Towards the Discovery of Biomarkers in CerebroSpinal Fluid by combining Peptide Ligand Library Treatment and Label Free Protein Quantification on a LTQ-Orbitrap

Anne Gonzalez de Peredo

Florence Roux-Dalvai

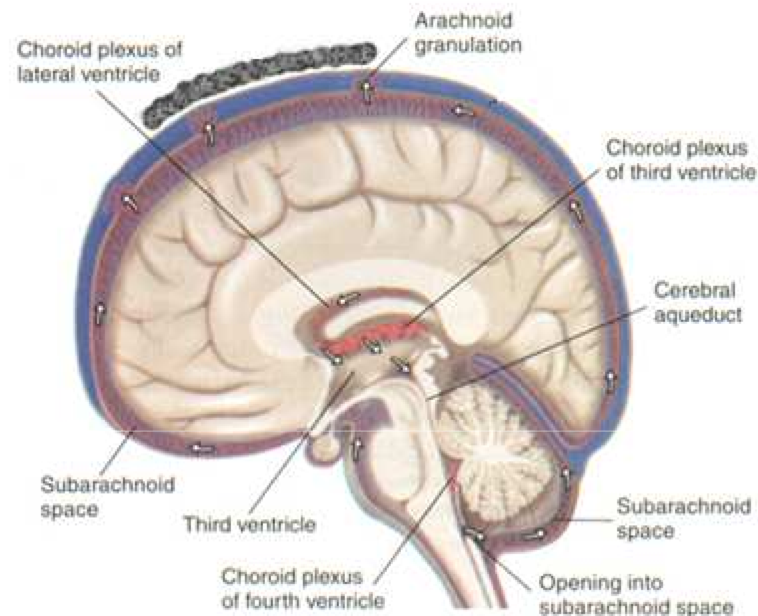
Emmanuelle Mouton

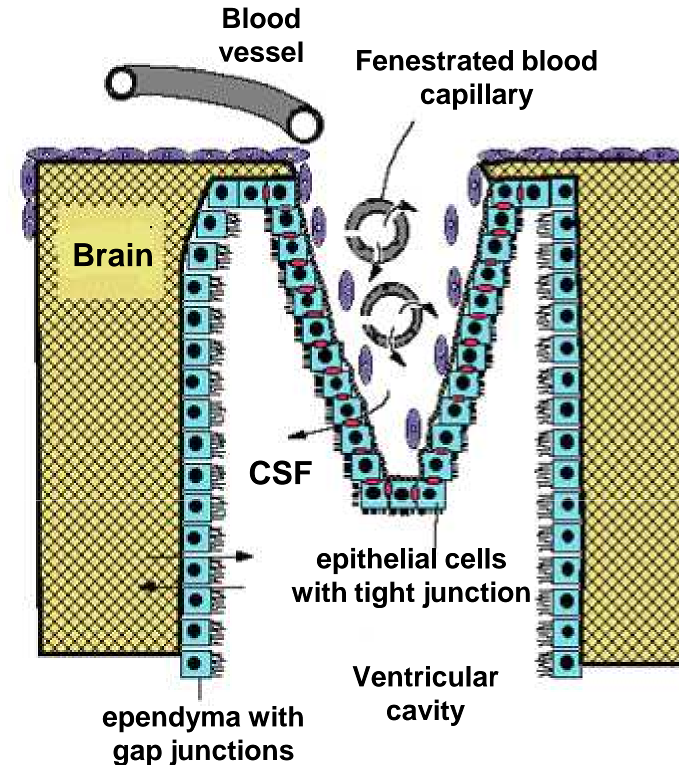
David Bouyssié

Bernard Monsarrat group

CNRS - IPBS
National Center for Scientific Research
Toulouse, France

About Cerebro-Spinal Fluid…





- Physical protection of the brain and metabolic function
- Secreted by the choroid plexus
- A small amount of CSF originates from the extracellular space of the brain

⇨ Potential biomarkers for neurological diseases

About Cerebro-Spinal Fluid…

Bottlenecks for proteomic studies on CSF:

Protein amounts

- Low protein concentration : 0.40 mg/ml (200x less than serum)
- Low available volume : Lumbar puncture = 1 to 2 mL

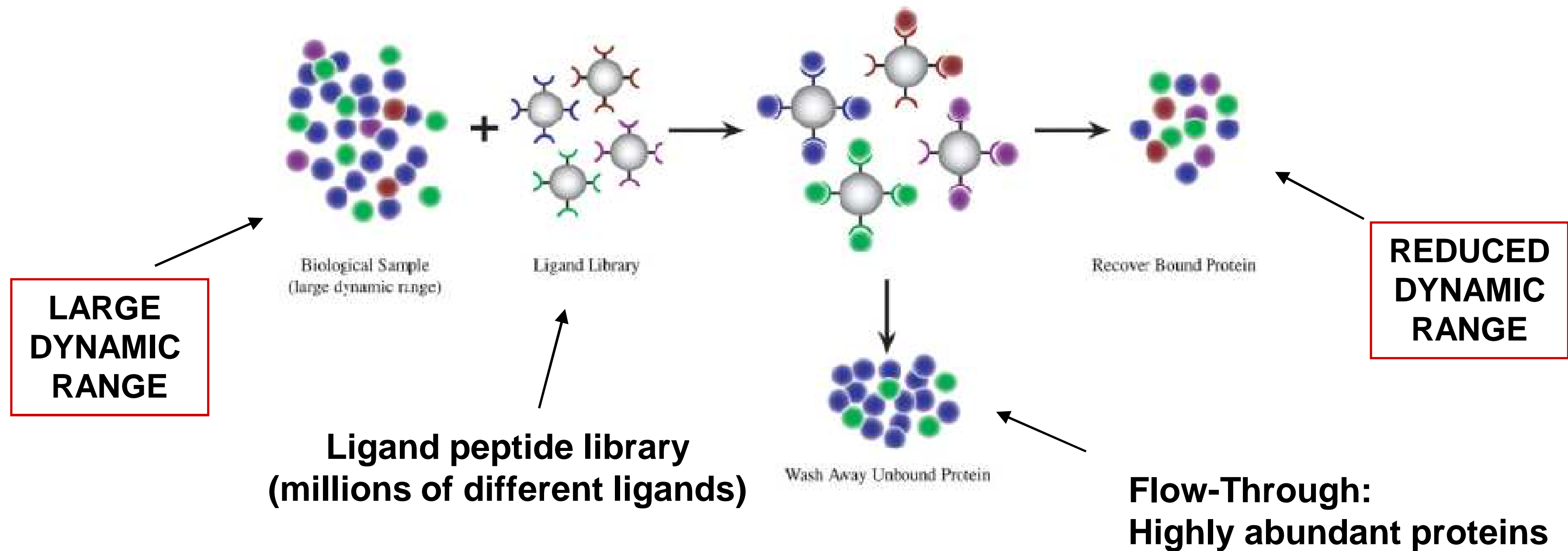
Dynamic range

- Very large range of protein concentrations : 10^{10}
- Albumin : 45% of total protein

Requirements for biomarker discovery using proteomic methods:

- Detection of a large number of proteins (low abundant proteins)
- Protein quantification with a good accuracy
- Analysis of a large number of patient samples

Reduction of sample dynamic range with ProteoMiner^{TM} beads



- Porous 60µm hydrophilic beads
- Hexapeptides synthetized by a combinatorial chemistry

- ~20^6 (64 millions) of different ligands
- Two libraries : NH2 terminus and COOH terminus

Optimal conditions for ProteoMiner^{TM} treatment

- The maximal ligand diversity is obtained with 1ml of beads
- A large overloading : typically 50-100X beads capacity (10mg/mL)

ProteoMiner^{TM} treatment of CSF

Applications:

- **Urine** (Castagna *et al. J Prot Research 2005*)
- **Platelets** (Guerrier *et al. J Prot Research 2007)*
- **Serum** (Sennels *et al. J Prot Research 2007)*
- **Red blood cell** (Roux-Dalvai *et al. Mol Cell Proteomics 2008)*

