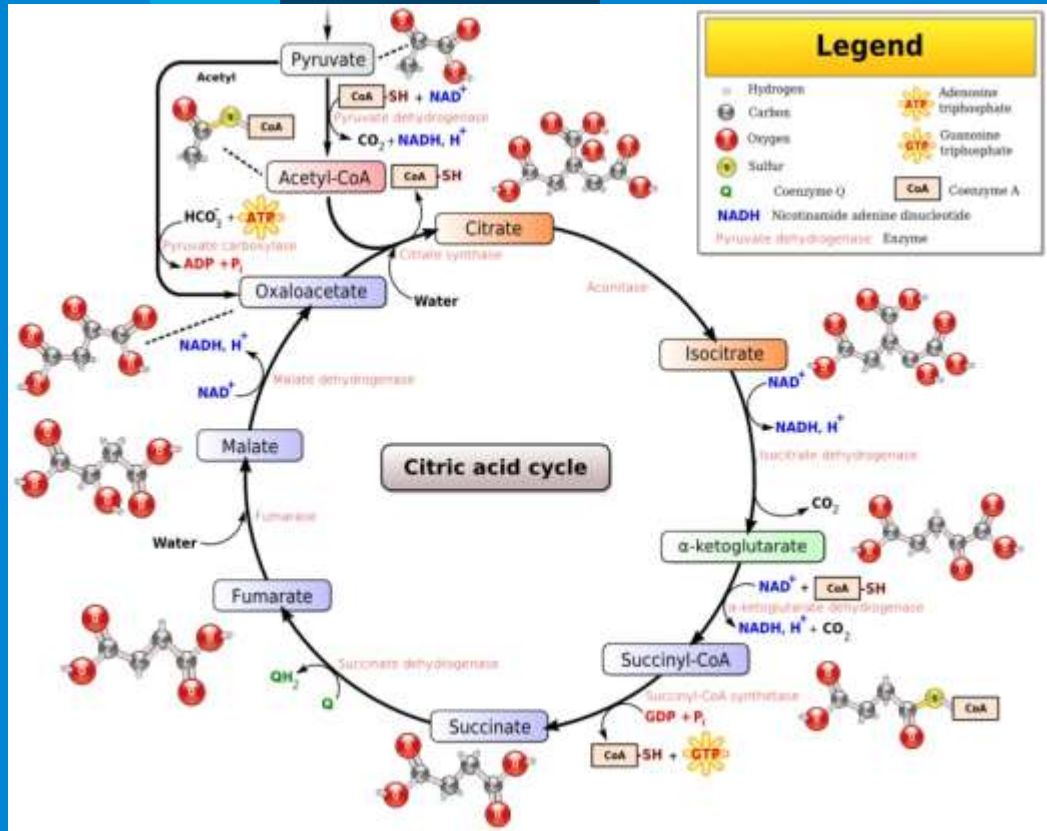
Herramientas y Soluciones en Ciencias -Omicas y perfilado

AGILENT SEMINAR

Universidad de Zaragoza

Jaume C. Morales
Iberia LCMS Product Specialist
Agilent Technologies



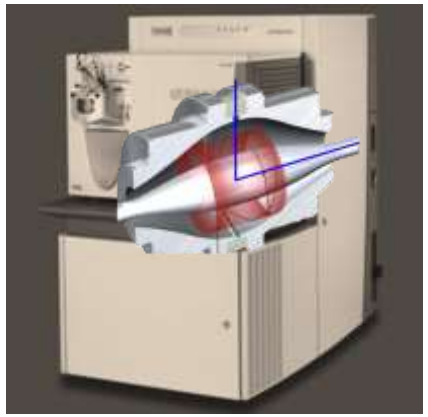
Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - *Preparación de datos previa al Análisis Estadístico diferencial.* ***Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder****".*
  - *¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil….* ***Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

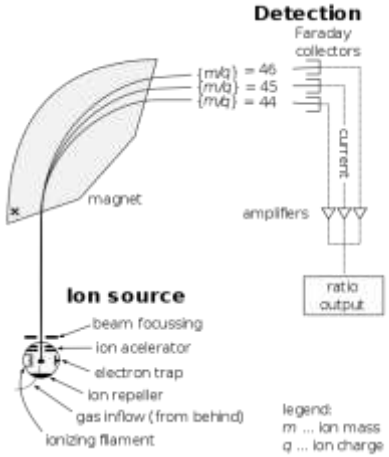
Replica of J. J. Thomson's third mass spectrometer.

Cyclotron



Orbi

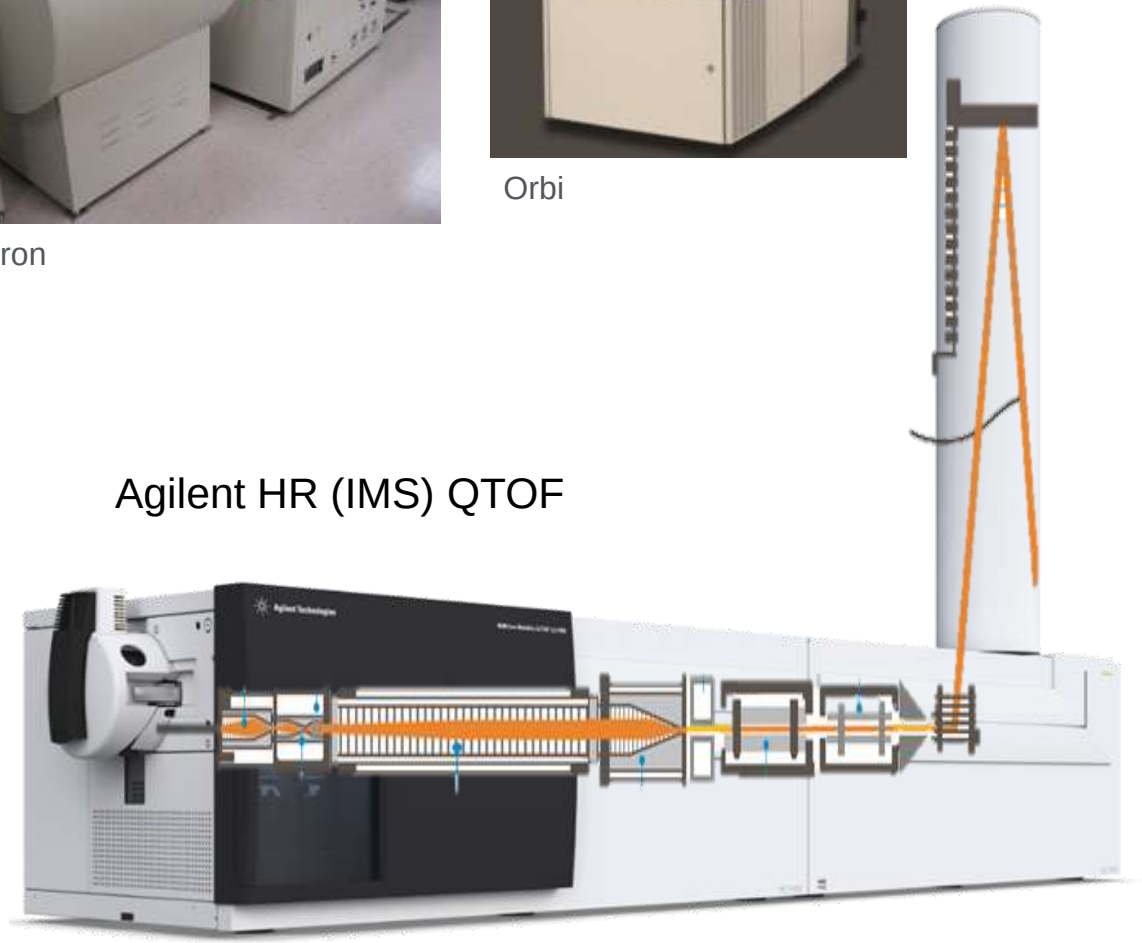
From 54th ASMS Conference on Mass Spectrometry



Magnetic Sector

Single quadrupole mass spectrometer used for John Fenn's Nobel Prize winning work on electrospray ionization

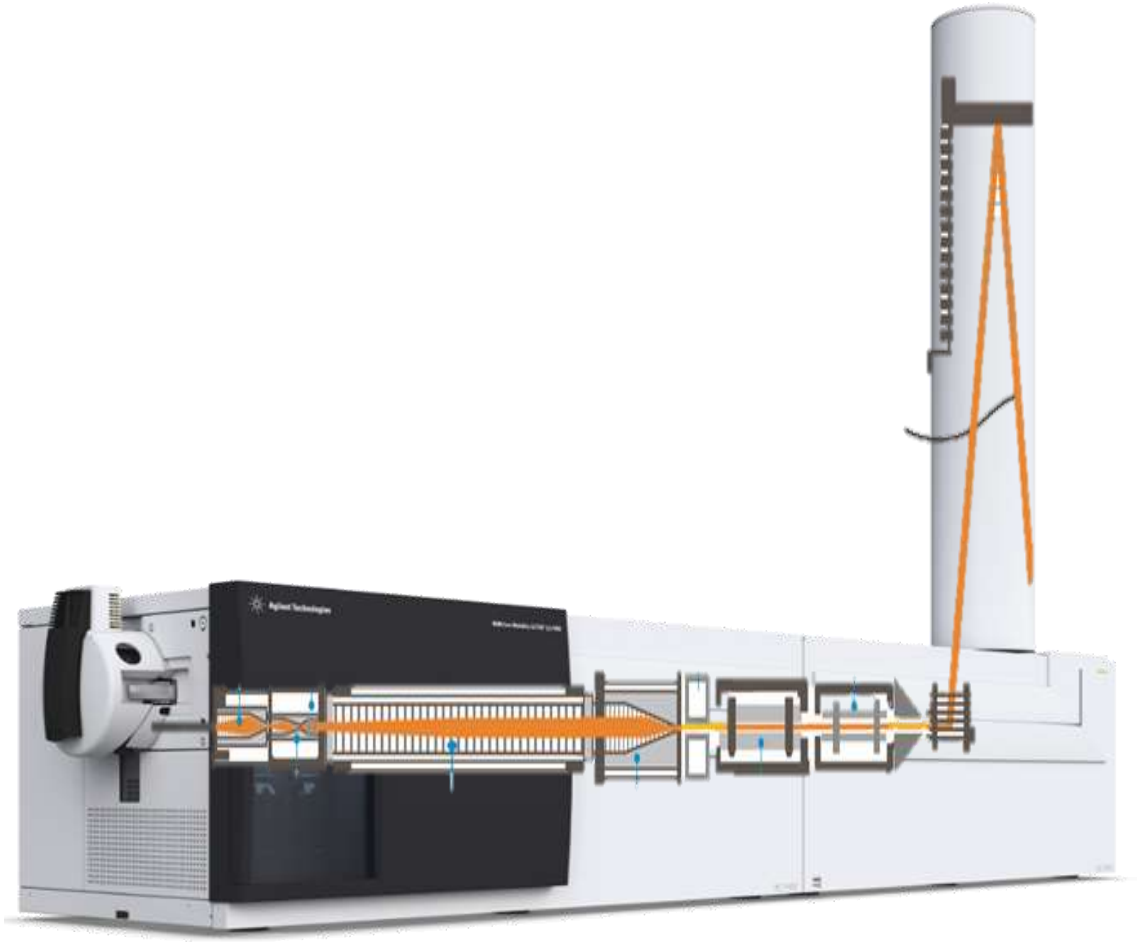
Agilent HR (IMS) QTOF



Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

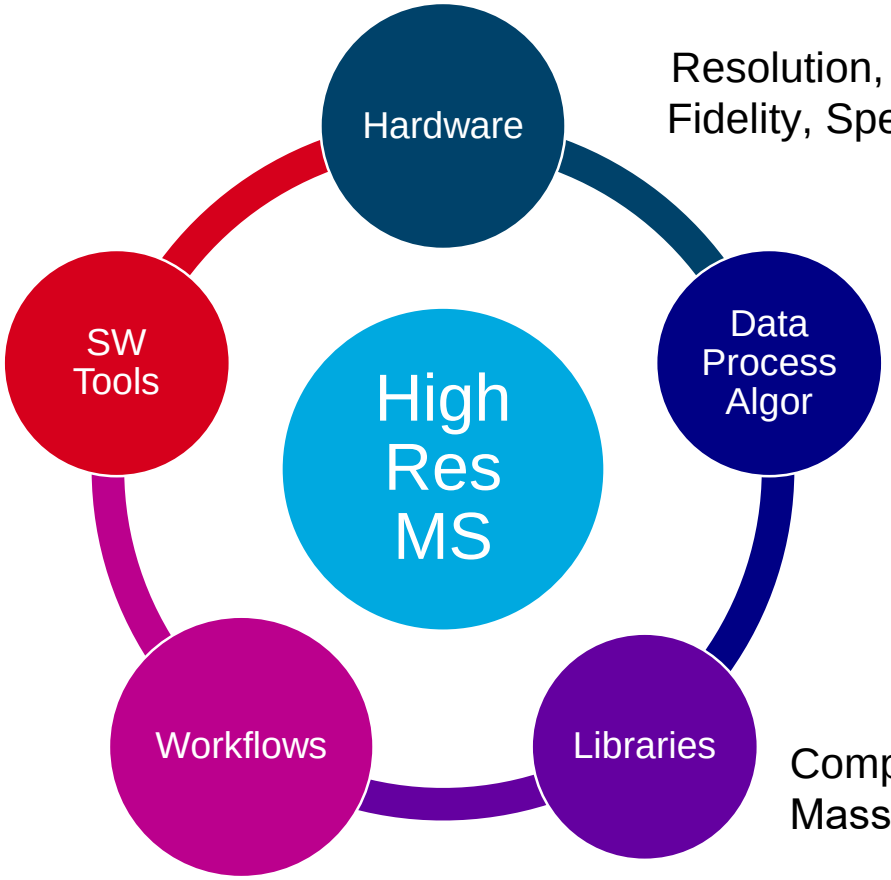
HRMS is absolutely differential MS technology. It allows for :

- Untarget Screening
- Target Screening
- Confirm Suspect compounds
- Sensitive Full Scan analysis
- Structural elucidation
- Identify Unknown Compounds
- -Omics disciplines (Metabolomis, Proteomics, …..)
- Profiling of samples
- Materials studies



Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

To work with HRMS is more than an instrument



Resolution, Accurate Mass, Isotopic Fidelity, Speed, Sensitivity, Dynamic Range….

Acquisition Modes and Data Processing Algorithms fit for purpose

Comprehensive ID Libraries with Exact Mass & MS/MS Spectra… “CCS”

Workflows to fast and easy performance

SW tools to identify Unknowns by Structure

Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

Agilent HRMS is based on Time Of Flight technology (TOF)

TOF is just an stop watch measuring time ions take to arrive to detector once they have been shot up at **PULSER**.

Lighter ions arrive before and the heaviest, later.

Reflectron optics provides larger flight path increasing resolution.

Time of flight is calibrated with known compounds (Tuning Mix) so time of ions contrasts with a calibration curve of t <-> m / z. This allows to know m/z with **High accuracy**.

On top of that, Spectral Peak Width is very narrow so we can differentiate very close spectral bands like Isotopic Pattern with High Fidelity.

In general, accurate mass instruments are those who can provide an error mass of < 5ppm.
MS systems based on Quadrupoles have mass error measurements of about > 150ppm.

Agilent systems can provide Mass error <1ppm or bellow

Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

RESOLUTION vs Mass Error

Spectral Resolution allows to get low mass error, on top of other advantages.

Unfortunately Mass error is not proportional to Resolution :

Other Fourier T. systems	Res. ~ 200.000	Mas error <1 ppm
QTOF systems	Res. ~ 60.000	Mas error <0.8 ppm

To identify/confirm a compound through HRMS we do not only use Mass error.

Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

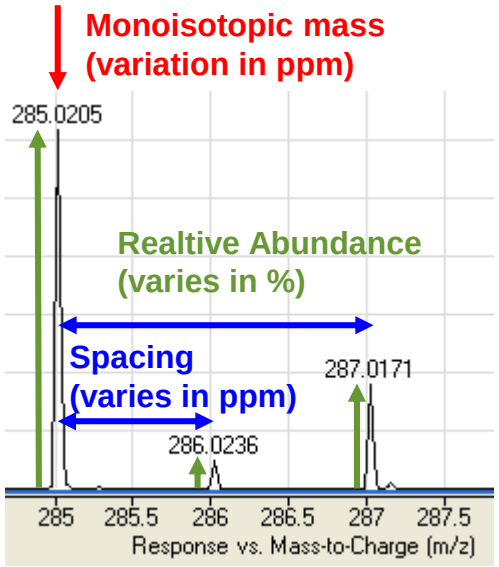
When we identify/confirm a compound through HRMS we do it by :

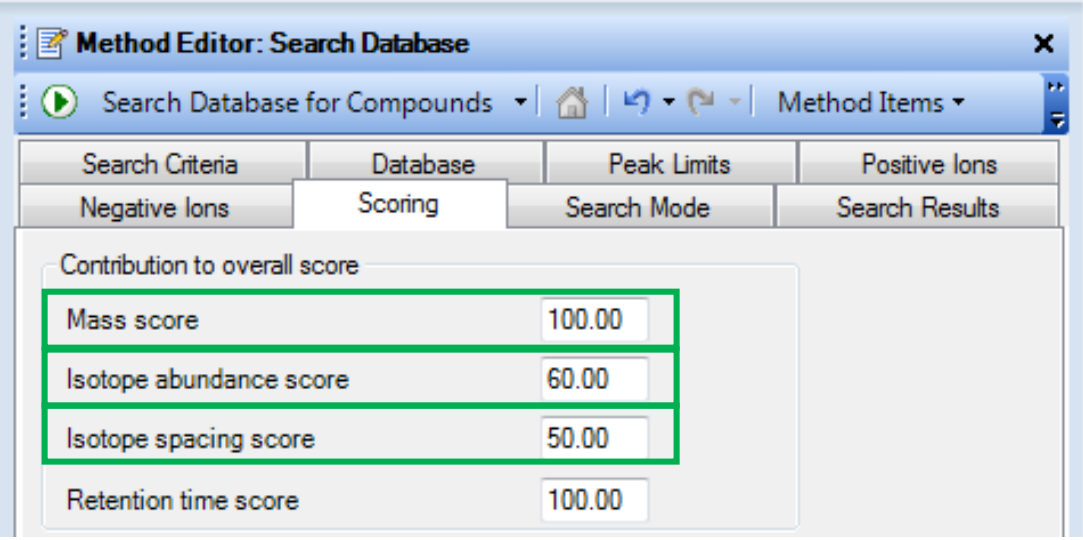
1. **Mass error**
2. **Isotopic Pattern**
3. **MS/MS Spectra**
4. **Rt**
5. **CCS (*IMS*)**

Spectral Resolution allows to get low mass error, on top of other advantages. <u>BUT we don't identify with Resolution itself.</u>

Isotope distribution = isotope ratio accuracy
→ Needs to be <5% even at good mass accuracy to reduce the number of potential database hits or empirical formulae

Scoring based on





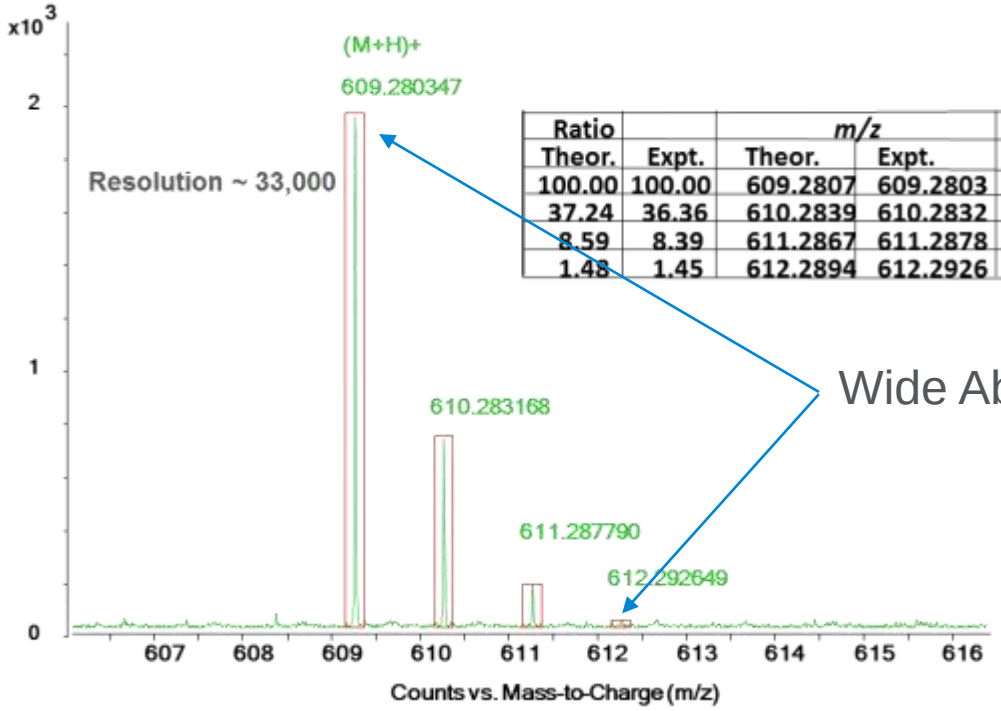
Technical features and Analytical Approaches with LC High resolution Mass Spectrometry.

SO everybody knows HRMS provides very low Mass error measurement BUT…

What other features are important on a HRMS and Why?

- Isotopic Fidelity. Better ID confidence
- Speed. UHPLC & MS/MS Coverage
- Sensitivity.
- MS/MS. Better ID confidence, Diff. Modes
- Dynamic Range. Better ID confidence
- **All at the same time!!!**

Dynamic Range of > 100.000 allows for Better ID



Ratio		m/z		Diff.
Theor.	Expt.	Theor.	Expt.	(ppm)
100.00	100.00	609.2807	609.2803	-0.51
37.24	36.36	610.2839	610.2832	-1.25
8.59	8.39	611.2867	611.2878	1.77
1.48	1.45	612.2894	612.2926	5.29

Wide Abund. Range

In-spectrum Dynamic Range: >4 Orders

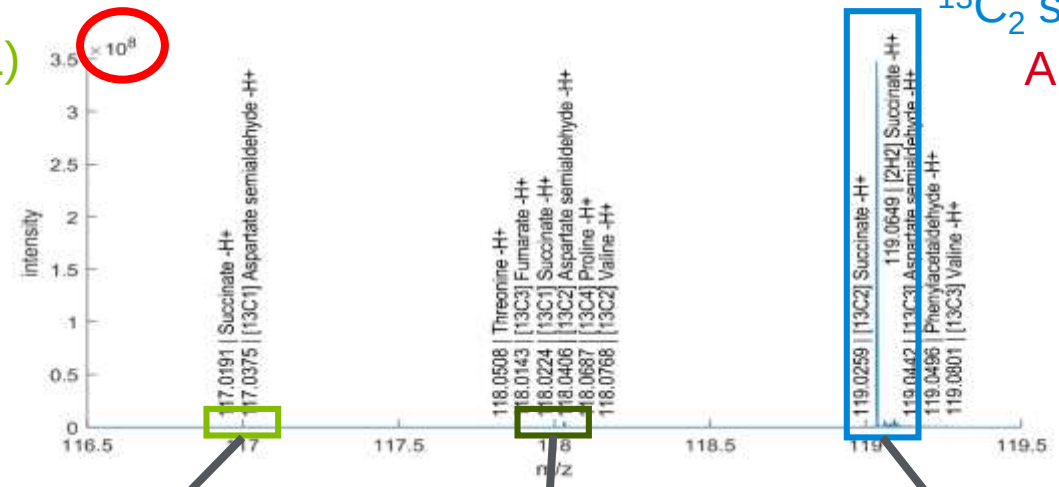
Stable label $^{13}C_2$ succinate (1µM) spiked into E. coli extract

succinate (*m/z* 117.0191)

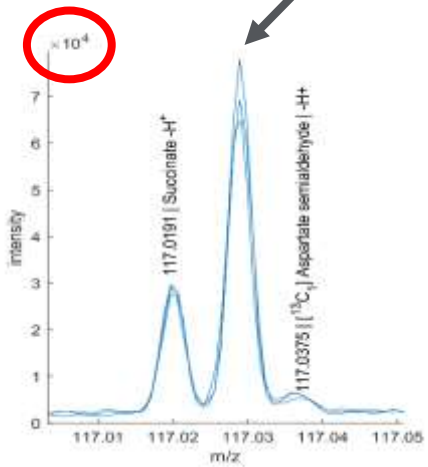
Abundance ~ 3 x 10^4

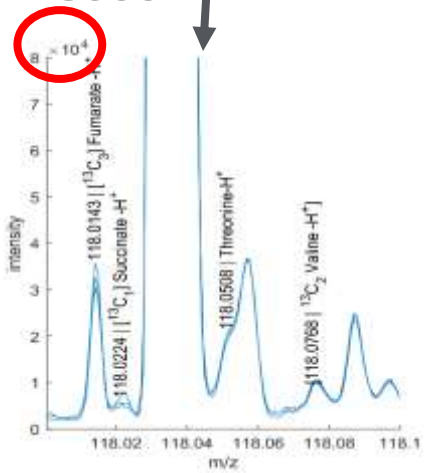
$^{13}C_2$ succinate (*m/z* 119.0259).

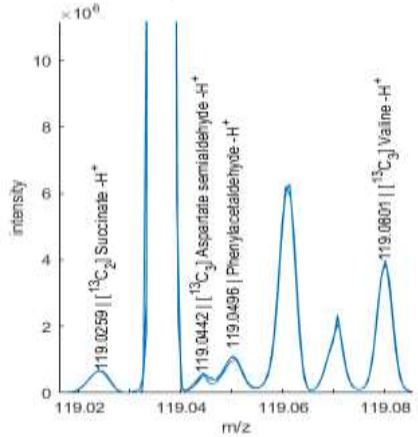
Abundance ~ 3 x 10^8



magnification: **5000x** **5000x** **100x**







HRMS QTOF Approaches
The Rumsfeld Conundrum

… *“as we know, there are **known knowns**; there are things we know that we know. There are **known unknowns**; that is to say, there are things that we now know we don't know. But there are also **unknown unknowns** – there are things we do not know we don't know. “*

United States Secretary of Defense Donald Rumsfeld

Screening Definitions

Target Screening. - Four Dimensions of Identification (4D-ID) QQQ Emulation

Compounds identified through Accurate Mass, Retention Time, Isotope Pattern and Fragment Confirmation. The quant method is developed via MassHunter Qual and Libraries (PCDLs). Acquisition is in All Ions MS/MS. **Standards Used**.

Suspect Screening - Propose and Identify. WIDEST Screening Approach

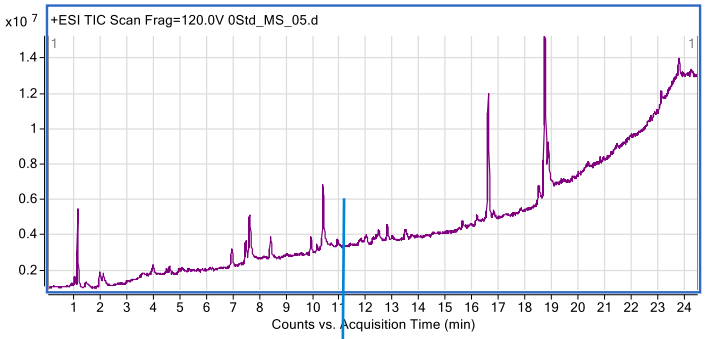
A Suspect List PCDL can be created from the encyclopaedic Master PCDLs. Compounds can be found and proposed using *Find-By-Formula*. The proposed list becomes the target or preferred (directed) list for target MS/MS or auto MS/MS. The resulting data can be searched against MS/MS spectra in the Suspect PCDL or Structure Correlated using MSC against the same PCDL if no MS/MS spectra are available.

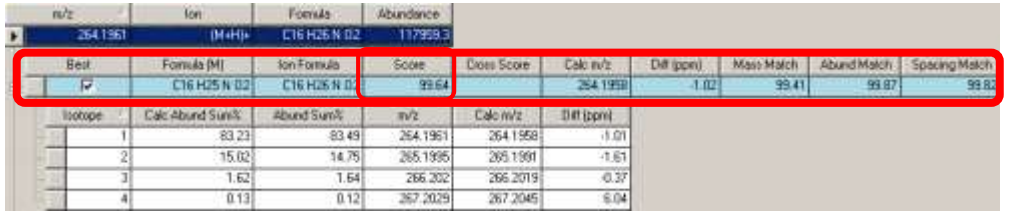
Discovery Profiling - Find, Compare and Identify

Profiling uses Molecular Feature Extraction (MFE) which is a naïve feature finding algorithm, in order to ***find*** **& characterize** compounds in a data set. These compounds can then be compared to determine the absence, presence or up- or down-regulation of the compounds from sample group to samples group. Profiling can be performed in single sample comparisons, two sample groups or multiple sample groups and conditions. Significant compounds can be imported into Acquisition in order to generate MS/MS data which is priceless in further compounds identification by formula generation, library search or structural correlation.

MassHunter QUAL : Find By Formula (MS)

Find by Formula (MS) análisis TARGET

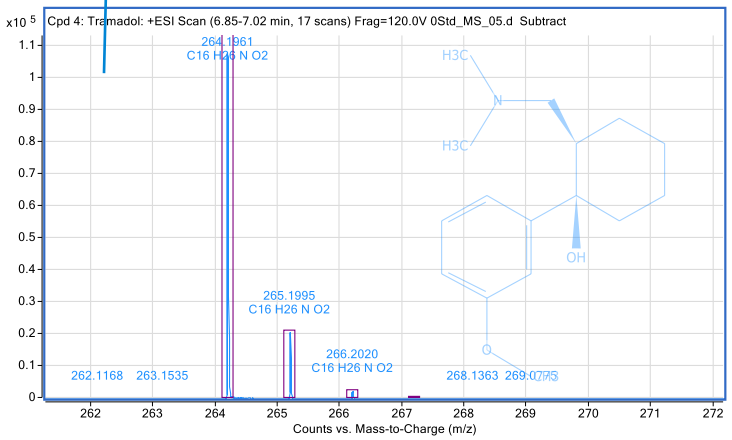




m/z	Ion	Formula	Abundance
264.1961	(M+H)+	C16 H26 N O2	117959.3

Best	Formula [M]	Ion Formula	Score	Cross Score	Calc m/z	Diff [ppm]	Mass Match	Abund Match	Spacing Match
☑	C16 H25 N O2	C16 H26 N O2	99.64		264.1958	-1.02	99.41	99.87	99.83

Isotope	Calc Abund Sum%	Abund Sum%	m/z	Calc m/z	Diff [ppm]
1	83.23	83.49	264.1961	264.1958	-1.01
2	15.02	14.75	265.1995	265.1991	-1.61
3	1.62	1.64	266.202	266.2019	-0.37
4	0.13	0.12	267.2029	267.2045	6.04



¿¿C16H25NO2 ??

Masa= 263.1885Da

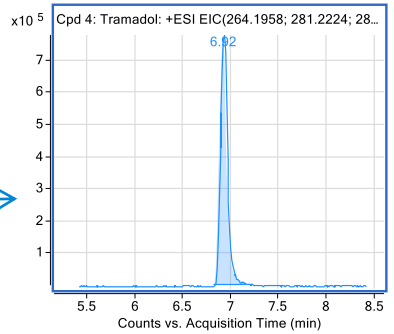
MassHunter Automático

Iones posibles :

H^+ m/z = 263.1885
Na^+ m/z = 264.1958
$NH4^+$ m/z = 286.1778

Extracción de Cromatograma de la Información MS de todas las especies iónicas posibles :

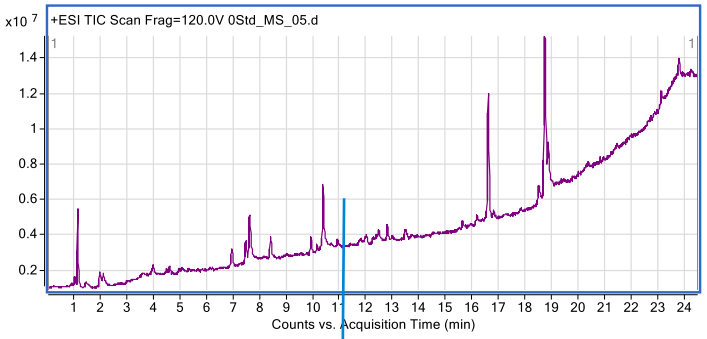
XIC : 263.1885 + 264.1958 + 286.1778

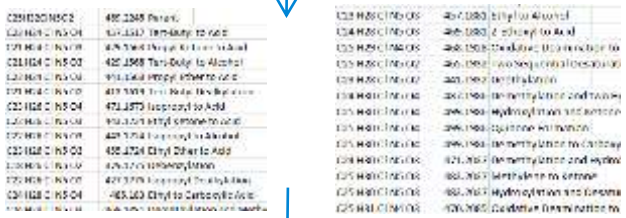


Extracción del espectro de MS del compuesto

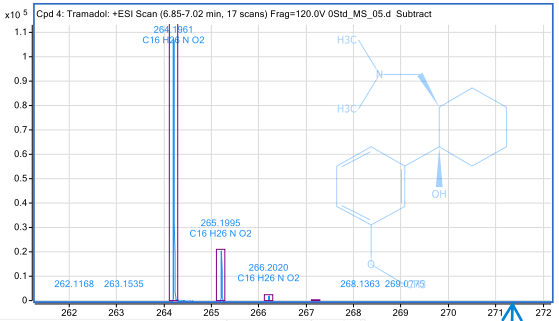
MassHunter QUAL : Find By Formula (MS)

Find by Formula (MS) análisis TARGET



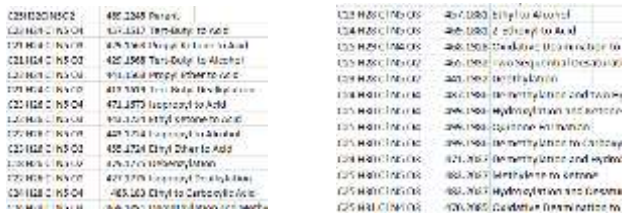


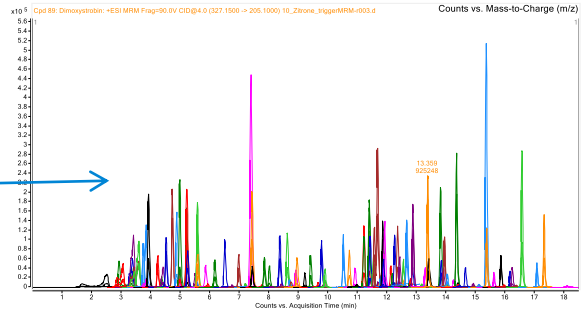
MassHunter Automático



Iones posibles : H^+ , Na^+ , $NH4^+$

Extracción de Cromatograma de la Información MS de todas las especies iónicas posibles :



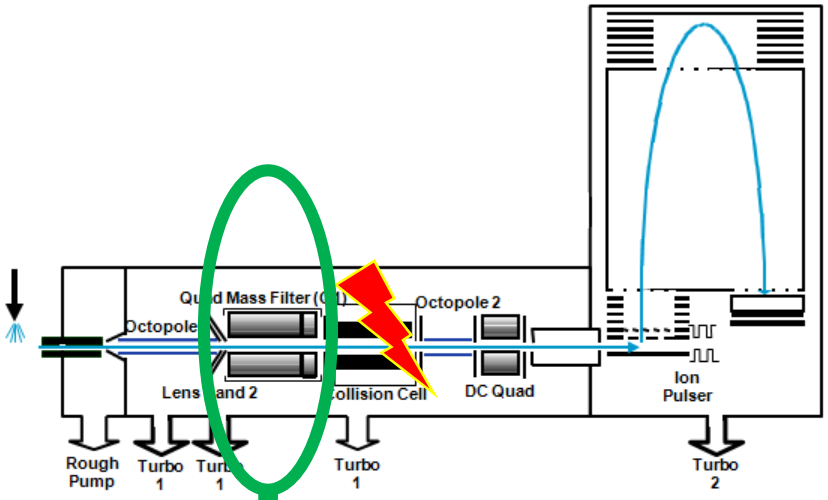


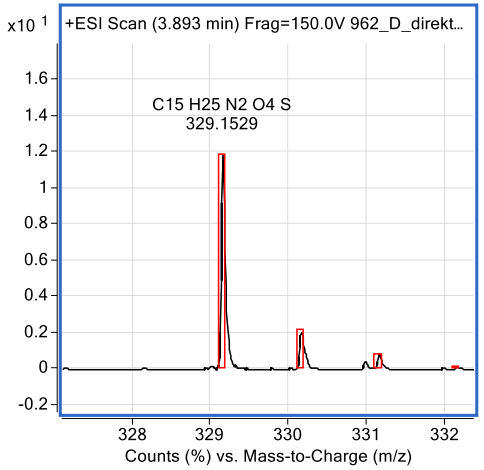
Extracción del espectro de MS de cada compuesto

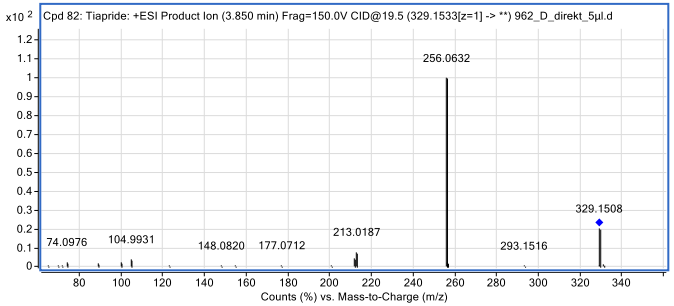
Target & suspect Screening
All Ions Acquisition

Untarget Acquisition
Target Process

QTOF
ALL-IONS MS & MS/MS Mode







Quadrupole doesn't isolate any ion. NO filtering

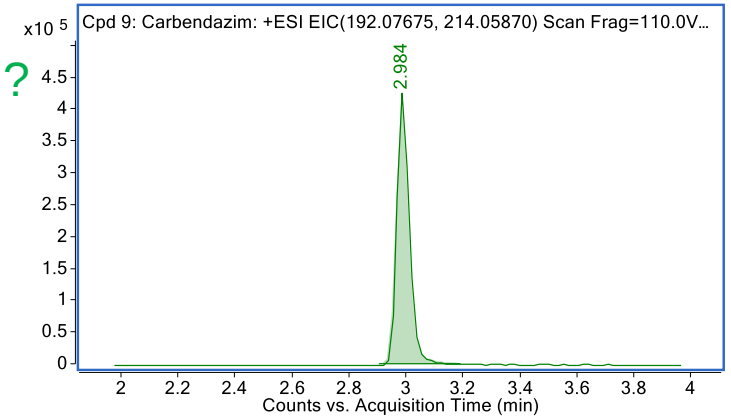
Collision Cell alternates Energies

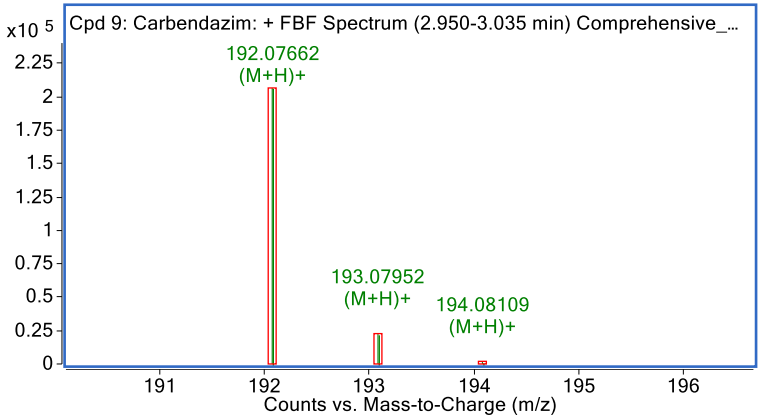
All Ions MS/MS – Screening with MS & MS/MS

Untarget Acquisition Target Process

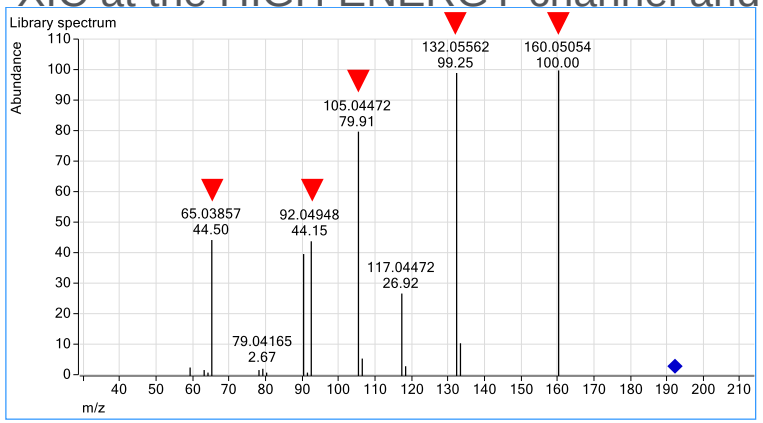
- All-Ions Data Process Algorithm is at first stage looking for compounds in a PCDL according Accurate Mass and Isotopic Pattern only at the **LOW ENERGY** Channel.

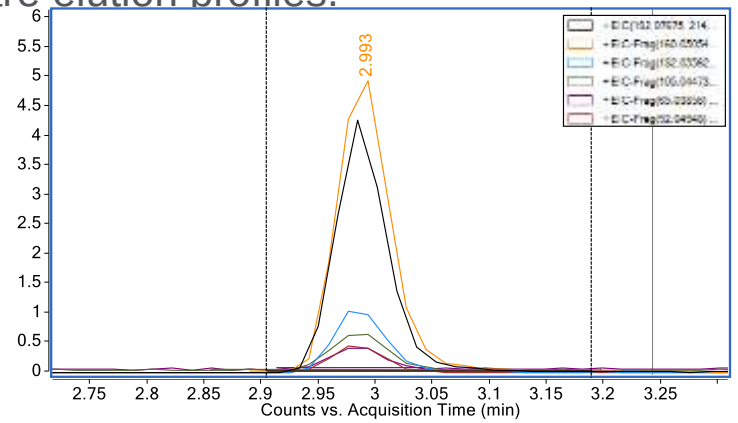
Score >90 ?





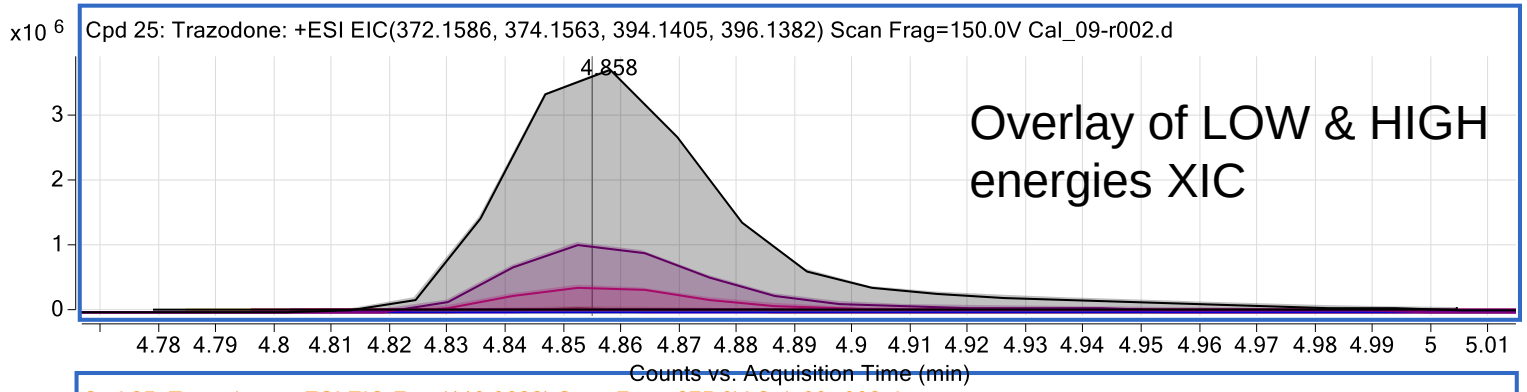
- Only if Score (user setting) is high enough, it queries PCDL Library to list known fragments in MS/MS spectra of compounds to extract their XIC at the HIGH ENERGY channel and compare elution profiles.

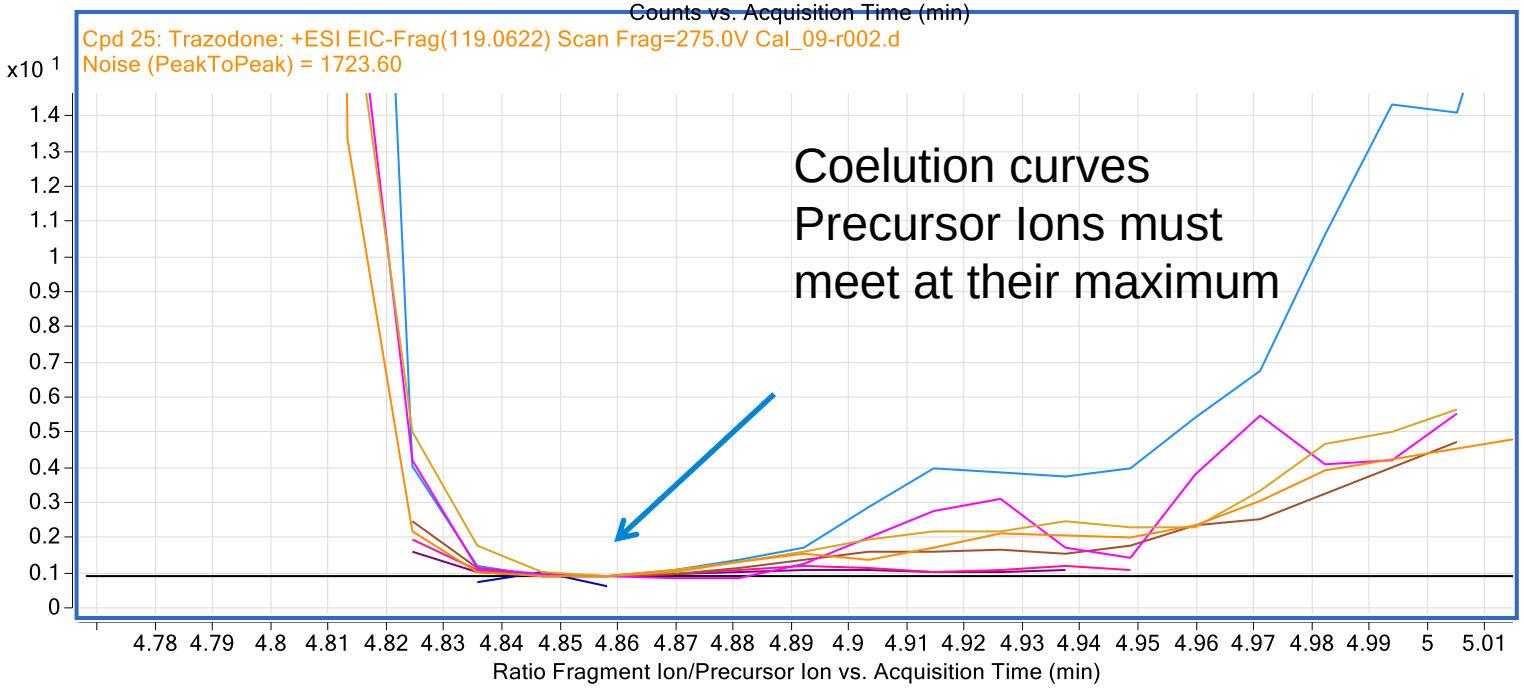




All Ions MS/MS :
Chromatografic Confirmation : Coelution Score

Untarget Acquisition
Target Process





All Ions Screening
Curated or PCDL Libraries

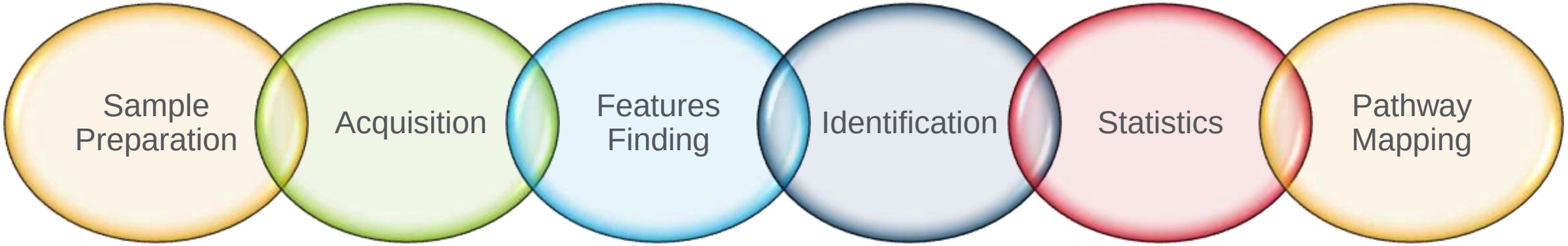
PCDLs by Compounds & Spectra

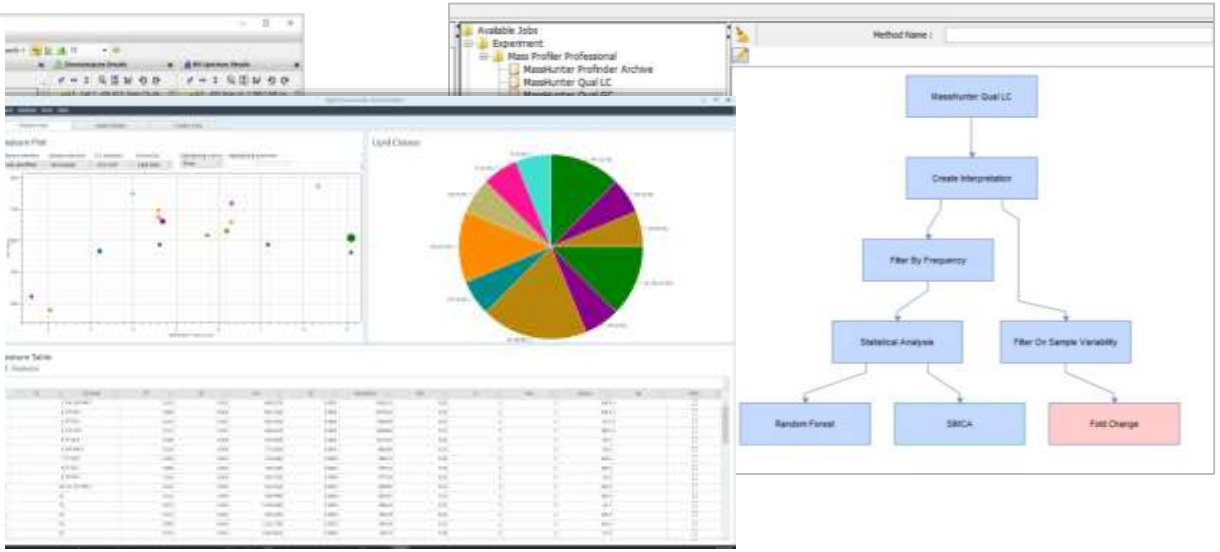
LC/MS PCDL	Market	PCDL	Compounds with AM MS/MS Spectra	Total number of Spectra	Compounds with RTs
Forensic Toxicology	Forensic Toxicology	>9,200	>3,900	>13,500	0
Pesticides	Food Safety / Environmental	>1,700	>800	>2,700	0
Veterinary Drugs	Food Safety	>2,100	>1,500	>5,200	>120
Mycotoxins	Food Safety	>450	>300	>1,300	0
Water Contaminants	Environmental	>1,400	>1,000	>3,900	>260
METLIN*	Metabolomics / Lipidomics	>79,600**	>9,400	>32,000	>680
NIST 2014 MS/MS	General	>9,300	>9,300	>234,000	0

*METLIN numbers exclude tri- and quatra- peptides in the online METLIN
** Plus 168k theoretical

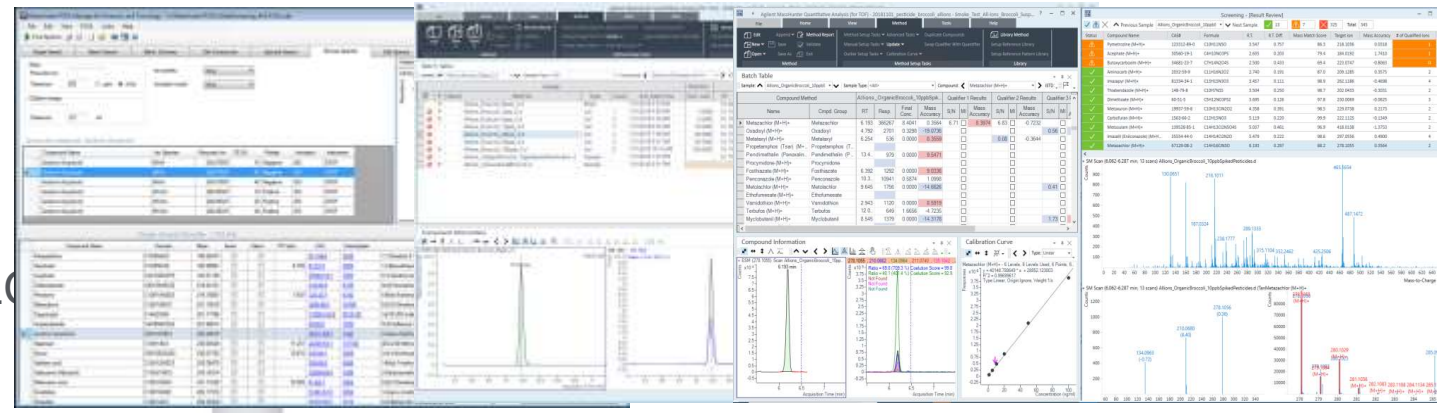
6546 LC/Q-TOF

Improving the Agilent Workflows: Metabolomics





Improving the Agilent Workflows: Food Safety

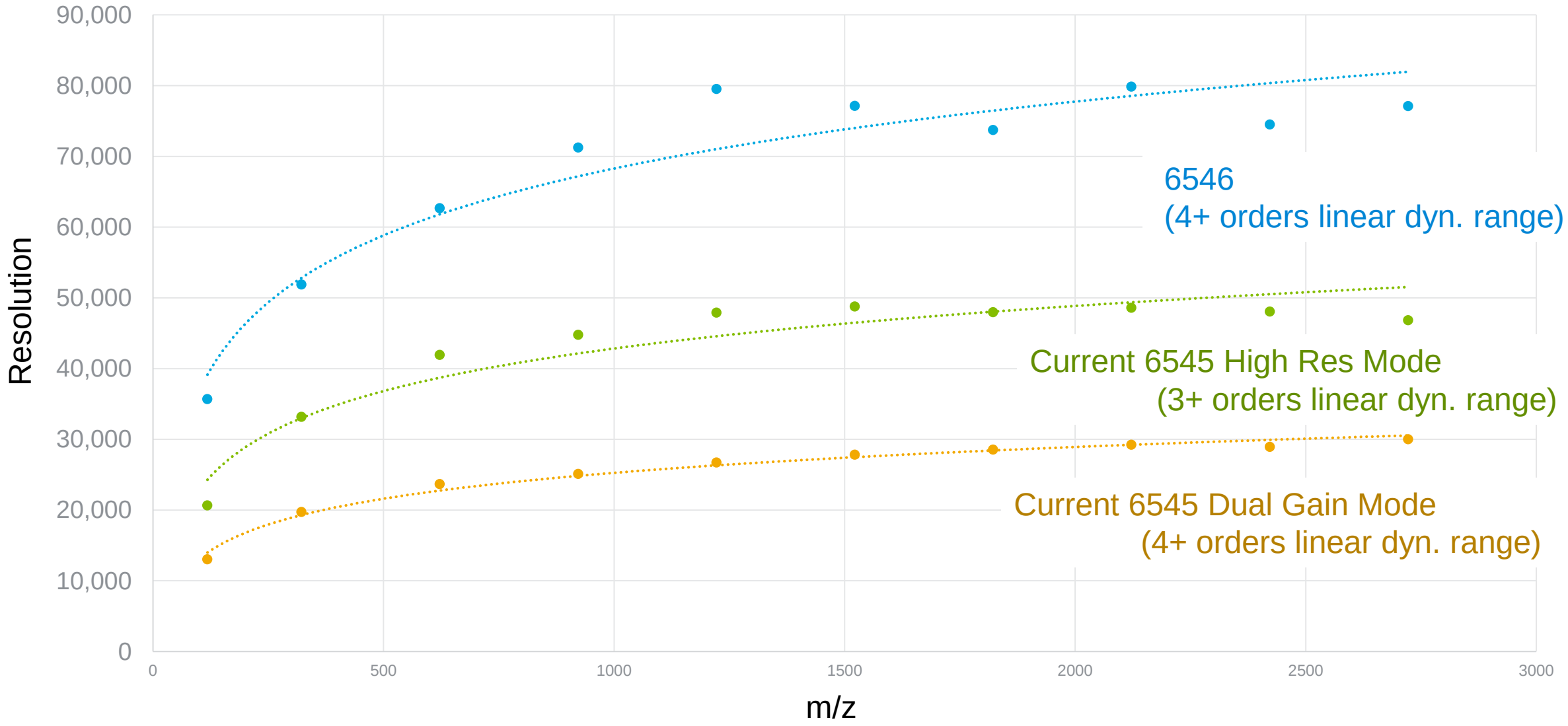


6546

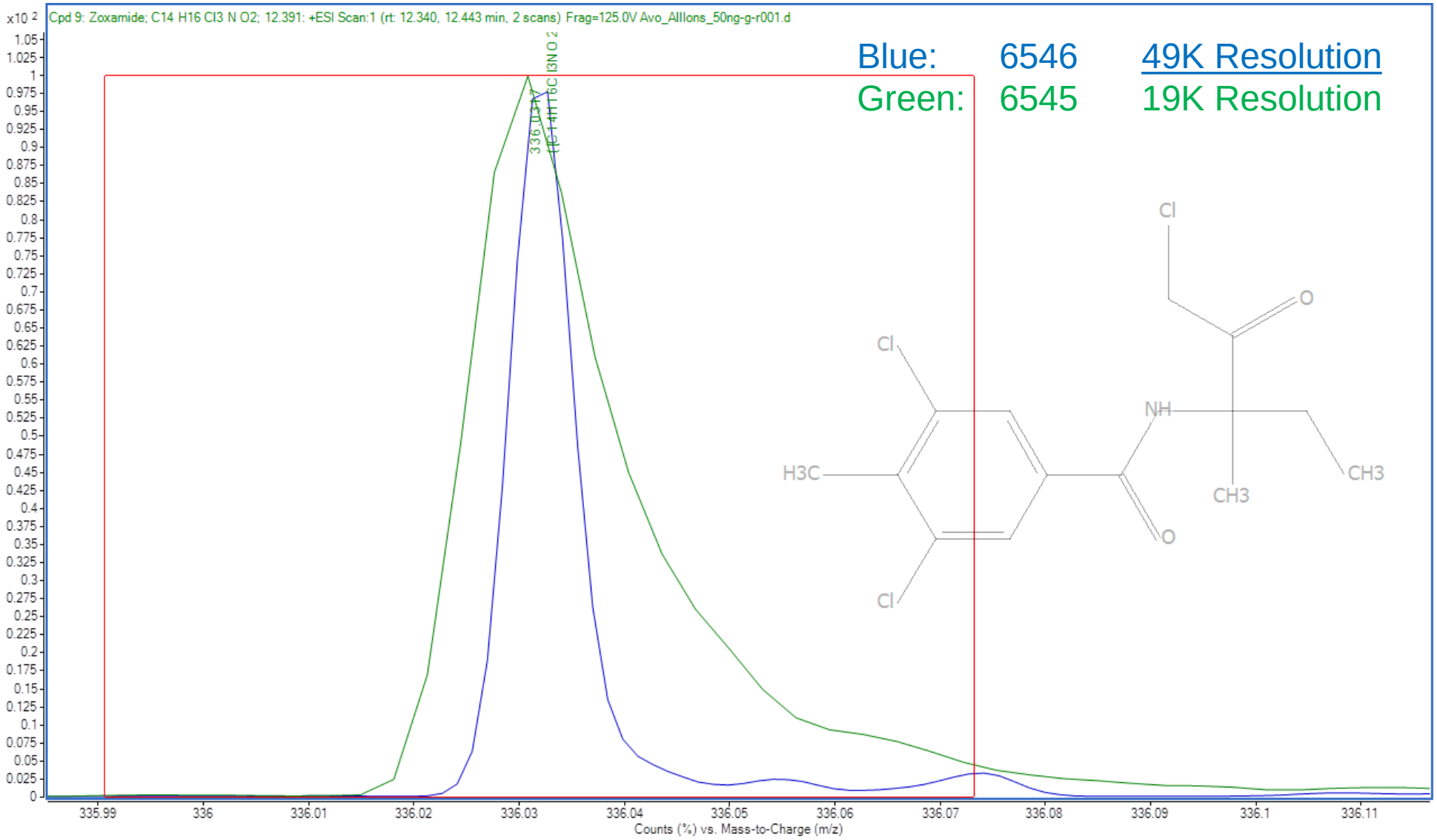
Like the 6545…with a lot more muscle!

