ASMS Metabolomics Short Course

Gary J. Patti
Michael and Tana Powell Professor
Senior Director, Center for Mass Spectrometry & Metabolic Tracing
Director, Clinical Research Core in Medicine
Director, Cancer Metabolomics Core
Director of Faculty Affairs, Chemistry
Dean's Fellow of Entrepreneurship and Advancement
Departments of Chemistry, Genetics, Medicine
Washington University in St. Louis
Chief Scientific Officer, Panome Bio
gjpattij@wustl.edu
pattilab.wustl.edu

LinkedIn

Gary Patti
Professor, Washington University | CSO, Panome Bio

Bluesky

Gary Patti
@gjpattij.bsky.social
metabophile (muh-tab-uh-fil'): a person obsessed with all things metabolism, metabolomics

ASMS Metabolomics Short Course

Equipment: ~20 mass spectrometers (Agilent, Bruker, SCEIX, Thermo, and Waters)
Untargeted, targeted, GC/MS, imaging, and isotope tracing analysis

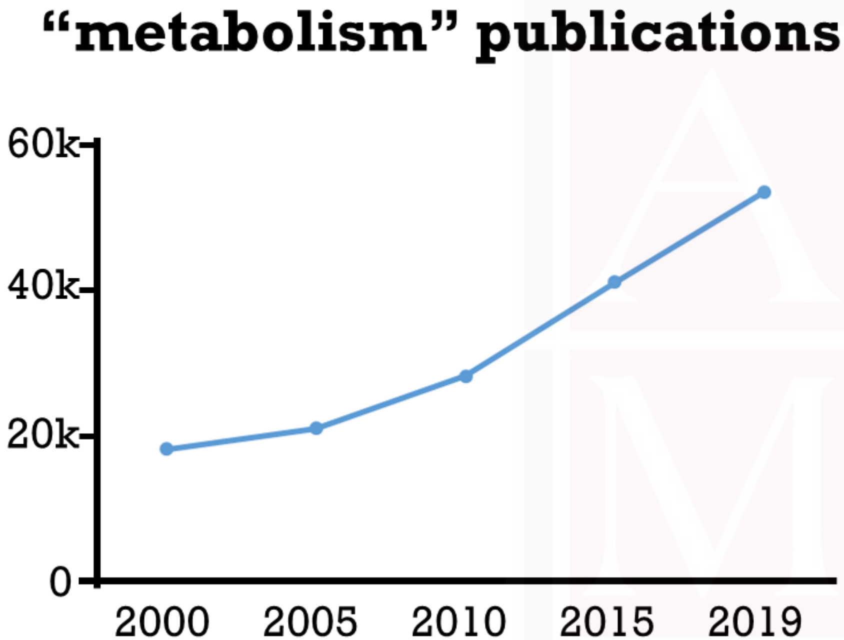
"As to methods, there may be a million and then some, but principles are few. The man who grasps principles can successfully select his own methods. The man who tries methods, ignoring principles, is sure to have trouble."

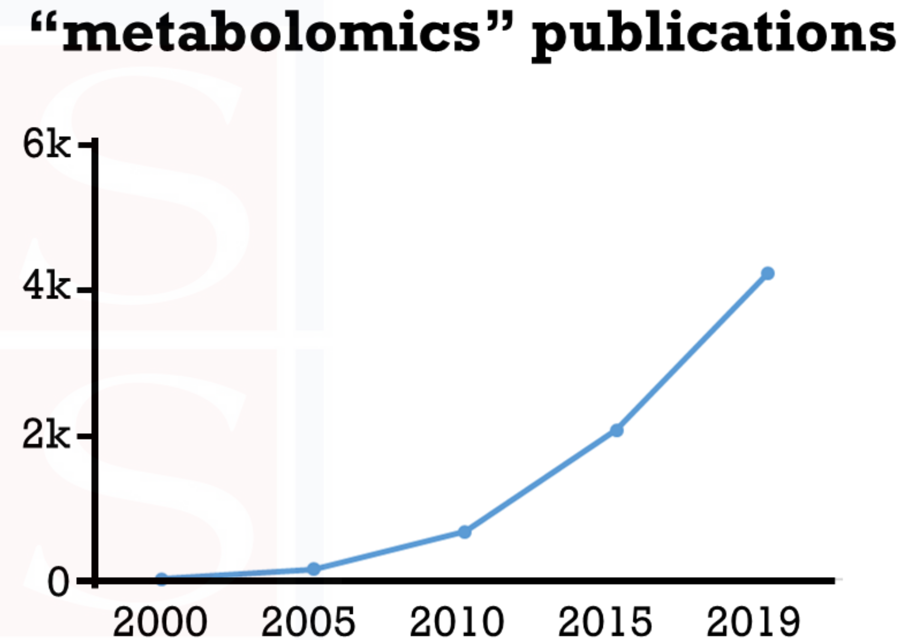
Emerson

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Overview

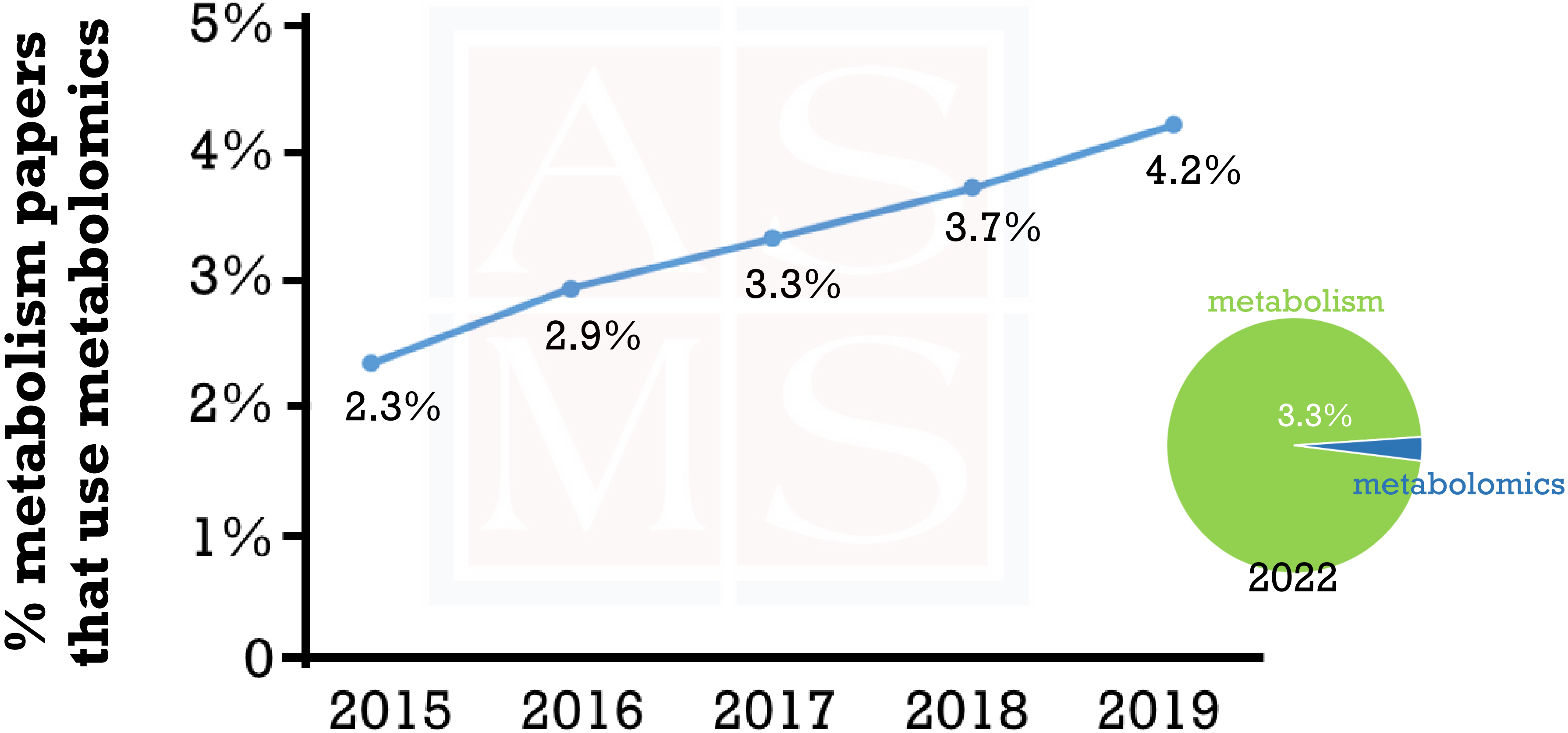
Metabolomics: chicken or egg?





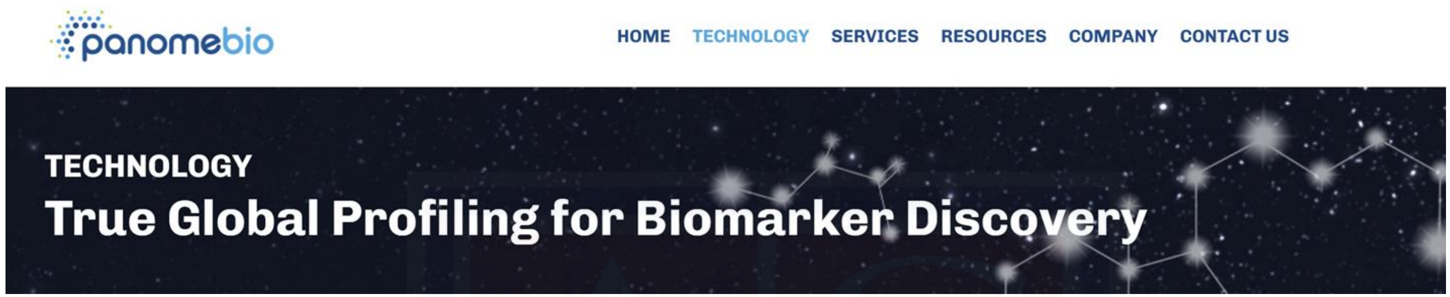
Web of Science, 1/2/2023

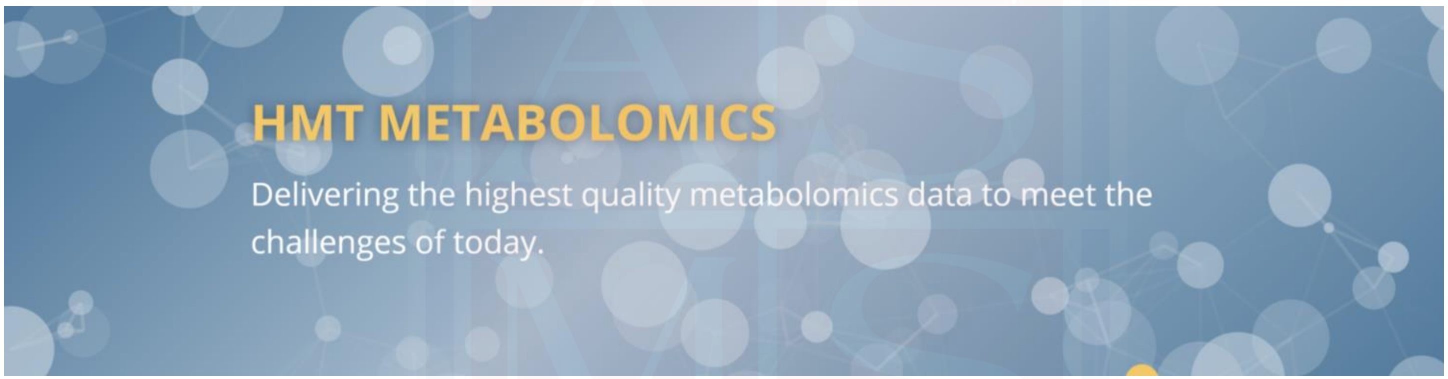
Metabolomics: room for growth



Web of Science, 1/2/2024

Metabolomics Resource Cores





Metabolomics Resource Cores

CONTACT NAME	INSTITUTION	SERVICE DESCRIPTION
O. Fiehn	Univ. of CA Davis	Comprehensive Metabolomics
R. Yost	Univ. of Florida	Integrated Metabolomics
R. Higashi	Univ. of Kentucky	Stable Isotope-Resolved Metabolomics
C. Burant	Univ. of Michigan	Comprehensive Metabolomics
K. S. Nair	Mayo Clinic	Metabolomics
S. Sumner	RTI	Comprehensive Metabolomics
G. Patti	WashU in Stl	Metabolomics and Isotope Tracing

Cores, cores, and more cores!

- Many/most institutions have a metabolomics core(s)
- Beware of their experience and credibility
- Just because a lab has a mass spectrometer, doesn’t mean that they can run a metabolomics core
- Expertise in proteomics does not equate to experience in metabolomics
- If you try one core unsuccessfully, might be good to try another: not all metabolomics cores are equal!

A good problem (mostly)

- **Surging interest in metabolism**

- Software solutions mostly available (>200 free) but require resources and training

- Widespread availability of technology has made accessible to most (with caveat that quality is issue)

- =>Lots of biologists/clinicians have untargeted metabolomics data, but cannot interpret it because:

  (i) Poor data quality, (ii) ID barrier, (iii) not versed in metabolism

- Sometimes "bad" data can be worse than no data

ID bottleneck: exciting opportunity or annoying barrier?

ID bottleneck: exciting opportunity or annoying barrier?

E. coli sample

ID bottleneck: exciting opportunity or annoying barrier?

E. coli sample

25,342 total signals

ID bottleneck: exciting opportunity or annoying barrier?

E. coli sample

25,342 total signals

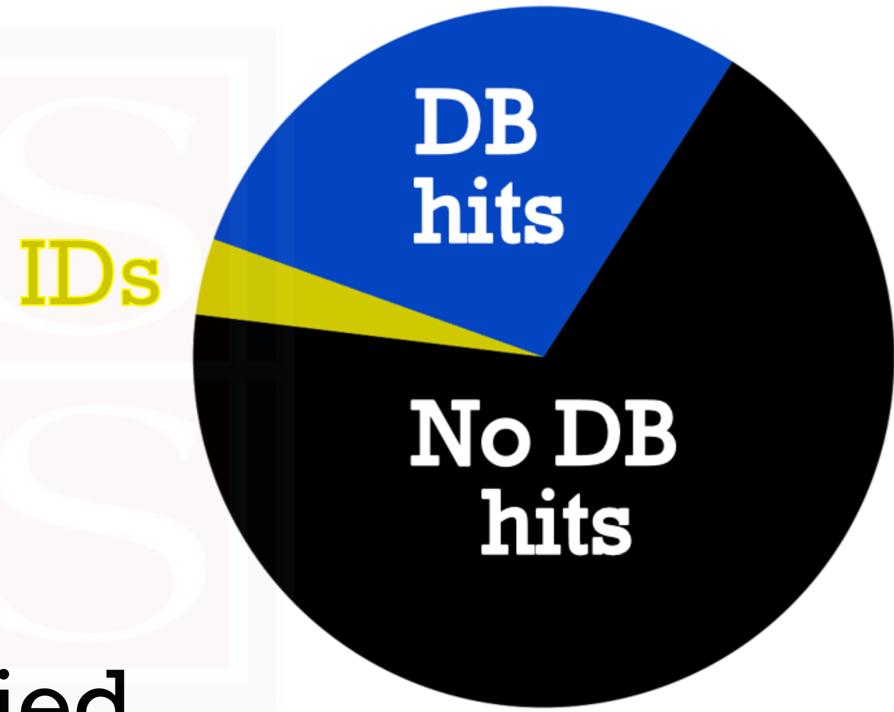
<1000 signals identified

ID bottleneck: exciting opportunity or annoying barrier?

E. coli sample

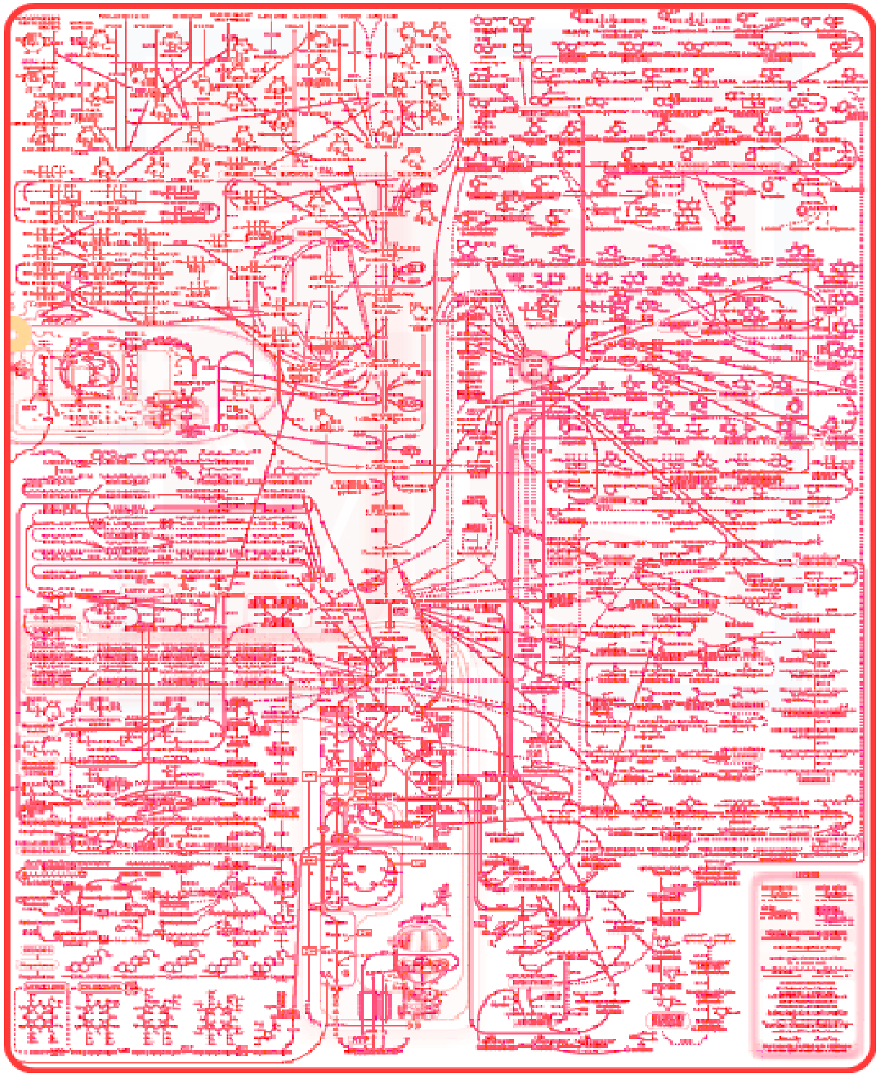
25,342 total signals

<1000 signals identified



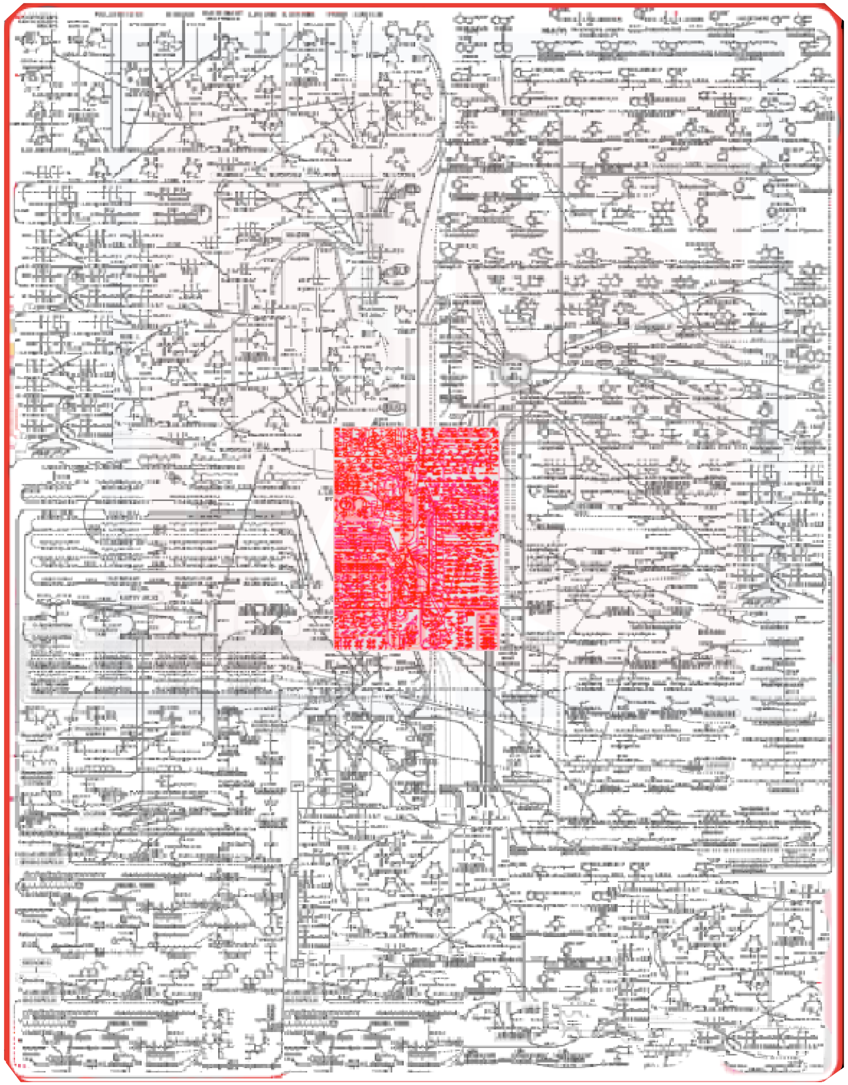
ID bottleneck: exciting opportunity or annoying barrier?

ID bottleneck: exciting opportunity or annoying barrier?

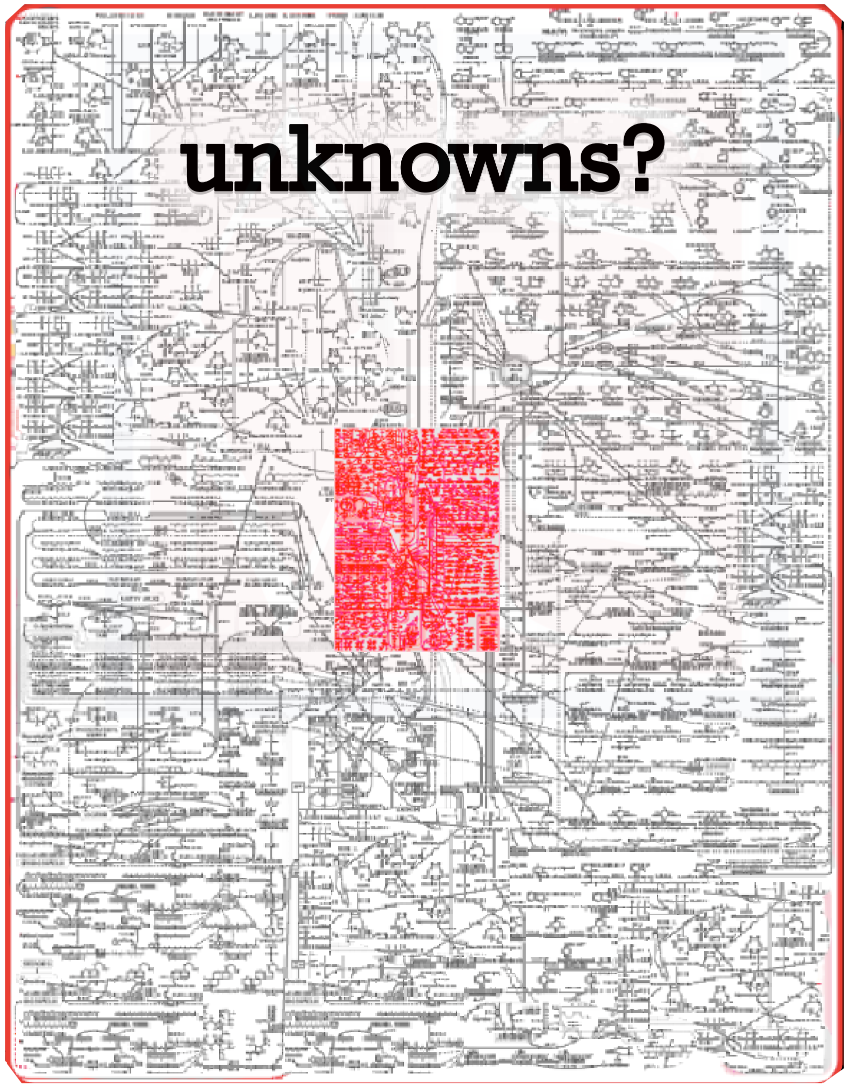


ID bottleneck: exciting opportunity or annoying barrier?

ID bottleneck: exciting opportunity or annoying barrier?



ID bottleneck: exciting opportunity or annoying barrier?



How big is the metabolome?

1800's: amino acids discovered

1897: Buchner: cellular fermentation

1904: Knoop theorizes ***β-oxidation***

1930's: Warburg: respiratory chain

1940: Meyerhof and Leloir: ***glycolysis***

1944: Lehninger demonstrates **β-oxidation**

1930-1940: Warburg, Lipmann: ***Pentose Phosphate Pathway***

1947: Cori's receive Nobel Prize for ***Cori Cycle***

1950's: Krebs describes ***Urea Cycle***

1953: Krebs receives Nobel Prize for ***TCA Cycle***

1961: Calvin receives Nobel Prize for ***Calvin Cycle***

How big is the metabolome?

1800's: amino acids discovered

1897: Buchner: cellular fermentation

1904: Knoop theorizes ***β-oxidation***

1930's: Warburg: respiratory chain

1940: Meyerhof and Leloir: ***glycolysis***

1944: Lehninger demonstrates **β-oxidation**

1930-1940: Warburg, Lipmann: ***Pentose Phosphate Pathway***

1947: Cori's receive Nobel Prize for ***Cori Cycle***

1950's: Krebs describes ***Urea Cycle***

1953: Krebs receives Nobel Prize for ***TCA Cycle***

1961: Calvin receives Nobel Prize for ***Calvin Cycle***

How big is the metabolome?

1800's: amino a

1897: Buchne

1904: Kno

1930's:

1940

19

athway

le

Cycle

1965, Chain et. al

Landmarks and Perspectives in Biochemical Research

*"The elucidation of the pathways of metabolism is one of the most important tasks of functional biochemistry. Very great progress has been made in this field, and **we are now familiar with** the essential steps of most of the important metabolic pathways."*

How big is the metabolome?

1800's: amino a

1897: Buchne

1904: Kno

1930's:

1940

19

athway

le

Cycle

1965, Chain et. al

Landmarks and Perspectives in Biochemical Research

"The elucidation of the pathways of metabolism is one of the most important tasks of functional biochemistry. Very great progress has been made in this field, and we are now familiar with the essential steps of most of the important metabolic pathways "

How big is the metabolome?

1800's: amino acids discovered

1897: Buchner: cellular fermentation

1904: Knoop theorizes ***β-oxidation***

1930's: Warburg: respiratory chain

1940: Meyerhof and Leloir: ***glycolysis***

1944: Lehninger demonstrates **β-oxidation**

1930-1940: Warburg, Lipmann: ***Pentose Phosphate Pathway***

1947: Cori's receive Nobel Prize for ***Cori Cycle***

1950's: Krebs describes ***Urea Cycle***

1953: Krebs receives Nobel Prize for **TCA *Cycle***

1961: Calvin receives Nobel Prize for ***Calvin Cycle***

How big is the metabolome?

1800's: amino

1897: Buchn

1904: Kn

1930's:

194

19

2012, Nature Cell Biology Reviews

Metabolomics: the apogee of the omics triology

"our understanding of metabolism is evolving much like our notion of physics evolved in the early twentieth century with the emergence of experimental results such as the photoelectric effect, which could not be explained by Newtonian laws. Ultimately, the ideas that emerged from this disparity resulted in a new set of principles for understanding physical phenomena known as quantum mechanics."

Pathway

cle

n Cycle

Best technology for untargeted metabolomics?

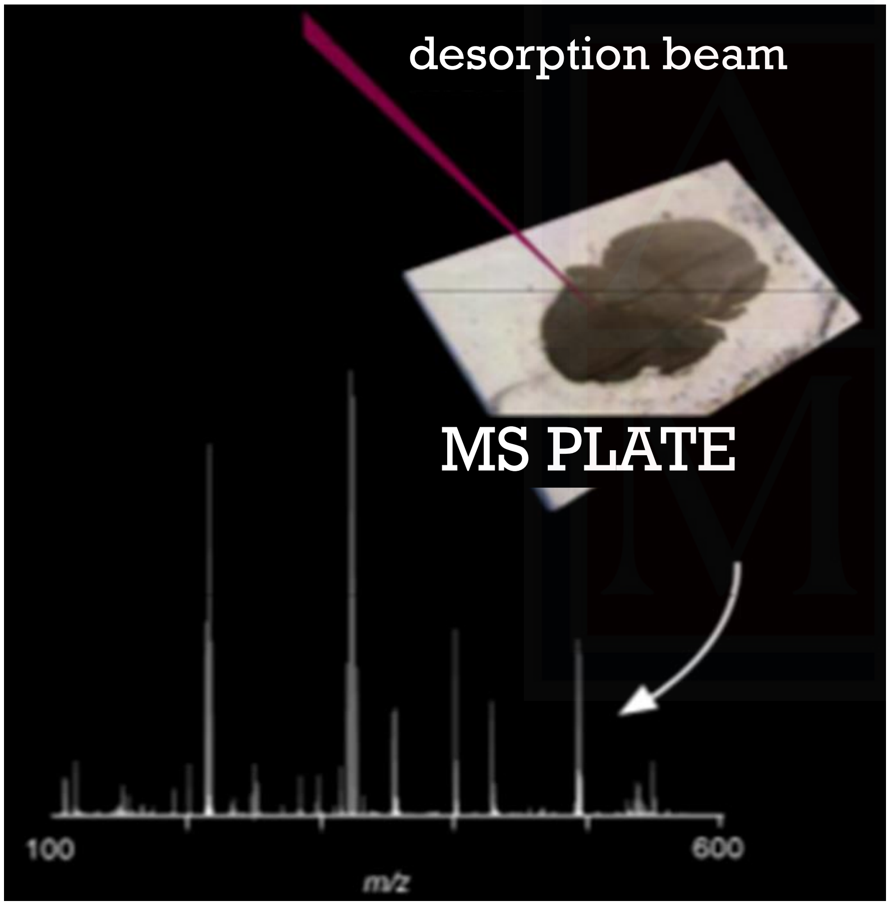
NMR *vs.* *GC/MS* *vs.* *LC/MS*

Best technology for untargeted metabolomics?

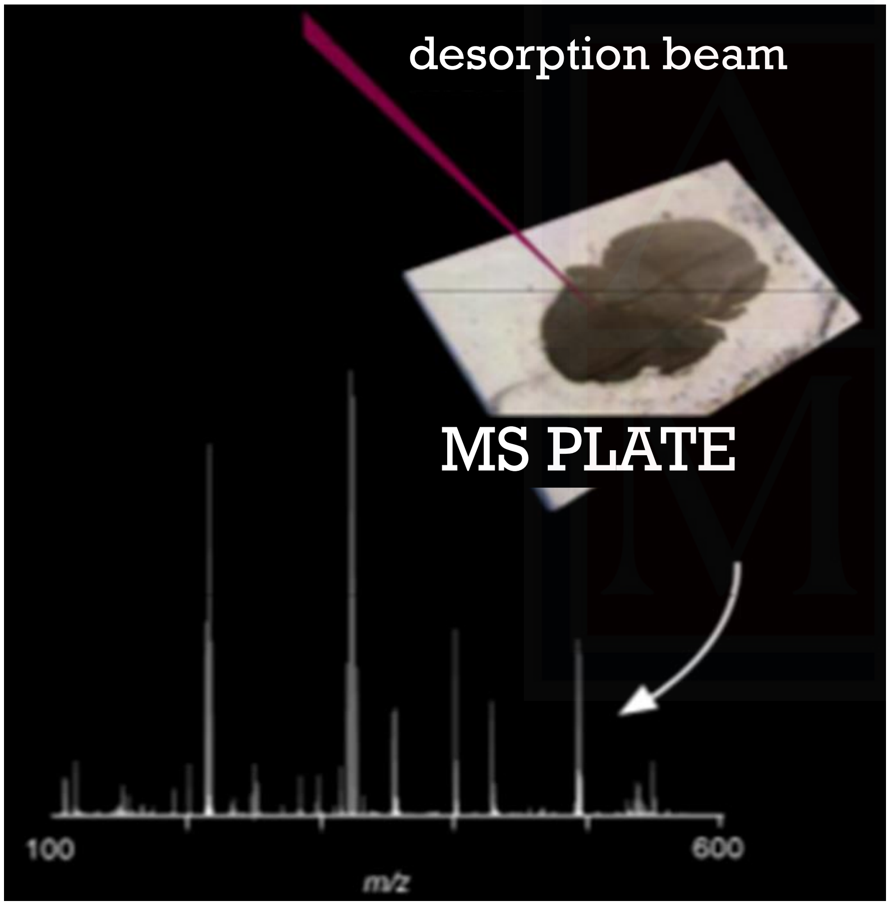
NMR vs. GC/MS vs. LC/MS

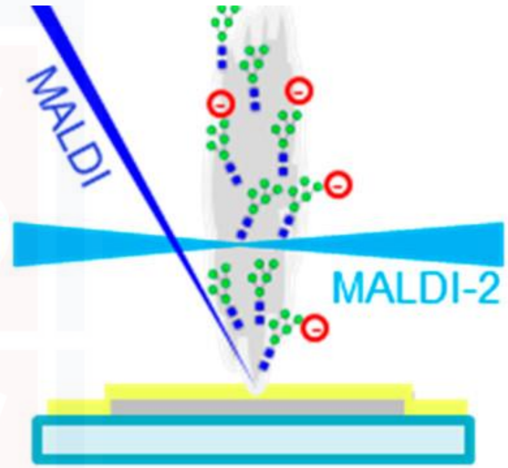
- **Each has unique strengths**
- **Peak numbers is classic argument, but it's fundamentally flawed**
- **LC/MS most comprehensive→ doesn't mean it's the best for <u>your</u> experiment**
- **"Peaks" is a bad metric…much more later**

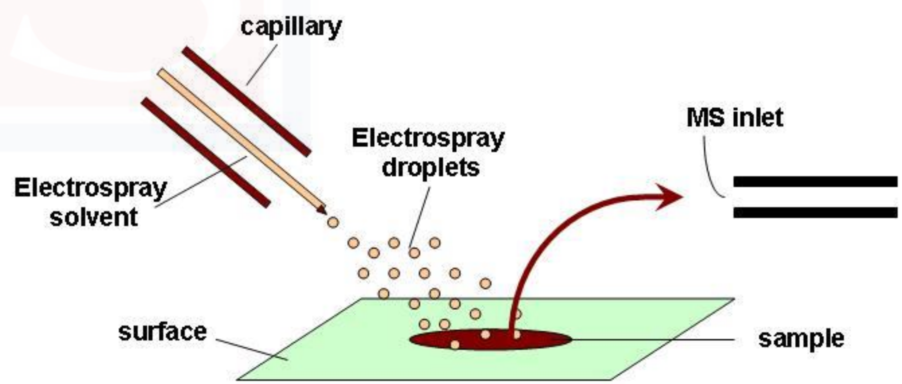
Best technology for untargeted metabolomics?



Best technology for untargeted metabolomics?



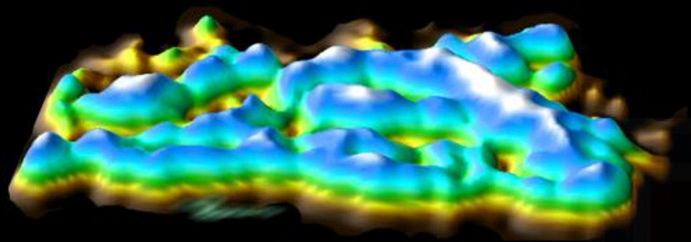




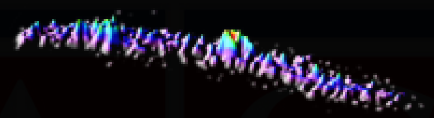
Best technology for untargeted metabolomics?

- **Imaging generally results in detection of many fewer metabolites than other technologies**
- **MALDI (and its variations), nanostructure initiator mass spectrometry (NIMS), desorption electrospray ionization (DESI)**
- **Each has advantages and disadvantages**
  - **Matrix interference? Commercialized? Spatial resolution?**

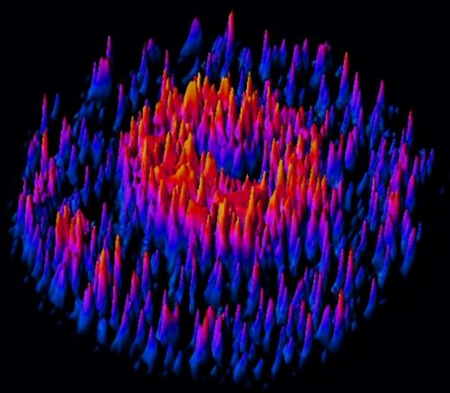
Brain



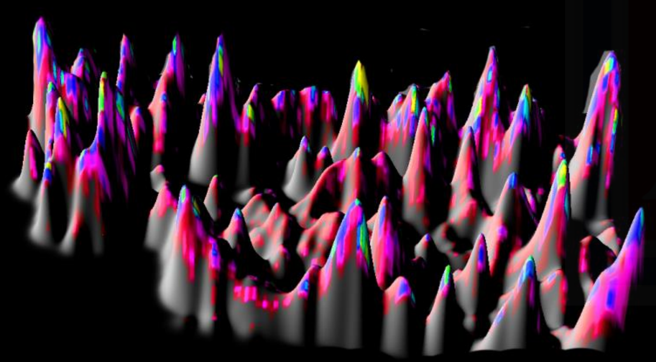
C. elegans



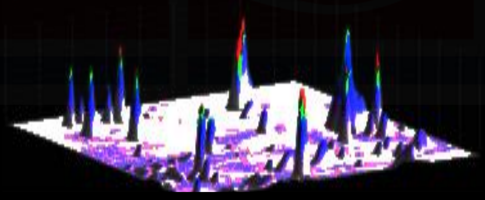
Plants



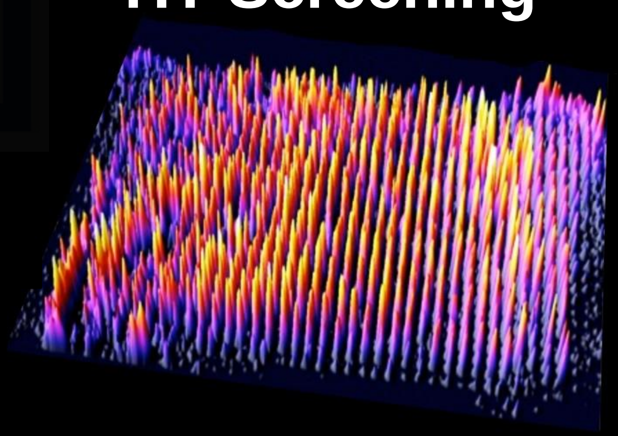
Breast Cancer



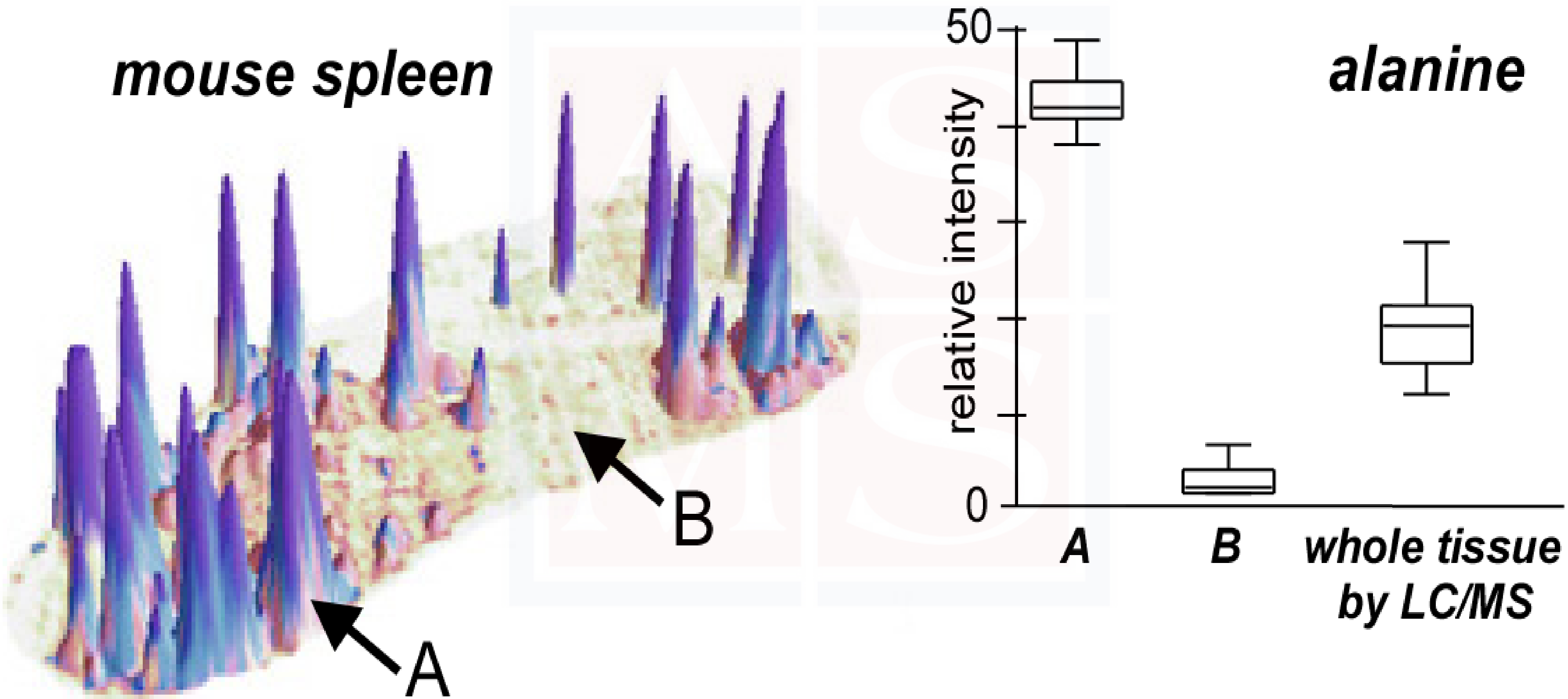
Single
Cancer Cells



HT Screening

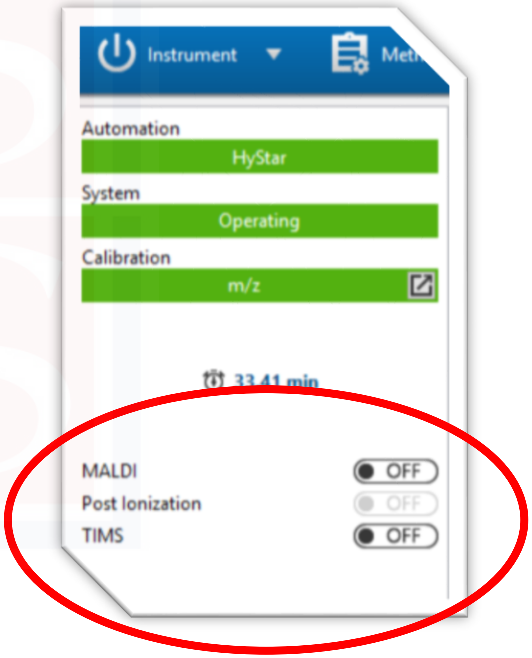


The "Averaging Effect"

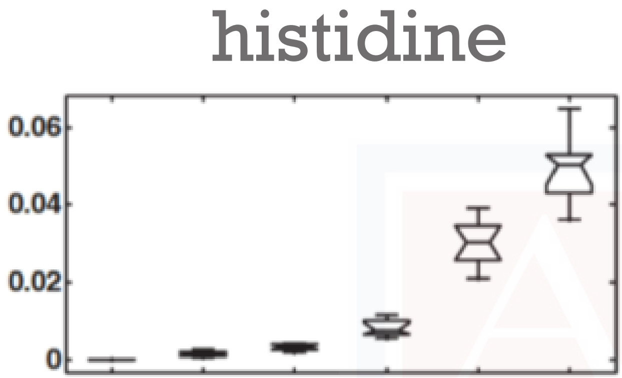


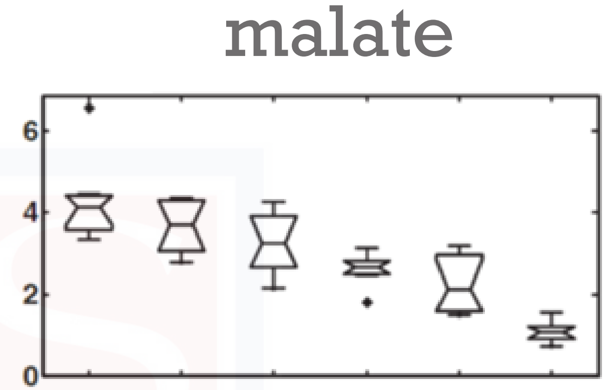
Best technology for untargeted metabolomics?

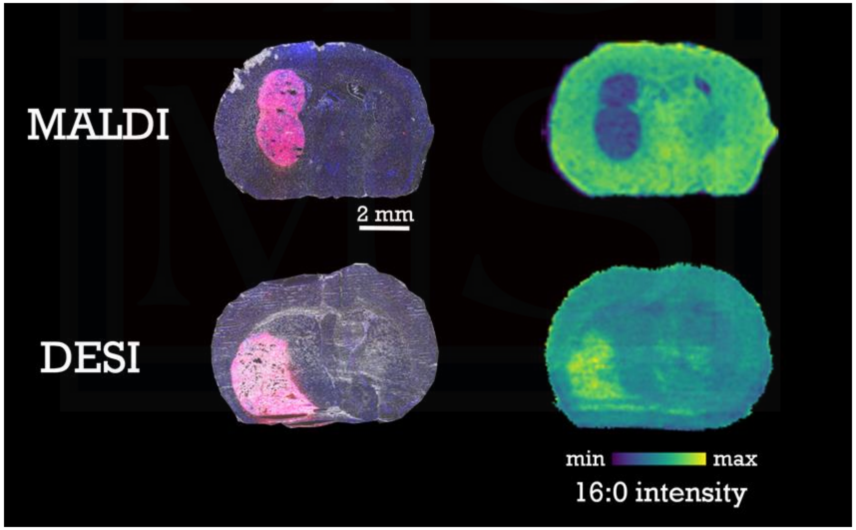
simple toggle switch:
on timsTOF flex:



Beware of matrix effects



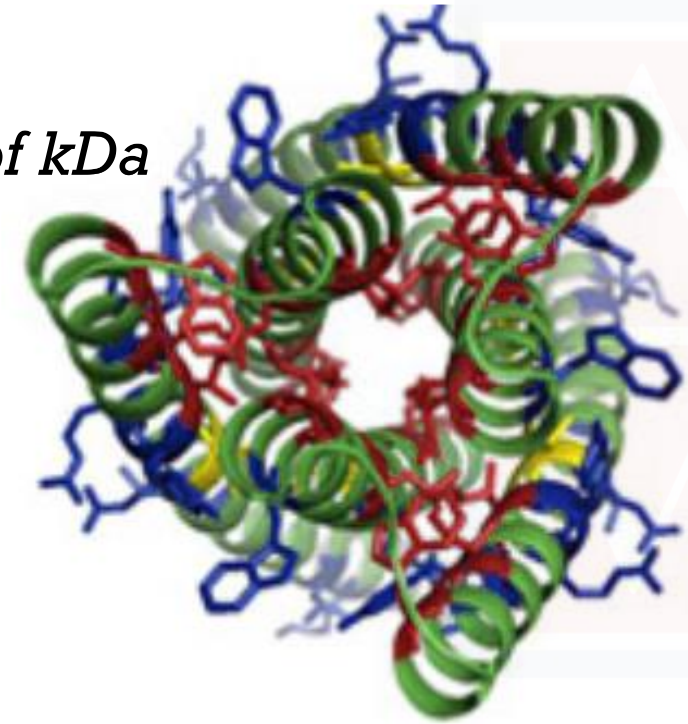




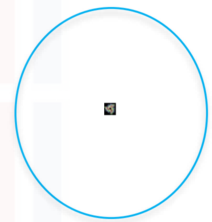
Goodacre et al, Rapid Comm in MS 2007
Schwaiger-Haber et al, Nat Comm 2023

Metabolomics is fundamentally different from proteomics

10s of kDa



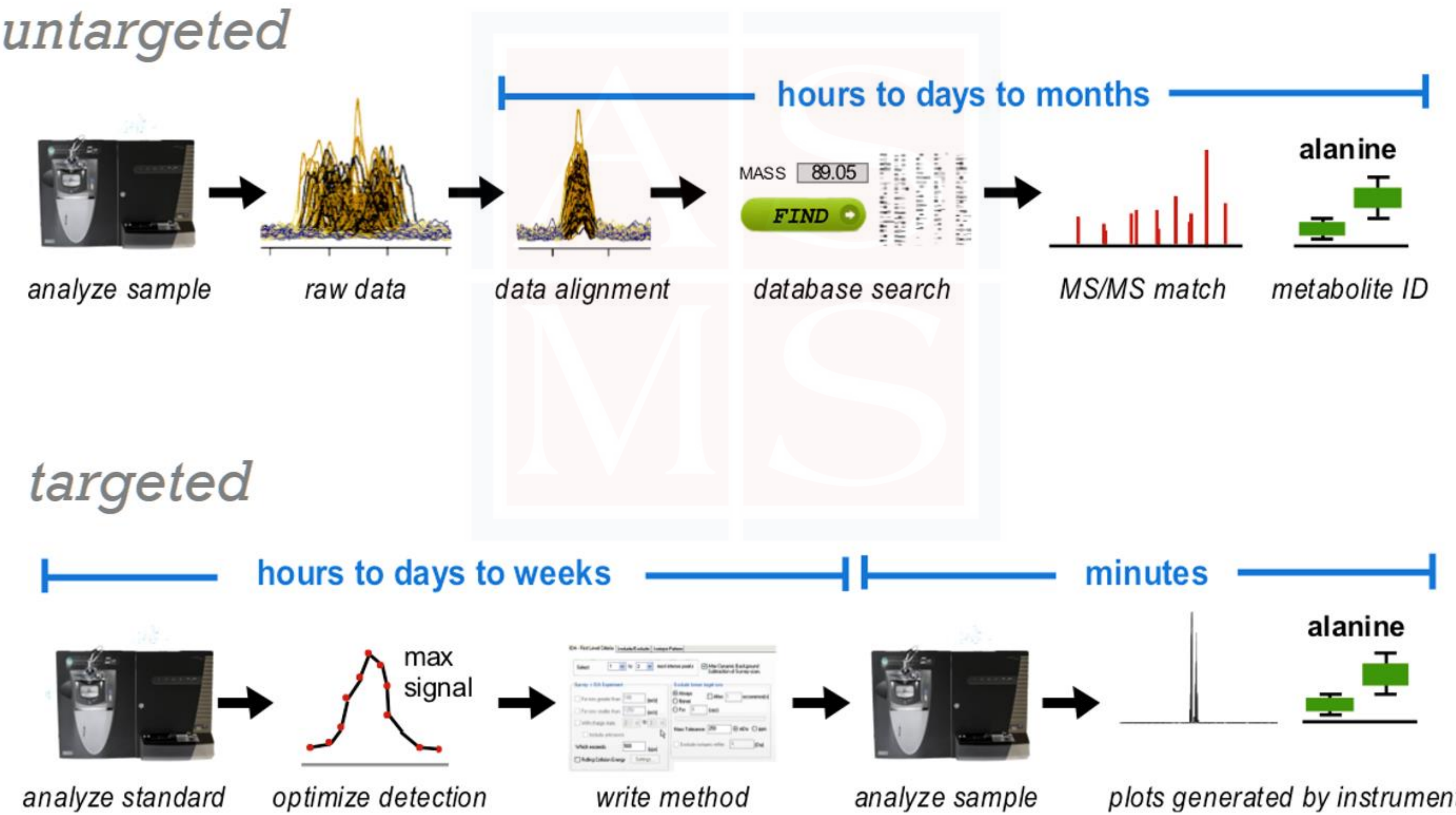
1 kDa



20k proteins,
digest,
1 M peptides

diff chromatography, workflows, IDs, etc.

untargeted vs. targeted metabolomics



LC/MS untargeted workflow (1)

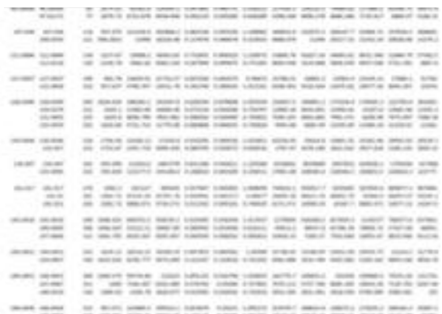
LC/MS untargeted workflow (1)

metabolite extraction → LC/MS → data processing → features

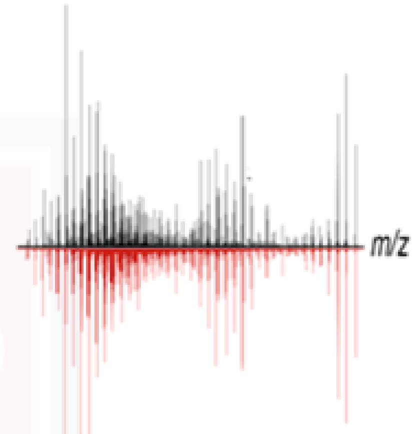
obtaining data relatively routine

LC/MS untargeted workflow (2)

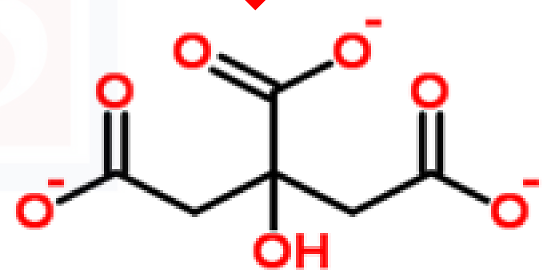
LC/MS untargeted workflow (2)



database



MS2



identification

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Objectives & Exp. Design

Experimental Objectives?

Experimental Objectives?

1.) Biomarkers

2.) Unknowns

3.) Disease pathogenesis

4.) Metabolic regulation

Experimental Objectives?

1.) Biomarkers

-most common application

2.) Unknowns

-low efficiency

3.) Disease pathogenesis

-requires rich understanding of metabolism

4.) Metabolic regulation

-isotopes often needed

Experimental Objectives?

We are exposed to millions of diff. chemicals over our lifetime — what impact do they have on human health?

Experimental Objectives?

Active Ingredient	*Purpose*
Pyrithione Zinc 1.0%	Anti-dandruff

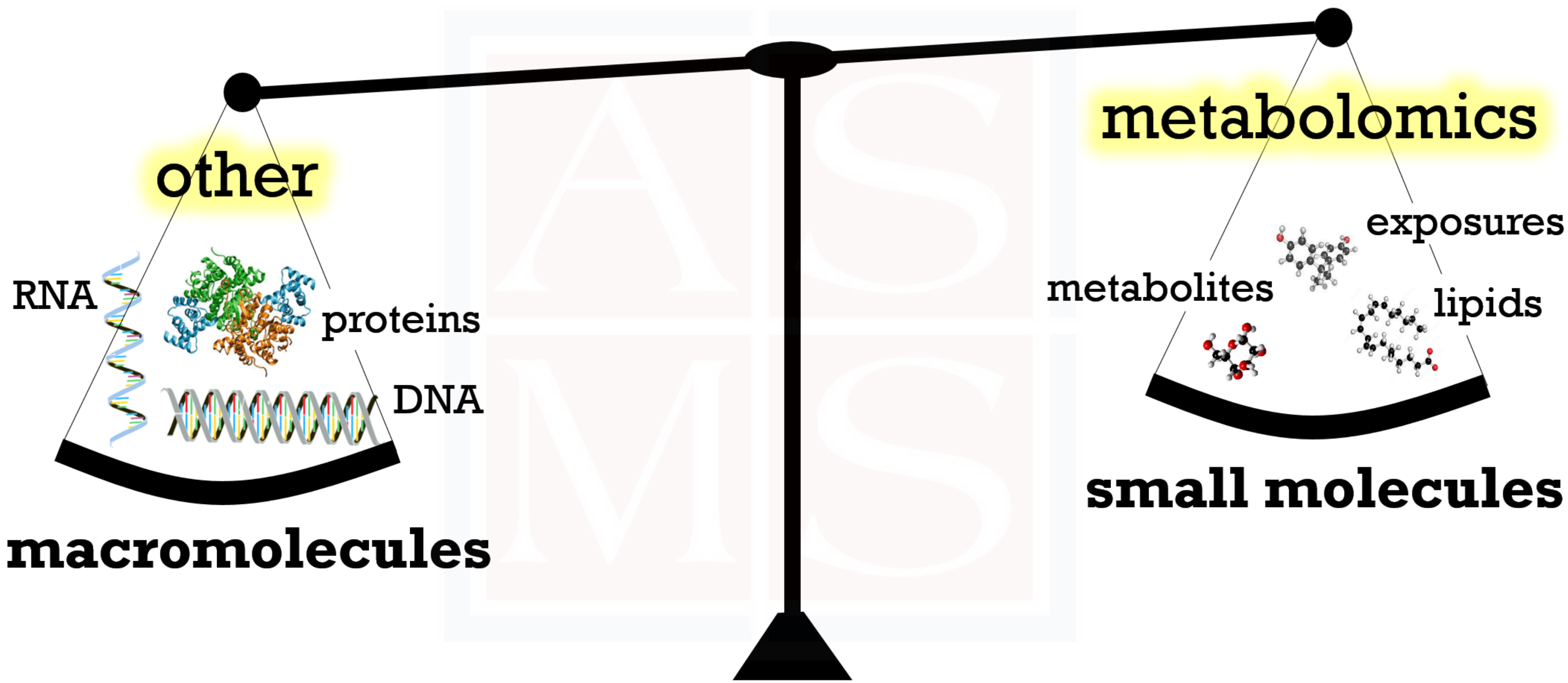
Inactive Ingredients

Aqua (Water, Eau), Sodium Cocoyl Isethionate, Sodium Lauroamphoacetate, Acrylates Copolymer, Ethylhexyl Olivate, Stearyl Alcohol, Carthamus Tinctorius (Safflower) Seed Oil, Allium Sativum (Garlic) Bulb Extract, Pyrus Malus (Apple) Fruit Extract, Camellia Sinensis Leaf Extract, Panthenol, Citrus Nobilis (Mandarin Orange) Peel Oil, Foeniculum Vulgare (Fennel) Oil, Lavandula Angustifolia (Lavender) Oil, Mentha Citrata Oil, Orbignya Oleifera Seed Oil, Rosmarinus Officinalis (Rosemary) Leaf Oil, Salvia Sclarea (Clary) Oil, Vanillin, Cocos Nucifera (Coconut) Oil, Laminaria Saccharina Extract, Squalane, Caprylyl Glycol, Yeast Ferment Extract, Guar Hydroxypropyltrimonium Chloride, Glycerin, Sodium Isethionate, Lactobacillus Ferment, Octenidine HCl, Sodium Polynaphthalenesulfonate, Xanthan Gum, Hexamidine Diisethionate, Ethylhexylglycerin, Coconut Alcohol, Raphanus Sativus (Radish) Root Extract, Arginine, Glycolic Acid, Polyquaternium-73, Pentylene Glycol, Propylene Glycol, Sodium Chloride, Potassium Sorbate, Sodium Benzoate, Citric Acid, Linalool

Experimental Objectives?

Strategy: measure the exposome from large cohorts and then correlate with health state

Experimental Objectives?



Experimental Objectives?

1.) Biomarkers

-most common application

2.) Unknowns

-low efficiency

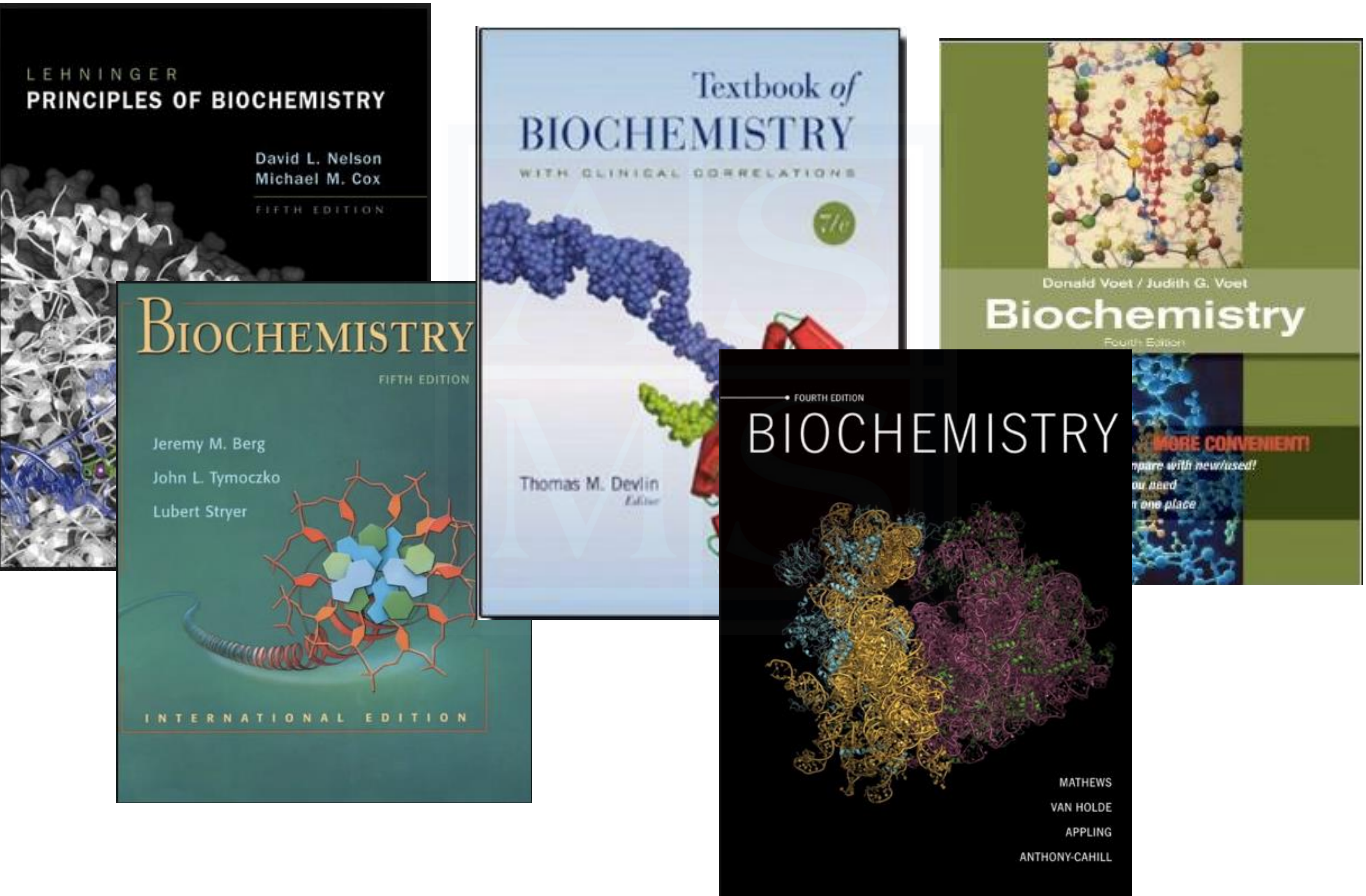
3.) Disease pathogenesis

-requires rich understanding of metabolism

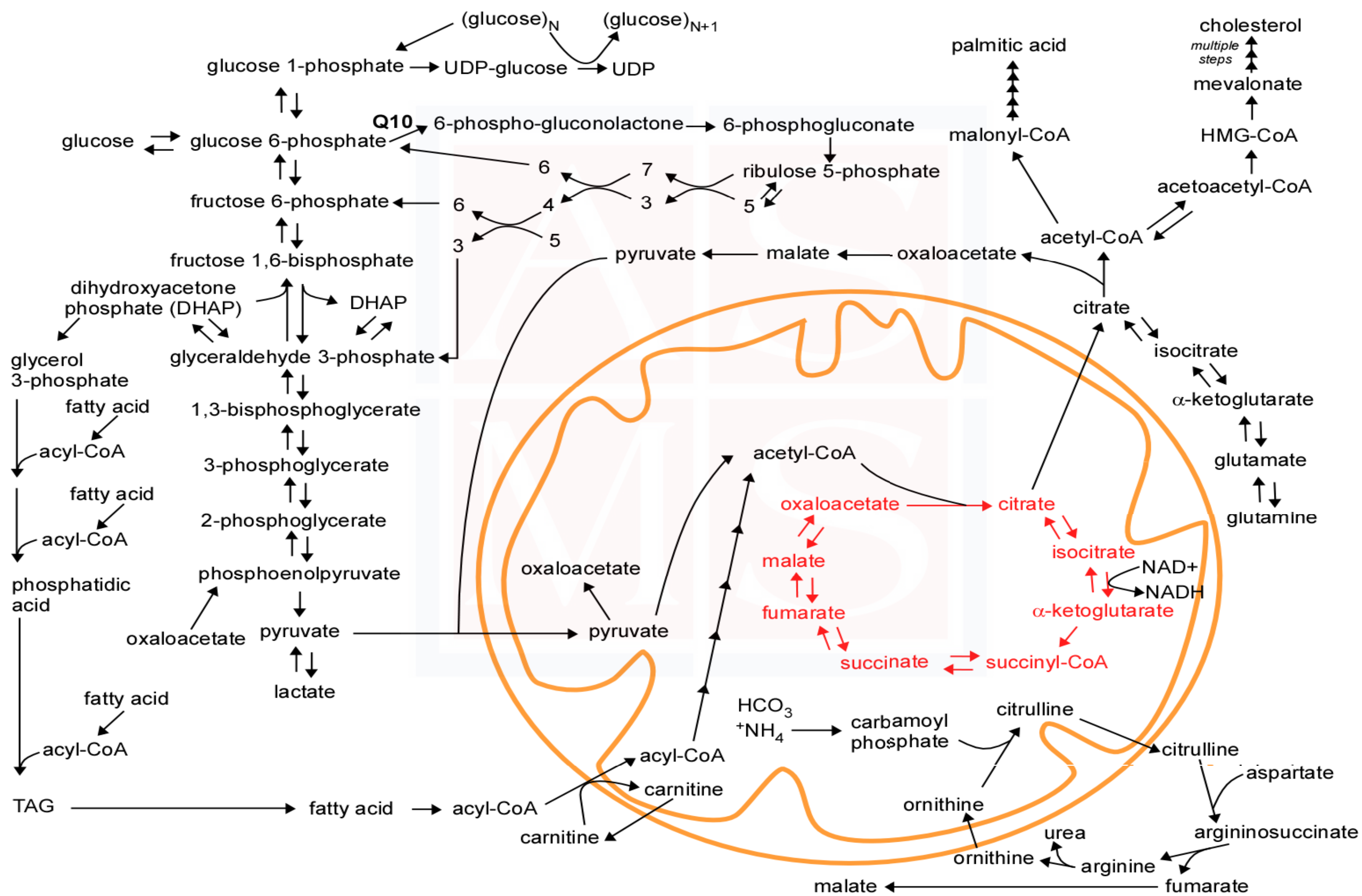
4.) Metabolic regulation

-isotopes often needed

Not covered in short course



Insights into disease pathogenesis



Untargeted

Typically QTOF or Orbi based

Informatic heavy

Extraction/chromatography challenging

Work on “back end”

Provides global info

Targeted

Historically QqQ based

Generally easier
-little/no informatics

Work on “front end”

Provides less info
(but often sufficient)

Untargeted

Typically QTOF or Orbi based

Informatic heavy

Extraction/chromatography challenging

Work on “back end”

Provides global info

Targeted **HYBRID** **Untargeted**

Historically QqQ based

Generally easier
-little/no informatics

Work on "front end"

Provides less info
(but often sufficient)

Typically QTOF or Orbi based

Informatic heavy

Extraction/chromatography challenging

Work on "back end"

Provides global info

Targeted

HYBRID

Untargeted

Historically QqQ based

Generally easier
-little/no informatics

Work on “front end”

Provides less info
(but often sufficient)

Typically QTOF or Orbi based

Informatic heavy

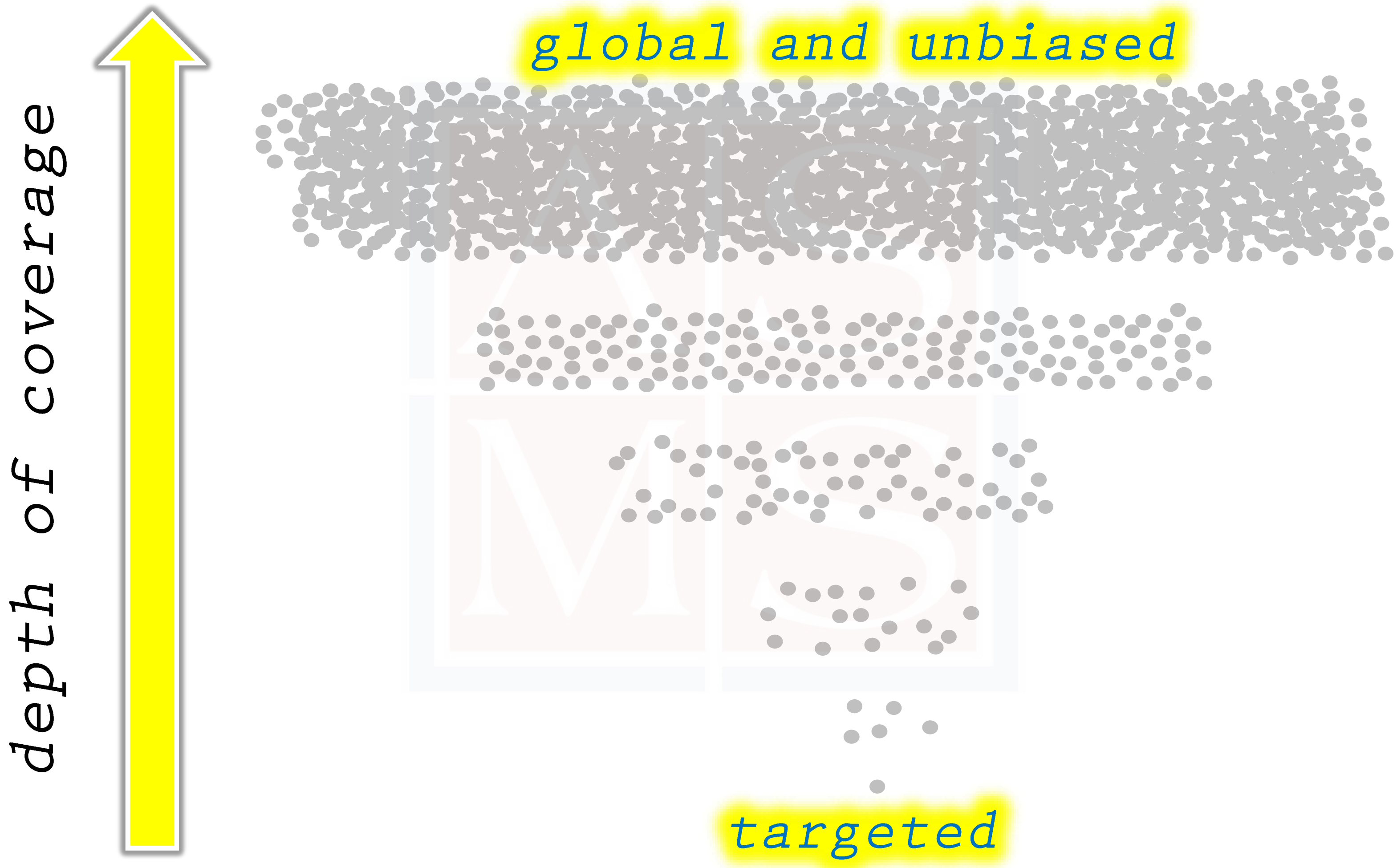
Extraction/chromatography challenging

Work on “back end”

Provides global info

focus of short course

Untargeted empowers discovery



Untargeted empowers discovery

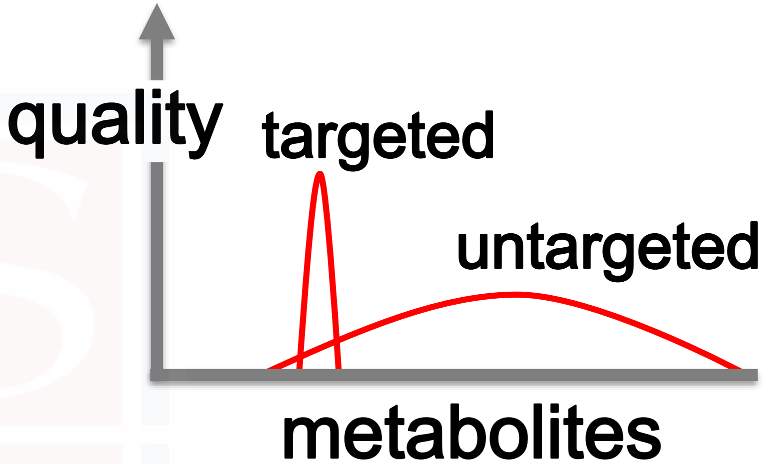
"the streetlight effect"

Untargeted does not mean unplanned

Untargeted does not mean unplanned

- Easy to generate data,
  hard to interpret.

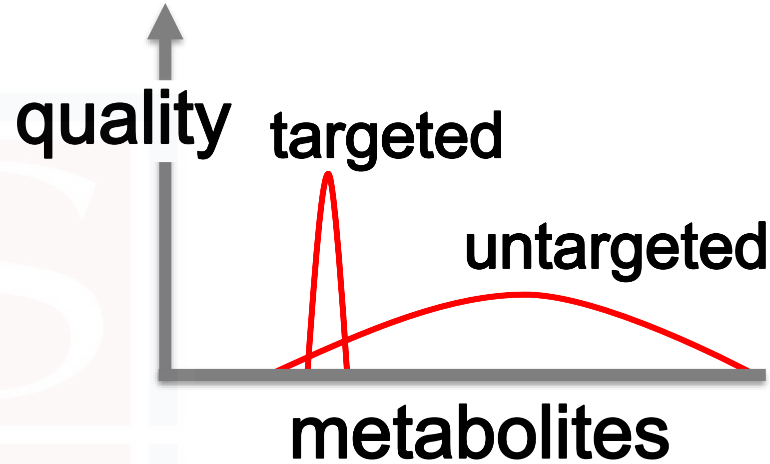
Untargeted does not mean unplanned



- Easy to generate data, hard to interpret.

- No experiment is comprehensive.

Untargeted does not mean unplanned



- Easy to generate data, hard to interpret.
- No experiment is comprehensive.
- Do you have a hypothesis that can be tested with a different experiment?

Untargeted metabolomics is challenging

Untargeted metabolomics is challenging

- Try to talk yourself out of it
- If you can do a different exp, it's probably better
- If no other design is possible, plan your exp very carefully

Experimental Design: Considerations

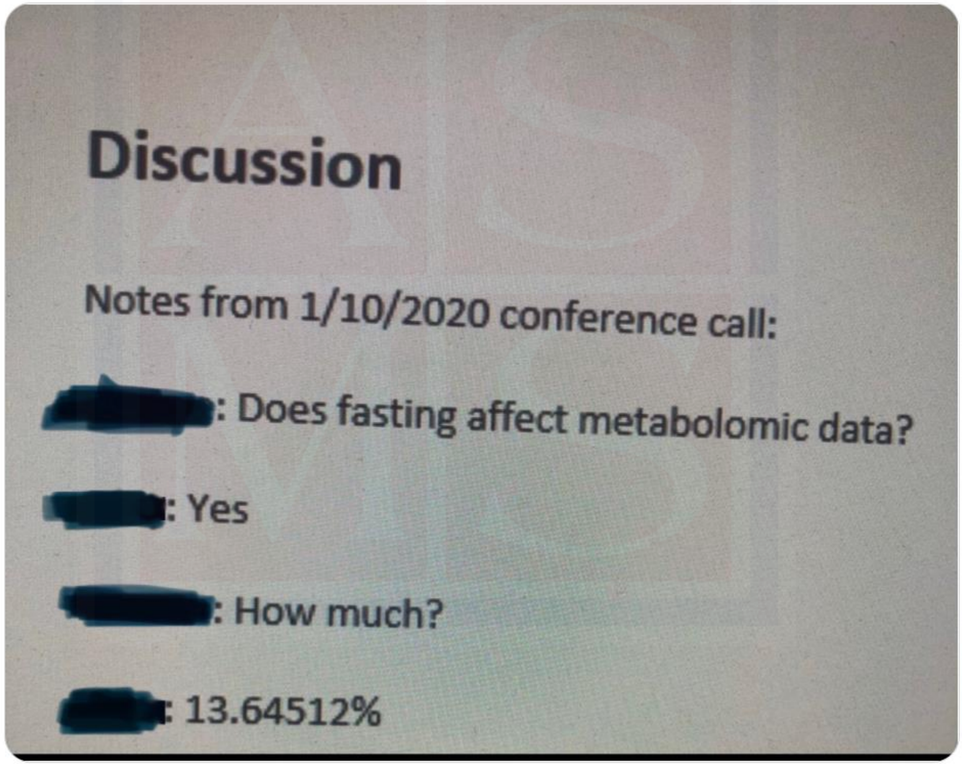
1. What should you compare?

- healthy vs disease
  - biomarkers, disease mech, therapeutic targets
- on drug vs. off drug
  - drug mode of action, drug metabolism
- wildtype vs knockout
  - metabolic effects of protein

Experimental Design: Considerations

Gary Patti
@gjpattij

When the person taking minutes for a conference call isn't a scientist and misses the dry humor.



10:29 AM · Jan 11, 2020 · Twitter Web App

View Tweet activity

Experimental Design: Considerations

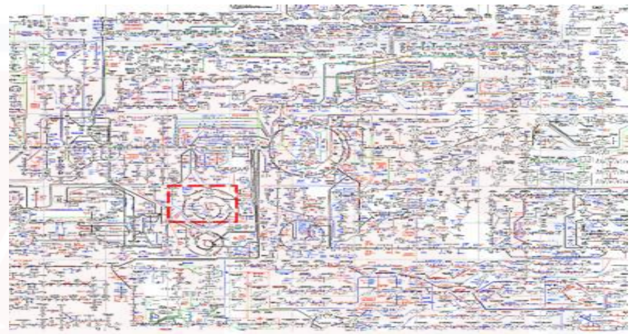
2. What type of sample?

sample type	pro	con
cells in culture	well controlled, high throughput, cost effective, one cell type	physiological relevance
plants, animals	more bio relevant than cell culture	multicellularity, less cost effective
people*	physiological relevance	large variability, cost, healthy samples can be challenging to obtain, IRB paperwork, hard to control variables (environment, diet, medications, exercise, stress, etc)

Experimental Design: Considerations

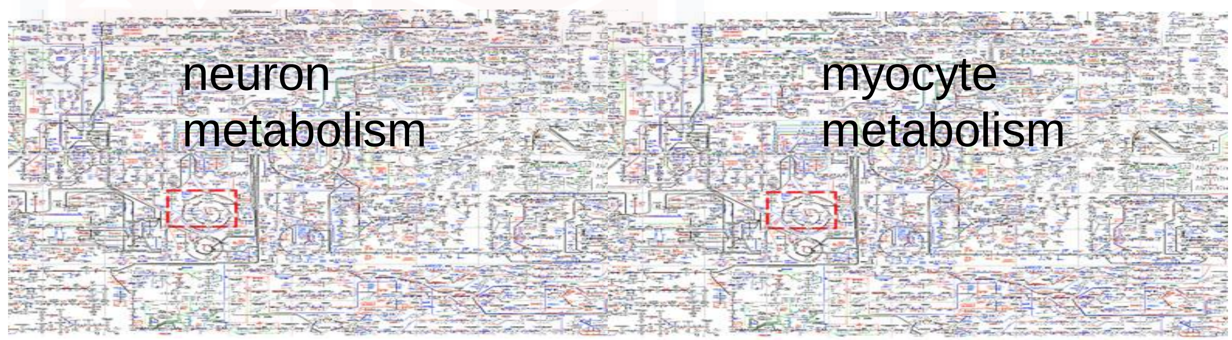
2. What type of sample?

Cells in culture: data only needs to be input into one set of pathways



e.g., muscle tissue

Animals: data needs to be input into multiple pathways



Experimental Design: Considerations

2. What type of sample?

sample type	pro	con
cells in culture	well controlled, high throughput, cost effective, one cell type	physiological relevance
plants, animals	more bio relevant than cell culture	multicellularity, less cost effective
people*	physiological relevance	large variability, cost, healthy samples can be challenging to obtain, IRB paperwork, hard to control variables (environment, diet, medications, exercise, stress, etc)

* Need large sample cohorts to average out variability (how large?)
* Individual sample runs are short enough that analysis of large cohorts is feasible

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods
- Rate-limiting step in data acquisition?

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods

- Rate-limiting step in data acquisition?
  chromatography

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods
- Rate-limiting step in data acquisition?

chromatography

"putting a horse carriage in front of a race car"

Richard Gross, WU

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods

- Rate-limiting step in data acquisition?
  chromatography

Advantage: SPEED

"putting a horse carriage in front of a race car"

Richard Gross, WU

shotgun lipidomics, MALDI-based approaches, flow-injection analysis, NMR, ….

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods
- Rate-limiting step in data acquisition?
  chromatography
- Short separation times make large-scale studies practical

Sreekumar et al. (Nature 2009) used 16-min run to analyze >200 tissue, plasma, and urine samples

Wang et al. (Nature Medicine 2011) used a 30-min run to analyze >1500 plasma samples

Wang et al. (Nature 2011) used a 14.5-min run to analyze 2000 plasma samples

Kurland et al., (J Proteome Res 2011) used a 10-min run to perform untargeted metabolomics

Experimental Design: Considerations

2. What type of sample?

- Intimately related to experimental methods
- Rate-limiting step in data acquisition?

  chromatography
- Short separation times make large-scale studies practical

Sreekumar et al. (Nature 2009) used 16-min run to analyze >200 tissue, plasma, and urine samples

Wang et al. (Nature Medicine 2011) used a 30-min run to analyze >1500 plasma samples

Wang et al. (Nature 2011) used a 14.5-min run to analyze 2000 plasma samples

Sebastani et al., (Cell Reports 2024) used 30-min run to perform untargeted metabolomics on 10k plasma

trade-off between coverage and speed

Experimental Design: Considerations

2. What type of sample?

- Biofluids (typically 100 μL is sufficient)
  plasma, CSF, urine, sputum, tears, etc.

- Tissues (typically 5-10 mg is sufficient)
  brain, *spinal cord*, liver, heart, kidney, muscle, prostate, etc.

Experimental Design: Considerations

2. What type of sample?

- Biofluids (typically 100 μL is sufficient)
  plasma, CSF, urine, sputum, tears, etc.

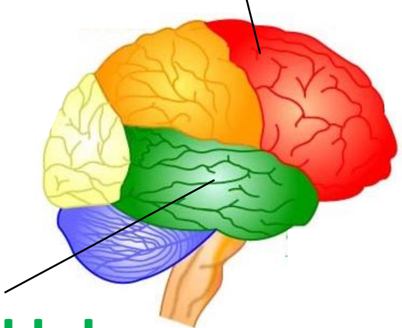
- Tissues (typically 5-10 mg is sufficient)
  brain, *spinal cord*, liver, heart, kidney, muscle, prostate, etc. Dissect if possible.

frontal lobe

temporal lobe

dysregulated features after morphine withdrawal

frontal lobe



temporal lobe

Experimental Design: Considerations

3. Analytical vs. biological replicates

Experimental Design: Considerations

3. Analytical vs. biological replicates

Analytical replicates: repeating the analysis on the identical sample

Biological replicates: repeating the analysis on another animal from the same bio group.

biological variability>>>analytical variability
(do not conflate during data processing)

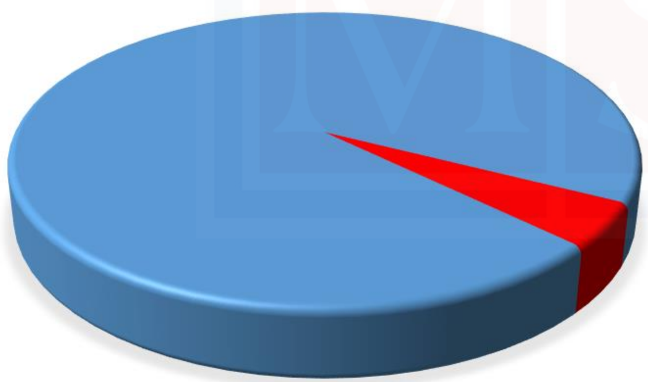
NOTE: only pool samples for quality control or MS/MS analyses

Experimental Design: Considerations

3. Analytical vs. biological replicates

Experience says do not do pilot studies with small numbers of biological replicates (e.g., 2 vs 2)

* About same time to perform analysis
* Metabolite differences may not be real



total experiment time

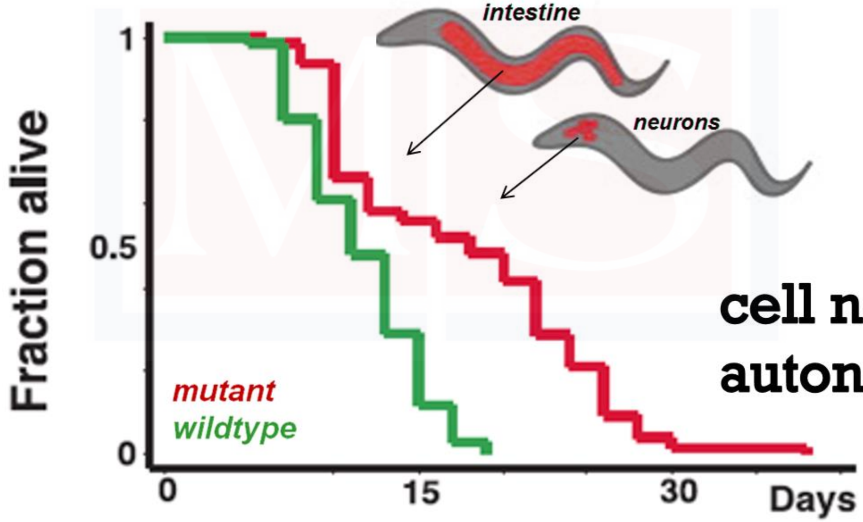
Experimental Design: Considerations

4. Using multiple sample groups

Experimental Design: Considerations

4. Using multiple sample groups

Ex: metabolomics to find aging signals

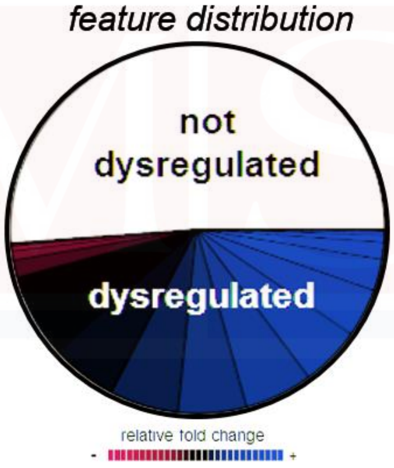


cell non-autonomous

Experimental Design: Considerations

4. Using multiple sample groups

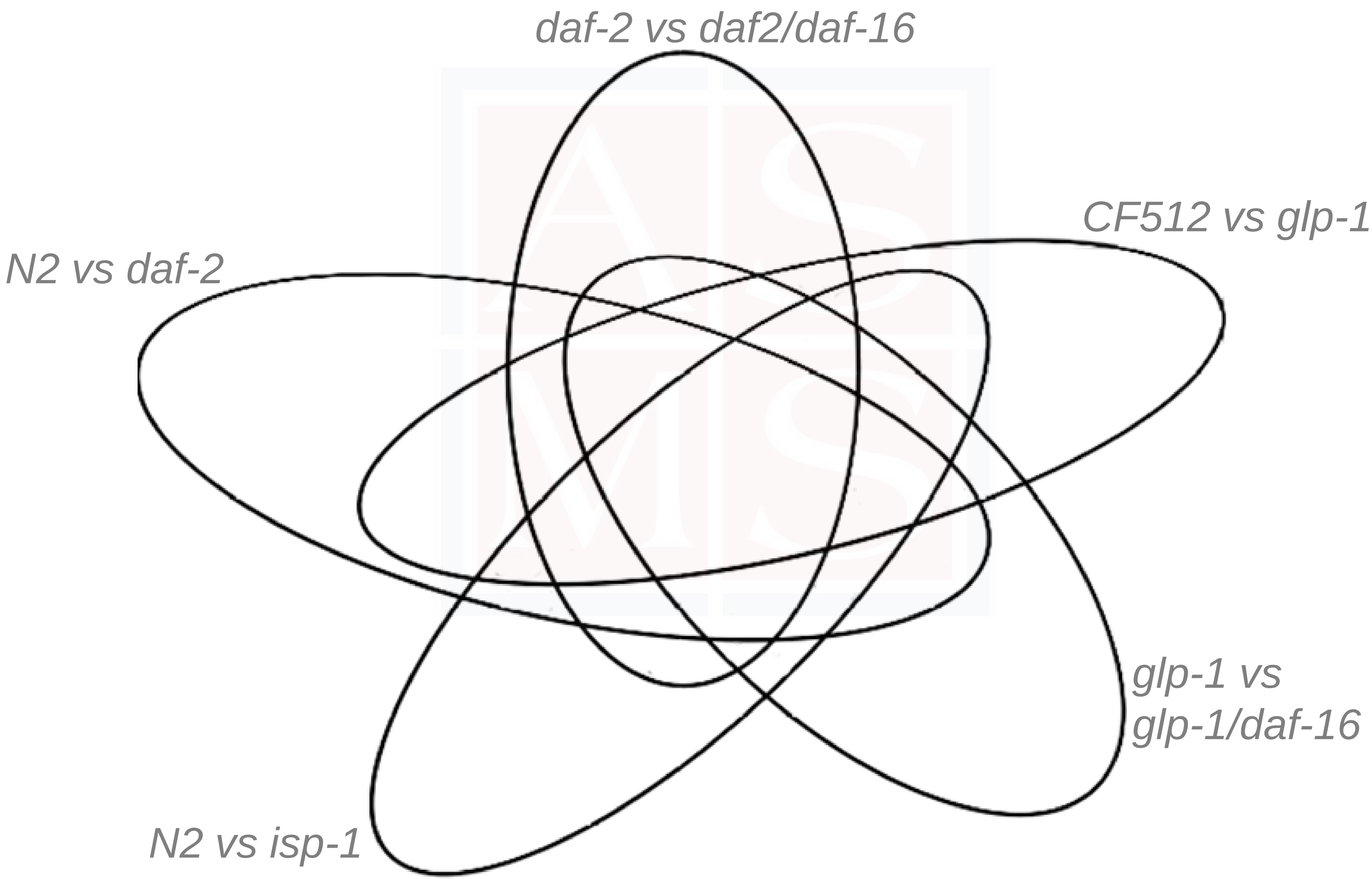
Ex: metabolomics to find aging signals



glp-1 *mutants vs.* CF512 *controls*

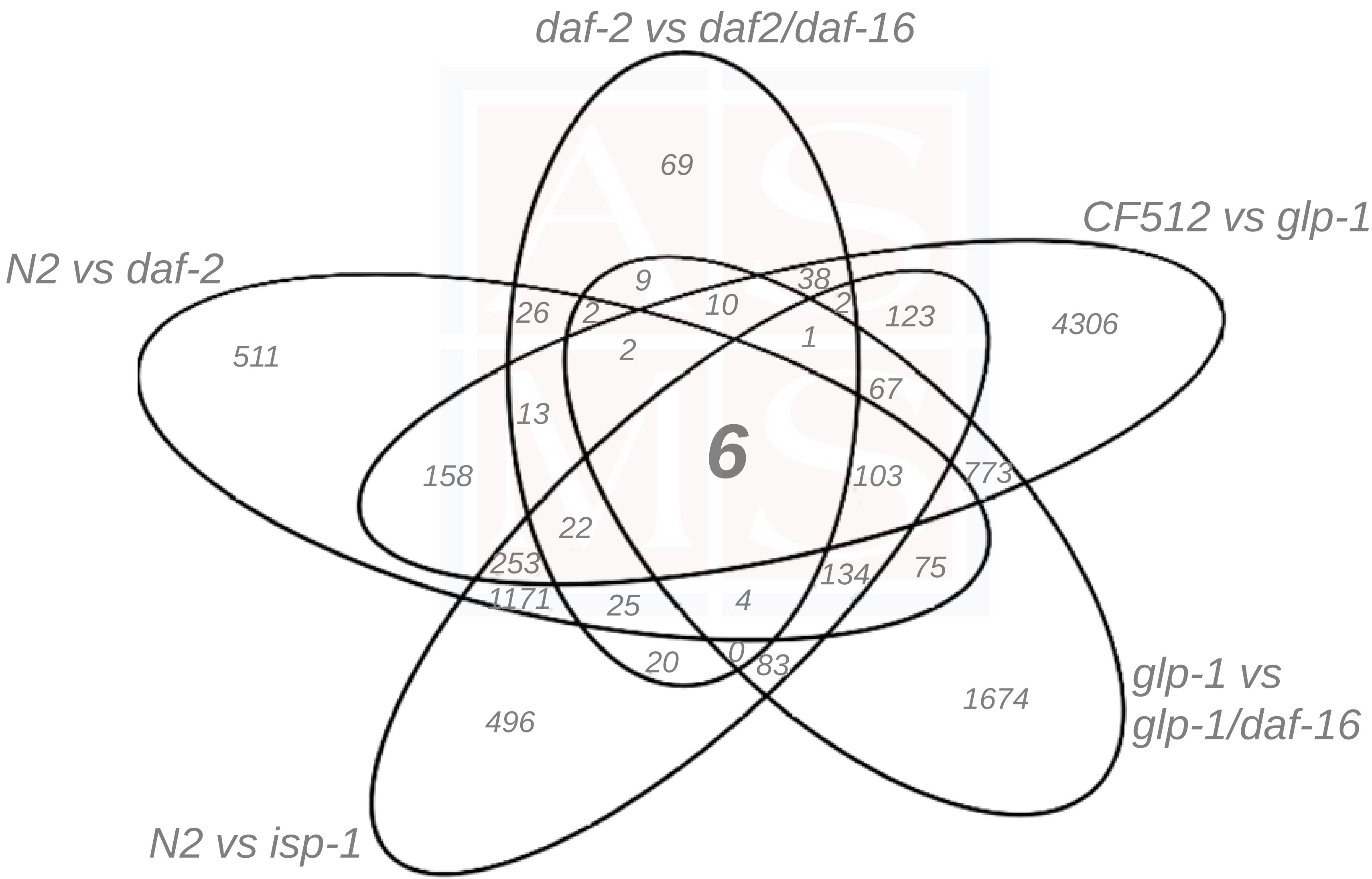
Experimental Design: Considerations

4. Using multiple sample groups



Experimental Design: Considerations

4. Using multiple sample groups



Experimental Design: Considerations

5. Choice of instrument and chromatography

- Choices bias coverage (much more later)
- HILIC-MS and RPLC-MS are most popular
- GC/MS may be best for some analyses (steroids, hormones, etc.)

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation

Experimental Design: Considerations

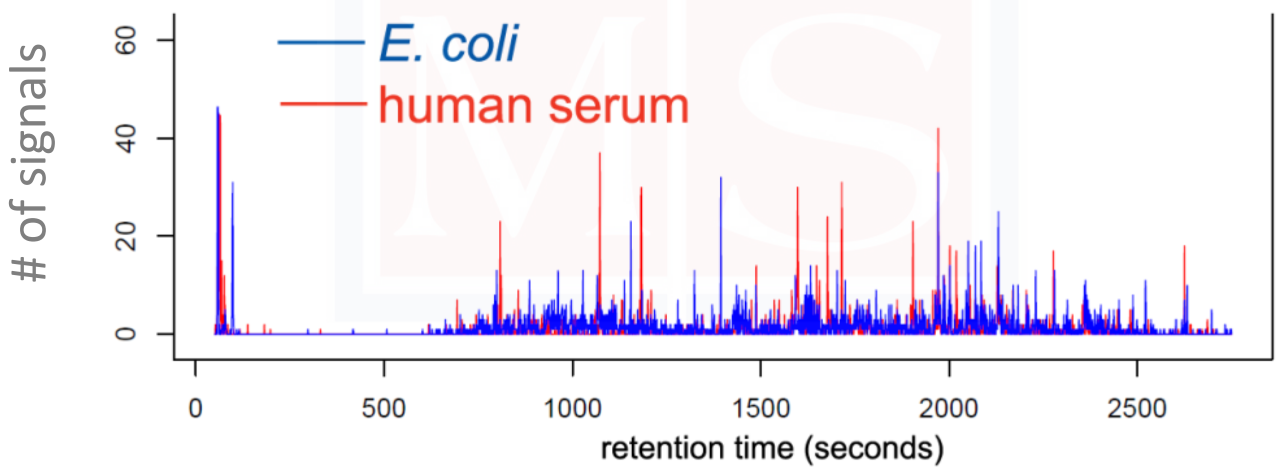
5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)



50 min C18 run
isotopes removed

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)

Experimental Design: Considerations

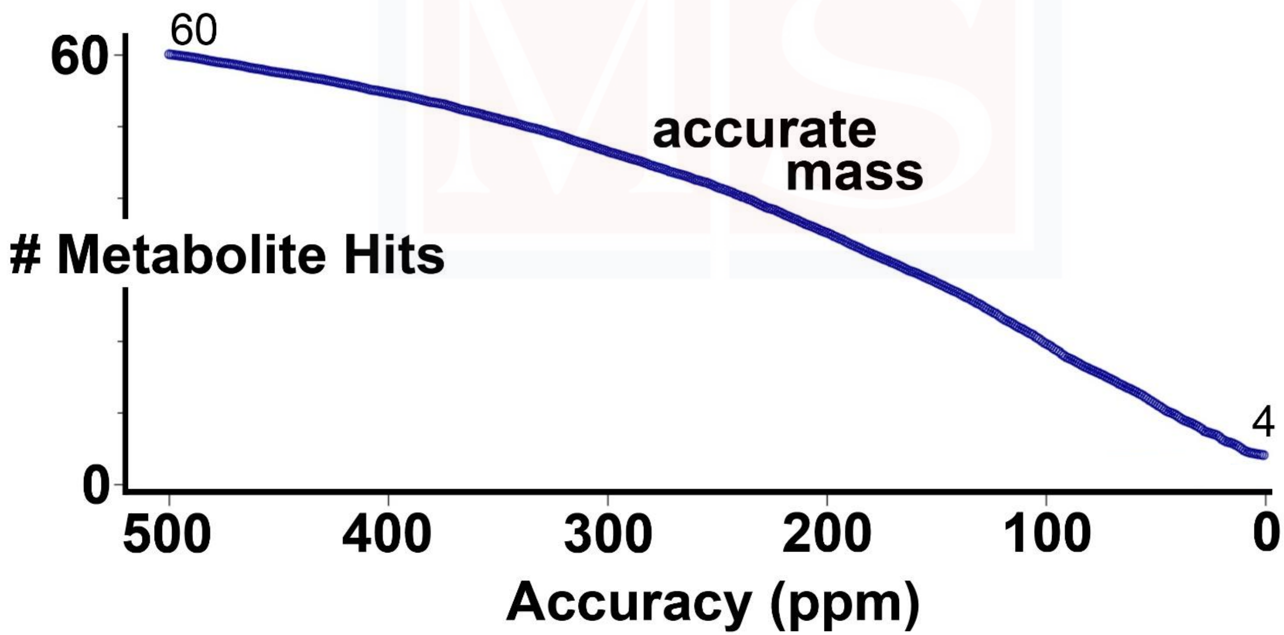
5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)

Experimental Design: Considerations

5. Choice of instrument and chromatography

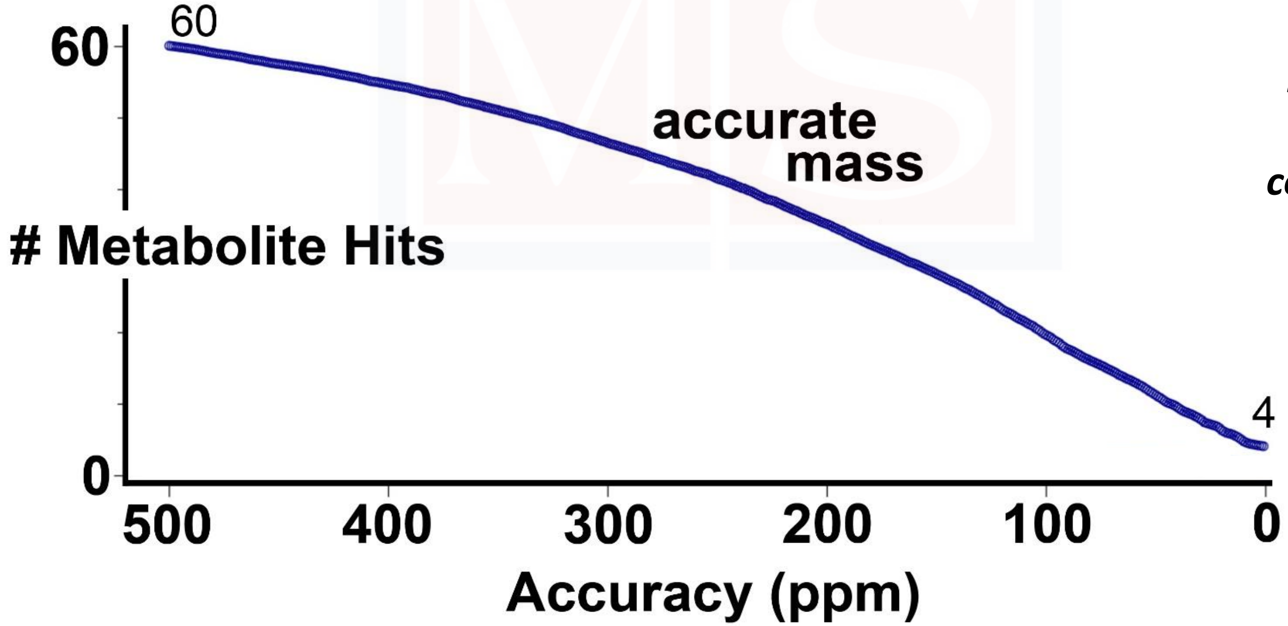
- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)



Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)



Ignores adducts, fragments, contaminants, and role in isotope tracer analysis

Experimental Design: Considerations

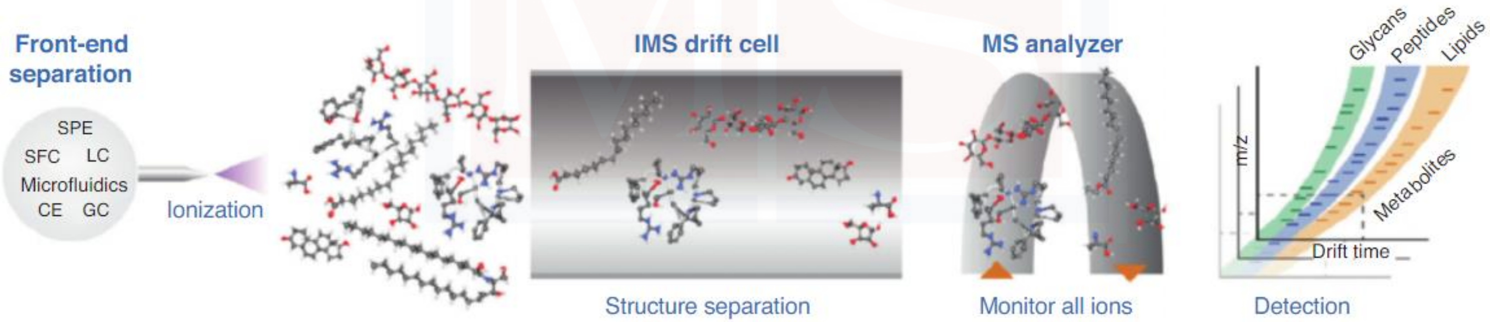
5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)



Limited by ion suppression when used without LC
Multidimensional software still lacking

Reisdorph et al., Anal Chem 2017
Williams et al., Anal Bioanal Chem 2015
Zhang et al., Clinical Mass Spec 2016
Kyle et al., Analyst 2016

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)

Experimental Design: Considerations

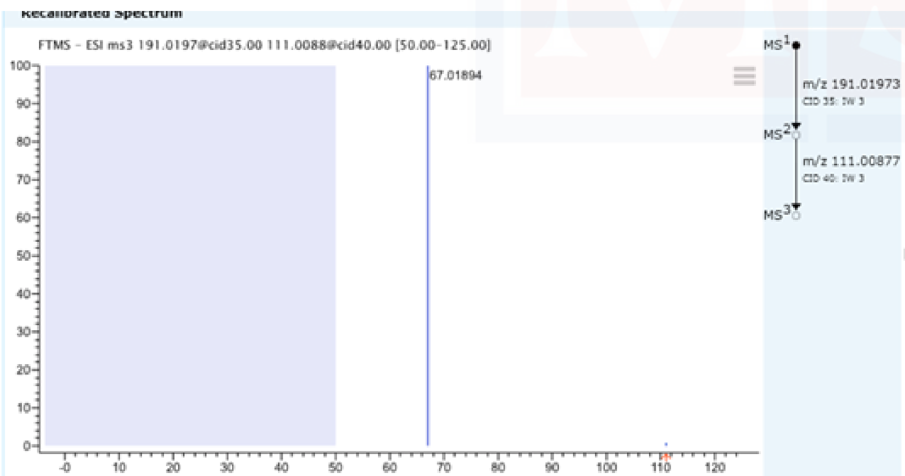
5. Choice of instrument and chromatography

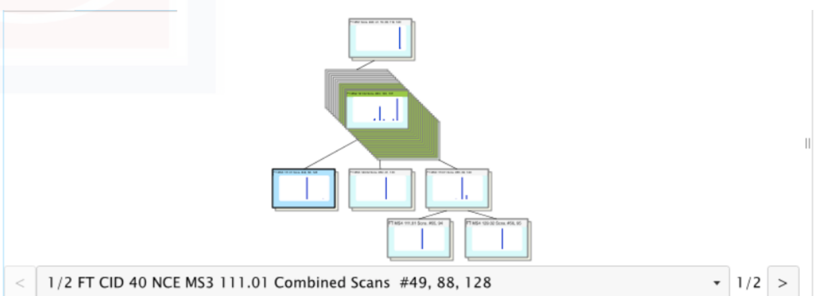
- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)
  - MS^n (FTMS, Thermo Tribrid)

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)
  - MS^n (FTMS, Thermo Tribrid)





Experimental Design: Considerations

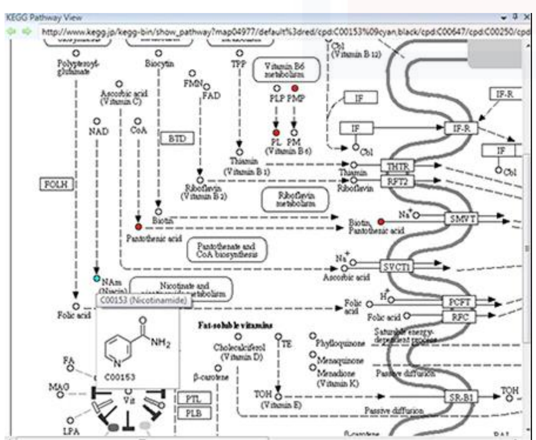
5. Choice of instrument and chromatography

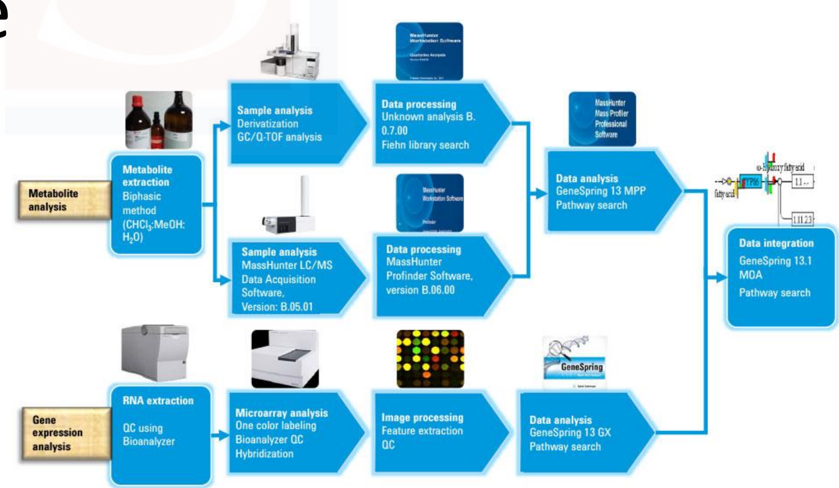
- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)
  - MS^n (FTMS, Thermo Tribrid)

Experimental Design: Considerations

5. Choice of instrument and chromatography

- Some considerations for MS instrumentation
  - Acquisition speed (SCIEX ZenoTOF/Astral)
  - Mass accuracy (Orbitraps)
  - Ion mobility (Waters, Agilent, Bruker)
  - MS^n (FTMS, Thermo Tribrid)
  - Vendor software





- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Evaluating performance

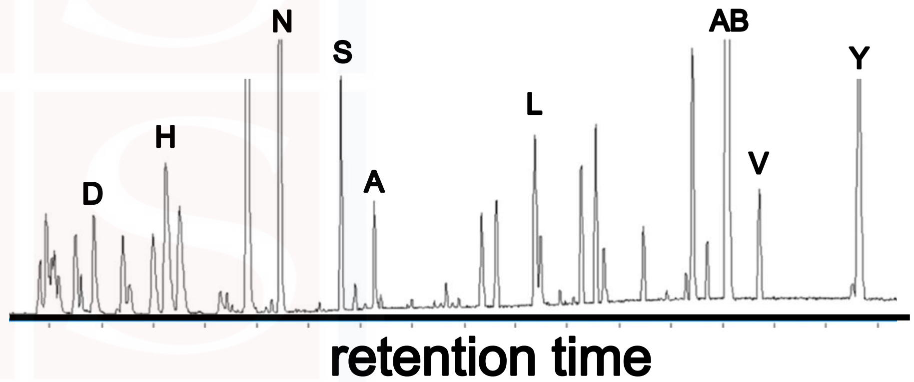
How do we choose methods?

ideal situation:

complete metabolite parts list

932/1781 = 52%

metabolite A	☑	metabolite P	☐
metabolite B	☐	metabolite Q	☑
metabolite C	☑	metabolite R	☐
metabolite D	☑	metabolite S	☑
metabolite E	☑	metabolite T	☐
metabolite F	☐	metabolite U	☐
metabolite G	☐	metabolite V	☑
metabolite H	☑	metabolite W	☐
metabolite I	☐	metabolite X	☐
metabolite J	☐	metabolite Y	☑
metabolite K	☐	metabolite Z	☐
metabolite L	☑	metabolite AA	☐
metabolite M	☑	metabolite AB	☑
metabolite N	☑	metabolite AC	☐
metabolite O	☐	metabolite AD	☐



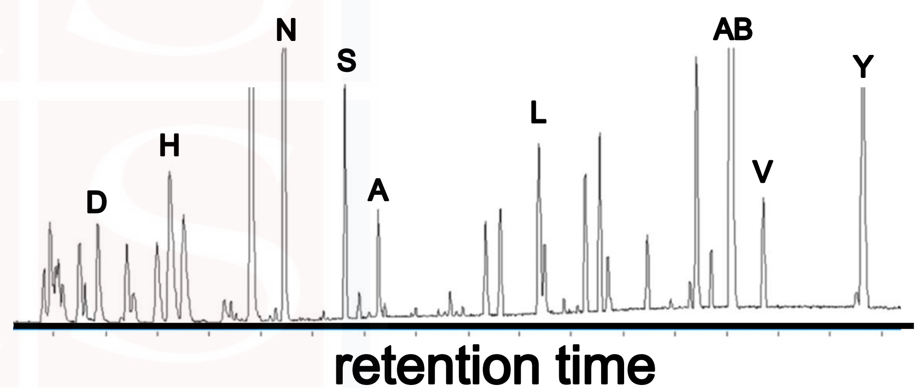
How do we choose methods?

ideal situation:

complete metabolite parts list

932/1781 = 52%

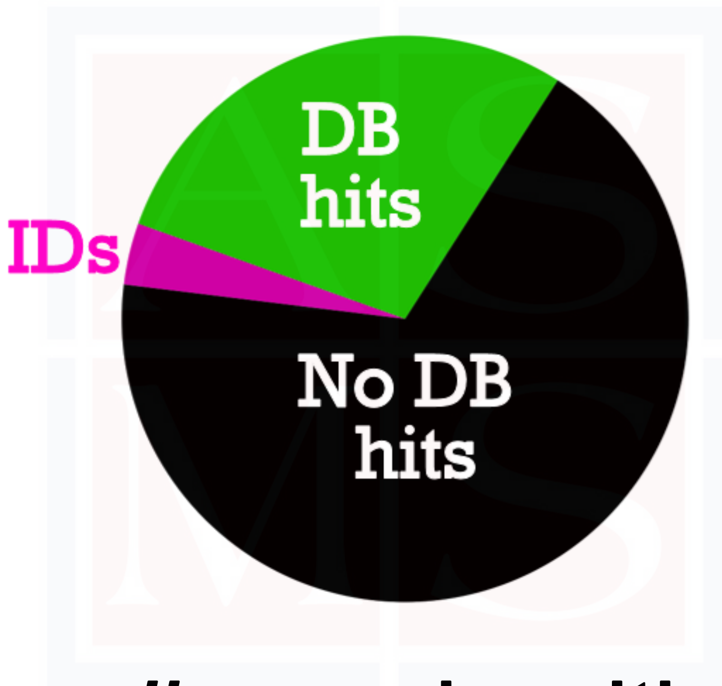
metabolite A	☑	metabolite P	☐
metabolite B	☐	metabolite Q	☑
metabolite C	☑	metabolite R	☐
metabolite D	☑	metabolite S	☑
metabolite E	☑	metabolite T	☐
metabolite F	☐	metabolite U	☐
metabolite G	☐	metabolite V	☑
metabolite H	☑	metabolite W	☐
metabolite I	☐	metabolite X	☐
metabolite J	☐	metabolite Y	☑
metabolite K	☐	metabolite Z	☐
metabolite L	☑	metabolite AA	☐
metabolite M	☑	metabolite AB	☑
metabolite N	☑	metabolite AC	☐
metabolite O	☐	metabolite AD	☐



Why not?

How do we choose methods?

the challenge



E. coli sample with ~25k features, <1000 identified

How do we choose methods?

Alternatives?

How do we choose methods?

Alternatives?

feature: an ion with a unique mass-to-charge ratio and retention time

How do we choose methods?

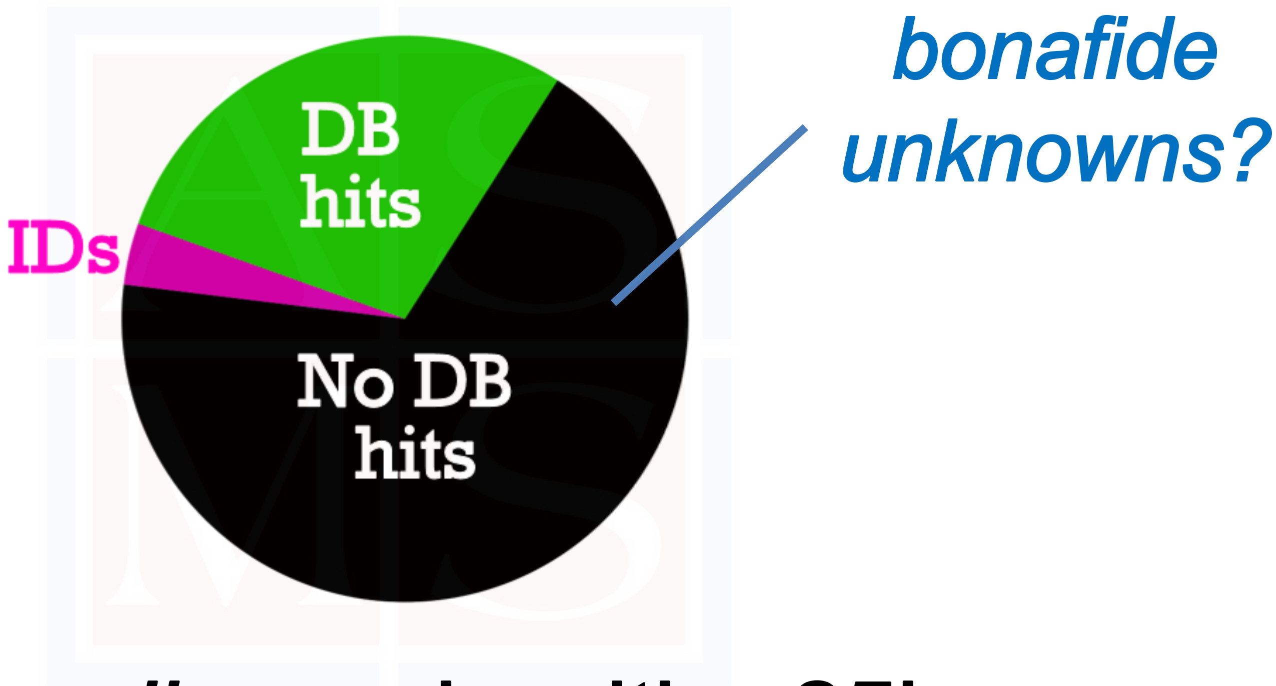
Alternatives?

feature: an ion with a unique mass-to-charge ratio and retention time

- *Historically, benchmarking based on feature numbers*
- *Assumes feature number is directly correlated to metabolite number*
- *Can select for worse methods*

How do we choose methods?

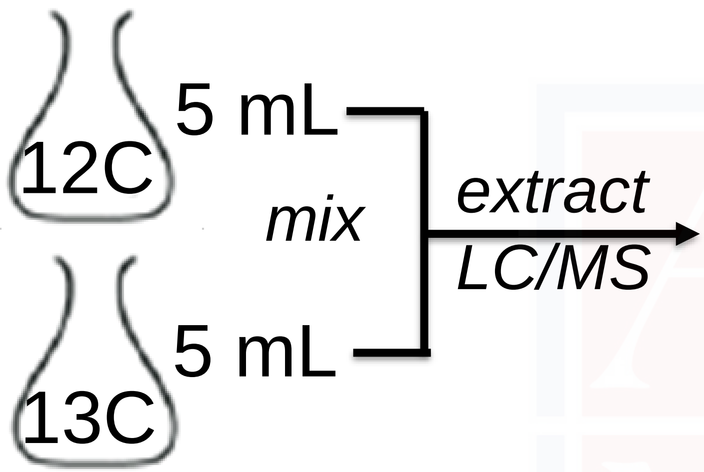
the challenge



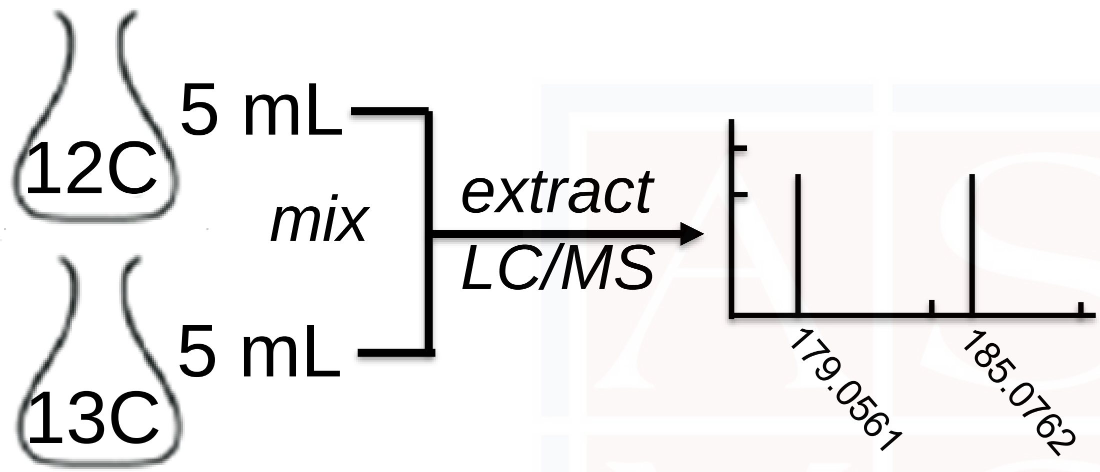
E. coli sample with ~25k features, <1000 identified

Credentialing features in *E. coli*

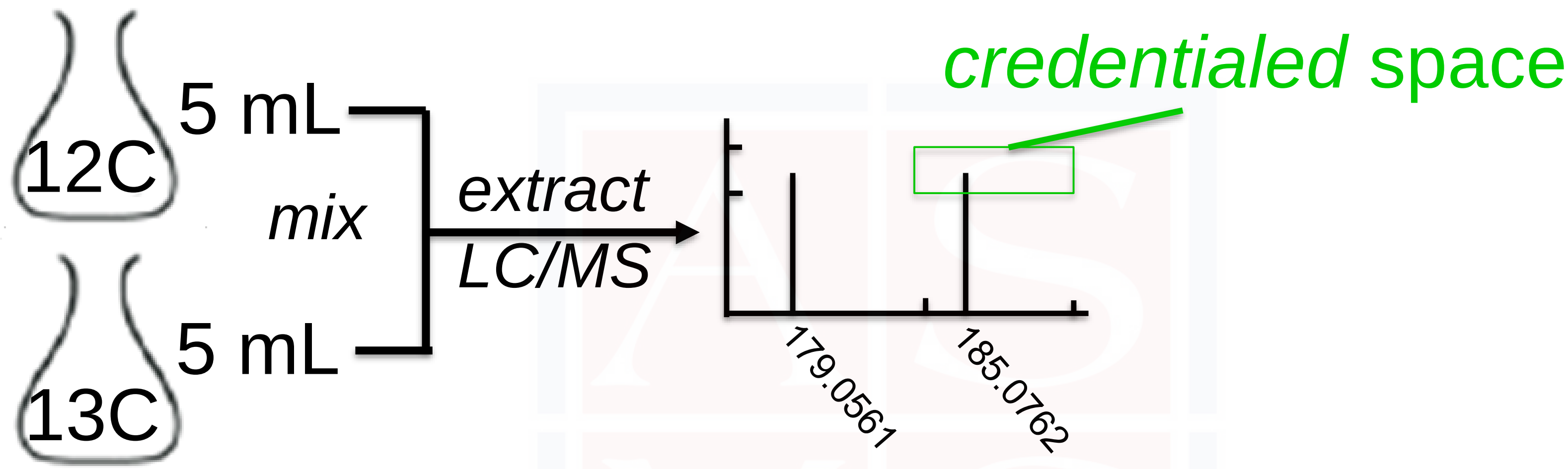
Credentialing features in *E. coli*



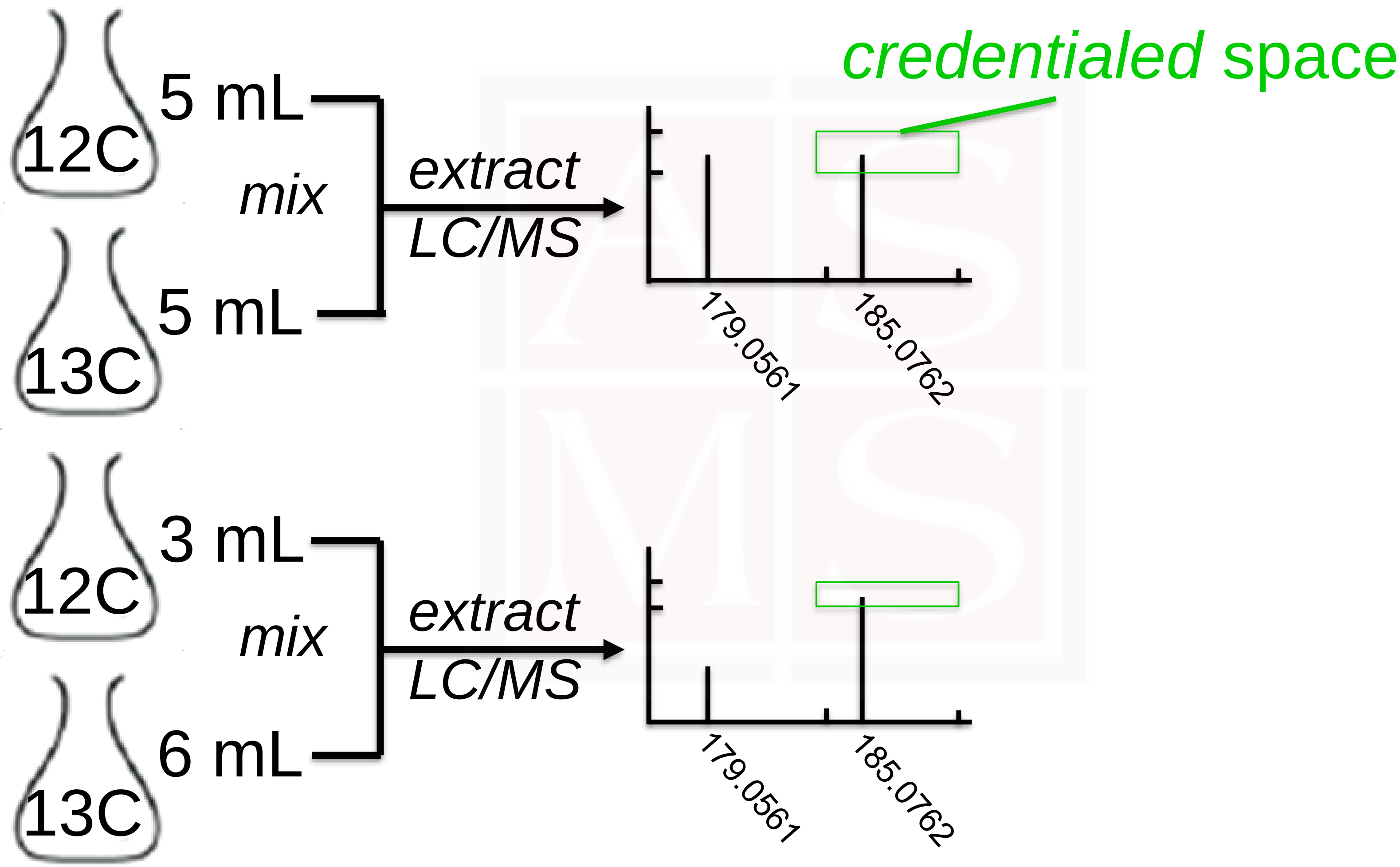
Credentialing features in *E. coli*



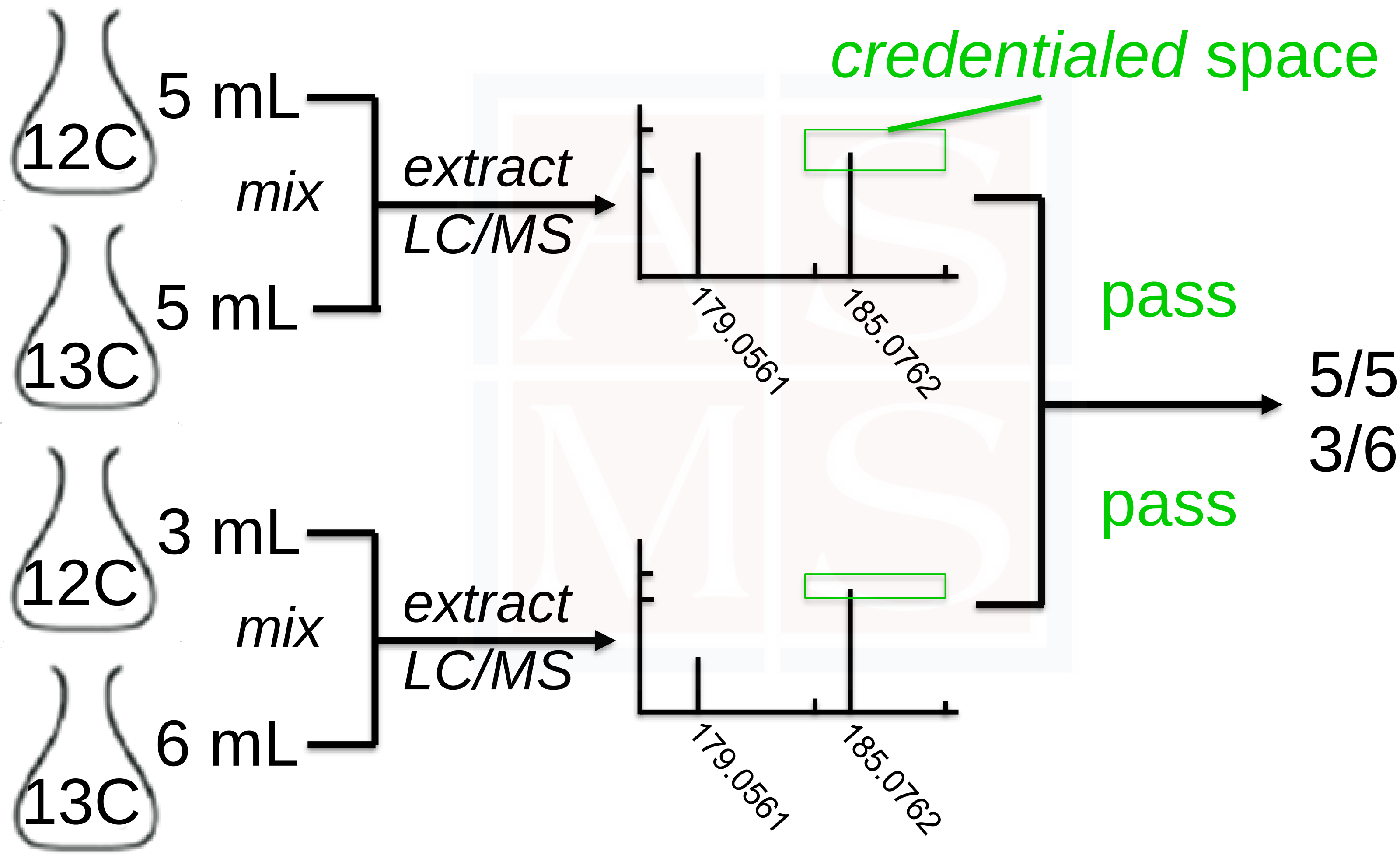
Credentialing features in *E. coli*



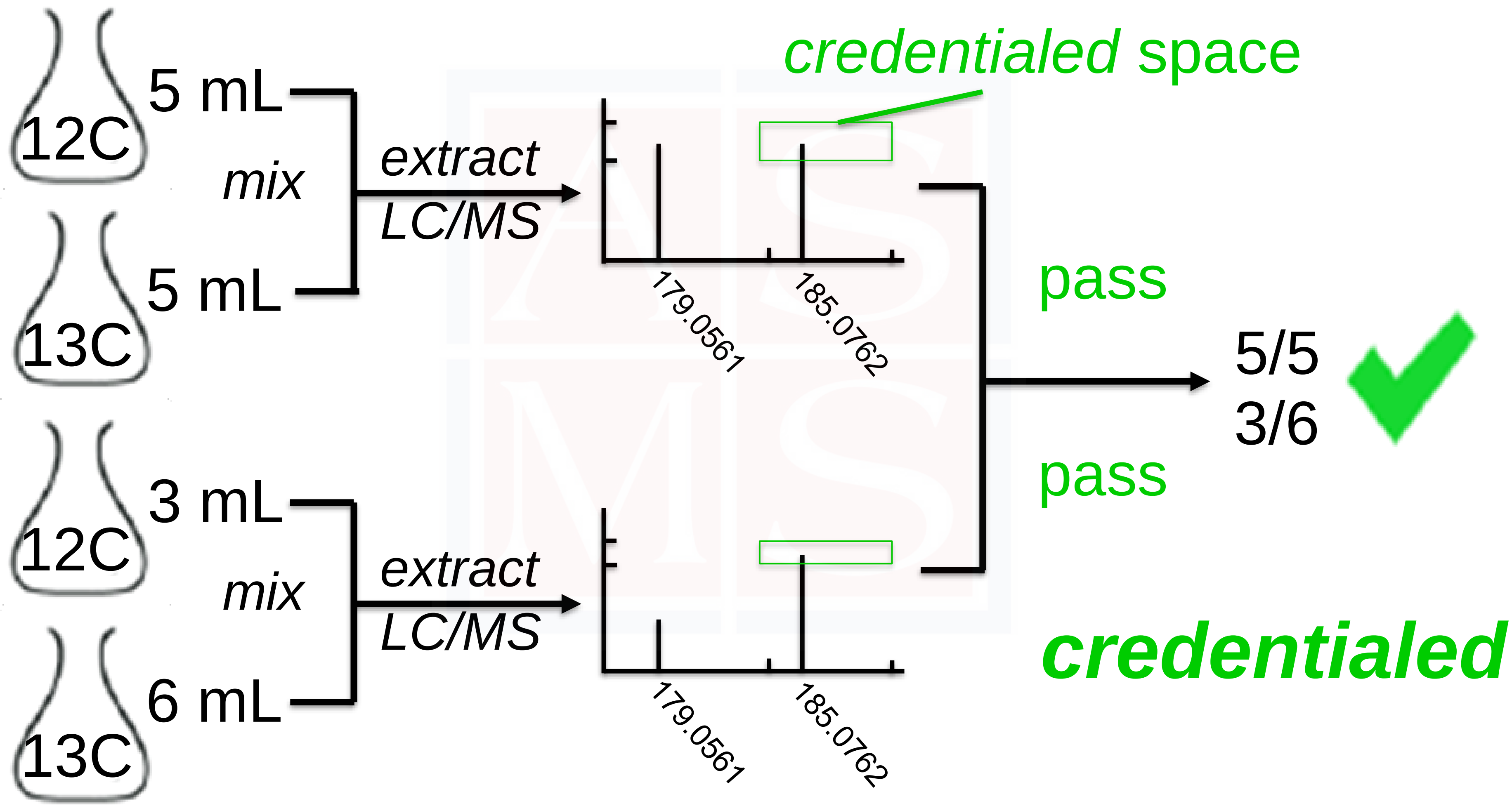
Credentialing features in *E. coli*



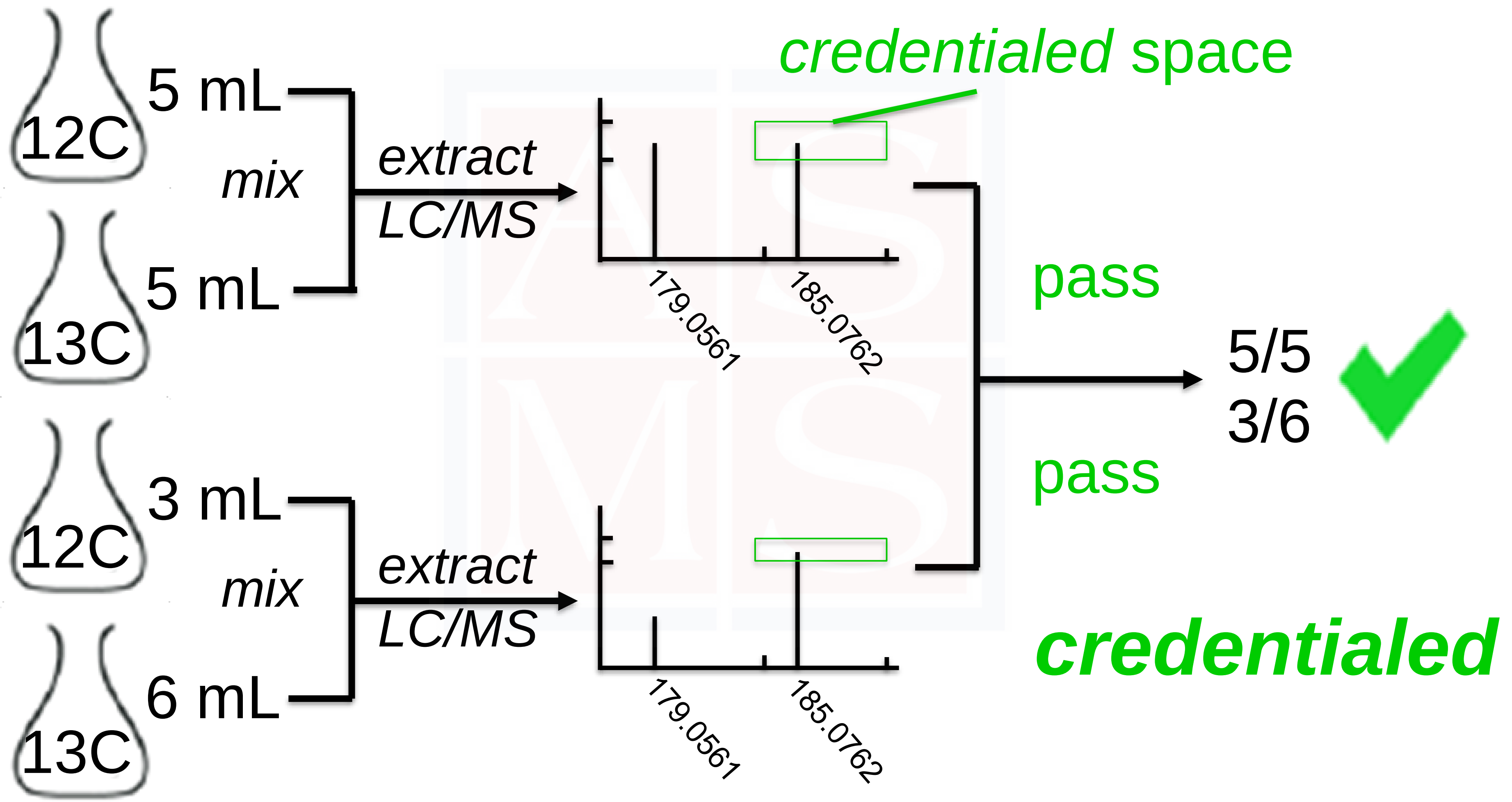
Credentialing features in *E. coli*



Credentialing features in *E. coli*



Credentialing features in *E. coli*



removes 100s-1000s of features

Cambridge Isotope Laboratories, Inc.
isotope.com

METABOLIC RESEARCH

Credentialed *E. coli* Cell Extract Kit

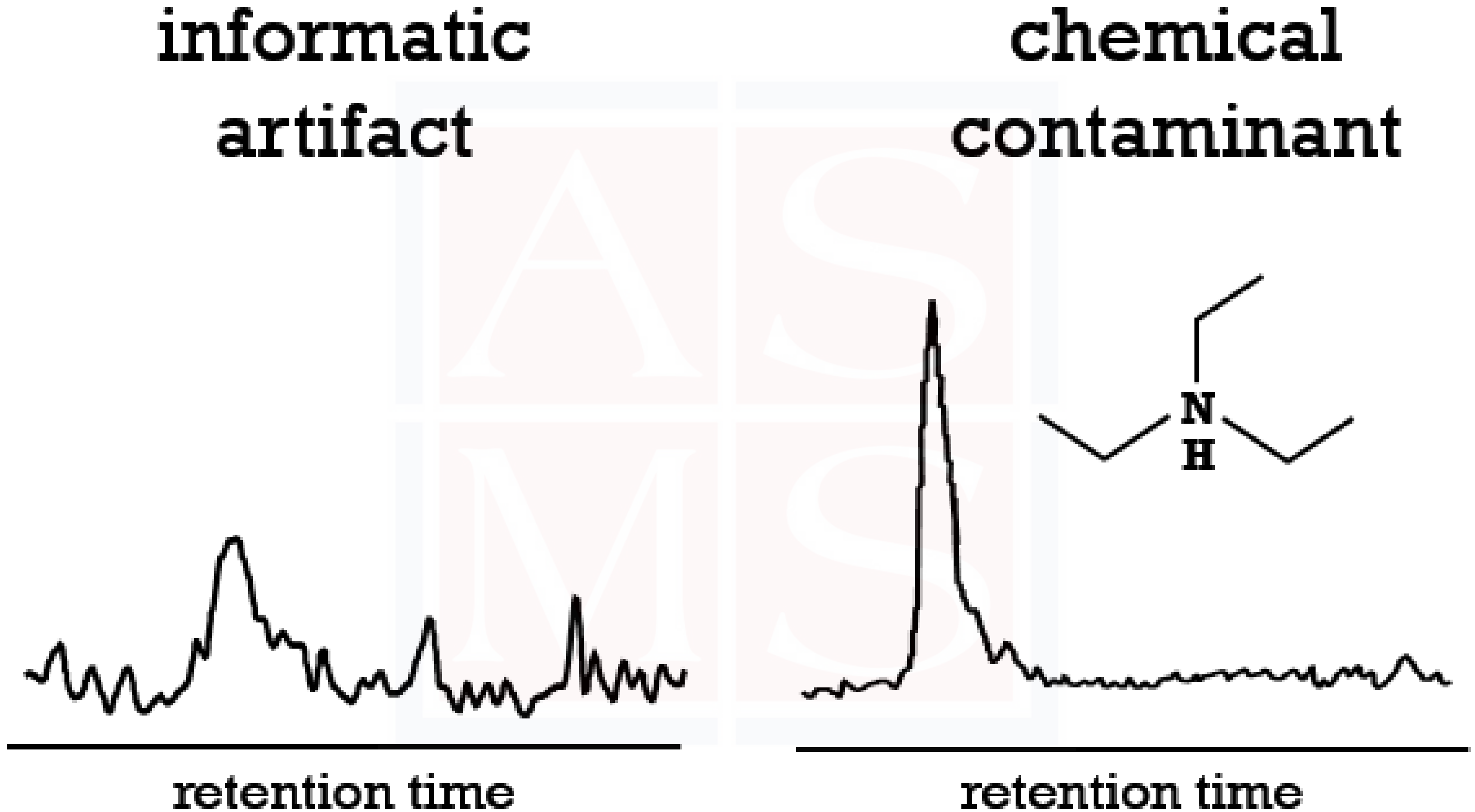
for Benchmarking and Optimizing Methods

Cambridge Isotope Laboratories, Inc.
isotope.com

Credentialed *E. coli* Cell Extract Kit for Benchmarking and Optimizing Untargeted MS Methods

User's Manual

Artifacts vs contaminants



Credentialing vs feature counting

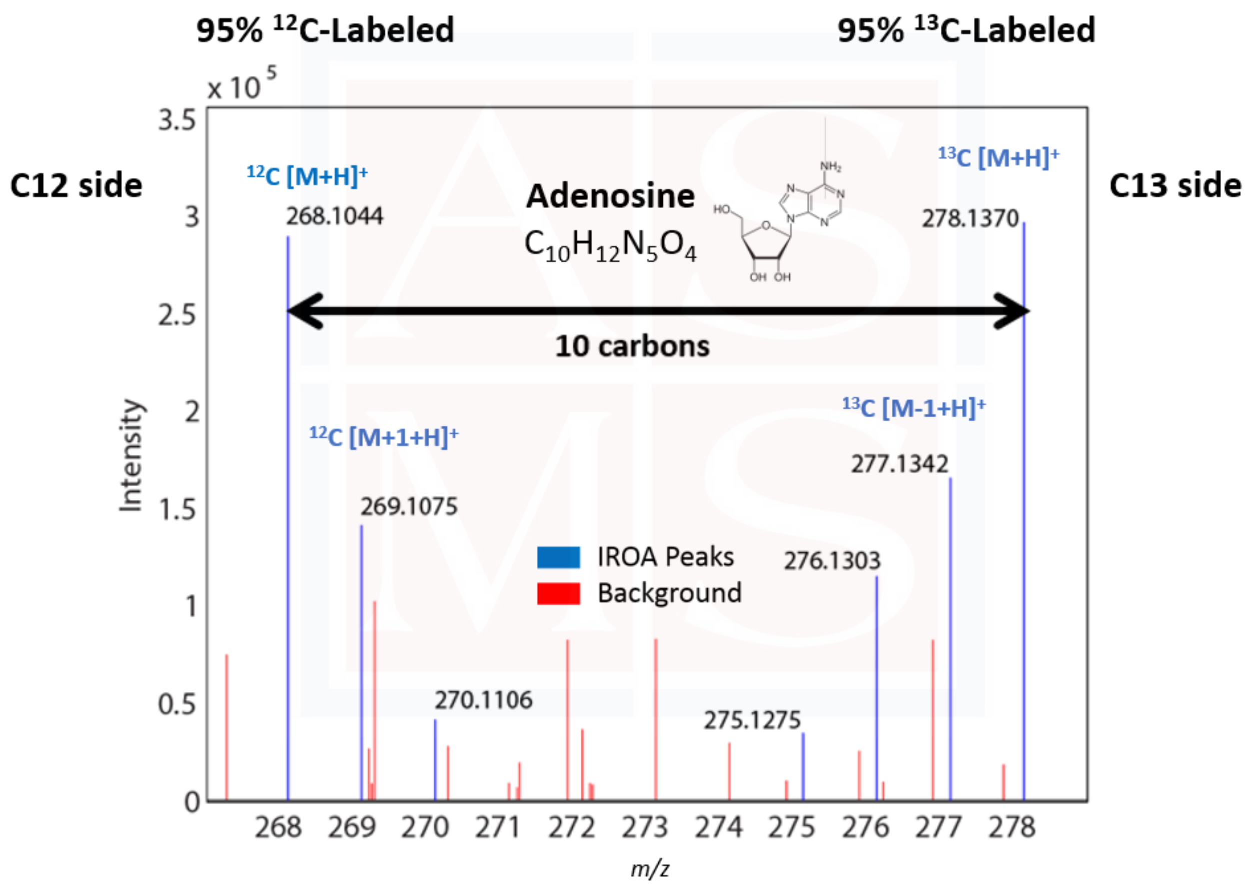
Table 1

Comparison of *credentialed* features between methods

Method	No. of signals[a]	No. credentialed	No. of standards
XBridge C_{18}	4166	173	10
CORTECS C_8	13,180	837	27
CORTECS T3	14,970	679	21

[a]Only highly abundant features were counted.

