METABOLÓMICA

Coral Barbas
Danuta Dudzik
Mª Fernanda Rey-Stolle
Francisco J. Rupérez
Antonia García

ÍNDICE

1. Introducción a la metabolómica.

2. Enfoques analíticos en la metabolómica.
   - Flujo de trabajo del estudio metabolómico
   - Procedimiento de Control de Calidad y Garantía de Calidad en la metabolómica

3. Procesamiento de datos e identificación de metabolitos
   - Flujo del procesamiento de datos
   - Tratamiento de datos en la metabolómica no dirigida
   - Identificación de metabolitos
   - Análisis estadístico

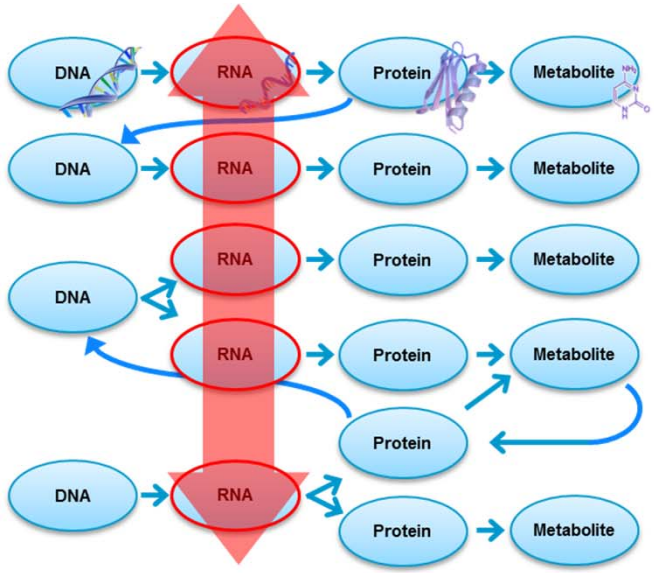
4. Análisis de datos
   - De la identificación de datos a las rutas metabolómicas
   - Validación de biomarcadores

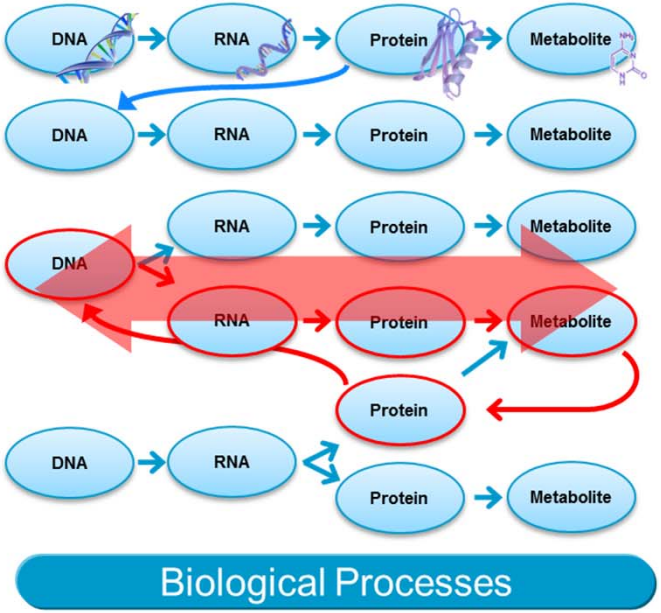
5. Sesiones prácticas
   - Metabolómica dirigida y no dirigida
   - Metabolómica con herramientas gratuitas en línea

Metabolómica

La investigación «ómica», nuevo campo emergente (que incluye la genómica, proteómica y metabolómica), se ocupa de la caracterización completa de los pequeños metabolitos moleculares presentes en sistemas biológicos.

Ómica y Biología de Sistemas





Anexo: Definición de Metabonómica

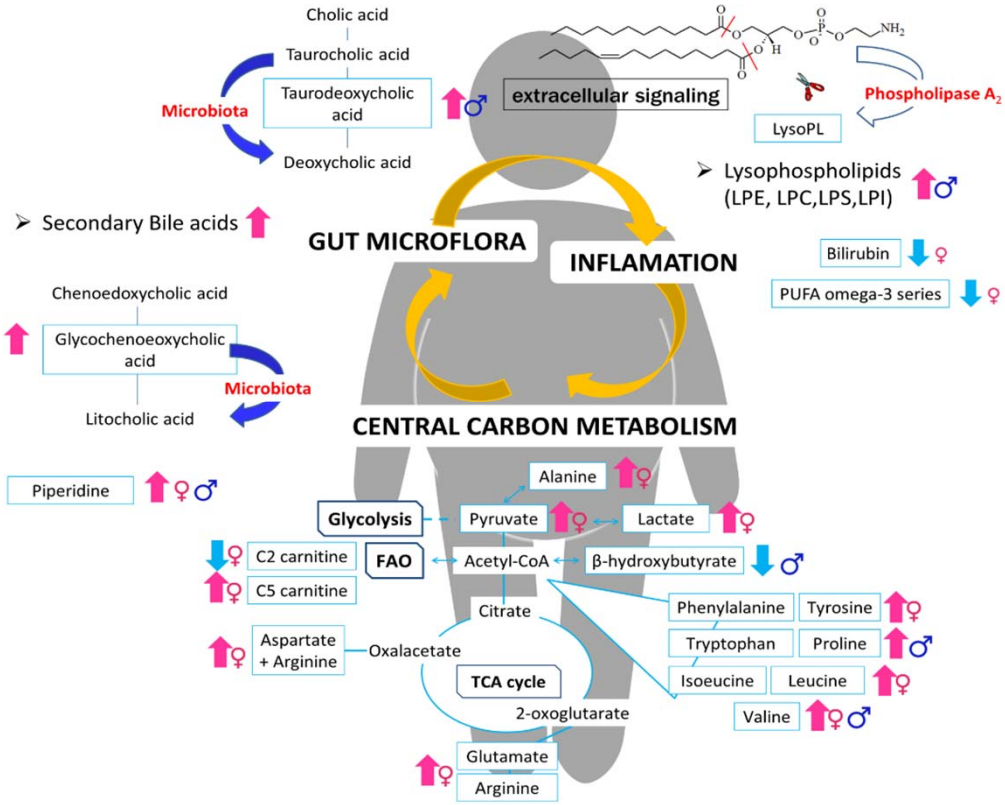
- Medida de la respuesta metabólica dinámica y multiparamétrica de un sistema vivo a estímulos patofisiológicos o modificaciones genéticas (Nicholson, 1999)
  - medida cuantitativa de la respuesta metabólica «total» en relación con el tiempo a estímulos patofisiológicos (nutricionales, xenobióticos, quirúrgicos o tóxicos)
- MetaboLómica - el cuadro / MetaboNómica - la película
- Hoy en día, todo es Metabolómica

Definición de Metaboloma

- «… el **conjunto completo de metabolitos**/sustancias intermedias con bajo peso molecular, que dependen de un contexto y que varían en función de la fisiología, el estado de desarrollo o patológico de la célula, tejido, órgano u organismo...» (Oliver 2002)
- Origen: Endometaboloma, Microbioma, Xenobioma, Nutribioma…
- Naturaleza: Glicoma, lipidoma, esfingolipidoma, peptidoma…
- Metaboloma ↔ Fenotipo

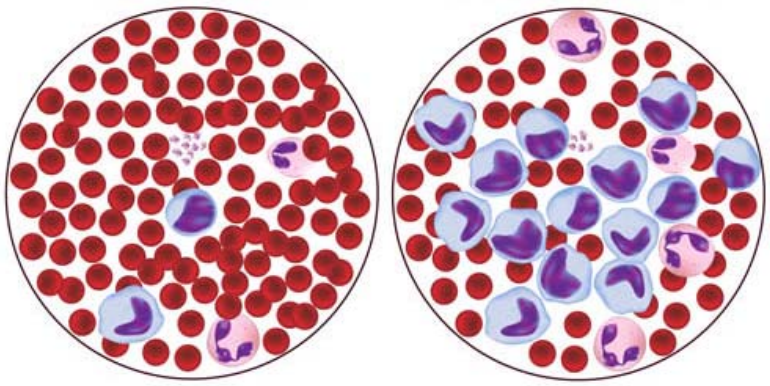
Lo que puede aportar la metabolómica (I)

- Perspectiva global del estatus metabólico y de los eventos bioquímicos globales asociados a un sistema celular o biológico.
  - Situaciones patológicas sin mecanismo conocido, como la relación entre la obesidad y la resistencia a la insulina



Lo que puede aportar la metabolómica (II)

- Identificación (propuesta) de nuevos biomarcadores, importantes en el proceso de descubrimiento de nuevos fármacos o como herramientas de diagnóstico in vitro.
  - Por ejemplo, nuevos biomarcadores diagnósticos para la agresividad en la leucemia linfática crónica



Utility of validated metabolites as biomarkers of aggressive state of CLL

Metabolite	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Acetylcarnitine	0.695	43.2	93.0	86.1	62.1
Butyrylcarnitine	0.548	10.8	98.0	84.4	52.4
Hexanoylcarnitine	0.690	27.0	96.0	87.1	56.8
Octanoylcarnitine	0.651	29.7	95.0	85.6	57.5
Decanoylcarnitine	0.662	27.0	94.0	81.8	56.3
Palmitoylcarnitine	0.719	40.5	94.0	87.1	61.2
Dodecanamide	0.497	8.1	100.0	100.0	52.1
Hexadecanamide	0.516	5.4	100.0	100.0	51.4
Oleamide	0.600	18.9	96.0	82.5	54.2
Linoleamide	0.672	16.2	98.0	89.0	53.9
Acylcarnitines[a]	0.743	32.4	95.0	86.6	58.4
FAA[b]	0.662	13.9	96.0	77.6	52.7
Acylcarnitines and FAA	0.750	54.0	89.0	83.1	65.9

AACLifeSci

CEU Universidad San Pablo

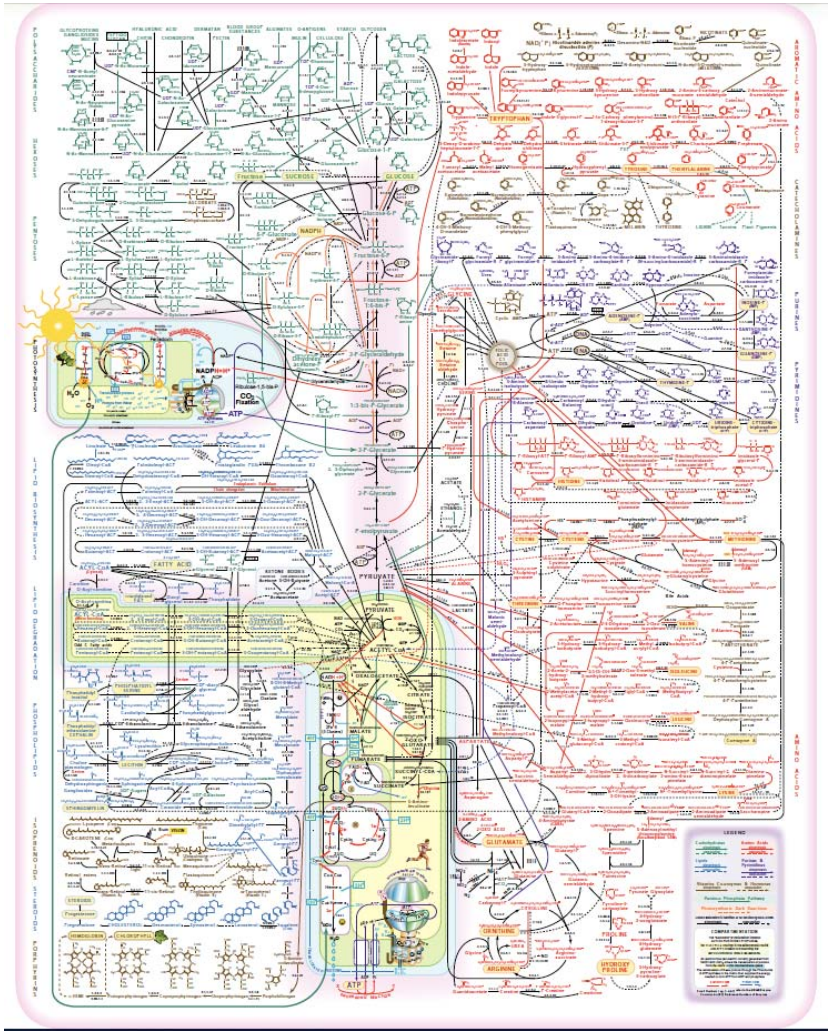
La metabolómica sirve para...

- buscar diferencias metabólicas entre grupos de muestras (caso *vs*. control; pre-tratamiento *vs*. post-tratamiento; una condición *vs*. otra)
- identificar compuestos que sean significativos y proponer los mecanismos
- averiguar información sobre el fenotipo
- observar los efectos de un tratamiento
- encontrar nuevos objetivos farmacológicos

La metabolómica NO es...

- un método para revelar el destino de un metabolito o un fármaco
- un método de cuantificación
- el uso de un simple kit para cuantificar un grupo de metabolitos (se requiere NMR, MS...)
- posible sin una comparación simultánea de muestras

Definición de Metabolismo



El grupo completo de procesos (bio)químicos dentro de un orgánulo. célula, tejido, órgano u organismo, esenciales para la vida

Enfoques analíticos en la metabolómica

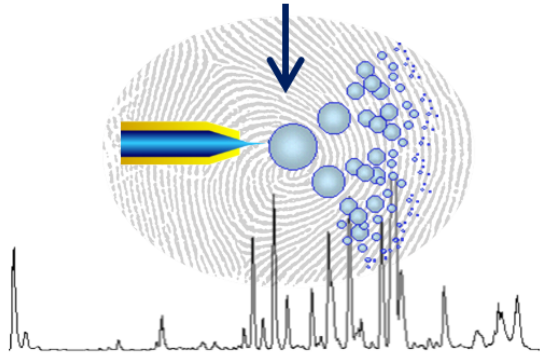
AACLifeSci

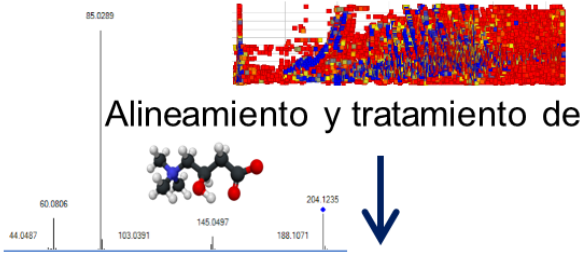
CEU Universidad San Pablo

Estrategias metabolómicas

No dirigida

Muestras



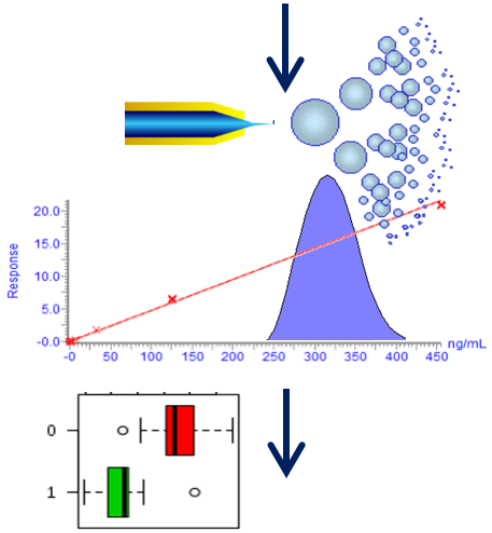


Alineamiento y tratamiento de datos

Confirmación con MS/MS a partir de los estándares

Dirigida

Muestras y Estándares de metabolitos



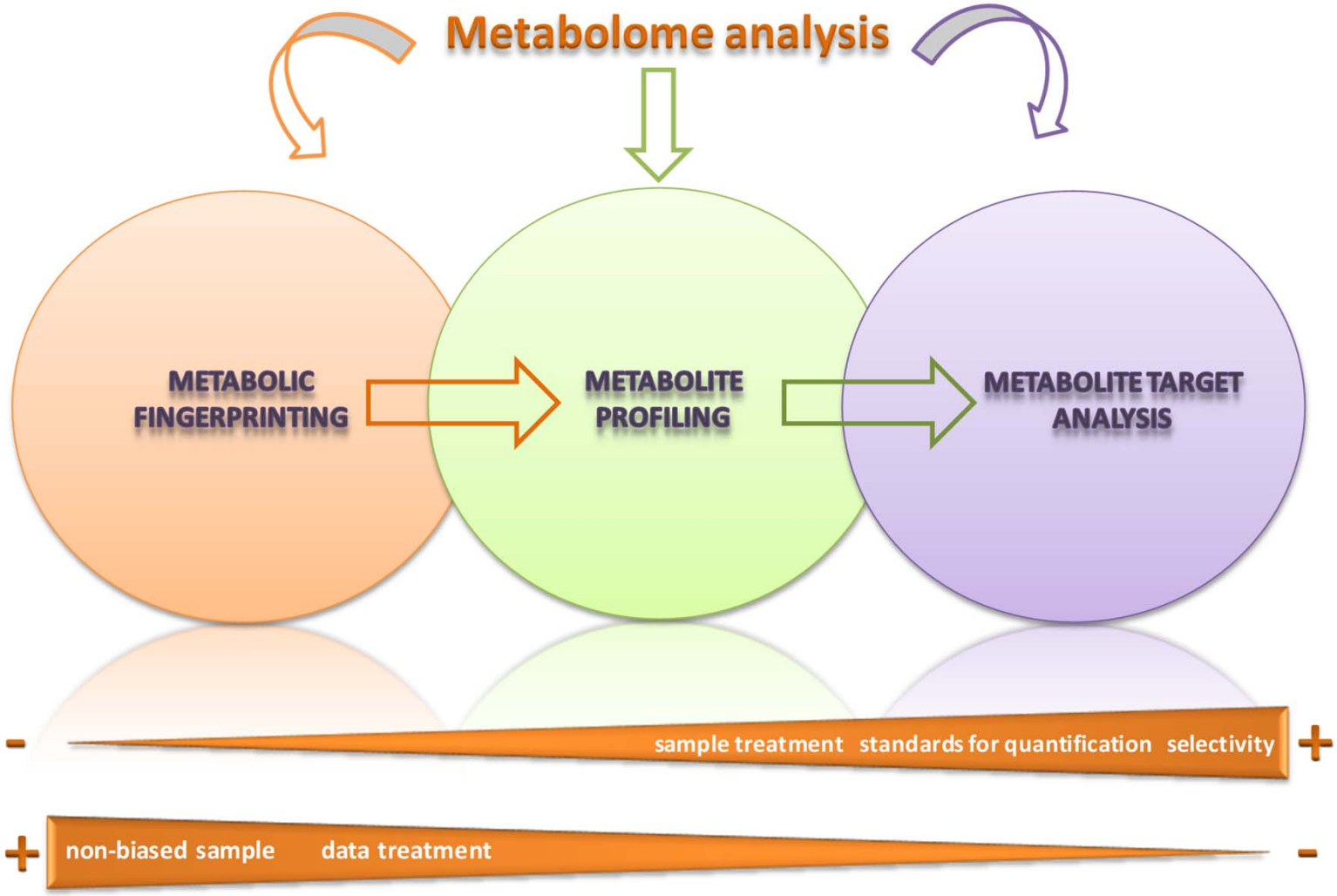
Análisis de datos/Cuantificación

Comparación de muestra y estándares

Perfil metabólico global

Cuantificación de metabolitos específicos

Tres maneras de hacer metabolómica



FLUJO DE TRABAJO EN METABOLÓMICA

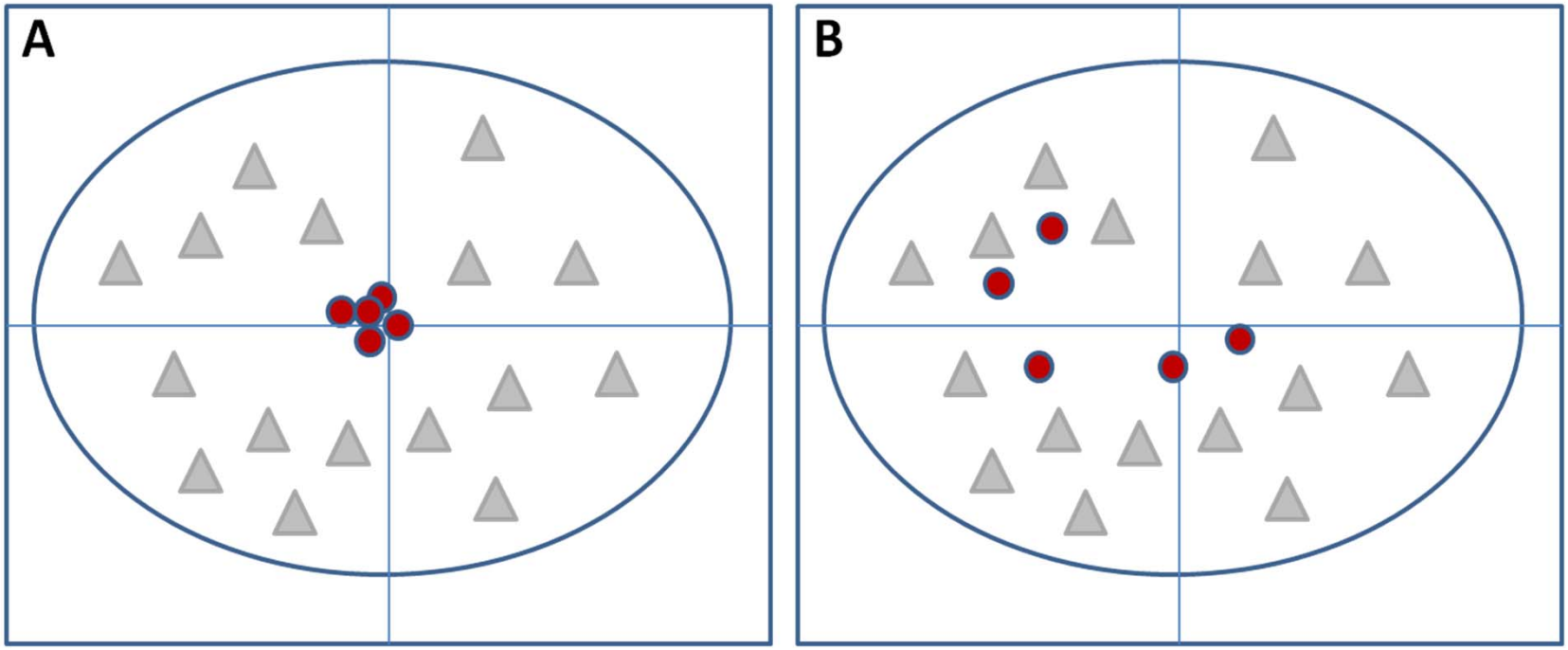
TÉCNICAS ANALÍTICAS

- GC/MS: Compuestos polares pequeños
  - Principalmente solubles en agua (algunos hidrofóbicos)
  - Tratamiento de muestras: Derivatización
  - Fragmentación reproducible - bases de datos
- NMR
  - Soluble en agua
  - Prácticamente sin tratamiento de muestra
  - LOD alto
- LC/MS
  - metabolitos de tamaño pequeño a grande (<1500 Da), y de polaridad media a nula
- CE/MS: Compuestos polares pequeños-medianos
  - Aminoácidos, acilcarnitinas, poliaminas, etc.
  - Sin derivatización

Plataformas analíticas basadas en la MS en la metabolómica

Técnica Analítica	Aplicación	Ventajas	Desventajas
GC-MS	Separación, identificación, y cuantificación de metabolites poco polares, volátiles y térmicamente estables	Elevada resolución, posibilidad de librerias de metabolites de amplio espectro para identificación.	Imporibilidad de analizar metabolitos termoestables. Se require la derivatización de metabolitos no volátiles de elevado peso molecular
LC-MS	Separación, identificación, y cuantificación de numerosos grupos de metabolitos, dependiendo del tipo de columna y fase móvil	Alta sensibilidad, alta capacidad de carga de muestra, no necesaria derivatización, posibilidad de analizar compuestos termolábiles.	Disponibilidad limitada de librerías comerciales, restricción de eluyentes en LC, efecto matriz, limitado potencial de identificación a menos que se utilice como detección MS-MS
CE-MS	Separación, identificación, y cuantificación de metabolites polares e ionizados, volúmenes de muestra muy pequeño.	Elevada resolución y análisis rápido, útil para muestras biológicas complejas, incluso con muy poco volumen	Disponibilidad limitada de librerías comerciales. Incompatibilidad de BGE, límites de detección. Limitado potencial de identificación a menos que se utilice como detección MS-MS

