

Establishment of a quality control mixture for benchmarking LC-MS based dereplication protocols in NP research

Miwa Kubo

**School of Pharmaceutical Sciences, University of Geneva,
Switzerland**



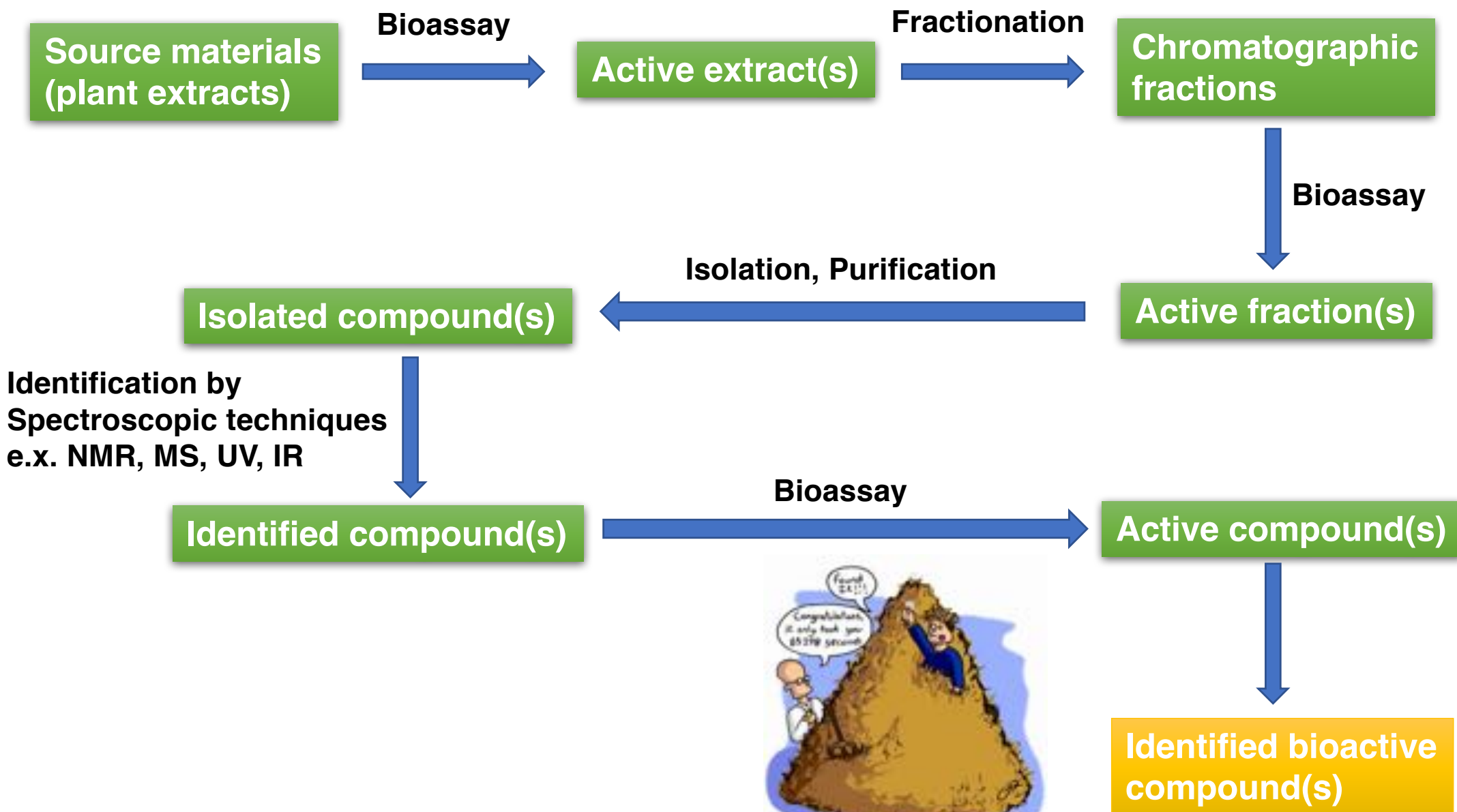


Prof. Jean-Luc Wolfender

**Section des sciences pharmaceutiques
University of Geneva - University of Lausanne
Phytochemistry and bioactive natural products**

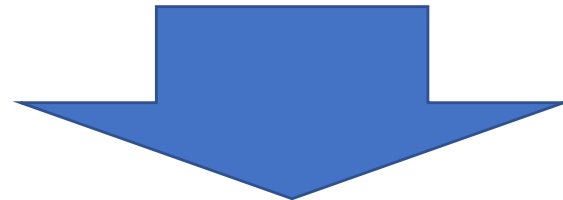
He is currently developing innovative MS- and NMR-based **metabolomics strategies** in the frame of projects related to phytochemistry, microbial interactions and phytotherapy. He is specialised in the *de novo* structure identification of biomarkers at the microgram scale and is using a miniaturised approach that combines activity-based HPLC profiling and high content information bioassays such as those involving zebrafish.

Overview of a bioassay-guided natural product drug discovery process



Issue of classical method

Can not secure sufficient material
Lose activity
Isolation of known compound

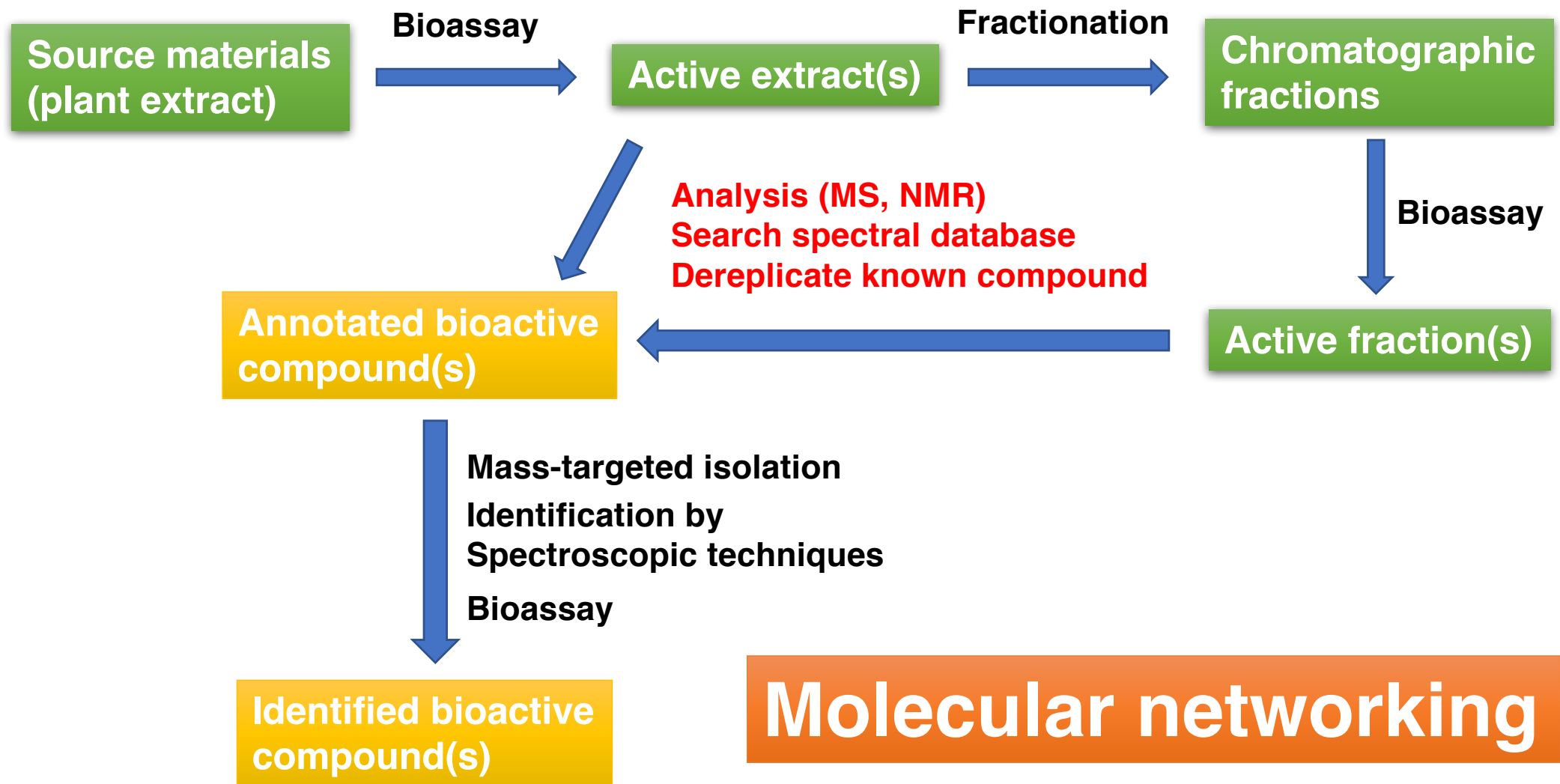


Metabolomics

Dereplication

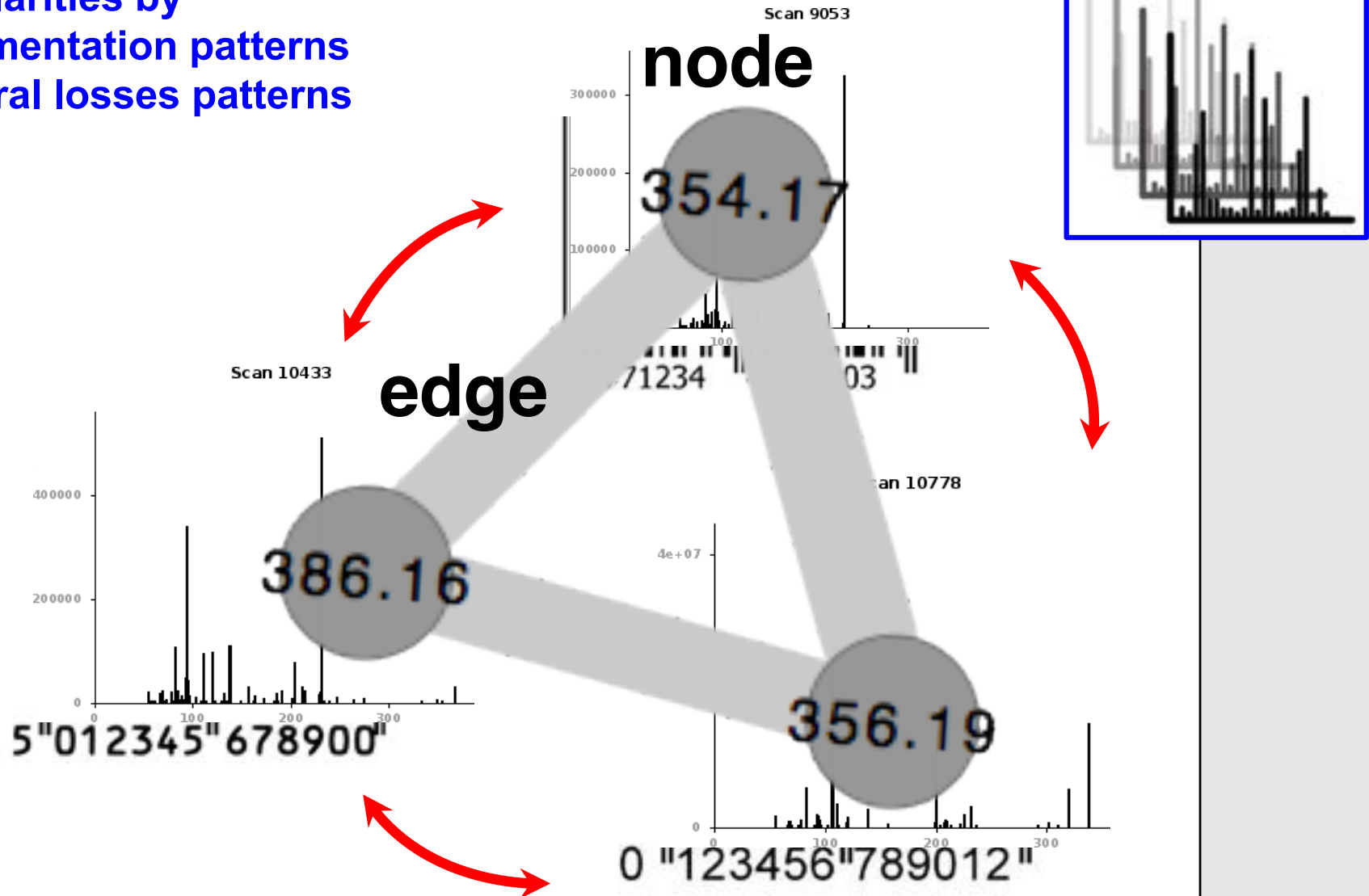
- The early identification of known compounds without time-consuming manual structure elucidation
- Spotting unknowns of interest could also enter in the definition of dereplication

Overview of a mass-guided natural product drug discovery process



MS/MS network for interpretation

Similarities by
fragmentation patterns
neutral losses patterns

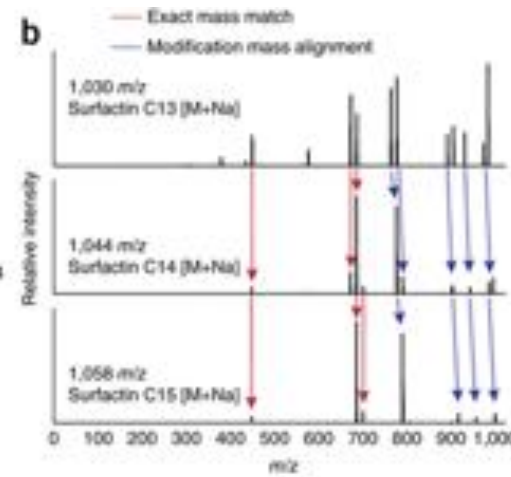
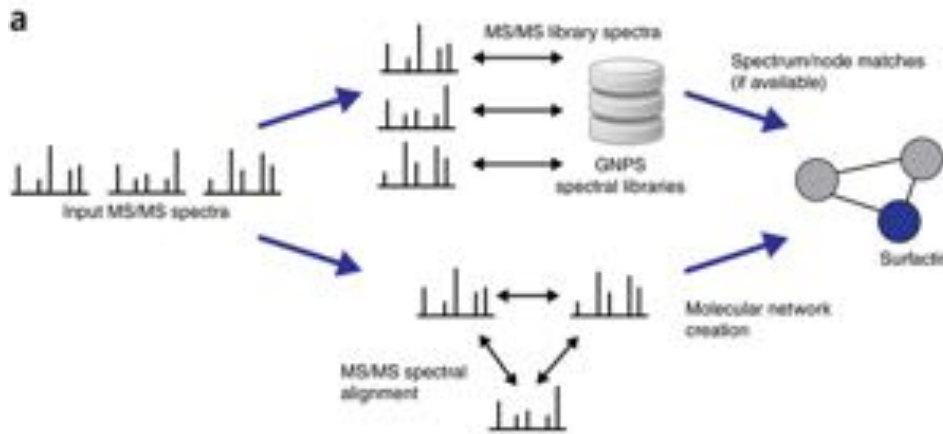
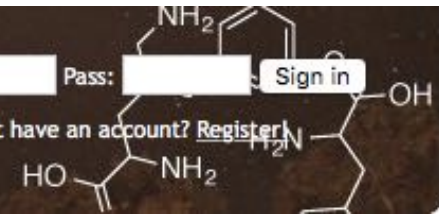


GNPS: Global Natural Products Social Molecular Networking

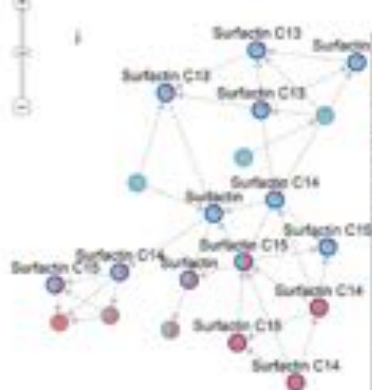
MassIVE Datasets | Documentation | Forum | Contact

User: Pass: [Sign in](#)

Don't have an account? [Register](#)



<https://gnps.ucsd.edu/>



Node Labels

- Cluster index
- parent mass
- LibraryID
- EventDate
- Peptide

Edge Labels

- cosine
- Default
- None

Edge Highlight

- Edge Score
- Edge Score 40

Node MS2 Peaks Highlight

- Library MS2
- User MS2

Cluster index 1881

parent mass 1021.05

LibraryID Surfactin C13

number of spectra 8

DefaultGroup 61.62

precursor charge 1

Peptide

RT info 66.65, n = 34.34

[Cluster/Spectra](#) [Custom Settings](#)

Align Spec

Score: 0.87

Starting

- G1
- G2
- G3
- G4
- G5
- G6

Node Size

Default

Node Color

Default

Draw Pts **Plot**

Cluster index 1884

parent mass 1053.25

LibraryID Surfactin C13

number of spectra 14

DefaultGroup 61.62

precursor charge 1

Peptide

RT info 61.65, n = 17.56

[Cluster/Spectra](#) [Custom Settings](#)

[View All Spec](#) [View Match](#)

Cluster index 1890

parent mass 1067.36

LibraryID Surfactin C13

number of spectra 28

DefaultGroup 61.62

precursor charge 1

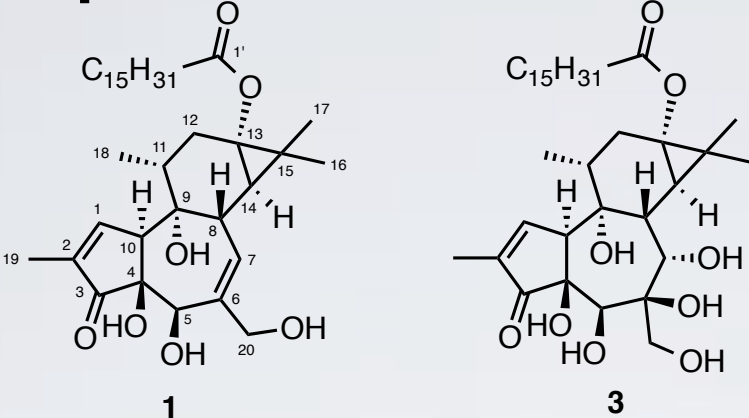
Peptide

RT info 127.62, n = 56.57

[Cluster/Spectra](#) [Custom Settings](#)

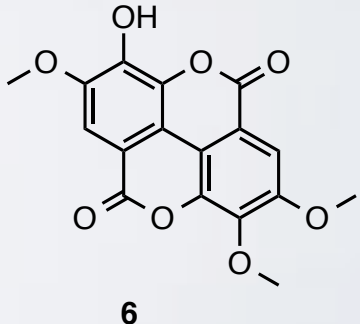
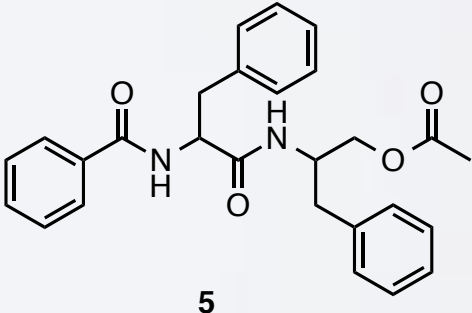
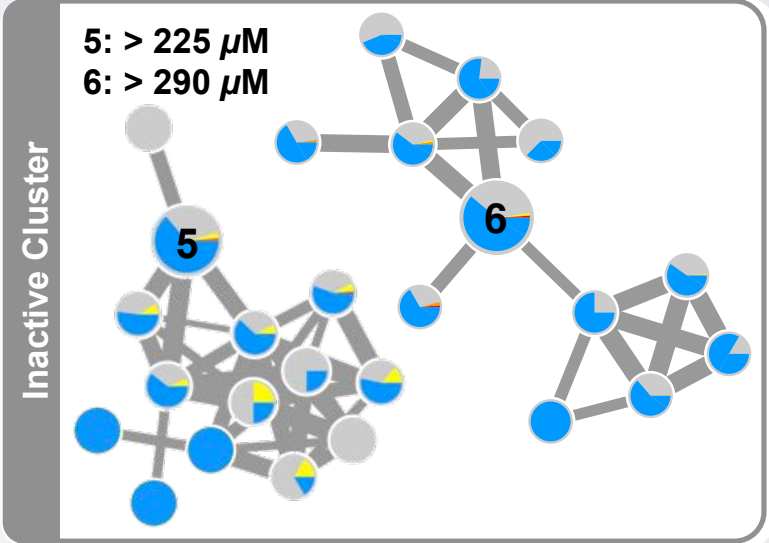
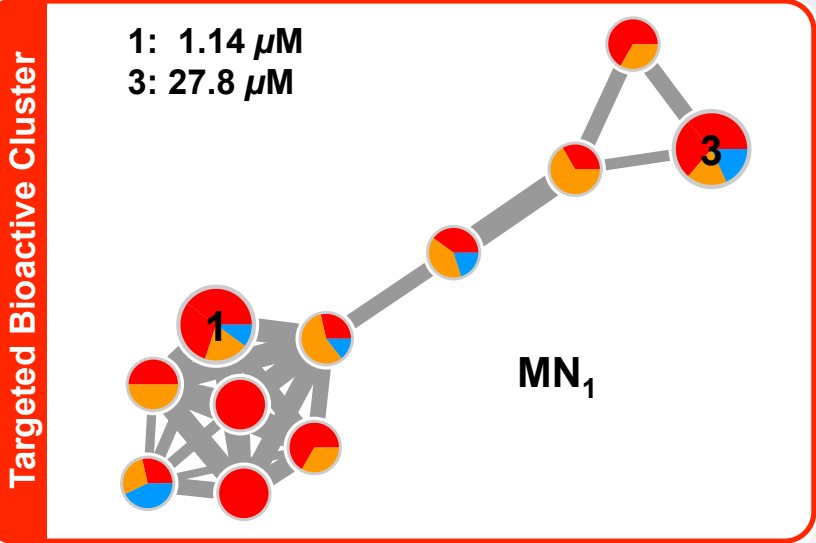
[View All Spec](#) [View Match](#)

Bio-chemical input



Extracts IC₅₀

- 0 - 5 $\mu\text{g.mL}^{-1}$
- 5 - 20 $\mu\text{g.mL}^{-1}$
- 20 - 50 $\mu\text{g.mL}^{-1}$
- > 50 $\mu\text{g.mL}^{-1}$
- n.d.



Number of natural products (NPs)

- existing ~unknown
- known ~300000
- with associated MSMS data ~25000



Number of natural products (NPs)

Dereplication

Informatics
methods

- existing ~unknown
- known ~300000
- with associated MSMS data ~25000



MAGMa



Generation of the **ISDB** (*in-silico* database)

Universal Natural Products Database



250000 entries



170000 SMILES



340 files of 500 SMILES

HPC cluster Baobab

Baobab is a High Performance Computing (HPC) cluster available to all UNIGE's researchers.



Baobab

~2400 compute cores
(Intel Sandy Bridge @ 2.20GHz)
10TB of RAM

Baobab
login node

cfm-id



12 hours



metadata



UNPD-ISDB

> 170 000 *in silico*
MS/MS spectra

Generation of the **ISDB** (*in-silico* database)



263446 entries



221771 SMILES + metadata



455 files of 500 SMILES

HPC cluster Baobab

Baobab is a High Performance Computing (HPC) cluster available to all UMIGE's researchers.



Baobab

~2400 compute cores
(Intel Sandy Bridge @ 2.20GHz)
10TB of RAM



cfm-id



12 hours



metadata



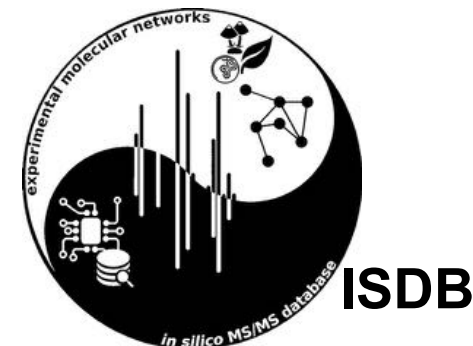
DNP-ISDB

> 220 000 *in silico*
MS/MS spectra

Number of natural products (NPs)

Dereplication

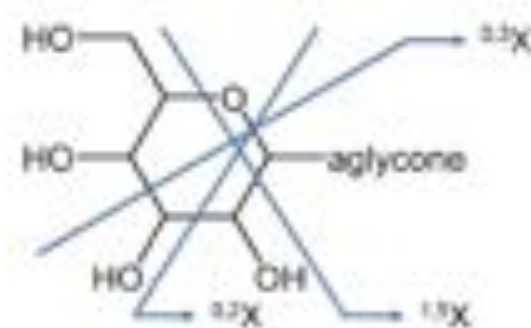
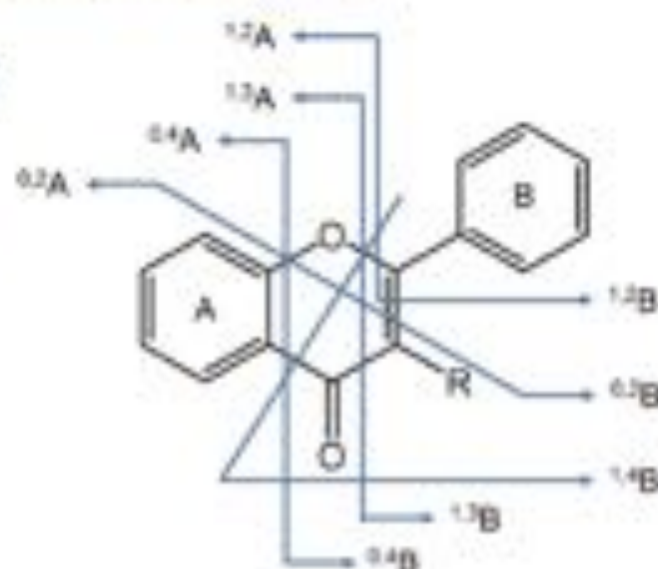
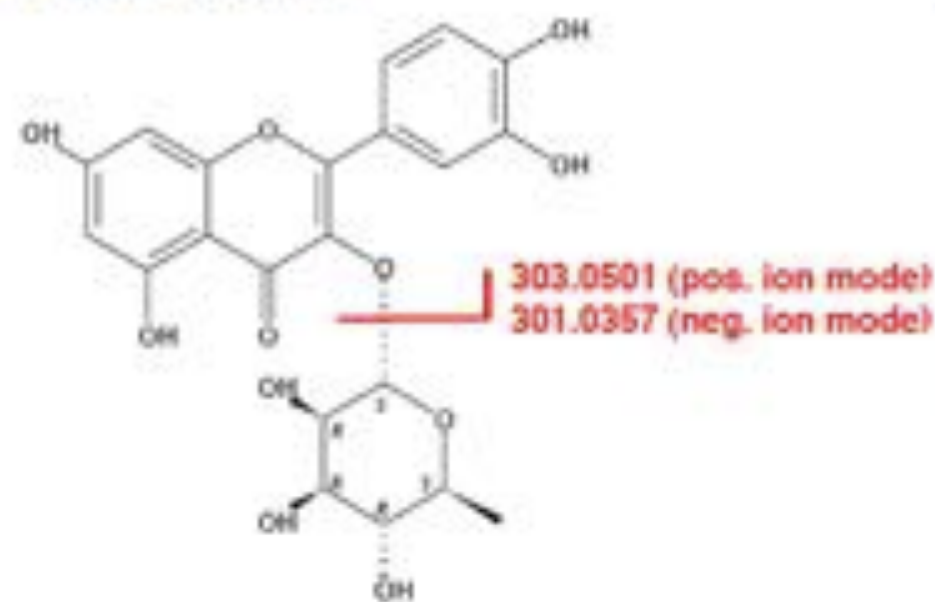
- existing ~unknown
- known ~300000
- with associated MSMS data ~25000



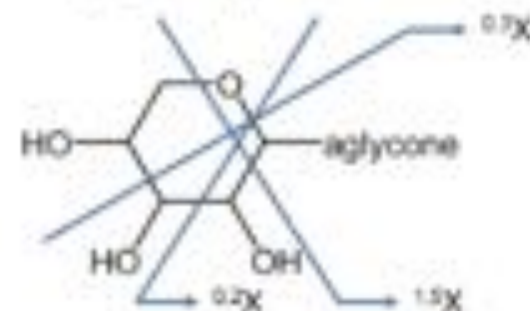
Fragmentation of flavonoids

- typical fragmentation of flavonoids and flavonoid-glycosides

- Quercetin



1,2X: $[M \pm H - 120]^+$
1,3X: $[M \pm H - 90]^+$
1,4X: $[M \pm H - 134]^+$



1,2X: $[M \pm H - 50]^+$
1,3X: $[M \pm H - 60]^+$
1,4X: $[M \pm H - 104]^+$



Collaboration with Pierre Fabre laboratories in the frame of their Open Innovation program

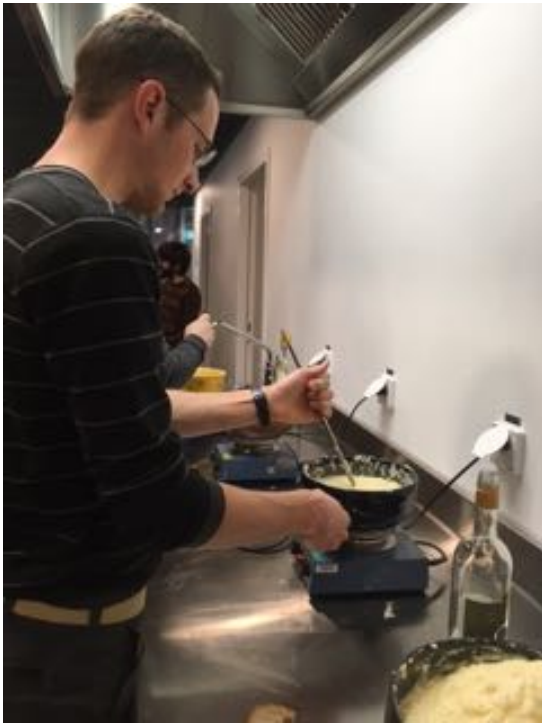


<http://nature-open-library-pierre-fabre.force.com/index>

A screenshot of the Nature Open Library website. The background is a blurred image of purple flowers. In the top left corner, there is a logo for "NATURE OPEN LIBRARY" with the tagline "AN OPEN INNOVATION INITIATIVE BY PIERRE FABRE". In the top right corner, there is the "Pierre Fabre" logo and a "Login" link. The main text in the center reads "Pierre Fabre Laboratories open their 15 000 plant library." followed by "Let's explore the potential of plants for Life Sciences." and a green "Contact us" button.

