

Developing, transferring, sharing, combining, and bridging  
global and targeted quantitative methods and data  
in a platform-independent manner thanks to Skyline

**Christine Carapito**

Laboratory of Bio-Organic Mass Spectrometry  
CNRS / Strasbourg University  
Director: A. Van Dorsselaer  
ccarapito@unistra.fr

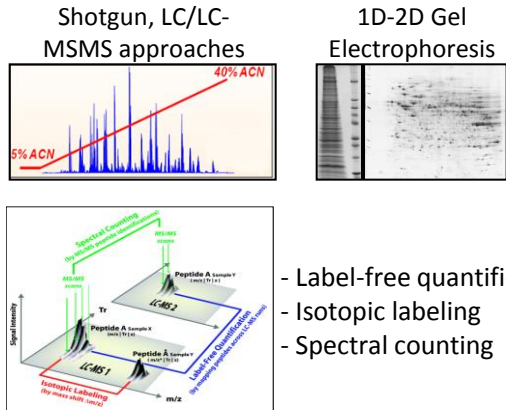
2<sup>nd</sup> Skyline User Group Meeting ASMS 2013  
June 8<sup>th</sup>, 2013



UNIVERSITÉ DE STRASBOURG

# From Global to Targeted Proteomics Approaches

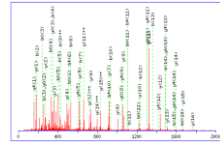
## Global, Discovery Proteomics



- Label-free quantification
- Isotopic labeling
- Spectral counting

From Mueller, L. N., et al., 2008

**LC-MS/MS**

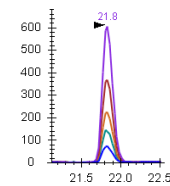


Qualitative	Quantitative
500-2000 identified proteins	Poorly reproducible, approx. quantitation

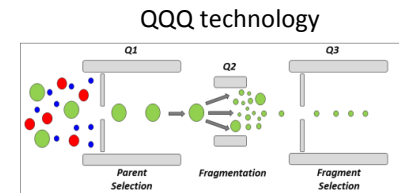
**Proteins of interest**

Qualitative	Quantitative
10-100 candidate proteins	Precise reproducible, absolute quantitation

**LC-SRM**



## Targeted Proteomics



Heavy labeled synthetic standards



# Examples of applications from our lab

## Proteome and Metaproteome Analysis of Arsenic-Resistant Bacteria and Bacterial Communities

Collaboration with Bertin P. and Ploetze F., Strasbourg University

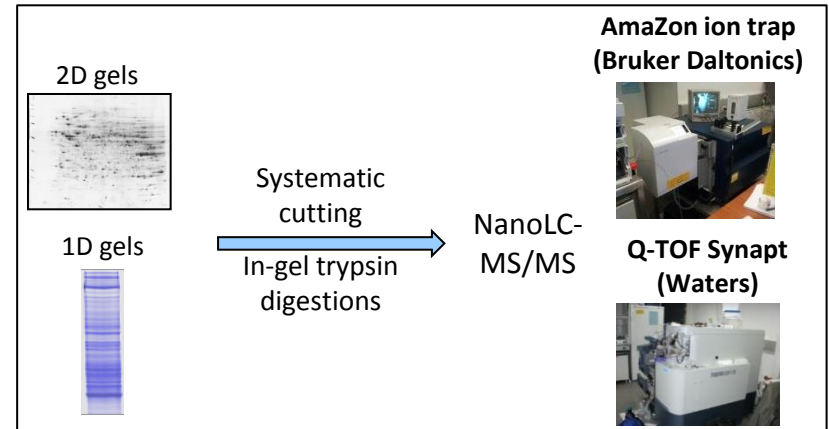
Acid mine drainage (AMD) of the Carnoules mine (south of France) characterized by acid waters containing high concentrations of arsenic and iron.



Sediment analysis:

- Metagenome sequencing of the community
- Metaproteome analysis using the metagenome data

### From Global/Discovery Proteomics :



➔ Identification of ~900 proteins among which interesting candidate proteins involved in arsenic bioremediation

# Examples of applications from our lab

## Proteome and Metaproteome Analysis of Arsenic-Resistant Bacteria and Bacterial Communities

Collaboration with Bertin P. and Ploetze F., Strasbourg University

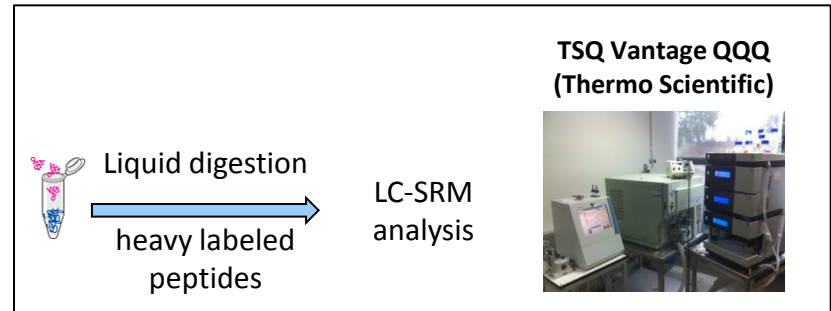
Acid mine drainage (AMD) of the Carnoules mine (south of France) characterized by acid waters containing high concentrations of arsenic and iron.



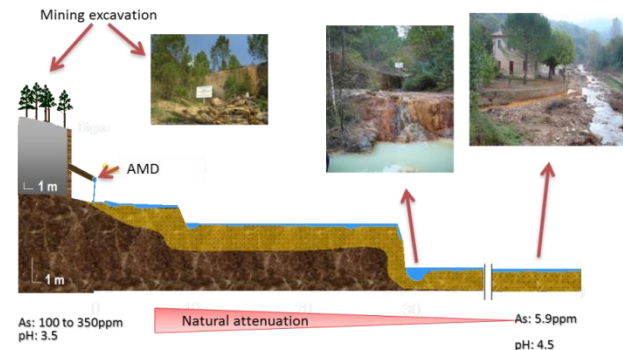
Sediment analysis:

- Metagenome sequencing of the community
- Metaproteome analysis using the metagenome data

### To Targeted Proteomics :



LC-SRM assay for accurate quantification of targeted proteins in sediments over the watercourse and seasons.



