Developing Approaches for Non-Targeted Screening of Complex Sample Matrices Using HR-MS

Ann M. Knolhoff and Timothy R. Croley

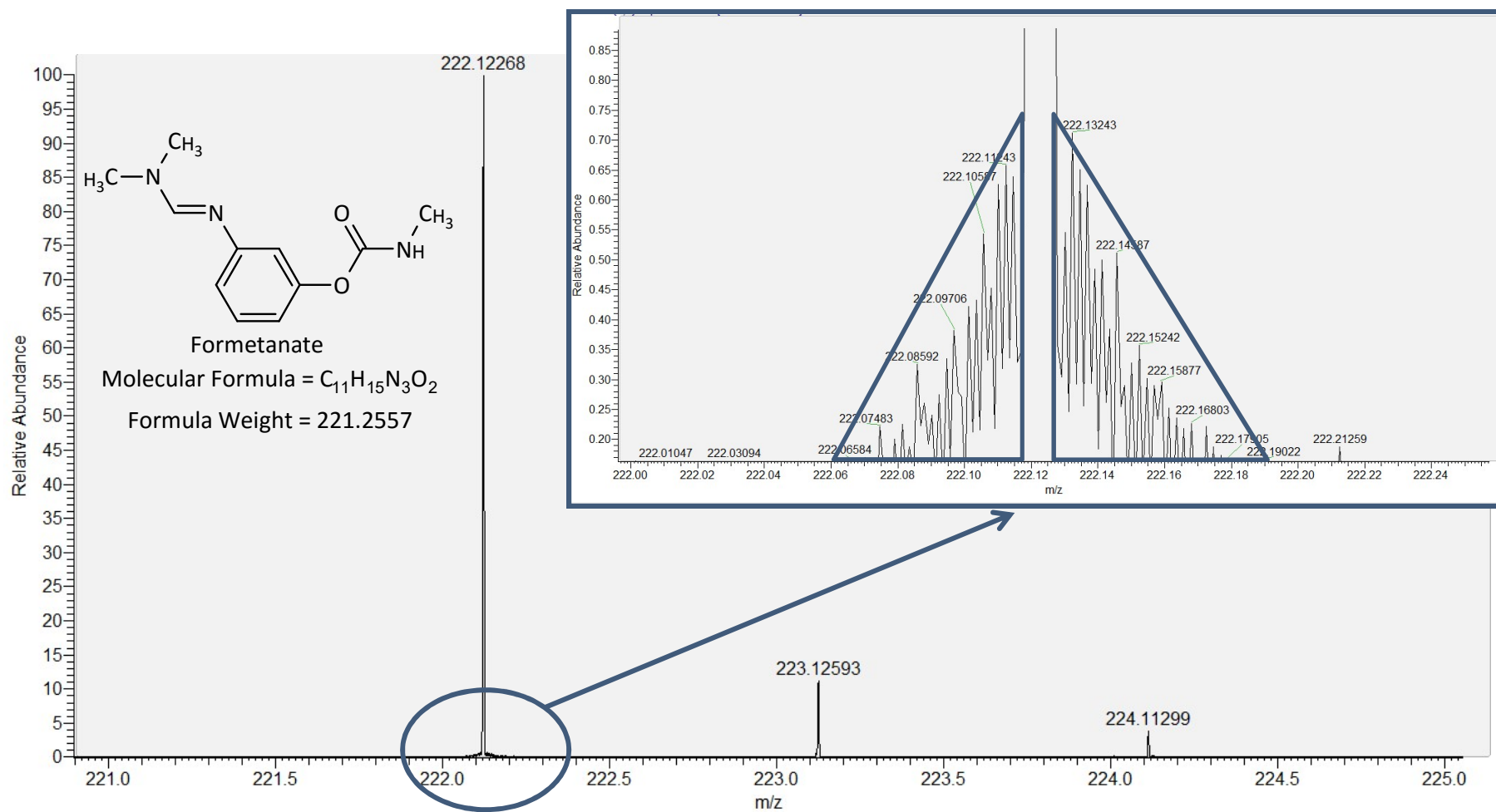Center for Food Safety and Applied Nutrition

Office of Regulatory Science

OFFICE OF
REGULATORY
SCIENCE

Benefits of Non-Targeted Screening

- Globalization of the food supply

- Potential Hazards
    - Contamination
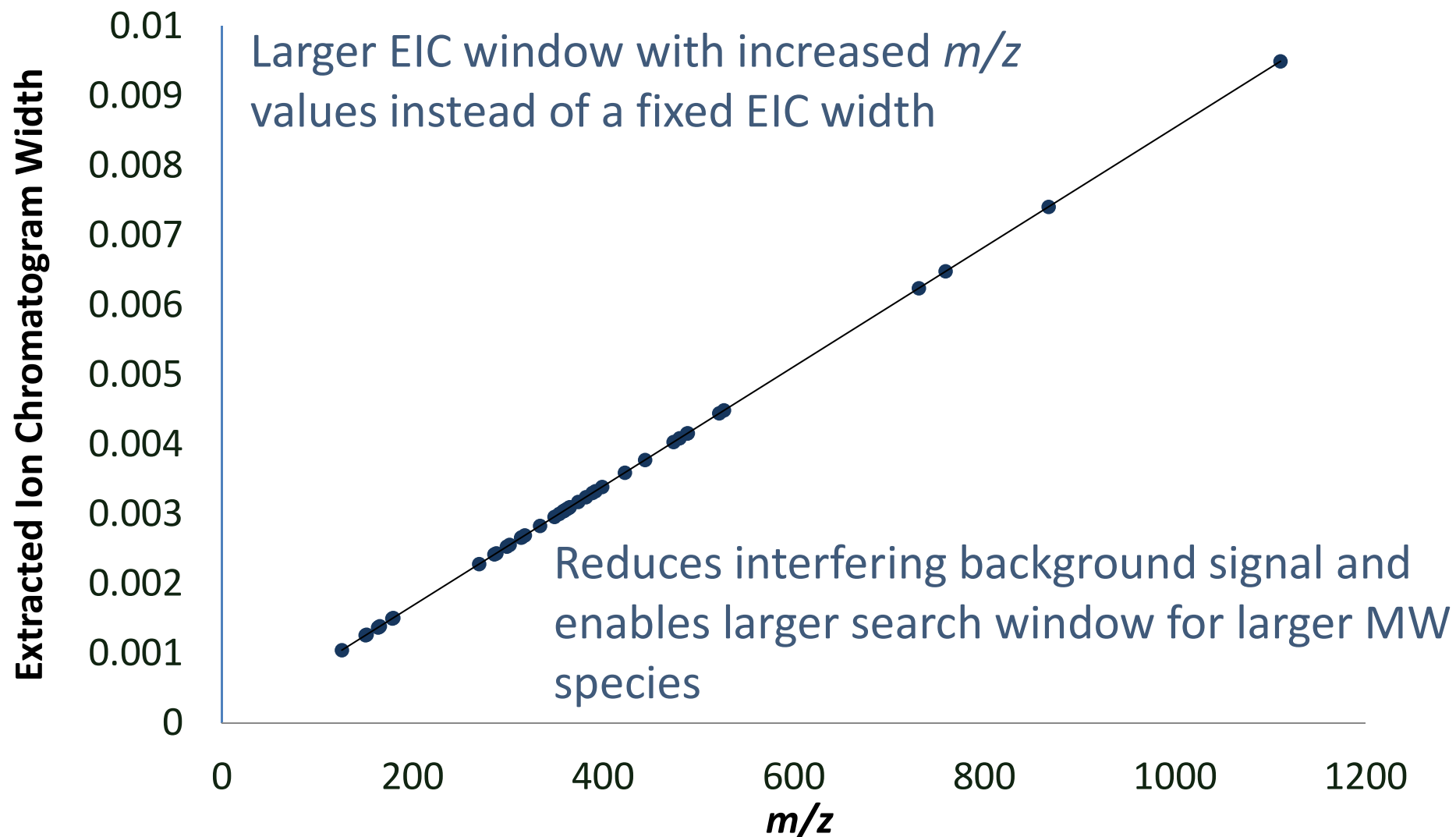   - Adulteration
    - Food fraud