Procedimiento de Control de Calidad y Garantía de Calidad en la metabolómica



A: QC (puntos rojos) agrupadas

B: QC dispersas

TRATAMIENTO DE DATOS EN METABOLÓMICA:
Procesamiento de señales

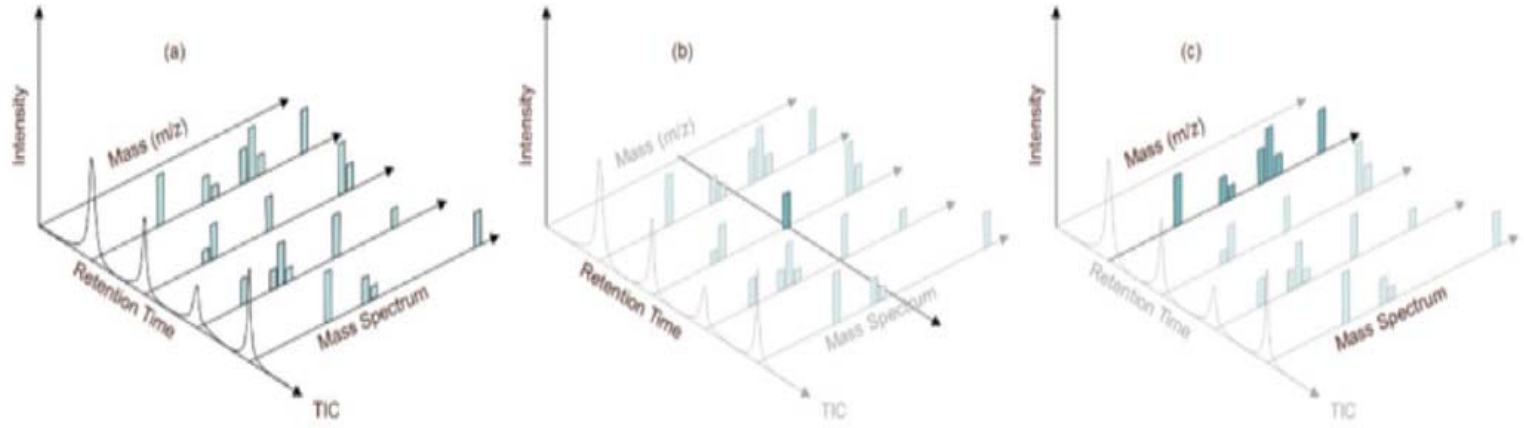
AACLifeSci
CEU Universidad San Pablo

Procesamiento de datos brutos (primarios)	Preprocesamiento de datos	Pretratamiento de datos	Tratamiento de datos
• Conversión de archivos de datos brutos	• Reducción del ruido • Detección de picos • Alineamiento • Filtrado de datos • Valores perdidos (en blanco)	• Normalización • Transformación • Escalado	• Identificación de metabolitos • Análisis estadístico: Multivariante & Univariante

FLUJO DEL PROCESAMIENTO DE DATOS

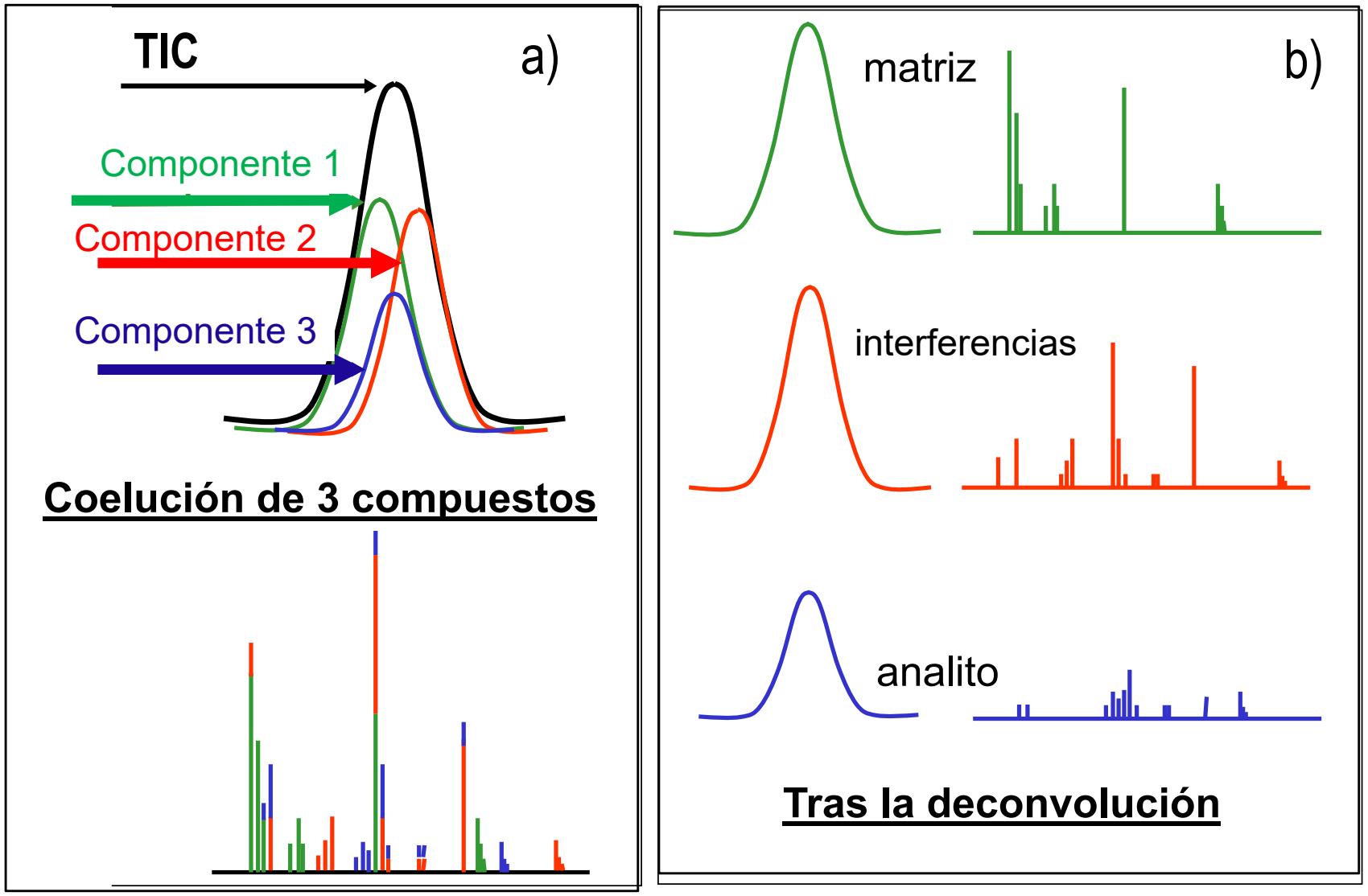
TÉCNICA ANALÍTICA: GC-MS

- Cromatografía de gases acoplada a la espectrometría de masas
- Criterio de referencia
  - Altamente sensible y reproducible
  - Información: Calidad y Cantidad
  - Librerías de espectros para propósitos de identificación
  - El 10-20% de los compuestos conocidos se pueden analizar mediante GC
    - Alta relevancia metabólica



(a) Datos 3D de GC/MS; (b) Cromatograma de ion extraído para el ion seleccionado
(c) Un único punto de datos en el tiempo proporciona un único espectro de masas
extraído de Chromatography Today

Deconvolución



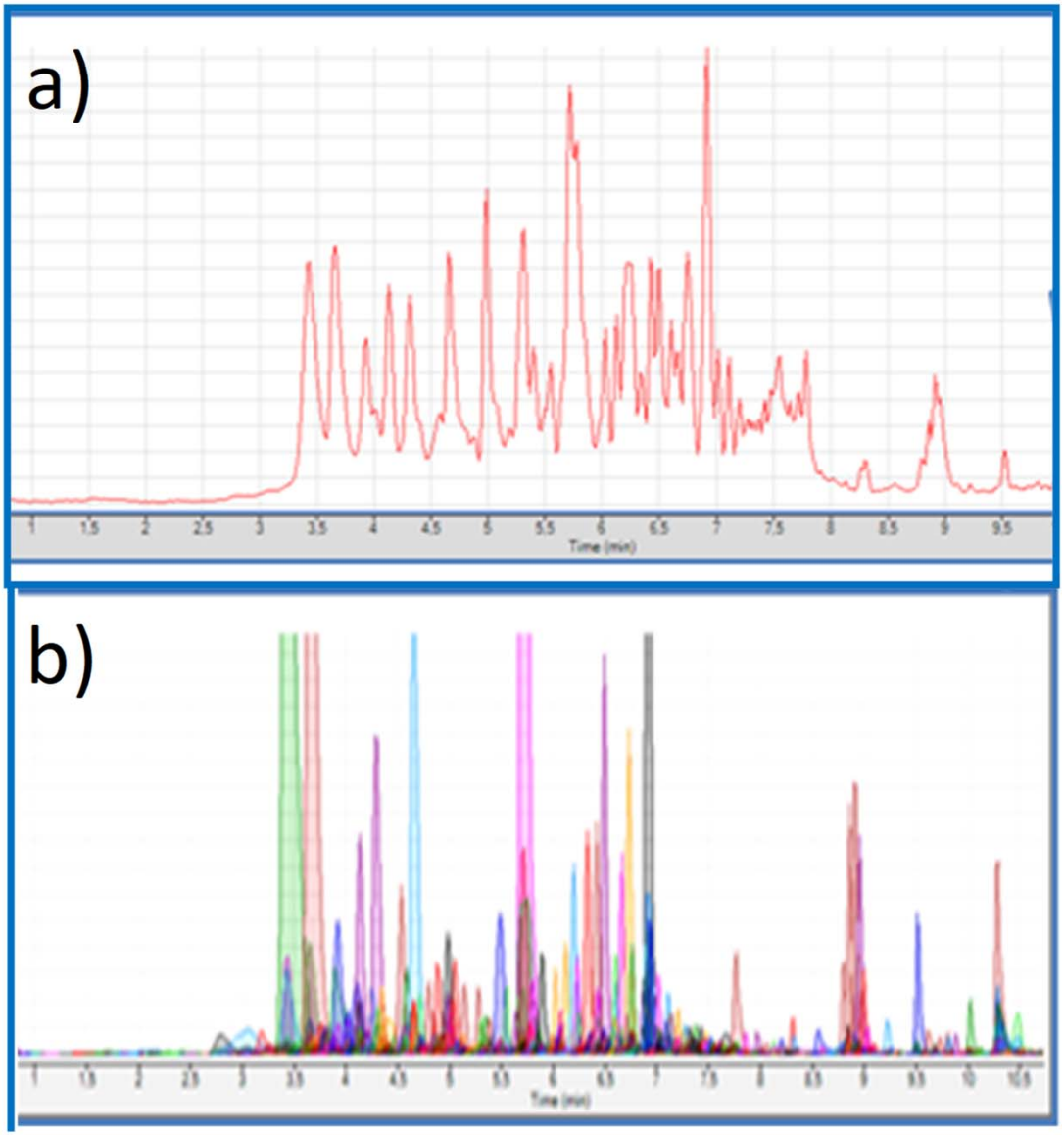
a) Antes y b) Después del proceso de deconvolución

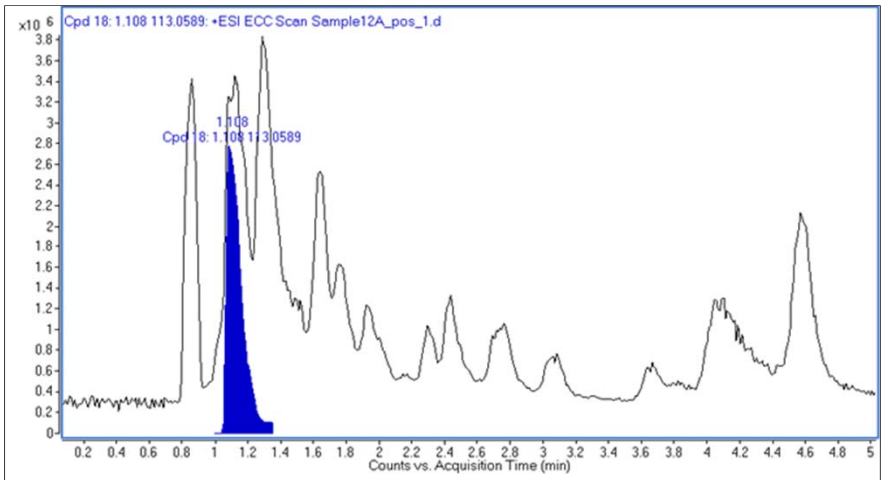
extraído de https://www.agilent.com/cs/library/Support/Documents/f05017.pdf

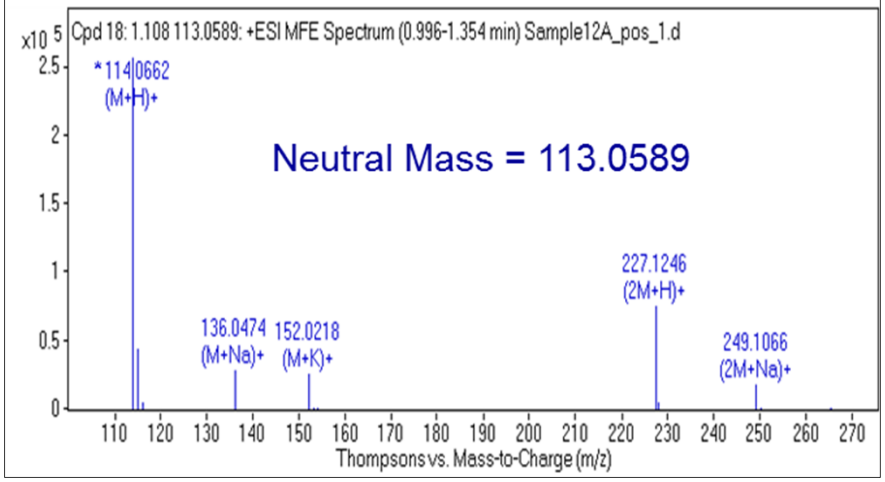
Deconvolución en LC-ESI-MS y CE-ESI-MS

- Métodos basados en picos
- A la hora de determinar si los distintos iones son del mismo “posible compuesto” (*feature*) el *Molecular Feature Extractor* (Agilent) considera la exactitud de las medidas de masas para iones relacionados en un grupo por su estado de carga, distribución isotópica, y las posibles relaciones químicas.
- También puede considerar iones vinculados como aductos: aductos de protón, sodio, potasio y amonio en la ionización positiva o pérdida de un protón, aductos con formiato, etc., en el modo de ionización negativa.

Después de la deconvolución

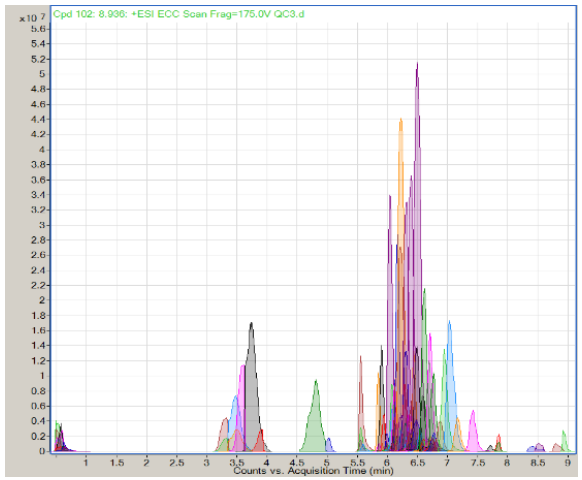






a) Cromatograma de Iones Totales
b) Cromatogramas de cada uno de los compuestos obtenidos después de la deconvolución

¿Cromatograma o lista de características?



Automatically Show Columns

Show/Hide	Saturated	RT	m/z	Mass	Polarity	Ions	Height	Area	Vol	Quality Score
☑	S	0.514	280.0923	279.085	Positive	5	2740215		14234373	100
☑	S	0.517	203.0526	202.0454	Positive	6	4045612		20286412	80
☑		0.526	140.0682	139.0609	Positive	3	1530642		7519633	100
☑		0.529	136.0482	135.041	Positive	2	1187260		6016909	100
☑	S	0.57	162.1126	161.1053	Positive	7	2758926		19465836	100
☑	S	0.57	304.2998	303.2926	Positive	4	3021606		14360408	100
☑		0.58	114.0664	113.0591	Positive	4	549599		6985240	80
☑		0.614	175.1192	174.1119	Positive	3	760396		3860452	91.6
☑		0.625	156.0768	155.0696	Positive	7	1055485		6005360	100
☑		0.646	170.0927	169.0854	Positive	4	604901		3355886	100
☑	S	3.29	520.3389	519.3316	Positive	7	3243984		36560884	80.2
☑		3.298	544.3388	543.3315	Positive	4	1339590		12508261	87
☑	S	3.483	520.3389	519.3316	Positive	11	5382675		86088232	85.2
☑	S	3.5	544.3389	543.3316	Positive	7	2216628		36501788	100
☑	S	3.573	496.3389	495.3316	Positive	6	8281163		80624664	86.1
☑	S	3.746	496.3388	495.3316	Positive	13	11742118		200075408	80
☑	S	3.89	522.3545	521.3472	Positive	8	2204035		17763280	87
☑	S	4.801	524.3702	523.3629	Positive	12	6862920		111854136	100
☑		5.027	524.3702	523.3629	Positive	6	1308318		8748089	87
☑		5.559	163.0393	162.032	Positive	4	1349239		5379156	100
☑	S	5.561	391.2839	390.2767	Positive	15	7367381		51796644	100
☑	S	5.561	149.0235	148.0162	Positive	4	2915086		12886771	100
☑		5.597	338.3418	337.3345	Positive	9	787867		4058574	100
☑	S	5.849	675.5425	674.5353	Positive	7	6821314		39496860	100
☑		5.857	627.5339	626.5266	Positive	3	712763		3673685	100
☑	S	5.904	701.5583	700.551	Positive	8	8615328		59468044	100
☑	S	5.941	689.558	688.5507	Positive	8	3178377		18867332	100
☑		5.991	754.5372	753.5299	Positive	4	737180		4127601	100
☑		5.995	730.5369	729.5297	Positive	7	1506513		8484420	100

Chromatogram Results | MS Spectrum

Preprocesamiento de datos

- Alineamiento
  - Los desplazamientos de los picos se observan a través del eje del RT
  - Dos grupos:
    - se alinean los datos antes de la detección de picos
    - métodos de alineamiento basados en el pico: se alinean en todas las muestras los picos espectrales detectados
    - programas:
      - MetaboAnalyst (metaboanalyst.ca)
      - mzmine and mzmine2 (http://mzmine.sourceforge.net/)
      - metAlign
      - BinBase (fiehnlab.ucdavis.edu)
      - xcms and xcms2 (Scripps)
      - metaXCMS (Scripps)
      - XCMS Online (Scripps)
- Valores en blanco
  - Problemas en nuevos análisis
  - Distintas estrategias
    - Sustituir por la mitad del mínimo, por media/mediana, el k vecino-más-cercano (KNN), el PCA probabilístico (PPCA), el método PCA bayesiano (BPCA), o la Descomposición en Valores Singulares (SVD), etc.
- Filtración
  - Variables de tamaños muy pequeños - detectados usando la media o la mediana
  - Variables que son prácticamente constantes - detectadas usando la desviación estándar (SD)
  - Variables que muestran baja repetibilidad - medidas usando la muestra QC

Pretratamiento de datos

- Normalización
  - Normalización **específica de muestra** (i.e. peso, volumen)
  - Normalización por **suma** o **mediana**
  - Normalización por **muestra** de referencia
  - Normalización mediante muestra **combinada** a partir de un grupo de control
  - Normalización por **característica** de referencia
  - Normalización por **cuantiles**
- Transformación de datos
  - Transformación **logarítmica**
  - Transformación de **raíz cúbica**
- Escalado de datos
  - Centrado a la **media**
  - **Autoescalado** (centrado a la media y dividido por la desviación estándar de cada variable)
  - **Escalado Pareto** (centrado a la media y dividido por la raíz cuadrada de la desviación estándar de cada variable)
  - **Escalado por rango** (centrado a la media y dividido por el rango de cada variable)

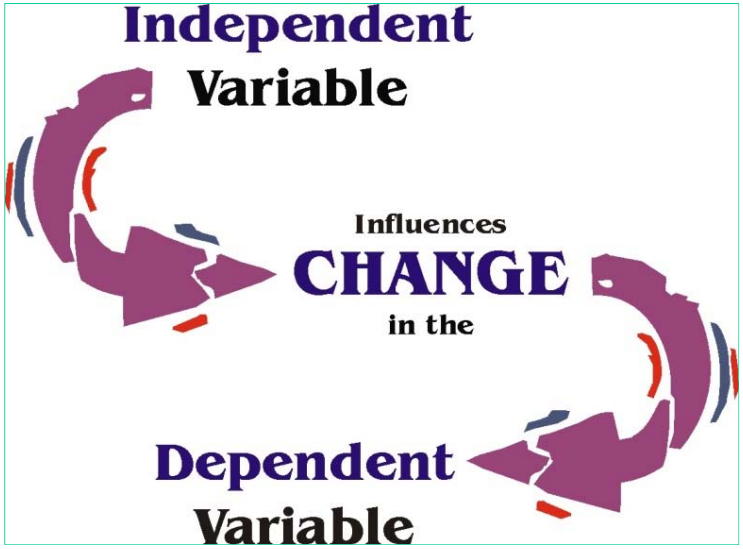
Estadística para Metabolómica

Objetivos:

- detectar diferencias entre grupos de muestras a nivel químico
- categorizar compuestos de acuerdo a su importancia relativa para la diferenciación de muestras

VARIABLES

- **variable dependiente**: representa el resultado o el efecto, o se evalúa para comprobar si existe un efecto, p. ej.: abundancia de metabolito
- **variable independiente**: representa las entradas o las causas, o se evalúan para determinar si son las causas, p. ej.: condiciones de tratamiento dentro del experimento

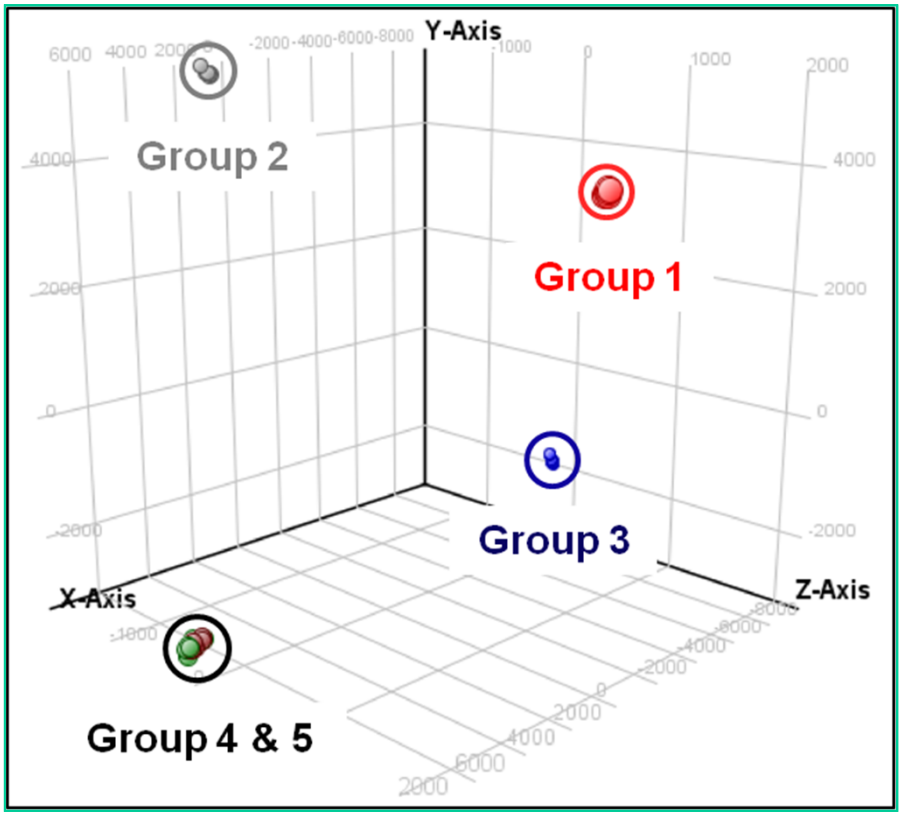


TIPOS

- **Análisis univariante UVA:**
  - Distribución normal: Prueba *t* de Student, ANOVA
  - Distribución no normal: prueba de la U de Mann-Whitney, Kruskal-Wallis
- **Análisis multivariante MVA:** PCA, PLS-DA, OPLSDA

PCA

- utilizado como herramienta en análisis exploratorio de datos
- cada punto representa gráficamente cada muestra medida
- el algoritmo no tiene conocimiento de las asociaciones en grupo de las muestras - análisis *no supervisado*
- el primer compuesto principal explica la mayor parte de la variación
- las cargas de compuestos indican el impacto del compuesto en el análisis
- cada punto es la suma de las cargas de compuestos de una muestra
- la densidad de la agrupación refleja la variación de las muestras



Predicción de clases

un algoritmo que utiliza datos previos para predecir los resultados de observaciones futuras

