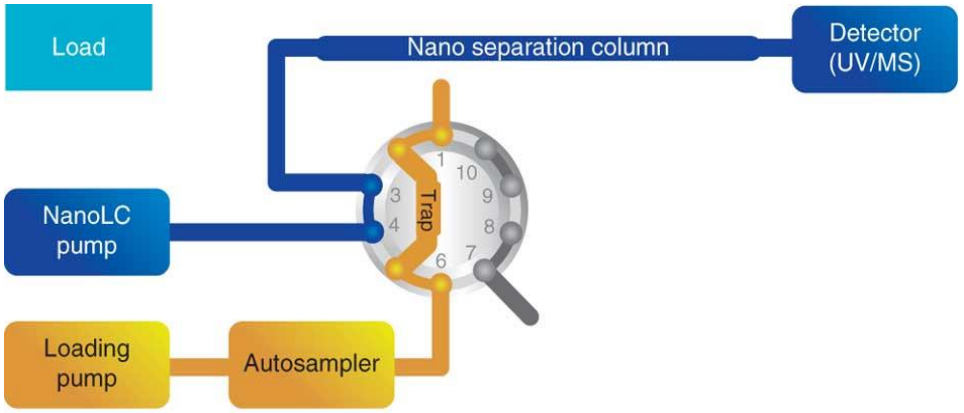
Data Independent Acquisition: DIA

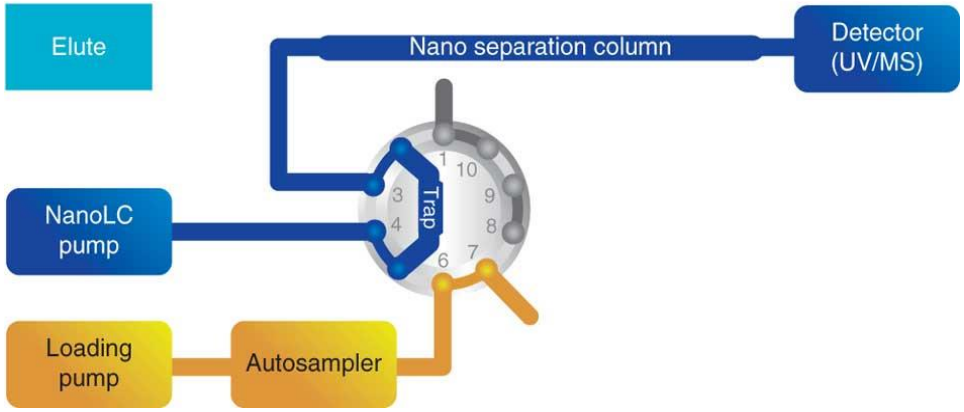


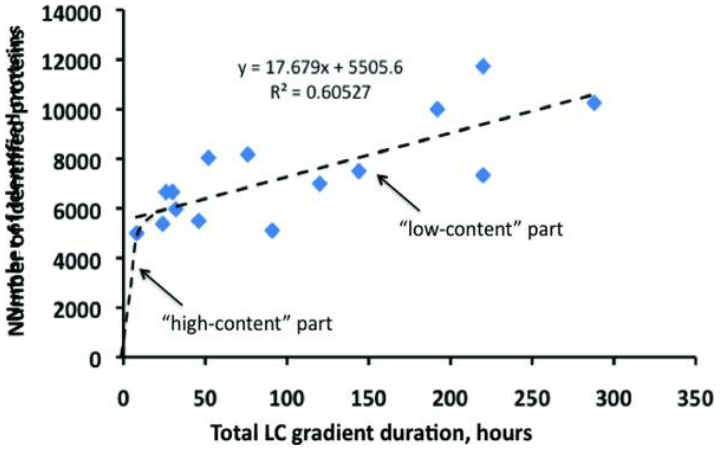
Contribution of nano-HPLC



Peptides separation by nano-LC





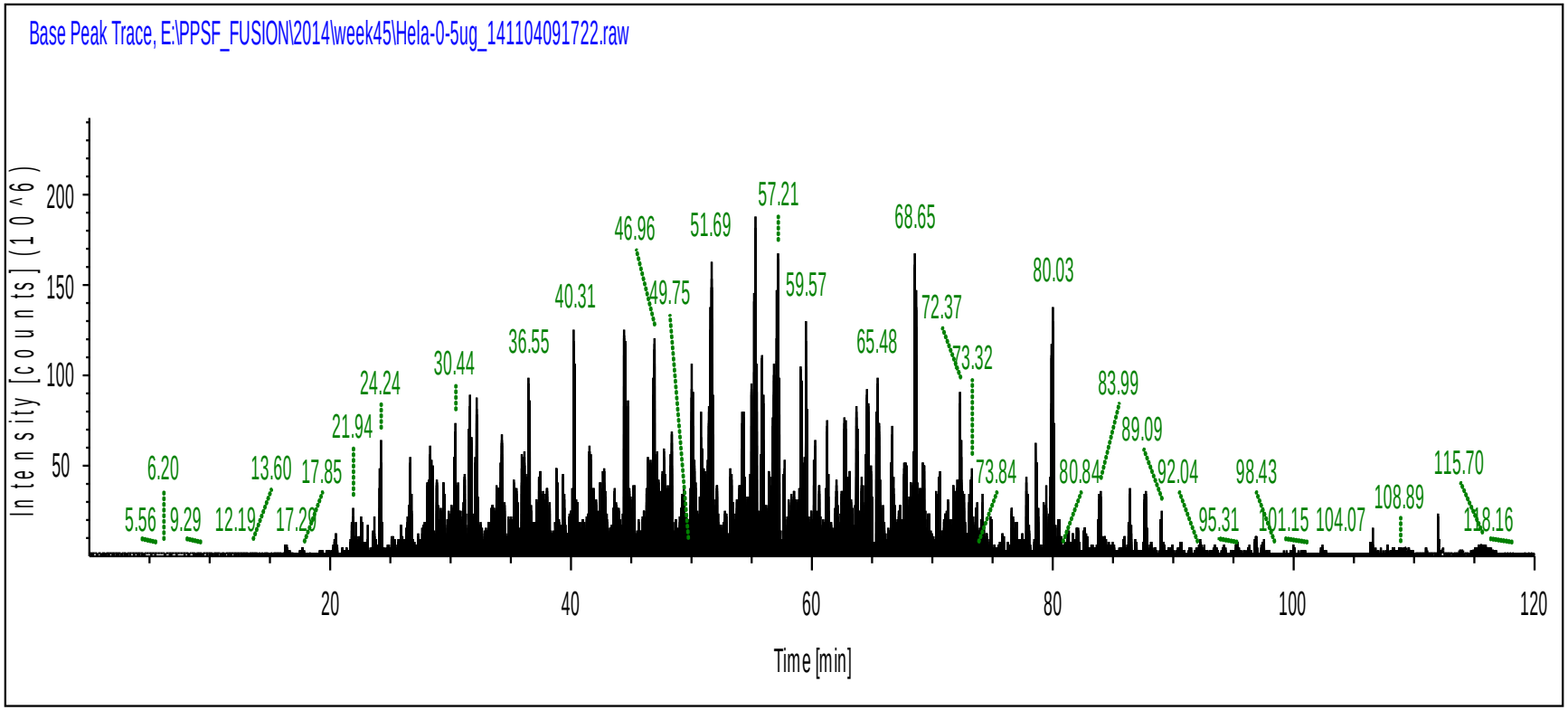


Peptides nano-LC separation

- It is impossible to resolve all species in a proteomics sample using only one separation method
- Multidimensional separation - two or more independent (ñorthogonalò) separation techniques coupled together for the analysis of a single sample.

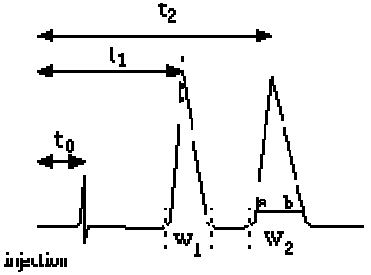
Separation method	Separation by:
Reversed phase	Hydrophobicity
Ion exchange, IsoElectroFocusing (IEF)	Net charge, Isoelectric point
Size exclusion, SDS Gel Electrophoresis	Size, molecular weight
Affinity chromatography	Specific functional groups

Total ion current (TIC) Hela tryptic digest (0.5 µg of total proteins)

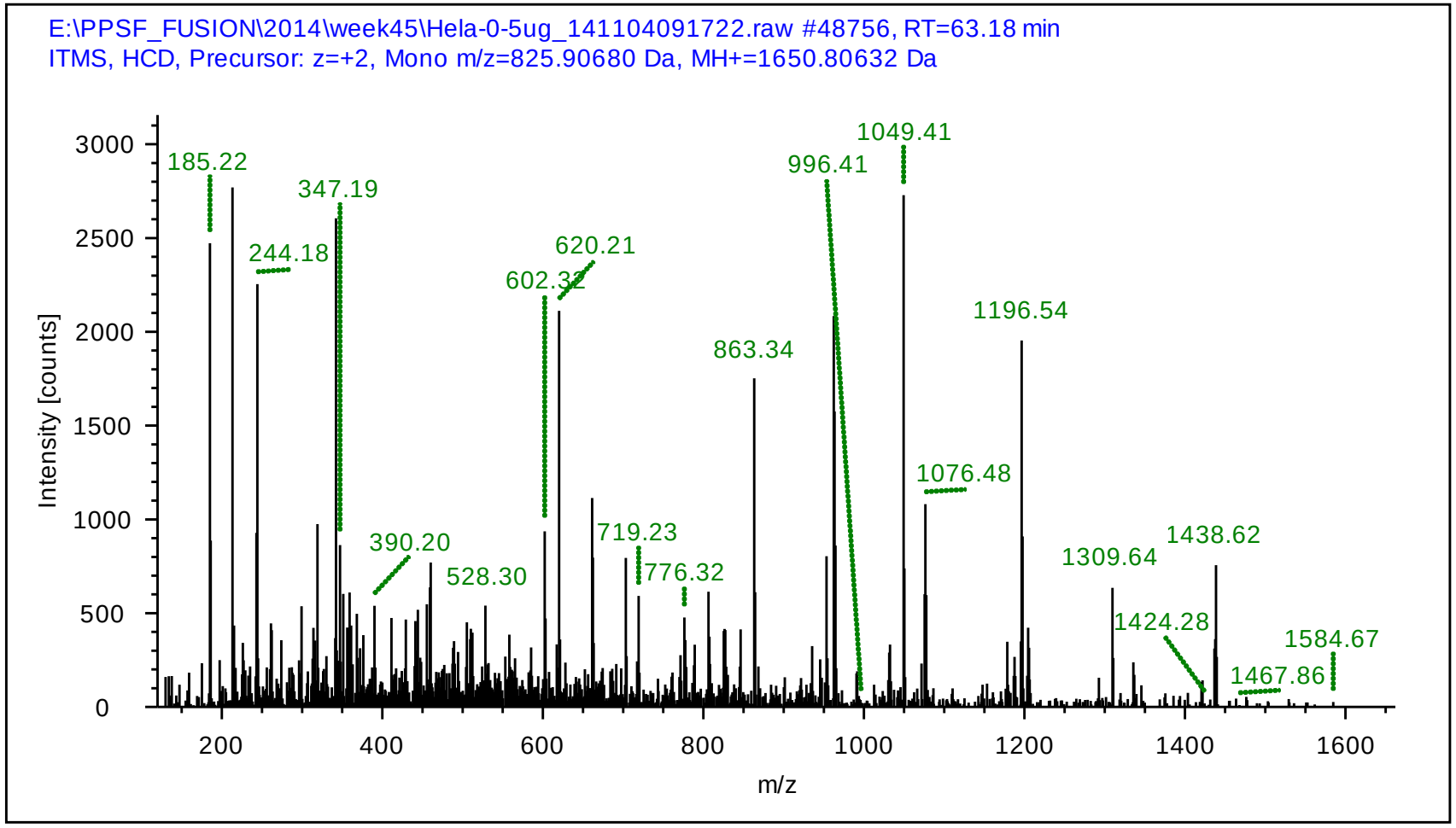


$$R = \frac{2(t_2 - t_1)}{(w_2 - w_1)}$$

with t1 and t2 the retention time and w1 and w2 peak widths at mid-height



90000 MSMS in 2h gradient (C18 RPC)



Proprietary MS data formats

Company	Extension	File type
Agilent Bruker	.D (folder)	Agilent MassHunter, Agilent ChemStation, or Bruker BAF/YEP/TDF data format
Agilent/Bruker	.YEP	instrument data format
Bruker	.BAF	instrument data format
Bruker	.FID	instrument data format
Bruker	.TDF	timsTOF instrument data format
ABI/Sciex	.WIFF	instrument data format
ABI/Sciex	.t2d	4700 and 4800 file format
Waters	.PKL	MassLynx peak list format
Thermo PerkinElmer	.RAW*	Thermo Xcalibur PerkinElmer TurboMass
Micromass**/Waters	.RAW* (folder)	Waters MassLynx
Chromtech Finnigan***	.DAT	Finnigan ITDS file format; MAT95 instrument data format
VG		MassLab data format
Finnigan***	.MS	ITS40 instrument data format
Shimadzu	.QGD	GCMSSolution format
Shimadzu	.qgd	instrument data format
Shimadzu	.lcd	QQQ/QTOF instrument data format
Shimadzu	.spc	library data format
Bruker/Varian	.SMS	instrument data format
Bruker/Varian	.XMS	instrument data format
ION-TOF	.itm	raw measurement data
ION-TOF	.ita	analysis data
Physical Electronics/ULVAC-PHI	.raw*	raw measurement data
Physical Electronics/ULVAC-PHI	.tdc	spectrum data

Open MS data formats

JCAMP-DX

This format was one of the earliest attempts to supply a standardized file format for data exchange in mass spectrometry. JCAMP-DX was initially developed for infrared spectrometry. JCAMP was officially released in 1988. JCAMP was found impractical for today's large MS data sets, but it is still used for exchanging moderate numbers of spectra.

ANDI-MS or netCDF

The Analytical Data Interchange Format for Mass Spectrometry is a format for exchanging data. ANDI was initially developed for chromatography-MS data and therefore was not used in the proteomics gold rush where new formats based on XML were developed.

mzData

mzData was the first attempt by the Proteomics Standards Initiative (PSI) from the Human Proteome Organization (HUPO) to create a standardized format for Mass Spectrometry data. This format is now deprecated, and replaced by mzML.

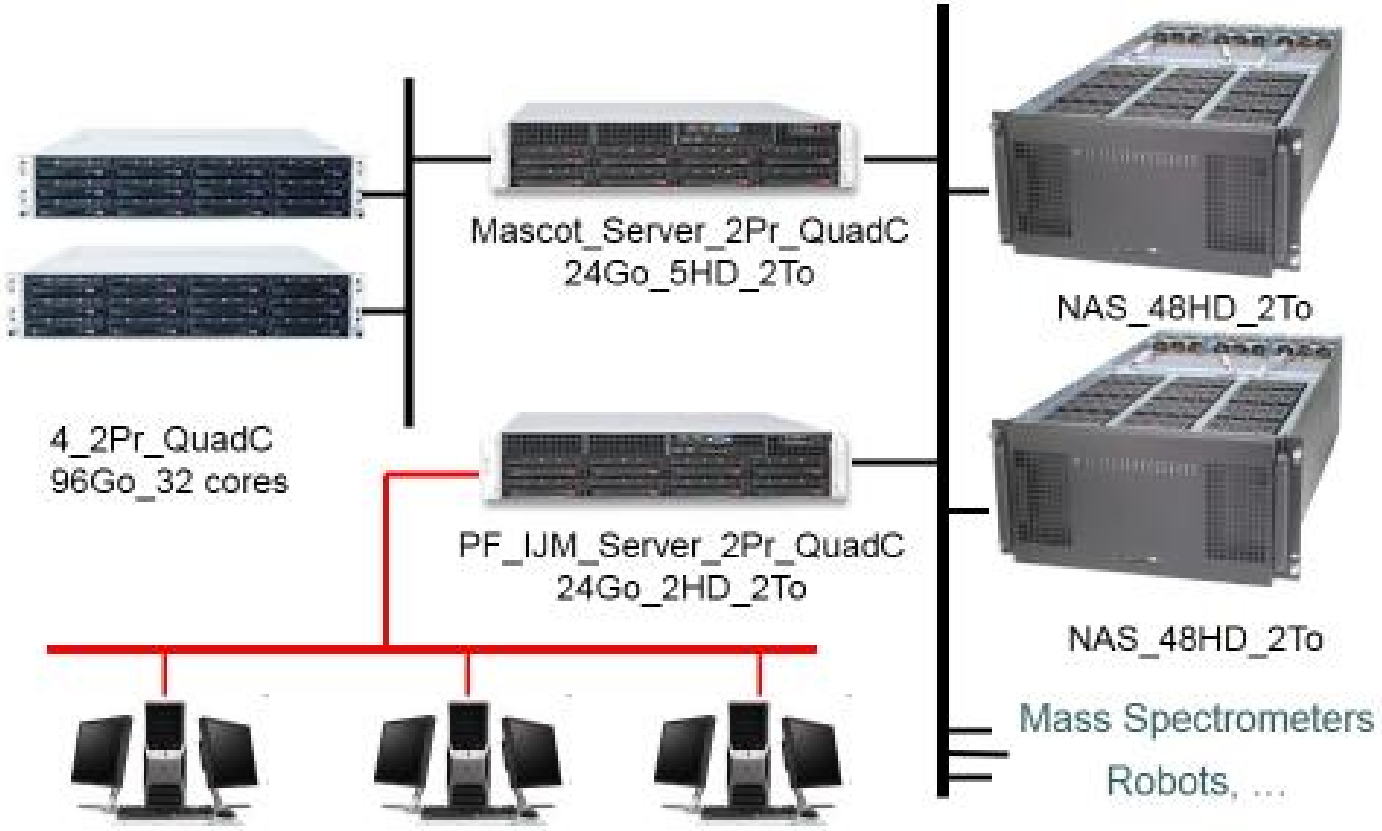
mzXML

mzXML is a XML (eXtensible Markup Language) based common file format for proteomics mass spectrometric data. This format was developed at the Seattle Proteome Center/Institute for Systems Biology while the HUPO-PSI was trying to specify the standardized mzData format, and is still in use in the proteomics community.

mzML

As two formats (mzData and mzXML) for representing the same information is an undesirable state, a joint effort was set by HUPO-PSI, the SPC/ISB and instrument vendors to create a unified standard borrowing the best aspects of both mzData and mzXML, and intended to replace them. The first specification was published in June 2008. This format was officially released at the 2008 American Society for Mass Spectrometry Meeting, and is since then relatively stable with very few updates. On 1 June 2009, mzML 1.1.0 was released. There are no planned further changes as of 2013.

Saving data and servers



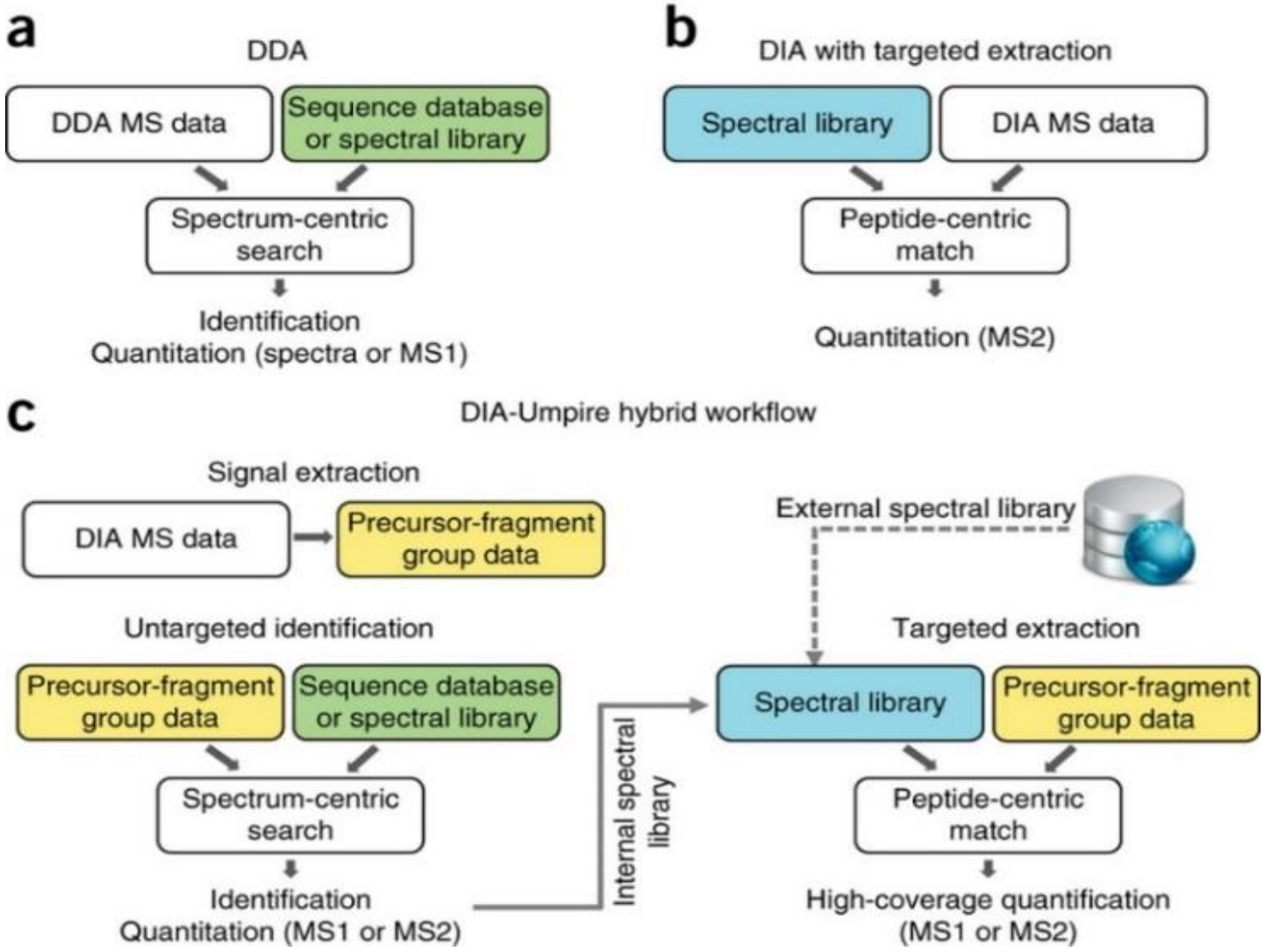
Progenesis QI for proteomics

PROTEOME DISCOVERER 1.4
Mass Informatics Platform for Protein Scientists
Thermo SCIENTIFIC

Byonic™
PROTEIN METRICS INC.

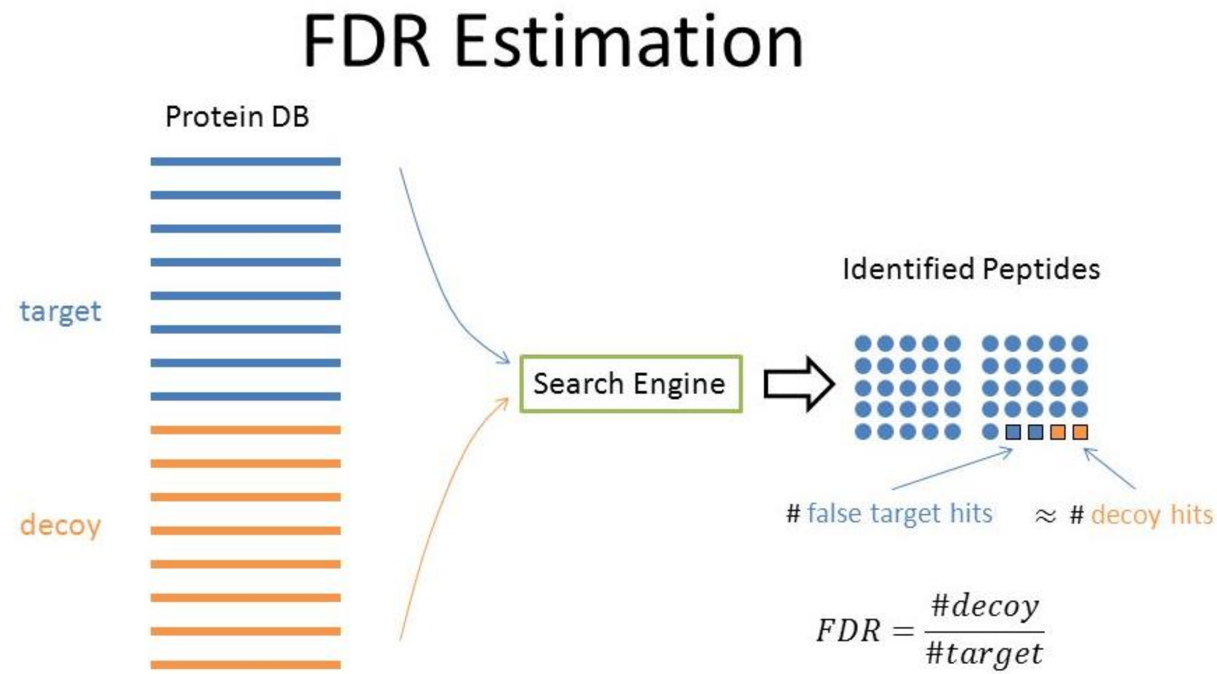
{MATRIX SCIENCE}

Search engine



Search engines and validation of peptides and proteins identifications

[illegible] (%) = ([illegible]) [illegible] 2 / [illegible] ([illegible] + [illegible])



Critical importance of mass accuracy for database searches

Expressed as Da or as ppm (10 ppm = 0,001%1 ppm = 0,0001%)

