METABO-OPEN 2021

virtual training school in
metabolomics fair data management

Quality Control and Randomization in Metabolomics. Untargeted Metabolomic Workflow: From sample preparation to mass spectrometry analysis.

Dr Marynka Ulaszewska- Tarantino

Center for Omics Sciences,
Protemics and Metabolomics Facility
San Raffaele Scientific Institute, Milano, Italy

COST Action
CA17111
INTEGRAPE

The Untargeted Metabolomics Workflow

RESEARCH QUESTION & STUDY DESING

DATA GENERATION

LC-MS
GC-MS
NMR ...

FILE CONVERSION OPEN FORMAT

raw→mzXML,mzML
mz5, mgf, text, ms1,
cms1,ms2

PRE-PROCESSING:

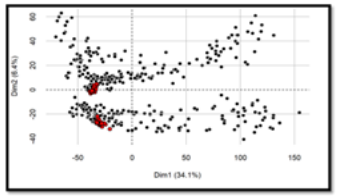
Peak picking
↓
Grouping and retention time correction
↓
Quality check
↓
Imputation of missing values

METADATA GENERATION

Soil, Climate
Genetics,
Age
Topography

EXPLORATIVE ANALYSIS

batch effect, intensity loss, ouliers

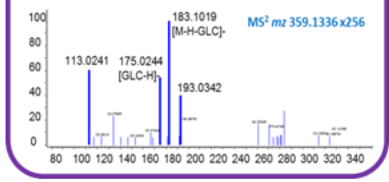


DATA TRANSFORMATION

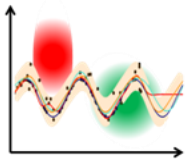
Intensity
Peak area

Log_{10}
scaling

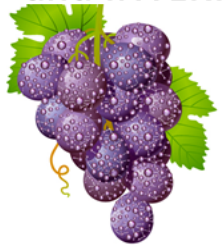
ANNOTATIONS

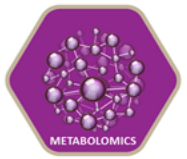


DATA ANALYSIS & STATISTICAL MODELLING



RESULTS VALIDATION and INTERPRETATION





METABO-OPEN 2021

METABO-OPEN 2021

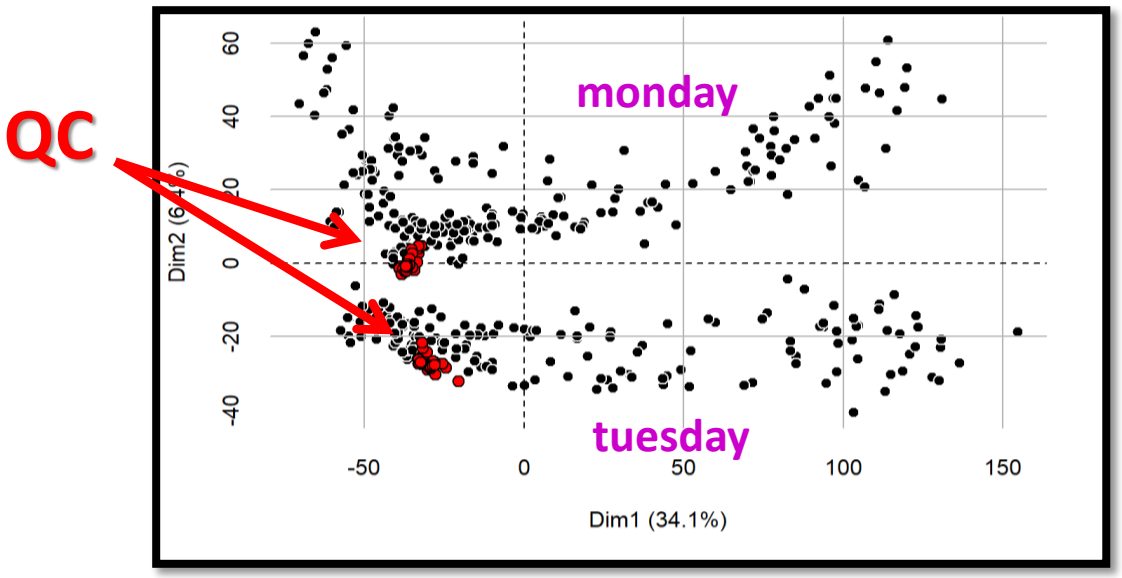
Quality Control and Randomization in Metabolomics