Objective:

- establish a molecular cartography of the PF library in order to prioritize high added values molecules (specific structures or bioactivities)**
- shift from classical bioguided fractionation to mass-directed approaches**



Methodology

Thermo Fisher
Orbitrap Q Exactive
Focus Hybrid MS



Waters
ACQUITY UPLC I-Class
PLUS



ANALYSIS

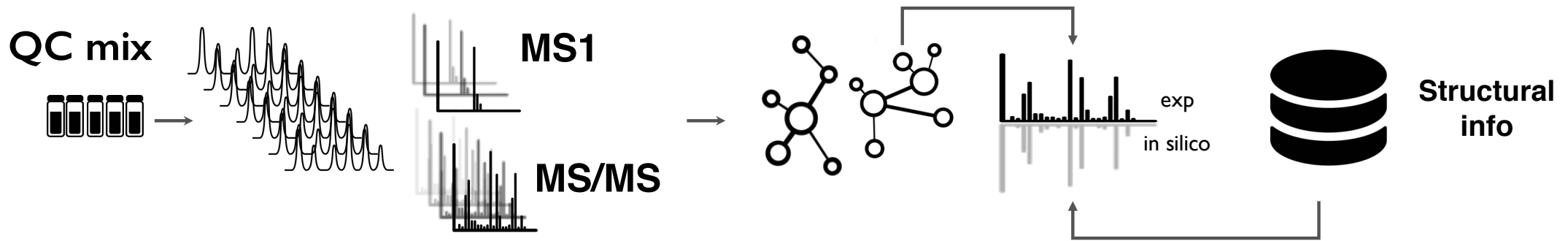


METABOLITE
PROFILING

Quality Control (QC)

The QC allow the user to evaluate and benchmark their LC and LC-MS chromatography system before analysis of critical analytes.

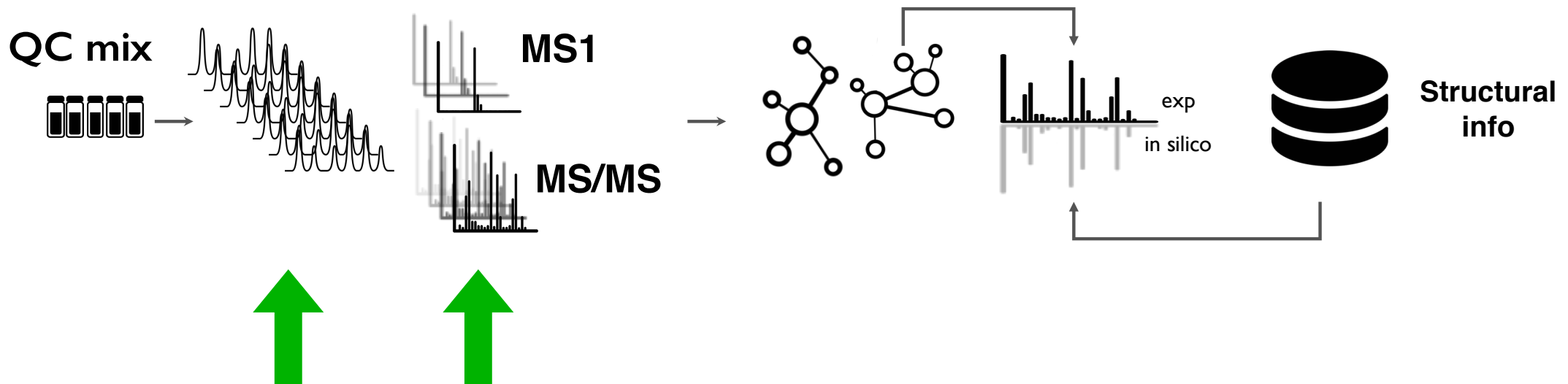
Objective:
Establish a universal QC mix for the quality evaluation of NPs profiling



Objective:

Establish a universal QC mix for the quality evaluation of NPs profiling

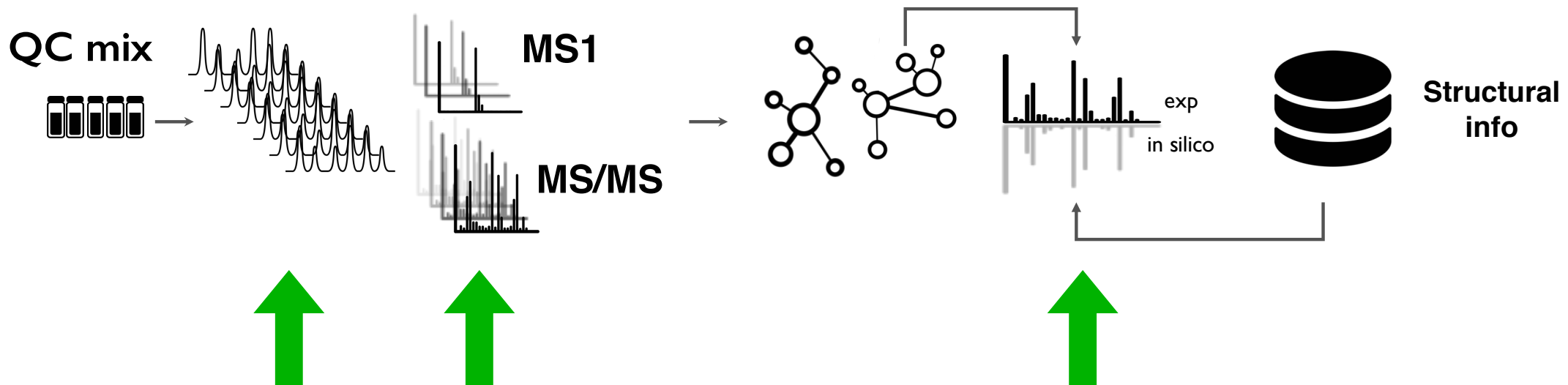
- at the analytical level (chromatography and MS acquisition)



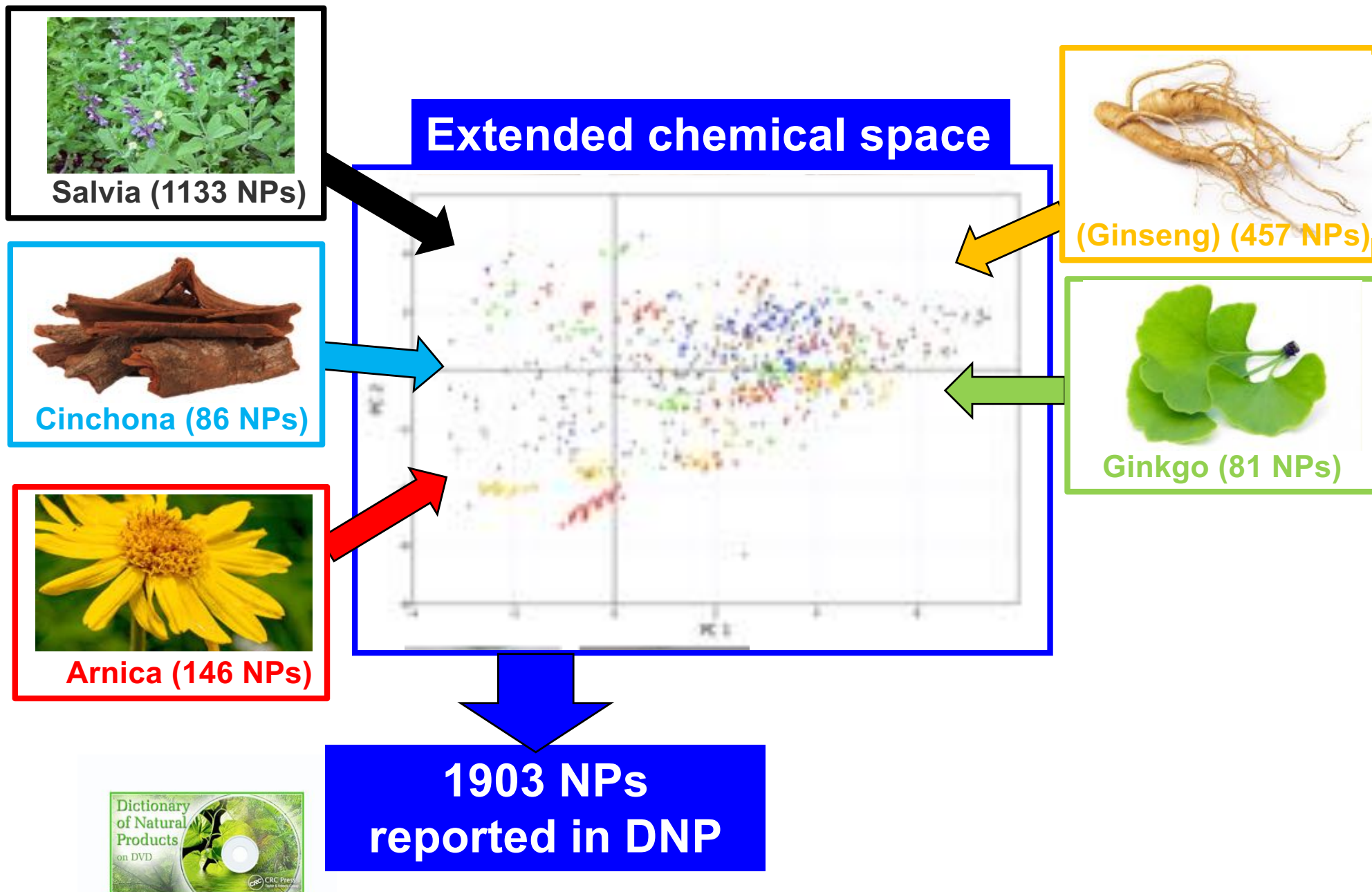
Objective:

Establish a universal QC mix for the quality evaluation of NPs profiling

- at the analytical level (chromatography and MS acquisition)
- at the data treatment and analysis level (metabolite identification)

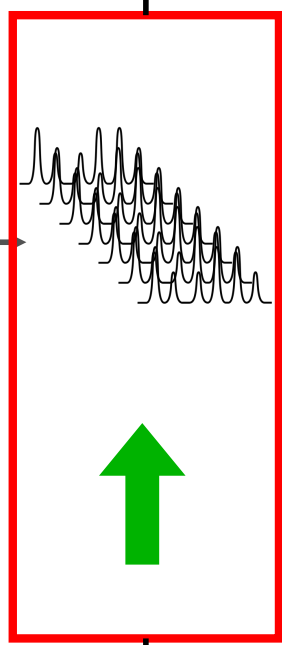


Chemical space of secondary metabolites of the QC-mix of 5 herbs

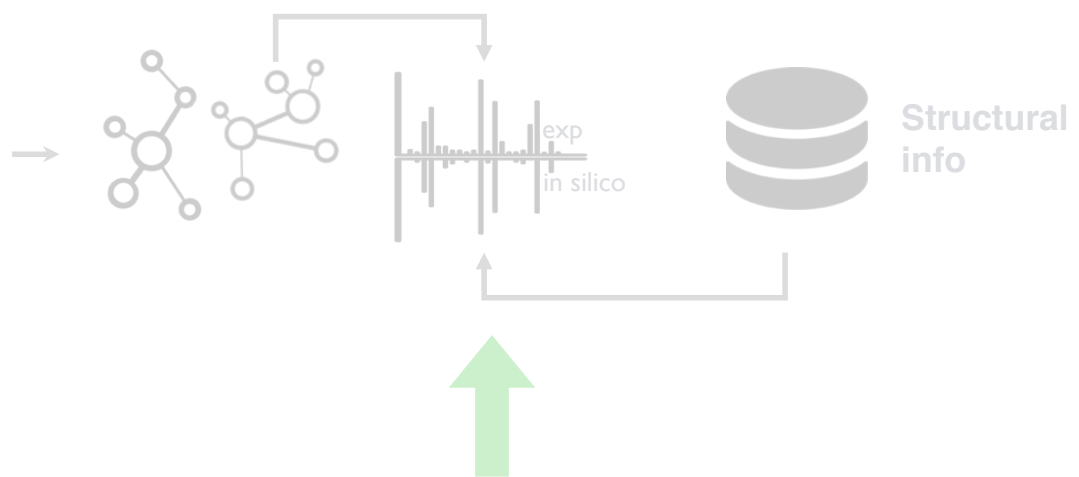


Method
UPLC, C18 BEH Acquity

QC mix

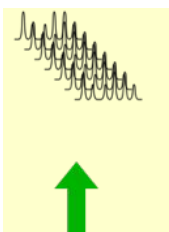
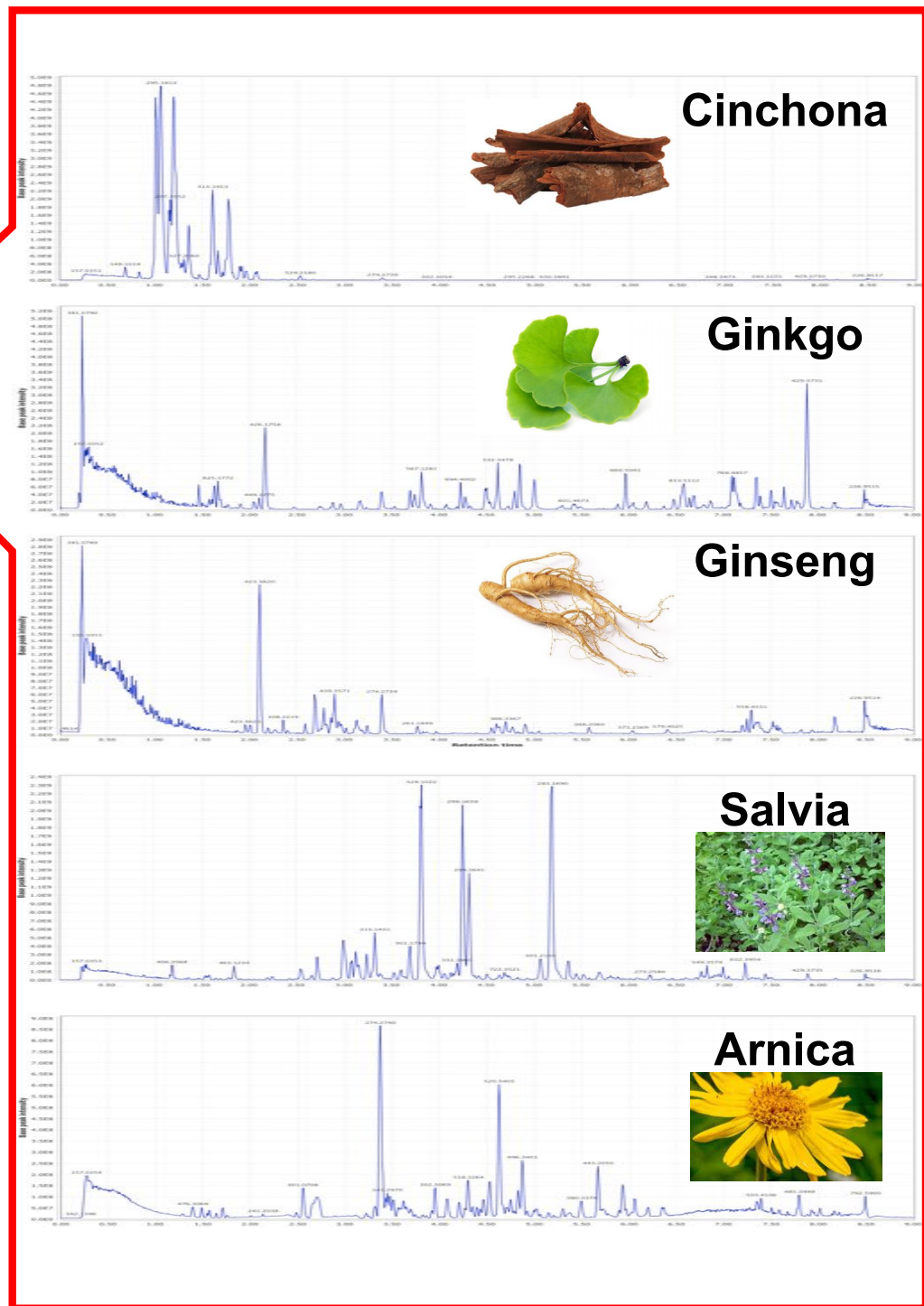
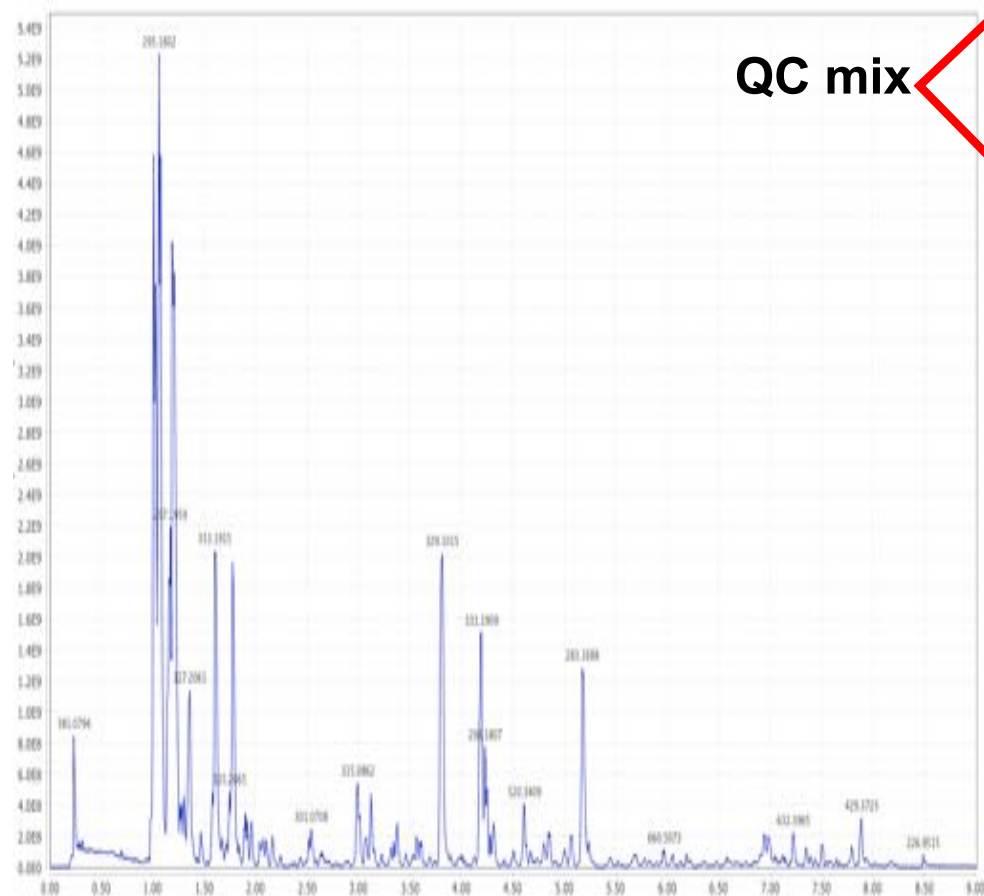



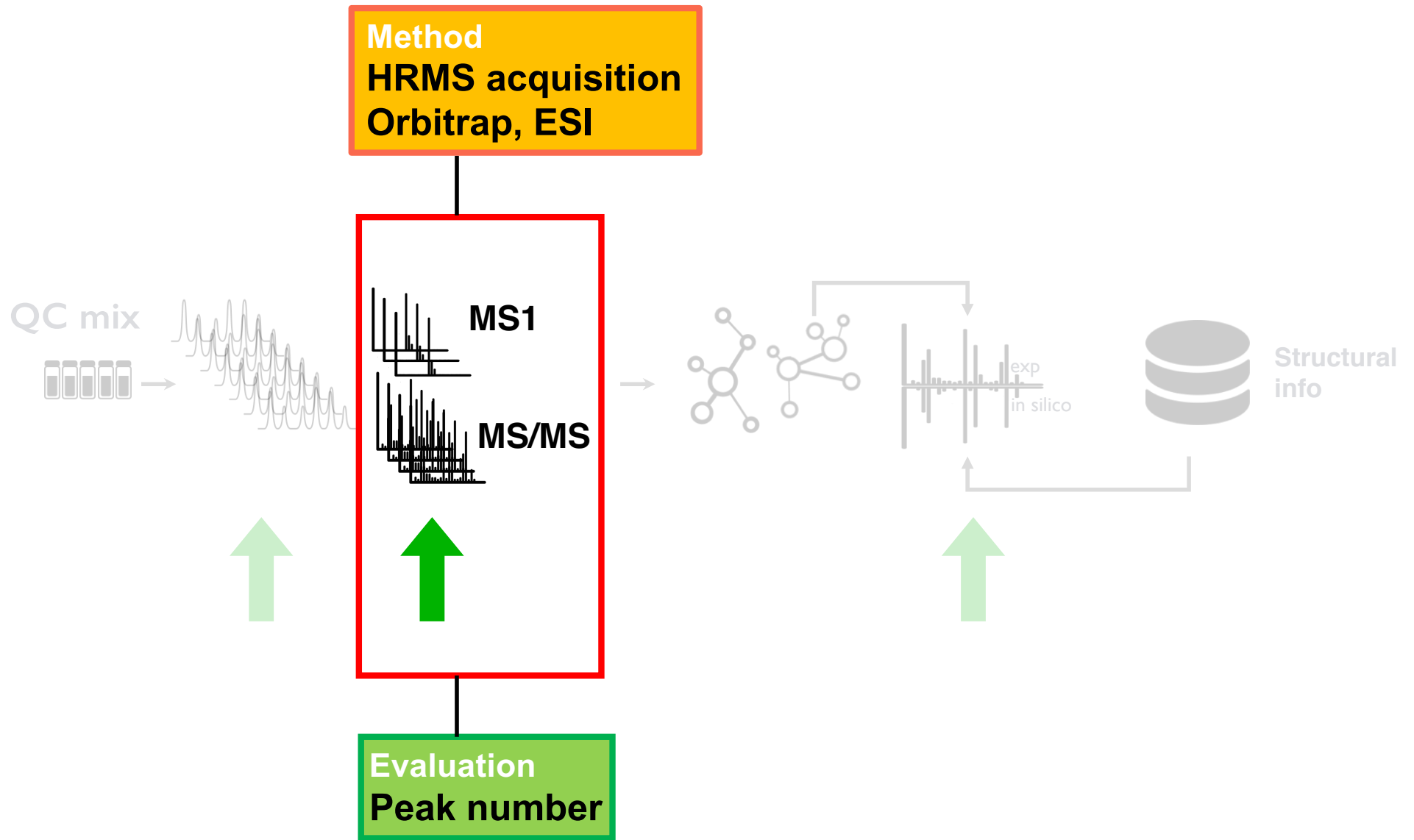
Evaluation
Peak capacity



QC mix?

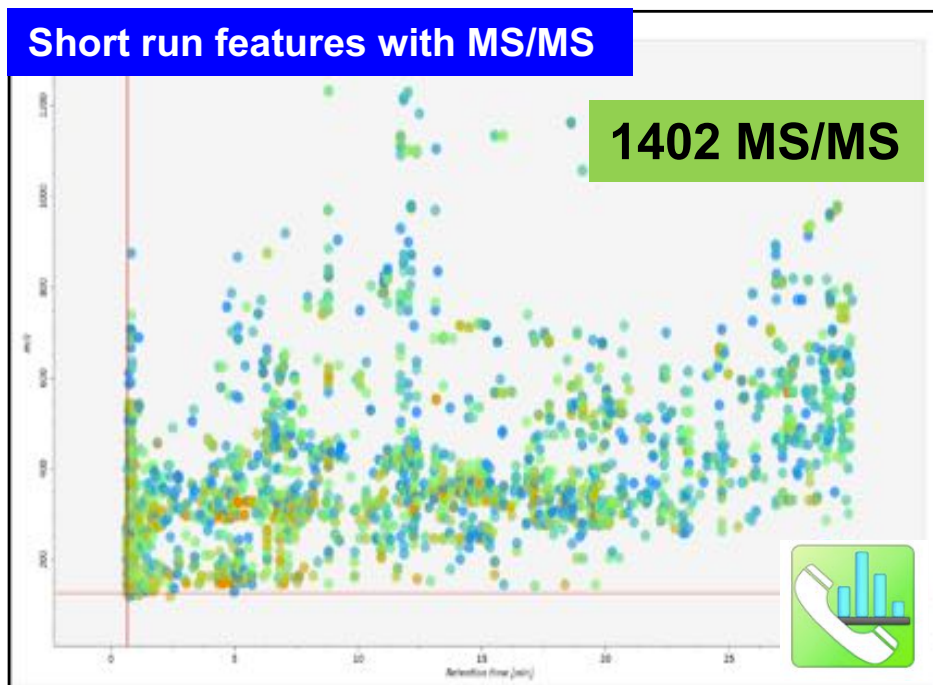
UHPLC-HRMS profiling of the QC mix and 5 herbs



