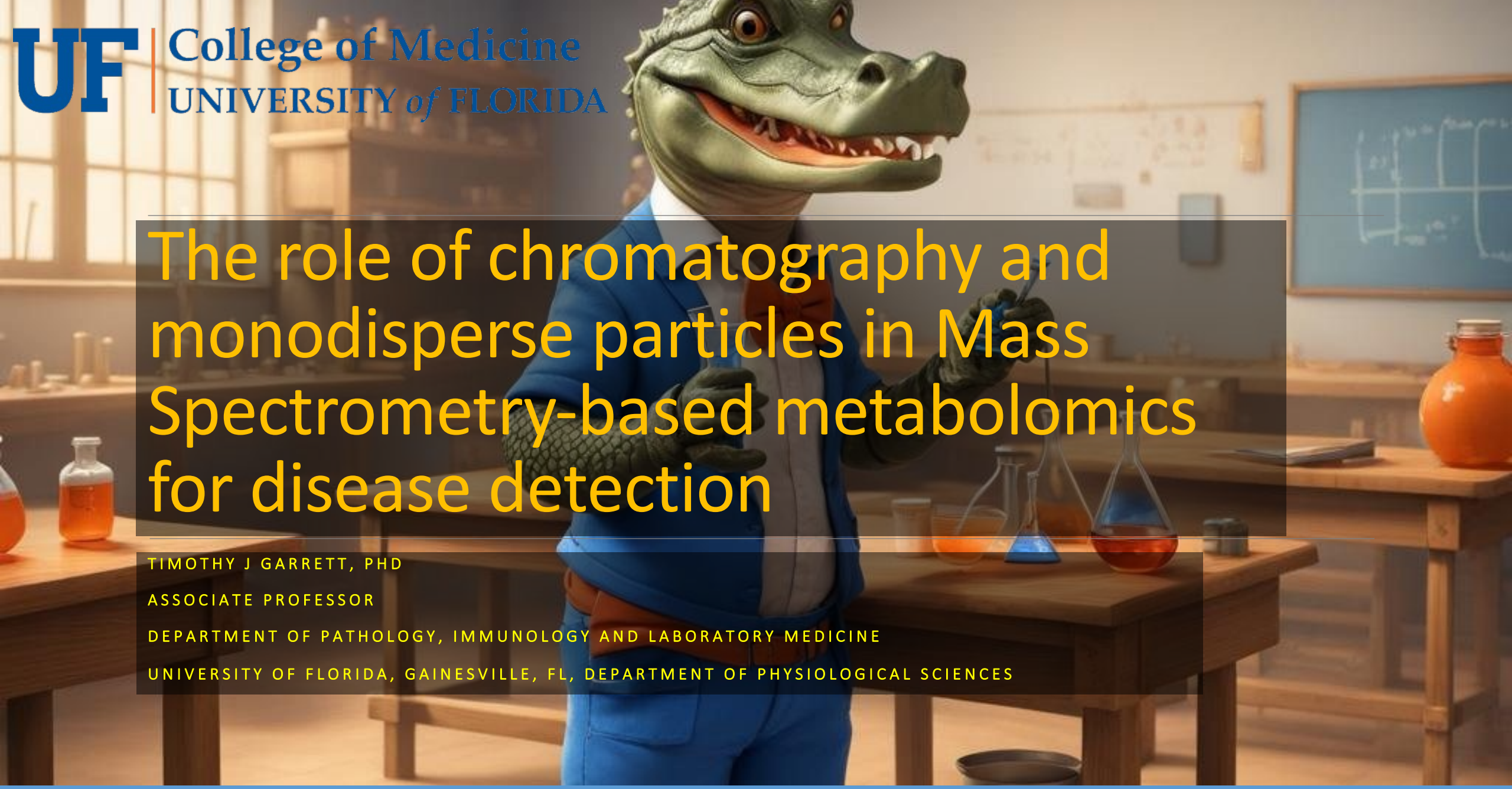


UF College of Medicine
UNIVERSITY of FLORIDA

The role of chromatography and monodisperse particles in Mass Spectrometry-based metabolomics for disease detection

TIMOTHY J GARRETT, PHD

ASSOCIATE PROFESSOR

DEPARTMENT OF PATHOLOGY, IMMUNOLOGY AND LABORATORY MEDICINE

UNIVERSITY OF FLORIDA, GAINESVILLE, FL, DEPARTMENT OF PHYSIOLOGICAL SCIENCES

Background photo credit to Don Chace

Outline

- The Southeast Center for Integrated Metabolomics
- The ‘Omics
- Metabolomics
- Monodisperse particle columns and metabolites
- Meningioma, machine learning and ‘Omics

Goals and Mission of Research and the Center

Provide metabolomic and lipidomic analyses to the broader scientific community

Focus on driving technology and high-throughput service analysis

Build open-source tools to aid in analysis and interpretation

Provide grants and collaborative opportunities in metabolomics/lipidomics

Train scientists/postdocs/graduate/undergraduate students in metabolomic and lipidomic tools and technologies

Metabolomic analyses to the broader community

approximate locations

Services and Research

10 instruments (HRMS and QQQ)

Targeted quantitative assays

- Tryptophan metabolites (7 analytes)
- Purines for food analysis(8 analytes)
- Mononucleotides (5 analytes)
- TMAO, choline, betaine
- Amino acids
- Acylcarnitines
- Metformin

Method development

Metabolomics via LC-HRMS/MS

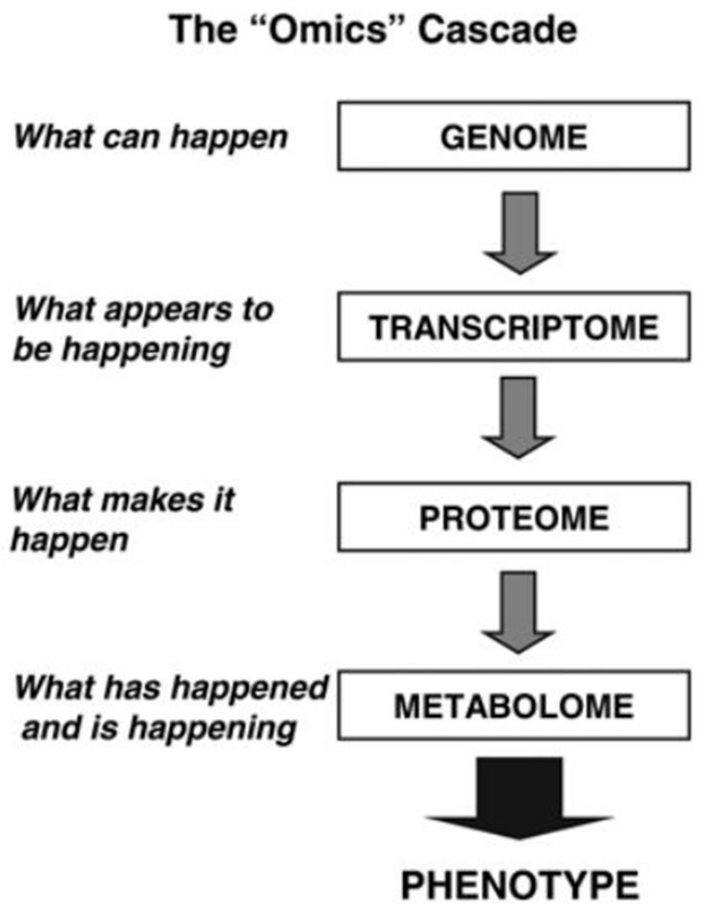
- HILIC and RP methods

Lipidomics via LC-HRMS/MS

Bioinformatics

- Experienced in univariate, multivariate and machine learning

The era of Omics technology



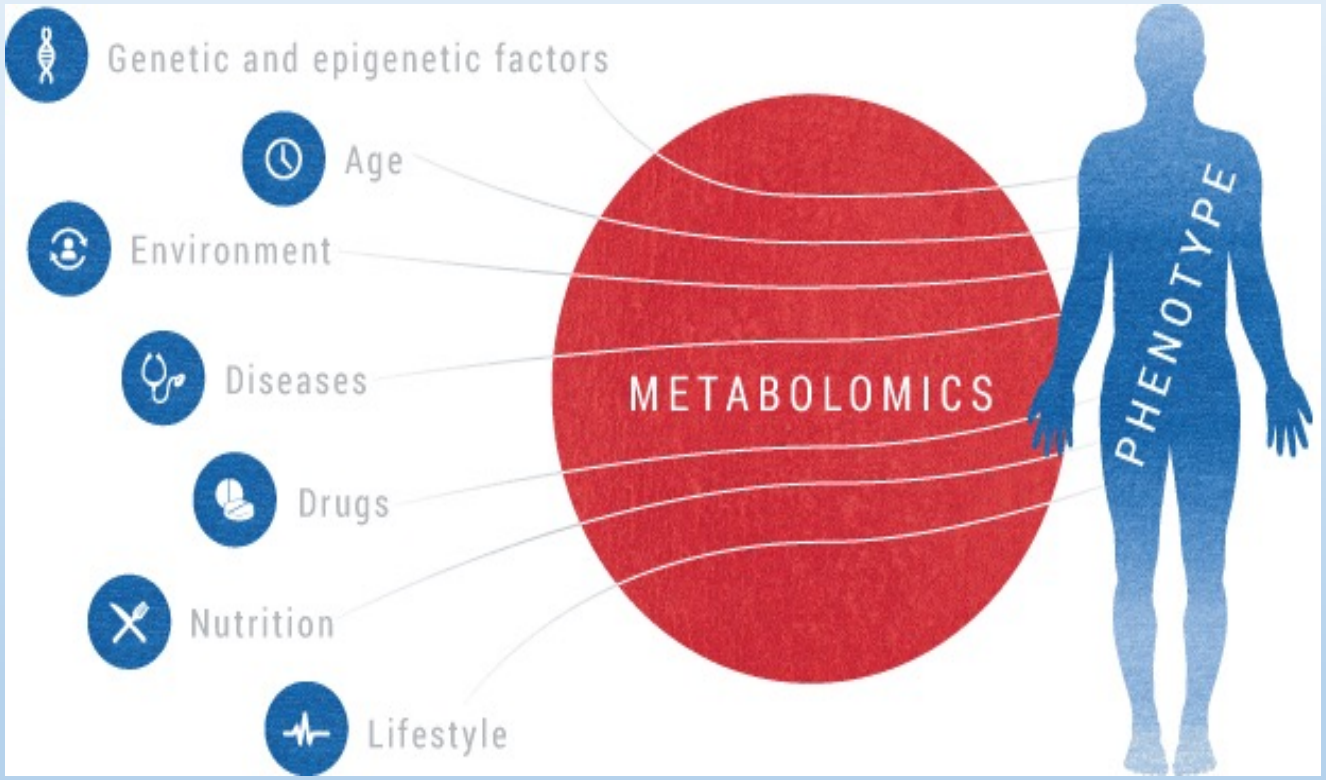
Metabolome is the collection of small molecules in cells, tissue, plasma, urine, tissue, etc.

Metabolomics is the measure of those metabolites

The comprehensive characterization of small molecules in a given system

Mass Spectrom Rev. 2007 Jan-Feb;26(1):51-78

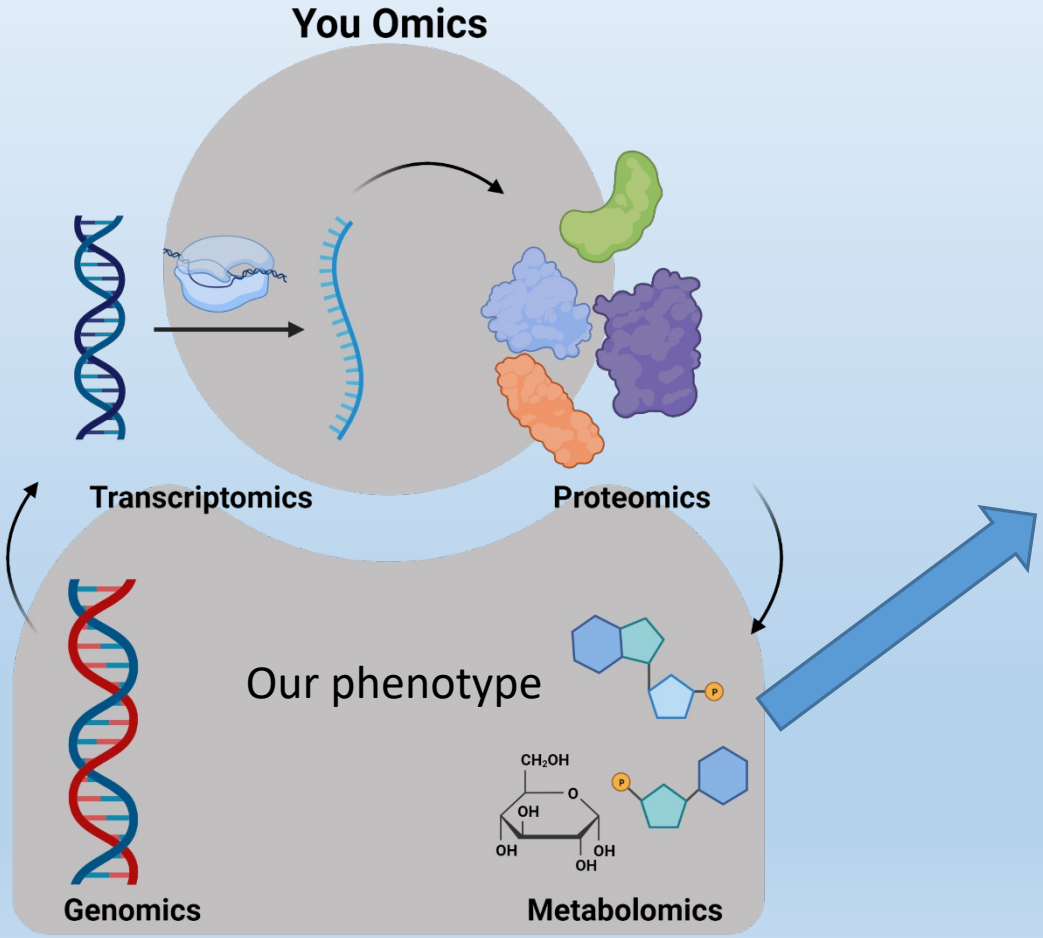
Metabolomics

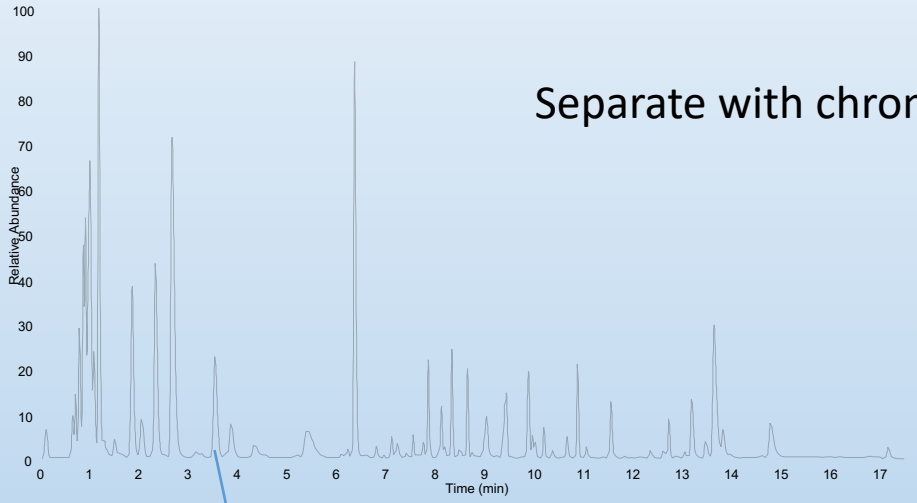


We are diverse and so our metabolome is reflective of that!

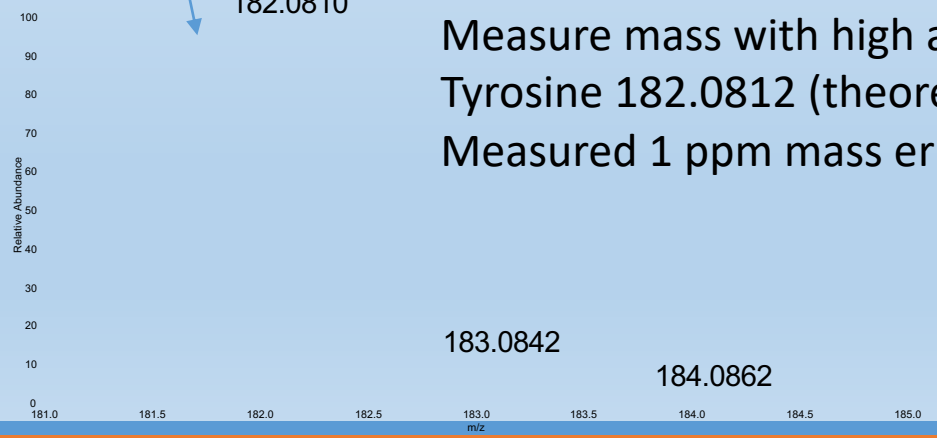
http://www.metanomics-health.com/en/why-metabolomics.html

The ‘Omics of you





Separate with chromatography



Measure mass with high accuracy
Tyrosine 182.0812 (theoretical)
Measured 1 ppm mass error

Global Metabolomics by LC-HRMS

Identify metabolites based on *m/z* values

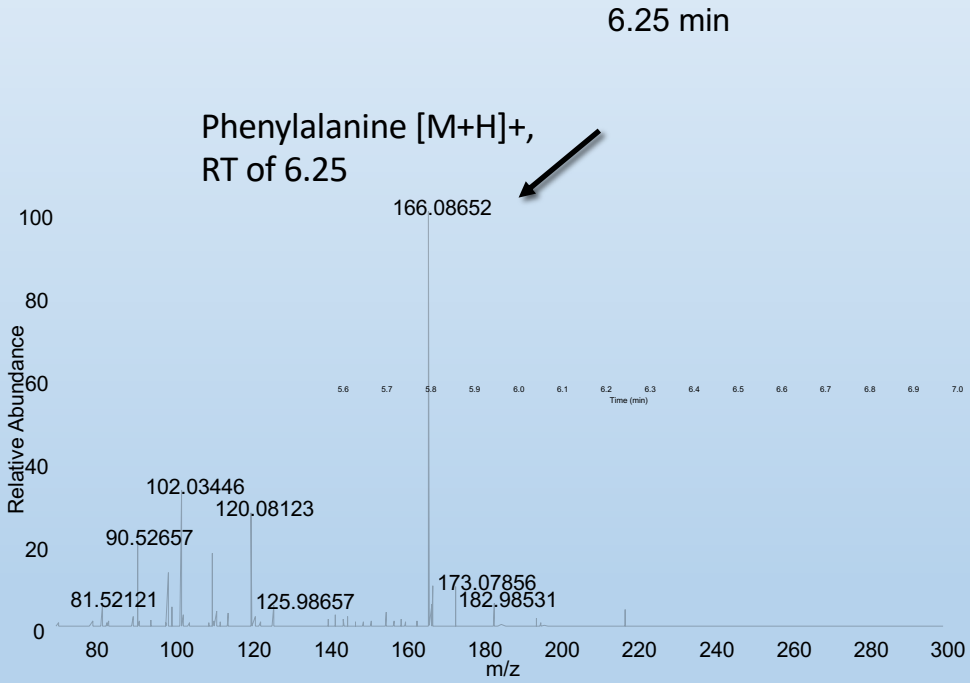
- Couple with chromatography to use m/z and RT pairs to improve number of metabolites detectable
- High mass accuracy is important for identification

