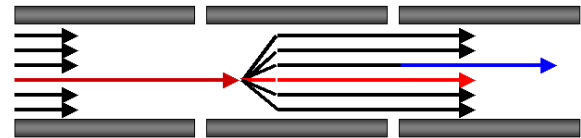# Skyline

Targeted Proteomics Environment

Signal processing for quantitative proteomics

Brendan MacLean
MacCoss Lab

UNIVERSITY *of* WASHINGTON

Spectrum-based Quantification

▸ Uses peptide spectrum matches already calculated

▸ Spectral counting
  ▸ Statistically weak
  ▸ Protein quantification – not for PTM studies
▸ Isobaric tags
  ▸ Measure many conditions at once
  ▸ Statistically better than spectral counting
  ▸ Requires complex sample prep

▸ Not possible to measure absence

▸

Chromatography-based Quantification

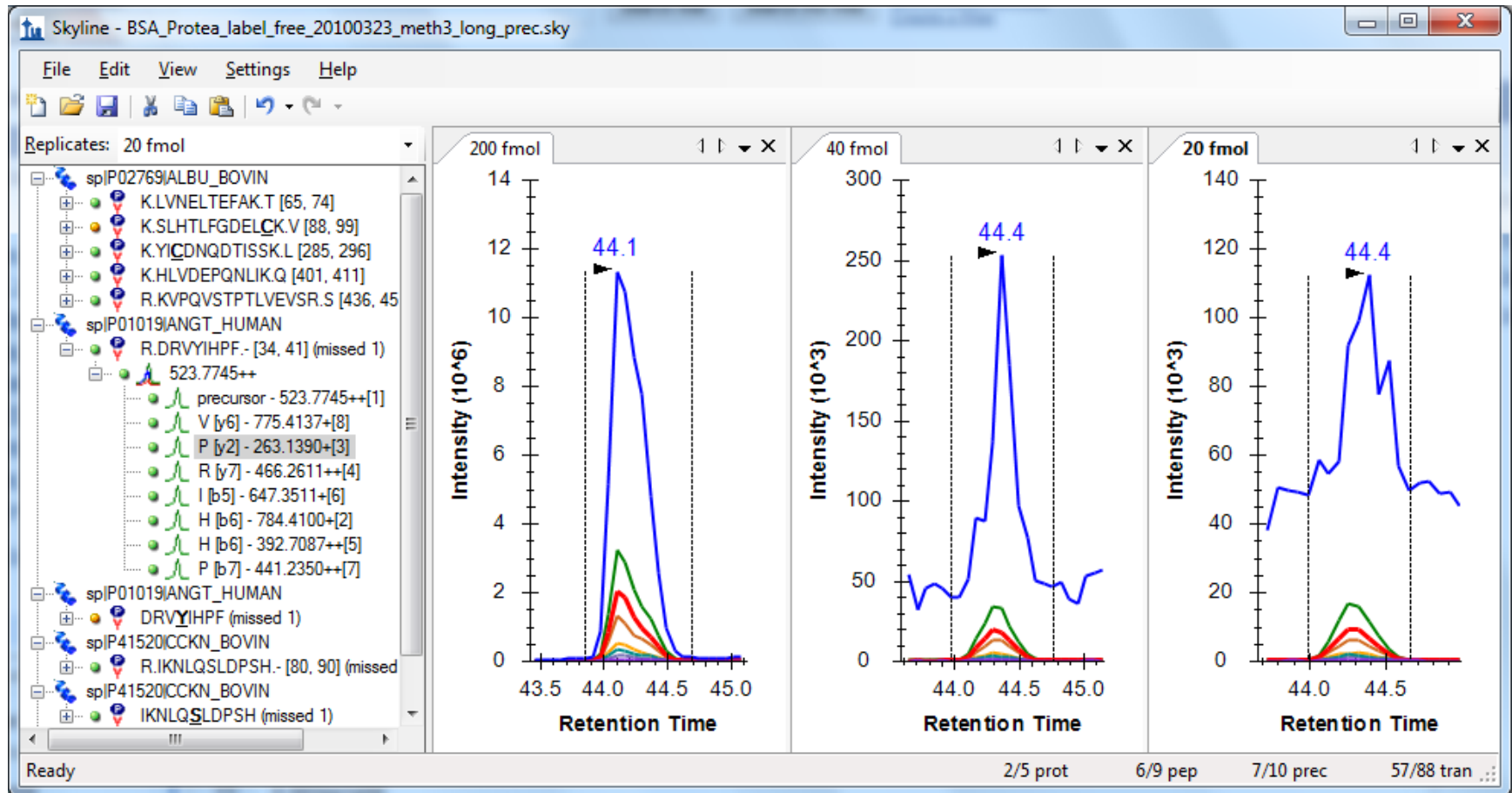▸ Requires new tools and data processing

▸ **SRM**

▸ MS1 chromatogram extraction

▸ Targeted MS/MS

▸ Data independent acquisition (DIA/SWATH)

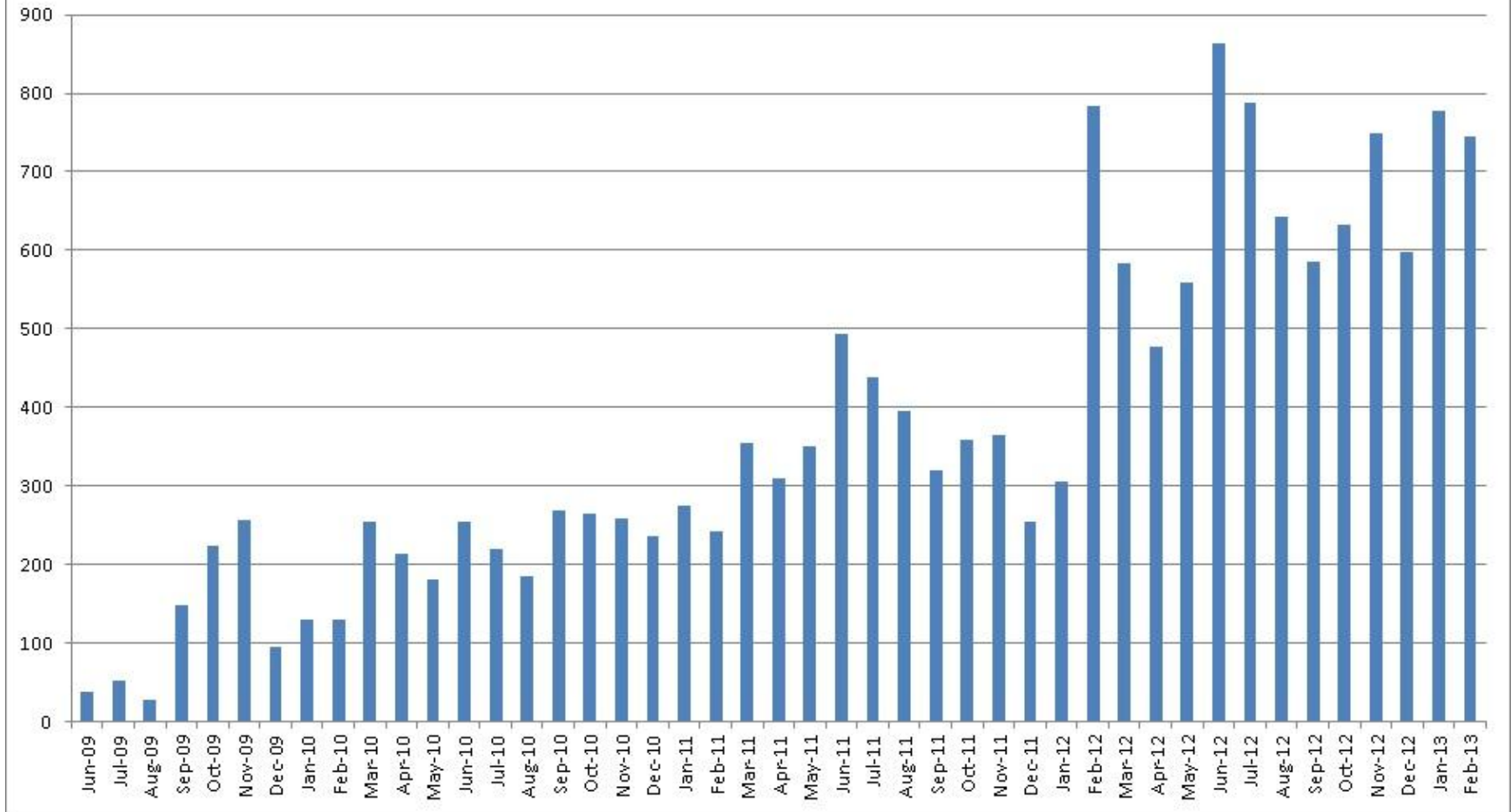| Acquisition	Targeted	Survey
Selective	Targeted-MS/MS	DIA
Sensitive	SRM	MS1

When Selectivity Becomes Sensitivity

Over 16,000 Skyline Installations

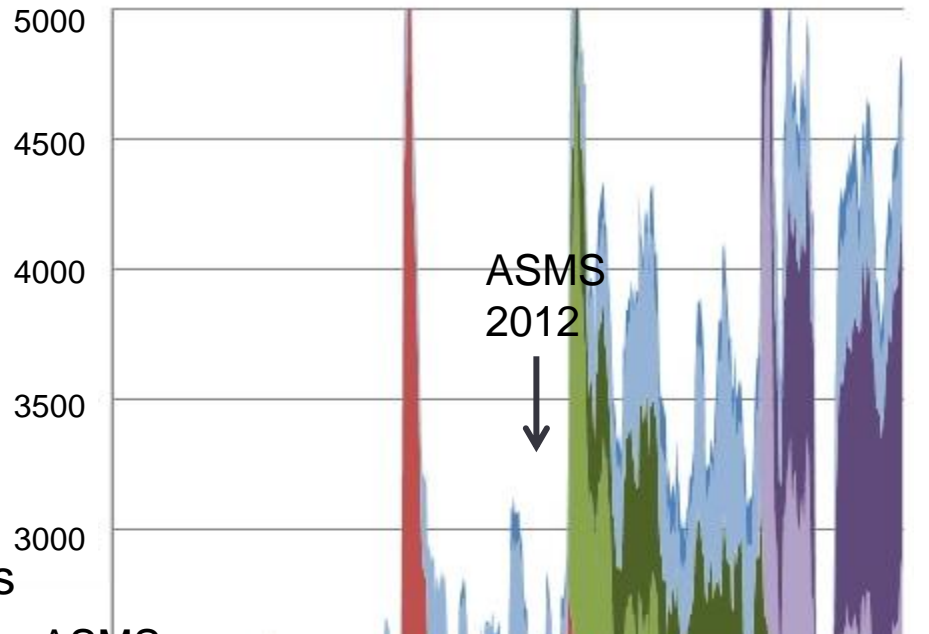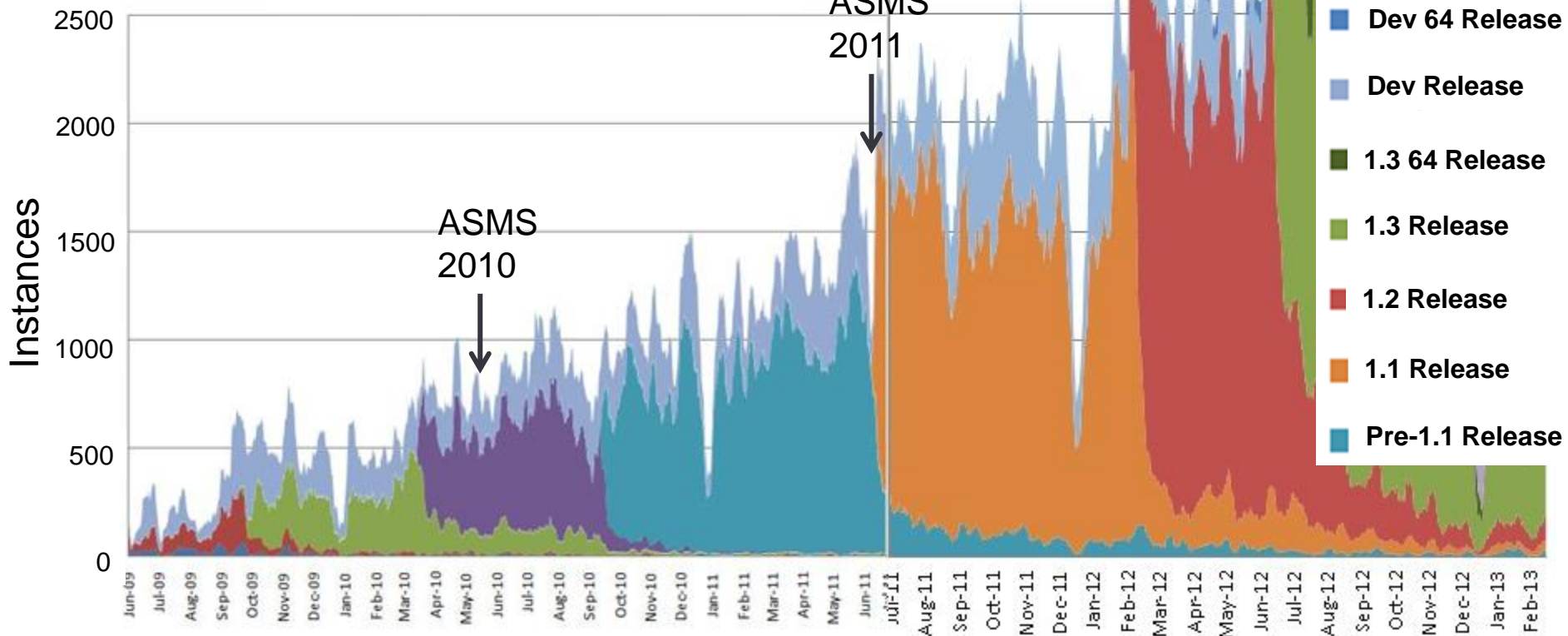## New Skyline Installations by Month

