

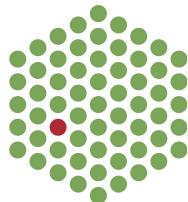
Untargeted Metabolomics workflow and spectral libraries

Prasad Phapale

EMBL Heidelberg, Germany

IITB live webcast

EMBL



Metabolomics

Relevance to biological response/ Phenotype

WHAT RESPONSES ARE POSSIBLE

GENOMICS



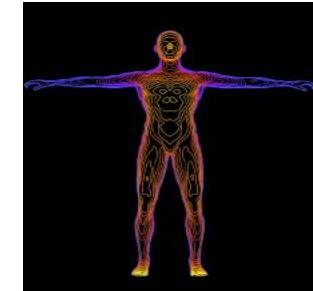
INITIATION OF RESPONSE

TRANSCRIPTOMICS



Environment

Functional Metabolomics



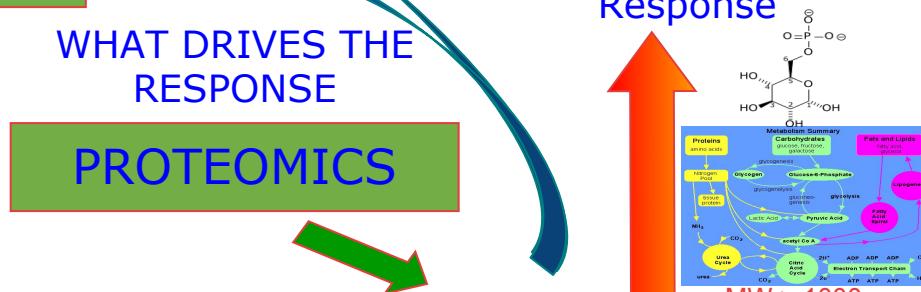
PHENOTYPE/
Drug or disease
Response

WHAT DRIVES THE RESPONSE

PROTEOMICS

THE ACTUAL RESPONSE

METABOLOMICS



EMBL Metabolomics Core Facility



Research areas

Developmental Biology
Metabolism and Signaling dynamics

Cell biology
Cancer metabolism

EMBL Metabolomics Core Facility

Structural and Computational
Spatial metabolomics
Metabolic network

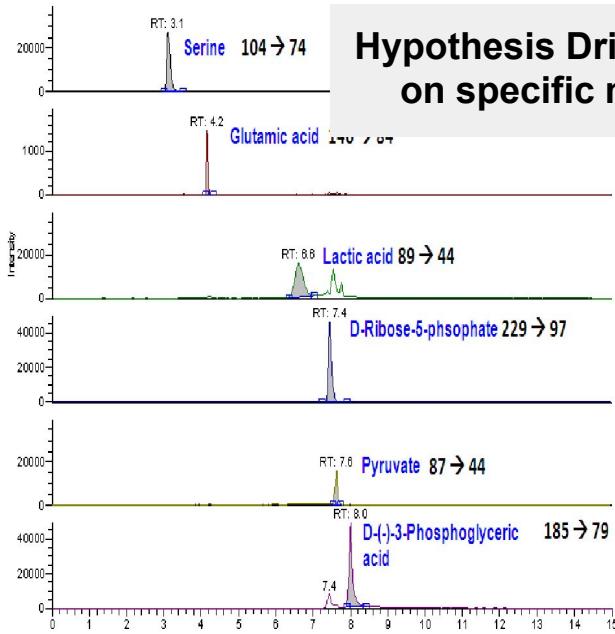
Genome biology
System microbiology,
Protein-metabolite interactions



UniversitätsKlinikum Heidelberg

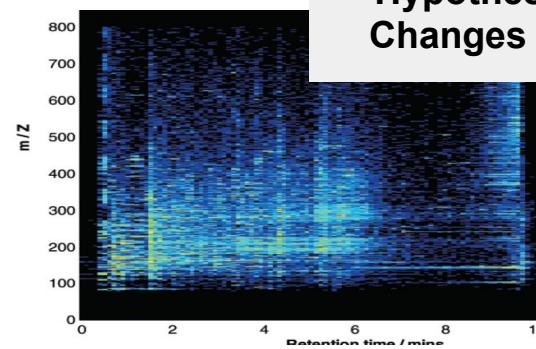
Metabolomic approaches

Targeted Metabolomics



Biomarker quantification

Untargeted Metabolomics

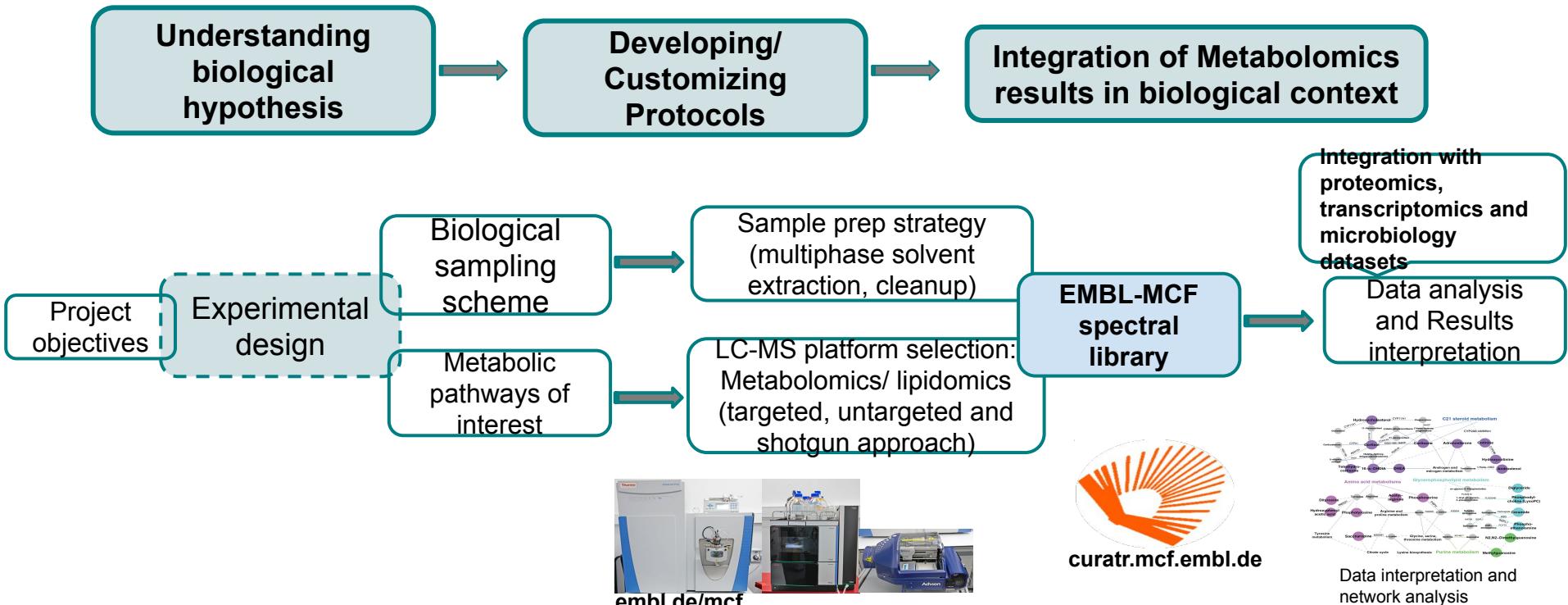


Hypothesis Generating:
Changes in metabolome



Biomarker discovery

Workflow management



Quantitative Untargeted Metabolomics

Sample amount

Metabolite extraction

Instrumental analysis (LC-MS)

Considerations

Experimental design, sample type, collection method, stability

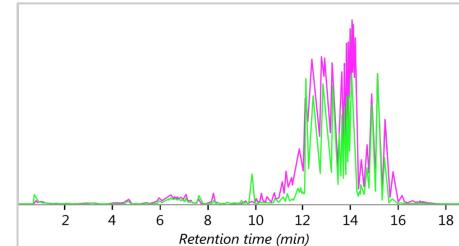
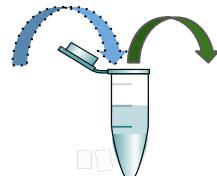
Cell numbers, weight, volume, biomass, protein/DNA concentration