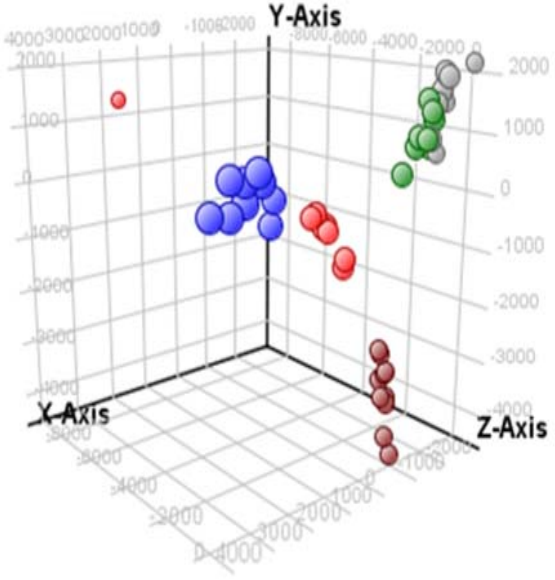
- el algoritmo tiene conocimiento de las asociaciones en grupo de las muestras - análisis *supervisado*
- algoritmos habituales
  - **Análisis discriminante por mínimos cuadrados parciales (PLS-DA)**
  - Máquinas de soporte vectorial
  - Árbol de decisión
  - Bayesiano ingenuo
  - Red neuronal

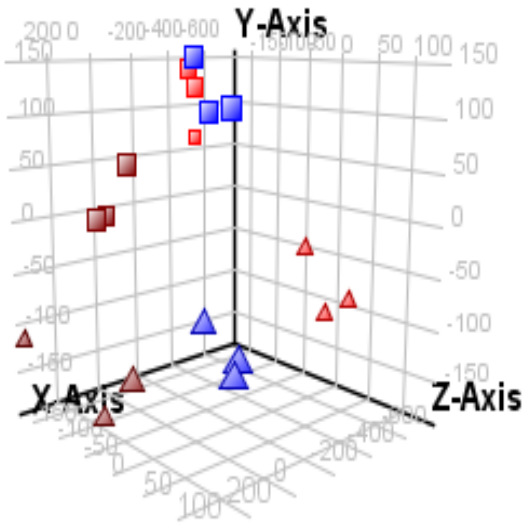
Predicción de clases: PLS-DA

Análisis Discriminante por Mínimos Cuadrados Parciales
Análisis Discriminante por Proyección en Estructuras Latentes

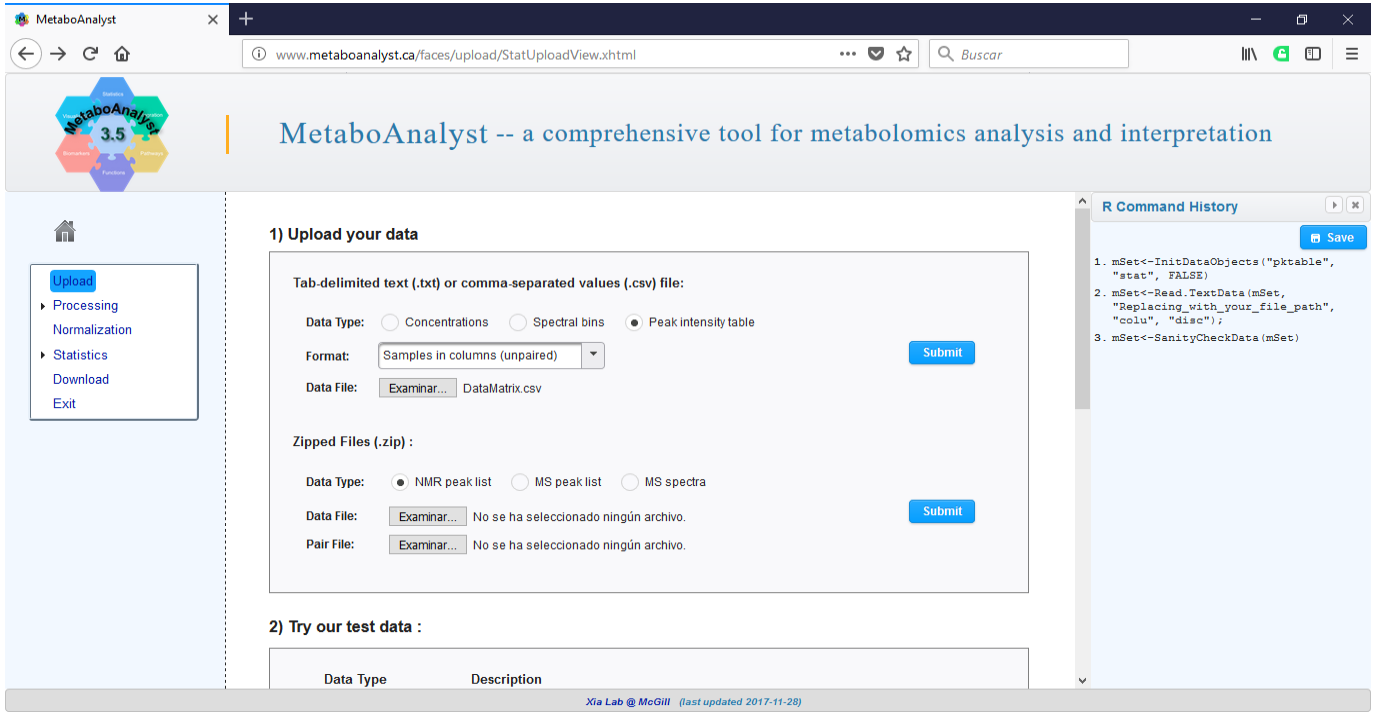
un método estadístico que guarda cierta relación con el análisis de compuestos principales (PCA), pero siendo un análisis *supervisado*

- crea un modelo de regresión lineal proyectando las variables predichas y observables en un nuevo espacio
- muy adecuado para cuando existen más indicadores (compuestos) que observaciones (muestras)
- cada compuesto posee un *t-score* que representa su impacto en la predicción
- se asigna un valor de confianza de la predicción al ejecutar el modelo





Análisis estadístico univariante y multivariante



Predicción de clases: Validación del modelo

se evalúa la exactitud de la regla de predicción que se crea y proporciona una indicación de modelos de sobreajuste:

dejando uno fuera (*leave one out*)

- se utilizan todas las muestras del conjunto de entrenamiento excepto una para construir la regla de predicción
- utilizando esta regla, se predice la clase de muestra que se ha dejado fuera
- la muestra se devuelve al conjunto de entrenamiento y se deja fuera una muestra diferente, volviendo a elaborar la regla de predicción con las muestras que quedan
- se repite este proceso hasta que se haya predicho cada muestra del conjunto de entrenamiento exactamente una vez
- se recuenta entonces el número de predicciones correctas e incorrectas para determinar el porcentaje de éxito

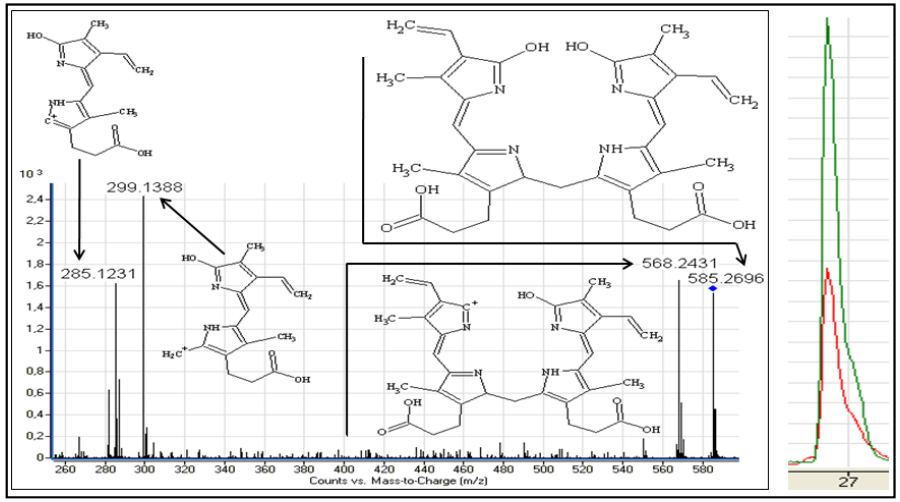
N - fold

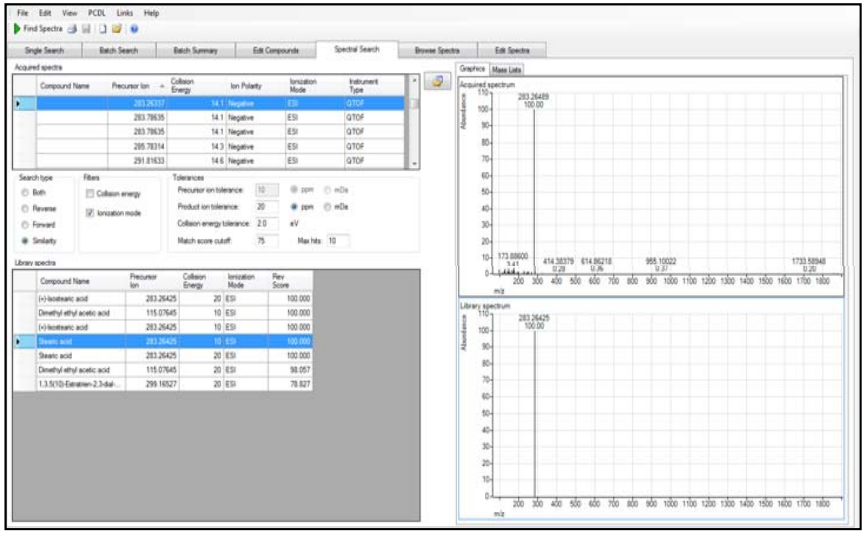
1. las muestras del conjunto de entrenamiento se dividen aleatoriamente en *N* subconjuntos iguales, manteniendo una frecuencia de clases relativa
2. *N*-1 subconjuntos se combinan entonces para el entrenamiento y el conjunto sobrante se utiliza para el ensayo
3. se repite paso a paso dejando un grupo fuera cada vez
4. se repite el paso 1, 2, 3 *M* veces
5. cada muestra se predice *M* veces y se comunica la clase que haya sido predicha más veces a lo largo de esas *M* veces en los resultados de validación

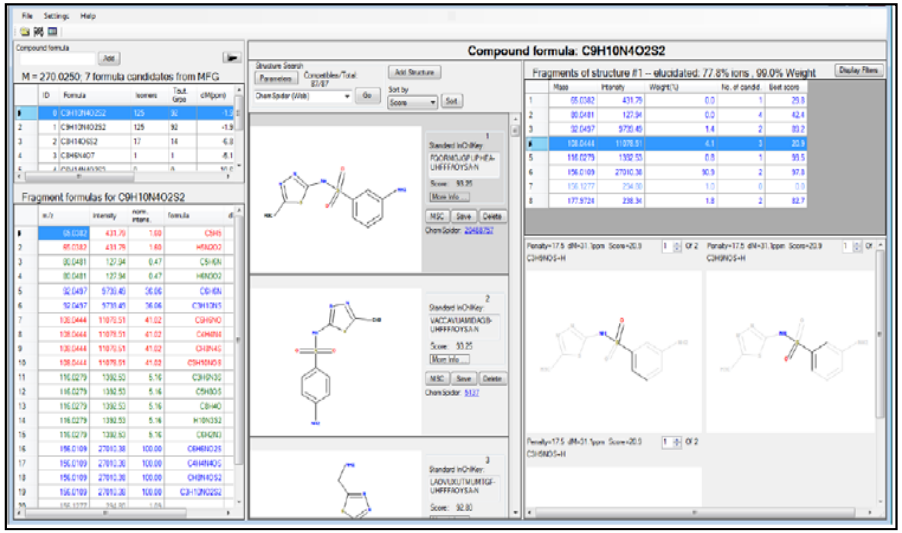
Identificación

1. Búsqueda en base de datos con medida de masa exacta
2. Búsqueda en base de datos comparando el patrón isotópico
3. Búsqueda en base de datos comparando el patrón isotópico y el tiempo de retención
4. Búsqueda en librerías de MS/MS
5. Búsqueda en librerías de MS/MS y tiempo de retención

Confianza







CLASIFICACIONES DE BASES DE DATOS

- **Basadas datos espectrales**
  - Principalmente pequeñas moléculas y no solo metabolitos
  - NMR
  - MS o MS/MS
- **Basadas en información de compuestos**
  - Nombre de compuesto, estructuras, propiedades físicas, identificación
- **Basadas en base de datos de rutas metabólicas**
  - Rutas de metabolitos, xenobióticos, proteínas y de señalización
- **Base de datos metabolómica completa**
  - Una combinación de las anteriores

Lista de Bases de Datos en 2018

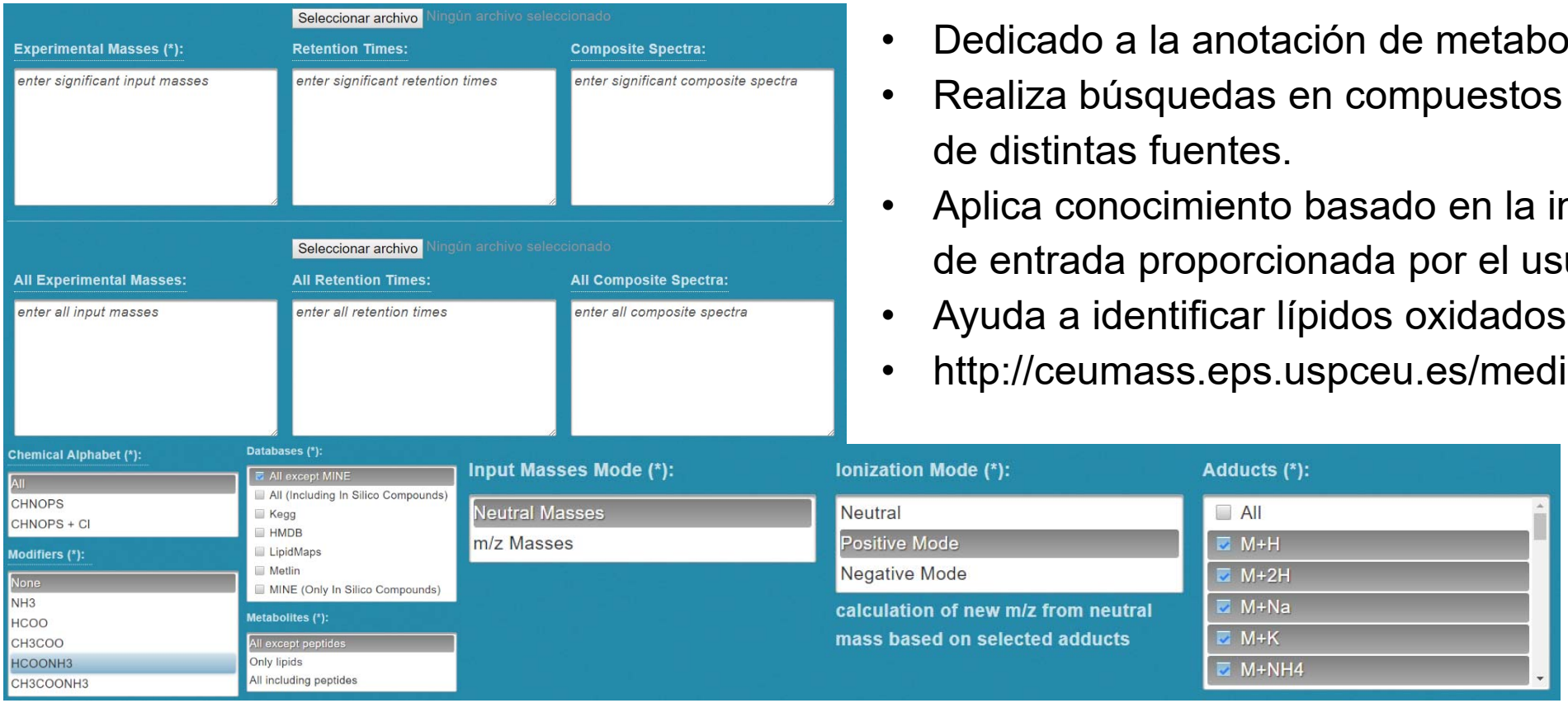
AACLifeSci

CEU Universidad San Pablo

Nombre	URL	Nombre	URL
ARALIP	http://aralip.plantbiology.msu.edu/pathways/pathways	KEGG	http://prime.psc.riken.jp/?action=metabolites_index
AtIPD	http://www.atipd.ethz.ch/	KEGG Glycan	http://www.genome.jp/kegg/glycan/
BiGG	http://bigg.ucsd.edu/	KNApSAcK	http://prime.psc.riken.jp/?action=metabolites_index
BioCyc	http://biocyc.org/	LipidMaps	http://www.lipidmaps.org/
BioNumbers	http://bionumbers.hms.harvard.edu/	MarkerDB	http://www.markerdb.ca/users/sign_in
BML-NMR	http://www.bml-nmr.org/	MassBank	http://www.massbank.jp/
BioMagResBank	http://www.bmrb.wisc.edu/metabolomics/	MetaboAnalyst	http://www.metaboanalyst.ca/MetaboAnalyst/
BMDB	http://www.cowmetdb.ca/cgi-bin/browse.cgi	MetaboLights	http://www.ebi.ac.uk/metabolights/index
ChEBI	http://www.ebi.ac.uk/chebi/	MetaCrop	http://metacrop.ipk-gatersleben.de/apex/f?p=269:111:
ChEMBL	https://www.ebi.ac.uk/chembl/about#	MetaCyc	http://metacyc.org/
ChEBI	http://www.ebi.ac.uk/chebi/	METAGENE	http://www.metagene.de/program/a.prg
ChemMine	http://chemminedb.ucr.edu/	METLIN	https://metlin.scripps.edu/index.php
ChemSpider	http://www.chemspider.com/	MMCD	http://mmcd.nmrfam.wisc.edu/
CCD	http://ccd.chemnetbase.com/intro/index.jsp#about	mzCloud	https://mzcloud.org/
CSF Metabolome Database	http://www.csfmetabolome.ca/	OMIM	http://www.ncbi.nlm.nih.gov/omim/
CyberCell Database	http://ccdb.wishartlab.com/CCDB/	OMMBID	http://ommbid.mhmedical.com/
DrugBank	http://www.drugbank.ca/	Oryzabase	http://www.shigen.nig.ac.jp/rice/oryzabase/
ECMDB	http://www.ecmdb.ca/	PepBank	http://pepbank.mgh.harvard.edu/
ExPaSy Pathways	http://web.expasy.org/pathways/	PharmGKB	http://www.pharmgkb.org/
Fiehn GC-MS Database	http://fiehnlab.ucdavis.edu/Metabolite-Library-2007/	PMN	http://www.plantcyc.org/
FooDB	http://www.foodb.ca	PubChem	http://pubchem.ncbi.nlm.nih.gov/
GMDB	http://gmd.mpimp-golm.mpg.de/	Reactome	http://www.reactome.org/
HMDB	http://metabolomics.pharm.uconn.edu/iimdb/	RiceCyc	http://pathway.gramene.org/gramene/ricecyc.shtml
HumanCyc	http://www.genome.jp/kegg/	Serum Metabolome Database	http://www.serummetabolome.ca/
IIDMB	http://www.genome.jp/kegg/glycan/	SetupX & BinBase	http://fiehnlab.ucdavis.edu/projects/binbase_setupx

AACLifeSci

CEU Universidad San Pablo



- Dedicado a la anotación de metabolitos.
- Realiza búsquedas en compuestos unificados de distintas fuentes.
- Aplica conocimiento basado en la información de entrada proporcionada por el usuario.
- Ayuda a identificar lípidos oxidados.
- http://ceumass.eps.uspceu.es/mediator

