UNIVERSITÉ PARIS SUD

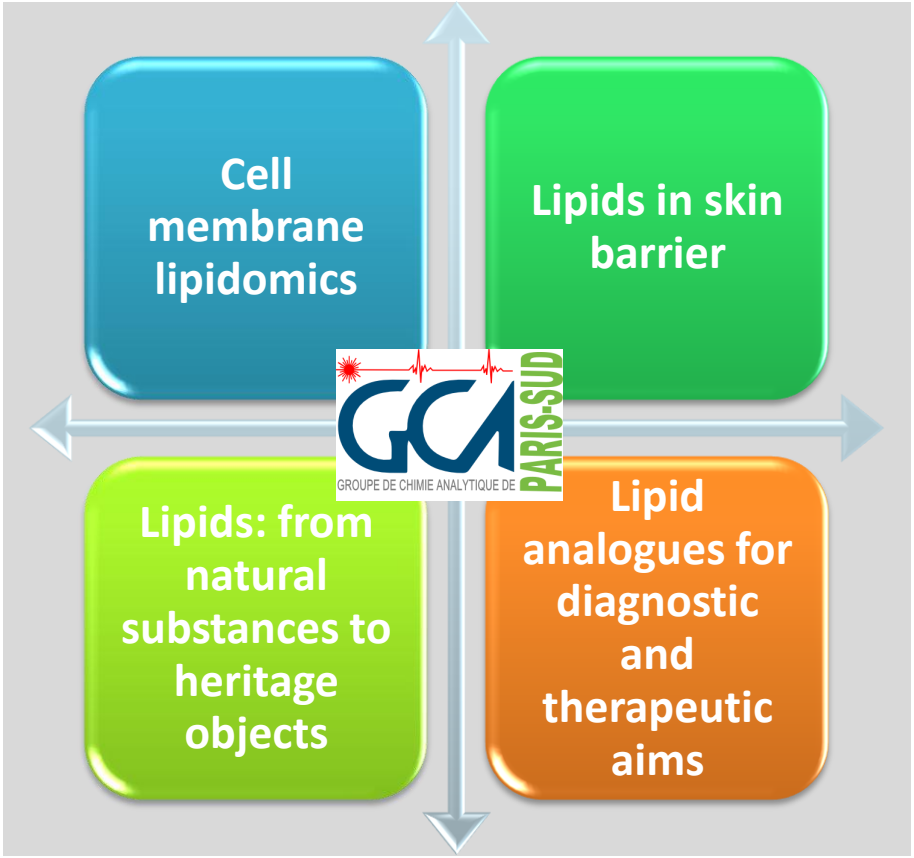
FACULTÉ DE PHARMACIE

GROUPE DE CHIMIE ANALYTIQUE DE PARIS-SUD

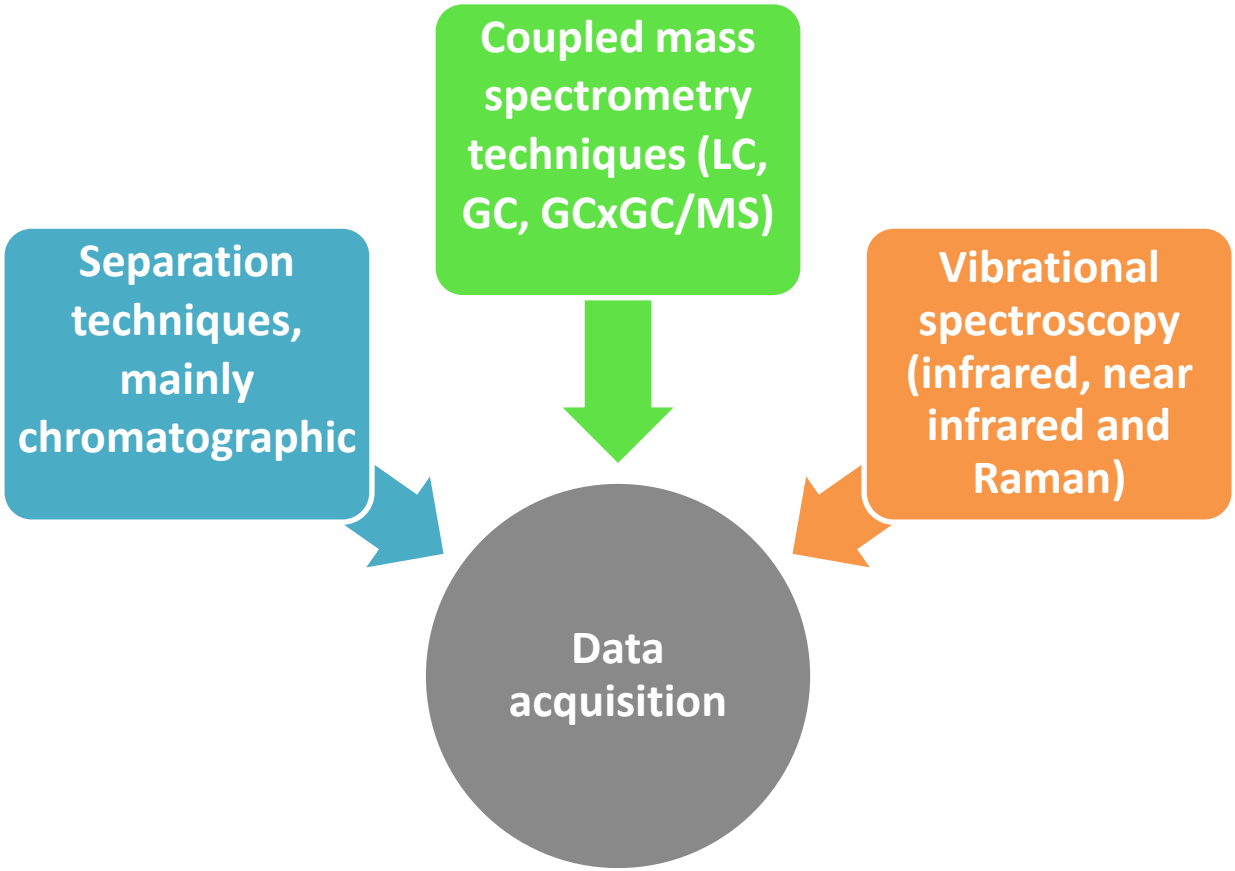
Analytical chemistry: from data (pre)-treatment optimization to data mining, data fusion and big data

EA4041, Groupe de Chimie Analytique de Paris-Sud ‘GCAPS’

Four principle themes



Different analytical tools



Data (pre)-processing
Multivariate analysis
Chemometric techniques

Schedule

10h00-10h30:
«Analytical chemistry: cell membrane lipidomics and data analysis», Sana Tfaili

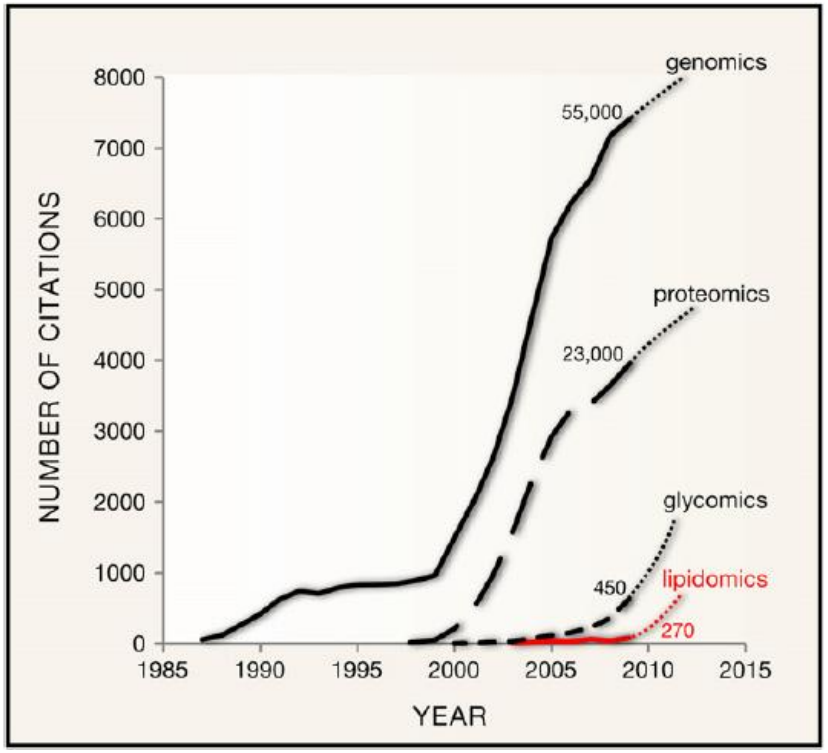
10h30-11h00:
«Analytical chemistry and chemometrics: a tool for skin physiological and physiopathological charecterization», Ali Tfayli

11h00-11h10:
«Analytical chemistry: from data (pre)-treatment optimization to data mining, data fusion and big data», S&A Tfai(y)li

Analytical chemistry: cell membrane lipidomics and data analysis

Sana Tfaili

Lipidomics



Cell

Volume 143, Issue 6, 10 December 2010, Pages 888–895

Primer

Lipidomics: New Tools and Applications

Markus R. Wenk[1, 2]

- Recent comparing to other omics
- Borrows heavily from metabolomics
- Specific analysis of lipids

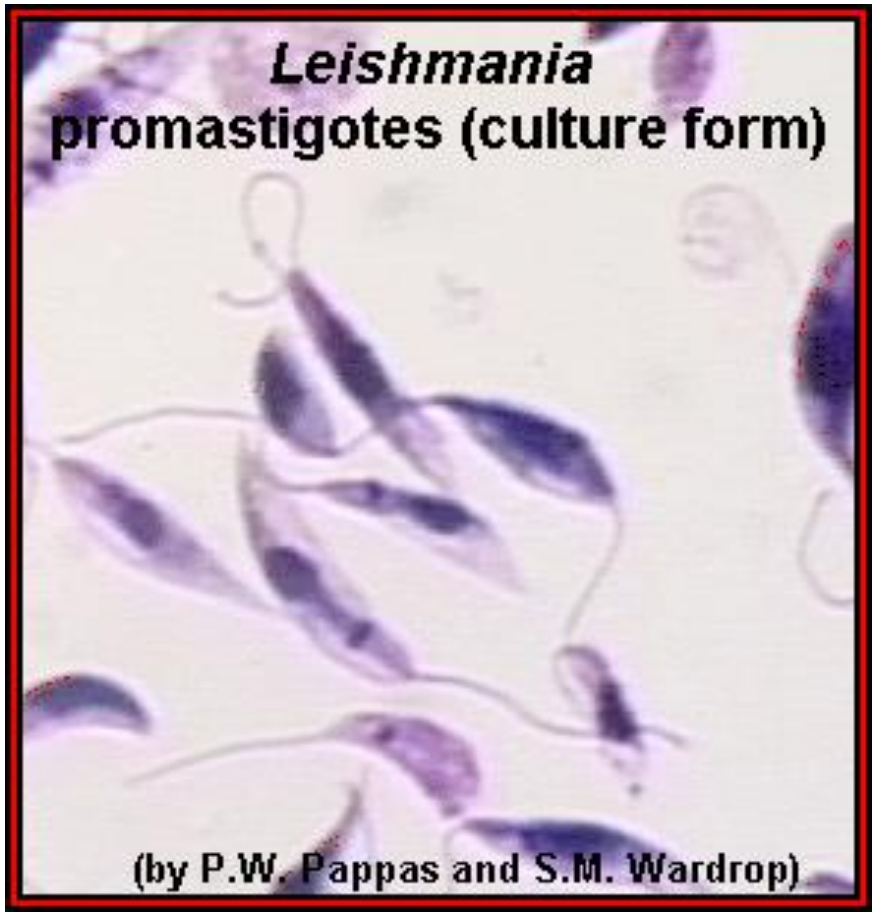
Analytical chemistry: cell membrane lipidomics and data analysis

Infectious diseases

•Leishmaniasis

•Impact of treatment on the lipid composition of membranes Leishmania donovani

Collaboration UMR BioCIS / Ph. Loiseau



Analytical chemistry: cell membrane lipidomics and data analysis

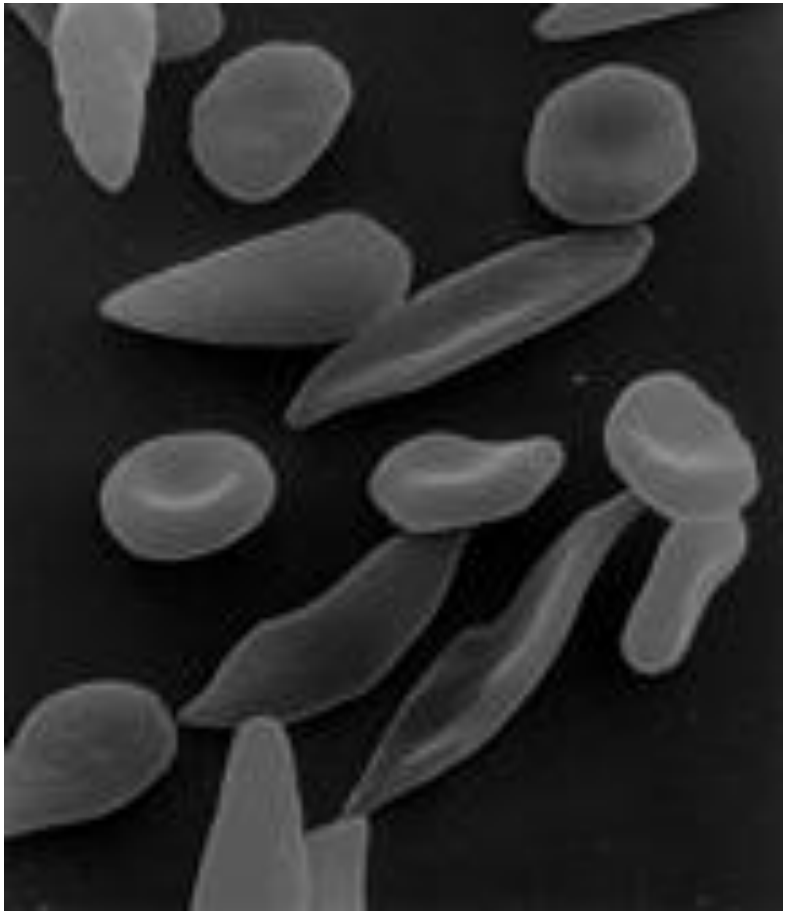
Hereditary diseases and RBCs

- Lipid composition and sickle cell disease

Olivier Blanc-Brude Paris Centre de Recherche Cardiovasculaire (PARCC)

- Lipid composition and Gaucher disease (LETIAM)

Institut National de Transfusion Sanguine (Pr. Le Van Kim et Dr. M Franco)
Service de Neuro-Pédiatrie de l'Hôpital Trousseau (Pr. T. Billette de Villemeur et Dr C. Mignot)



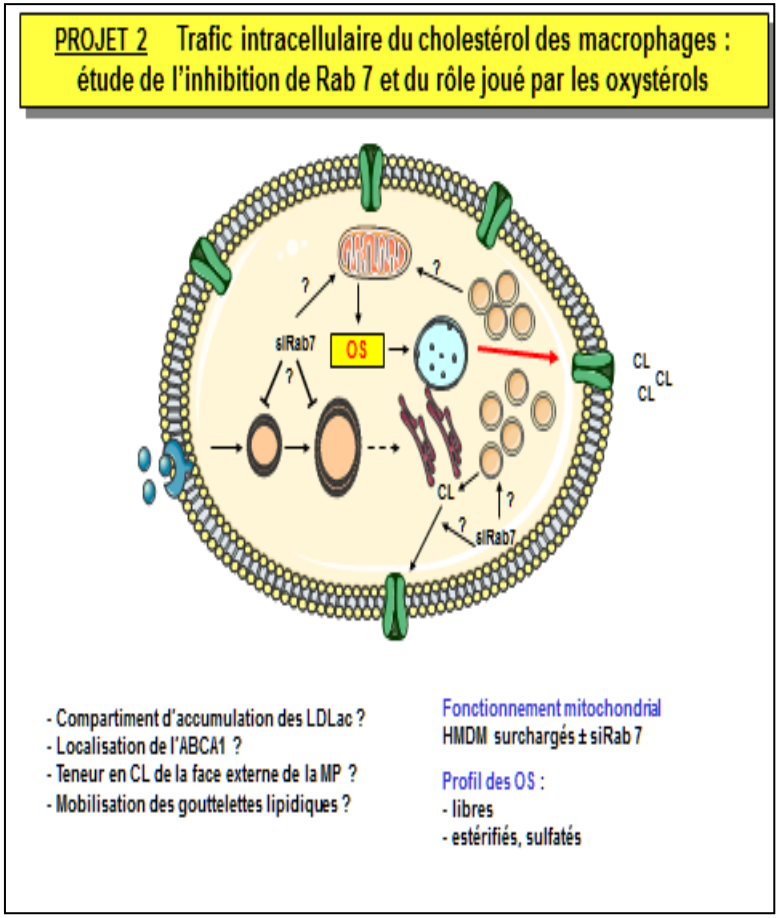
Analytical chemistry: cell membrane lipidomics and data analysis

Lipidomics macrophages & Atherosclerosis

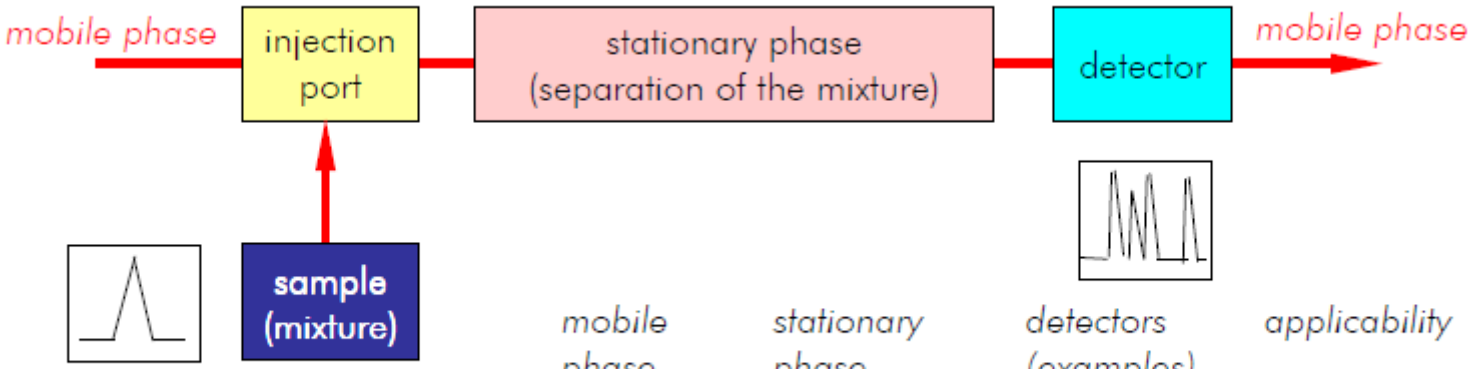
- Impact of membrane incorporation of w3 PUFA on cholesterol efflux from macrophages

- Intracellular trafficking of cholesterol from macrophages:study of the inhibition of Rab7 and the role of oxysterols

Collaboration EA4529 (cross-cutting theme in the future unit Lip (Sys) ²)