>150 citations

Skyline Use
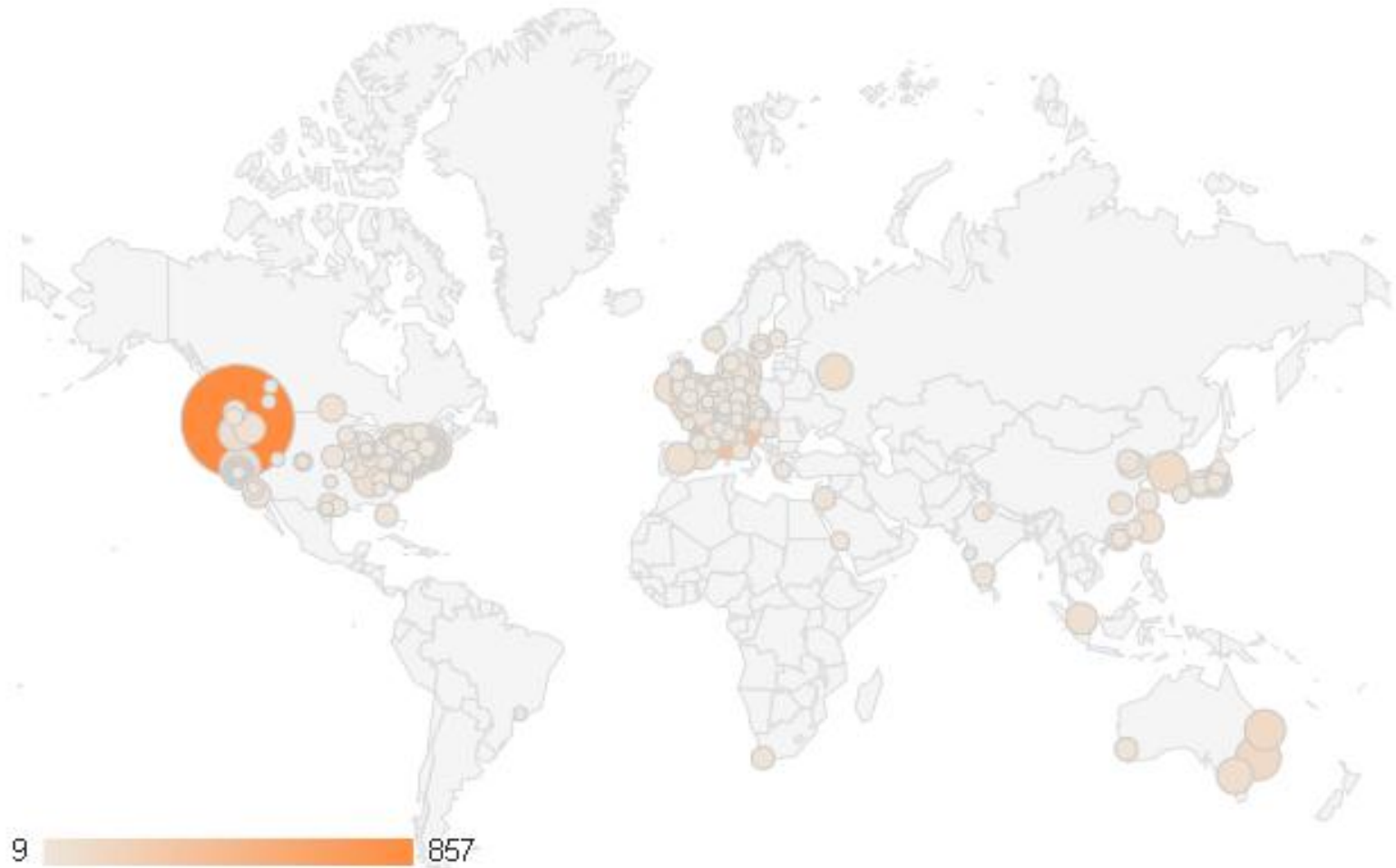
Skyline Instances Started Trailing 7 Days

ASMS 2012

ASMS 2011

ASMS 2010

Instances

Legend:
- Dev 64 Release
- Dev Release
- 1.3 64 Release
- 1.3 Release
- 1.2 Release
- 1.1 Release
- Pre-1.1 Release

Skyline Web Site Visits (past 3 months)

9 | 857

Getting Started

- Freely available & Open Source
  http://proteome.gs.washington.edu/software/skyline
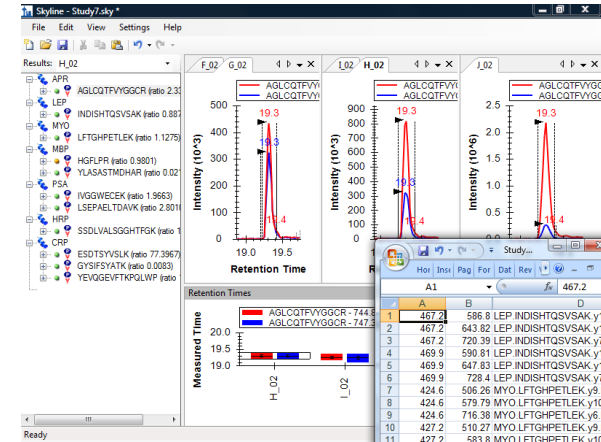- Self-updating web installation
- 3 instructional videos (25 minutes each)
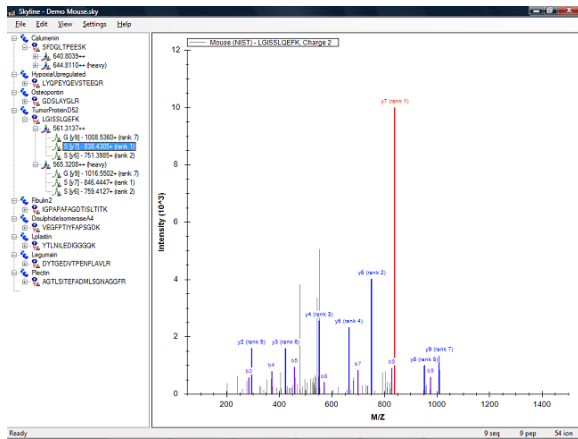
Building Methods   Results & Refinement   Existing Experiments
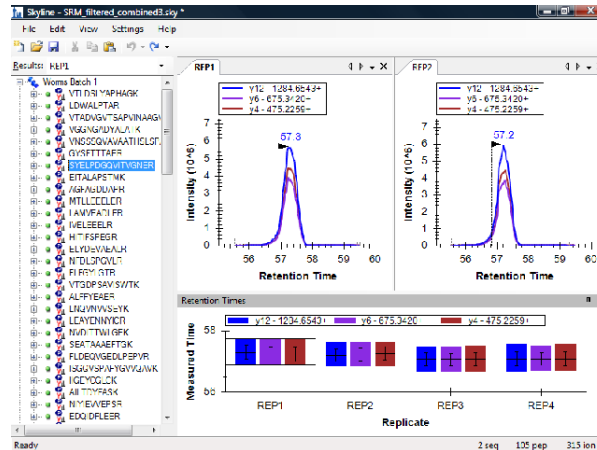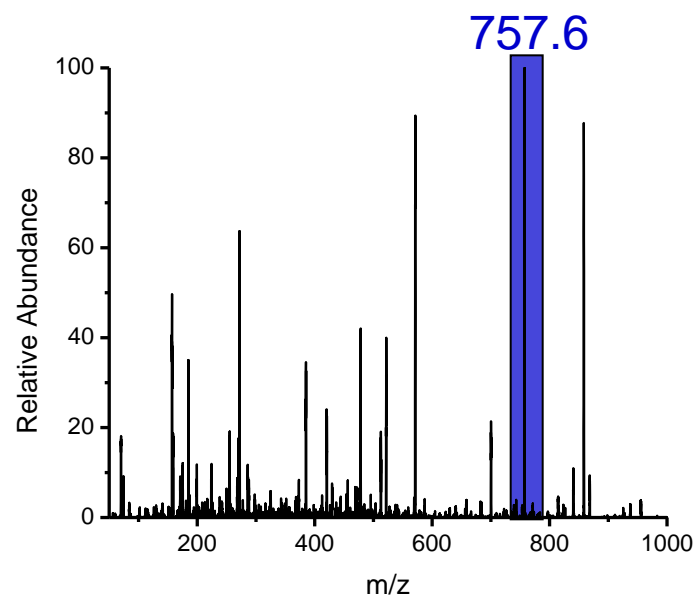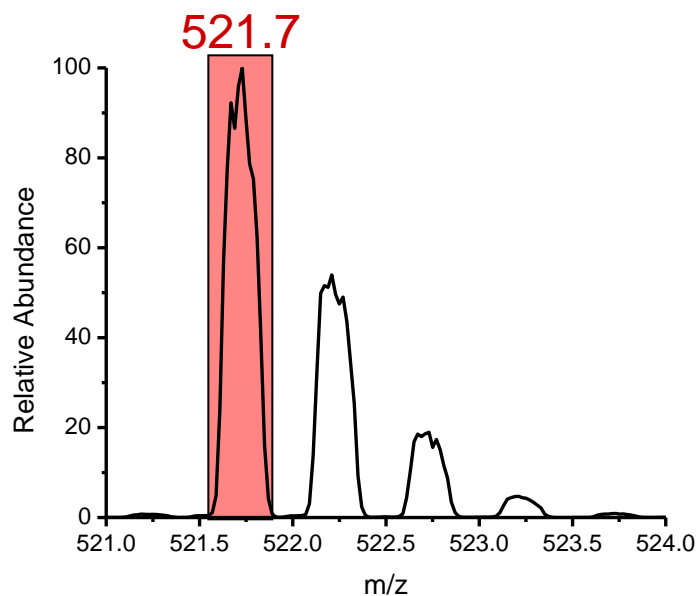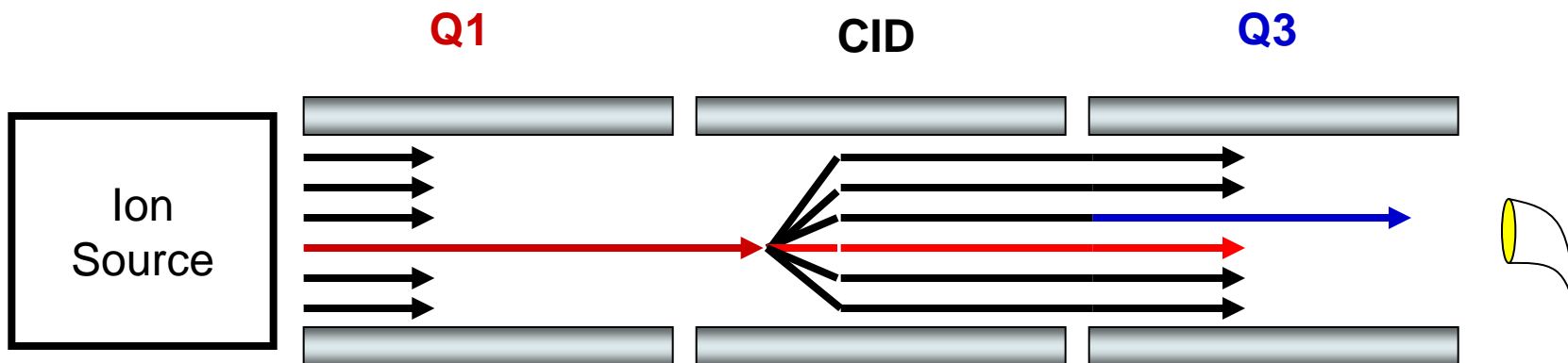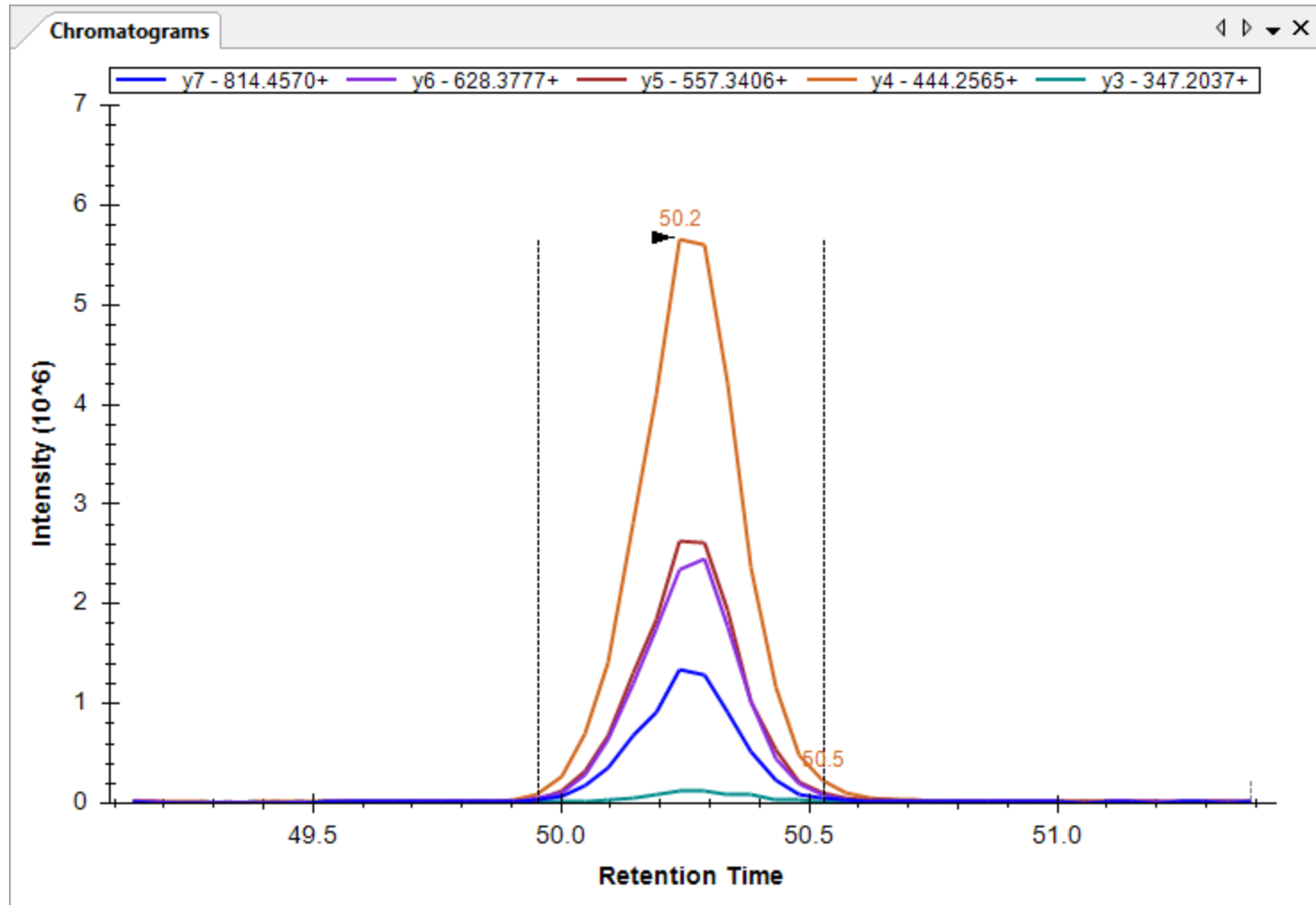