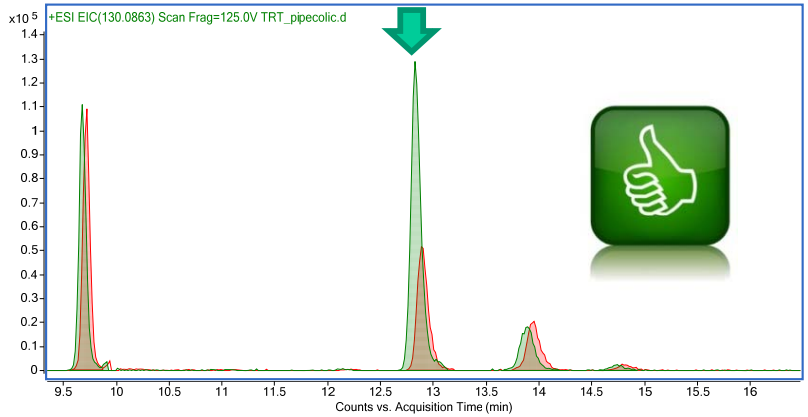
AACLifeSci

CEU Universidad San Pablo

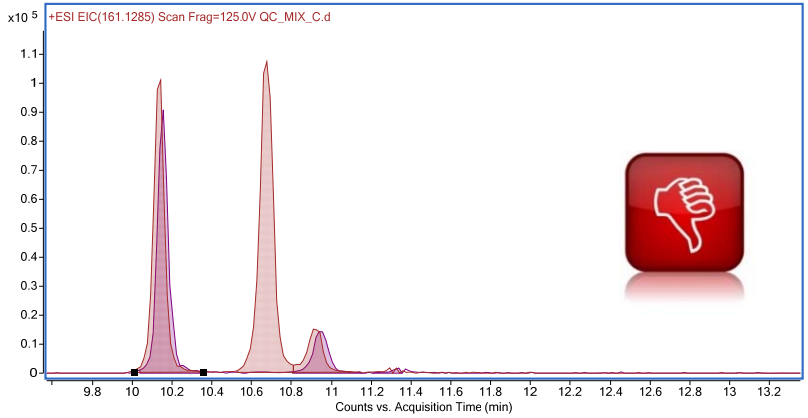
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	LIST OF COMPOUNDS															
2	Experimental mass	Identifier	Adduct	PPM Erro	Molecular Weigh	Name	Formula	CAS	Kegg	HMDB	LipidMaps	Metlin	PubChem	InChIKey	Pathways	
3	399.3367	17732	M+H	5	399.3349	L-palmitoylcarnitine	C23H45NO4				LMFA07070079			XOMRRQXKHMYMOC-NRFANRHFSA-N		
4	399.3367	17751	M+H	5	399.3349	O-palmitoylcarnitine	C23H45NO4				LMFA07070098			XOMRRQXKHMYMOC-UHFFFAOYSA-N		
5	399.3367	17657	M+H	5	399.3349	Palmitoylcarnitine	C23H45NO4	2364-67-2	C02990	HMDB0000222	LMFA07070004	36667	11953816	XOMRRQXKHMYMOC-OAQYLSRUS	Fatty acid Metabolism	Fatty a
6	399.3367	0	M+Na	0	0	No compounds found for experimental mass 399.3367 and adduct: I										
7	399.3367	13442	M+NH4	5	382.3083	methyl 9-butylperoxy-10,12-octadecadienoate	C23H42O4				LMFA01040036	74461		XVZNKFOXXKMPCH-IZAPIVEJSA-N		
8	399.3367	13443	M+NH4	5	382.3083	methyl 13-butylperoxy-9,11-octadecadienoate	C23H42O4				LMFA01040037	74462		GUWYTNFTPALBOJ-AEPWDTSSSA-N		
9	399.3367	95769	M+NH4	5	382.3083	Lepidiumterpenyl ester	C23H42O4			HMDB0036865		91868	11794320	PASMASQJCDKBJK-UHFFFAOYSA-N		
10	399.3367	53861	M+NH4	5	382.3083	MG(0:0/20:2(11Z,14Z)/0:0)	C23H42O4			HMDB0011544		62328	53480964	PMJSUEZTCFTBMD-HZJYTTRNSA-N		
11	399.3367	52646	M+NH4	5	382.3083	MG(20:2(11Z,14Z)/0:0/0:0)	C23H42O4			HMDB0011574		62356	53480983	QRBGFYBOCBYOSN-KDTZXJSHSA-N		
12	399.3367	80631	M+NH4	5	382.3083	Persenone B	C23H42O4			HMDB0035959		91145		NLXNQLZUOMHEHB-ISLYRVAYSA-N		
13	399.3367	0	M+H-H2O	0	0	No compounds found for experimental mass 399.3367 and adduct: I										
14	421.3169	16968	M+H	6	421.3192	Gamma-linolenyl carnitine	C25H43NO4			HMDB0006318	LMFA07010893	58389	53477819	VDPMHVILMVSAQX-BAHSRKMSSA-N		
15	421.3169	96332	M+H	6	421.3192	Alpha-linolenyl carnitine	C25H43NO4			HMDB0006319		58390	53477821	DFVGGGHKDAHYIU-UHMZJXMFSA-N		
16	421.3169	126612	M+H	9	421.3205	AGELASINE	C26H39N5					43731				
17	421.3169	138401	M+H	9	421.313	Latanoprost ethyl amide-d4	C25H35D4NO4					36571				
18	421.3169	17732	M+Na	0	399.3349	L-palmitoylcarnitine	C23H45NO4				LMFA07070079			XOMRRQXKHMYMOC-NRFANRHFSA-N		
19	421.3169	17751	M+Na	0	399.3349	O-palmitoylcarnitine	C23H45NO4				LMFA07070098			XOMRRQXKHMYMOC-UHFFFAOYSA-N		
20	421.3169	17657	M+Na	0	399.3349	Palmitoylcarnitine	C23H45NO4	2364-67-2	C02990	HMDB0000222	LMFA07070004	36667	11953816	XOMRRQXKHMYMOC-OAQYLSRUS	Fatty acid Metabolism	Fatty a
21	421.3169	1294	M+NH4	6	404.2927	1alpha,25-dihydroxy-21-nor-20-oxavitamin D3 / 1alpha,25-dihydroxy-2	C25H40O4				LMST03020029	41970		AOMQZQQKWZZTNV-OQGZJSIESA-N		
22	421.3169	1295	M+NH4	6	404.2927	1alpha,25-dihydroxy-24-nor-22-oxavitamin D3 / 1alpha,25-dihydroxy-2	C25H40O4				LMST03020030	41971		UMRLCGLUMINBCT-NPNXQCSOSA-N		
23	421.3169	2650	M+NH4	6	404.2927	7b-Hydroxy-3-oxo-5b-cholan-24-oate	C25H40O4				LMST04070028	57939		XHRLTYUHWGHDCJ-JFRFJXPMSA-N		
24	421.3169	117990	M+NH4	6	404.2927	Androstane-3,17-diol dipropionate;5alpha-Androstane-3alpha,17bet	C25H40O4	4350-14-5	C14624			70213	134572	FWAYUWSHYLKUEY-LFWYWRBCSA-N		
25	421.3169	86607	M+NH4	6	404.2927	11'-Carboxy-gamma-chromanol	C25H40O4			HMDB0012517			53481453	IITULCXNOMOXAH-YYULODDRSA-N		
26	421.3169	89679	M+NH4	6	404.2927	MG(0:0/22:5(7Z,10Z,13Z,16Z,19Z)/0:0)	C25H40O4			HMDB0011556		62339	53480971	LRBJILYYDLUADN-JLNKQSITSA-N		
27	421.3169	50261	M+NH4	6	404.2927	MG(22:5(4Z,7Z,10Z,13Z,16Z)/0:0/0:0)	C25H40O4			HMDB0011585		62367	53480993	HDIQCISTZKHUDO-AJWITYRPSA-N		
28	421.3169	105373	M+NH4	6	404.2927	MG(0:0/22:5(4Z,7Z,10Z,13Z,16Z)/0:0)	C25H40O4			HMDB0011555		62338	53480970	NPZWSBAEZZLYQU-WMPRHZDHSA-N		
29	421.3169	56824	M+NH4	6	404.2927	MG(22:5(7Z,10Z,13Z,16Z,19Z)/0:0/0:0)	C25H40O4			HMDB0011586		62368	53480994	IDSLCYURGAOTOA-YAWQMZOFSA-N		
30	421.3169	17695	M+H-H2O	5	439.3298	3-hydroxylinoleoylcarnitine	C25H45NO5				LMFA07070042			WQYXCASYXUFNSI-UTJQPWESSA-N		
31	315.2424	17712	M+H	5	315.241	Decanoylcarnitine	C17H33NO4			HMDB0000651	LMFA07070059		10245190	LZOSYCMHQXPBFU-UHFFFAOYSA-N		
32	315.2424	126591	M+H	5	315.241	L-Hexanoylcarnitine n-butyl ester	C17H33NO4					3549				
33	315.2424	17659	M+H	5	315.241	O-decanoyl-R-carnitine	C17H33NO4	3992-45-8	C03299	HMDB0062631	LMFA07070006	36669	11953821	LZOSYCMHQXPBFU-OAHLLOKOSA-N		
34	315.2424	0	M+Na	0	0	No compounds found for experimental mass 315.2424 and adduct: M										
35	315.2424	14277	M+NH4	5	298.2144	8E-Heptadecenedioic acid	C17H30O4				LMFA01170053	74325		VDTSYDDUGXHLDV-OWOJBTEDSA-N		
36	315.2424	111819	M+NH4	5	298.2144	Plakortic acid	C17H30O4		C17158			71590	10402441	ZCLJFHUIADAYRQ-CMDGGOBGSA-N		
37	315.2424	0	M+H-H2O	0	0	No compounds found for experimental mass 315.2424 and adduct: M										
38	337.2234	0	M+H	0	0	No compounds found for experimental mass 337.2234 and adduct: I										
39	337.2234	17712	M+Na	2	315.241	Decanoylcarnitine	C17H33NO4			HMDB0000651	LMFA07070059		10245190	LZOSYCMHQXPBFU-UHFFFAOYSA-N		
40	337.2234	126591	M+Na	2	315.241	L-Hexanoylcarnitine n-butyl ester	C17H33NO4					3549				
41	337.2234	17659	M+Na	2	315.241	O-decanoyl-R-carnitine	C17H33NO4	3992-45-8	C03299	HMDB0062631	LMFA07070006	36669	11953821	LZOSYCMHQXPBFU-OAHLLOKOSA-N		
42	337.2234	1060	M+NH4	6	320.1988	testolic acid	C19H28O4		C01618		LMST02020081	57817	439534	KMUJXIPRPXRPTP-DZBHQSCQSA-N		
43	337.2234	116962	M+NH4	6	320.1988	10beta-Hydroxy-6beta-isobutyrylfuranoeremophilane	C19H28O4		C09685			67884	442377	WWBNAQLGGJMUOJ-BIGGFVEDSA-N		
44	337.2234	130809	M+NH4	6	320.1988	(±)a-CMBHC	C19H28O4	7083-09-2				44834				
45	337.2234	53386	M+NH4	6	320.1988	[8]-Gingerdione	C19H28O4	77334-06-6		HMDB0039276		93864	14440537	QDSRAFNZQKMHPZ-UHFFFAOYSA-N		
46	337.2234	42665	M+NH4	6	320.1988	5'-Carboxy-alpha-chromanol	C19H28O4			HMDB0012798			53481524	QWPNLYBAEZJBMI-IMEVZPHKSA-N		

Sheet0

Confirmación mediante Adición estándar



Ácido pipecólico

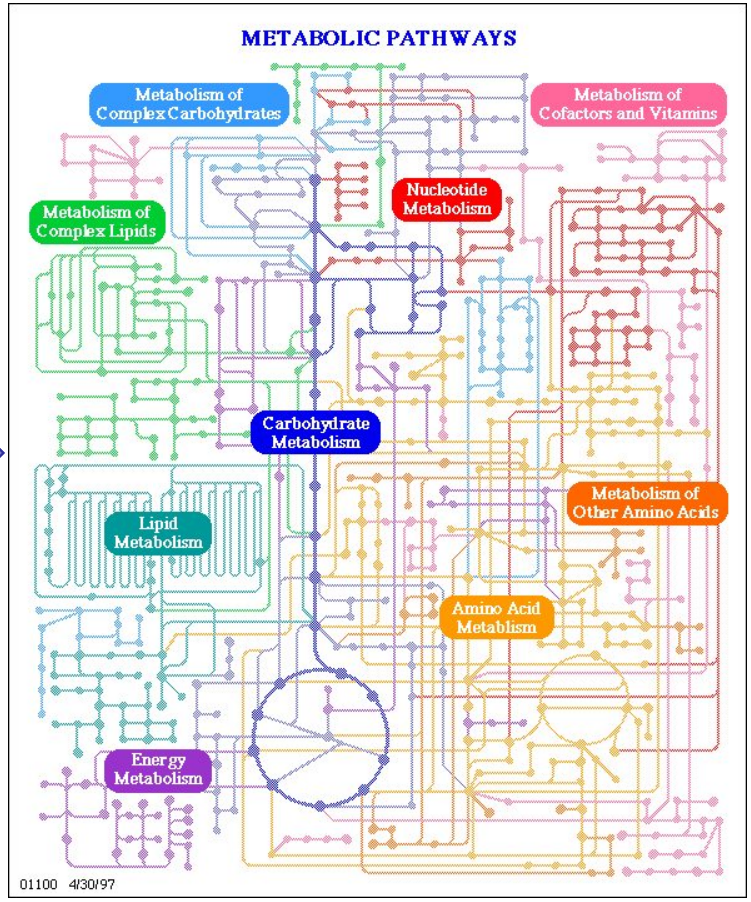


Metil-lisina

De Listas a Rutas

metabolómica

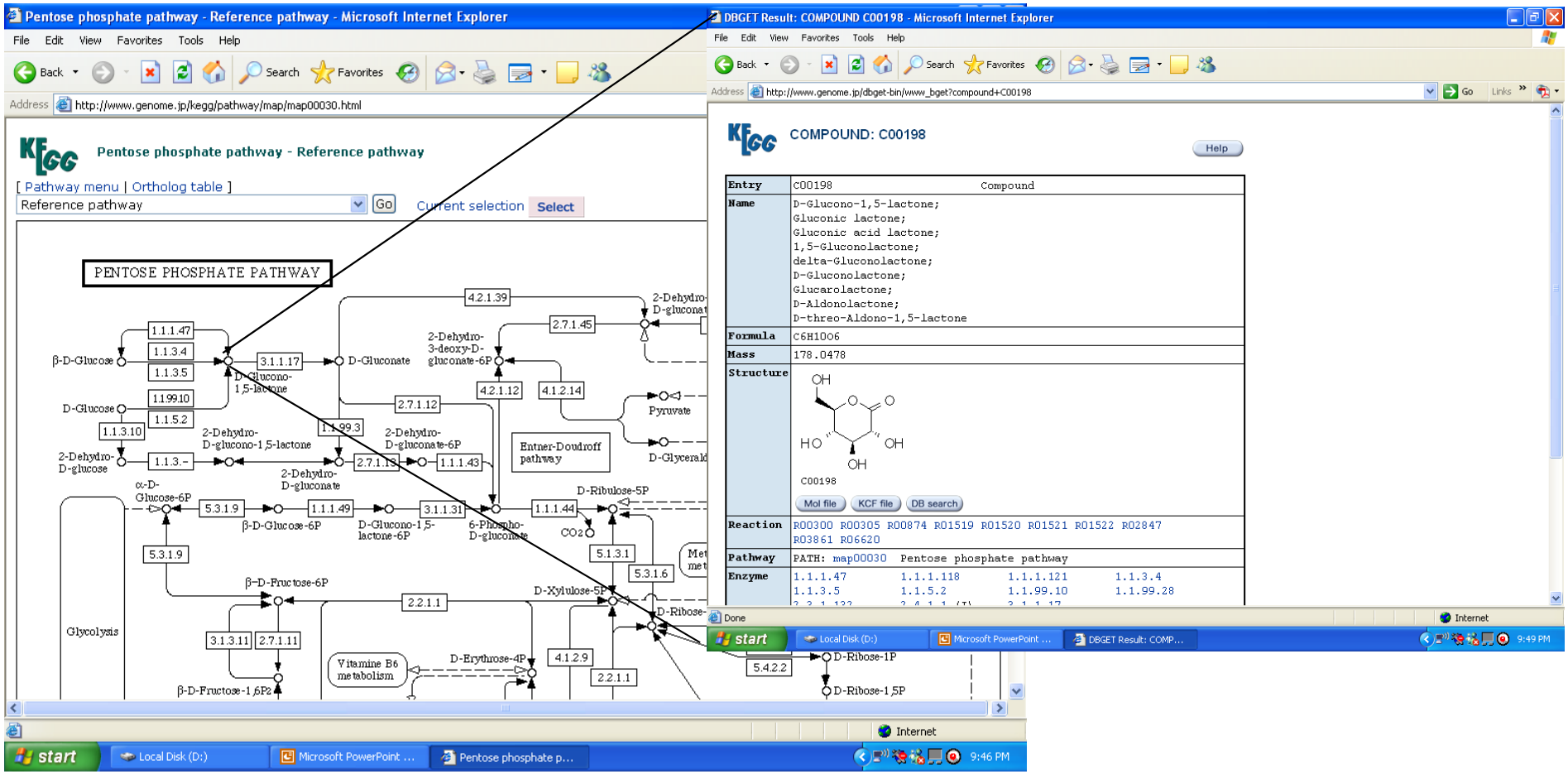
Compound	Retention Time (min)	Conc. in Urine (µM)	Compound	Retention Time (min)	Conc. in Urine (µM)
Dns-o-phospho -L-serine	0.92	<D.L. *	**Dns-Ile**	6.35	25
Dns-o-phospho -L-tyrosine	0.95	<D.L.	**Dns-3-aminosalicylic acid**	6.44	0.5
Dns-adnosine monophosphate	0.99	<D.L.	**Dns-pipecolic acid**	6.50	0.5
Dns-o-phosphoethanolamine	1.06	16	**Dns-Leu**	6.54	54
Dns-glucosamine	1.06	22	**Dns-cystathionine**	6.54	0.3
Dns-o-phospho -L-threonine	1.09	<D.L.	**Dns-Leu-Pro**	6.60	0.4
Dns-6-dimet hylamine purine	1.20	<D.L.	**Dns-5-hydroxylysine**	6.65	1.6
Dns-3-methyl -histidine	1.22	80	**Dns-Cystine**	6.73	160
Dns-taurine	1.25	834	**Dns-N-norleucine**	6.81	0.1
Dns-carnosine	1.34	28	Dns-5-hydroxydopamine	7.17	<D.L.
Dns-Arg	1.53	36	**Dns-dimethylamine**	7.33	293
Dns-Asn	1.55	133	**Dns-5-HIAA**	7.46	18
Dns-hypotaurine	1.58	10	**Dns-umbelliferone**	7.47	1.9
Dns-homocarnosine	1.61	3.9	Dns-2,3-diaminoproprionic acid	7.63	<D.L.
Dns-guanidine	1.62	<D.L.	**Dns-L-ornithine**	7.70	15
Dns-Gln	1.72	633	**Dns-4-acetyamidophenol**	7.73	51
Dns-allantoin	1.83	3.8	**Dns-procaine**	7.73	8.9
Dns-L-citrulline	1.87	2.9	**Dns-homocystine**	7.76	3.3
Dns-1 (or 3 -)-methylhistamine	1.94	1.9	**Dns-acetaminophen**	7.97	82
Dns-adenosine	2.06	2.6	**Dns-Phe-Phe**	8.03	0.4
Dns-methylguanidine	2.20	<D.L.	**Dns-5-methyo xysalicylic acid**	8.04	2.1
Dns-Ser	2.24	511	**Dns-Lys**	8.16	184
Dns-aspartic acid amide	2.44	26	Dns-aniline	8.17	<D.L.
Dns-4-hydroxy -proline	2.56	2.3	**Dns-leu-Phe**	8.22	0.3
Dns-Glu	2.57	21	**Dns-His**	8.35	1550
Dns-Asp	2.60	90	Dns-4-thialysine	8.37	<D.L.
Dns-Thr	3.03	157	Dns-benzylamine	8.38	<D.L.
Dns-epinephrine	3.05	<D.L.	**Dns-1-ephedrine**	8.50	0.6
Dns-ethanolamine	3.11	471	**Dns-tryptamine**	8.63	0.4
Dns-aminoadipic acid	3.17	70	Dns-pyrydoxamine	8.94	<D.L.
Dns-Gly	3.43	2510	Dns-2-methyl -benzylamine	9.24	<D.L.
Dns-Ala	3.88	593	**Dns-5-hydroxytrptophan**	9.25	0.12
Dns-aminolevulinic acid	3.97	30	**Dns-1,3-diaminopropane**	9.44	0.23
Dns-r-amino-butyric acid	3.98	4.6	**Dns-putrescine**	9.60	0.5
Dns-p-amino-hippuric acid	3.98	2.9	**Dns-1,2-diaminopropane**	9.66	0.1
Dns-5-hydro xymethyluricil	4.58	1.9	**Dns-tyrosinamide**	9.79	29
Dns-tryptophanamide	4.70	5.5	**Dns-dopamine**	10.08	140
Dns-isoguanine	4.75	<D.L.	**Dns-cadaverine**	10.08	0.08
Dns-5-aminopentanoic acid	4.79	1.6	**Dns-histamine**	10.19	0.4
Dns-sarcosine	4.81	7.2	**Dns-3-methoxy -tyramine**	10.19	9.2
Dns-3-amino -isobutyrate	4.81	85	**Dns-Tyr**	10.28	321
Dns-2-aminobutyric acid	4.91	17	Dns-cysteamine	10.44	<D.L.



Bases de datos de rutas metabólicas

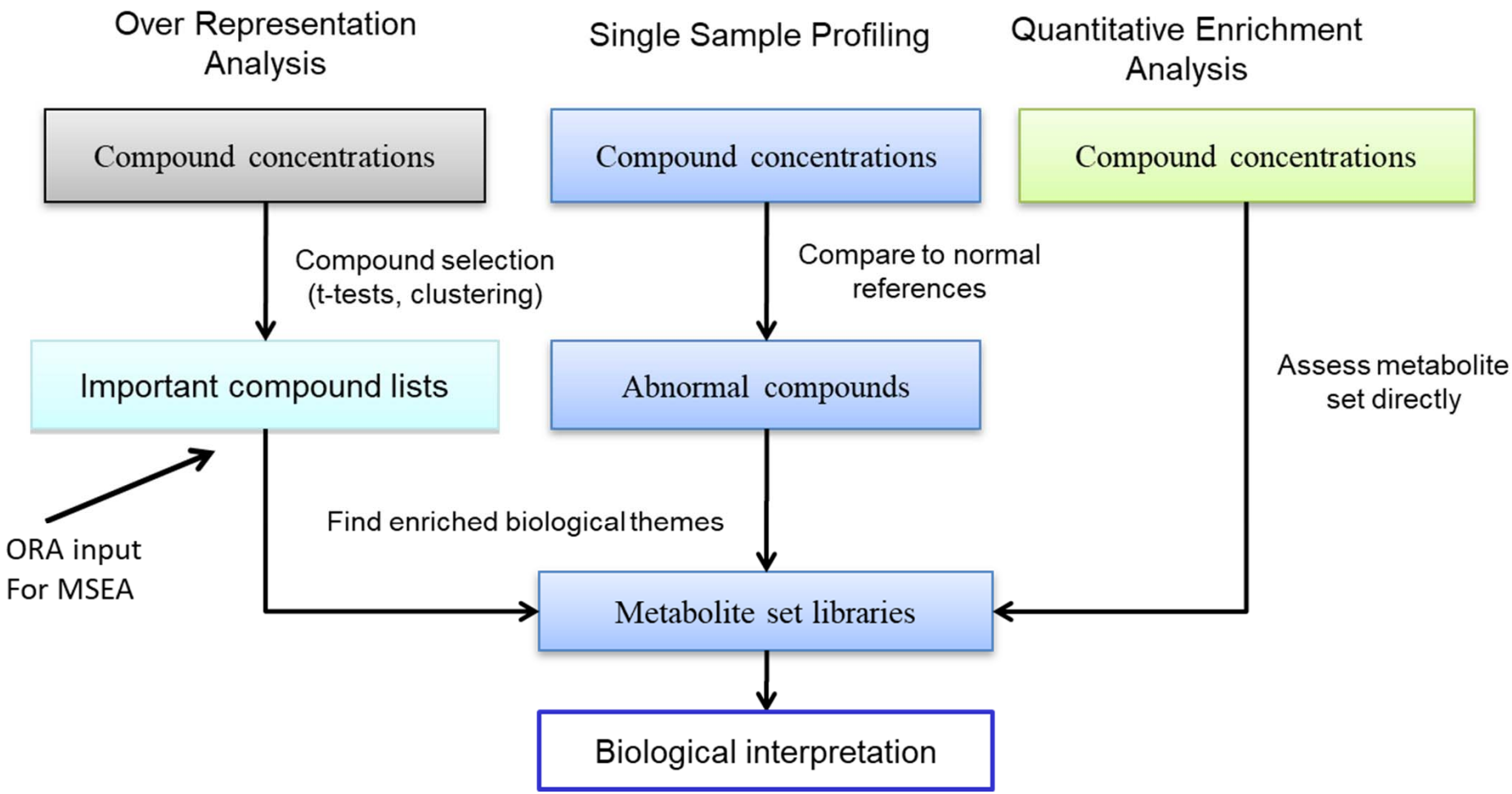
- Excelente fuente de datos biológicos que sirve para relacionar metabolitos con genes, proteínas y enfermedades, señalando eventos y procesos
- Ofrece distintas herramientas que permiten la visualización y el mapeo de genes/metabolitos
- Normalmente cubre múltiples especies
- KEGG (www.genome.jp/kegg/), BioCyc/MetaCyc (https://biocyc.org/), SMPDB (www.smpdb.ca), Reactome (www.reactome.org), WikiPathways (http://www.wikipathways.org)...
- *«De forma estricta, se podría afirmar que no existen las rutas metabólicas... son solo redes.»* (WikiPathways.org)

KEGG – Enciclopedia de Genes y Genomas de Kioto

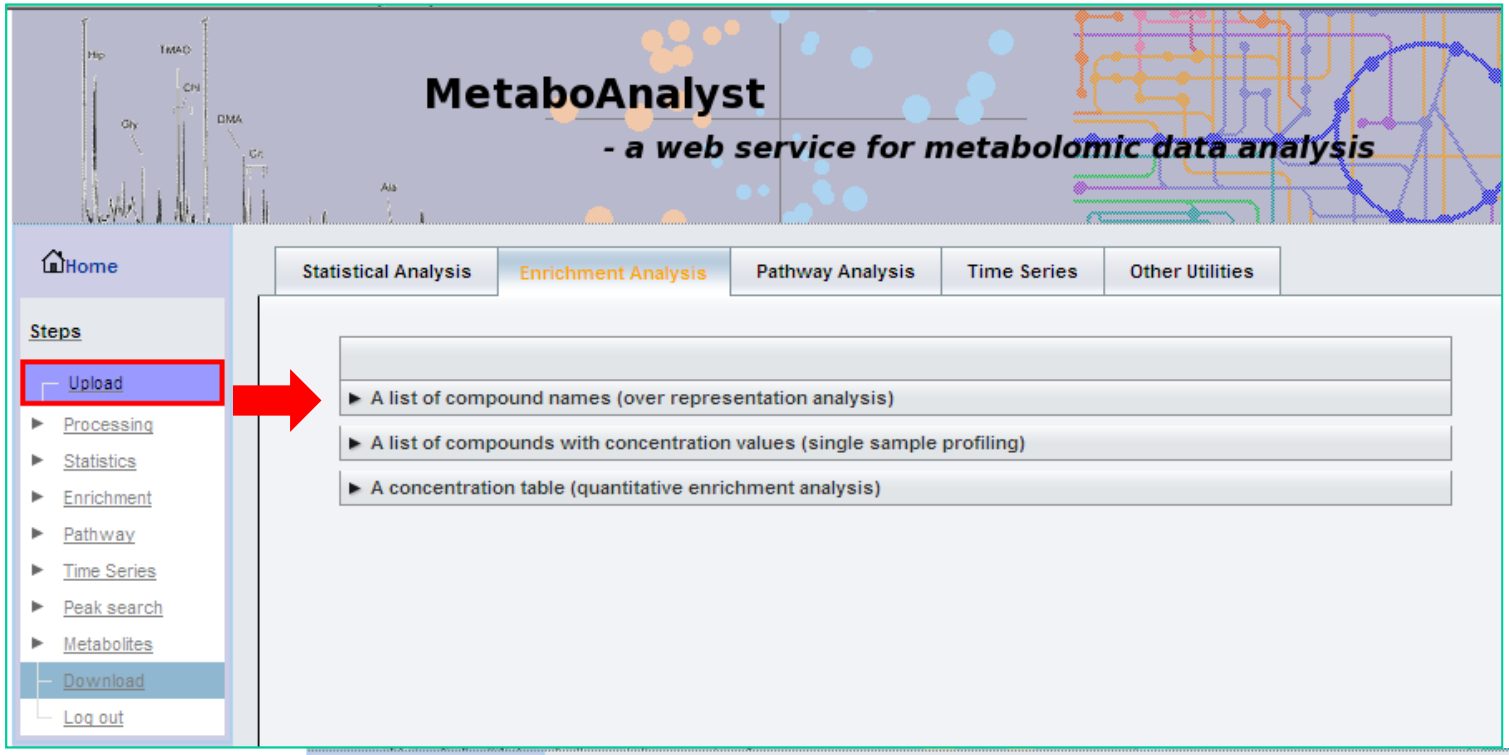


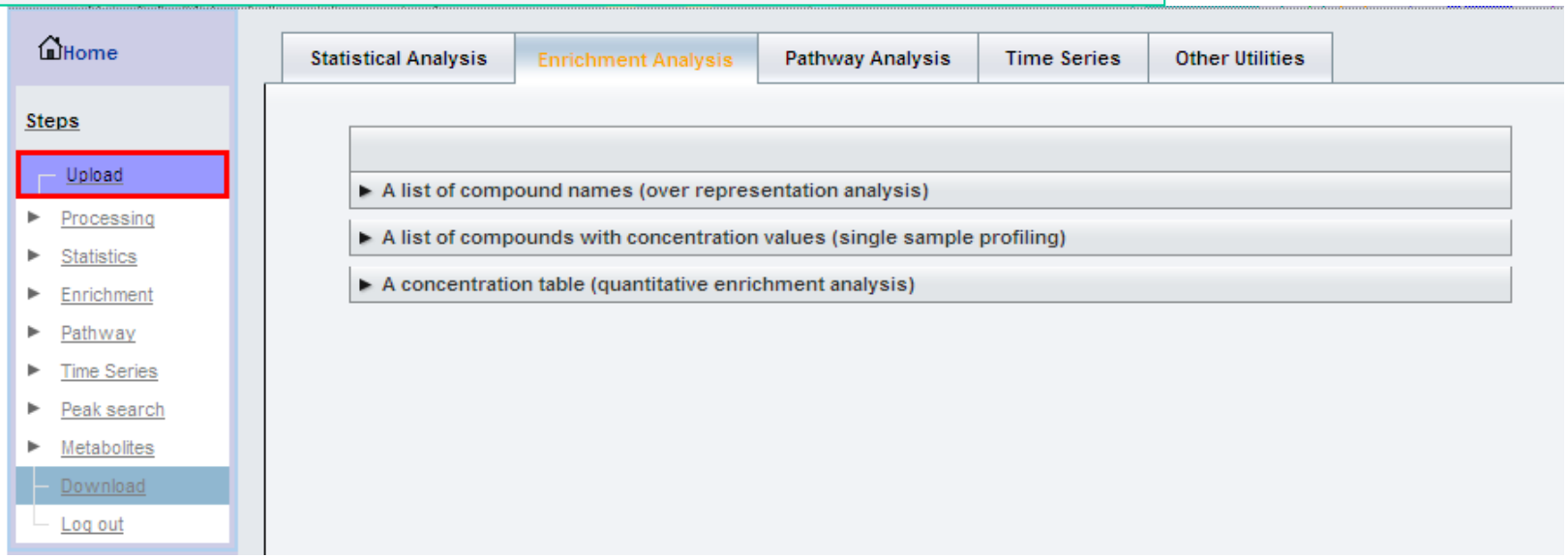
http://www.genome.jp/kegg/

El enfoque MSEA (*Metabolite Set Enrichment Analysis*)