Application to CSF

CSF

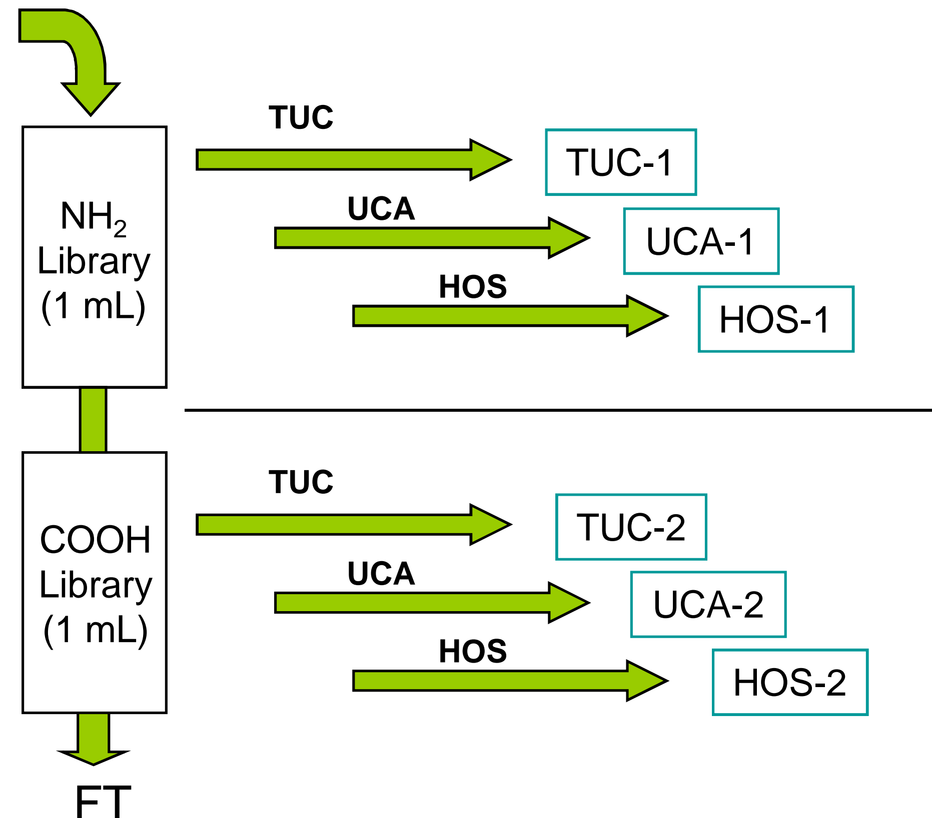
Pool of different patients

3 sources (Paris, Grenoble, Toulouse)

~ 1000 ml

770mg total protein on 1mL of beads

(overloading 77x)

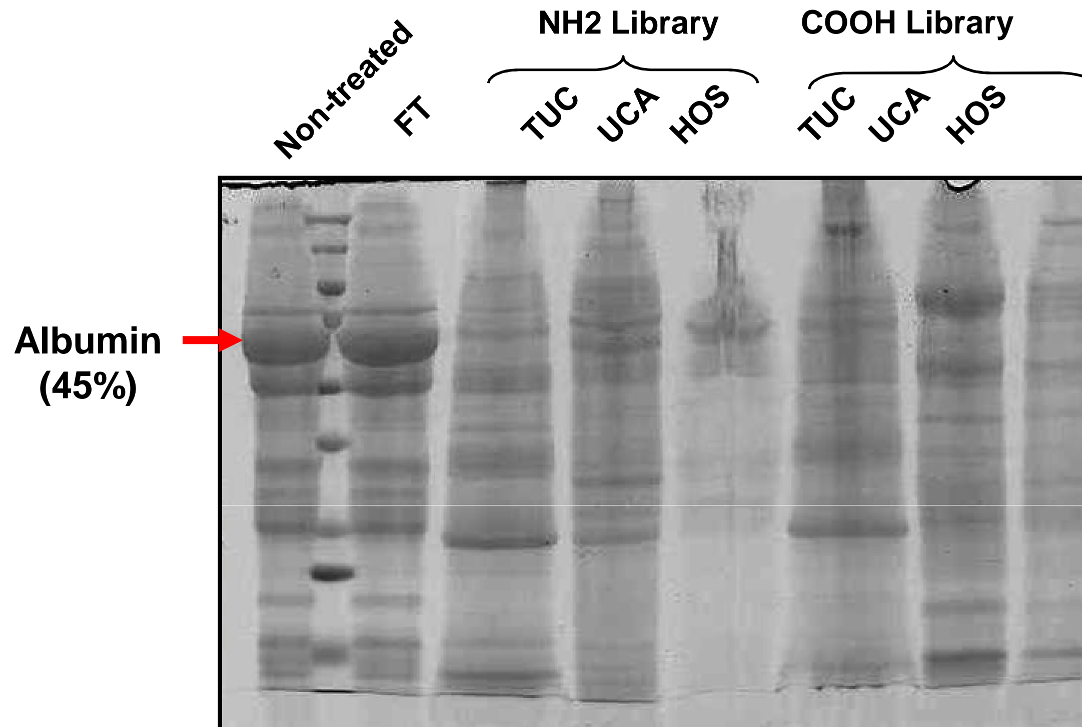


- **Serially connected columns carrying the two libraries**

- **Three different elutions for each column (10mg protein in total)**

TUC : Thiourea, Urea, CHAPS
UCA : Urea, Citric Acid
HOS : Hydro-Organic Solution

ProteoMiner^{TM} treatment of CSF



SDS-PAGE gel
150µg of protein for each sample → **Cut 20 bands per lane** → **Trypsin digestion** → **NanoLC-MSMS on LTQ-Orbitrap**

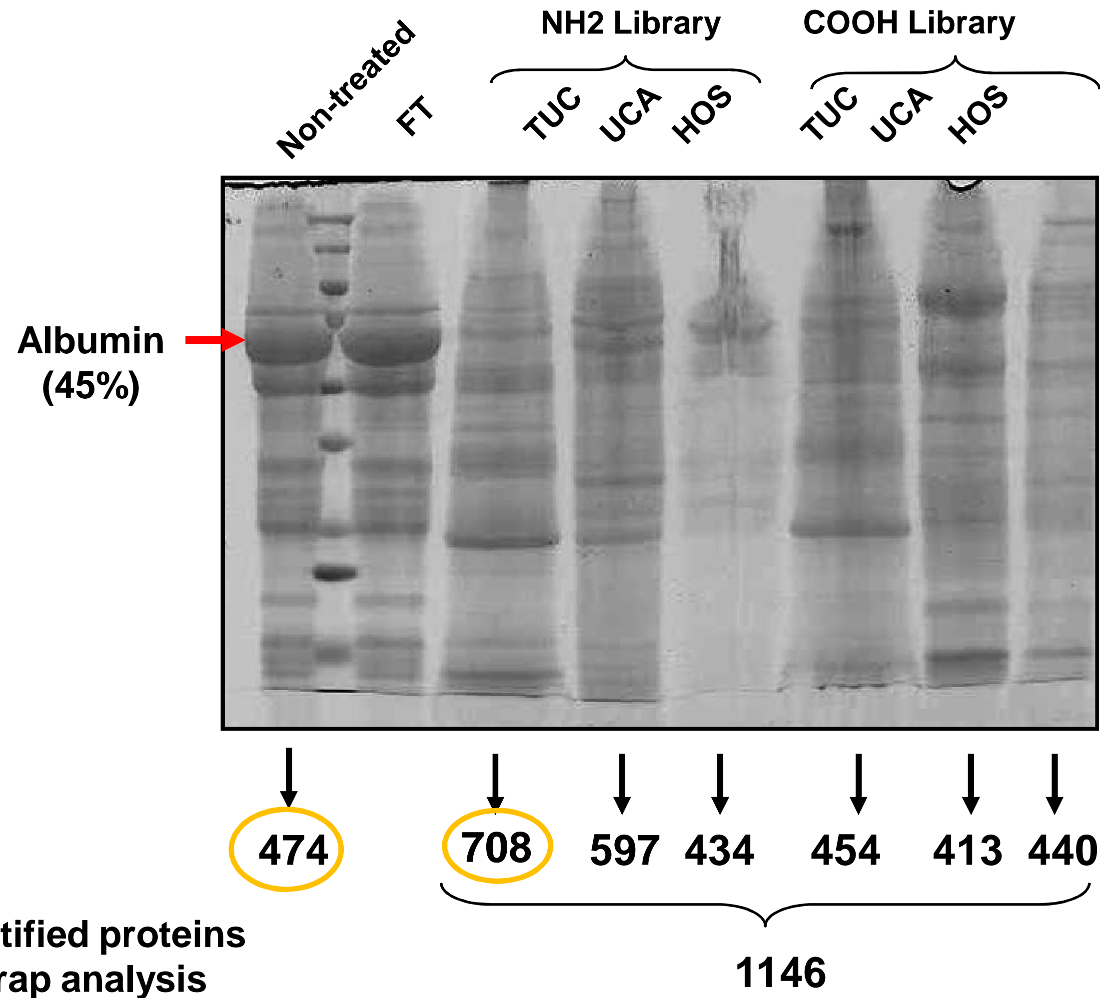
↓

MASCOT database search

↓

Protein validation with MFPaQ

ProteoMiner^{TM} treatment of CSF



Number of identified proteins
by LTQ-Orbitrap analysis

Proteomic analysis of CSF by nanoLC-MS/MS

Identification of low-abundant neuronal proteins

Proteins associated with the **GO term «neurogenesis»**:

Proteins in yellow:

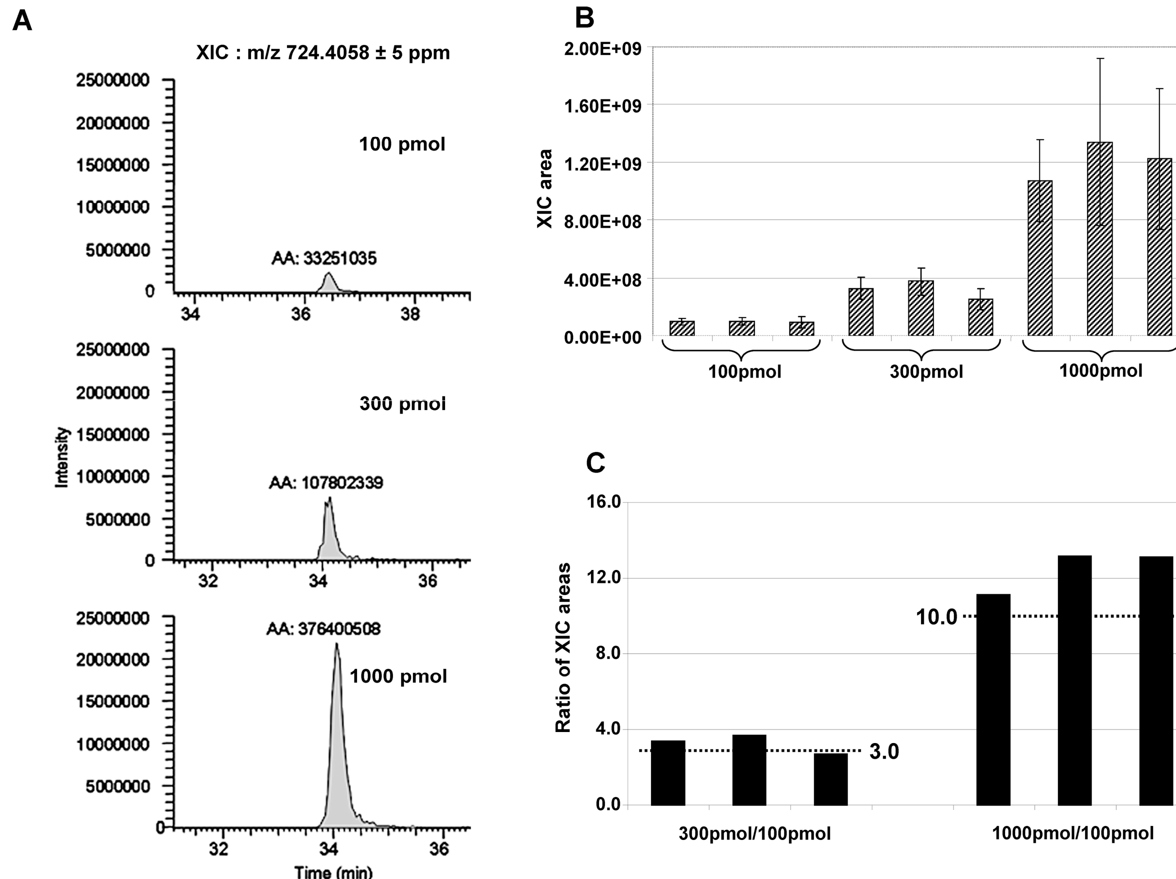
- Identified only after Proteominer treatment
- Low-abundant proteins
- Involved in specific neuronal functions: control of axonal extension, dendrite outgrowth
- May constitute potential biomarkers of neuronal pathologies

Description	Best score	#Peptides	#Total MS/MS
Apolipoprotein E	2268.62	38	3317
Clusterin	1498.54	26	2366
Serpin F1	1470.22	25	880
Angiotensinogen	1401.99	16	877
Azurocidin	863.55	14	361
14-3-3 protein gamma	882.08	19	219
14-3-3 protein eta	530.12	19	106
NRCAM protein	549.21	26	85
Agrin	621.66	28	78
Neural cell adhesion molecule L1-like protein	565.02	27	65
Metalloproteinase inhibitor 2	214.33	10	41
Contactin-2	354.53	21	39
Laminin subunit beta-1	313.14	12	39
Isoform 1 of Neurexin-1-alpha	199.82	10	33
Tenascin-R	157.67	6	26
Protein S100-A6	224.51	3	25
Isoform A22 of Neuropilin-2	122.09	8	25
ubiquitin and ribosomal protein S27a	104.95	6	18
Meteorin	244.61	6	18
Proprotein convertase subtilisin/kexin type 9	289.69	7	16
Semaphorin-3B	242.89	7	13
SLIT and NTRK-like protein 1	57.15	8	12
Neuropilin 1	72.65	2	10
Protein S100-B	165.32	4	7
Myotrophin	79.7	2	6
Delta and Notch-like epidermal growth factor-related receptor	100.76	2	6
Neurofilament medium polypeptide	120.68	5	5
Neurexin-3-beta	35.84	4	4
Spondin-2	73.81	4	4
Palmitoyl-protein thioesterase 1	59.89	1	2
SLIT and NTRK-like protein 4	76.84	1	2
Alpha-soluble NSF attachment protein	64.85	1	1

Effect of Proteominer on quantitative information

Quantification of a spiked protein

Replicate Proteominer treatments on RBC lysate spiked with increasing amount of ADH protein



Roux-Dalvai *et al.*

Mol Cell Proteomics 2008

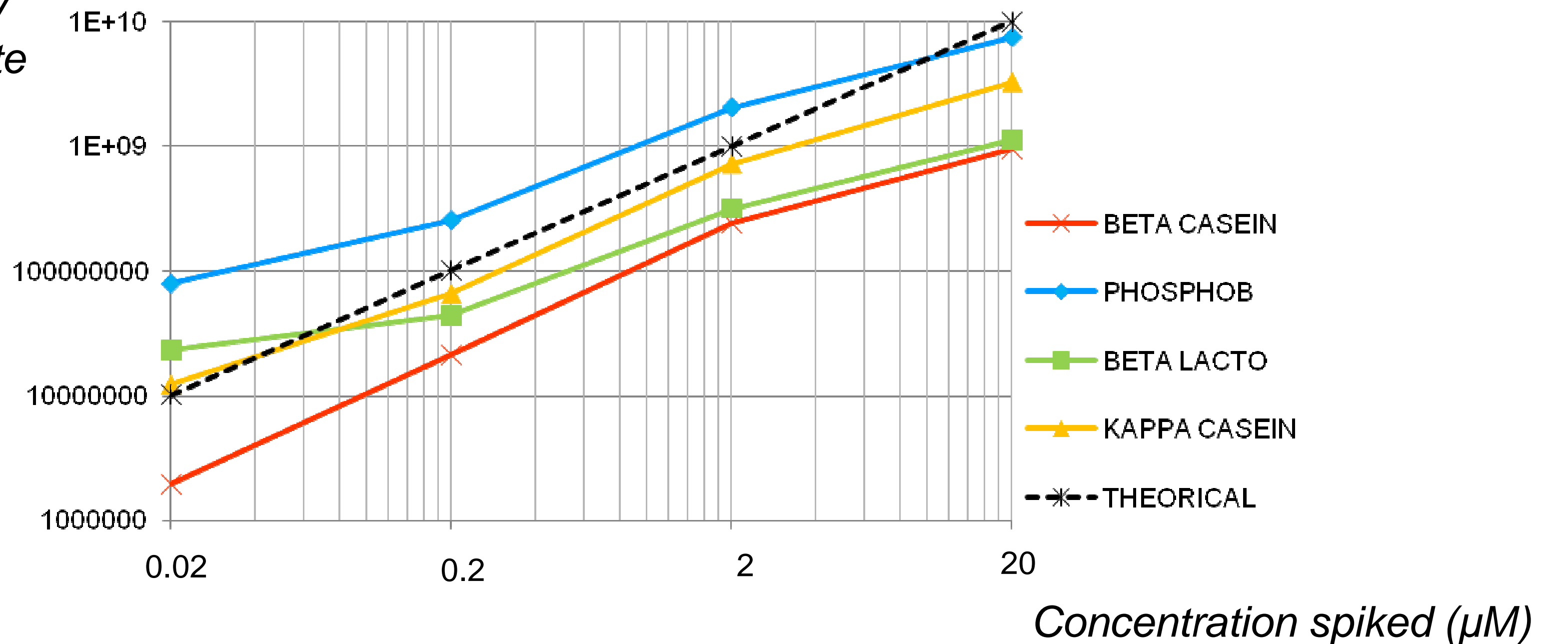
- Good reproducibility between the different equalization replicates
- Ratios are closed to what expected
- Relative protein quantification after equalization is possible if the protein does not saturate the beads

