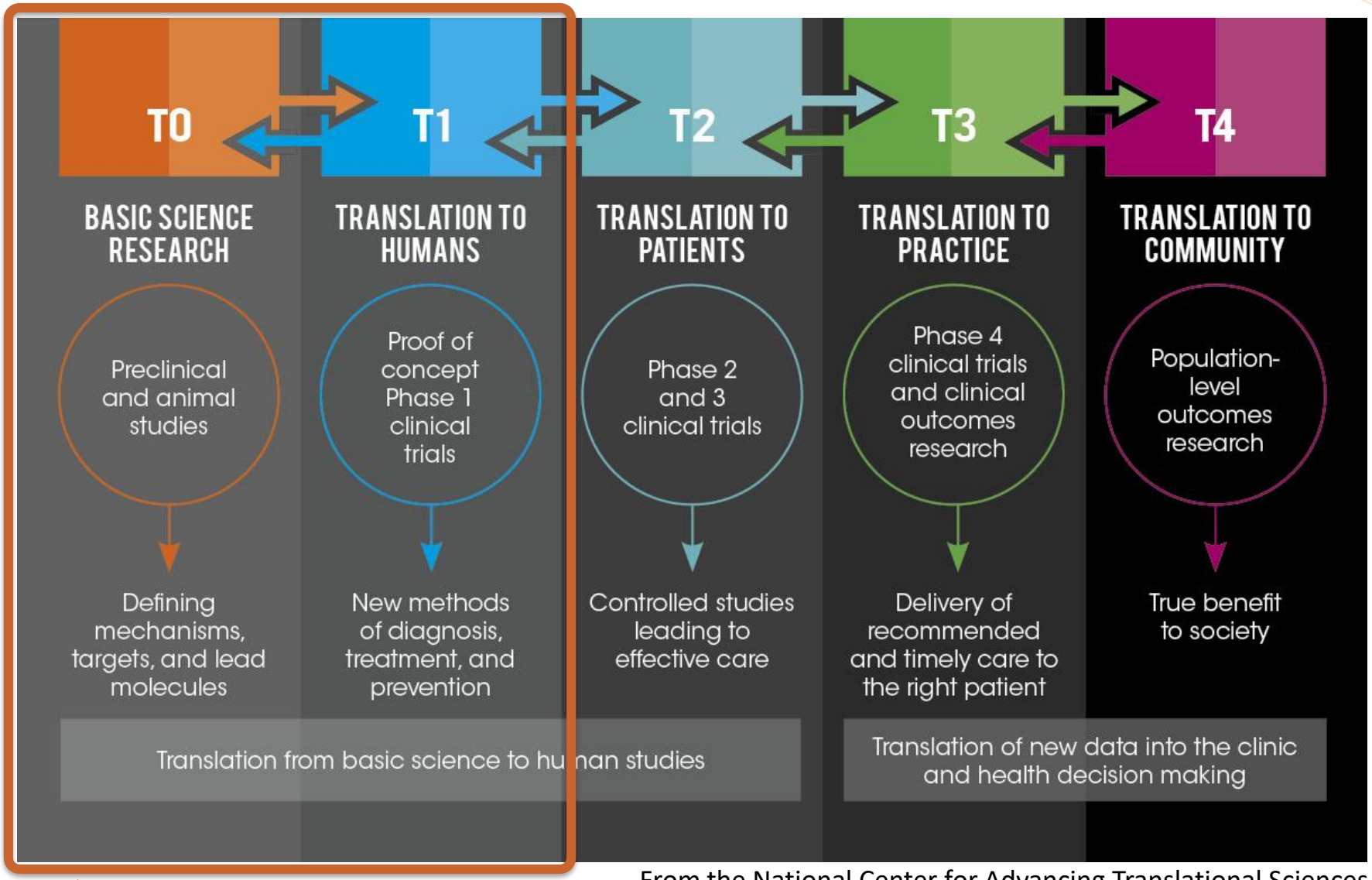
Omics hub

BIOASTER
MICROBIOLOGY TECHNOLOGY INSTITUTE

Technological integration in metabolomics & proteomics for translational research

