Effectively Dealing with Transition Selection and Data Analysis for Multiplexed Quantitative SRM-MS Assays across Multiple Vendor Instruments

Susan Abbatiello, Ph.D.

Skyline User's Meeting

May 20, 2012

BROAD INSTITUTE

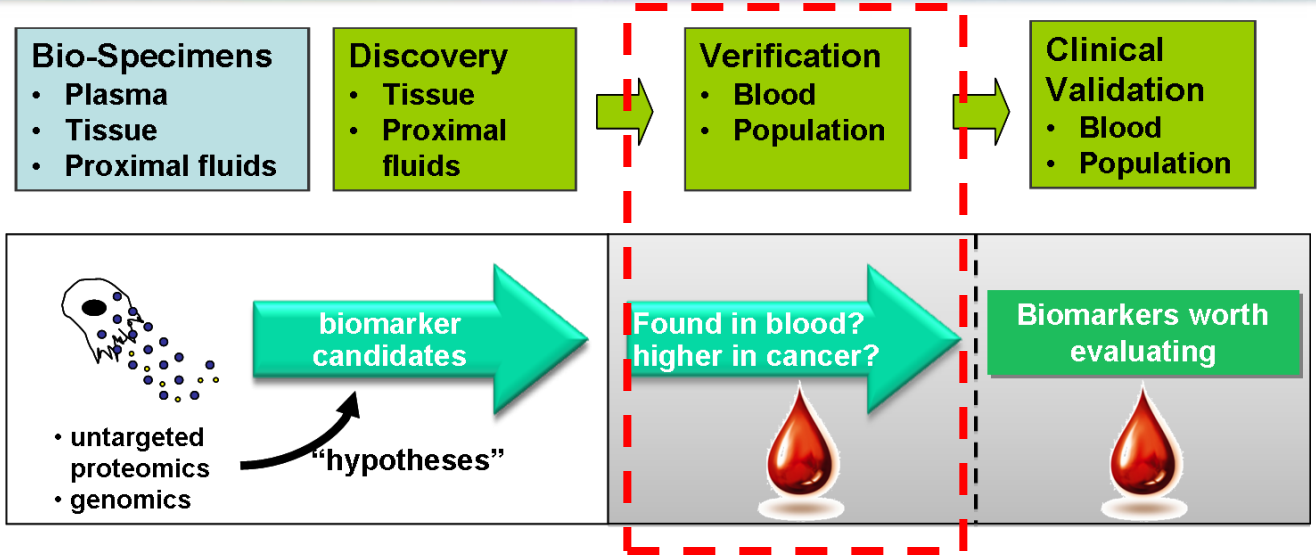
CLINICAL PROTEOMIC TECHNOLOGIES FOR CANCER

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

NATIONAL INSTITUTES OF HEALTH

NATIONAL CANCER INSTITUTE

CPTAC – Clinical Proteomic Technologies Assessment for Cancer



NCI established CPTAC October 2006 to Support Biomarker Development

- **Evaluate and standardize proteomic verification platforms for analysis of cancer-relevant proteomic changes in human clinical specimens.**

Is SID-MRM-MS Technology Reproducible, Transferrable, and Sensitive? Yes! – Especially with Skyline!

Candidate Protein Biomarkers

↓

Define "Signature peptides" for candidate biomarkers

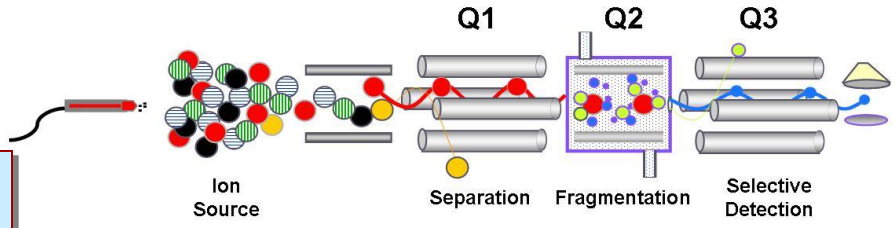
↓

Synthesize $^{13}C/^{15}N$-labeled versions of signature peptides for use as internal standards

→ **Spike heavy ($^{13}C_6$)-labeled peptides**

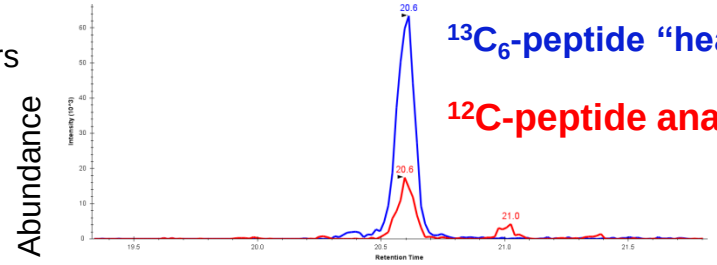
Endogenous ^{12}C signature peptides

↓ **MRM-MS**



- Observed ratio gives precise, relative quantitation across samples
- 10's to 100's peptides can be simultaneously quantified

↓ **Ratio ^{13}C-peptide to ^{12}C-peptide by SID-MRM-MS**



$^{13}C_6$-peptide "heavy standard"

^{12}C-peptide analyte

Whiteaker, et al, JPR 2007……Breast cancer
Keshishian et al, MCP, 2007 and 2009…Cardiovascular markers
Hoofnagle et al, Clin. Chem. 2008……Thyroglobulin
Addona et al, Nat. Biotech. 2009……..……Interlab study
Kuhn et al, Clin Chem 2009……IL-33, Troponin I
Williams et al, JPR 2009……C-Reactive Protein
Ossola et al, Methods Mol. Bio., 2011…..Glycated peptides
Selevsek et al, Proteomics, 2011……..Urine proteins

Study 9S: Participants, Platforms, and Objectives

Prior to analyzing complex samples, are LC-MRM-MS systems running in optimal condition?

Michrom Mix
6 bovine
proteins,
digested

50 fmol/uL

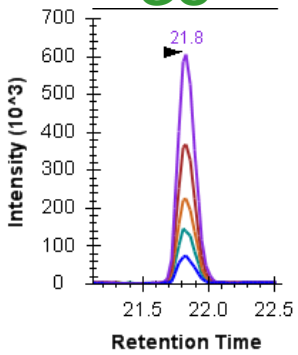
Site 52, ABI 4000 QTRAP
Site 56, ABI 4000 QTRAP
Site 56A, ABI 5500
Site 56B, Agilent 6460 ChipCube
Site 73, ABI 4000 QTRAP
Site 32, ABI 4000 QTRAP
Site 95, ABI 4000 QTRAP
Site 98, ABI 4000 QTRAP
Site 86, ABI 4000 QTRAP
Site 86A, Waters Xevo
Site 90, Agilent 6410 ChipCube
Site 65, Thermo Vantage
Site 54, ABI 4000 QTRAP
Site 19, ABI 4000 QTRAP
Site 19A, Agilent 6410 ChipCube
Site 20, Thermo TSQ Quantum

12 Laboratories
4 MS Vendors
7 MS models
5 LC models

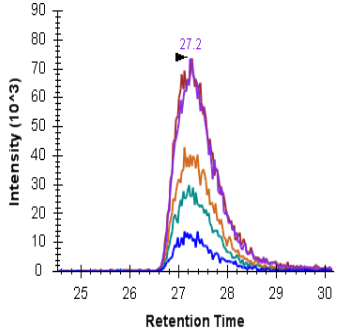
Establish Instrument Specific Ranges for

- o **RT Variability**
- o **Peak Area**
- o **Peak Width**
- o **Carry over**
- o **Column conditioning**

GO



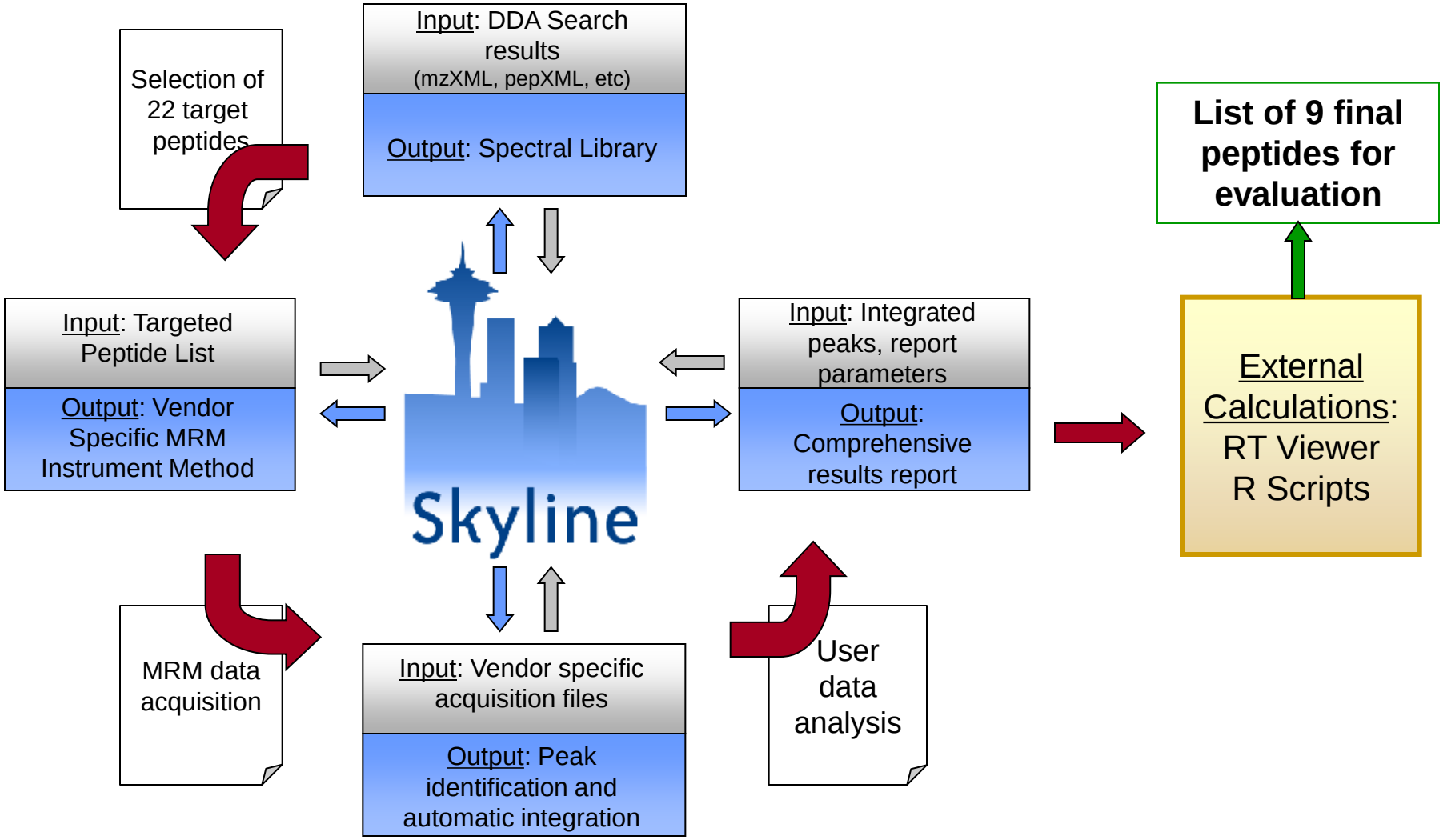
No GO



Define Pass/Fail Criteria

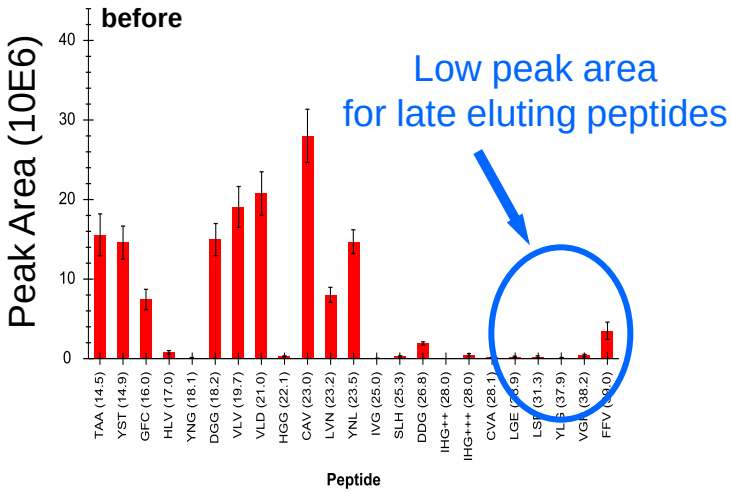
Development of a System Suitability Protocol for Multiple Instrument Platforms