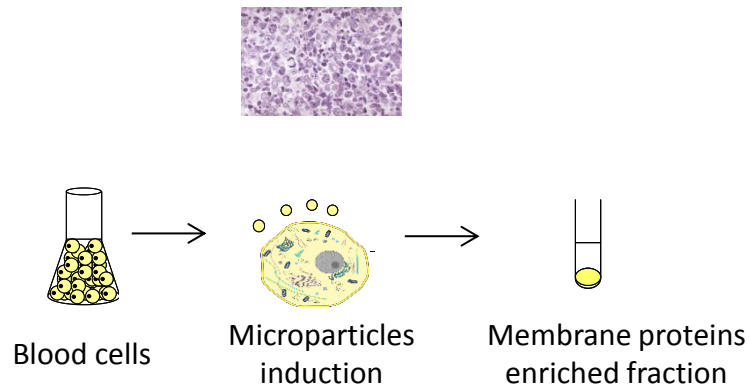
- Carapito C., *et al.* (2006) *Biochimie* 88: 595-606  
Muller D., *et al.* (2007) *PLoS Genet* 3: e53  
Weiss S., *et al.* (2009) *Biochimie* 91: 192-203  
Bruneel O., *et al.* (2011) *Microb Ecol* 61: 793-810  
Bertin P.N., *et al.* (2011) *ISME J.* 5:1735-1747  
Halter D., *et al.* (2011) *Res Microbiol* 162: 877-887  
Halter D., *et al.* (2012) *ISME J.* 6: 1391-1402

# Examples of applications from our lab

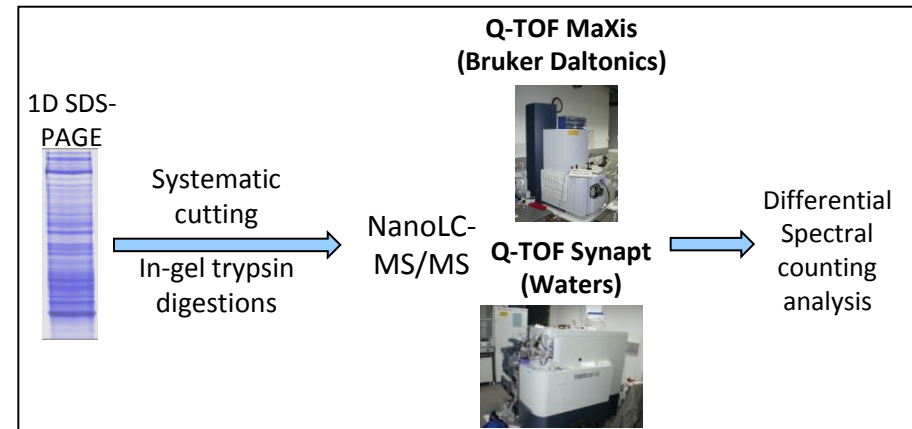
## B-cells lymphoma biomarker discovery

Sarah Lennon, Christine Carapito, Laurent Miguet, Luc Fornecker, Laurent Mauvieux, Alain Van Dorsselaer, Sarah Cianferani  
Collaboration with Institute of Hematology and Immunology, Strasbourg University

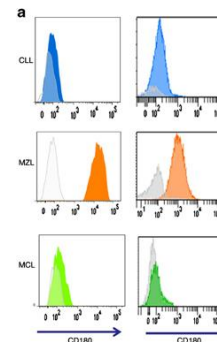
B-cell Lymphoma: Blood disease characterized by a proliferation of B lymphocytes



### From Global/Discovery Proteomics :



➔ Identification of 2 robust candidate biomarkers: CD148 and CD180



Validated by flow cytometry  
(on 1 epitope) on > 500 samples

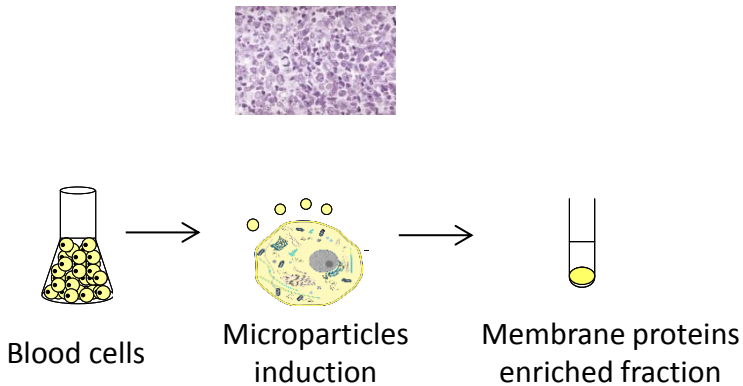
- Miguet L. *et al.*, (2006) *Proteomics* 6: 153-171
- Miguet L. *et al.*, (2007) *Subcell Biochem* 43: 21-34
- Miguet L. *et al.*, (2009) *J Proteome Res* 8: 3346-3354
- Miguet L. *et al.*, (2013) *Leukemia* Epub ahead of print

# Examples of applications from our lab

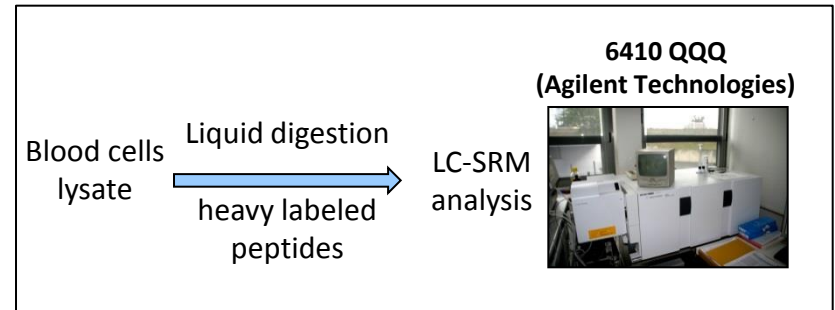
## B-cells lymphoma biomarker discovery

Sarah Lennon, Christine Carapito, Laurent Miguet, Luc Fornecker, Laurent Mauvieux, Alain Van Dorsselaer, Sarah Cianferani  
Collaboration with Institute of Hematology and Immunology, Strasbourg University