Extraction efficiency, recovery, quenching

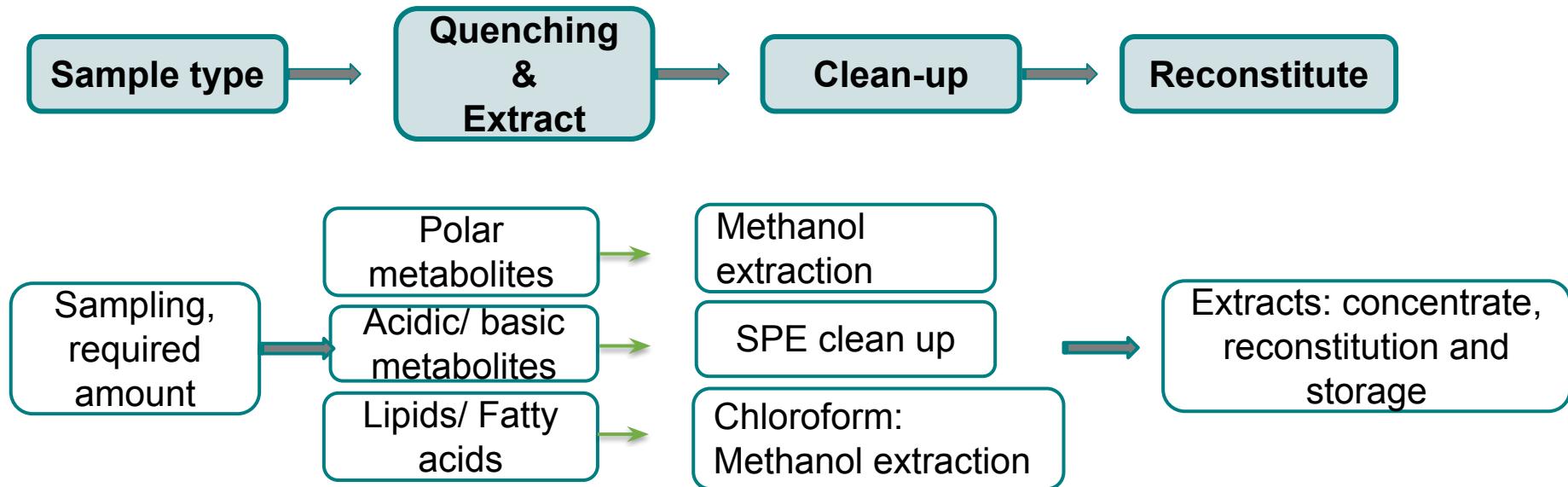
Internal standards

Reproducibility, sensitivity

Quality controls

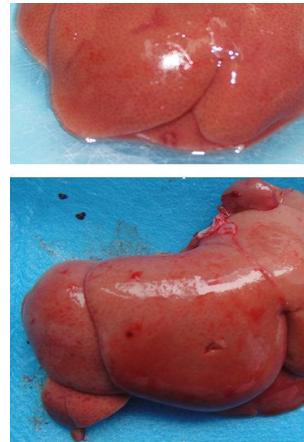


Workflow: Sample prep

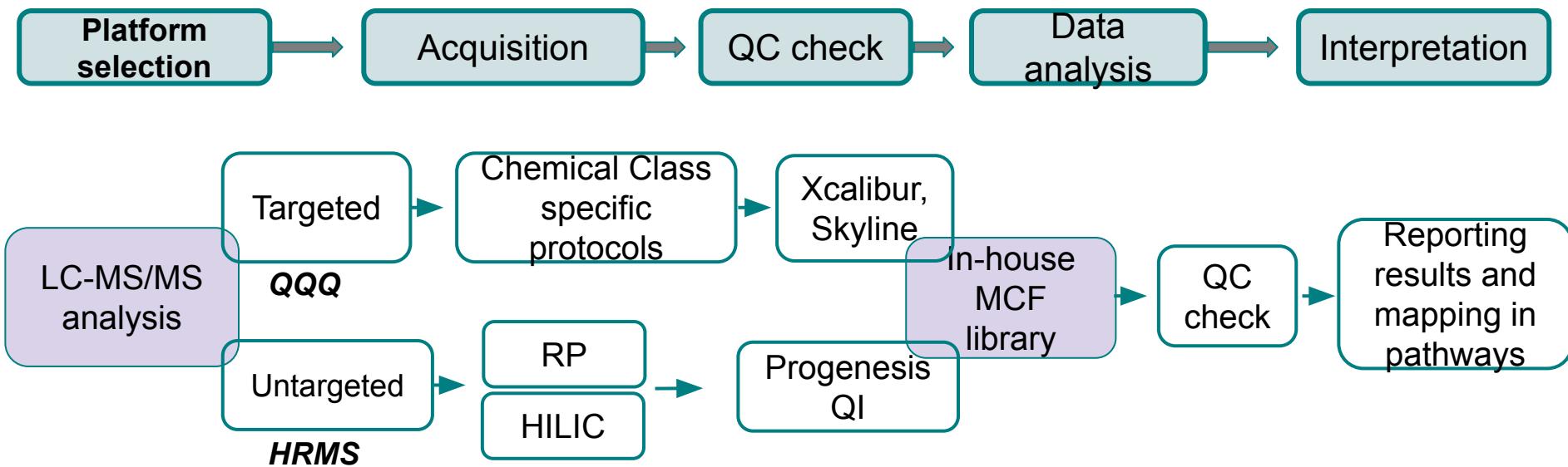


Sample specific considerations

- Cell washing
 - Quenching
 - Suspension vs Adherent
- Collection, storage
 - Weight
 - Homogenization
- Collection, storage
 - Aliquots
 - Processing, throughput



Workflow: LC-MS/MS analysis



LC-MS analysis sequence

- Randomize analysis sequence
- LC-MS run sequence
 1. Blank (*minus sample preparation*)
 2. System suitability (*standard mixture to evaluate instrument condition*)
 3. Standard mixtures (*for targeted analysis*)
 4. IS mixture (*for targeted and non-targeted analysis*)
 5. Test pooled QC samples 5 injections (*for stabilizing system*)
 6. Blank (*for checking carryover*)
 7. Samples (1 to 5)
 8. Pooled QC and blank
 9. Samples (5 to 10)
 10. Pooled QC and blank (*continues...*)
 11. System suitability (*end of analysis system evaluation*)
 12. Standard mixtures (*end of analysis system evaluation for targeted analysis*)
- Using PCA plots for QC injections
- Calculating % CV for IS spiked in all samples
- Variation within replicate injections

*

Quality Control (QC) overview

QC system 1: Injecting QCs after each 6 injections (20% injection of total analysis sequence)

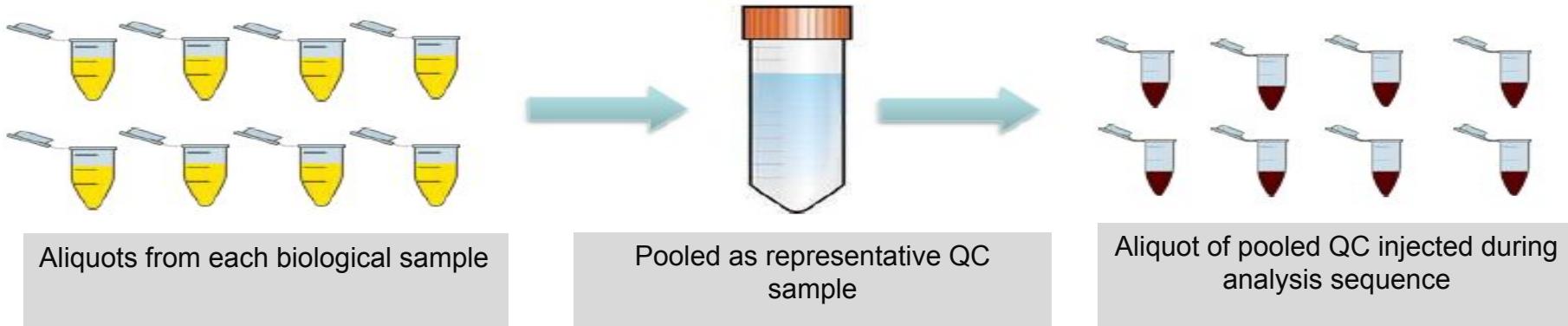
QC system 2: Internal standard spiking

QC system 3: Replicates (n=2, 3 or 5 technical replicates)



