EXPANDING THE SCOPE OF TILE-BASED GC×GC–TOFMS DATA ANALYSIS

Robert E. Synovec, Caitlin N. Cain, Lina Mikaliunaite

Department of Chemistry, Box 351700
University of Washington
Seattle, WA 98195, USA

15th Multidimensional Chromatography Workshop
Friday, January 12, 2024

OUTLINE

- The Goals and Challenges for GC×GC-TOFMS Data Set Discovery-Based Analyses
- Revisiting a Cycling Yeast Metabolite Data Set with Tile-based Fisher-ratio Analysis
- Discovery of Biomarkers for Knee Inflammation
- Impact of Pacu Fish Environmental Conditions on their Metabolome
- Comparative Analysis with One Sample per Class: Tile-based 1v1 Analysis
- Tile-based Variance Ranking for Unsupervised Analyses: PCA and PLS of Fuels

Goals for the Analysis of GC×GC-TOFMS Data Sets

GOALS:

- Translate Data into Actionable Information. Discover *Important Differences* in Chemical Composition in Complex Samples, using either Supervised or Unsupervised Experimental Designs.
- Identify & Quantify the Discovered Compounds, often with *Deconvolution / Decomposition*.
- Translate into *Targeted Analysis Methods* for routine testing.

APPLICATION AREAS

- Fuels
- Food Quality, Safety and Security
- Metabolomics
- Industrial Discovery and Quality Control (eg., Biotech)
- Health Biomarker Discovery…. Clinical Applications
- Feed Stock Evaluation (eg., Biomass to Fuel)
- Forensics, Impurity Profiling
- Environmental Analysis

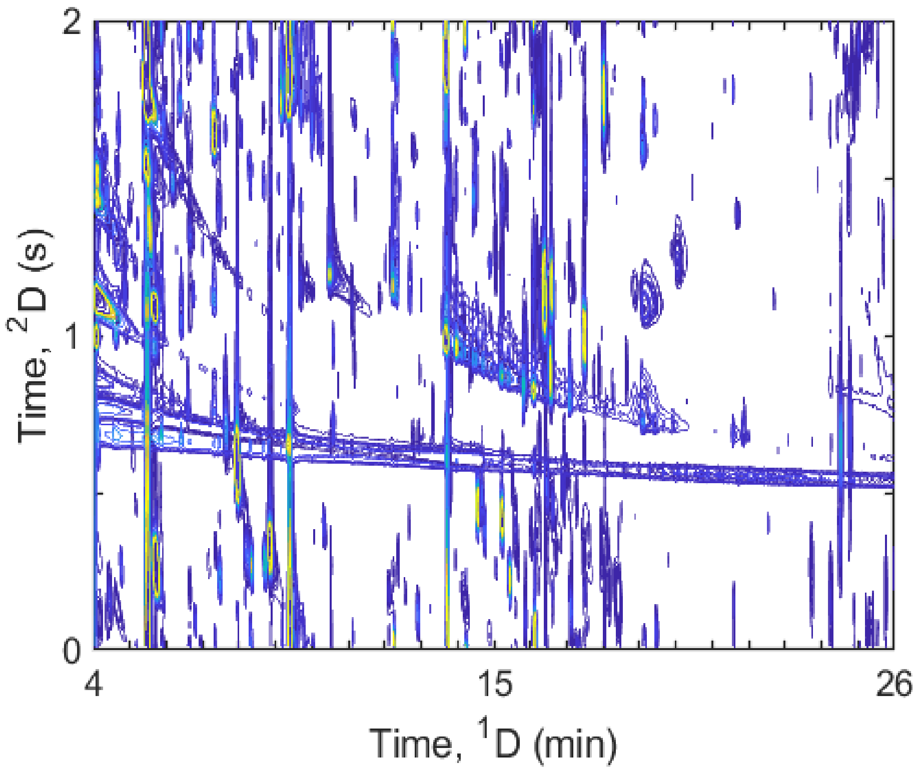
"Recent Advances in GC×GC and Chemometrics to Address Emerging Challenges in Nontargeted Analysis,"
T. J. Trinklein[#], C.N. Cain[#], G.S. Ochoa, S. Schöneich, L. Mikaliunaite, R.E. Synovec, *Anal. Chem.*, 2023, 95, 264−286.

The Analytical Challenge for Comparative Analysis of GC×GC-TOFMS Data Sets

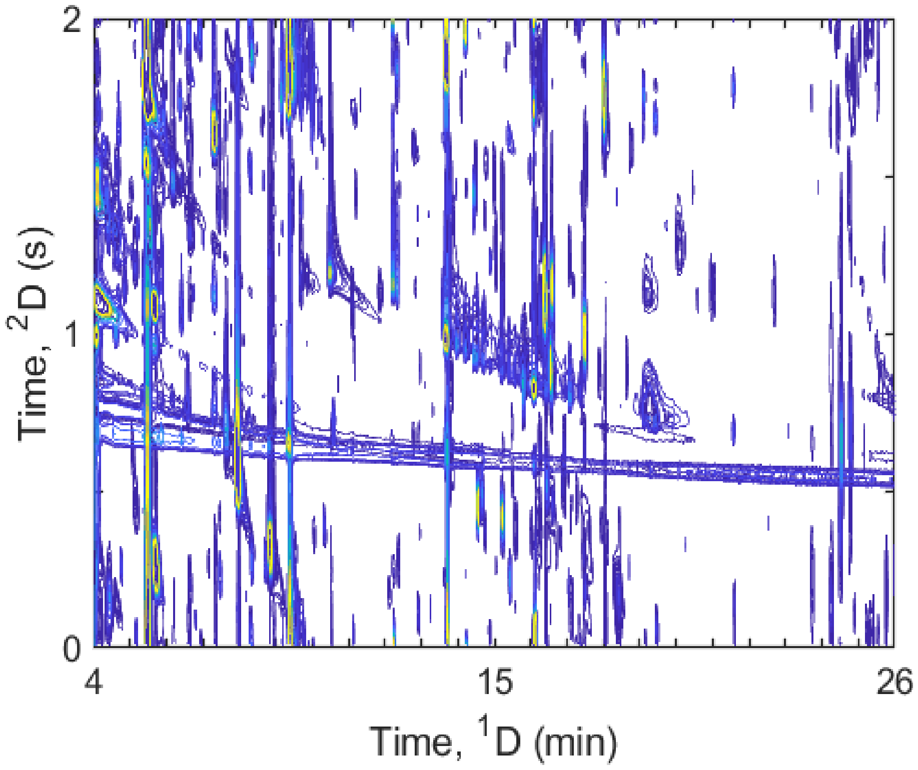
How do we confidently discover the chemically relevant differences between the two (or more) sample classes?

For GC×GC-TOFMS, the data analysis software must overcome run-to-run retention time shifting, while leveraging the 2D separation and mass spectrum data dimensions to pinpoint key analytes using an informative metric.

GC×GC Chromatogram
Class A Sample

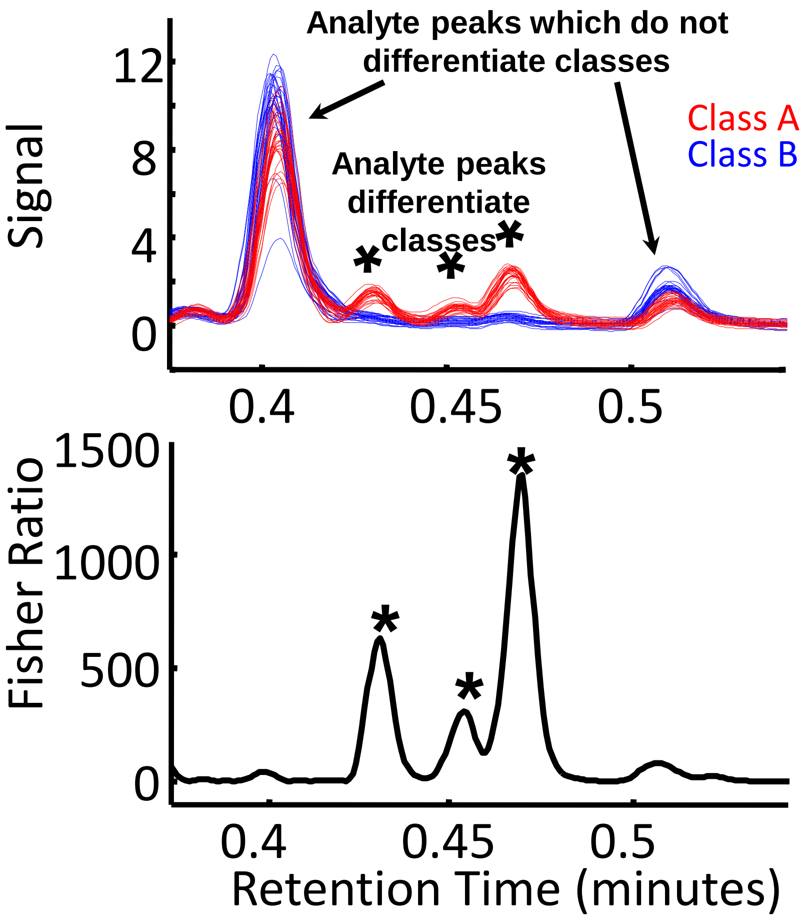


GC×GC Chromatogram
Class B Sample



"These samples look the same to me!"

