Chimeric Spectra and Data-Driven DIA analysis

reSpect, DISCo, Quantic

ISB

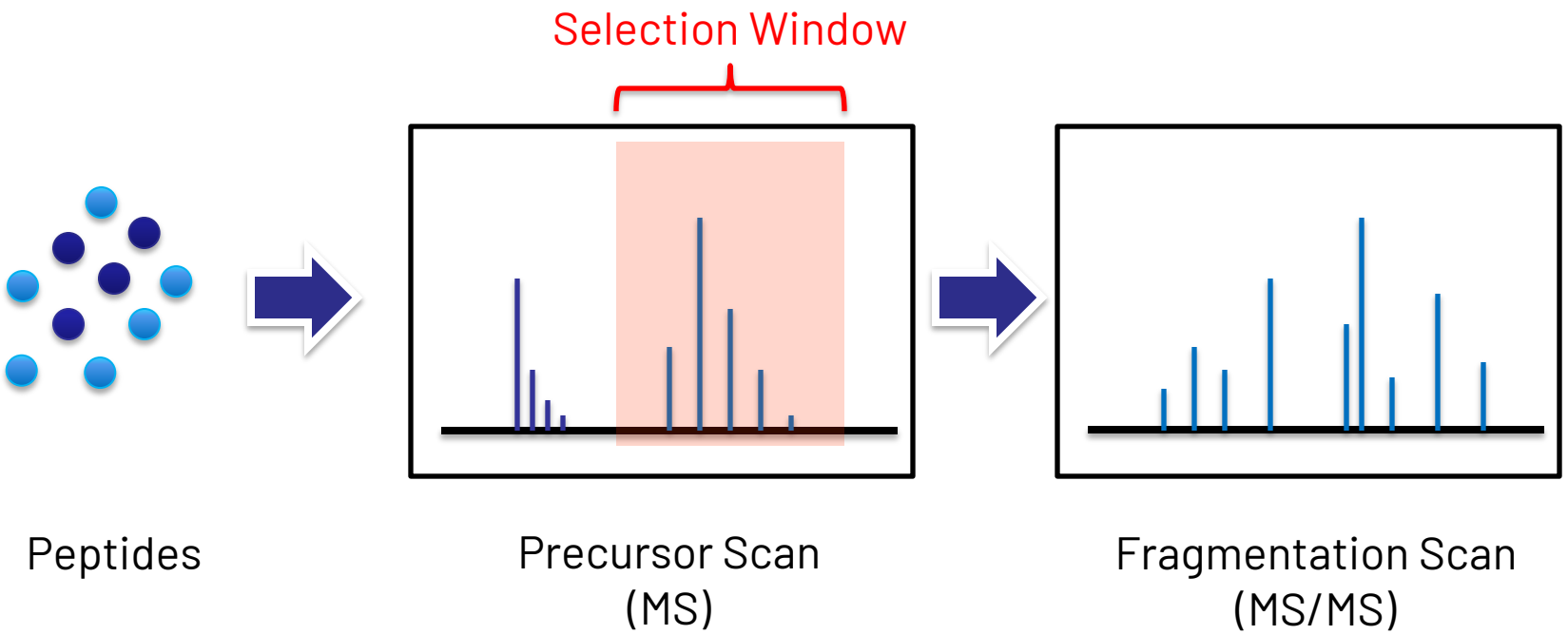
reSpect

Finding chimeric peptides from high resolution MS/MS spectra

ISB

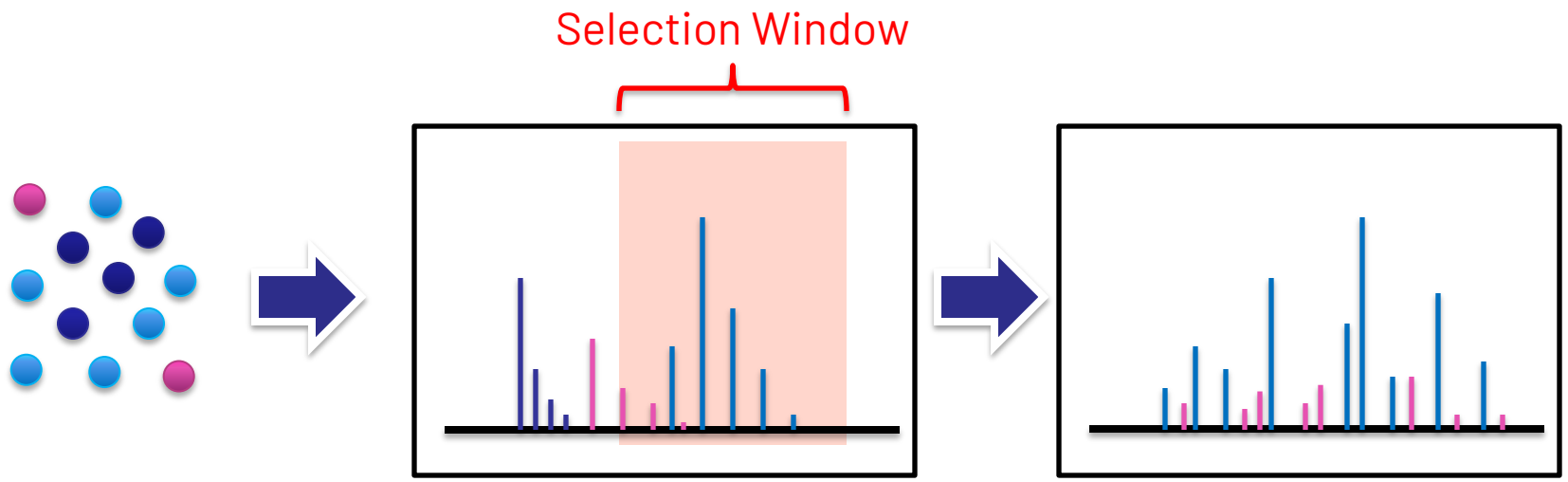
MS/MS

- **Standard Shotgun (MS/MS) Workflow**



Chimeric MS/MS

- **Chimeric spectra contain multiple precursor ions**



Selection Window includes two peptides!

Fragments ions are seen for all precursor ions

Factors Contributing To Chimeric Spectra

- **Chimeric spectra are a common occurrence in LC-MS/MS**
  - ***proportional*** to sample complexity
  - ***proportional*** to selection window size