- Simultaneous Hi Resolution, Extended Dynamic Range (10Ghz)
- Higher resolution (>60k @2722*m/z*, >30k @118*m/z*)
- DIA Quadrupole-resolved All-Ions (Q-RAI)
- Capillary gate valve
- Same sensitivity, isotopic fidelity, robustness from the 6545

Tune Mix – Measured data

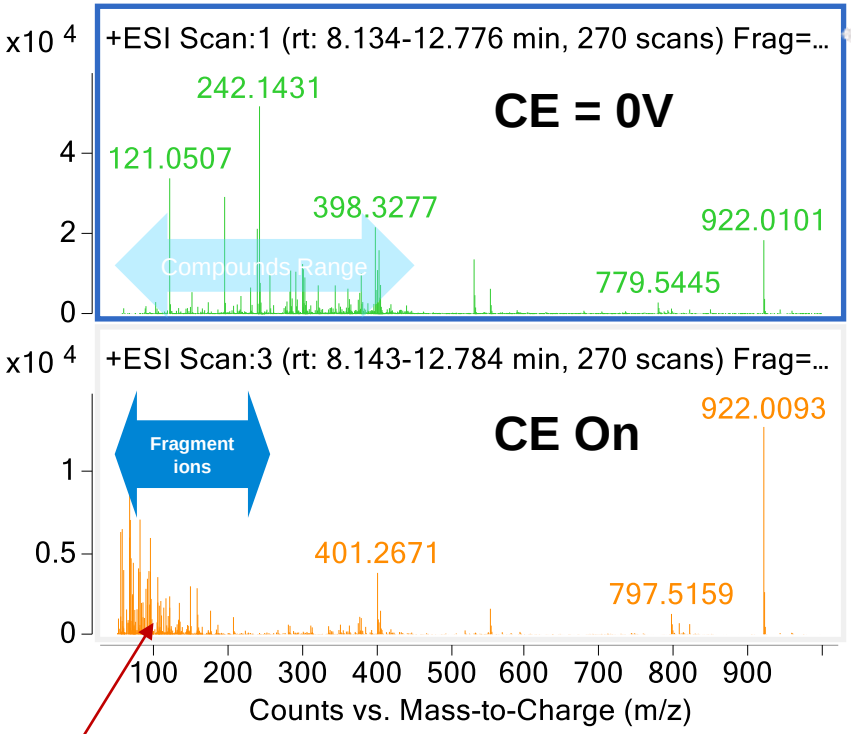


Resolution of 6546A Q-TOF System



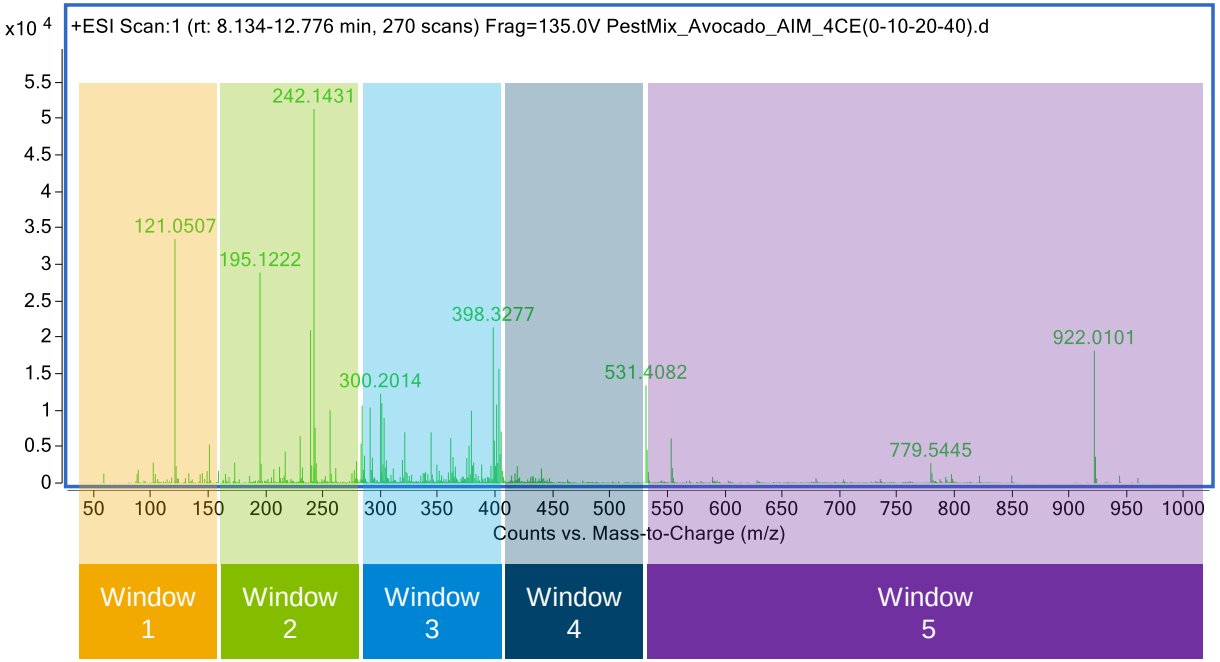
Data Independent Analysis (DIA)

Current: All Ions MS/MS



Too much interference in complex matrix at low m/z when CE applied during All Ions MS/MS

Silhouette: Q-RAI
(Quadrupole resolved All-Ions)

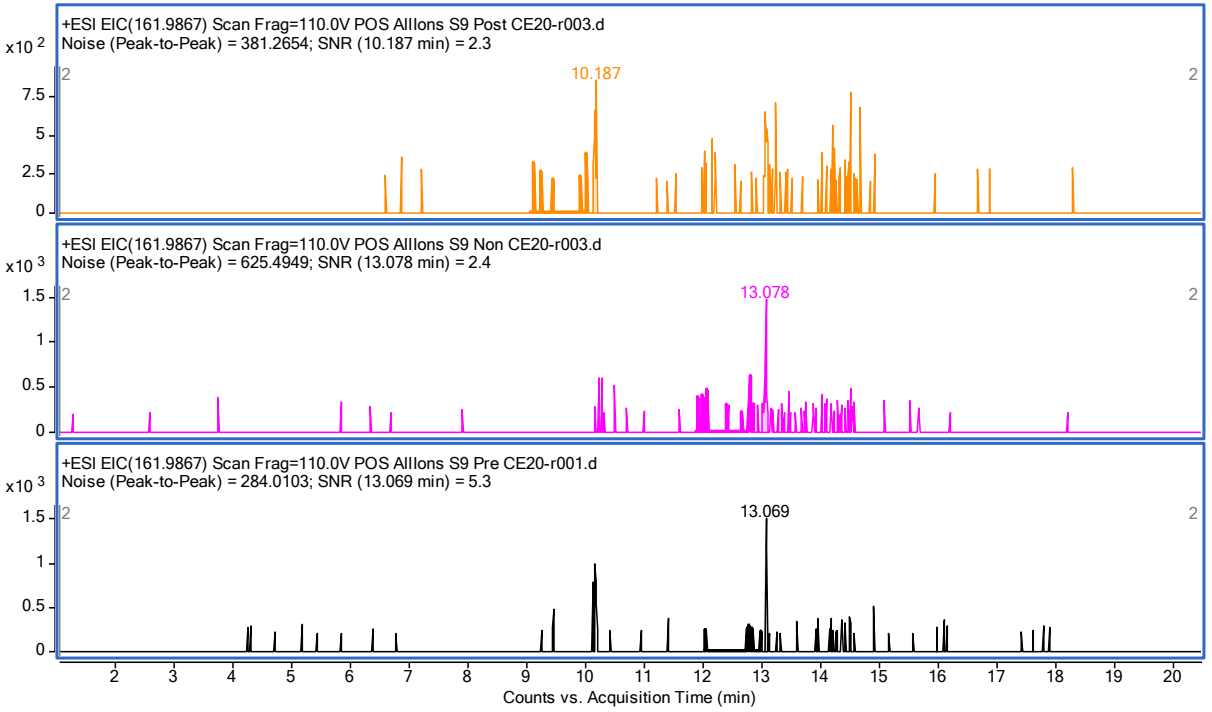


Similar to All Ions, but using the quadrupole to take sequential windows of the mass range to reduce complexity of the MS/MS spectra.

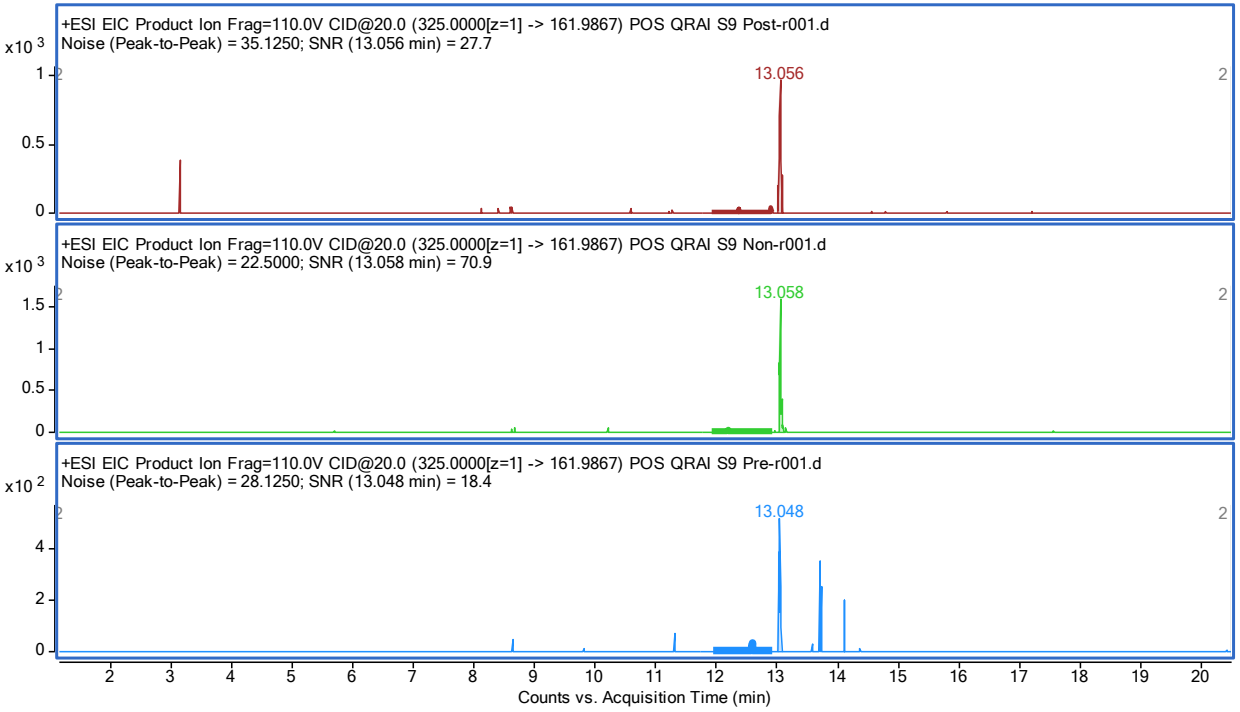
Qualifier Ion

Triclocarban [M+H]+ 314.9853 m/z

All Ions

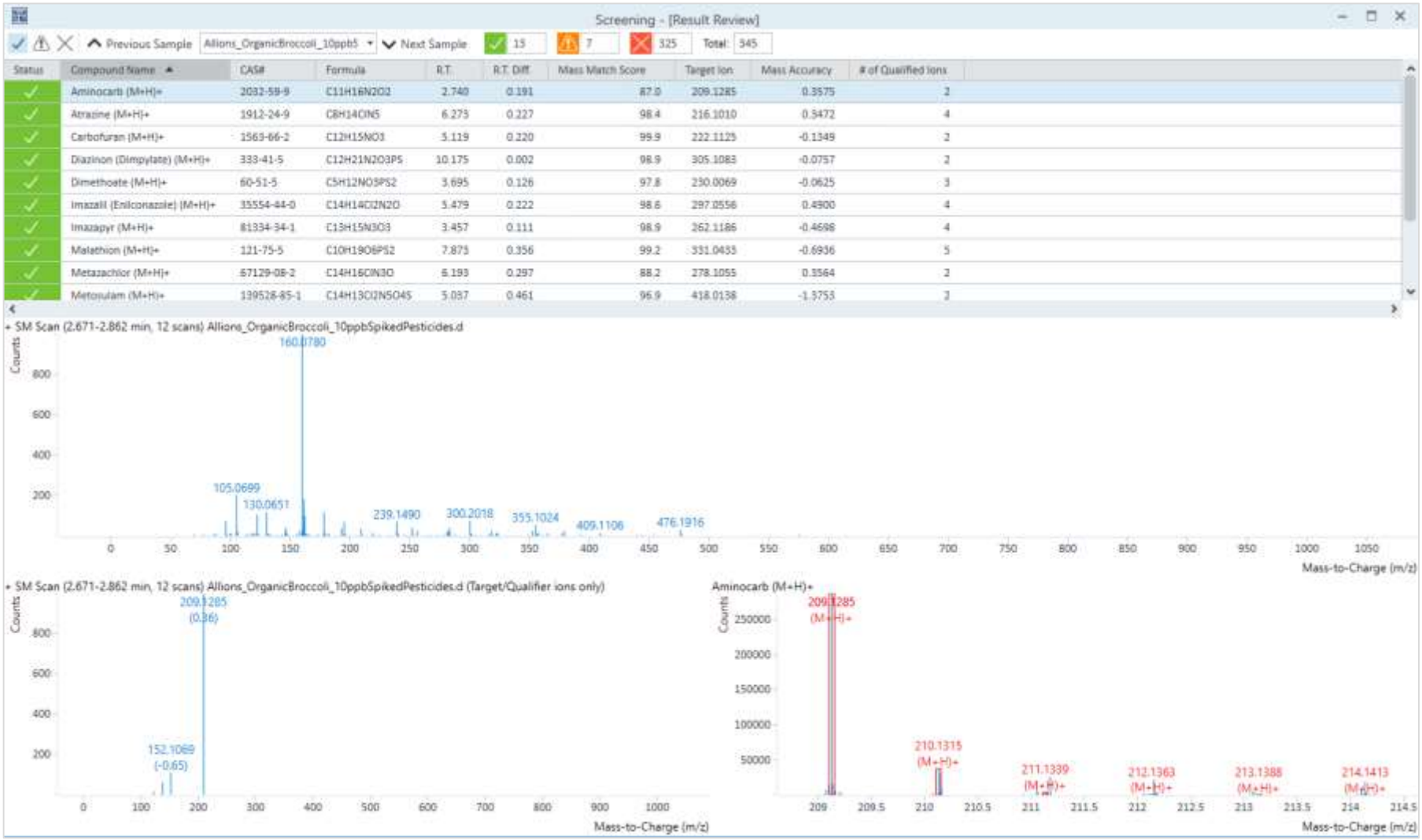


Q-RAI (Quadrupole resolved All Ions)



New: Quant 10 – New viewing tool: LC-QTOF Screener

Ease of use: improving Qual/Quant approaches



Review by “Acceptance”

Screening - [Result Review]

Previous Sample | Allions_OrganicBroccoli_10ppbS | Next Sample | 13 | 7 | 325 | Total: 345

Status	Compound Name	CAS#	Formula	R.T.	R.T. Diff.	Mass Match Score	Target Ion	Mass Accuracy	# of Verified Ions
✓	Dimethoate (M+H)+	60-51-5	C5H12NO3PS2	3.695	0.126	97.8	230.0069	-0.0625	3
✓	Imazalil (Enilconazole) (M...	35554-44-0	C14H14Cl2N2O	5.479	0.222	98.6	297.0556	0.4900	4
✓	Imazapyr (M+H)+	81334-34-1	C13H15N3O3	3.457	0.111	98.9	262.1186	-0.4698	4
✓	Malathion (M+H)+	121-75-5	C10H19O6PS2	7.873	0.356	99.2	331.0433	-0.6936	5

- Number of Verified Ions
- Target Ion Mass Accuracy
- Mass Match Score
- RT Difference

SANTE/11813/2017 Guidelines

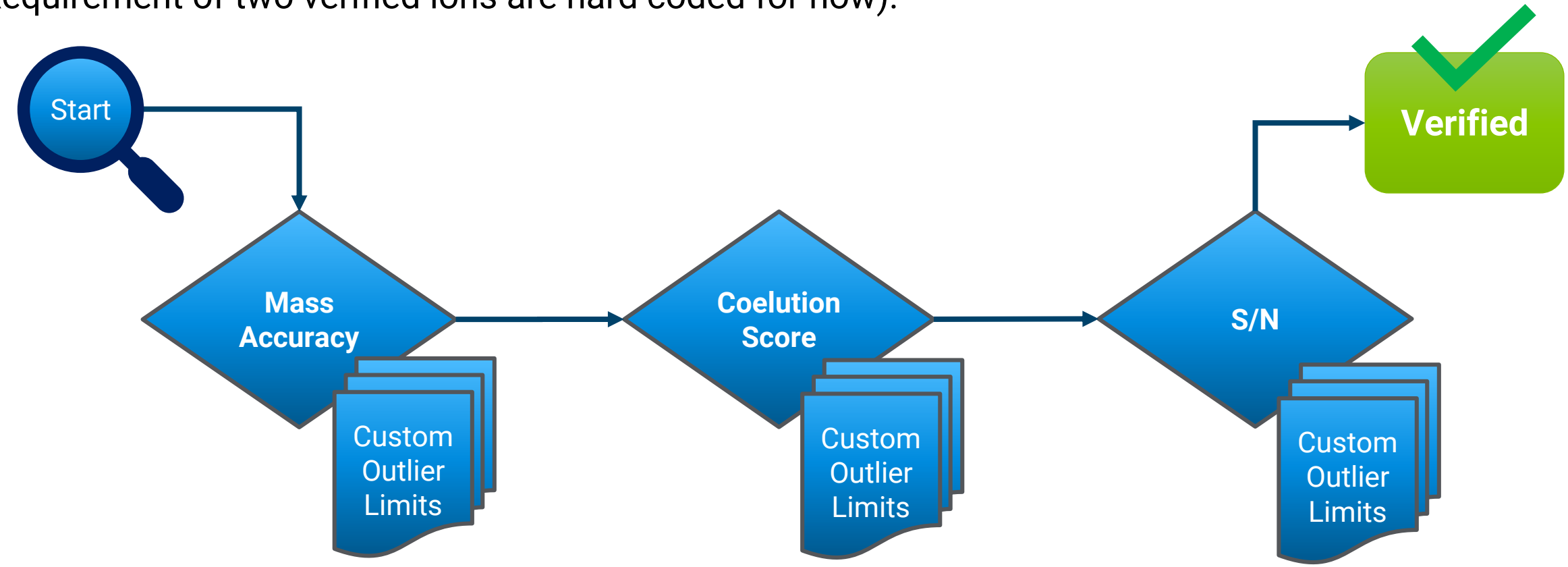
MS detector/Characteristics		Acquisition	Requirements for identification	
Resolution	Typical systems (examples)		minimum number of ions	other
Accurate mass measurement	High resolution MS: (Q-)TOF (Q-)Orbitrap FT-ICR-MS sector MS	full scan, limited m/z range, SIM, fragmentation with or without precursor-ion selection, or combinations thereof	2 ions with mass accuracy ≤ 5 ppm[a, b, c]	S/N ≥ 3[d] Analyte peaks from precursor and/or product ion(s) in the extracted ion chromatograms must fully overlap. Ion ratio: see D12

[a] preferably including the molecular ion, (de)protonated molecule or adduct ion
[b] including at least one fragment ion
[c] < 1 mDa for m/z < 200
[d] in case noise is absent, a signal should be present in at least 5 subsequent scans

Number of Verified Ions Flowchart

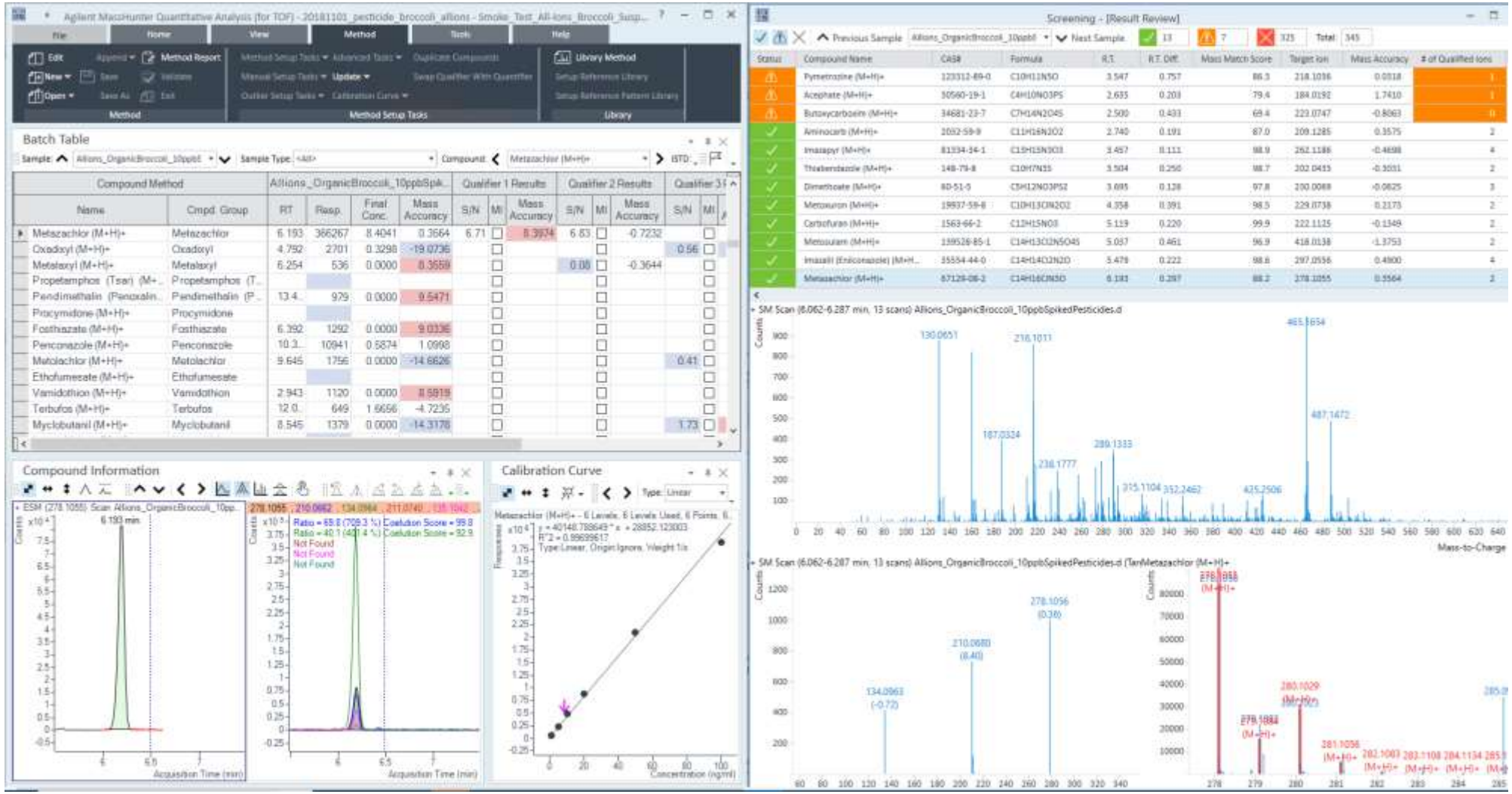
Each extracted ion feature to be evaluated against custom outlier limits

(Requirement of two verified ions are hard coded for now).



For example (≤ 5 ppm) (≥ 70) (≥ 3)

Example Results of LC-QTOF Screener



Broccoli sample (fortified)

14 cpds @ 10 ppb

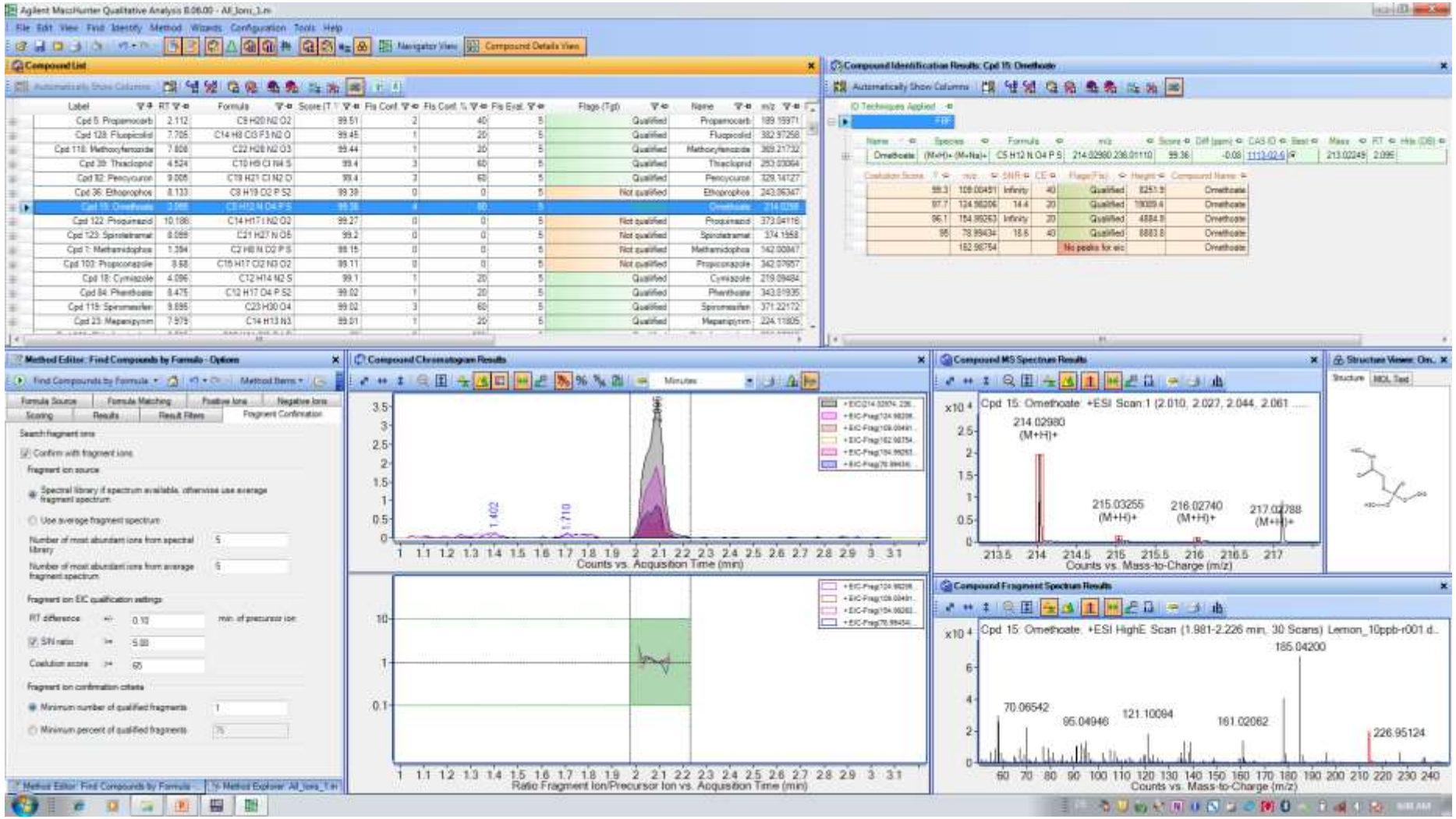
=> all identified

Target	
Aminocarb	●
Diazinon	●
Dimethoate	●
Imazalil	●
Malathion	●
Metazachlor	●
Molinate	●
Pyraclostrobin	●
Thibendazole	●
Suspect	
Atrazine	●
Carbofuran	●
Imazapyr	●
Metosulam	●
Metoxuron	●

Qualitative screening using All Ions MS/MS

Results overview of a Suspect Screening

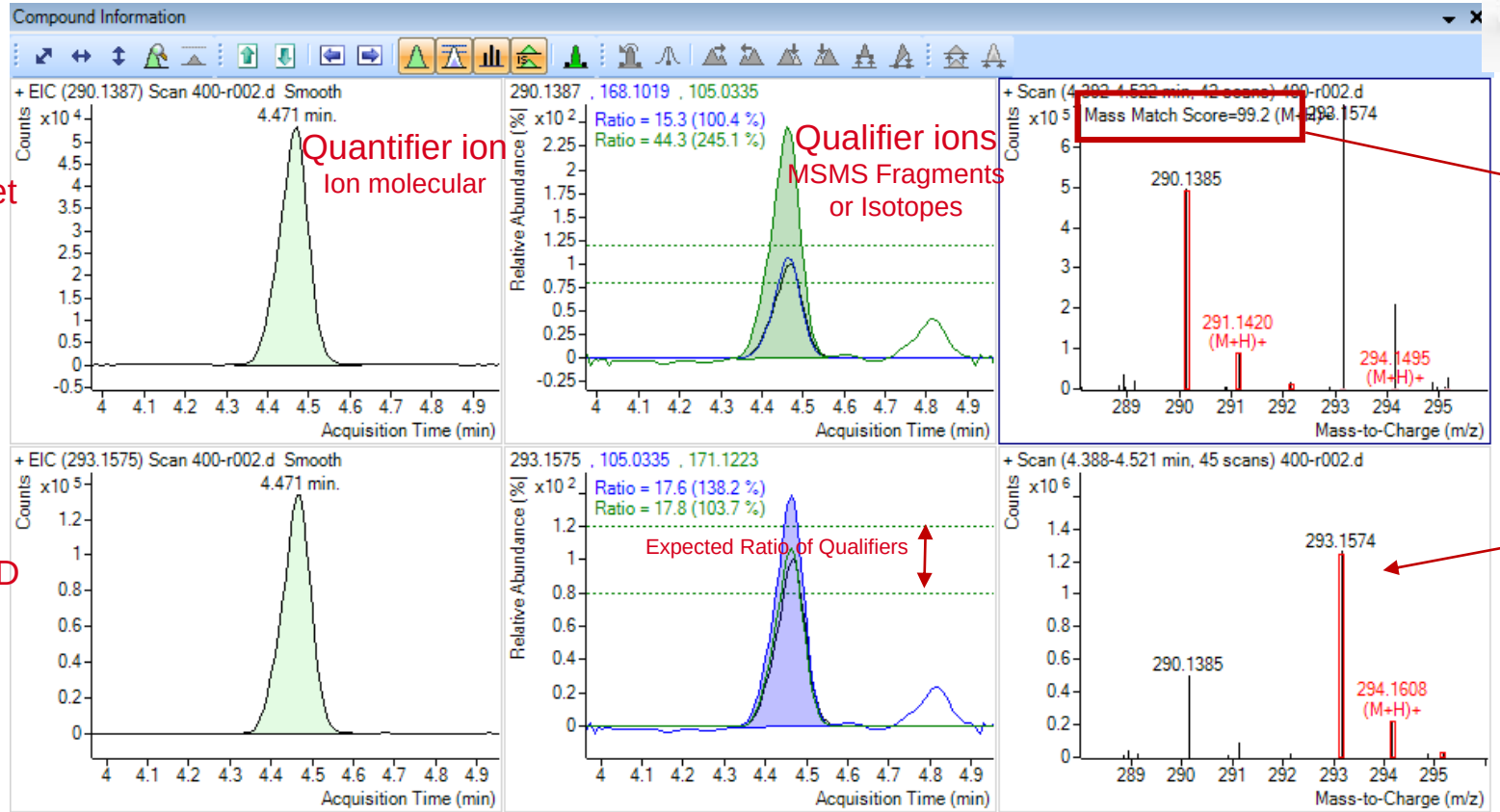
Untarget Acquisition
Target Process



Quantitative screening using All Ions MS/MS

Results overview of a Quant Screening with Standards

Untarget Acquisition
Target Process



Score based on accurate Mass and Isotopic Profile

Red boxes are theoretical Isotopic abundances

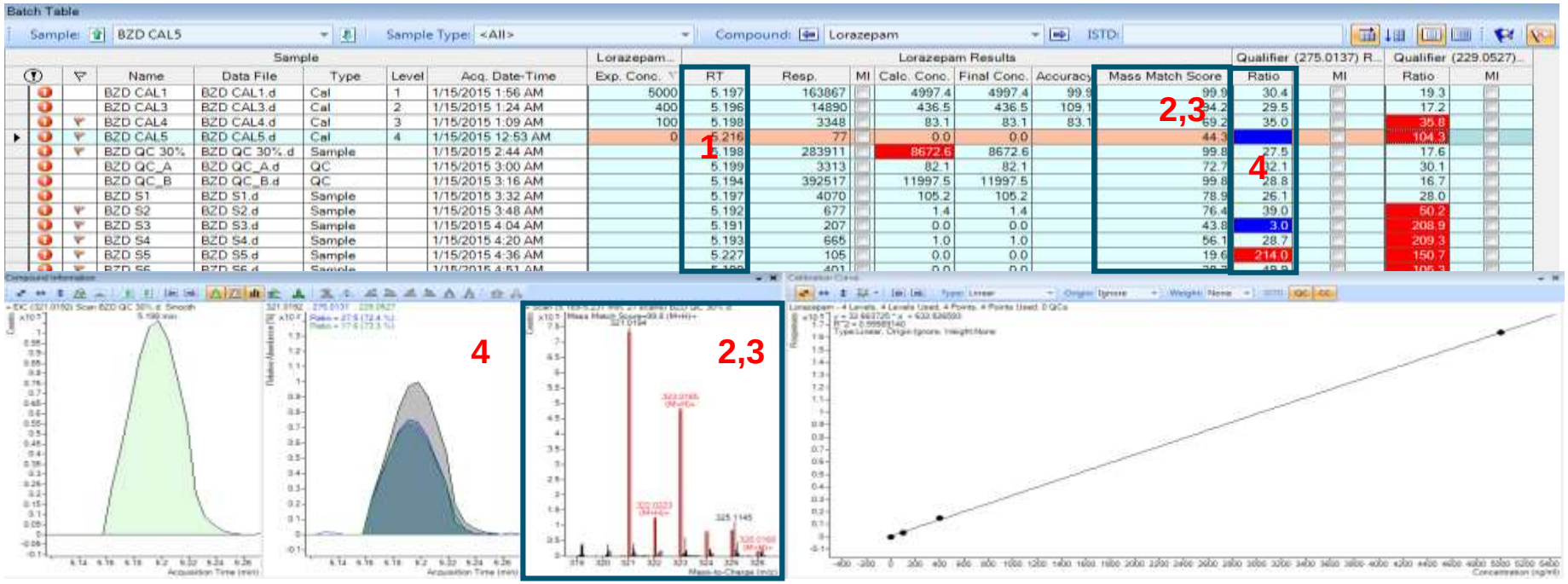
All Ions data in MassHunter Quant

Quantitative screening using All Ions MS/MS

4D-ID Confianza en los resultados

Untarget Acquisition
Target Process

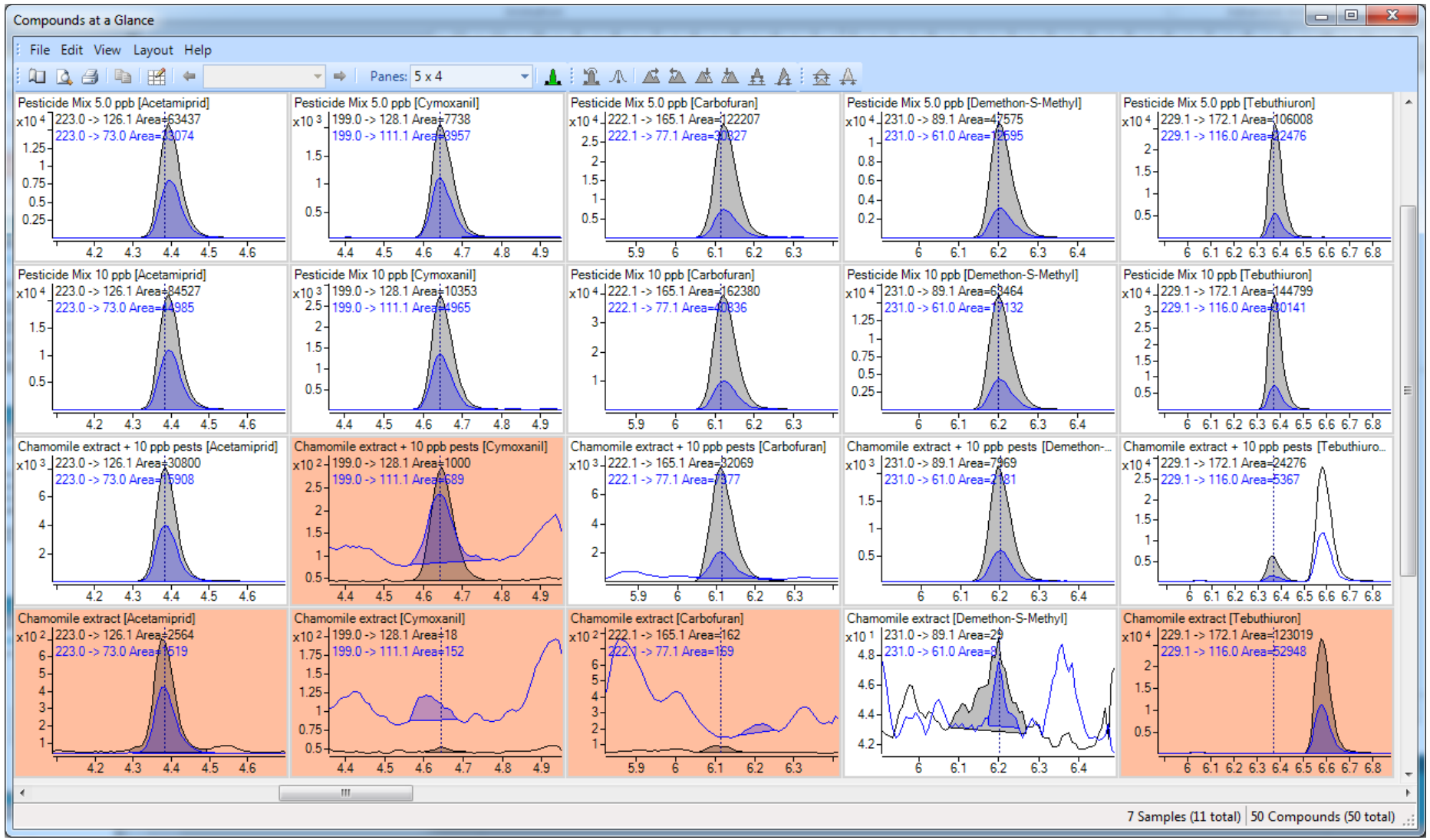
1. Tiempo de retención
2. Masa Exacta
3. Perfil Isotópico del compuesto
4. Fragmentos MSMS como Ion Qualifiers



Quantitative screening using All Ions MS/MS

Compounds at a glance

Untarget Acquisition
Target Process



Screening Definitions

Target Screening. - Four Dimensions of Identification (4D-ID) QQQ Emulation

Compounds identified through Accurate Mass, Retention Time, Isotope Pattern and Fragment Confirmation. The quant method is developed via MassHunter Qual and Libraries (PCDLs). Acquisition is in All Ions MS/MS. Standards Used.

Suspect Screening - Propose and Identify WIDEST Screening Approach

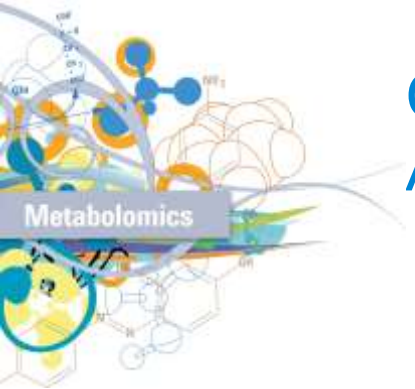
A Suspect List PCDL can be created from the encyclopaedic Master PCDLs. Compounds can be found and proposed using *Find-By-Formula*. The proposed list becomes the target or preferred (directed) list for target MS/MS or auto MS/MS. The resulting data can be searched against MS/MS spectra in the Suspect PCDL or Structure Correlated using MSC against the same PCDL if no MS/MS spectra are available.

Discovery Profiling - Find, Compare and Identify

Profiling uses Molecular Feature Extraction (MFE) which is a naïve feature finding algorithm, in order to ***find* & characterize** compounds in a data set. These compounds can then be compared to determine the absence, presence or up- or down-regulation of the compounds from sample group to samples group. Profiling can be performed in single sample comparisons, two sample groups or multiple sample groups and conditions. Significant compounds can be imported into Acquisition in order to generate MS/MS data which is priceless in further compounds identification by formula generation, library search or structural correlation.

Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - *Preparación de datos previa al Análisis Estadístico diferencial.* ***Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder****".*
  - *¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil….* ***Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*



Chemometric strategy for Life Science -omics and Food Profiling. Agilent proposal Workflows in different scenarios

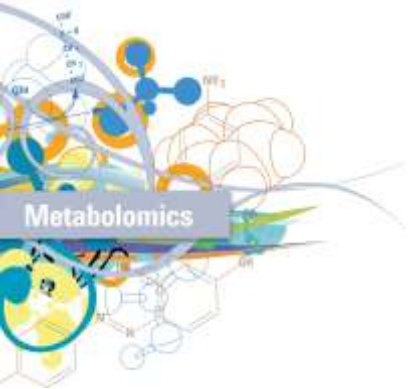
Metabolomics, as a discipline to find a **differential metabolite** to correlate with one or multiple independent variables on epidemiological studies has at least two major approaches :

- **Holistic or Untarget Metabolomics** : Top down approach with massive data to find differences, needs advanced mathematical tools.

- **Reductionist or Target Metabolomics** : Easiest approach looking for differences in just a limited list of compounds

Genomics/Transcriptomics was first of the –Omics discipline to use advanced Chemometric strategies to resolve large data sets problems.

Untarget Metabolomics and furthermore other omics disciplines like Foodomics, Glycomics, Petrolomics, etc..…… leverages from the tools developed in different applications and scenarios.



Chemometric strategy for Life Science -omics and Food Profiling. Agilent proposal Workflows in different scenarios

Metabolomic Studies Introduction

- Classical Metabolomics **objectives on research** area are to find **new Biomarkers** for **early** diseases diagnosis, classify patients,…. There are wo important phases :
  - Biomarkers **Discovery** phase will require **massive profiles** of metabolites; hundreds/ **thousands of metabolites** to follow up.
  - Biomarkers **Validation & Diagnosis** phase will require usually to follow up a **reduced number of metabolites;** a few ones/dozens of metabolites to follow up.

- Following such Metabolomics Methodology or approach, other disciplines made progress :
  - Exposomics : as study of the disease-causing effects of environmental factors.
  - Foodomics : Food and Nutrition domains through the application of omics technologies including Nutrigenomics and Nutrigenetics
  - Profiling in general. Food, Materials, etc….