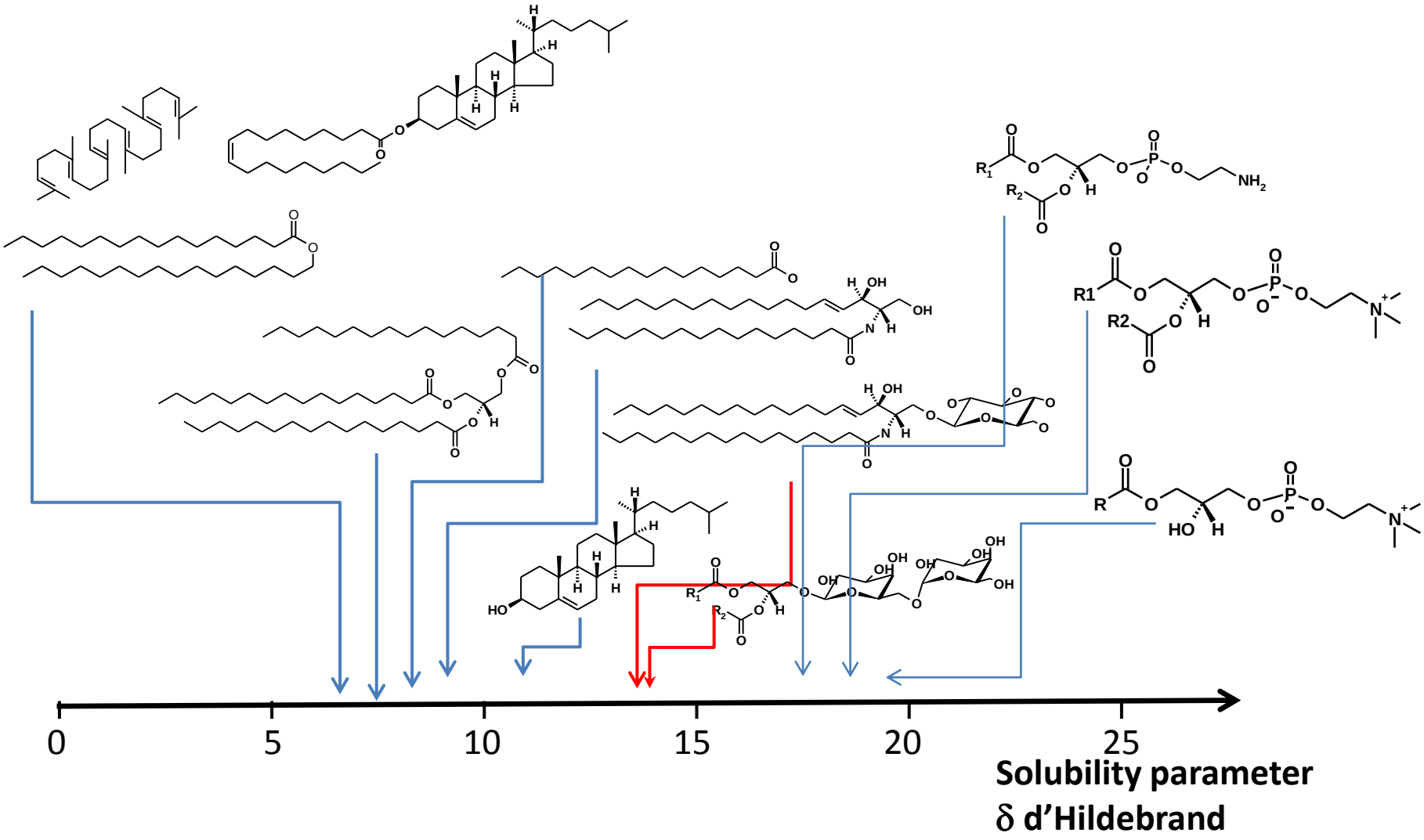


Chromatography



	mobile phase	*stationary phase*	*detectors (examples)*	*applicability*
Gas chromatography (GC):	gases (H_2, N_2, Ar, He)	capillary or packed columns with	TCD, FID, MSD	volatile organic compounds, permanent gases
High pressure liquid chromatography (HPLC):	H_2O, organic solvents	substituted siloxanes	UV-VIS, RI, MSD	organic compounds and inorganic salts
Thin layer chromatography (TLC):	vapor of organic solvents	impregnated plates	optical detection (UV-VIS, fluorescence)	organic compounds

Lipids and polarity

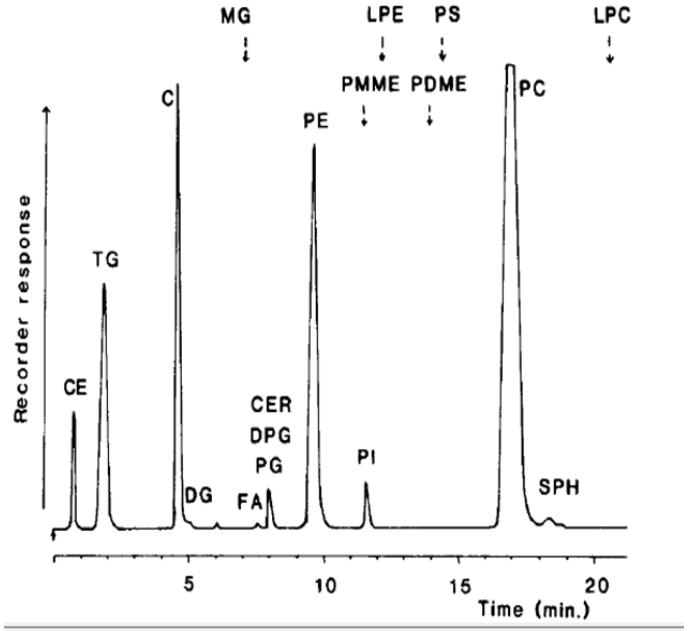


Normal Phase / Reversed Phase Liquid chromatography LC

LIPID CLASSES ANALYSIS

Lipides de foie de rat

Colonne: Spherisorb 3µm 150x5mm

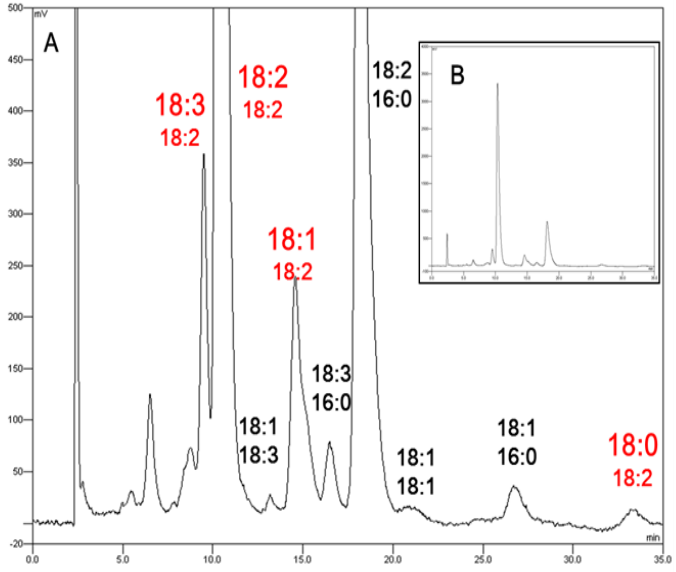


Rapid separation and quantification of lipid classes by high performance liquid chromatography and mass (light-scattering) detection. W W Christie, *1985 The Journal of Lipid Research, 26, 507-512.*

Lipids are separated according to their polar moeties.

Digalactosyldiglycéride

Colonne Hypercarb®

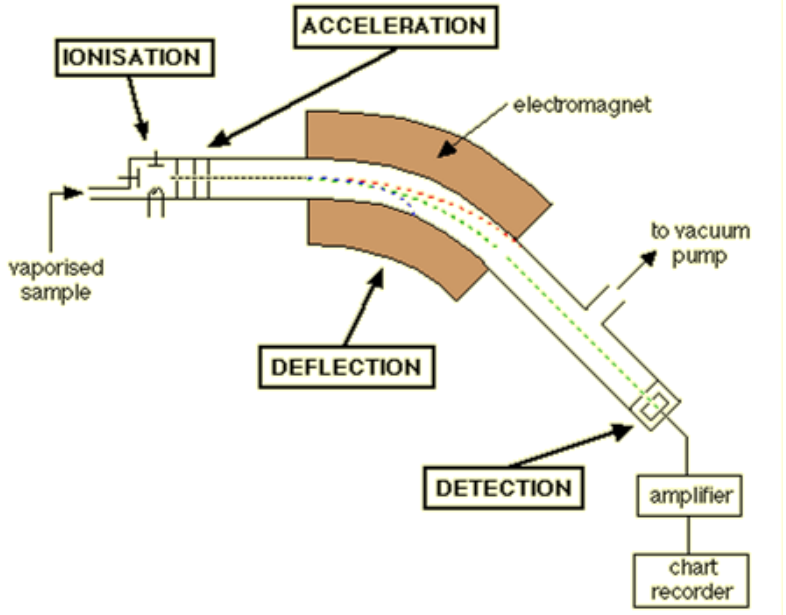


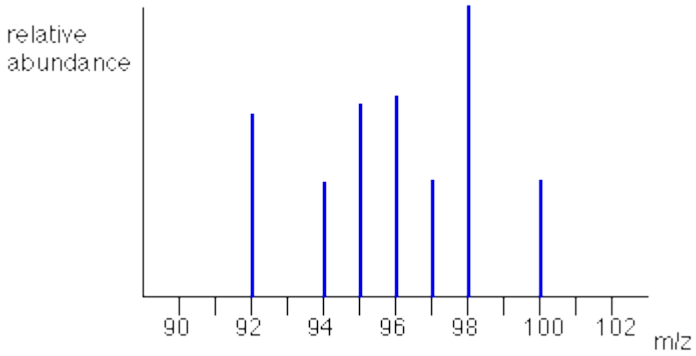
Deschamps F, Gaudin K, Baillet A & Chaminade P. Wheat digalactosyldiacylglycerol molecular species profiling using porous graphitic carbon stationary phase. *JOURNAL OF SEPARATION SCIENCE* (2004) **27**: p. 1322.

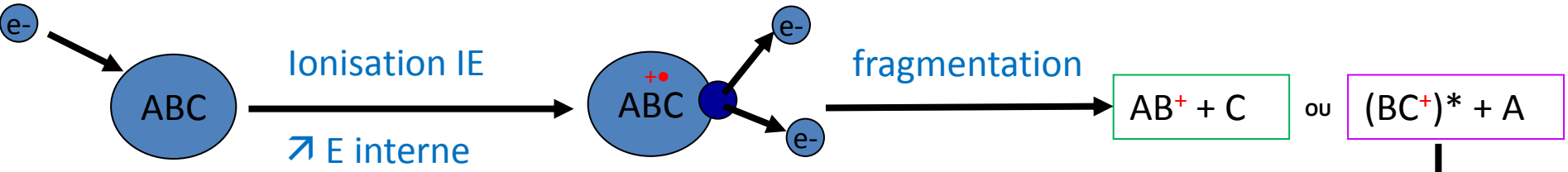
Lipids are separated according to their chain length & number and position of double bonds

LIPID MOLECULAR SPECIES ANALYSIS

Mass spectrometry MS

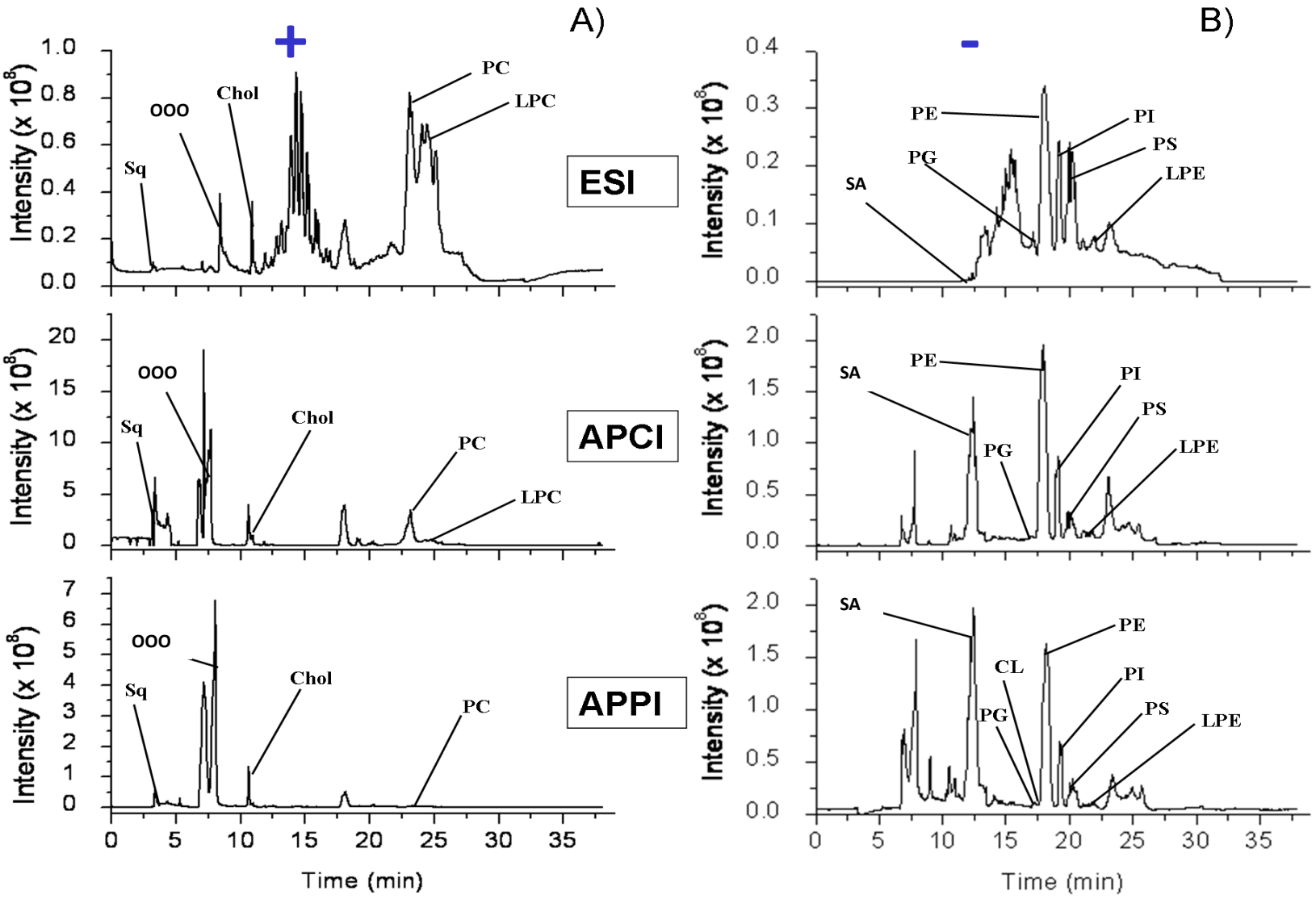






Mass spectrometry by electronic impact (GC-MS)

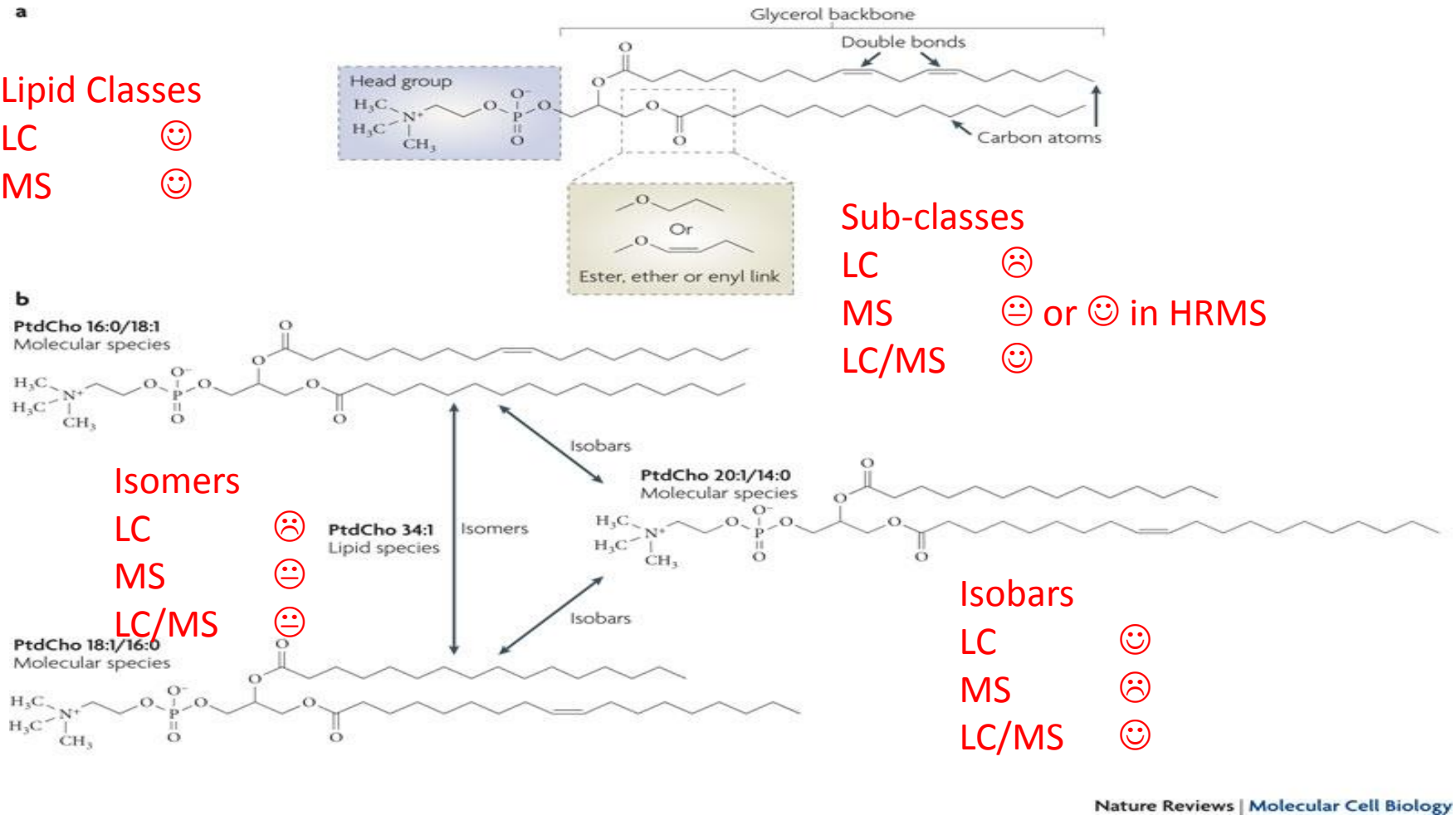
LC-MS acquisitions: Different ionization modes



Laurent Imbert, PhD thesis, GCAPS,EA4041, 2012, Univ. Paris Sud

Lipid molecular species

Lipid Classes
LC ☺
MS ☺



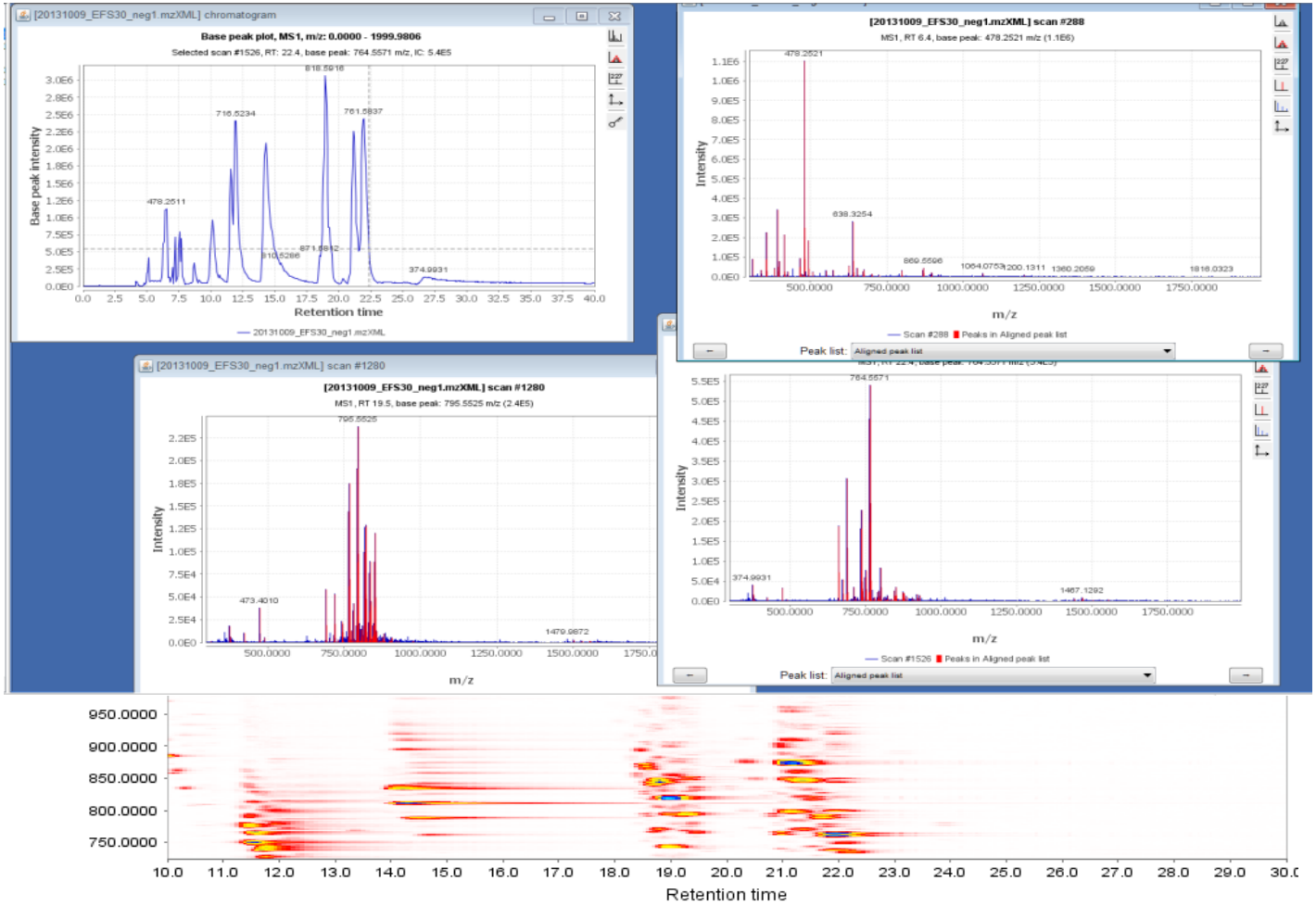
Sub-classes
LC ☹
MS 😐 or ☺ in HRMS
LC/MS ☺