Tools were created to handle workflow and data

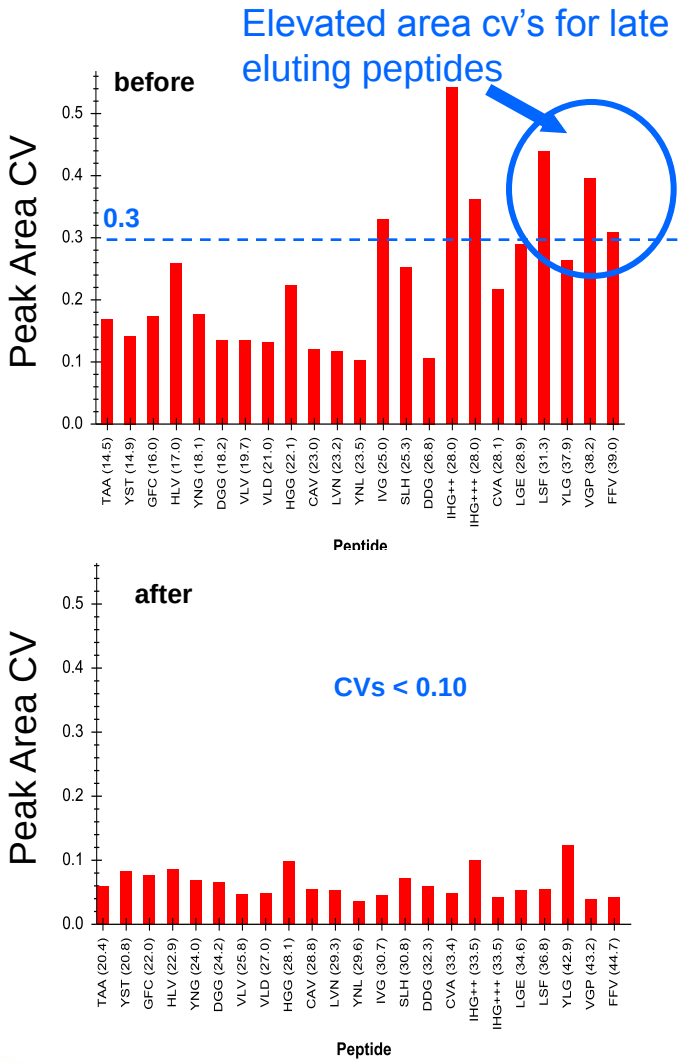


Problems Can Be Visualized Early: Peak Area Stability in Skyline

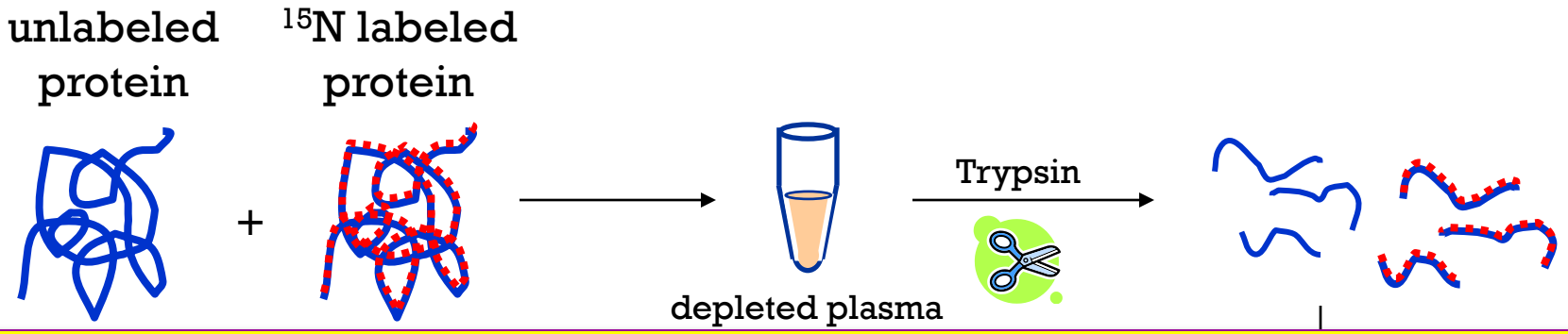
Peak area stability over 10 replicates Site Z



Peak area CV over 10 replicates Site Z



CPTAC VWG Study 9 – Targeting 34 Proteins in Depleted Plasma, 125 Peptide Targets

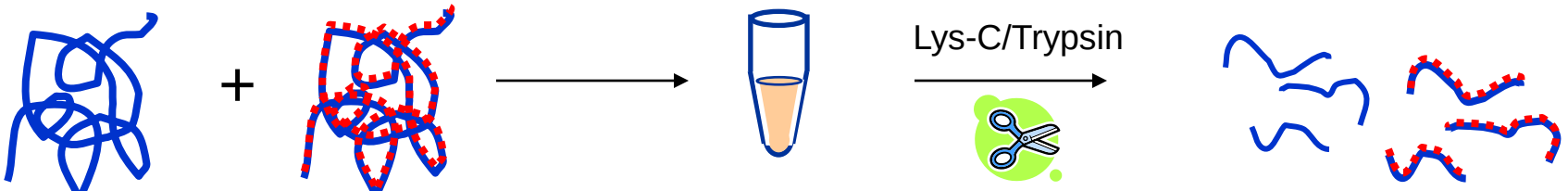


Goals:

- Demonstrate cancer relevancy
- Prove feasibility of > 100-plex (34 proteins) assays in plasma
- Improve LOD and LOQ by depleting abundant proteins
- Demonstrate true quantitative accuracy and evaluate depletion/digestion recovery using heavy labeled proteins
- Conduct blinded verification study to assess accuracy, precision and reproducibility across multiple sites and instrument platforms
- Evaluate system suitability test in context of this large-scale inter-lab study

34 proteins, 1095 transitions, 9 participating sites, 14 instruments, 4 vendors

Peptide and Transition Selection is Streamlined using Skyline



DDA LC-MS/MS, Database search

Skyline

Spectral Library

Selection of 123 target peptides

Skyline

Top 5 Product Ions

MRM-MS Data Acquisition

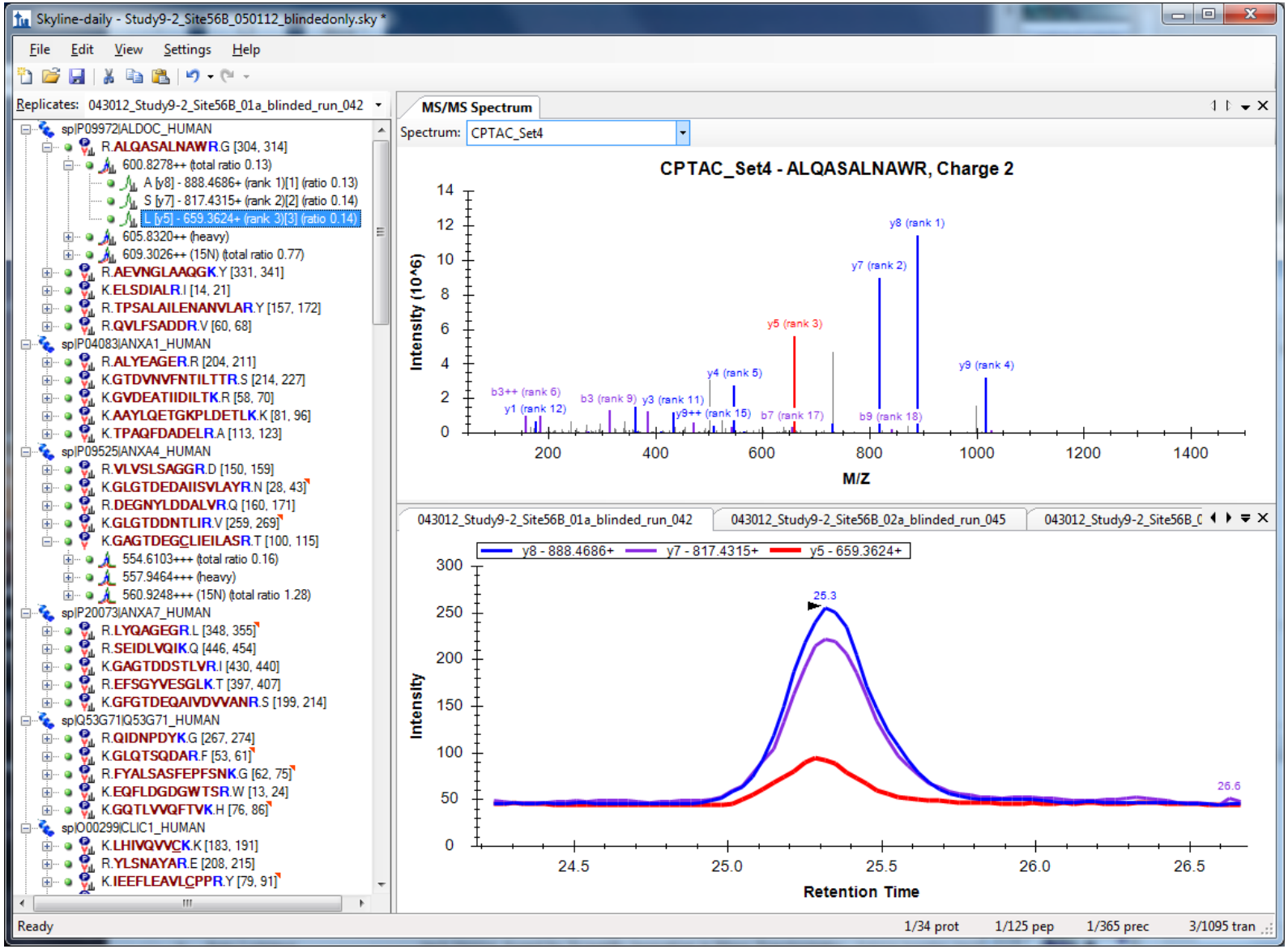
Skyline

- Selection of best 3 ions
- CE Calculation

Final Transition List
L/H: 750 Transitions
L/H/15N: 1095 Transitions

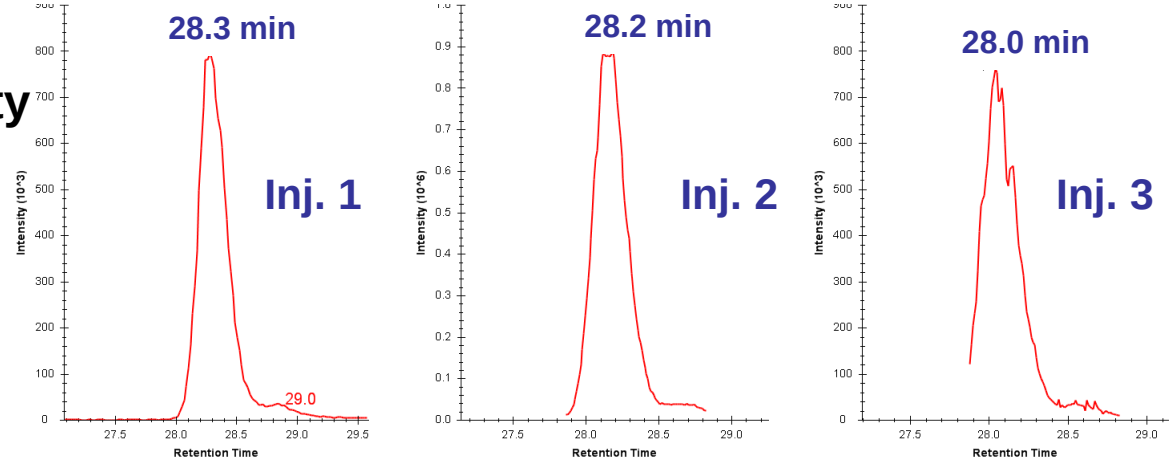
CE publication: B. MacLean et al, 2010, Anal Chem

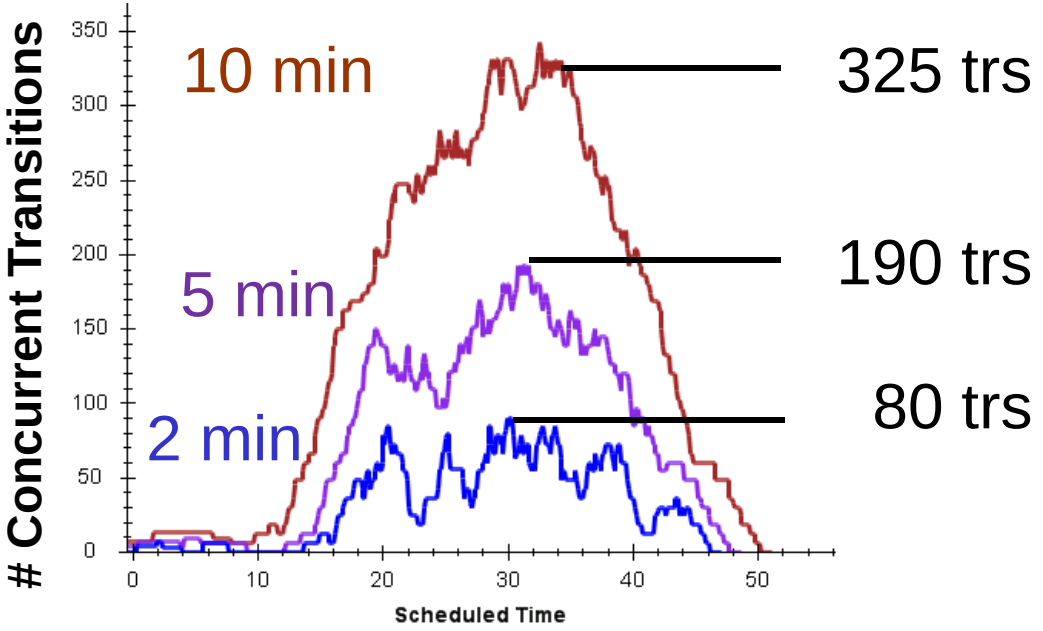
Spectral Libraries Focus Peptide and Transition Selection