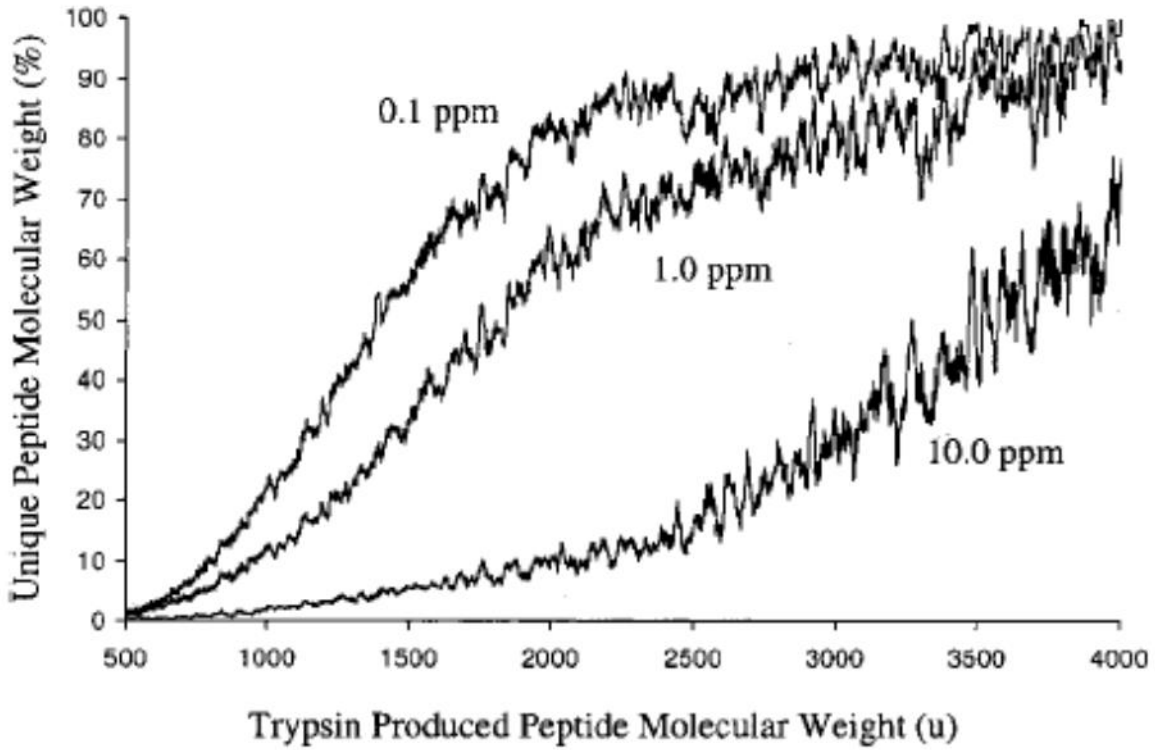
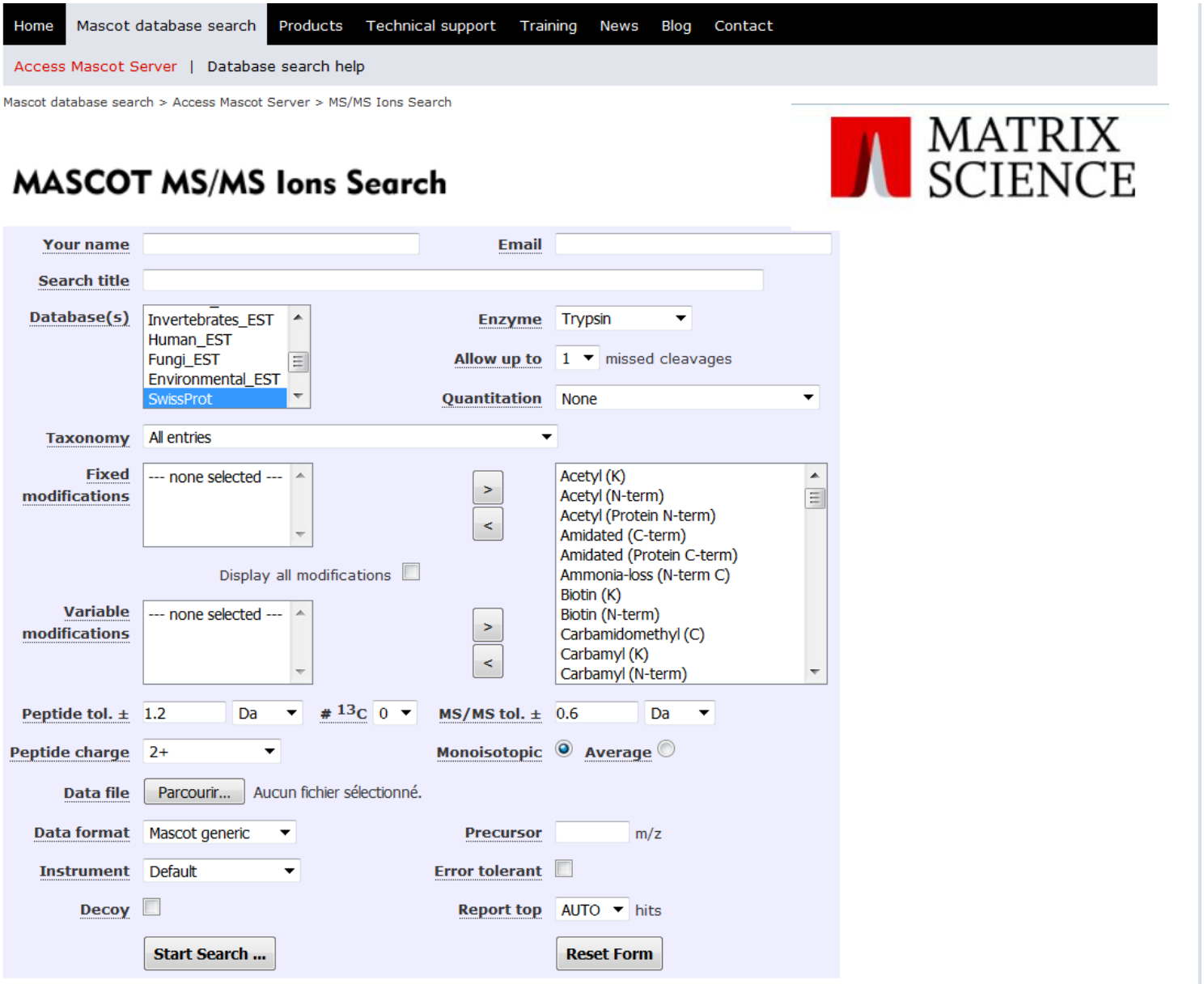
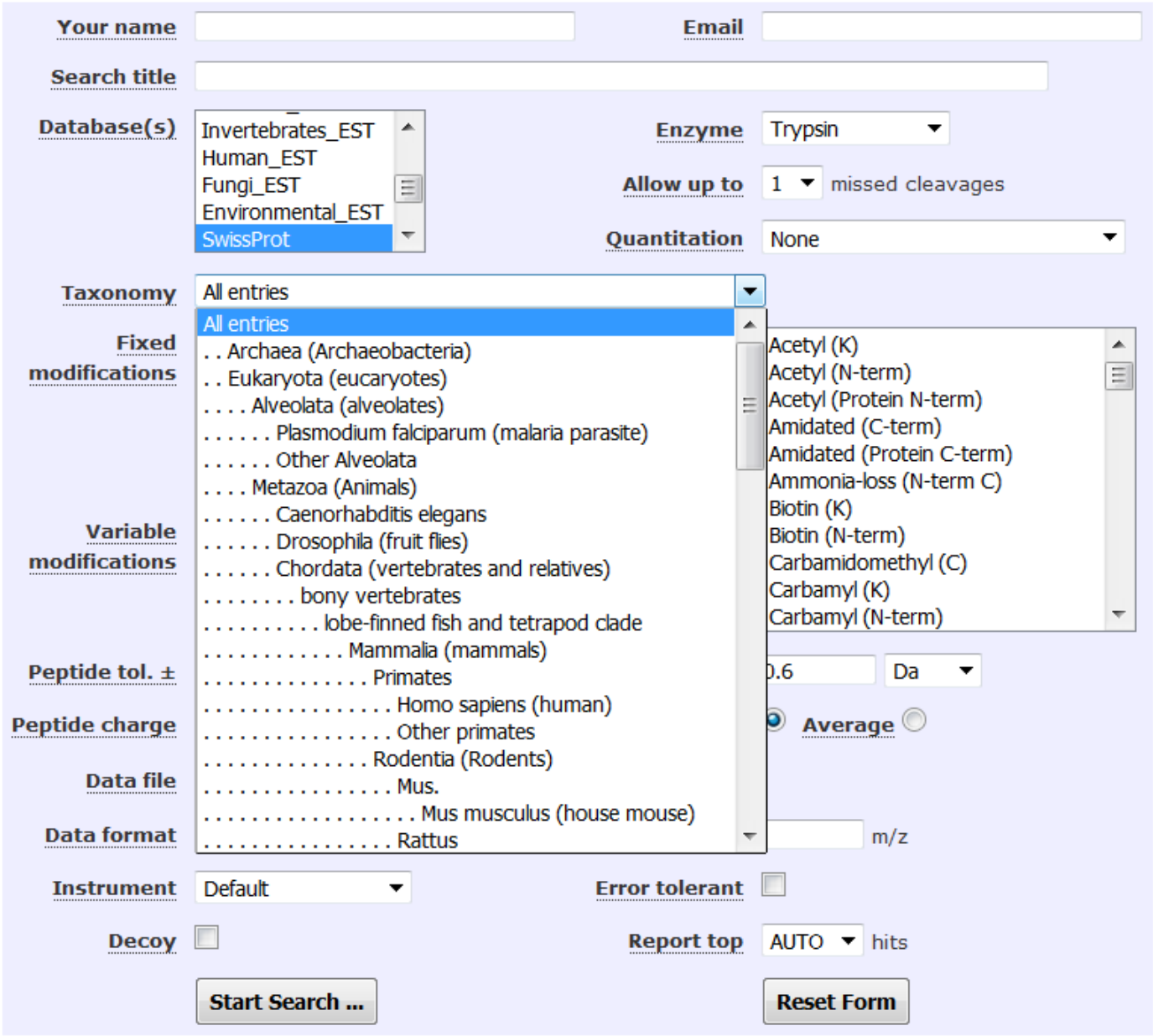


Figure 1. All possible unique peptide molecular weights after digestion of all yeast proteins in the National Center for Biotechnology Information at a mass accuracy of 0.1, 1.0, and 10.0 ppm.

A database search engine : Mascot



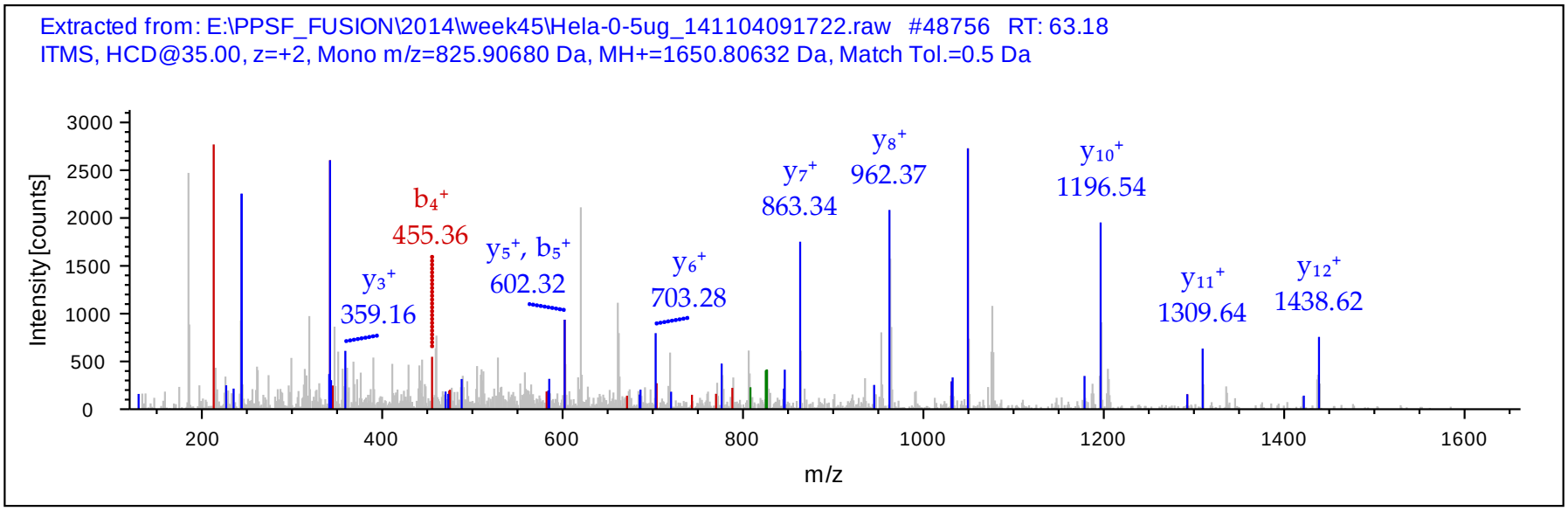
MASCOT MS/MS Ions Search



Search engine output formats

File name	File content
Processed peak lists	Heavily processed form of mass spectrometry data, usually derived from raw data files via various (semi-) automatic steps, e.g.: centroiding, deisotoping and charge deconvolution. These files are formatted in plain text, with typical formats like **dta**, **pkl**, **ms2** or **mgf**.
Search engine output files	These files contain the data and metadata generated by the software (called search engines) used for performing the identification and quantification of peptides and proteins. **Each search engine has its own specific output file format**. The outputs are typically formatted in either plain text or XML. mzIdentML - provides a common format for the export of identification results from any search engine. mzQuantML - provides a common format for the export of quantification results from any search engine. mzTab - represents both identification and basic quantification results. To allow a full representation of the processed results in the PRIDE database and in the PX tool, the search engine output files need to be converted to PRIDE XML. PRIDE Converter and PRIDE Converter 2 are the two tools developed by the PRIDE team to make this conversion possible.
Protein/peptide identifications	Proteomics mass spectra can be matched to peptides or proteins, resulting in identifications for those spectra. Typically a spectrum is considered to have been identified if the score attributed to a peptide or protein match qualifies against an *a priori* or *a posteriori* defined threshold. In the case of fragmentation spectra, the initial identification will consist of a peptide sequence; subsequent steps will derive a list of proteins from the identified peptides. The protein assembly step can be a discernible process with its own input and output files, or it can be implicit in the overall identification software.

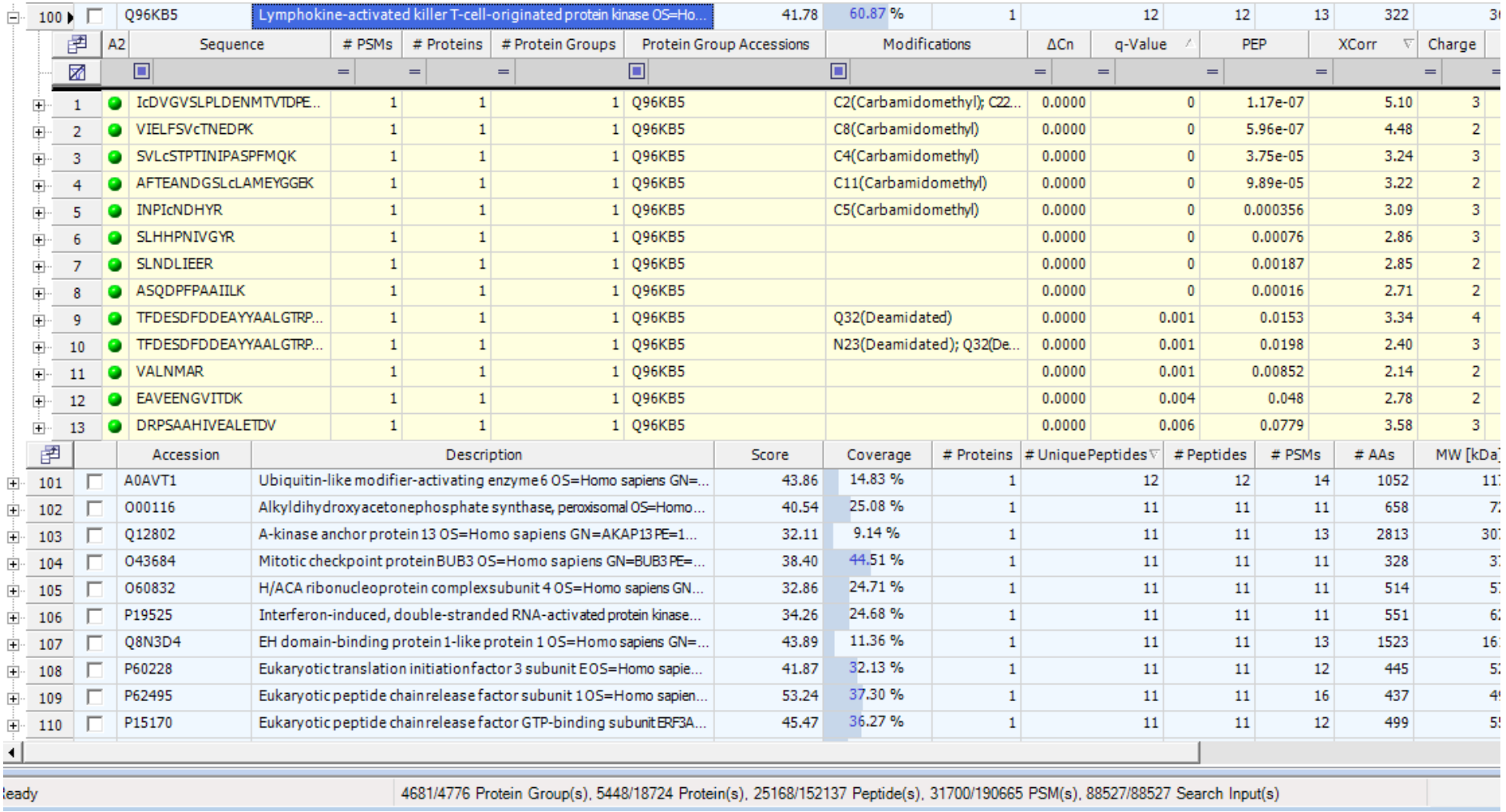
31700 MS/MS spectra interpreted!!!!



Sequence: VIELFSVCTNEDPK, C8-Carbamidomethyl (57.02146 Da)
Charge: +2, Monoisotopic m/z: 825.90680 Da (+0.95 mmu/+1.15 ppm), MH+: 1650.80632 Da, RT: 63.18 min,
Identified with: Sequest HT (v1.3); XCorr:4.48, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3

Protein references (1):
- Lymphokine-activated killer T-cell-originated protein kinase OS=Homo sapiens GN=PBK PE=1 SV=3 - [TOPK_HUMAN]

5448 identified proteins



		Q96KB5	Lymphokine-activated killer T-cell-originated protein kinase OS=Ho...	41.78	60.87 %	1	12	12	13	322	3

	A2	Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	ΔCn	q-Value	PEP	XCorr	Charge
1		IcDVGVSLPLDENMTVTDPE...	1	1	1	Q96KB5	C2(Carbamidomethyl); C22...	0.0000	0	1.17e-07	5.10	3
2		VIELFSVcTNEDPK	1	1	1	Q96KB5	C8(Carbamidomethyl)	0.0000	0	5.96e-07	4.48	2
3		SVLcSTPTINIPASPFMQK	1	1	1	Q96KB5	C4(Carbamidomethyl)	0.0000	0	3.75e-05	3.24	3
4		AFTEANDGSLcLAMEYGGEK	1	1	1	Q96KB5	C11(Carbamidomethyl)	0.0000	0	9.89e-05	3.22	2
5		INPIcNDHYR	1	1	1	Q96KB5	C5(Carbamidomethyl)	0.0000	0	0.000356	3.09	3
6		SLHHPNIVGYR	1	1	1	Q96KB5		0.0000	0	0.00076	2.86	3
7		SLNDLIEER	1	1	1	Q96KB5		0.0000	0	0.00187	2.85	2
8		ASQDPFPAAIILK	1	1	1	Q96KB5		0.0000	0	0.00016	2.71	2
9		TFDESDFDDEAYYAALGTRP...	1	1	1	Q96KB5	Q32(Deamidated)	0.0000	0.001	0.0153	3.34	4
10		TFDESDFDDEAYYAALGTRP...	1	1	1	Q96KB5	N23(Deamidated); Q32(De...	0.0000	0.001	0.0198	2.40	3
11		VALNMAR	1	1	1	Q96KB5		0.0000	0.001	0.00852	2.14	2
12		EAVEENGVITDK	1	1	1	Q96KB5		0.0000	0.004	0.048	2.78	2
13		DRPSAAHIVEALETDV	1	1	1	Q96KB5		0.0000	0.006	0.0779	3.58	3

	Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	# AAs	MW [kDa]
101	A0AVT1	Ubiquitin-like modifier-activating enzyme 6 OS=Homo sapiens GN=...	43.86	14.83 %	1	12	12	14	1052	11
102	O00116	Alkyldihydroxyacetonephosphate synthase, peroxisomal OS=Homo...	40.54	25.08 %	1	11	11	11	658	7
103	Q12802	A-kinase anchor protein 13 OS=Homo sapiens GN=AKAP13 PE=1...	32.11	9.14 %	1	11	11	13	2813	30
104	O43684	Mitotic checkpoint protein BUB3 OS=Homo sapiens GN=BUB3 PE=...	38.40	44.51 %	1	11	11	11	328	3
105	O60832	H/ACA ribonucleoprotein complex subunit 4 OS=Homo sapiens GN...	32.86	24.71 %	1	11	11	11	514	5
106	P19525	Interferon-induced, double-stranded RNA-activated protein kinase...	34.26	24.68 %	1	11	11	11	551	6
107	Q8N3D4	EH domain-binding protein 1-like protein 1 OS=Homo sapiens GN=...	43.89	11.36 %	1	11	11	13	1523	16
108	P60228	Eukaryotic translation initiation factor 3 subunit E OS=Homo sapie...	41.87	32.13 %	1	11	11	12	445	5
109	P62495	Eukaryotic peptide chain release factor subunit 1 OS=Homo sapien...	53.24	37.30 %	1	11	11	16	437	4
110	P15170	Eukaryotic peptide chain release factor GTP-binding subunit ERF3A...	45.47	36.27 %	1	11	11	12	499	5

eady

4681/4776 Protein Group(s), 5448/18724 Protein(s), 25168/152137 Peptide(s), 31700/190665 PSM(s), 88527/88527 Search Input(s)

Key questions in proteomics

- What is the protein content of my biological sample?

  => problem of identification

- What is the abundance of my protein of interest?

  => quantification problem

- Relative question: What are the protein abundance variations of the proteomes studied?

- What are the partners of my protein of interest?

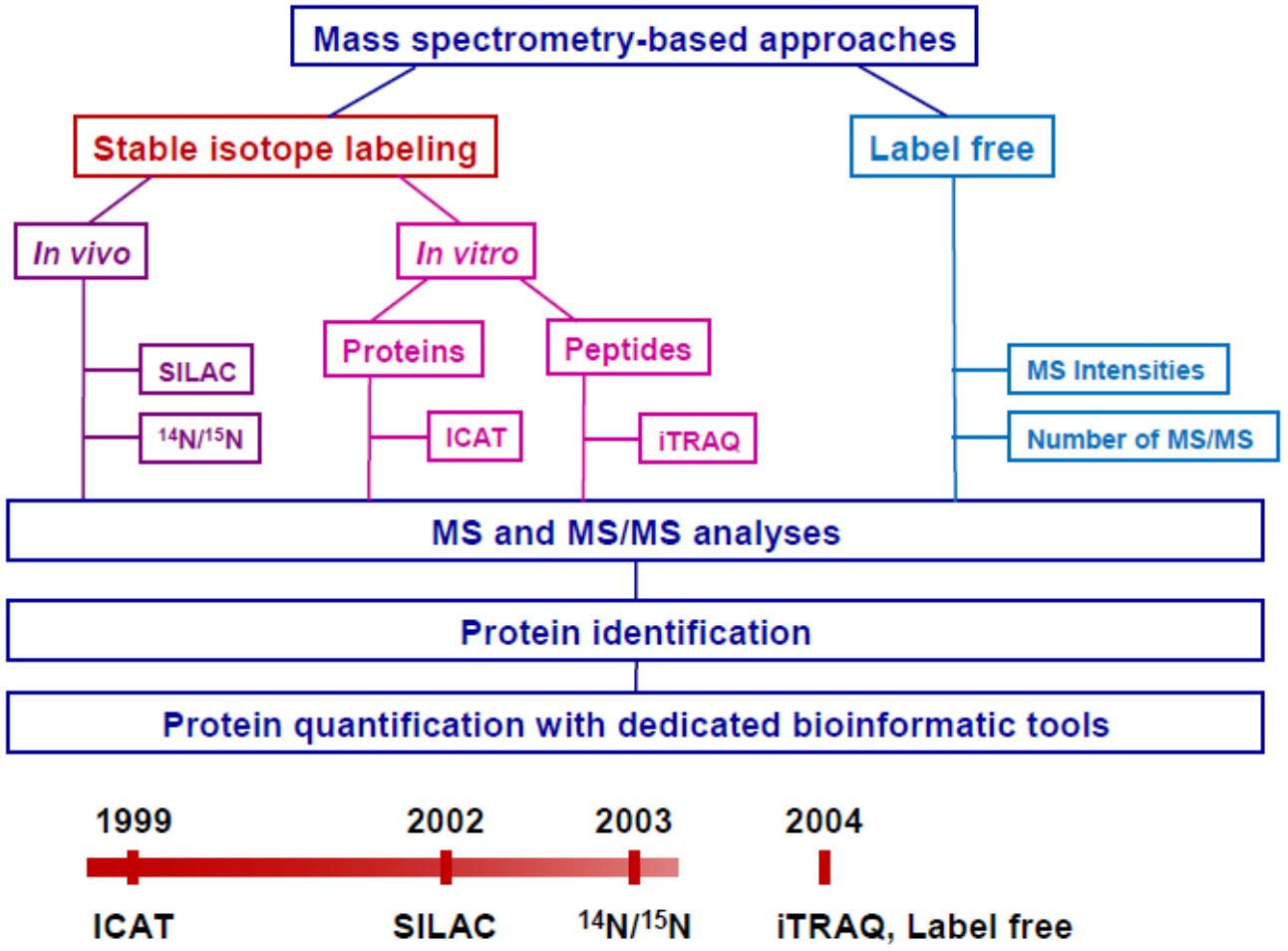
- Are there any signature proteins related to a particular biological process?

  => biomarkers identifications and quantifications

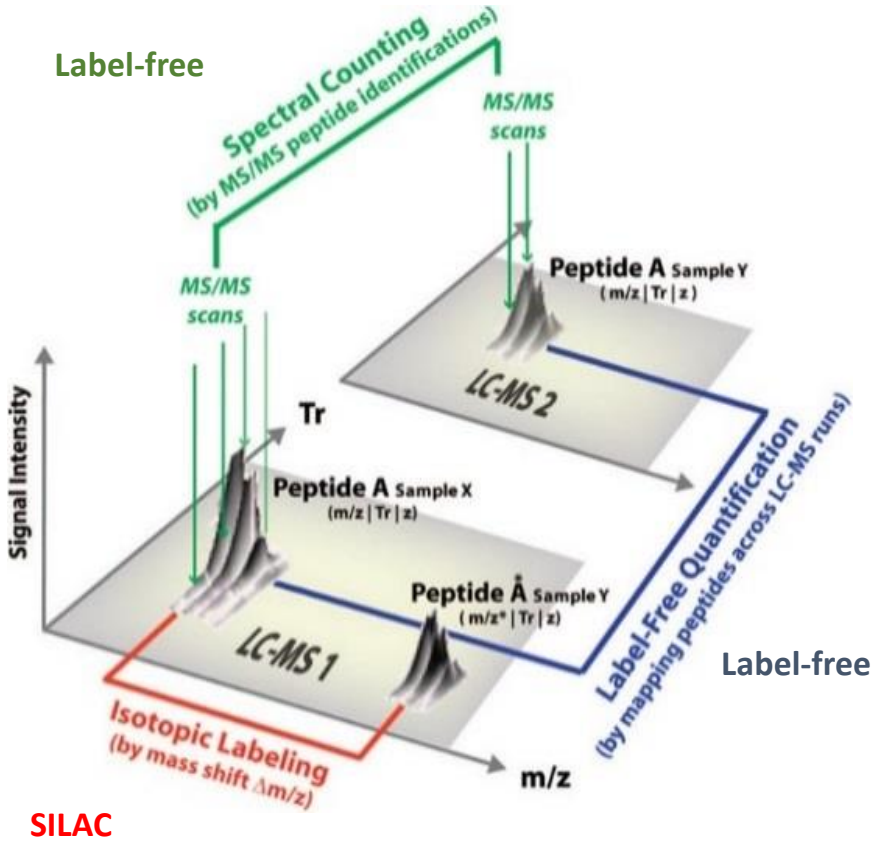
Quantitative proteomics

- Relative quantification
  - Stable isotopes labelling
  - Label-free
  - Metabolic labeling
- Absolute quantification

Quantitative proteomics

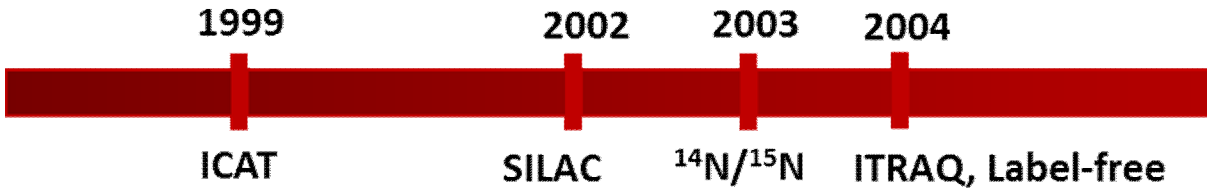


Quantitative proteomics in bottom-up

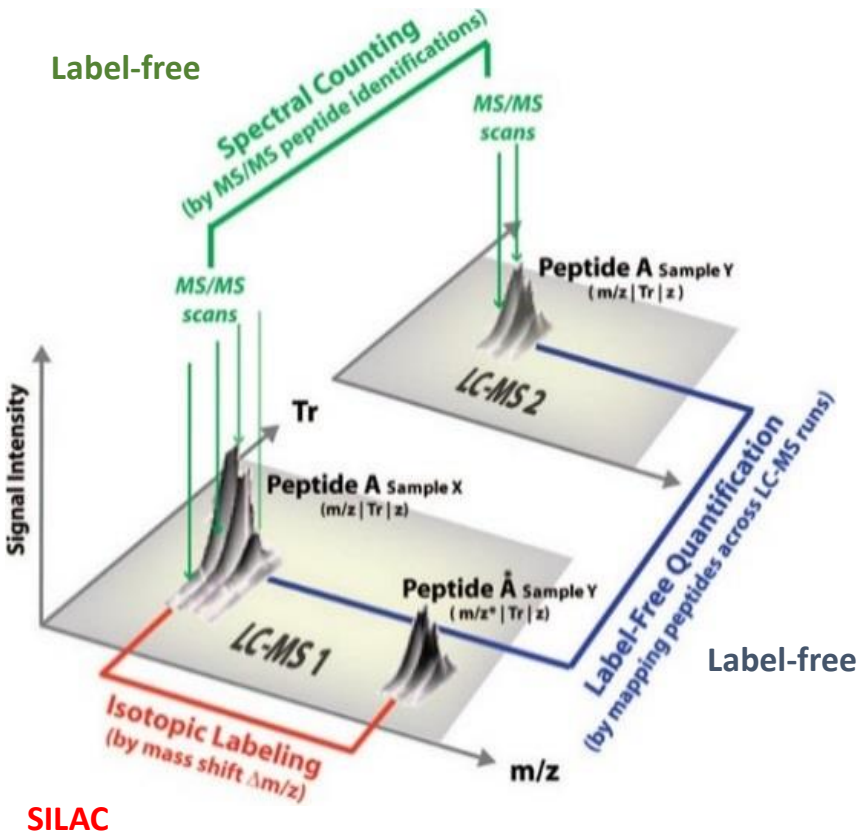


Advantages/Limitations:

- ❑ Label-free
- ❑ Metabolic Labeling (SILAC, $^{14}N/^{15}N$ – ^{13}C labeling)
- ❑ Chemical labeling (TMT,ITRAQ)



Quantitative proteomics: label-free

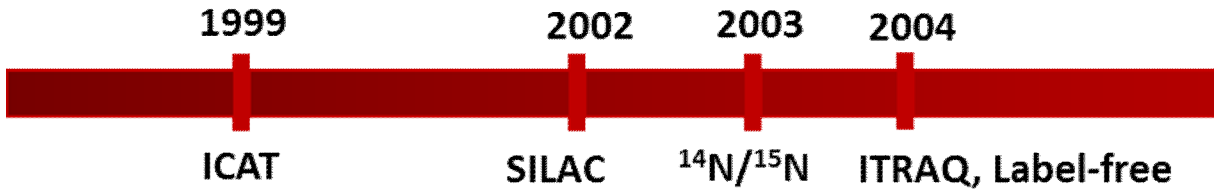


Advantages/Limitations:

❑ Label-free:

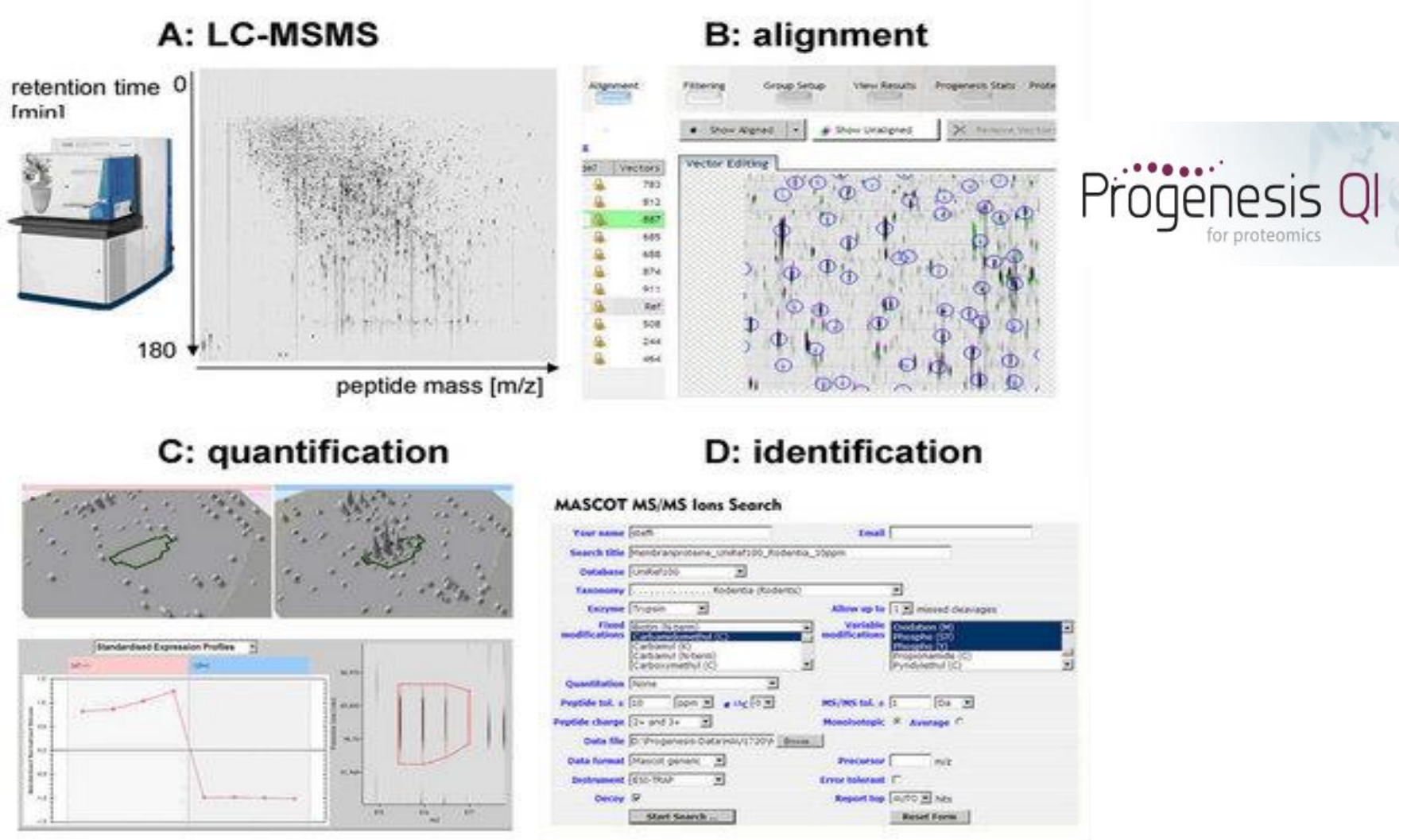
- Simplicity
- Number of identifications
- Reproductibility between runs
- Number of samples to run