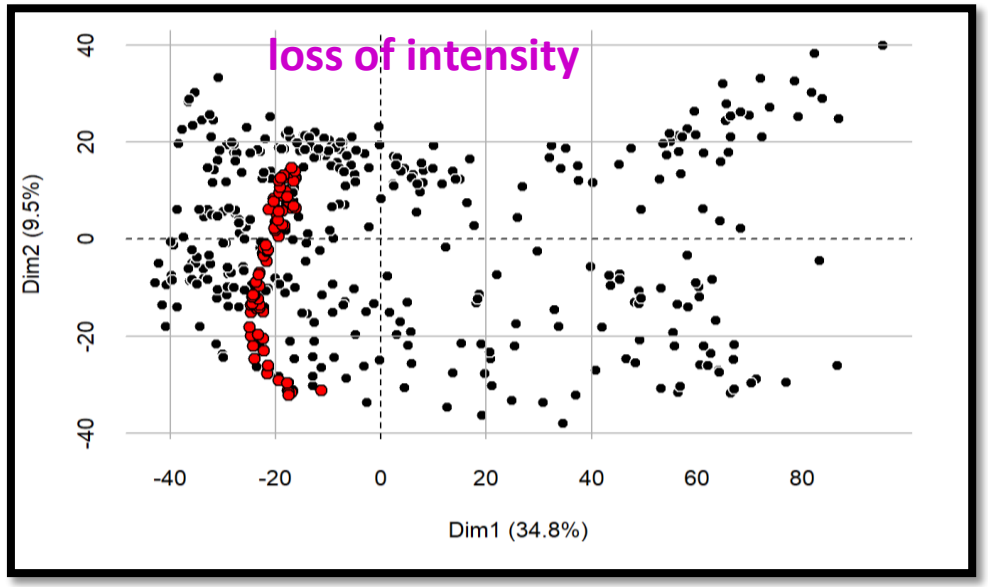
The repeated analysis of QC samples serves several purposes:

i) the general **monitoring** of the performance of the **analytical system**, for example, concerning retention time (*t*R) and signal intensity stability, mass calibration, etc.;

ii) the **determination** of a method's overall **precision** (including sample preparation, intraday and interday) if several QC sample aliquots are prepared independently per batch

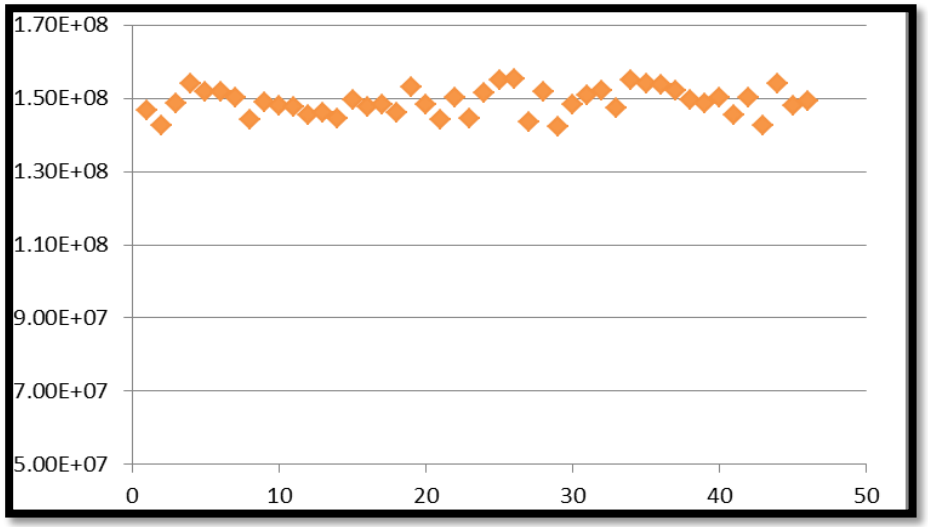
iii) the calculation of QC sample-based drift correction functions aiming to remove systematic trends and batch effects