Intensity of peak represents concentration

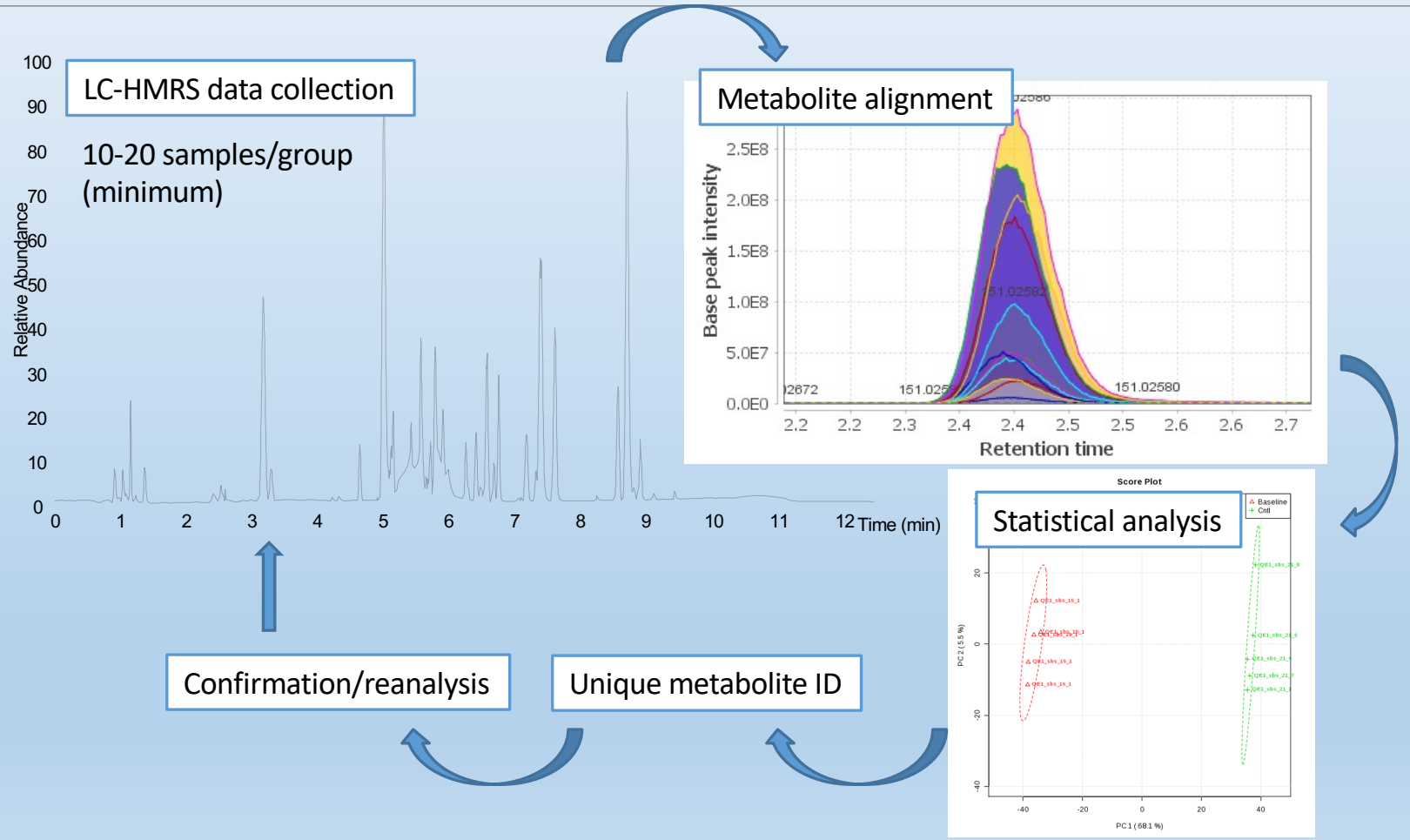
- *m/z* peak or chromatographic peak

Identification requires additional information

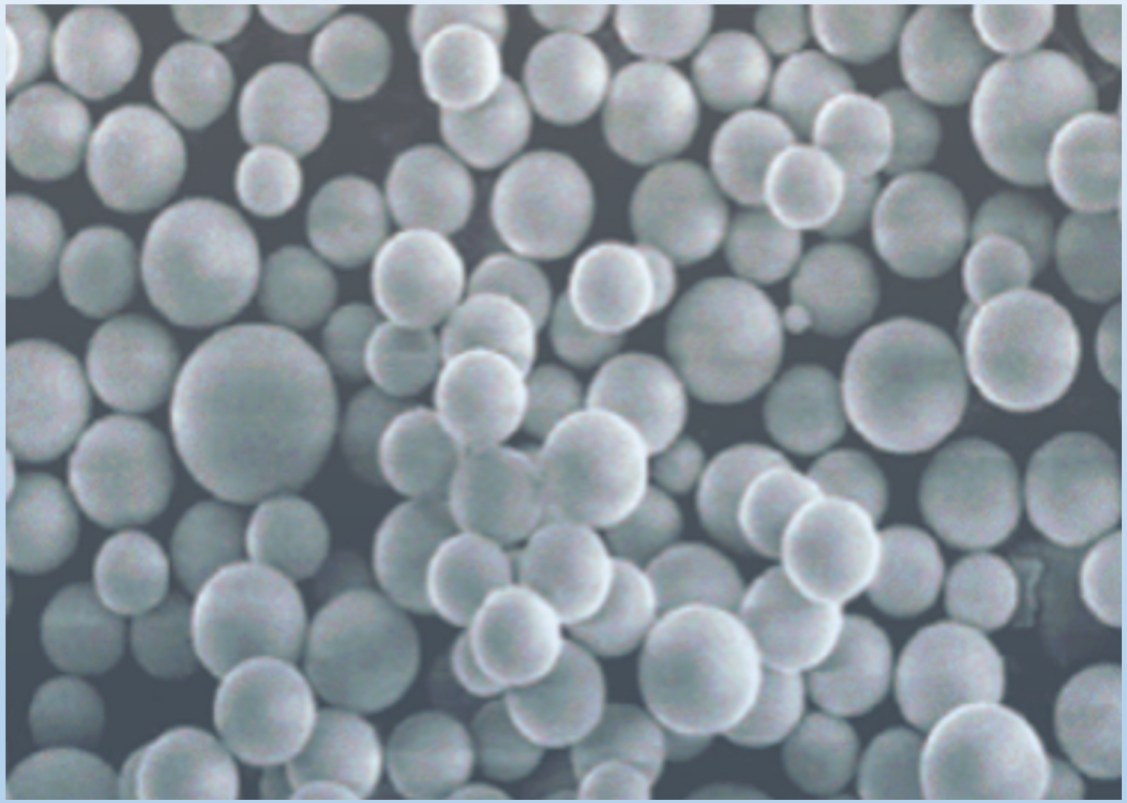
- RT match
- *m/z* match
- MS/MS match
- Not just a match to a metabolite library



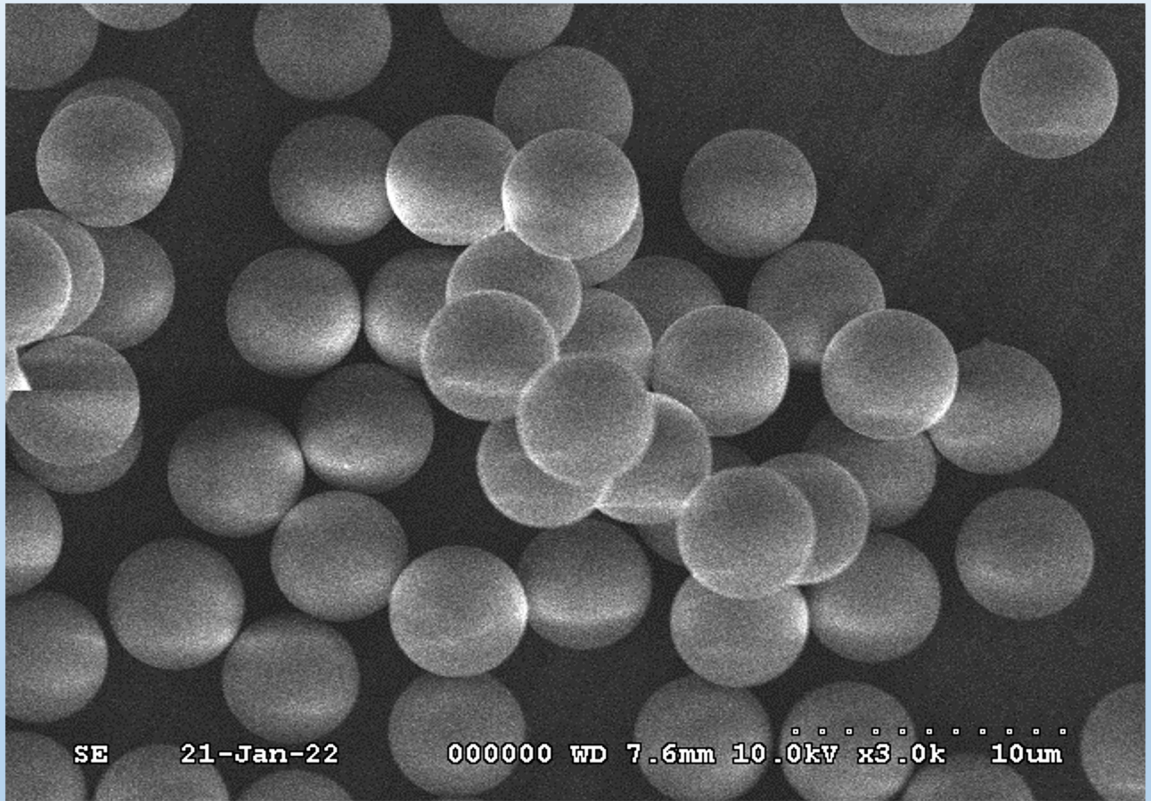
Metabolomic workflow



Polydisperse vs Monodisperse

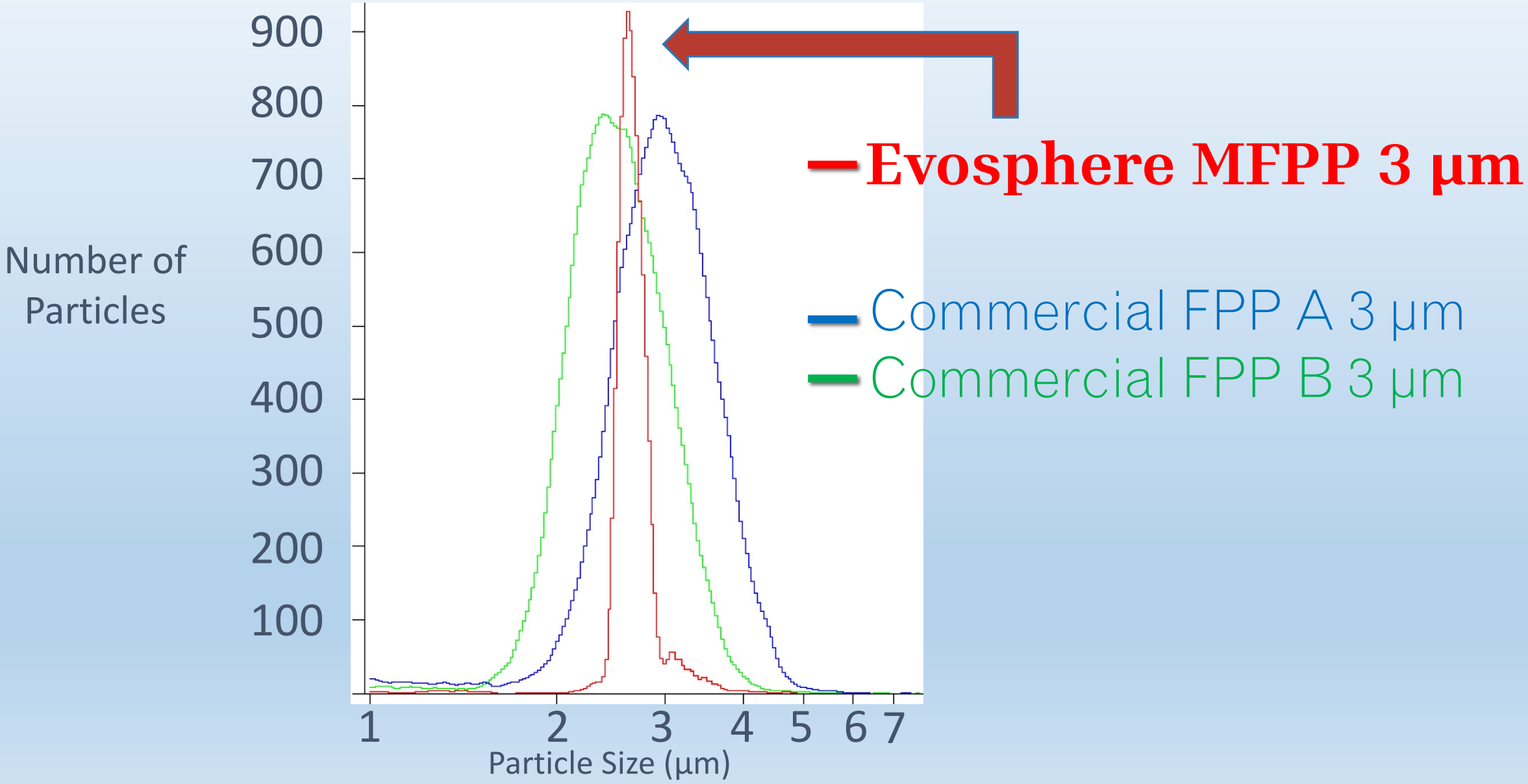


Polydisperse
Fully Porous Particle



Monodisperse
Fully Porous Particle

Particle Size Distribution Comparison



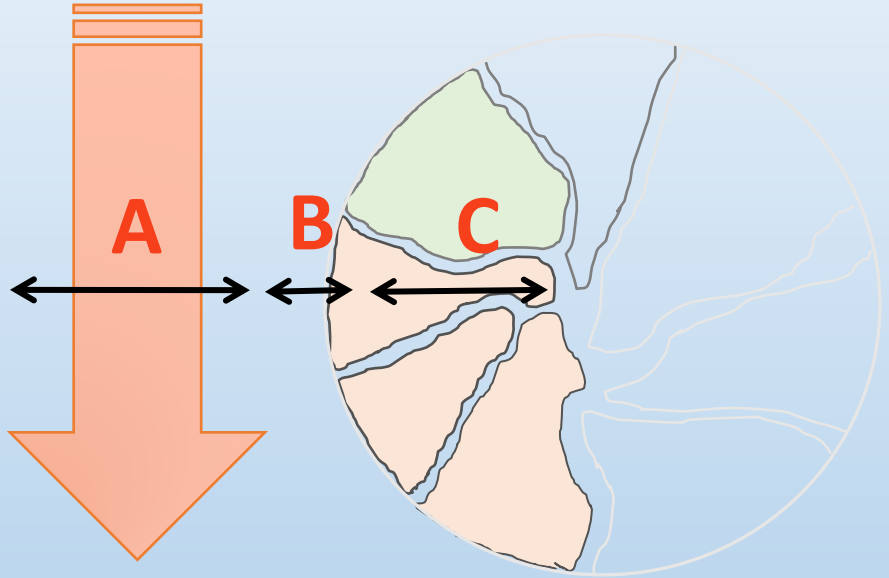
How does MFPP impact band broadening?

H: Height Equivalent to a Theoretical Plate

A: Eddy Diffusion

B: Longitudinal Diffusion

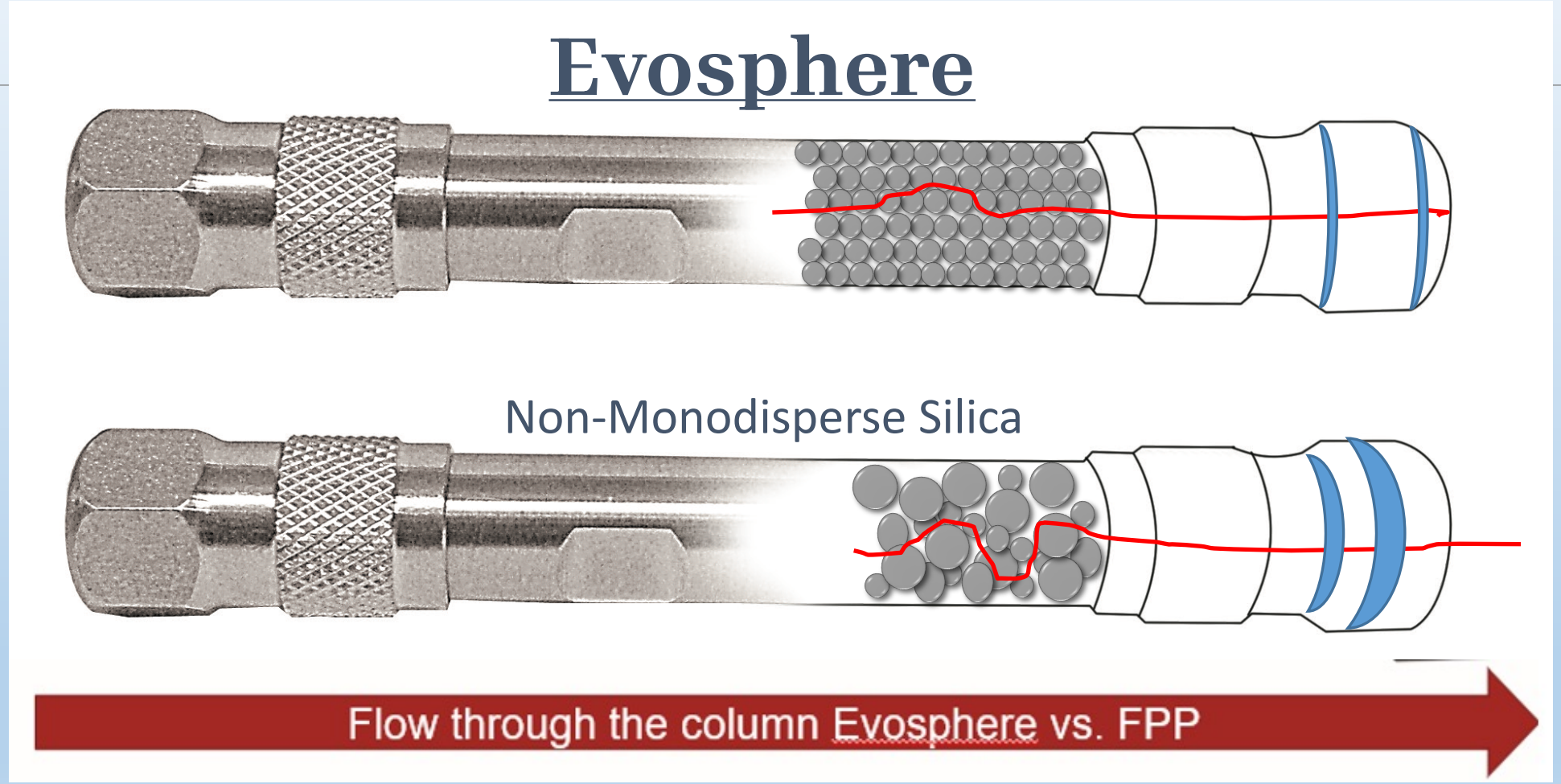
C: Resistance to Mass Transfer



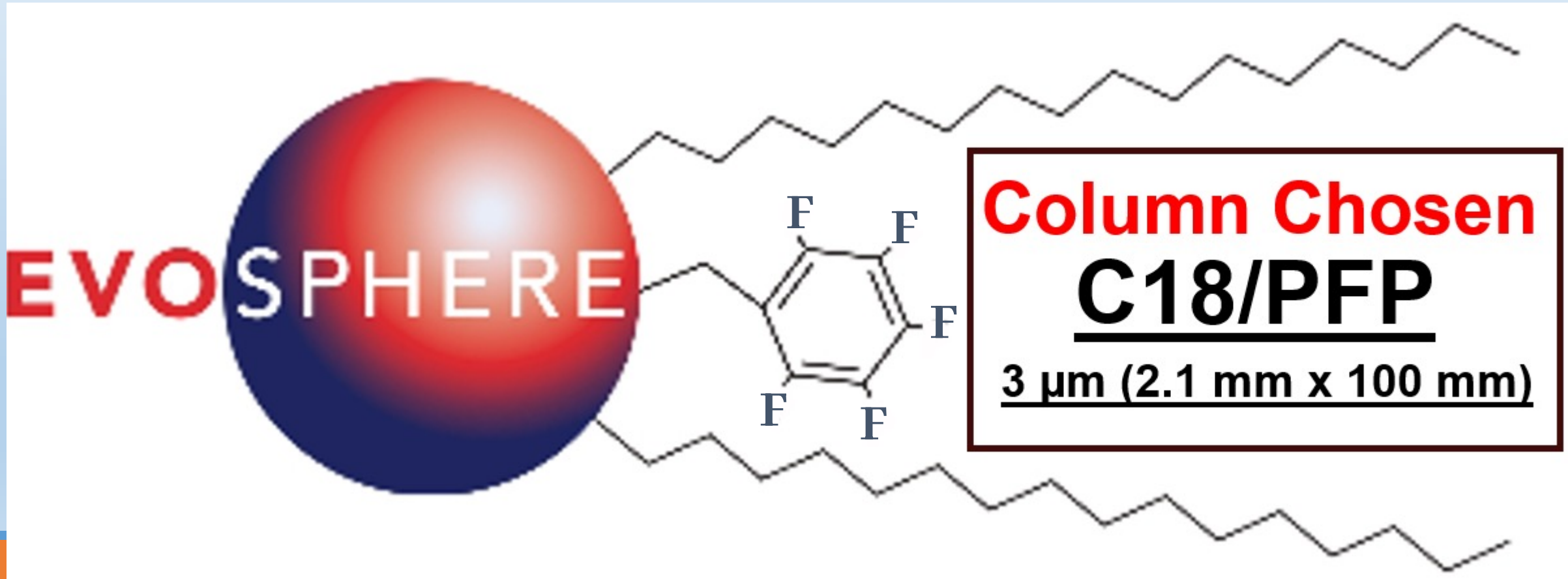
Van Deemter Equation

$$H = A + \frac{B}{u} + Cu$$

What does this look like visually through the column?