Review for Label-free and yeasts:
Leger et al. Methods Mol Biol (2016)



Quantitative proteomics without labeling

Quantification label-free basée sur les intensités MS



Quantitative proteomics without labeling : results

Experiment Design

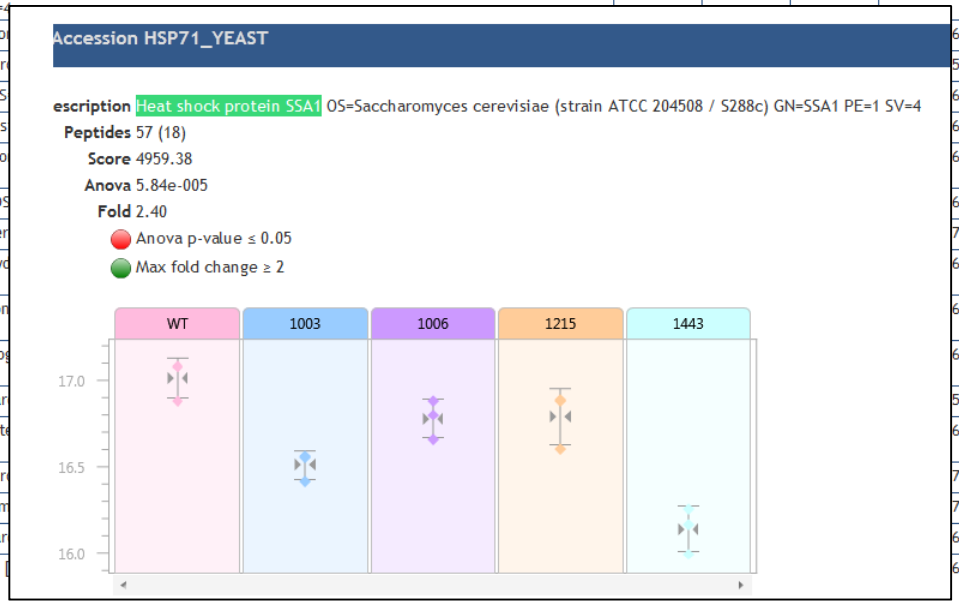
Condition	WT	1003	1006	1215	1443
Replicates	3	3	3	3	3

Proteins

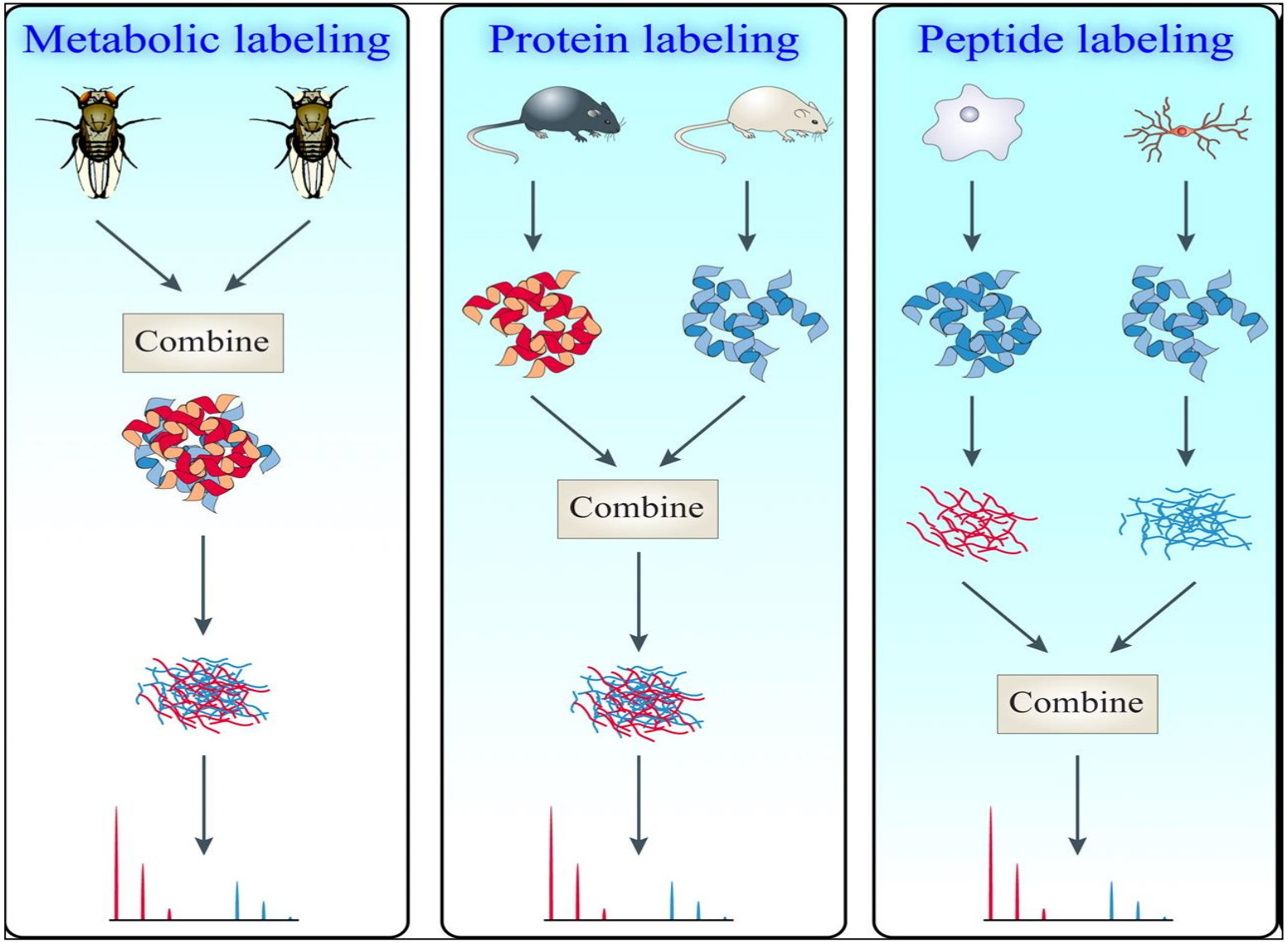
Protein building options

Protein grouping **Group similar proteins**
Protein quantitation **Using only features with no protein conflicts**

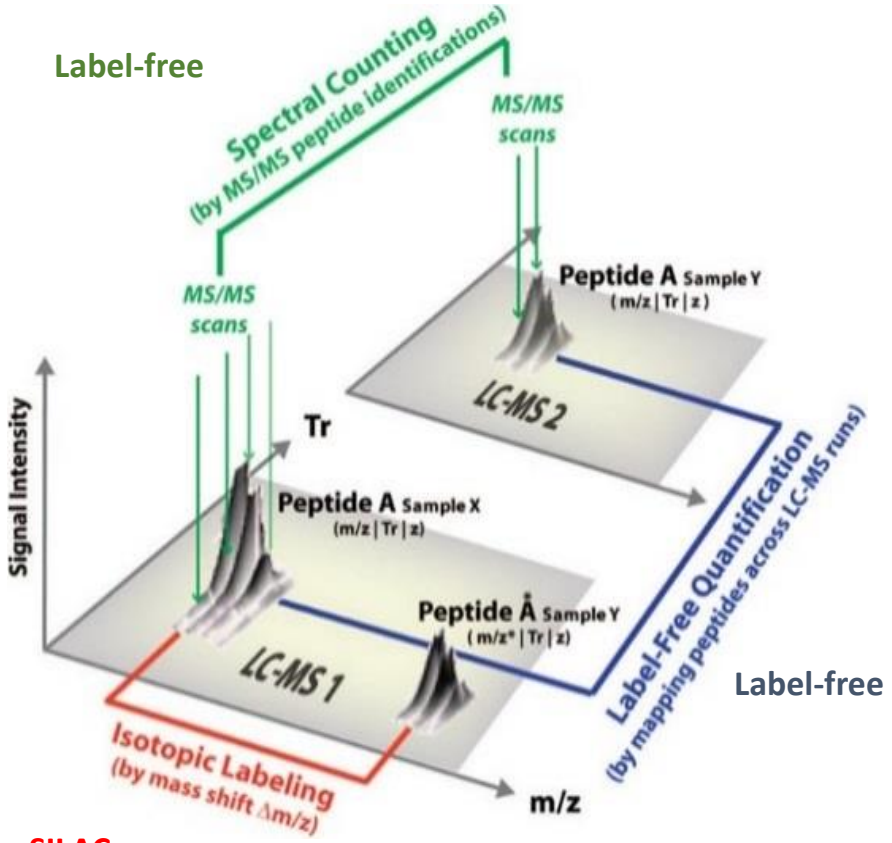
Accession	Peptides	Score	Anova (p)*	Fold	Tags	Description	Average Normalised Abundances				
							WT	1003	1006	1215	1443
HSP71_YEAST	57 (18)	4959.38	5.84e-005	2.40		Heat shock protein SSA1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=SSA1 PE=1 SV=4	1.23e+007	7.42e+006	9.72e+006	9.87e+006	5.13e+006
EF2_YEAST	69	4650.11	7.26e-004	2.16		Elongation factor 2 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=EFT1 PE=1 SV=1	3.40e+007	2.39e+007	2.89e+007	3.28e+007	1.58e+007
FAS1_YEAST	74 (71)	4506.21	9.72e-003	2.33		Fatty acid synthase subunit beta OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=FAS1 PE=1 SV=2	1.03e+007	6.55e+006	8.51e+006	1.01e+007	1.53e+007
EF3A_YEAST	58 (44)	3816.91	7.06e-006	3.32		Elongation factor 3A OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YEF3 PE=1 SV=4	2.88e+007	1.29e+007	2.33e+007	2.72e+007	8.66e+006
METE_YEAST	46	3373.13	1.85e-006	13.40		5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=MET6 PE=1 SV=4	4.34e+006	5.55e+006	3.77e+006	5.16e+006	5.05e+007
HS104_YEAST	53	3190.19	6.84e-004	2.29		Heat shock protein 104 OS=Saccharo					3.51e+006
HSP75_YEAST	40 (1)	3062.55	1.21e-006	28.16		Heat shock protein SSB1 OS=Saccharo					1.05e+004
HSP7F_YEAST	39 (32)	2658.69	2.58e-004	2.01		Heat shock protein homolog SSE1 OS					3.50e+006
ENO1_YEAST	31 (15)	2367.12	6.07e-005	2.19		Enolase 1 OS=Saccharomyces cerevis					1.46e+007
ATPA_YEAST	32	2341.09	3.17e-006	2.59		ATP synthase subunit alpha, mitocho SV=5					3.66e+006
SYLC_YEAST	37	2176.12	1.52e-006	2.01		Leucine--tRNA ligase, cytoplasmic OS					1.82e+006
HXKA_YEAST	29 (28)	2162.35	3.17e-004	2.88		Hexokinase-1 OS=Saccharomyces cer					4.61e+006
ALDH6_YEAST	30	2091.58	4.85e-004	2.15		Magnesium-activated aldehyde dehyd GN=ALD6 PE=1 SV=4					2.50e+006
ATPB_YEAST	28	2015.82	4.45e-006	2.39		ATP synthase subunit beta, mitochon SV=2					4.05e+006
G3P1_YEAST	31 (21)	1986.15	8.75e-005	4.16		Glyceraldehyde-3-phosphate dehydrog PE=1 SV=3					1.79e+007
HSP74_YEAST	26 (12)	1750.55	0.04	2.68		Heat shock protein SSA4 OS=Sacchar					5.24e+005
PUR92_YEAST	28 (22)	1725.94	6.21e-007	7.07		Bifunctional purine biosynthesis prote PE=1 SV=2					9.63e+006
ADH1_YEAST	24 (17)	1689.13	5.88e-004	2.62		Alcohol dehydrogenase 1 OS=Sacchar					1.03e+007
HSP26_YEAST	18	1538.64	2.83e-006	2.31		Heat shock protein 26 OS=Saccharom					8.43e+006
SAHH_YEAST	27	1535.76	2.79e-006	3.51		Adenosylhomocysteinase OS=Sacchar					1.31e+007
PCKA_YEAST	20	1515.31	3.42e-009	9.67		Phosphoenolpyruvate carboxykinase [SV=2					6.25e+005



Quantitative proteomics with labeling



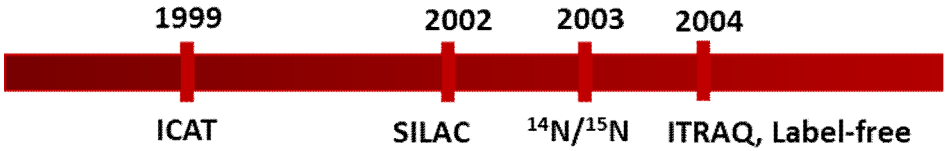
Quantitative proteomics: metabolic labeling



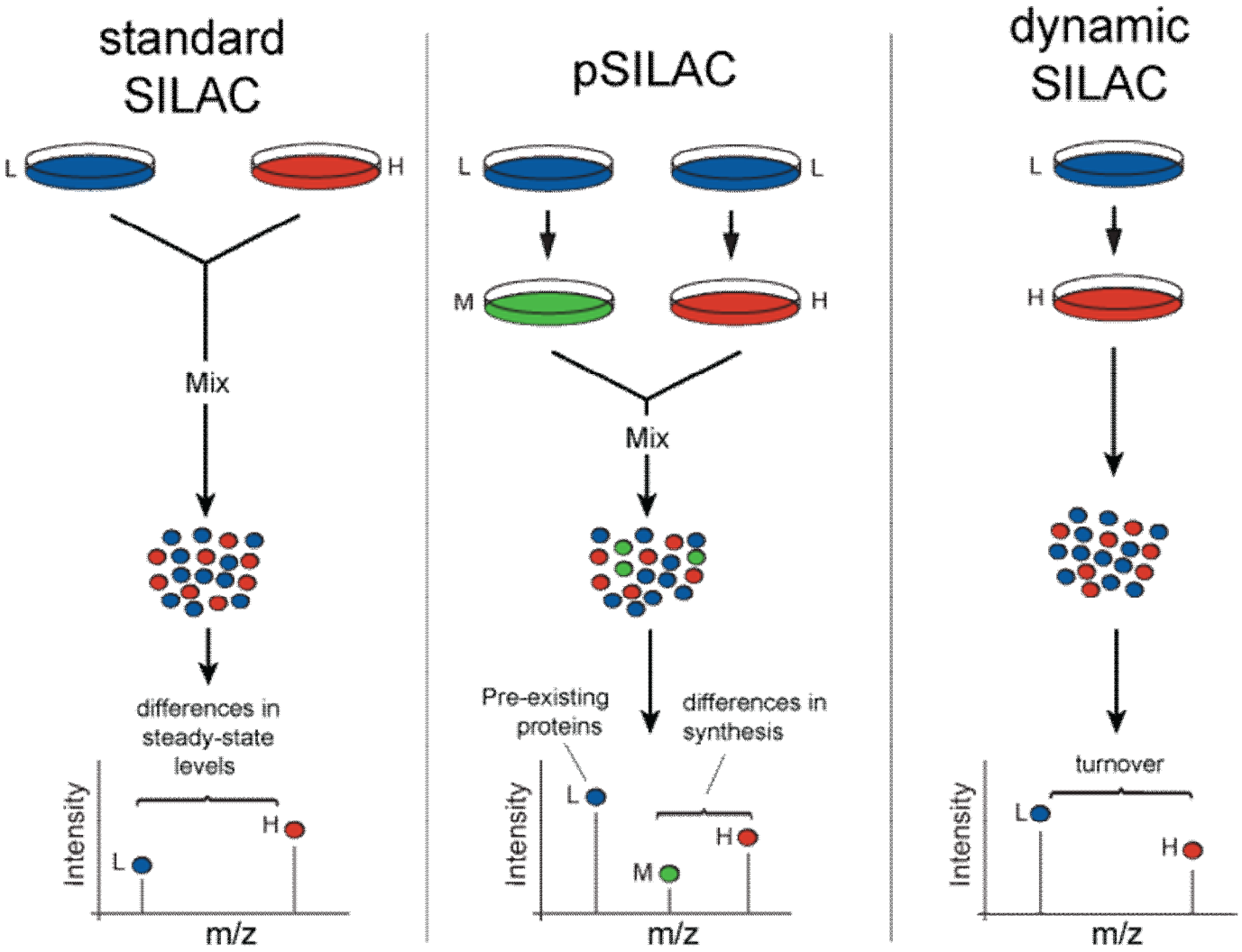
Advantages/Limitations:

- ❑ SILAC:
  - Multiplexing
  - Reproductibility
  - **2 peaks instead of 1 to analyze by the MS (for 2 samples)**
  - Less identifications
  - Partial labeling
  - Arginine/proline conversion (use of mutants)
  - Trypsin exclusively

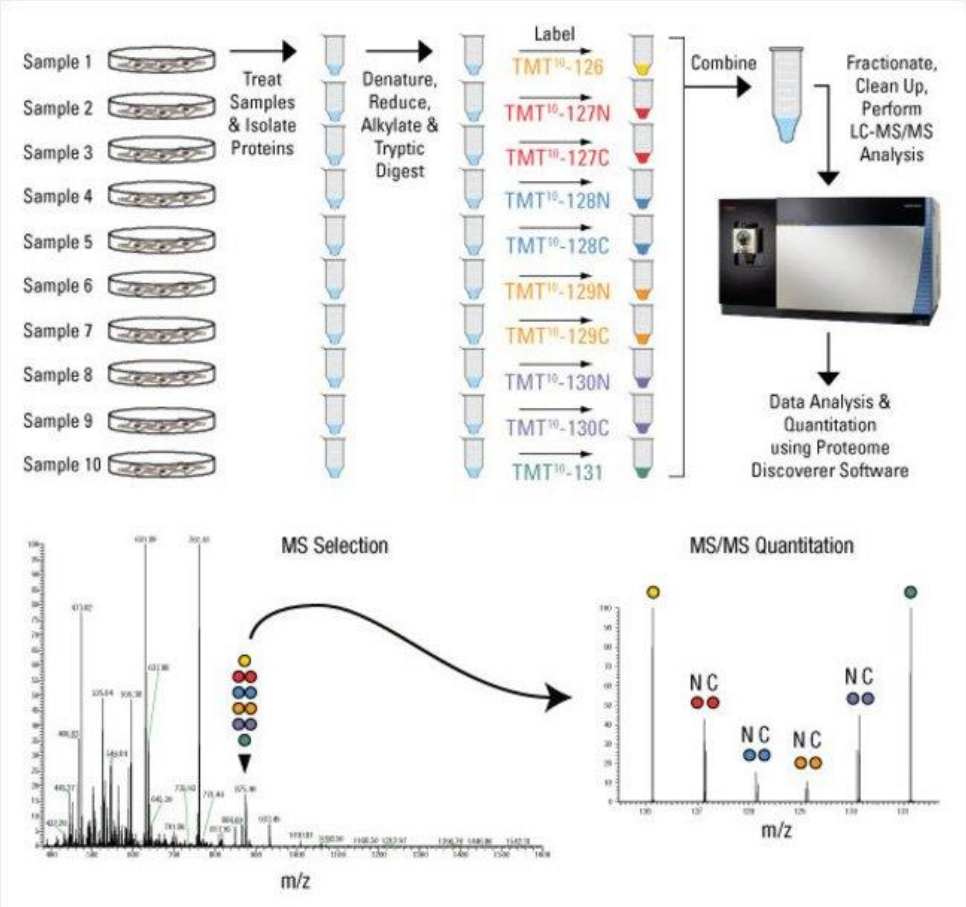
- ❑ $^{14}N/^{15}N$ – ^{13}C labeling:
  - Multiplexing
  - Reproductibility
  - **2 peaks instead of 1 to analyze by the MS (for 2 samples)**
  - Less identifications and quantifications
  - Partial labeling
  - **Variable mass shift between heavy and light forms**



SILAC approaches



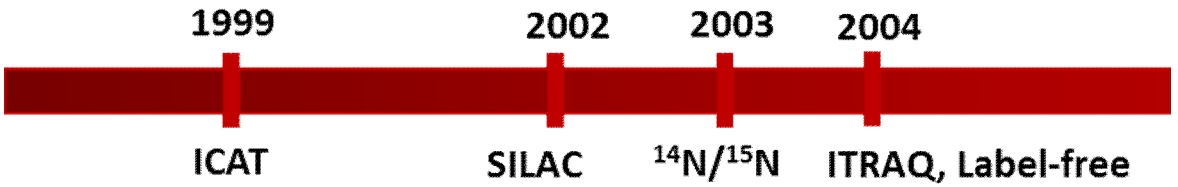
Quantitative proteomics: chemical labeling



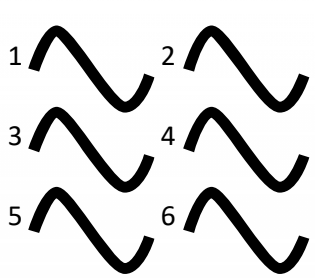
Advantages/Limitations:

❑ Chemical labeling (TMT, ITRAQ)

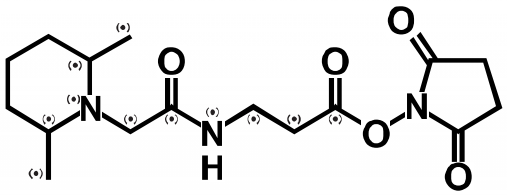
- Multiplexing (until 11plex)
- Reproductibility
- Quantification in MS2 or MS3
- **1 peak instead of N (for N samples) to analyse in MS**
- Amount of materials for the peptide labeling
- Need of resolution in MS2 for quantifications
- Incomplete labeling
- Less identifications and quantifications



TMT labeling : principles



Peptides



Labeled Peptides

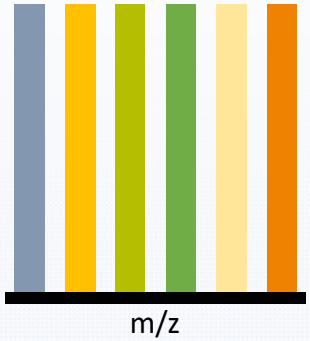
LC-MS/MS

Full MS

One Signal

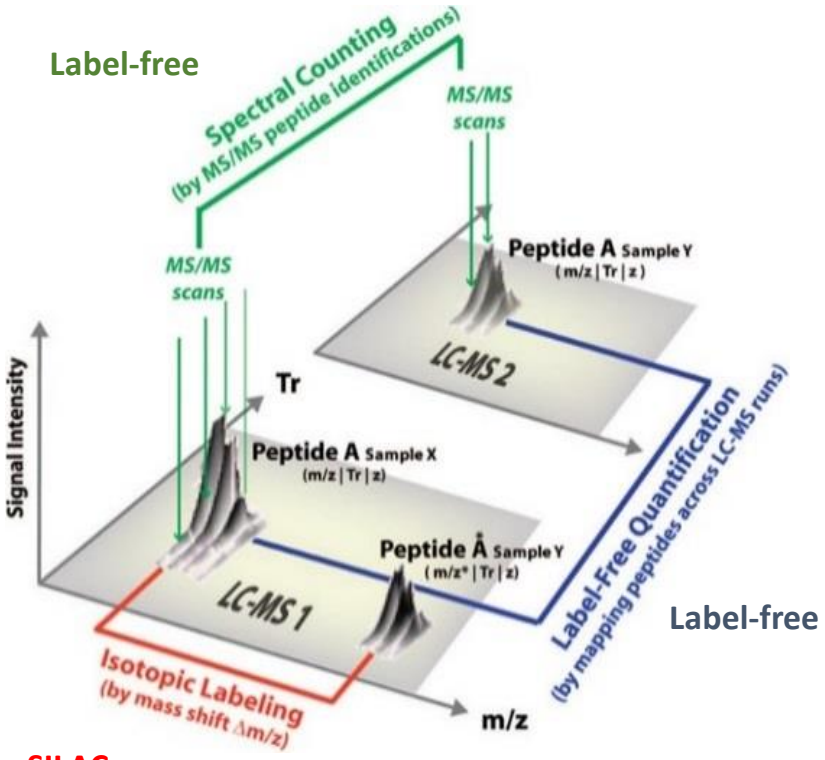
HCD FRAGMENTATION

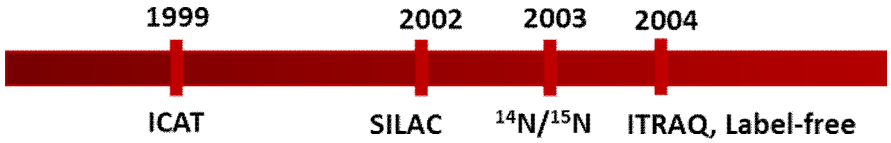
MS/MS



6 Reporter Ions

Quantitative proteomics in bottom-up





Advantages/Limitations:

- ❑ Label-free:
  - Simplicity
  - Number of identifications
  - Reproductibility between runs
  - Number of samples to run

- ❑ SILAC:
  - Multiplexing
  - Reproductibility
  - **2 peaks instead of 1 to analyze by the MS (2 samples)**
  - Less identifications
  - Partial labeling
  - Arginine/proline conversion (use of mutants)
  - Trypsin exclusively

- ❑ $^{14}N/^{15}N - ^{13}C$ labeling:
  - Multiplexing
  - Reproductibility
  - **2 peaks instead of 1 to analyze by the MS (2 samples)**
  - Less identifications and quantifications
  - Partial labeling
  - **Variable mass shift between heavy and light forms**

Metabolic labeling

- ❑ Chemical labeling (TMT,ITRAQ)
  - Multiplexing
  - Reproductibility
  - 1 peak instead of N to analyse in MS (N samples)
  - Amount of materials for the peptide labeling
  - Need of resolution in MS2 for quantifications
  - Incomplete labeling
  - Less identifications and quantifications

Quantitative proteomics without labeling : export

Protein level!