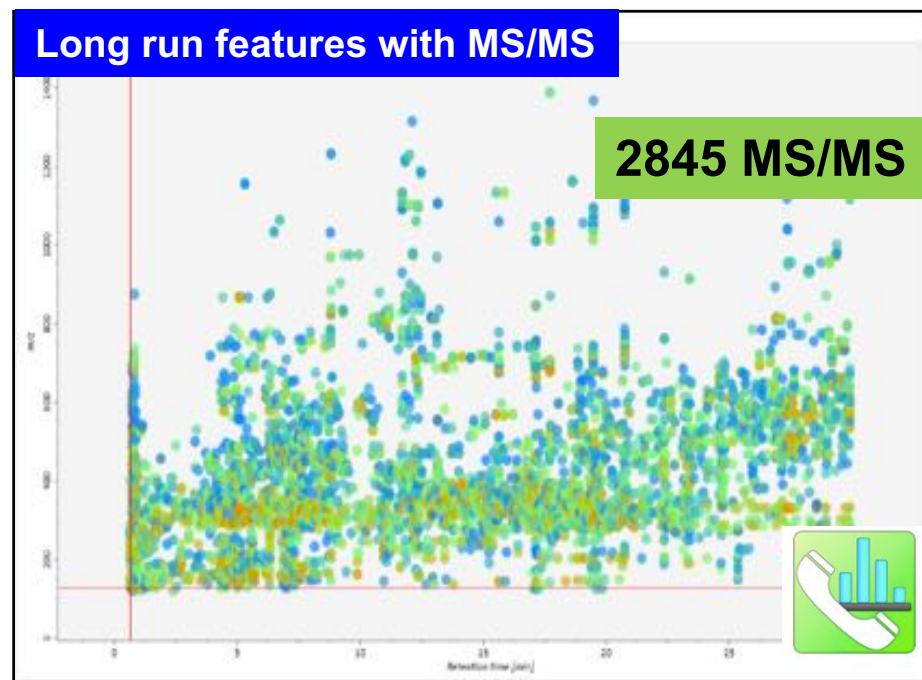


MS/MS ion map of the QC mix of 5 herbs

Assess peak number



2 x 50 mm, C18 BEH Acquity
column, 9 min



2 x 150 mm, C18 BEH Acquity
column, 30 min

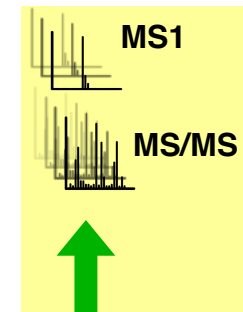
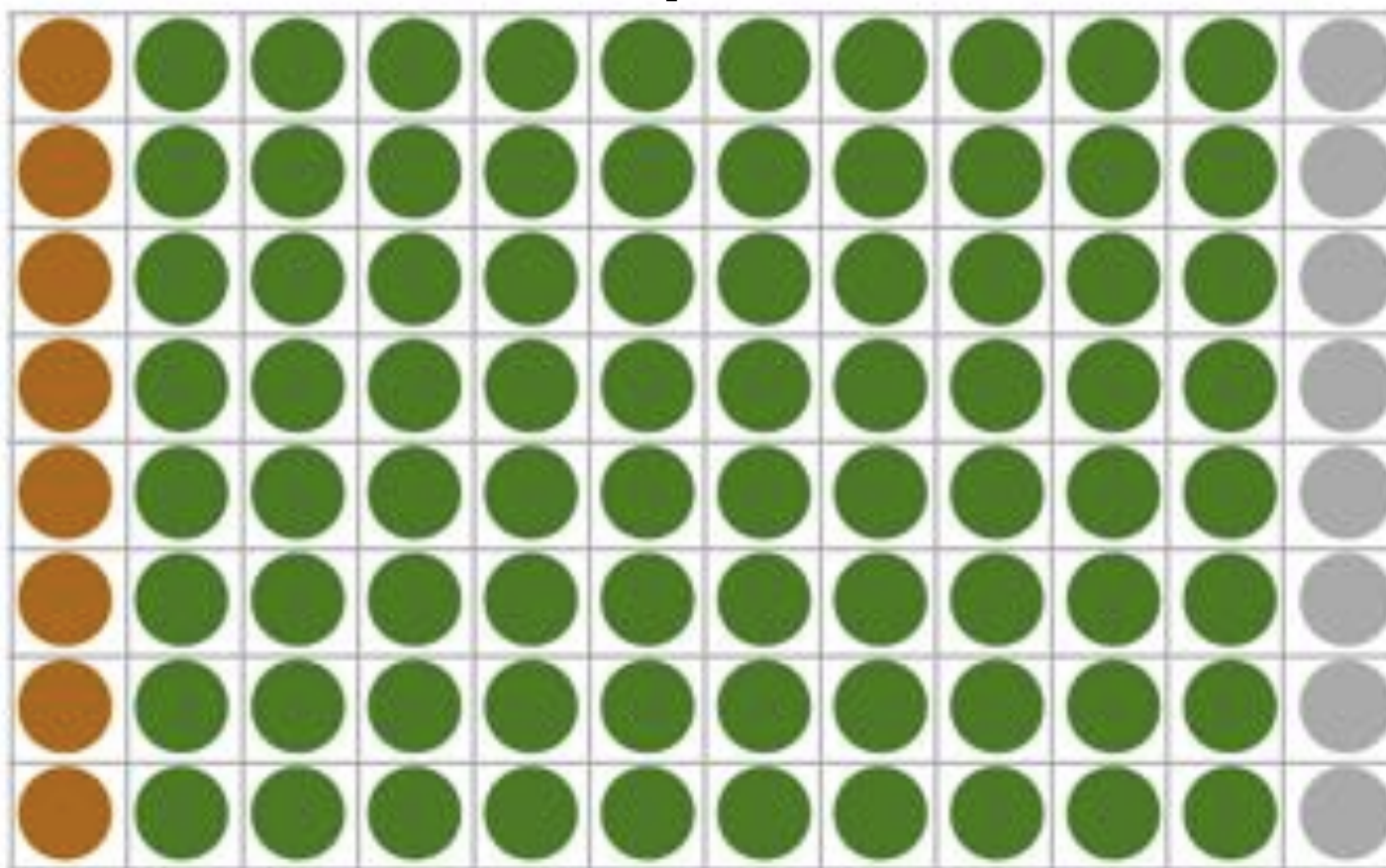


Plate plan

QCmix

samples

blank



Analyzed so far:

480 extracts (6 plates containing 80 extracts)

Visualization in Cytoscape

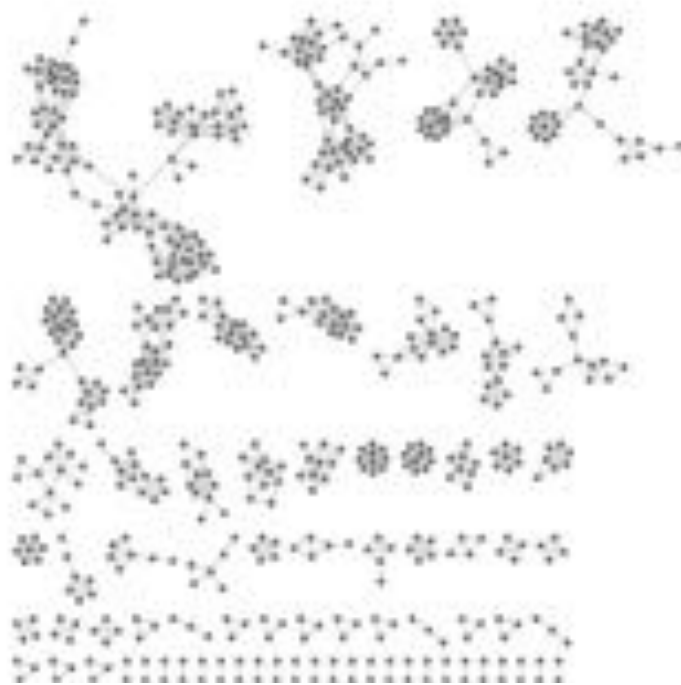


Cytoscape

<http://cytoscape.org/index.html>



python



2 new things in the MN annotations :

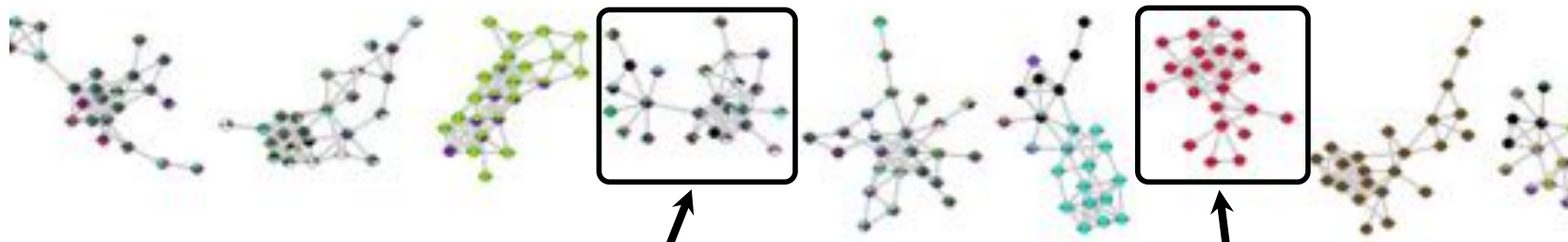
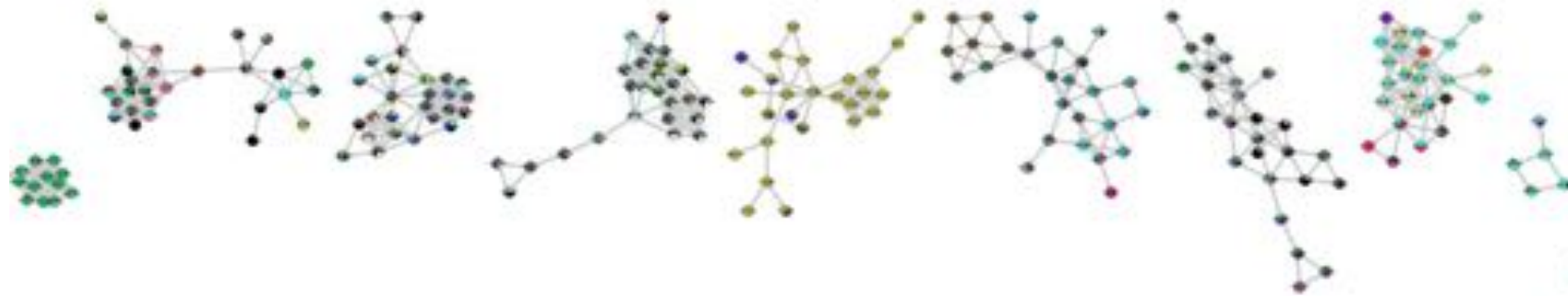
- exact mass and RT info by MzMine pretreatment
- new script gives all additional metadata of annotation

b0149c7badb74b7e913be915c7423630.pairsinfo

Table Panel

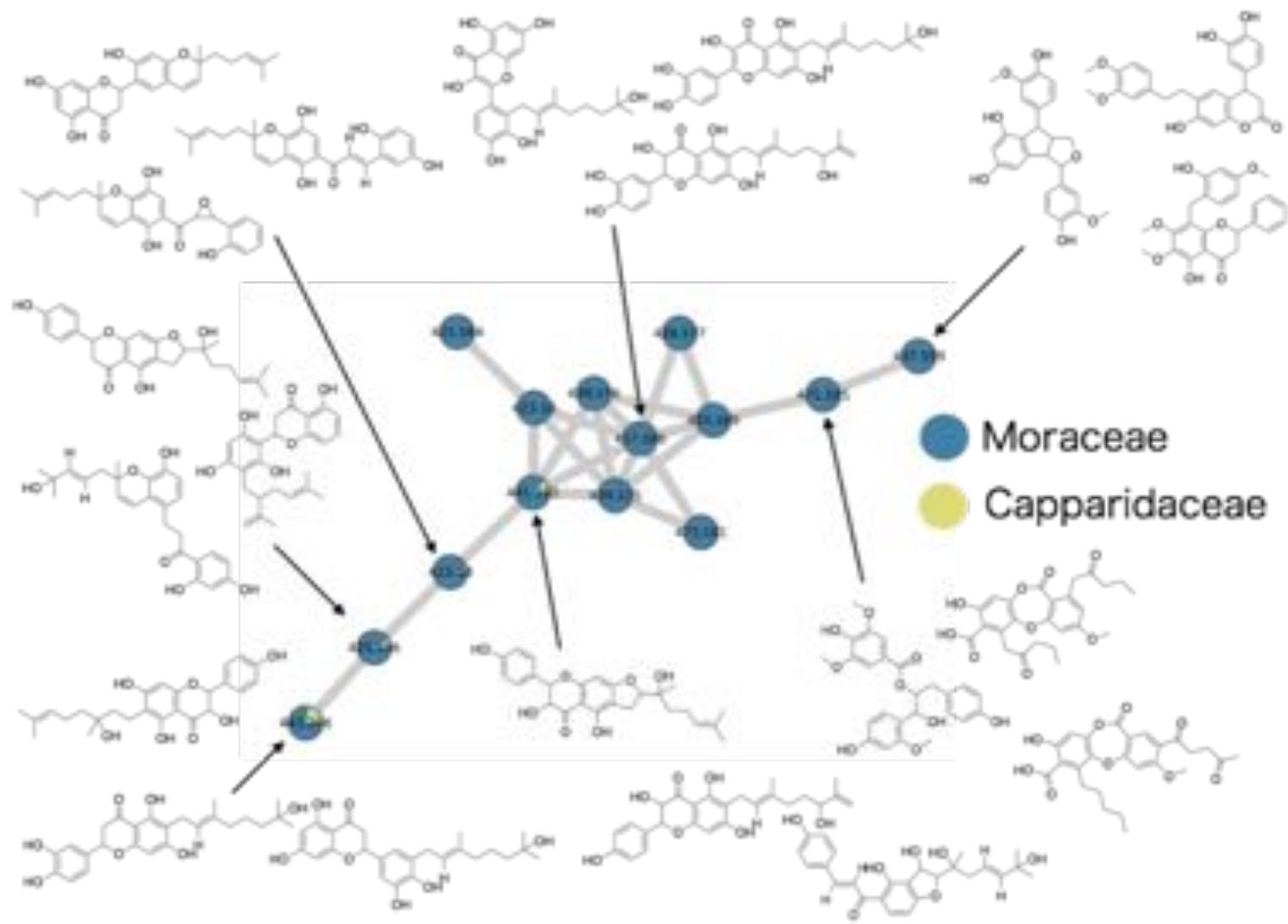
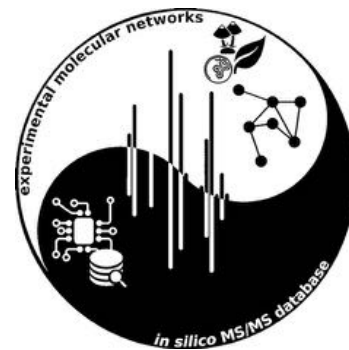
SMILES_DNP	DNP_IDs	Molecular_Formula_DNP	Compound_Types_DNP	Biological_Source_DNP	Molecule_Name_DNP	
[CCCCCCCCCCCC...	[PQF4]-A, OYY...	[C34H67NO5, C34H67NO5, C3...	[VA7200 2Q7020, VA7200 2Q...	[Constit. of the radix of Angel...	[2-Amino-7-octadecene-1...	
Molecule_Name_DNP	Biological_Use_DNP	Spectral_Score_DNP	IS_libname	IS_match_mzerror	SMILES_DNP_QC	DN
Amino-7-octadecene-1...	[]	[0.300686, 0.289631, 0.28...	[New_DNP_full_p...	[0, 0, 0]	[]	[]

Addition of a taxonomical layout according to the group mapping. Done at the Family level



Common to different botanical Family

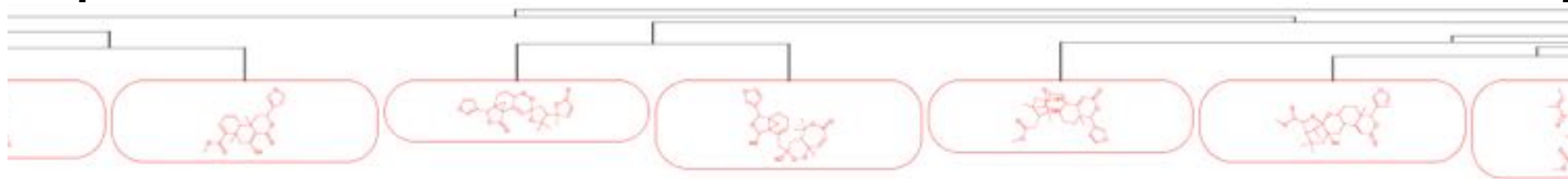
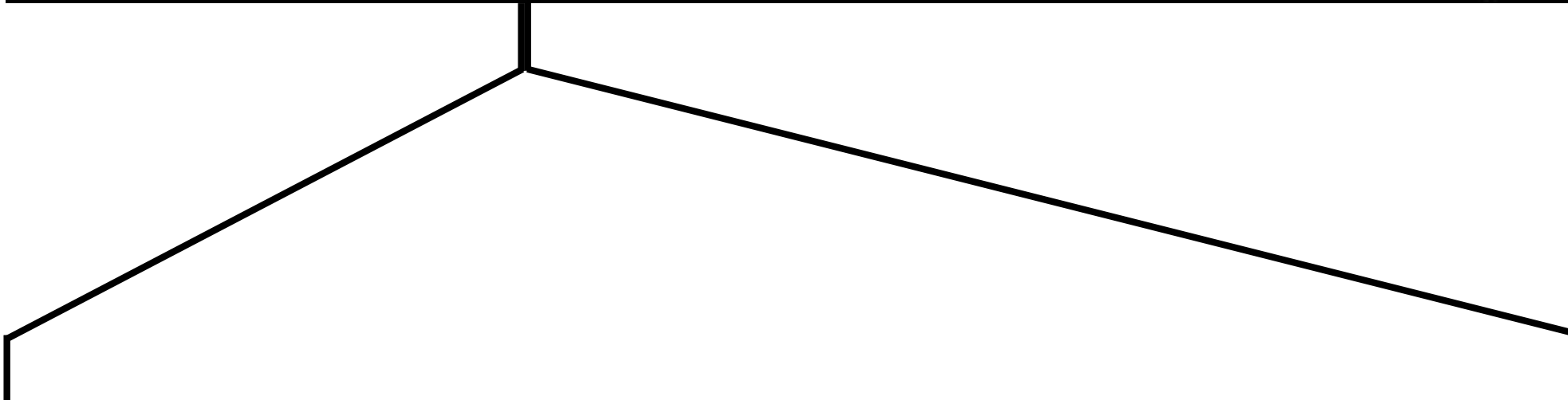
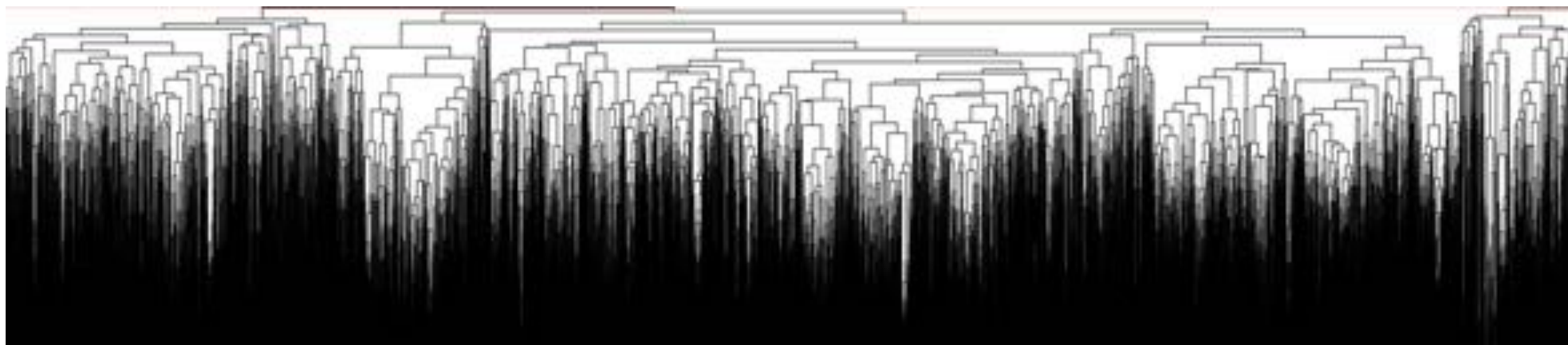
Found in a unique botanical Family



The 7500 structures were loaded in Scaffold Hunter for structural organisation and queries.

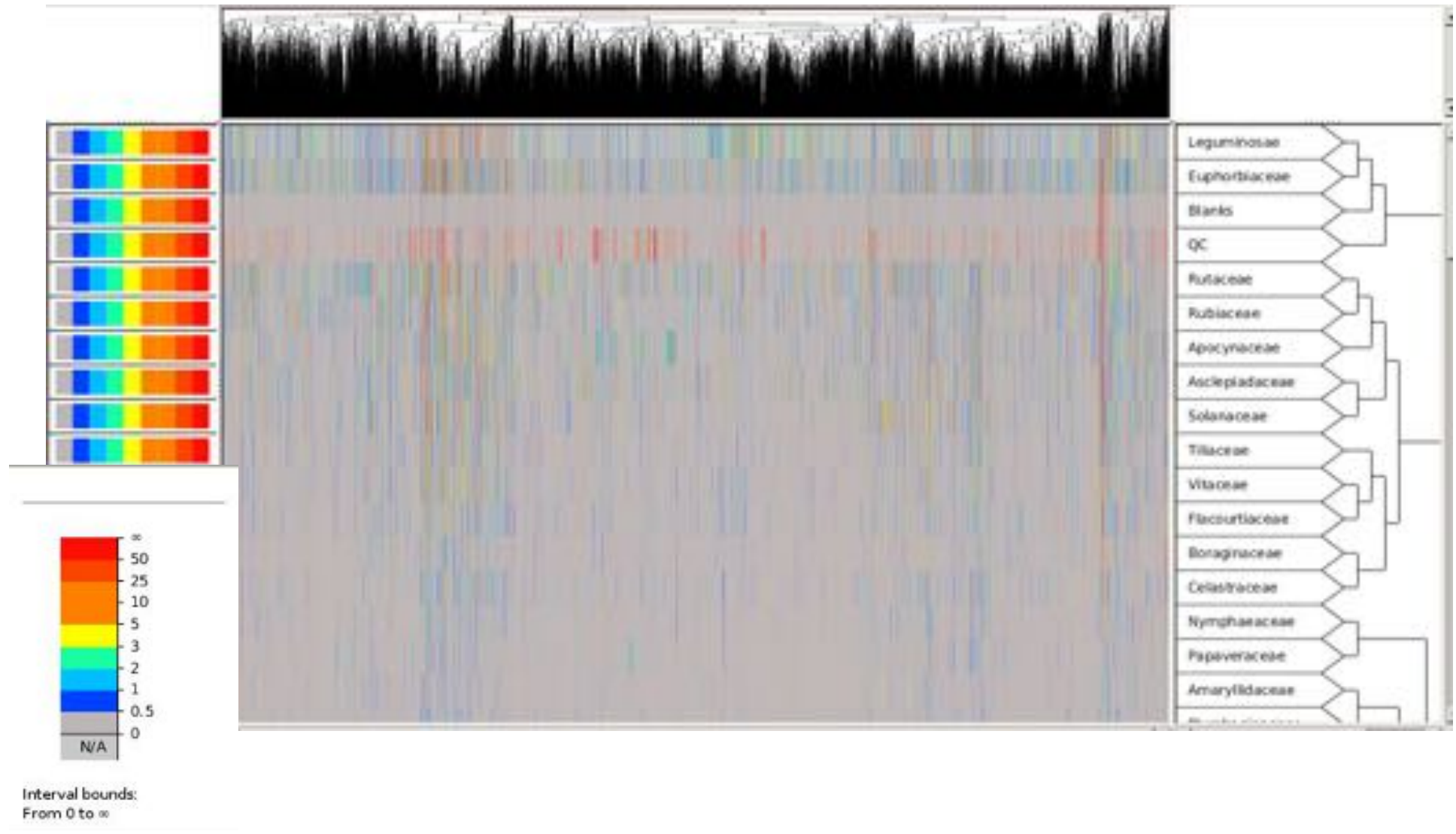


<http://scaffoldhunter.sourceforge.net/>



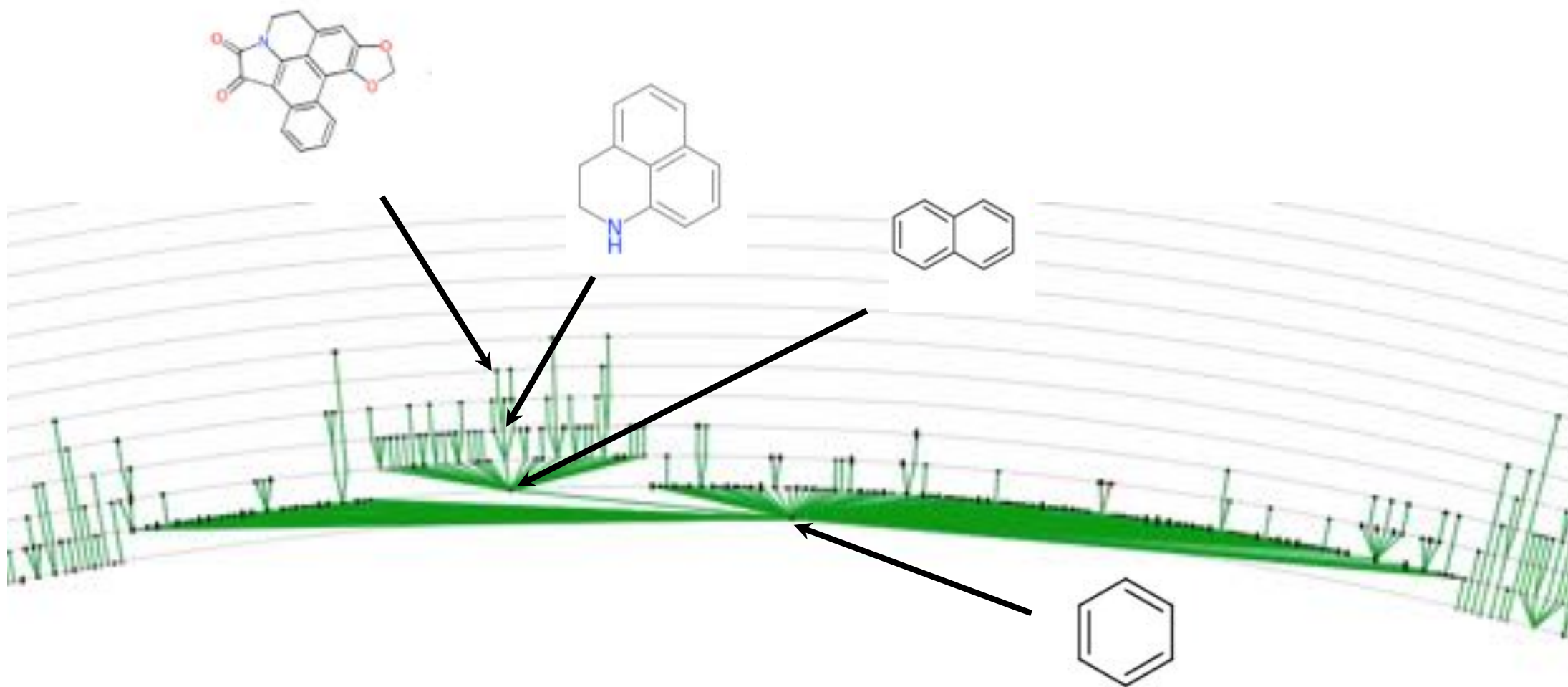
Tanimoto clustering

Heat mapping

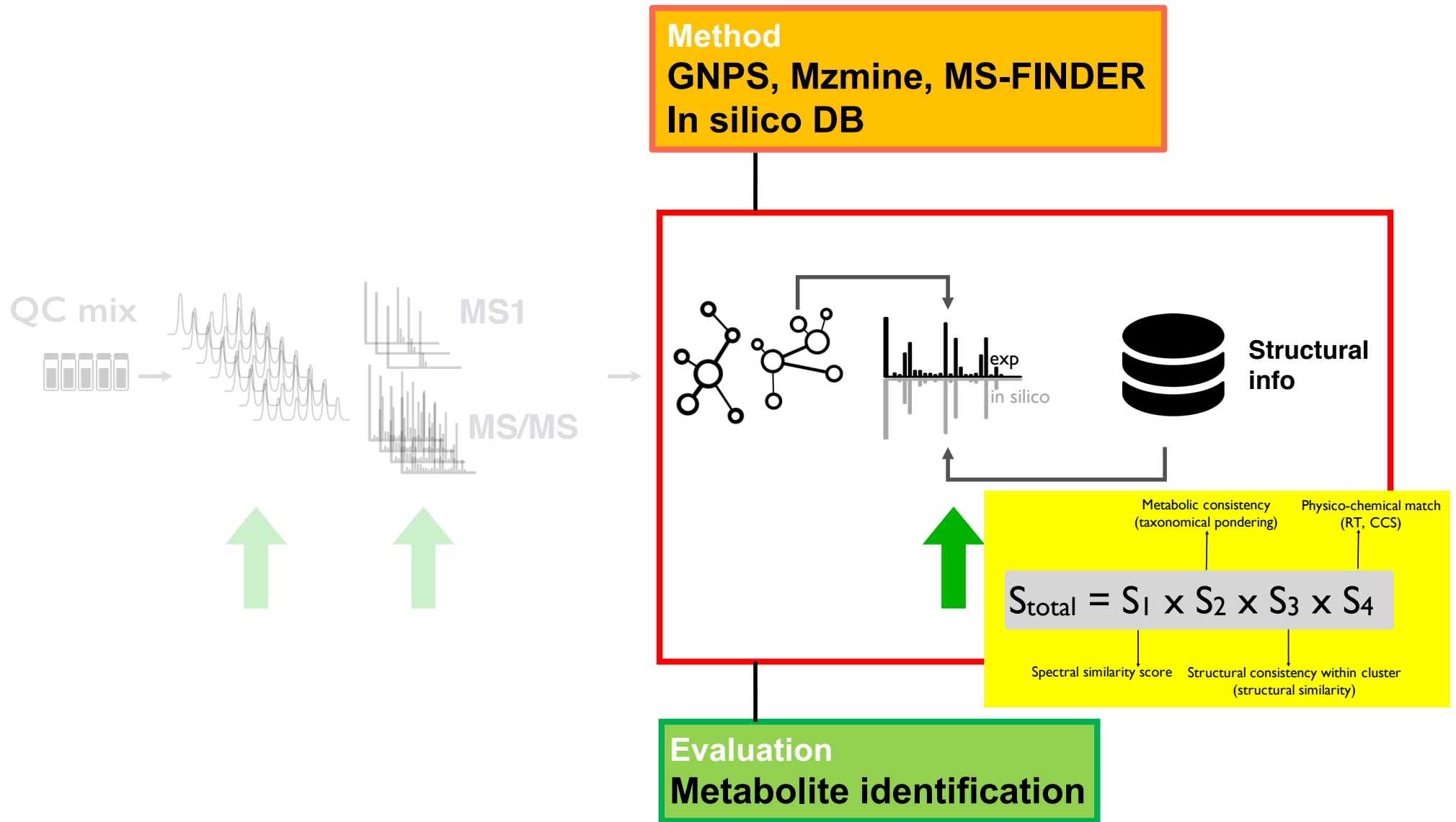


Scaffold view of the 7500 compounds

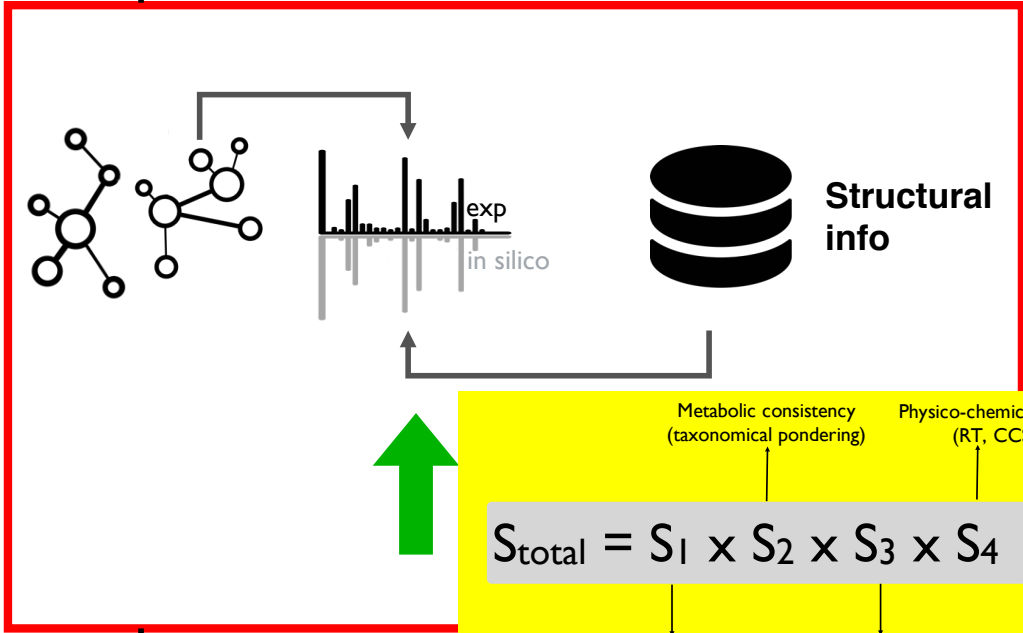




**Useful for finding common structure
and interesting material based on pharmacophore**



Method
GNPS, Mzmine, MS-FINDER
In silico DB



$$S_{total} = S_1 \times S_2 \times S_3 \times S_4$$

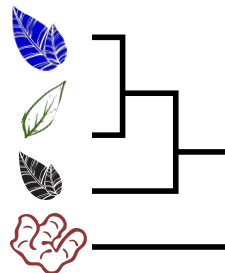
Metabolic consistency (taxonomical pondering) Physico-chemical match (RT, CCS)