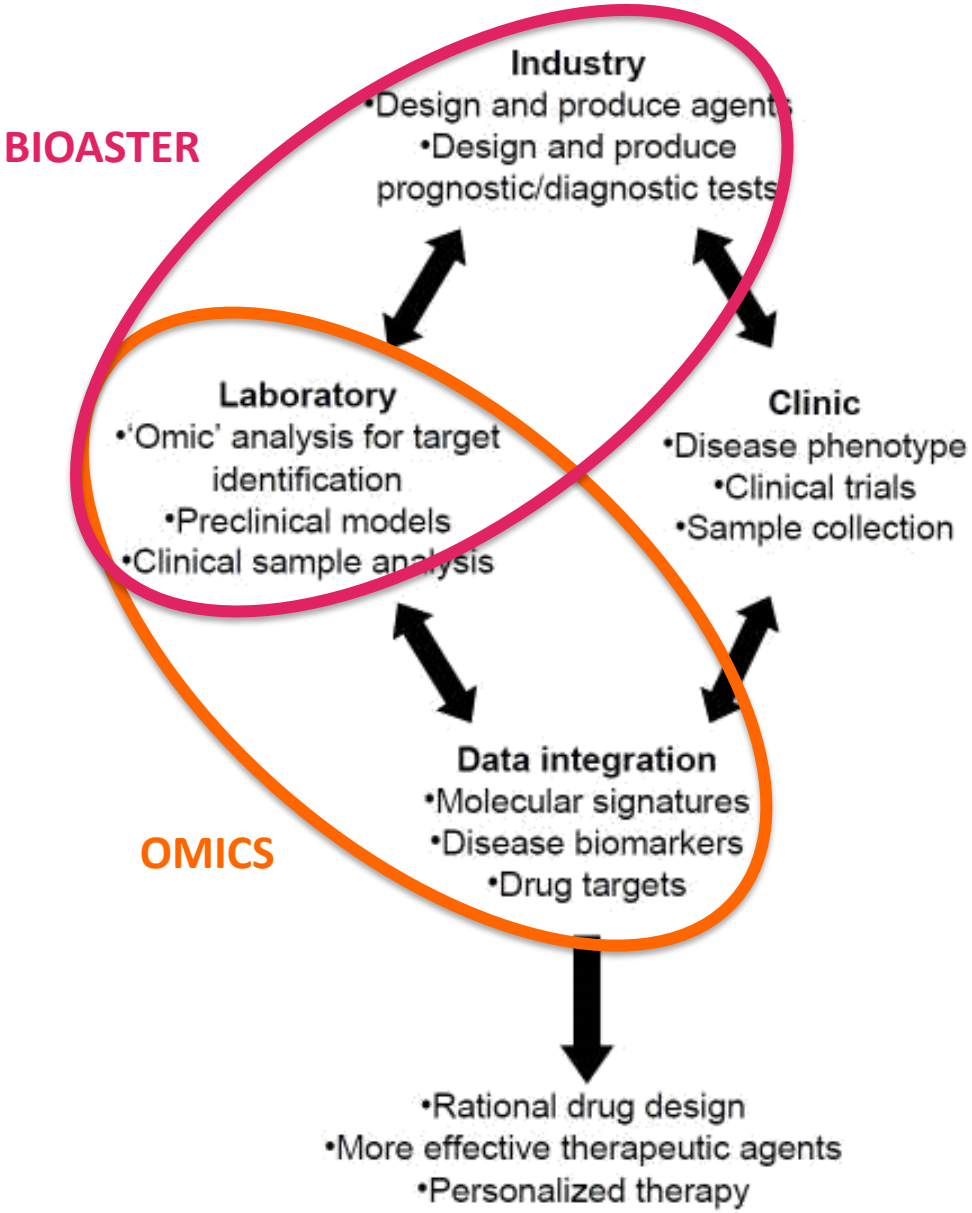
Frédéric BEQUET, PhD
Head of Omics hub

Translational research: Translation Spectrum



From the National Center for Advancing Translational Sciences

Translational research – A central role for Omics



EM Goldblatt &W-H Lee, *Am J Transl Res.* 2010

Translational research –Main challenges

Omics need to address two main questions in translational projects:

The **understanding of biological mechanisms** (diseases, treatments, microbiota), by addressing both the microbes and the host response

- Samples diversity/complexity
- Comprehensive analysis

The identification and valida **biomarkers & signatures** for diagnostics or therap

- overy
- umber/low volume of samples

TECHNOLOGIES INTEGRATION

Bioaster need to address industry problematics:

- Budget
- Time constrains

Comprehensive analyses

Integration of NMR & MS for **comprehensive** metabolomics/proteomics

NMR

polar metabolites

- Quantitative
- High-throughput
- Non destructive
- Reproducible
- Low sensitivity

Complementarity of metabolic species

LC-HRMS

non polar metabolites

- High sensitivity
- High coverage
- High dynamic range
- Low-throughput
- Metabolite-specific

proteins

- High sensitivity
- High coverage
- High dynamic range
- Complex matrixes
- Low-throughput

Complementarity of properties

Complementarity of technologies

Structure elucidation or confirmation

Metabonomics

NMR for polar metabolites analysis

Equipment:

600 MHz Avance III HD NMR from Bruker

- 1,7mm / 5mm cryogenic inverse probes for 1H/13C/15N/P12 detection
- SampleJet for automated sample analysis

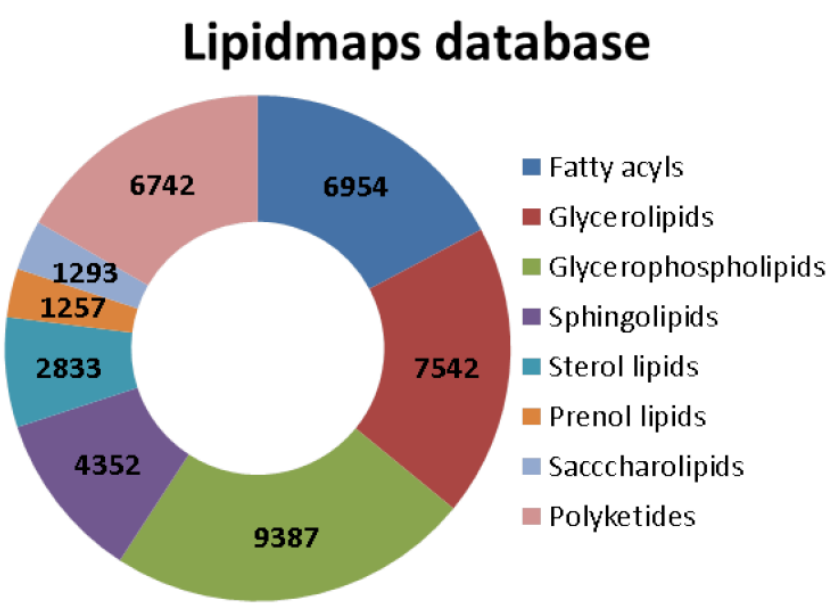
Analysis:

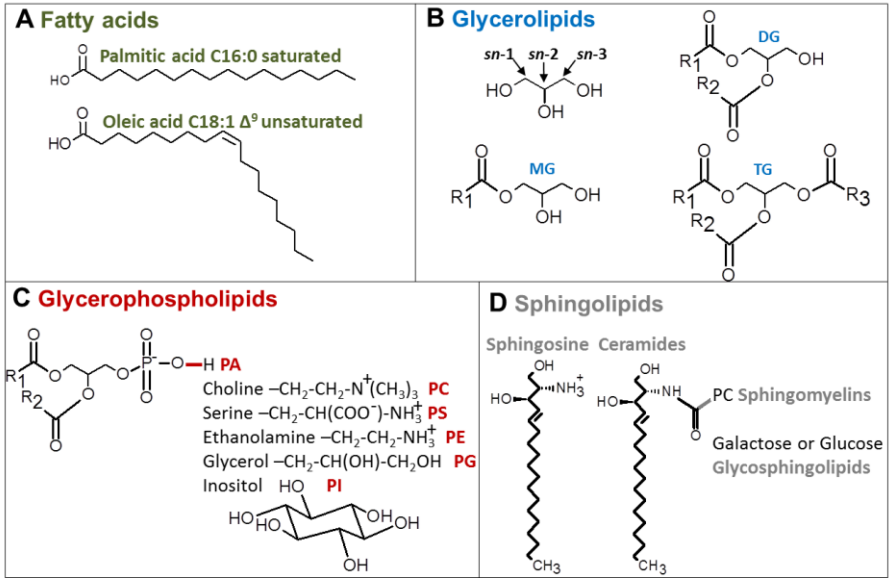
1D H^1 NMR, NOESY sequence

Coverage of all canonical pathways
(glycolysis, TCA cycle, pentose-phosphate way…)