Metabolomic Studies Introduction

- **Metabolomics analysis** to get metabolites profiles are based on:
  - **Chromatography/Mass Spectrometry (LC/MS,** GC/MS, CE/MS**): for all kinds of metabolites** (minoritarian & majoritarian ones). LC/MS is also used for Proteomics.
  - **Nuclear Magnetic Resonance (NMR): only for majoritarian metabolites.**

- **LC/MS, GC/MS "versus" NMR**
  - **Sensitivity: LC/MS, GC/MS** are able to **detect metabolites at much lower concentration than NMR.**
    - **MS** sensitivity is **> 10^6** times better than **NMR**.
    - **MS** requires typically > 1-100**pg** (10^{-12} g) metabolite **NMR** >200**µg** - 5mg
  - **Sample state:**
    - **LC/**MS requires **liquid samples** (or solid dissolved on aqueous or organic solvent). GC/MS also accepts gaseous samples.
    - NMR accepts liquid & solid samples.
  - **Number of Spectras/sample:**
    - **LC-GC/**MS: thousands of MS spectra/sample NMR: 1 NMR spectra/sample

A Comprehensive Metabolomics Workflow

Agilent LCMS, CEMS and GCMS

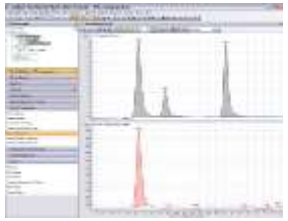
Separate & Detect	Feature Finding & Data Prepare	Data Prep. & Statistics	Identify	Pathway Analysis // Profiling

CE-LC-TOF/QTOF
CE-LC-QQQ

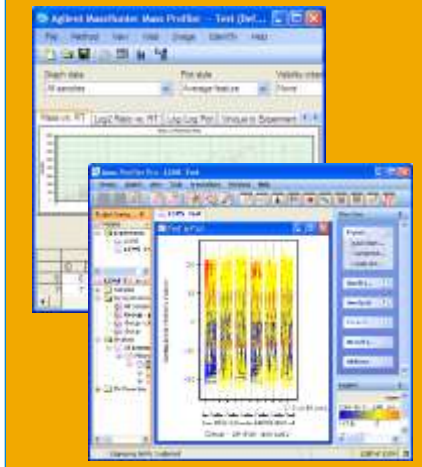
MassHunter Profinder

MassHunter Qual

MassHunter Quant



Mass Profiler (Professional)

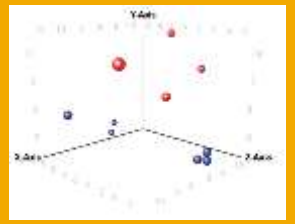


Statistics Visualization

ID Browser

Annotation & Identification

Pathway Analysis

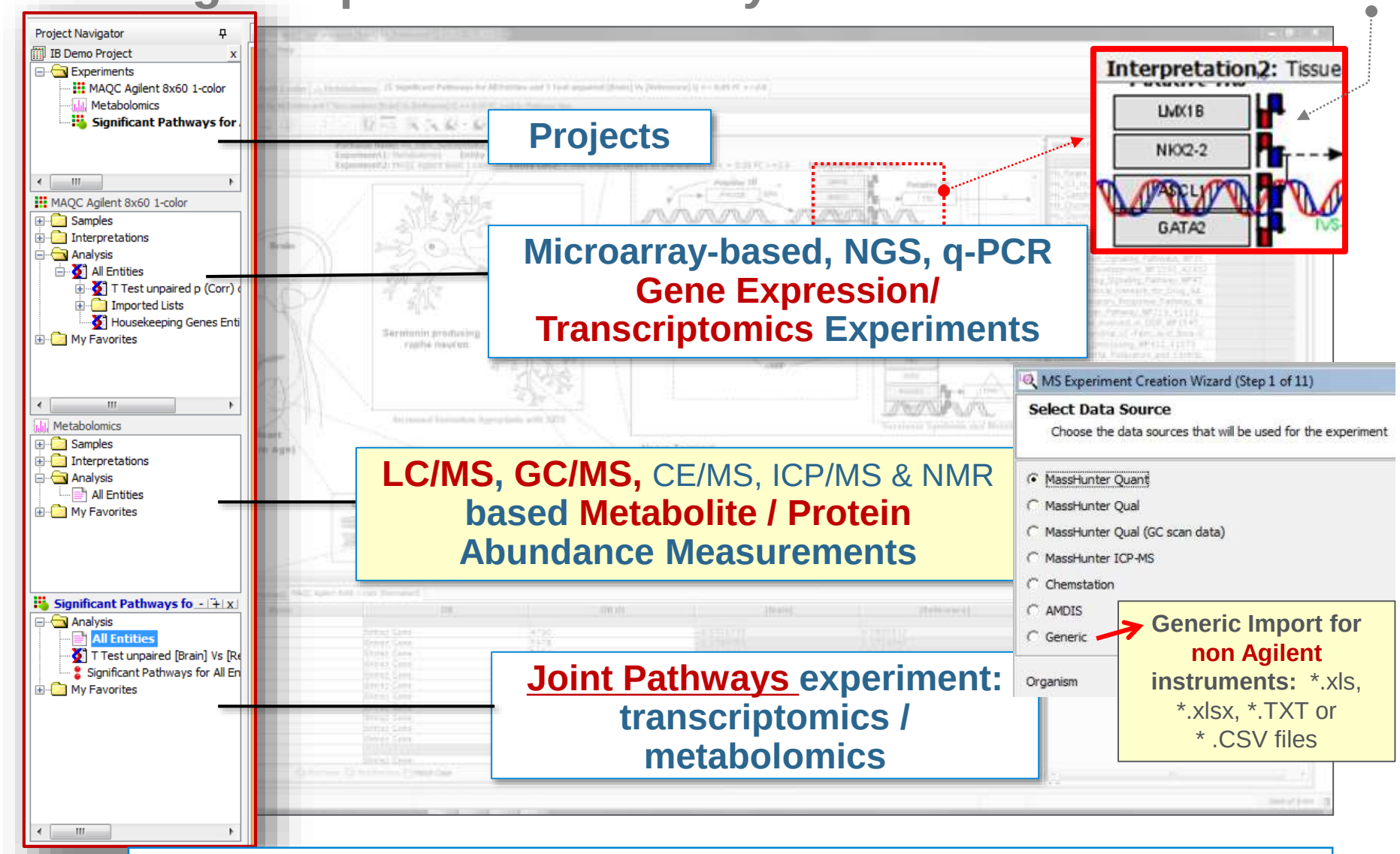


Profiling

LCMS, CEMS and GCMS Data can be analyzed together in the same project

Multi-Omics Open Platform: Mass Profiler Professional

Expression changes represented directly on routes



Projects

Microarray-based, NGS, q-PCR Gene Expression/ Transcriptomics Experiments

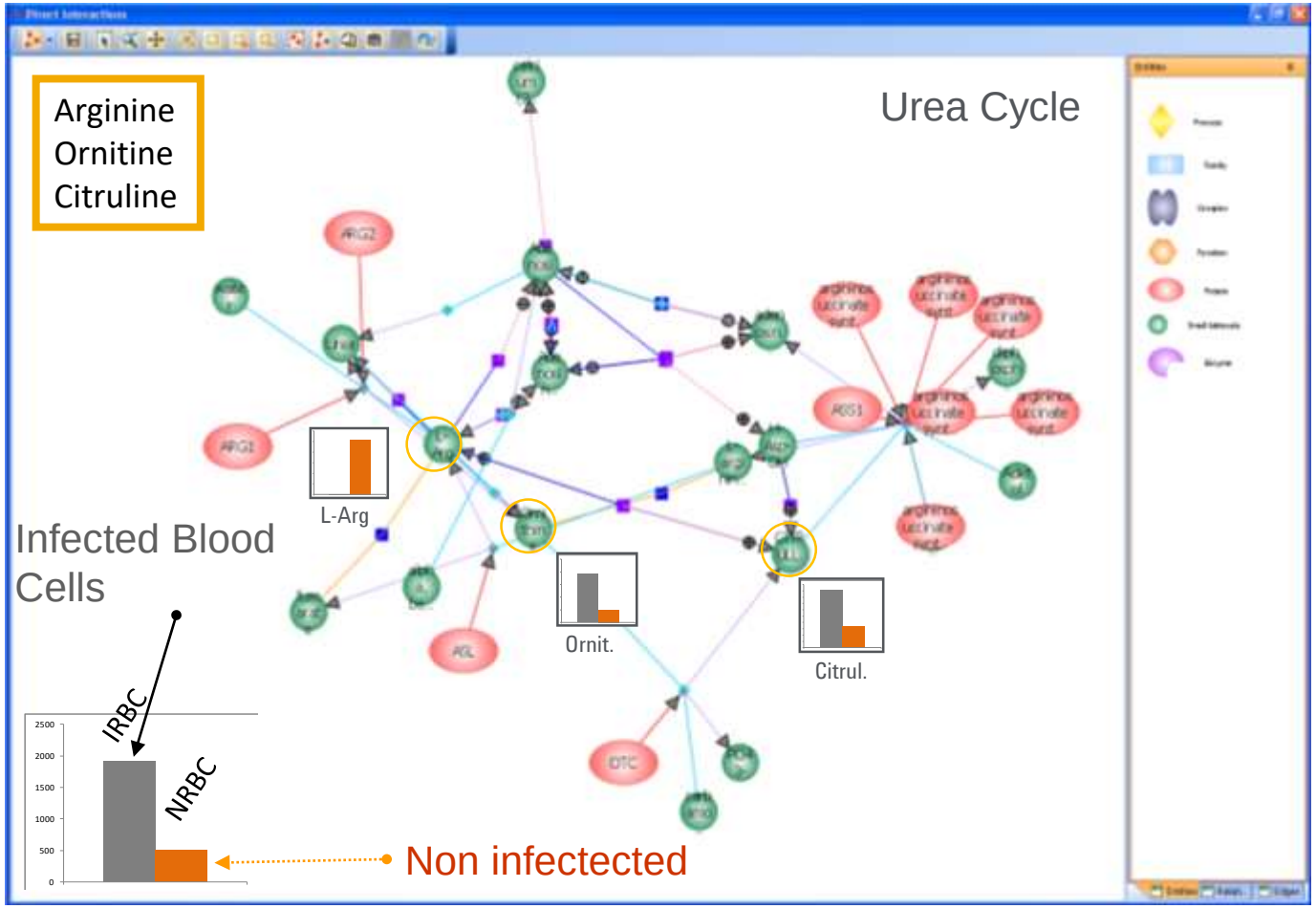
LC/MS, GC/MS, CE/MS, ICP/MS & NMR **based Metabolite / Protein Abundance Measurements**

Joint Pathways experiment: transcriptomics / metabolomics

Generic Import for non Agilent instruments: *.xls, *.xlsx, *.TXT or * .CSV files

Enrichment Analysis on curated pathways and computationally – derived networks

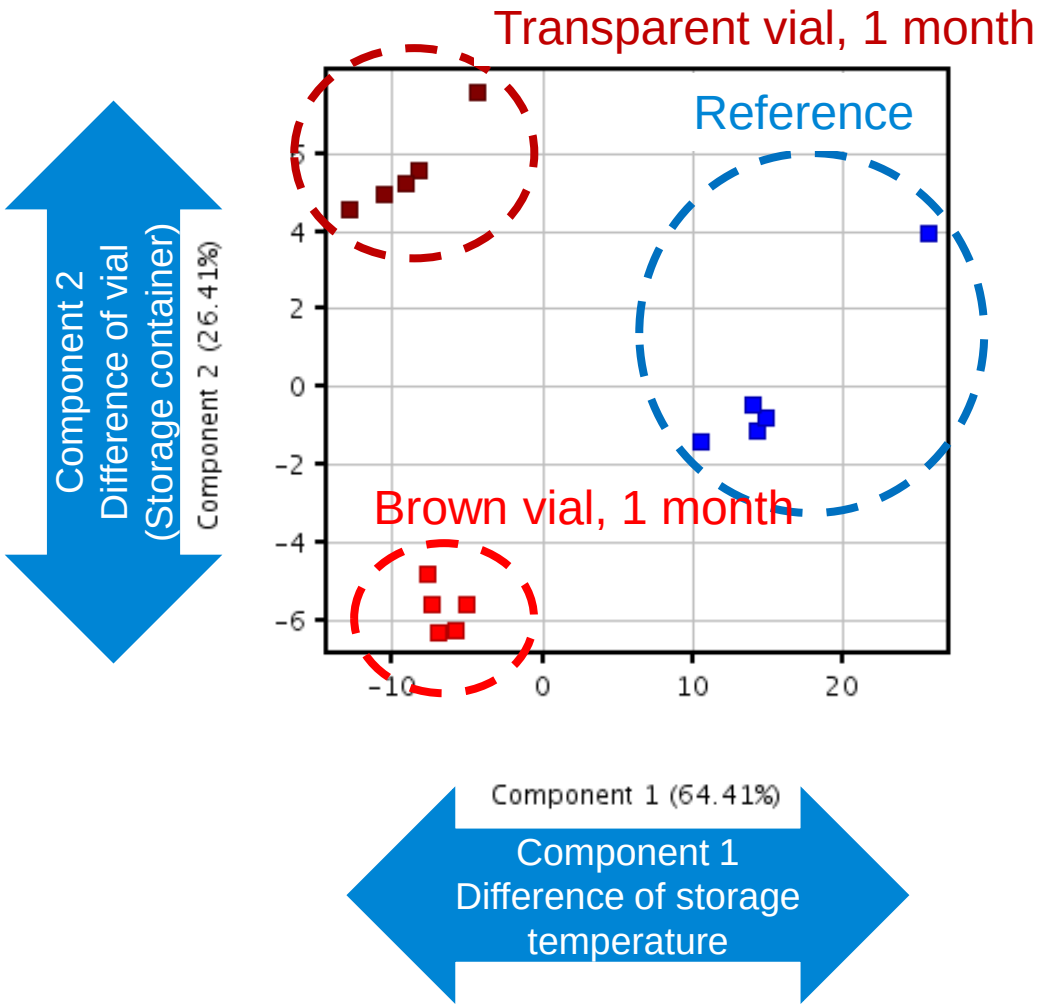
Differential Abundances of 3 Metabolites of Arginase Route (urea cycle) in Malaria Infected Red Blood Cells (RBS/ erythrocytes).



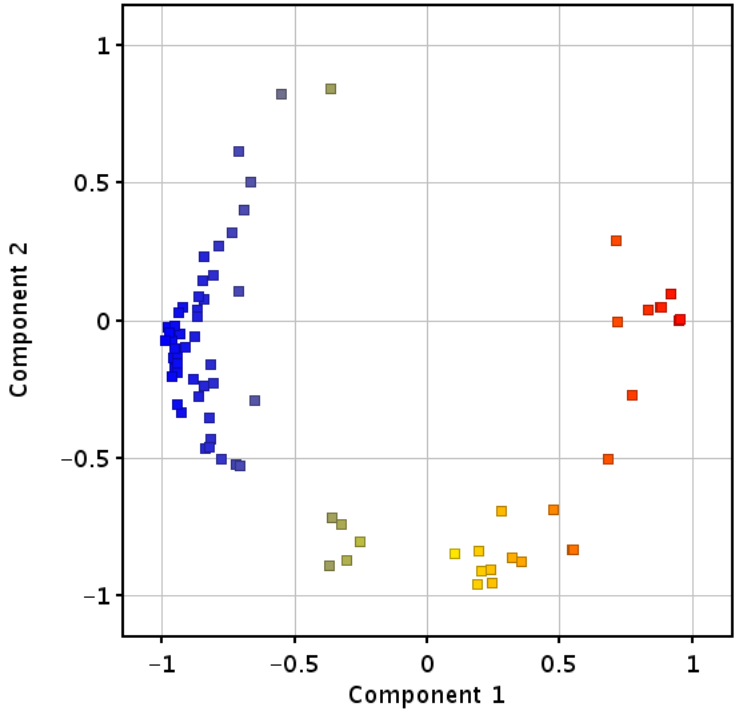
Abstract RBS SAMPLE PREPARATION:

1. - Centrifuge blood at 4ºC 2min (citrate as anticoagulant) and remove on ice the supernatant to get the erythrocytes.
2. - Wash with PBS (phosphate buffered saline) to remove the external erythrocytes metabolites.
3. - "Quench" (-25➔37ºC) and lyse the cell membrane to release the internal erythrocytes metabolites.
4. - Add aqueous phase modifier (methanol) at -20ºC.
5. -Add organic phase (chloroform) -25ºC.
6. – Do Liquid-liquid extraction at different pH's.
7. - Evaporate in vacuum and do and aqueous extract reconstitution.

Sake (Japanese liquor) deterioration test
PCA 2D Score Plot and Loading Plot



PCA Score Plot



PCA Loading Plot

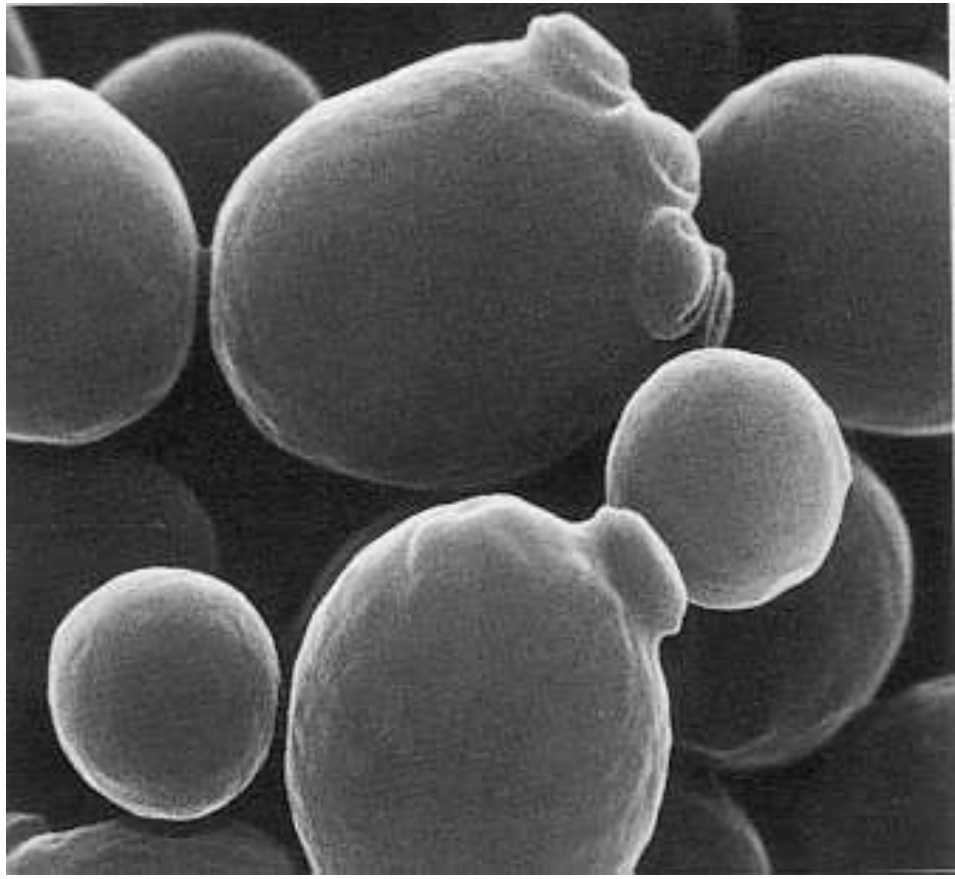
A point in the loading plot corresponds to a compound

Agilent
Omics approaches examples

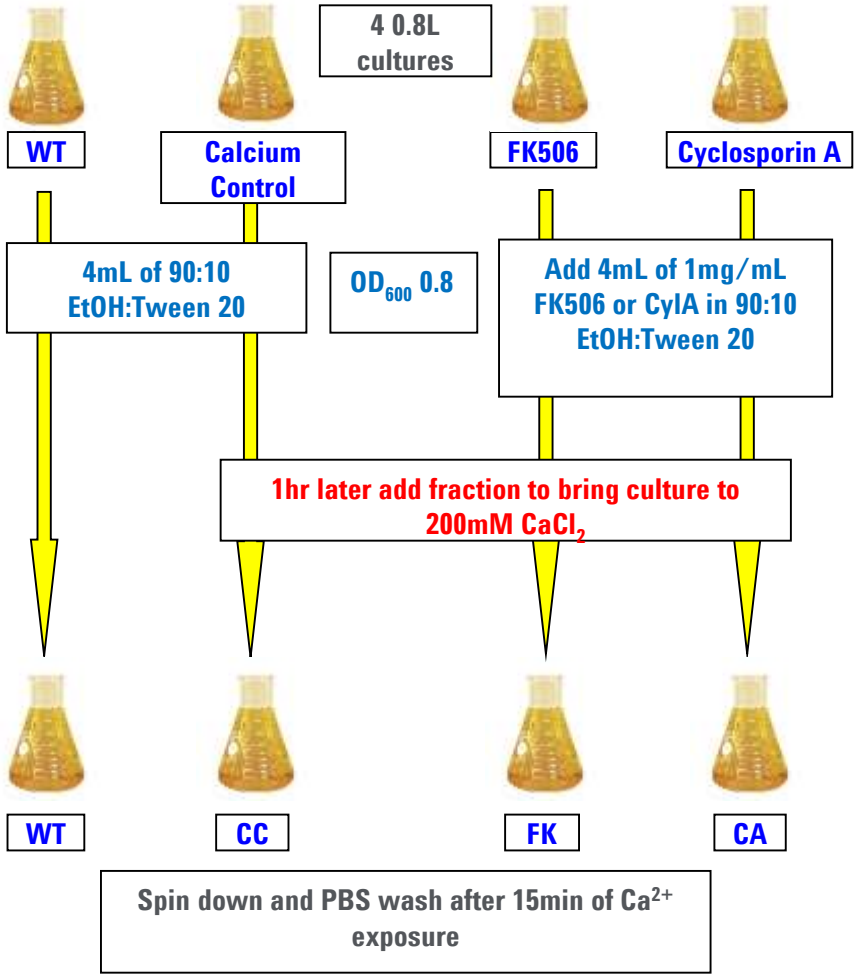
Metabolomics & Food Profiling

Effect of two different drugs on Ca regulation pathway

- *Baker's Yeast is an Ideal Model Organism for Studying Pathways*
- *Saccharomyces cerevisiae* is a widely used model organism
- Biochemistry and pathways are extensively studied
- Fully sequenced genome
- Ideal for "multi-omics" studies with the goal of facilitating research for other organisms.



Yeast Metabolomics :



Experimental Design

Treatments:

- Wild type (WT) - no treatment
- Calcium control (CC) - $CaCl_2$
- FK - FK506 and $CaCl_2$
- CA - Cyclosporin A and $CaCl_2$

Extraction :

Wet mill with 5:3:3 $CHCl_3$:CH_3OH:H_2O. Only the aqueous is analyzed

Metabolomics Workflow

Acquire Data

- Analyze metabolomics samples
- Use accurate mass LC/MS to analyze samples

Analyze Data

- Mine data using MassHunter
- Analyze in Mass Profiler Professional

Identify Compounds

- Identify metabolites using ID Browser
- Search METLIN to annotate metabolites

Pathway Analysis

- Map results to pathways using Pathway Architect

TOF/Q-TOF For Discovery Metabolomics

Goal – Detect all metabolites

Data is acquired in Full Scan

Metabolite tracking uses retention time and mass or mass fragments

Statistical analysis is used to find differential metabolites (features)

Feature identification is required for biological interpretation

Ideal for discovery metabolomics

Spectral quality

- Accurate mass
- Good mass resolution
- 5 orders of dynamic range
- High isotope ratio fidelity
- Maintains performance at high acquisition speeds
- Sensitive

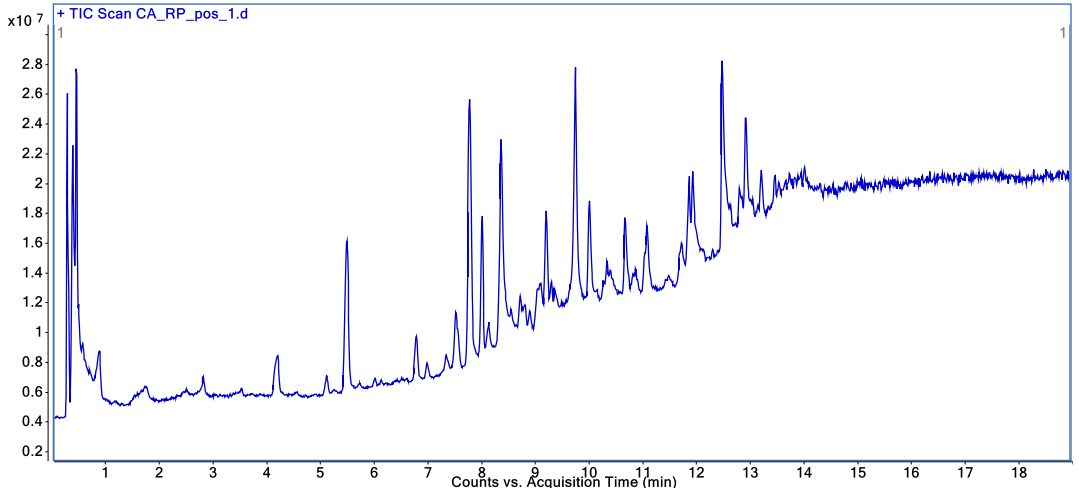
Compound identification

- Accurate mass, isotope ratio
- MS/MS with accurate mass, isotope ratio (Q-TOF only)

LC/MS Analysis of Metabolites in Stressed Yeast

ESI (+) by RP

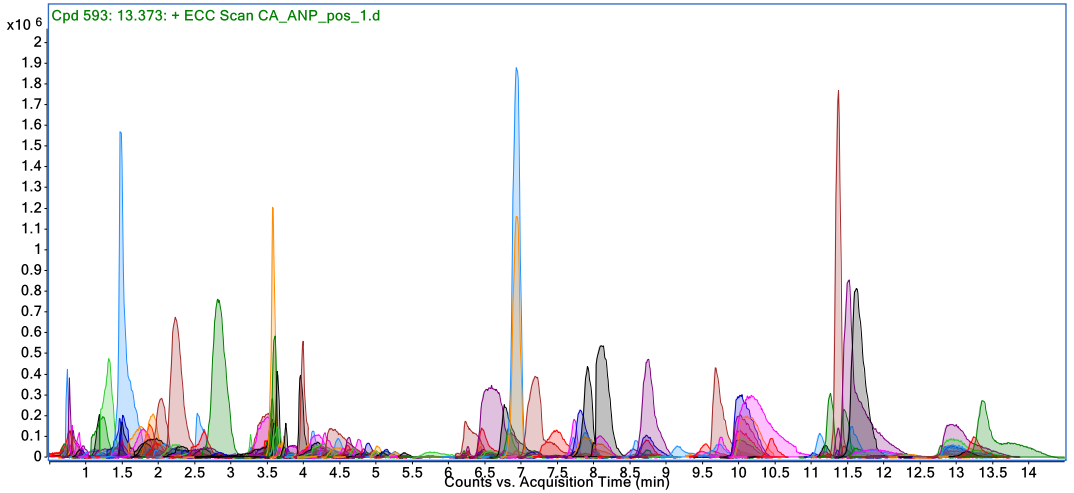
Polar compounds elute close to the void volume



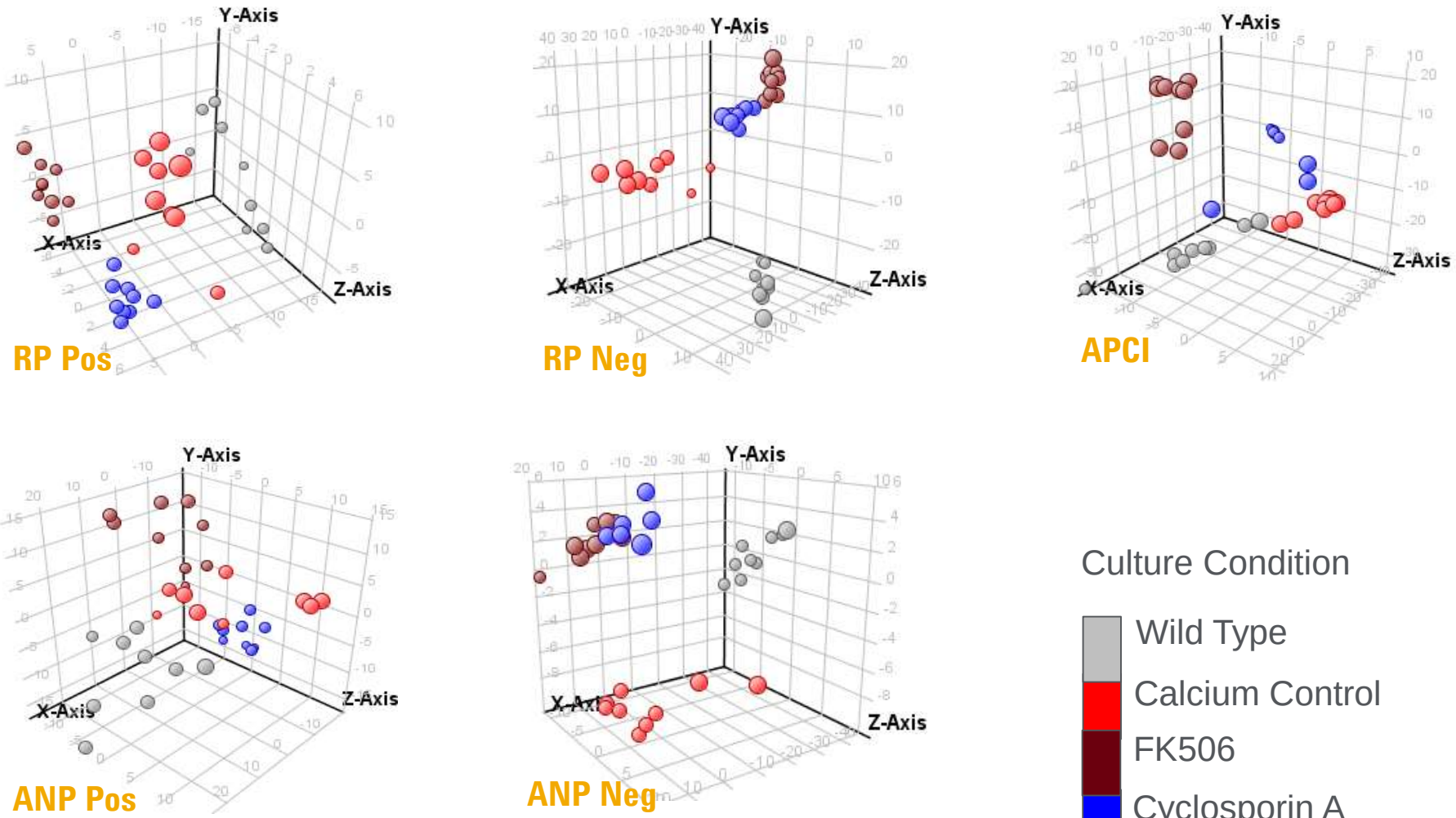
ESI (+) by ANP

ANP separates polar compounds that were in the void volume with RP

Using both ANP and RP yields more comprehensive coverage



PCA Plots for Yeast Metabolites Using Different Analytical Methodologies



Increasing Your Confidence in Compound Identification

Confident compound identification is crucial for pathway visualization!

Increasing confidence in identification

Compound identification data

Accurate Mass (AM)	AM + Isotope Pattern (IP)	AM + Retention Time (AMRT) + IP	MS/MS Library	MS/MS Library + AMRT

Summary of Yeast Metabolomics Analyses

Differential Features with METLIN Database Annotation

Polar metabolites

- ANP chromatography
- ESI +/-

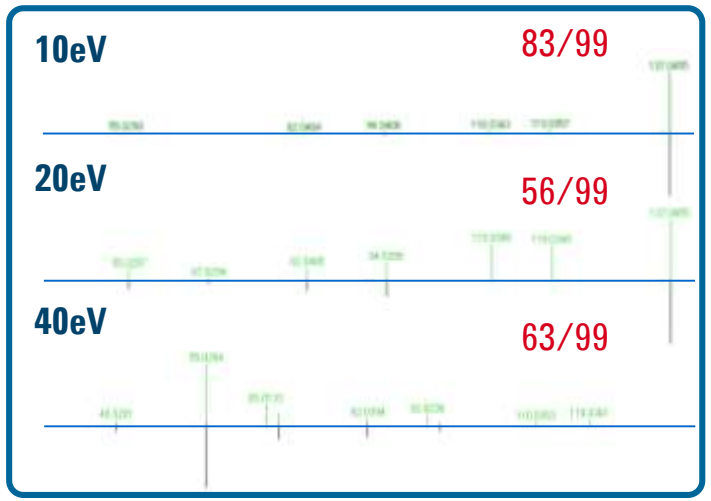
Non-polar metabolites

- RP chromatography
- ESI +/-
- APCI

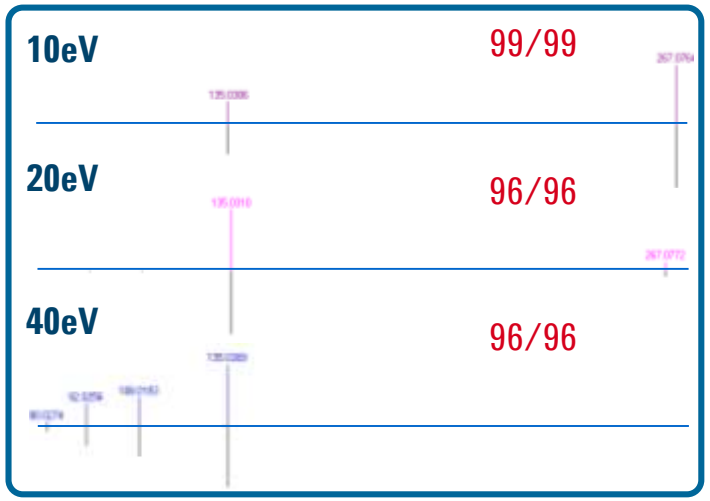
	Number of Features			
	After QC	METLIN Matches	p<0.05 Cutoff	METLIN Matches
RP-ESI pos	300	112	158	79
RP-ESI neg	523	141	418	115
RP-APCI pos	364	48	333	37
ANP-ESI pos	492	155	145	113
ANP-ESI neg	276	88	213	63

MS/MS Identification Using the Agilent METLIN PCDL Library

Hypoxanthine *m/z* 137.0458 (+)



Inosine *m/z* 267.0740 (-)

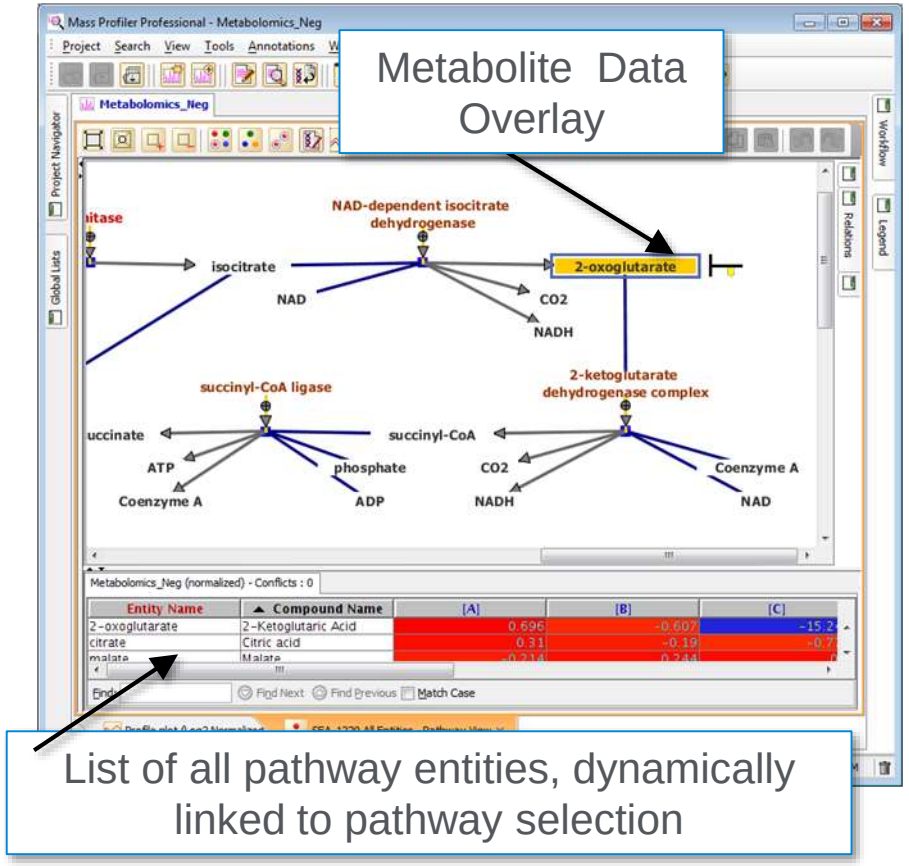


MS/MS spectra library matches:

- MS/MS spectra obtained at 10, 20 and 40eV collision energies
- Matched to METLIN PCDL library spectra
- Displayed as acquired spectra mirrored above library spectra

Pathway Architect

Pathway Architect is an optional module in MassProfiler Professional



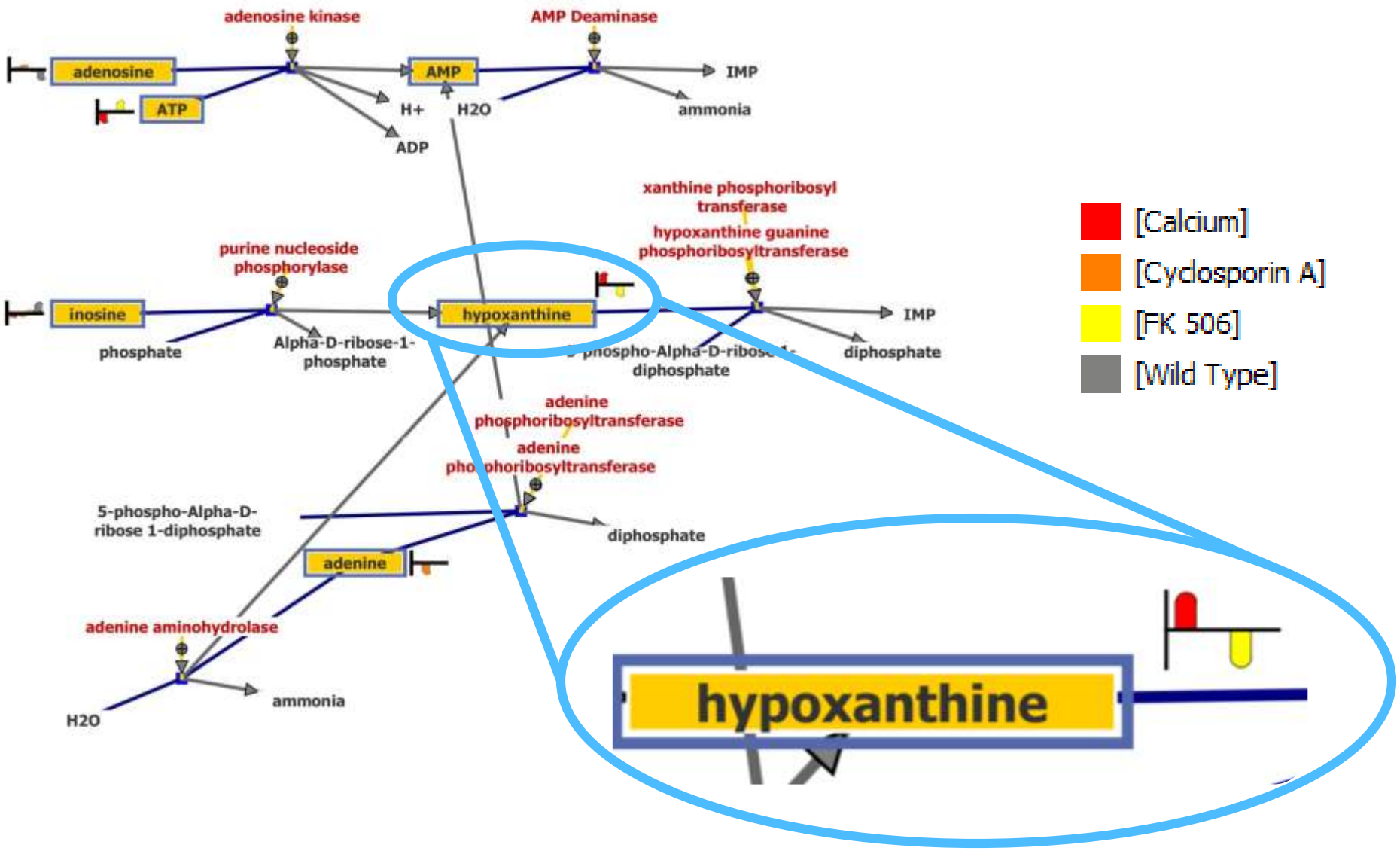
Map and visualize data from one or two types of -omic data on pathways

Search, browse and filter pathways

Supports biological pathways from publicly available databases

- WikiPathways
- BioCyc
- Supported pathway formats
  - BioPAX 3 – Pathway Commons, Reactome, NCI Nature Pathway
  - GPML – PathVisio –custom drawing
- Export compound list from pathways

One of the Stress Activated Pathways in Yeast



Introduction to Food Profiling

The questions contaminant testing won't answer:

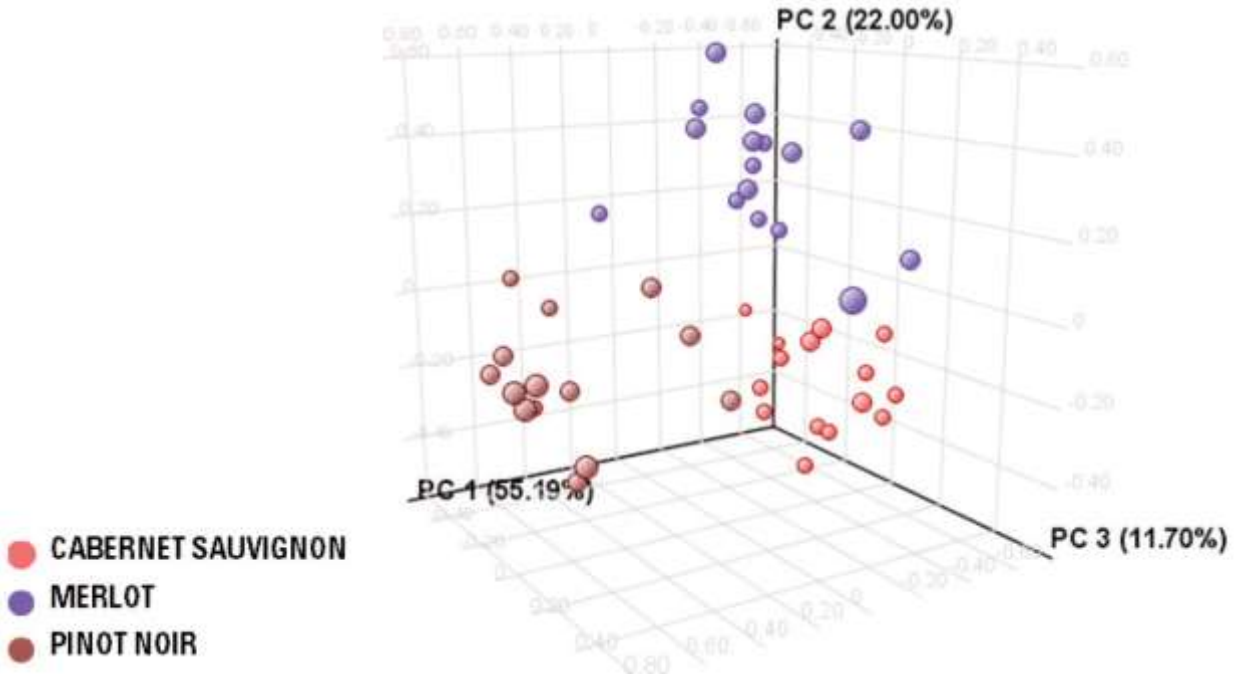
- Is this wine cabernet or pinot noir?
- Is this Olive Oil really "Extra Virgin"?
- Is this rice from Japan or from somewhere else?
- How did changing my growing or fermentation process affect my food product?

To answer these questions, we need
Food Profiling

Authenticity: Food Type

Determining Wine Varietal by LC/MS

- 45 red wine samples used to create a prediction model
  - 15 Cabernet
  - 16 Merlot
  - 14 Pinot Noir
- Wines sampled varied in geographic origin and vintage
- 5 additional wines which were not part of the original sample set were correctly classified using model



Application note **5990-8451**

Determining Wine Varietal by LC/MS

- **45 Red wines**
- **3 Varieties:** Cabernet Sauvignon (15), Merlot (16), Pinot Noir (14)
- **11 different countries:** Czech Republic, Slovakia, France, Italy, Macedonia, Bulgaria, Hungary, Australia, Chile, Germany EE.UU.
- **Harvests:** 2004 – 2008

SET OF VERY VARIED SAMPLES

Instruments used

Jet Stream ESI source

Multimode ion source

Eclipse Plus C18 (2.1×100, 1.8µm)

HILIC Plus C18 (2.1×100, 3.5µm)

Agilent Technologies
1200 RRLC system

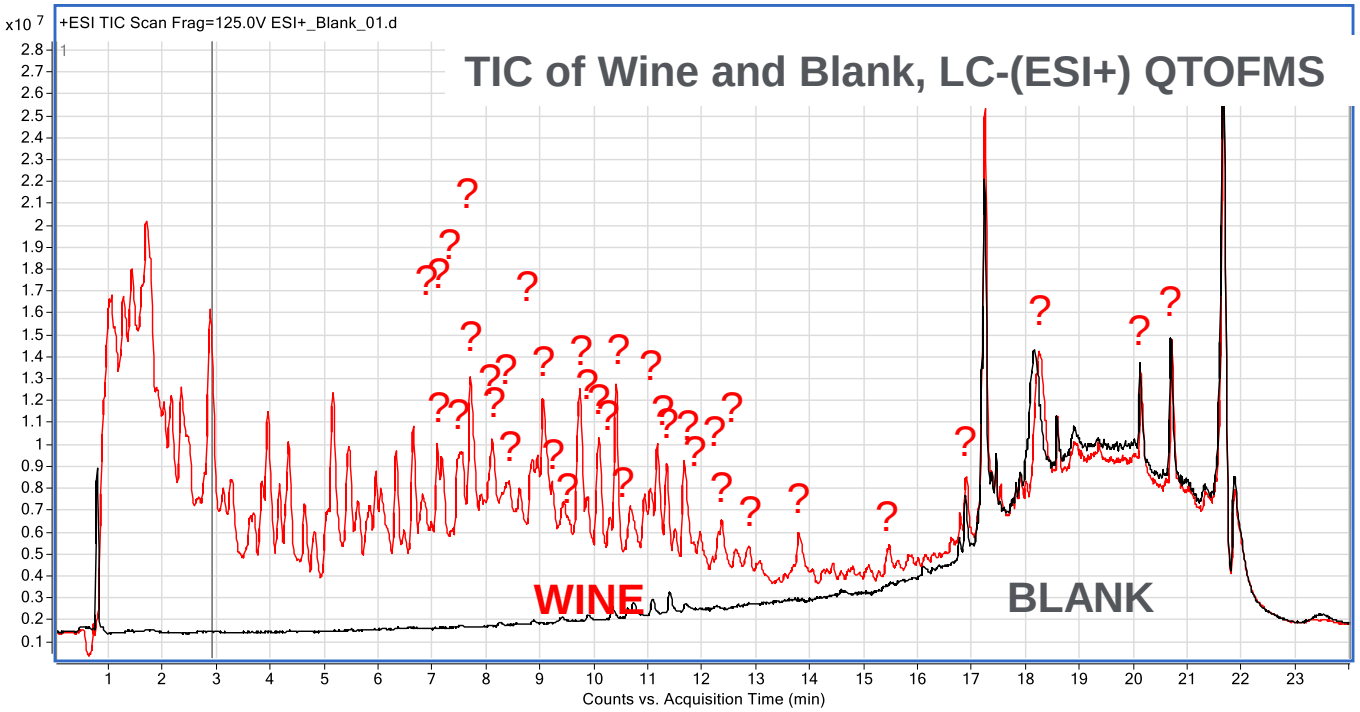
Agilent Technologies
6530 Accurate-Mass Q-TOF LC/MS

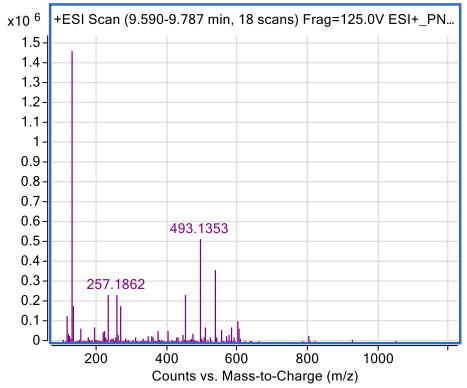
Micro-filtration is the only Sample Prep used

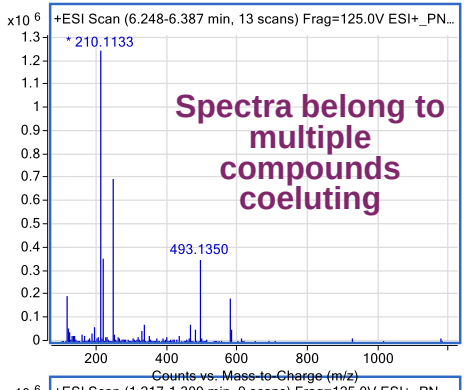
Ondrej Lacina[a], Lukas Vaclavik[a], Jana Hajslova[a], Jerry Zweigenbaum[b]

[a] Institute of Chemical Technology Prague, Czech Republic [b] Agilent Technologies, Wilmington, DE, USA

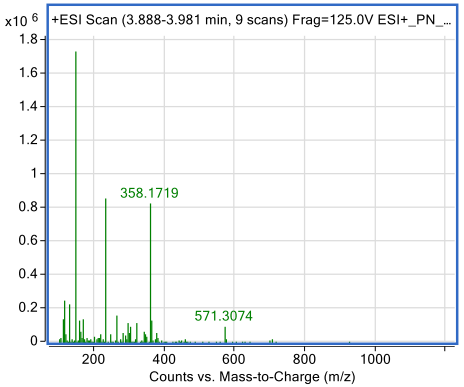
Data Processing:

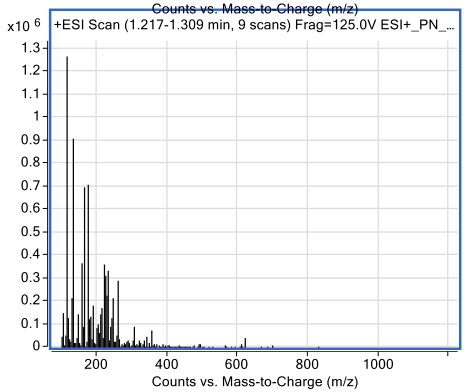






Spectra belong to multiple compounds coeluting

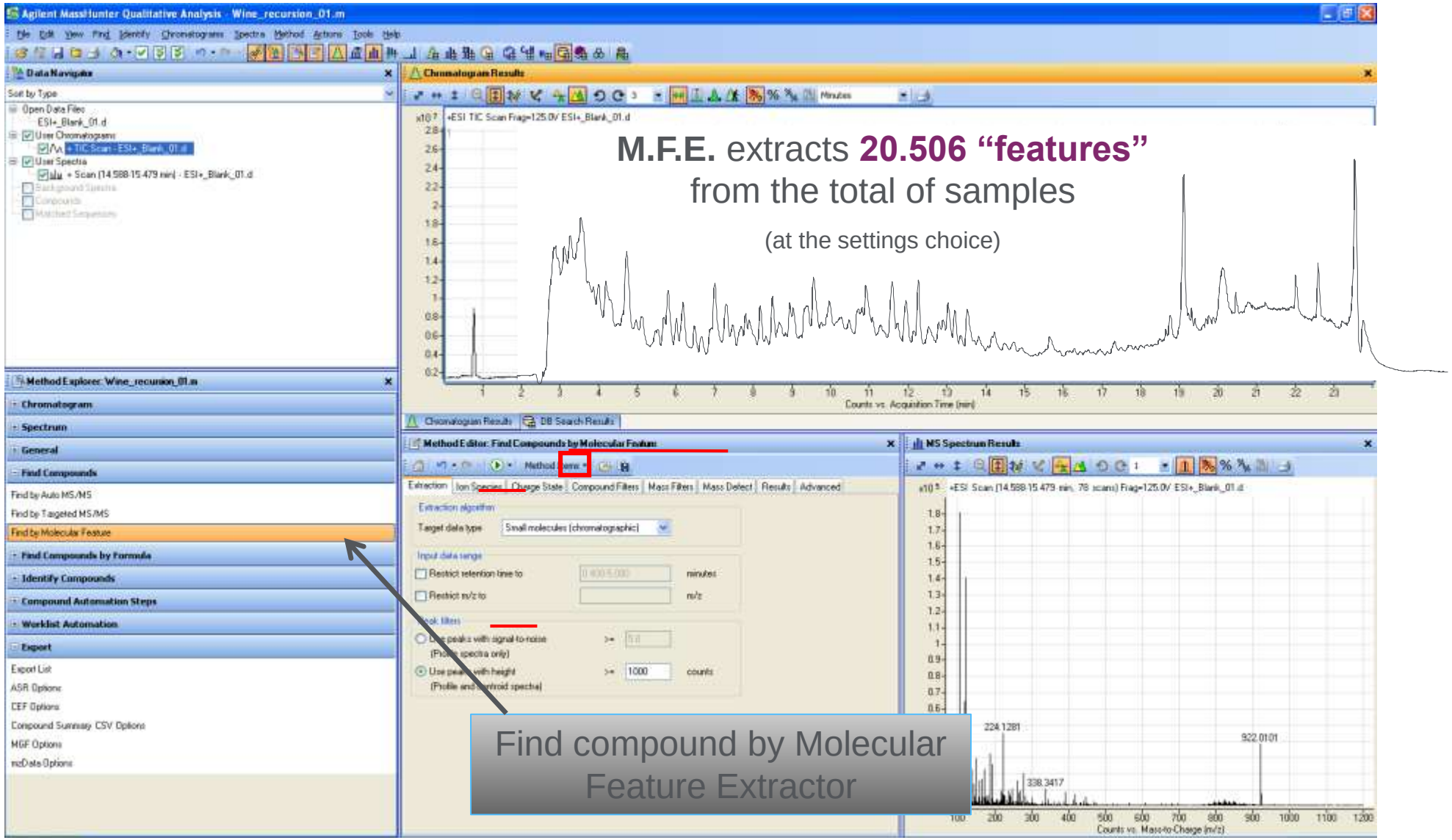




- **Very complex data sets.**
- **Masked minority compounds.**

A deconvolution software is needed to characterize all the ionized compounds.