Retention Time Scheduling: A Necessity for >100 Transitions

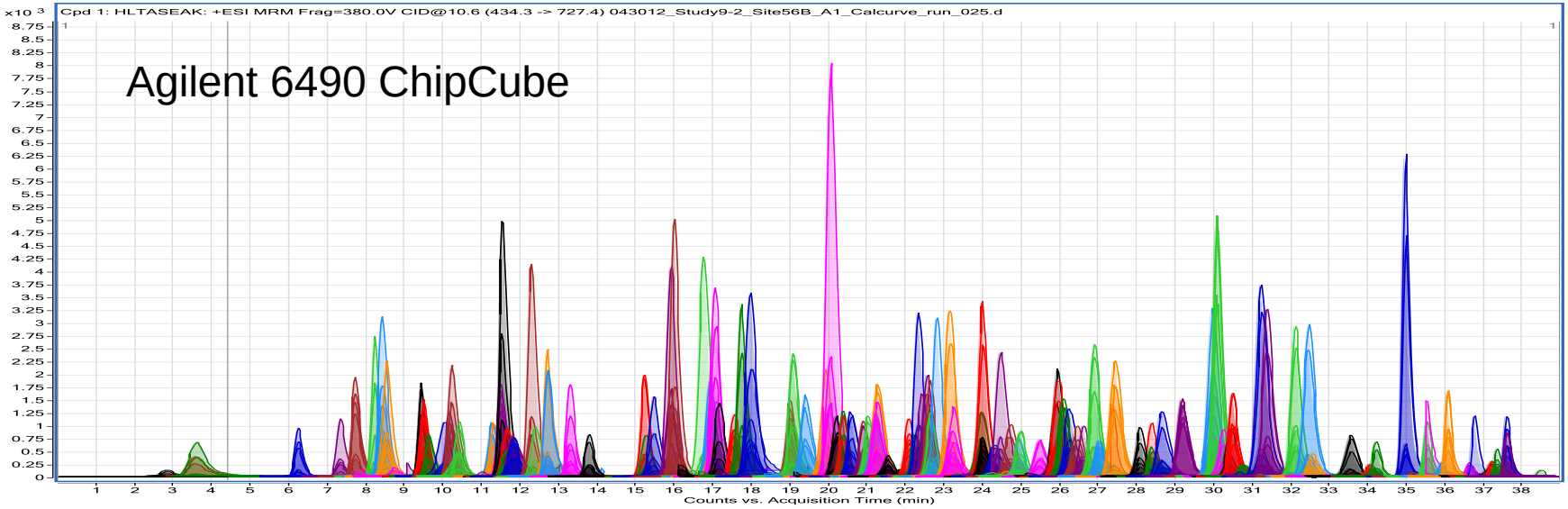
- Scheduling puts rigorous demands on RT reproducibility
- Peak width and RT drift are often limiting factors
- Different peptides shift to various degrees.

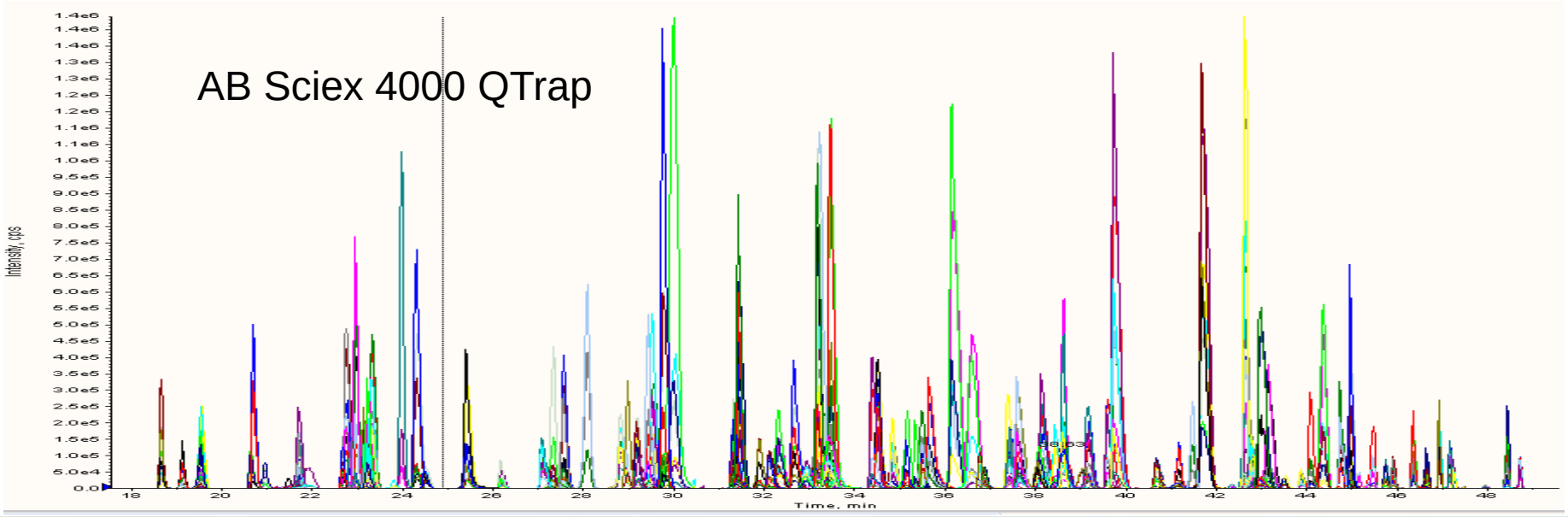




- Large numbers of transitions require narrow RT windows or longer cycle times
- Cycle times may be governed by chromatographic peak width
- Skyline helps gauge number of concurrent transitions based on RT window

Retention Time Scheduling for 1095 Transitions is Challenging – and Different from System to System





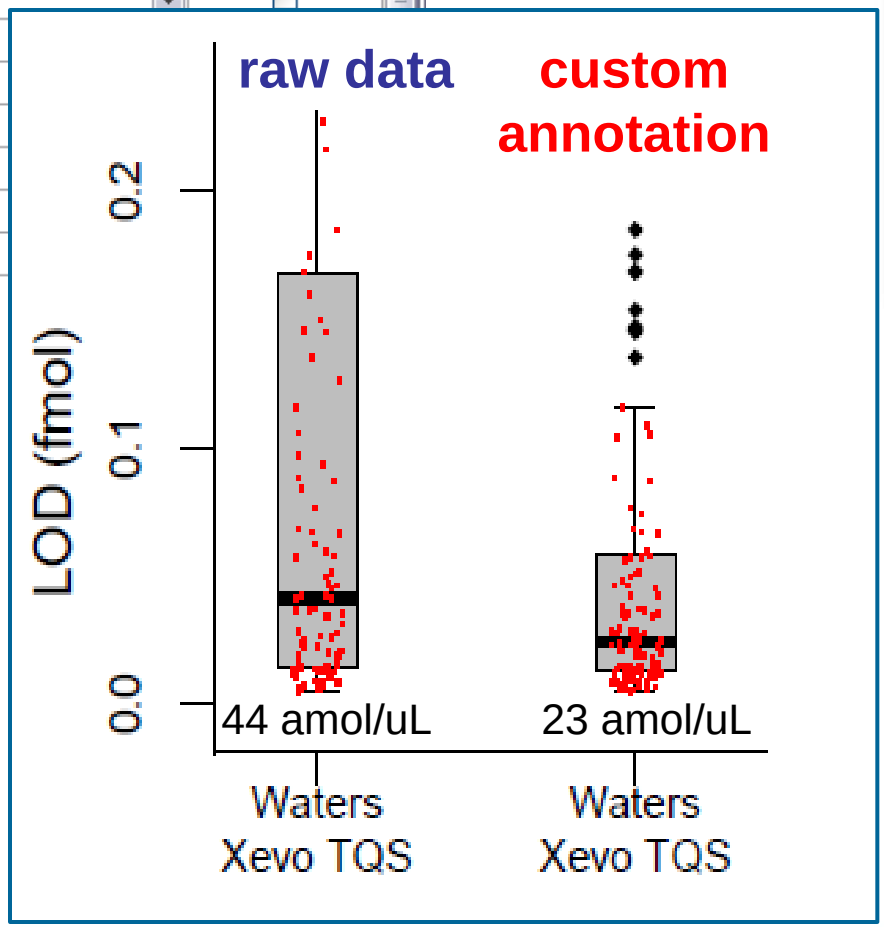
Data Quality Filtering and Custom Annotation by Operators for Data Sets Improves LOD

Results Grid

Total Area	signal quality	RT scheduling problem	general remarks	do not use
110472	peak tailing / poor peak shape			☑
125381				☐
127262		peak outside scheduling window.		☑
138775				☐
138484				
170053				
174726				
192156				
203884				
210323				

very weak signal
peak tailing / poor peak shape
interference
shoulder
transition missing
poor spray
narrow peak (not enough points)

Skyline



Automated version = "AuDIT"

Flags potentially bad transitions

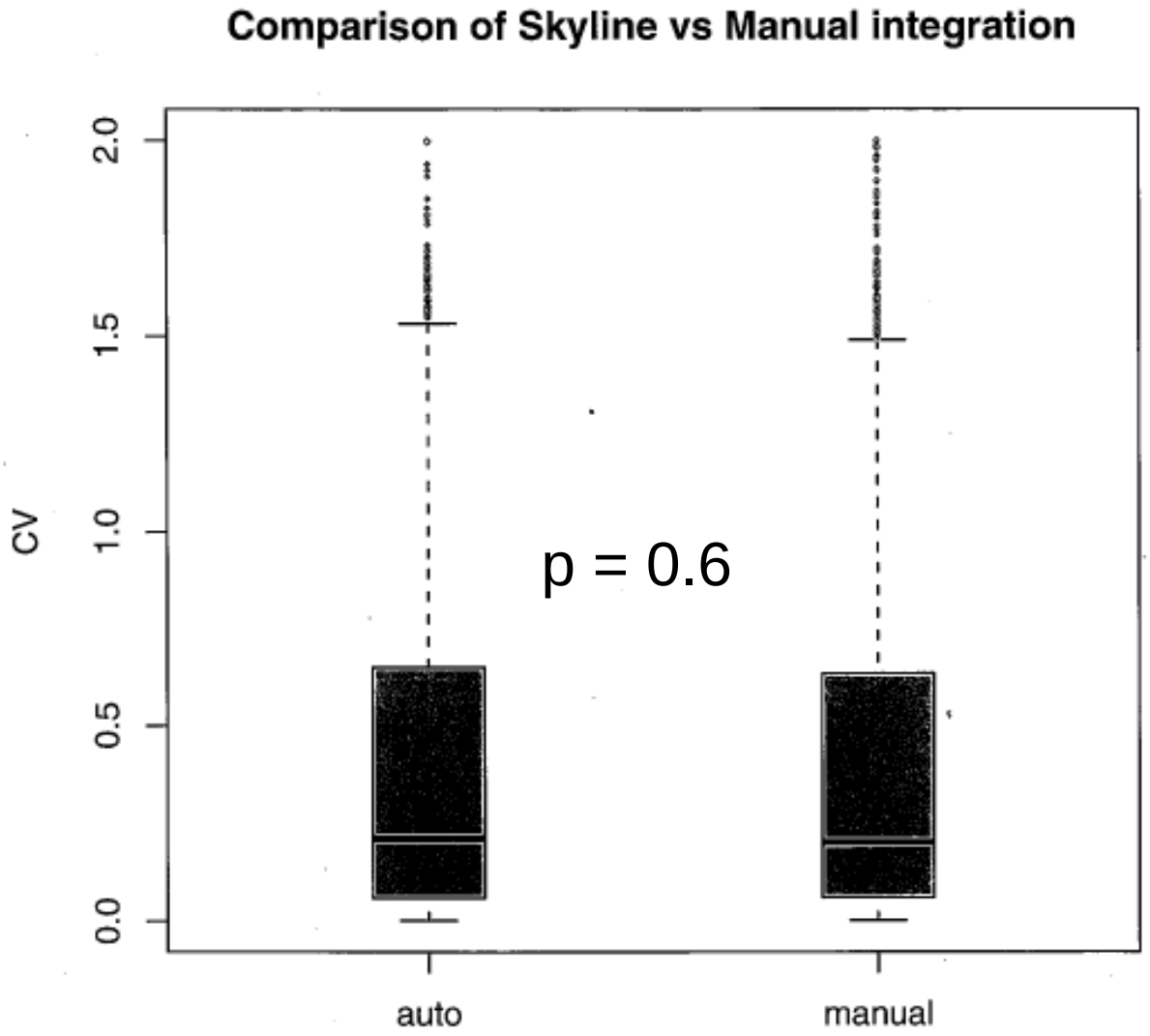
- **poor peak shape**
- **interferences**
- **missing data**

Reduces manual inspection to questionable data

Reduces subjectivity in data analysis

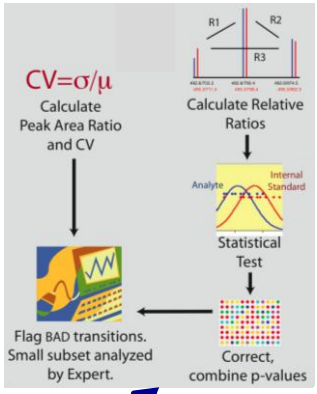
(Abbatiello, Mani et al. Clin. Chem. 2010)