Isotope Ratio Outlier Analysis



IROA Technologies

Analysis without IDs is not very useful

- ***Do not benchmark method by feature number***

- ***Do not try to compare samples qualitatively by percentage of feature changes***

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Sample prep and extraction

From intact samples to metabolomic analysis

purify/wash/ normalize sample

quench metabolism

dry samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

Lots of possibilities for
each step….
infinite number of
permutations for
overall workflow

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

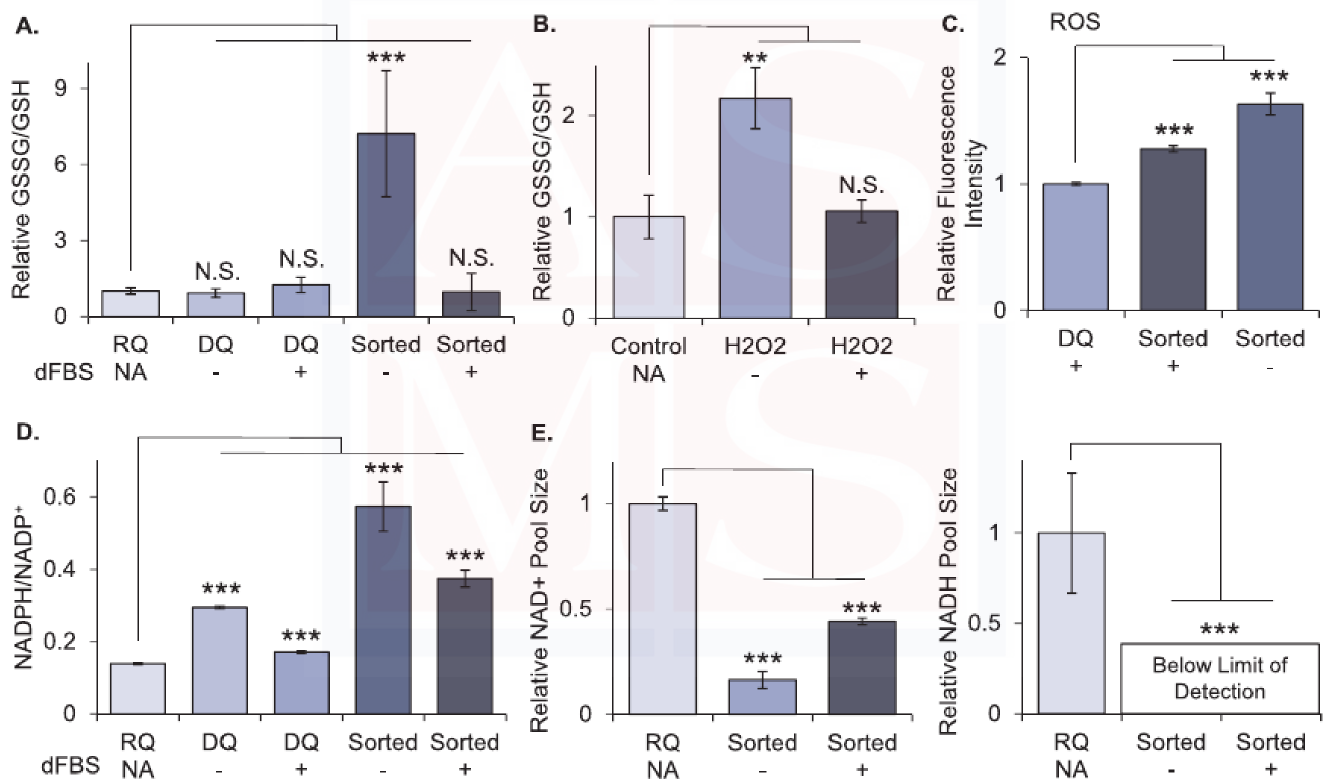
extract

concentrate

Lots of possibilities for
each step….
infinite number of
permutations for
overall workflow

Purifying, washing, and normalizing samples

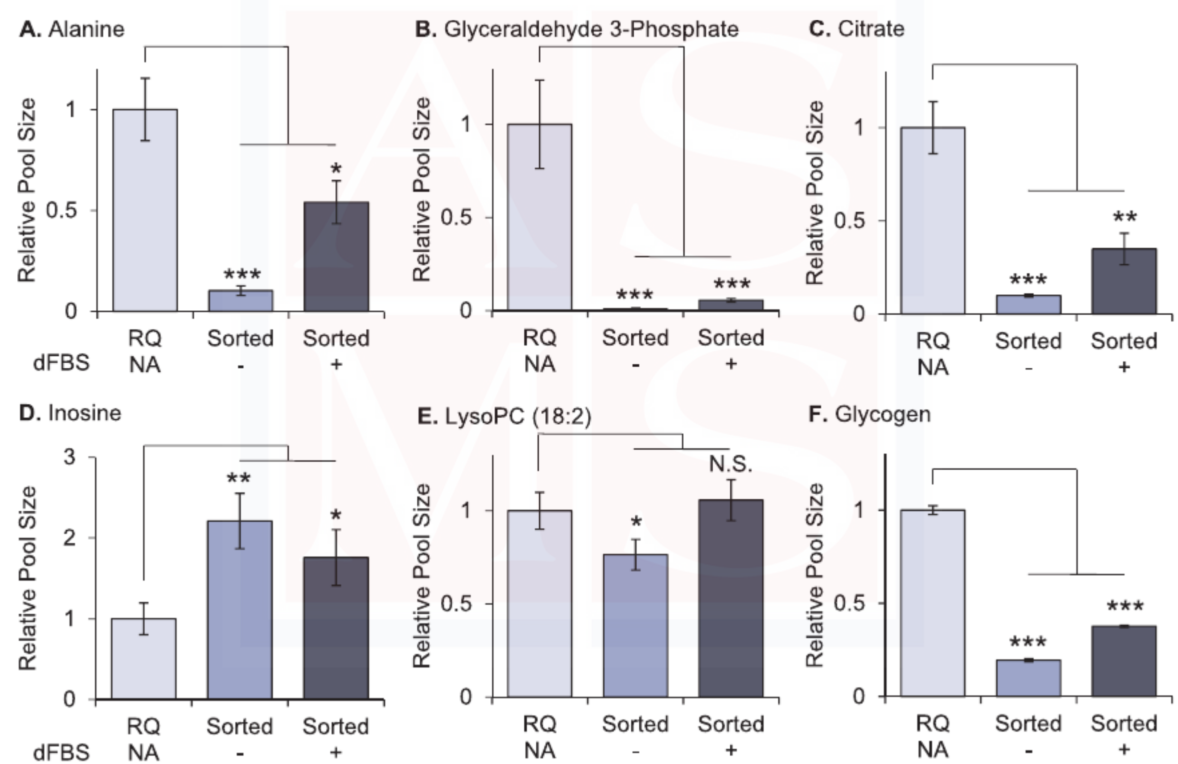
Tempting to sort samples to purify cell populations, but this alters the metabolome.



Llufrio et al., Redox Biology 2018

Purifying, washing, and normalizing samples

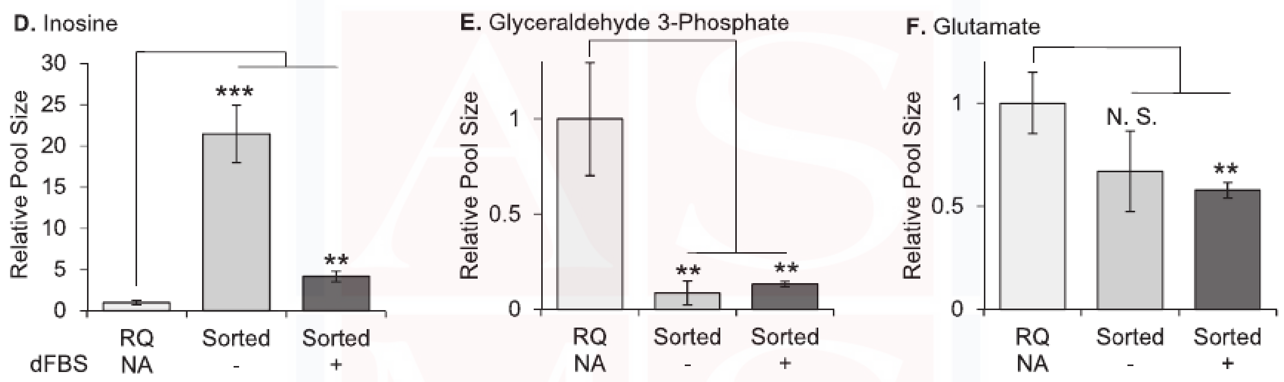
Tempting to sort samples to purify cell populations, but this alters the metabolome.



Llufrio et al., Redox Biology 2018

Purifying, washing, and normalizing samples

Tempting to sort samples to purify cell populations, but this alters the metabolome.

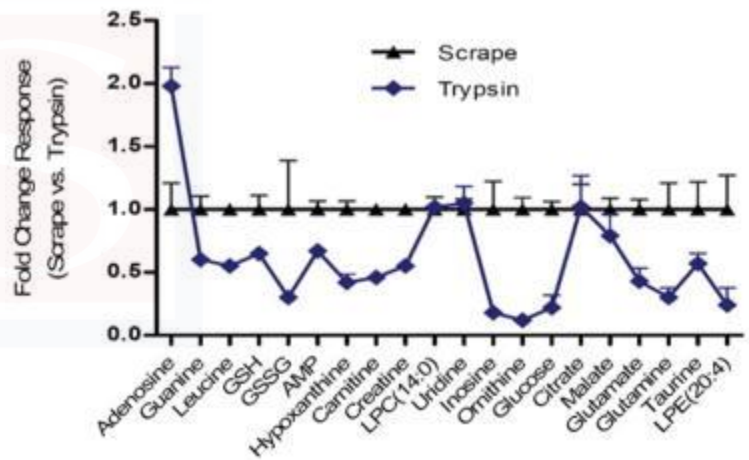


NOTE: cannot "normalize" out changes

Llufrio et al., Redox Biology 2018

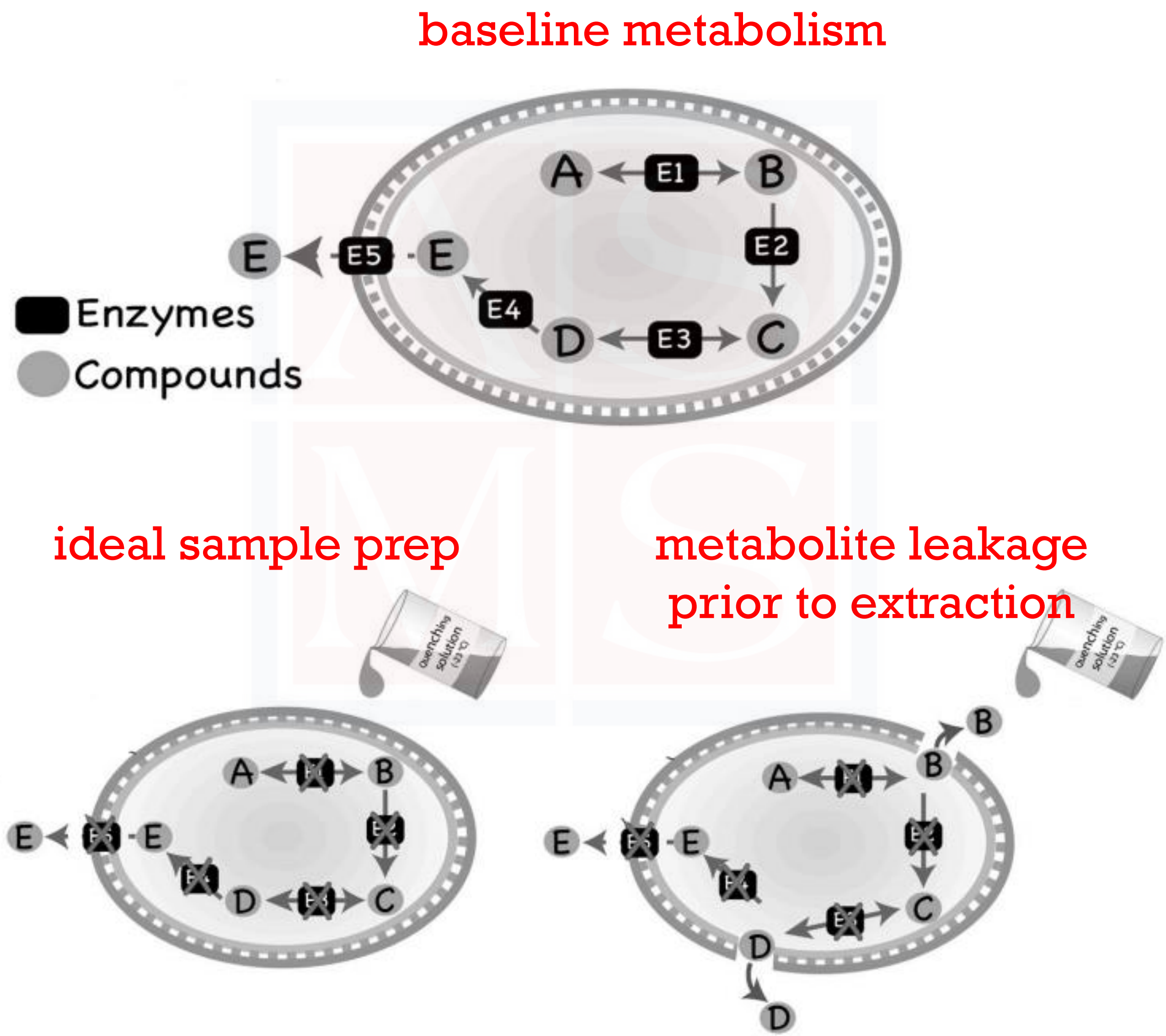
Purifying, washing, and normalizing samples

- In cell culture, after removing media, cells often rinsed with PBS.
- Can take significant time if have large number of samples.
- Trysinization
  - <u>pro</u>: enables cell count and analysis of protein conc.
  - <u>con</u>: metabolite leakage
- Cell scraping
  - <u>pro</u>: less leakage
  - <u>con</u>: no normalization



Gonzalez et al., Anal Bioanal Chem 2014

Purifying, washing, and normalizing samples



Pinu et al., Metabolites 2017

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

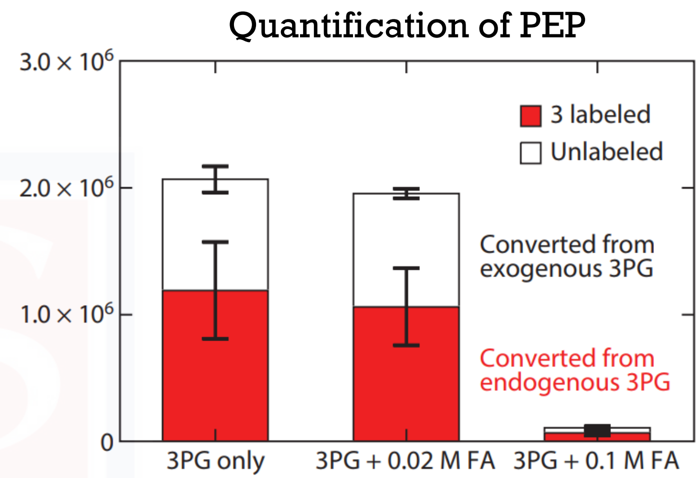
extract

concentrate

<u>quenching (def)</u>: rapid inactivation of biochemical and enzymatic activity in a biological system

Quenching: the faster, the better

- Some metabolites turnover in < 1 s (e.g., ATP and glucose 6-phosphate).
- <u>Suspension cells</u>: pelleting generally too slow, fast filtration followed by placing filter disc in quenching solution better.
- <u>Adherent cells</u>: best to add quench solution directly to culture flask (limitation is starts extraction).
- <u>Tissues</u>: liquid nitrogen or smashing against cold metal plates (Wohlenberger clamp).



HEK293 cells grown in $^{13}C_6$-glucose to completely label glycolytic metabolites. Unlabeled 3-phosphoglycerate (3PG) was added to the extraction solvent of 80:20 methanol:H2O at -70 C. FA: formic acid. Phosphoenolpyruate (PEP) is made from 3PG unless 0.1 FA added.

Lu et al., Annu Rev Biochem 2017. 86:277

Quenching: the faster, the better

Table 1. Different techniques for quenching

Organism	Quenching method	Leakage	Separation of media	References
Microbes, cultured cells	Liquid nitrogen	No	No	Tiziani *et al.*, 2009
	Cold methanol	Yes	Yes	Koek *et al.*, 2006
	Cold ethanol	Yes	Yes	Ewald *et al.*, 2009
	Perchloric acid	Yes	Yes	Koek *et al.*, 2006
Yeast	Cold methanol	Yes	Yes	Koning and Dam, 1992
	Methanol base buffer (methanol + 10 mM tricine buffer)	Yes	Yes	Castrillo *et al.*, 2003
	Liquid nitrogen	No	No	Mashego *et al.*, 2003
	Pre-cooled stainless steel beads	No	Yes	Theobald *et al.*, 1997
Fungi	Chilled water	Yes	Yes	Matsuzaki *et al.*, 2008
	Liquid nitrogen	No	No	Hajjaj *et al.*, 1998
	Cold methanol	Yes	Yes	Mashego *et al.*, 2007
Insects	Cold methanol	Yes	NA	Bratty *et al.*, 2011
	Liquid nitrogen	No	NA	Williams *et al.*, 2010
	Cold acetonitrile	Yes	NA	Pedersen *et al.*, 2008
Plants	Liquid nitrogen	No	No	Kim *et al.*, 2009
Animals	Liquid nitrogen	No	NA	Wang *et al.*, 2011
	Cold methanol	Yes	NA	Stentiford *et al.*, 2005

NA, not applicable.

NOTE: beware that quenching solvents can degrade some metabolites

Phytochem Anal 2014, 25, 291-306

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

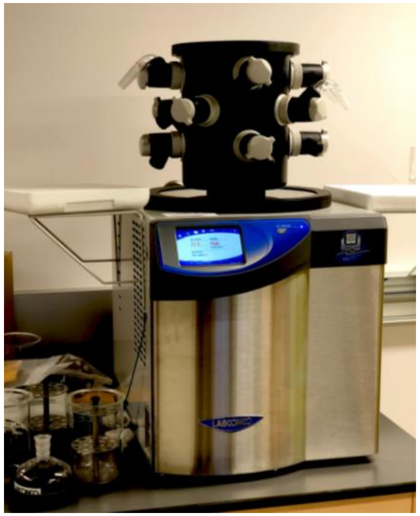
dry
samples

homogenize

extract

concentrate

Methods for drying samples down



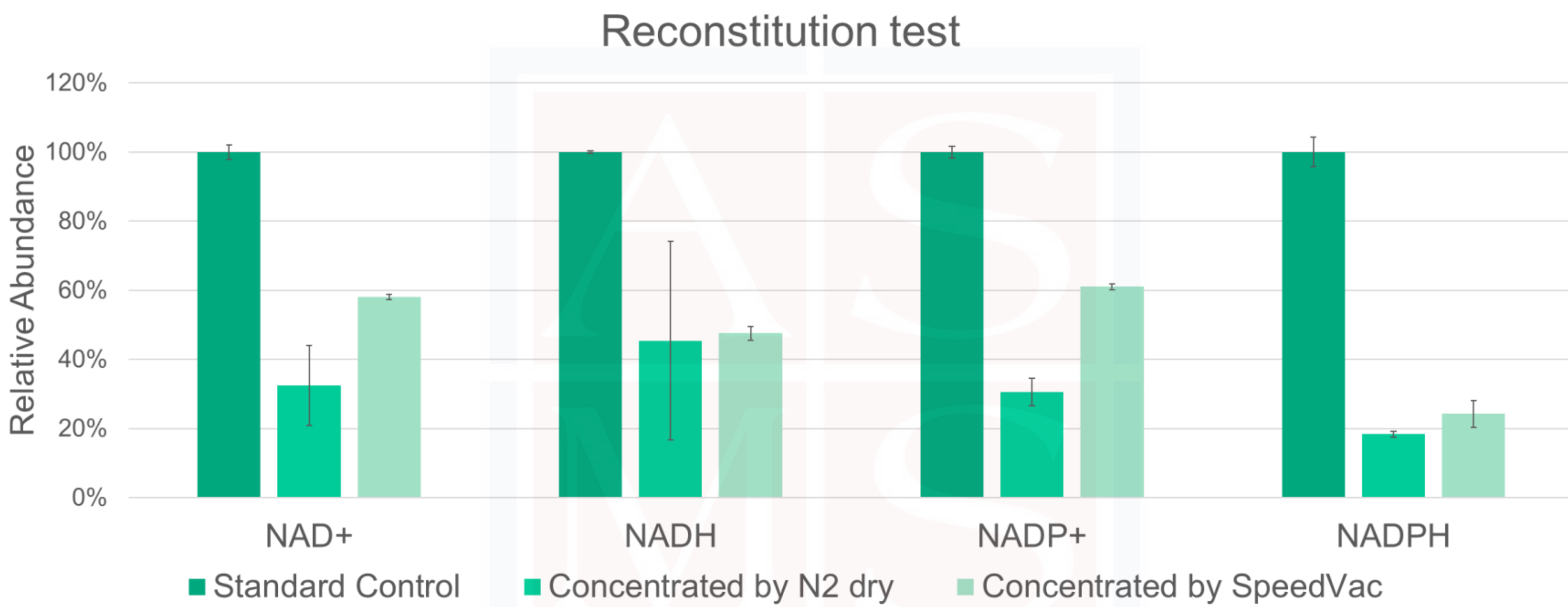
1. **Lyophillizor:**
   provides opportunity to normalize by mass

2. **N2 evaporator**

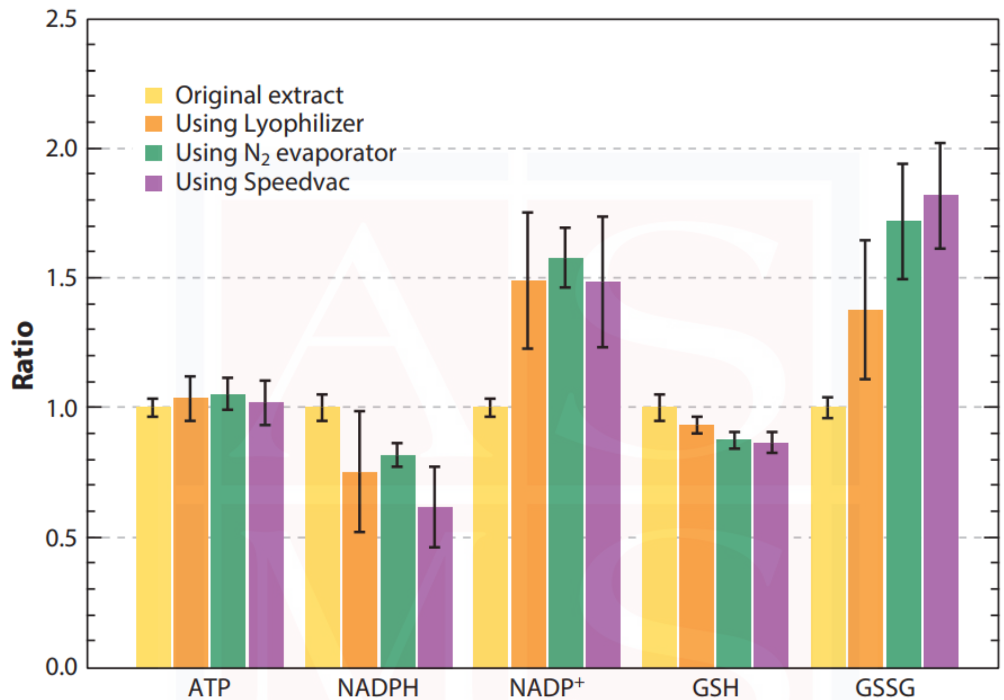
3. **SpeedVac**

Methods for drying samples down

Degradation of redox pairs is observed after metabolite reconstitution



Methods for drying samples down



Beware of redox-active species

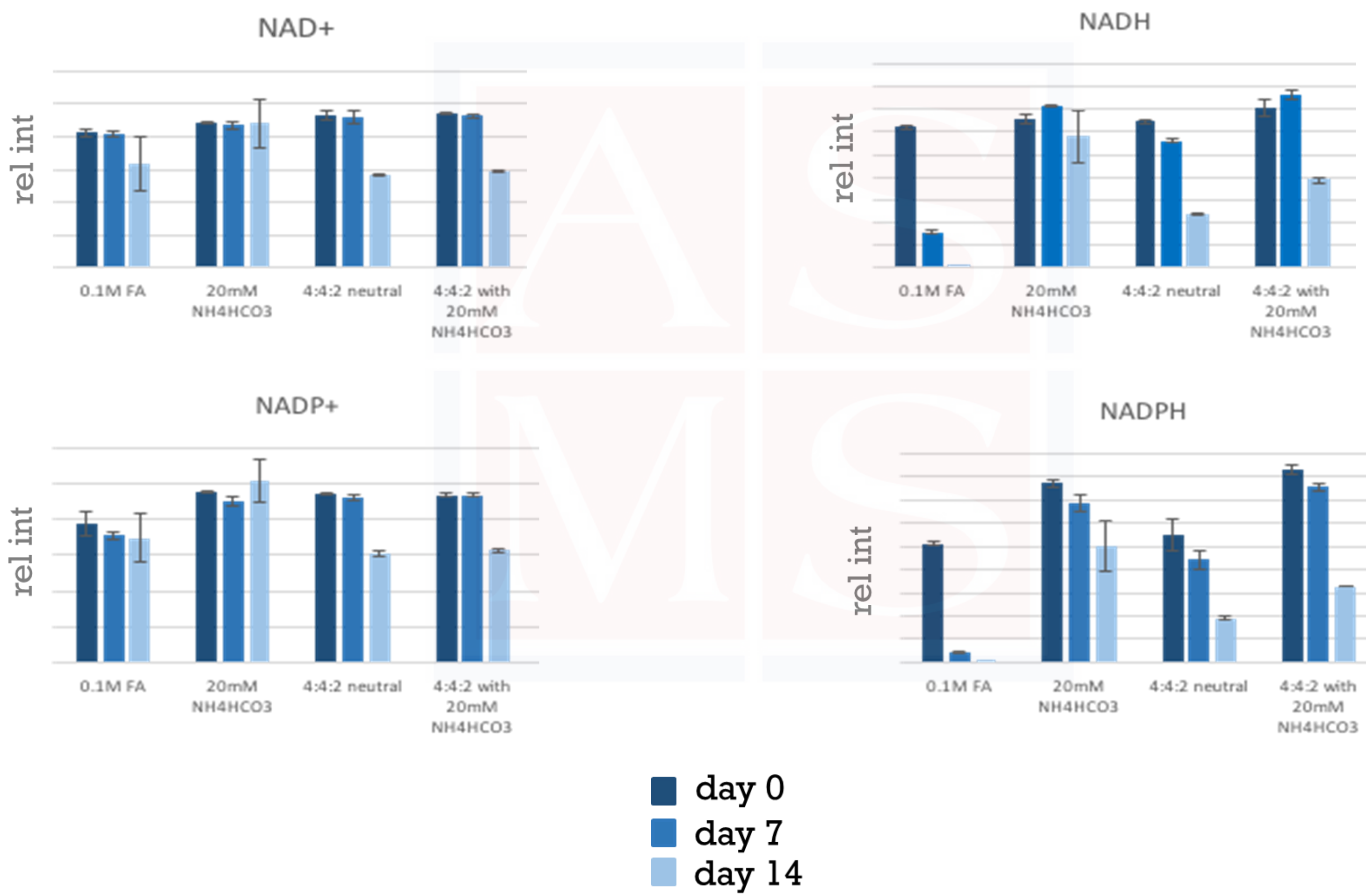
30 mg of mouse liver extracted with 40:40:20 acn:methanol:water. LC/MS signals compared before and after drying.

Most metabolites unaffected.

Lu et al., Annu Rev Biochem 2017. 86:277

Methods for drying samples down

Degradation of redox pairs is observed after long-term storage in solution



From intact samples to metabolomic analysis

purify/wash/ normalize sample

quench metabolism

dry samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

Homogenization: direct contact?

- Rate of extraction is inversely proportionate to particle size of sample

Direct contact methods: mortar and pestle, probe sonicator

Strong homogenization, but risk of contamination and variability

No direct contact methods: freeze-thaw cycles, bath sonciator.

Weak homogenization, but reduced risk of contamination and variability

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/ normalize sample

quench metabolism

dry samples

homogenize

extract

concentrate

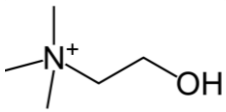
Major objectives of the extraction step

1.) Remove macromolecules

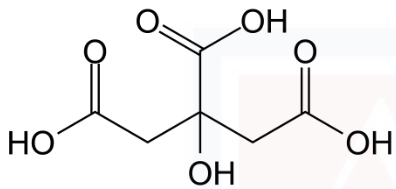
2.) Comprehensive coverage (for untargeted metabolomics)

3.) Quantitative accuracy and reproducibility

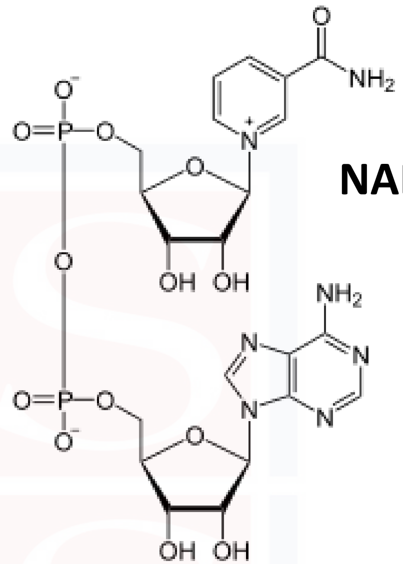
The challenge: physicochemical diversity



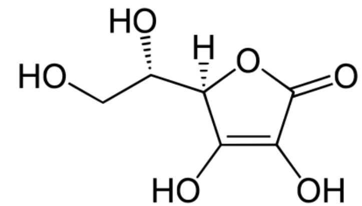
choline



citric acid

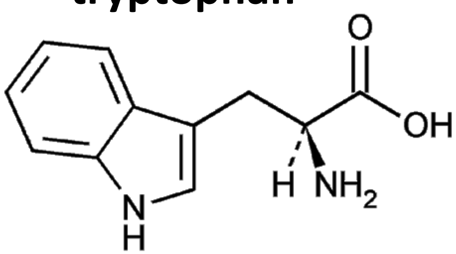


NAD+

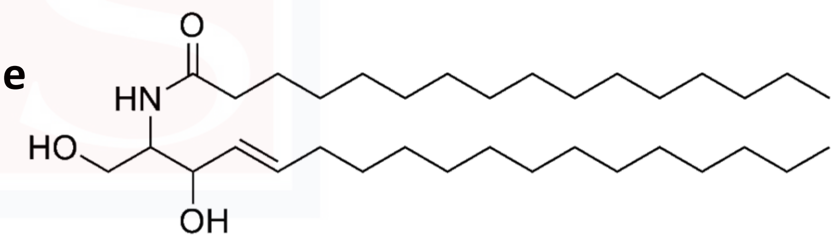


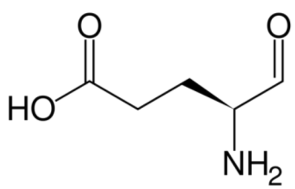
ascorbic acid

tryptophan

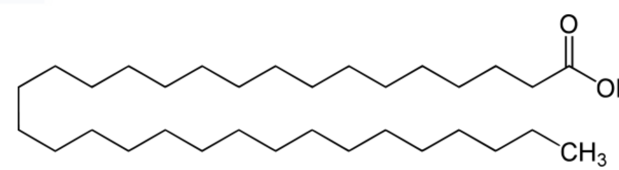


C16 ceramide





glutamate



lacceroic acid

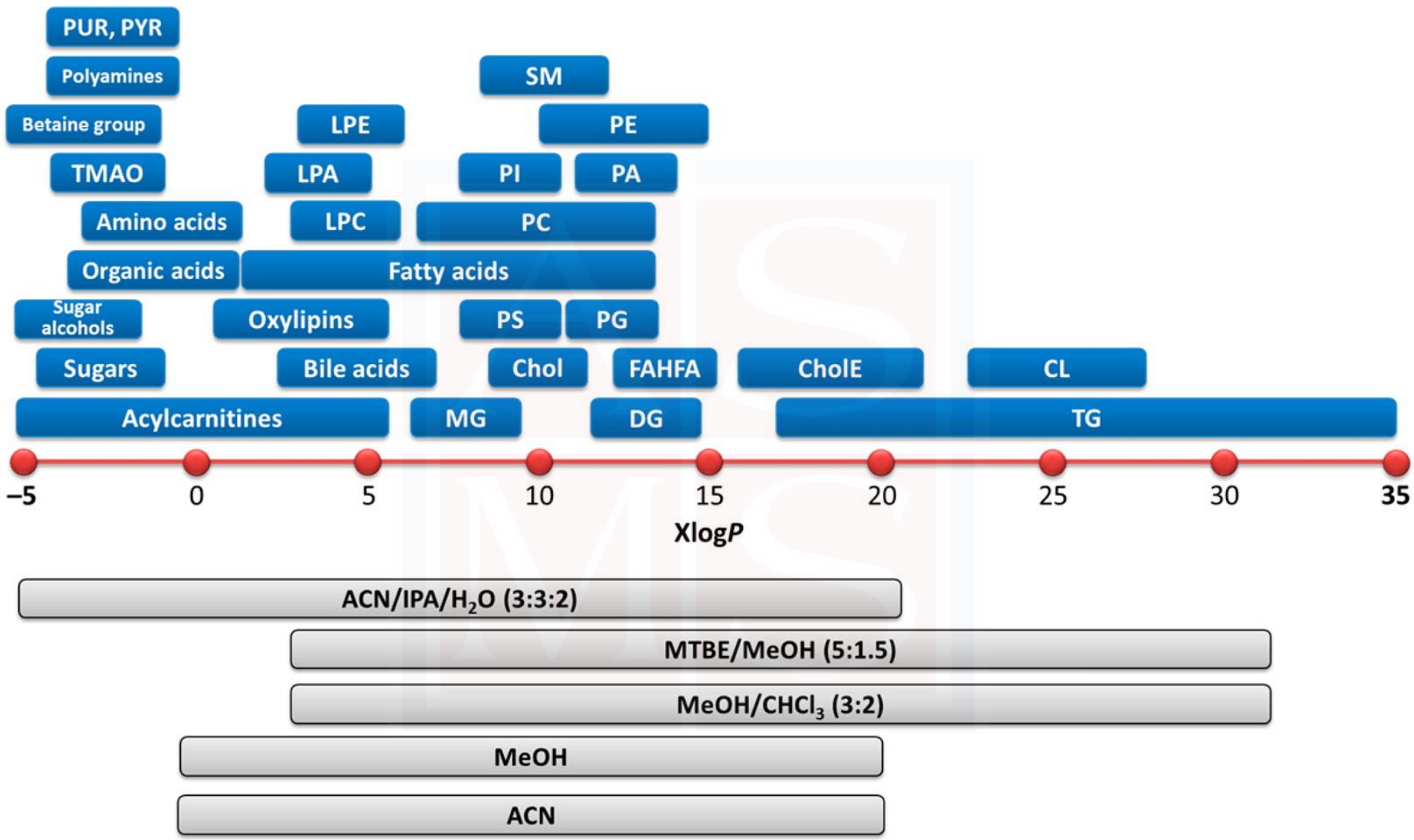
Major factors influencing coverage and reproducibility

- Chemical stability and reactivity
  * for example, degradation of thermo instable metabolites
  * consider reactivity of solvent, temperature, pH

- Solubility
  * for example, polar metabolites in organic solvent
  * can be assessed by logP

Logarithm of the partition coefficient (logP):
the ratio of the concentration of an un-ionized metabolite in octanol to the concentration of the un-ionized metabolite in water

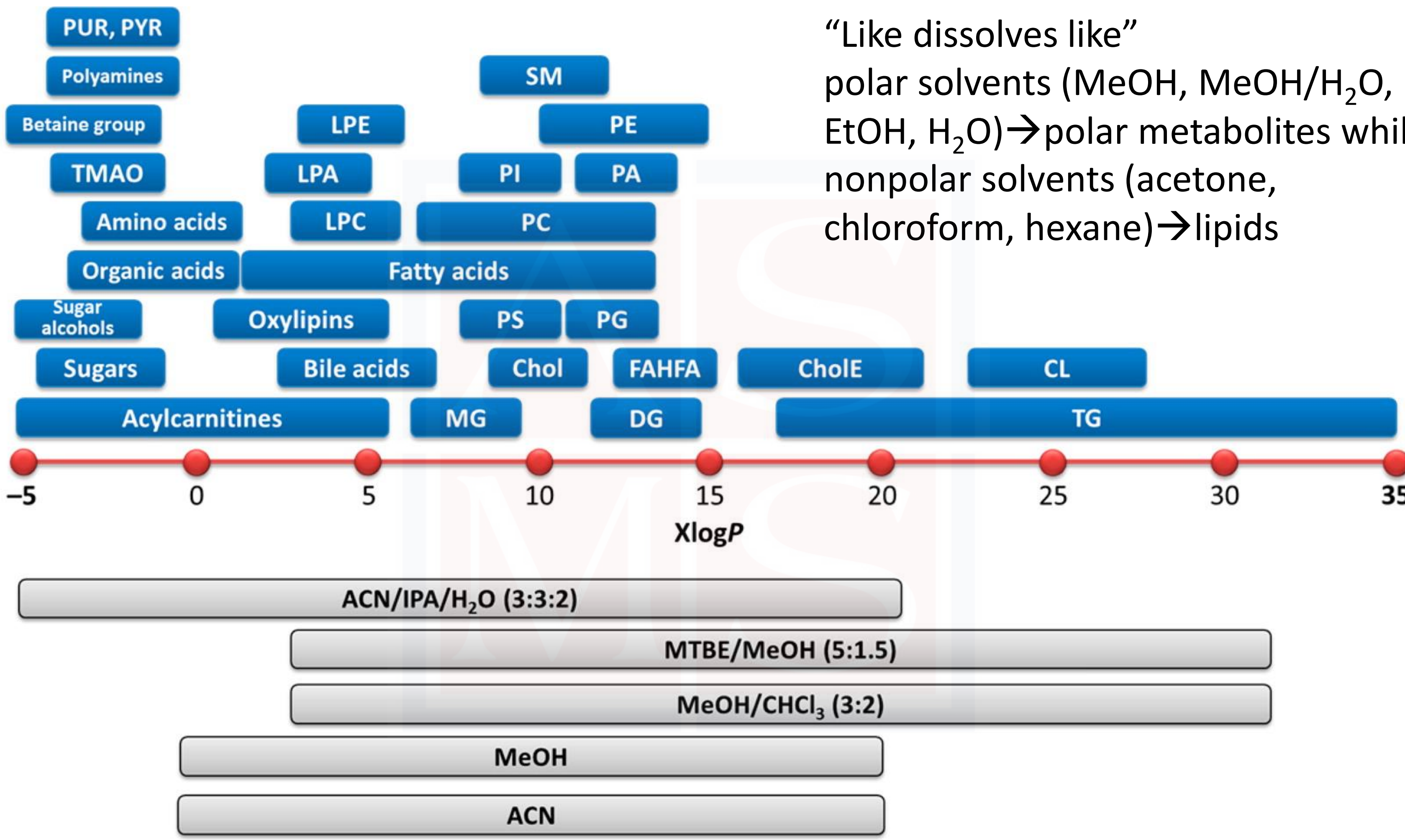
Solvent choices are critical



Cajka and Fiehn, Anal Chem 2016

Cer, ceramides; Chol, cholesterol; CholE, cholesteryl esters; CL, cardiolipins; DG, diacylglycerols; FAHFA, fatty acid esters of hydroxyl fatty acids; LPA, lysophosphatidic acids; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamines; MG, monoacylglycerols; PA, phosphatidic acids; PC, phosphatidylcholines; PE, phosphatidylethanolamines; PG, phosphatidylglycerols; PI, phosphatidylinositols; PS, phosphatidylserines; PUR, purines; PYR, pyrimidines; SM, sphingomyelins; TG, triacylglycerols; TMAO, trimethylamine N-oxide.

Solvent choices are critical



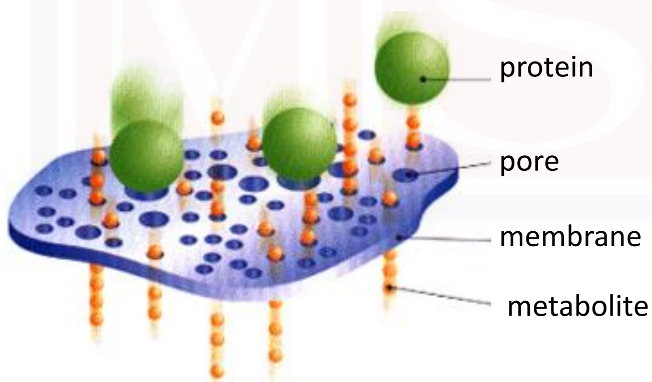
"Like dissolves like"
polar solvents (MeOH, MeOH/H_2O, EtOH, H_2O)→polar metabolites while nonpolar solvents (acetone, chloroform, hexane)→lipids

Cajka and Fiehn, Anal Chem 2016

Cer, ceramides; Chol, cholesterol; CholE, cholesteryl esters; CL, cardiolipins; DG, diacylglycerols; FAHFA, fatty acid esters of hydroxyl fatty acids; LPA, lysophosphatidic acids; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamines; MG, monoacylglycerols; PA, phosphatidic acids; PC, phosphatidylcholines; PE, phosphatidylethanolamines; PG, phosphatidylglycerols; PI, phosphatidylinositols; PS, phosphatidylserines; PUR, purines; PYR, pyrimidines; SM, sphingomyelins; TG, triacylglycerols; TMAO, trimethylamine N-oxide.

Variables to manipulate during extraction

- **Solvents** (see above)
- **Temperature** (consider thermo stability and efficiency)
- **pH** (acid stable vs base stable compounds; e.g., acid hydrolyzes sugars)
- **Molecular-weight filter** (metabolite loss, reproducibility)

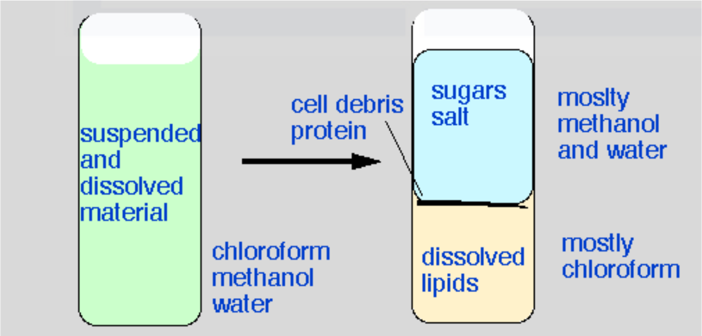


Different strategies for introducing solvent

<u>Monophasic extractions</u>:
1:1 methanol →sonicate, vortex, homogenize (SVH) →supernatant (metabolites)

<u>Sequential monophasic extractions</u>:
1:1 chloroform:methanol→sonicate, vortex, homogenize→supernatant (nonpolar metabolites)→subject precipitate (polar metabolites) to 1:1 methanol:water→ supernatant (polar metabolites)

<u>Biphasic extractions</u>:
1:1 chloroform:methanol (1 phase)→sonicate, vortex, homogenize→1 water→1:1:1 chloroform:methanol:water (upper phase water - polar metabolites; lower phase chloroform - nonpolar metabolites)



Different strategies for introducing solvent

Method 1:

1. Add 200-400uL hot MeOH (80˚C) to the sample
2. Incubate for 5 min at 80˚C (oven).
3. Centrifuge at 13,000 rpm for 15 min at 4˚C
4. Keep supernatant.
5. Add 100-200uL hot MeOH (80 ˚C) to the pellet.
6. Incubate for 5 min at 80˚C
7. Centrifuge at 13,000 rpm for 15 min at 4˚C
8. Pool supernatants and analyze

Method 2

1. Add 200-400ul of cold 5% MPA/1mM EDTA/0.1% FA to the sample
2. 1 min in liquid nitrogen
3. Thaw
4. Sonicate 5 min
5. centrifuge 15 min at 13.000rpm
6. Analyze supernatant

Note: most metabolomic extractions have been optimized by either using a small set of targeted compounds or by counting features

Method 3:

1. Add 300-400 uL hot 80% MeOH/20% 1mM HEPES, 1mM EDTA (pH 7.0) (80˚C)
2. 5 min at 80 ˚C
3. Vortex
4. 1 min in liquid nitrogen
5. Thaw
6. 1h at -80 ˚C
7. centrifuge 15 min at 13.000rpm
8. Analyze supernatant

Method 4:

1. Add 400ul of cold acetone to the sample.
2. 1 min in liquid nitrogen
3. thaw
4. sonicate 10 min
5. 1h at -20 C
6. centrifuge 15 min at 13.000rpm
7. Keep supernatant
8. Add 200ul of cold MetOH/water/formic acid (86.5/12.5/1.0) to the pellet
9. sonicate 15 min
10. 1h at -20 C
11. centrifuge 15 min at 13.000rpm
12. Keep supernatant and pool it with the first (step 8). Discard pellet (proteins).
13. Dry out (SpeedVac) supernatant.
14. Re-dissolve sample in 100ul of 95% ACN/water. Sonicate tubes for 10min and leave them 1h at 4 C.
15. Centrifuge 10 min at 13.000rpm

Method 5:

1. Add 200-300 uL of cold 1mM HEPES/ 1mM EDTA (pH 7.0) to the sample
2. 1 min in liquid nitrogen
3. Sonicate 5 min in cold water
4. Vortex 10 sec (repeat 2-4 three times)
5. Centrifuge 15 min at 13.000rpm. Keep supernatant.
6. Centrifuge at 13,000 g with Microcon® YM-3.
7. Analyze the filtrate

Method 6:

1) Cell Lysing + Extraction: Add cold MeOH: ACN: Water 2:2:1 to the pellet/tissue
   i) Repeat 3 times
      a) Vortex (30s)
      b) Liquid N2 bath (1m)
      c) Allow to thaw in sonicator (10s)
      d) Bath sonicate at 25°C (10m)
   ii) Store samples at -20°C (1-2h or overnight)
   iii) Centrifuge at 14K RPM and 4°C
   iv) Transfer supernatant to new Eppendorfs

2.) Dry with speedvac
   i) No heating / Manual Run / Ramp 4

3) Add water:acetonitrile 1:1 to residue
   i) Repeat 2 times
      a) Bath sonicate at 25°C
      b) Vortex 1m
   ii) Store samples at 4°C (1h or overnight)
   iii) Centrifuge at 14K RPM and 4°C

4) Transfer supernatant to LC vials

5) Store supernatant at -80°C in LC vials for MS analysis

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

homogenize

extract

concentrate

From intact samples to metabolomic analysis

purify/wash/
normalize
sample

quench
metabolism

dry
samples

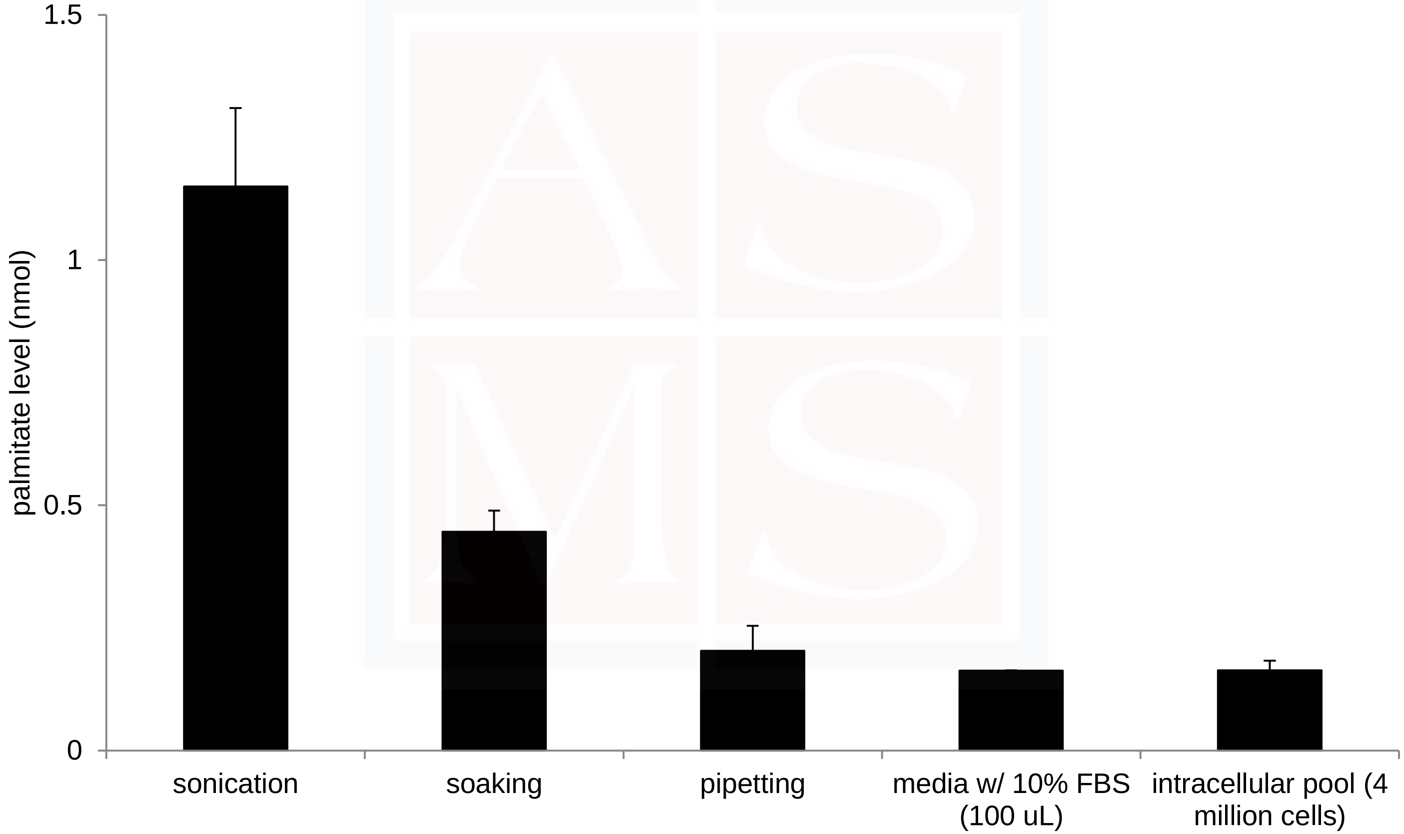
homogenize

extract

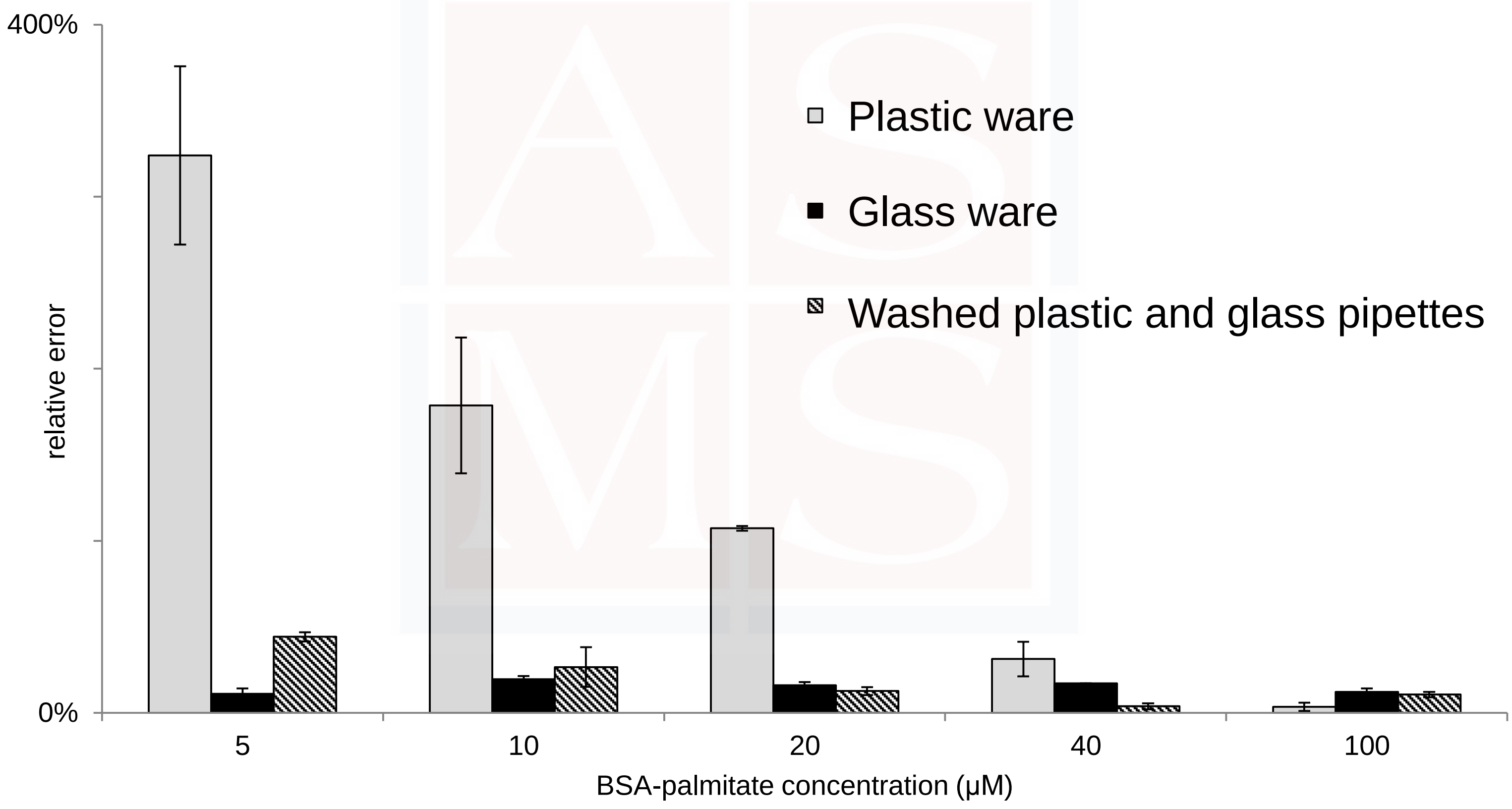
concentrate

Extraction usually dilutes metabolites, concentration uses same techniques as drying step

Beware of contamination
(e.g., due to plastic slip agents)



Beware of contamination
(e.g., due to plastic slip agents)



- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

*Separating metabolites**

by chromatography and MS

Why is chromatography important in metabolomics?

Required?
NMR, MALDI, FIA, shotgun lipidomics

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data

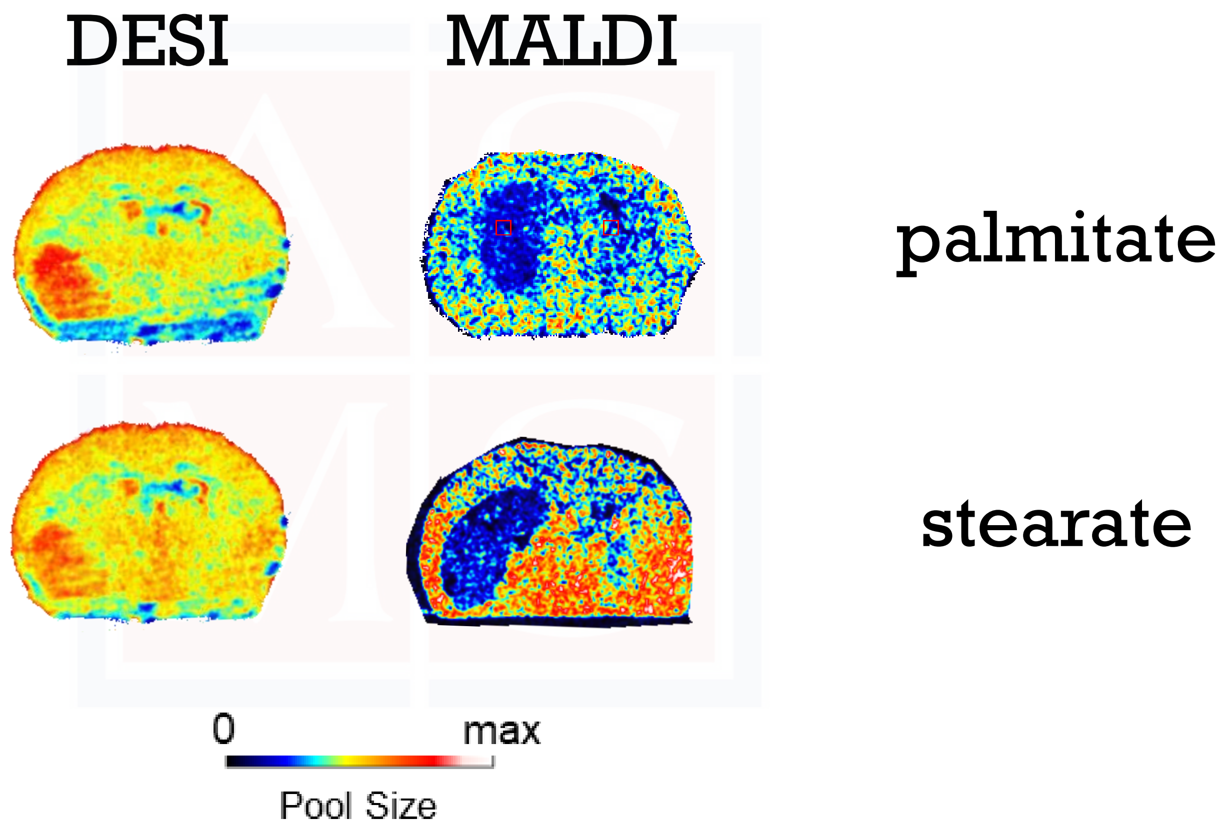
Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data

Matrix effects can be misleading



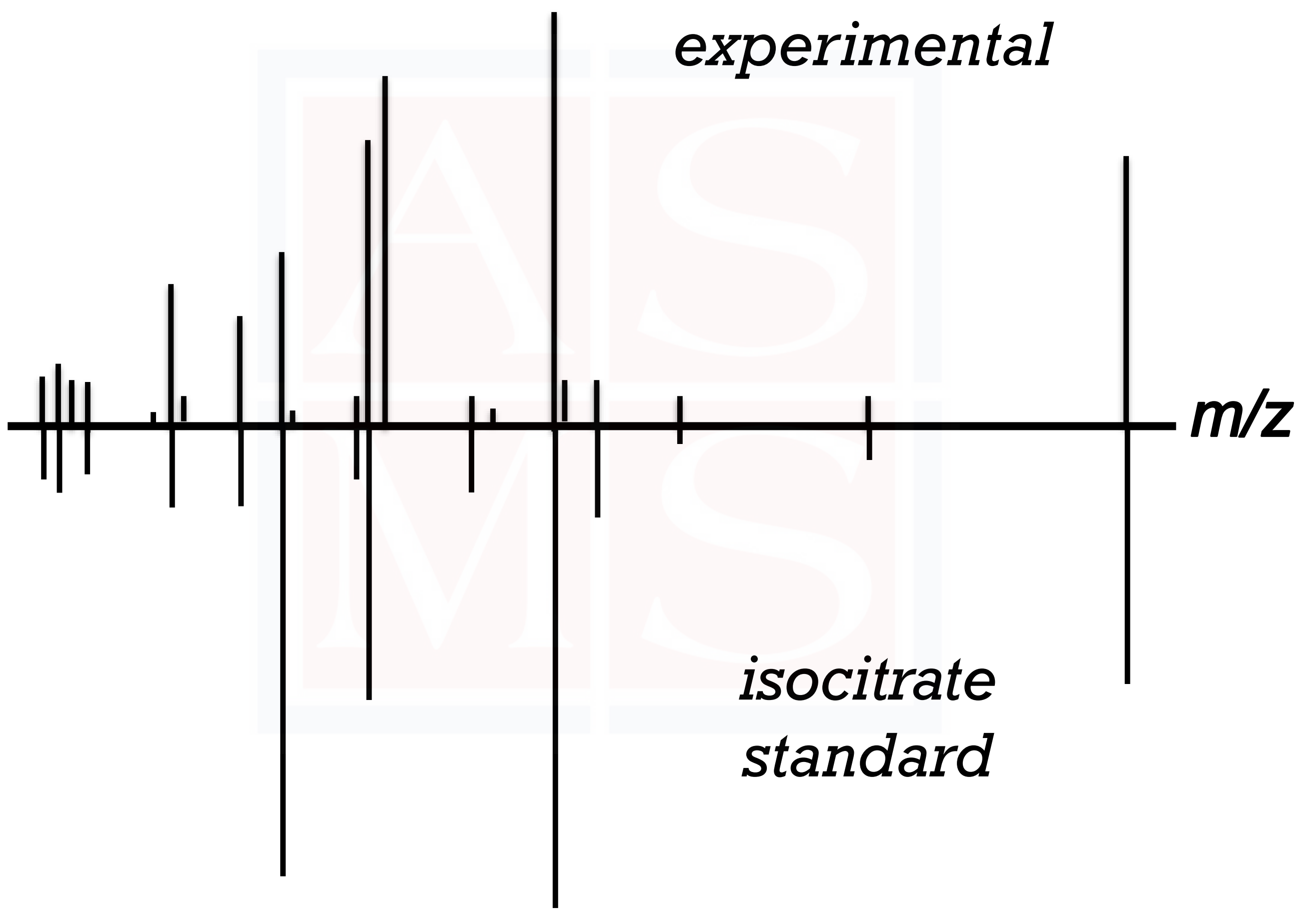
Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

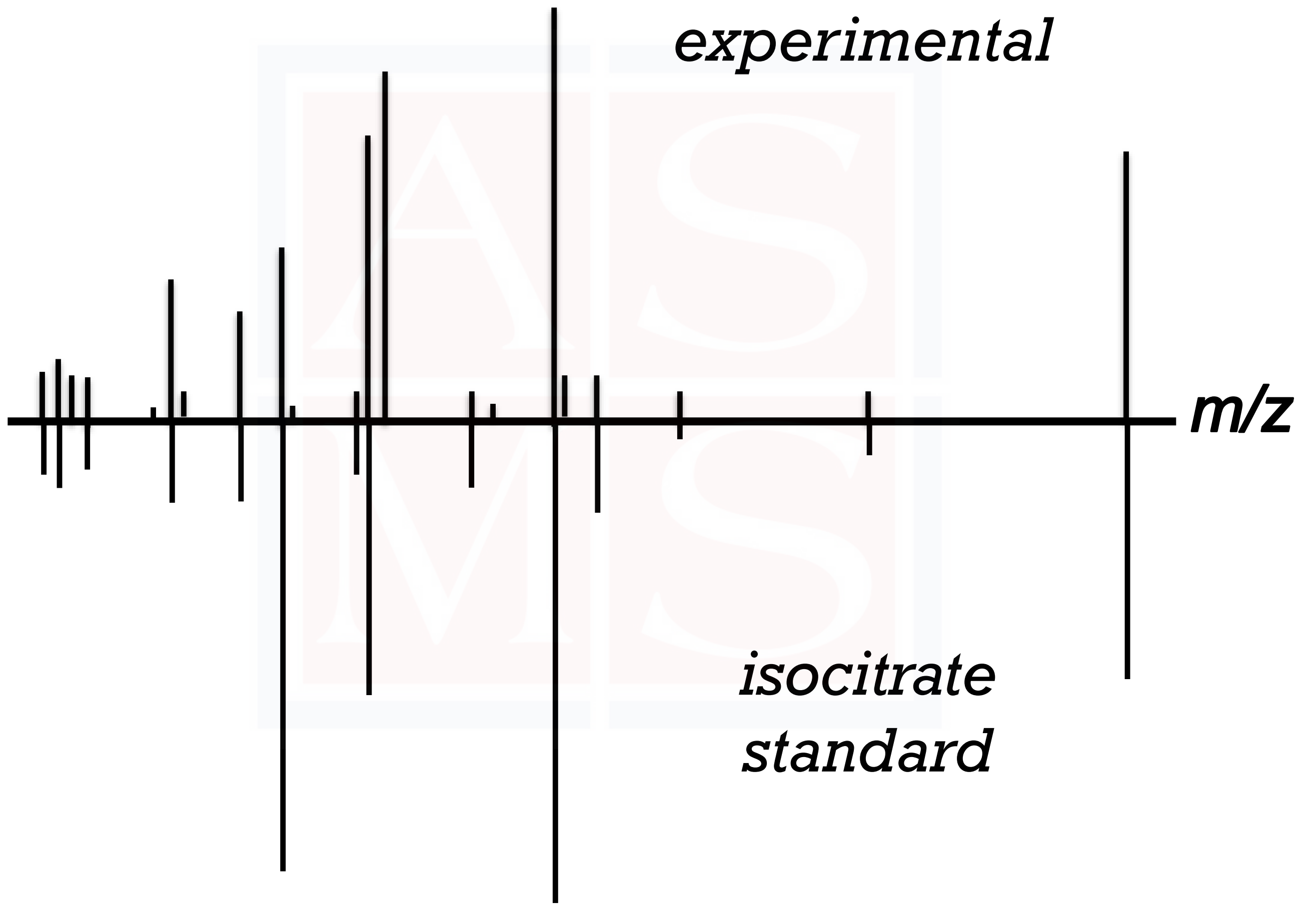
2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data

Metabolomic MS/MS Data

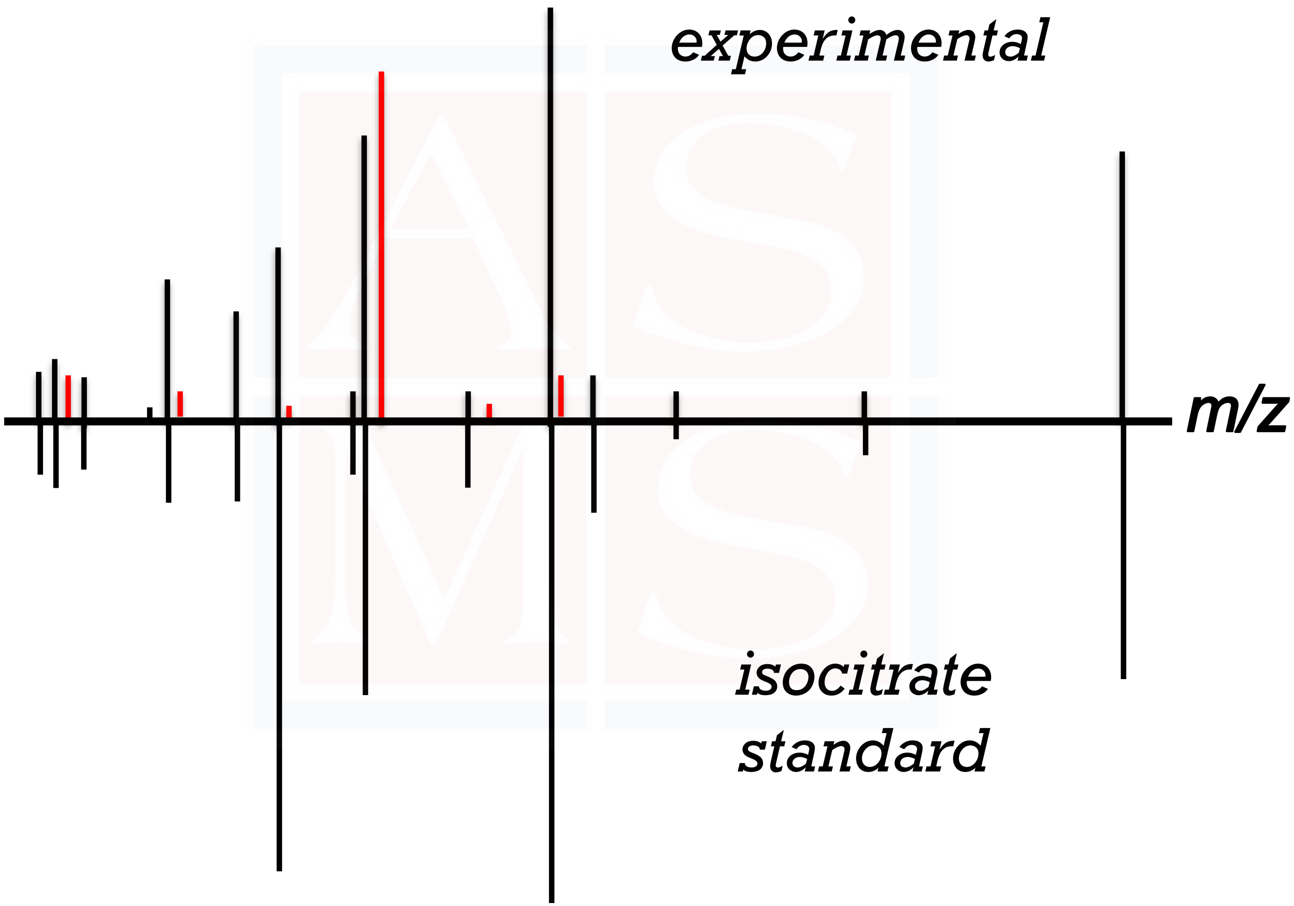


Metabolomic MS/MS Data



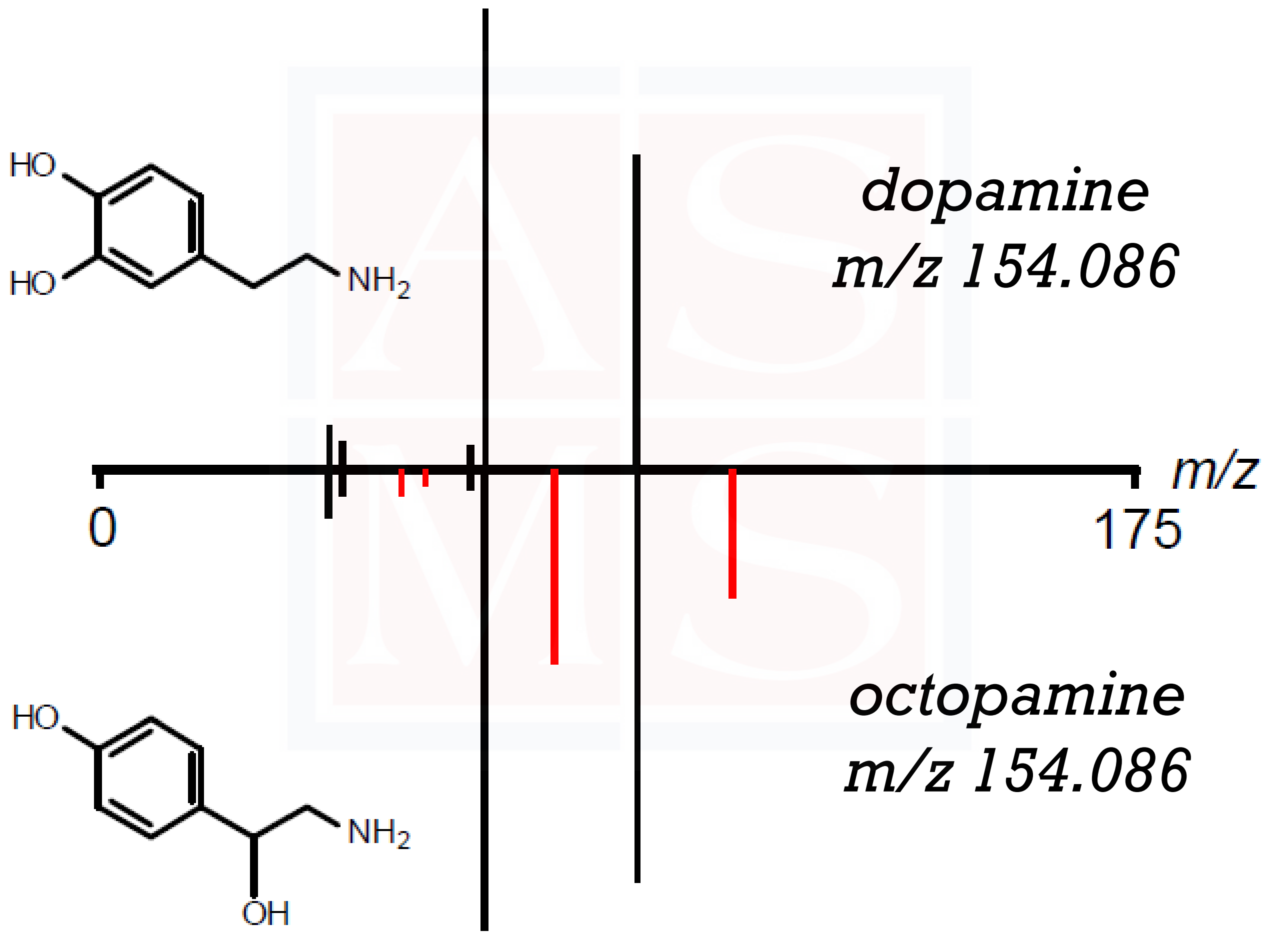
identification?

Metabolomic MS/MS Data

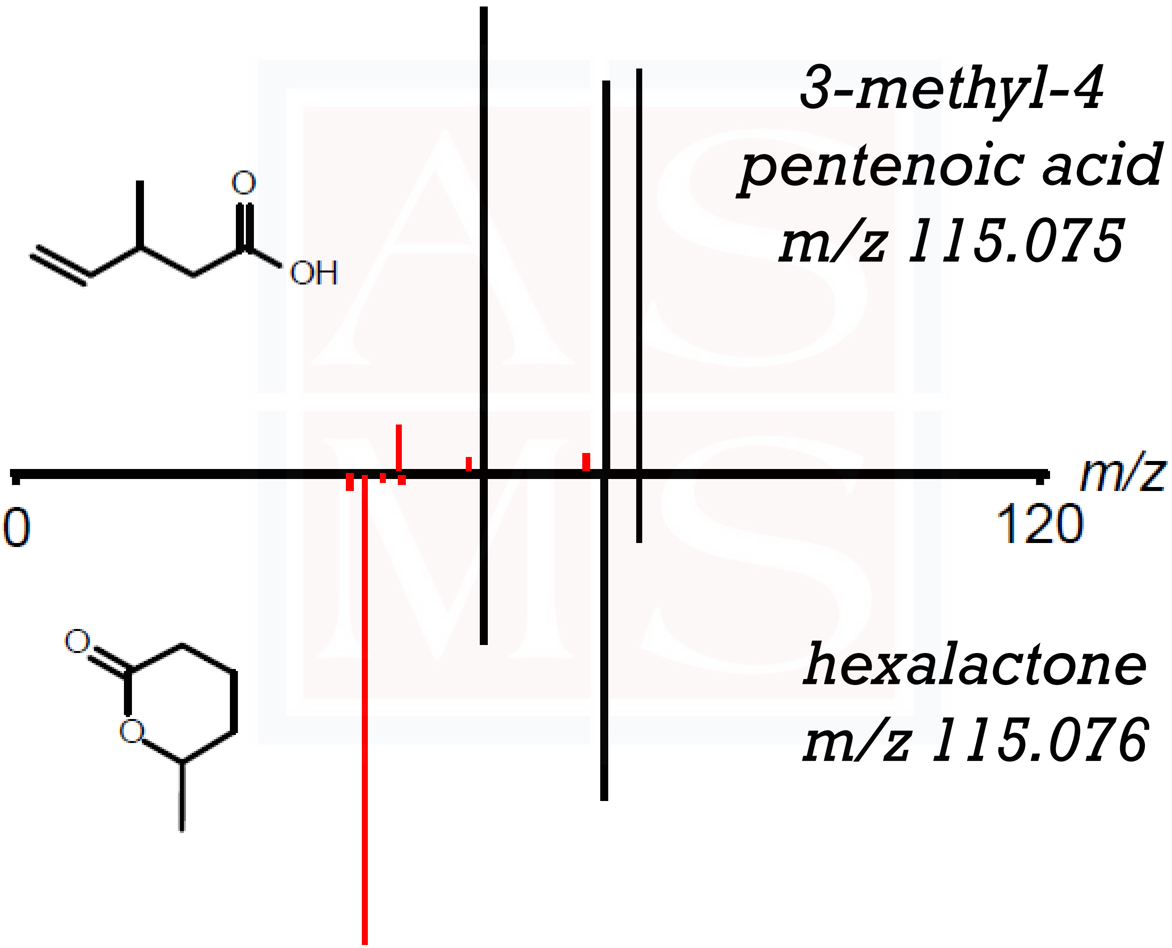


identification?

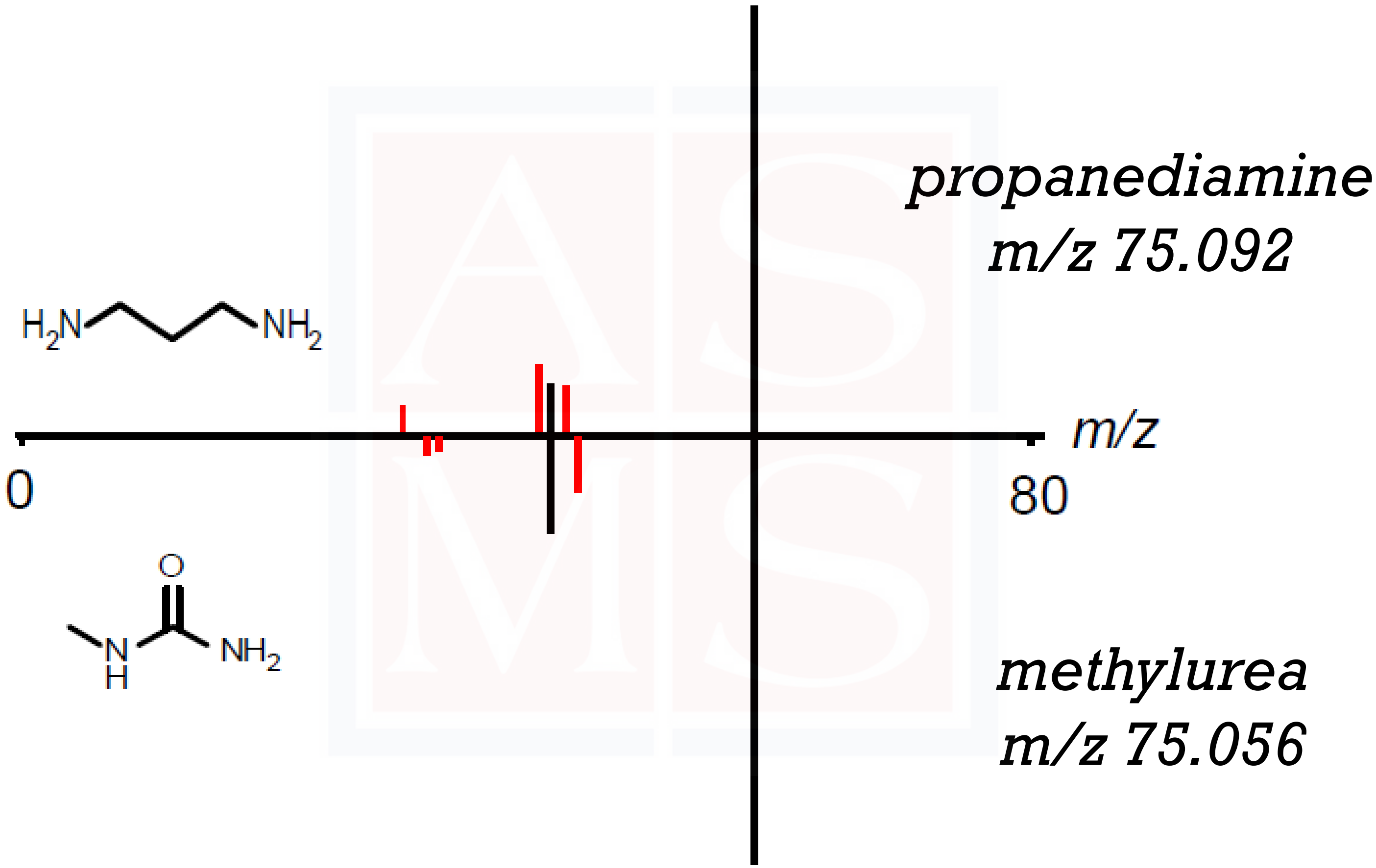
Metabolomic MS/MS Data



Metabolomic MS/MS Data



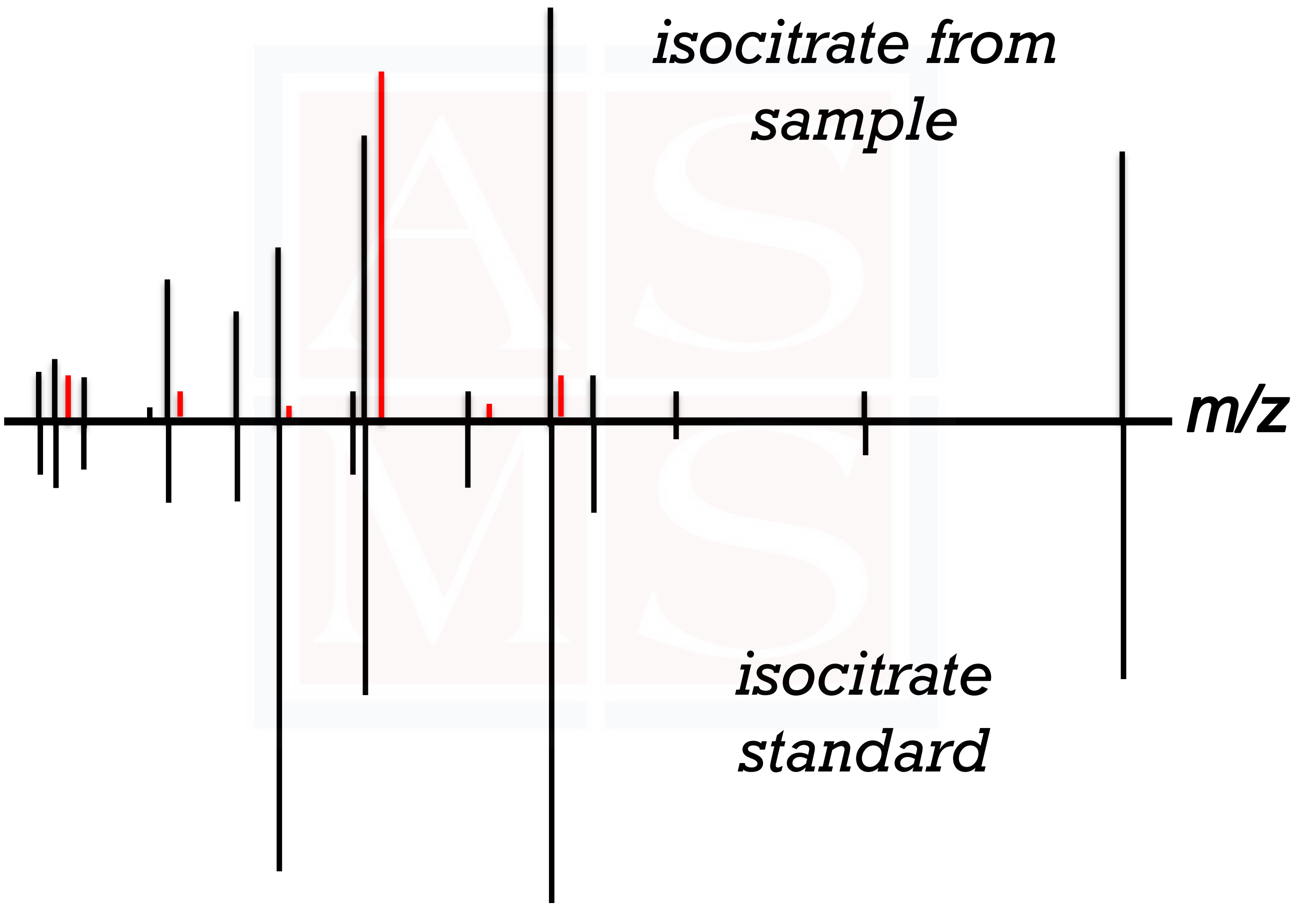
Metabolomic MS/MS Data



The Challenge of metabolomic MS/MS data

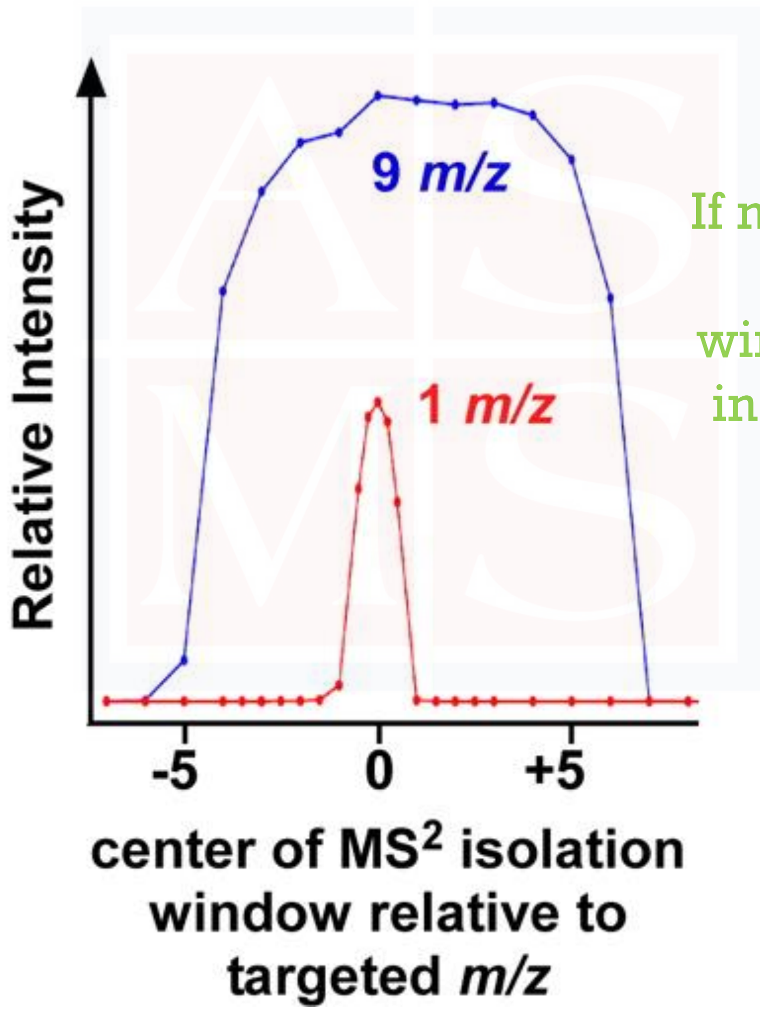
- *Most metabolite MS^2 are not predictable*
- *The size of the metabolome is unknown*
- *Therefore, metabolite MS^2 are not intuitive*
- *MS^2 data must match exactly to support ID*

Metabolomic MS/MS Data



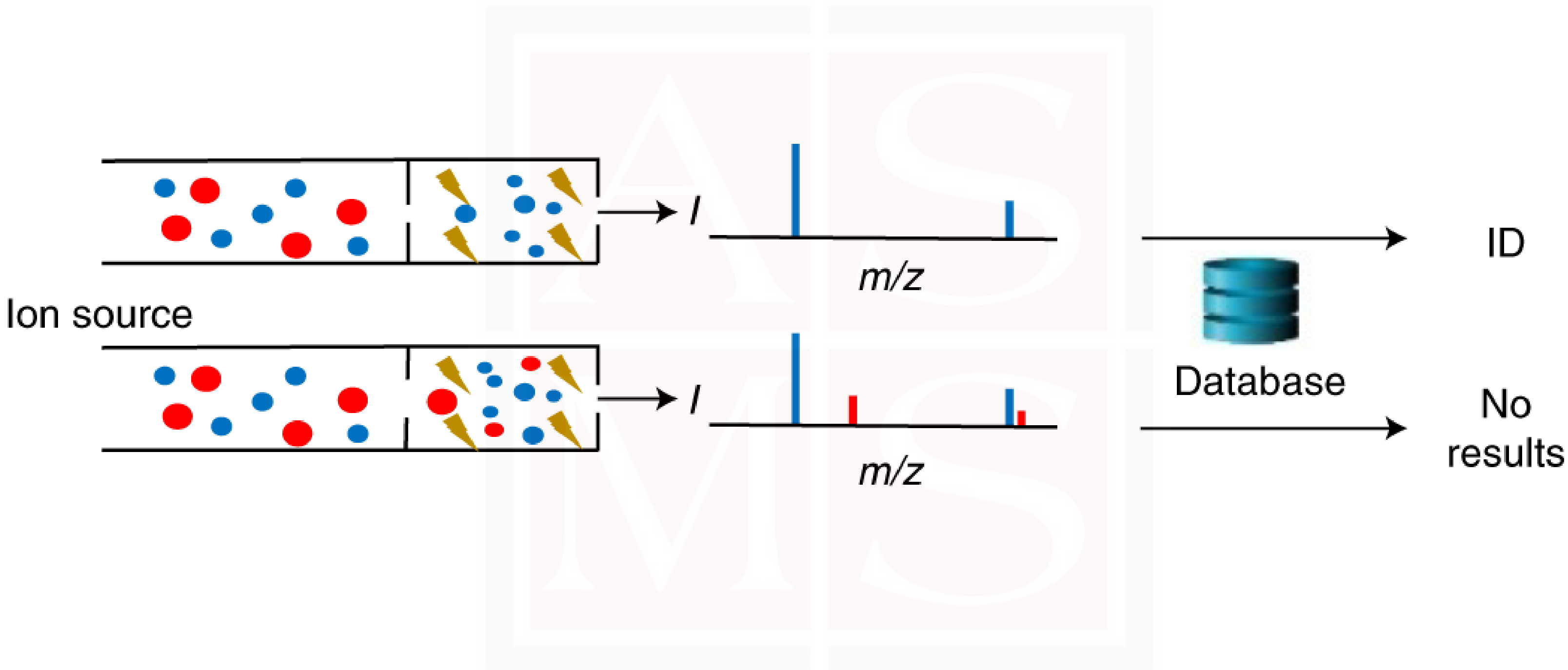
Why don't they match?

Source of MS/MS contamination (i.e., "chimeric" spectra")



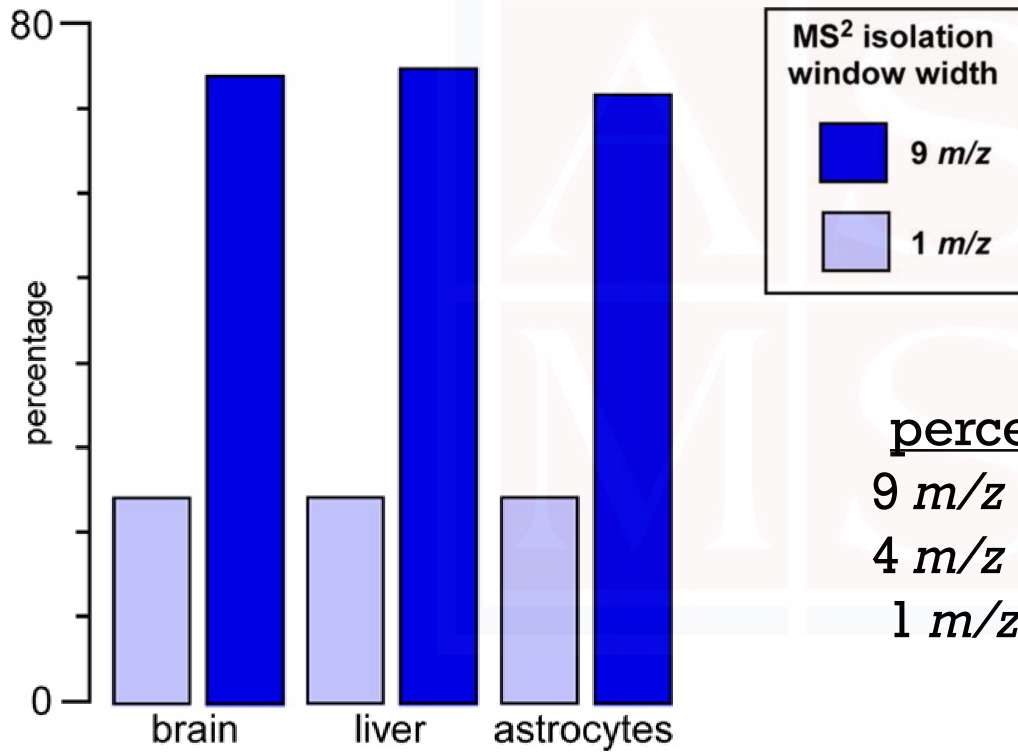
If metabolites coelute and are within isolation window, then they end up in collision cell together for fragmentation

Source of MS/MS contamination (i.e., "chimeric" spectra")



How frequently do compounds in metabolomic experiments fit these criteria?

percent contaminated scans



<u>percent contaminated scans</u>

9 *m/z* isolation window: ~73%

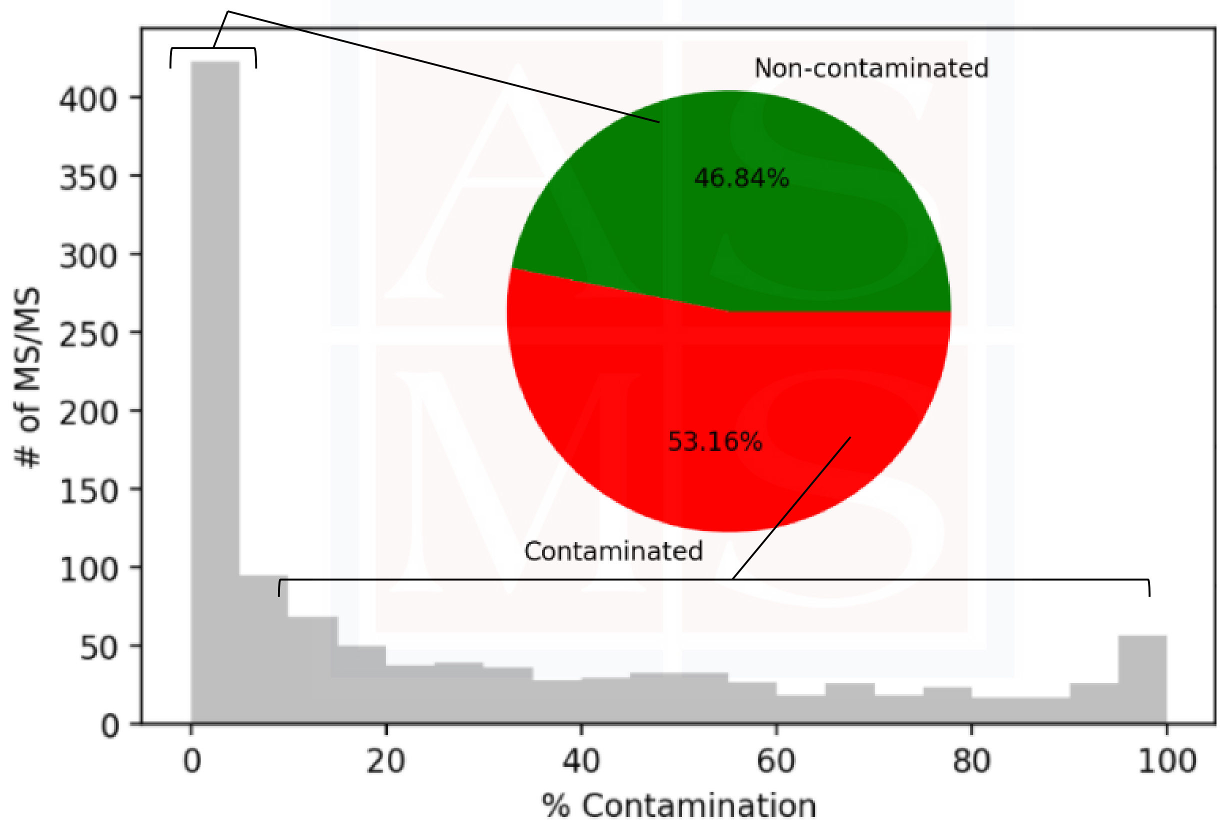
4 *m/z* isolation window: ~41%

1 *m/z* isolation window: 14%

Nikolskiy et al, Anal Chem 2013

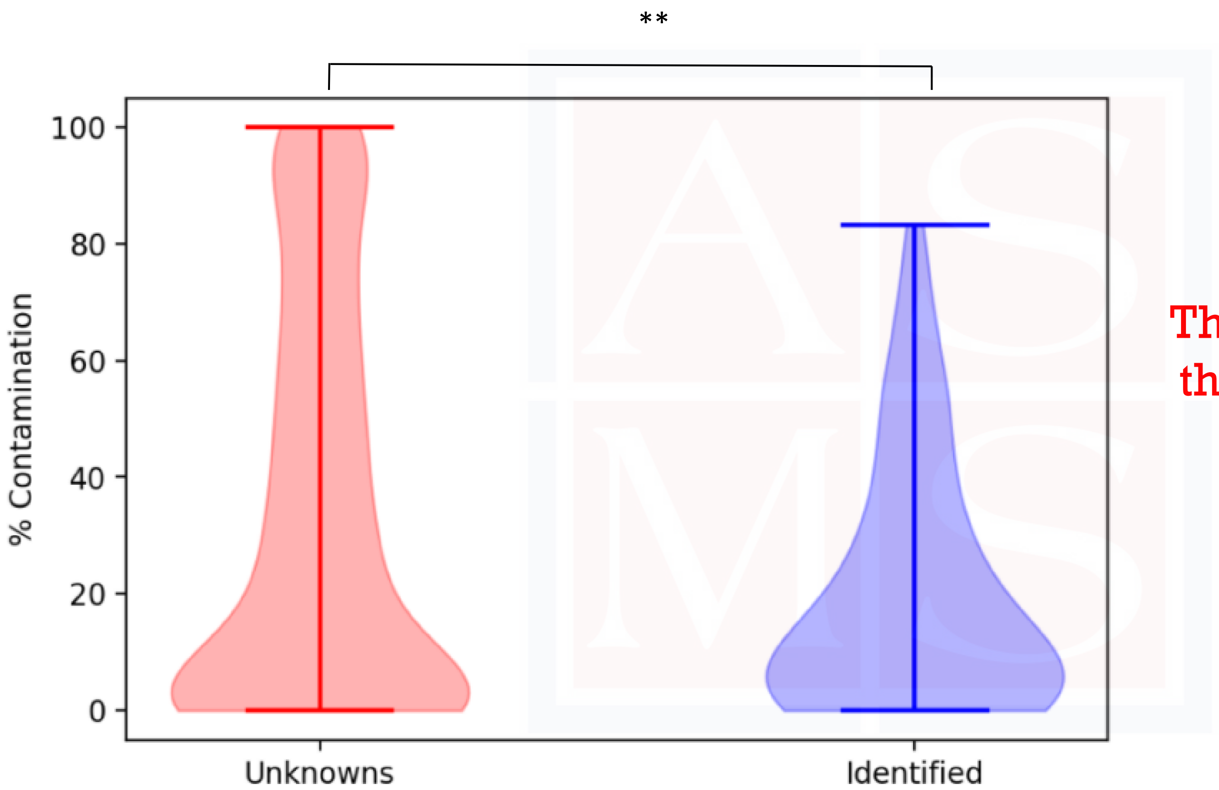
How frequently do compounds in metabolomic experiments fit these criteria?

human plasma, 1 m/z isolation window



Stancliffe et al, Nature Methods 2021
Lawson et al, Anal Chem 2017

How frequently do compounds in metabolomic experiments fit these criteria?



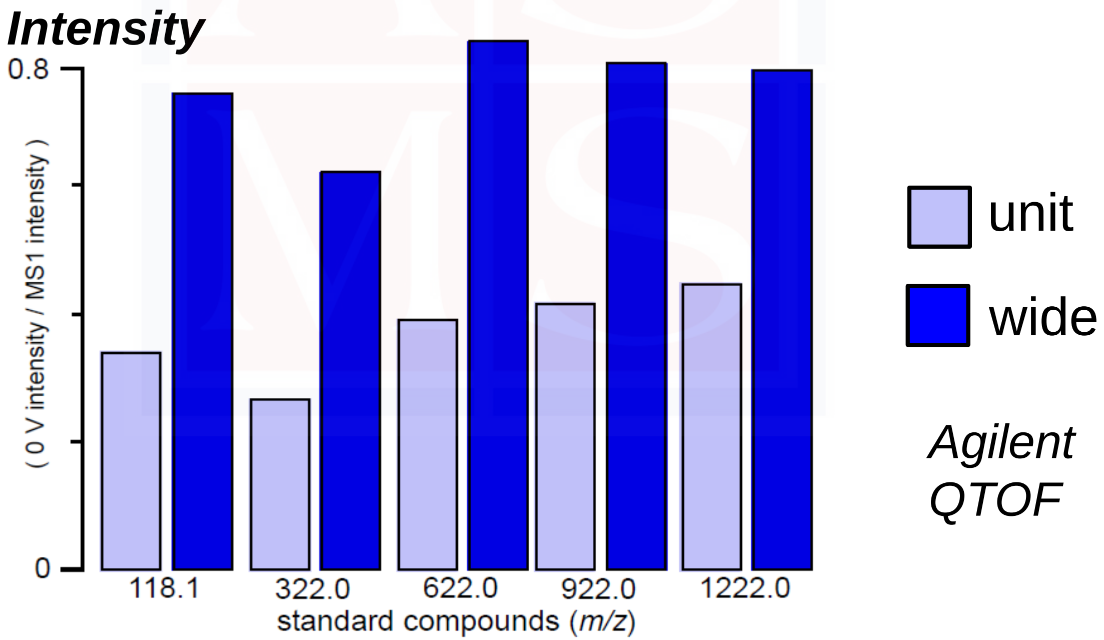
The more contamination, the less likely to identify

Stancliffe et al, Nature Methods 2021
Lawson et al, Anal Chem 2017

Why not always use narrow MS/MS isolation window in metabolomics?

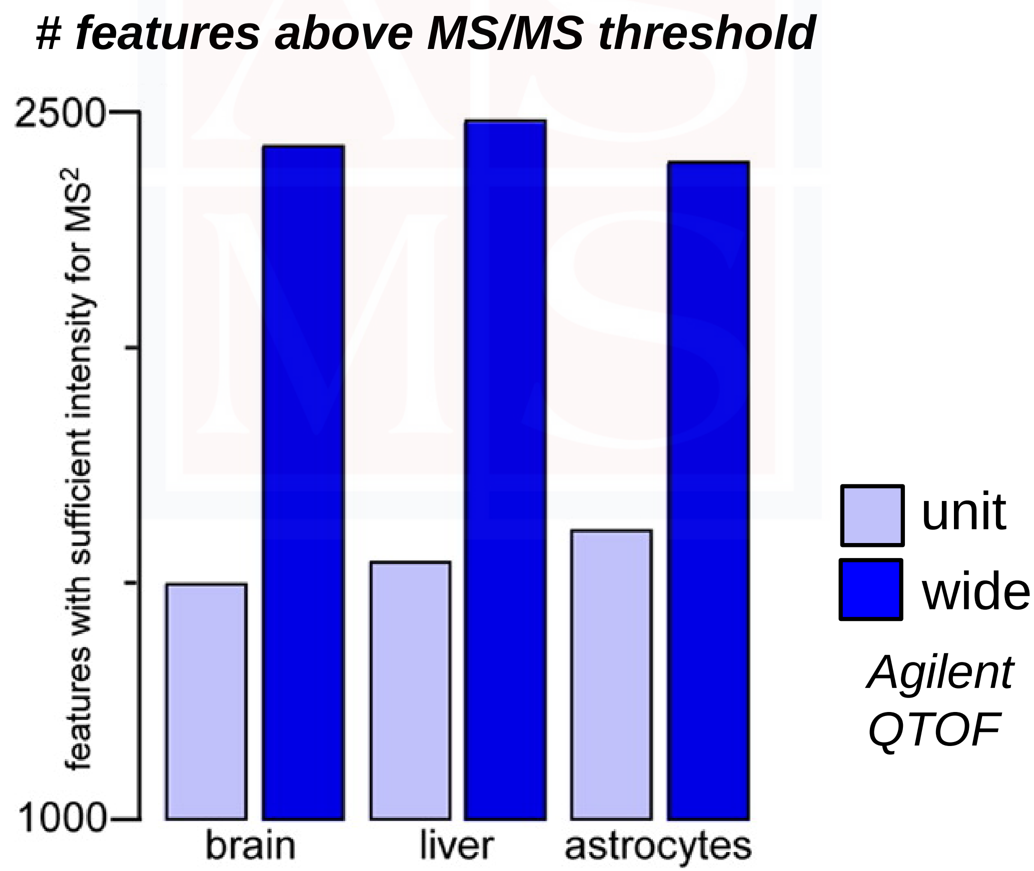
Why not always use narrow MS/MS isolation window in metabolomics?

- Misses isotope patterns
- Sensitivity vs. specificity

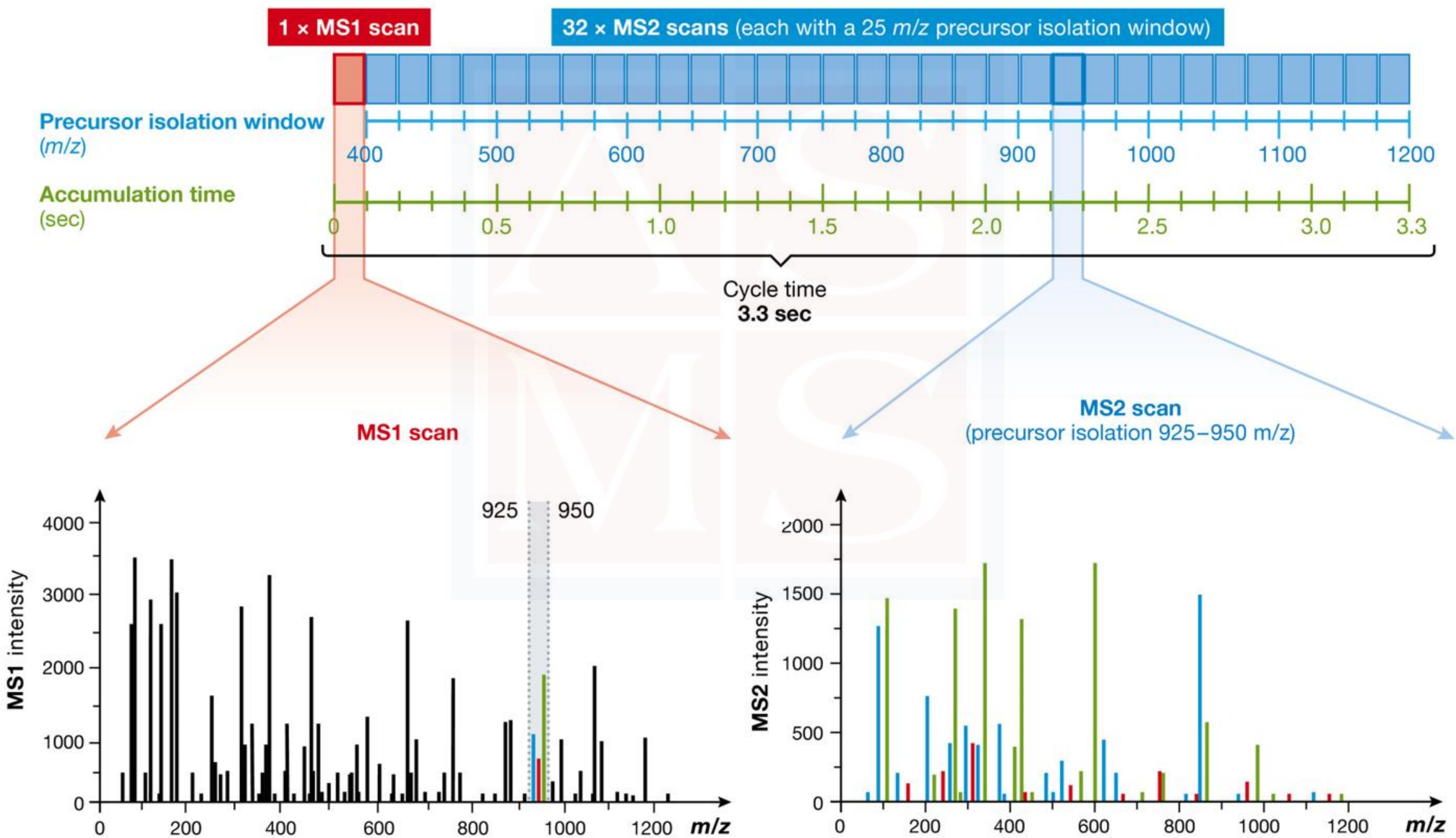


Why not always use narrow MS/MS isolation window in metabolomics?

- Fewer features accessible to MS/MS



SWATH-MS/MS acquisition in untargeted metabolomics

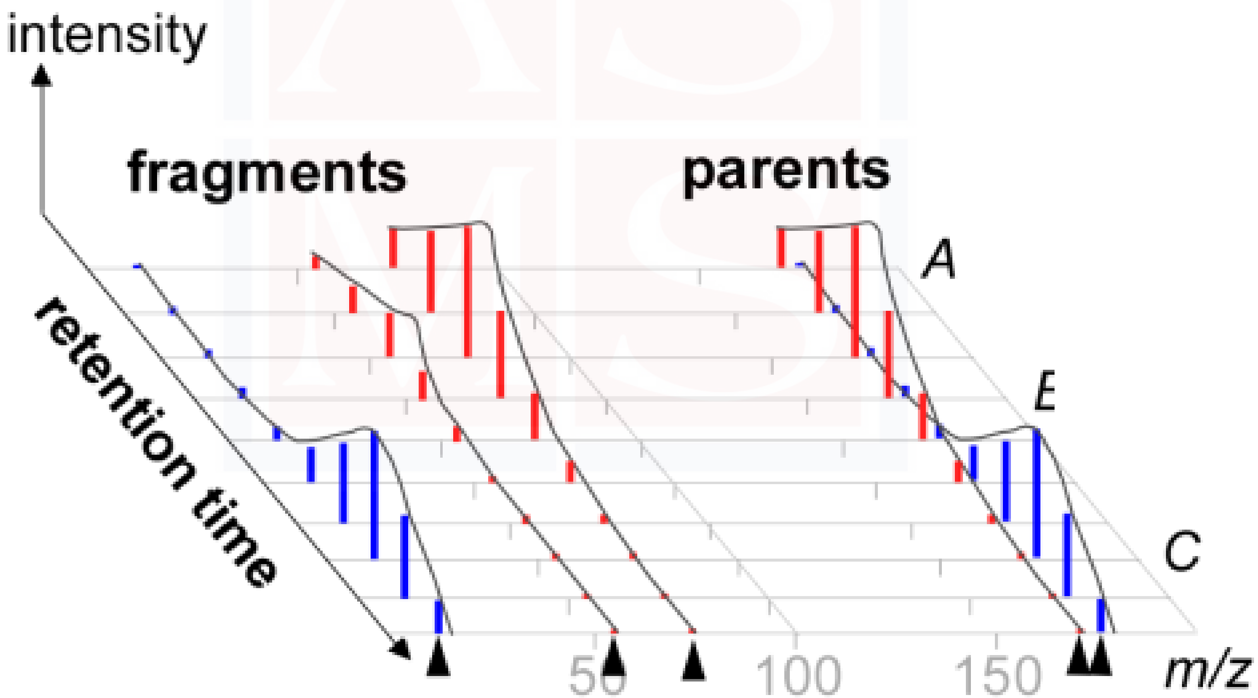


Ludwig et al., Molecular Systems Biology, 2018

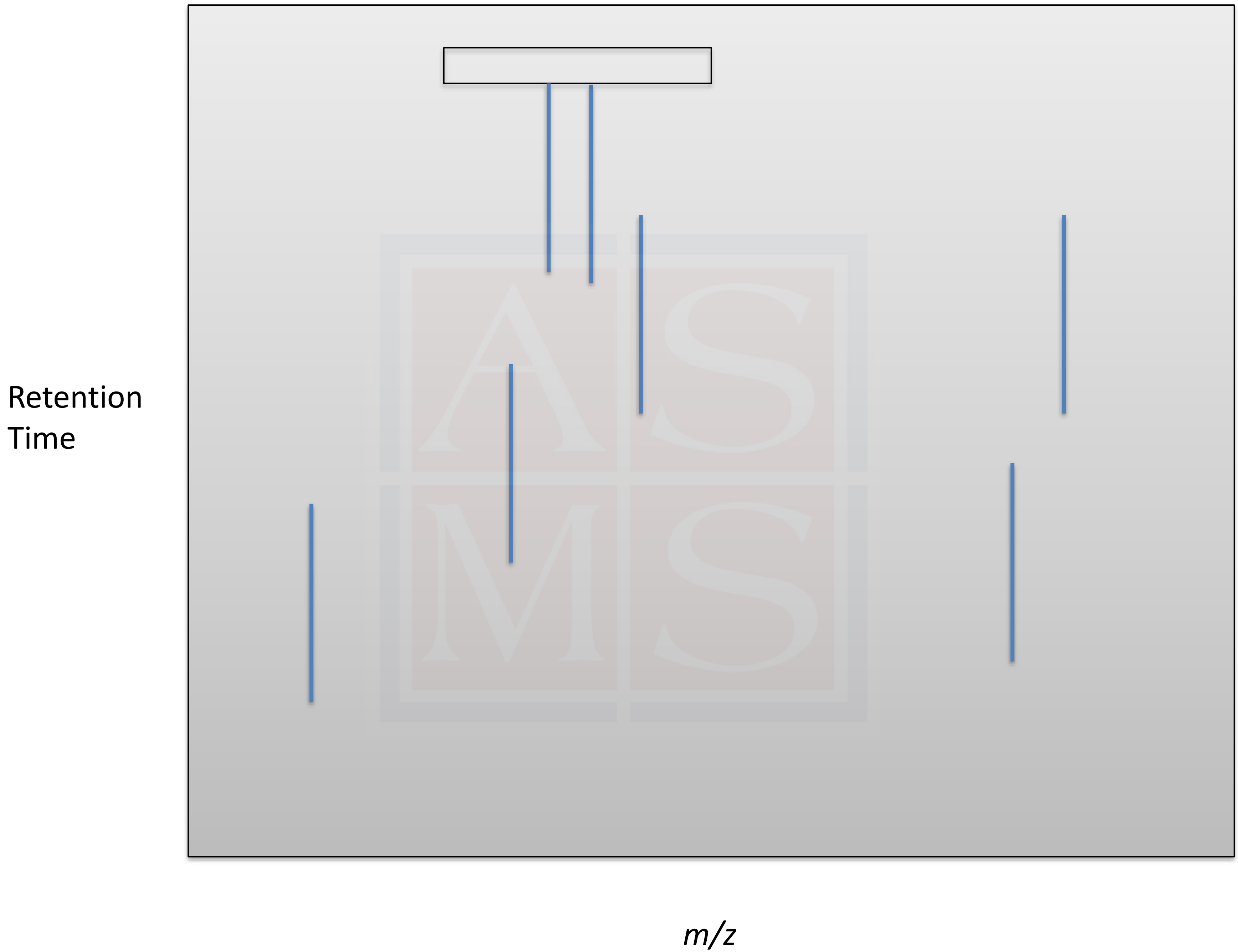
Untargeted metabolomic analysis with chimeric MS/MS data

decoMS2 and MSDIAL

Basic principle: RT deconvolution



Nikolskiy et al., Anal Chem 2013
Tsugawa et al., Nature Methods, 2015

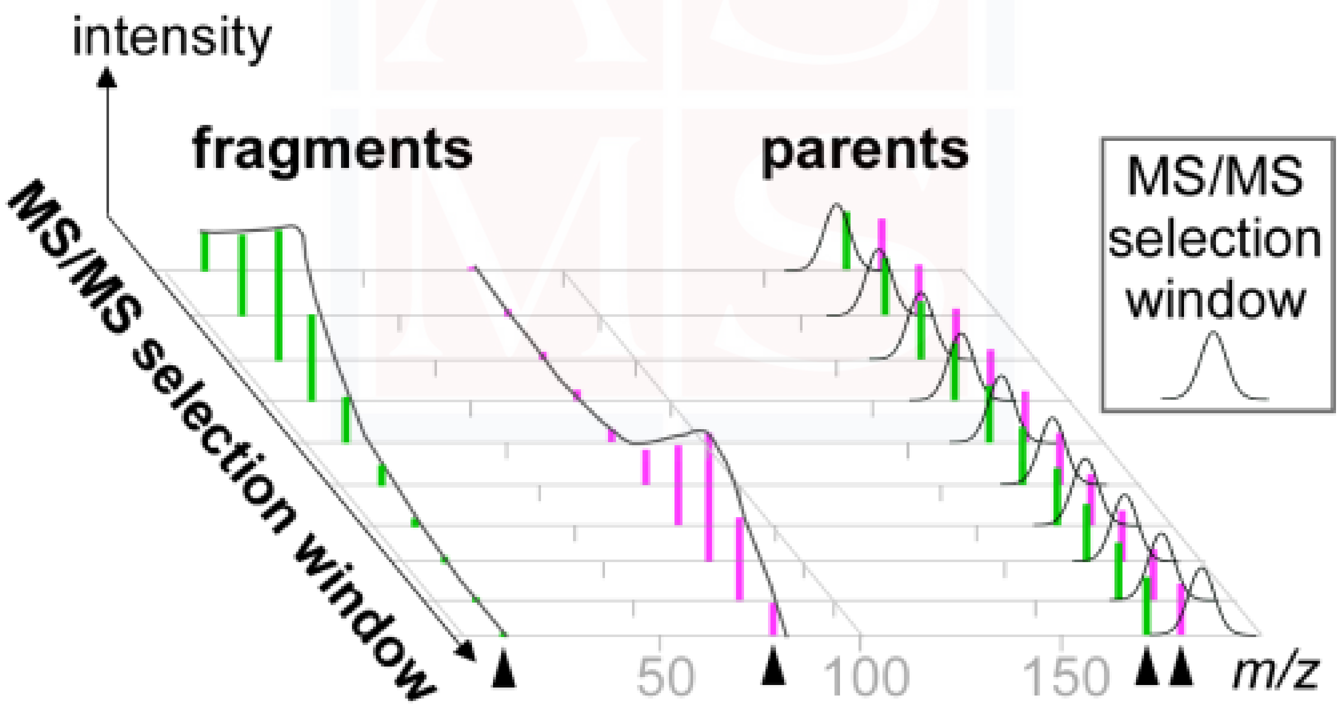


Retention Time
m/z

Untargeted metabolomic analysis with chimeric MS/MS data

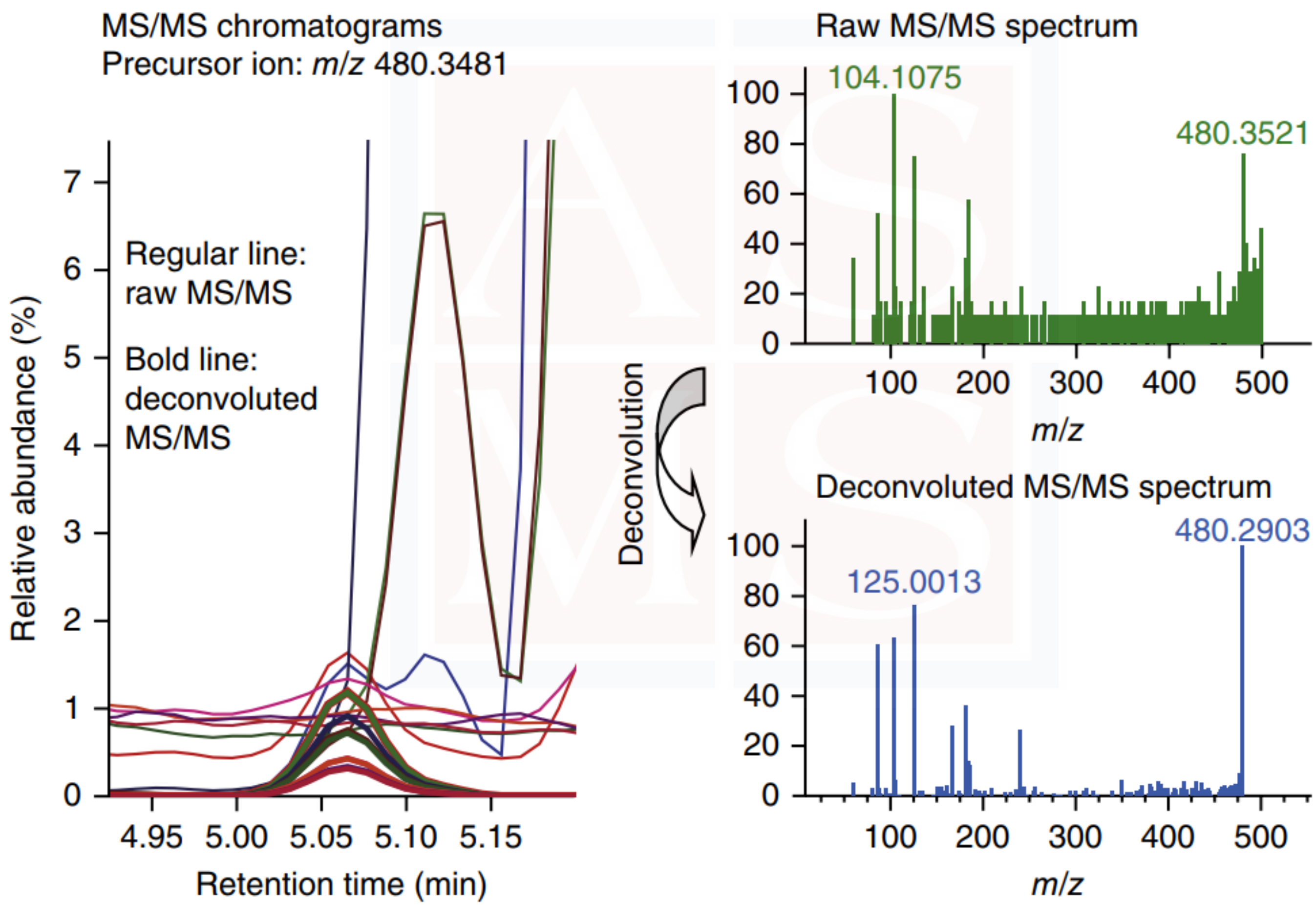
decoMS2

Experimental deconvolution with sliding MS/MS windows



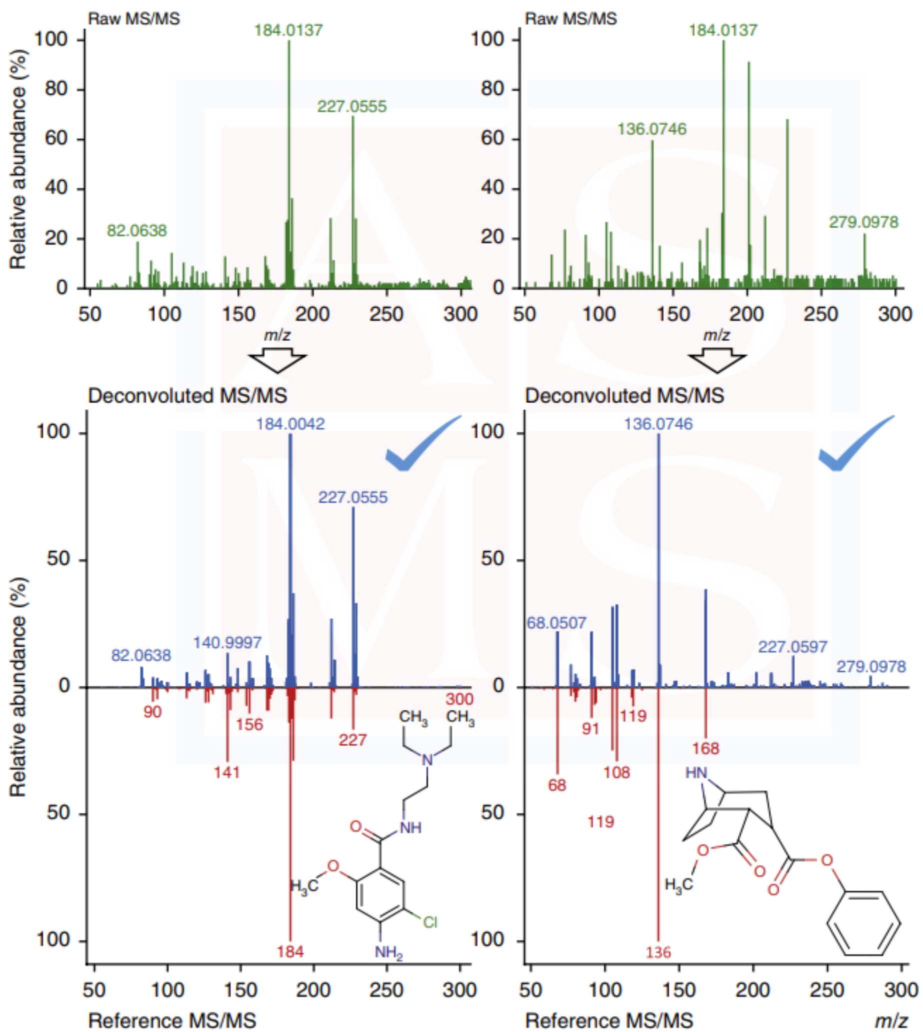
Nikolskiy et al., Anal Chem 2013

SWATH-MS/MS acquisition in untargeted metabolomics



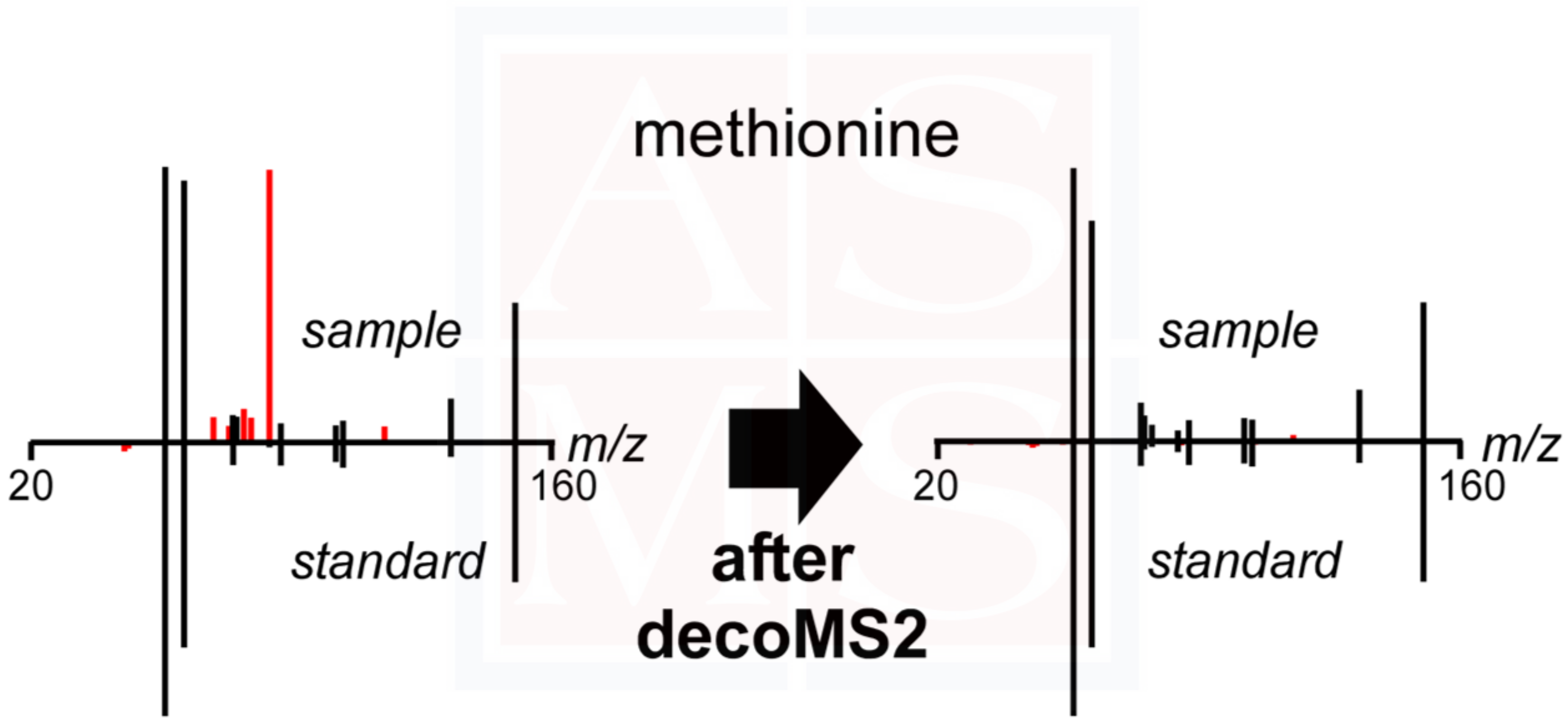
Tsugawa et al., Nature Methods, 2015

Identifications from Data Independent MS/MS SWATH untargeted metabolomics exp.



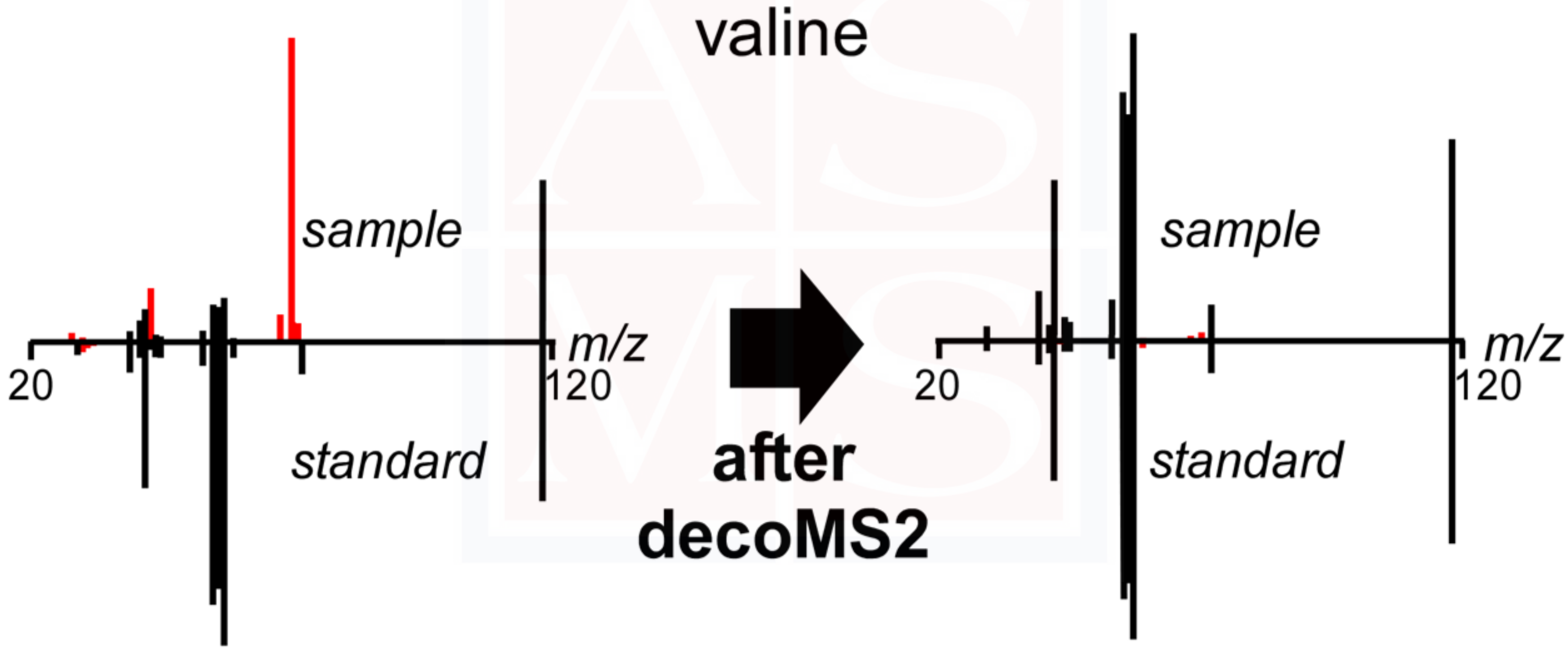
Tsugawa et al., Nature Methods, 2015

Identifications from chimeric MS/MS data in untargeted metabolomics



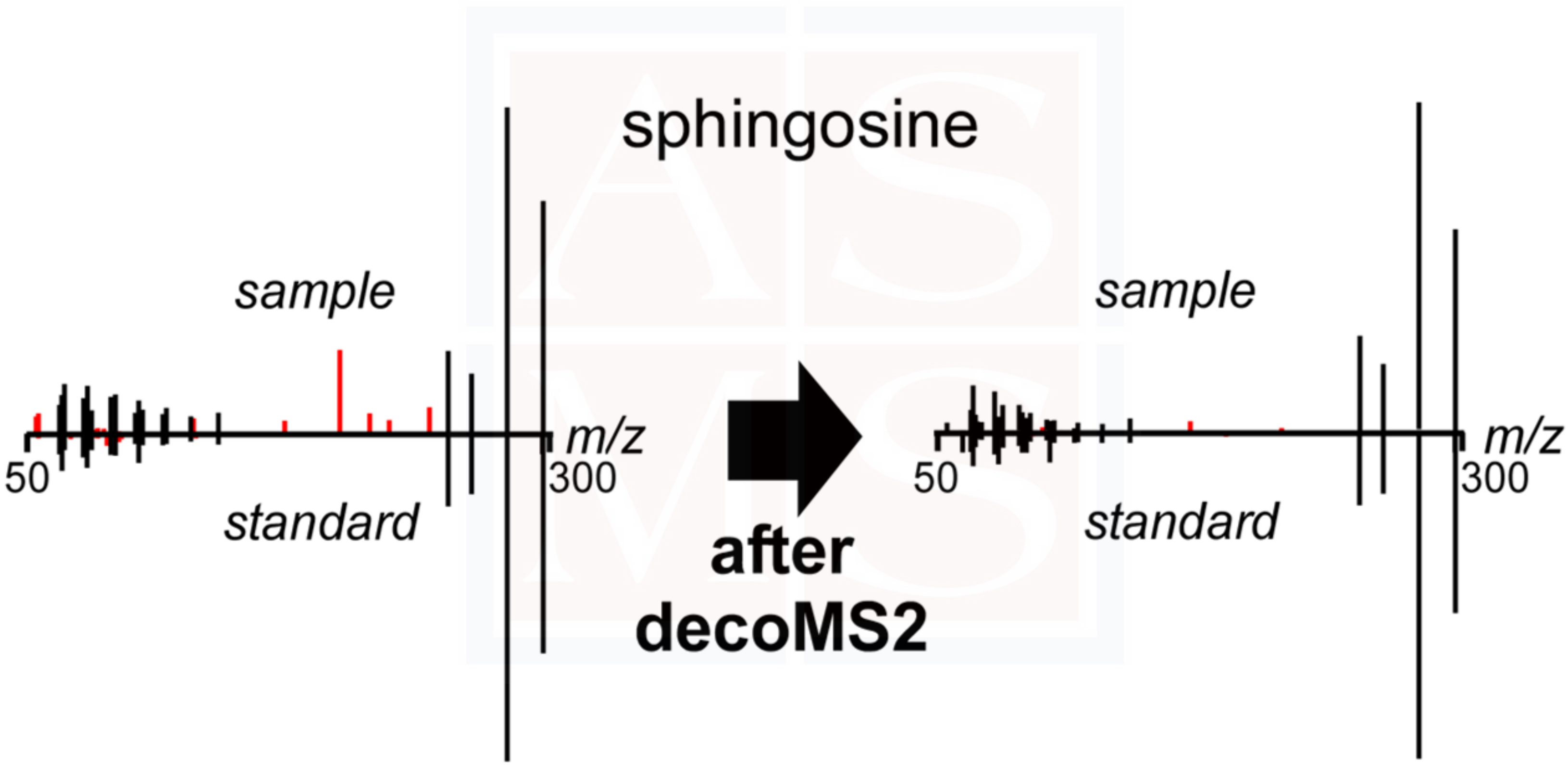
Nikolskiy et al., Anal Chem 2013

Identifications from chimeric MS/MS data in untargeted metabolomics



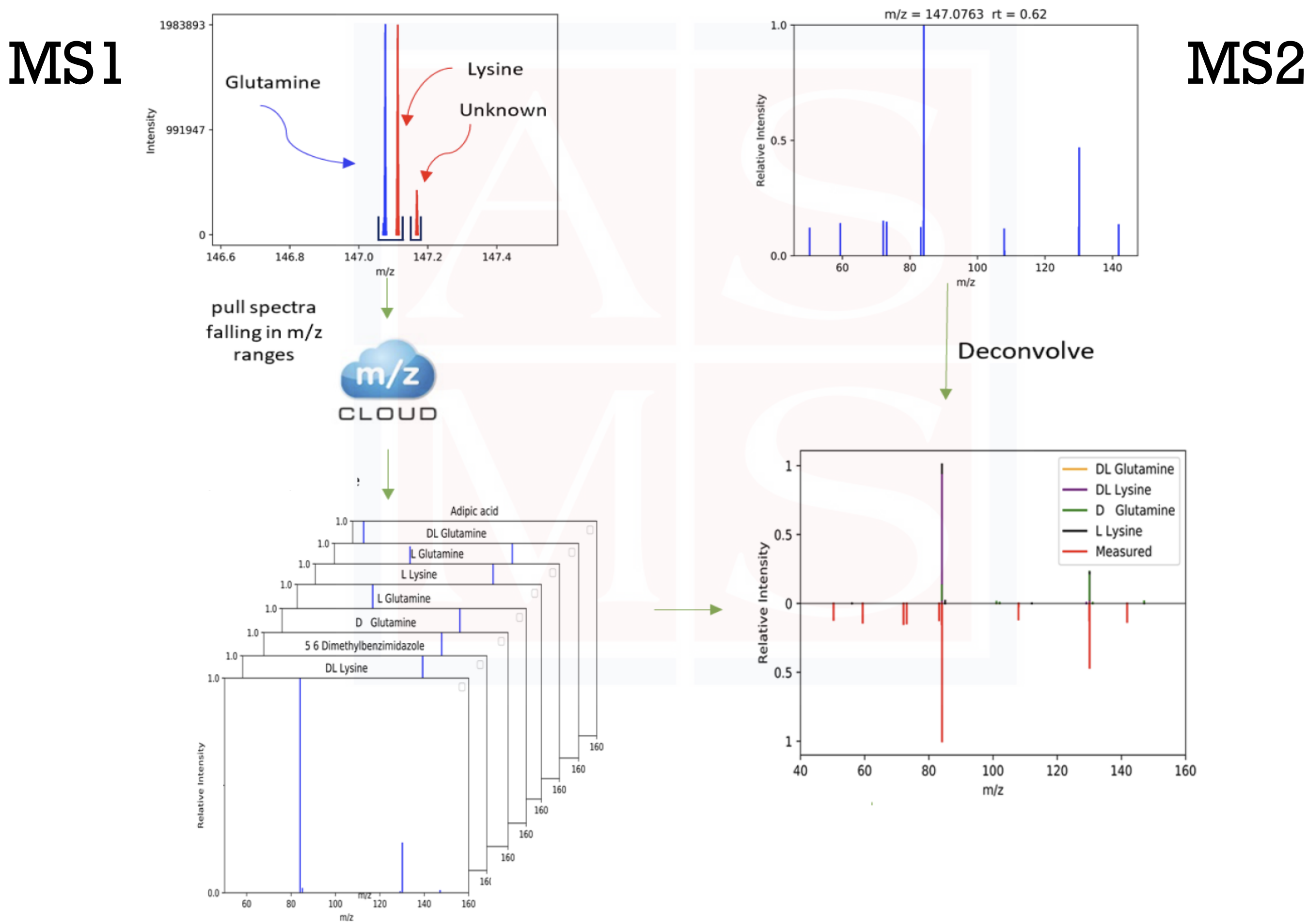
Nikolskiy et al., Anal Chem 2013

Identifications from chimeric MS/MS data in untargeted metabolomics



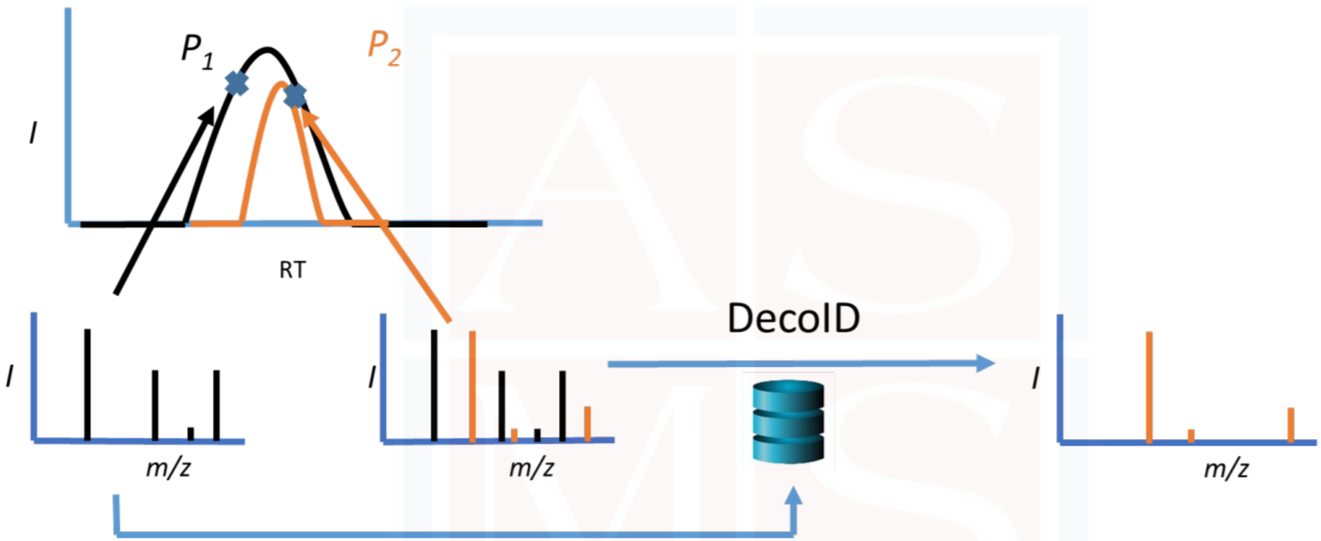
Nikolskiy et al., Anal Chem 2013

DecoID: identifying metabolites from chimeric MS/MS data by the "reverse" approach



Stancliffe et al., Nature Methods 2021

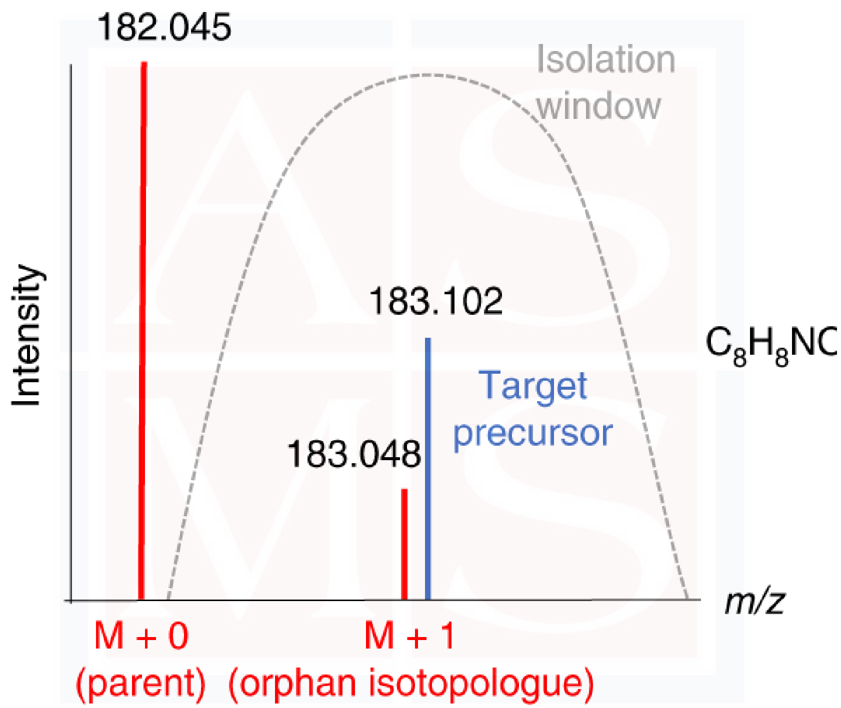
DecoID: identifying metabolites from chimeric MS/MS data by the "reverse" approach



Dealing with molecules that
do not have MS/MS in libraries

Stancliffe et al., Nature Methods 2021

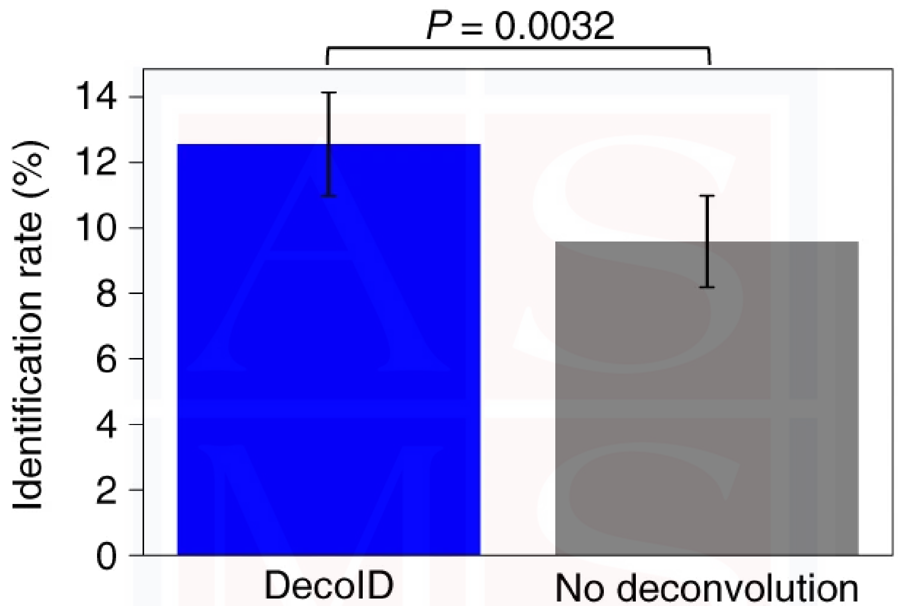
DecoID: identifying metabolites from chimeric MS/MS data by the "reverse" approach



Dealing with molecules that do not have MS/MS in libraries

Stancliffe et al., Nature Methods 2021

DecoID: identifying metabolites from chimeric MS/MS data by the "reverse" approach

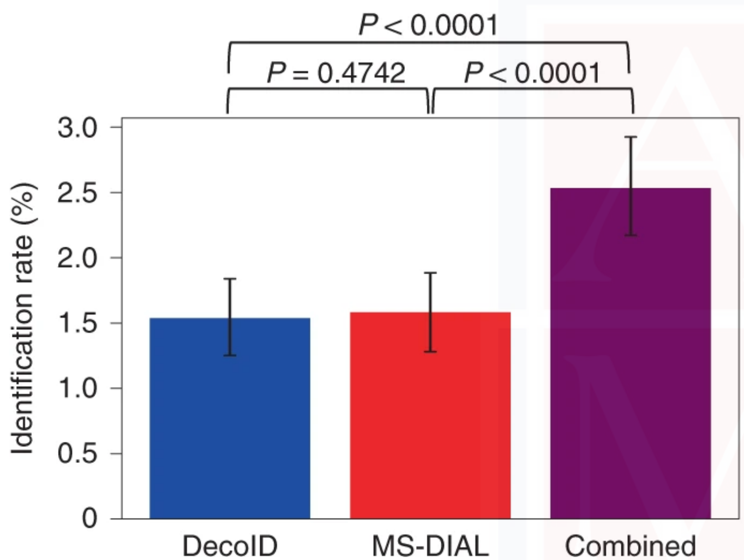


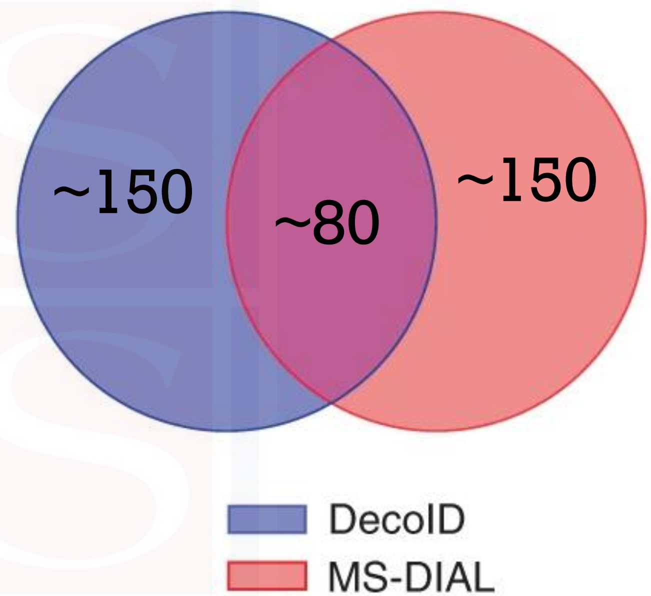
DecoID increases identifications from human plasma

(backwards compatible with all MS/MS data)

Stancliffe et al., Nature Methods 2021

Deconvolution of chimeric MS/MS data in metabolomics by using a combined approach





data from human plasma

Stancliffe et al., Nature Methods 2021

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

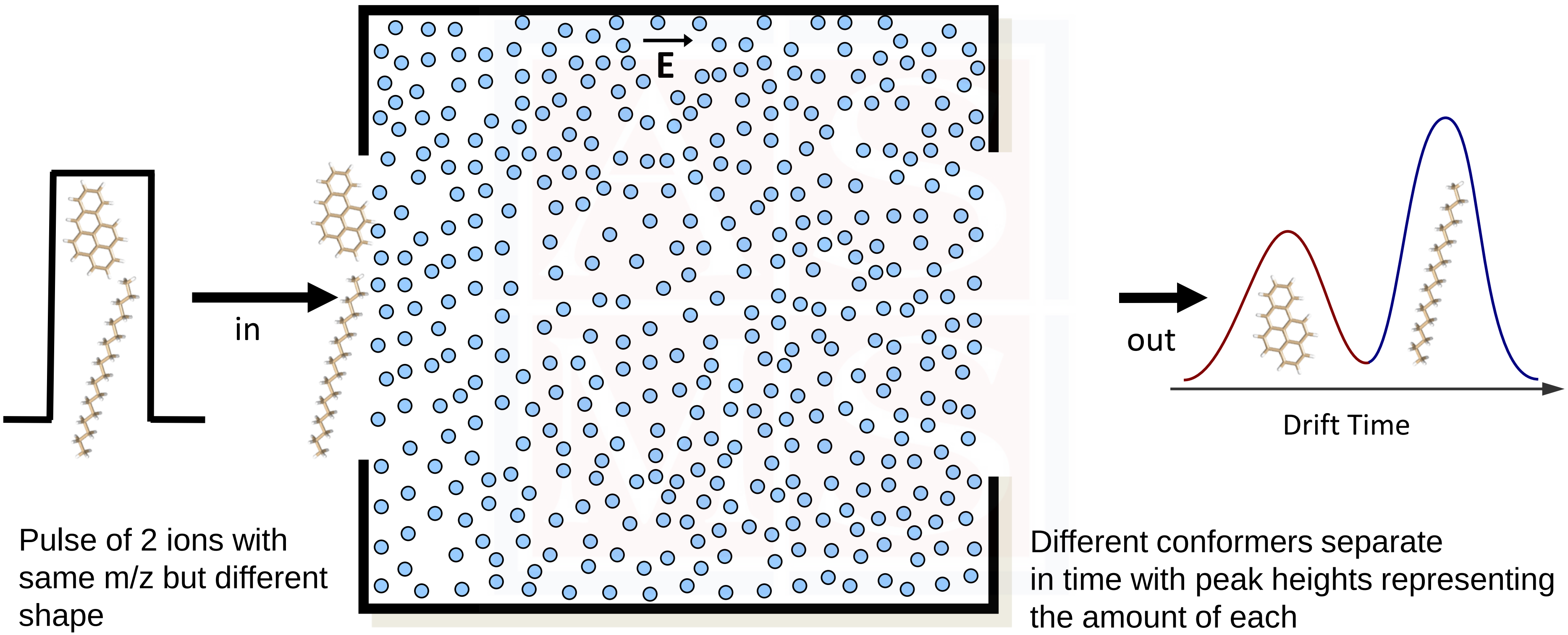
3.) achieve high-quality MS/MS data

Can ion mobility replace chromatography?

