Modern Mass Spec tools for food profiling to ensure authenticity and quality

Dr. Joerg Riener
EMEA GC/MS Food Segment Scientist

Agilent Technologies

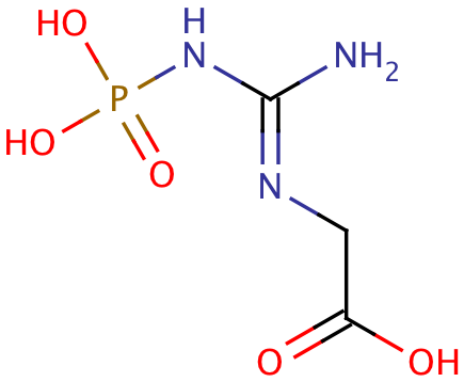
Chemometric Multivariate Analysis

Goal: Detect observed phenotype differences at the chemical level

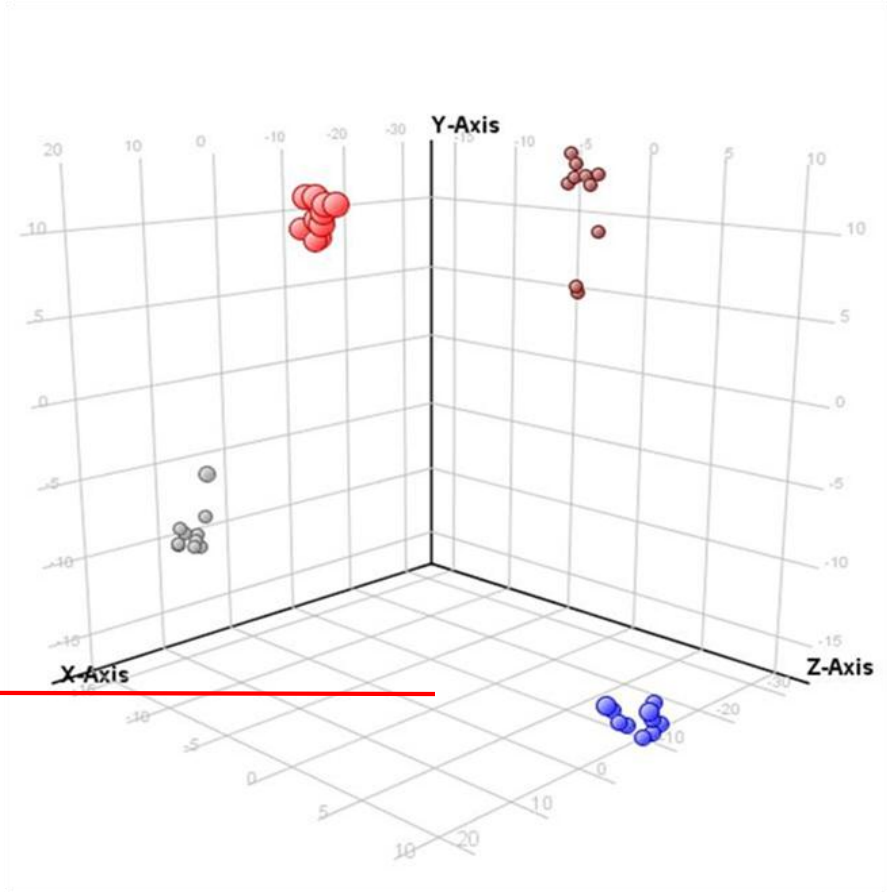
Multivariant analysis is a means to detect differences between phenotypes at the chemical level

Multivariant analysis can rank metabolite importance to phenotype

Does not explain underlying biology



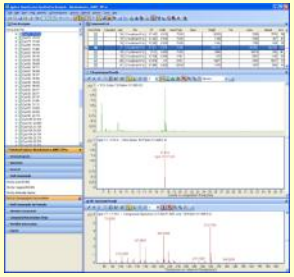
Phosphoguanidinoacetate



Separate & Detect	Feature Finding Quantitate	Alignment & Statistics	Identify	Pathways

GCMS

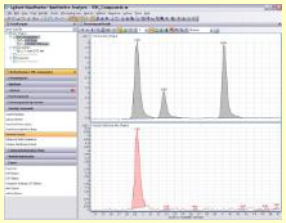
GC-QTOF/MSD
GC-QQQ



MassHunter Qual
AMDIS* or *Find by chromatographic deconvolution

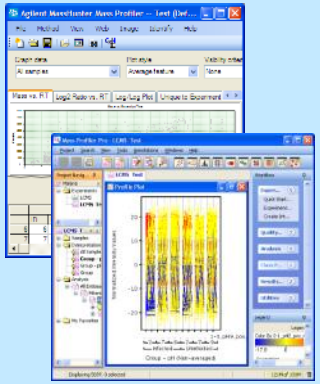
LCMS

LC-TOF/QTOF
LC-QQQ

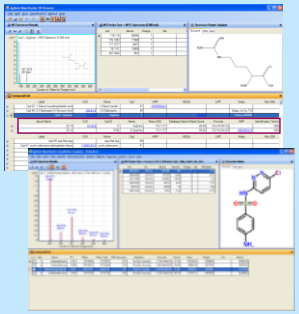


MassHunter Qual
***MFE*,**
***Find by Formula*,**
Find by Ion

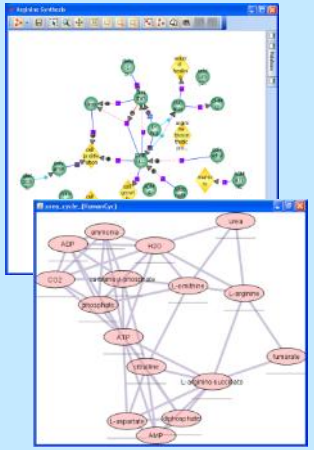
Mass Profiler (MP)



Mass Profiler Professional (MPP)



ID Browser



Pathway module Cytoscape

LC/MS and GC/MS systems are already used to identify organic and inorganic contaminants in feed, food and other food stuff.

Chemometric – Un-Targeted or Targeted Analysis

Un-Targeted - Looking for all metabolites and chemicals

- Data is acquired in full scan mode
- Metabolite tracking uses retention time and EI spectra
- Statistics used to find interesting metabolites (features)
- Features must be identified to make biological sense

Targeted - Looking for known metabolites only

- Data is acquired in SIM (MS) or MRM (MS/MS) mode
- Absolute quantitation - Need external and internal standards
- Statistics used to confirm interesting metabolites

Sample Analysis – GC/MS

GC/MS – Best for routine analysis

Advantages

- No ionization suppression
- Identification by EI searchable libraries

Disadvantages

- Volatile analytes - derivatization required
- Molecular ion often missing

GC/MS/MS – Best for targeted

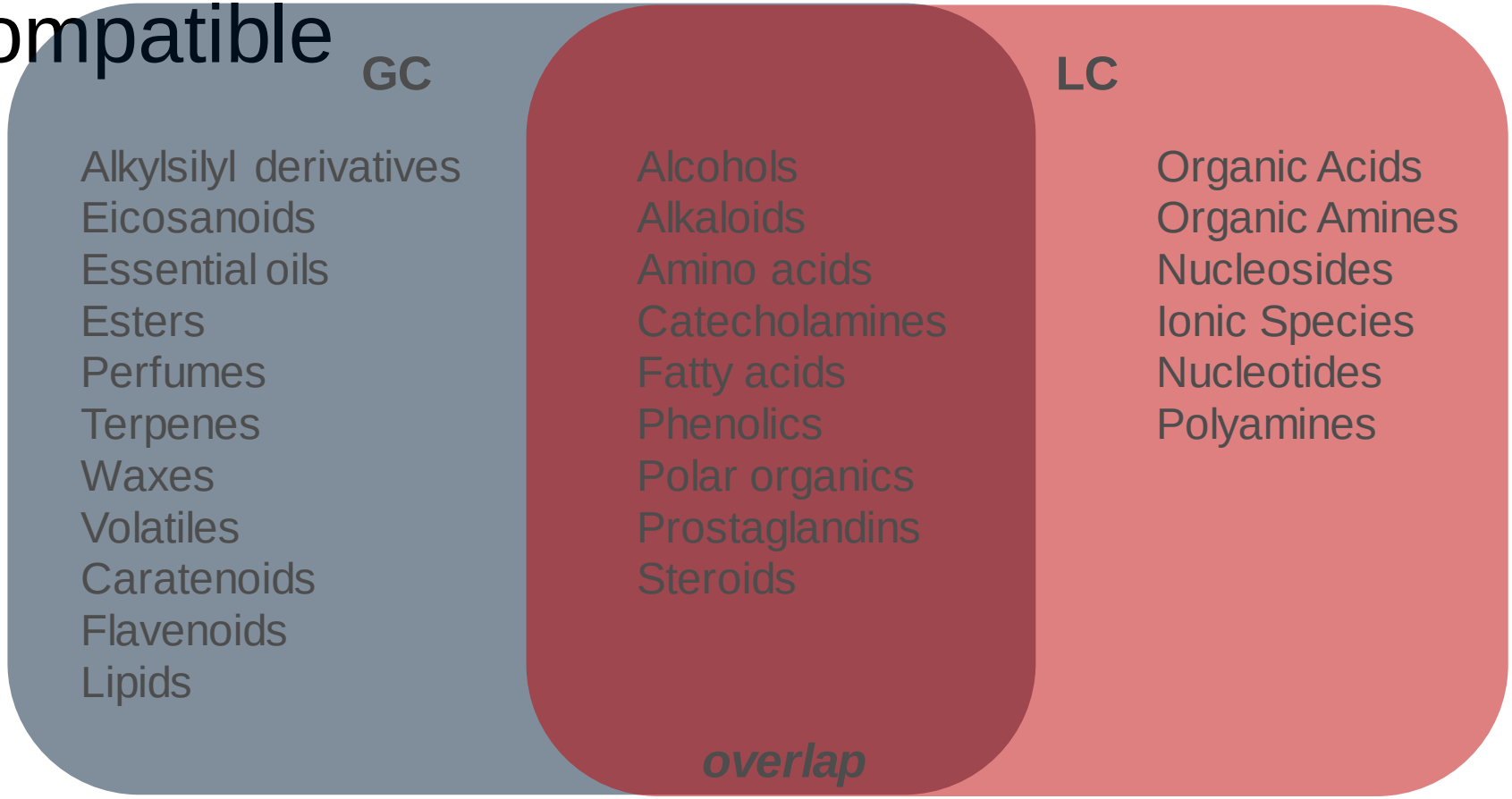
Advantages

- Lower detection limits
- Greater selectivity

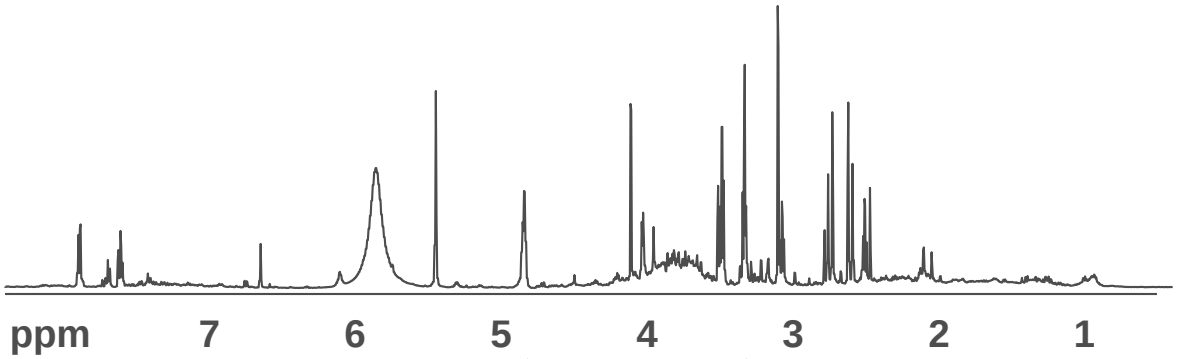
Disadvantages

- Volatile analytes - derivatization required
- Molecular ion often missing

Classes of chemicals and the analytical techniques with which they are most compatible