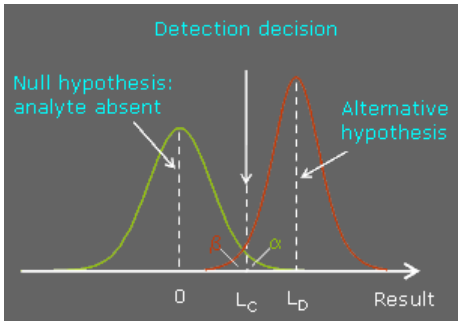
Automatic Integration as Good as Manual Intervention (but takes less time)

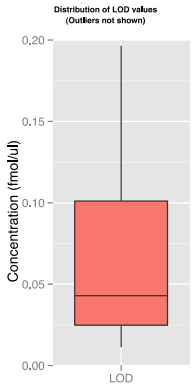


QuaSAR Overview:

Quantitative Statistical Analysis of Reaction Monitoring Results

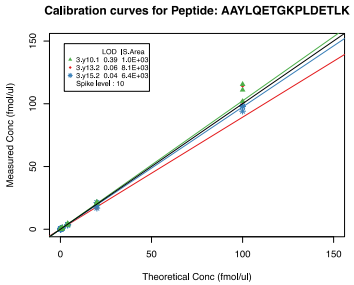


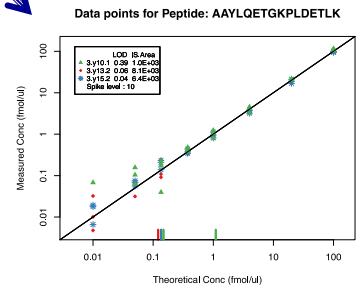




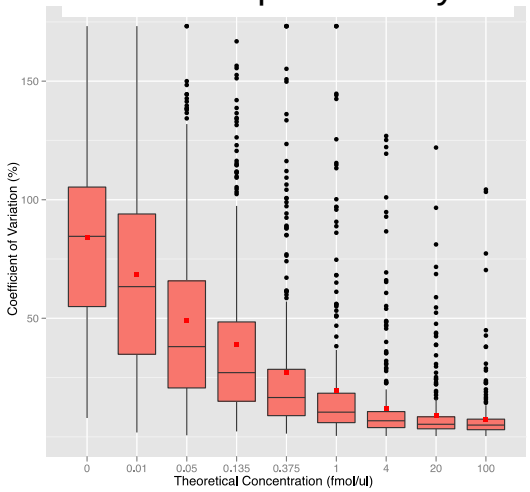
Pre-process & Filter

Concentration & sample info





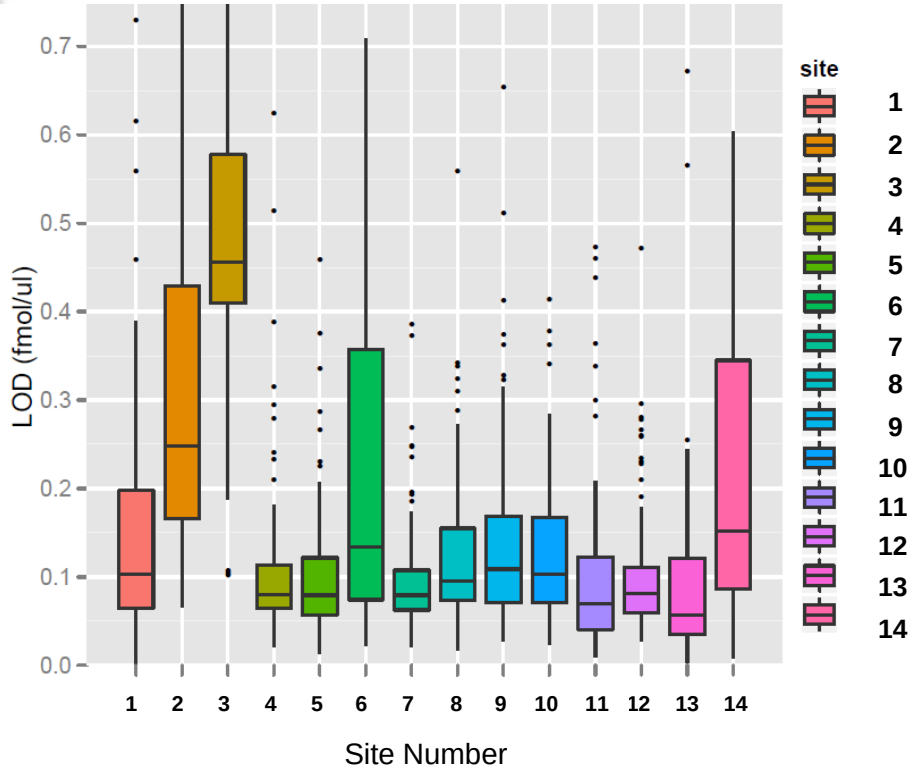
Overall Reproducibility



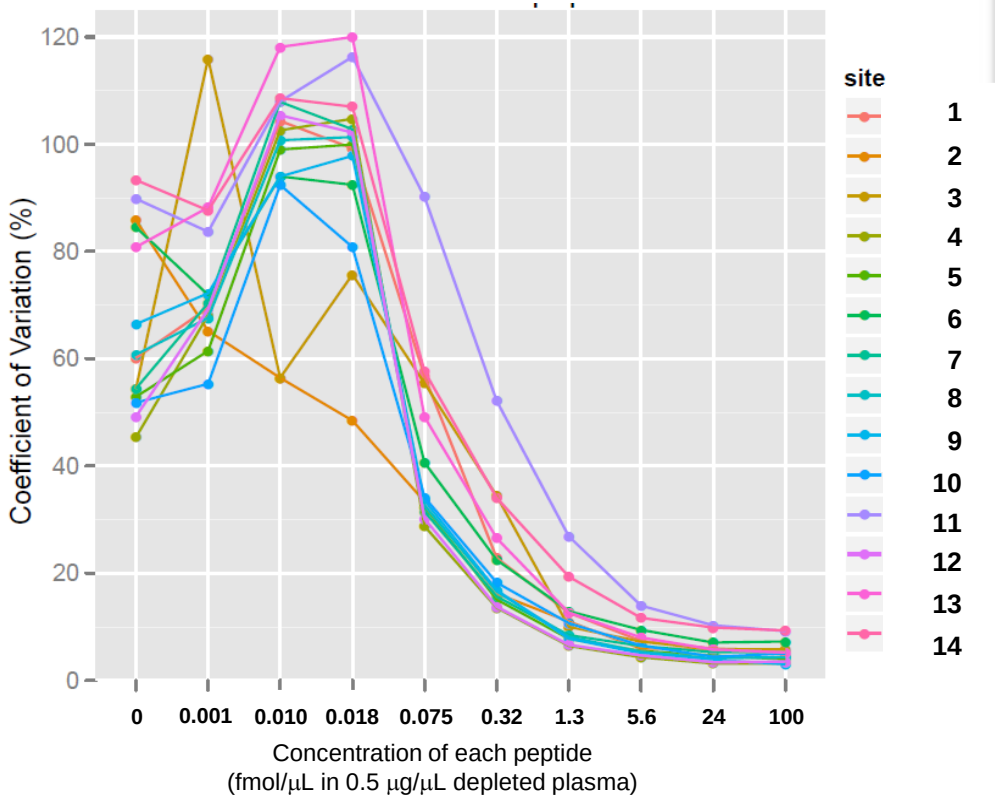
Poster ThP12, #284

Outcome of CPTAC Study 9 is Promising for the Use of Highly Multiplexed SID-MRM-MS Assays

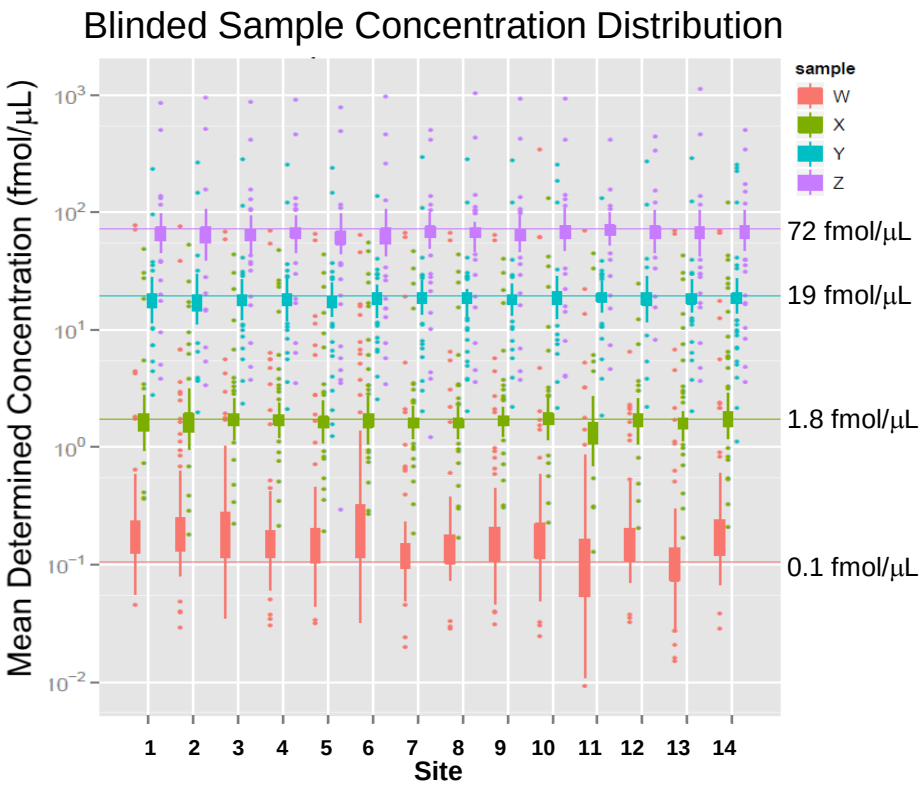
LOD Distribution for all peptides across Sites, Study 9.1

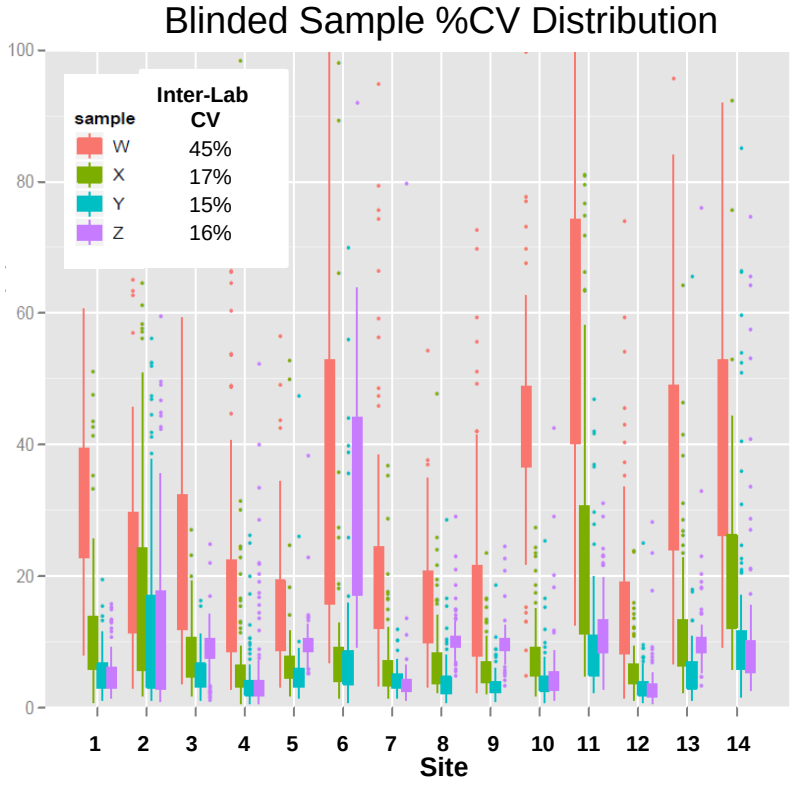


Median CV at each Concentration, Study 9.1



Good Reproducibility and Accuracy is Demonstrated Through Blinded Samples

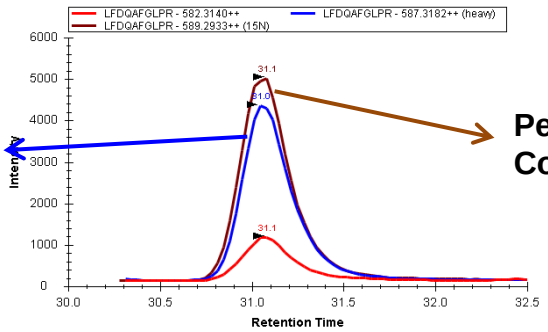




^{15}N Protein Standards Improve Quantitative Accuracy

^{13}C/^{15}N Peptide Internal Standards

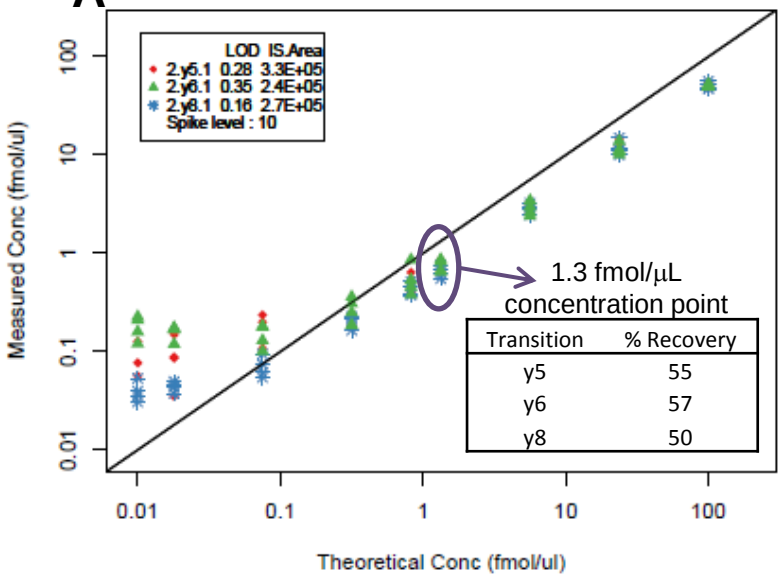
$$\text{Peptide Conc}_{(13C/15N)} = \frac{\text{Light Peak Area}}{^{13}\text{C}/^{15}\text{N Peak Area}} \times \frac{10\ \text{fmol}}{\mu\text{L}}$$