Isomers
LC ☹
MS 😐
LC/MS 😐

Isobars
LC ☺
MS ☹
LC/MS ☺

[Lipidomics: coming to grips with lipid diversity]
Andrej Shevchenko & Kai Simons
Nature Reviews Molecular Cell Biology 11, 593-598 (August 2010)

LC-MS data matrix



Data = 3D matrix

need to concatenate and "unfold" files

In the data matrix:

Objects (lines) = sample

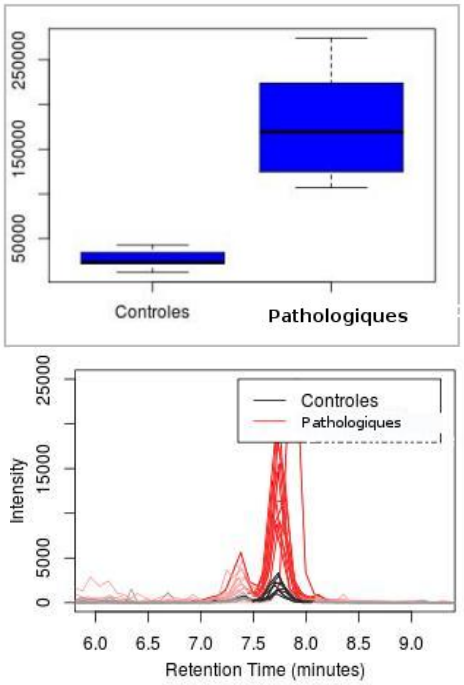
Variables (columns) = couple (Tr, m/z)

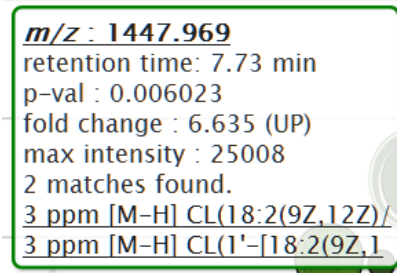
LC-MS data matrix processing

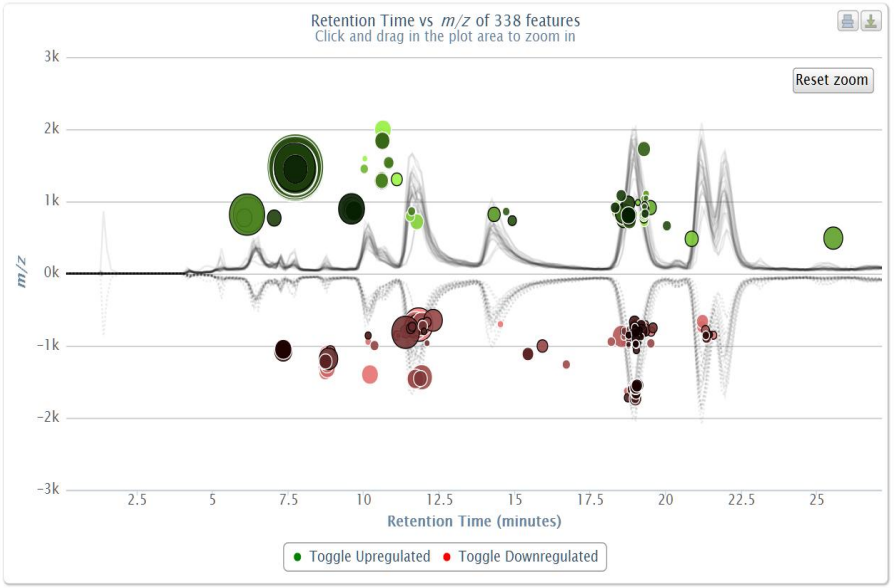
Univariate statistical analysis using XCMS online:

Paired Student t-test between the two groups of signals

Evident significant difference between the intensities of the ions.



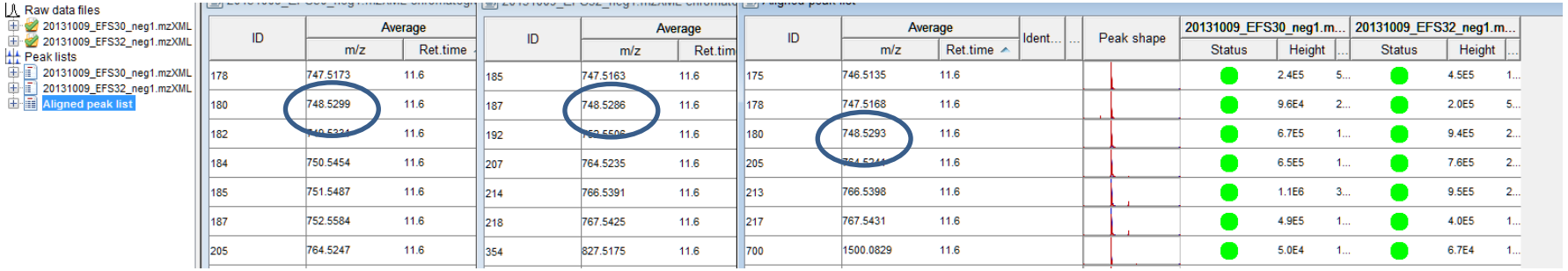




p-Value : 0-0.001 | fold change : 1.5 - 129.990 | *m/z* Range of 0 - 2197 | Retention Range of 1-42 | max intensity : 0-4444805.475 | Radius Scale : Fold Change

LC-MS data matrix processing

Alignment tools using MzMine

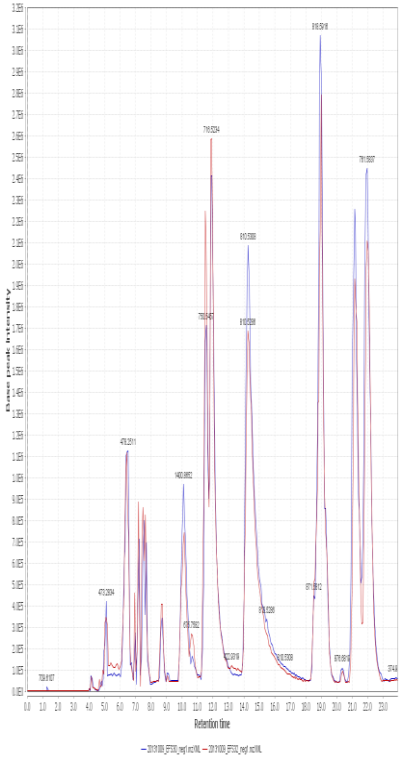


Step 1: file by file, detection of ions (> threshold) scan by scan

Step 2: ion chromatogram generation, file by file

Step 3: file by file, peak detection table (ion; rt, intensity) (data file)

Step 4: Alignment: Setting a tolerance window (m/z, rt) based in general on the first chromatogram. (data set)

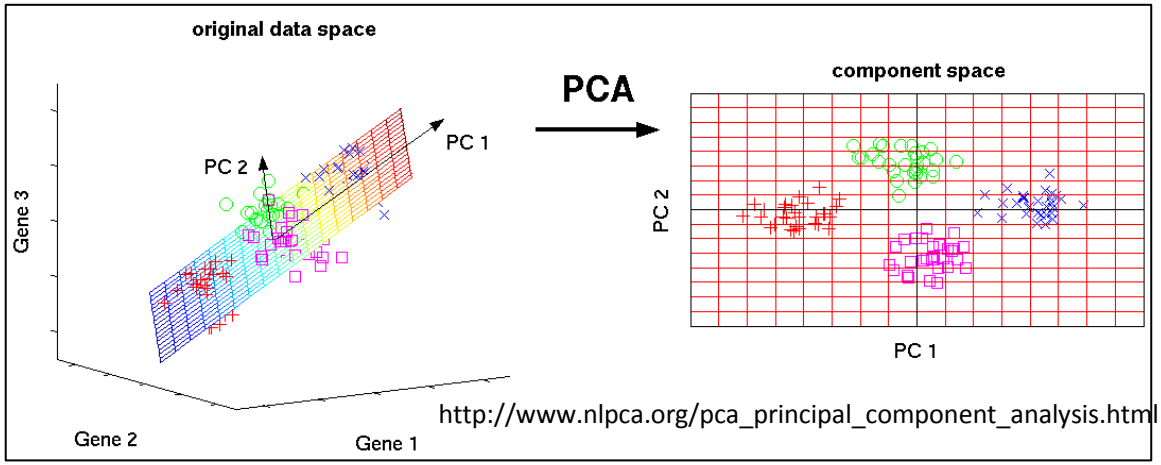


LC-MS data matrix processing

Principal component analysis

Unsupervised method

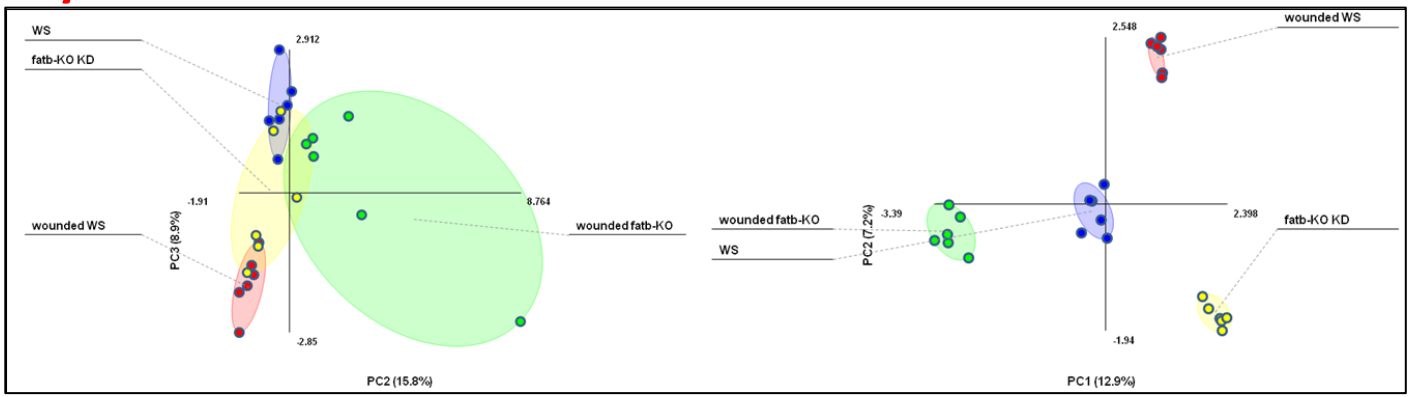
Data mining



http://www.nlpca.org/pca_principal_component_analysis.html

PLS Discriminant analysis

Supervised method

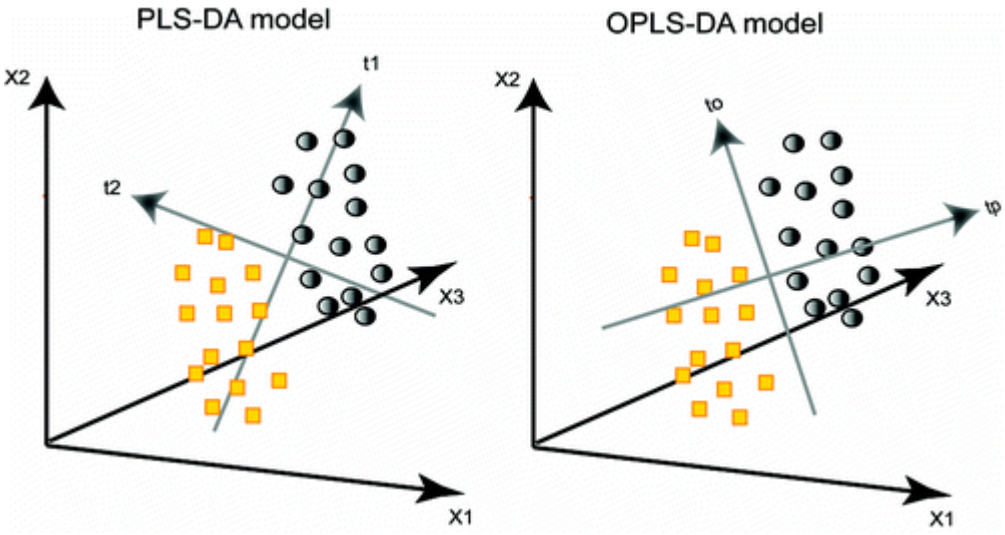


http://fiehnlab.ucdavis.edu/staff/kind/Statistics/Concepts/OPLS-PLSDA

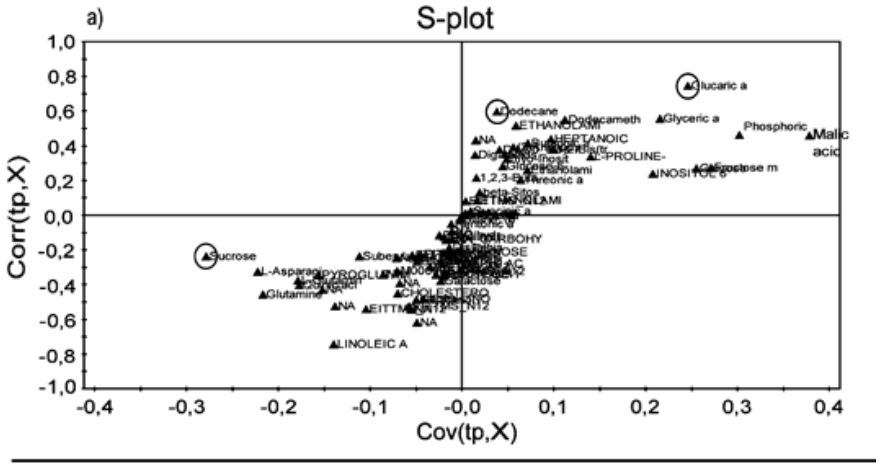
LC-MS data matrix processing

Orthogonal PLS-DA (OPLS-DA)

Maximizes the discrimination between two classes in its first component



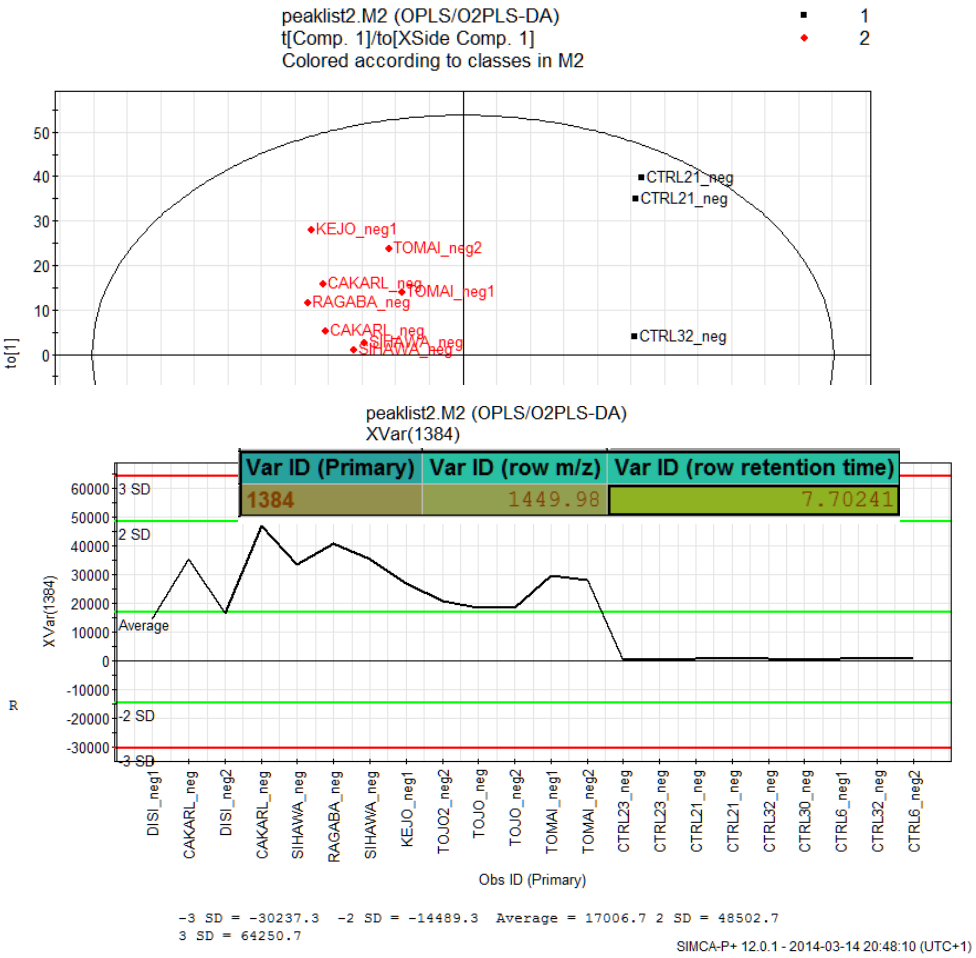
S-Plot expresses the relationship between the original variables (rt, m/z) and scores on the selected axis.

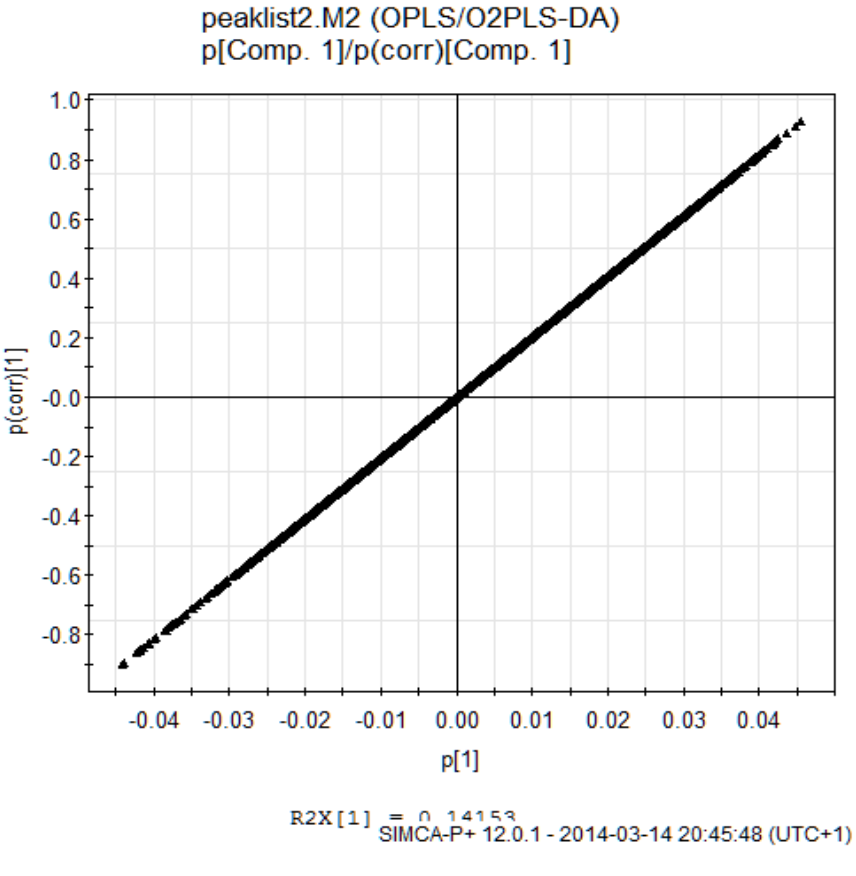


Published in: Susanne Wiklund; Erik Johansson; Lina Sjöström; Ewa J. Mellerowicz; Ulf Edlund; John P. Shockcor; Johan Gottfries; Thomas Moritz; Johan Trygg; Anal. Chem. 2008, 80, 115-122.