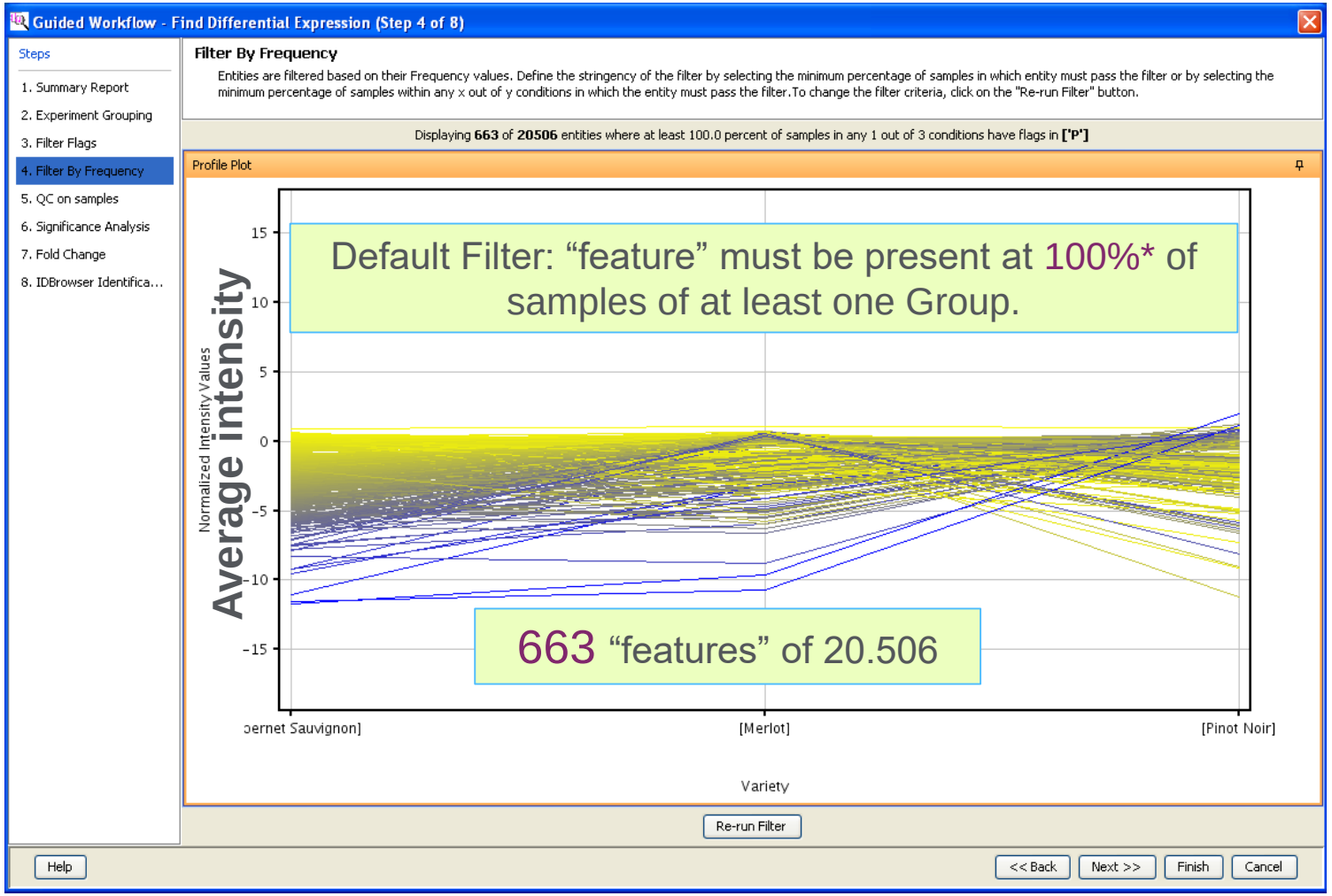
Data Extraction: “Find By Molecular Feature”



M.F.E. extracts **20.506 “features”** from the total of samples

(at the settings choice)

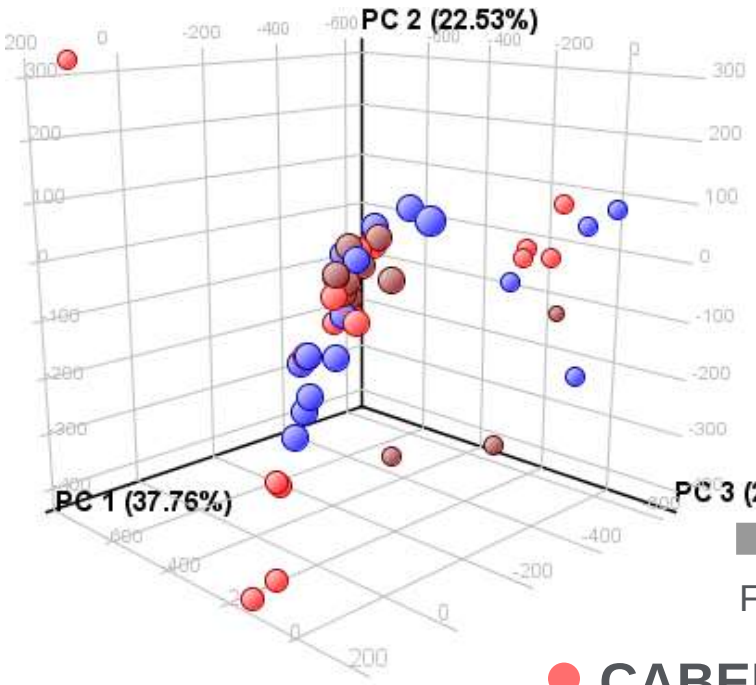
Data Processing with Agilent Mass Profiler Profesional: Filtering By Frequencies



*** Filtering by 50% “features” would increase from 663 to 3600**

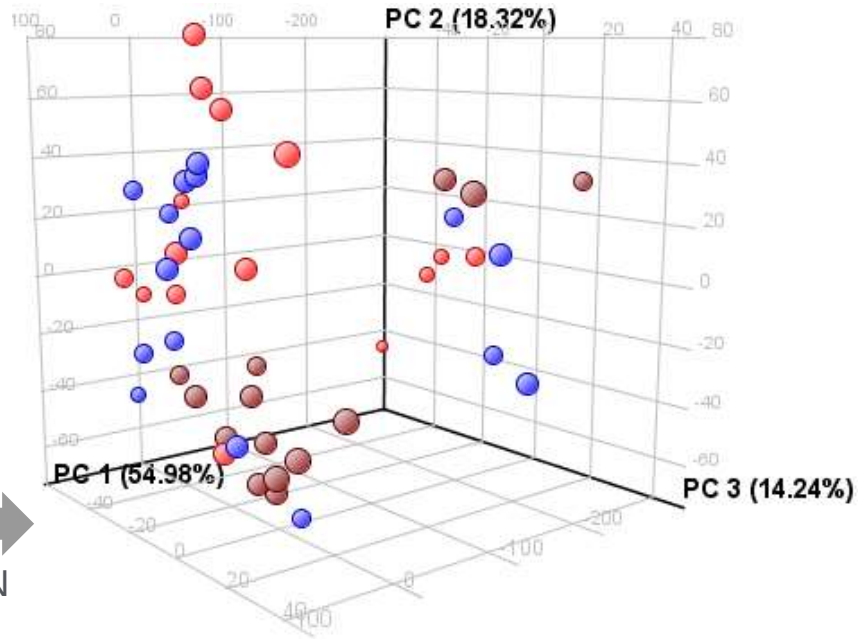
Data Processing with Agilent Mass Profiler Professional: Filtering By Frequencies + PCA

PCA of the data
Initial features (20506)



FILTRATION

PCA of the data
Features filtered by Frequencies (3600)



- **CABERNET SAUVIGNON**
- **MERLOT**
- **PINOT NOIR**

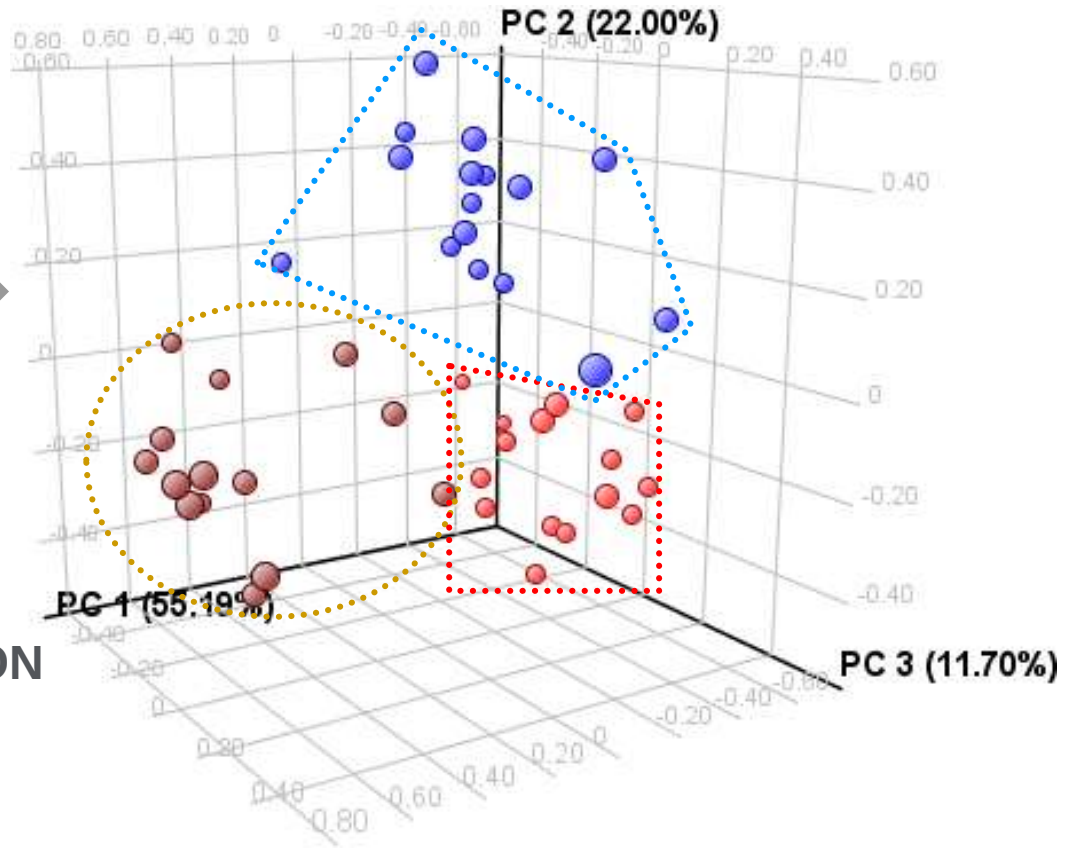
Data Processing with Agilent Mass Profiler Professional: Filtering By Frequencies + PCA + ANOVA

PCA of Data:

Features after ANOVA (p≤0.05) & Fold Change (≥2.0): 26

A good filter of data is crucial for a good Classification fit

FILTRATION



- CABERNET SAUVIGNON
- MERLOT
- PINOT NOIR

Class Prediction Model Validation

Number of samples used for model valiation: **45**

Class Prediction (Step 3 of 5)

Validation Algorithm Outputs

The validation tables provide the result of the model validation step. The prediction is compared with the true values of the samples. If many mistakes are made in the prediction, press the "Back" button to make changes to the model.

Confusion Matrix

	[Cabernet Sau...	[Merlot] (Predi...	[Pinot Noir] (Pr...	Accuracy
(True) [Cabernet Sauvignon]	15	0	0	100.000
(True) [Merlot]	1	14	1	87.500
(True) [Pinot Noir]	0	0	14	100.000
Overall Accuracy				95.556

- During model validation, 2 MERLOT samples were incorrectly classified.
- All the Cabernet Sauvignon & Pinot Noir were correctly classified.
- **The prediction reliability of the model determined to be of 95.6%***

The model classified correctly 5 of the blind samples (2 CS, 1M, 2 PN).

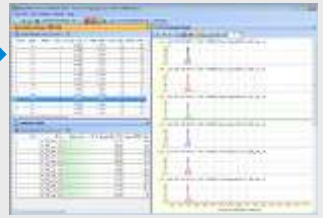
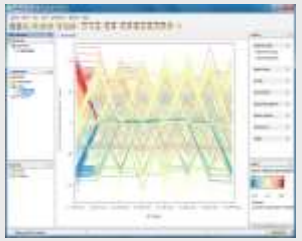
* ANOVA used $p \leq 0.05$.

Food Authenticity Analysis with MPP and MassHunter Classifier

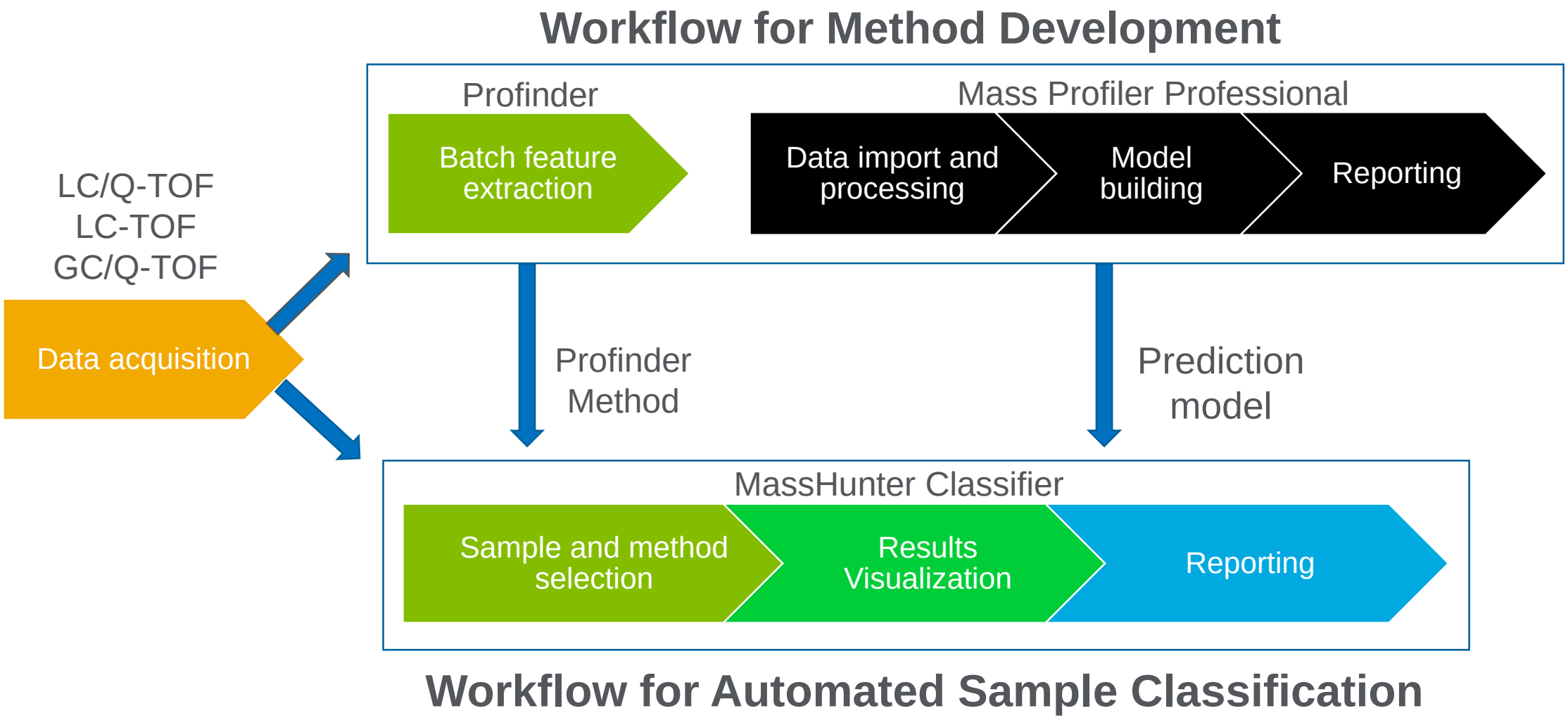
Challenges in Food Authenticity Testing

- Meeting regulatory requirements
- Availability of authentic samples
- Speed of analysis
- Extensive method development required
- Extensive validation required

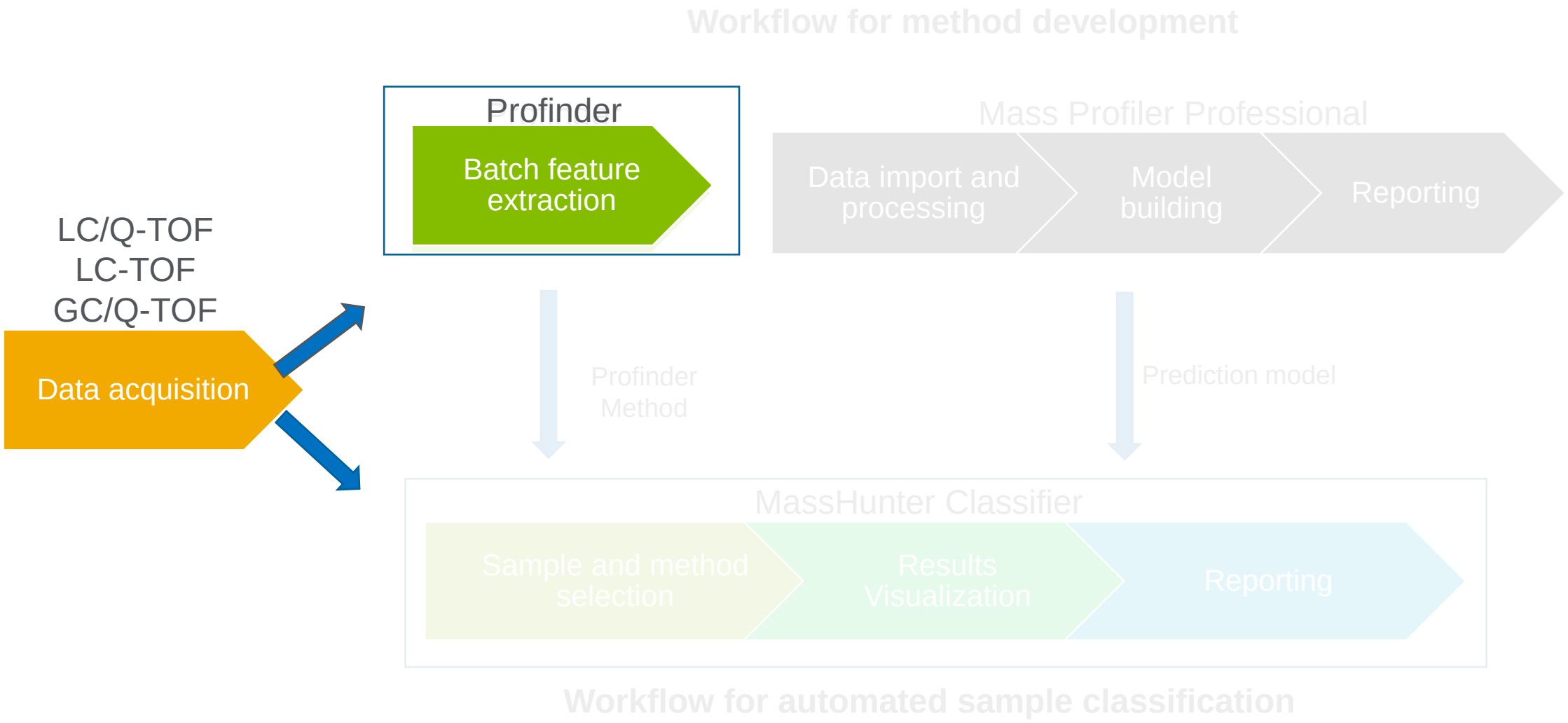
Agilent’s Food Authenticity Workflow

Sample Preparation	Separation	Separate and Detect	Feature Finding	Sample Classification
QuEChERS and Clean up materials Multi-class QuEChERS concept	Infinity II Series 2DLC Broad column portfolio	 LC/MS  GC/MS	 MassHunter Profinder  MassHunter Unknowns Analysis	 Mass Profiler Professional (MPP)

Food Authenticity Workflow



Food Authenticity Workflow



Method Development in MassHunter Profinder

Batch Feature Extraction

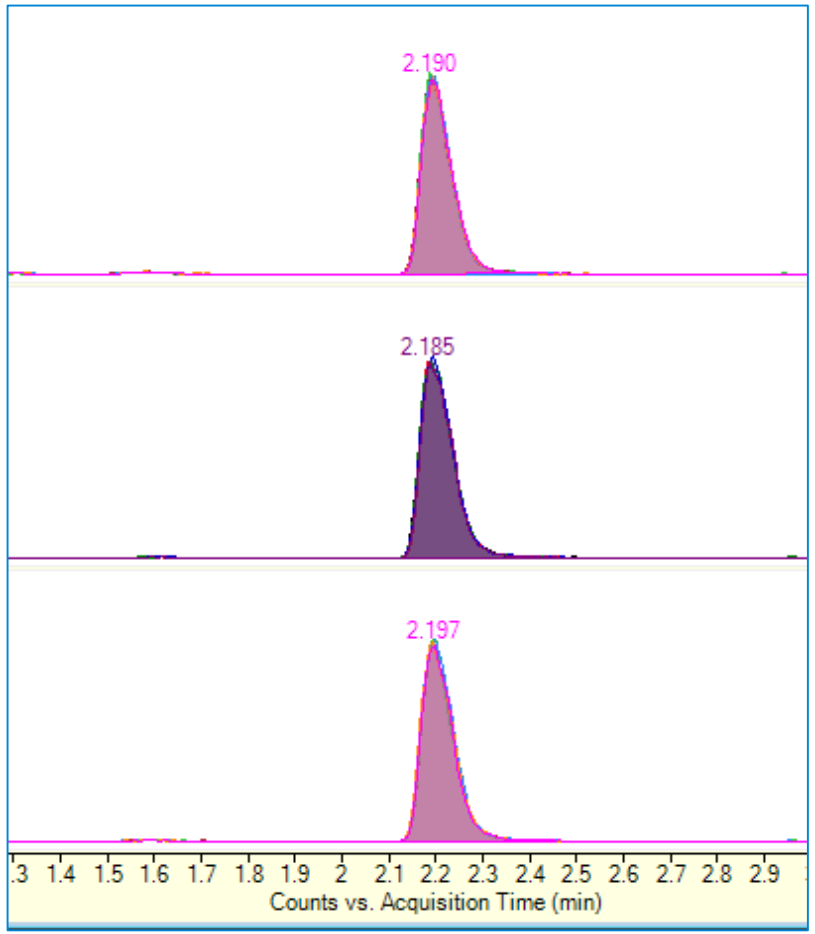
Import sample files in Profinder

↓

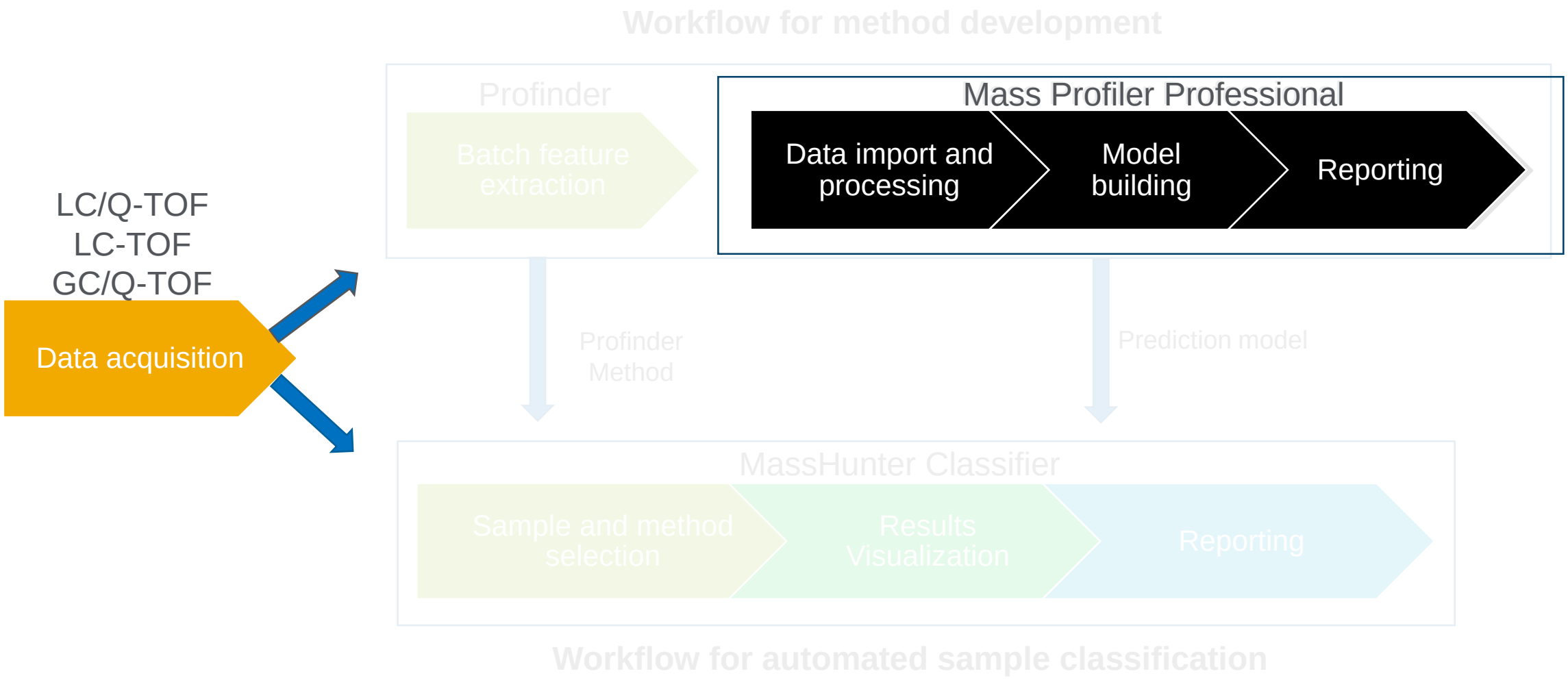
Select Batch molecular feature extraction
or Batch recursive feature extraction for small molecules

↓

Build and save the method file



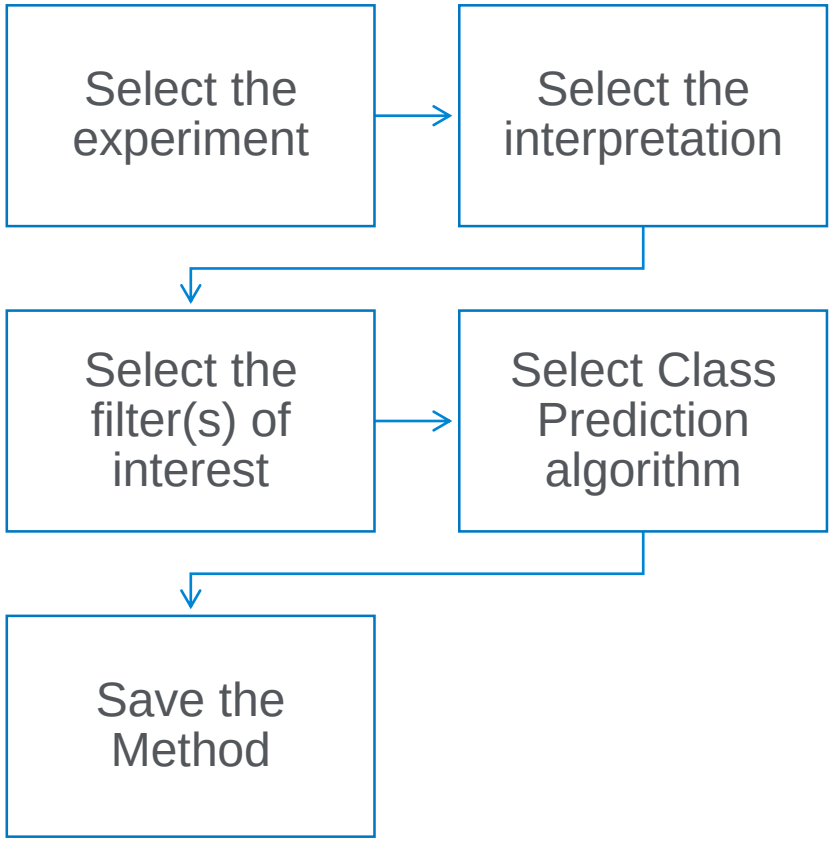
Food Authenticity Workflow



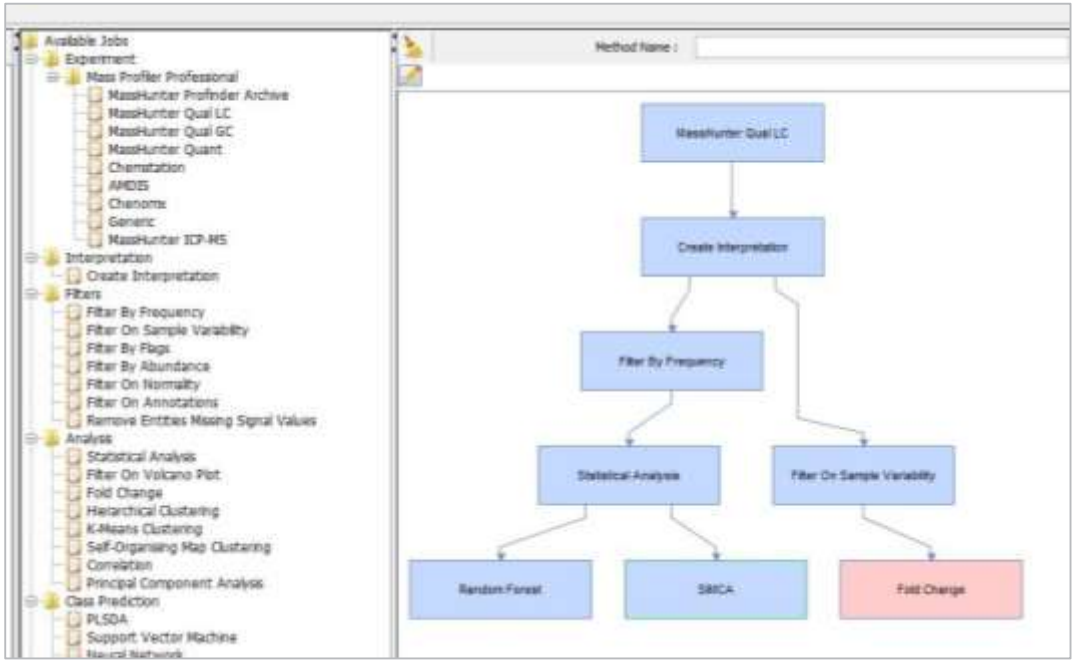
Method Development in MPP

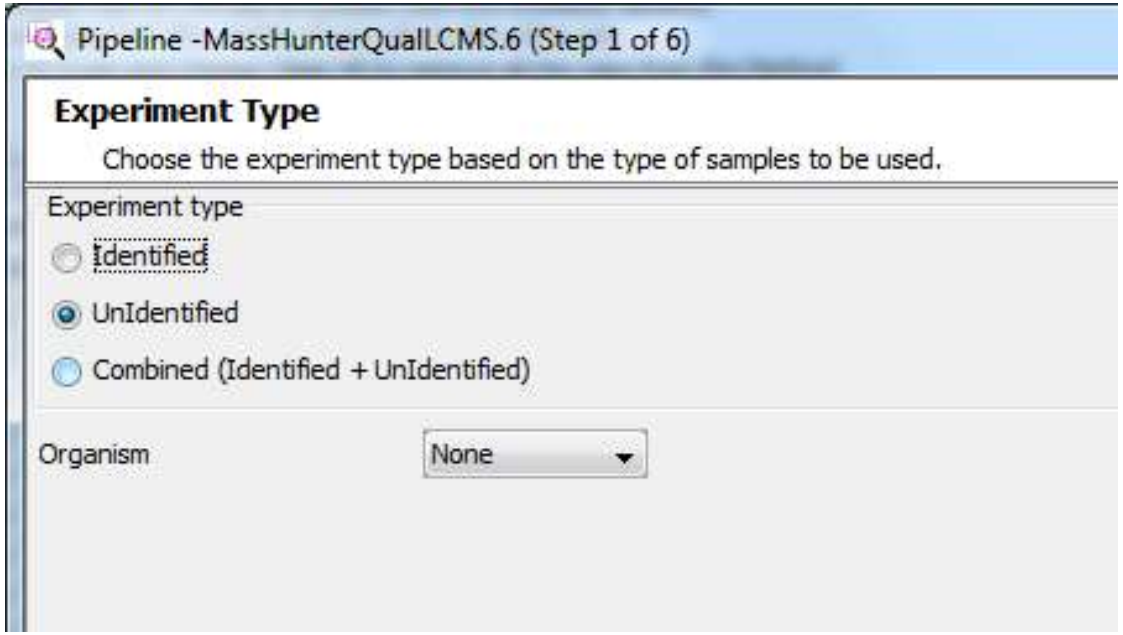
Method manager

- Allows creation, running and monitoring of a method
- Drag and drop the tasks of choice in the drawing area to create a method
- Advanced features like normalization, missing value imputation, etc. can also be configured from the same user interface
- Method can be started as well as exported from the tool
- Exported method gets saved as .m file and can be shared



Creating a Method





Select the tasks of choice and connect them to create a method

Select advanced parameters

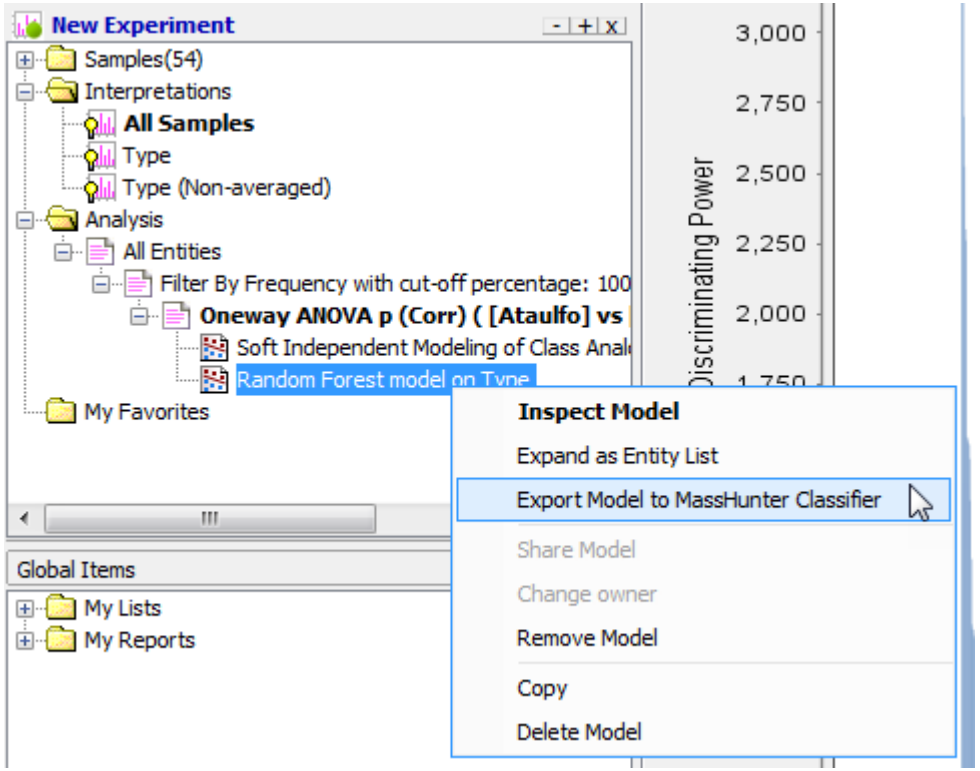
Class Prediction Algorithms

- Random Forest
- New Linear Discriminant Analysis
- SIMCA
- Partial least Squares Discrimination
- Support Vector Machine
- Naive Bayes
- Decision Tree
- Neural Network

Export Model to MassHunter Classifier

Exported model is saved in method folder for further use in MassHunter Classifier

For mango data, a prediction model using Random forest was created.



Automated Sample Classification

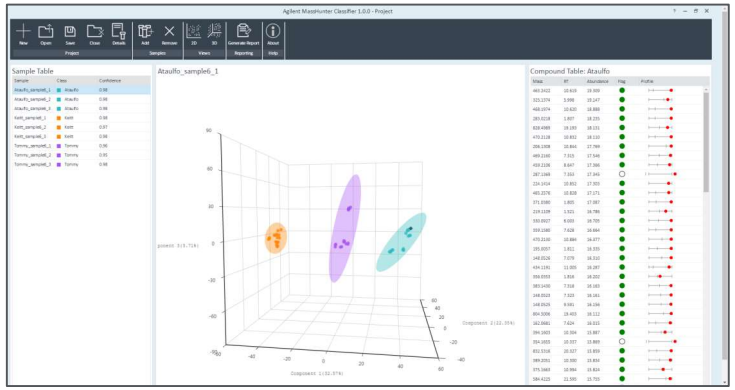
MassHunter Classifier

Automated Sample Classification Workflow

Sample collection

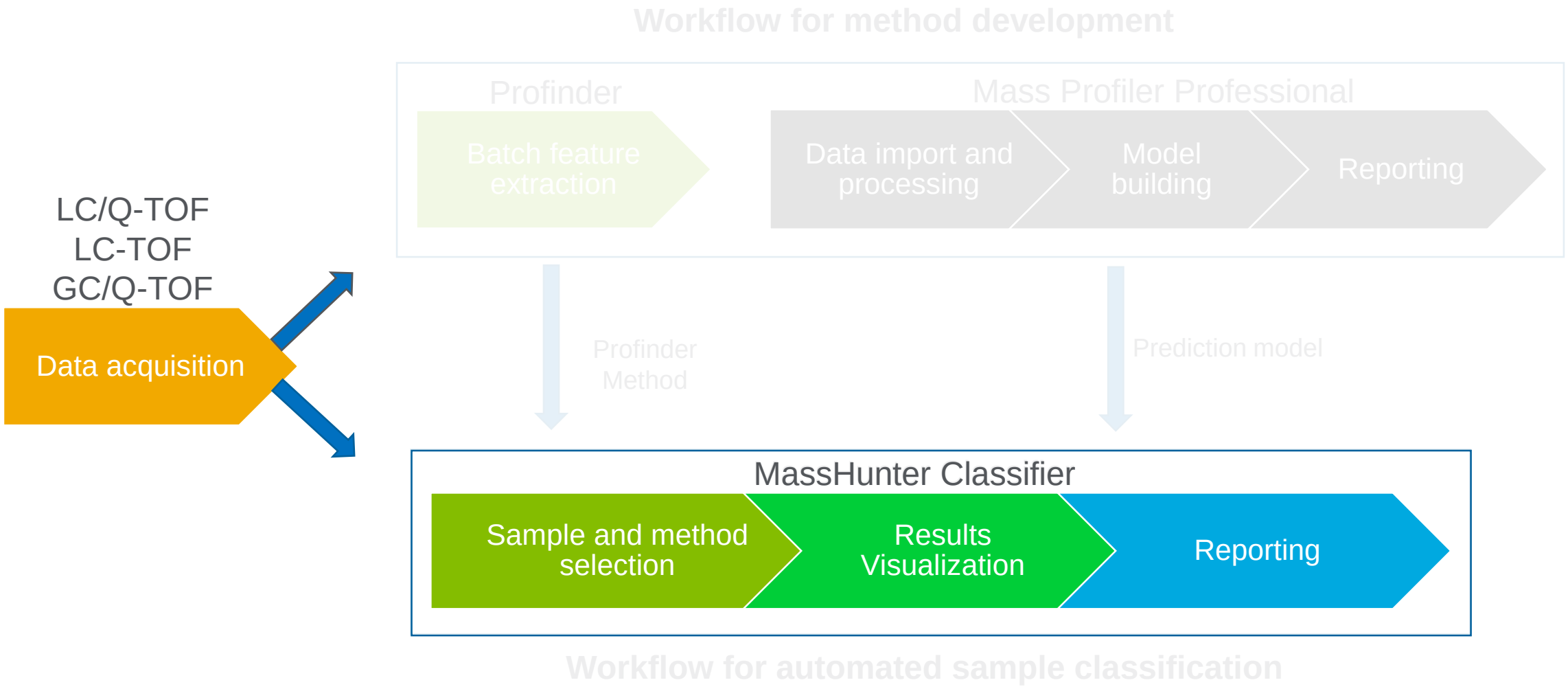
Sample preparation

Detection

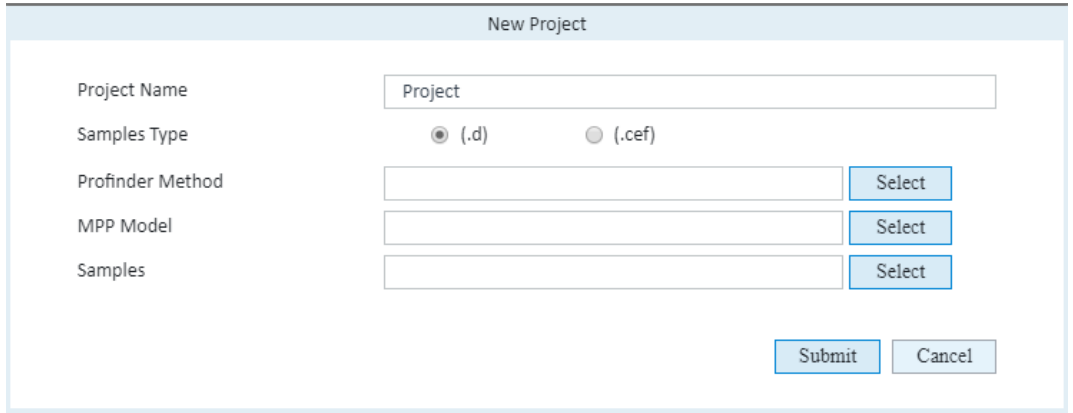


MassHunter Classifier

Food Authenticity Workflow



Project Creation in MassHunter Classifier

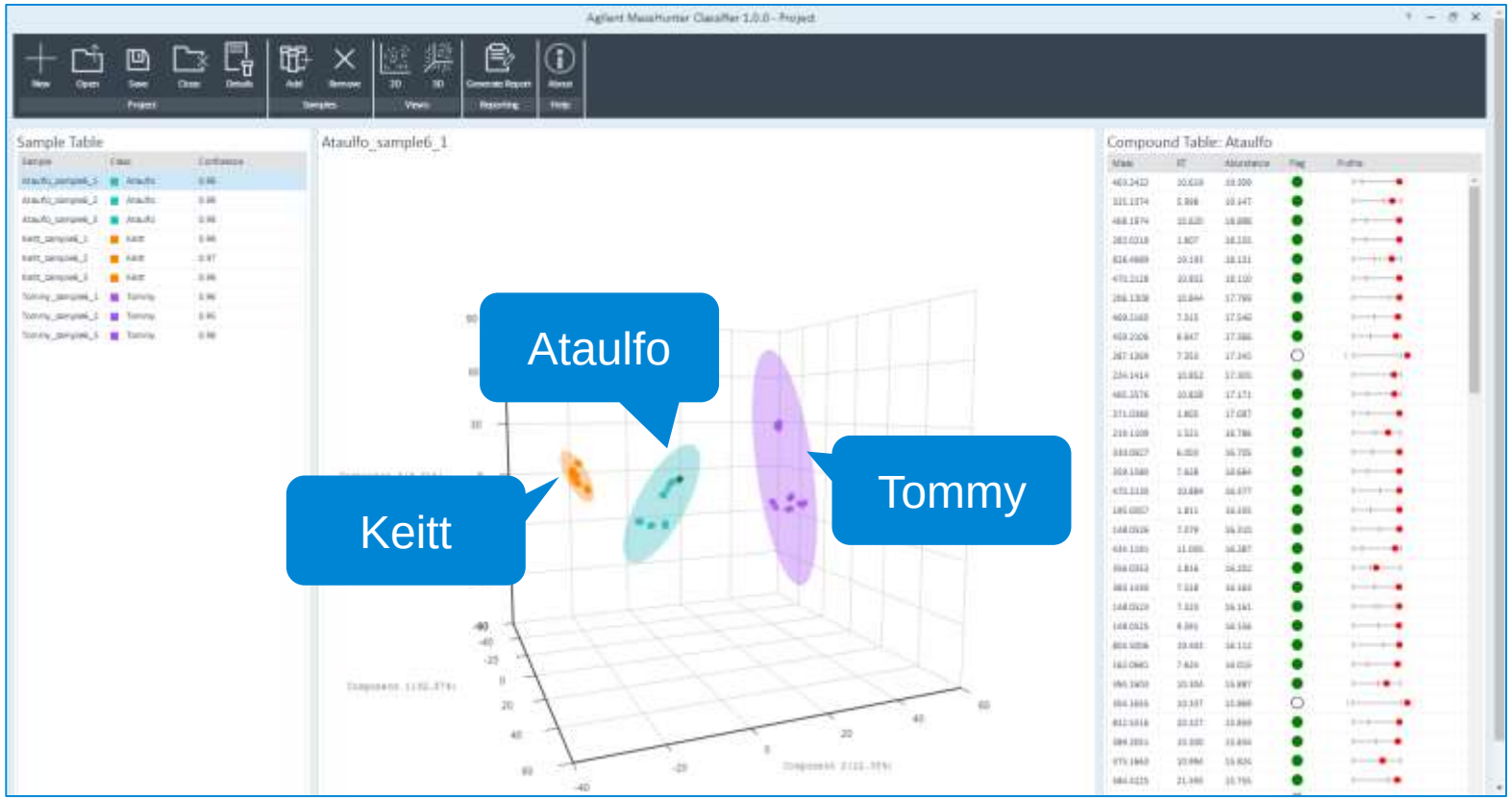


Import Method, Model and Samples

- Profinder method is required when user is working with .d files
- At a time only one MPP model can be imported. To run prediction with another model, a separate project can be created.
- One or more unknown samples can be imported for prediction

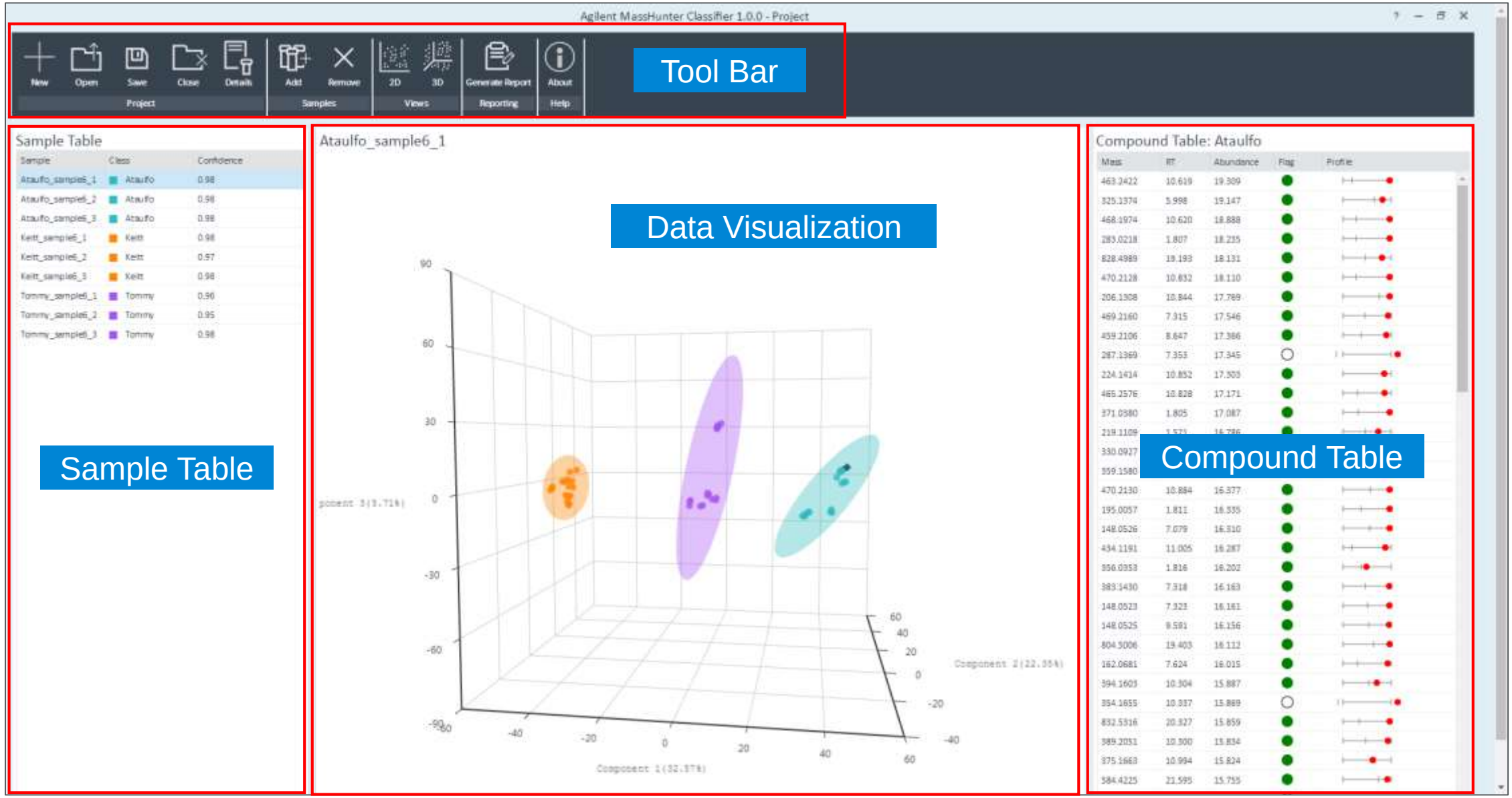
MassHunter Classifier

A simple tool for automated sample classification



- Simple, uncluttered user interface for routine analysis
- Run batch feature extraction and class prediction on previously created classification models
- The tool quickly processes the samples to be classified and user is presented with visualization options such as Principal component Analysis (PCA) to review the results
- User can interactively review the sample classification results and associated compound data before choosing to save a PDF report

MassHunter Classifier User interface



Sample Table

Sample Table

Sample	Class	Confidence
Ataulfo_sample6_1	Ataulfo	0.98
Ataulfo_sample6_2	Ataulfo	0.98
Ataulfo_sample6_3	Ataulfo	0.98
Keitt_sample6_1	Keitt	0.98
Keitt_sample6_2	Keitt	0.97
Keitt_sample6_3	Keitt	0.98
Tommy_sample6_1	Tommy	0.96
Tommy_sample6_2	Tommy	0.95
Tommy_sample6_3	Tommy	0.98

- Sample table shows the predicted class to which unknown sample belongs and the confidence score
- All the samples here are classified with higher than 95% confidence

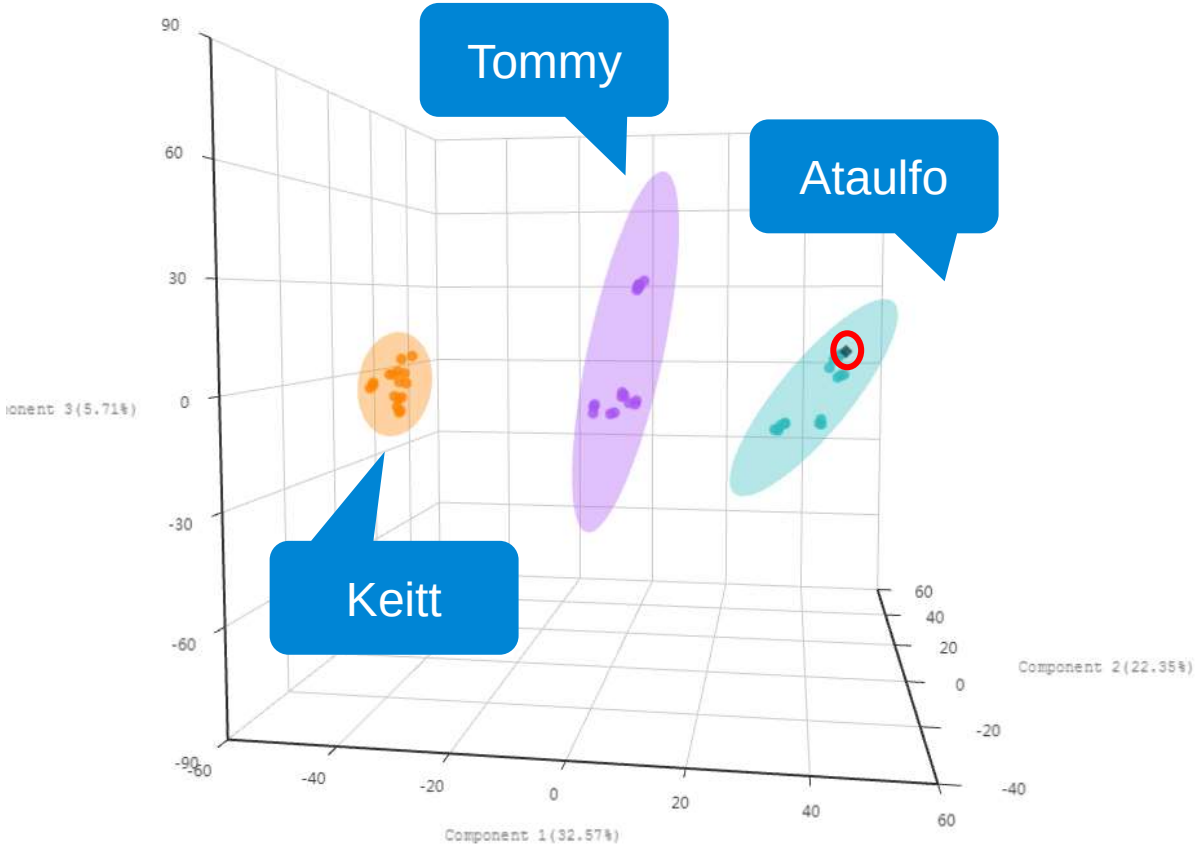
Compound Table

- Compound table shows the list of metabolites used in classification
- Flag column provides information on if a compound contributed to sample classification
- Additional information on compound is seen in profile plot
- Red dot indicates the abundance of the compound in the unknown sample with respect to the abundance distribution of the compound in the training samples for the predicted class

Compound Table: Ataulfo

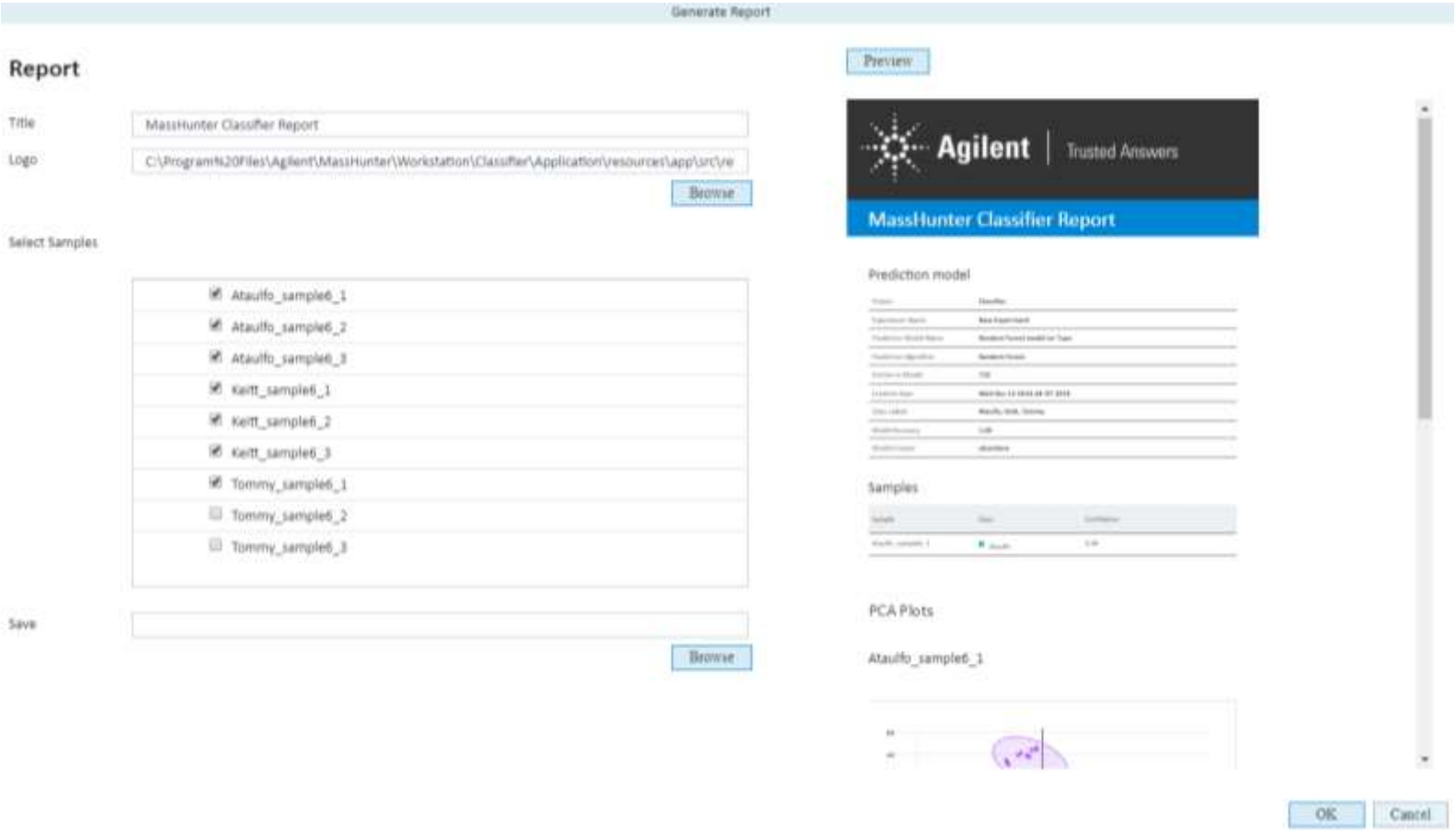
Mass	RT	Abundance	Flag	Profile
463.2422	10.619	19.309	●	
325.1374	5.998	19.147	●	
468.1974	10.620	18.888	●	
283.0218	1.807	18.235	●	
828.4989	19.193	18.131	●	
470.2128	10.832	18.110	●	
206.1308	10.844	17.769	●	
469.2160	7.315	17.546	●	
459.2106	8.647	17.366	●	
287.1369	7.353	17.345	○	
224.1414	10.852	17.303	●	
465.2576	10.828	17.171	●	
371.0380	1.805	17.087	●	

3D-PCA



- Confidence ellipses in the PCA are drawn for model training data
- Dots on the PCA plot are training samples used for model building
- Diamond structured sample is the unknown Ataulfo sample in 3D PCA plot

Report Generation



- Information on classified samples can be recorded in a report
- Captured information includes name of model used, samples tested as well as per sample data in a PDF file

Summary

- One stop solution for food authenticity analysis
- Streamlined workflow for batch processing and model building
- Class prediction algorithms to support various applications
- Brand new simplified application for automated sample classification
- Separation in “Scientist” part and “Technician” part
- Standalone software which utilizes our already established platform of Profinder and MPP

Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - *Preparación de datos previa al Análisis Estadístico diferencial.* ***Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder"****.*
  - *¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil….* ***Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

From experiment design to conclusions, a long way to help scientists. Agilent Tools and workflows to better decisions making at Integrated Biology :

Complementing different analytical technologies LCMS, GCMS, CEMS, ICPMS. Data Acquisition modes.

