Mass Spectrometry Facility Advanced Analysis Centre



## **METABOLOMIC PROFILING**

• Identify and Quantify



## **METABOLOME COMPLEXITY**





## Mass Spectrometry Based Metabolite Profiling/Metabolomics Workflow

- Experiment design
- Method development
- Sample preparation
- Data acquisition
- Data processing & analysis





- Recovery rate
- Clean-up, SPE, filtering, centrifugation
- Ion suppression
- Sample handling
- Extraction solvents, vials, caps PTFE/silicon
- Eluents, chromatography column

## **Method development**

### **Sample preparation**



🖻 🞉 🚽 🖿 🎒 🦄 • 🗹 関 💙 🔊 - (°) × ∴ Chromatogram Results 🚰 Data Navigator Sort by Data File 2 Feb14-19-FP(1)(B)-negMS2.d x10<sup>5</sup> - V User Chromatograms Image: Image 3.9-User Spectra 3.8-Background Spectra 3.7-Compounds 3.6-Matched Sequences Feb14-19-FP(1)(E)-negMS2.d 3.5-User Chromatograms Comparing 3.4-VA - BPC Scan 3.3-User Spectra 3.2-Background Spectra 3.1-Compounds Different Matched Sequences 3-Feb14-19-FP(1)(H)-negMS2.d 2.9-User Chromatograms 2.8-User Spectra Extraction 2.7-Background Spectra Compounds 2.6-Matched Sequences 2.5-**Solvents for** Feb14-19-FP(1)(W)-negMS2.d 2.4-User Chromatograms 2.3-V AA - BPC Scar User Spectra 2.2-LC-MS Background Spectra 2.1-Compounds 2-Matched Sequences 1.9-1.8-1.7-



## **Method development**

### **Sample preparation**

GC/MS Chromatogram of silylated tobacco extract



Moldoveanu & David, "Derivatization Methods in GC and GC/MS" IntechOpen, 2018

**GC-MS** 

### **Data acquisition**

- Randomized samples
- Precolumn
- Blank run, "Junk" run, Pooled QC
- Internal standard(s)
- Mass calibration, LockMass, LockSpray, Rt reference standards
- Data-dependent and data-independent acquisition (DDA & DIA)

### **Data processing & analysis**

## PEAK-PICKING

- > DECONVOLUTION, ADDUCTS
- ISOTOPIC PATTERN
- > ALIGNMENT
- **GAP FILLING**
- > NORMALIZATION
- > MS2 ANALYSIS
- > ANNOTATION, ID
- > INTERPRETATION