Progenesis QI for proteomics

													Normalized abundance							
													Sain						Patient	
Accession	Peptide cou	Unique pep	Confidence	Anova (p)	q Value	Max fold ch	Power	Highest me	Lowest mea	Mass	Description	1845007-F1	1845007-F3	1845007-F5	1845007-F7	1845007-F9	1845007-F1	1845007-F2	1845007-F4
P40197	12	12	545.23	2.48E-07	4.74E-05	32.042418	1	Patient	Sain	60.921	Platelet glycoprotein V OS	1447.8857	898.47877	2114.249	3517.982	2506.9091	1186.7152	33178.605	73995.78
P02776	2	2	134.77	2.63E-07	4.74E-05	78.817355	1	Patient	Sain	10.838	Platelet factor 4 OS=Homc	3812.4369	3755.8358	1044.3911	2939.4867	3862.0883	1110.5549	333829.66	68071.51
Q13201	6	6	221.4	5.33E-07	6.40E-05	72.663027	1	Patient	Sain	138.023	Multimerin-1 OS=Homo sa	160.48528	939.93933	155.7663	732.57482	572.39757	196.60688	75519.303	23299.12
P04114	336	334	28302.18	7.99E-06	0.0006593	2.1384481	1	Patient	Sain	515.283	Apolipoprotein B-100 OS=	14117253	10601958	8472023.6	9982572.2	10056625	11898407	24415094	1744463
P07996	42	42	2678.71	9.16E-06	0.0006593	143.08362	1	Patient	Sain	129.3	Thrombospondin-1 OS=Hc	11604.367	10346.75	4059.3393	5652.5533	2121.1343	2367.1543	2303007.2	83175.08
Q15485	7	7	289.57	1.42E-05	0.0007087	24.427918	0.9999997	Patient	Sain	33.98	Ficolin-2 OS=Homo sapier	8778.176	6132.5206	10104.107	8194.6312	8690.2929	1737.9715	185819.2	105024.9
P10720	3	3	208.63	1.49E-05	0.0007087	36.552808	0.9999997	Patient	Sain	11.545	Platelet factor 4 variant O	32881.787	11758.269	1451.6395	5448.2604	4187.5362	4196.5766	546497.96	239272.0
Q12884	2	2	70.47	1.58E-05	0.0007087	Infinity	0.9999996	Patient	Sain	87.657	Prolyl endopeptidase FAP	0	0	0	0	0	0	961.55592	106.477
Q15061	1	1	3.67	3.71E-05	0.0014832	23.811179	0.9999905	Sain	Patient	74.843	WD repeat-containing prc	567513.12	751475.75	852125.74	1189728.8	625225.18	78135.037	46681.126	36585.06
P04075	8	8	371.85	4.58E-05	0.0016488	4.2937733	0.9999814	Patient	Sain	39.395	Fructose-bisphosphate al	33219.801	26234.171	28730.699	26696.908	33816.665	32024.353	128602.25	168571.3
P09486	10	9	463.23	6.25E-05	0.0020451	6.1746973	0.999953	Patient	Sain	34.61	SPARC OS=Homo sapiens (	9522.5101	37758.131	36721.019	20388.641	15117.423	28228.93	299004.24	97804.24
P10124	1	1	70.16	7.11E-05	0.002134	272.44302	0.9999321	Patient	Sain	17.641	Serglycin OS=Homo sapier	67.740323	0	958.26875	53.15204	160.31231	238.84011	105180.09	42246.24
P02775	6	6	388.98	8.21E-05	0.0022557	28.012451	0.9998996	Patient	Sain	13.885	Platelet basic protein OS=	61607.9	109352.83	105910.89	67216.894	60865.635	22752.188	3524260.5	1692399.5
Q9H1K0	2	2	12.94	8.77E-05	0.0022557	1.8919044	0.9998799	Patient	Sain	88.815	Rabenosyn-5 OS=Homo sa	871721.79	1194024.5	1103071.3	1121288.3	1020557.4	1361492.9	1506333.1	1757466.5
P35542	4	4	338.41	0.0001236	0.0029659	16.656374	0.9997122	Patient	Sain	14.737	Serum amyloid A-4 protein	192222.76	102064.74	9416.0169	8023.3014	46690.651	107524.48	1543185.2	1741696.5
P80188	6	6	223.58	0.0001847	0.0041561	9.1194461	0.9992657	Patient	Sain	22.574	Neutrophil gelatinase-as	2265.4484	1473.9176	6203.72	3676.0015	7827.3304	4603.3796	66633.817	26485.80
P02144	4	4	209.23	0.0002392	0.0050651	6.5032553	0.9987201	Patient	Sain	17.173	Myoglobin OS=Homo sapi	2754.3333	2223.1201	1035.6898	1813.5104	1462.164	5472.3528	17977.389	20813.25
P05067	5	5	179.76	0.000297	0.0059397	18.947351	0.9980148	Patient	Sain	86.888	Amyloid beta A4 protein C	408.20926	53.510036	338.118	494.28658	258.47482	516.31147	8561.0101	2161.7757
Q9NPH3	7	7	223.36	0.0003417	0.0064734	2.9534965	0.9973949	Patient	Sain	65.377	Interleukin-1 receptor acc	19269.326	23550.392	17918.705	19071.463	29165.24	20448.453	74326.28	82743.303
P22352	7	7	322.73	0.0005632	0.0101356	2.3655577	0.9936247	Patient	Sain	25.537	Glutathione peroxidase 3	356078.09	500176.63	411199.92	354267.54	608490.2	620677.21	1139291.1	1214566.8
A0A075B6I0	1	1	30.88	0.0005981	0.0102519	15.884721	0.9929515	Sain	Patient	12.806	Immunoglobulin lambda	14017.761	25223.891	6493.2483	70640.84	15600.774	36585.067	4641.3232	1487.298
P05155	33	33	2513.9	0.0007389	0.0120898	5.4023431	0.9900925	Patient	Sain	55.119	Plasma protease C1 inhib	741421.19	592478.75	377555.54	465885.04	2331779.9	7791080.5	11109177	10511766
A0A0C4DH2	2	2	100.38	0.0013344	0.0208833	3.8865369	0.9766266	Sain	Patient	12.999	Immunoglobulin heavy va	50500.146	64330.439	38938.111	121804.39	27881.723	32180.678	24057.435	17302.863
P14780	6	6	207.75	0.0014988	0.0221418	62.066206	0.972746	Patient	Sain	78.408	Matrix metalloproteinase	794.97907	136.39171	262.24899	313.84328	0	555.64619	58928.664	9636.858
P02649	27	27	1941.9	0.0015378	0.0221418	2.4710215	0.971824	Patient	Sain	36.132	Apolipoprotein E OS=Hom	2552396.1	1536331.8	853676.88	1079240.3	1188651.7	1327575.2	2857655.7	2177433.7
Q3C1V8	1	1	14.2	0.0019728	0.0268284	5.551993	0.9615339	Patient	Sain	25.917	Brain-specific homeobox	119.76184	56.50442	16.821931	8.5075688	3888.8817	85.190432	6077.0842	3199.583
P02652	11	11	995.34	0.0020124	0.0268284	2.4014067	0.9606031	Patient	Sain	11.168	Apolipoprotein A-II OS=Hc	34056410	26771849	14632439	15663194	42268641	26890744	89512485	43236635
P08571	12	12	937.27	0.0025308	0.0320335	1.6971067	0.9485492	Sain	Patient	40.051	Monocyte differentiation	797874.17	434080.26	499432.28	477652.53	733837.28	418725.29	352767.33	318079.37
P26927	31	31	1222.76	0.0025808	0.0320335	3.3123536	0.9474046	Patient	Sain	80.268	Hepatocyte growth factor-	190652.76	43028.557	155598.93	60550.701	75682.634	141408.77	141623.19	380083.05
P02655	6	6	819.16	0.0030504	0.0361789	4.7206762	0.9368237	Patient	Sain	11.277	Apolipoprotein C-II OS=Hc	1281811.8	751583.4	312951.49	1013499	603757.76	1691499.6	4859804.1	137623
P02671	8	8	369.69	0.0031158	0.0361789	191.8384	0.9353762	Patient	Sain	94.914	Fibrinogen alpha chain O	1792.4021	4843.2609	2616.5938	746.39308	685.49376	8280.2447	1254362.7	107899
P02763	17	17	1378.36	0.0036234	0.0407586	2.8533872	0.924364	Sain	Patient	23.497	Alpha-1-acid glycoprotein	67476432	40320887	40796528	53919933	106057189	35895348	25642029	2274357
Q14831	1	1	23.81	0.0038979	0.0425178	2.9788784	0.9185746	Sain	Patient	102.185	Metabotropic glutamate r	26519.187	83446.172	72384.114	34456.805	83227.769	36550.192	8908.5356	31759.88
P02647	32	32	2367.65	0.0040205	0.0425648	2.2688435	0.9160271	Patient	Sain	30.759	Apolipoprotein A-I OS=Ho	105663566	161836184	96430549	75237801	158864474	104736598	166481100	27802701
P01876	21	21	1557.26	0.0044316	0.0455765	3.5420267	0.9076475	Sain	Patient	37.631	Ig alpha-1 chain C region (	10958465	3390608.1	19371147	11129440	8232536	4863588.4	3403586.2	5212718.5
Q9UNW1	8	8	227	0.005259	0.0525841	1.7900311	0.8915219	Sain	Patient	55.016	Multiple inositol polypho	62640.486	55052.459	58667.71	30459.856	63293.849	79838.921	34467.455	30358.52
Q16853	2	2	81.52	0.0054293	0.0528194	6.1436713	0.8883207	Sain	Patient	84.568	Membrane primary amine	24451.336	3442.0472	3308.1744	4995.8869	4135.7996	9790.8939	236.61601	194.7495
P01833	8	8	219.66	0.006944	0.0656448	9.1555604	0.8614401	Sain	Patient	83.232	Polymeric immunoglobuli	14801.731	38882.249	8643.925	36538.175	31852.344	72666.846	1025.7634	2483.330
P27169	28	27	2064.1	0.0071604	0.0656448	1.8486226	0.8578184	Patient	Sain	39.706	Serum paraoxonase/aryle	3861076.2	5186138.5	2149700.4	5390878.2	2346566.5	3471134.2	5925120.8	6280114.4
P60174	1	1	28.75	0.0074	0.0656448	2.9166777	0.8538651	Patient	Sain	30.772	Triosephosphate isomera	4225.6237	1794.3666	1066.6861	1210.1584	3775.1162	2981.76	10398.334	4675.589
P00736	41	39	3067.37	0.0074771	0.0656448	1.4866701	0.8526056	Sain	Patient	80.067	Complement C1r subcomp	5434650.2	5169551.9	3906709.5	4169170.8	3458132.5	4641351.5	4547911.2	2638367.5
P09871	37	37	2588.95	0.0085558	0.0726349	1.93603	0.835606	Sain	Patient	76.635	Complement C1s subcomp	5789892.8	5490839.9	3882940.1	5662793.4	5033975.1	2280238.9	1956631.4	2302798.4
A0A0B4J1V0	3	3	104.52	0.0086769	0.0726349	2.7965659	0.833766	Sain	Patient	12.917	Immunoglobulin heavy va	92353.063	20397.701	46790.32	87239.628	43135.867	81618.911	38454.812	17117.343
P01619	4	4	262.21	0.0095577	0.0781901	2.30903	0.8207663	Sain	Patient	12.549	Immunoglobulin kappa va	409287.43	332471.99	228180.64	675002.86	220569.52	187876.28	139130.56	148976.36
Q08380	21	21	1485.84	0.0115602	0.0900161	2.528023	0.7934679	Sain	Patient	65.289	Galectin-3-binding protei	938408.99	614190.09	1091308.6	2622296.8	1676638.8	2590031.7	640186.94	906404.59
P06732	9	9	314.24	0.0117011	0.0900161	5.8938011	0.7916528	Patient	Sain	43.074	Creatine kinase M-type OS	7288.867	9691.0748	1004.0051	6516.3248	4494.7392	6004.4549	19154.379	89665.94

Quantitative proteomics without labeling : export

Peptide level!

Row	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2								Best identification									
3	Peptide ide	Ions used i	Ions	Deconvolu	Deconvolu	Retention t	Neutral ma	Score	Sequence	Modificatio	Accession	Description	Max fold ch	Highest me	Lowest me	Anova	Maximum C
4	72.73_130	2	2	#115;#1147	2;3	72.73305	1301.7079	75.54	SLDLDSIIAEVK		P04264	Keratin, typ	1.0924126	P2Y2-IFN	NT-P2Y2	0.7811603	106.87214
5	60.30_130	0	1	#6540	2	60.298133	1301.7084	20.29	SLDLDSIIAEVK		P04264	Keratin, typ	1.1406329	P2Y2-IFN	NT-P2Y2	0.8070283	50.874645
6	60.51_130	0	1	#26898	2	60.514783	1301.7088	16.04	SLDLDSIIAEVK		P04264	Keratin, typ	1.2779871	P2Y2-IFN	NT-P2Y2	0.3895883	123.29264
7	39.00_117	1	1	#150	2	39.001683	1178.5939	71	YEELQITAGR		P04264	Keratin, typ	1.2978820	P2Y2-IFN	NT-P2Y2	0.3403630	32.299921
8	56.96_138	0	1	#153	2	56.96445	1382.6833	87.18	SLNNQFASFIDK		P04264	Keratin, typ	2.0771715	P2Y2-IFN	NT-P2Y2	0.0601139	39.471601
9	55.69_147	0	1	#17867	2	55.687866	1474.7409	66.14	WELLQQVDTSTR		P04264	Keratin, typ	1.9501543	NT-P2Y2	P2Y2-IFN	0.0932066	52.594856
10	56.63_147	0	1	#200	2	56.633916	1474.7414	86.01	WELLQQVDTSTR		P04264	Keratin, typ	1.5717229	P2Y2-IFN	NT-P2Y2	0.0755519	27.940603
11	63.32_127	0	1	#215	2	63.3216	1276.7040	64.5	LALDLEIATYR		P04264	Keratin, typ	1.5188866	P2Y2-IFN	NT-P2Y2	0.9260598	107.53169
12	69.32_127	0	1	#34344	2	69.315416	1276.7034	27	LALDLEIATYR		P04264	Keratin, typ	Infinity	NT-P2Y2	P2Y2-IFN	0.3739009	173.20508
13	26.18_103	0	1	#292	2	26.184133	1032.5095	34.34	TLLEGEESR		P04264	Keratin, typ	1.9482978	P2Y2-IFN	NT-P2Y2	0.0374736	36.507447
14	32.38_873	0	1	#300	2	32.382833	873.49238	27.05	SLVNLGGSK		P04264	Keratin, typ	1.3133291	NT-P2Y2	P2Y2-IFN	0.4893575	94.329898
15	49.15_135	0	1	#323	2	49.146133	1356.6889	57.58	LNDLEDALQQAK		P04264	Keratin, typ	51.472069	P2Y2-IFN	NT-P2Y2	0.0309571	80.583779
16	70.37_199	2	2	#6877;#39	2;3	70.368508	1992.9697	70.42	THNLEPYFESFINNLR		P04264	Keratin, typ	1.2772453	P2Y2-IFN	NT-P2Y2	0.7562446	67.867555
17	40.10_126	0	1	#414	2	40.097966	1264.6298	48.42	TNAENEFVTIK		P04264	Keratin, typ	3.2632073	P2Y2-IFN	NT-P2Y2	0.0087539	33.807577
18	58.13_163	0	2	#3340;#60	2;3	58.132416	1637.8529	35.56	SLNNQFASFIDKVR		P04264	Keratin, typ	2.5974945	NT-P2Y2	P2Y2-IFN	0.1535334	169.86068
19	57.42_163	0	1	#23689	3	57.4163	1637.8526	7.92	SLNNQFASFIDKVR		P04264	Keratin, typ	4.4531173	P2Y2-IFN	NT-P2Y2	0.4425376	126.06406
20	23.27_112	0	1	#661	2	23.274633	1124.5353	29.8	AEAESLYQSK		P04264	Keratin, typ	1.3469793	P2Y2-IFN	NT-P2Y2	0.3103307	58.585204
21	29.02_139	0	2	#2369;#73	2;3	29.020275	1392.7249	33.58	TNAENEFVTIKK		P04264	Keratin, typ	2.1442453	NT-P2Y2	P2Y2-IFN	0.2880216	138.35432
22	39.21_114	0	1	#745	2	39.209283	1140.5125	37.04	DYQELMNTK		P04264	Keratin, typ	11.7282151	P2Y2-IFN	NT-P2Y2	0.0468781	80.410623
23	42.91_171	0	2	#806;#181	2;3	42.91125	1715.8444	97.44	QISNLQQSISDAEQR		P04264	Keratin, typ	9.0915362	P2Y2-IFN	NT-P2Y2	0.0145680	54.749022
24	61.26_171	0	1	#24914	2	61.2566	1715.8456	53.58	QISNLQQSISDAEQR		P04264	Keratin, typ	18.192896	P2Y2-IFN	NT-P2Y2	0.0951607	87.852537
25	26.92_101	0	1	#913	2	26.923366	1014.4330	37.9	DVDGAYM	[7] Oxidatic	P04264	Keratin, typ	1.1520109	P2Y2-IFN	NT-P2Y2	0.7045560	42.333170
26	44.18_129	0	1	#963	2	44.177	1299.5227	37.02	NMQDMVEDYR		P04264	Keratin, typ	5.4096899	P2Y2-IFN	NT-P2Y2	0.10518187	72.756971
27	17.30_133	0	2	#2112;#123	2;3	17.30085	1339.6613	31.13	SKAEAESLYQSK		P04264	Keratin, typ	1.4966523	NT-P2Y2	P2Y2-IFN	0.3931951	111.941052
28	44.19_131	0	1	#1434	2	44.189416	1315.5173	32.37	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	2.4577293	P2Y2-IFN	NT-P2Y2	0.1043870	61.658569
29	30.45_131	0	1	#18448	2	30.451316	1315.5176	12.88	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	4.2507130	P2Y2-IFN	NT-P2Y2	0.8643282	145.916134
30	32.97_131	0	1	#10520	2	32.96975	1315.5178	31.19	NMQDMVE	[5] Oxidatic	P04264	Keratin, typ	4.9859009	P2Y2-IFN	NT-P2Y2	0.10136127	79.512617
31	25.24_115	0	1	#6851	2	25.2391	1156.5075	21.94	DYQELMN	[6] Oxidatic	P04264	Keratin, typ	3.8834292	P2Y2-IFN	NT-P2Y2	0.3200131	106.08651
32	39.21_115	0	1	#1463	2	39.209283	1156.5076	36.97	DYQELMN	[6] Oxidatic	P04264	Keratin, typ	7.2419825	P2Y2-IFN	NT-P2Y2	0.0055966	40.549868
33	49.59_159	0	2	#3020;#21	2;3	49.586816	1598.8268	54.06	NKLNDLEDALQQAK		P04264	Keratin, typ	3.4440257	P2Y2-IFN	NT-P2Y2	0.0214419	56.125894
34	43.32_165	0	1	#2737	2	43.32145	1656.7860	88.97	SGGGFSSGSAGIINY		P04264	Keratin, typ	1.3066371	NT-P2Y2	P2Y2-IFN	0.4876225	76.268465
35	30.44_133	0	1	#11109	2	30.4389	1331.5123	40	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	2.8591550	P2Y2-IFN	NT-P2Y2	0.9872820	86.994767
36	32.98_133	0	1	#7069	2	32.97595	1331.5125	51.97	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	3.9961599	P2Y2-IFN	NT-P2Y2	0.0429856	63.482265
37	33.95_133	0	1	#5953	2	33.9499	1331.5123	21.46	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	4.3863708	P2Y2-IFN	NT-P2Y2	0.0479284	68.269541
38	44.19_133	0	1	#3079	2	44.189416	1331.5121	34.21	NMQDMVE	[2] Oxidatic	P04264	Keratin, typ	1.8459656	P2Y2-IFN	NT-P2Y2	0.2198095	58.508176
39	63.58_218	0	2	#3104;#21	3;4	63.580533	2183.1181	13.89	NKLNDLEDALQQAKE		P04264	Keratin, typ	1.0450035	P2Y2-IFN	NT-P2Y2	0.4491257	143.132614
40	25.51_331	0	1	#4319	3	25.505866	3311.3035	52.2	GSYGSGGSSYGSGG		P04264	Keratin, typ	1.7179400	NT-P2Y2	P2Y2-IFN	0.5663209	74.371373
41	87.38_234	0	1	#4991	3	87.382566	2348.2112	24.57	SLDLDSIIAEVKAQYED		P04264	Keratin, typ	9.9913535	NT-P2Y2	P2Y2-IFN	0.1911382	169.17974
42	22.90_238	0	2	#5595;#85	2;3	22.896225	2382.9452	113.92	GGGGGGYGSGGSS		P04264	Keratin, typ	1.7557570	NT-P2Y2	P2Y2-IFN	0.5203451	68.259991
43	43.10_153	0	1	#5759	3	43.099533	1538.7766	10.13	LLRDYQEL	[9] Oxidatic	P04264	Keratin, typ	1.0987312	P2Y2-IFN	NT-P2Y2	0.3322667	164.73966
44	48.76_228	0	1	#6371	3	48.7625	2285.1172	17.59	AEAESLYQSKYEELQI		P04264	Keratin, typ	4.7741061	NT-P2Y2	P2Y2-IFN	0.1646858	166.319914
45	31.92_121	0	1	#6957	2	31.920233	1216.6095	25.03	FVSTTYSGVTR		P04264	Keratin, typ	2.4180408	NT-P2Y2	P2Y2-IFN	0.1408007	110.80825
46	43.18_250	0	1	#7003	3	43.184383	2500.2444	12.04	SKAEAESLYQSKYEEI		P04264	Keratin, typ	9.2525599	NT-P2Y2	P2Y2-IFN	0.1134288	171.71833
47	65.09_214	0	2	#7383;#78	3;4	65.087133	2149.0705	7.93	THNLEPYFESFINNLRF		P04264	Keratin, typ	3.2877905	NT-P2Y2	P2Y2-IFN	0.1383409	147.172211
48	56.06_264	0	1	#7722	4	56.057916	2642.3253	25.36	QISNLQQSISDAEQRG		P04264	Keratin, typ	1.1843984	NT-P2Y2	P2Y2-IFN	0.2982029	160.94860
49	21.77_127	0	2	#20957;#9	2;3	21.772708	1276.6259	13.4	SDQSRLDSELK		P04264	Keratin, typ	3.5922894	P2Y2-IFN	NT-P2Y2	0.6285423	172.02585
50	21.98_127	0	2	#20244;#8	2;3	21.981825	1276.6259	12.16	SDQSRLDSELK		P04264	Keratin, typ	25.212269	NT-P2Y2	P2Y2-IFN	0.0581183	134.96240
51	33.81_184	0	1	#9322	3	33.810416	1843.9392	56.83	KQISNLQQSISDAEQR		P04264	Keratin, typ	16.3171846	P2Y2-IFN	NT-P2Y2	0.0953386	81.703149
52	29.46_130	0	1	#10434	3	29.459266	1301.6937	18.36	NSKIEISELNR		P04264	Keratin, typ	1.7721441	NT-P2Y2	P2Y2-IFN	0.2229925	171.626917
53	65.15_194	0	1	#12901	3	65.148333	1940.9804	31.54	LNDLEDALQQAKEDL		P04264	Keratin, typ	1.9951862	P2Y2-IFN	NT-P2Y2	0.5939525	101.208177
54	48.13_168	0	1	#15207	3	48.12795	1686.9627	14.78	SLVNLGGSKSISISVA		P04264	Keratin, typ	3.3680944	NT-P2Y2	P2Y2-IFN	0.1686845	95.185337
55	64.00_293	0	1	#16301	3	64.002766	2931.5090	24.16	FLEQQNQVLQTKWEL		P04264	Keratin, typ	1.8349425	NT-P2Y2	P2Y2-IFN	0.2092689	95.718791
56	67.05_136	2	2	#293;#154	2;3	67.045066	1360.7349	55.27	AQFEGIVTDLIR		P38646	Stress-70	1.2032339	NT-P2Y2	P2Y2-IFN	0.4012992	27.956632
57	53.15_124	0	1	#375	2	53.1508	1241.6731	82.2	DAGQISGLNVLR		P38646	Stress-70	1.3454871	NT-P2Y2	P2Y2-IFN	0.0617153	18.165128
58	45.32_144	0	1	#514	2	45.322833	1449.7108	73.47	TTPSVVAFTADGER		P38646	Stress-70	2.3213500	NT-P2Y2	P2Y2-IFN	0.0772035	63.180148

Row	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	Normalized abundance						Raw abundance						Spectral counts	
2	NT-P2Y2			P2Y2-IFN			NT-P2Y2			P2Y2-IFN			NT-P2Y2	
3	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	1914005-G	19140
4	5429611.5	2396907.2	6611636.4	1925290.3	11744671.9	2102461.8	3616651.2	2396907.2	3625930.6	4227100.8	2766728.9	4354021.4	1	
5	32369.598	24424.673	44472.259	26914.657	61118.7124	27474.576	21561.3117	24424.673	24389.321	59092.889	14397.925	56897.534	0	
6	0	239.60268	915.21617	685.04119	253.57420	537.22823	0	239.60268	501.91967	1504.0526	59.735265	1112.55443	0	
7	2168311.73	1268884.3	1670198.5	1752814.0	3032025.1	1843956.7	1444307.2	1268884.3	915964.49	3848417.7	714263.59	3818679.1	1	
8	1558842.5	1173669.8	667027.27	1971950.7	3133234.7	1956241.5	1038341.2	1173669.8	365808.77	4329546.8	738105.86	4051211.48	1	
9	14494.515	11166.4226	18652.090	4235.6873	11984.023	6503.1213	9654.7621	11166.4226	10229.1146	9299.7283	2823.1136	13467.416	0	
10	1417558.4	1012105.52	1125228.1	1469296.3	2452542.6	1665466.5	944232.28	1012105.52	617093.72	3225936.1	577753.10	3449040.9	1	
11	864082.21	724826.88	1167152.27	521744.34	3127963.1	536437.46	575563.08	724826.88	640085.58	1145523.7	736864.00	1110916.80	1	
12	14.238816	0	0	0	0	0	9.4844413	0	0	0	0	0	0	
13	565077.21	259522.08	435785.62	746444.76	855097.40	854063.07	376396.57	259522.08	238992.03	1638868.2	201437.95	1768692.6	0	
14	973040.00	428278.90	1248380.4	310007.98	1405023.6	302512.44	648139.61	428278.90	684632.43	680642.77	330985.78	626477.80	0	
15	16964.421	13808.149	11363.392	1084577.9	50893.909	1033353.3	11299.960	13808.149	6231.8724	2381261.8	11989.236	2139987.8	0	
16	1658628.7	1358222.1	2568785.9	1666749.1	4225490.5	1241989.3	1104808.6	1358222.1	1408764.6	3659456.5	995411.94	2572055.2	2	
17	200263.00	152635.86	195233.08	737212.53	365349.37	686106.31	133394.70	152635.86	107069.05	1618598.2	86066.487	1420868.3	1	
18	1727366.1	596070.09	2632978.7	10001.364	1883586.5	14564.299	1150594.4	596070.09	1443969.0	21958.650	443722.33	30161.436	1	
19	755.86267	51.810844	1046.3936	169.20377	1391.5291	6695.6456	503.47831	51.810844	573.85954	371.49794	327.80685	13866.1182	0	
20	307014.56	88304.354	170421.45	265433.31	233707.32	262899.95	204501.66	88304.354	93461.941	582776.19	55055.161	544443.67	0	
21	790199.03	284797.07	913521.64	129647.10	797307.22	419.82833	526349.67	284797.07	500990.35	284648.69	187824.14	869.42911	0	
22	22287.809	8061.3943	4609.8938	170095.05	25034.008	214878.75	14845.856	8061.3943	2528.1419	373454.80	5897.3393	444995.80	0	
23	80525.023	48208.510	28468.339	632154.40	175286.52	621765.60	53637.524	48208.510	15612.507	1387936.2	41292.791	1287624.2	0	
24	906.37638	572.13906	1161.37427	17940.903	1078.6449	29007.692	603.73513	572.13906	636.91683	39390.425	254.09975	60072.486	0	
25	214773.11	119855.00	211314.76	151315.33	311778.96	165837.86	143059.85	119855.00	115888.50	332222.70	73446.740	343436.25	0	
26	41620.224	39418.997	4597.4129	170284.34	46089.088	246894.21	27723.131	39418.997	2521.2971	373870.40	10857.349	511297.118	0	
27	513258.24	156505.95	458226.19	83212.117	575888.62	94574.909	341880.08	156505.95	251298.81	182697.64	135663.87	195856.67	1	
28	106908.68	106807.21	22697.983	181521.763	167031.23	232488.35	71211.6171	106807.21	12447.949	398542.89	39348.067	481463.80	1	
29	684.42363	482.26351	0	4420.9633	538.28889	0	455.89293	482.26351	0	9706.5139	126.80638	0	0	
30	4309.4306	3925.8455	240.77486	17307.818	7254.3231	17698.609	2870.5014	3925.8455	132.04491	38000.445	1708.9234	36652.330	0	
31	669.65506	3879.5554	1660.5098	1124.8844	5422.7530	17567.372	446.05563	3879.5554	910.65103	2469.7572	1277.4548	36380.548	0	
32	26729.187	19080.851	11637.3197	185290.16	76222.017	154520.58	17804.247	19080.851	6382.0983	406816.66	17955.857	319999.111	0	
33	77152.864	46552.268	22558.207	139778.40	178701.03	185255.27	51391.337	46552.268	12371.293	306892.62	42097.157	383648.06	0	
34	110914.265	126677.55	259589.59	60637.445	237850.75	82016.387	73879.725	126677.55	142363.22	133133.47	56031.241	169849.02	1	
35	6054.2507	4155.5090	239.46233	15986.224	13859.159	30.561395	4032.7219	4155.5090	131.325100	35098.801	3264.8454	63.290075	0	
36	14067.350	12956.219	2670.9282	45502.242	32843.838	40317.882	9370.2284	12956.219	1464.7811	99903.145	7737.1252	83494.939	0	
37	13716.945	11656.795	2032.8119	45995.698	31814.660	42404.945	9136.8243	11656.795	1114.82765	100986.56	7494.6786	87817.070	0	
38	99702.889	109409.43	25688.004	104197.910	208129.81	121105.621	66411.855	109409.43	14087.726	228773.32	49029.787	250799.54	0	
39	434480.29	255246.65	403625.75	49111.0875	1009995.3	83451.088	289406.27	255246.65	221355.03	107826.60	237927.74	172820.17	0	
40	506701.59	86213.468	506660.50	136436.97	395822.96	107794.72	337512.71	86213.468	277860.98	299556.29	93245.248	223233.79	0	
41	26239.056	24459.391	807996.47	1269.9058	84607.913	65.983585	17477.772	24459.391	443118.612	2788.1612	19931.349	136.64644	0	
42	225262.93	46806.045	263819.46	64228.046	181875.68	59114.1551	150047.10	46806.045	144682.95	141016.872	42845.019	122420.43	1	
43	65204.051	31593.975	53828.671	2322.2268	160104.87	3071.1616	43432.261	31593.975	29520.532	5098.6008	37716.402	6360.11711	0	
44	53745.273	10451.838	176527.83	801.44348	49086.413	535.18436	35799.597	10451.838	96810.782	1759.6215	11563.439	1108.32173	1	
45	1199.7422	2417.4885	2818.4457	309.48953	329.77740	2022.2581	799.14539	2417.4885	1545.6821	679.50448	77.686688	4187.9263	0	
46	129520.70	23289.565	518711.675	104.18766	72161.412	311.281277	86273.431	23289.565	284470.05	228.75082	16999.288	644.63730	0	
47	274798.91	210452.68	820986.60	19683.845	357489.89	20125.952	183042.89	210452.68	450242.61	43217.168	84215.007	41679.1515	2	
48	226679.15	87180.006	143727.56	7027.4777	368107.14	11210.6283	150990.44	87180.006	78822.570	15429.286	86716.145	23216.267	1	
49	3083.6152	29622.480	225.69068	536.69593	117762.95	0.8608249	2053.9886	29622.480	123.77249	1178.3509	27741.784	1.7826960	0	
50	104112.908	6238.1430	99568.646	954.90433	7083.3435	287.84502	69349.357	6238.1430	54605.090	2096.5548	1668.6452	596.10279	1	
51	402.23804	979.16092	245.52940	11581.8346	651.00156	14314.054	267.92979	979.16092	134.65237	25428.674	153.35846	29643.201	0	
52	37420.998	9628.3348	39195.949	296.07610	48371.1183	0	24926.037	9628.3348	21495.706	650.05443	11394.9348	0	0	
53	44198.142	25523.945	19854.455	15722.317	128330.41	34669.157	29440.276	25523.945	10888.5112	34519.376	30231.195	71796.904	0	
54	4251.0332	5400.7163	22426.338	1777.7076	5917.7732	1828.6249	2831.6030	5400.7163	12298.974	3903.0733	1394.0682	3786.9282	0	
55	62397.069	40013.378	28968.025	12026.297	50199.812	9372.0311	41562.538	40013.378	15886.543	26404.522	11825.725	19408.687	1	
56	1353499.7	1694073.7	1079049.5	1092631.5	1485075.2	851903.05	901562.89	1694073.7	591768.61	2398944.0	349843.80	1764219.46	1	
57	1165586.7	813513.84	951143.31	765983.34	741966.65	669881.20	776394.50	813513.84	521622.70	1681766.4	174787.39	1387267.5	1	
58	1335812.3	833758.52	1250862.6	223949.00	831506.69	418011.697	889781.41	833758.52	685993.73	491694.66	195880.61	865667.01	1	