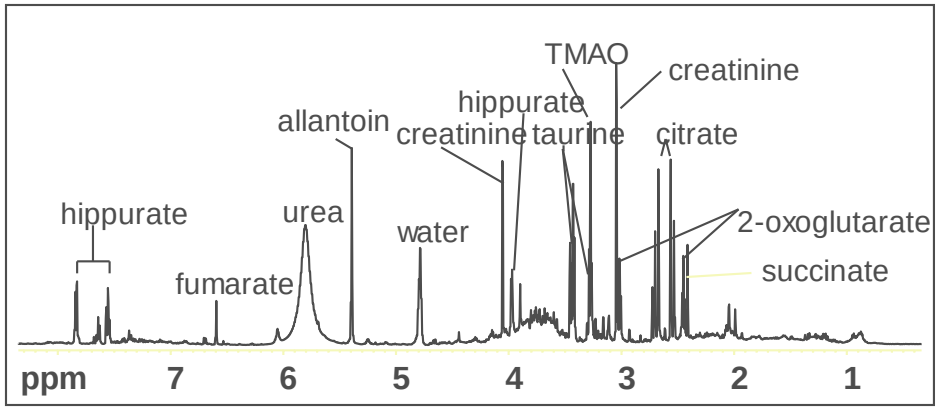


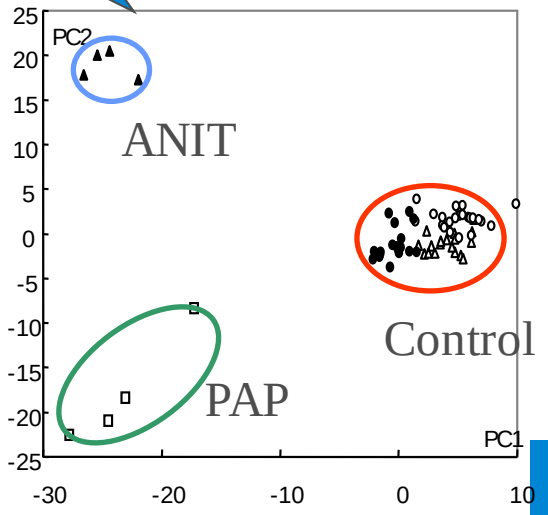
2 Routes to Chemometric



Quantitative Methods

Chemometric (Pattern) Methods





Main Challenge

A CASE STUDY ON CHAMOMILE FLOWERS AND OILS

What is Chamomile?

Chamomile: one name, different botanical species

German Chamomile

Matricaria chamomilla L.

Roman Chamomile

Chamaemelum nobile (L.) All.

Juhua

Chrysanthemum morifolium Ramat.

All chamomile species belong to the same *Asteraceae (Compositae)* family but different genera

Why Is Chamomile Important?

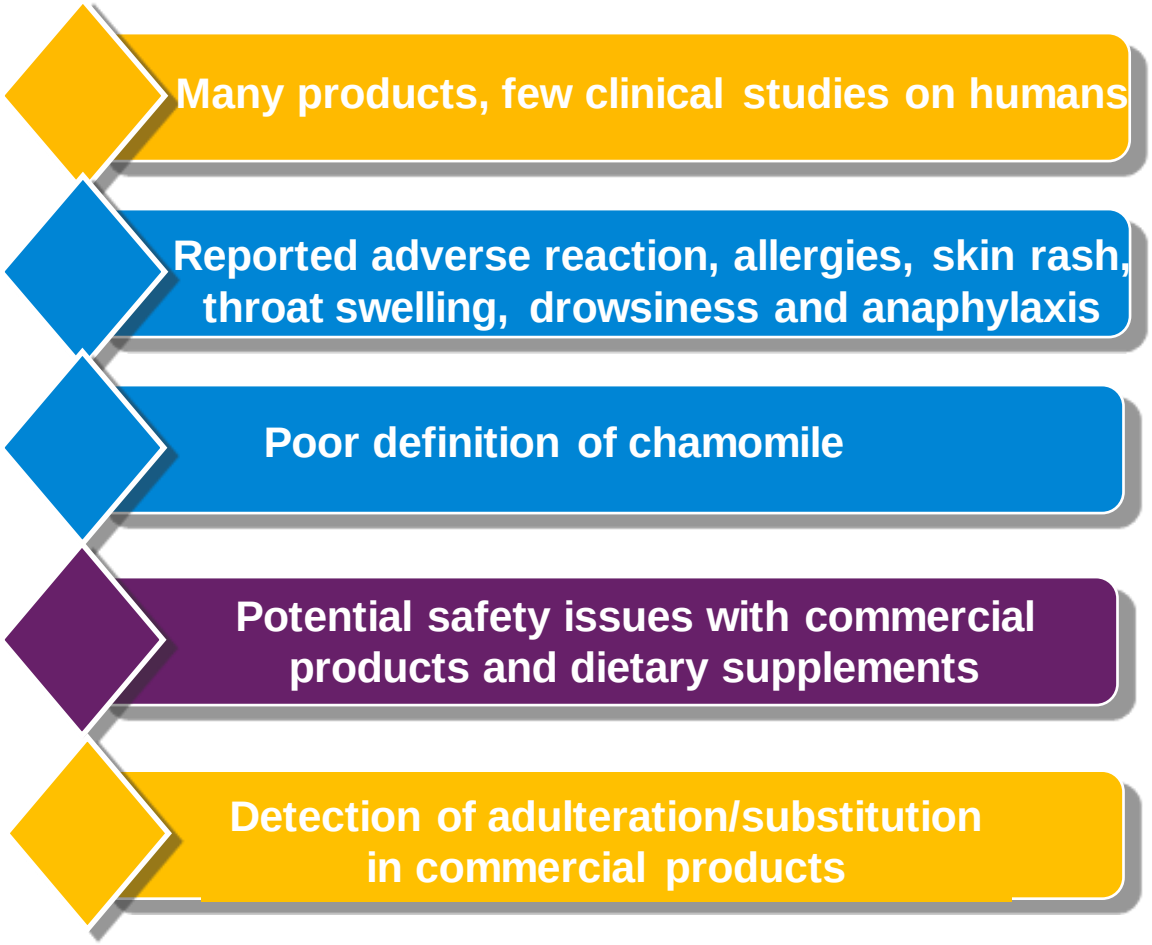
Beneficial Properties

- Anti-inflammatory
- Hay fever
- Wound healing and burn relief
- Gastro-intestinal disorders
- Tooth ache, ear ache
- Ulcers
- Antibacterial
- Antifungal
- Infections
- Common cold
- Laryngitis
- Anxiety relief
- Sleep disorders
- Cardiovascular diseases
- Muscle spasm
- Rheumatic pain
- Arthritis

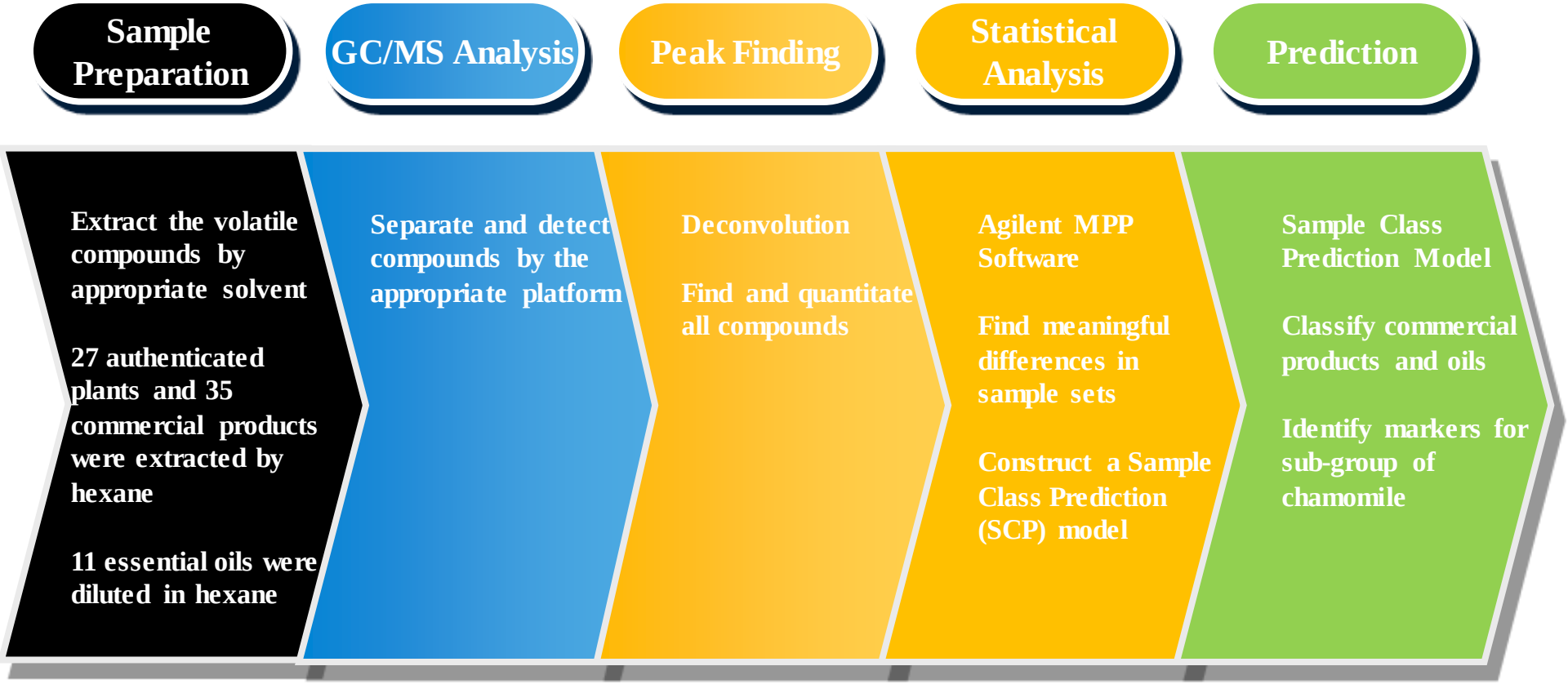
Commercial Products

- Cosmetics
- Aromatherapy
- Teas
- Lotions
- Herbal Beer
- Gargles
- Shampoos

Why Did We Study Chamomile?