→ ~15 min run

LC-MS for lipids analysis



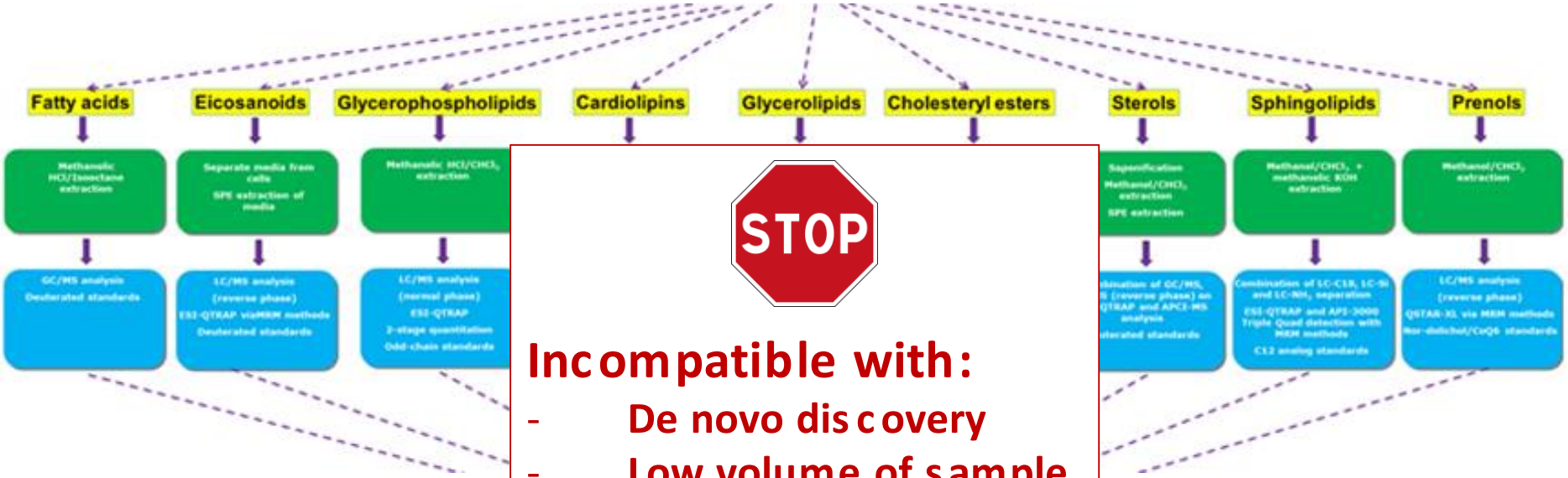


2 POSSIBLE STRATEGIES

targeted or untargeted analyses

Lipidomics

Targeted analyses



STOP

Incompatible with:

- **De novo discovery**
- **Low volume of sample**
- **High number of sample**
- **Time constrains**
- **Budget**

Adapted from the Lipid maps consortium)

Quantitative m **n** metabolites

→ Hypothesis driven

Lipidomics

LC-HRMS Untargeted analyses

= global unbiased analysis, **semi-quantitative** → hypothesis generating

Equipment:

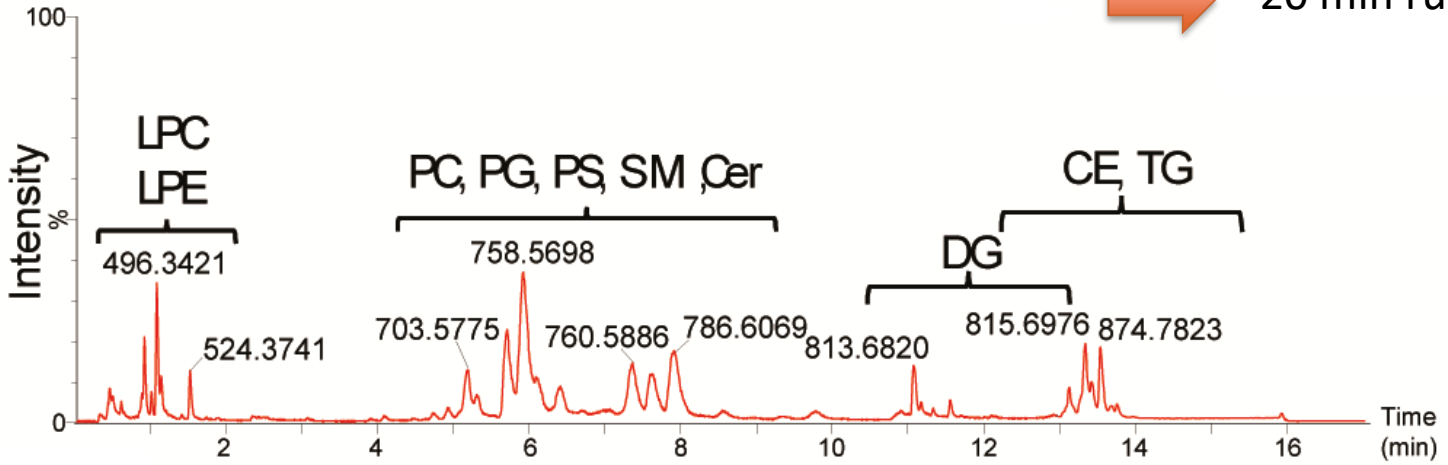
Thermo Scientific™ QExactive™ HF

- UHPLC Dionex 3000

Analysis:

- Full scan data acquisition (m/z ratio from 100 – 1250 Da), at a resolution of 70 000
- Automated Data Dependent Acquisition of MS/MS spectra, at a resolution of 35 000
- Two separate injections for acquisition in positive and negative ionization modes

→ ~20 min run (x2)