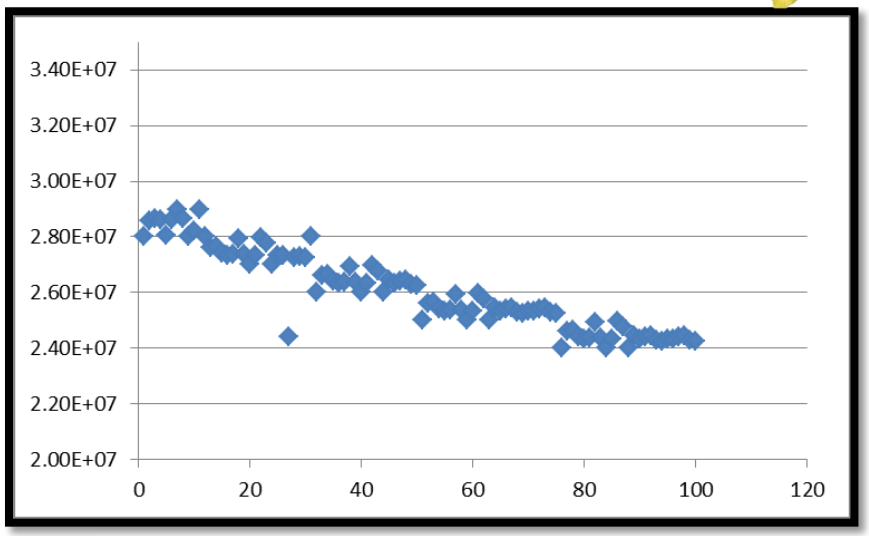


METABO-OPEN 2021

Intensity of the MS signal during injection queue for QC sample



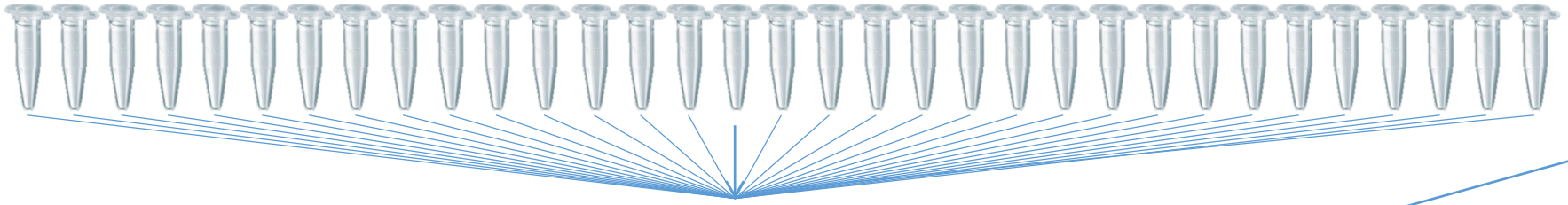
Expectation



Reality

METABO-OPEN 2021

Quality Control



QC pool sample

Consists of small aliquote (10 uL or 50 uL)
taken from each study samples
Pooled in one vial and injected along queue many times

Depending on retention time duration
it can be ever 5, 8 or10 samples

DAY1

x001_solvent
x002_solvent
x003_QC_equilibration_run
x004_QC_equilibration_run
x005_QC_equilibration_run
x006_QC_equilibration_run
x007_QC_equilibration_run
x008_Blank1
x009_Blank2
x010_Blank3
x011_solvent
x012_QC pooled
x013_QC pooled
x014_QC pooled
x015_QC pooled
x016_sample_GR
x017_sample_IT
x018_sample_IT
x019_sample_GR
x020_sample_ES
x021_QC pooled
x022_sample_GR
x023_sample_ES
x024_sample_GR
x025_sample_IT
x026_sample_IT
x027_QC pooled
x028_sample_ES
x029_sample_ES
x030_sample_IT
x031_sample_GR
x032_sample_IT
x033_QC pooled

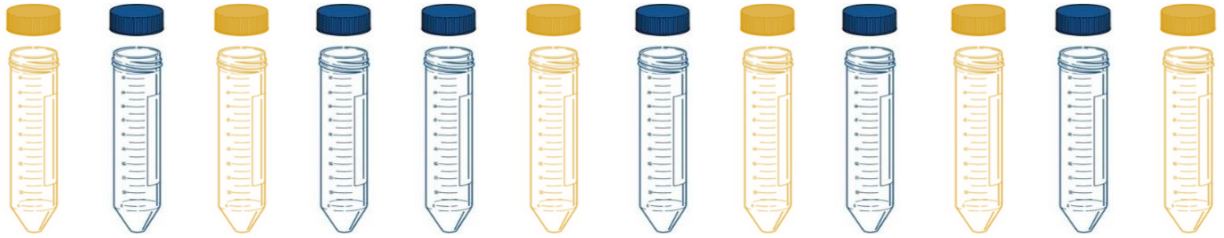
METABO-OPEN 2021

Quality Control and Randomization in Metabolomics

Monday – Extraction of control group

Tuesday – Extraction of treatment group

Randomize samples for extraction procedure



Link to randomize a sample list:
https://www.random.org/lists/

METABO-OPEN 2021

Quality Control and Randomization in Metabolomics

Link to randomize a sample list:
https://www.random.org/lists/

Randomize samples for injection queue

The queue of samples for injection can be very long and take **several days of injections**.

Randomize samples along queue so each day of analysis will consist of **balanced/similar number** of samples belonging to different **observation groups**

METABO-OPEN 2021

Quality Control and Randomization in Metabolomics

Randomization of the sample extraction and sample analysis sequence

- Complete randomization - samples are randomized over all batches or the entire measurement series. This means that the samples from all groups are exposed to all sources of analytical errors (random or systematic) to approximately the same extent

- Partial or groupwise randomization - samples belonging to particular subgroups are analyzed in a batchwise fashion (equal amount of samples in each batch). This approach minimizes analytical errors within the group but accepts potentially higher systematic offsets between the groups

Experiment of 1000 samples divided in batches usining partial randomization

Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
200 samples Balanced for: ~ Nº CTRL vs OBS ~ Timepoints in kinetics experiments of the same species/replicate should be kept together	200 samples Balanced for: ~ Nº CTRL vs OBS ~ Timepoints in kinetics experiments of the same species/replicate should be kept together	200 samples Balanced for: ~ Nº CTRL vs OBS ~ Timepoints in kinetics experiments of the same species/replicate should be kept together	200 samples Balanced for: ~ Nº CTRL vs OBS ~ Timepoints in kinetics experiments of the same species/replicate should be kept together	200 samples Balanced for: ~ Nº CTRL vs OBS ~ Timepoints in kinetics experiments of the same species/replicate should be kept together

METABO-OPEN 2021

Untargeted Metabolomic Workflow: From sample preparation to mass spectrometry analysis.

Entire berry

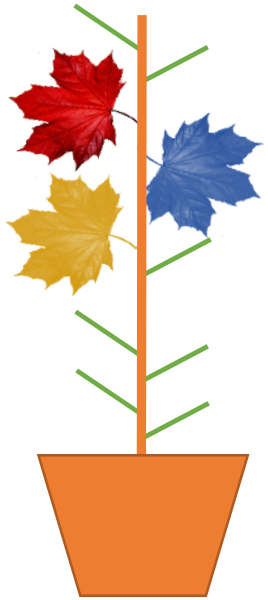
- Skin?
- Pulp?
- Seeds?