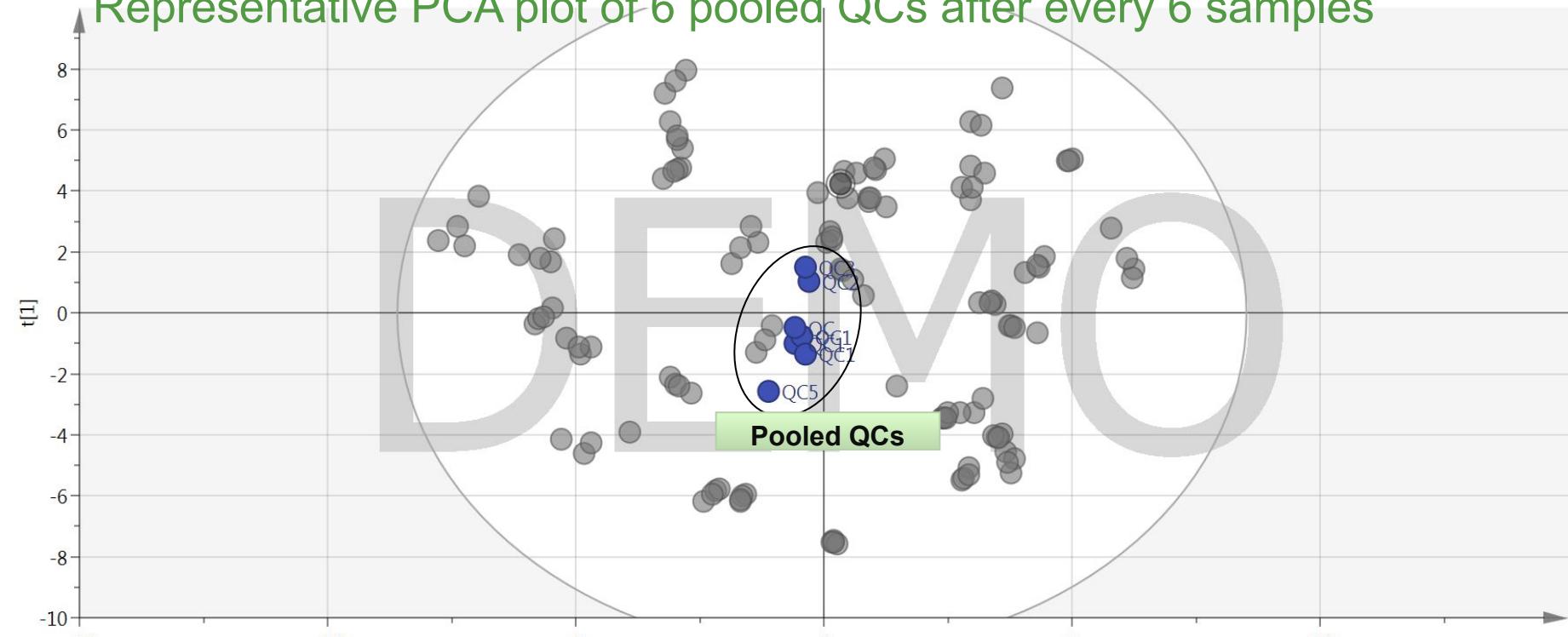
Acceptance of analysis sequence/ data
(PCA plot and % CV evaluation)

QC System 1: pooled QC



QC System 1: pooled QC

Representative PCA plot of 6 pooled QCs after every 6 samples



Data from metabolomic analysis of 35 suspension cell extract samples sequenced with 6 pooled QCs over 48 hrs of total runtime

QC System 2: Internal standards (IS) spiked in biological samples to track sample extraction, recovery and analysis variations

- Glycolytic metabolites, Amino acids (HILIC-LC-MS)

- 13C Creatinine and 13C pyruvate
 - Uridine-13C9,15N2 5'-triphosphate
 - 13C,15N Cell free amino acid mixture

- Lipids (RP lipid + ve mode)

- 13C labeled triolein
 - d5-DG ISTD Mix I (Avanti LM-6001)
 - Cardiolipin Mix I (Avanti LM-6003)
 - d5-TG ISTD Mix I (Avanti LM-6000)
 - PIP2 (37:4)
 - Fatty acids (RP lipid -ve mode)
 - 13C Palmitic acid
 - 13C Algal fatty acid extracts
 - Steroids, Purine, drugs (RP + ve mode)

- 13C3 Testosterone

- Sulfachloropyridazine

- Sulfadimethoxine

- Amitriptyline

QC System: acceptance criteria

- **Randomization:** Samples analysis sequence is randomized
- **System suitability:** Respective IS mixture (in previous slide) injected before analysis
- Blank and QC samples are 20 % of total batch
- **Replicates (n=3):** CV > 20 %
- **QC sample variation across analysis sequence:** CV > 20 % for targeted metabolites or at least 75 % features (for non-targeted analysis)
- **Spiked IS variation:** CV > 20 % for all spiked internal standards
- Data normalization if necessary using IS response or QC variation

Data analysis

Software

Commercial/ Freeware

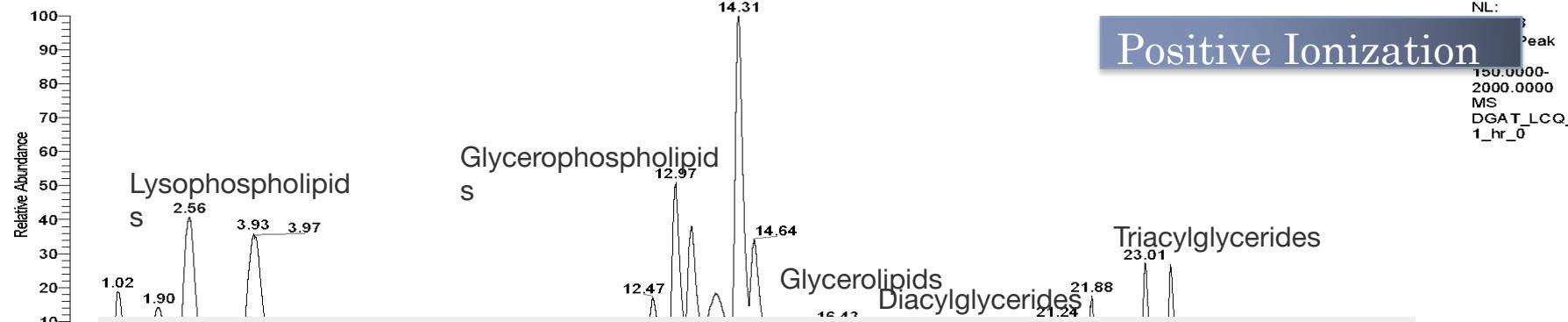
- Progenesis QI (Untargeted)
- Xcalibur
- Skyline (Targeted)

In-house developments (by Alexandrov Team)

- Optimus
- MCF spectral library (Curatr)

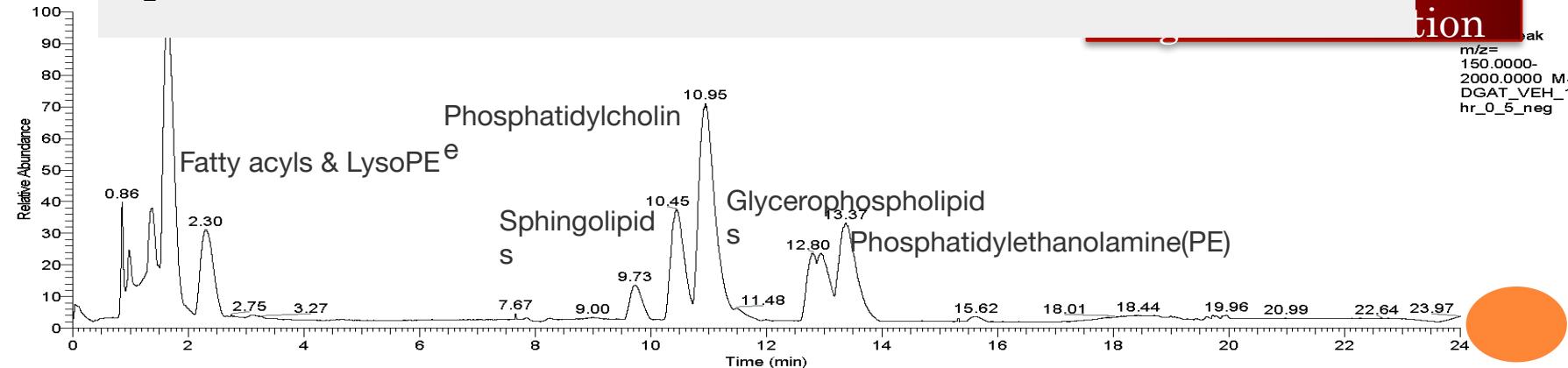
LIPIDOMIC PROFILING IN RAT AND HAMSTER PLASMA