An Overview of our Non-targeted Analysis Projects

- Chemical screening methods
  - Automated data processing
    - Network solution
- Data quality
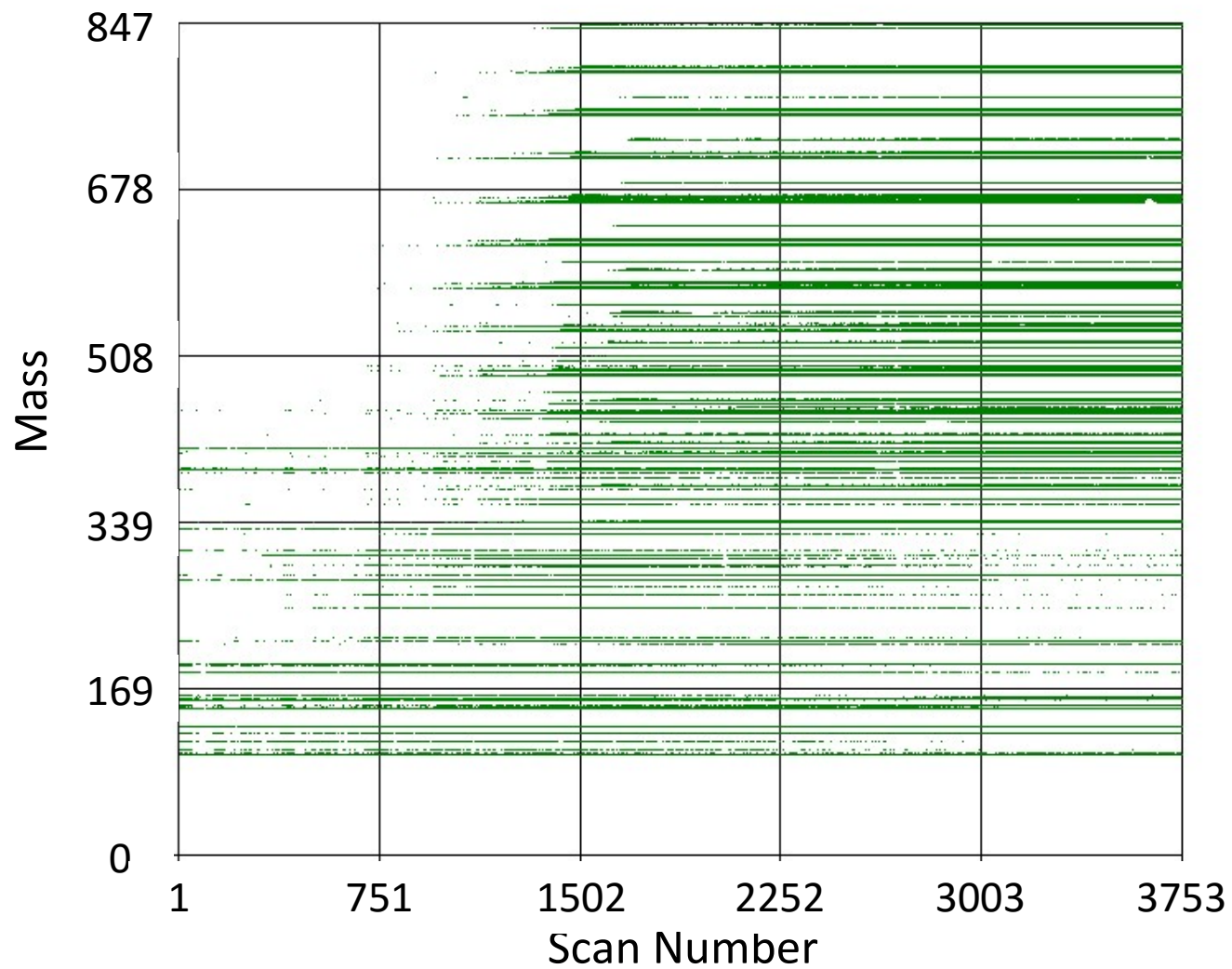- Implementation of chemometric models

Developed Software Features

Automated EIC Determination

Larger EIC window with increased *m/z* values instead of a fixed EIC width

Reduces interfering background signal and enables larger search window for larger MW species

Extracted Ion Chromatogram Width (y-axis: 0, 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01)

m/z (x-axis: 0, 200, 400, 600, 800, 1000, 1200)