Untargeted Metabolomics on Plasma Extract



Total Ion Chromatogram

Column Phase – Evosphere C18/PFP

Dimensions - 3 µm (2.1 mm x 100 mm)

Instrument - Thermo Q-Exactive with Dionex UHPLC

Sample – Plasma Extract

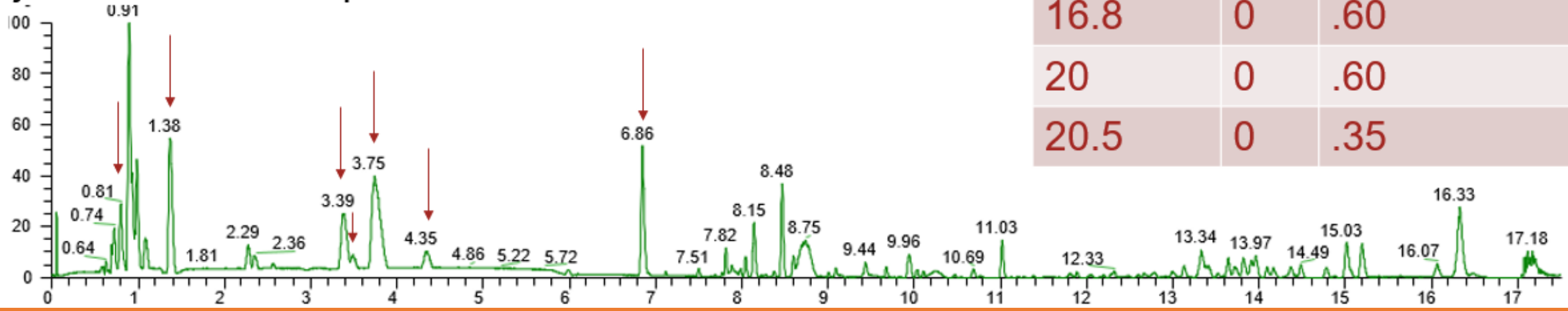
Mobile Phase A = 0.1% Formic Acid in H_2O

Mobile Phase B = Acetonitrile

Temperature = 25ºC

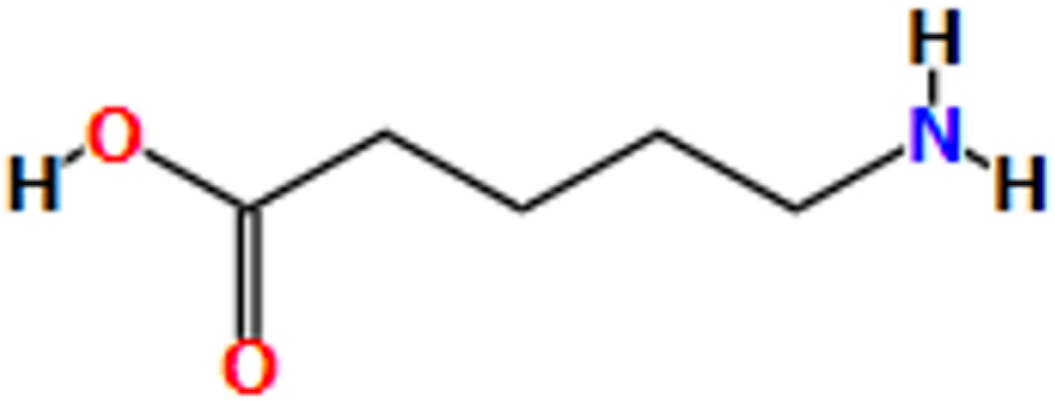
Injection volume = 2µL

Time	% B	Flow Rate (mL/min)
3 min	0	.35
13 min	80	.35
16 min	80	.35
16.5	0	.35
16.8	0	.60
20	0	.60
20.5	0	.35

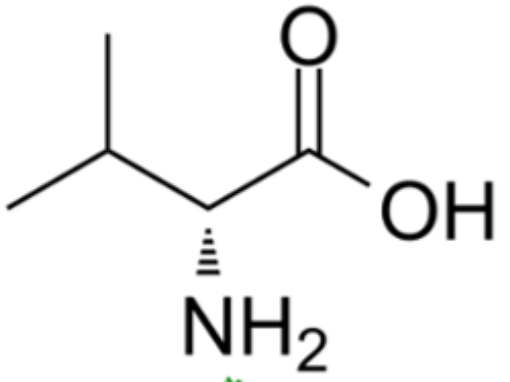


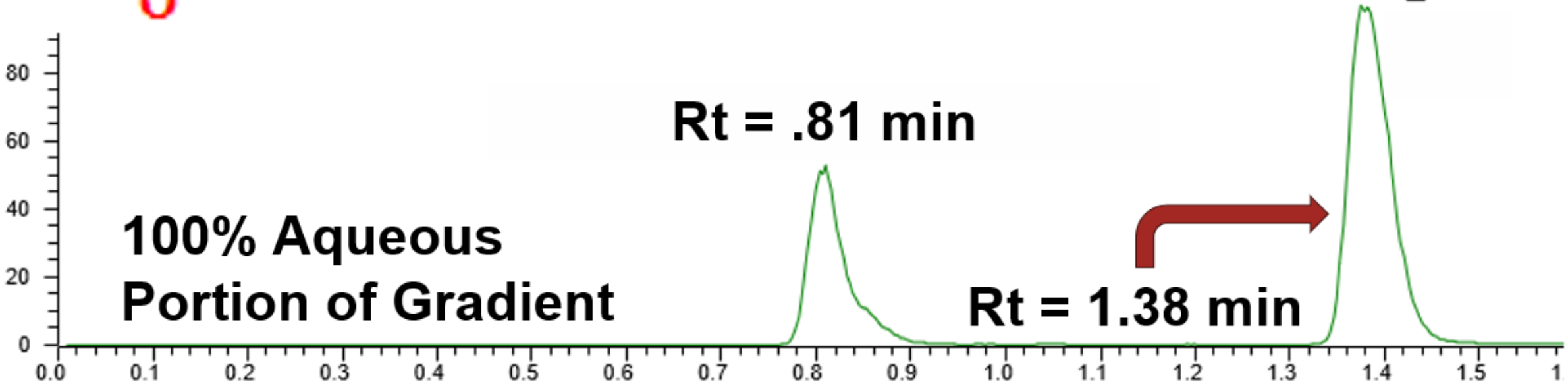
Polar Retention

5-Aminopentanoic Acid



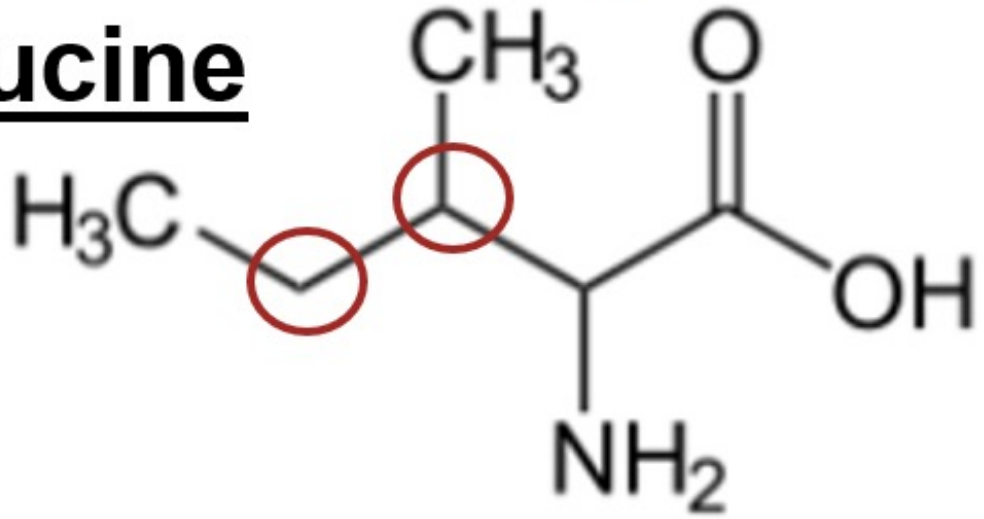
Valine



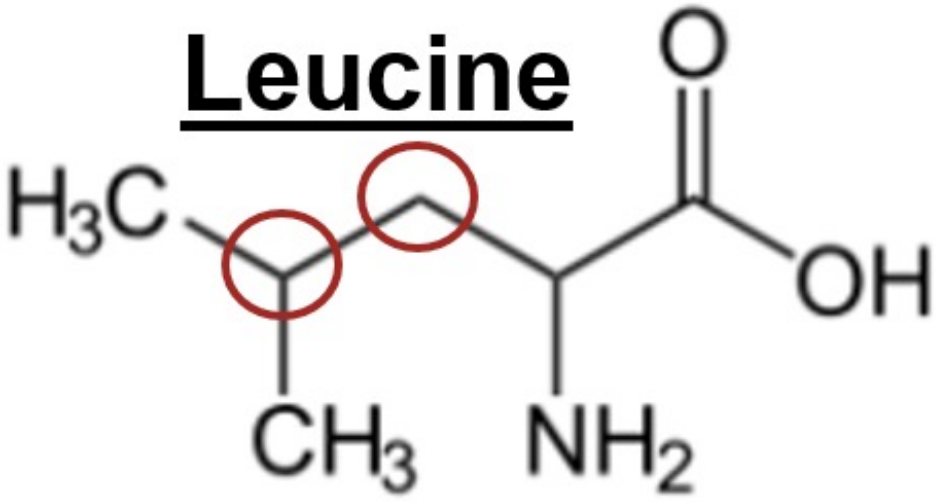


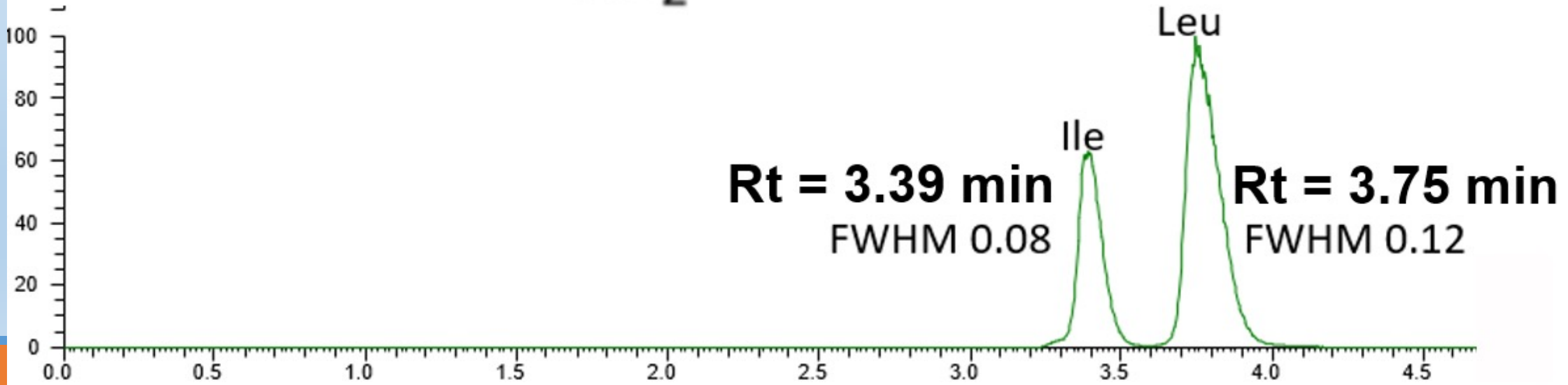
Methyl Position Switch

Isoleucine



Leucine

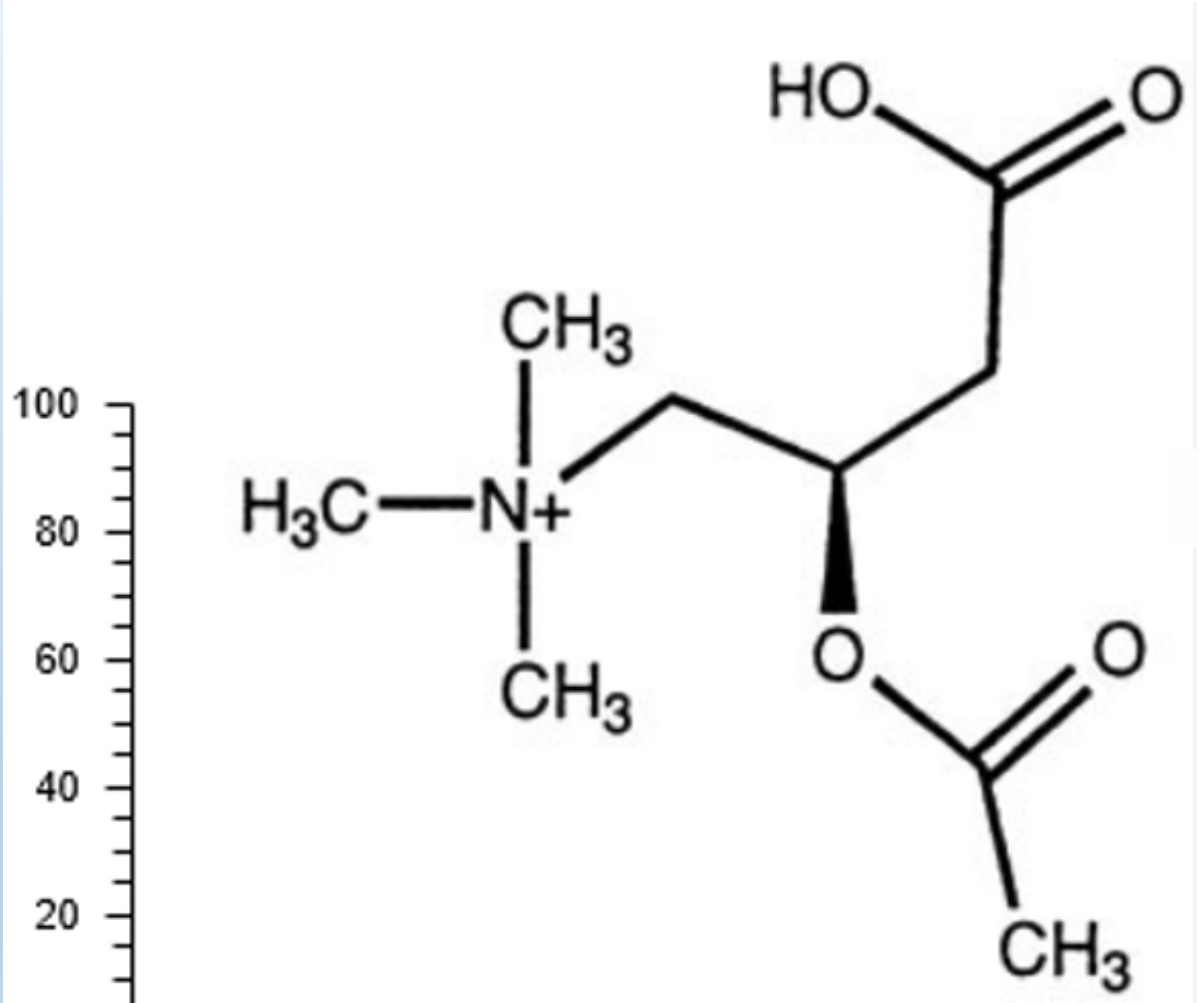




Rt = 3.39 min
FWHM 0.08

Rt = 3.75 min
FWHM 0.12

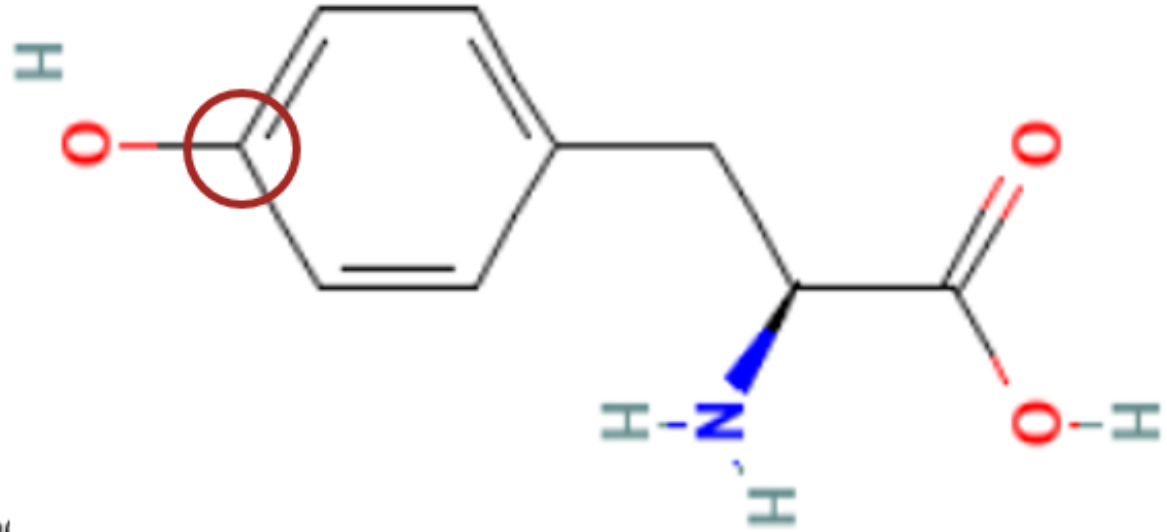
Optimal Peak Shape

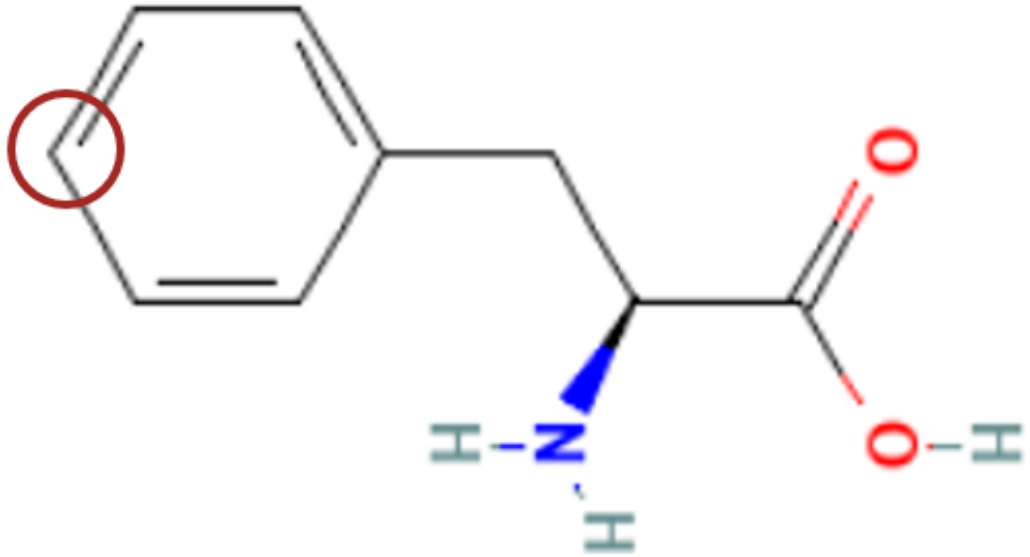


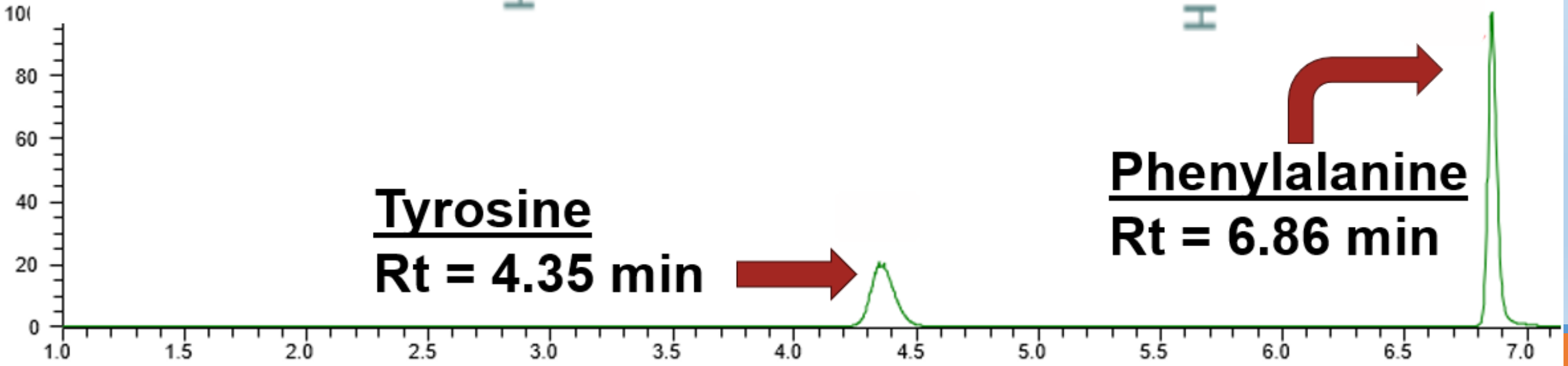
Acetyl-Carnitine

Rt = 3.51 min

Strong Isomeric Selectivity





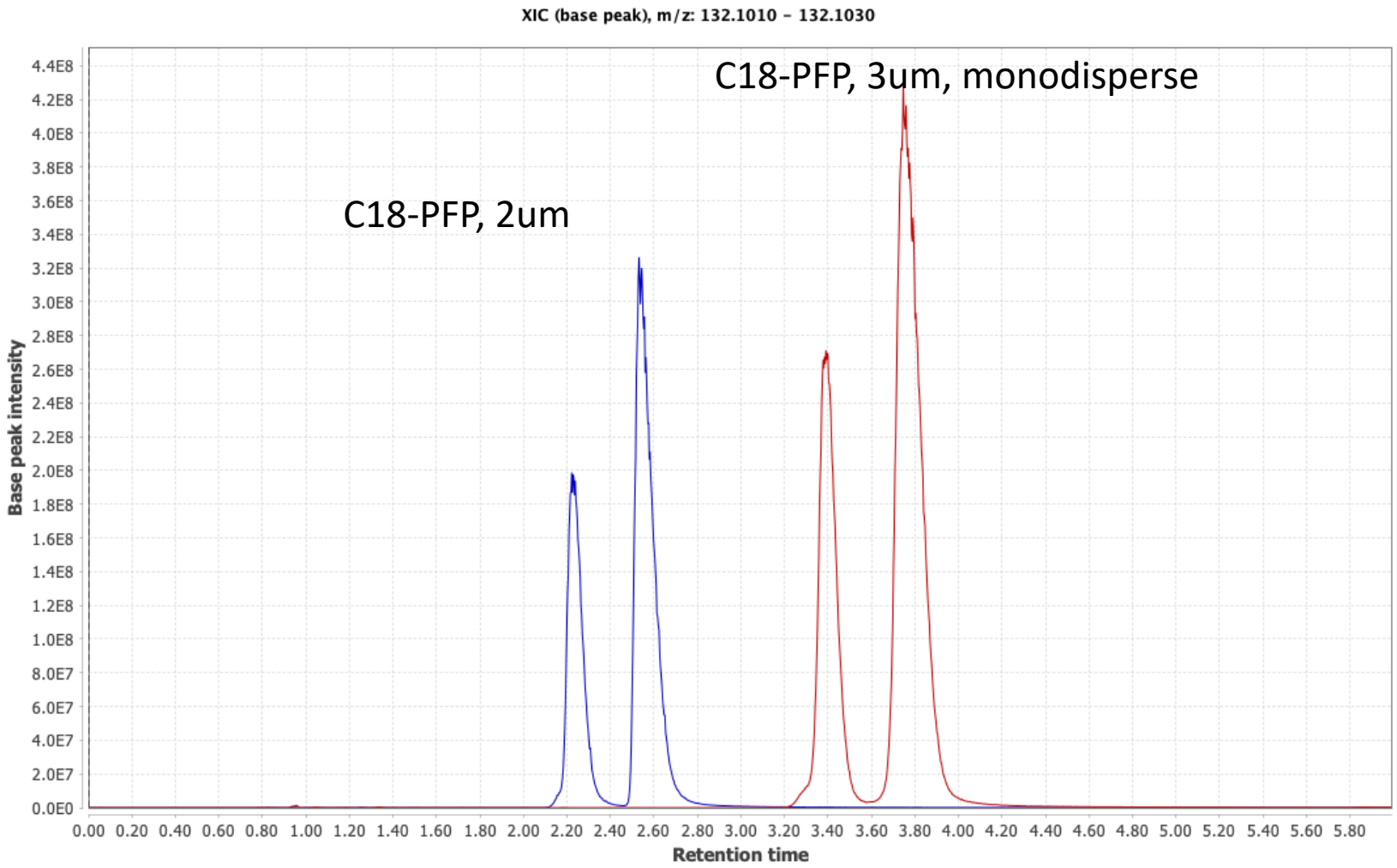


Column comparison

C18-PFP 100x2.1mm, 2um

C18-PFP 100x2.1, 3um (monodisperse)