B-cell Lymphoma: Blood disease characterized by a proliferation of B lymphocytes



## To Targeted Proteomics :



LC-SRM assay for absolute quantification of targeted proteins, following at least 10 peptides per protein (versus 1 epitope)

```
1 MKPAA REARL PPRSP GLRWA LPLLL LLLRL GQILC AGGTP SPIPD PSVAT
51 VATGE NGITQ ISSTA ESHK QNGTG TPQVE TMTSE DGESS GANDS LRTEP
181 QGSHG TGGAS QKTPS STGPP PPFDL KAVSI SPTMV LLTHK SNO7A ASEYK
151 YVWKH KMENE KTTIV VHQPW CNITG LAPAT SYVFS ITPGE GNETW GDRPV
281 IKVIT EPIPV SDLR V ALTGV RKAAL SWNSG NGTAS CRVLL ESIGS HEELT
251 QDSRL QWITS GLKPG VQYNI NPYLL QSNKT KGDPL GTEGG LDASN TERSR
381 AGSPT APVHD ESLVG PVQPS SQQSS RDTFV LLVEL EPGRV YNATV YSQAA
351 NGTEG QPDAI EFRTH AIQVF DVTAW NISRI SLLIL WKWSD NESSS NYTFK
481 IHVAG ETQSS NLNVS EPRAV IPGLR SSTFV NITVC PVLGD IEGTF GFLOV
451 HTPPV PVSDF RVTWV STTEI GLAWS SHDAE SFQMH IQTEG AGNSR VEITT
581 NQSI I IGLF PGTKY CFEIV PKGNP GTEGA SRTVC NRTVP SAVED IHVVY
551 VTTTE MWLDW KSPDG ASEYV YHLVZ ESKHG SNTS TYDKA ITLQG LIPGT
681 LYNIT ISPEV DHWAG DPKST AQYTR PSNVS NIDVS TINTA ATLSW QNFDD
651 ASPTY SVCLL LEKAG KSSNA TQWTF DIGIT DMTVF ELIPE SSYTV EIFAQ
781 VGDI KSLEP GRKSF CTDPA SWASF DCEVW PKEPA LVLKW TCPPG ANAGF
751 ELEVS SGAWN NATHL ESCSS ENGTE YRTEV TYLNF STSYN ISITT VSQKG
881 MAAPT RMTCT TGITD PPPPD GSPNI TSVSH NSKVV KFSGF EASHG PIKAY
851 AVLLT TGEAG HPSAD VLKYT YDDPK KGASD TYVTV LIRTE EKGRS QELSE
981 VLYEY IOKNI ESTTL GYYNG KLEPL GSYRA CVAGF TMTTF HPQNK GLIIG
951 AESYV SESRY SDAVS LPQDP GVICG AVFGC IFGAL VEIVT VGGFV WRKKR
1881 KDAKN NEVSF SQTKP KSKL IRVEN FEAYF KRQQA DSMCG FAEEY EDLKL
1851 VGISQ PKVAA ELAEN RGNKR YNNVL PYDIS RVKLS VQTHS TDDYI NANYM
1181 PGVHS KKDFI ATQGP LPNLT KDPNR MWEXK NYVAI IMLTK CVEQG RTKCE
1151 EYVPS KQKQD YGDDT VAMTS EIVLP EKIFR DFTMK NIXTS ESMLP QSFFH
1281 TSWPD HGVPD TTDLL INFRY LVKDY MKQSP PESPI LVNVS AGVGR TGTFI
1251 AIDLK IVQIE NENTV DWYGI VYDLR MHRPL MVQTE DQYVF LNQCQ LDIVR
1381 SQKDS KVLLI YQNTT AMTIV ENLAP VITFG KTMGY IA
```

Sequence coverage of CD148 (Q12913)

- Miguet L. *et al.*, (2006) *Proteomics* 6: 153-171
- Miguet L. *et al.*, (2007) *Subcell Biochem* 43: 21-34
- Miguet L. *et al.*, (2009) *J Proteome Res* 8: 3346-3354
- Miguet L. *et al.*, (2013) *Leukemia* Epub ahead of print

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest



2. Proteotypic peptides for proteins of interest



3. Transitions selection and optimisation



4. SRM analysis



5. Quantitative data interpretation



# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

Previous global/discovery proteomics experiments

+

Additional hypotheses, Biological observations or literature/data mining, ...

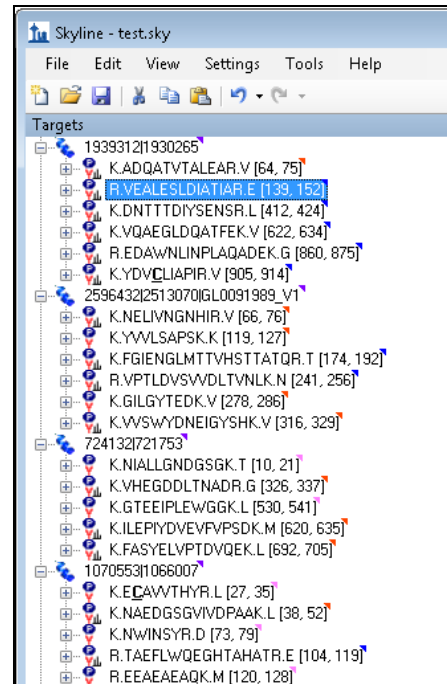
2. Proteotypic peptides for proteins of interest

3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Upload of targeted proteins  
(.fasta file)





# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

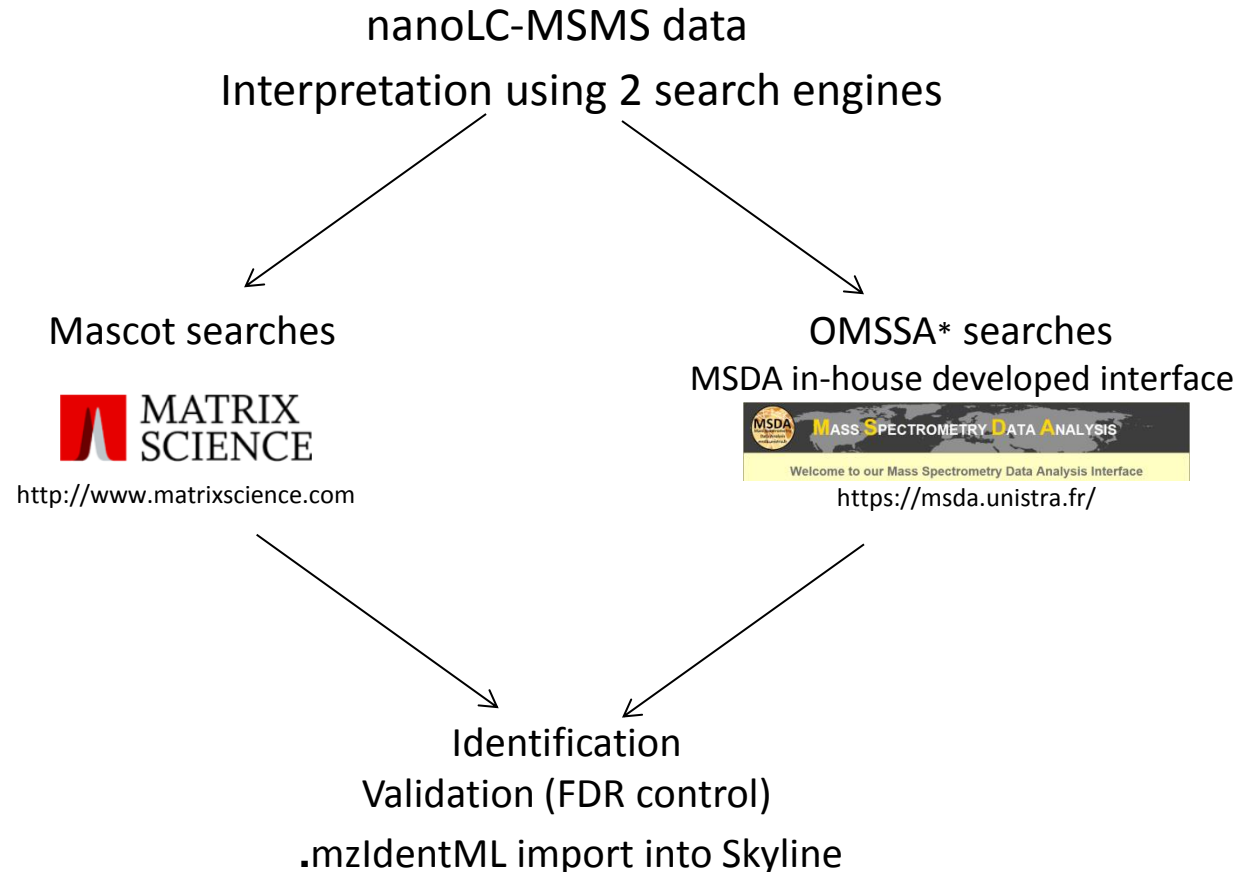
3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to identify **best flyers** and unique peptides :