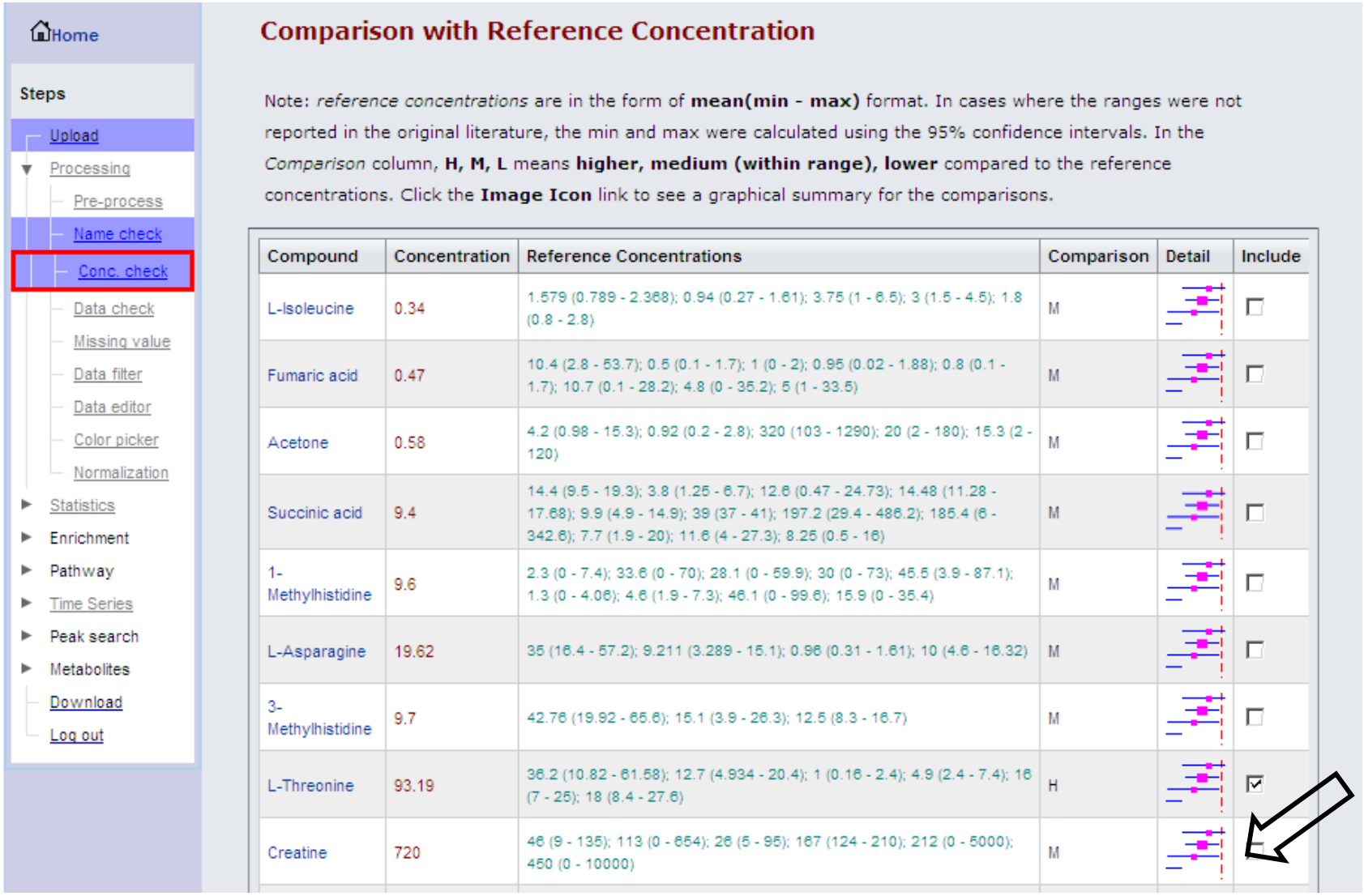


Empezar con una lista de compuestos

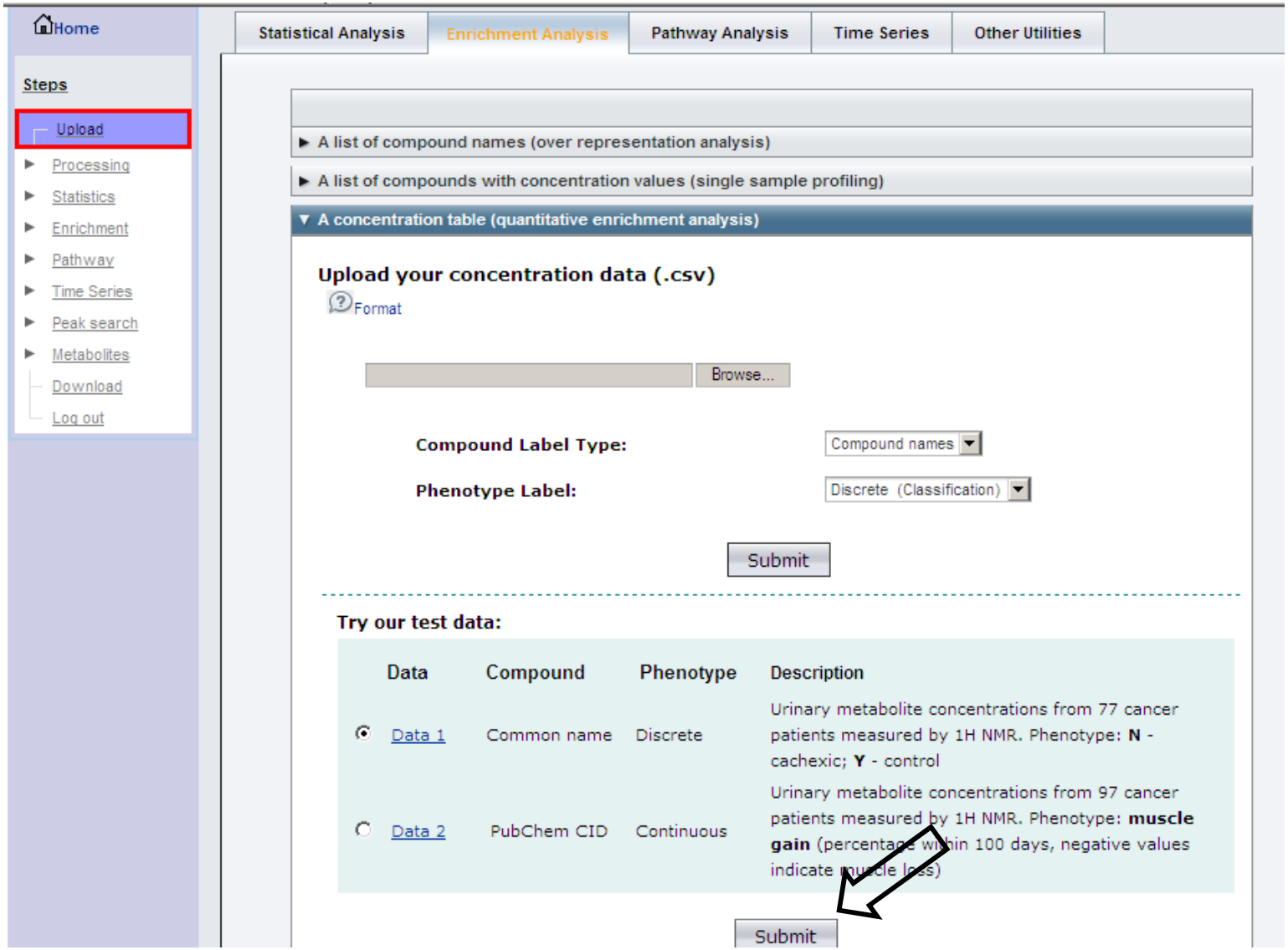




Comparación con la concentración



Análisis de Enriquecimiento Cuantitativo (*Quantitative enrichment analysis*)

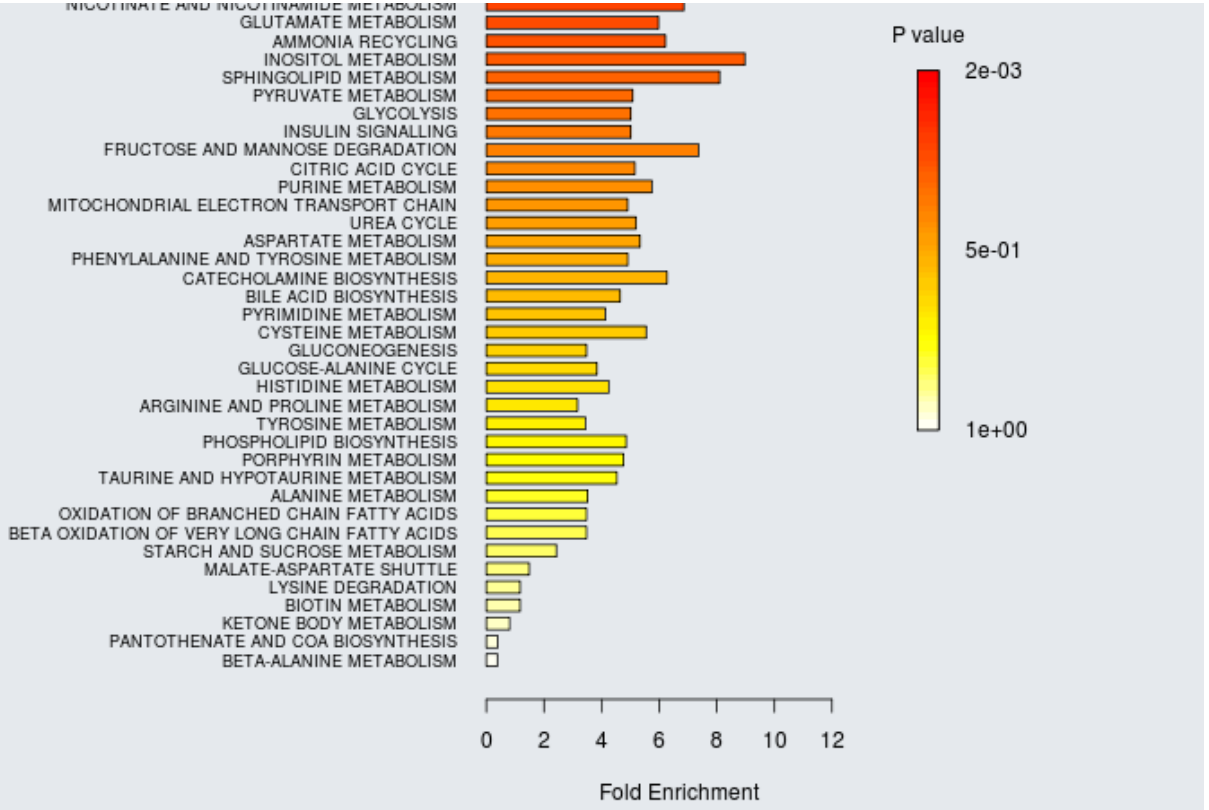


RESULTADO

AACLifeSci

CEU Universidad San Pablo

- Time Series
- Peak search
- Metabolites
- Download
- Log out

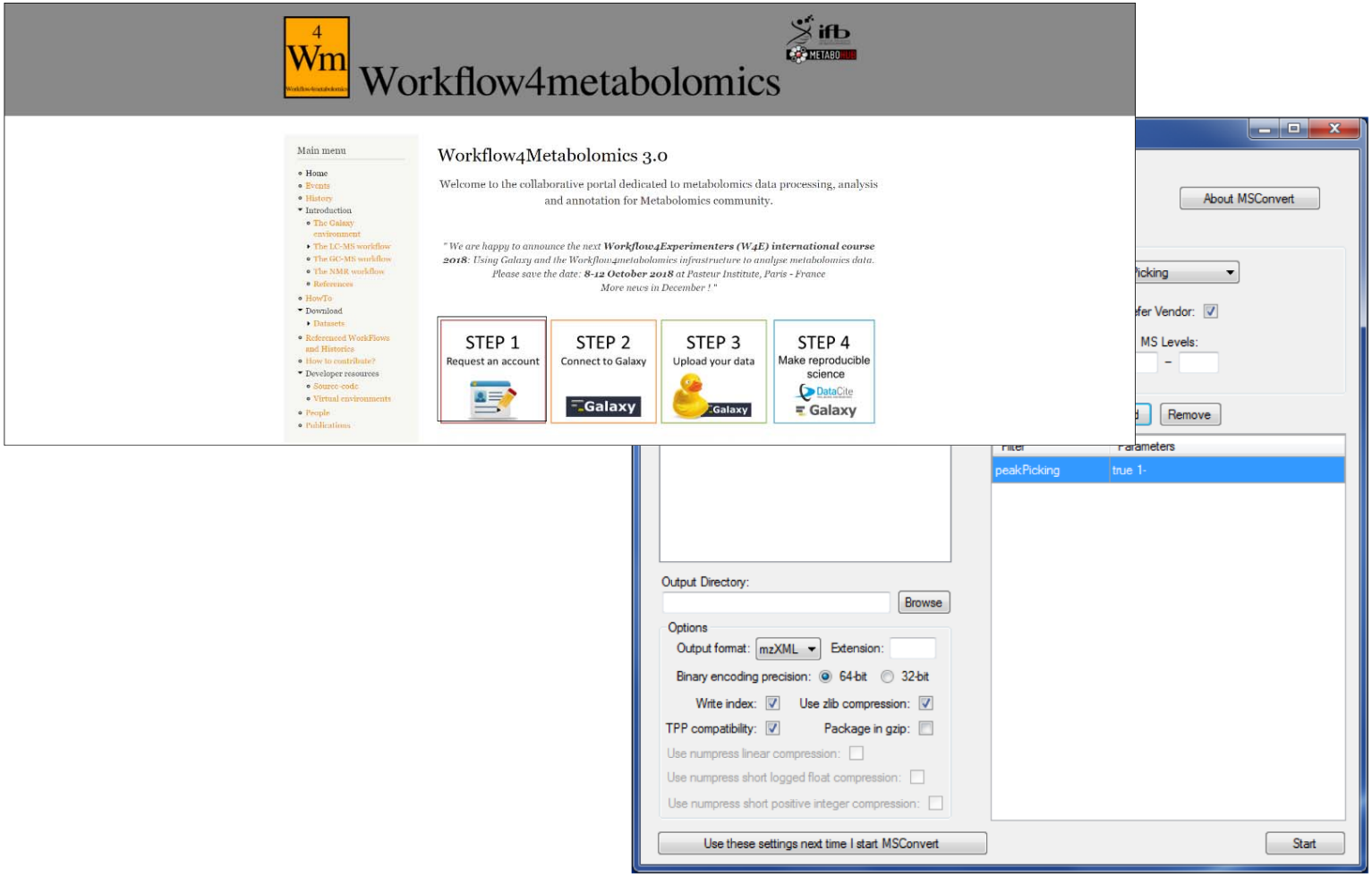


	Metabolite Set	Total	Hit	Statistic	Expected	P Value	Holm P	FDR	Details
	TRYPTOPHAN METABOLISM	34	2	15.088	1.3158	5.3712E-5	0.0024707	0.0020529	
	PROPANOATE METABOLISM	18	1	17.695	1.3158	1.3942E-4	0.0062741	0.0020529	
	BETAINE METABOLISM	10	2	14.311	1.3158	1.4515E-4	0.0063865	0.0020529	
	METHIONINE METABOLISM	24	4	11.386	1.3158	1.7852E-4	0.0076762	0.0020529	

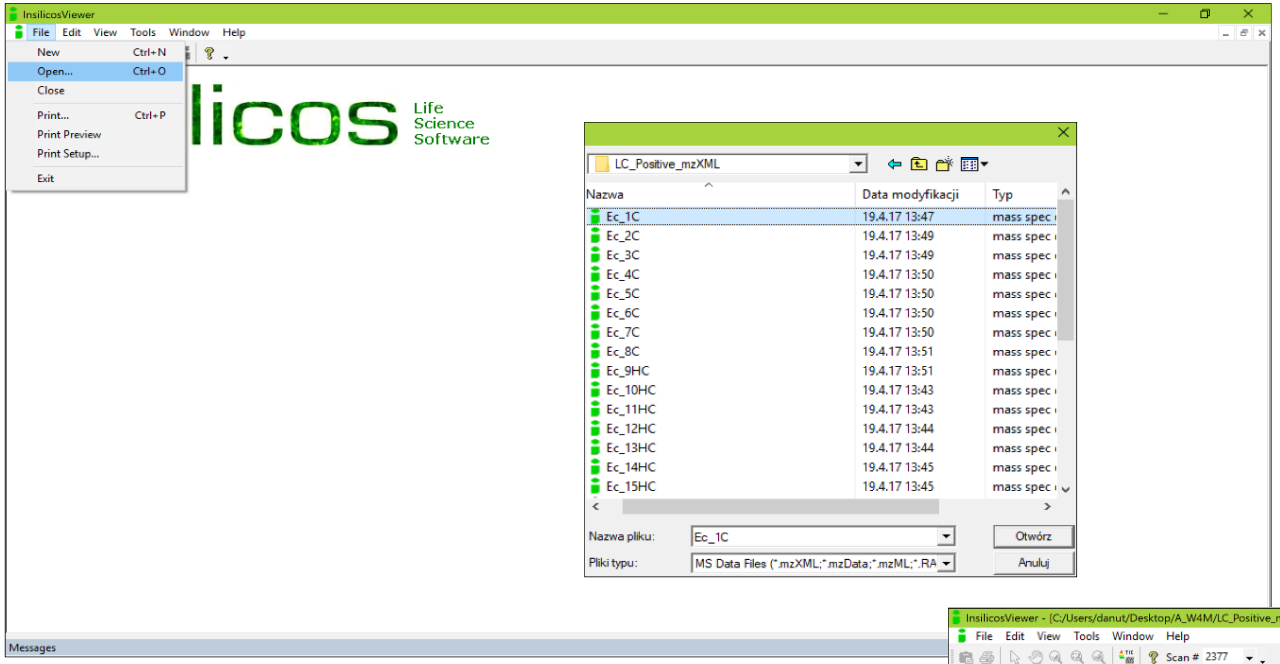
Metaboanalyst
Metabolic Pathway Analysis (MetPA)

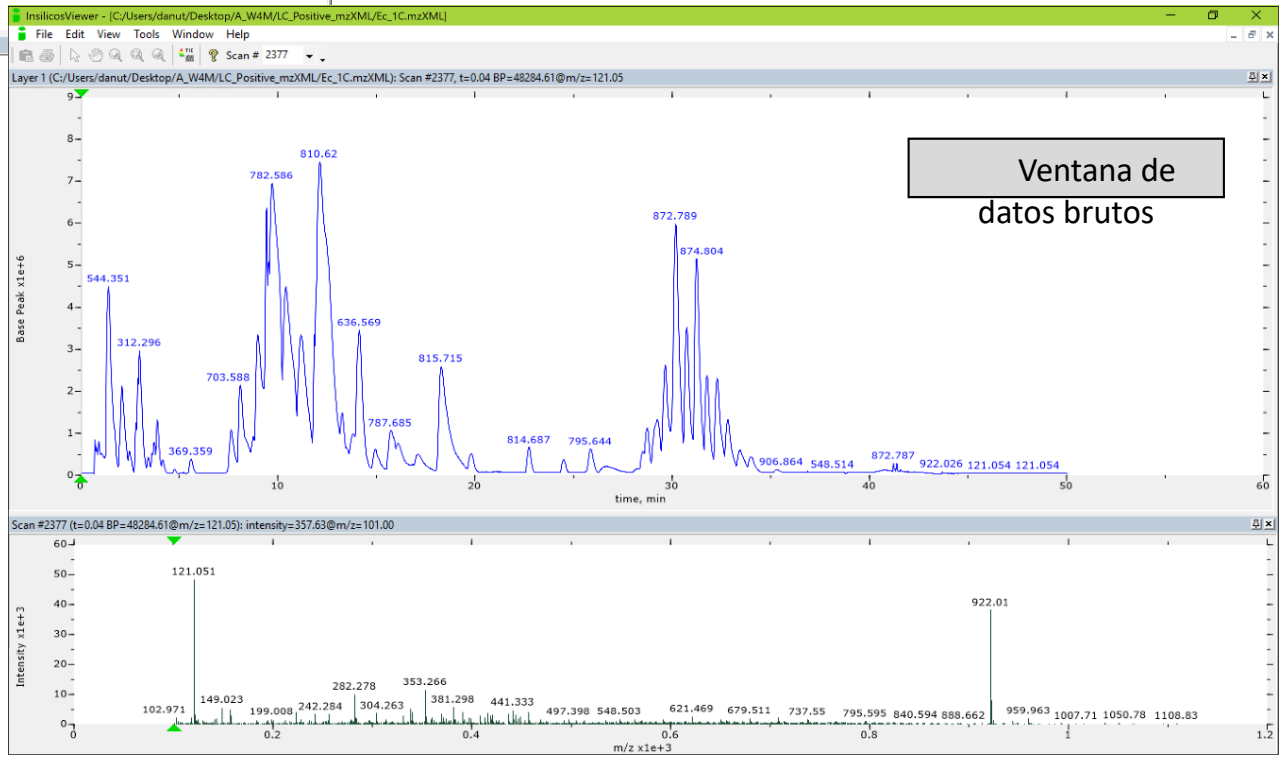
- Objetivo: extender y potenciar el MSEA para rutas metabólicas mediante
  - La consideración de las **estructuras de ruta metabólicas**
  - La visualización dinámica de las rutas metabólicas
- Actualmente admite aprox.1500 rutas que comprenden 17 organismos (basadas en KEGG)