Isoleucine and leucine shown as a comparison



Note the increased retention of both isoleucine and leucine on the monodisperse column (red trace)

Column comparison

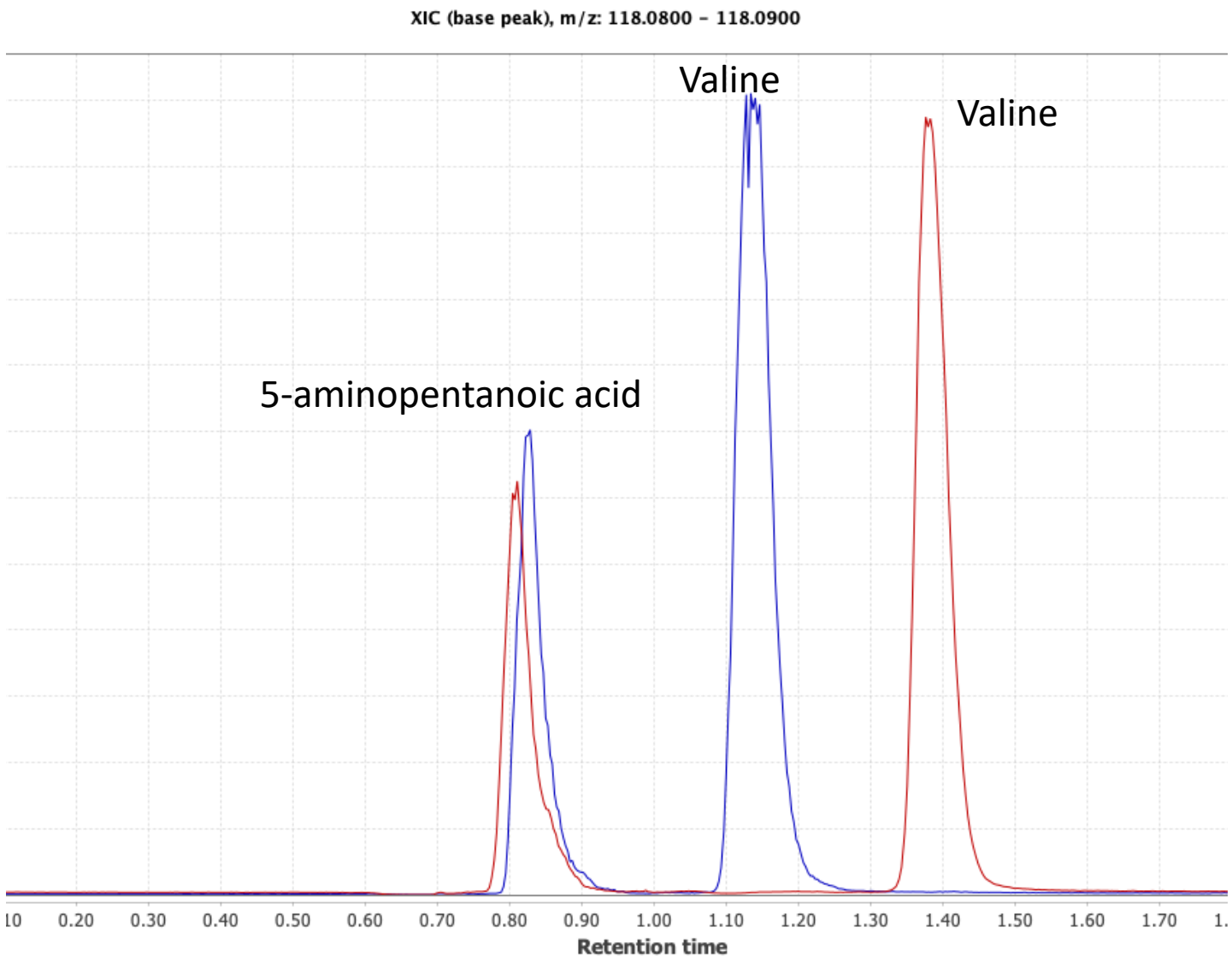
C18-PFP 100x2.1mm, 2um

C18-PFP 100x2.1, 3um (monodisperse)

Isomers at 118.086

5-aminopentanoic acid

Valine



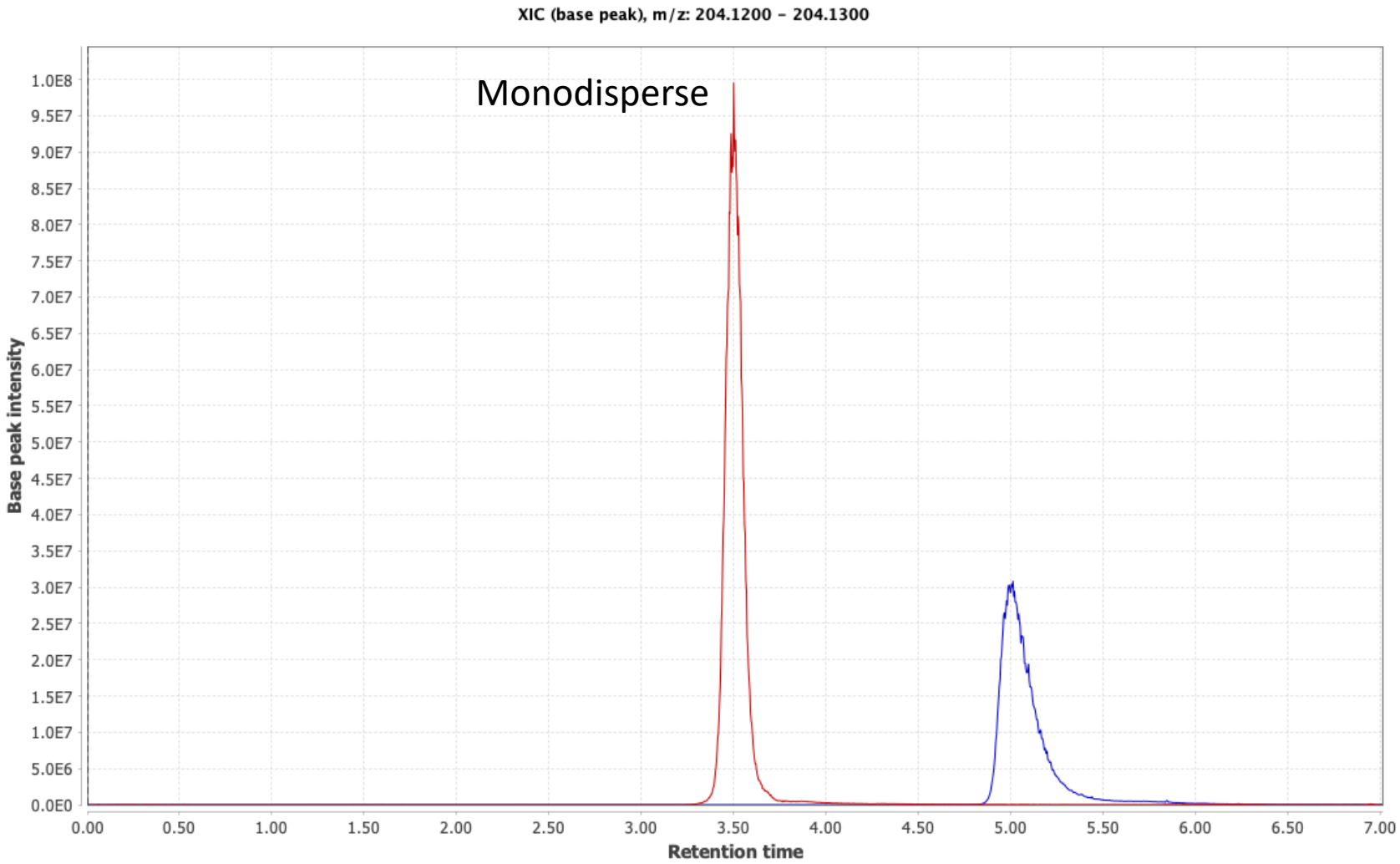
Monodisperse column in red trace

Column comparison

C18-PFP 100x2.1mm, 2um

C18-PFP 100x2.1, 3um (monodisperse)

Acetylcarnitine (204.123)



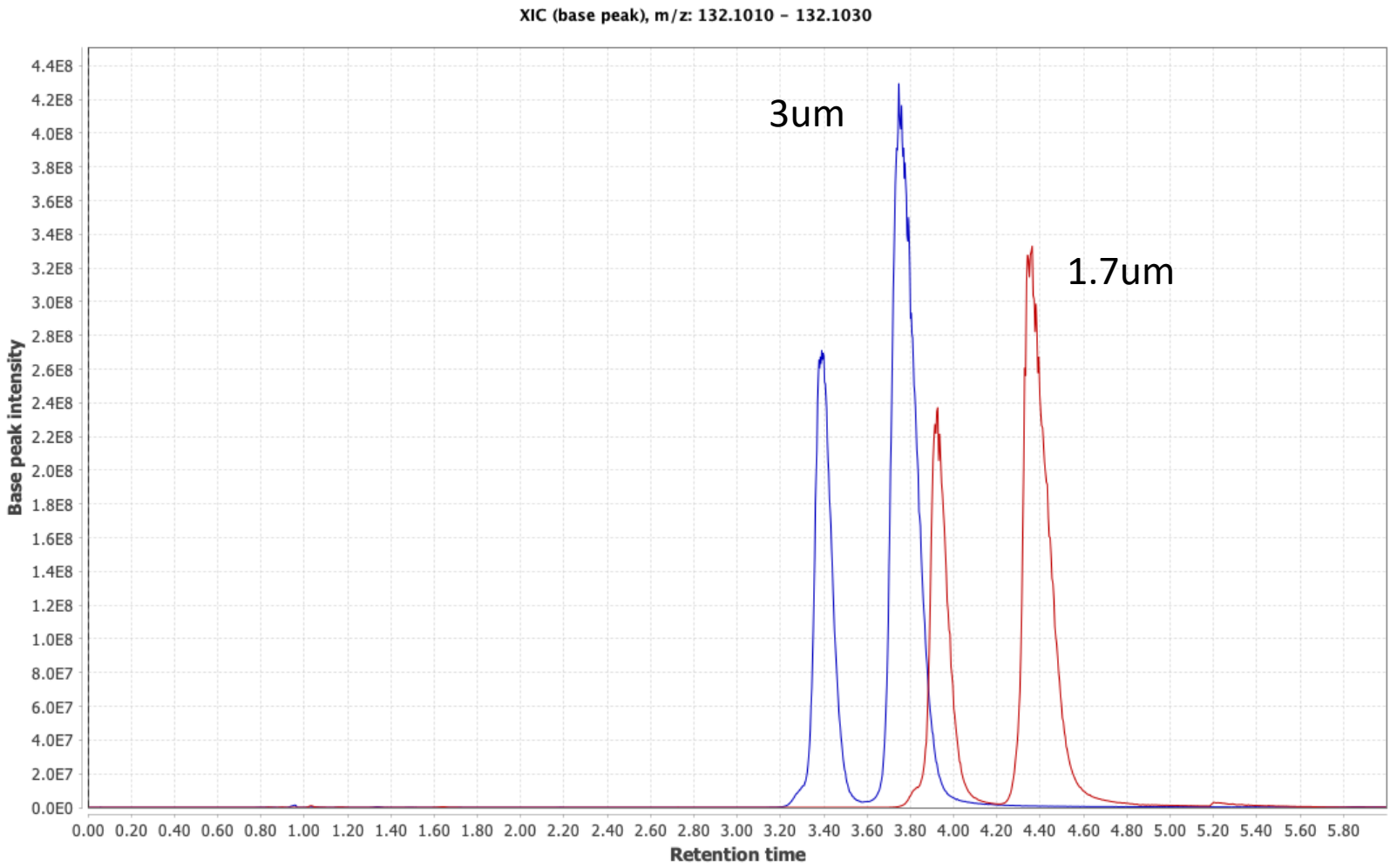
In this case, the monodisperse particle column results in superior peak shape

Column Comparison

C18-PFP 100x2.1mm, 3um (monodisperse)

C18-PFP 100x2.1, 1.7um (monodisperse)

Revisiting isoleucine and leucine



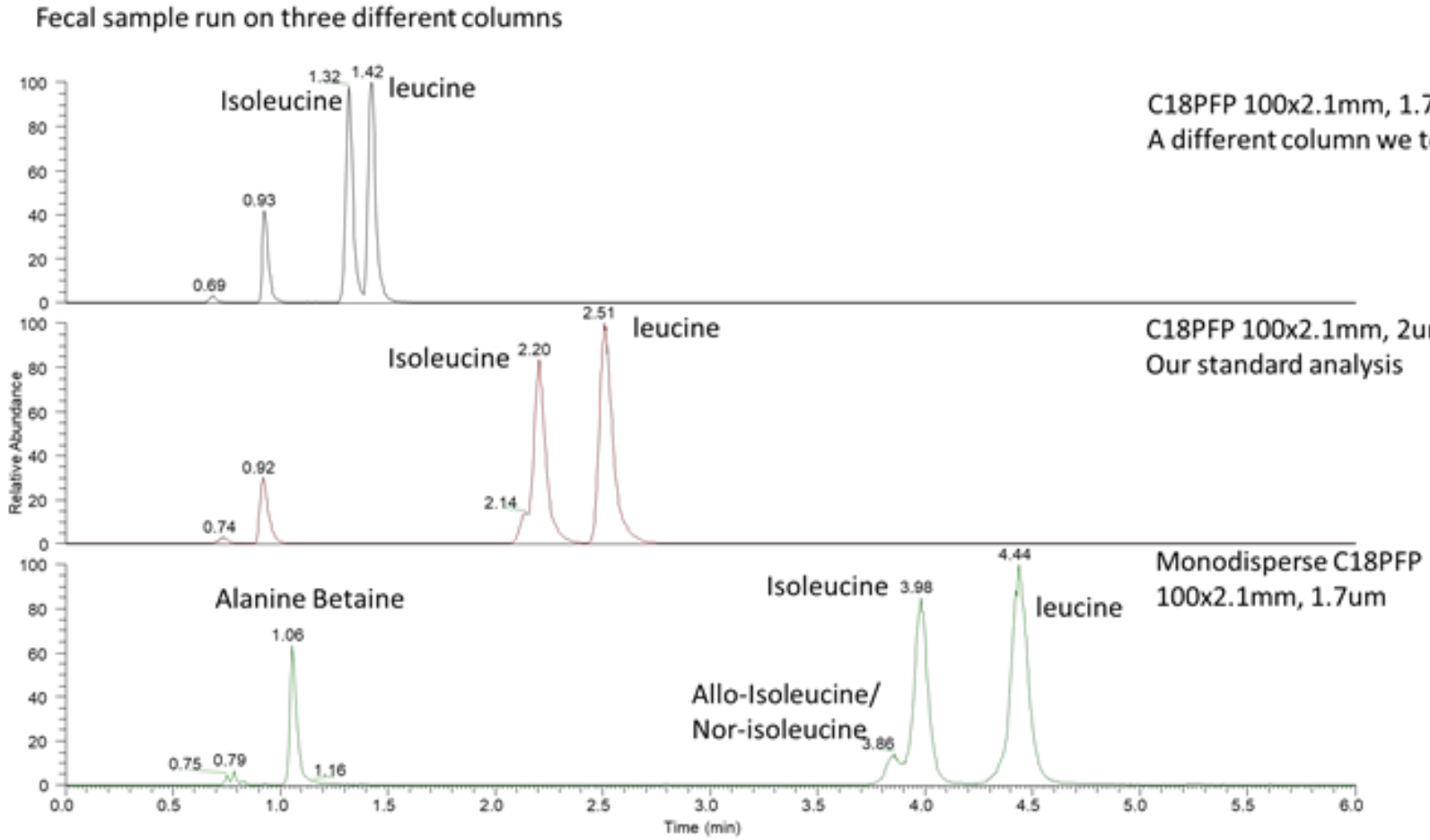
Slightly longer retention, tighter peaks and improve separation between the isomers with 1.7um

Sample type comparison

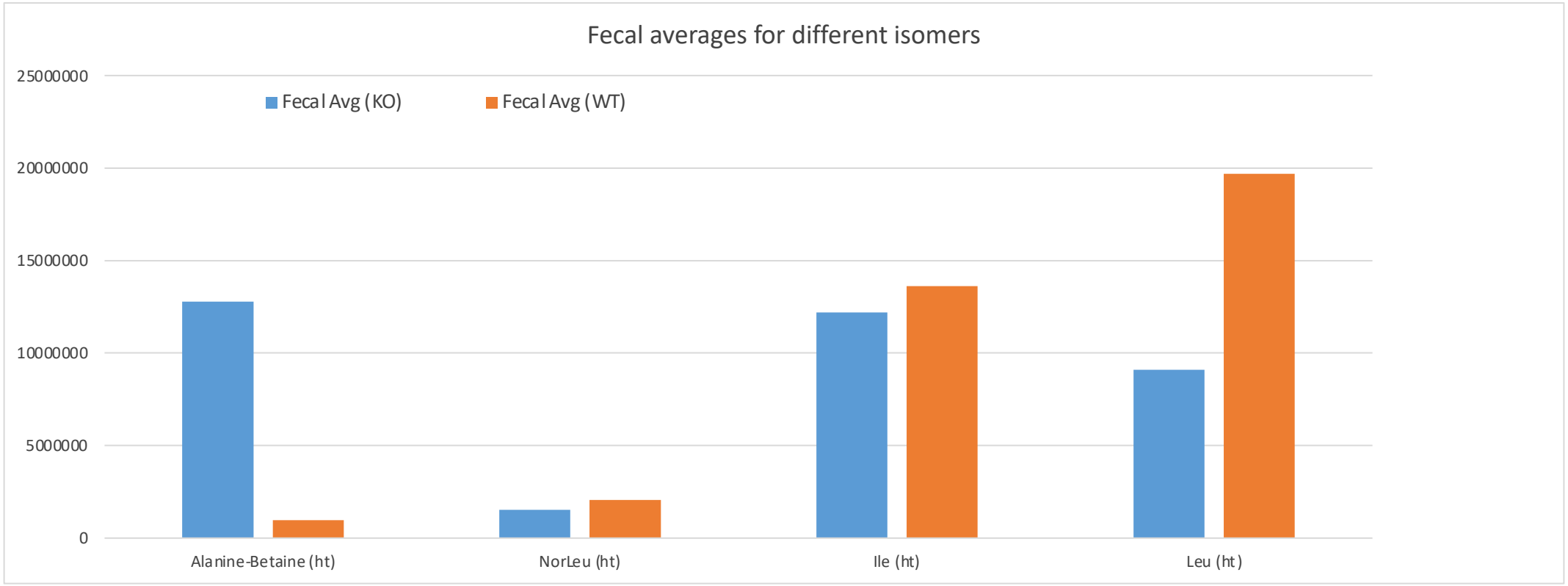
Fecal sample looking at isomers of 132.102

Isoleucine, Leucine, allo-isolueucine/nor-isoleucine

Alanine Betaine



Note the longer retention with monodisperse columns which helps to elucidate another isomer (alloisoleucine/norisolecuine)



Improved separation=better statistical analysis

Example of integrated ‘Omics

MENINGIOMA STUDY

METABOLOMICS, LIPIDOMICS AND AI

Hoda Safari Yazd, Sina Bazargani, Garrett Fitzpatrick, Richard A Yost, Jesse Kresak, Timothy J Garrett: Mass Spectrometry-Based Metabolomic and Lipidomic Characterization of Meningioma Grades using Machine Learning. *In preparation*

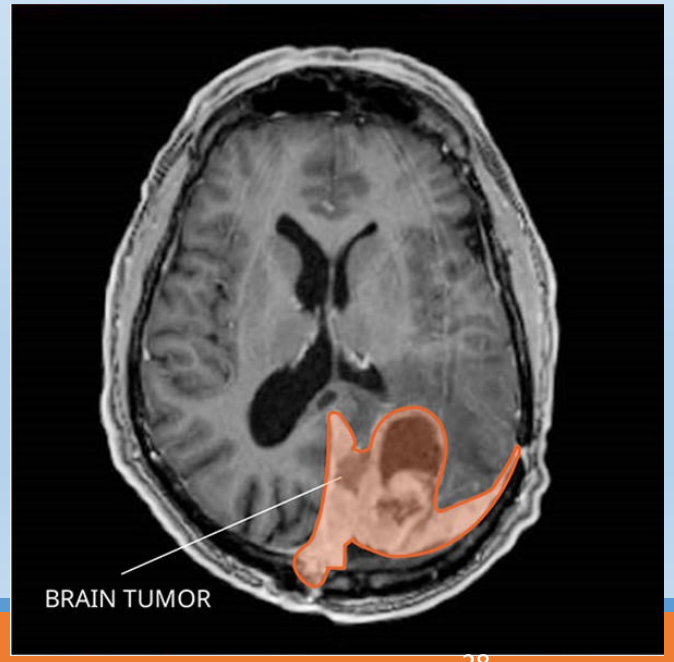
Meningioma Background

A usually noncancerous tumor that arises from the membranes surrounding the brain and spinal cord.