LC-MS data matrix processing

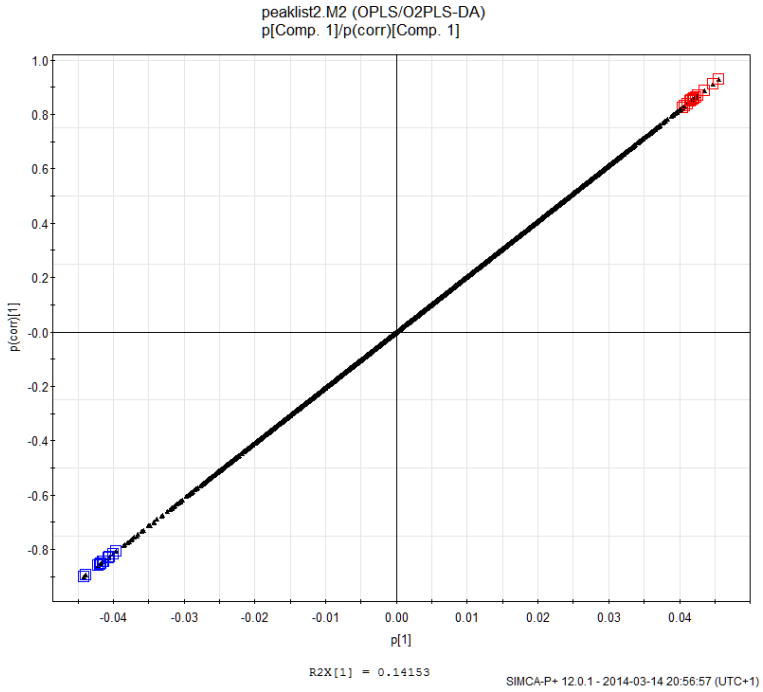
O-PLS and S-plot using SimcaP





LC-MS data matrix processing

S-plot / Selection of discriminate variables

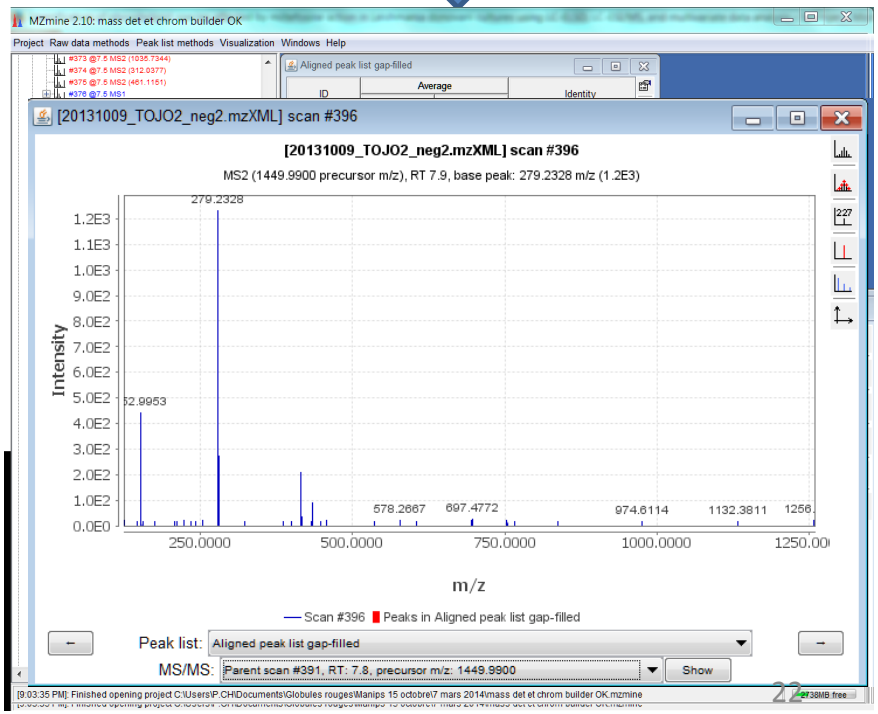


General List [M2]

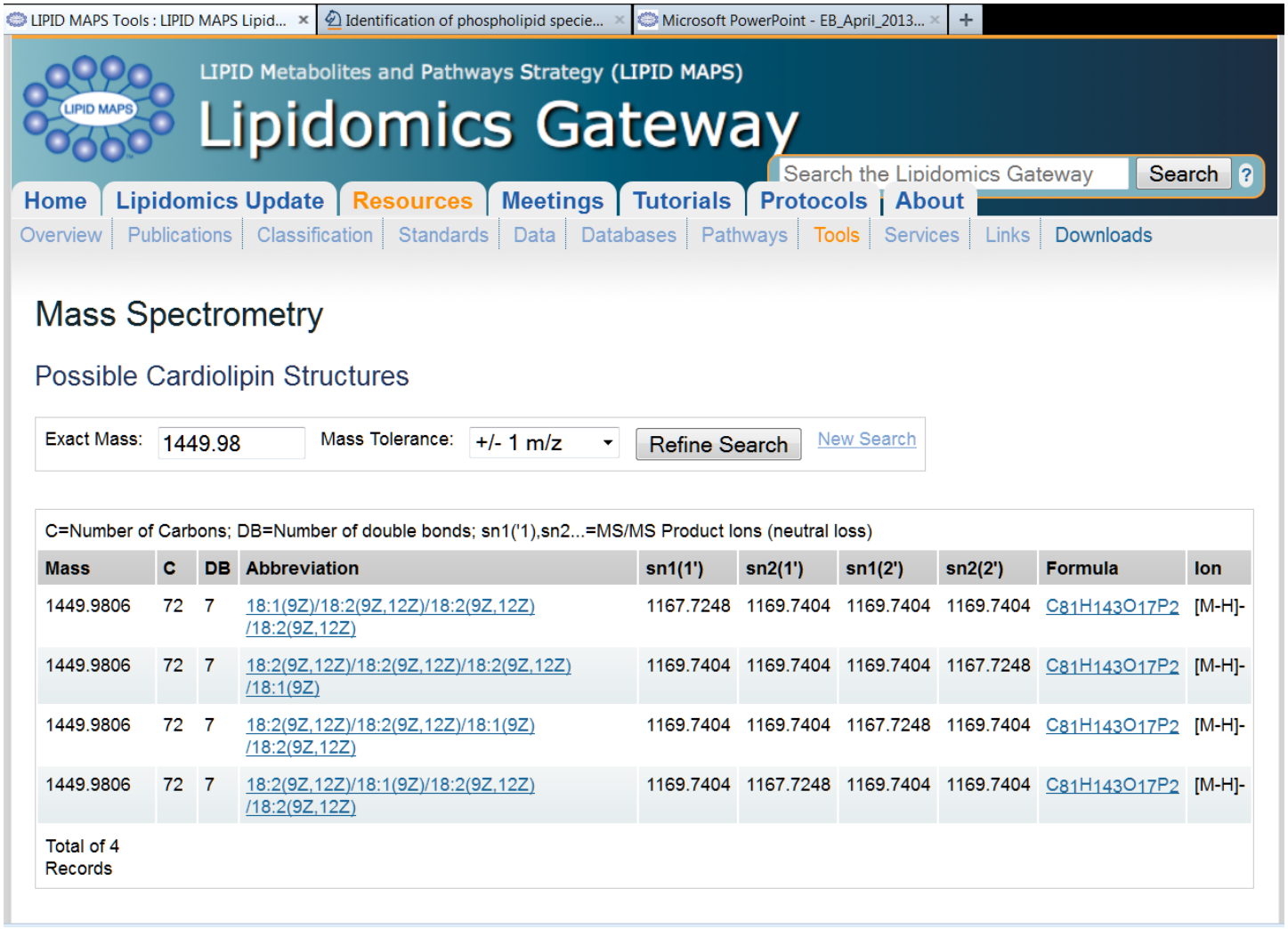
	1	2	3	4	5
1	Var ID (Primary)	Var ID (row m/z)	Var ID (row retention time)	M2.p[1]	M2.p(corr)[1]
2	1280	1577.11	19.0169	0.0454953	0.927955
3	1561	1577.11	19.0146	0.0447281	0.912305
4	2475	720.497	11.7137	0.0423014	0.862809
5	2512	898.579	19.0235	0.0419978	0.856616
6	2661	1551.1	19.0404	0.0435393	0.888057
7	2683	814.561	18.9809	0.0421245	0.8592
8	2684	1550.1	19.0077	0.0425956	0.86881
9	2957	901.554	19.0197	0.0419536	0.855716

General List [M2]

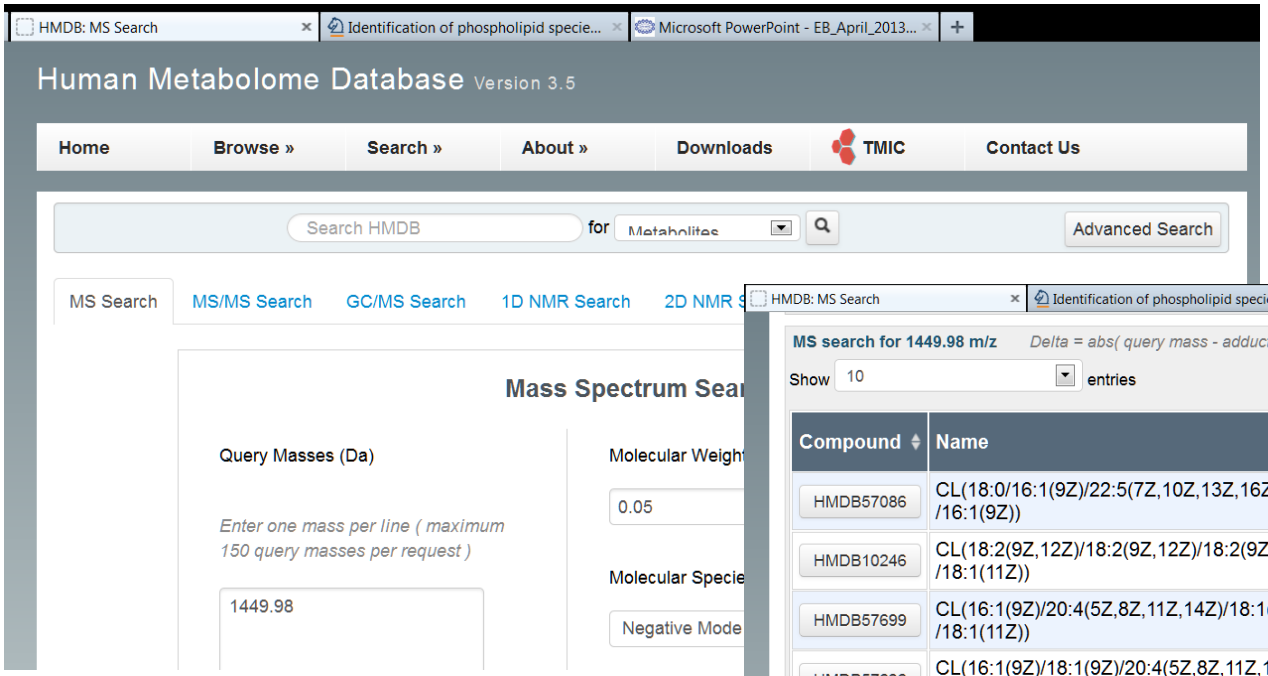
	1	2	3	4	5
1	Var ID (Primary)	Var ID (row m/z)	Var ID (row retention time)	M2.p[1]	M2.p(corr)[1]
2	294	807.588	18.7738	-0.0416959	-0.850459
3	305	804.61	18.7946	-0.0399648	-0.815149
4	838	830.627	18.6037	-0.0395783	-0.807266
5	1198	730.576	18.8086	-0.0407172	-0.830496
6	1302	834.608	18.6162	-0.0414102	-0.844632
7	1363	730.576	18.813	-0.0438618	-0.894635
8	1384	1449.98	7.70241	-0.0441028	-0.899551
9	1763	1452	7.71264	-0.0421506	-0.859734
10	1774	730.575	18.8107	-0.0406005	-0.828116
11	1830	780.567	18.8099	-0.0419063	-0.85475
12	2023	1449.98	7.70241	-0.0414215	-0.844861

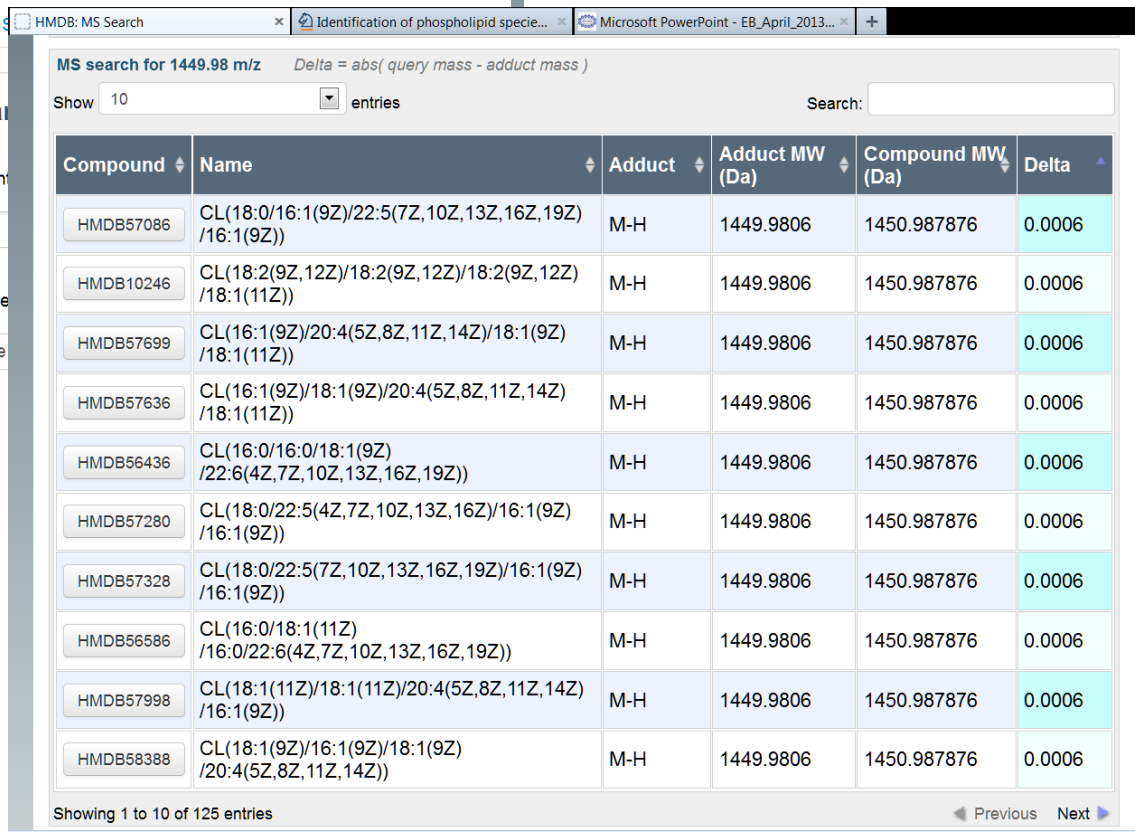


Online databases

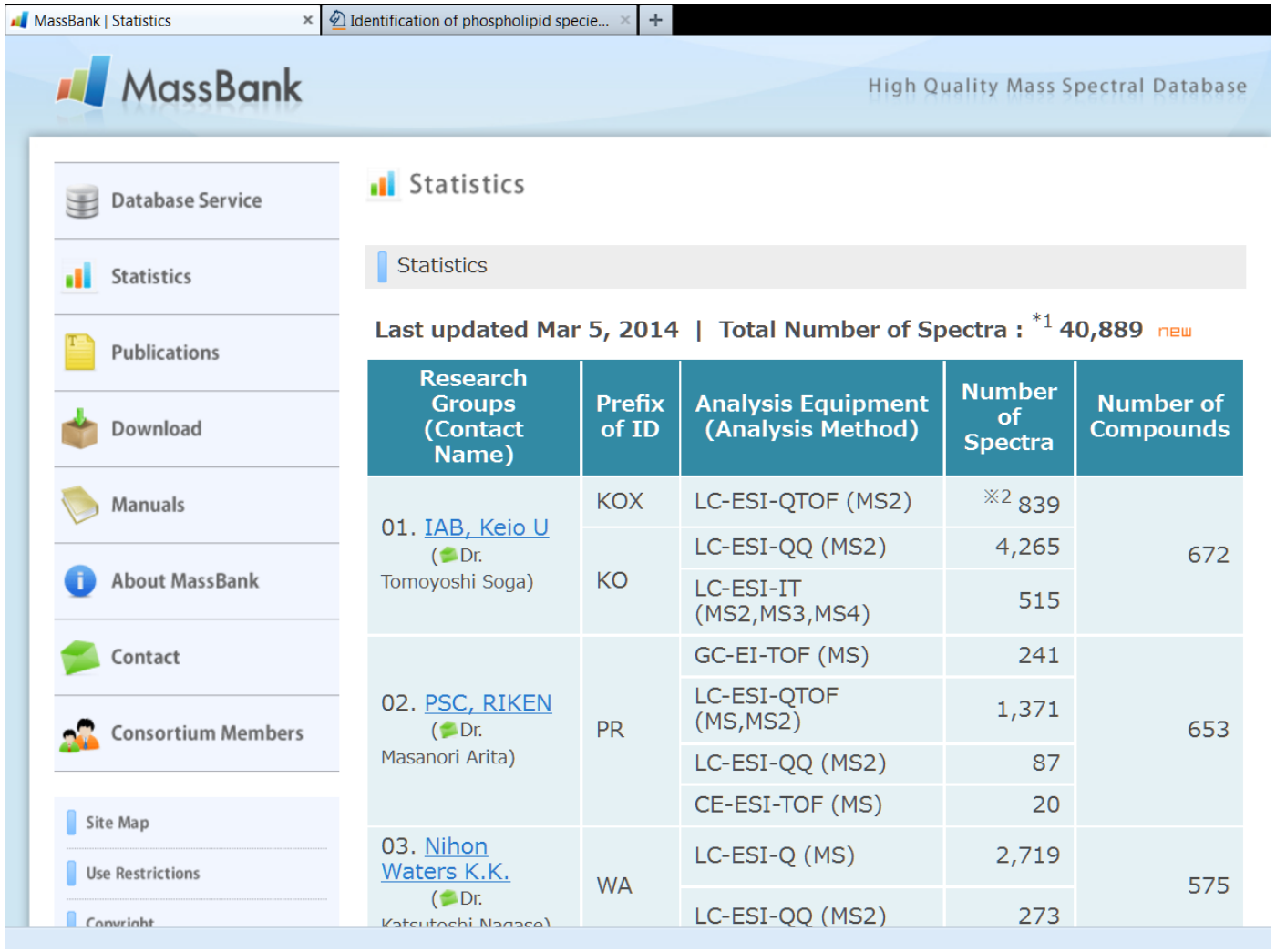


Online databases

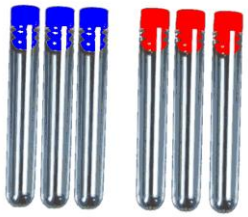




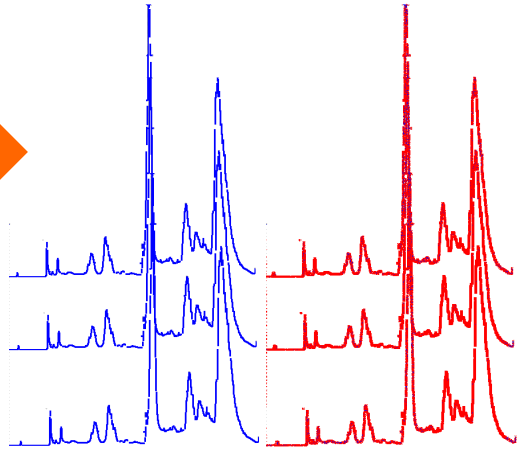
Online databases



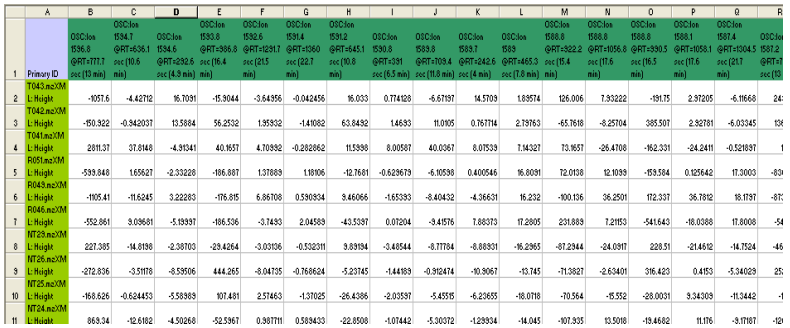
Chemometric tools for LC-MS lipidomics profiles analysis.