^{15}N Protein Internal Standards

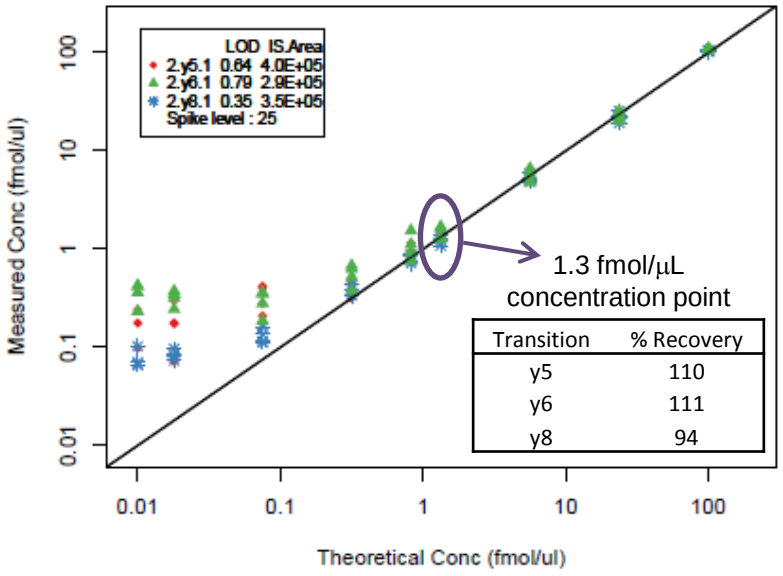
$$\text{Peptide Conc}_{(15N)} = \frac{\text{Light Peak Area}}{^{15}\text{N Peak Area}} \times \frac{25\ \text{fmol}}{\mu\text{L}}$$

A



Transition	% Recovery
y5	55
y6	57
y8	50

B

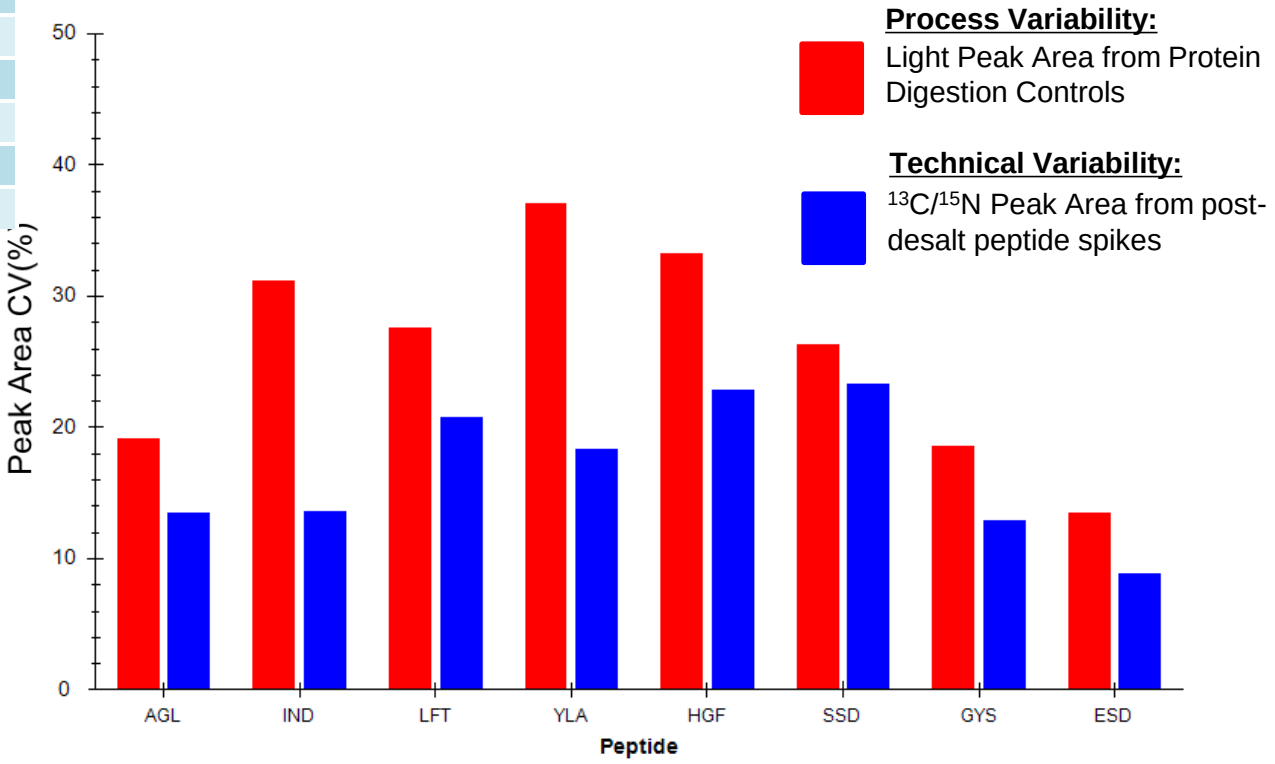


Transition	% Recovery
y5	110
y6	111
y8	94

Protein Digestion Controls Help Gauge Assay Variability

Control Proteins	# Peptide Targets
Aprotinin	1
C-reactive protein	2
Horseradish peroxidase	1
Leptin	1
Myelin basic protein	2
Myoglobin	1

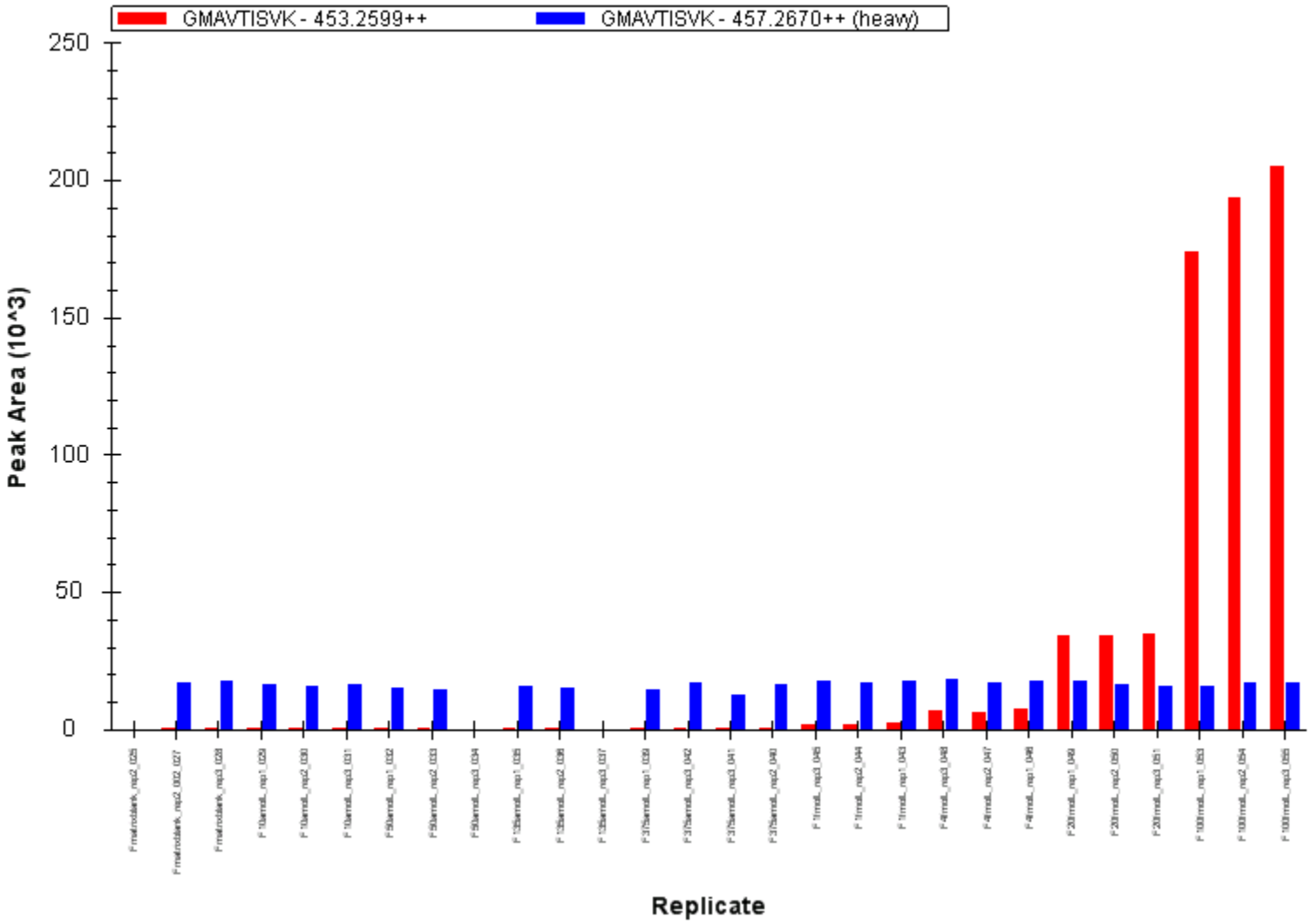
Technical and Process Variability Assessed From Digestion Controls and SIS Peptide Spikes