Can ion mobility replace chromatography?

Can ion mobility replace chromatography?

Ion mobility concept

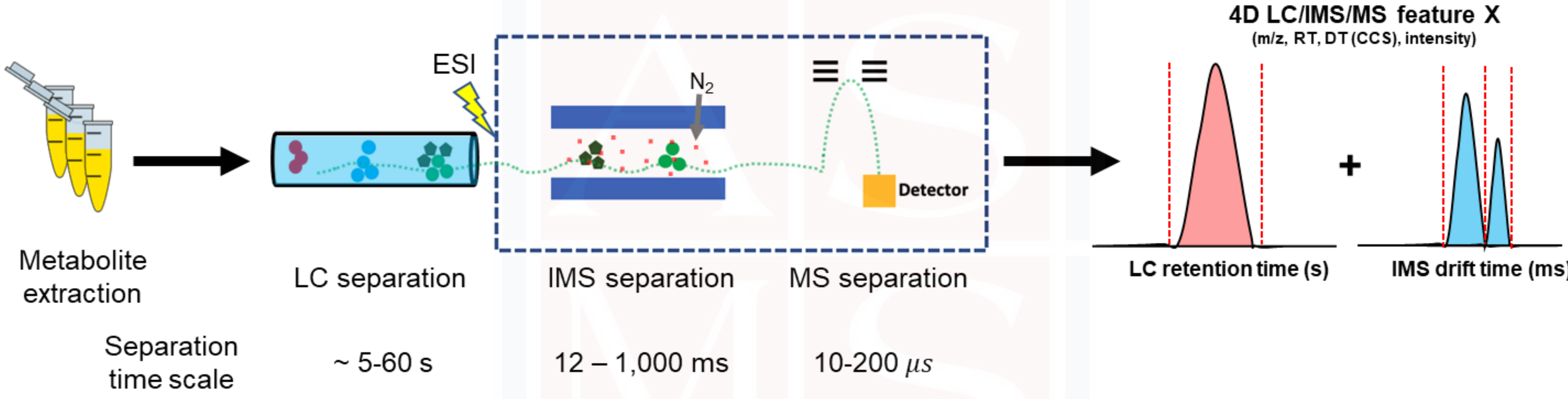


Pulse of 2 ions with same m/z but different shape

Different conformers separate in time with peak heights representing the amount of each

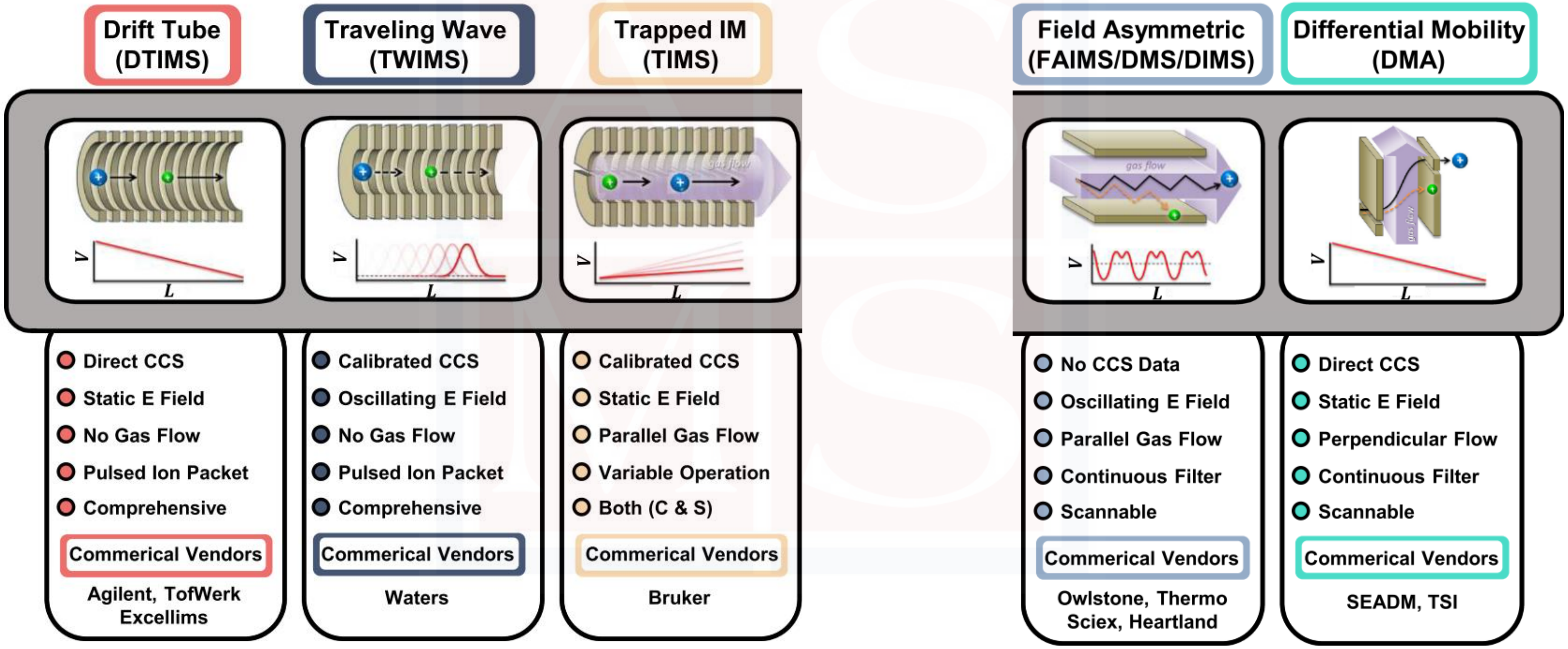
Drift Cell

Can ion mobility replace chromatography?



Can ion mobility replace chromatography?

- Temporally-dispersive system:
  - DTIMS/TWIMS(SLIM)/TIMS
- Spatially-dispersive system:
  - FAIMS/DMA



Journal of the American Society for Mass Spectrometry 30.11 (2019): 2185-2195

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data

Can ion mobility replace chromatography?

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers

3.) achieve high-quality MS/MS data √

Can ion mobility replace chromatography?

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation)

2.) provide retention-time identifiers √

3.) achieve high-quality MS/MS data √

Can ion mobility replace chromatography?

Why is chromatography important in metabolomics?

1.) reduction of matrix effects (quantitation) ?

2.) provide retention-time identifiers √

3.) achieve high-quality MS/MS data √

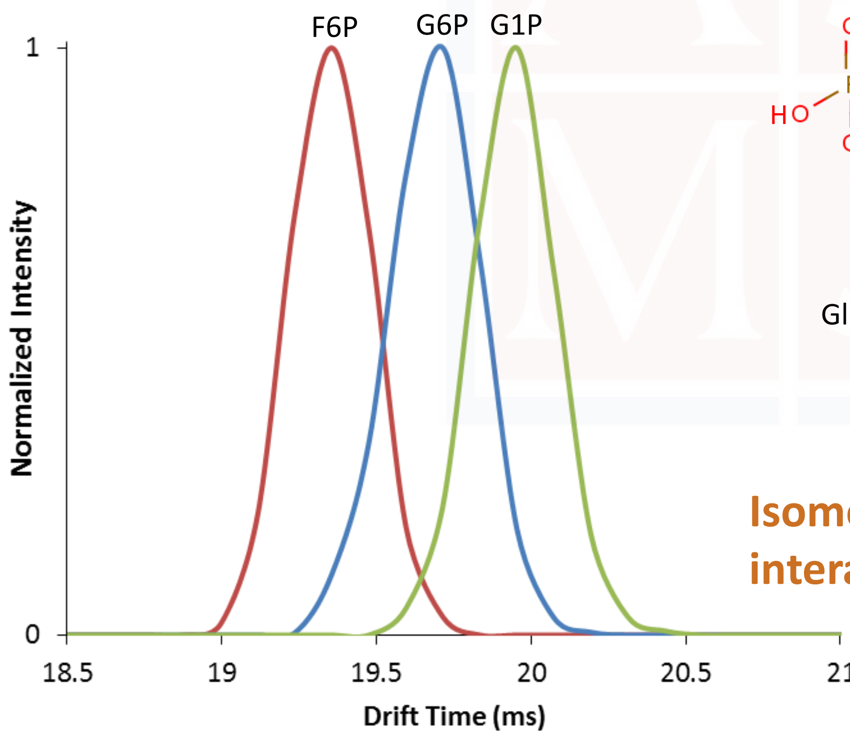
Can ion mobility replace chromatography?

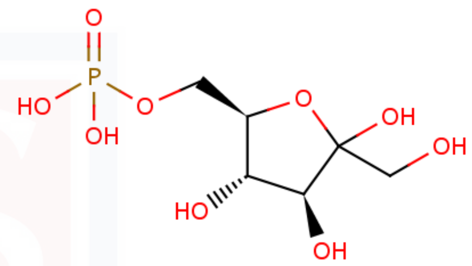
Can ion mobility replace chromatography?

- Excellent complement

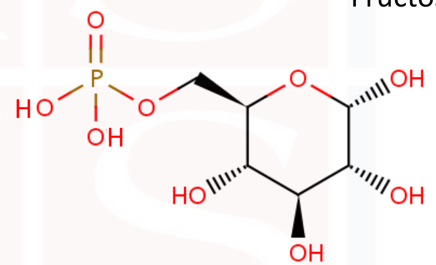
Can ion mobility replace chromatography?

- Excellent complement

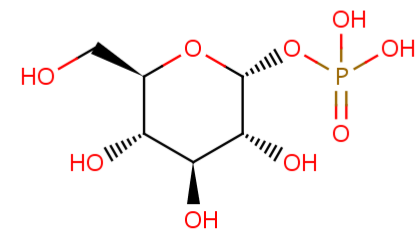




Fructose-6-phosphate (F6P)



Glucose-6-phosphate (G6P)



Glucose-1-phosphate (G1P)

Isomers difficult to separate with hydrophobic interaction liquid chromatography (HILIC)

Deprotonated form $[M - H]^-$ m/z = 259.02

X. Zhang, et al. Clinical Mass Spectrom. 2017.

Can ion mobility replace chromatography?

- Excellent complement
- Collision cross section related to shape and size of an ion
- Corresponds to area that collides w/drift gas
- Robust physiochemical property
- Can easily be compared between labs

Courtesy of Professor Kevin Pagel and Waters

Choosing the appropriate chromatography for untargeted metabolomics

- At this time, many researchers use RPLC and HILIC
- However, there are many variations of columns and gradients
- Most methods have only been evaluated with targeted methods or by counting features

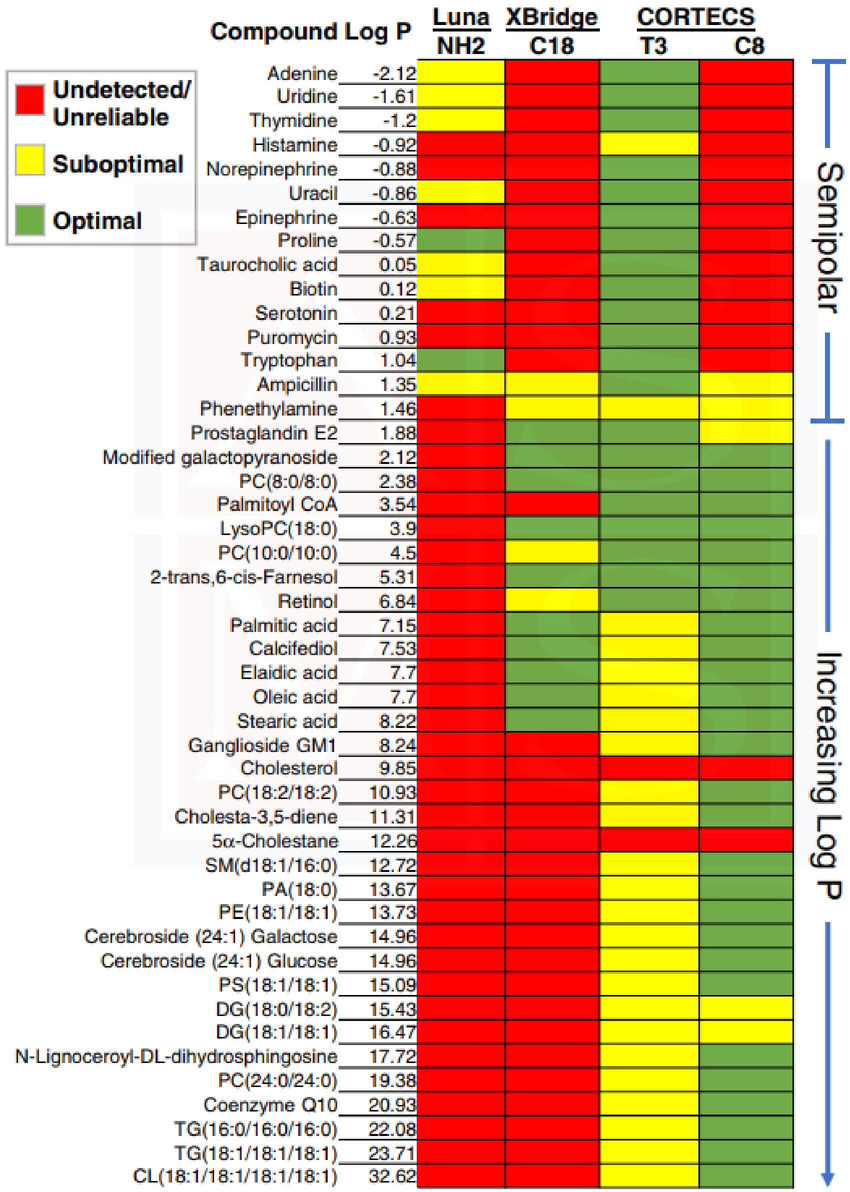
Choosing the appropriate chromatography for untargeted metabolomics

1. Reversed-phase LC
2. Hydrophillic interaction LC
3. Silica-hydride based LC
4. Mixed-mode LC

1. RPLC for untargeted metabolomics

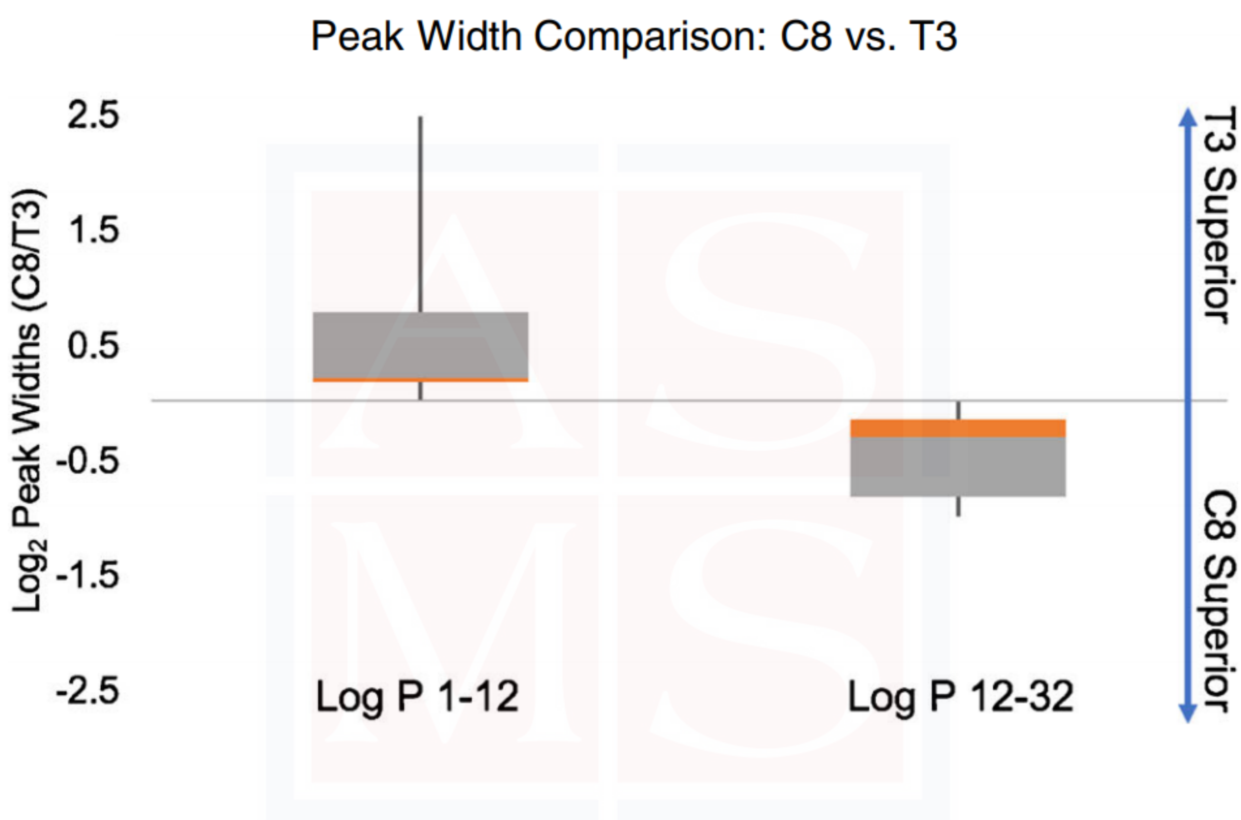
- Pro: Most robust and well understood chromatography
- Pro: Peak shapes tend to be better behaved than HILIC
- Con: Water-soluble metabolites (e.g., central carbon) come out in void volume

1. RPLC for untargeted metabolomics



Naser et al., Anal BioAnal Chem 2018

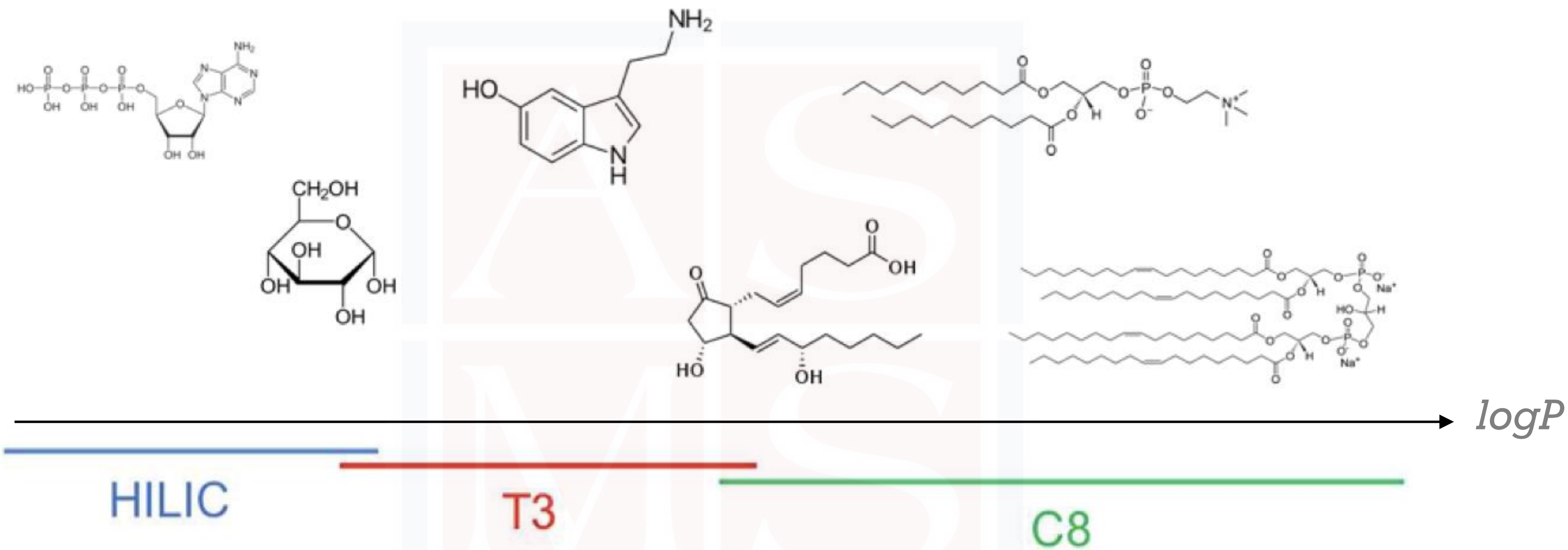
1. RPLC for untargeted metabolomics



20-min gradients

Naser et al., Anal BioAnal Chem 2018

1. RPLC for untargeted metabolomics



Naser et al., Anal BioAnal Chem 2018

1. RPLC for untargeted metabolomics

ION PAIRING

- Typically used in negative-mode to detect central carbon metabolites
- Metabolites form ionic interactions with counter ions that have lipid tails
- Tributylamine (TBA) at 10 mM and pH 4.95 most popular.

WARNING: TBA CONANMINATES LINES AND NEGATIVELY AFFECTS POSITIVE MODE ANALYSIS!

Choosing the appropriate chromatography for untargeted metabolomics

1. Reversed-phase LC

2. Hydrophillic interaction LC

3. Silica-hydride based LC

4. Mixed-mode LC

2. HILIC for untargeted metabolomics

2. HILIC for untargeted metabolomics

- Pro: Best strategy to separate highly polar metabolites (central carbon)
- Con: Not as well understood as RPLC
- Con: Peak shapes less well behaved compared to RPLC and cause informatic problems
- Con: Column bleeding problematic for some columns
- Con: Column lives to tend to be shorter (~150 injections) compared to RPLC (~1000 injections)
- Con: Much longer equilibration times than RPLC

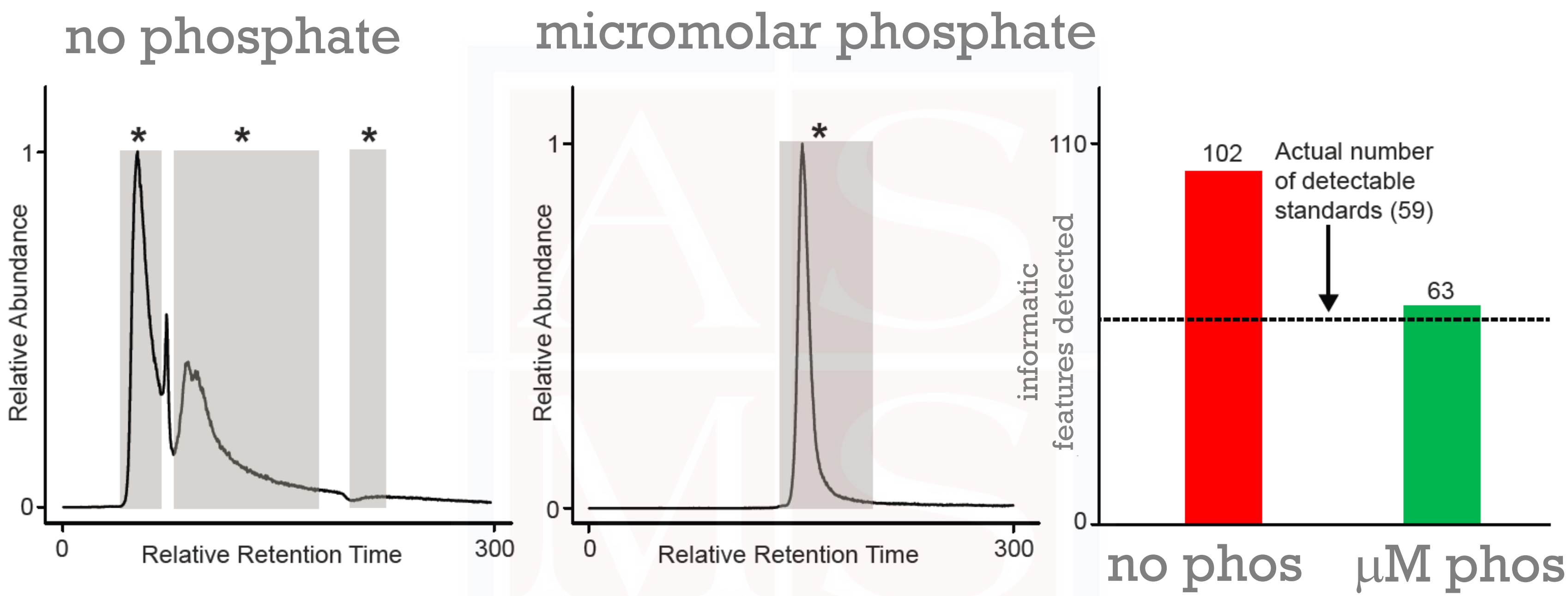
2. HILIC for untargeted metabolomics

- Retention mechanisms are:
  (i) liquid-liquid partitioning (water-layer formation)

  (ii) electrostatic interactions with point charges on silica and/or its derivatization

- Electrostatic interactions necessary to form water layer but can be problematic because of spread in adsorption energies → causes variable elution behaviors

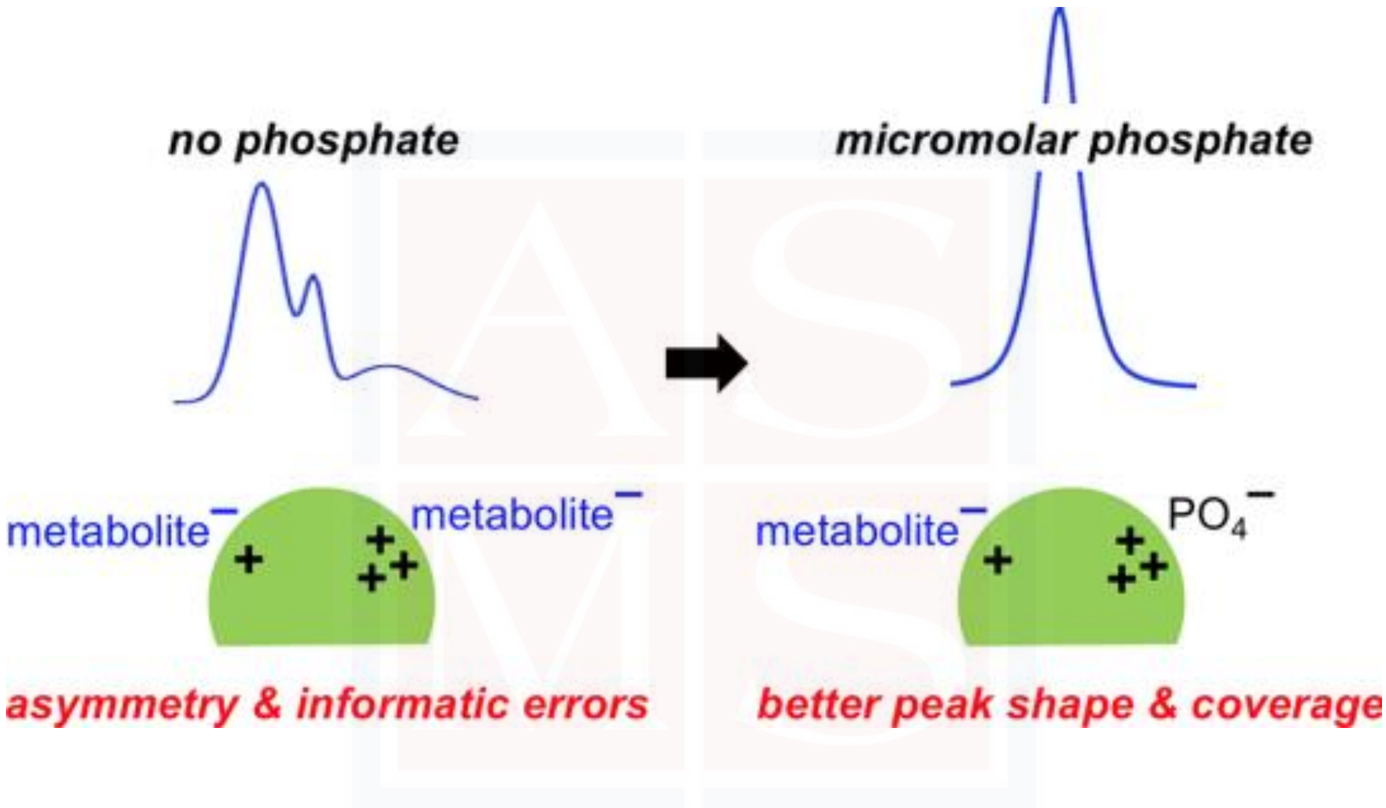
2. HILIC for untargeted metabolomics



Spalding et al., Journal of Proteome Research 2018
Trace Phosphate Improves ZIC-pHILIC Peak Shape, Sensitivity, and Coverage for Untargeted Metabolomics

Hsiao et al., Analytical Chemistry 2018
Improved LC/MS Methods for the Analysis of Metal-Sensitive Analytes using Medronic Acid as a Mobile Phase Additive

2. HILIC for untargeted metabolomics

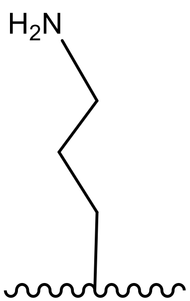


Spalding et al., Journal of Proteome Research 2018
Trace Phosphate Improves ZIC-pHILIC Peak Shape, Sensitivity, and Coverage for Untargeted Metabolomics

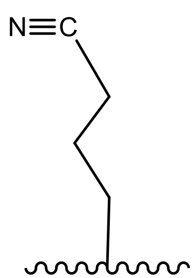
Hsiao et al., Analytical Chemistry 2018
Improved LC/MS Methods for the Analysis of Metal-Sensitive Analytes using Medronic Acid as a Mobile Phase Additive

2. HILIC for untargeted metabolomics

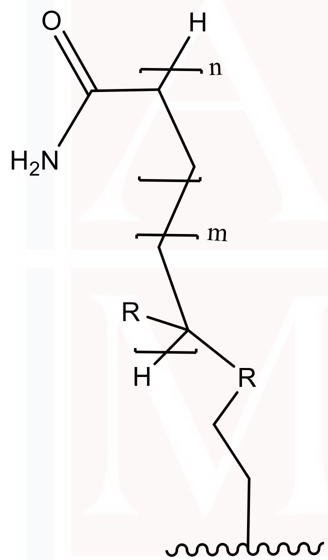
Most commonly used HILIC stationary phases have derivatize silica to enhance retention



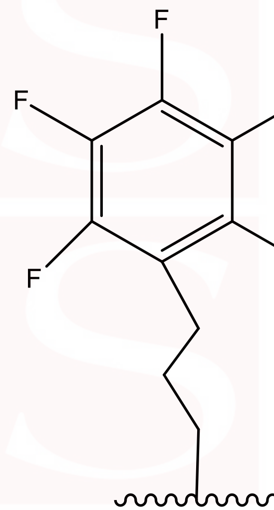
amino (Luna)



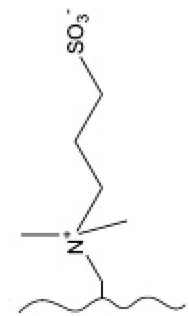
cyano



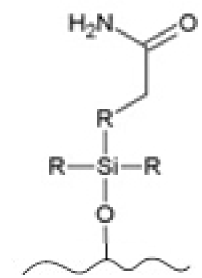
TSK-Gel Amide-80



pentafluoro-phenylpropyl



sulfobetaine (zwitterionic)

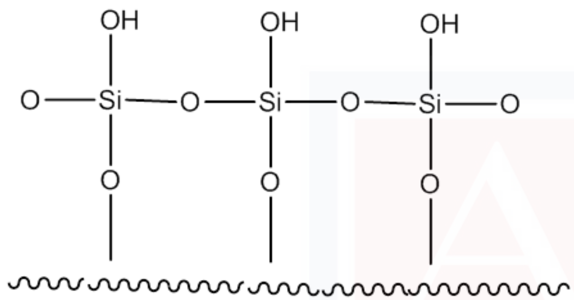


amide

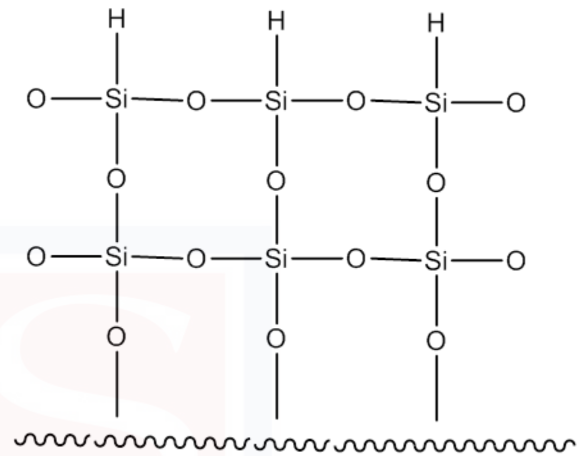
Choosing the appropriate chromatography for untargeted metabolomics

1. Reversed-phase LC
2. Hydrophillic interaction LC
3. Silica-hydride based LC
4. Mixed-mode LC

3. Silica-hydride for untargeted metabolomics



type-B silica



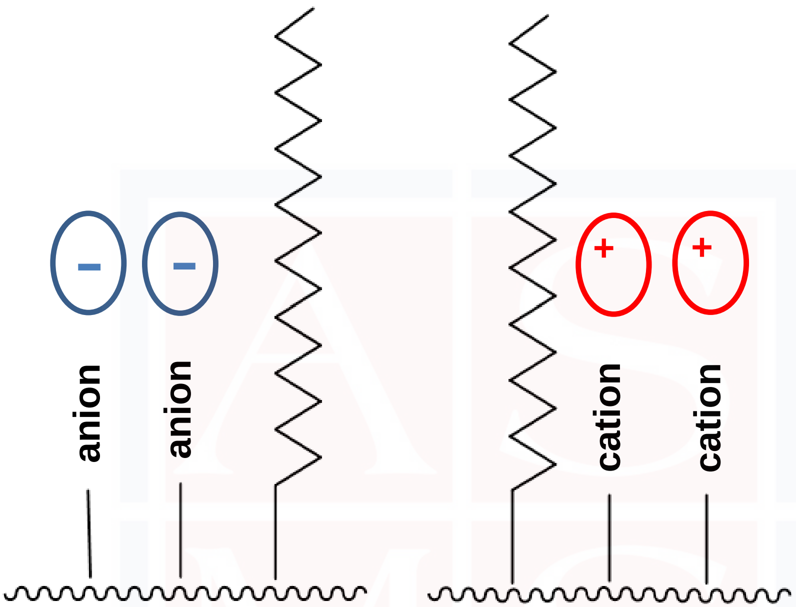
silica hydride

- Silica-hydride columns have less of a polar surface and show less of an attraction for water
- At high % water in mobile phase, RP properties dominate. At high % organic in mobile phase, hydrophilic compounds retained
- Diamond hydride columns have 2% bonded carbon moieties to retain lipids

Choosing the appropriate chromatography for untargeted metabolomics

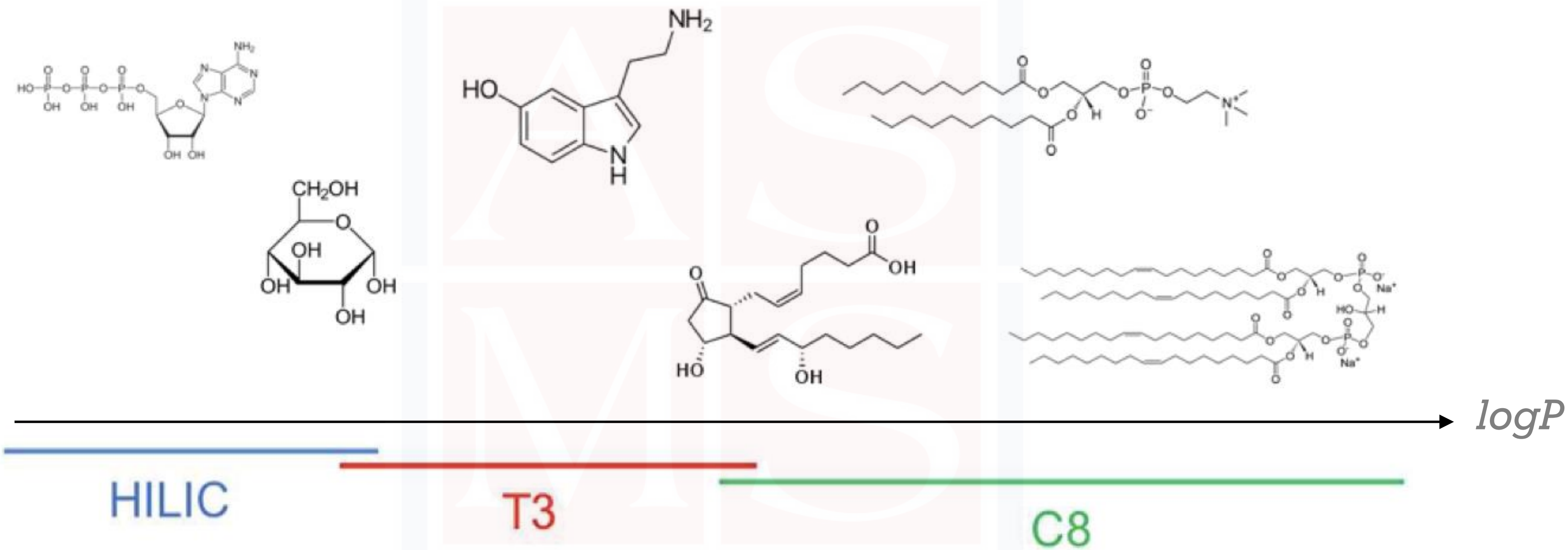
1. Reversed-phase LC
2. Hydrophillic interaction LC
3. Silica-hydride based LC
4. Mixed-mode LC

4. Mixed-mode for untargeted metabolomics



- Ion-exchange ligands blended with alkyl functional groups
- Scherzo SM-C18 and Acclaim Trinity P1

Combining stationary phases for comprehensive coverage



- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Principles of informatics

Informatics can be divided into two steps

1. Processing raw metabolomic data (software is required)

2. Analyzing software results (use databases)

Informatics can be divided into two steps

1. Processing raw metabolomic data (software is required)

2. Analyzing software results (use databases)

Step 1: processing raw data with software

Goals of data processing:

- find features

(aka "peak detection")

- group same features between samples

(aka "correspondence determination")

Step 1: processing raw data with software

Goals of data processing:

- find features

(aka "peak detection")

- group same features between samples

(aka "correspondence determination")

Hundreds of software options available

Step 1: processing raw data with software

Many options available ranging from customized R scripts to canned solutions

Step 1: processing raw data with software

Many options available ranging from customized R scripts to canned solutions

Freeware (GUI, complete workflow): XCMS Online, MZmine, MetAlign, MAVEN, MS-DIAL, MetaboAnalyst, and others…

Commercial (complete workflow): MassProfiler Professional (Agilent), Compound Discoverer (Thermo), PeakView (SCIEX), Markerview (SCIEX), MetabolitePilot (SCIEX), Progenesis (Nonlinear Dynamics/Waters), MarkerLynx (Waters), AMIX (Bruker), Profiler AM+ (Shimadzu),…

R/C/Python/MATLAB packages: XCMS, RAMclustR, CAMERA, FragPred, IPO, MetExtract, xMSannotatot, compMS2Miner, MIDcor, MetaboQC, mixOmics, LIQUID, mzunity, massPix, PIXiE, proFIA, MetaboAnalystR, warpgroup, ChemRICH, MetaboLyzer, and hundreds more….

Step 1: processing raw data with software

How do you decide?

- Important considerations include cost, ease of use, performance, speed, and data compatibility

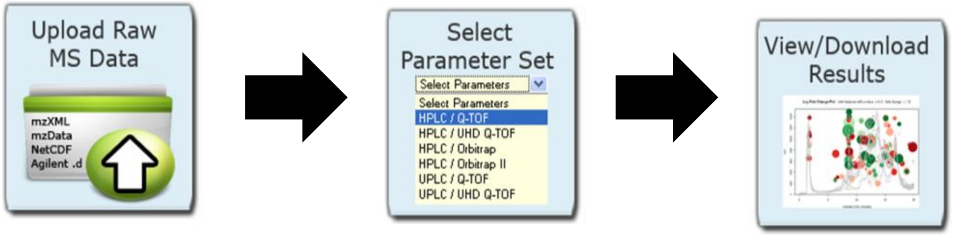
Step 1: processing raw data with software

How do you decide?

- Important considerations include cost, ease of use, performance, speed, and data compatibility

- XCMS was most popular in 2017*

→R-based (many diff. variations)
→ Implemented in Galaxy-M (facilitates integration with other software)
→Cloud-based (terrific resource developed & maintained by G. Siuzdak at Scripps)
→Cloud-based software, easy to use and compatible with most workflows



*according to International survey: Weber et al., Metabolomics 2017
*recent data shows increasing usage of MZmine and MSDIAL

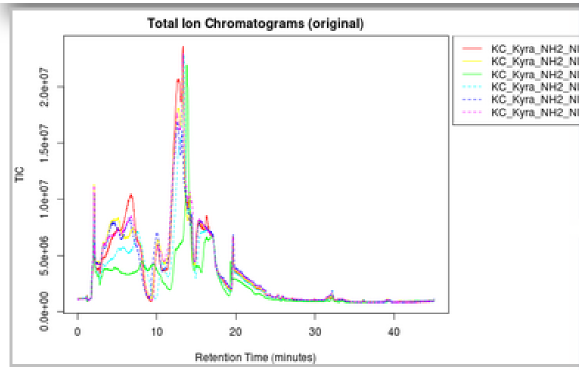
Example XCMS Online output

Scripps Center For Metabolomics XCMS Online

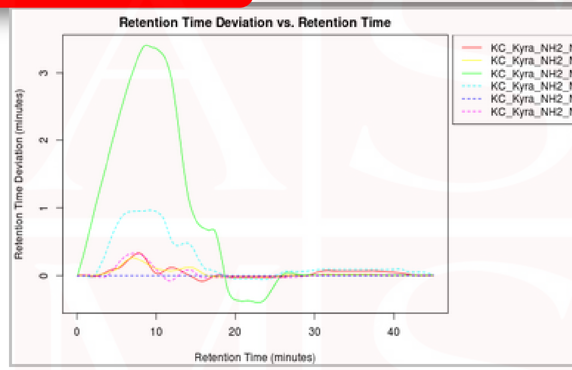
Job ID:	1021725	Job Name:	kevincho007_2014-04-26_05:58:46	Create Date:	2014-04-26 05:58:46
Parameter (ID#)	HPLC / Q-TOF (1)	Log:	View Log	Finish Date:	2014-04-26 06:20:47
Status:	job complete (See Warnings)	Total Aligned Features:	8,091	Experiment Type:	Pairwise
Datasets Used:	KK_Mice_WT_NH2_NEG(ID#77877) * vs. KK_Mice_KO_NH2_NEG(ID#77878)	Paired Samples:	False	Share Status:	Shared with 1 User(s) | [Stop sharing]

2014-04-26 06:18:57 : There are regions with poor chromatographic resolution. Feature annotations (CAMERA) were omitted for these regions.

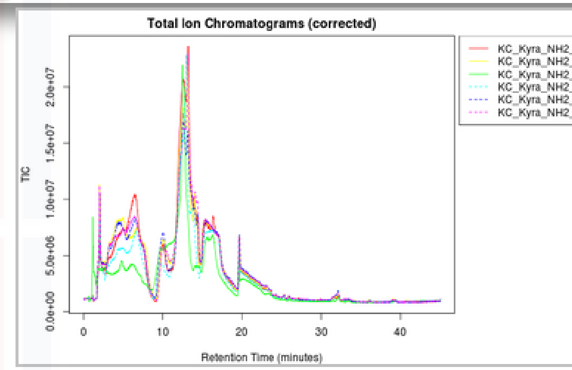
BROWSE RESULT TABLE | INTERACTIVE CLOUD PLOT | iPCA



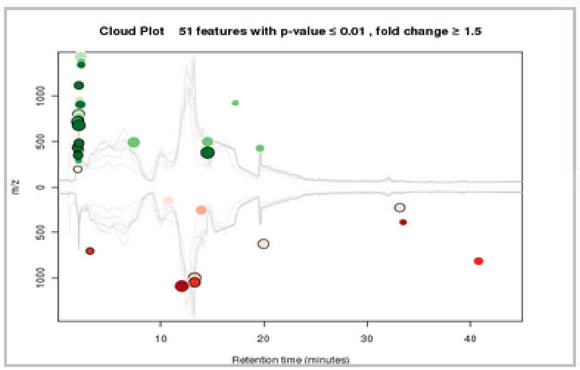
PNG PDF



PNG PDF

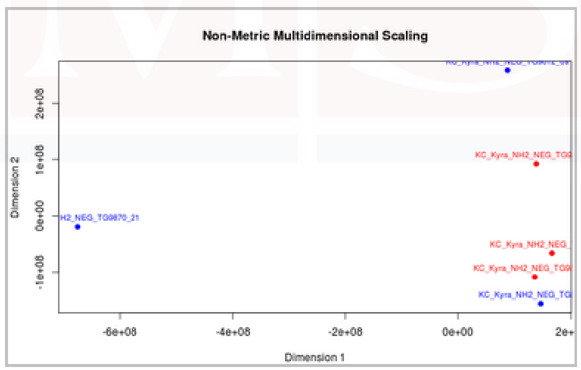


PNG PDF

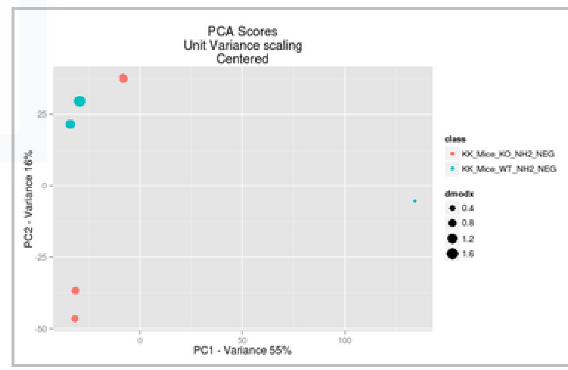


PNG PDF SVG

Interactive Version



PNG PDF



PNG PDF

More

Complete Downloadable Results (including diffreport, EIC's, boxplots, etc.) : results.zip

Example XCMS Online output

Scripps Center For Metabolomics XCMS Online

Quick Compound Search: [ ] Search Clear

Job#1021725 : kevincho007_2014-04-26_05:58:46

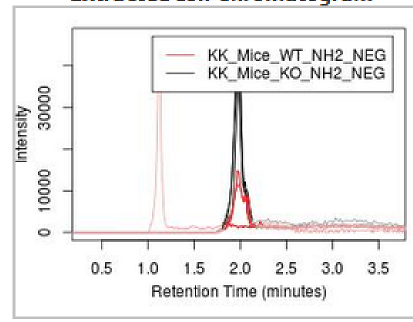
Columns | Hide isotopic peaks | Page 1 of 81 | 100 | View 1 - 100 of 8 091

Feature	UP/DOWN	fold chang	p-value	m/z	retention time	MaxInt	Ctrl(x̄)	Exp(x̄)	isotopes	adducts	feature g	Notes
1	UP	2.1	0.00008	347.1791	1.98	41,236	119,128	251,582	[258][M]+	[M+K]+ 308	54	
2	UP	4.0	0.00016	678.3123	2.05	3,968	2,943	11,877			10	
3	UP	2.1	0.00043	1,115.6273	2.06	825	1,719	3,536			10	
4	DOWN	8.2	0.00045	1,091.5924	12.04	964	3,732	150		[2M+Na+3]	2	
5	UP	1.7	0.00045	361.1923	1.97	14,573	32,987	57,639	[271][M]+	[M+Na]+ 33	54	
6	UP	3.0	0.00059	906.5352	2.19	3,203	18,658	55,429		[M+H]+ 905	26	
7	UP	2.3	0.00073	481.2039	2.07	12,750	41,475	94,857	[416][M]+		10	
8	UP	2.1	0.00087	1,344.8760	2.28	2,710	21,548	45,649	[902][M+1]+		122	
9	UP	4.9	0.00103	377.0750	14.54	14,416	25,065	122,159		[M+H-H20]	15	
10	UP	2.6	0.00109	431.1997	1.94	18,652	34,283	87,616	[349][M]+		40	
11	DOWN	1.8	0.00139	386.9846	33.49	1,007	14,725	8,193			312	
12	UP	1.2	0.00152	408.8461	19.61	3,166	19,429	22,809			6	
13	UP	3.2	0.00203	717.4303	1.92	3,449	8,085	26,144		[M+Na+K-H	40	
14	UP	2.2	0.00277	723.4338	2.00	2,331	10,864	24,131			10	
15	DOWN	2.5	0.00289	818.8290	40.82	14,881	245,779	98,396		[M+K]+ 779	151	
16	UP	2.6	0.00302	725.4329	2.02	3,216	10,910	28,014			10	
17	UP	2.2	0.00304	1,371.8961	2.28	1,884	14,670	31,679	[906][M+1]+		122	
18	DOWN	1.8	0.00307	704.3520	3.15	4,632	122,443	68,927			74	
19	UP	2.0	0.00311	902.5183	2.15	3,343	19,838	39,088			10	
20	UP	2.1	0.00314	701.4355	2.02	3,135	10,543	22,040			10	
21	DOWN	2.9	0.00329	1,049.5458	13.29	2,931	16,900	5,782	[826][M]+	[M+Na]+ 10	36	
22	UP	1.7	0.00384	290.1325	2.04	6,501	20,413	35,237	[171][M+1]+		10	
23	UP	2.2	0.00408	372.1831	1.95	23,684	74,590	164,601	[287][M+1]+		40	
24	UP	4.0	0.00461	494.9777	14.54	779	547	2,776		[M+Na]+ 47	15	
25	UP	1.9	0.00470	923.7675	17.24	1,161	7,662	14,334		[M+Na]+ 9(	66	
26	UP	2.0	0.00470	451.2206	2.07	21,044	41,183	82,707	[364][M+3]+		10	
27	UP	2.2	0.00515	424.8194	19.61	1,758	5,154	11,237			6	
28	UP	3.3	0.00518	757.4223	2.06	4,161	5,688	18,576	[667][M+2]+		10	
29	UP	2.4	0.00528	1,370.8881	2.33	1,776	15,805	37,225	[906][M]+	[M+Na]+ 13	122	

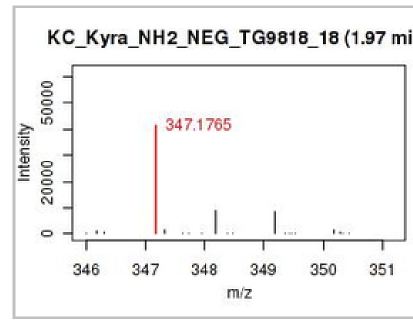
Columns | Export | Page 1 of 81 | 100 | View 1 - 100 of 8 091

Please click on a row to view feature details

Feature #1
m/z : 347.1791
Retention Time (min): 1.98
Extracted Ion Chromatogram



Mass Spectrum | Box-and-Whisker Plot



PPM	Name	Adduct	METLINID
1	Pergolide sulfone	M+H	1789
6	Spenolimycin	M+H	71966
8	Paroxetine	M+NH4	1710

Example XCMS Online output

Scripps Center For Metabolomics XCMS Online

Quick Compound Search: [ ] Search Clear

Job#1021725 : kevincho007_2014-04-26_05:58:46

Columns · Hide isotopic peaks · Page 1 of 81 · 100 · View 1 - 100 of 8 091

Feature	UP/DOWN	fold chang	p-value	m/z	retention time	MaxInt	Ctrl(x̄)	Exp(x̄)	isotopes	adducts	feature g	Notes
1	UP	2.1	0.00008	347.1791	1.98	41,236	119,128	251,582	[258][M]+	[M+K]+ 308	54	
2	UP	4.0	0.00016	678.3123	2.05	3,968	2,943	11,877			10	
3	UP	2.1	0.00043	1,115.6273	2.06	825	1,719	3,536			10	
4	DOWN	8.2	0.00045	1,091.5924	12.04	964	3,732	150		[2M+Na+3]	2	
5	UP	1.7	0.00045	361.1923	1.97	14,573	32,987	57,639	[271][M]+	[M+Na]+ 3	54	
18	DOWN	1.8	0.00307	704.3520	3.15	4,632	122,443	68,927			74	
19	UP	2.0	0.00311	902.5183	2.15	3,343	19,838	39,088			10	
20	UP	2.1	0.00314	701.4355	2.02	3,135	10,543	22,040			10	
21	DOWN	2.9	0.00329	1,049.5458	13.29	2,931	16,900	5,782	[826][M]+	[M+Na]+ 10	36	
22	UP	1.7	0.00384	290.1325	2.04	6,501	20,413	35,237	[171][M+1]+		10	
23	UP	2.2	0.00408	372.1831	1.95	23,684	74,590	164,601	[287][M+1]+		40	
24	UP	4.0	0.00461	494.9777	14.54	779	547	2,776		[M+Na]+ 47	15	
25	UP	1.9	0.00470	923.7675	17.24	1,161	7,662	14,334		[M+Na]+ 9	66	
26	UP	2.0	0.00470	451.2206	2.07	21,044	41,183	82,707	[364][M+3]+		10	
27	UP	2.2	0.00515	424.8194	19.61	1,758	5,154	11,237			6	
28	UP	3.3	0.00518	757.4223	2.06	4,161	5,688	18,576	[667][M+2]+		10	
29	UP	2.4	0.00528	1,370.8881	2.33	1,776	15,805	37,225	[906][M]+	[M+Na]+ 1	122	

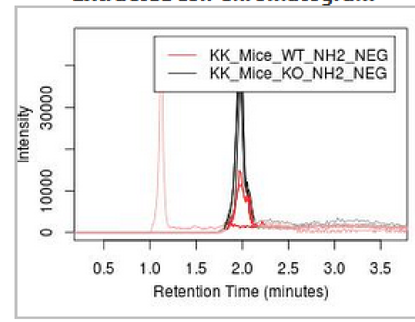
Columns · Export · Page 1 of 81 · 100 · View 1 - 100 of 8 091

Mine data manually

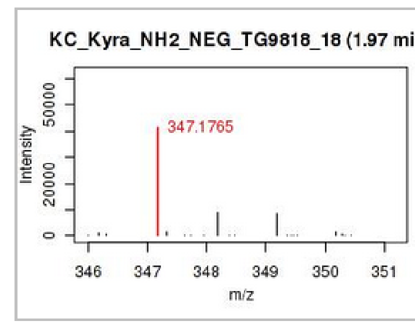
Design targeted MS/MS

Please click on a row to view feature details

Feature #1
m/z : 347.1791
Retention Time (min): 1.98
Extracted Ion Chromatogram



Mass Spectrum | Box-and-Whisker Plot



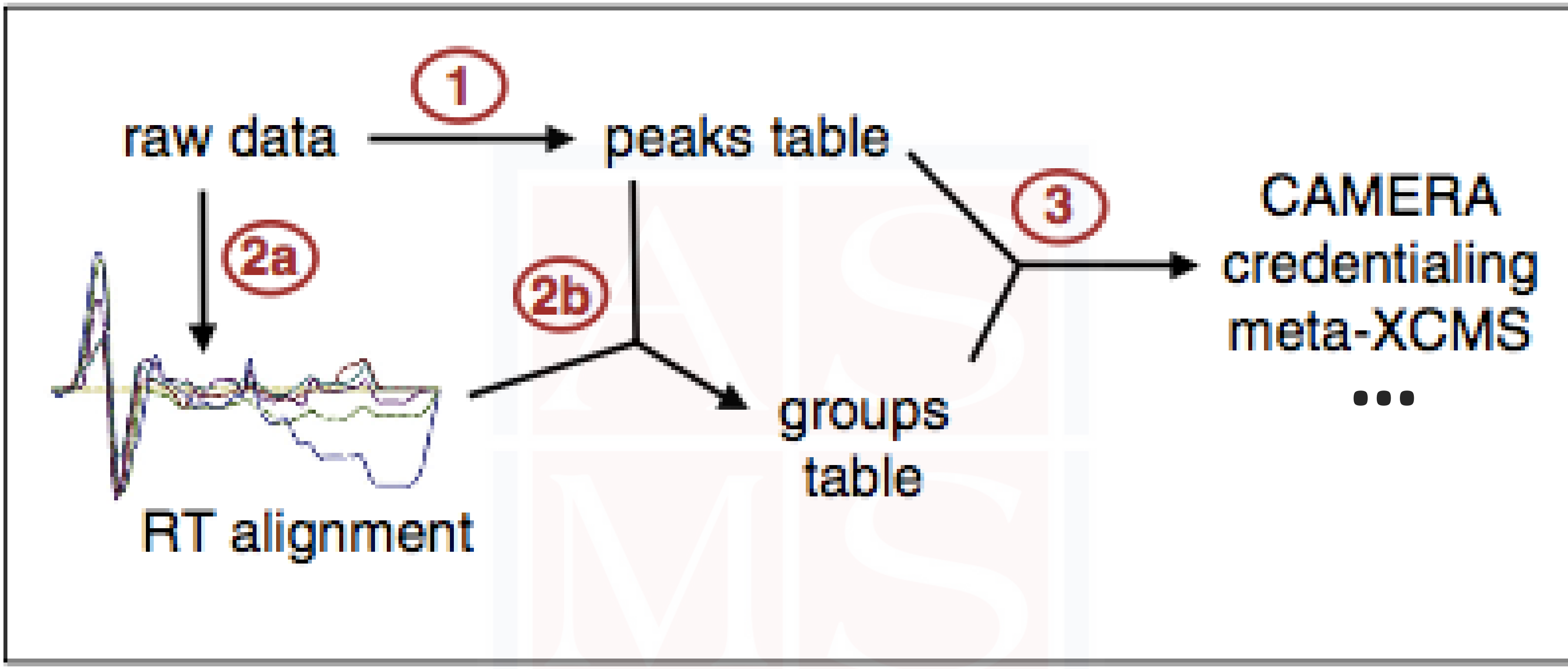
PPM	Name	Adduct	METLINID
1	Pergolide sulfone	M+H	1789
6	Spenolimycin	M+H	71966
8	Paroxetine	M+NH4	1710

Step 1: processing raw data with software

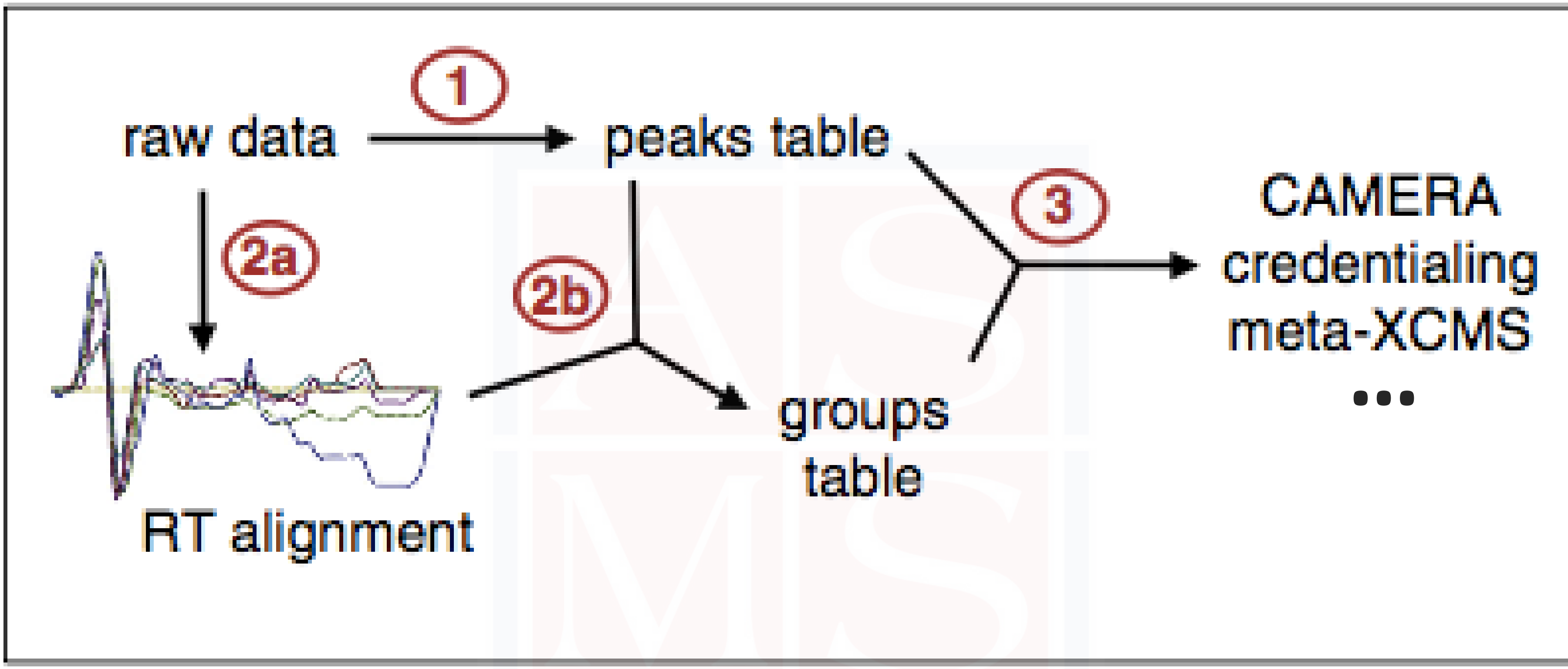
Challenges

- Different labs use different software solutions
- Different software platforms provide different results
- Results are highly setting dependent—poor understanding of the programs and/or data can lead to improper selection (ADV: vendor software)
- Some recent studies suggest using multiple orthogonal platforms
- Many additional software functionalities emerging that have not yet been incorporated into the canned "complete workflows"

Nuts and bolts of data processing



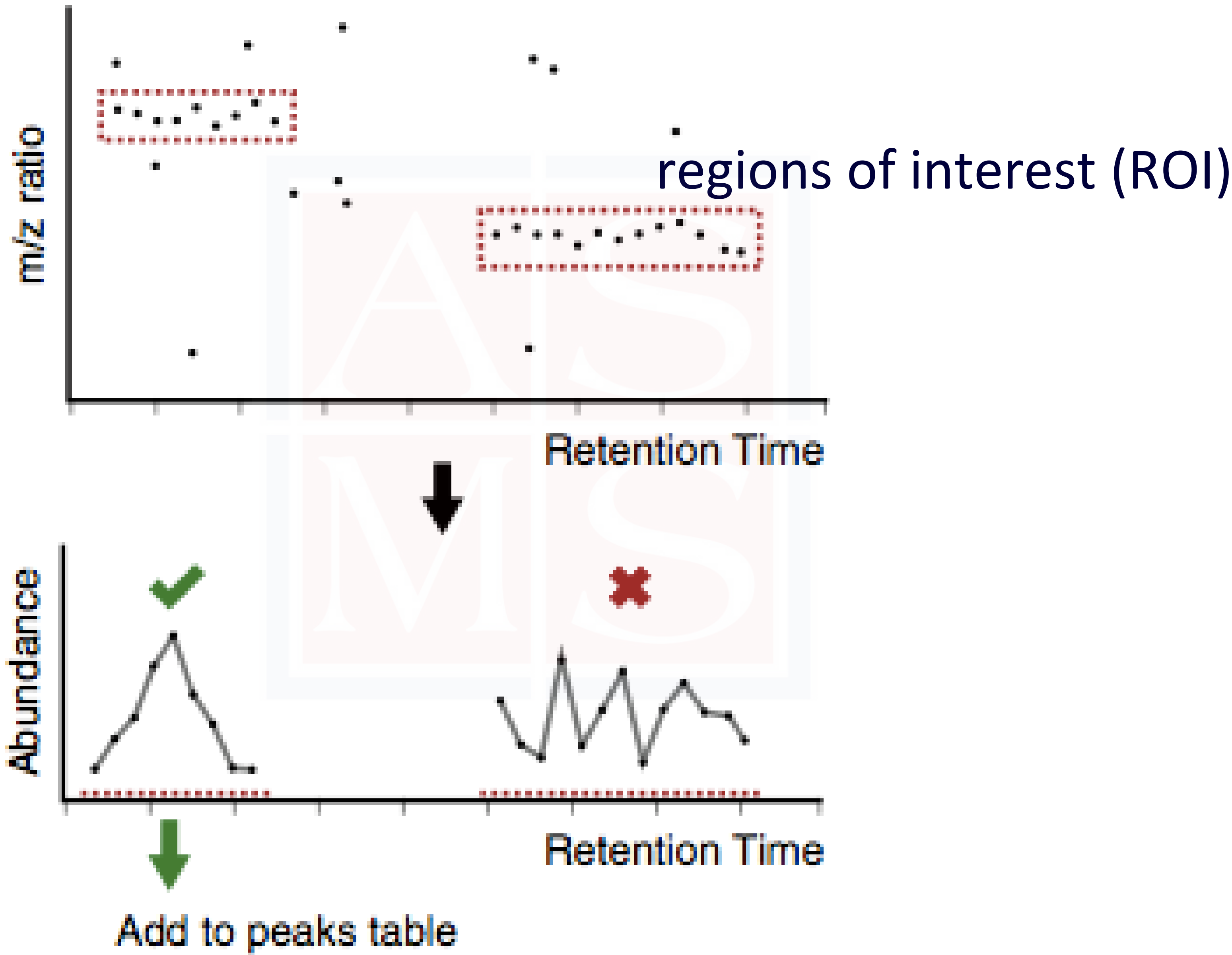
Nuts and bolts of data processing



1. **Feature detection (centWave*)**
2. **Correspondence determination (OBI-warp*)**
3. **Context-dependent analysis**

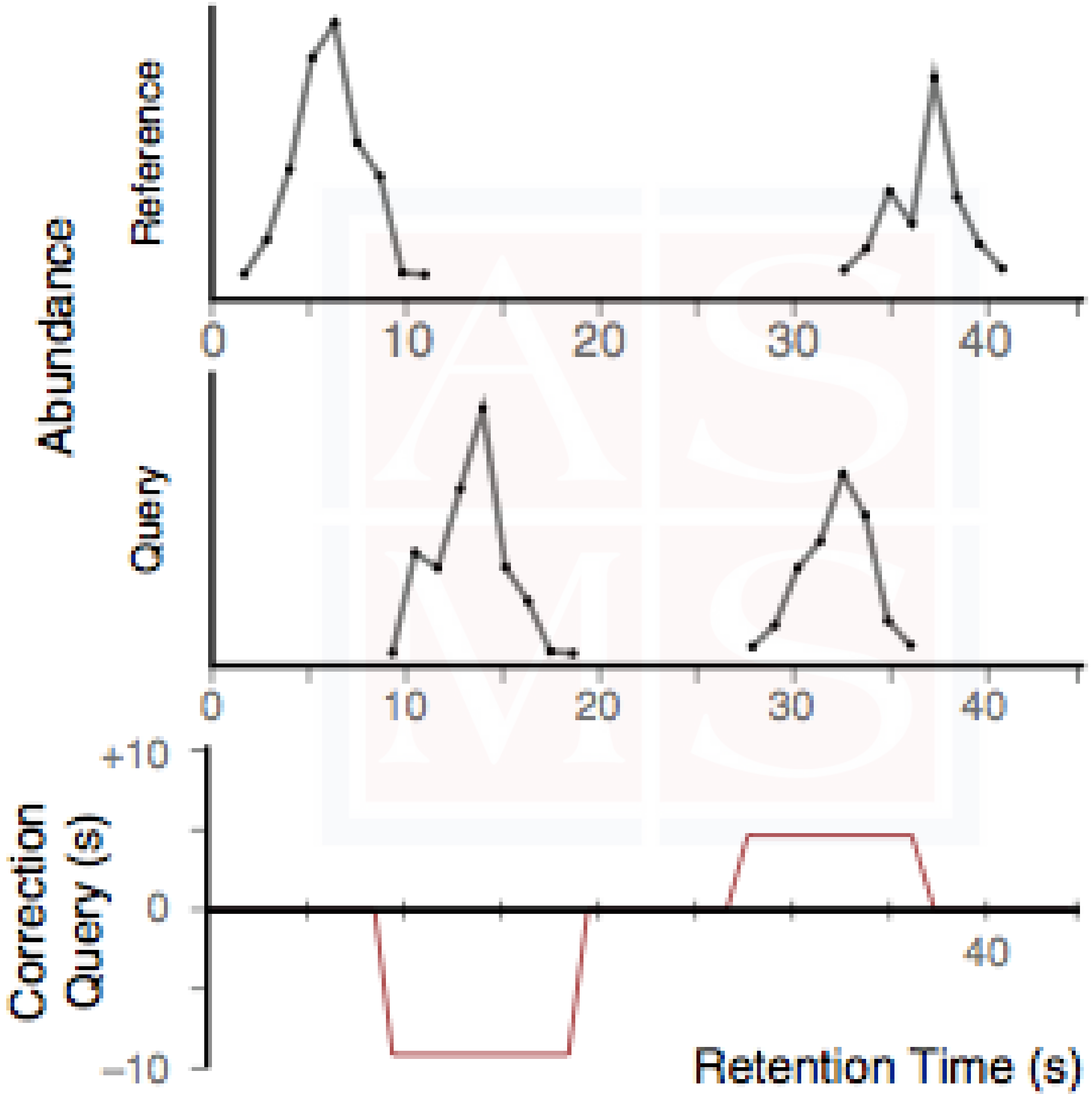
*popular algorithms, but not the only algorithms

Peak detection by centWave



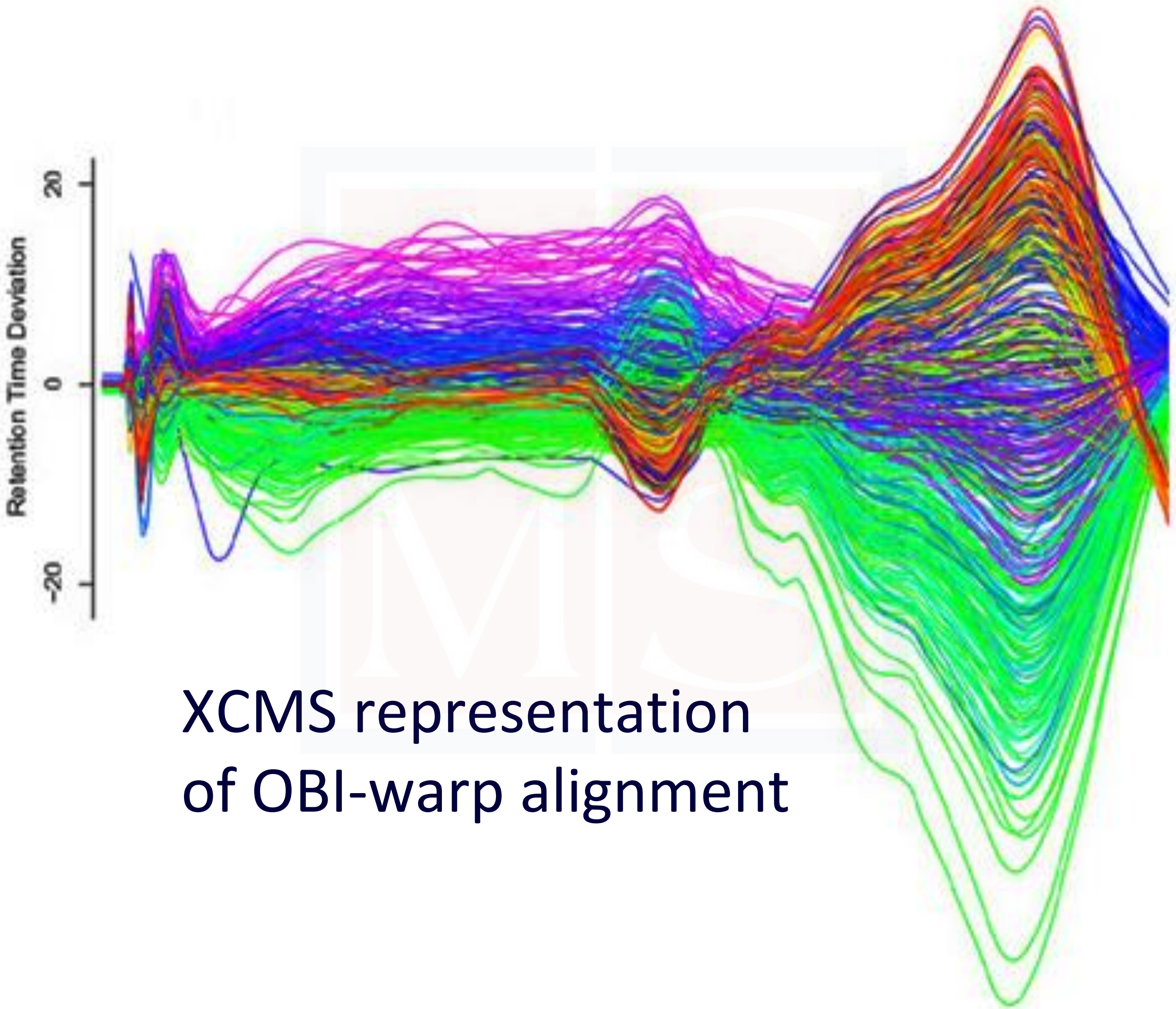
Tautenhahn, BMC Bioiformatics, 2008

Correspondence by OBI-warp



Prince and Marcotte, Anal Chem, 2006

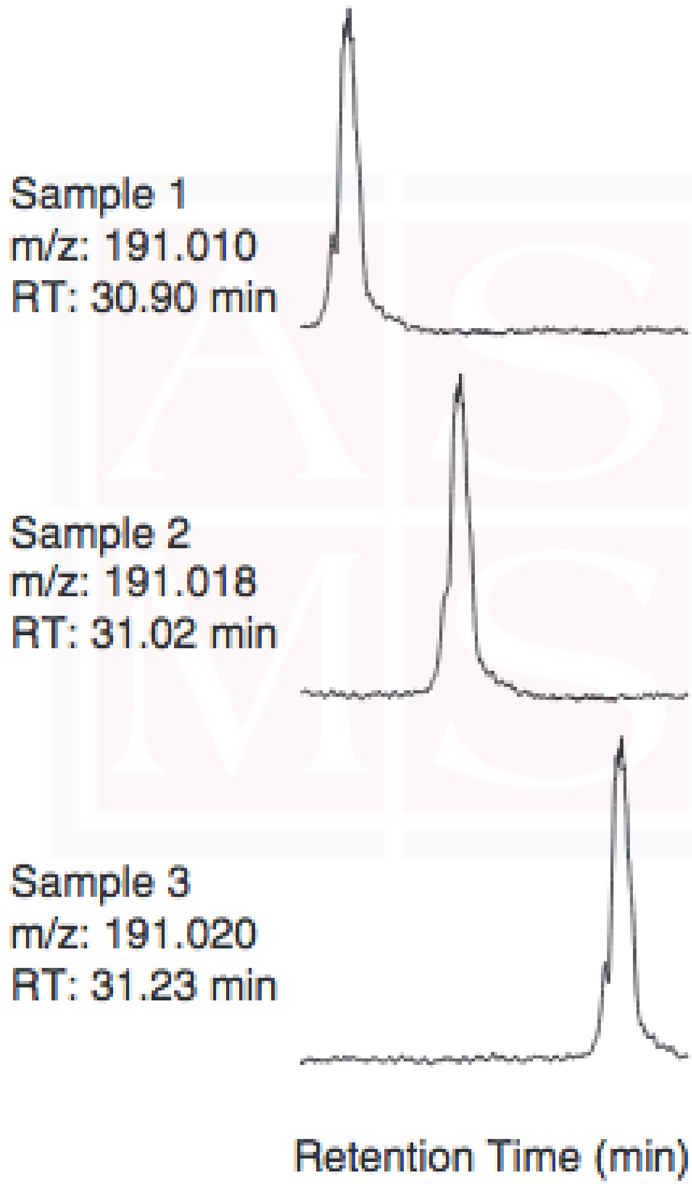
Correspondence by OBI-warp



XCMS representation
of OBI-warp alignment

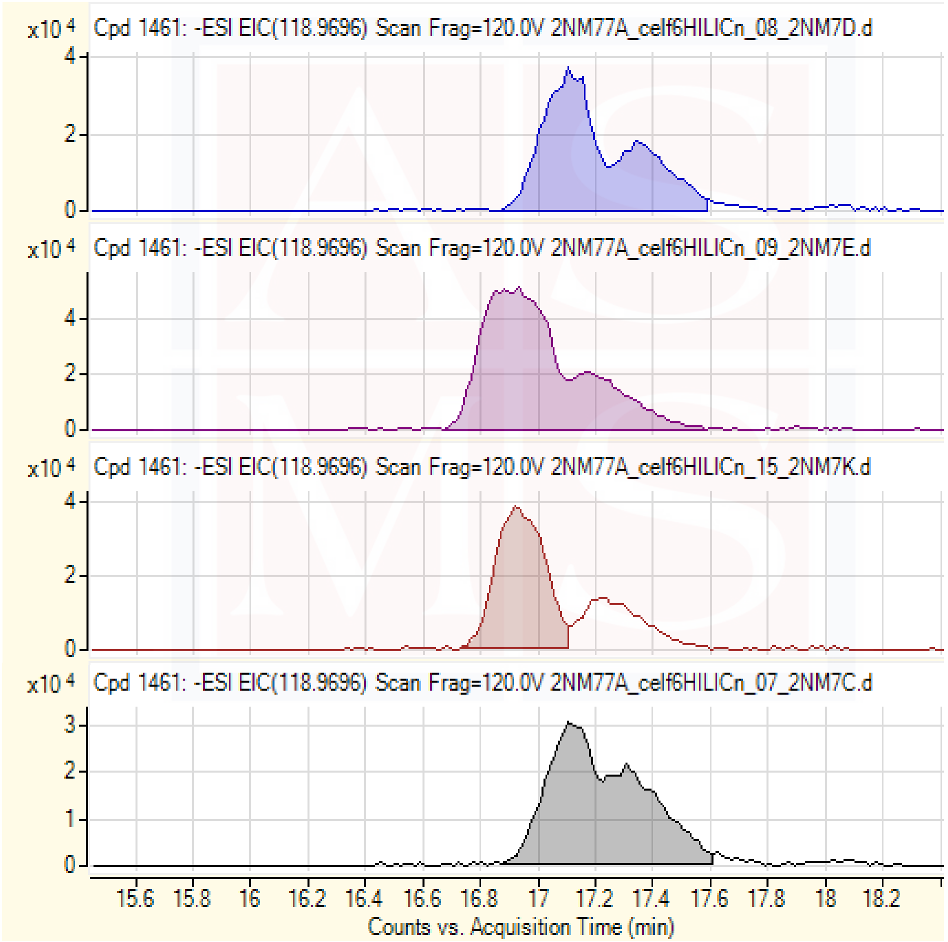
Prince and Marcotte, Anal Chem, 2006

Correspondence by OBI-warp

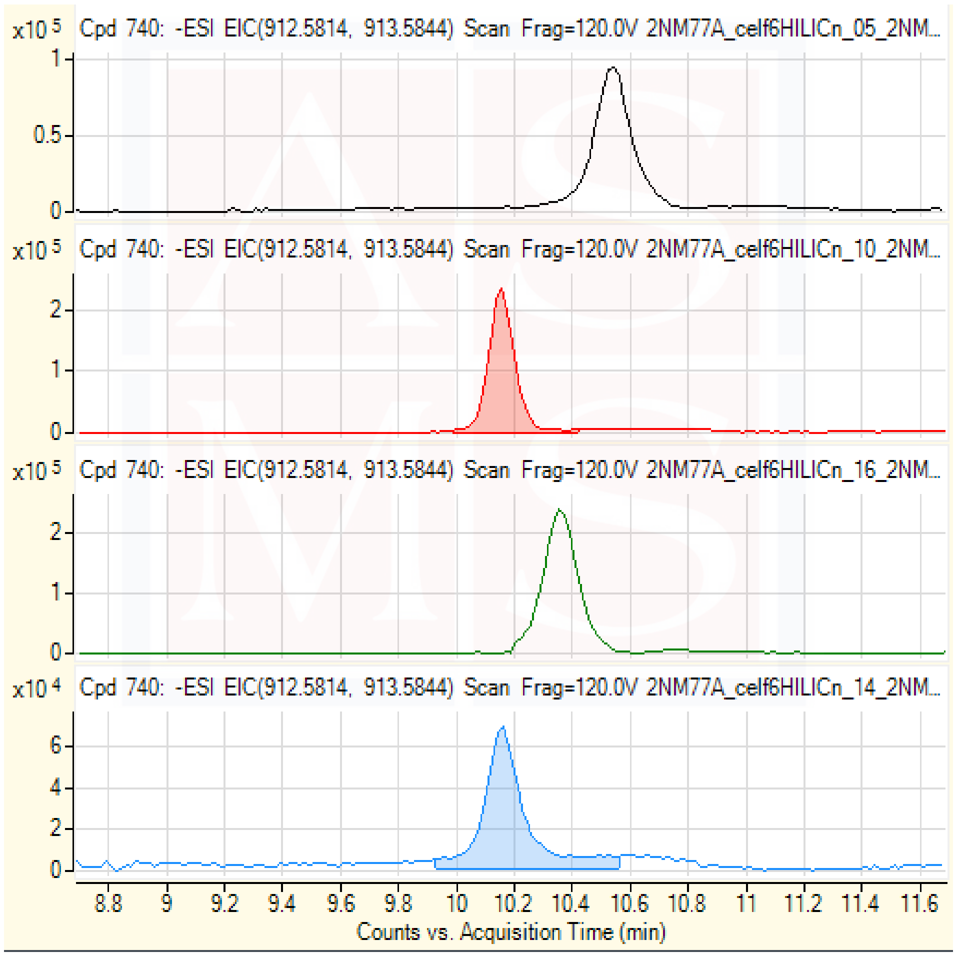


Informatic challenges: bounds

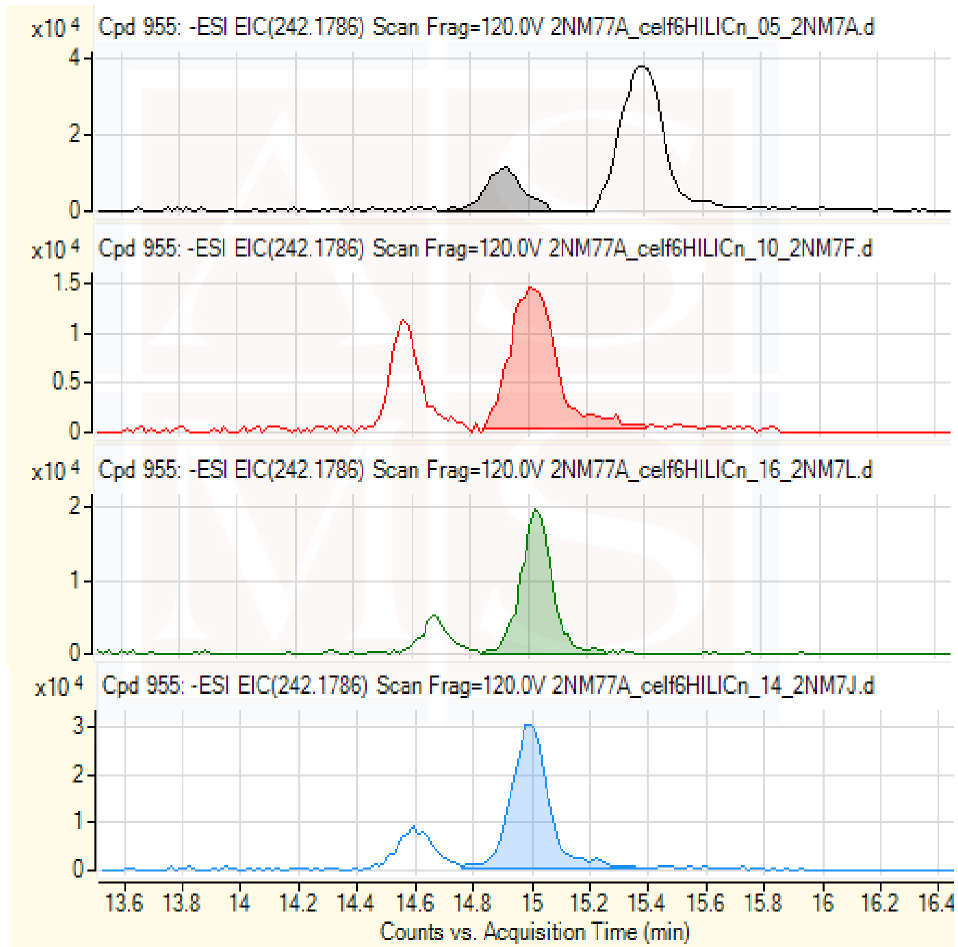
(NOTE: does not occur for all peaks but representative of challenges with *some* peaks, challenging for all software including vendors)



Informatic challenges: missing peaks



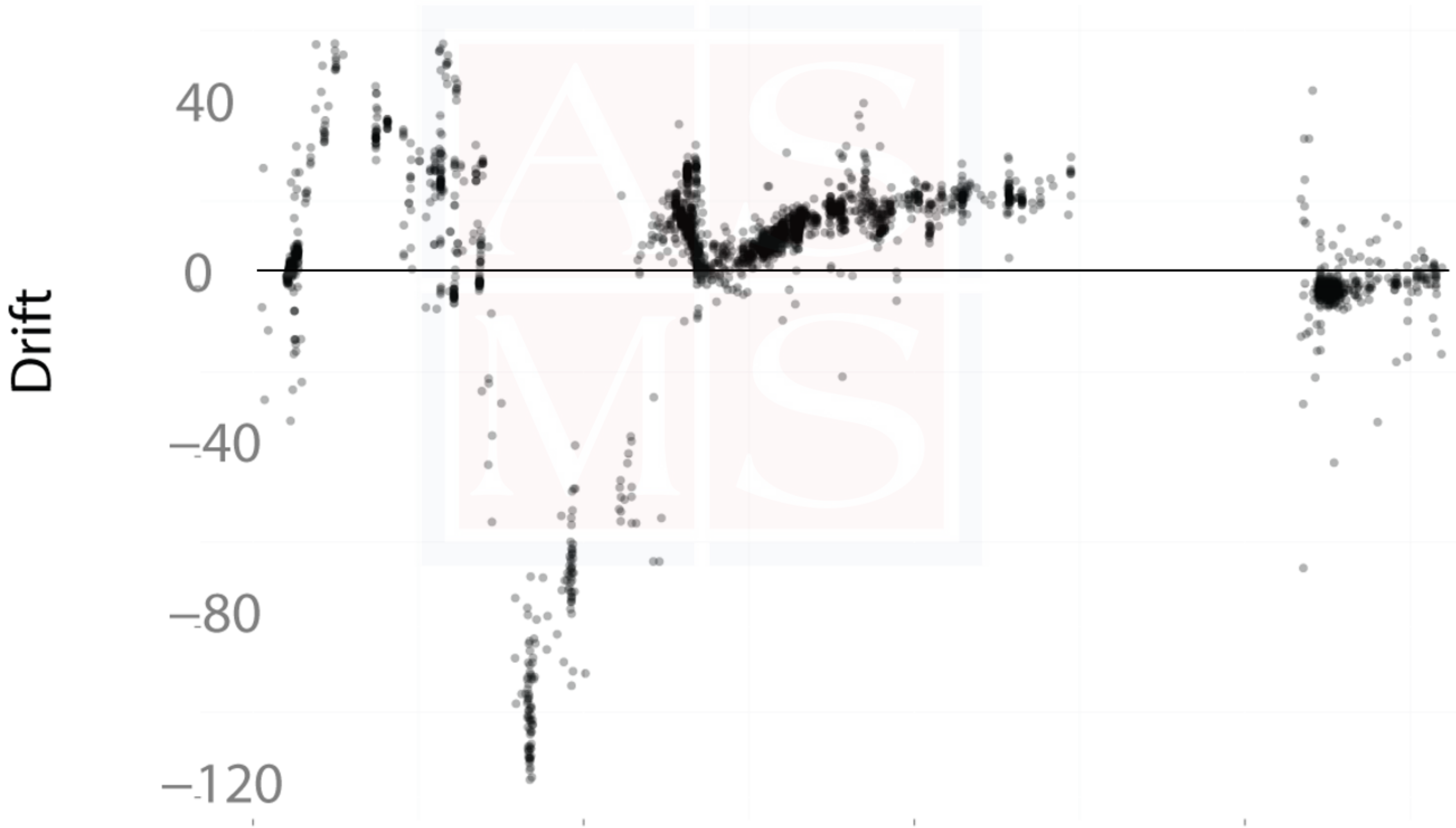
Informatic challenges: correspondence



Informatic challenges

Informatic challenges before correction

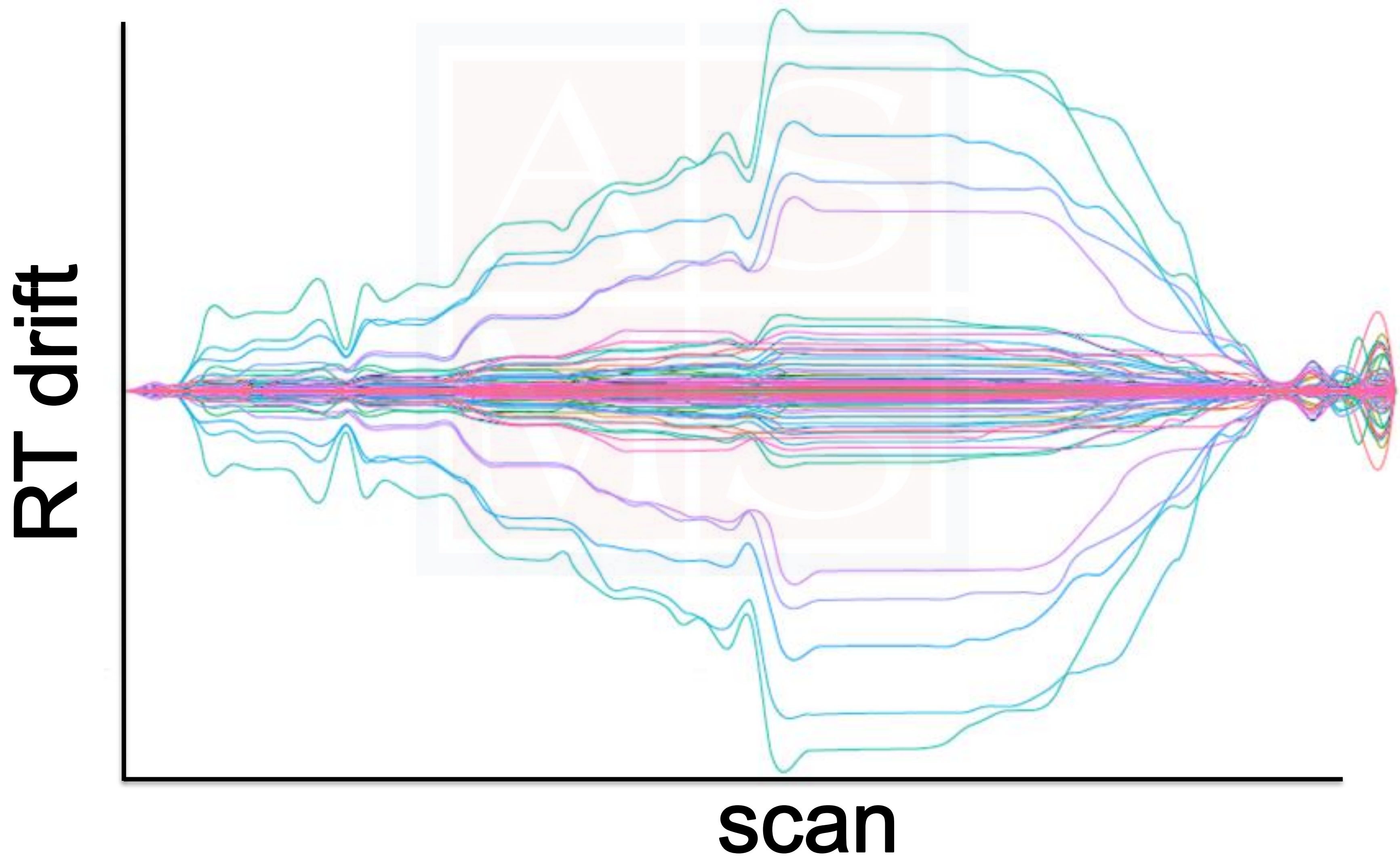
Informatic challenges before correction



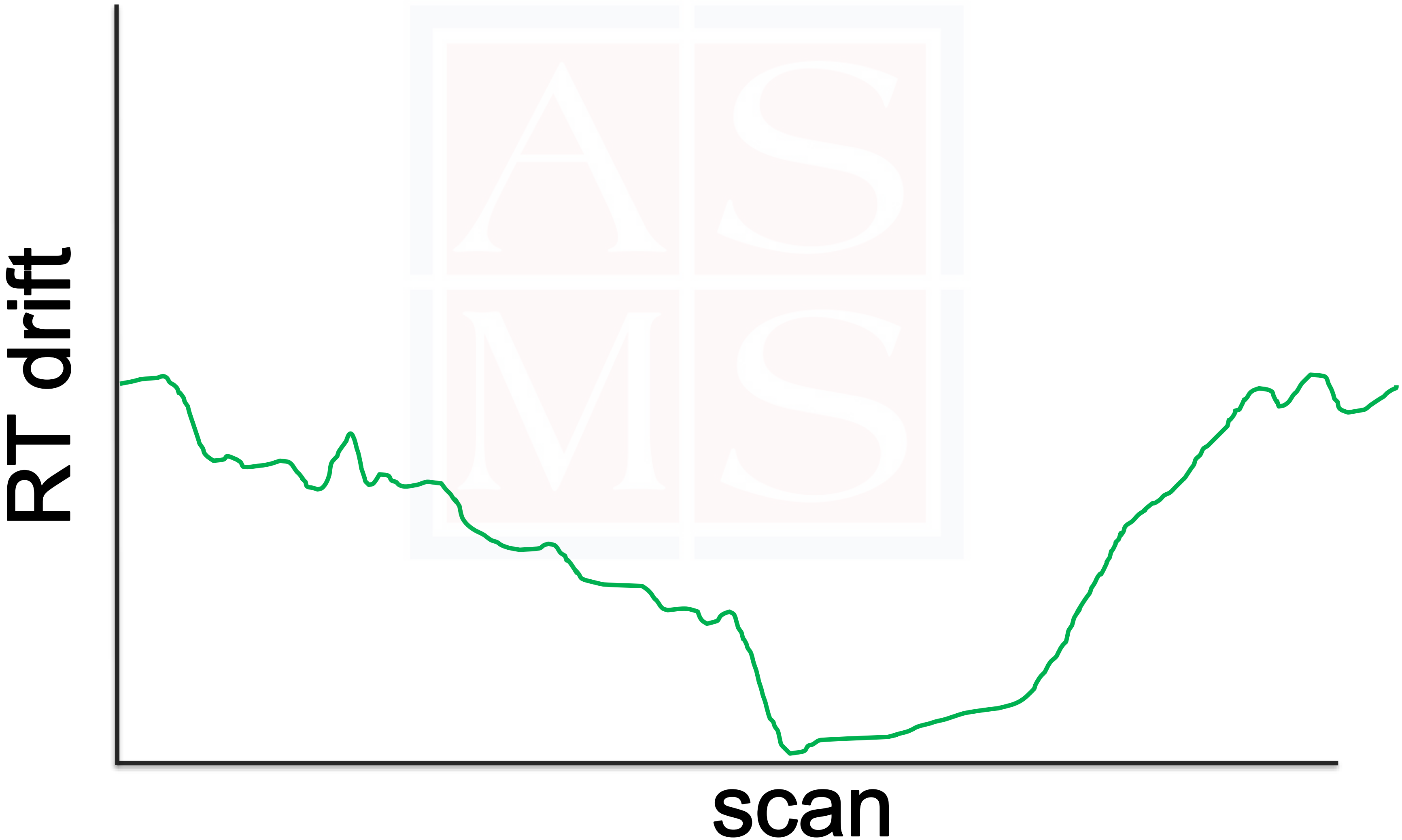
Informatic challenges

Informatic challenges
RT correction

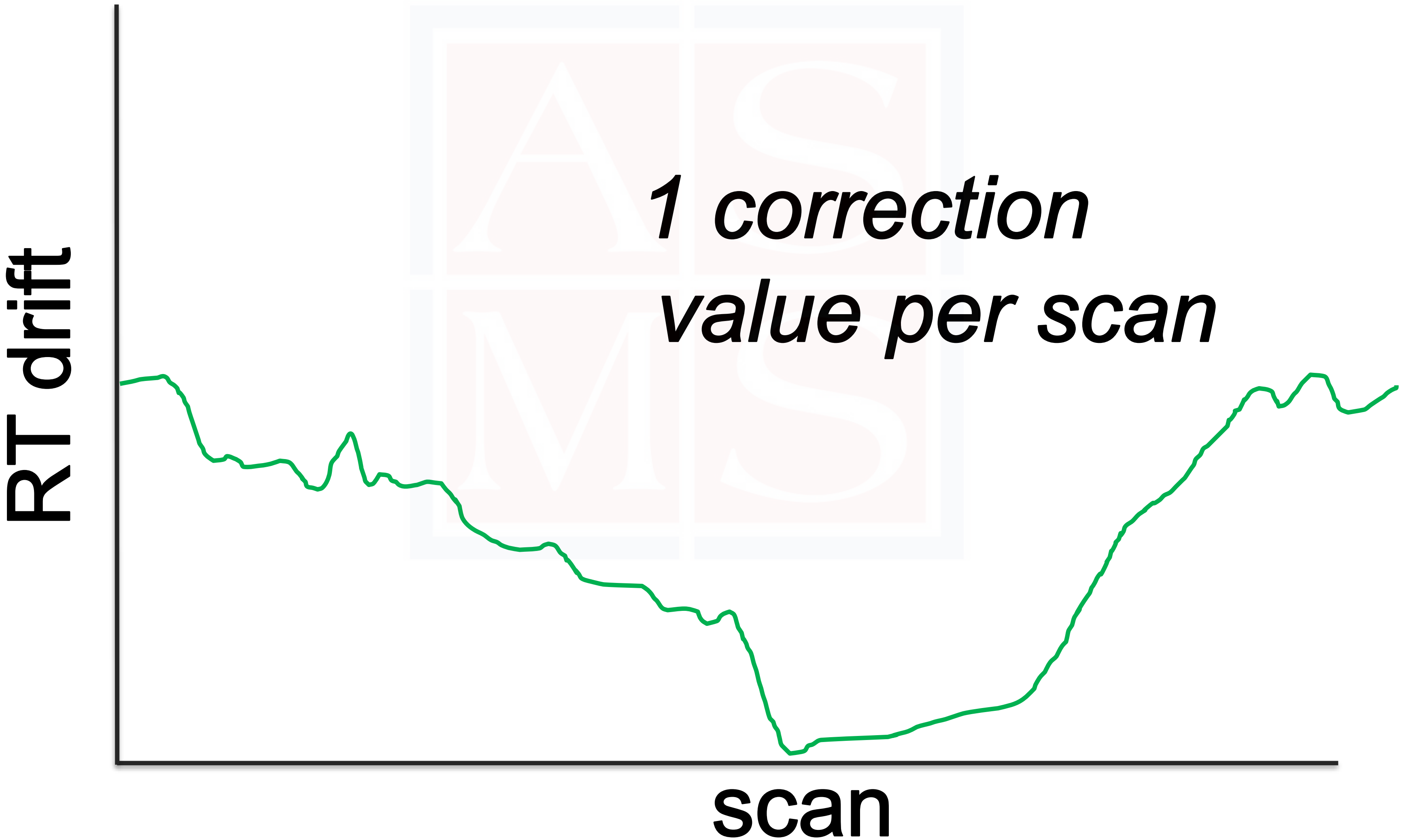
Informatic challenges
RT correction



Informatic challenges
RT correction



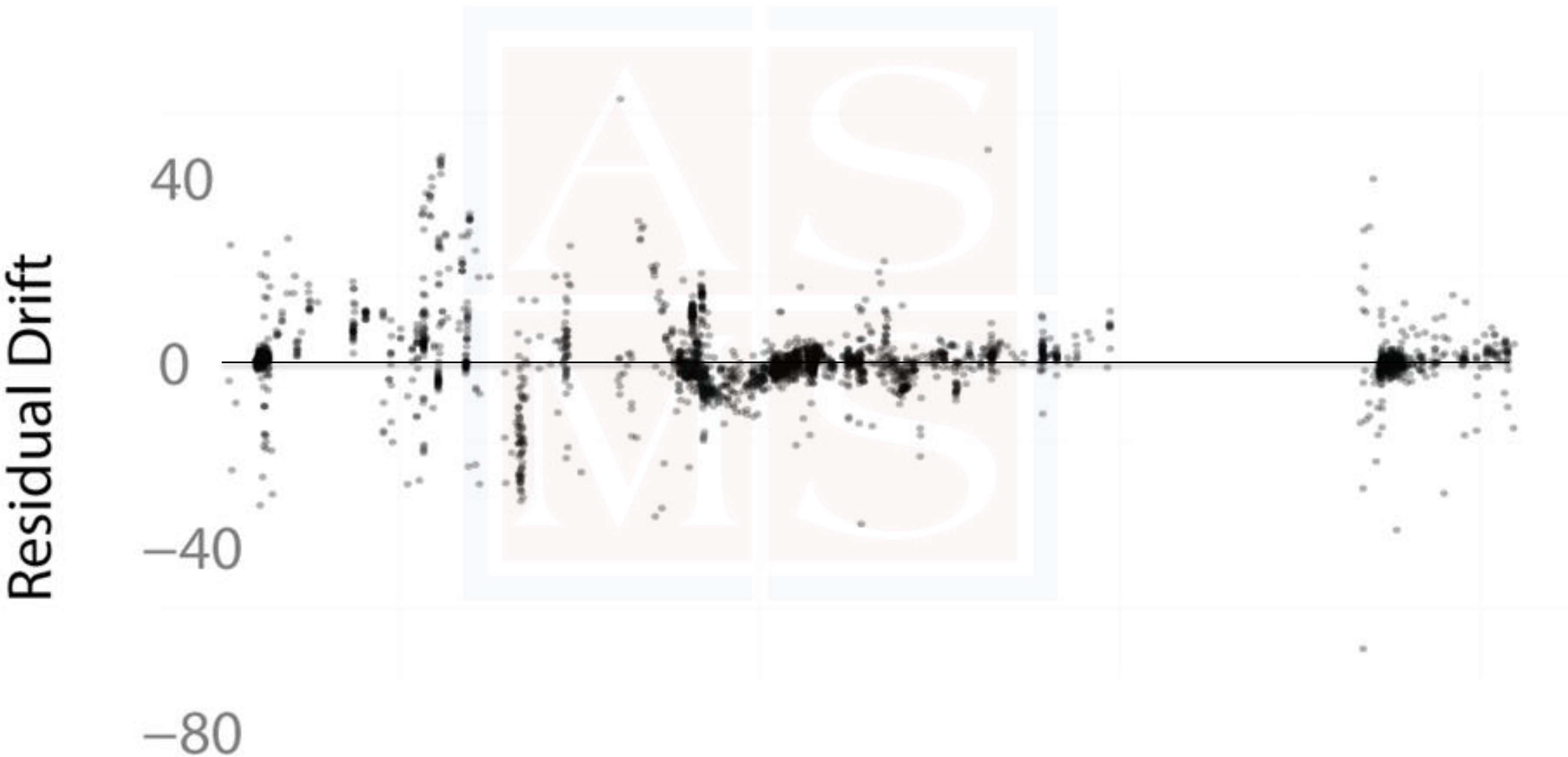
Informatic challenges
RT correction



Informatic challenges

Informatic challenges after correction

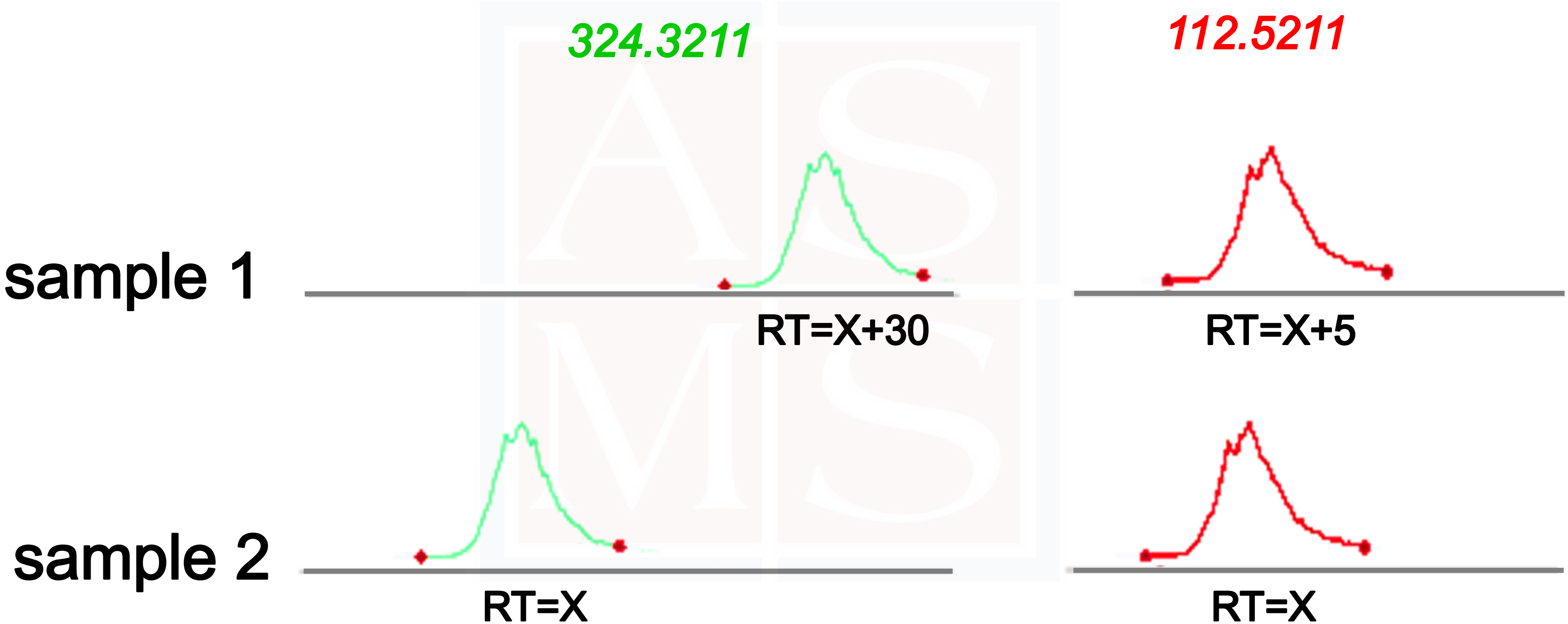
Informatic challenges after correction



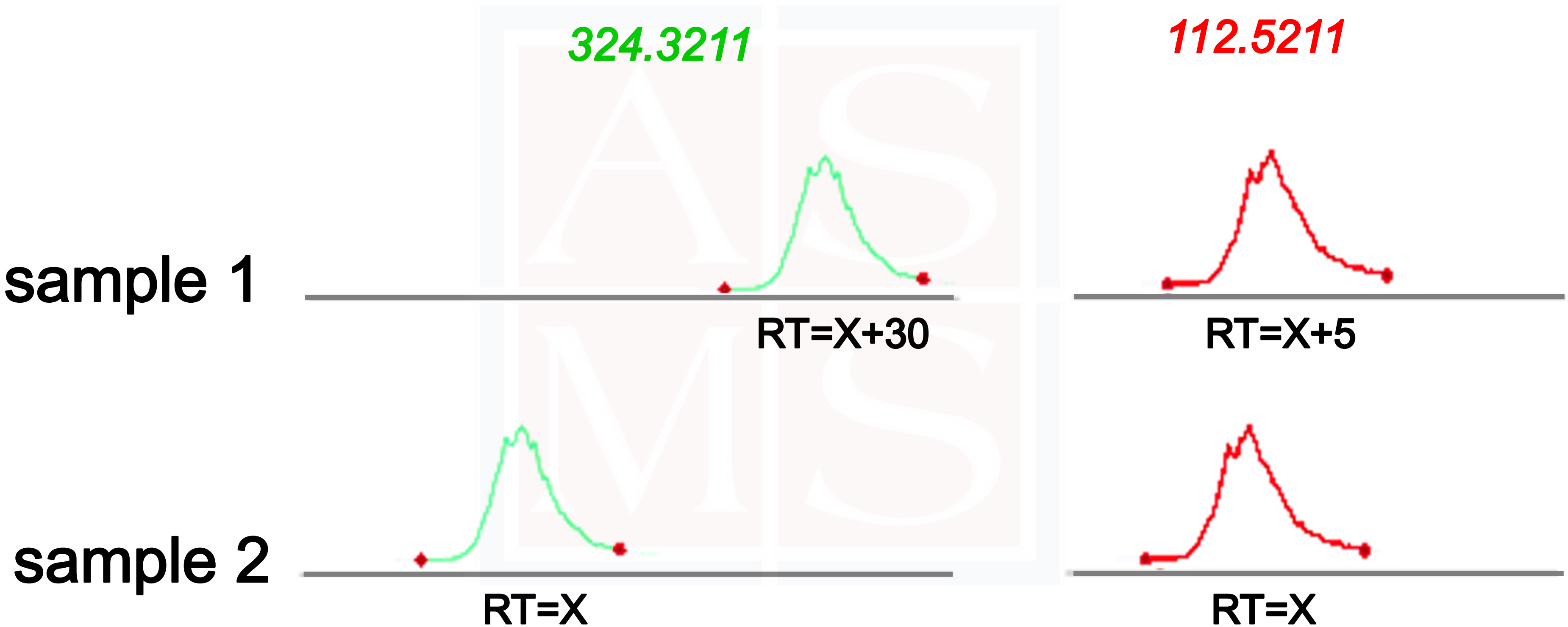
Informatic challenges

Informatic challenges compound-specific drift

Informatic challenges compound-specific drift



Informatic challenges compound-specific drift



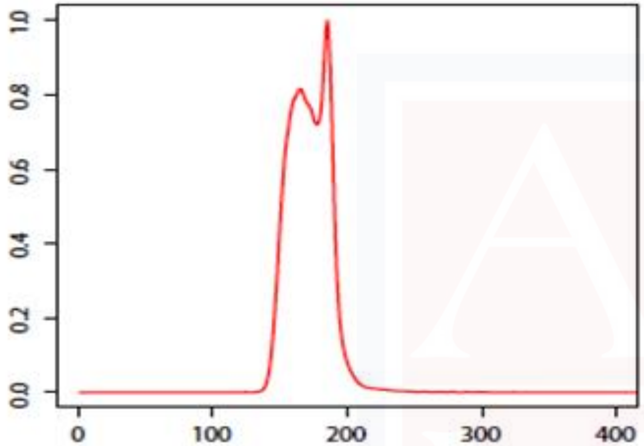
Do you shift the peaks at retention time x in sample 2 by 30 seconds or 5 seconds?

Warpgroup

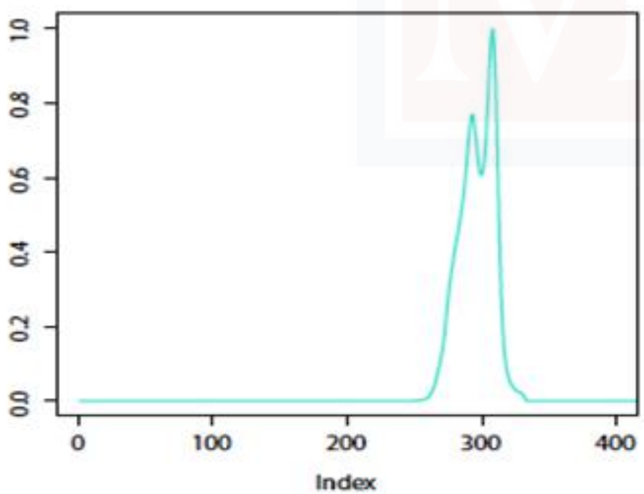
Warpgroup

before

reference



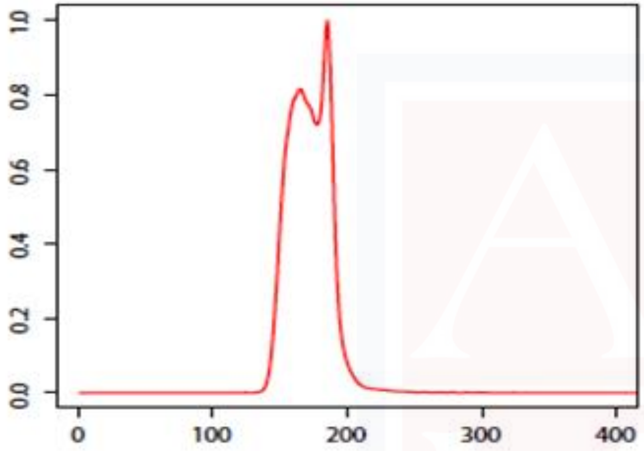
query



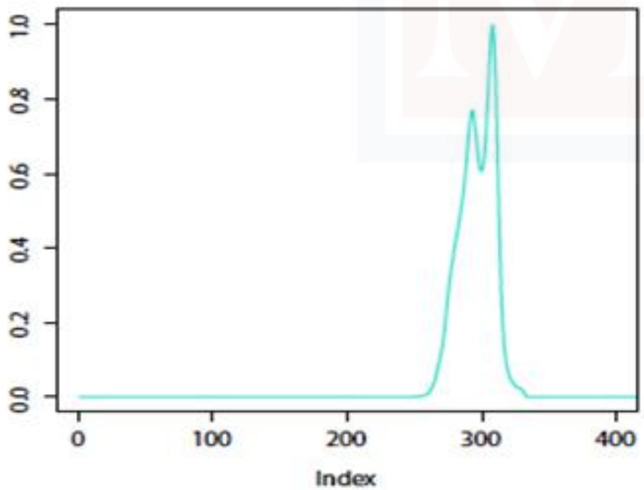
Warpgroup

before *after*

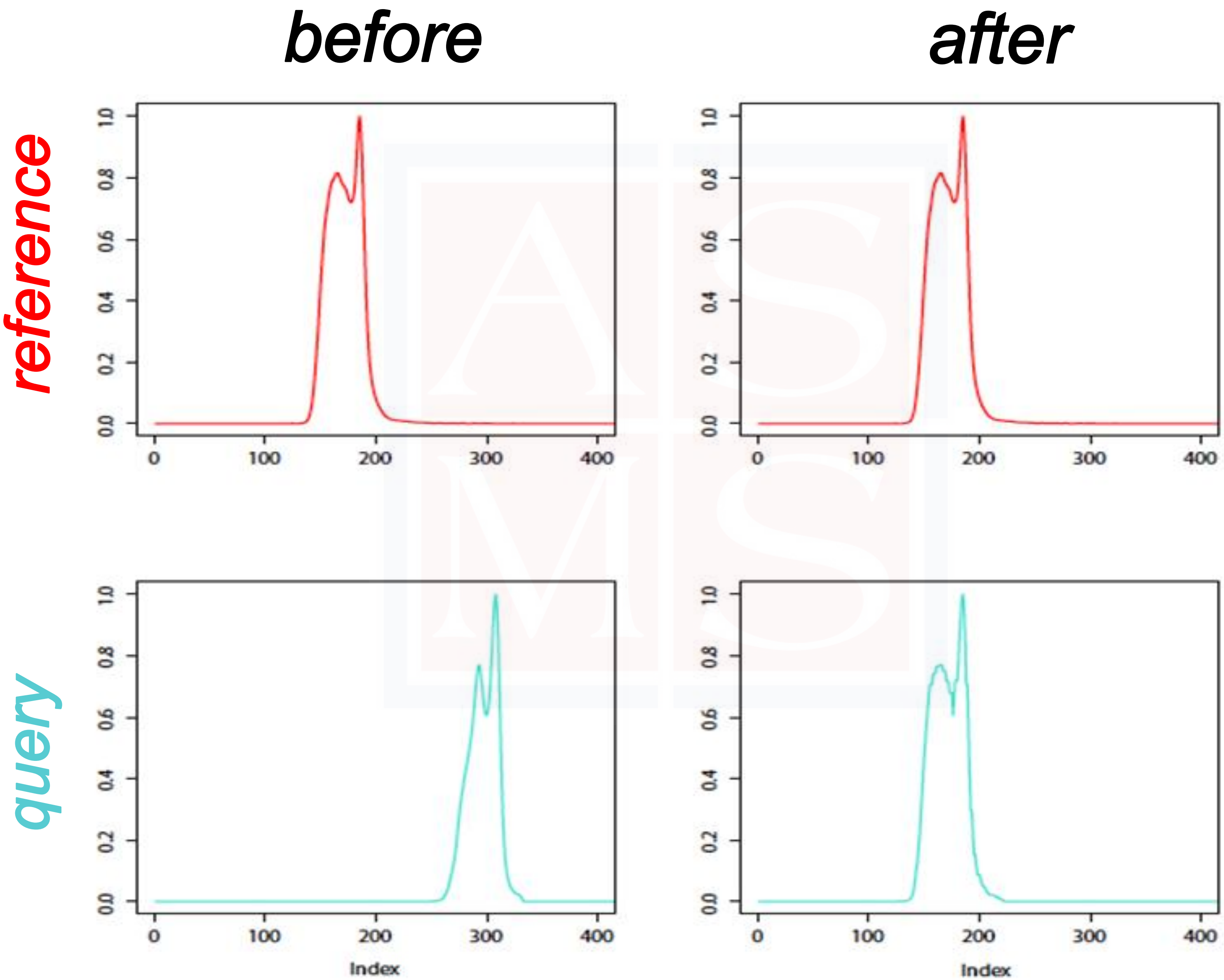
reference



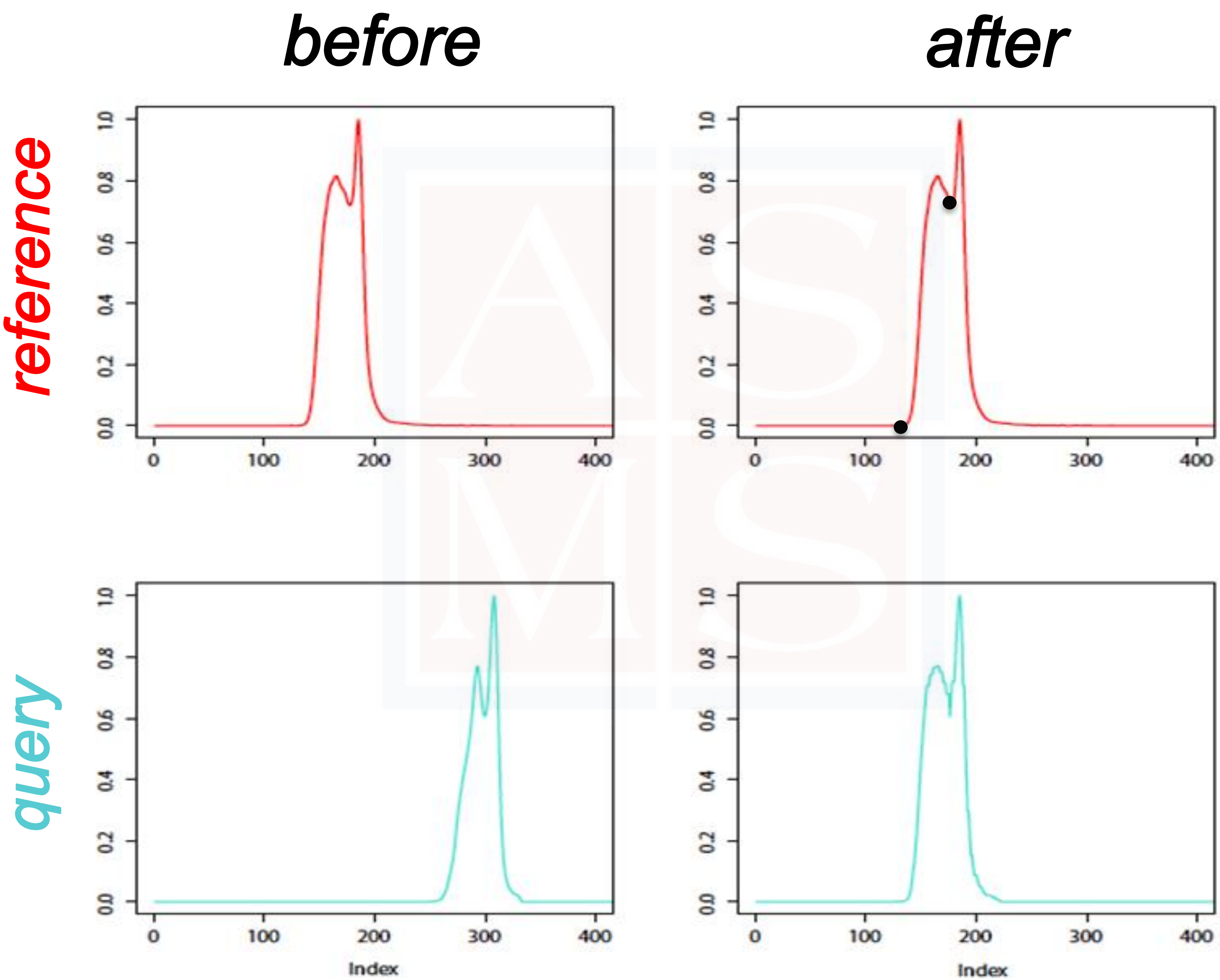
query



Warpgroup

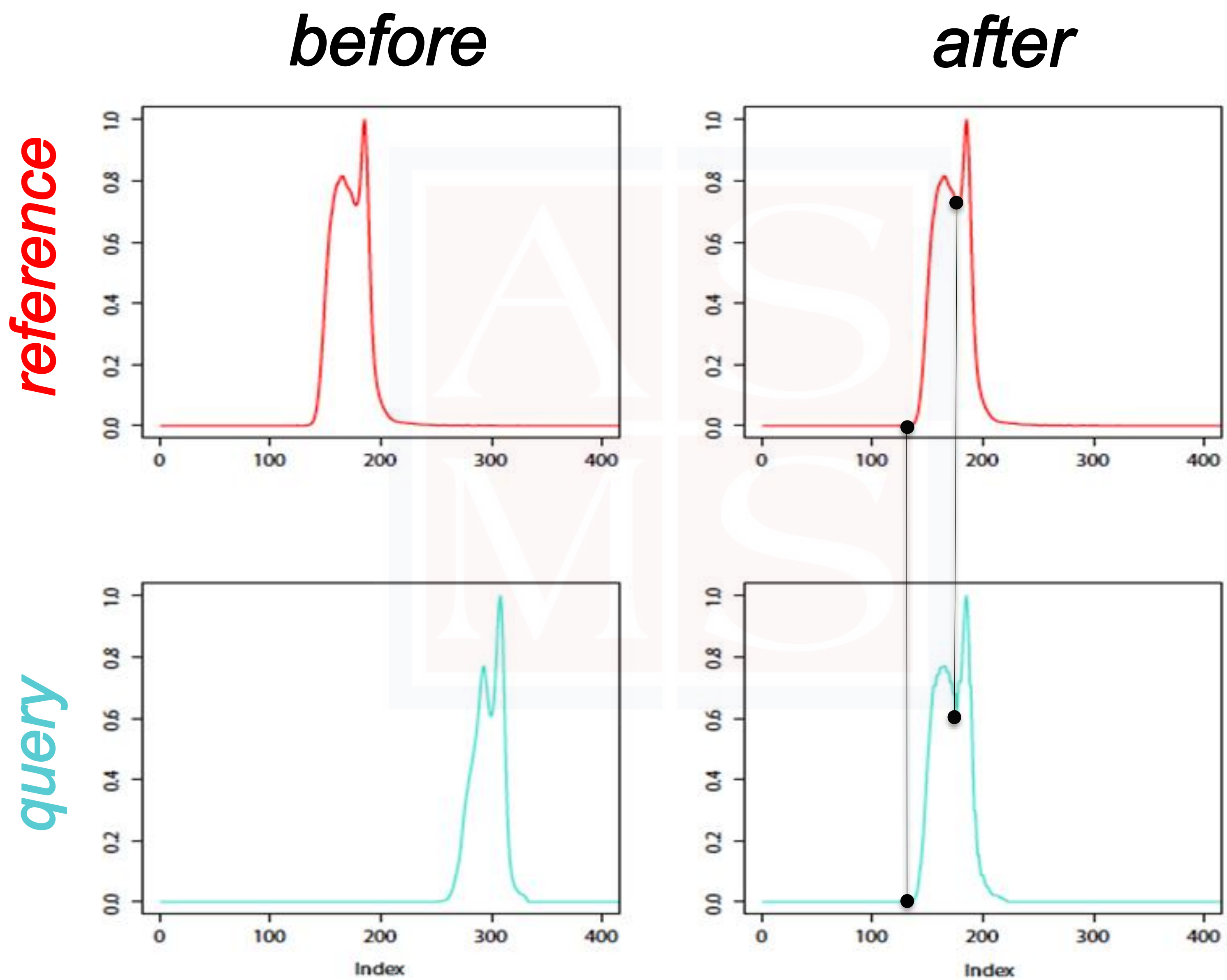

before
after
reference
query
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
Index

Warpgroup



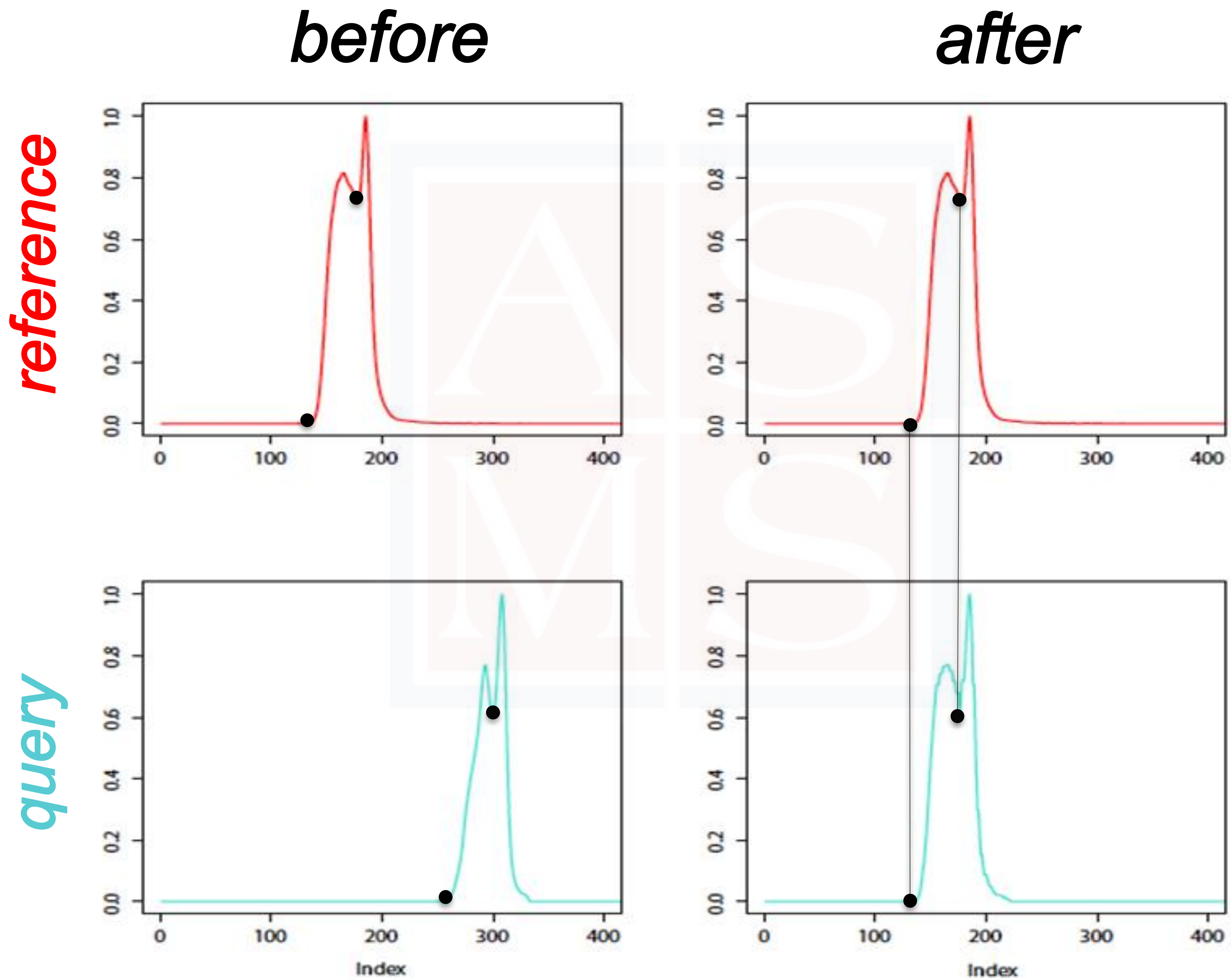
before
after
reference
query
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Index

Warpgroup

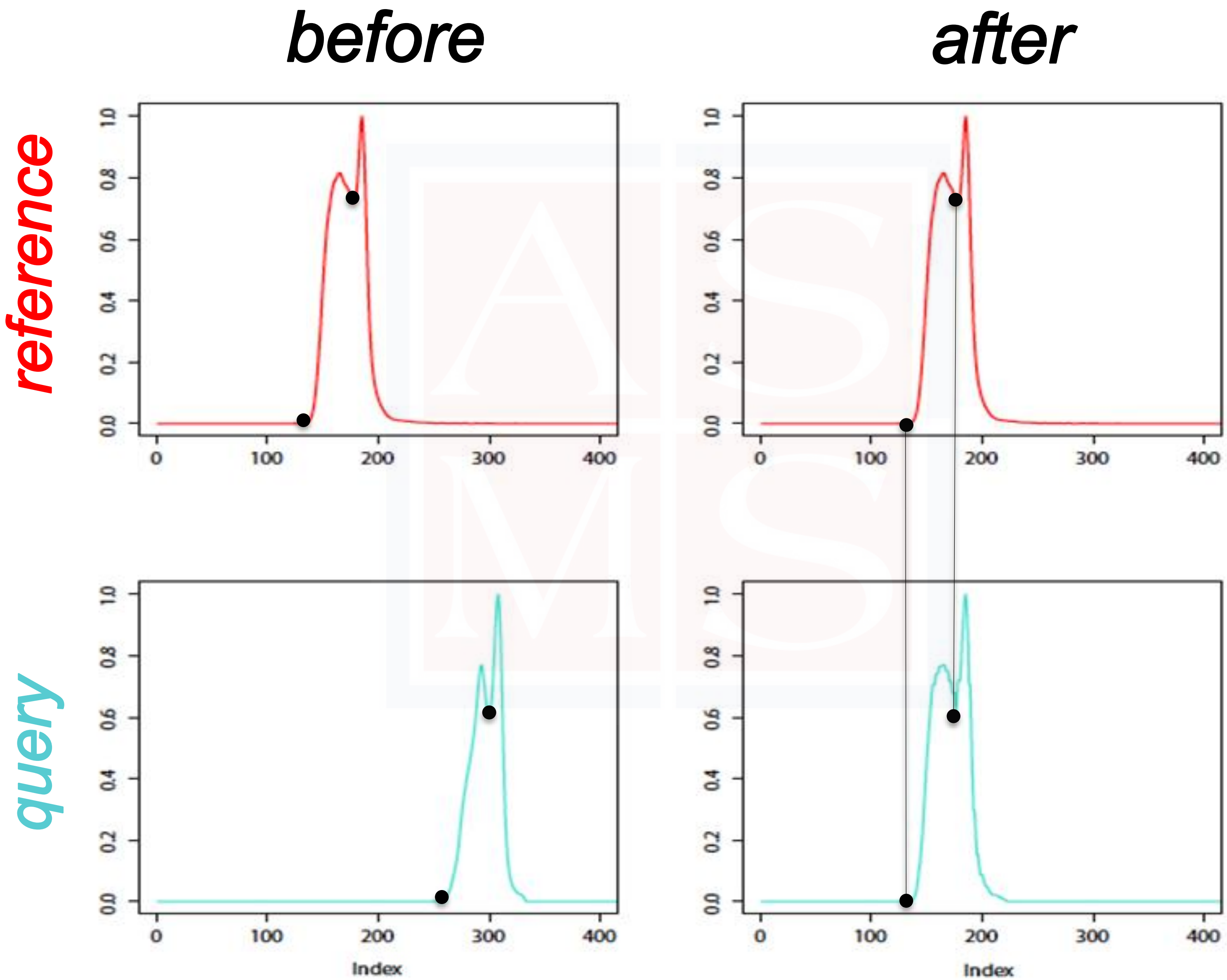


before
after
reference
query
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Index

Warpgroup


before
after
reference
query
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Index

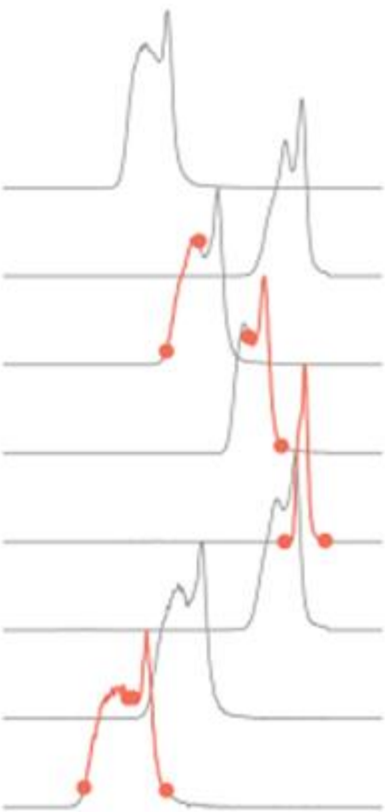
Warpgroup



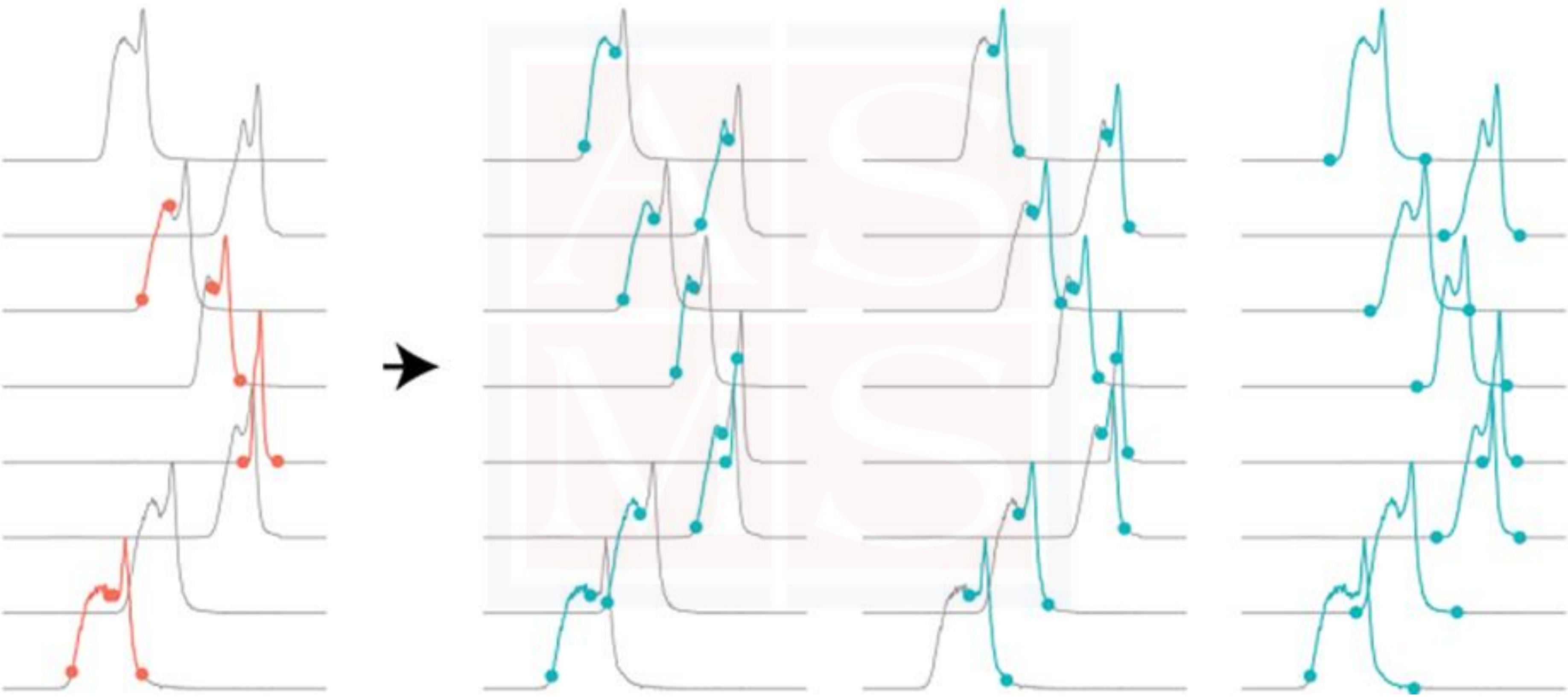
Mahieu, Bioinformatics, 2016

Warpgroup

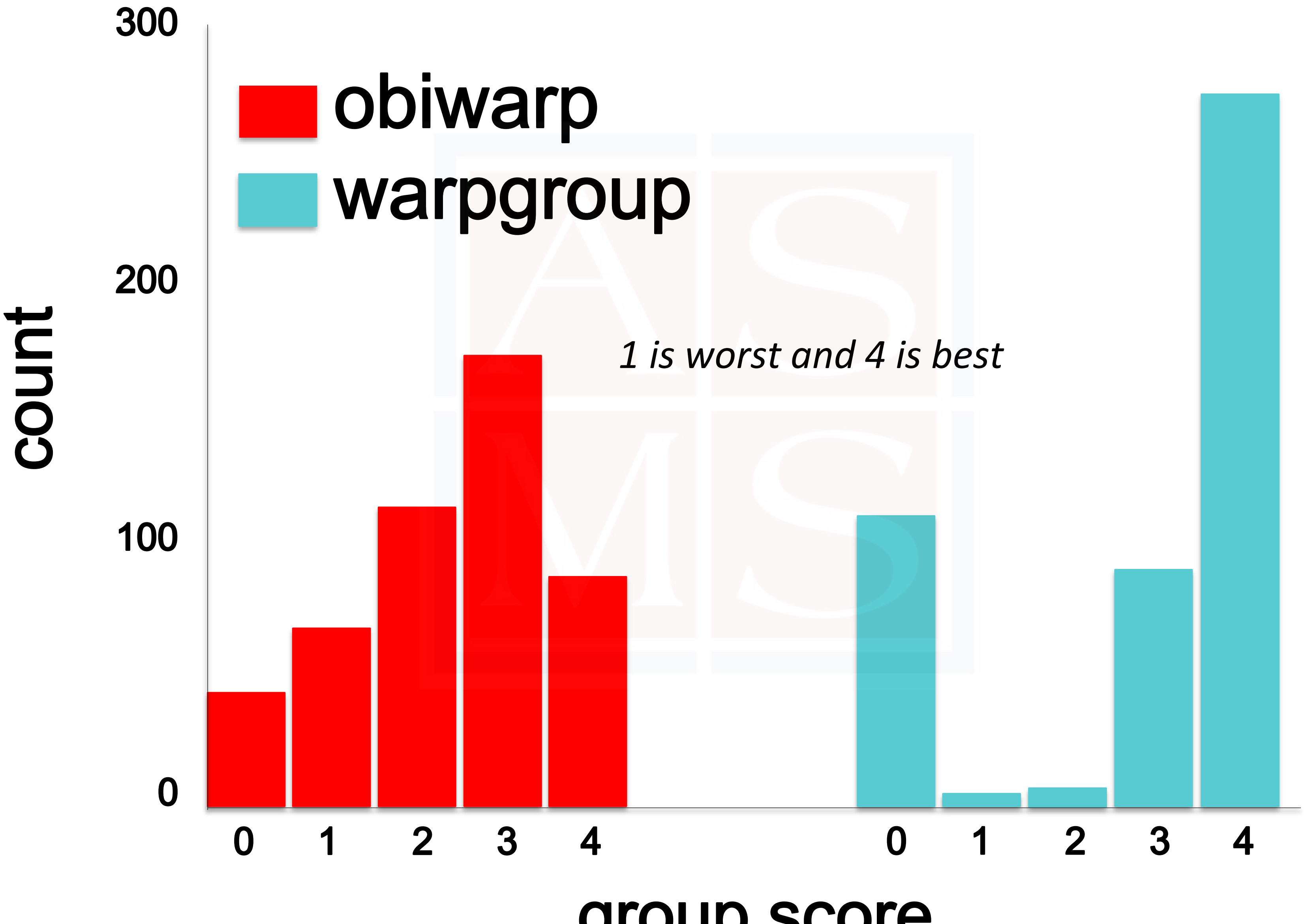
Warpgroup



Warpgroup



Warpgroup



Mahieu, Bioinformatics, 2016

Informatics can be divided into two steps

1. Processing raw metabolomic data (software is required)

2. Analyzing software results (use databases)

Informatics can be divided into two steps

1. Processing raw metabolomic data (software is required)

2. Analyzing software results (use databases)

Informatics can be divided into two steps

1. Processing raw metabolomic data (software is required)

2. Analyzing software results (use databases)

Ultimate goal is to unambiguously ID metabolites

Table 1. New confidence levels of compound annotations, as discussed by the Compound Identification work group of the Metabolomics Society at the 2017 annual meeting of the Metabolomics Society (Brisbane, Australia). The new addition refers to the 'Level 0' annotation; other levels remain as discussed by the Metabolomics Standards Initiative.

Confidence Level	Description	Minimum Data Requirements
Level 0	Unambigous 3D structure: Isolated, pure compound, including full stereochemistry	Following natural product guidelines, determination of 3D structure
Level 1	Confident 2D structure: uses reference standard match or full 2D structure elucidation	At least two orthogonal techniques defining 2D structure confidently, such as MS/MS and RT or CCS
Level 2	Probable structure: matched to literature data or databases by diagnostic evidence	At least two orthogonal pieces of information, including evidence that excludes all other candidates
Level 3	Possible structure or class: Most likely structure, isomers possible, substance class or substructure match	One or several candidates possible, requires at least one piece of information supporting the proposed candidate
Level 4	Unkown feature of insterest:	Presence in sample

Blazenovic et al., Metabolites 2018

Ultimate goal is to unambiguously ID metabolites

Table 1. New confidence levels of compound annotations, as discussed by the Compound Identification work group of the Metabolomics Society at the 2017 annual meeting of the Metabolomics Society (Brisbane, Australia). The new addition refers to the 'Level 0' annotation; other levels remain as discussed by the Metabolomics Standards Initiative.

