# Use of isotopes in metabolomics

**Stephen Barnes, PhD**

**03-02-18**

## Synopsis

- **Natural abundance isotopes**
  - **The value of the M+2 ion**
- **Tracing a metabolic pathway**
  - **$^{14}$C-labeling a precursor for qualitative analysis**
- **Isotope ratio outlier analysis**
  - **Discerning real metabolites**
  - **Identifying metabolites**
  - **Use in quality control/quality assurance**

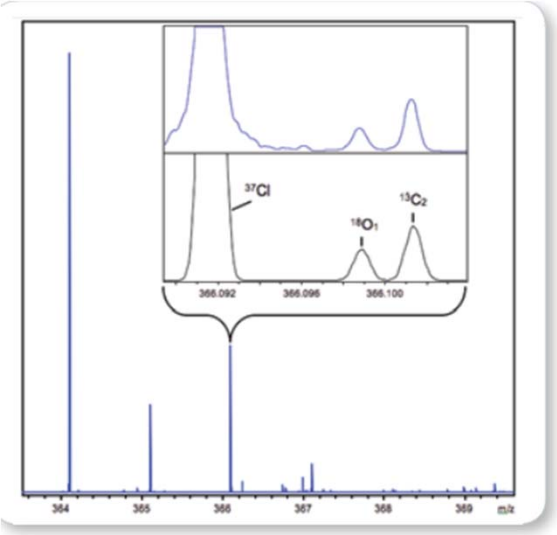
# Value of natural isotopes

- The natural abundance of isotopes enables the investigator to determine the charge state of an ion
  - The principal contribution to $[M+H]^+$ or $[M-H]^-$ isotope ions comes from $^{13}C$ (~1.1% of all carbon atoms)
  - The intensity of the $^{13}C$ isotope ion increases relative to the number of carbon atoms
  - There is often an observable $^{13}C_2$ isotope peak

# Value of the [M+/-H+2] peak

- The mass difference due to a nominal increase in mass of 2 contains a lot of information
  - These are isotopic mass differences for each of the common elements
    - $^{1}H_2/^{2}H_2$ 2 x 1.006277 = 2.012554 (0.012%)
    - $^{12}C_2/^{13}C_2$ 2 x 1.003355 = 2.006710 (1.078%)
    - $^{14}N_2/^{15}N_2$ 2 x 0.997035 = 1.994079 (0.364%)
    - $^{16}O_2/^{17}O_2$ 2 x 1.004217 = 2.008434 (0.038%)
    - $^{16}O_2/^{18}O_1$ 1 x 2.004246 = 2.004246 (0.205%)
    - $^{32}S_2/^{33}S_2$ 2 x 0.999387 = 1.998774 (0.752%)
    - $^{32}S_2/^{34}S_1$ 1 x 1.995796 = 1.995796 (4.252%)

# The importance of the M+2 ion



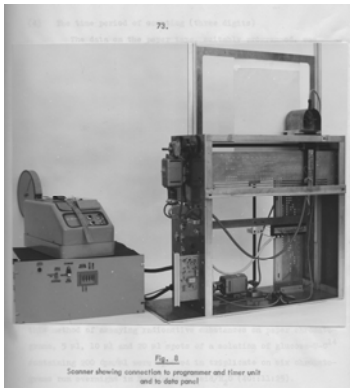

From Bruker

# Using isotopes to trace a pathway

- Early studies (1930s) used $^{2}H$, $^{13}C$ and $^{15}N$ labeling to map pathways
  - Limited to 1-200 *m/z* mass range
- 1950s/60s $^{14}C$-radiotracers
  - 2D-Paper or thin layer chromatography
  - Radio gas chromatography
    - labeling of specific carbon atoms

# Origins of practical metabolomics

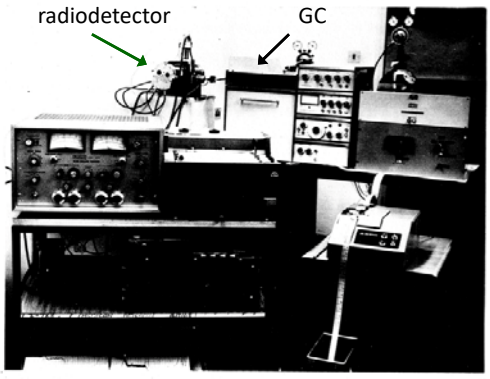
## Imperial College 1967-1970



Radio 2D-paper chromatography scanner with digitization of collected data

The room had 20 of these scanners – data analyzed by a central computer (in 1968)

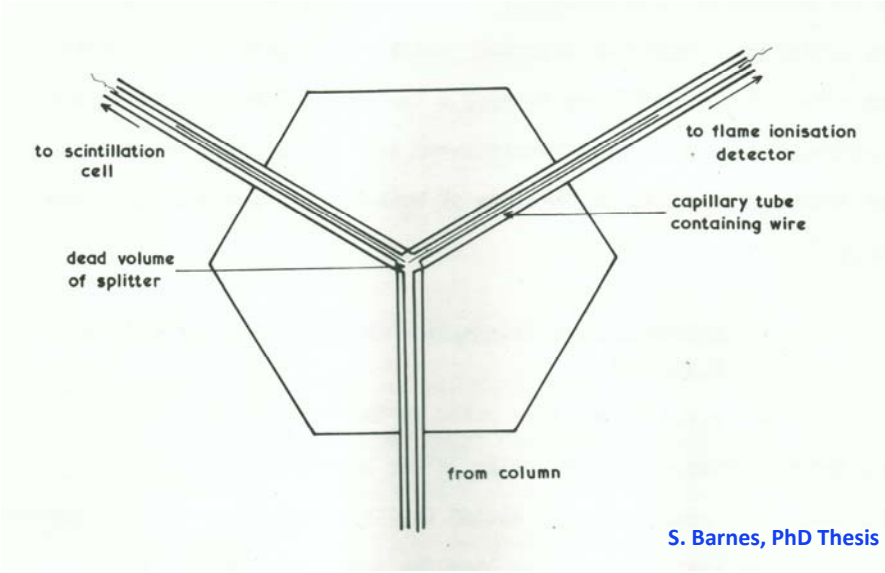
Courtesy of K.R. Mansford, PhD




Radio gas-liquid chromatography with digitization of collected data

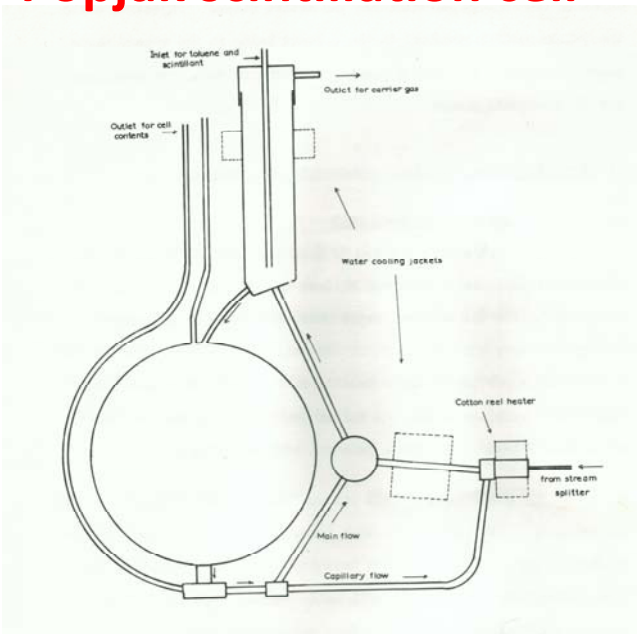
Developed this for my PhD work (1967-1970) to study glucose metabolism in acellular slime moulds