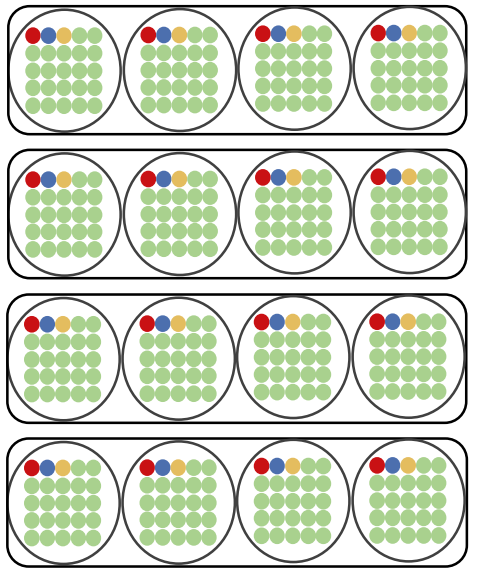
Plant Leaves

- Entire Leave?
- Discs

Wine

- Fermentation trials
- Final product





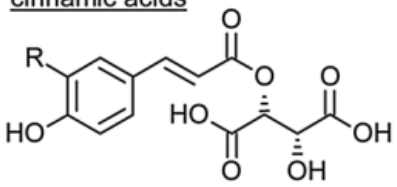
METABO-OPEN 2021

Untargeted Metabolomic Workflow: From sample preparation to mass spectrometry analysis.

Untargeted profiling of:

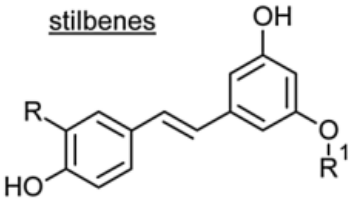
- Lipidome
- Carotenoids
- Phenolics Compounds
- Volatile Compounds
- Sugars

cinnamic acids



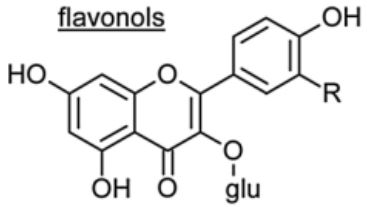
trans-caftaric acid	R = OH
trans-fertaric acid	R = OCH_3
trans-coutaric acid	R = H

stilbenes



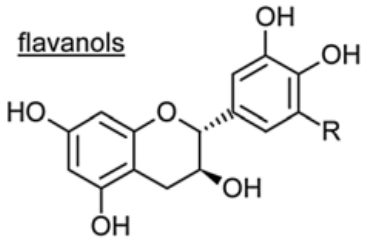
trans-resveratrol	R, R^1 = H
trans-piceid	R = H, R^1 = glucose
astringin	R = OH, R^1 = glucose
piceatannol	R = OH, R^1 = H

flavonols



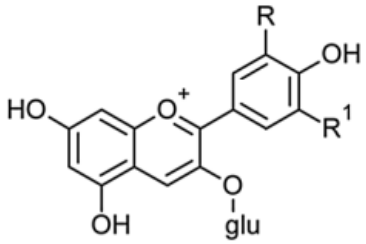
keampferol 3-*O*-glu	R = H
quercetin 3-*O*-glu	R = OH

flavanols



(+)-catechin	R = H
(+)-gallocatechin	R = OH

anthocyanins

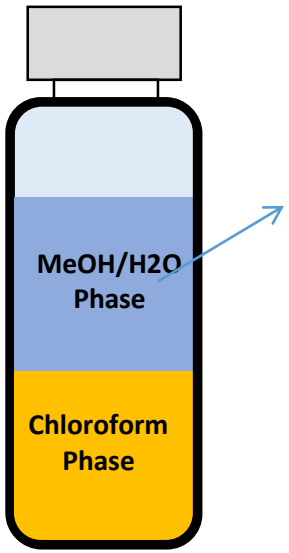


pelargonidin 3-*O*-glu	R,R^1 = H
delphinidin 3-*O*-glu	R,R^1 = OH
cyanidin 3-*O*-glu	R = H, R^1 = OH
peonidin 3-*O*-glu	R = OCH_3, R^1 = H
petunidin 3-*O*-glu	R = OCH_3, R^1 = OH
malvidin 3-*O*-glu	R,R^1 = OCH_3

METABO-OPEN 2021

Multiple phenolic classses extraction protocol

1) 1 g of frozen powder of ground berries should be weighed into 15-ml amber vials
2) Add Internal Standards IS: 50 µl o-coumaric acid (2 mg/ml MeOH)
3) Add 1.2 ml of H_2O/ CH_3OH (1:2)
4) Add 0.8 ml of $CHCl_3$ (chloroform)
5) Vortex for 1 min
6) Shake for 15 min at room temperature using an orbital shaker (Grant-Bio Rotator PTR-60)
7) Centrifuged at 4 °C and 1000 g for 10 min.