Meningiomas account for 37% of primary brain tumors in the US.

Meningiomas are classified into WHO Grades 1, 2, and 3.

	Grade I	Grade II	Grade III
Frequency	75%	20-35%	1-3%
Treatment	Gross total resection	Gross total resection +/- Radiotherapy	Gross total resection +/- Radiotherapy
Survival	Same as age-matched controls	modest decrease	18-40 months
Recurrence	5 year - 5%	5 year - 40%	Frequent



Purpose

Investigate the metabolomic profile of meningiomas

- Grade classification with 2 and 3 is currently difficult
- Compare low-grade and high-grade meningiomas
- Find new biomarkers capable of differentiating different stages of meningioma's
- Identify potential metabolites which may correlate with disease free and overall survival

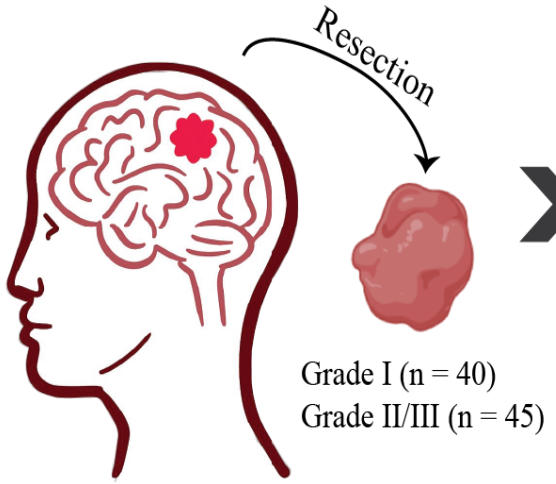
Overarching goal to use metabolomic data to identify biomarkers for disease diagnostics

Collaboration with anatomical pathology (Jesse Kresak, MD)

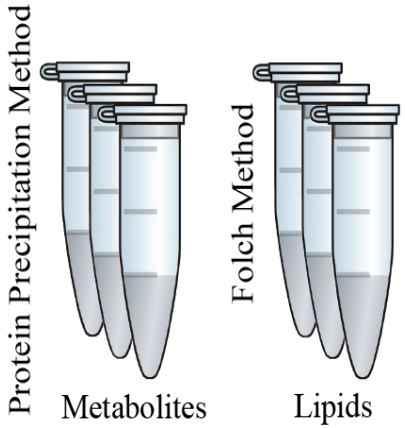
Study Design and Workflow

1. Experimental and Laboratory Analysis

Meningioma Tissue Collection

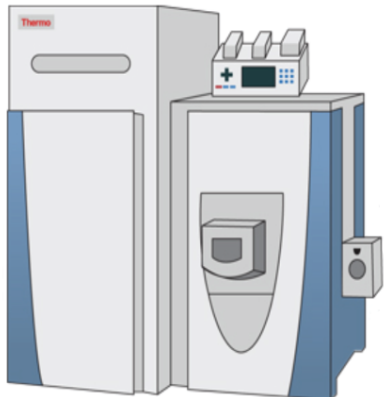


Sample Preparation

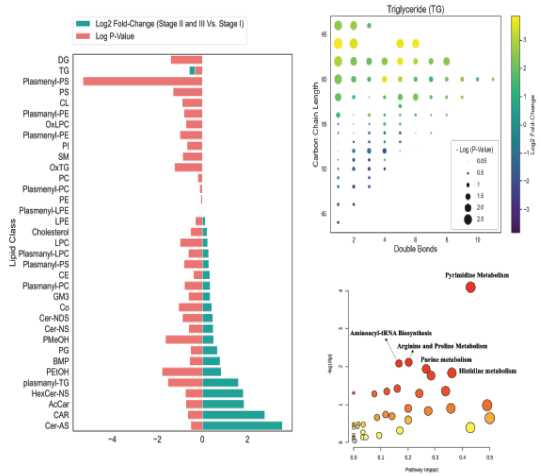


UHPLC-HRMS Analysis

Untargeted Metabolomics
Untargeted Lipidomics



Statistical and Pathway Analysis

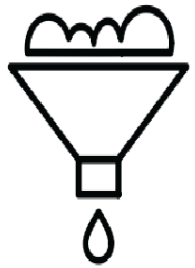


2. Artificial Intelligence and Statistical Analysis

Data Pre-Processing & Cleaning

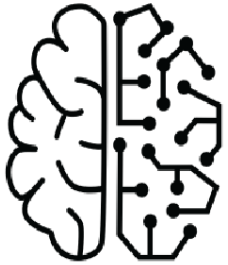
- Normalizing
- Transforming
- Scaling

Feature Selection & Dimension Reduction



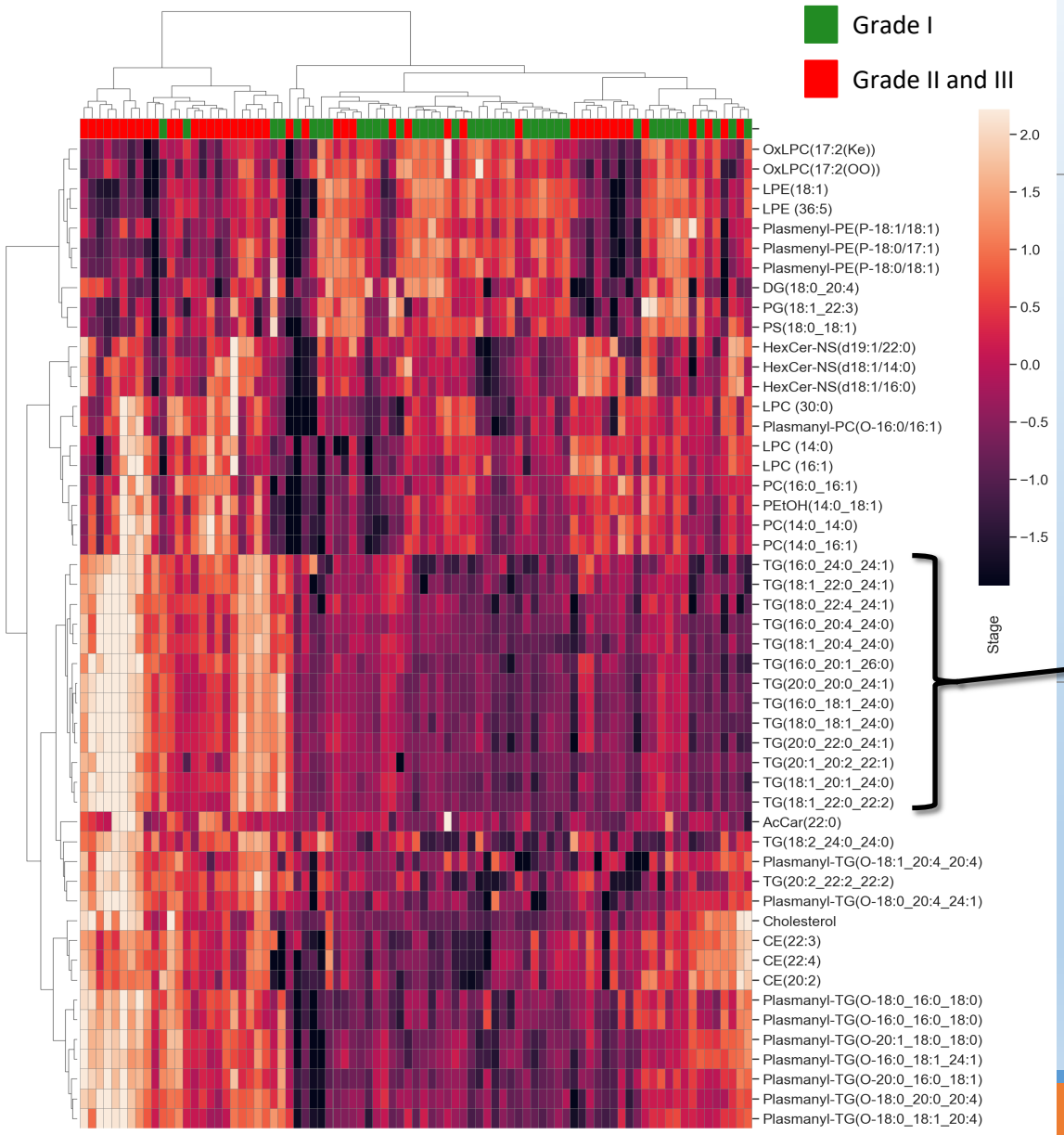
- Python's Scikit-Learn
- ExtraTreesClassifier
- Top 50 Features

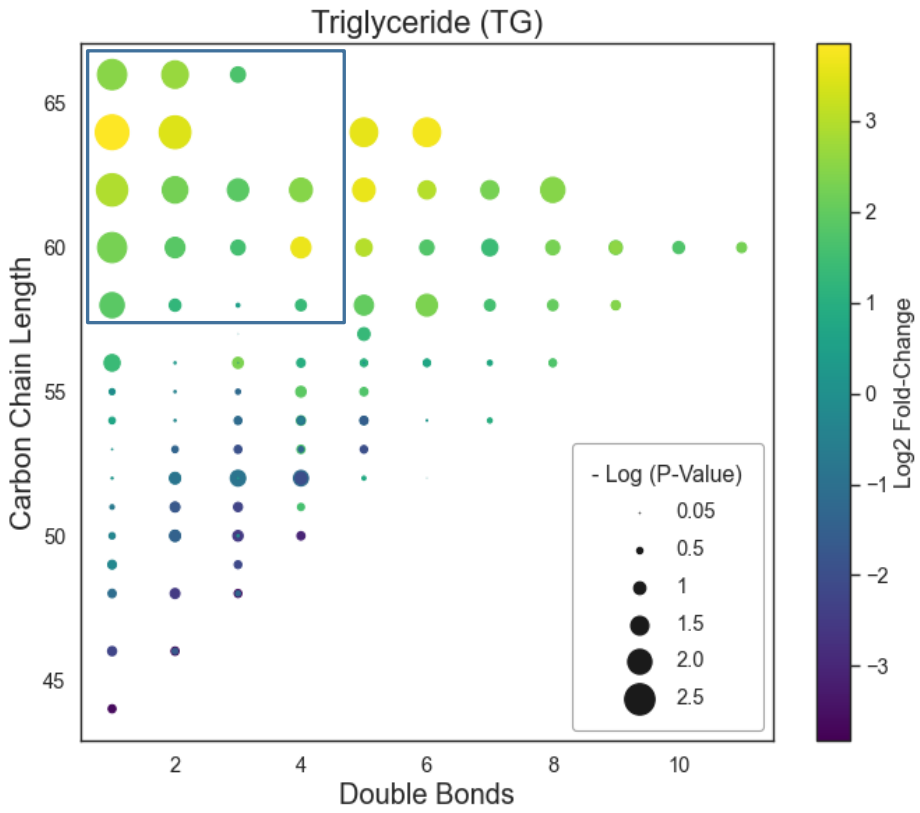
Machine Learning & Model Training



- Distinguishing Important Metabolites
- Identifying Potential Biomarkers
- Meningioma Grade Classifying

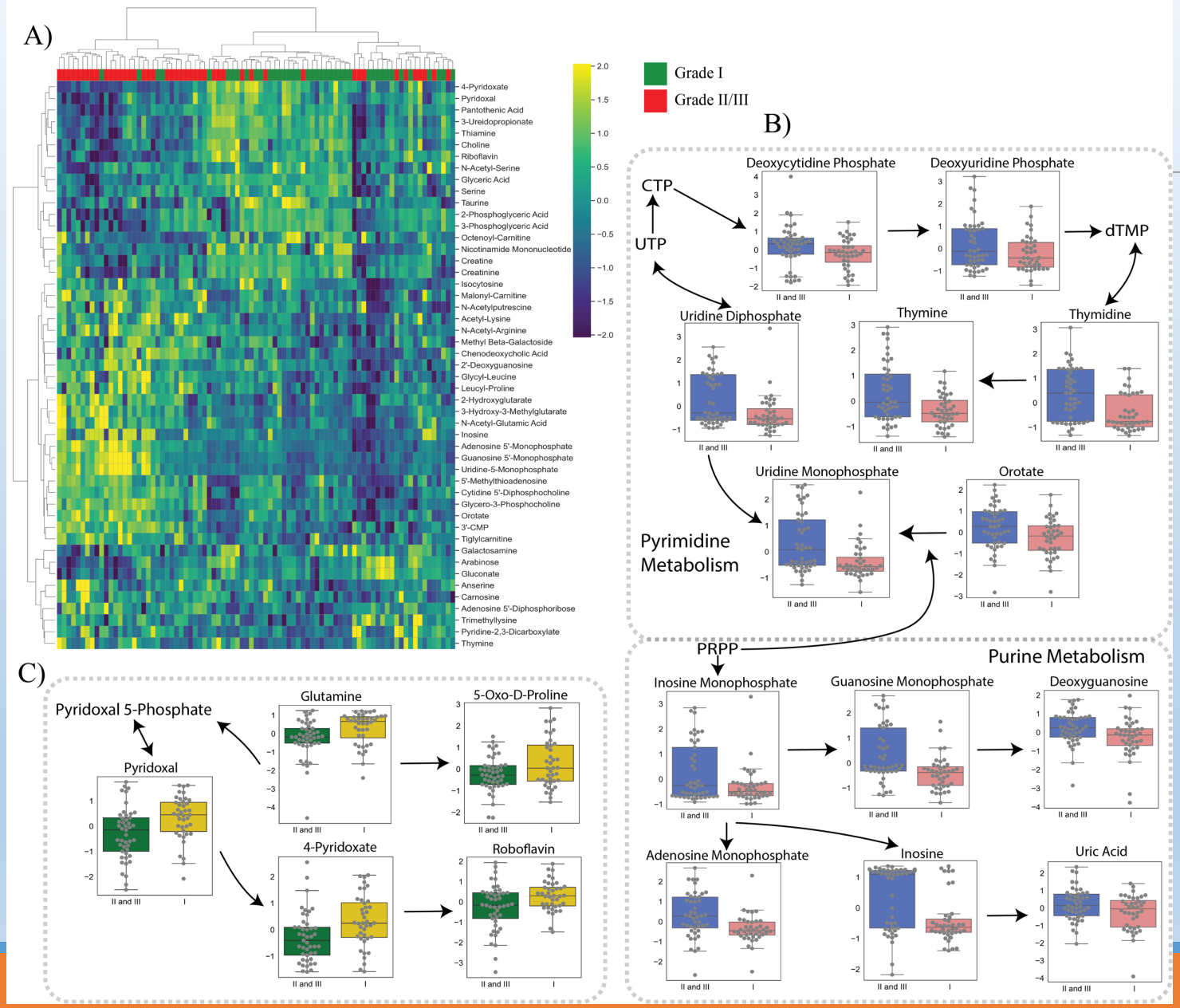
Lipidomics





- Key Results:
  - TG Levels **higher** in grade 2 and 3
  - **Higher** levels of long chain TGs in 2/3
- Cancer cells require lipids for growth
  - Obtain fatty acids from lipogenesis

Metabolomics



Pathways implicated

Key Results:

- Pyrimidine and Purine metabolism is <u>upregulated</u> in grade II/III
- Vitamin B6 metabolism is <u>downregulated</u> in grade II/III

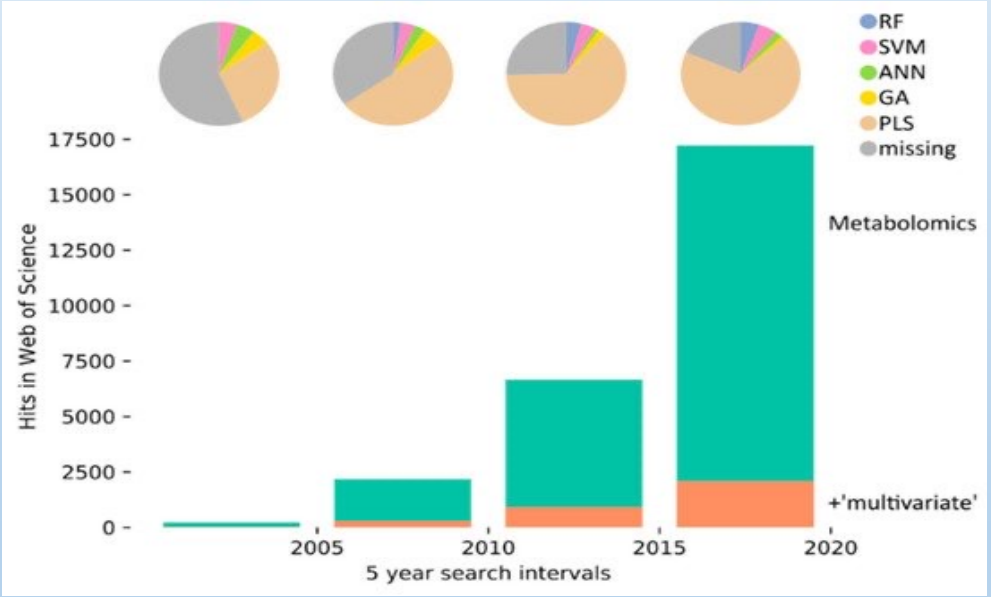
Multi-Omics Data Analysis

MS Metabolomics data analysis is complicated:

- Large data environment
- Nonlinear data
- Heterogeneous data

Machine learning methods applied to MS-based multi-omics ease data analysis and can support clinical decisions, guide metabolic engineering, and stimulate fundamental biological discoveries

The integration of multiple omics levels will enhance our understanding of the interactions among the different biological layers



History of machine learning (ML) in metabolomics

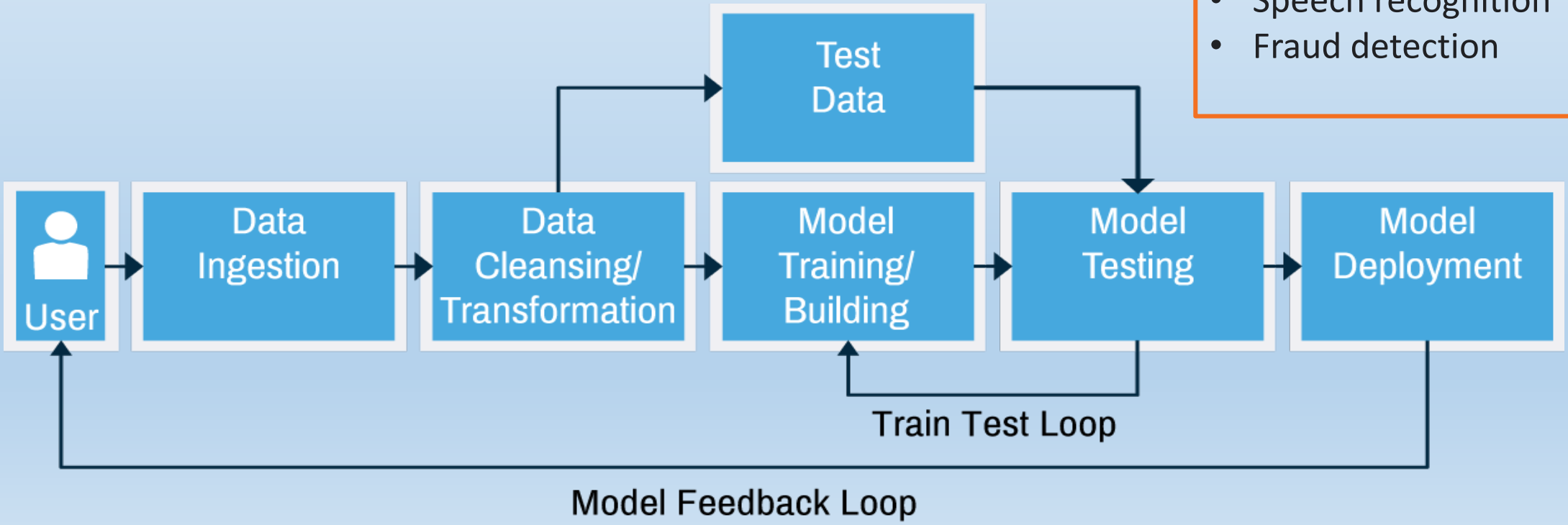
Liebal, U. W.; Phan, A. N. T.; Sudhakar, M.; Raman, K.; Blank, L. M., Machine Learning Applications for Mass Spectrometry-Based Metabolomics. *Metabolites* **2020,** *10* (6).