Recalibration Using Background Masses

Yogurt,
No Recalibration

Yogurt,
Thermo Lock
Mass Calibration

Yogurt,
Recalibrated
with Developed
Software

Mass accuracy (ppm)

m/z

● 2000 pg   ● 100 pg
● 500 pg   ● 10 pg
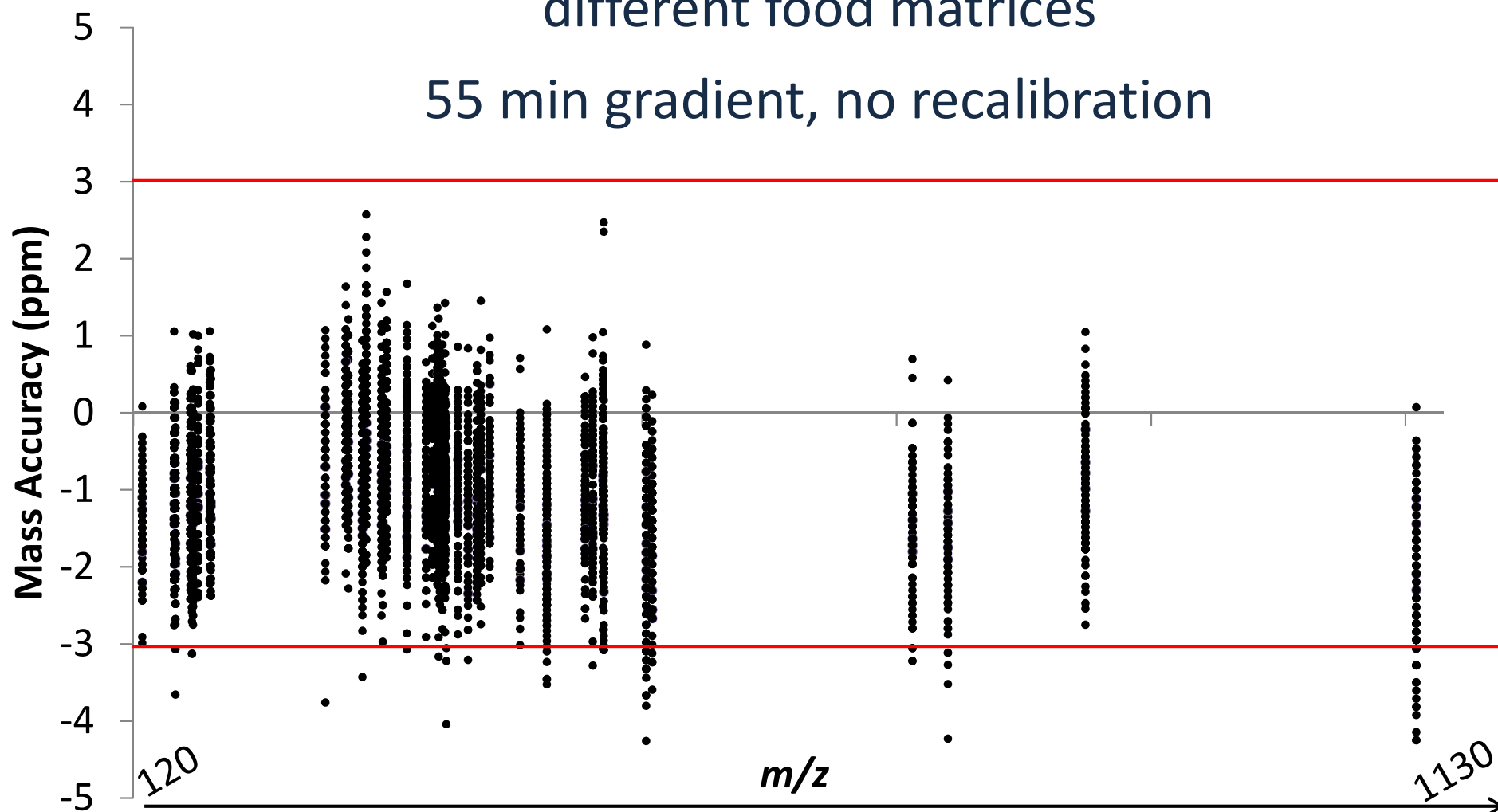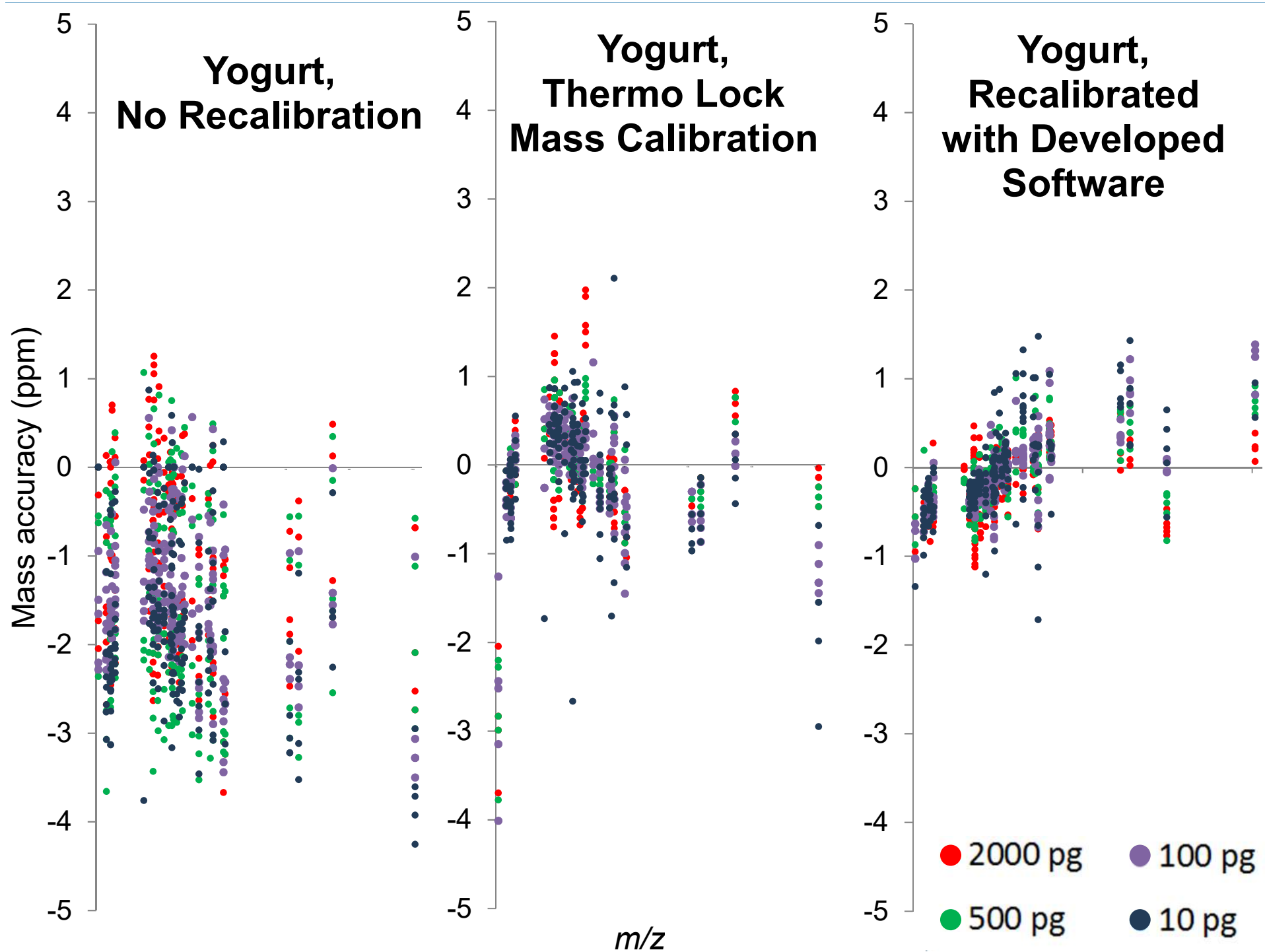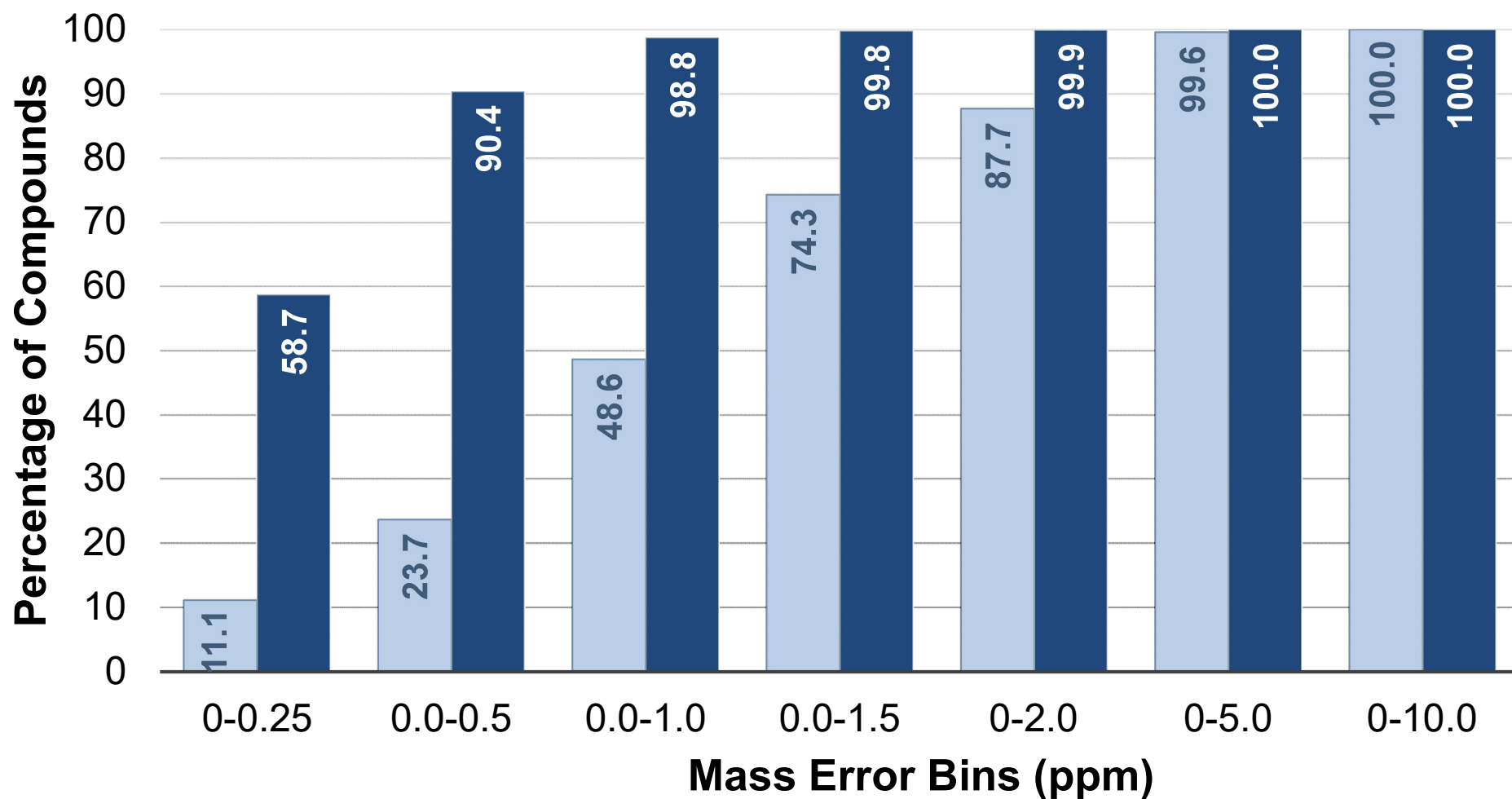
Advantages to Developed Software