- For LC HRMS, at first stage, where ID is still not necessary, Full Scan Acquisition or No Data Dependent MS/MS such All Ions is generally the choice of acquisition mode. At second stage, where more ID confidence is needed, a Target acquisition mode could be of interest to get reliable and pure MS/MS spectra of all differential compounds.

- Identification of differential compounds is one of the biggest CHALLENGE, a real bottleneck. Use of analytical technology with strong ID power is crucial to afford such challenge.

- For LCMS & CEMS where Ionization sources are based basically on ESI, **HRMS** is key due to their Qualitative power. (AM, IP, MSMS, *CCS*).

- For GC techniques where Electron Impact ionization (EI) is such an universal technique with universal libraries, a Single Quadrupole is a very robust and convenient technique. Recent developments of **soft EI** allows to keep some Molecular Ion intact so GCQTOF is also an interesting technique combining EI spectra and HRMS advantages.

- For ICPMS, due to the very low list of possible compounds (periodic table) it is feasible to perform Target Acquisition for all the elements.

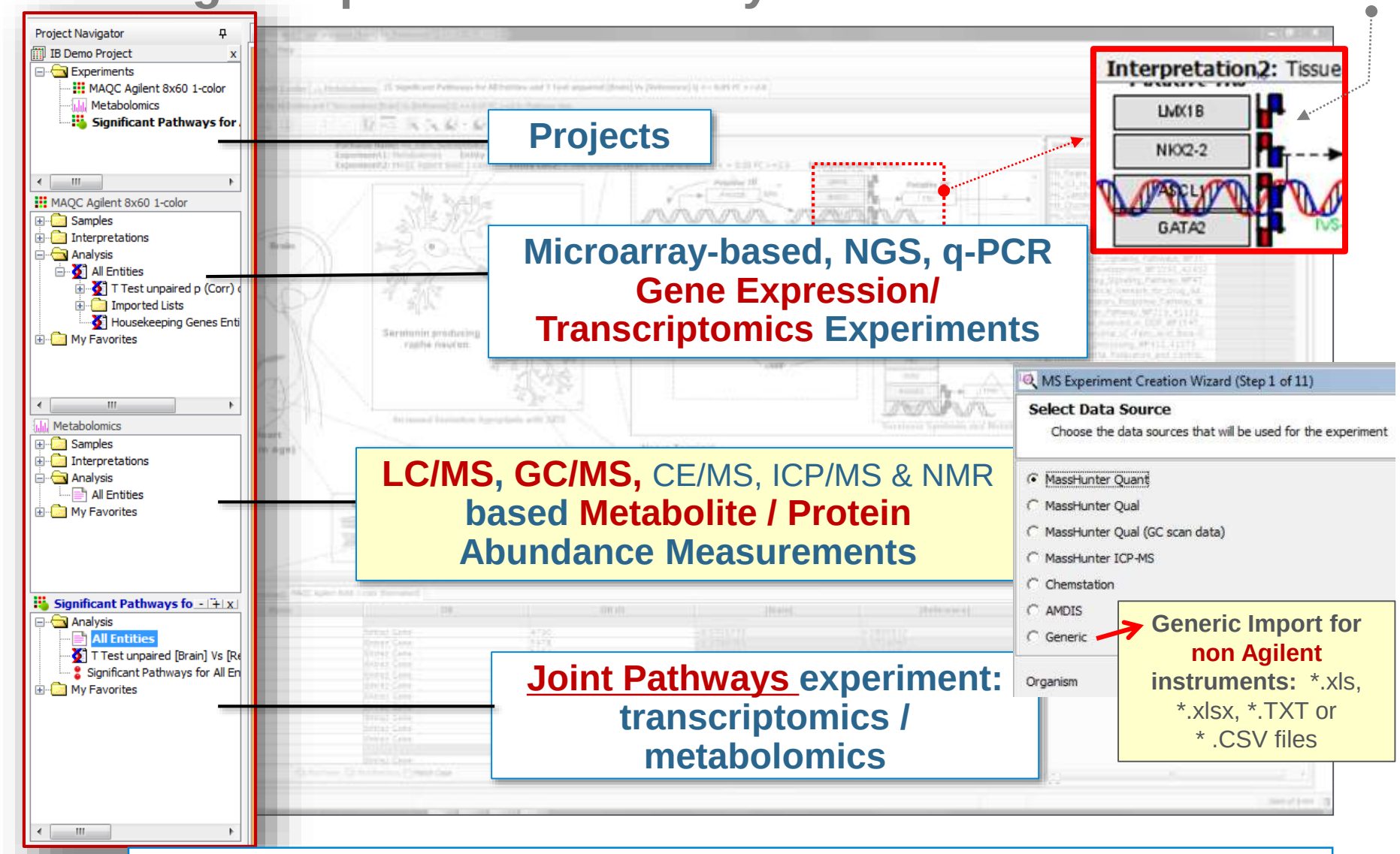
From experiment design to conclusions, a long way to help scientists. Agilent Tools and workflows to better decisions making at Integrated Biology :

Complementing different analytical technologies LCMS, GCMS, CEMS, ICPMS. Data Acquisition modes.

- The choice of separation technique is biasing the kind of compounds we will find in terms of polarity.
- Volatile compounds would be difficult to get robust information on a LCMS system as well as other non-polar compounds. GC is in this case the best choice.
- In the other hand, highly polar compounds would be difficult to retain on a regular RP column. An alternative is to use HILIC columns or Capillary Electrophoresis (CE).
- For a comprehensive project considering different techniques it is interesting to have a **Tool** who can handle all these different experiments on the same Data Treatment project pointing out same organism.

Multi-Omics Open Platform: Mass Profiler Professional

Expression changes represented directly on routes



Projects

Microarray-based, NGS, q-PCR Gene Expression/ Transcriptomics Experiments

LC/MS, GC/MS, CE/MS, ICP/MS & NMR **based Metabolite / Protein Abundance Measurements**

Joint Pathways experiment: transcriptomics / metabolomics

Generic Import for non Agilent instruments: *.xls, *.xlsx, *.TXT or * .CSV files

Enrichment Analysis on curated pathways and computationally – derived networks

Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - *Preparación de datos previa al Análisis Estadístico diferencial.* ***Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder"****.*
  - *¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil….* ***Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

Deconvoluting Data and visualization tools

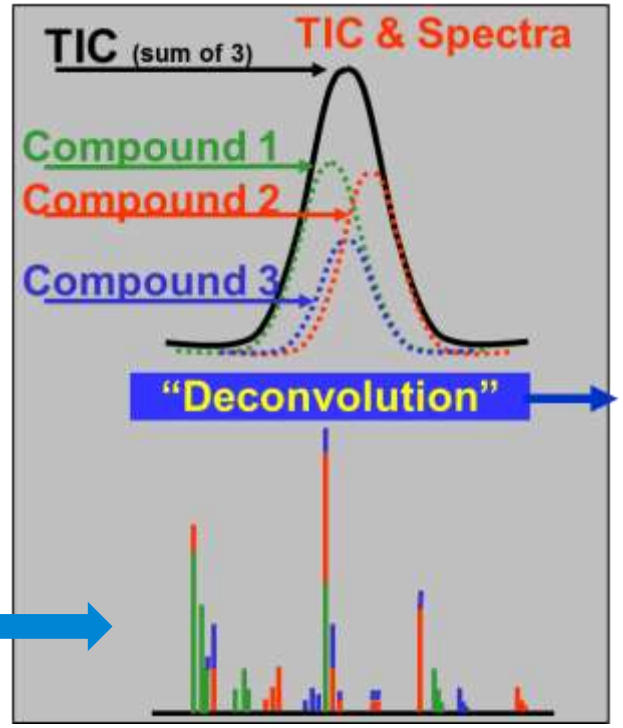
How does Agilent algorithms get compounds/features lists from a FullScan acquisition.

Due Full Scan acquisition, a **deconvolution technique** is needed in order to characterize all possible compounds eluted and ionized on the source.

Different Ionization sources (ESI for LC/CE & EI for GC) need different type of algorithms to extract and characterize compounds (features).

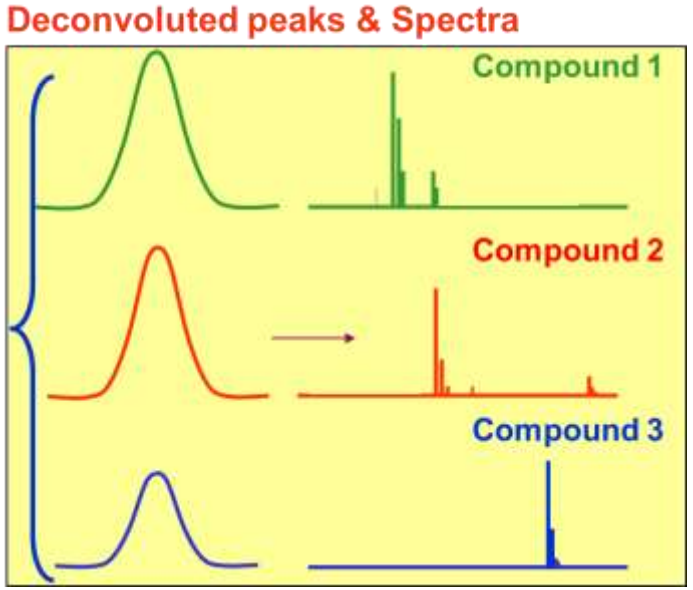
Electron Impact (GCMS) is a destructive ionization technique so Deconvolution is based on EI fragments.

The black TIC has three components underneath it. If you take a spectrum at the apex of the black TIC peak, you will see a mixture of three components, as the spectrum shown here on the right with all the green, red, and blue colors.



Mass Hunter Qual & Quant have a dedicated Deconvolution Algorithm for this type of data, an improved AMDIS version.

The MassHunter Deconvolution can pull out these individual components from the total ion chromatogram. So after deconvolution, we can get clean spectra of the matrix, the interference, and the target compound as shown on the right side of the slide.

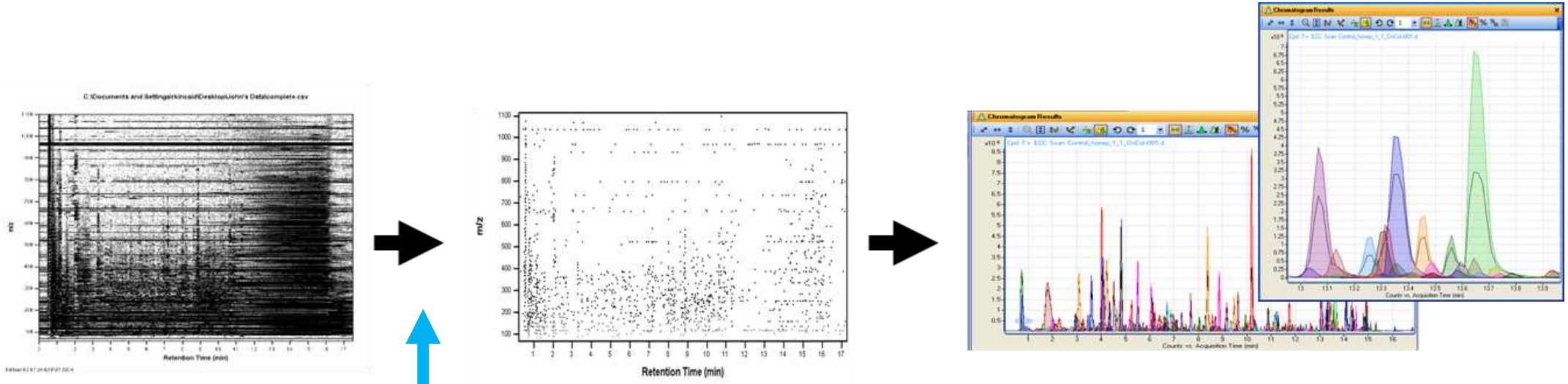


Deconvoluting Data and visualization tools

How does Agilent algorithms get compounds/features lists from a FullScan acquisition.

Due Full Scan acquisition, a **deconvolution technique** is needed in order to characterize all possible compounds eluted and ionized on the source.

For LC HRMS Agilent developed **Molecular Feature Extraction** (MFE).
MFE is an advanced deconvolution algorithm working at both spectra and chromatographic level.



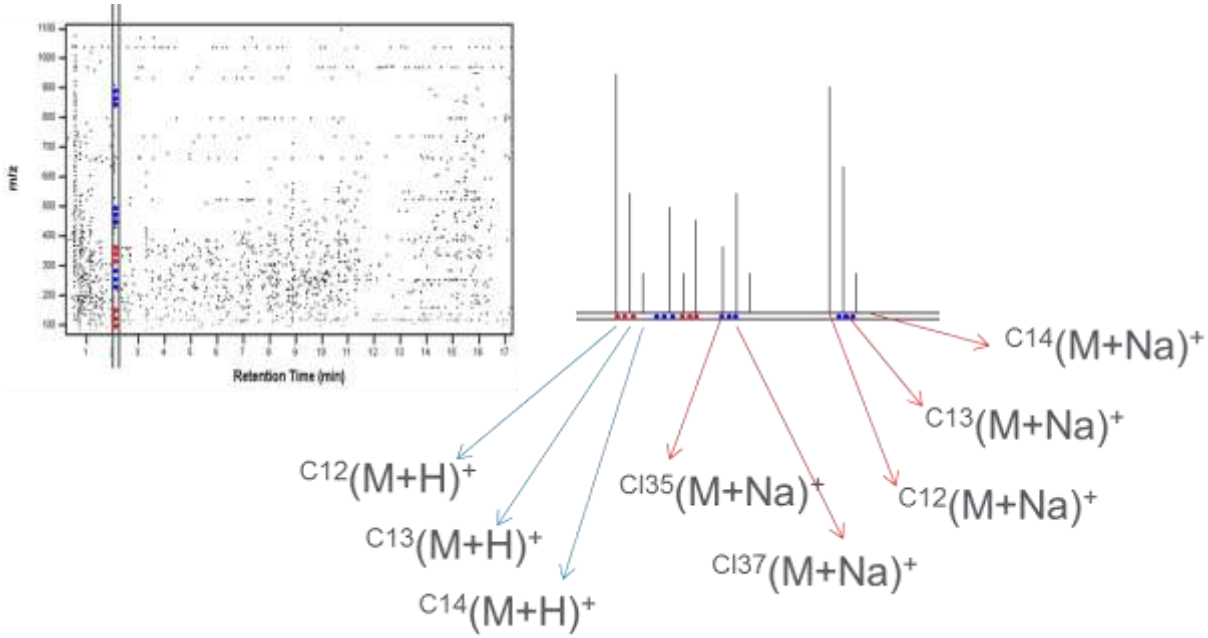
First step is to remove all the m/z which are present along the full chromatogram. This is fix background.

Deconvoluting Data and visualization tools

How does Agilent algorithms get compounds/features lists from a FullScan acquisition.

Molecular Feature Extraction (MFE).

MFE is an advanced deconvolution algorithm working at both spectra and chromatographic level.



Extraction Works in 3 dimensions : Abundance, RT, m/z

Groups ionic species of same compound based on accurate mass.

Combines signals with chemical relationship (isotopes, adducts, multiple charges, dimers) and "molecular features" (= compounds)

Creates a list of mass intra spectra.

Mass
1071.50627
1050.40797
829.39682
885.40936
1137.49286
1137.48985
1442.63618
1290.59812
1033.4674
664.36901
784.39024
299.37294
299.70937
897.48015
745.33915
973.45314
1672.76497
688.36669
1531.78037
897.38411
788.46911
1501.6093
1398.6487
1248.6202
921.48555
757.41799
911.40588
1120.46384
1748.66086

Different smart settings are used for data mining and noise discrimination.

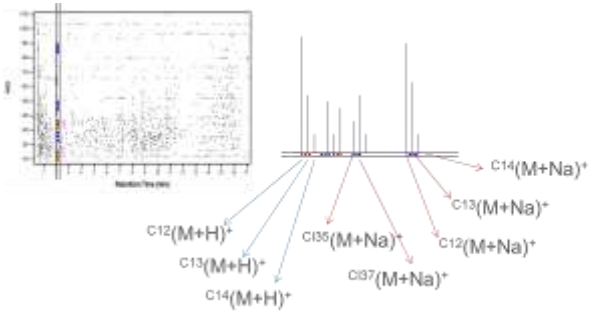
This process is done spectrum by spectrum, only those mass who show a gaussian peak along the time are considered as possible compounds (features)

Deconvoluting Data and visualization tools

How does Agilent algorithms get compounds/features lists from a FullScan acquisition.

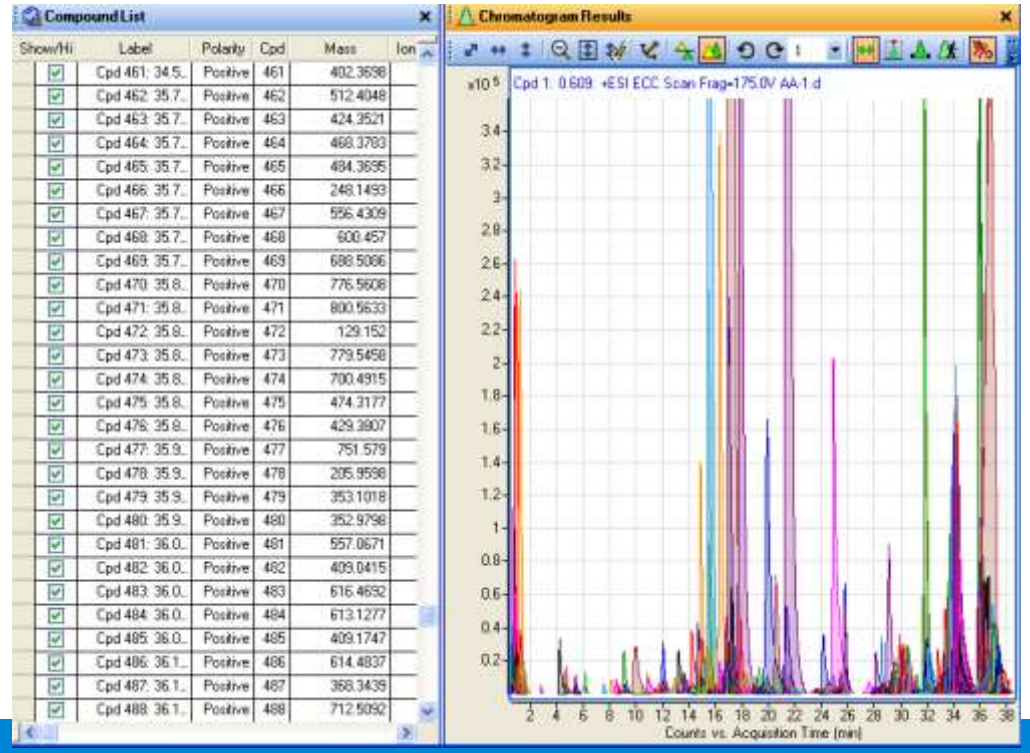
Molecular Feature Extraction (MFE).

MFE is an advanced deconvolution algorithm working at both spectra and chromatographic level.



MFE creates a list of possible compounds (features) characterized by Rt, Abundance, Acc.Mass & Isotopic Pattern.

This data is ready for Chemometric processing among different groups.



Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - *¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil….* ***Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

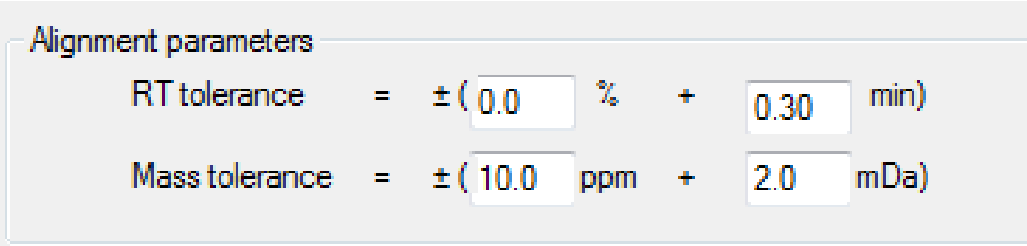
Data prepare for Statistical Analysis.

Alignment, Normalization, Baselining with Mass Hunter ProFinder.

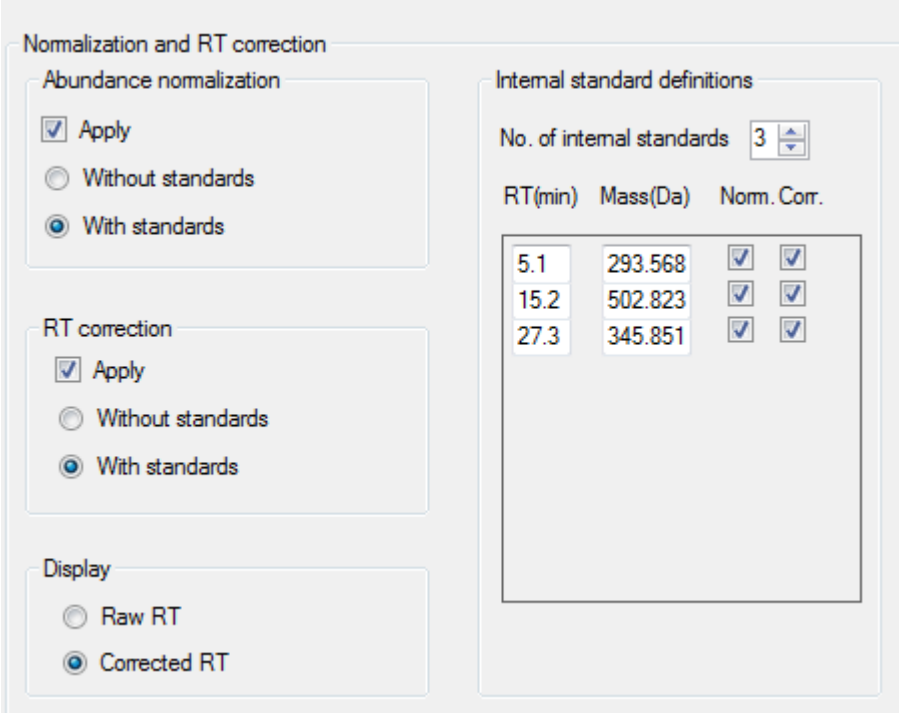
Statistical analysis requires for multiple replicates, both technical and natural

This statistical test requires to prepare data before tests.

Alignment of **Rt** and **Mass** along the different replicates is mandatory to avoid to skip possible compounds due to some instrumental deviation.



But also Normalization and Rt correction can be performed using Standards or Not.

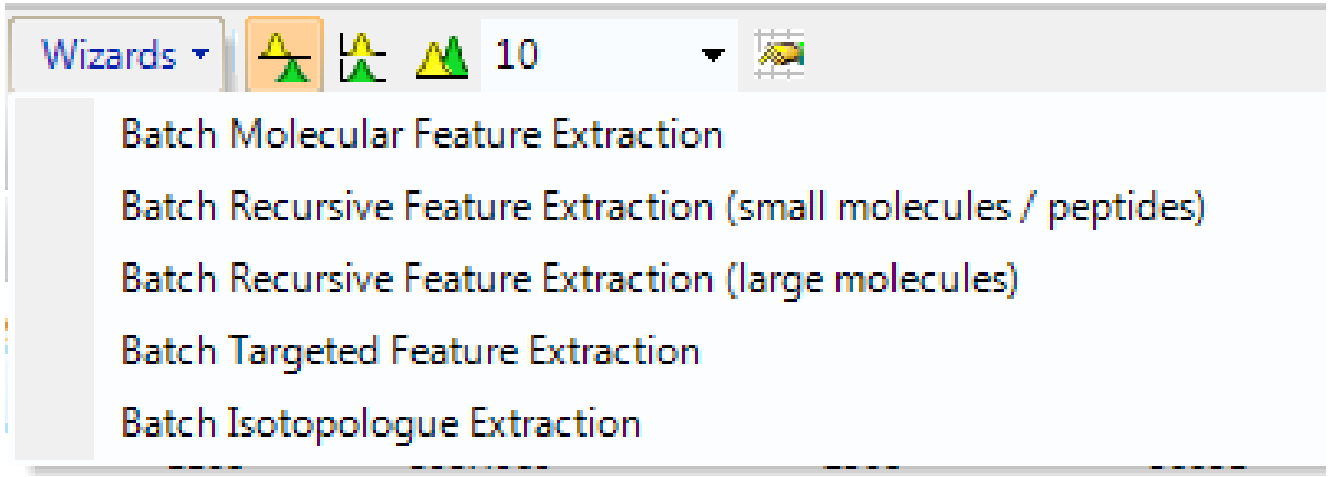


Data prepare for Statistical Analysis.

Alignment, Normalization, Baselining with Mass Hunter ProFinder.

Statistical analysis requires for multiple replicates, both technical and natural

Mass Hunter Profinder is a NEW advanced smart tool combining Deconvolution, Data prepare and Recursive Analysis for Small or Large Molecules. It is a Wizard assisted program



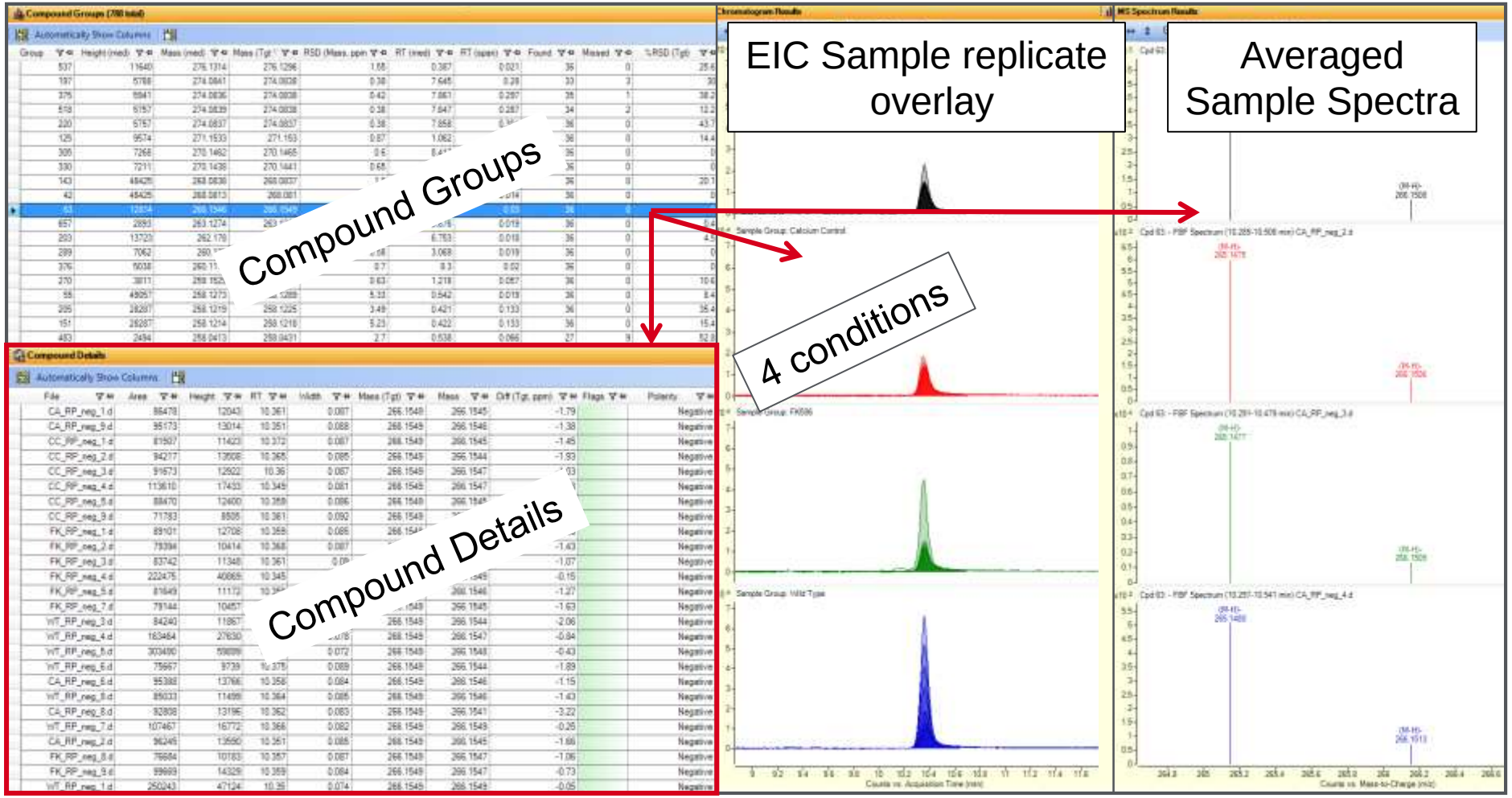
Automatically Deconvolutes with the right technique depending on Data loaded :

- Fragments deconvolution for GCMS
- MFE for LCQTOF

If desired Aligns, Corrects, Normalize and **makes a recursive Analysis**.

Four Profinder Windows:

Compound centric visualization and editing of results



Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - ***¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil…. Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - *Mass Profiler professional.* ***Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

Do I need recursive Analysis?

Please, make it easy … Comprehensive Data treatment to avoid missing compounds.

Why should I perform Recursive Analysis?

MFE extracts many possible compounds (features) using some filter settings to avoid noise and artifacts.

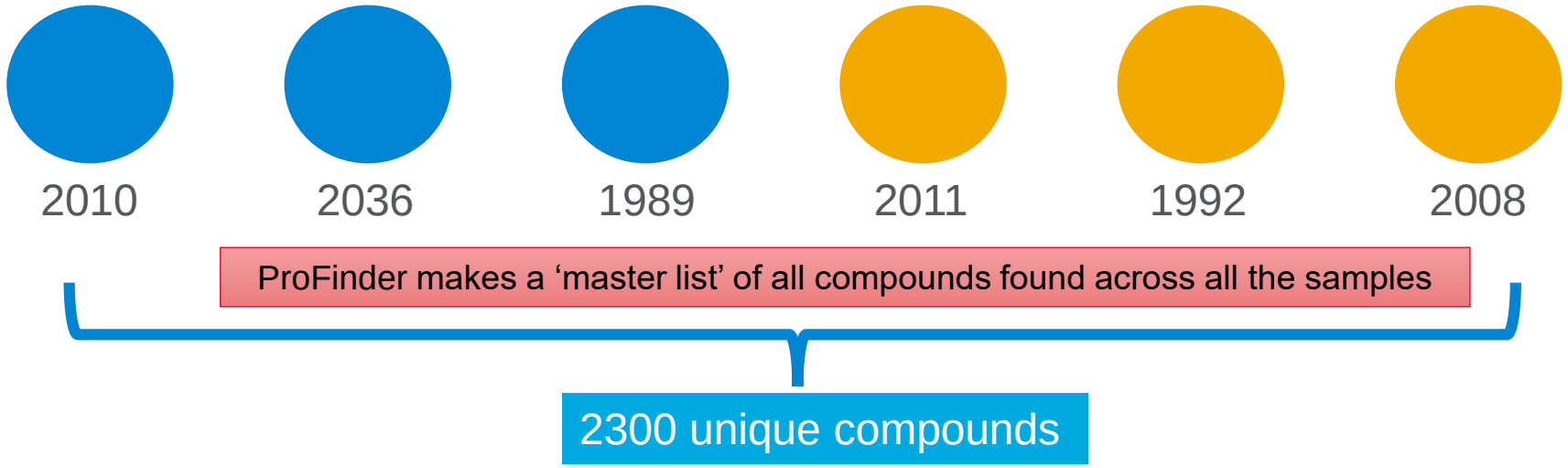
It could be possible some minor compounds would skip MFE on any replicate due to interference or any other reason.

Recursive allows to go deeper in the data and extract without noise filters from data on other replicates.

- Feature extraction using MFE (works on 3D data set)
  finds targeted AND untargeted OR unknowns
- Alingment & Data prepare
  features found across ALL samples, but might be missing in some
- Feature extraction Find by Ions (**works EIC based**)
  finds targeted metabolites only (all or significant only), **finds lower level missing features**

Recursive data extraction in Pro Finder

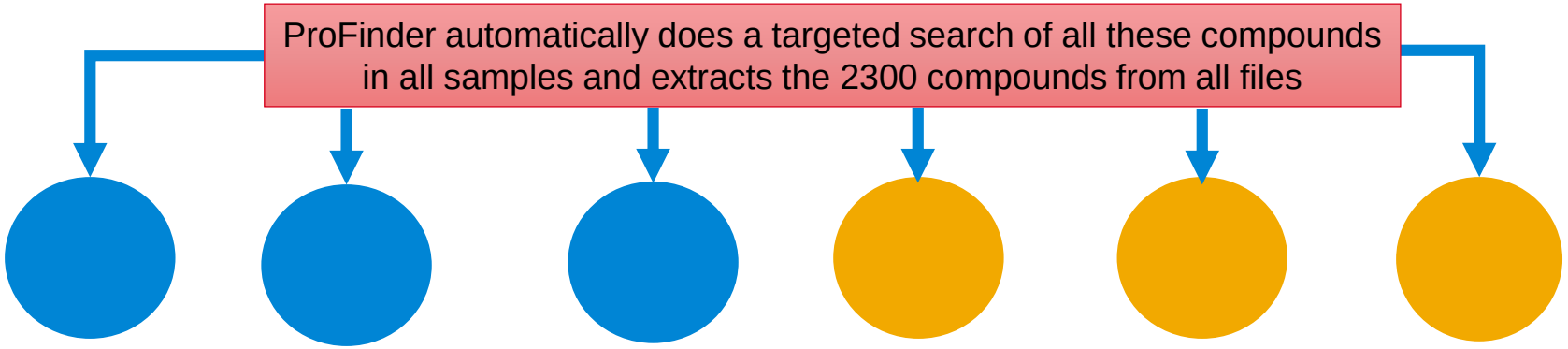
1. Unbiased feature extraction



Unbiased feature detection will always find different numbers of compounds

Peak-picking/feature detection algorithms requires you to set a threshold level

2. Targeted feature extraction



Targeted feature extraction has no threshold, it will extract right down into the noise

Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - ***¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil…. Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - ***Mass Profiler professional. Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - *Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos?* ***Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

A Comprehensive Metabolomics Workflow

Agilent LCMS, CEMS and GCMS

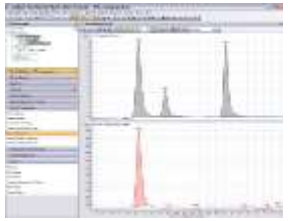
Separate & Detect	Feature Finding & Data Prepare	Data Prep. & Statistics	Identify	Pathway Analysis // Profiling

CE-LC-TOF/QTOF
CE-LC-QQQ

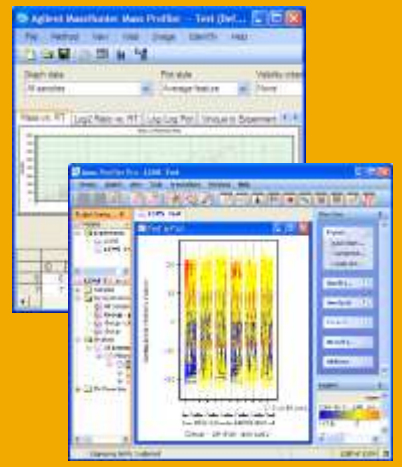
MassHunter Profinder

MassHunter Qual

MassHunter Quant



Mass Profiler (Professional)

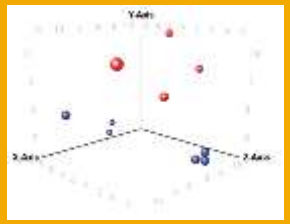


Statistics Visualization

ID Browser

Annotation & Identification

Pathway Analysis



Profiling

LCMS, CEMS and GCMS Data can be analyzed together in the same project

Mass Profiler professional.

Differential Analysis using Interpretations, clustering, PCA, PLRS, model of prediction

- Agilent Mass Profiler Professional (MPP) software is a powerful chemometrics platform
- Designed to exploit the high information content of mass spectra (MS) data
- Can be used in any MS-based differential analysis to determine relationships among two or more sample groups and variables.
- Provides advanced statistical analysis and visualization tools for GC/MS, LC/MS, CE/MS, and ICP-MS data analysis.
- Also integrates smoothly with Agilent MassHunter Workstation, Profinder, Spectrum Mill, and ChemStation software.
- Is the only platform that provides integrated identification/ annotation of compounds and integrated pathway analysis for metabolomic and proteomic studies.

Mass Profiler Professional
MPP Terminology

Project – Workspace, container of Experiments

Experiment – Collection of samples acquired under same instrument method.

Parameter – Variable in the experiment (p.e. Time, Temp, Infected)

Condition – One or more samples representing a common biologic status (p.e. Time 14h)

Interpretation – Samples grouping based on Conditions. (p.e. Time vs Temp)

Entity – Molecular Entity from which we know Rt, Mass and Abundance. It can be "Identified" or "Not Identified".

***Technology* – Registry or container of all data acquired under a simple Technology: Metabolomics, Transcriptomics, Proteomics….**

Mass Profiler Professional
MPP Terminology

Sample	Grape	Country of Origin
1	Cabernet Sauvignon	USA
2	Cabernet Sauvignon	USA
3	Cabernet Sauvignon	France
4	Cabernet Sauvignon	France
5	Merlot	USA
6	Merlot	USA
7	Merlot	France
8	Merlot	France
9	Pinot Noir	USA
10	Pinot Noir	USA
11	Pinot Noir	France
12	Pinot Noir	France

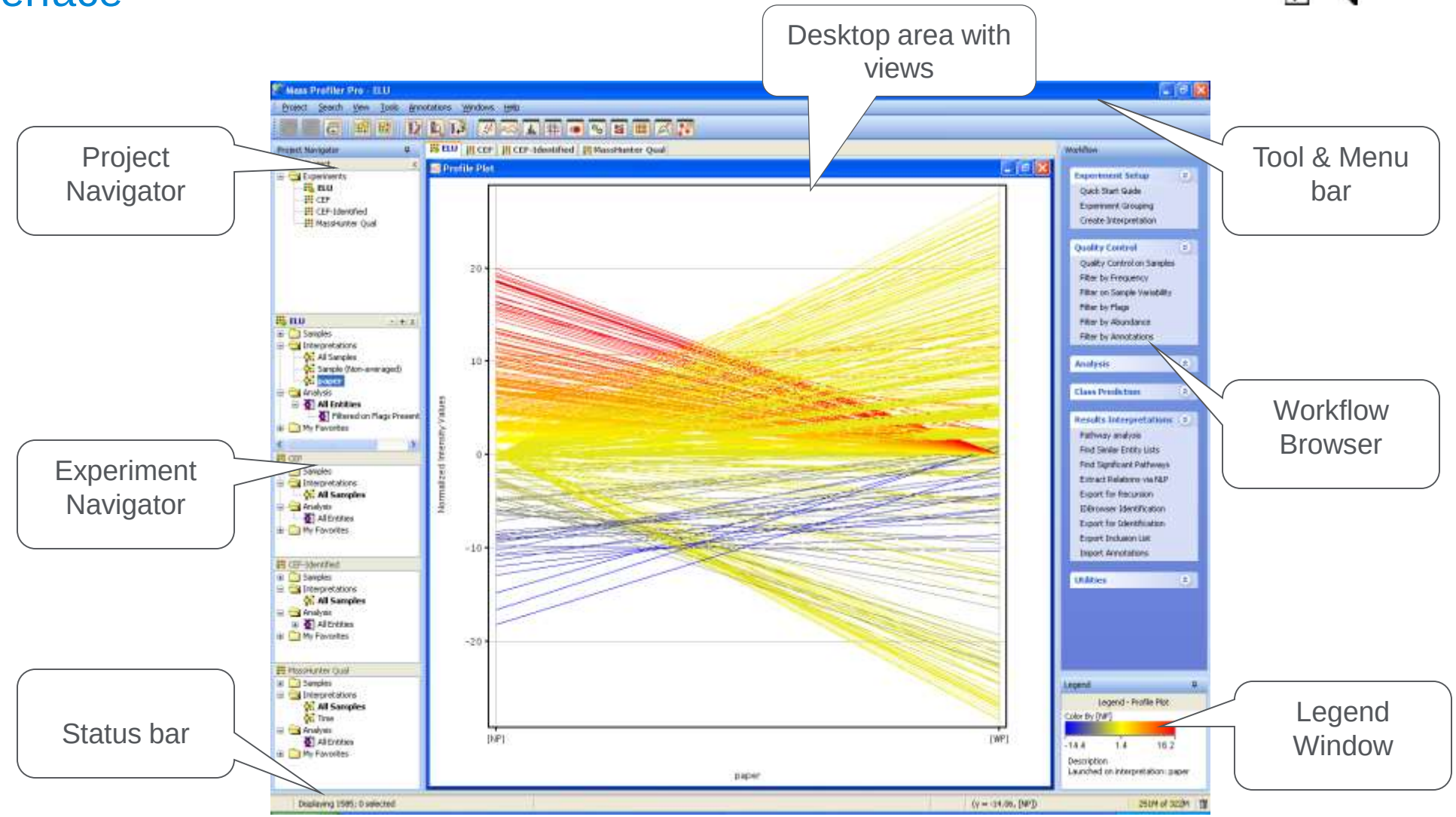
Interpretation 1: Define samples by parameter Grape
Condition1: Cabernet Sauvignon (samples 1-4)
Condition2: Merlot (samples 5-8)
Condition3: Pinot Noir (samples 9-12)

Interpretation 2: Define samples by parameter Country
Condition1: USA (samples 1,2,5,6,9,10)
Condition2: France (samples 3,4,7,8,11,12)

Interpretation 3: Define samples by parameter Grape and Country
Condition1: Cabernet Sauvignon-USA (samples 1,2)
Condition2: Cabernet Sauvignon-France (samples 3,4)
Condition3: Merlot-USA (samples 5,6)
Condition4: Merlot-France (samples 7,8)
Condition5: Pinot Noir-USA (samples 9,10)
Condition6: Pinot Noir-France (samples 11,12)

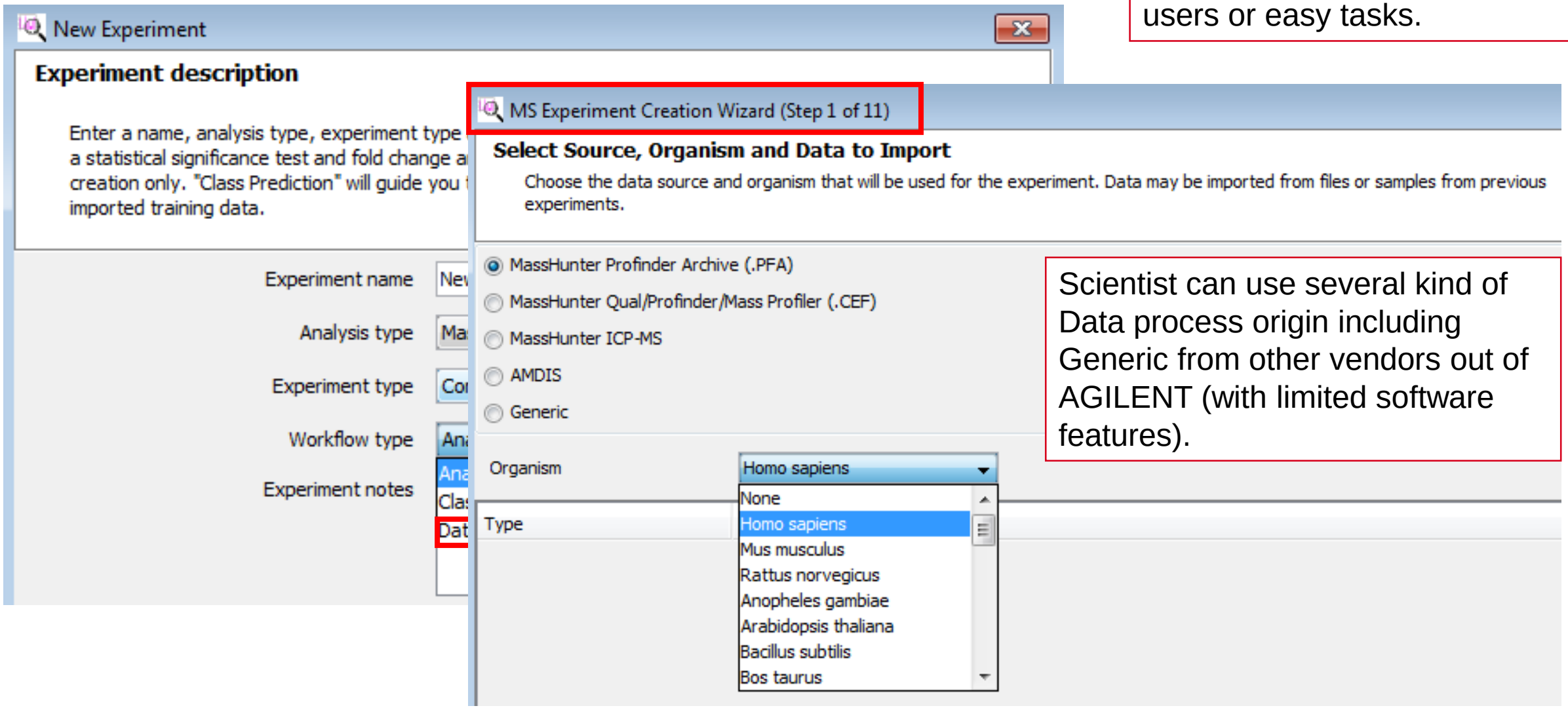
Mass Profiler Professional

MPP Interface



Mass Profiler Professional
Experiment Creation

There are different choices when creating an experiment including Wizard for novel users or easy tasks.

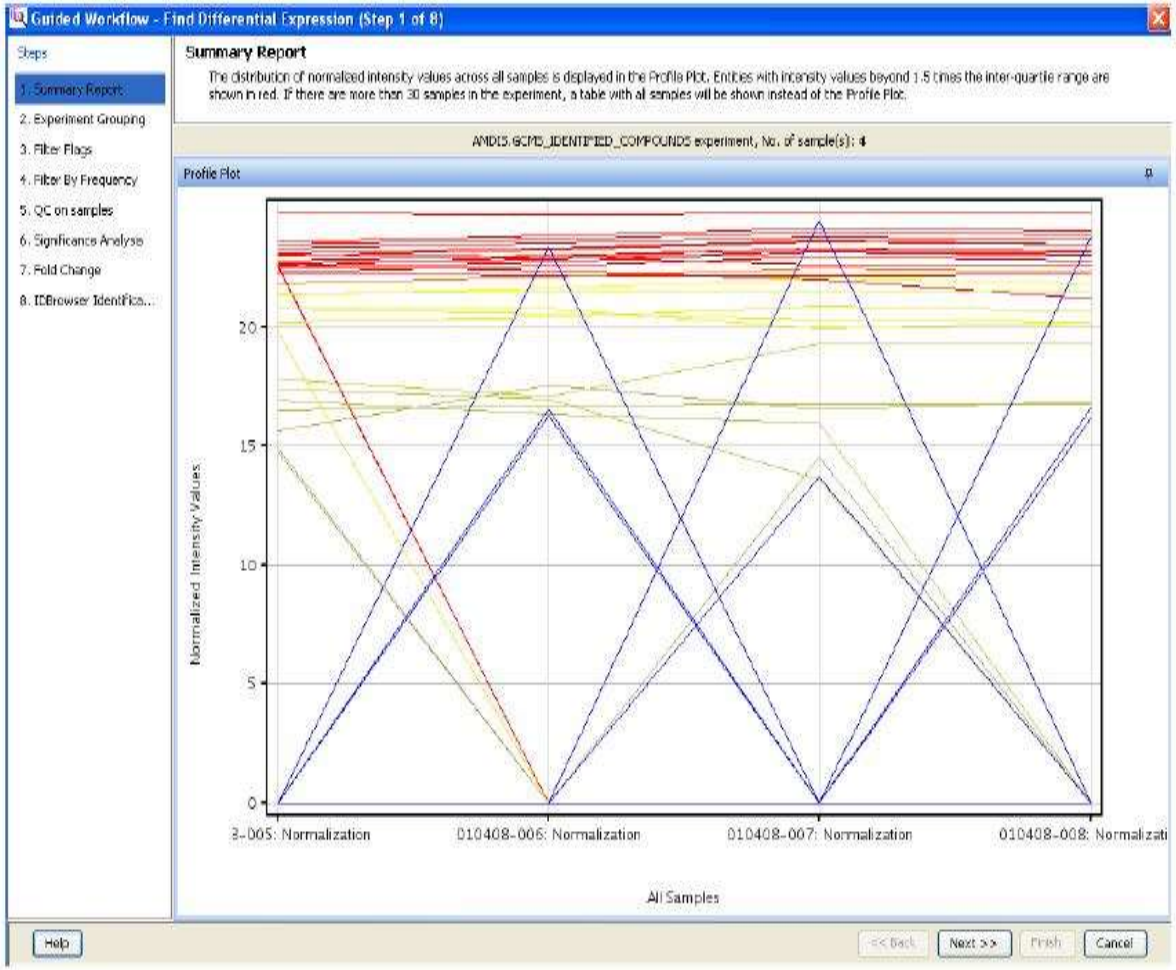


Scientist can use several kind of Data process origin including Generic from other vendors out of AGILENT (with limited software features).