1. Building of Peptide Spectral Libraries generated from global proteomics data



# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

3. Transitions selection and optimisation

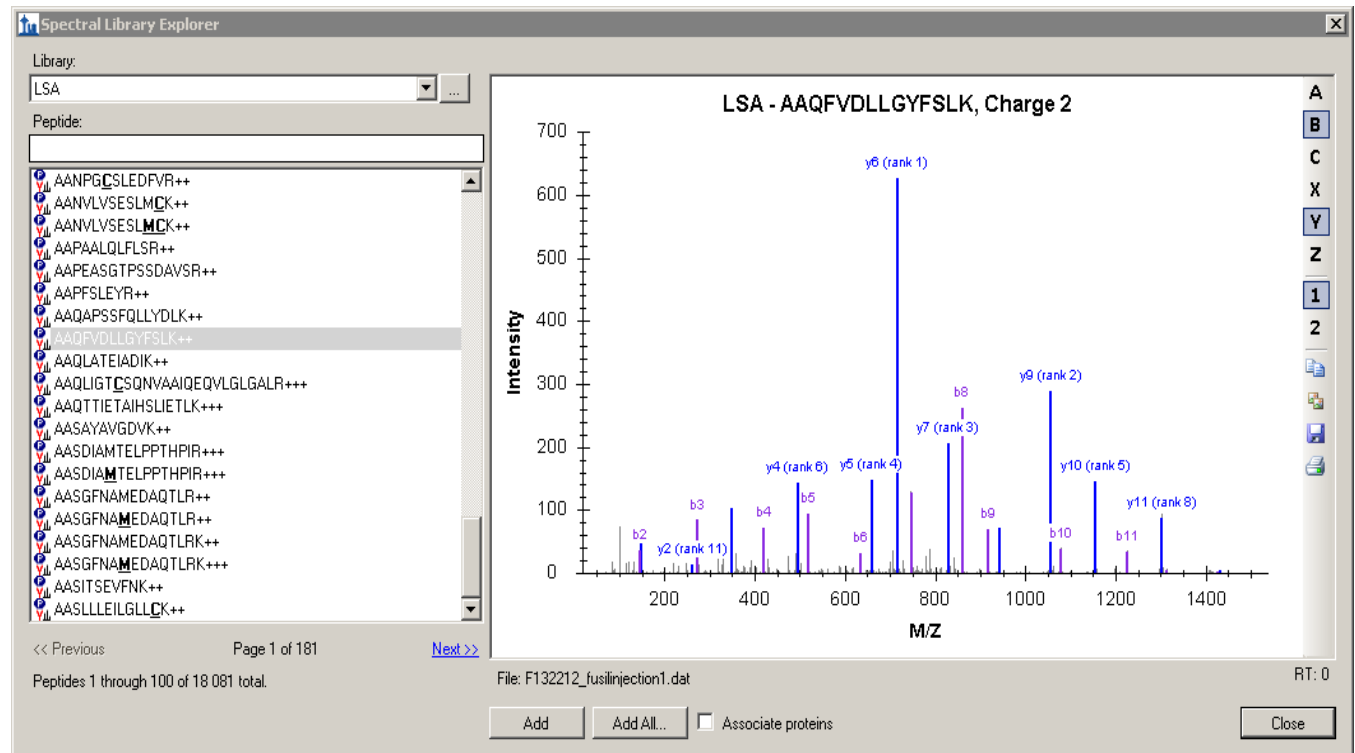
4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to identify **best flyers** and unique peptides :

1. Building of Peptide Spectral Libraries generated from global proteomics data

## Spectral Library Explorer



# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

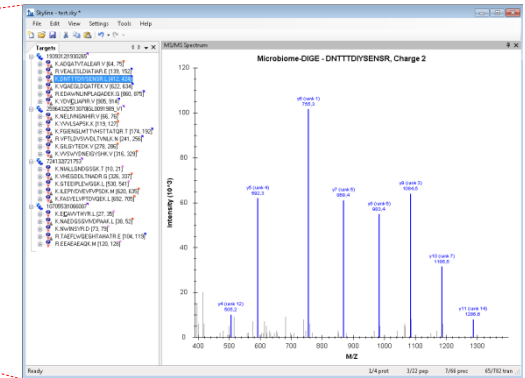
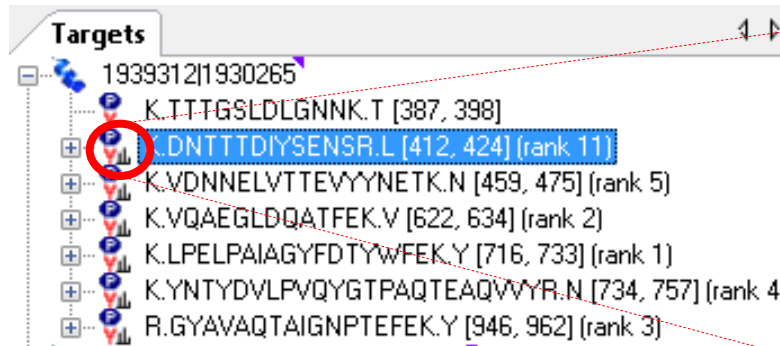
3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to identify **best flyers** and unique peptides :

1. Building of Peptide Spectral Libraries generated from global proteomics data



- Among all possible peptides of the proteins of interest, several have already been seen in global proteomics experiments and are likely the best candidates
- Ranking of peptides added (Expect values, picked intensity, spectrum count)

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to identify best flyers and **unique peptides** :