  - influenced by chromatography conditions

Database search algorithms only report one peptide per spectrum!

Search Algorithms Do Not Find Chimeras

- **Selection window of the instrument is *much wider* than the peak resolution**
  - Search algorithms use only the measured precursor mass
    - For example: ±25 ppm

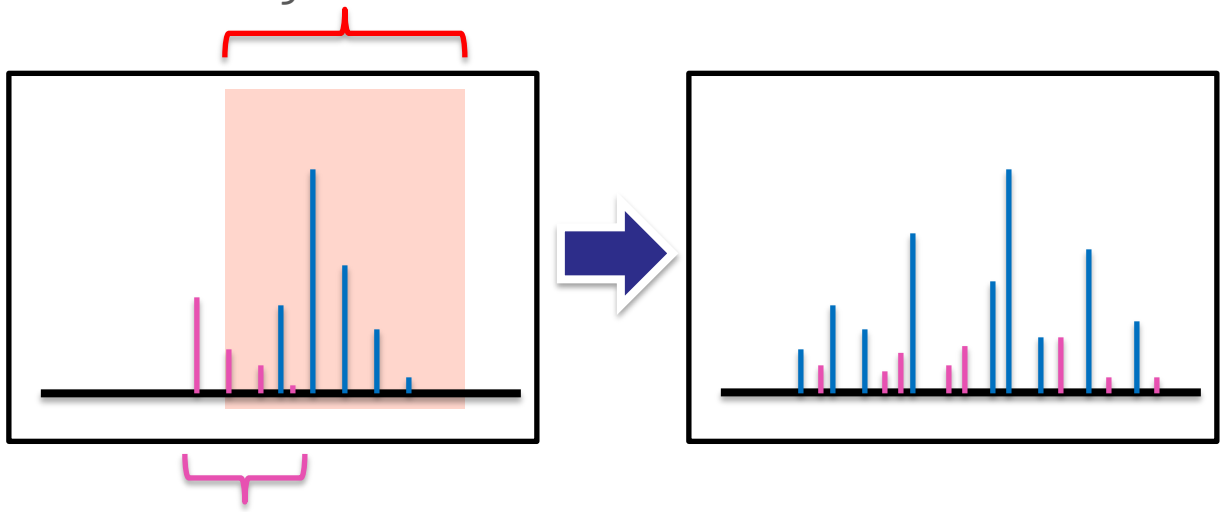
- **Selection window may include *peptides outside the search mass***
  - Those peptides never considered during search
    - For example: different charge leads to mass outside ±25 ppm tolerance

Chimeric MS/MS

Selected Ion (A):
Mass = 615.74
Charge = 3+

Database Search:
Peptides ~ 1844.22 Da

Peptide (A) Found!!