Effect of Proteominer on quantitative information

Quantification of a panel of proteins spiked at high concentrations

If many proteins saturate their recognition bead, it will not be possible to quantify them after Proteominer treatment

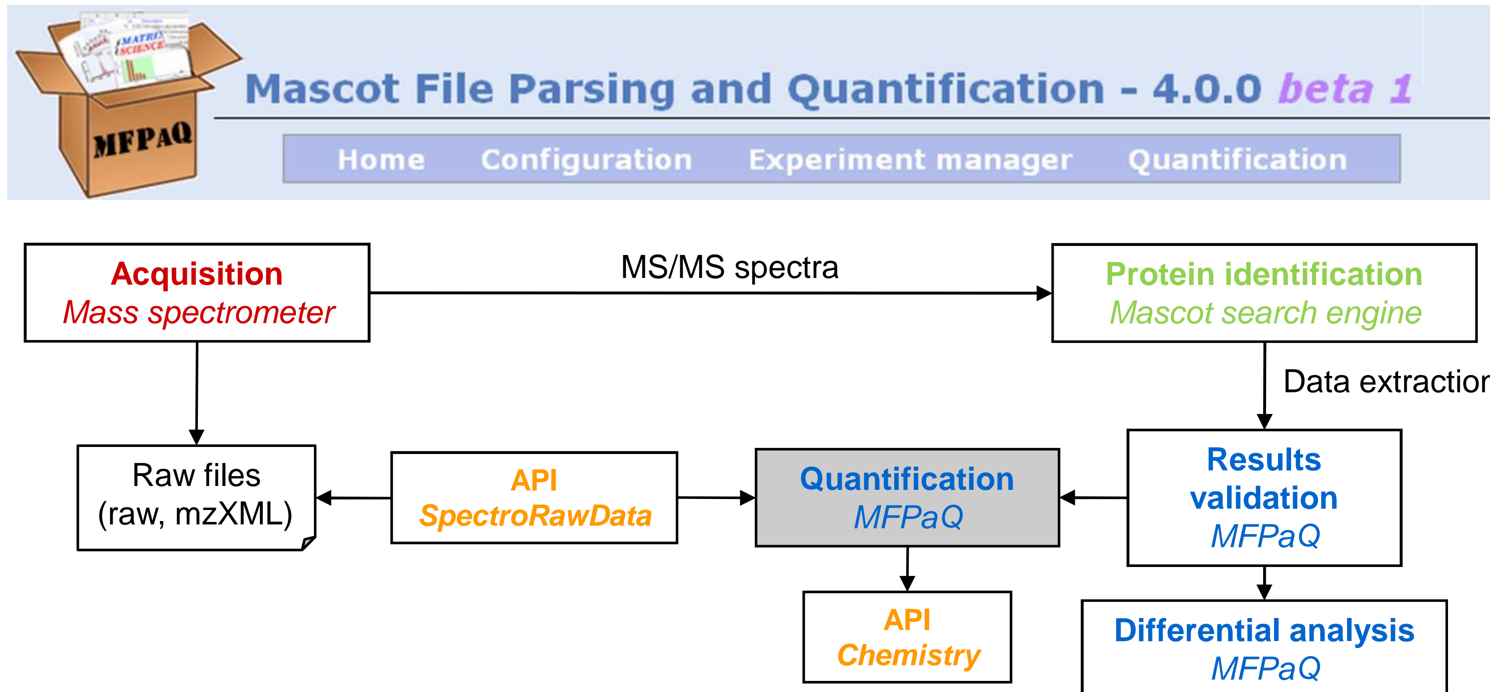
- Test performed on several proteins (myoglobin, beta and kappa casein, beta-lactoglobulin, phosphorylase B, carbonic anhydrase and alcohol dehydrogenase)
- Spiked at very high amounts in serum (up to 20µM), equalization in triplicate



- Relatively good linear response at intermediate concentrations
- Slight saturation effect was observed at very high concentrations

Protein quantification

Development of bioinformatic tools for differential quantitative analysis by nanoLC-MS



Quantification:

- Start from identified and validated proteins

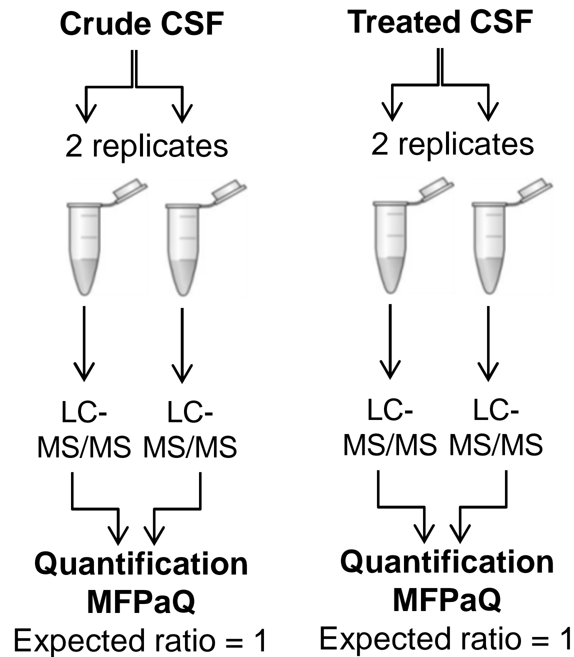
- The *m/z* and retention time (RT) of each identified peptide ion are used to extract XIC signal for each peptide in the MS survey scans

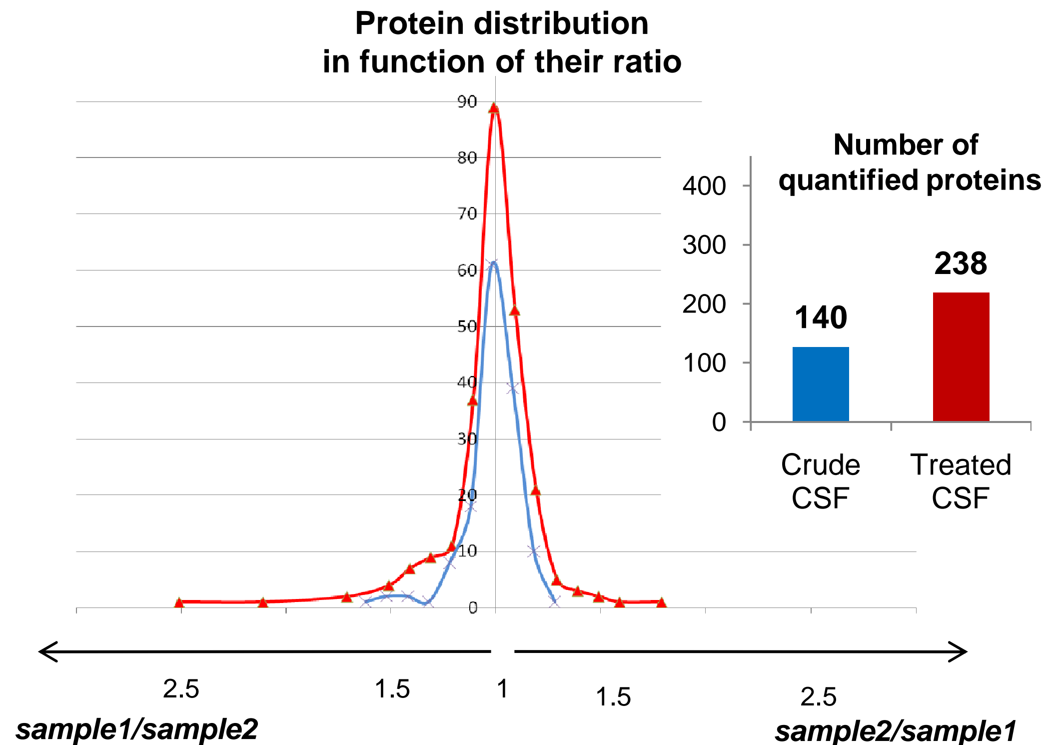
(+) Quick and robust (only confident peptides are extracted)