What Workflow Did We Use?

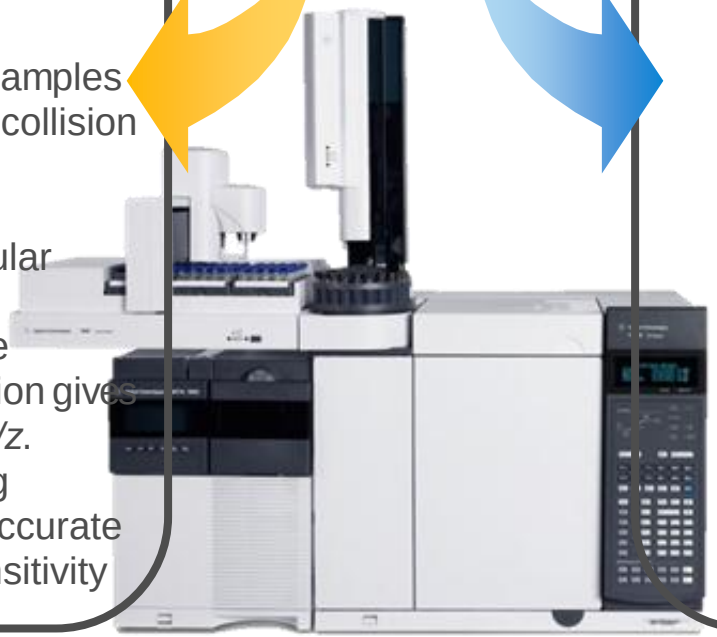


How Did We Study Chamomile?

GC/MS Analysis

Advantages:

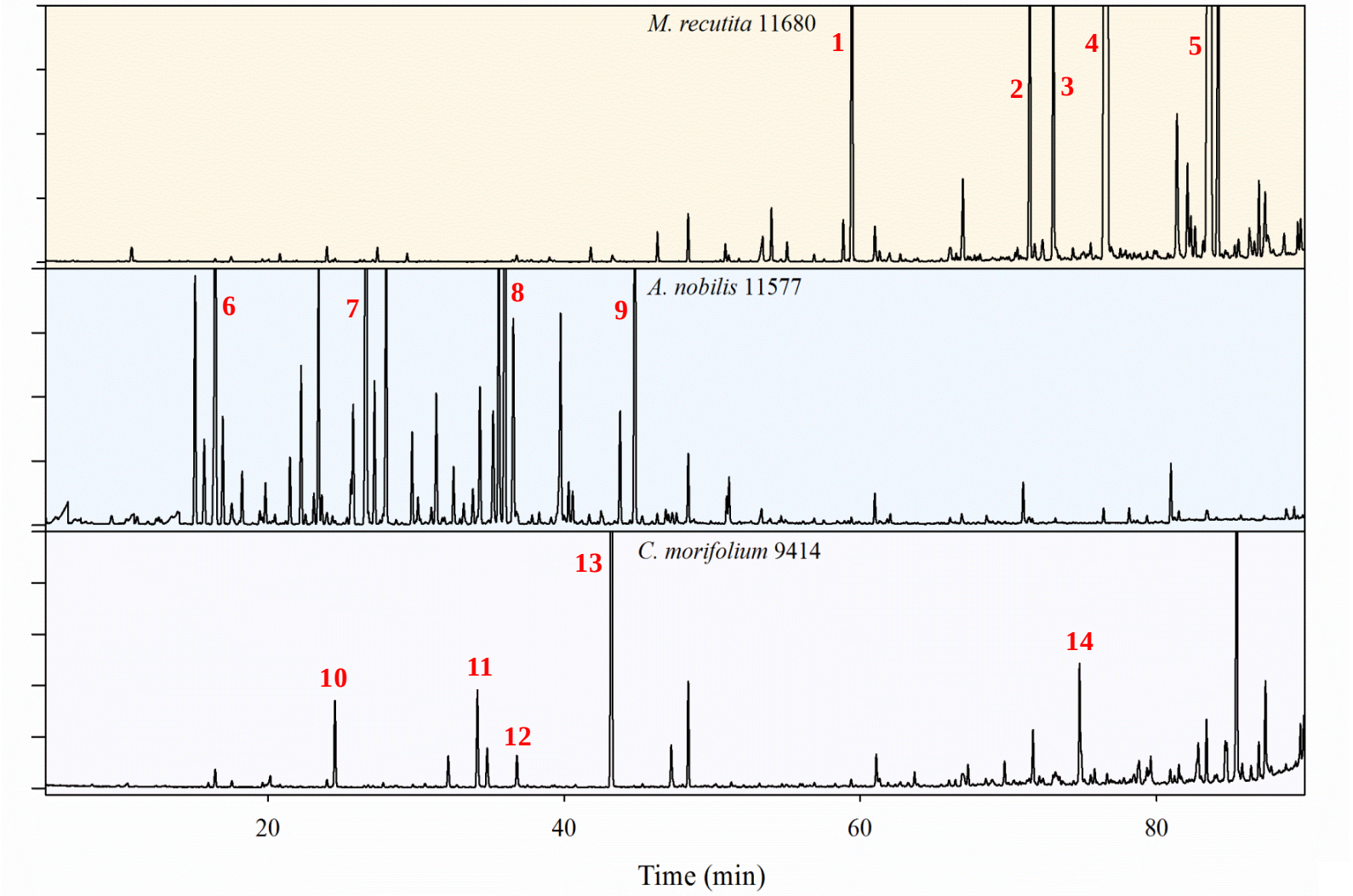
- GC resolves complex samples
- Ionization produced by collision with 70 eV electrons
- Fragmentation pattern characteristic of molecular structure
- Many libraries available
- GC/QToF instrumentation gives 2-5 ppm accuracy in *m/z*.
- Selected ion monitoring GC/QQQ allows very accurate quantitation at high sensitivity



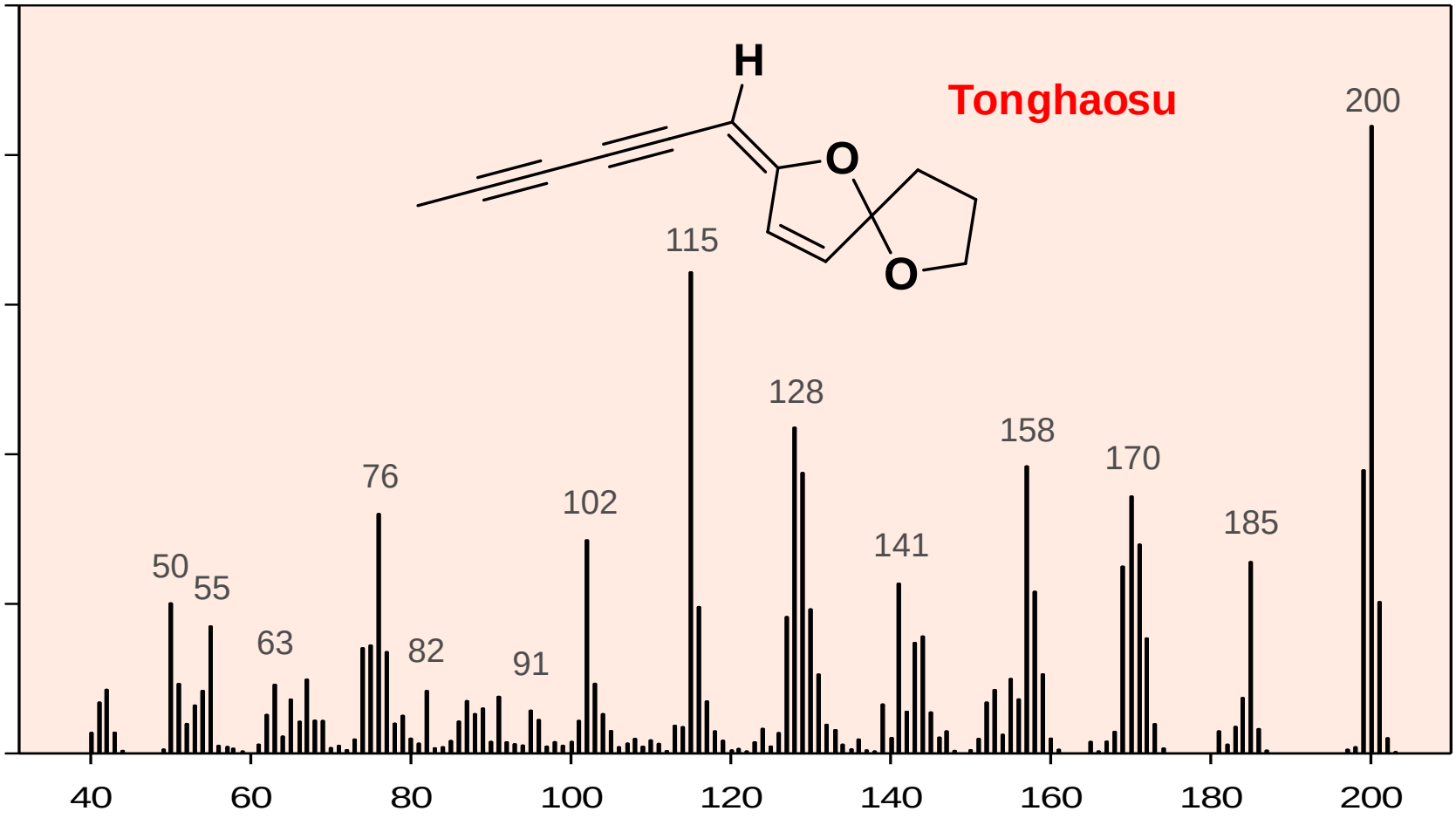
Disadvantages:

- Sample must be volatile (300 ºC)
- Complex and expensive instrumentation
- Fragmentation sometimes destroys molecular ion.