Spectral similarity score Structural consistency within cluster (structural similarity)

Evaluation
Metabolite identification

Development of hypothesis metascore

Metabolic consistency
(taxonomical pondering)

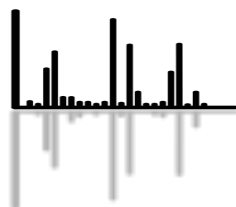


Physico-chemical match
(RT, CCS)

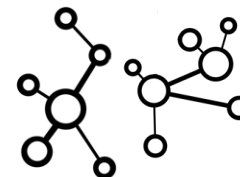


$$S_{\text{total}} = S_1 \times S_2 \times S_3 \times S_4$$

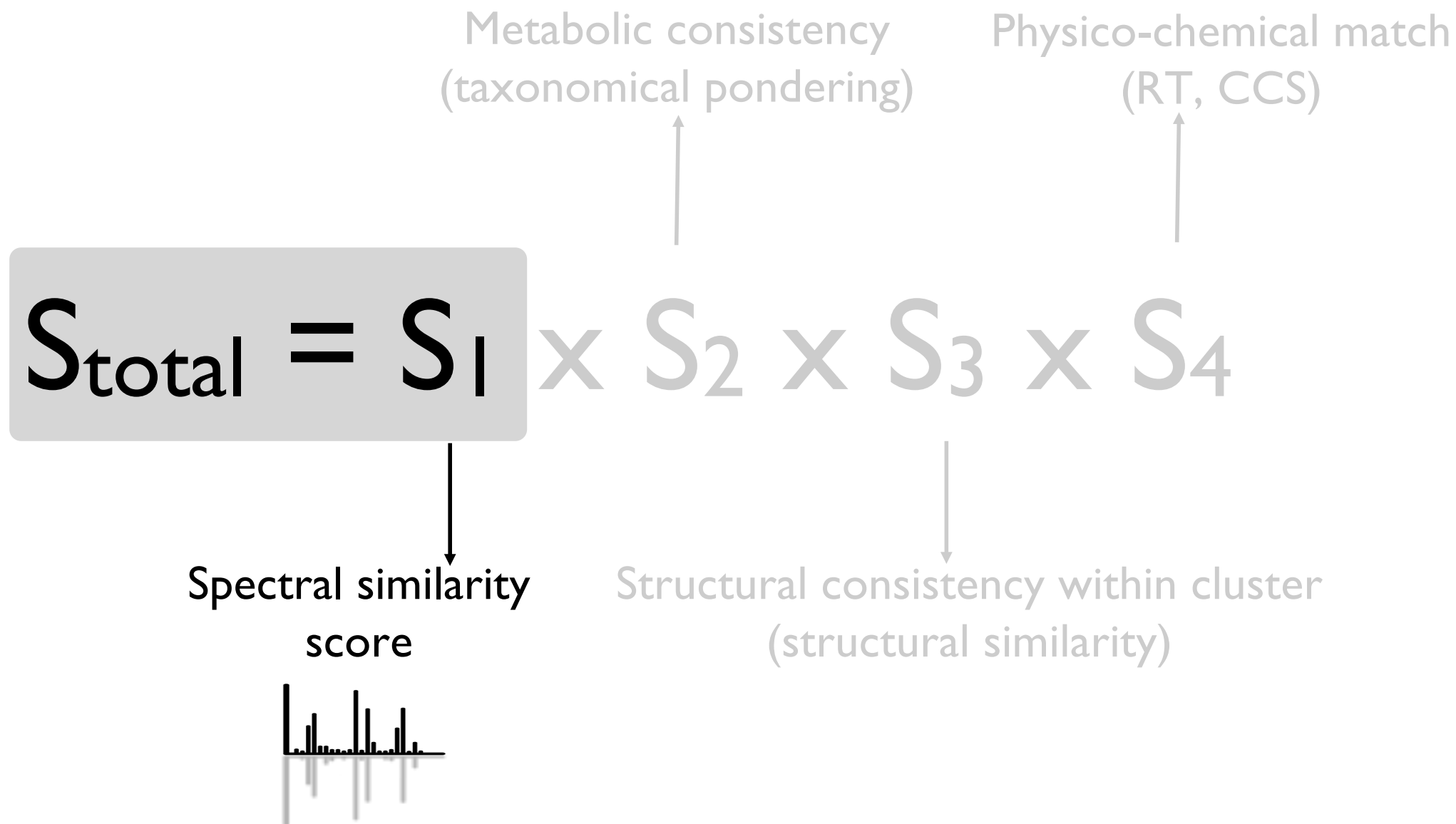
Spectral similarity
score



Structural consistency
within cluster
(structural similarity)

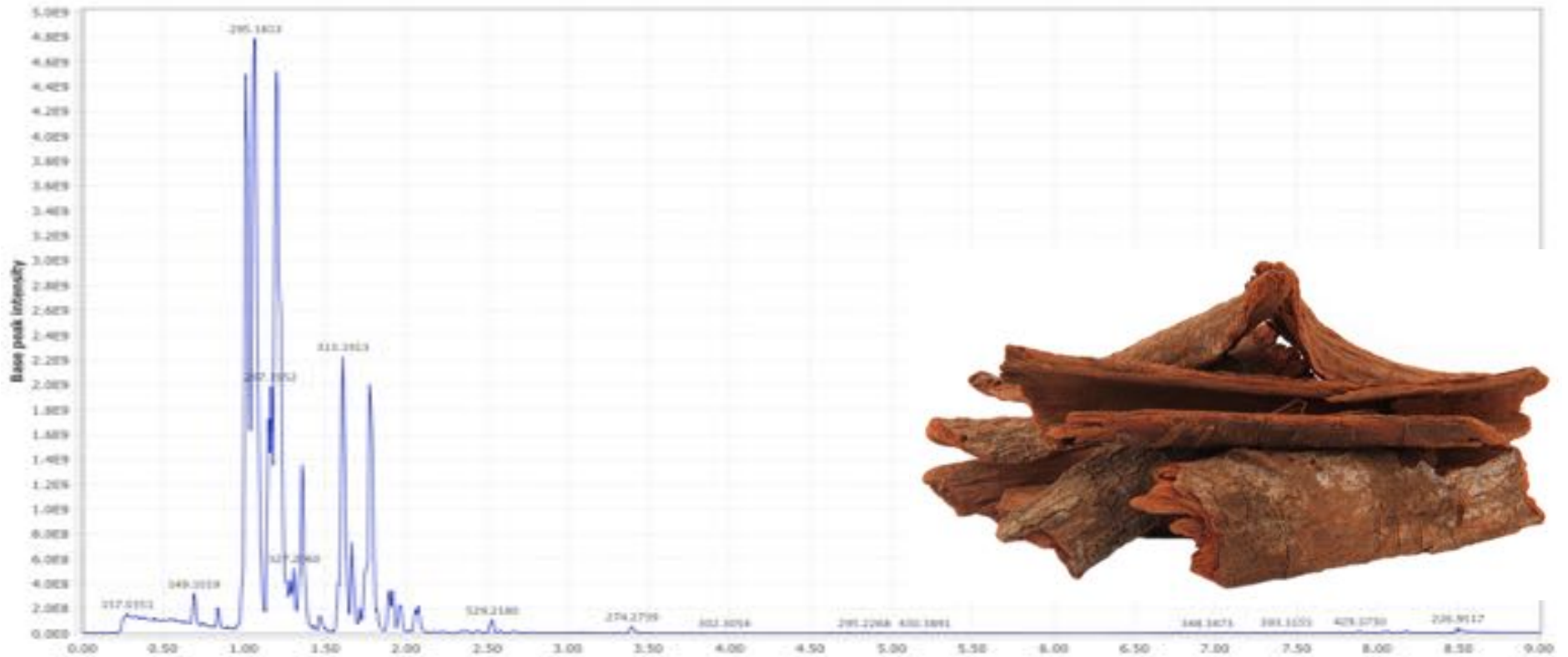


Development of hypothesis metascore



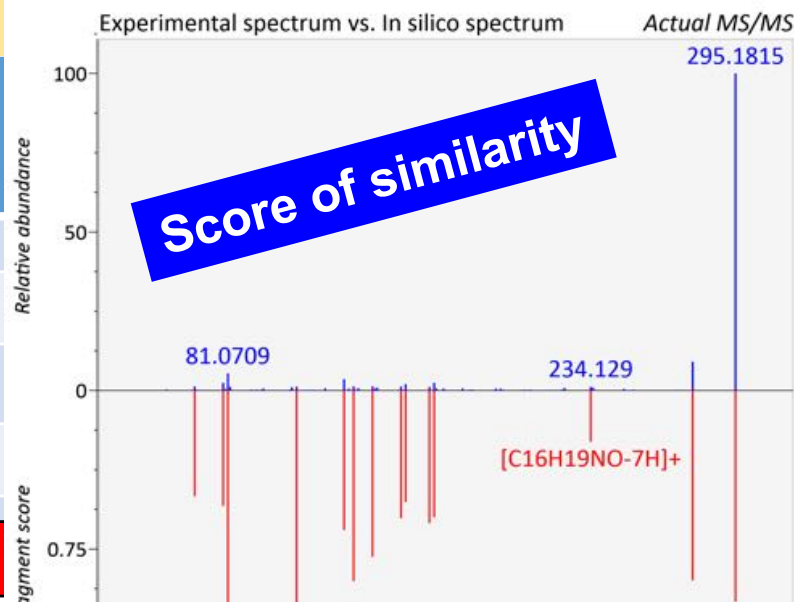


Cinchona pubescens

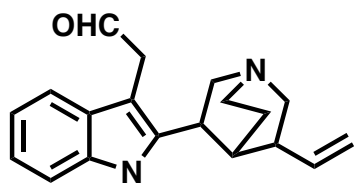


Dereplication results of *C. pubescens* with **score of spectral similarity**

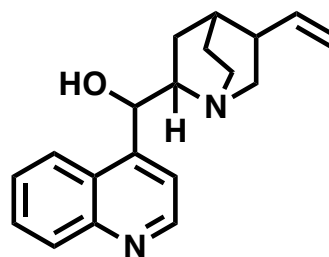
No	m/z
1	295.18134



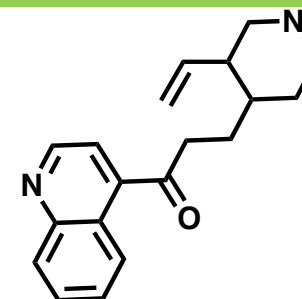
MS2	
Top5 hit structure in MS-Finder	Score
OGY41	7.4705
GPF00	7.088
CFP87,80	6.9084
OTG15	6.6168
CFB99	6.5119



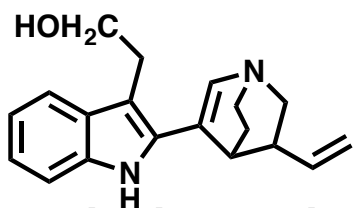
JYL65



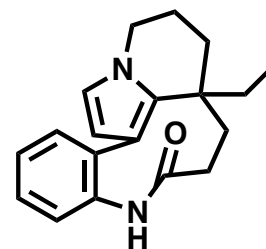
CFP80, CFP87



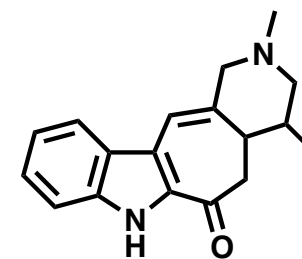
CFB99



apocinchonamine
(OGY41)



rhazinilam
(GPF00)

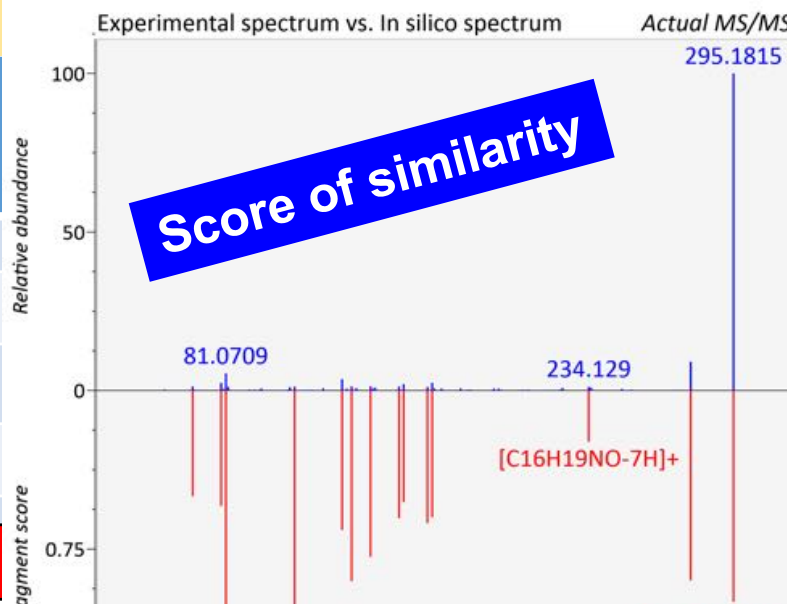


6,16-dihydro-20-episilicine
(OTG15)

$$S_{\text{total}} = S_I$$

Dereplication results of *C. pubescens* with **score of spectral similarity**

No	m/z
1	295.18134

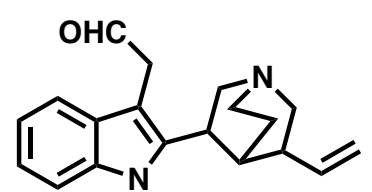


MS2	
Top5 hit structure in MS-Finder	S ₁
OGY41	7.4705
GPF00	7.088
CFP87,80	6.9084
OTG15	6.6168
CFB99	6.5119

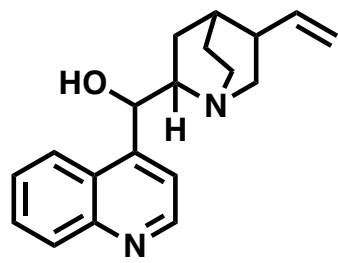


- 1
- 2
- 3
- 4
- 5

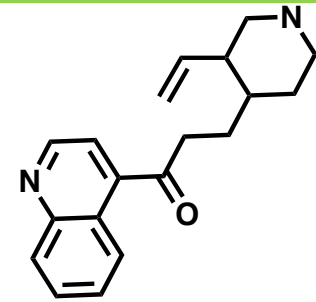
MS/MS scoring whole DB



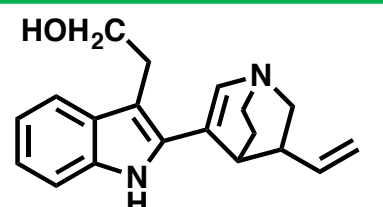
JYL65



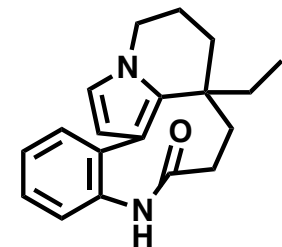
CFP80, CFP87



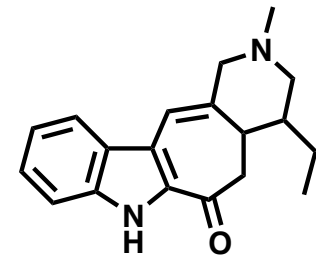
CFB99



apocinchonamine (OGY41)



rhazinilam (GPF00)

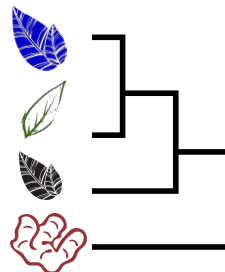


6,16-dihydro-20-episilicine (OTG15) CGJ39: No structure information

$$S_{\text{total}} = S_1$$

Development of hypothesis metascore

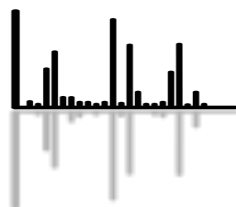
Metabolic consistency
(taxonomical pondering)



Physico-chemical match
(RT, CCS)

$$S_{\text{total}} = S_1 \times S_2 \times S_3 \times S_4$$

Spectral similarity
score

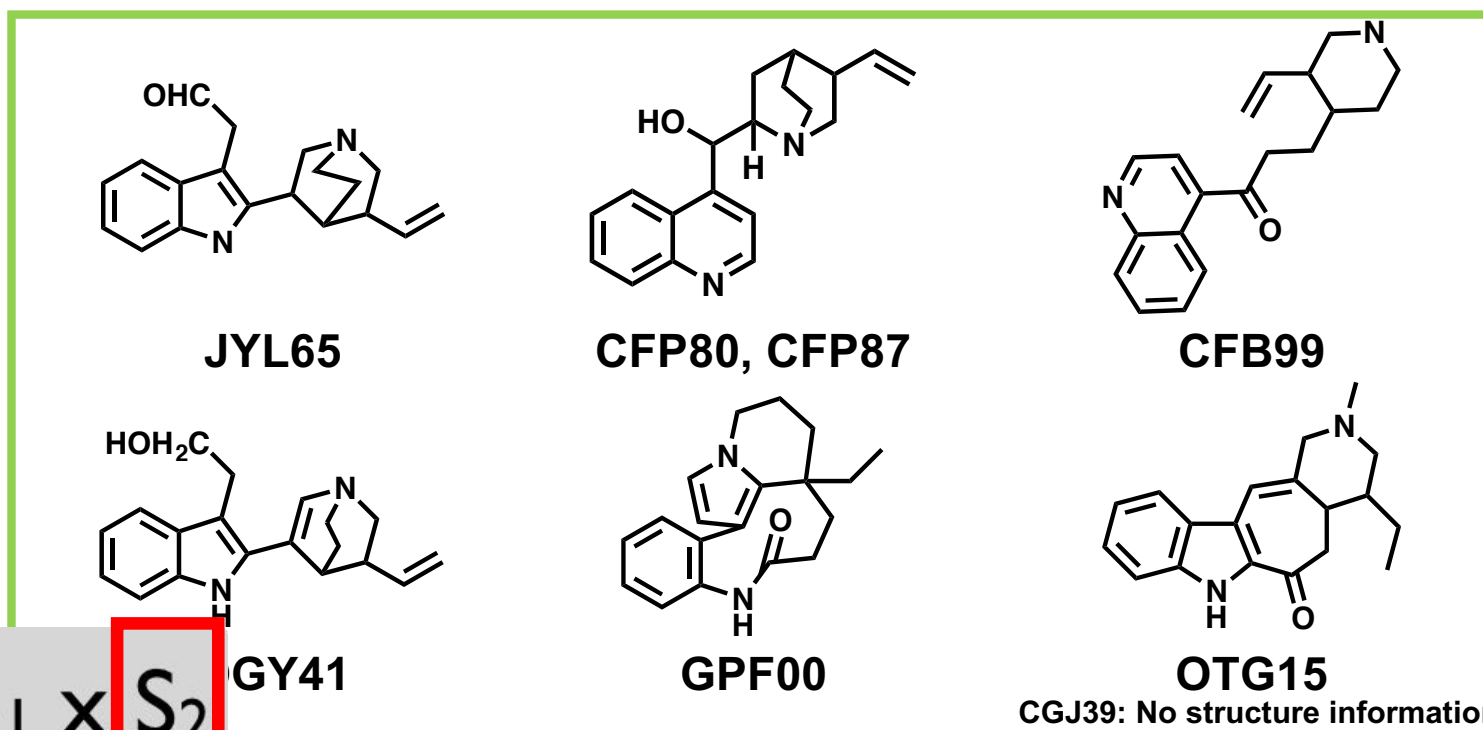


Structural consistency within cluster
(structural similarity)

Dereplication results of *C. pubescens* with **Metascore**

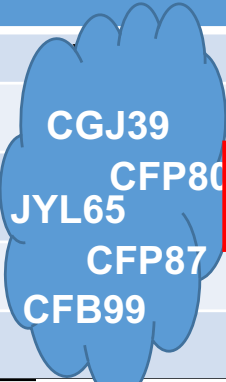
No	MZmine	Top5 hit structure in MS_Finder	S ₁	genus	point*	
1 C19H22N2O	CGJ39 CFP80 JYL65 CFP87 CFB99	OGY41	7.471	1	No IF	1
		GPF00	7.088	2	another family	1
		CFP87,80	6.908	3	<i>C. officinalis</i> and all <i>Cinchona</i> spp.	1.6
		OTG15	6.617	4	another family	1
		CFB99	6.512	5	<i>Succirubra</i> (<i>pubescens</i>)	1.6

1.6 point: same sp
 1.4 point: same genus
 1.2 point: same family
 1.0 point: another family and No IF

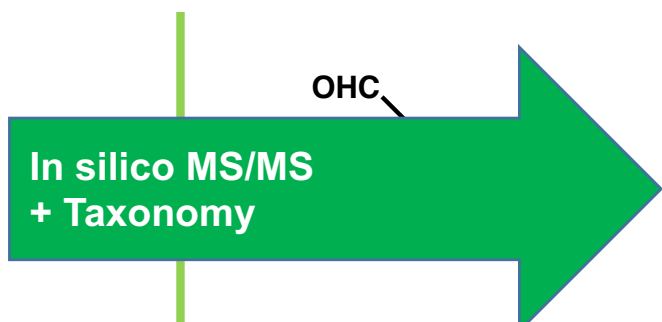


$$S_{\text{total}} = S_1 \times S_2$$

Dereplication results of *C. pubescens* with **Metascore**

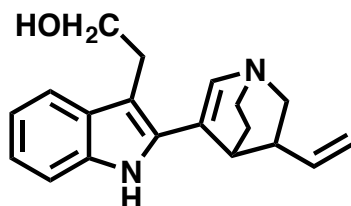
No	MZmine	Top5 hit structure in MS_Finder	S_1	genus	point*	Top5 hit structure in MS_Finder	Metascore Score_taxonomically
1 C19H22N2O		OGY41	7.471	1	No IF	OGY41	7.471
		GPF00	7.088	2	another family	GPF00	7.088
		CFP87,80	6.908	3	<i>C. officinalis</i> and all <i>Cinchona</i> spp.	CFP87,80	11.053
		OTG15	6.617	4	another family	OTG15	6.617
		CFB99	6.512	5	<i>uccirubra</i> (<i>pubescens</i>)	CFB99	10.419

*1.6 point: same sp, 1.4 point: same genus, 1.2 point: same family, 1.0 point: another family and No IF

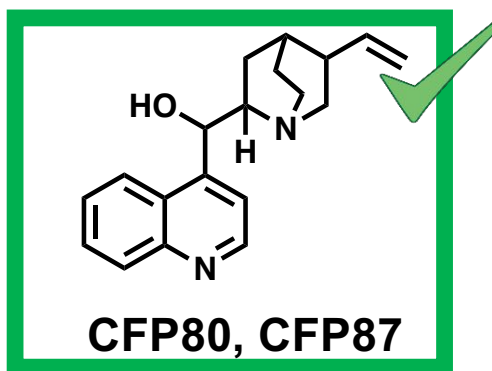


OHC

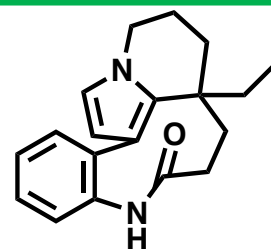
JYL65



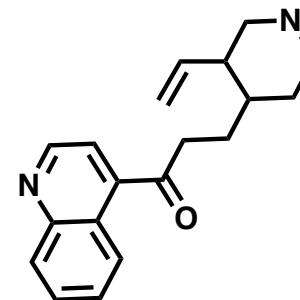
OGY41



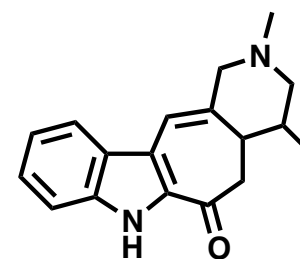
CFP80, CFP87



GPF00



CFB99



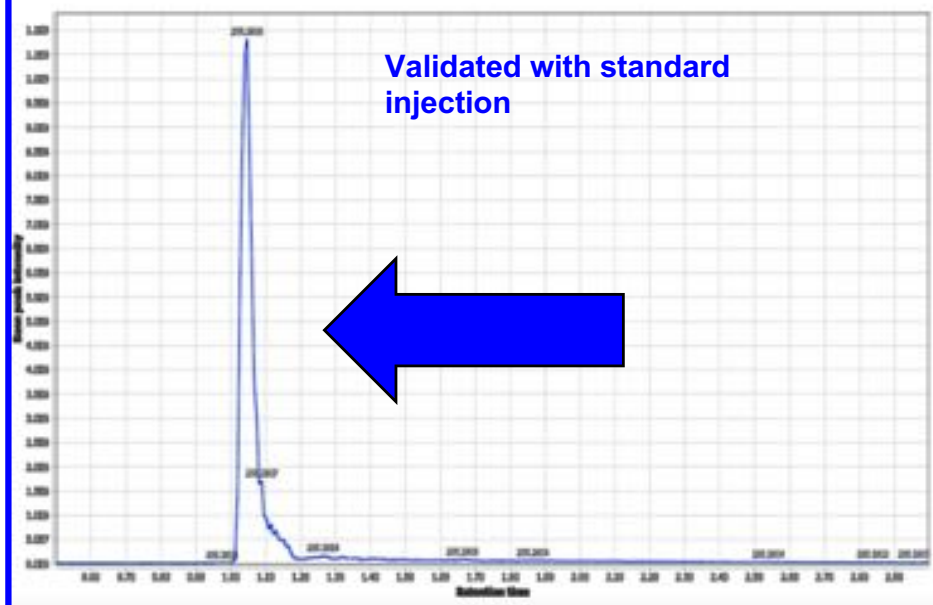
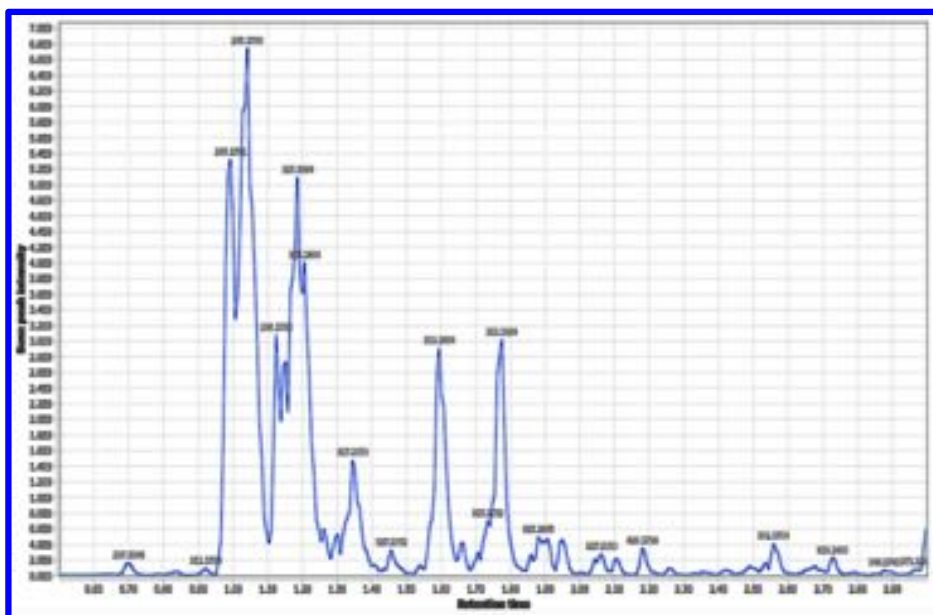
OTG15