Mass Profiler Professional

Guided Workflow Analysis – Step 1 : Summary Report

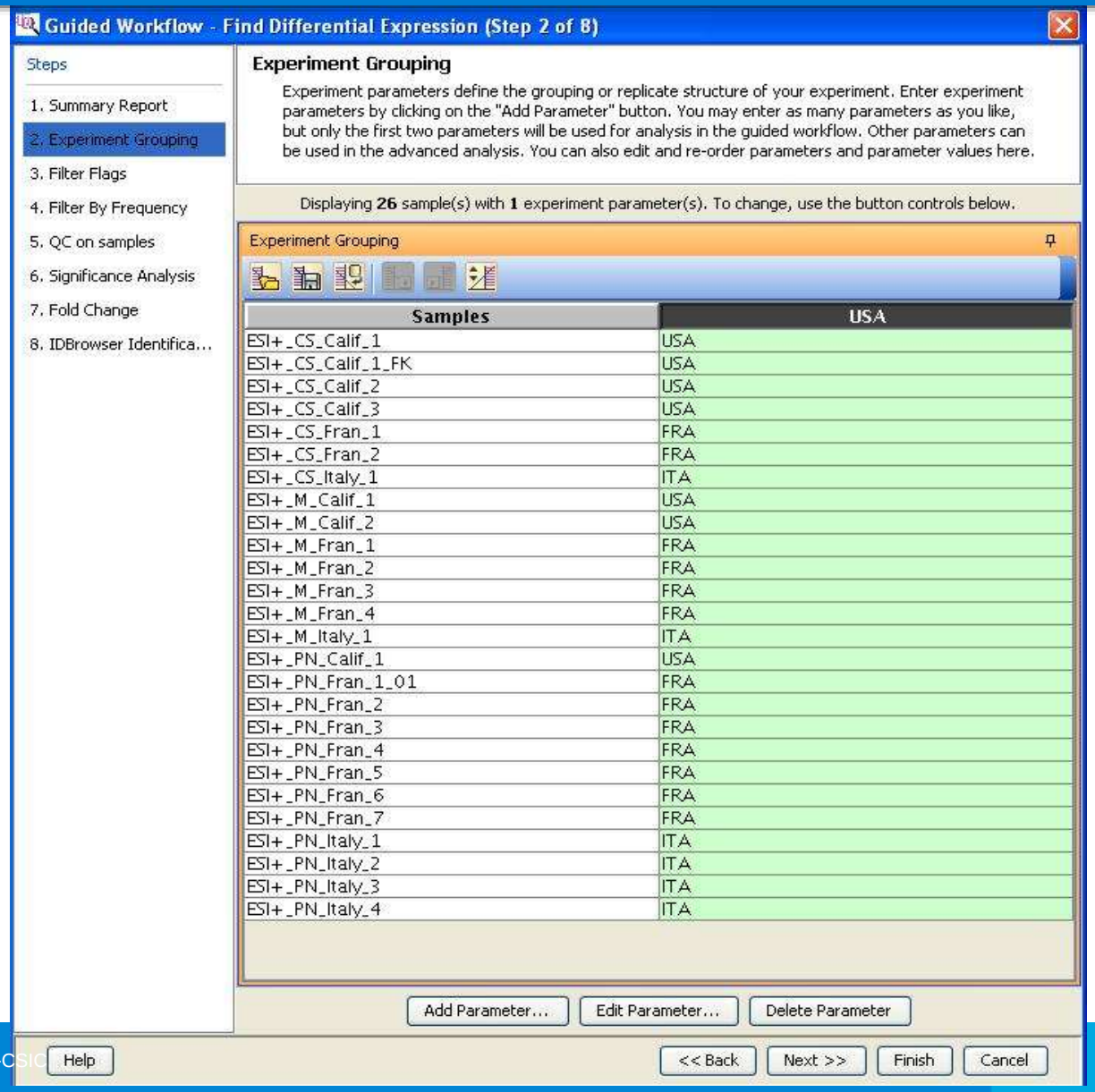
1. Summary Report
2. Experiment Grouping
3. Filter on Flags
4. Filter by Frequency
5. QC on Samples
6. Significance Analysis
7. Fold Change
8. ID Browser identification



If < 30 Samples then Profile Plot (shown above) is Generated Baseline

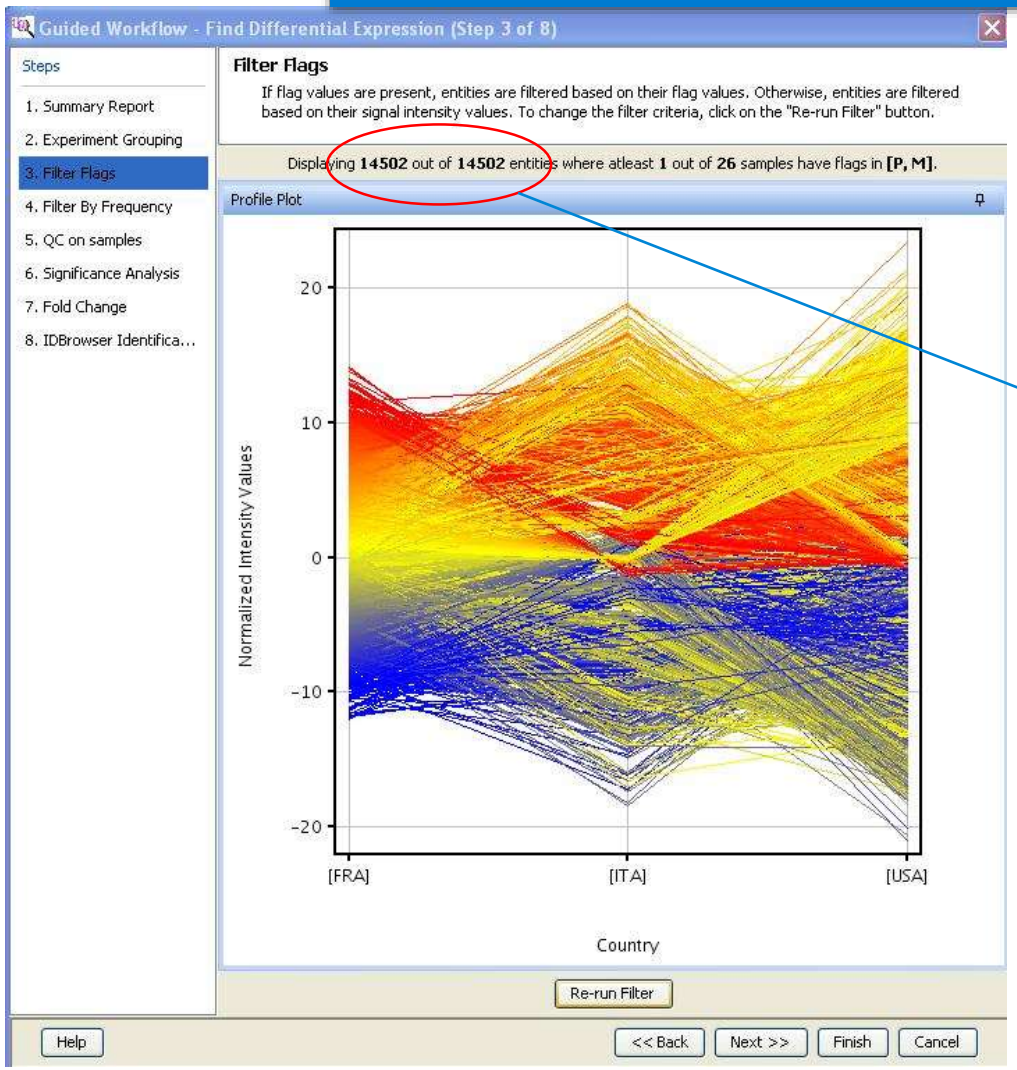
Mass Profiler Professional

Guided Workflow Analysis – Step 2: Grouping



Mass Profiler Professional

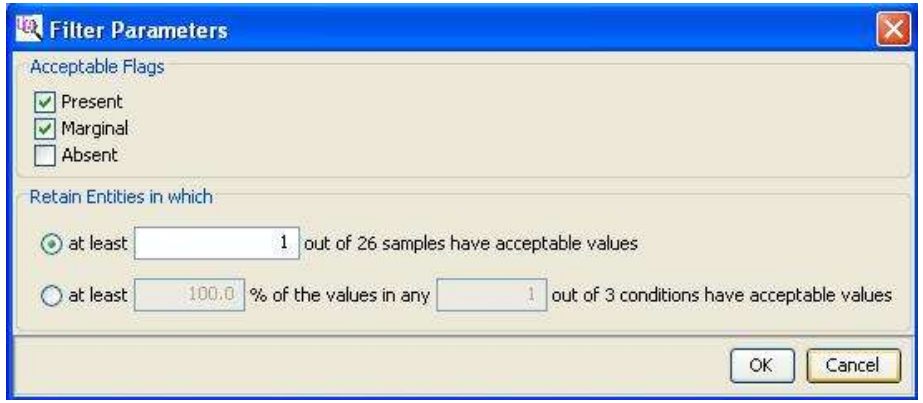
Guided Workflow Analysis – Step 3: Filter on Flags



- Can be used to filter out entities which are rarely detected, therefore not very reliable

- Similar to filter by frequency

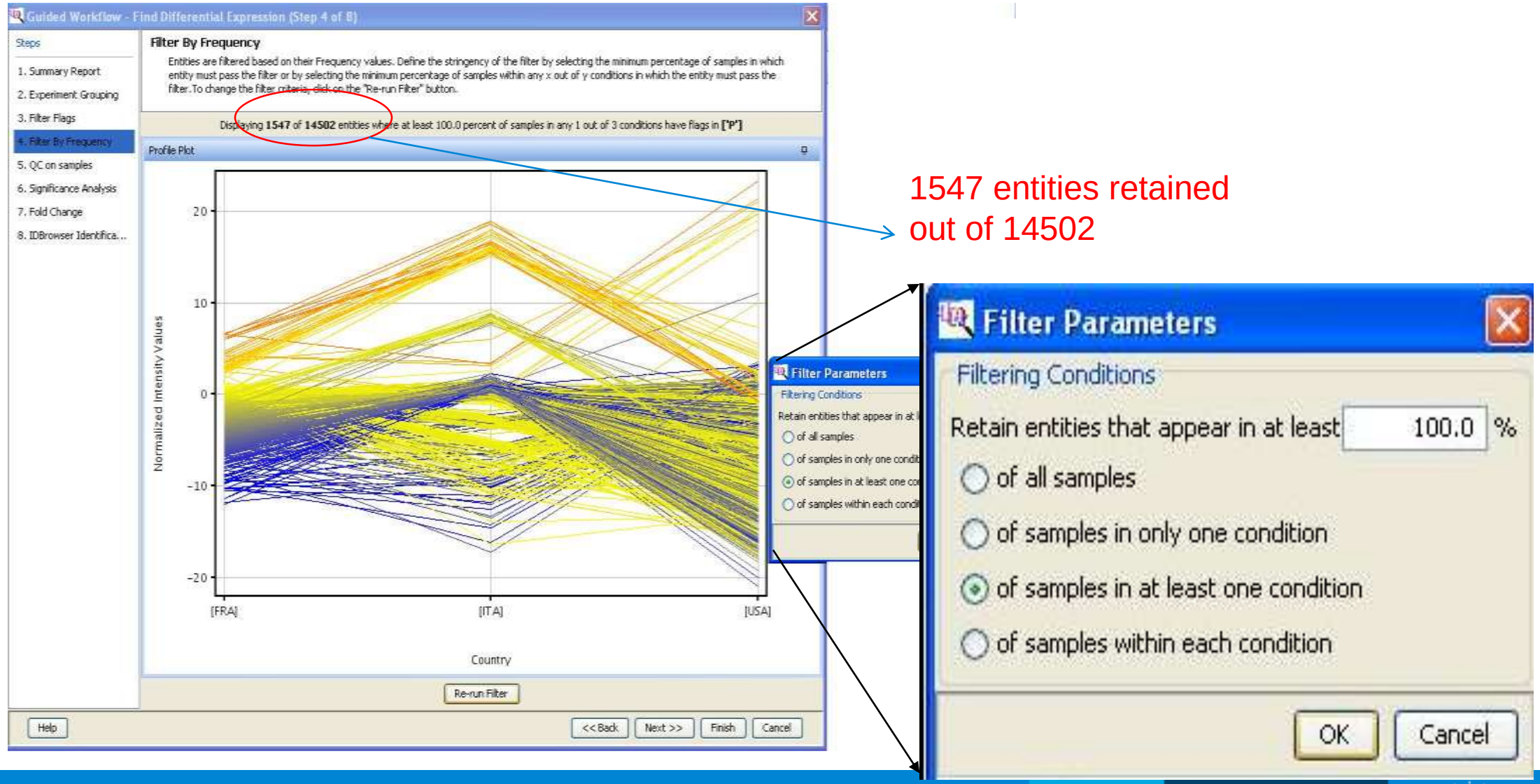
14502 entities retained out of 14502



Filter on “Present” and “Marginal” (saturated entity)

Mass Profiler Professional

Guided Workflow Analysis – Step 4: Filter by Frequency

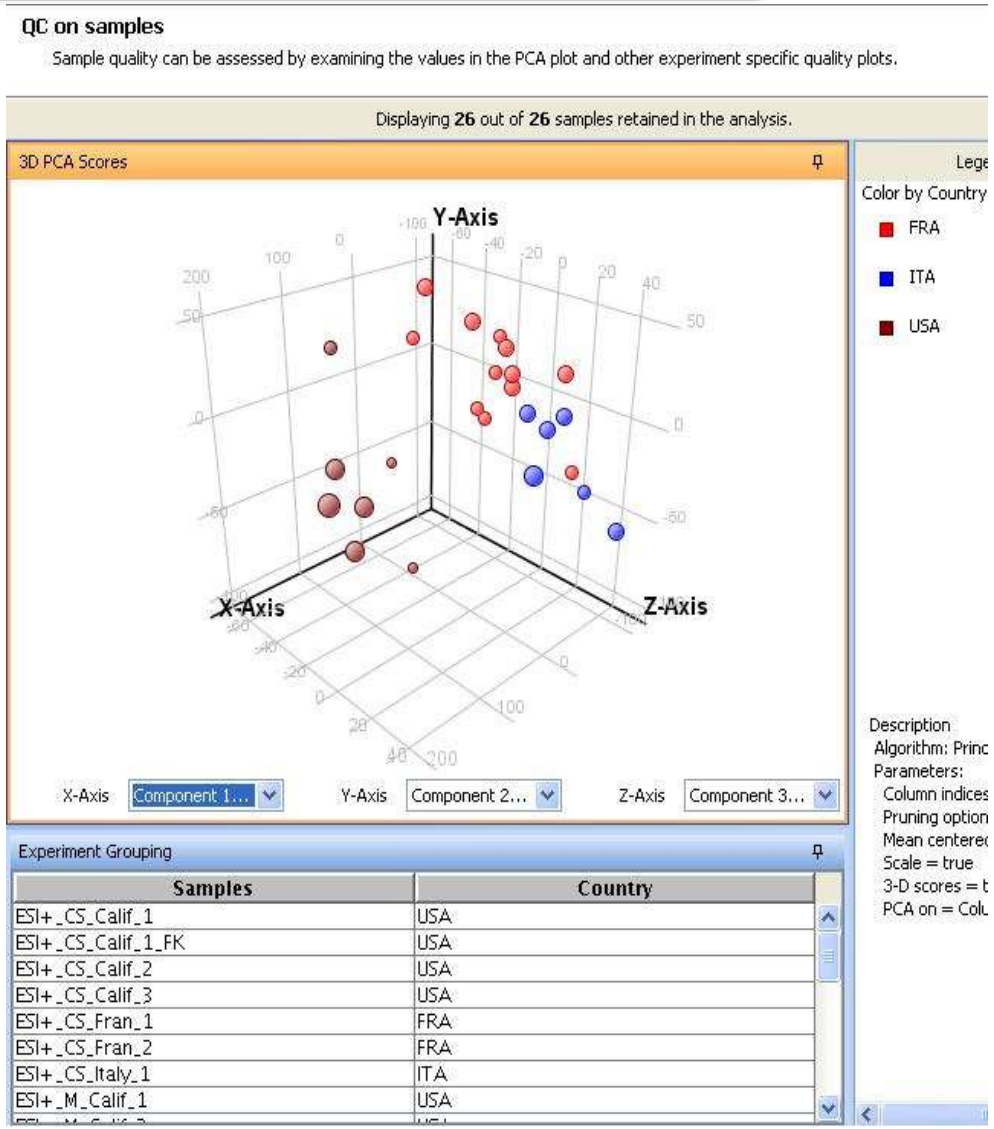


1547 entities retained
out of 14502

Mass Profiler Professional

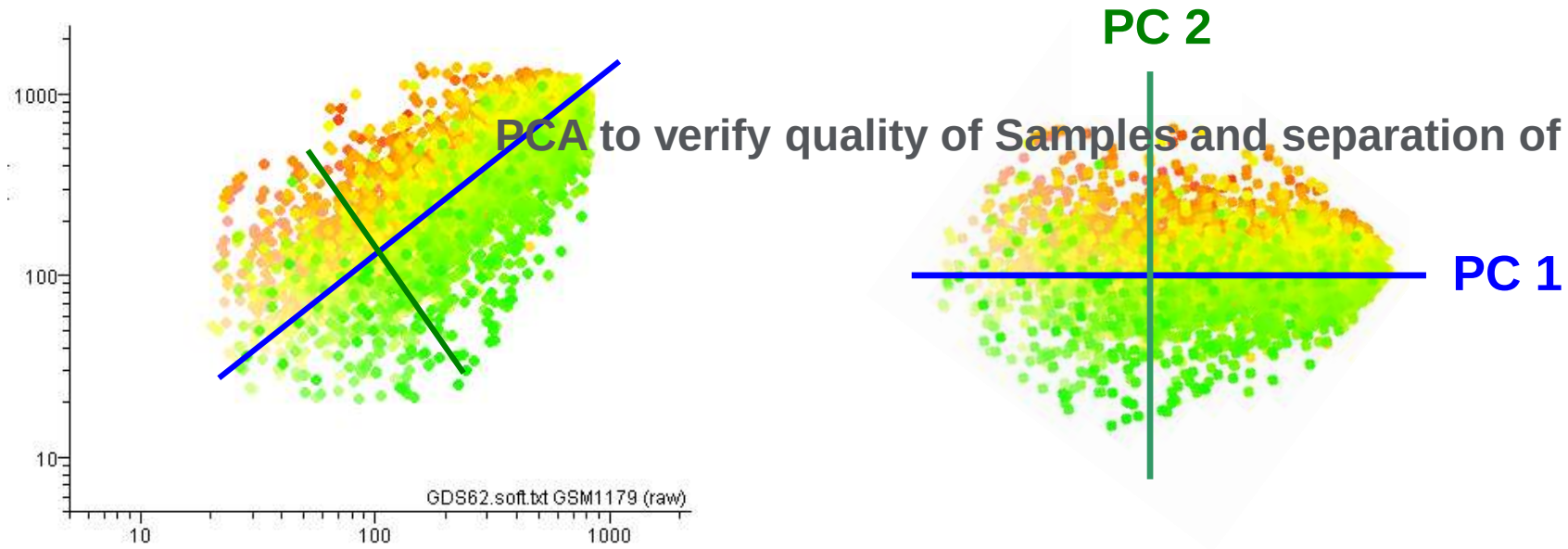
Guided Workflow Analysis – Step 5: QC on Samples

- By default, each sample is plotted according to its values for the first three Principal Components
- Principal Components are vectors that capture the most variance in the data.
- Assumption: samples within an experimental condition should be more similar to each other than to those from different conditions.
- Expect to see samples from the same experimental condition to group closer to each other than to samples of a different condition



Mass Profiler Professional

PCA is a variable reduction Method



PCA to verify quality of Samples and separation of groups

An eigenvalue-eigenvector decomposition is performed on the covariance matrix of the entity expression values around zero

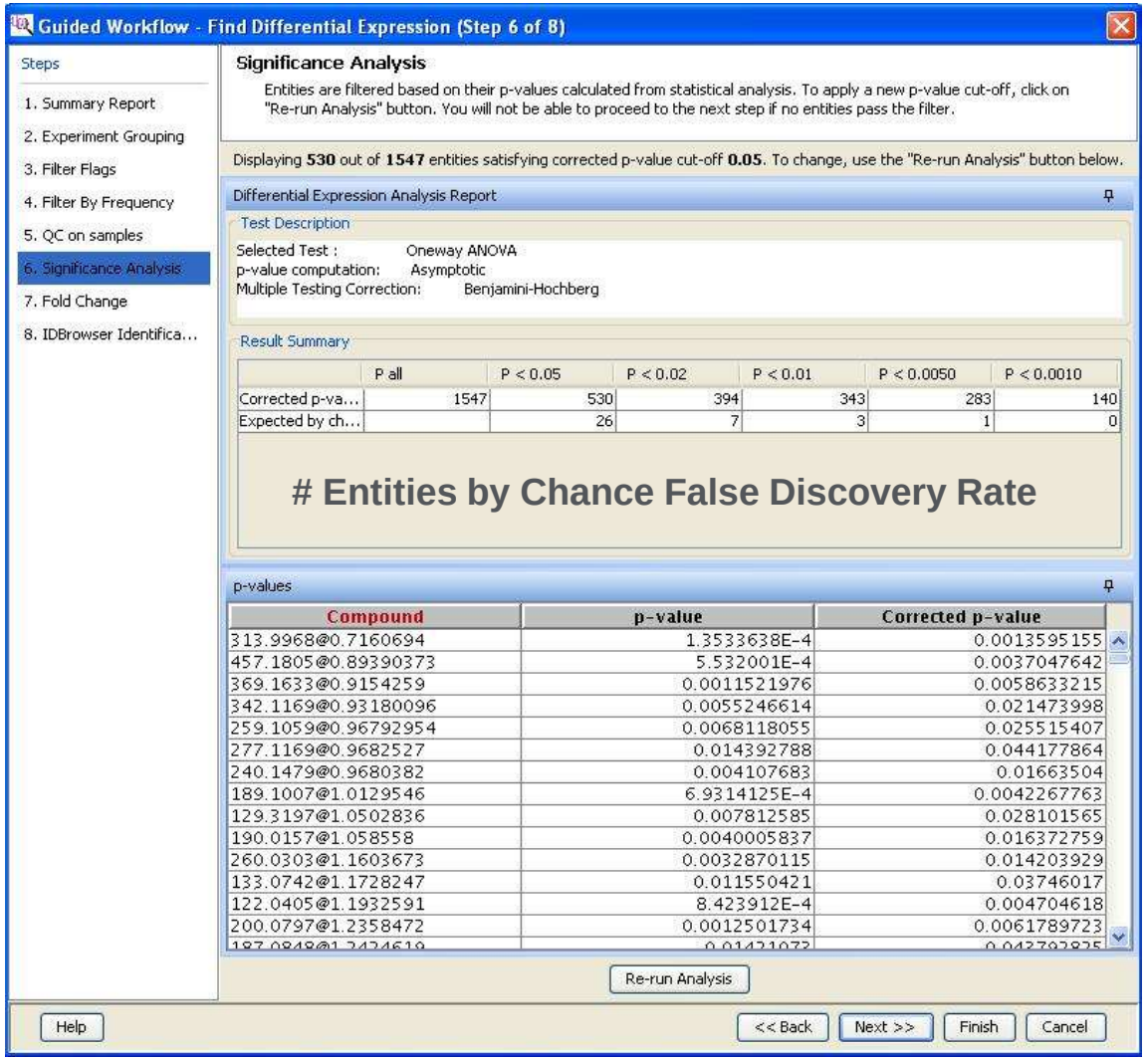
The eigenvector corresponding to the largest eigenvalue is called the first principal component

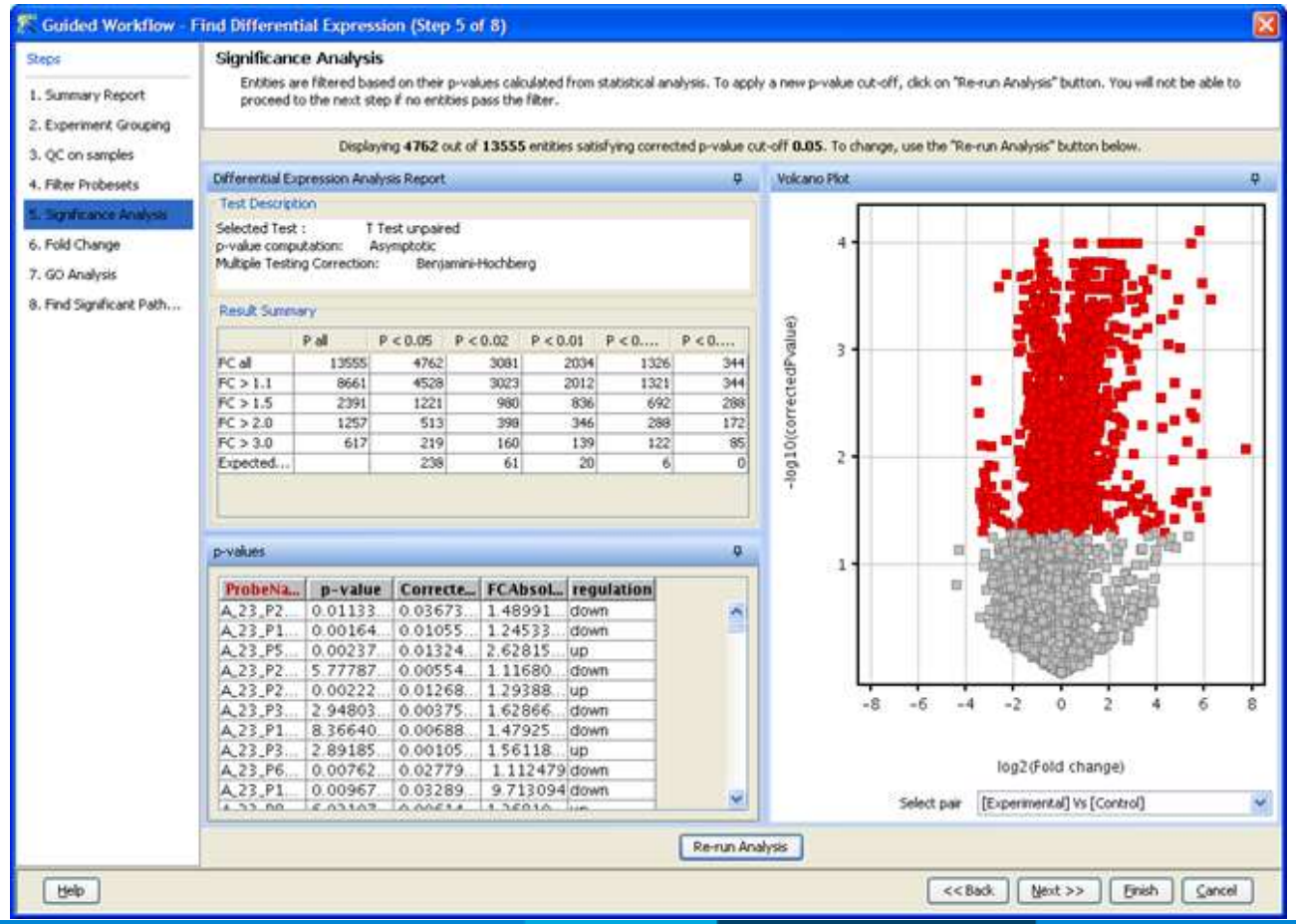
Successive principal components are eigenvectors corresponding to each smaller eigenvalue

Mass Profiler Professional

Guided Workflow Analysis – Step 6: Significance Analysis

Depending upon the experimental grouping, **Mass Profiler Professional** performs either T-test or Analysis of Variance (ANOVA) based on the samples.





Mass Profiler Professional

Guided Workflow Analysis – Statistical Tests

<u>One-way Tests:</u> Compare conditions defined by a single parameter (i.e Grape)

T-test

Merlot ⟺ Pinot Noir

ANOVA

Merlot ⟺ Pinot Noir ⟺ Caber

<u>*N*-way Tests:</u> Compare conditions defined by 2 or more parameters

2-way ANOVA

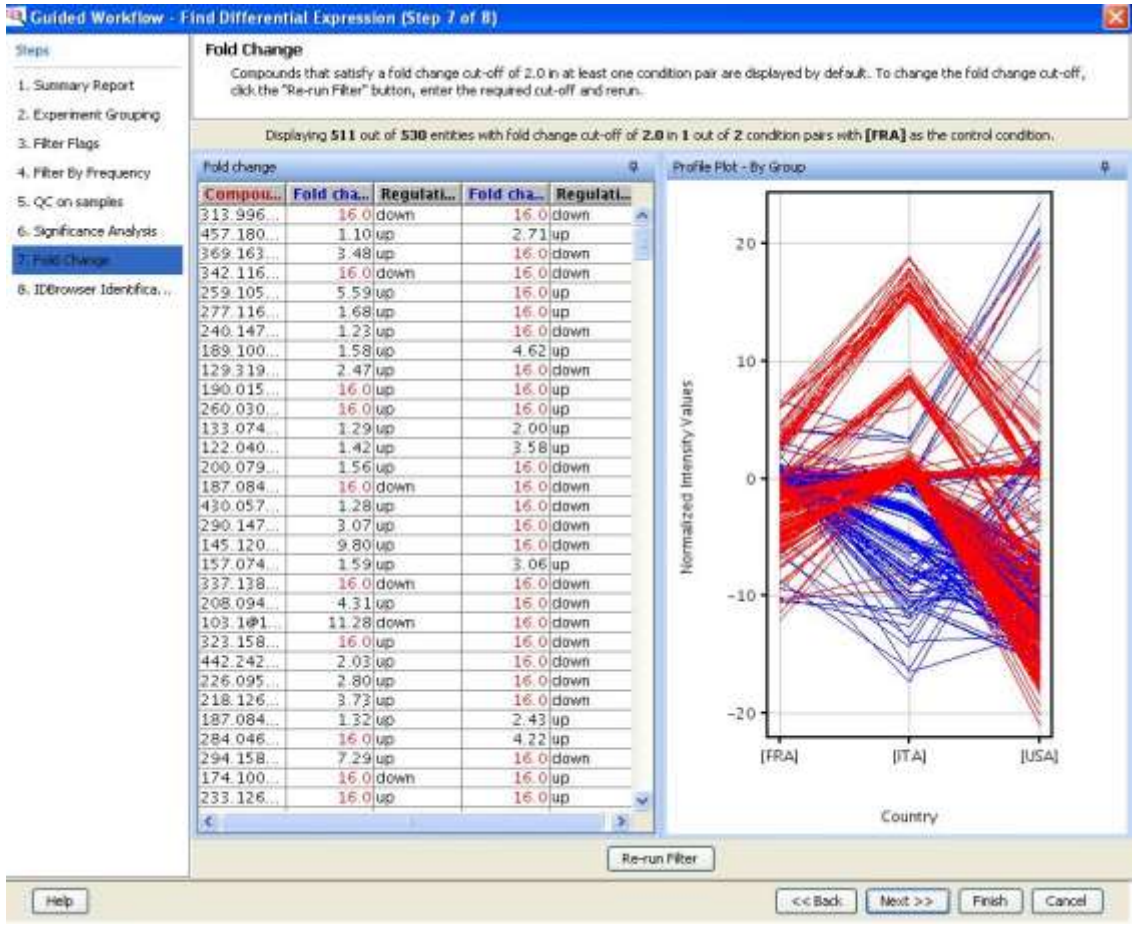
<u>Grape</u>		<u>Country</u>
Merlot	X	USA
Pinot Noir		France

3-way ANOVA

<u>Grape</u>		<u>Country</u>		<u>Vintage</u>
Merlot	X	USA	X	2004
Pinot Noir		France		2006

Mass Profiler Professional

Guided Workflow Analysis – Step 7: Fold Change / Volcano



Filter Sets Fold Change > 2.0.. To more

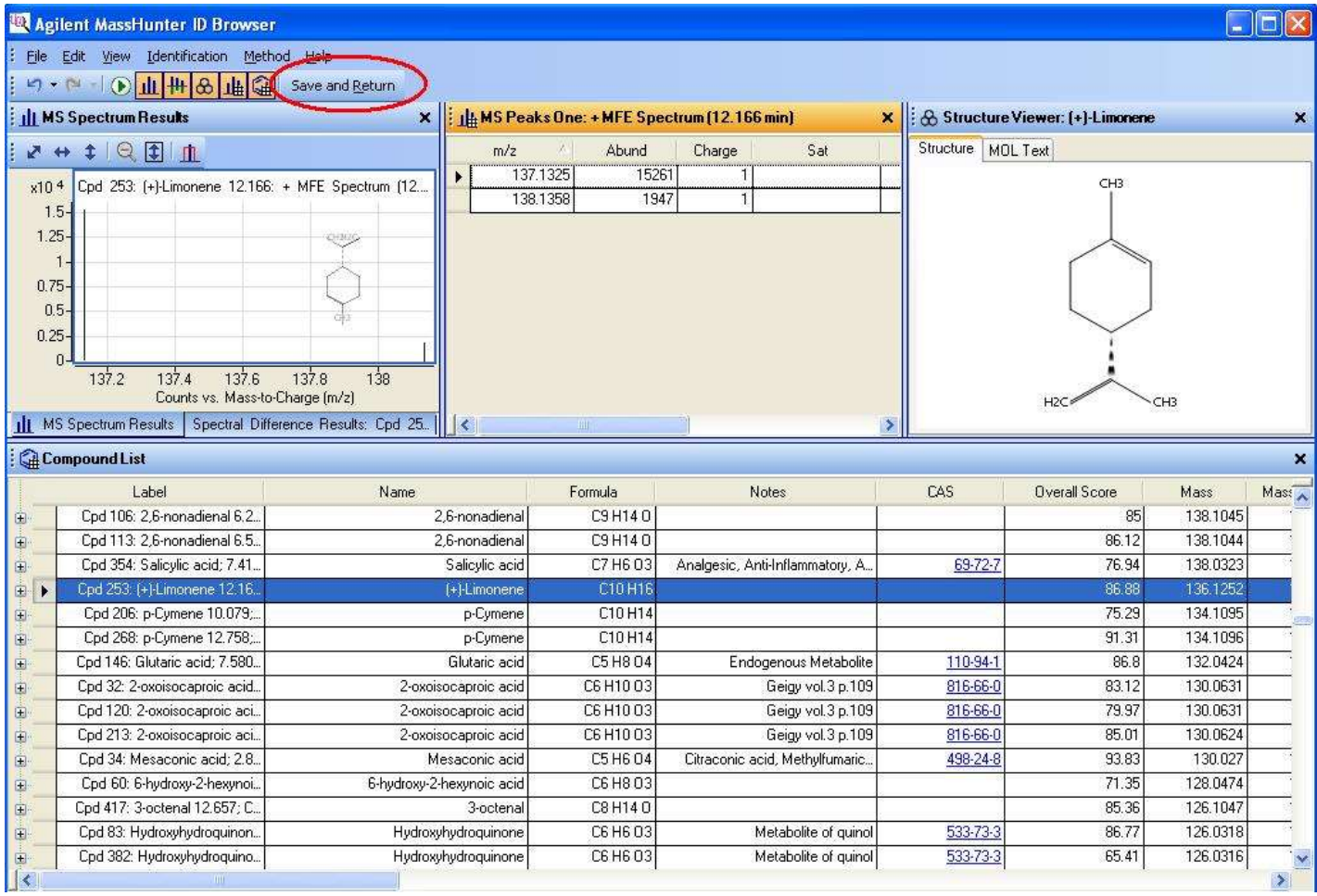
Mass Profiler Professional

Guided Workflow Analysis – Step 8: ID Browser

Once we have a list of differential features, we need to Identify them by any library at the first attempt.

GCMS data is more easy due to available universal EI libraries.

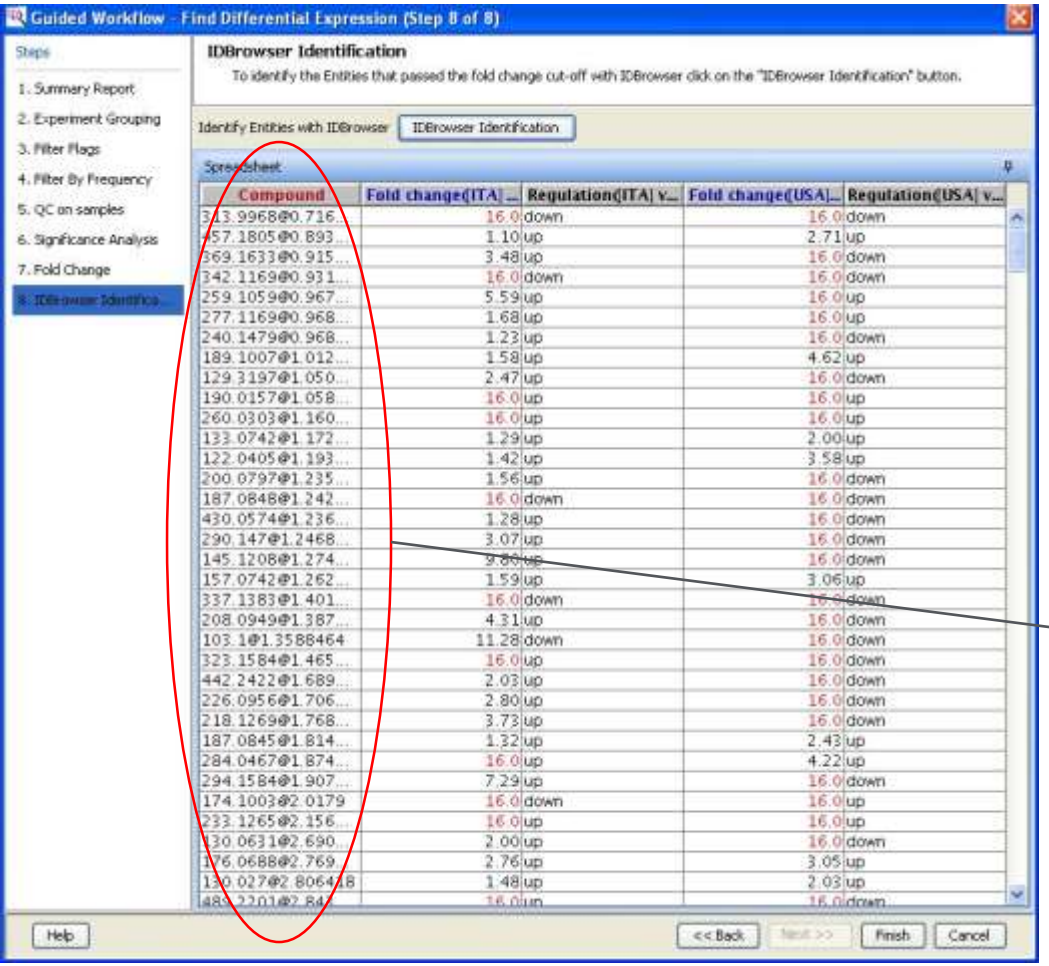
For LCHRMS Vendor or user should create their own libraries for a confident and reliable ID.



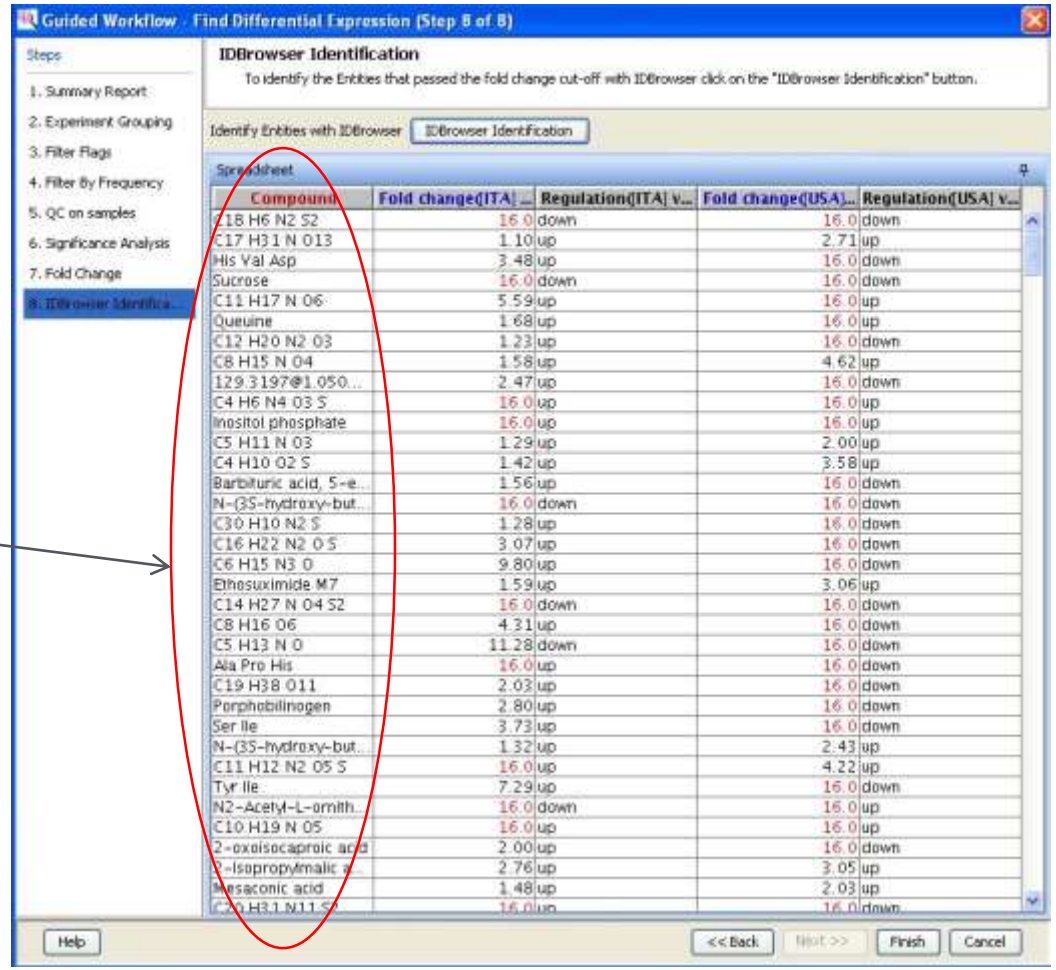
Entities will be annotated upon return to MPP Program

Mass Profiler Professional

Guided Workflow Analysis – Step 8: ID Browser - Return



Compound	Fold change(ITA) ...	Regulation(ITA) v...	Fold change(USA)...	Regulation(USA) v...
313.9968@0.716...	16.0	down	16.0	down
457.1805@0.893...	1.10	up	2.71	up
369.1633@0.915...	3.48	up	16.0	down
342.1169@0.931...	16.0	down	16.0	down
259.1059@0.967...	5.59	up	16.0	up
277.1169@0.968...	1.68	up	16.0	up
240.1479@0.968...	1.23	up	16.0	down
189.1007@1.012...	1.58	up	4.62	up
129.3197@1.050...	2.47	up	16.0	down
190.0157@1.058...	16.0	up	16.0	up
260.0303@1.160...	16.0	up	16.0	up
133.0742@1.172...	1.29	up	2.00	up
122.0405@1.193...	1.42	up	3.58	up
200.0797@1.235...	1.56	up	16.0	down
187.0848@1.242...	16.0	down	16.0	down
430.0574@1.236...	1.28	up	16.0	down
290.147@1.2468...	3.07	up	16.0	down
145.1208@1.274...	9.80	up	16.0	down
157.0742@1.262...	1.59	up	3.06	up
337.1383@1.401...	16.0	down	16.0	down
208.0949@1.387...	4.31	up	16.0	down
103.1@1.3588464	11.28	down	16.0	down
323.1584@1.465...	16.0	up	16.0	down
442.2422@1.689...	2.03	up	16.0	down
226.0956@1.706...	2.80	up	16.0	down
218.1269@1.768...	3.73	up	16.0	down
187.0845@1.814...	1.32	up	2.43	up
284.0467@1.874...	16.0	up	4.22	up
294.1584@1.907...	7.29	up	16.0	down
174.1003@2.0179	16.0	down	16.0	up
233.1265@2.156...	16.0	up	16.0	up
130.0631@2.690...	2.00	up	16.0	down
176.0688@2.769...	2.76	up	3.05	up
130.027@2.8064[illegible]	1.48	up	2.03	up
[illegible]	16.0	up	16.0	down



Compound	Fold change(ITA) ...	Regulation(ITA) v...	Fold change(USA)...	Regulation(USA) v...
C18 H6 N2 S2	16.0	down	16.0	down
C17 H31 N 013	1.10	up	2.71	up
His Val Asp	3.48	up	16.0	down
Sucrose	16.0	down	16.0	down
C11 H17 N 06	5.59	up	16.0	up
Queuine	1.68	up	16.0	up
C12 H20 N2 03	1.23	up	16.0	down
C8 H15 N 04	1.58	up	4.62	up
129.3197@1.050...	2.47	up	16.0	down
C4 H6 N4 03 S	16.0	up	16.0	up
Inositol phosphate	16.0	up	16.0	up
C5 H11 N 03	1.29	up	2.00	up
C4 H10 02 S	1.42	up	3.58	up
Barbituric acid, 5-e...	1.56	up	16.0	down
N-(3S-hydroxy-but...	16.0	down	16.0	down
C30 H10 N2 S	1.28	up	16.0	down
C16 H22 N2 O S	3.07	up	16.0	down
C6 H15 N3 0	9.80	up	16.0	down
Ethosuximide M7	1.59	up	3.06	up
C14 H27 N 04 S2	16.0	down	16.0	down
C8 H16 06	4.31	up	16.0	down
C5 H13 N 0	11.28	down	16.0	down
Ala Pro His	16.0	up	16.0	down
C19 H38 011	2.03	up	16.0	down
Porphobilinogen	2.80	up	16.0	down
Ser Ile	3.73	up	16.0	down
N-(3S-hydroxy-but...	1.32	up	2.43	up
C11 H12 N2 05 S	16.0	up	4.22	up
Tyr Ile	7.29	up	16.0	down
N2-Acetyl-L-ornith...	16.0	down	16.0	up
C10 H19 N 05	16.0	up	16.0	up
2-oxoisocaproic acid	2.00	up	16.0	down
2-Isopropylmalic a...	2.76	up	3.05	up
Mesaconic acid	1.48	up	2.03	up
C20 H31 N11 S2	16.0	up	16.0	down

Entities annotated return to MPP Program

Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - ***¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil…. Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - ***Mass Profiler professional. Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - ***Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos? Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - *Análisis de rutas Metabólicas a través de "Pathways Analysis".* ***Biología integrada e interpretación biológica de mis datos.Pathways Analysis****.*
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

So, what are my compounds of interest…. Where can Identify them?

Curated or Free MS/MS Libraries. Agilent METLIN PCDL.

PCDLs by Compounds & Spectra

LC/MS PCDL	Market	PCDL	Compounds with AM MS/MS Spectra	Total number of Spectra	Compounds with RTs
Forensic Toxicology	Forensic Toxicology	>9,200	>3,900	>13,500	0
Pesticides	Food Safety / Environmental	>1,700	>800	>2,700	0
Veterinary Drugs	Food Safety	>2,100	>1,500	>5,200	>120
Mycotoxins	Food Safety	>450	>300	>1,300	0
Water Contaminants	Environmental	>1,400	>1,000	>3,900	>260
METLIN*	Metabolomics / Lipidomics	>79,600**	>9,400	>32,000	>680
NIST 2014 MS/MS	General	>9,300	>9,300	>234,000	0

*METLIN numbers exclude tri- and quatra- peptides in the online METLIN

**** Plus 168k theoretical**

Why is Curation Important?
More isn’t Always Better

An open-source database may contain over a million spectra

But how many of those spectra are:

- Relevant to your application?
- Duplicates, triplicates…?
- Collected under questionable conditions?
- Contain inaccurate metadata?

Not all Databases and Libraries are Created Equal

Non curated or poorly curated databases and libraries costs you time, money and uncertatinty

Questions Laboratories don’t want to ask:

- Which of these entries is my molecule?
- Are my identification match scores meaningful?
- Can I trust this source?
- There must be valuable information here.., but where?
- Is this even relevant to my application?

How to Identify the Right Database and Library

Agilent made a huge effort to create Databases and Libraries to ensure results are reliable

Advantages include:

- Developed by experts
- Designed to high standards
- Turn-Key and fully integrated

The PCDL Curation Process

There are four primary steps in the development of an Agilent database or library:

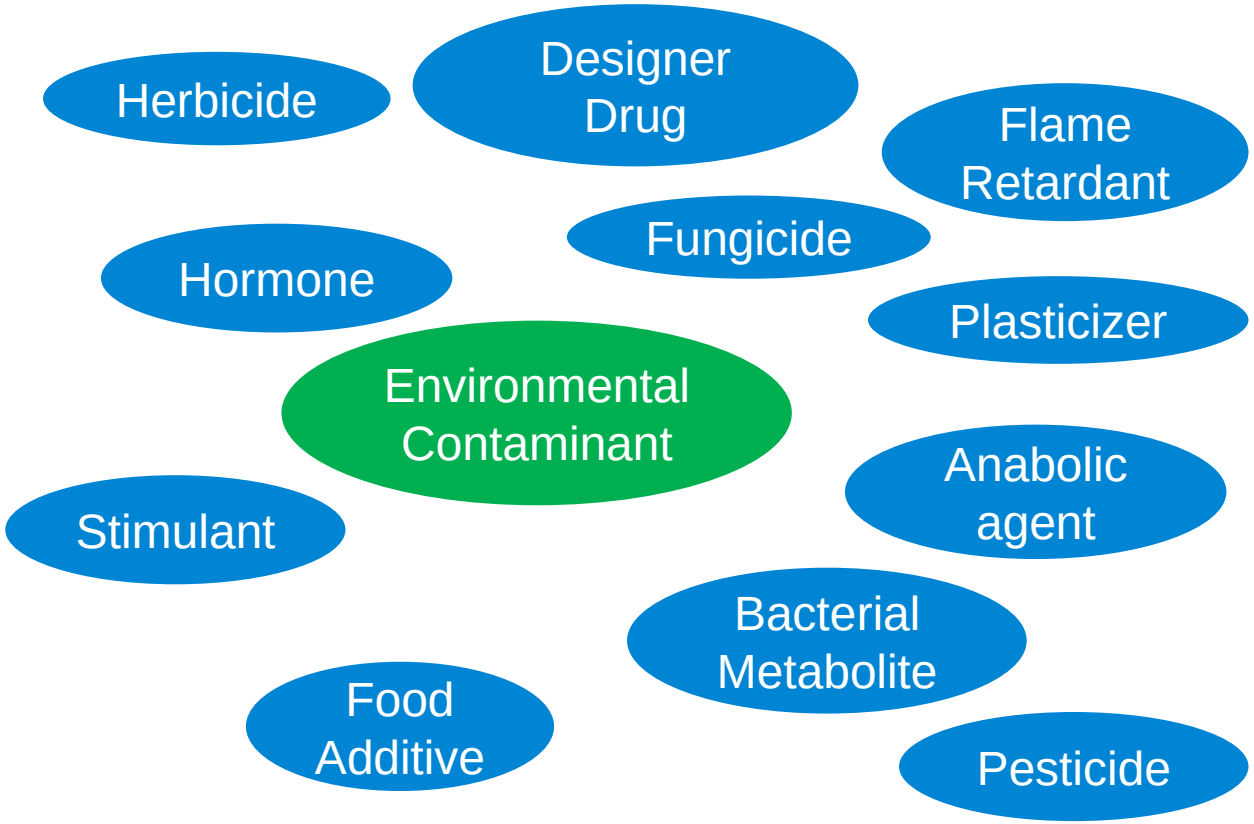
1. Identify the target compound list in collaboration with leading experts
2. Create the compound database with verified information for each target entry
3. Collect mass spectral data using high purity reference standards
4. Create the library with spectra curated according to Agilent's rigorous quality control curation process

Collaborations with Leading Experts

What's Relevant – the Importance of Targeted Compound Lists

When searching for trace compounds…

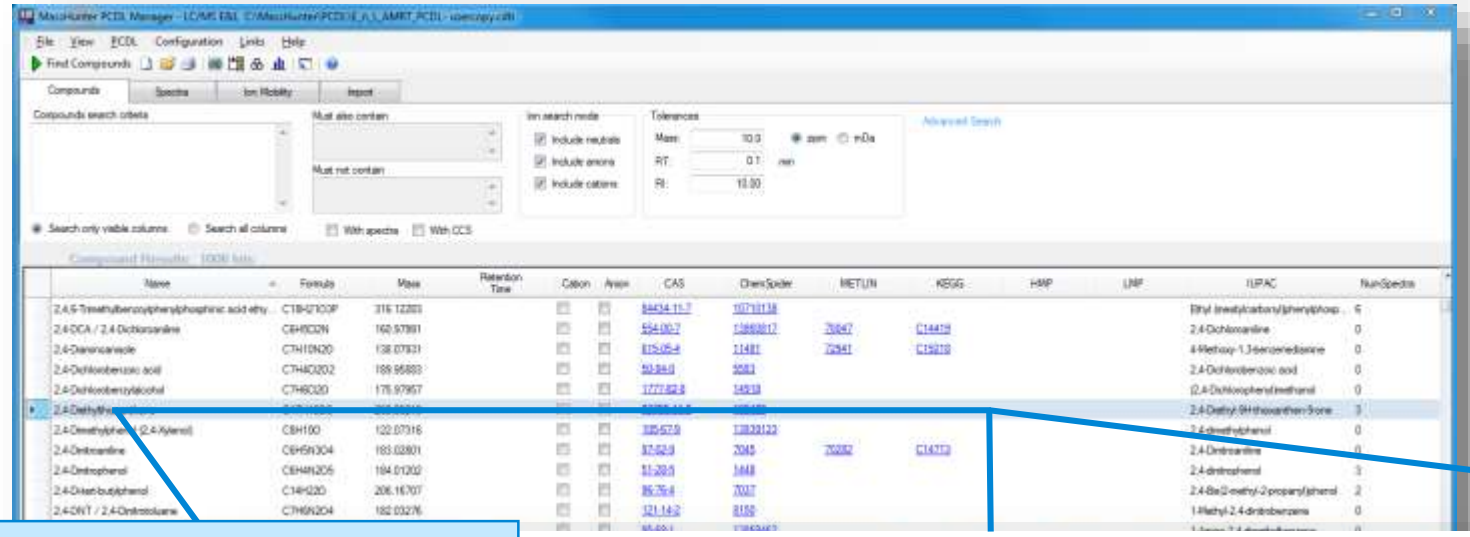
…do you really look everywhere?