Evolution of Tile-based Fisher Ratio (F-ratio) Analysis for Supervised Comparisons



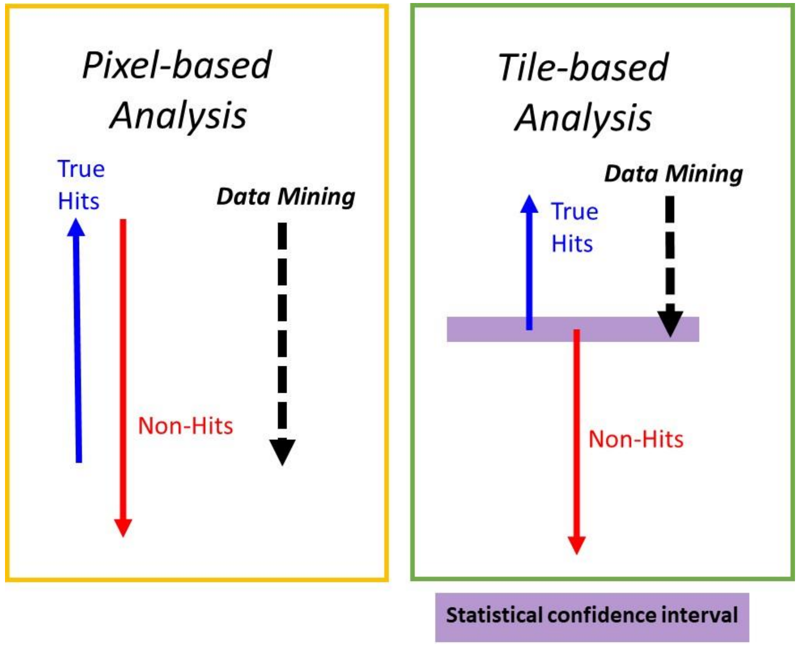
F-ratio Analysis provides a ranked hitlist of analytes that are likely to be statistically different in concentration (p-value < 0.05) between sample classes.

$$Standard\ \ F-ratio = \frac{Between\ Class\ Variance}{\sum(Within\ Class\ Variance)}$$

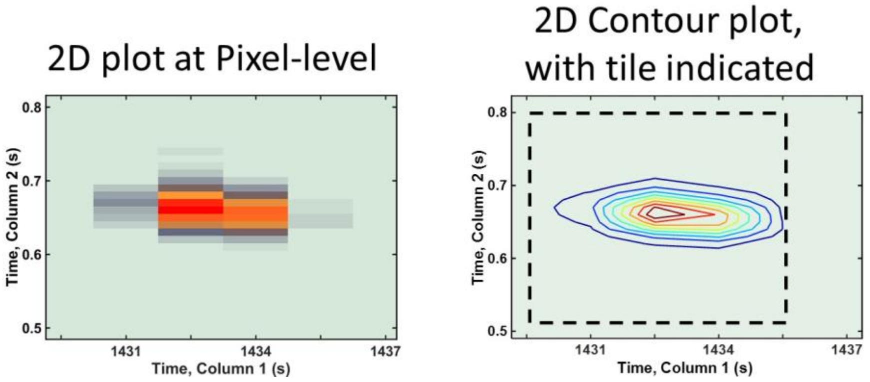
- **K.J. Johnson, R.E. Synovec, J. Chemom. Intell. Lab. Syst., 2002, 60, 225-237.**
- **L.C. Marney, W.C. Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Talanta 2013, 115, 887–895.**
- **B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Anal. Chem. 2015, 87, 3812–3819.**
- **N.E. Watson, B.A. Parsons, R.E. Synovec, J. Chromatogr. A, 2016, 1459, 101-111.**
- **B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, J. Chromatogr. A, 2016, 1440, 179-190.**
- **B.C. Reaser, B.W. Wright, R.E. Synovec, Anal. Chem., 2017, 89, 3606-3612.**

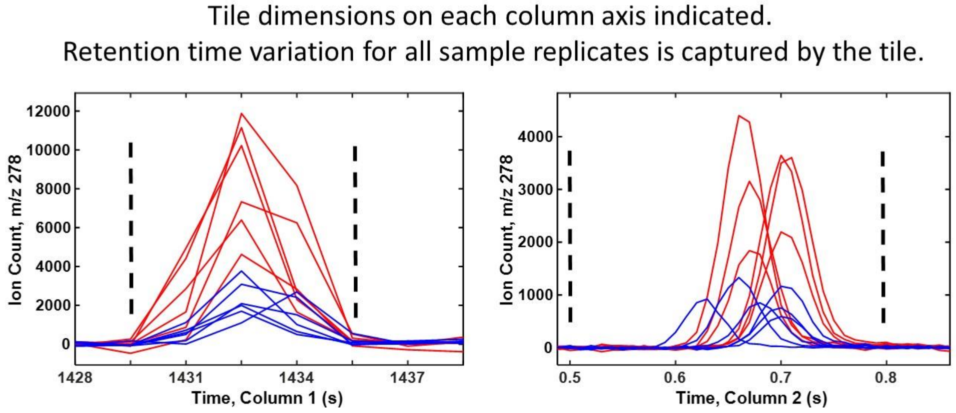
Hit List

Hit #	Fisher Ratio	Analyte
1	High	AAA
N	Low	ZZZ

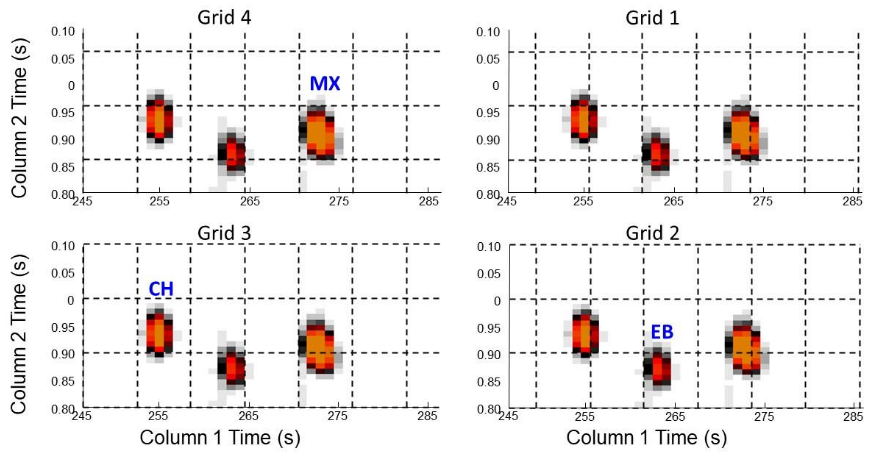


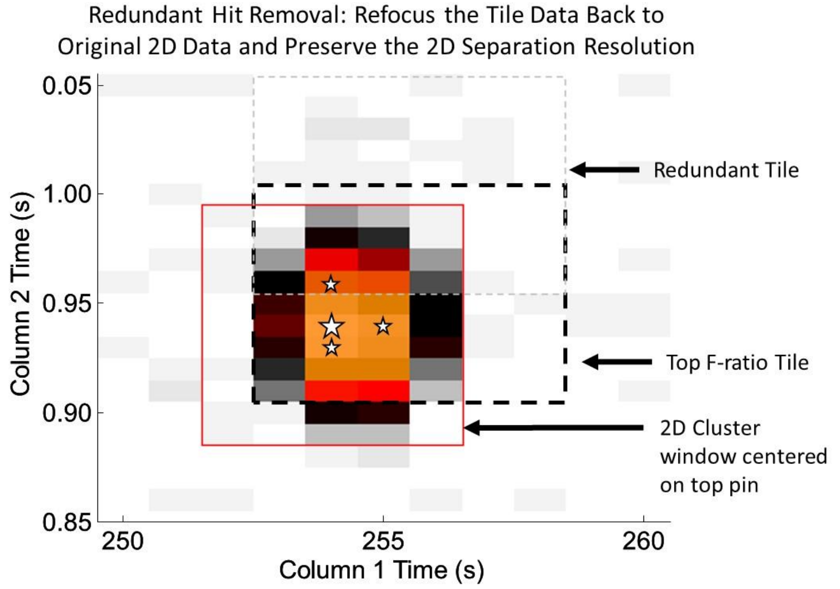
Tile-based Fisher Ratio (F-ratio) Analysis:
Minimization of Run-to-run Retention Time Shifting Impact





Four Tile Grids to Capture All 2D Peaks: Smart Binning





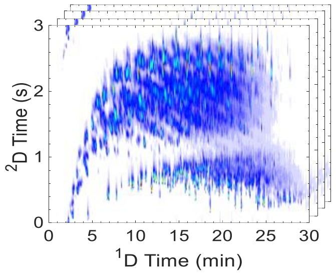
Multiple Hits reduced to one Hit per analyte with original 2D resolution !

Tile-based Fisher Ratio Analysis

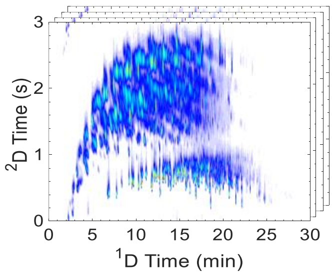
ChromaTOF Tile

Supervised, discovery-based analysis……….