CGJ39: No structure information

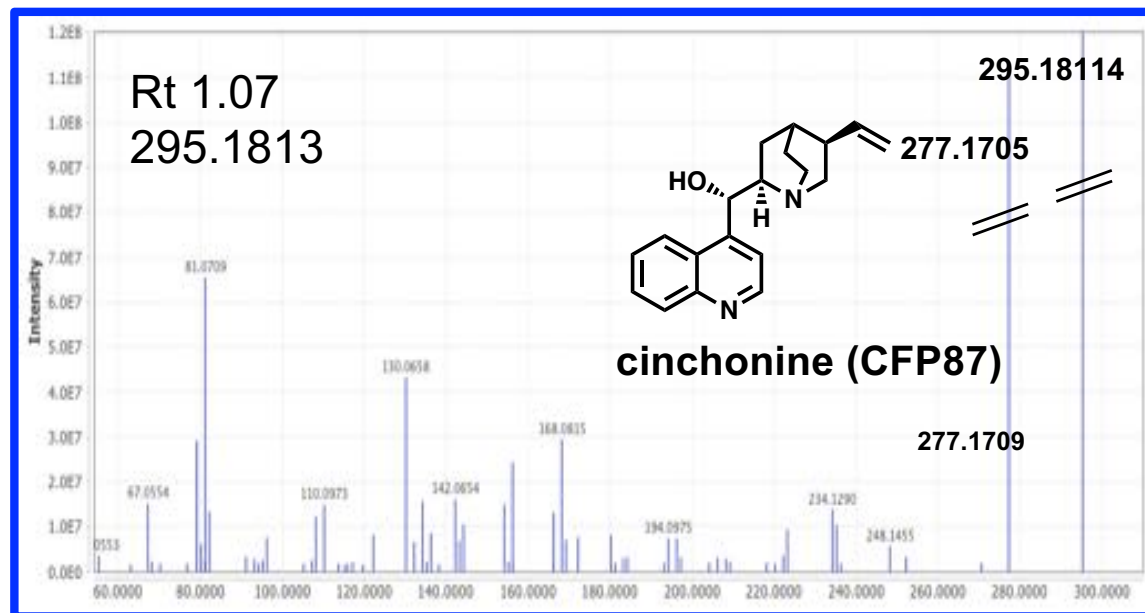
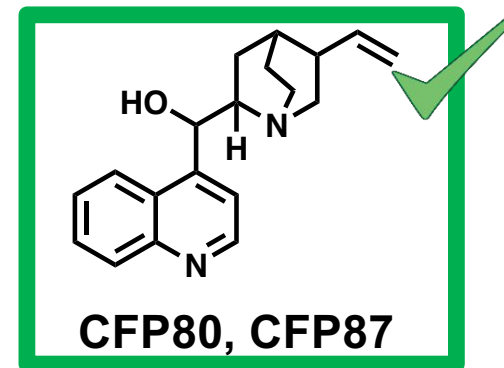


$$S_{\text{total}} = S_1 \times S_2$$

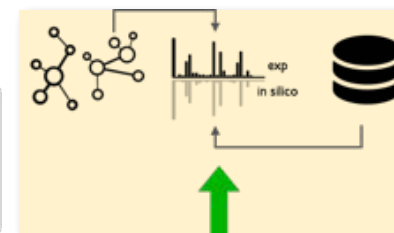
Validated with standard injection



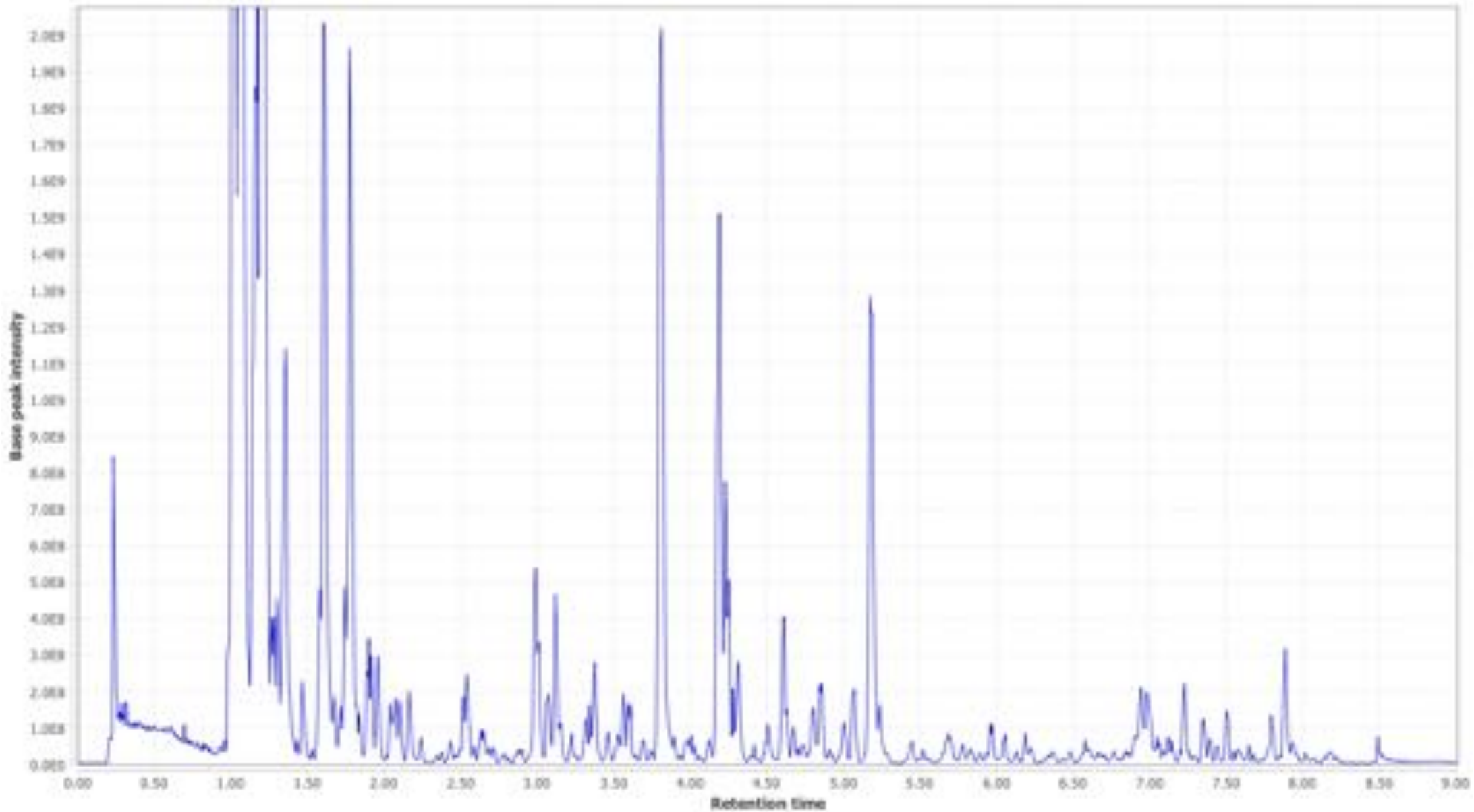
In silico MS/MS
+ Taxonomy



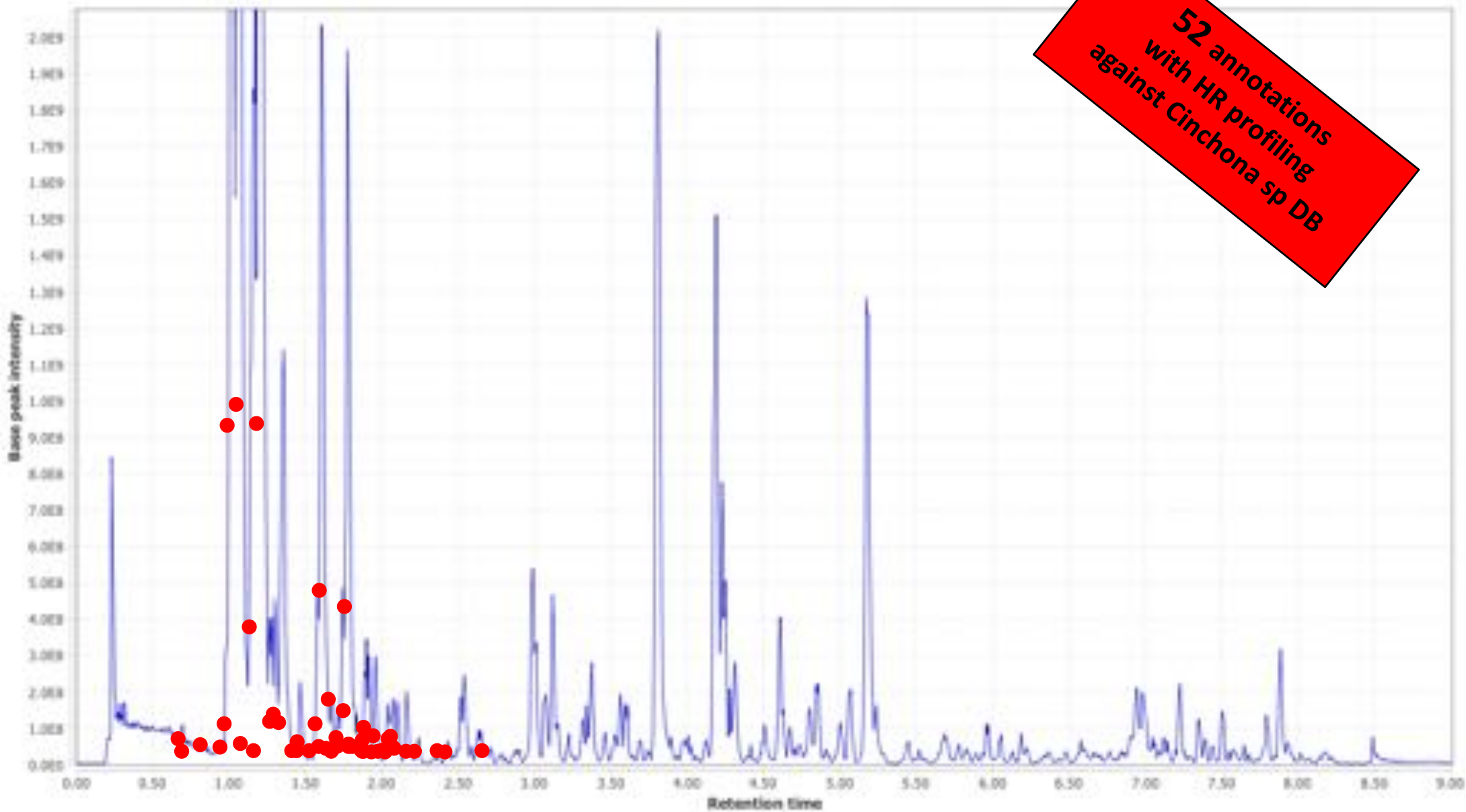
$$S_{\text{total}} = S_1 \times S_2$$



UHPLC-HRMS profiling of the QC mix of 5 herbs

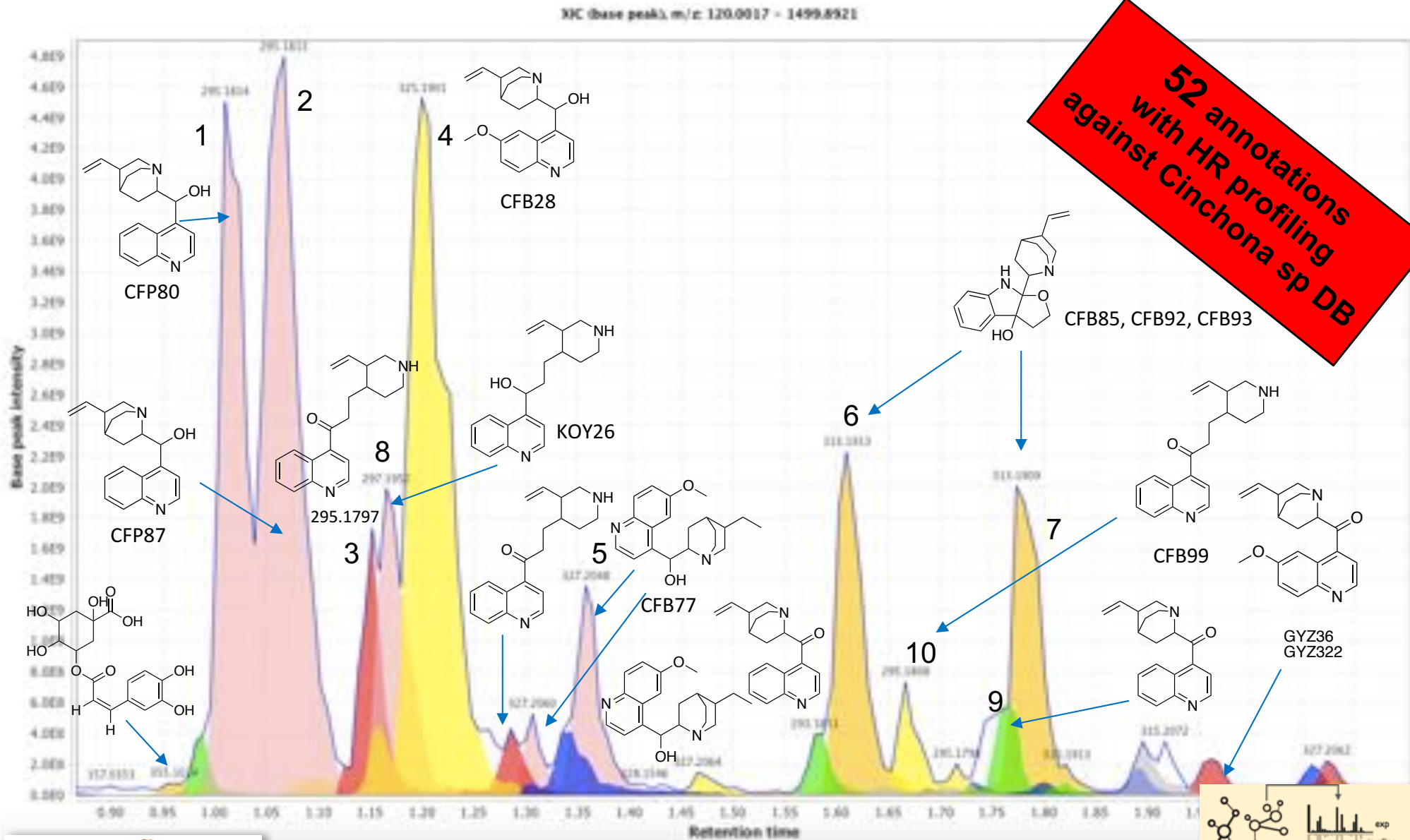


UHPLC-HRMS profiling of the QC mix of 5 herbs

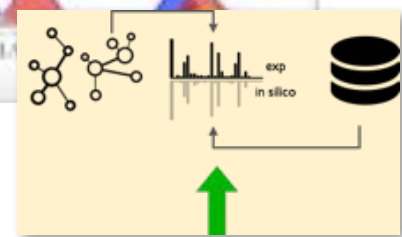


52 annotations
with HR profiling
against Cinchona sp DB

Alkaloids annotated from *C. pubescens*

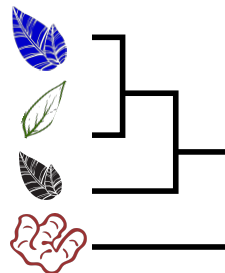


$$S_{\text{total}} = S_1 \times S_2$$



Development of hypothesis metascore

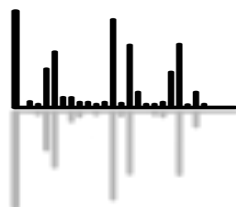
Metabolic consistency
(taxonomical pondering)



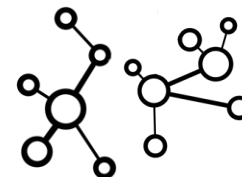
Physico-chemical match
(RT, CCS)

$$S_{\text{total}} = S_1 \times S_2 \times S_3 \times S_4$$

Spectral similarity
score

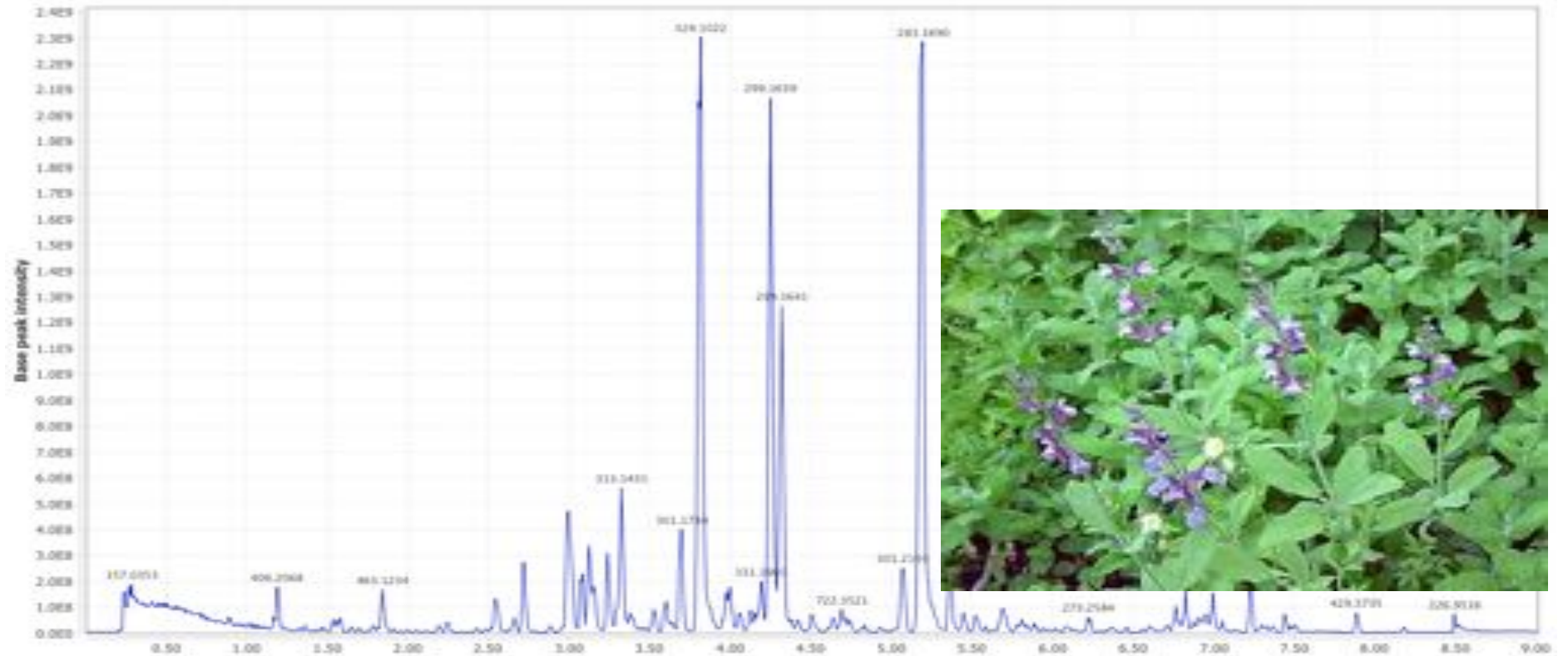


Structural consistency
within cluster
(structural similarity)

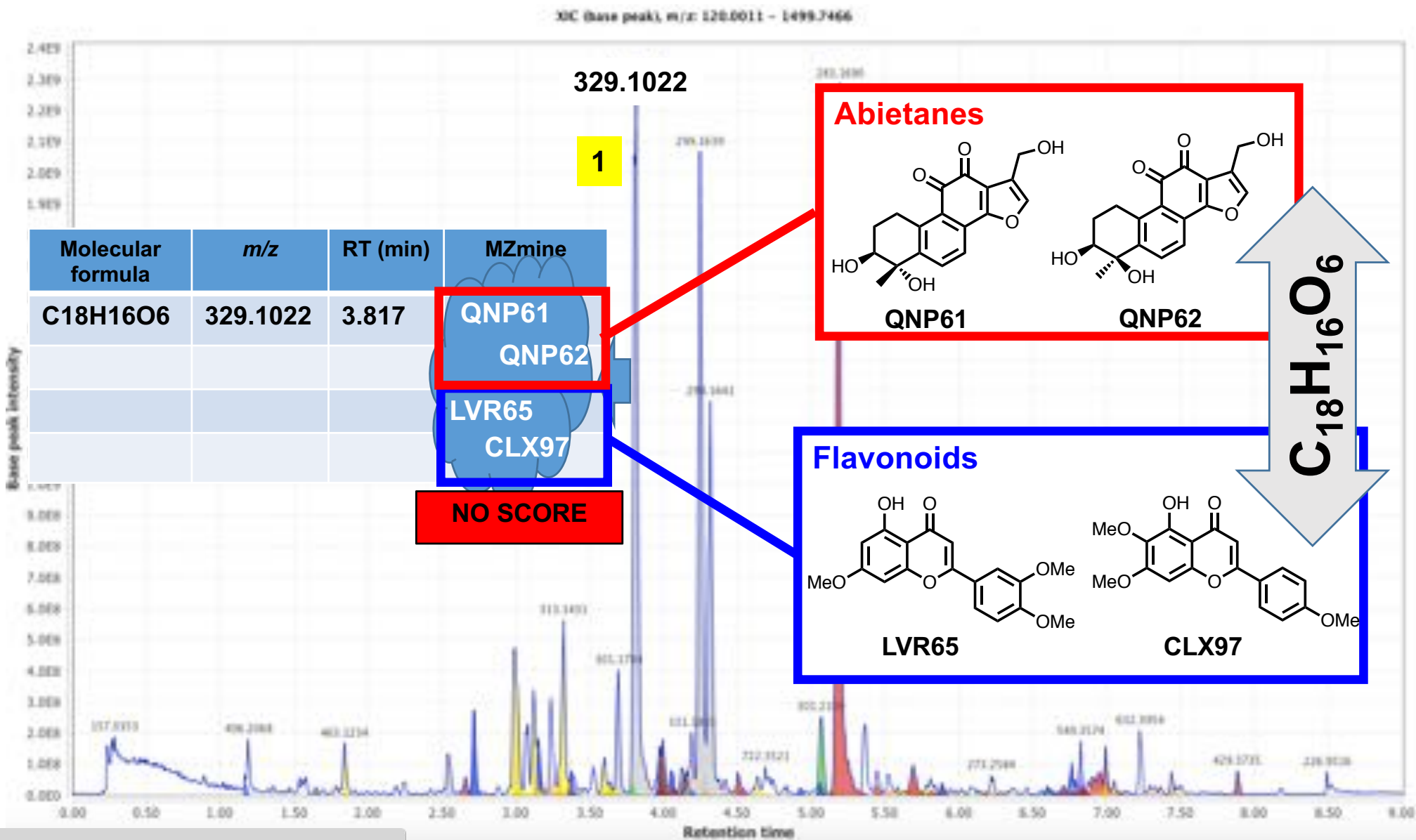




Salvia officinalis

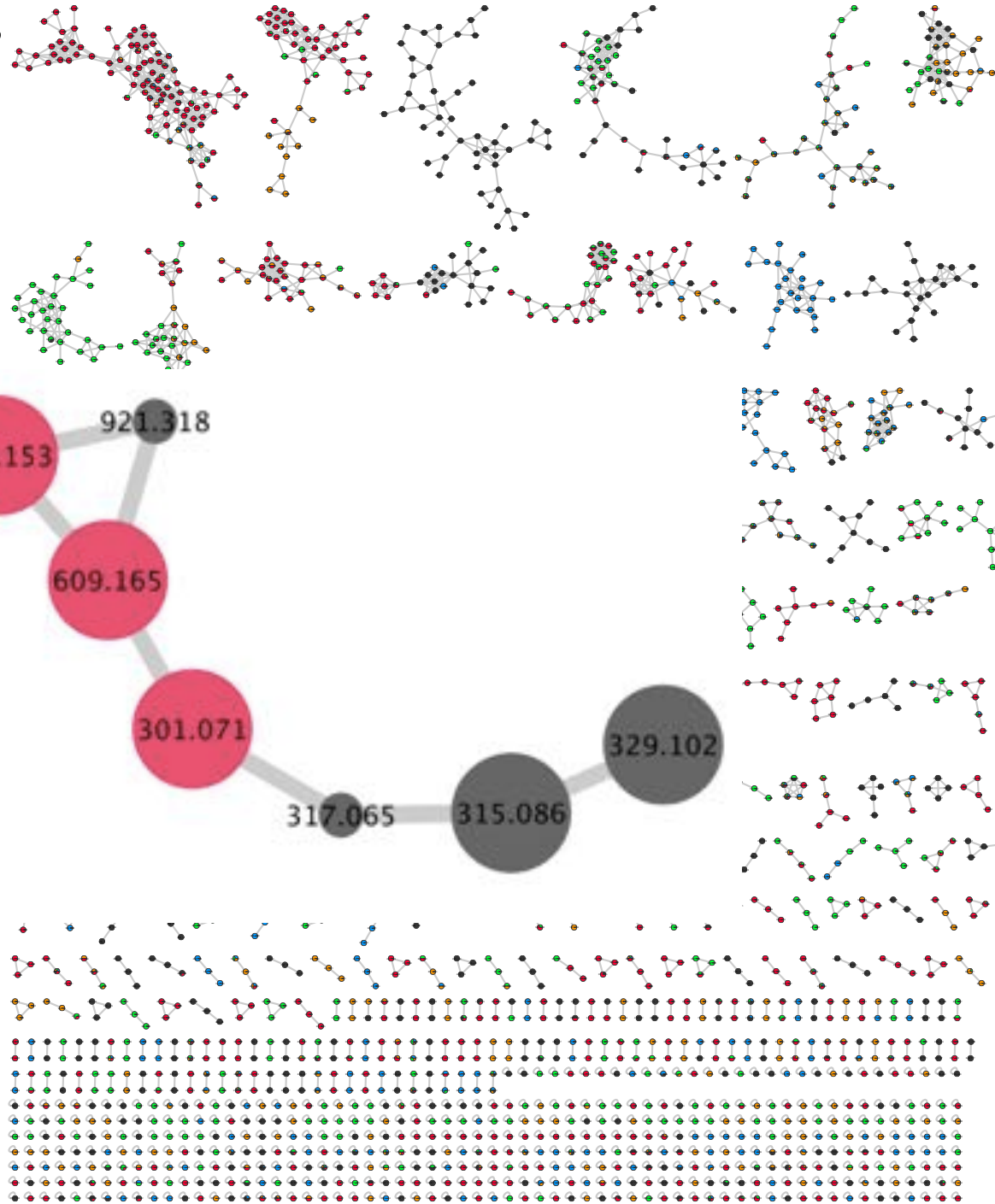


HRMS-ESI (+) of the MeOH extract of *S. officinalis* with annotation restricted to *Salvia spp.*