1914005_peptideMeasurement_NTP2

Quantitative proteomics without labeling : export

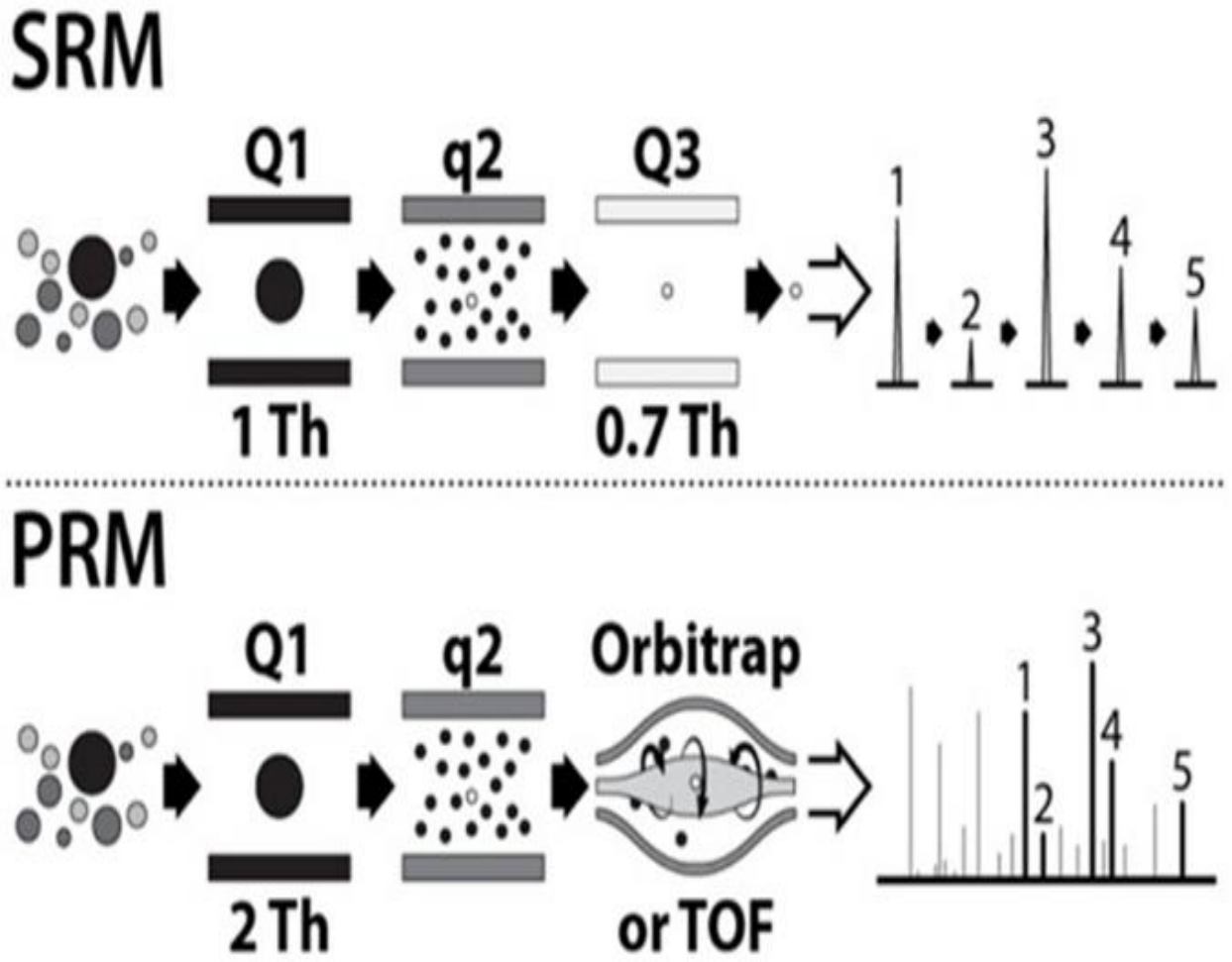
aifMsms.txt	18-Oct-18 6:45 PM	Fichier TXT	0 Ko
allPeptides.txt	18-Oct-18 6:47 PM	Fichier TXT	63,279 Ko
evidence.txt	18-Oct-18 6:44 PM	Fichier TXT	5,765 Ko
matchedFeatures.txt	18-Oct-18 6:45 PM	Fichier TXT	0 Ko
modificationSpecificPeptides.txt	18-Oct-18 6:44 PM	Fichier TXT	2,281 Ko
msms.txt	18-Oct-18 6:44 PM	Fichier TXT	26,108 Ko
msmsScans.txt	18-Oct-18 6:45 PM	Fichier TXT	15,918 Ko
msScans.txt	18-Oct-18 6:47 PM	Fichier TXT	2,211 Ko
mzRange.txt	18-Oct-18 6:47 PM	Fichier TXT	95 Ko
Oxidation (M)Sites.txt	18-Oct-18 6:44 PM	Fichier TXT	125 Ko
parameters.txt	18-Oct-18 6:44 PM	Fichier TXT	2 Ko
peptides.txt	18-Oct-18 6:44 PM	Fichier TXT	2,741 Ko
proteinGroups.txt	18-Oct-18 6:43 PM	Fichier TXT	1,108 Ko
summary.txt	18-Oct-18 6:44 PM	Fichier TXT	2 Ko
tables.pdf	18-Oct-18 6:45 PM	Adobe Acrobat D...	40 Ko

Protein IDs	Majority protein IDs	Peptide counts (all)	Peptide counts (razor+unique)	Peptide counts (unique)	Fasta headers	Number o
C1_00060W_A	C1_00060W_A	10	10	10	>C1_00060W_A translated using codon table 12 (512 amino acids) Verified ORF; (orf19.6109) Transcriptiona	
C1_00070W_A	C1_00070W_A	12	12	12	>C1_00070W_A translated using codon table 12 (362 amino acids) Verified ORF; (orf19.6105) Mevalonate dip	
C1_00110W_A	C1_00110W_A	14	14	14	>C1_00110W_A translated using codon table 12 (540 amino acids) Verified ORF; (orf19.6099) Chaperonin-con	
C1_00140W_A	C1_00140W_A	5	5	5	>C1_00140W_A translated using codon table 12 (1018 amino acids) Verified ORF; (orf19.6092) Kelch repeat	
C1_00150C_A	C1_00150C_A	1	1	1	>C1_00150C_A translated using codon table 12 (622 amino acids) Verified ORF; (orf19.6091) Beta-arrestin-	
C1_00160C_A	C1_00160C_A	4	4	4	>C1_00160C_A translated using codon table 12 (400 amino acids) Verified ORF; (orf19.6090) Putative nucle	
C1_00170W_A	C1_00170W_A	9	9	9	>C1_00170W_A translated using codon table 12 (579 amino acids) Verified ORF; (orf19.6086) Putative 2-iso	
C1_00180W_A	C1_00180W_A	3	3	3	>C1_00180W_A translated using codon table 12 (200 amino acids) Verified ORF; (orf19.6085) Ribosomal prot	
C1_00210C_A	C1_00210C_A	2	2	2	>C1_00210C_A translated using codon table 12 (384 amino acids) Verified ORF; (orf19.6082) Ortholog(s) ha	
C1_00220W_A;C4_04530C_A	C1_00220W_A	5;1	5;1	5;1	>C1_00220W_A translated using codon table 12 (544 amino acids) Verified ORF; (orf19.6081) Gl	
C1_00320W_A	C1_00320W_A	2	2	2	>C1_00320W_A translated using codon table 12 (261 amino acids) Uncharacterized ORF; (orf19.6076) Ortholo	
C1_00330C_A	C1_00330C_A	2	2	2	>C1_00330C_A translated using codon table 12 (182 amino acids) Uncharacterized ORF; (orf19.6075) Putativ	
C1_00340W_A	C1_00340W_A	4	4	4	>C1_00340W_A translated using codon table 12 (248 amino acids) Verified ORF; (orf19.6074) Essential prot	
C1_00380C_A	C1_00380C_A	7	7	7	>C1_00380C_A translated using codon table 12 (745 amino acids) Uncharacterized ORF; (orf19.6071) Ortholo	
C1_00400W_A	C1_00400W_A	7	7	7	>C1_00400W_A translated using codon table 12 (382 amino acids) Uncharacterized ORF; (orf19.6068) Putativ	
C1_00410C_A	C1_00410C_A	13	13	12	>C1_00410C_A translated using codon table 12 (542 amino acids) Uncharacterized ORF; (orf19.6066) Hexadec	
C1_00420W_A	C1_00420W_A	7	7	7	>C1_00420W_A translated using codon table 12 (323 amino acids) Uncharacterized ORF; (orf19.6065) RNA pol	
C1_00440W_A	C1_00440W_A	11	11	11	>C1_00440W_A translated using codon table 12 (478 amino acids) Uncharacterized ORF; (orf19.6063) Putativ	
C1_00450C_A	C1_00450C_A	2	2	2	>C1_00450C_A translated using codon table 12 (150 amino acids) Uncharacterized ORF; (orf19.6062.3) Mitoc	
C1_00460W_A	C1_00460W_A	1	1	1	>C1_00460W_A translated using codon table 12 (106 amino acids) Verified ORF; (orf19.6062) Putative TIM23	
C1_00480C_A	C1_00480C_A	4	4	4	>C1_00480C_A translated using codon table 12 (751 amino acids) Uncharacterized ORF; (orf19.6060) YEF3-su	
C1_00490C_A	C1_00490C_A	2	2	2	>C1_00490C_A translated using codon table 12 (119 amino acids) Verified ORF; (orf19.6059) Putative gluta	
C1_00500C_A	C1_00500C_A	3	3	3	>C1_00500C_A translated using codon table 12 (342 amino acids) Uncharacterized ORF; (orf19.6058) Putativ	
C1_00560W_A	C1_00560W_A	2	2	2	>C1_00560W_A translated using codon table 12 (390 amino acids) Verified ORF; (orf19.6052) Putative co-ch	
C1_00590W_A	C1_00590W_A	14	14	14	>C1_00590W_A translated using codon table 12 (426 amino acids) Uncharacterized ORF; (orf19.6047) Transla	
C1_00610W_A	C1_00610W_A	3	3	3	>C1_00610W_A translated using codon table 12 (590 amino acids) Verified ORF; (orf19.6045) Phosphatidylse	

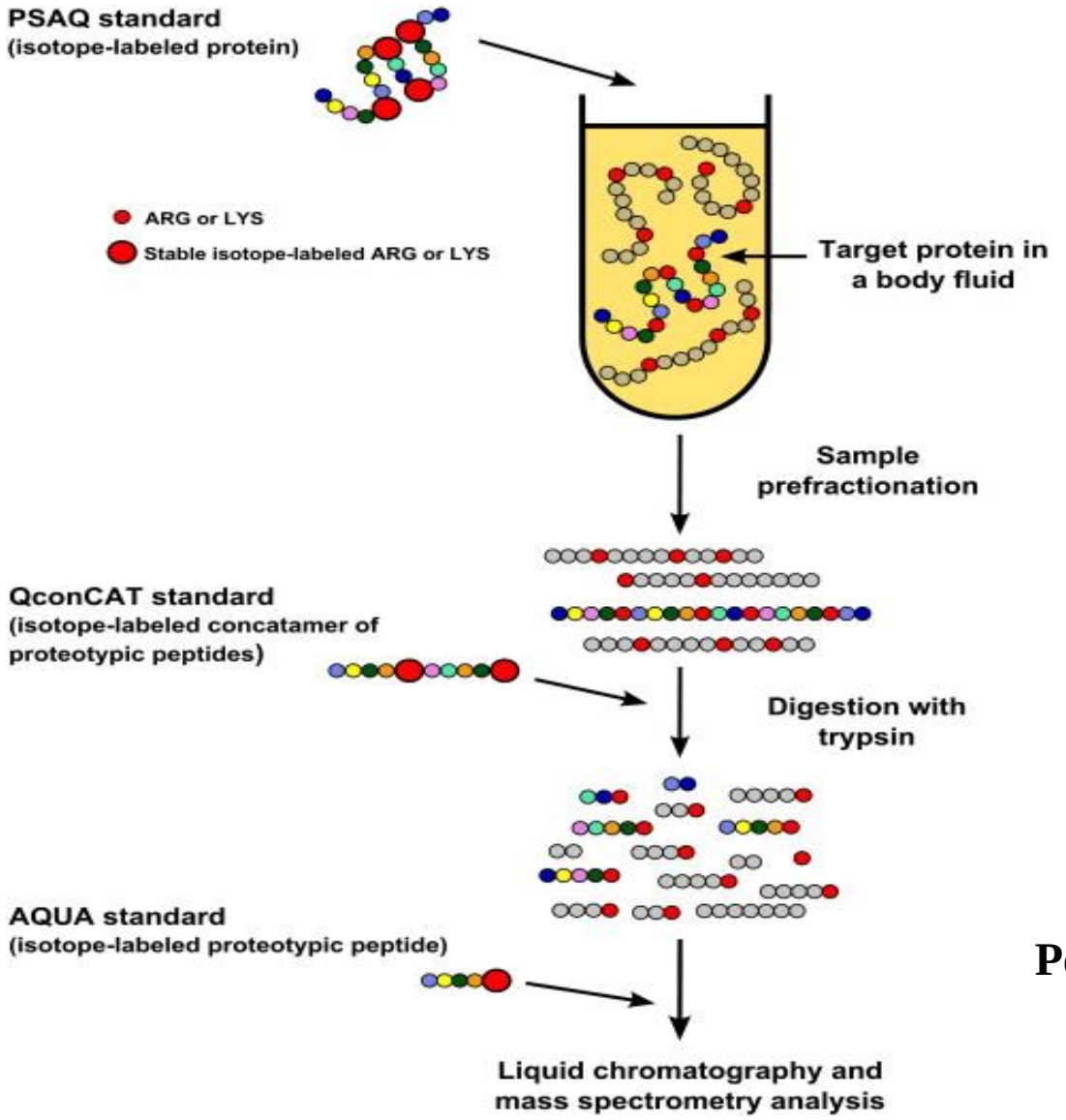
Quantification output formats

File name	File content
Protein/peptide quantification	Protein/peptide expression values can also be obtained from an MS--based proteomics experiment and then this data and metadata is used for performing the quantification analysis of peptides and proteins.
Metadata	A term used to describe data that provides additional information about a particular data set. This information can include how, when and where the data set was generated and what standards were used. In the proteomics context the addition of metadata such as peptide and protein identifications and quantification of their expression values gives meaning to a simple collection of mass spectra output files.

Targeted proteomics : PRM mode



Absolute quantification



PSAQ

QconCAT

Peptides AQUA

How can I use these data?

													Normalized abundance							
													Sain						Patient	
	Accession	Peptide cou	Unique pep	Confidence	Anova (p)	q Value	Max fold ch	Power	Highest me	Lowest mea	Mass	Description	1845007-F1	1845007-F3	1845007-F5	1845007-F7	1845007-F9	1845007-F1	1845007-F2	1845007-F4
4	P40197	12	12	545.23	2.48E-07	4.74E-05	32.042418	1	Patient	Sain	60.921	Platelet glycoprotein V OS	1447.8857	898.47877	2114.249	3517.982	2506.9091	1186.7152	33178.605	73995.785
5	P02776	2	2	134.77	2.63E-07	4.74E-05	78.817355	1	Patient	Sain	10.838	Platelet factor 4 OS=Homo	3812.4369	3755.8358	1044.3911	2939.4867	3862.0883	1110.5549	333829.66	68071.511
6	Q13201	6	6	221.4	5.33E-07	6.40E-05	72.663027	1	Patient	Sain	138.023	Multimerin-1 OS=Homo sa	160.48528	939.93933	155.7663	732.57482	572.39757	196.60688	75519.303	23299.125
7	P04114	336	334	28302.18	7.99E-06	0.0006593	2.1384481	1	Patient	Sain	515.283	Apolipoprotein B-100 OS=	14117253	10601958	8472023.6	9982572.2	10056625	11898407	24415094	1744463
8	P07996	42	42	2678.71	9.16E-06	0.0006593	143.08362	1	Patient	Sain	129.3	Thrombospondin-1 OS=Ho	11604.367	10346.75	4059.3393	5652.5533	2121.1343	2367.1543	2303007.2	83175.083
9	Q15485	7	7	289.57	1.42E-05	0.0007087	24.427918	0.9999997	Patient	Sain	33.98	Ficolin-2 OS=Homo sapier	8778.176	6132.5206	10104.107	8194.6312	8690.2929	1737.9715	185819.2	105024.97
10	P10720	3	3	208.63	1.49E-05	0.0007087	36.552808	0.9999997	Patient	Sain	11.545	Platelet factor 4 variant O	32881.787	11758.269	1451.6395	5448.2604	4187.5362	4196.5766	546497.96	239272.09
11	Q12884	2	2	70.47	1.58E-05	0.0007087	Infinity	0.9999996	Patient	Sain	87.657	Prolyl endopeptidase FAP	0	0	0	0	0	0	961.55592	106.477
12	Q15061	1	1	3.67	3.71E-05	0.0014832	23.811179	0.9999905	Sain	Patient	74.843	WD repeat-containing prc	567513.12	751475.75	852125.74	1189728.8	625225.18	78135.037	46681.126	36585.068
13	P04075	8	8	371.85	4.58E-05	0.0016488	4.2937733	0.9999814	Patient	Sain	39.395	Fructose-bisphosphate al	33219.801	26234.171	28730.699	26696.908	33816.665	32024.353	128602.25	168571.3
14	P09486	10	9	463.23	6.25E-05	0.0020451	6.1746973	0.999953	Patient	Sain	34.61	SPARC OS=Homo sapiens (	9522.5101	37758.131	36721.019	20388.641	15117.423	28228.93	299004.24	97804.241
15	P10124	1	1	70.16	7.11E-05	0.002134	272.44302	0.9999321	Patient	Sain	17.641	Serglycin OS=Homo sapier	67.740323	0	958.26875	53.15204	160.31231	238.84011	105180.09	42246.241
16	P02775	6	6	388.98	8.21E-05	0.0022557	28.012451	0.9998996	Patient	Sain	13.885	Platelet basic protein OS=	61607.9	109352.83	105910.89	67216.894	60865.635	22752.188	3524260.5	1692399.5
17	Q9H1K0	2	2	12.94	8.77E-05	0.0022557	1.8919044	0.9998799	Patient	Sain	88.815	Rabenosyn-5 OS=Homo sa	871721.79	1194024.5	1103071.3	1121288.3	1020557.4	1361492.9	1506333.1	1757466.9
18	P35542	4	4	338.41	0.0001236	0.0029659	16.656374	0.9997122	Patient	Sain	14.737	Serum amyloid A-4 protei	192222.76	102064.74	9416.0169	8023.3014	46690.651	107524.48	1543185.2	1741696.5
19	P80188	6	6	223.58	0.0001847	0.0041561	9.1194461	0.9992657	Patient	Sain	22.574	Neutrophil gelatinase-as	2265.4484	1473.9176	6203.72	3676.0015	7827.3304	4603.3796	66633.817	26485.807
20	P02144	4	4	209.23	0.0002392	0.0050651	6.5032553	0.9987201	Patient	Sain	17.173	Myoglobin OS=Homo sapi	2754.3333	2223.1201	1035.6898	1813.5104	1462.164	5472.3528	17977.389	20813.25
21	P05067	5	5	179.76	0.000297	0.0059397	18.947351	0.9980148	Patient	Sain	86.888	Amyloid beta A4 protein C	408.20926	53.510036	338.118	494.28658	258.47482	516.31147	8561.0101	2161.7757
22	Q9NPH3	7	7	223.36	0.0003417	0.0064734	2.9534965	0.9973949	Patient	Sain	65.377	Interleukin-1 receptor acc	19269.326	23550.392	17918.705	19071.463	29165.24	20448.453	74326.28	82743.303
23	P22352	7	7	322.73	0.0005632	0.0101356	2.3655577	0.9936247	Patient	Sain	25.537	Glutathione peroxidase 3	356078.09	500176.63	411199.92	354267.54	608490.2	620677.21	1139291.1	1214566.8
24	A0A075B6I0	1	1	30.88	0.0005981	0.0102519	15.884721	0.9929515	Sain	Patient	12.806	Immunoglobulin lambda	14017.761	25223.891	6493.2483	70640.84	15600.774	36585.067	4641.3232	1487.298
25	P05155	33	33	2513.9	0.0007389	0.0120898	5.4023431	0.9900925	Patient	Sain	55.119	Plasma protease C1 inhib	741421.19	592478.75	377555.54	465885.04	2331779.9	7791080.5	11109177	10511766
26	A0A0C4DH2	2	2	100.38	0.0013344	0.0208833	3.8865369	0.9766266	Sain	Patient	12.999	Immunoglobulin heavy va	50500.146	64330.439	38938.111	121804.39	27881.723	32180.678	24057.435	17302.865
27	P14780	6	6	207.75	0.0014988	0.0221418	62.066206	0.972746	Patient	Sain	78.408	Matrix metalloproteinase	794.97907	136.39171	262.24899	313.84328	0	555.64619	58928.664	9636.8582
28	P02649	27	27	1941.9	0.0015378	0.0221418	2.4710215	0.971824	Patient	Sain	36.132	Apolipoprotein E OS=Hom	2552396.1	1536331.8	853676.88	1079240.3	1188651.7	1327575.2	2857655.7	2177433.7
29	Q3C1V8	1	1	14.2	0.0019728	0.0268284	5.551993	0.9615339	Patient	Sain	25.917	Brain-specific homeobox	119.76184	56.50442	16.821931	8.5075688	3888.8817	85.190432	6077.0842	3199.5835
30	P02652	11	11	995.34	0.0020124	0.0268284	2.4014067	0.9606031	Patient	Sain	11.168	Apolipoprotein A-II OS=Ho	34056410	26771849	14632439	15663194	42268641	26890744	89512485	43236635
31	P08571	12	12	937.27	0.0025308	0.0320335	1.6971067	0.9485492	Sain	Patient	40.051	Monocyte differentiation	797874.17	434080.26	499432.28	477652.53	733837.28	418725.29	352767.33	318079.37
32	P26927	31	31	1222.76	0.0025808	0.0320335	3.3123536	0.9474046	Patient	Sain	80.268	Hepatocyte growth factor-	190652.76	43028.557	155598.93	60550.701	75682.634	141408.77	141623.19	380083.05
33	P02655	6	6	819.16	0.0030504	0.0361789	4.7206762	0.9368237	Patient	Sain	11.277	Apolipoprotein C-II OS=Ho	1281811.8	751583.4	312951.49	1013499	603757.76	1691499.6	4859804.1	137623
34	P02671	8	8	369.69	0.0031158	0.0361789	191.8384	0.9353762	Patient	Sain	94.914	Fibrinogen alpha chain O	1792.4							
35	P02763	17	17	1378.36	0.0036234	0.0407586	2.8533872	0.924364	Sain	Patient	23.497	Alpha-1-acid glycoprotein	6747(							
36	Q14831	1	1	23.81	0.0038979	0.0425178	2.9788784	0.9185746	Sain	Patient	102.185	Metabotropic glutamate r	26519							
37	P02647	32	32	2367.65	0.0040205	0.0425648	2.2688435	0.9160271	Patient	Sain	30.759	Apolipoprotein A-I OS=Ho	10566							
38	P01876	21	21	1557.26	0.0044316	0.0455765	3.5420267	0.9076475	Sain	Patient	37.631	Ig alpha-1 chain C region	1095(							
39	Q9UNW1	8	8	227	0.005259	0.0525841	1.7900311	0.8915219	Sain	Patient	55.016	Multiple inositol polypho	62640							
40	Q16853	2	2	81.52	0.0054293	0.0528194	6.1436713	0.8883207	Sain	Patient	84.568	Membrane primary amine	24451							
41	P01833	8	8	219.66	0.006944	0.0656448	9.1555604	0.8614401	Sain	Patient	83.232	Polymeric immunoglobuli	14801							
42	P27169	28	27	2064.1	0.0071604	0.0656448	1.8486226	0.8578184	Patient	Sain	39.706	Serum paraoxonase/aryle	38610							
43	P60174	1	1	28.75	0.0074	0.0656448	2.9166777	0.8538651	Patient	Sain	30.772	Triosephosphate isomera	4225.(							
44	P00736	41	39	3067.37	0.0074771	0.0656448	1.4866701	0.8526056	Sain	Patient	80.067	Complement C1r subcomp	54346							
45	P09871	37	37	2588.95	0.0085558	0.0726349	1.93603	0.835606	Sain	Patient	76.635	Complement C1s subcomp	57898							
46	A0A0B4J1V0	3	3	104.52	0.0086769	0.0726349	2.7965659	0.833766	Sain	Patient	12.917	Immunoglobulin heavy va	92353							
47	P01619	4	4	262.21	0.0095577	0.0781901	2.30903	0.8207663	Sain	Patient	12.549	Immunoglobulin kappa va	40928							
48	Q08380	21	21	1485.84	0.0115602	0.0900161	2.528023	0.7934679	Sain	Patient	65.289	Galectin-3-binding protei	938408.99	614190.09	1091308.6	2622296.8	1676638.8	2590031.7	640186.94	906404.59
49	P06732	9	9	314.24	0.0117011	0.0900161	5.8938011	0.7916528	Patient	Sain	43.074	Creatine kinase M-type O	7288.867	9691.0748	1004.0051	6516.3248	4494.7392	6004.4549	19154.379	89665.94

Filtered data?

Peptide filters

- ❑ Unique peptides
- ❑ FDR <0.01 (or scoring)
- ❑ p-value <0.05 (t-tests, Anova)
- ❑ Scoring for PTMs localization (e.g ptmRS)

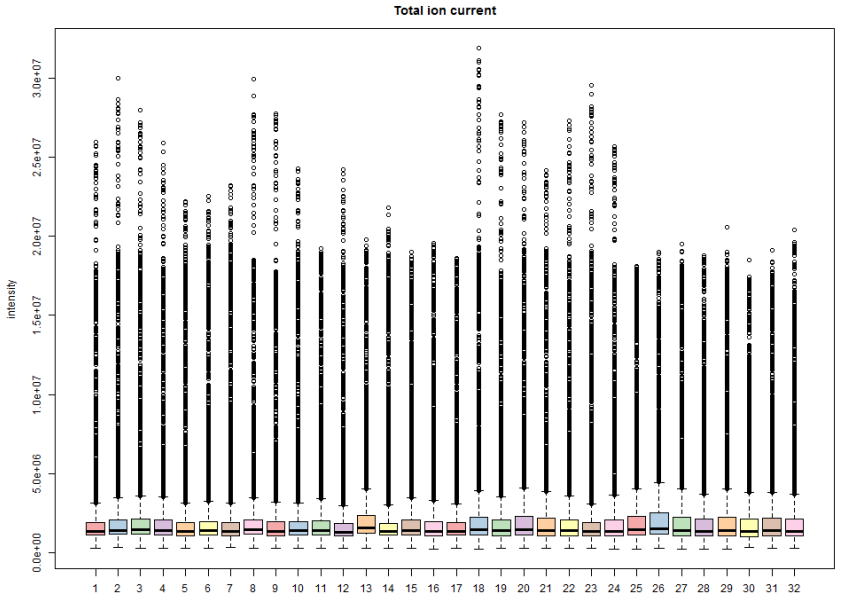
Protein filters

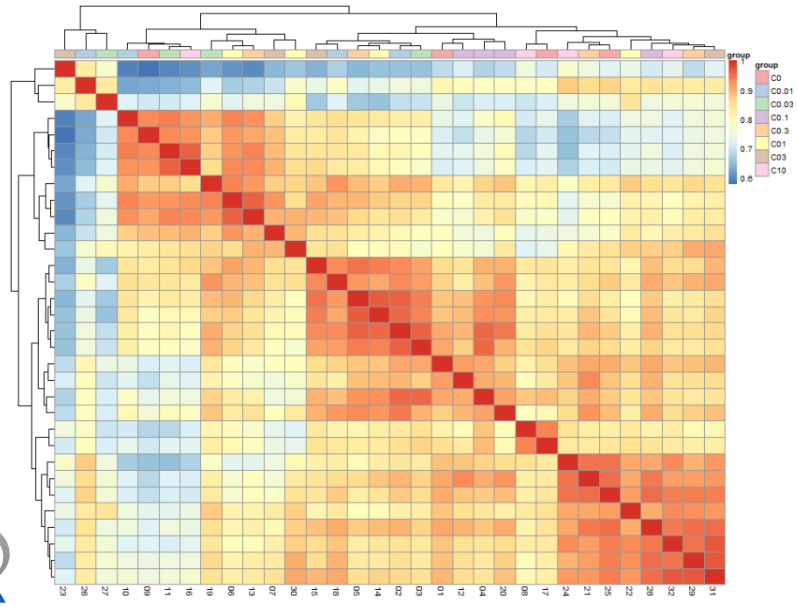
- ❑ 2 peptides by protein
- ❑ Unique accession
- ❑ FDR <0.01 (or scoring)
- ❑ p-value <0.05 (t-tests, Anova)

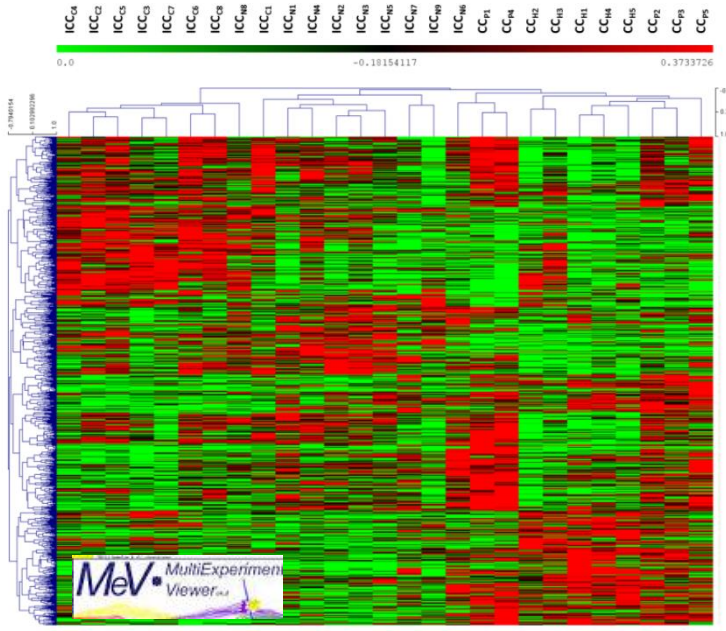
	Peptides						
Case 1	**1**	**2**	**3**	**4**	**5**	**Parsimony Reported?**	**Grouping Reported?**
Protein A	X	X	X	X	X	Yes	Yes
Protein B	X	X				No	No
Protein C			X	X	X	No	No
Case 2	**1**	**2**	**3**	**4**	**5**	**Parsimony Reported?**	**Grouping Reported?**
Protein D	X	X				Yes	Yes
Protein E		X	X			No	Yes
Protein F			X	X		Yes	Yes

No score ratio to compare a protein abundance between 2 conditions!
No score ratio to compare two different proteins in 1 sample!