RT: 0.00 - 30.00 SM: 15B

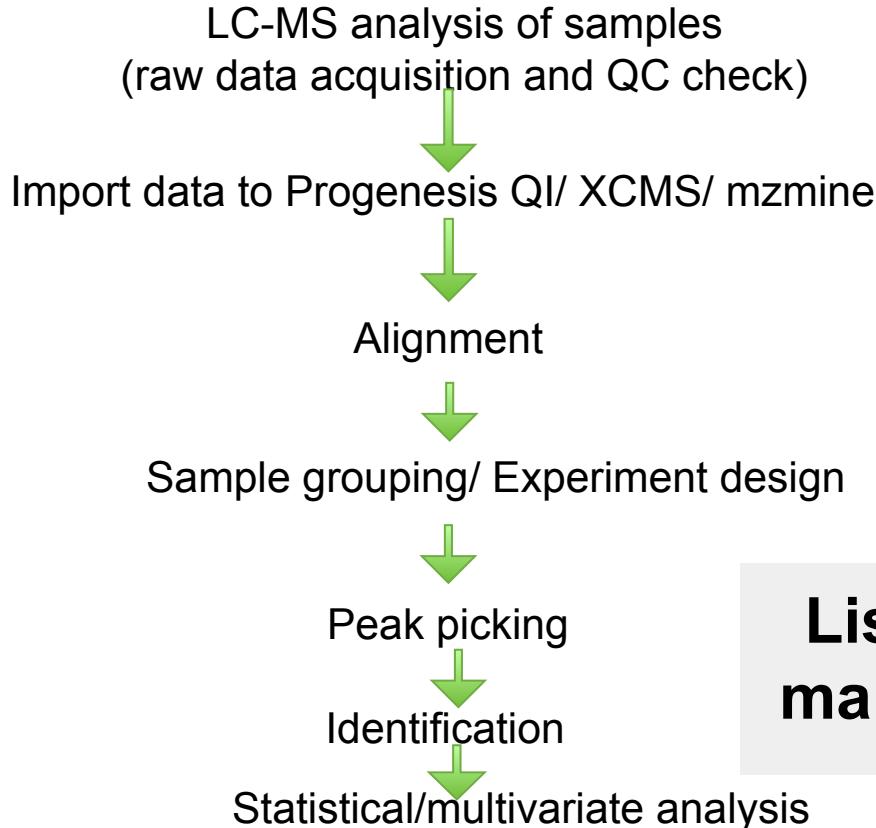


Positive Ionization

Spectra → Features (reliable signals of unique m/z and RT pairs)



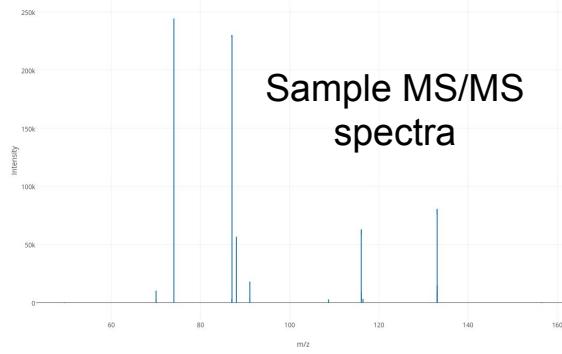
Data analysis: Non-targeted profiling



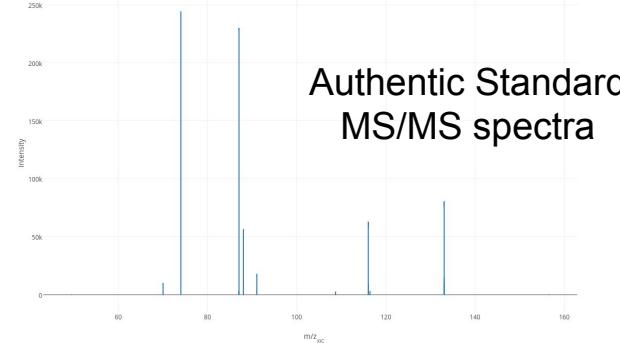
**List of Metabolite
markers of interest**

Metabolite Identification

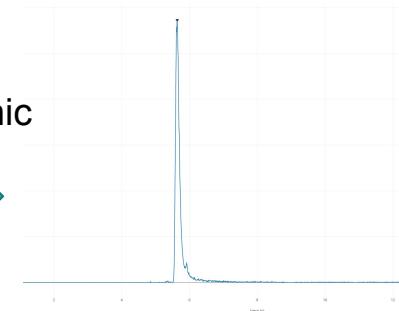
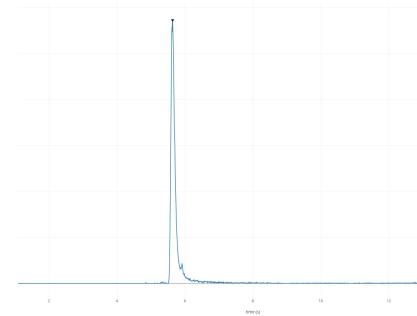
Level 1



Spectral Matching



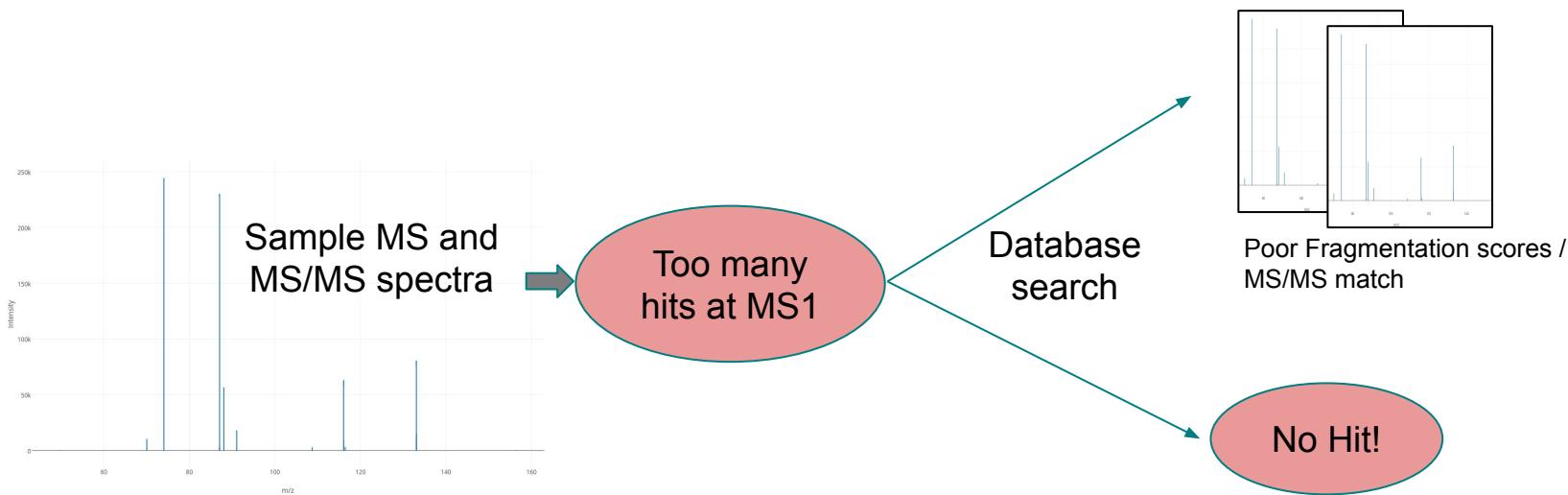
Matching chromatographic profiles



L.W. Sumner et al; Metabolomics (2007)