- Support board and issues list
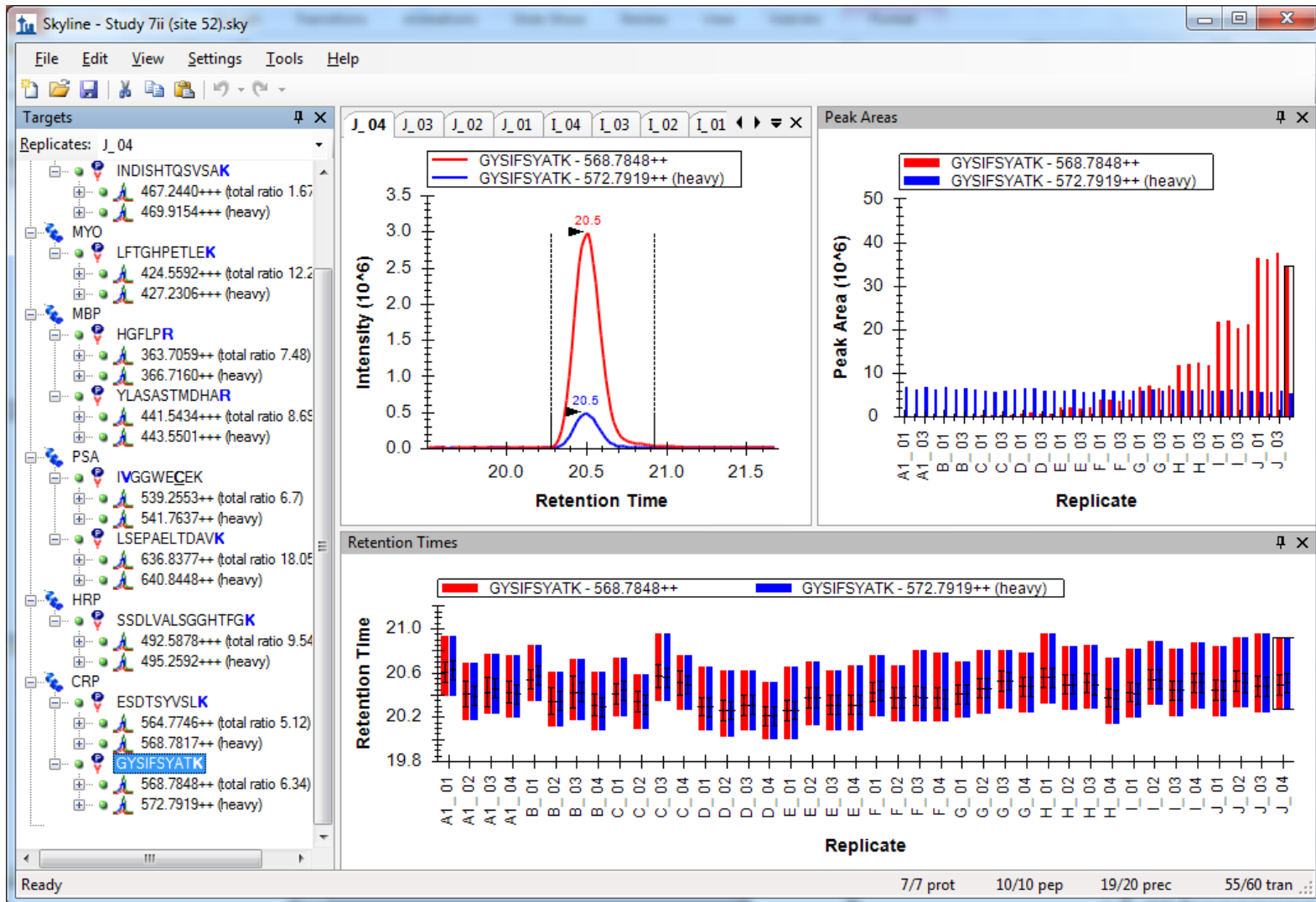- 11 tutorials & 3 instructional videos

Selected Reaction Monitoring

Selected Reaction Monitoring
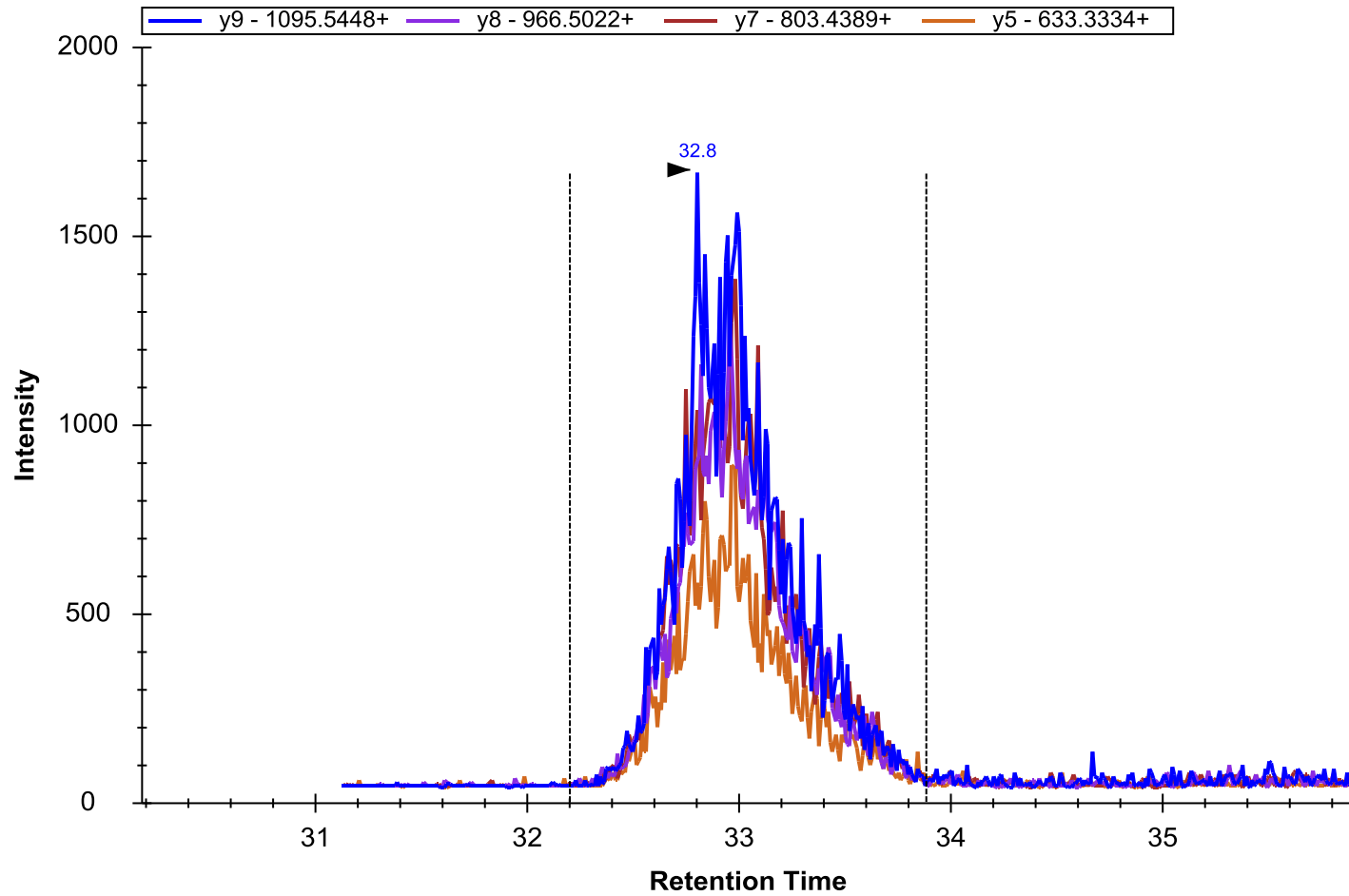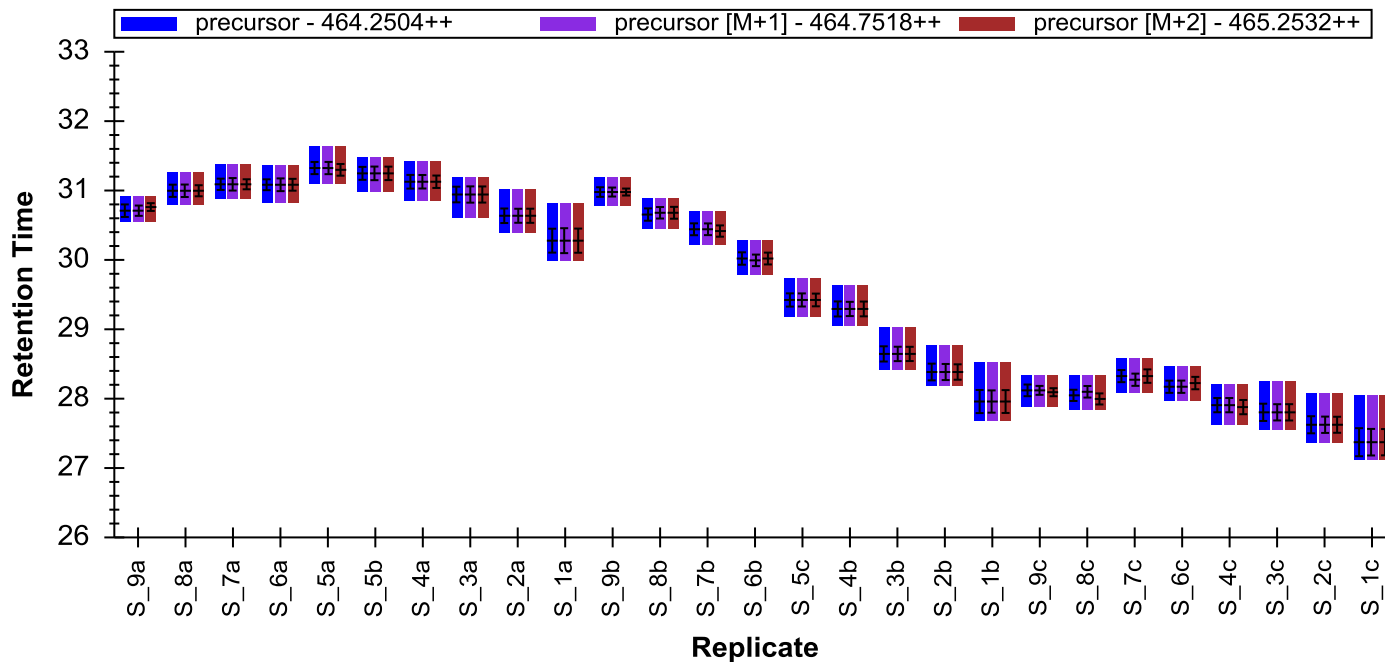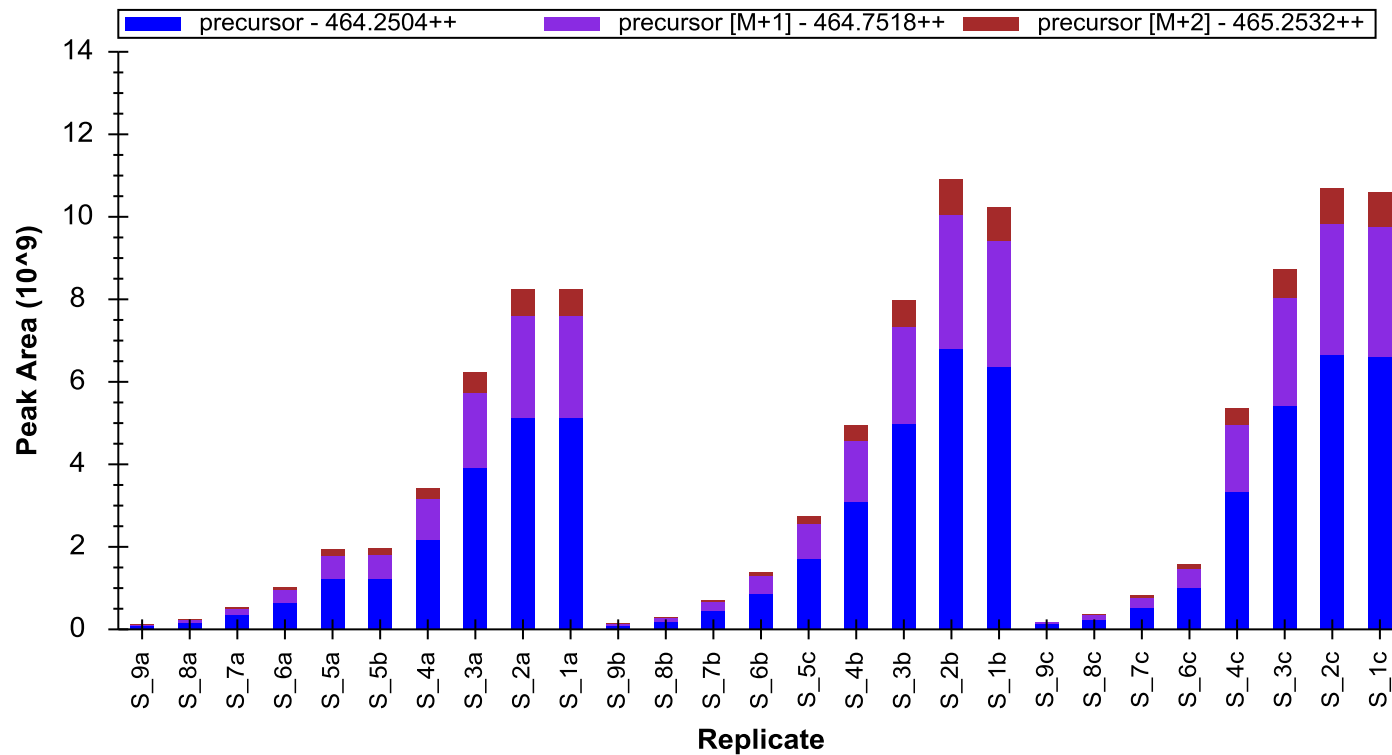
Depends on Chromatography