Proteomics

LC-HRMS for proteomics

= Bottom-up approach

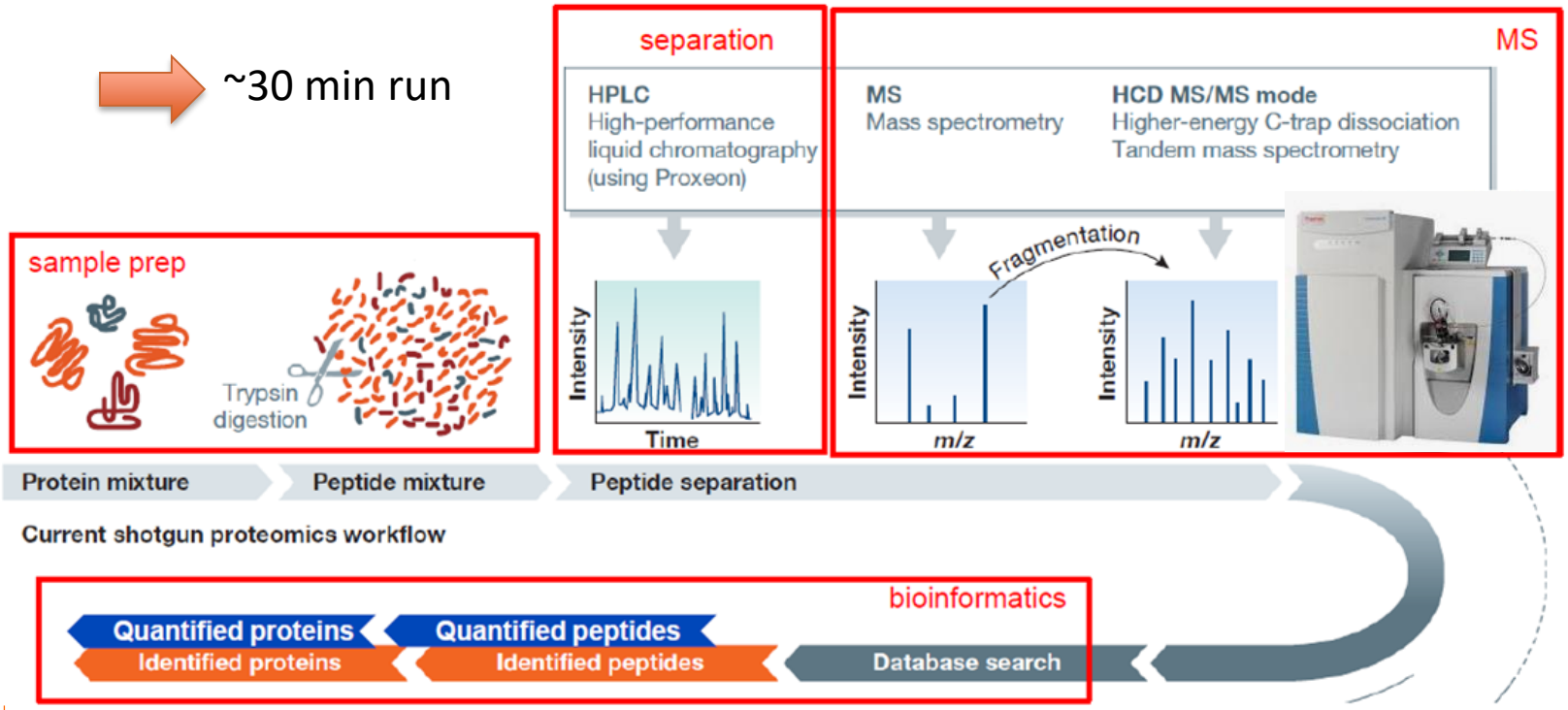
Equipment:

2 Thermo Scientific™ QExactive™

- UHPLC Dionex 3000

Analysis:

- Full scan data acquisition (m/z ratio from 400 – 1250 Da), at a resolution of 70 000
- Automated Data Dependent Acquisition of Top20 MS/MS spectra, at a resolution of 17 500



High number/low volume of samples

Integration of automated sample preparation for:

- Combination of sample preparation for MS/NMR
  - Robotic head with both long needles (NMR) and tips (MS)
- Fast, accurate, robust extraction
- High throughput for high number of samples (e.g. clinical trials)

- Sample Management System (LIMS)
  - Internal solution: NoE

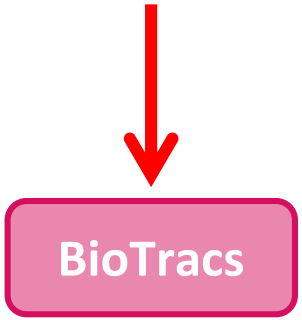
NoE
Network of Experiments

De Novo discovery

Integration of bioinformatics for:

- MS & NMR data combination
- Automated data processing
- Automated metabolites/proteins identification
- Automated QC evaluation
- Automated biostatistical analyses

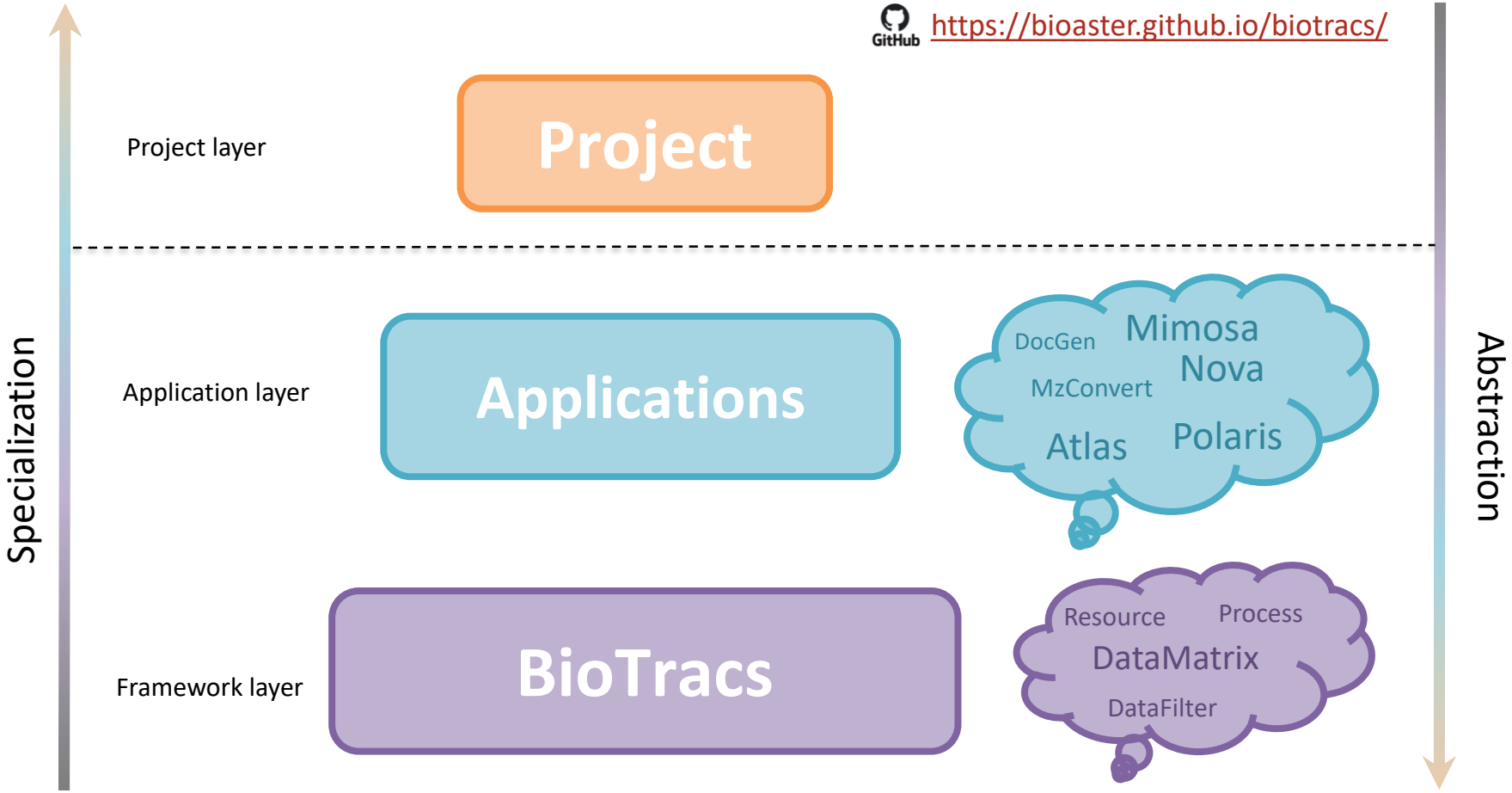
No existing tool reaching all these specifications