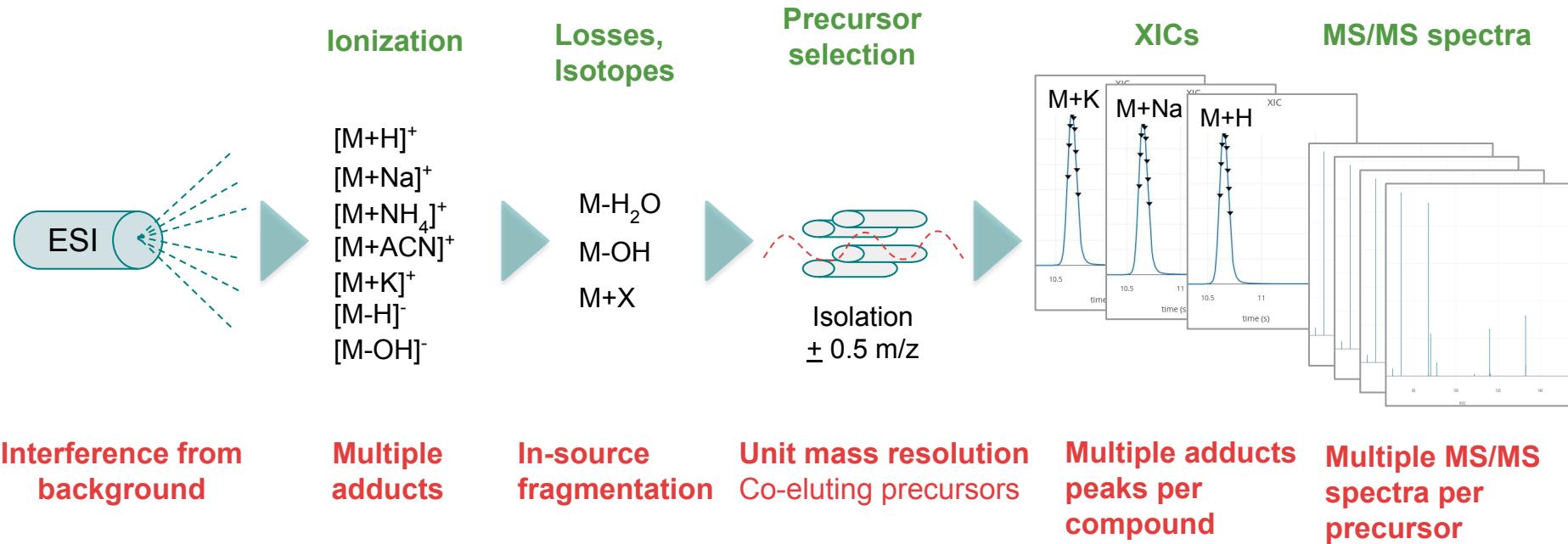
Limited Availability of authentic standards

Spectral database search for Untargeted metabolomics



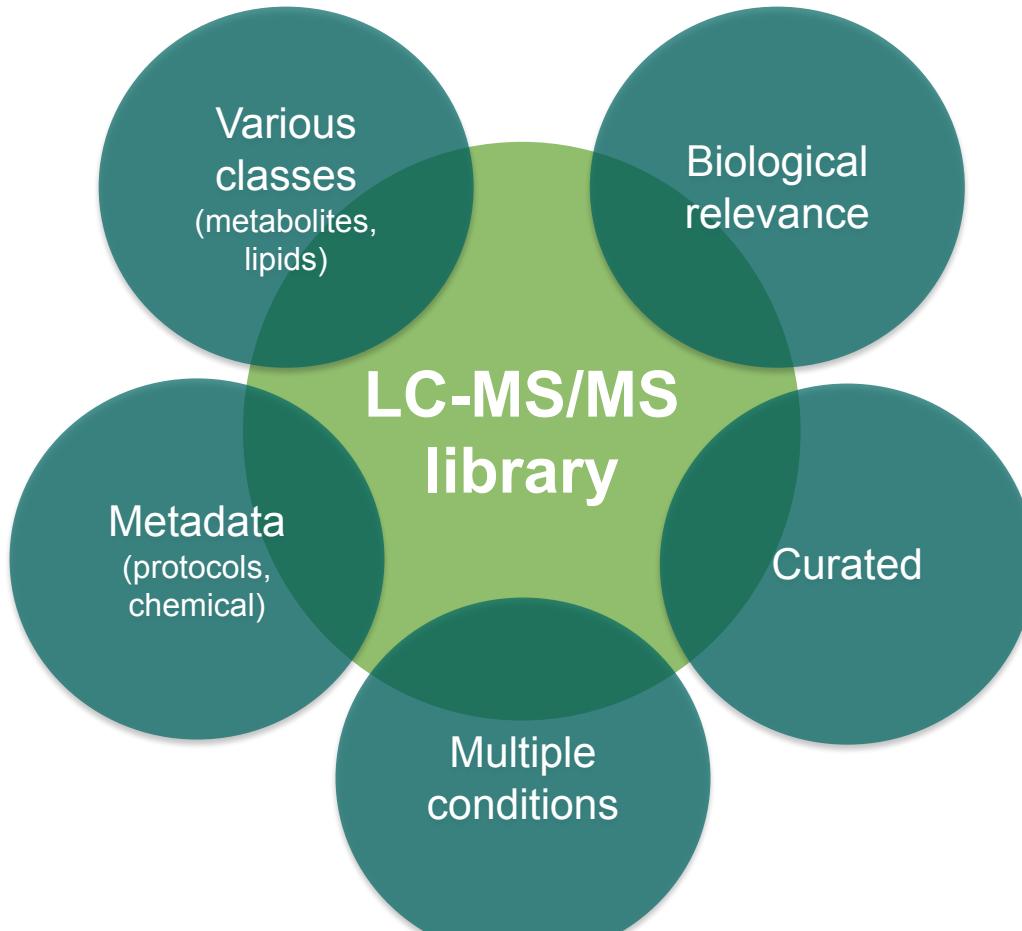
- Noisy features/ peaks
- Public libraries are not specific to Metabolomics study of your interest
- **Non-transferable LC-MS/MS spectra: different LC-MS conditions**
- False positive hits for exogenous compounds

Challenges of Metabolite ID in LC-MS/MS data



Curated open-access LC-Orbitrap-MS/MS spectral library

LC-MS/MS spectral library for metabolomics



Available online ESI-MS/MS spectral libraries

| Library | Compounds with spectra | Total MS/MS spectra | Orbitrap | LC info | Open access |
|---------|------------------------|---------------------|----------|---------|-------------|
| mzCloud | 6,585 | 2,045,858 | | | |

No fully open-access LC-Orbitrap-MS/MS library for endogenous metabolites

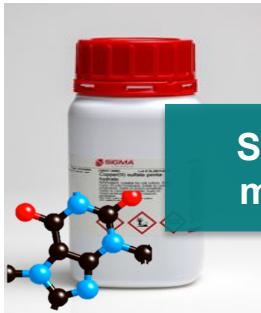
| | | | | | |
|----------|--------|--------|--|--|--|
| Massbank | 15,828 | 43,734 | | | |
| GNPS | 10,587 | 18,263 | | | |
| HMDB | 3,500 | 7228 | | | |