Confidence Level	Description	Minimum Data Requirements
Level 0	Unambigous 3D structure: Isolated, pure compound, including full stereochemistry	Following natural product guidelines, determination of 3D structure
Level 1	Confident 2D structure: uses reference standard match or full 2D structure elucidation	At least two orthogonal techniques defining 2D structure confidently, such as MS/MS and RT or CCS
Level 2	Probable structure: matched to literature data or databases by diagnostic evidence	At least two orthogonal pieces of information, including evidence that excludes all other candidates
Level 3	Possible structure or class: Most likely structure, isomers possible, substance class or substructure match	One or several candidates possible, requires at least one piece of information supporting the proposed candidate
Level 4	Unkown feature of insterest:	Presence in sample

Blazenovic et al., Metabolites 2018

Ultimate goal is to unambiguously ID metabolites

level	atoms	atomic connections	relative stereo-chemistry	chirality	metabolite ID class	metabolite information level
A	✓	✓	✓	✓	• a single, defined enantiomer • or a single, defined achiral metabolite	• known molecular formula • known structure • known stereochemistry if present • known chirality if present • e.g. tartaric acid (2R,3R)-2,3-dihydroxybutanoic acid
B	✓	✓	✓	X	• one of two enantiomers	• known molecular formula • known structure • known relative stereochemistry if present • unknown chirality • e.g. tartaric acid: enantiomer undefined (2R,3R or 2S, 3S)
C	✓	✓	X	X	• one of a number of stereoisomers • e.g. E/Z geometric or cis-/trans- ring isomers	• known molecular formula • known molecular structure • unknown relative stereochemistry • e.g. tartaric acid diastereomer undefined: 2R, 3S, 2R, 3R or 2S, 3S • oleic acid (Z-isomer) or elaidic acid (E-isomer of 9-octadecanoic acid)
D	✓	X	X	X	• one of a number of positional isomers	• known molecular formula • known functional groups • unknown structure • e.g. $C_{18}H_{34}O_2$: 9-octadecanoic acid or 7-octadecanoic acid
E	✓	X	X	X	• one of a number of possible compounds of known molecular formula	• known molecular formula • unknown structure • e.g. C_2H_6O: dimethylether or ethanol
F	X	X	X	X	• specific spectral features defining a structural class	• unknown molecular formula • known structural class
G	X	X	X	X	• specific spectral features	• unknown molecular formula

Proposed reporting standards for metabolite annotation and identification (proposed by the Metabolite Identification Task Group of the Metabolomics Society)

Ultimate goal is to unambiguously ID metabolites

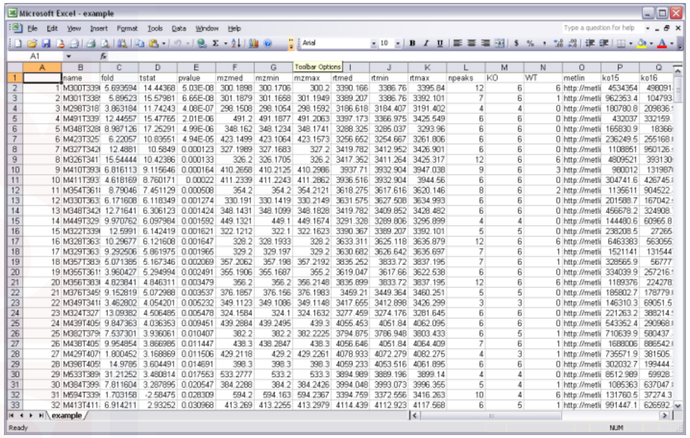
level	atoms	atomic connections	relative stereo-chemistry	chirality	metabolite ID class	metabolite information level
A	✓	✓	✓	✓	• a single, defined enantiomer • or a single, defined achiral metabolite	• known molecular formula • known structure • known stereochemistry if present • known chirality if present • e.g. tartaric acid (2R,3R)-2,3-dihydroxybutanoic acid
B	✓	✓	✓	X	• one of two enantiomers	• known molecular formula • known structure • known relative stereochemistry if present • unknown chirality • e.g. tartaric acid: enantiomer undefined (2R,3R or 2S, 3S)
C	✓	✓	X	X	• one of a number of stereoisomers • e.g. E/Z geometric or cis-/trans- ring isomers	• known molecular formula • known molecular structure • unknown relative stereochemistry • e.g. tartaric acid diastereomer undefined: 2R, 3S, 2R, 3R or 2S, 3S • oleic acid (Z-isomer) or elaidic acid (E-isomer of 9-octadecanoic acid)
D	✓	X	X	X	• one of a number of positional isomers	• known molecular formula • known functional groups • unknown structure • e.g. $C_{18}H_{34}O_2$: 9-octadecanoic acid or 7-octadecanoic acid
E	✓	X	X	X	• one of a number of possible compounds of known molecular formula	• known molecular formula • unknown structure • e.g. C_2H_6O: dimethylether or ethanol
F	X	X	X	X	• specific spectral features defining a structural class	• unknown molecular formula • known structural class
G	X	X	X	X	• specific spectral features	• unknown molecular formula

Isolation of metabolite, chiral chromatography, > 2 orthogonal pieces of data with standards

Proposed reporting standards for metabolite annotation and identification (proposed by the Metabolite Identification Task Group of the Metabolomics Society)

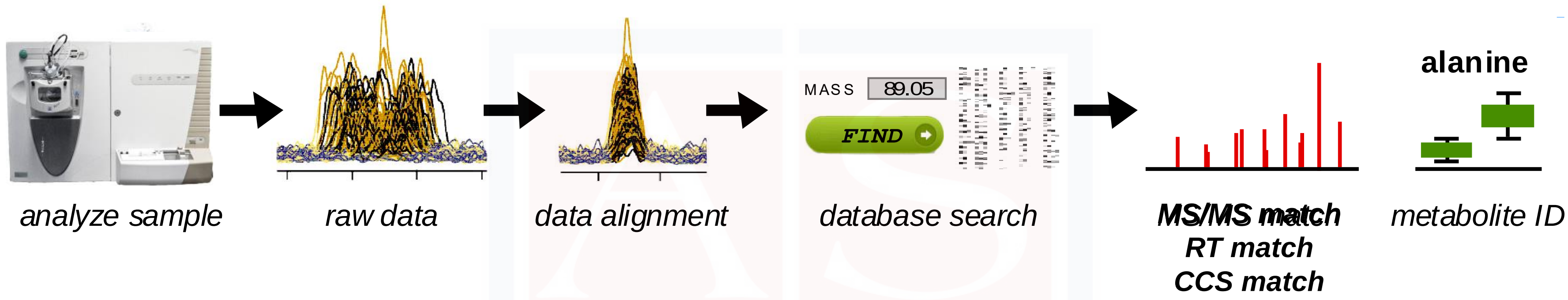
Software typically only provides a list of *m/z* values

- Automated identification workflows are emerging (data-dependent MS/MS, data-independent MS/MS, biology-dependent MS/MS, etc.)
- Conventionally, perform profiling in MS1 and then search m/z values in databases to get "leads"
- Targeted validation of leads by comparing MS2 spectra, retention time, and collision cross section to standards



Barupal et al., BMC Bioinformatics 2012

Identifying metabolomic signals



a “conventional” workflow

Identifying metabolomic signals

Table 2
Summary of the most widely used mass spectral databases in metabolomics

Database	Pros	Cons
HMDB [31]	- Public - Mass spectral data on ~9500 chemical standards - Spectral data are downloadable	- Mixed collision energies and instrument types
METLIN [53]	- Public - Curated mass spectral data on >13,000 chemical standards - Over 63,500 high-resolution MS/MS spectra	- Only Q-TOF data - Spectral data are not downloadable
LipidSearch [92]	- Over 1.5 million lipid ions and their predicted fragment ions - Includes lipid adduct ions and MS^n fingerprints - Data are stored in XML files	- Commercial license required - Developed for Orbitrap technology - *In silico* generated MS/MS library - Overlap with LipidBlast is unclear
LipidBlast [62]	- Over 200,000 tandem mass spectra covering 25 lipid classes - Publicly available - Spectral data are downloadable	- *In silico* generated library using heuristic modeling of tandem mass spectra - "One-third rule" limitation: developed with mostly ion-trap tandem mass spectra - Does not allow batch search of precursor ions - Overlap with LipidSearch unclear
LipidMaps [60]	- Over 40,000 unique lipid structures - Spectral data are downloadable	- MS/MS spectra only predicted in negative or positive ionization mode - MS/MS spectra only available for one adduct per lipid
mzCloud [93]	- Public - Highly curated MS/MS and MS^n spectral information - Spectral peaks are structurally annotated	- Low number of metabolites - Spectral data are not downloadable - Only Orbitrap spectra
Wiley 10th [94]	- Largest mass spectral library commercially available - 719,000 spectra (>950,000 spectra if combined with NIST 14) - Over 638,000 compounds (>760,000 compounds if combined with NIST 14) - Compatible with most instrument manufacturers	- Commercial license required - Only 70 eV EI mass spectra - Beyond metabolomic applications
MaConDa [95]	- Public database of ~200 contaminants in mass spectrometry - Theoretical and experimental spectral records detected across several MS platforms - Downloadable	- Has no MS/MS data
MassBank [29]	- Public - Mass spectra from different MS setups - Approximately 19,000 MS1 and 28,000 MS2 and MS^n spectra - Spectral data are downloadable	- Not sufficiently curated
NIST 14 [63]	- 234,284 ESI MS/MS spectra of 9344 chemical standards - Large number of MS/MS spectra from adducts - MS/MS spectra recorded using multiple high- and low-resolution instruments - Curated collection of 276,259 EI mass spectra from 242,477 unique compounds - 387,463 measured Kovats or Lee retention index information from 82,337 chemical standards	- Commercial license - Lack of additional identifiers to external database resources
GMD [30]	- Public - Over 2500 EI mass spectra and retention index information - Spectral data are downloadable	- Data derived primarily from plant materials
FiehnLib [67]	- Over 2200 EI and retention indices for >1000 metabolites	- Commercial license with Agilent Technologies and LECO Corporation - Data derived primarily from plant materials
ReSpect [96]	- Public - More than 9000 MS/MS spectra corresponding to >3600 metabolites: ~38% literature data ~12% Q-TOF MS/MS ~50% QqQ MS/MS - Merged spectra (same as MassBank) - Curated record data - Downloadable	- Only Q-TOF and QqQ MS data - Mainly phytochemicals (plant metabolomics) - High degree of redundancy with MassBank
GNPS [85]	- Public - 8853 MS/MS spectra - MS/MS of adducts - MS/MS of unidentified structures - Downloadable	- Very few spectra in negative ionization - Limited spectrum information - No spectral clean-up/noise removal - "Gold standard" is not comparable with reference databases

Vinaixa et al., Trends in Anal Chem 2016

Identifying metabolomic signals

Table 2
Summary of the most widely used mass spectral databases in metabolomics

Database	Pros	Cons
	~50% QqQ MS/MS - Merged spectra (same as MassBank) - Curated record data - Downloadable	
GNPS [85]	- Public - 8853 MS/MS spectra - MS/MS of adducts - MS/MS of unidentified structures - Downloadable	- Very few spectra in negative ionization - Limited spectrum information - No spectral clean-up/noise removal - "Gold standard" is not comparable with reference databases

<u>NOTES:</u>

METLIN (developed & maintained by G. Siuzdak, Scripps):

biggest MS2 library

HMDB (developed & maintained by D. Wishart, Alberta):

metabocards

NMR

MoNA (maintained by O. Fiehn, UC Davis)

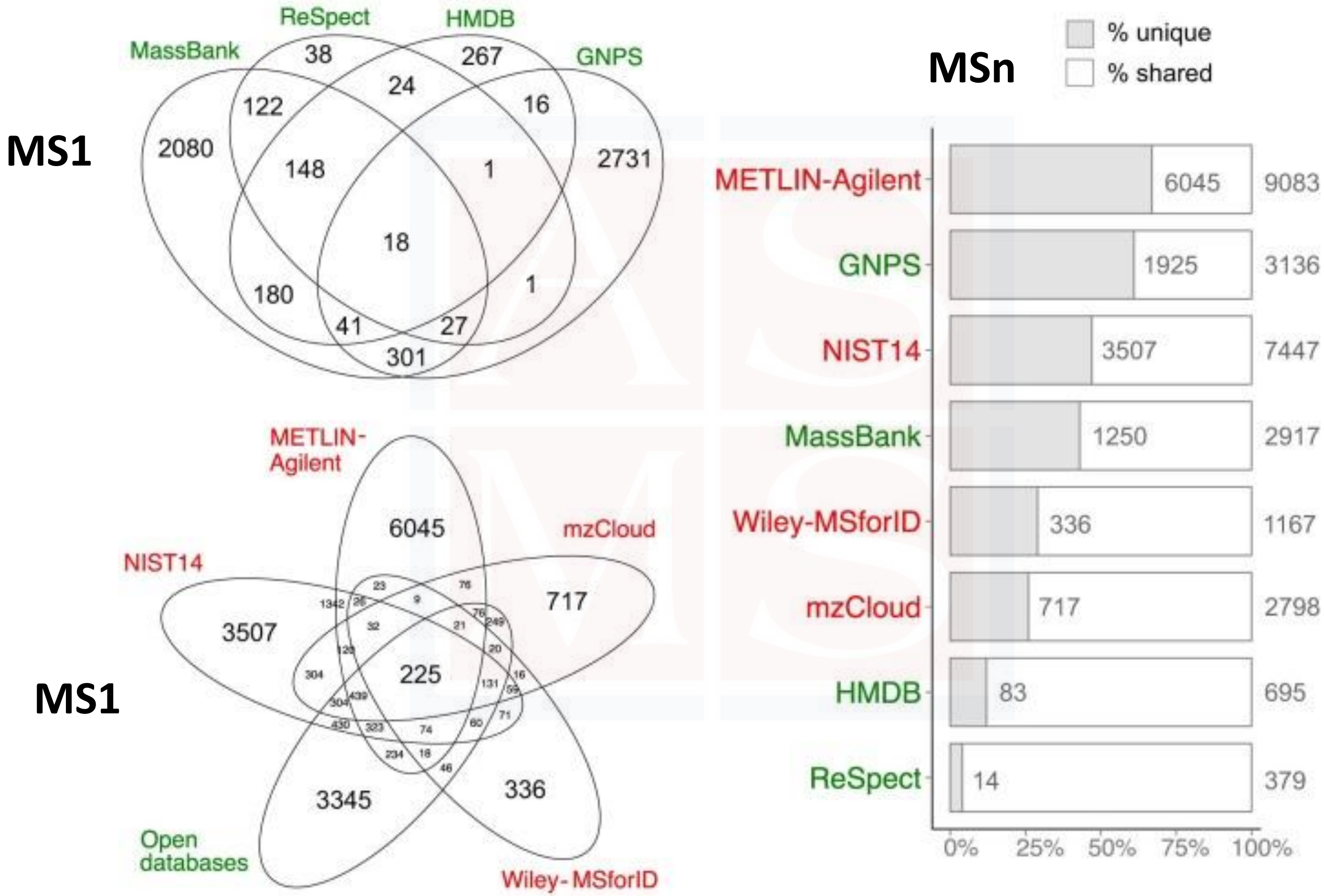
2M spectra, deposited by 106 labs

mzCloud (Thermo):

MSn

Vinaixa et al., Trends in Anal Chem 2016

Identifying metabolomic signals



Vinaixa et al., Trends in Anal Chem 2016

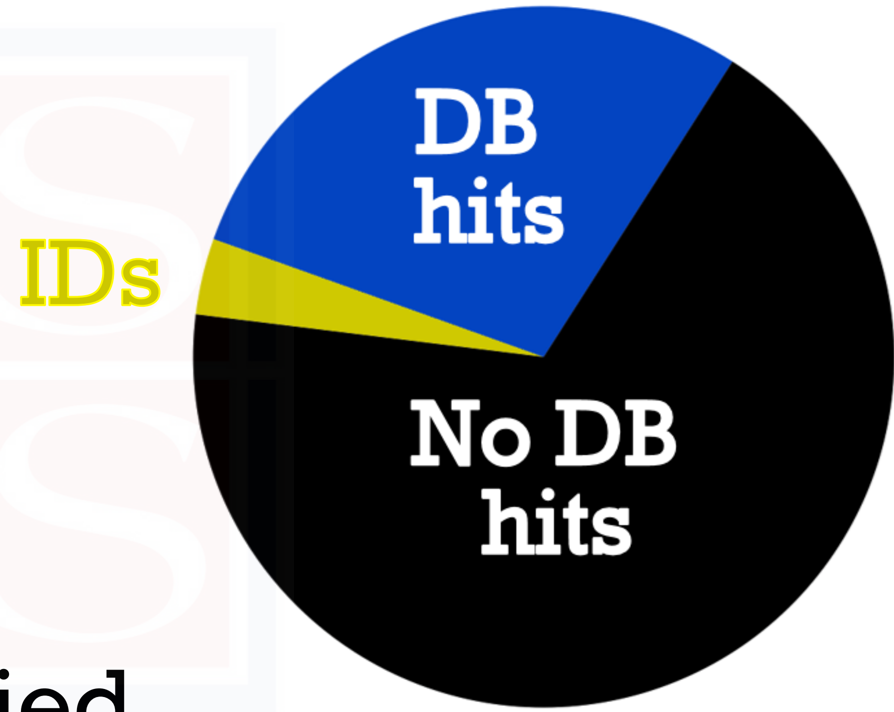
Identifying metabolomic signals

Identifying metabolomic signals

E. coli sample

25,342 total signals

<1000 signals identified



Gap between canned software solutions & complete annotation

- Signal ID widely recognized as bottleneck of untargeted metabolomics

- In 2015, it was found that only 1.8% of spectra in untargeted metabolomics can be annotated[1]

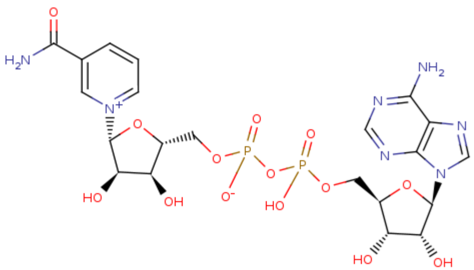
- Similar trends seen in data from public repositories[2]

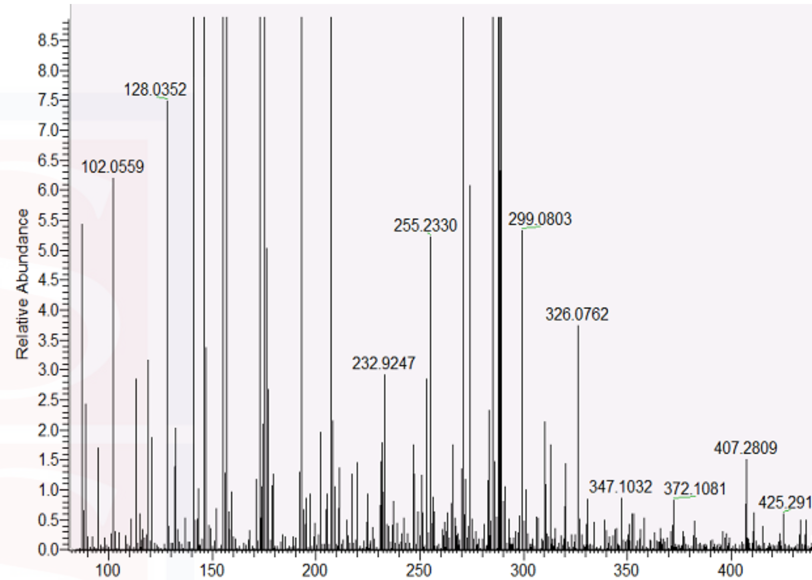
- Some signals that cannot be identified are due to contaminants/artifacts and degenerate signals not currently annotated using conventional workflows

[1]da Silva et al., PNAS 2015
[2]Blazenovic et al., Metabolites 2018

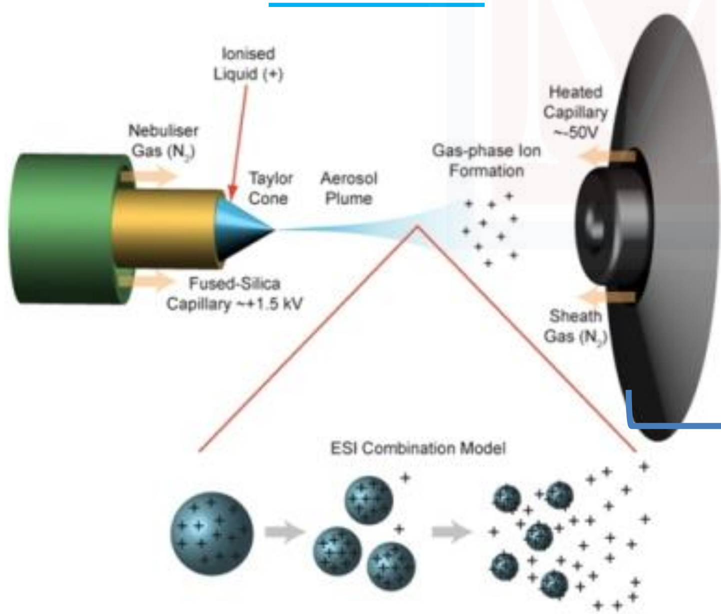
Sources of signal degeneracy

1. Isotopes





2. Adducts



3. Fragments

4. Harmonics
5. Excited States
6. Isomers

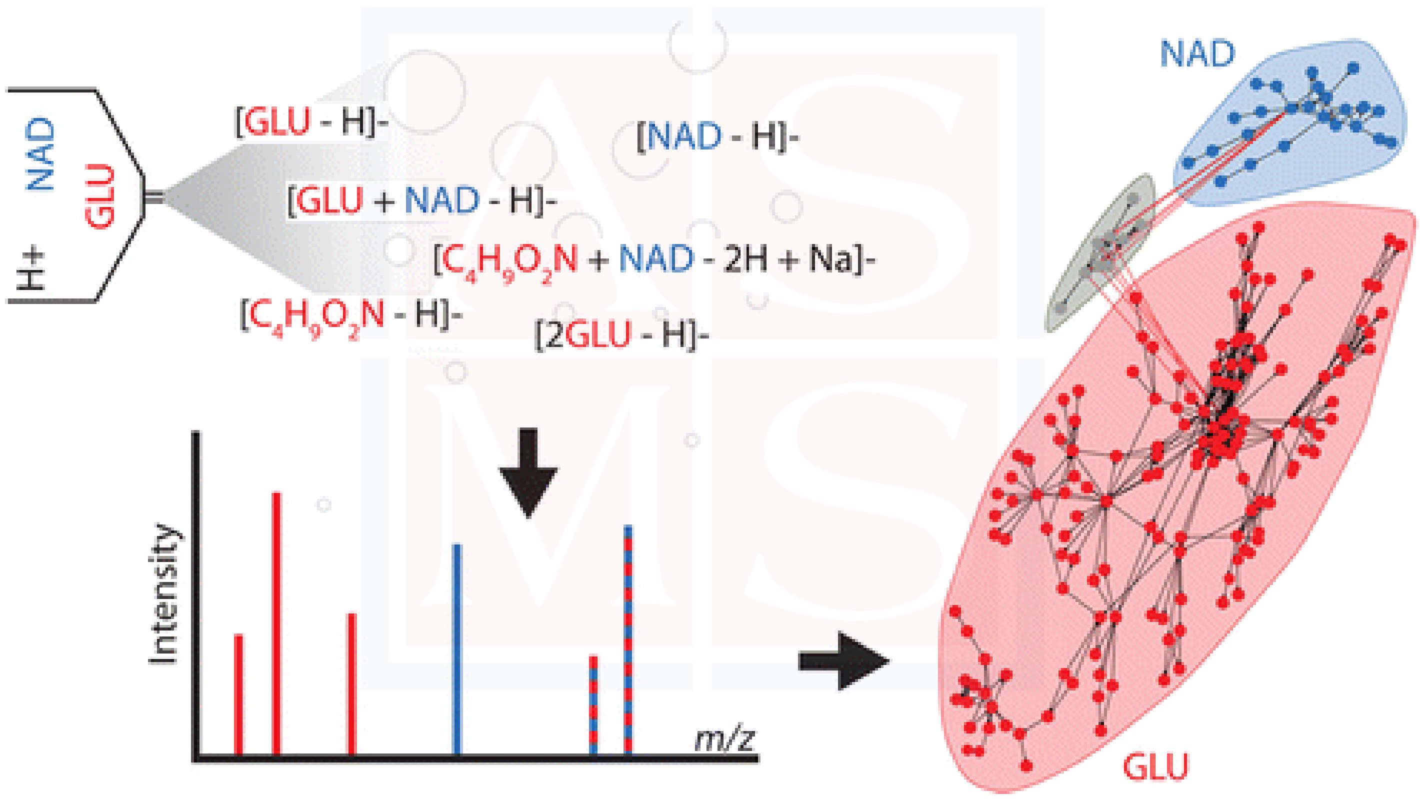
Sources of signal degeneracy

1. **Isotopes**
   - A single analyte is a mixture of formula due to the natural abundance of heavy isotopes.
   - Isotopes exist prior to analysis

2. **Adducts**
   - During evaporation and ion formation two species may form a single ion held together by noncovalent interactions.
   - Any present species can participate (including contaminants).

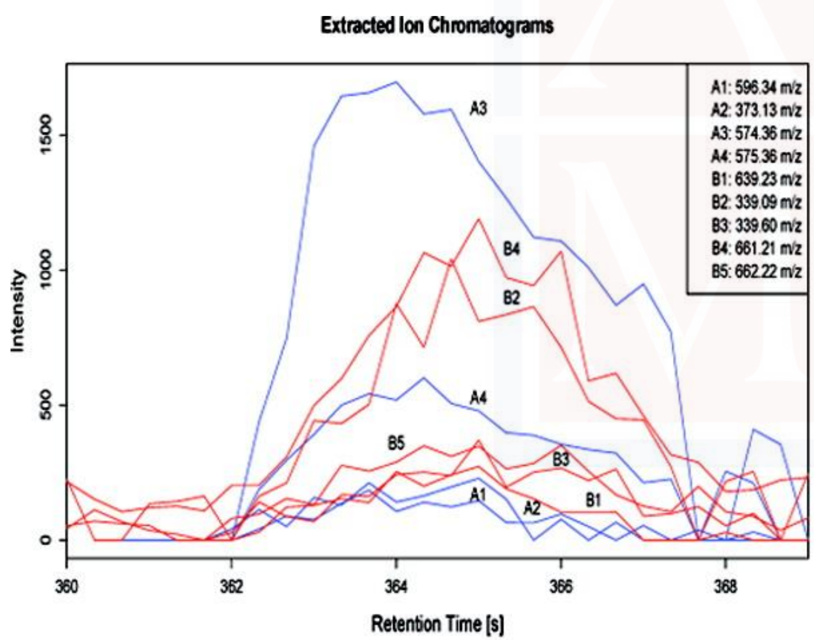
3. **Fragments**
   - As ions transition from atmospheric pressure they are pulled through surrounding neutral gas, collisions imparting KE. They further accelerate to supersonic speeds as the ion cloud expands at 10E-5 torr.
   - Fragments are limited to subsets of the formula present after adduct formation.

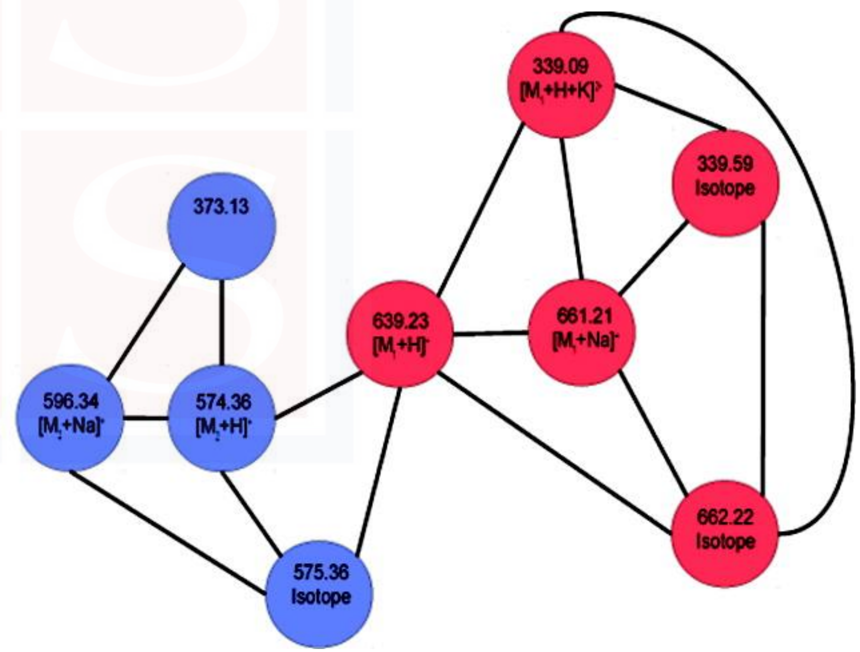
Annotating signal degeneracy



Annotating signal degeneracy

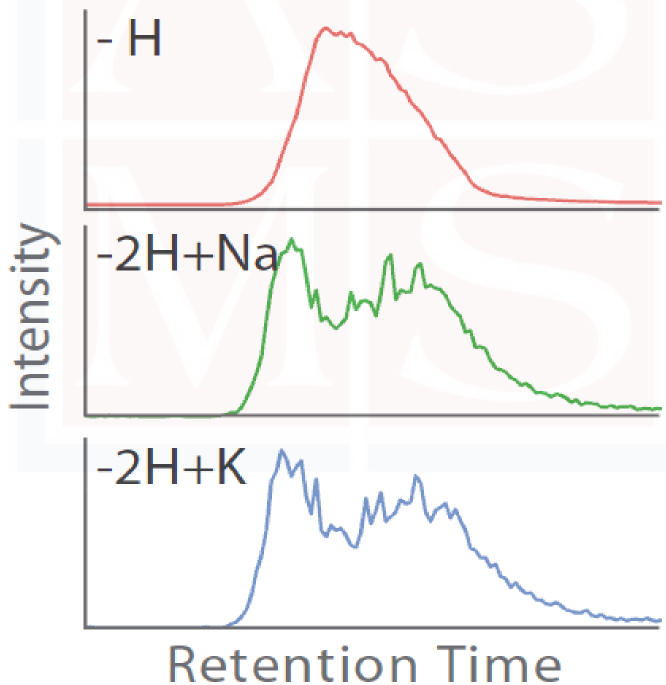
Clustering of peaks based on chromatographic shape alone can be limiting





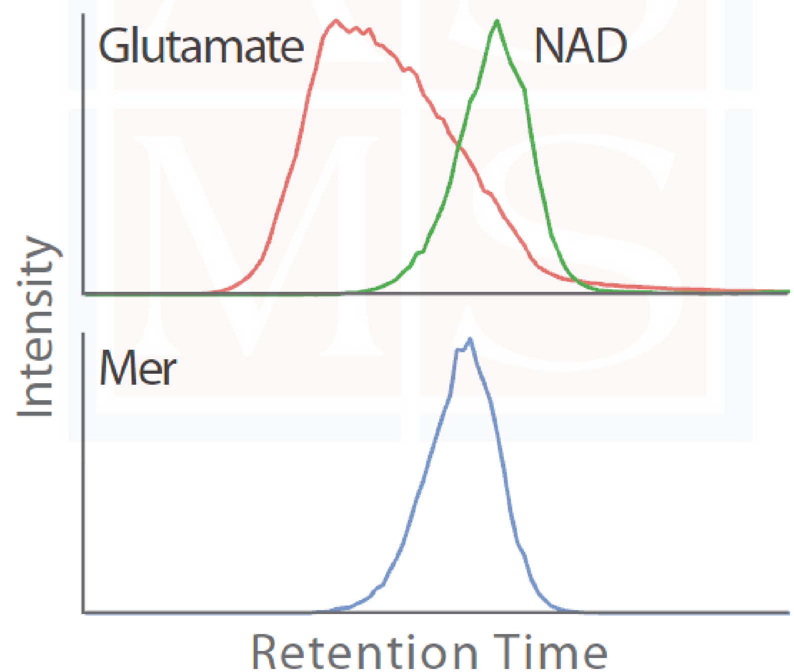
Annotating signal degeneracy

Clustering of peaks based on chromatographic shape alone can be limiting



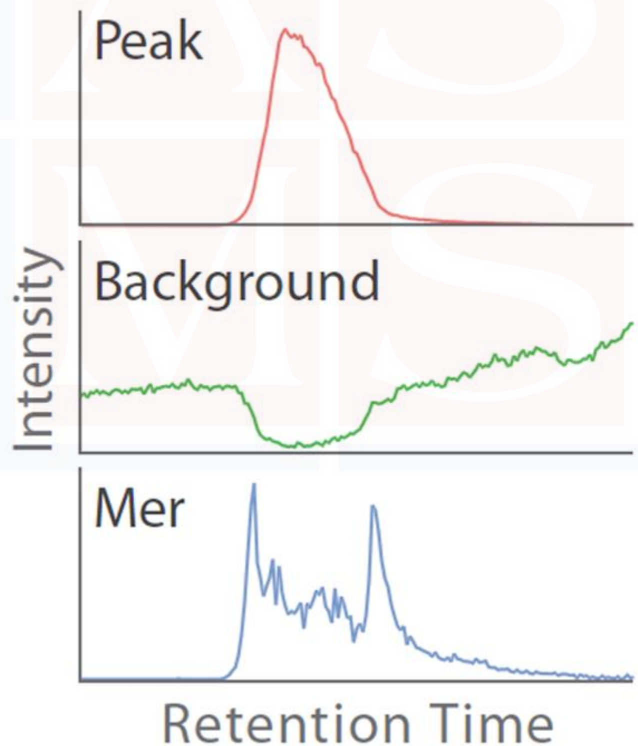
Annotating signal degeneracy

Clustering of peaks based on chromatographic shape alone can be limiting



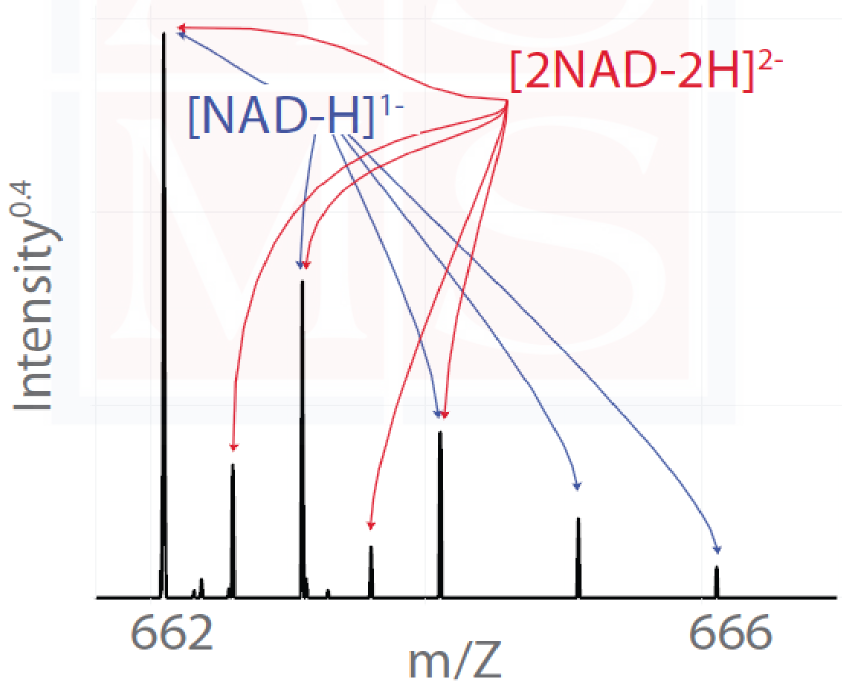
Annotating signal degeneracy

Clustering of peaks based on chromatographic shape alone can be limiting



Annotating signal degeneracy

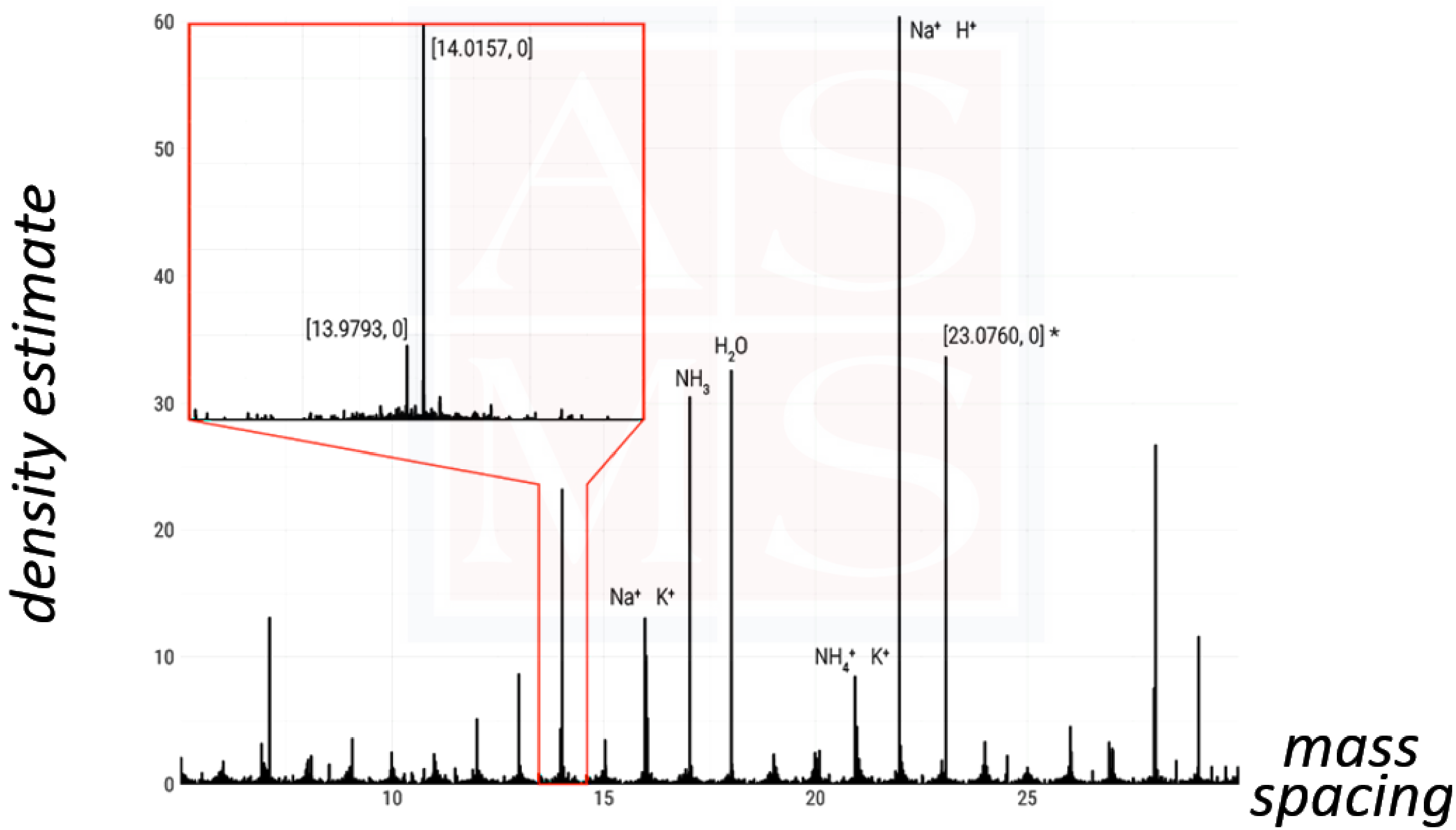
Clustering of peaks based on chromatographic shape alone can be limiting



Annotating signal degeneracy

- Annotation of isotopic fine structure is important
- Relationships can exist between more than two peaks
- The chromatographic profiles of degenerate signals may not match

Annotating signal degeneracy



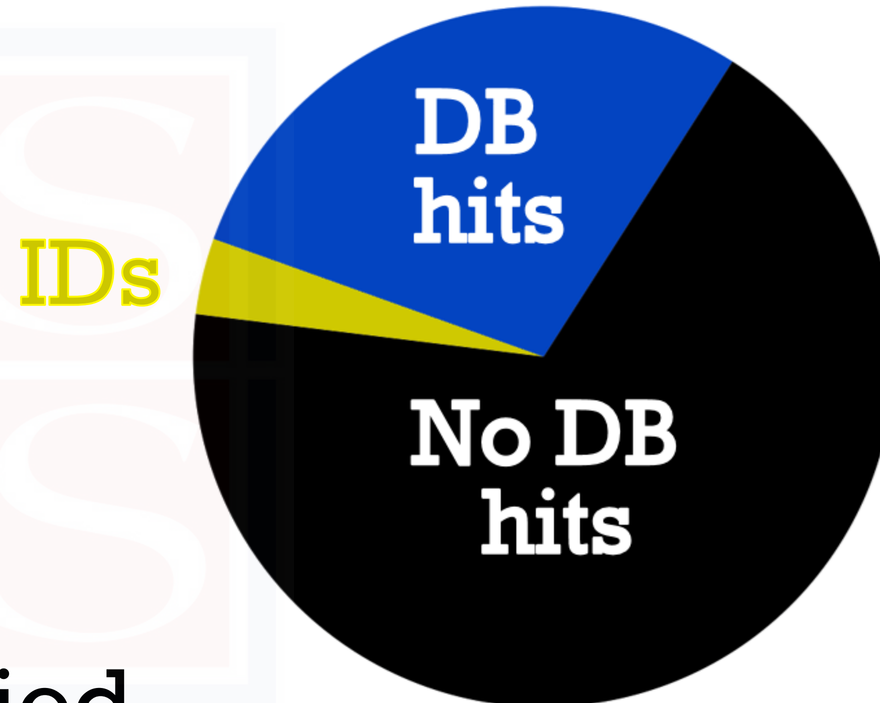
Brown et al., Analyst 2009
Mahieu et al., Anal Chem 2016

Identifying metabolomic signals

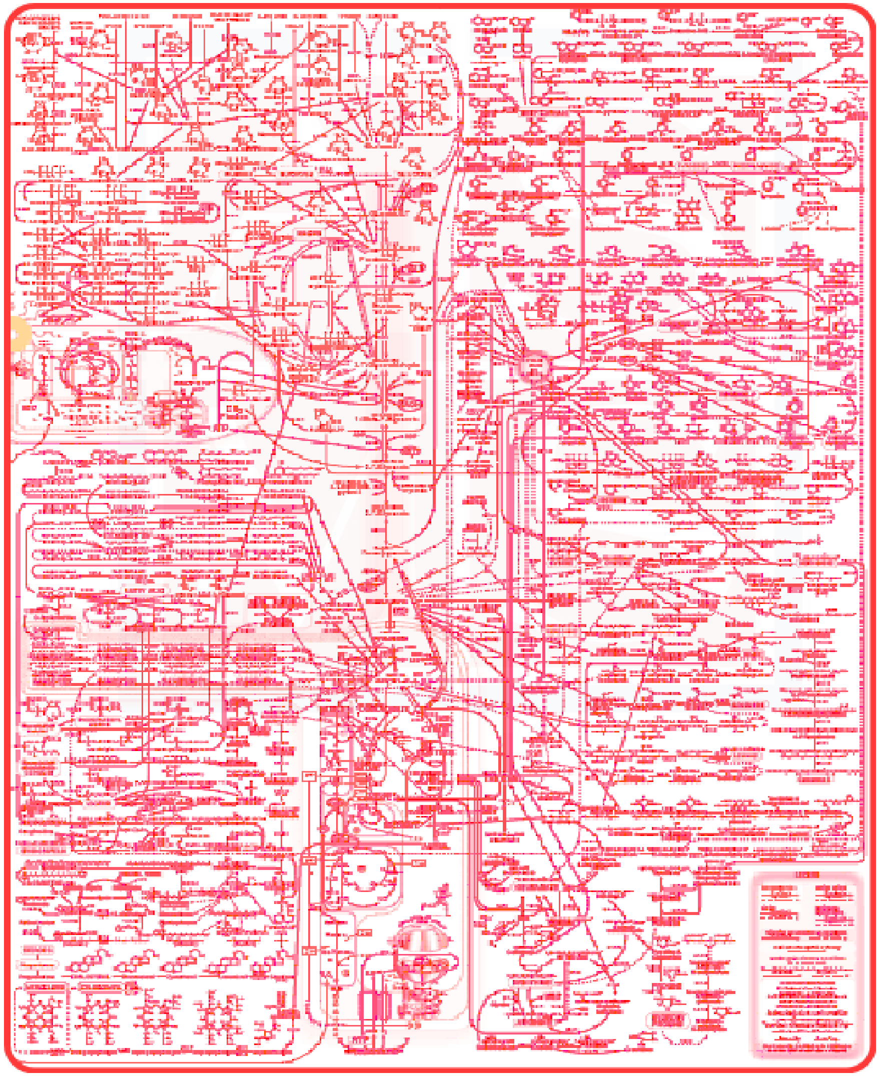
E. coli sample

25,342 total signals

<1000 signals identified

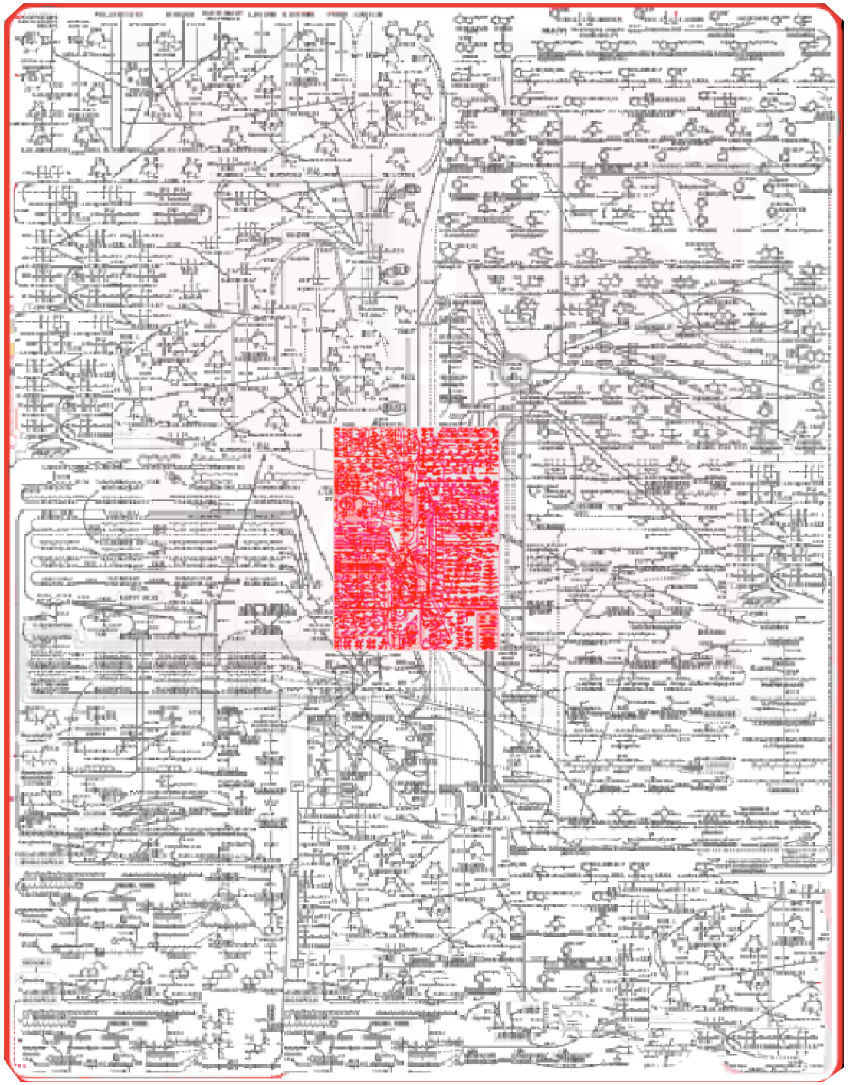


Identifying metabolomic signals

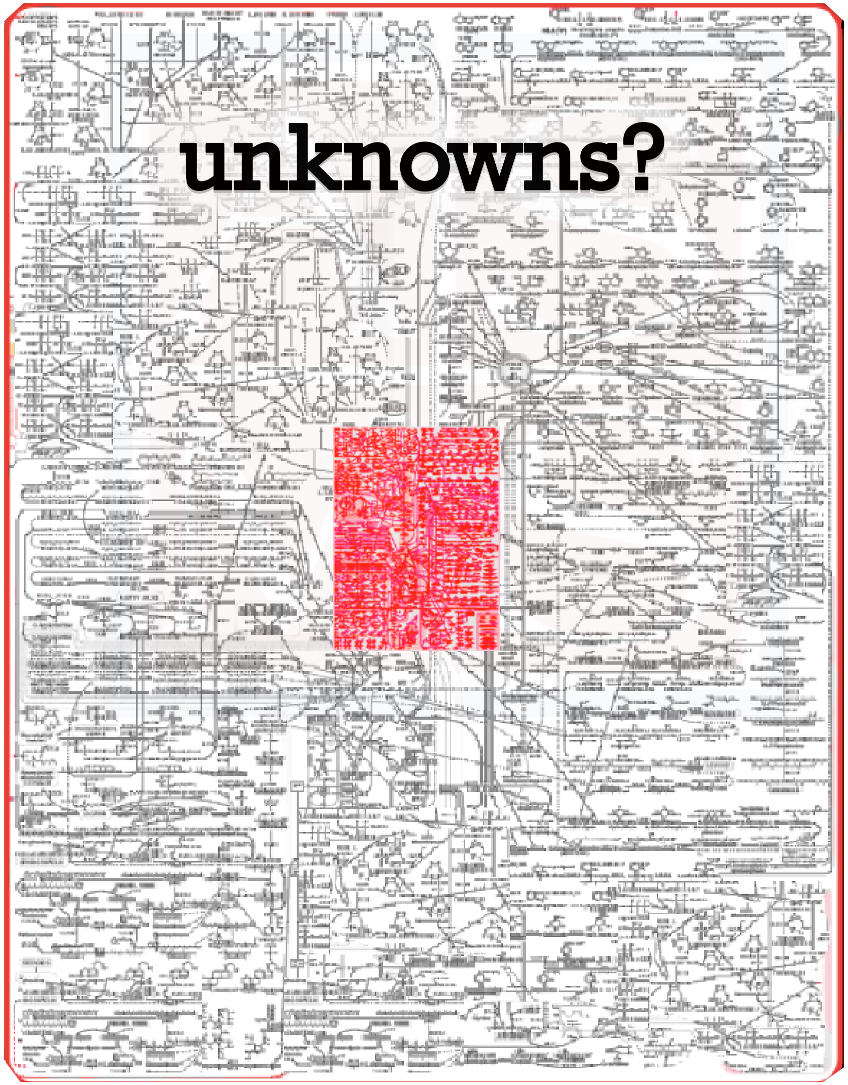


Identifying metabolomic signals

Identifying metabolomic signals

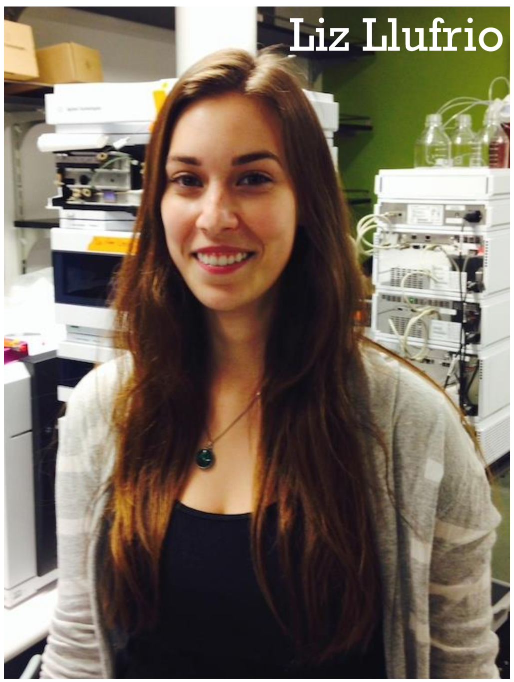


Identifying metabolomic signals



unknown ID: a cautionary anecdote

m/z = 809.1550

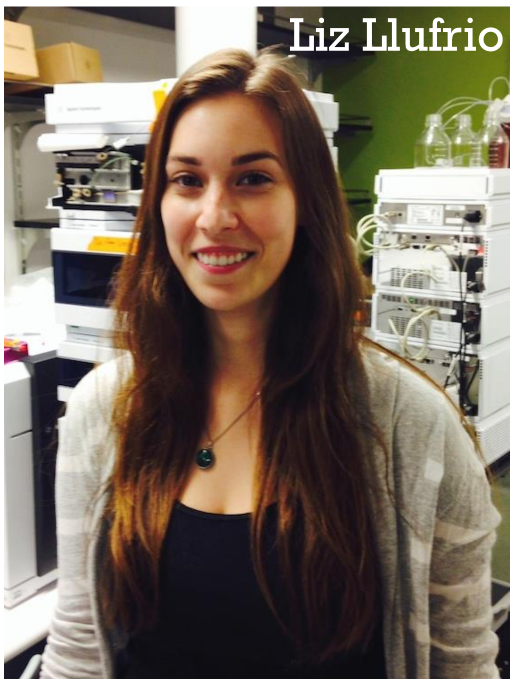


Sept. 2012

no DB hits
fold change
408 fragment
M+5 labeling
novel cmpd

unknown ID: a cautionary anecdote

m/z **= 809.1550**



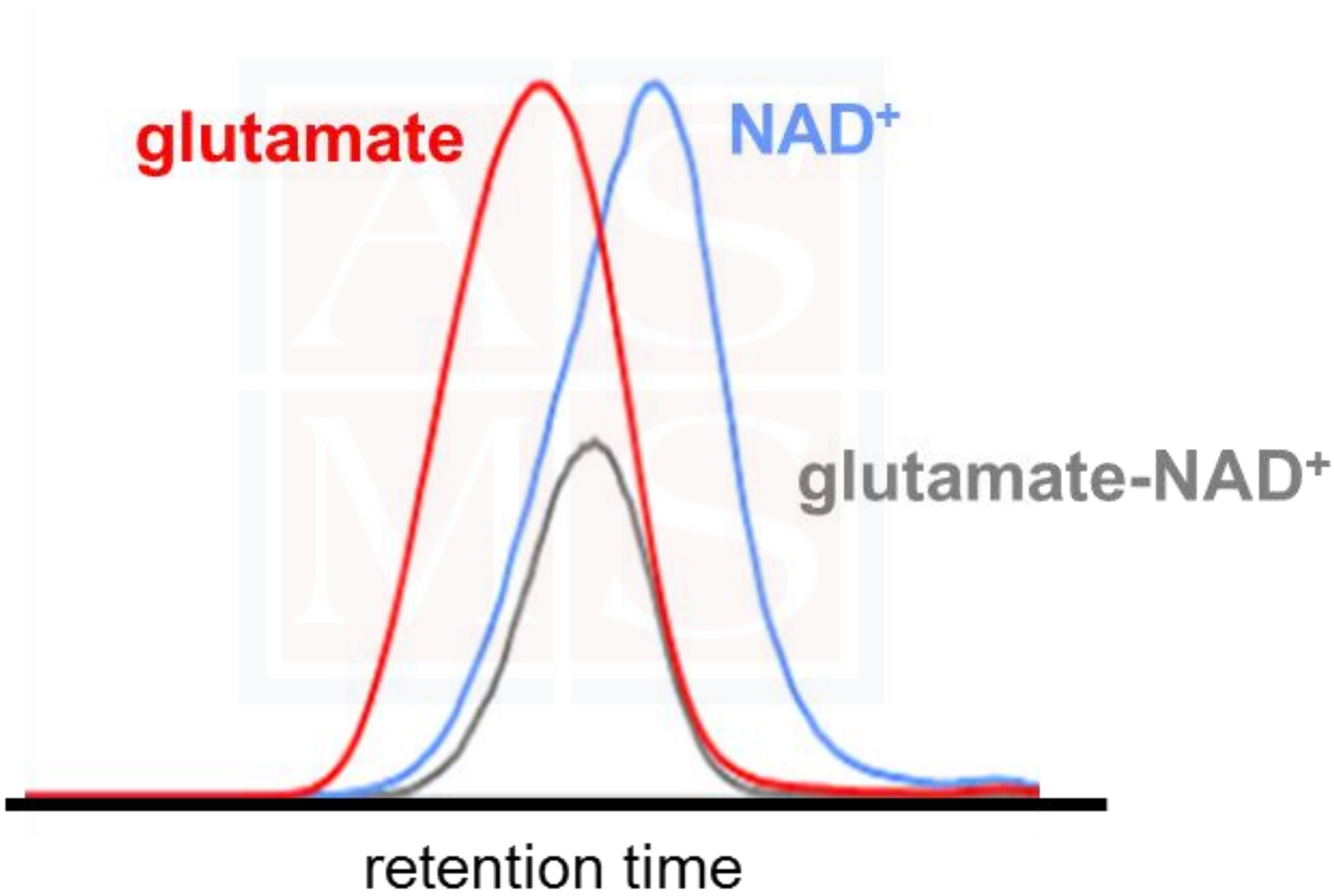
<u>Sept. 2012</u>

no DB hits
fold change
408 fragment
M+5 labeling
novel cmpd

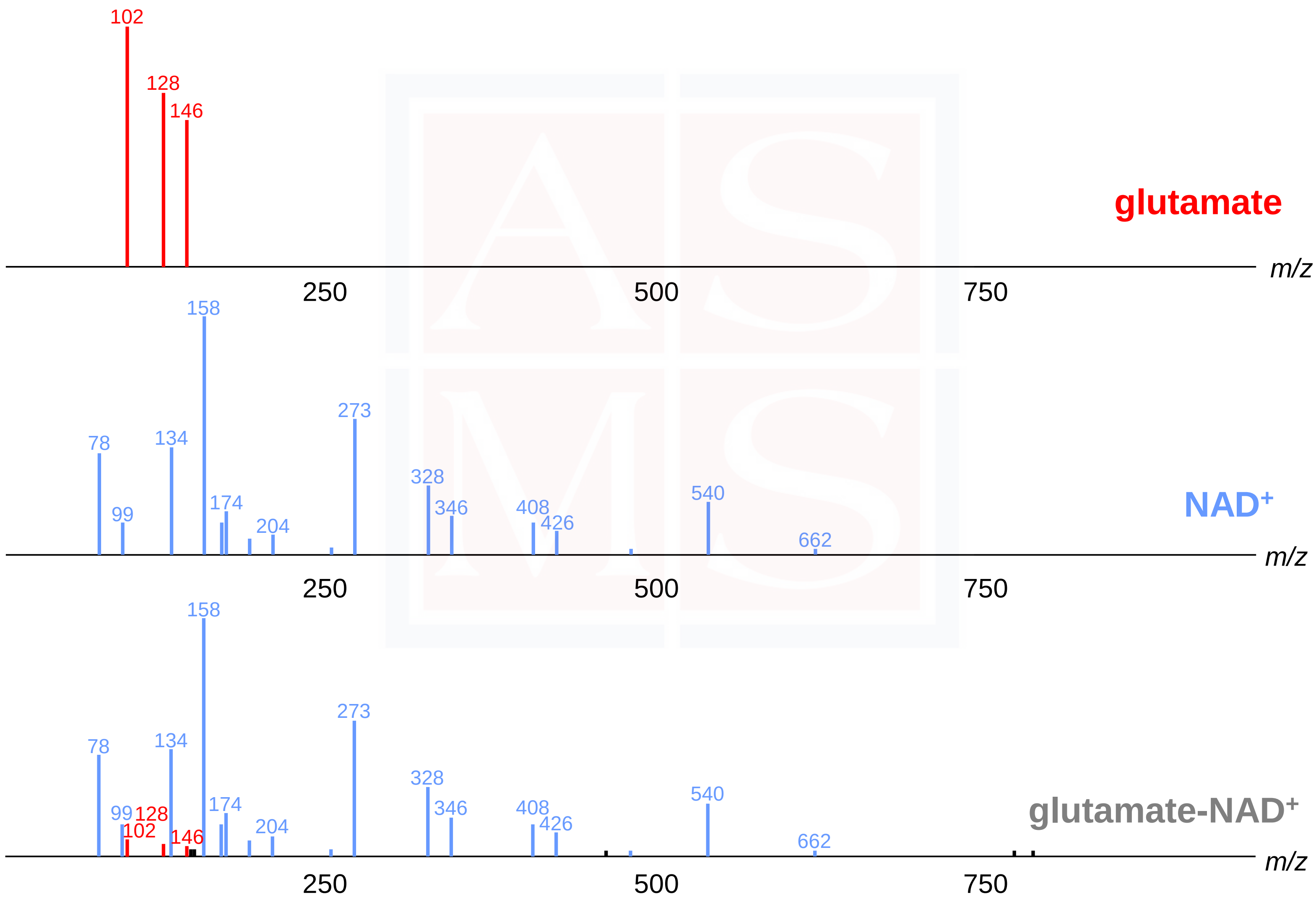
<u>June 2015</u>

heteromer
glu-NAD

unknown ID: a cautionary anecdote



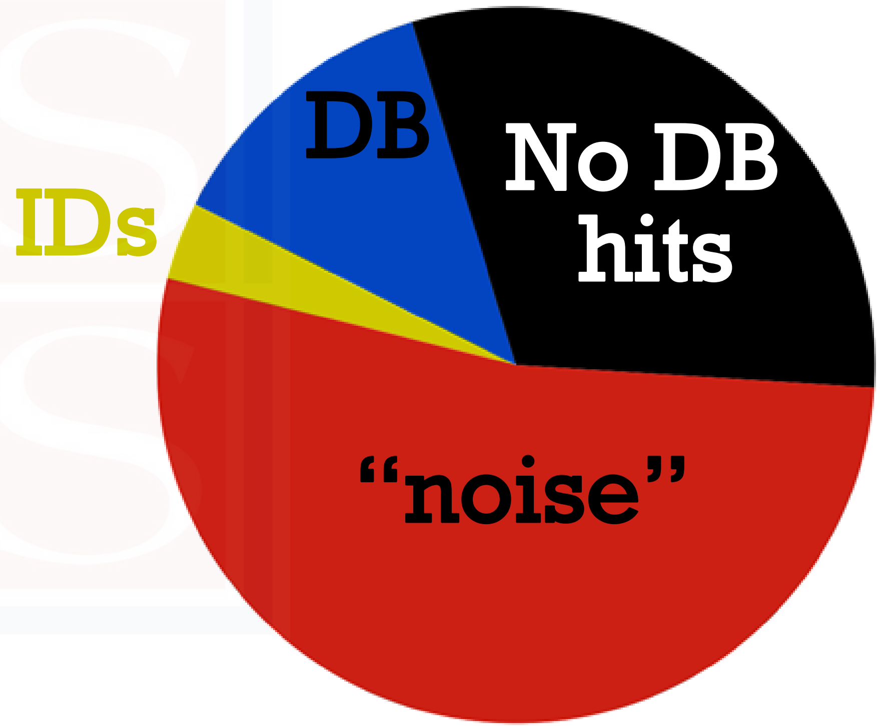
unknown ID: a cautionary anecdote



Identifying metabolomic signals

E. coli sample

~50% artfct/contmn



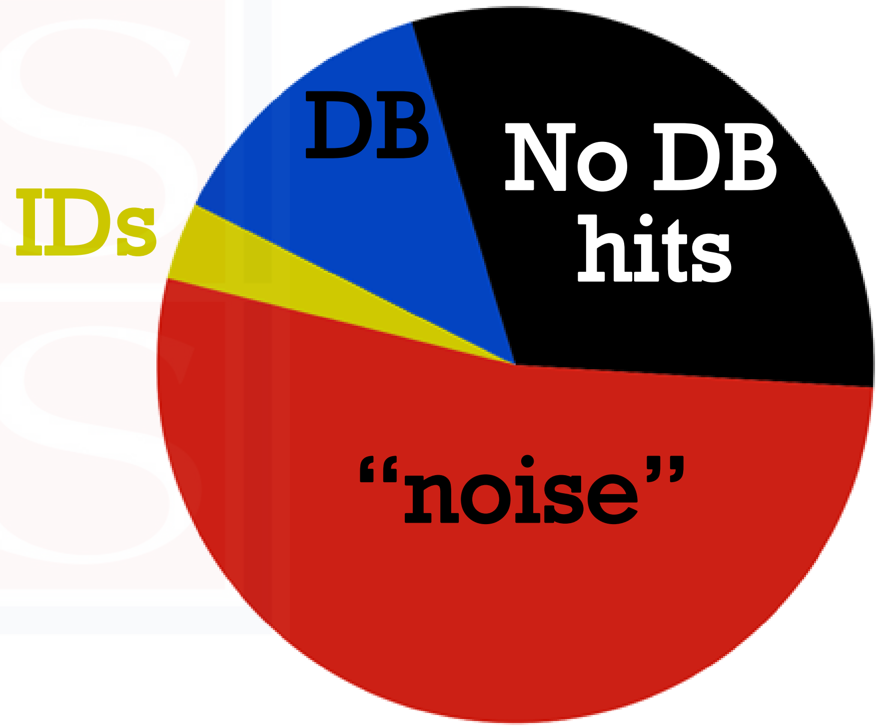
Mahieu et al., Anal Chem 2017

Identifying metabolomic signals

E. coli sample

~50% artfct/contmn

1000s of redundant



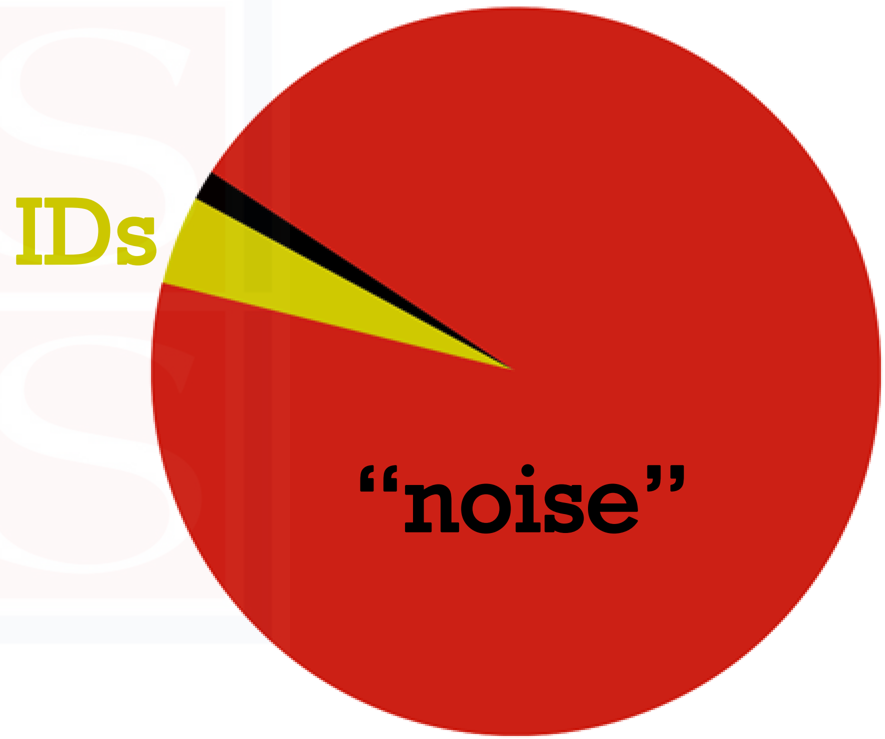
Mahieu et al., Anal Chem 2017

Identifying metabolomic signals

E. coli sample

~50% artfct/contmn

1000s of redundant



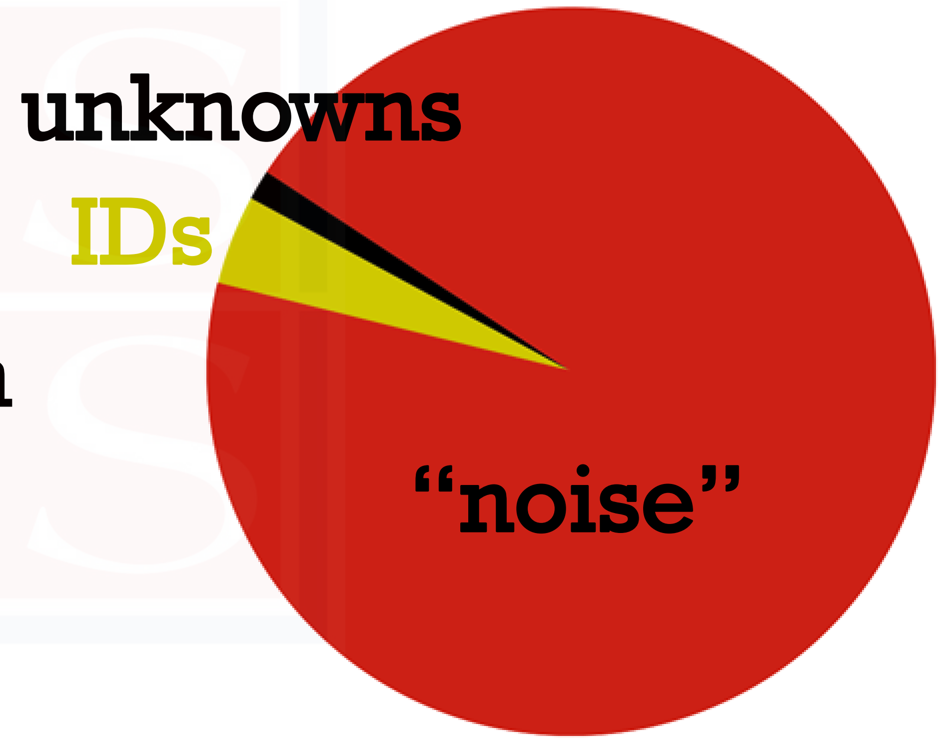
Mahieu et al., Anal Chem 2017

Identifying metabolomic signals

E. coli sample

~50% artfct/contmn

1000s of redundant



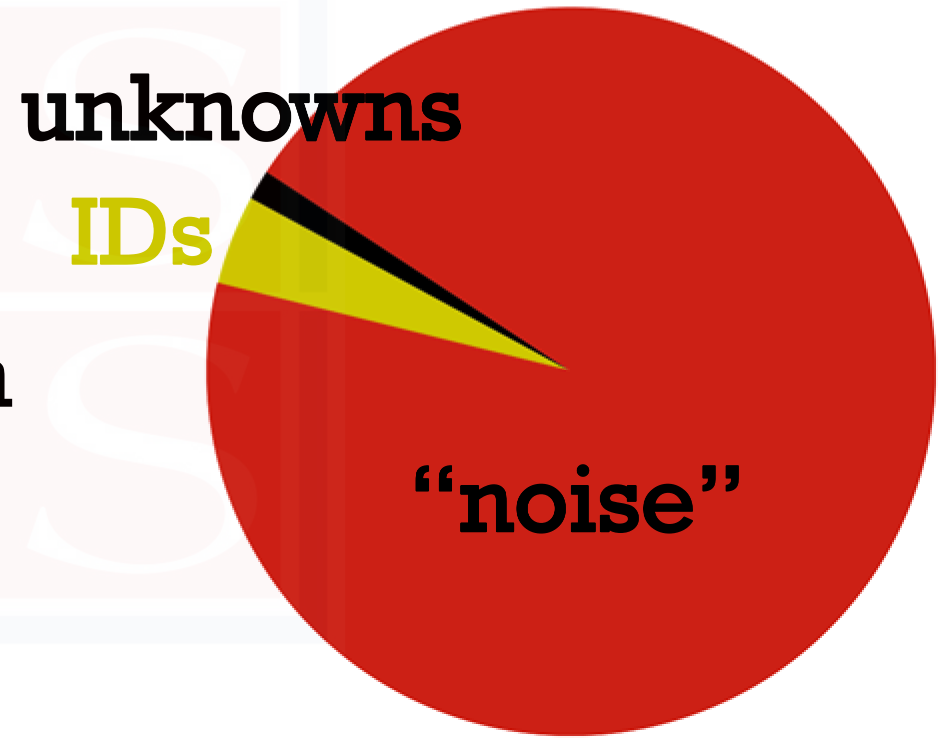
Mahieu et al., Anal Chem 2017

Identifying metabolomic signals

E. coli sample

~50% artfct/contmn

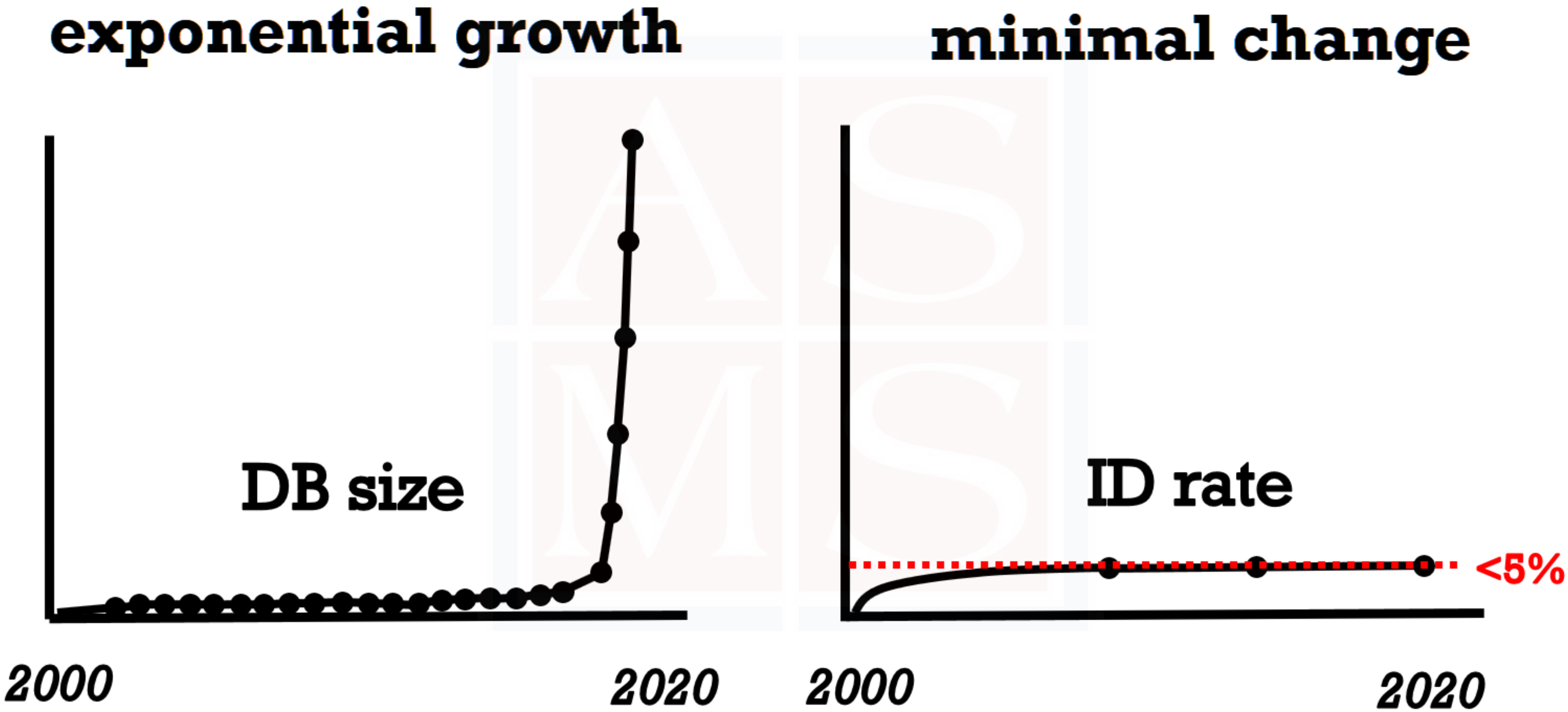
1000s of redundant



NOTES: (1) this is one experiment, may not typify your experiment
(2) this only speaks to the number of unknowns in this data set,
not the total number of metabolites in *E. coli*

Mahieu et al., Anal Chem 2017

Identifying metabolomic signals

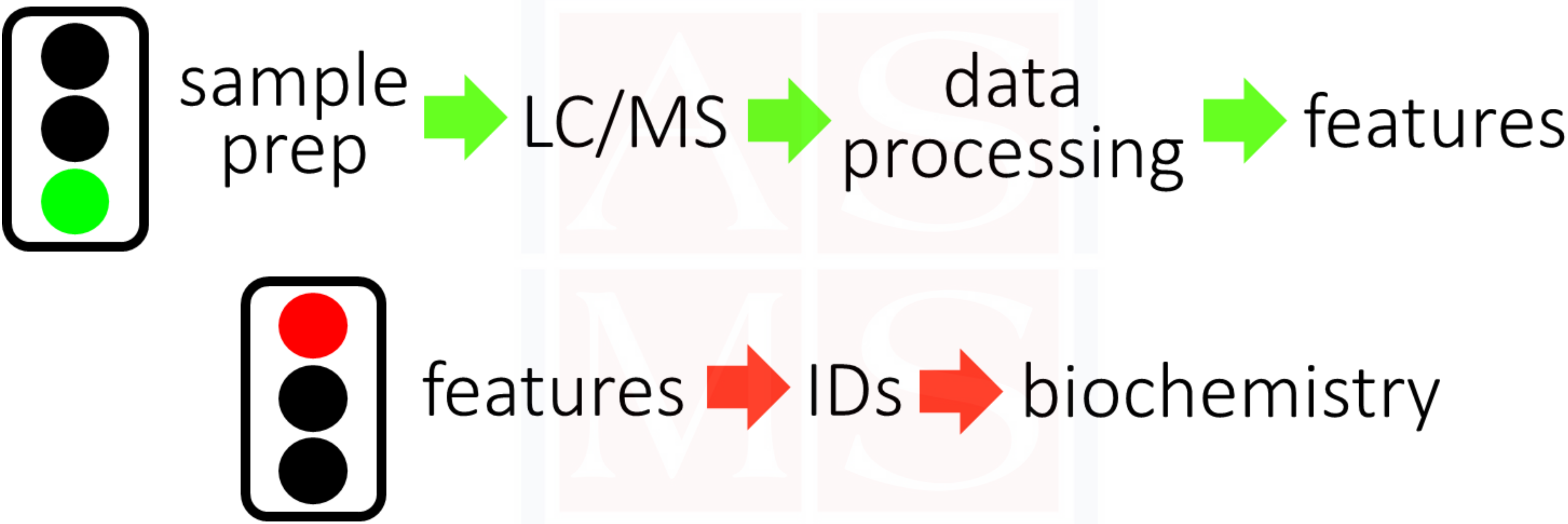


Implications of artifacts, contaminants, and signal degeneracy

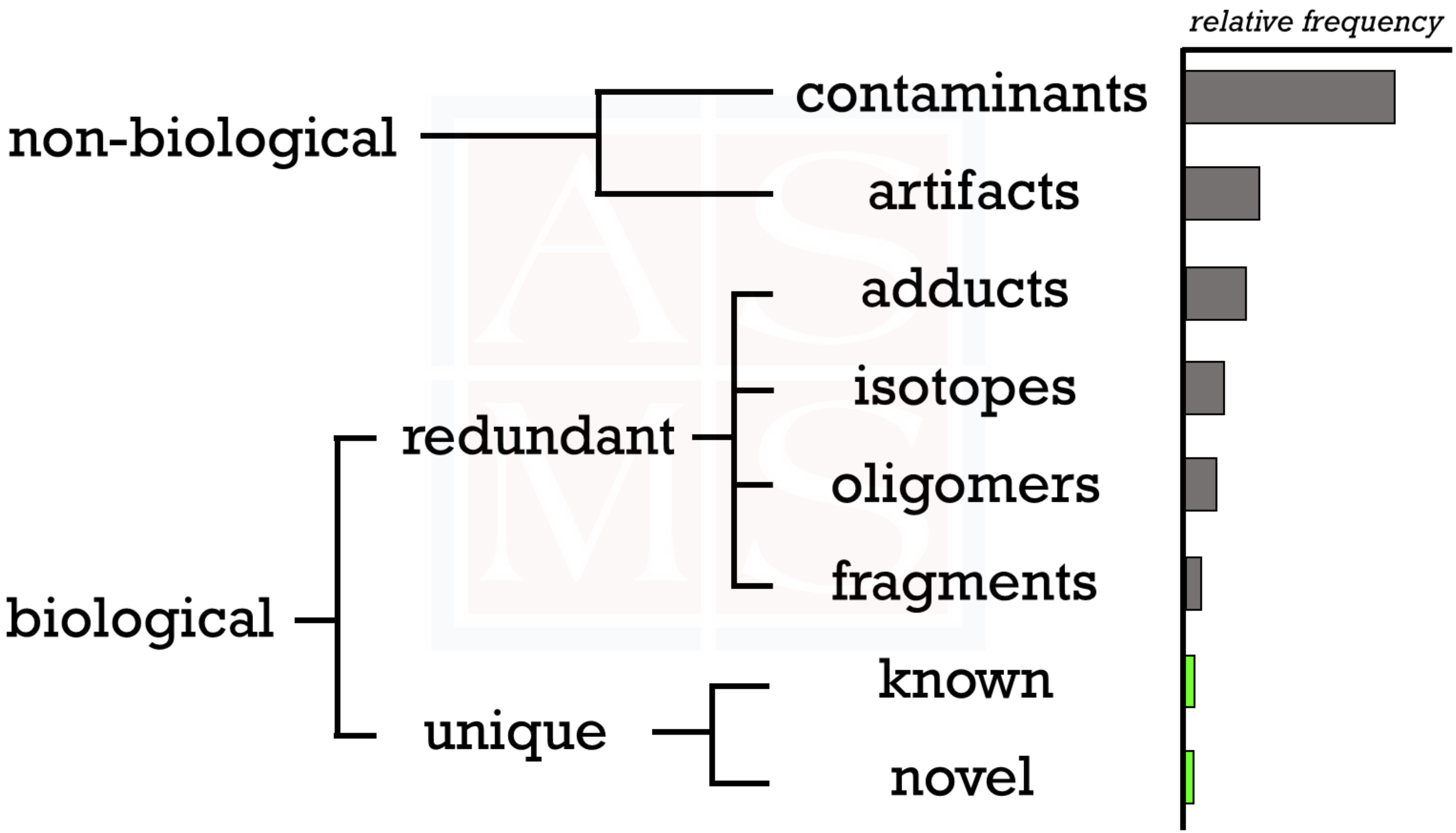
Caution should be taken when analyzing features prior to metabolite ID. Examples:

- Cell type A is significantly different than cell type B because 55% of features are dysregulated.
- Method A is better than method B because 15% more features are detected.
- Inferring pathway dysregulation on basis of MS1 only.

"Easy" does not mean robust



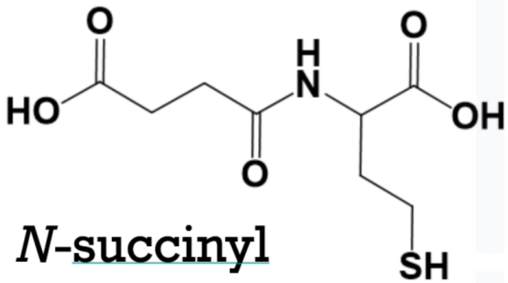
Roadmap of metabolomics data



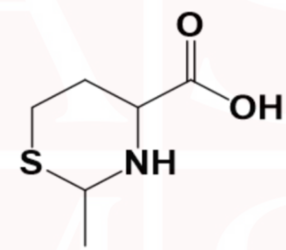
Sindelar et al., JACS 2020

Roadmap of metabolomics data

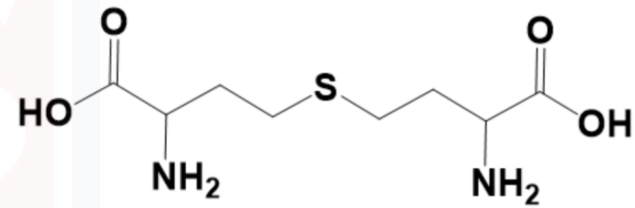
novel compounds from *E. coli*



N-succinyl homocysteine



2-methyl-1,3-thiazinane-4-carboxylic acid



homolanthionine

Jim Edwards

Chris Arnatt

Software for annotating features

Software tools commonly used for metabolite annotation

Tool	Annotation level	Software type	Website	References
CAMERA	Level 4	R Package	http://bioconductor.org/packages/release/bioc/html/CAMERA.html	Kuhl et al. (2012)
MZedDB	Level 4	Web App	http://maltese.dbs.aber.ac.uk:8888/hrmet/index.html	Draper et al. (2009)
Rdisop	Level 4	R Package	http://bioconductor.org/packages/release/bioc/html/Rdisop.html	–
SIRIUS	Level 4	CLI, GUI	https://bio.informatik.uni-jena.de/software/sirius	Kim et al. (2016)
MI-PACK	Level 3	CLI	http://www.biosciences-labs.bham.ac.uk/viant/mipack	Weber and Viant (2010)
PUTMEDID-LCMS	Level 3	CLI	http://www.mcisb.org/resources/putmedid.html	Brown et al. (2011)
ProbMetab	Level 3	R Package	http://labpib.fmrp.usp.br/methods/probmetab	Silva et al. (2014)
MetAssign–mzMatch	Level 3	R Package	http://mzmatch.sourceforge.net/index.php	Daly et al. (2014)
MetFrag	Level 2a	Web App	http://c-ruttkies.github.io/MetFrag	Ruttkies et al. (2016)
CFM-ID	Level 2a	CLI, Web App	https://sourceforge.net/projects/cfm-id/	Allen et al. (2014)
FingerID	Level 2a	Web App	https://github.com/icdishb/fingerid	Heinonen et al. (2012)
MAGMa	Level 2a	Web App	http://www.emetabolomics.org/magma	Ridder et al. (2013)
MyCompoundID	Level 2a	Web App	http://mycompoundid.org/mycompoundid_IsoMS	Li et al. (2013)
BATMAN	NMR	R Package	http://batman.r-forge.r-project.org	Hao et al. (2012)
Bayesil	NMR	Web App	http://bayesil.ca	Ravanbakhsh et al. (2015)
MetaboMiner	NMR	CLI	http://wishart.biology.ualberta.ca/metabominer	Xia et al. (2008)
SpinAssign	NMR	Web App	http://prime.psc.riken.jp/?action=nmr_search	Chikayama et al. (2010)
COLMAR	NMR	Web App	http://spin.ccic.ohio-state.edu/index.php/colmar	Zhang et al. (2009)

CLI command line interface, *GUI* graphical user interface

Spicer et al., Metabolomics 2017

Software for annotating features

Software tools for the post-processing of metabolomics data

Tool	Instrument data type	Software type	Website	References
batchCorr	LC–MS	R Package	https://gitlab.com/CarlBrunius/batchCorr	Brunius et al. (2016)
crmn	LC–MS, GC–MS	R Package	https://cran.r-project.org/web/packages/crmn/	Redestig et al. (2009)
EigenMS	LC–MS	CLI	https://sourceforge.net/projects/eigenms	Karpievitch et al. (2014)
KMDA	MS	R Package	https://cran.r-project.org/web/packages/KMDA/	Zhan et al. (2015)
metabolomics	MS, NMR	R Package	https://cran.r-project.org/web/packages/metabolomics/	De Livera et al. (2012)
metabomxtr	LC–MS, GC–MS	R Package	https://www.bioconductor.org/packages/release/bioc/html/metabomxtr.html	Nodzenski et al. (2014)
Metabnorm	NMR	R Script	https://sourceforge.net/projects/metabnorm	Jauhiainen et al. (2014)
MetabR	LC–MS	R Script	http://metabr.r-forge.r-project.org/	Ernest et al. (2012)
MetNorm	LC–MS, GC–MS, NMR	R Package	https://cran.r-project.org/web/packages/MetNorm/	Livera et al. (2015)
MSPrep	LC–MS	R Package	https://sourceforge.net/projects/msprep/	Hughes et al. (2014)
muma	MS, NMR	R Package	https://cran.r-project.org/web/packages/muma/	Gaude et al. (2013)

CLI command line interface

Spicer et al., Metabolomics 2017

Data interpretation

<u>Biomarkers</u>	<u>Mechanism</u>
• May be single cmpd	• Pathway interpretation
• Sig. technical burden	• Requires follow-up exp
• Best with untargeted	• Works well w/targeted

SUCCESS DEPENDS ON STUDY GOALS

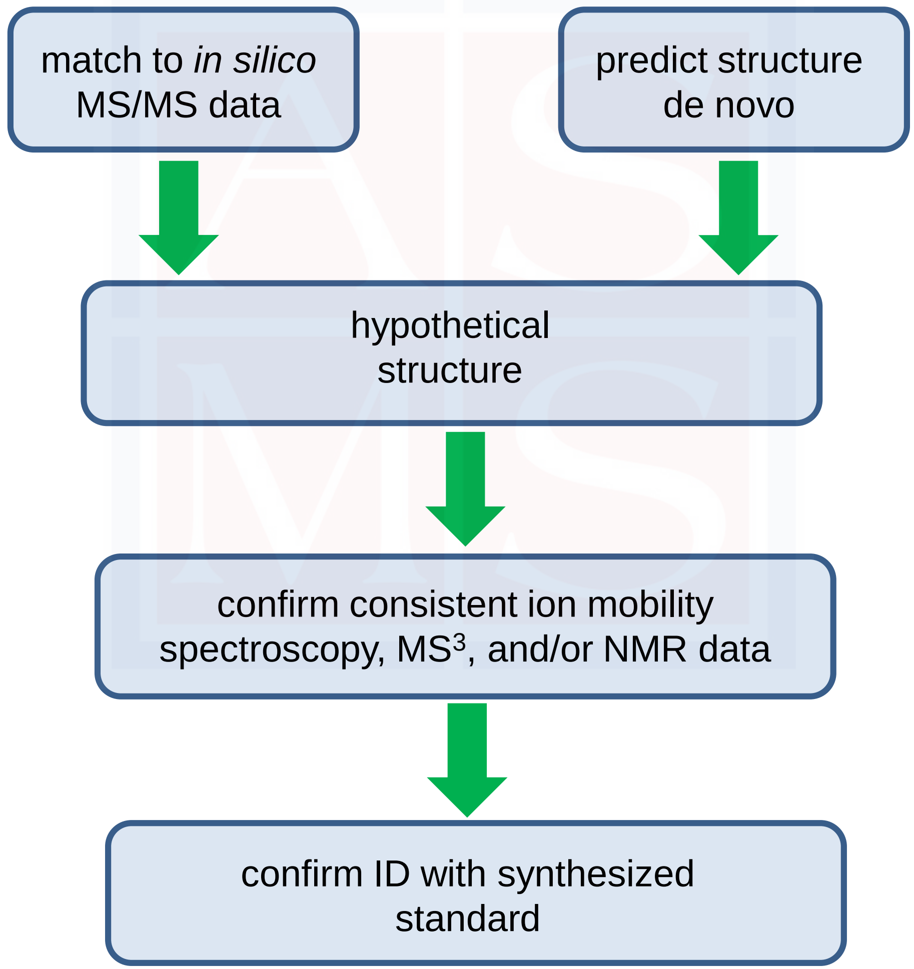
Data interpretation

- One cmpd in pathway may be altered
- Some metabolites easier to interpret

 *lactate, GSH to GSSH, acylcarnitines*

 *erythrose 4-phosphate, PC 18:2/22:6*

How to ID an unknown metabolite?

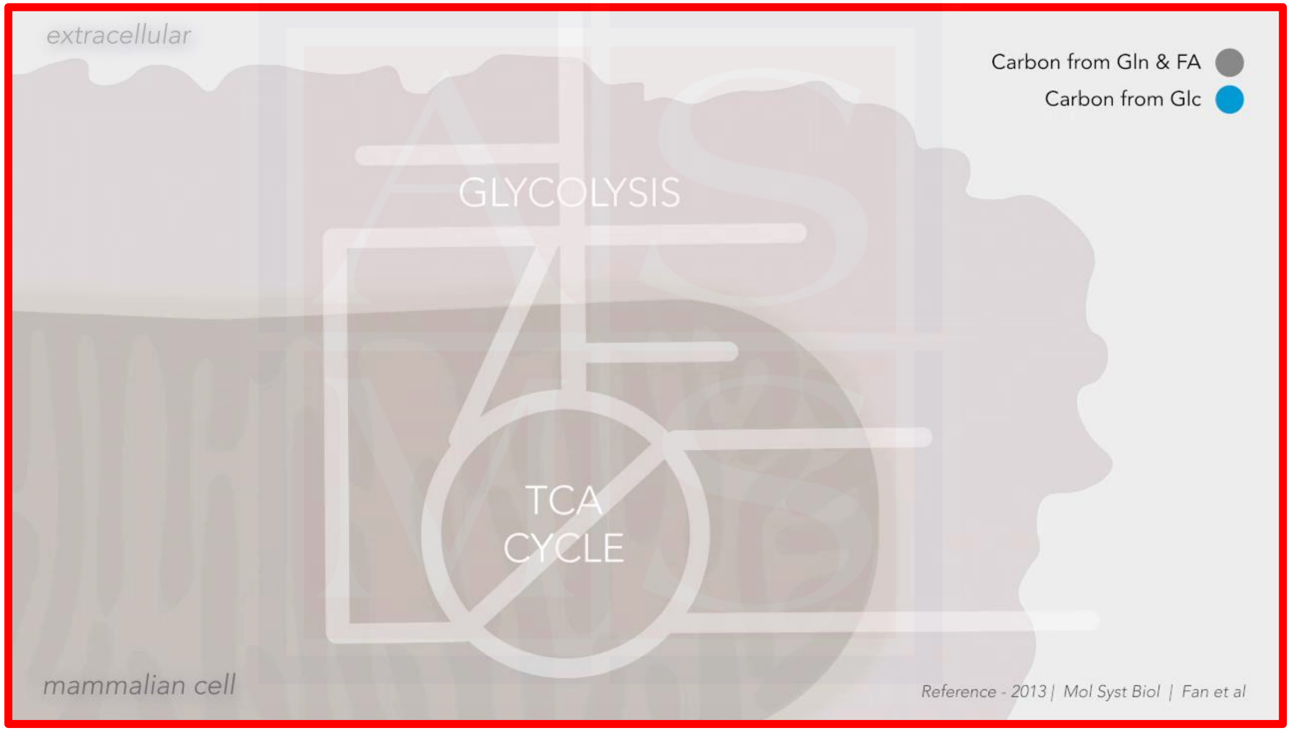


- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

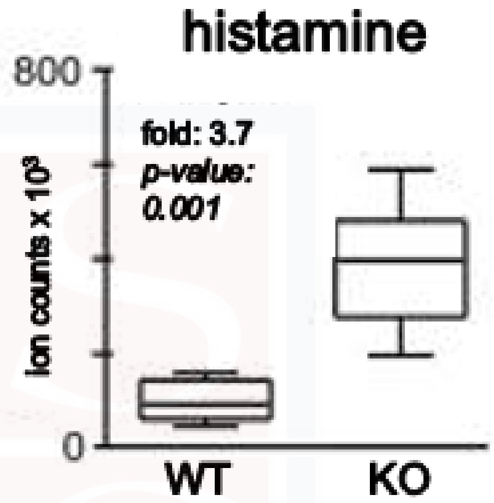
Stable isotope tracer analysis

Benefits of using isotope tracers

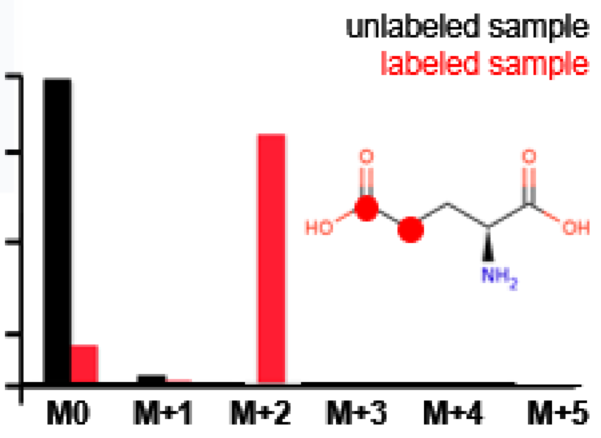


Metabolite profiling strategies

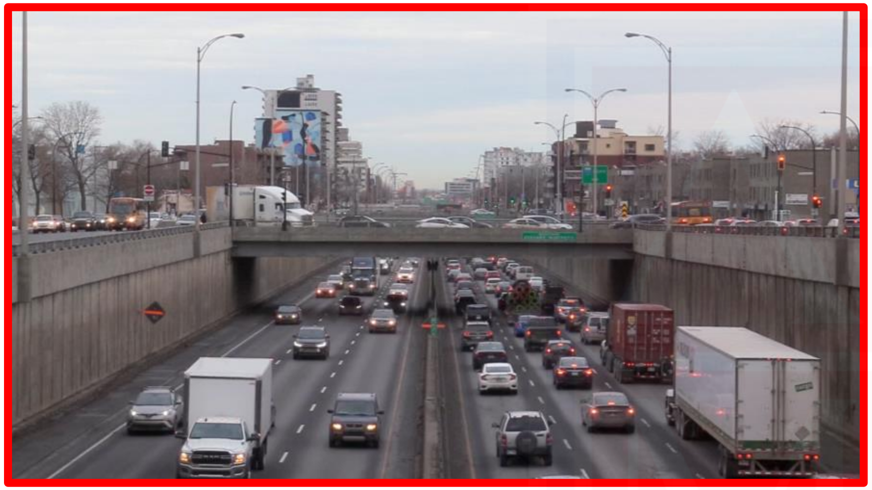
pool-size measurements



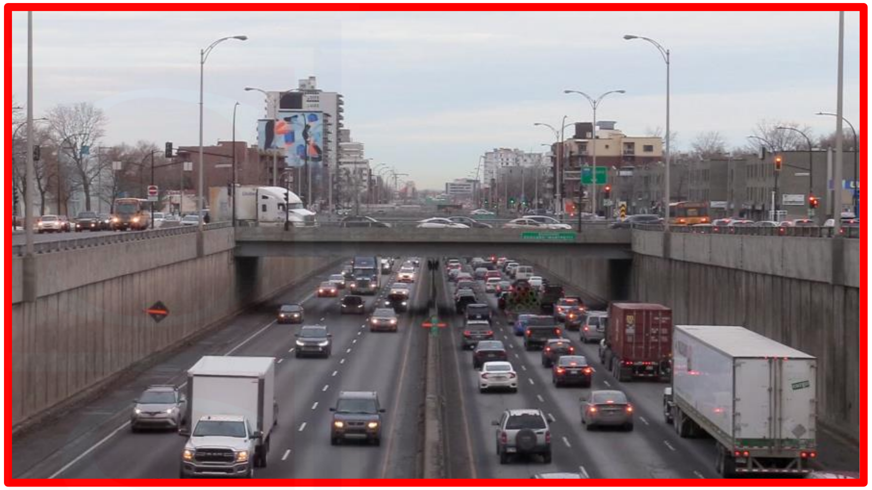
isotope-tracer analysis



Metabolomics vs metabolic flux



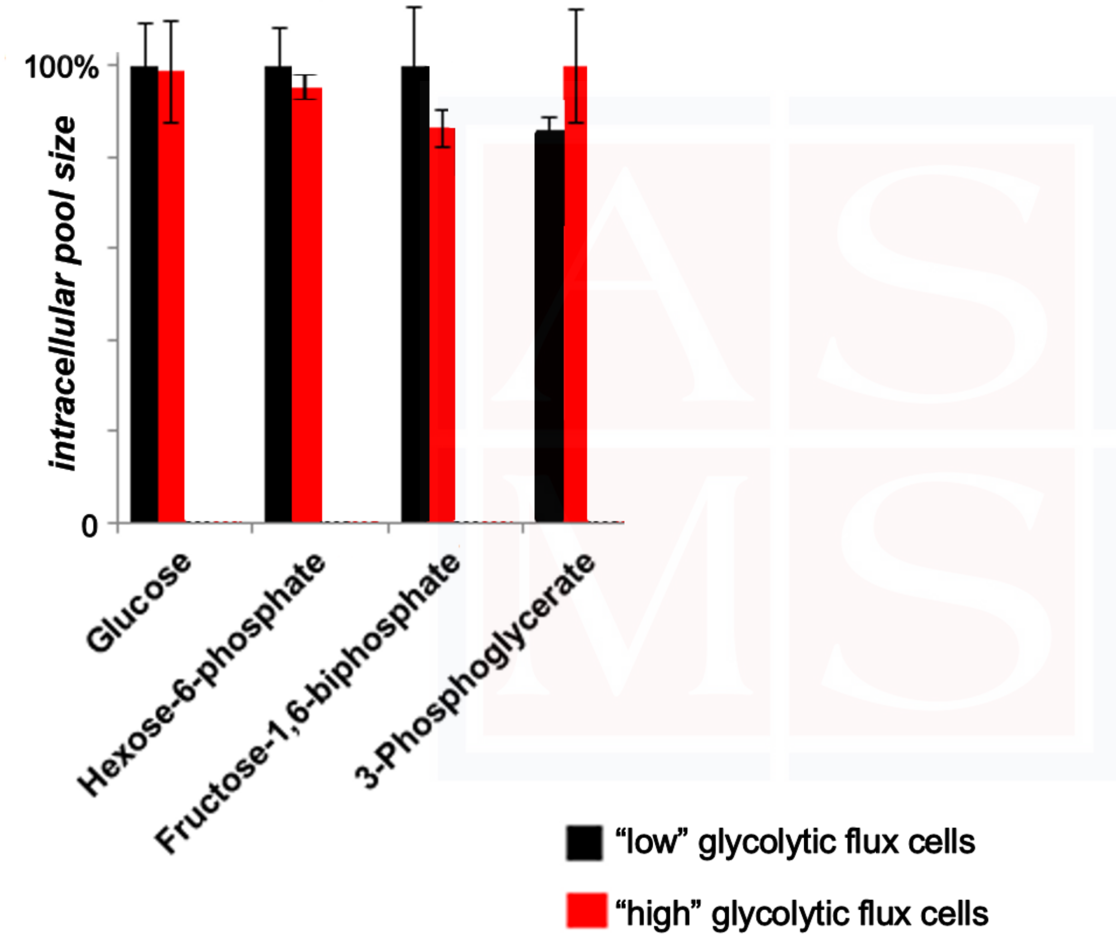
metabolite levels



metabolic fluxes*

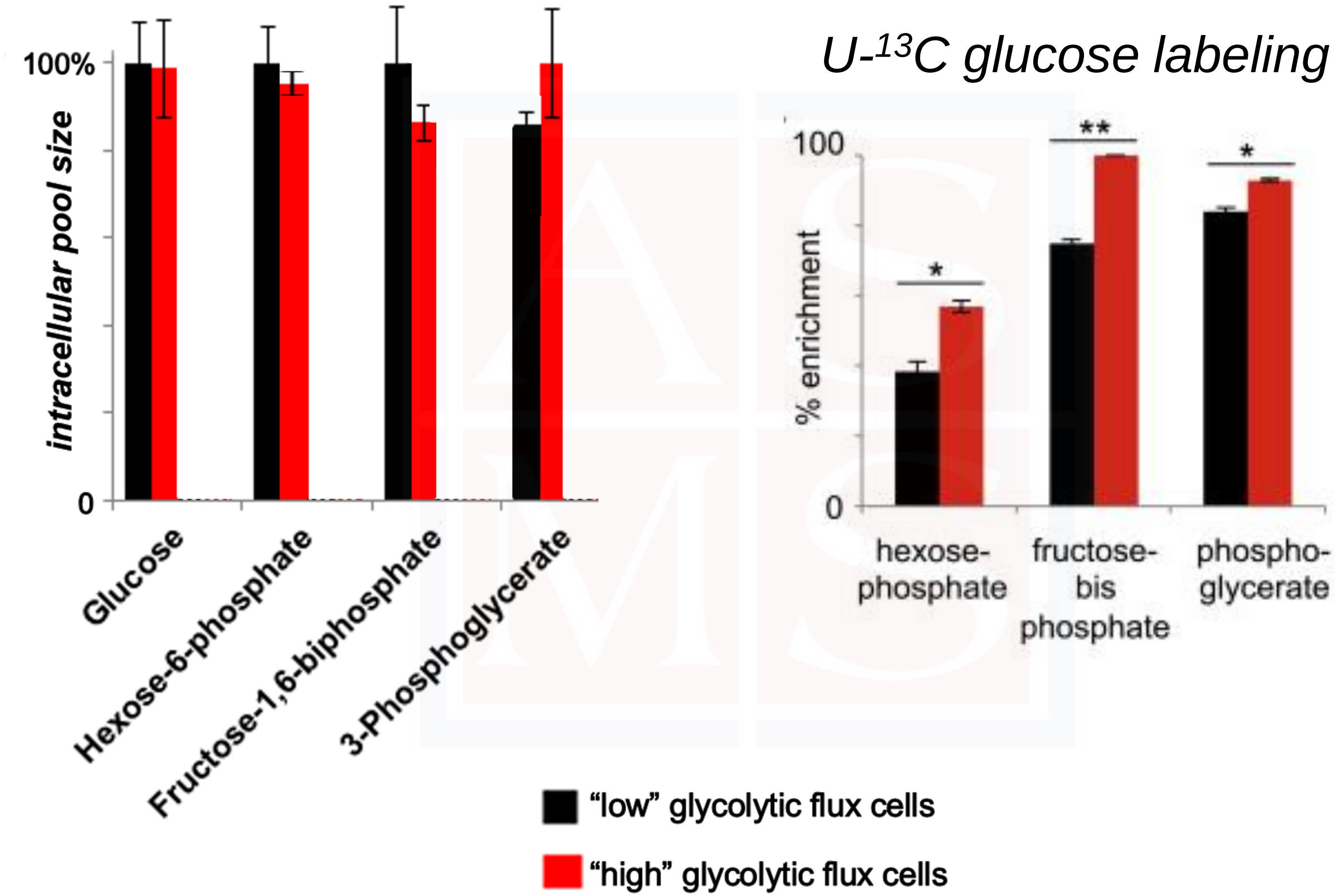
*when reviewing slides as PDF, please note that this is a video

Metabolomics vs metabolic flux



Yao et al, Cell Chemical Biology 2016

Metabolomics vs metabolic flux



Yao et al, Cell Chemical Biology 2016

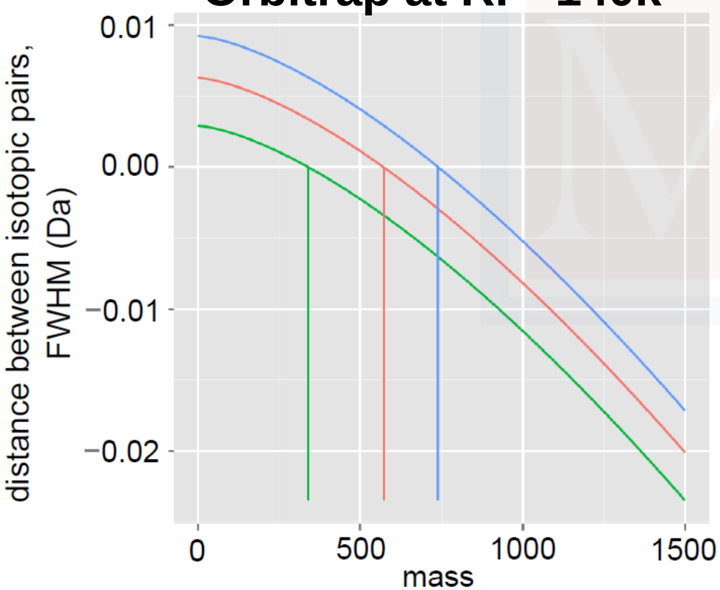
What samples can you label?

Sample	Pro	Con
micro-organisms	Defined media, highly developed computational models, less compartmentalization	Dynamic labeling requires short pulses (sec), limited relevance to disease
cells in culture	Pure cell types, easy to manipulate, glucose- and glutamine-free media, minimal label required (affordable)	Serum creates background of unlabeled material, serum constitution is variable, question of physiological relevance
plants, animals, and patients	Physiological relevance, small sample sizes amenable to MS	Expensive, multiple cell types that cannot deconvolve, computational models limited, high background levels of unlabeled material

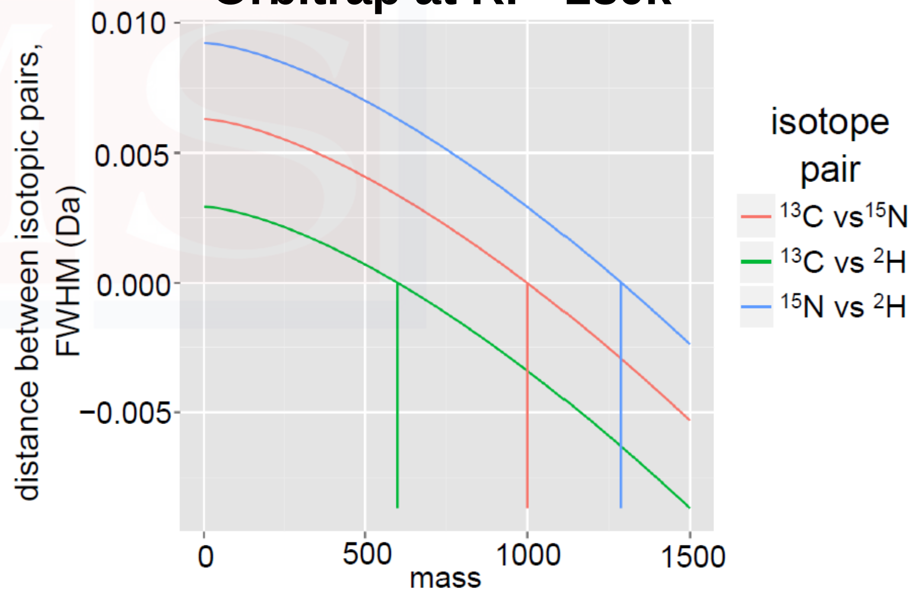
What isotope tracer should you use?

- ^{13}C most widely used and typically most informative.
- ^{15}N, ^{2}H, ^{18}O also used.
- Multiple isotopes also possible if can resolve.

Orbitrap at RP=140k

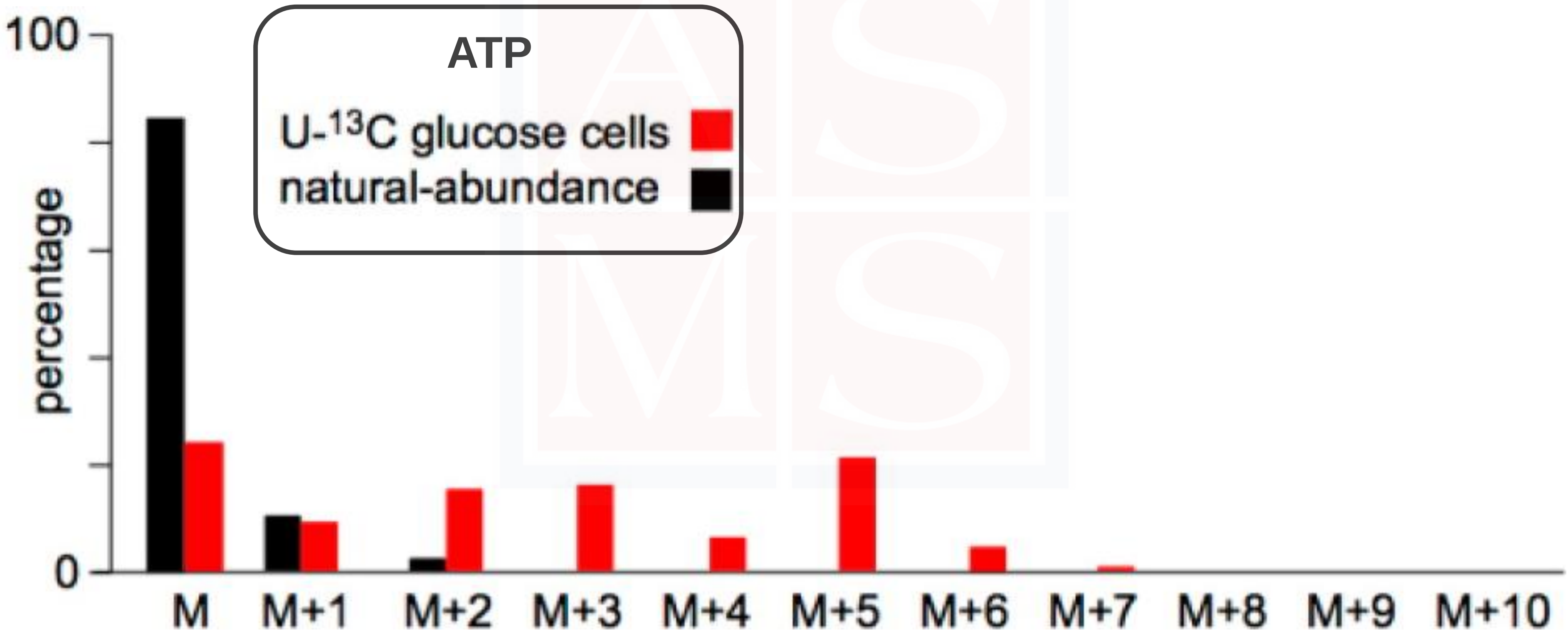


Orbitrap at RP=280k

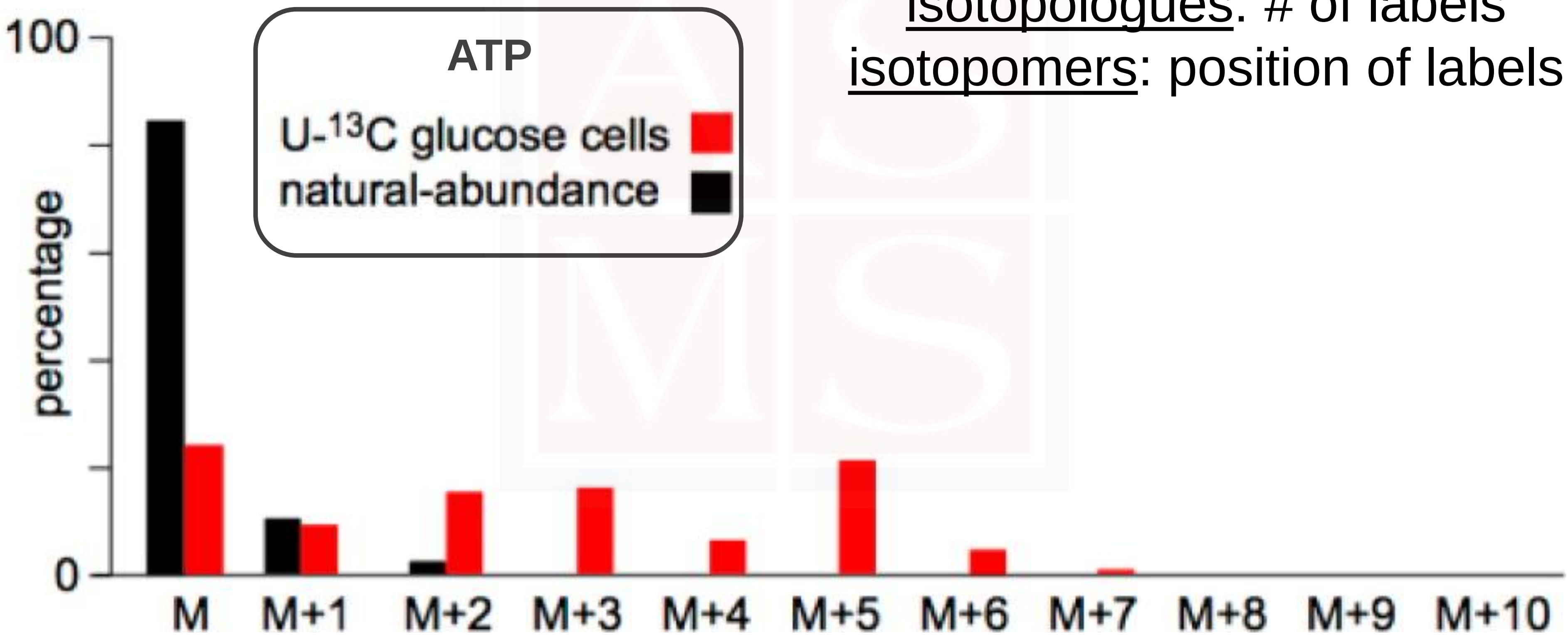


Nature of MS labeling data

Nature of MS labeling data



Nature of MS labeling data

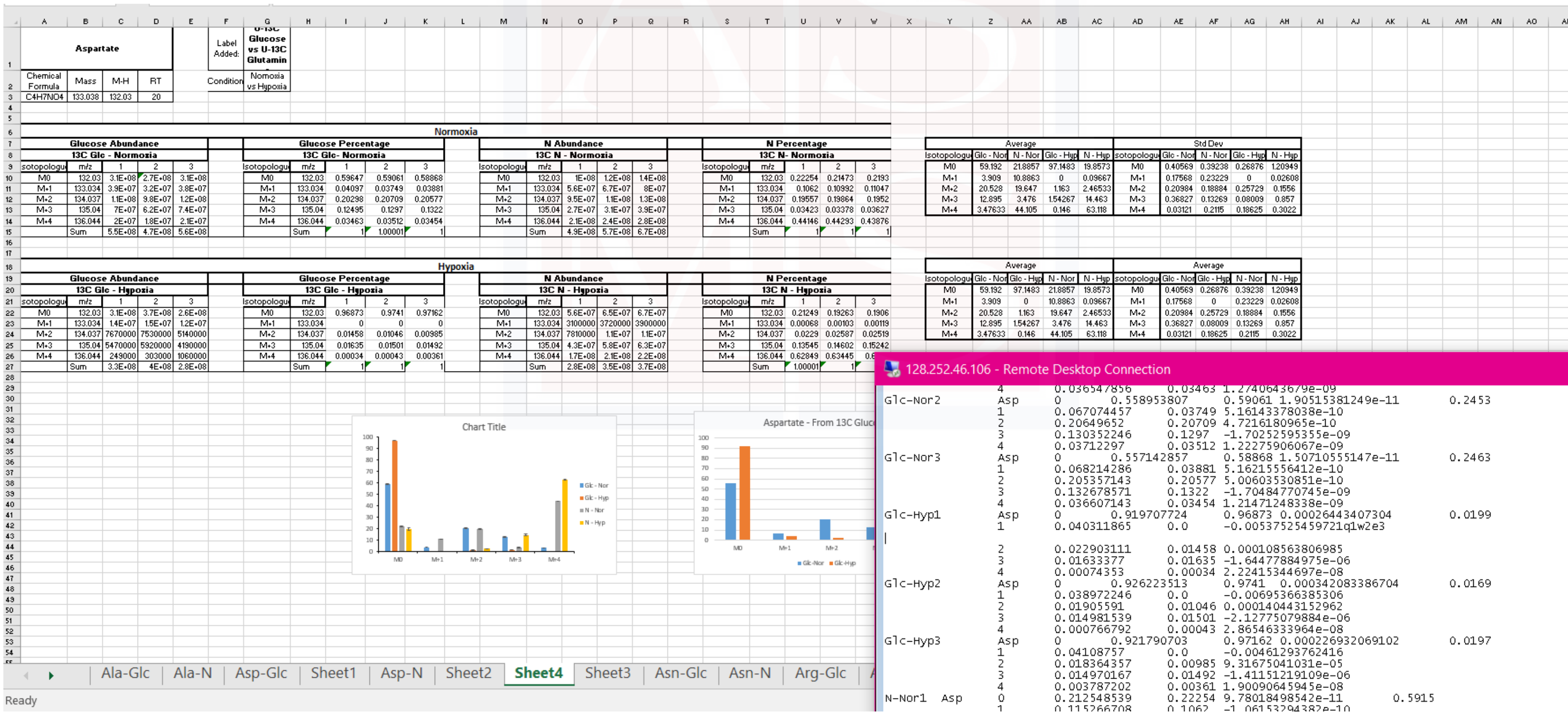


<u>isotopologues</u>: # of labels
<u>isotopomers</u>: position of labels

MS1 reports isotopolgues

Creating isotopologue plots

- Manual inspection: can be time intensive



Creating isotopologue plots

- Manual inspection: can be time intensive

  e.g., Palmitate has 17 isotopologues. If you have two sample groups with five replicates each, that is 170 EICs to inspect.

Creating isotopologue plots

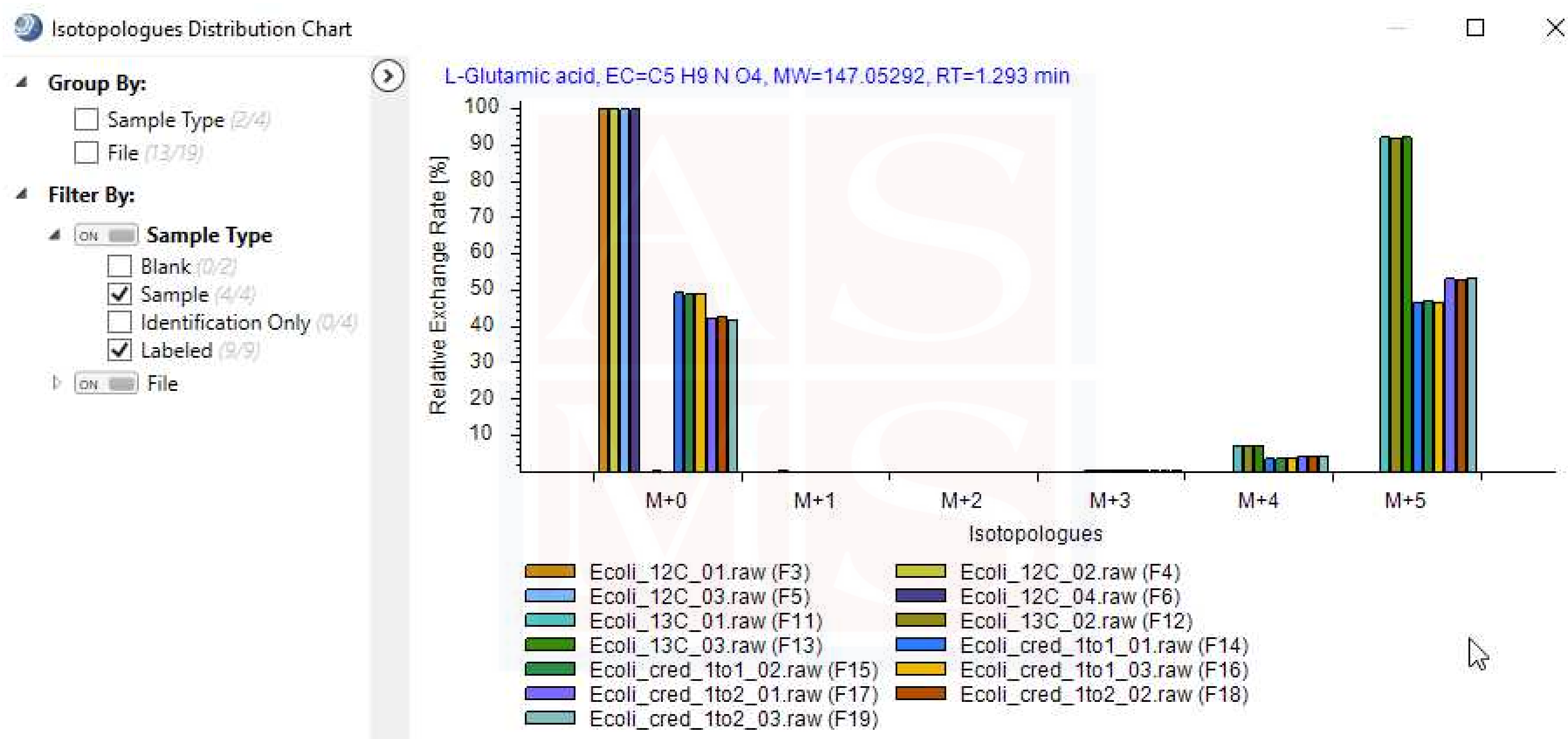
- Manual inspection: can be time intensive

  e.g., Palmitate has 17 isotopologues. If you have two sample groups with five replicates each, that is 170 EICs to inspect.

- Multiple vendor options

  - Agilent: VistaFlux
  - Thermo: Compound Discoverer 3.0

Creating isotopologue plots



Thermo’s CD 3.0

https://mycompounddiscoverer.com/

Step 1: correct for natural abundance

- If calculated manually, then can use stand-alone software such as isocor.
- Within commercial software, there is a correction button.
- It is not acceptable to simply subtract the natural-abundance MS data from labeled data
- Remember to correct for derivatization agents

Step 2: correct for isotopic impurity of the tracer

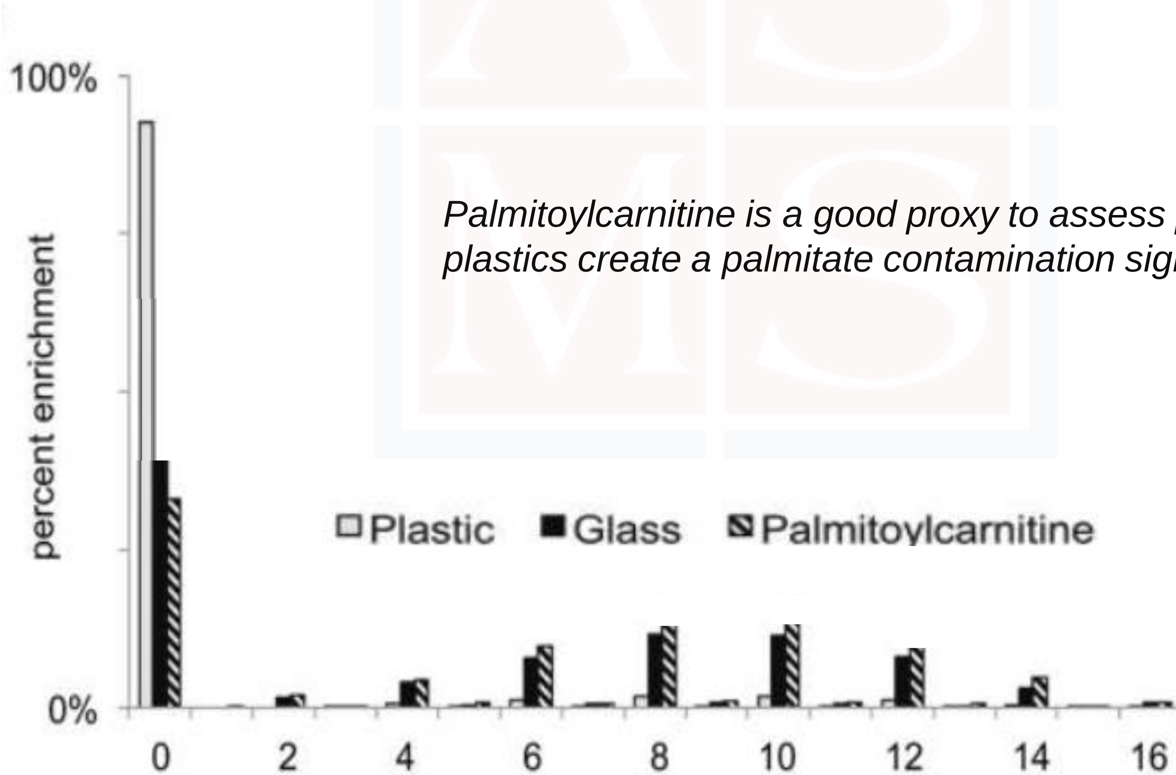
- Often this is negligible.

- An isotopic purity of 99% means that there is a 1% chance that a given carbon is ^{12}C instead of ^{13}C.

- With commercial software, there is a user-entry box to correct.

Note: beware of background contributions to unlabeled pool

- When using plastics, palmitate background is common and contributes to unlabeled signal.



Palmitoylcarnitine is a good proxy to assess palmitate labeling when plastics create a palmitate contamination signal.

Two general questions often considered with isotope tracers

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

2.) What is the flux of a metabolic pathway?

(flux: material flow per unit time)

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

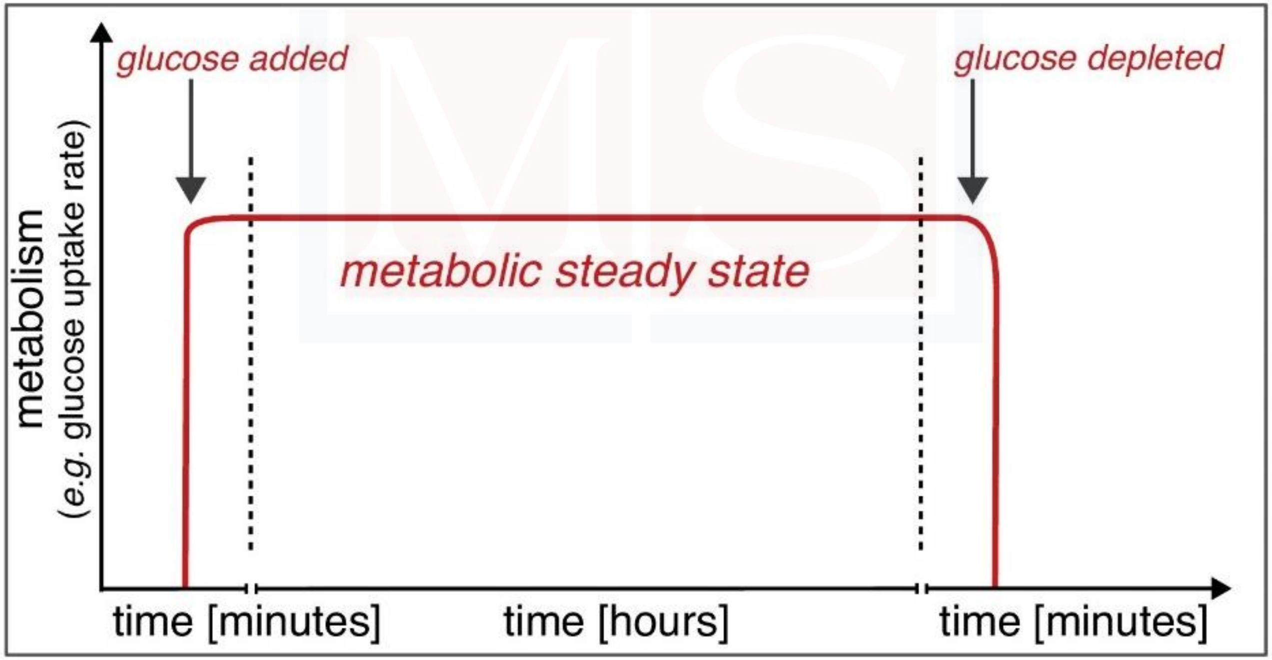
2.) What is the flux of a metabolic pathway?

(<u>flux</u>: *material flow per unit time*)

NOTE*: Just because an experiment uses a stable isotope tracer does not mean its experimental output is "flux".*

Some definitions relevant to isotope tracer analysis

Metabolic steady state: When intracellular metabolite concentrations and metabolic fluxes are constant with time. (e.g., continuous cultures maintain constant nutrient conditions throughout the experiment).

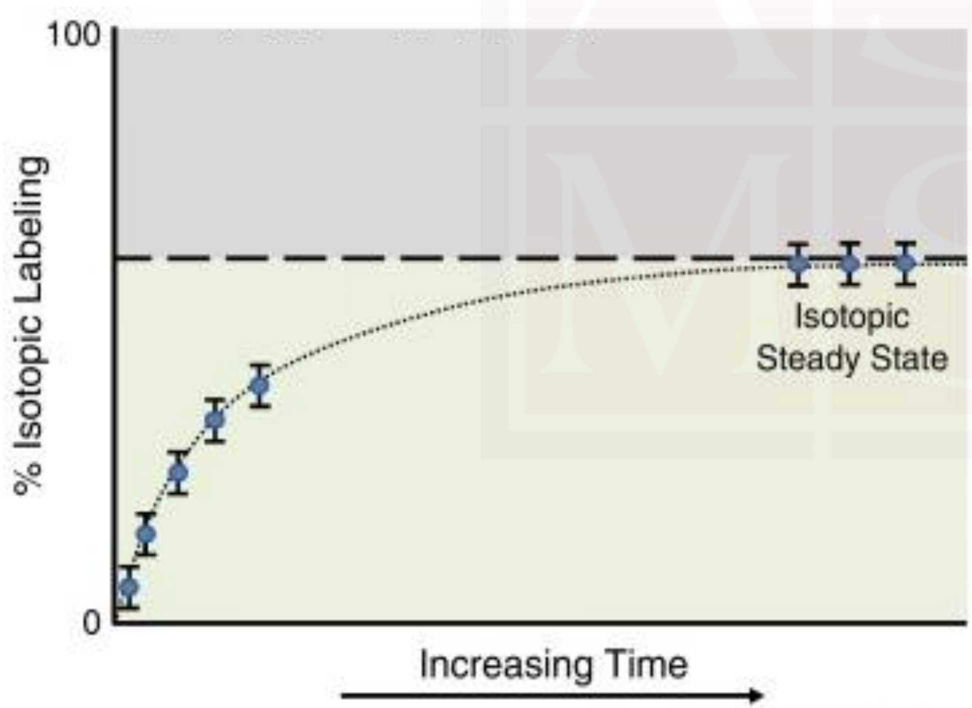


Some definitions relevant to isotope tracer analysis

Pseudo-steady state: When intracellular metabolite concentrations and metabolic fluxes change minimally with time. This is generally assumed during exponential growth phase.

Some definitions relevant to isotope tracer analysis

Isotopic steady state: When the labeling of a metabolite is constant with time.



For mammalian systems:
glycolysis, ~10 min
TCA cycle, ~2 h
nucleotides, ~15 h

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

2.) What is the flux of a metabolic pathway?

(flux: material flow per unit time)

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

2.) What is the flux of a metabolic pathway?

(flux: material flow per unit time)

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

Use fully labeled tracers because positionally labeled tracers are complicated by differential pathway usage.

Two general questions often considered with isotope tracers

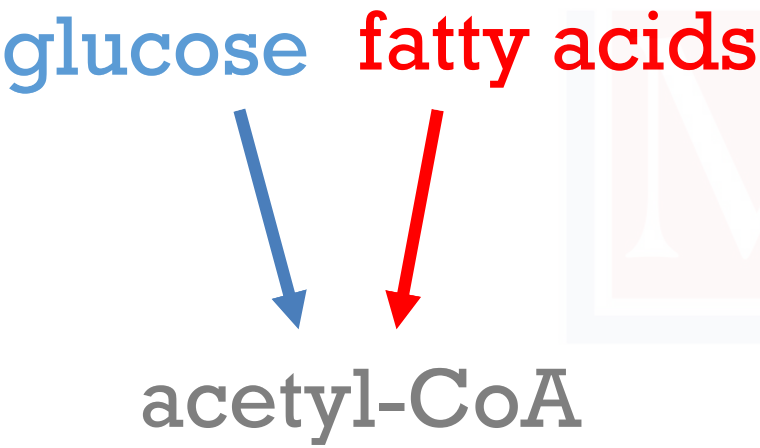
1.) What contribution does a nutrient make to the biosynthesis of X?

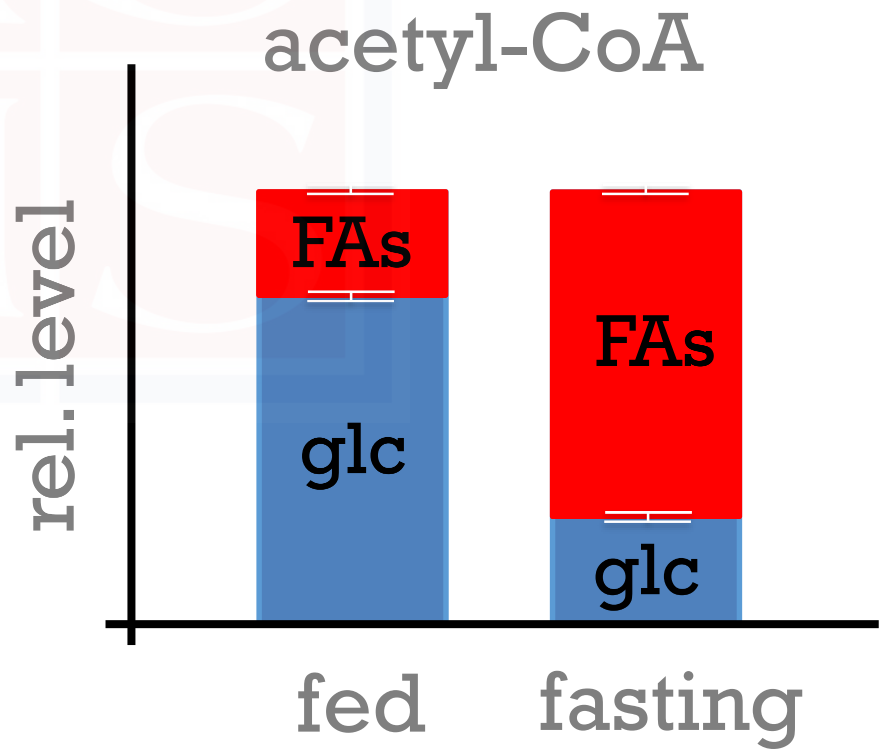
Use fully labeled tracers because positionally labeled tracers are complicated by differential pathway usage.

Easiest to do at isotopic steady state

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?





Puchalska et al, Cell Met 2019

increasing difficulty with MS

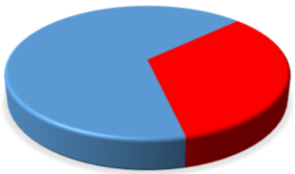
1. How much palmitate is labeled?
2. How much lipid is labeled?
3. How much label goes to palmitate or lipids?

increasing difficulty with MS

1. How much palmitate is labeled?

2. How much lipid is labeled?

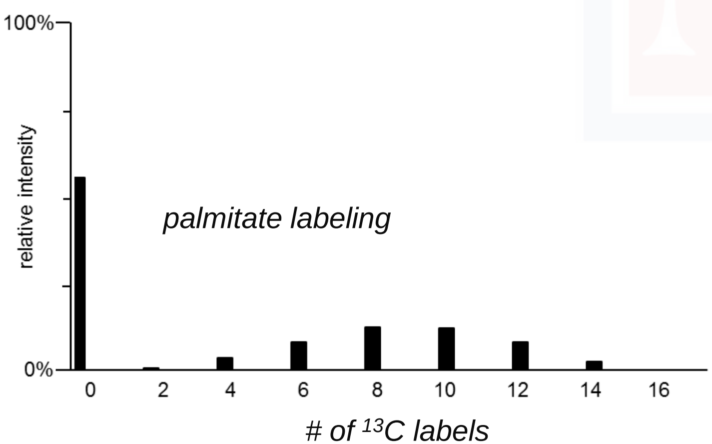
3. How much label goes to palmitate or lipids?



unlabeled palmitate

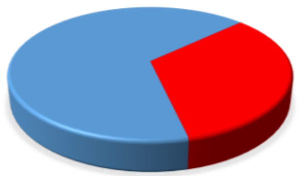
labeled palmitate

LC/MS can easily determine the percentage of a specific molecule that is labeled.



increasing difficulty with MS

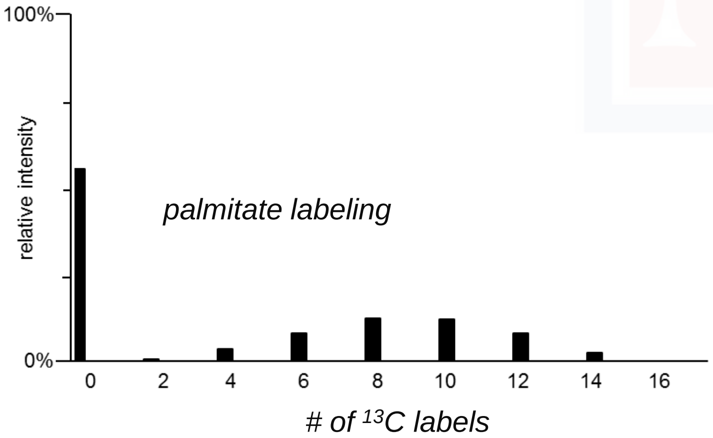
1. How much palmitate is labeled?



unlabeled palmitate

labeled palmitate

LC/MS can easily determine the percentage of a specific molecule that is labeled.



2. How much lipid is labeled?

unlabeled lipid

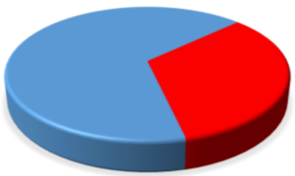
labeled lipid

Determining how much of a molecular class of compounds is labeled requires quantitating many compounds.

3. How much label goes to palmitate or lipids?

increasing difficulty with MS

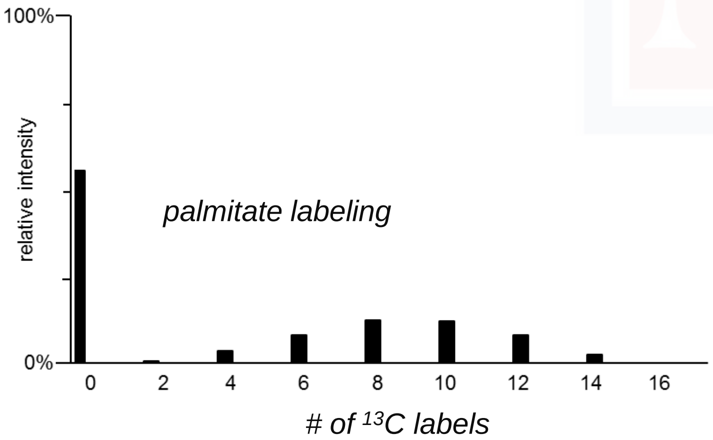
1. How much palmitate is labeled?



unlabeled palmitate

labeled palmitate

LC/MS can easily determine the percentage of a specific molecule that is labeled.



2. How much lipid is labeled?

unlabeled lipid

labeled lipid

Determining how much of a molecular class of compounds is labeled requires quantitating many compounds.

3. How much label goes to palmitate or lipids?

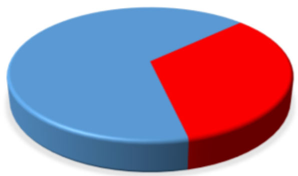
Labeled molecules that are not lipids

labeled lipids

Determining how the label is partitioned requires quantitating all labeled molecules.

increasing difficulty with MS

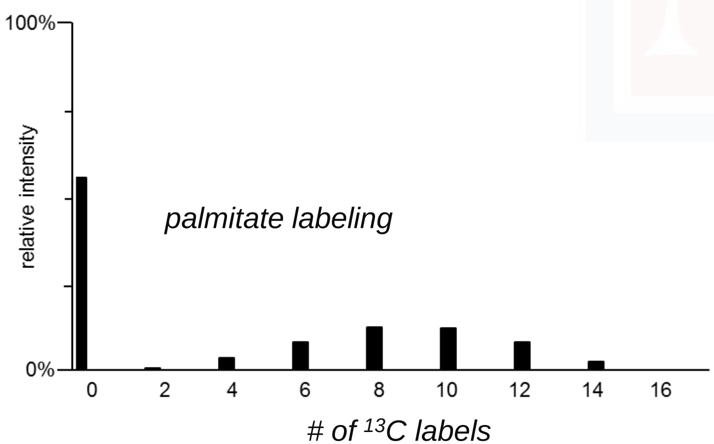
1. How much palmitate is labeled?



unlabeled palmitate

labeled palmitate

LC/MS can easily determine the percentage of a specific molecule that is labeled.



2. How much lipid is labeled?

unlabeled lipid

labeled lipid

Determining how much of a molecular class of compounds is labeled requires quantitating many compounds.

3. How much label goes to palmitate or lipids?

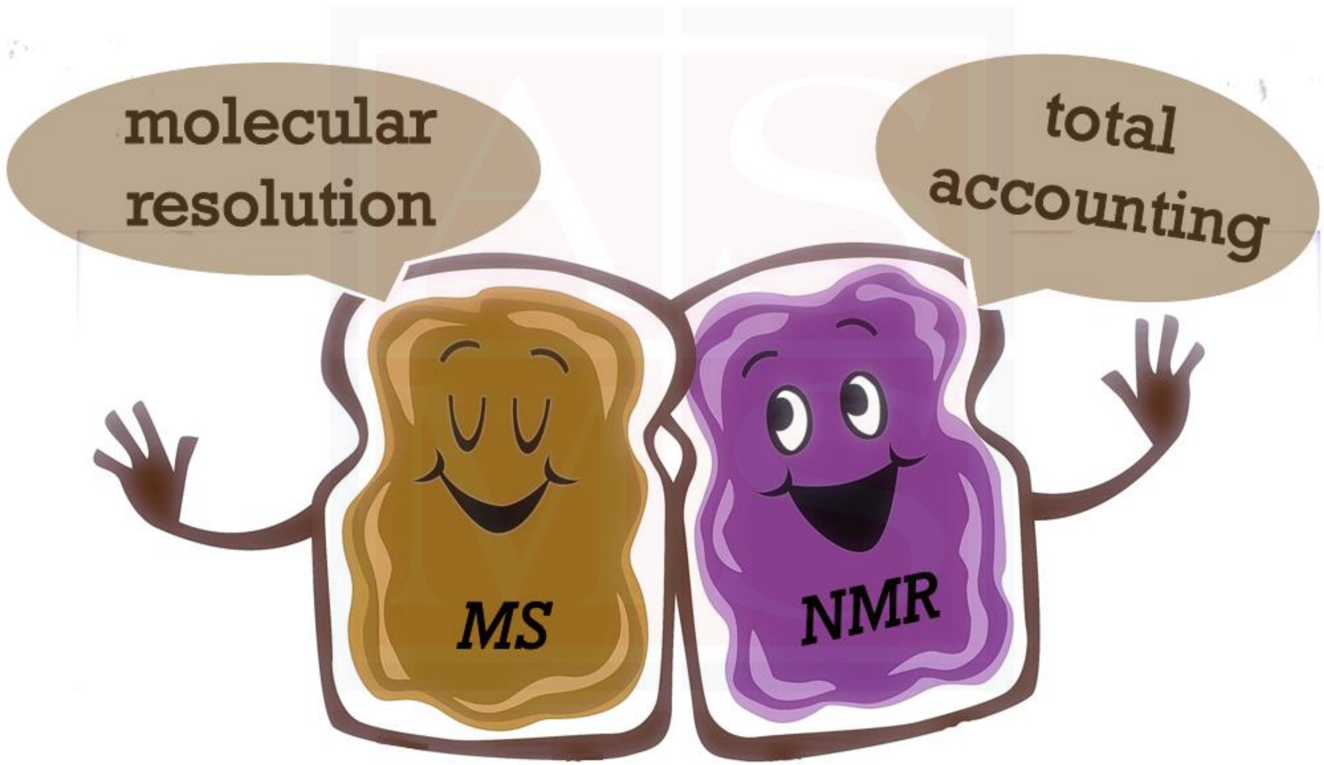
Labeled molecules that are not lipids

labeled lipids

Determining how the label is partitioned requires quantitating all labeled molecules.

NMR

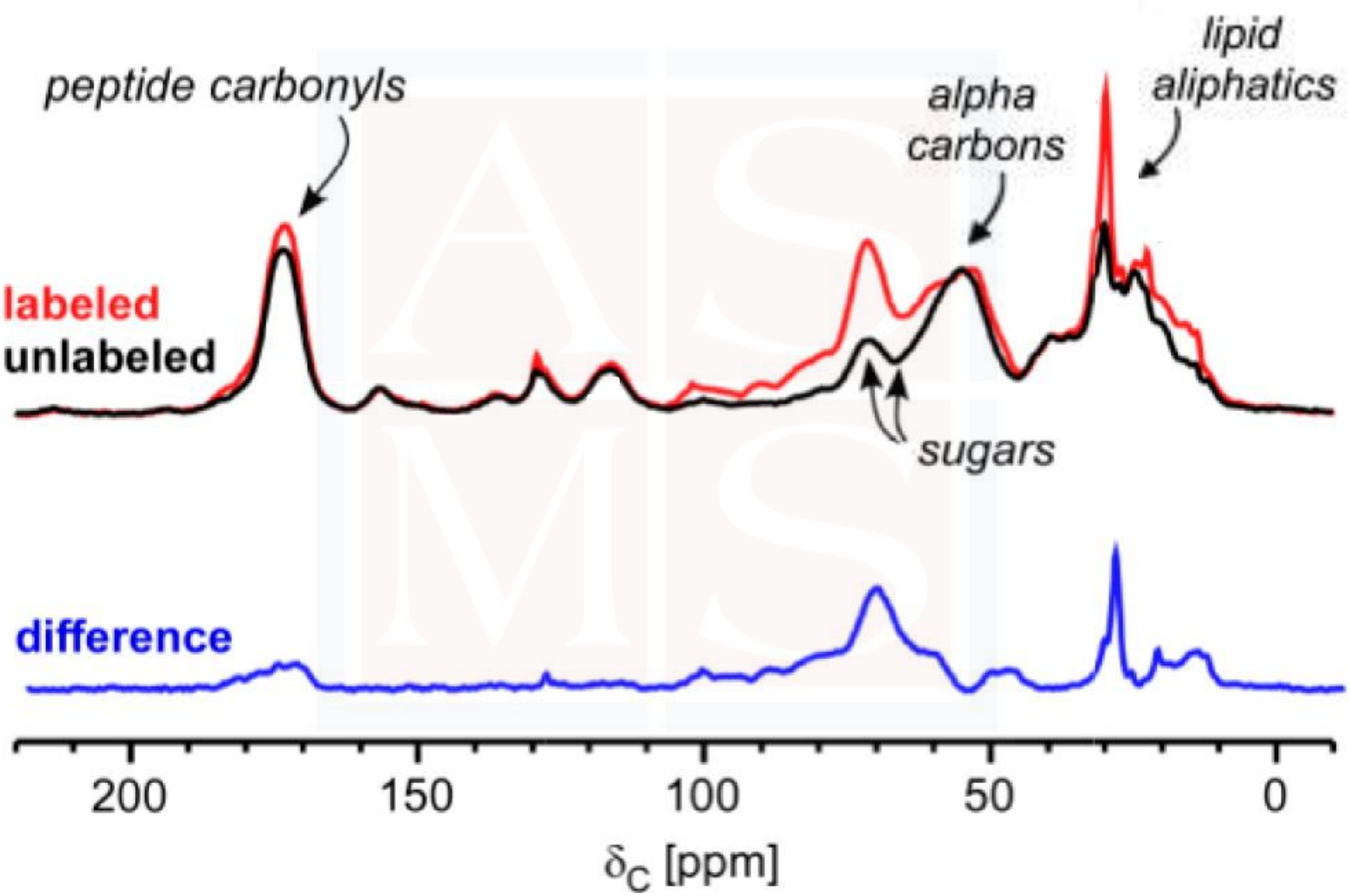
Why can't we be friends?



Nutrient contributions by solid-state NMR

intact cells and tissues

Nutrient contributions by NMR



^{13}C CPMAS NMR of H460 cells

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

2.) What is the flux of a metabolic pathway?

(flux: material flow per unit time)

Two general questions often considered with isotope tracers

1.) What contribution does a nutrient make to the biosynthesis of X?

2.) What is the flux of a metabolic pathway?

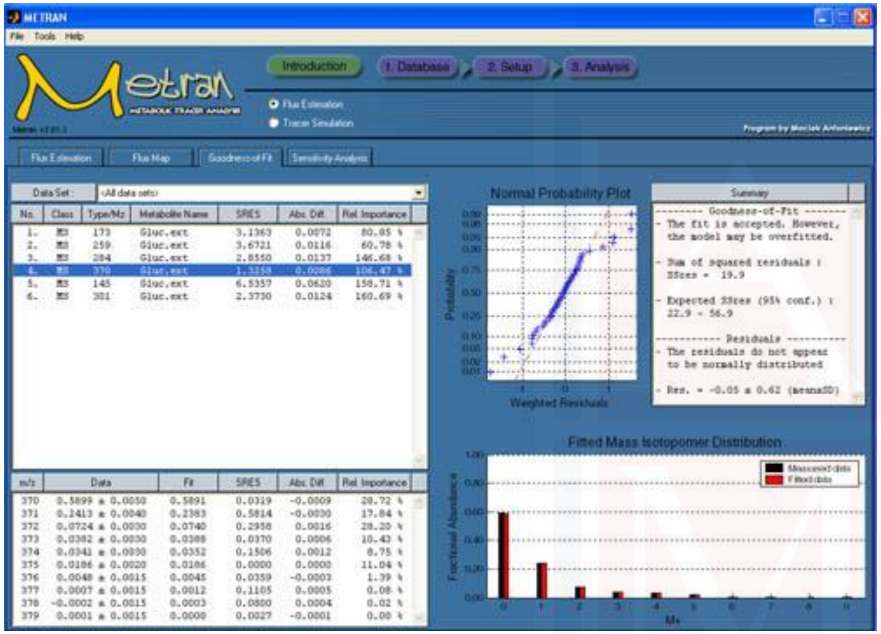
(flux: material flow per unit time)

Formal metabolic flux analysis

- Computationally intensive, requires model.
- Input isotope labeling patterns (free metabolites or proteic amino acids), metabolite concentrations, nutrient uptake rates, and/or metabolite excretion rates.
- Can be isotopic steady state (concentration independent) or dynamic (concentration dependent).
- Usually preferred over flux balance analysis, which requires an objective function (most useful for *E. coli*).
- Well-established software programs available.