$$S_{\text{total}} = S_1 \times S_2 \times S_3$$

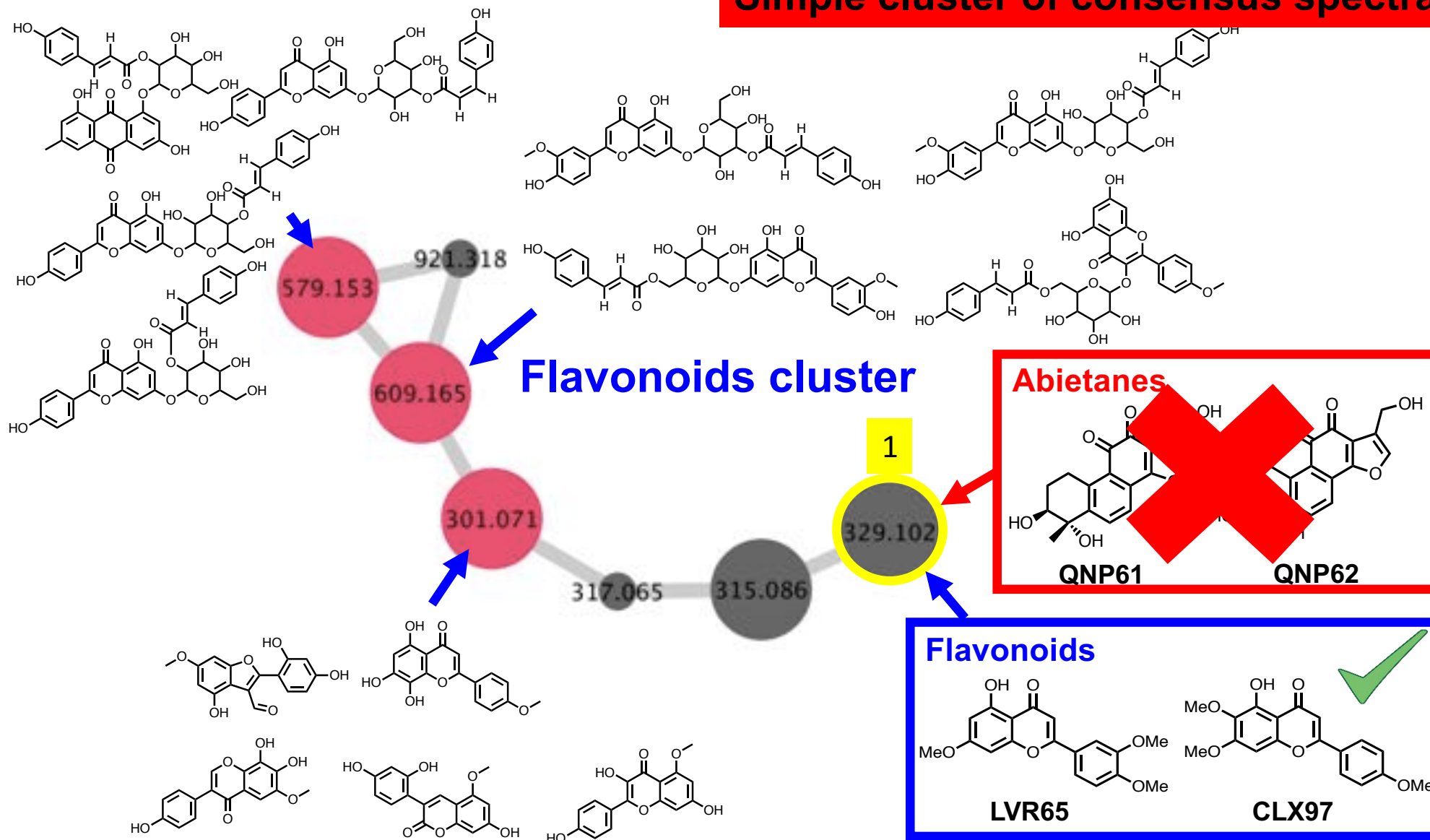
Molecular networking profiles of QCmix



- Cinchona
- Arnica
- Ginseng
- Ginkgo
- Salvia

Molecular networking profiles of *S. officinalis*

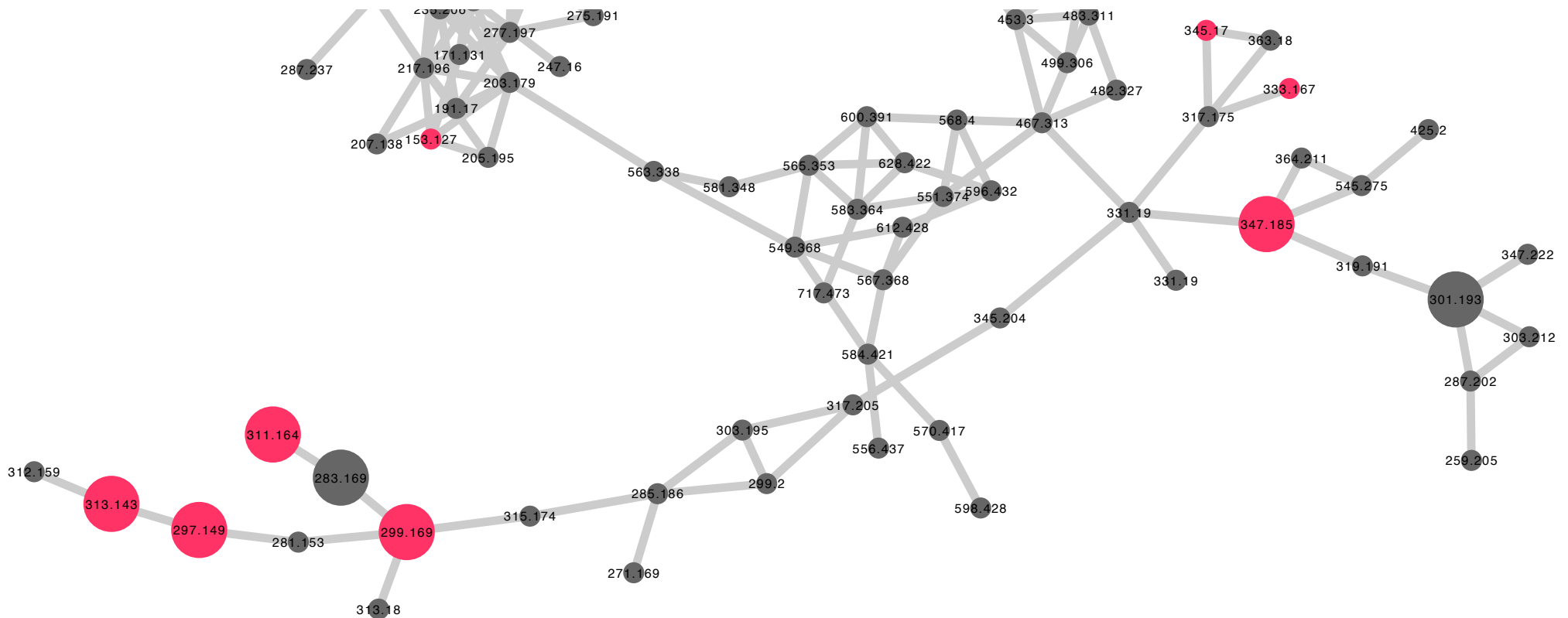
Simple cluster of consensus spectra



$$S_{\text{total}} = S_1 \times S_2 \times S_3$$

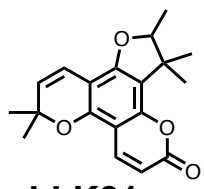
Manual comparison of structure types

Molecular networking profiles of *S. officinalis*

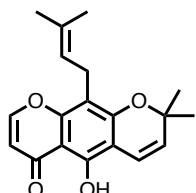


$$S_{\text{total}} = S_1 \times S_2 \times S_3$$

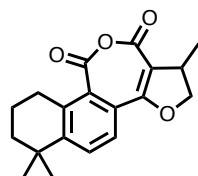
m/z 313.143



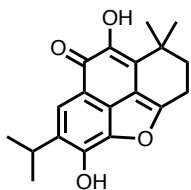
LLK81



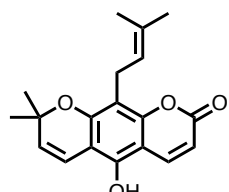
NNF29



HPX83

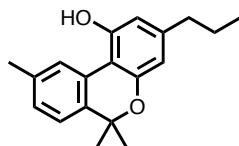


OOW07

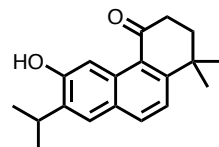


CLT75

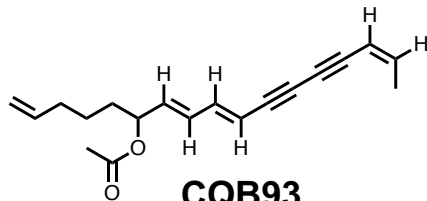
m/z 283.169



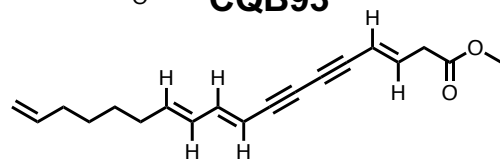
CMK79



QCY13

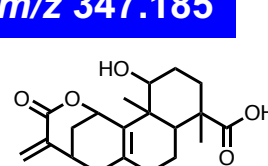


CQB93

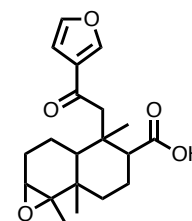


LKT25, LLR47

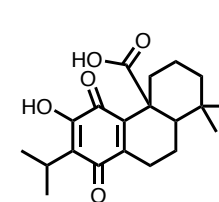
m/z 347.185



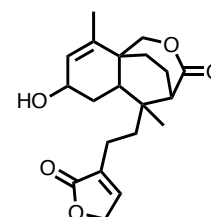
NQN19



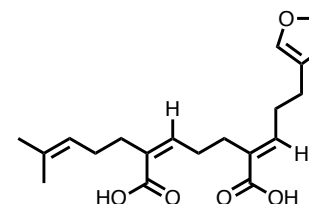
LFM65



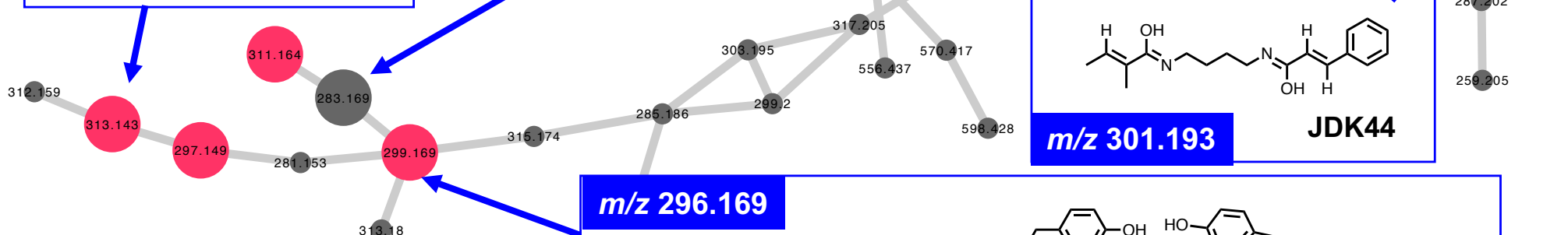
LGC37



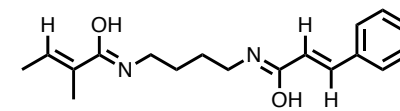
MTG27



PSH50

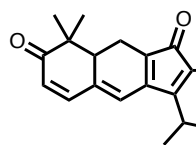


m/z 301.193

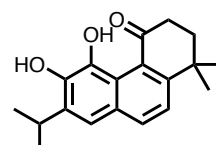


JDK44

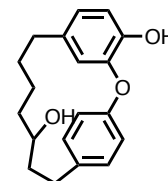
m/z 296.169



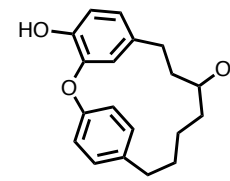
PCX46



JNS32



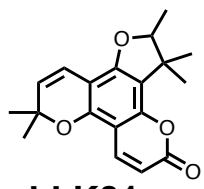
HGT33



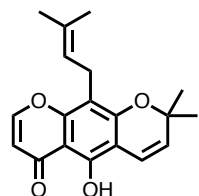
MYB70, MYB68

$S_{total} = S_I$

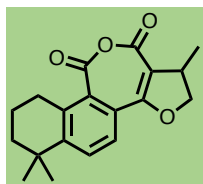
m/z 313.143



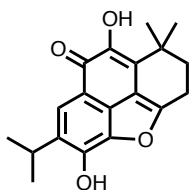
LLK81



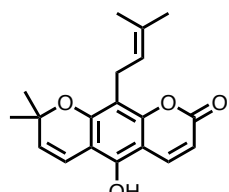
NNF29



HPX83

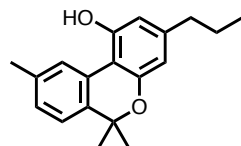


OOW07

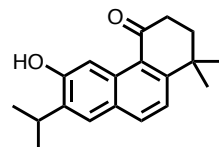


CLT75

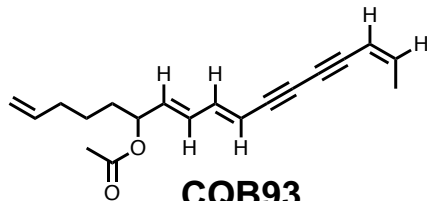
m/z 283.169



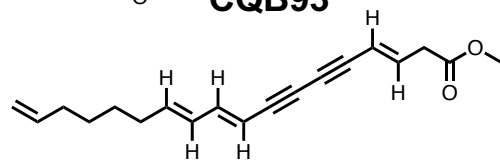
CMK79



QCY13

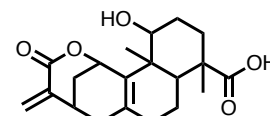


CQB93

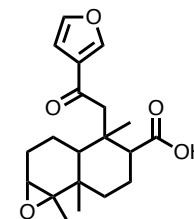


LKT25, LLR47

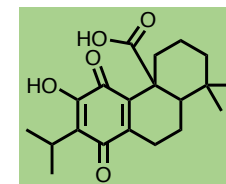
m/z 347.185



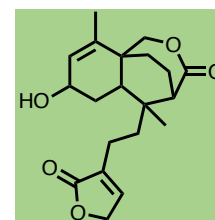
NQN19



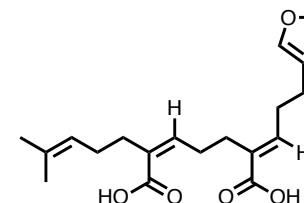
LFM65



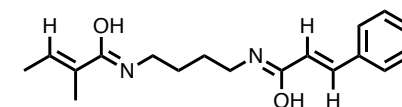
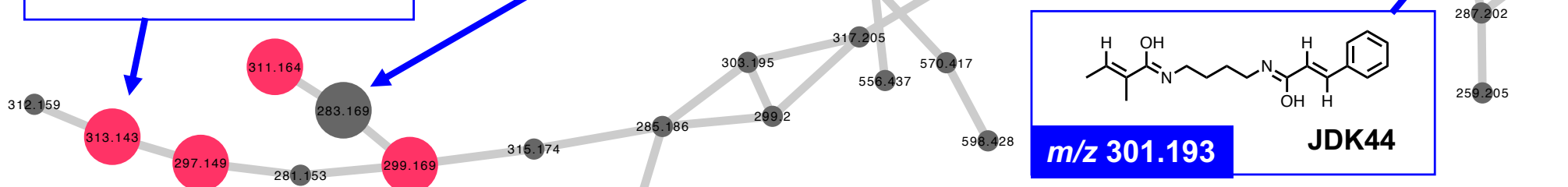
LGC37



MTG27



PSH50

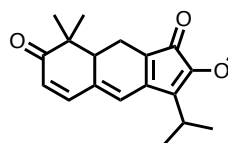


m/z 301.193

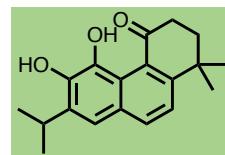
JDK44

Match with Labiatae
(Lamiaceae)

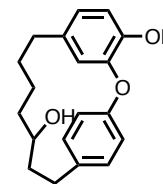
m/z 296.169



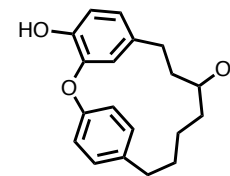
PCX46



JNS32



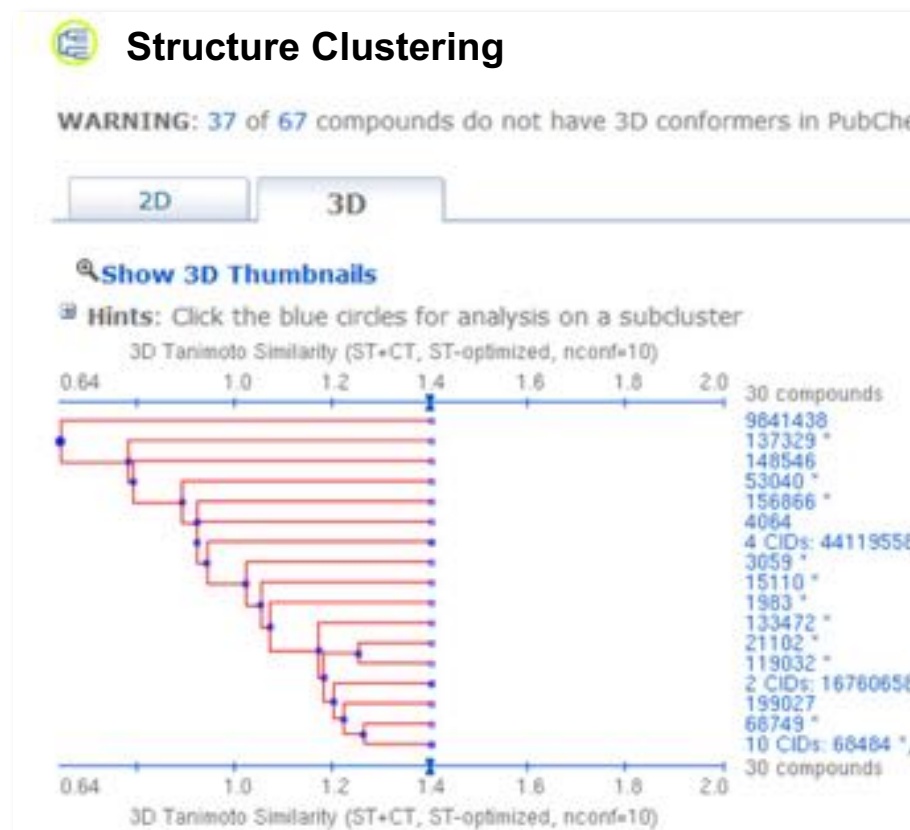
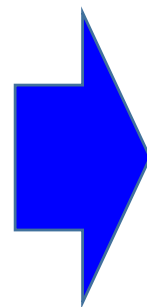
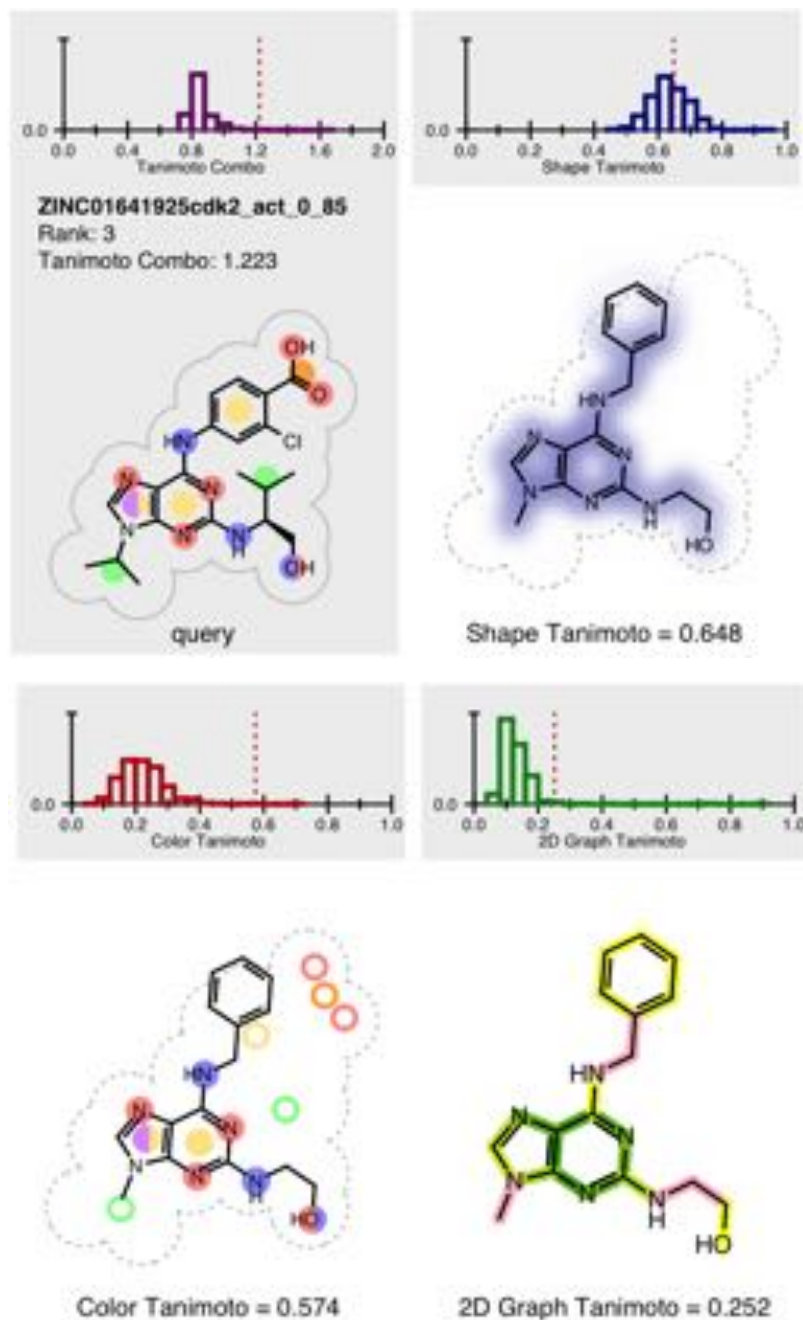
HGT33



MYB70, MYB68

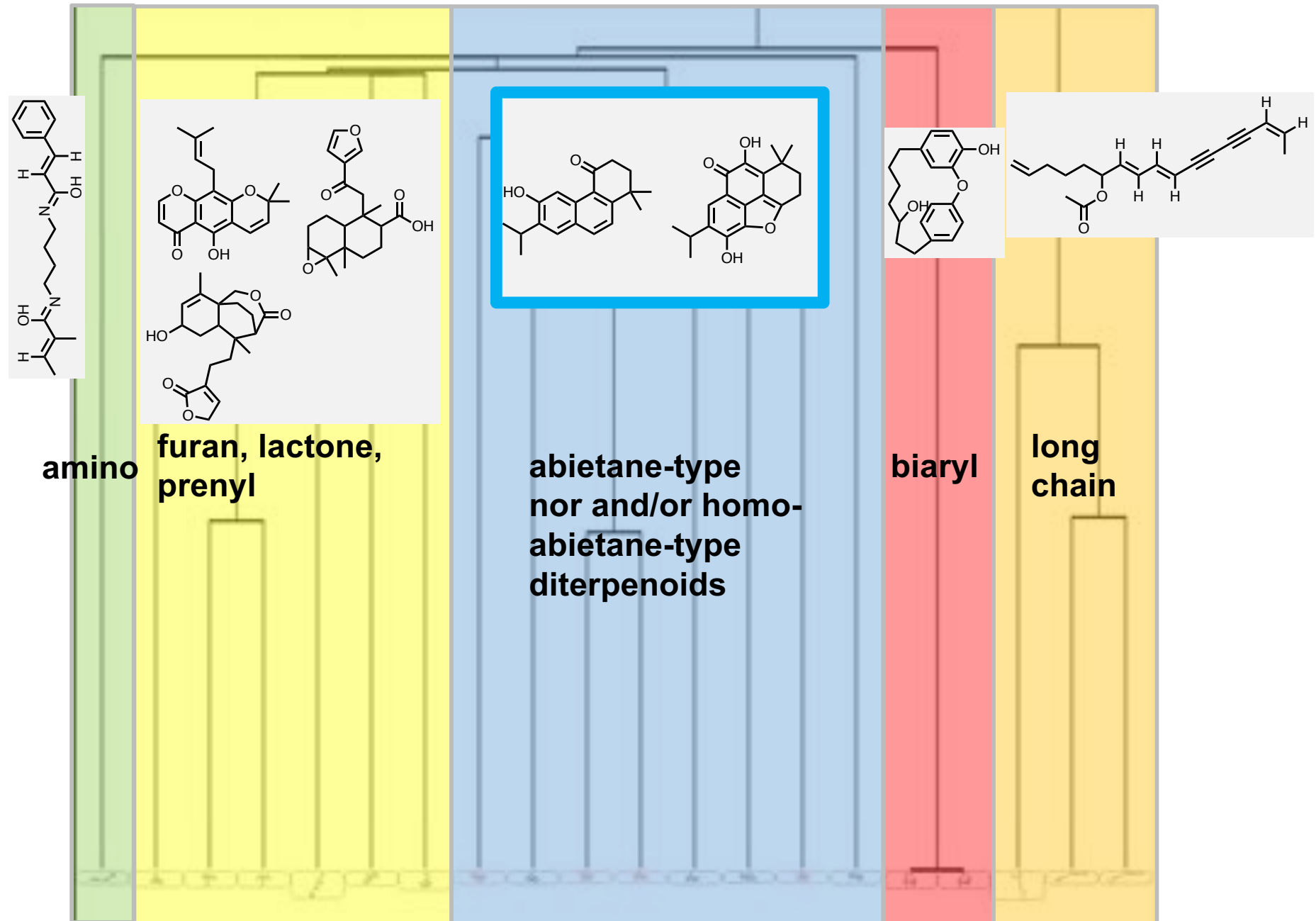
$$S_{\text{total}} = S_1 \times S_2$$

Structural consistency within clusters using Tanimoto scores

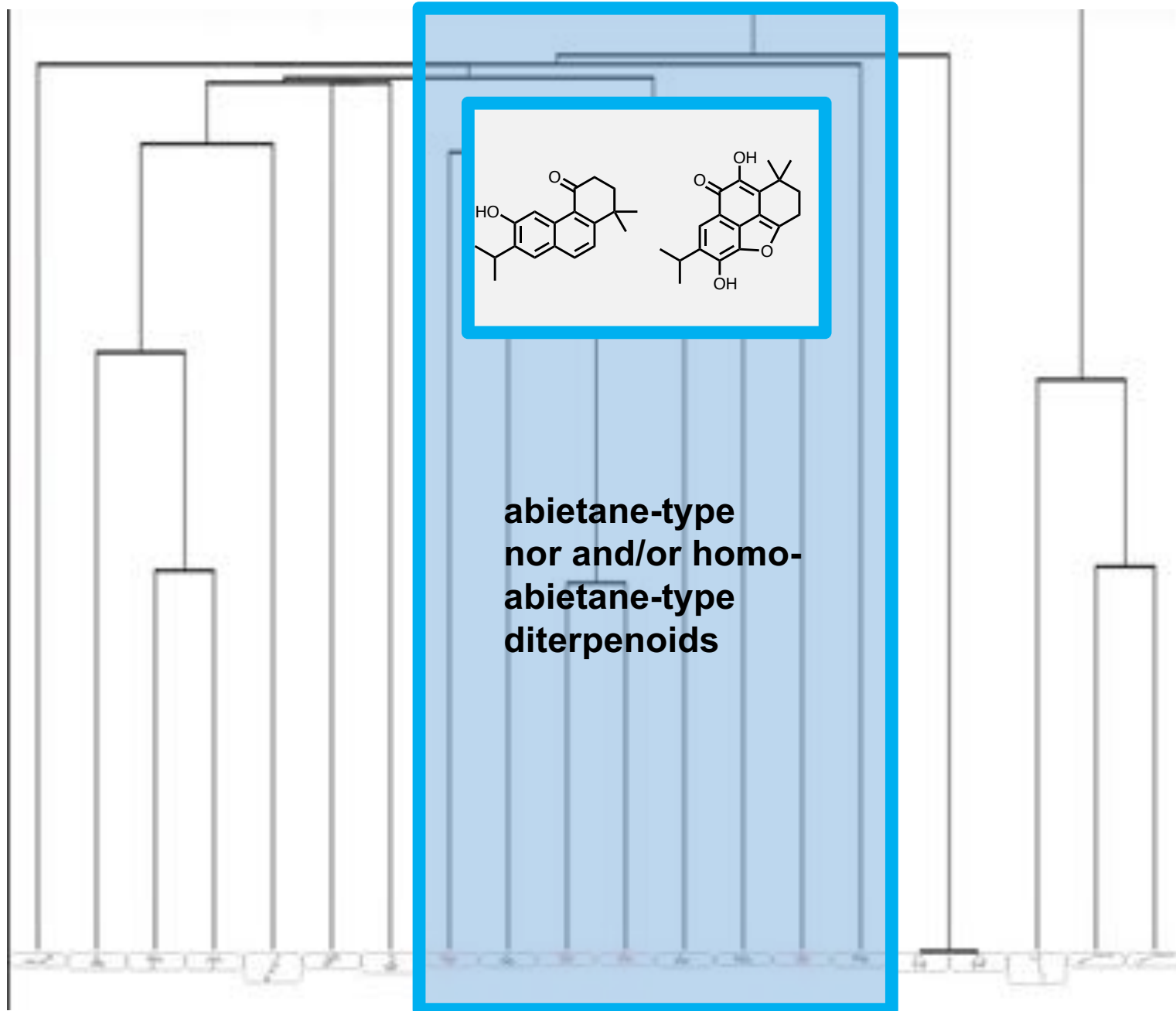


Assigning similarity of structures by substructure search

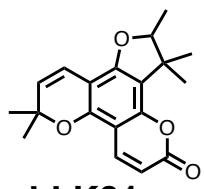
Tanimoto dendrogram of a MN *Salvia* cluster