(-) Only proteins that have been identified by MS/MS in at least one of the samples can be quantified

Quantitative MS analysis

Comparison of two replicate LC-MS runs on CSF samples





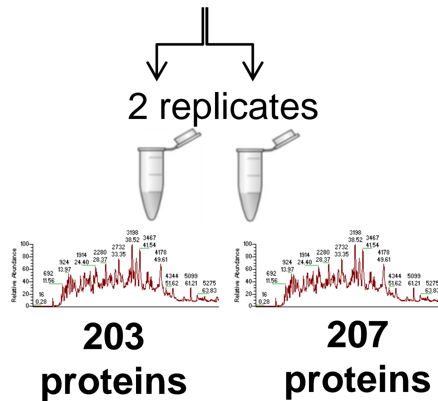
- The label free quantification between 2 replicate samples displays a ratio of 1 for most of the proteins
- More proteins can be quantified in Proteominer treated samples
- Drawback : the analysis of each sample in a single run (no fractionation) decreases the number of identified/quantified proteins.

Quantitative MS analysis

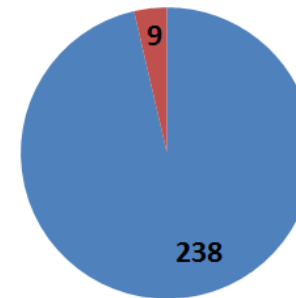
Label-free quantification using an identification database

See poster: *E.Mouton-Barbosa et al.*

Proteominer treated CSF



Direct label free quantification

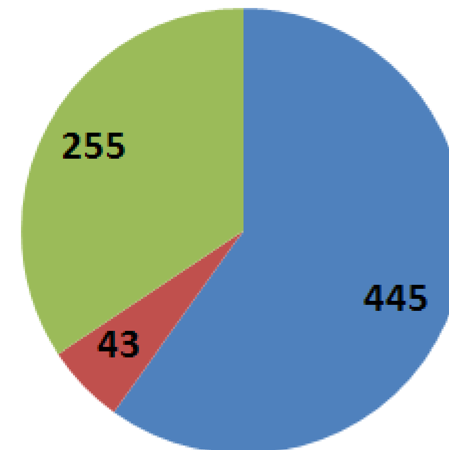


Identification database

Obtained by fractionation of the sample on 1D gel

743 proteins
5632 peptide ions

Label free quantification through an identification database



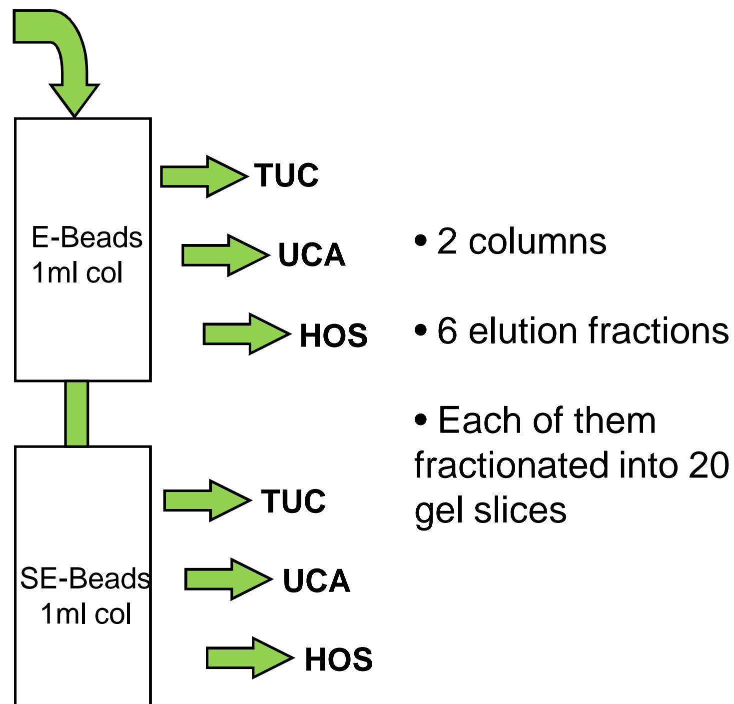
- The use of an identification database increases by a factor 2 the number of quantified proteins
- Best results are obtained using equalized samples vs equalized database

Miniaturization and reproducibility on low CSF volumes

Extensive proteomic analysis of CSF

Pool of different patients ~ 1000 ml

770mg total protein
Bead capacity 10mg } overloading of 77X



- 2 columns
- 6 elution fractions
- Each of them fractionated into 20 gel slices

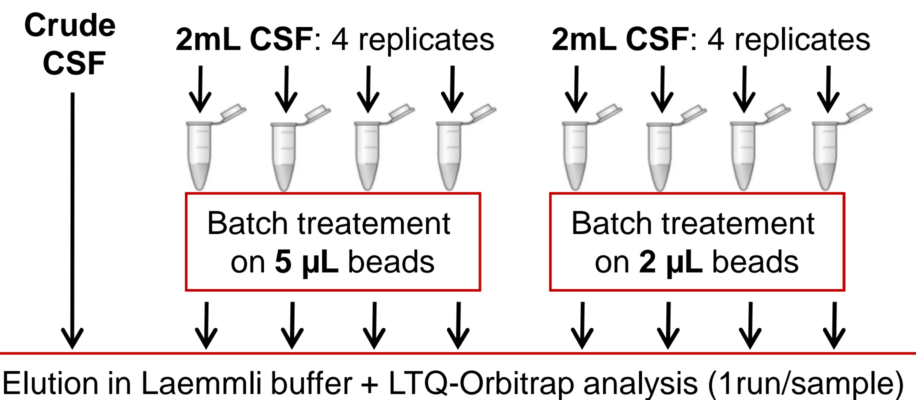
Clinical studies

Total protein concentration in CSF: 0.40 mg/ml (45% Albumin)

2 mL lumbar puncture: < 1mg of protein

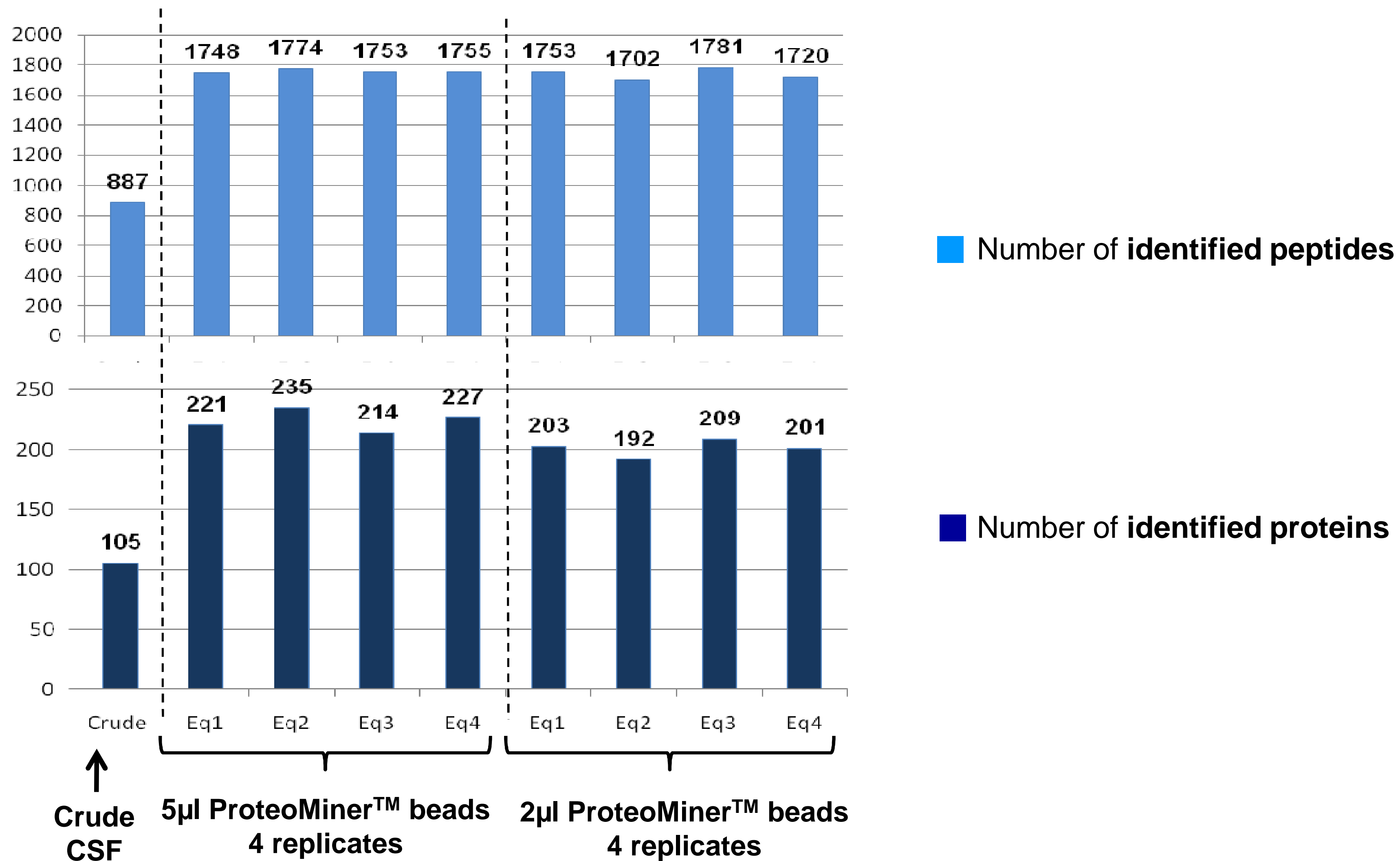
- In order to keep the overloading ratio at 50x, the beads volume must be decreased
- For large series of samples, extensive fractionation of the equalized sample is difficult