- While not strictly non-targeted screening, it provides:
  - Automated processing
  - Screening against a known library notifies the analyst that potentially problematic compounds are or are not detected in the sample
  - Recalibration should yield better compound matches and fewer/more accurate generated molecular formulae

Peak Coalescence

NL: 4.30E7
TIC
MS

TIC

NL: 6.96E6
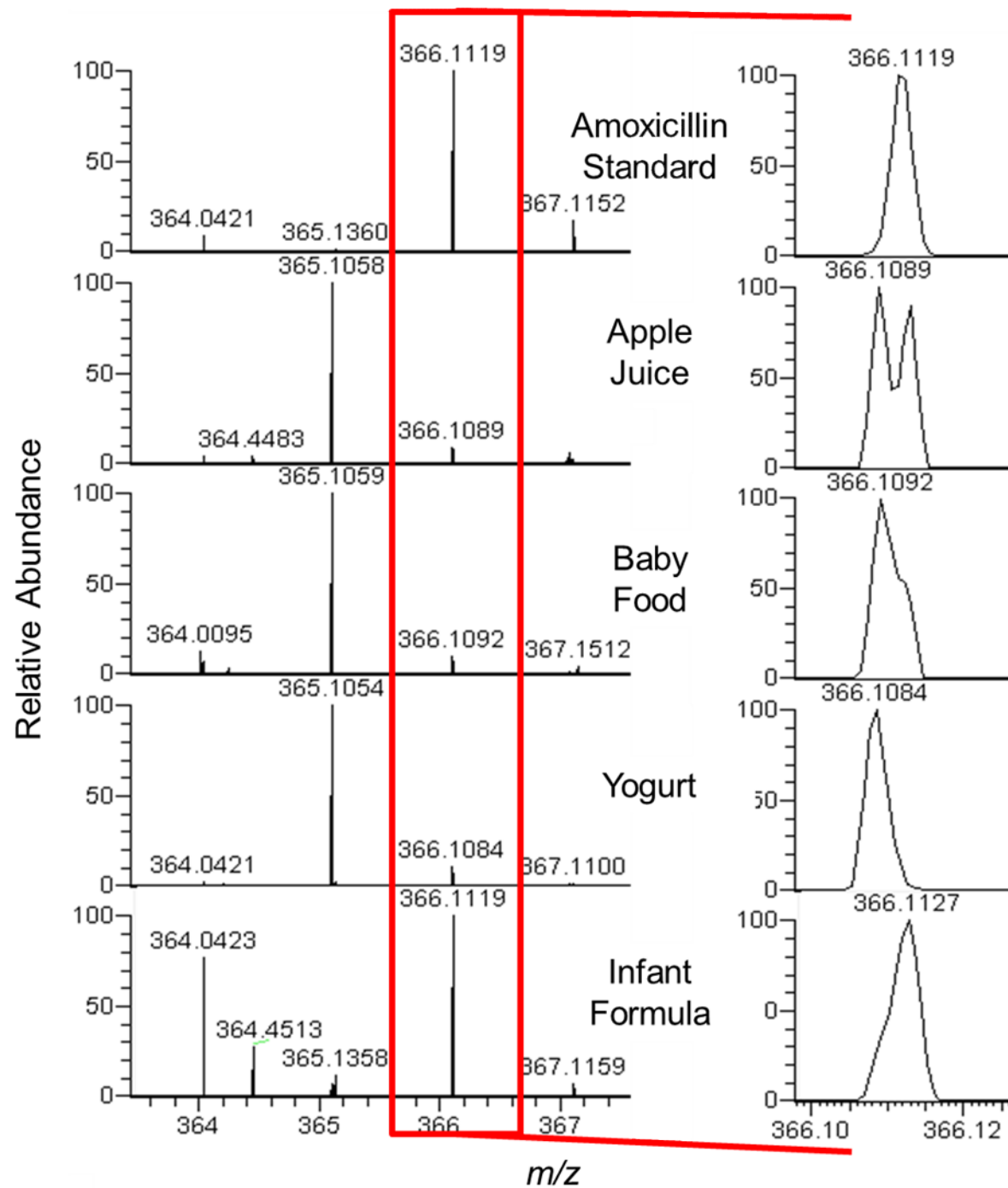m/z 314.07043-
314.07357
MS

Triazophos
m/z_{theor} **314.071177**

NL: 1.60E7
m/z 314.08043-
314.08357
MS

Hexaconazole
m/z_{theor} **314.082144**

Relative Abundance

Time (min)