Depends on Chromatography

Importance of Run Order

▸ Potential confounding variable in mass spec
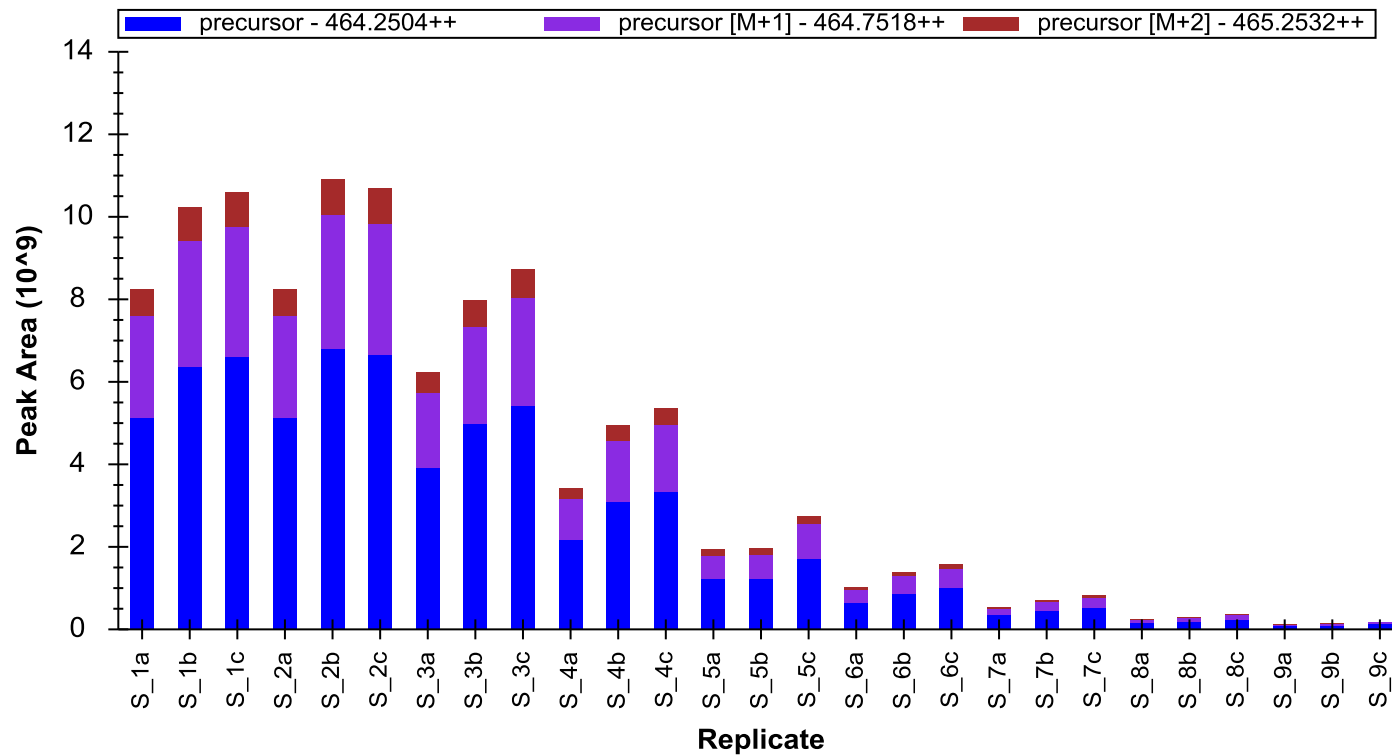
▸ System degradation over time

▸ Carryover

Importance of Run Order

▶ Acquired order

Importance of Run Order

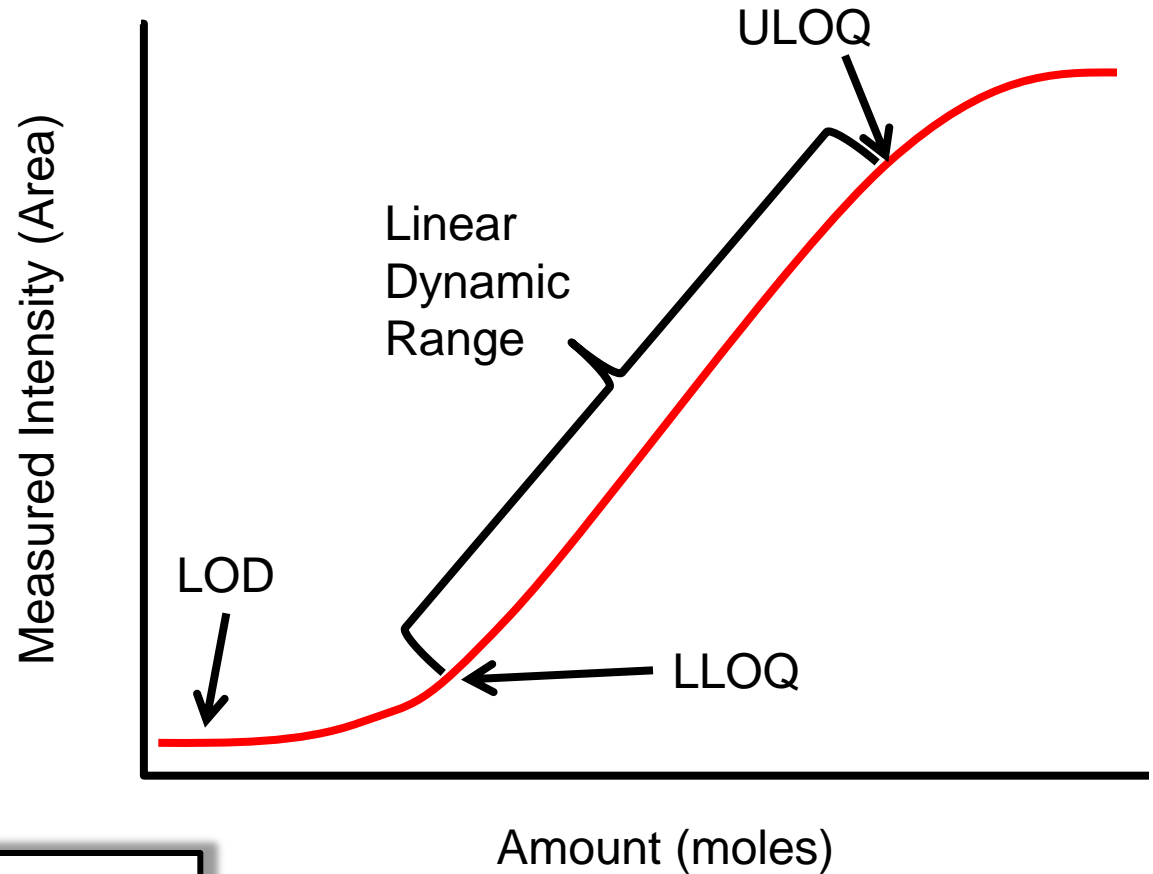▸ Grouped by concentration, then acquired order

Randomization & Blocking

▸ Very important to grouped studies

▸ Because of implications of run order

▸ Avoid

   ▸ Running all healthy on day 1

   ▸ Running all diseased on day 2

Accurate Quantification Requires Linear Response

LOD: Limit of Detection
LLOQ: Lower Limit of Quantitation
ULOQ: Upper Limit of Quantitation

▶ **Tutorial**

Labeled Reference Peptides

▶ High-accuracy quantification

  ▶ Normalize away some technical variance

▶ Certainty of peptide being measured

▶ Interference detectable

▶ Expensive

▶ Use for accurate measurement

▶ Not necessarily for differential discovery

▶ Or some hybrid

▶

Labeled Reference Peptides

▸ Extensive support in Skyline
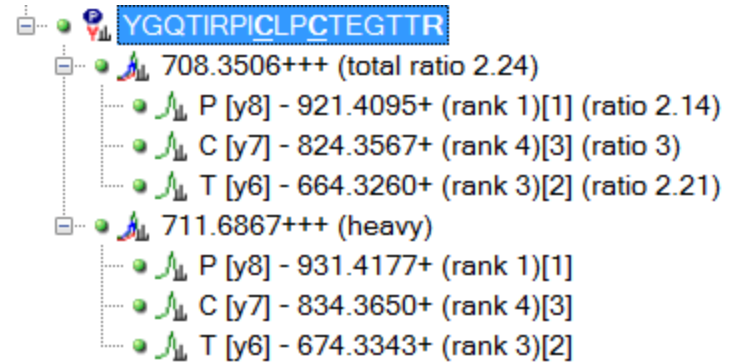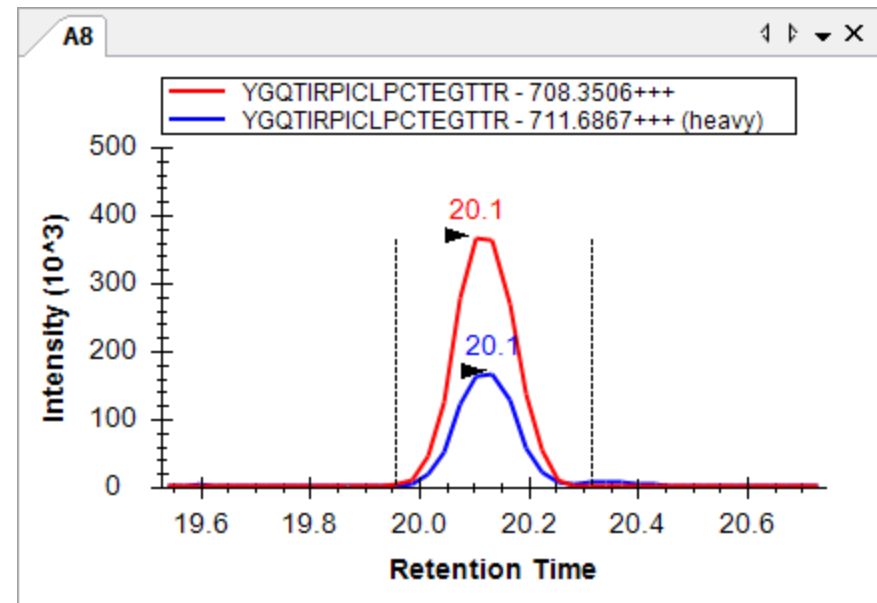
Use Prior Knowledge

- Relative ion abundance
  - Spectral libraries
- Retention time
  - iRT

- Accept some identification uncertainty
- Measure the same thing in every replicate

MS/MS Spectral Library Sources

‣ Global Proteome Machine

‣ MacCoss Lab

‣ NIST

‣ Peptide Atlas

‣ Build your own from peptide search results
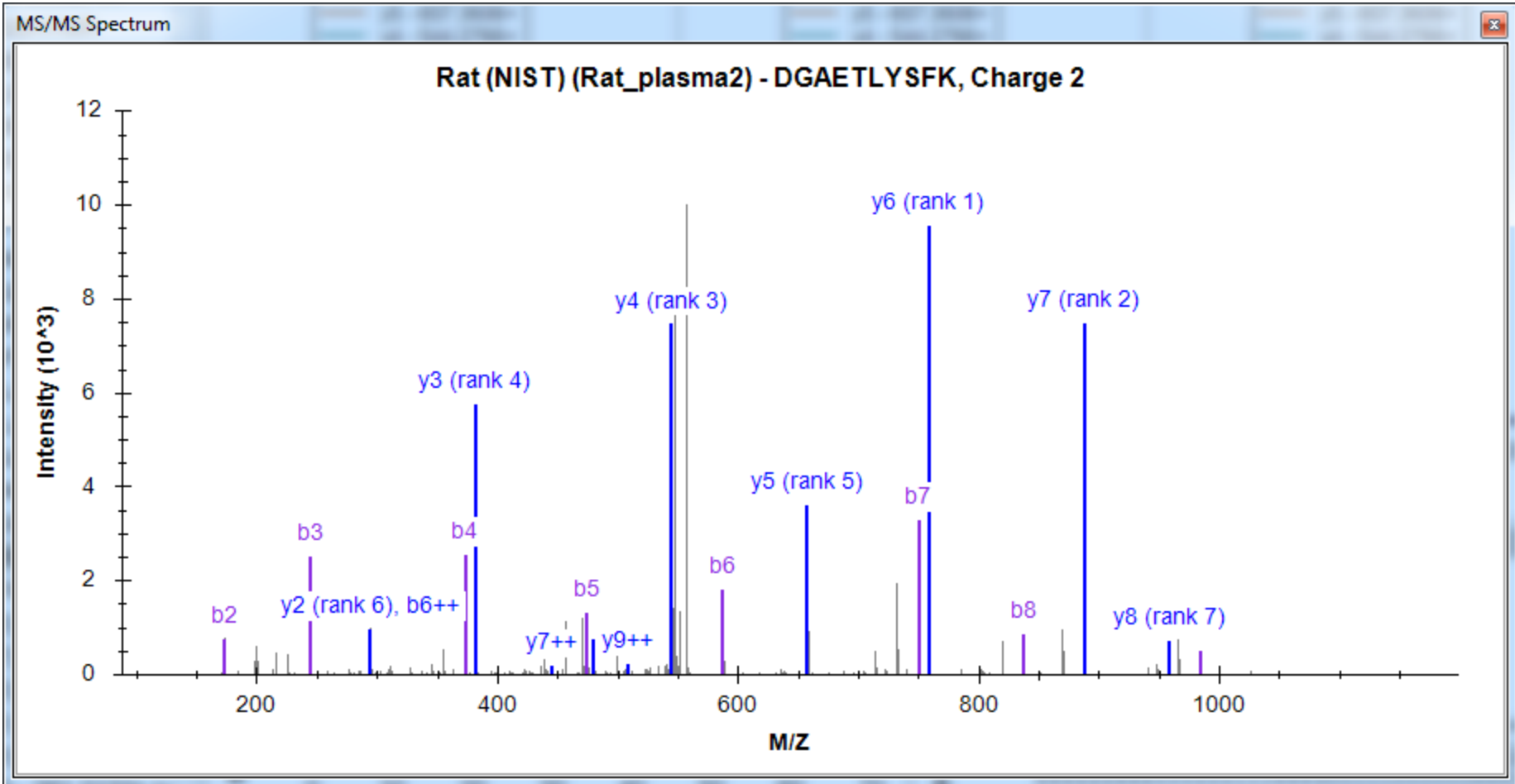
‣ Mascot
‣ Myrimatch / IDPicker
‣ OMSSA
‣ MaxQuant Andromeda
‣ PRIDE XML
‣ Protein Pilot
‣ Protein Prospector