# Stream splitter for radio GC

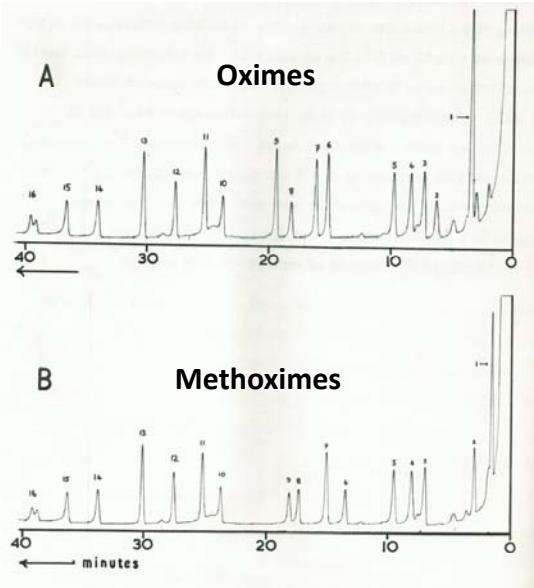



S. Barnes, PhD Thesis

# Popjak scintillation cell




# GC of glycolytic and Krebs cycle intermediates

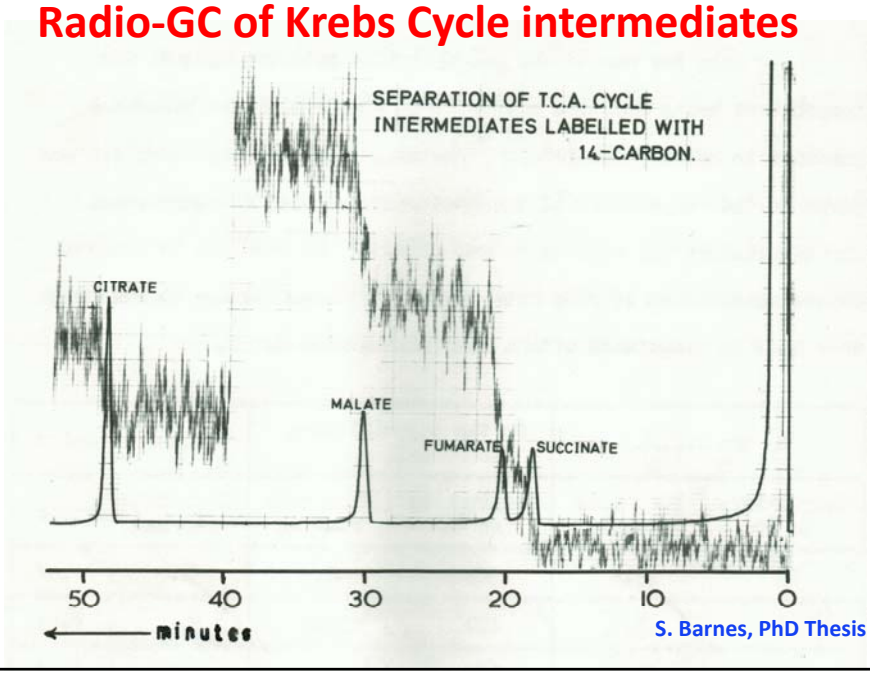



**Temperature programming of TMS ester/ethers on a 5' x ¼ inch packed column of Chromosorb W coated with OV-1 liquid phase**

**1=pyruvate , 2=?? , 3=phosphate , 4=succinate , 5=fumarate, 6=oxaloacetate, 7=malate, 8=αKG, 9=hexadecane, 10=αGP, 11=citrate, 12=α–D-glucose, 13=β–D-glucose, 14=docosane, 15=F6P, 16=G6P**