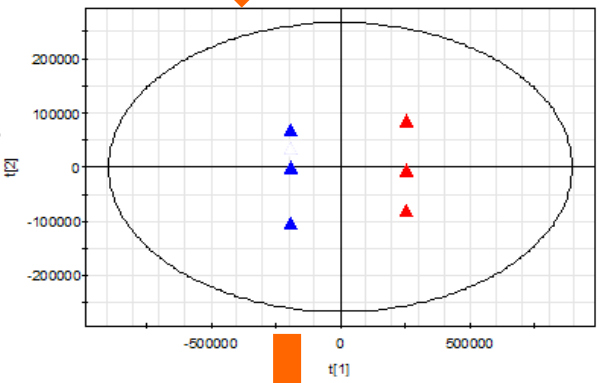
Samples



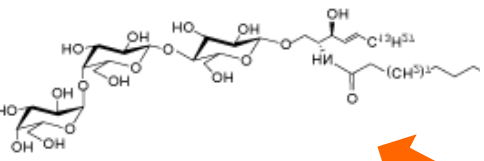
LC-MS profiles

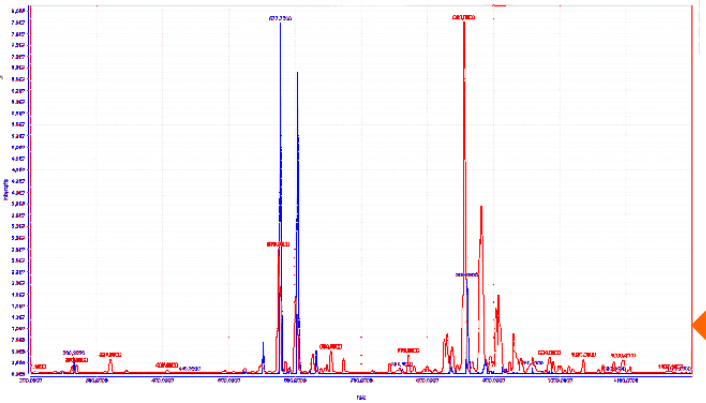


Statistical analysis

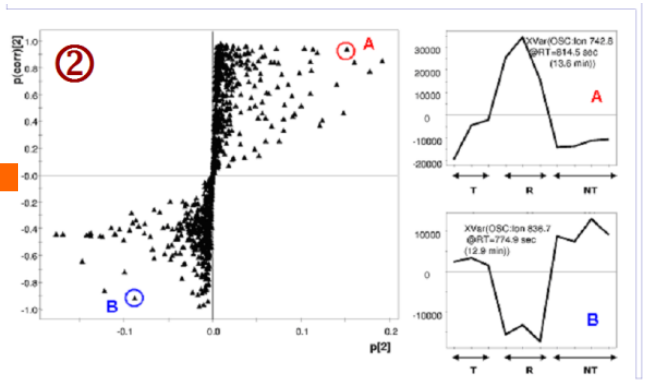


Biomarkers





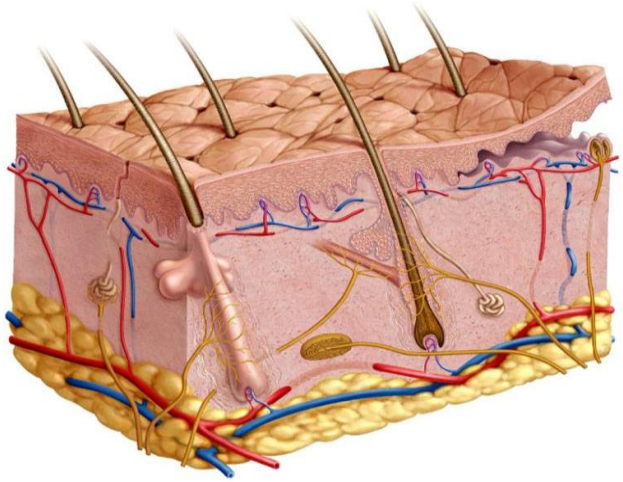
Highlighting specific MS signals of the 2 groups



Analytical chemistry and chemometrics: a tool for skin physiological and physiopathological characterization

Ali Tfayli

General structure of the skin



Largest organ in human body

Epidermis

- Stratum Basale
- Struatum Spinosum
- Stratum Granulosum
- **Stratum Corneum**

Dermis

- Superficial
- Deep

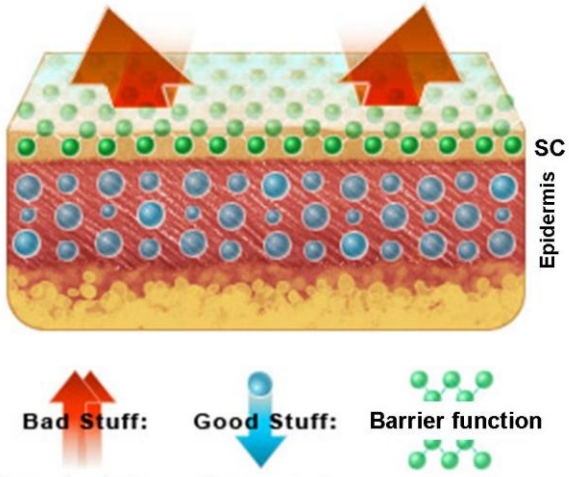
Barrier function

Sensation

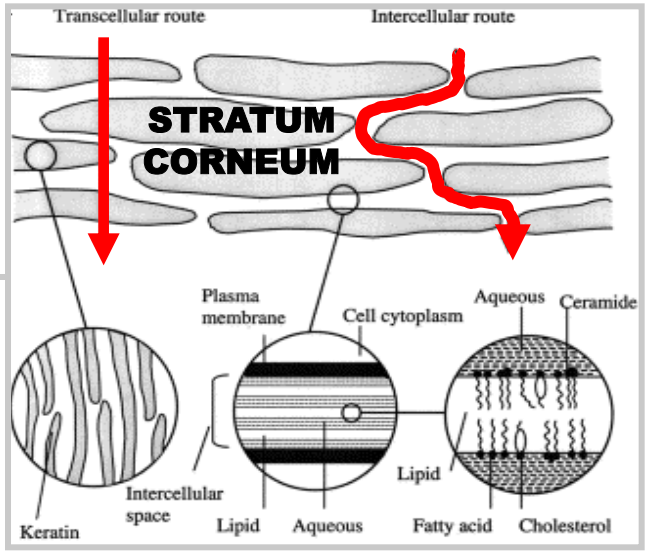
Heat regulation

Secretion end excretion

Skin Barrier function



SC structure

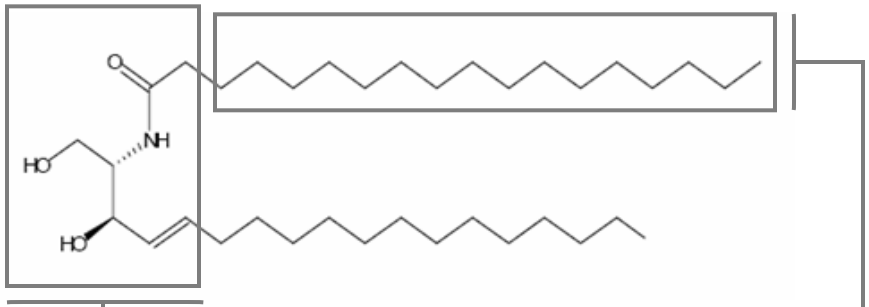


Barrier against

- External insult
- Water loss

- Keratin
- Intercellular lipids

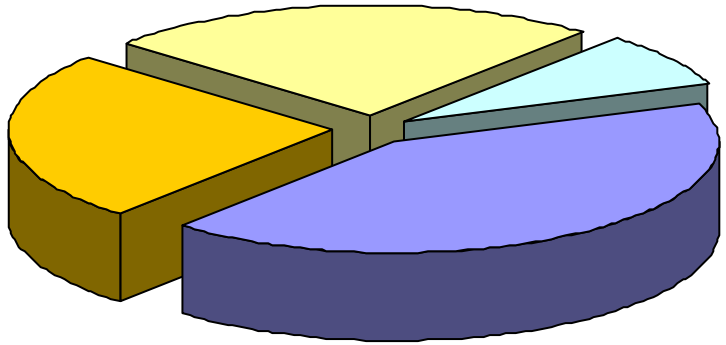
Skin Barrier function



Structure of Ceramides

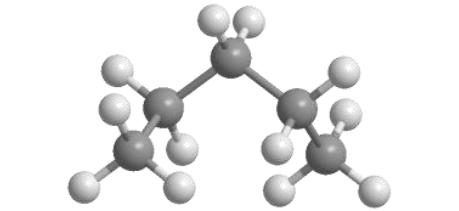
- Different polar heads
- Different chain lengths
- Presence of double bonds

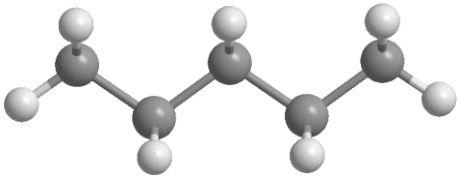
Composition of SC lipids



- CERAMIDES
- CHOLESTEROL
- FATTY ACIDS
- OTHERS

Skin Barrier function



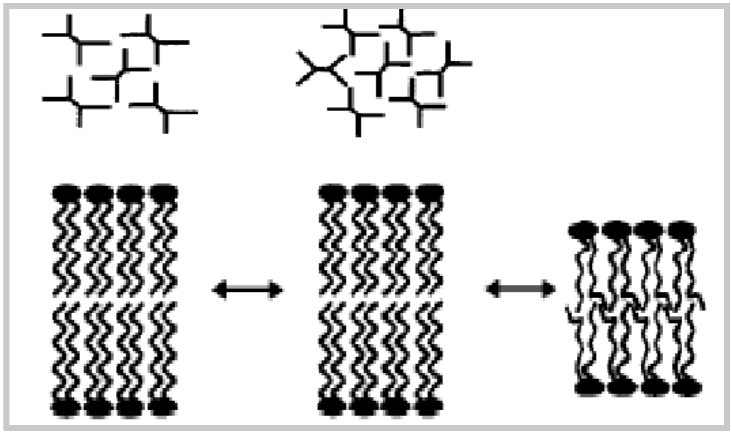


Conformational order

- *trans*
- *gauche*

Lateral packing of SC lipids

Orthorhombic | Hexagonal | Disordered



HIGHLY ORGANIZED

GOOD BARRIER PROPERTIES

Dermatology

Characterization of skin barrier

-Physiological status

-Physiopathological status

Skin aging

Skin hydration and dry skin diseases

Mechanical stress

Toxicology

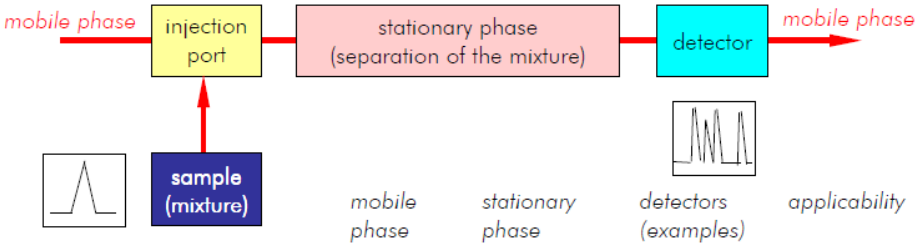
Cosmetology

Analyses of skin barrier

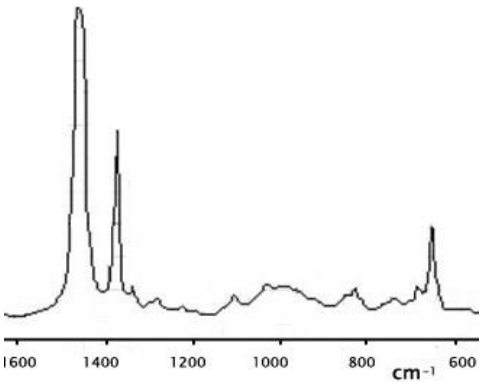
Organization, lateral packing

Composition and profiling

Separative techniques - mass



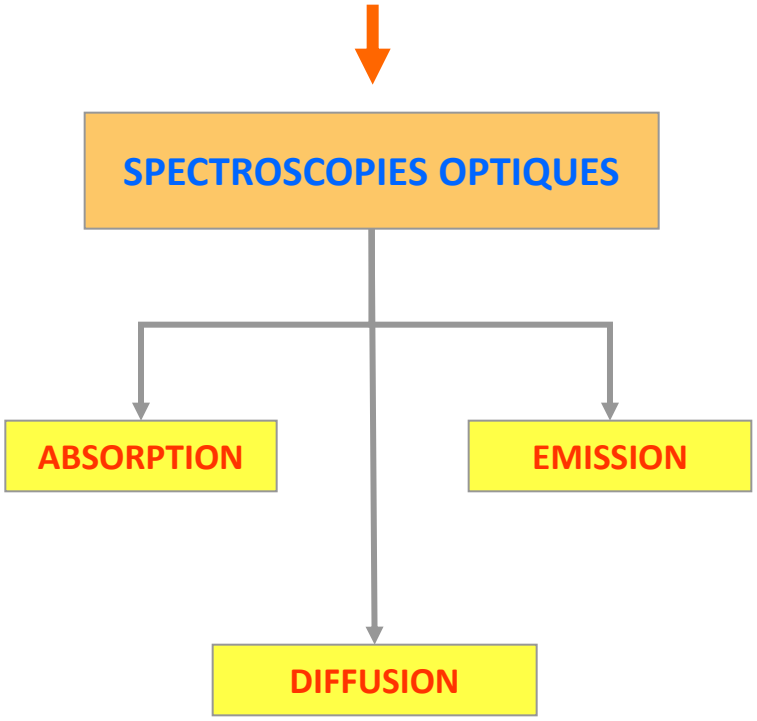
Vibrational spectroscopies: Infrared and Raman



Vibrational spectroscopies

INTERACTION

RAYONNEMENT - MATIERE



$$\bar{\nu} = \frac{\nu}{C} = \frac{1}{2\pi C}\sqrt{\frac{k}{\mu}}$$

$\bar{\nu}$ = nombre d'onde, ν = fréquence, μ = masse réduite

$$\mu = \frac{m_A . m_B}{m_A + m_B}$$

$$\mu_{tot} = \mu_{p(0)} + \sum_{n=1}^{3N-6}\left(\frac{d\mu_p}{dQ_n}\right)Q_0 \cos\nu_n t + \alpha_0(E_0 \cos\nu_0 t) + \frac{1}{2}E_0\sum_{n=1}^{3N-6}\left(\frac{d\alpha_0}{dQ_n}\right)Q_0\left[\cos(\nu_0 - \nu_n)t + \cos(\nu_0 + \nu_n)t\right]$$