Transfer into a 5-ml flask. The upper aqueous methanolic phase

8) Add again 1.2 mL of of H_2O/ CH_3OH (1:2)
9) Vortex for 1 min
10) Shake for 15 min at room temperature using an orbital shaker
11) Centrifuged at 4 °C and 1000 g for 10 min.

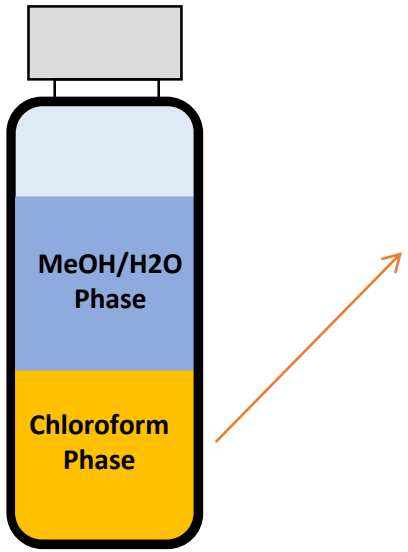
The aqueous-methanolic phase was collected and combined with the previous one, brought to a final volume of 5 mL with Milli-Q water,

12) filtered with a 0.2 µm PTFE filter (Millipore).

Vrhovsek U, Masuero D, Gasperotti M, Franceschi P, Caputi L, Viola R, et al. A versatile targeted metabolomics method for the rapid quantification of multiple classes of phenolics in fruits and beverages. J Agric Food Chem. 2012;60:8831–40.

METABO-OPEN 2021

Carotenoids extraction protocol - From the previous vial you recover the chloroform phase



The chloroform phase of the extraction solution is collected in separate vial.

1. Add 20 µL of trans-β-apo-8′-carotenal (25 µg/mL) was used as internal standard.
2. Add 10 µL of a 0.1 % triethylamine solution to prevent rearrangement of carotenoids.
3. Dry samples with N2, and resuspended in 50 µL of ethyl acetate.
4. Transfer to dark vials for analysis

Vrhovsek U, Masuero D, Gasperotti M, Franceschi P, Caputi L, Viola R, et al. A versatile targeted metabolomics method for the rapid quantification of multiple classes of phenolics in fruits and beverages. J Agric Food Chem. 2012;60:8831–40.

Wehrens R, Carvalho E, Masuero D, De Juan A, Martens S. High-throughput carotenoid profiling using multivariate curve resolution. Anal Bioanal Chem. 2013;405:5075–86.

METABO-OPEN 2021

Sugars – derivatization protocol

Take 0.1 g of fresh leaf or berry powder

1. Add 1 mL of cool (−20°C) extraction solvent: isopropanol/acetonitrile/water (3:3:2 v/v/v).
2. Add Internal Standards: 20 µL aliquot of palmitic-D3, nicotinic-D4, and glucose-D7 (1000 mg/L)
3. Vortex for 10 s and shake at 4°C for 5 min and Centrifuged at 12,000 g for 2 min at 5°C.

Take the supernatant to separate vial, and repeat extraction.

Merged two supernatants and re-suspend in a final volume of 5 mL using the extraction solvent.

4. Place a total of 250 µL of supernatant in a 1.5 mL Eppendorf tube
5. Evaporated to dryness under N2, and re-suspended in 500 µL of acetonitrile/water (50:50 v/v),
6. Vortex for 10 s, sonicate and centrifuge at 12,000 g for 2 min.
7. Transfer supernatant into a 1.5 mL Eppendorf tube and dried out under N2.
8. Derivatization: Add 20 µL of methoxamine hydrochloride in pyridine (20 mg/mL) to inhibit cyclization of reducing sugars and shaken at 30°C for 1 h;
9. Then add 80 µL of N-methyl-Ntrimethylsilyl-trifluoroacetamide with 1% trimethylchlorosilane for trimethylsilylation of acidic protons and shaken at 37°C for 30 min.
10. Add 5 µL of a solution containing decane and heptadecane (1000 mg/L) to monitor the chromatographic analysis and the instrumental conditions.
11. Transfer such extract into vials for analysis.

Chitarrini, Soini, Riccadonna et al Identification of Biomarkers for Defense Response to Plasmopara viticola in a Resistant Grape Variety Front. Plant Sci., 05 September 2017 | https://doi.org/10.3389/fpls.2017.01524

METABO-OPEN 2021

Volatile Compounds VOCs

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Quantitative metabolic profiling of grape, apple and raspberry volatile compounds (VOCs) using a GC/MS/MS method

Urska Vrhovsek*,[1], Cesare Lotti[1], Domenico Masuero, Silvia Carlin, Georg Weingart, Fulvio Mattivi

Research and Innovation Centre, Edmund Mach Foundation (FEM), Food Quality and Nutrition Department, Via E. Mach 1, 38010 San Michele all'Adige, Italy

1. Add 80 mL of water MilliQ, 0.5 g of gluconolacton and 50 L of internal standard n-heptanol (100 mg/L) into 30 g of deep frozen sample powder.
2. Homogenise with ultra-turrax at 21000 rpm for 3 min
3. centrifuge at 21,000 rpm at 5 °C for 5 min.
4. Recover supernatant and centrifuge again for 10 min and filtrated through rapid paper filter.
5. Bring the solution to an exact volume of 110 mL with milliQ water.