Yes

Limited

No

Creating a spectral library: Our workflow



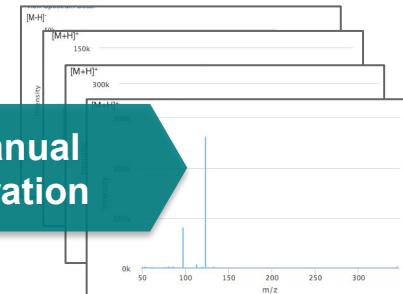
Standard mixtures



LC-MS method development



Analysis of standard mixtures



Manual curation

800 standards
10-15 in a mix

8 LC-MS protocols
ESI +ve, -ve, HILIC, C18

Evaluate XICs
Curate MS/MS spectra

20,000 €

3 months of development

300 h analysis time

10 min per spectrum

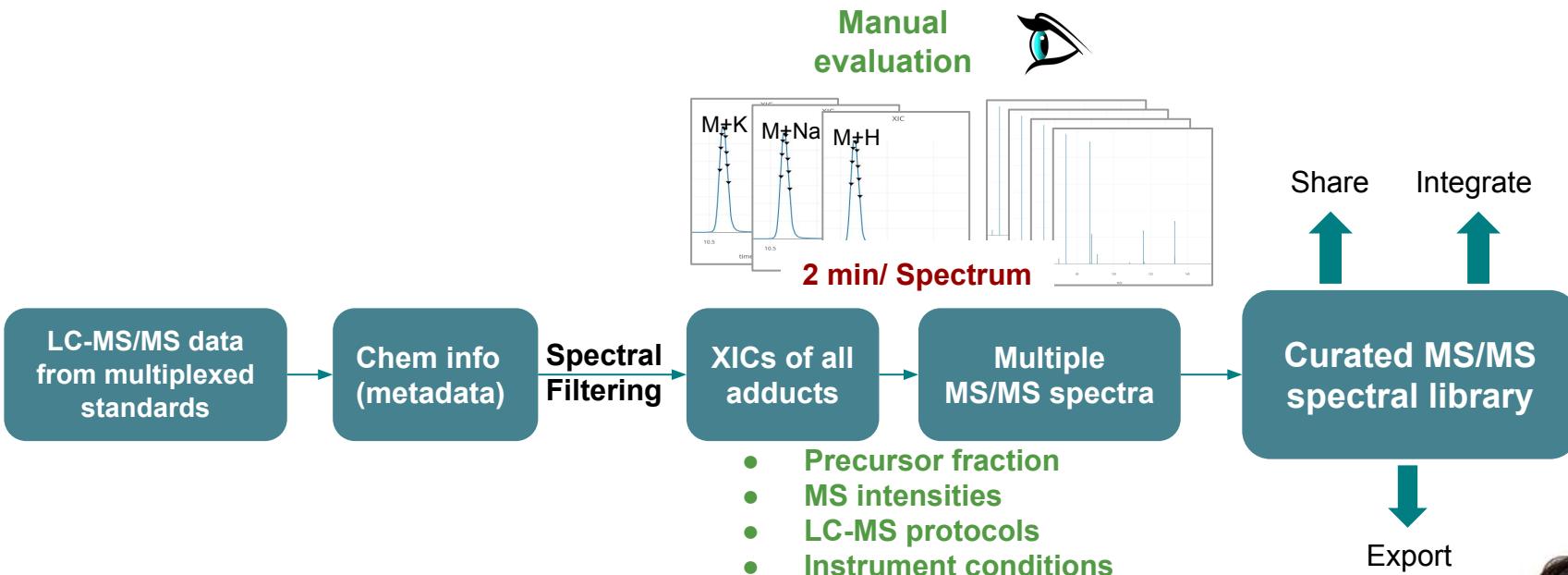
Resources

Time

Efforts

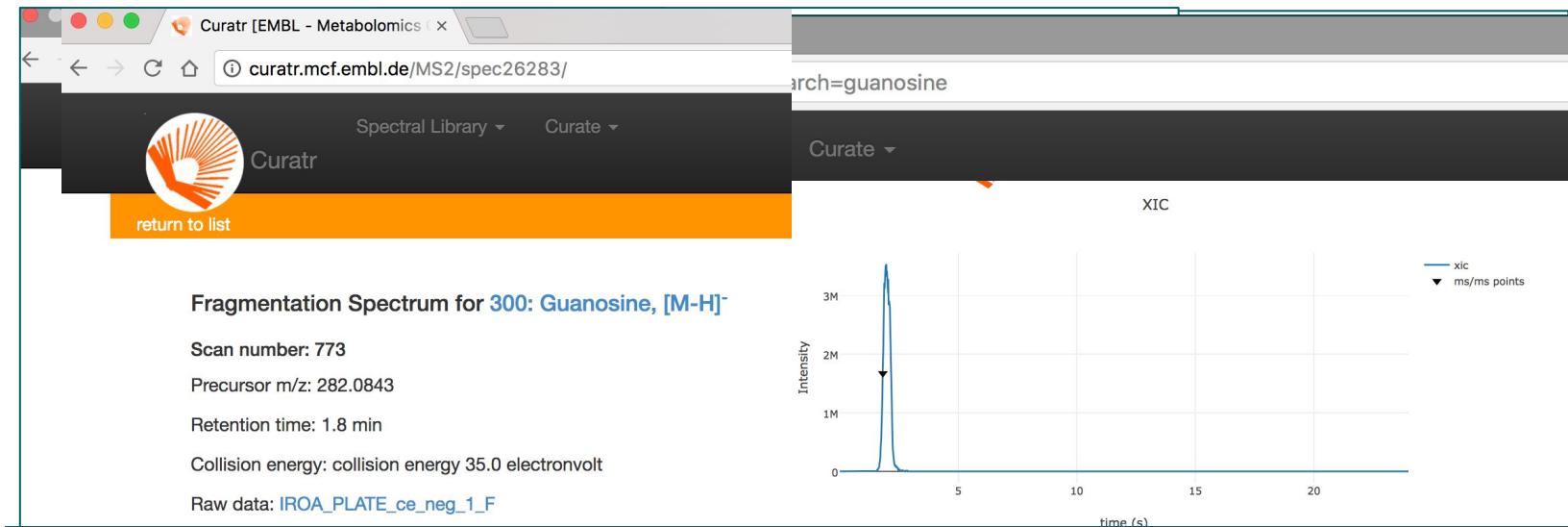
Our solution: web application Curatr

<https://github.com/alexandrovteam/curatr>



Andrew
Palmer

Web interface of Curatr

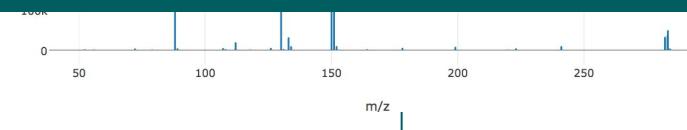


No Additional Software
Curate online
Easy Visualization
Fast

Date added: April 8, 2016

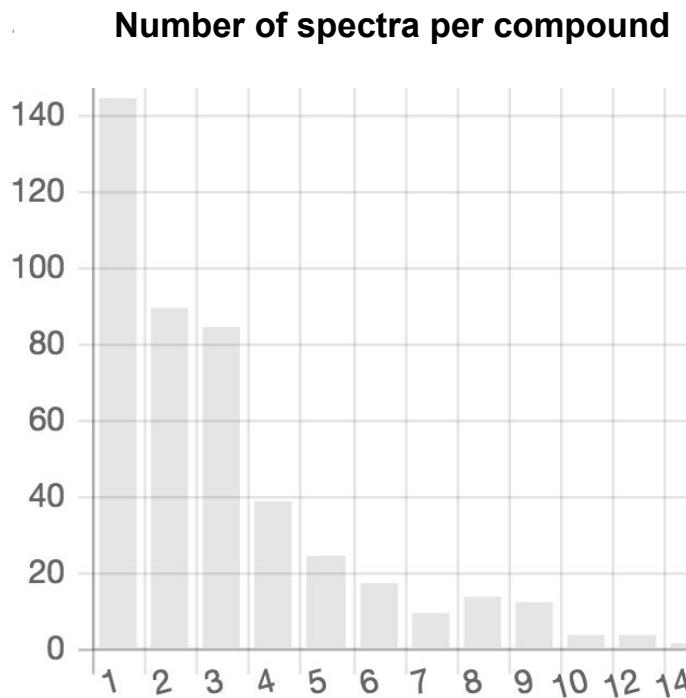
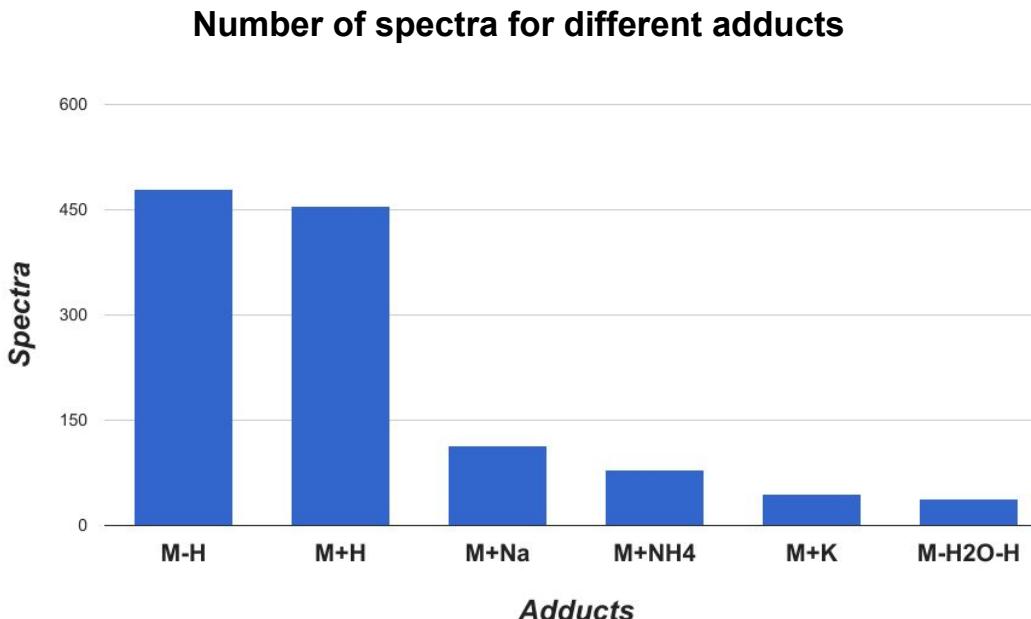
Date curated: April 10, 2016