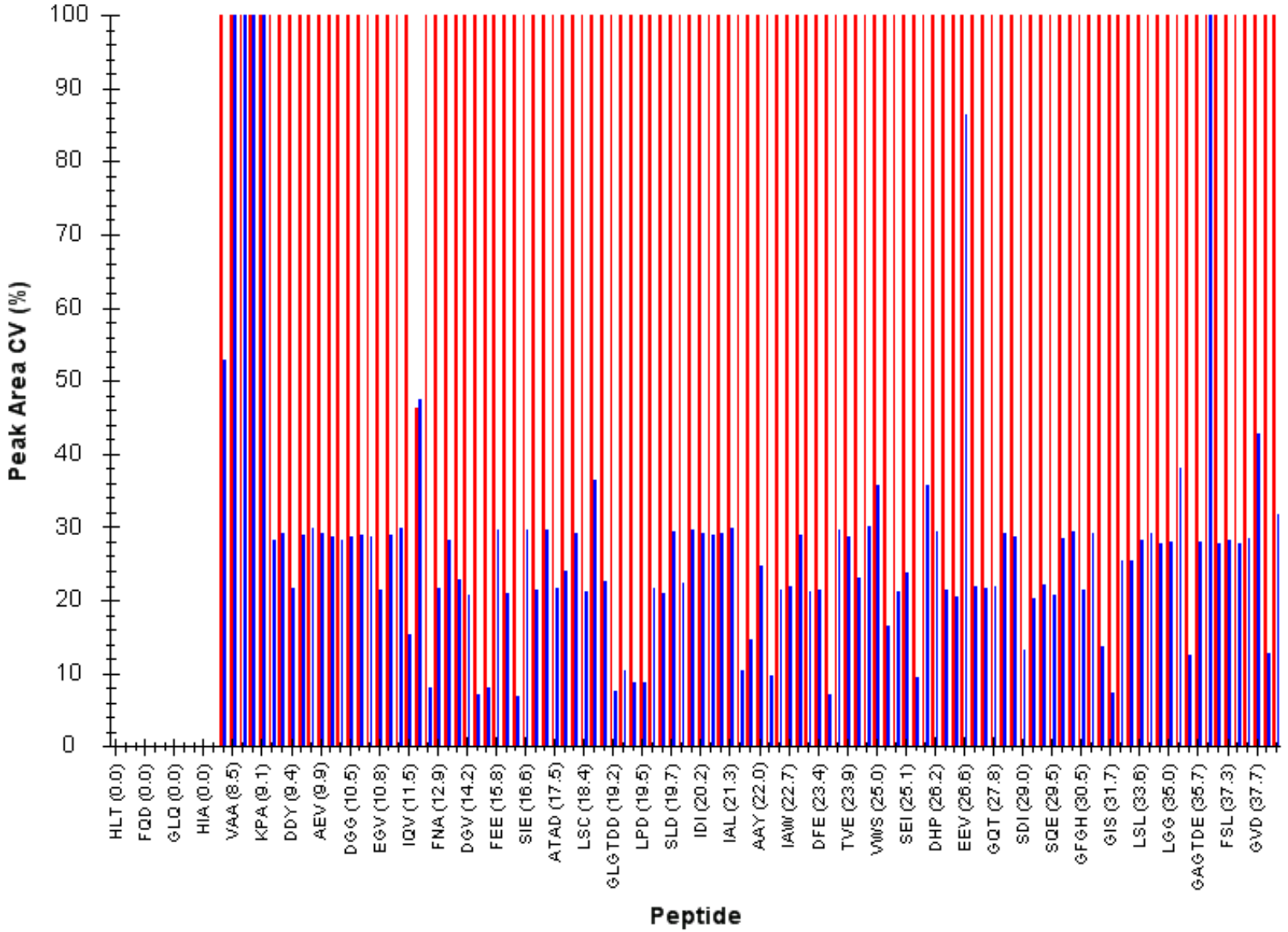
Poster MP01, #004

Skyline Facilitates Rapid Data Analysis Through Overview Plots

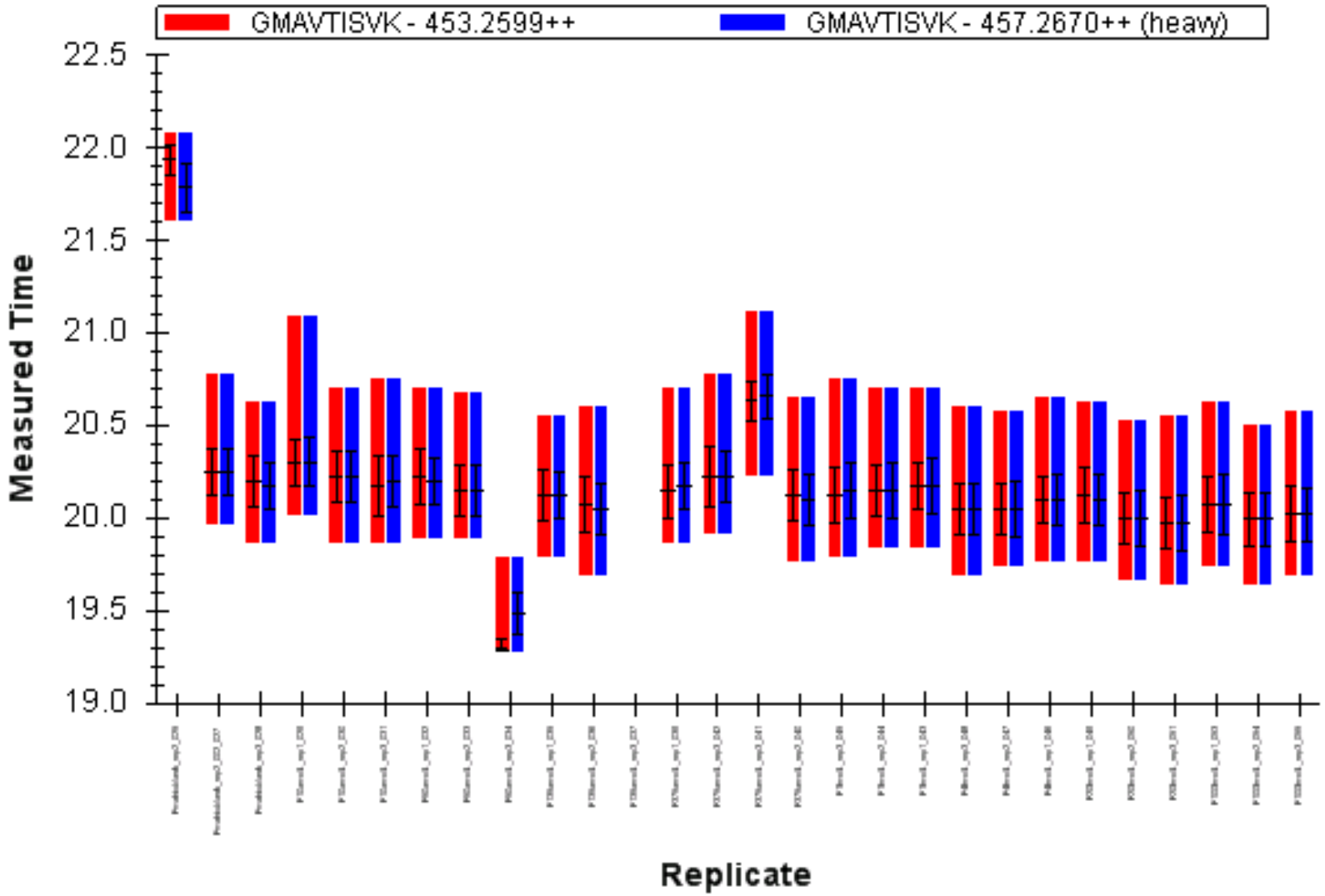
Peak Area Replicate View, Light and Heavy



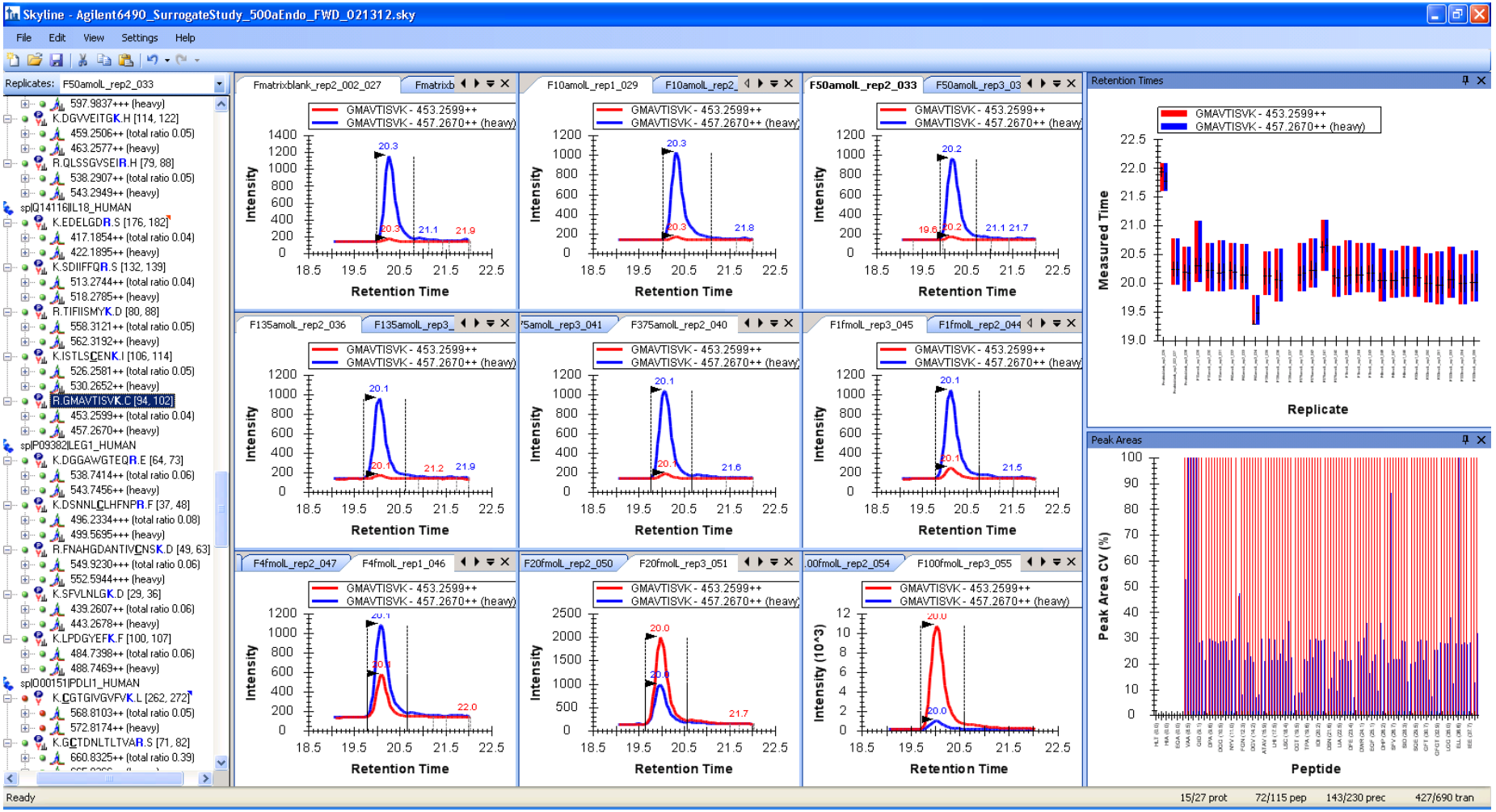
Peak Area CV Plots Provide Quick Assessment of Reproducibility Across a Series of Samples



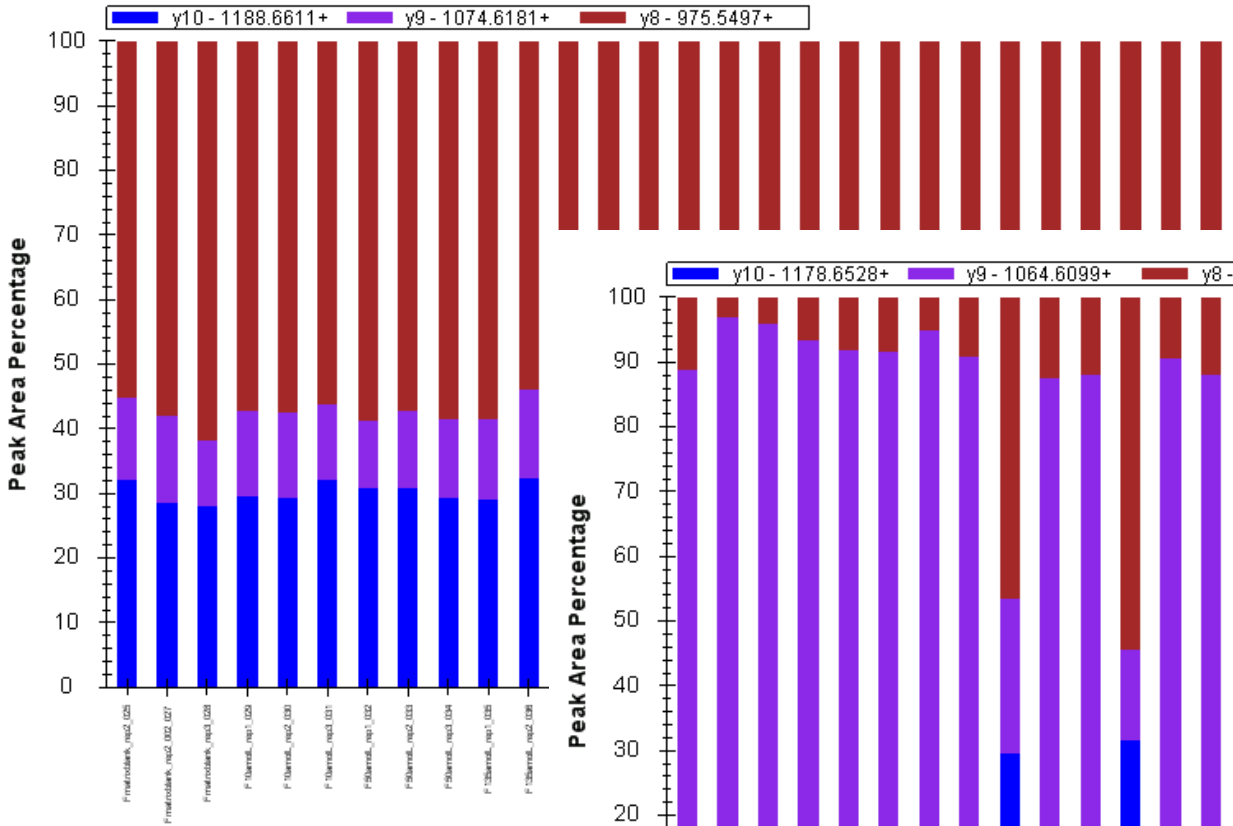
Retention Time Reproducibility Plots Show Trends in Retention Time