What Can We Get From GC/MS?

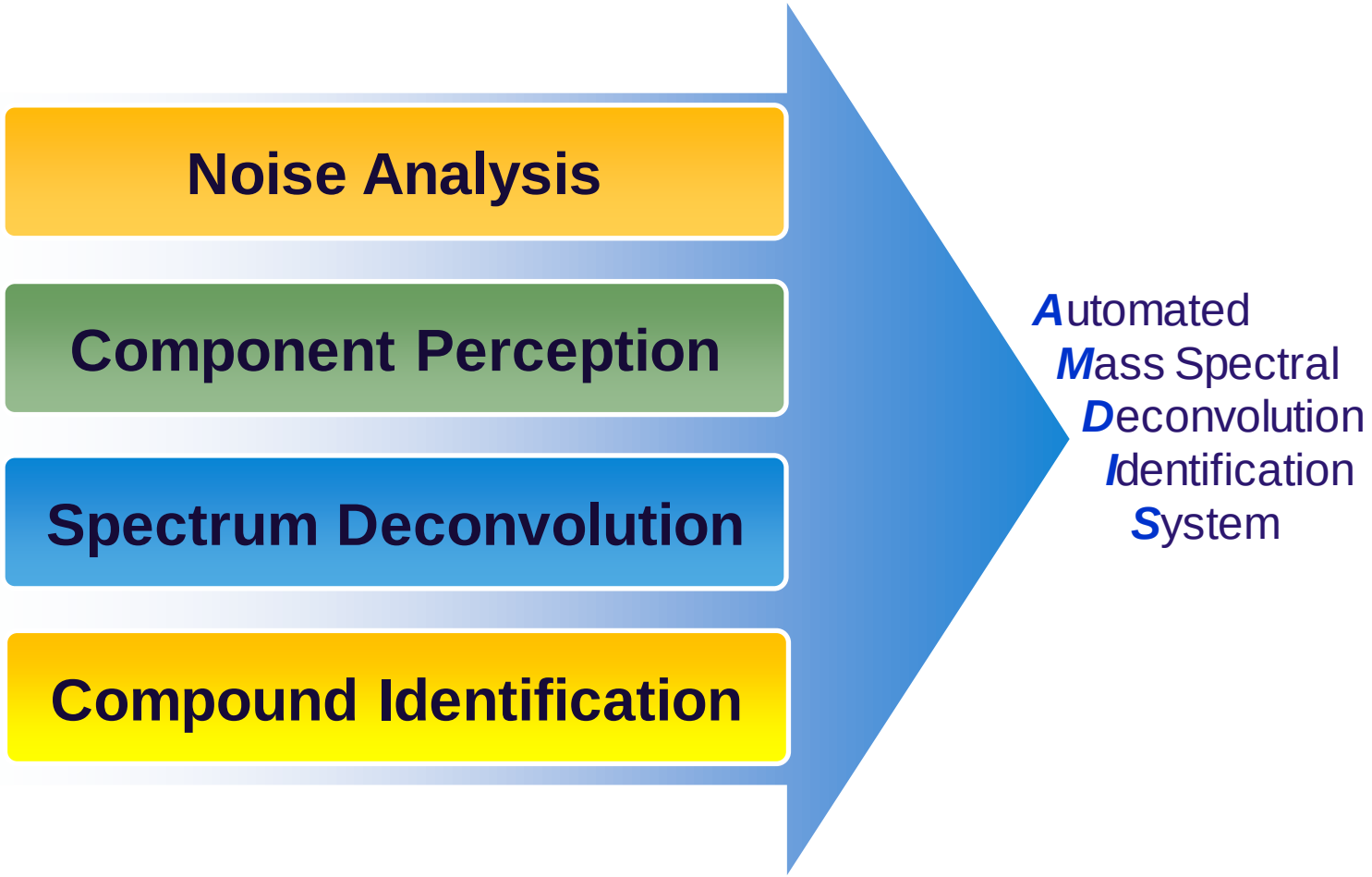


What Can We Get From GC/MS?

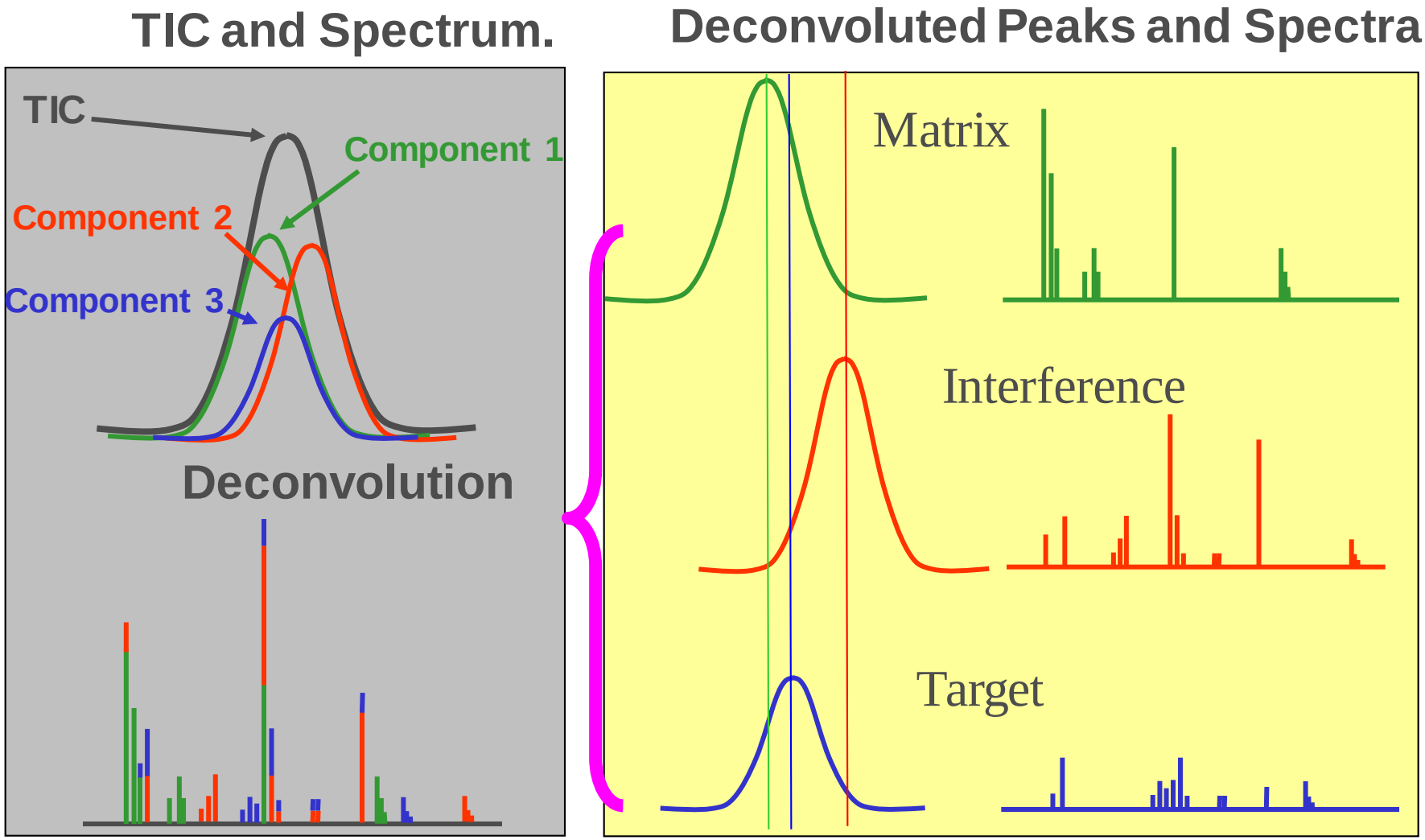


How Did We Study Chamomile?

Preparation Data for Statistical Analysis Using *AMDIS*

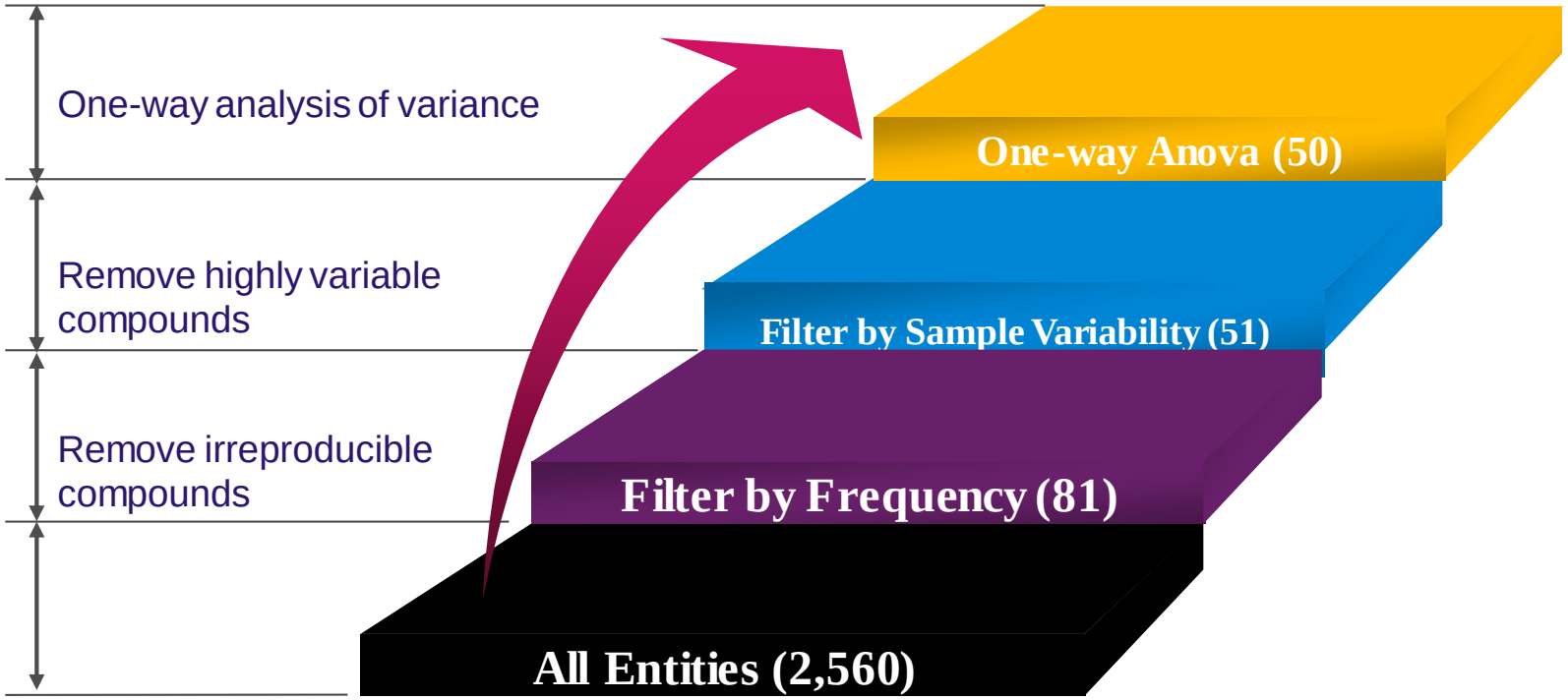


Pulls Out Individual Components and Their Spectra



How Did We Study Chamomile?