The SPE: ENV+ cartridges, 1 g (Biotage, Sweden).

1. Pre-condition the cartridge with 15 mL methanol followed by 20 mL of water.
2. Load the aqueous extract onto the cartridge, and wash with 15 mL of water.
3. Elute the free aromatic compounds with 30 mL dichloromethane. Add 60 mL of pentane
4. Elute the bound aromatic compounds (i.e. the glycosides) with 30 mL of methanol.

Vrhovsek, Lotti, Masuero, et al 2014. Quantitative metabolic profiling of grape, apple and raspberry volatile compounds (VOCs) using a GC/MS/MS method

METABO-OPEN 2021

Separation columns and instrumental analysis

RP C8 Column -> Lipidomics
RP C18 Column -> Various Phenolics Classess
RP C30 Column -> Carotenoids
HILIC Column -> Polar Metabolites (i.e amino acids)
VF-WAXms capillary column -> VOCs
Rxi-5Sil MS Columns fused silica -> Sugars

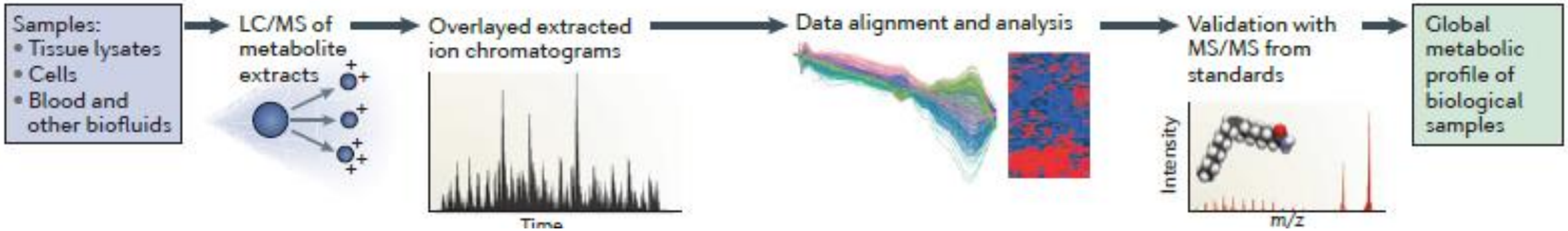
LC	GC
Mobile phase	
Liquid: Methanol, Acetonitrile, Water, 2-propanol, with buffers	Gases: Hellium, Nitrogen, Argon
Stationary phase	
Silica, C-8, C-18, modified silica, BEH Amide, Zirconia and titania stationary phases, Alumina, Porous graphitic carbon	Dimethylpolysiloxane, 5% Phenyl 95% dimethyl arylene siloxane, 14% Cyanopropyl-phenyl 86% dimethyl polysiloxane
Column lenghts, temperature and flows, retention time duration	
5cm; 10 cm, 15 cm Usually constant 30-60 C Shot gun, 5 minutes up to 90 minuts	10m, 100m Gradient cycle 50 C -250 C 60 minutes, 90 minutes and longer

Basing on your needs a different mobile phases can be used and different types of column

Data acquisition modes

b Untargeted metabolomics

Question:
What is the global metabolic profile of a sample?



NATURE REVIEWS | **MOLECULAR CELL BIOLOGY; April 2012**

Gary J. Patti, Oscar Yanes and Gary Siuzdak Metabolomics: the apogee of the omics trilogy

High Resolution Mass Spectrometer TOF and Orbitrap technologies

GC×GC-HR-TOFMS

METABO-OPEN 2021

Data acquisition modes

General Strategies:

Full Scan injections → stat.analysis → re-injection for MS^n

(Manual setting for selected m/z features)

Full Scan injections + DDA injections→ stat.analysis

The DDA mode allows for automatic acquisition of MS^n;

Thermo Scientific Orbitrap Fusion Tribrid LC-MS (I) with Thermo Scientific Dionex UltiMate 3000 Series UHPLC.

METABO-OPEN 2021

MS/MS data acquisition modes

DDA: DATA DEPENDENT ACQUISITION

dd-MS² (Data Dependent MS²) by Thermo, **IDA** (Information Dependent Acquisition) by ABSciex, **AutoMS** by Bruker, **Auto MS/MS** by Agilent