**S. Barnes, PhD Thesis**

# Radio-GC of Krebs Cycle intermediates

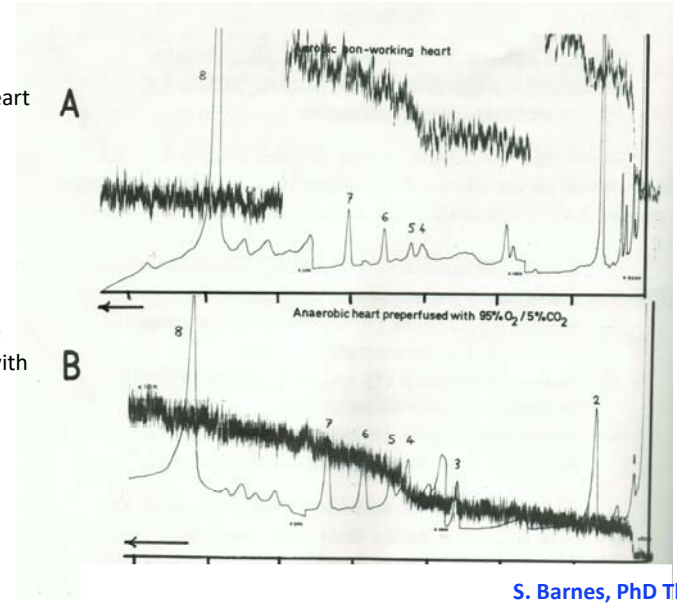



S. Barnes, PhD Thesis

# Radio GC analysis of beating heart

Aerobic beating heart

Anerobic beating heart perfused with 95% $O_2$/5% $CO_2$

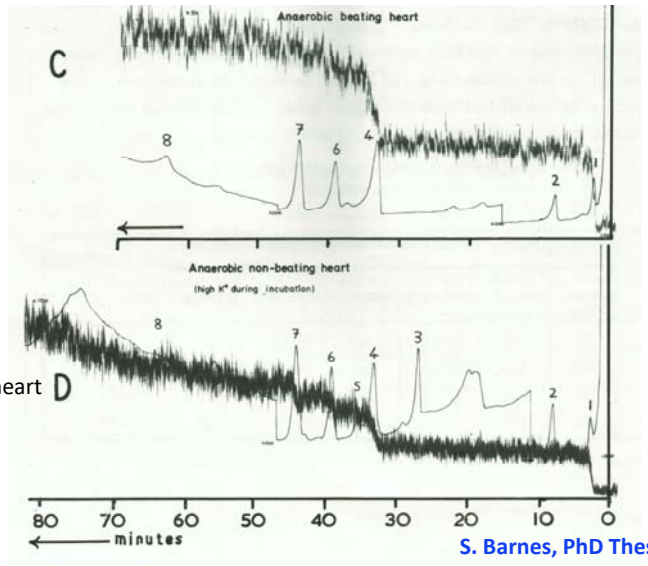



S. Barnes, PhD Thesis

# Radio GC analysis of anerobic heart

Anerobic beating heart

Anerobic non-beating heart

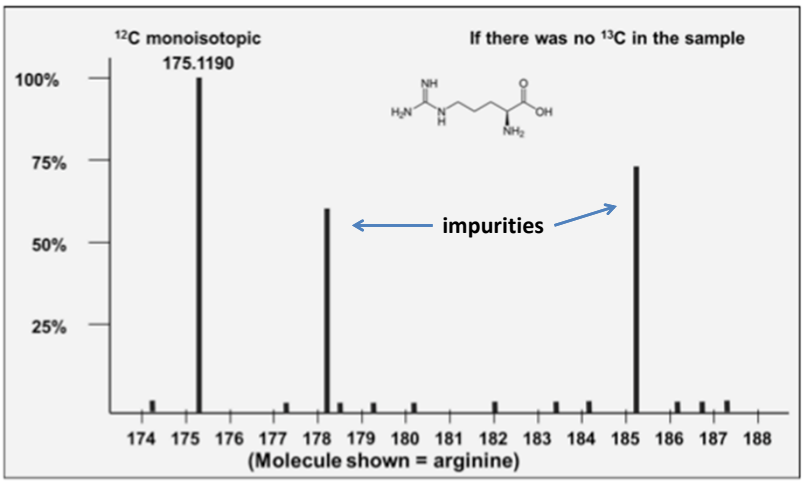



S. Barnes, PhD Thesis

# Tracking metabolites with IROA

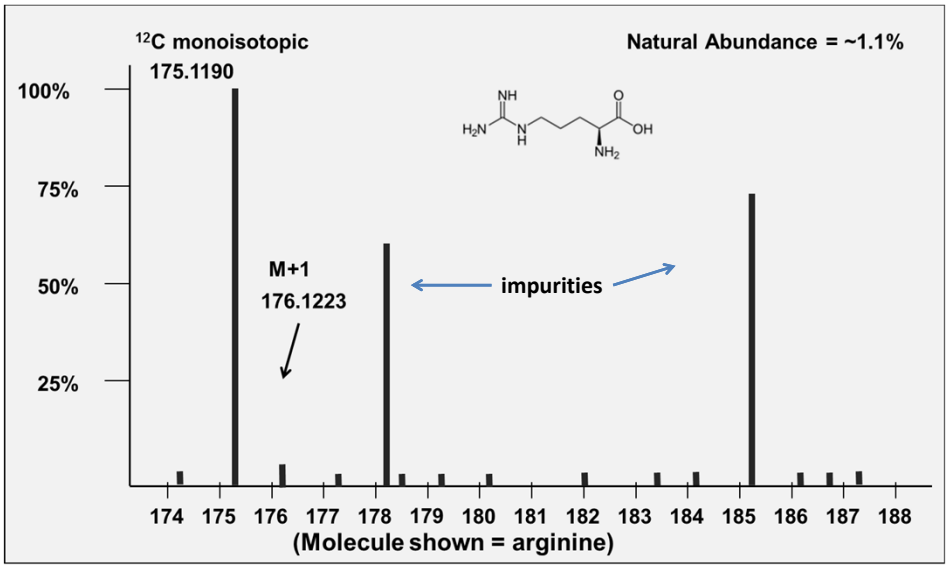
- **Isotope ratio outlier analysis (IROA)**
  - **Not used for flux analysis, but rather to create a unique signal for metabolites**
  - **Used for LC-MS (and possibly GC-MS)**
  - **Designed to distinguish between metabolites of interest and background signals**
  - **Requires uniform labeling at the 95% and 5% $^{13}$C-enrichment levels**