Test on a pool of CSF divided into 2mL aliquots



Efficiency of equalization ? Reproducibility ?

Equalization on low CSF volumes – Efficiency?

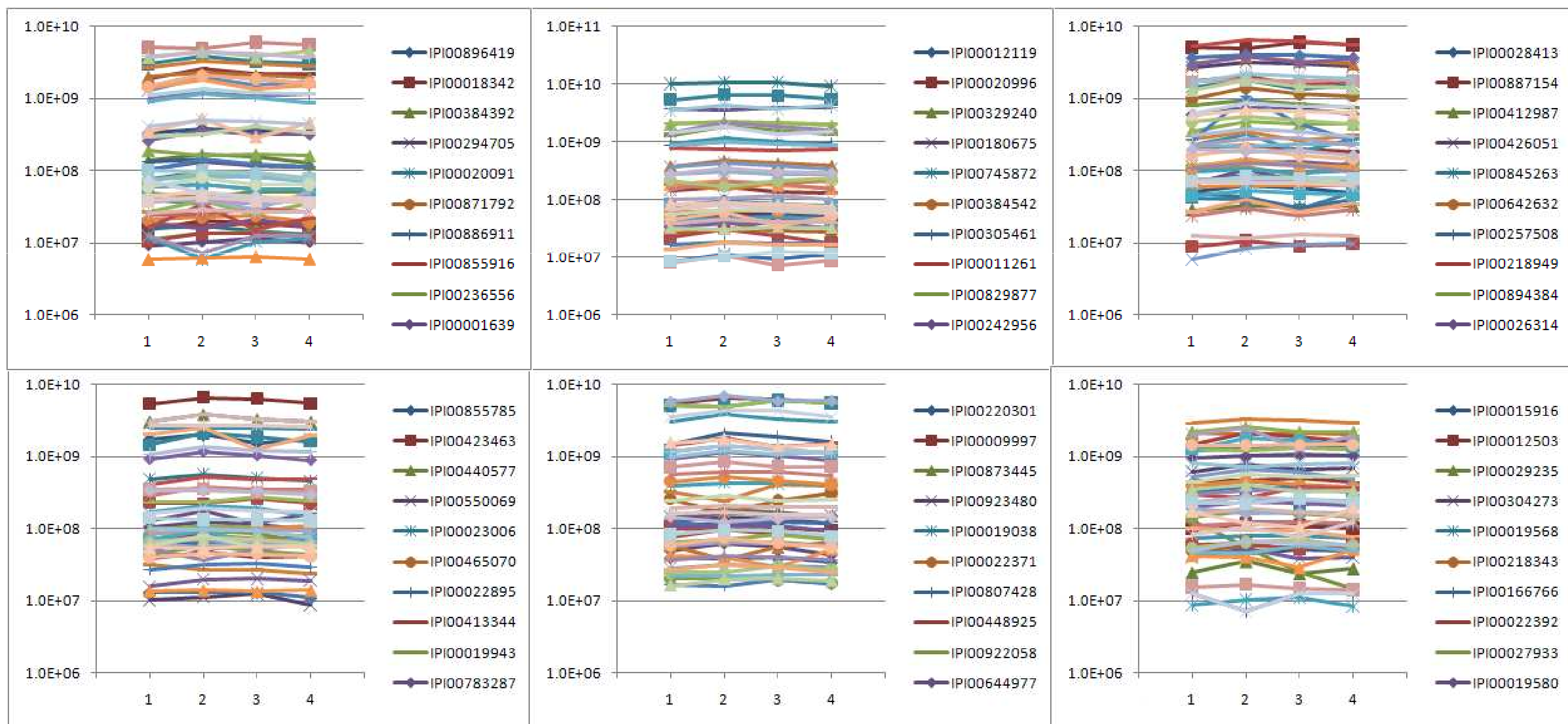


- **Number of proteins identified increases by a factor 2**
- **The treatment is efficient with low beads volumes (5µL better)**

Equalization on low CSF volumes – Reproducibility?

- Plots of the PAI (protein abundance index) for the 406 CSF proteins quantified in the 4 replicates equalization treatments on 5 µL of beads

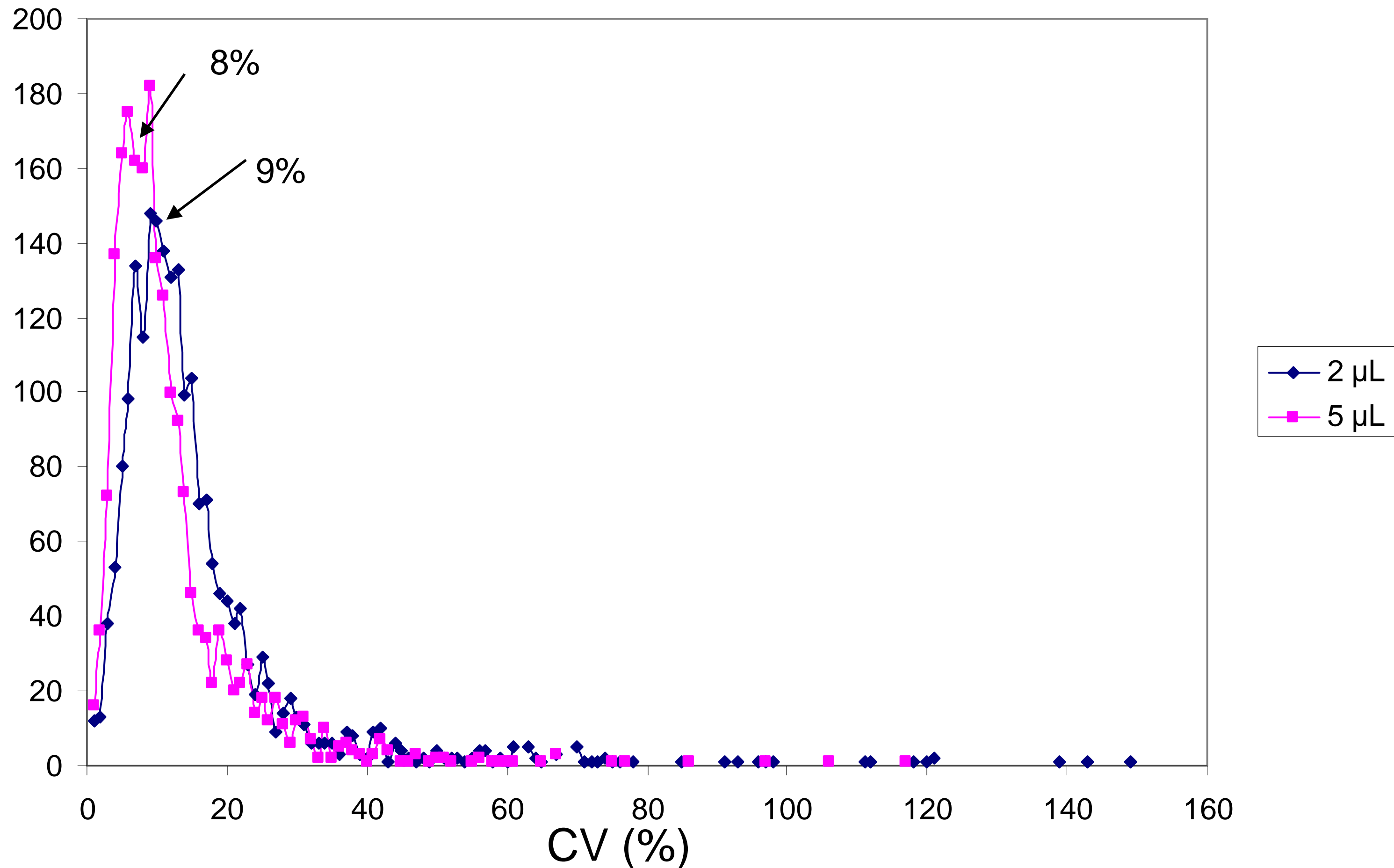
- PAI = log of average MS signal response for the three most intense tryptic peptides of a protein (Silva et al, MCP, 2006)