TOF - IDA

TOF – 8 ions: Total Cycle Time 1,25 sec

Full Scan HR 250ms	8 MS^2 120 ms each

TOF – 20 ions: Total Cycle Time 1,30 sec

Full Scan HR 250ms	20 MS^2 50 ms each

TOF – 50 ions: Total Cycle Time 1,50 sec

Full Scan HR 250ms	50 MS^2 20 ms each

ORBITRAP - DDA

DDA TOP4: Total Cycle Time 2,8 sec

Full Scan HR, Res. 60,000; ca 1sec	4 MS^2 Res.7500 each; 0,45s each

DDA TOP7: Total Cycle Time 4,15 sec

Full Scan HR, Res. 60,000; ca 1sec	7 MS^2 Res.7500; 0,45s each

DDA TOP12: Total Cycle Time 5,6 sec

HR 7500 ca 0,25 ms	12 MS^2 Res.7500 each; 0,45s each

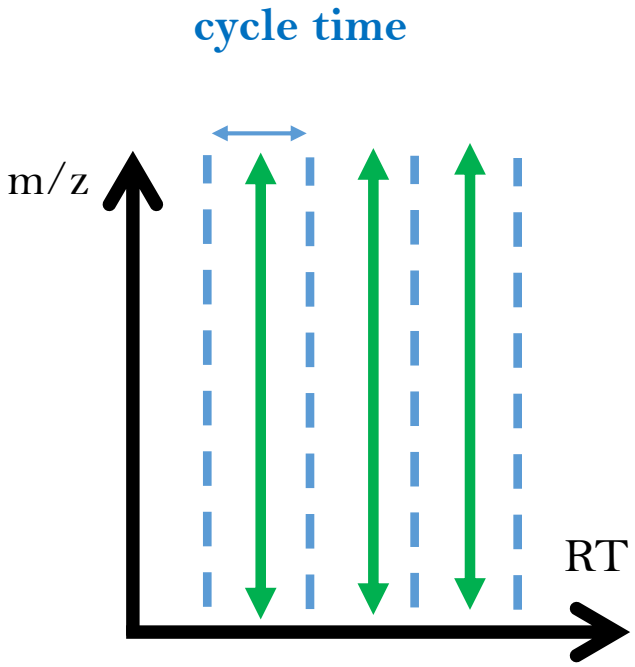
METABO-OPEN 2021

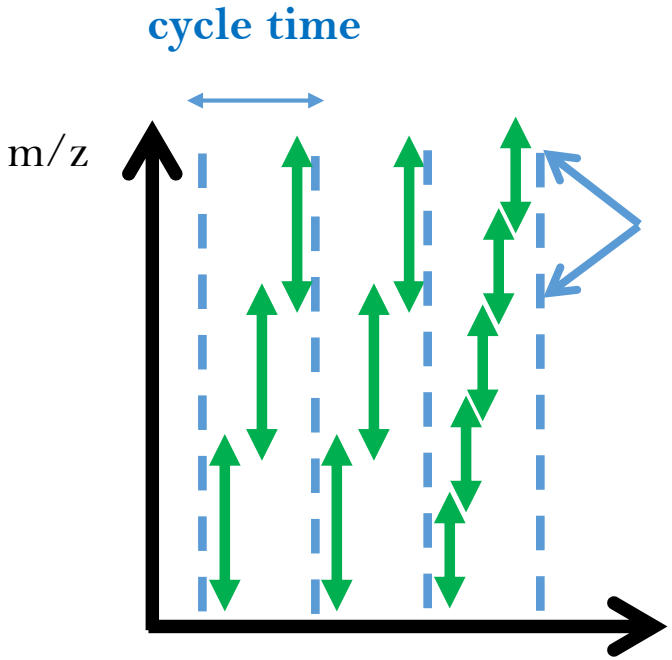
MS/MS data acquisition modes

DIA: DATA INDEPENDENT ACQUISITION

MS^E by Waters; **AI** (All Ions) by Agilent; **MSc^2** (Leco);
AIF-MS^2 (All-Ion fragmentation) by Thermo Fisher Scientific;
vDIA (Variable Data Independent) by Thermo Fisher Scientific;
bbCID (Borad Band Collision Induced Dissiociation) by Bruker

SWATH MS^{ALL} (Sequential Window Acquisition of All Theoretical Mass Spectra) by ABSciex