Software programs for formal metabolic flux analysis

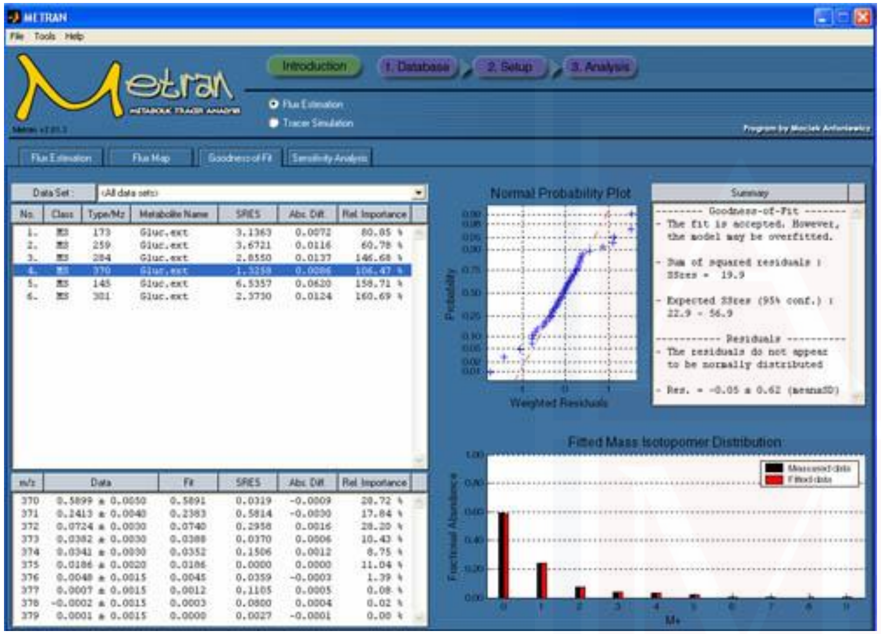
METRAN



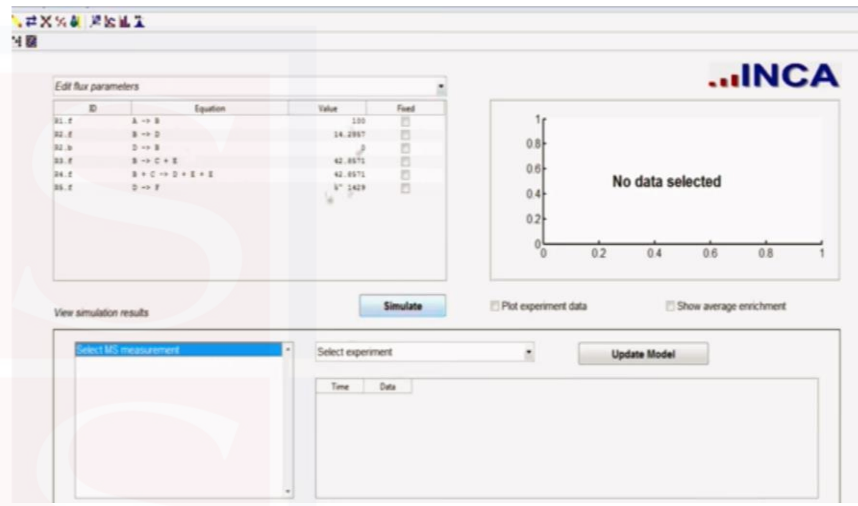
- free for academic users
- intuitive graphical user interface
- user-defined network models
- confidence intervals of fluxes

Software programs for formal metabolic flux analysis

METRAN



Isotopomer Network Compartmental Analysis (INCA)



- free for academic users
- intuitive graphical user interface
- user-defined network models
- confidence intervals of fluxes

- free for academic users
- intuitive graphical user interface
- high flexibility
- confidence intervals of fluxes

Assumptions of formal metabolic flux analysis*

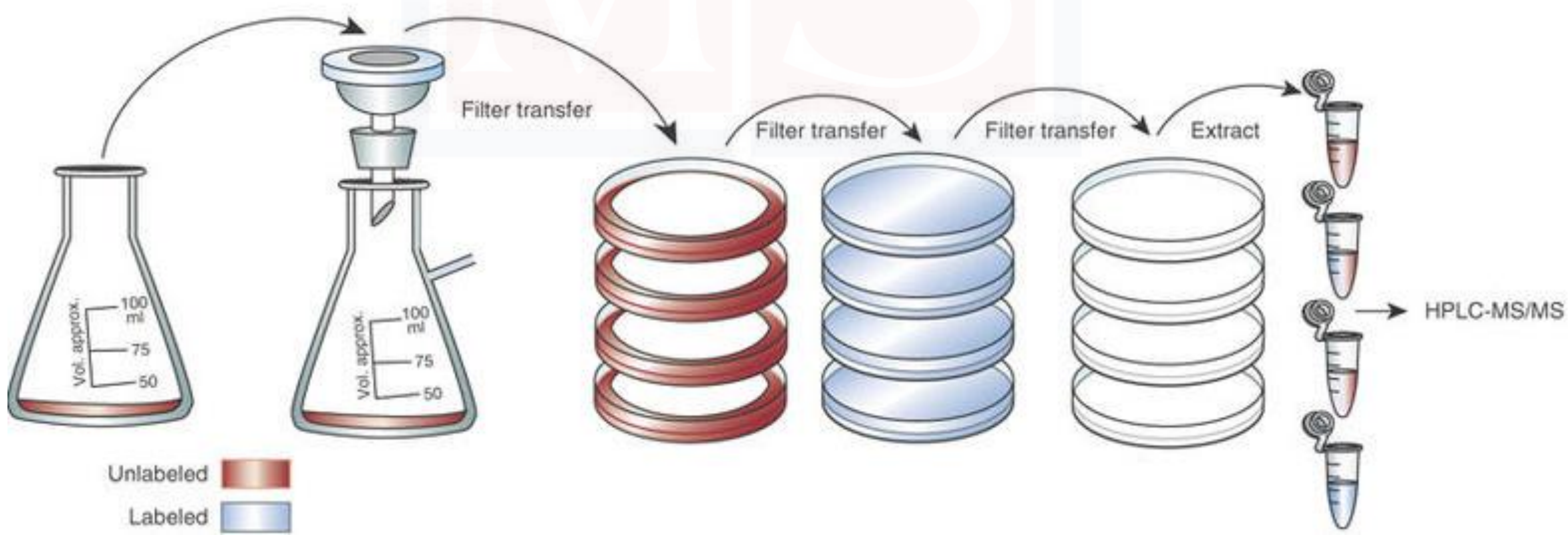
- Metabolite fluxes are constant during labeling exp
- No kinetic isotope effect
- No metabolite channeling
- Homogenous mixing within compartments
- Homogeneous cell populations
- No turnover of macromolecules (protein breakdown...)

*if incorrect, must adjust models

Other experimental methods to get quantitative flux information

Kinetic flux profiling:

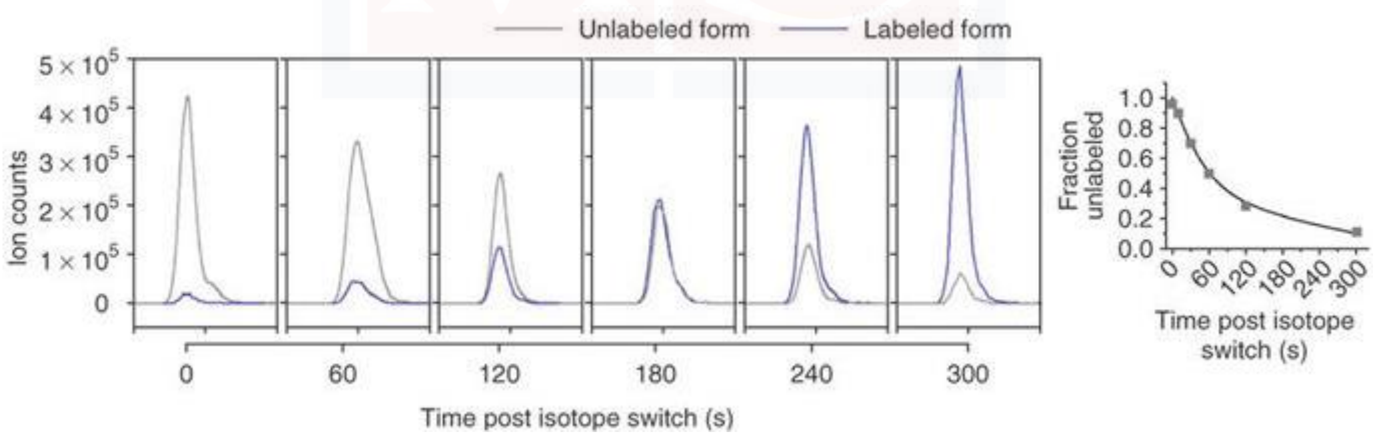
- More experimentally demanding than MFA at isotopic steady state because need multiple time points (sometimes <1 min after label is introduced).



Other experimental methods to get quantitative flux information

<u>Kinetic flux profiling</u>:

- More experimentally demanding than MFA at isotopic steady state because need multiple time points (sometimes <1 min after label is introduced).



Other experimental methods to get quantitative flux information

Isotopomer spectral analysis (ISA):

- Much simpler computational approach for condensation biosynthesis reactions of the stoichiometry: nA→B, where n is an integer >1.
  (e.g., 8 acetyl-CoA→palmitate,
  18 acetyl-CoA→cholesterol)

Other experimental methods to get quantitative flux information

Isotopomer spectral analysis (ISA):

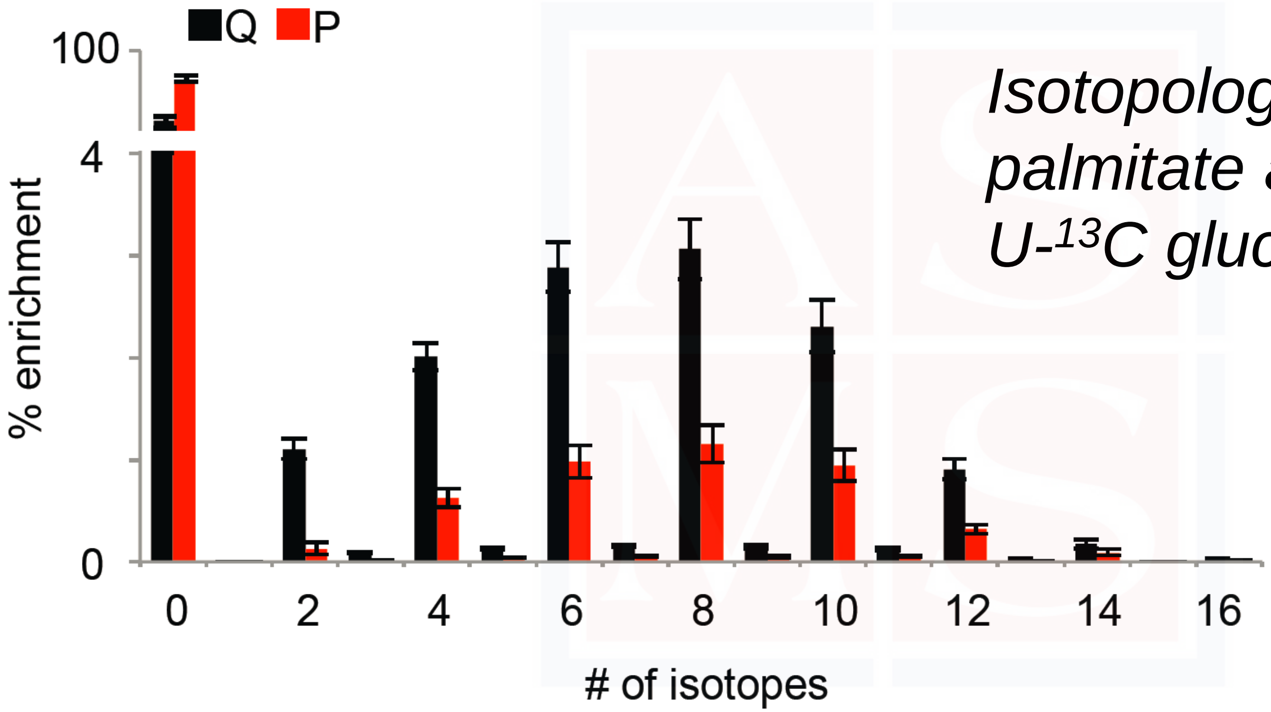
- Much simpler computational approach for condensation biosynthesis reactions of the stoichiometry: nA→B, where n is an integer >1.
  (e.g., 8 acetyl-CoA→palmitate,
  18 acetyl-CoA→cholesterol)

Input: labeling pattern of B
Output: **g(t)**, fraction of B synthesized from A during exp. and **D**, enrichment of precursor A pool

Qualitative analysis of metabolic fluxes

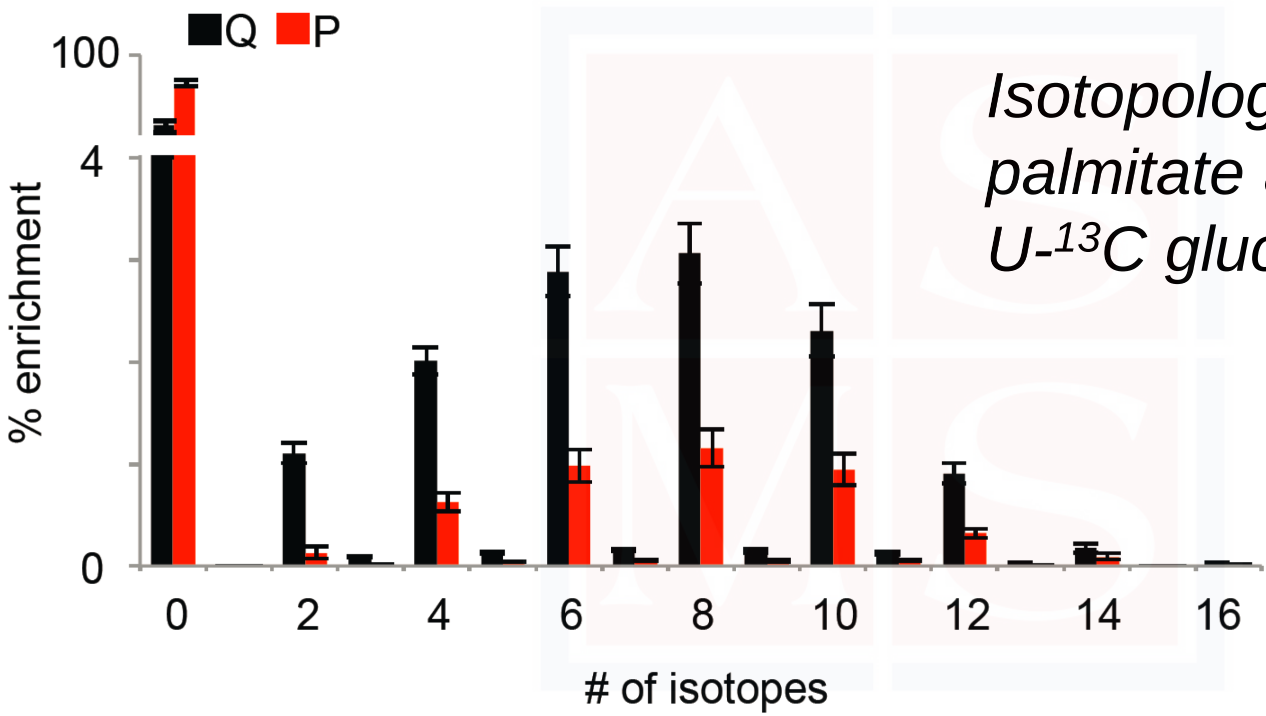
- Commonly used
- Often requires deep understanding of metabolism
- May not require any models or software programs
- Frequently uses positionally labeled nutrients for easier to interpret results

Qualitative analysis of metabolic fluxes



Isotopologue distribution of palmitate after labeling with U-^{13}C glucose for 6 h.

Qualitative analysis of metabolic fluxes



Isotopologue distribution of palmitate after labeling with U-^{13}C glucose for 6 h.

More labeling of a compound does not necessarily mean higher flux! Only indicates an alteration in the associated flux distribution.

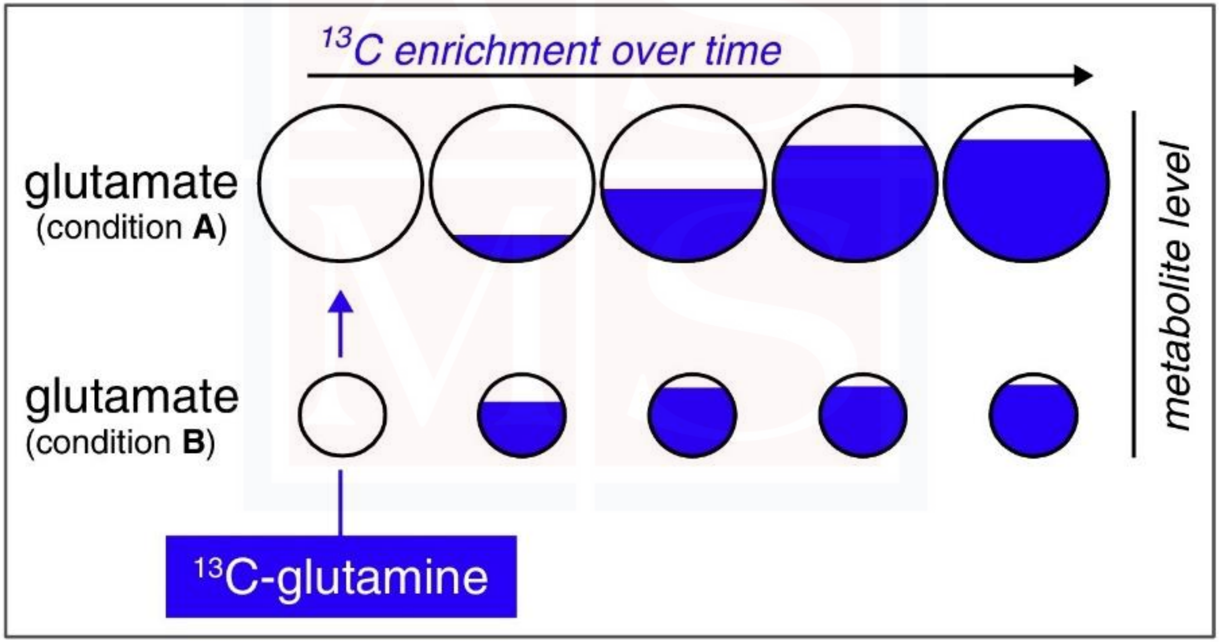
Qualitative analysis of metabolic fluxes

<u>Some explanations for increased palmitate labeling</u>:

- Increased lipogenic flux from glucose
- Decreased lipogenic flux from another unlabeled substrate (e.g., glutamine).
- Increased uptake of unlabeled palmitate from media
- Decrease in palmitate pool size
- Some combination of the above

Qualitative analysis of metabolic fluxes

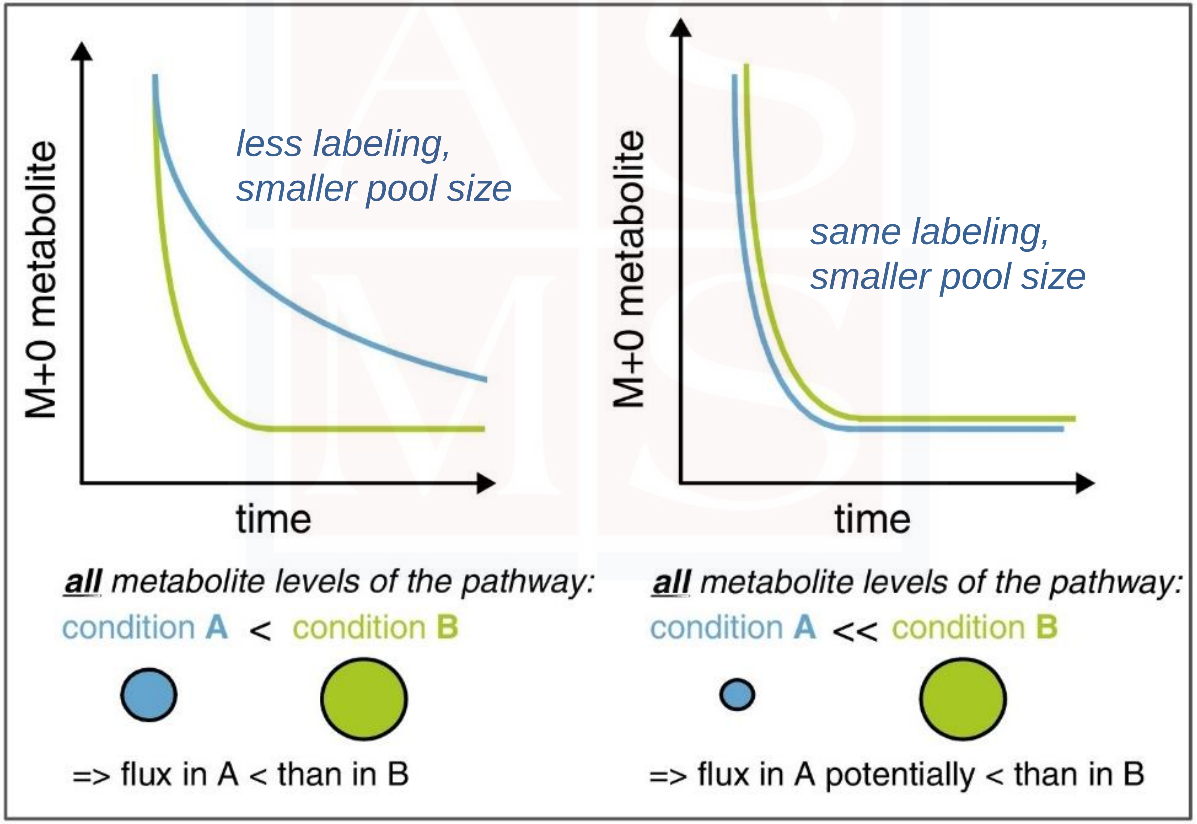
Effect of pool size on isotopic labeling pattern



Buescher et al, COBT 2015

Qualitative analysis of metabolic fluxes

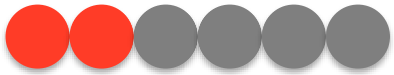
Effect of pool size on isotopic labeling pattern



Buescher et al, COBT 2015

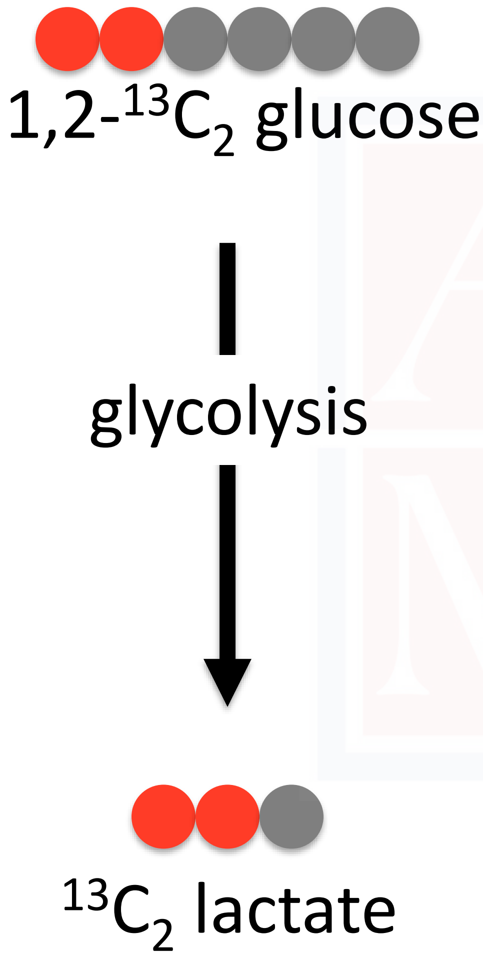
<u>Qualitative flux, example 1</u>: Relative flux of pentose phosphate pathway overflow to glycolysis

Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis

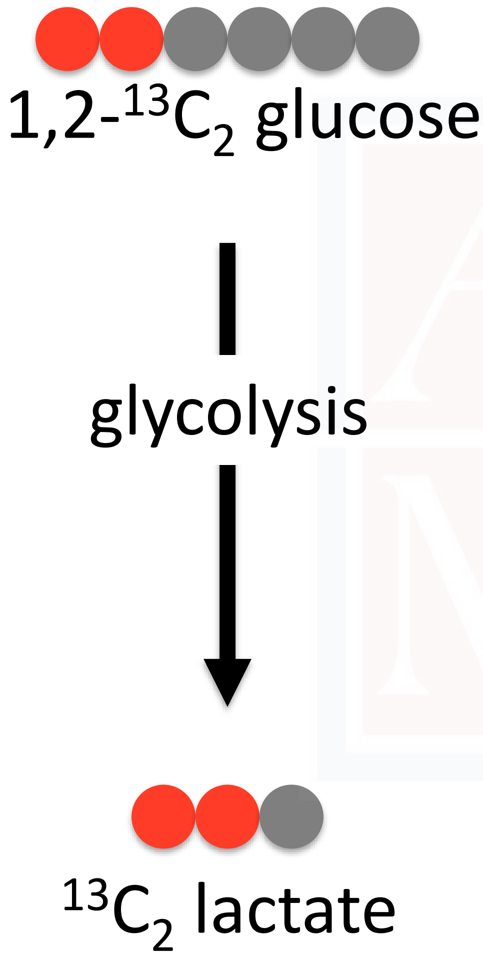


1,2-$^{13}C_2$ glucose

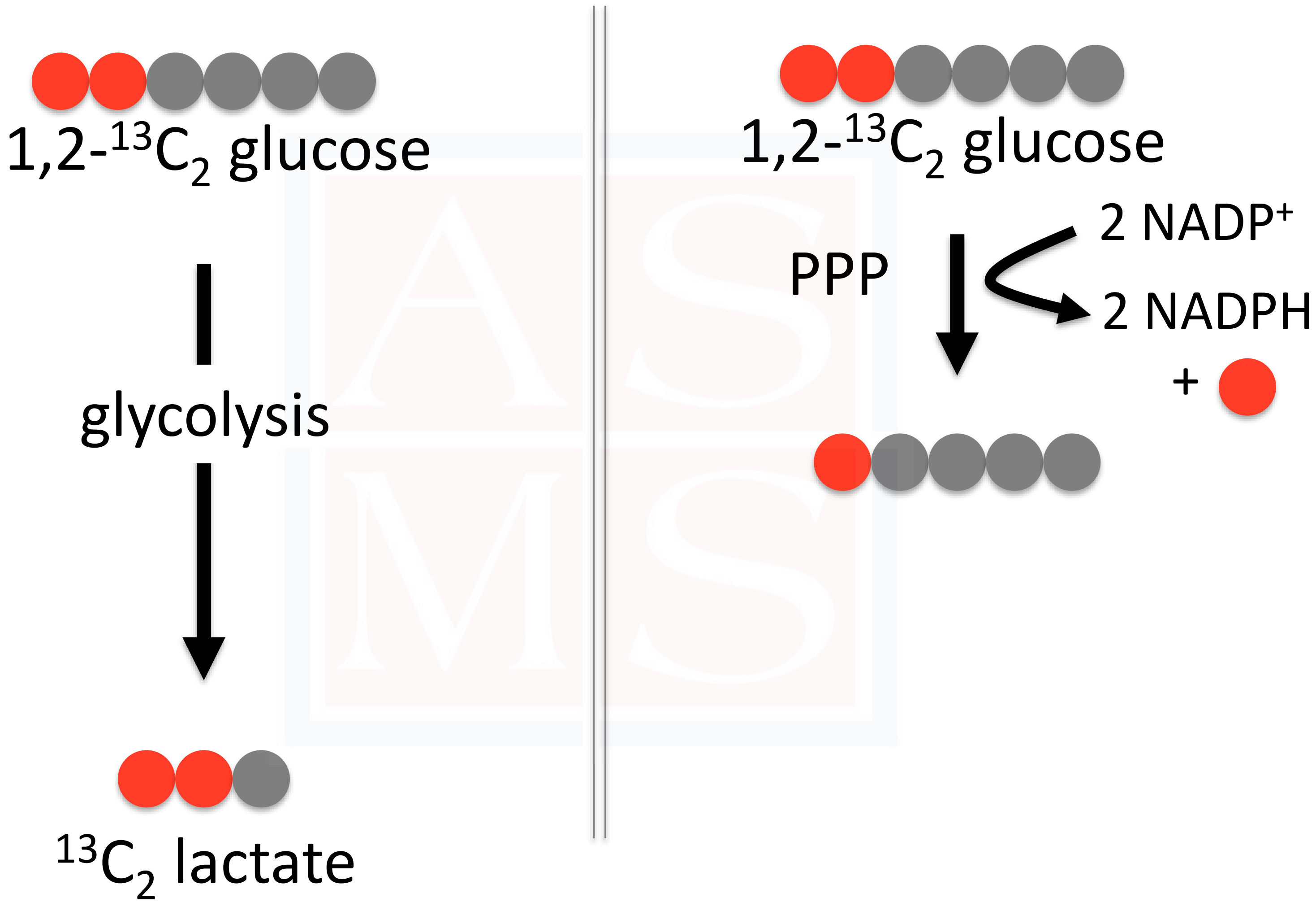
Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis



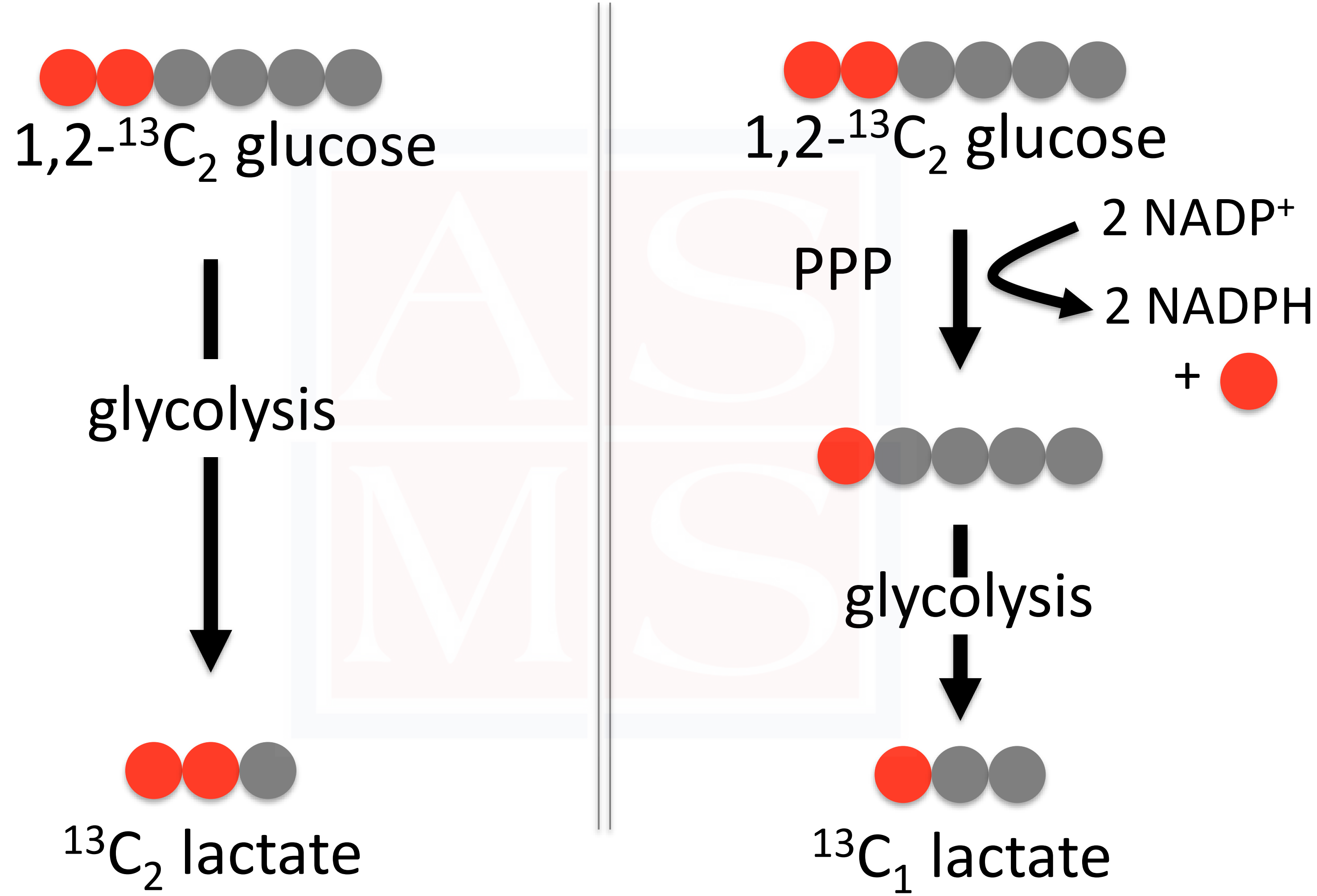
Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis



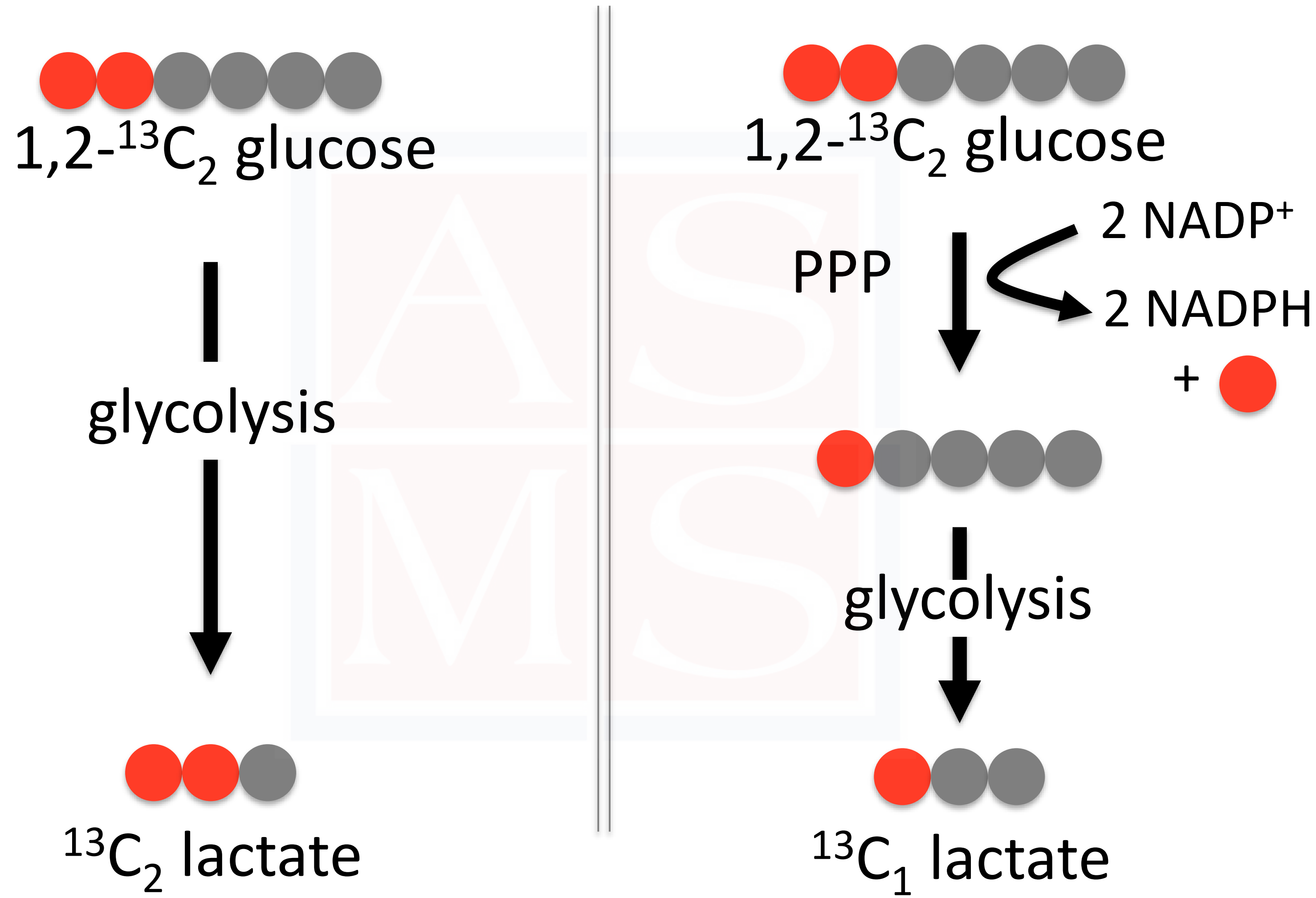
<u>Qualitative flux, example 1</u>: Relative flux of pentose phosphate pathway overflow to glycolysis



Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis

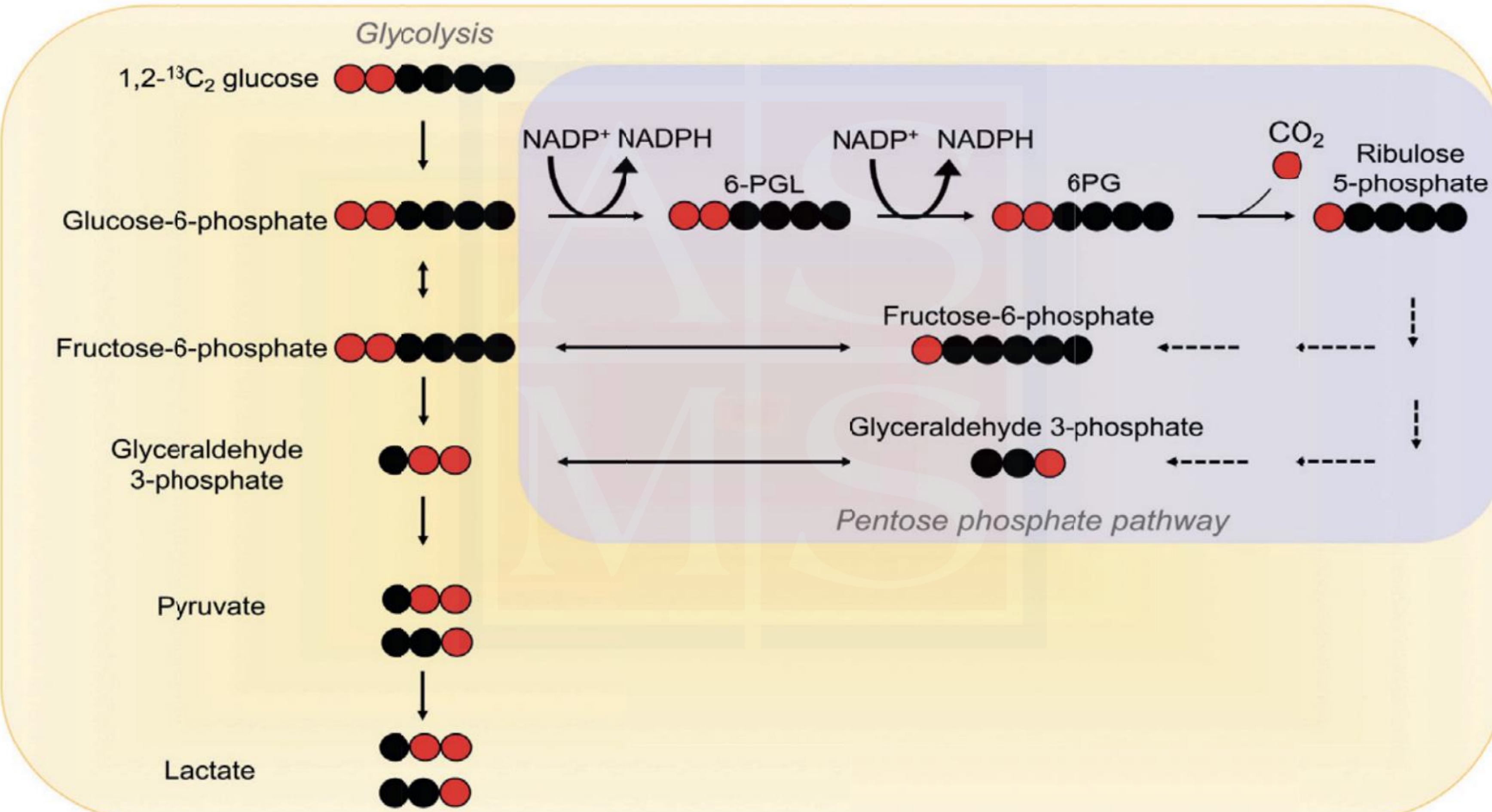


Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis



Limitation: assumes non-oxidative PPP is feeding carbon back into glycolysis

Qualitative flux, example 1: Relative flux of pentose phosphate pathway overflow to glycolysis



Others:

Table 2. Isotopic Tracers for Measuring Pathway Activities

Application	Tracer	Metabolite readouts	Interpretation
Pentose phosphate pathway (PPP)			
PPP overflow	[1,2-^{13}C]glucose	Lactate M+1, M+2	Flux through the combined oxidative and non-oxidative PPP generates M+1 lactate from [1,2-^{13}C]glucose, while glycolysis generates only M+2 lactate (Lee et al., 1998). LacM+1 / LacM+2 reflects ratio of PPP overflow to glycolysis.
Source of ribose (oxidative versus non-oxidative branch of PPP)	[1,2-^{13}C]glucose	Ribose phosphate M+1, M+2	The oxPPP make M+1 ribose phosphate; the non-oxPPP makes M+2. Ratio of M+1/M+2 depends on the gross flux (net flux + exchange flux) of each branch: Reversibility of the non-oxPPP can make M+2 even if all net ribose production is by oxPPP.
Glycolysis, TCA and gluconeogenesis			
Glycolytic rate	[U-^{13}C]glucose	FBP Dihydroxyacetone phosphate 3-phosphoglycerate	Higher flux yields faster labeling. Labeling results should be confirmed by glucose uptake and lactate excretion measurements.
Reversibility of glycolysis	50%: 50% mix of [U-^{12}C]: [U-^{13}C] glucose	Glucose-6-phosphate M+3 FBP M+3	Feeding a mixture of labeled and unlabeled glucose results in unlabeled and M+3 triose phosphates. Reversibility of aldolase produces M+3 FBP. Fructose bisphosphatase activity yields M+3 glucose-6-phosphate (Park et al., 2016).
Gluconeogenesis	[U-^{13}C]lactate [U-^{13}C]glutamine	Glucose M+2, M+3 Glucose-6-phosphate M+2, M+3 3-phosphoglycerate M+2, M+3	Lactate and glutamine are major TCA feedstocks. Flux from TCA to glycolysis catalyzed by PEPCK results in triose phosphate labeling. Fructose bisphosphatase activity then makes labeled hexose phosphates.
Pyruvate carboxylase contribution to TCA	[3-^{13}C]glucose [1-^{13}C]pyruvate	Aspartate M+1 Malate M+1	C1 of pyruvate comes from glucose C3/C4. Pyruvate C1 is lost in making acetyl-CoA, but can enter TCA via pyruvate carboxylase which makes M+1 oxaloacetate and thus M+1 aspartate and M+1 malate (Sellers et al., 2015).
Reductive carboxylation ("backwards" TCA flux)	[U-^{13}C]glutamine [1-^{13}C]glutamine	Citrate M+5, Malate M+3 or Citrate M+1, Malate M+1	Reductive carboxylation of α-ketoglutarate (derived from labeled glutamine) produces M+5 citrate from [U-^{13}C]glutamine and M+1 citrate from [1-^{13}C]glutamine, and subsequent ATP citrate lyase produces M+3 or M+1 malate, respectively (Yoo et al., 2008)
TCA carbon sources	[U-^{13}C]nutrients	Succinate Malate Citrate α-ketoglutarate	Carbon enrichment (number of ^{13}C atoms versus total carbon atoms) reflects carbon contribution from the nutrient; useful *in vivo* with correction for circulating nutrient enrichment (Davidson et al., 2016; Faubert et al., 2017; Hui et al., 2017)
Biosynthesis			
Acetyl-CoA sources	[U-^{13}C]glucose [U-^{13}C]glutamine [U-^{13}C]acetate	Fatty acids (saponified) Acetyl amino acids	Fatty acids (e.g., palmitate) are made from stochastic condensation of labeled and unlabeled acetyl-CoA. Acetyl group labeling can be inferred by binomial fitting of fatty acid labeling or by comparing steady-state labeling of acetyl-amino acids and the corresponding free amino acids.
De novo fatty acid biosynthesis	$^{2}H_2O$	Fatty acids (saponified)	$^{2}H_2O$ labels newly synthesized fat directly and via NADPH, with 21 potential deuterium per palmitate (Lee et al., 1994; Zhang et al., 2017).

Jang et al., Cell 2018

Others:

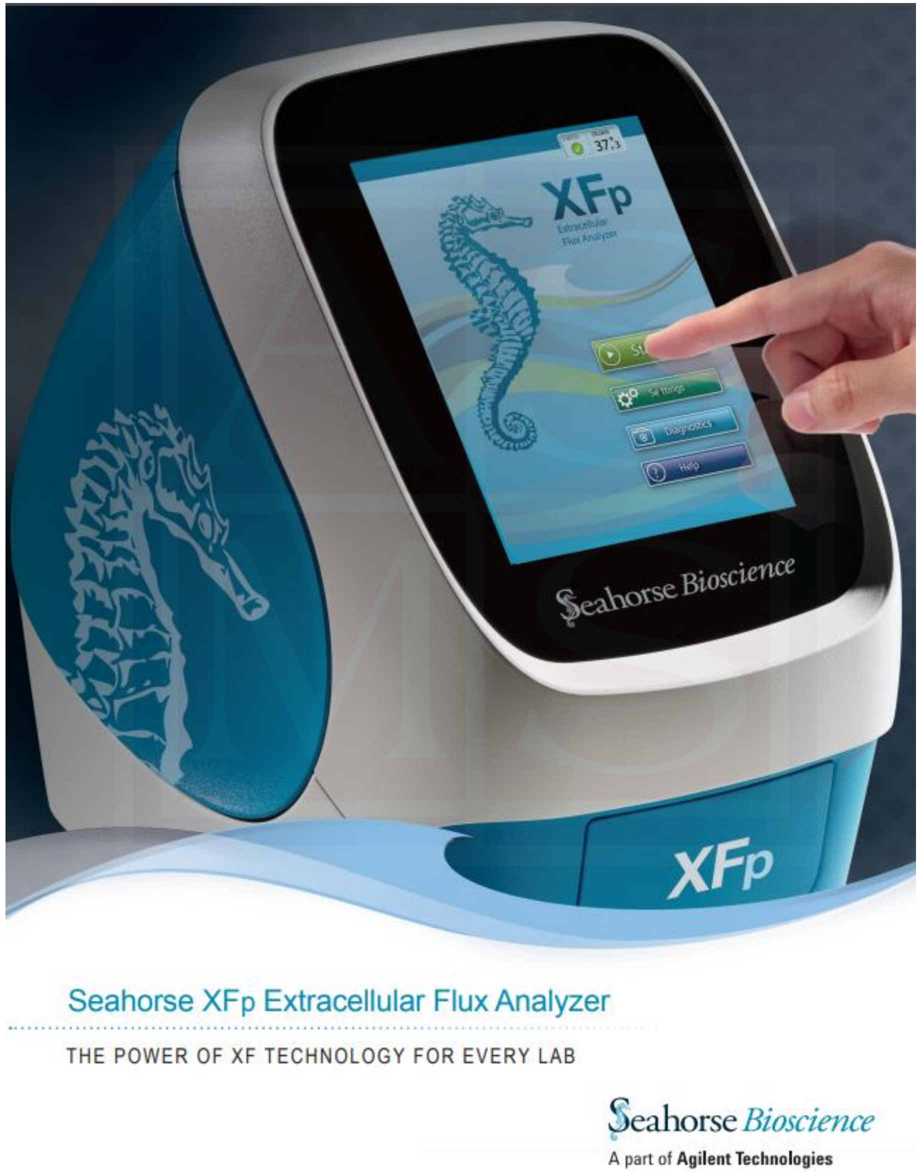
Table 2. *Continued*

Application	Tracer	Metabolite readouts	Interpretation
Purine biosynthesis	[U-^{13}C]glycine	ATP M+2 GTP M+2	Purine ring contains a glycine moiety. Newly synthesized purines are M+2.
Pyrimidine biosynthesis	[U-^{13}C]bicarbonate [U-^{15}N]glutamine [U-^{13}C]glutamine	UTP UDP-glucose	Pyrimidines are made from carbonyl phosphate (which contains one bicarbonate and one glutamine nitrogen) and aspartate (which typically contains glutamine nitrogen and carbon (Strong et al., 1983).
Protein synthesis	$^{2}H_2O$ [U-^{13}C]essential amino acids	Amino acids (hydrolyzed from protein)	^{2}H from $^{2}H_2O$ incorporates into non-essential amino acids (Busch et al., 2006). Essential AA are directly incorporated.
One-carbon metabolism			
De novo synthesis of serine	[U-^{13}C]glucose	Serine M+3	Serine is made from glucose via the glycolytic intermediate 3-phosphoglycerate. Fraction of serine M+3 indicates fraction serine made by *de novo* synthesis (Locasale et al., 2011)
Source of folate 1C units	[3-^{13}C]serine [U-^{13}C]glycine [U-^{13}C]sarcosine [U-^{13}C]formate	dTTP M+1 ATP M+1, M+2, M+3, M+4 Formyl-methionine M+1 Formate M+1	dTTP contains a 1C unit from cytosolic methylene-THF. Purine rings contain two 1C units from cytosolic formyl-THF. Formyl-methionine contains a 1C unit from mitochondrial formyl-THF. Excess 1C units are secreted as formate (Ducker et al., 2016). Note that purine rings also contain an intact glycine; thus, ATP M+2 may be from glycine not 1C.
Location of serine catabolism to make cytosolic 1C units	[2,3,3-^{2}H]serine	dTTP M+1, M+2	Direct cytosolic production of methylene-THF by SHMT1 yields dTTP M+2. The more circuitous route from mitochondrial SHMT2 yields dTTP M+1 (Herbig et al., 2002; Ducker et al., 2016).
Methylation through SAM	[Methyl-^{13}C,^{2}H$_3$] methionine	Methylated lysine (free or on histones)	Histones are methylated by SAM with the methyl group from methionine (Zee et al., 2010).
Redox metabolism			
NADH production from GAPDH	[4-^{2}H]glucose	NADH M+1 Lactate M+1 (compare to NAD, pyruvate)	GAPDH transfers the ^{2}H of glyceraldehyde-3-phosphate, derived from [4-^{2}H]glucose, to NADH. The ^{2}H can then be transferred to lactate by LDH (Lewis et al., 2014).
NADPH sources	[1-^{2}H]glucose [3-^{2}H]glucose [4-^{2}H]glucose [2,3,3-^{2}H]serine	NADPH (compare to NADP) Fatty acids (saponified) 2-hydroxyglutarate	The oxPPP makes NADPH from [1-^{2}H]glucose (G6PD) and [3-^{2}H]glucose (PGD) (Fan et al., 2014). Malic enzyme and isocitrate dehydrogenase make NADPH from malate and isocitrate, which can be labeled indirectly via [4-^{2}H]glucose (Liu et al., 2016). Folate metabolism makes NADPH from ^{2}H-serine. ^{2}H can be transferred to fatty acids or 2-hydroxyglutarate (whose production can be induced by mutant IDH expression) (Lewis et al., 2014).
Hydrogen-deuterium exchange between NADPH and water	$^{2}H_2O$	NADPH (compare to NADP) Fatty acids (saponified)	NADPH redox-active hydrogen undergoes water exchange catalyzed by Flavin enzymes. Knowledge of the fraction of NADPH undergoing exchange is required to determine the quantitative contribution of the oxPPP and other NADP reduction pathways (Zhang et al., 2017).
Glutathione biosynthesis	[U-^{13}C]glycine [U-^{13}C]glutamine	Glutathione	Glutathione is made from glutamate, cysteine, and glycine. Glutamine is a main source of glutamate (Mak et al., 2017).

Jang et al., Cell 2018

Easier way to measure flux?

Easier way to measure flux?



Easier way to measure flux?

Seahorse XF analyzers

Measures two things:

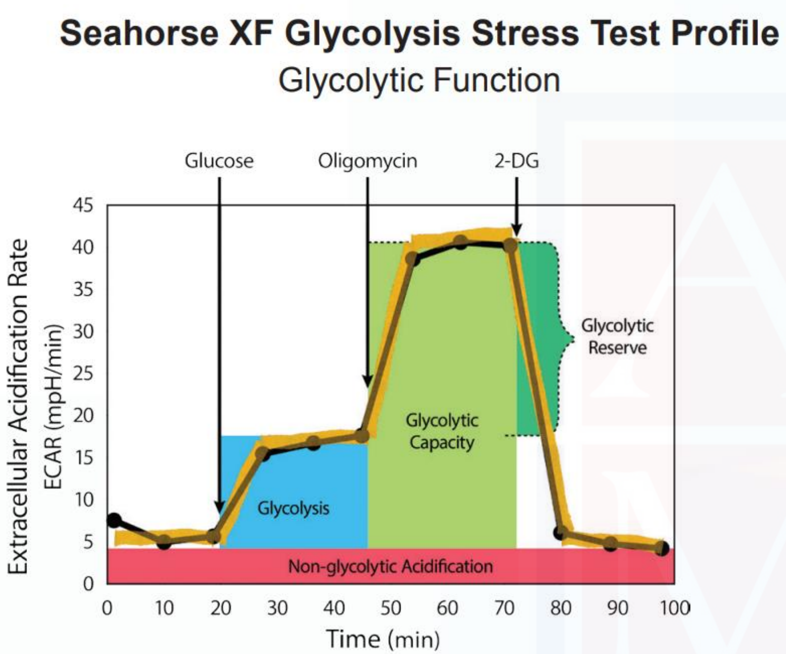
(1) ECAR (extracellular acidification rate)

- Lactate excretion
- Not exceptionally helpful if have a mass spectrometer

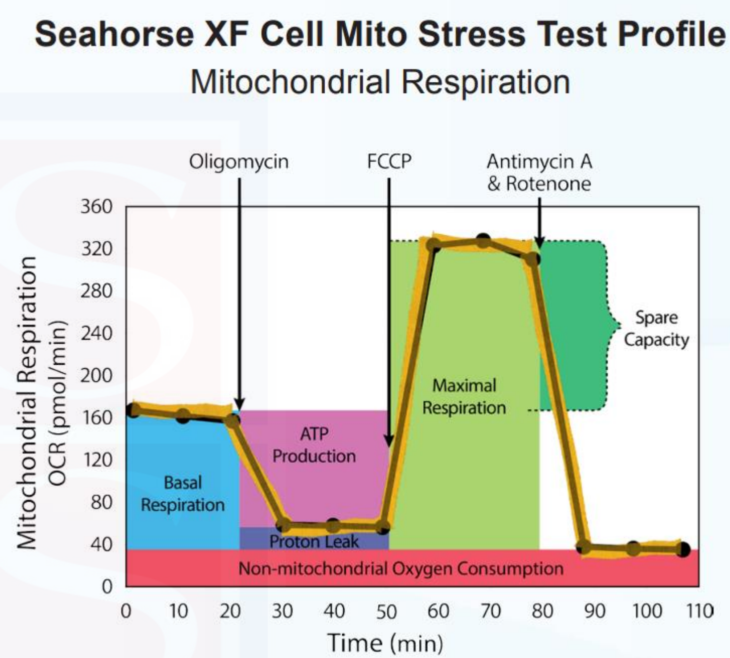
(2) OCR (oxygen consumption rate)

- Other respirometers available (OROBOROS)
- Insightful and complementary to MS or NMR data

Easier way to measure flux?



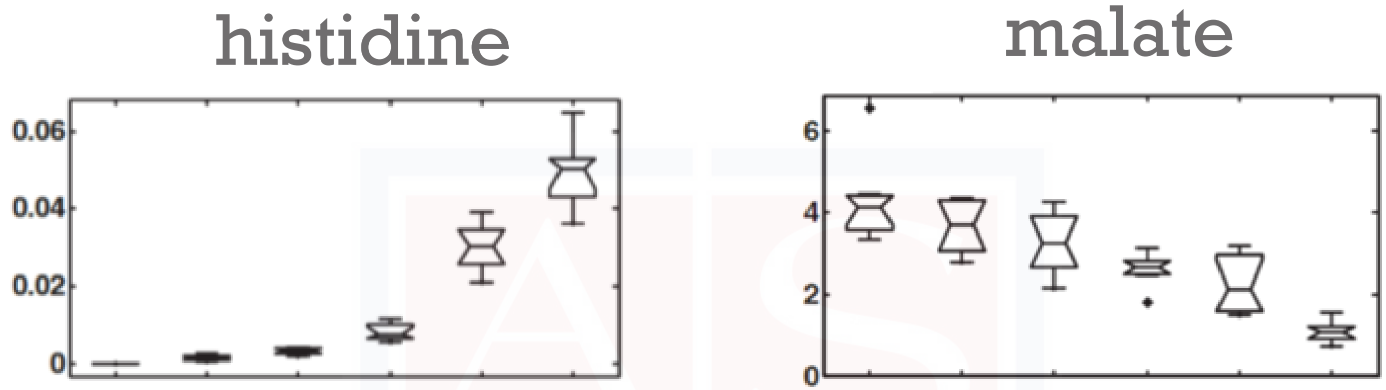
The Seahorse XF Glycolysis Stress Test Profile illustrates the three key parameters of glycolytic function: glycolysis, glycolytic capacity, and glycolytic reserve.

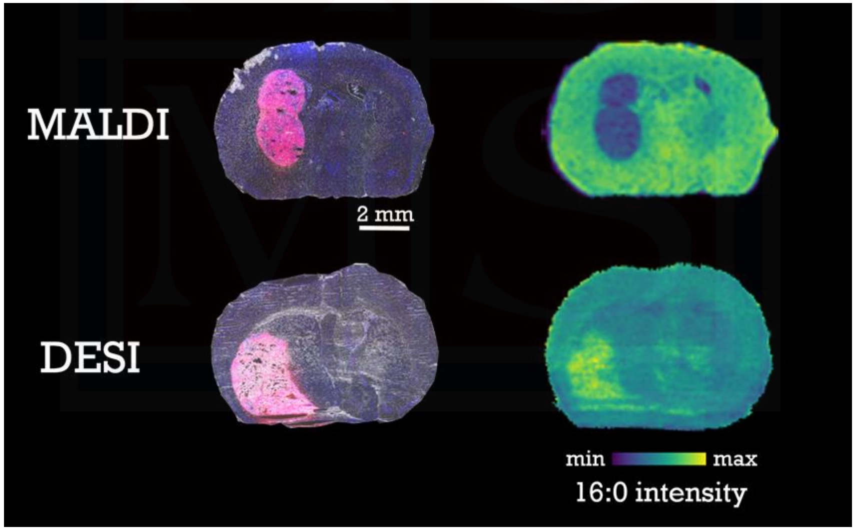


The Seahorse XF Cell Mito Stress Test Profile illustrates the key parameters of mitochondrial function: basal respiration, ATP production, proton leak, maximal respiration, and spare respiratory capacity.

Spatial analysis of metabolic flux

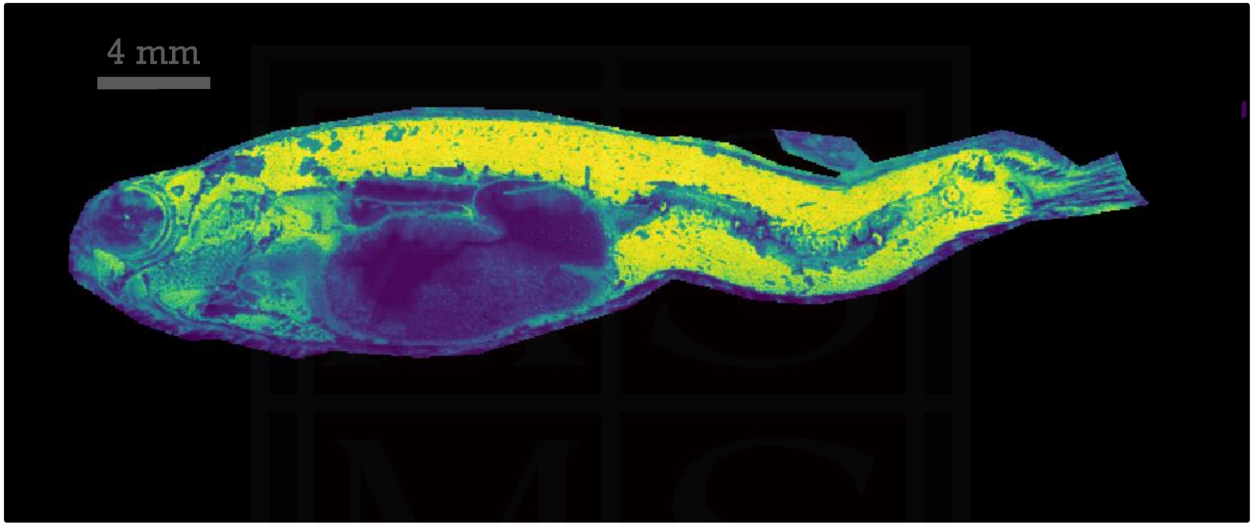
Spatial analysis of metabolic flux





REMINDER: COMPLICATIONS OF MATRIX EFFECTS

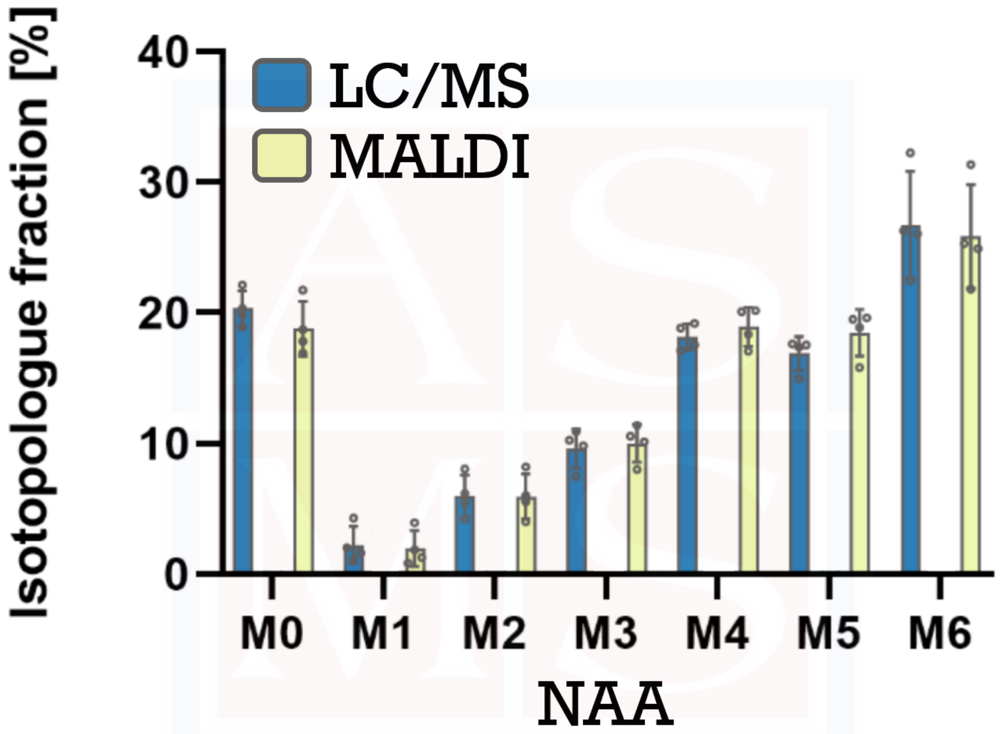
Spatial analysis of metabolic flux



diff. organs have diff. matrix

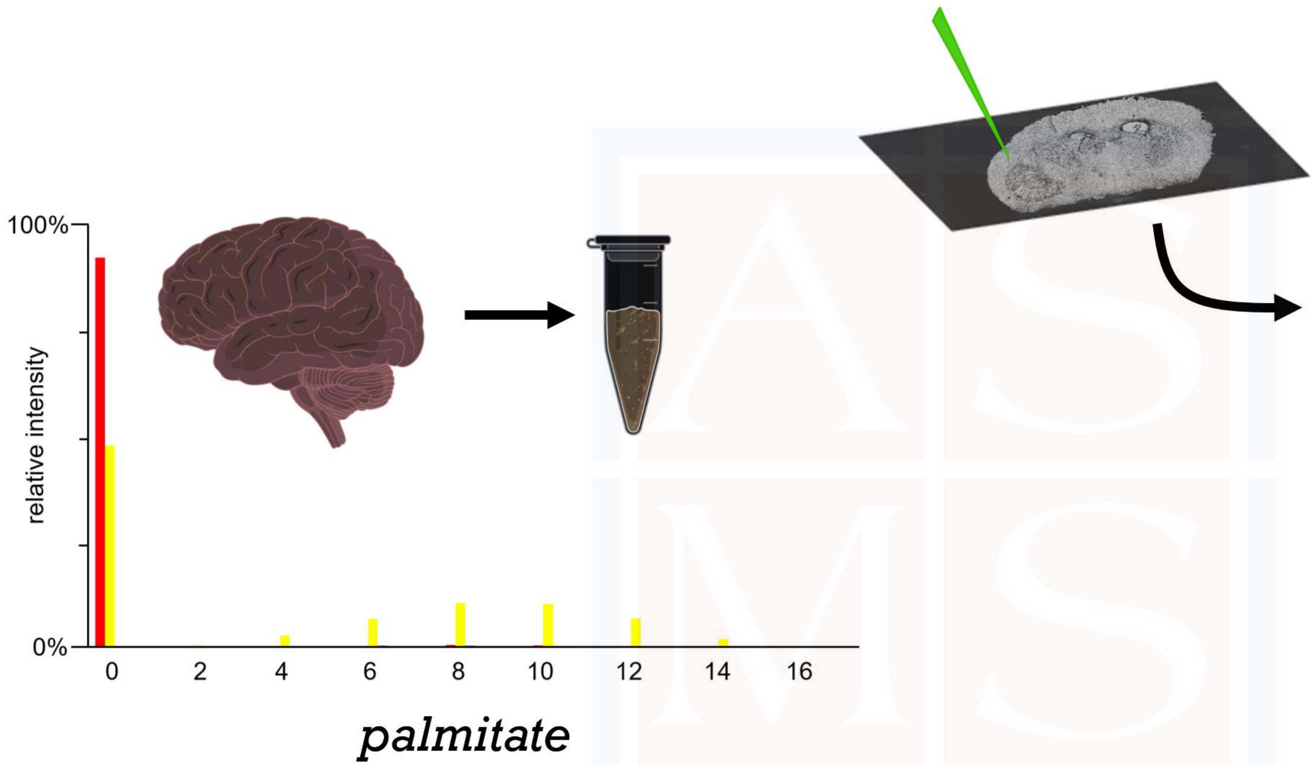
REMINDER: COMPLICATIONS OF MATRIX EFFECTS

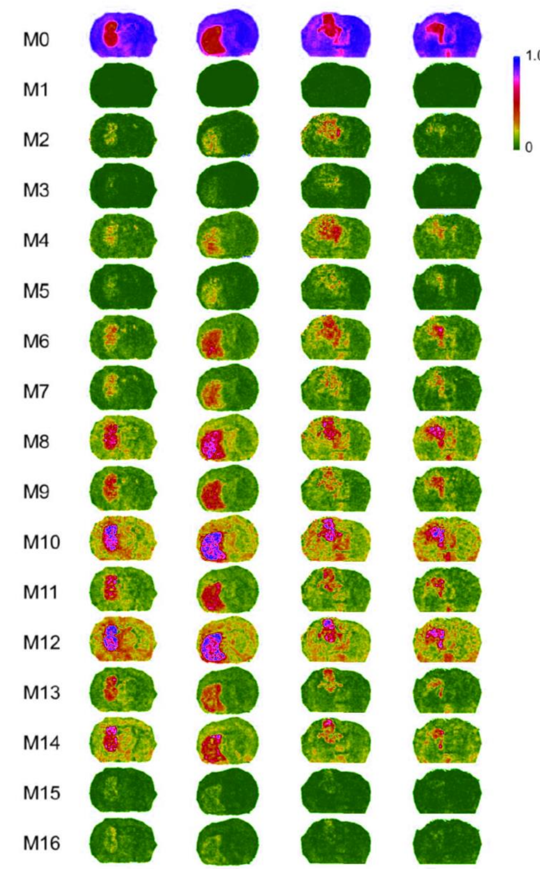
Spatial analysis of metabolic flux



A possible solution to matrix effects during imaging is to focus on labeling form tracers

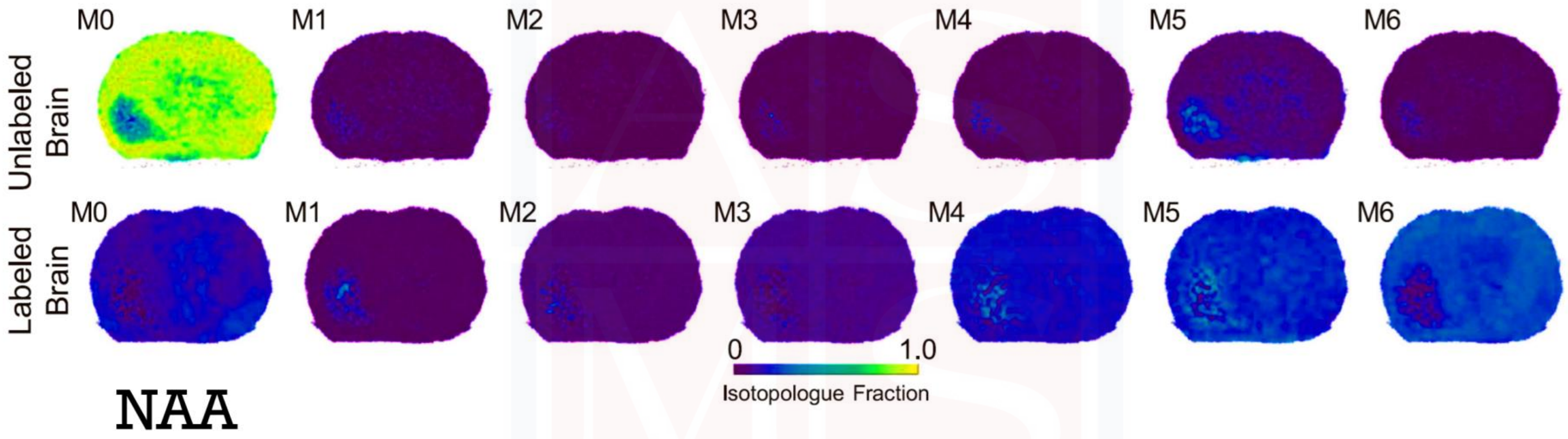
Spatial analysis of metabolic flux





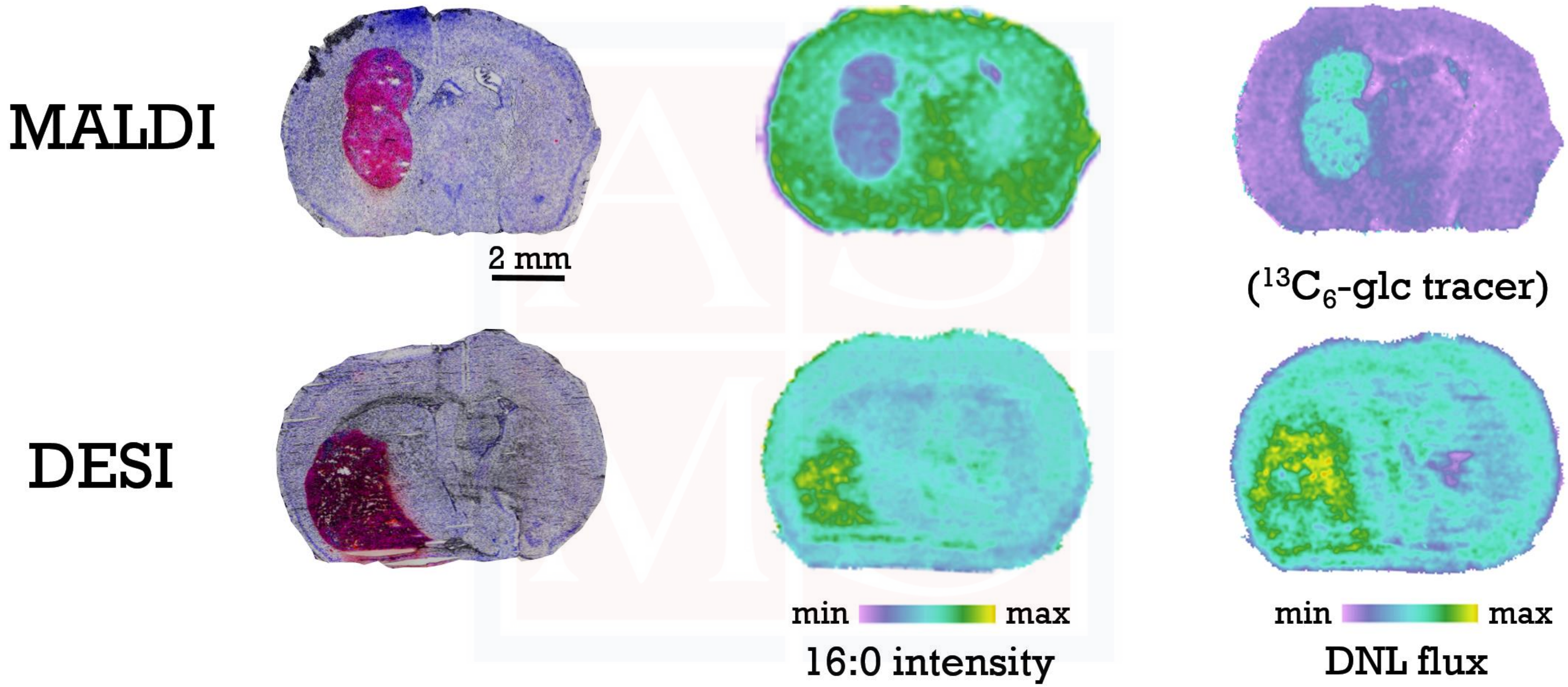
Schwaiger-Haber et al, Nat Comm 2023

Spatial analysis of metabolic flux



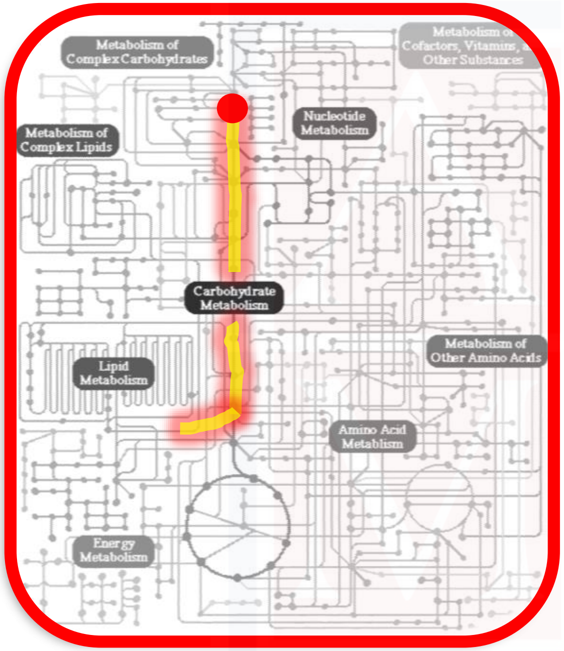
Main limitation: isotope interferences

Spatial analysis of metabolic flux

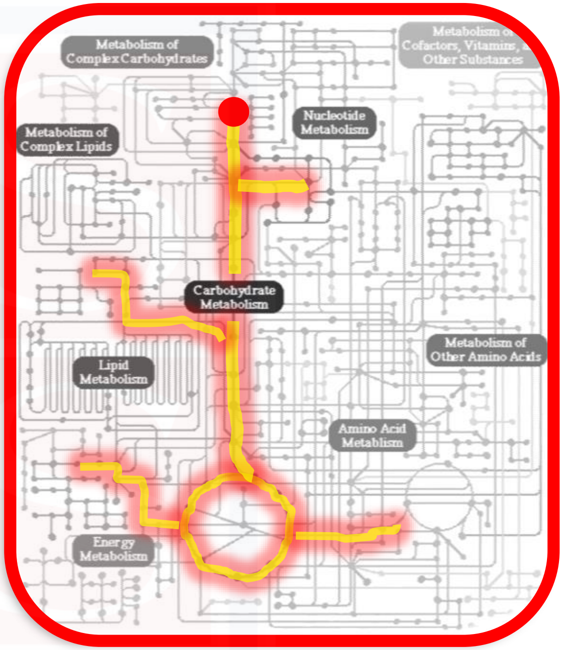


Schwaiger-Haber et al, Nat Comm 2023

Global tracking of isotope tracers with untargeted metabolomics



hypoxia



normoxia

Llufrio et al, Nature Protocols 2019

Global tracking of isotope tracers with untargeted metabolomics

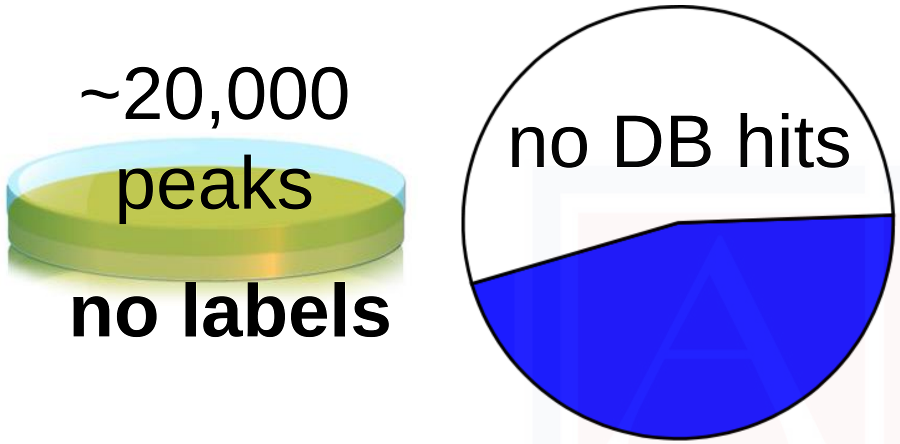
Appropriate for:

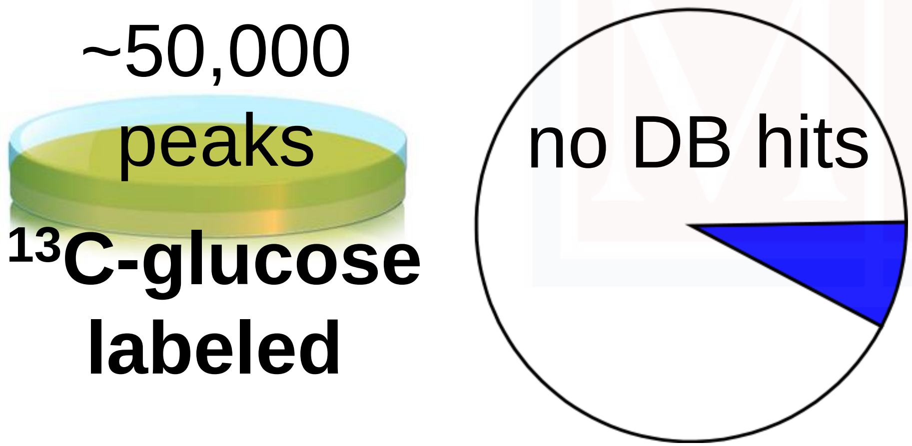
- Analyses when do not know tracer fate (e.g., drugs, unknown metabolites, etc.)
- Finding unexpected differences in tracer fates between multiple samples

Inappropriate for:

- Formal metabolic flux analysis
- Targeted analysis of specific pathways

Global analysis of isotopes: the challenge



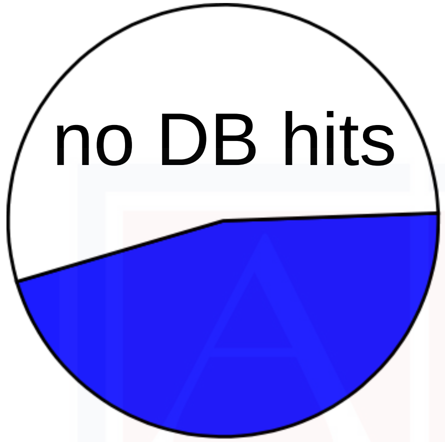


Global analysis of isotopes: the challenge

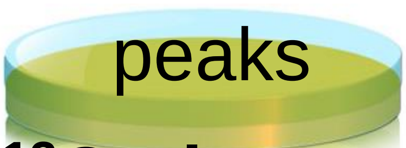
~20,000

peaks

no labels

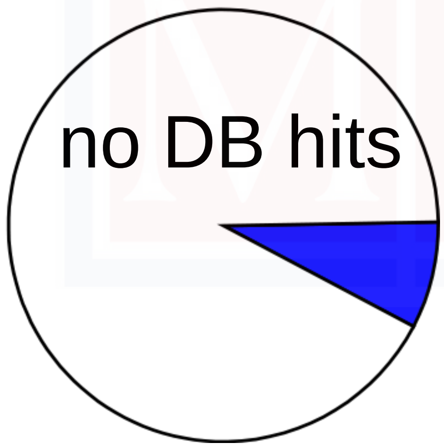


~50,000



peaks

^{13}C-glucose labeled



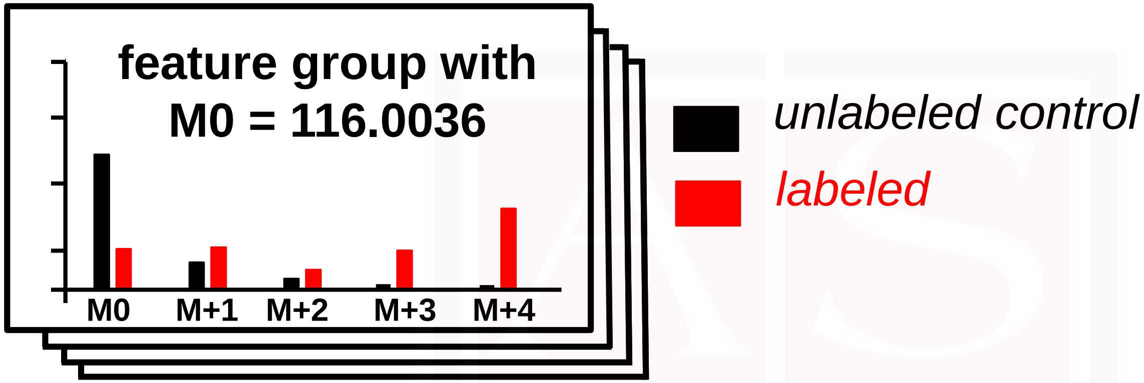
software to “collapse”

1. X^{13}CMS (R)
2. NTFD (C++)
3. geoRge (R)
4. MetExtract (C++)
5. CD 3.0 (GUI, Thermo)

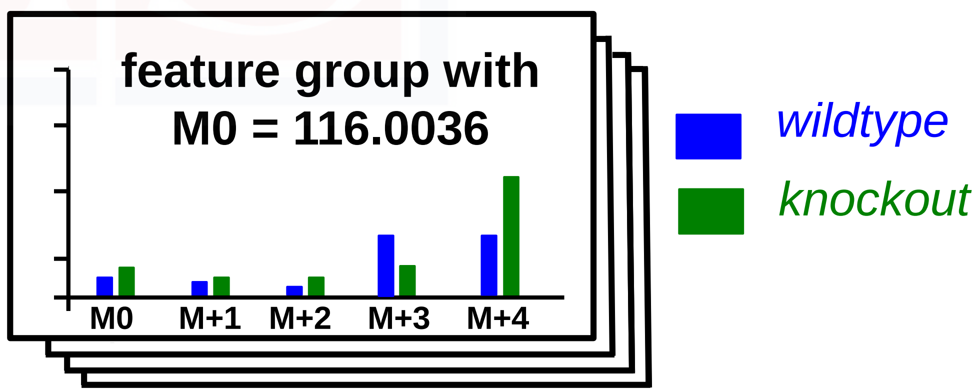
1. Huang et al., Anal Chem 86(3) pgs 1632-1639
2. Hiller et al., Anal Chem 82(15) pgs 6621-6628
3. Capellades et al., Anal Chem 88(1) pgs 621-628
4. Bueschl et al., Bioinformatics 24 (5) pgs 736-738

Global analysis of isotopes: objectives

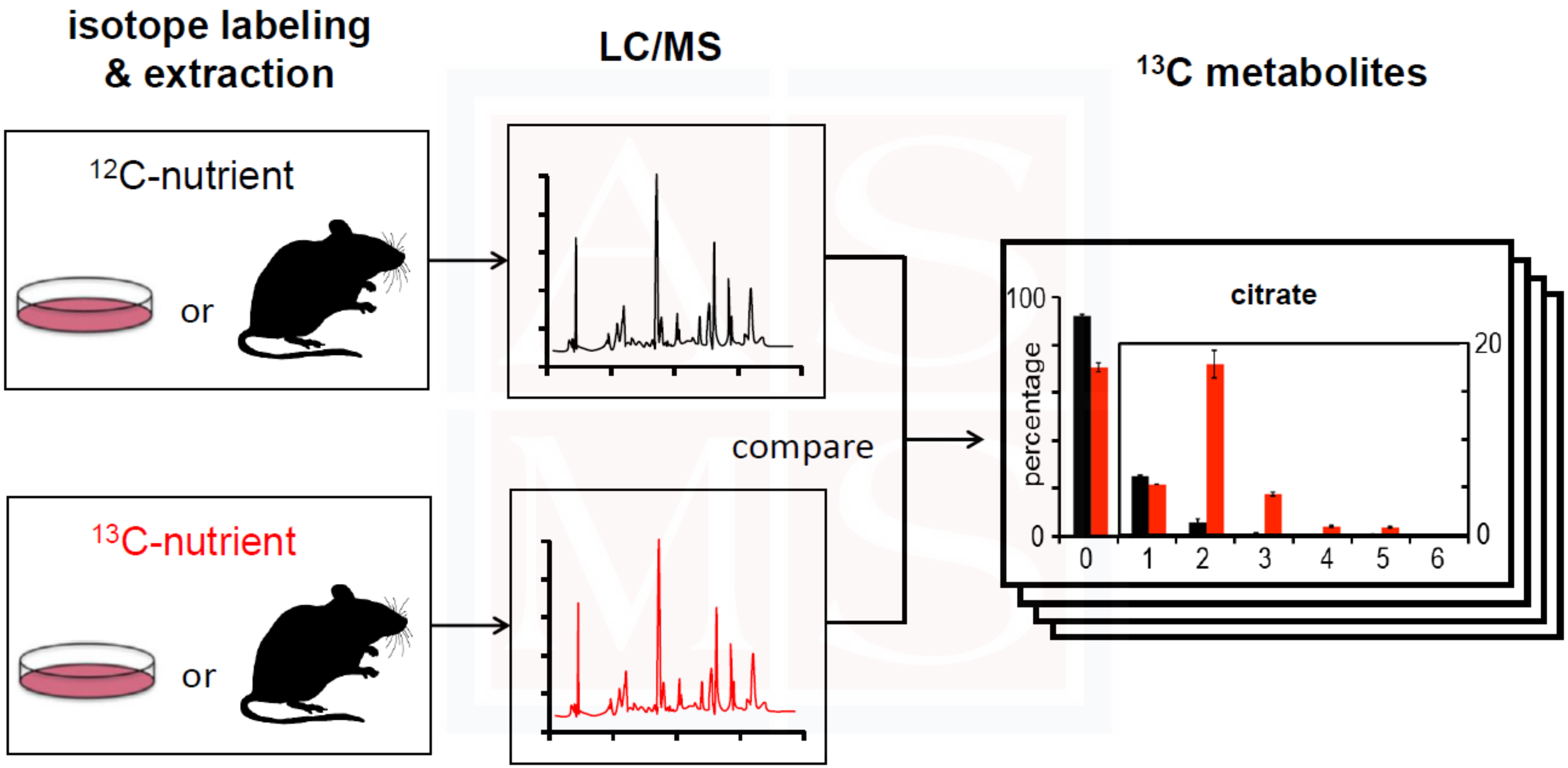
Where does label go?



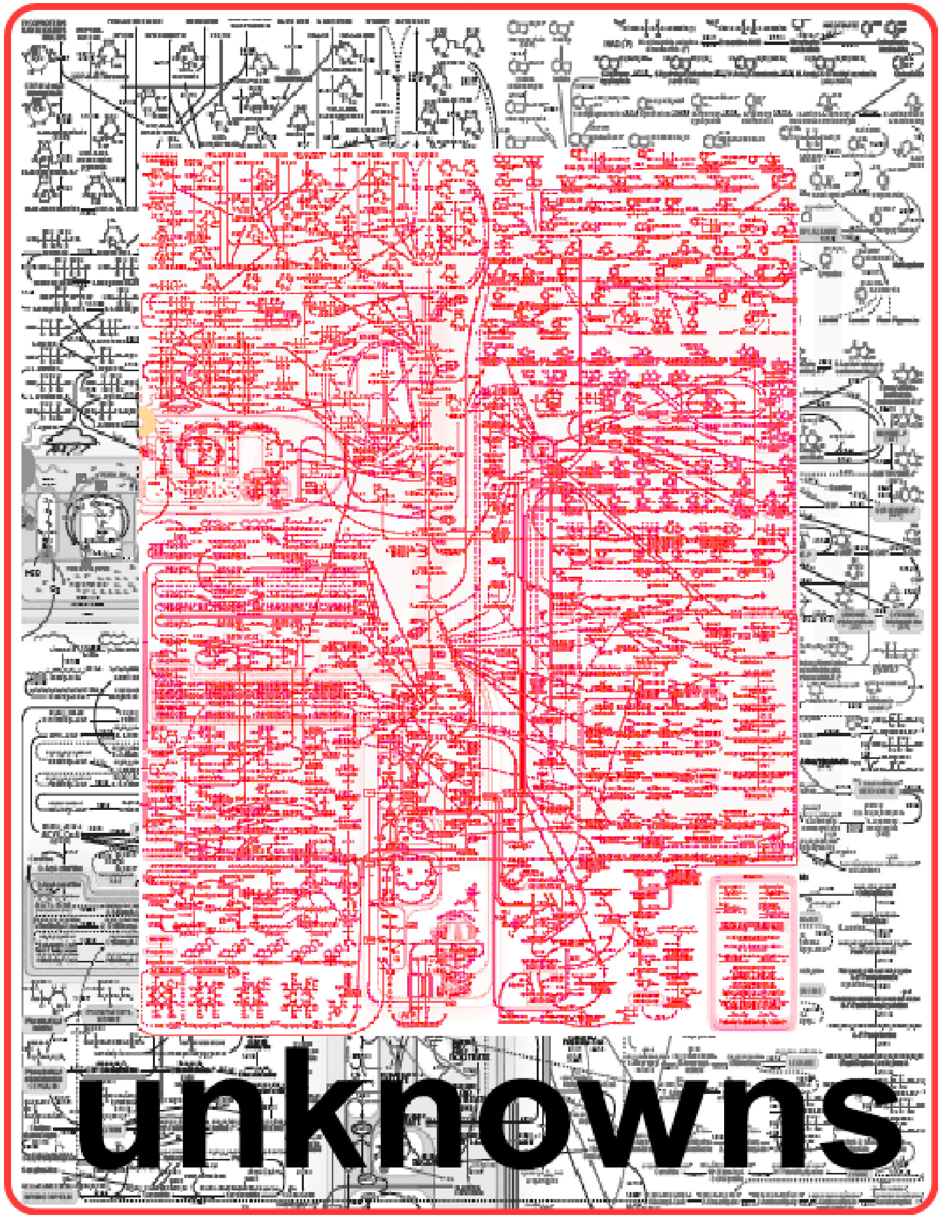
Does fate of label change w/stress?



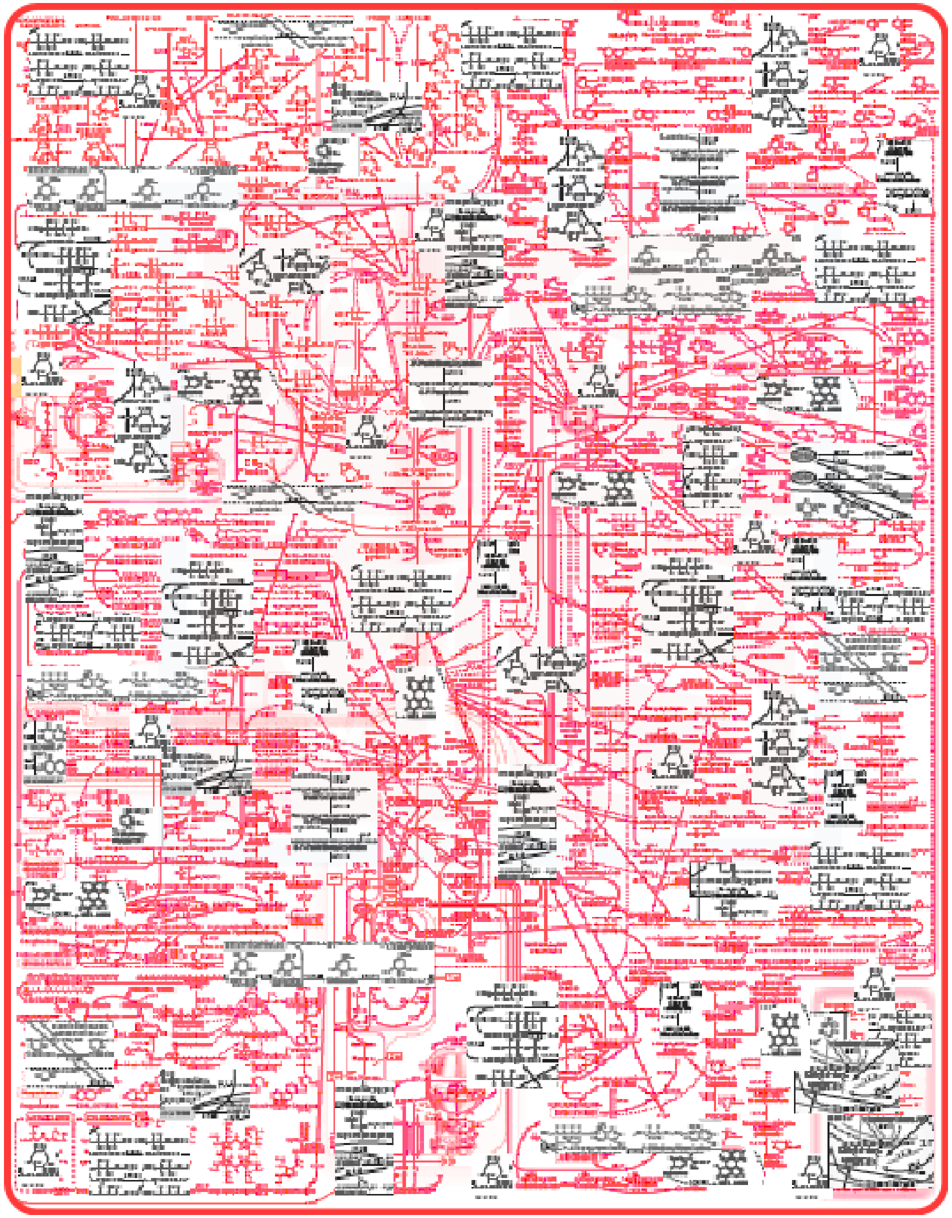
Global analysis of isotopes: exp design



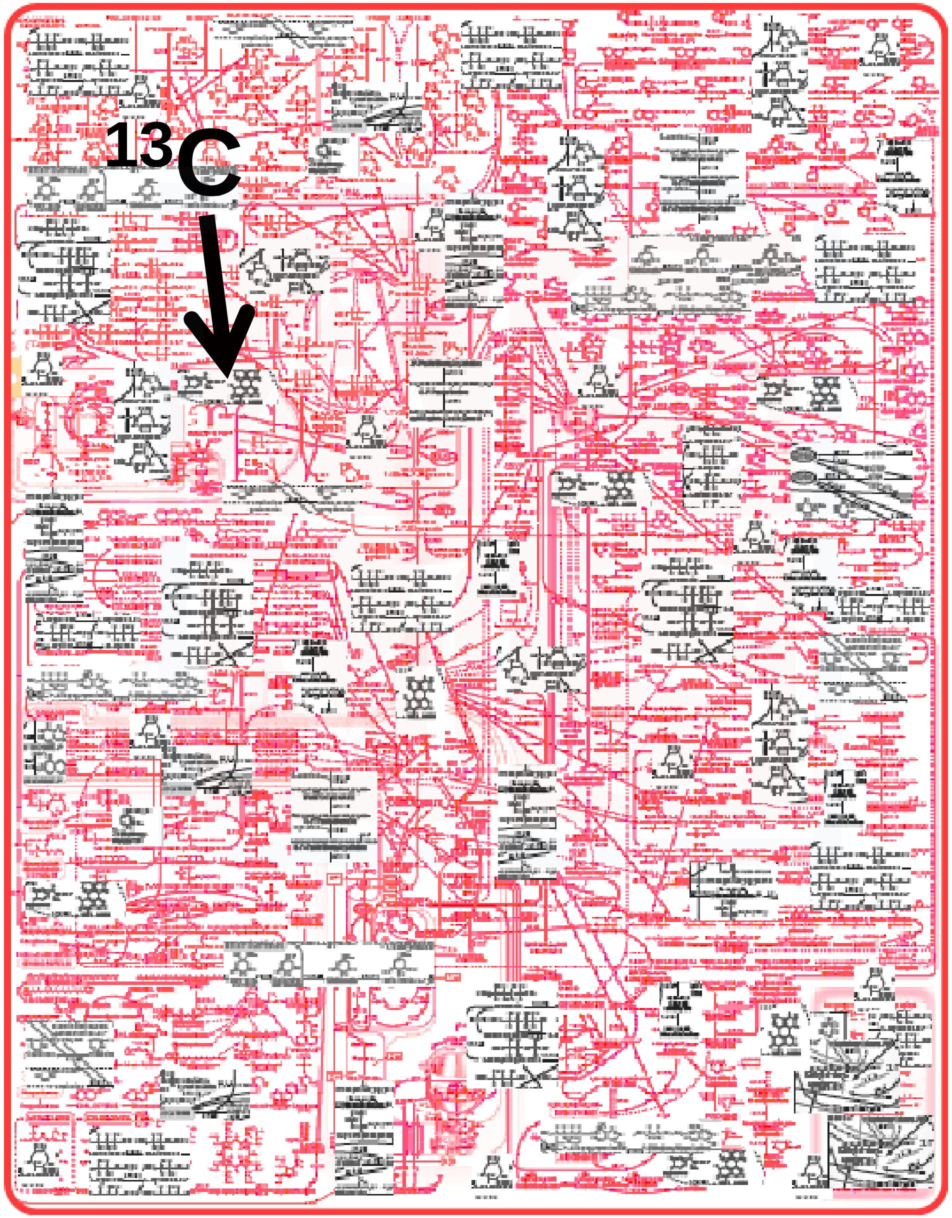
The “dark matter” of the metabolome



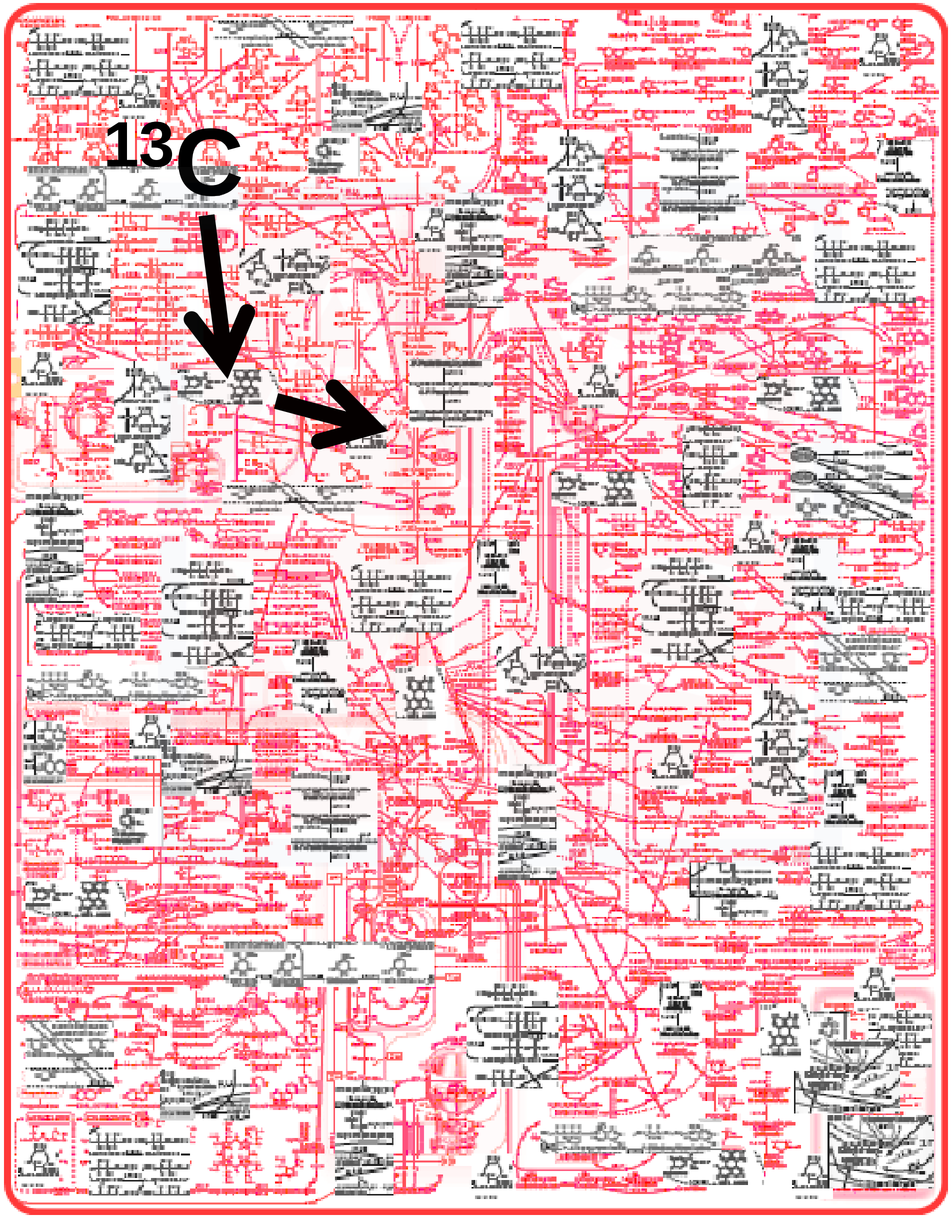
The “dark matter” of the metabolome



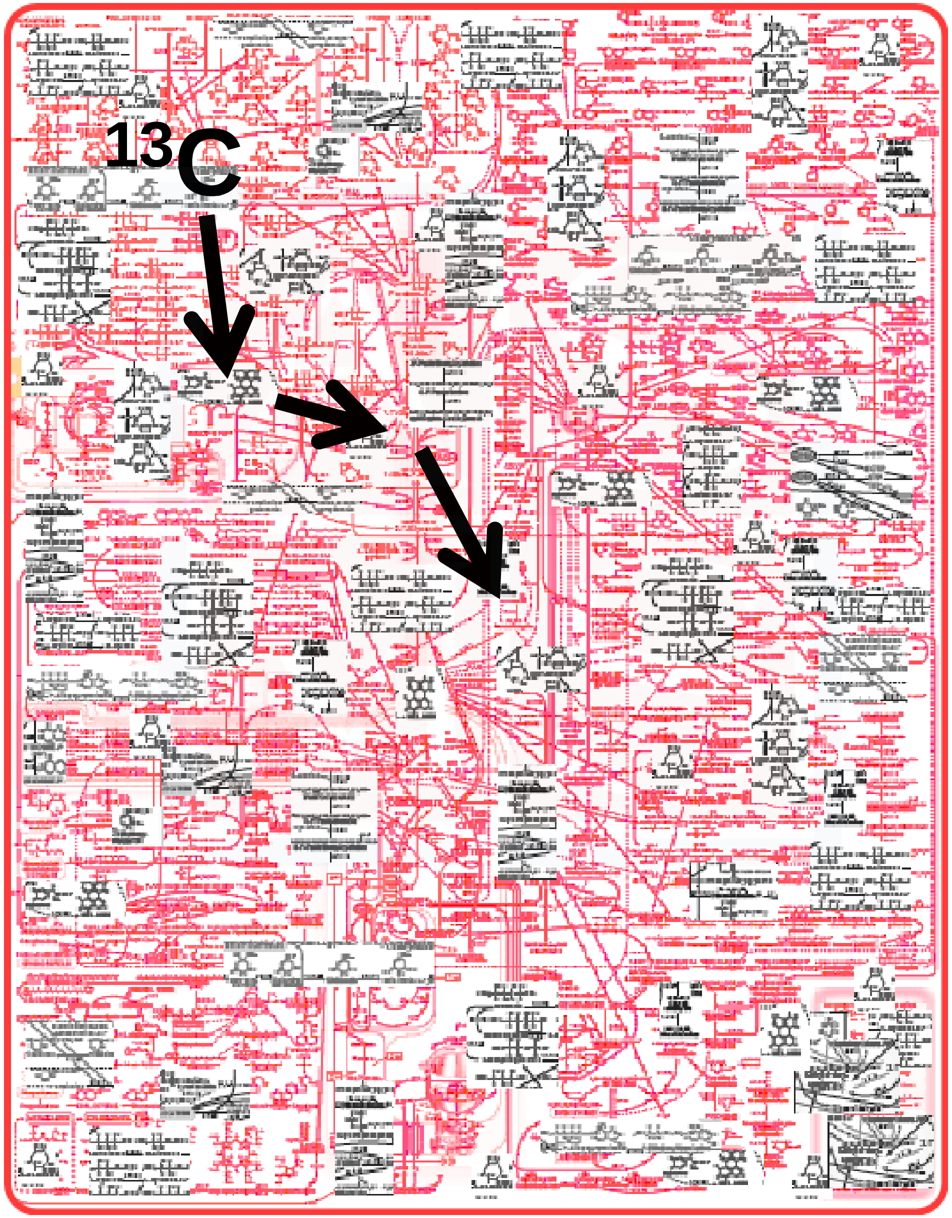
The “dark matter” of the metabolome



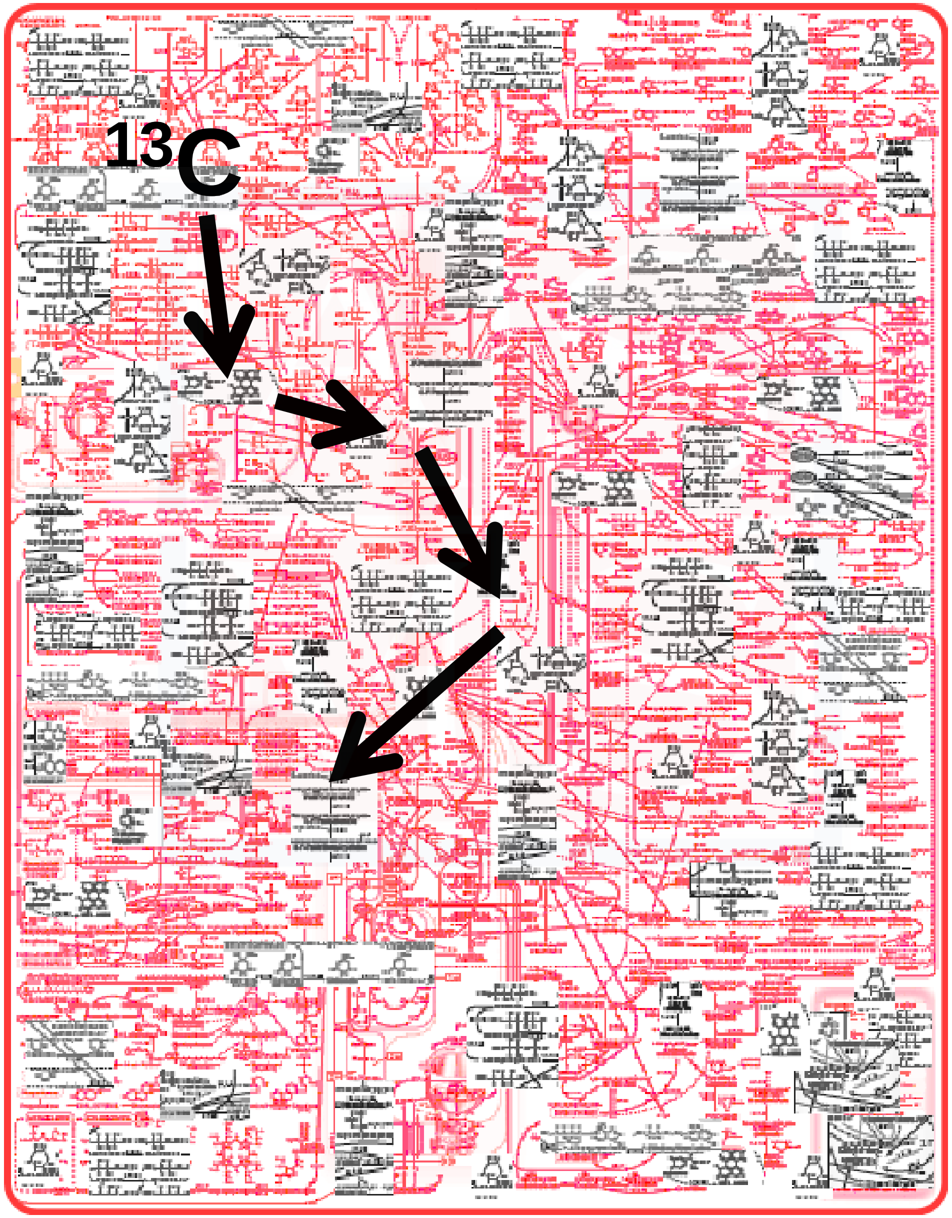
The "dark matter" of the metabolome



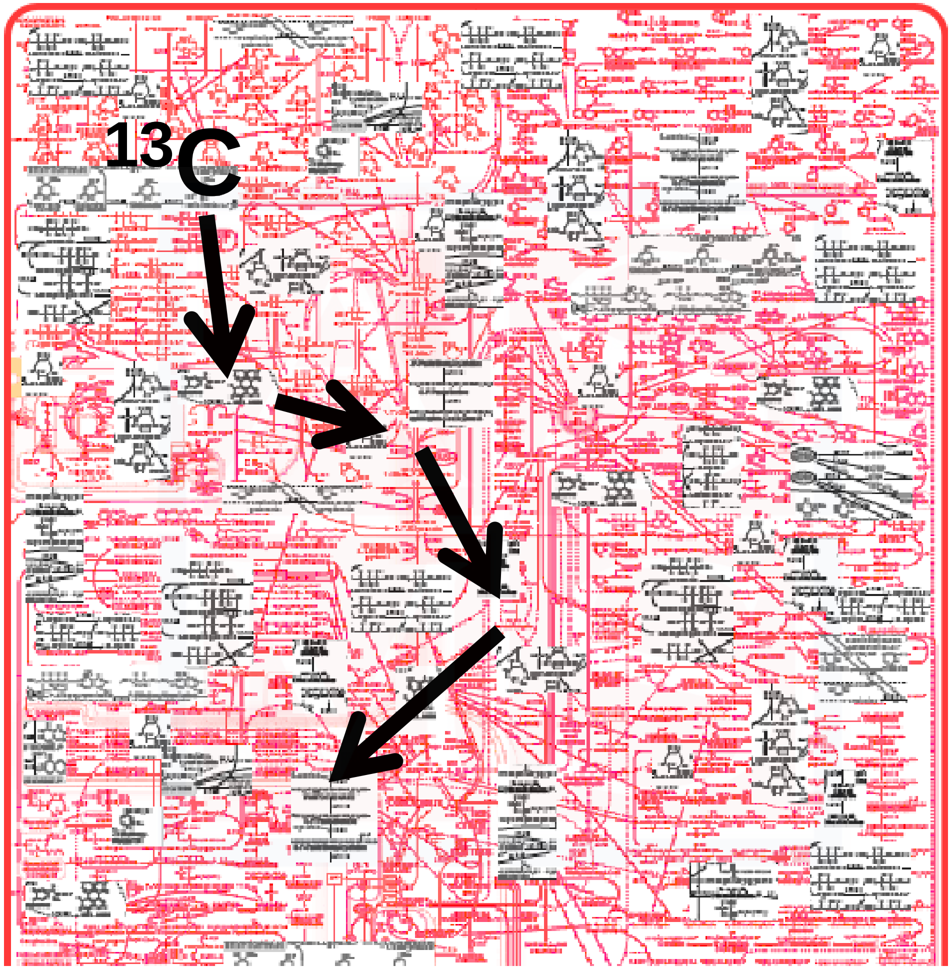
The “dark matter” of the metabolome



The “dark matter” of the metabolome

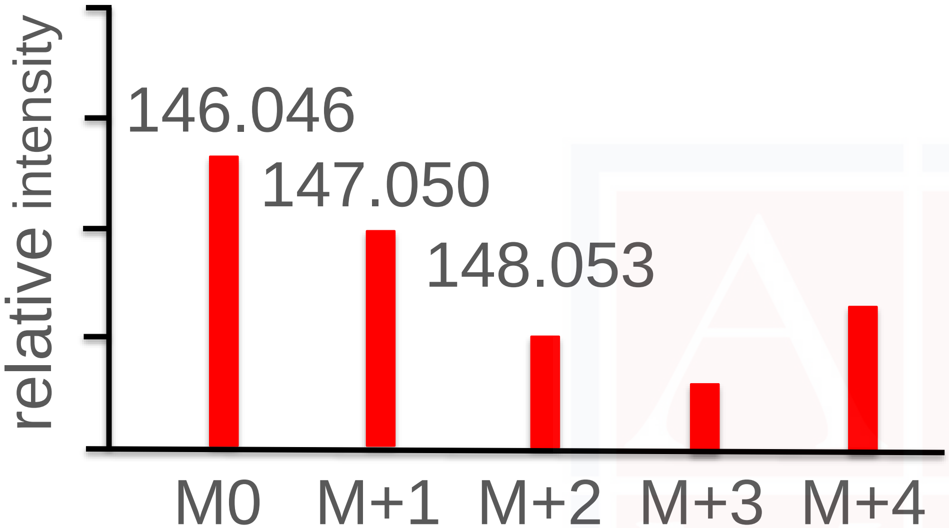


The "dark matter" of the metabolome

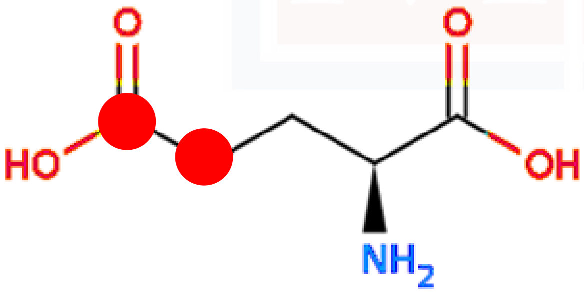


untargeted tracking of specific isotope labels

Global analysis of labeled metabolites

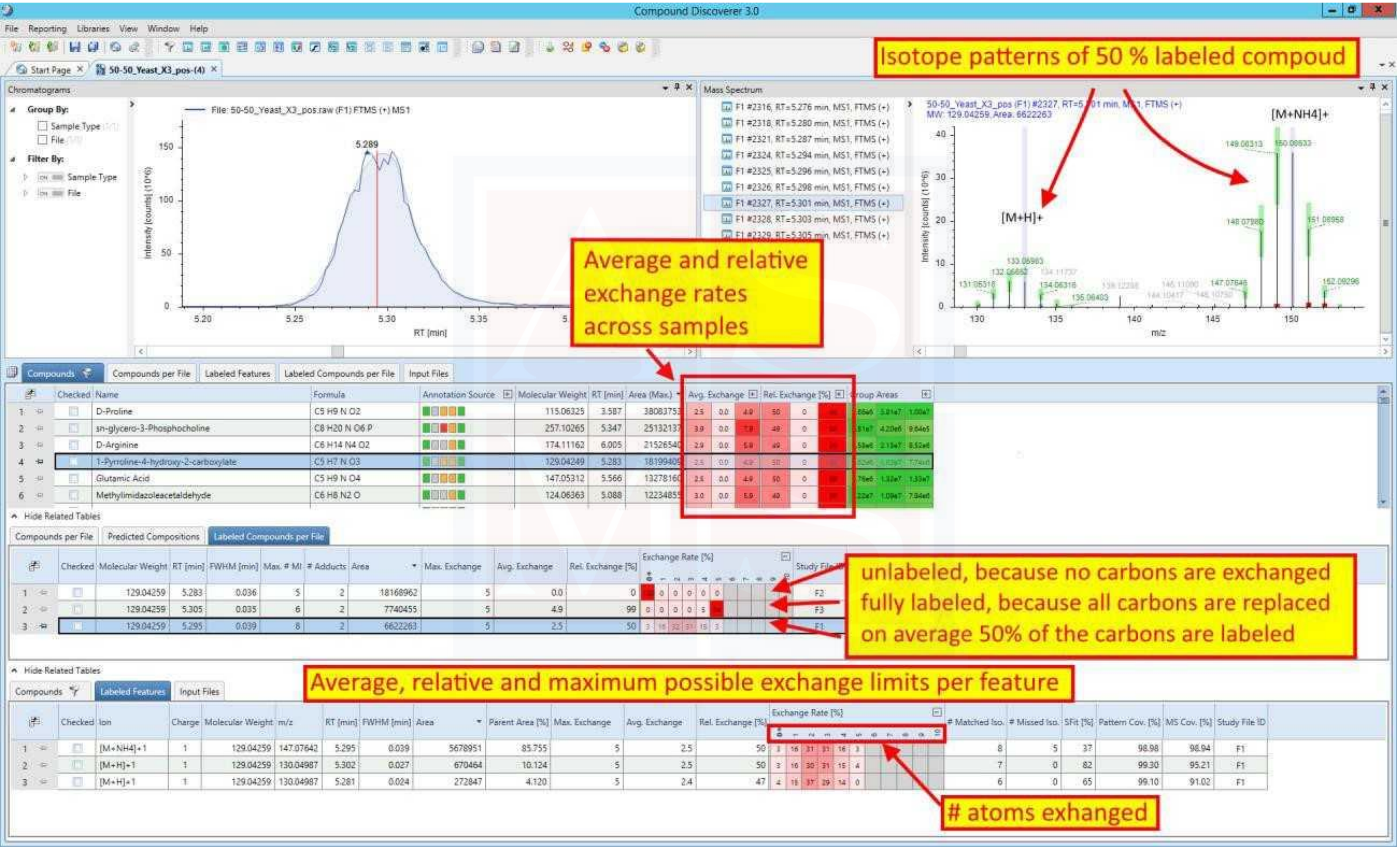


isotopologues (MS^1)

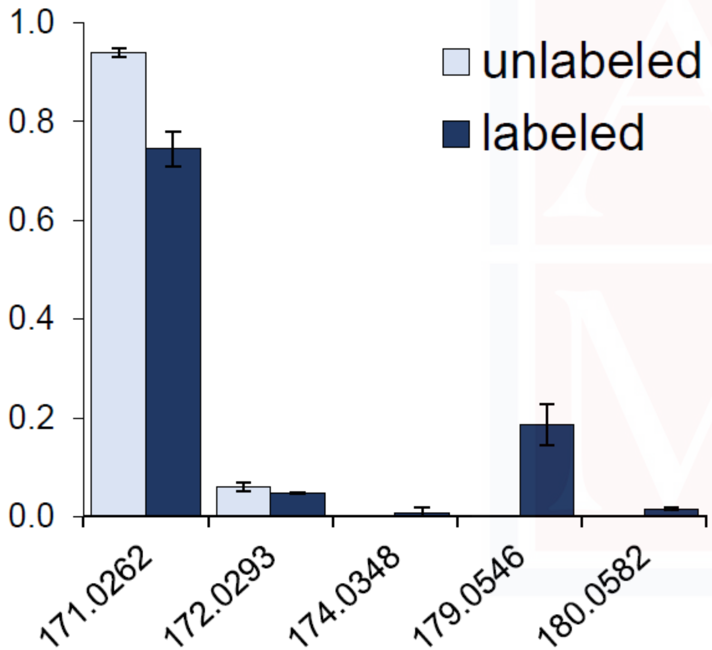


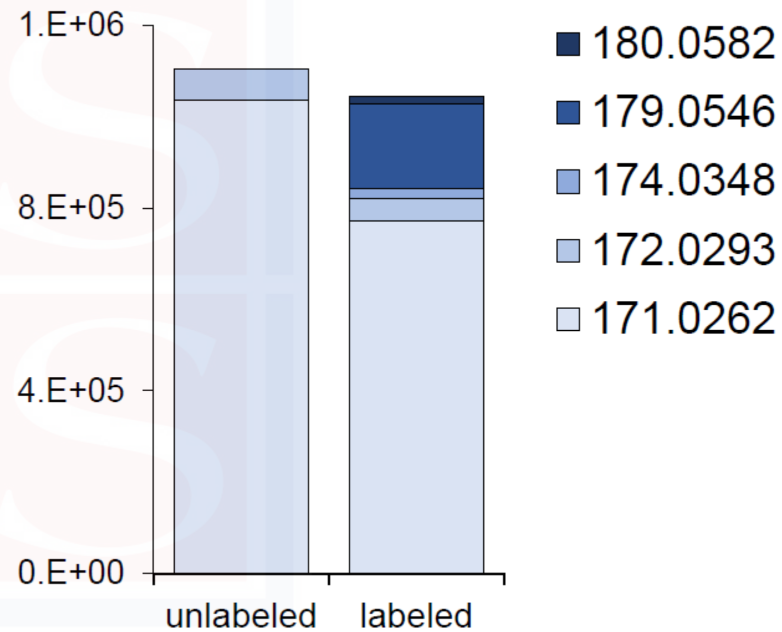
isotopomers (MS^2)

Thermo Compound Discoverer 3.0 output

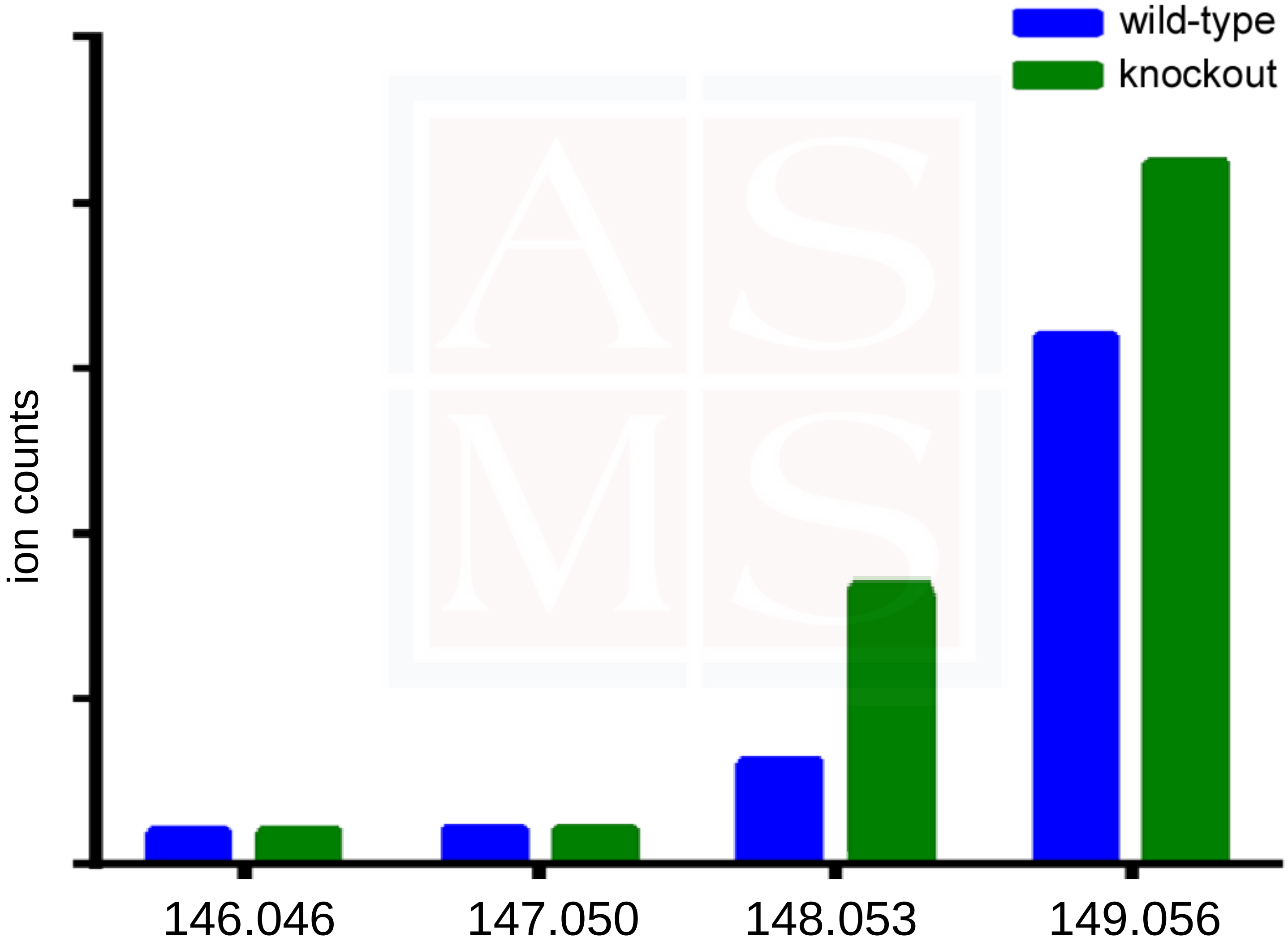


X^{13}CMS output

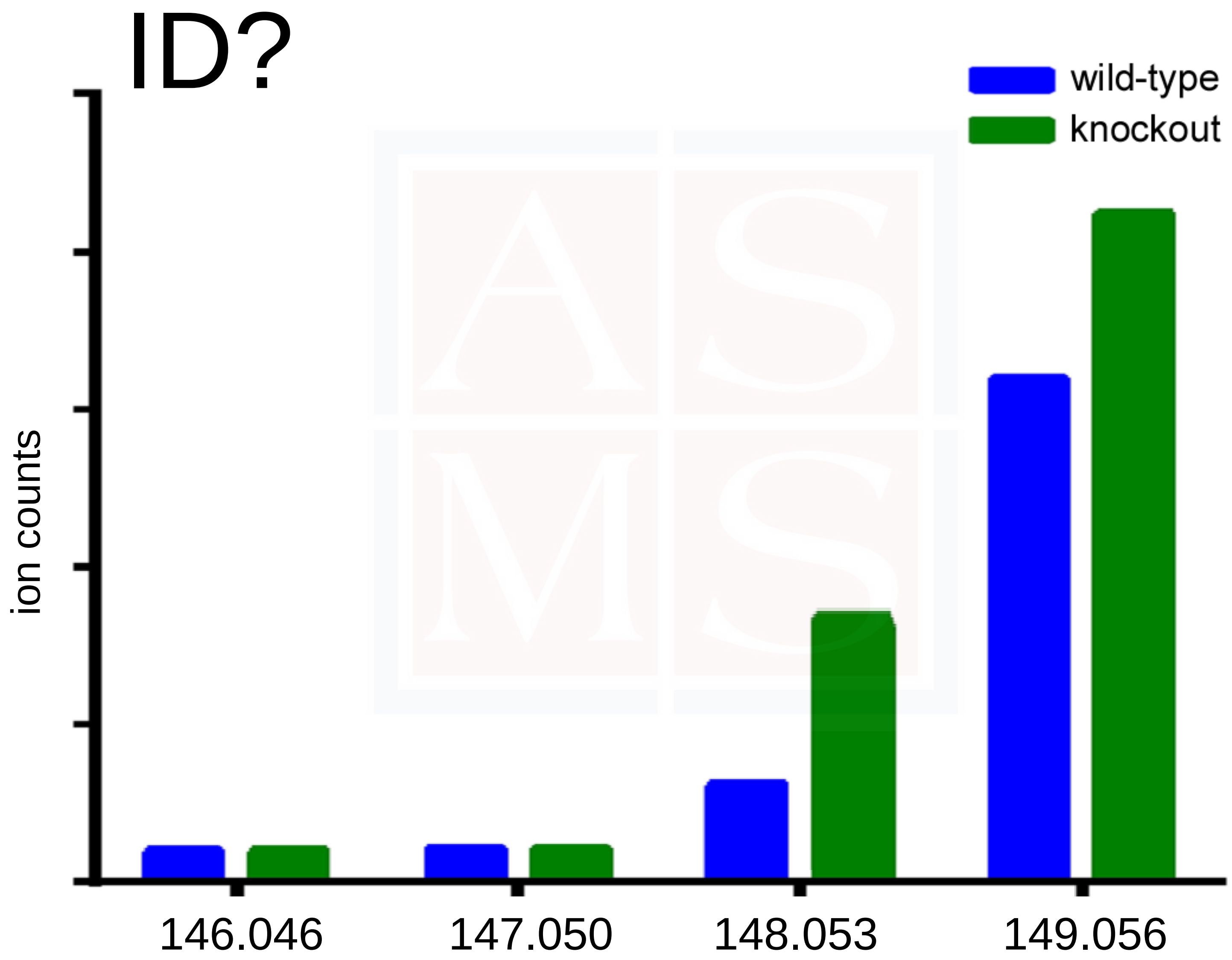




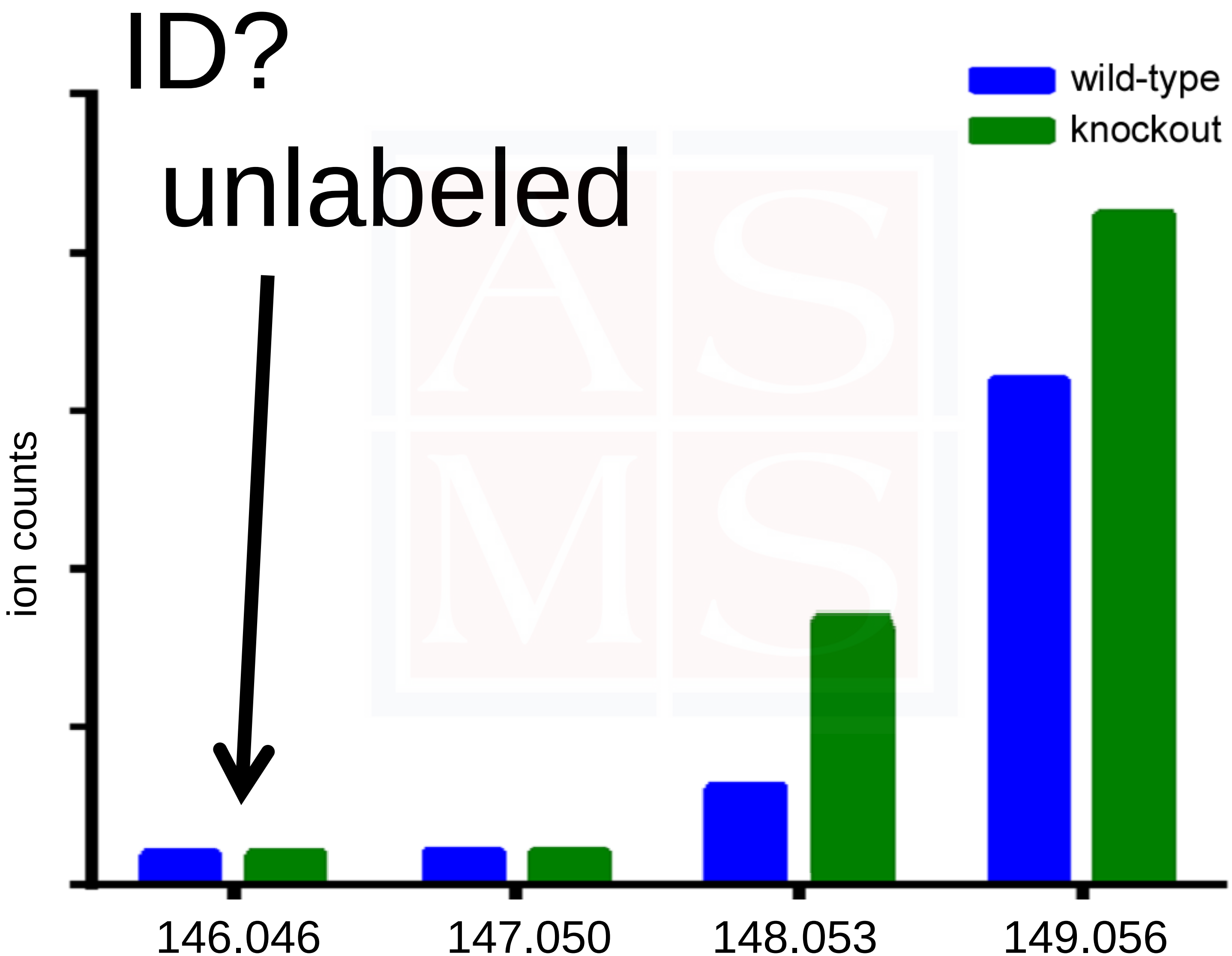
X^{13}CMS output



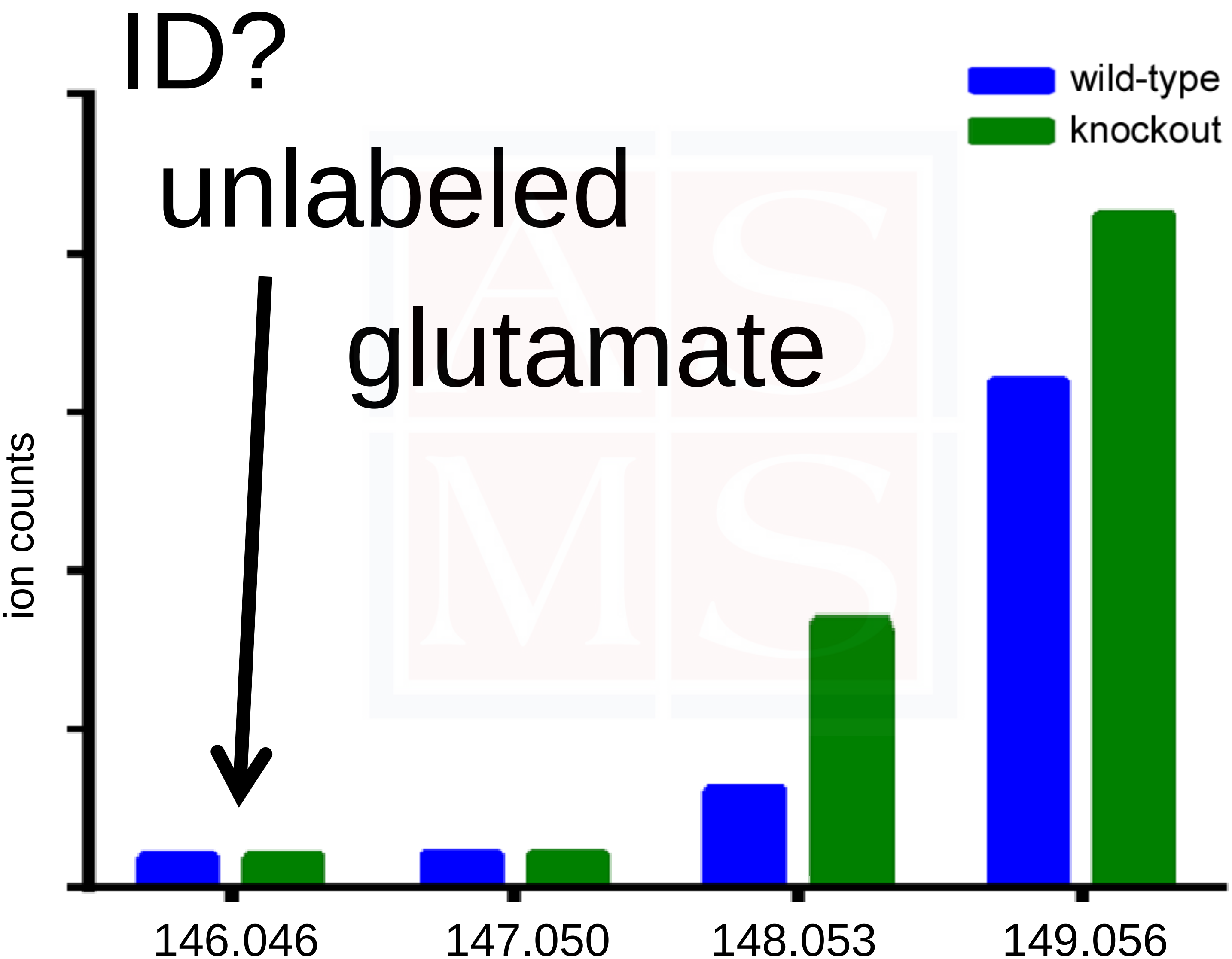
$X^{13}CMS$ output



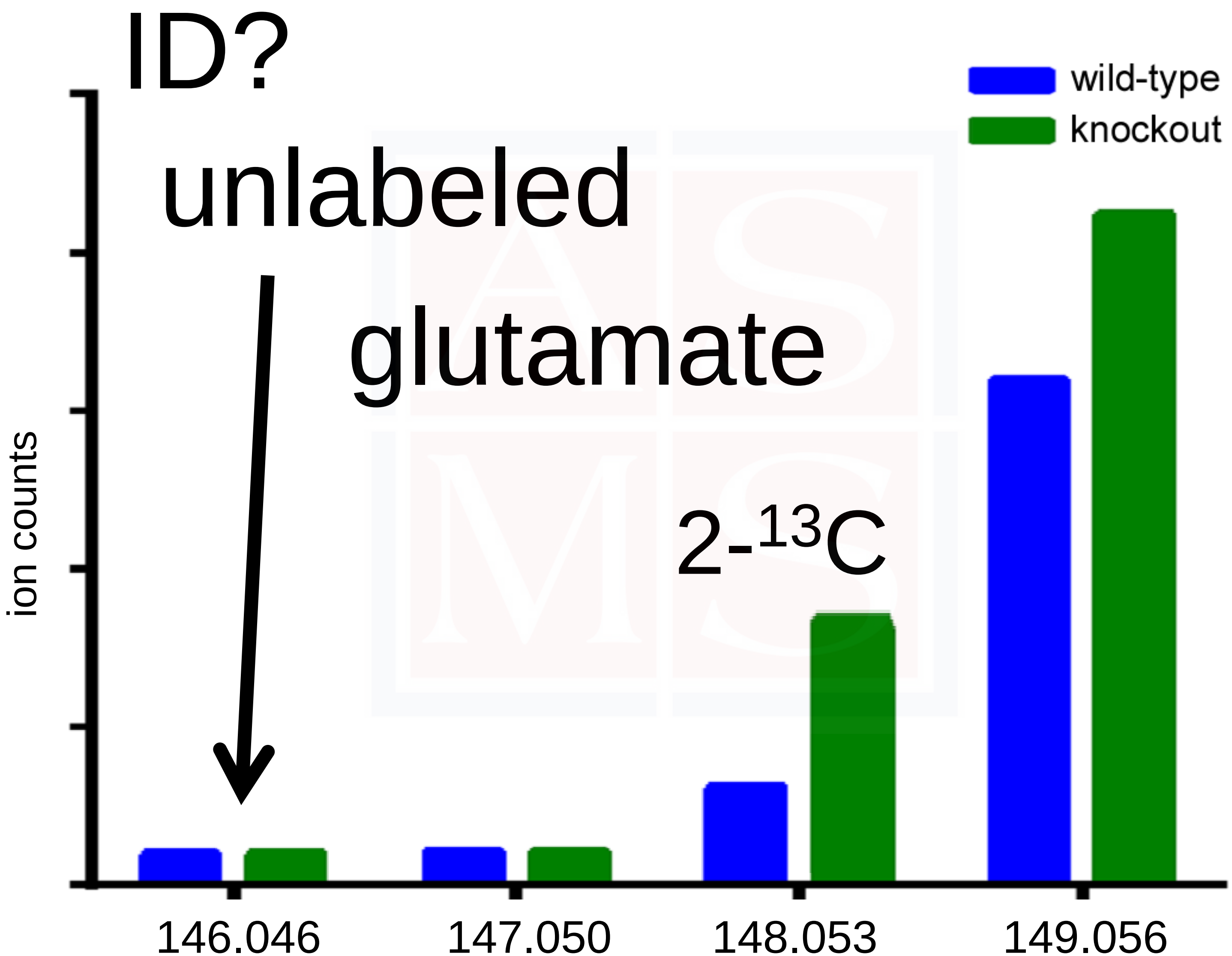
X^{13}CMS output



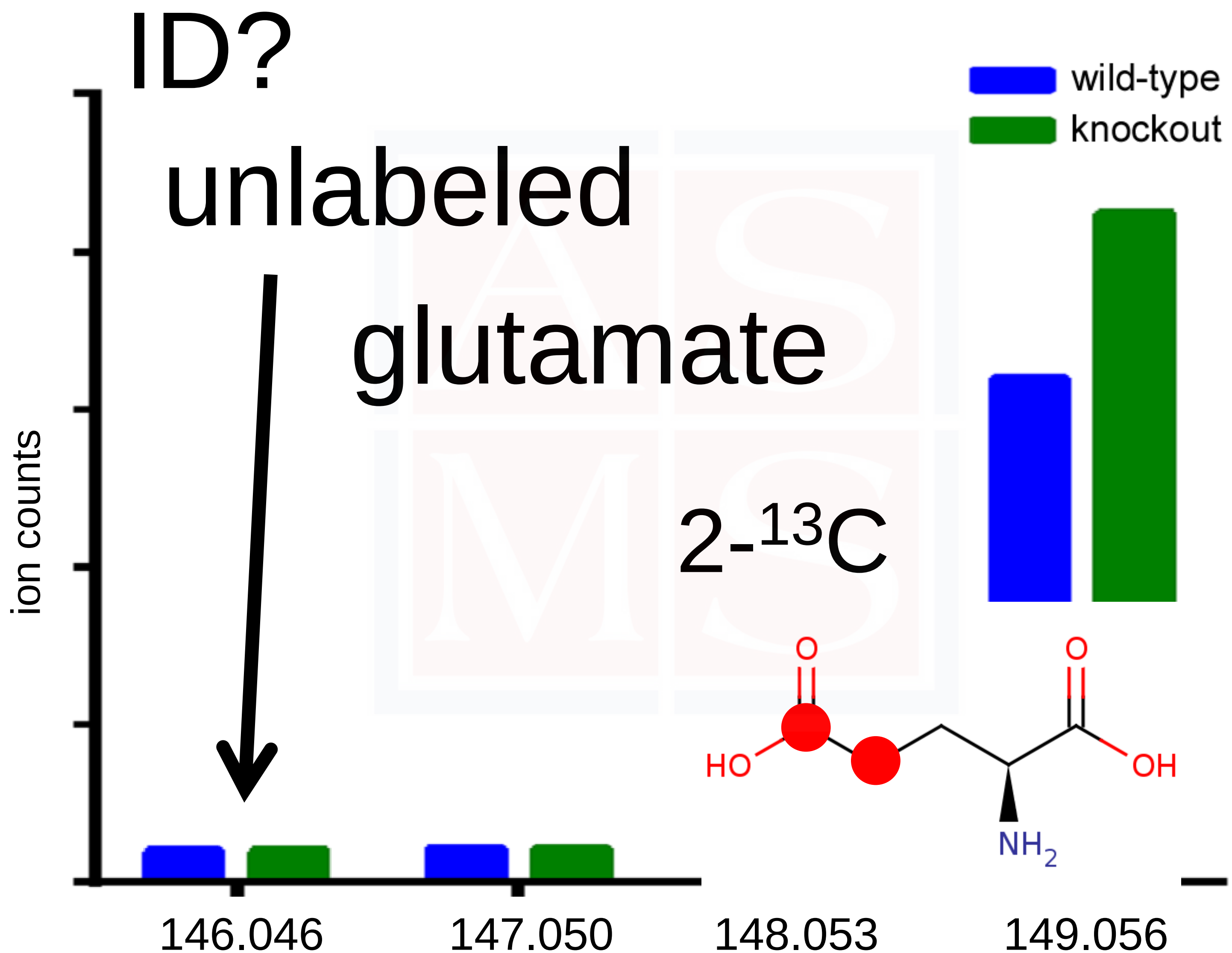
X^{13}CMS output



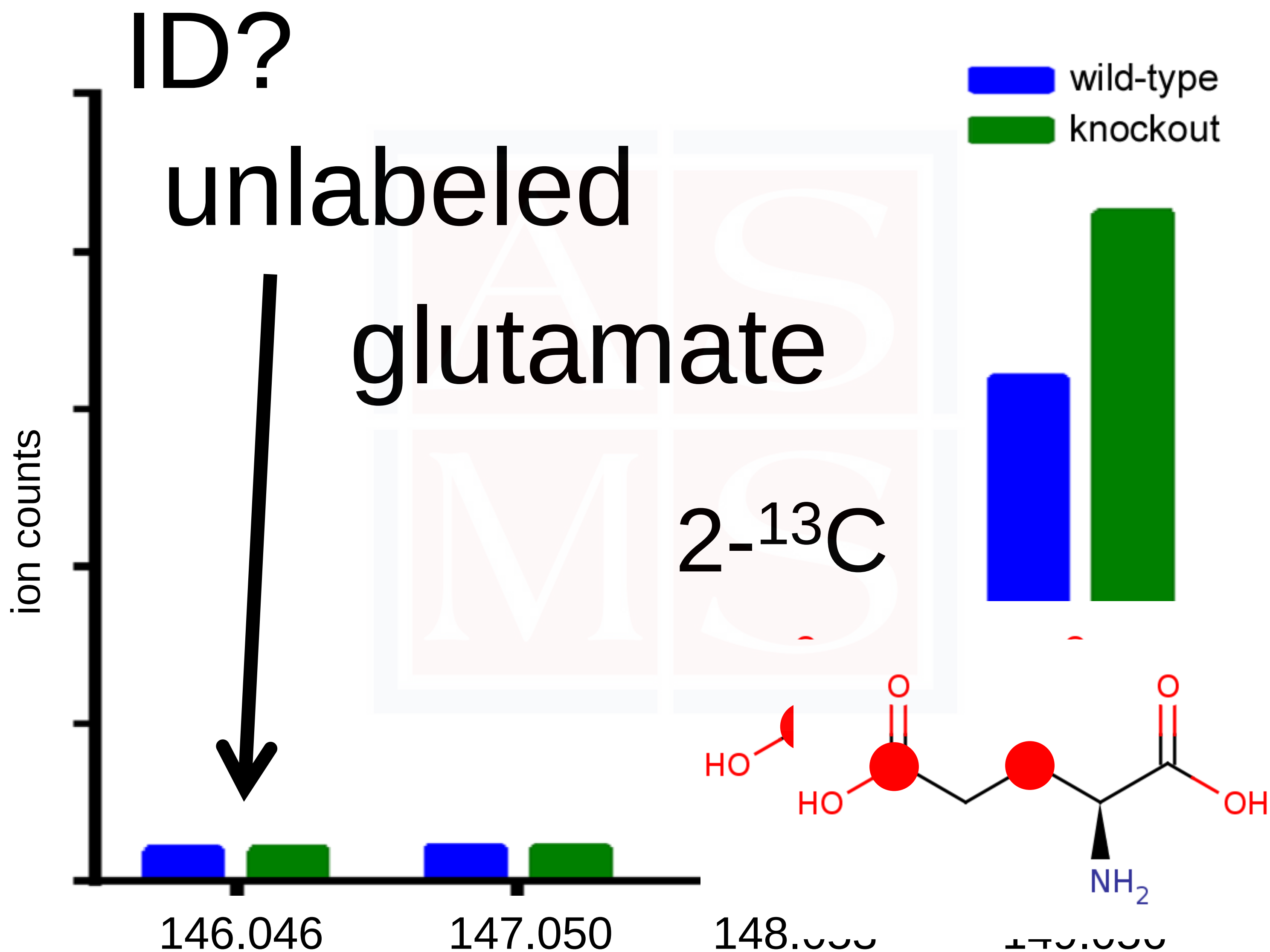
X^{13}CMS output



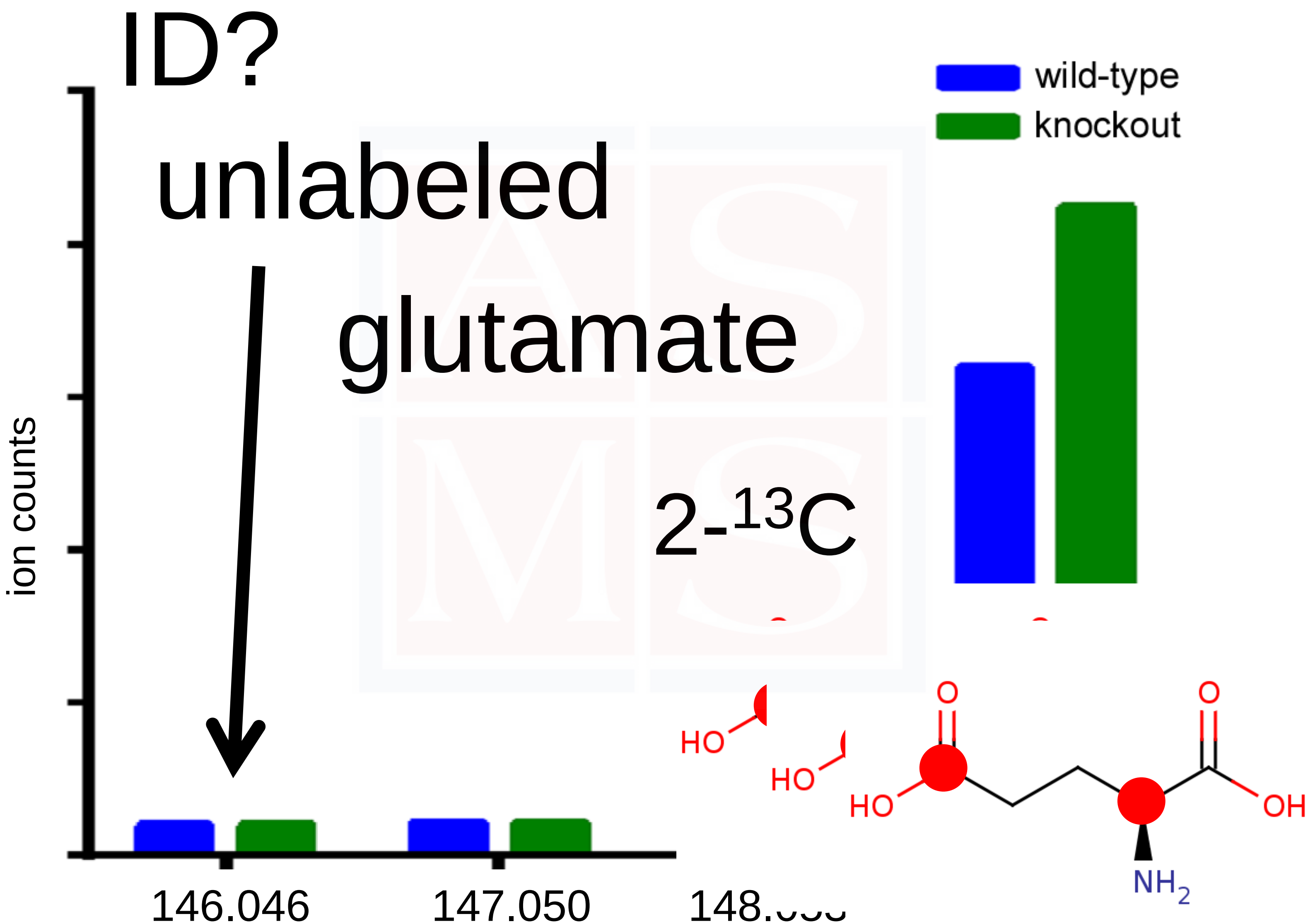
X^{13}CMS output



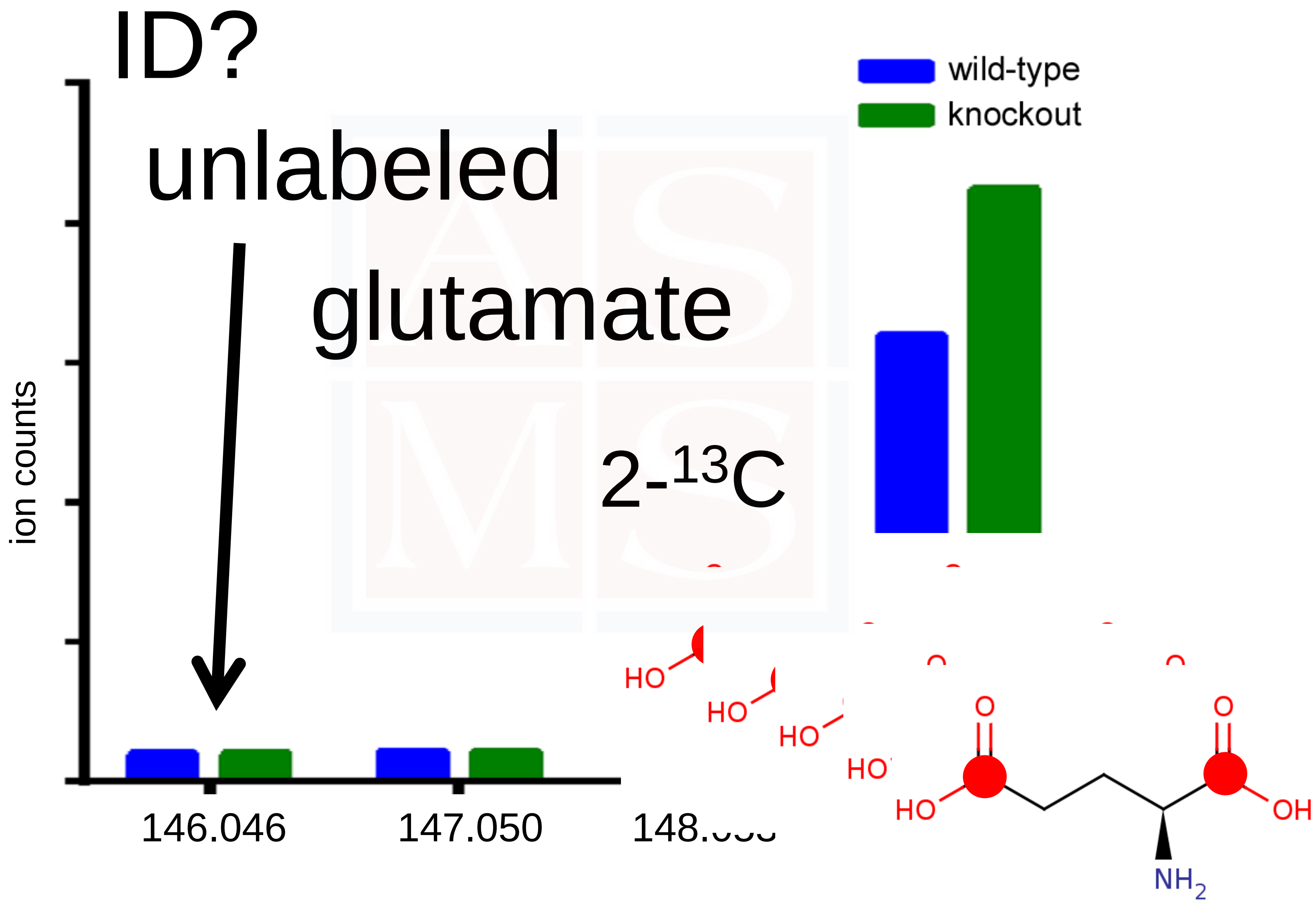
X^{13}CMS output



$X^{13}CMS$ output

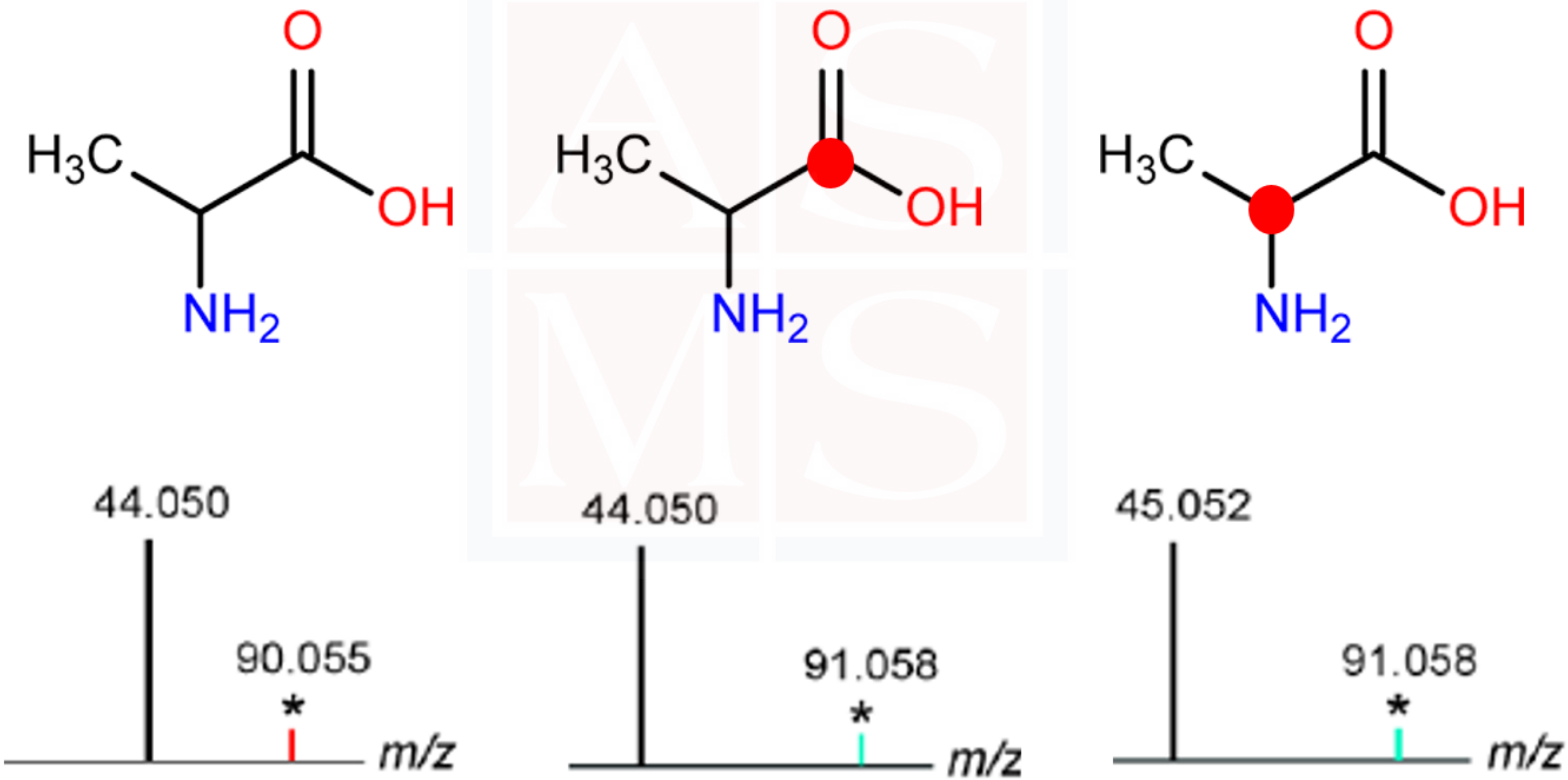


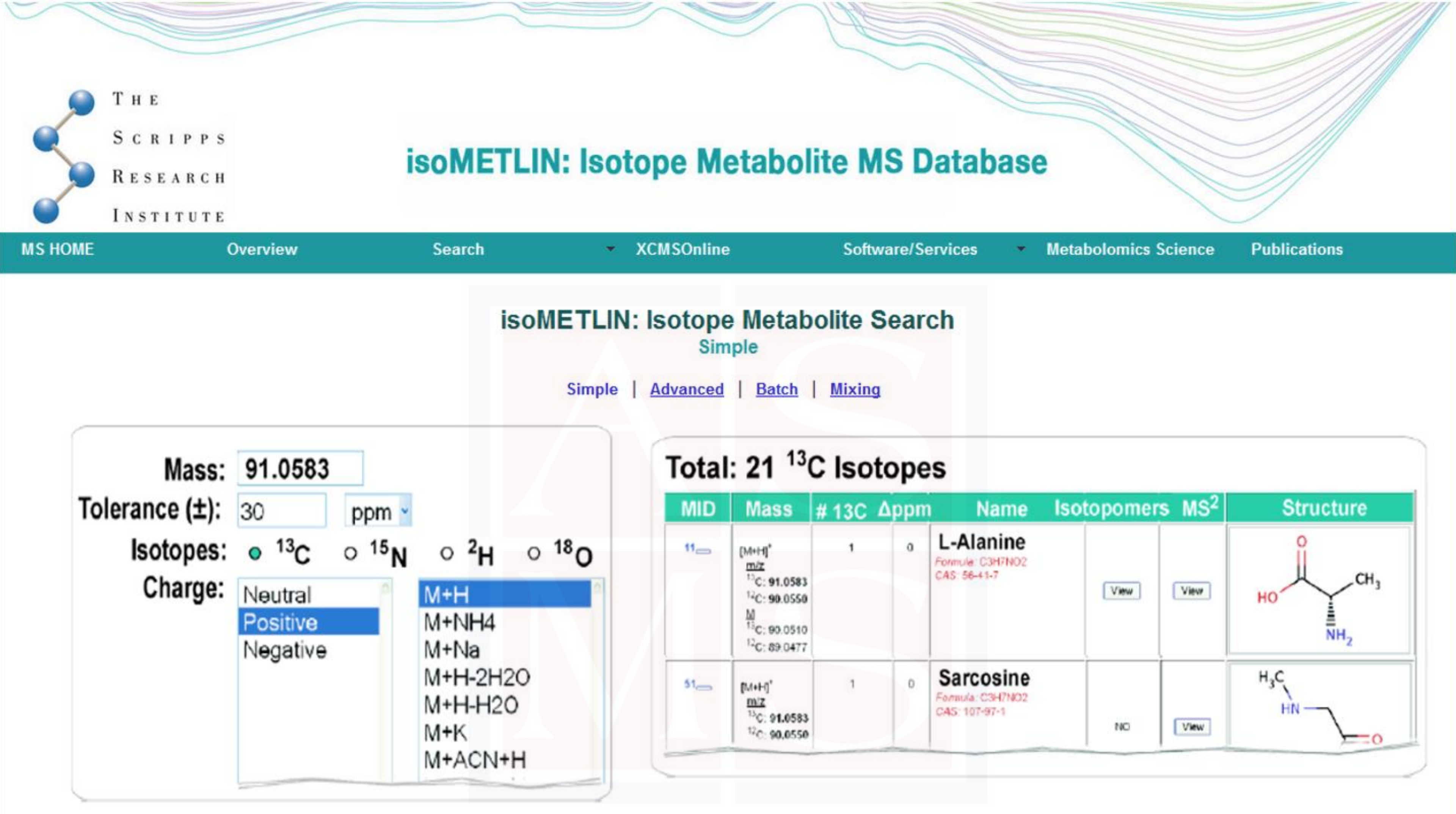
$X^{13}CMS$ output



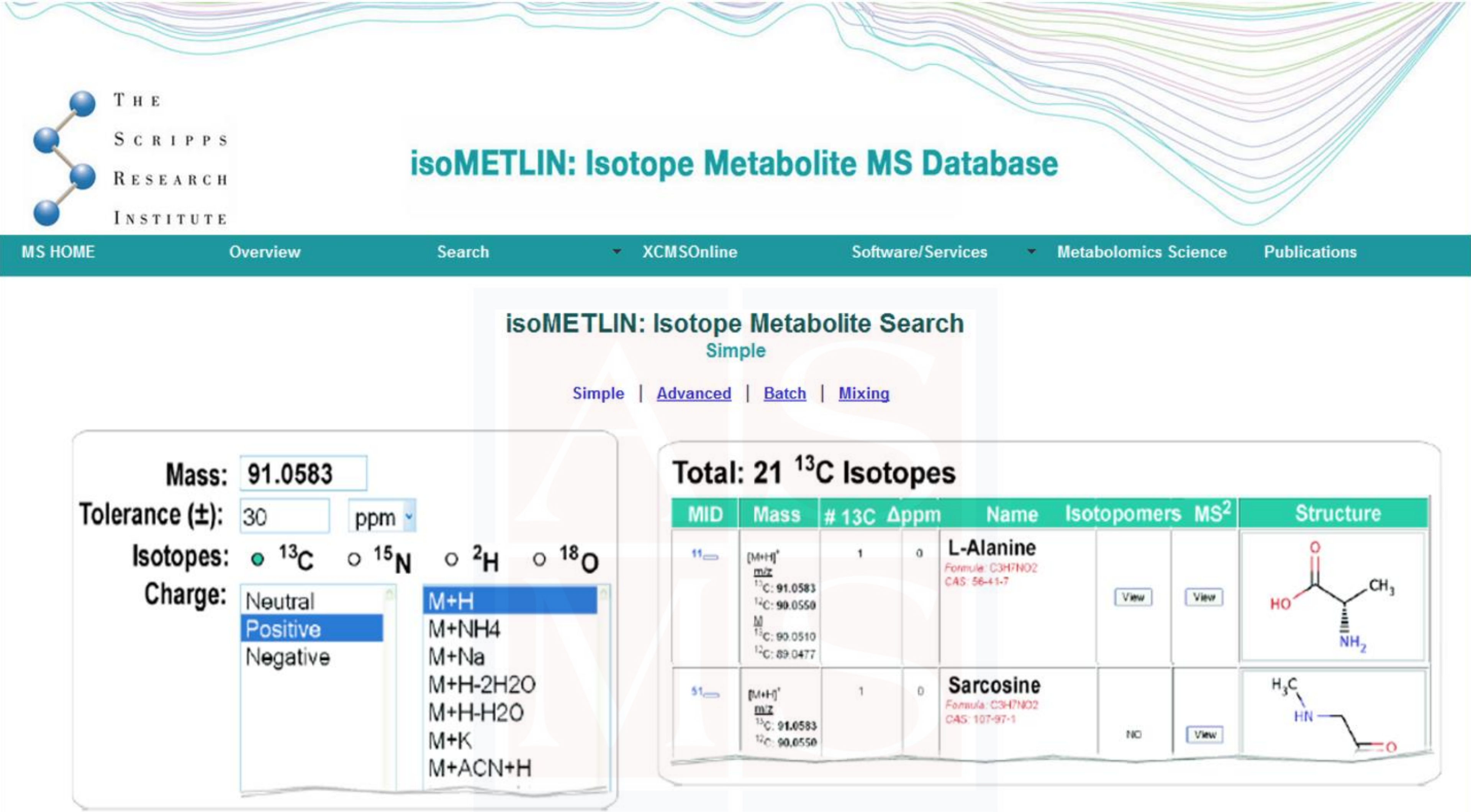
Discriminating isotopomers by fragmentation data