# All $^{12}C$ in arginine $[M+H]^+$



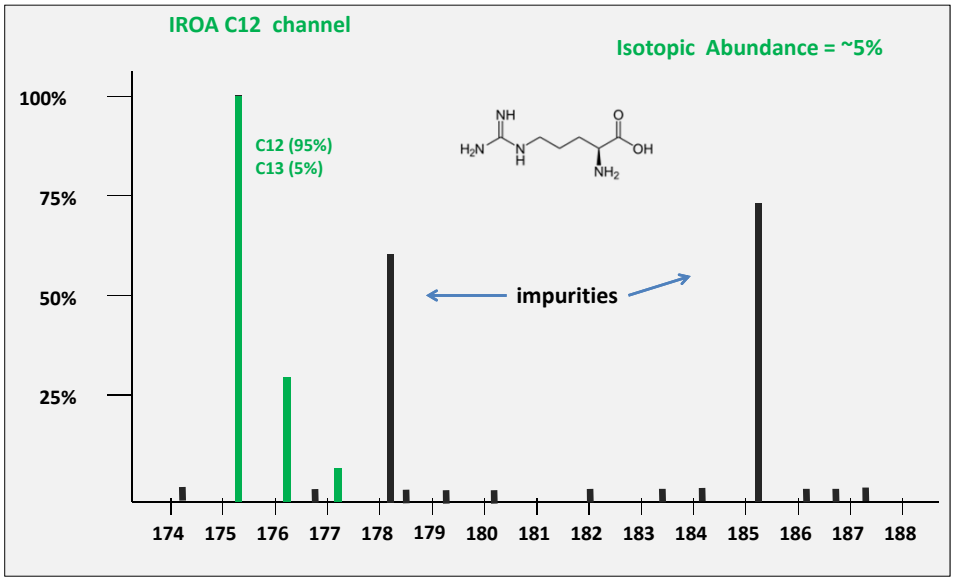

IROA Technologies

# Natural abundance of $^{13}C$ in arginine



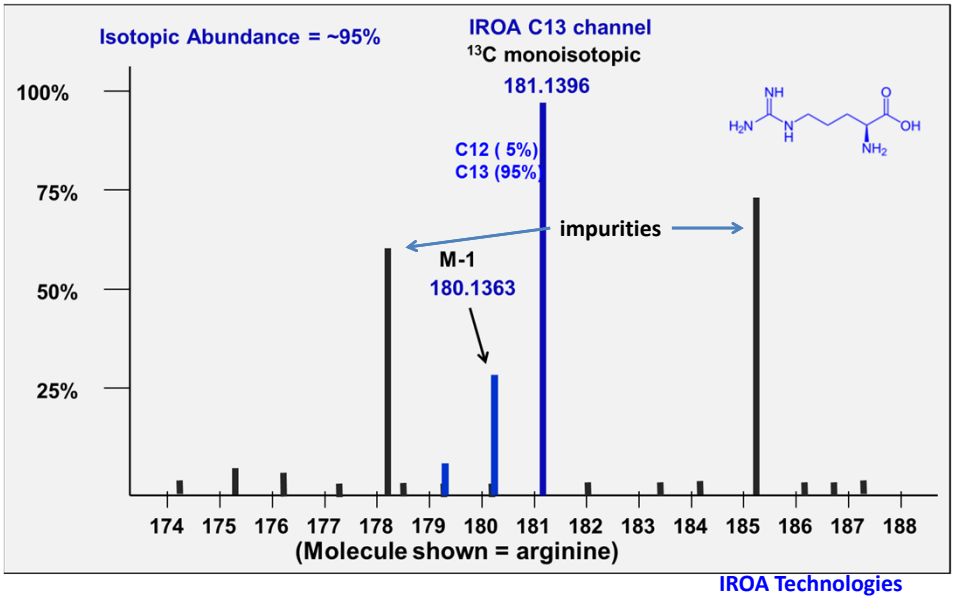

IROA Technologies

# Making $^{13}C$ abundance = 5%



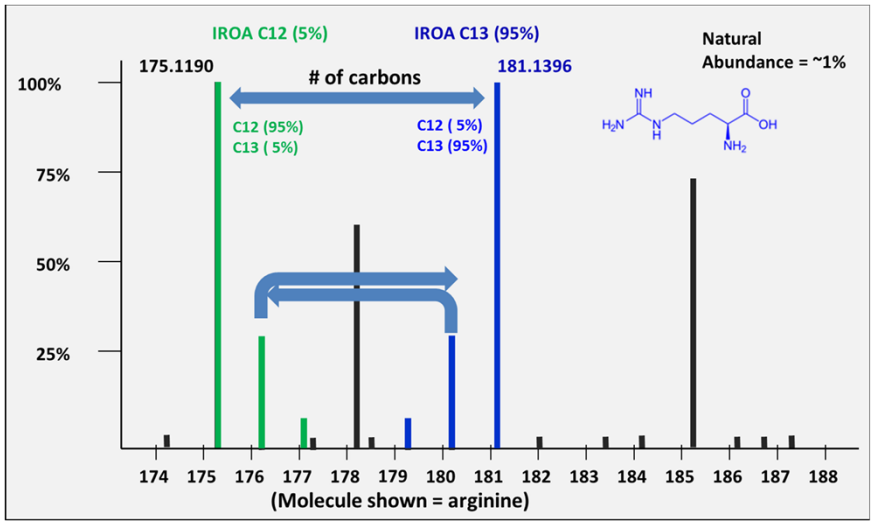

IROA Technologies

# Making $^{13}C$ abundance = 95%



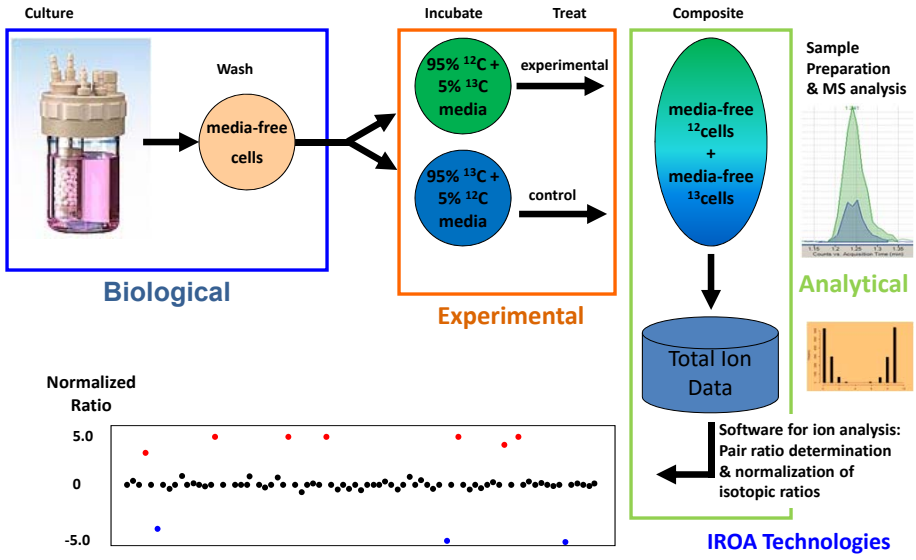

IROA Technologies

# Pairing the 5% and 95% $^{13}$C-labeling



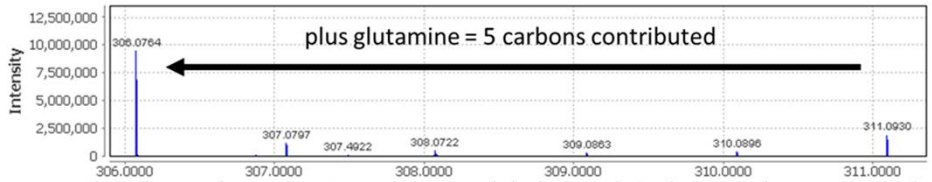

IROA Technologies

# The IROA approach

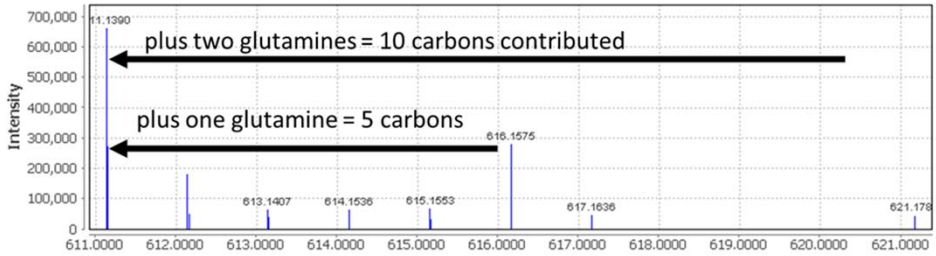



IROA Technologies

# The span of isotopes = # carbon atoms



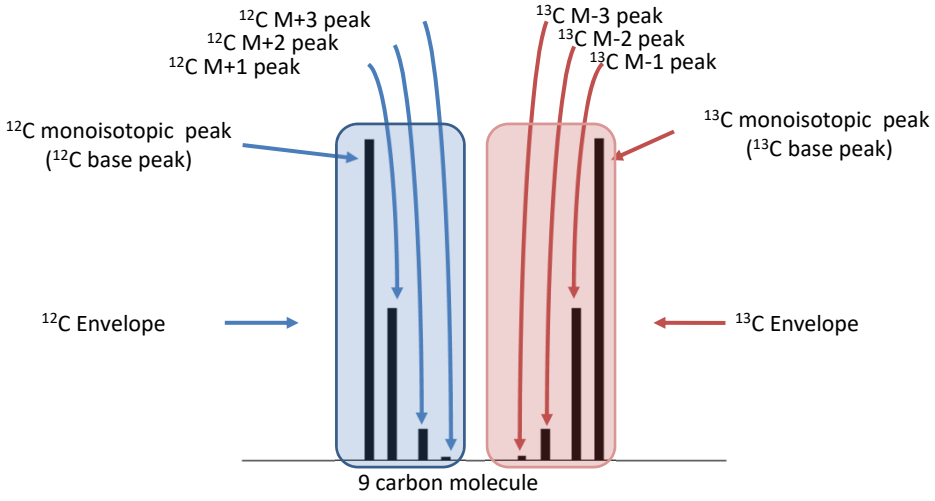

Glutathione shows the sizes of the non-labeled, and singly labeled species, which inherits a five carbon fragment from glutamine.




Oxidized glutathione shows the sizes of the non-labeled, singly labeled, and doubly labeled species. Each inheriting a five carbon fragment from glutamine.