- **All the proteins were present in all 4 replicates**
- **Protein abundance profiles are constant over the 4 equalization treatments**

Equalization on low CSF volumes - Reproducibility

Distribution of Coefficients of Variation on the peptides intensities (4 replicates)



- Coefficients of variation are slightly lower for equalization on 5µL beads than on 2µL beads.
- Good reproducibility of ProteoMiner™ equalization on small beads volumes

Conclusion and Applications

- The ProteoMiner^{TM} treatment allows to **increase the number of proteins detected in CSF** by nanoLC-MS/MS (X2 proteins identified when the analysis is performed in one run)
- **Quantitative information is conserved** after ProteoMiner^{TM} treatment on a panel of test proteins
- The treatment can be performed on **low volumes of CSF with good reproducibility**
- Perspectives: clinical studies for biomarkers identification of CNS pathologies

Acknowledgements

CNRS-IPBS Toulouse

Emmanuelle Mouton
Florence Roux-Dalvai
David Bouyssié
Odile Schiltz
Bernard Monsarrat

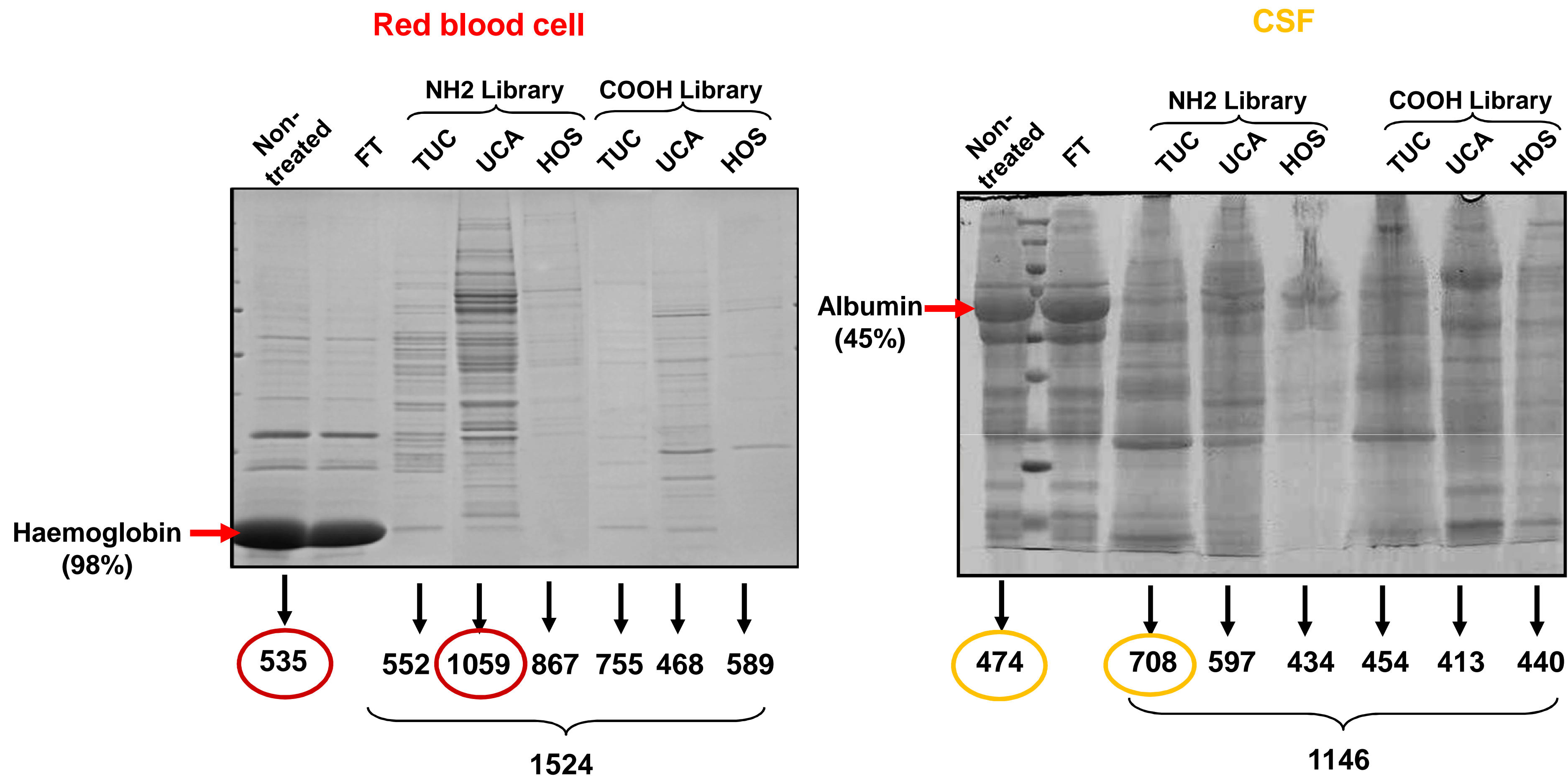
Bio-Rad Gif-sur-Yvette

Egisto Boschetti
Luc Guerrier

INSERM Grenoble
François Berger
CHU Purpan Toulouse
Eric Schmidt

This work is supported by « Institut National du Cancer »