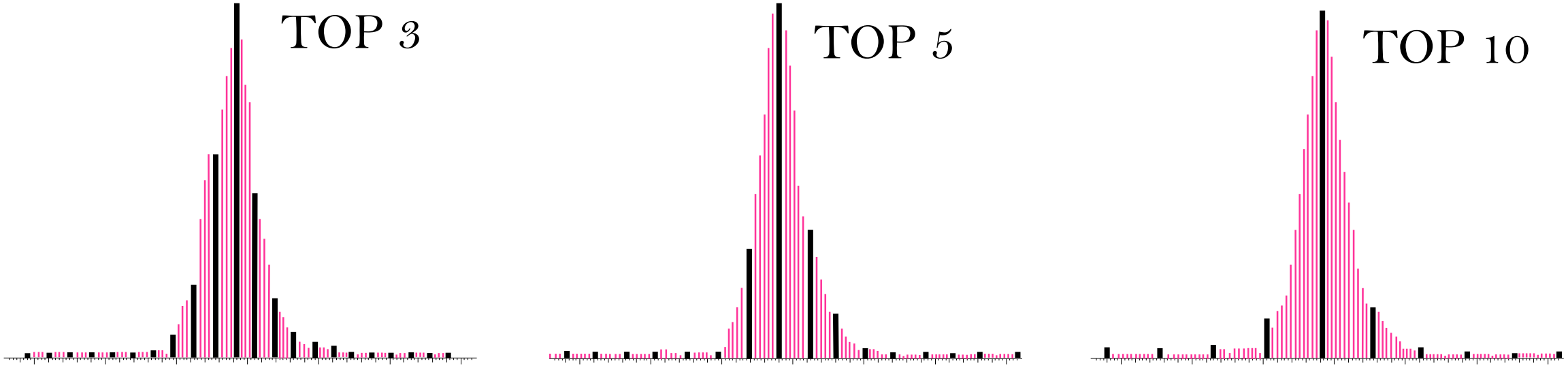


SWATH width window
i.e 25 Da, 50 Da or others

METABO-OPEN 2021

MS/MS data acquisition modes

Full Scan and DDA scans



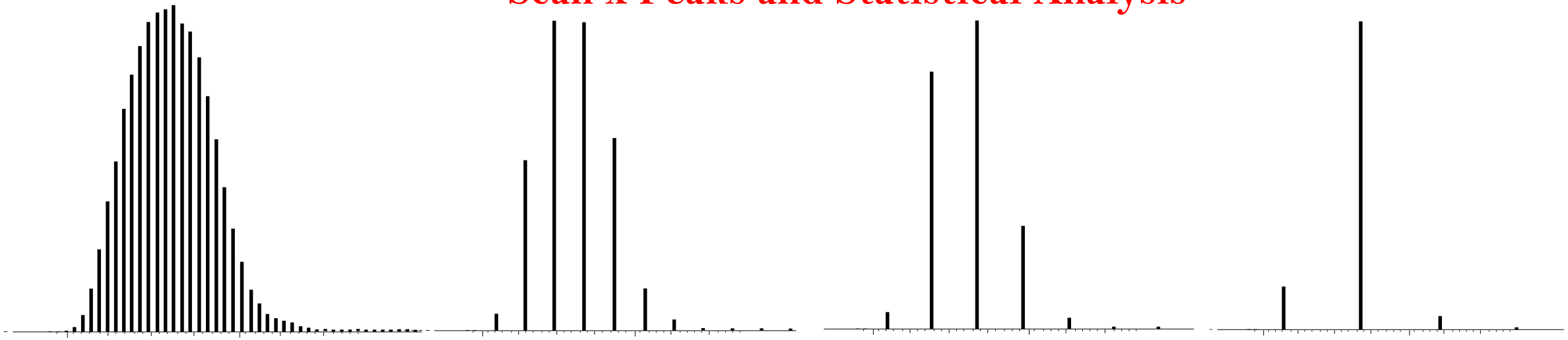
Strategy depends on you: **TOP 3, TOP 5, TOP 10**

Dear Orbitrap:
*Please do **1 Full Scan**, kindly check m/z features, and take 3/5 or 10 the most intensive ions and do fragmentation scan.*

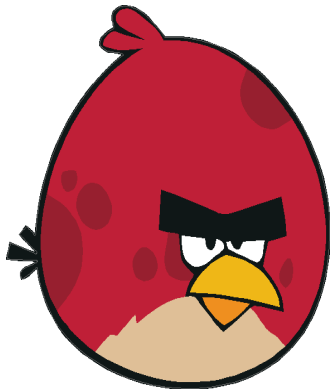
METABO-OPEN 2021

MS/MS data acquisition modes

Scan x Peaks and Statistical Analysis



Information from Full Scan spectra is introduced to statistical analysis –
it is clear why DDA is hard to be used for this issue



Biostatistian
won't be happy

TIP:

DDA is good strategy for **automatic MSn spectra acquisition.**

However**, statistic analysis** must be done on **Full Scan data**, thus, check which TOP2, or TOPx would be the best for your chromatography!

The Workflow

RESEARCH QUESTION & STUDY DESING

DATA GENERATION

LC-MS
GC-MS
NMR ...

METADATA GENERATION

Soil, Climate
Genetics,
Age
Topography

FILE CONVERSION OPEN FORMAT

raw→mzXML,mzML
mz5, mgf, text, ms1,
cms1,ms2

PRE-PROCESSING:

Peak picking
↓
Grouping and retention time correction
↓
Quality check
↓
Imputation of missing values

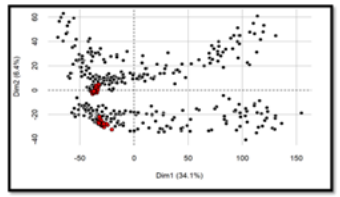
DATA TRANSFORMATION

Intensity
Peak area

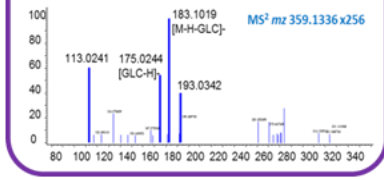
Log_{10}
scaling

EXPLORATIVE ANALYSIS

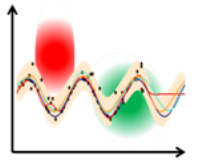
batch effect, intensity loss, ouliers



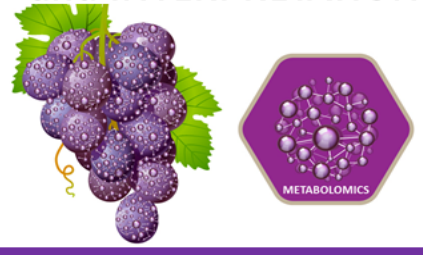
ANNOTATIONS



DATA ANALYSIS & STATISTICAL MODELLING



RESULTS VALIDATION and INTERPRETATION



METABO-OPEN 2021

METABO-OPEN 2021

virtual training school in
metabolomics fair data management

Thank you for your kind attention

COST Action
CA17111
INTEGRAPE