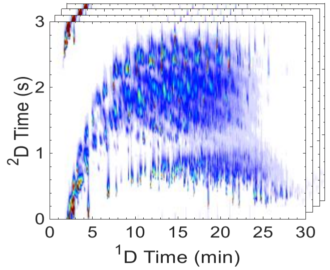
Sample Class 1



Sample Class 2

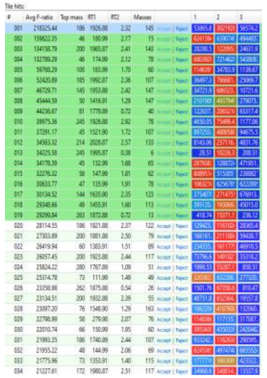


Sample Class 3

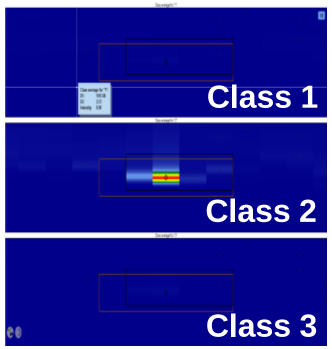


Apply ChromaTOF Tile

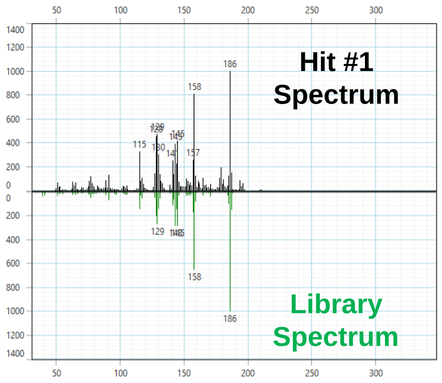
Hit Table



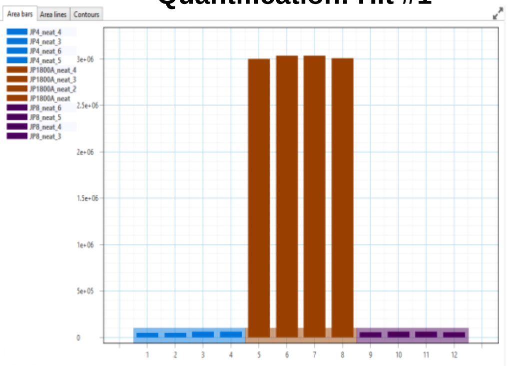
Hit #1 Selective Signal



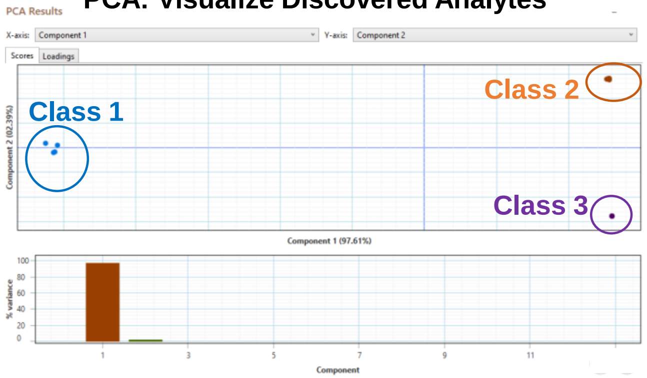
Identification: Hit #1



Quantification: Hit #1



PCA: Visualize Discovered Analytes



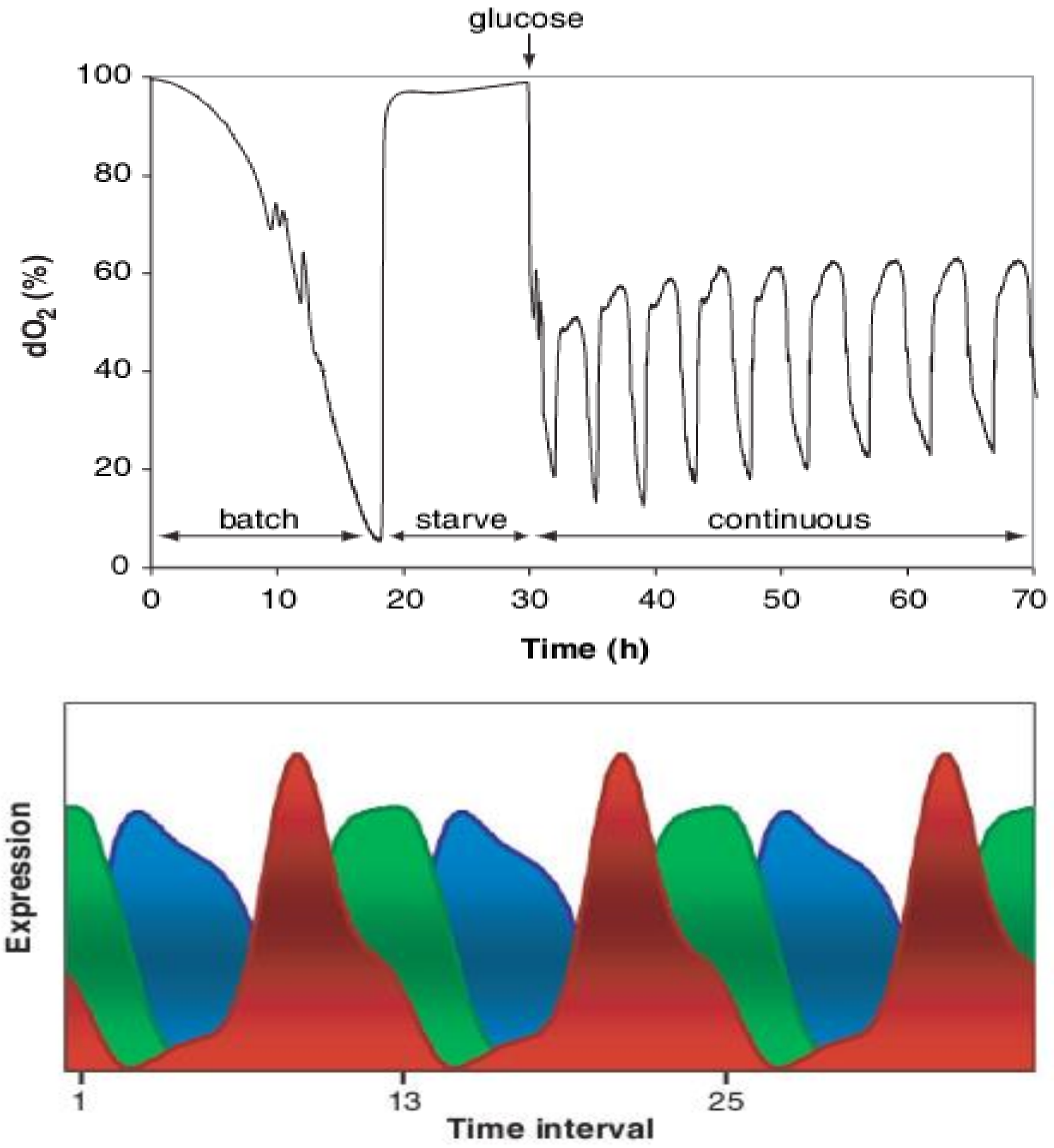
Yeast Metabolic Cycle (YMC)

In 2005, Ben Tu, Steve McKnight and coworkers reported an ultradian cycle in yeast cells where the molecular oxygen (dO_2) exhibited robust oscillation with a period of 5 hours. ~57% of yeast genes were shown to experience periodic behavior.*
**(Saccharomyces cerevisiae - prototrophic yeast strain CEN.PK)*

Our Research: Study the effects of ultradian cycle on the yeast metabolome.

- *How does the metabolome relate to the genome?*
- *Provide insight into genomic, proteomic and metabolomic pathways*

*Collaboration with the Steve McKnight group at the U. Texas Southwestern Medical Center and Ted Young group, U. Washington



B. P. Tu, et. al. *Science* 310 (2005) 1152.

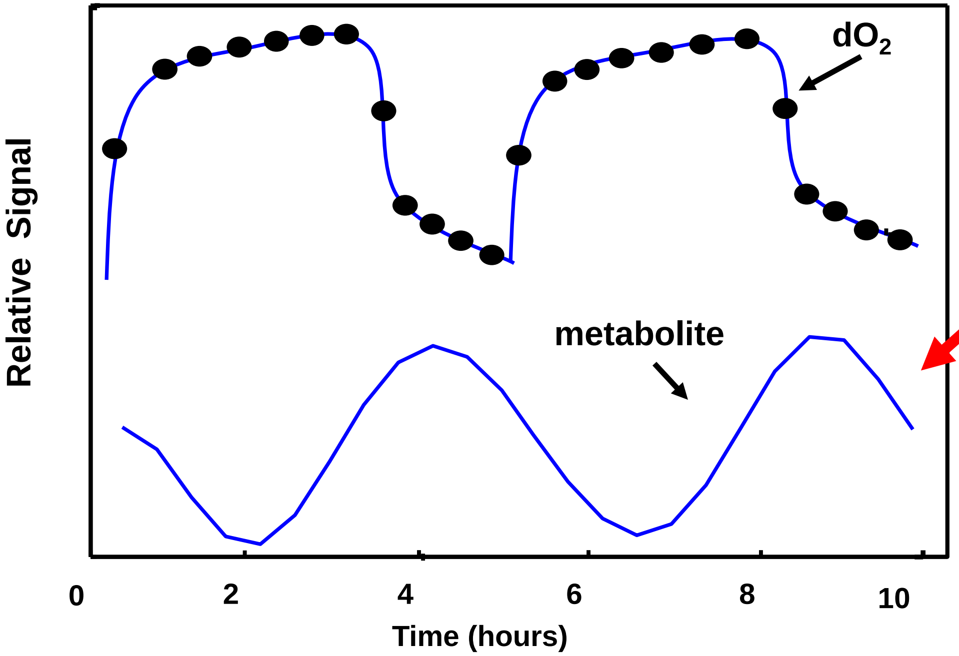
Yeast Metabolic Cycle (YMC) and GC×GC-TOFMS

Production and Consumption of Molecular Oxygen for Saccharomyces cerevisiae - prototrophic yeast strain CEN.PK

Samples collected every

25 min over 10 hr:

24 Samples ➔ 24 Classes!



Can cycling metabolite patterns be readily observed?

What do they look like and how do they relate to dO_2?

85 Metabolites Discovered using "prior" software (non-tiling, based on determining the max Signal ratio)

- B. Tu, R. E. Mohler, J. C. Liu, K. M. Dombek, E. T. Young, R. E. Synovec, S. L. McNight, Proc. Nat. Acad. Sci., 2007, 104, 16886 – 16891.
- R. E. Mohler, B. P. Tu, K. M. Dombek, J. C. Hoggard, E. T. Young, R. E. Synovec, J. Chromatogr. A, 2008, 1186, 401–411.

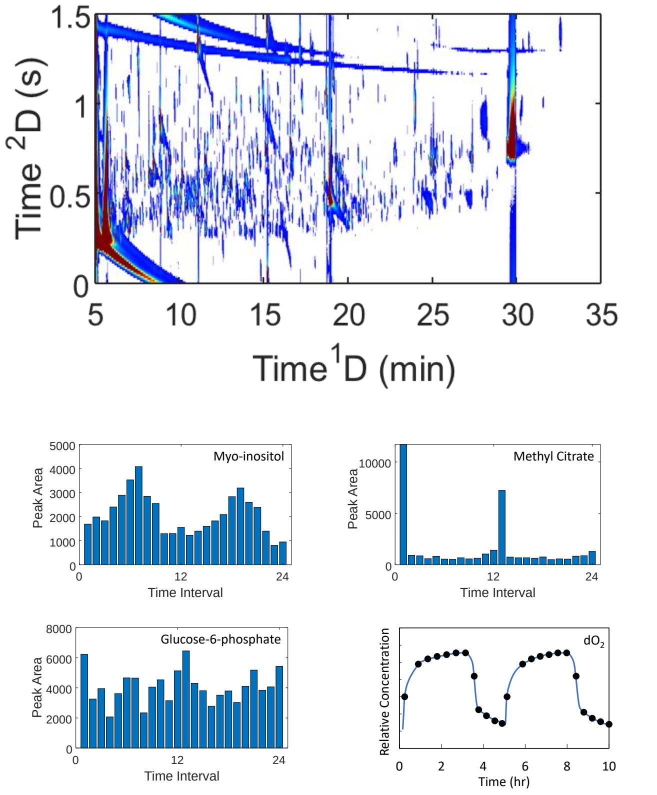
Recent Study using ChromaTOF Tile version of Tile-based F-ratio software: 210 Metabolites Discovered!
This study serves as validation that ChromaTOF Tile has the ability to handle many Sample Classes

- L. Mikaliunaite, R. E. Synovec, Talanta, 2022, 244, 123396.