Machine Learning

Machine learning is the process of teaching a computer system how to make accurate predictions when fed data.

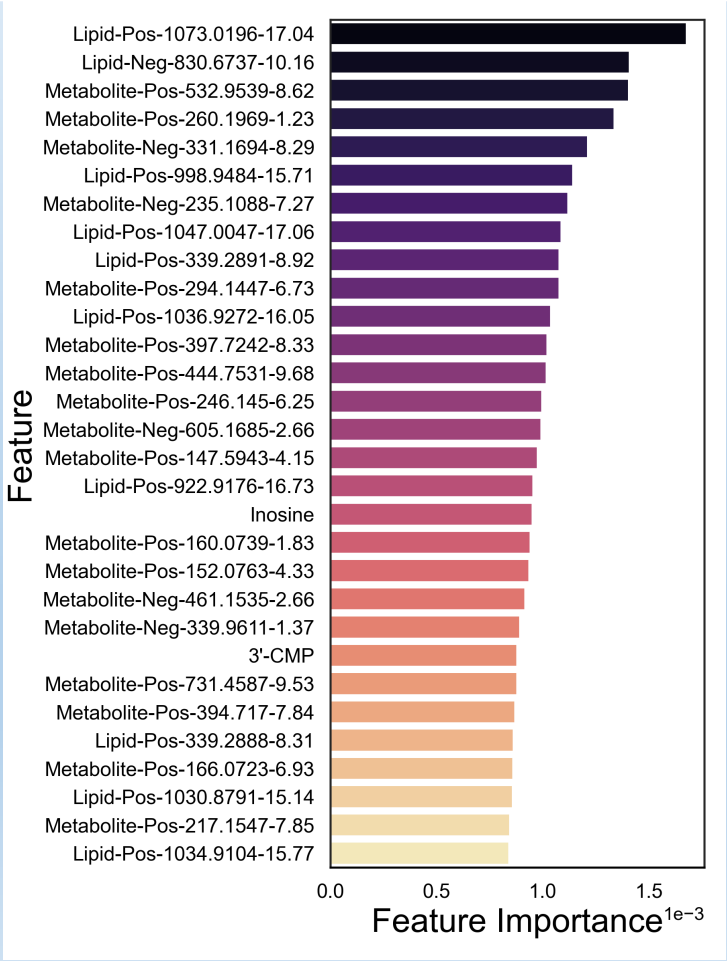
Machine Learning Applications:

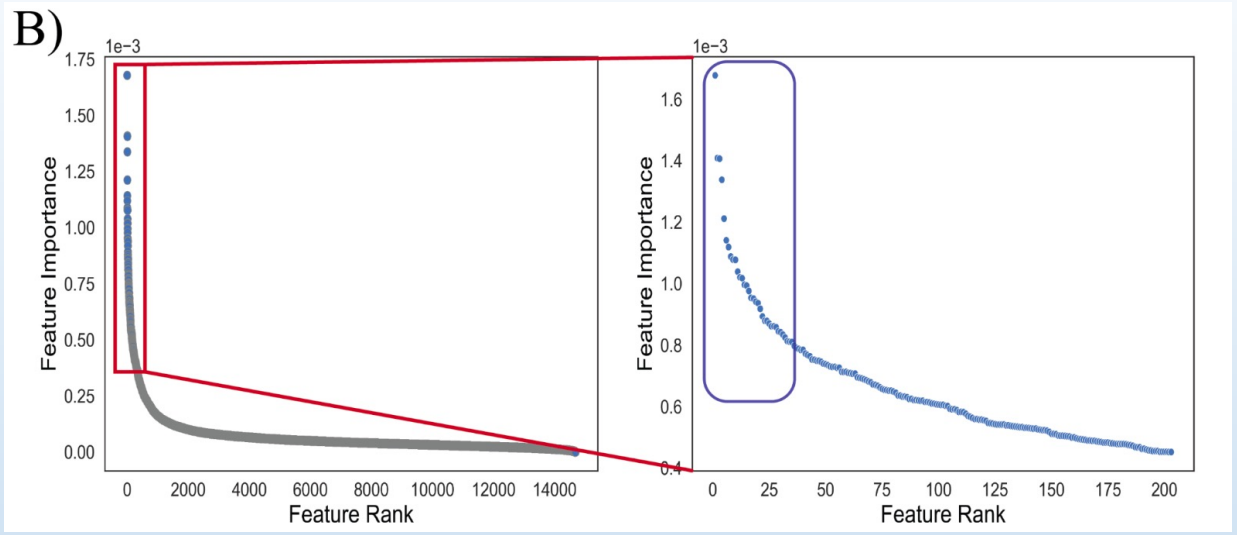
- Face detection
- Handwriting recognition
- Computer vision
- **Healthcare**
- Voice interfaces
- Speech recognition
- Fraud detection

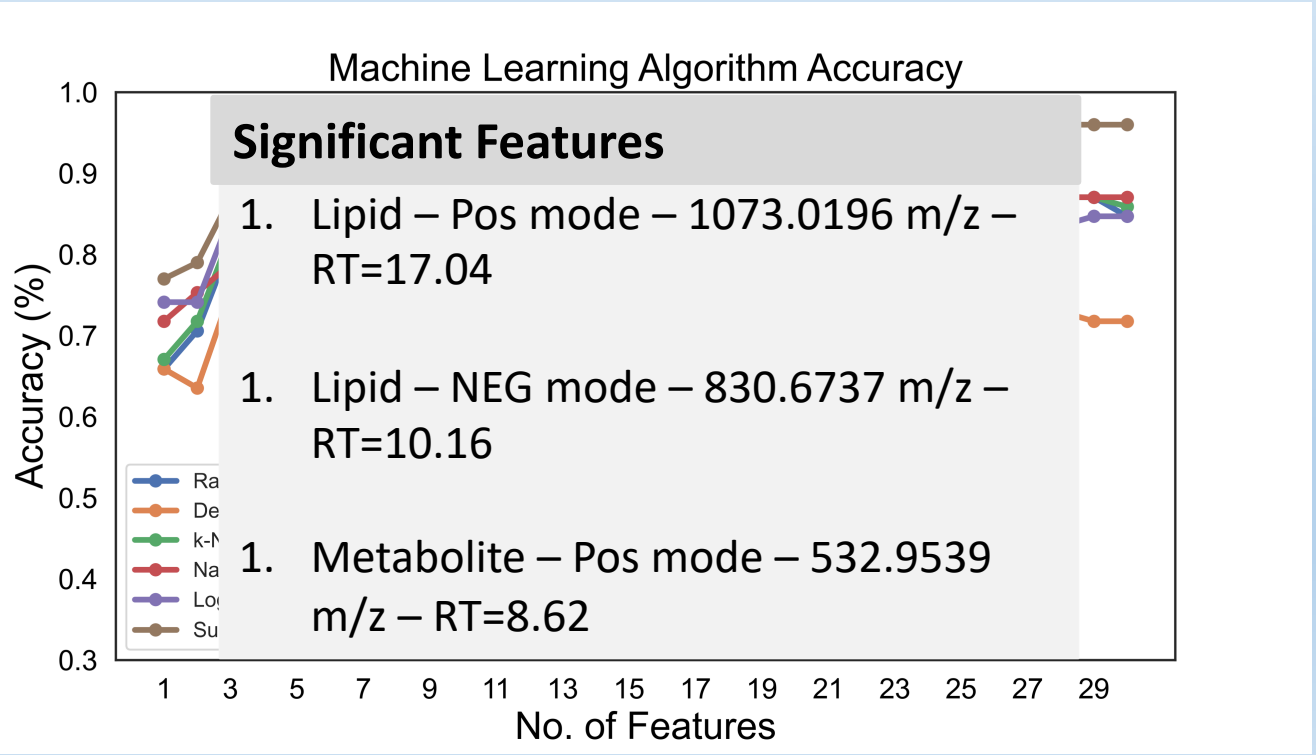


Feature Selection

- **> 17,000 Features detected**
- **Python Scikit-Learn Package**
- **ExtraTrees Package**
  - Classification and regression based on an ensemble of decision trees.







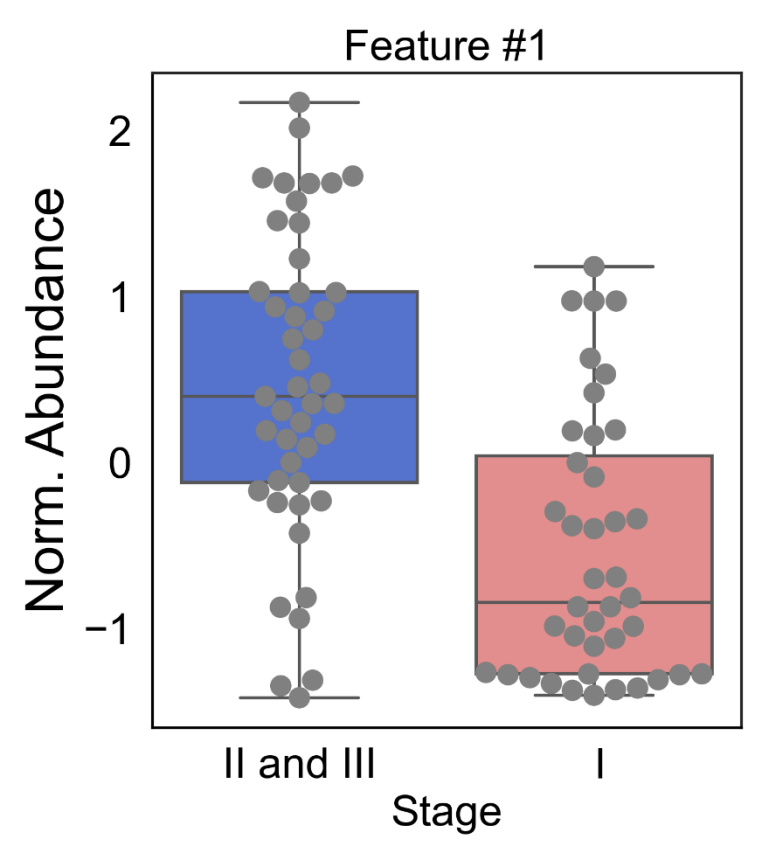
Significant Features

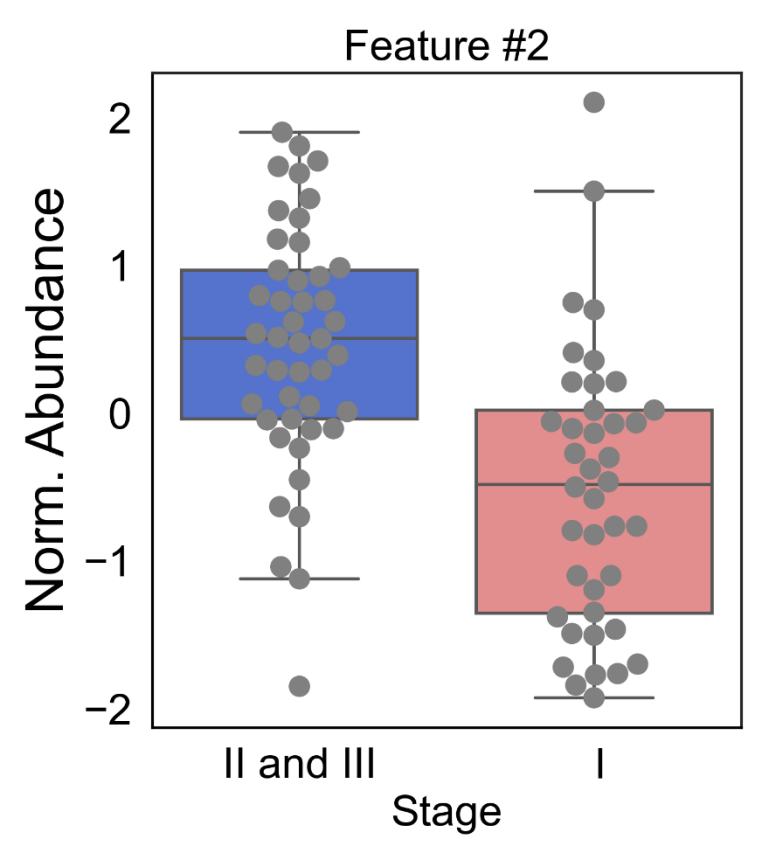
1. Lipid – Pos mode – 1073.0196 m/z – RT=17.04

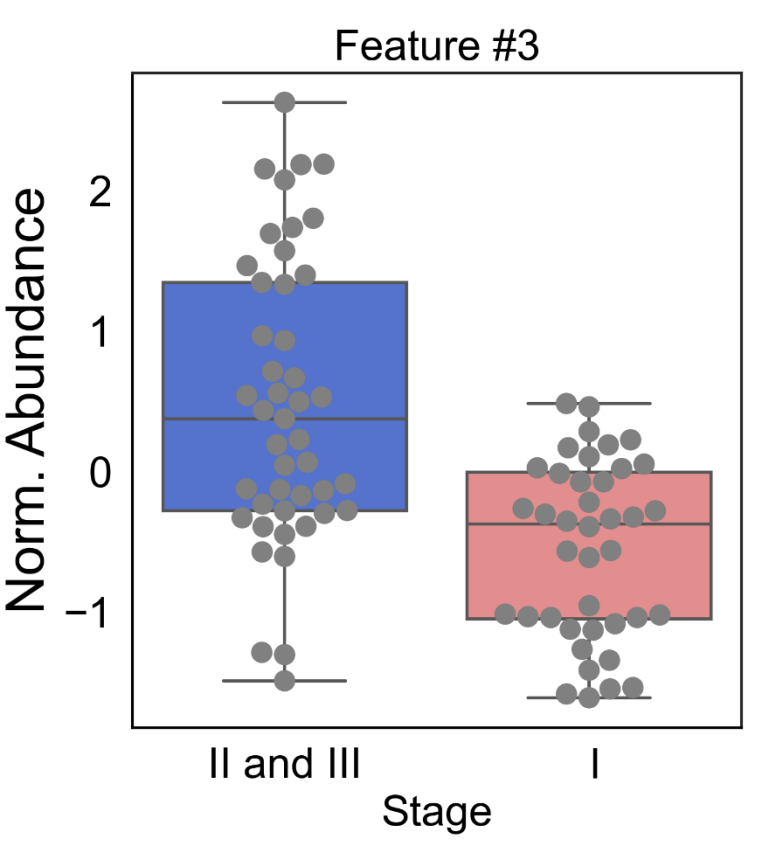
1. Lipid – NEG mode – 830.6737 m/z – RT=10.16

1. Metabolite – Pos mode – 532.9539 m/z – RT=8.62

Response for 3 of the features

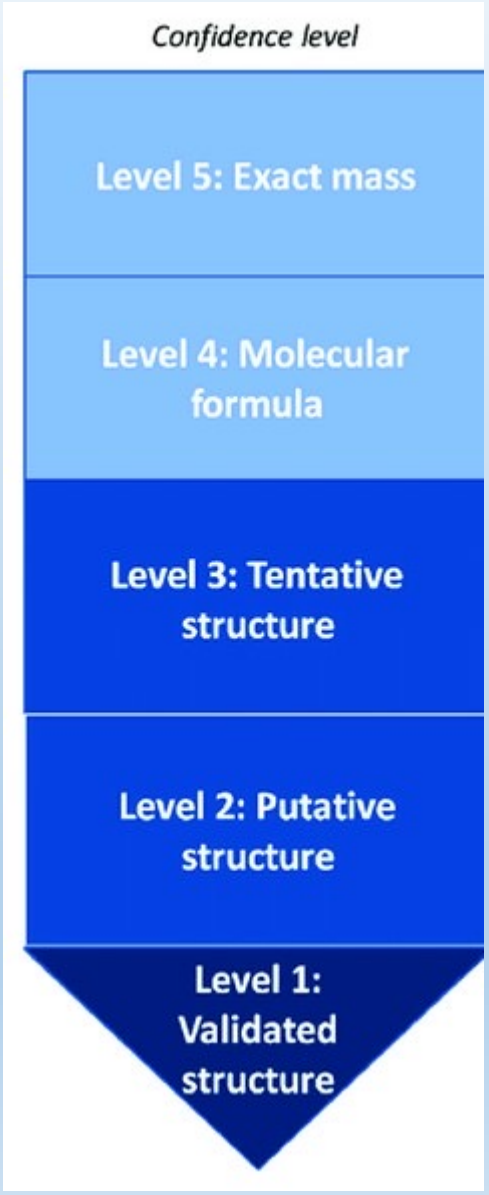






All are elevated in Stage 2/3

Unknown Identification



- Metabolomics/chemical database search based on the exact mass
- Fragmentation spectrum
- Standard check (if available)

Significant Features

1. Lipid – Pos mode – 1073.0196 m/z – RT=17.04
1. Lipid – NEG mode – 830.6737 m/z – RT=10.16
1. Metabolite – Pos mode – 532.9539 m/z – RT=8.62

1. TG(18:2)(24:0)(24:0)
2. GalCer(d18:0/22:0)
3. Peptide with 14 Amino-Acid

Summary of Meningioma

Combine machine learning using lipids and metabolites identified 5 unknowns to classify Grade 1 vs Grade 2/3

- MS/MS interpretation was used to identify 2 of the 5 unknowns as lipids
- 1 of the unknowns is a small peptide from the metabolite analysis
- 2 are still unknown, these will require NMR for fraction collection

Cannot classify Grade 2 vs Grade 3 because of small sample sizes

- Only had 15 grade 3 and it is still difficult by current histopathology to classify grade 2 and 3

Biological interpretation is on going

- Lipids are harnessed by cancer cells
- The increased TG content could be localized to lipid droplet accumulation in tumors
- The longer chain, but not significantly increased PUFA content could be significant
- 24:0 and 24:1 were associated with TG

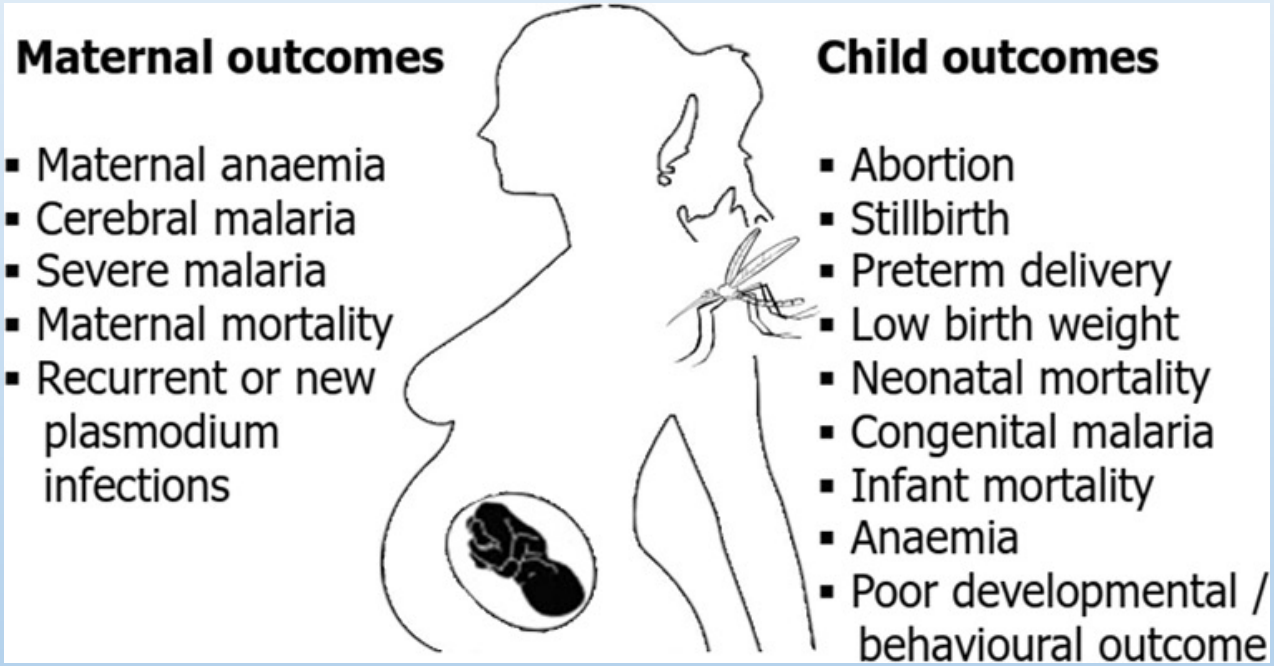
DBS analysis

- MS analysis of Phenylalanine started in 1990, and the first validation article published in 1993 which established DBS of phenylalanine and tyrosine for the detection of PKU (phenylketonuria)
- The original extraction and derivatization method was robust and flexible
- Electrospray ionization allowed for automated analysis to enable rapid scale up and screening applications-Newborn screening (Sensitivity)
- Neutral Loss Scanning/Precursor ion scanning and SRM used in data acquisition (Specificity)
- Flow Injection Analysis (FIA) for rapid analysis and reduced cost (Speed)
- However, standard DBS has limited precision of ~15% RSD