SESIÓN PRÁCTICA. VISUALS_1



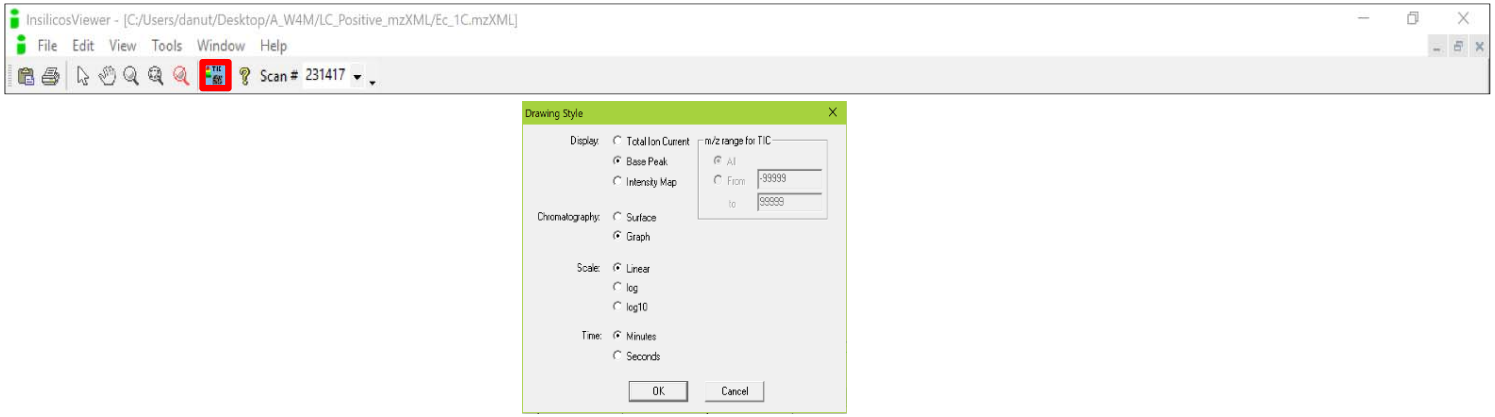
SESIÓN PRÁCTICA. VISUALS_2

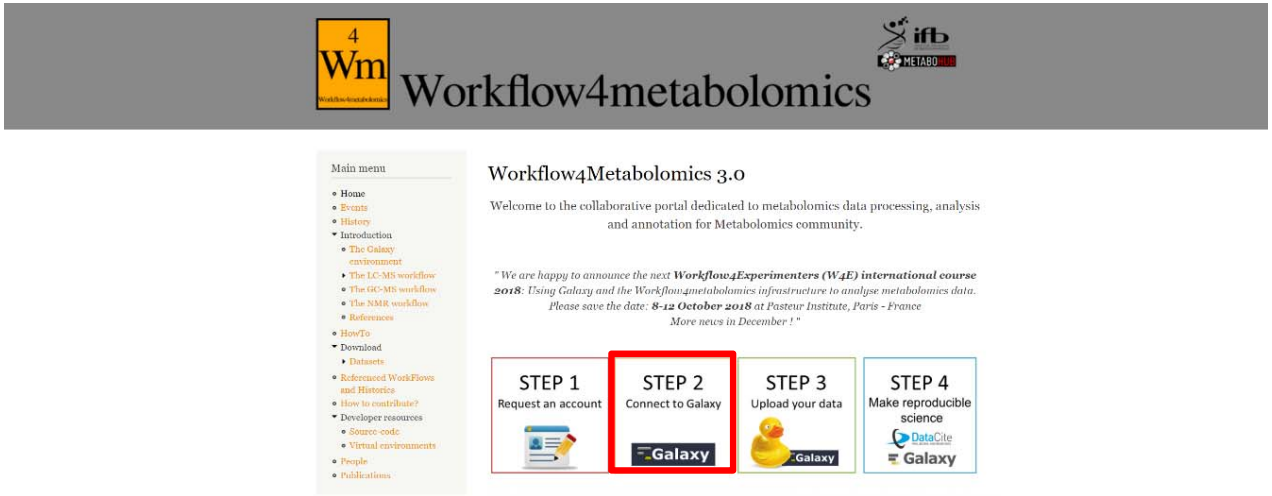




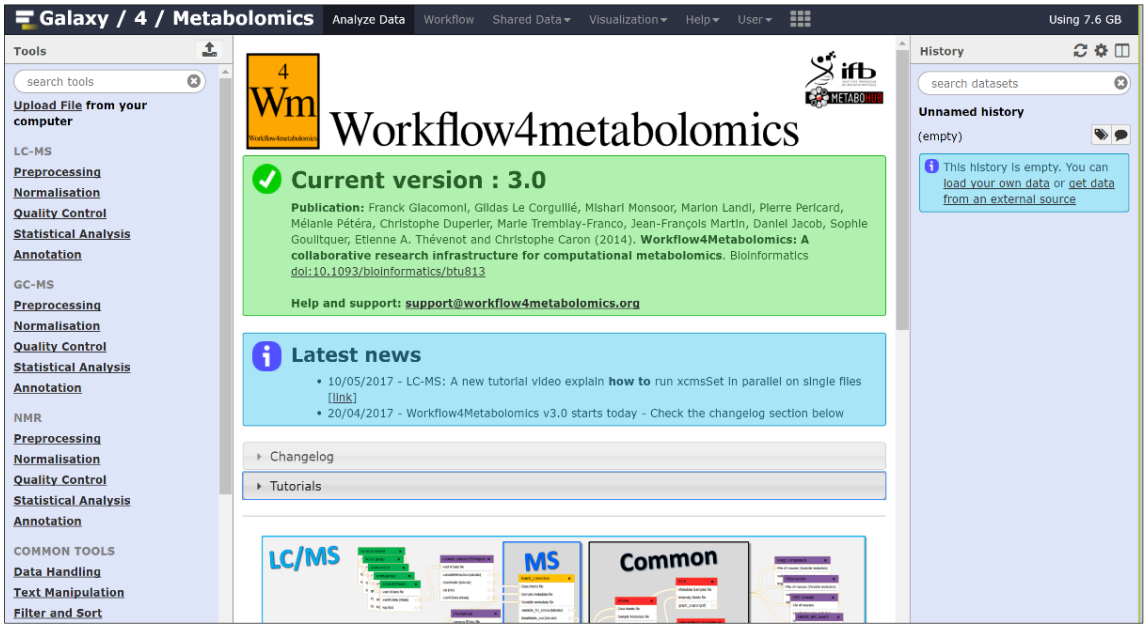
Ventana de datos brutos

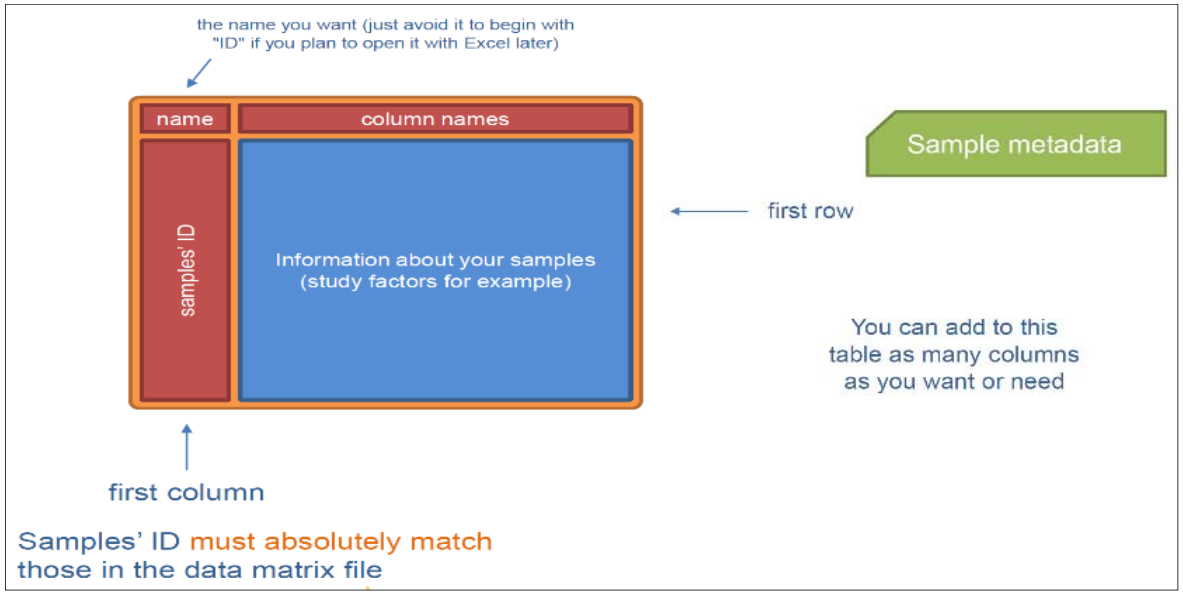
SESIÓN PRÁCTICA. VISUALS_3





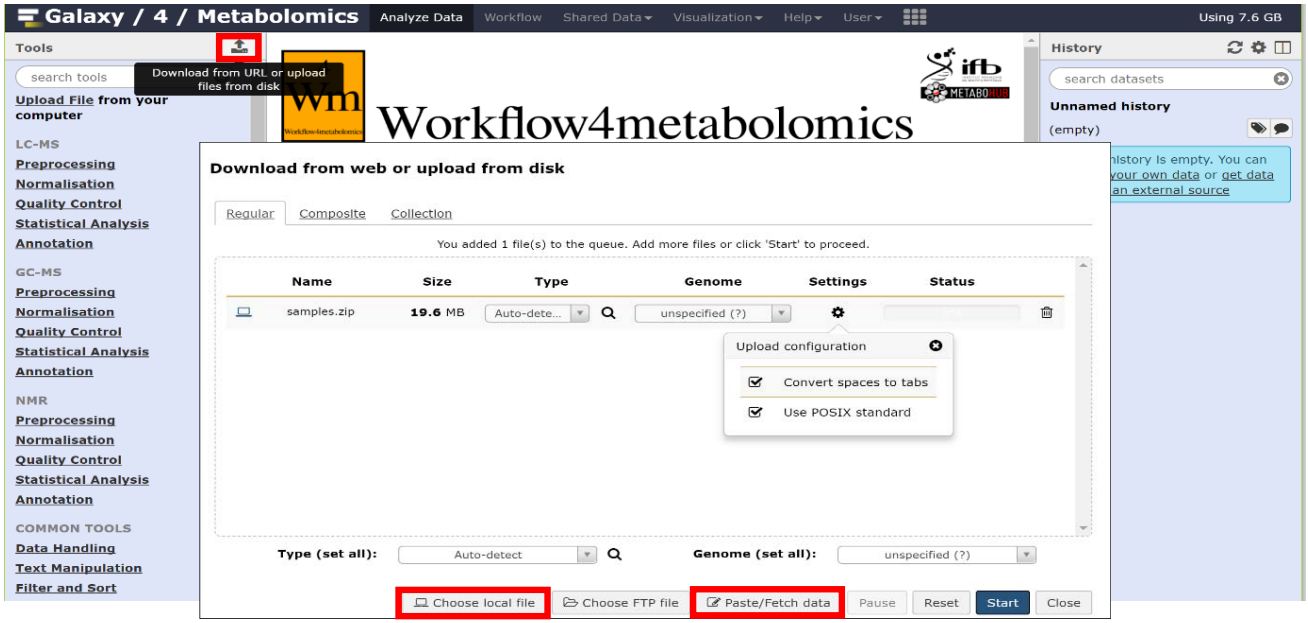
SESIÓN PRÁCTICA. VISUALS_4



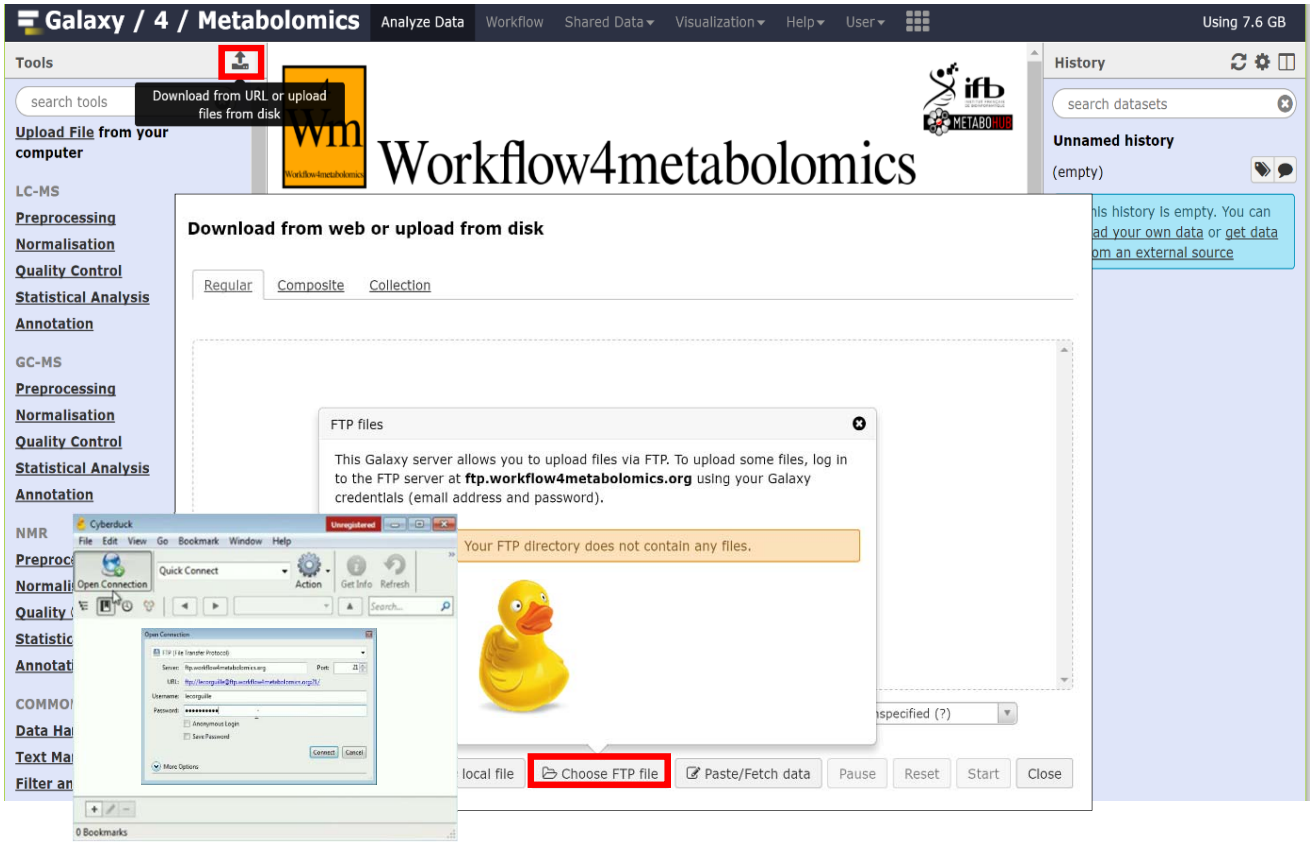


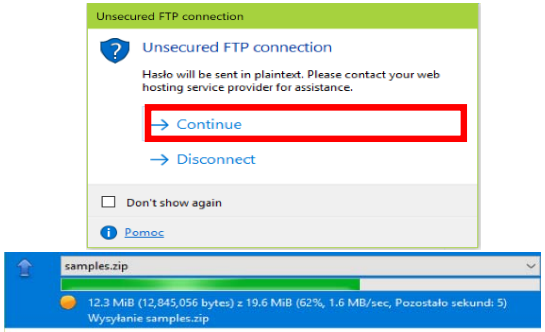
SESIÓN PRÁCTICA. VISUALS_5

sampleName	class	polarity	sampleType	batch	injectionOrder	diet
QC	one	positive	pool	B1	1	NA
C1	one	positive	sample	B1	7	C
HC3	one	positive	sample	B1	10	HC
BL	one	positive	blank	B1	12	NA
…	…	…	…	…	…	…