Computational method for untargeted determination of cycling yeast metabolites using GC×GC-TOFMS and ChromaTOF Tile

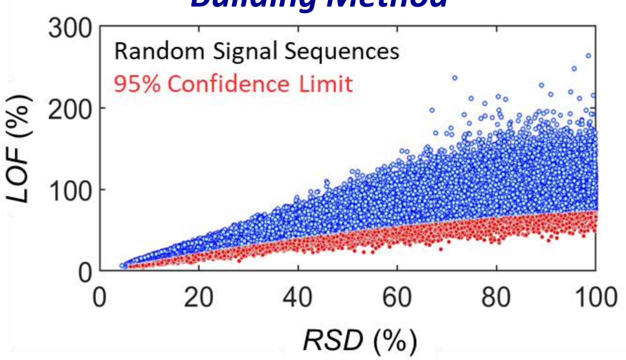
Lina Mikaliunaite

GC × GC – TOFMS

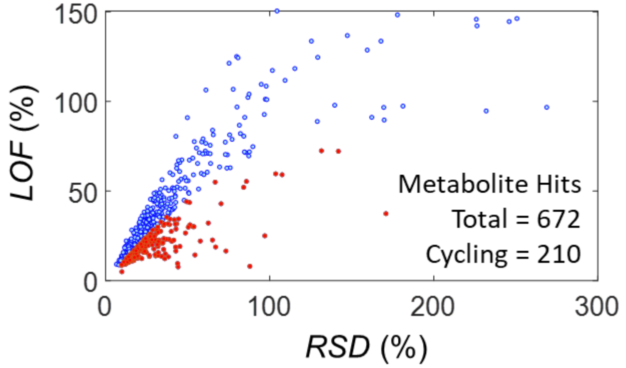


Computational Method

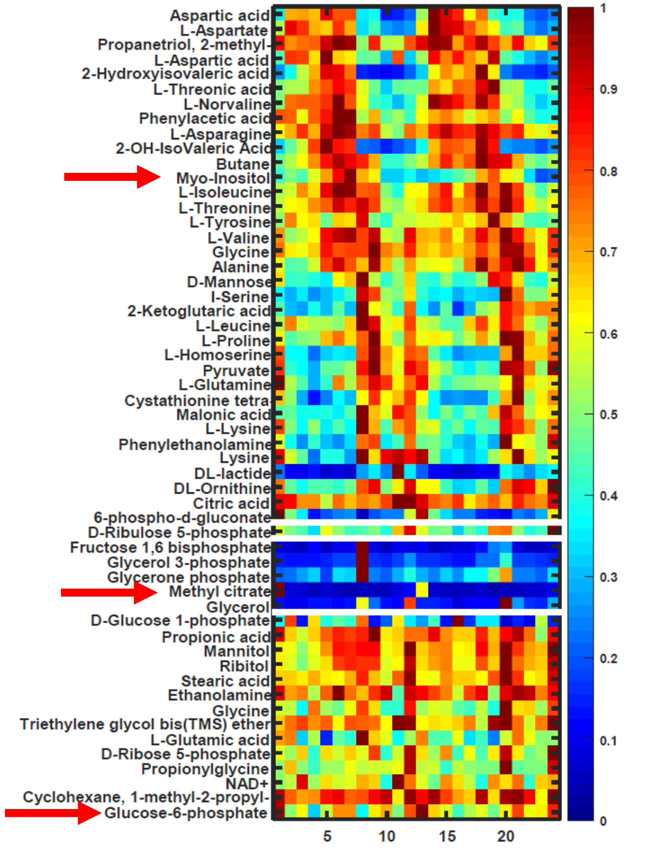
Building Method



Applying Method



55 Identified Cycling Metabolites



• L. Mikaliunaite, R. E. Synovec, Talanta, 2022, 244, 123396.

Beyond Standard Tile-based F-Ratio Analysis
What other Metrics can be Implemented to Rank a Hit List?

- **Number of Components > to >> Number of Peaks (TIC)**

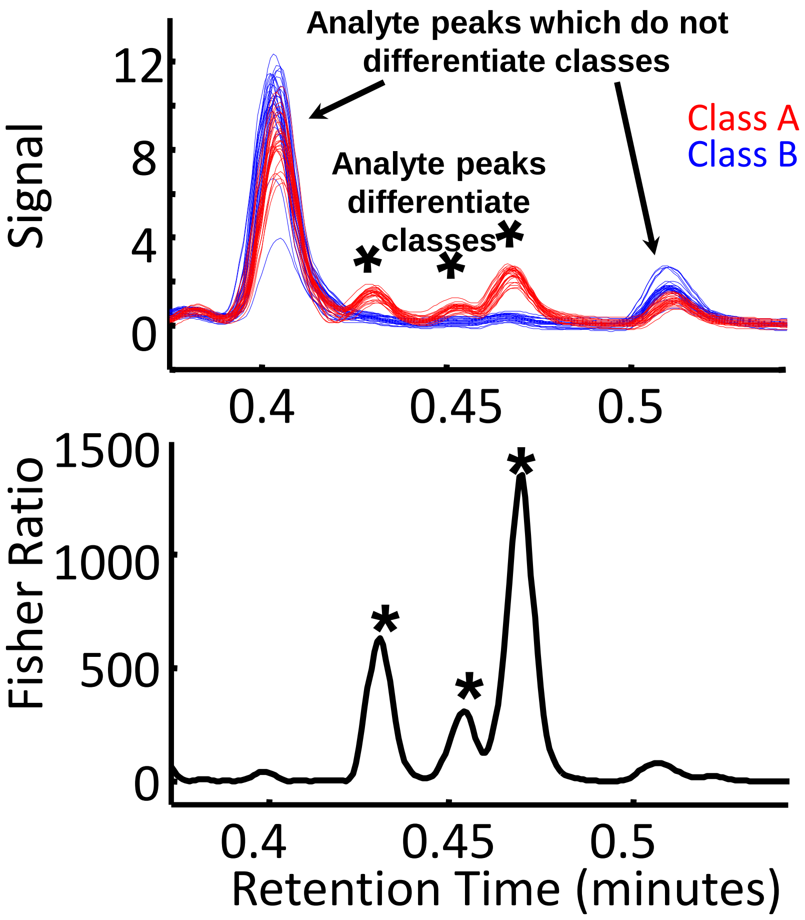
- **The Metric (e.g. F-ratio) Simply Ranks the Components in a Hit List**

- **The Tiling + Hit List Generation Essentially Finds all of the Components, and the Analyst Utilizes the Components Toward the Top of the Hit List**

- **Depending upon Expt Design, what other Metrics can be Implemented to Rank the Hit List?**

- **Fortunately, the Tiling Steps are Independent of the Metric Implemented**

Evolution of Fisher Ratio (F-ratio) Analysis for Supervised Comparisons



F-ratio Analysis provides a ranked hitlist of analytes that are likely to be statistically different in concentration (p-value < 0.05) between sample classes.

$$Standard\ \ F - ratio = \frac{Between\ Class\ Variance}{\sum(Within\ Class\ Variance)}$$

- **K.J. Johnson, R.E. Synovec, J. Chemom. Intell. Lab. Syst., 2002, 60, 225-237.**
- **L.C. Marney, W.C. Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Talanta 2013, 115, 887–895.**
- **B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Anal. Chem. 2015, 87, 3812–3819.**
- **N.E. Watson, B.A. Parsons, R.E. Synovec, J. Chromatogr. A, 2016, 1459, 101-111.**
- **B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, J. Chromatogr. A, 2016, 1440, 179-190.**
- **B.C. Reaser, B.W. Wright, R.E. Synovec, Anal. Chem., 2017, 89, 3606-3612.**

$$Control\text{-}Normalized\ F - ratio = \frac{Between\ Class\ Variance}{Controls\ within\ class\ variance}$$

- **S.E. Prebihalo, G.S. Ochoa, K.L. Berrier, K.J. Skogerboe, K.L. Cameron, J.R. Trump, S.J. Svoboda, J.K. Wickiser, R.E. Synovec, Anal. Chem. 2020, 92, 15526–15533.**

analytical chemistry

pubs.acs.org/ac Article

Control-Normalized Fisher Ratio Analysis of Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry Data for Enhanced Biomarker Discovery in a Metabolomic Study of Orthopedic Knee-Ligament Injury

Sarah E. Prebihalo, Grant S. Ochoa, Kelsey L. Berrier, Kristen J. Skogerboe, Kenneth L. Cameron, Jesse R. Trump, Steven J. Svoboda, J. Kenneth Wickiser, and Robert E. Synovec*

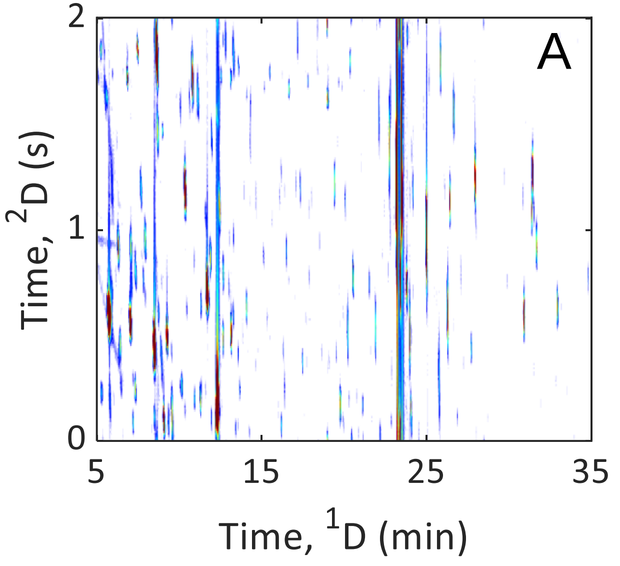
Cite This: *Anal. Chem.* 2020, 92, 15526−15533 Read Online

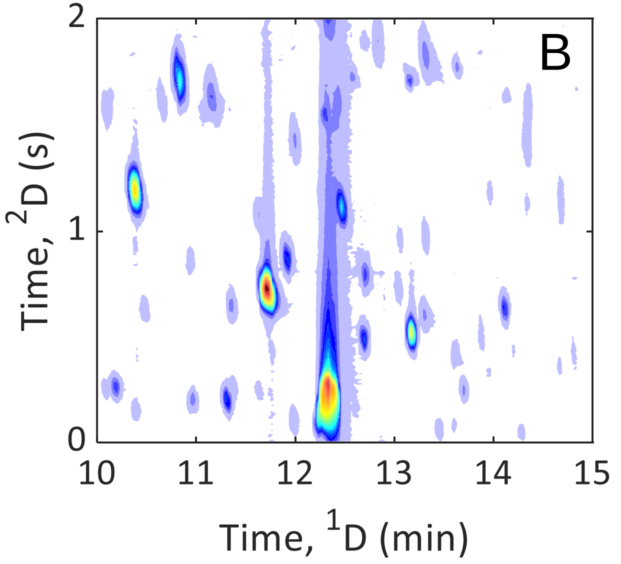
Sarah Prebihalo

- Study ACL Injury and post-traumatic osteoarthritis development
- Use GC×GC-TOFMS and Chemometrics to find differences between 30 injured patients and 30 non-injured controls
- Collaborators: US Military Academy (West Point) and Keller Army Community Hospital

Are there biomarkers associated with the ACL injury that can improve therapeutic intervention and/or early diagnosis?