An in-house bioinformatics solution for complex analytical workflows

BioTracs

BioTracs is a computation framework that allows implementing complex analytic workflows while ensuring *traceability* and *transparency* in computational processes.



BioTracs-Mimosa/Polaris

BioTracs-Mimosa/Polaris are transversal computational applications based on BioTracs framework for the analysis of *MS metabolomics & proteomics data.*

BioTracs-Mimosa/Polaris

Mass spectrometry data processing

BioTracs Mimosa
Mass spectrometry data processing

Raw data

BioTracs-MzConvert
Raw files conversion

BioTracs-OpenMS
Feature extraction

BioTracs-Metadata
Sample data annotation

Metadata

NoE
Network of Experiments

Processed data

Feature processing

- ✓ Filtering
- ✓ Gap filling
- ✓ QC CV filtering
- ✓ Intra and inter-batch drift correction
- ✓ Quality control assessment
- ✓ Feature grouping

BioTracs-Mira
Metabolite/Proteins identification
(LipidMatch, Mascot)

LipidMatch flow

BioTracs-Atlas
Statistics & Machine learning

BioTracs-NOVA
NMR data processing

Workflow for a (pre)clinical study with low volumes

from 150µl of Plasma

Samples Management (LIMS)

Documented Biological Samples

Robotics
BSL2/BSL3

Automated sample preparation

NMR
High Resolution MS

~1850 lipid features
~750 metabolites features
~800 putative proteins

Untargeted Analysis
DISCOVERY

Bioinformatics

Data processing

Discriminant analyses

Markers ID

Biomarker Candidates
SIGNATURES

~850 lipids
~250 metabolites
~280 proteins

Database
Biostatistics

Advanced analyses

Pathway enrichment
Functional analyses

MECHANISM UNDERSTANDING

Data integration

Limitations and sources of improvement

Metabolomics

- Limitation of NMR sensitivity
  - Problematic for intracellular metabolomics
- Limited number of metabolites addressed

Proteomics

- Place for improvement in proteins identification

Limitations and sources of improvement

Metabolomics

- Limitation of NMR sensitivity
  - Problematic for intracellular metabolomics
- Limited number of metabolites addressed

→ **HILIC**

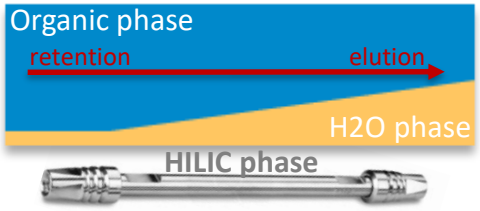
Proteomics

- Place for improvement in proteins identification

Hydrophilic Liquid Interaction Chromatography (HILIC) Workflow

Sample preparation
Polar metabolites extraction

Metabolomics: Hilic approach

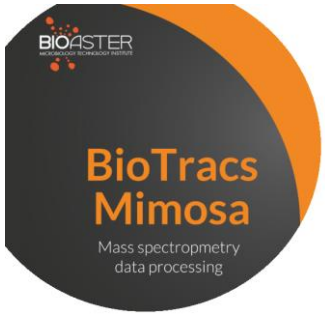


Thermo Scientific™
UHPLC Vanquish™

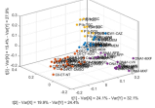
Q Exactive HF ™

Bioinformatics

Data pre-processing

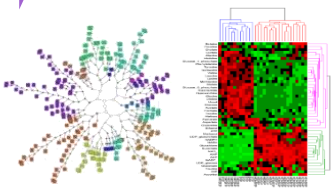


Discriminant analyses



Markers: identification

Pathway analysis

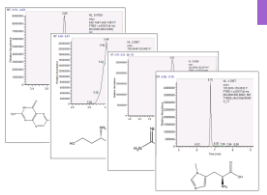


"HilicMatch"

Biological matrix

Standards injected
n≈700

Ion extraction



In house database

m/z
RT
Type of ion
Correlation with other ions

- 700 new accessible metabolites, in a scalable database
- Complementary to NMR analyses