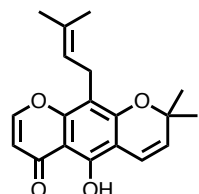
Tanimoto dendrogram of a MN *Salvia* cluster



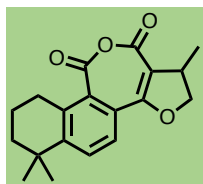
m/z 313.143



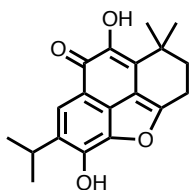
LLK81



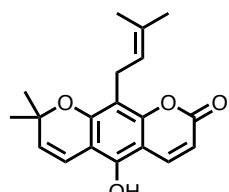
NNF29



HPX83

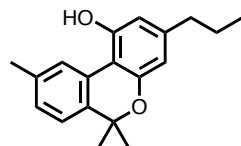


OOW07

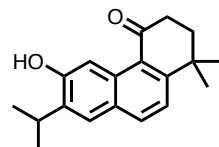


CLT75

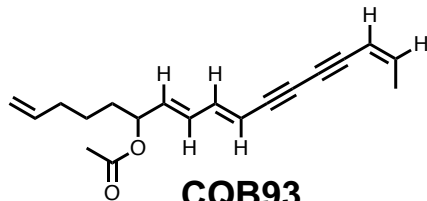
m/z 283.169



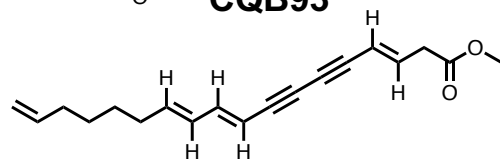
CMK79



QCY13

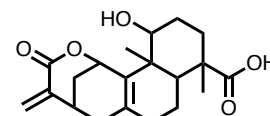


CQB93

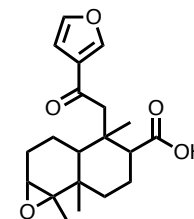


LKT25, LLR47

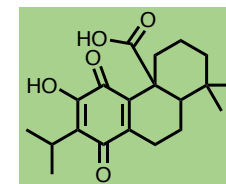
m/z 347.185



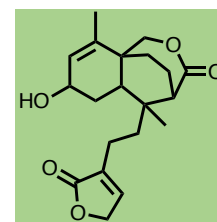
NQN19



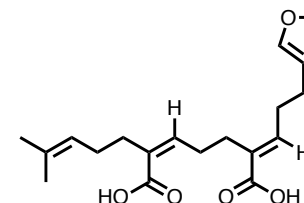
LFM65



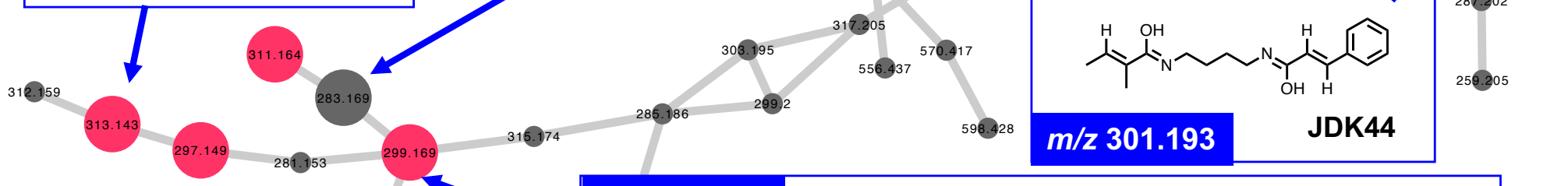
LGC37



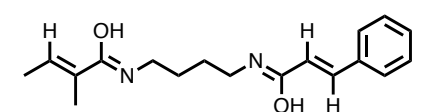
MTG27



PSH50

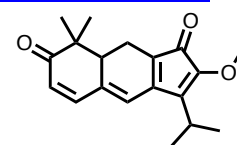


m/z 301.193

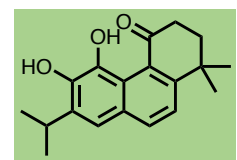


JDK44

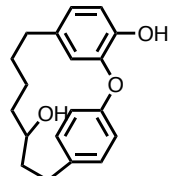
m/z 296.169



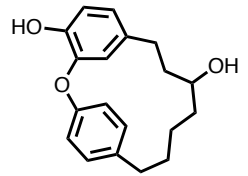
PCX46



JNS32



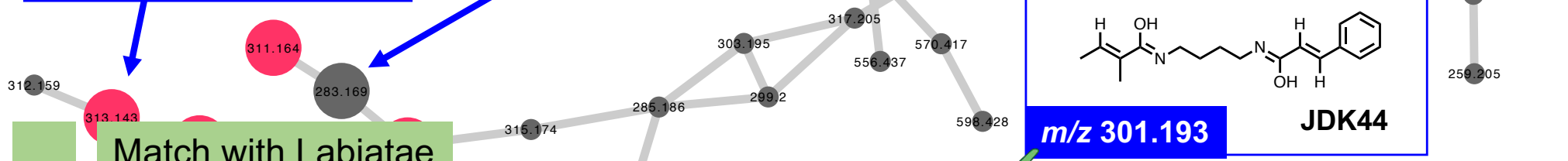
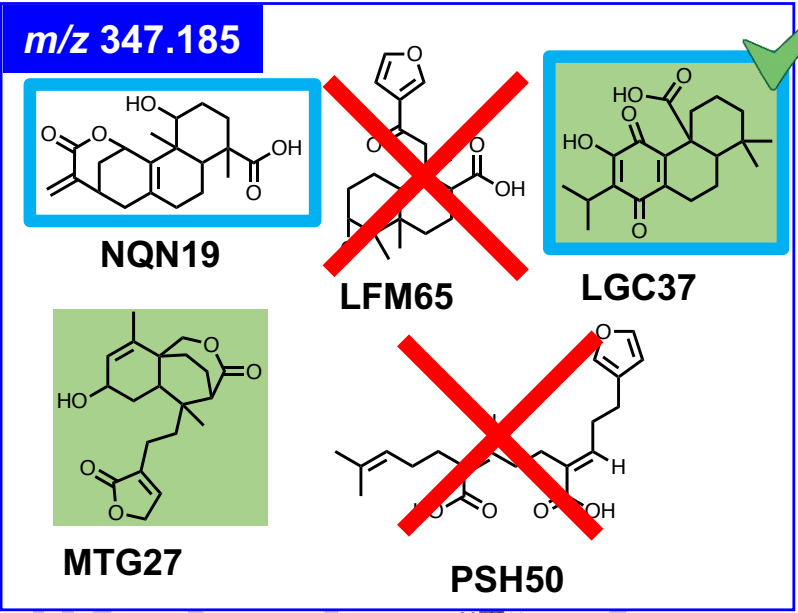
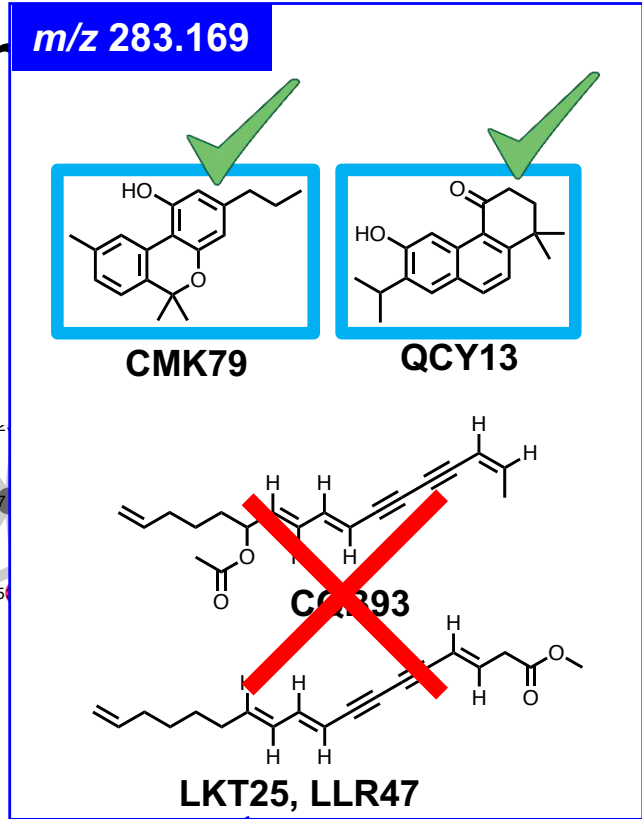
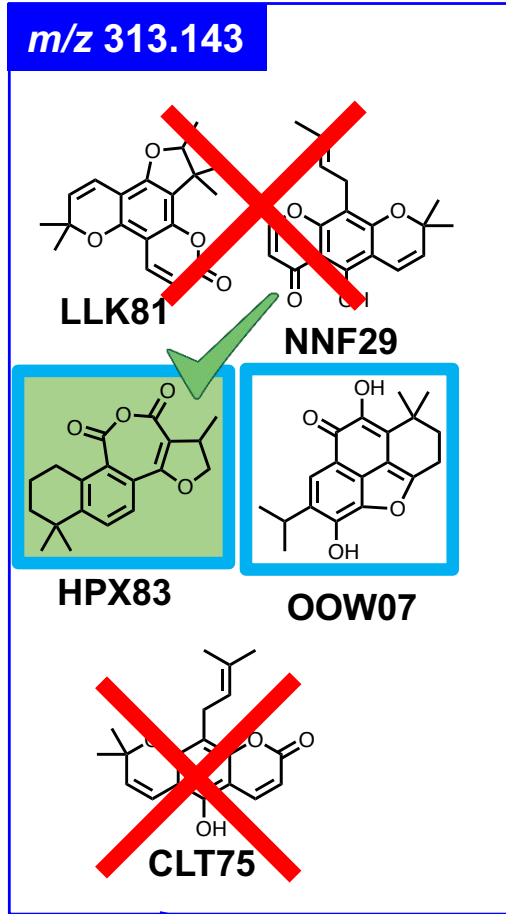
HGT33



MYB70, MYB68

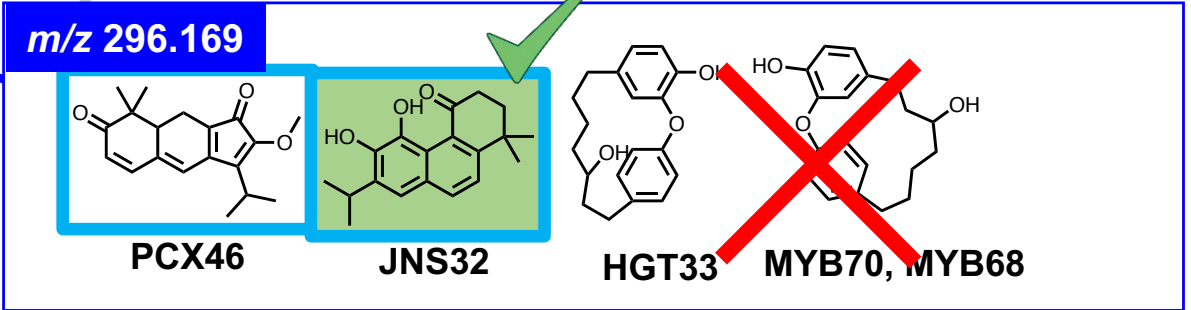
Match with Labiatae (Lamiaceae)

$S_{total} = S_1 \times S_2$



Match with Labiatae (Lamiaceae)

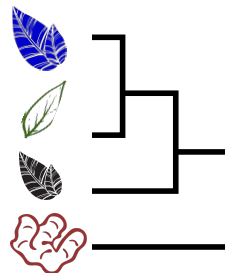
Tanimoto consistency



$$S_{total} = S_1 \times S_2 \times S_3$$

Development of hypothesis metascore

Metabolic consistency
(taxonomical pondering)

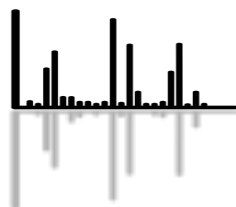


Physico-chemical match
(RT, CCS)

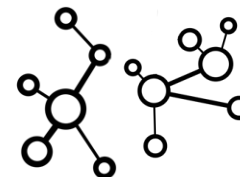


$$S_{\text{total}} = S_1 \times S_2 \times S_3 \times S_4$$

Spectral similarity
score

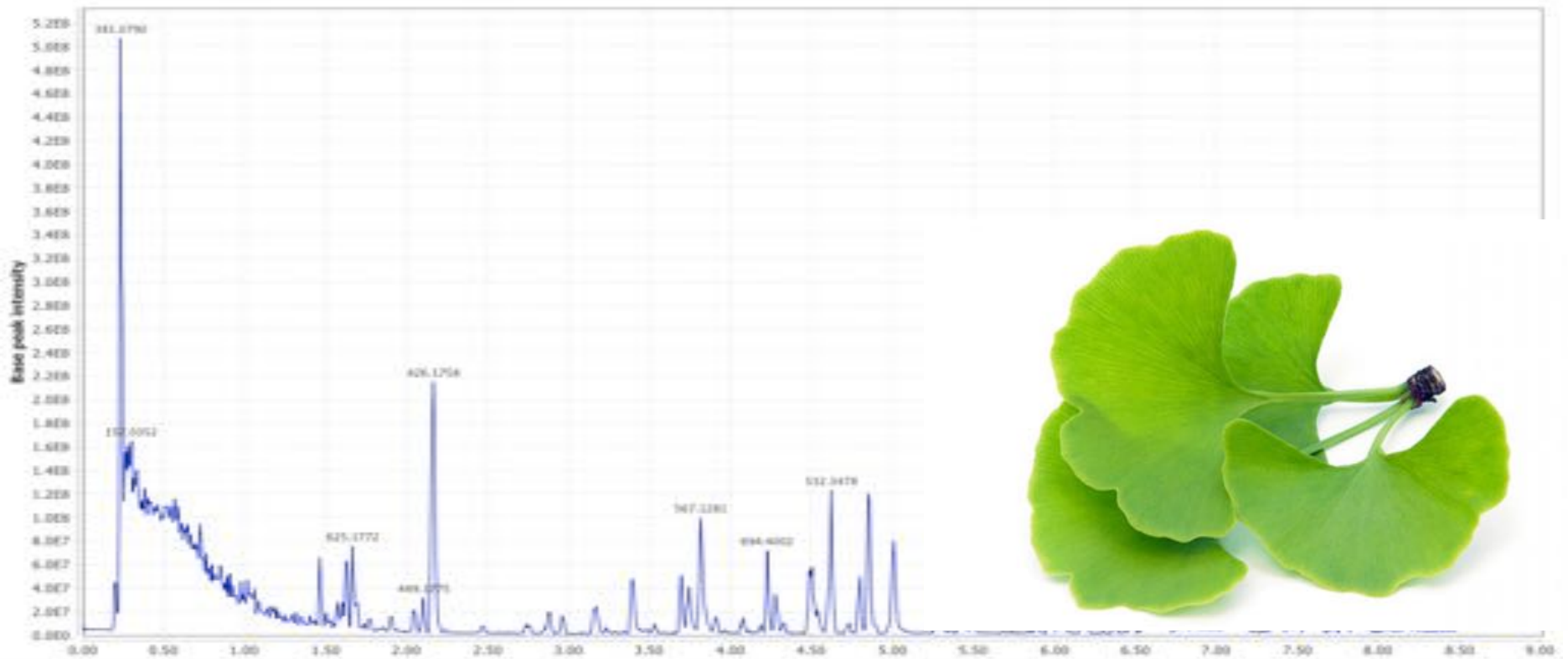


Structural consistency
within cluster
(structural similarity)

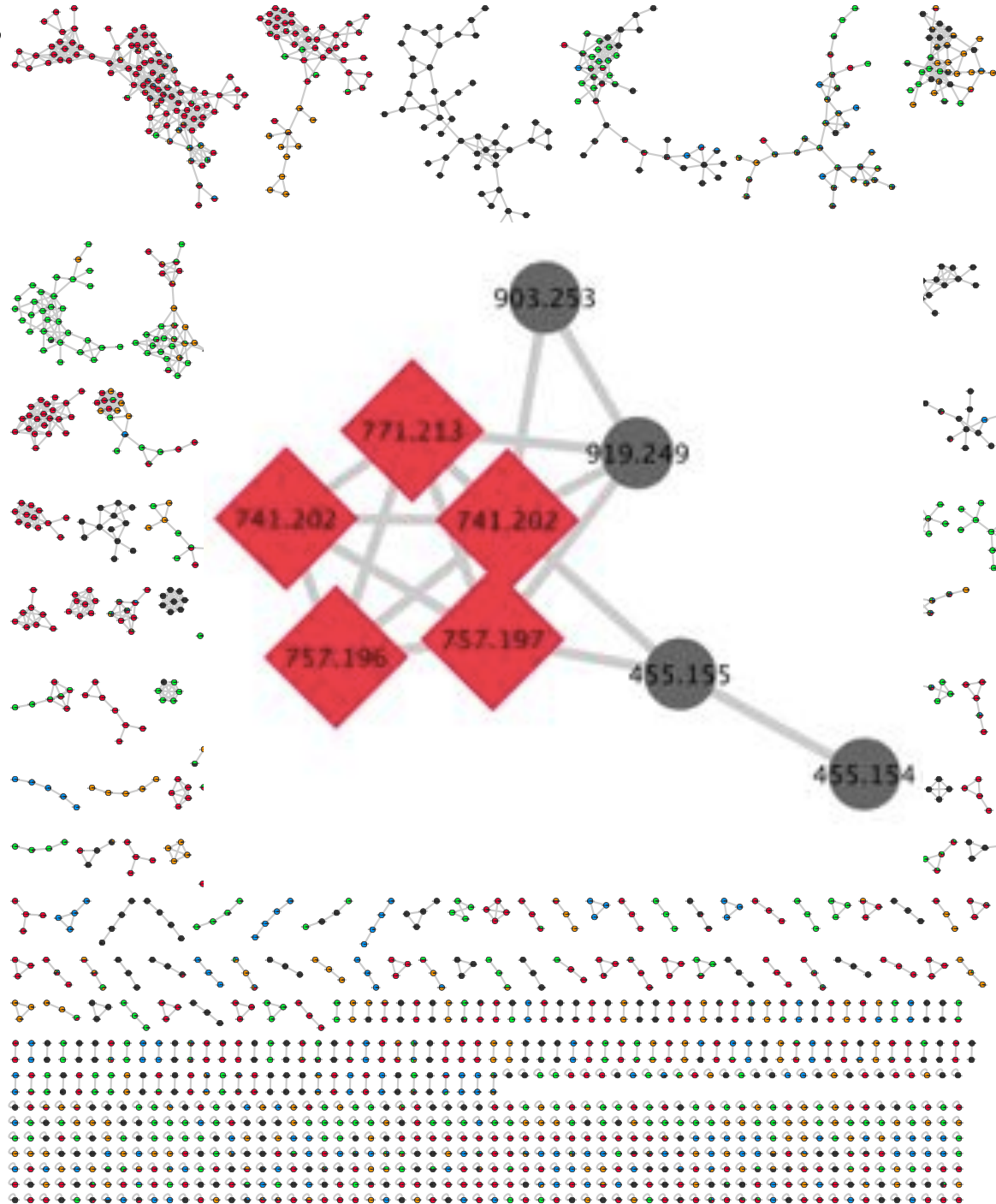




Ginkgo biloba

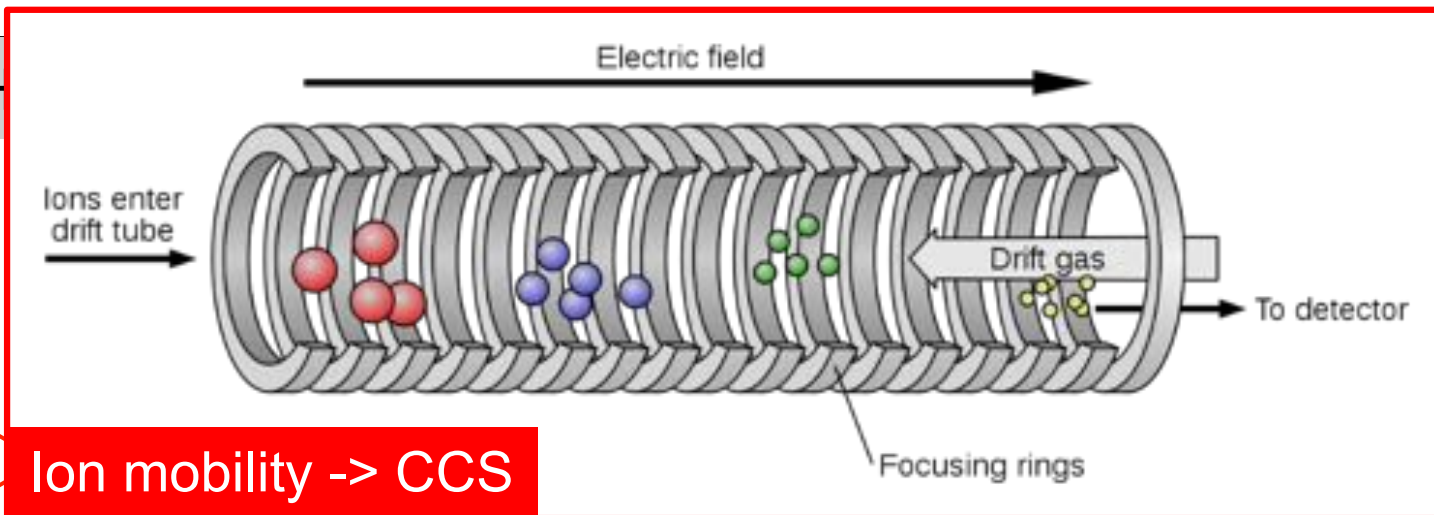


Molecular networking profiles of QCmix

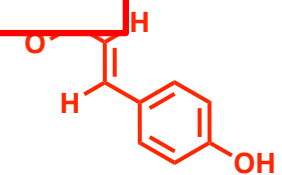
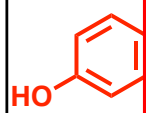


- Cinchona
- Arnica
- Ginseng
- Ginkgo
- Salvia

Molecular network of the MeOH extracts of *Ginkgo biloba*

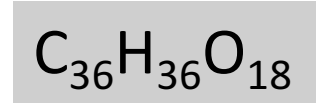
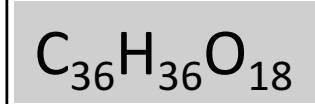


Ion mobility -> CCS

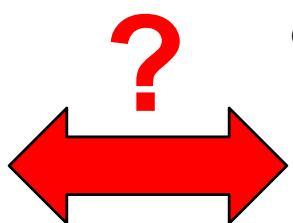


Score: 0.565(rank1)

Score: 0.571(rank1)

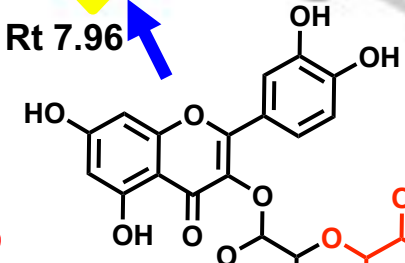
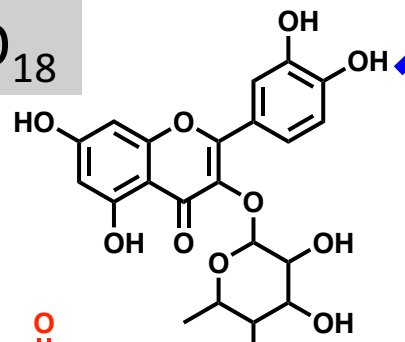
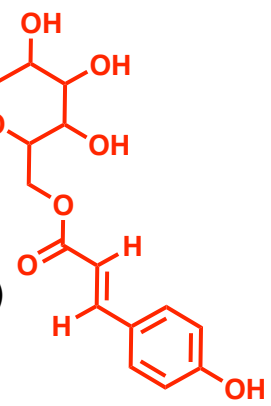
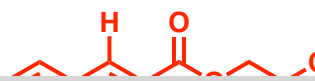


$$S_{total} = S_1 \times S_2 \times S_3 \times S_4$$



RT ?
CCS?