Representative Chromatogram Time-of-Injury Patient



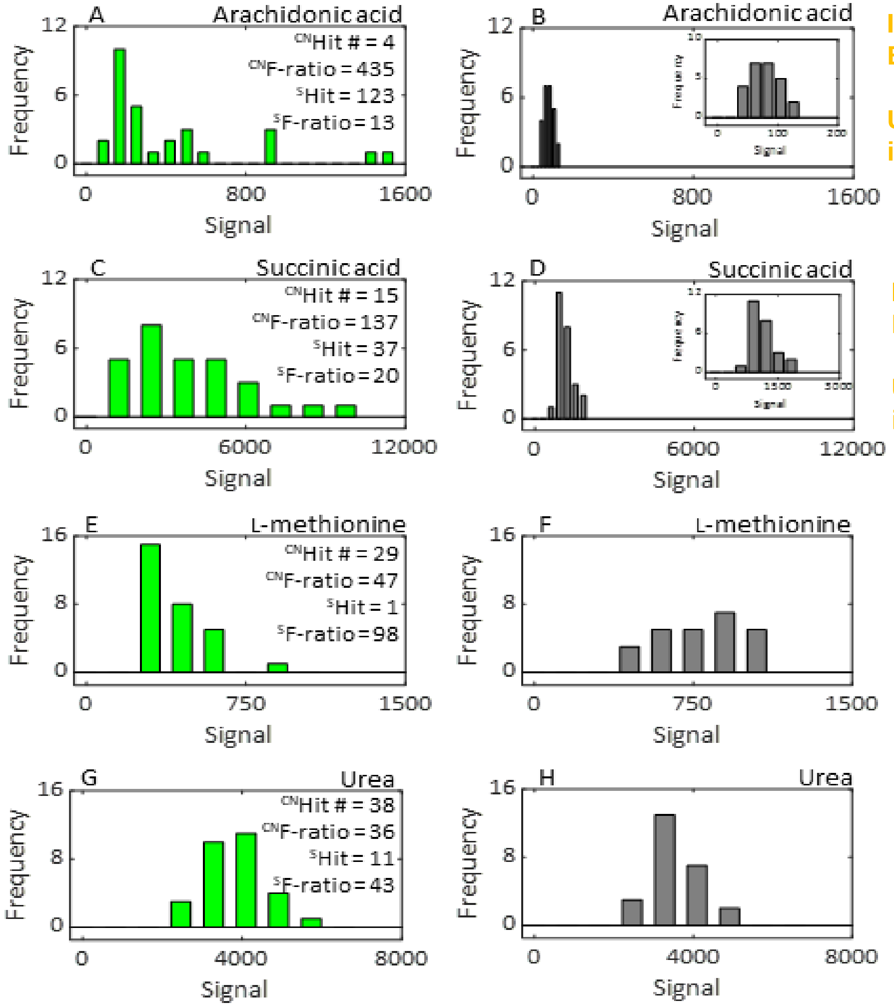


F-ratio Analysis Summary 30 patients vs. 30 controls

ID	^{CN}Hit #	^{CN}F-ratio	^{S}Hit #	^{S}F-ratio	[P]/[C]	Var #
Naproxen	1	6075	254	7	22.8	1
Phenylalanine*	3	473	12	42	2.44	2
Arachidonic acid	4	435	123	13	5.51	3
Palmitic acid*	5	403	20	32	3.16	4
Mannose	6	299	2	74	3.01	5
Stearic acid*	7	295	19	33	3.73	6
Glycine	8	285	5	51	1.28	7
Linoleic acid	9	280	16	34	4.38	8
Metabolite Unk 6	10	247	147	11	2.10	9
Glutamine	11	222	24	30	2.36	10
Glutamic acid*	12	203	18	33	2.87	11
L-lysine*	13	165	10	44	4.46	12
Metabolite Unk 1	14	162	4	54	3.44	13
Succinic acid*	15	137	37	20	3.31	14
CoA fragment	16	133	15	34	3.22	15
Metabolite Unk 9	21	63	109	13	3.77	16
Lactic acid*	23	60	21	32	2.42	17
Phospho-D-glu-conate	24	56	31	23	1.37	18
Metabolite Unk 10	25	53	36	21	1.56	19
Malonic acid*	26	52	3	59	0.65	20
L-histidine	27	49	30	23	1.96	21
Metabolite Unk 11	28	48	130	12	0.73	22
L-methionine*	29	47	1	98	0.46	23
L-threonine	30	44	28	25	2.50	24
Glycerol*	36	40	9	45	0.60	25
Glycopyranose	66	23	13	37	0.62	26
Metabolite Unk 2	85	21	22	31	0.35	27
Metabolite Unk 4	168	10	26	28	3.19	28
Pyroglutamic acid*	206	8	8	46	2.44	29

- Standard F-ratio
- Control-Normalized F-ratio
- Both F-ratio methods

Signal Distributions for 4 Notable Metabolites (selective m/z)



Inflammation Biomarker

Up-regulated in patients

Inflammation Biomarker

Up-regulated in patients

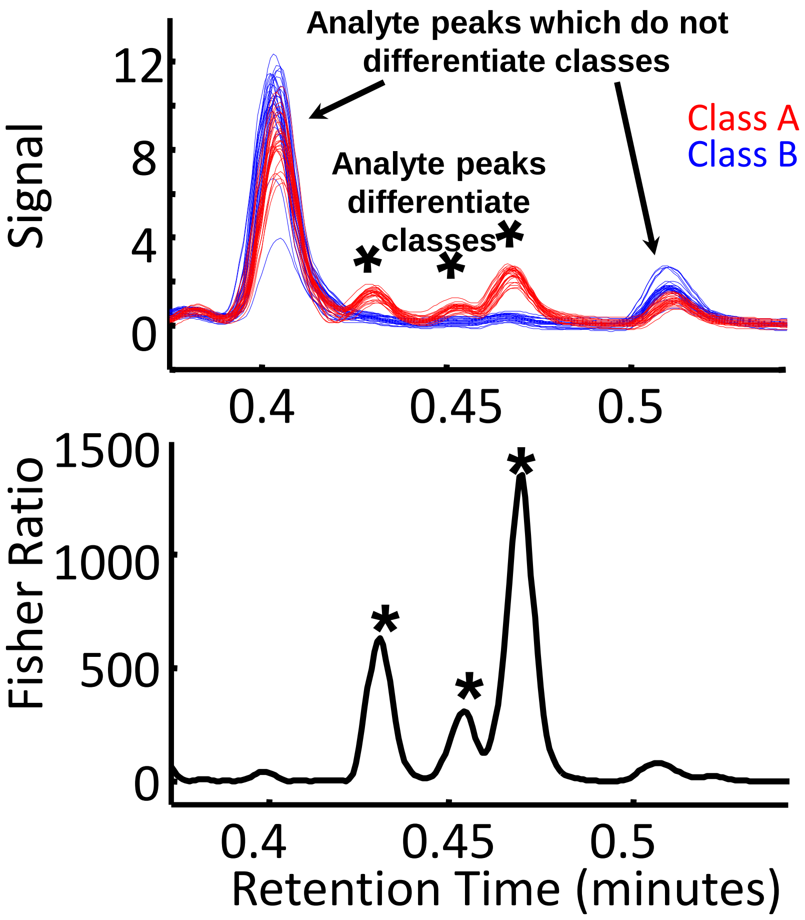
Down-regulated in patients

Native “control” metabolite

Patients

Controls

Evolution of Fisher Ratio (F-ratio) Analysis for Supervised Comparisons



F-ratio Analysis provides a ranked hitlist of analytes that are likely to be statistically different in concentration (p-value < 0.05) between sample classes.

$$Standard\ \ F-ratio = \frac{Between\ Class\ Variance}{\sum(Within\ Class\ Variance)}$$

- K.J. Johnson, R.E. Synovec, J. Chemom. Intell. Lab. Syst., 2002, 60, 225-237.
- L.C. Marney, W.C. Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Talanta 2013, 115, 887–895.
- B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Anal. Chem. 2015, 87, 3812–3819.
- N.E. Watson, B.A. Parsons, R.E. Synovec, J. Chromatogr. A, 2016, 1459, 101-111.
- B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, J. Chromatogr. A, 2016, 1440, 179-190.
- B.C. Reaser, B.W. Wright, R.E. Synovec, Anal. Chem., 2017, 89, 3606-3612.

$$Control\text{-}Normalized\ F-ratio = \frac{Between\ Class\ Variance}{Controls\ within\ class\ variance}$$

- S.E. Prebihalo, G.S. Ochoa, K.L. Berrier, K.J. Skogerboe, K.L. Cameron, J.R. Trump, S.J. Svoboda, J.K. Wickiser, R.E. Synovec, Anal. Chem. 2020, 92, 15526–15533.

$$Minimum\ Variance\ Optimized\ \ F-ratio = \frac{Between\ Class\ Variance}{Mimimum\ within\ class\ variance}$$

- S. Schöneich, G.S. Ochoa, C.M. Monzon, R.E. Synovec, J. Chromatogr. A, 2022, 1667, 462868.