Discriminating isotopomers by fragmentation data





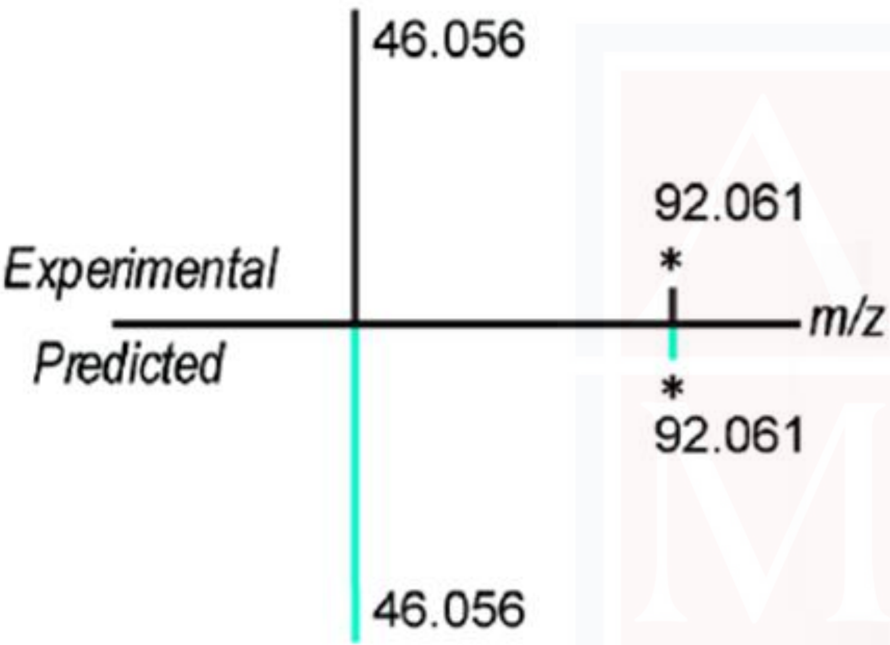
Cho et al., Anal Chem (2014)



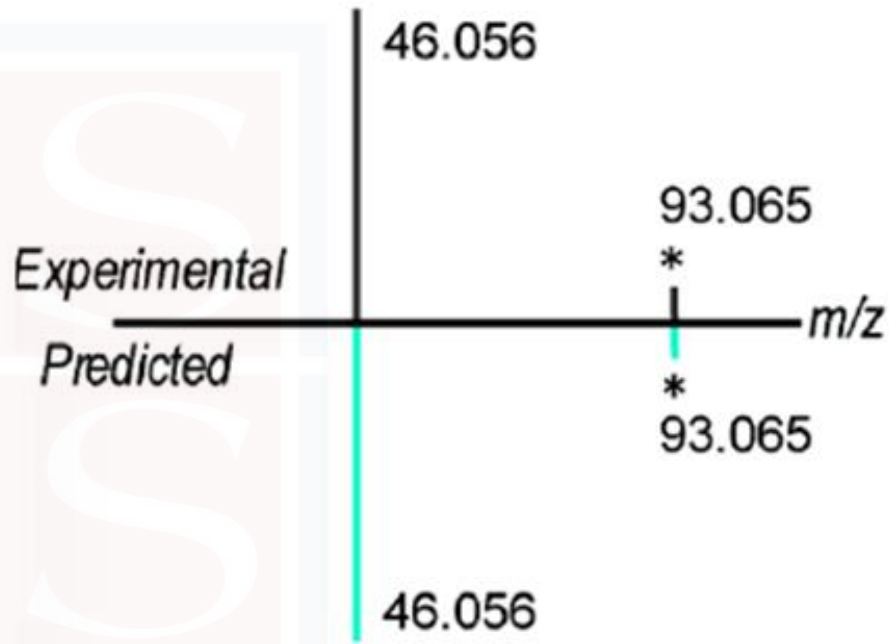
MS/MS data on ~700 isotopomers

Cho et al., Anal Chem (2014)

Predicting isotopomer patterns



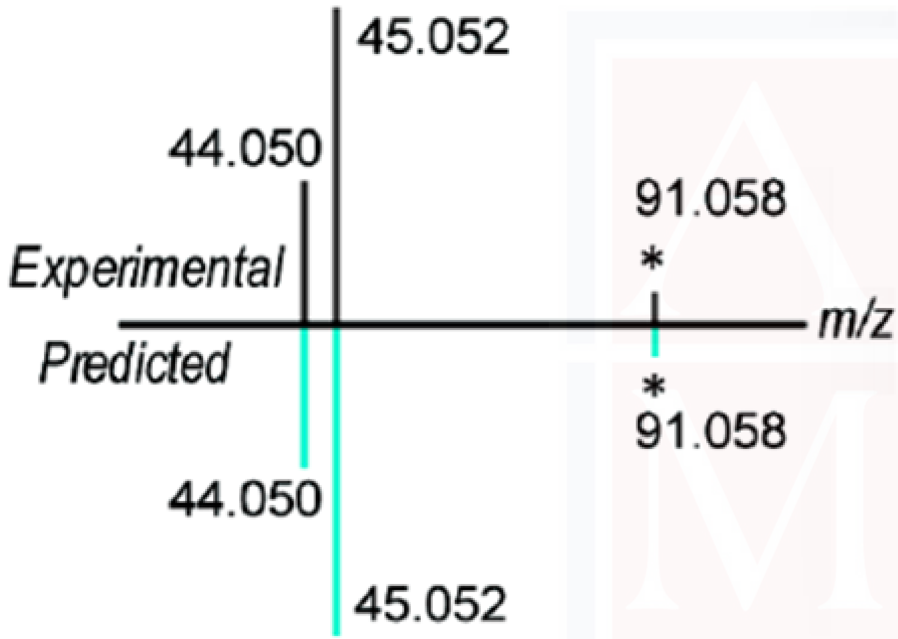
2,3-^{13}C alanine



1,2,3-^{13}C alanine

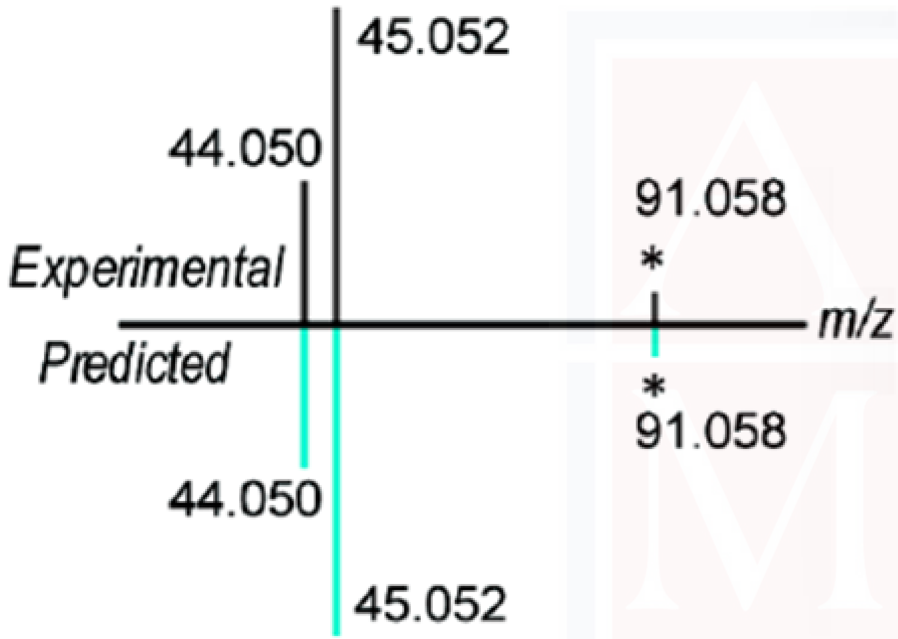
Computational mixing function

Computational mixing function

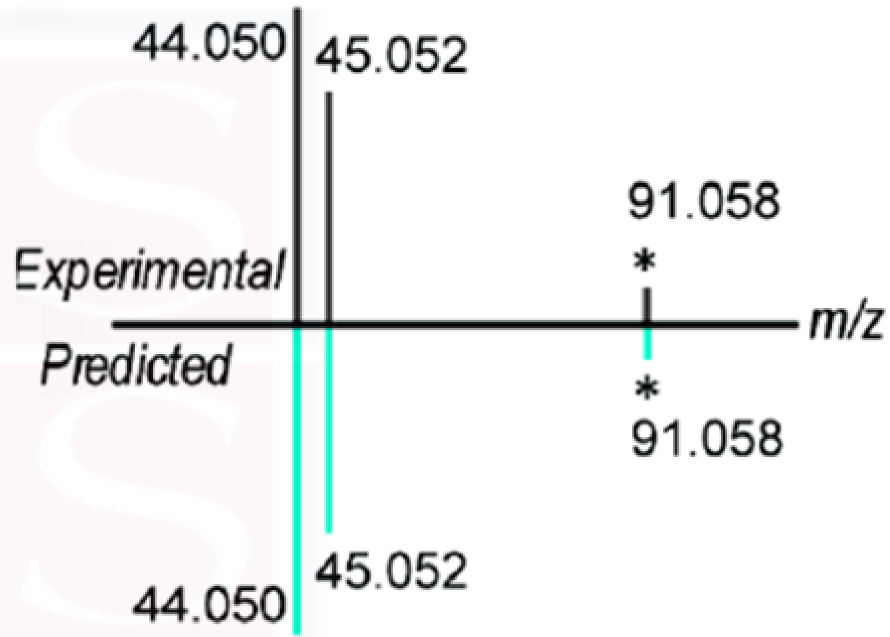


1-^{13}C: 2-^{13}C: 3-^{13}C
alanine @ 1:1:1

Computational mixing function



1-^{13}C: 2-^{13}C: 3-^{13}C
alanine @ 1:1:1



1-^{13}C: 2-^{13}C: 3-^{13}C
alanine @ 3:1:1

The Scripps Research Institute

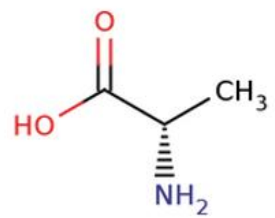
isoMETLIN: Isotope Metabolite MS Database

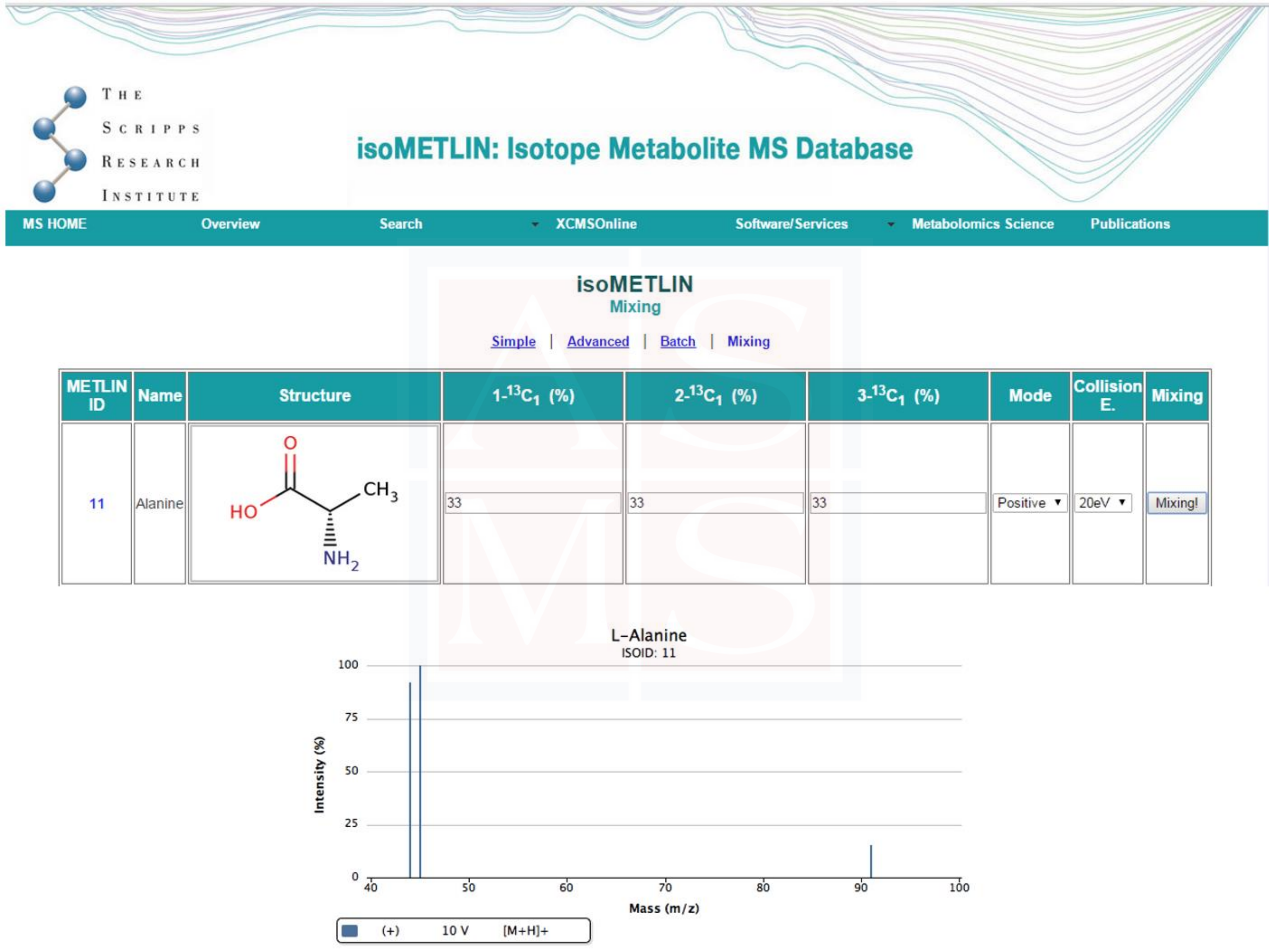
MS HOME | Overview | Search | XCMSOnline | Software/Services | Metabolomics Science | Publications

isoMETLIN

Mixing

Simple | Advanced | Batch | Mixing

METLIN ID	Name	Structure	1-$^{13}C_1$ (%)	2-$^{13}C_1$ (%)	3-$^{13}C_1$ (%)	Mode	Collision E.	Mixing
11	Alanine		33	33	33	Positive ▾	20eV ▾	Mixing!



THE SCRIPPS RESEARCH INSTITUTE
isoMETLIN: Isotope Metabolite MS Database
MS HOME
Overview
Search
XCMSOnline
Software/Services
Metabolomics Science
Publications
isoMETLIN
Mixing
Simple | Advanced | Batch | Mixing
METLIN ID
Name
Structure
1-13C1 (%)
2-13C1 (%)
3-13C1 (%)
Mode
Collision E.
Mixing
11
Alanine
O
HO
CH3
NH2
33
33
33
Positive
20eV
Mixing!
L-Alanine
ISOID: 11
Intensity (%)
Mass (m/z)
(+)
10 V
[M+H]+

Application 1: 2-hydroxyglutarate

Known pathways

$$\text{isocitrate} \xrightarrow[\text{NADP}^+ \;\rightarrow\; \text{NADPH}]{\text{WT IDH1}} \alpha\text{-KG} + CO_2$$

Application 1: 2-hydroxyglutarate

Known pathways

$$\text{isocitrate} \xrightarrow[\text{NADP}^+ \rightarrow \text{NADPH}]{\text{WT IDH1}} \alpha\text{-KG} + CO_2$$

$$\alpha\text{-KG} \xrightarrow[\text{NADPH} \rightarrow \text{NADP}^+]{\text{mutant IDH1}} \text{2-HG}$$

<u>Application 1</u>: 2-hydroxyglutarate

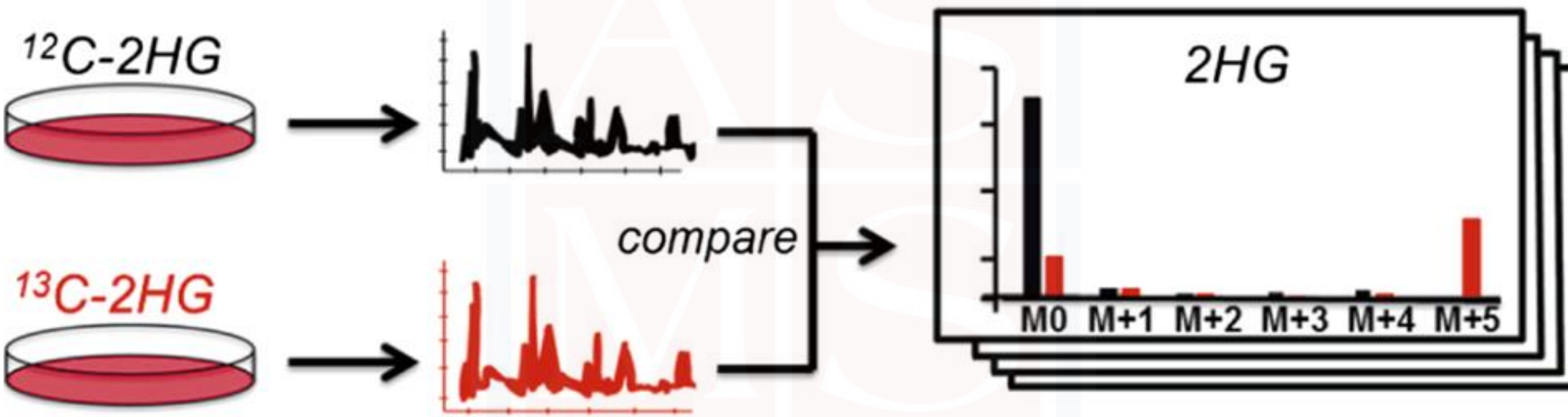
Known pathways

$$\text{isocitrate} \xrightarrow[\text{NADP}^+ \;\to\; \text{NADPH}]{\text{WT IDH1}} \alpha\text{-KG} + CO_2$$

$$\alpha\text{-KG} \xrightarrow[\text{NADPH} \;\to\; \text{NADP}^+]{\text{mutant IDH1}} \text{2-HG} \longrightarrow ?$$

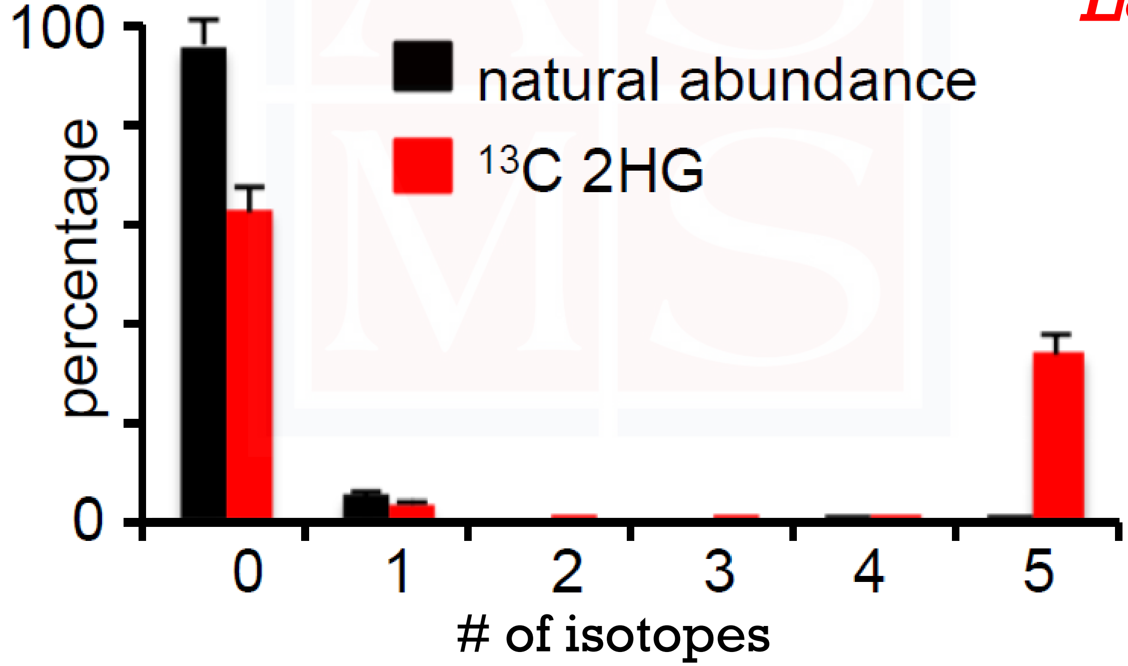
Application 1: 2-hydroxyglutarate

Is 2HG metabolized?



Application 1: 2-hydroxyglutarate

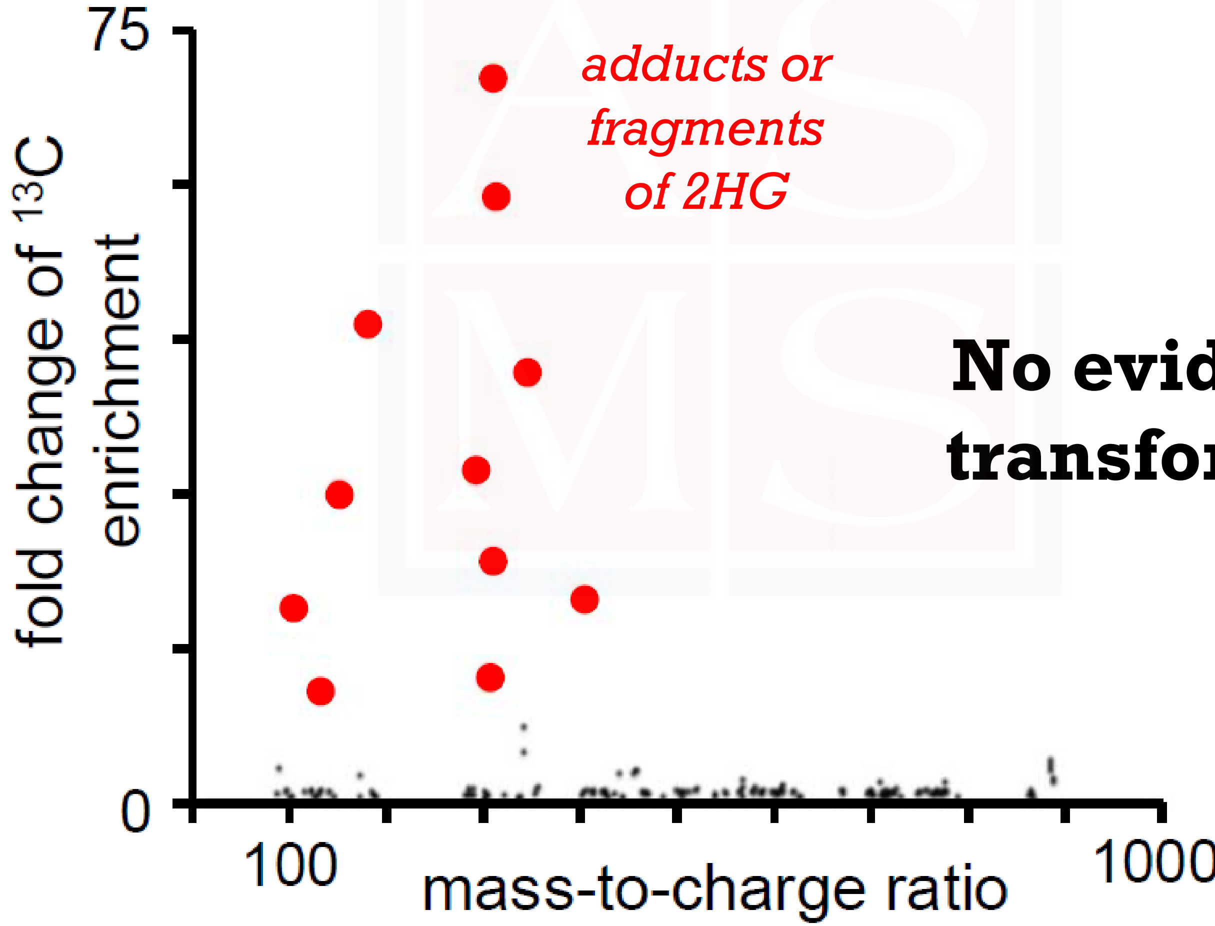
Is 2HG metabolized?



Label got in cells

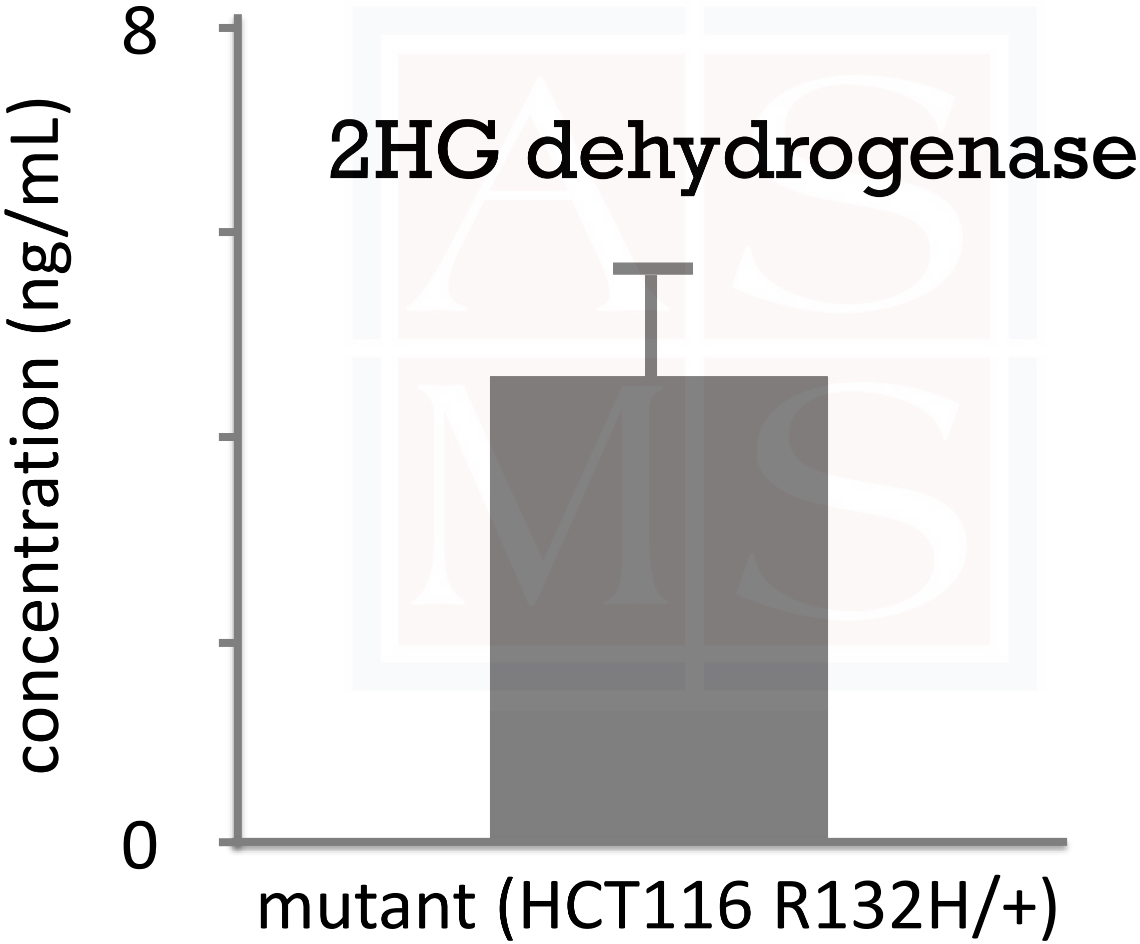
<u>Application 1</u>: 2-hydroxyglutarate

Is 2HG metabolized?



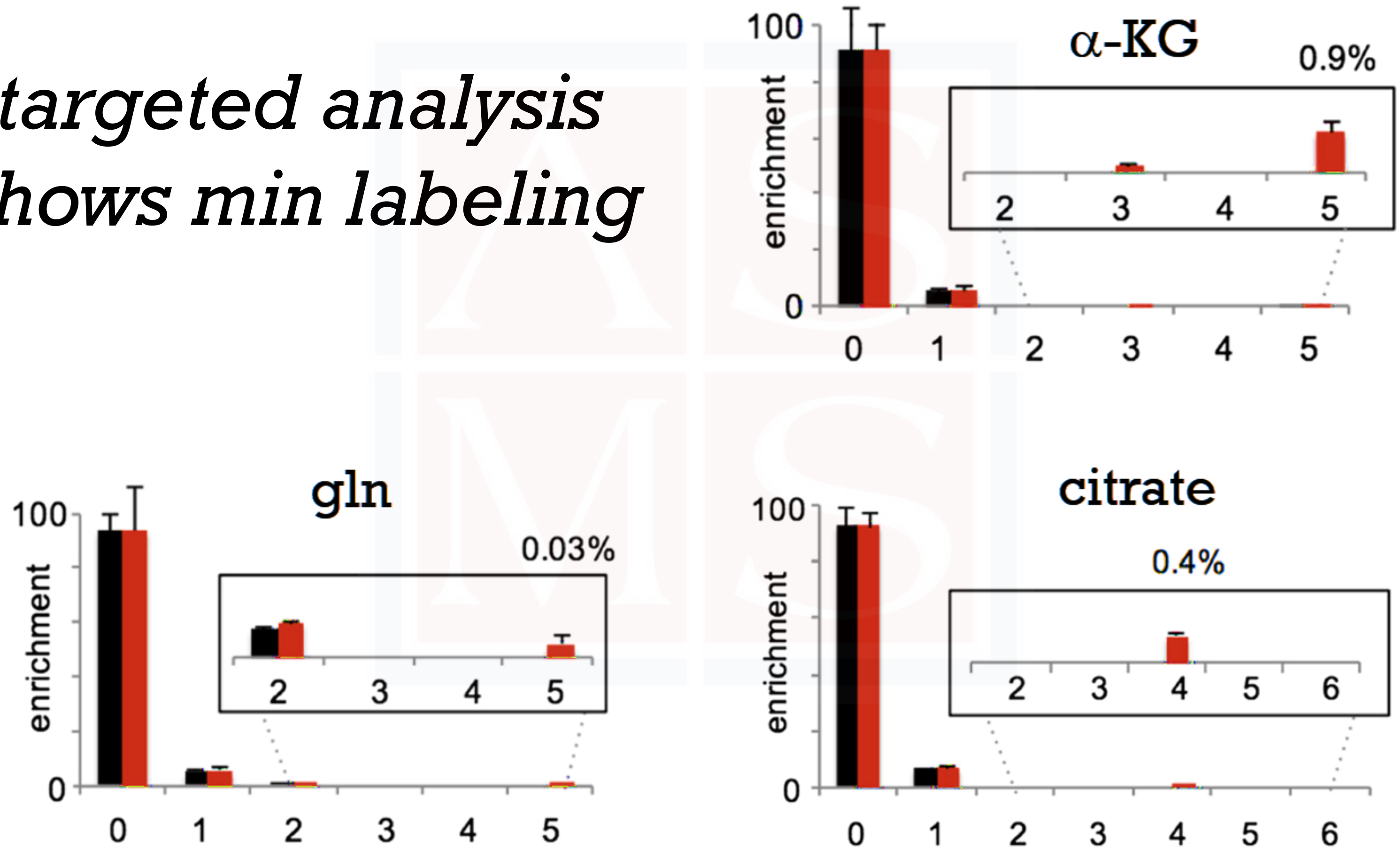
No evidence of transformation

Application 1: 2-hydroxyglutarate

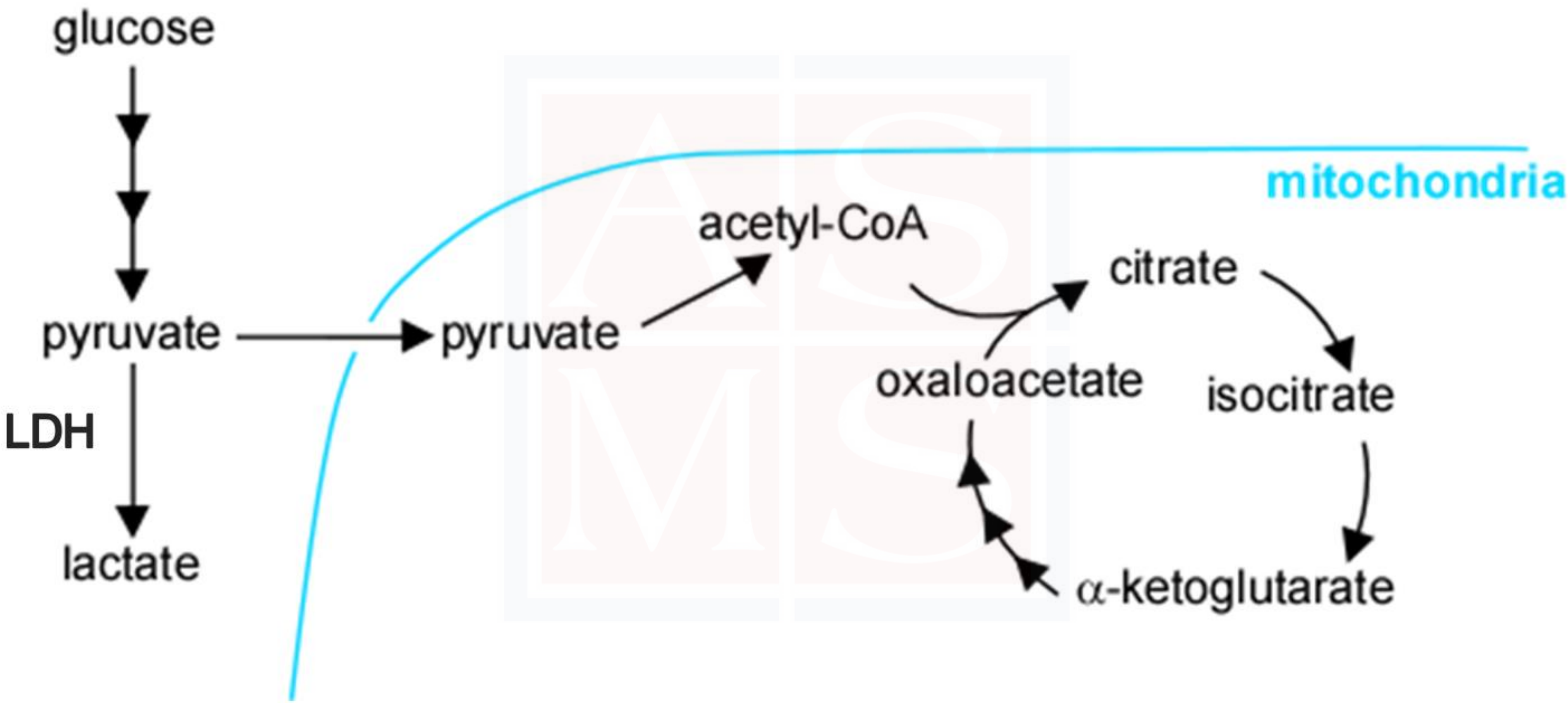


Application 1: 2-hydroxyglutarate

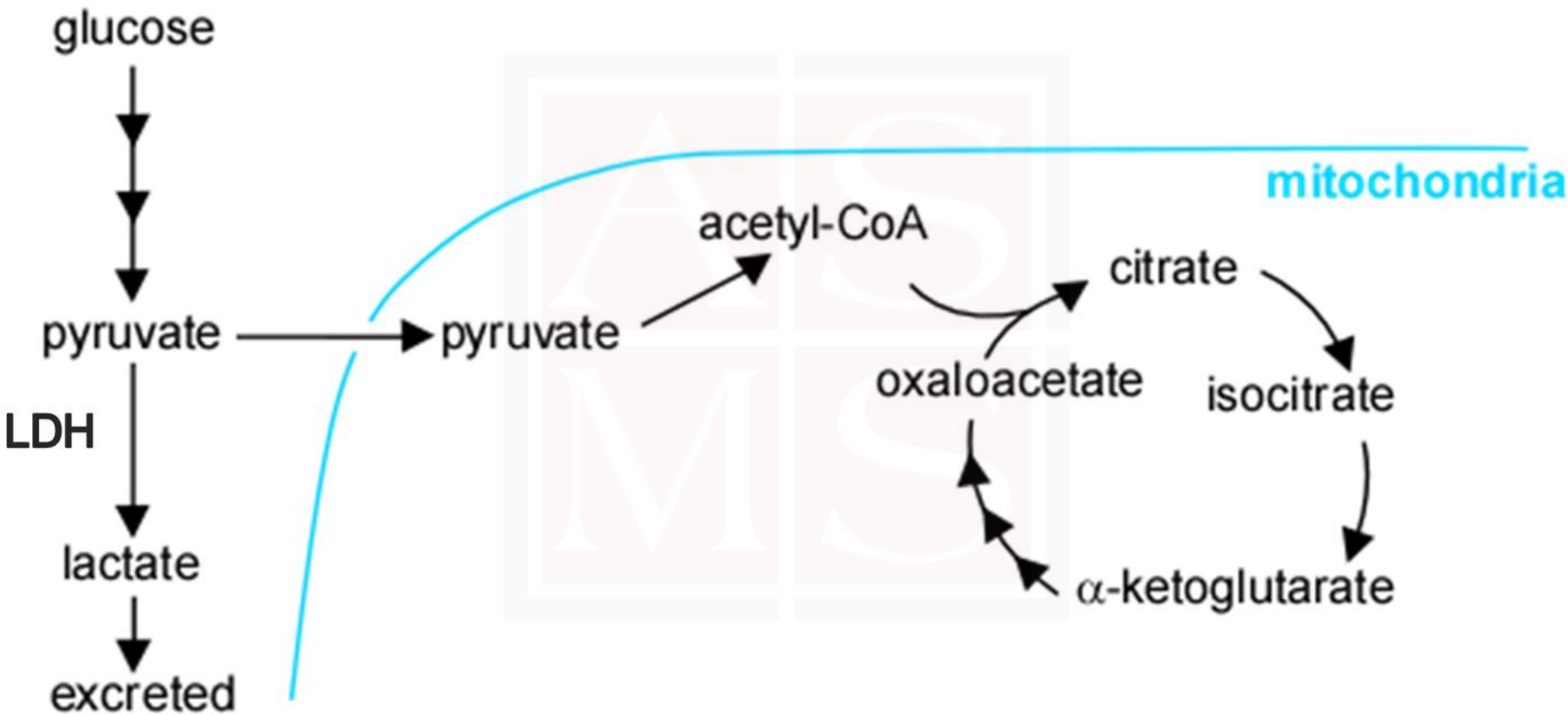
targeted analysis shows min labeling



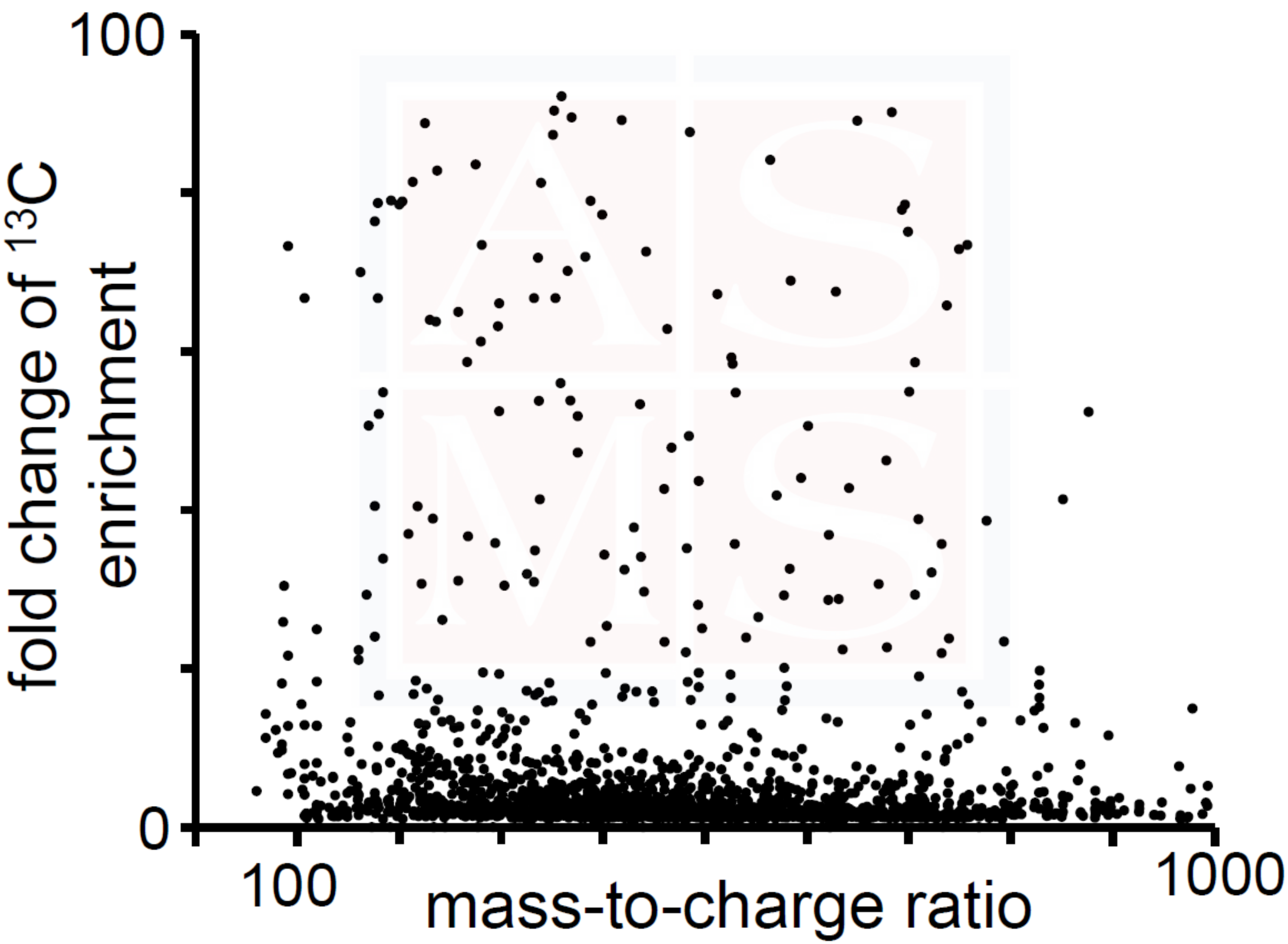
Application 2: lactate



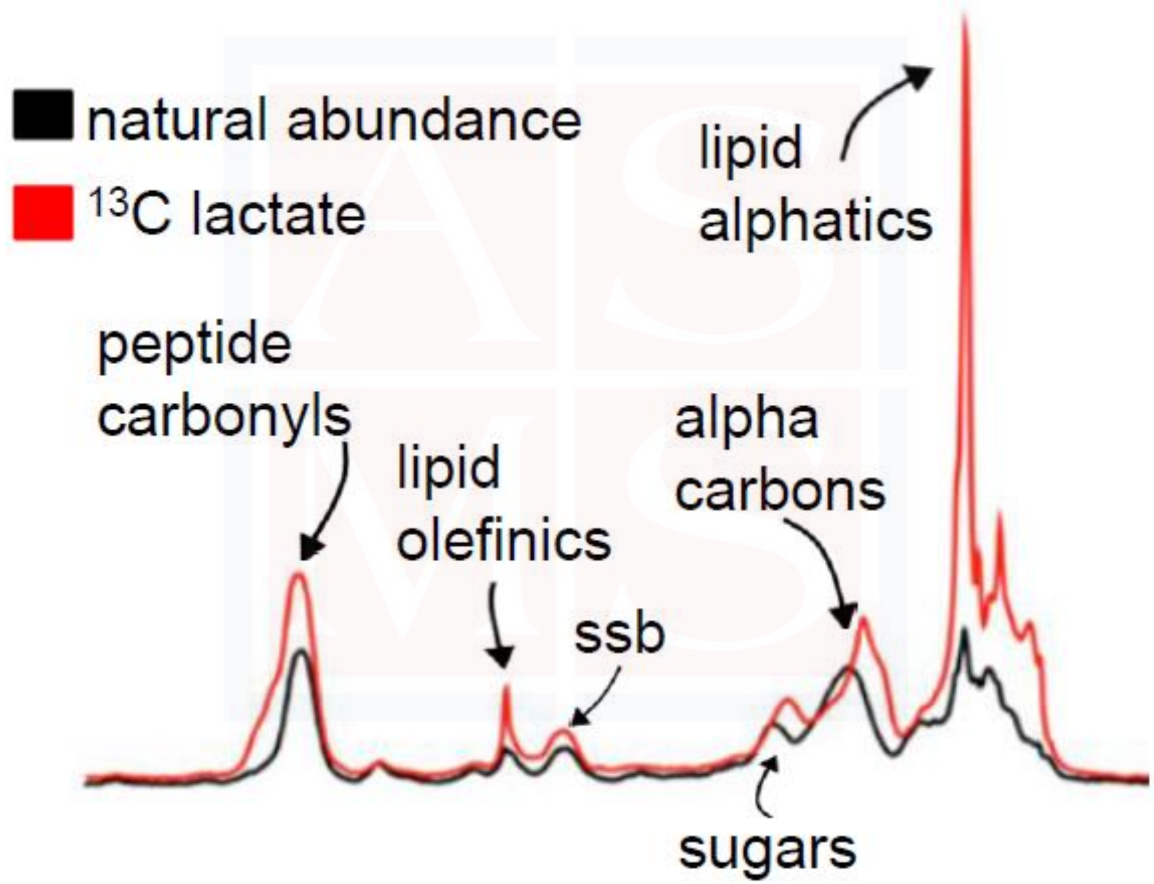
Application 2: lactate



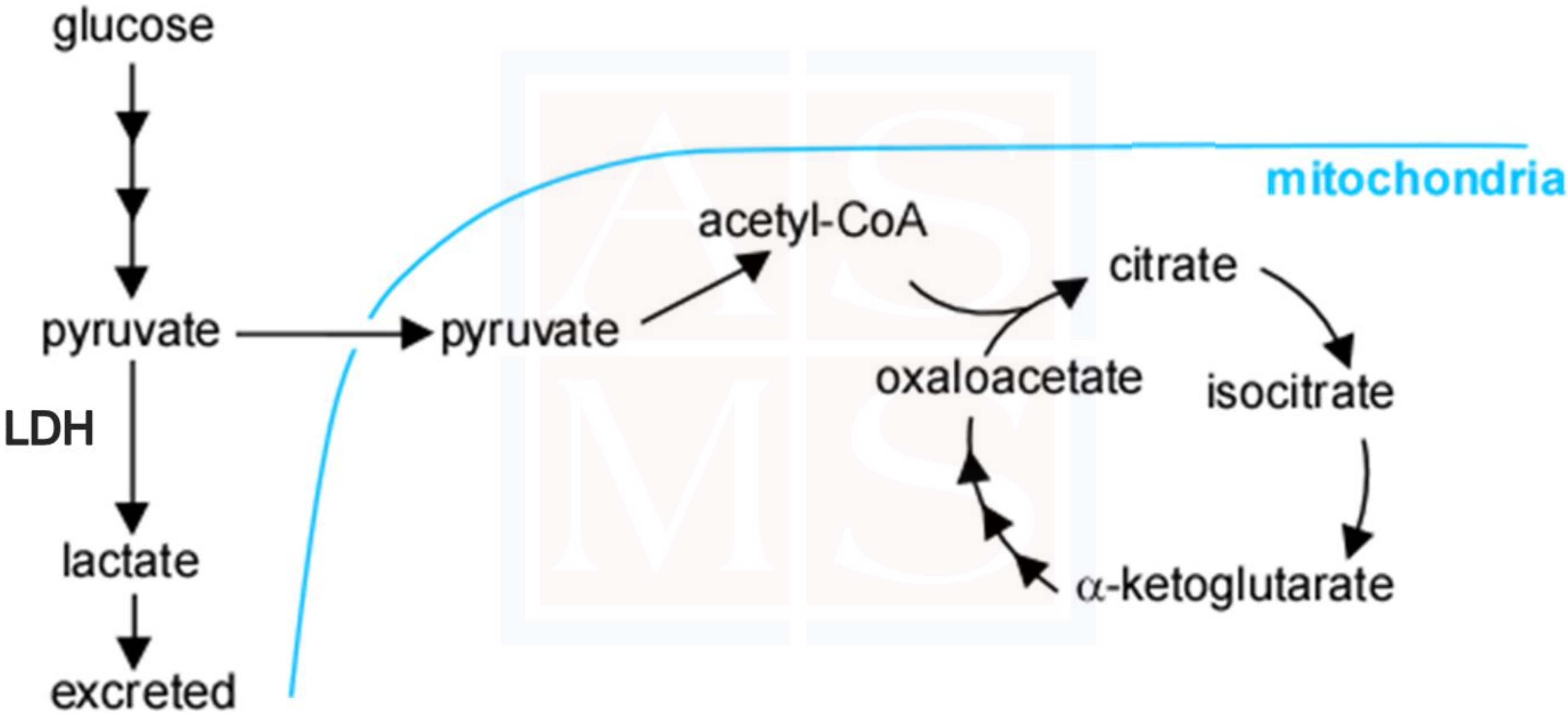
Application 2: lactate



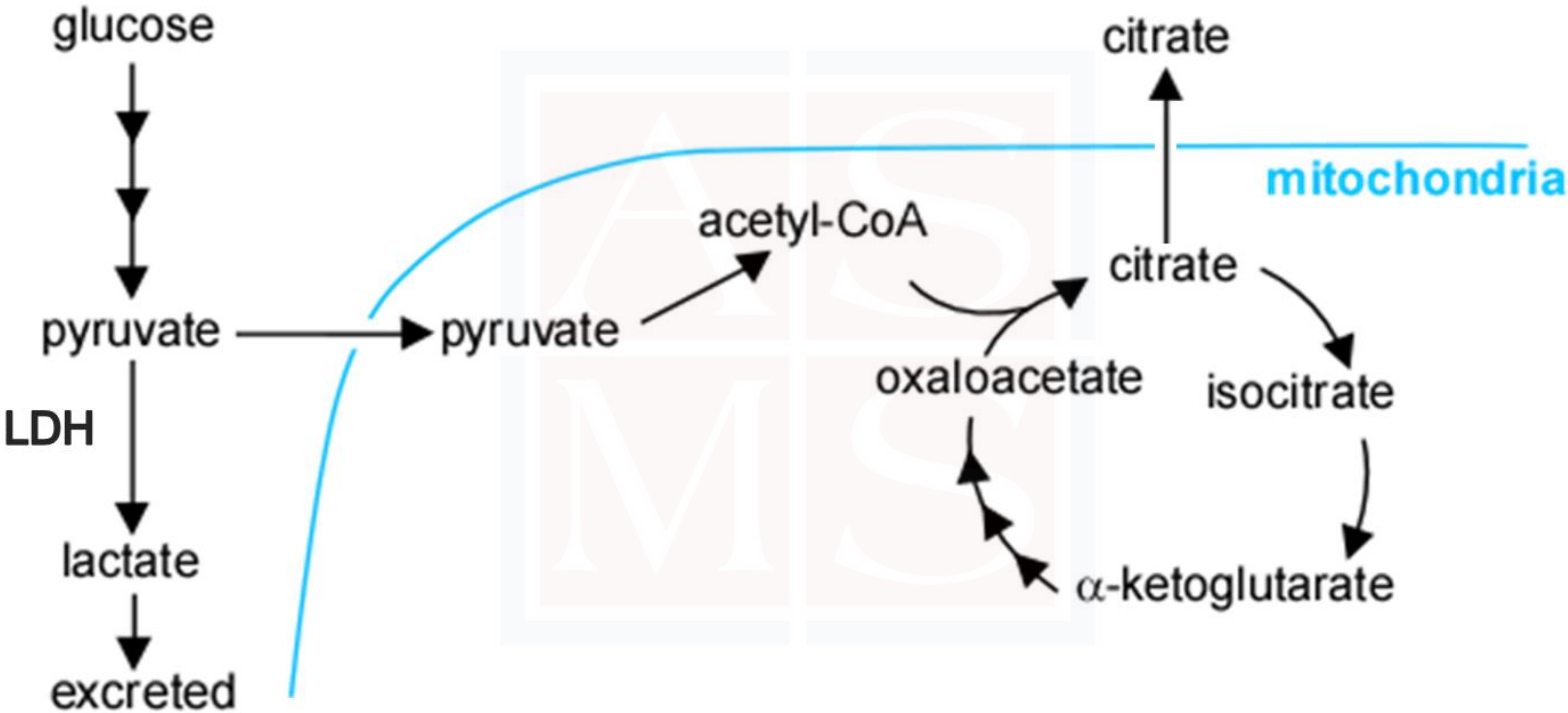
Application 2: lactate



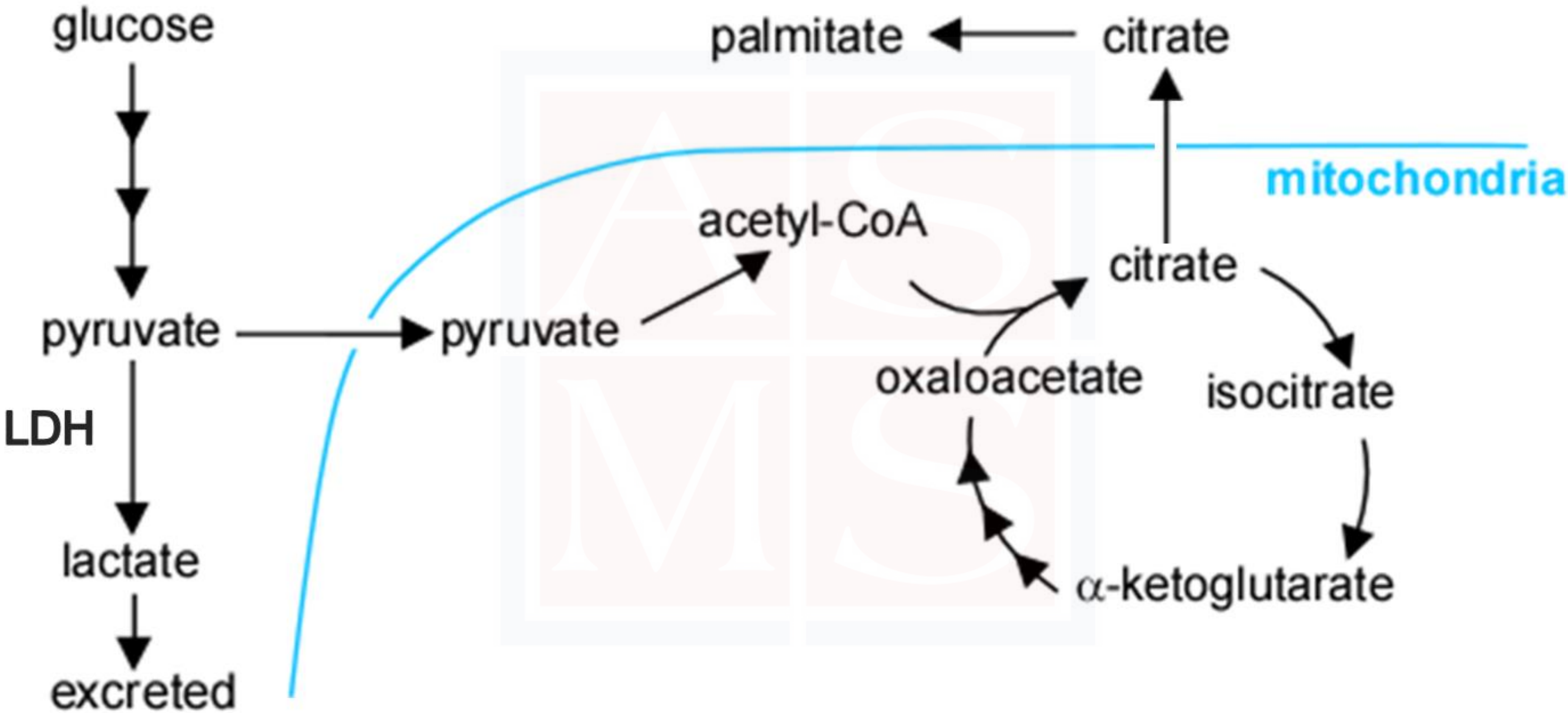
Application 2: lactate



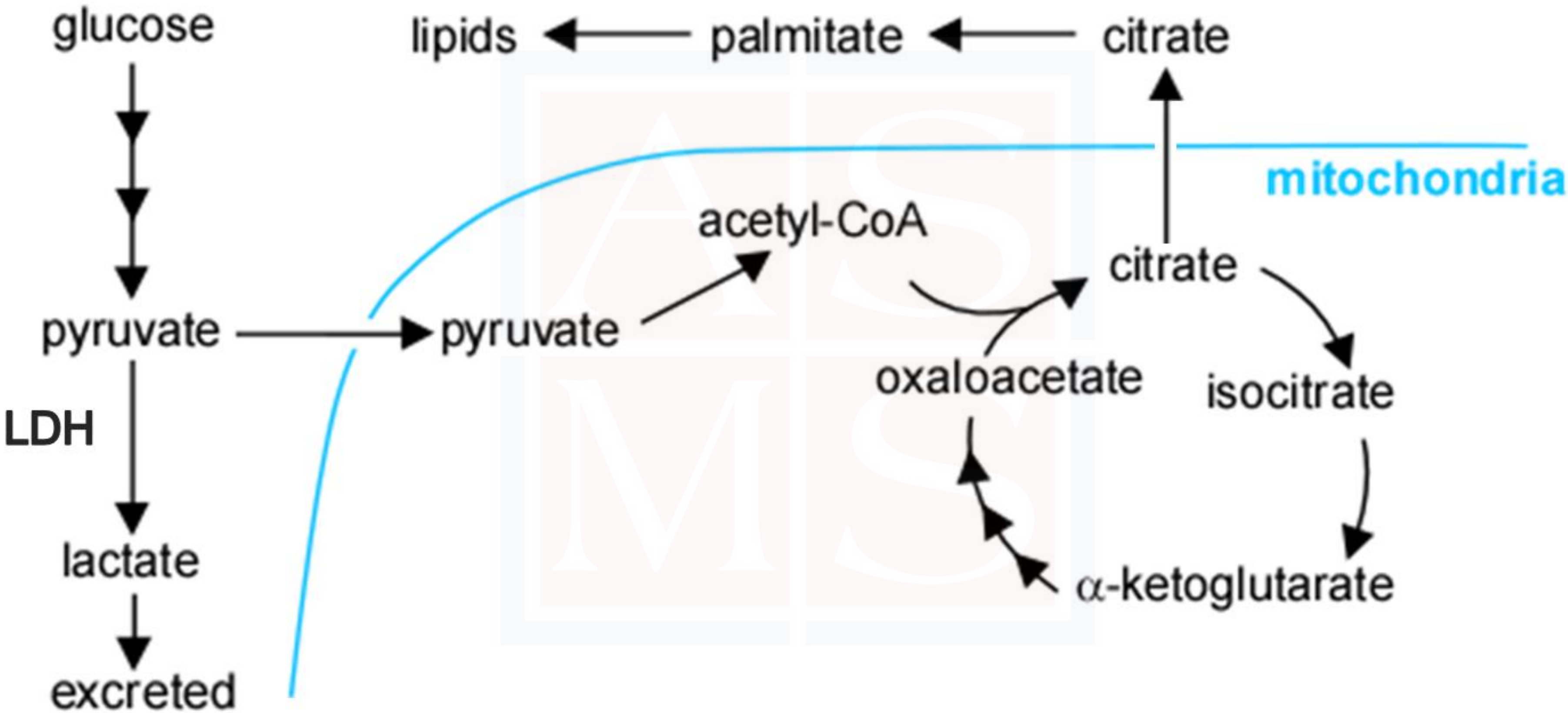
Application 2: lactate



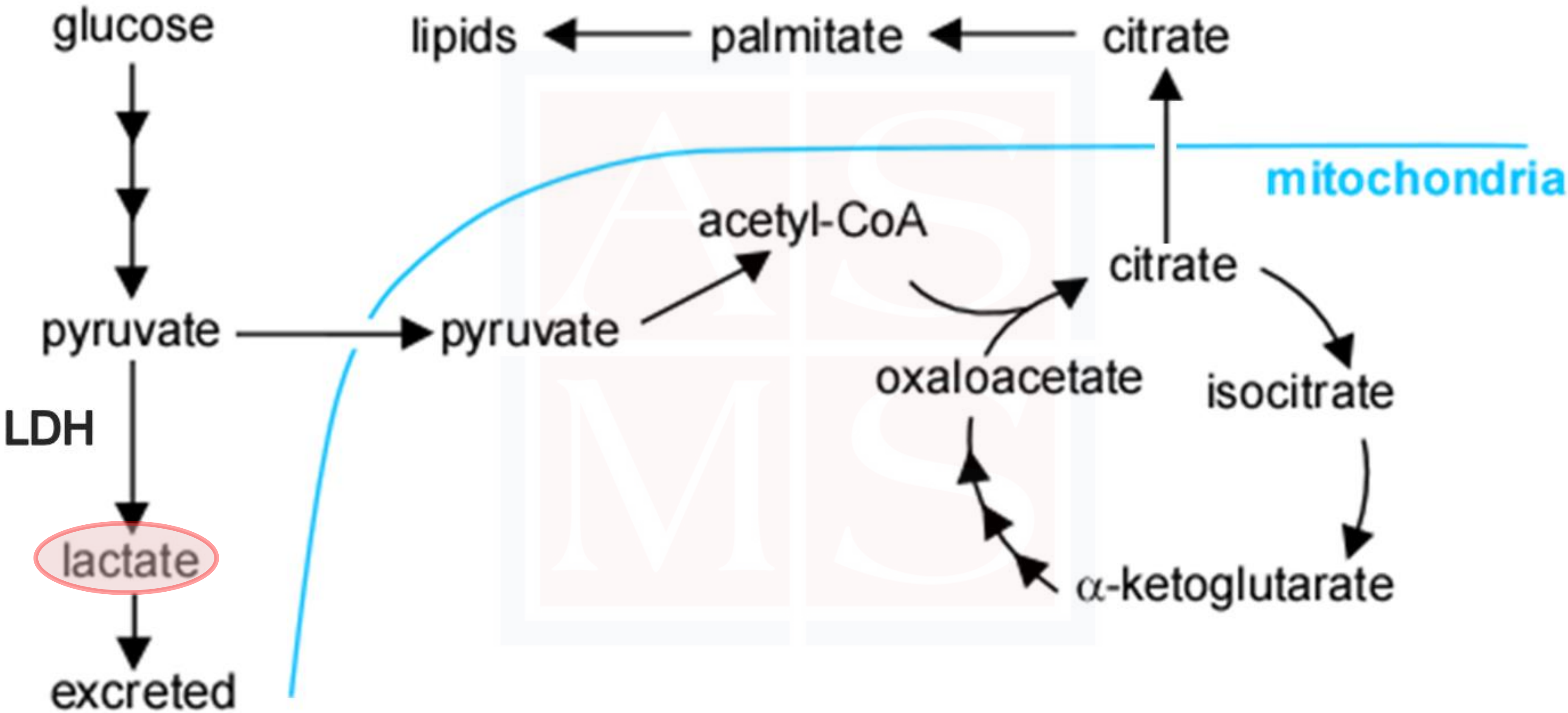
Application 2: lactate



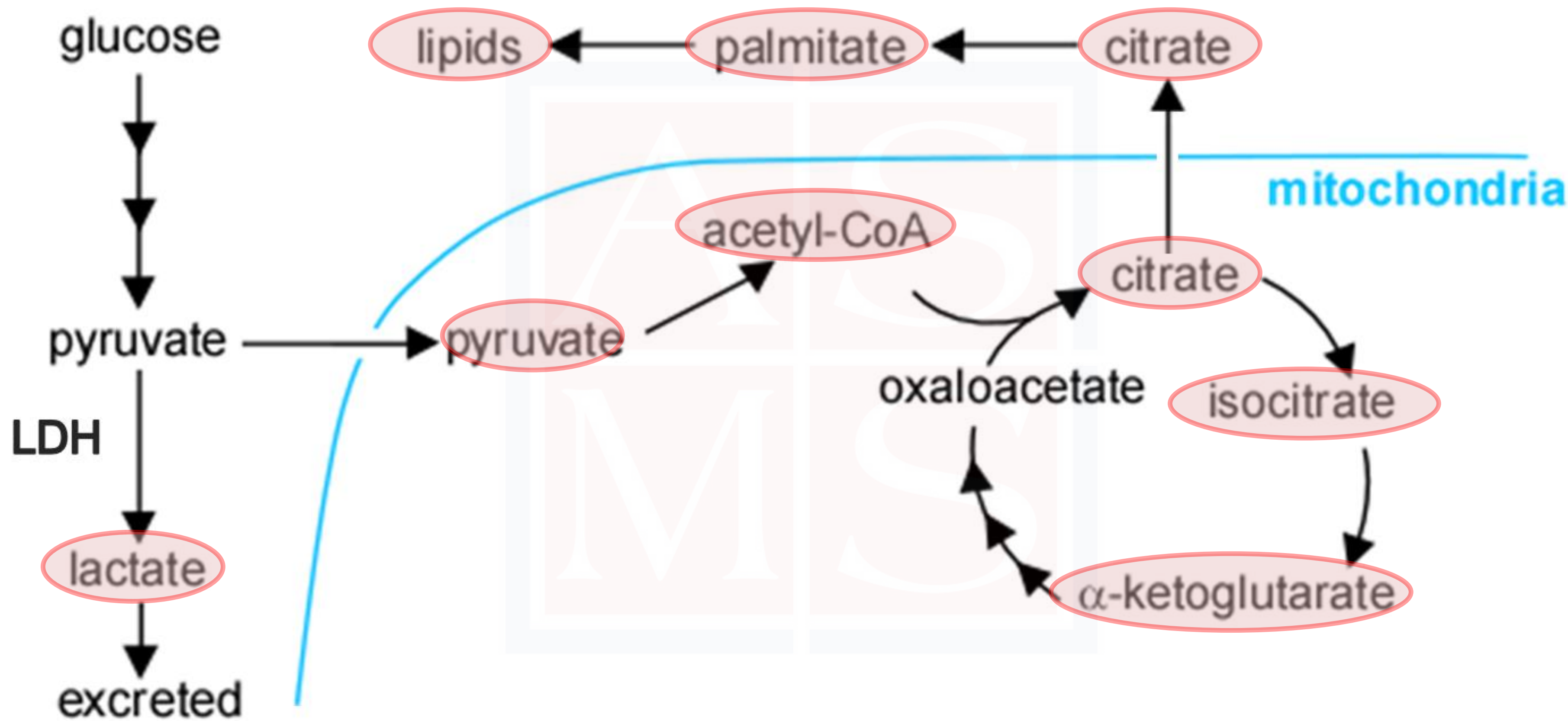
Application 2: lactate



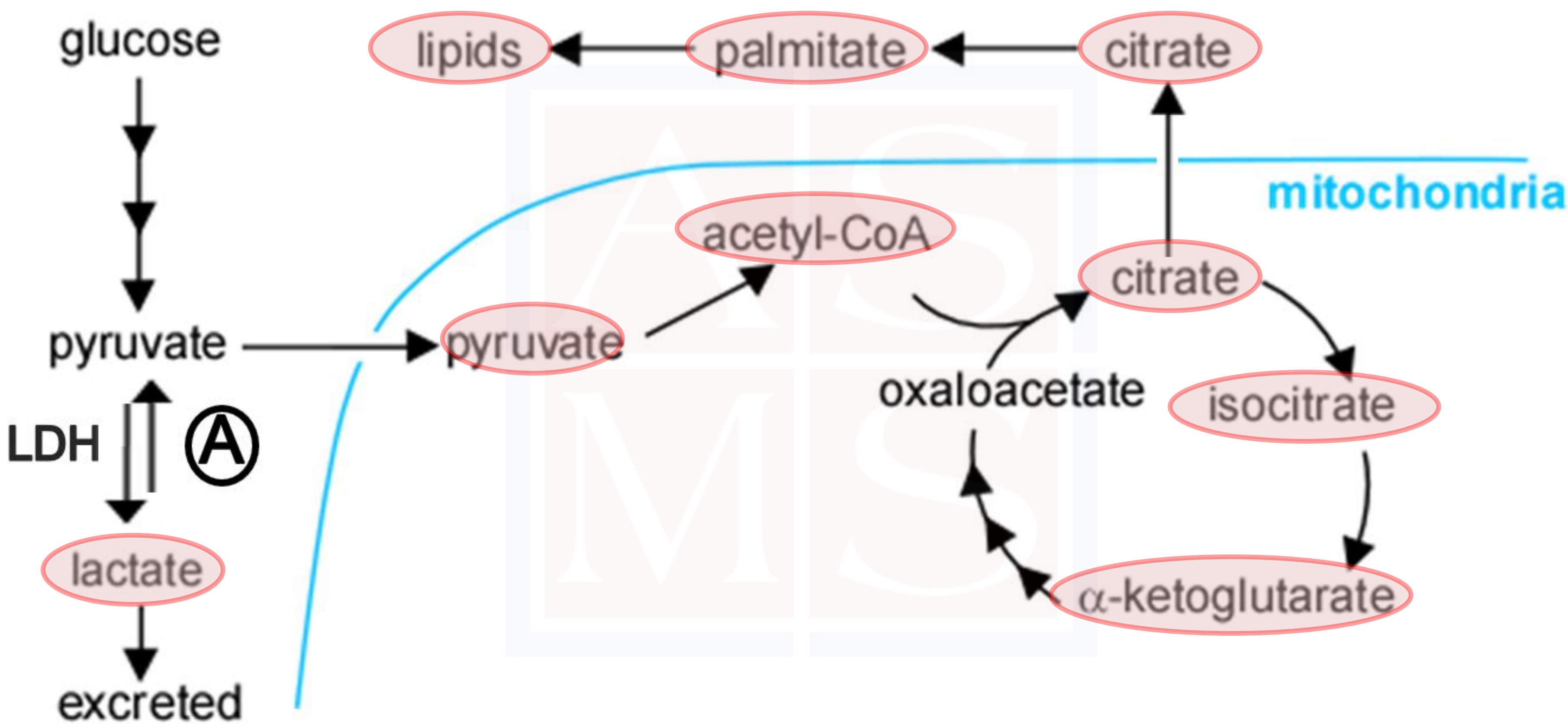
Application 2: lactate



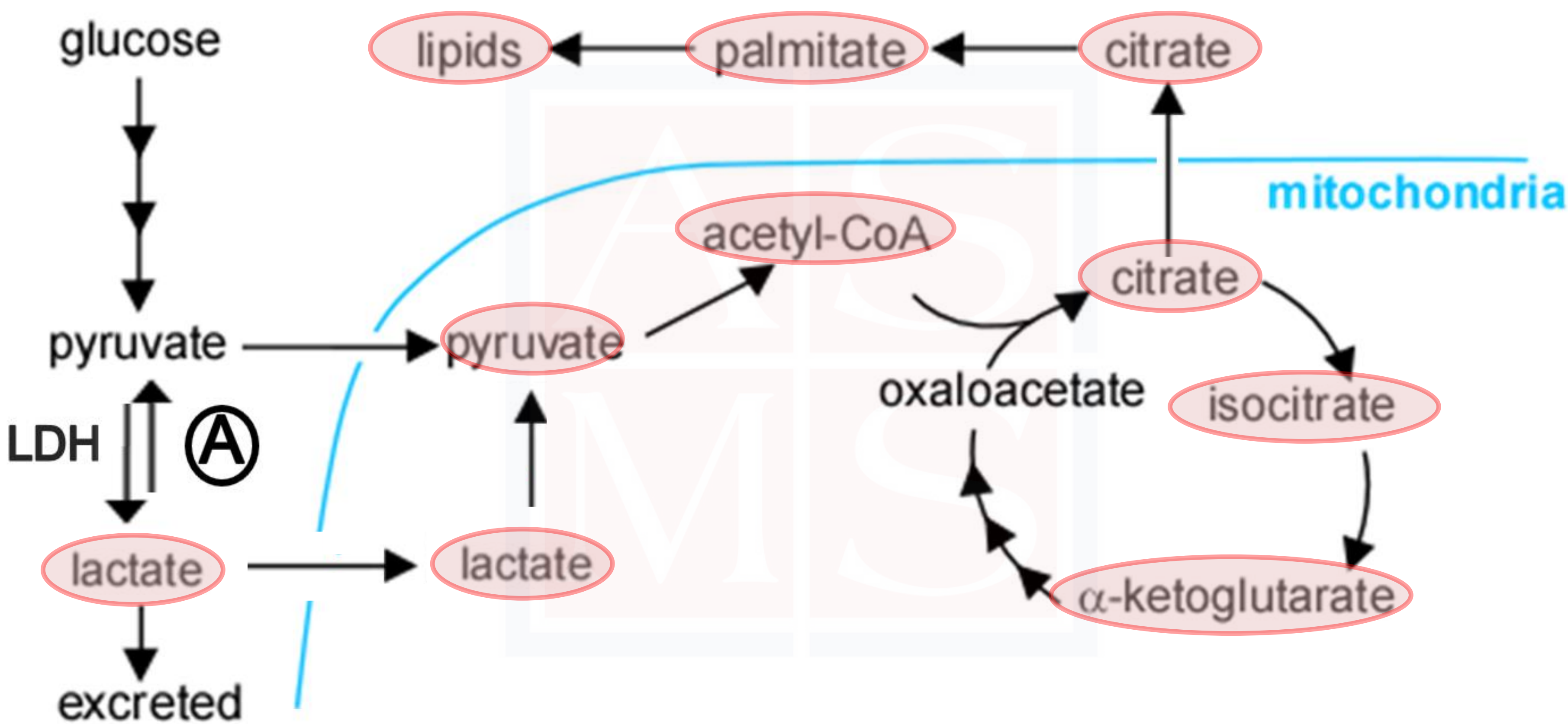
Application 2: lactate



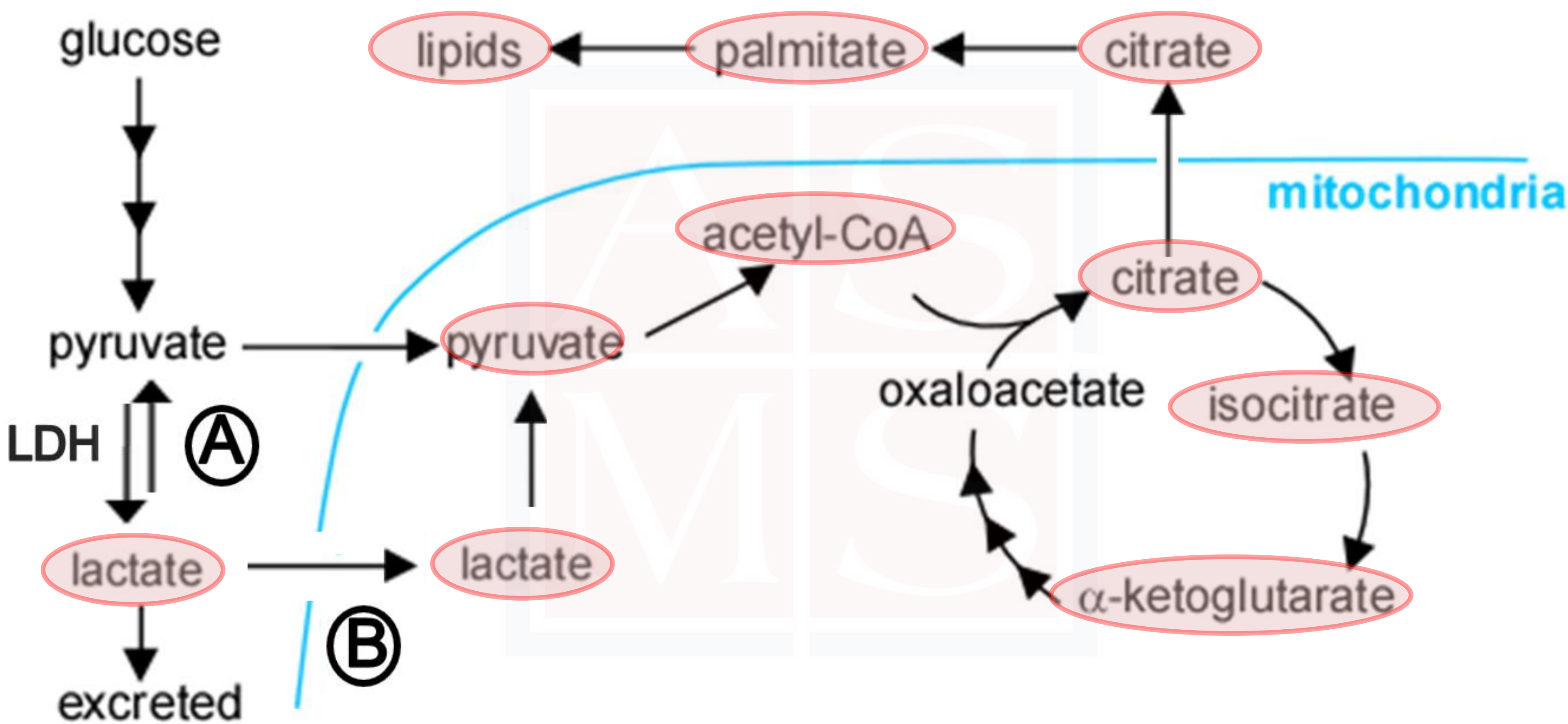
Application 2: lactate



Application 2: lactate

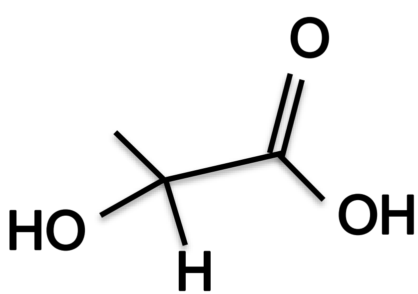


Application 2: lactate



Application 2: lactate

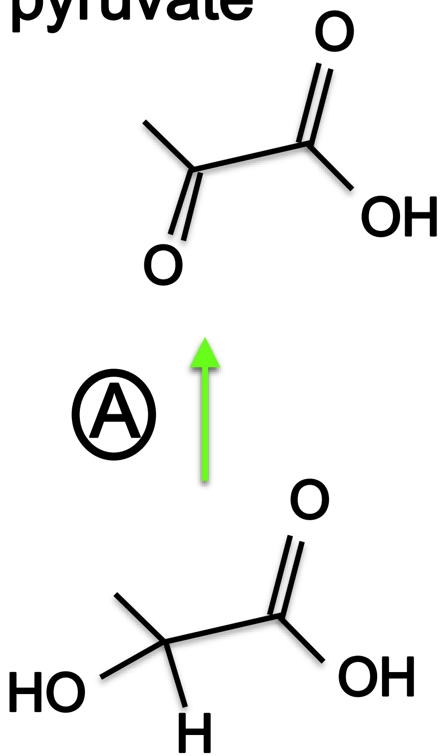
Application 2: lactate



lactate

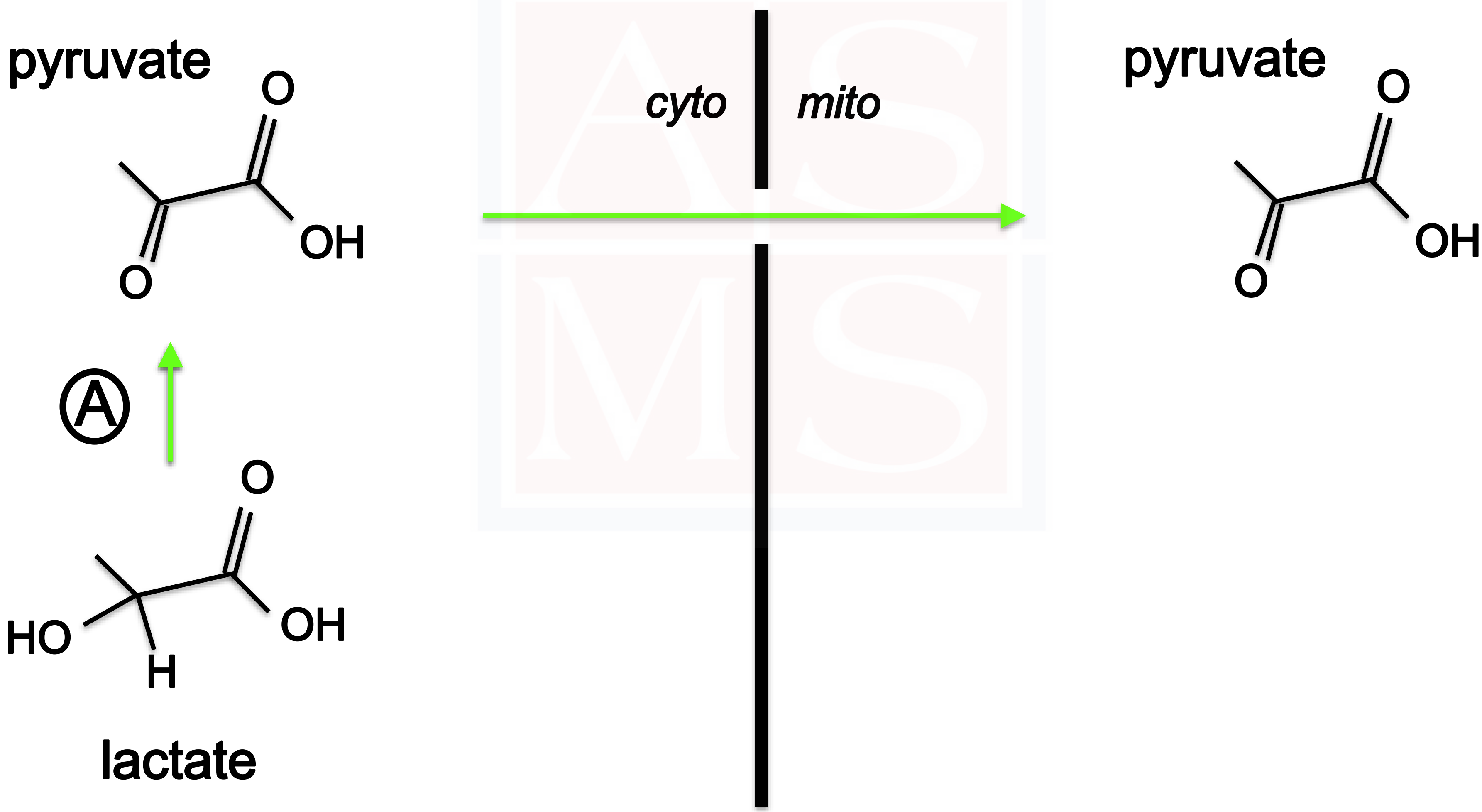
Application 2: lactate

pyruvate

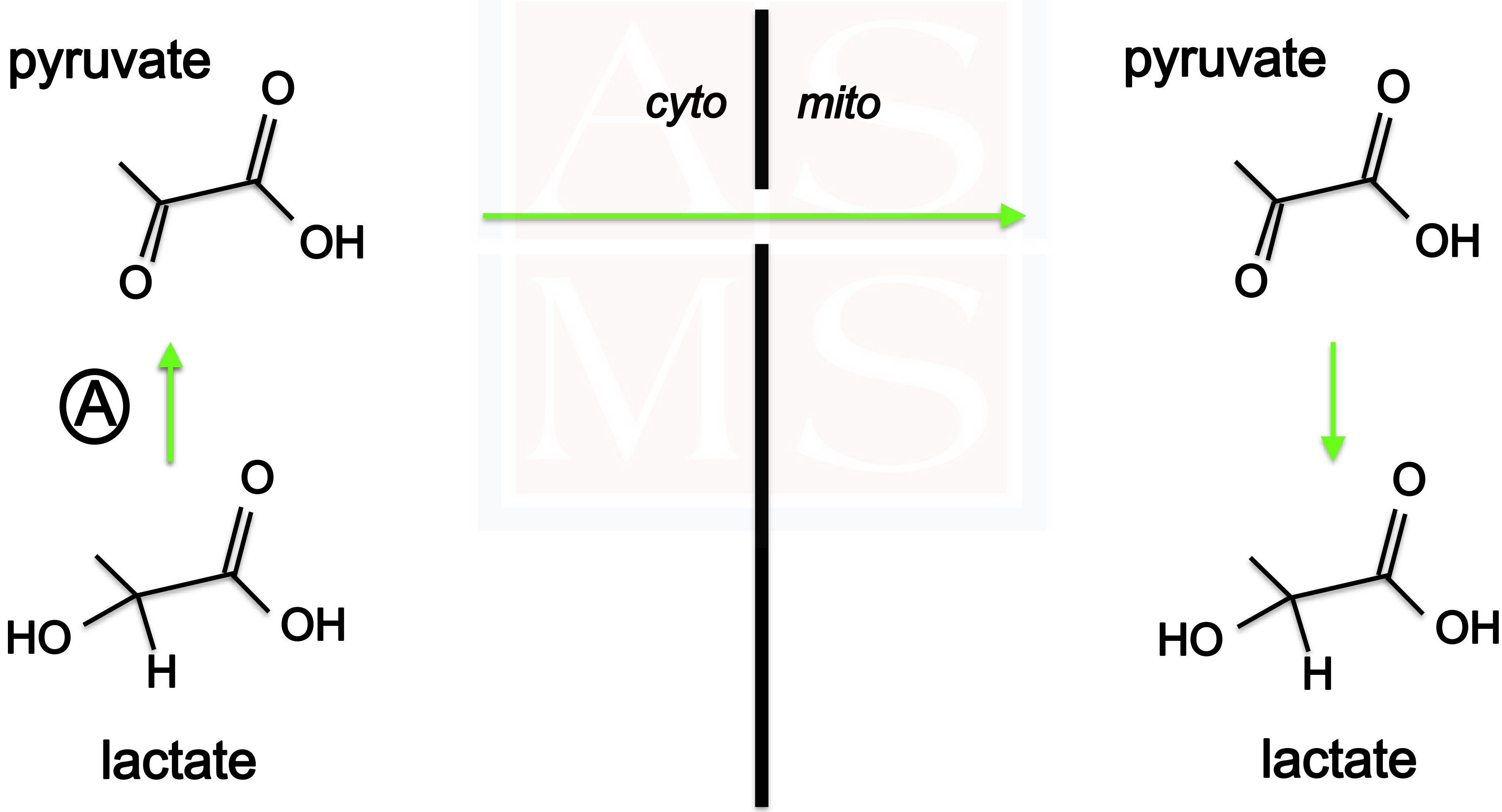


lactate

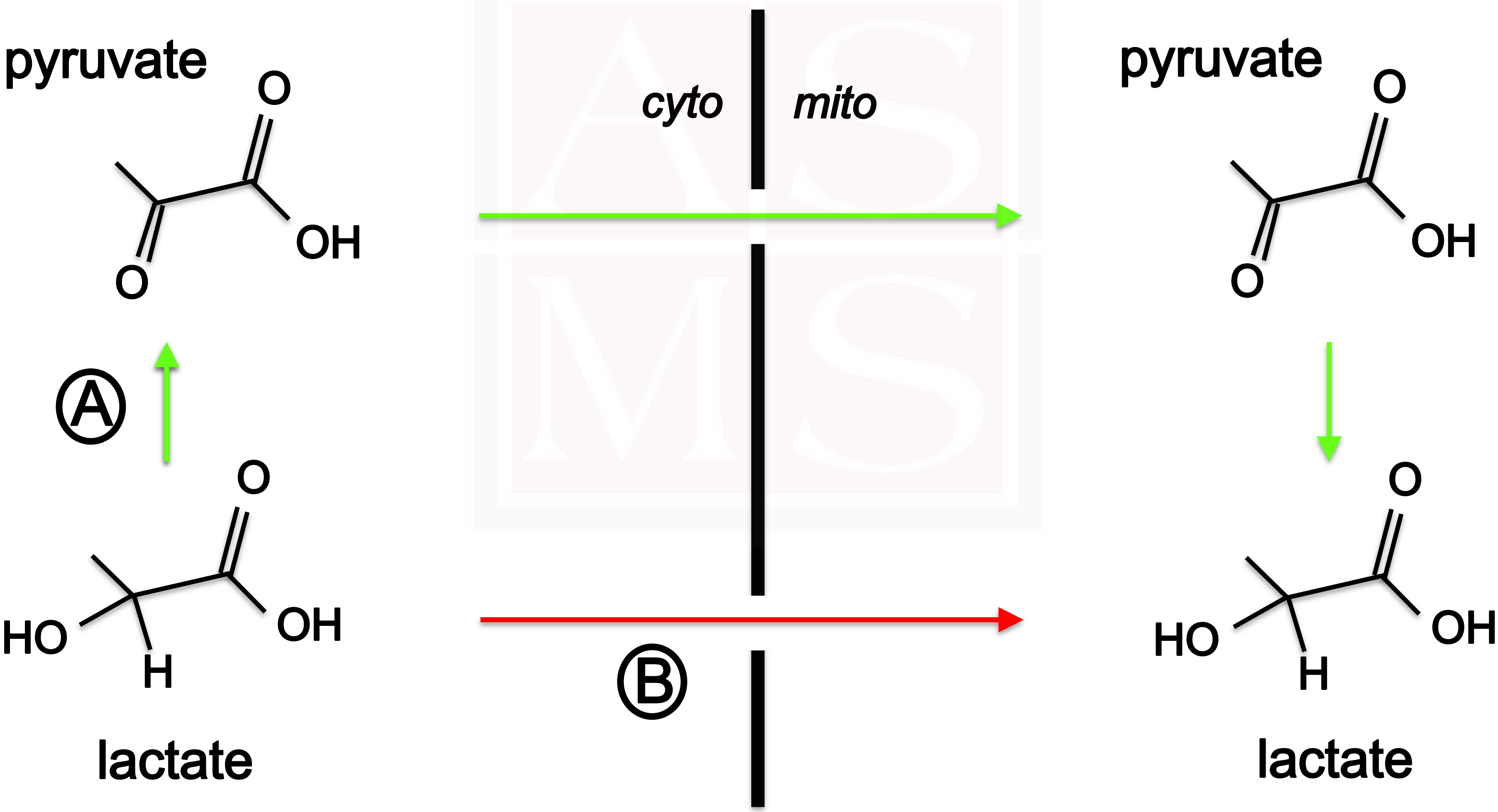
Application 2: lactate



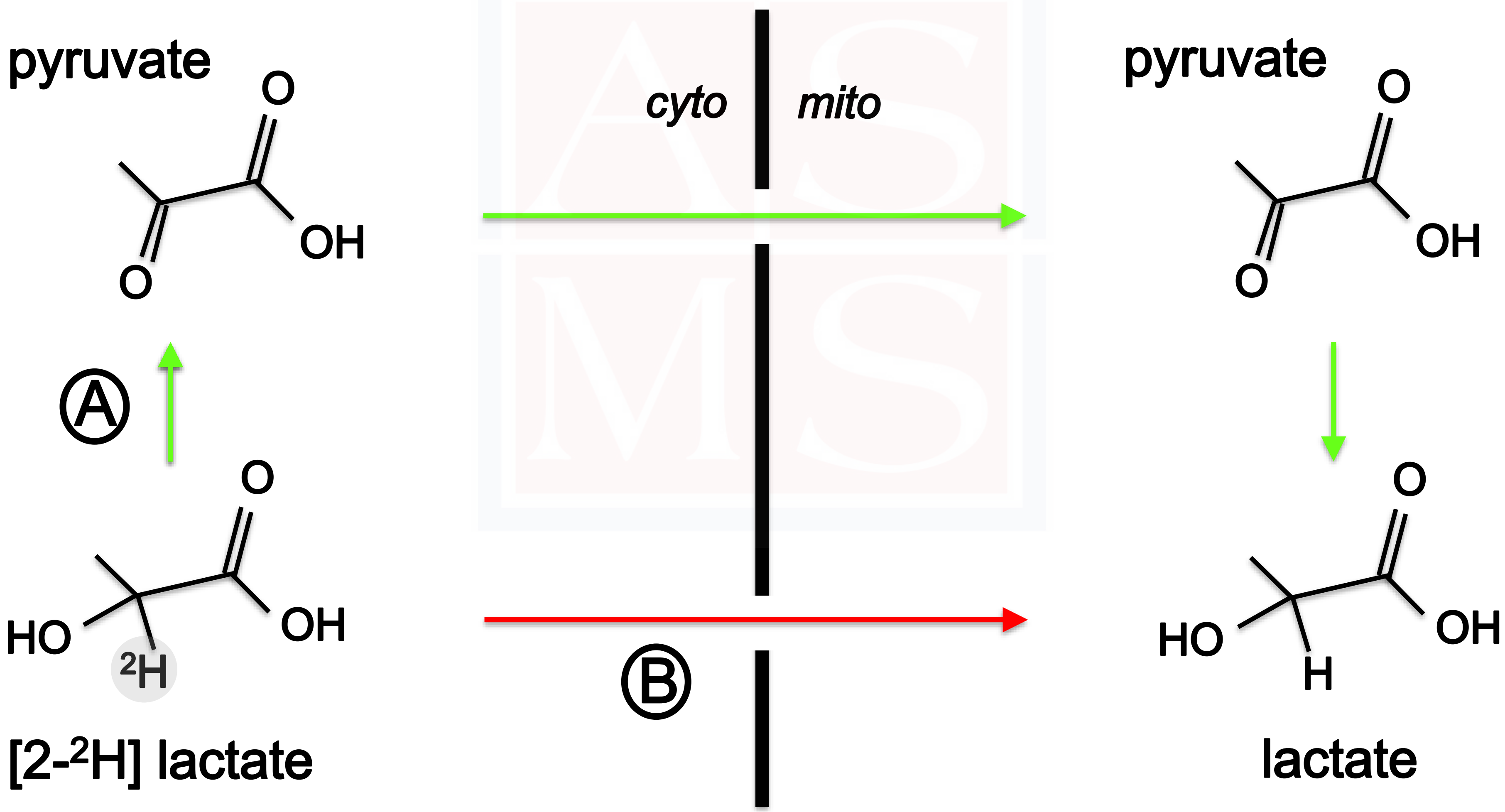
Application 2: lactate



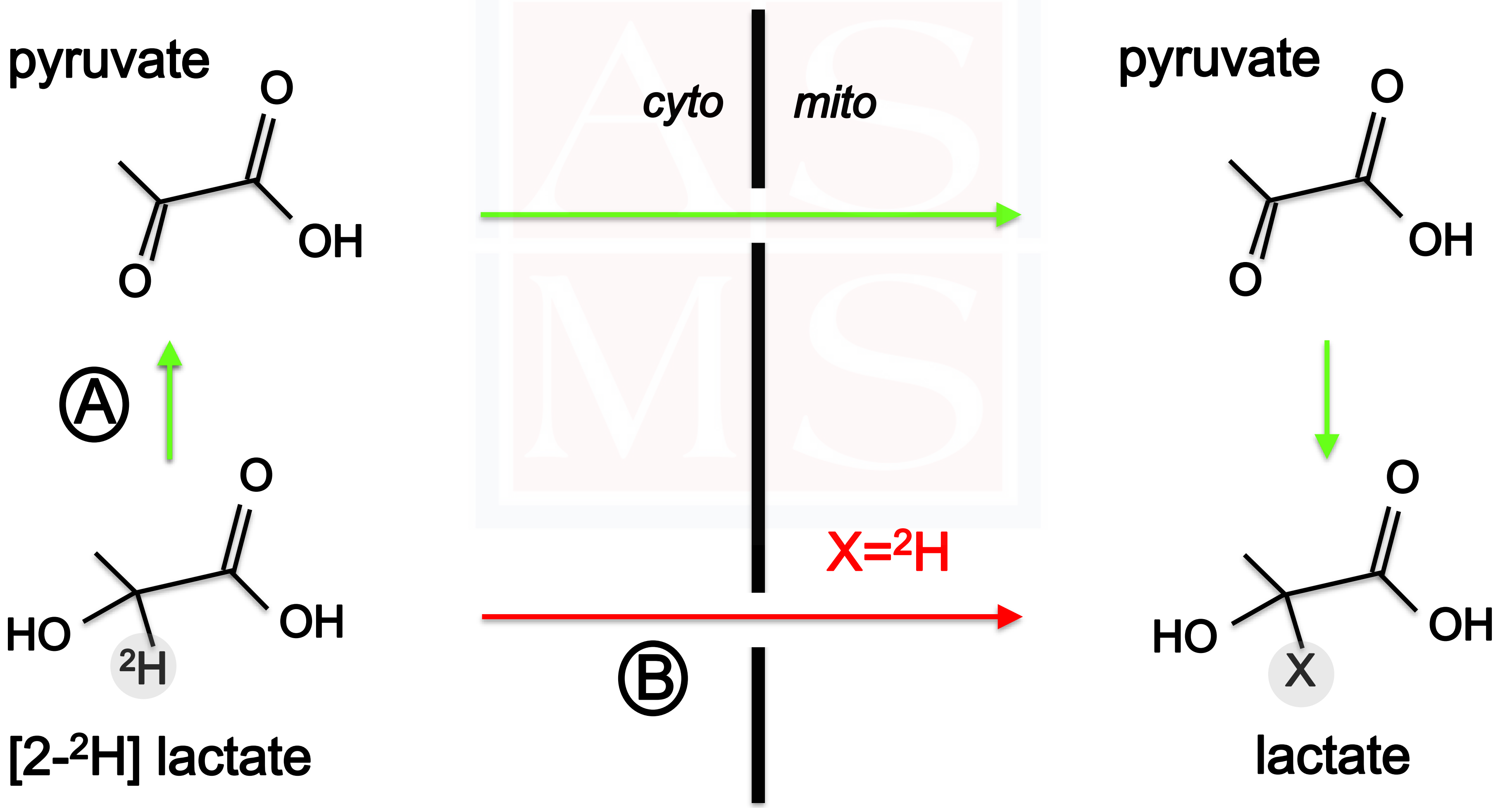
Application 2: lactate



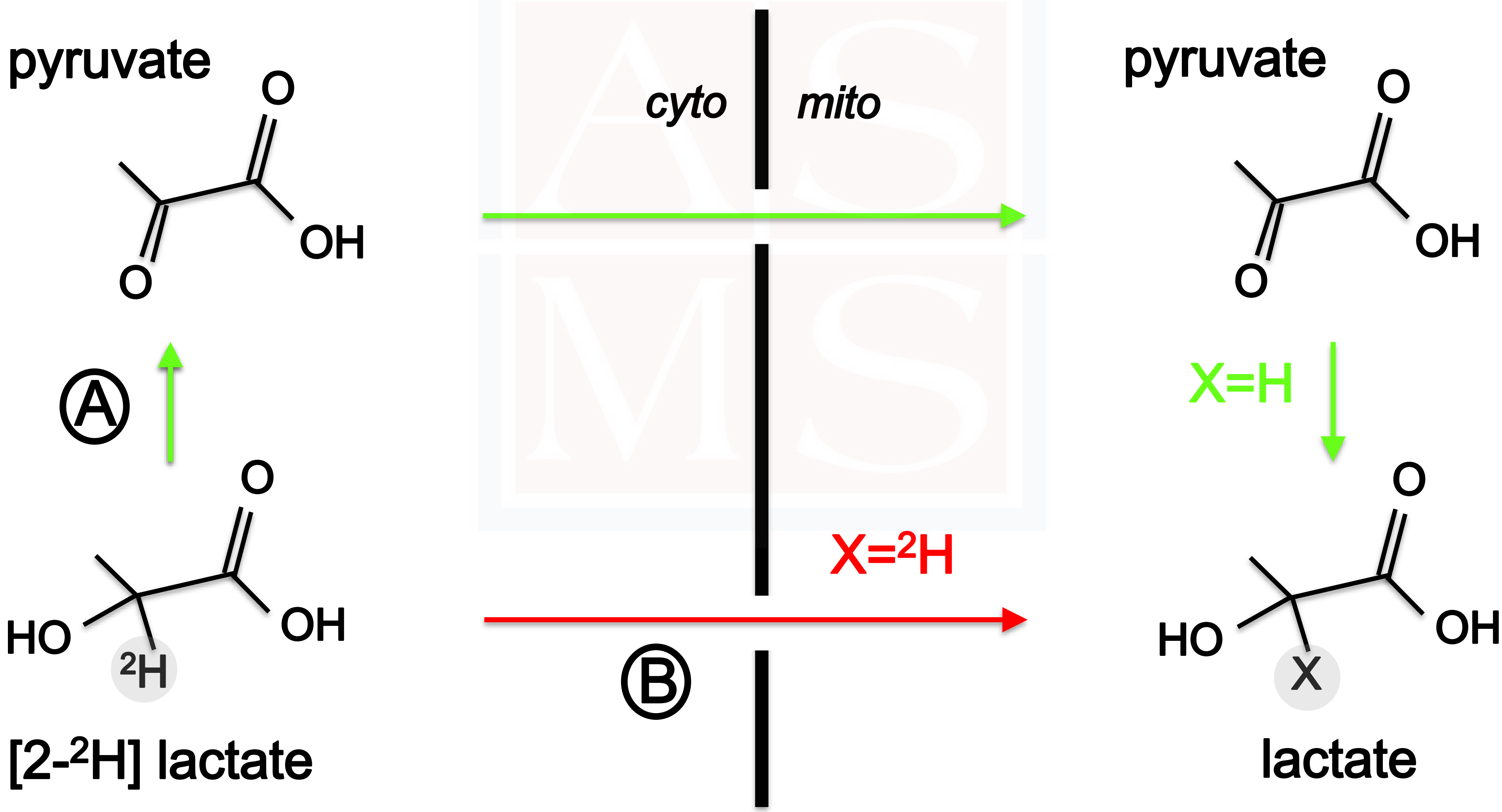
Application 2: lactate



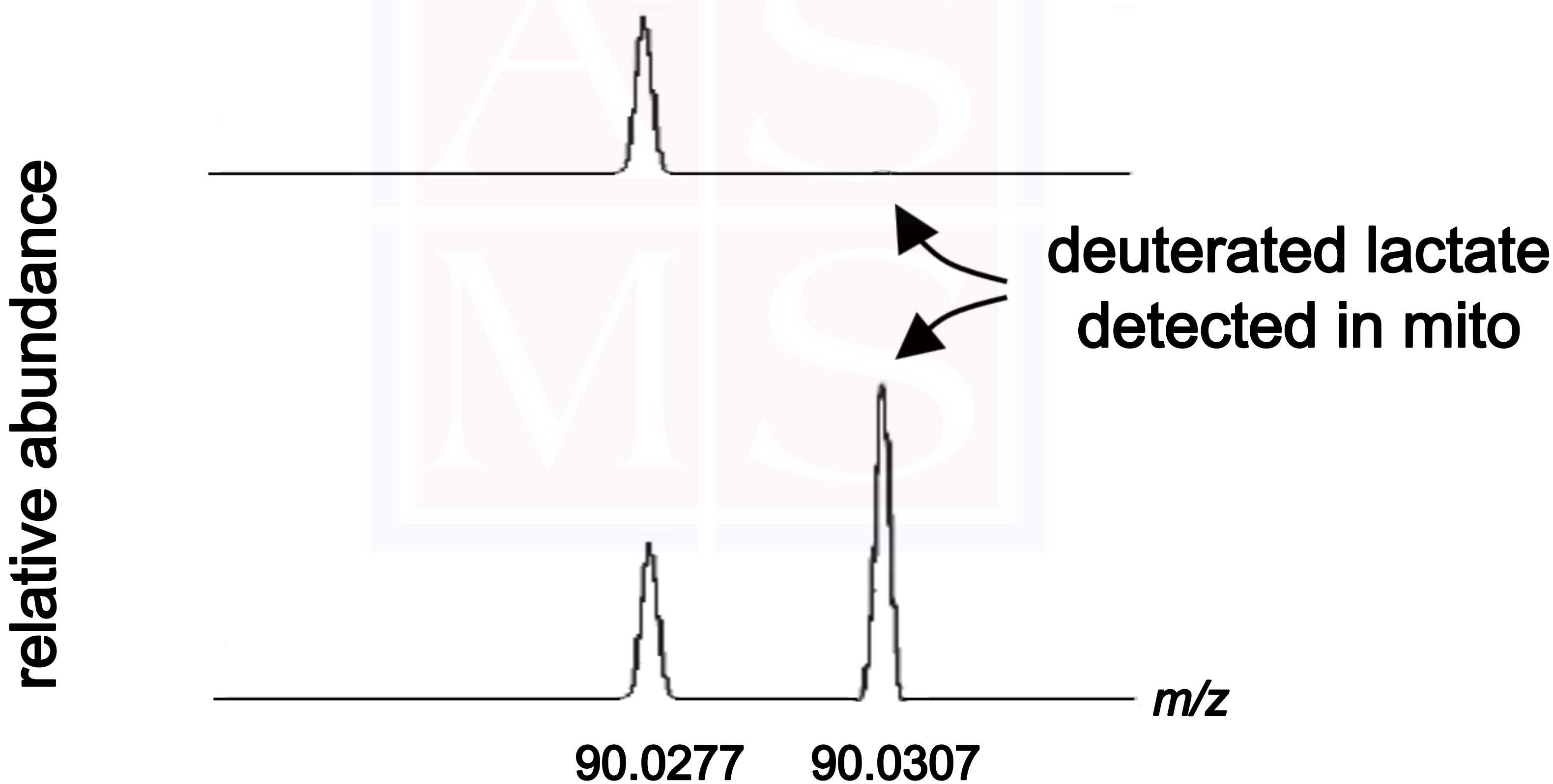
Application 2: lactate



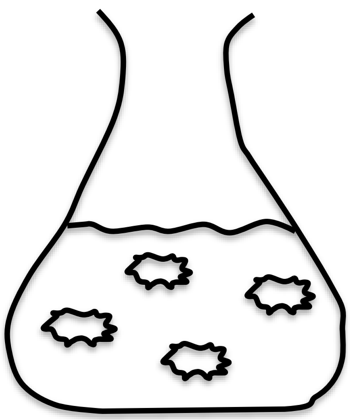
Application 2: lactate



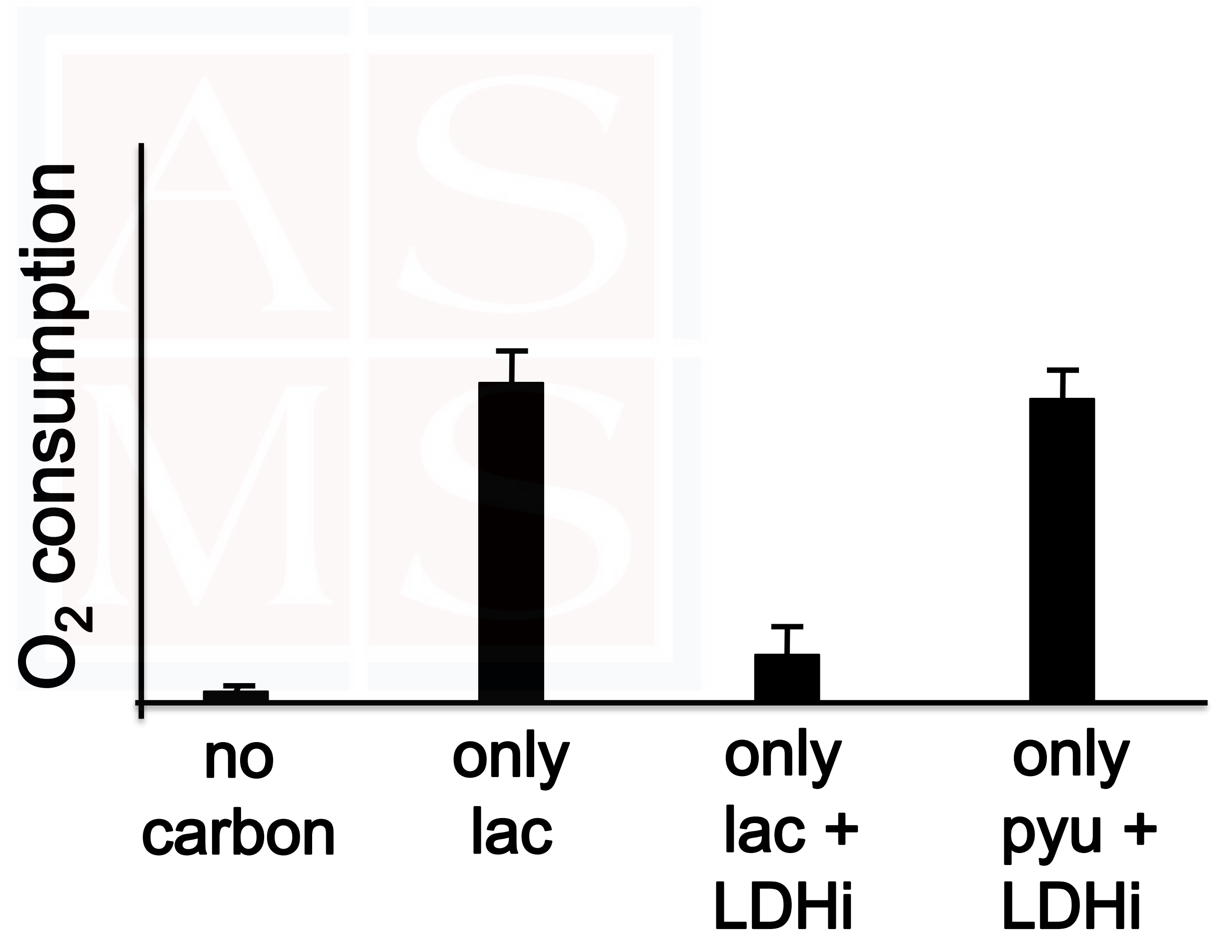
Application 2: lactate



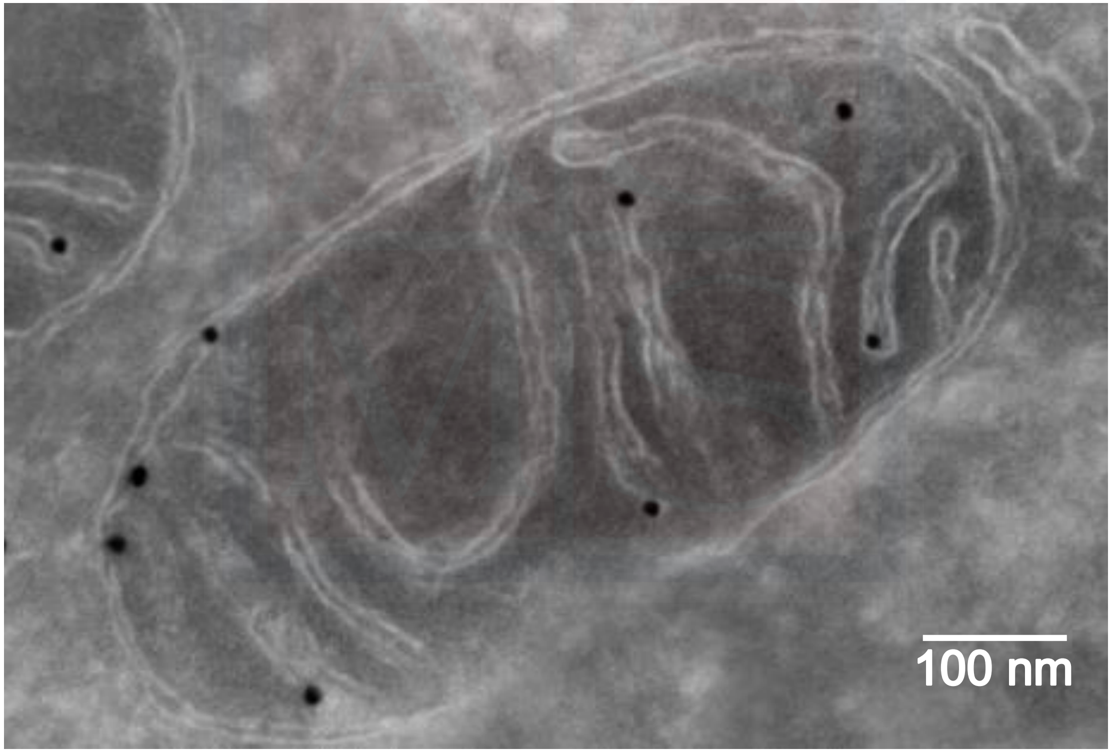
Application 2: lactate



isolated mitochondria

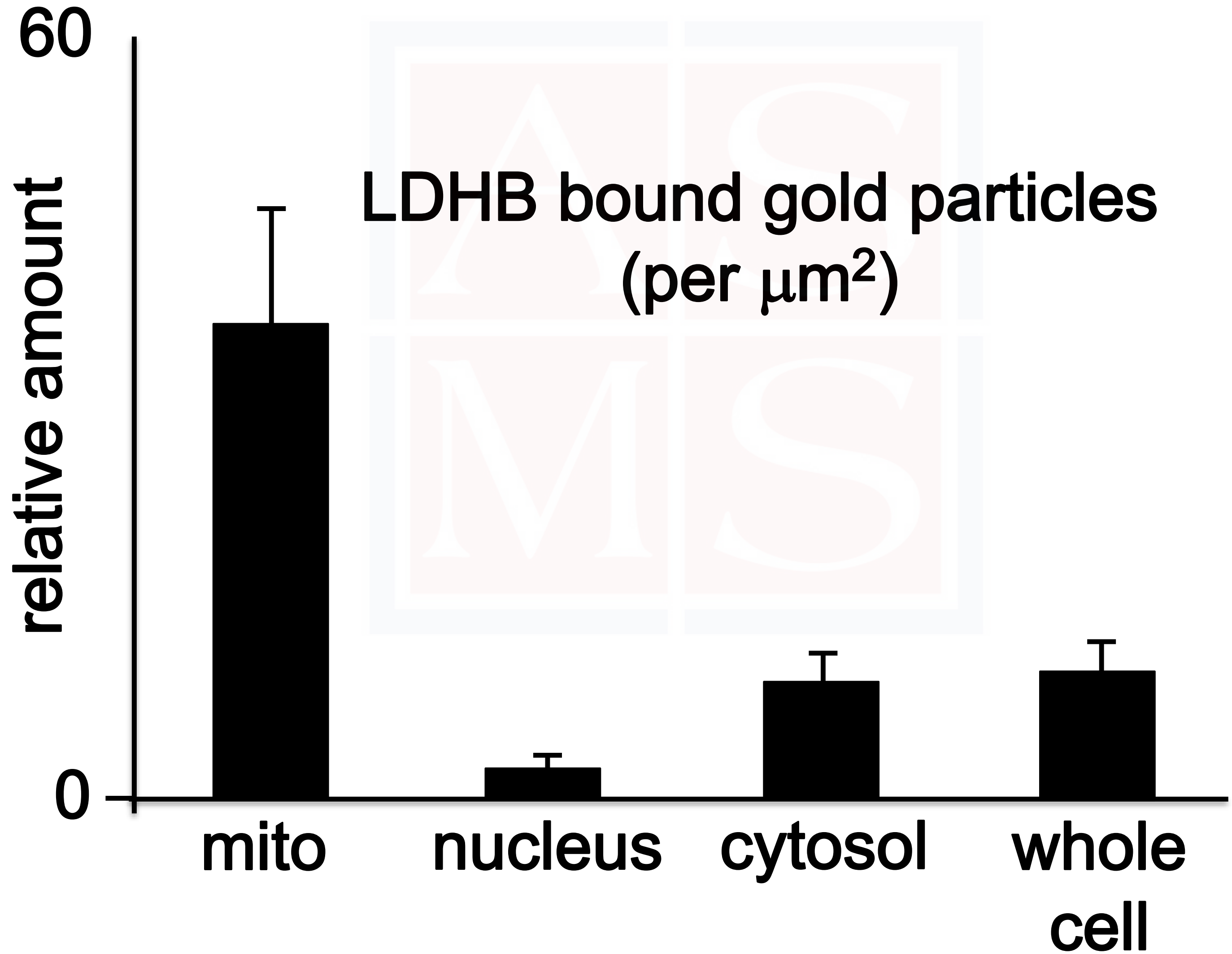


<u>Application 2:</u> lactate

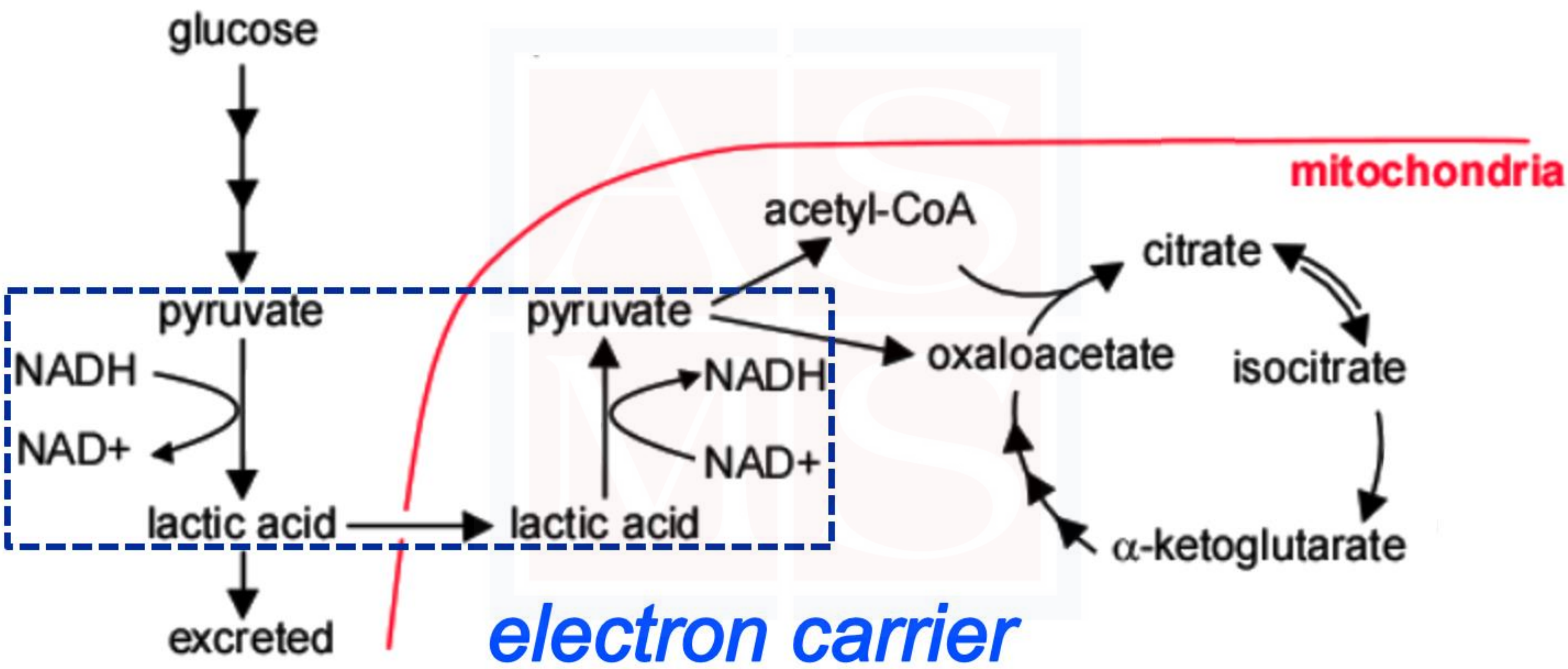


EM of LDHB

Application 2: lactate



Application 2: lactate



ASMS Metabolomics Short Course

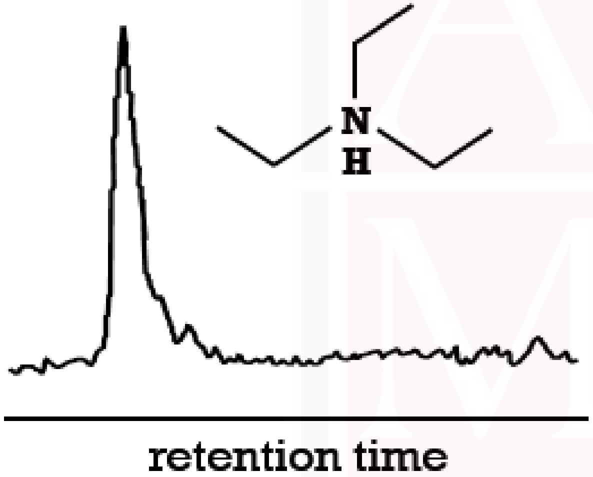
- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

Advanced workflows

Common experimental practices

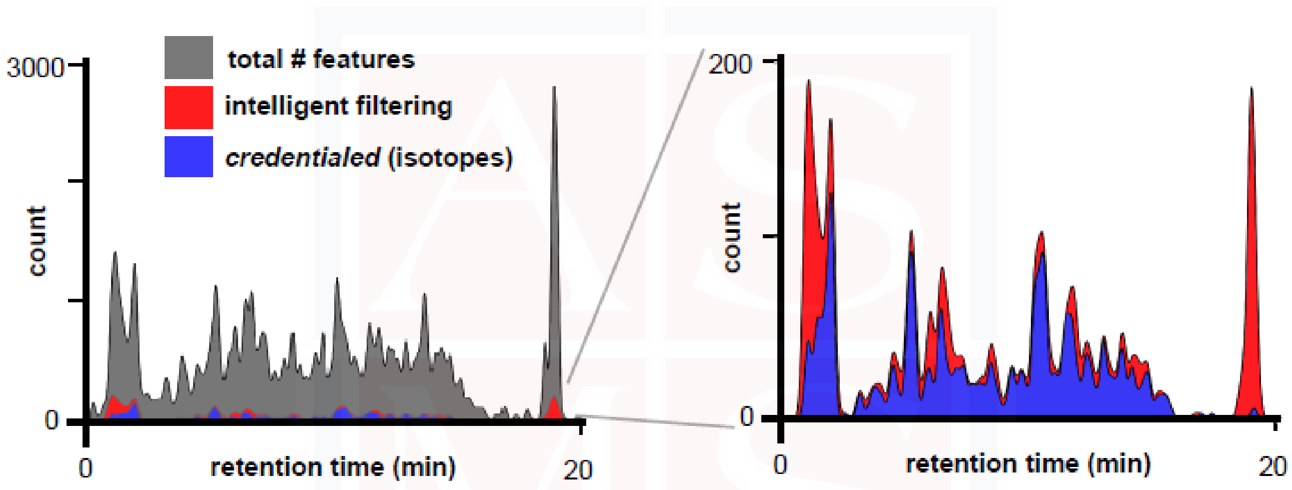
BLANK SUBTRACTION



chemical contaminants are prevalent in data

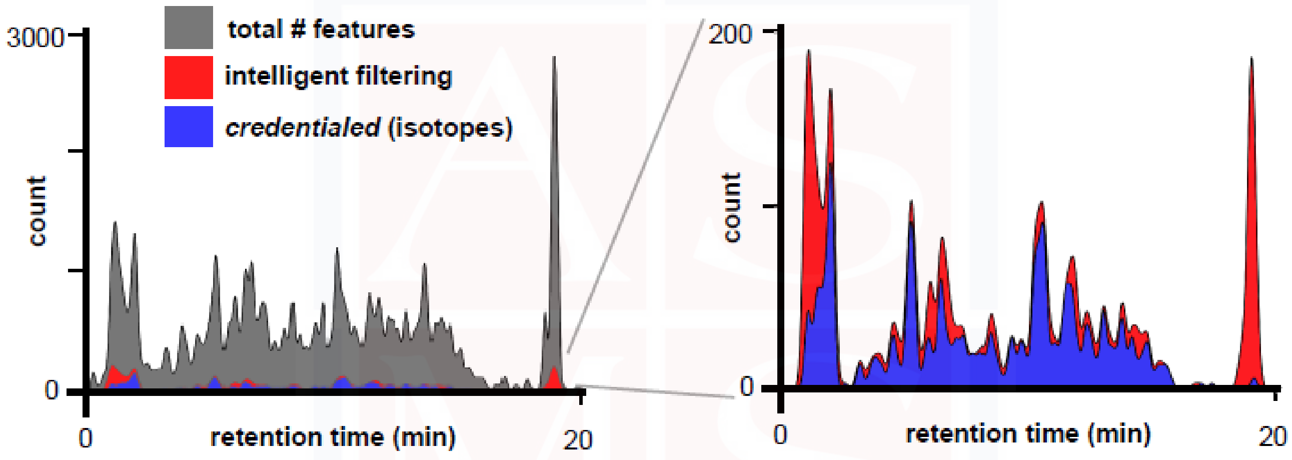
Common experimental practices

BLANK SUBTRACTION



Common experimental practices

BLANK SUBTRACTION



Common experimental practices

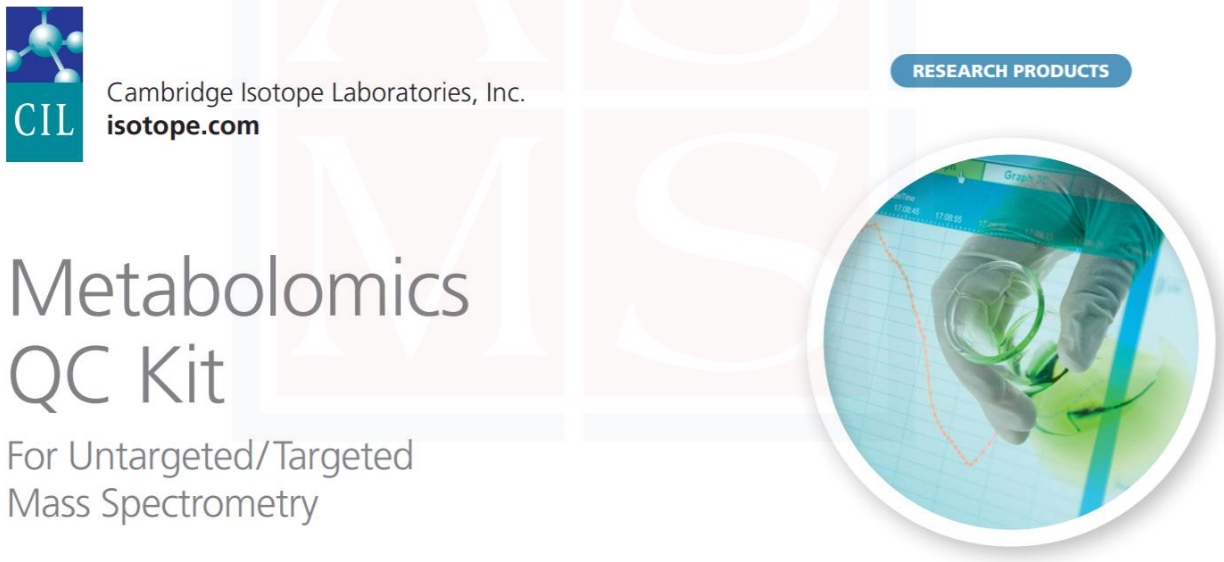
QUALITY-CONTROL SAMPLES

- *Reference material (plasma)*
- *Pooled samples*

Common experimental practices

INTERNAL STANDARDS

- *Often spiked into samples*
- *Typically isotopically labeled*



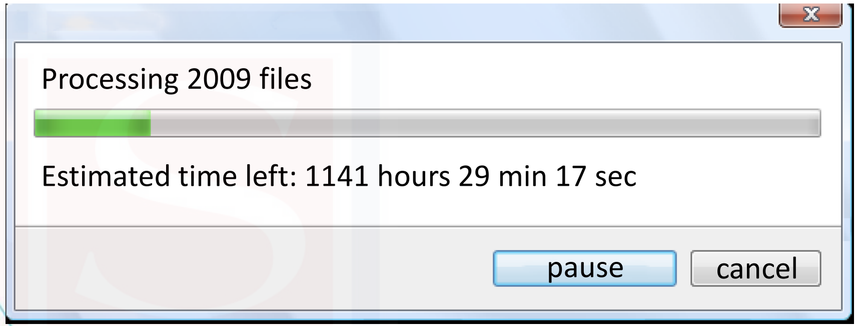
Advanced workflows for metabolomics

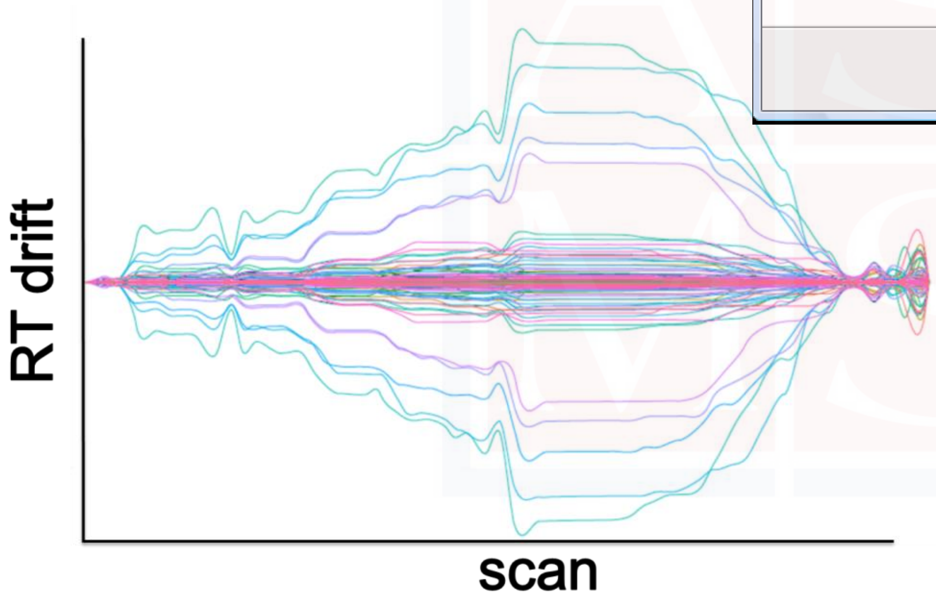
1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