Dried blood spot applications

- Newborn screening (NBS) is most commonly known application
- Other areas of use
  - Toxicology
  - Infectious disease
  - Therapeutic drug monitoring
- Home sampling/testing had increased awareness with COVID-19 and the number of products increased rapidly especially lateral flow assays
- Desire for use in multi-omics especially in areas where collecting liquid samples can be difficult
- Need for small sample sizes that are stable

Malaria and pregnancy



DOI: 10.5772/66342

- Malaria affects both the mother and baby
- Pregnant woman have a higher risk of infection
- Primigravida women are at higher risk for adverse pregnancy outcomes

https://www.ncbi.nlm.nih.gov/books/NBK1710/

Studies of malaria during pregnancy in Kenya

In 2001, established NIH-supported, hospital-based lab in western Kenya

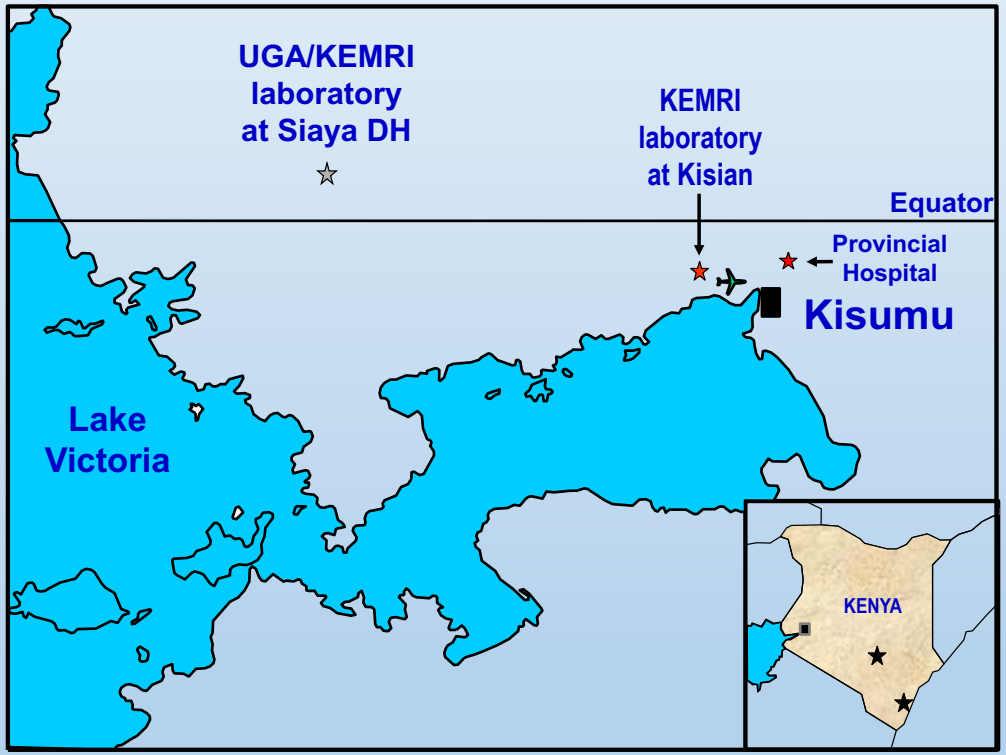
Partnered with the Kenya Medical Research Institute and Kenya Ministry of Health

Over 6 years, recruited >1000 parturient women and ~200 non-pregnant mothers

Collected clinical data, placental blood and tissue, and peripheral blood

Collection of samples in remote locations is difficult and DBS are commonly used here

Collaboration with Dr. Julie Moore, UF



Eunice Kennedy Shriver NICHD National Institute of Child Health & Human Development

National Institute of Allergy and Infectious Diseases

Studies of malaria during pregnancy in Kenya

Gravidity	PM+	LBW infant	HIV+
1	98/402 (24.4%)	73/400 (18.3%)	58/401 (14.5%)
2	37/244 (15.2%)	25/242 (10.3%)	52/241 (21.6%)
≥3	51/400 (12.8%)	22/400 (5.5%)	124/395 (31.4%)

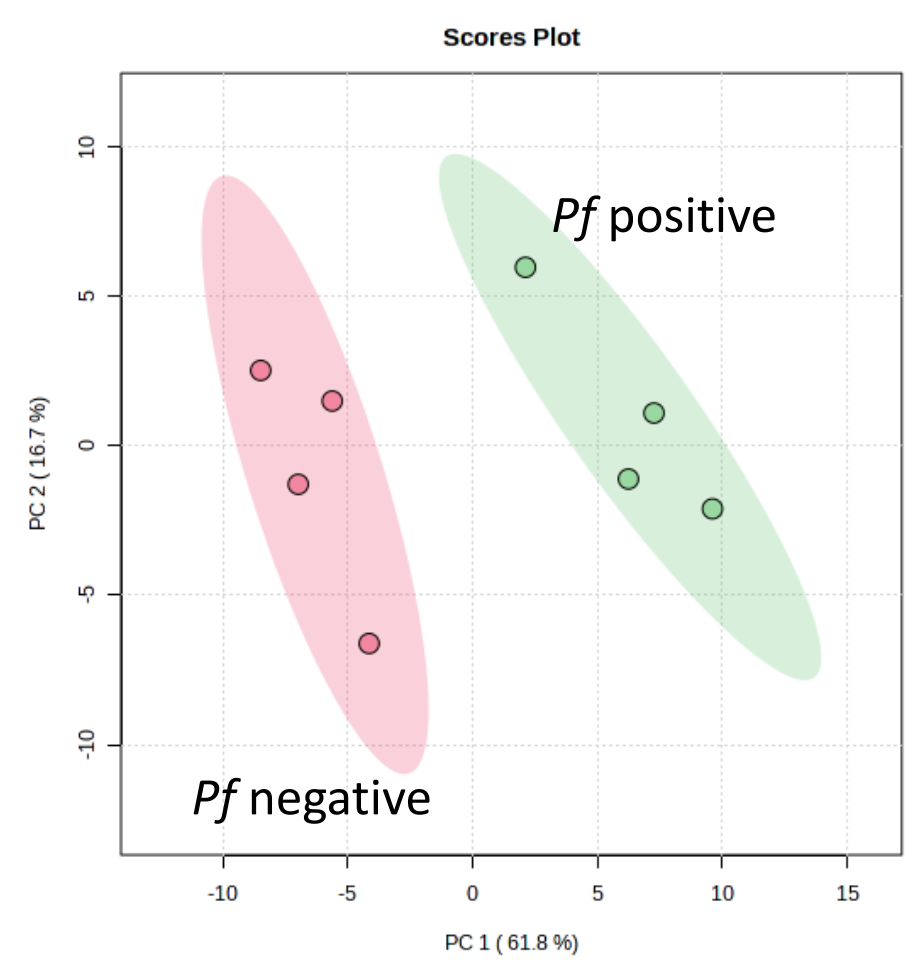
- Finding a primigravid or secundigravid placenta devoid of signs of malaria (i.e., hemozoin deposition) is challenging in this setting
- Placenta manages to support successful pregnancy most of the time
- Fundamental question: what is happening when it *isn't* successful?

Examples of placental blood samples collected

Dried blood spots

- Blood samples collected were spotted on paper to simplify transport
- MS analysis from dried blood spots is commonly performed for newborn screening, but not as common in lipidomic or metabolomic profiling studies
- Spots were punched out and extracted for lipids and metabolites (Folch)
- Analyses conducted on a Thermo Q-Exactive with reversed-phase separation
- Iterative exclusion MS/MS analysis for increased lipid coverage
- Compared the lipid expression from malaria positive vs control

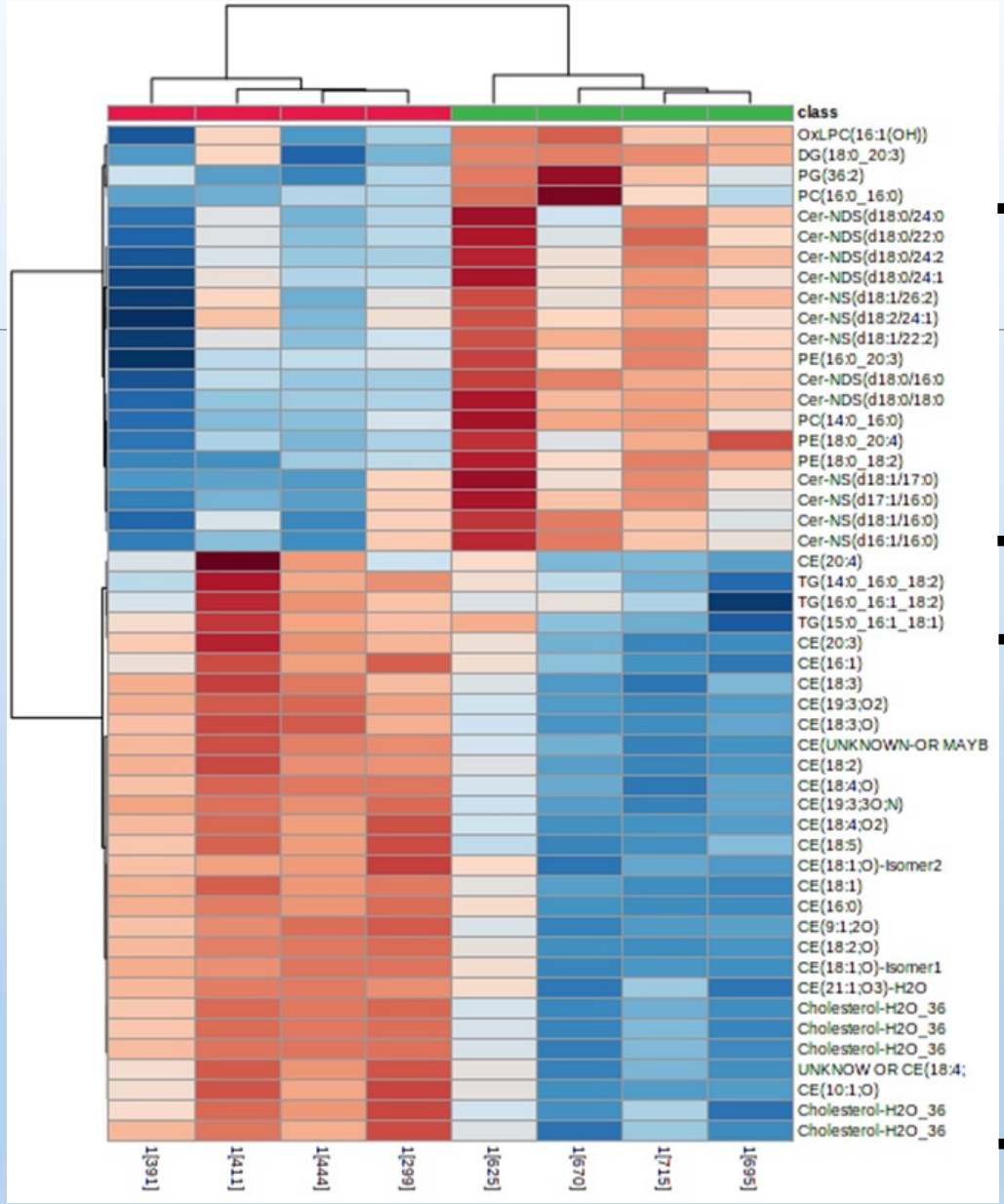
Principal components analysis



- 4 biological replicates in this pilot study
- Clear separation using PCA is observed, primarily along PC1 with ~62% of the variance explained by this component
- Biological variability could be related to small changes in blood volume in each blood spot or other factors
- Pilot analysis shows feasibility is good so a larger study can now be conducted

Lipid expression

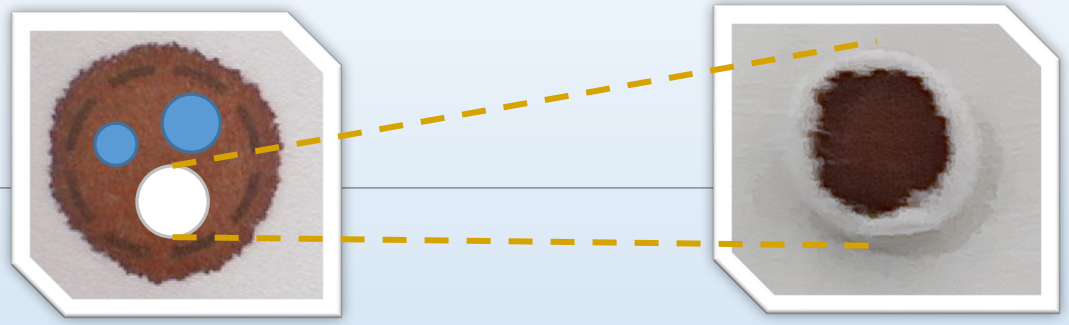
- Ceramides are elevated in Malaria positive
- Cholesterol esters are depleted in malaria-positive patients as well as some triglycerides
  - Decrease of CEs in the host could translate to an increase in CEs and cholesterol in the parasite for growth and replication, but this was not tested
- *Plasmodium* species lack the ability to synthesis cholesterol *de novo* and can uptake it from the local environment
- This is a pilot so drawing early conclusions should be done carefully
- Just collected data on nearly 100 subjects as a follow up



The Dried Blood Spot
Revisited

Classic Guthrie

To a Precise blood spot**



50 µL – Whatman 903 / Ahlstrom 226
± 15%

10 µL – Ahlstrom 222
± 5%

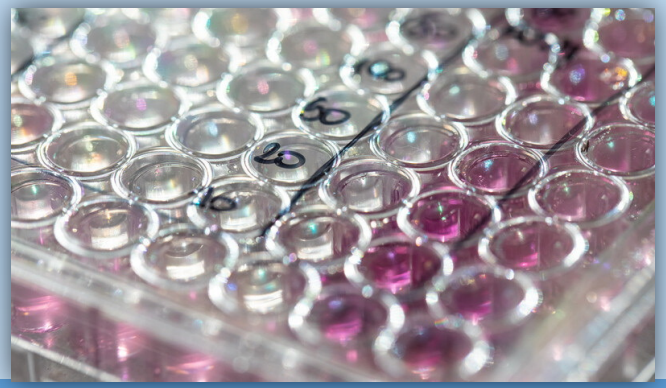
3.2 mm ~ 3.4 µL

4.8 mm ~ 7.6 µL

6.3 mm ~ 8.9 µL

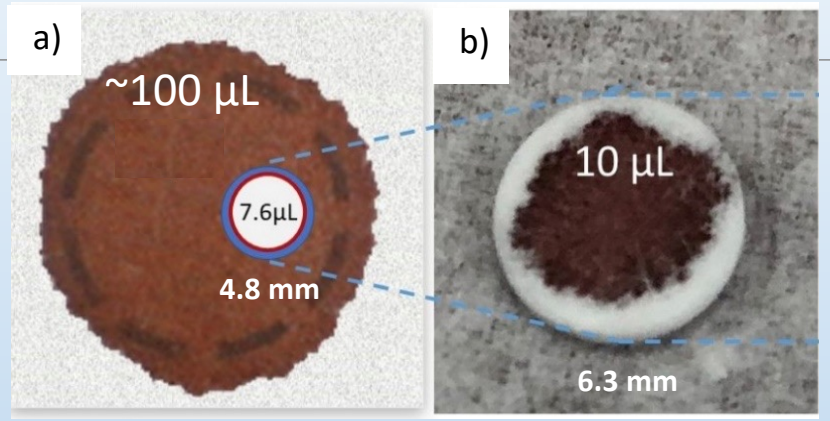
6.3 mm = 10 µL

For size comparison



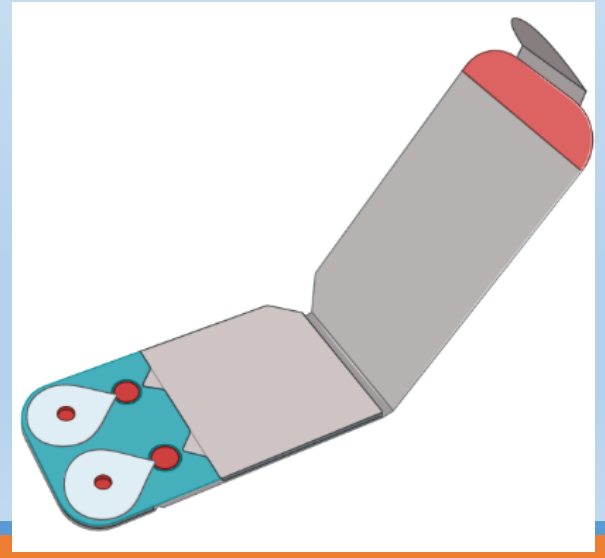
Microwell plate diameter is 8.6 mm

Standard DBS vs quantitative DBS



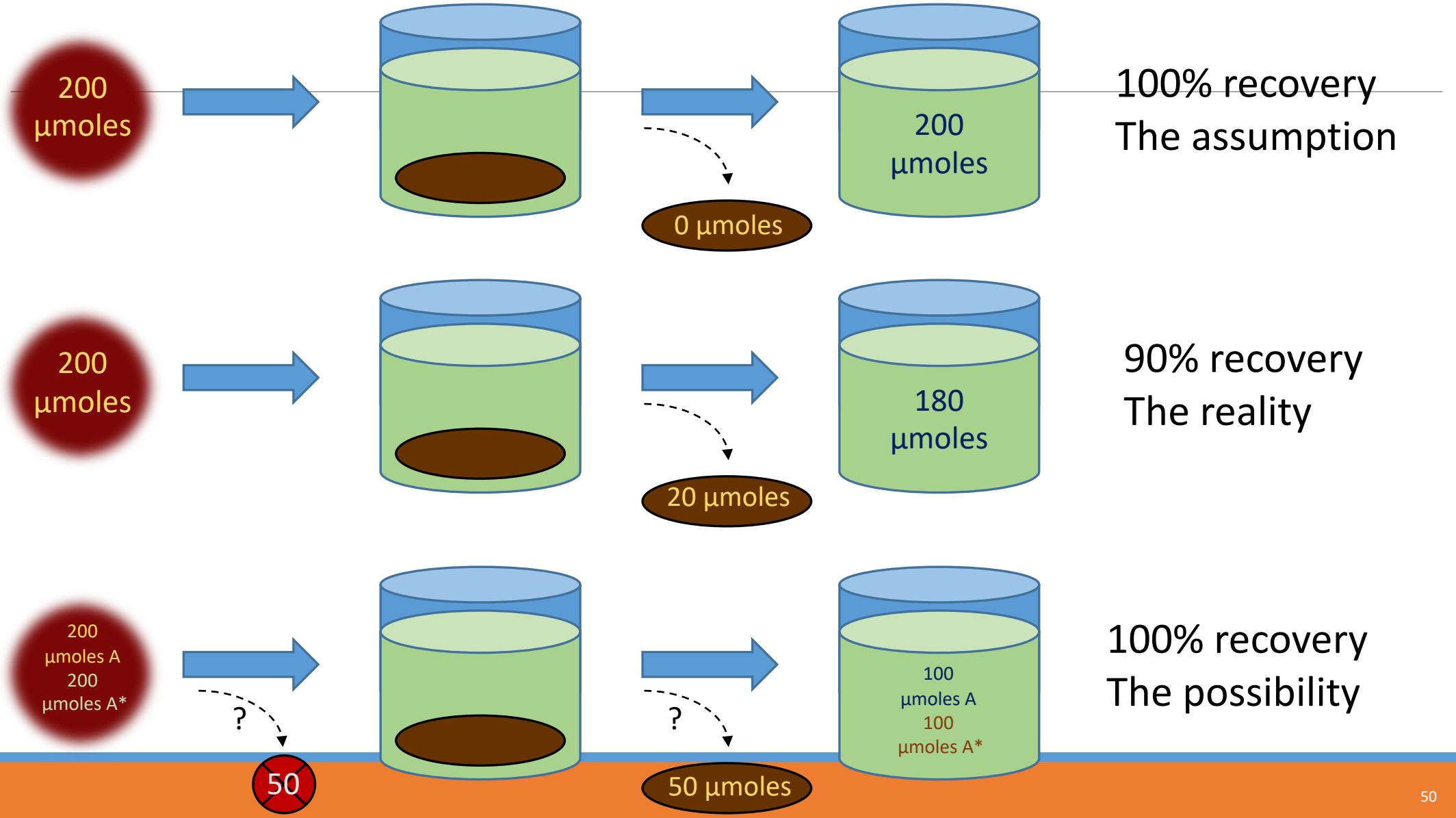
Standard Guthrie 10 µL spot

- Standard DBS (classic Guthrie Card, newborn photo)
  - A dried volume, not precise, but routine and cheap
- qDBS (quantitative DBS)
  - A blood spot with a precise volume collected and dried (e.g. 10 µL)
  - Still fingerstick