Metabolite Determination in Pacu Fish by GC×GC-TOFMS and ^{MVO}F-Ratio Analysis

Sonia Schöneich

- Pacu (*Piaractus mesopotamicus*) are an important food source in South American countries such as Brazil and Argentina

- Traditionally, fish are tank-bred and reared where they can be monitored closely

- A new practice is rotating the same land area for farming rice and pacu

- **Could herbicides and pesticides used in rice farming affect the fish metabolome?**

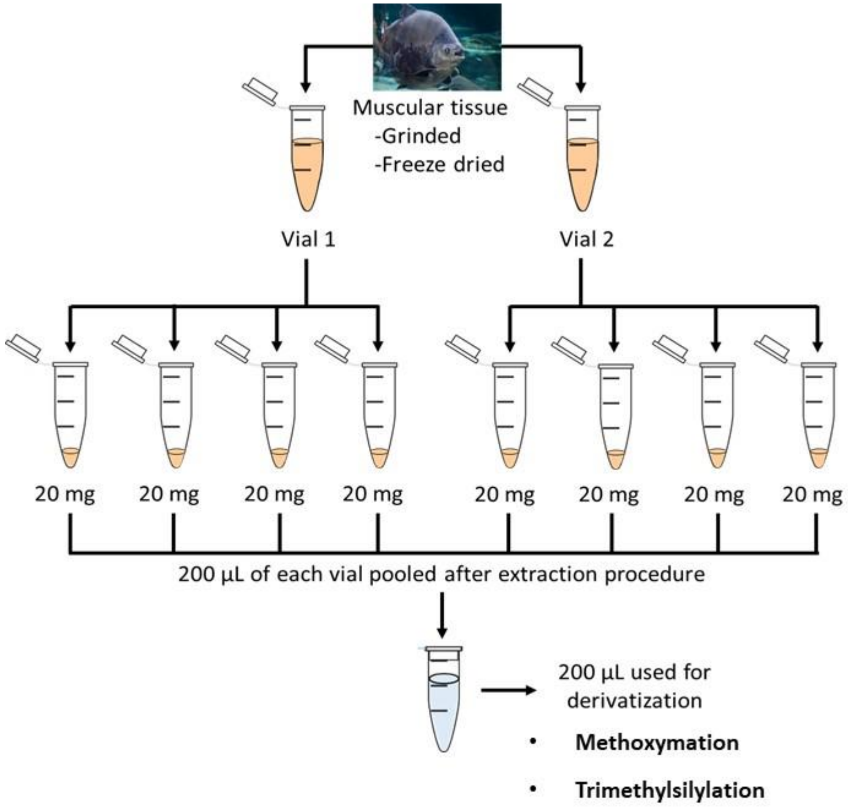
Tank-raised Fish

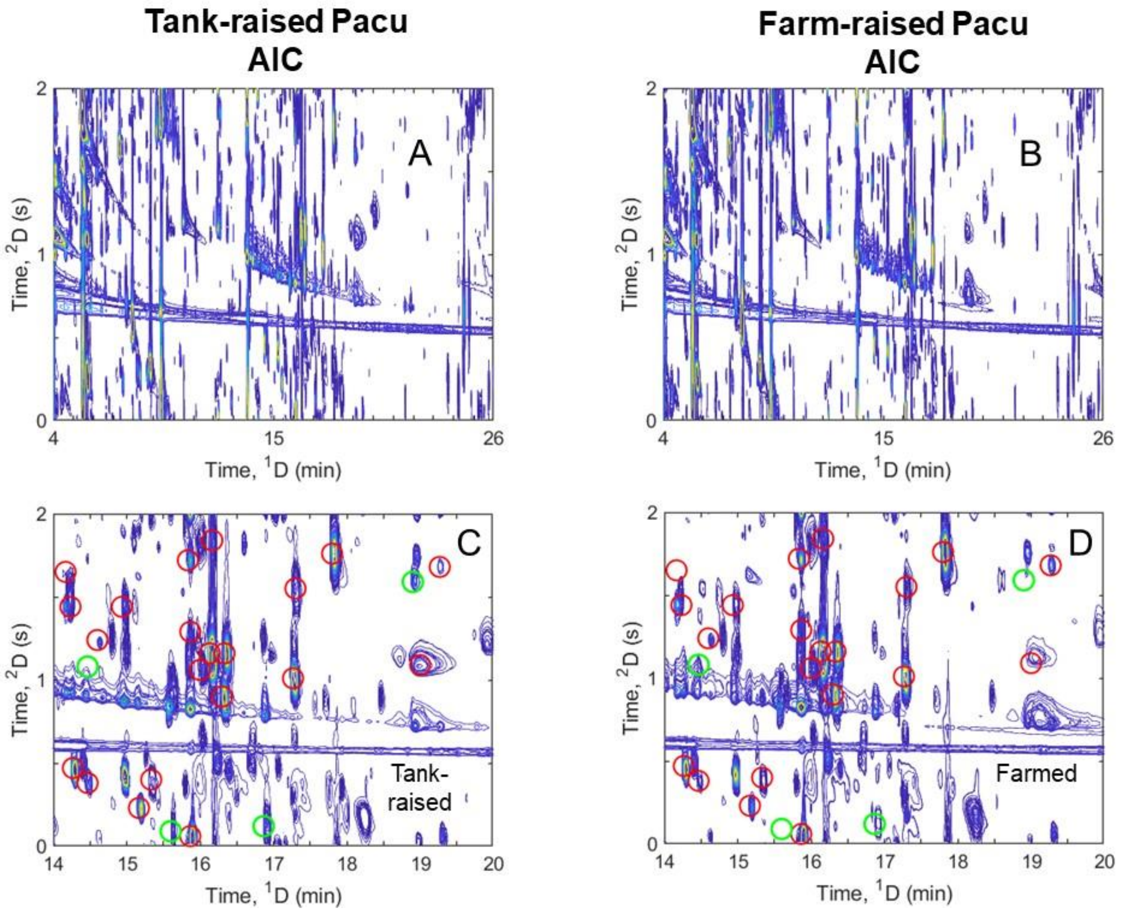
Farm-raised Fish & Rice

Photo Credit: G. Wicki, L. Luchini, L. Romano, and S. Panne Huidobro, Stock densities, growth, and survival for pacu

- **C. Monzón, S. Schöneich, R.E. Synovec, Microchem. J. 164, 2021, 106004.**
- **S. Schöneich, G.S. Ochoa, C. Monzón, R.E. Synovec, J. Chromatogr. A, 2022, 1667, 462868.**

Experimental Design, Sample Preparation, and Data Collection





- 10 Farm-raised fish and 10 Tank-raised fish were obtained
- Each pooled sample per each fish was analyzed in duplicate, resulting in 40 chromatograms

- Analytical ion chromatograms (AIC) using *m/z* 73, 117, 174, and 217, corresponding to trimethylsilyl derivatives, alcohols, carboxylic acids, primary amines, and sugars
- Analyte peaks circled in green have readily visible differences that are easily found by F-ratio analysis, while peaks circled in red possess much less obvious differences, but are also discovered by ^{MVO}F-ratio analysis

F-Ratio Analysis Hitlist Results

Some Interesting Hits

2-butanol (Hit# 2) – associated with sweet, flowery odor in fish paste

γ-butyrolactone (Hit# 3) – sold as "fish tank cleaner" used in synthesis of pesticides and metabolizes into to 4-hydroxybutyric acid (Hit# 28)

L-ornithine (Hit# 7) – important for tissue repair and immune response

2,3-butanediol (Hit# 13) – associated with fishy odor

L-isoleucine (Hit# 26) – plays role in purine biosynthesis as precursor to aspartic acid

4-hydroxybutyric acid (Hit# 28) – affects central nervous system (sedation, memory impairment)

Hitlist: Top 30 Analytes (110 Total)

Analyte	F-ratio	Hit#	*m/z*	[F]/[T]	p-value
propylene glycol	27802	1	117	0.17	5.8E-04
2-butanol	19113	2	117	0.029	1.8E-05
γ-butyrolactone	13456	3	86	0.12	0.011
tyrosine	5951	4	179	0.051	0.021
DL-phenylalanine	2098	5	74	0.080	1.6E-04
scyllo-inositol	1222	6	318	0.15	0.0022
L-ornithine	816	7	174	0.029	0.049
glycerol	751	8	74	1.9	1.1E-05
D-glucose	750	9	135	0.67	5.9E-05
succinic acid	702	10	147	0.15	0.096
glycolic acid	687	11	204	0.50	9.9E-07
2,3-butanediol	591	13	117	2.9	2.3E-07
caproic acid	551	14	73	0.24	9.4E-07
glycerol-3-phosphate	465	15	357	0.45	0.065
N-(hydroxymethyl)trifluoroacetamide	390	17	170	3.1	1.7E-05
2,2'-methylenediphenol	389	18	171	0.27	0.0041
mannose	372	19	160	0.63	2.2E-04
sarcosine	310	20	116	0.14	0.0088
Pyruvate	289	21	174	0.47	0.012
methylaminoheptane	271	24	170	3.0	5.2E-06
D-alloisoleucine	239	25	86	0.41	0.0023
L-iosleucine	226	26	218	0.18	9.6E-04
4-hydroxybutyric acid	211	28	147	2.7	3.1E-06
bromosuccinate	204	30	73	0.33	8.3E-06
L-leucine	200	31	73	0.27	6.1E-05
methionine	190	32	147	0.28	0.017
L-serine	187	33	116	0.41	5.3E-04
N,N-dimethylethanolamine	136	38	58	2.1	6.7E-07
mercaptoacetic acid	132	39	221	0.44	9.1E-04
asparagine	122	41	80	0.53	2.8E-06

[F]/[T] = [Conc in Farm-raised] / [Conc in Tank-raised]

Pacu Fish Study Observations

- **Of the 110 analyte hits, 70 expressed a concentration ratio statistically different than 1 ($p < 0.05$)**
- **A majority of the changing analytes (54 out of 70) that are important for normal biological functioning of pacu fish were significantly downregulated in the farmed fish**
- **The downregulation of these analytes suggests the integrated farming system possibly impacts the pacu fish quality**