Advanced workflows for metabolomics

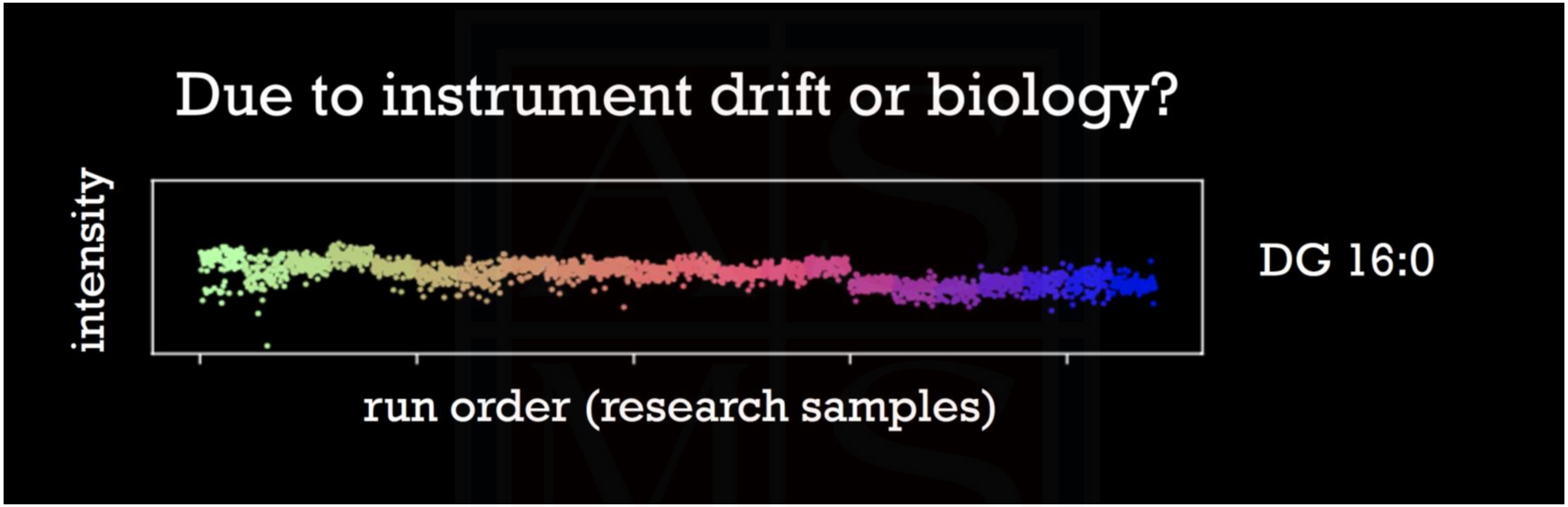
1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

1. Large-scale analyses

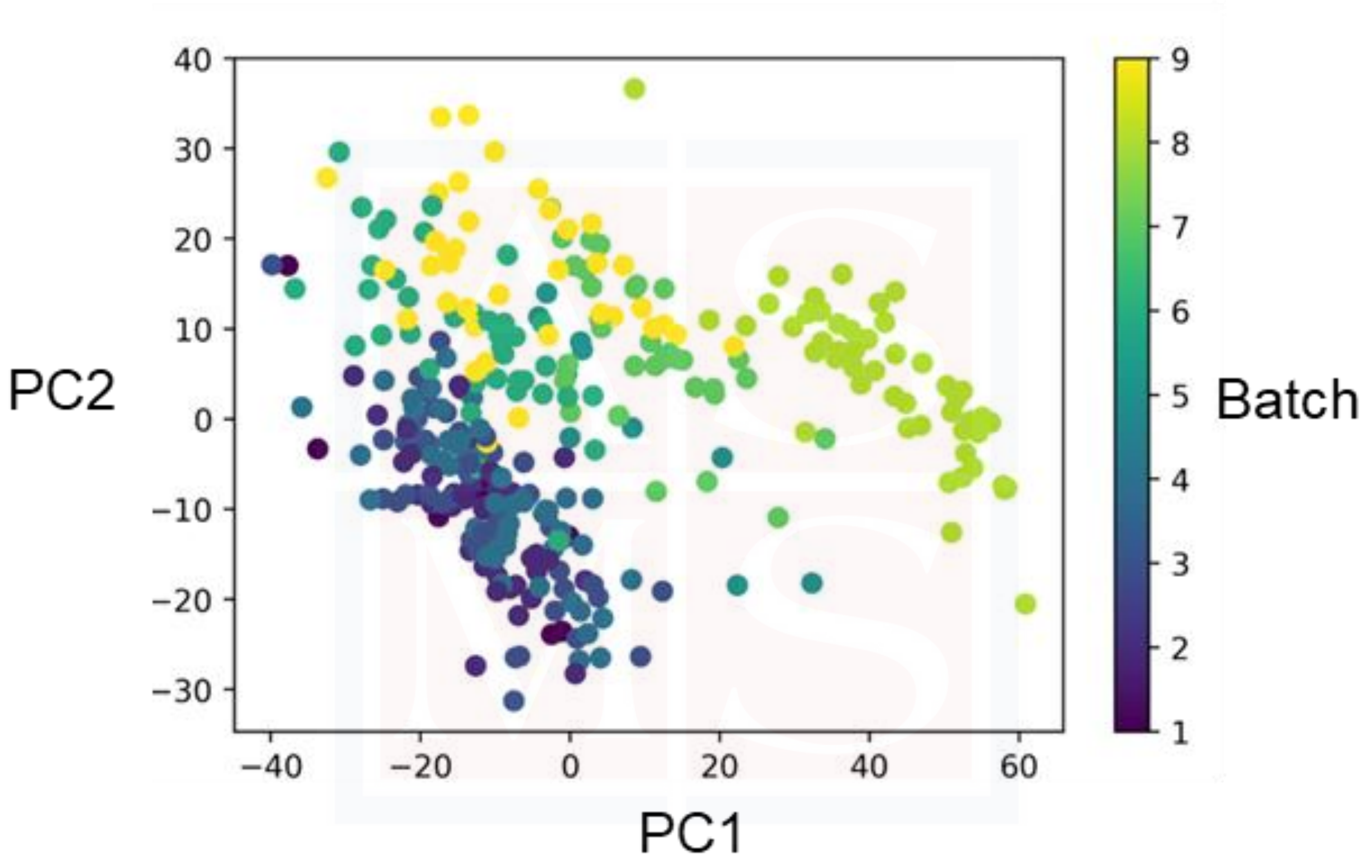




1. Large-scale analyses



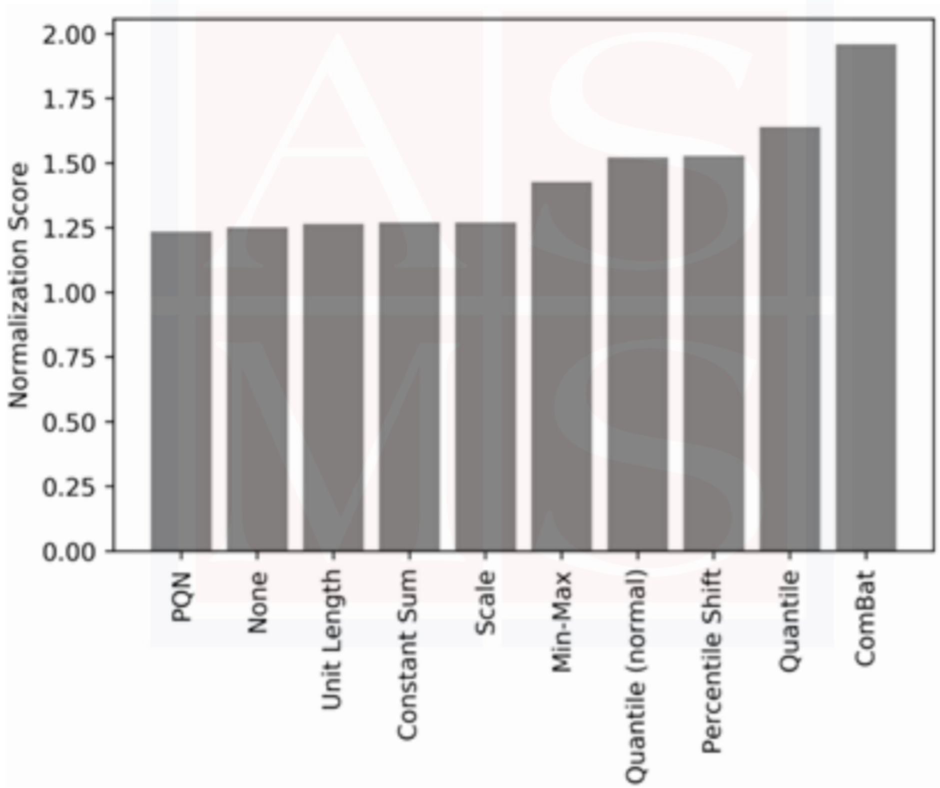
1. Large-scale analyses



differences due to analytical variability instead of biological variability

1. Large-scale analyses

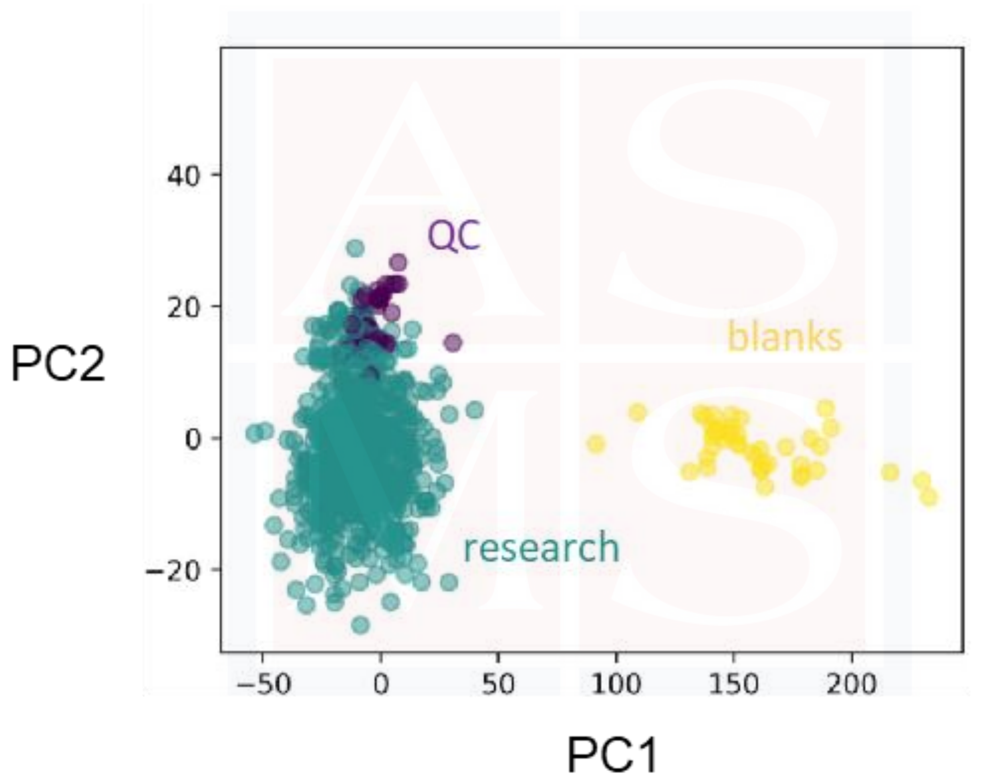
often requires batch correction



Sindelar, Stancliffe, et al., Cell Reports Medicine, 2021

1. Large-scale analyses

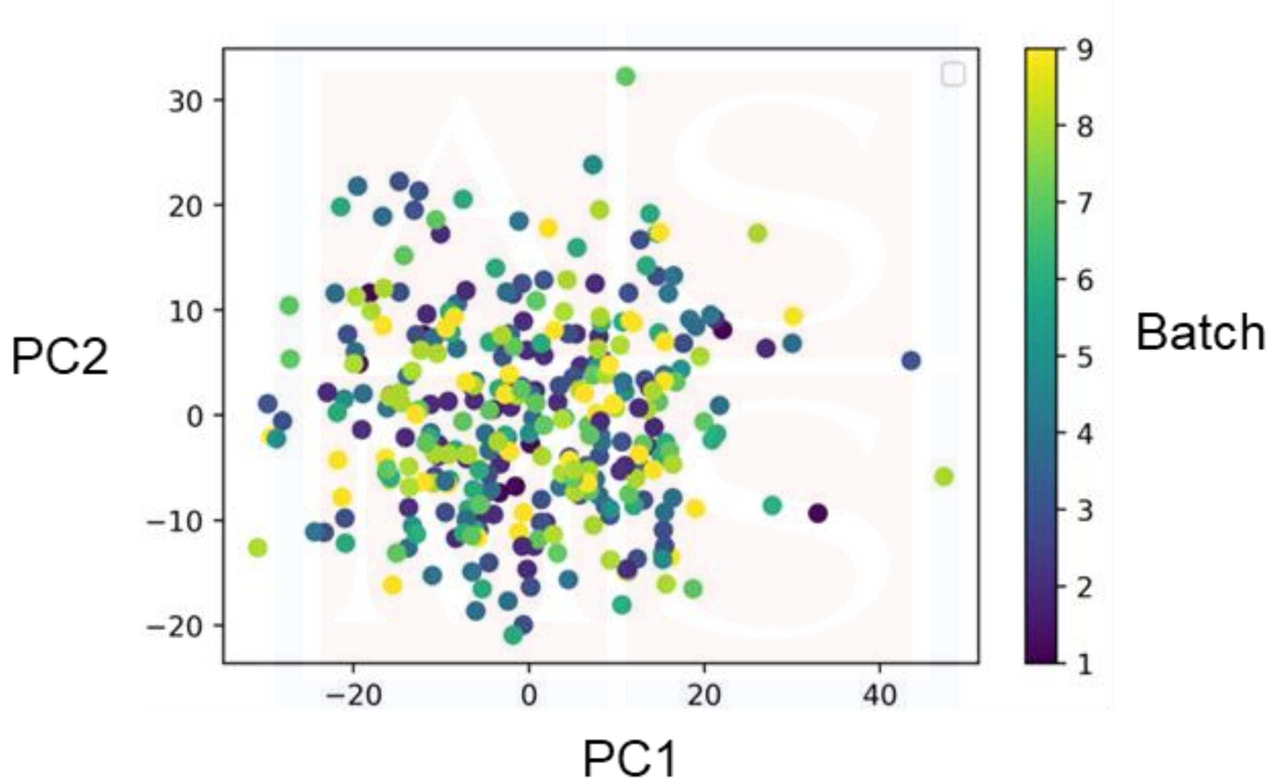
after batch correction



Sindelar, Stancliffe, et al., Cell Reports Medicine, 2021

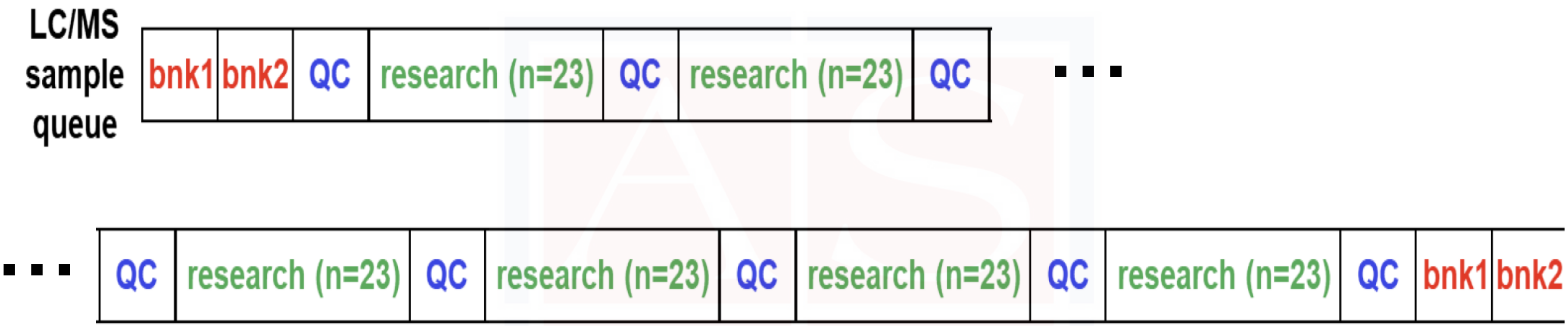
1. Large-scale analyses

after batch correction



Sindelar, Stancliffe, et al., Cell Reports Medicine, 2021

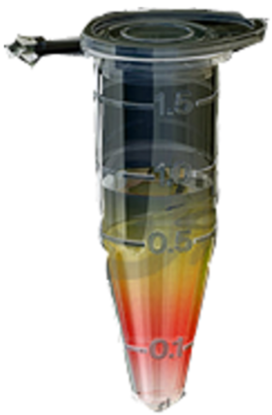
1. Large-scale analyses

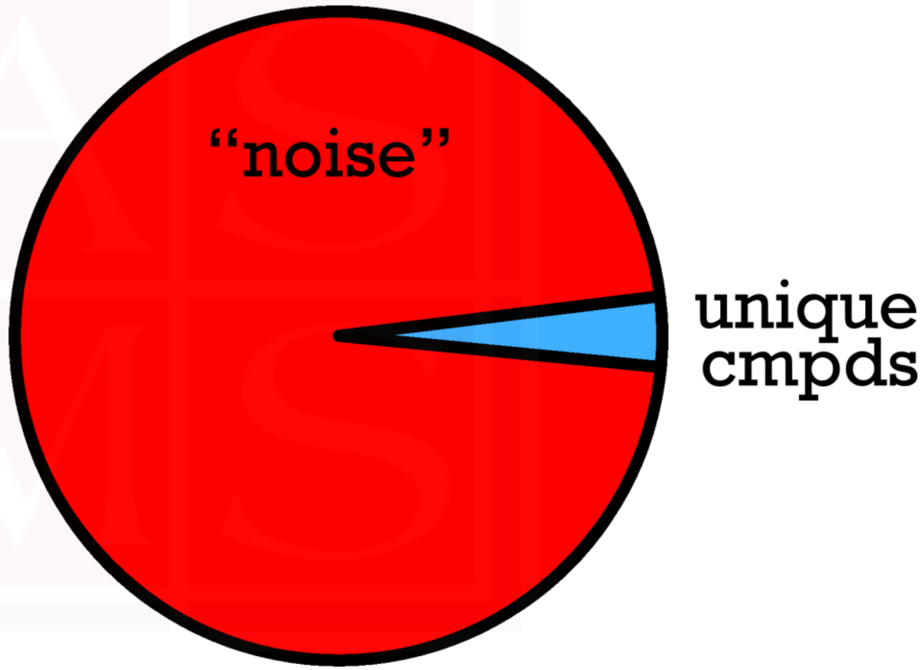


CV	# of metabolites
<2%	20
2-5%	417
5-10%	179
10-15%	76
15-20%	10
20-25%	5
>30%	0
Total	707

1. Large-scale analyses

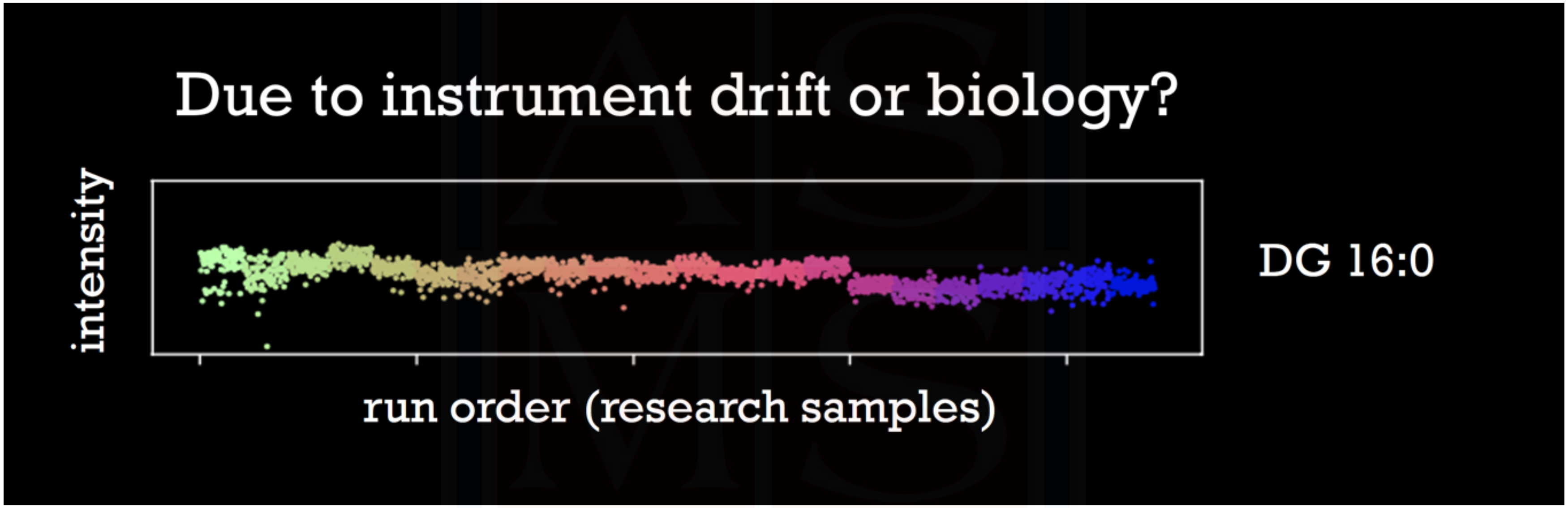
pooled sample





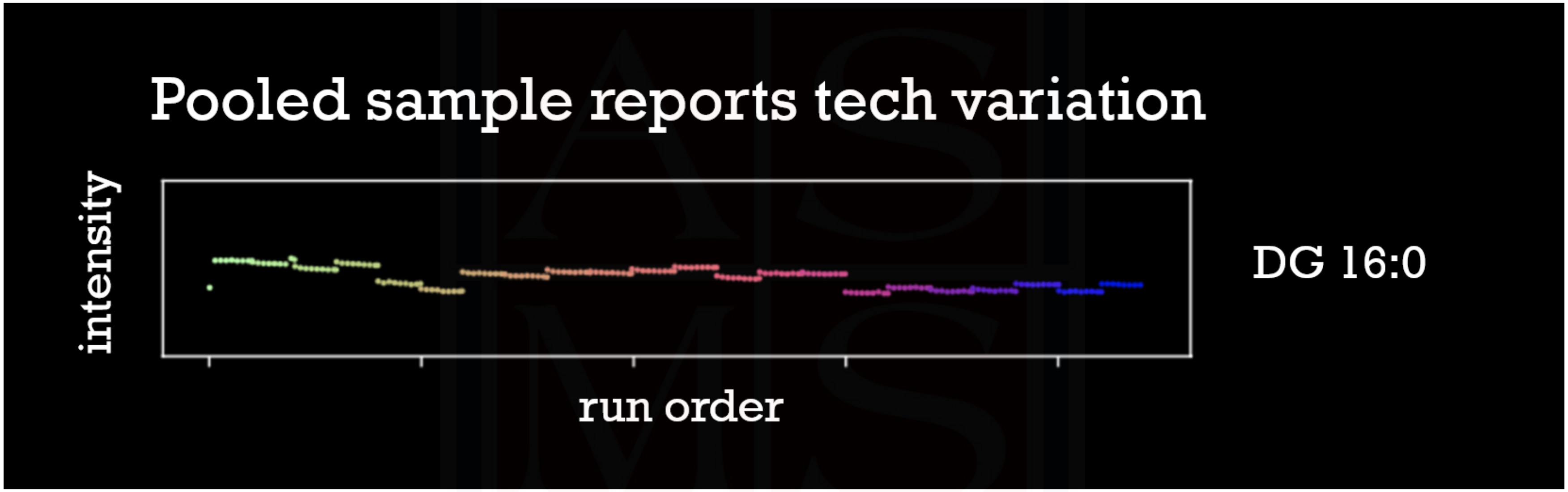
Stancliffe, Schwaiger-Haber et al., Anal Chem, 2022

1. Large-scale analyses



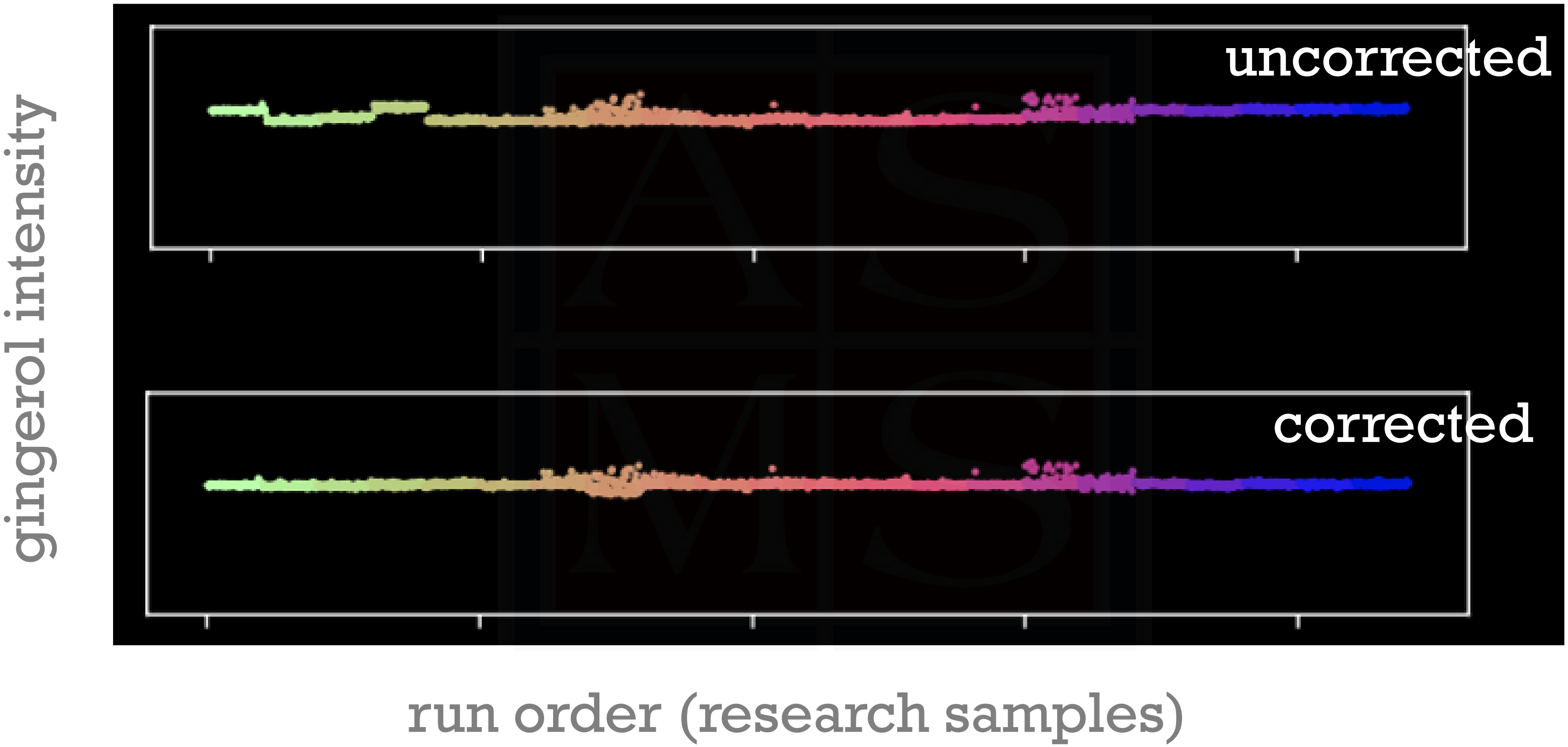
Stancliffe, Schwaiger-Haber et al., Anal Chem, 2022

1. Large-scale analyses



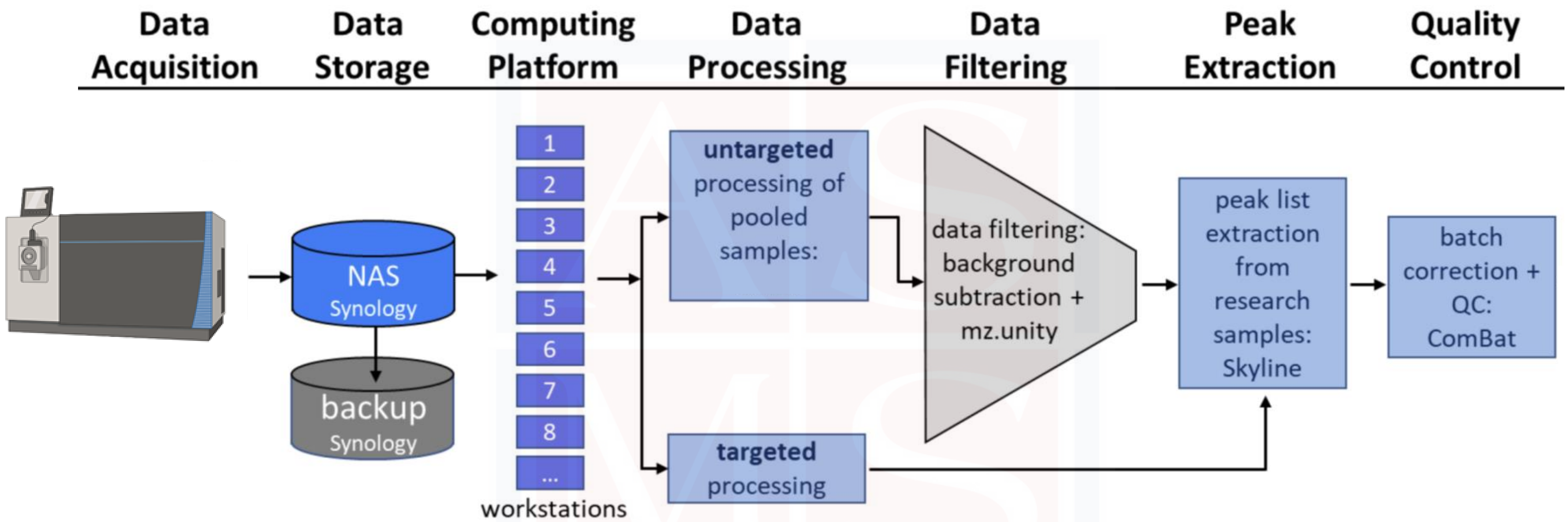
Stancliffe, Schwaiger-Haber et al., Anal Chem, 2022

1. Large-scale analyses



Stancliffe, Schwaiger-Haber et al., Anal Chem, 2022

1. Large-scale analyses



Advanced workflows for metabolomics

1. Large-scale analyses
   - *several hundred to thousands of samples*

2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*

3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*

4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

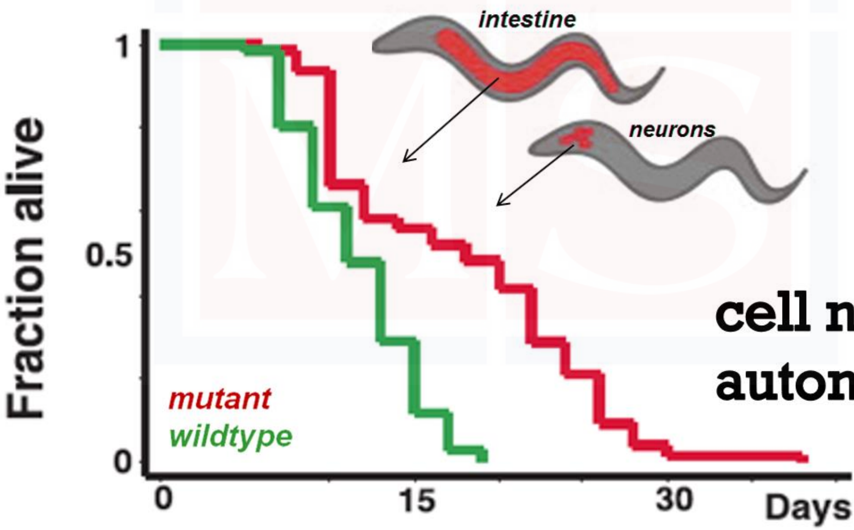
Advanced workflows for metabolomics

1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

2. Variations in experimental design

example: meta-analysis

metabolomics to find aging signals

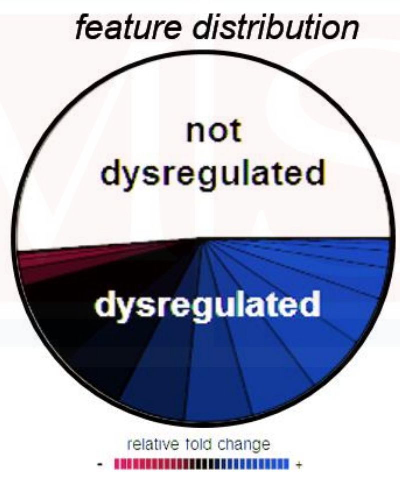


cell non-autonomous

2. Variations in experimental design

example: meta-analysis

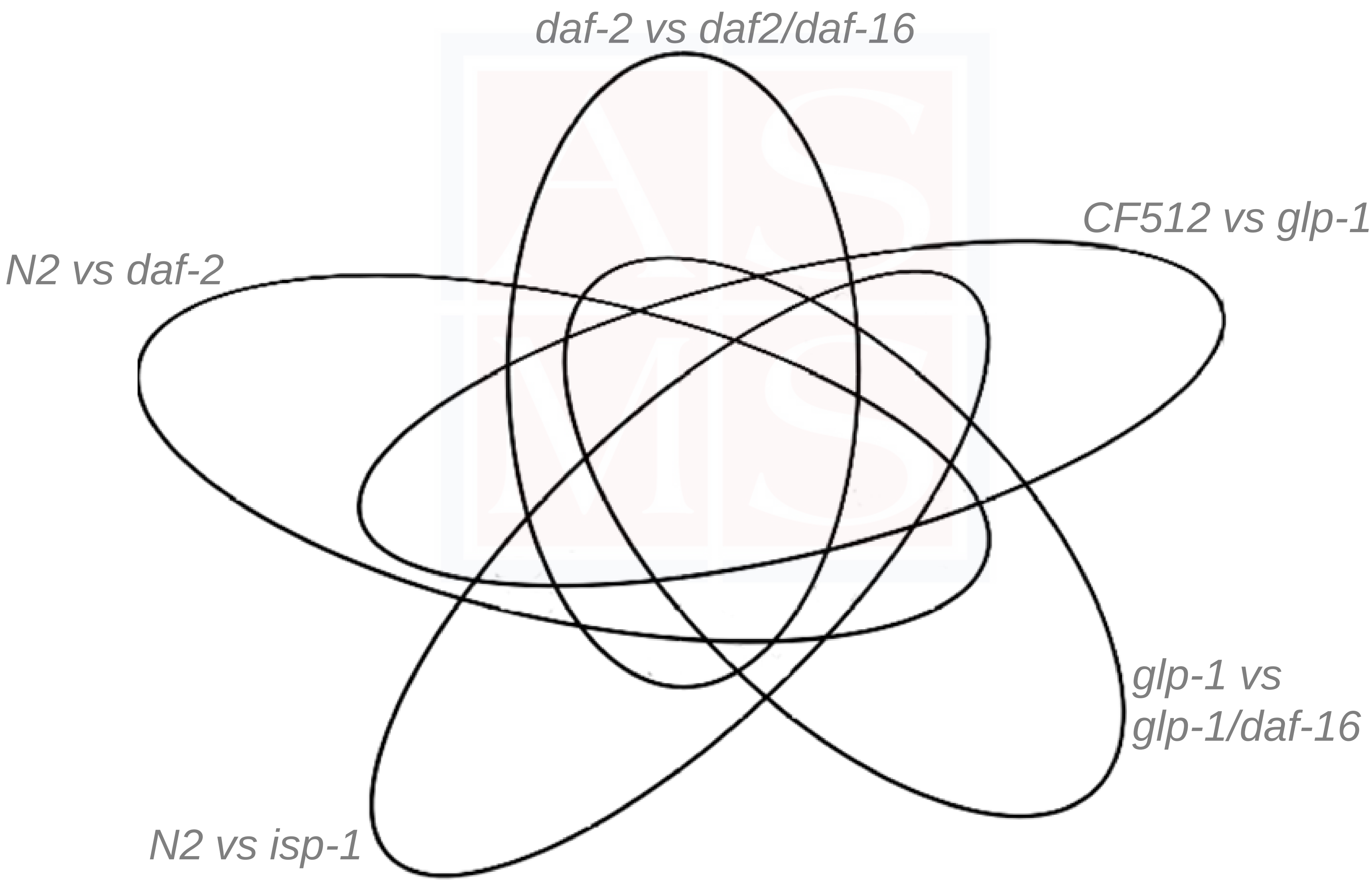
metabolomics to find aging signals



glp-1 *mutants vs.* CF512 *controls*

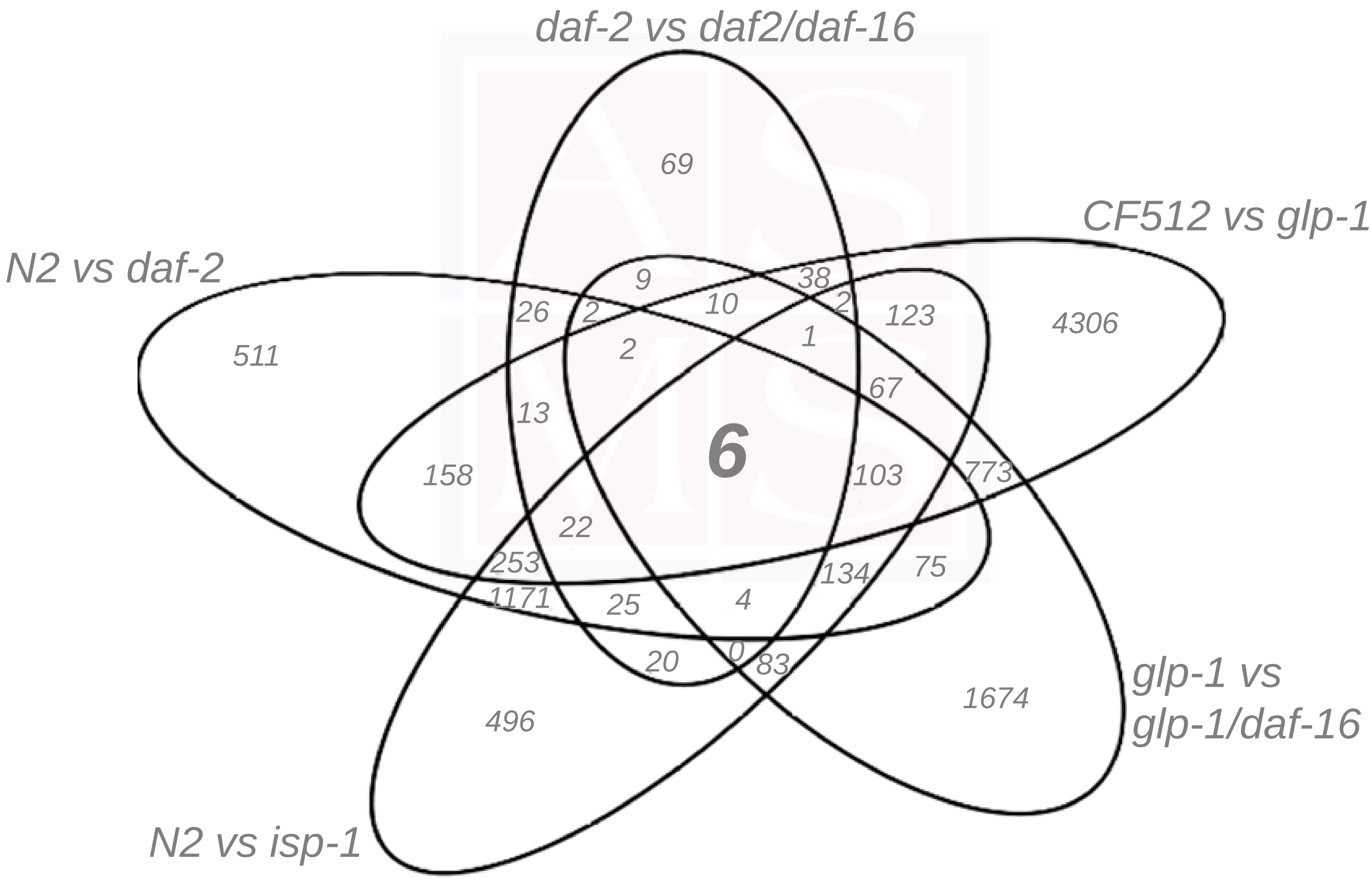
2. Variations in experimental design

example: meta-analysis



2. Variations in experimental design

example: meta-analysis



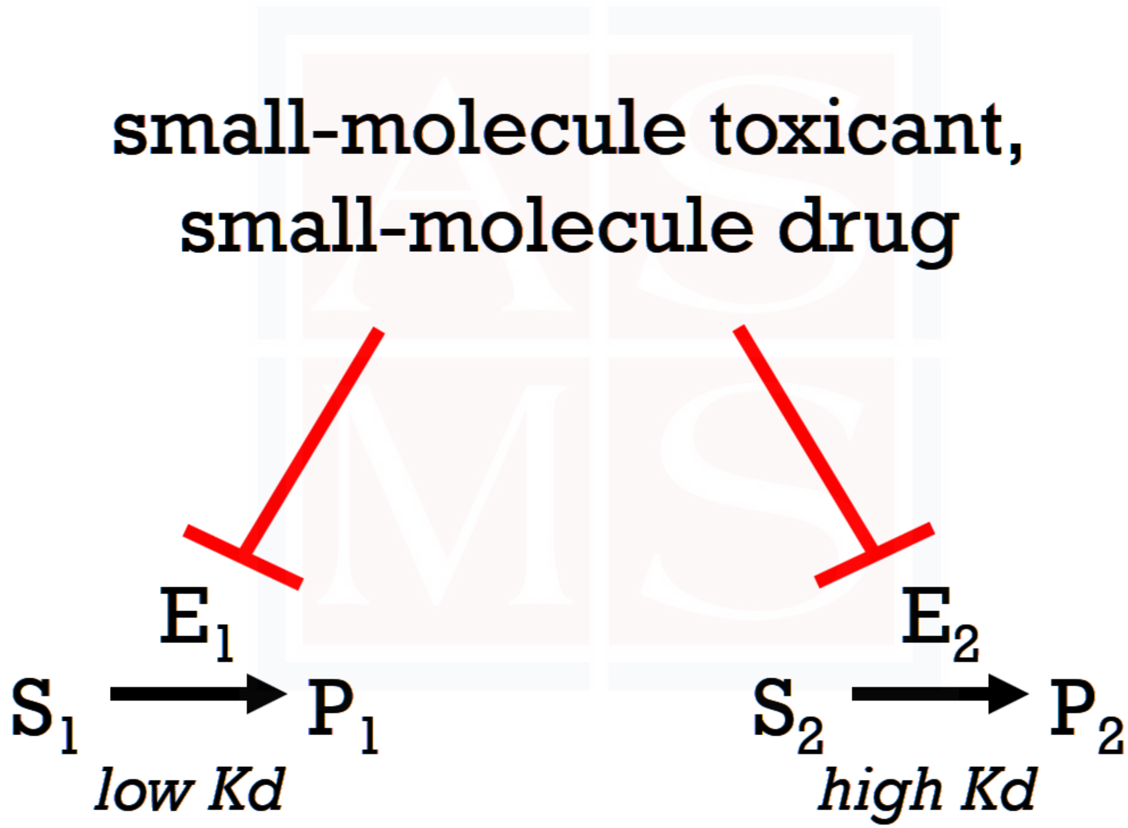
2. Variations in experimental design

2. Variations in experimental design

example: dose-response metabolomics

2. Variations in experimental design

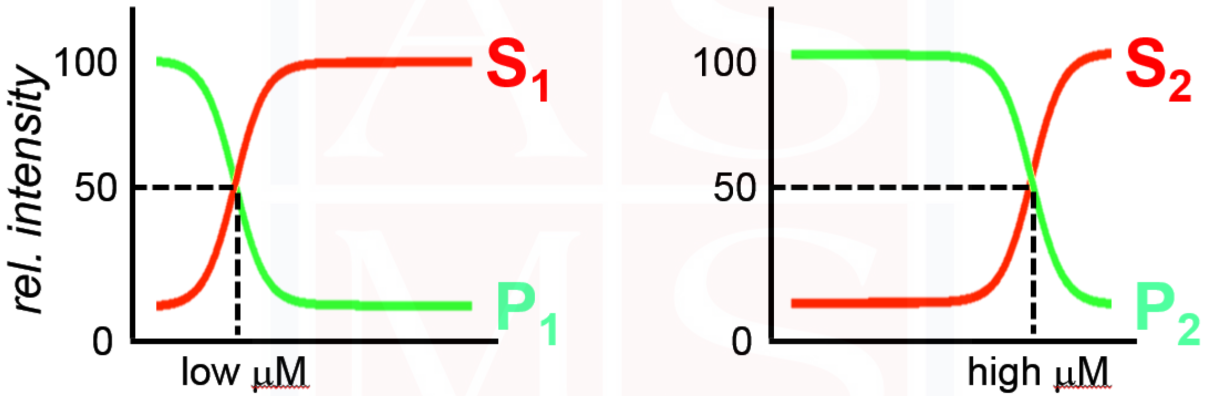
example: dose-response metabolomics



Mestres et al., Nat Biotech 2008, 26 (9)

2. Variations in experimental design

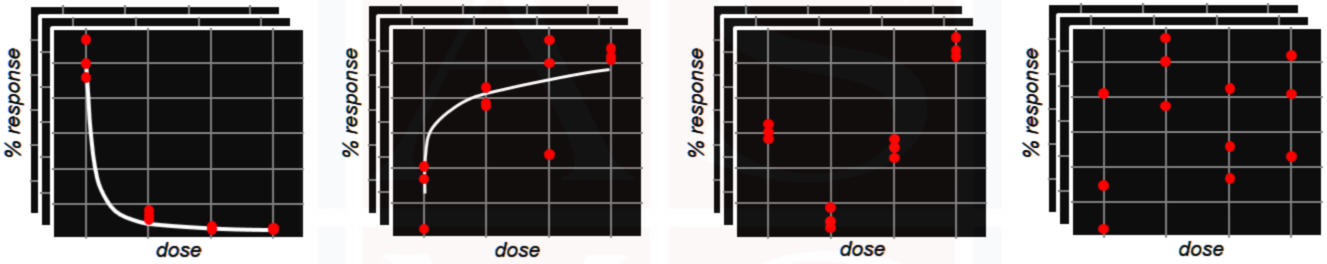
example: dose-response metabolomics



at high concentrations, target effects conflated

2. Variations in experimental design

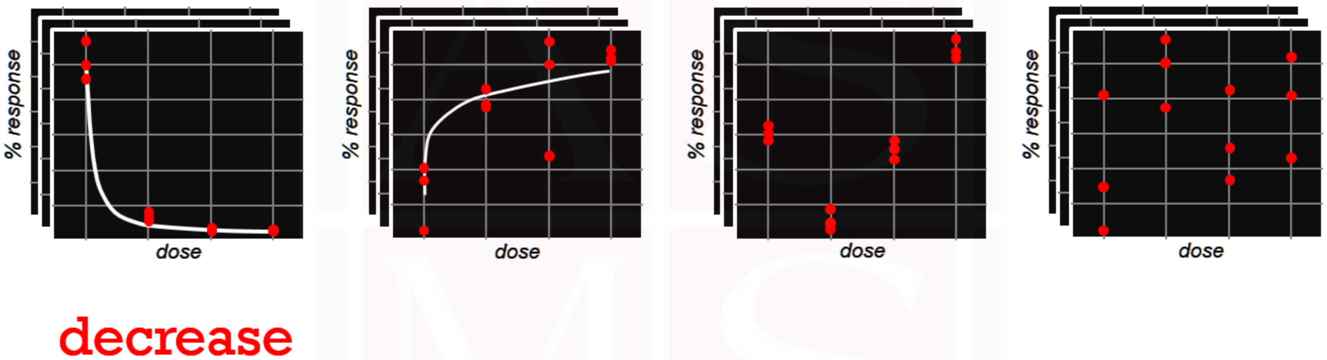
example: dose-response metabolomics



TOXcms: Yao et al, Anal Chem 2020

2. Variations in experimental design

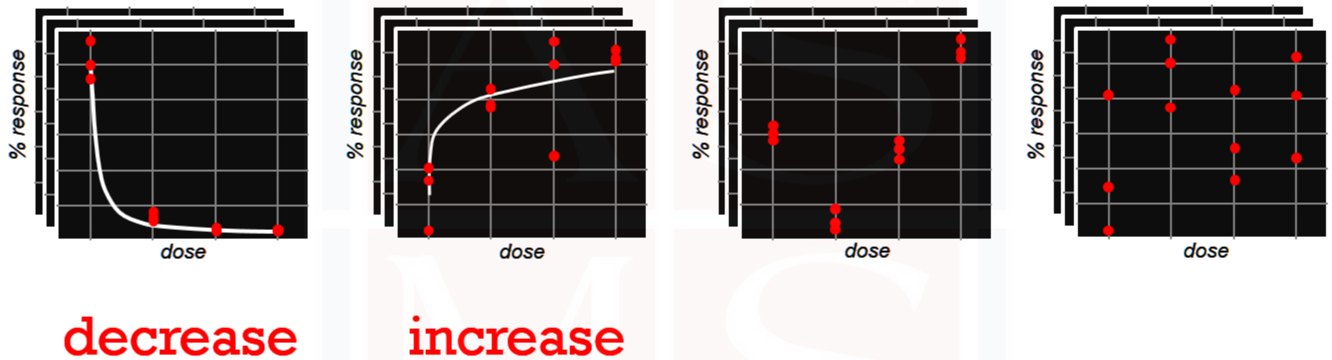
example: dose-response metabolomics



TOXcms: Yao et al, Anal Chem 2020

2. Variations in experimental design

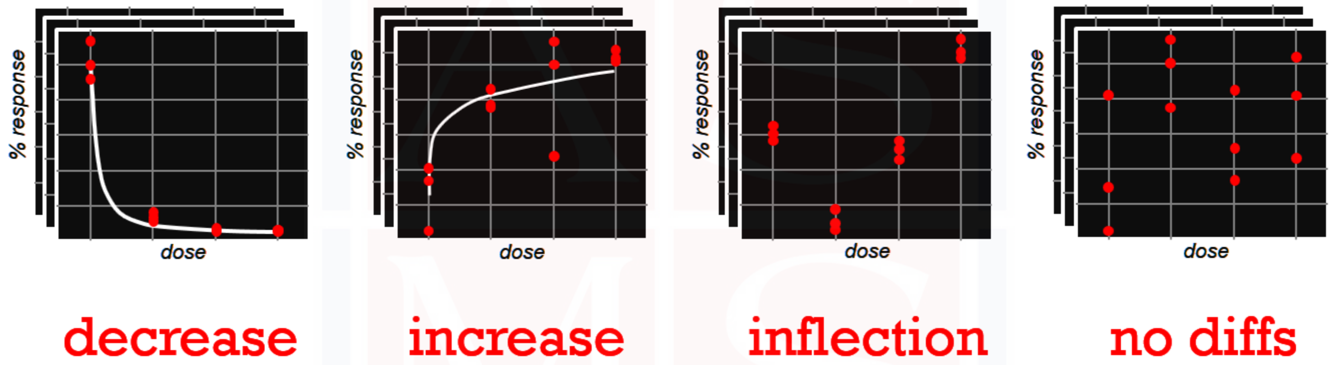
example: dose-response metabolomics



TOXcms: Yao et al, Anal Chem 2020

2. Variations in experimental design

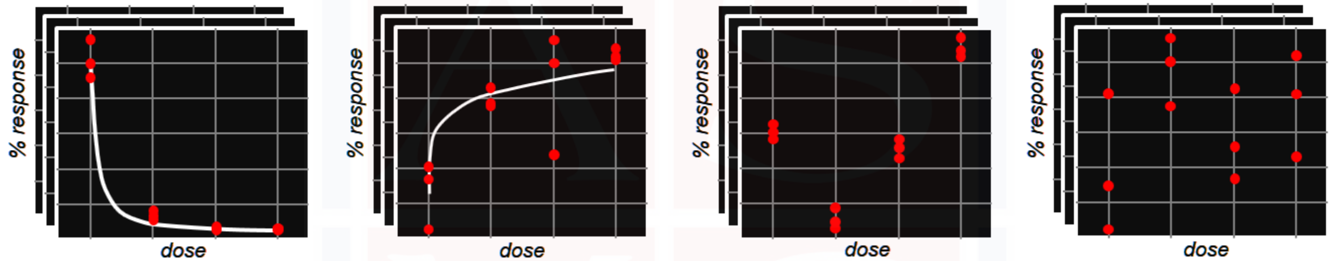
example: dose-response metabolomics



TOXcms: Yao et al, Anal Chem 2020

2. Variations in experimental design

example: dose-response metabolomics



decrease increase inflection no diffs

ED_{50} calculated

TOXcms: Yao et al, Anal Chem 2020

2. Variations in experimental design

<u>Application: etomoxir</u>

clinical trial: 2007

heart disease, psoriasis, cancer

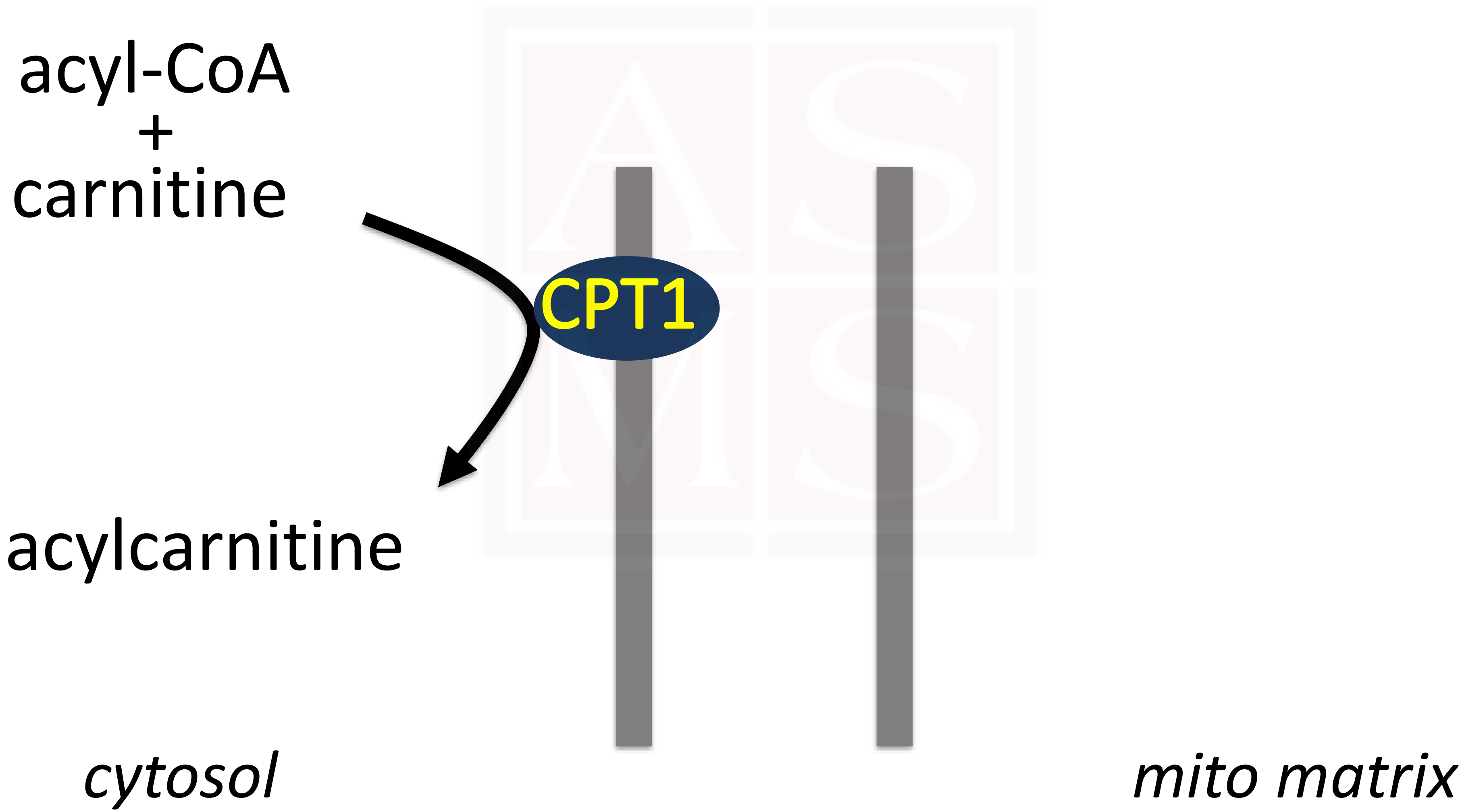
>10,000 publications

2. Variations in experimental design

expected mechanism of etomoxir

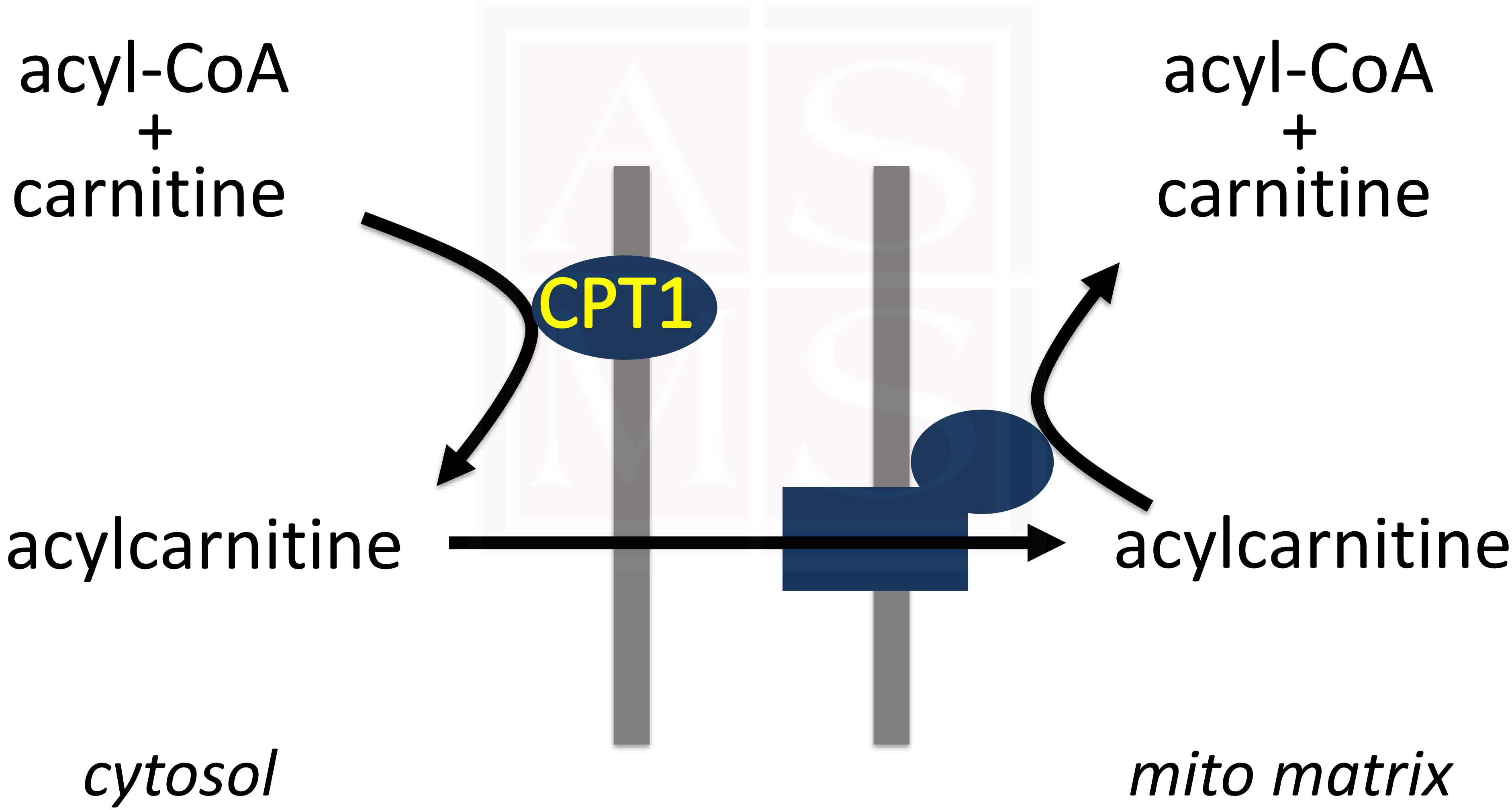
2. Variations in experimental design

expected mechanism of etomoxir



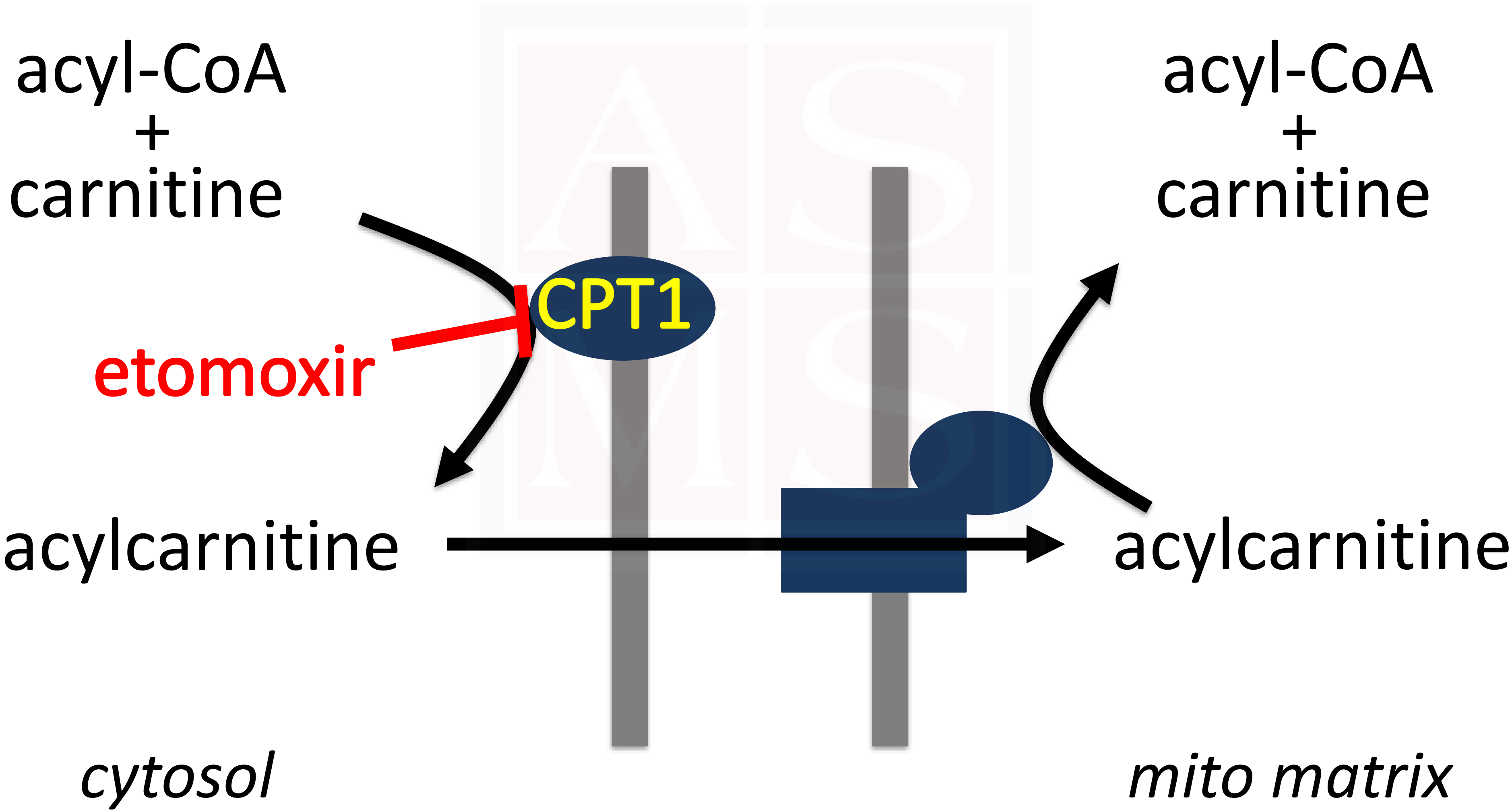
2. Variations in experimental design

expected mechanism of etomoxir



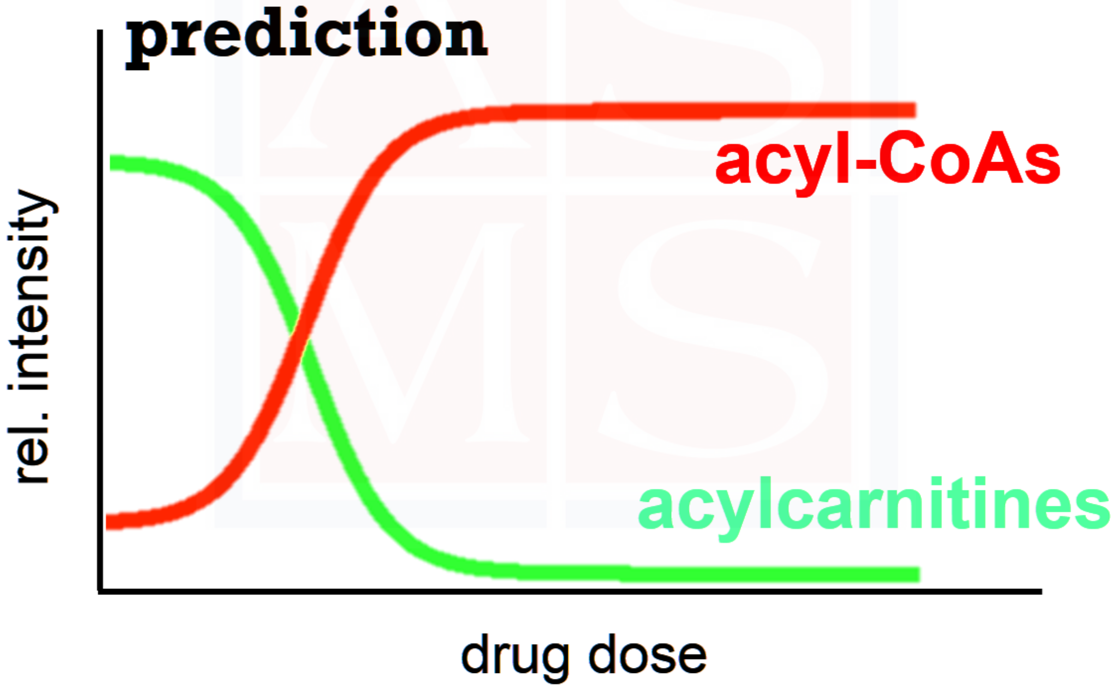
2. Variations in experimental design

expected mechanism of etomoxir



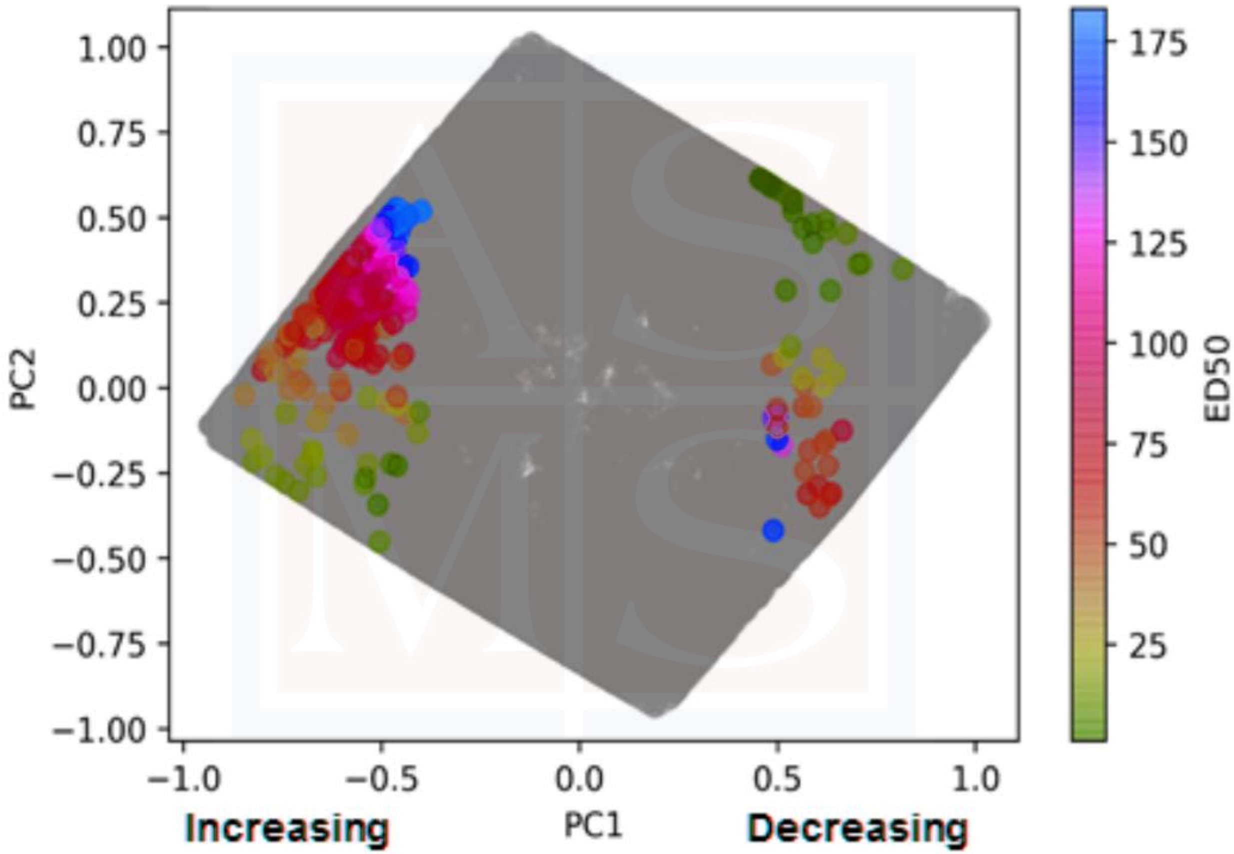
2. Variations in experimental design

expected mechanism of etomoxir



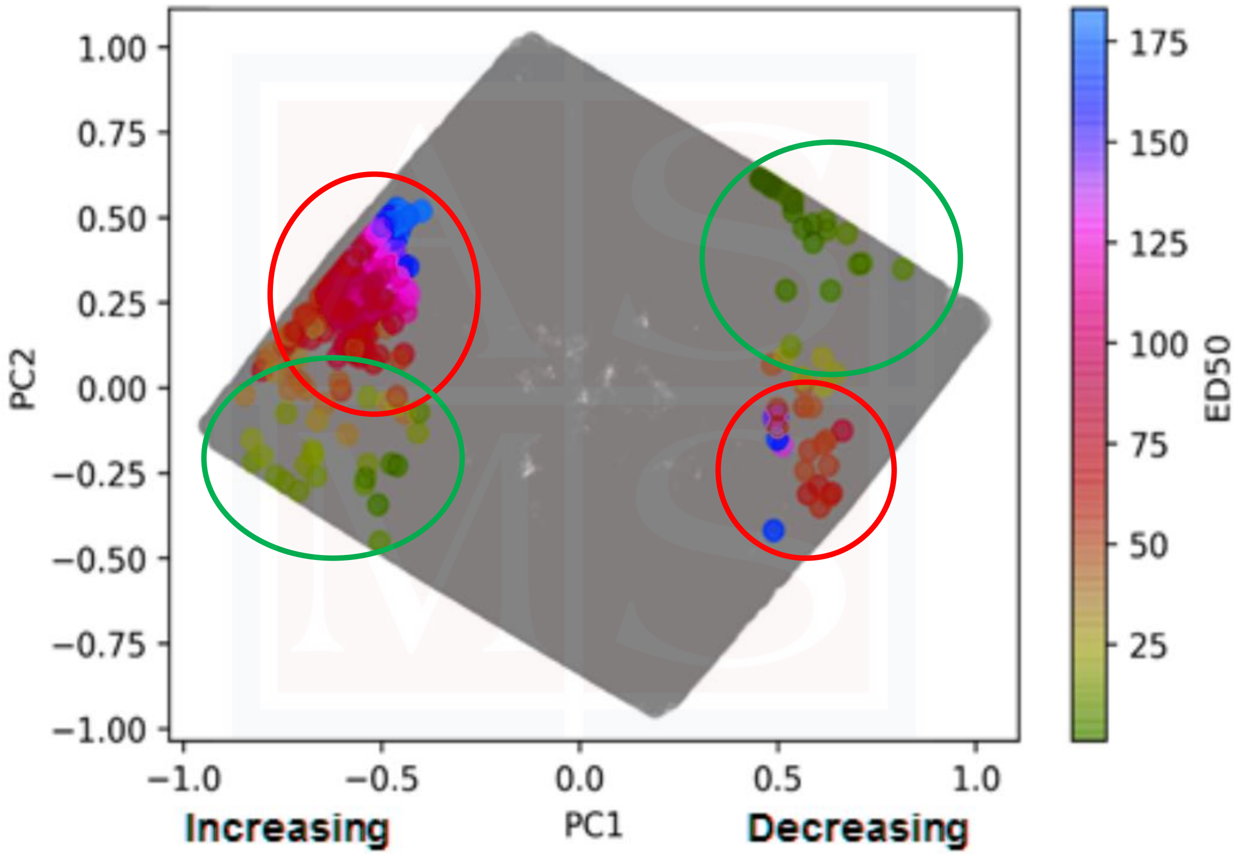
2. Variations in experimental design

dose-response data for etomoxir



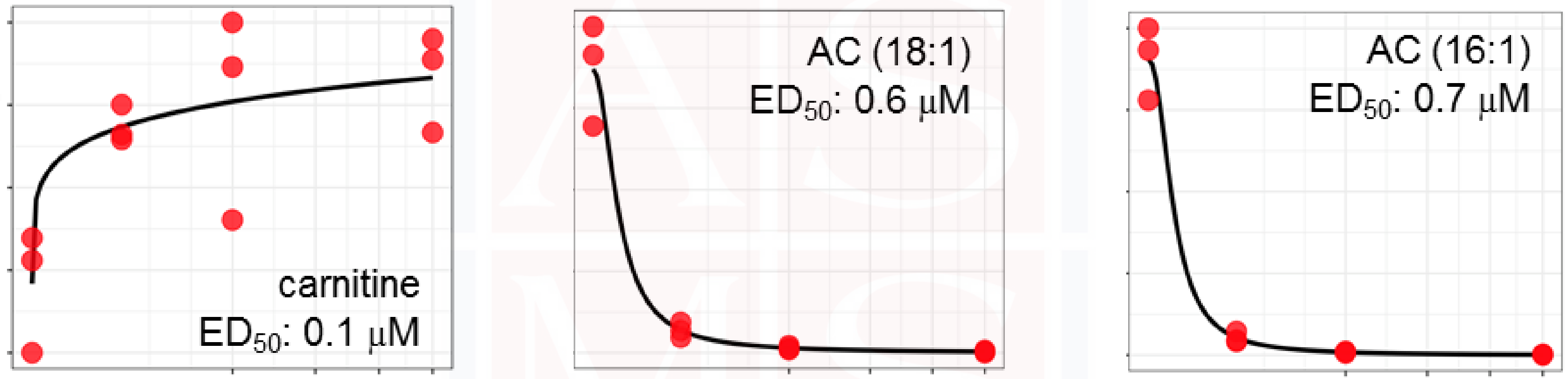
2. Variations in experimental design

dose-response data for etomoxir



2. Variations in experimental design

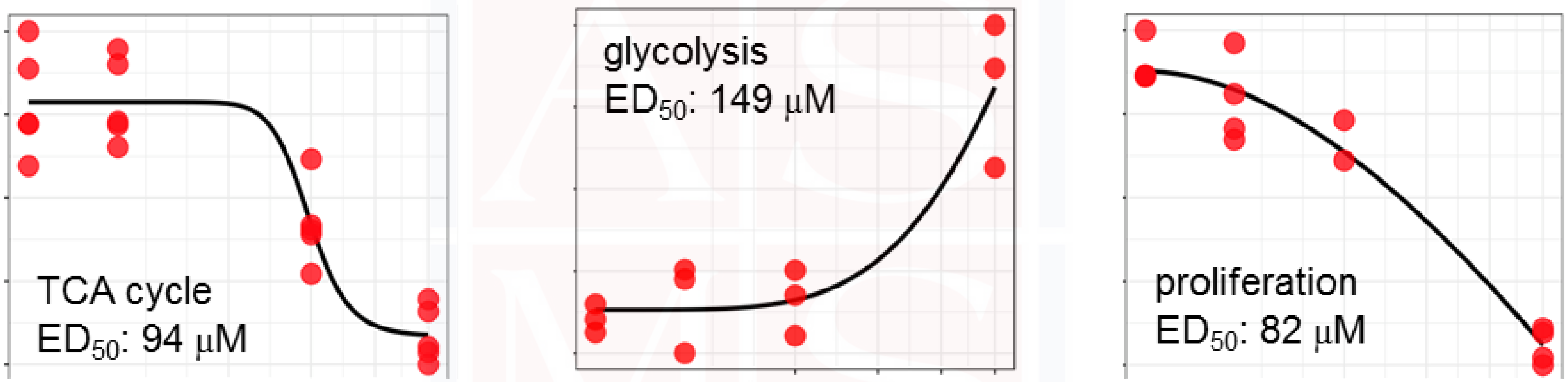
dose-response data for etomoxir



fatty acid ox inhibited at low conc.

2. Variations in experimental design

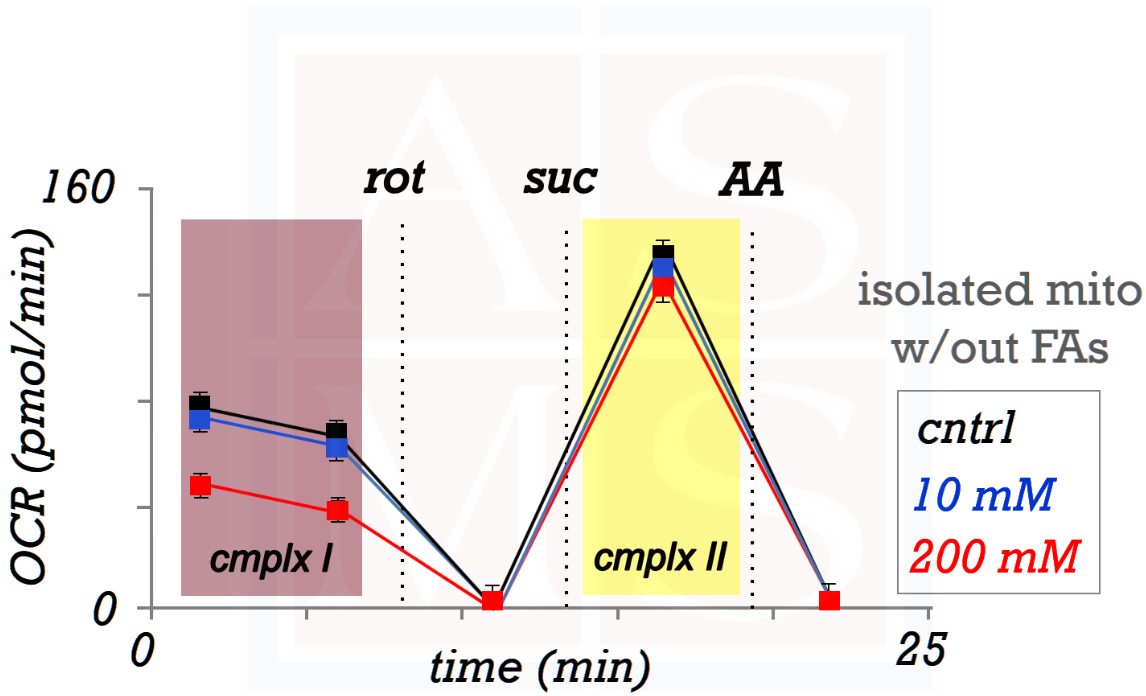
dose-response data for etomoxir



respiration inhibited at high conc.

2. Variations in experimental design

Off target of etomoxir is complex I



Advanced workflows for metabolomics

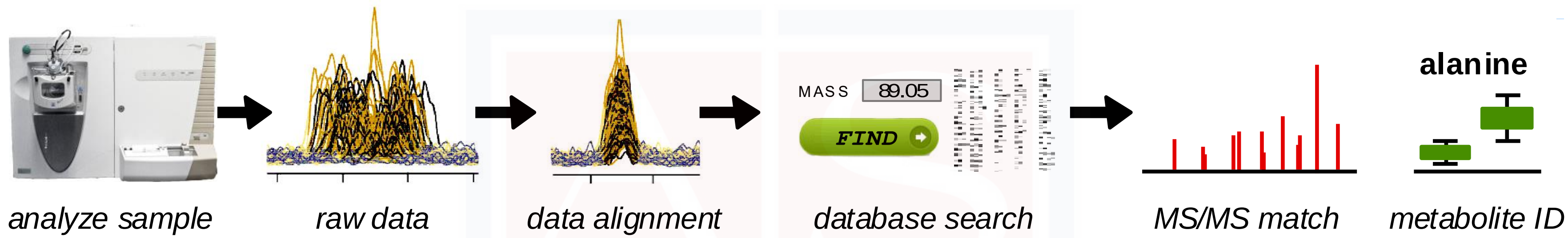
1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

Advanced workflows for metabolomics

1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

3. Improved identification workflows

"Conventional"

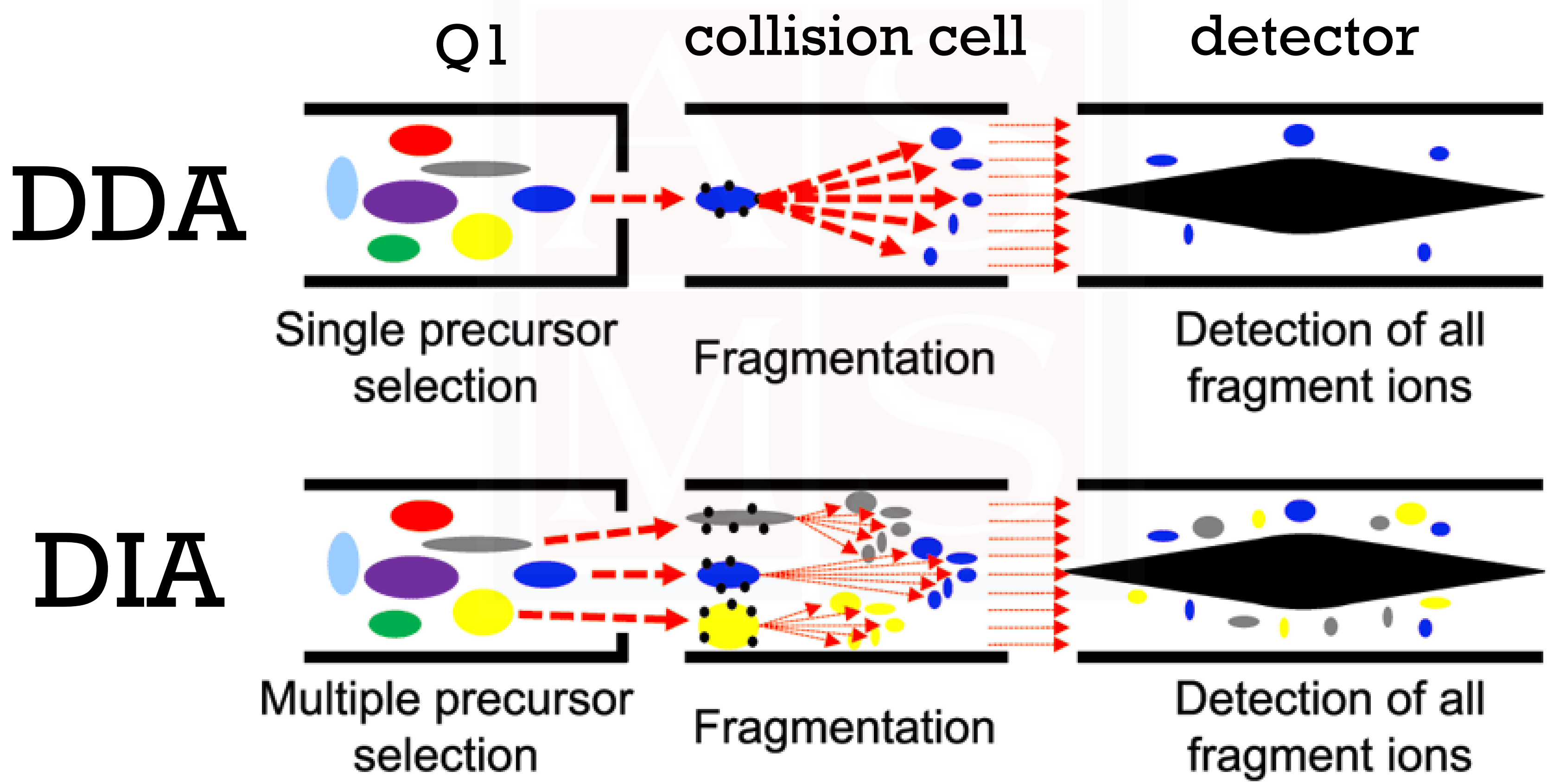


- Time intensive, information limited
- MS2 data<<MS1 data
- Alternative is brute-force approach where get MS2 on all signals during profiling

  - data independent analysis (DIA), SWATH-like acquisition, All-Ions MS/MS (Agilent), LC/MS^{e} (Waters)

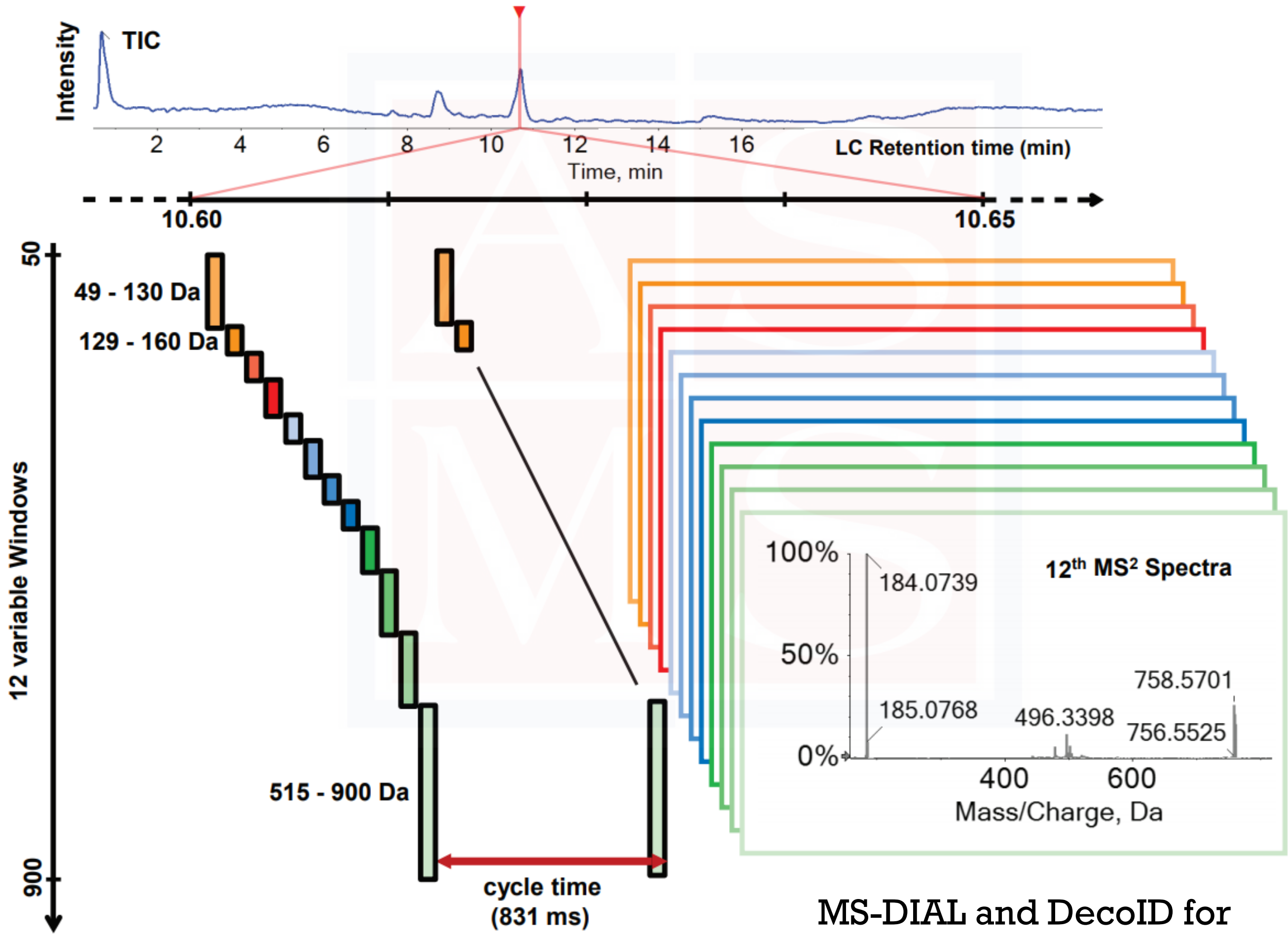
3. Improved identification workflows

DDA vs DIA methods



3. Improved identification workflows

SWATH for DIA

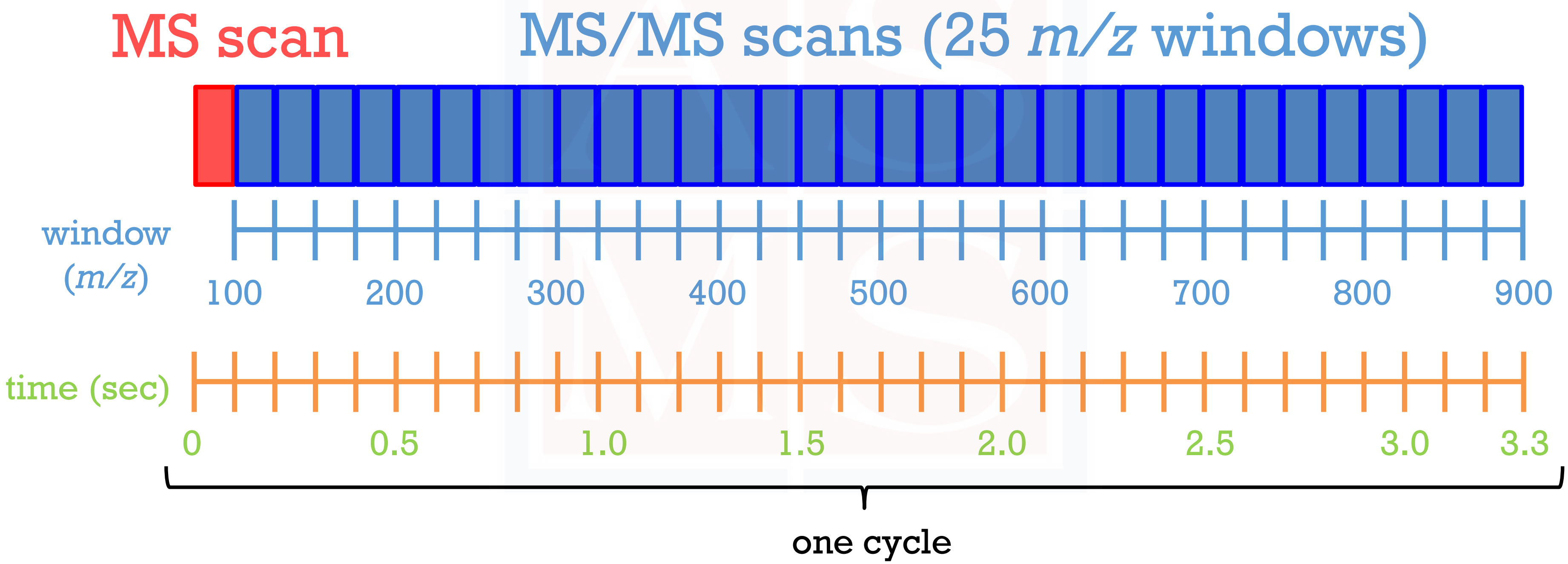


MS-DIAL and DecoID for deconvolution of data

Tsugawa et al., Nature Methods 2015
Stancliffe et al., Nature Methods 2021

3. Improved identification workflows

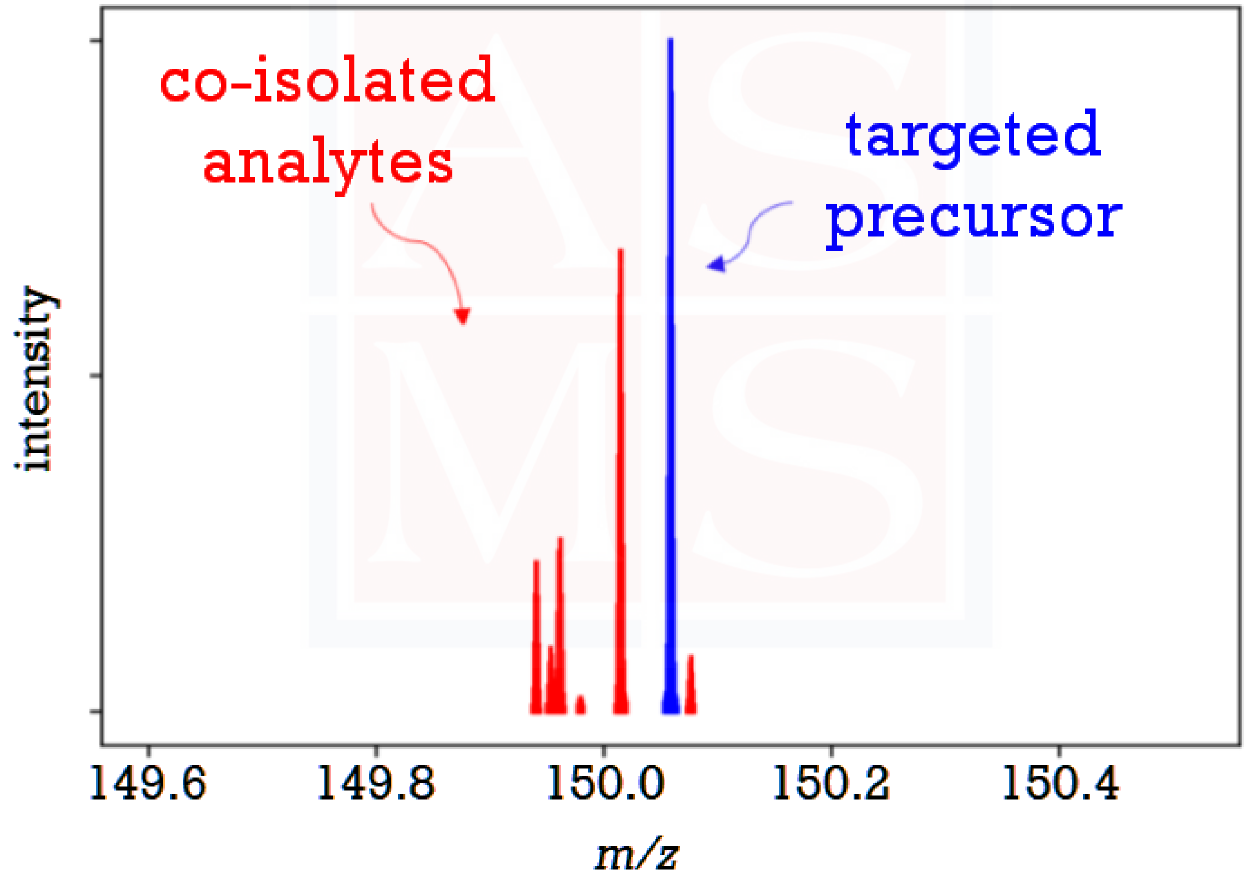
SWATH for DIA



Tsugawa et al., Nature Methods 2015
Stancliffe et al., Nature Methods 2021

3. Improved identification workflows

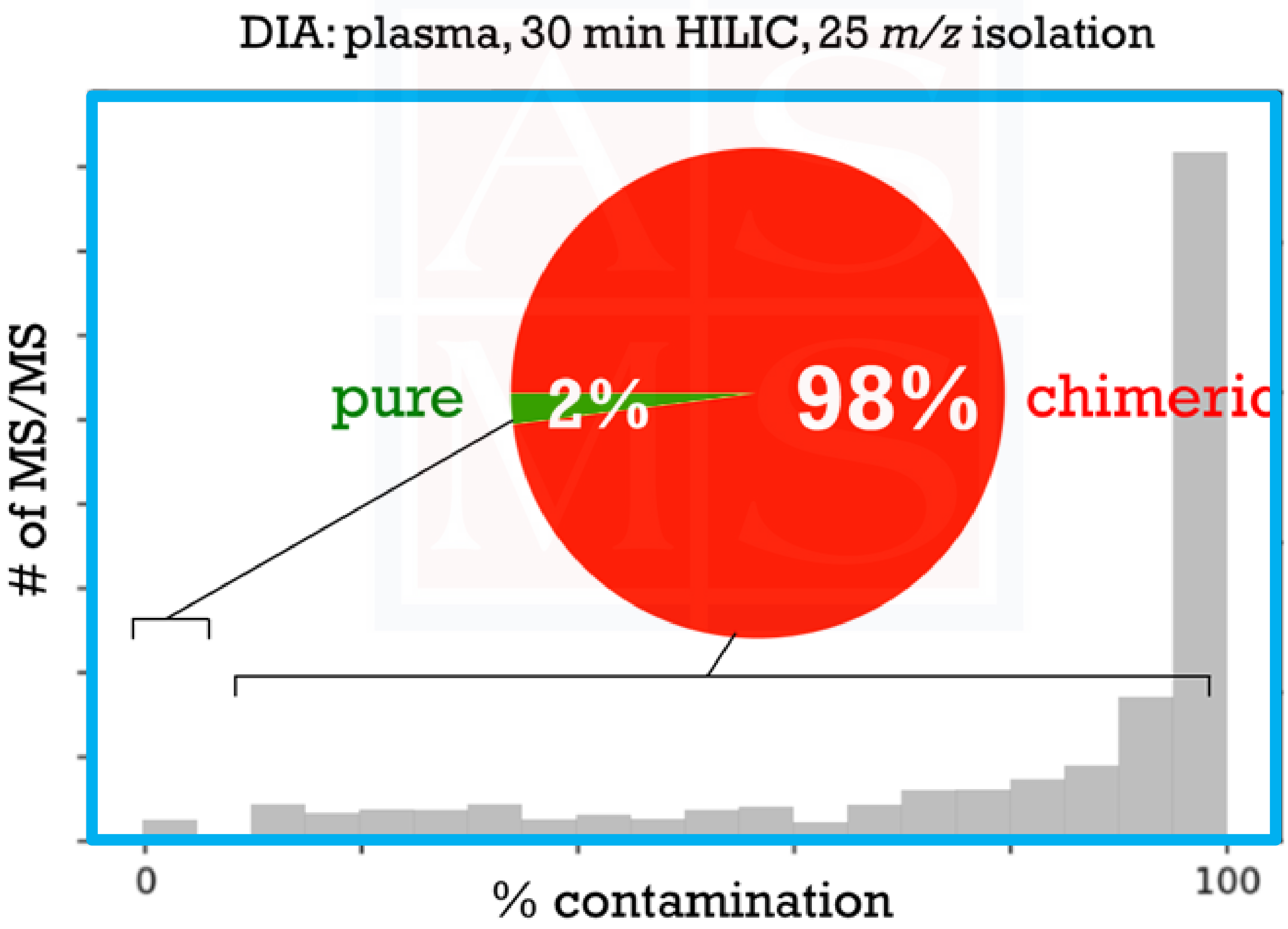
SWATH for DIA



Lawson et al., Anal Chem (89, 4) 2017

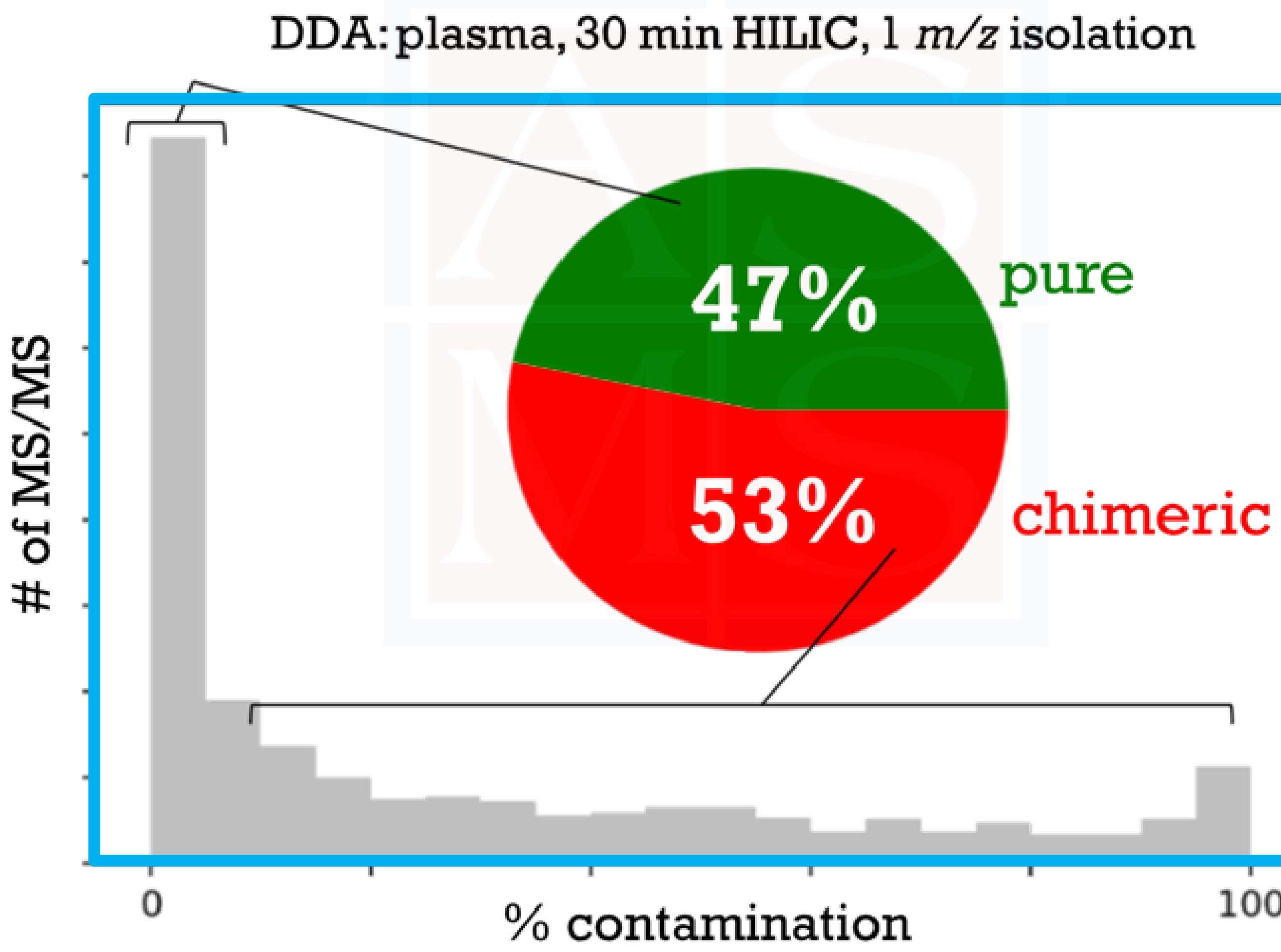
3. Improved identification workflows

SWATH for DIA



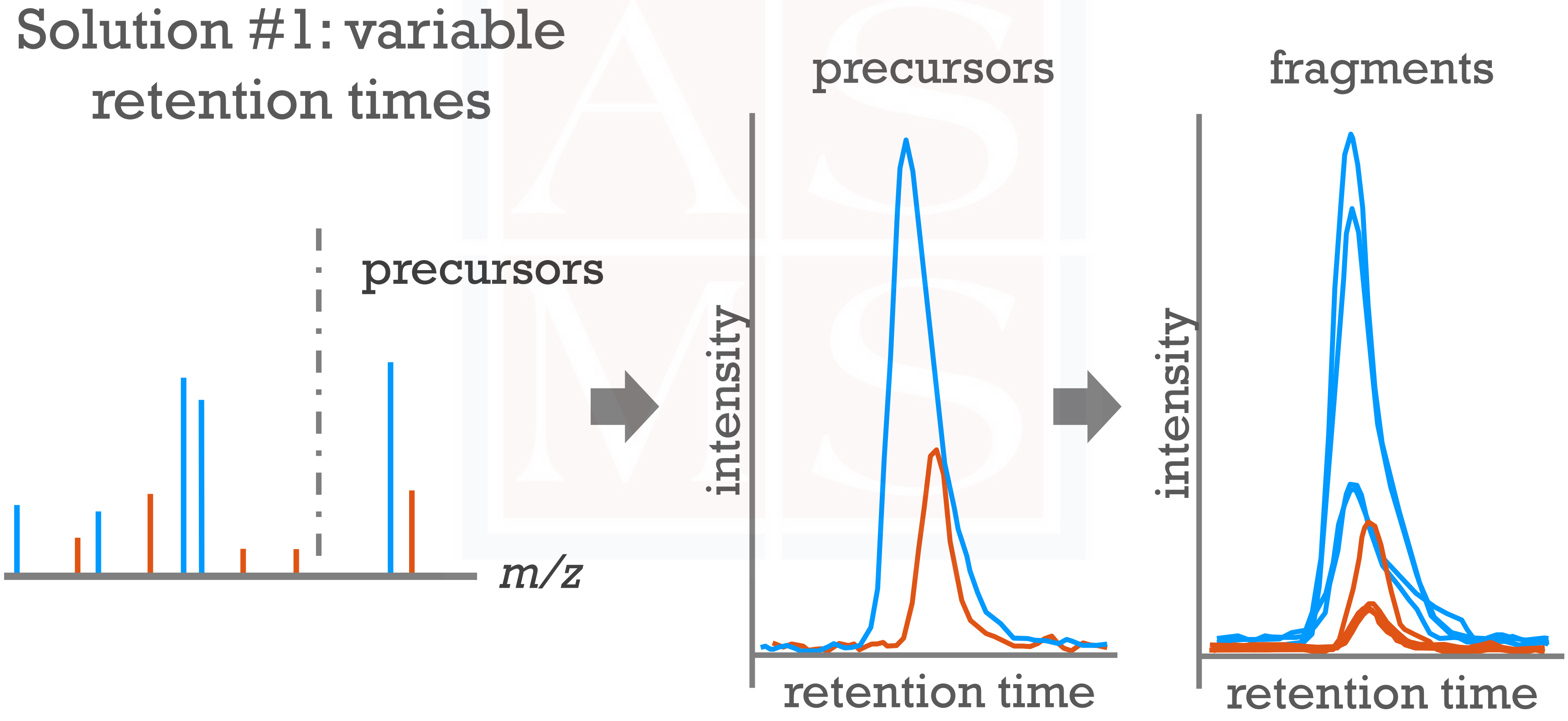
3. Improved identification workflows

SWATH for DIA



3. Improved identification workflows

SWATH for DIA

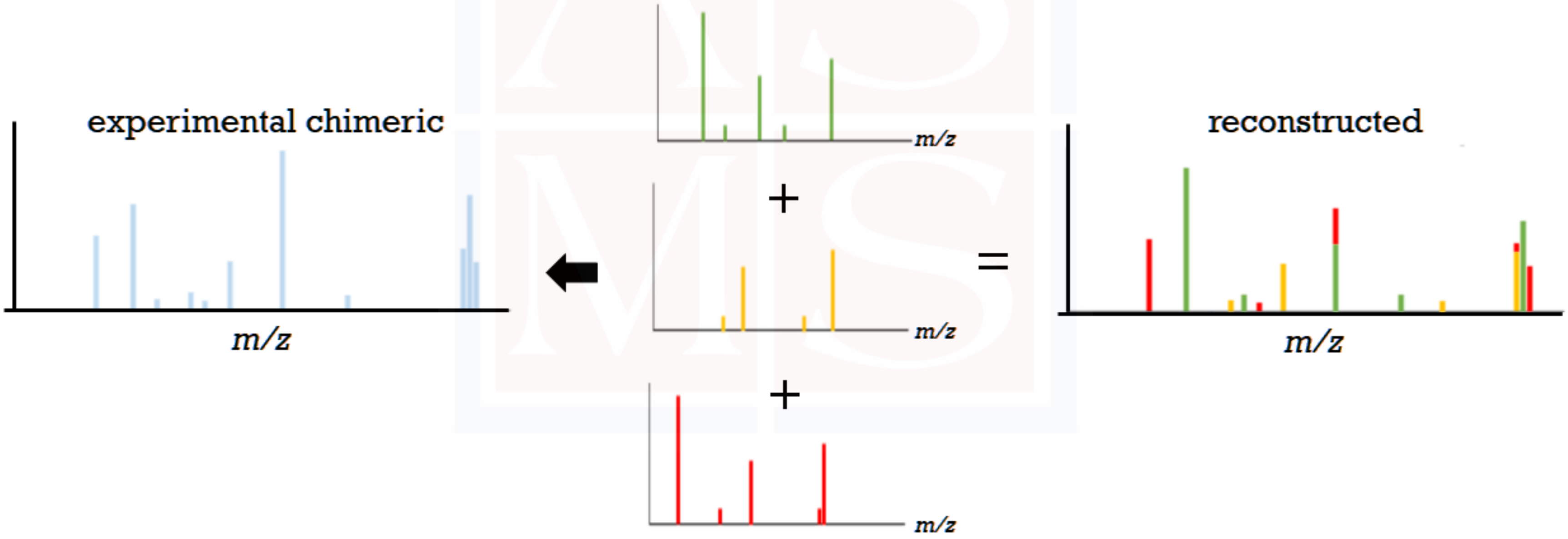


DecoMS2: Nikolskiy et al., Anal Chem (85, 16) 2013
MS-DIAL: Tsugawa et al., Nature Methods (12) 2015

3. Improved identification workflows

SWATH for DIA

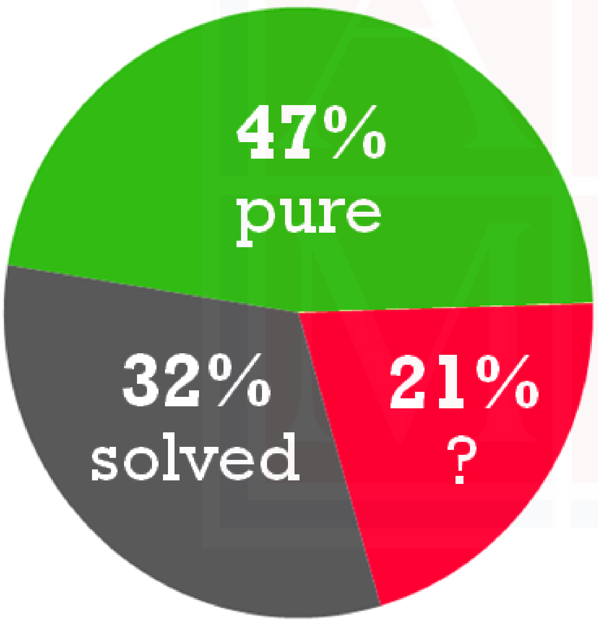
Solution #2: back-end deconvolution



DecoID: Stancliffe et al, Nature Methods 2021

3. Improved identification workflows

SWATH for DIA



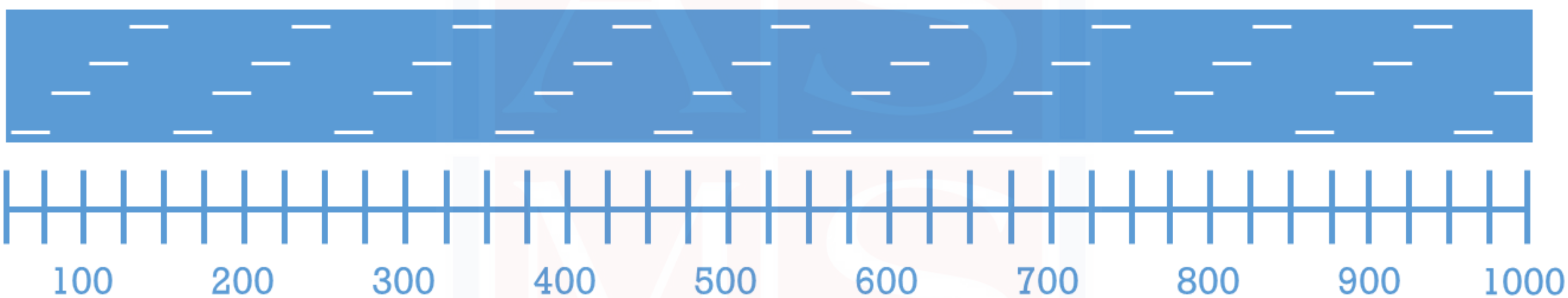
unresolved: how do we ID these compounds?

3. Improved identification workflows

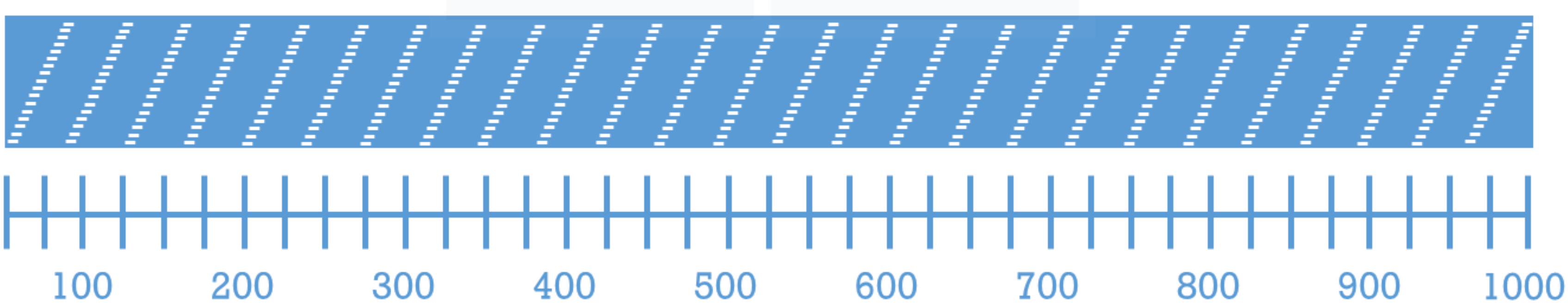
SWATH for DIA

MS/MS windows

SWATH 25 *m/z* windows (minimal overlap)



ZT SCAN 5.1 *m/z* windows (0.1 *m/z* increments)



3. Improved identification workflows

Why DIA?

(+) MS/MS collected for every single feature

(-) Almost all of the MS/MS data collected is chimeric and requires computational deconvolution*

(+) SWATH has more specificity than other DIA approaches

*Newer instruments such as the Astral and ZenoTOF can do DIA experiments with small MS/MS isolation windows

3. Improved identification workflows

3. Improved identification workflows

- standard DDA
  ions with highest intensity selected for MS/MS as time allows

3. Improved identification workflows

- standard DDA
  ions with highest intensity selected for MS/MS as time allows

- iterative DDA
  *after a high-intensity ion is targeted for MS/MS, it is moved to exclusion list**

*Koelmel et al., JASMS 2017

3. Improved identification workflows

- standard DDA
  ions with highest intensity selected for MS/MS as time allows

- iterative DDA
  *after a high-intensity ion is targeted for MS/MS, it is moved to exclusion list**

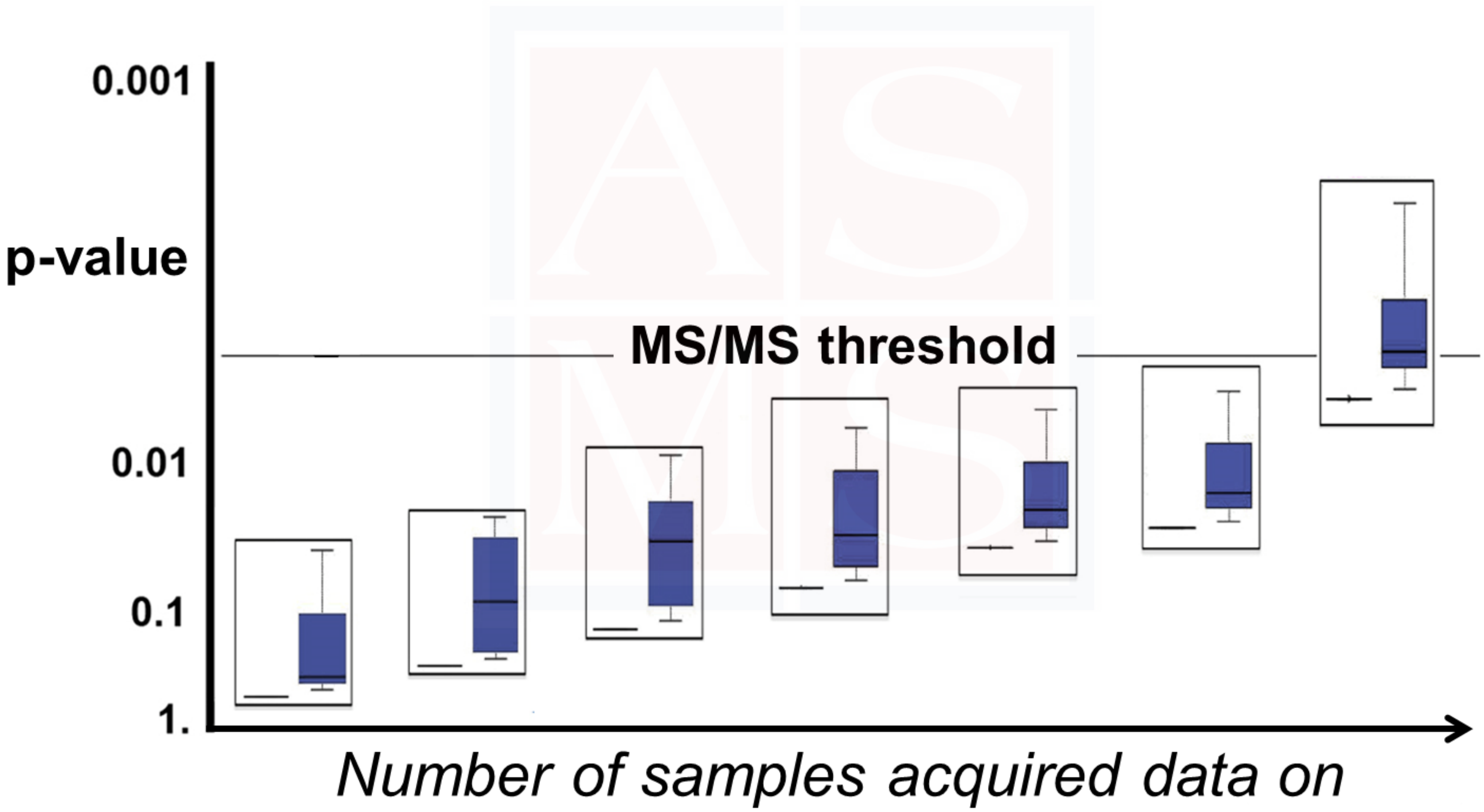
- intelligent DDA
  features selected for MS/MS on the basis of statistics and/or biology

*Koelmel et al., JASMS 2017

3. Improved identification workflows

3. Improved identification workflows

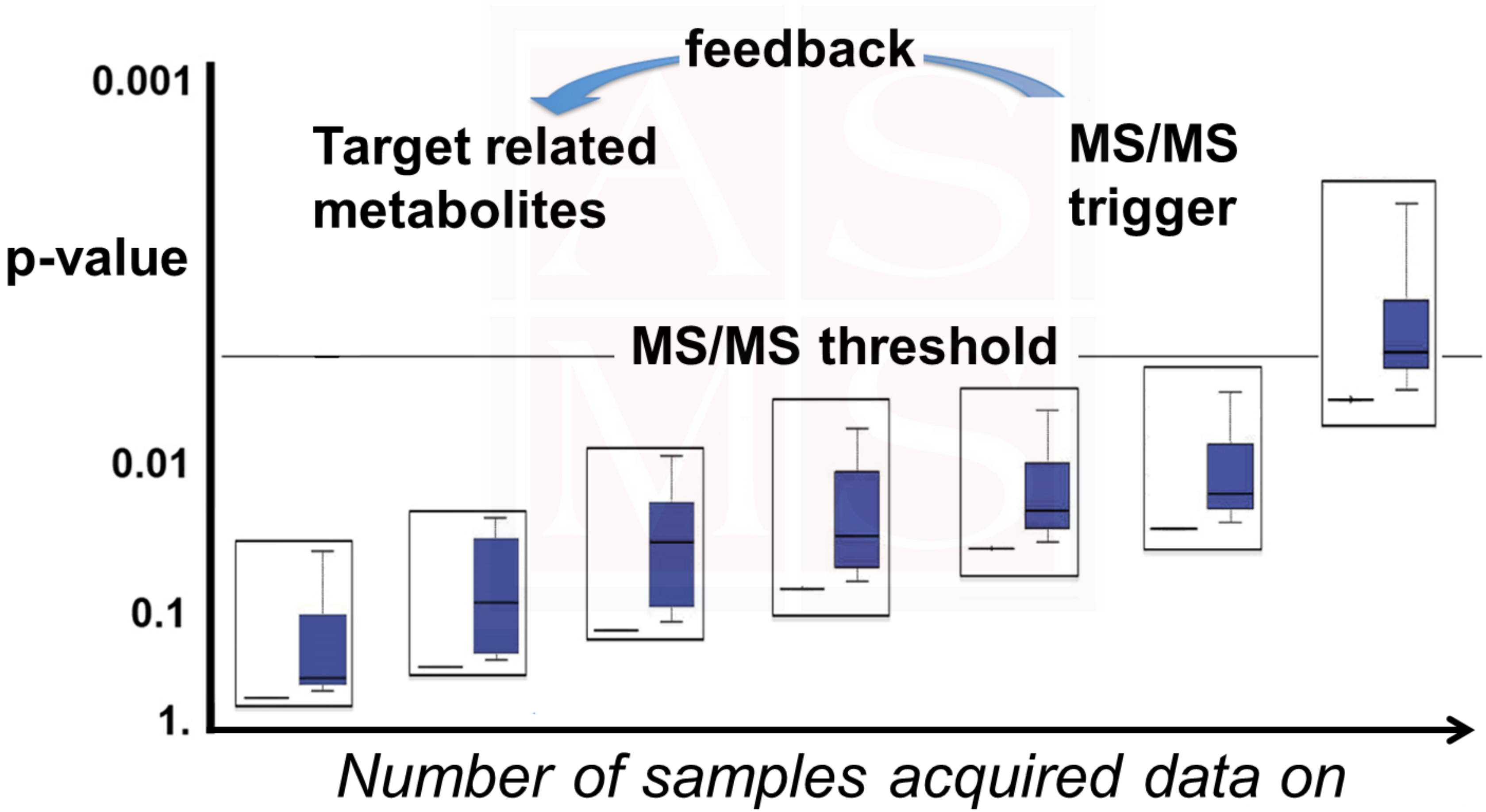
Intelligent DDA: targeting MS/MS intelligently



Rinehart et al., Nat Biotechnol. 2014

3. Improved identification workflows

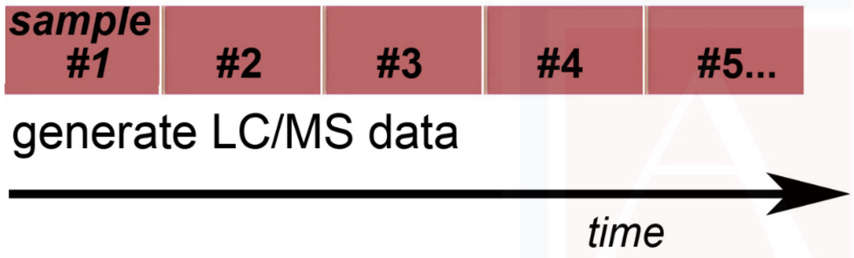
Intelligent DDA: targeting MS/MS intelligently



Rinehart et al., Nat Biotechnol. 2014

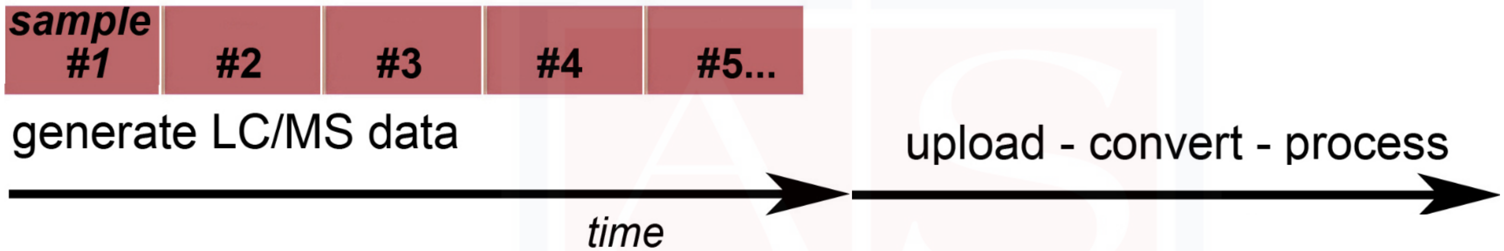
3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



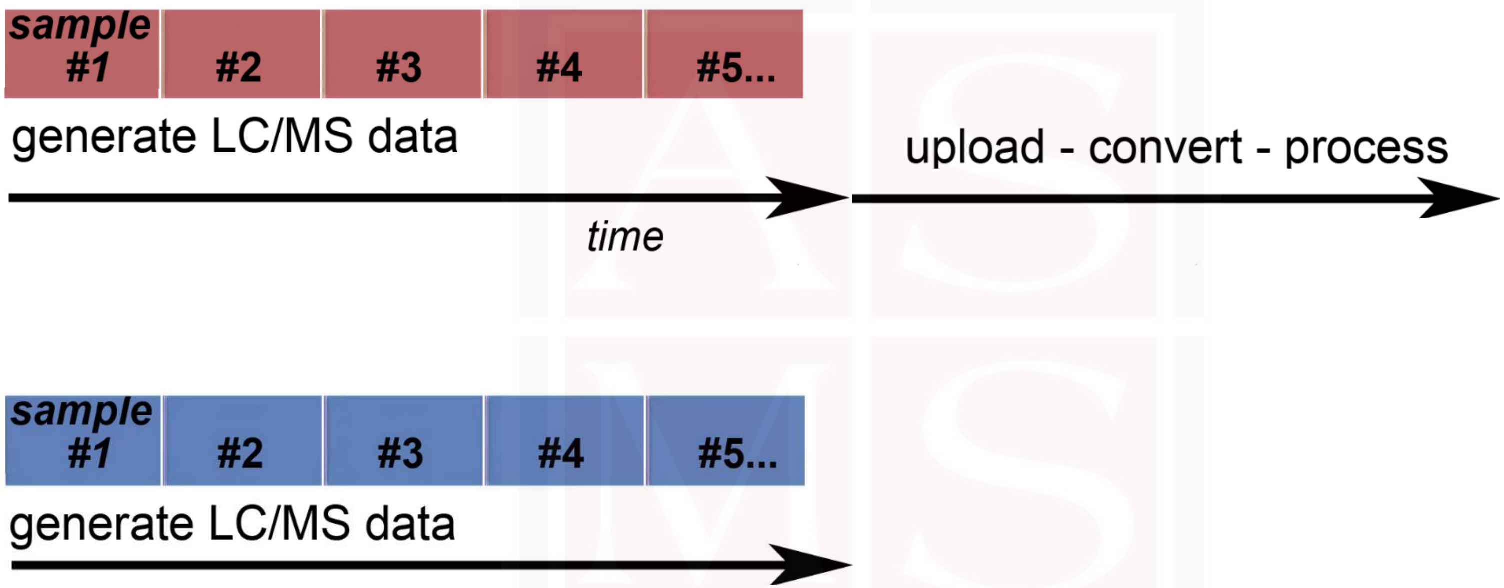
3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



3. Improved identification workflows

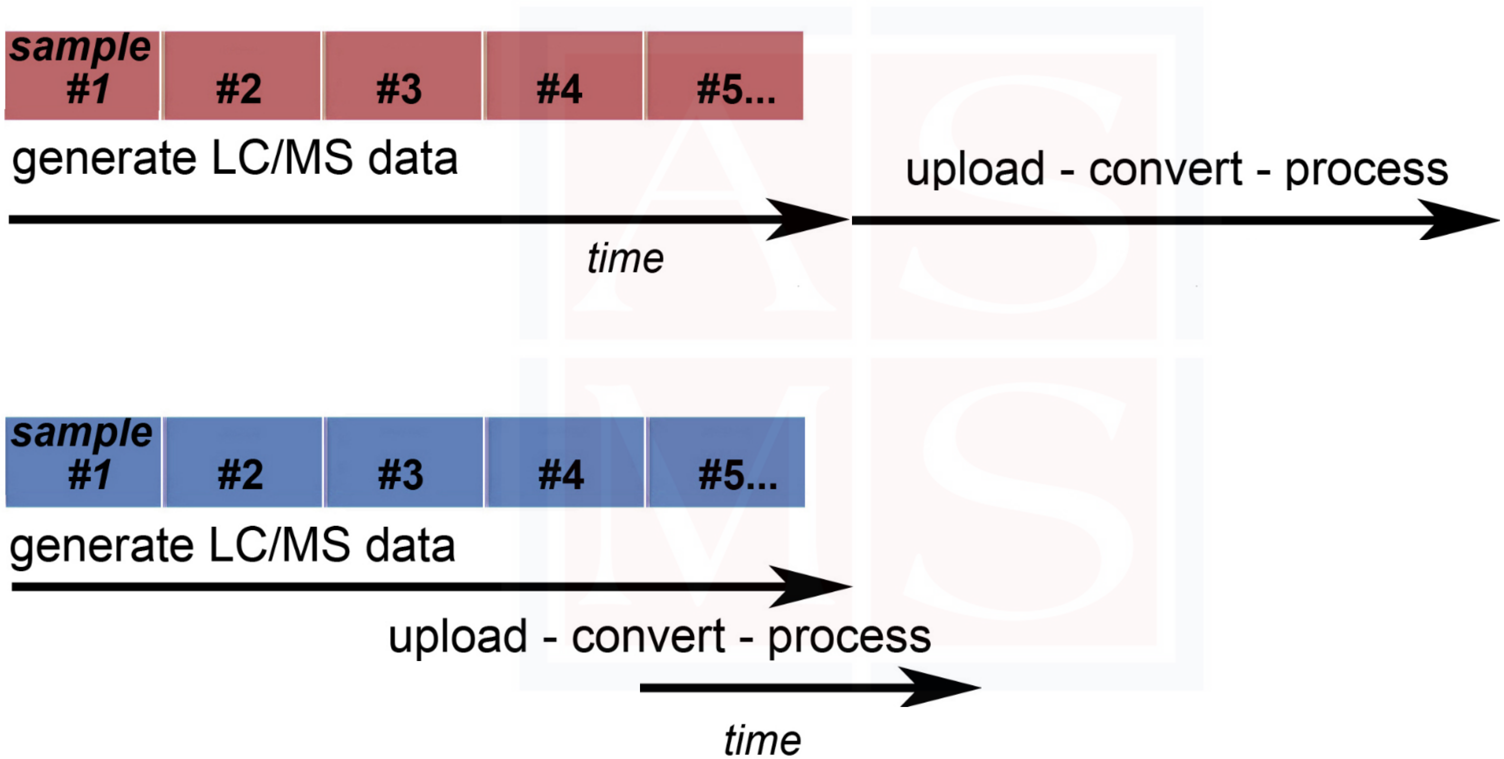
Intelligent DDA: targeting MS/MS intelligently



Rinehart et al., Nat Biotechnol. 2014

3. Improved identification workflows

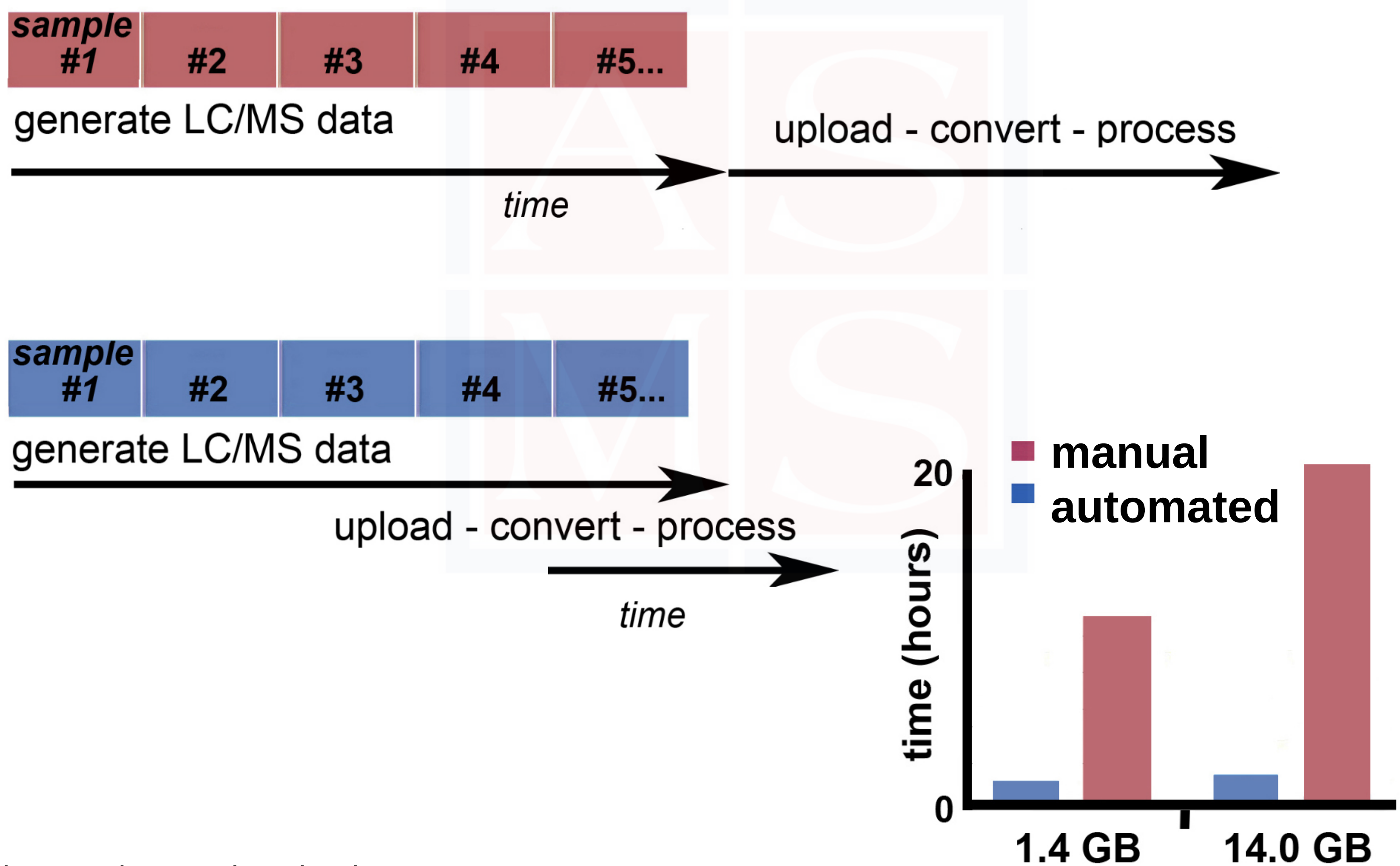
Intelligent DDA: targeting MS/MS intelligently



Rinehart et al., Nat Biotechnol. 2014

3. Improved identification workflows

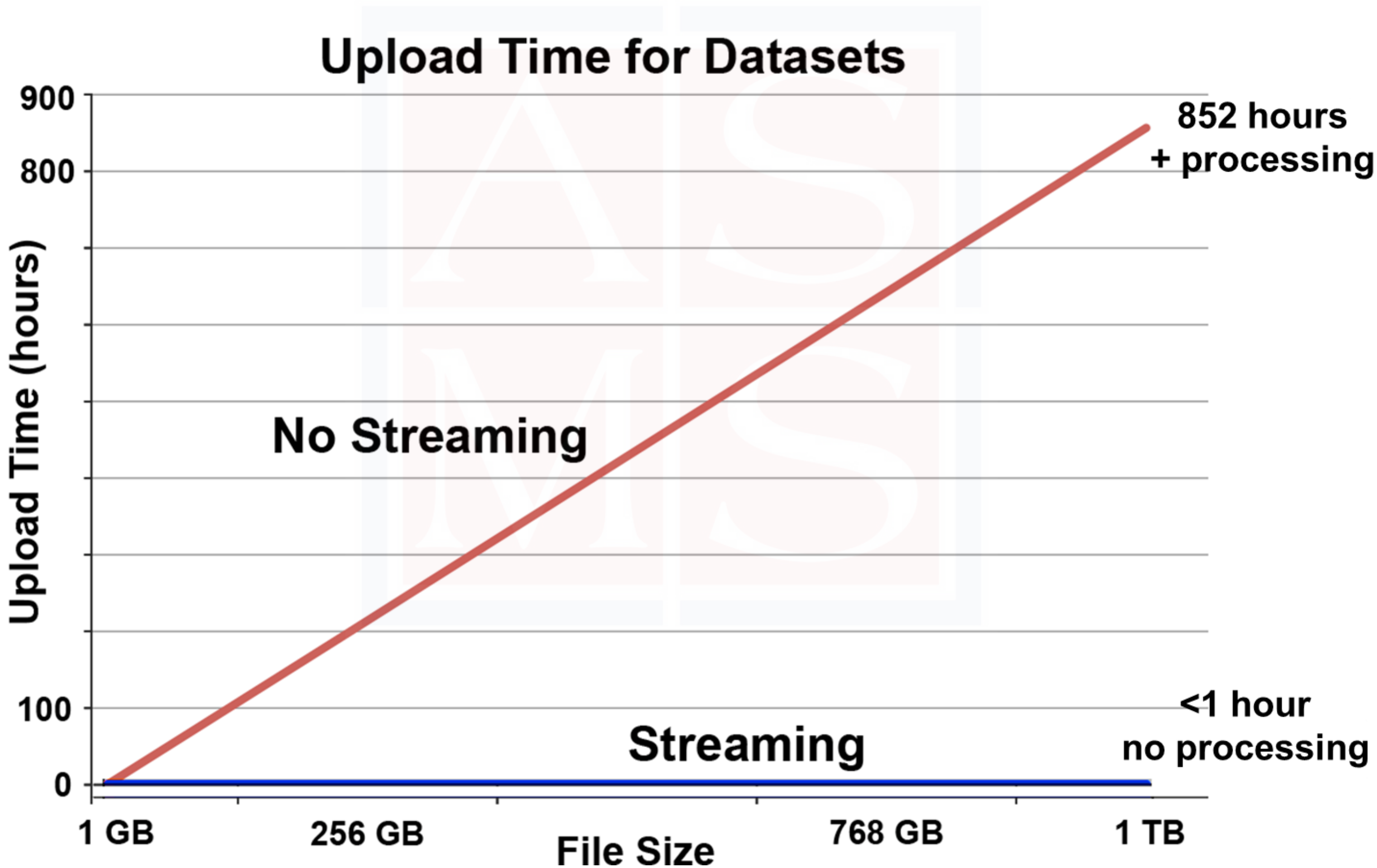
Intelligent DDA: targeting MS/MS intelligently



Rinehart et al., Nat Biotechnol. 2014

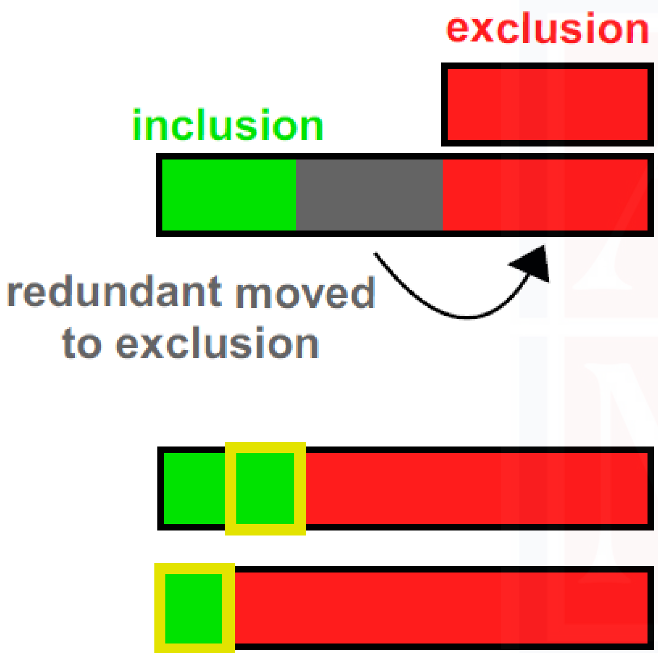
3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



Step 1: build exclusion list from blank

Step 2: build inclusion list from research sample

Step 3: generate MS2 from research sample

run 1

run 2

...

as needed

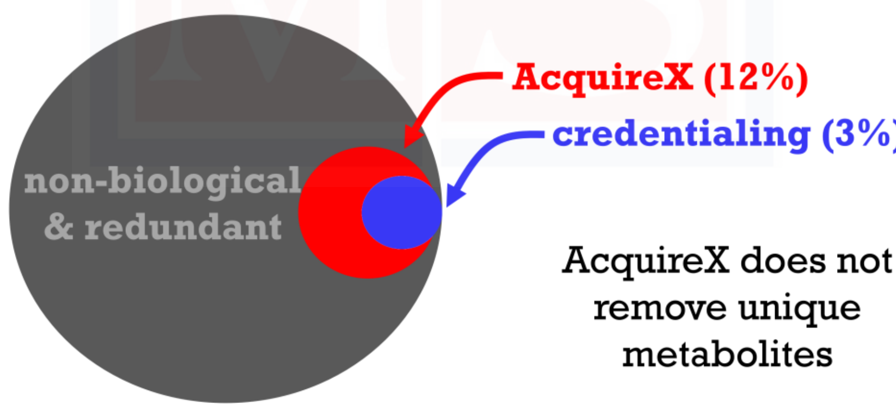
AcquireX on IQ-X Orbitrap instrument

3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently

AcquireX Thermo Fisher SCIENTIFIC

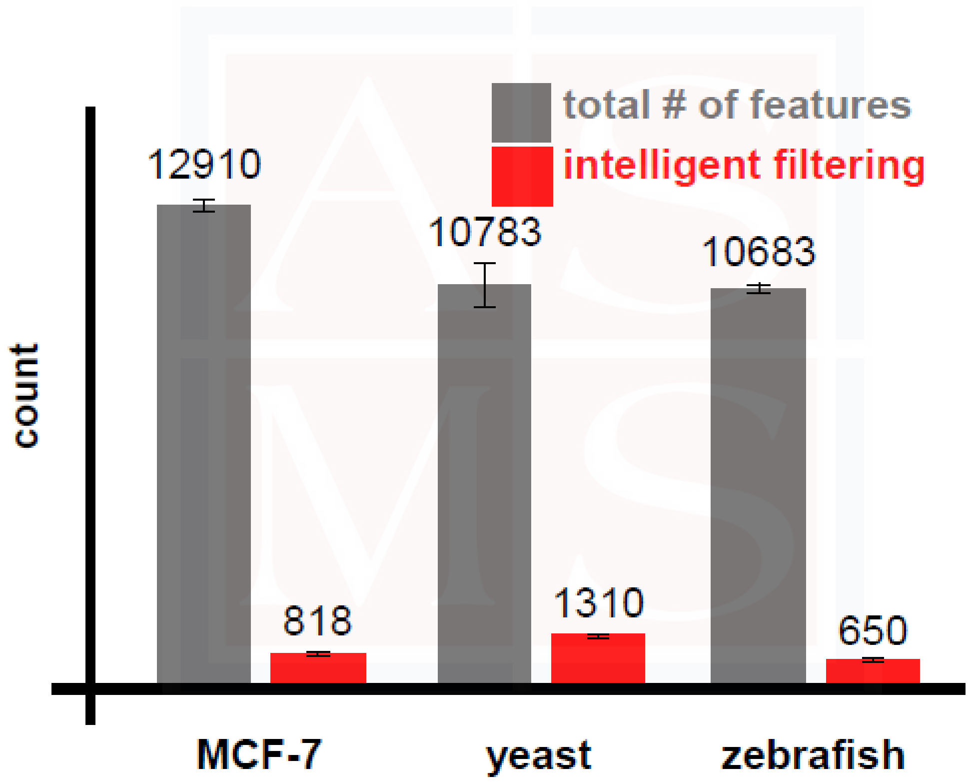
isotopes, adducts,
blank subtraction



AcquireX does not remove unique metabolites

3. Improved identification workflows

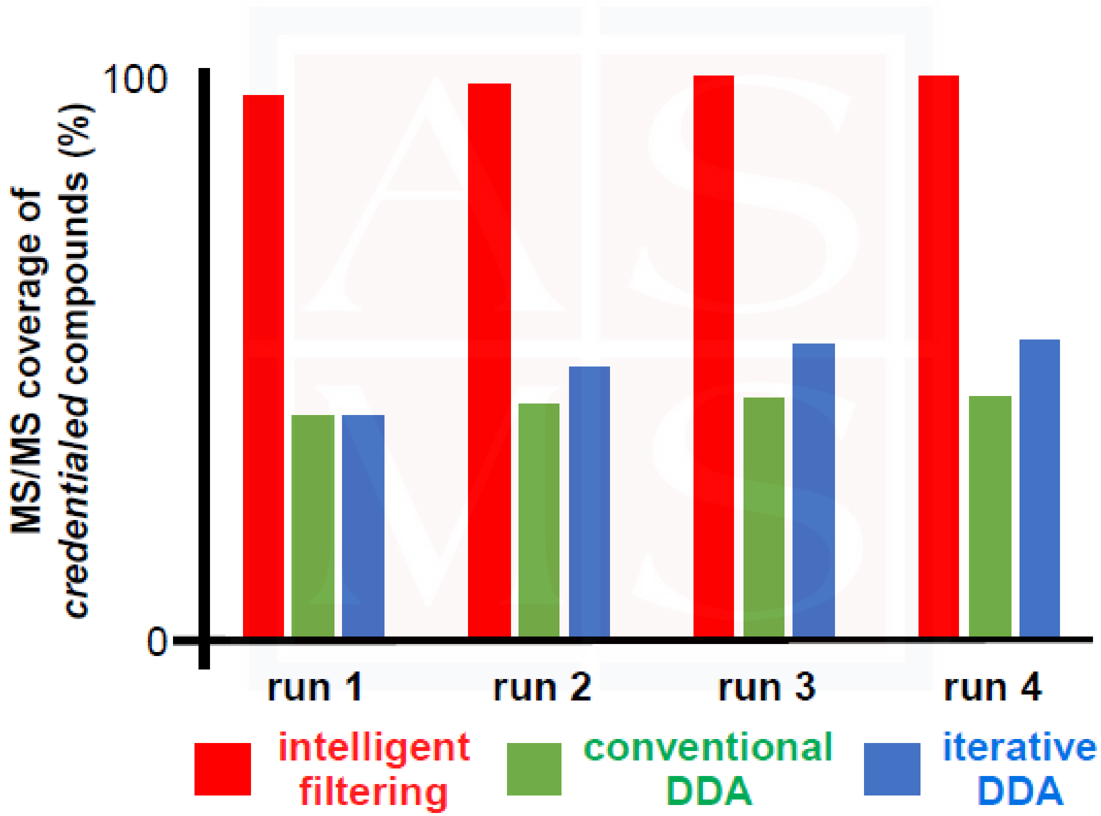
Intelligent DDA: targeting MS/MS intelligently