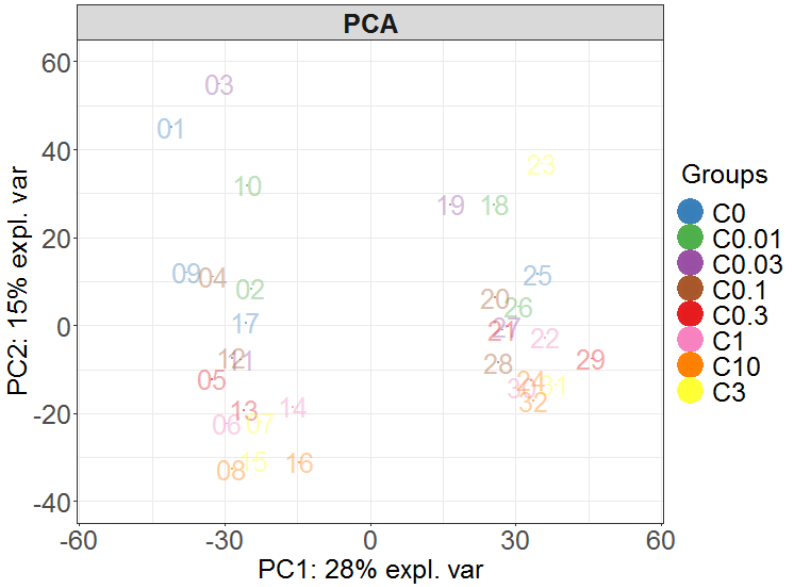
Unsupervised analyses



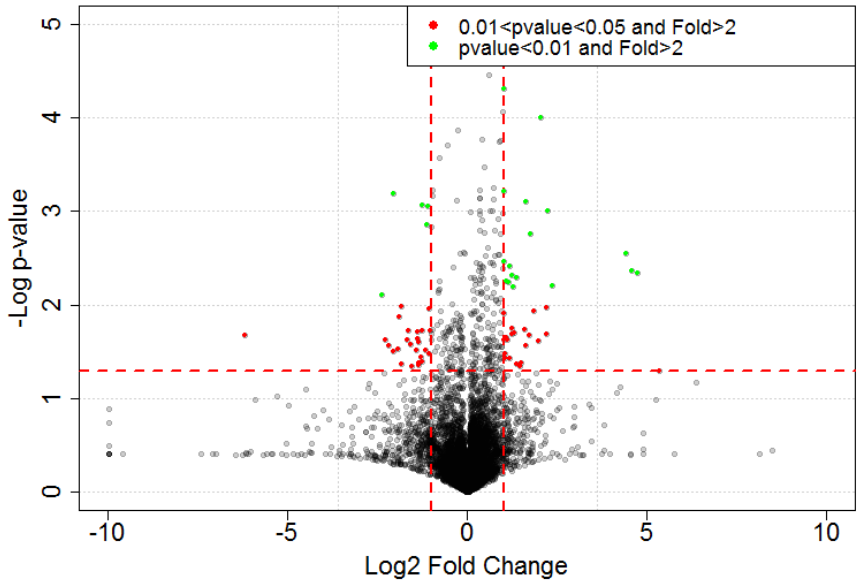


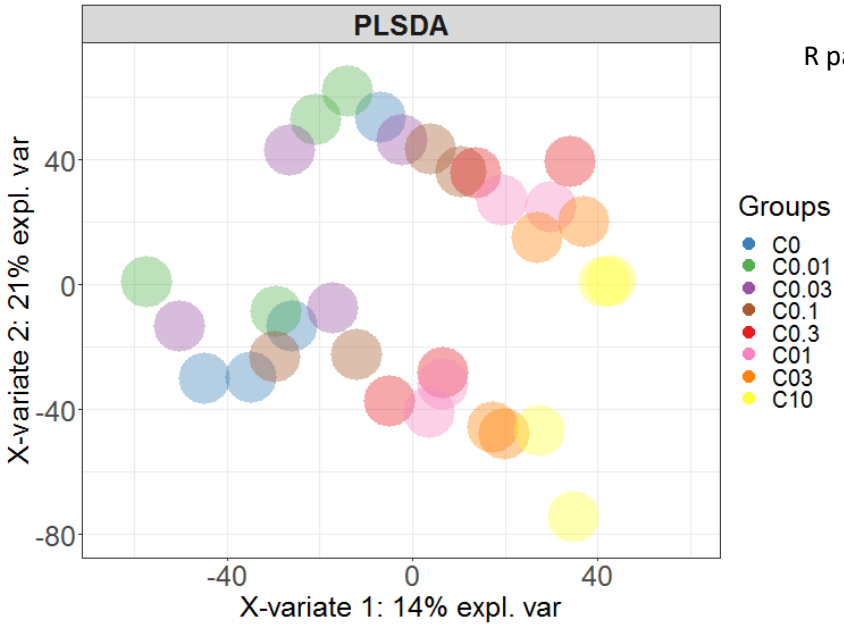


R packages:
pcaMethods
mixOmics
factoextra

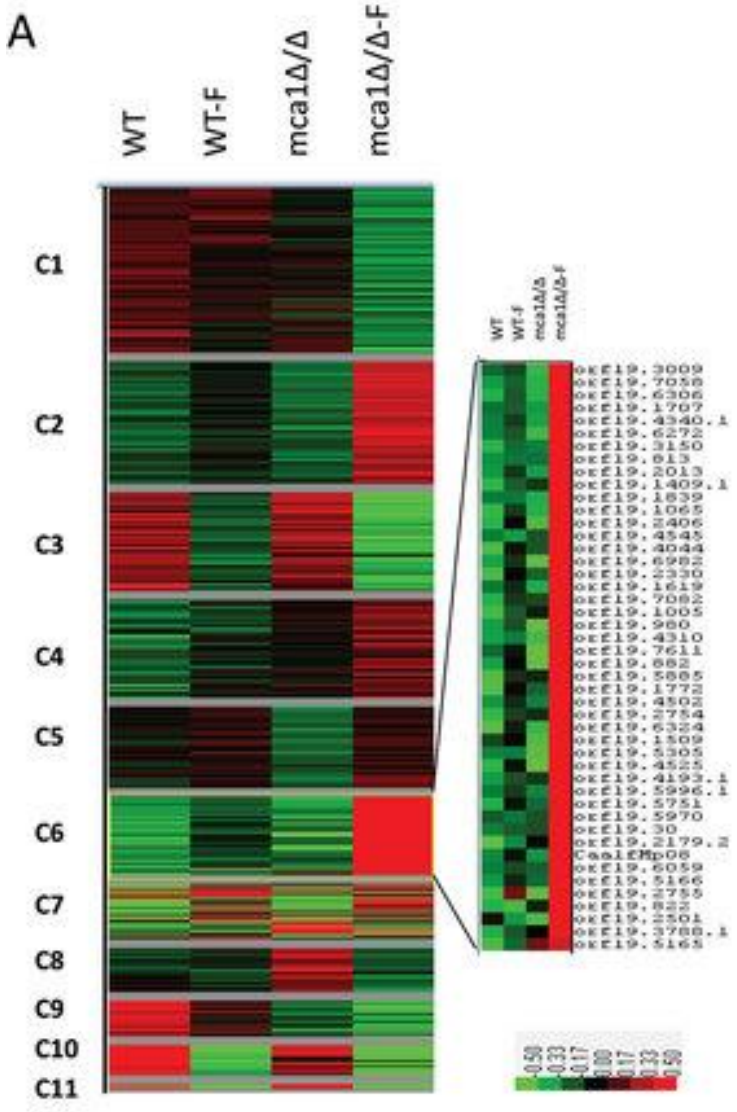


Supervised analyses





R packages:
ropls

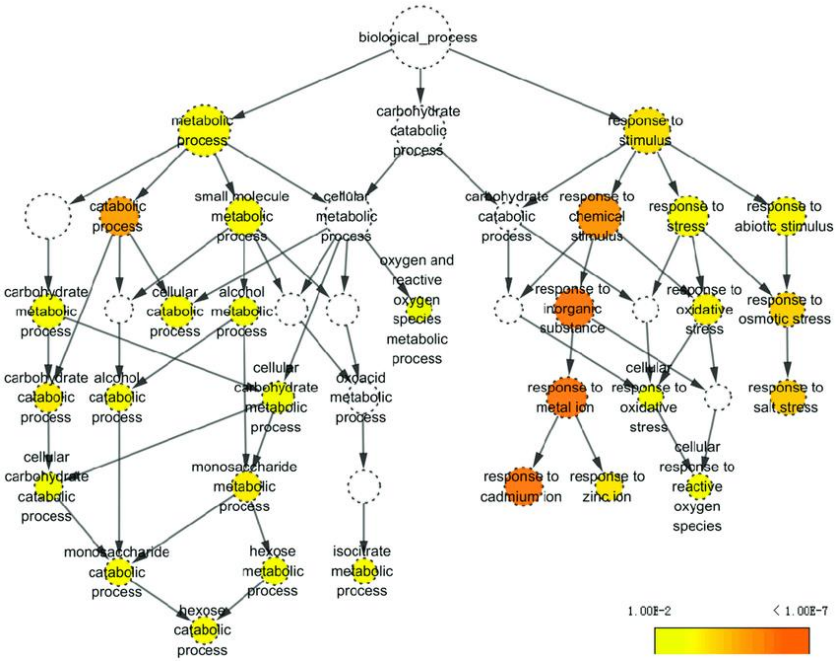


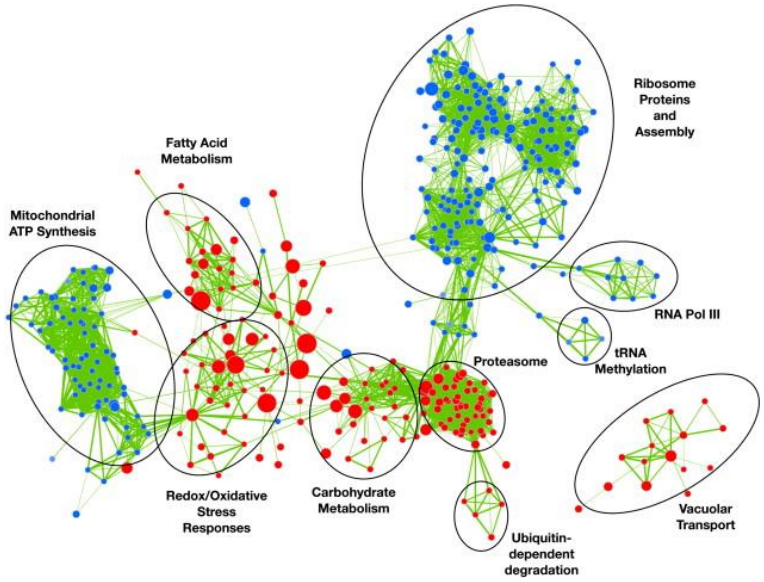
Leger et al. Mol Cell Proteomics 2015

Gene Ontology analyses

- Cytoscape plugins: BINGO, CLUEGO
- PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System
- GSEA (Gene Set Enrichment Analyses)
- David software
- Tools dedicated to specific species:
  - Candida Genome database (CGD)
  - Saccharomyces Genome Database (SGD)

Fisher's exact test / hypergeometric test





GENEONTOLOGY
Unifying Biology

Key questions in proteomics

- What is the protein content of my biological sample?

  => problem of identification

- What is the abundance of my protein of interest?

  => quantification problem

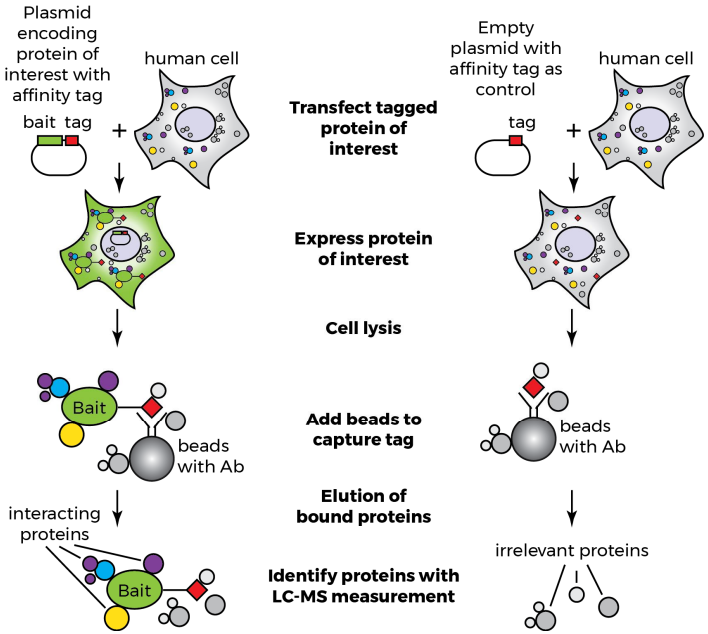
- Relative question: What are the protein abundance variations of the proteomes studied?

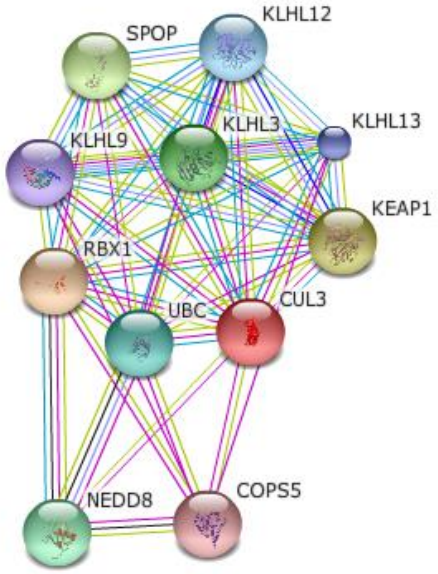
- What are the partners of my protein of interest?

- Are there any signature proteins related to a particular biological process?

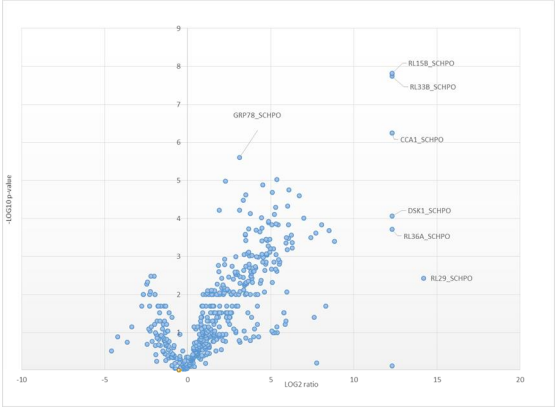
  => biomarkers identifications and quantifications

Affinity purification MS (AP-MS)





Description	Score A3	Score B3	Score C3
Cullin-3 OS=Homo sapiens GN=CUL3 PE=1 SV=2 - [CUL3...		1172.08	547.08
Cullin-associated NEDD8-dissociated protein 1 OS=Homo...	0.00	394.42	0.00
E3 ubiquitin-protein ligase RBX1 OS=Homo sapiens GN=R...		251.21	123.68
Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV...		199.36	136.63
Kelch-like protein 20 OS=Homo sapiens GN=KLHL20 PE=...		164.62	78.37
Kelch-like protein 13 OS=Homo sapiens GN=KLHL13 PE=...		158.68	86.54
Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE...		150.00	143.23
Kelch-like protein 42 OS=Homo sapiens GN=KLHL42 PE=...		149.87	62.55
ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 P...		148.12	57.72
COP9 signalosome complex subunit 5 OS=Homo sapiens...		142.70	29.32
COP9 signalosome complex subunit 8 OS=Homo sapiens...		135.68	37.43
Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B P...	37.89	135.00	103.98
Kelch-like protein 9 OS=Homo sapiens GN=KLHL9 PE=1 S...		131.61	110.90
ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 P...		131.31	78.50
Kelch-like protein 15 OS=Homo sapiens GN=KLHL15 PE=...		130.71	35.01
Kelch-like protein 12 OS=Homo sapiens GN=KLHL12 PE=...		127.87	23.09
Elongation factor Tu, mitochondrial OS=Homo sapiens GN...		112.52	119.93
Kelch-like protein 8 OS=Homo sapiens GN=KLHL8 PE=2 S...		110.71	28.11
COP9 signalosome complex subunit 6 OS=Homo sapiens...		106.43	25.42



Ilektra Kouranti (HEGP)

Key questions in proteomics

- What is the protein content of my biological sample?

  => problem of identification

- What is the abundance of my protein of interest?

  => quantification problem

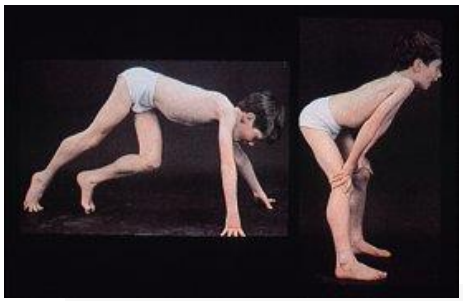
- Relative question: What are the protein abundance variations of the proteomes studied?

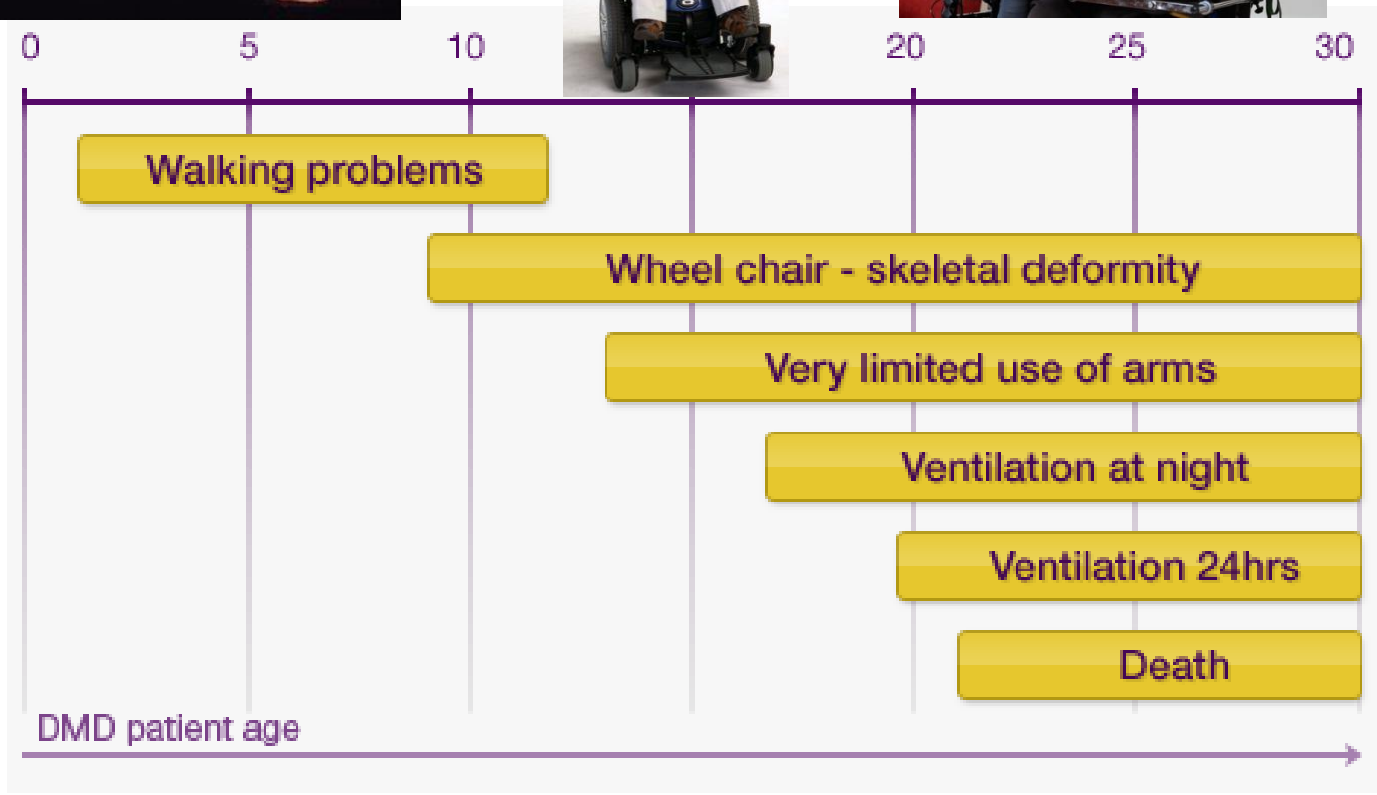
- What are the partners of my protein of interest?

- Are there any signature proteins related to a particular biological process?

  => biomarkers identifications and quantifications

Biomarkers: applications to Duchenne dystrophy





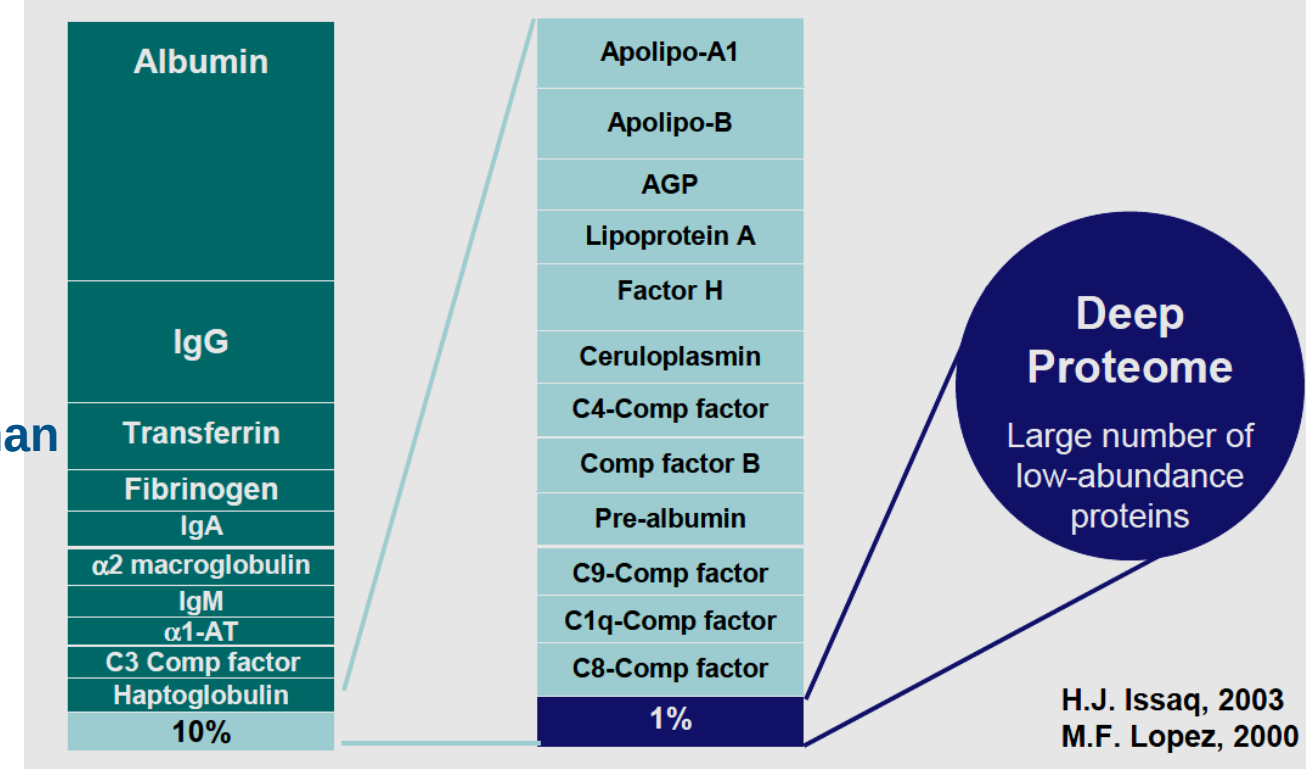
Å**Death** of DMD patients usually occurs ~ 30s

Serum: a ñtrickyò fluid for Mass Spectrometry

ŏ Serum : Mixture of proteins with different ranges of proteins concentration (from mg/ml to pg/ml)

ŏ 99% of serum proteome = 20 major proteins

ŏ 1% remaining = more than one thousand proteins



- Albumin : ~40 mg/ml (60% of serum proteome)

- C-reactive protein: ~1 µg/ml (40 000 times less than albumin)

- FGF-9 : ~400 pg/ml (100 000 000 times less than albumin)

Biomarkers: applications to Duchenne dystrophy

Description	No. of peptides	Score	ANOVA (*p*-value)	Fold change DMD/Healthy
Titin	23	1469.0	3.88E-06	37.4
Uromodulin	13	777.4	3.96E-03	5.5
Cubilin	10	576.8	2.55E-03	−2.3
Nuclear transport factor 2	5	356.9	1.05E-04	5.8
TNF-receptor superfamily member 16	4	308.7	4.07E-05	3.3
Myosin-1	3	265.3	8.66E-04	39.4
Fibulin-2	3	256.7	1.75E-03	2.9
β-galactosidase	6	253.7	1.23E-03	−2.4
Complement C1r subcomponent-like protein	5	235.8	3.47E-05	2.7
Aminopeptidase	3	213.7	2.71E-03	2.4

From: **"Proteomics profiling of urine reveals specific titin fragments as biomarkers of Duchenne muscular dystrophy", *J. Rouillon, A. Zocevic, T. Leger, C. Garcia, J-M. Camadro, B. Udd, B. Wong, L. Servais, T. Voit, F. Svinartchouk,* 2014 *Neuromuscular Disorders***

Biomarkers: applications to Duchenne dystrophy

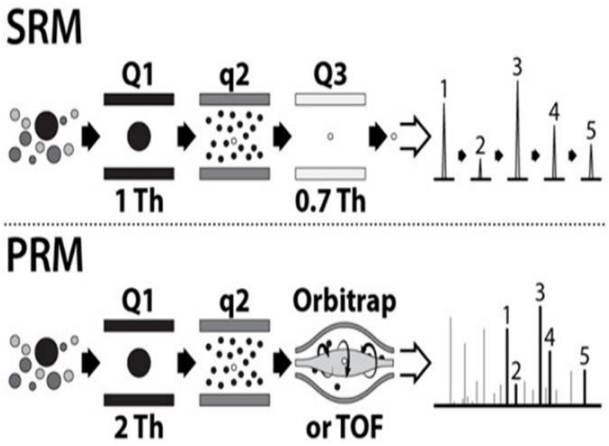
No. accession	Description	Localization	Peptides	Score	ANOVA (P-value)	Fold change
MYG_HUMAN	Myoglobin	Cytoplasm	4	195	2.7e-03	234.8
MYOM2_HUMAN	MYOM2	Myofibril	10	390	9.8e-05	100.1
MYOM3_HUMAN	**MYOM3**	**Myofibril**	**11**	**491**	**1.5e-05**	**49.7**
TPIS_HUMAN	Triosephosphate isomerase	Cytoplasm	3	128	2.3e-03	48.4
AATC_HUMAN	Aspartate aminotransferase	Cytoplasm	3	75	4.7e-04	45.7
KCRM_HUMAN	**CK-M**	**Cytoplasm**	**15**	**849**	**2.9e-05**	**39.8**
MYH7_HUMAN	**Myosin-7**	**Myofibril**	**11**	**520**	**2.2e-05**	**38.3**
ENOB_HUMAN	**β-enolase**	**Cytoplasm**	**4**	**178**	**7.4e-05**	**34.8**
G6PI_HUMAN	Glucose-6-phosphate isomerase	Cytoplasm/Secreted	4	130	1.6e-03	29.5
CAH3_HUMAN	**Carbonic anhydrase 3**	**Cytoplasm**	**5**	**182**	**8.6e-05**	**23.9**
FLNC_HUMAN	Filamin-C	Myofibril	4	145	4.3e-04	19.4
ALAT1_HUMAN	**Alanine aminotransferase 1**	**Cytoplasm**	**4**	**127**	**3.0e-05**	**15.6**
ALDOA_HUMAN	**Fructose-bisphosphate aldolase A**	**Cytoplasm**	**15**	**729**	**9.3e-05**	**14.2**
KPYM_HUMAN	**Pyruvate kinase PKM**	**Cytoplasm**	**16**	**845**	**1.1e-05**	**12.8**
TITIN_HUMAN	Titin	Myofibril	14	495	1.9e-03	10.8
VINC_HUMAN	**Vinculin**	**Cytoplasm/Membrane**	**2**	**74**	**7.2e-05**	**10.3**
PYGM_HUMAN	Glycogen phosphorylase, muscle form	Cytoplasm	8	257	6.1e-04	9.9
LDHA_HUMAN	L-lactate dehydrogenase A chain	Cytoplasm	8	378	9.1e-04	9.5
HPT_HUMAN	Haptoglobin	Secreted	29	1867	1.5e-04	7.6
HBD_HUMAN	Haemoglobin subunit δ	Cytoplasm	3	100	5.1e-03	6.2
LDHB_HUMAN	**L-lactate dehydrogenase B**	**Cytoplasm**	**10**	**598**	**2.4e-05**	**5.4**
HBB_HUMAN	Haemoglobin subunit β	Cytoplasm	7	552	8.0e-03	3.6
HBA_HUMAN	Haemoglobin subunit α	Cytoplasm	7	407	5.3e-03	3.4
TPM2_HUMAN	Tropomyosin βchain	Myofibril	5	170	2.0e-02	2.6

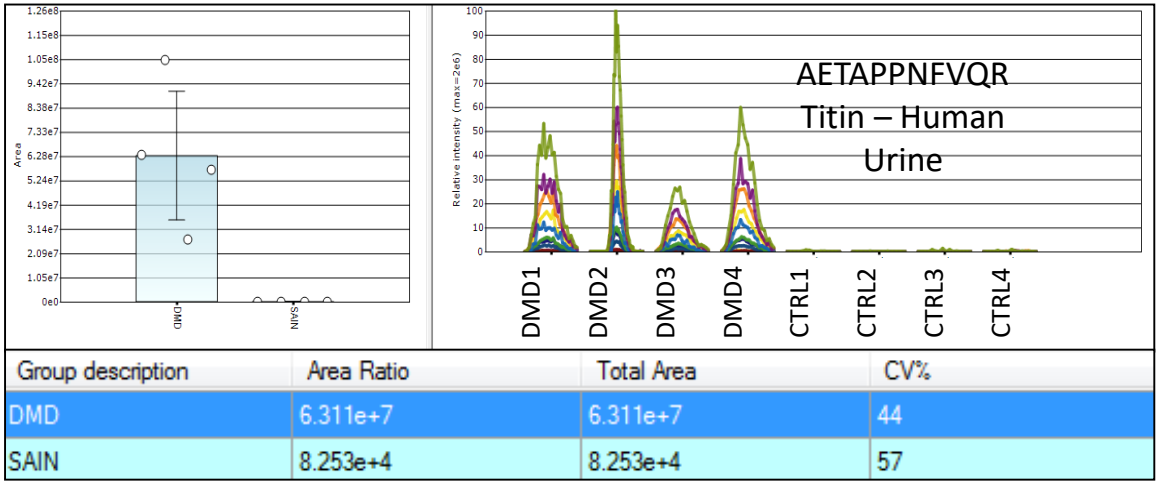
GENETHON
INNOVER POUR GUERIR

Rouillon, J., Zocevic, A., Poupiot, J., Amor, F., **Léger, T., Garcia, C., Camadro, J.M.**, Wong, B., Cosette, J., ML Coenen-Stass, A., McClorey, G., C Roberts, T., JA Wood, M., Servais, L., Voit, T., Richard, I., Svinartchouk, F. (2015). *Serum proteomic profiling reveals specific MYOM3 fragments as biomarkers of Duchenne muscular dystrophy with applications for the follow-up of gene therapy treatment in a mouse model of muscular dystrophies.* ï Human Mol. Genetics

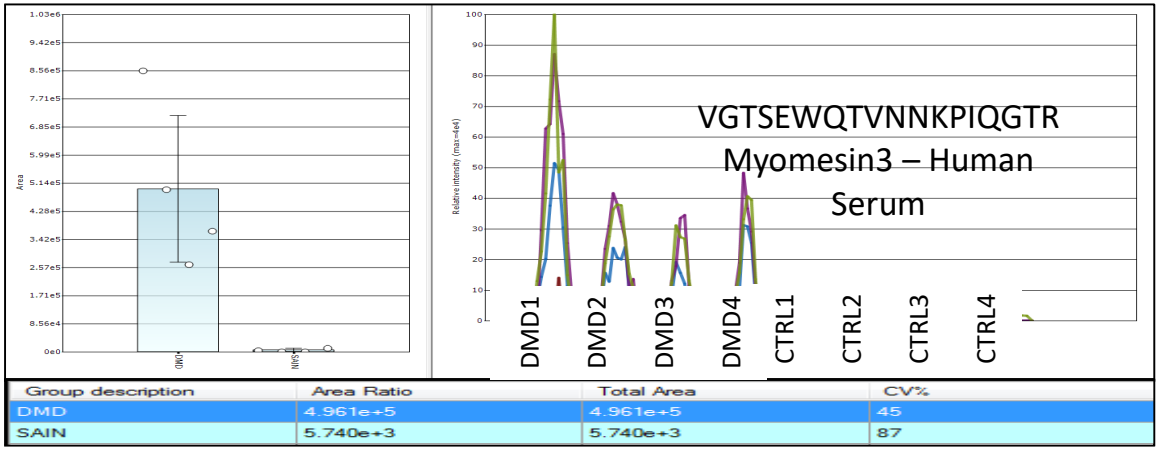
Jeremy Rouillon; Aleksandar Zocevic; **Thibaut Léger**; **Camille Garcia; Jean-Michel Camadro**; Bjarne Udd; Laurent Servais; Thomas Voit; Fedor Svinartchouk. (2014). *Proteomics profiling of urine reveals specific titin fragments as biomarkers of Duchenne muscular dystrophy.* Neuromuscular disorders.