314.08185

316.07881

315.08522

< 2 ppm

314.07797

316.07859

315.08474

~15 ppm

314.07225

315.07510
316.06772

< 2 ppm
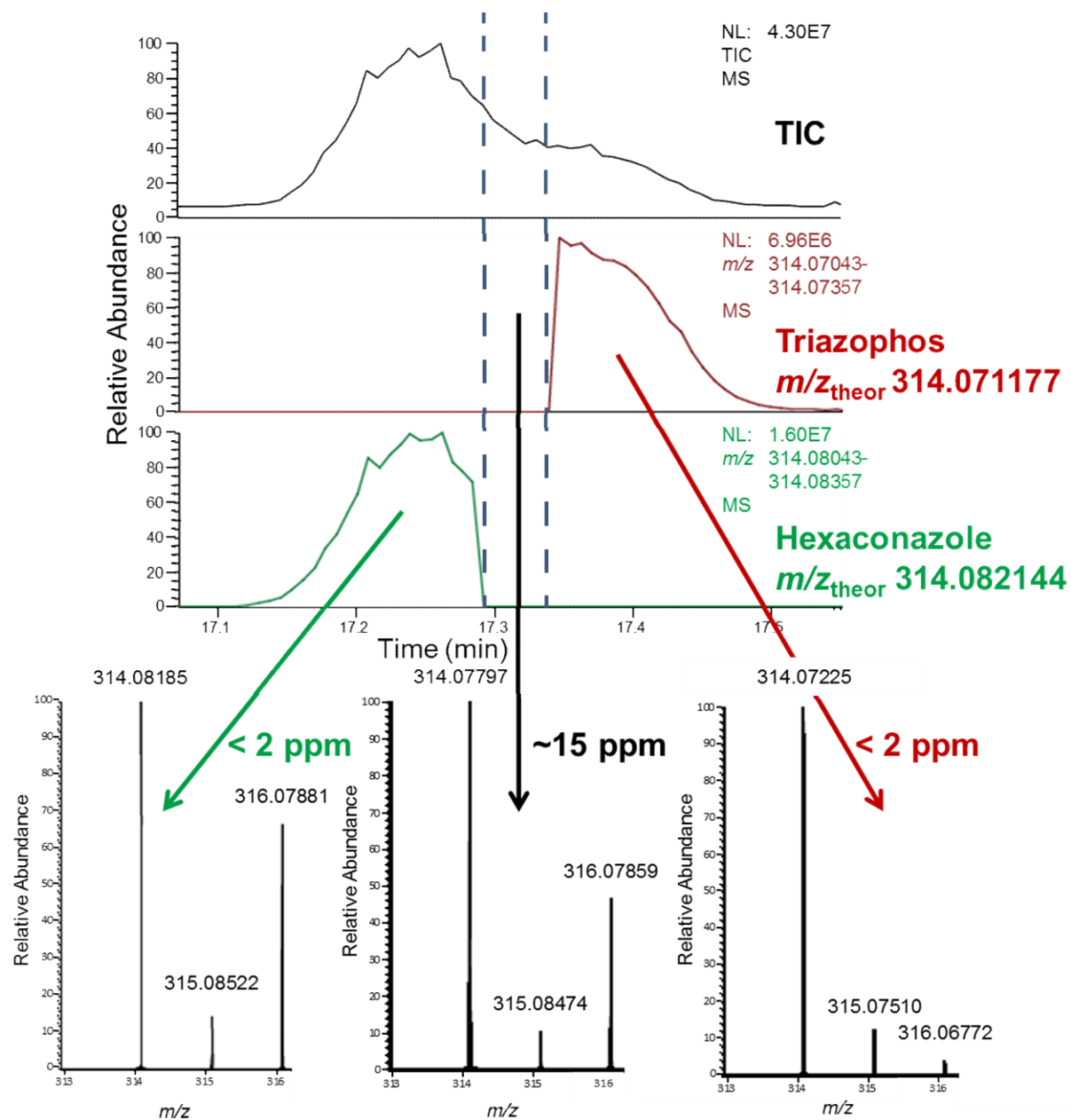
Peak shape

Impaired Data Quality: Example 3

Orbitrap A+1 Deviation Versus Monoisotopic Peak Intensity
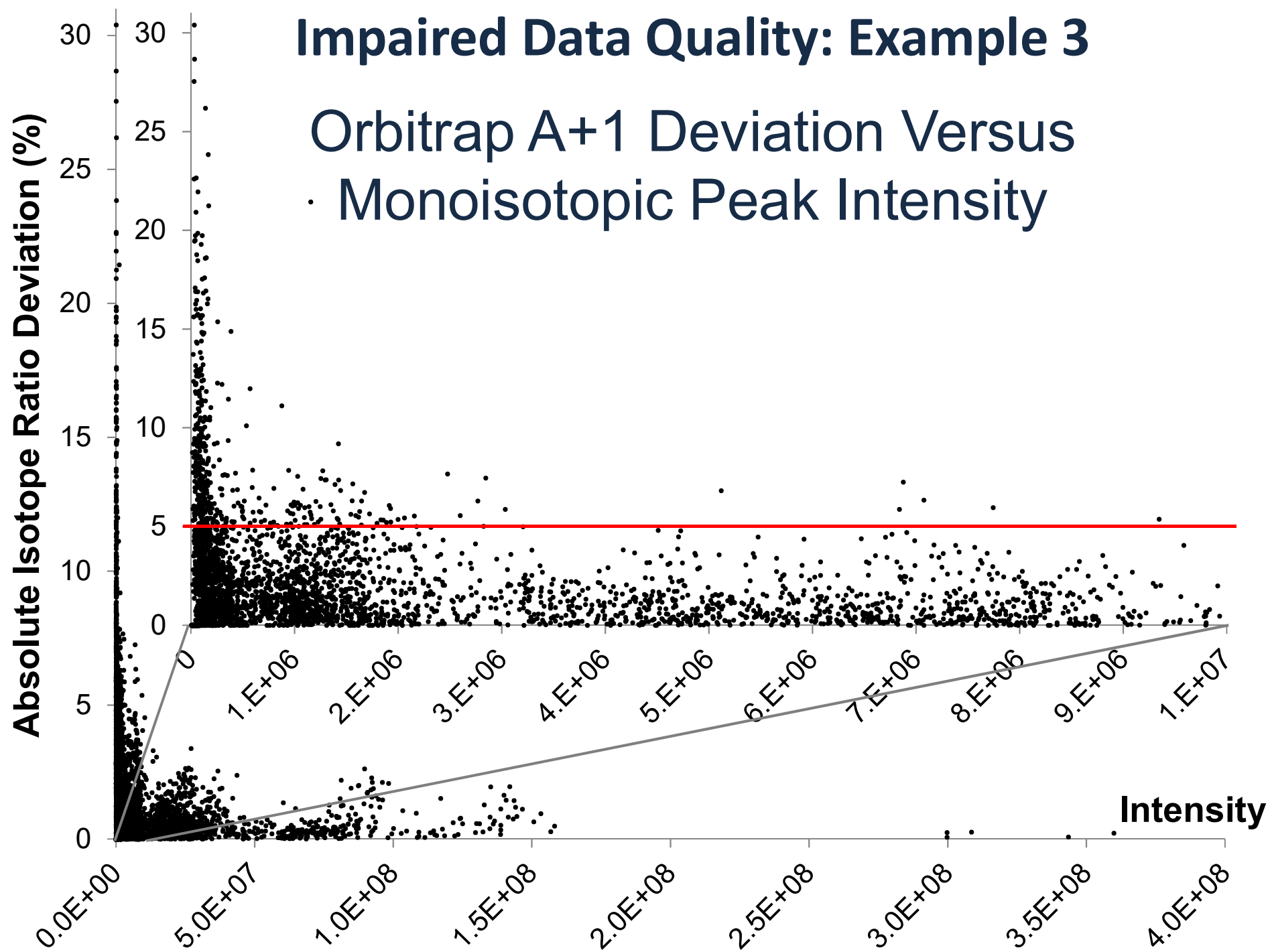
Finding Molecular Features and Chemometric Analysis

- LC/HR-MS of complex sample matrices result in large data sets
- ID of every detected compound in a sample is lengthy
- Strategies
  - Control VS new set of samples
  - Identification of outlier samples from well defined sample pool
- Data quality, method development, and data filtering are critical