Workflow in place
Database under construction

Limitations and sources of improvement

Metabolomics

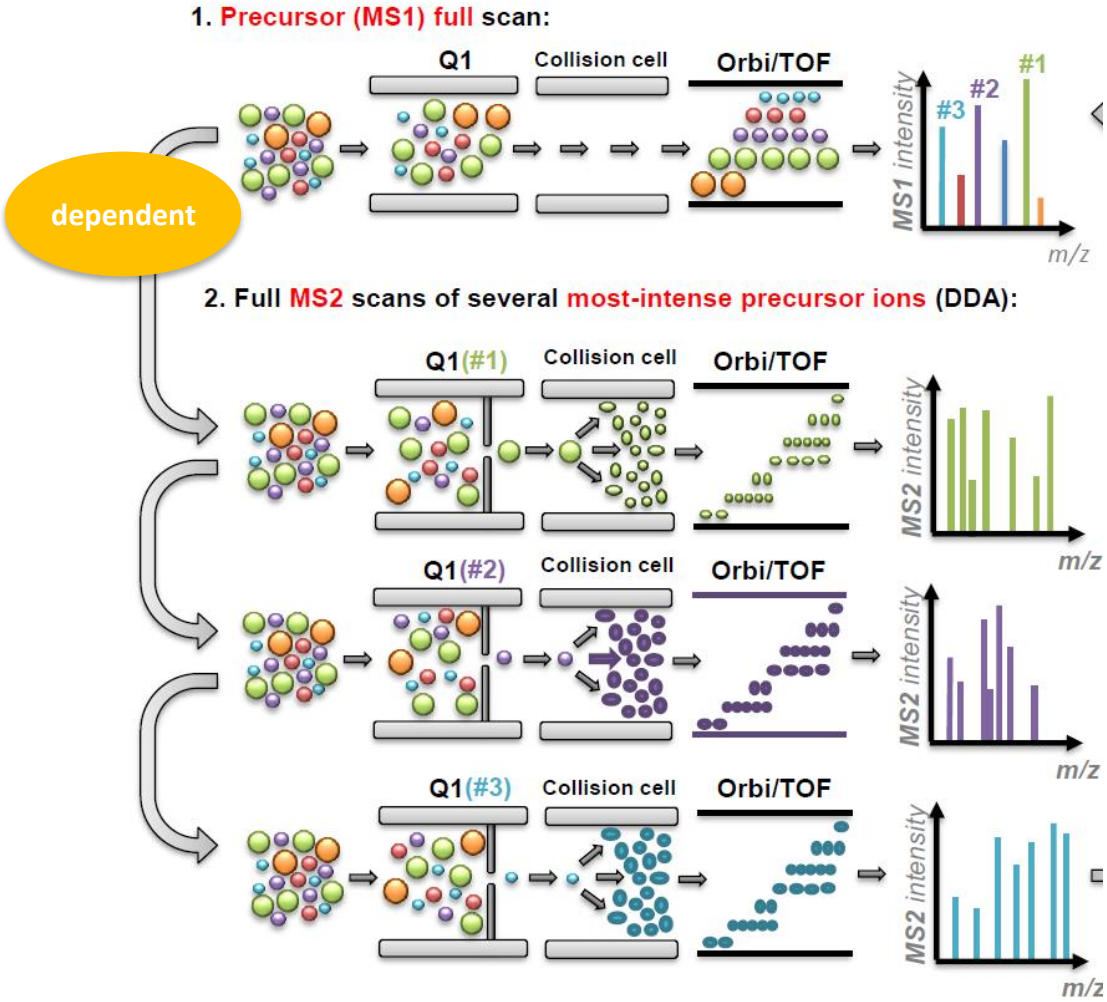
- Limitation of NMR sensitivity
  - Problematic for intracellular metabolomics
- Limited number of metabolites addressed

Proteomics

- **Place for improvement in proteins identification**

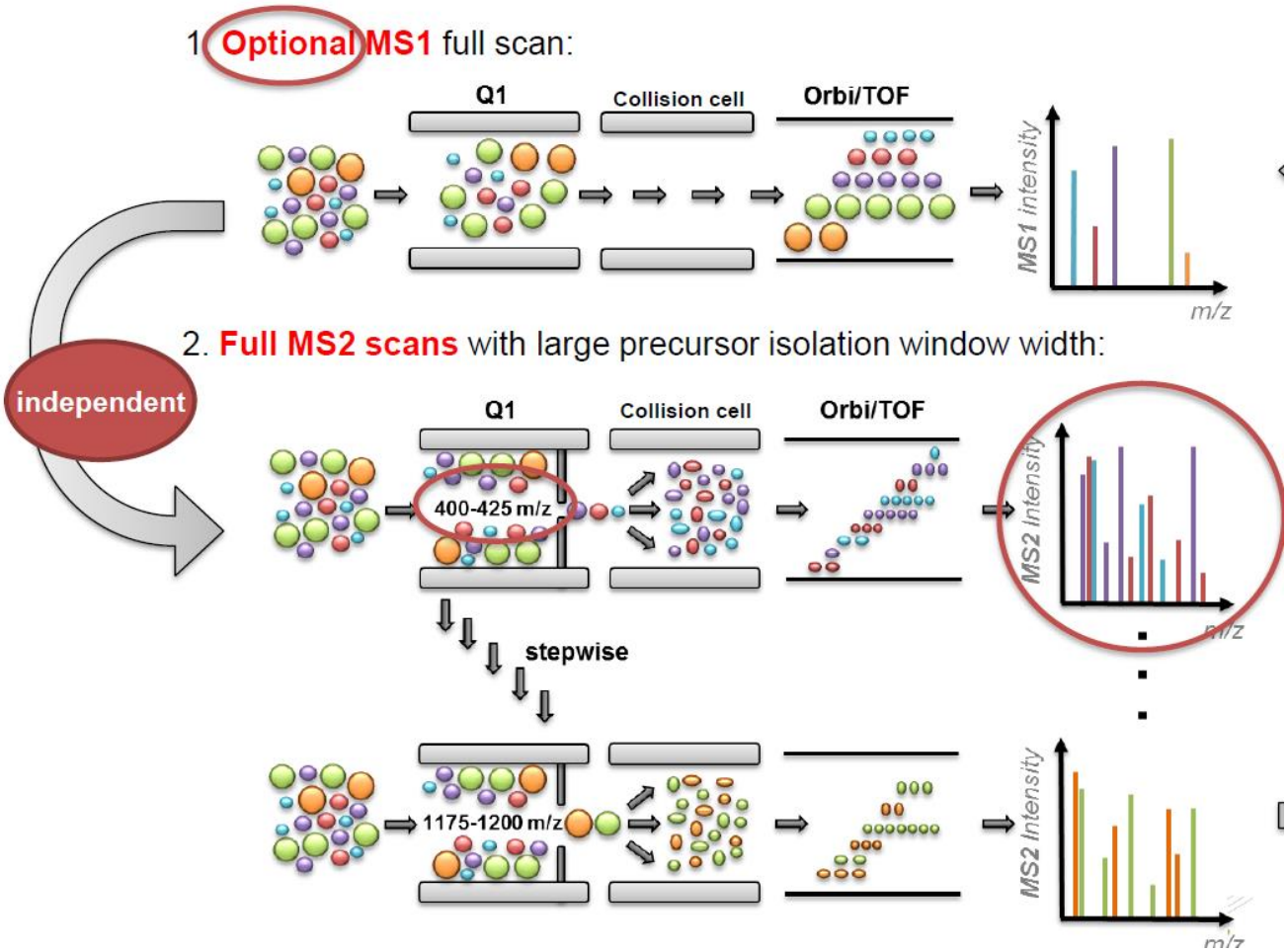
→ **DIA**

DDA: Data Dependent Acquisition



- Identification is done based on fragments ions matched onto *in silico* digested protein databases