Vibrational spectroscopies

Jablonski diagram

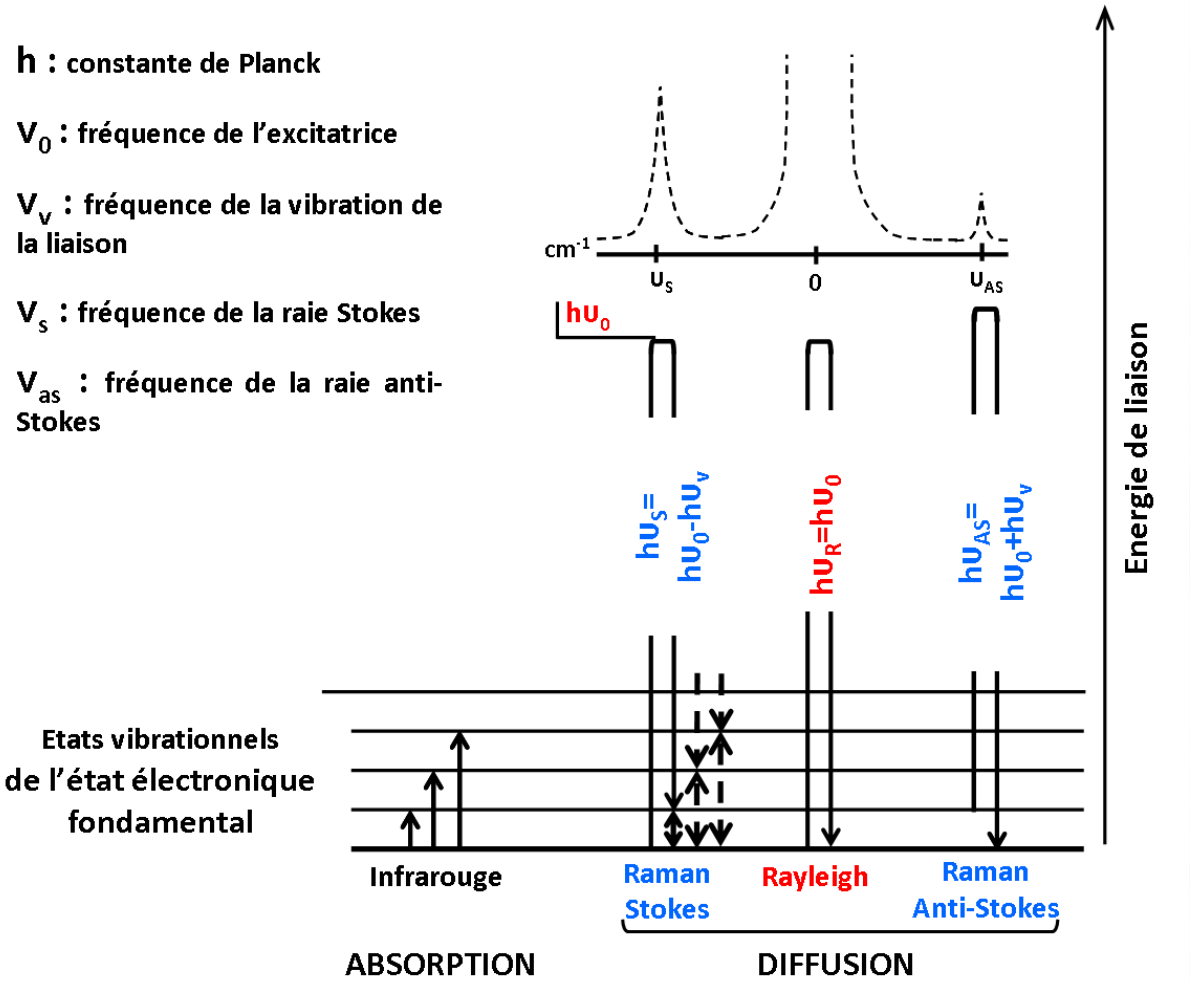
h : constante de Planck

V_0 : fréquence de l'excitatrice

V_v : fréquence de la vibration de la liaison

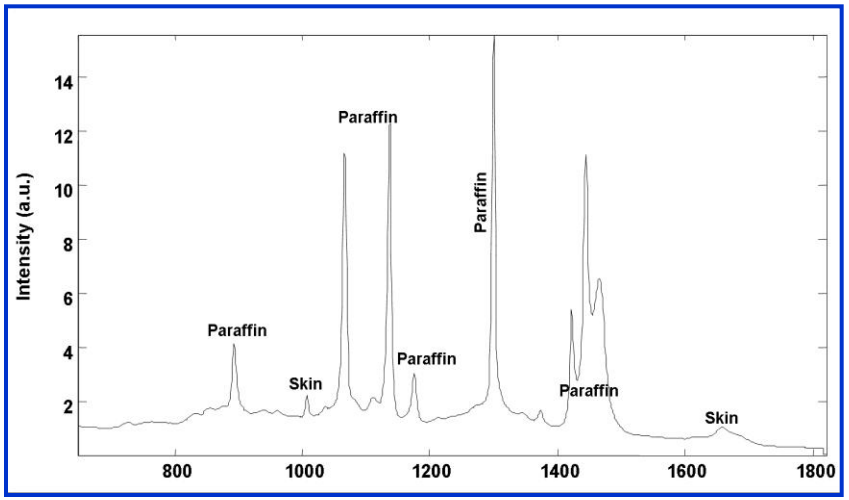
V_s : fréquence de la raie Stokes

V_{as} : fréquence de la raie anti-Stokes

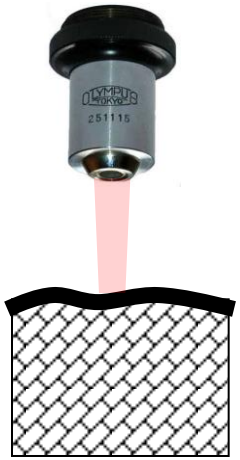


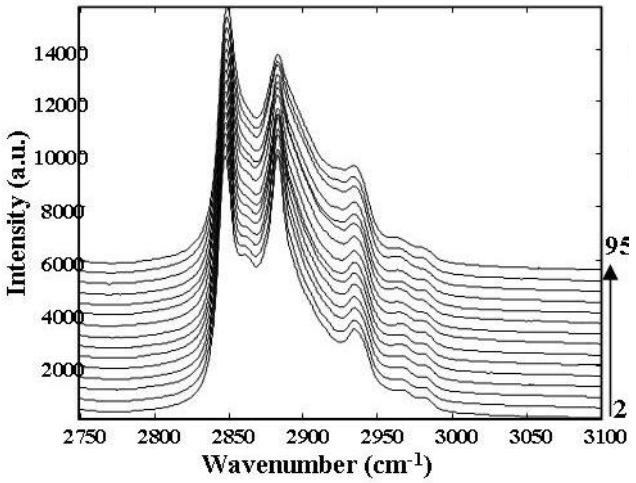
Vibrational signal collection

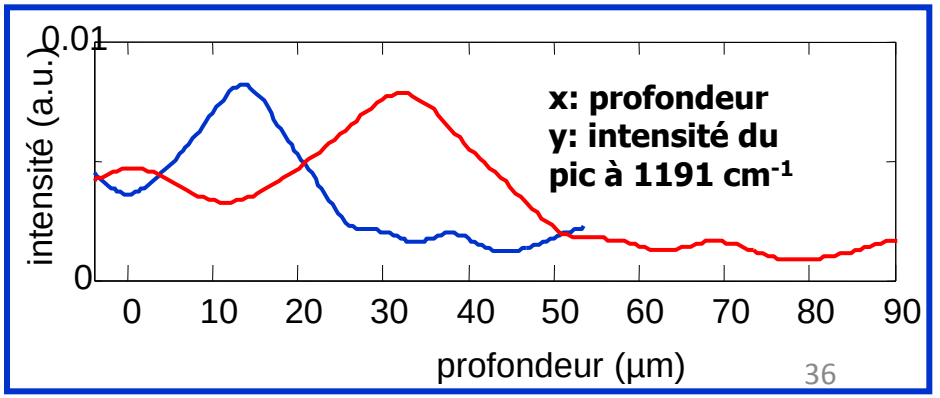
Individual spectral collection



In depth spectral collection

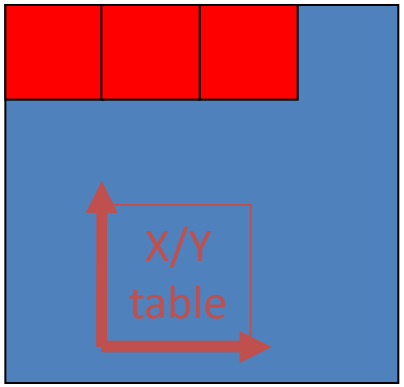




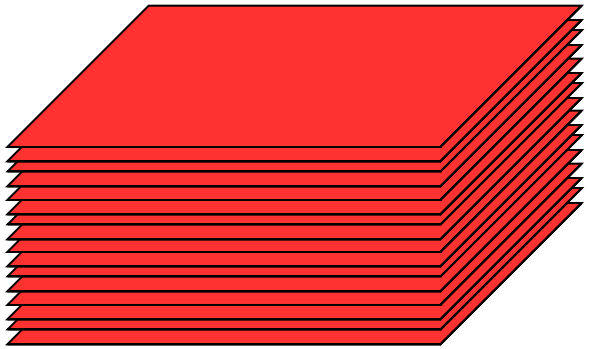


Vibrational signal collection

2D spectral mapping



3D spectral mapping



Data pre-processing

-Dark current

-CCD response correction

-Optical components contribution

-Smoothing

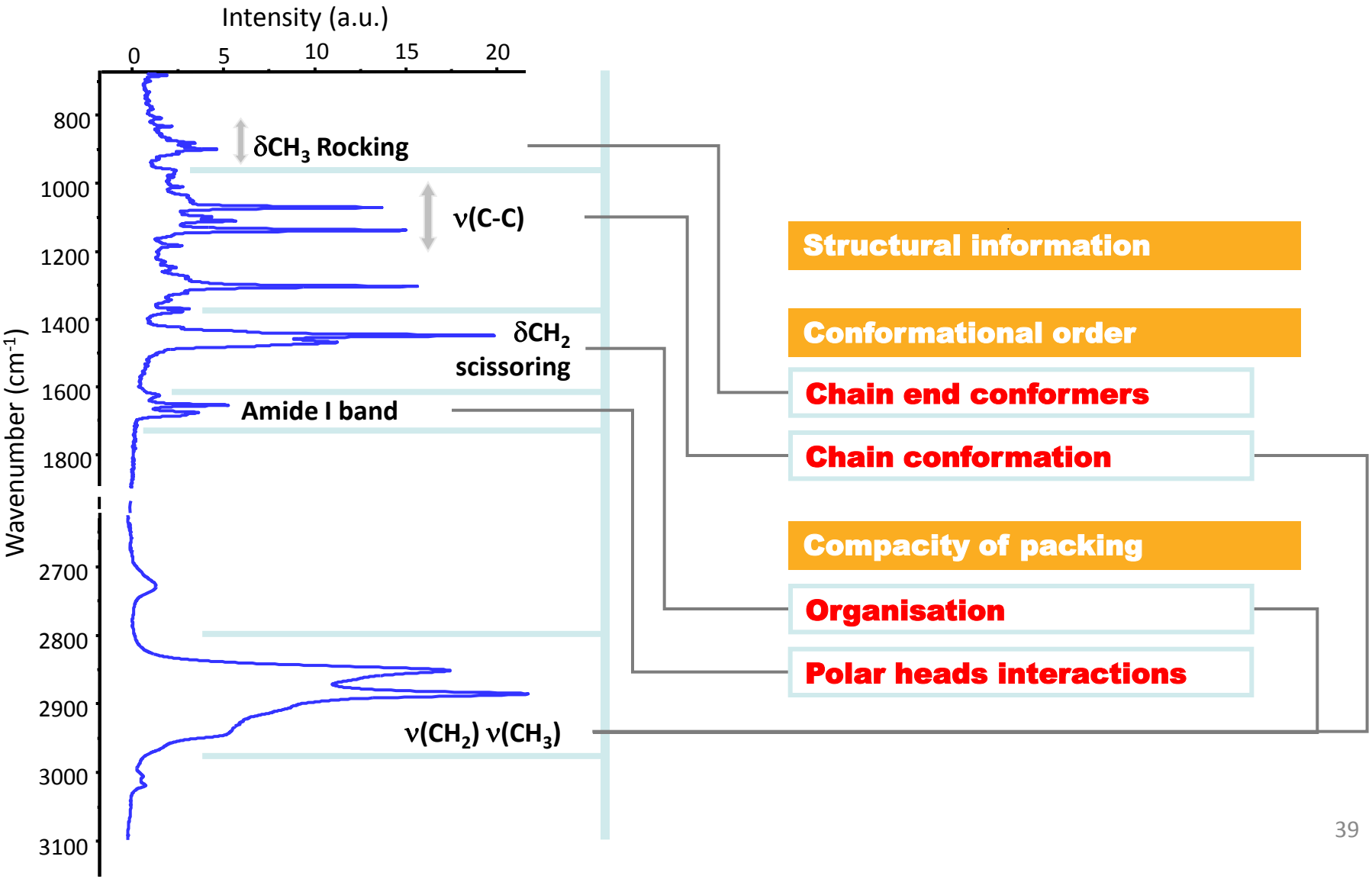
-Baseline correction

-Normalization

Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

Raman descriptors of SC barrier



Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

Skin aging

Ex vivo et in vivo

Volontaires

13 F & 7 M (22 à 64 ans)

Méthode

Surface nettoyée

1ère acquisition *in vivo*

Extraction des lipides

2ème acquisition *in vivo*

In vivo

Spectres *in vivo* avant l'extraction

–

Spectres *in vivo* après l'extraction

=

Signal in vivo des lipides

Ex vivo

Pas de séparation

TFAYLI A. *et al.* **EJD 2012**

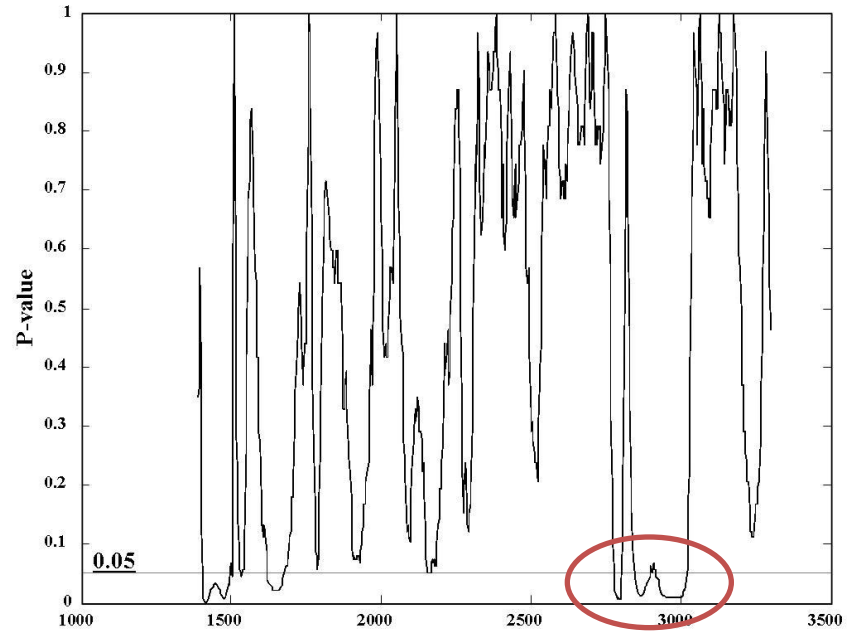
Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

Skin aging

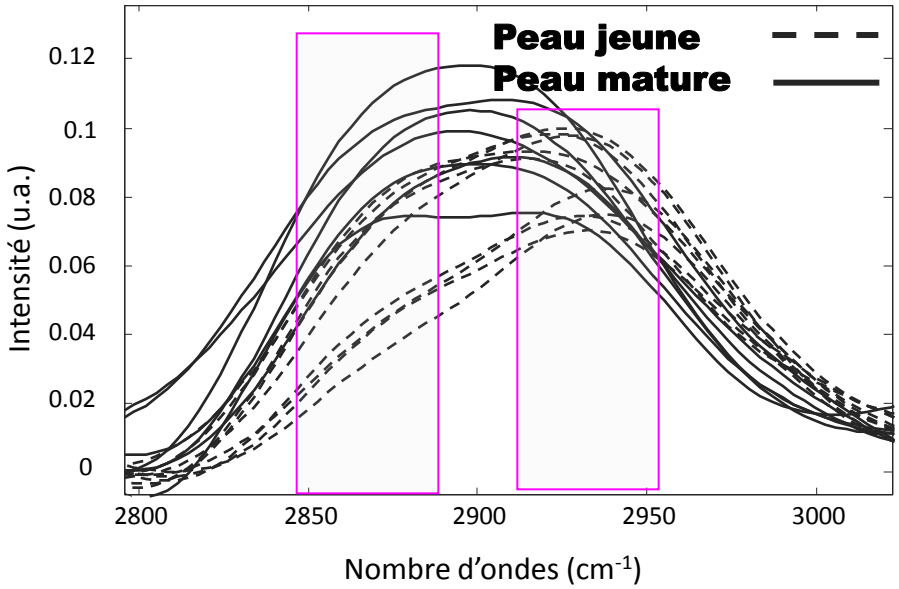
in vivo

kruskal Wallis



2845-3020 cm^{-1}

Observations directes



TFAYLI A. *et al.* **EJD 2012**

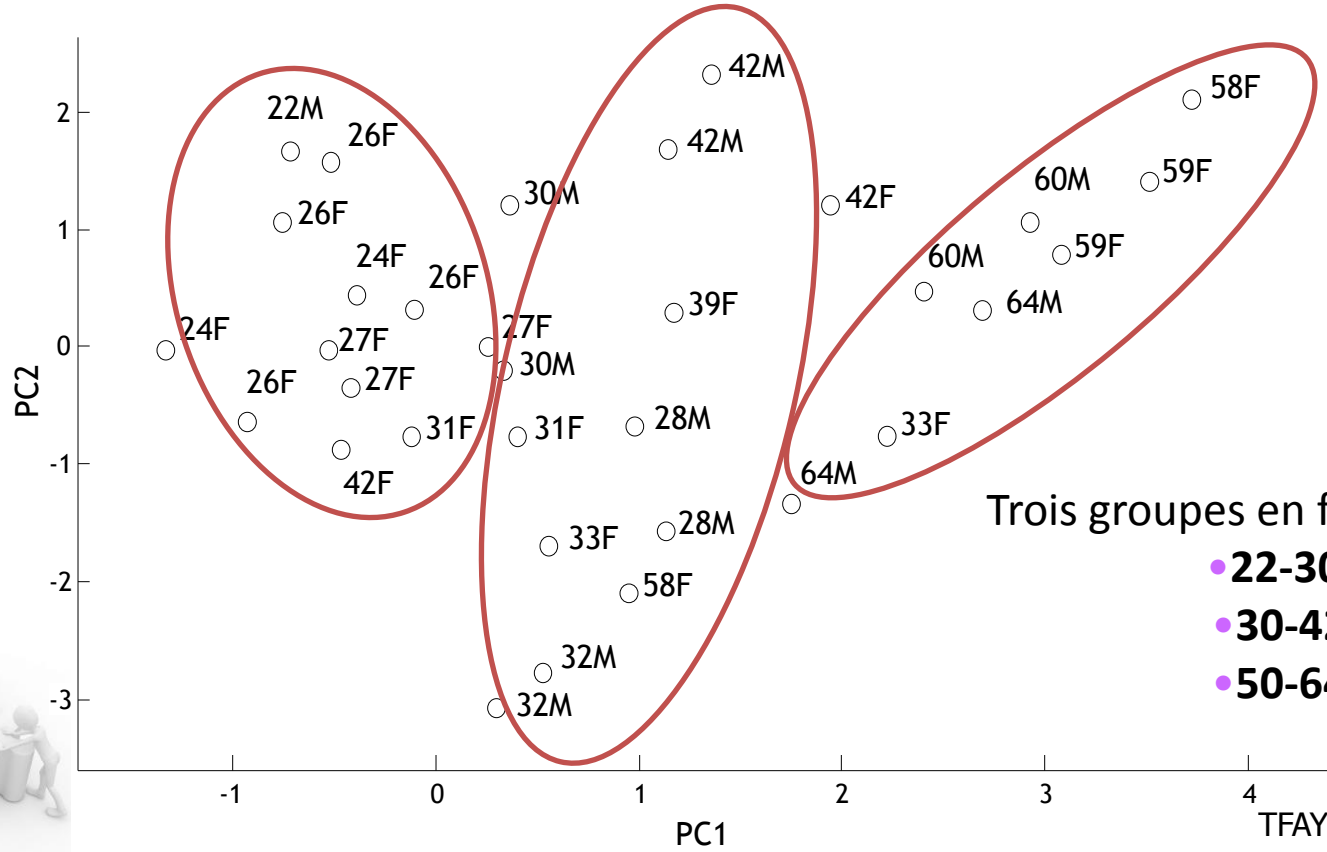
Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

Skin aging

in vivo

Analyse en composantes principales sur la gamme: 2845-3020 cm^{-1}



Trois groupes en fonction de l'âge:

- **22-30 ans**
- **30-42 ans**
- **50-64 ans**

TFAYLI A. *et al.* **EJD 2012**

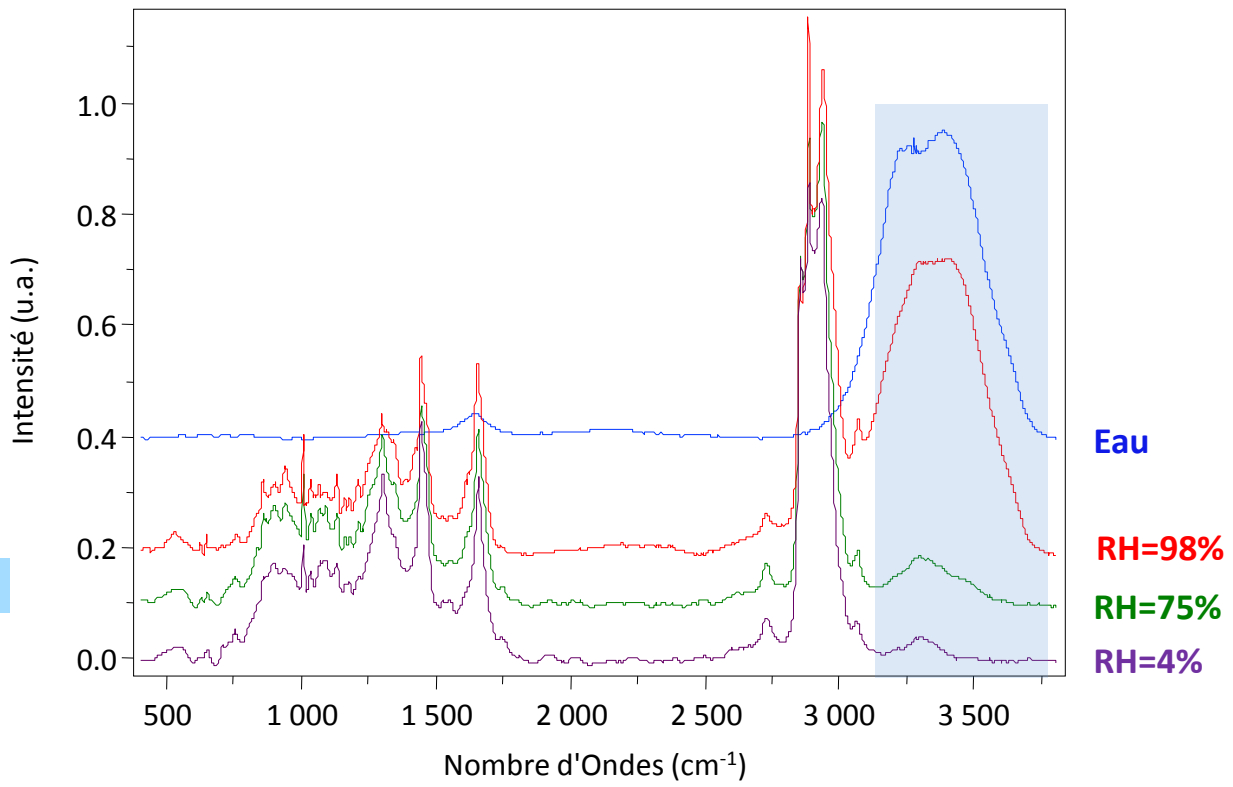
Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

Skin hydration / dry skin diseases

Ex vivo

Water structure



RH : 2.5% » 75%

VYUMVUHORE R. *et al.* **Analyst, 2013**

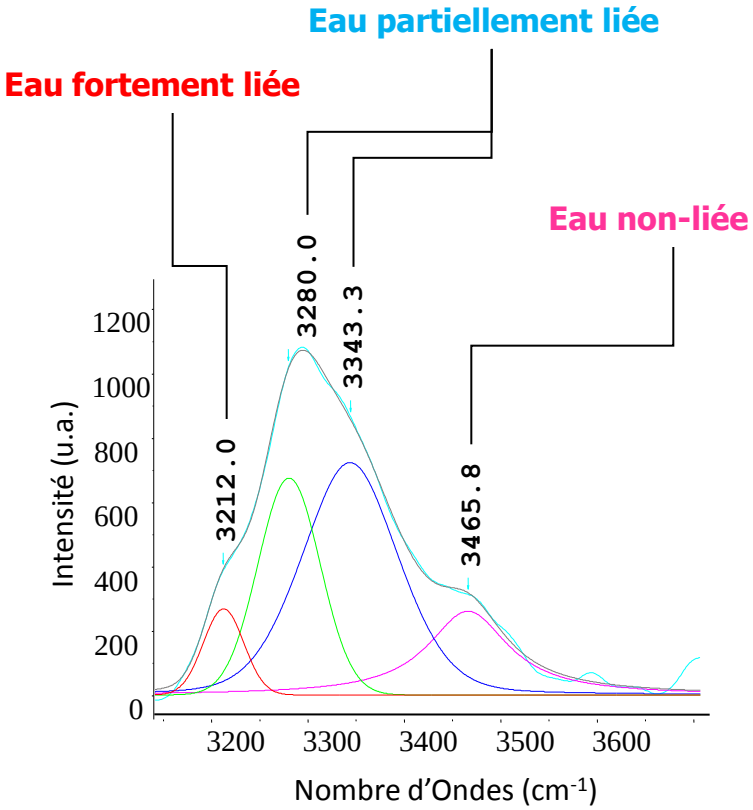
Vibrational spectroscopies VS barrier function

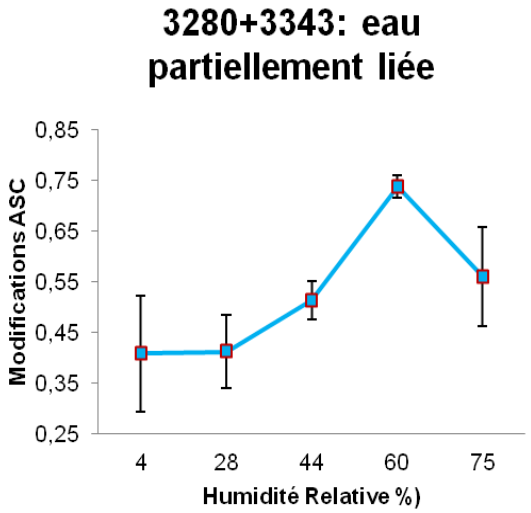
Physiological / Physiopathological status