Data Analysis Method Optimization

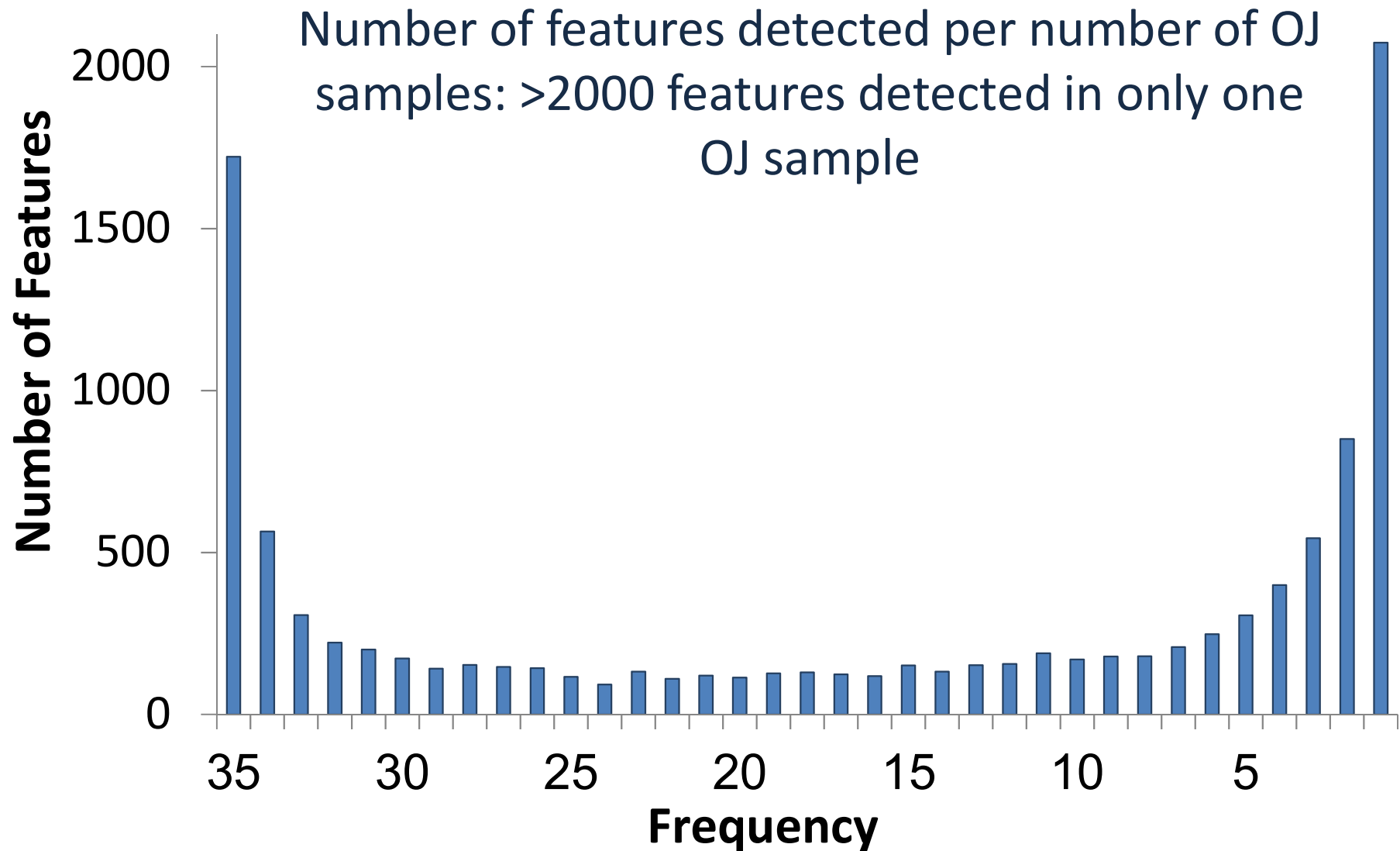## Molecular feature extraction tested with QC standard mixture and food matrix spikes

Method Parameter	QC compounds detected	All compounds present in 4/5 10 ppb spikes	All compounds present in 4/5 100 ppb spikes	All compounds present in 4/5 500 ppb spikes
No peak limit	No	-----------------	-----------------	-----------------
Blank removal	No	-----------------	-----------------	-----------------
50 quality score	Yes	Yes	Yes	Yes
2-29 min RT restriction	Yes	No	No	Yes
4-20 min RT restriction	Yes	Yes	No	Yes

Data Filtering

Number of features detected per number of OJ samples: >2000 features detected in only one OJ sample