Targeted proteomics PRM





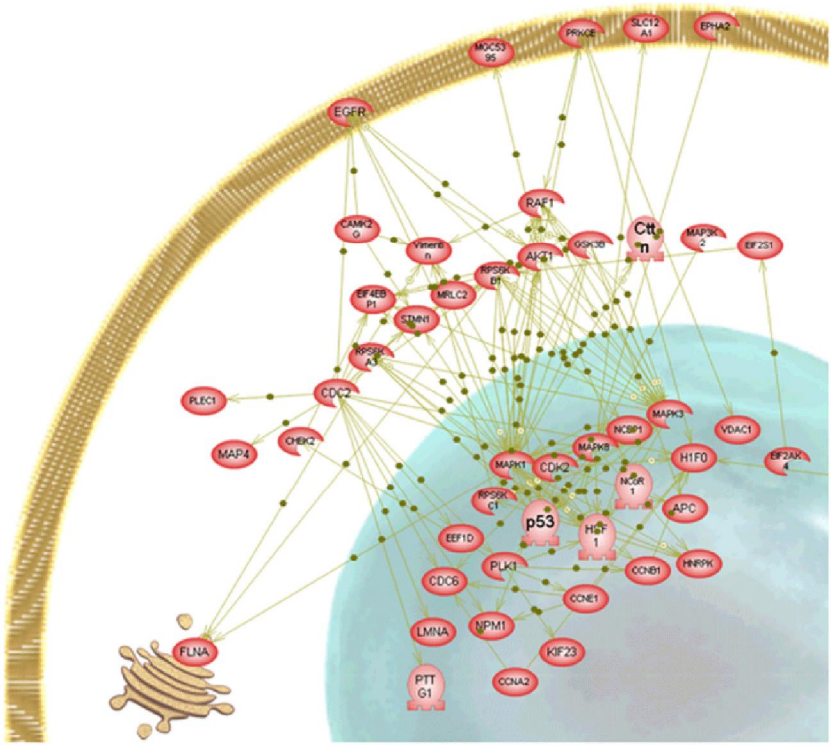
Group description	Area Ratio	Total Area	CV%
DMD	6.311e+7	6.311e+7	44
SAIN	8.253e+4	8.253e+4	57



Group description	Area Ratio	Total Area	CV%
DMD	4.961e+5	4.961e+5	45
SAIN	5.740e+3	5.740e+3	87

Why study PTMS?

- Cells can rapidly respond to stimuli and perturbations
- Important cellular mechanisms are tightly controlled
- Often, diseases (e.g. cancer) are due to aberrantly activated proteins
  - Protein expression is much too slow for quick adaption
  - PTMs are crucial regulator
  - MS-based proteomics allows to analyze complex networks of post-translationally modified proteins



http://www.omicsonline.org/ArchiveJPB/2010/April/03/JPB-03-113.php

PTMs *in vivo*

- **Phosphorylation** (Ser, Thr, Tyr; +80 Da)
  - Phosphorylation is one of the most important PTMs
  - A key event in signaling
  - Catalyzed by kinases/phosphatases
- **Glycosylation** (Asn, Ser, Thr)
  - marks proteins for degradation
  - s for degradation
- **Glycation** (Asn, Ser, Thr)
  - marks proteins for degradation
  - s for degradation
- **Ubiquitination** (Lys; +114 Da)
  - marks proteins for degradation
- **Proteolytic cleavage**
- **Acetylation** (N-termini and Lys +42 Da)
  - often combined with removal of protein initial Met

Others: oxidations, methylations, sumoylations, glutathionylations…

PTMs characterization: techniques

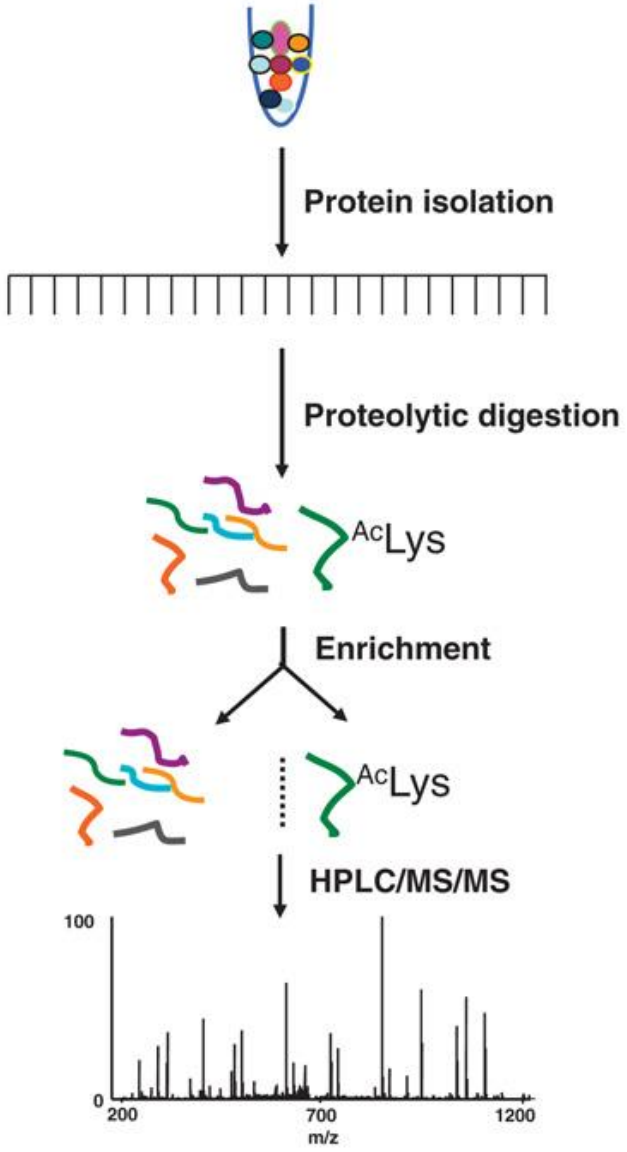
Techniques for detection and identification of PTM substrates

Method	*In vitro* or *in vivo*	Case studies	Advantages	Disadvantages
Radioactive isotope labeling	*In vitro* or *in vivo*	^{32}P (pSer, pThr, pTyr) ^{3}H, or ^{14}C for ^{Ac}Lys or ^{Me}K	Reagents accessible	Inconvenience/hazard low sensitivity
Western blotting	*In vitro* or *in vivo*	pTyr, ^{Ac}Lys or ^{Me}K	Good affinity	Moderate sensitivity
Peptide/protein array	*In vitro*	pSer/Thr/Tyr, ^{Ac}Lys or ^{Me}K	Rapid, global scale	Possibly non-specific, low sensitivity, requires verification
MS-proteomics	*In vitro*	pSer/Thr/Tyr, ^{Ac}R or ^{Me}K	Specific, global scale	Need enrichment methods

^{Ac}R, ^{Me}K, pSer, pThr, and pTyr, represent acetylated arginine, methyllysine, phosphorylated serine, threonine, and tyrosine residues, respectively.

Proteomics. 2009 Oct; 9(20): 4632–4641.

Workflow for PTMs characterization



Cell / Tissue sample
- Cell/tissue lysate
- Organelle prep
- Protein complex

Protein fractionation
- SDS PAGE
- Chromatography
- Precipitation

Peptide fractionation
- RP-HPLC
- SCX or SAX
- IEF

PTM enrichment
- ^{Ac}Lys-peptides: Antibody
- Phosphopeptides: IMAC/TiO2
- Glycopeptides: Lectin
- Chemical derivatization

HPLC-MS/MS & bioinformatics
- Peptide identification
- PTM localization
- PTM quantitation

Proteomics. 2009 Oct; 9(20): 4632–4641.

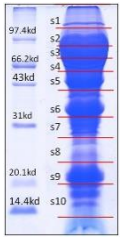
Quantitative proteomics and phosphorylations

CONDITION 1:
Mast cells inactivated
CO-IP PY20

Immunoaffinity

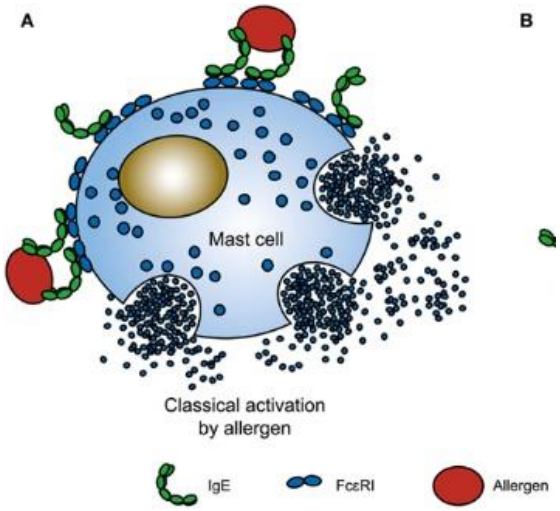
1 Hour

SDS-PAGE



Trypsin digestion overnight
(phosphate buffer)

CONDITION 1:
Mast cells activated (BC4)
CO-IP PY20

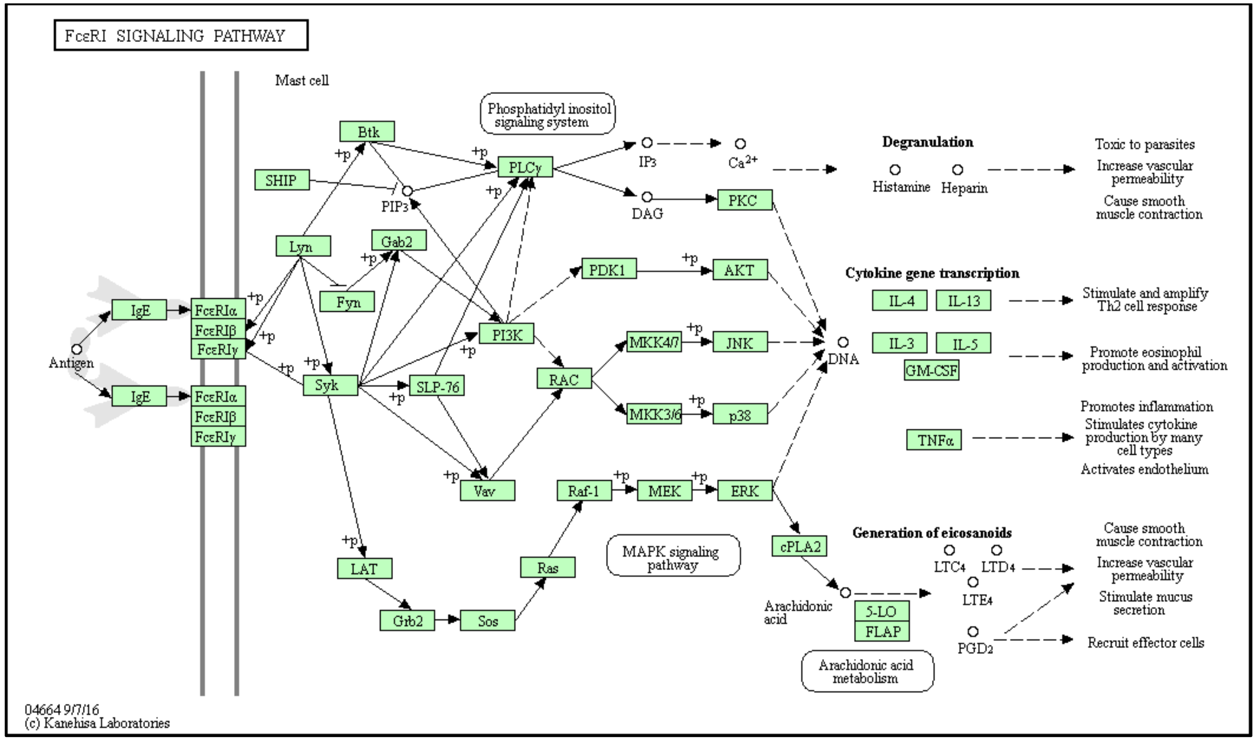


MATRIX SCIENCE

LC-MS/MS

Differential phosphorylated peptide and protein identification

		Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	MH+ [Da]	phosphoRS Site Probabilities	A4	IonScore A4	Exp Value A4	B4	IonScore B4
935	Q9JM80	Phosphoprotein associated with glycosphingolipid-enriched mic	nal transducer activity	membrane	; response to stimulus		763.10	646.62	45.9	4.81				
953	P20411	High affinity immunoglobulin epsilon receptor subunit gamma C	nal transducer activity	ell surface; membrane	to stimulus; transport	Pf11628	573.65	352.21	9.8	6.00				
963	Q64725	Tyrosine-protein kinase SYK OS=Rattus norvegicus GN=Syk P	nal transducer activity	cleus; organelle lumen	to stimulus; transport	Pf00017; Pf00069; Pf07714	835.28	323.20	71.5	8.15				
964		Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	MH+ [Da]	phosphoRS Site Probabilities	A4	IonScore A4	Exp Value A4	B4	IonScore B4
965		LLTLEDNELGSGNFGTVK	2	1	1	Q64725		1906.97726		High	107	1.50125E-10		
966		DESEQTVLIGSK	4	1	1	Q64725		1305.65435		High	84	2.49757E-08	High	67
967		ELNGTYAISGGR	7	1	1	Q64725		1237.61919		High	55	2.38967E-05	High	58
968		ADENYYK	3	1	1	Q64725		902.38894		High	46	0.000100972	High	23
969		NVLLVTQHYAK	8	1	1	Q64725		1285.72683		High	44	0.000137239	High	31
970		ISDFGLSK	6	1	1	Q64725		866.46208		High	42	0.000234007	High	40
971		LIATTAHEK	7	1	1	Q64725		983.55124		High	37	0.000403124	High	37
972		LRNYYYDVVN	4	1	1	Q64725		1318.64326		High	37	0.001486192	ediu	24
973		YLEESNFVHR	4	1	1	Q64725		1293.62261		High	36	0.00176677	High	27
974		MGCPPGCPR	4	1	1	Q64725	M1(Oxidation); C3(Carbamidon	1047.41679		High	35	0.000734469	High	41
975		GSEVTAMLEK	2	1	1	Q64725		1064.52982		High	34	0.002136328		
976		GSEVTAMLEK	4	1	1	Q64725	M7(Oxidation)	1080.52370		High	34	0.002154235	ediu	20
977		EVYLDRK	2	1	1	Q64725		922.49926		High	33	0.001493862		
978		ALRADENYYK	2	1	1	Q64725		1242.61182		High	30	0.007003353		
979		VLTVPCQK	5	1	1	Q64725	C6(Carbamidomethyl)	944.52278		Medium	25	0.020099256		
980		GSEVTAMLEKGER	1	1	1	Q64725	M7(Oxidation)	1422.68938		Medium	25	0.027162273		
981		LRNYYYDVVN	6	1	1	Q64725	Y4(Phospho)	1398.61077	Y(4): 93.9; Y(5): 5.7; Y(6): 0.4	Medium	24	0.023646873		
982		TGPFEDLKENLIR	2	1	1	Q64725		1531.81073		Medium	22	0.025122601		
983		GSEVTAMLEKGER	2	1	1	Q64725		1406.69446		Medium	20	0.067444564		
984		KPFNRPPGVQPK	1	1	1	Q64725		1364.77961		Medium	19	0.035807446		
985		LLTLEDNELGSGNFGTVKK	1	1	1	Q64725		2035.07124		Low	19	0.093821803		
986		WYAPECINYFK	1	1	1	Q64725	C6(Carbamidomethyl)	1490.67815		Low	17	0.171057347		
987		NYYYDVVN	2	1	1	Q64725		1049.45671		Low	16	0.121475657		
988		MGCPPGCPR	1	1	1	Q64725	C3(Carbamidomethyl); C7(Carb	1031.42131		Low	15	0.108098357		
989		MPWFHGNISR	1	1	1	Q64725	M1(Oxidation)	1260.59497		Low	13	0.348397318		
990		YLQQNR	1	1	1	Q64725		821.42674		Low	12	0.268819476		
991		NYLGGFALSVAHNR	1	1	1	Q64725	Y2(Phospho)	1598.76055	Y(2): 99.9; S(9): 0.1				Low	11
992	Q5U2U2	Crk-like protein OS=Rattus norvegicus GN=Crkl PE=1 SV=1 - [	protein binding	m; cytosol; membrane	development	Pf00017; Pf00018; Pf07653	234.55	96.32	33.8	6.74				
1005	P60868	40S ribosomal protein S20 OS=Rattus norvegicus GN=Rps20 F	tural molecule activity	sm; cytosol; ribosome	metabolic process	Pf00338	172.28	36.35	13.4	9.94				



Proteomics workflows

Top ï Down

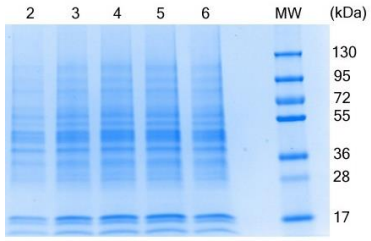
(< 50000 Da in particular conditions)

Proteins

Peptides

HPLC

PROTEASES

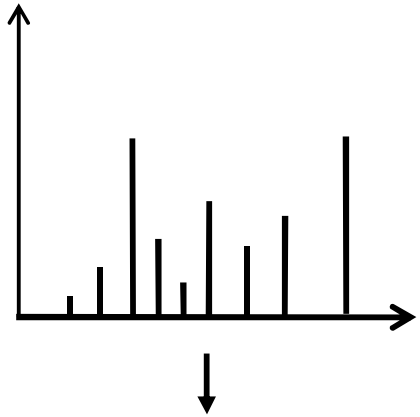


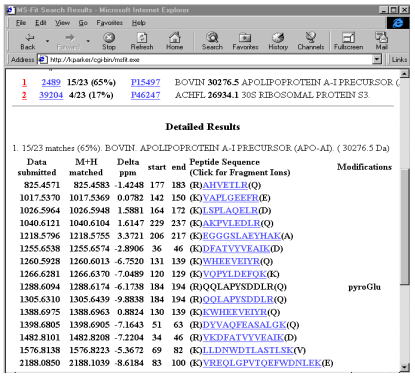
Bottom ï Up

(< 4000 Da)

MS acquisition

- MALDI TOF/TOF
- Orbitrap



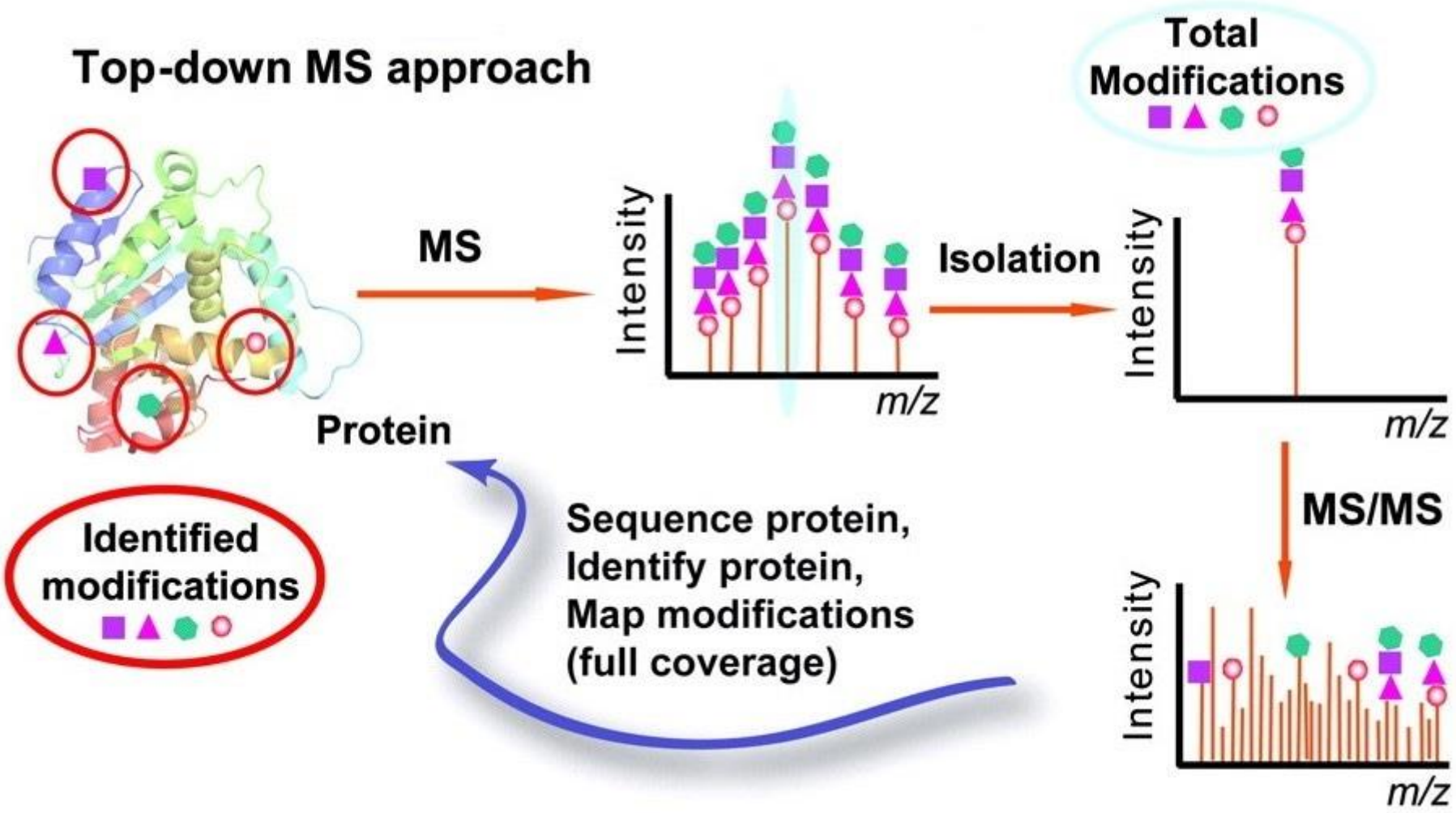


Identification
Quantification

The strategy is dependent of the sample complexity

Search engine (Mascot, Sequest, Peaks, Maxquant, OMSSA, Prosight, Byonik)

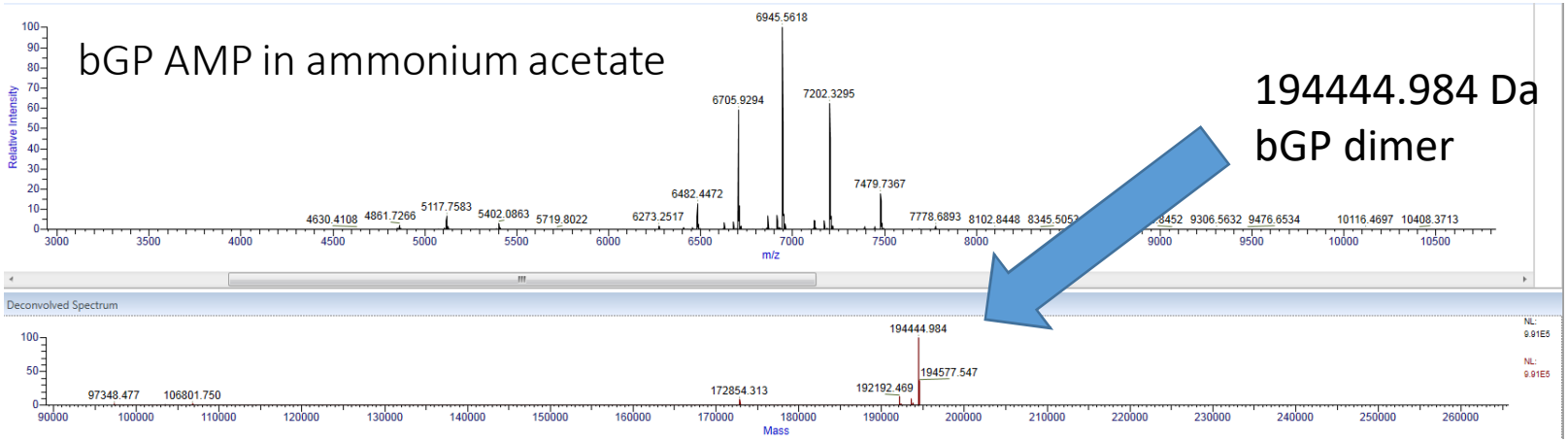
TOP DOWN proteomics for PTMs characterization

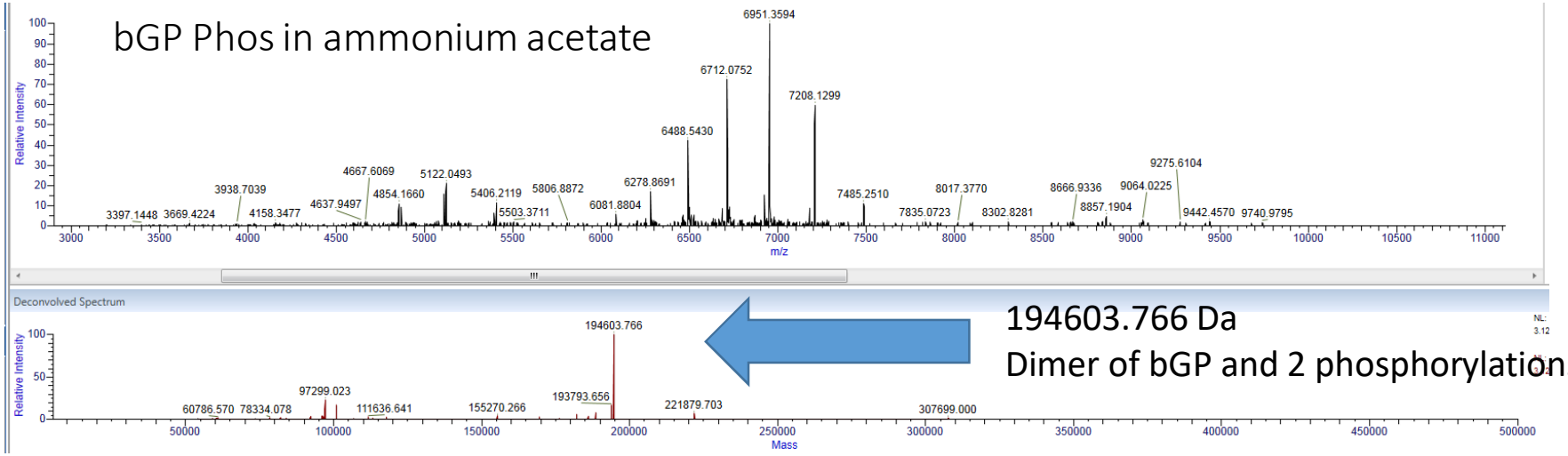


Challenges in TOP-DOWN proteomics

Challenges	Innovations
1. Protein solubility Conventional surfactant (e.g. SDS) not compatible with MS	**Develop new top-down MS compatible surfactant**
2. Proteome complexity Intact protein chromatography under-developed	**Develop novel multi-dimensional chromatography for intact protein separation**
3. Proteome dynamic range Difficulty in detecting low abundant proteins	**Develop novel nanomaterials for enriching low abundant proteins**
4. Protein MS data interpretation Software for top-down proteomics underdeveloped	**Develop user-friendly and versatile software interface**