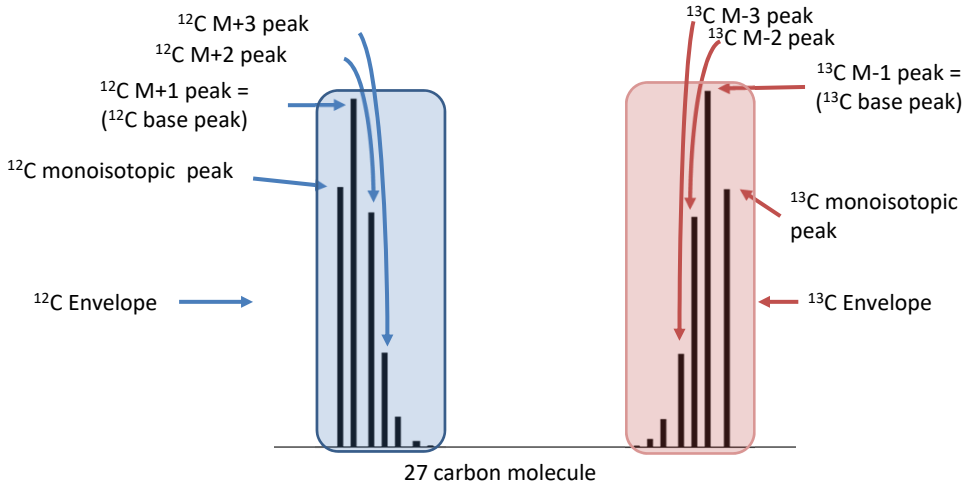
IROA Technologies

# The IROA peak pattern



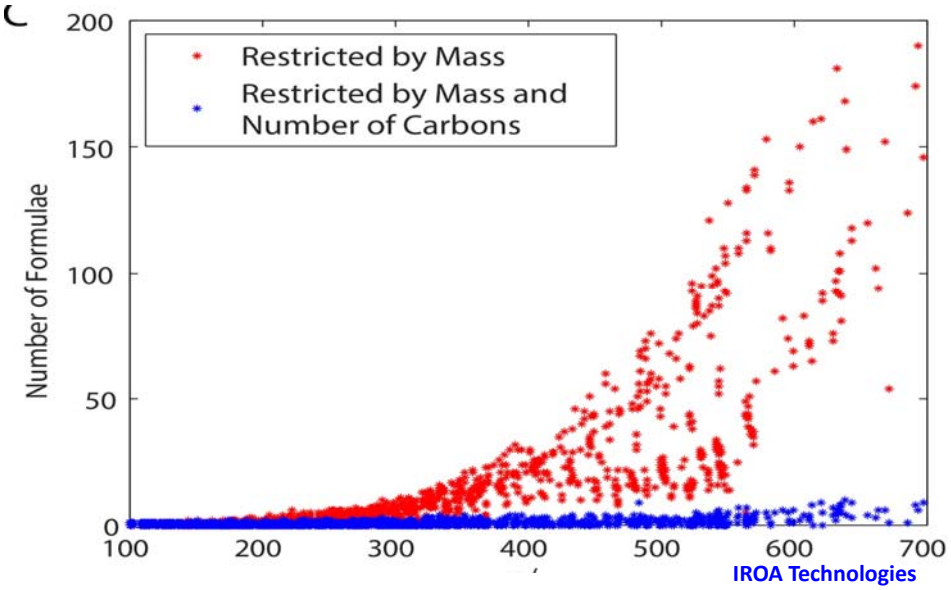

IROA Technologies

# IROA profile of a bigger metabolite



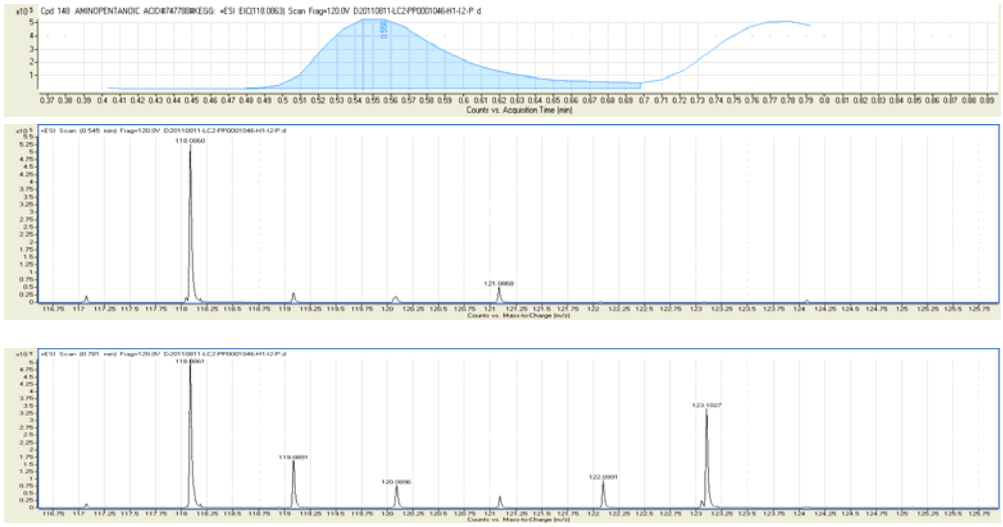

IROA Technologies

# Value of knowing the carbon #



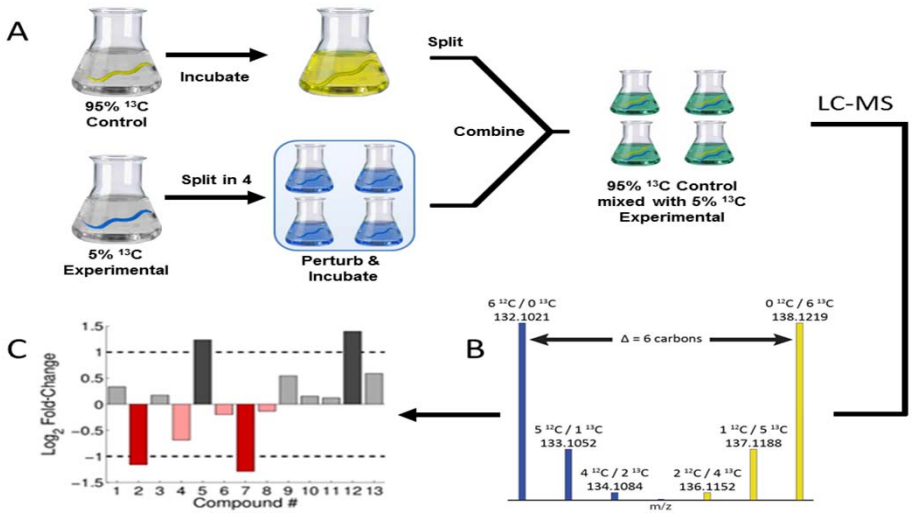

IROA Technologies

# Avoiding metabolite artifacts with IROA



IROA Technologies

# IROA with *C. elegans*

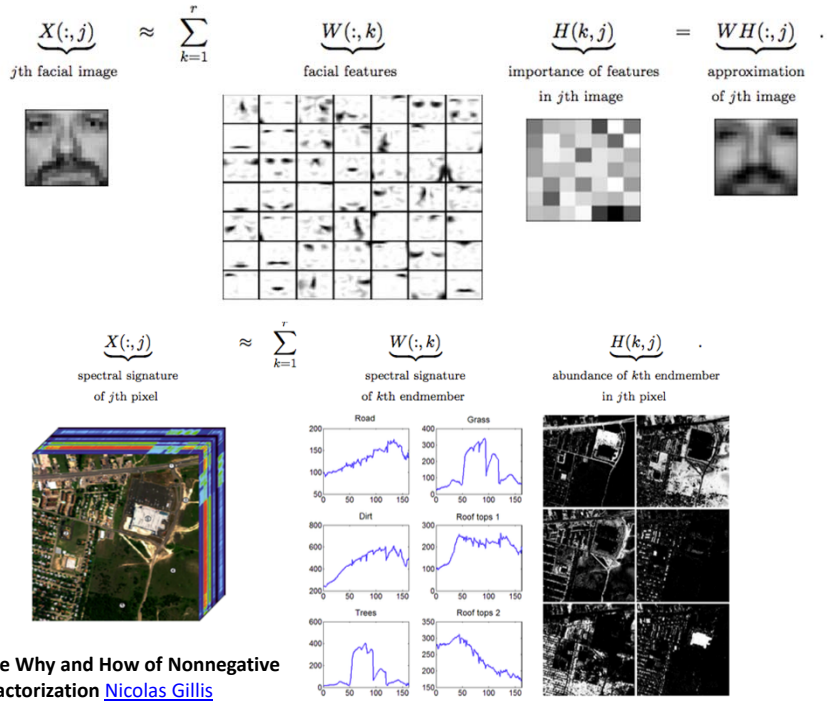



IROA Technologies

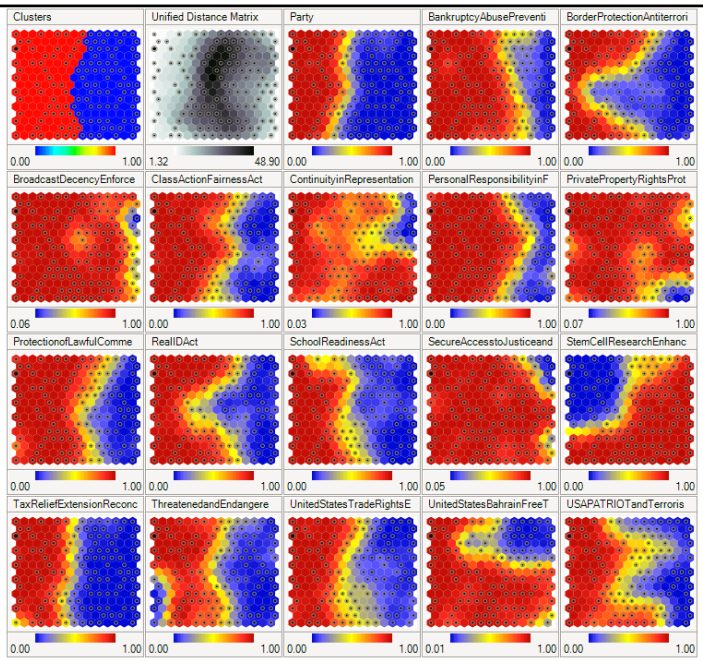
# Effect of a toxin on *C. elegans*

- **742 strong IROA peak pairs were found**
  - **314 named / 428 formula determined**
  - **Positive and negative mode LC**
  - **Thermo Orbi-trap @ 70K resolution**
- **Strong response signature determined**
  - **Basic statistics, PCA, Random Forest, non-negative matrix factorization (NMF), self-organizing map (SOM)**
  - **74 compounds were considered significant by at least 3 of these methods.**

**IROA Technologies**




From: **The Why and How of Nonnegative Matrix Factorization** Nicolas Gillis




https://en.wikipedia.org/wiki/Self-organizing_map#/media/File:Synapse_Self-Organizing_Map.png

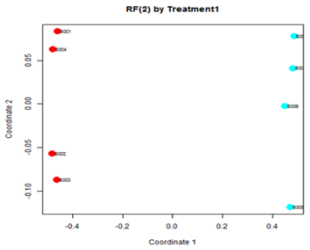
# Ions significantly affected by the toxin

| Name | | p.value | F-Value treatment² |
|---|---|---|---|
| L-LYSINE | P000009 | 7.89E-05 | 89.71 |
| Possibly $C_5H_7N_3O_9S$ | P000018 | 3.06E-05 | 124.99 |
| L-ARGININE | P000019 | 0.000131 | 74.84 |
| Possibly $C_5H_9NO_{11}$ | P000025 | 0.000182 | 66.63 |
| UNK *m/z* 369.2215 RT 0.58 | P000040 | 2.19E-05 | 140.24 |
| SACCHAROPINE | P000046 | 7.23E-05 | 92.51 |
| L-THREONINE | P000051 | 2.64E-05 | 131.52 |
| L-GLUTAMIC ACID | P000053 | 1.09E-06 | 389.79 |
| 4-OXOPROLINE | P000054 | 1.74E-05 | 151.81 |
| Possibly $C_4H_5NO$ | P000058 | 1.8E-05 | 150.26 |
| L-VALINE | P000060 | 0.000262 | 58.37 |
| CITRULLINE | P000061 | 3.15E-05 | 123.67 |
| 4-METHYLENE-L-GLUTAMINE | P000062 | 0.000169 | 68.40 |
| L-METHIONINE S-OXIDE | P000065 | 7.55E-06 | 202.32 |