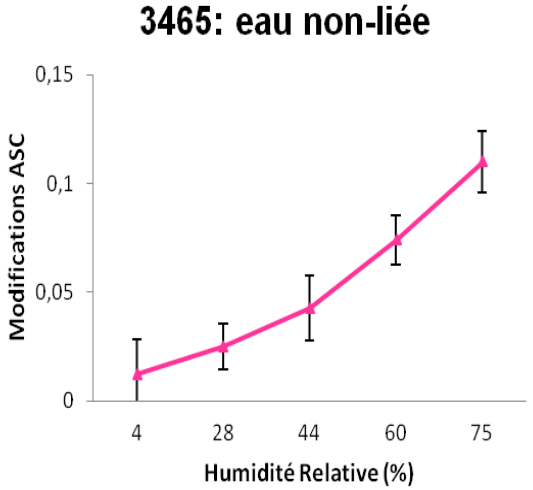
Skin hydration / dry skin diseases

Water structure

Ex vivo







VYUMVUHORE R. *et al.* **Analyst, 2013**

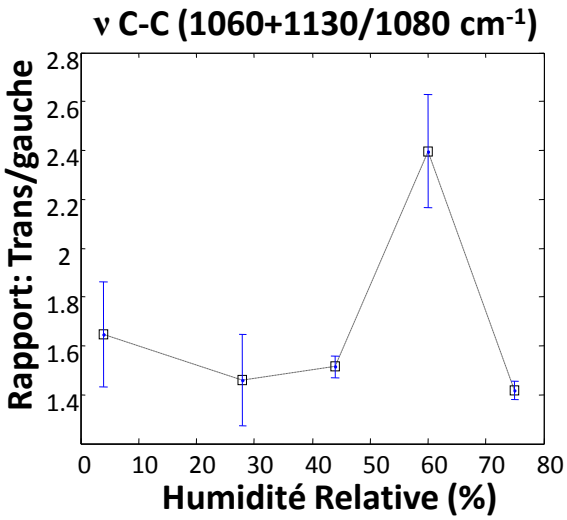
Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

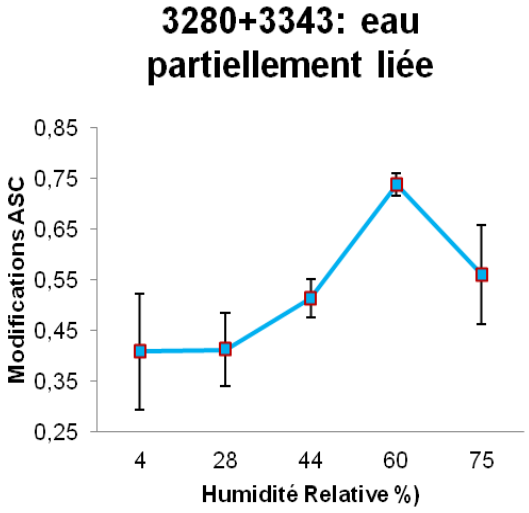
Skin hydration / dry skin diseases

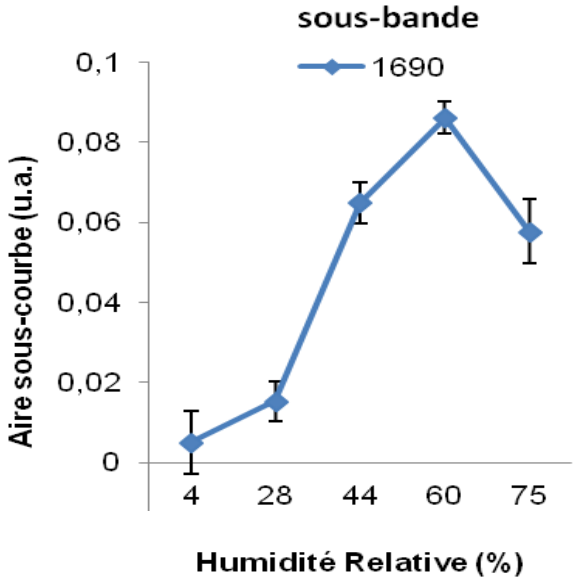
Ex vivo

Conformation des chaînes lipidiques



Structure secondaire de la kératine: bande Amide I





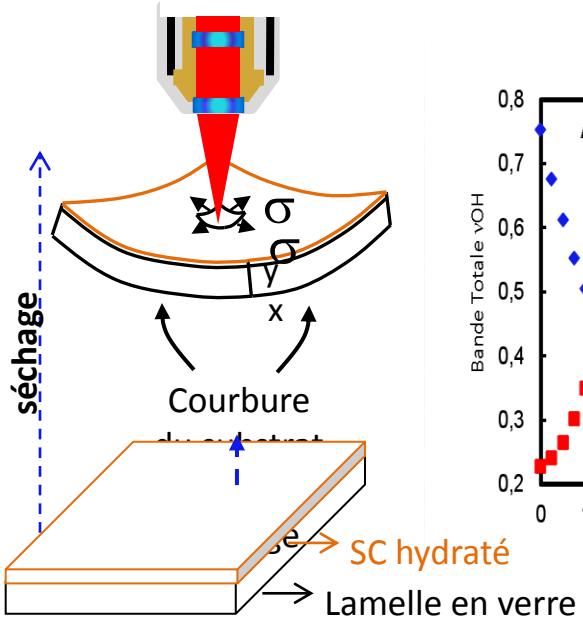
VYUMVUHORE R. *et al.* **Analyst, 2013**

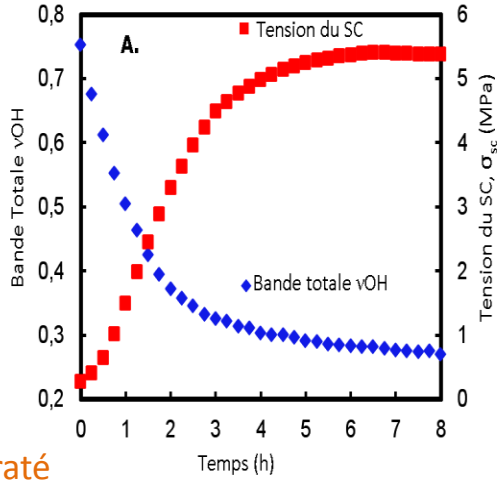
Vibrational spectroscopies VS barrier function

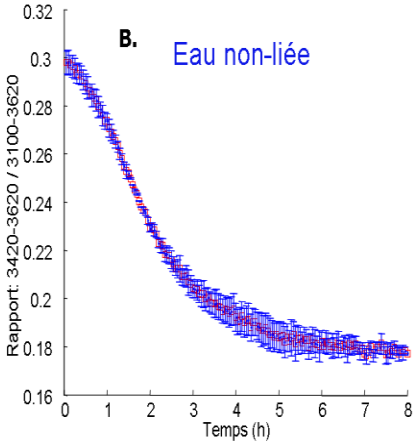
Physiological / Physiopathological status

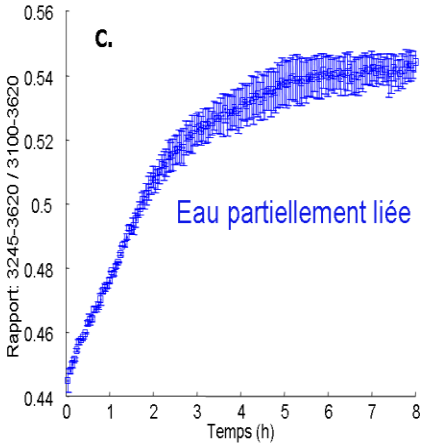
Hydration and mechanical stress

Ex vivo









Tension mécanique du SC ⟺ **Bande totale vOH**

VYUMVUHORE R. *et al.* **J. Biophotonics, 2014**

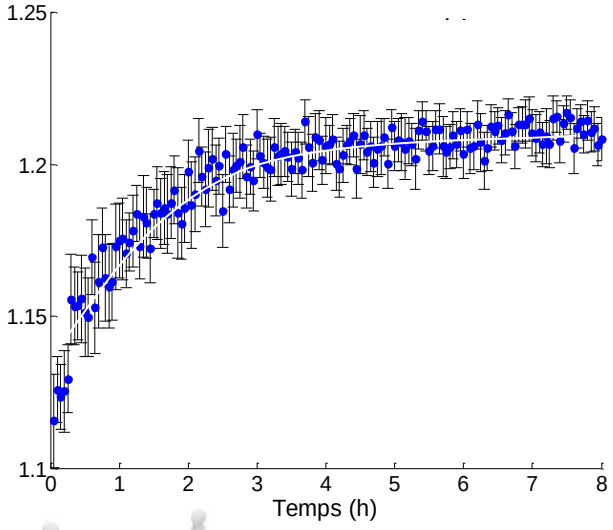
Vibrational spectroscopies VS barrier function

Physiological / Physiopathological status

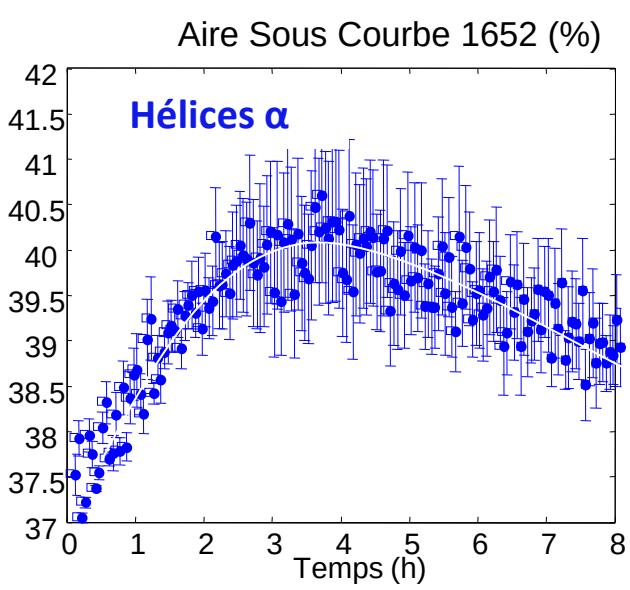
Hydration and mechanical stress

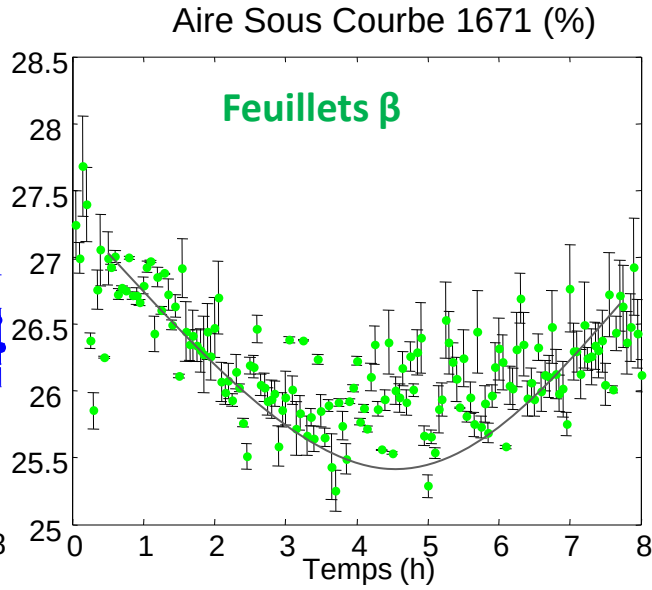
Ex vivo

Conformation des lipides



Structures secondaires des protéines

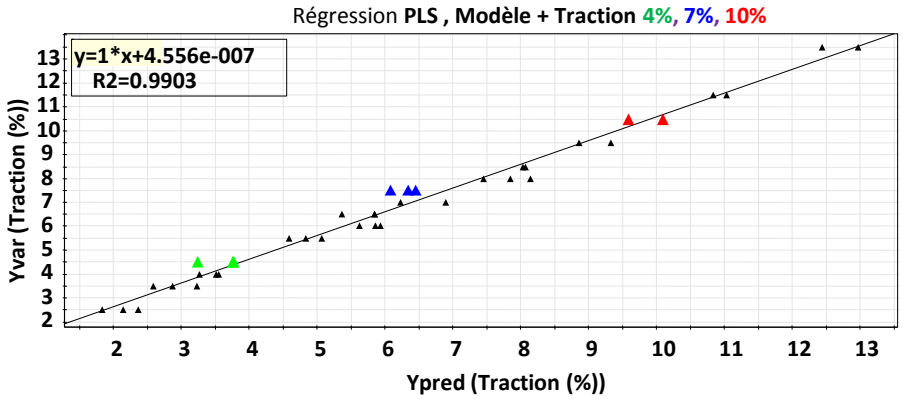




VYUMVUHORE R. *et al.* **J. Biophotonics, 2014**

SC mechanical strains

Ex vivo



Analyse statistique multivariée

⇕

Pics discriminants

⇓

Interprétation structurale

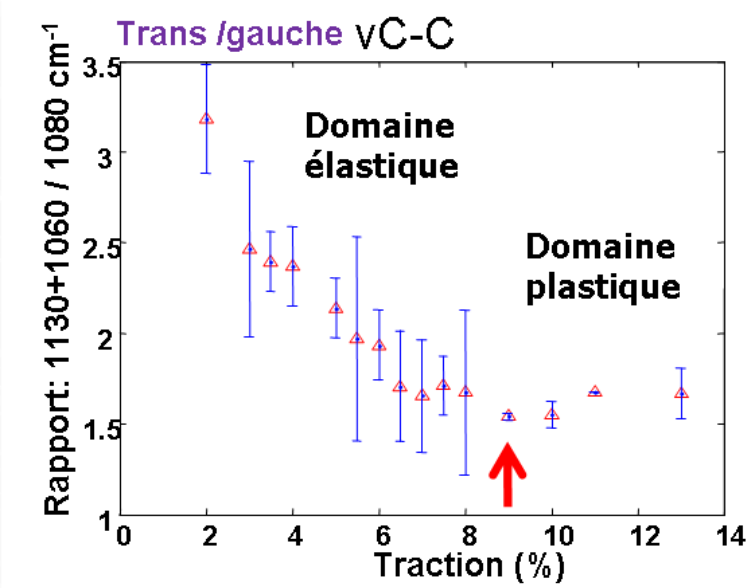
⇓

État moléculaire lié au stress mécanique

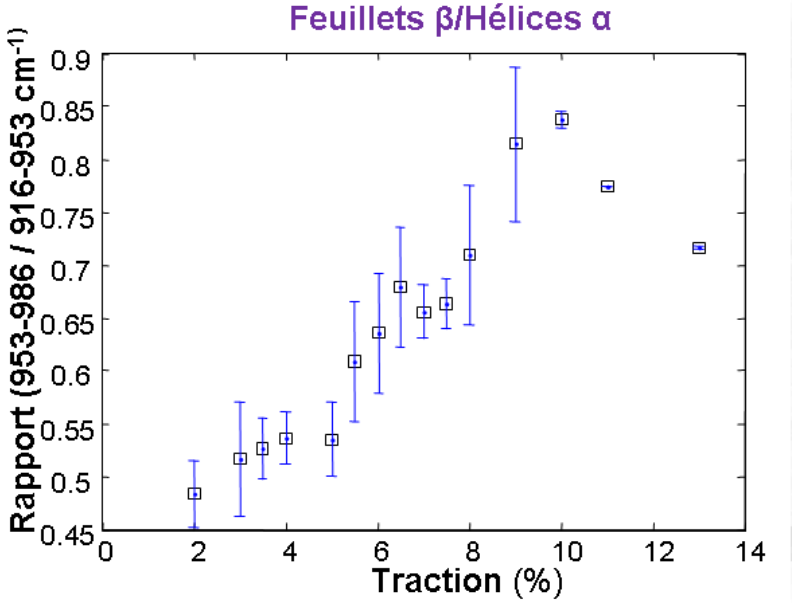
VYUMVUHORE R. *et al.* **J. Raman spectroscopy, 2013**

SC mechanical strains

Compacité des lipides



Structure secondaire des protéines



Deux domaines

zone élastique

zone plastique

VYUMVUHORE R. *et al.* **J. Raman spectroscopy, 2013**

Physiological status

"QR code" of the skin

ANR CARE

In vivo

Analyse multi-paramétrique du SC ⟺ **Patients**

pH	**: pH mètre**
PIE	**: Tewl-mètre**
Hydratation globale du SC	**: Cornéomètre**
Composition lipidique	**: Chromatographies**
Information moléculaire + profondeur	**: Raman**

11 volontaires F sains âgés de 57 à 62 ans

Analyse sur bras et mollet

1. **Identification des relations entre les différents paramètres du Stratum Corneum**
2. **Développement d'un outil multi-informationnel**

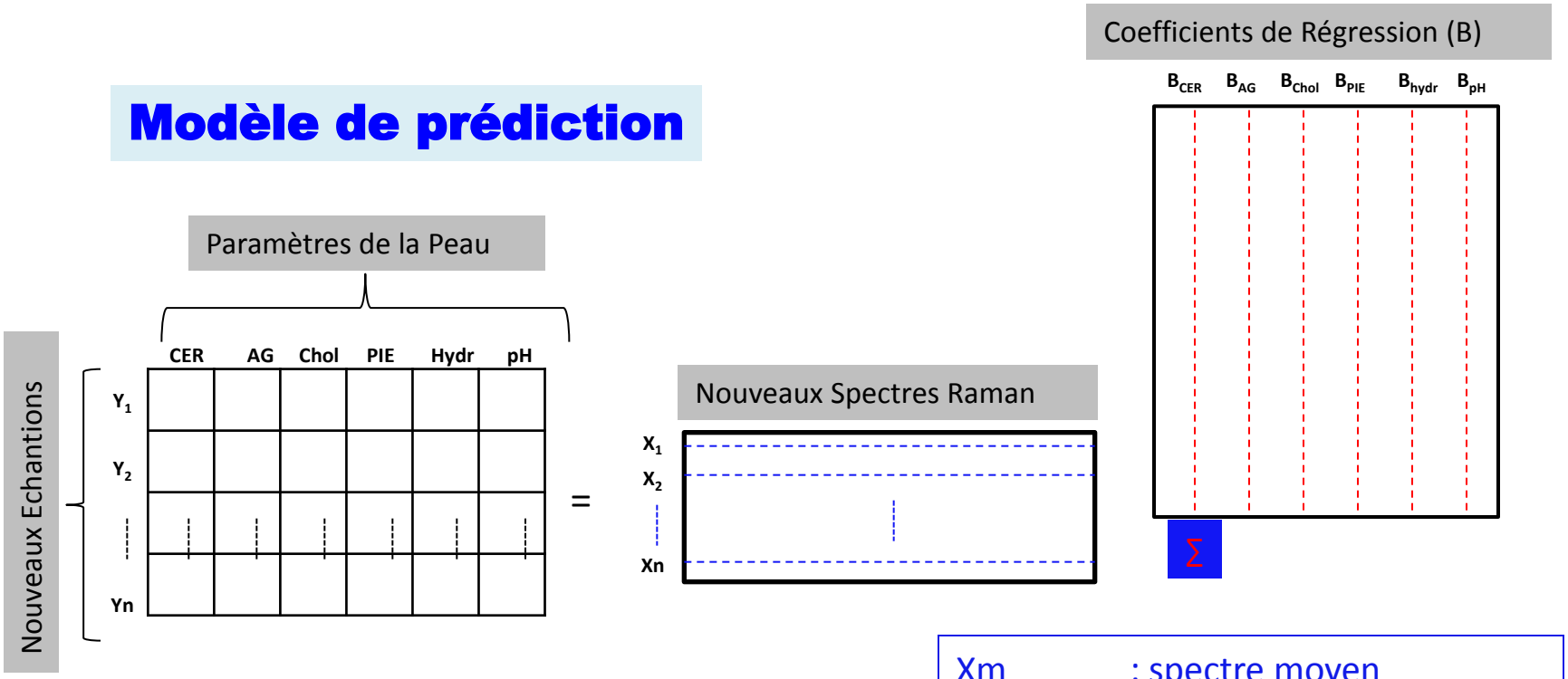
VYUMVUHORE R. *et al.* **JBO, 2014**

Physiological status

"QR code" of the skin

ANR CARE

Modèle de prédiction



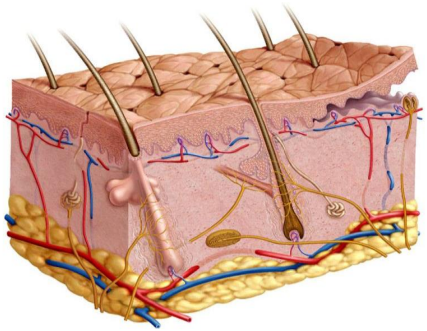
Données centrées-réduites; besoin de:

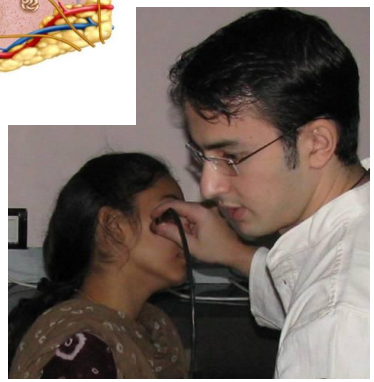
Xm	: spectre moyen
Xstd	: écart-type des spectes
Ym	: moyenne de Y
Ystd	: écart-type de Y