ProteoMiner^{TM} treatment of Red Blood Cell lysate and CSF

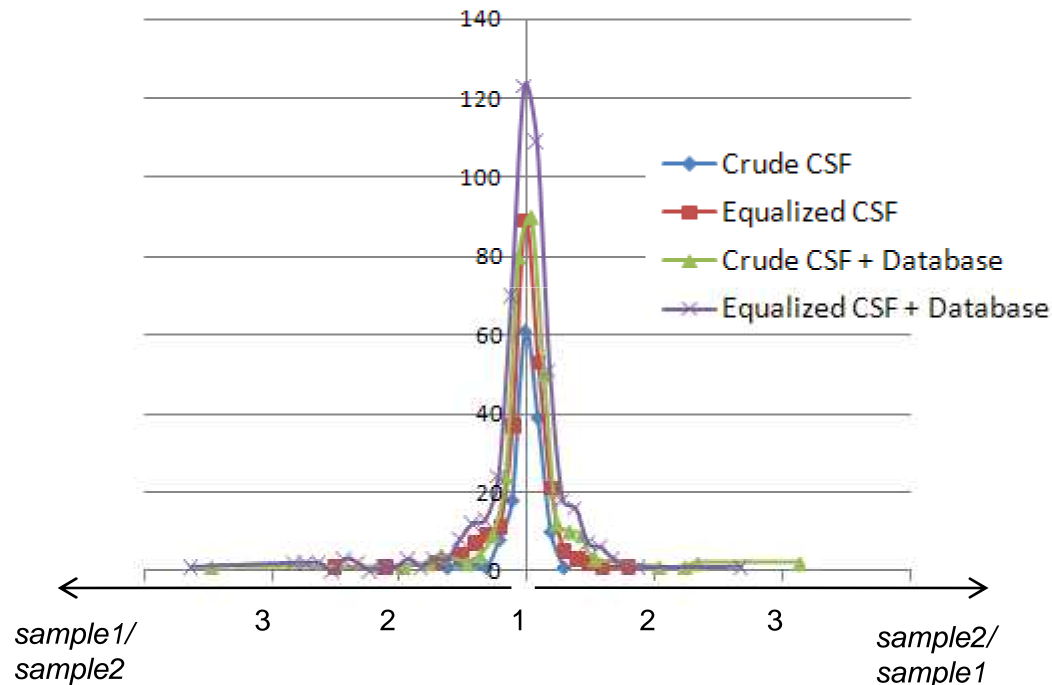


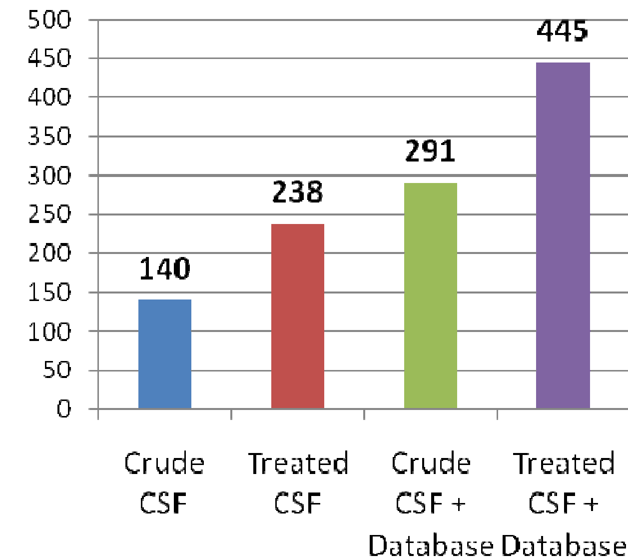
Number of identified proteins
by LTQ-Orbitrap analysis

Quantitative MS analysis

Comparison of two replicate LC-MS runs on CSF samples

- Comparison of 2 replicate sample (treated or not / + or – identification database)
- Each sample is analysed in a single run

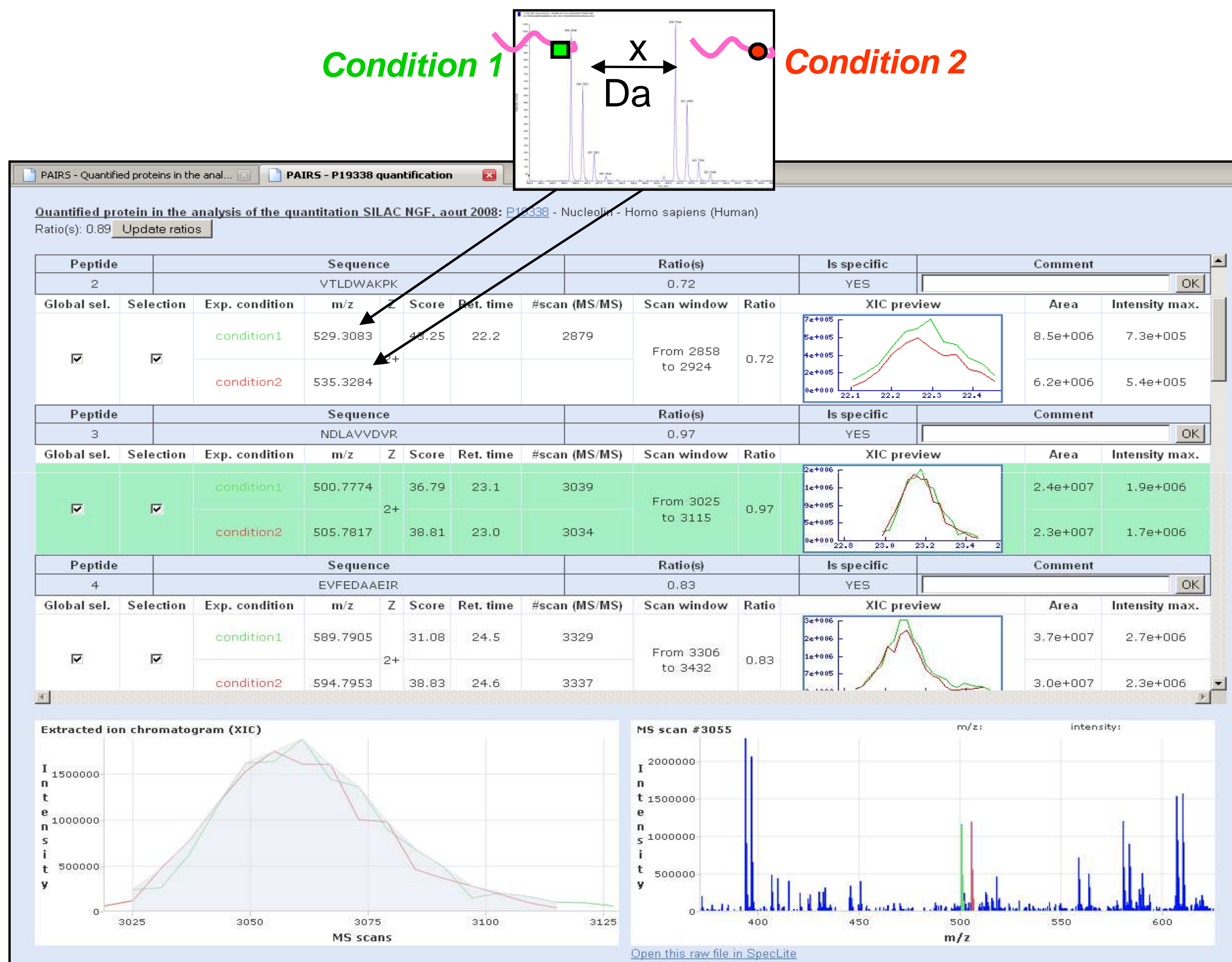




- **The use of an identification database helps to increase the number of quantified proteins**
- **Best results are obtained using equalized samples vs equalized database**

Quantitative MS analysis

MFPaQ software: isotopic labeling

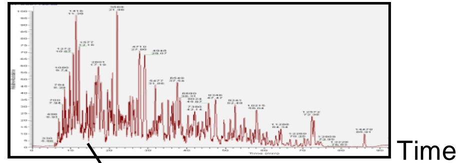


• Extraction of XICs for isotopically labeled peptide pairs in the same run (SILAC, ICAT, $^{14}N/^{15}N$)

Quantitative MS analysis

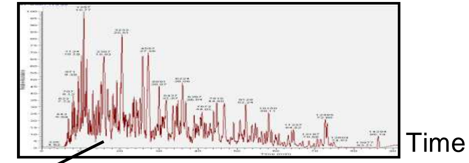
MFPaQ software: label-free

Condition 1



Time

Condition 2



Time

PAIRS - Quantified proteins in the anal... | **PAIRS - IPI00011937 quantificati...**

Quantified protein in the analysis of the quantitation LFssbanque: IPI00011937 - Gene_Symbol=PRDX4 Peroxiredoxin-4

Ratio(s): 0.99 Update ratios

Peptide	Sequence	Ratio(s)	Is specific	Comment
1	SVDETLR	1.00	NO (1)	OK

Global sel.	Selection	Exp. condition	m/z	Z	Score	Ret. time	#scan (MS/MS)	Scan window	Ratio	XIC preview	Area	Intensity max.
☑	☑	condition1	410.2154	2+	31.25	9.7	1056	From to	1.00		3.8e+008	3.2e+007
		condition2	410.2154			9.9					3.8e+008	3.1e+007

Peptide	Sequence	Ratio(s)	Is specific	Comment
2	GLFIIDDK	0.97	NO (1)	OK

Global sel.	Selection	Exp. condition	m/z	Z	Score	Ret. time	#scan (MS/MS)	Scan window	Ratio	XIC preview	Area	Intensity max.
☑	☑	condition1	460.7594	2+		33.4		From to	0.97		7.1e+008	3.2e+007
		condition2	460.7594		32.26	33.8	5818				6.9e+008	3.2e+007

Peptide	Sequence	Ratio(s)	Is specific	Comment
3	LVQAFQYTDK	0.95	YES	OK

Global sel.	Selection	Exp. condition	m/z	Z	Score	Ret. time	#scan (MS/MS)	Scan window	Ratio	XIC preview	Area	Intensity max.
☑	☑	condition1	606.8187	2+	39.22	21.4	3432	From to	0.95		7.5e+007	5.4e+006
		condition2	606.8186		52.37	21.9	3540				7.1e+007	5.1e+006

- Extraction of XICs for the same peptide ion detected in two LC-MS/MS parallel runs