| L-PROLINAMIDE | P000085 | 0.000227 | 61.56 |
| STACHYDRINE | P000102 | 4.75E-05 | 107.19 |
| UNK *m/z* 206.0368 RT 0.71 | P000114 | 0.000251 | 59.35 |
| N-ACETYLPUTRESCINE | P000122 | 8.96E-07 | 417.06 |
| EPSILON-CAPROLACTAM | P000123 | 1.29E-08 | 1731.72 |
| 2-AMINO-OCTANOIC ACID | P000131 | 0.000213 | 62.99 |
| UNK m/z 345.1258 RT 0.97 | P000141 | 0.000111 | 79.36 |
| Possibly $C_{10}H_{19}N_2O_5P_2$ | P000151 | 0.000154 | 70.78 |
| CYS-GLY | P000152 | 0.000116 | 78.29 |
| URATE | P000156 | 0.000222 | 62.02 |
| Possibly $C_{13}H_{16}N_5OPS$ | P000218 | 1.1E-05 | 177.82 |

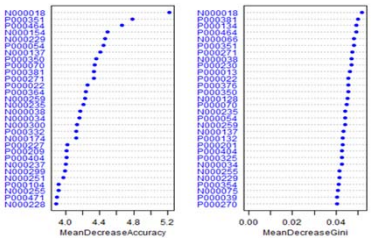
**IROA Technologies**

# Multivariate statistics

A




B



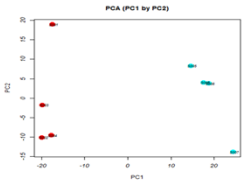

C

| Name | | X1.HP | heatShock | MeanDecreaseAccuracy | MeanDecreaseGini |
|---|---|---|---|---|---|
| Possibly C8H18N6PS2 | N000018 | 5.165779 | 5.091283 | 5.219694 | 0.051833 |
| UNK MZ 197.0968 RT 0.55 | P000351 | 4.61306 | 4.651274 | 4.787018 | 0.047917 |
| Possibly C11H14N3O6 | P000464 | 4.527905 | 4.623332 | 4.660738 | 0.04925 |
| Possibly C11H23N5PS | N000154 | 4.310867 | 4.345237 | 4.49079 | 0.038917 |
| UNK MZ 373.1256 RT 0.60 | N000229 | 4.384942 | 4.313313 | 4.461684 | 0.0415 |
| 4-OXOPROLINE | P000054 | 4.307818 | 4.380554 | 4.447013 | 0.044083 |
| D-ALANYL-D-ALANINE | N000137 | 4.357447 | 4.298196 | 4.403362 | 0.043417 |
| UNK MZ 160.0958 RT 0.55 | P000350 | 4.194825 | 4.331486 | 4.35738 | 0.045333 |
| S-ADENOSYLMETHIONINE | P000070 | 4.247356 | 4.247356 | 4.345614 | 0.04475 |
| UNK MZ 224.1142 RT 0.78 | P000381 | 4.245944 | 4.309512 | 4.333807 | 0.049917 |
| Possibly C6H13N2S2 | P000271 | 4.286066 | 4.238801 | 4.330957 | 0.047167 |
| Possibly C5H10N4O5P | P000022 | 4.237845 | 4.104634 | 4.253873 | 0.045583 |
| Possibly C14H29N11O3S | P000364 | 4.207787 | 4.123607 | 4.236865 | 0.036583 |
| Possibly C18H18N11S3 | N000259 | 4.115656 | 4.119522 | 4.230318 | 0.044 |
| N-FORMIMINO-L-GLUTAMATE | N000235 | 4.115656 | 4.091016 | 4.208769 | 0.044167 |