Overall Workflow

Feature extraction

↓

Confirm extraction of compounds
in QC & matrix spikes

◇ Y

Limit to 2/total features amongst all samples

↓

Limit to 60%(initial) or 80% (repeat) of features
within sample grouping

↓

Unpaired t-test

↓

2-fold increase in suspect sample group

↓ Only on repeat analysis

Bin features with the same RT and choose
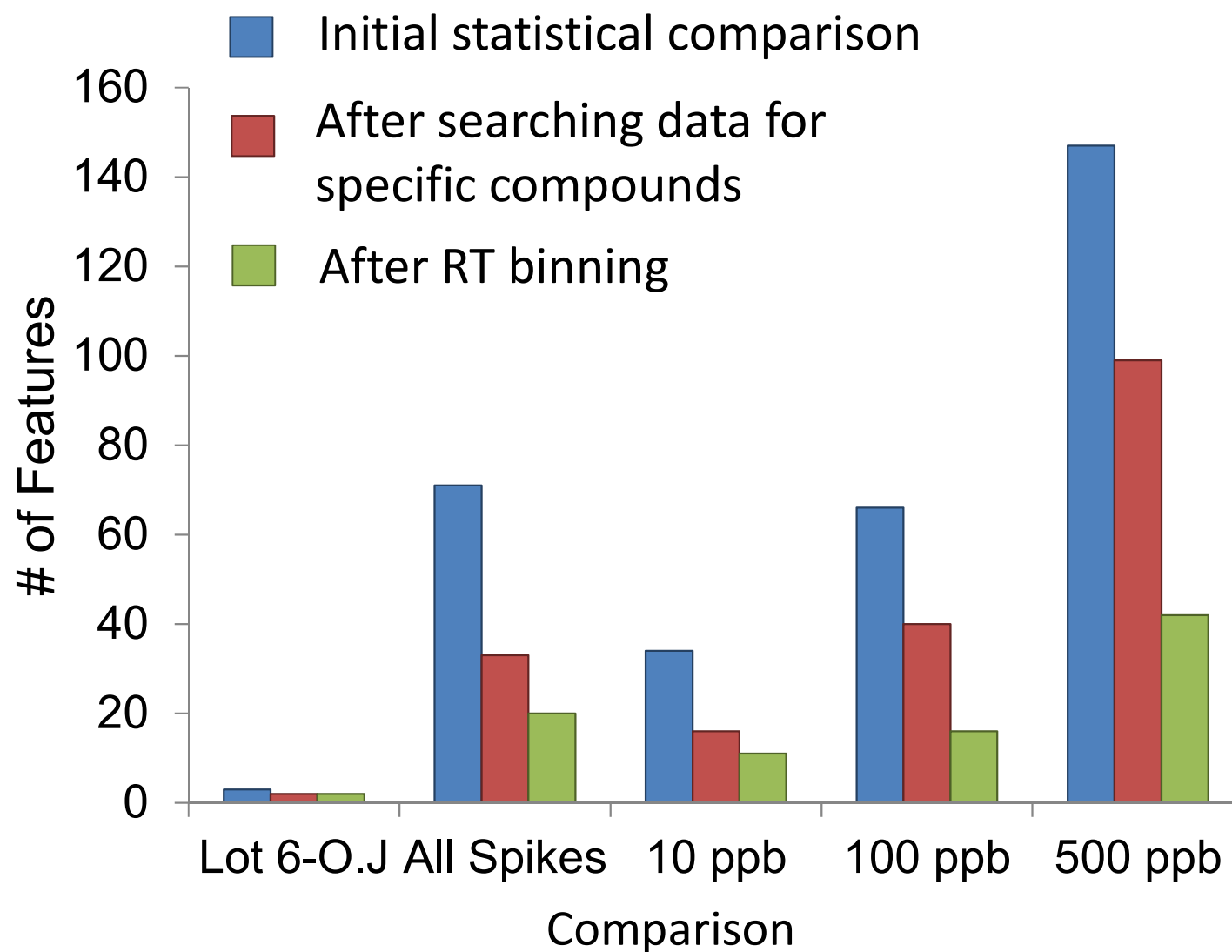most abundant feature

Optimize
method

◇ N

Search data for
specific compounds

Data Reduction

Number of Compounds Found for Molecular Formulae

Generated Molecular Formula	ChemSpider	SciFinder	PubChem	Metlin
$C_8H_8N_2O_2$	515	1414	1	1
$C_{15}H_{30}O_2S_2$	1	36	0	0
$C_{18}H_{21}NO_3$	3974	9675	15	13
$C_{16}H_{10}N_7O_3$	2	0	0	0
$C_{17}H_{25}NO_{10}$	21	129	1	0
$C_{21}H_{26}N_2O_3$	6984	11430	3	9
$C_{29}H_{47}N_5OS_4$	0	0	0	0
$C_{21}H_{23}NO_6$	1786	2414	0	8
$C_{22}H_{25}NO_6$	1721	2301	2	3
$C_{22}H_{22}O_9$	114	360	1	17

Factors That Can Influence Chemometric Analyses

- Blanks
- QC standard
- Matrix spike
- Sample replicates
- Heterogeneity between samples
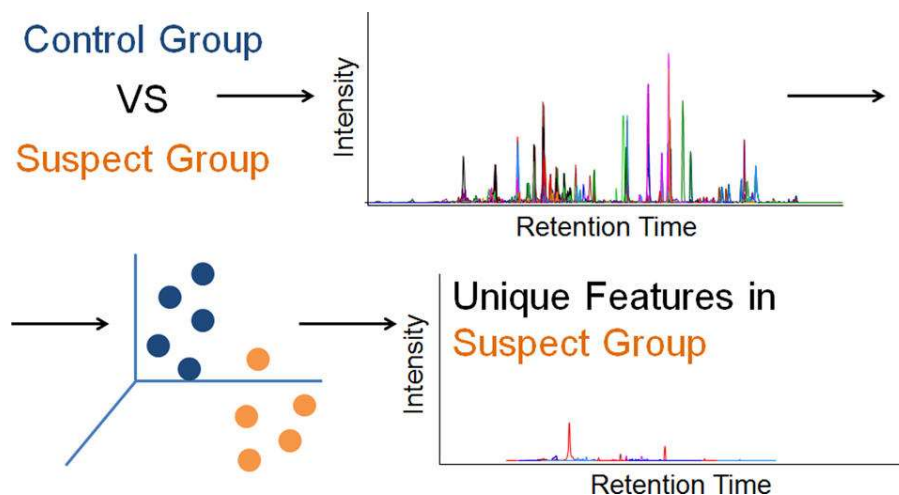
Work in Progress

- Improving chemical coverage
  - Different sample preparation
  - Positive and negative ion mode
- Data quality
  - Resolving power
  - Inclusion of MS/MS scans
  - Chromatographic influence on molecular feature detection
  - Centroided data
  - Polarity switching

Conclusions

- Screening against a substance list is not non-targeted screening, but it enables higher throughput analysis and the detection of known hazardous compounds.
- Data quality is influential in the feasibility of non-targeted workflows. To ensure sufficient data quality, consider:
  - Ion abundance
    - Improved isotopic abundance measurements
  - Chromatography and resulting peak capacity
    - Reduces ion suppression, peak coalescence, and mass accuracy errors
  - Indicators of impaired data quality
    - Peak width and intensity
    - Coeluting compounds with similar *m/z* values
- Method development and data reduction are critical for chemometric analyses of complex matrices to minimize the number of compounds that require ID.

Acknowledgments

- FDA
  - Caitlin Kneapler
  - John Callahan
  - Samantha Farris
- Virscidian
  - Mark Bayliss
  - Stephane Murphy
  - Joe Simpkins
- Agilent
  - Jerry Zweigenbaum

References

- Knolhoff, A. M.; Callahan, J. H.; Croley, T. R. *J. Am. Soc. Mass Spectrom.* 2014, 25, 1285-1294.
- Knolhoff, A. M.; Zweigenbaum, J. A.; Croley, T. R. *Anal. Chem.* 2016, 88, 3617-3623.
- Croley, T.; White, K.; Callahan, J.; Musser, S. *J. Am. Soc. Mass Spectrom.* 2012, 23, 1569-1578.