‣ Proteome Discoverer (MSF)
‣ Scaffold – mzIdentML / MGF
‣ Spectrum Mill
‣ TPP – pepXML / mzXML files – Peptide Atlas
‣ X! Tandem
‣ Waters MSe

Library Spectra

▸ Choosing transitions   ▸ Peak validation
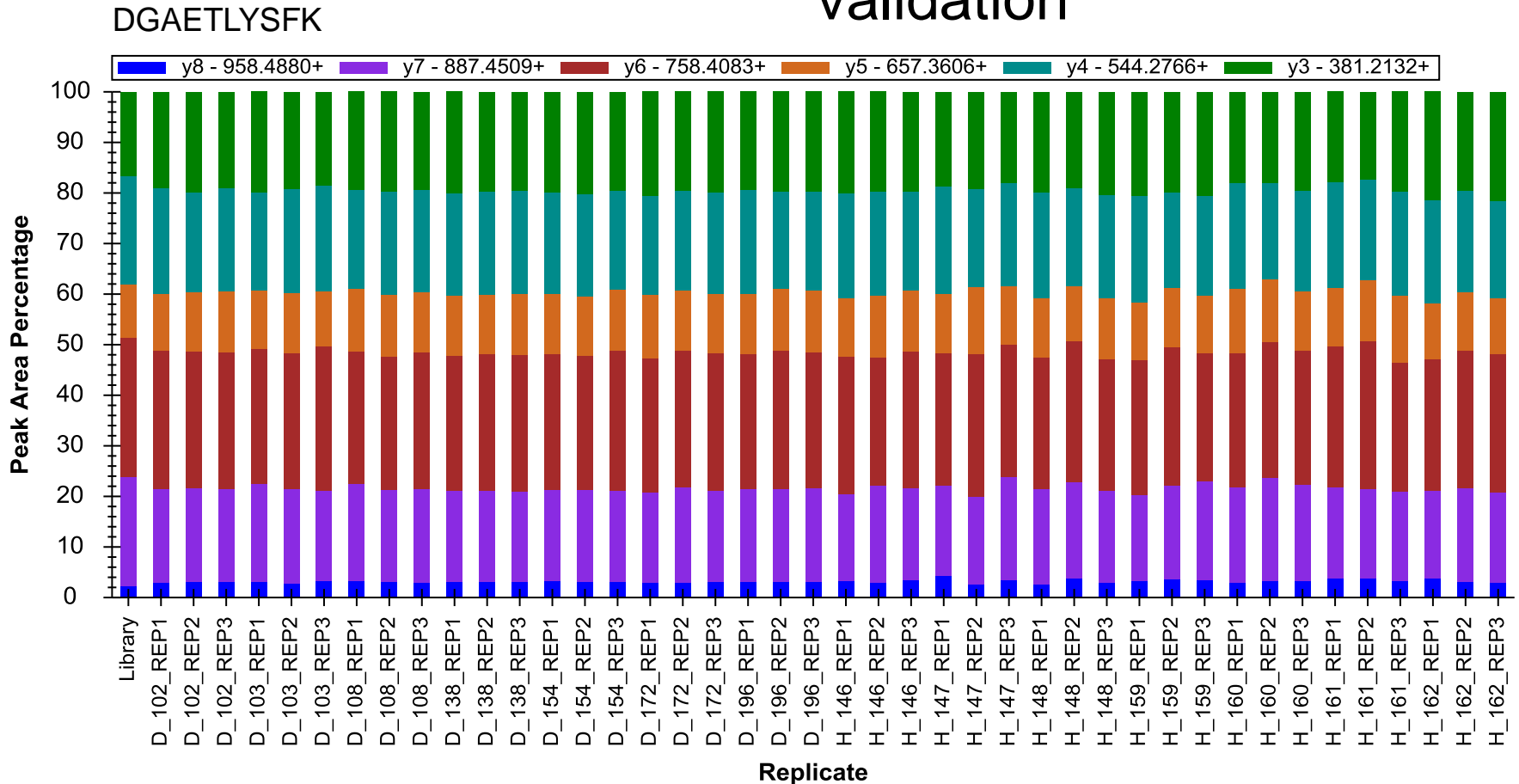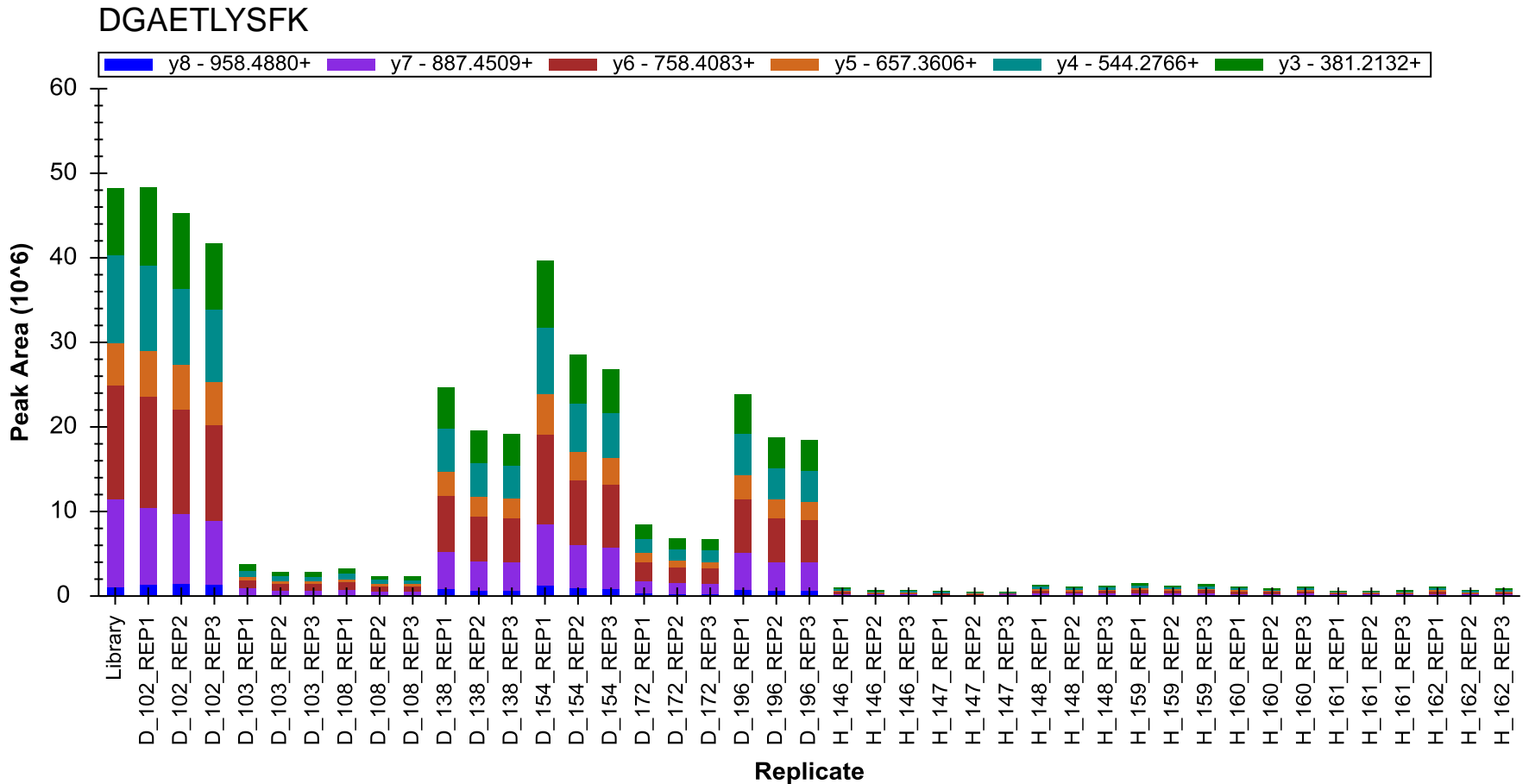
Normalized Ion Abundance

▸ Peak validation

▸ Replicate measurement validation

DGAETLYSFK

Un-normalized Ion Abundance

DGAETLYSFK

iRT Retention Time Prediction

Escher, C. et al. Proteomics. 2012.

Picking What to Measure

- Proteins – Biological question
- Peptides
  - Measure and refine
  - Chromatogram libraries (Panorama)
- Transitions
  - Spectral libraries
  - Measure and refine
  - Chromatogram libraries (Panorama)

Panorama Knowledge Base for Skyline

- ▶ Experiment and assay repository software
- ▶ panoramaweb.org – hosted at UW
- ▶ Exporting chromatogram libraries

- ▶ Vagisha Sharma (UW) and Josh Eckels (LabKey Software)

Panorama
A targeted proteomics knowledge base for Skyline

Experimental Data

- Human plasma experiment – ovarian cancer
  - Ruth Hüttenhain at ETH Zürich
- 81 subjects (15 benign, 66 cancer)
- No technical replicates
- Method design
  - Proteins from literature
  - Labeled reference peptides
  - All targets measured by MS/MS to choose transitions
  - iRT values calculated

- http://skyline.gs.washington.edu/tutorials/20130310_USHUPO_Workshop.zip

System Changes
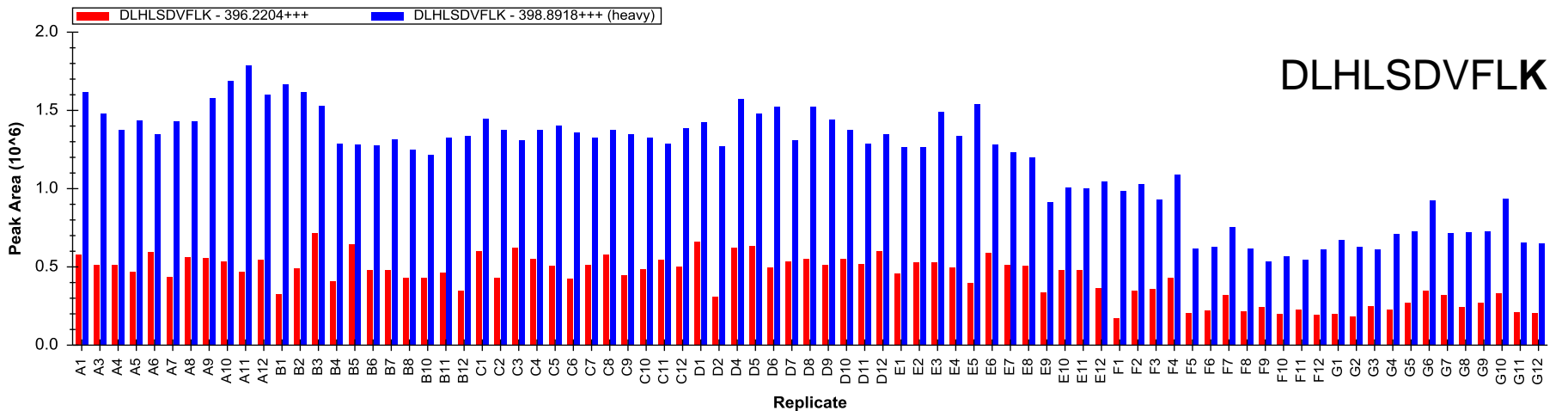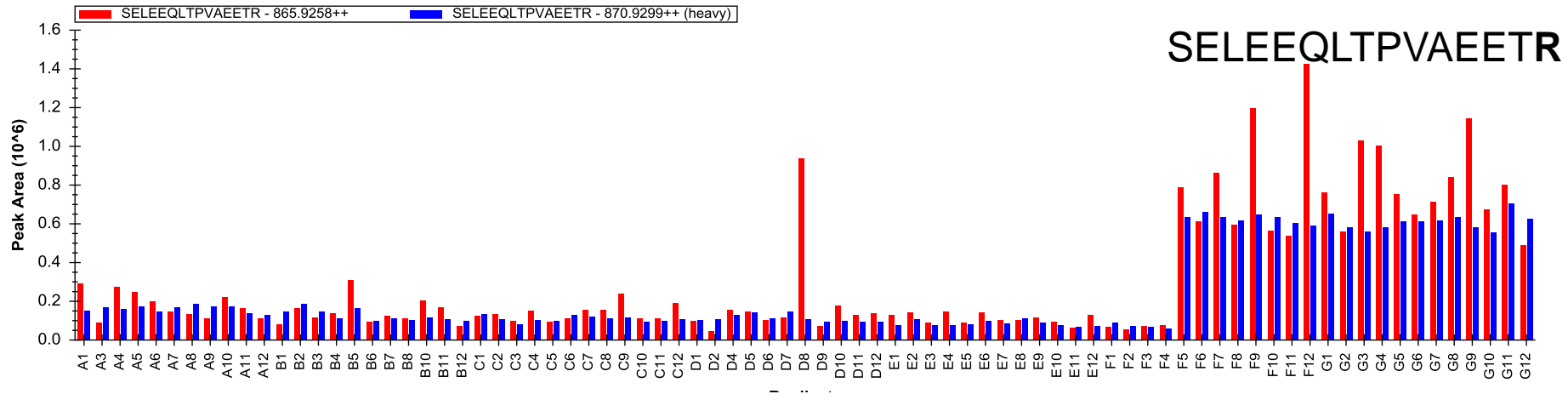
17 to 22-Feb-2012