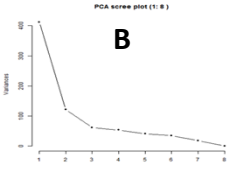
IROA Technologies

# PCA analysis of toxin's effect

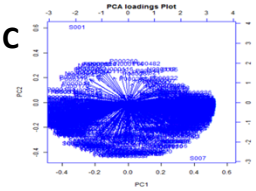
A




B




C




D

| | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| P000123 | -0.04867 | -0.00122 | -0.0028 | 0.002634 | -0.00263 |
| P000305 | -0.04864 | -0.01333 | 0.003947 | -0.00022 | 0.000461 |
| P000328 | -0.04854 | -0.01474 | -0.00285 | 0.001186 | 0.002235 |
| P000065 | -0.04839 | -0.01076 | -0.00949 | 0.016892 | 0.002798 |
| N000034 | -0.04838 | -0.00625 | 0.004756 | -0.00207 | -0.00469 |
| P000268 | -0.04834 | -0.01365 | 0.013402 | -0.00273 | -0.00146 |
| P000122 | -0.04832 | -0.00499 | 0.000535 | 0.0041 | 0.008176 |
| P000485 | -0.04814 | -0.01292 | 0.013305 | -0.00659 | 0.01036 |
| N000006 | -0.04792 | -0.00964 | -0.02169 | 0.005208 | 0.003092 |
| N000068 | -0.04775 | -0.00849 | 0.026794 | -0.00377 | -0.00419 |
| P000304 | -0.04768 | -0.02207 | 0.008128 | -0.00055 | 0.004545 |
| N000057 | -0.04759 | -0.00101 | 0.027516 | -0.00376 | -0.00763 |
| N000174 | -0.04753 | 0.005826 | 0.00295 | 0.003997 | -0.0041 |
| P000152 | -0.04751 | -0.01938 | 0.01513 | -0.00054 | 0.003671 |
| P000141 | -0.04749 | -0.00612 | 0.024997 | -0.00648 | -0.00814 |

IROA Technologies

# Summary of most likely metabolites

| Name | | Stats1 | RFTop1 | RFTop2 | NMF3 | NMF4 | NMF5 | NMF6 | Count |
|---|---|---|---|---|---|---|---|---|---|
| UNK *m/z* 160.0958 RT 0.55 | P000350 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| UNK *m/z* 197.0968 RT 0.55 | P000351 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| UNK *m/z* 216.0852 RT 0.61 | N000034 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| D-ALANYL-D-ALANINE | N000137 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| Possibly $C_{25}H_{34}N_4O_5$ | N000174 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| UNK *m/z* 373.1256 RT 0.60 | N000229 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 2-AMINO-OCTANOIC ACID | P000131 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_6H_8N_4O_3$ | P000354 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| UNK *m/z* 510.2122 RT 0.68 | P000373 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| UNK *m/z* 224.1142 RT 0.78 | P000381 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_6H_9NO_5P$ | P000410 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_{11}H_{14}N_3O_6$ | P000464 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_6H_4N_2O_6P$ | P000471 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_5H_{12}N_2O_7PS$ | N000006 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| Possibly $C_{11}H_{23}N_5PS$ | N000154 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 6 |
| D-GLUCOSE | N000228 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| UNK *m/z* 548.2037 RT 0.63 | N000232 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |
| GLYCERATE | N000237 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 6 |

**IROA Technologies**

# From Chris Beecher, IROA Technologies

https://www.uab.edu/proteomics/metabolomics/workshop/2017/videos/beecher_day3.html

IROA TECHNOLOGIES
...making metabolic profiling easy®

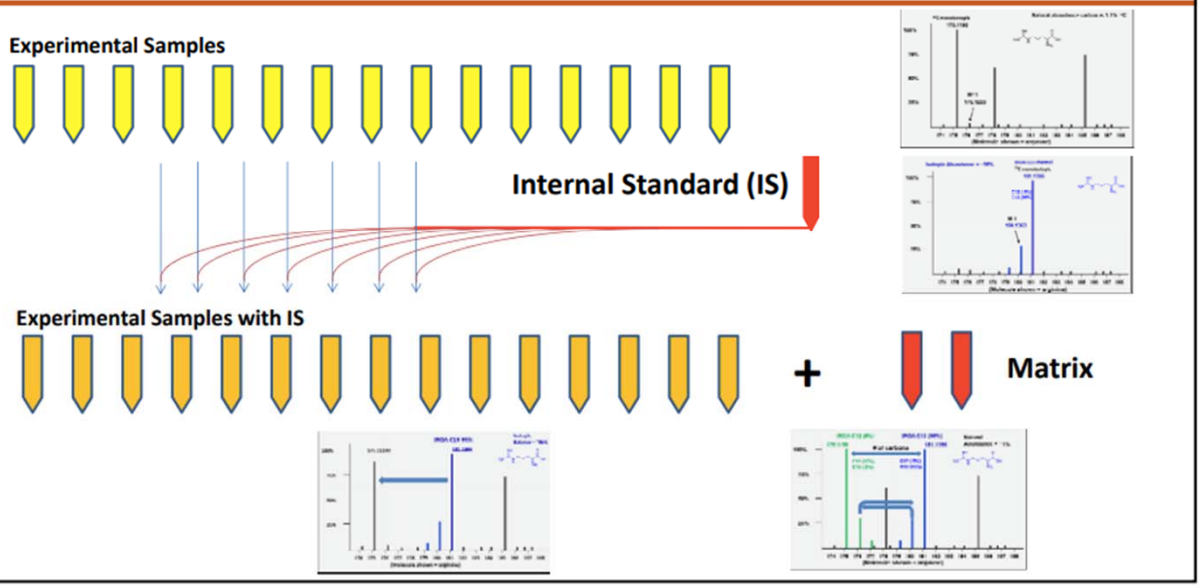
## The IROA Workflow

- The IROA Workflow is an extension of the Phenotypic Protocol in which:
  - a defined IROA-based Internal Standard (IS) is used in any type of experimental or clinical sample, and
  - an equally defined QA/QC sample (Matrix) is analyzed daily.
- Together they make a systematic measurement system that is completely reproducible across sample types, instrument types, and overcomes time-induced variance.

The IROA Workflow is the basis of "Clinical Metabolomics".