1. Building of Peptide Spectral Libraries generated from global proteomics data
2. Defining a Background proteome

Upload a background proteome as a database .fasta file

Included	Peptide	sp Q43423 AN32C_	sp Q95626 AN32D_	sp Q92688 AN32B_	sp Q92688-2 AN32I_
<input checked="" type="checkbox"/>	ELVLDNSR	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	EPEDEGEDDD	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	LLPQLTYLDGYDR	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	VSGGLEVLAEK	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	DLSTIEPLK	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Protein Name: sp|P39687|AN32A\_HUMAN  
Description: Acidic leucine-rich nuclear phosphoprotein 32 family member A OS=Homo sapiens GN=ANP32A PE=1 SV=1  
Sequence: MEMGRRHLELRNRTPSDVV**ELVLDNSR**SNEGKLEGLTDFEELFLSTINVGLTSIANLPKLNKLLKLELSDNRVSGGLEVLAEKCPNLTHLNLGNIKIKDLSTIEPLKLENLKSLLDFNCEVTNLNDYRENVFKLLPQLTYLDGYDRDDKEAPDSDAEGYVEGLDDEEEDDEEEYDEDQVVEDEEEDDEEEEGEEEDVSGEEEDDEEGYNDGEVDDDEEELGEEERGGKRRKEPEDEGEDDD

R.VLDAGDPTSR.G [328, 337]
R.NSHLSAYPALEGALHR.S [458, 473]
sp P06734 FCER2_HUMAN
sp P08575 PTPRC_HUMAN
sp P02768
<b>sp P02768 ALBU_HUMAN</b> Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2
<b>sp P02768-2 ALBU_HUMAN</b> Isoform 2 of Serum albumin OS=Homo sapiens GN=ALB

- Allows to easily visualise **unique** / **shared** peptides (much faster than performing BLAST alignments)
- Especially important for discriminating isoforms that are present/added in the background proteome

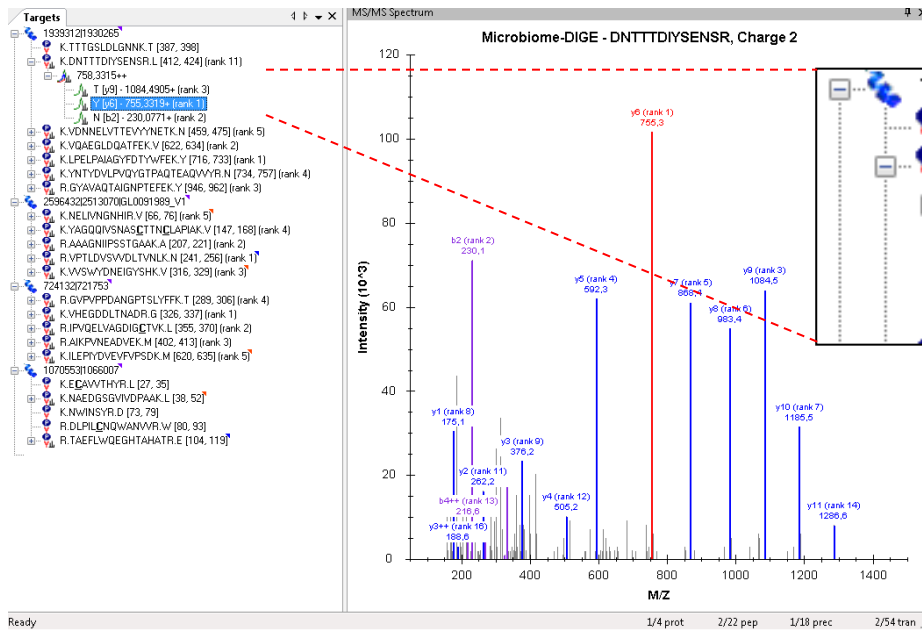
# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

Useful functionalities to select **the best** (specific (no interferences) and sensitive) **transitions** / peptides :

1. Again Peptide Spectral Libraries

2. Proteotypic peptides for proteins of interest



3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

- Spectral libraries built on LC-MSMS data acquired on heavy labeled synthetic standard peptides (for yet unseen peptides)
- Transition ranking + many adjustable filters

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

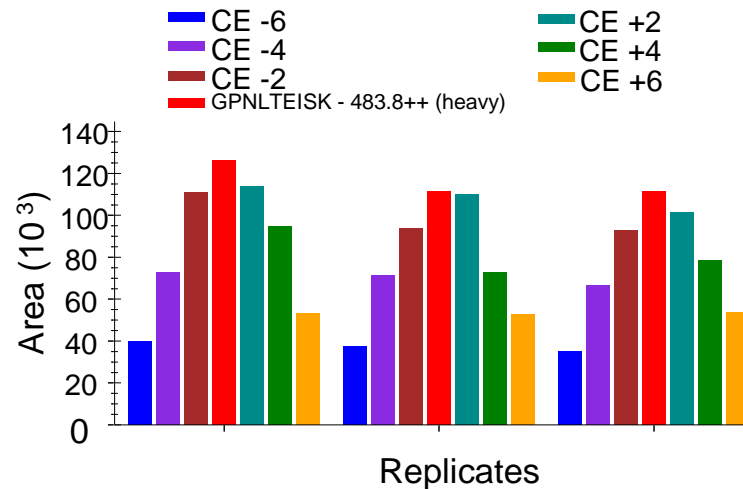
3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to select **the best** (specific (no interferences) and sensitive) **transitions** / peptides :

1. Again Peptide Spectral Libraries
2. Collision energy optimisation



Easily possible thanks to :

- Automatic collision energy optimisation methods setup with different CE steps
- Availability of heavy labeled standard peptides

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

3. Transitions selection and optimisation

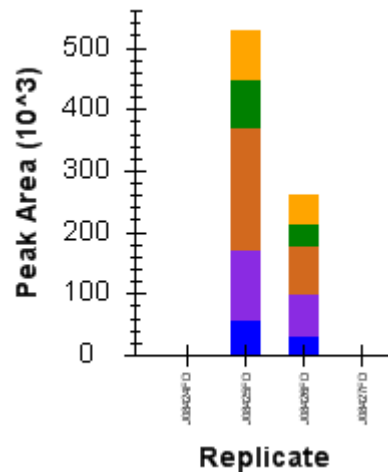
4. SRM analysis

5. Quantitative data interpretation

Useful functionalities to select **the best** (specific (no interferences) and sensitive) **transitions** / peptides :

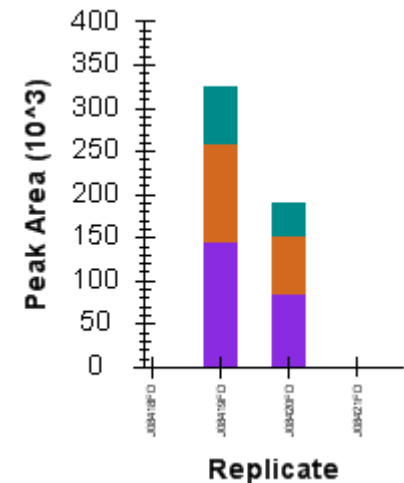
1. Again Peptide Spectral Libraries
2. Collision energy optimisation