Curation – Building the Compound Database

Chemical List

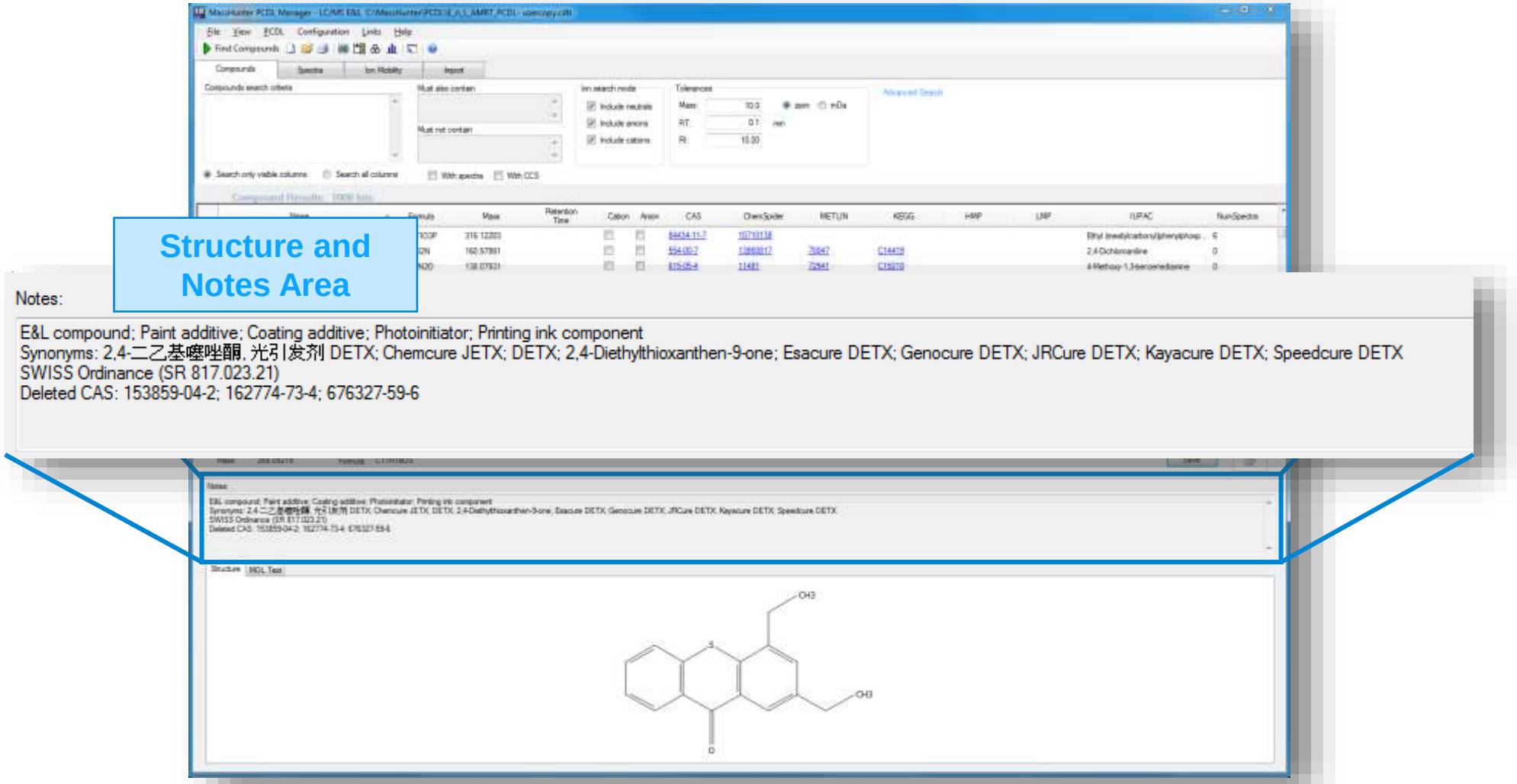
Common Name & Compound Information



1006 hits

	Formula	Mass	Retention Time	Cation	Anion	CAS	ChemSpider	METLIN
phosphinic acid ethy...	C18H21O3P	316.12283				84434-11-7	10710138	
2,4-DCA / 2,4-Dichloroaniline	C6H5Cl2N	160.97991				554-00-7	13860817	70047
2,4-Diaminoanisole	C7H10N2O	138.07931				615-05-4	11481	72941
2,4-Dichlorobenzoic acid	C7H4Cl2O2	189.95883				50-84-0	5583	
2,4-Dichlorobenzylalcohol	C7H6Cl2O	175.97957				1777-82-8	14918	
2,4-Diethylthioxanthone	C17H16OS	268.09219				82799-44-8	109489	
2,4-Dimethylphenol (2,4-Xylenol)	C8H10O	122.07316				105-67-9	13839123	
2,4-Dinitroaniline	C6H5N3O4	183.02801				97-02-9	7045	70282
2,4-Dinitrophenol	C6H4N2O5	184.01202				51-28-5	1448	
2,4-Di-tert-butylphenol	C14H22O	206.16707				96-76-4	7037	
2,4-DNT / 2,4-Dinitrotoluene	C7H6N2O4	182.03276				121-14-2	8150	
2,4-Xylidine (2,4-Dimethylaniline)	C8H11N	121.08915				95-68-1	13869462	

Curation – Building the Compound Database



Collection and Curation of MS/MS Spectra

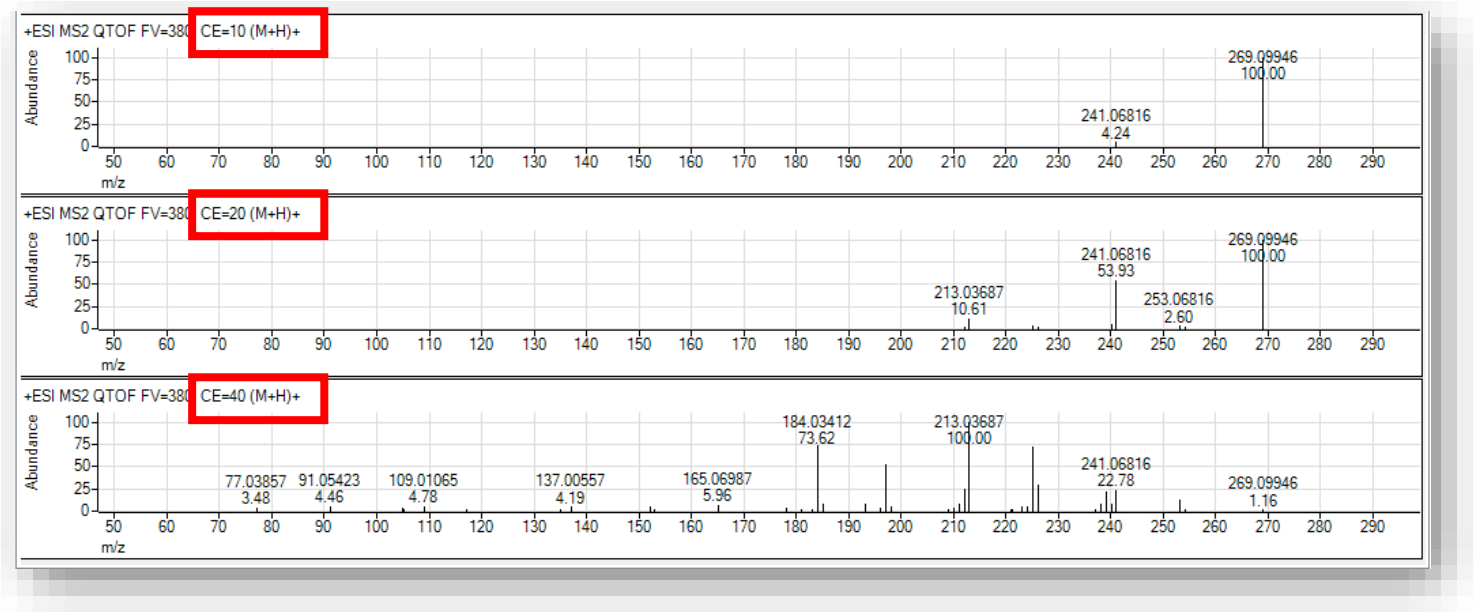
Carefully designed experiments and curation protocols

Data collection

- Flow Injection Analysis (FIA) of pure standards or purified isolates
- Inclusion of commonly analyzed adduct species
- Collected at multiple collisions energies, ion modes, and ion species

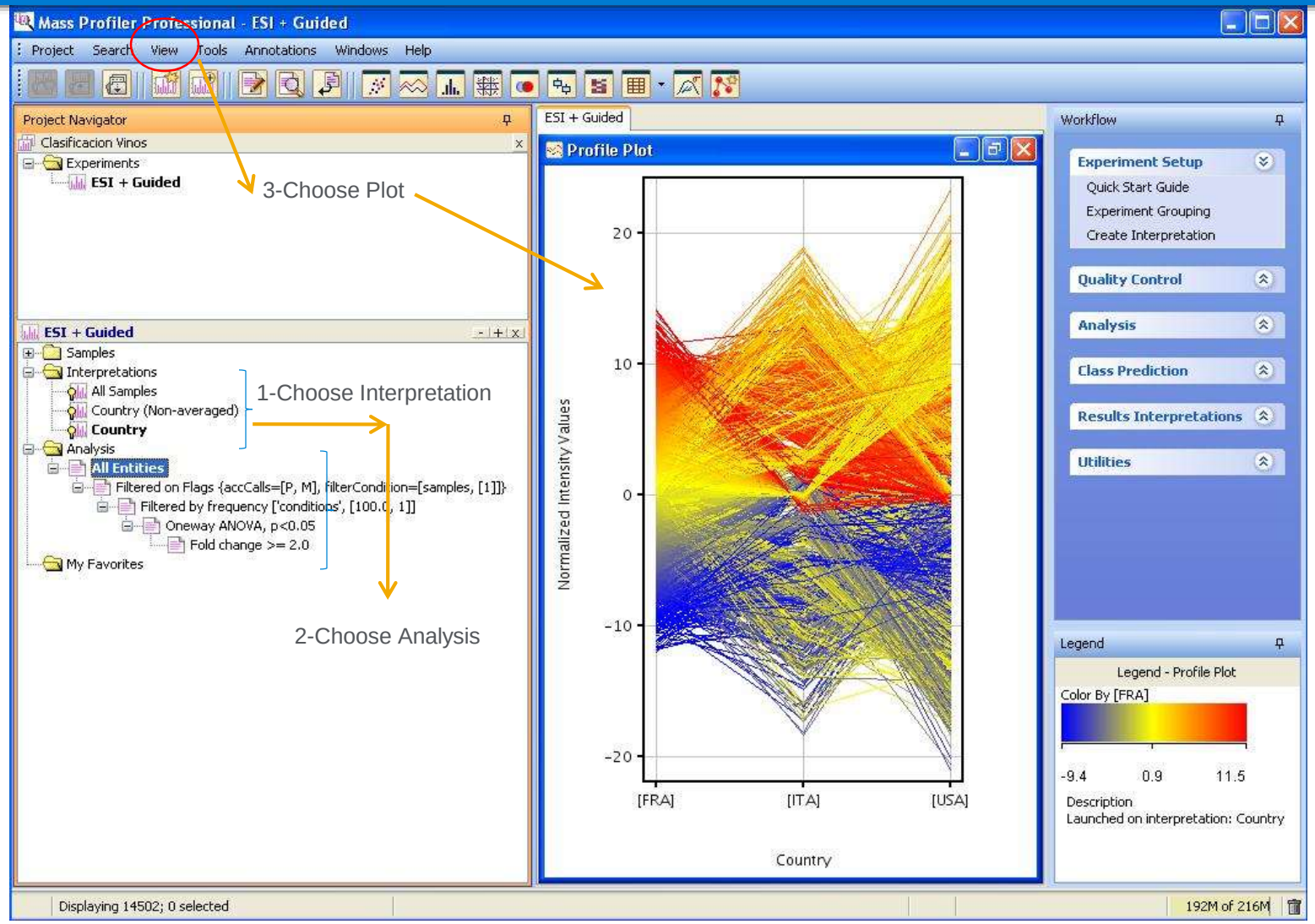
Curation

- Correction to Theoretical accurate mass
- Filtered for signal intensity and curated for spectrum noise and chemical impurities



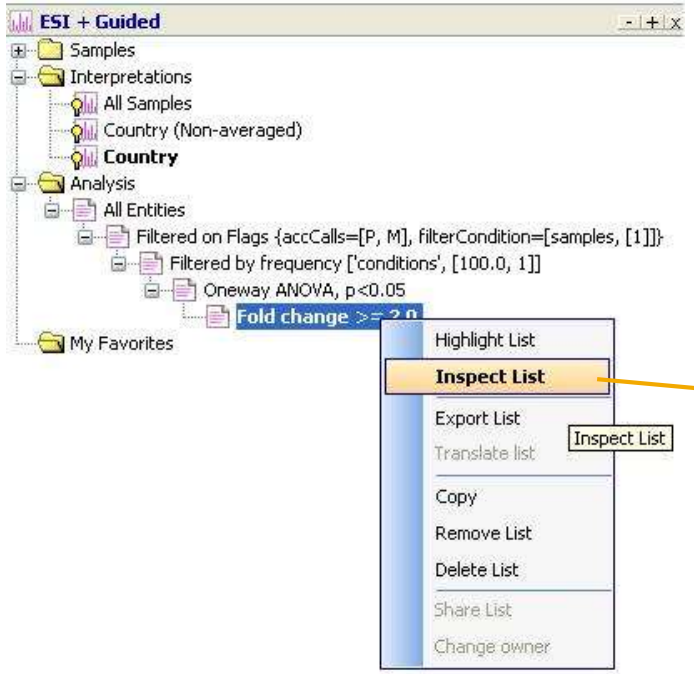
Mass Profiler Professional

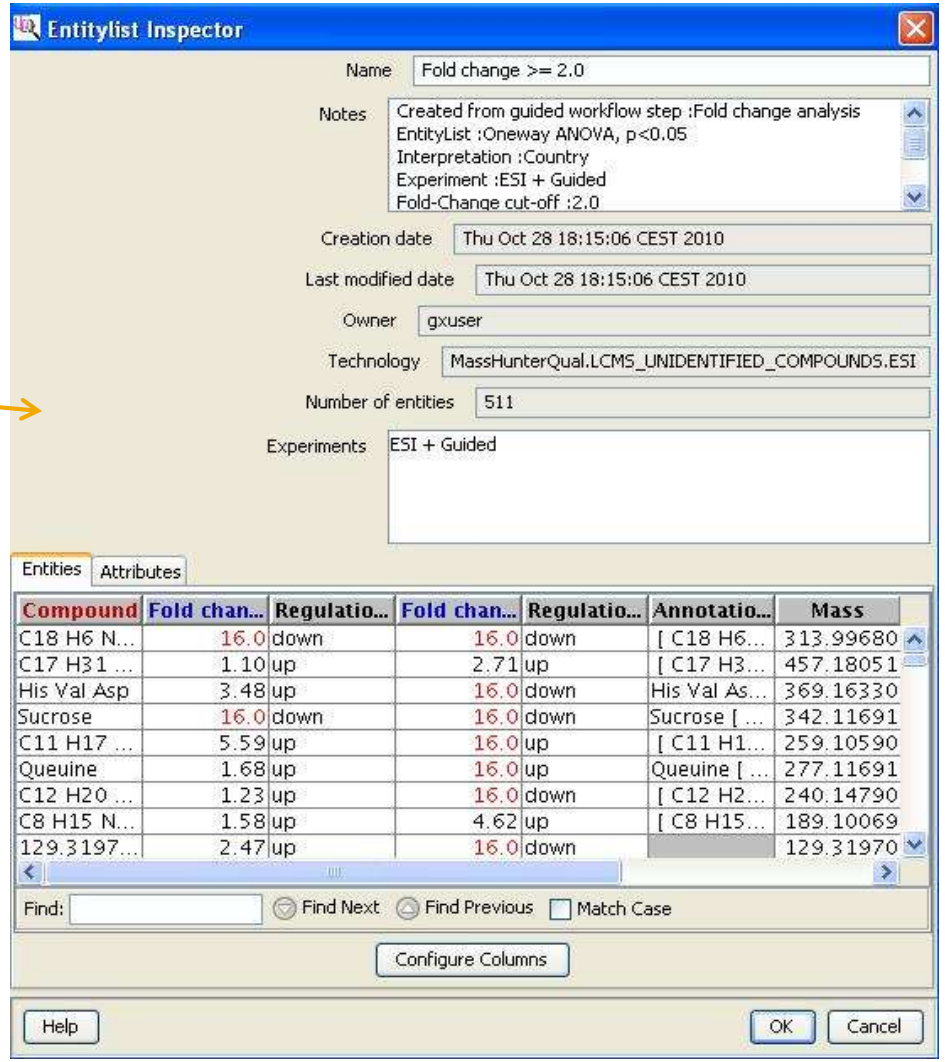
After Guided Worklfow



Mass Profiler Professional

Listing Entities

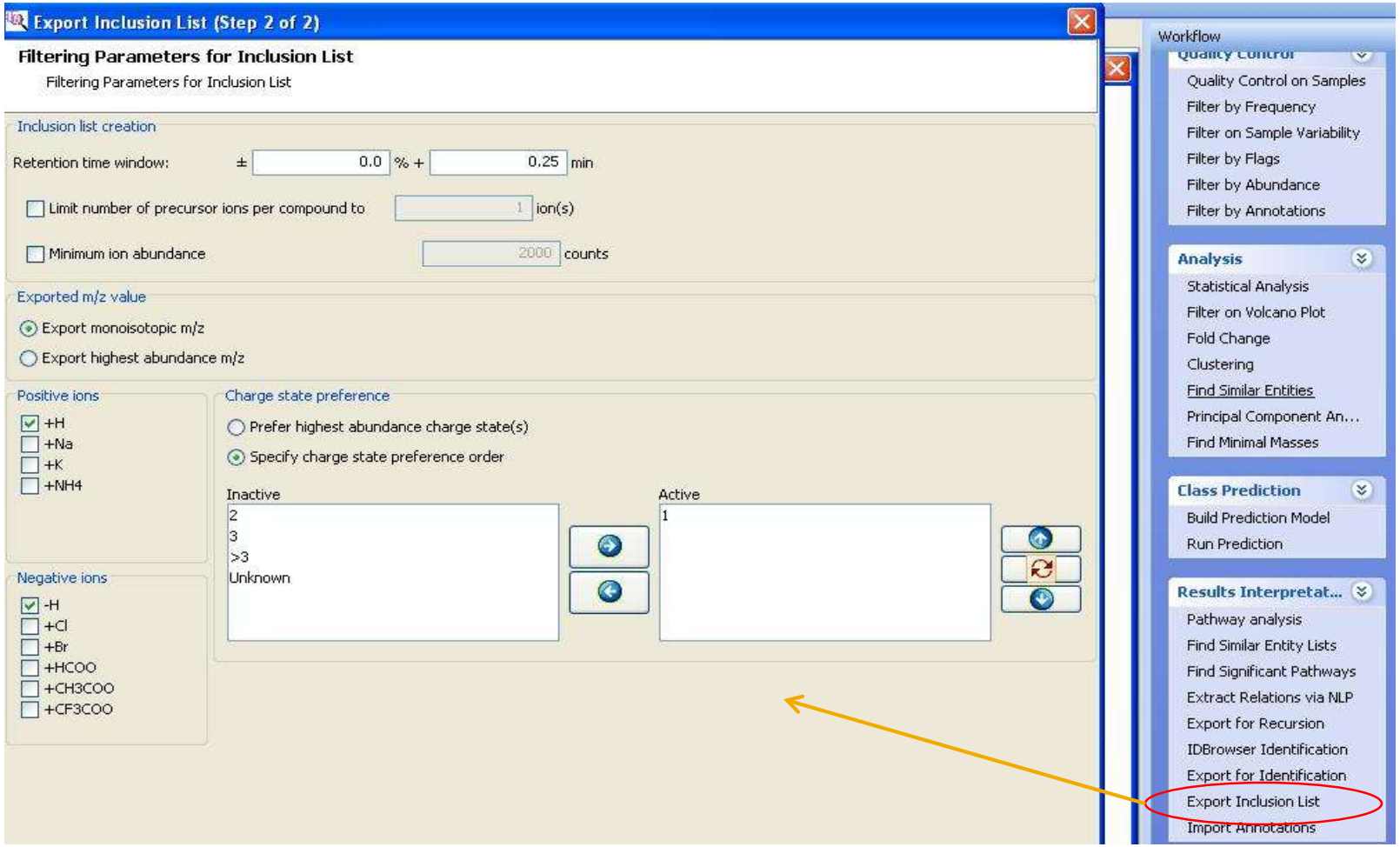




Compound	Fold chan...	Regulatio...	Fold chan...	Regulatio...	Annotatio...	Mass
C18 H6 N...	16.0	down	16.0	down	[C18 H6...	313.99680
C17 H31 ...	1.10	up	2.71	up	[C17 H3...	457.18051
His Val Asp	3.48	up	16.0	down	His Val As...	369.16330
Sucrose	16.0	down	16.0	down	Sucrose [...	342.11691
C11 H17 ...	5.59	up	16.0	up	[C11 H1...	259.10590
Queuine	1.68	up	16.0	up	Queuine [...	277.11691
C12 H20 ...	1.23	up	16.0	down	[C12 H2...	240.14790
C8 H15 N...	1.58	up	4.62	up	[C8 H15...	189.10069
129.3197...	2.47	up	16.0	down		129.31970

Mass Profiler Professional

Exporting Entities

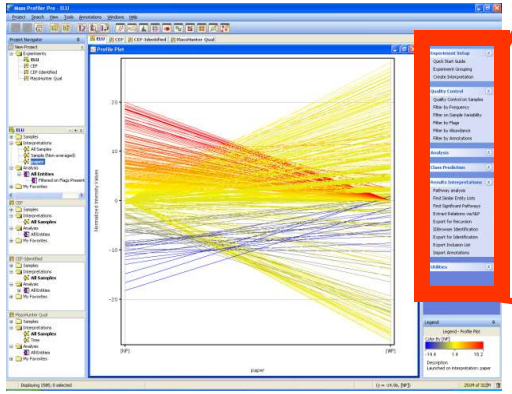


Inclusion list for Target MS/MS on QTOF

Mass Profiler Professional

MPP Advanced Workflow

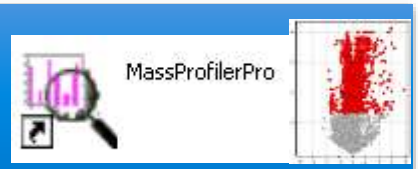
- Use of Experiment Browser
- User defined Interpretations
- User defined Analysis
- Free use of Advanced tools & Utilities



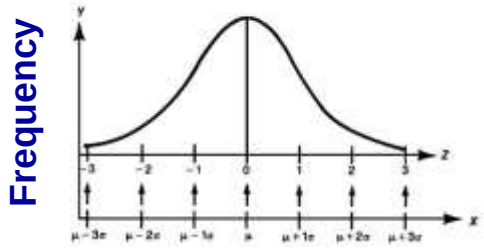
Experiment Browser is designed following usual workflow

Mass Profiler Professional

Analysis- Parametric Test



Parametric test



Bell shaped distribution symmetrical about the mean

MPP assume that you have sampled from populations where expression of Entity X follows a normal distribution

Non- Parametric test

Does not assume normal distribution
Does not assume equal variances
Ranks the order of normalized data across conditions for analyses

With small number of replicates, non-parametric tests have less statistical power than the corresponding parametric tests

P-value Calculation Methods

Asymptotic Method

Permutation Method

Mass Profiler Professional

Multiple Testing Correction . P-value Cut-off

What type of error you are more comfortable with

Type I Error (false positive): Calling entities differentially expressed when they really are not

Type II Error (false negative): Not calling entities differentially expressed when they really are

Choice of cut-off trades off between type I and type II errors

5HT1c	0.002364
NFL	0.002649
NMDA2C	0.017181
aFGF	0.027544
GRa3	0.041179
actin	0.045342
nAChRd	0.046372
EGFR	0.0468
bFGF	0.087842
5HT2	0.106591
Brm	0.137903
SOD	0.147089
mGluR2	0.174708
IGF.I	0.223558
SC2	0.274809
trkC	0.288776
mGluR1	0.313801
SC6	0.343059
CNTFR	0.354717
pre-GAD67	0.366955
BDNF.rat	0.417615
GDNF	0.421125
IP3R2	0.421308
L1	0.443525
GAD67	0.462416
H2AZ	0.561907
IP3R1	0.573717
MK2	0.630177
CCO2	0.640797
mGluR3	0.654866
PDGFa	0.659352
IGF.II	0.683554
CNTF	0.690512
nAChRe	0.701041
IGFR2	0.728141
GAP43	0.732078
ODC	0.745628
SC1	0.74575
NT3	0.78811
PTN	0.795557
trk	0.82403
mGluR5	0.8305
cjun	0.839991
Ins2	0.841945
MAP2	0.851833
neno	0.879299
GRb1	0.888485
TCP	0.892361
GRb2	0.900601
S100beta	0.930265

- p-value=0.05

Truly differentially expressed

Unchanged between populations

Mass Profiler Professional

Multiple Testing Correction

1 entity = 1 individual test

- 10000 entities = 10000 tests
- With *p*-value = 0.05 and analyzing 10000 entities: 500 entities (0.05 x 10000) likely to appear significant by chance
  - Number of false positives increases proportionally to number of tests being performed

Performing multiple testing correction further decreases the number of false positives

Bonferonni FWER
Bonferonni Holm FWER
Benjamini Hochberg FDR
No Correction

More false negatives

More false positives

> **Family-wise Error Rate (FWER) - Very conservative and does not tolerate any false positives**

> **False Discovery Rate (FDR) - False positives a percentage of called entities**

> **None - False positives a percentage of entities being tested**

Mass Profiler Professional

Multiple Testing Correction . Post Hoc Tests

Decreases rate of false positives, only available in MPP for one-way ANOVA tests

MPP test options:

Tukey's Honestly Significant Difference (HSD) test

Student-Newman-Keuls (SNK) test

Pink boxes contain entities whose expression does *not* differ significantly between the two conditions

Test Description

SNK Post Hoc test. Entities found to be differentially expressed a
in the blue boxes, while entities found not to be differentially exp
represented in the orange boxes. To save entities of interest as
select one or multiple boxes and click on the "Union" or "Intersec

Result Summary

Group Name	[Pinot Noir]	[Cab Sav]	[Merlot]
[Pinot Noir]	39	38	12
[Cab Sav]	1	39	25
[Merlot]	27	14	39

Union Intersection

Blue boxes contain entities whose expression differs significantly between the two conditions

Mass Profiler Professional

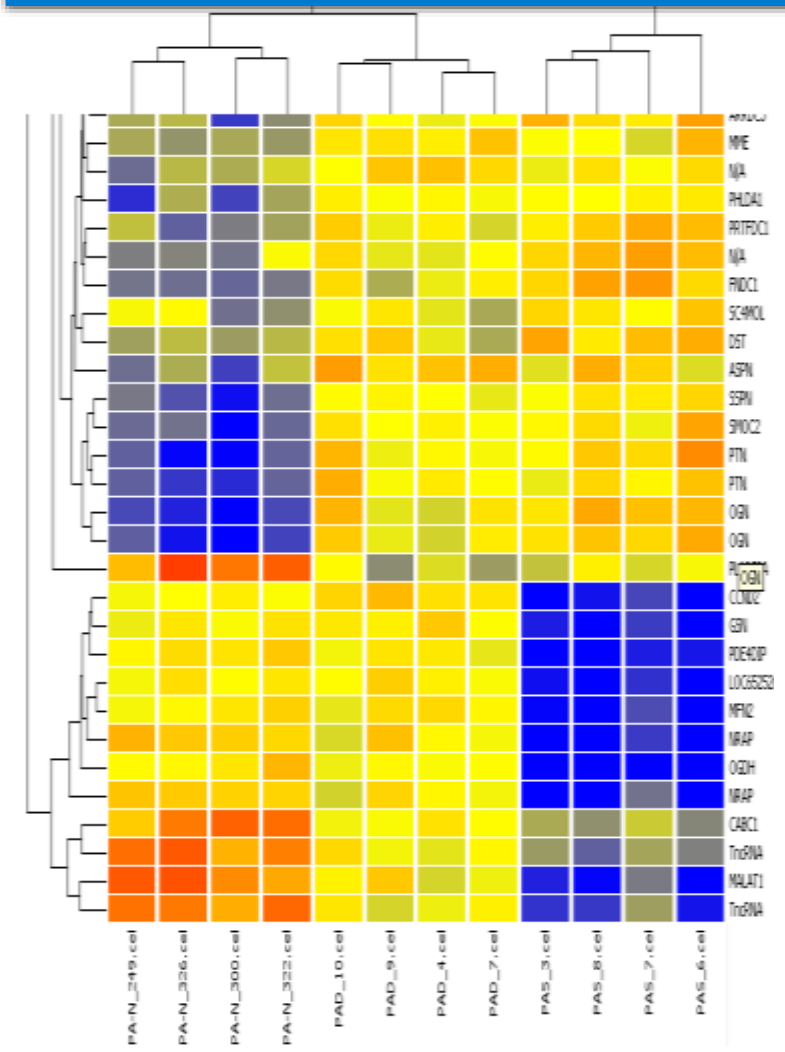
CLUSTERING

Clustering is an *unsupervised* method for identifying patterns within datasets.

1) What do you want to cluster together?
2) What similarity metric to select?
3) What clustering algorithm will be applied?

Mass Profiler Professional

Hierarchical CLUSTERING



Overview:

- Hierarchical clustering algorithm can be used to group entities and conditions based on the similarity of their expression profiles
- Performing Hierarchical clustering on both entities and conditions result in a 2-dimensional dendrogram
- Most similar profiles are joined together into a group and groups are further joined in a tree structure until all data forms a single group

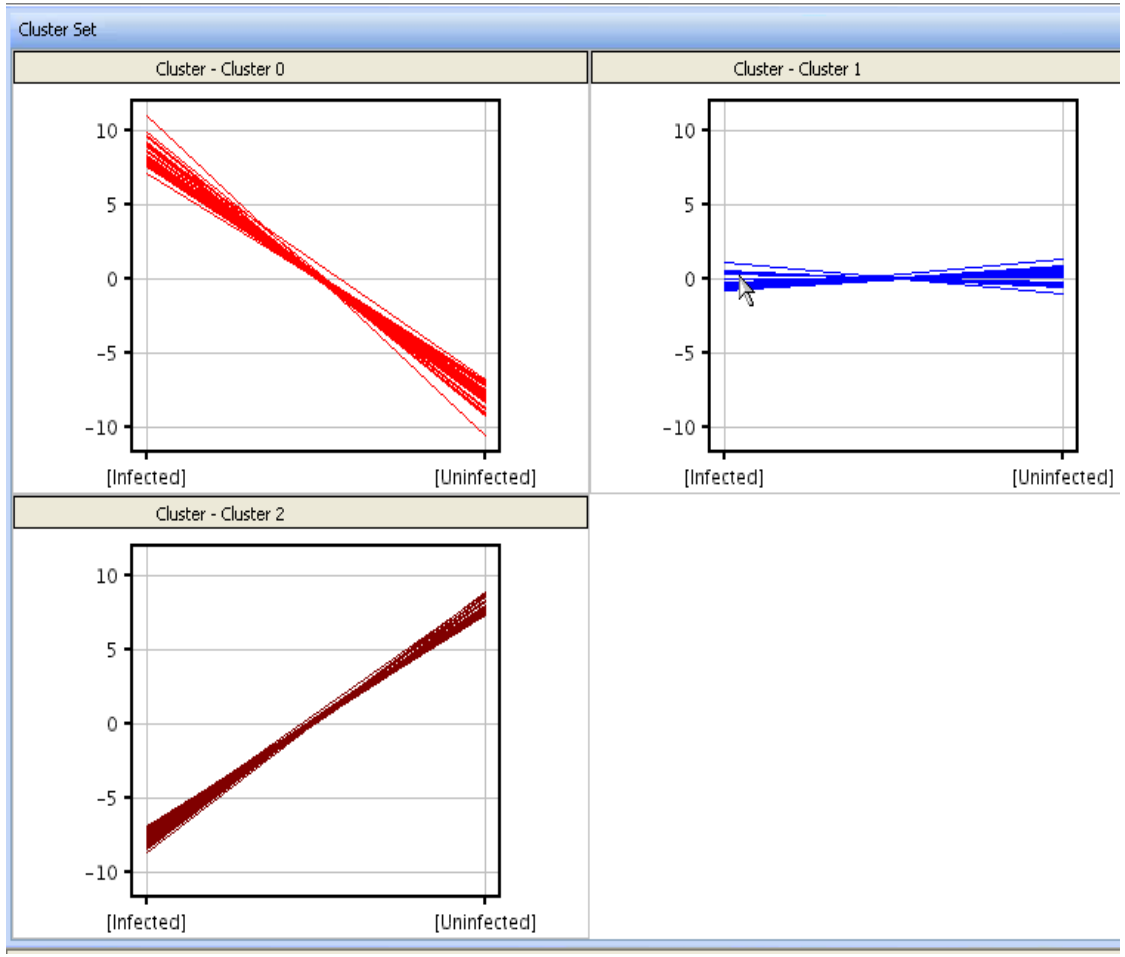
Mass Profiler Professional

K-means CLUSTERING

Overview:

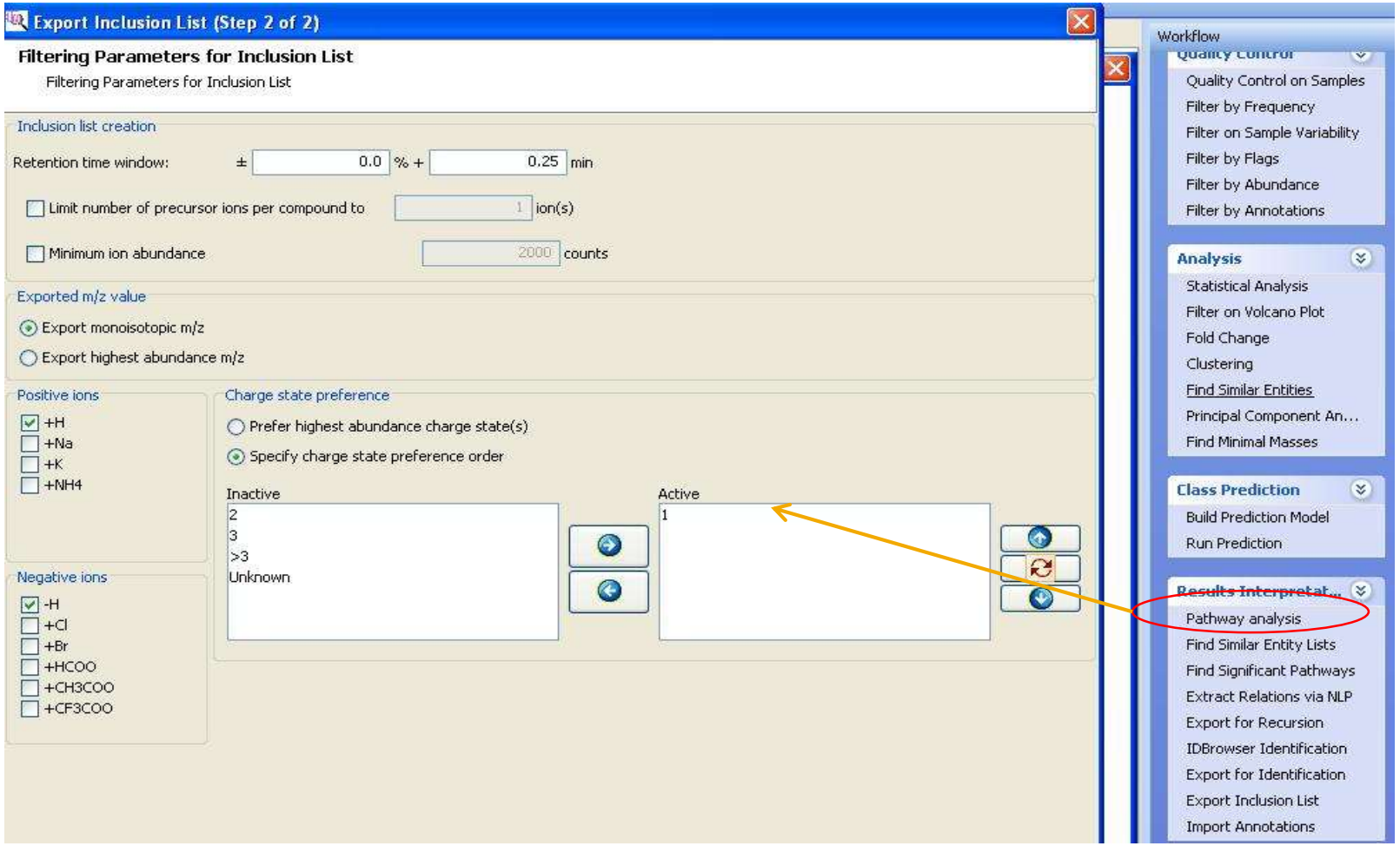
- User chooses K, the number of clusters to partition selected entities or conditions into
- Algorithm attempts to minimize intra-cluster variability and maximize inter-cluster variability

Groups of Compounds that behave similarly



Mass Profiler Professional

Patways Analysis



Inclusion list for Target MS/MS on QTOF

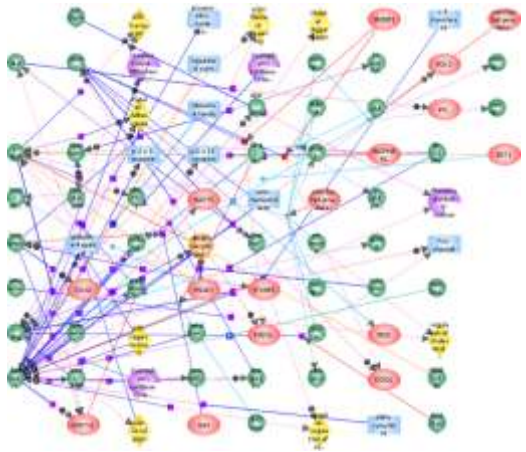
Agenda

- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - ***¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil…. Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - ***Mass Profiler professional. Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - ***Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos? Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - ***Análisis de rutas Metabólicas a través de "Pathways Analysis". Biología integrada e interpretación biológica de mis datos.Pathways Analysis.***
  - *¿Cuál es mi próximo experimento?* ***La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

Mass Profiler Professional

Pathway Analysis

Pathway Anaylsis allows for Finding Biological Relevance of differencial entities



Two types of pathway analysis in MPP:

1. <u>Find Significant Pathways</u>:

   Is there a significant enrichment of my entities of interest in a particular pathway?

2. <u>Pathway Analysis</u>:

   How do my entities of interest interact in a biochemical network?

Mass Profiler Professional

Pathway Analysis

BioPAX (Biological Pathway Exchange) is a standard pathway data exchange format.

Pathways in the biopax format will have the extension .owl

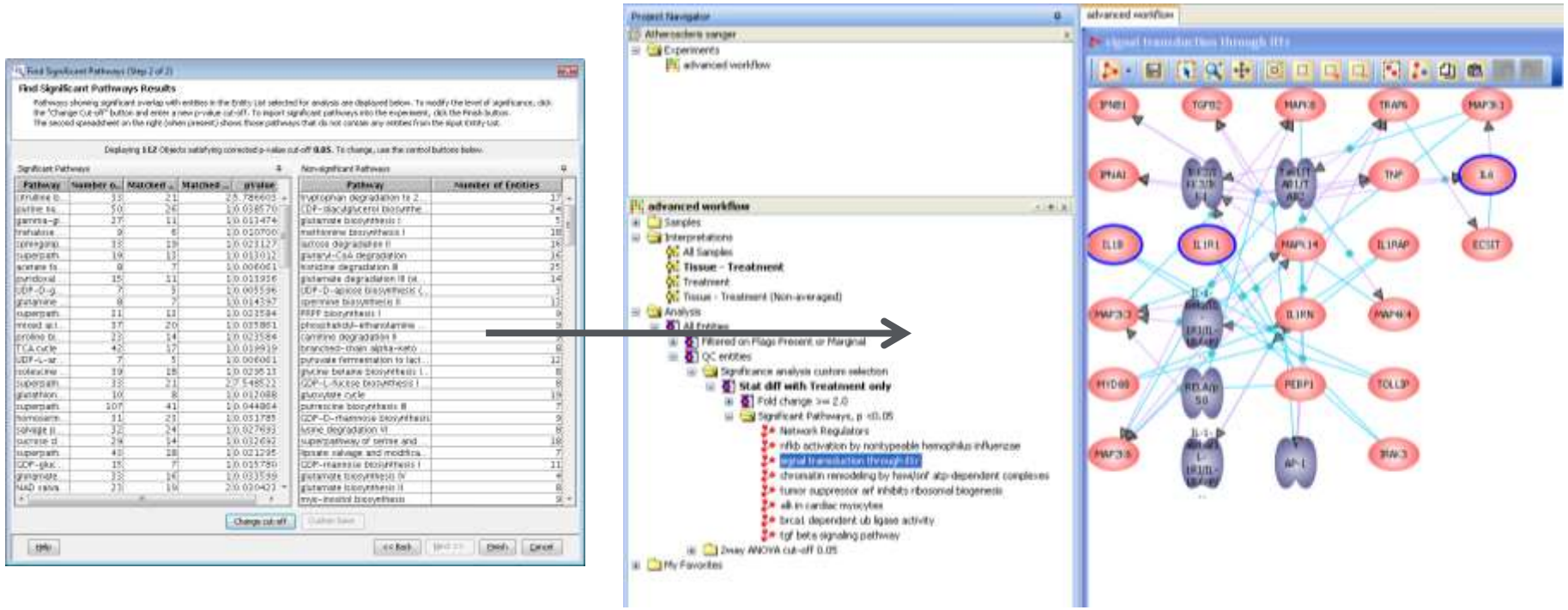
MPP users can import pathway data standard pathway sites in BioPAX level 1 or 2 format

www.pathguide.org is a useful website which list repositories of pathways

Databases for any organism of interest can be created using the Biopax files- Rice, zebra fish, chimpanzee, dog

Mass Profiler Professional

Pathway Analysis – Find Significant Pathway

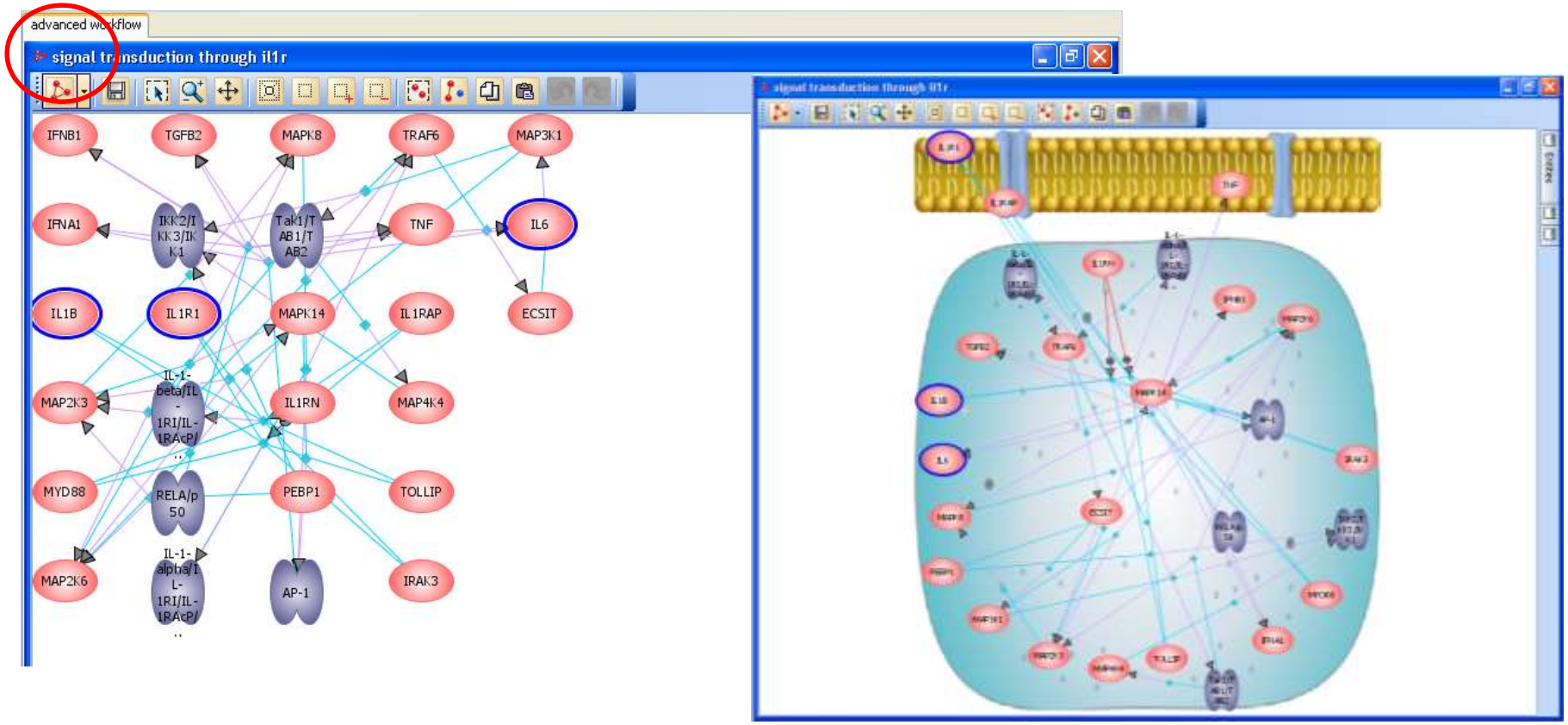


Is there a significant enrichment of my entities of interest in a particular pathway?

Analysis will be performed on every pathway that has been imported into MPP for the matching organism and every pathway created in MPP

Mass Profiler Professional

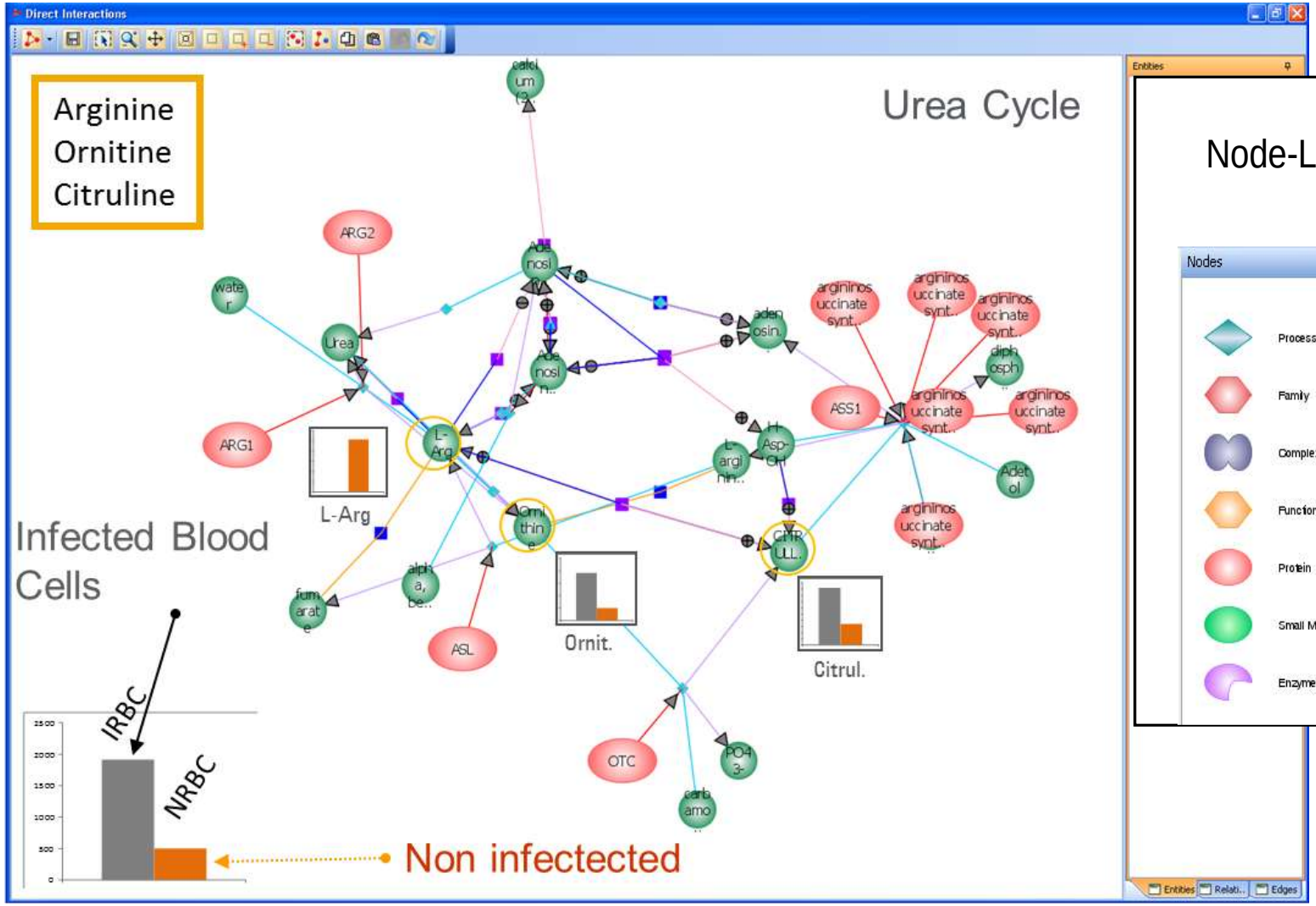
Pathway Viewer

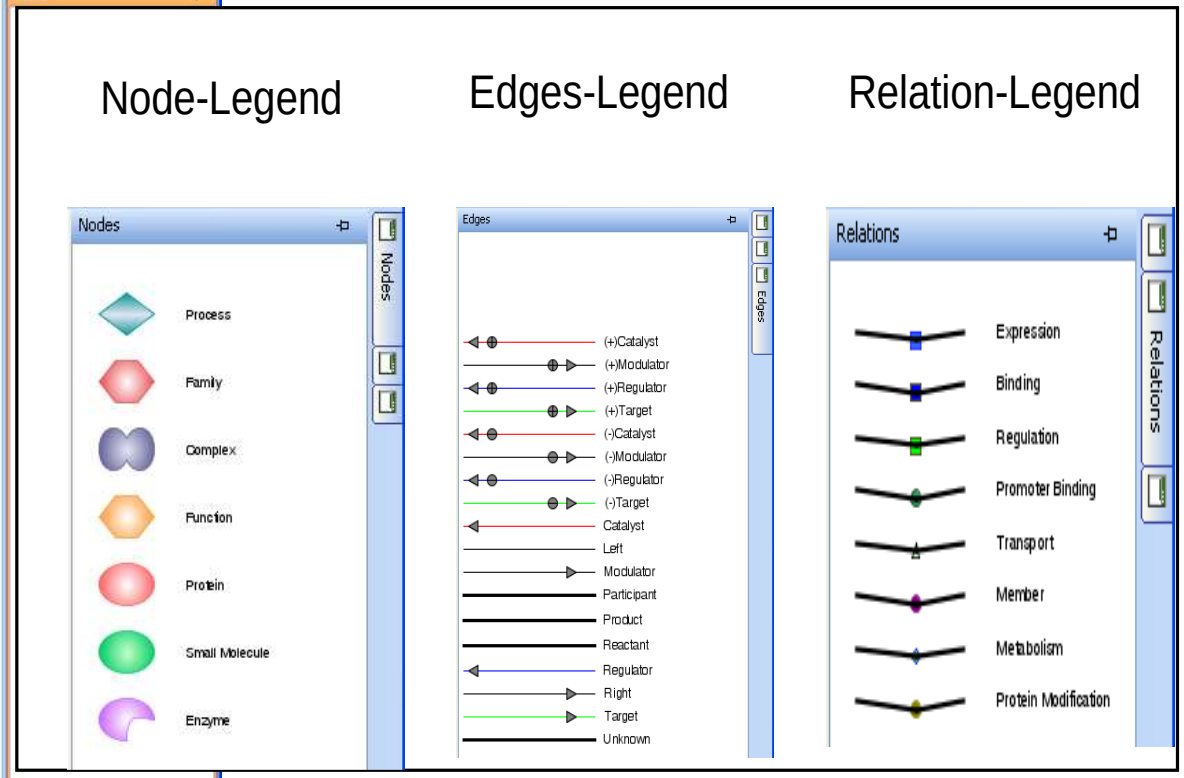


Layout of entities can be changed – 6 options including cellular view

Mass Profiler Professional

Pathway Viewer

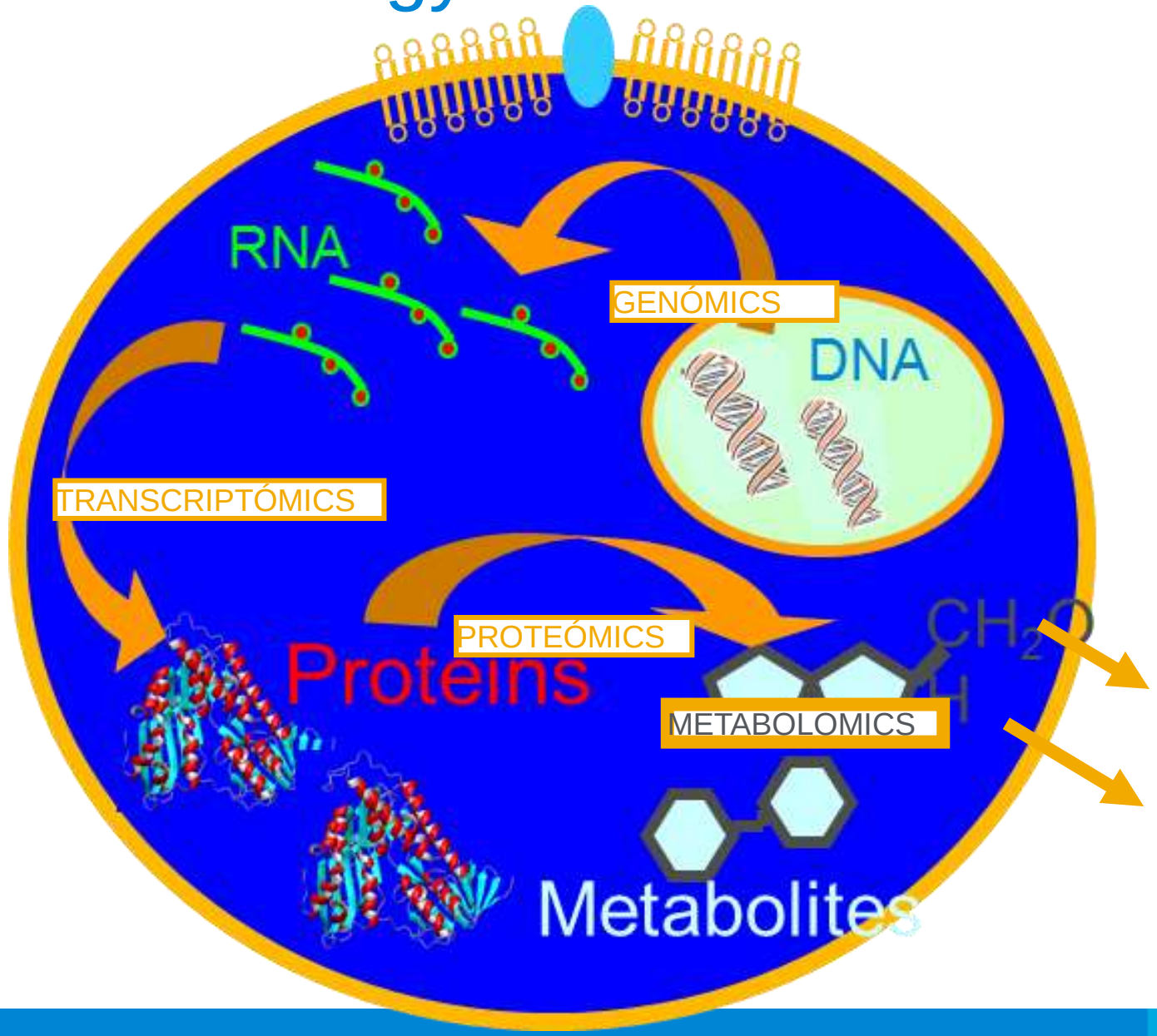


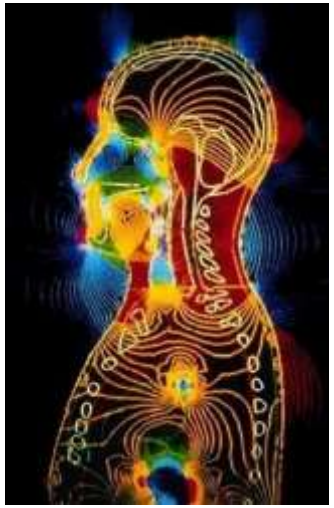


Agenda

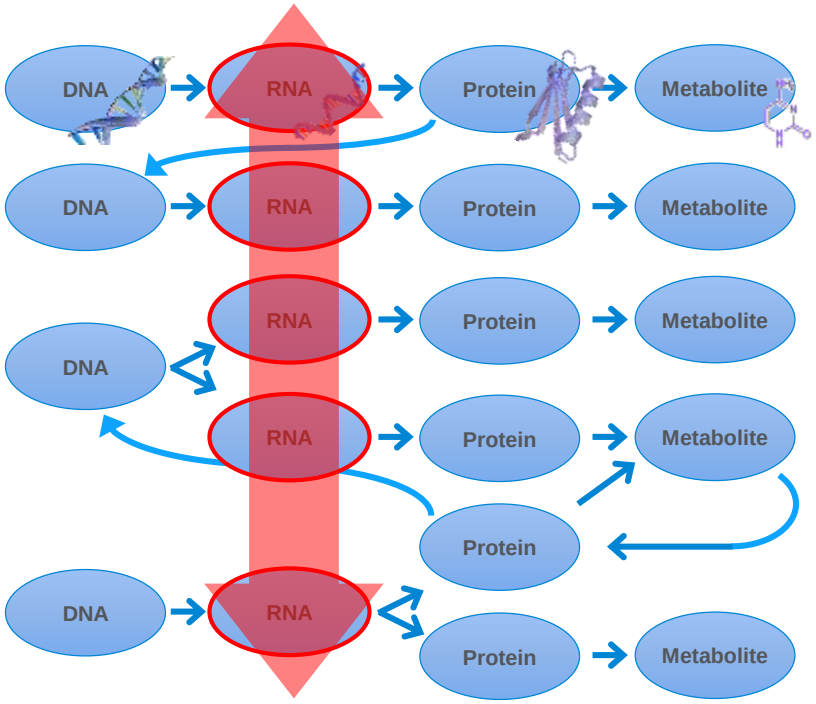
- *Enfoques y estrategias analíticas que nos permiten las ultimas tecnologías de LC Mass Spectrometry de alta resolución.* ***UHPLC QTOF technical details***
- ***Agilent proposal Workflows in different scenarios.*** *Diferentes estrategias inherentes a investigación en el ámbito de Ciencias de la Vida, así como para el perfilado de Alimentos, Materiales, Procesos…*
- ***Herramientas de Agilent y flujos de Trabajo para tomar mejores decisiones en un entorno de Biología integrada.*** *Del diseño experimental a las conclusiones, un largo camino para ayudar al investigador..* ***:***
  - ***Datos según modos de Adquisición.*** *Complementariedad de las diferentes tecnologías LCMS, GCMS, CEMS, ICPMS*
  - ***Deconvolución de datos y herramientas de visualización****. Como funcionan los algoritmos de Agilent para extraer información de compuestos de un Full Scan.*
  - ***Preparación de datos previa al Análisis Estadístico diferencial. Alineamiento, Normalización, "Baselining" con "Mass Hunter ProFinder".***
  - ***¿Necesito análisis recursivo a través de iteración? Por favor hágamelo fácil…. Exhaustivo tratamiento de datos para evitar la Perdida de compuestos.***
  - ***Mass Profiler professional. Análisis Diferencial a través de Interpretaciones, Clustering, PCA, PLRS, modelos de predicción***
  - ***Así, ¿Cuales son mis compuestos diferenciales de interés? ¿Como puedo identificarlos? Librerías empíricas de espectros MS/MS. Agilent METLIN PCDL.***
  - ***Análisis de rutas Metabólicas a través de "Pathways Analysis". Biología integrada e interpretación biológica de mis datos.Pathways Analysis.***
  - ***¿Cuál es mi próximo experimento? La potencia del enfoque de la Biología integrada.***
- ***Movilidad Iónica.*** *Una nueva dimensión para extracción de datos más selectiva en muestras complejas. Una nueva herramienta de identificación*
- ***Fluxómica. Fácil y rápida*** *visualización de la incorporación de sustratos marcados isotópicamente en una ruta metabólica a través de "VistaFlux".*
- *Método llave en mano para el análisis Metabolómico dirigido en rutina de los* ***metabolitos del Ciclo Central de Carbono***
- *Sinergias con la medida In-vivo del Metabolismo celular con "Seahorse".*