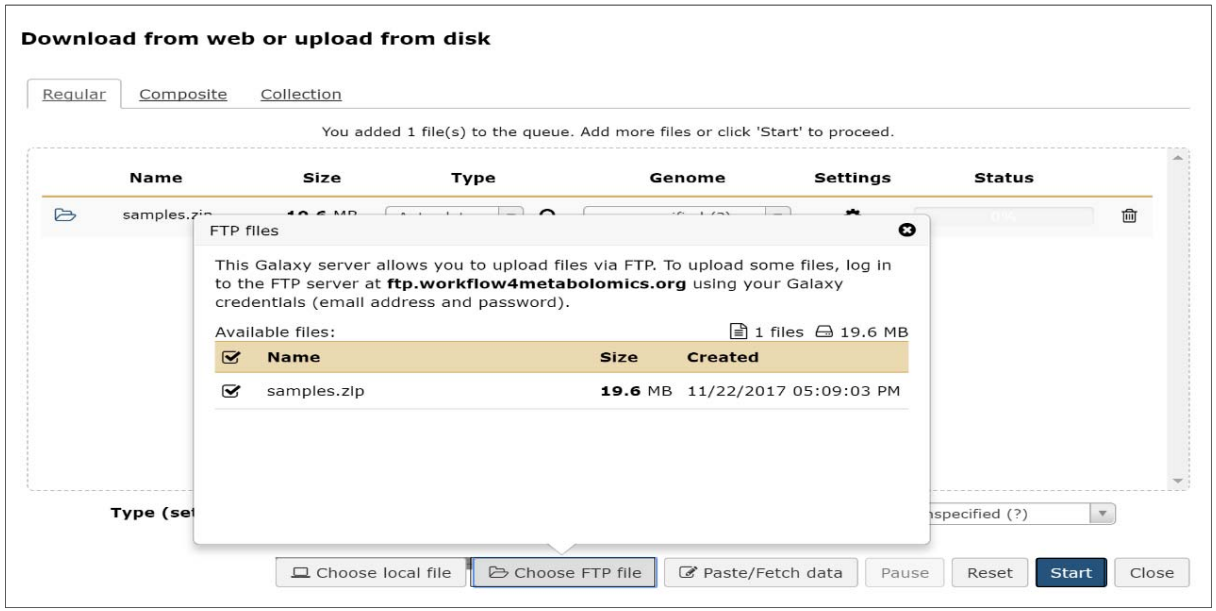


SESIÓN PRÁCTICA. VISUALS_6

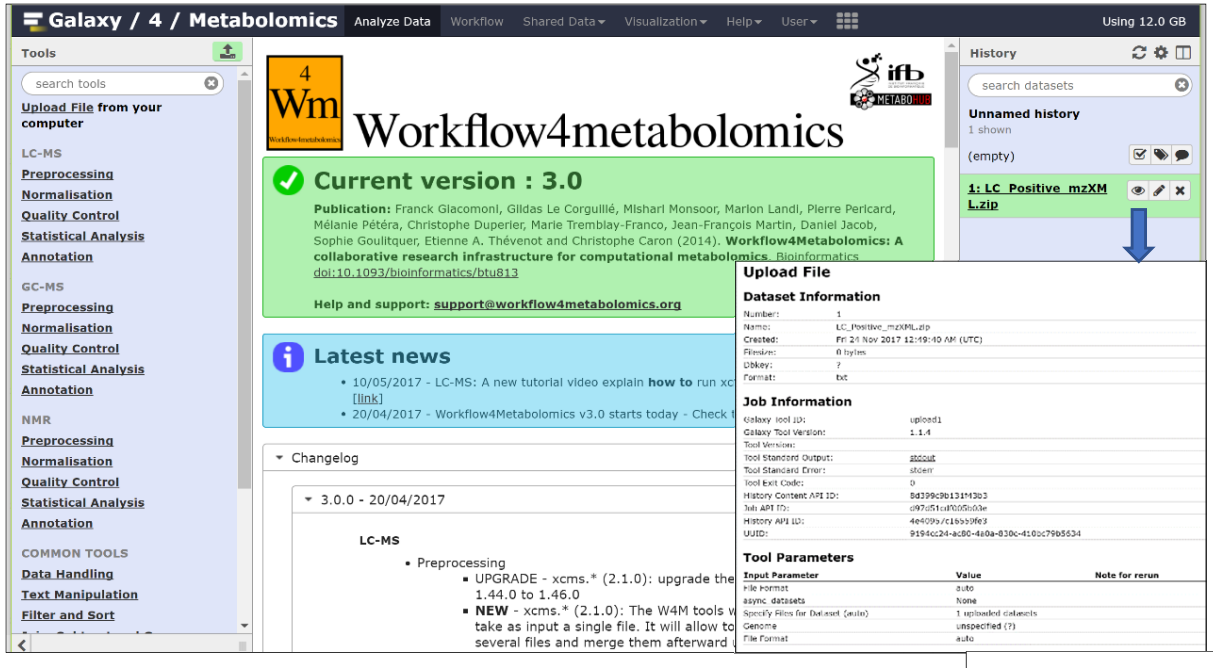


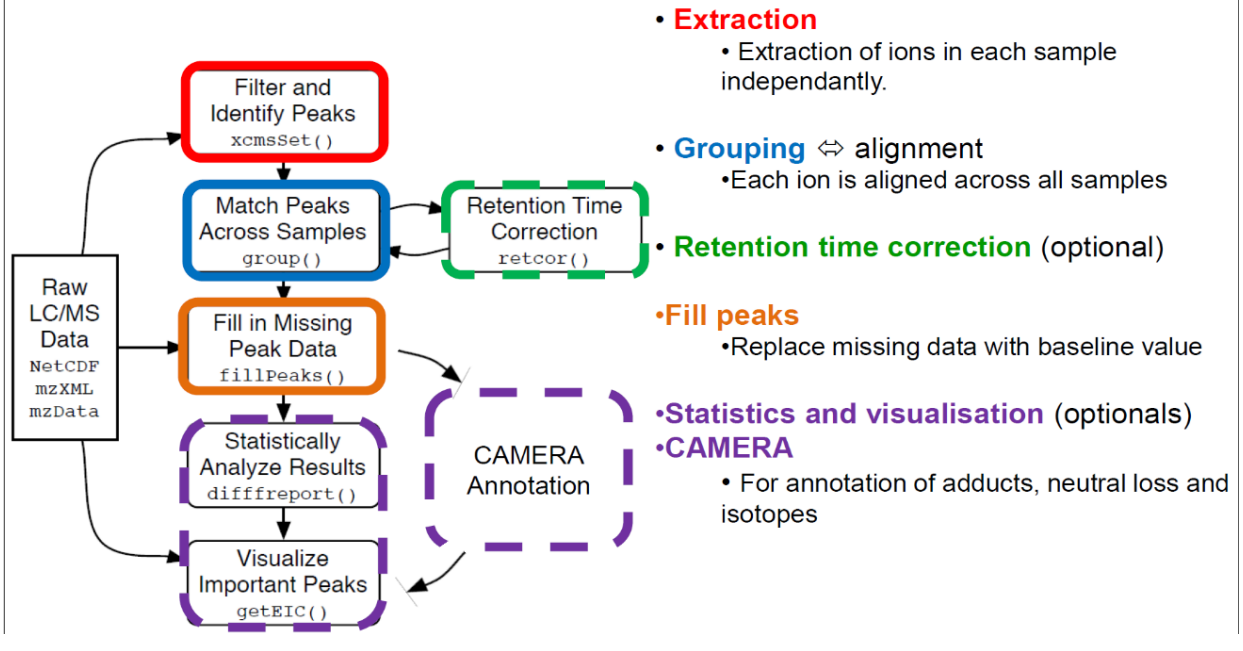


SESIÓN PRÁCTICA. VISUALS_7

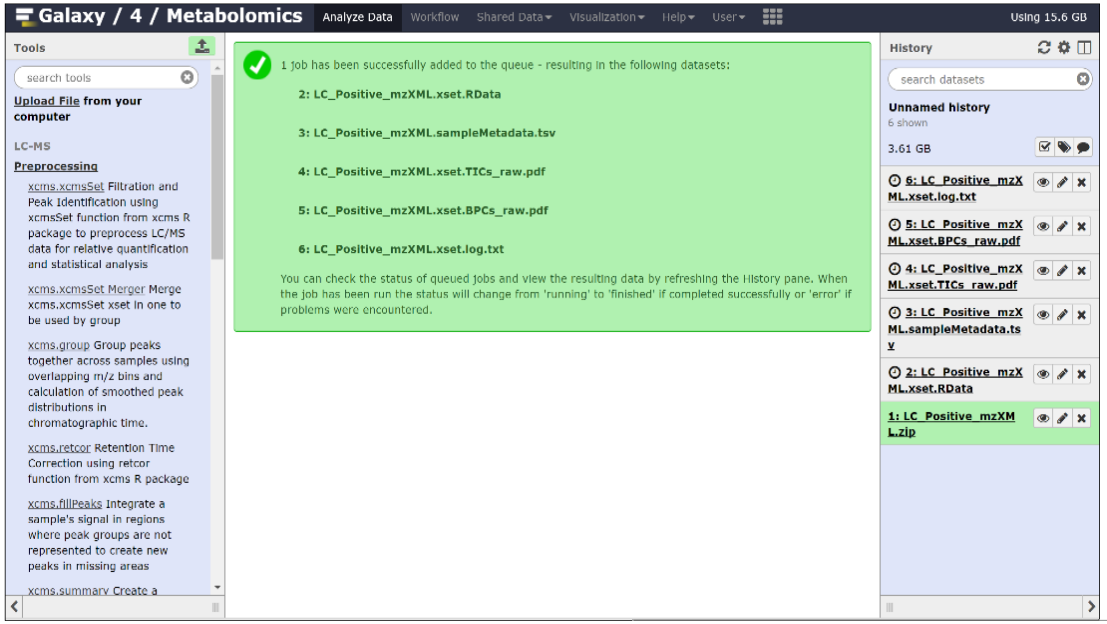


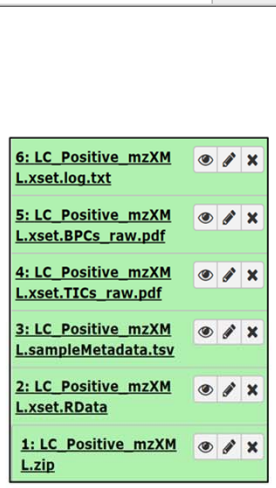
SESIÓN PRÁCTICA. VISUALS_8





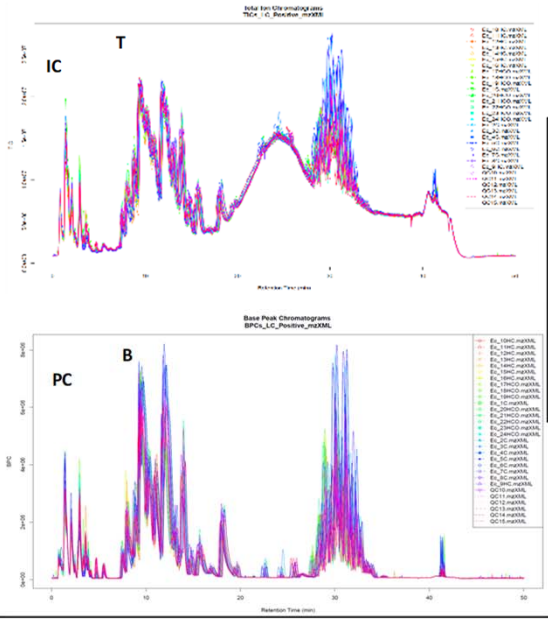
SESIÓN PRÁCTICA. VISUALS_9

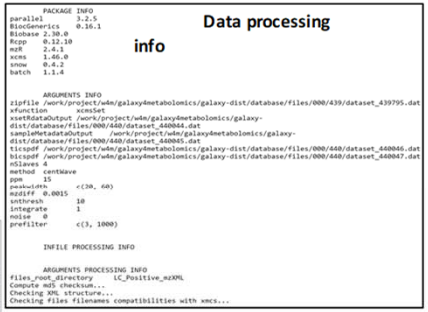




sample

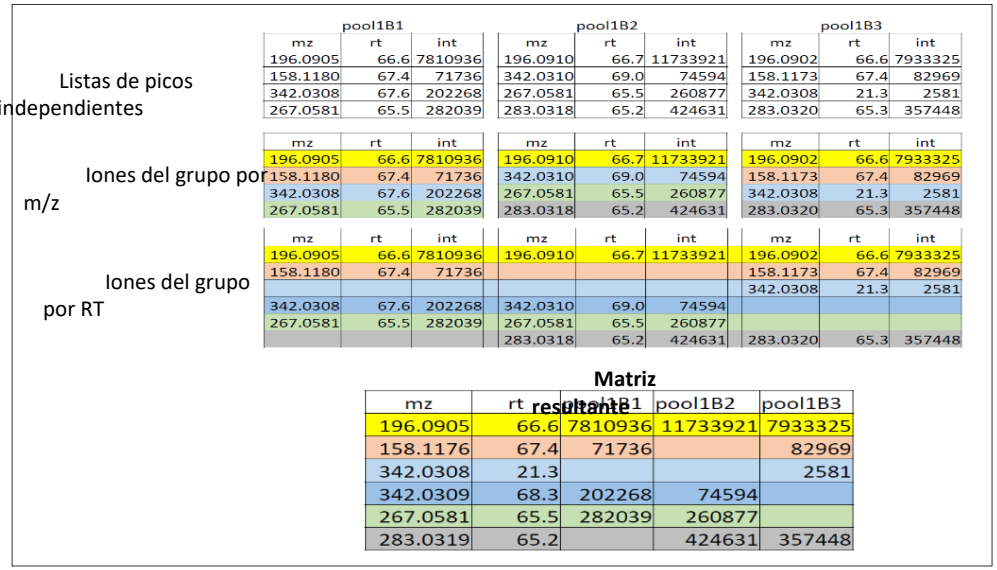
1 Metadata		3
sampleMetadata	class	polarity
Ec_10HC	LC_Positive_mzXML	positive
Ec_11HC	LC_Positive_mzXML	positive
Ec_12HC	LC_Positive_mzXML	positive
Ec_13HC	LC_Positive_mzXML	positive
Ec_14HC	LC_Positive_mzXML	positive
Ec_15HC	LC_Positive_mzXML	positive
Ec_16HC	LC_Positive_mzXML	positive
Ec_17HCO	LC_Positive_mzXML	positive
Ec_18HCO	LC_Positive_mzXML	positive
Ec_19HCO	LC_Positive_mzXML	positive
Ec_1C	LC_Positive_mzXML	positive
Ec_20HCO	LC_Positive_mzXML	positive
Ec_21HCO	LC_Positive_mzXML	positive
Ec_22HCO	LC_Positive_mzXML	positive
Ec_23HCO	LC_Positive_mzXML	positive
Ec_24HCO	LC_Positive_mzXML	positive
Ec_2C	LC_Positive_mzXML	positive
Ec_3C	LC_Positive_mzXML	positive
Ec_4C	LC_Positive_mzXML	positive
Ec_5C	LC_Positive_mzXML	positive
Ec_6C	LC_Positive_mzXML	positive
Ec_7C	LC_Positive_mzXML	positive
Ec_8C	LC_Positive_mzXML	positive
Ec_9HC	LC_Positive_mzXML	positive
QC10	LC_Positive_mzXML	positive
QC11	LC_Positive_mzXML	positive
QC12	LC_Positive_mzXML	positive
QC13	LC_Positive_mzXML	positive
QC14	LC_Positive_mzXML	positive
QC15	LC_Positive_mzXML	positive





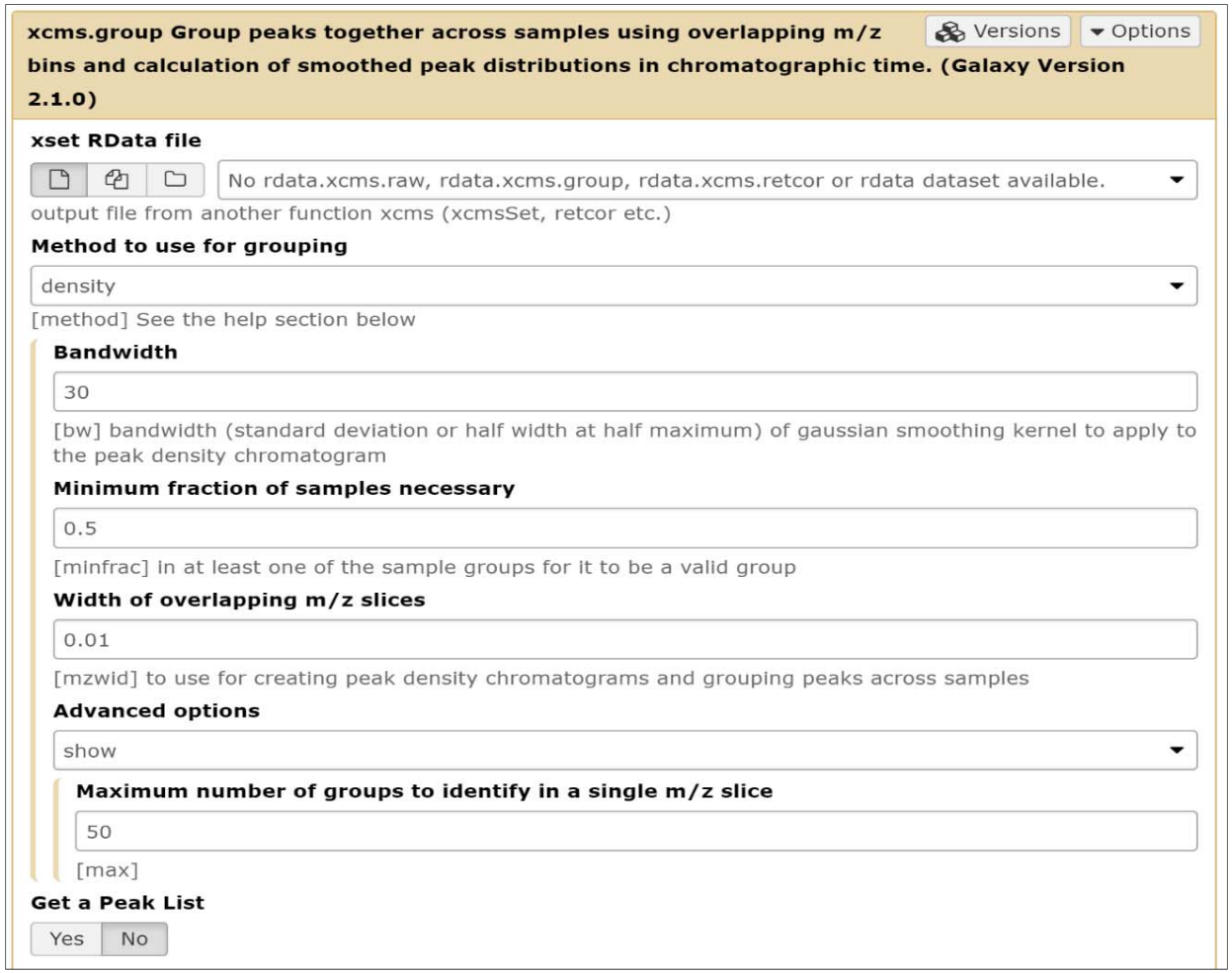
SESIÓN PRÁCTICA. VISUALS_10

xcms.group Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.



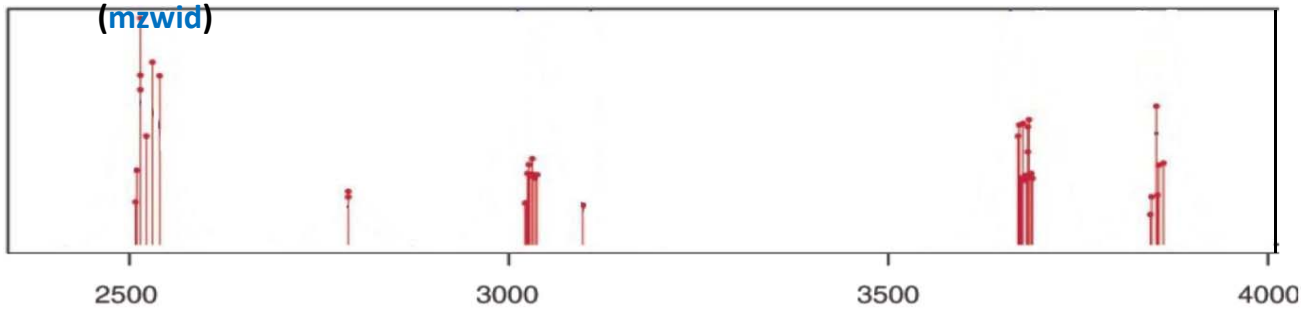
Parameter : num + label	Format
Or : RData file | rdata.xcms.raw	
Or : RData file | rdata.xcms.retcor	

SESIÓN PRÁCTICA. VISUALS_11

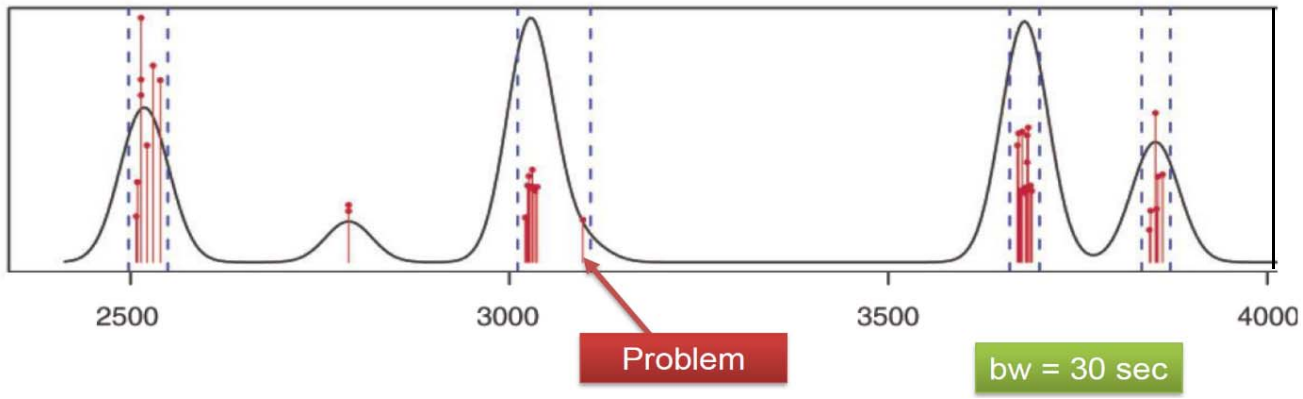


SESIÓN PRÁCTICA. VISUALS_12

Agrupación de picos en *bin* de masa: 337.975 – 338.225 m/z (mzwid)

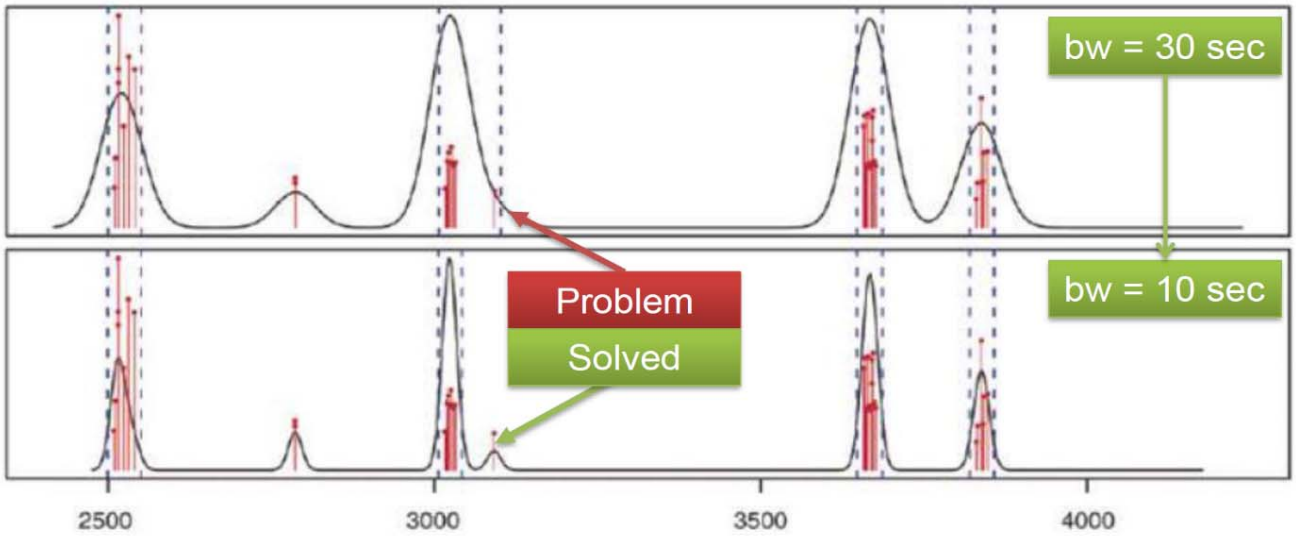


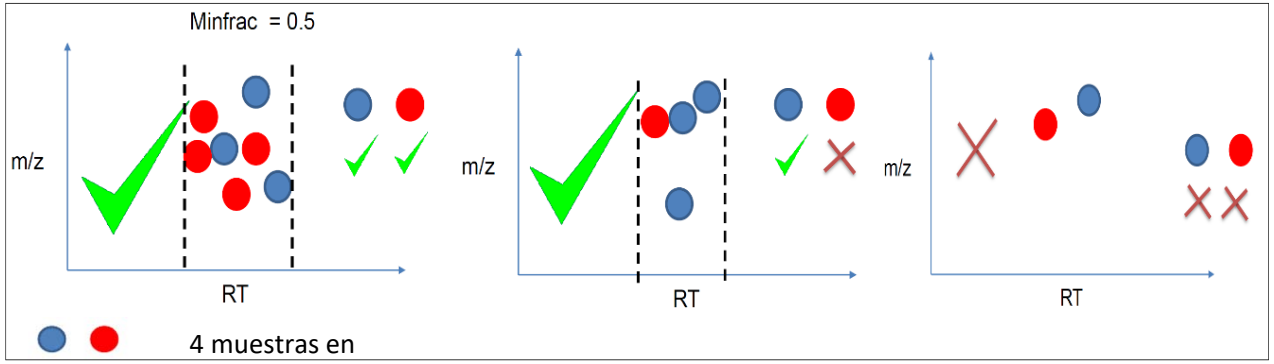
Agrupación de picos en *bin* de masa: 337.975 – 338.225 m/z (mzwid)



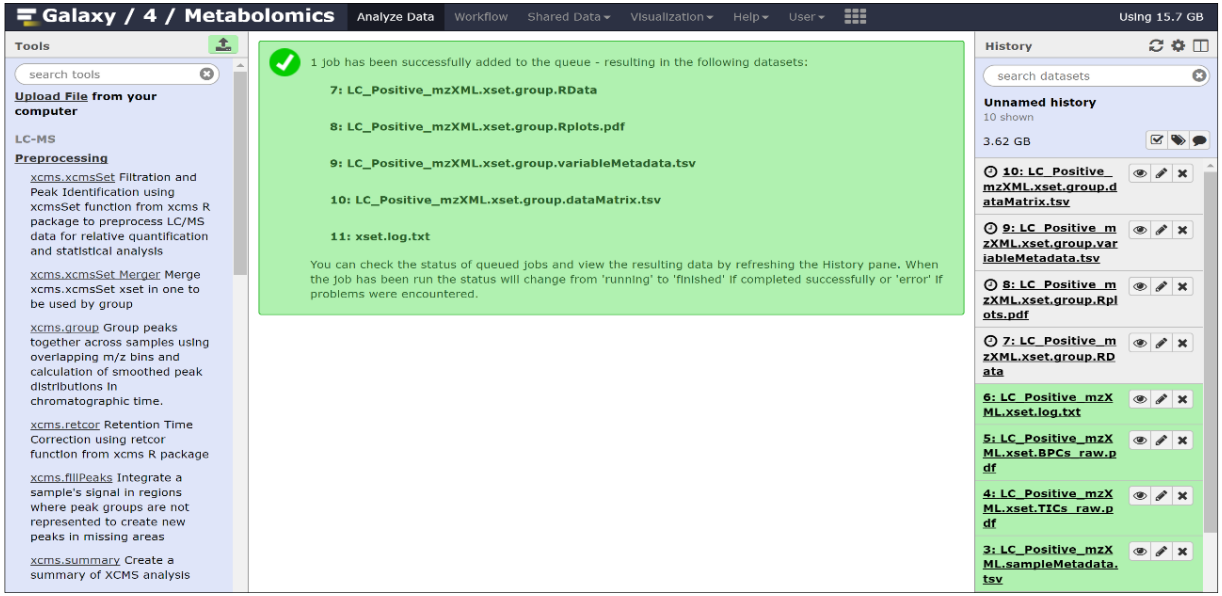
SESIÓN PRÁCTICA. VISUALS_13

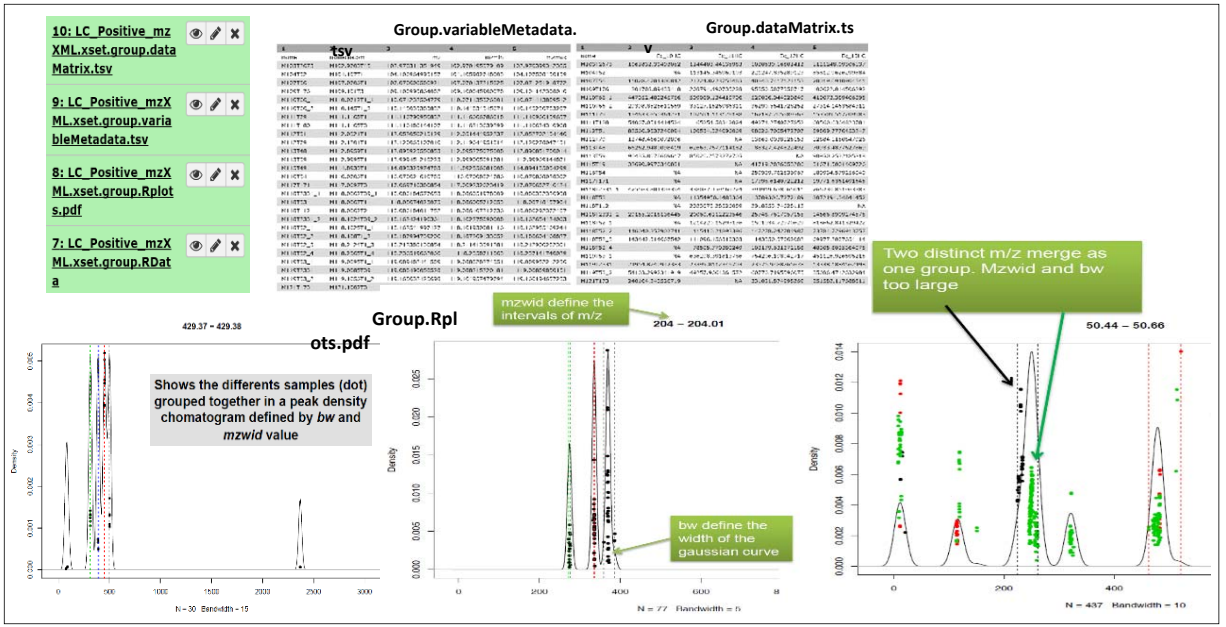
Agrupación de picos en *bin* de masa: 337.975 – 338.225 m/z (mzwid)