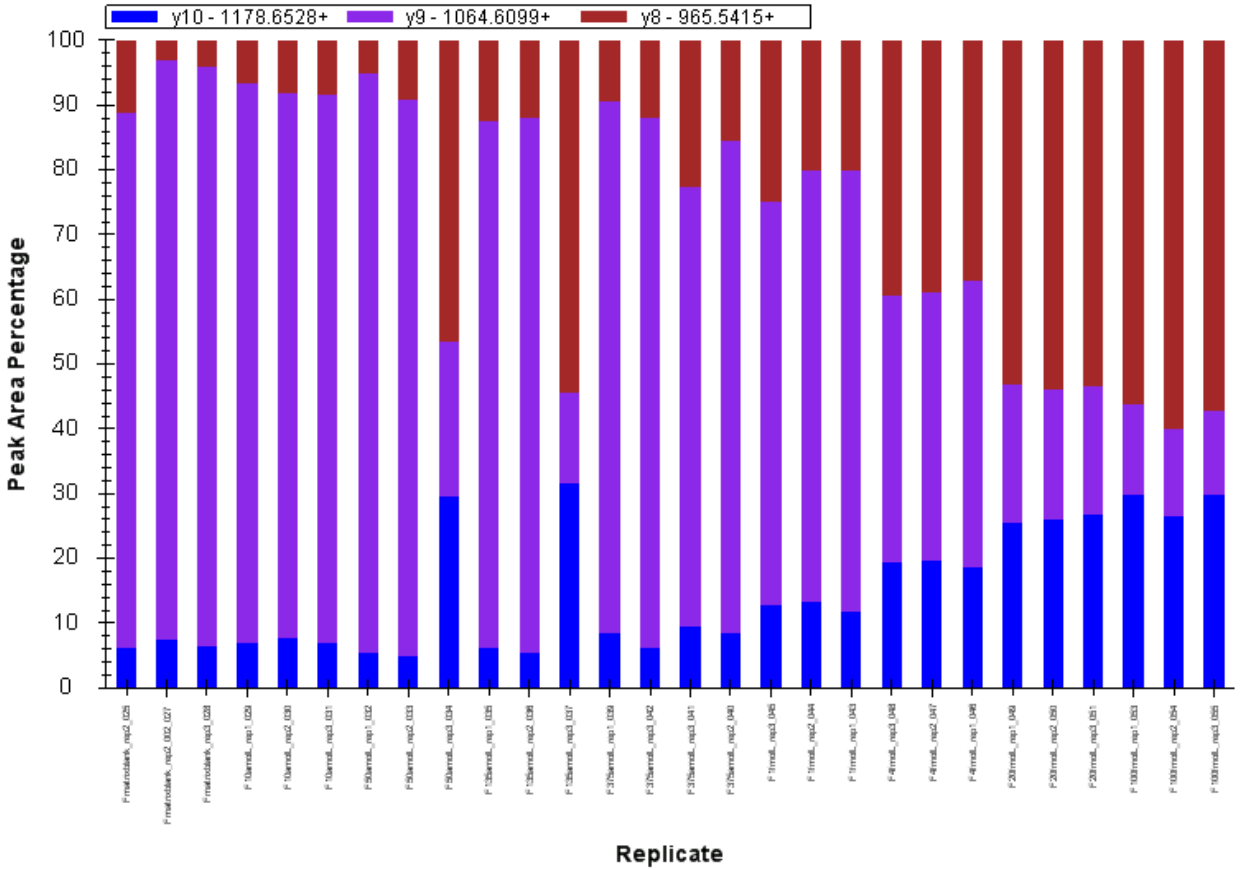
Quick View of All Replicates



Interference Visualization



Heavy Peptide Transitions



Light Peptide Transitions

Summary

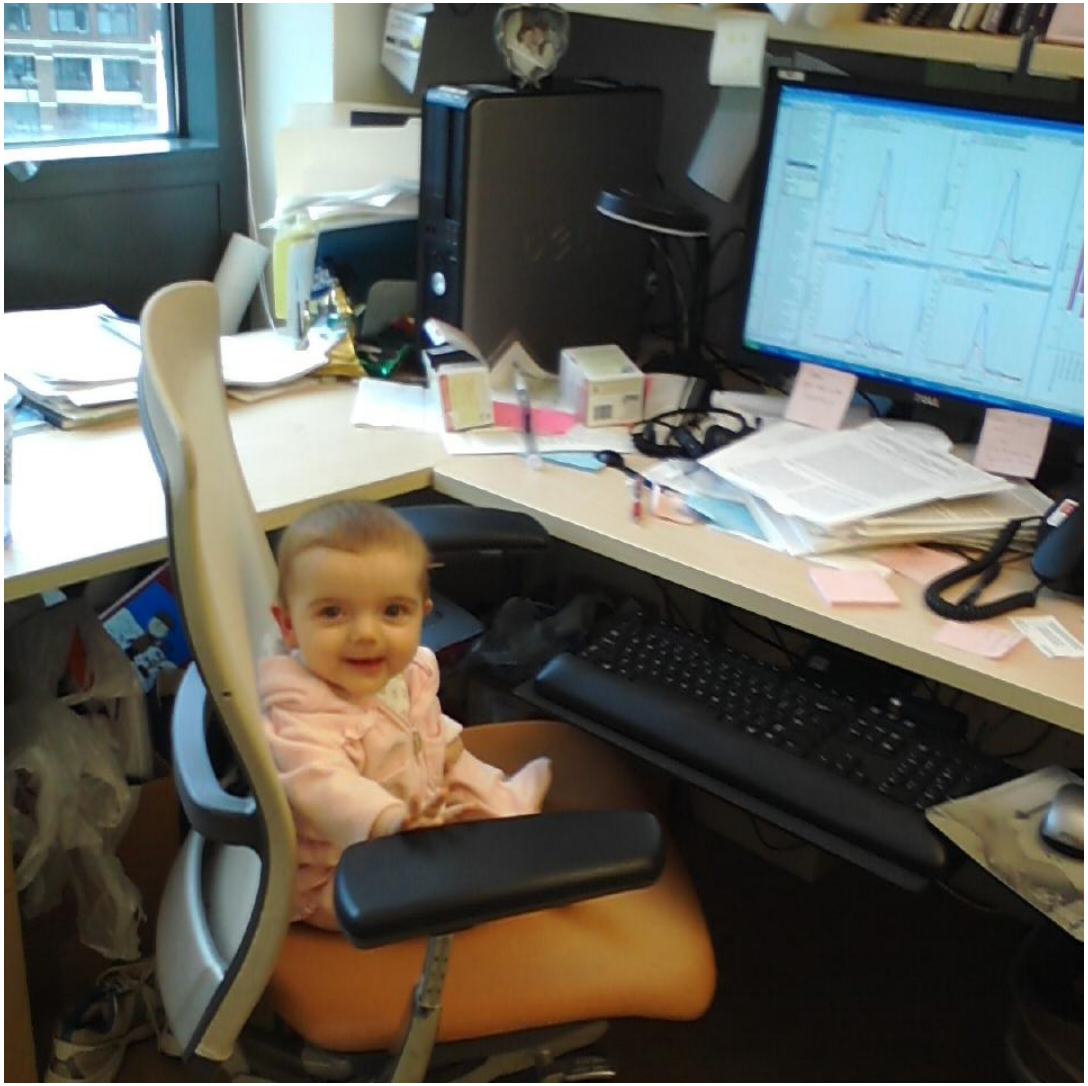
- First large-scale interlab study to include 15N protein reagents and >100 peptide targets (>350 peptide forms)
- Sensitivity improvement from previous study by using depleted plasma, adjusting the gradient
- Transition selection and MS method transfer across 4 instrument platforms facilitated through Skyline
- Peak Area and Retention Time views help quickly assess data quality
- Customizable reports from Skyline enable down-stream processing, helps remove subjectivity of data evaluation, and increases data analysis throughput
- Skyline helps maintain objective processing of data, requiring less manual tweaking
- It’s free, it is easy, and it will process your data

CPTAC VWG Participants & Acknowledgements

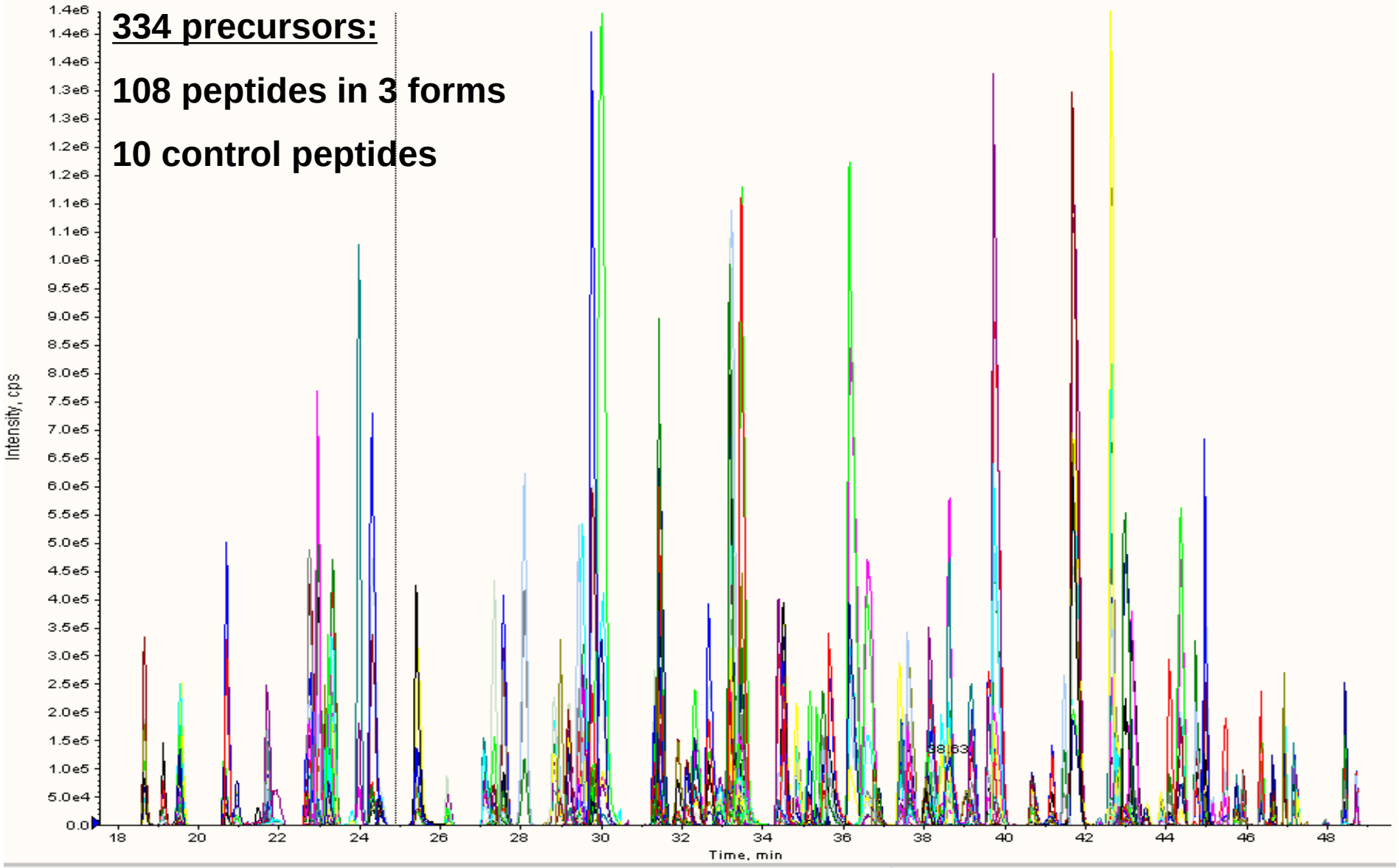
Broad Institute: Susan Abbatiello, Terri Addona, Steven A. Carr, Hasmik Keshishian, D.R. Mani, Michael Burgess, James Markell
Buck Institute for Age Research: Michael P. Cusack, Bradford W. Gibson. Jason M. Held, **Birgit Schilling**
Fred Hutchinson Cancer Research Center: Amanda G. Paulovich, Jeffrey R. Whiteaker, Shucha Zhang
Indiana University: Mu Wang, Jong-Won Kim, Jimsan You
Massachusetts General Hospital: Steven J. Skates
Memorial Sloan-Kettering Cancer Center: Paul Tempst, Mousumi Ghosh
National Cancer Institute: Emily Boja Tara Hiltke, Christopher Kinsinger, Mehdi Mesri, Henry Rodriguez, Robert Rivers
NISS: Xingdong Feng, Nell Sedransk, Jessie Xia
NIST: Paul Rudnick
New York University: Thomas A. Neubert, Åsa Wahlander, Sofia Waldemarson, Pawel Sadowski, John Lyssand
Plasma Proteome Institute: N. Leigh Anderson
Purdue University: Charles Buck, Fred Regnier, Dorota Inerowicz, Vicki Hedrick
University of California, San Francisco: Simon Allen, Susan J. Fisher, **Steven C. Hall**,
University of North Carolina: David Ransohof
University of Victoria: Christoph H. Borchers, Angela Jackson, Derek Smith
University of Washington: Michael MacCoss, Brendan MacLean, Daniela Tomazela
Vanderbilt University: Daniel Liebler, Kent Shaddox, Corbin Whitwell, Lisa Zimmerman