Score: 0.540(rank1)



Rt 9.39

Rt 8.58

Rt 8.80

Rt 7.96

741.202

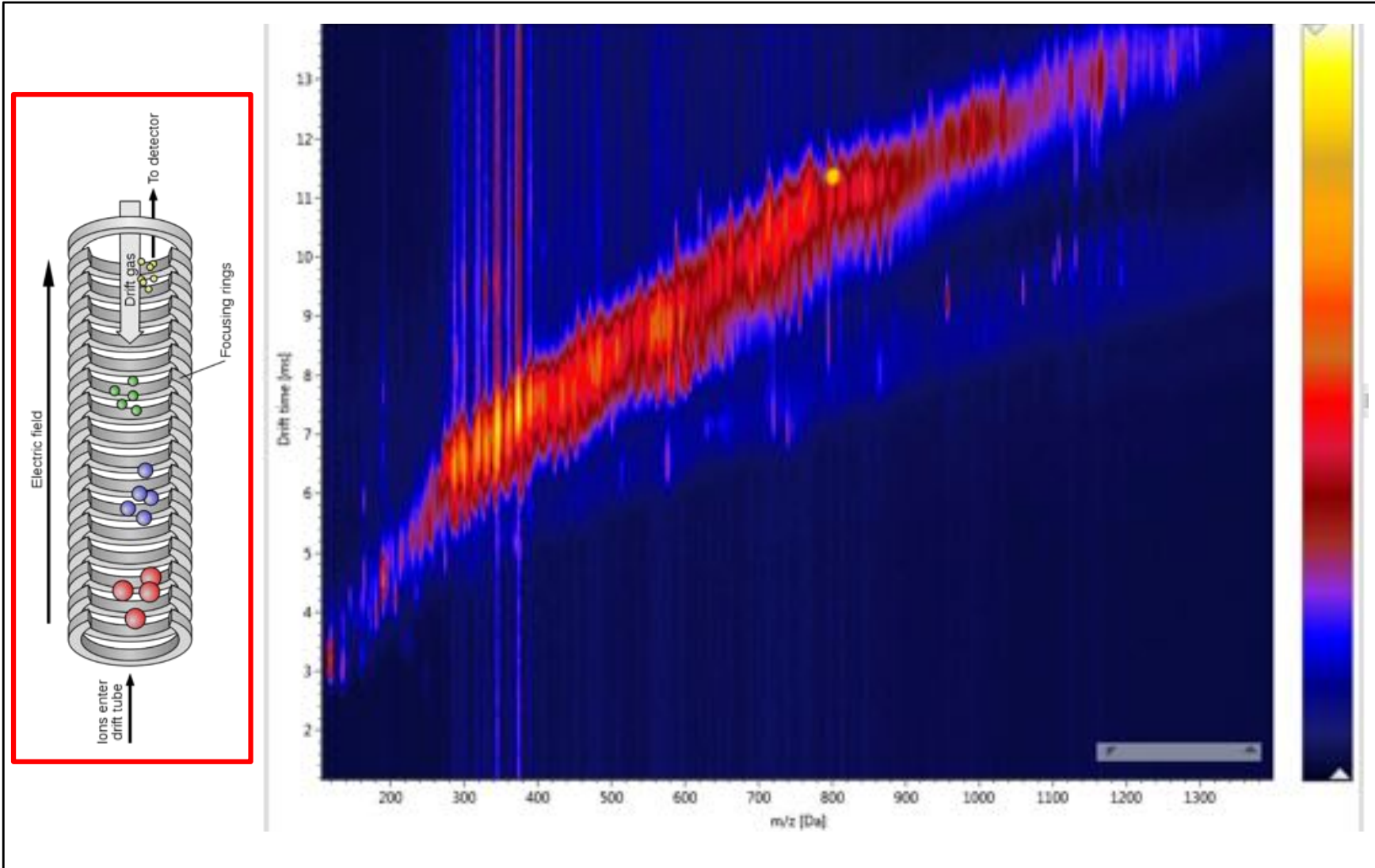
757.196

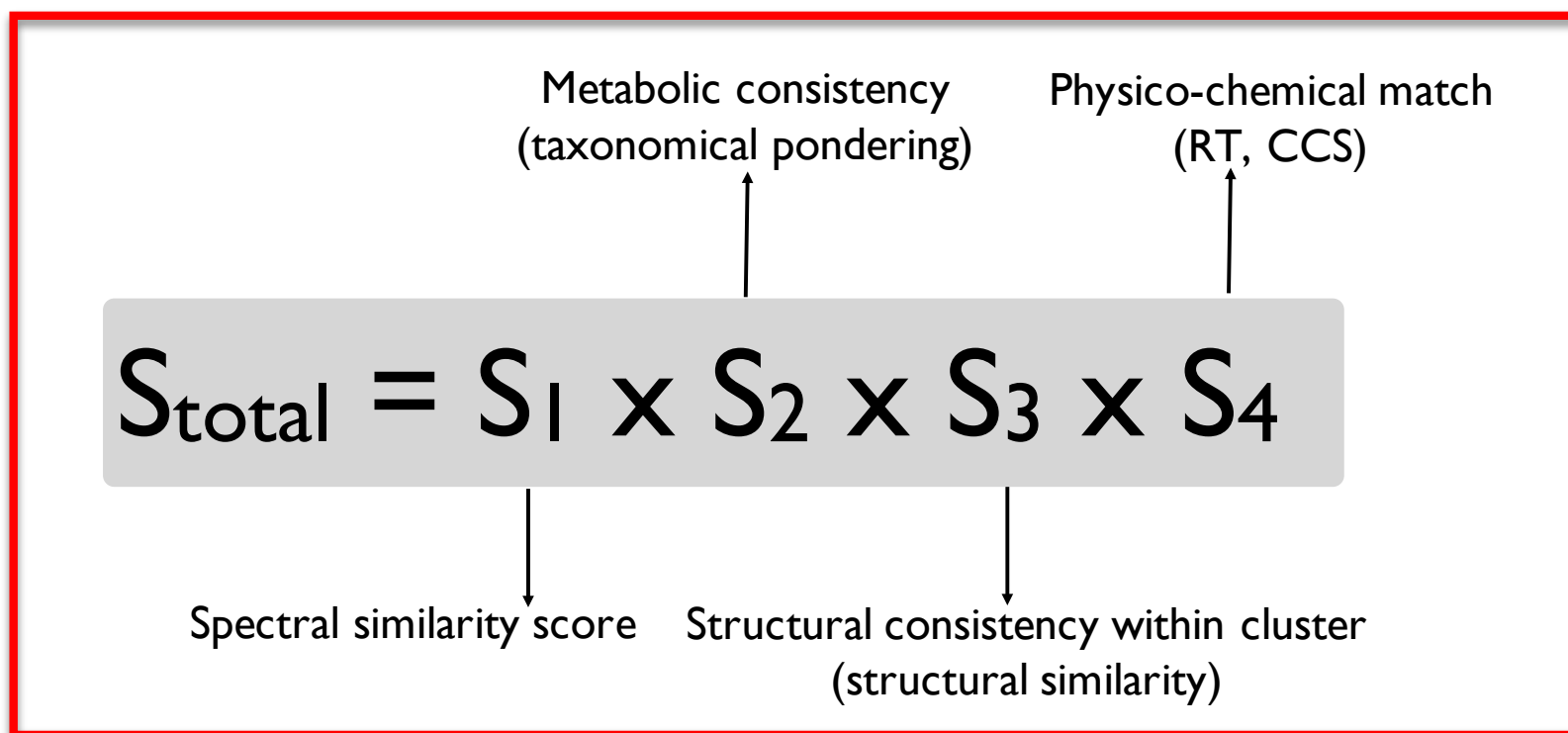
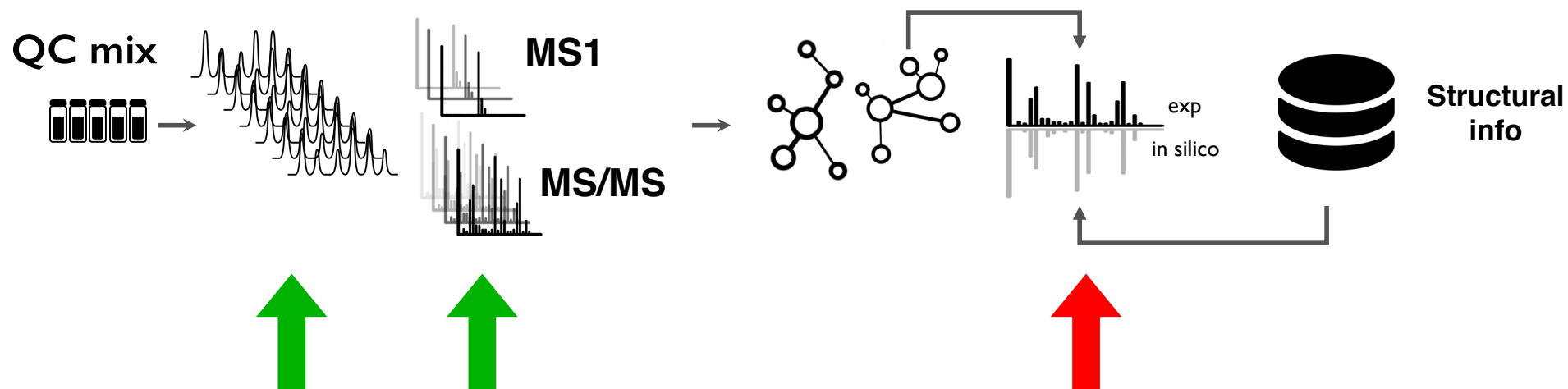
757.197

455.155

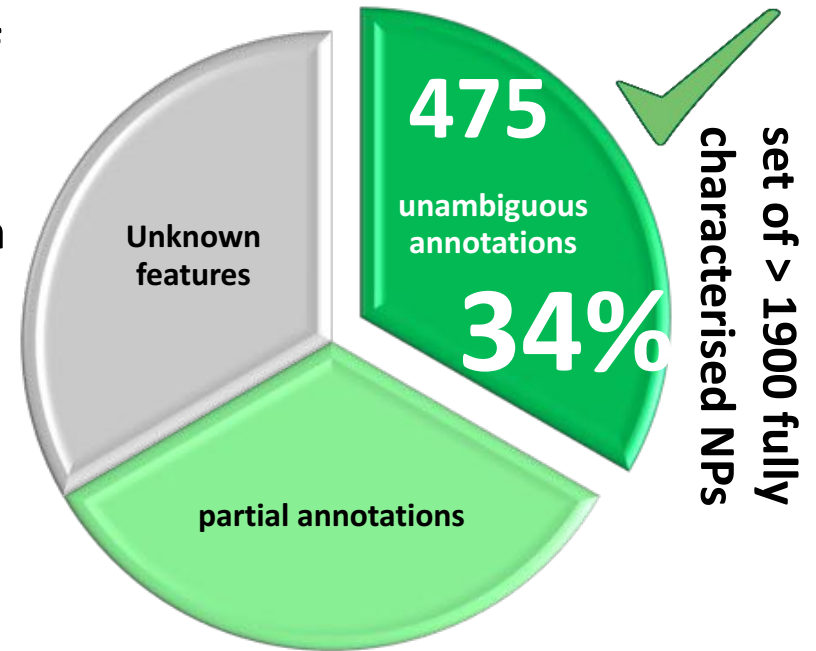
455.154

Ion mobility for IDA HRMS/MS





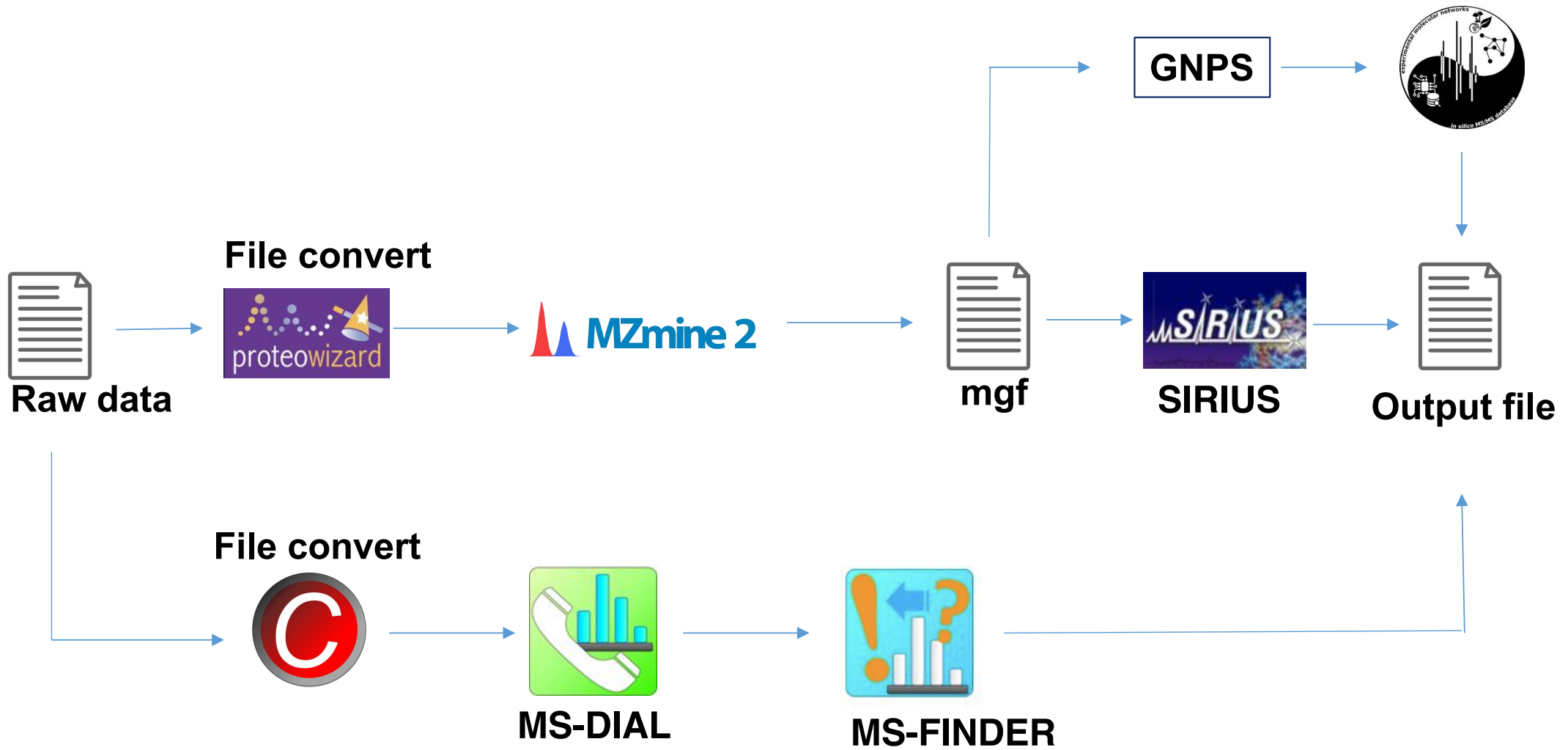
Goal Assessment of metabolite ID capacity of given platform with a given dereplication workflow



To do:

- ❑ appropriate pondering
- ❑ confirm annotation by standard analysis
- ❑ implement the metascore system automatically

Work flow





Output file



SMILES

PlantCyc(Plant)
ChEBI
UNPD(NPs)
NANPDB(NPs)
KNApSAcK(NPs)
DNP



CRC code

— DNP



Short InChIKey

— PubChem (biological db)

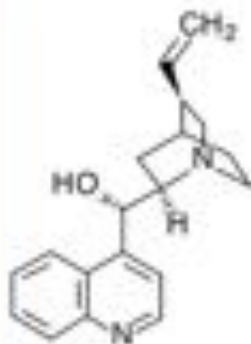


COMPOUND: C06528

Help

Entry	C06528	Compound
Name	Cinchonine; (9S)-Cinchonan-9-ol	
Formula	C ₁₉ H ₂₂ N ₂ O	
Exact mass	294.1732	
Mol weight	294.3908	

Structure



C06528

Mol file

KCF file

DB search

SMILES

O[C@H]([C@@H]1C[C@H]2[C@H](C[N@@]1CC2)C=C)C3=CC=NC4=CC=CC=C34

InChI

InChI=1S/C₁₉H₂₂N₂O/c1-2-13-12-21-10-8-14(13)11-18(21)19(22)16-7-9-20-17-6-4-3-5-15(16)17/h2-7,9,13-14,18-19,22H,1,8,10-12H2/t13-,14-,18-,19-/m0/s1

InChIKey

KMPWYEUPVWOPIM-LSOMNZGLSA-N



Output file



SMILES

PlantCyc(Plant)
 ChEBI
 UNPD(NPs)
 NANPDB(NPs)
 KNApSAcK(NPs)
 DNP



CRC code

— DNP



Short InChIKey

— PubChem (biological db)

Conversion

Short InChIKey

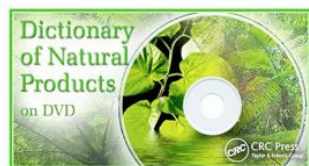
-export information
 -search in MetaData (DNP)



Annotated results file

Work on Excel

← limitation



MS-Finder rank new	MS-Finder (pos)	InChIKey	score S1	Biological_Source	Compound_Types	point_sp	point_genus	point_family	total_point	score S2
1	HDC47-U	YXHVCZZLWZYHSA	8.1314	Constit. of Ginkgo b	DA7300 DB2340 DG0040 VA	0.2	0.2	0.2	1.6	13.0102
	KZV56-S	KEQVPNIOSICVBG-	7.8132	Constit. of Juniperus	VS5550 ZN2000	0	0	0	1	7.8132
	KGG90-Y	DEPBIQJKIVZYES-Z	7.7299	Constit. of Cladiella	VS6440 ZT1000	0	0	0	1	7.7299
	GZG27-V	UEOBFNCQTNUCC	7.7261	Constit. of Ozoroa n	DA7300 DG0040 VA7300 VG	0	0	0	1	7.7261
	NRL91-E	SVXPDLNSXIXYEA-	7.7234	Constit. of Myrosper	VS5800 ZQ3840	0	0	0	1	7.7234



PlantCyc(Plant)
ChEBI
UNPD(NPs)
NANPDB(NPs)
KNApSAcK(NPs)
DNP



— DNP



— PubChem (biological db)



Short InChIKey



CSV



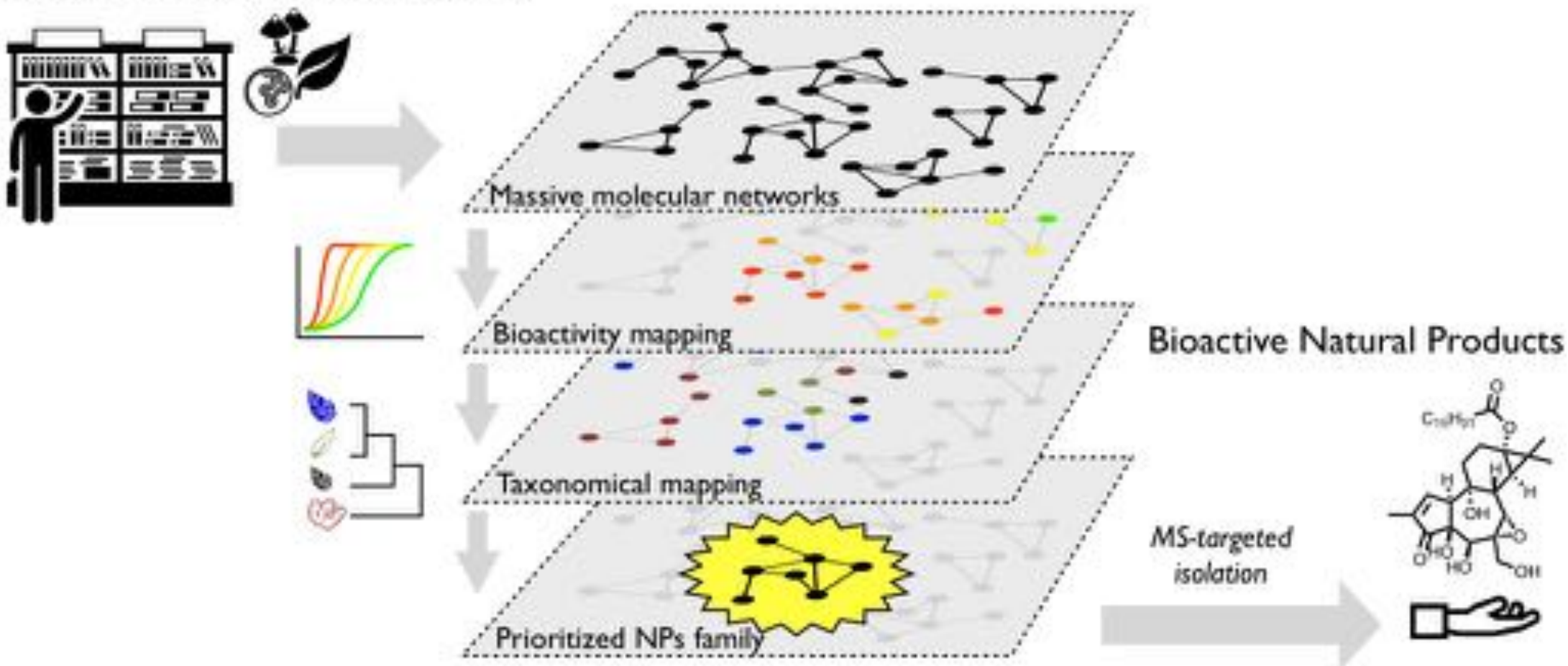
InstantJchem

Annotated
results file

...	Cdid	Structure	Mol Weight	Formula	File path	File name	Title	MS1 count	MSMS count	PRECURSO	PRECURSO	Rank	Structure rank	Total score	Databases	Ontology	Short InChIKey
1	1		328.50	C22H3...	C:\User...	Peak I...	Unkno...	752	54	329.25	[M+H] ⁺	1	C0001...	6.93	ECMDB...	NA	PHENP...
2	2		328.50	C22H3...	C:\User...	Peak I...	Unkno...	752	54	329.25	[M+H] ⁺	2	O-15~...	6.90	ChEBI...	NA	QGNJR...
3	3		328.50	C22H3...	C:\User...	Peak I...	Unkno...	752	54	329.25	[M+H] ⁺	3	C0001...	6.85	ECMDB...	NA	DVQKS...
4	4		328.50	C22H3...	C:\User...	Peak I...	Unkno...	752	54	329.25	[M+H] ⁺	4	UNPD3...	6.76	ECMDB...	NA	CGCM...
5	5		328.50	C22H3...	C:\User...	Peak I...	Unkno...	752	54	329.25	[M+H] ⁺	5	HMDB...	6.74	BMDB...	NA	MBMB...

Bioactive Natural Products prioritization pipeline

Natural Products Extracts Library

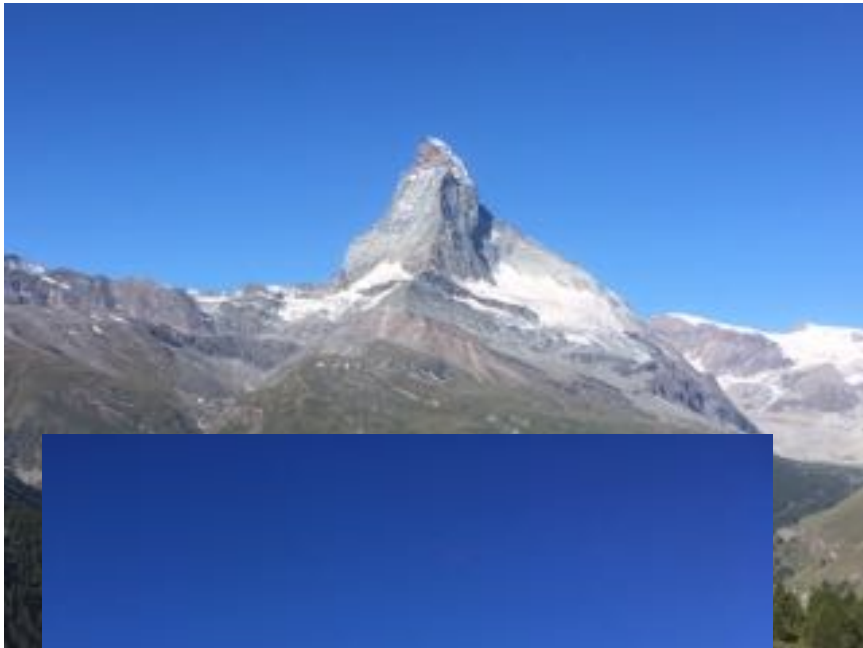






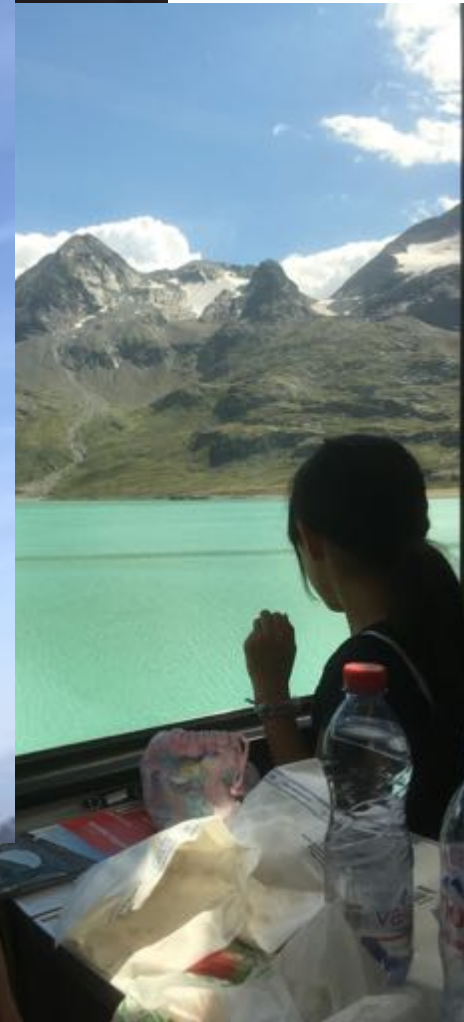
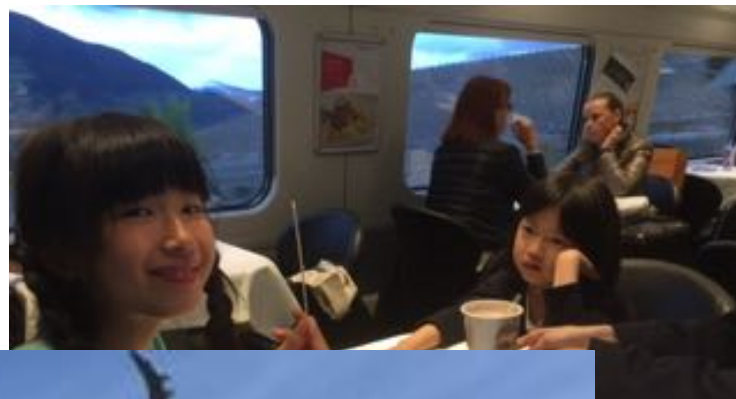




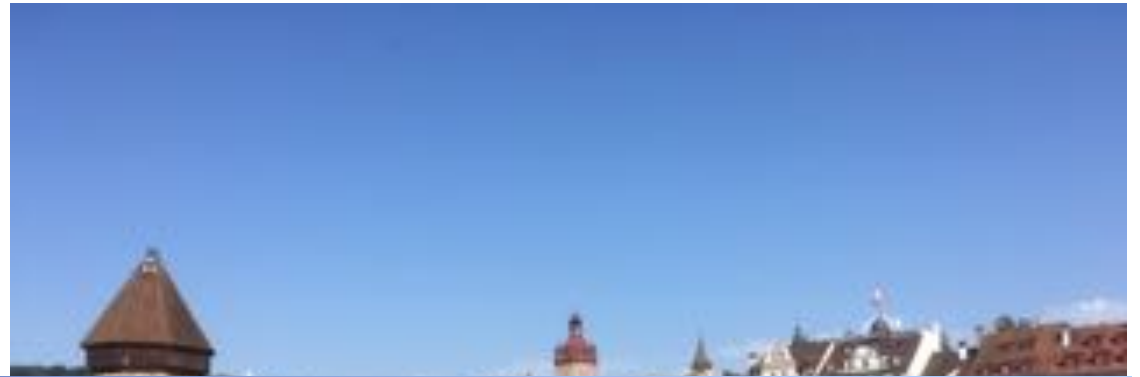






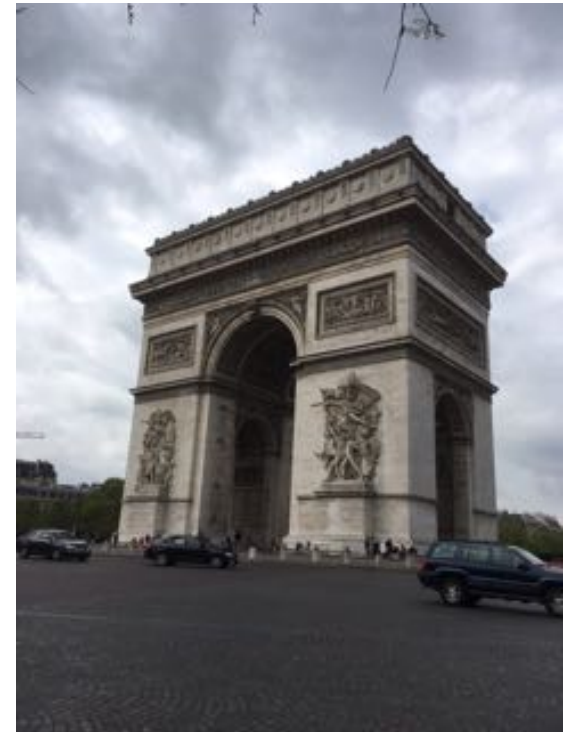
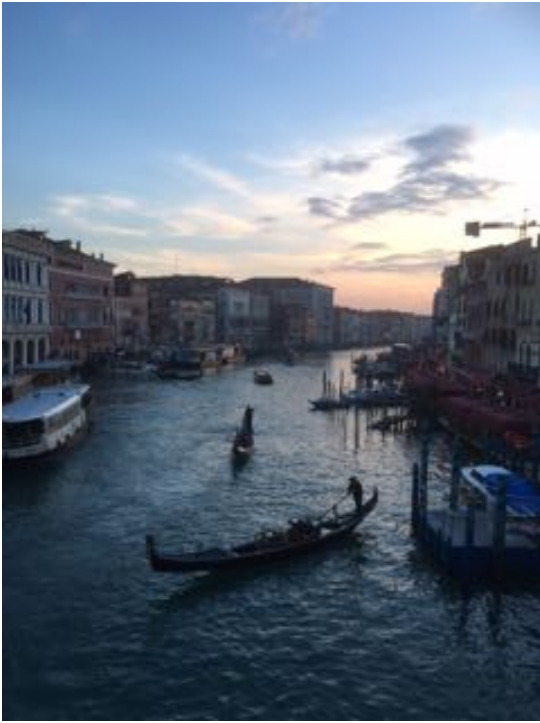








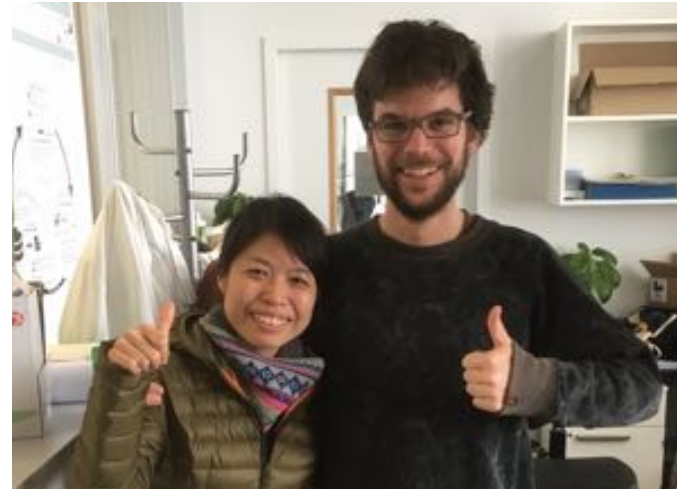




Acknowledgement



Prof. Jean-Luc Wolfender



Dr. Pierre-Marie Allard

**Prof. Yoshiyasu Fukuyama
Prof. Yoshinori Asakawa
Dr. Kenichi Harada
Dr. Masato Murasaki**

**Mrs Corinne Galland (Welcome center)
My friends and my family**

公益財団法人 山田科学振興財団
公益財団法人 永井記念薬学国際交流財団



The best friends @ Jean-Luc's lab



Merci beaucoup!