AcquireX on IQ-X Orbitrap instrument

3. Improved identification workflows

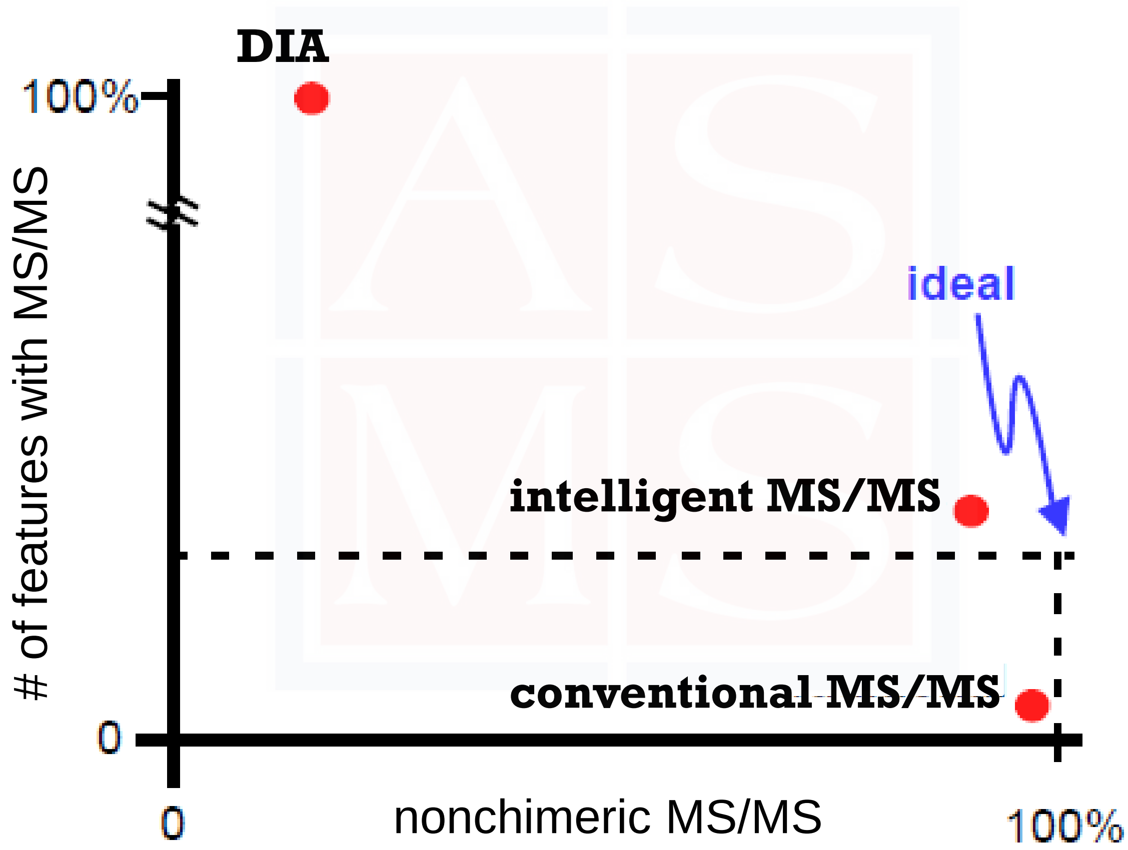
Intelligent DDA: targeting MS/MS intelligently



AcquireX on IQ-X Orbitrap instrument

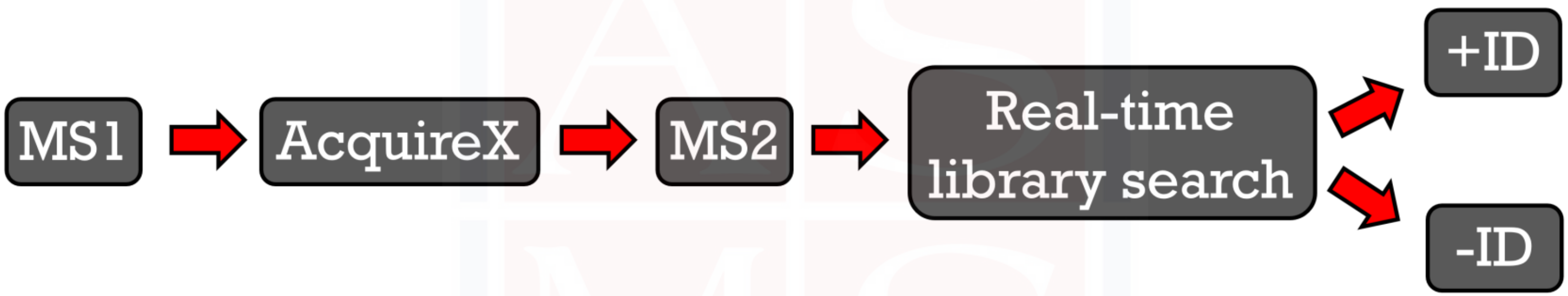
3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



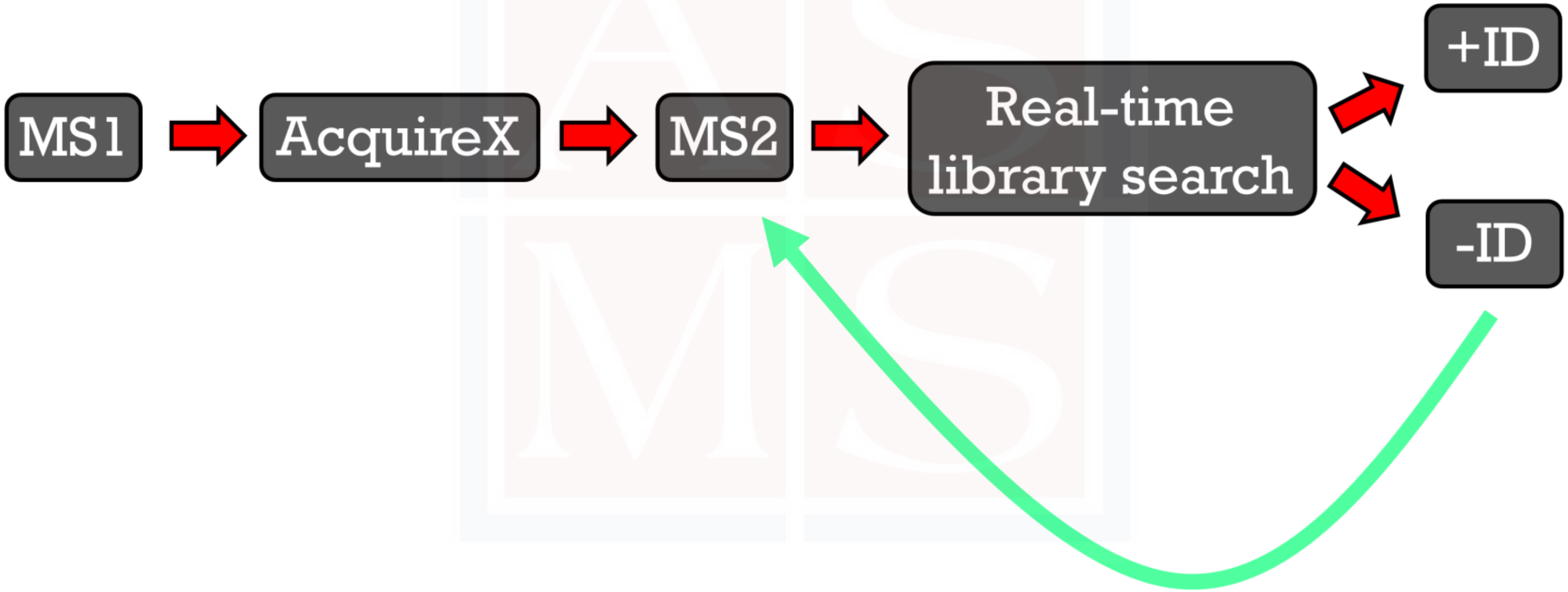
3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently



3. Improved identification workflows

Intelligent DDA: targeting MS/MS intelligently

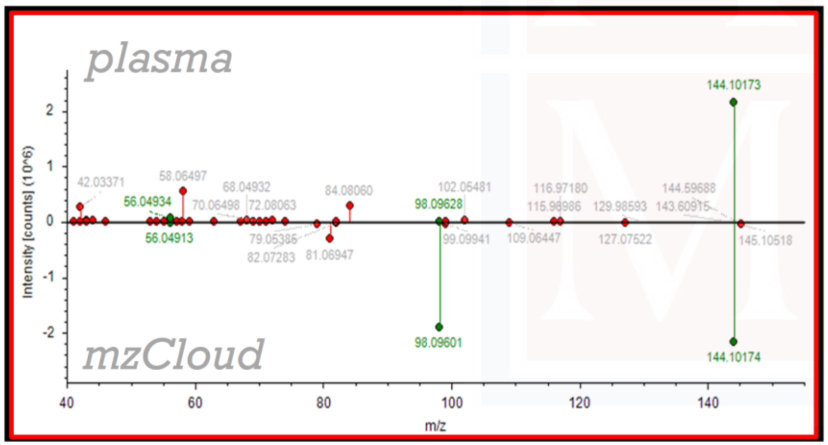


*IQ-X Orbitrap: high coverage **<u>and</u>** quality*

3. Improved identification workflows

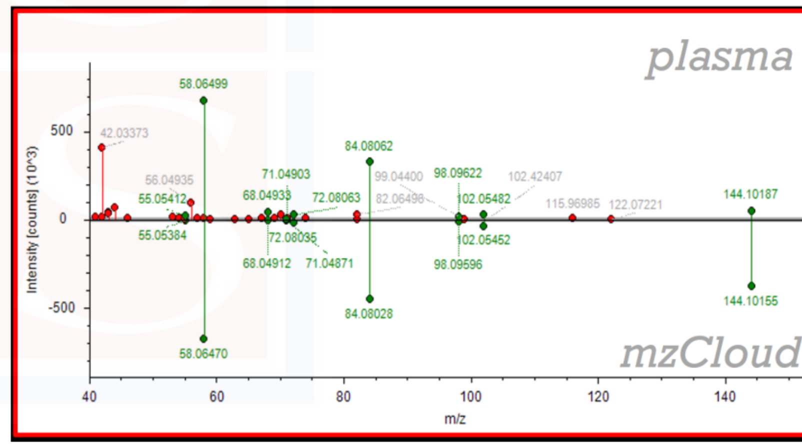
Intelligent DDA: targeting MS/MS intelligently

example: IQ-X data for stachydrine



under-fragmented

low confidence



higher voltage

more confidence

Advanced workflows for metabolomics

1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

Advanced workflows for metabolomics

1. Large-scale analyses
   - *several hundred to thousands of samples*
2. Variations in experimental design
   - *meta-analysis, dose-response metabolomics*
3. Improved identification workflows
   - *DIA, SWATH, DDA, iterative DDA, AcquireX*
4. Expanding targeted analyses
   - *predicting MRMs, barcoding metabolomics*

4. Expanding targeted analyses

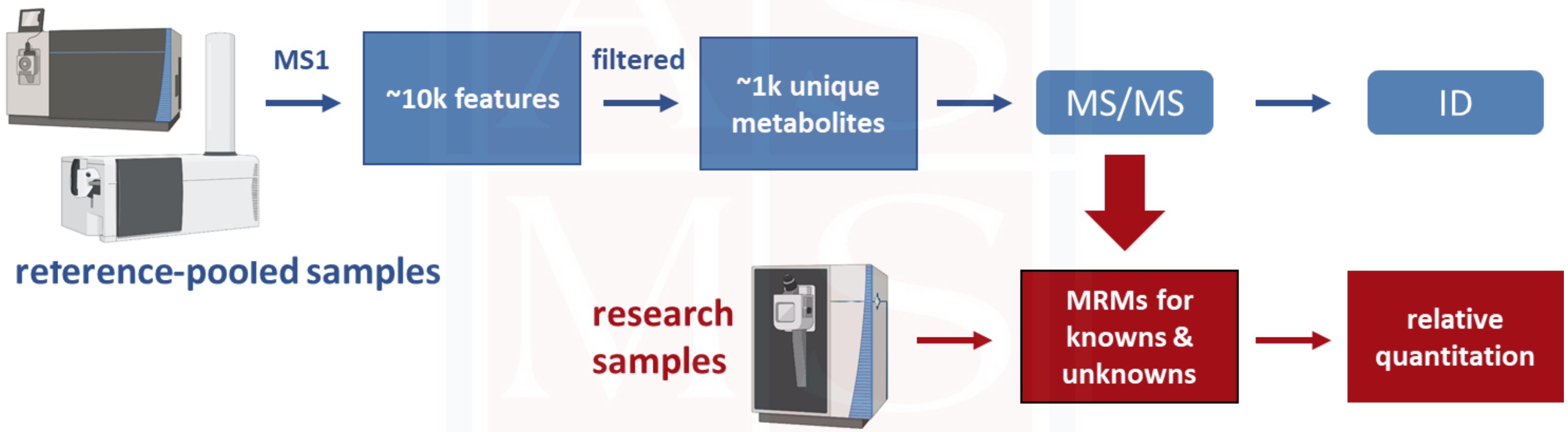
Targeted metabolomics using QqQ at Untargeted Scale

4. Expanding targeted analyses

Targeted metabolomics using QqQ at Untargeted Scale

4. Expanding targeted analyses

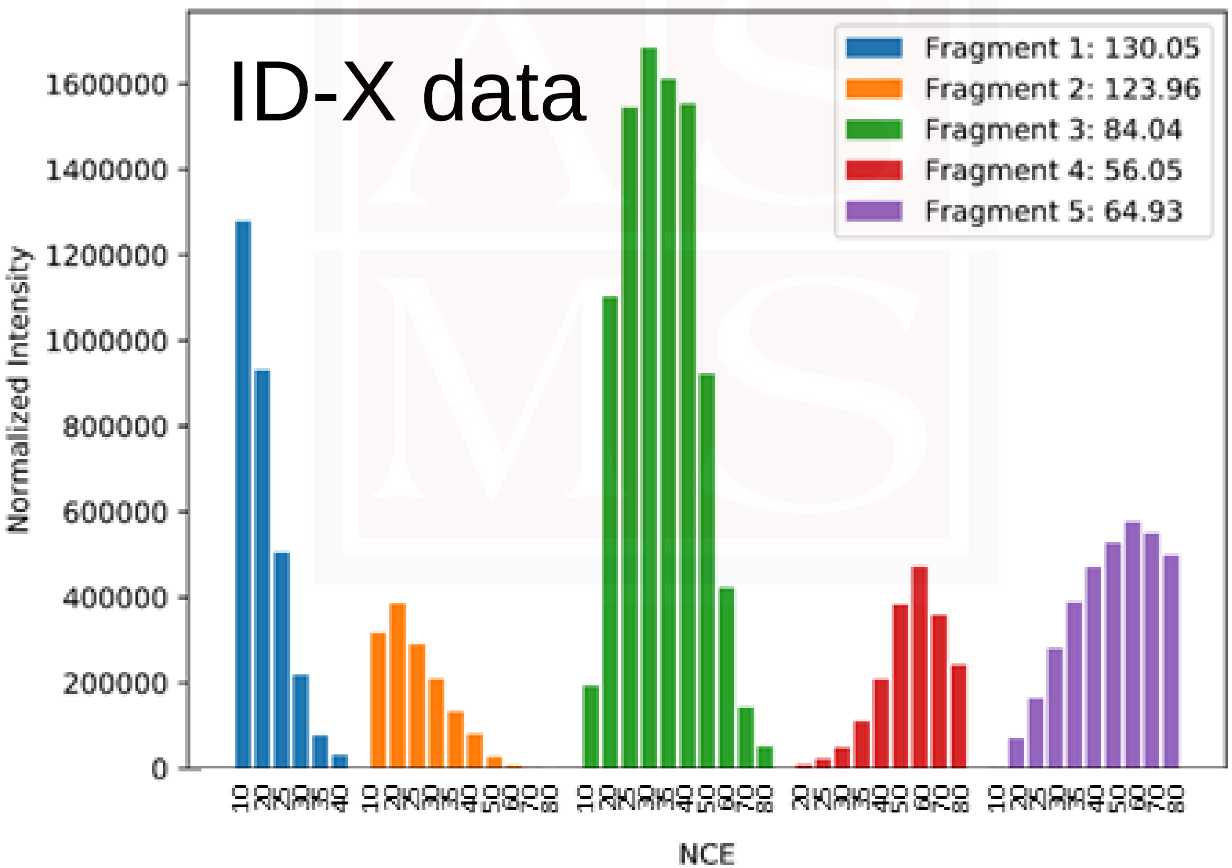
Targeted metabolomics using QqQ at Untargeted Scale



Schwaiger-Haber et al., ACS Measurement Science 2021

4. Expanding targeted analyses

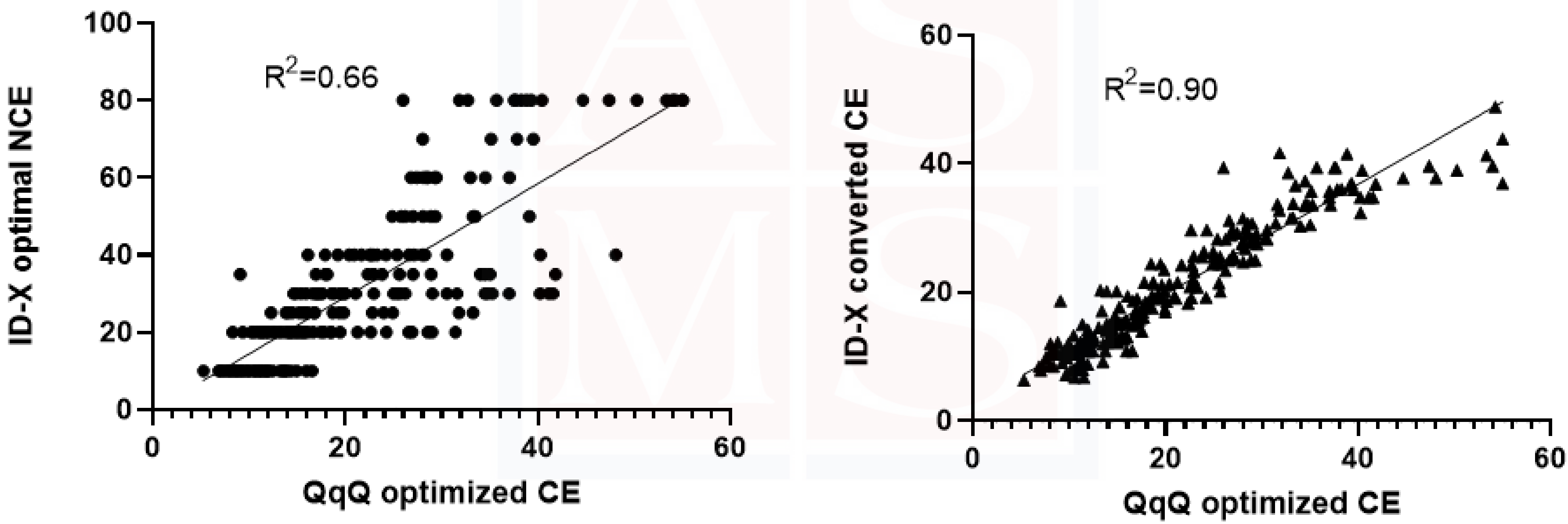
Targeted metabolomics using QqQ at Untargeted Scale



Schwaiger-Haber et al., ACS Measurement Science 2021

4. Expanding targeted analyses

Targeted metabolomics using QqQ at Untargeted Scale



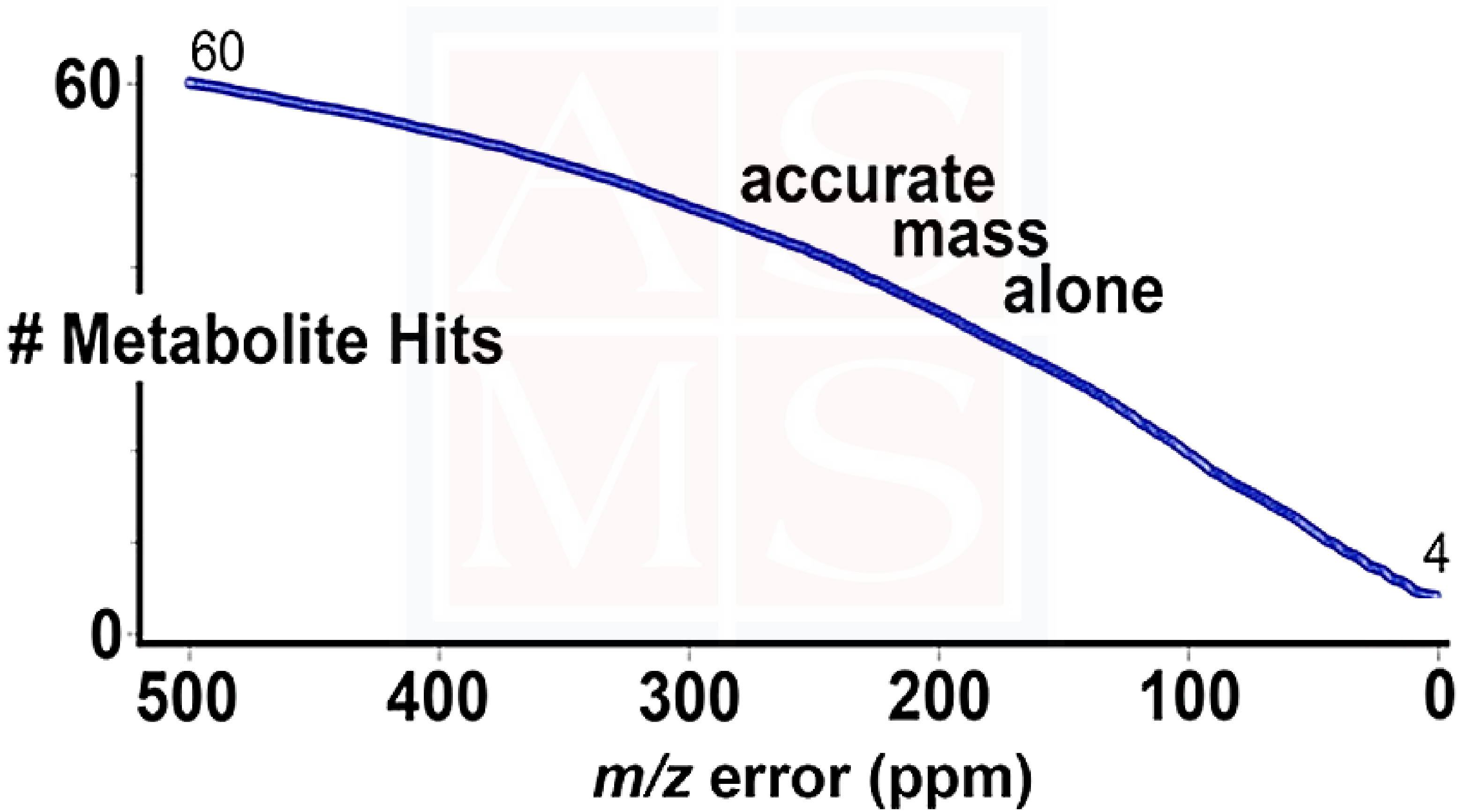
Schwaiger-Haber et al., ACS Measurement Science 2021

4. Expanding targeted analyses

Bar coding metabolomics

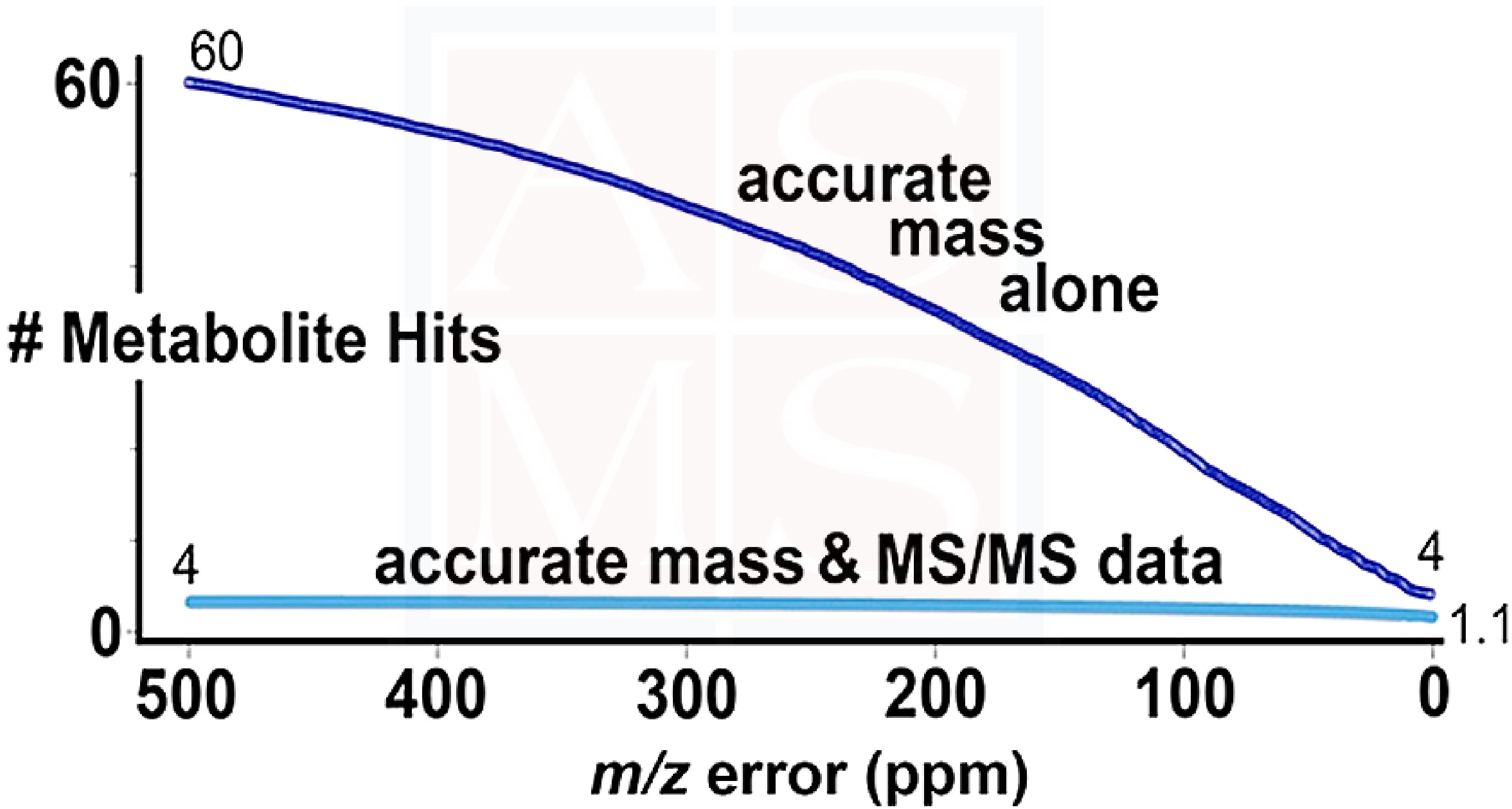
4. Expanding targeted analyses

Bar coding metabolomics



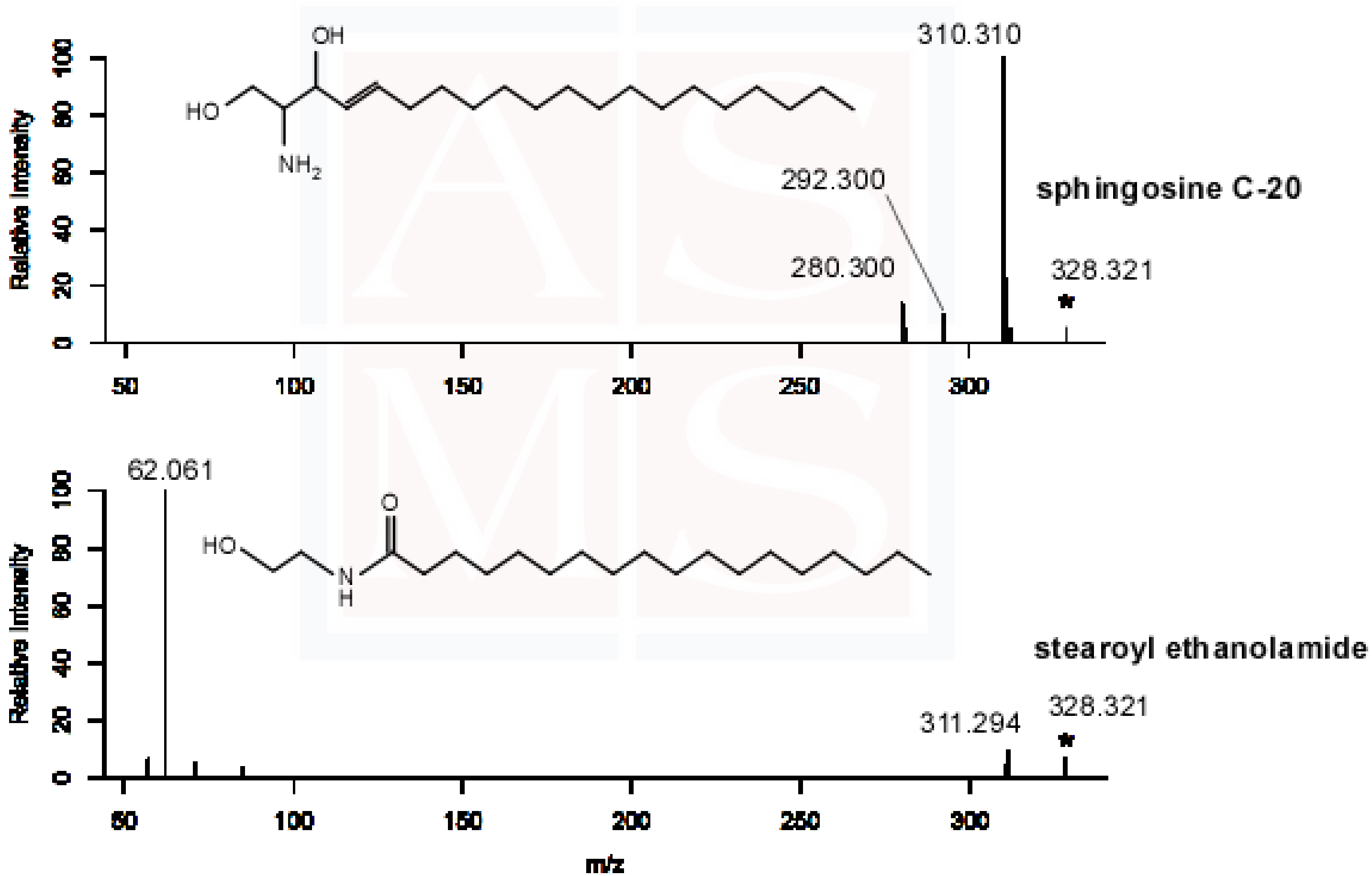
4. Expanding targeted analyses

Bar coding metabolomics



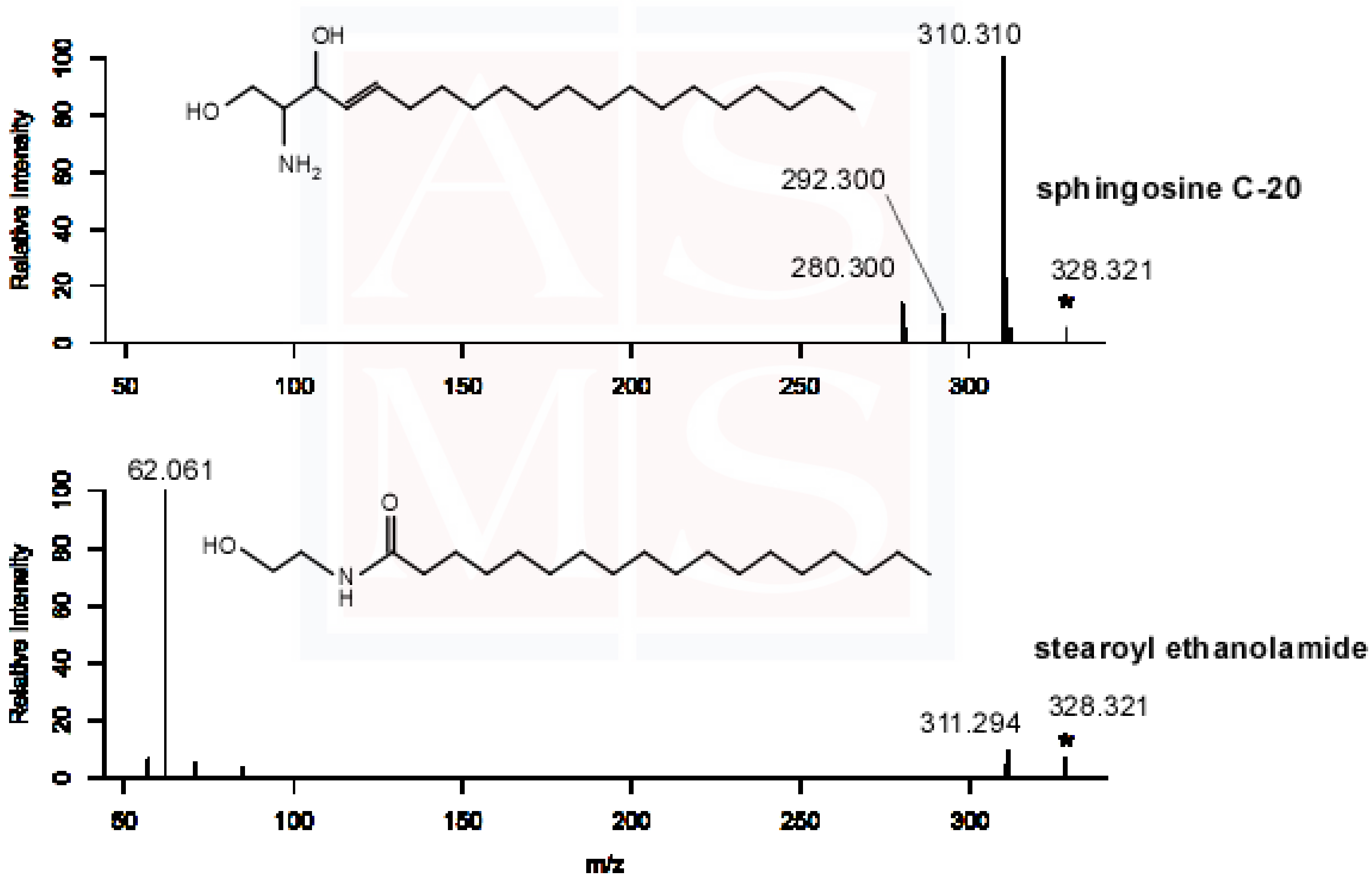
4. Expanding targeted analyses

Bar coding metabolomics



4. Expanding targeted analyses

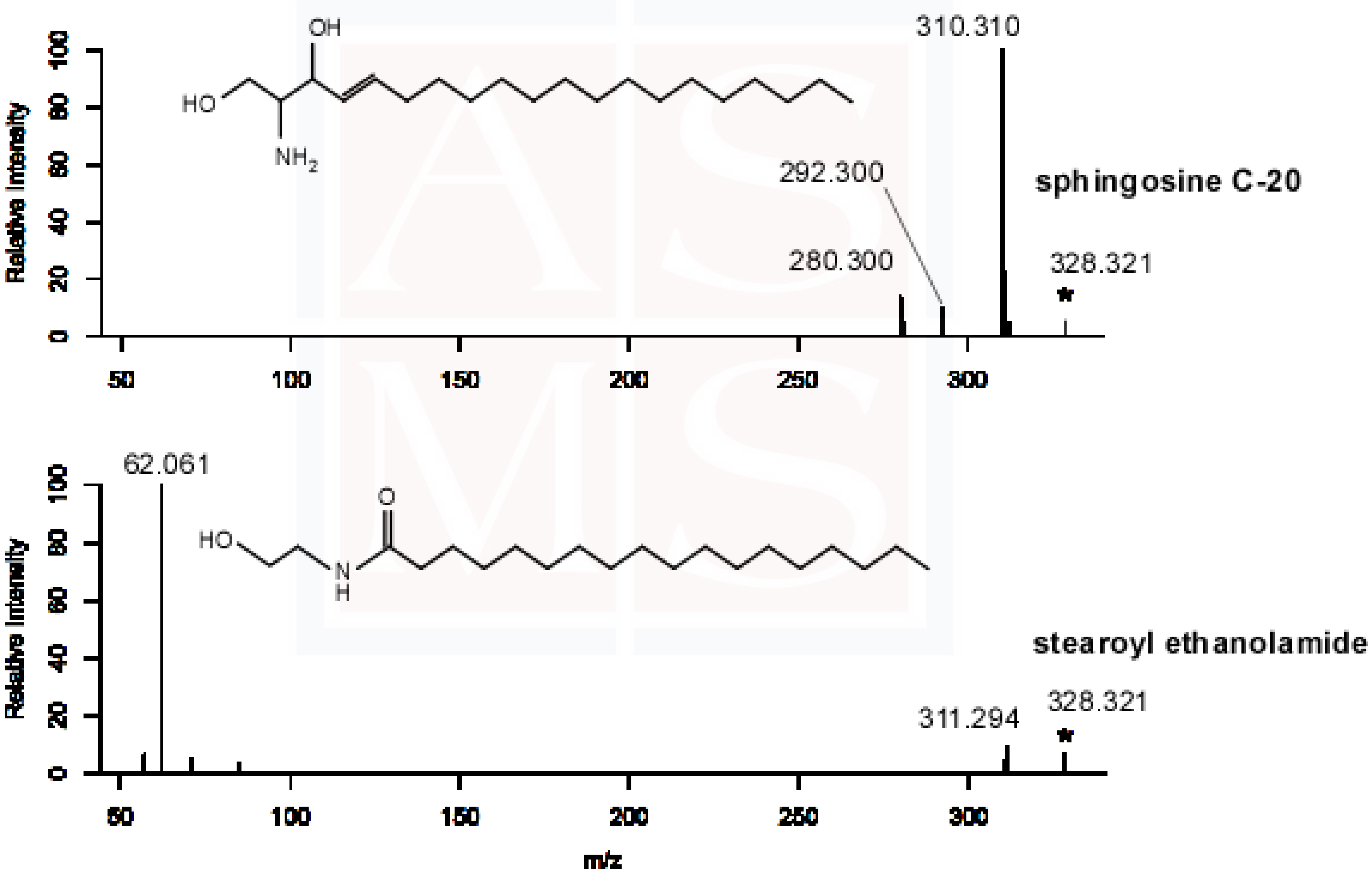
Bar coding metabolomics



328.3210 vs 328 → 62 + 44

4. Expanding targeted analyses

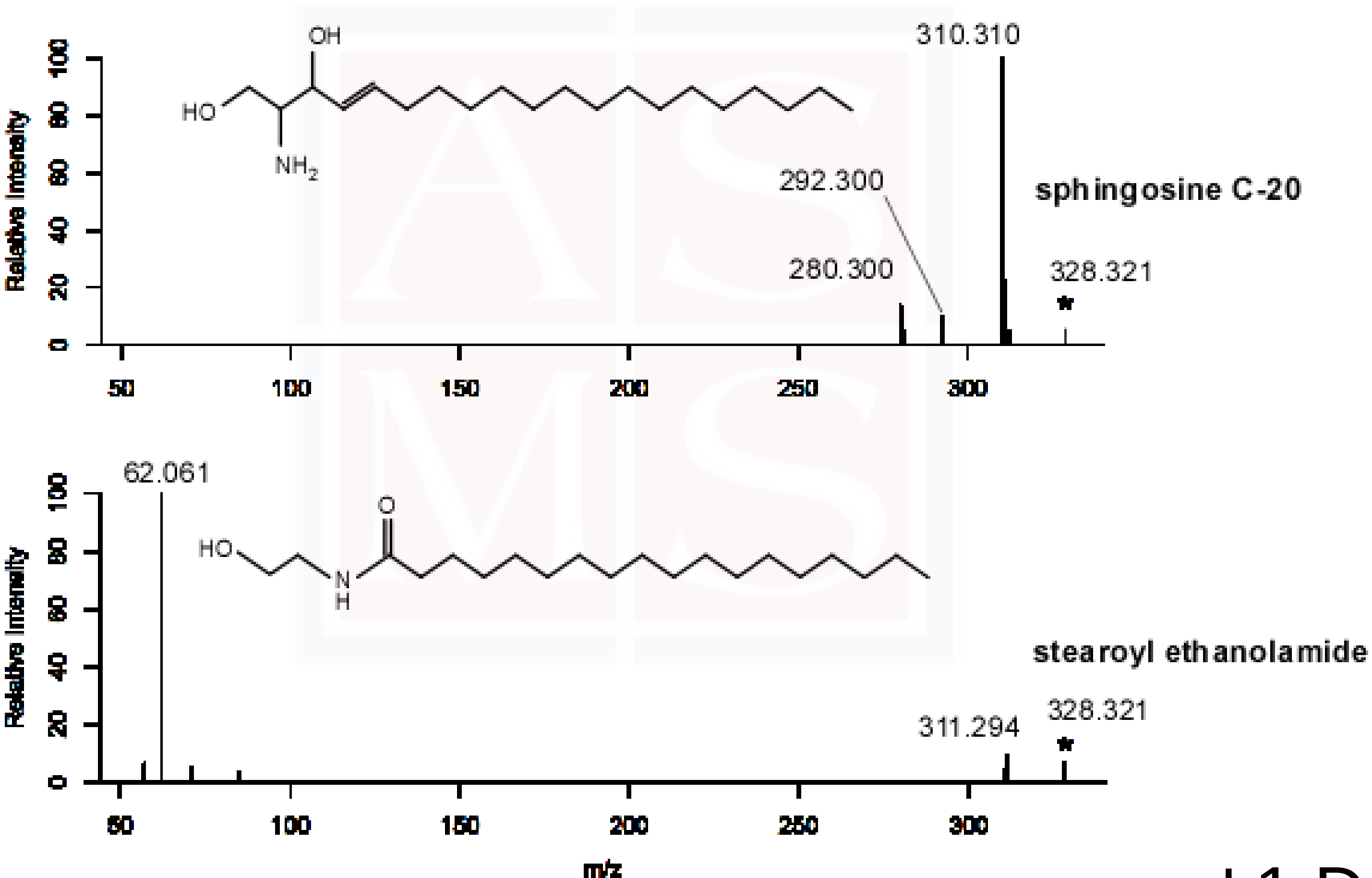
Bar coding metabolomics



328.3210 vs 328 → 62 + 44

4. Expanding targeted analyses

Bar coding metabolomics

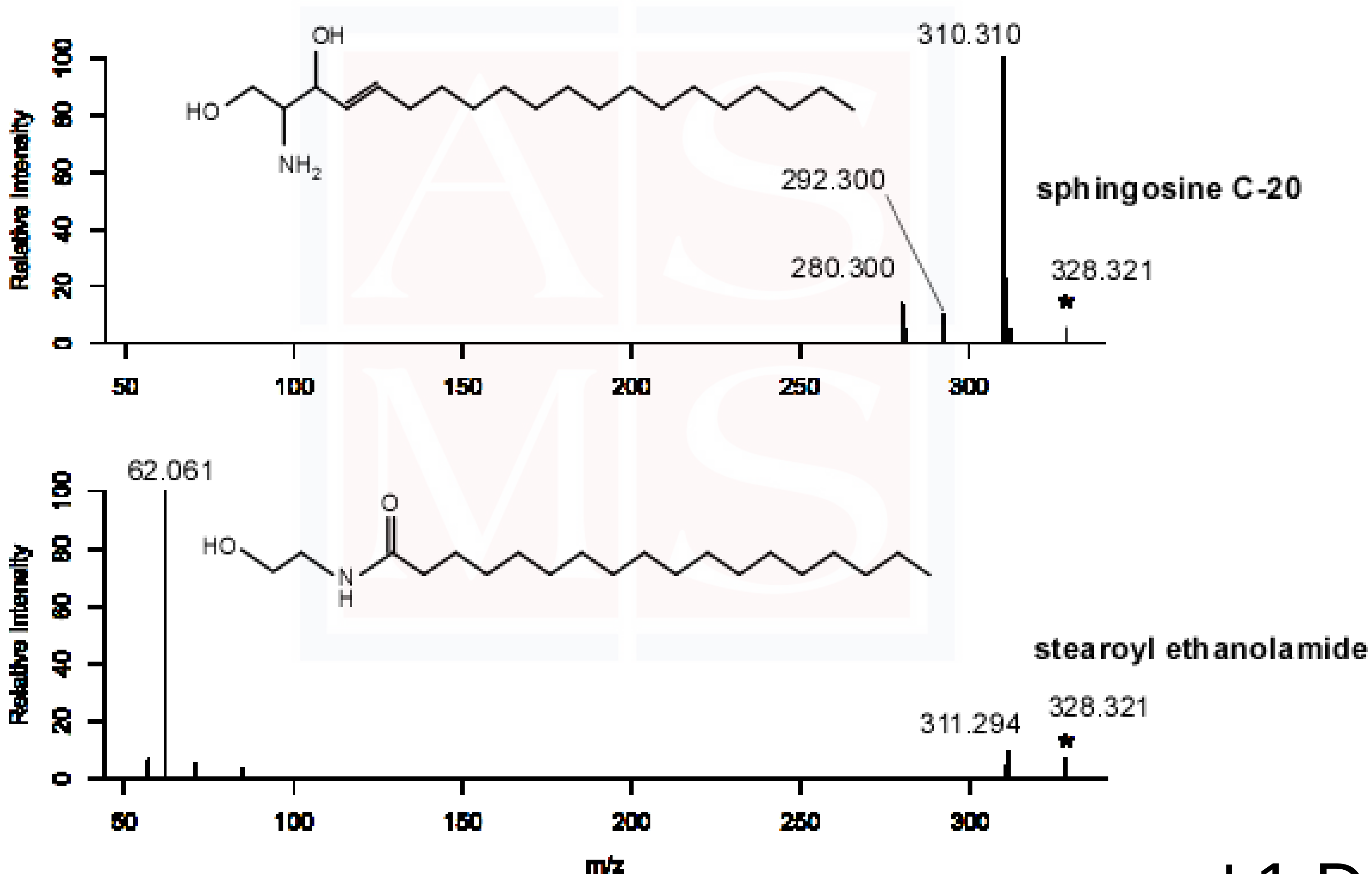


±1 Da

328.3210 vs 328 → 62 + 44

4. Expanding targeted analyses

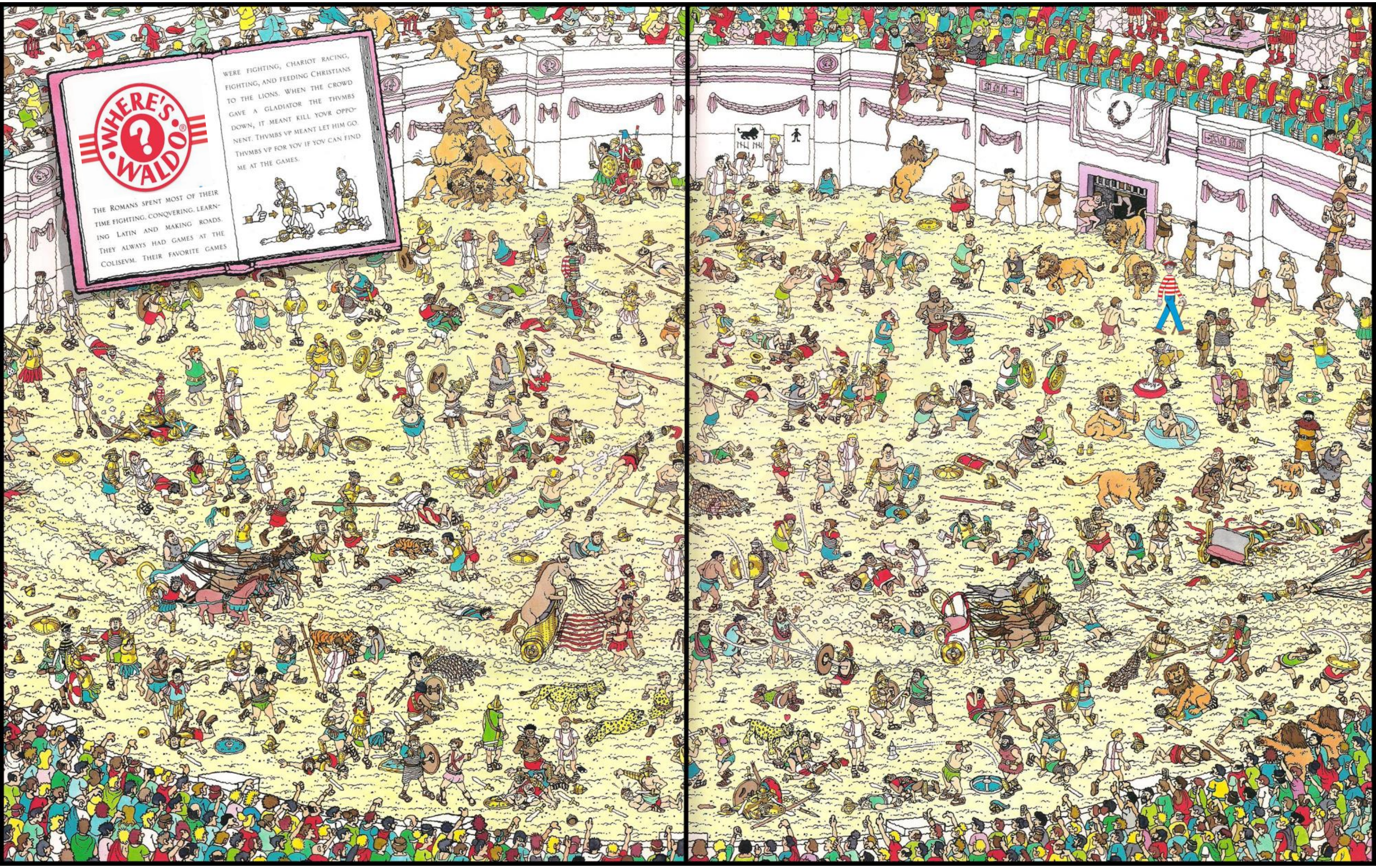
Bar coding metabolomics

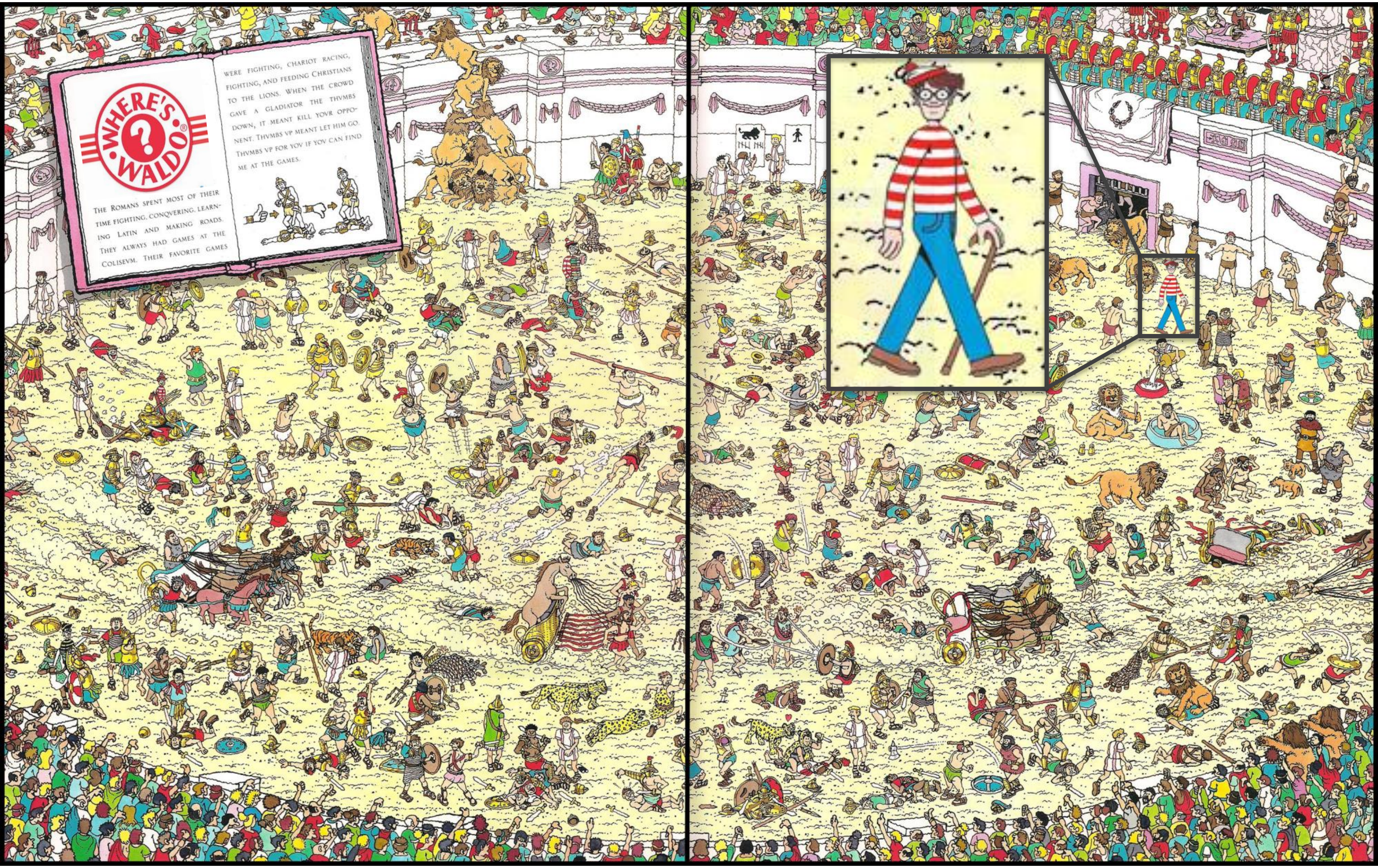


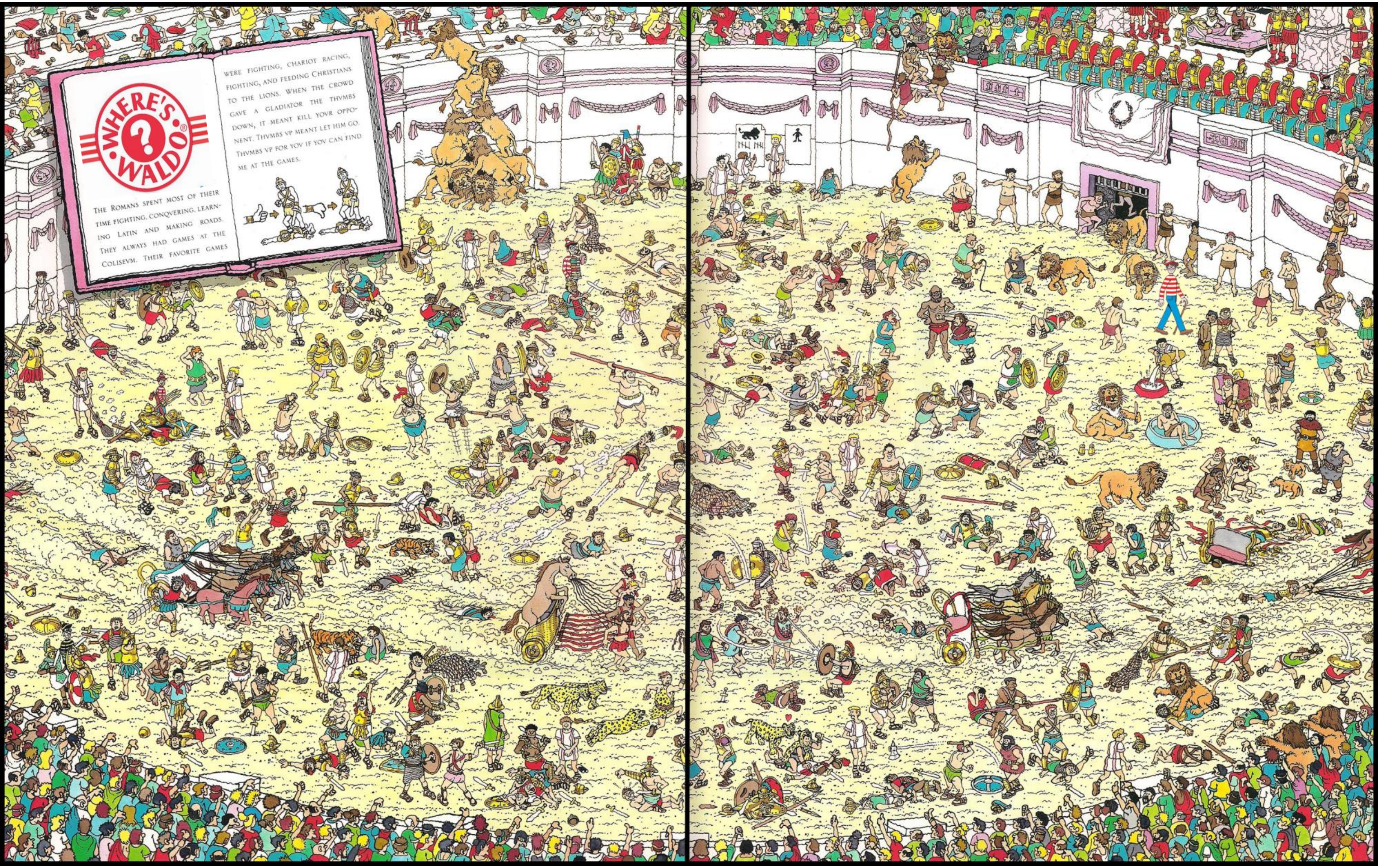
328.3210 vs 328 → 62 + 44

±1 Da

±4 Da


WHERE'S WALDO®
THE ROMANS SPENT MOST OF THEIR TIME FIGHTING, CONQVERING, LEARNING LATIN AND MAKING ROADS. THEY ALWAYS HAD GAMES AT THE COLISEVM. THEIR FAVORITE GAMES WERE FIGHTING, CHARIOT RACING, FIGHTING, AND FEEDING CHRISTIANS TO THE LIONS. WHEN THE CROWD GAVE A GLADIATOR THE THVMBS DOWN, IT MEANT KILL YOVR OPPONENT. THVMBS VP MEANT LET HIM GO. THVMBS VP FOR YOV IF YOV CAN FIND ME AT THE GAMES.


WHERE'S WALDO
THE ROMANS SPENT MOST OF THEIR TIME FIGHTING, CONQVERING, LEARNING LATIN AND MAKING ROADS. THEY ALWAYS HAD GAMES AT THE COLISEVM. THEIR FAVORITE GAMES WERE FIGHTING, CHARIOT RACING, FIGHTING, AND FEEDING CHRISTIANS TO THE LIONS. WHEN THE CROWD GAVE A GLADIATOR THE THVMBS DOWN, IT MEANT KILL YOVR OPPONENT. THVMBS VP MEANT LET HIM GO. THVMBS VP FOR YOV IF YOV CAN FIND ME AT THE GAMES.

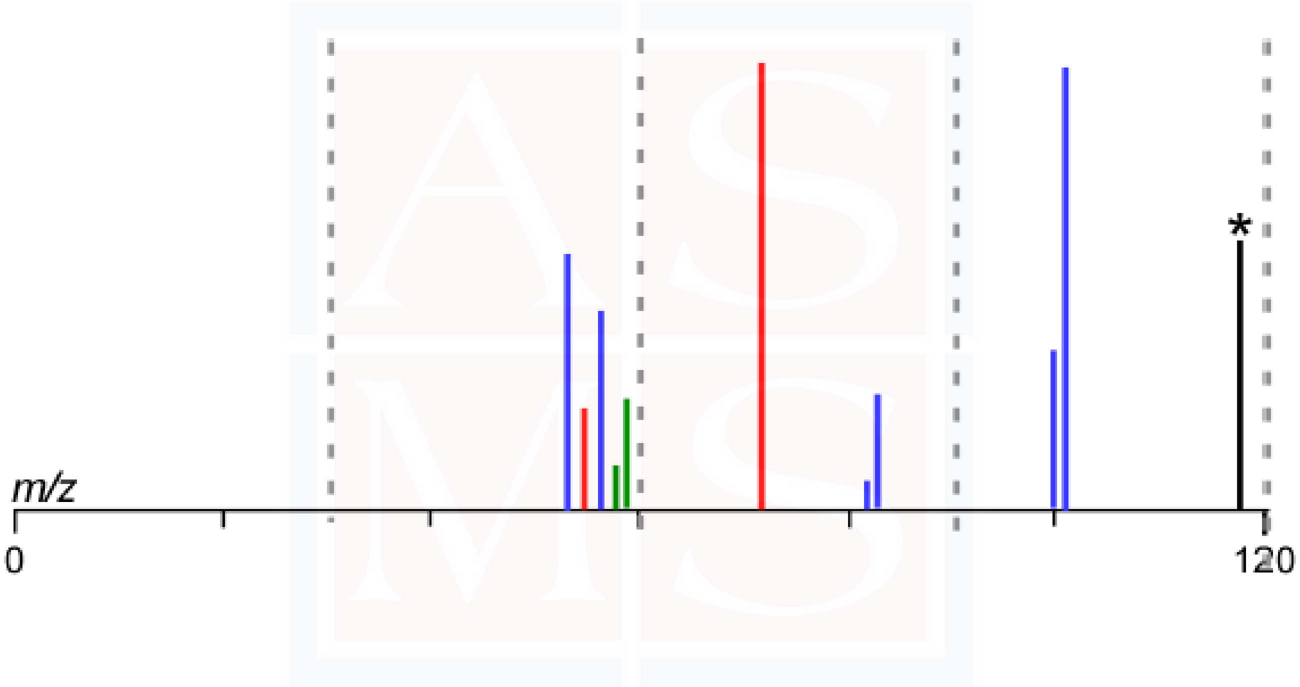

WHERE'S WALDO®
THE ROMANS SPENT MOST OF THEIR TIME FIGHTING, CONQVERING, LEARNING LATIN AND MAKING ROADS. THEY ALWAYS HAD GAMES AT THE COLISEVM. THEIR FAVORITE GAMES WERE FIGHTING, CHARIOT RACING, FIGHTING, AND FEEDING CHRISTIANS TO THE LIONS. WHEN THE CROWD GAVE A GLADIATOR THE THVMBS DOWN, IT MEANT KILL YOVR OPPONENT. THVMBS VP MEANT LET HIM GO. THVMBS VP FOR YOV IF YOV CAN FIND ME AT THE GAMES.


WHERE'S WALDO


WHERE'S WALDO

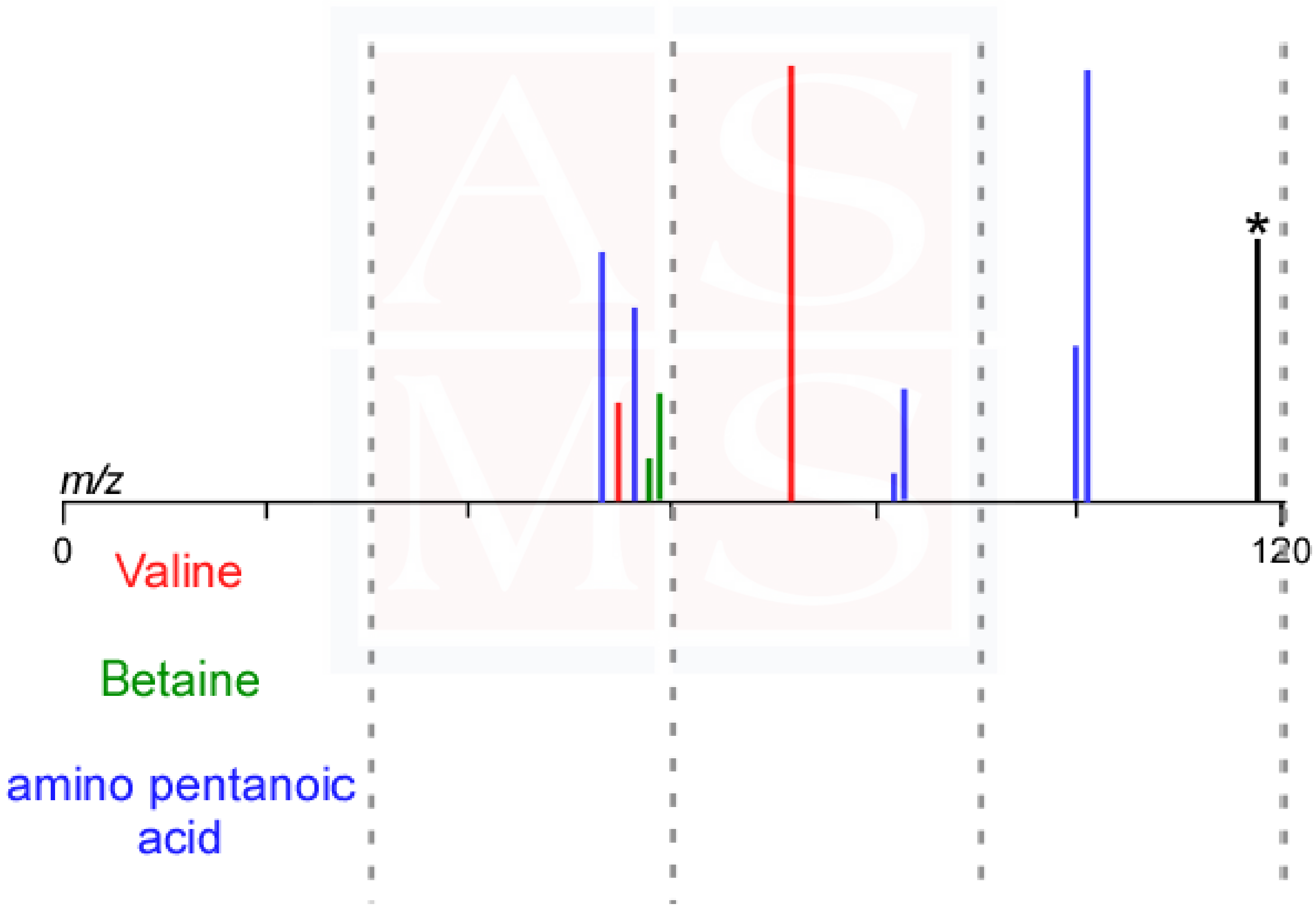
4. Expanding targeted analyses

Bar coding metabolomics



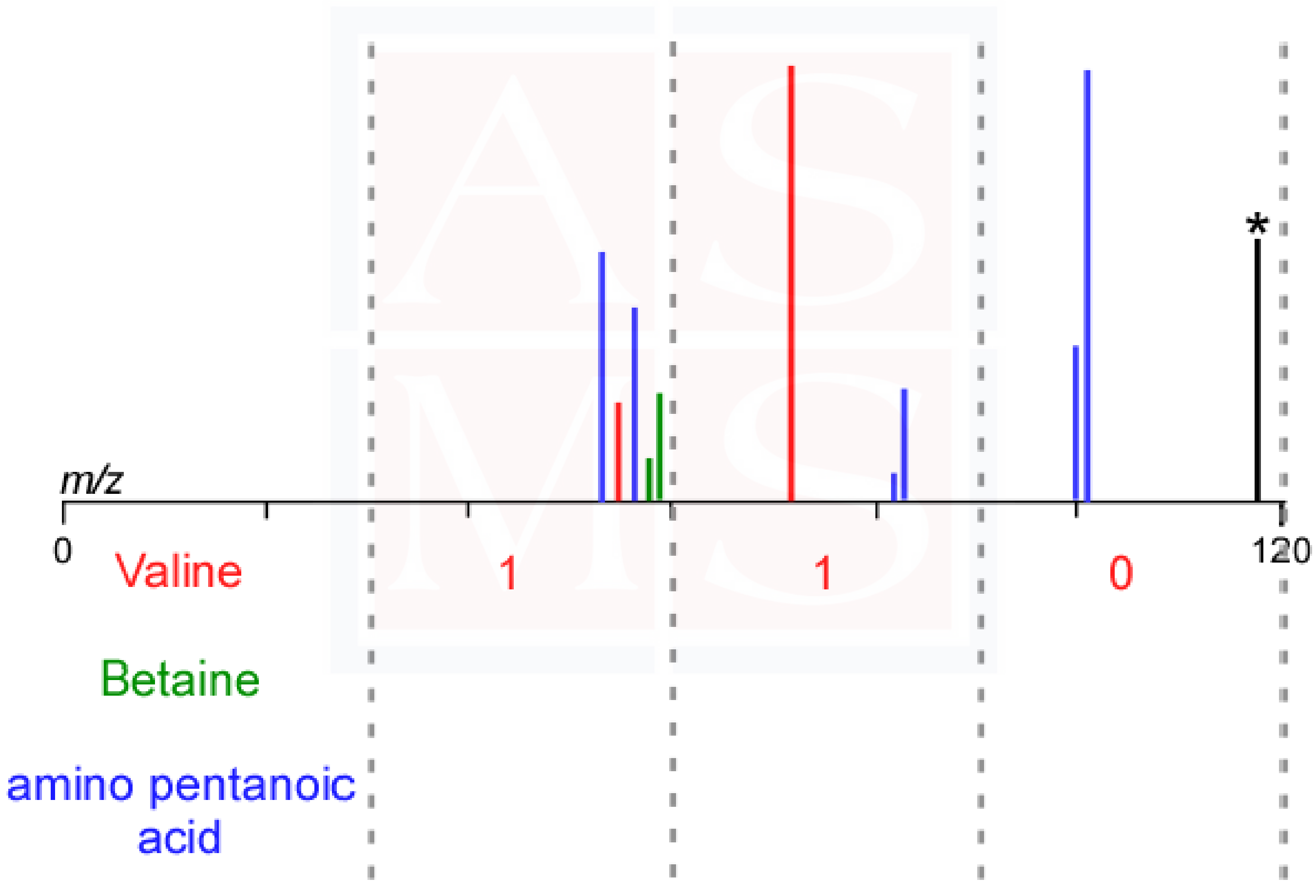
4. Expanding targeted analyses

Bar coding metabolomics



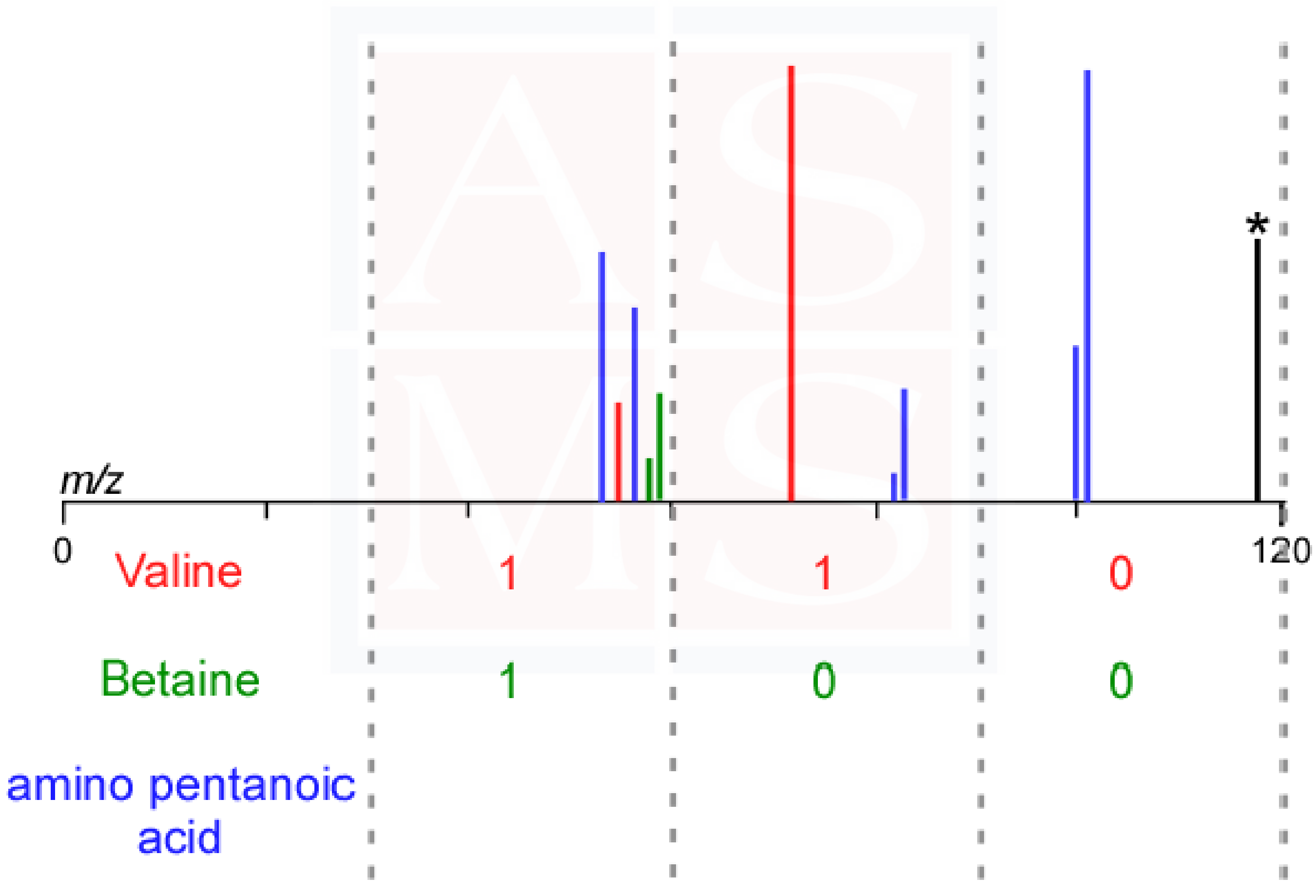
4. Expanding targeted analyses

Bar coding metabolomics



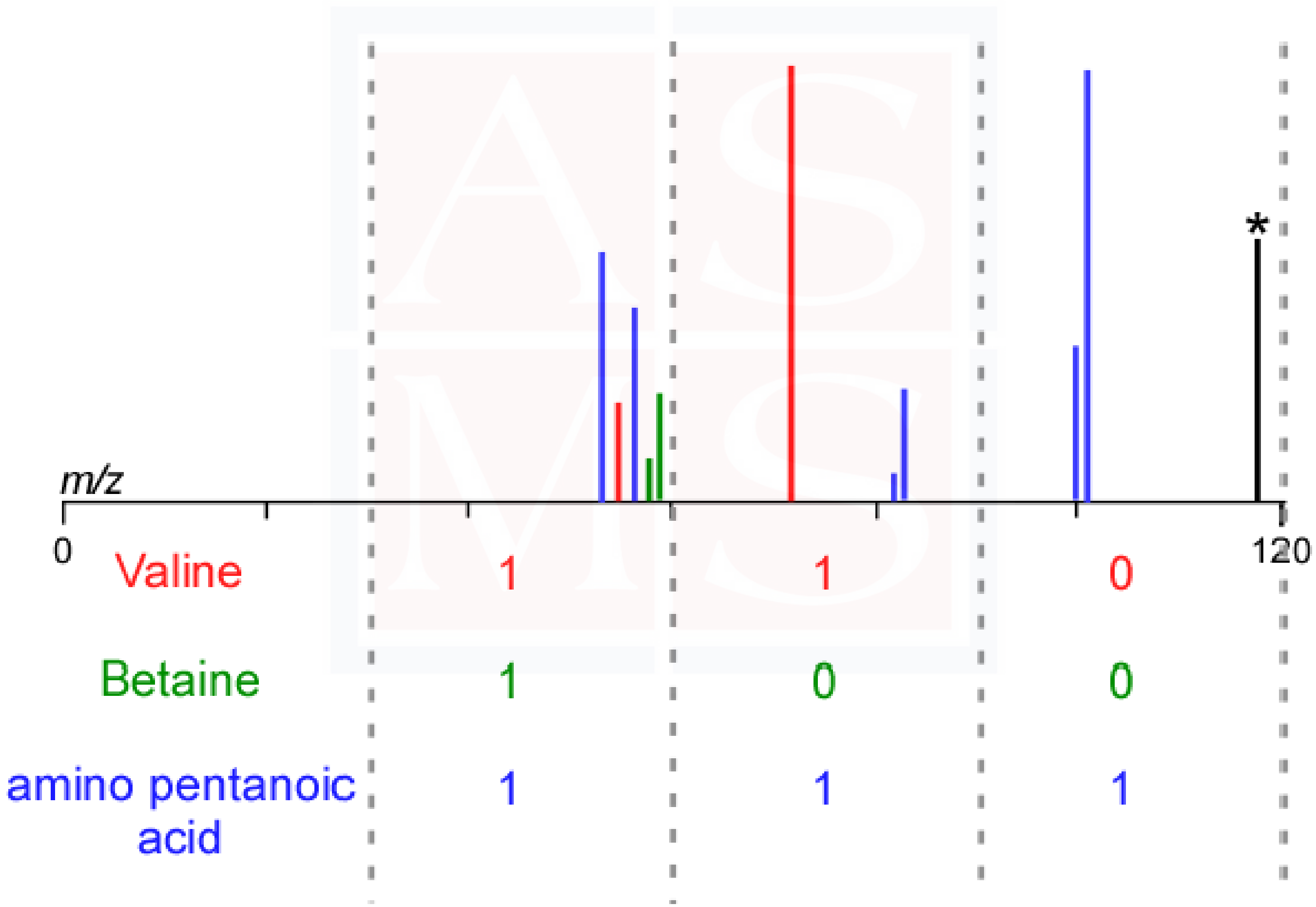
4. Expanding targeted analyses

Bar coding metabolomics



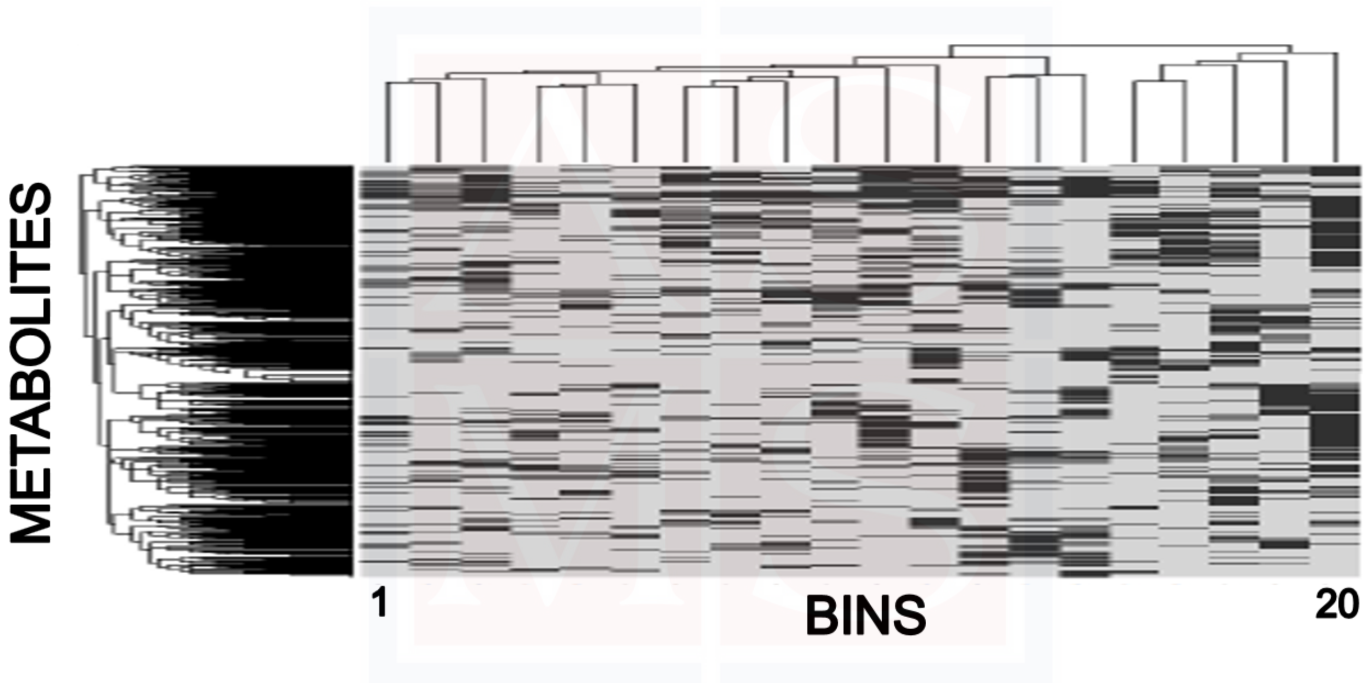
4. Expanding targeted analyses

Bar coding metabolomics



4. Expanding targeted analyses

Bar coding metabolomics



20 bins optimal encode 1,048,576 metabolites in theory and 241,081 in practice

4. Expanding targeted analyses

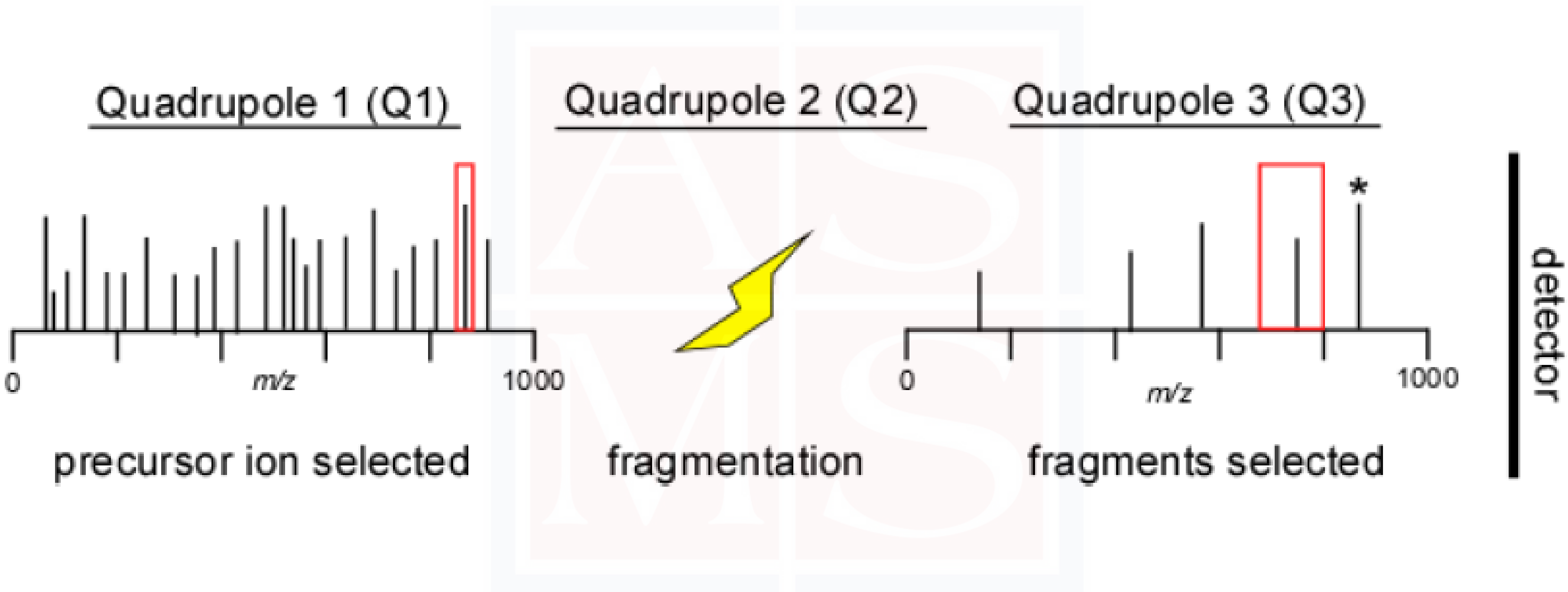
Bar coding metabolomics

Bin ID	Lower limit *(m/z)*	Upper limit *(m/z)*	Bin ID	Lower limit *(m/z)*	Upper limit *(m/z)*
1	37.0	41.5	11	110.0	116.2
2	42.0	46.6	12	120.0	126.4
3	55.0	59.9	13	129.0	135.6
4	65.0	70.1	14	136.0	142.8
5	72.0	77.3	15	144.0	150.9
6	79.0	84.4	16	149.0	156.1
7	84.0	89.6	17	159.0	166.3
8	86.0	91.6	18	177.0	184.7
9	91.0	96.7	19	197.0	205.2
10	98.0	103.9	20	262.0	271.7

* Indicates collision energy of 20 V. all other bins use 40 V

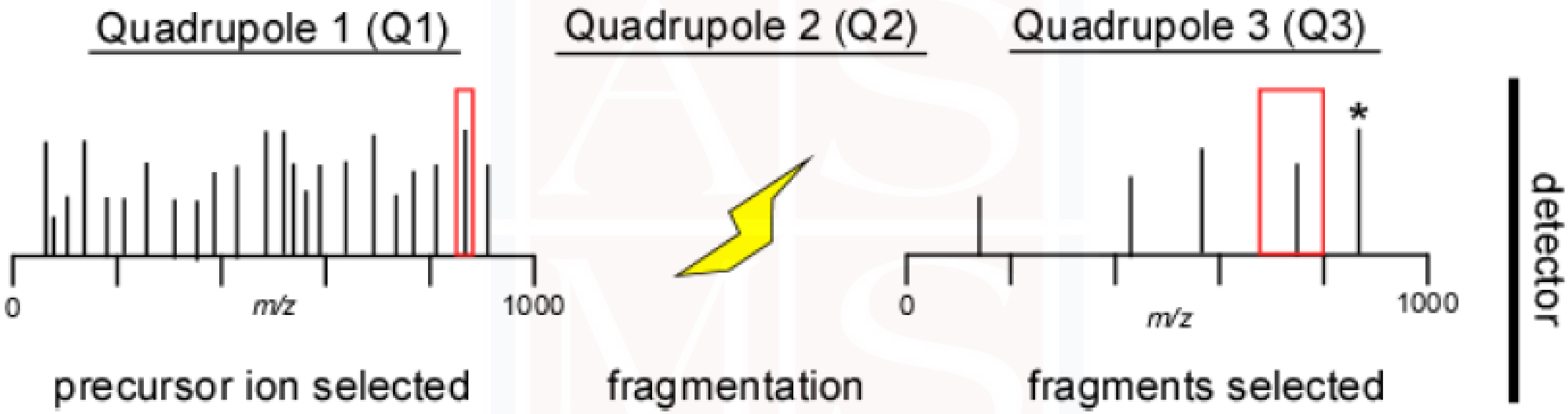
4. Expanding targeted analyses

Bar coding metabolomics



4. Expanding targeted analyses

Bar coding metabolomics



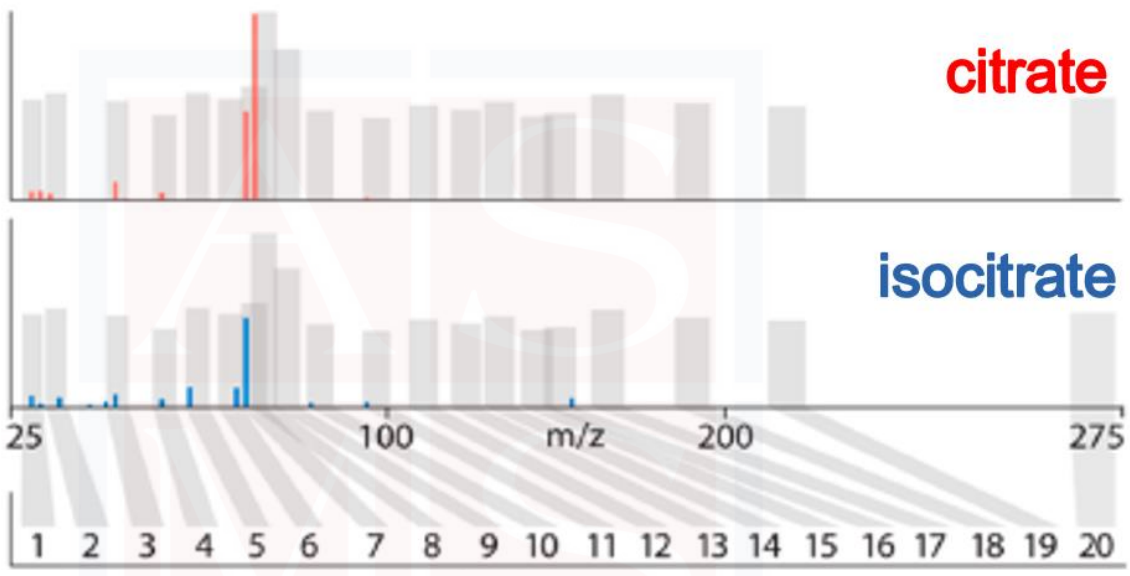
narrow isolation (0.1 Da)

wide-isolation (9 Da)

4. Expanding targeted analyses

Bar coding metabolomics

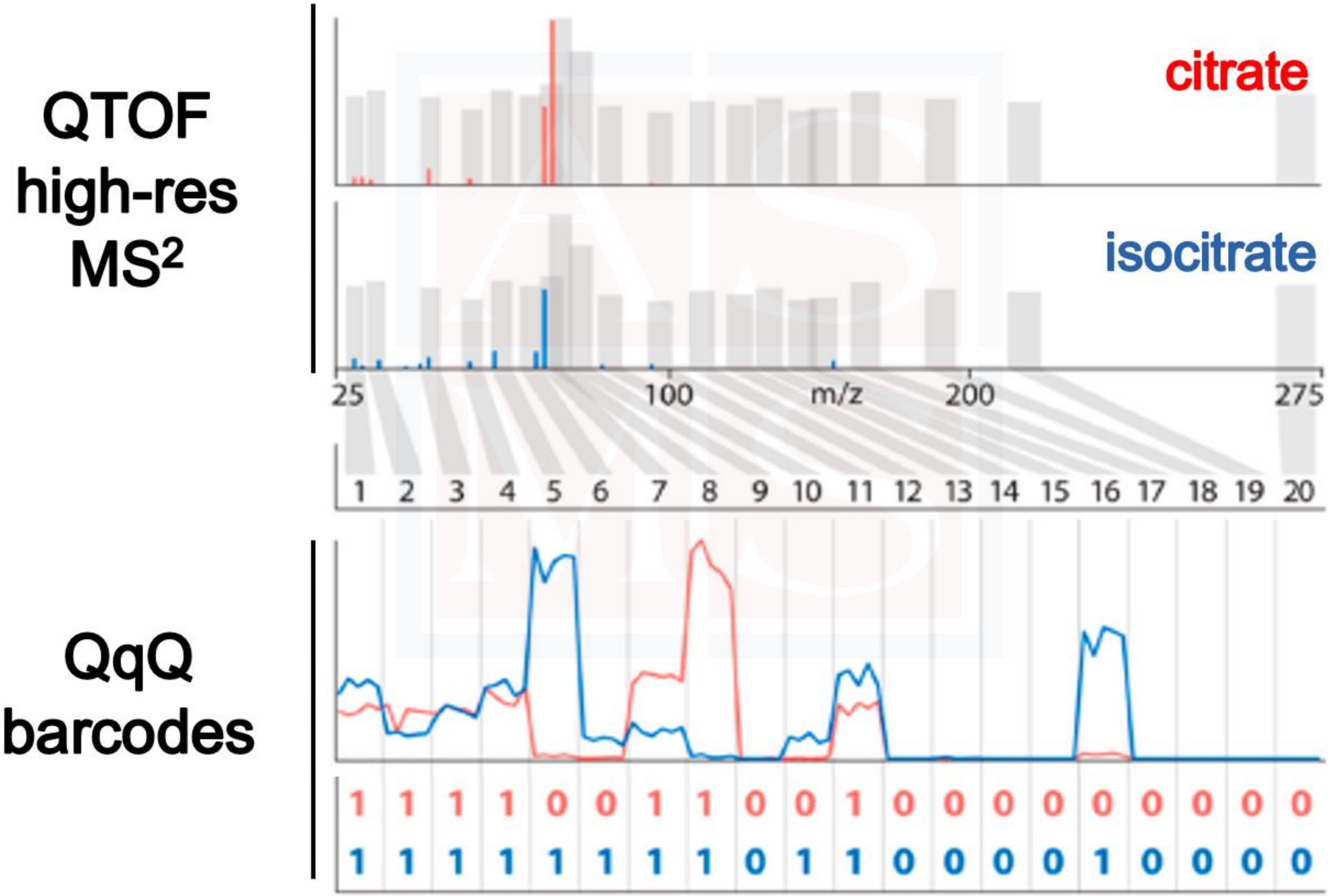
QTOF
high-res
MS^2



Spalding et al., Anal Chem (2016)

4. Expanding targeted analyses

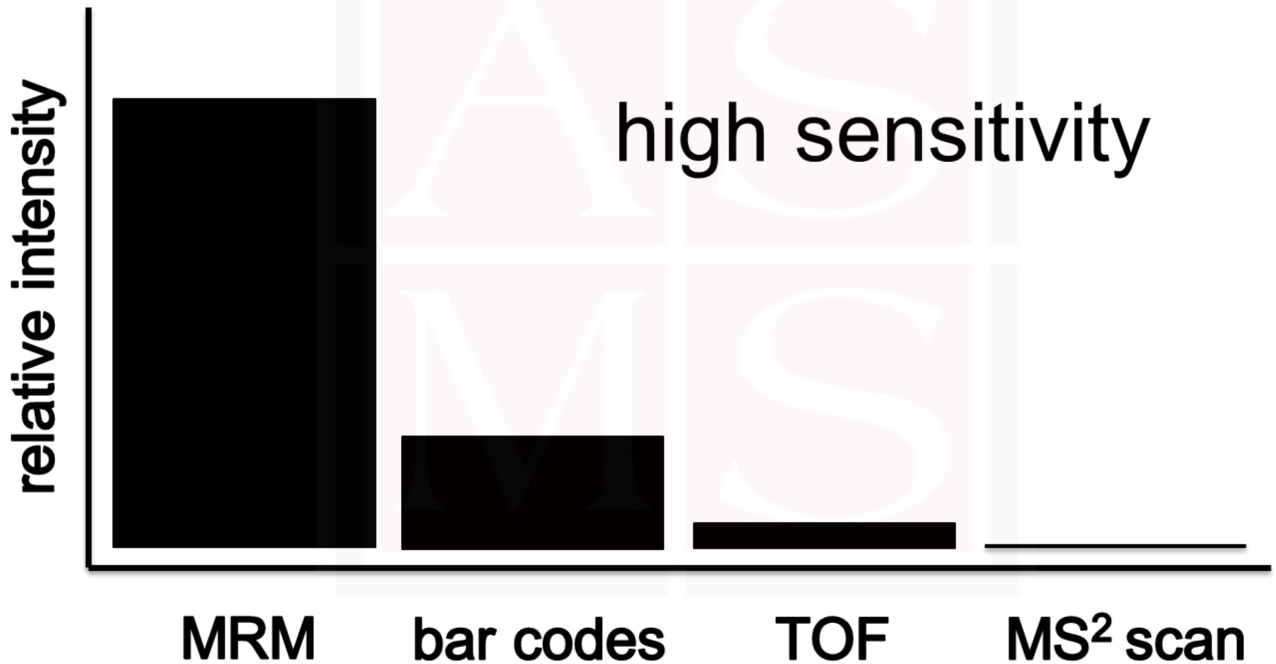
Bar coding metabolomics



Spalding et al., Anal Chem (2016)

4. Expanding targeted analyses

Bar coding metabolomics



Spalding et al., Anal Chem (2016)

4. Expanding targeted analyses

Bar coding metabolomics

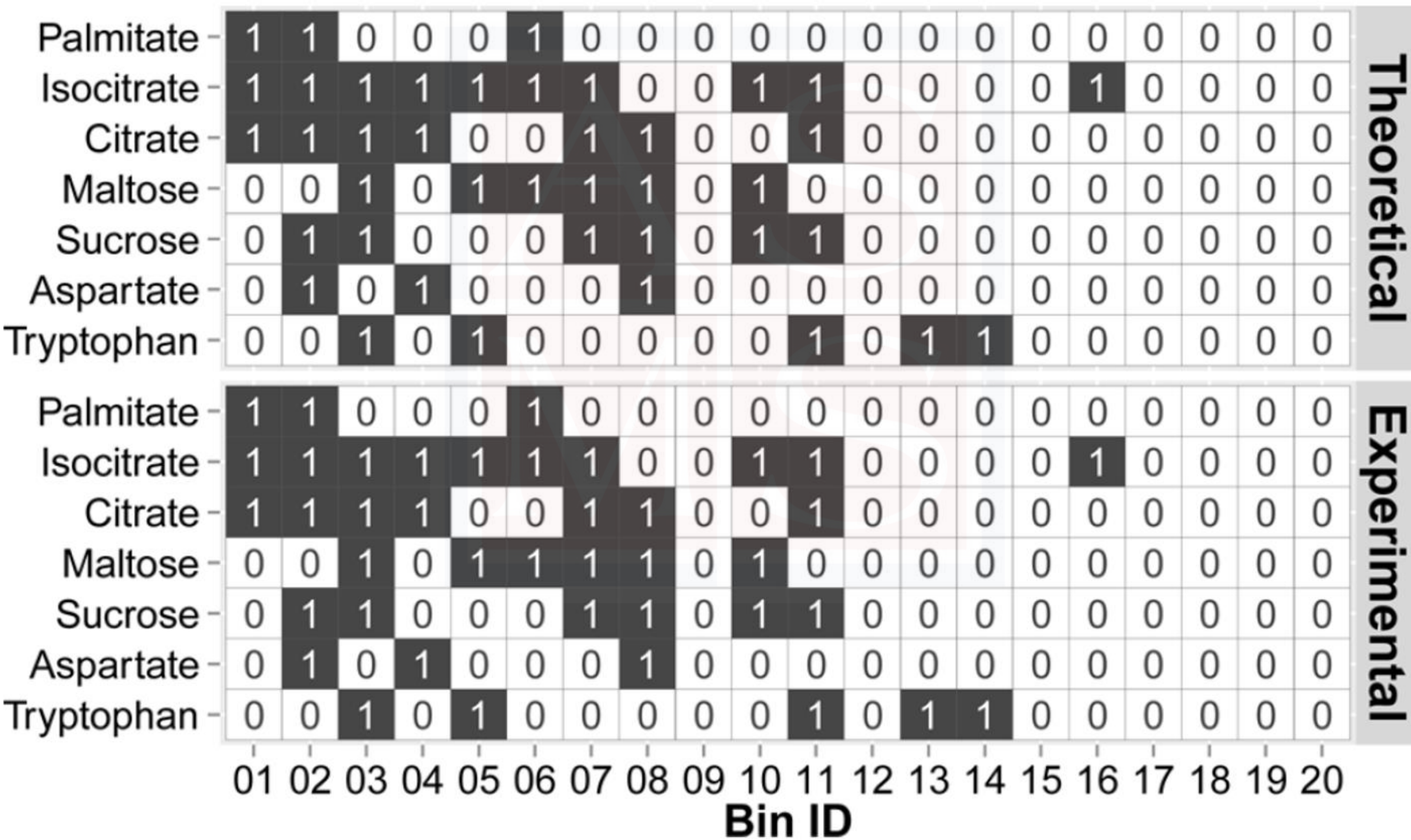
Theoretical

Palmitate	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Isocitrate	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0
Citrate	1	1	1	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
Maltose	0	0	1	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
Sucrose	0	1	1	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0
Aspartate	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Tryptophan	0	0	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0

Spalding et al., Anal Chem (2016)

4. Expanding targeted analyses

Bar coding metabolomics



Spalding et al., Anal Chem (2016)

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

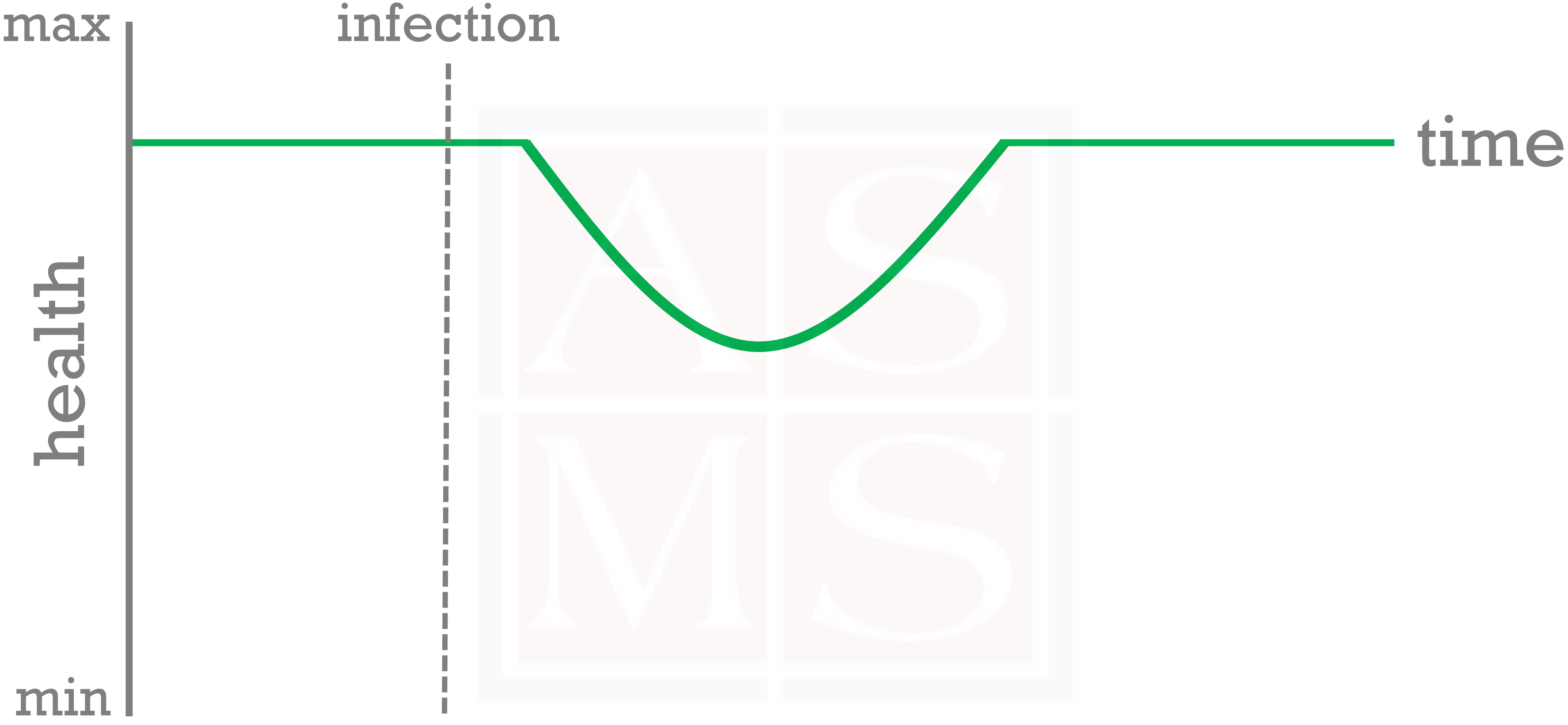
ASMS Metabolomics Short Course

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*

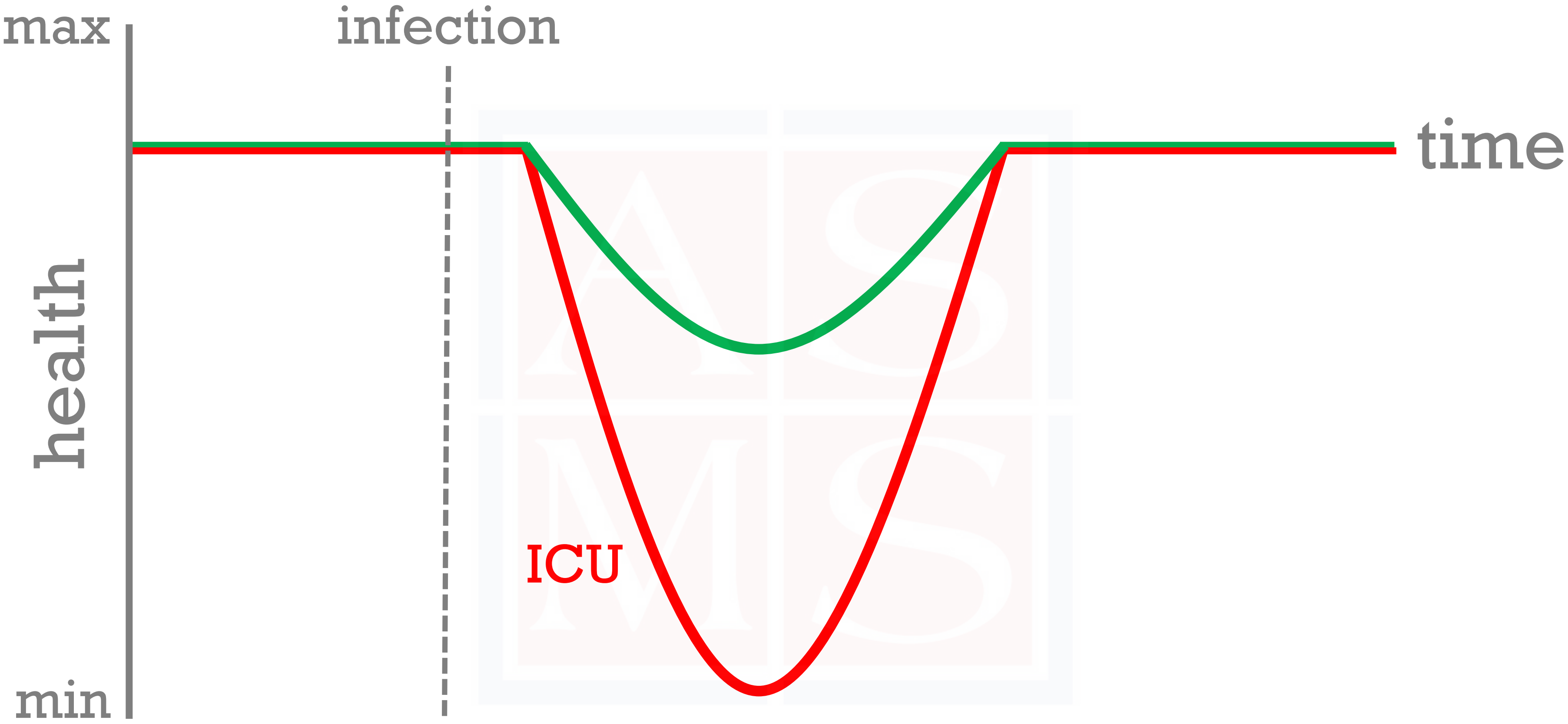
Applications

1. COVID-19

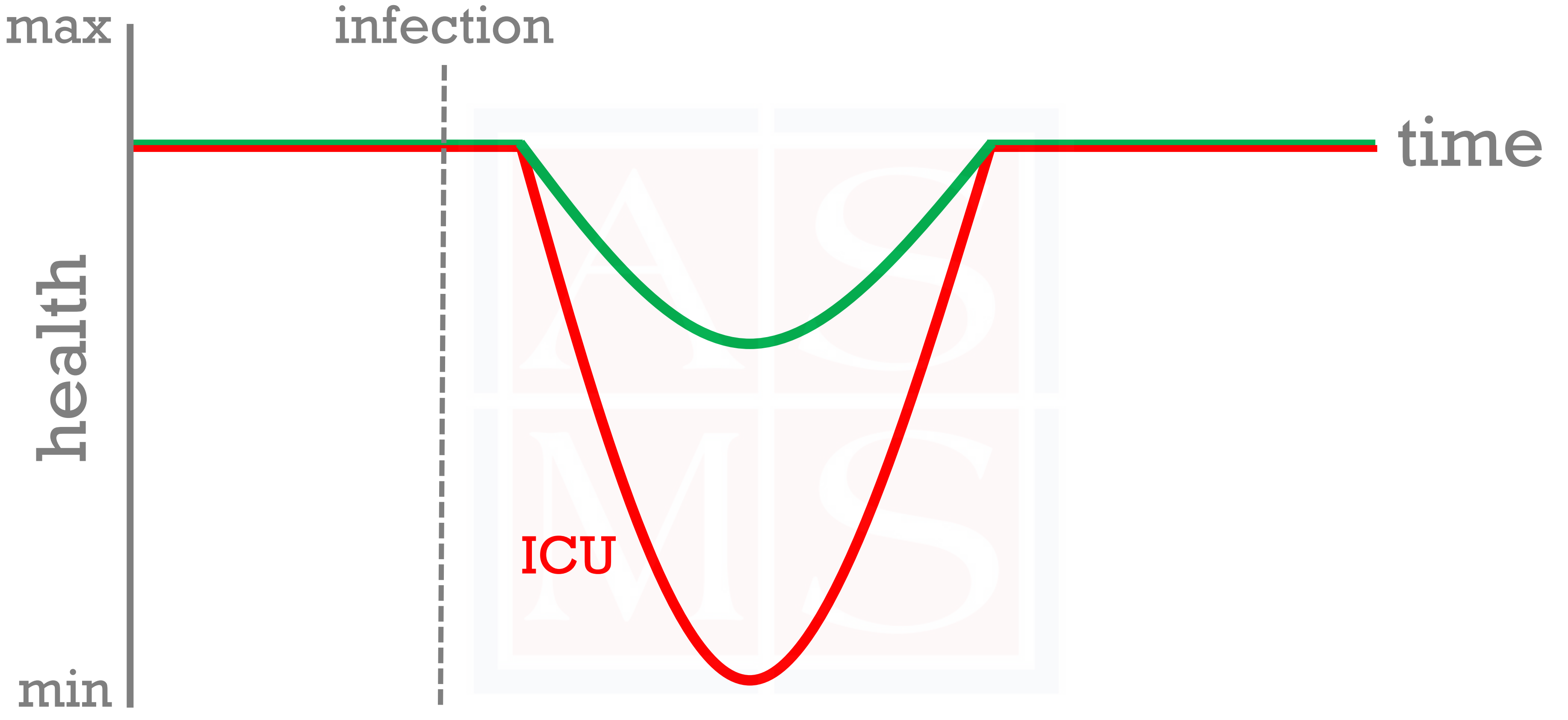
1. COVID-19



1. COVID-19

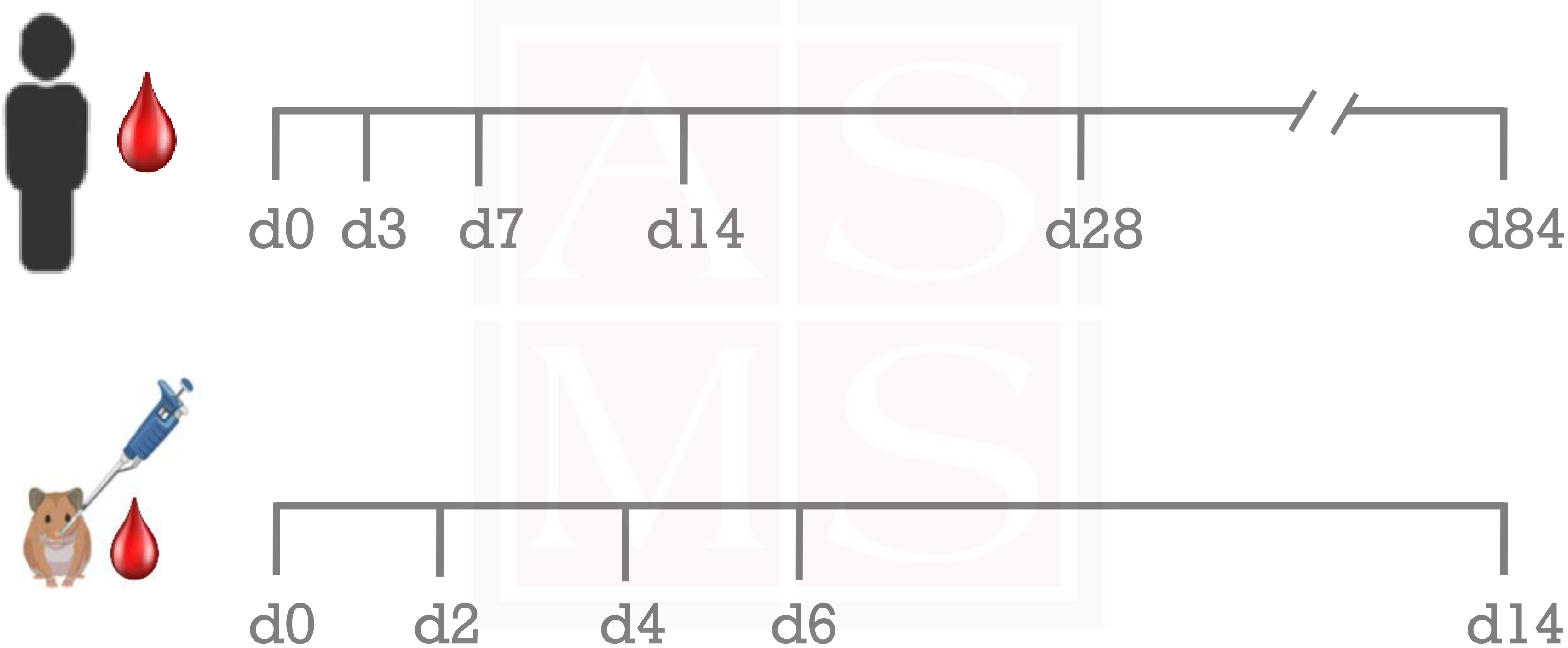


1. COVID-19

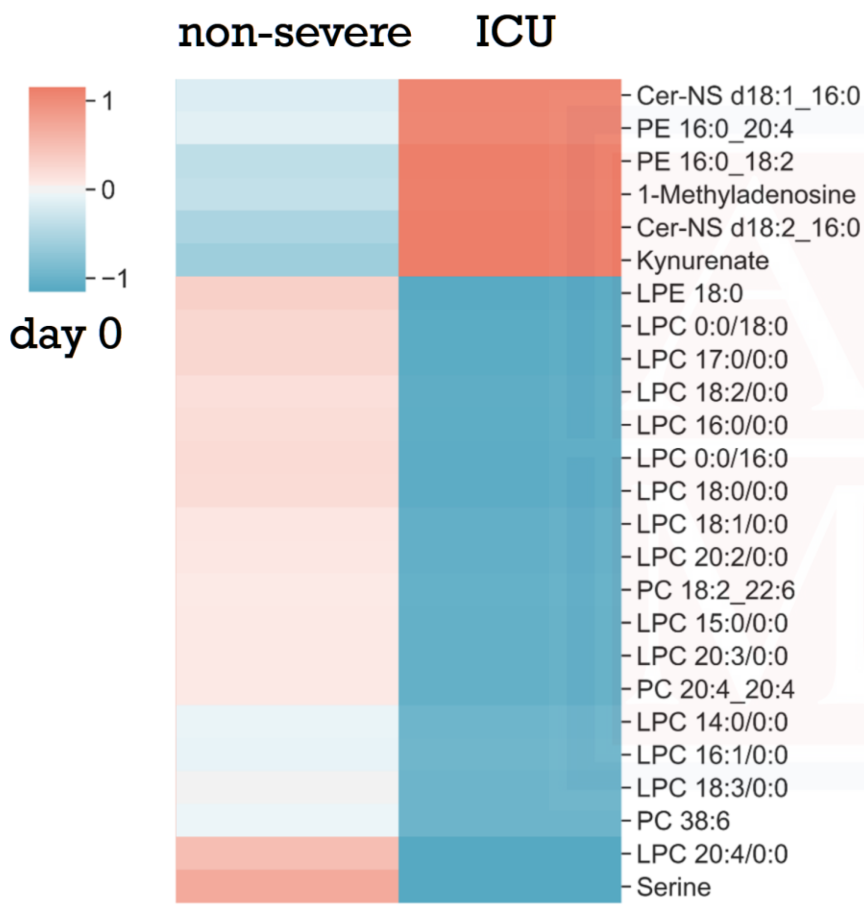


Can metabolism predict who goes to ICU?

1. COVID-19

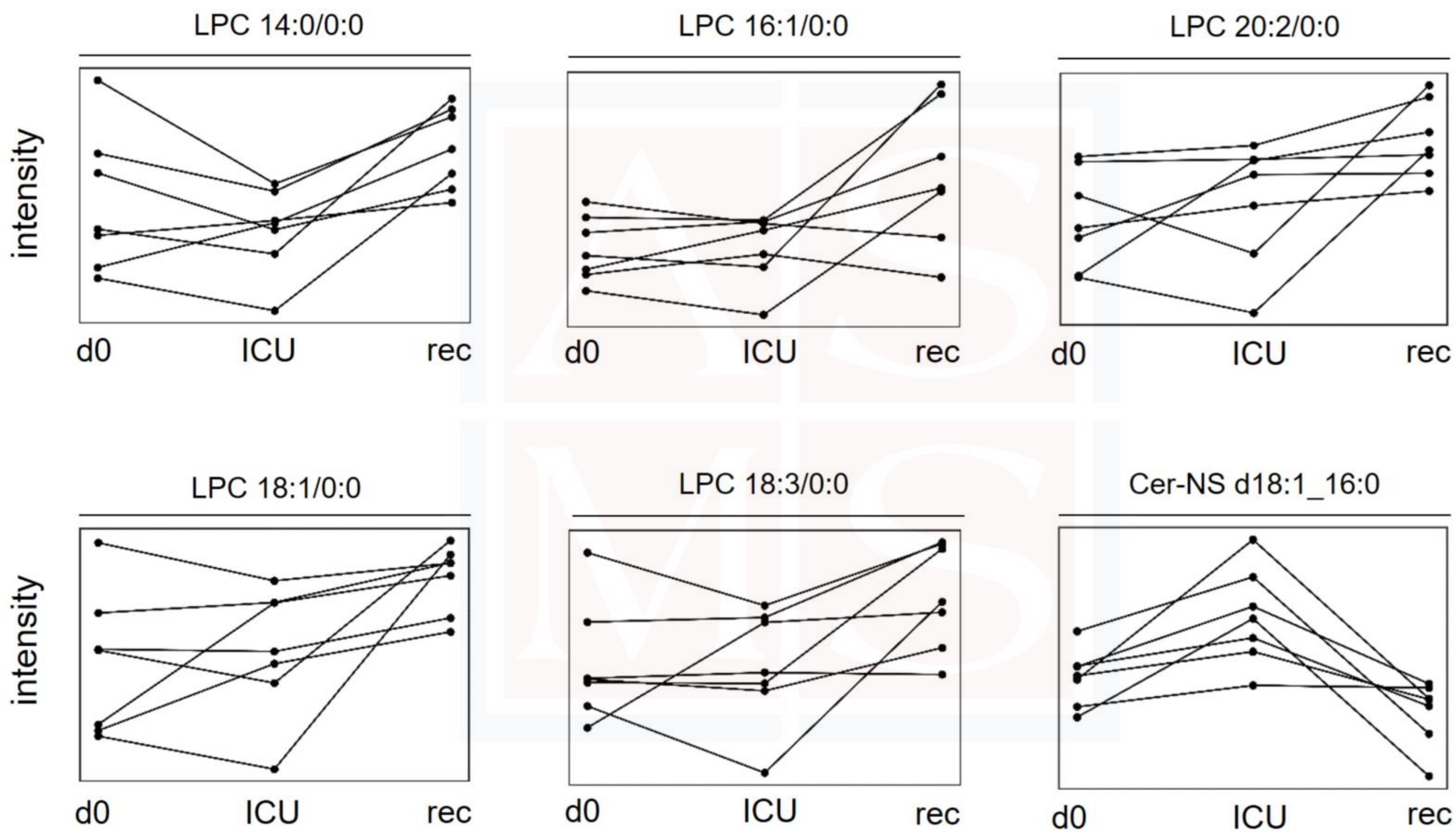


1. COVID-19



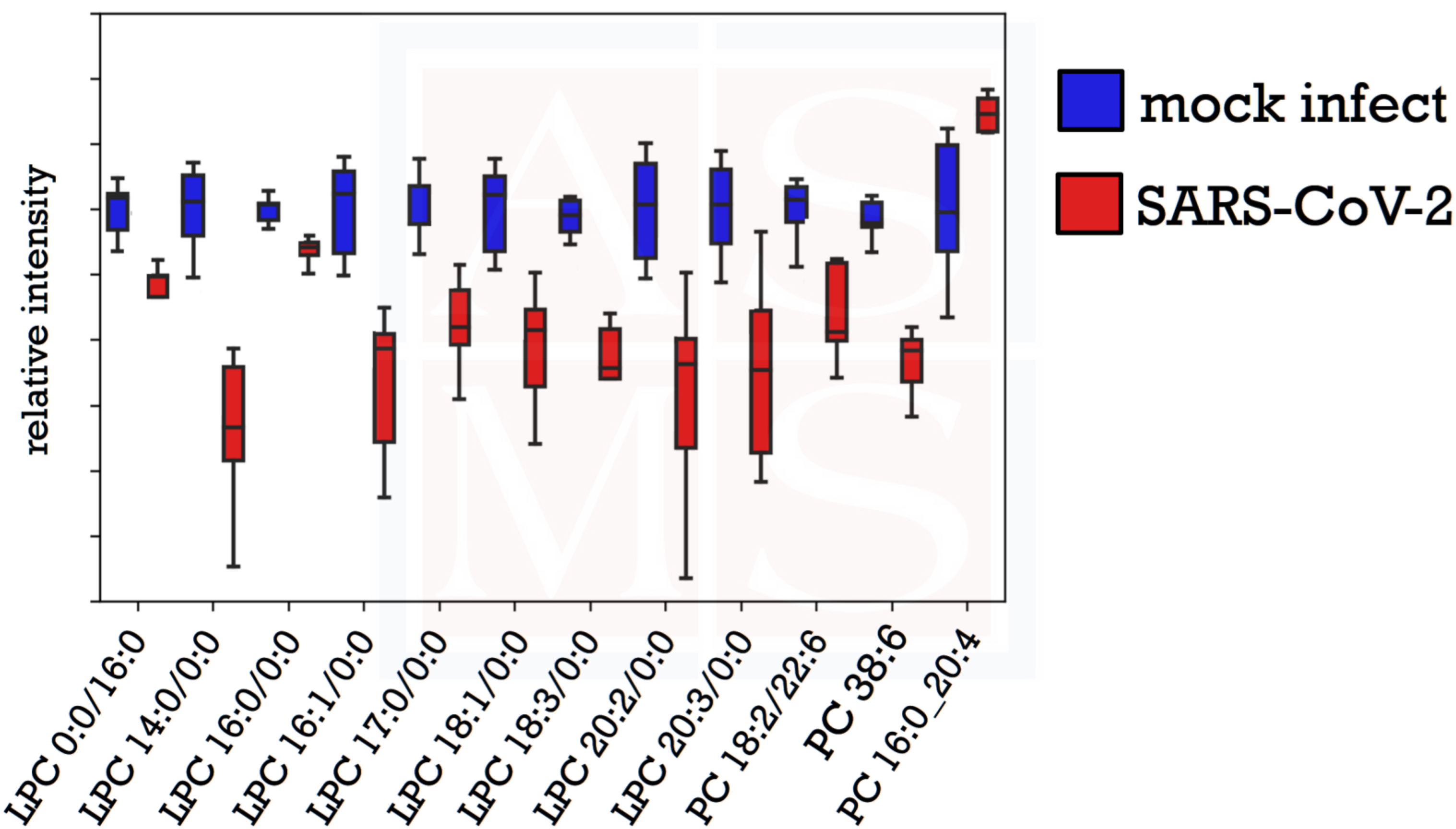
- 25 metabolites predict ICU admission
- Some previously reported
- Lysophosphatidyl -cholines

1. COVID-19


LPC 14:0/0:0
LPC 16:1/0:0
LPC 20:2/0:0
LPC 18:1/0:0
LPC 18:3/0:0
Cer-NS d18:1_16:0
intensity
intensity
d0
ICU
rec

1. COVID-19

patterns conserved in hamsters



1. COVID-19

- Relatively large patient cohorts
- Longitudinal analysis of people
- Validation of results in hamster

2. Chronic pain of neuropathic origin

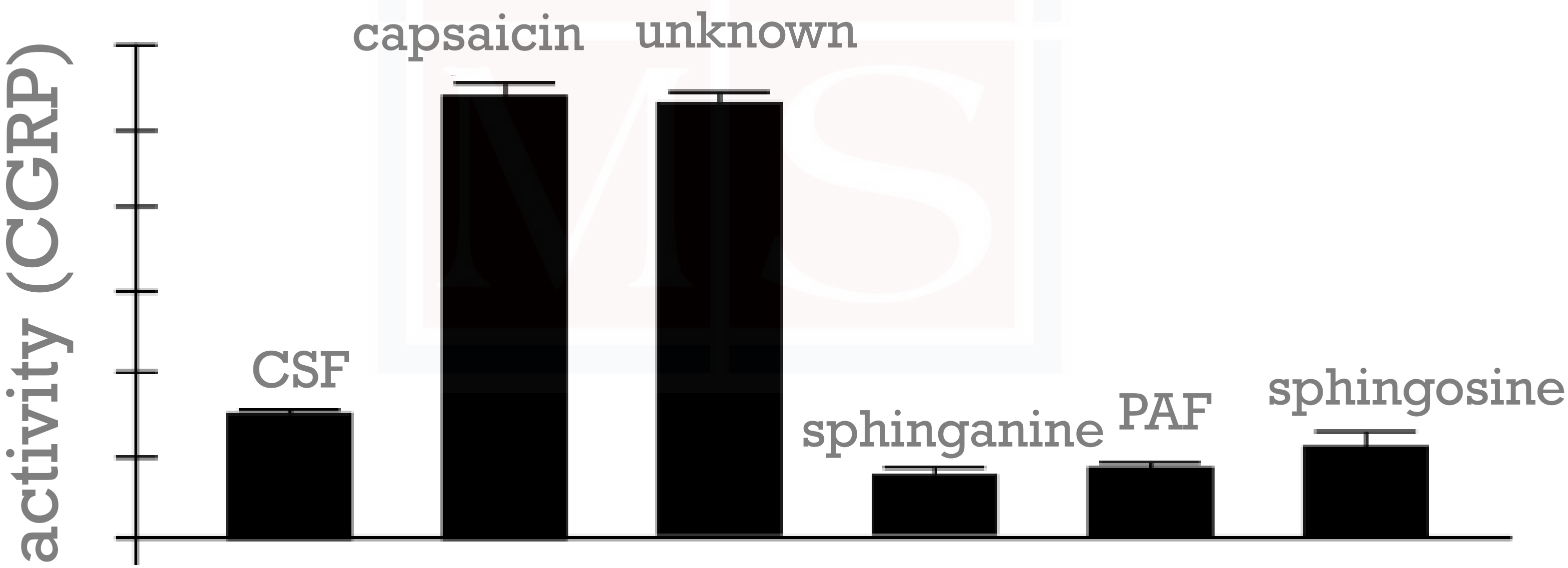
Rat model of pain (TNT)

~800 altered metabolites

2. Chronic pain of neuropathic origin

Rat model of pain (TNT)

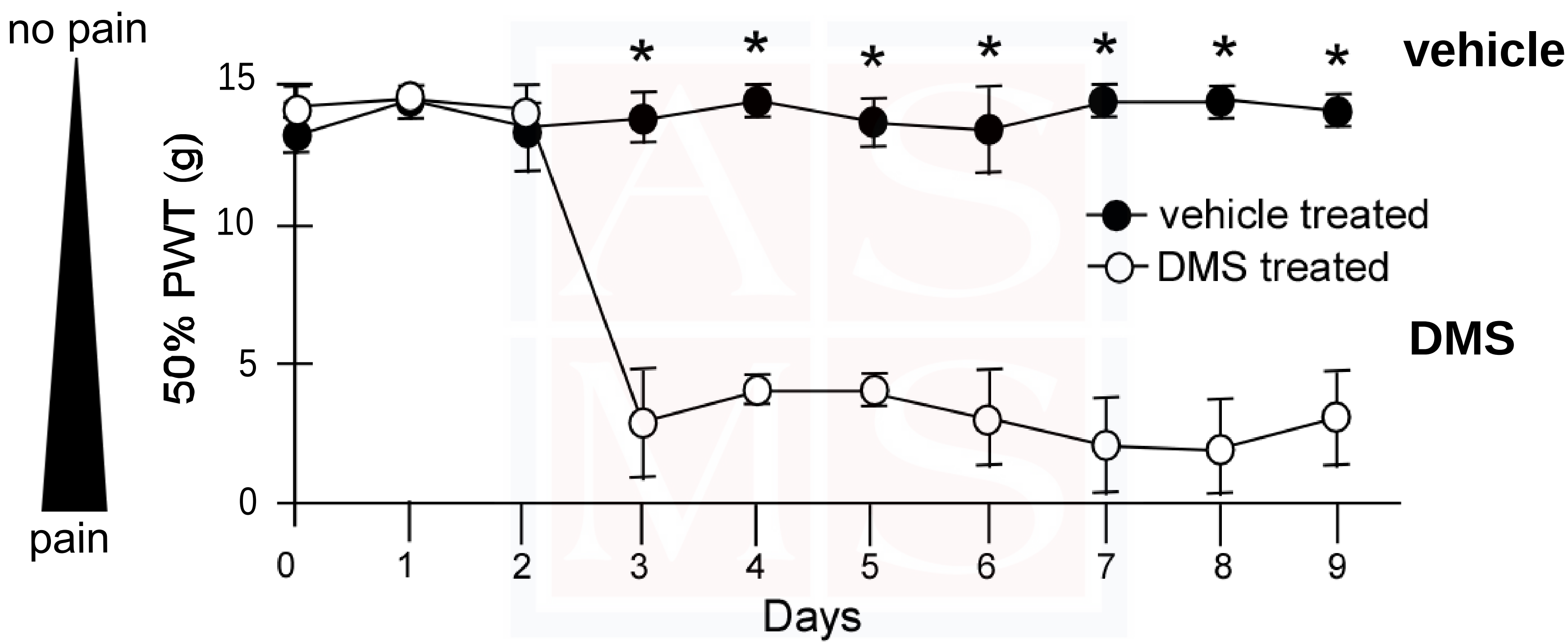
~800 altered metabolites



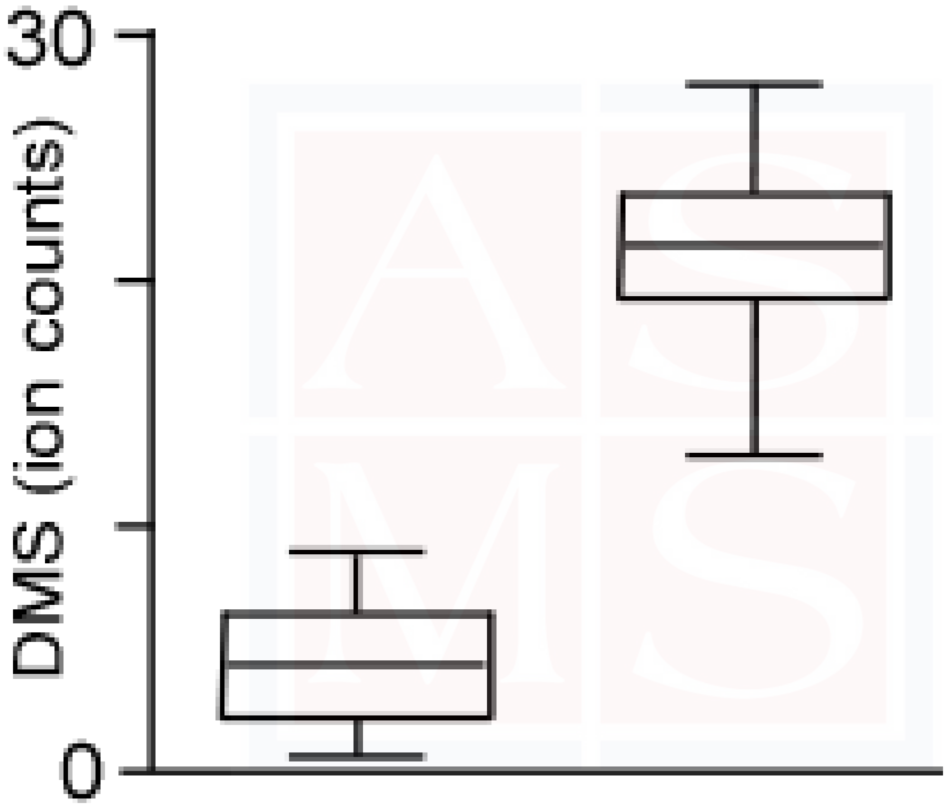
2. Chronic pain of neuropathic origin

Dimethylsphingosine (DMS)

2. Chronic pain of neuropathic origin



2. Chronic pain of neuropathic origin



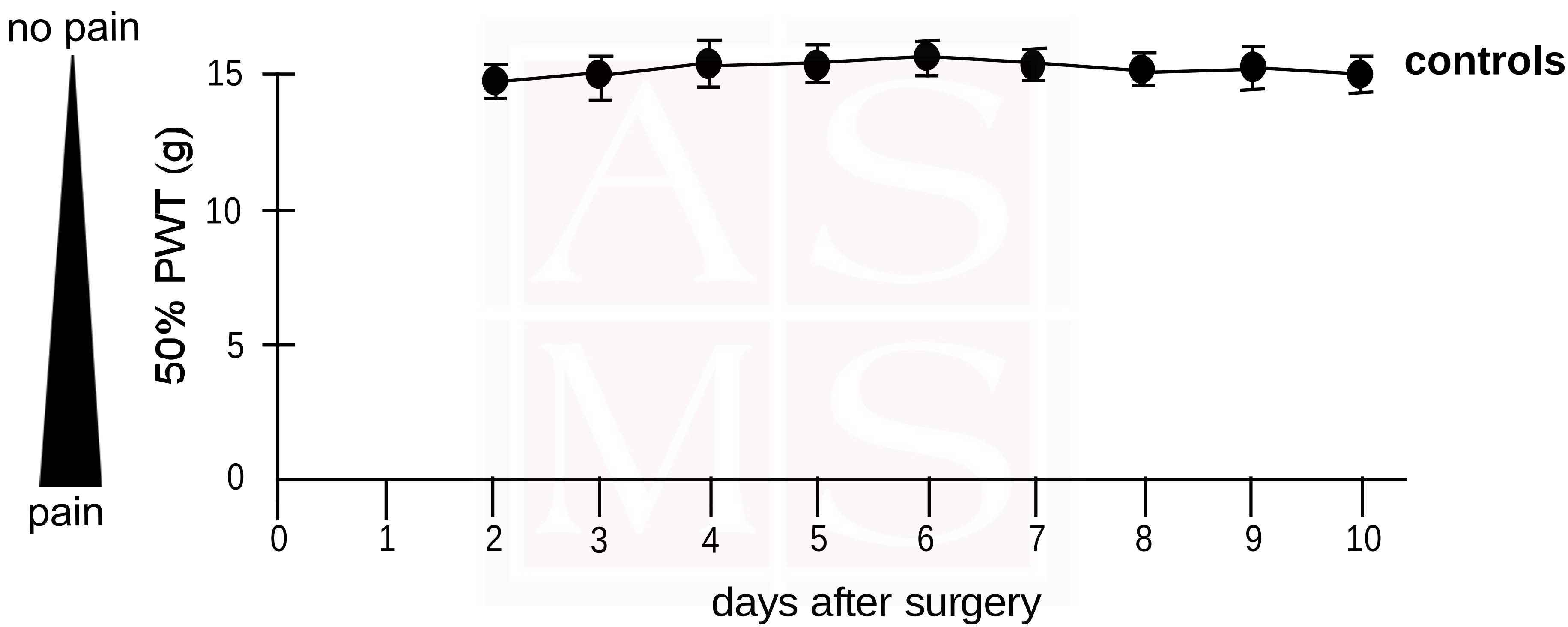
fold change 3.4
p-value 0.007

white matter from control patient

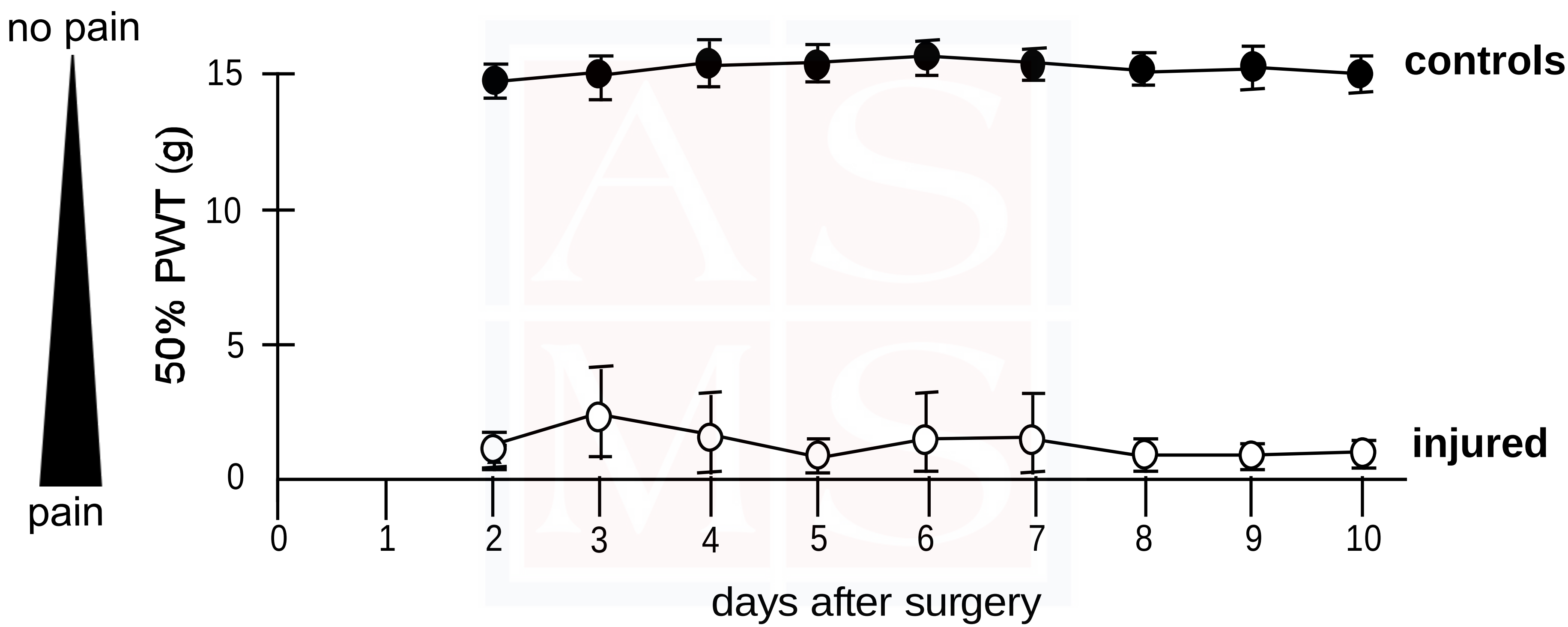
white matter from chronic pain patient

2. Chronic pain of neuropathic origin

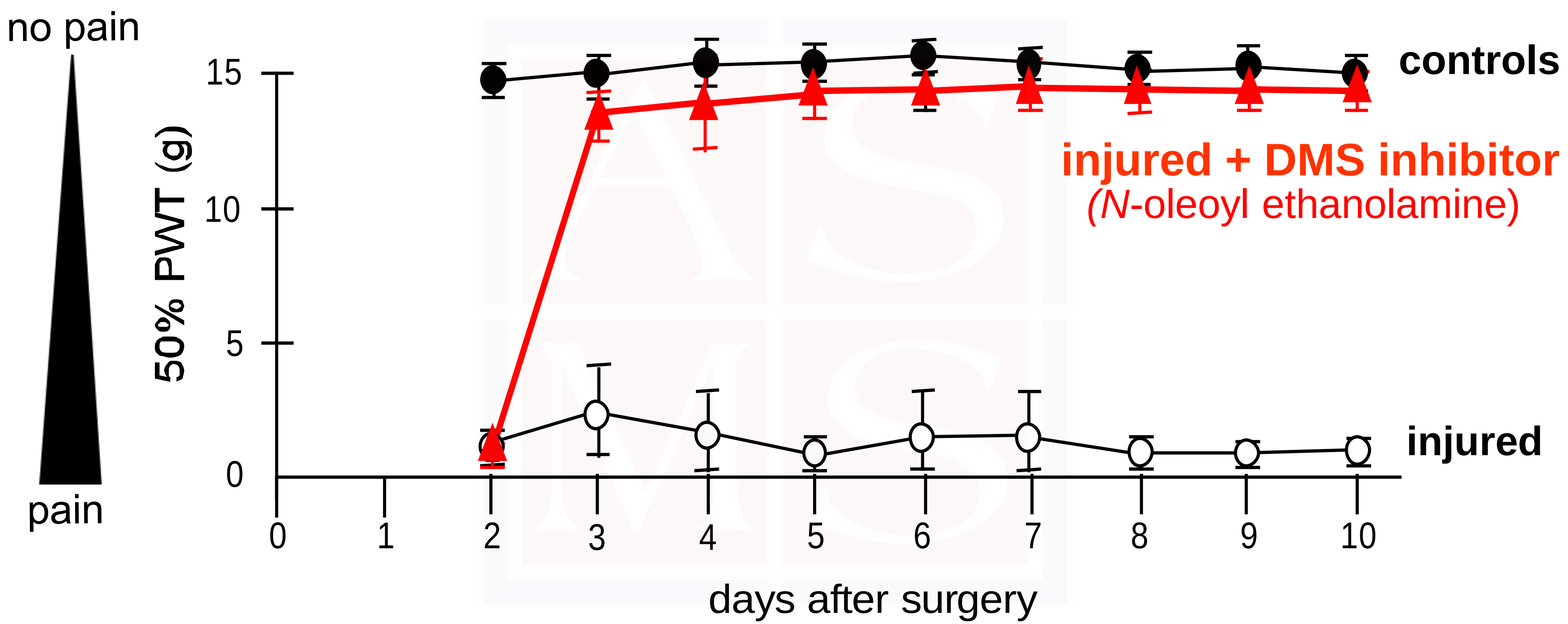
2. Chronic pain of neuropathic origin



2. Chronic pain of neuropathic origin



2. Chronic pain of neuropathic origin



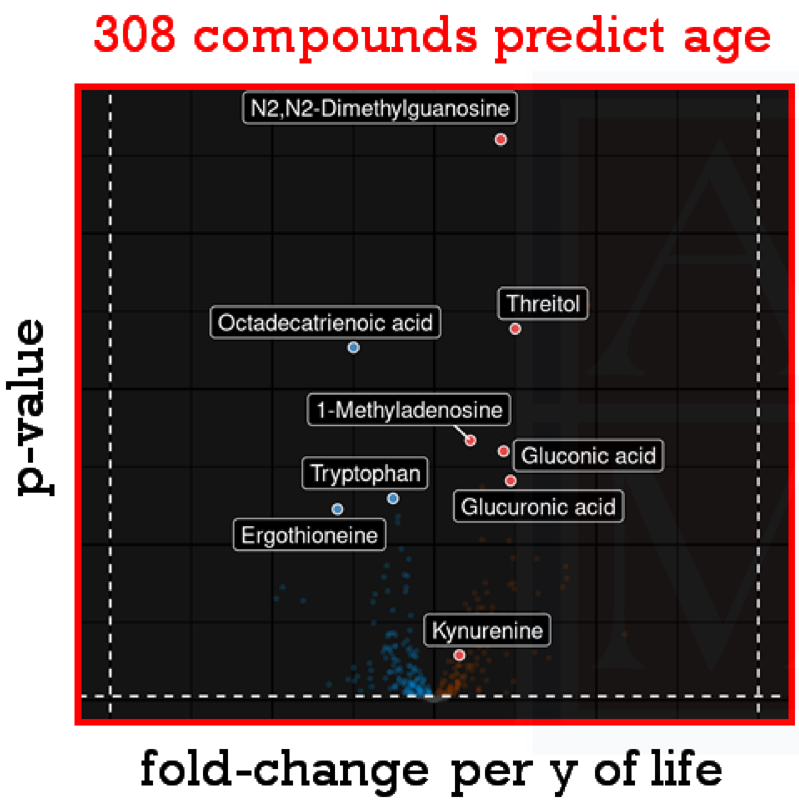
2. Chronic pain of neuropathic origin

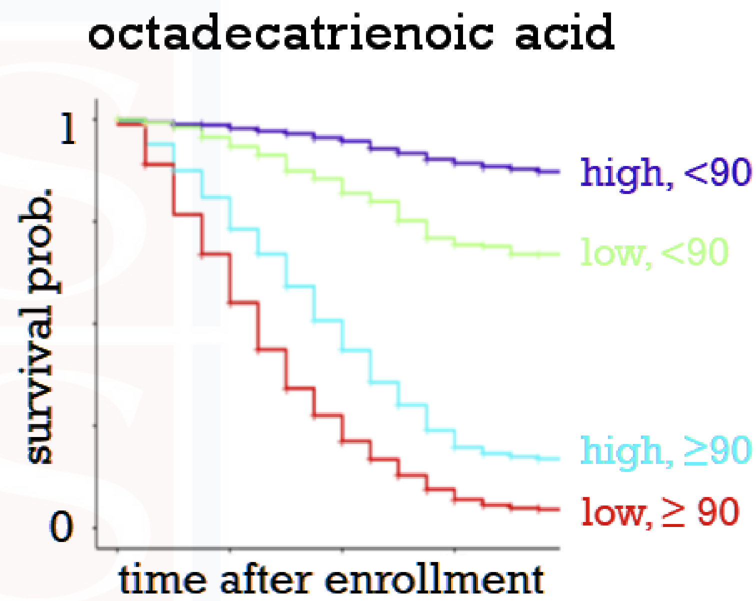
- Activity screen to test unknowns
- Behavioral validation of unknown
- Therapeutic relevance of pathway

3. Healthy aging

- Project to understand healthy aging
- Non-targeted analysis: >30k LC/MS

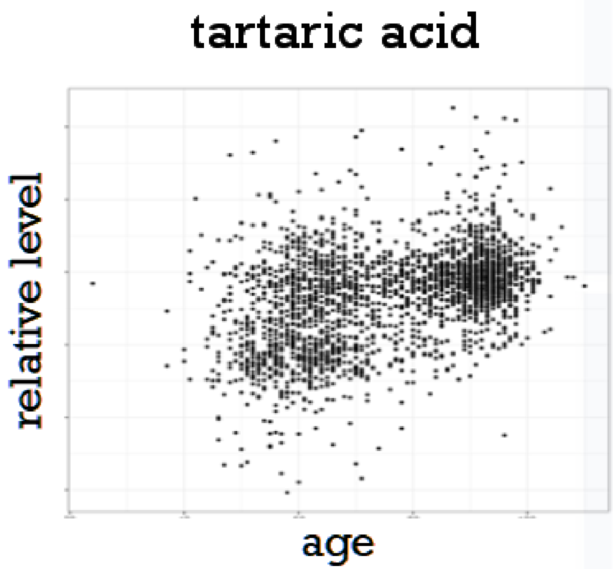
3. Healthy aging

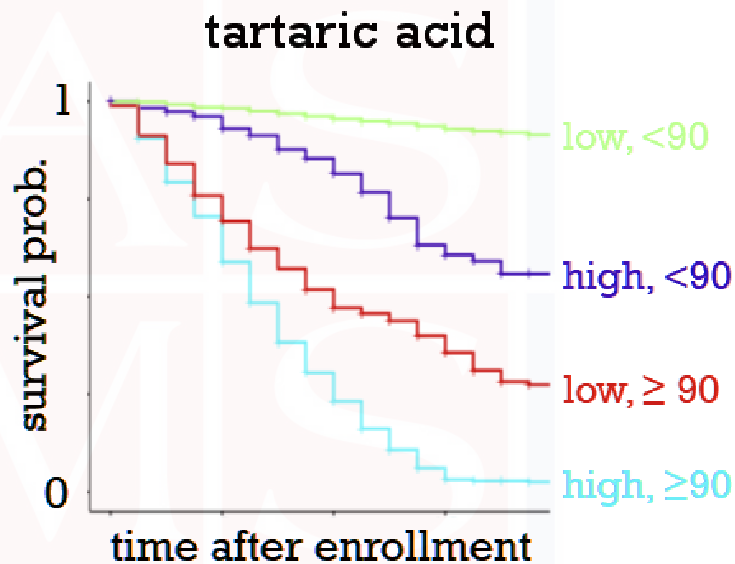




Sebastiani et al: Cell Reports 2024

3. Healthy aging





Sebastiani et al: Cell Reports 2024

- *Overview*
- *Objectives and exp. design*
- *Evaluating performance*
- *Sample prep. and extraction*
- *Separating metabolites*
- *Principles of informatics*
- *Stable isotope tracer analyses*
- *Advanced workflows*
- *Applications*