SESIÓN PRÁCTICA. VISUALS_14

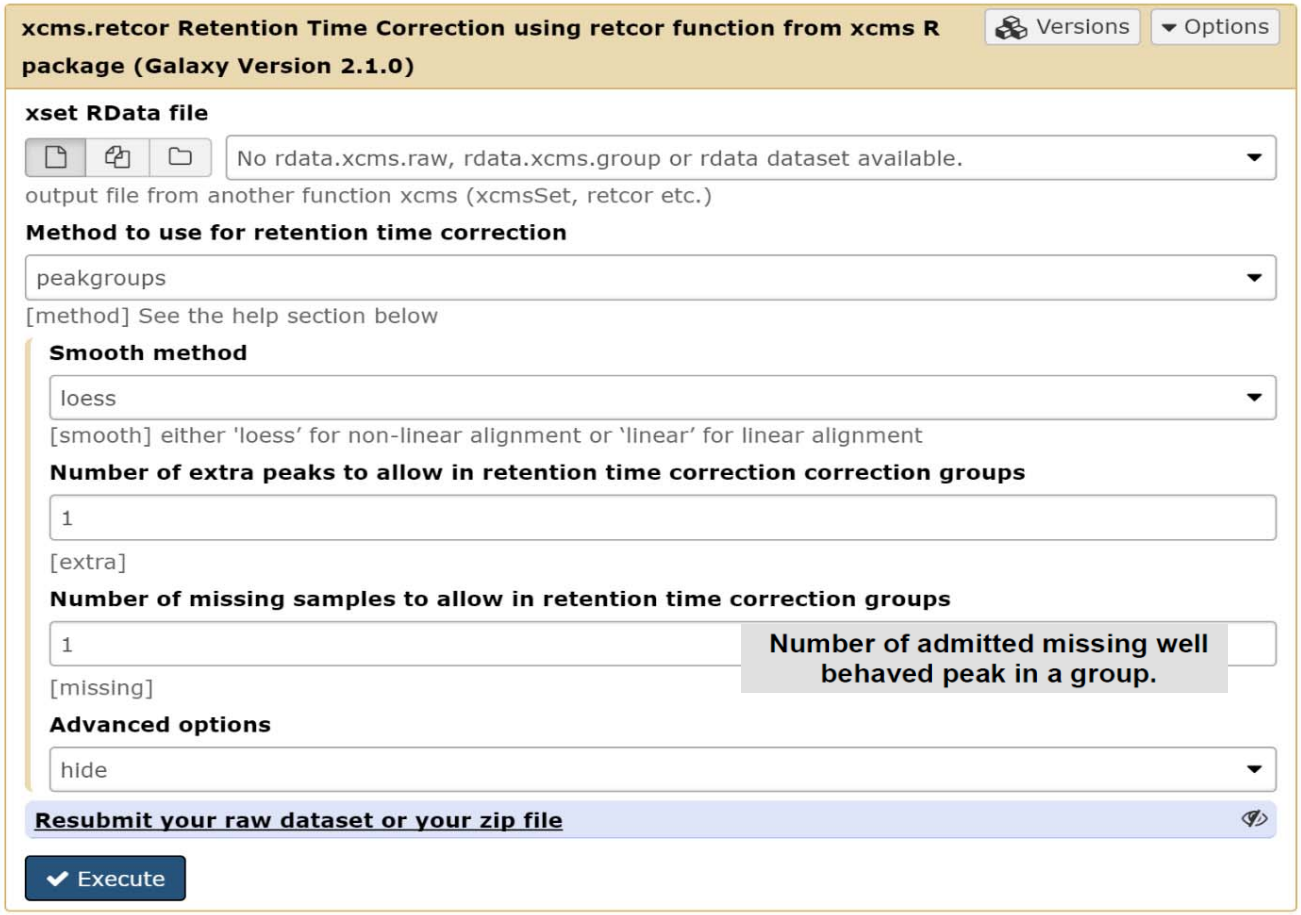




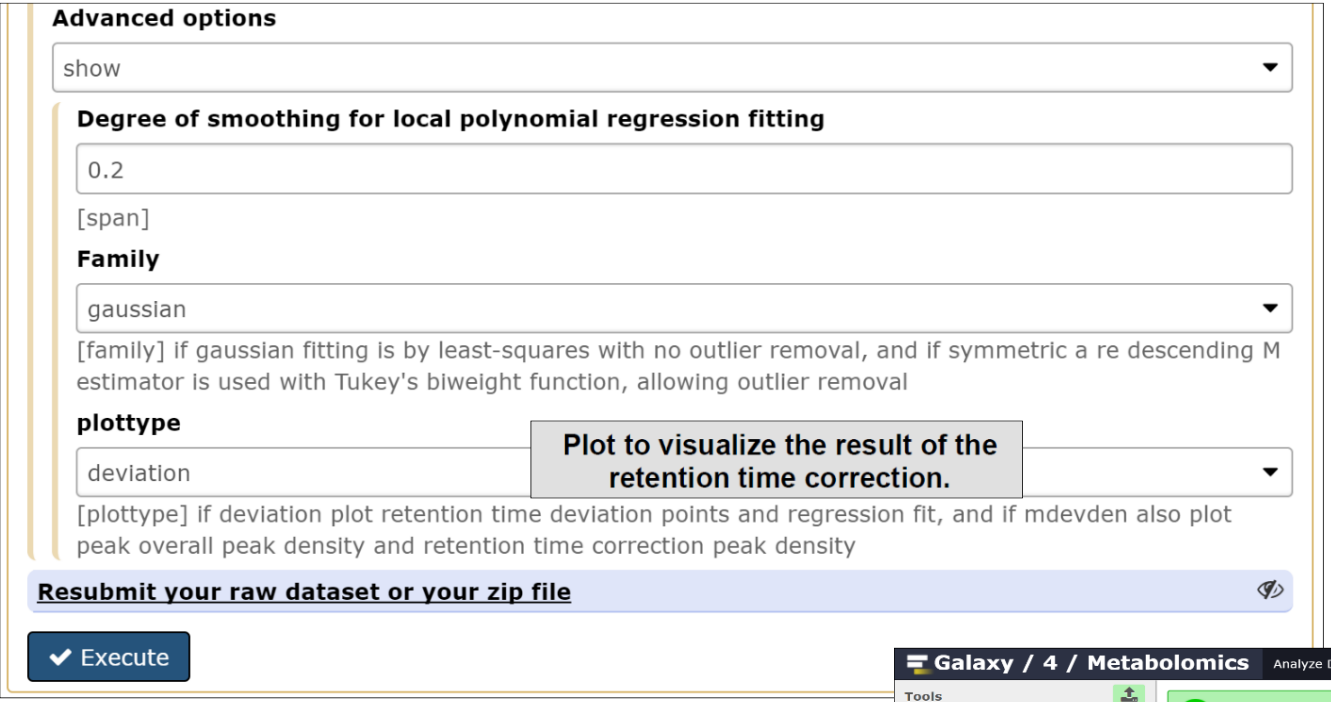
SESIÓN PRÁCTICA. VISUALS_15

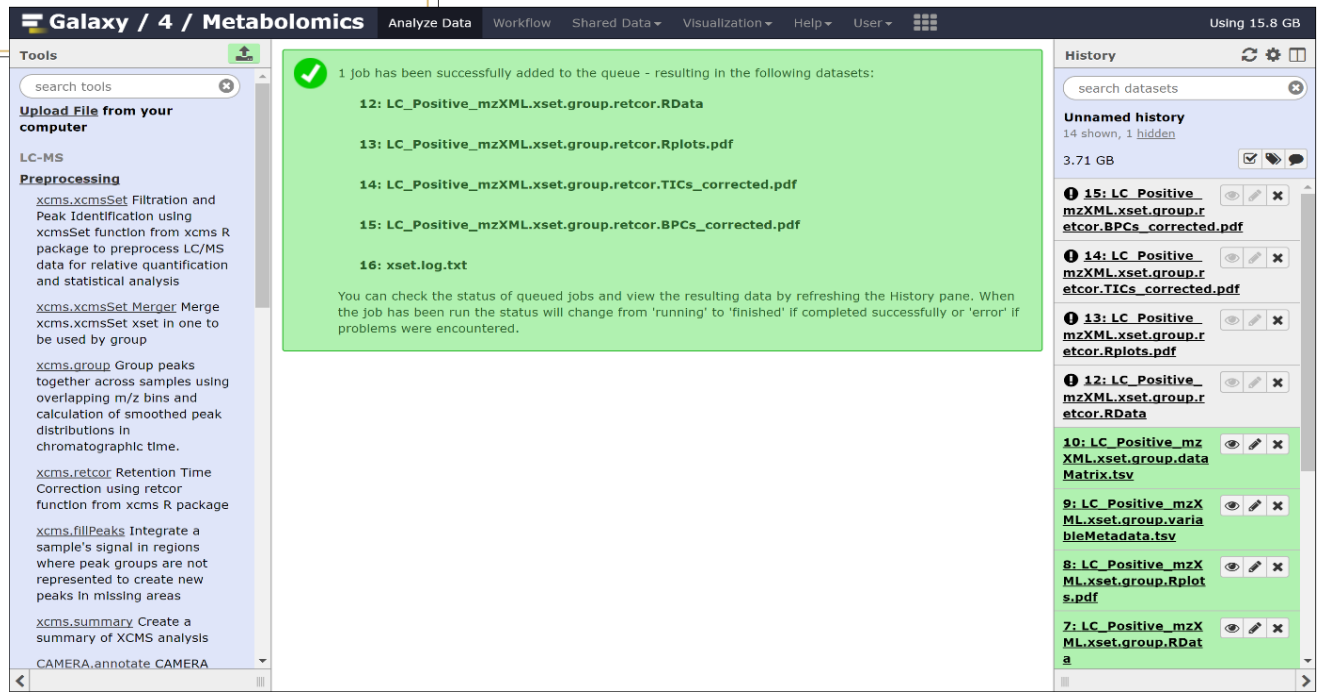
xcms.retcor Retention Time Correction using retcor function from xcms R package

Parameter : num + label	Format
1 : RData file	rdata.xcms.group

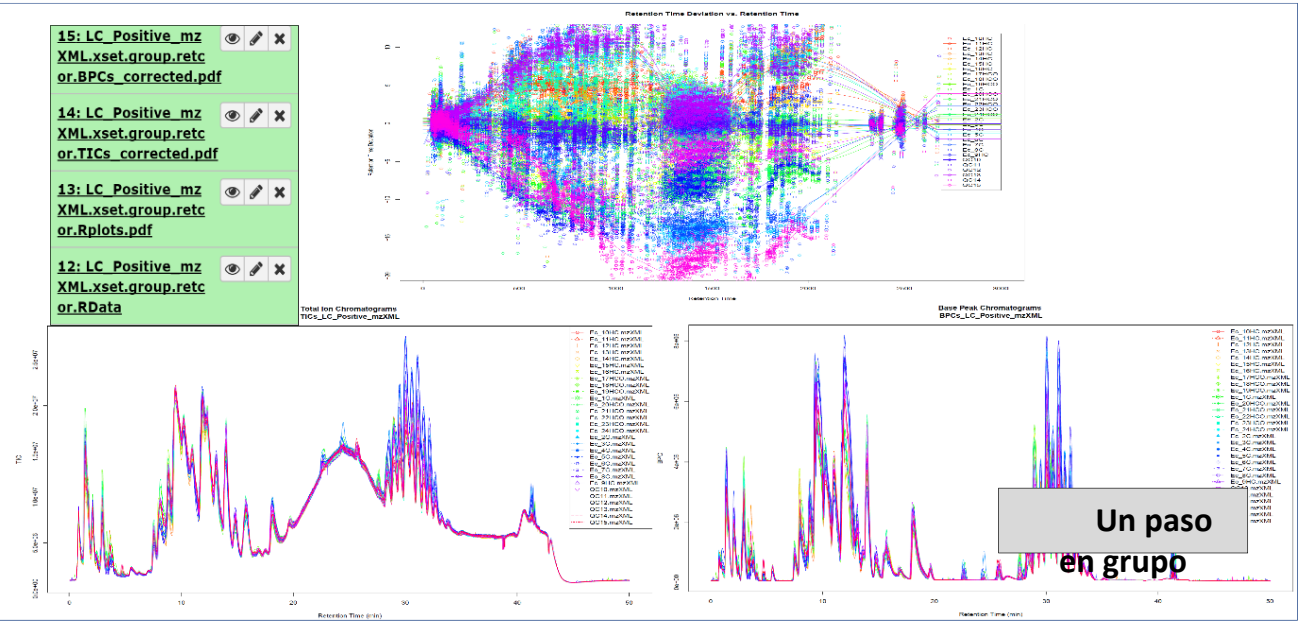


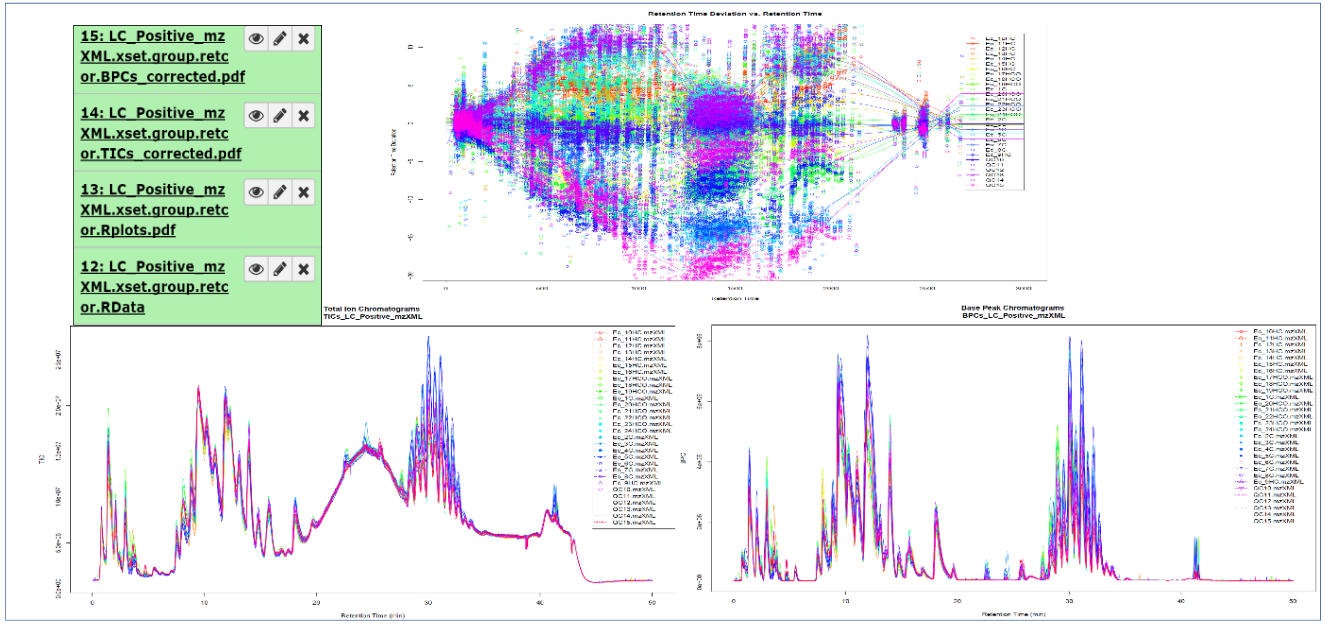
SESIÓN PRÁCTICA. VISUALS_16



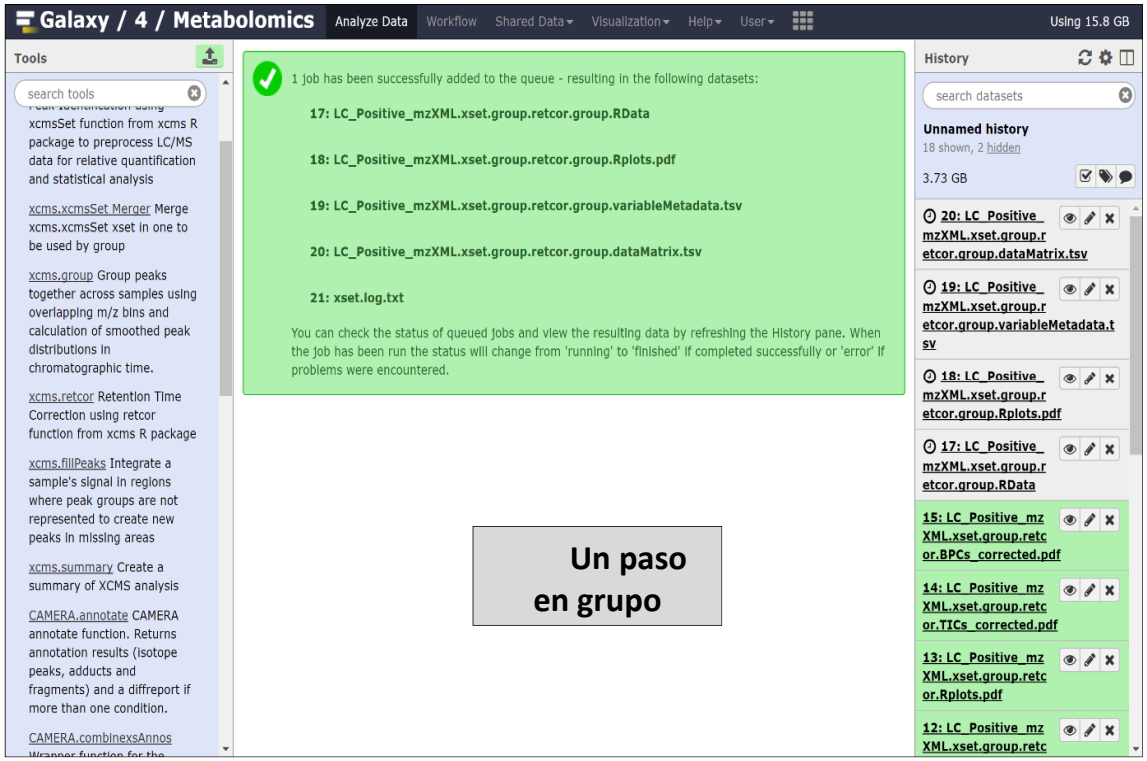


SESIÓN PRÁCTICA. VISUALS_17





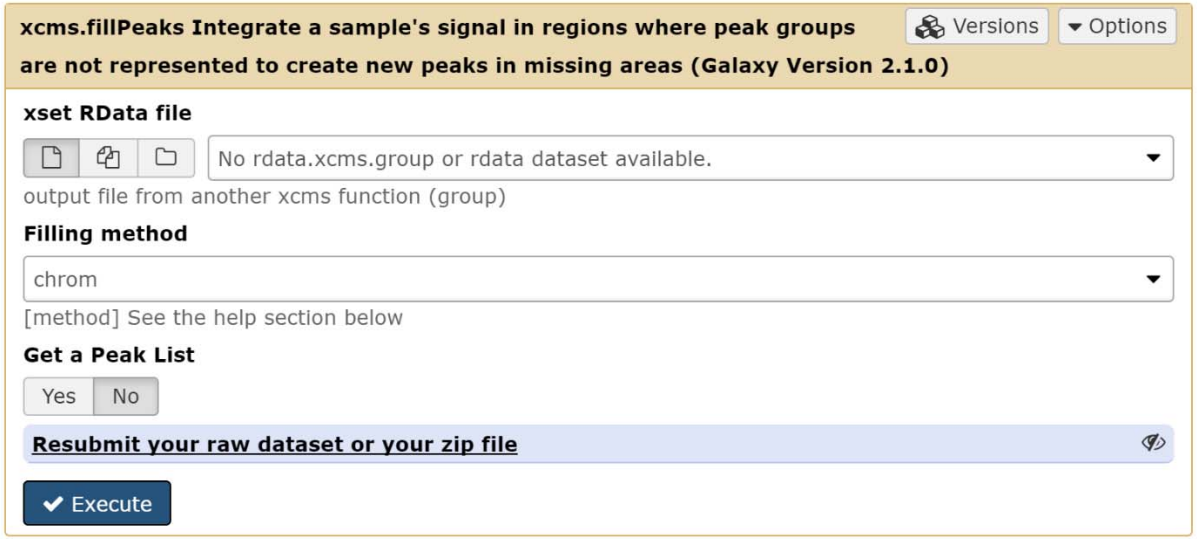
SESIÓN PRÁCTICA. VISUALS_18



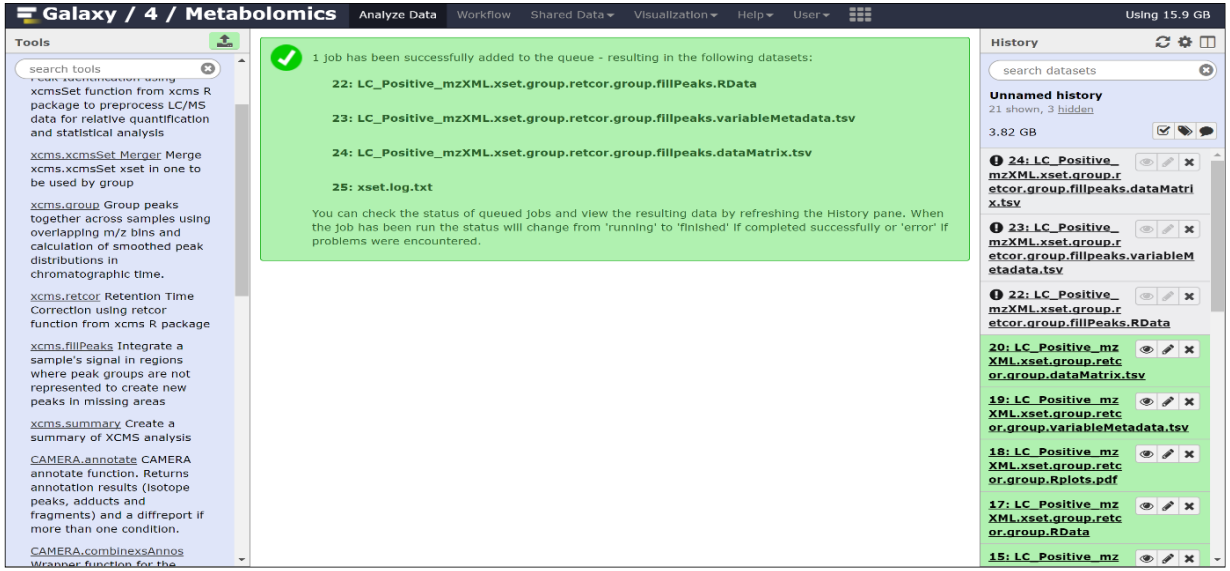
Un paso en grupo

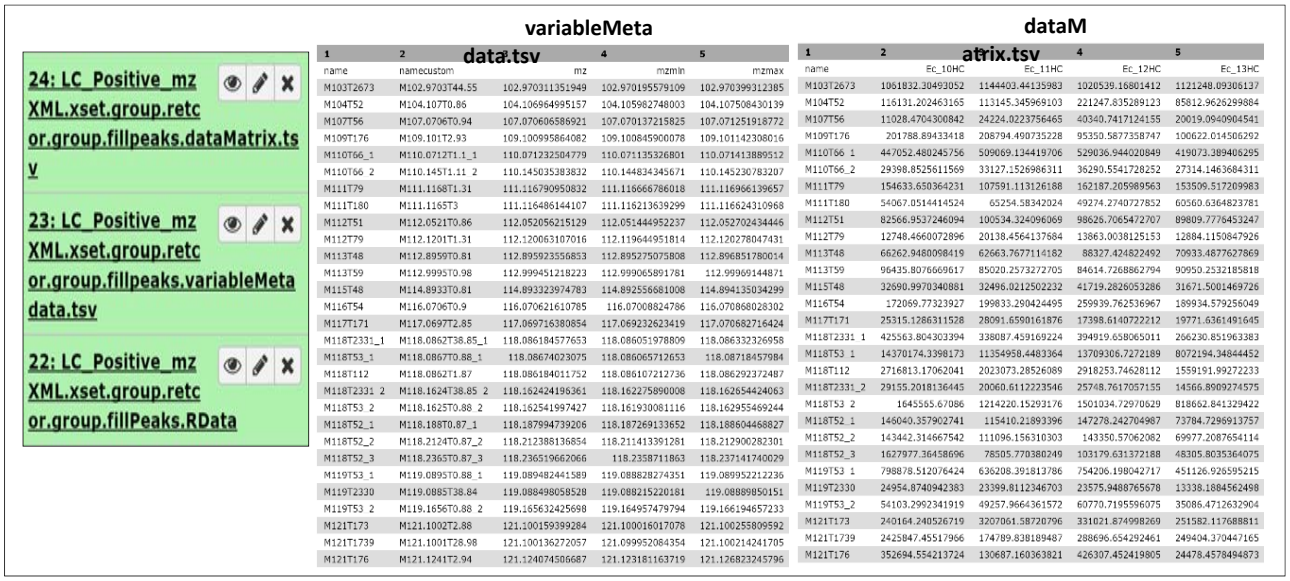
xcms.fillPeaks Integrate a sample's signal in regions where peak groups are not represented to create new peaks in missing areas

Parameter : num + label	Format
1 : RData file	rdata.xcms.group

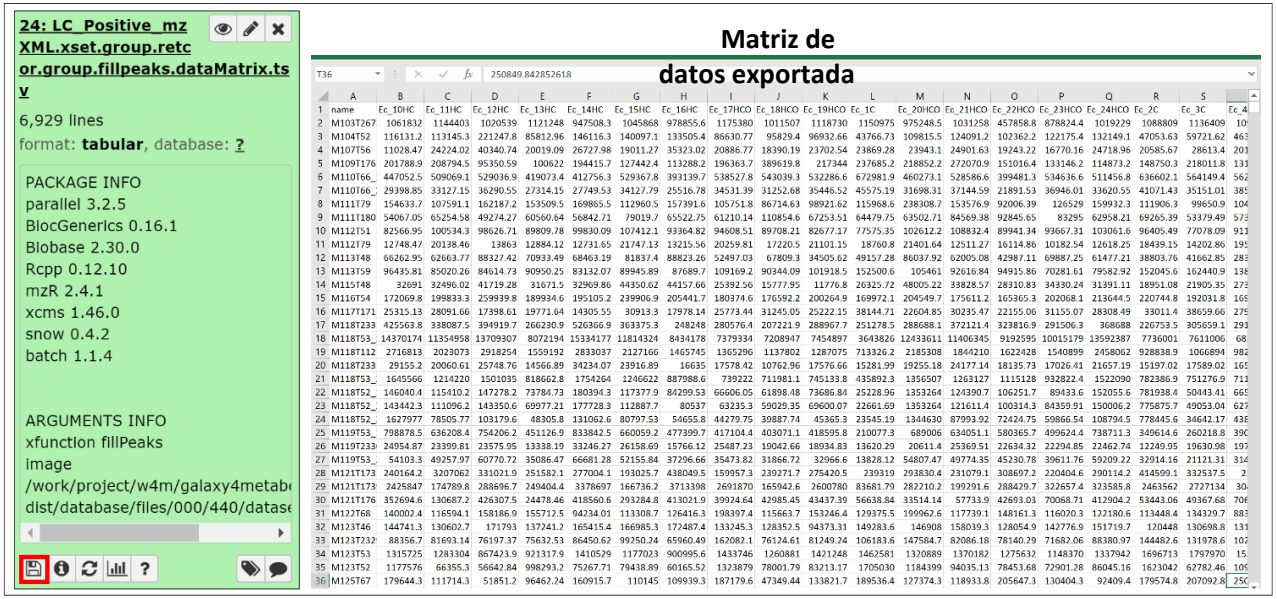


SESIÓN PRÁCTICA. VISUALS_19

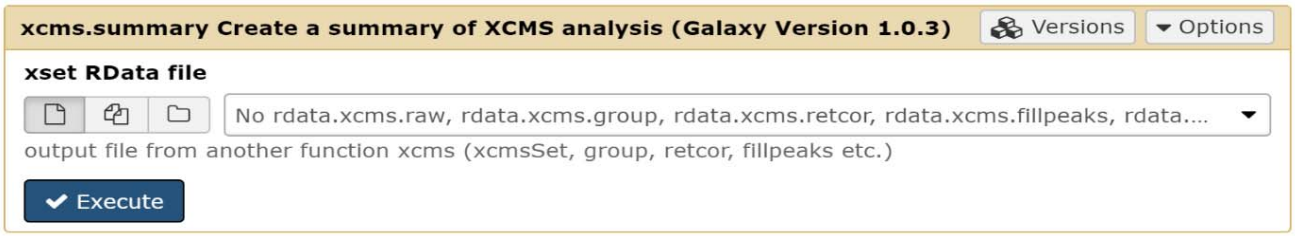




SESIÓN PRÁCTICA. VISUALS_20



xcms.summary Create a summary of XCMS analysis



Este proyecto ha sido financiado con ayuda de la Comisión Europea.

AACLifeSci

Erasmus+

CEU Universidad San Pablo