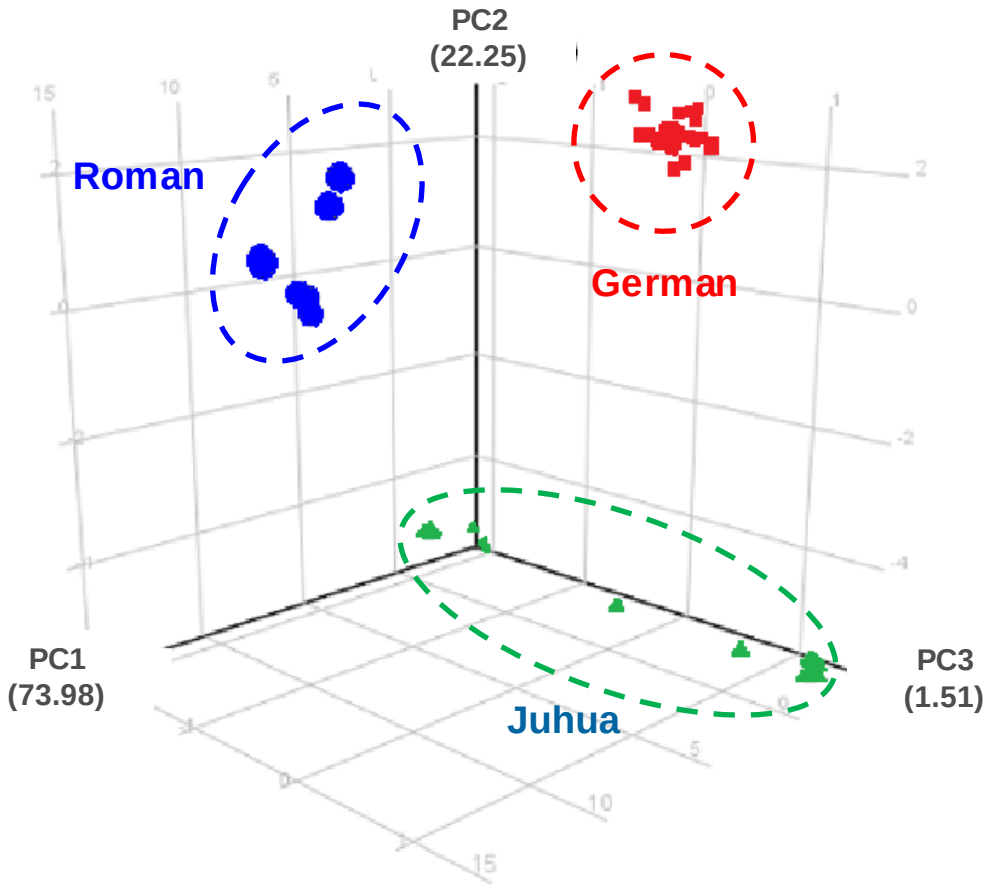
Find Meaningful Differences in Sample Sets Using Agilent *Mass Profiler Professional*



- To explore the most characteristic markers representing different chamomiles
- To reduce the dimensionality of the data

How Did We Study Chamomile?

Principal Component Analysis (PCA)



Uses for:

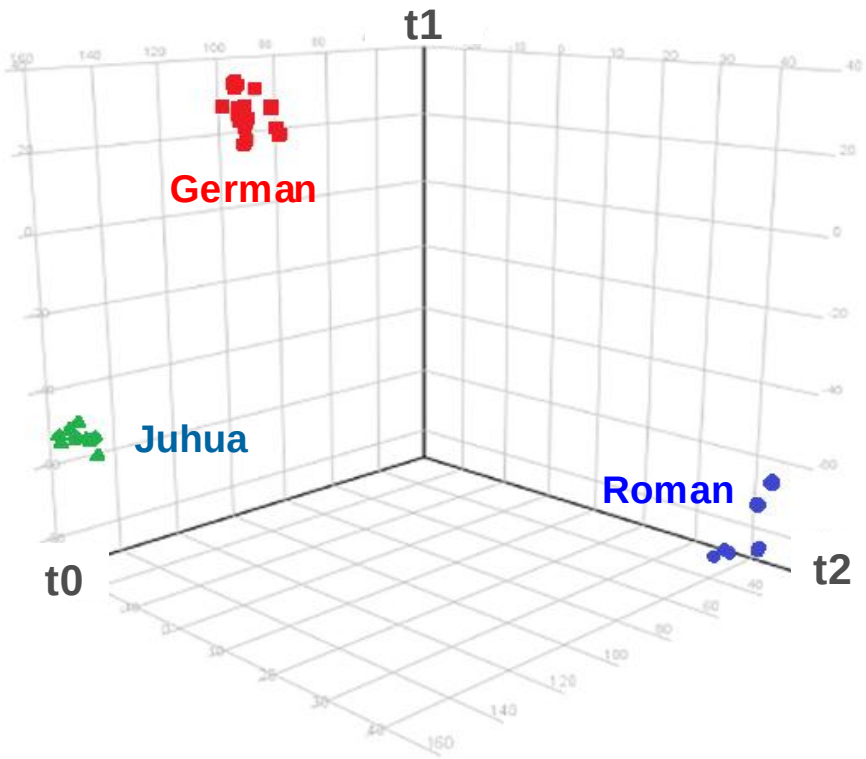
- Data Visualization
- Data Reduction
- Data Classification
- Trend Analysis

Solved Problems in the Study:

- How many unique "sub-sets" are in the samples?
- How are they similar / different?
- What are the underlying factors that influence the samples?
- Which measurements are needed to differentiate?

How Did We Study Chamomile?

Sample Class Prediction Model – Partial Least Squares Analysis (PLS-DA)



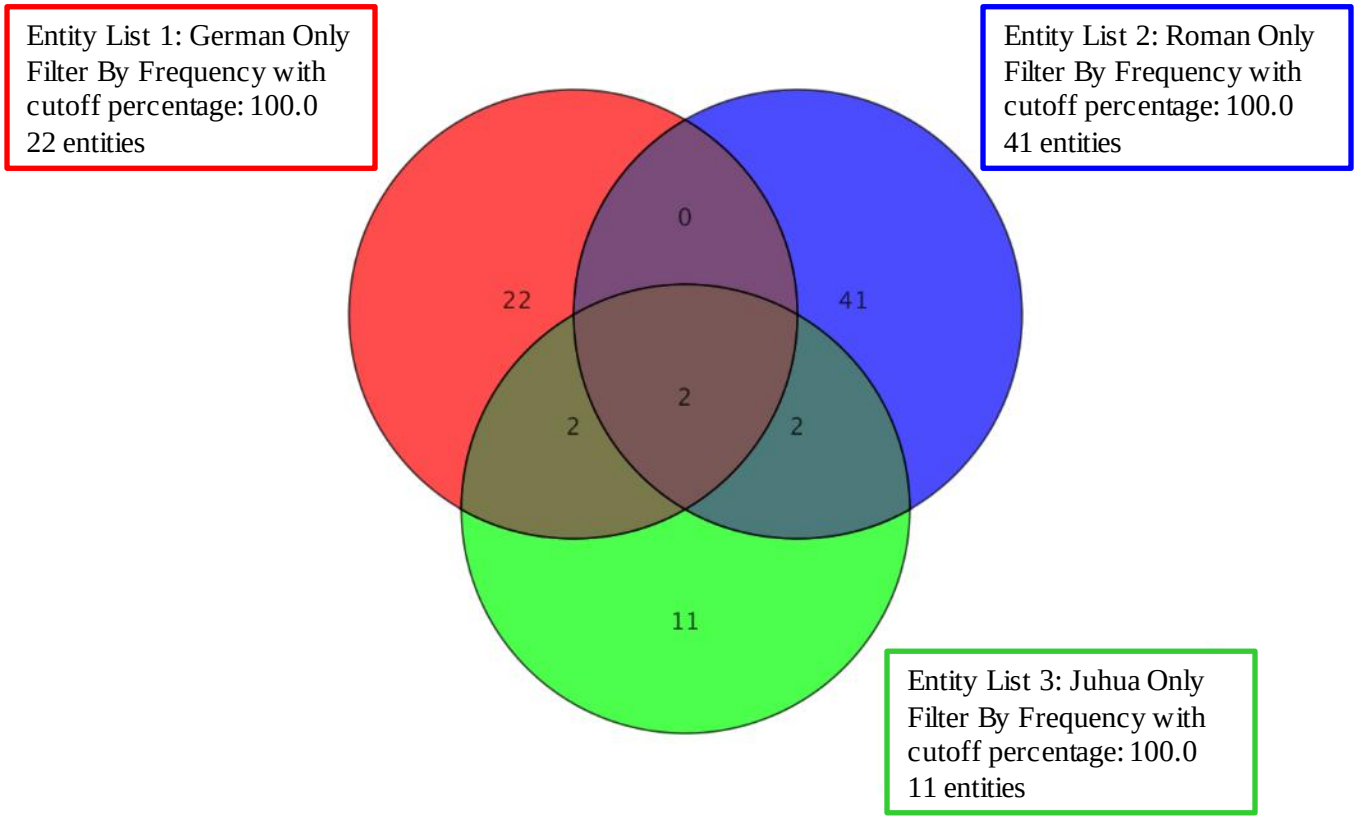
	German	Roman	Juhua	Accuracy (%)
Model Training				
German	15	0	0	100.0%
Roman	0	4	0	100.0%
Juhua	0	0	8	100.0%
Recognition Ability (%)				100.0%
Model Validation				
German	4	0	0	100.0%
Roman	0	4	0	100.0%
Juhua	0	0	4	100.0%
Prediction Ability (%)				100.0%

How Did We Study Chamomile?