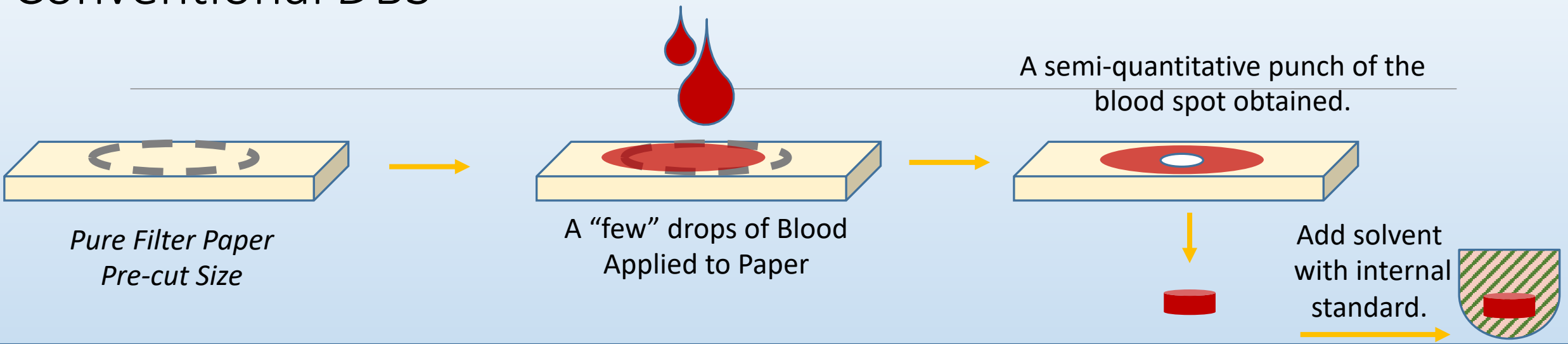
https://capitainer.se/technology/

Extraction and Recovery

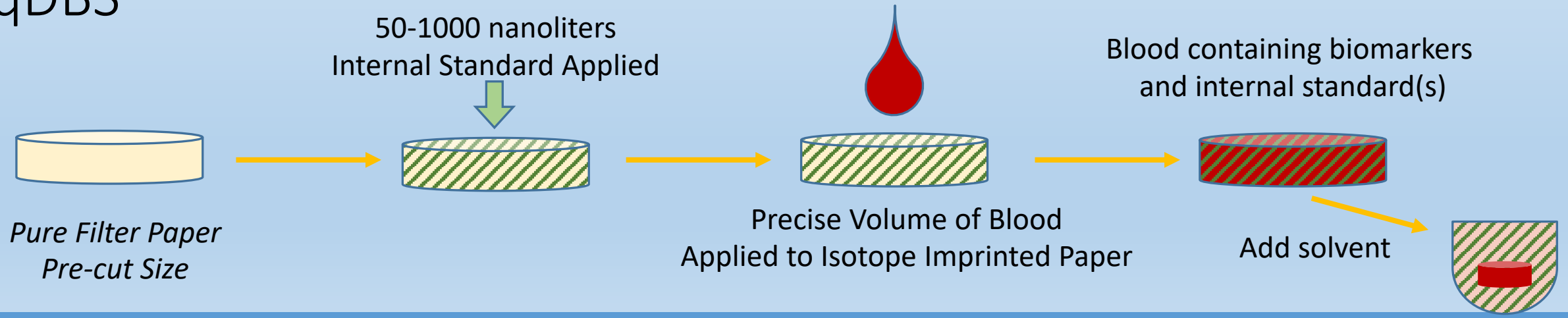


Internal quantitative DBS – How it works

Conventional *D*BS

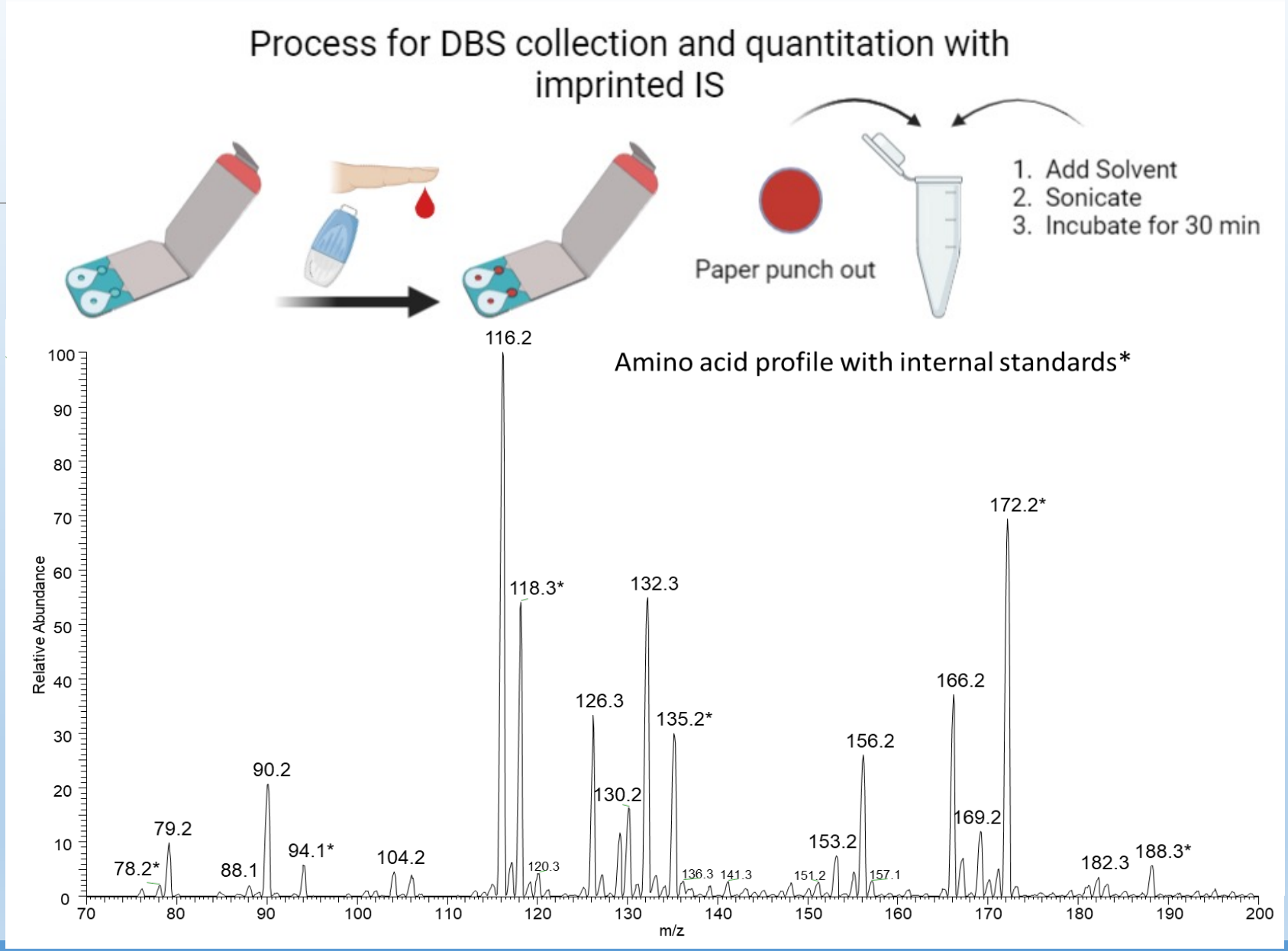


iqDBS

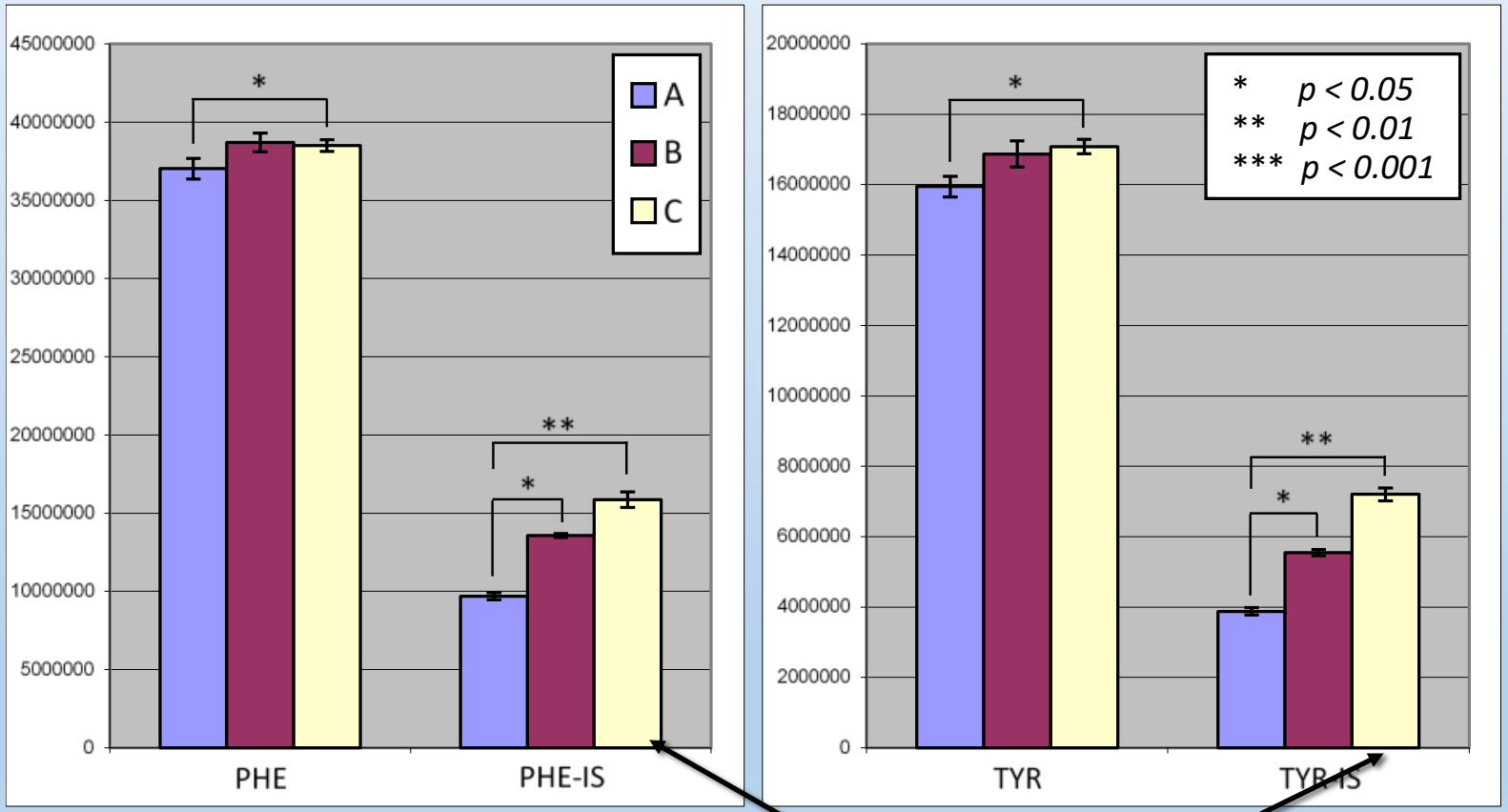


A stable isotope blood spot

- Internal quantitative dried blood spot (iqDBS)
- Precise volume of blood
- Mixed with a precise concentration of stable isotopic internal standard (CIL)



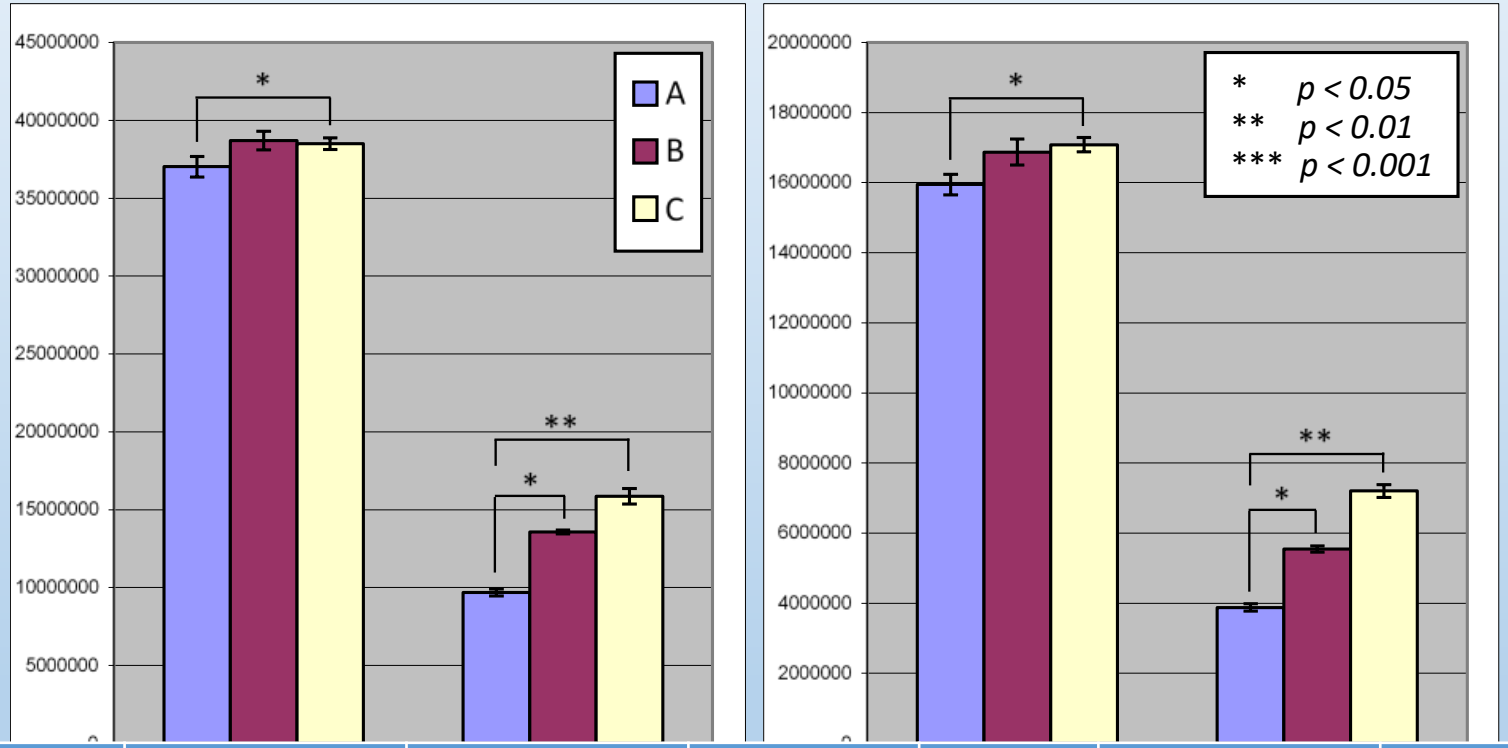
Analysis with 3 different IS extraction approaches



- A = Stable isotope IS spiked in the card, plasma added
- B = IS spiked into plasma and then added to collection card
- C = Plasma added, IS in Methanol for extraction (traditional DBS extraction)

IS in MeOH does not account for extraction efficiency from paper, hence signal is higher for the IS

Analysis with 3 different IS approaches

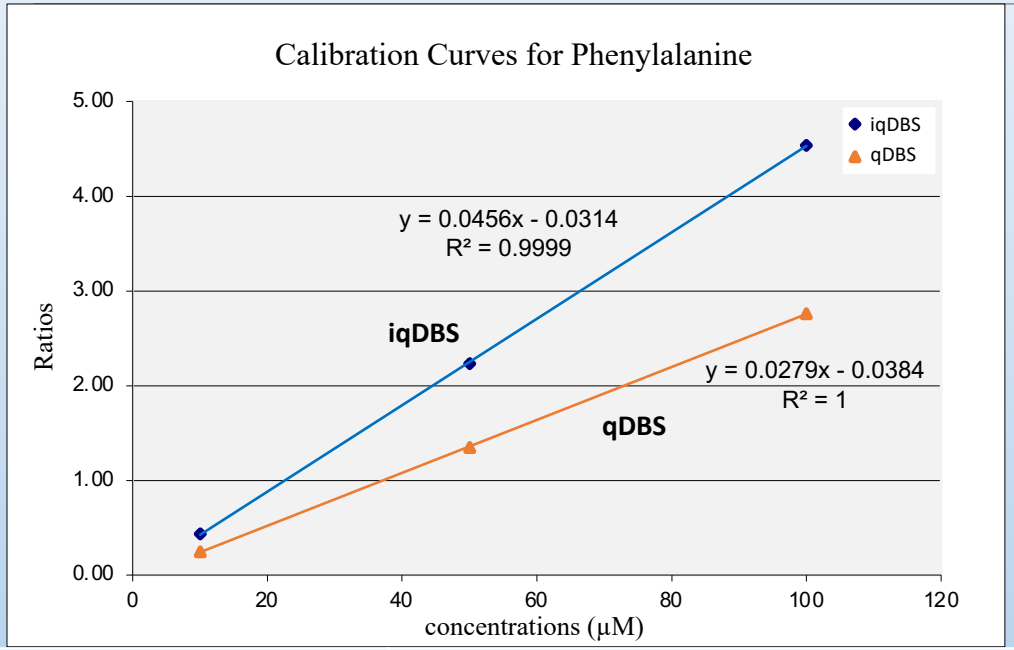


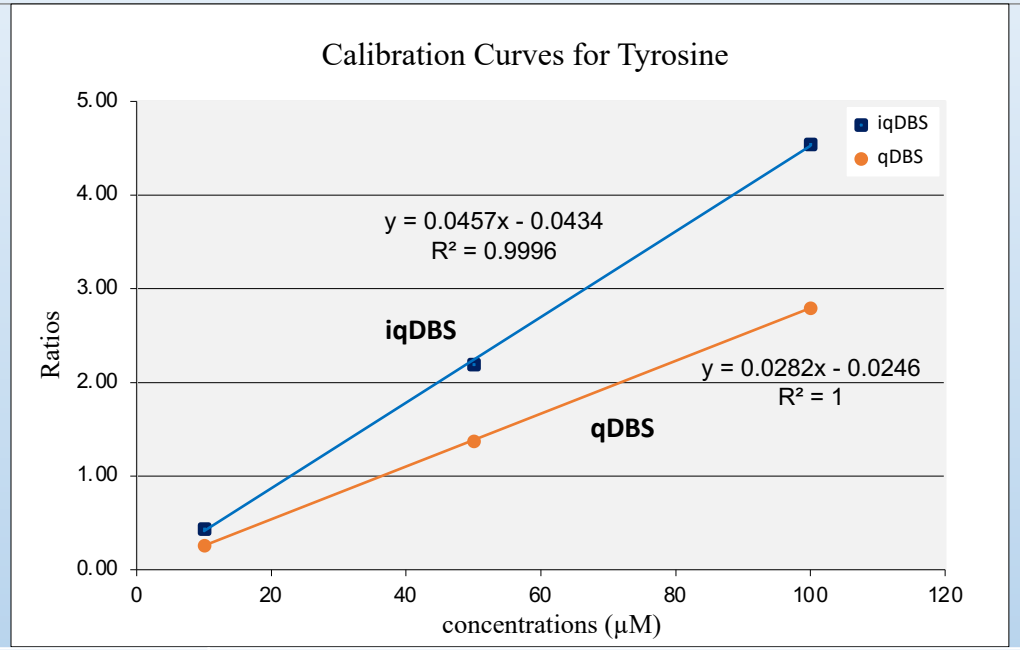
- A = Stable isotope IS spiked in the card, plasma added
- B = IS spiked into plasma and then added to collection card
- C = Plasma added, IS in Methanol for extraction (traditional DBS extraction)

(n=3)	PHE	PHE-IS	PHE/IS	RSD	TYR	TYR-IS	TYR/IS	RSD
A	37032244	9672365	3.83	3.0%	15944088	3878135	4.11	3.6%
B	38702563	13564783	2.86	0.8%	16873603	5548202	3.06	1.4%
C	38505270	15863641	2.43	4.4%	17081405	7200376	2.37	3.0%

IS in MeOH does not account for extraction efficiency from paper, hence signal is higher for the IS

Quantitative Results PHE and TYR- 2 IS conditions





(n=3)	iqDBS				qDBS			
Concentration (µM)	PHR/IS*	RSD	TYR/IS*	RSD	PHE/IS	RSD	TYR/IS	RSD
10	0.44	2.8%	0.44	4.5%	0.25	1.4%	0.26	1.6%
50	2.23	1.7%	2.19	1.0%	1.35	1.9%	1.38	3.1%
100	4.54	4.0%	4.55	4.5%	2.76	2.6%	2.80	2.8%
Plasma	3.39	1.5%	3.04	3.1%	2.24	2.4%	2.01	2.0%
Calculated concentrations	PHE	75 µM	TYR	67 µM	PHE	82 µM	TYR	72 µM

Future work

Evaluate long-term stability

- Amino acids tested so far and no loss over a month
- Acylcarnitines to test next since they are known to degrade

Test in a large patient cohort

- Currently looking a different opportunities

Establish a Multi-Omic study

- Tested Metabolomic and lipidomic profiling, but need to add proteomics

Samples sent to Africa for a larger infectious disease study for genomics

Work with a company to manufacture the internal standardized cards

Internal Standard(s) in the card

Mixes directly with a precise volume of collected blood

- Currently, blood is collected, dried and extracted with IS added during extraction
- This fails to account for extraction efficiency from the card, introducing error

If degradation occurs, IS will degrade similar; therefore, concentration will still be accurate regardless of any breakdown

- Currently conducting breakdown studies

Multiple stable isotopes can be added to the card

- We have tested up to 500 for untargeted metabolomics using the IROA yeast standards

The lab process is simplified, internal standards are in the card, no punch out as entire card with IS can be placed in a well plate and extracted

Large screening studies can be more readily standardized since any errors in IS preparation are eliminated

Acknowledgements

Hoda Safari Yazd, PhD
William Clapp, PhD
Jesse Kresak, PhD
Manasi Kamat, PhD
Joel McCray
Sina Feizbakhsh, PhD
Wenqian Li
Lizzi Flammer
Jae Hwan Lee
Abdulkarim Alfaez
Michael Christopher
Geoff Faden
Ed Faden

NIH National Institutes of Health

UF PathLabs UNIVERSITY of FLORIDA

macmod SMARTER CHROMATOGRAPHY

TRAORDINARY JOY!

https://garrettlab.pathology.ufl.edu/

Thank you for your attention