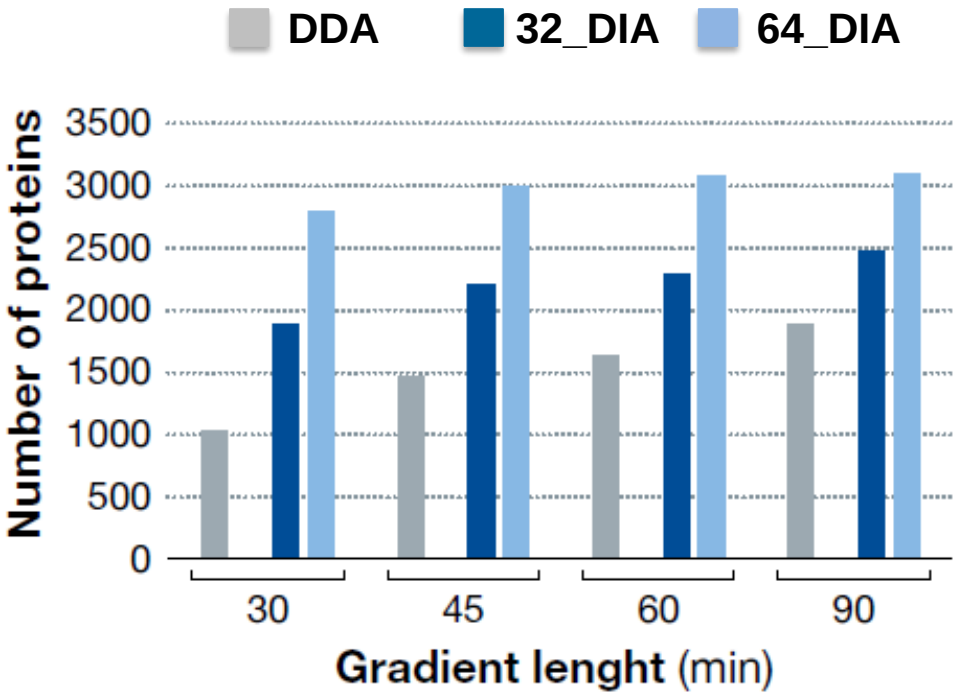
DIA: Data Independent Acquisition



- DIA is an acquisition method that acquires fragment ions (MS2) spectra in an unbiased fashion, without requiring the detection of peptide precursor ions in an MS1 survey scan (as in DDA).
- Identification is done using a spectral libraries, built up after fractionation (5-40 fractions) of pooled samples of the study (e.g. 20 fractions acquired in DDA mode)

DDA Vs DIA

- Comparison of results acquired either in DDA or DIA (32 & 64 windows) on HEK cell lysates (n=3), and impact of the LC gradient length (same instrument).



DIA outperforms DDA in terms of detectable peptides and associated proteins as well as measurement reproducibility (Bruderer et *al.*, 2015 & 2017, Kelstrup et *al.*, 2018).