Funding: National Cancer Institute

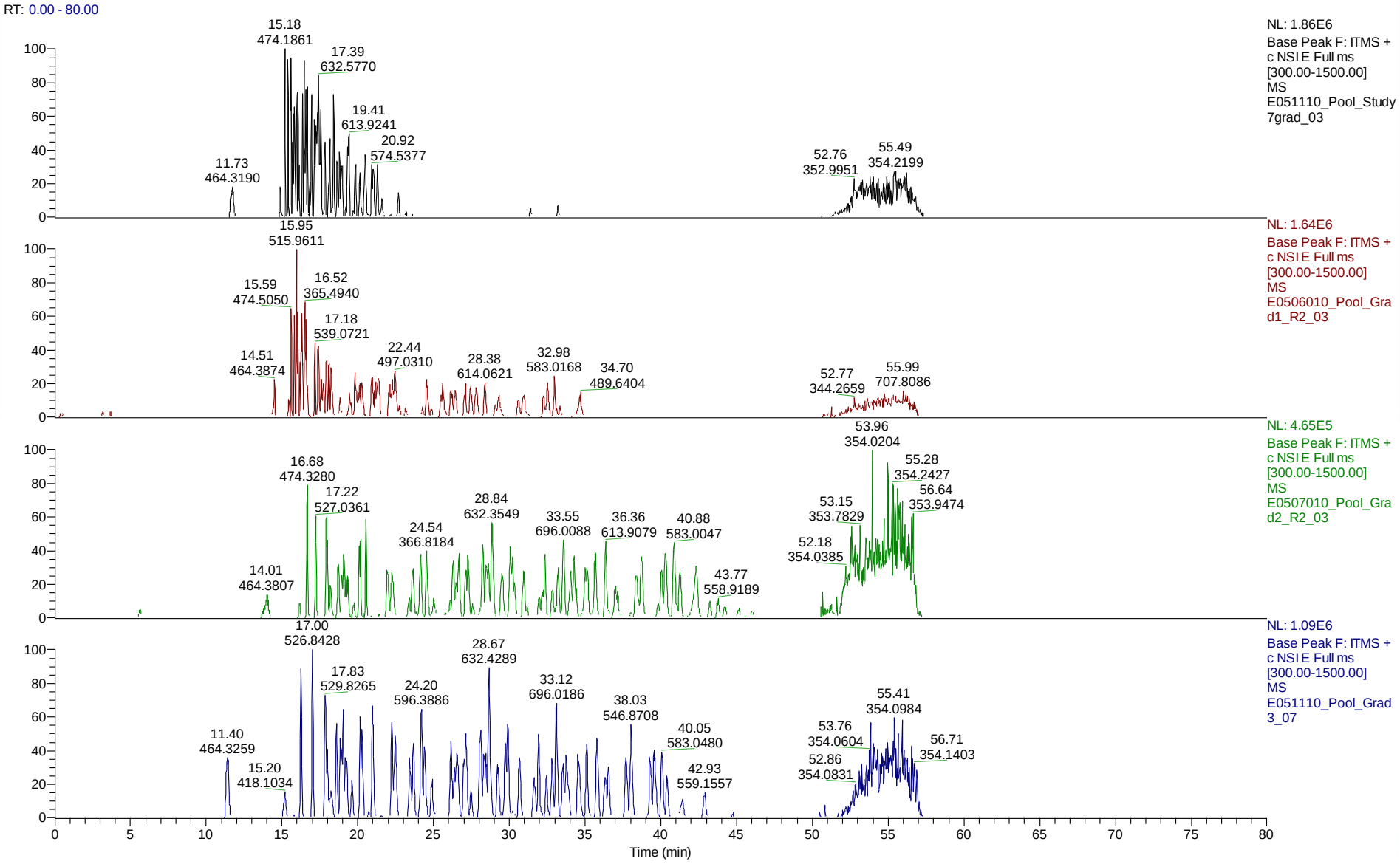
Skyline…
So easy a baby can do it



1000 Q1/Q3 Pairs – AB Sciex 4000 QTRAP



Gradient Optimization will Improve Sensitivity and Data Acquisition



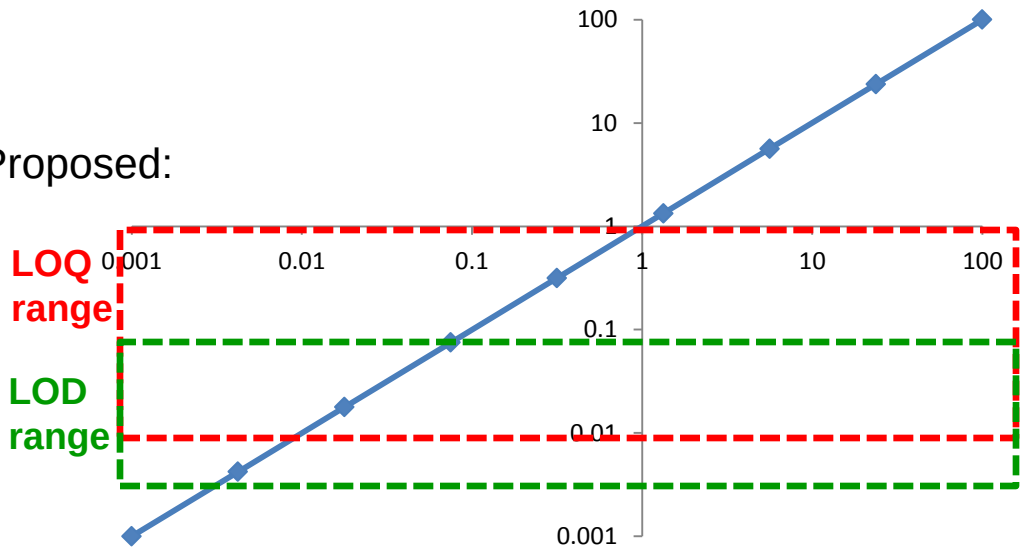
LOD/LOQ Calculations: How Many Points in the Curve are Needed?

What is the ideal concentration range?

$$LOD = \bar{s}_{blank} + t_{0.95} \times (\sigma_{blank} + \sigma_{low})/\sqrt{n}$$

(fmol/µL)
250
113
51
23
10
4.6
2.0
0.9
0.42
0.19
0.09
0.04
0.017
0.008
0.004
0.002

Proposed:

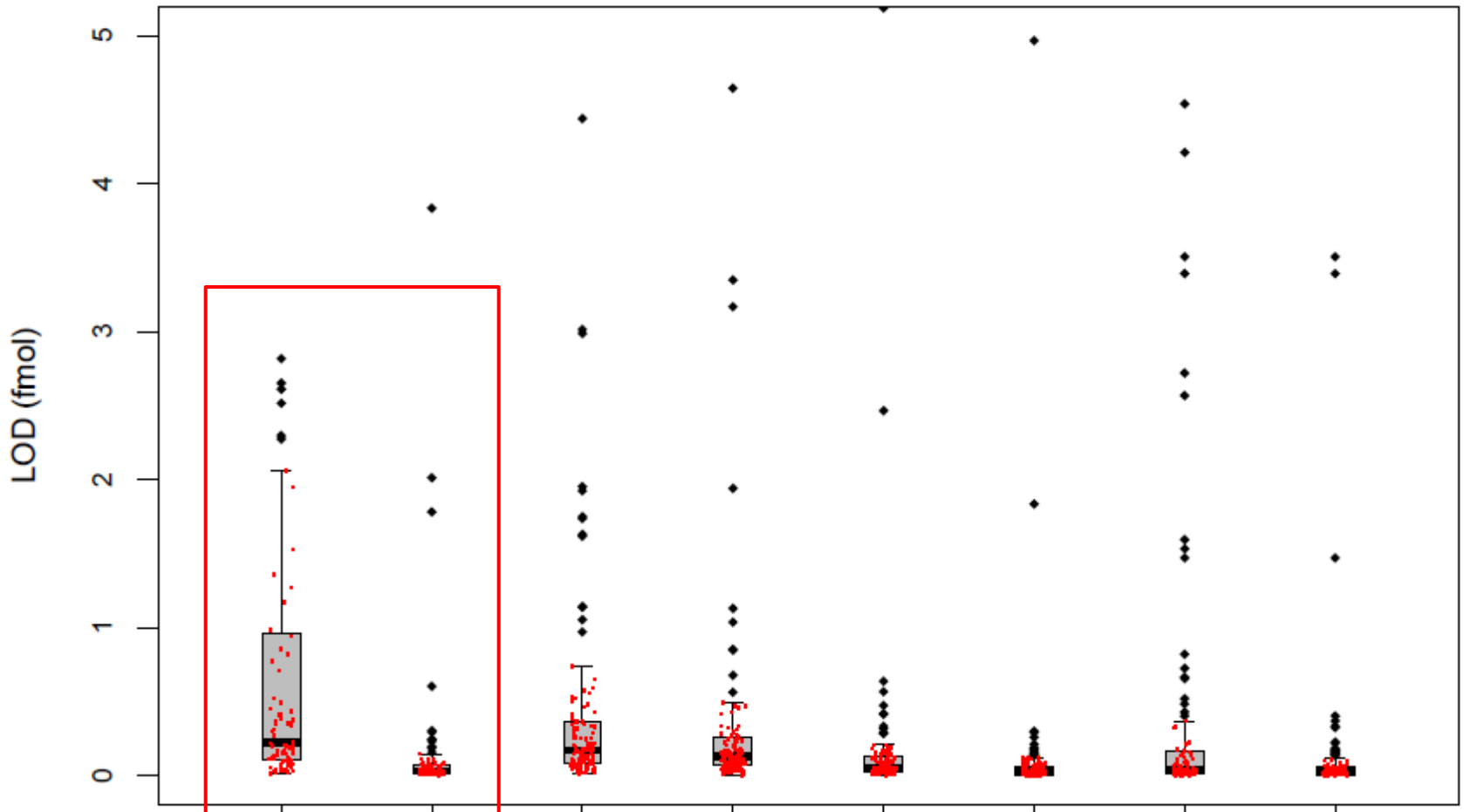


- Generate preliminary curves (16 pts)
- Pick a range and number of points to cover most peptides

Linnet & Kondratovich, (2004) Clin Chem

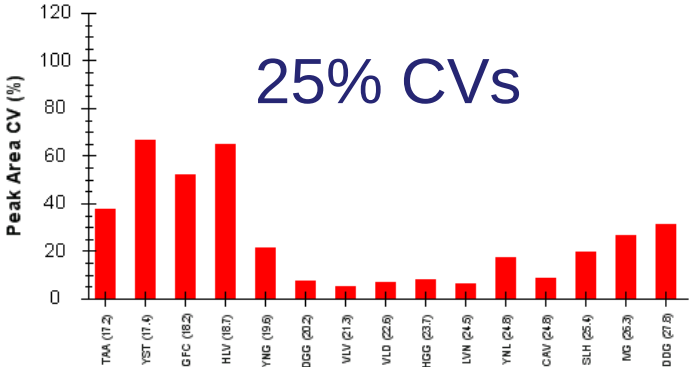
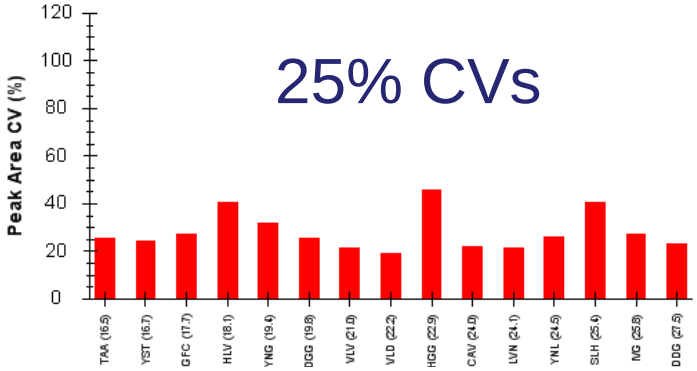
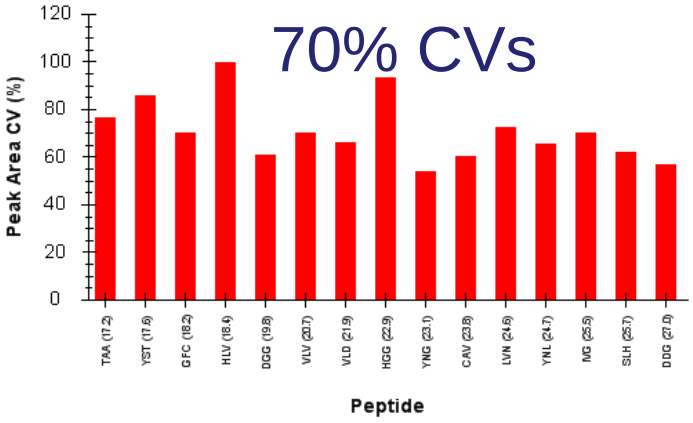
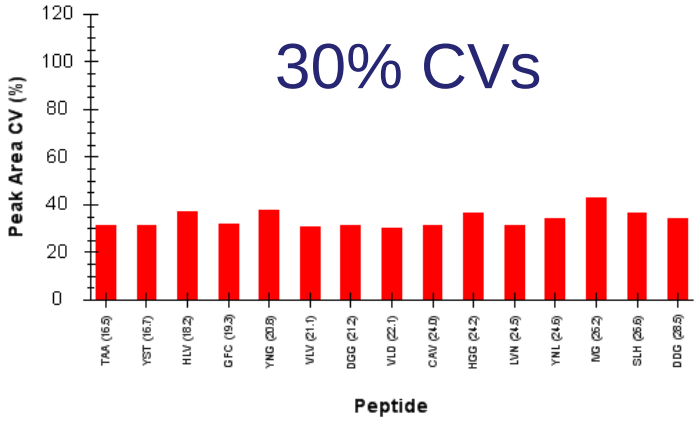
Keshishian et al, (2009) MCP

16 Point Curve at Selected CPTAC Sites Shows Good Reproducibility and Sensitivity



	AB Sciex 5500	AB Sciex 4000	Agilent 6410	Agilent 6460	Thermo Vantage	Waters Xevo	Waters Xevo TQS	Waters Xevo TQS
Median LOD (fmol/uL)	0.22	0.032	0.17	0.13	0.055	0.027	0.044	0.023
Outliers	15	16	15	15	15	12	23	19

LOD is Highly Dependent Upon System Performance: Chromatography and Ionization

	QTRAP 5500 Median LOD: 220 amol	4000 QTRAP Median LOD: 32 amol
Pre-Assay System Suitability Runs (5)		
Throughout Assay System Suitability Runs (24)		

Unstable ESI was a major factor in poor detection and reproducibility

System Suitability assessment detects poor system performance