VVSQYHELVVQAR



After optimisation / Equation prediction

LVLEVAQHLGESTVR



After optimisation / Equation prediction

➔ Increased sensitivity for specific peptides

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

3. Transitions selection and optimisation

4. SRM analysis

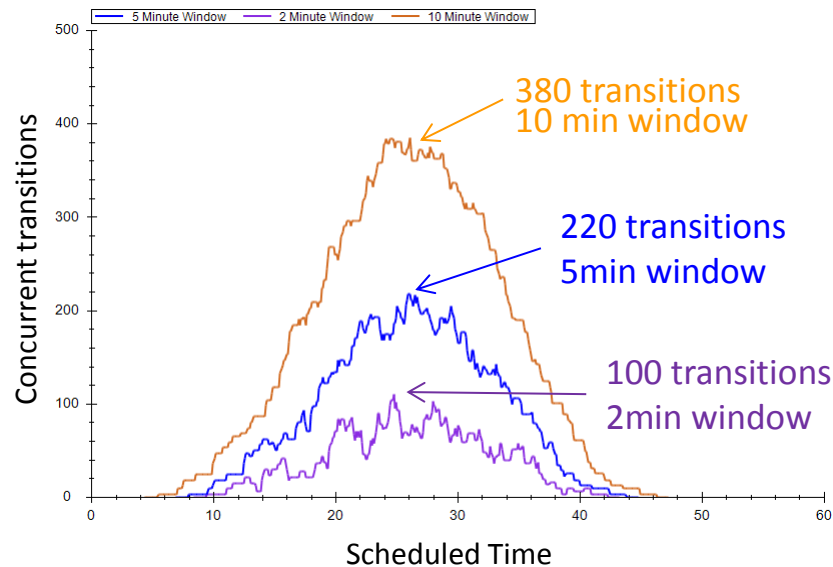
5. Quantitative data interpretation

Useful functionalities to setup up the acquisition methods:

1. Vendor specific method export from a generic Skyline file
2. Retention time scheduling et retention time prediction tools

**Time scheduling is challenging but mandatory for multiplexing!**

- Requires precisely controlled chromatography
- Retention times need to be highly reproducibility
- Peak width and retention time shifts limit the multiplexing.



Use of Retention Time reference (iRT) peptides, spiked in all samples

Escher C, Reiter L, MacLean B, Ossola R, Herzog F, Chilton J, MacCoss M.J, Rinner O  
Proteomics 2012, 12(8): 1111-1121.



# Targeted quantitative proteomics workflow using SRM-MS

## Retention time prediction

1. List of proteins of interest



2. Proteotypic peptides for proteins of interest



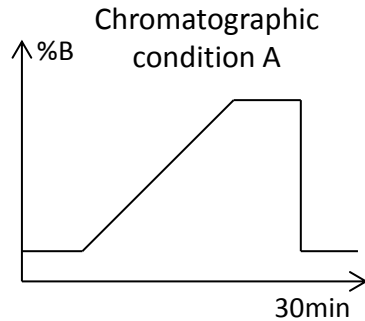
3. Transitions selection and optimisation



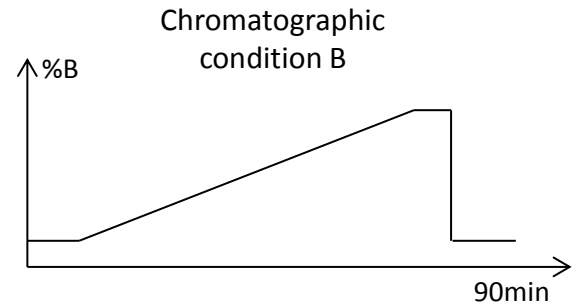
4. SRM analysis



5. Quantitative data interpretation



Gradient change,  
Column change,  
System change, ...



# Targeted quantitative proteomics workflow using SRM-MS

## Retention time prediction

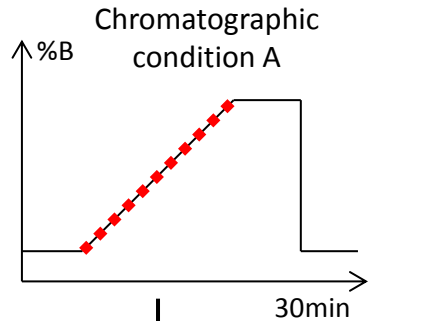
1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

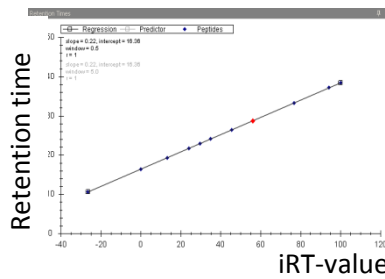
3. Transitions selection and optimisation

4. SRM analysis

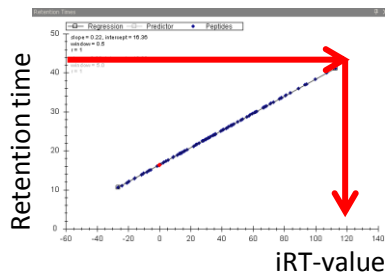
5. Quantitative data interpretation



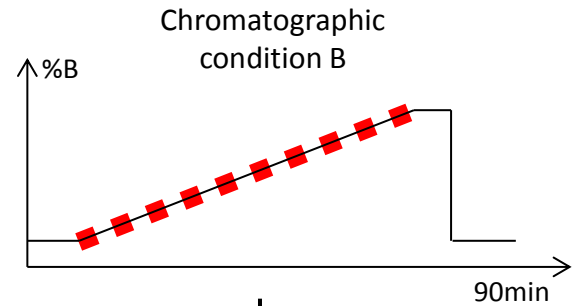
iRT measured in condition A



Calculator

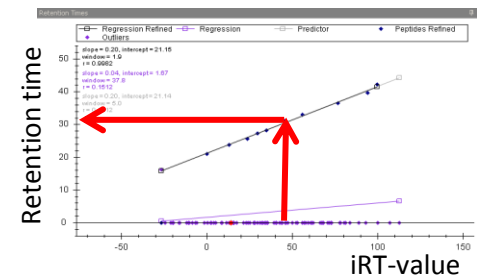


Determination of iRT values for the peptides of interest



iRT measured in condition B

Predictor



Export of scheduled SRM method

# Targeted quantitative proteomics workflow using SRM-MS

## Retention time prediction

1. List of proteins of interest



2. Proteotypic peptides for proteins of interest



3. Transitions selection and optimisation



4. SRM analysis



5. Quantitative data interpretation

- Gain of time for determining peptides' retention times
- Less sample consumption
- Easy change in chromatography type and scale (nanoLC ↔ microLC ↔ LC)
- Easy method transfer inside the laboratory and with collaborating laboratories