DIA Workflow

Samples

- Pool samples
- Trypsin digestion
- **Extensive fractionation**

Data Dependent Analysis

BioTracs-Polaris

Generation of MSF search result file to serve as Spectral library

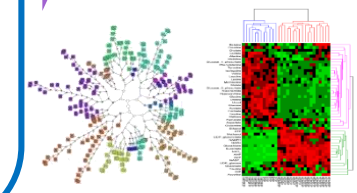
Data Independent Analysis

Spectronaut™

And

Skyline

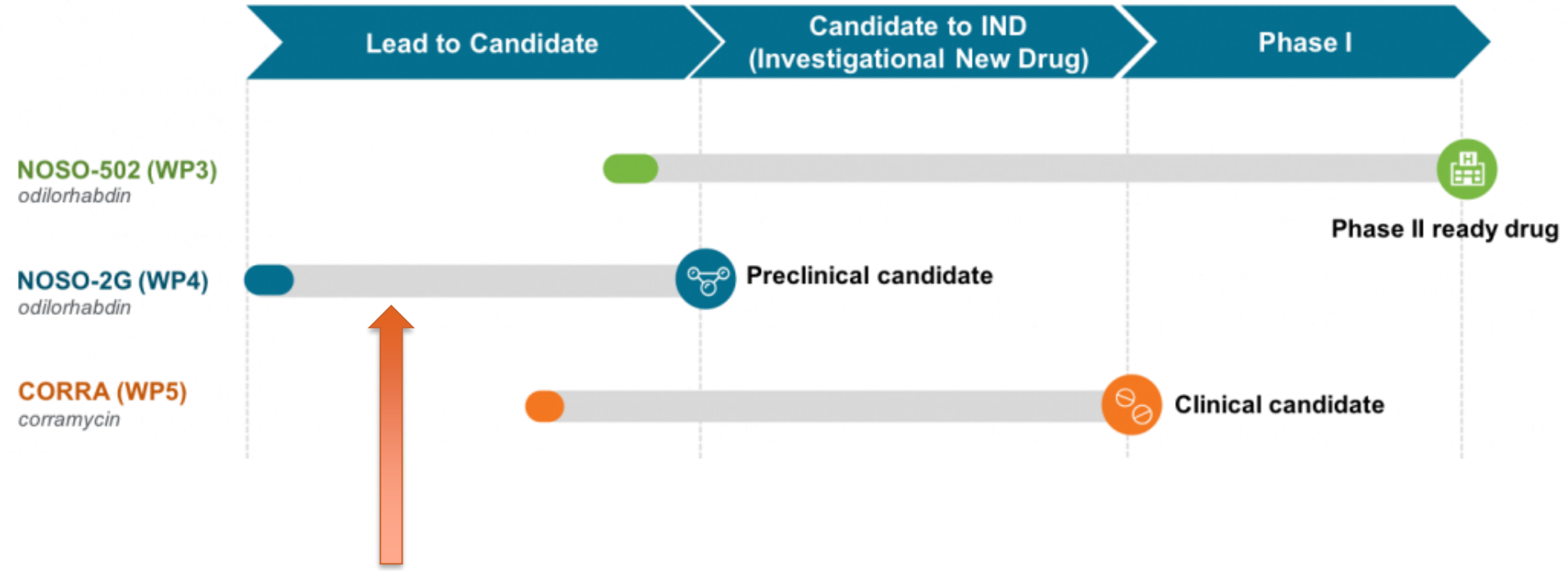
Pathway analysis



Fully operational

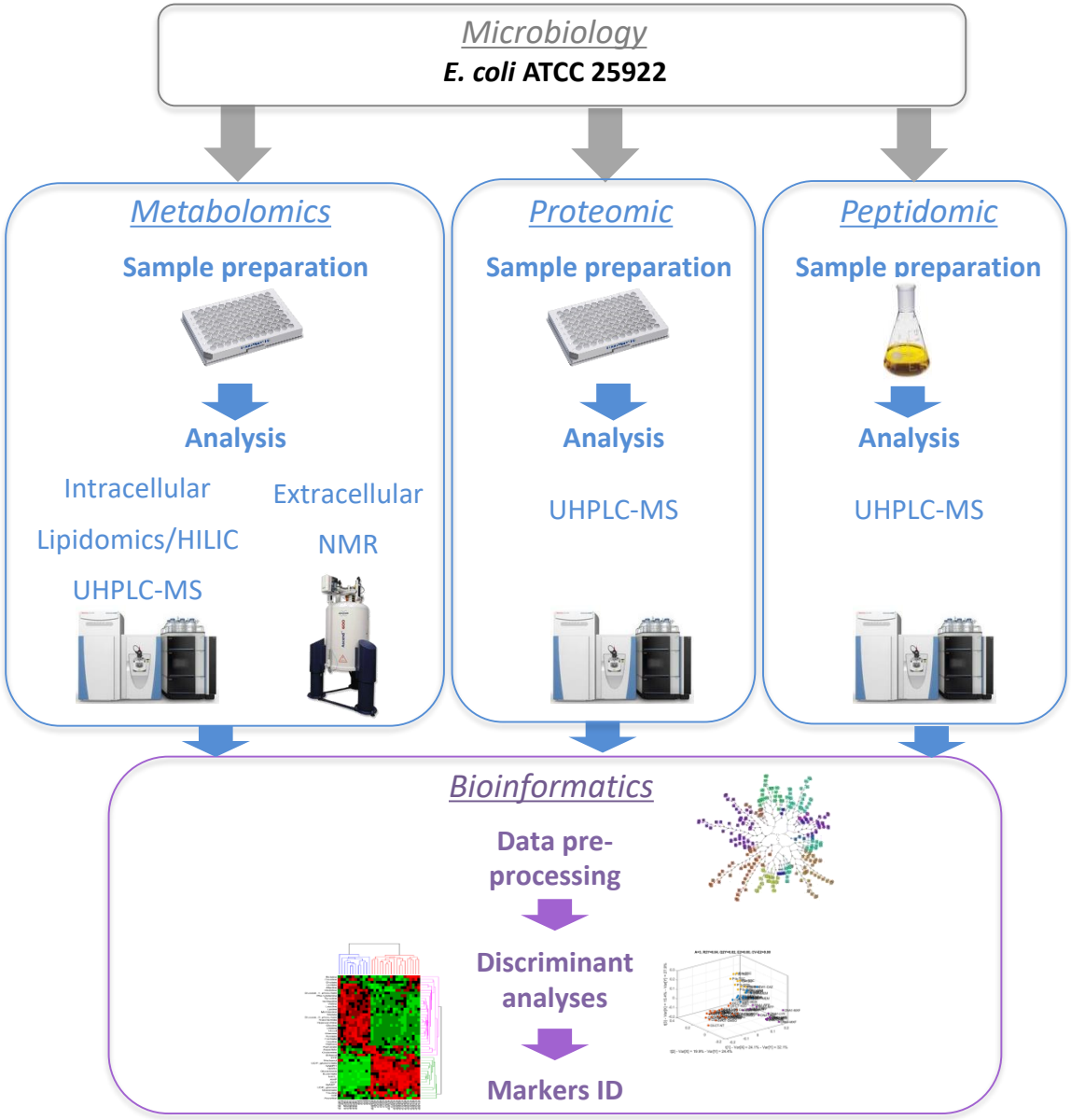
Application in an European Project – GNA NOW

Under the global umbrella of IMI's AMR Accelerator, the **Gram-Negative Antibacterials NOW** (GNA NOW) Consortium is a six-year project aimed at bringing together key European and private experts in **antibiotic discovery and development**.



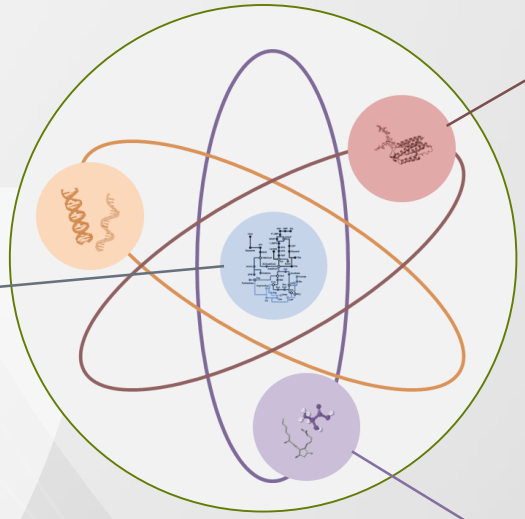
BIOASTER MICROBIOLOGY TECHNOLOGY INSTITUTE → **Metabolomics & Proteomics for Mode of Action deciphering**

Omics strategy



Project ongoing
Results soon !

Omics hub – Acknowledgments



Proteomics

Xavier Meniche

Mathieu Cyrille

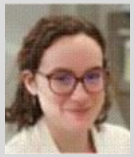
Bioinformatics

Joséphine AbiGhanem

Viet Dung Tran

Metabolomics

Andrei Bunescu



Magali Sarafian

Emeline Biliaut

AND THANK YOU FOR YOUR ATTENTION !