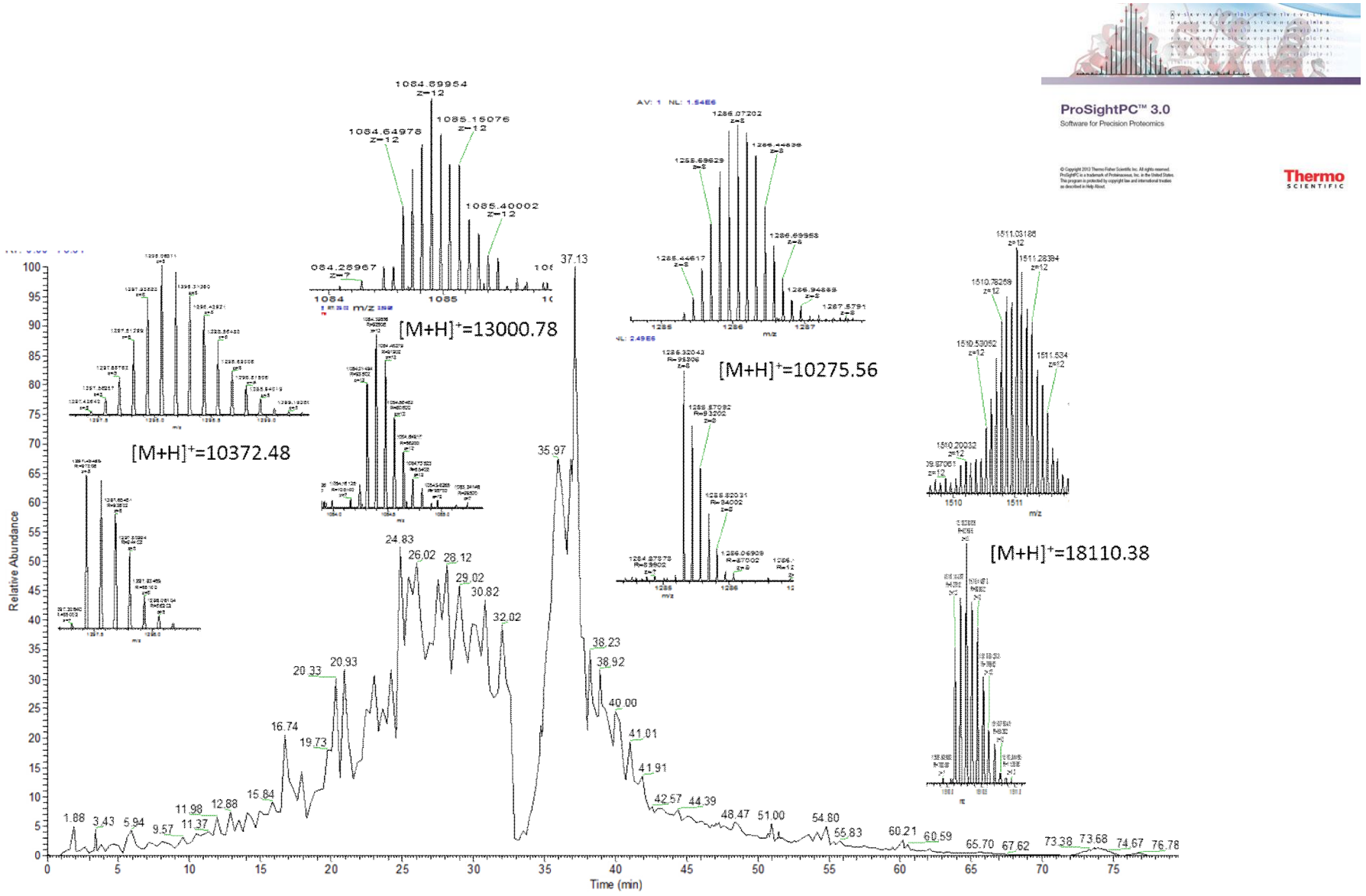
Analysis in intact protein mode: human brain glycogen phosphorylase



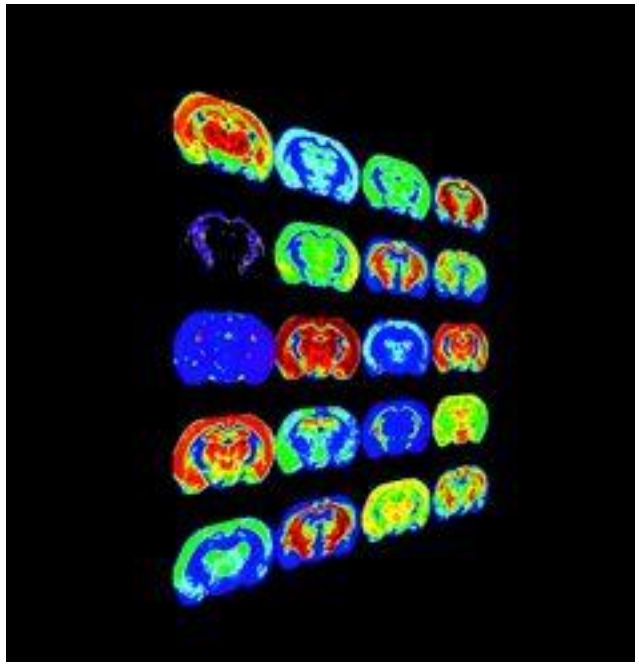


Crystal structure of human brain glycogen phosphorylase. Cécile Mathieu, Ines de la Sierra-Gallay, Romain Duval, Ximing Xu, Angélique Cocaign, Thibault Léger, Jean-Michel Camadro, Catherine Etchebest, Ahmed Haouz, Jean-Marie Dupret, Fernando Rodrigues-Lima. Under review.

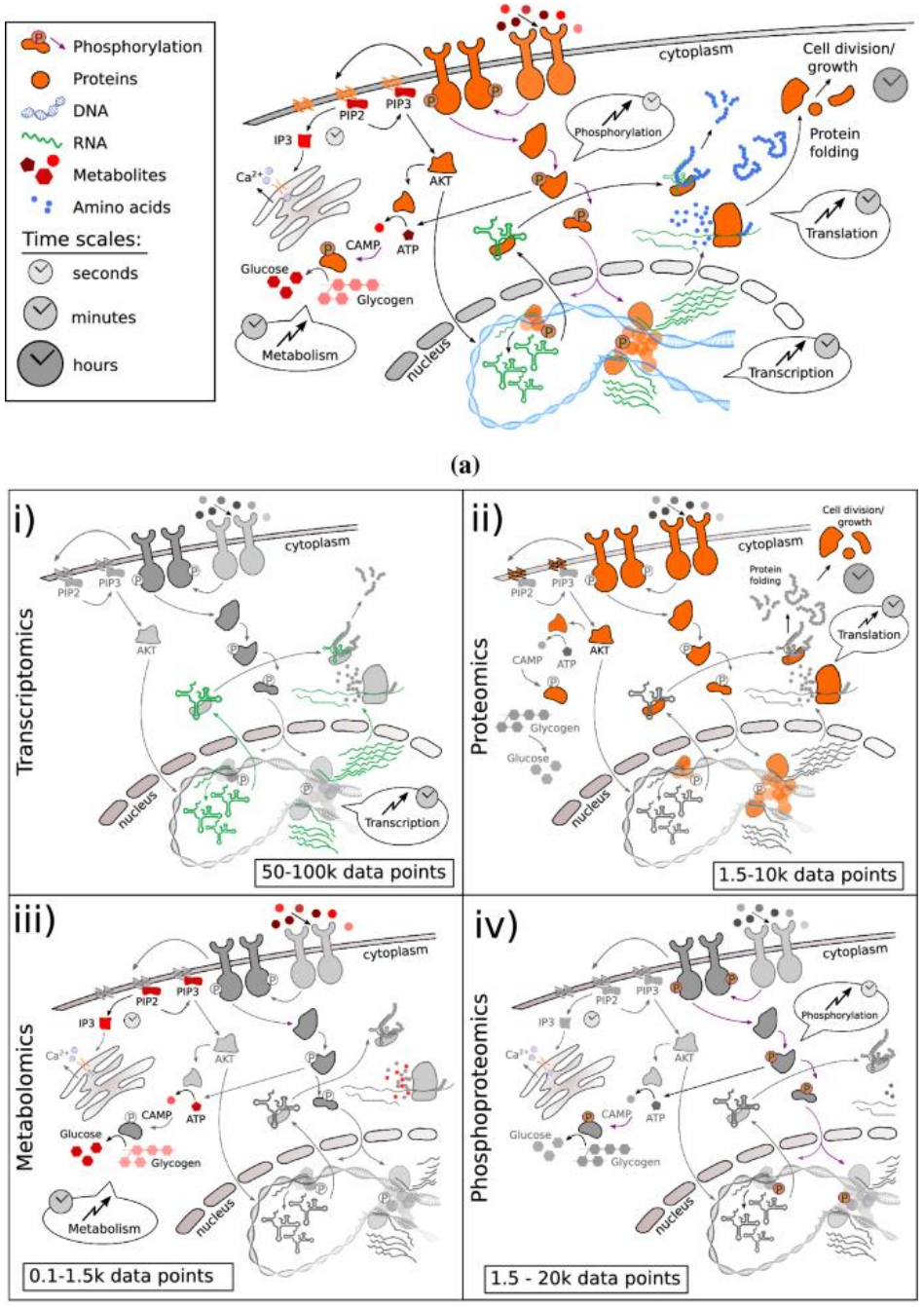
Analysis in intact protein mode



Emerging MS technologies

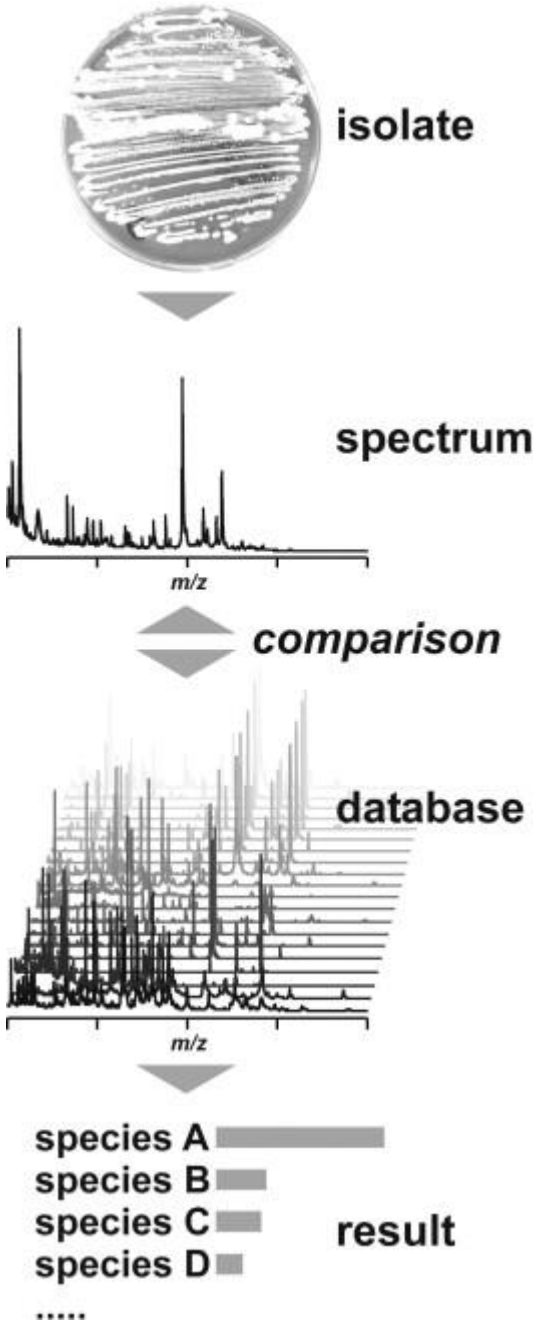


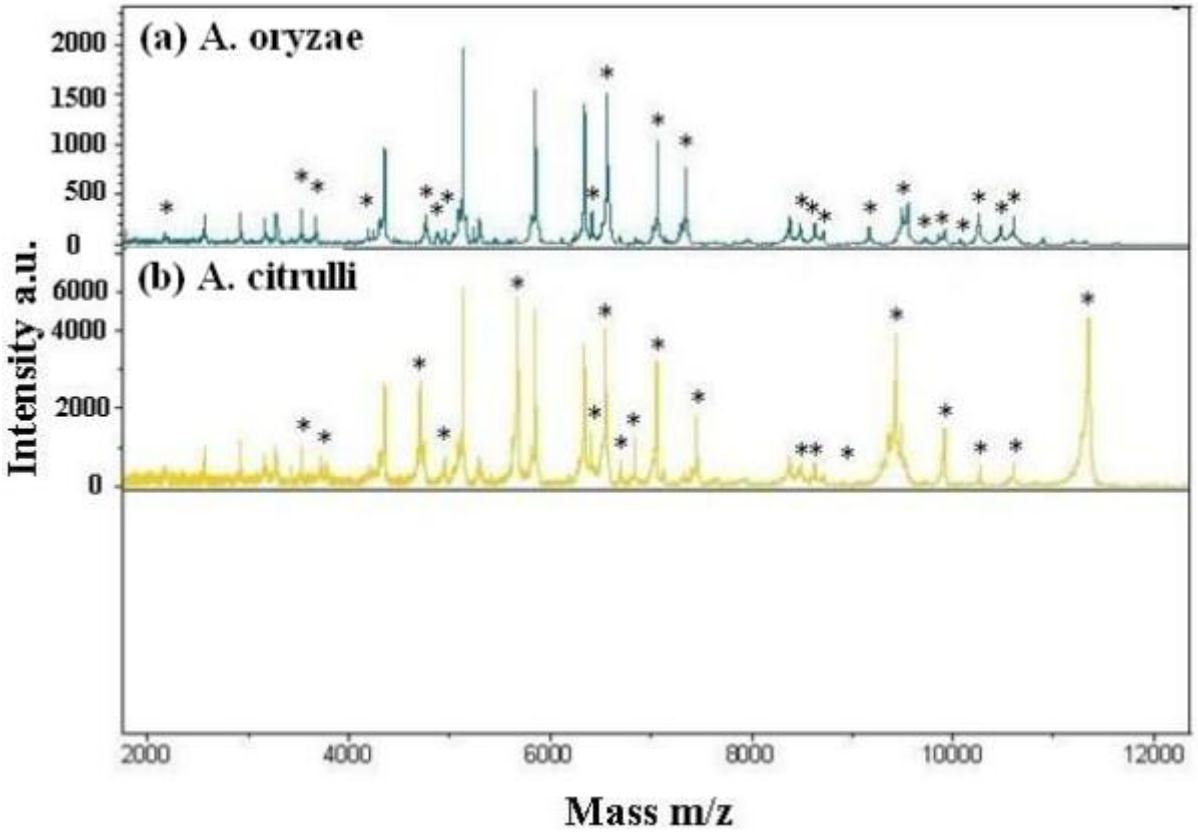
Multi-omics



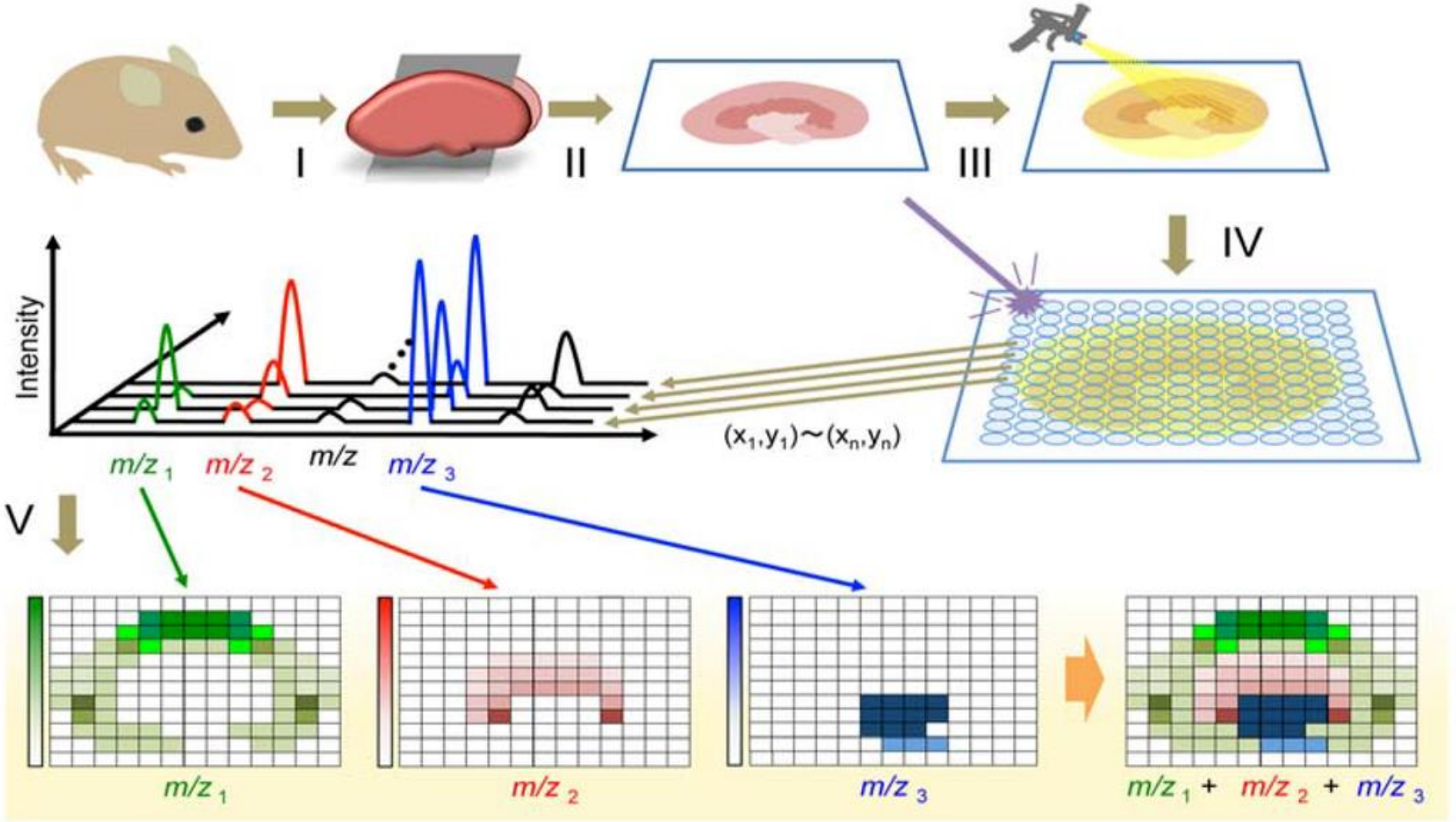
Canzler et al., Prospects and challenges of multi-omics data integration in toxicology 2020

MS identification in Bacteriology (Biotyper)



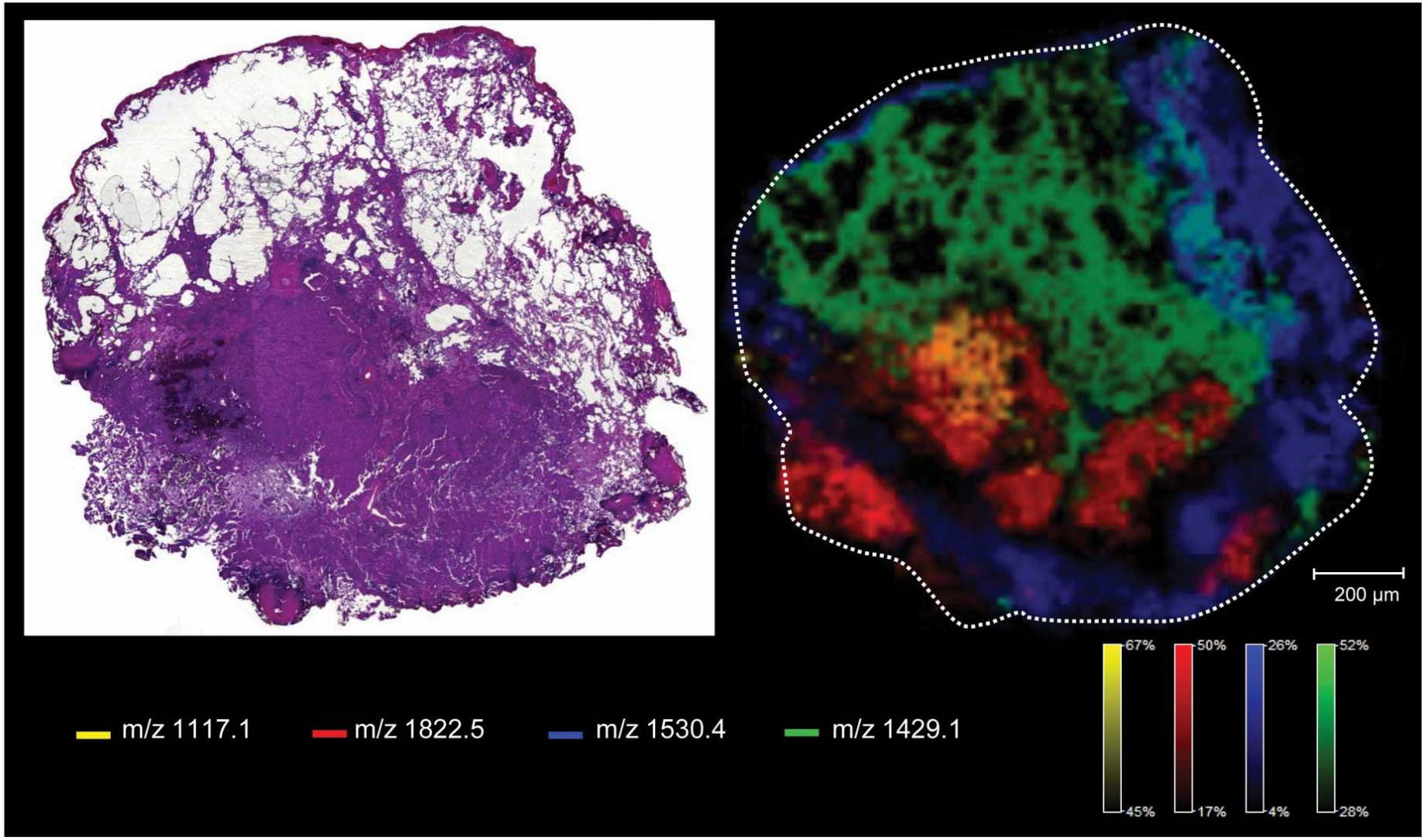


Mass spectrometry imaging



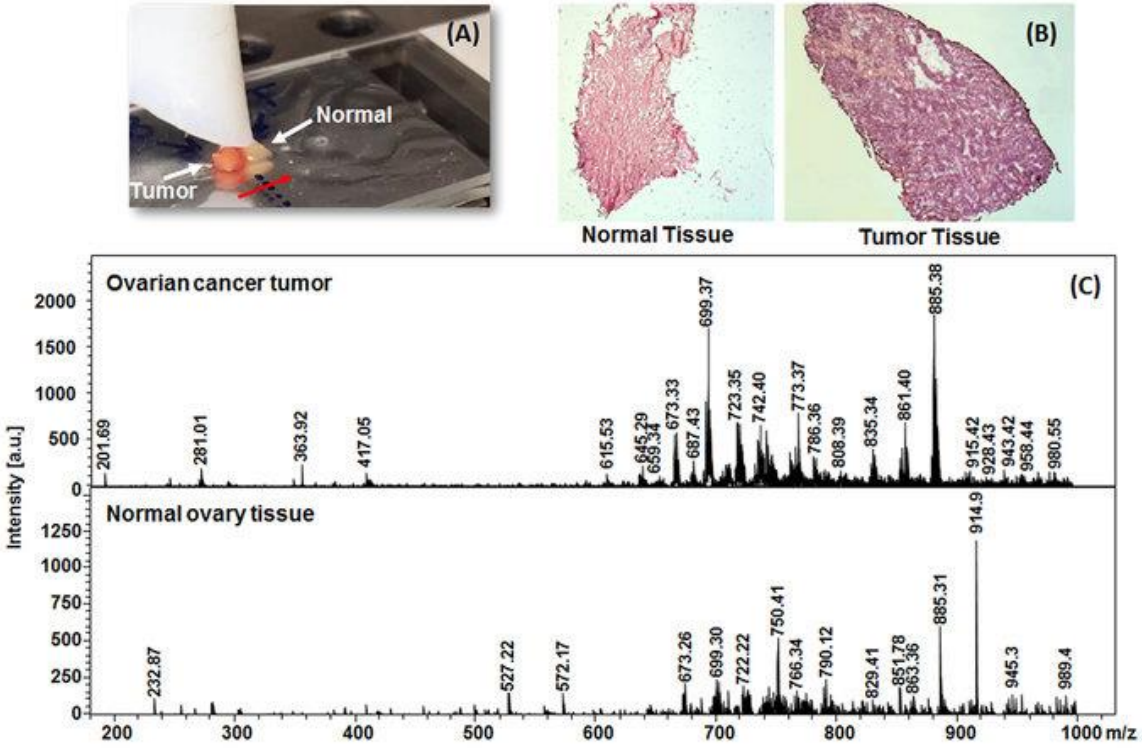
I Sacrifice and organ dissection
II Cryosectioning and moving to ITO glass slide
III Matrix deposition
IV MALDI laser 2D scanning
V Reconstruction of intensity image

Mass spectrometry imaging

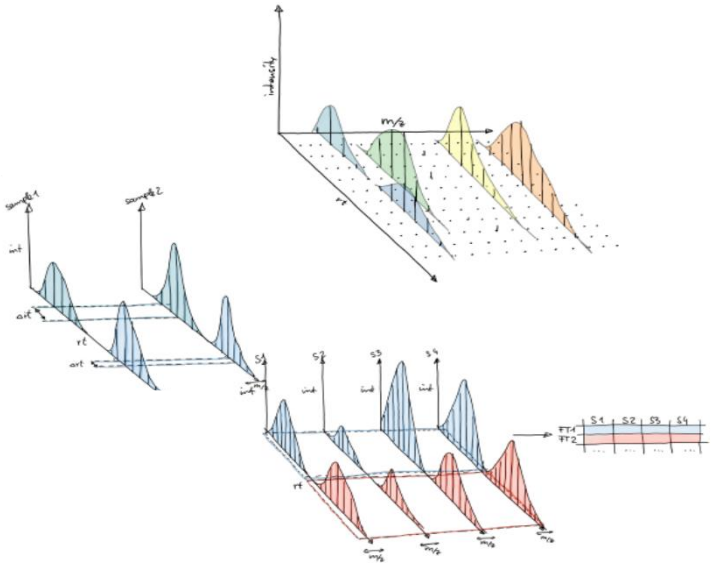


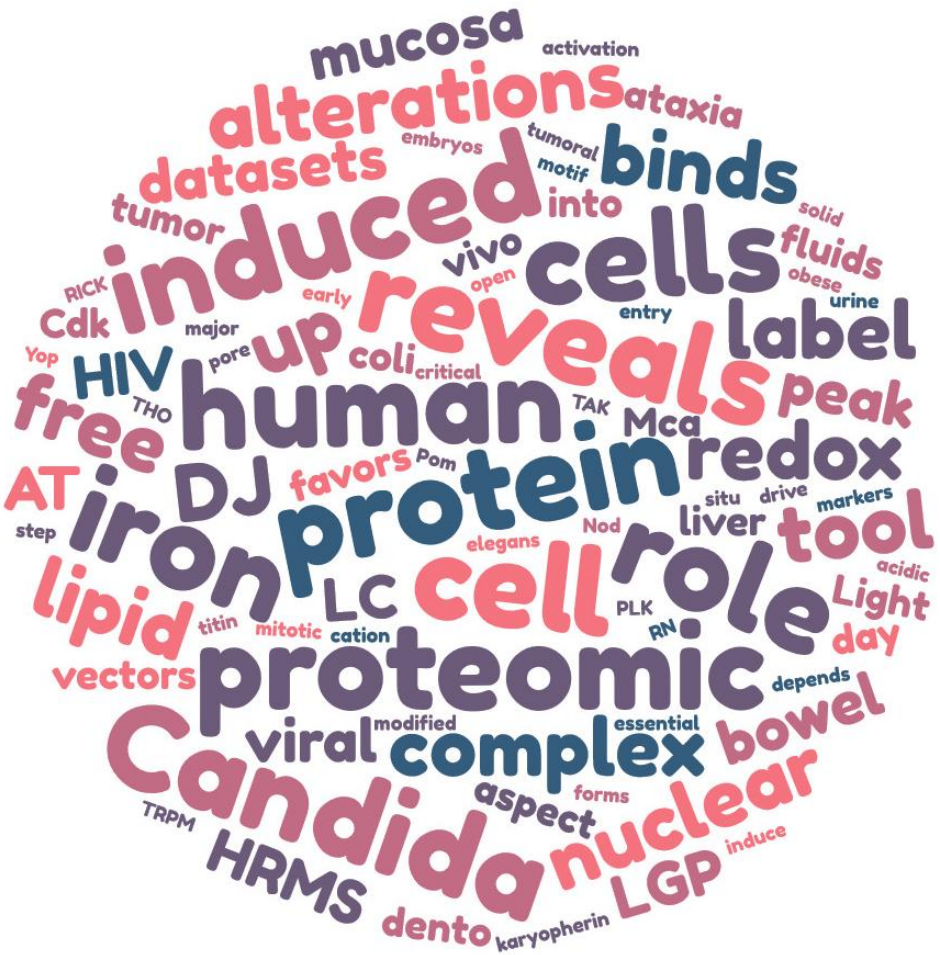
SPIDERMASS

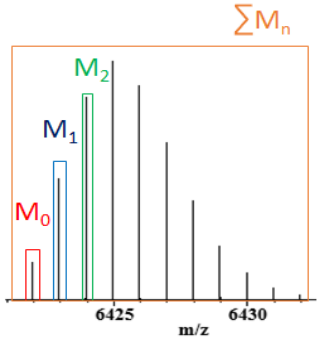
PRISM
Protéomique • Réponse Inflammatoire • Spectrométrie de Masse
INSERM U 1192 – Université de Lille



Thank you for your attention!







anses