16 to 17-Mar-2012

System Changes Normalized

System Changes

Hard to See Statistical Difference

- GELS
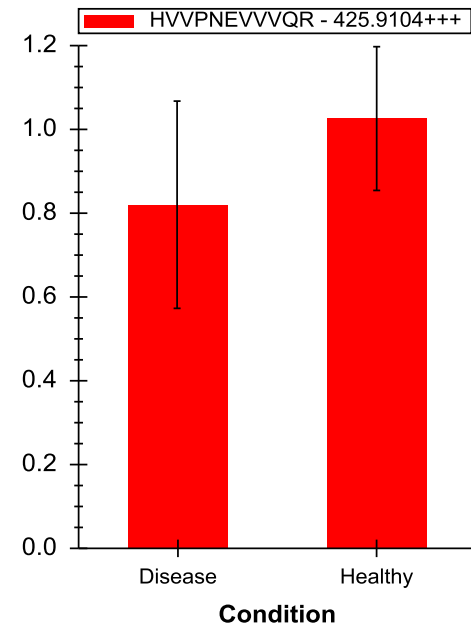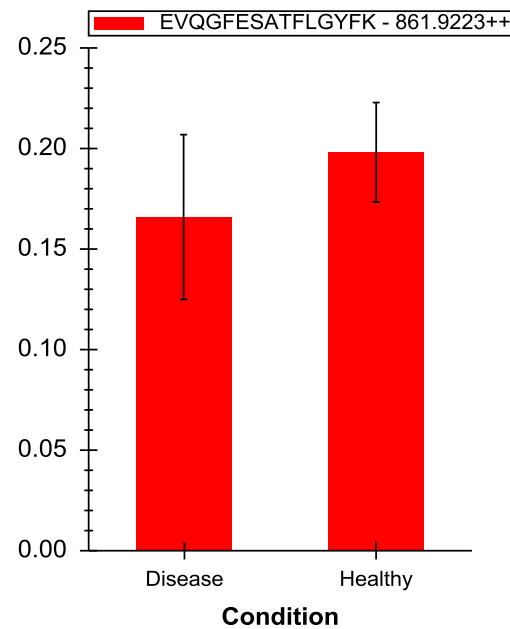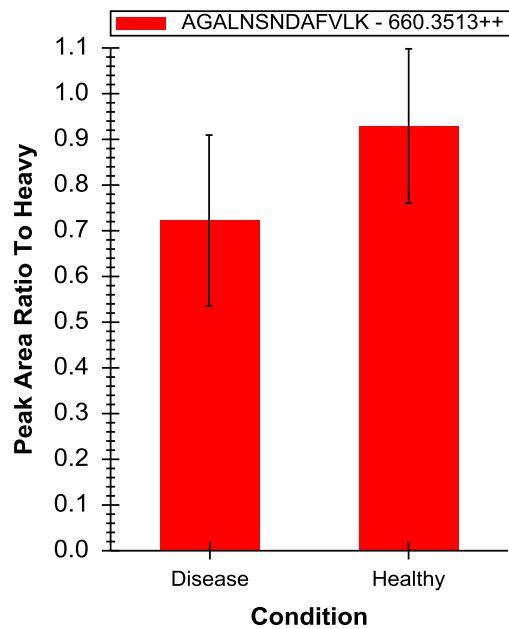
Hard to See Statistical Difference

‣ HPT
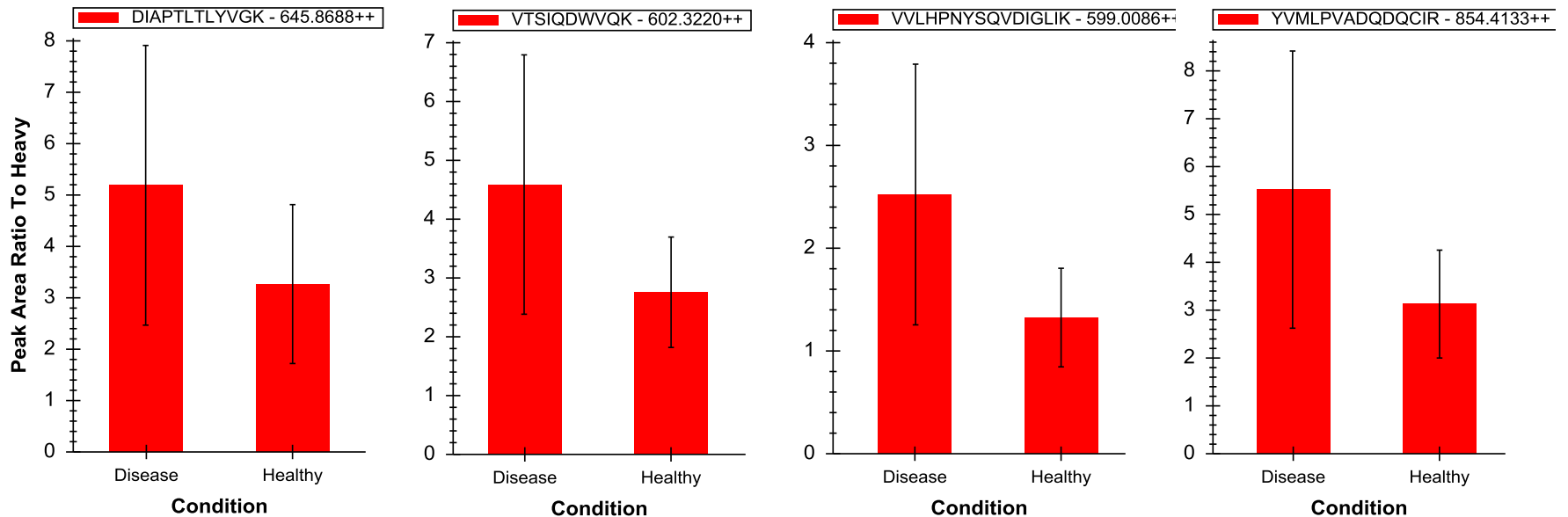
Supporting a Broader Hypothesis
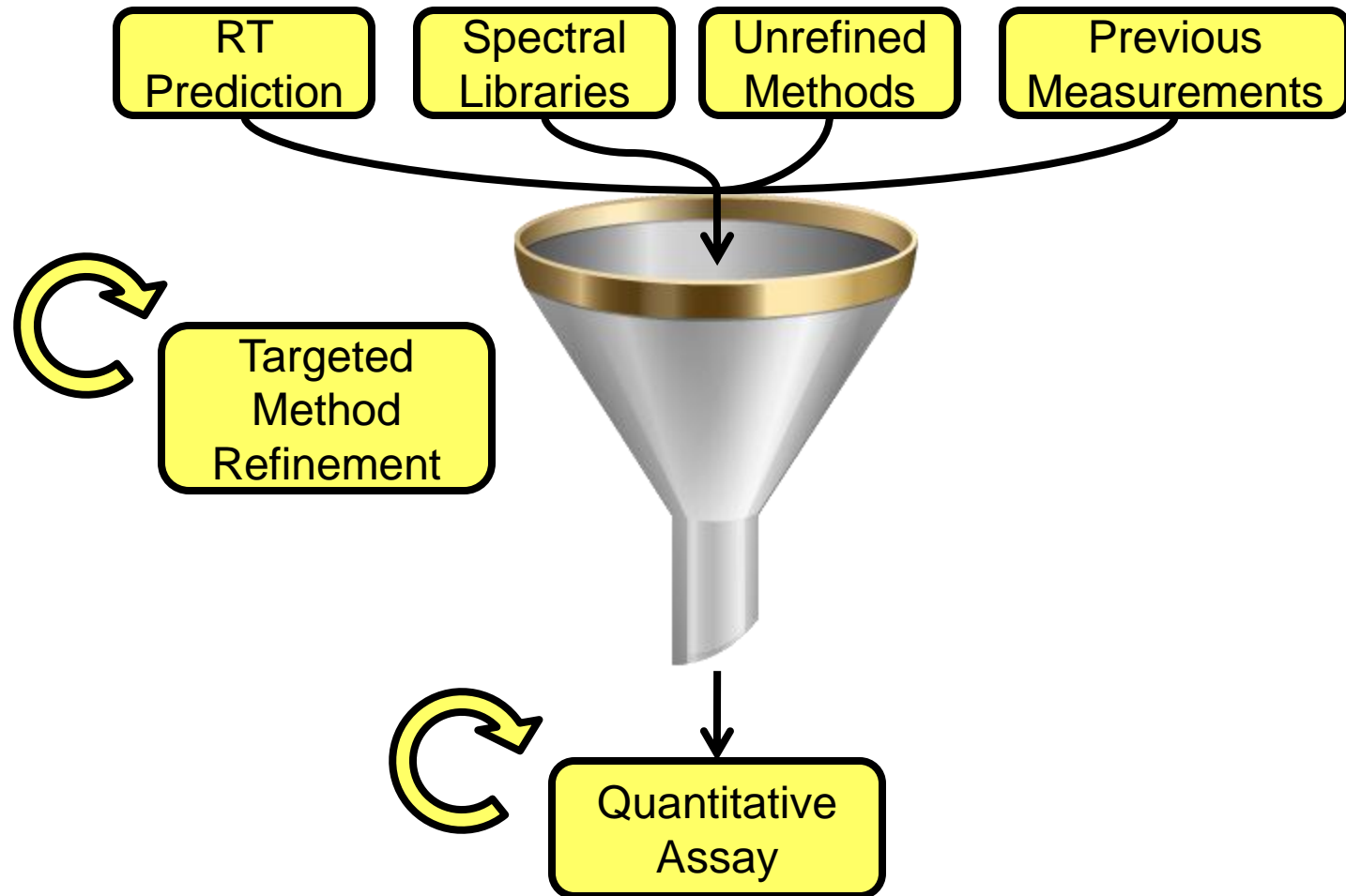
Targeted Method Refinement
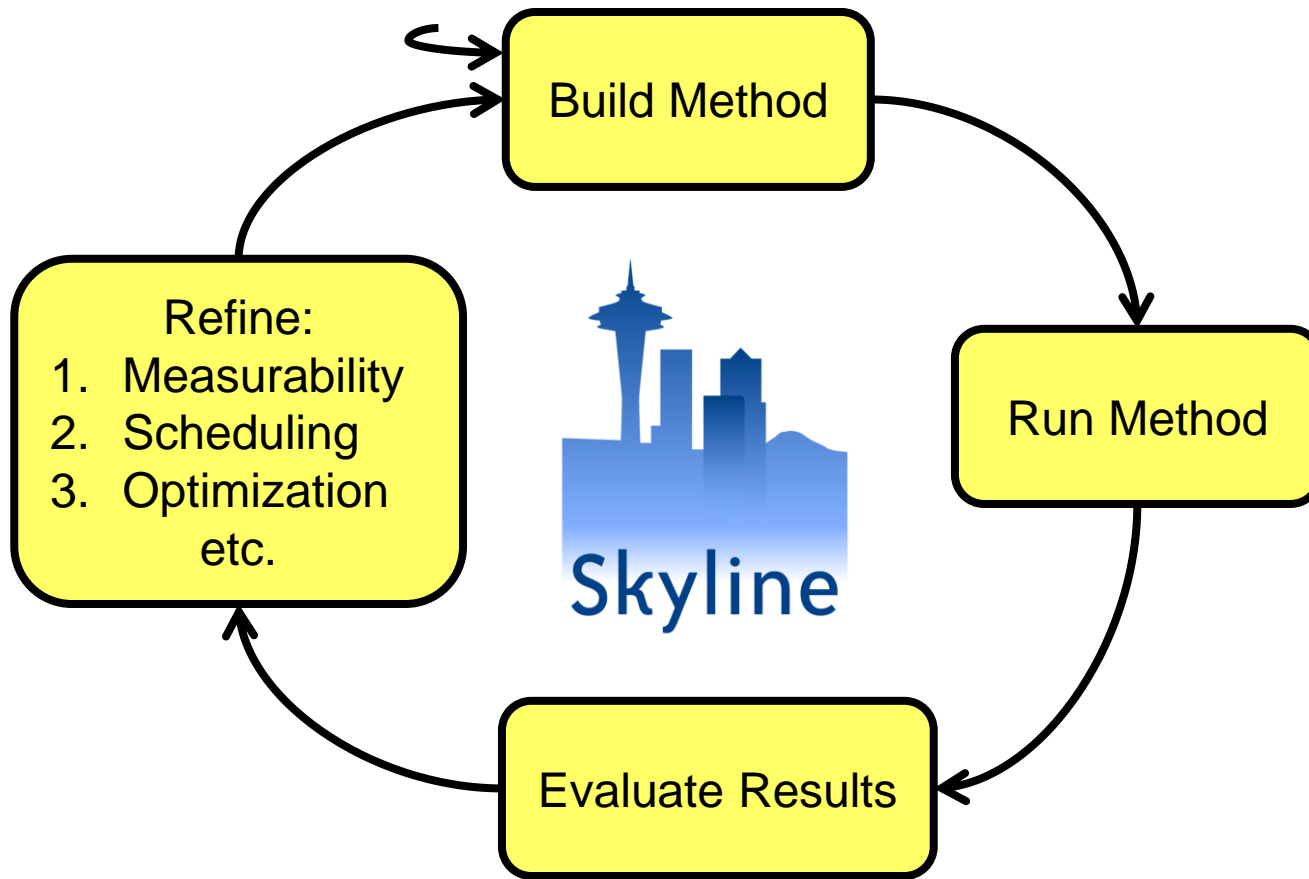
Identifying Plasma Proteins with Altered Levels in a Dahl Salt Sensitive Rat Model