Classical Biology Process



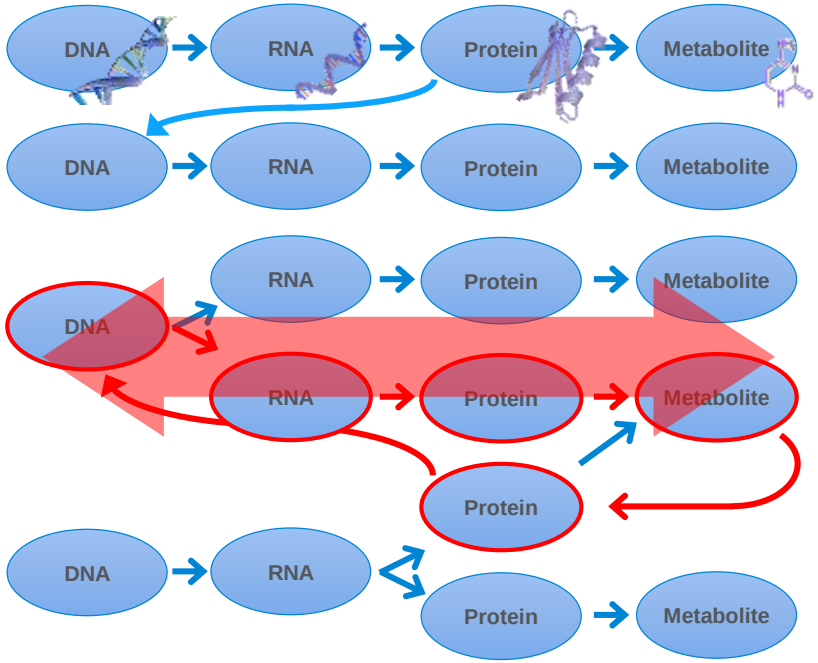


The Biology Challenge



"-Omics"

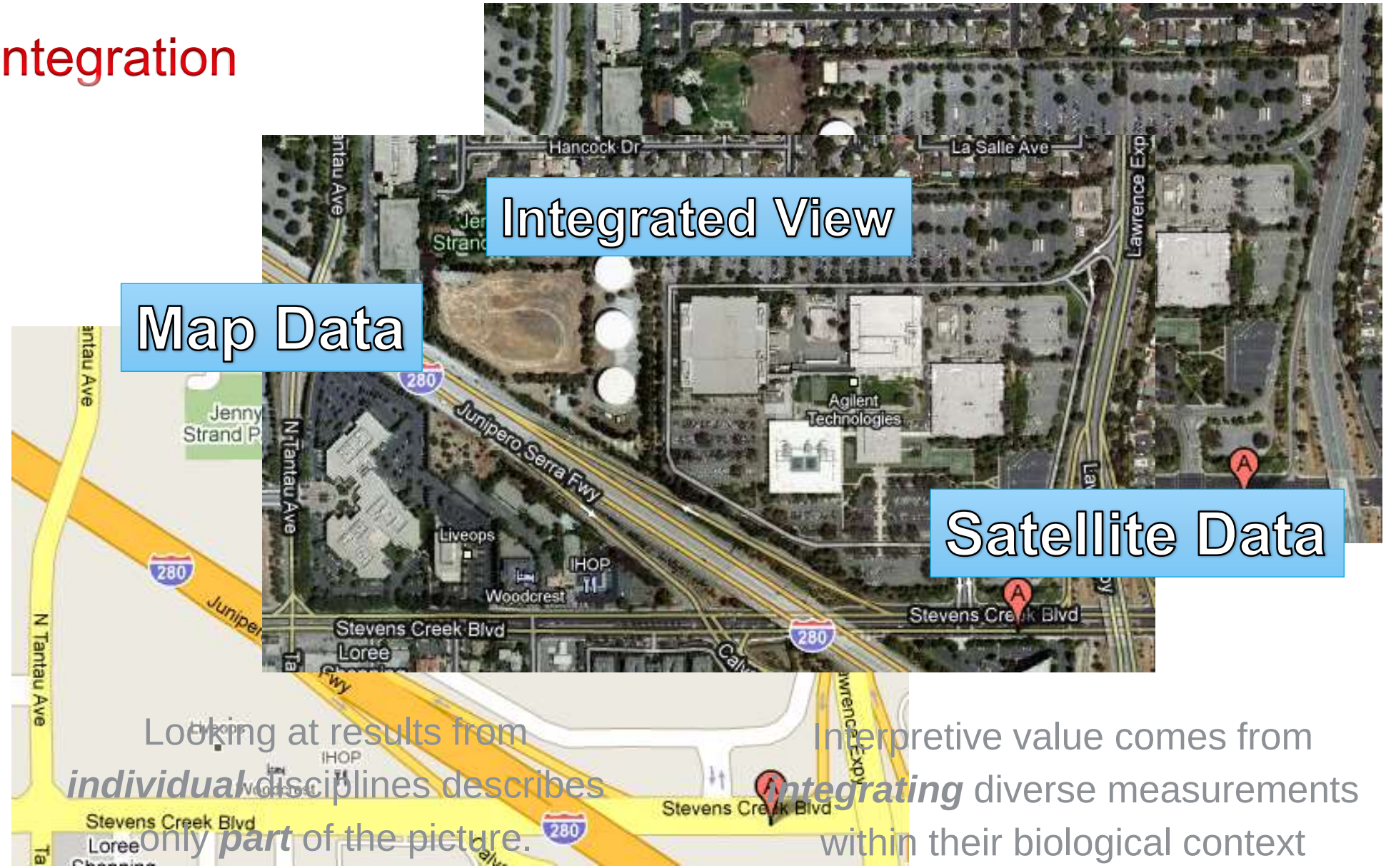
"Classical Biology" approach



Biological Processes

"Integrated Biology" approach

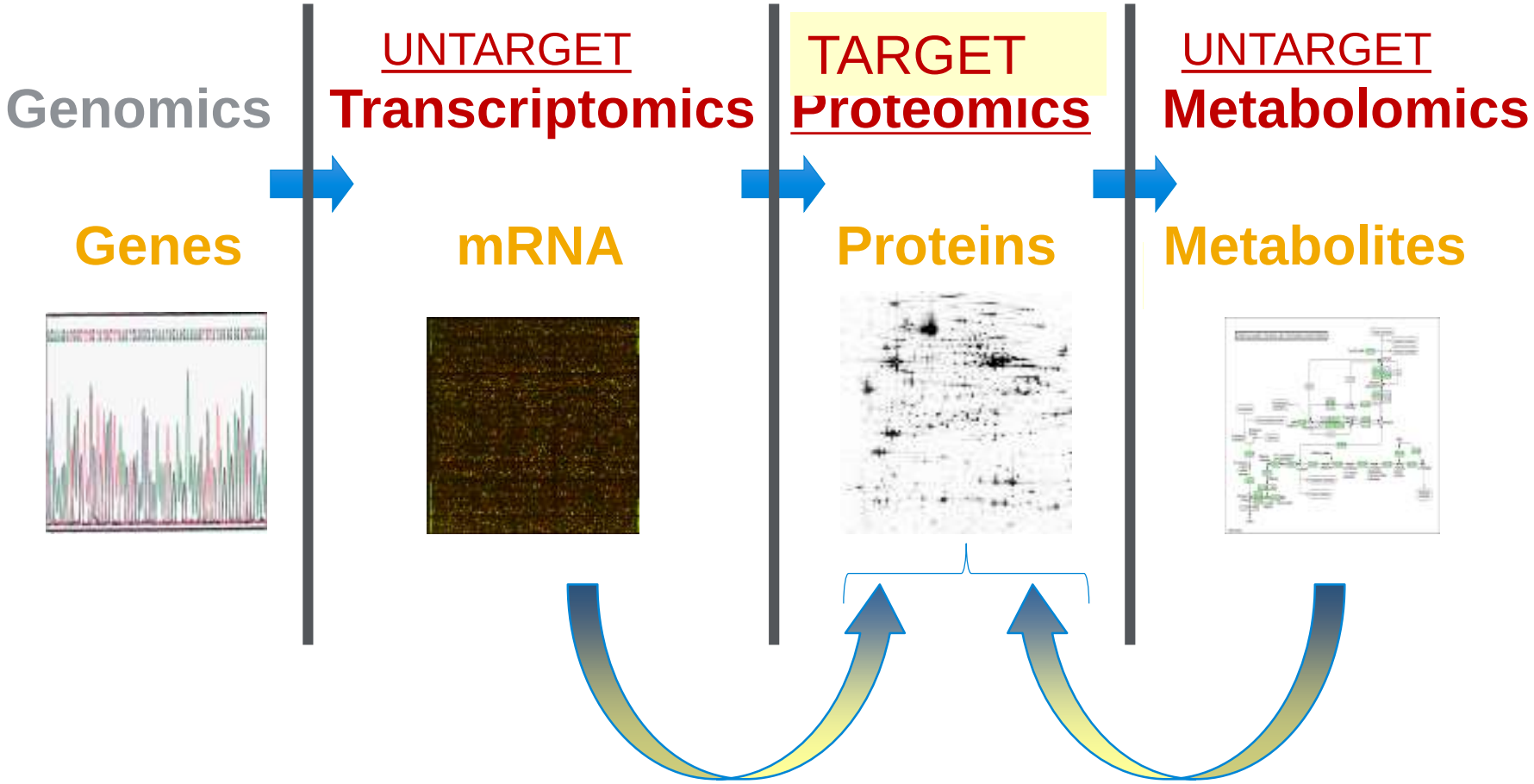
IBS
Data Integration



Looking at results from ***individual*** disciplines describes only ***part*** of the picture.

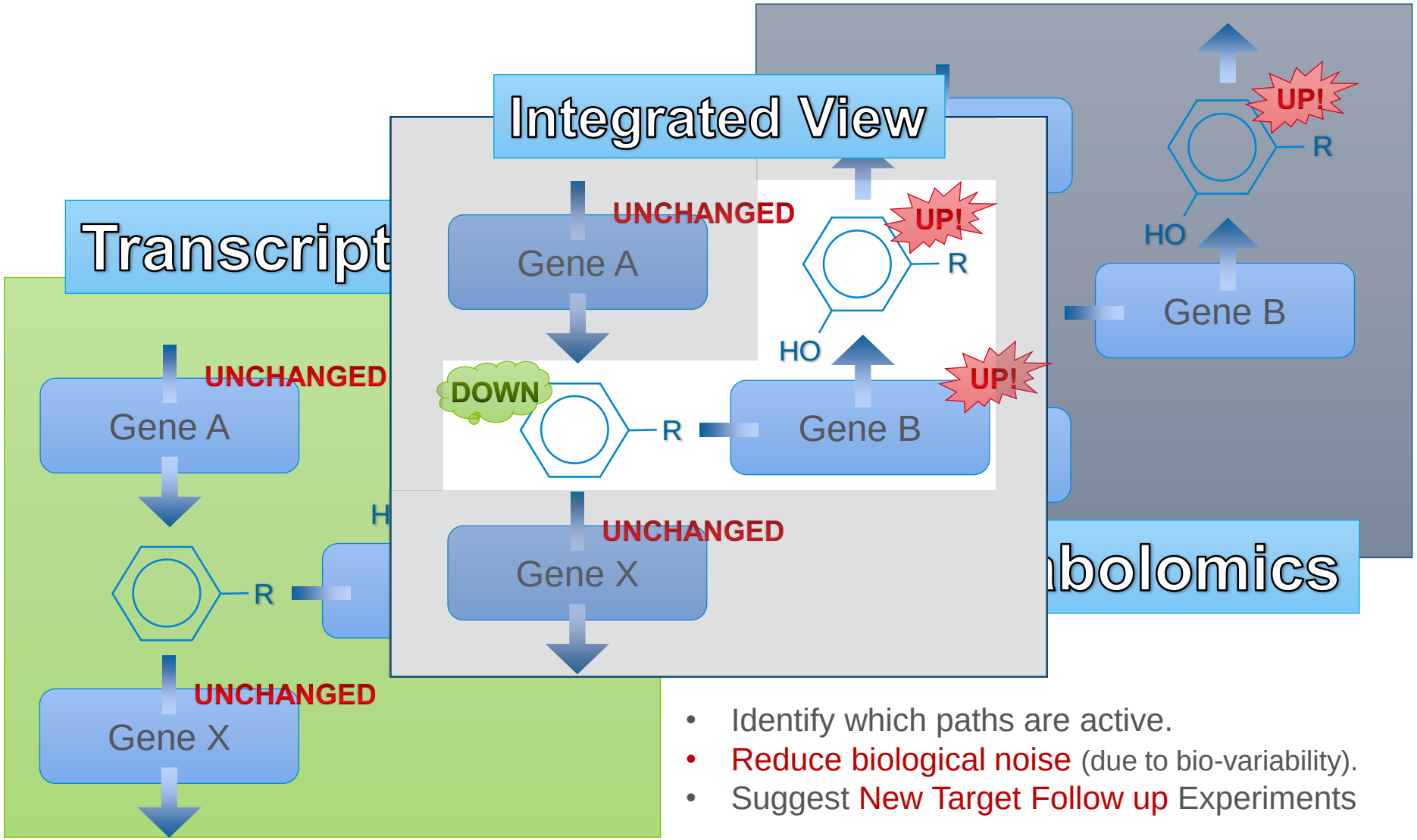
Interpretive value comes from ***integrating*** diverse measurements within their biological context

Workflow Strategy to Address Integrated Biology Studies



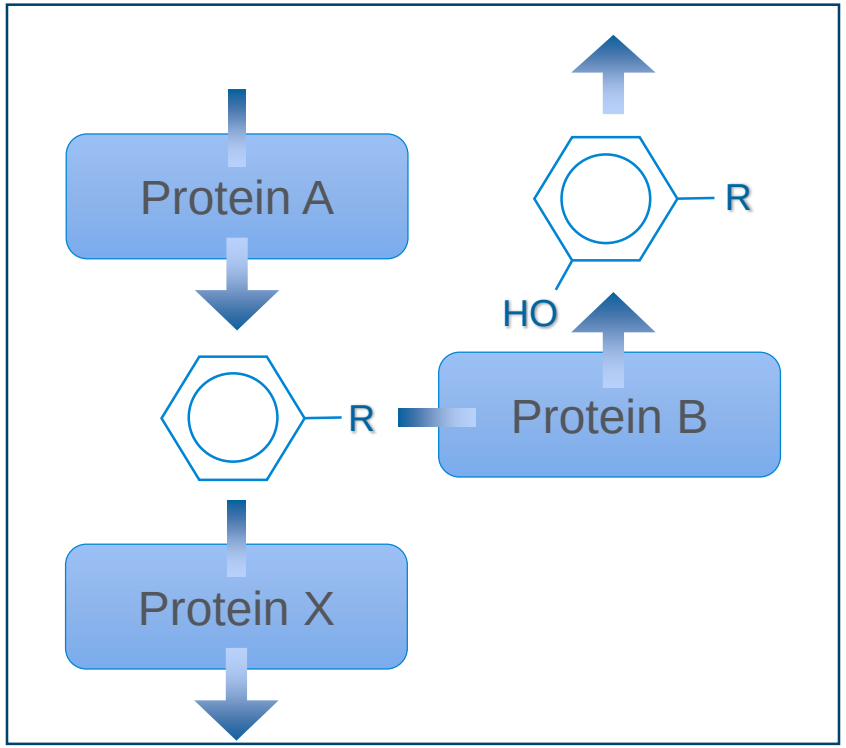
Multi-omics approaches helps a lot to reduce "biological samples noise" and IMPROVES RELIABILITY OF FINDINGS.

Metabolic Pathways: the Key Link for multi -omics Studies



- Identify which paths are active.
- Reduce biological noise (due to bio-variability).
- Suggest New Target Follow up Experiments

Integrating Biological Analysis Using Pathways



Sources

- **WikiPathways**
- **BioCyc/MetaCyc**
- **Generalized BioPax**
- **KEGG**

Platforms

- **GeneSpring**
- **Mass Profiler Professional**
- **Pathway Architect**

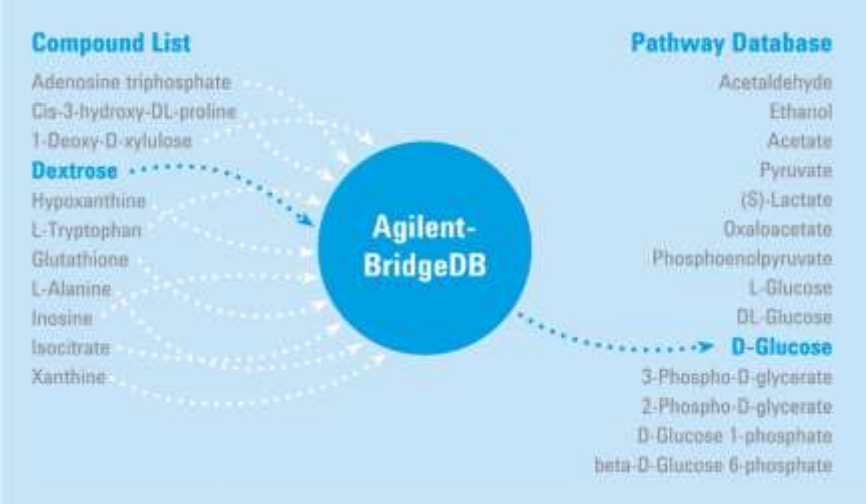
Agilent-BridgeDB

Resolving the Mapping Problem Between Databases

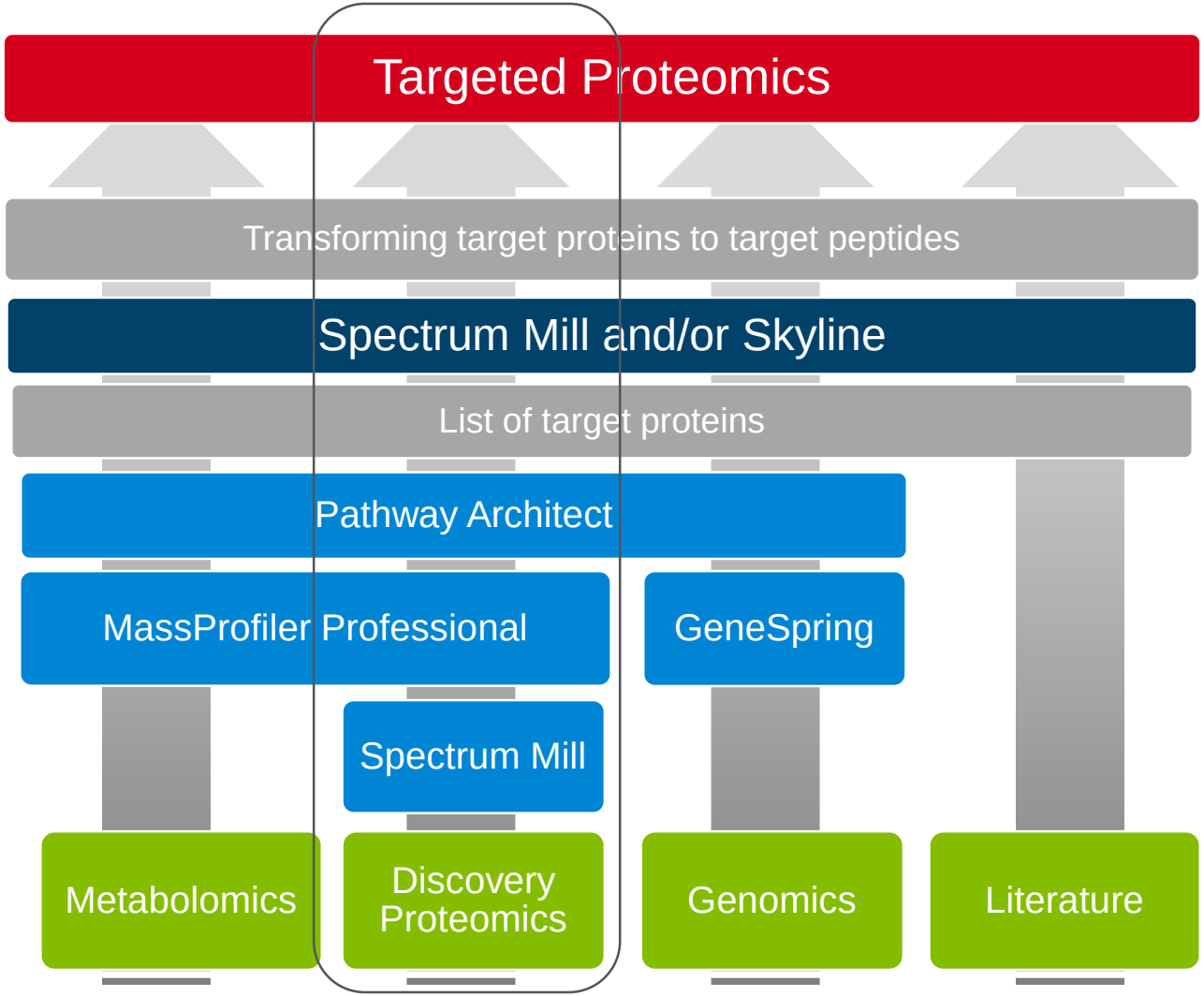
Solves the translation problem of identification names to pathway database names

Automatic – does not require user intervention

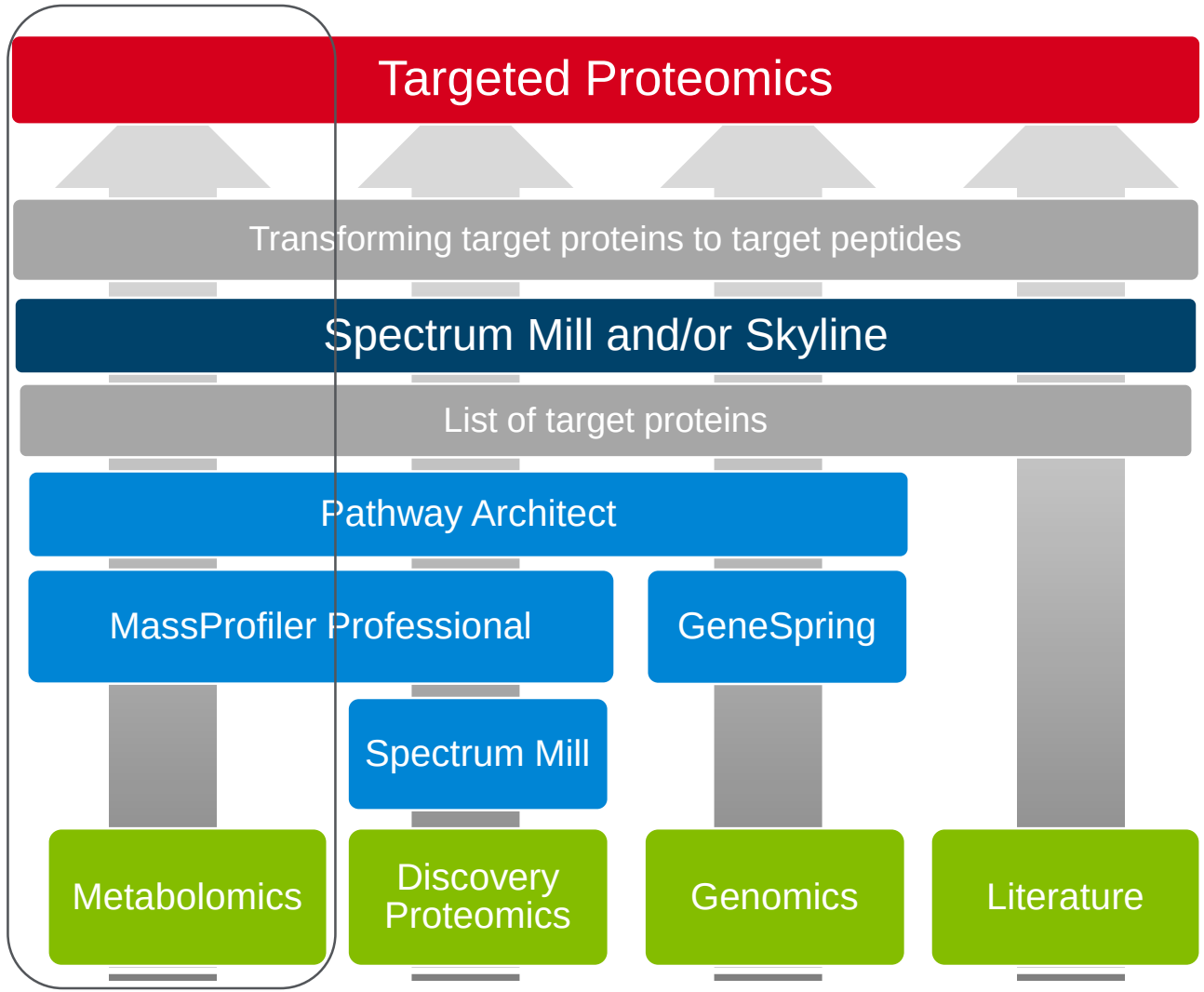
- Metabolites Identifiers
  - KEGG, MetaCyc, PubChem. LMP, HMDB, ChEBI, and CAS
- Proteins Identifiers:
  - Swiss-Prot, UniProt, and UniProt/TrEMBL
- Genes Identifiers:
  - Entrez Gene, GenBank, Ensembl, EC Number, RefSeq, UniGene, HUGO, HGNC, and EMBL



Biology-directed Workflows to Targeted Proteomics



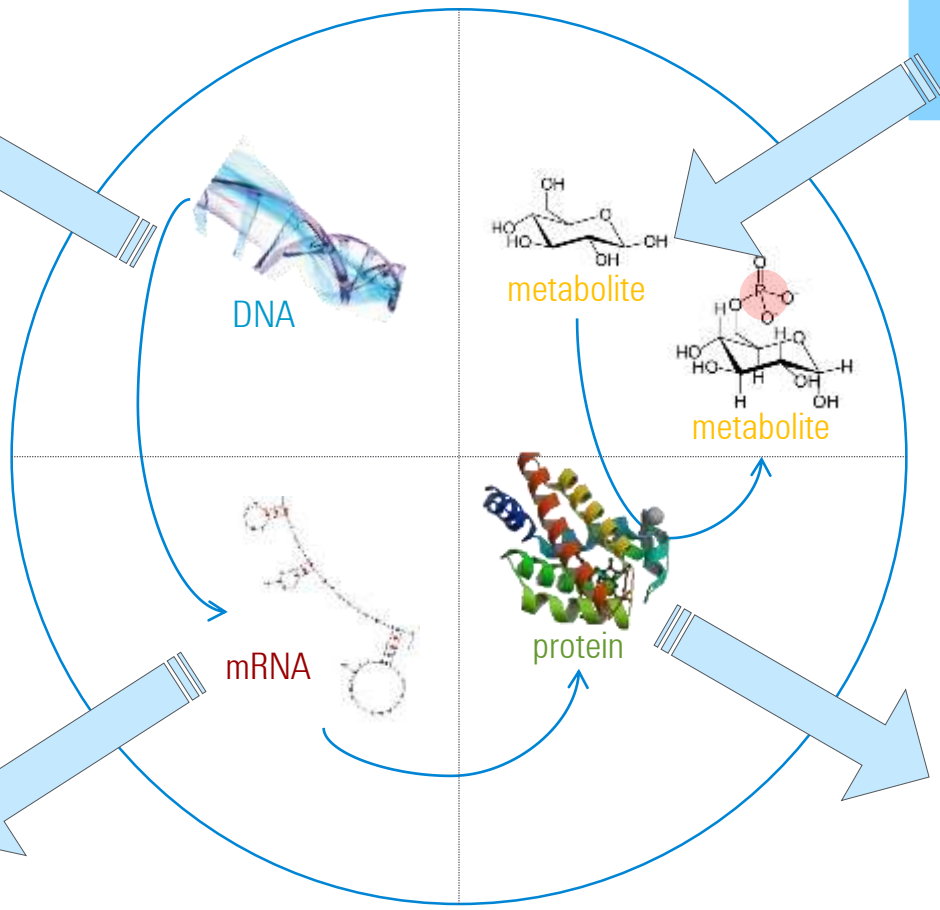
Biology-directed Workflows to Targeted Proteomics



What is the Next Experiment?

From a metabolic pathway I've identified…

- Missense or nonsense mutation?



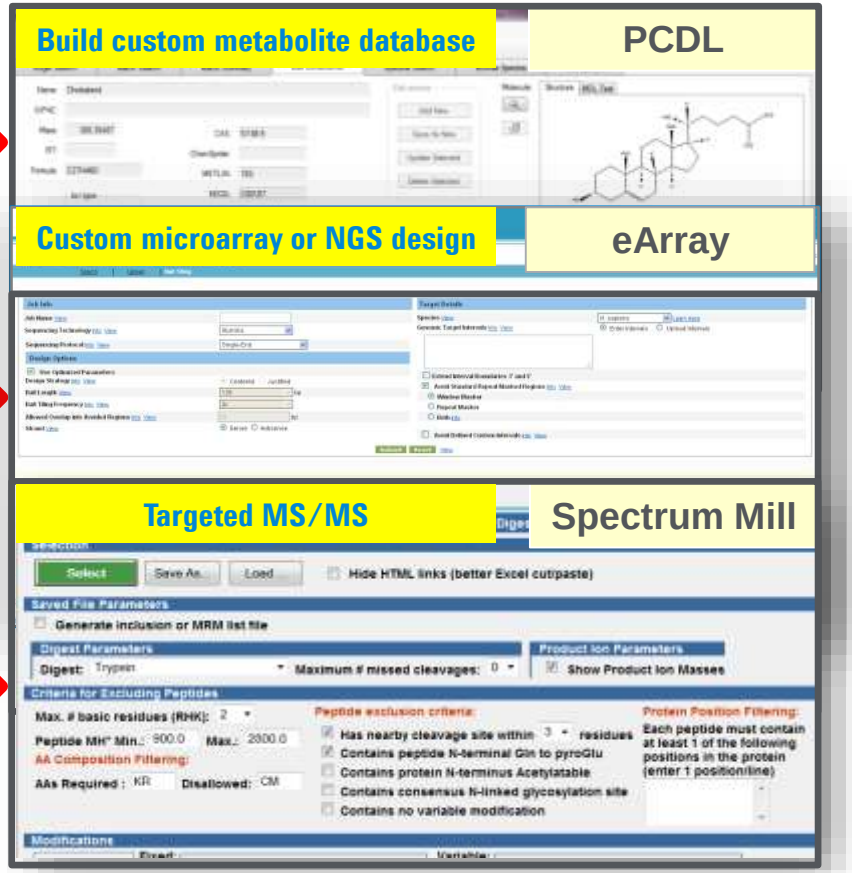
- Differential Protein abundances
- Targeted analyses

- Transcriptional regulation?
- Splice variants?

Pathway Directed Experiment Creation

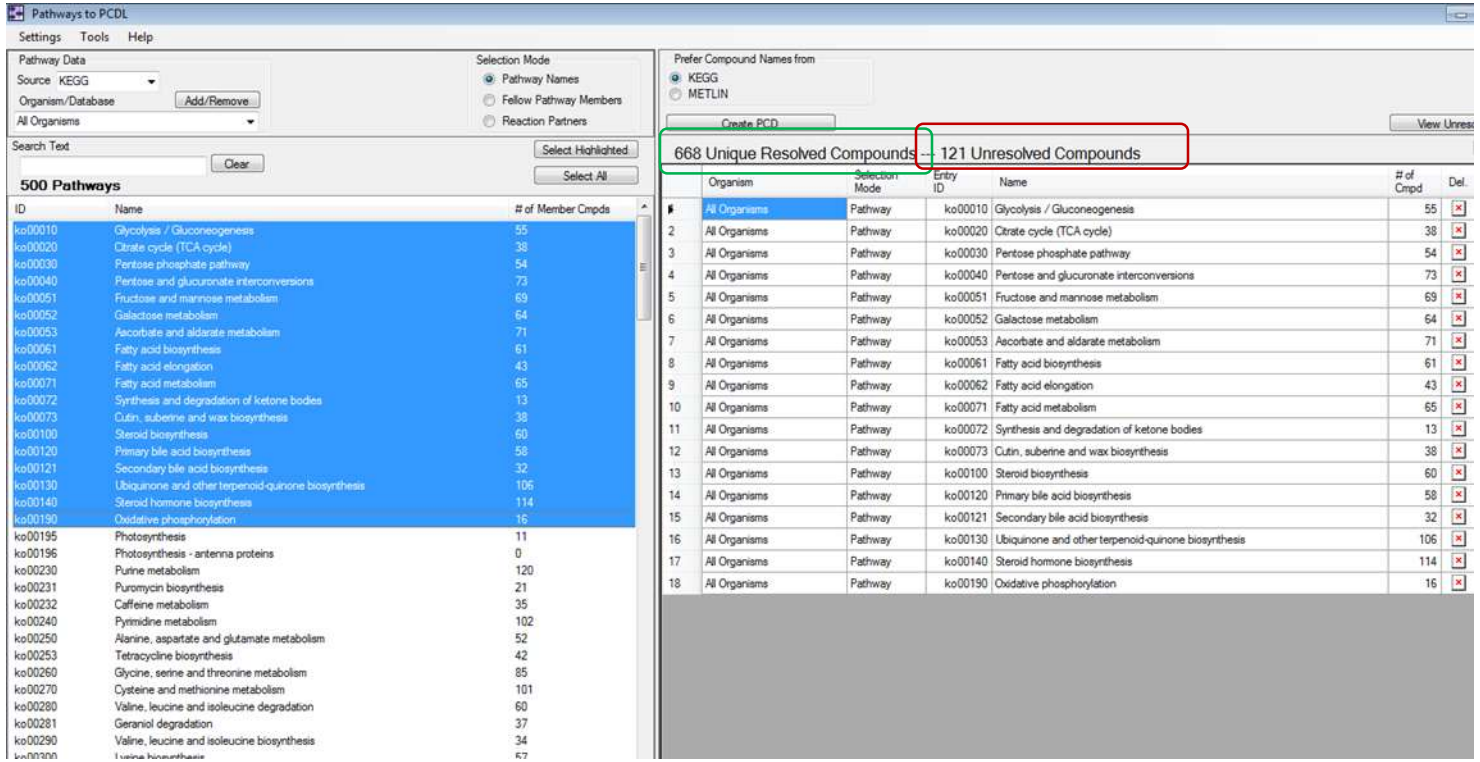
Propose new experiments based on pathway analysis

- Re-examine acquired untargeted metabolomics data based on pathway analysis
- Design new experiments (metabolite, protein or genes) based on pathway results interpretation

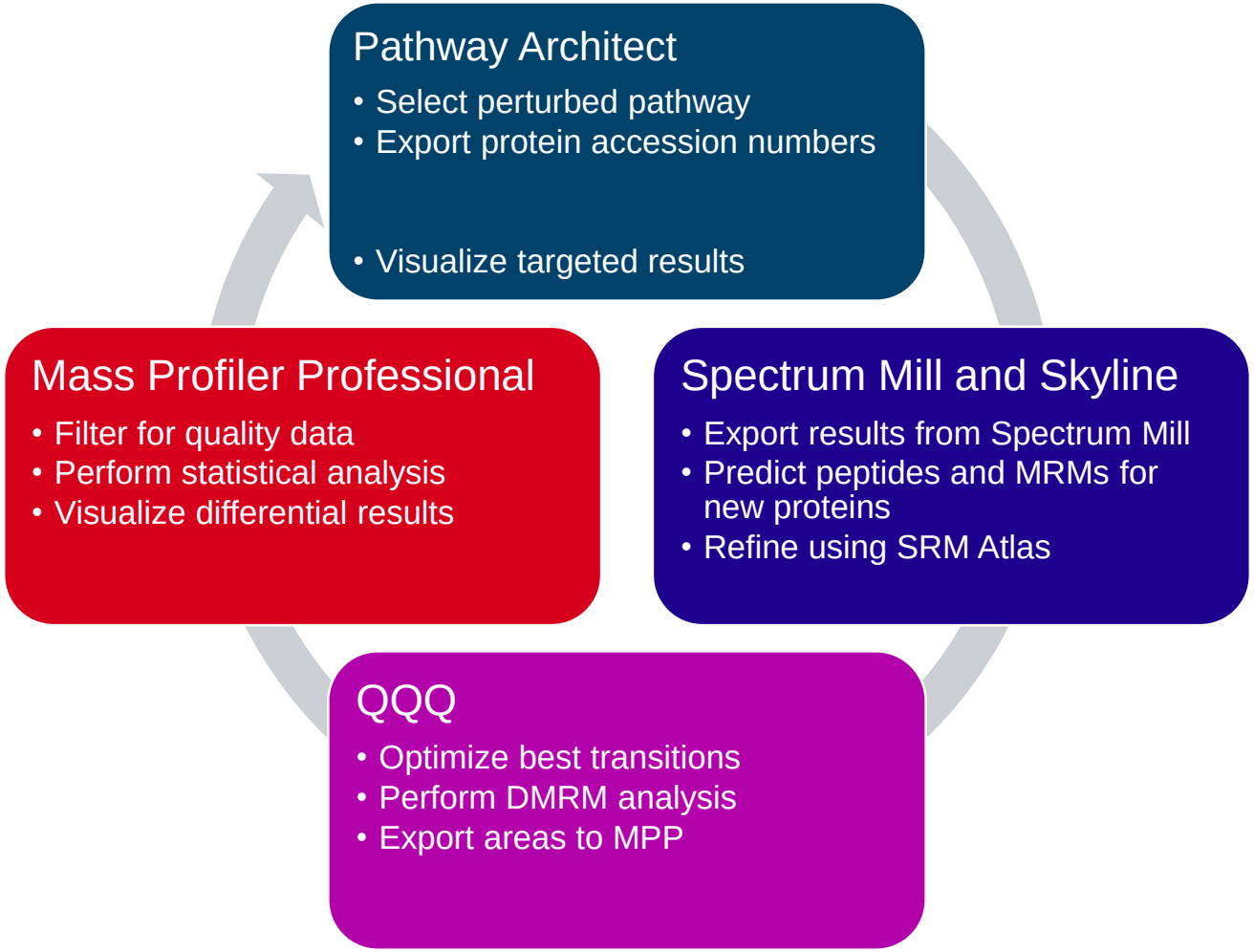


Pathways to PCDL: Create custom databases

Select an Organism → Select Pathway(s) → Create new subset KEGG .cdb



Targeted Proteomics: Using Pathway-Directed Information to Inform the Next Experiment



Targeted Proteomics: Agilent 6495 QQQ with iFunnel Technologies

Outstanding sensitivity with iFunnel

Excellent standard flow performance with AJS + 1290

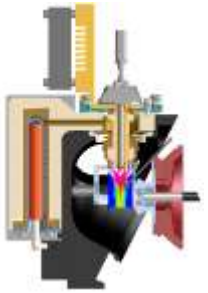
Routine, robust nano LC with HPLC-Chip/QQQ

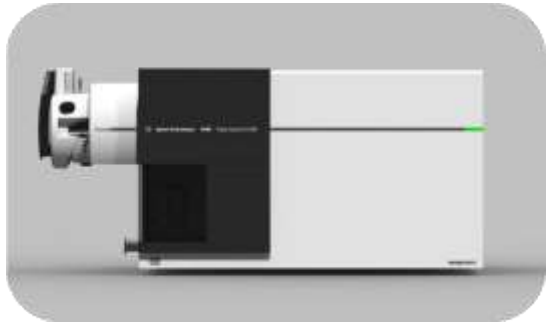
tMRM mode

Skyline workflow manuals

New Skyline Automation tool

Skyline workflow manuals

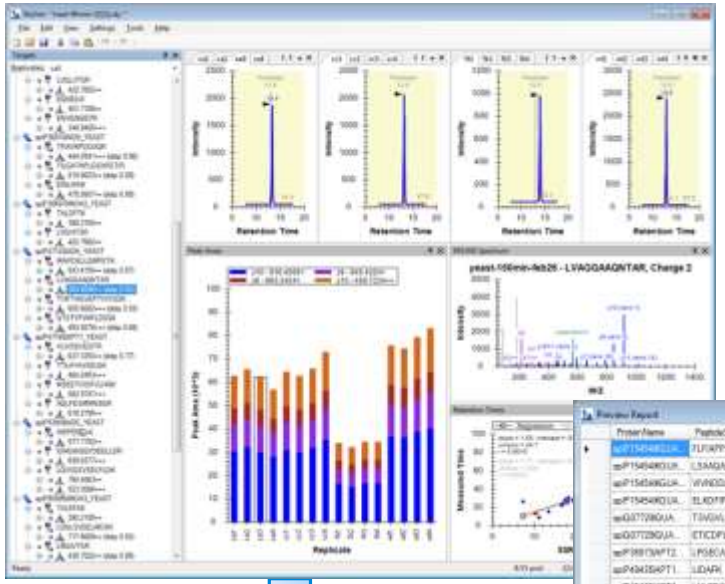




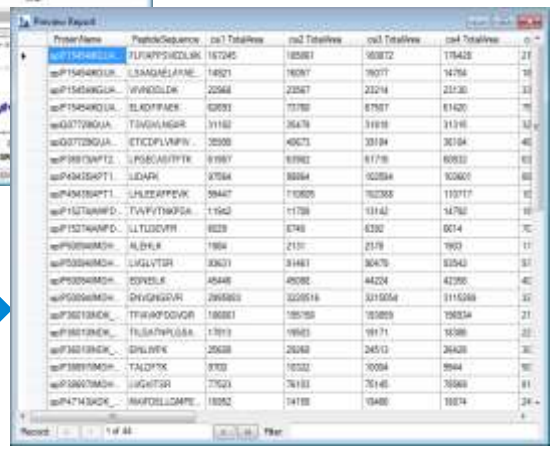
Skyline – MPP Data Exchange

Importing Targeted Results

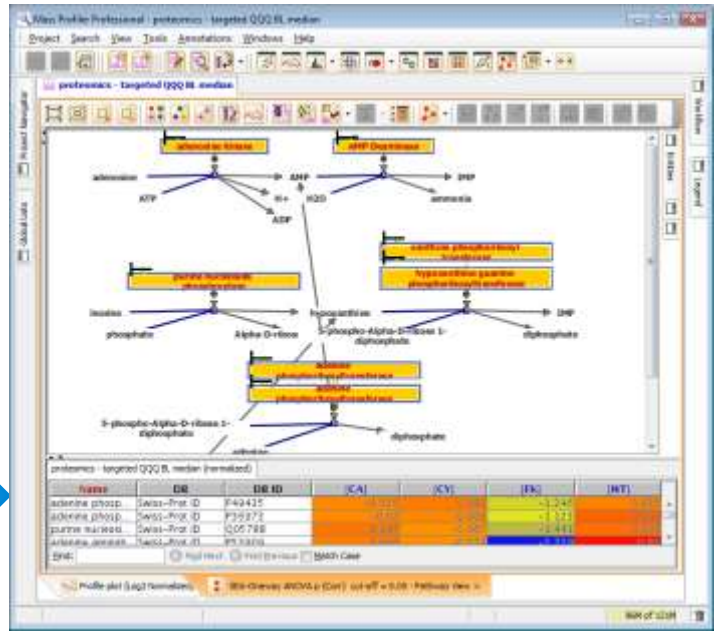
Exporting Protein Areas From Skyline to MPP



Review and process QQQ results in Skyline

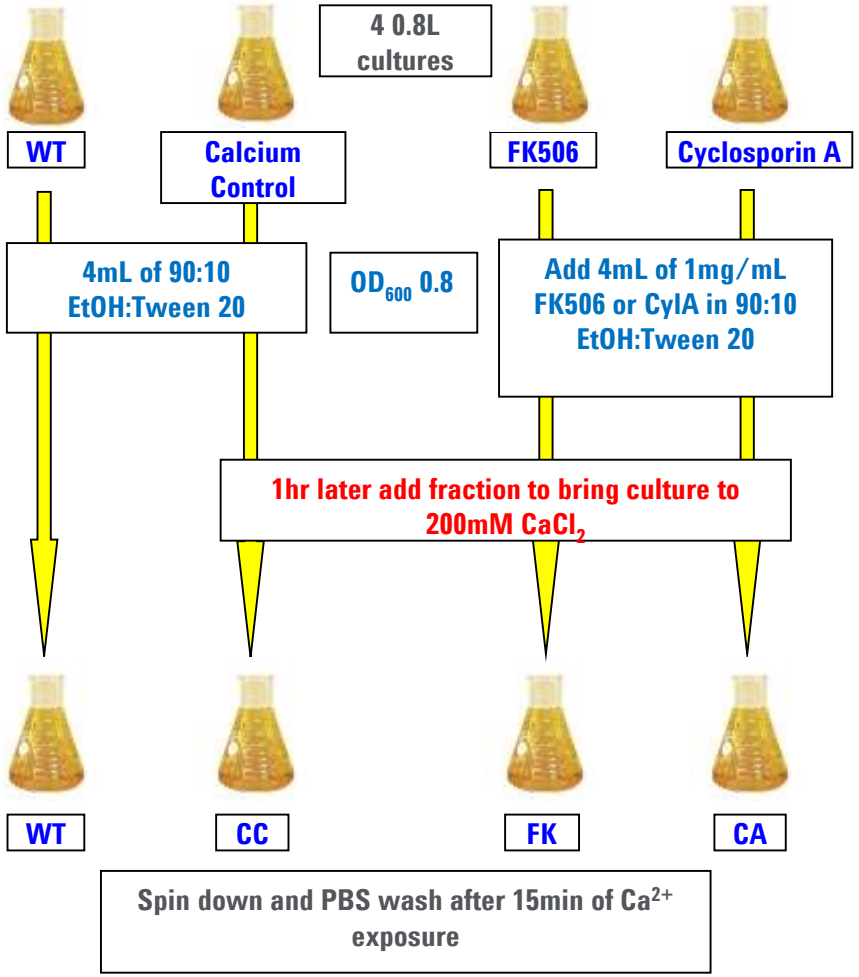


Export results to MPP



Pathway visualization in Pathway Architect

Yeast Metabolomics :



Experimental Design

Treatments:

- Wild type (WT) - no treatment
- Calcium control (CC) - $CaCl_2$
- FK - FK506 and $CaCl_2$
- CA - Cyclosporin A and $CaCl_2$

Extraction :

Wet mill with 5:3:3 $CHCl_3$:CH_3OH:H_2O. Only the aqueous is analyzed

Metabolomics Workflow

Acquire Data

- Analyze metabolomics samples
- Use accurate mass LC/MS to analyze samples

Analyze Data

- Mine data using MassHunter
- Analyze in Mass Profiler Professional

Identify Compounds

- Identify metabolites using ID Browser
- Search METLIN to annotate metabolites

Pathway Analysis

- Map results to pathways using Pathway Architect

Metabolomics Workflow

Acquire Data

- Analyze metabolomics samples
- Use accurate mass LC/MS to analyze samples

Analyze Data

- Mine data using MassHunter
- Analyze in Mass Profiler Professional

Identify Compounds

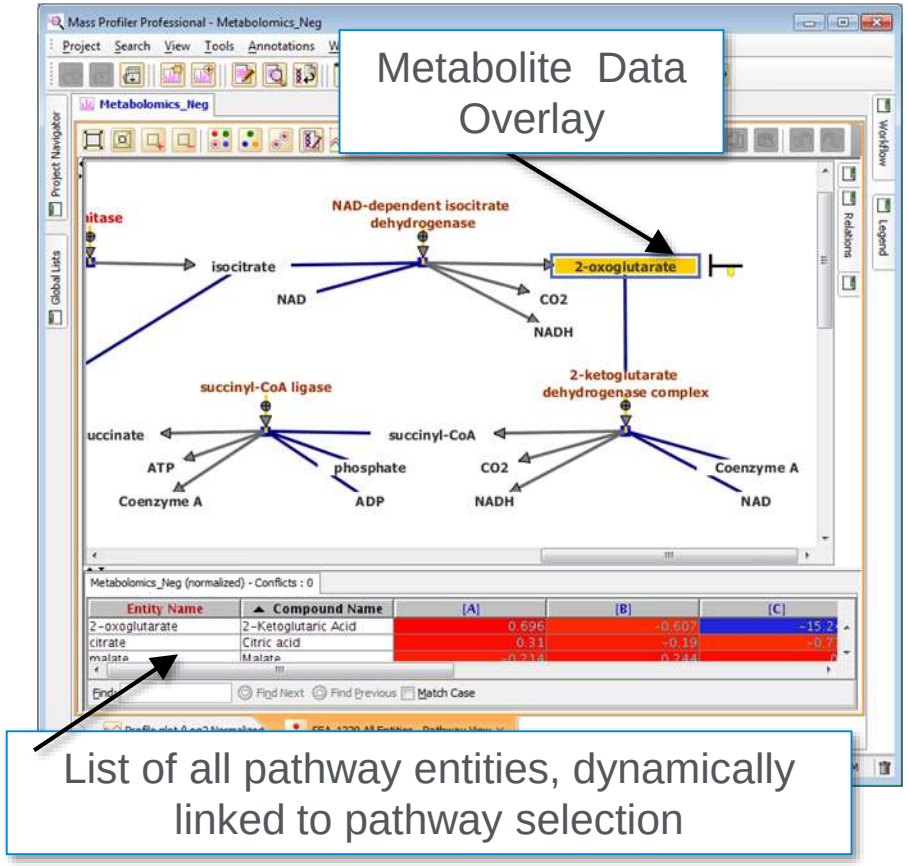
- Identify metabolites using ID Browser
- Search METLIN to annotate metabolites

Pathway Analysis

- Map results to pathways using Pathway Architect
- Export protein list for next experiment

Pathway Architect

Pathway Architect is an optional module in MassProfiler Professional



Map and visualize data from one or two types of -omic data on pathways

Search, browse and filter pathways

Supports biological pathways from publicly available databases

- WikiPathways
- BioCyc
- Supported pathway formats
  - BioPAX 3 – Pathway Commons, Reactome, NCI Nature Pathway
  - GPML – PathVisio –custom drawing
- Export compound list from pathways