Tank-raised fish

Versus

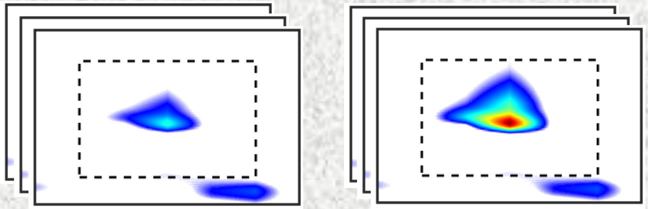
Farm-raised Fish & Rice

Tile-based 1v1 Analysis ➡ Fold-Change

Non-Targeted Discovery Analysis Comparing Two Chromatograms

Caitlin Cain

Tile-based F-ratio Analysis

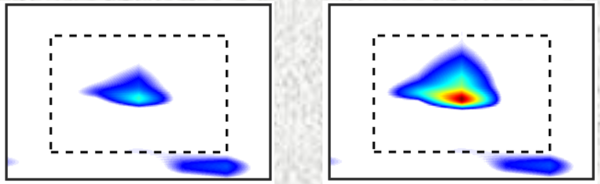


$$F - ratio = \frac{Between\ Class\ Variance}{\sum(Within\ Class\ Variance)}$$

Replicates may not always be available due to sample, time, and/or expense limitations

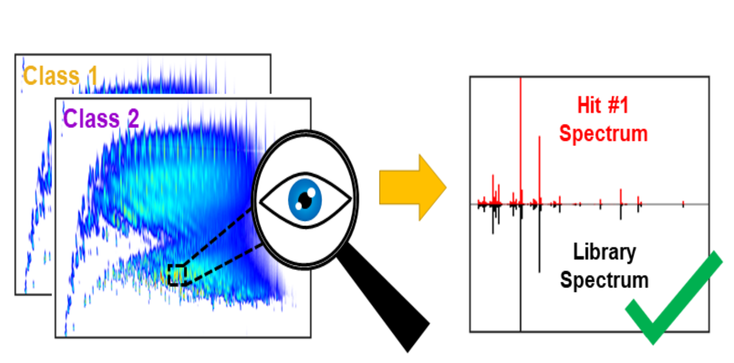
Tile-based 1v1 Analysis

$$Rank\ Metric, RM = \frac{|Class\ 2 - Class\ 1|}{Class\ 2 + Class\ 1}$$



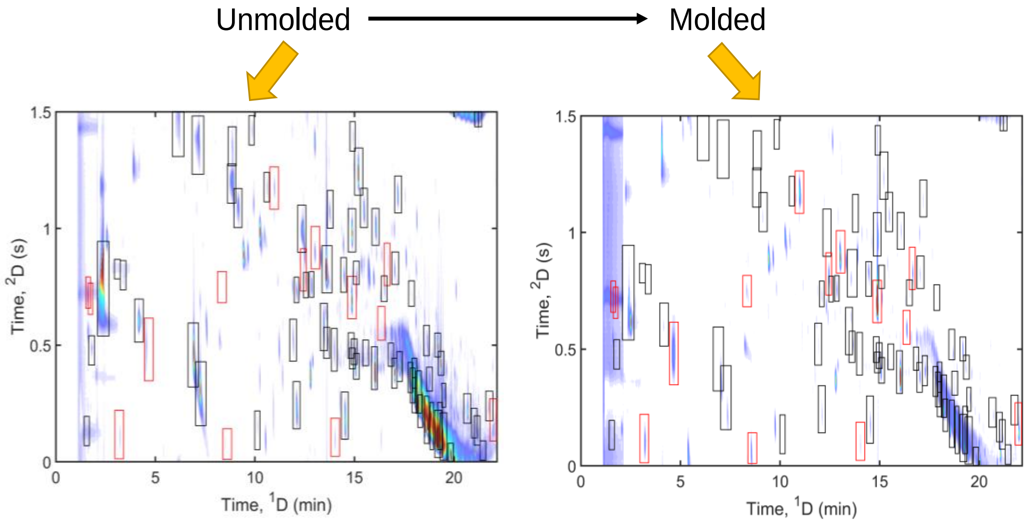
Overcomes issues associated with pixel-based subtraction plots

Obtain high quality mass spectra, much better than chemometric decomposition methods



- Cain, C. N.; Ochoa, G. S.; Trinklein, T. J.; Synovec, R. E. *Anal. Chem. 2022,* 94, 5658–5666.
- Ochoa, G.S.; Sudol, P.E.: T.J. Trinklein, T.J.; Synovec, R.E. *Talanta*, 2022, 236, 122844.
- Humston, E. M.; Knowles, J. D.; McShea, A.; Synovec, R. E. *J. Chromatogr. A*, 2010, *1217*, 1963-1970.

Pairwise Analysis of Cacao Samples

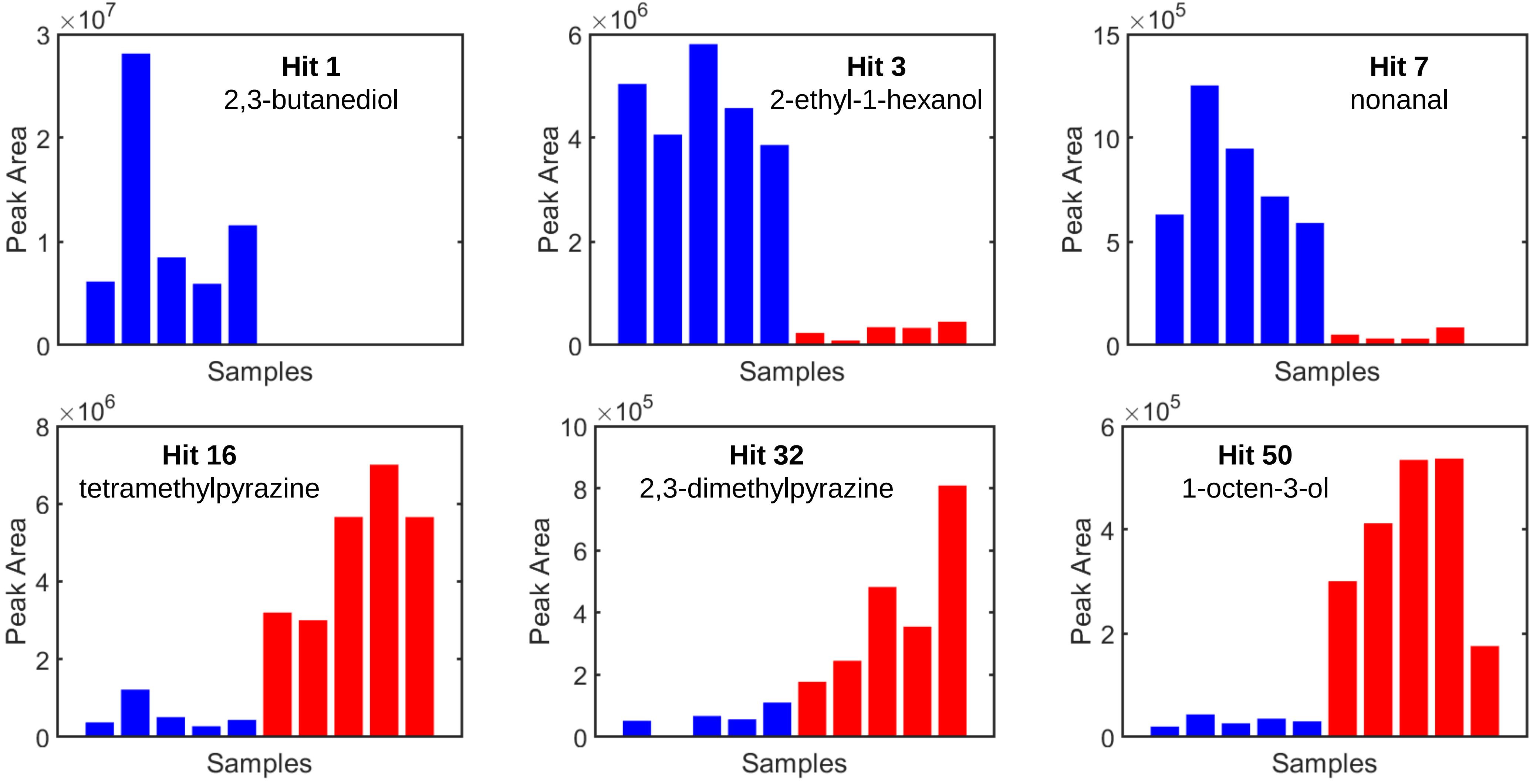


- Black rectangles: peaks upregulated in unmolded
- Red rectangles: peaks upregulated in molded

Cacao Bean Volatiles Affecting the Flavor Profile

Unmolded
Molded

1v1 Applied to 5 Separate Paired Analyses



Sweet, creamy, citrus-like aromas

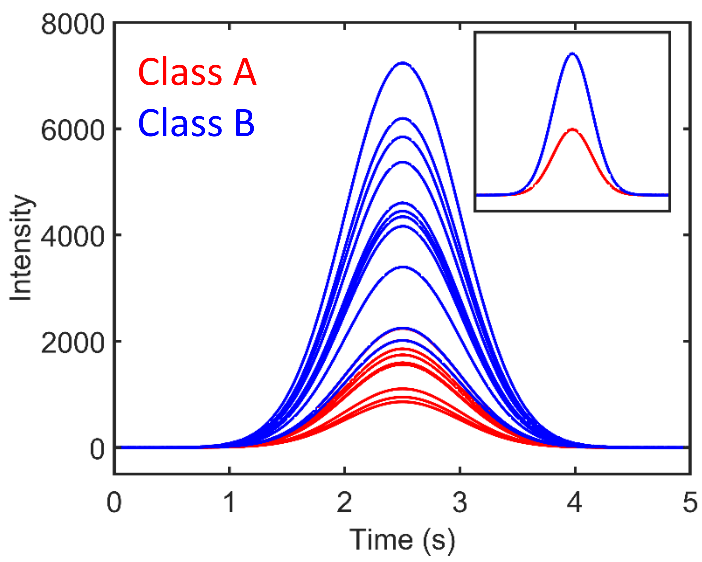
Earthy, moldy aromas

Cain, C. N.; Trinklein, T. J.; Ochoa, G. S.; Synovec, R. E. *Anal. Chem.*, **2022**, *94*, 5658-5666.