7 Rats on a Low Salt Diet

7 Rats on a High Salt Diet

109 proteins reported in literature as related to heart disease

The Unrefined Method

- 109 Proteins

- 2165 Peptides

- 12,194 Transitions

- 151 Sample injections
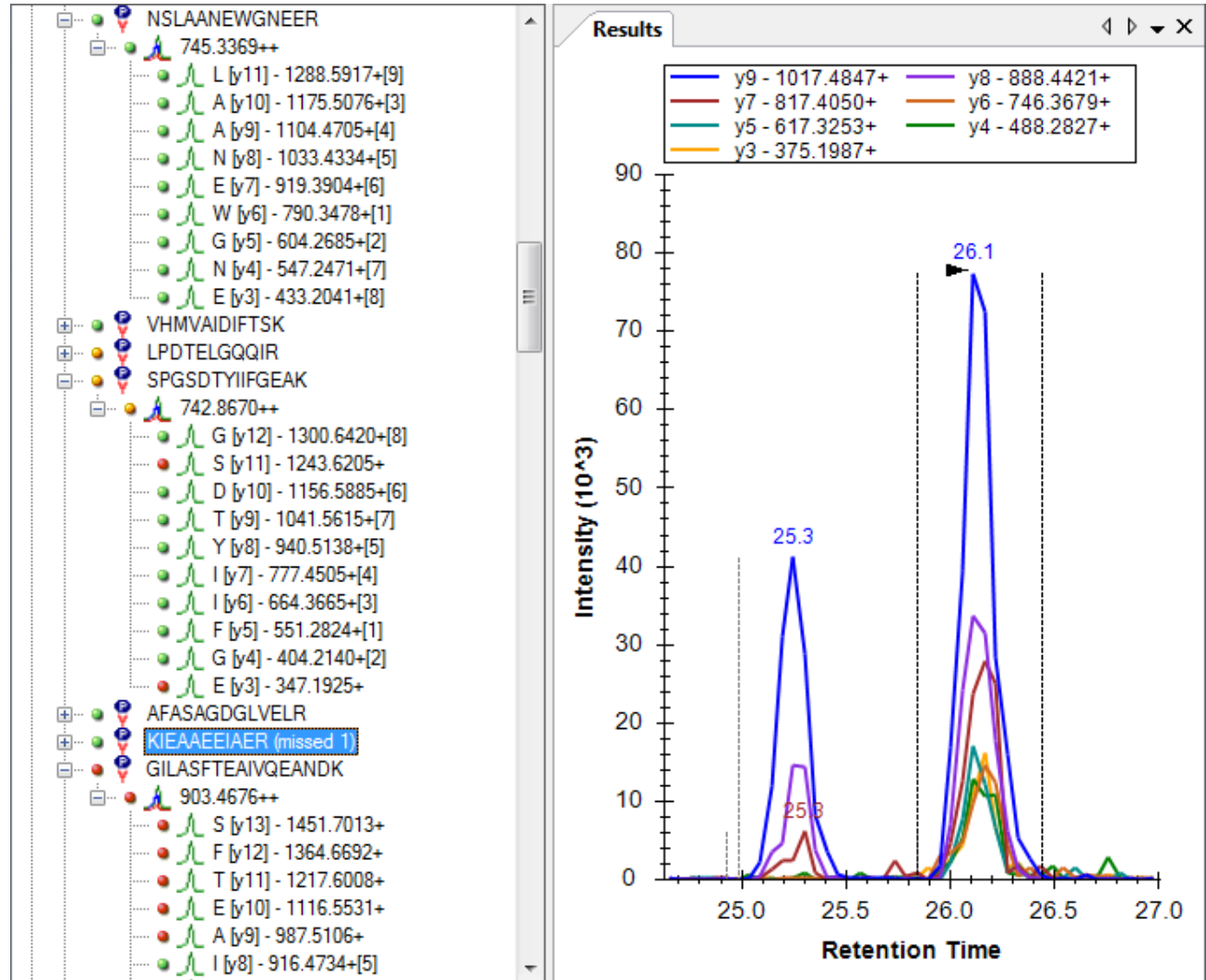
Scheduling Multi-Replicate Assays

- 49 Proteins

- 135 Peptides

- 780 Transitions

- 6 Unscheduled Injections
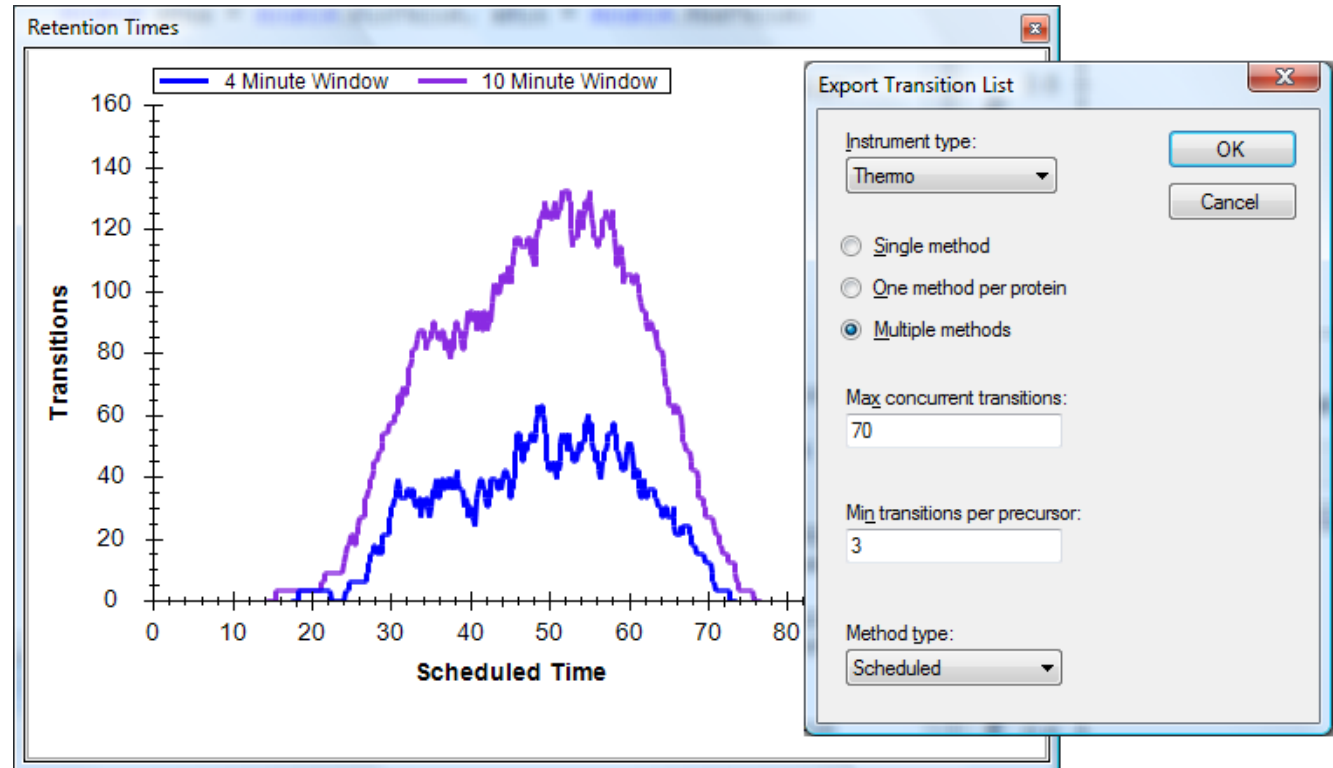
Experimental Data

- Rat plasma experiment
  - Daniela Tomazela at University of Washington
- 14 subjects (7 diseased, 7 healthy)
- Technical triplicate
- Method design
  - Proteins from literature
  - No labeled reference peptide
  - Sparse spectral library coverage
  - Exhaustive measurement of tryptic peptides
  - Exhaustive measurement of y3 – y(n-1) product ions
  - Refinement

- http://skyline.gs.washington.edu/tutorials/20130310_USHUPO_Workshop.zip

Running Multiple Replicates

▶ 42 sample injections

AIEDYVNEFSAR
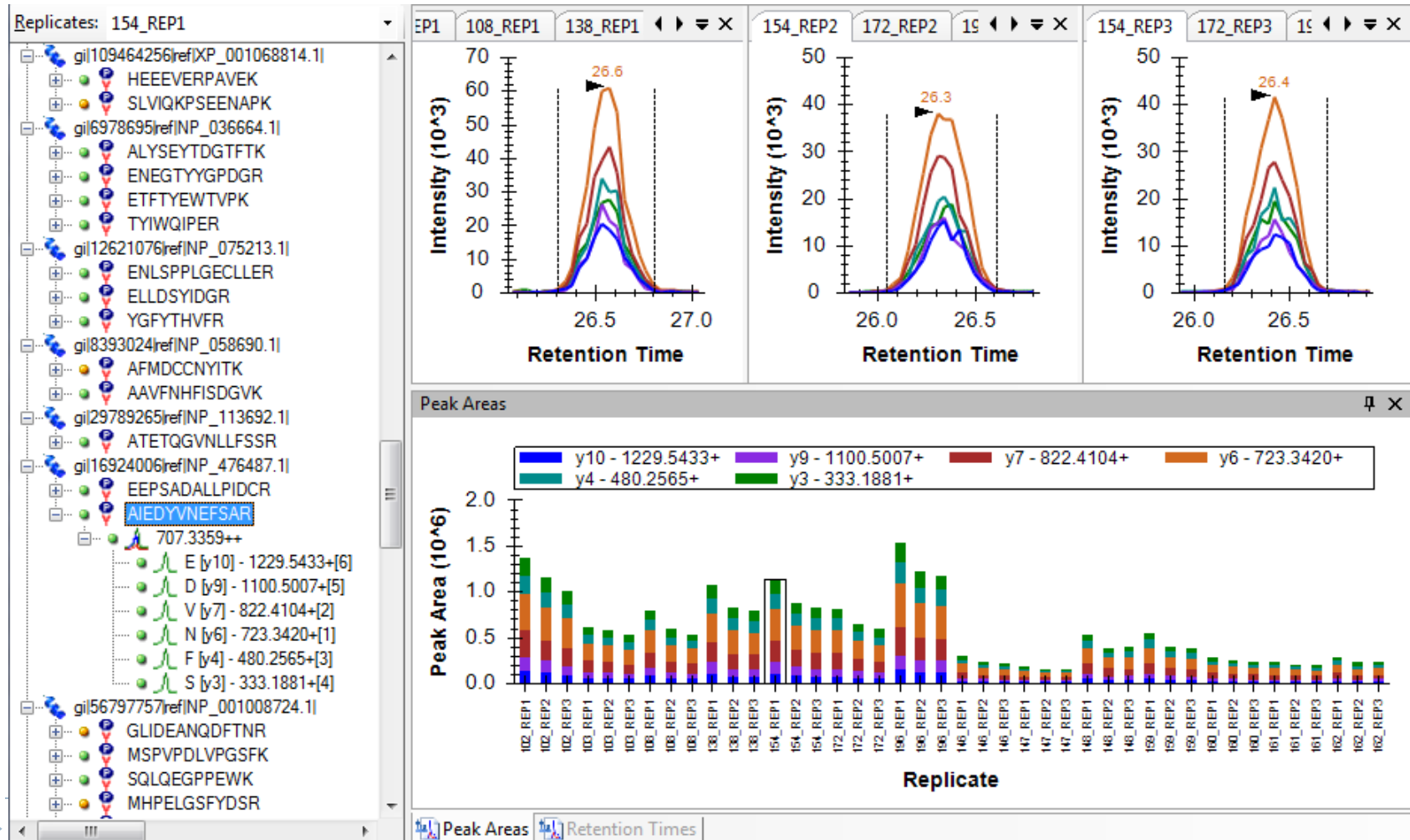
Truncated and Missing Peaks

Replicate Annotations

Custom Reports

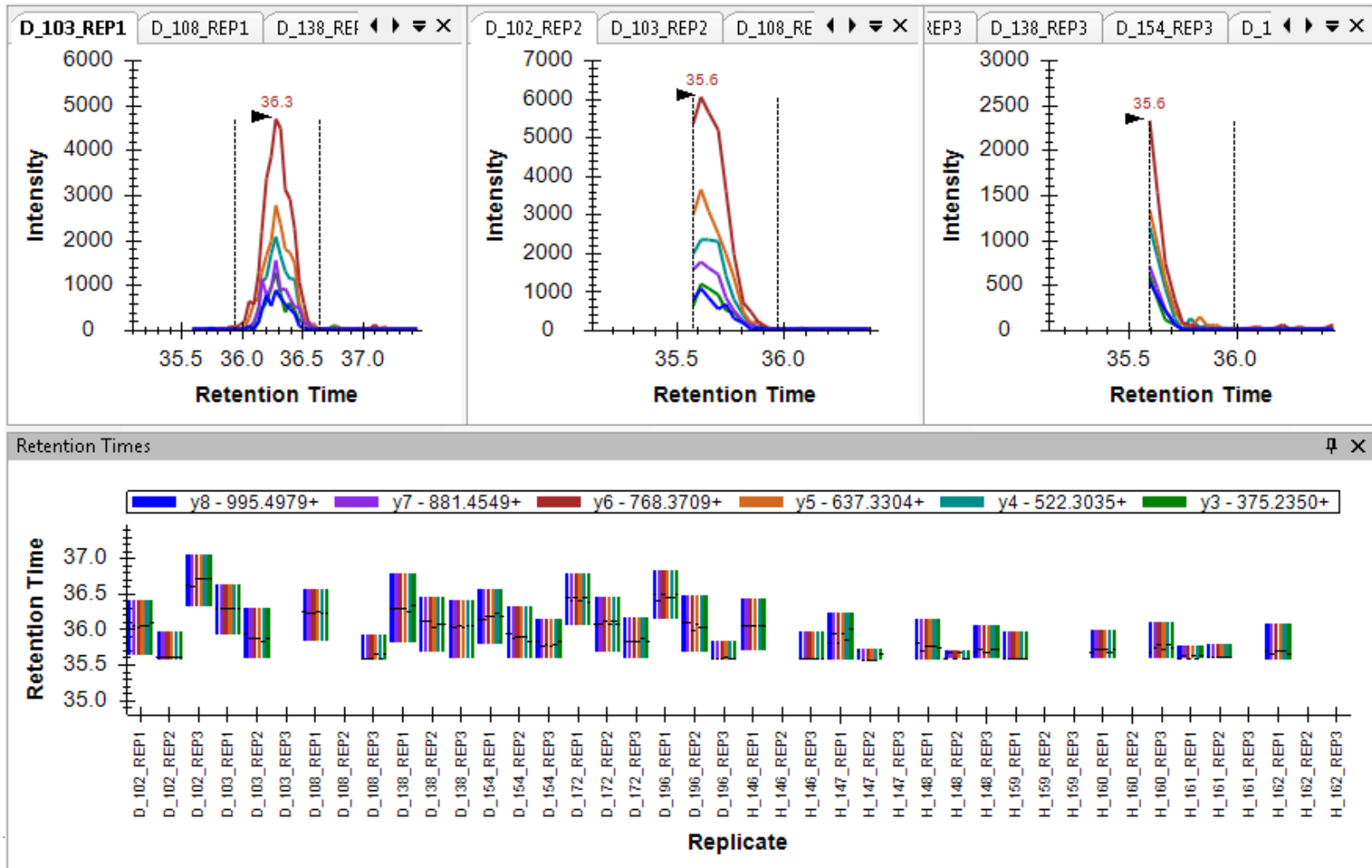| ProteinName	PeptideSequence	Precurs	Frag	Produ	IsotopeLa	Condition	Bio	Run	Area
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	1	1	14516
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	1	15	9607
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	1	29	7480
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	2	2	5692
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	2	16	5953
NP_036629	CSLPRPWALTFSYGR	2	y10	1	light	Disease	2	30	649
	FSYGR	2	y10	1	light	Disease	3	3	10476
	FSYGR	2	y10	1	light	Disease	3	17	3952
	FSYGR	2	y10	1	light	Disease	3	31	3165
	FSYGR	2	y10	1	light	Disease	4	4	9830
	FSYGR	2	y10	1	light	Disease	4	18	10671
	FSYGR	2	y10	1	light	Disease	4	32	6369
	FSYGR	2	y10	1	light	Disease	5	5	15037
	FSYGR	2	y10	1	light	Disease	5	19	9128
	FSYGR	2	y10	1	light	Disease	5	33	6918
	FSYGR	2	y10	1	light	Disease	6	6	11991
	FSYGR	2	y10	1	light	Disease	6	20	8630
	FSYGR	2	y10	1	light	Disease	6	34	6896
	FSYGR	2	y10	1	light	Disease	7	7	13061
	FSYGR	2	y10	1	light	Disease	7	21	12258
	FSYGR	2	y10	1	light	Disease	7	35	9037
	FSYGR	2	y10	1	light	Healthy	8	8	7891
	FSYGR	2	y10	1	light	Healthy	8	22	3362
	FSYGR	2	y10	1	light	Healthy	8	36	4448