Curator: prasad

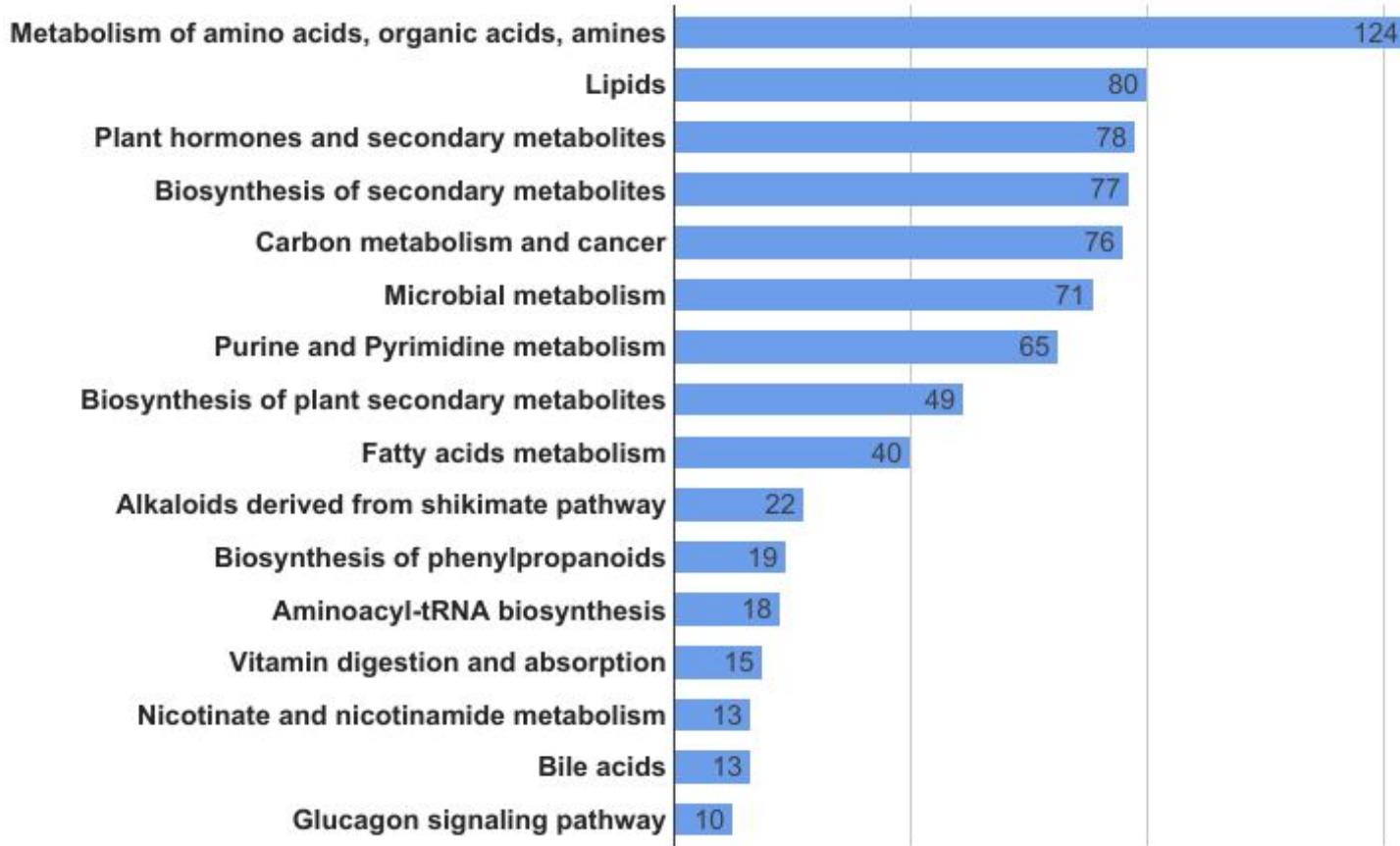


EMBL-MCF spectral library: Statistics

1,543 spectra for 439 compounds



Covered KEGG pathways

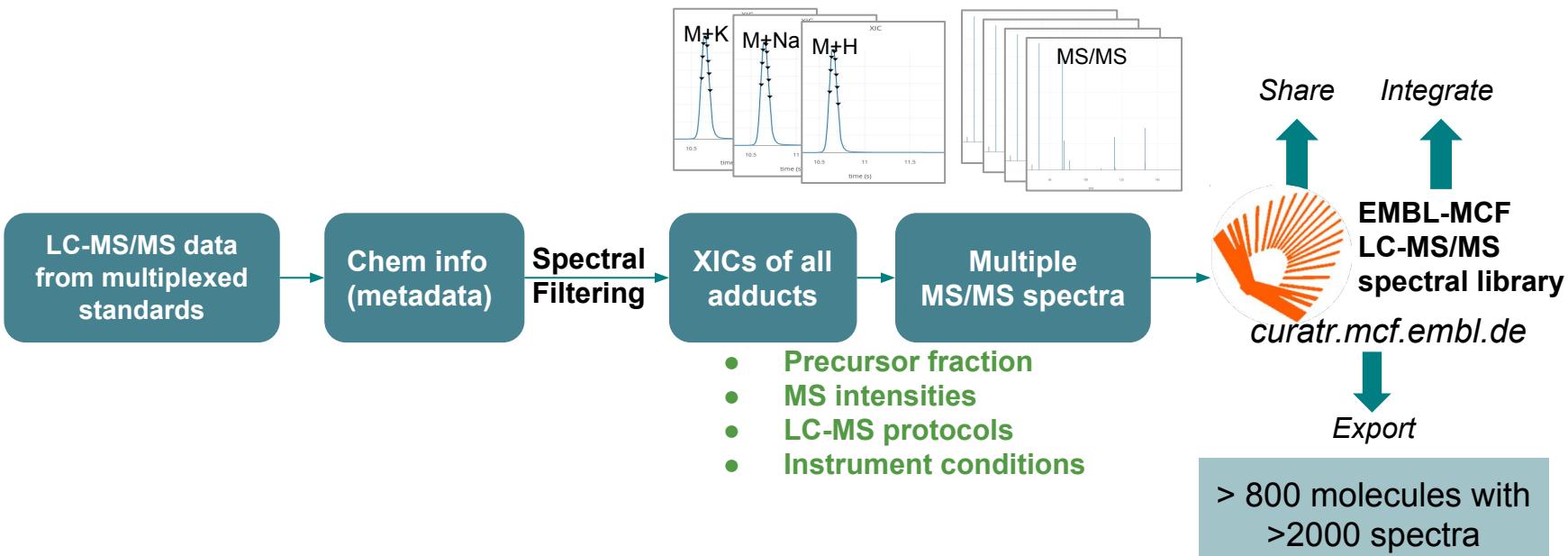


Open access, export and sharing

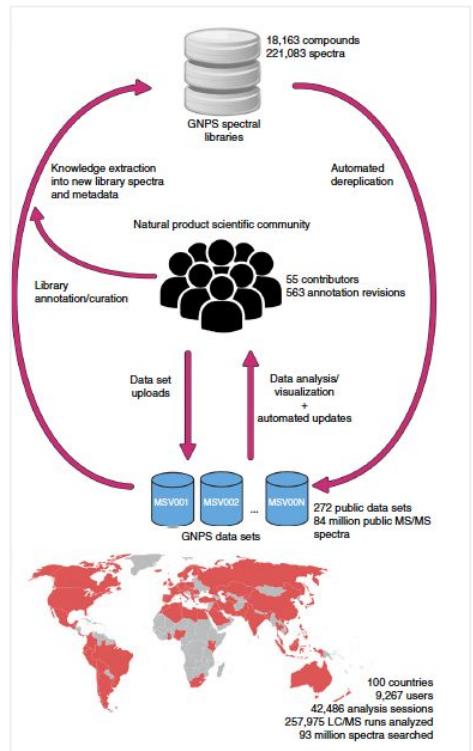
<http://curatr.mcf.embl.de>

- Open access
 - Creative Commons license (CC-BY)
- Downloadable in various formats
 - .TSV (as spreadsheet)
 - .MGF (spectral search)
 - **.MSP** (Text-based NIST)
- Integrated with spectral repositories
 - GNPS: included
 - MetaboLights: export
 - MassBank EU: on-going

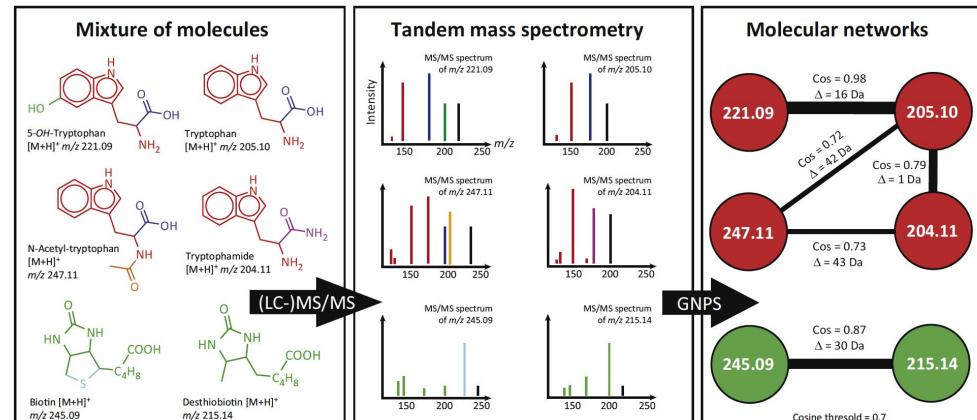
EMBL-MCF LC-MS/MS spectral library



Integration with ‘Global Natural Product Social Molecular Networking (GNPS)’