NCNPR Accession Code	Product Information from the Label	Predicted	Confidence Measure
2061	Roman chamomile	German	0.47
3670	Chamomile flower	German	0.92
3998	Chamomile extracts	German	0.53
4903	Chamomile powder	German	0.90
5770	Chamomile powder	German	0.93
9357	Chamomile flowers	German	0.82
9359	Chamomile flowers	German	0.84
9362	Chamomile flowers	German	0.84
9364	Chamomile flowers	German	0.92
9382	Chamomile Organic Tea (Leaves and flowers)	German	0.94
9383	Herbal Chamomile & Fruit Tea (Rosehips, chamomile, orange peel, lemon peel & lemon myrtle)	German	0.72
9384	Chamomile Herb Tea	German	0.58
9385	Organic Tea	German	0.81
9386	Carrington Tea-Chamomile	German	0.75
9387	Chamomile Herbal Tea	German	0.91
9388	Chamomile Herb Dietary Supplement	German	0.89
9389	Chamomile Herbal Tea	German	0.61
9390	Chamomile Herbal Tea	German	0.92
9391	Chamomile Herbal Tea	German	0.77
9393	Whole German Chamomile Flowers	German	0.87
9423	Chamomile Herbal Dietary Supplement	Juhua	0.83
9424	Chamomile Herbal Dietary Supplement	Juhua	0.84
9425	Chamomile Herbal Dietary Supplement	Juhua	0.60
9426	Chamomile Herbal Dietary Supplement	Juhua	0.86
9428	Chamomile Herbal Dietary Supplement	Juhua	0.82
9429	Chamomile Herbal Dietary Supplement	Juhua	0.81
9432	Chamomile Herbal Dietary Supplement	Juhua	0.99

How Did We Study Chamomile?

Data Evaluation – Venn Diagram



How Did We Study Chamomile?

Markers Identified From Venn Diagram

Entities		Tentative NIST Identification	Molecular Weight	CAS Number
m/z	t_R (min)			
Roman Chamomile				
71.0	15.10	Isobutyric acid, isobutyl ester	144	97-85-8
71.0	23.42	Isobutyric acid, 2-methylbutyl ester	158	2445-69-4
55.0, 83.0	26.64	Butyl Butenoate[a,b]	156	54056-51-8
83.0	39.75	3-Methyl-2-butenoic acid, 3-methylbut-2-enyl ester	168	299309
100.0	44.75	Hexyl Butenoate	324	60129-26-2
German Chamomile				
205.0	66.94	Spathulenol	220	77171-55-2
143.0	71.43	α-Bisabolol oxide B[a,b,c]	238	26184-88-3
93.0, 141.0	73.04	α-Bisabolol[a,b,c]	222	515-69-5
143.0	76.07	Bisabolol oxide A[a,b,c]	238	22567-36-8
128.0	83.70	*E*-1,6-Dioxaspiro[4.4]non-3-ene, 2-(2,4-hexadiynylidene)-	200	50257-98-2
Juhua				
95.0	36.82	Borneol	154	10385-78-1
132.0	61.06	α-Curcumene	202	644-30-4
91.0	67.27	Caryophyllene oxide	220	1139-30-6
105.0, 121.0	69.75	Alloaromadendrene oxide	220	156128
204.0	71.69	Eudesm-7(11)-en-4-ol	222	473-04-1

A CASE STUDY ON CHARACTERIZATION OF OLIVE OIL

Authentication of “Extra Virgin” Purity

Goals for Olive Oil Characterization Study

Initial screening with high-res TOF

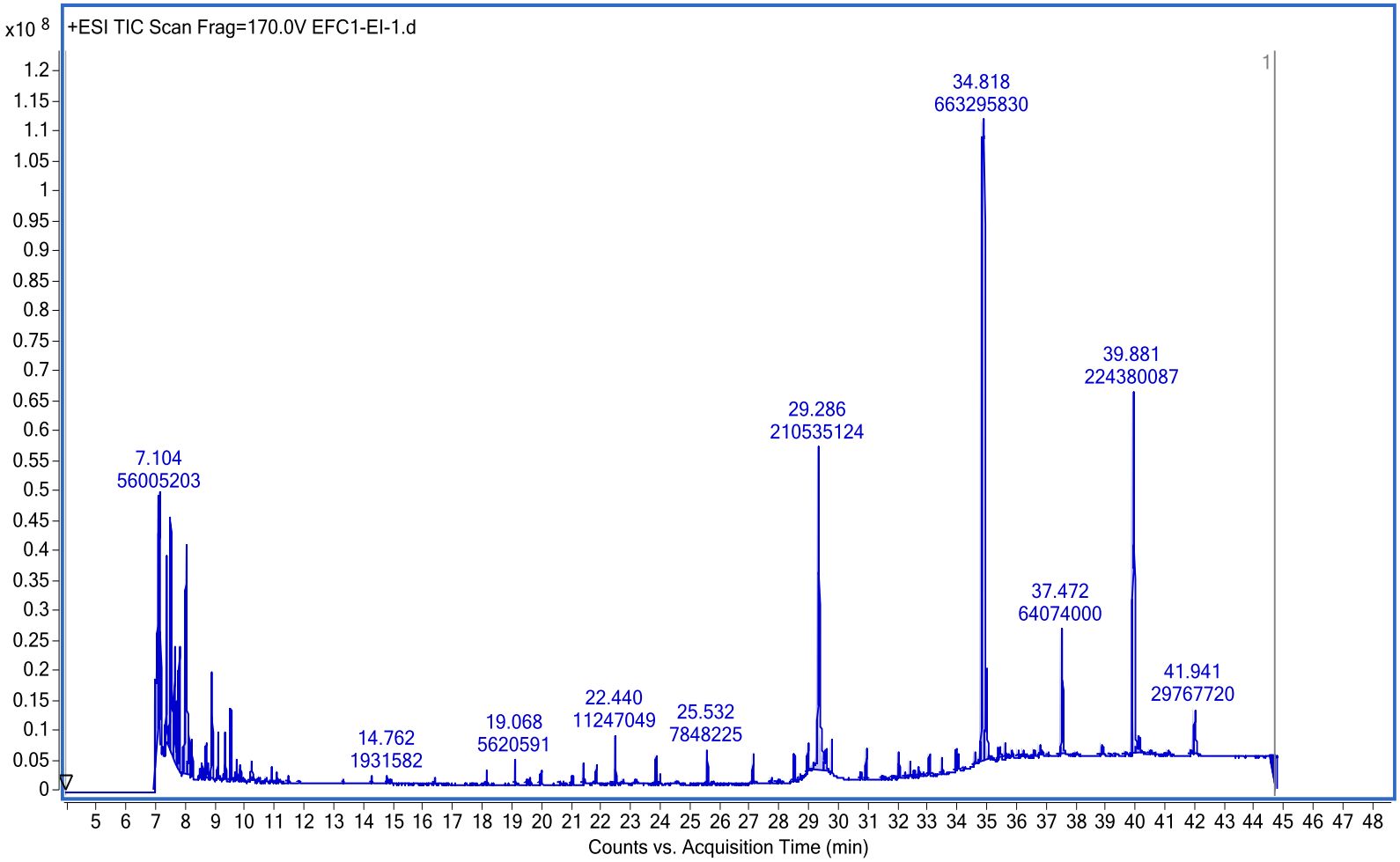
Confirm Molecular Ion with PCI

Mass Profiler Professional (MPP) analysis of data

- Generate statistical data on representative samples
- Build Class Prediction model based on sensory Pass/Fail information
- Test model on additional samples of known quality

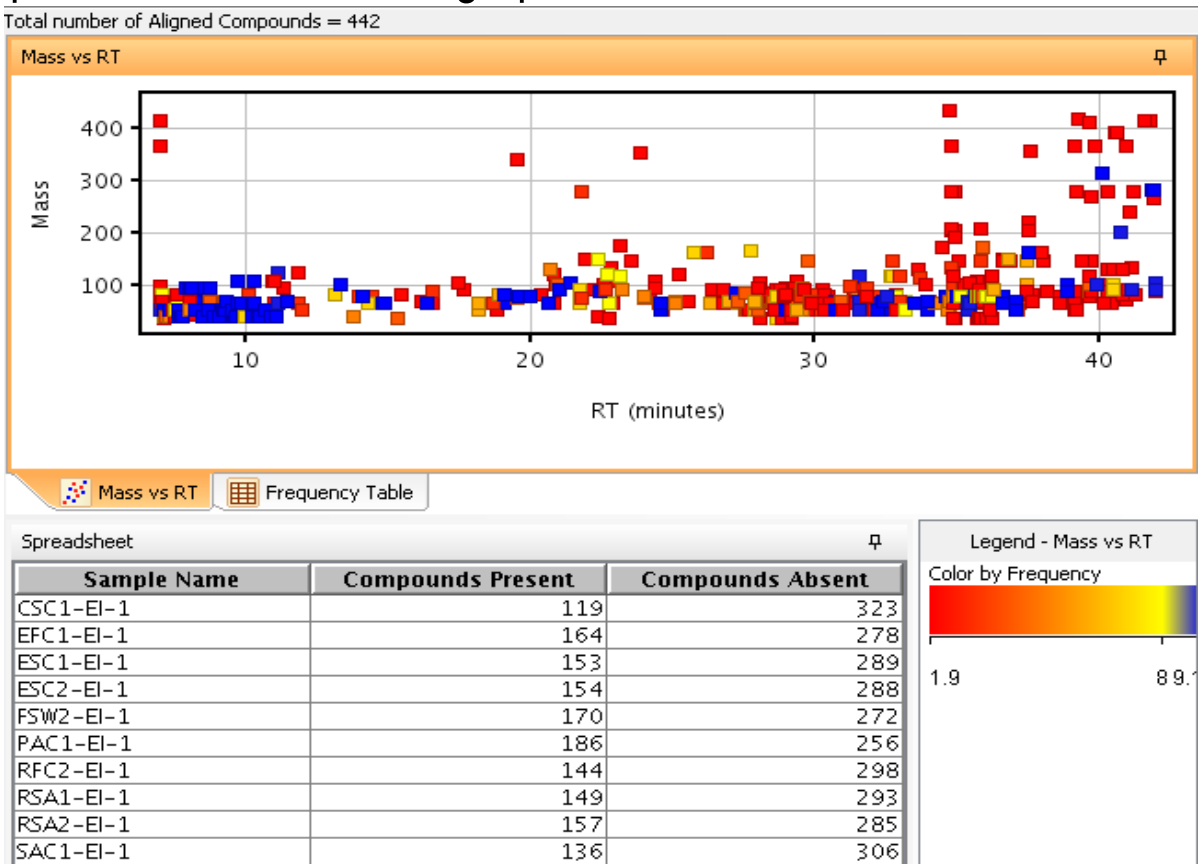
Total Ion Current of Extra Virgin Olive Oil