and Name I	-													_
Auto Remote Serve	and the	NO BIO	A. Calor In Li	and warm	Name 1									
and the second s										Advantation and Antipageters				
over carrier E										A . I . THAN A REPORT	1	of 11 th 11 12	00 1 1	
re' y a masel y					146 Y # 170	CAR INNUE				with Sample Drive Cardinaport A	100 - 640	1-18P Space (330)	All and D. W. M.	24
		10	41.00	14	144		la.	1.92	4.74	45		widow.		
	1	- 974	101.007	< 0.	5.40	81 4	(la)	1.041	1104	10	1		(MAG)	10.0
34	£	Long of	18.048	16.42	0.94	84 2	15	6274	141	3	- 14	and the second	ser per	asta
	15	Physpires	1, 1,48	1.4	2.64	38 4	10	8.842	3.074	25	+121 (bel)	1-THE Spectrum (E. Sub-I	Lations CA, PP, reg.	24
											2	.05.		
_	_	_					_		-			-	1	-
A									-	10	_		324	1419
Bite Calument	Constant State										100	and the second states	and some the state of the	
¥* 5	10.00 78	1.00 2.0	1942. 214	Sevent Y*	740 24	DATA ANY VA	#* ¥	* 100 YA	109 54			H.C.		· · ·
A.89,002.14		10010	1014		16.978	-244	6.8	0.04		and several sea Calcul und	-	when		
A.M. (ma. 24		11112	108		107.325	43		N 012		10			101404	
100 million		100.00	-		2.00					0.0-		and the second second	and the second	
a BP age \$1	- 31	1000	7100		100		17				x104 Car	1 FM Spectrum () See (	NT MILES, My MIL	44
1. Mil. 1990. Ref.	1.1	Sec.14	U-m		16.30			1 84		18-	2	14.554		
A RP and Tot	10.00	10704	6763		W 808	24	4.9	1 0.04		14		1		
1,000,000,000	412	Numu .	14.402		N.68	-14	11	<ul> <li>6121</li> </ul>		22			254	1422
AP. 46.34	34	0.96	4071		10.00	-02	- 17	% 645a			100.00	- FRI Danne Provi	off south and the	24
100 mg 1 mg	- 615	71.08	11.00		10.000		- 13	1 0.007	_		-	31.4		-
A	- 54	100	100.00		× 80	-41		1 107		115		14.04		
W. market	0.0	11000	10.0		1 10	20				1/5* Servic Sour PCRI	1		10.40	
C 497 res 1.4	8.18	END.	8/21		10,000	12	11	0 0.04			- 4	in the second second	- 100	
1.17.14.24	33	387.3	1255		× 64	-1.8	6.7	5 115			100 100	1 HP Sector 5 284	(#0====================================	34
2,80,00,74	in T	2020	107.6		52 827	- 39	1.1	N 6101			-1	and the second		
1,17,16,314	6.5	192.8	0409.		10.04	-1.9	11	9 809		P1				-
5 M 140 314	21.55	Table	147407		14° 304	4.6	1.7	4 6.DH	_				10.000	- 40
A DE LAS AND		1000			10,000			000			100 000	1. The Summer Strates	Information West	24
a dia mandata		2004	100		1.1.1			110			14	MAC.		-
1.87 141.64	10.14	274%	760		4.82	-0.4	- 13	P 8424	_	1 Aller Aller	-	14048		
A. M. March	8.28	254.8	1104		10.004	1.0	11	0. 0100					.77.	.27
N.W./46.54	12.4	2010	7908		W 80	-18	6.7	5 6494		all to family firmer firmer	-		a. 901	
SP2421	11.8	2019	1041		167,0526		17	7, 6178		2	101.04	-19 Sector 0300	(400 mil) (A, 17, 16,	34
AUT, 46.51		1952	1968		w 80	-15	- 63	<ul> <li>6.64</li> </ul>			3-	100		
000		2000	- X.		10,000			0.04			1			
A	2.4	- 275	1000		100			101					100,003	
AP and by	7.8	200.00	10,1004		N Ad	4.5	- 10	1 100			120 100	1 - Fel Same (1961	Water Ca. W. on	24
AT 140.44	218	3004	hills.		147.052	41	13	0 0100			-11	84,05		-
10,0035	14	264	1844		10 Mail	4.0	1.3	<ul> <li>F124</li> </ul>				1404		
Pres La	31	2101	0.92		147392	33	- 17	B 0.016					10.00	36.0
17,000,14	4.14	2010	11,11,4		147.962	-74		a (104					- 100 ·	-

## **Data processing & analysis**

## Potentially significant molecular features discriminate between two groups



#### **MULTIPLE MODELS FOR CONSENSUS BUILDING**

#### Univariate analysis

- volcano plots
- fold change analysis
- t-test

#### **Multivariate analysis**

- principal component analysis (PCA),
- partial least squares-discriminant analysis (PLS-DA)
- hierarchical clustering (HC)
- support vector machines (SVM)
- random forests (RF)

## FROM MS DATA TO MOLECULAR INSIGHT

Metabolite identification remains a major challenge in untargeted MS-based metabolomics

#### **MS1 Level**

## MOLECULAR FORMULA DETERMINATION

Golden Rules (Kind & Fiehn, Bioinformatics 2007)

#### MS2 Level or El ionization at MS1

**STRUCTURE ELUCIDATION** Shattering Pattern & Fragment Mass Measurement

# FROM MS DATA TO MOLECULAR INSIGHT





Differential analysis of multiple sample sets & statistical tools ANOVA, PCA, volcano plots, hierarchical trees, SOMs. METLIN Database

# FROM MS DATA TO MOLECULAR INSIGHT



NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY U.S. DEPARTMENT OF COMMERCE



350,704 electron ionization (EI) spectra 306,643 compounds, 43,774 replicate spectra 447,289 retention index (RI) values

#### **SUGGESTED RESOURCES FOR LC/GC-MS DATA PROCESSING & ANALYSIS**









100 XCMS Rdisop Molecular formula LC/MS and GC/MS Data calculations Analysis 10 CAMERA MetShot Collection of annotation related High throughput prioritized acquisition and processing of methods for mass spectrometry data tandem mass spectra MassBank MassBank MetFrag MeiFrag Tandem MS Reference Database In-silico Metabolite identifikation AMDIS MetFamily Metermine MS Search NIST For the identification of regulated metabolite families.

## **RESOURCES FOR MASS SPECTROMETRY BASED MOLECULAR ID**



Picture credit: Jarmusch et al., Natural Product Reports, 2021



**Knowledgebase and Database ID** 

ç	<b>Confirmation of</b>	Matching
Nou	presence	obtained data
$\mathbf{\underline{\mathbf{Z}}}$		

Networking ID *de novo* 

Unknown

# What's in my sample?

"The hardest thing of all is to find a black cat in a dark room, especially if there is no cat"

Confucius



cat???