IROA TECHNOLOGIES
...making metabolic profiling easy®

## IROA-based Workflow



IROA TECHNOLOGIES
...making metabolic profiling easy®

## What is the IROA-IS

- The IS is a 95% U-$^{13}$C-labeled complex Internal Standard
- The IS has a standard concentration of 1000+ **identified** and curated compounds for **co-location in an experimental sample.**
- The IS has enough compounds to provide for a Retention Time (RT) ladder that allows **alignment** of all peaks in the chromatogram.
- The IS may be used to **normalize** the samples against one another.
- The IS allows **day-to-day, or even instrument-to-instrument variances to be eliminated.**

IROA TECHNOLOGIES
...making metabolic profiling easy®

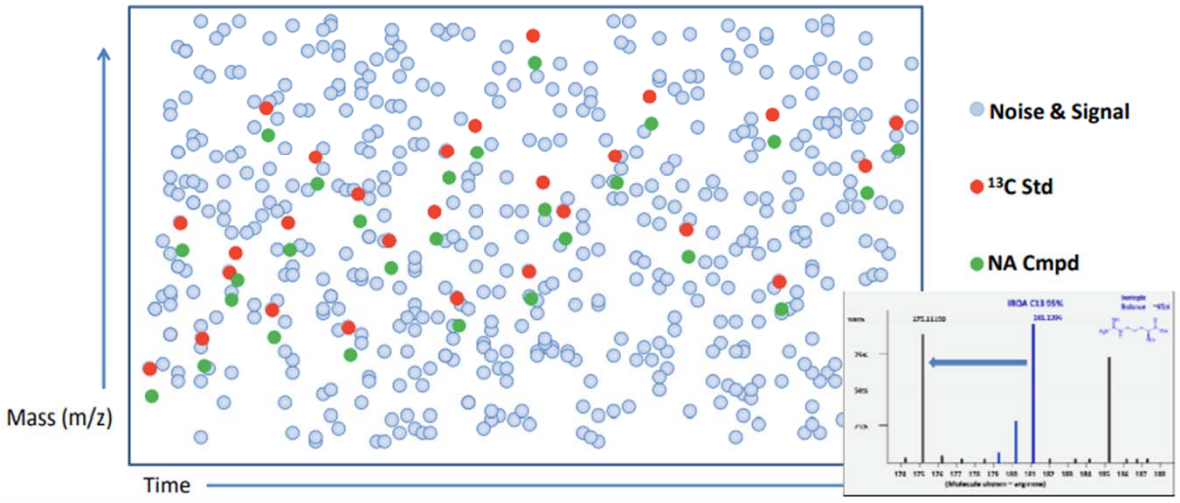
## What is Matrix?

- The Matrix sample is a made from the same material as the IS but rather than having a natural abundance partner it is paired with a perfectly matched IROA 5% U-$^{13}$C sample.
  - The almost perfect balance of the 5% and 95% chemical composition,
  - The completely defined nature of the Matrix sample, and
  - The absolute reproducibility of the Matrix sample.
- Provide a way to compare day-to-day analytical performance on all parts of the analytical process, and
- Provide a daily mapping of all compounds found in the IS so that their complete identification is always assured.

IROA TECHN
...making metabolic p

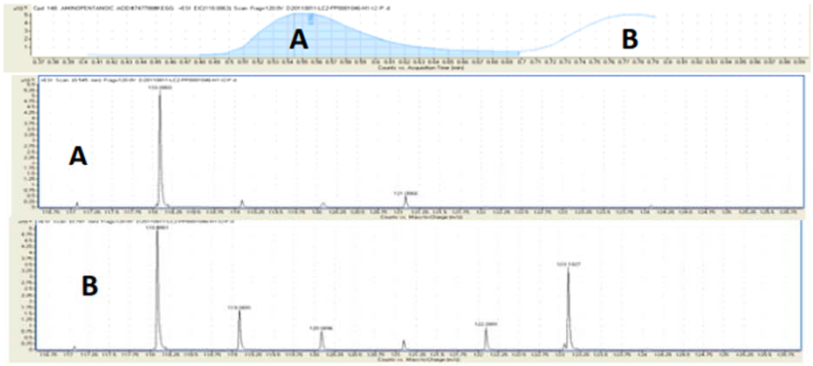
# The Experimental Sample Process

## Identify, Align, Co-locate, Measure & Normalize




IROA TECHNOLOGIES
...making metabolic profiling easy®

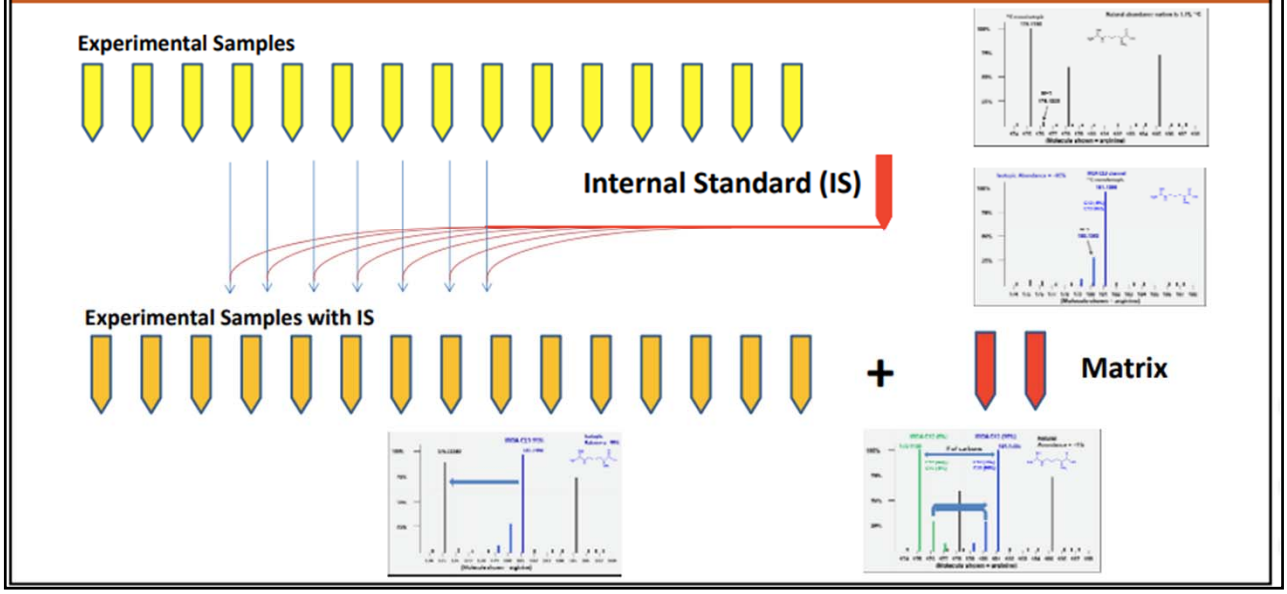
# Artifact-based errors are avoided




IROA identifies **A** as an artifact and **B** as a peak of biological origin, preventing an artifact from creating an error despite a "better time signature" on **A**.

IROA TECHNOLOGIES
...making metabolic profiling easy®

## IROA-based Workflow




IROA TECHNOLOGIES
...making metabolic profiling easy®

## Summary of IROA-based Workflow

1) Cost-effective simultaneous quantitative measurement of several hundred biochemicals in a single analytical run through the use of IROA Internal Standards (IS) and Matrix (M);
   i. IS & M provide high level QC for accurate and reproducible results;
   ii. IS enables removal of false data (all noise and artifacts);
   iii. IS enables precise quantitation through complex software algorithms;
   iv. IS allows for normalization of samples to overcome sample-to-sample variation;
   v. Once normalized, IS provides a map that can be used for compounds that are not in the IS.
2) All of the above are completely software automated and may be done with minimal (or even no) human interaction.

IROA TECHNOLOGIES
...making metabolic profiling easy®

# The 1500 peaks in Matrix (pos mode)

While finding them is completely automated, we are currently examining each one, and annotating it

- We have built a database to collect all of this information

- We are using this to directly tackle the problem of the percent of peaks that are "knowns", and what percent are "unknowns", i.e. not just fragments, adducts, etc. of "knowns"

*We will be successful because the Matrix is a pure IROA mixture, i.e. we can discriminate between real compounds and artifacts