About 150 peaks with relative area filter of 0.1% of largest peak



Statistical Data on Olive Oil Samples

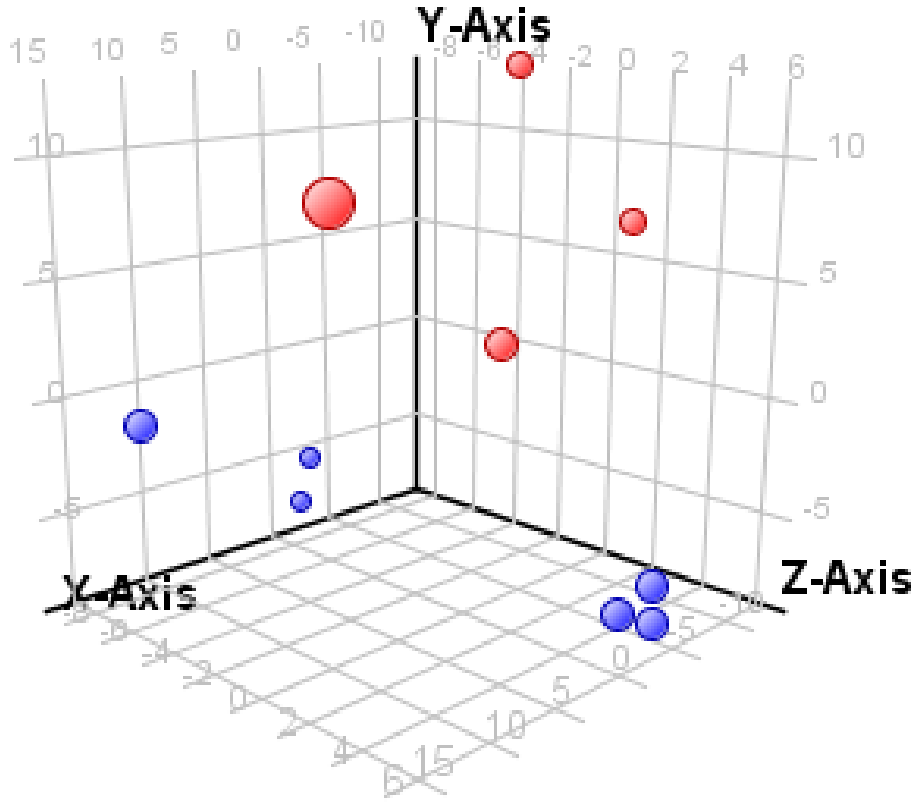
442 Unique Compounds from Chromatographic Deconvolution



Sample Name	Compounds Present	Compounds Absent
CSC1-EI-1	119	323
EFC1-EI-1	164	278
ESC1-EI-1	153	289
ESC2-EI-1	154	288
FSW2-EI-1	170	272
PAC1-EI-1	186	256
RFC2-EI-1	144	298
RSA1-EI-1	149	293
RSA2-EI-1	157	285
SAC1-EI-1	136	306

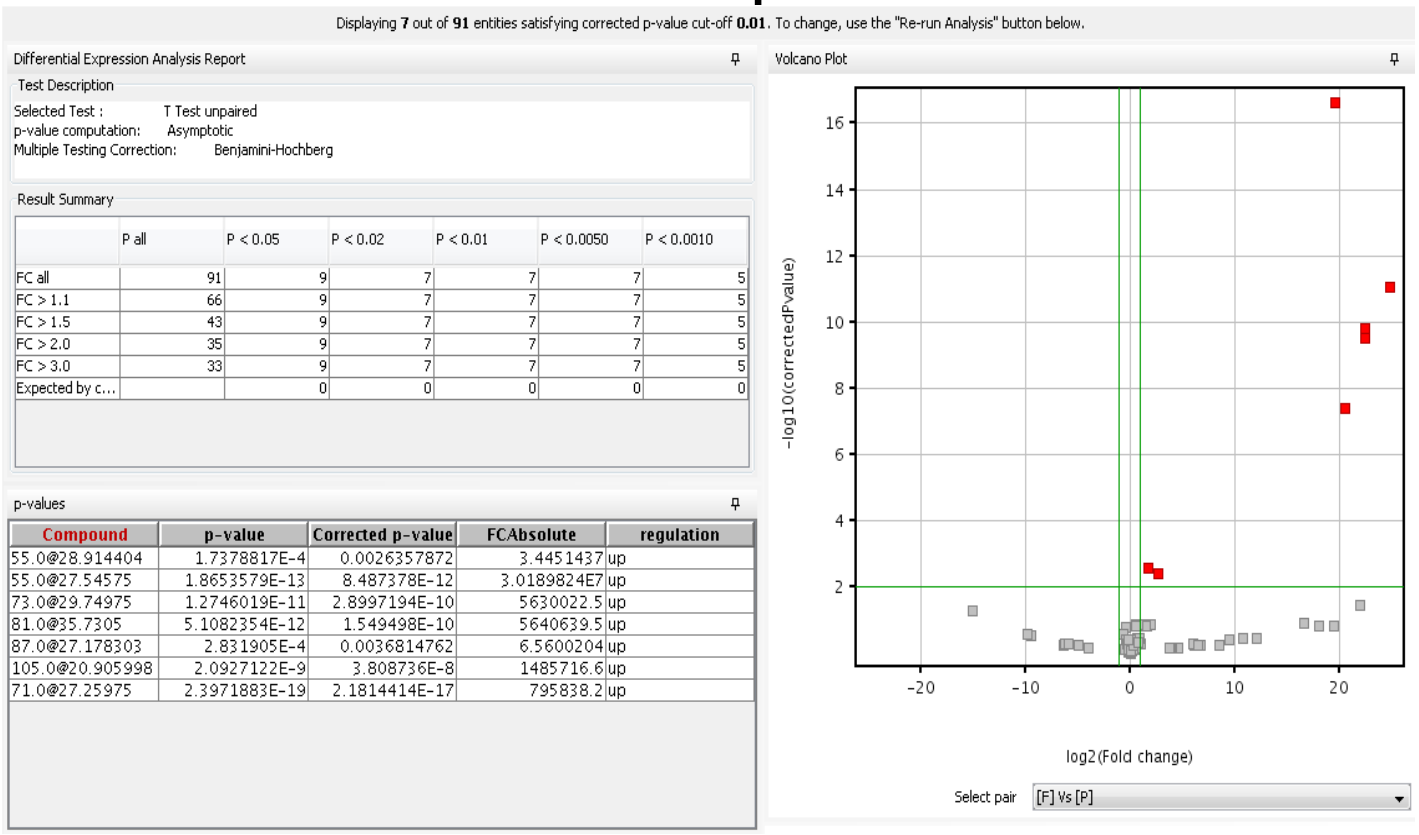
Most occurred only once or twice and were filtered out by MPP
Each sample is defined by passed/failed sensory evaluation

Principle Component Analysis (PCA) Data Clusters



Samples initially evaluated by sensory tests:
"Failed" marked red
"Passed" marked blue

Fold-Change Analysis Reveals Compounds Accumulated in “Failed” Samples



Displaying **7** out of **91** entities satisfying corrected p-value cut-off **0.01**. To change, use the "Re-run Analysis" button below.

Differential Expression Analysis Report

Test Description

Selected Test : T Test unpaired
p-value computation: Asymptotic
Multiple Testing Correction: Benjamini-Hochberg

Result Summary

	P all	P < 0.05	P < 0.02	P < 0.01	P < 0.0050	P < 0.0010
FC all	91	9	7	7	7	5
FC > 1.1	66	9	7	7	7	5
FC > 1.5	43	9	7	7	7	5
FC > 2.0	35	9	7	7	7	5
FC > 3.0	33	9	7	7	7	5
Expected by c...		0	0	0	0	0

p-values

Compound	p-value	Corrected p-value	FCAbsolute	regulation
55.0@28.914404	1.7378817E-4	0.0026357872	3.4451437	up
55.0@27.54575	1.8653579E-13	8.487378E-12	3.0189824E7	up
73.0@29.74975	1.2746019E-11	2.8997194E-10	5630022.5	up
81.0@35.7305	5.1082354E-12	1.549498E-10	5640639.5	up
87.0@27.178303	2.831905E-4	0.0036814762	6.5600204	up
105.0@20.905998	2.0927122E-9	3.808736E-8	1485716.6	up
71.0@27.25975	2.3971883E-19	2.1814414E-17	795838.2	up

Volcano Plot

The Volcano Plot on the right shows fold-change for each entity on the x-axis and significance on the y-axis.