Tile-Based Feature Selection Metrics

Supervised
Known Classes

Metrics = *F-ratio* & *RM*



High F-ratio

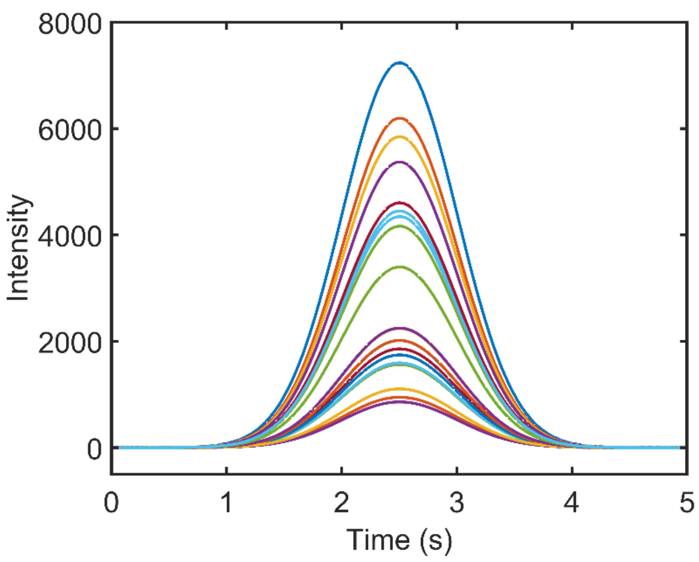
Inset: High 1v1 *RM*

Unsupervised – No Classes

Metric = *RSD*^{2}

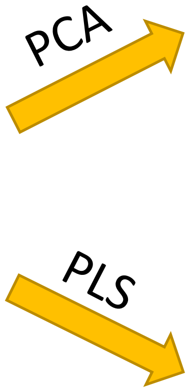
total relative signal variance

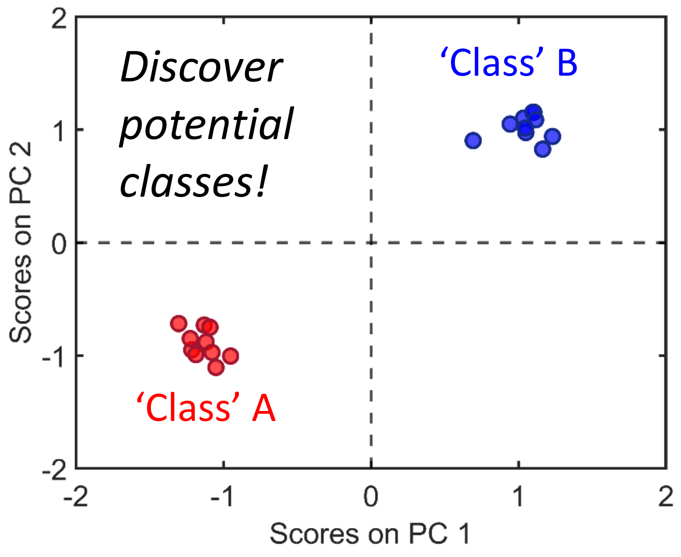
Coefficient of Variance

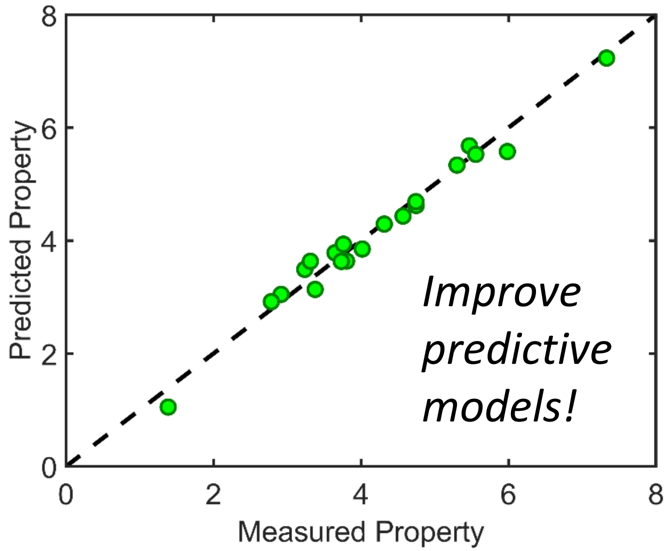


High *RSD*^{2}

Might be a relevant feature in the data set!

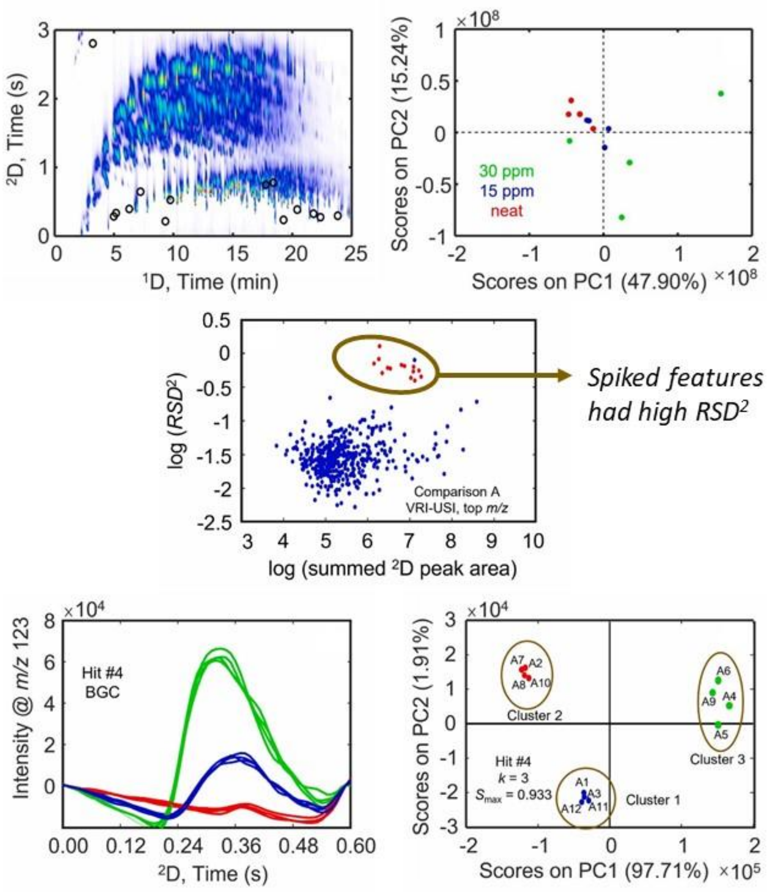




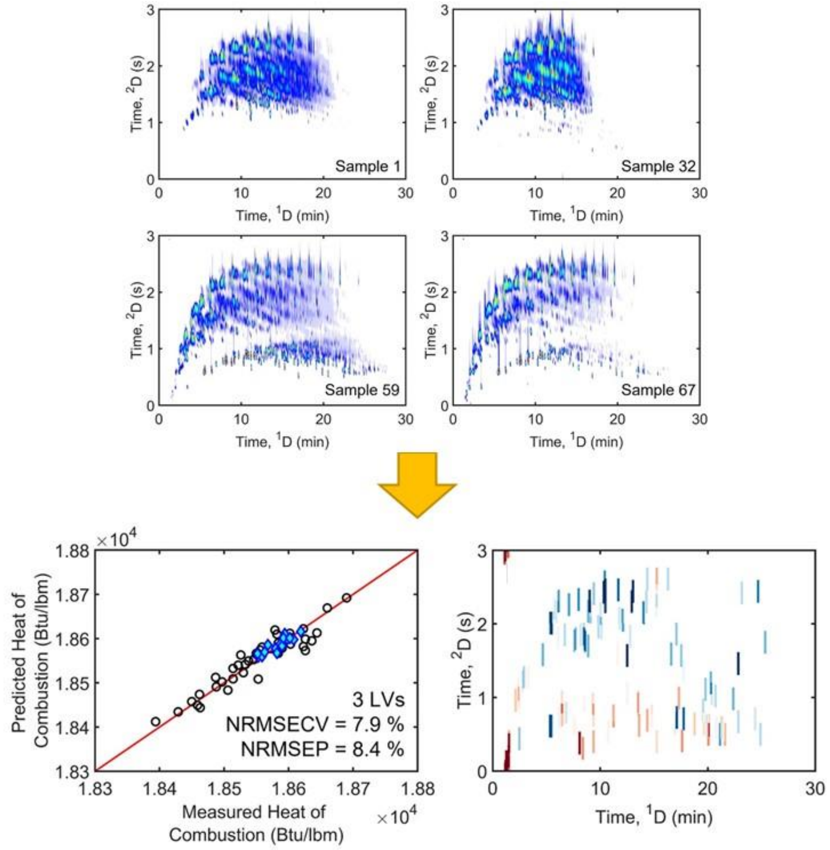


Tile-based Variance Ranking (RSD^2) Analyses of Jet and Rocket Fuels

Discover potential sample classes with PCA



Improve predictive modeling with PLS



Paige Sudol

Grant Ochoa

- **P.E. Sudol, G.S. Ochoa, C.N. Cain, R.E. Synovec, *Anal. Chim. Acta*, 2022, 1209, 339847.**
- **C.N. Cain#, P.E. Sudol#, K.L. Berrier#, R.E. Synovec, *Talanta*, 2021, 233, 122495.**

Caitlin Cain

- **C.N. Cain, G.S. Ochoa, R.E. Synovec, *J. Chromatogr. A* 1694 (2023) 463920.**
- **K.L. Berrier#, C.E. Freye#, M.C. Billingsley, R.E. Synovec, *Energy Fuels*, 2020, 34, 4084-4094.**

CONCLUSION

We are developing sample comparison software tools and analytical methodology to gain a deeper understanding of the subtle differences in the chemical composition between various complex samples using both supervised and unsupervised experimental designs. With the information gained, we can take action to meet the challenges and objective of a wide variety of research studies. Specifically, we are using compound class focused sample preparation and GC×GC–TOFMS coupled with tile-based analysis software tools to achieve our goals.

Thank You!

Synovec Lab

Gas Chromatography, Liquid Chromatography, and Mass Spectrometry, with Multi-Dimensional Data Analysis

Seattle's Best Chromatography

Principal Investigator
- Dr. Robert Synovec

Current/Recent Lab Members
- Peri Abdigali
- Naod Semere Berhe
- Caitlin Cain
- Austin Dobrecevich
- Robert Halvorsen
- Jakob Klein
- Owen Lee
- Wenjing Ma
- Arty Manafe
- Haylee Meissner
- Lina Mikaliunaite
- Cassandra Padilla

- Dr. Sonia Schöneich (PhD 2023)
- Dr. Tim Trinklein (PhD 2023)
- Dr. Grant Ochoa (PhD 2023)
- Vlada Olkovych (MS 2023)