Edit Report

Report Name: MSstats2 Input   Preview...

- Peptides
- Results
- ProteinName
- ProteinDescription
- ProteinSequence
- ProteinNote

Add >

ProteinName
PeptideSequence
PrecursorCharge
FragmentIon
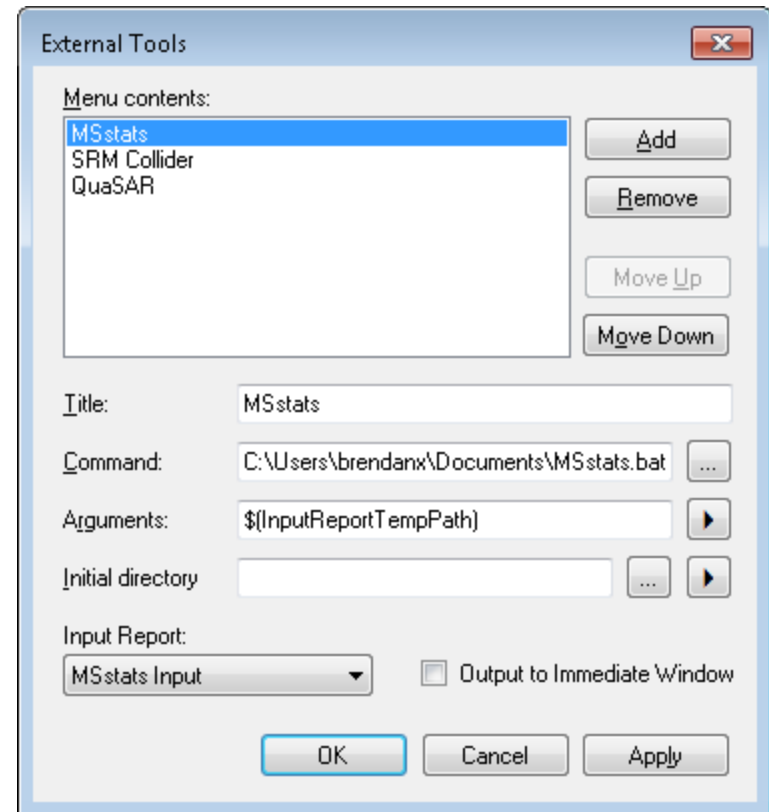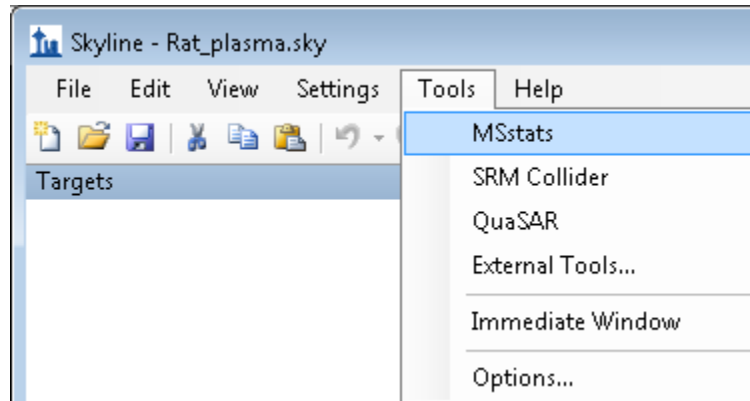ProductCharge
IsotopeLabelType
Condition
BioReplicate
Run
Area

☐ Pivot Replicate Name   ☐ Pivot Isotope Label

OK   Cancel

▶ **Tutorial**

External Tool MSstats2

External Tool – Web Page

Downstream Analysis with Statistical Tools

▸ Analysis of reports with R – MSstats2
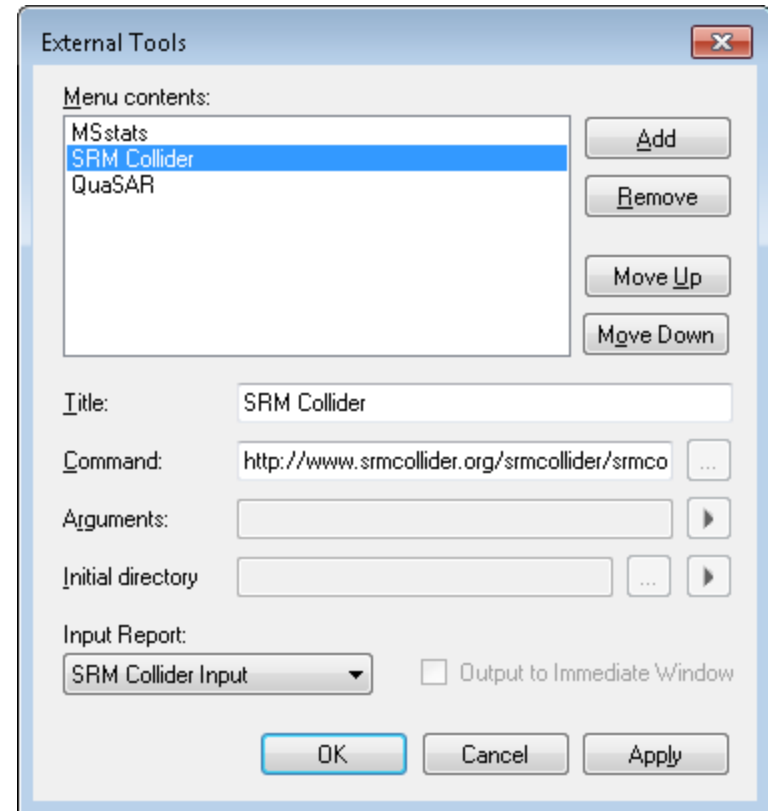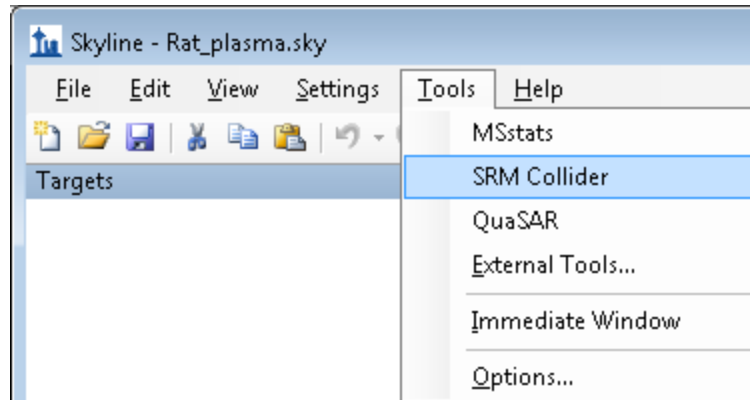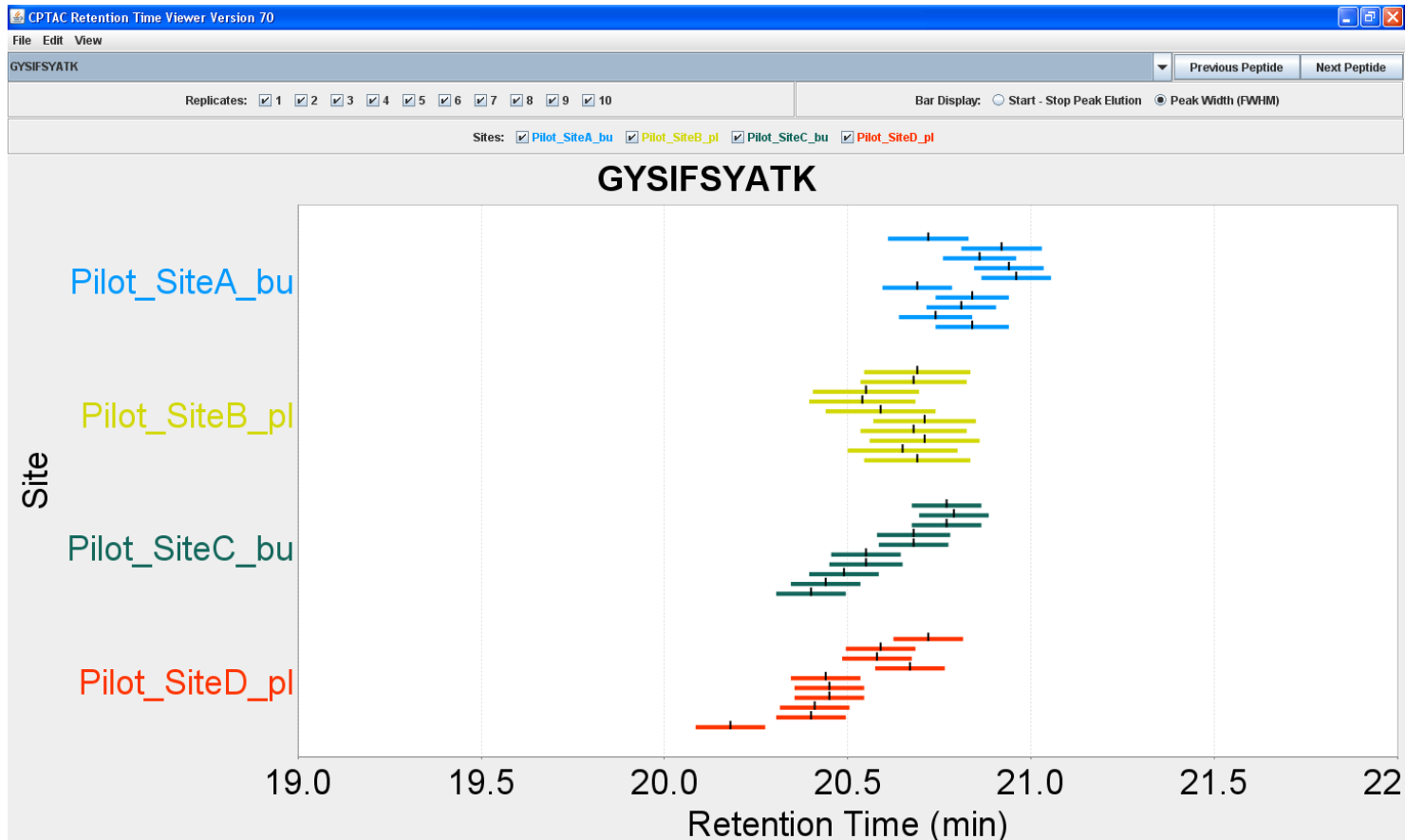
Deeper Analysis with Custom Tools

▸ Analysis of reports with Retention Time Viewer (Java program)

Collaborators:

- **U. of Wa.**
  - Eva Baker
  - Jarrett Egertson
  - Jimmy Eng
  - Sonia Ting
- **Broad Institute**
  - Sue Abbatiello
  - Steve Carr
  - Jake Jaffe
  - D. R. Mani
- **Buck Institute**
  - Birgit Schilling
  - Matthew Rardin
  - Brad Gibson

- **Duke**
  - Will Thompson
  - Arthur Moseley
- **IMSB**
  - Rudolph Aebersold
  - Christina Ludwig
  - Olga Schubert
  - Hannes Röst
  - Ludovic Gillet
  - Lars Malstrom
  - Lucia Espona Pernas

- **Vanderbilt**
  - Matthew Chambers
  - Amy Ham
  - Daniel Liebler
  - David Tabb

Instrument Vendor Support

- AB Sciex
  - Fadi Abdi
  - David Cox
  - Christie Hunter
  - Brent Lefebvre
- Agilent Technologies
  - GRANT
  - Christine Miller
  - Joe Roark
  - Pat Perkins
- Bruker
  - GRANT
  - Carsten Baessmann
  - Stephanie Kaspar
  - Marius Kallhardt

- Thermo-Scientific
  - GRANT
  - Markus Kellmann
  - Andreas Kuehn
  - Vlad Zabrouskov
- Waters
  - Laurence Firth
  - James Langridge
  - Roy Martin
  - Kieran Neeson
  - Keith Richards