Molecular Networking: A Graph-Based Tool to Explore Spectral Similarity in Liquid Chromatography–Tandem Mass Spectrometry (LC-MS/MS) Data from Molecular Mixtures



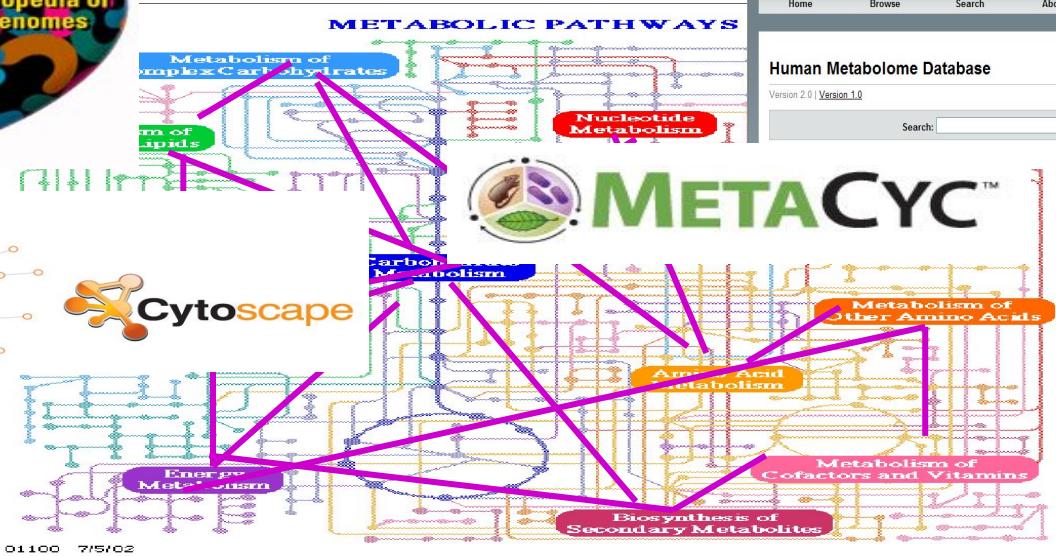
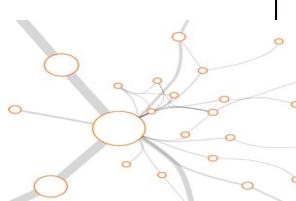
Trends in Pharmacological Sciences

Figure 1. Molecular networks from LC-MS/MS spectra of tryptophan and biotin derivatives (from EMBL MCF spectral library on GNPS). The interactive view of the molecular network can be visualized directly on GNPS via the following link <http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=0415e5eef1d68149297b8aef26a480af9>.

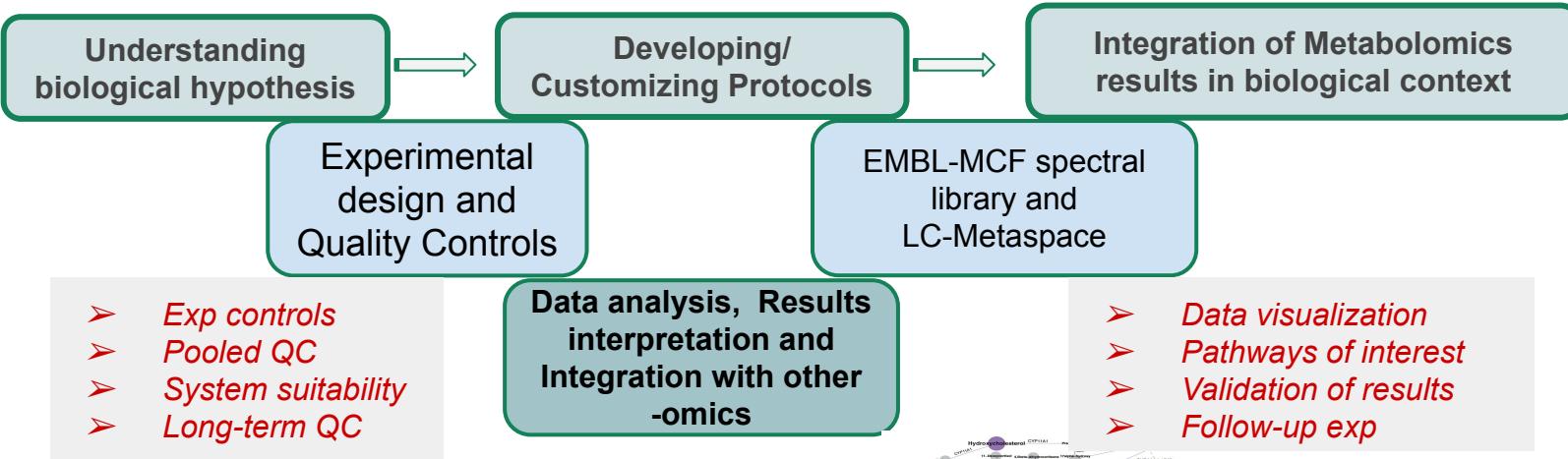
Quinn et al. CellPress-Trends in Pharmacological Sciences (2016)

METABOLIC PATHWAY MAPPING FOR IDENTIFIED METABOLITES

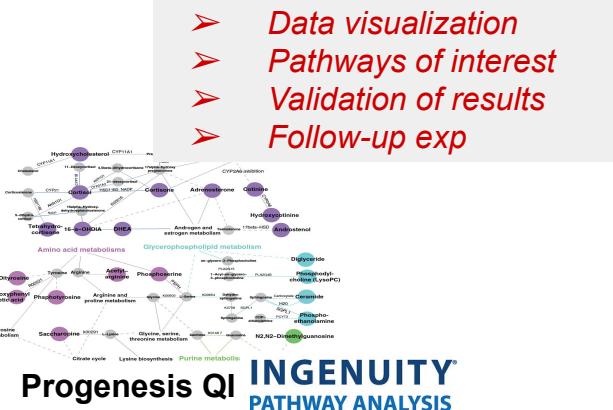
- Visualize biochemical network of metabolites
- Understand functional role of metabolites and their pathways including their association enzyme, protein and genomic information
- System biology understanding of metabolomic data



Workflow management



EMBL-MCF spectral library



Progenesis QI INGENUITY®
PATHWAY ANALYSIS

Resources

E-learning protocols
On [YouTube](#) EMBL channel →

- EMBL-EBI MetaboLights repository
- metabolomics-forum.com

The image displays two screenshots. The top screenshot shows the EMBL YouTube channel page for 'Metabolomics workflow overview (short tutorials)'. It features a video thumbnail of a man speaking, a 'PLAY ALL' button, and a list of five short tutorial videos. The bottom screenshot shows the homepage of the MetaboLights repository, which includes sections for 'Study', 'Compound Library', and 'Training'.

Metabolomics workflow overview (short tutorials)

5 videos • 256 views • Last updated on Apr 26, 2019

1 Sample preparation: Introduction (tutorial 1/5)
European Molecular Biology Laboratory (EMBL)
1:26

2 Metabolomics experimental design (tutorial 2/5)
European Molecular Biology Laboratory (EMBL)
3:12

3 Sample preparation: Metabolite extraction (tutorial 3/5)
European Molecular Biology Laboratory (EMBL)
3:25

4 Metabolomics platforms at EMBL (tutorial 4/5)
European Molecular Biology Laboratory (EMBL)
2:35

5 EMBL-MCF LC-Orbitrap-MS/MS spectral library (tutorial 5/5)
European Molecular Biology Laboratory (EMBL)
3:27

Study

BROWSE

ORCID SEARCH

METABOLIGHTS LABS
COMING SOON

Compound Library

COMPOUNDS

SPECIES

Training

TRAINING ONLINE

QUICK TOUR

Acknowledgement

Alexandrov Team/ MCF

Theodore Alexandrov

Andrew Palmer

Maria Naumenko

MCF users

Dorrestein Lab - UCSD

Past group members

Thank you

'Heidelberg Molecular Life Sciences (HMLS) Investigator Award' for Advion Nanomate instrument

NIH Common Fund Metabolomics program for free ¹³C standards

EMBL Core Facilities