Compounds Associated with "Failed" Sensory Test

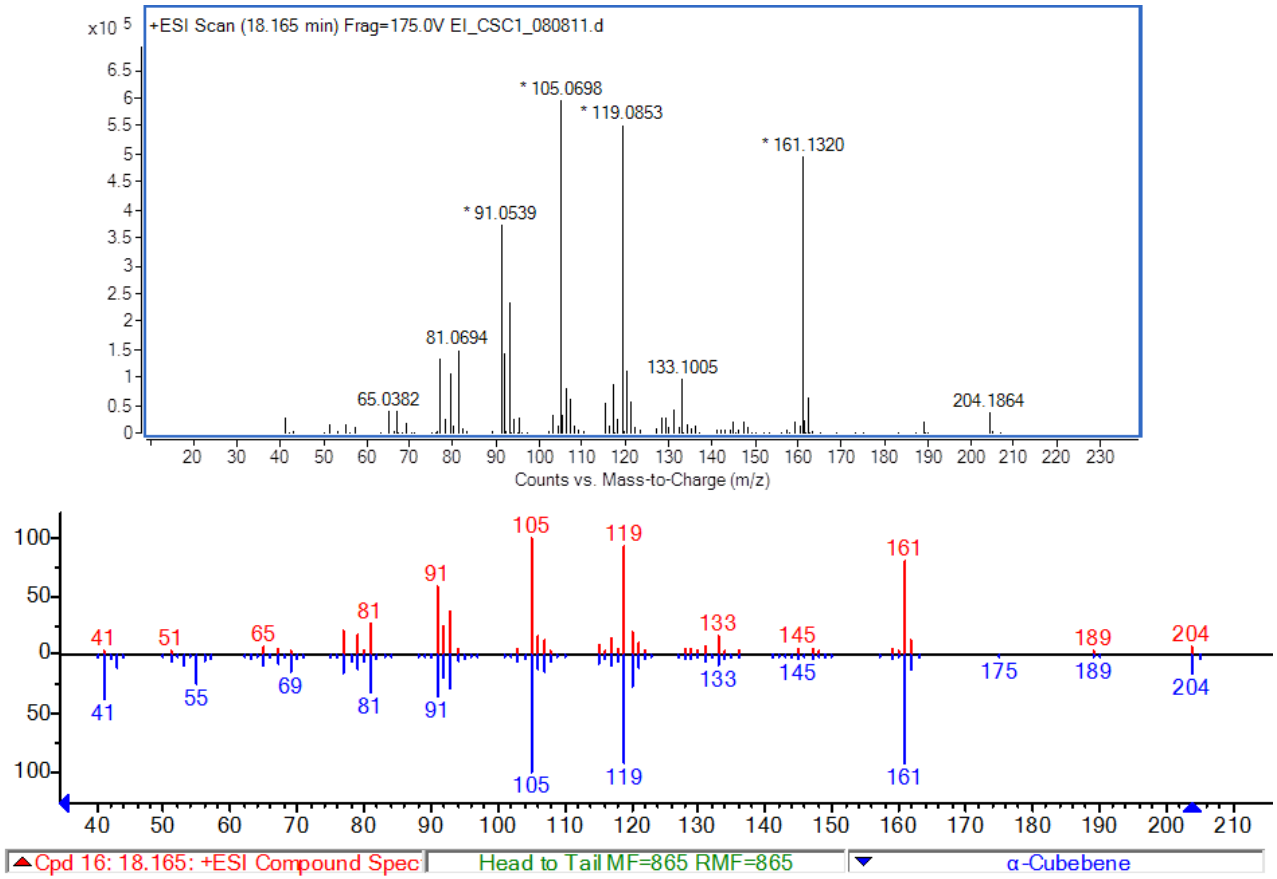
Tentative NIST ID	NIST Match	Formula	CAS	Odor	Source
n-Hexadecanoic acid	789	C16H32O2	57-10-3	Faint Oily	Bedoukian Research
Octadecanoic acid, ethyl ester	703	C20H40O2	111-61-5	Waxy	The Good Scents Company
2,6,10,14,18,22-Tetracosahexaene, 2,6,10,15,19,23-hexamethyl-	831	C30H50	111-02-4	Floral	The Good Scents Company
α-Cubene	880	C15H24	17699-14-8	Herbal	The Good Scents Company
None: confirmed by PCI data as 2H-Pyran-2-one, tetrahydro-6-nonyl-	574??	C14H26O2	2721-22-4	Waxy	The Good Scents Company

Lower NIST match factor indicates that tentative NIST ID may not be correct (actual compound not included in combined NIST Wiley database), but data gives a reasonable estimate of empirical formula

				EI [M*]+			PCI [M+H]+		
Tentative NIST ID	NIST Match	Formula	CAS	Calculated	Measured	Mass Error (PPM)	Calculated	Measured	Mass Error (PPM)
n-Hexadecanoic acid	789	C16H32O2	57-10-3	256.2397	256.2385	4.683115	257.2475	257.2470	1.9437
Octadecanoic acid, ethyl ester	703	C20H40O2	111-61-5	312.3023	312.3008	4.803039	313.3101	313.3091	3.1917
2,6,10,14,18,22-Tetracosahexaene, 2,6,10,15,19,23-hexamethyl-	831	C30H50	111-02-4	410.3907	410.39037	0.816295	411.3985	411.3987	0.4861
α-Cubene	880	C15H24	17699-14-8	204.1873	204.1883	4.897464	205.1951	205.1945	2.9240
None: confirmed by PCI data as 2H-Pyran-2-one, tetrahydro-6-nonyl-	574??	C14H26O2	2721-22-4	226.1927	ND	ND	227.2006	227.1987	8.3627

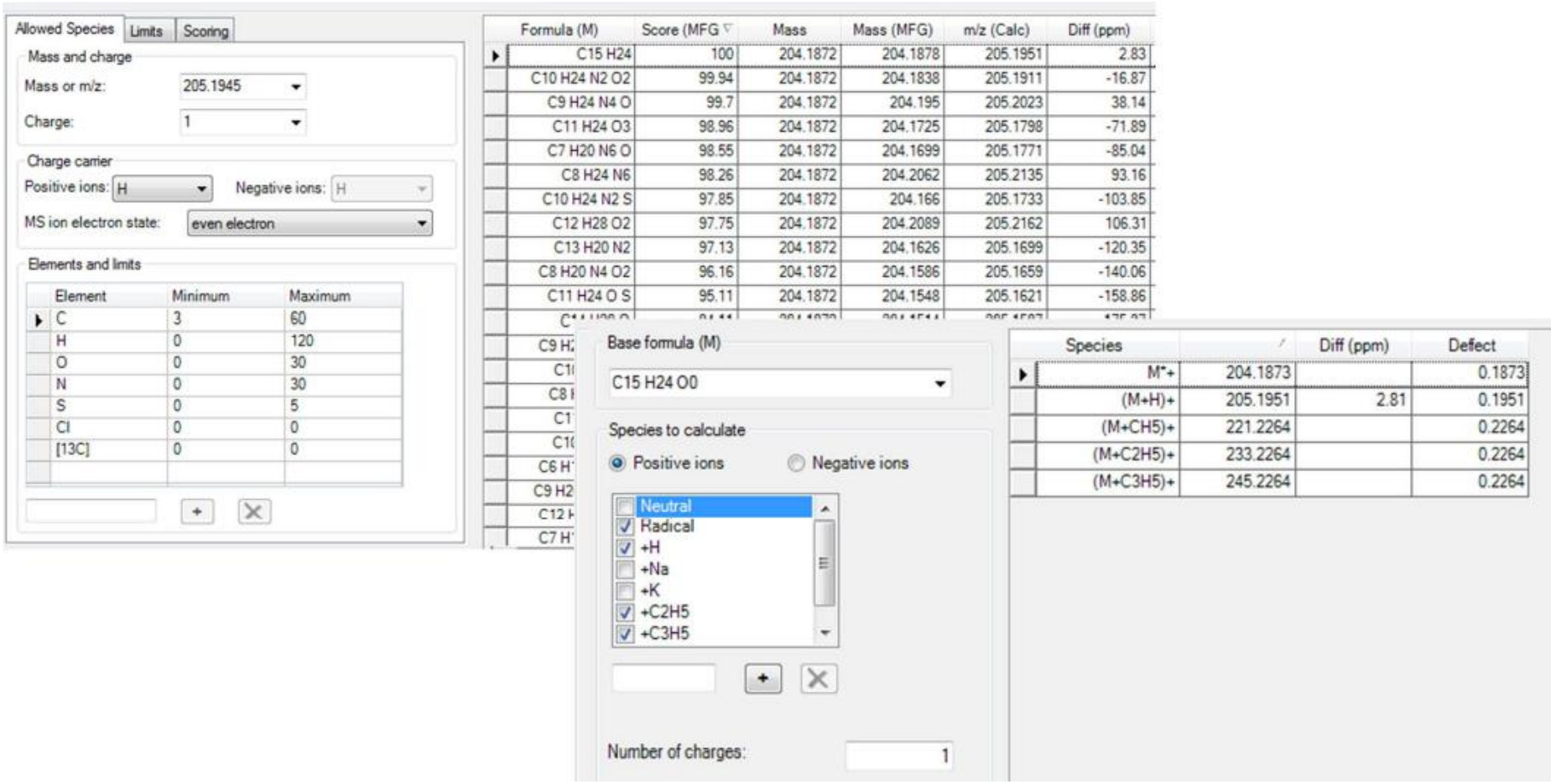
Compound with lowest NIST match also has highest deviation (ppm) for PCI

EI Library Search of Un-regulated Compounds



Commercial unit mass EI spectral libraries like Wiley and NIST can be searched using accurate mass EI TOF data to identify compounds.

PCI Confirmation of Molecular Ion



Mass and formula calculators used to determine empirical formula

Compounds Associated with "Failed" Sensory Test

Tentative NIST ID	NIST Match	Formula	CAS	Odor	Source
n-Hexadecanoic acid	789	C16H32O2	57-10-3	Faint Oily	Bedoukian Research
Octadecanoic acid, ethyl ester	703	C20H40O2	111-61-5	Waxy	The Good Scents Company
2,6,10,14,18,22-Tetracosahexaene, 2,6,10,15,19,23-hexamethyl-	831	C30H50	111-02-4	Floral	The Good Scents Company
α-Cubene	880	C15H24	17699-14-8	Herbal	The Good Scents Company
None: confirmed by PCI data as 2H-Pyran-2-one, tetrahydro-6-nonyl-	574??	C14H26O2	2721-22-4	Waxy	The Good Scents Company

Lower NIST match factor indicates that tentative NIST ID may not be correct (actual compound not included in combined NIST Wiley database), but data gives a reasonable estimate of empirical formula

				EI [M*]+			PCI [M+H]+		
Tentative NIST ID	NIST Match	Formula	CAS	Calculated	Measured	Mass Error (PPM)	Calculated	Measured	Mass Error (PPM)
n-Hexadecanoic acid	789	C16H32O2	57-10-3	256.2397	256.2385	4.683115	257.2475	257.2470	1.9437
Octadecanoic acid, ethyl ester	703	C20H40O2	111-61-5	312.3023	312.3008	4.803039	313.3101	313.3091	3.1917
2,6,10,14,18,22-Tetracosahexaene, 2,6,10,15,19,23-hexamethyl-	831	C30H50	111-02-4	410.3907	410.39037	0.816295	411.3985	411.3987	0.4861
α-Cubene	880	C15H24	17699-14-8	204.1873	204.1883	4.897464	205.1951	205.1945	2.9240
None: confirmed by PCI data as 2H-Pyran-2-one, tetrahydro-6-nonyl-	574??	C14H26O2	2721-22-4	226.1927	ND	ND	227.2006	227.1987	8.3627

Compound with lowest NIST match also has highest deviation (ppm) for PCI

What Did We Learn?

- Chemometrics can be used to analyze large, complex (3-D) data sets MUCH faster than manual analysis
- Mass Profiler Professional allows easy statiscally analysis of MS data.
- With AUTHENTICATED samples, an accurate sample class prediction model can be developed and verified
- The SCP model can subsequently be used to analyze samples in an automated manner w/o reanalysis of the authenticated samples
- Chemometric analysis can be used to identify potential markers for different type of samples

Thank you for your attention!

Any Questions?

joerg_riener@agilent.com