# Targeted quantitative proteomics workflow using SRM-MS

1. List of proteins of interest

2. Proteotypic peptides for proteins of interest

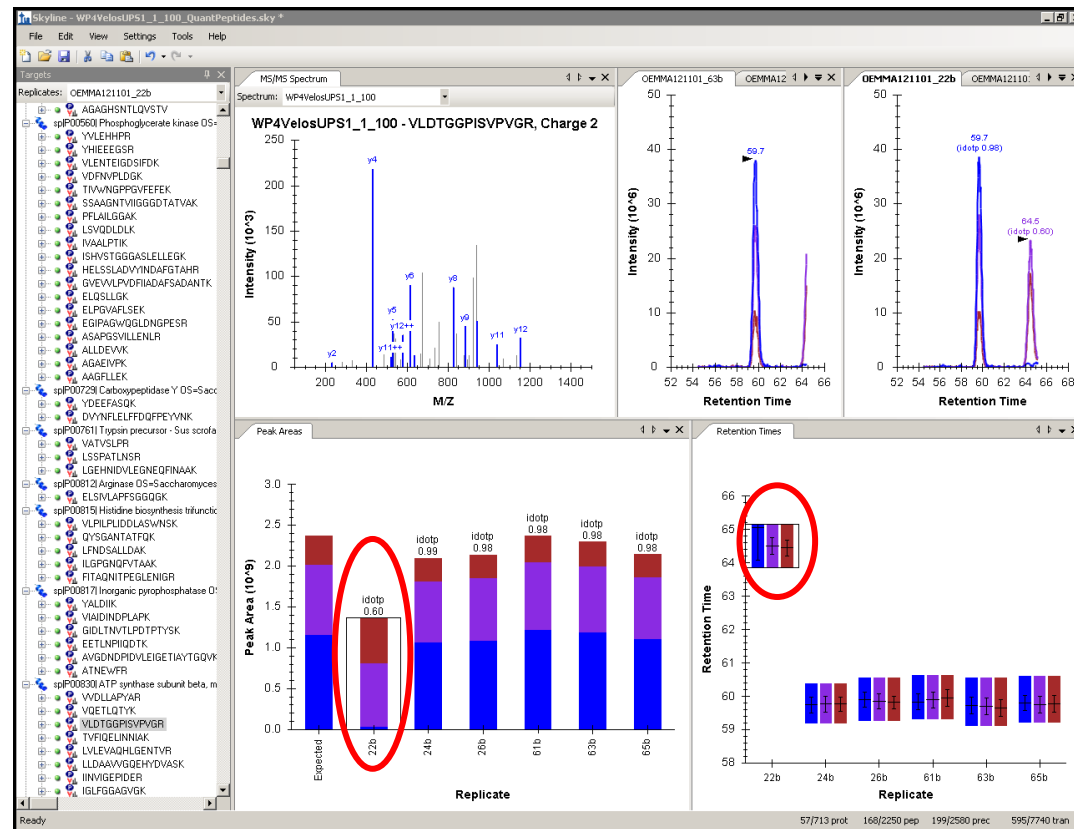
3. Transitions selection and optimisation

4. SRM analysis

5. Quantitative data interpretation

Useful functionalities for quantitative data interpretation:

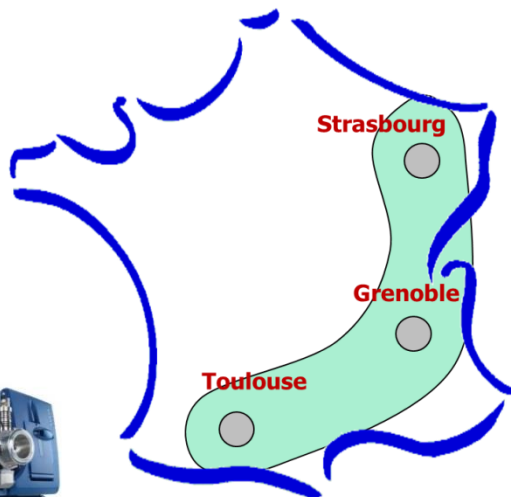
- All Skyline views
- Easy data checking: manual verification is possible, in a fast and efficient way
- View of all replicates
- Visualisation of interferences
- Flexible and rich export templates



# An **inter-laboratory** performance evaluation standard

48 human proteins (Universal Proteomics Standard UPS1)  
spiked into a yeast cell lysate background +  
iRT reference peptides

Weekly injections over 6 months:



G6410  
(Agilent  
Technologies)



TSQ Vantage  
(Thermo)



Q-Trap (ABSciex)



Q-Trap (ABSciex)

Data processing/exchange with Skyline!  
Definition of a series of criteria to meet  
for System OK/Not OK:

- Signal intensities (Peak areas)
- Peak widths
- Retention time
- Peak distribution

Allows us to check:

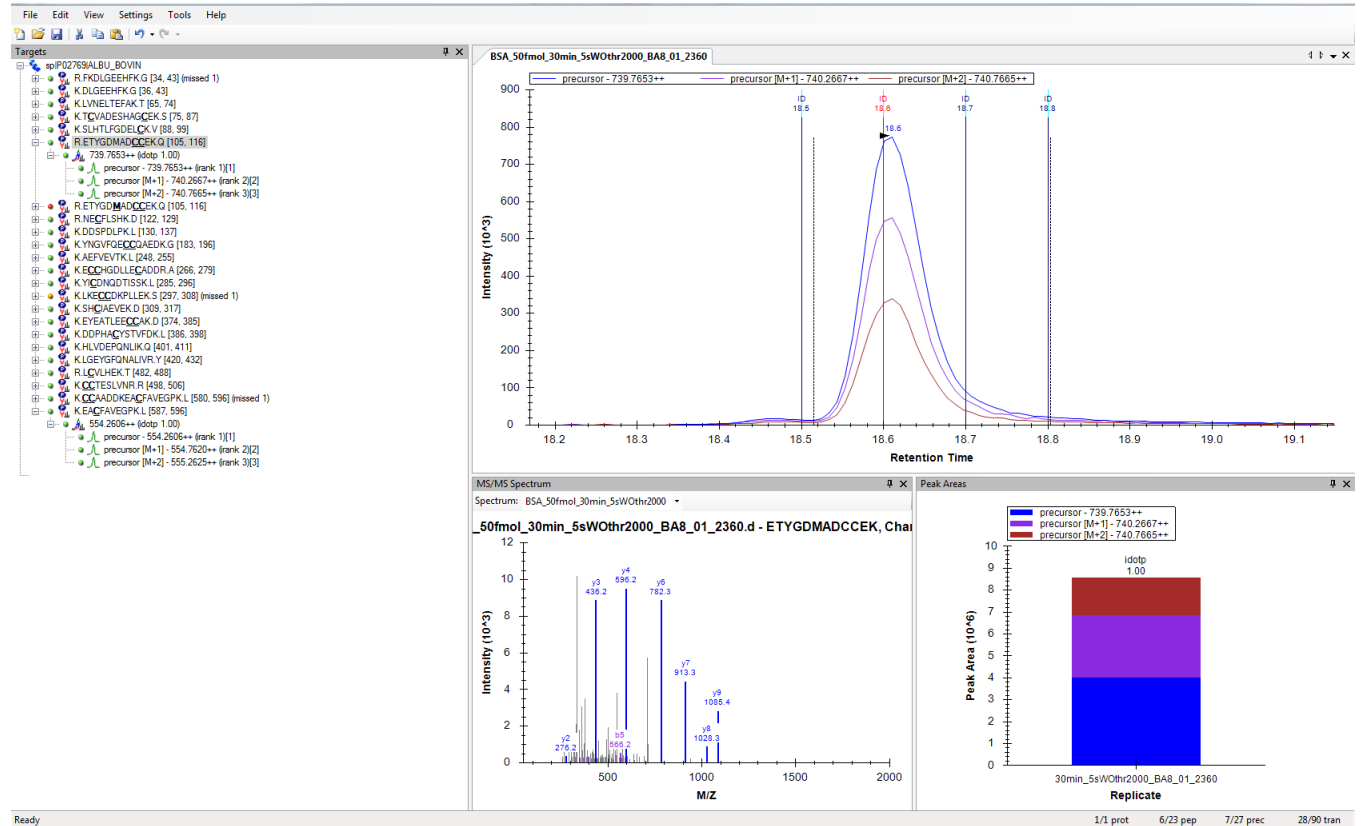
- Multiplexing capability (689 transitions)
- Signal fluctuations
- Retention time variability
- Platform comparisons
- Robustness over time
- Peptide storage over time, ...

# Global/Discovery proteomics approaches with Skyline

## MS1- filtering



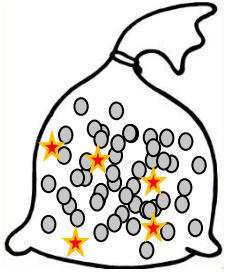

Q-TOF MaXis and  
Q-TOF Compact  
(Bruker Daltonics)



➔ Even easier integration of full-scan/discovery results with follow-up targeted experiments !



# From Global to Targeted Proteomics Approaches

## Global/Discovery Proteomics

Qualitative	Quantitative
	
500-2000 identified proteins	Poorly reproducible, approx. quantitation

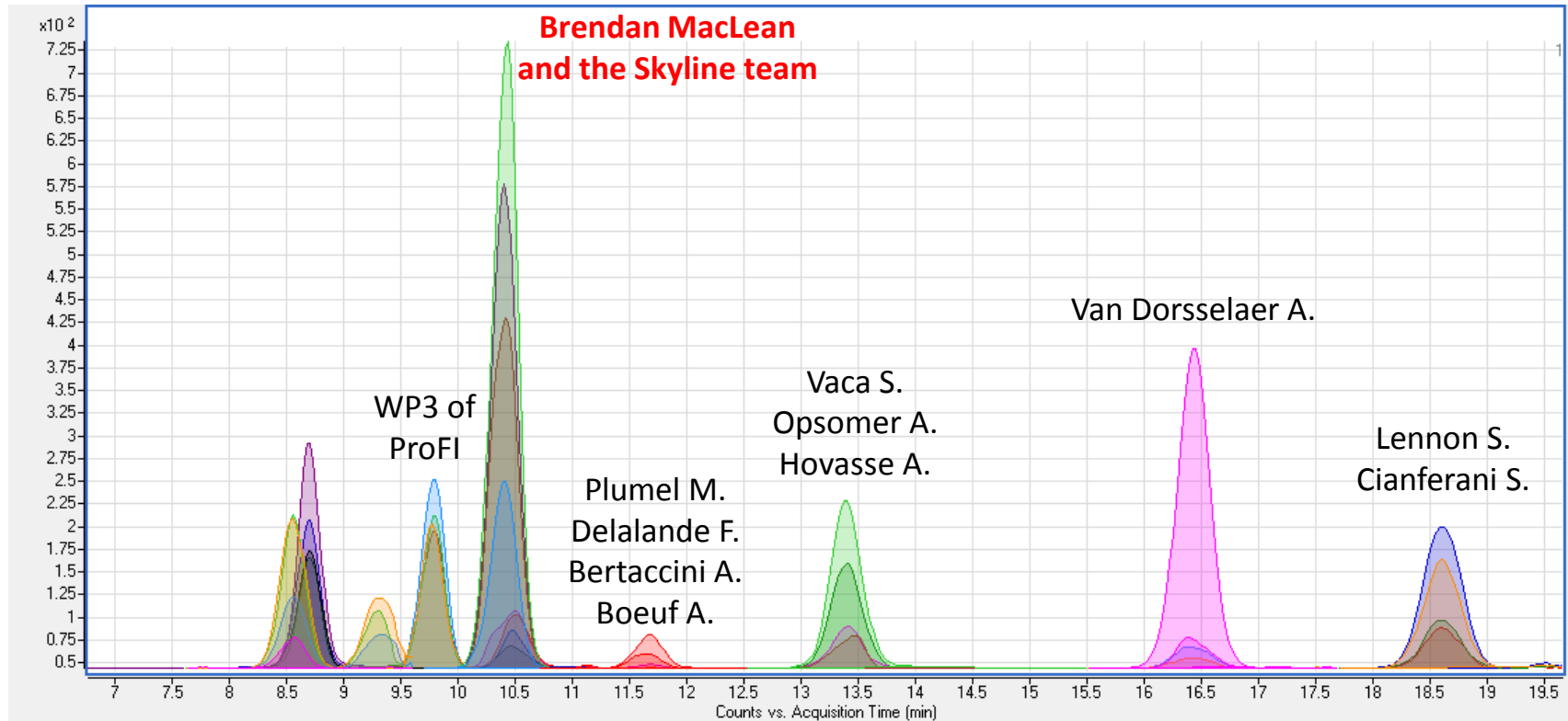


## Targeted Proteomics

Qualitative	Quantitative
	
10-100 candidate proteins	Precise reproducible, absolute quantitation



# Thanks !



WP3 of the French Proteomics Infrastructure (Garin J.) :

-Grenoble : Benama M., Adrait A., Ferro M.

-Strasbourg : Opsomer A., Vaca S., Hovasse A., Schaeffer C., Carapito C.

-Toulouse : Garrigues L., Dalvai F., Stella A., Bousquet M.P., Gonzales A.