VYUMVUHORE R. *et al.* **JBO, 2014**

Raman et peau synthétique

Sources de peau





Peau humaine

Restriction

disponibilité

Manque de reproductibilité

Peau animale

Restriction **Interdiction**

Peau synthétique

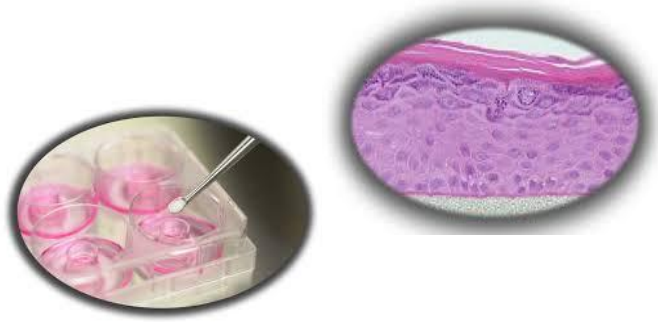
Culture de keratinocytes

Raman et peau synthétique

Evaluation des peaux

Validité en tant que substits

- Morphologie
- marqueurs biomécaniques
- Tests d'irritation
- Tests de phototoxicité
- Composition protéique
- Composition lipidique
- Perméabilité



Les classes lipidiques majoritaires sont présentes

Perméabilité PLUS ÉLEVÉE

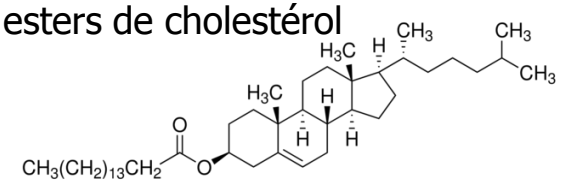
Comparaison de la composition et de l'organisation des lipides

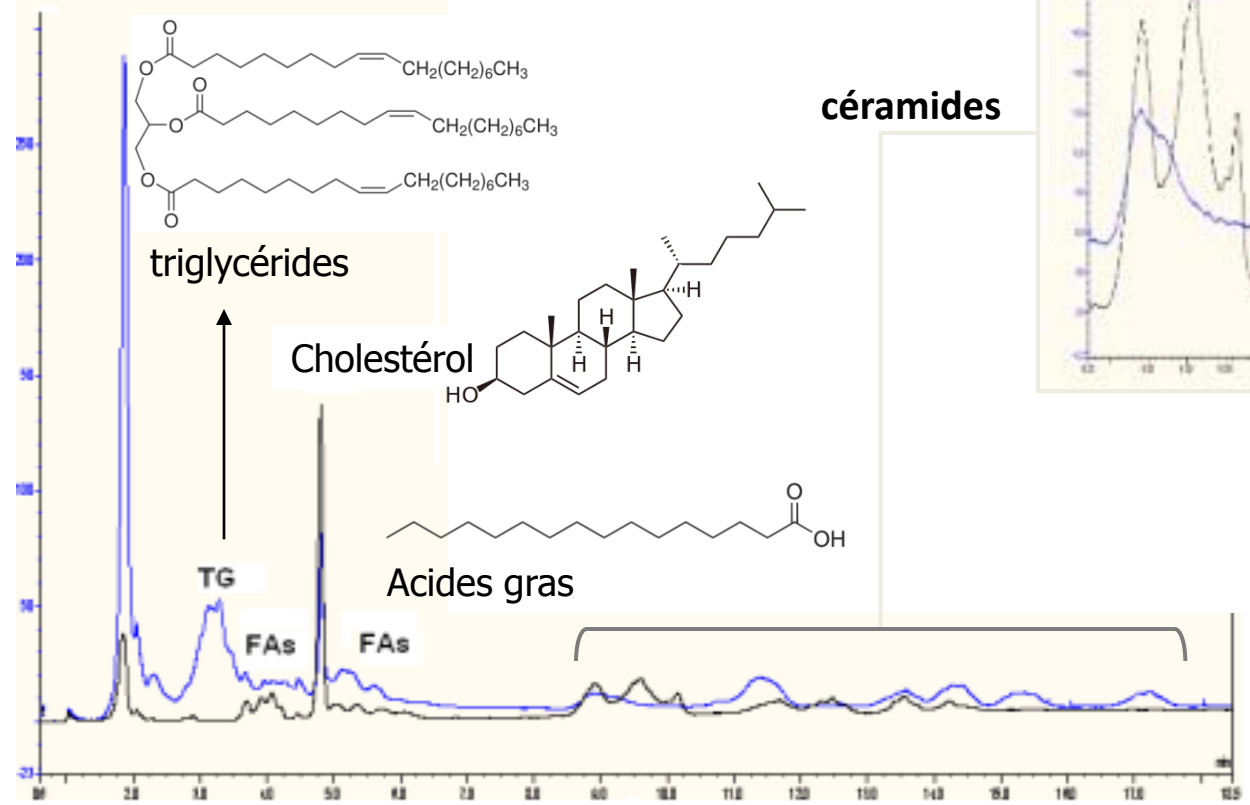
Raman et peau synthétique

Composition des lipides

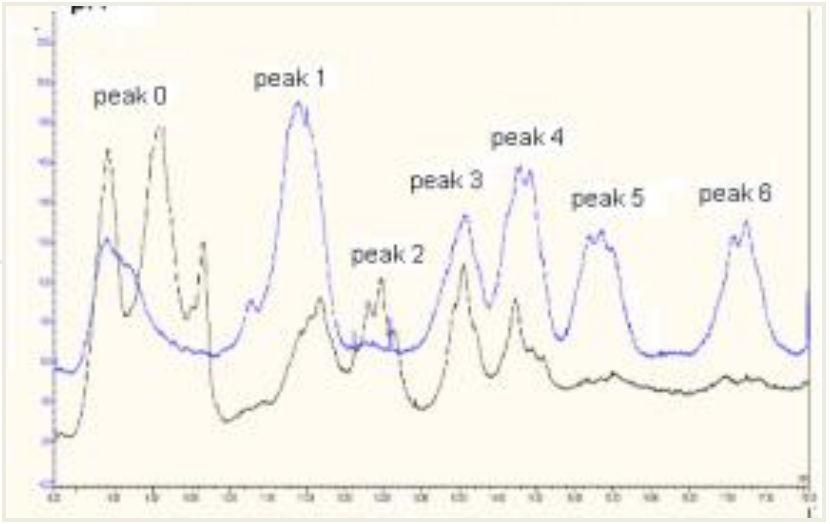
Composition globale

esters de cholestérol





céramides



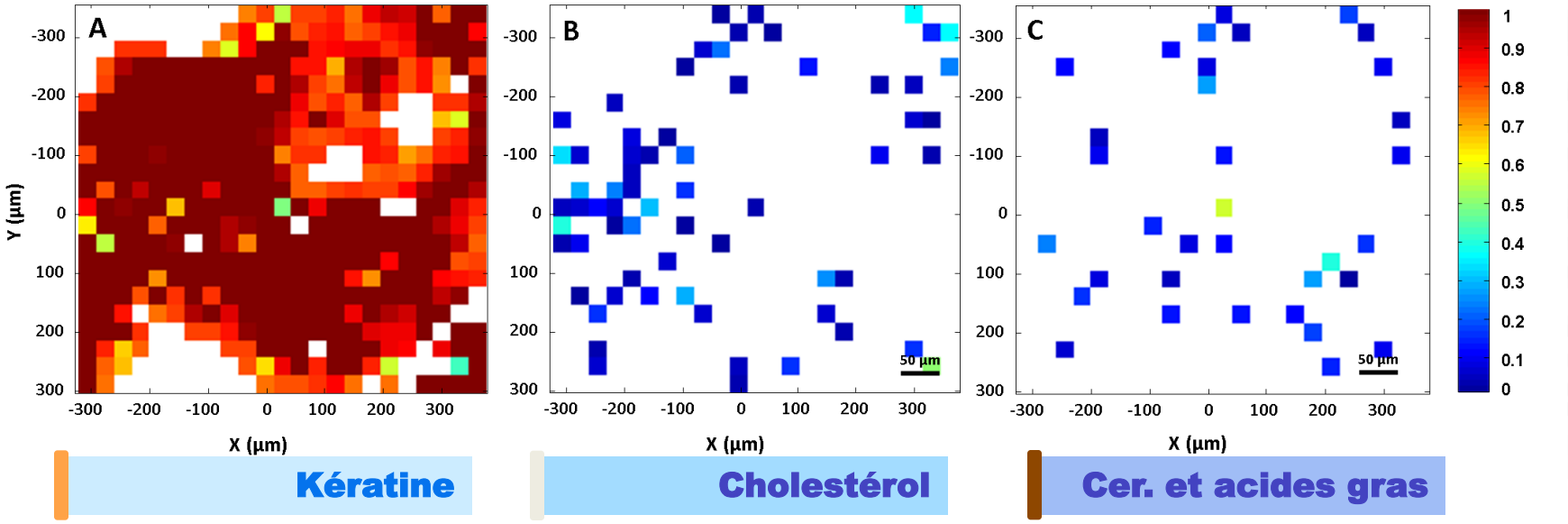
Peau humaine: bleu

Peau synthétique: noir

Raman et peau synthétique

Imagerie Raman + NCLS

Stratum corneum reconstruit



Kératine | **Cholestérol** | **Cer. et acides gras**

-Taille de l'image: 600*600 µm²
-Taille du pixel: 4 µm
-Pas: 20 µm

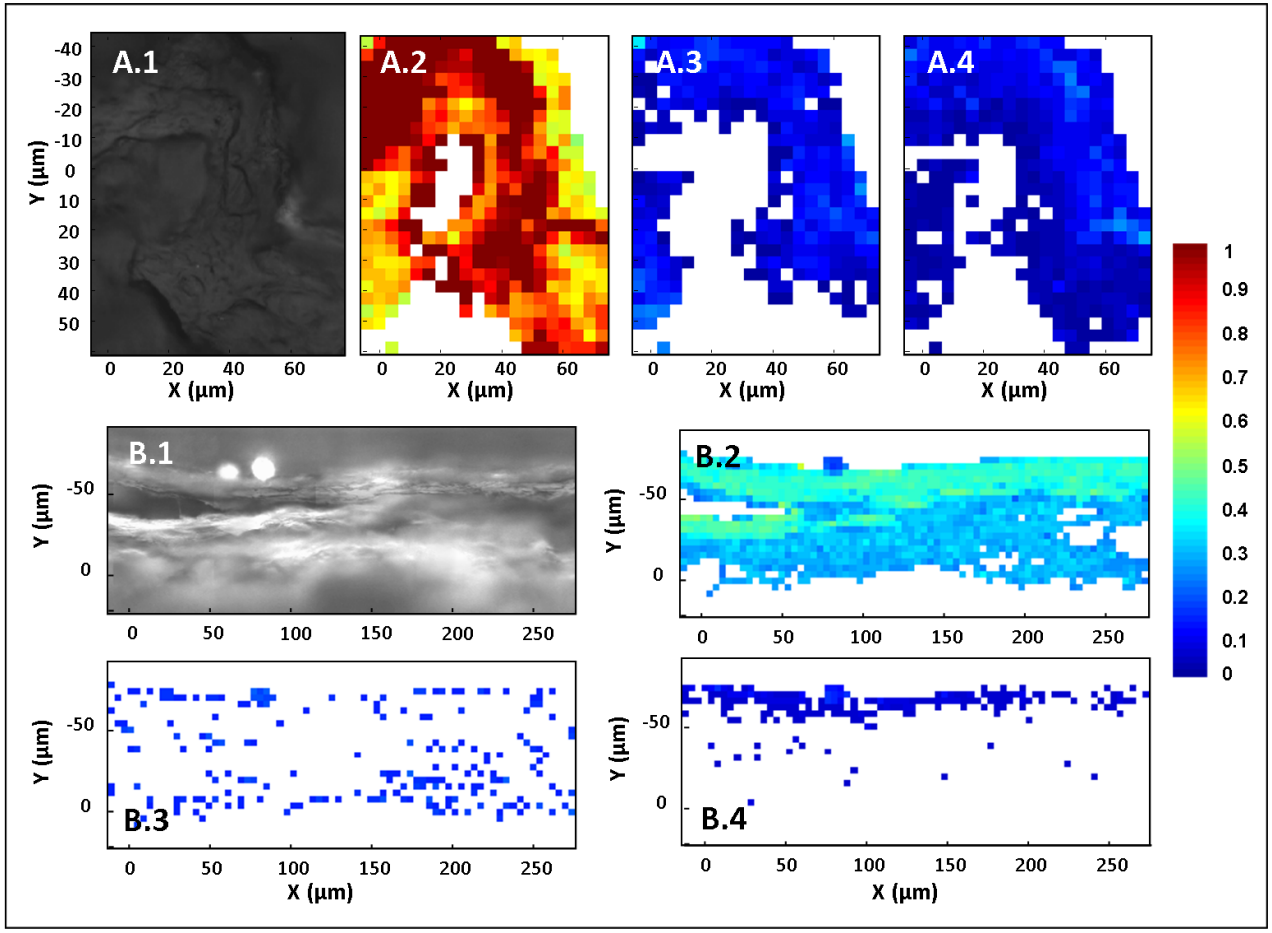
Distribution hétérogène des lipides

Raman et peau synthétique

Coupes de SC: imagerie Raman + NCLS

Stratum corneum humain

Stratum corneum synthétique



1. Lumière blanche
2. Kératine
3. Cholestérol
4. Cer. et acides gras

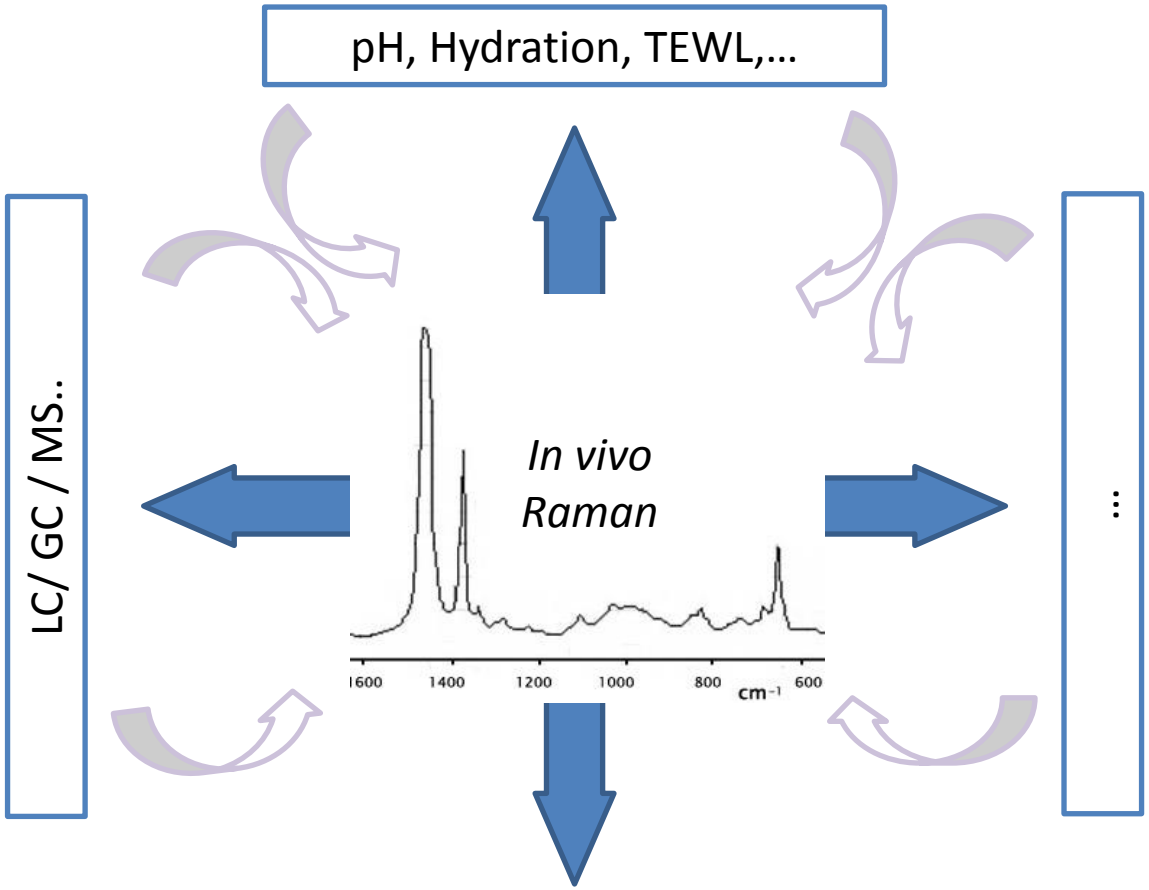
GO RAIBH MAITH AGAT

شكرا

Analytical chemistry: from data (pre)-treatment optimization to data mining, data fusion and big data

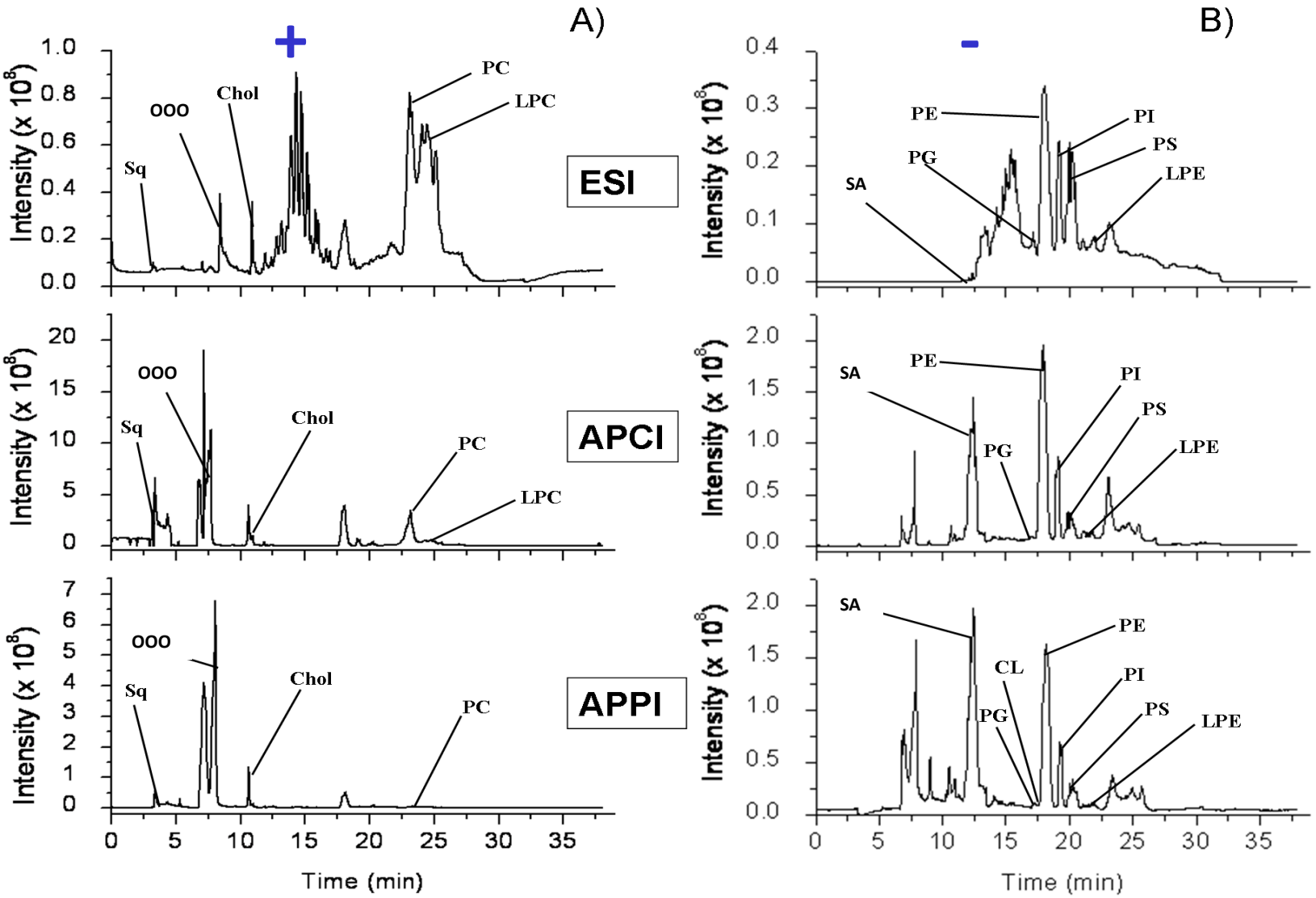
Sana et Ali Tfai(y)li

Raman analyses: correlation with LC/GC/MS.., biometric data



Physiopathological state: Diagnosis

LC-MS acquisitions: Different ionization modes (Data fusion?)



Laurent Imbert, PhD thesis, GCAPS,EA4041, 2012, Univ. Paris Sud

Data fusion and data mining

Perspectives:

Data fusion between:

- RPLC and NPLC in chromatography
- Different ionization modes in mass spectrometry
- Increase the separation dimensionality LCxLC MS… (new treatment approach)
- Between Separation techniques, coupled mass spectrometry
- Between several techniques: Raman, IR, separative techniques
- Multi-block analysis (specific algorithms for data processing and fusion)?

Additional analysis will increase the time for data processing: other approaches for data processing ?

Virtual data project (LAL): work on a cloud and save the image
Buy cores (possible demand through a project in process)

Data storage

Perspectives:

DATA ARE NOT CENTRALIZED.

- centralize data
- Data storage and management
- Results storage and management
- Generate our own databases