Chimeric Ion (B):
Mass = 614.85
Charge = 2+

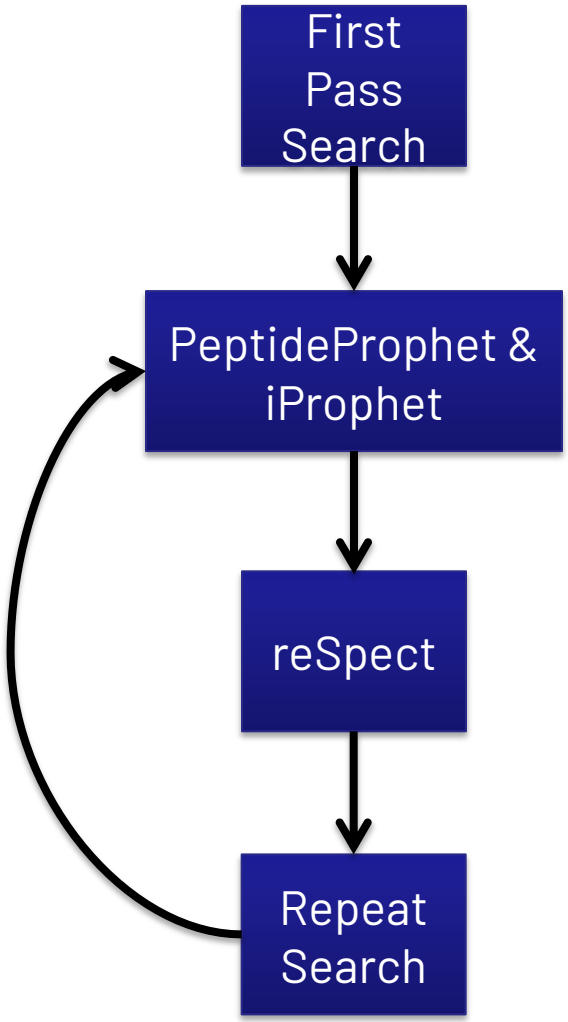
Not Searched:
Peptides ~ 1227.70 Da

Peptide (B) NOT Found!!

reSpect

- **reSpect**
  - Allows identification of chimeric peptides from spectra.
  - Works with existing search algorithms.

reSpect: The Approach



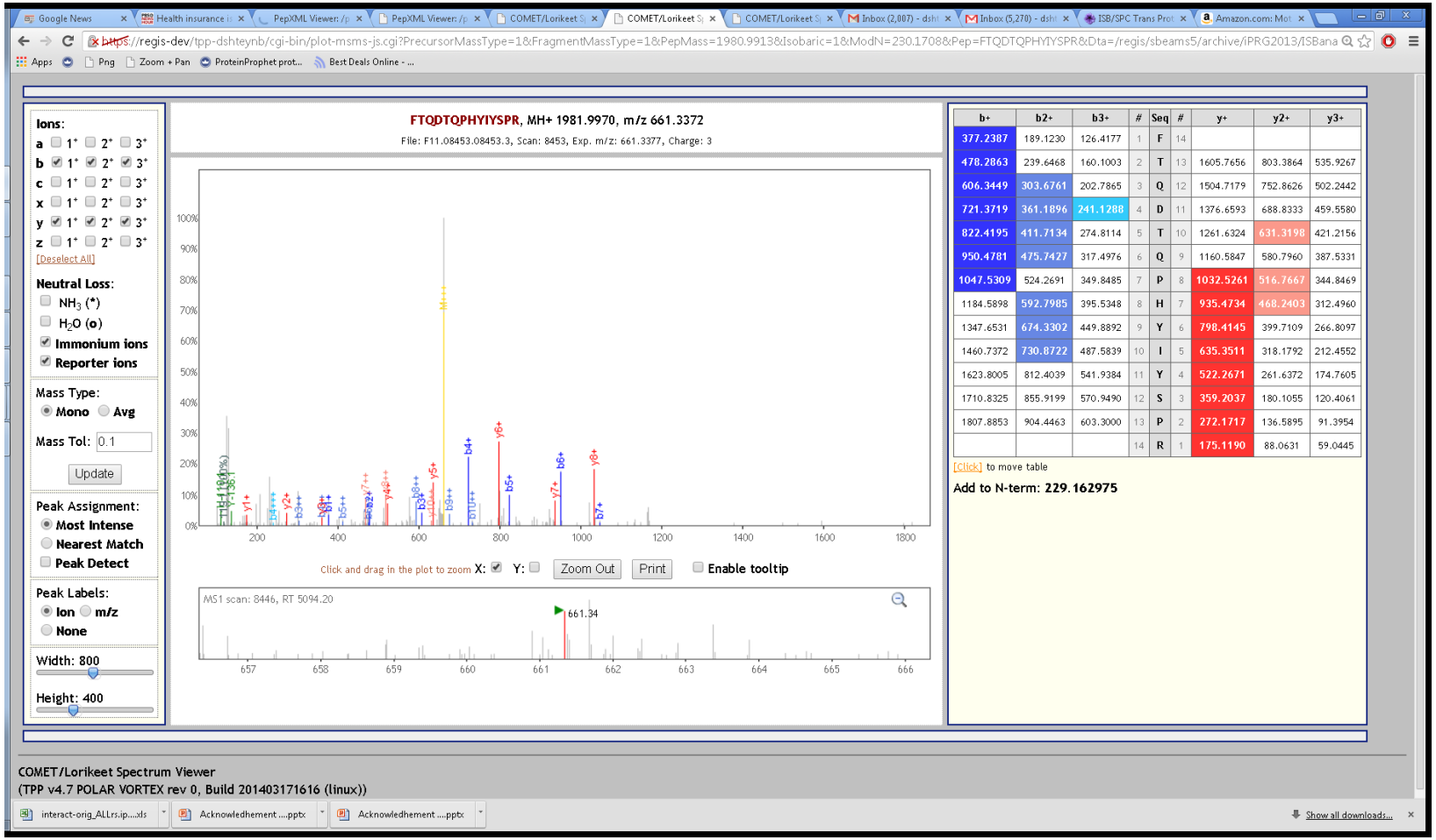
Search algorithms using standard, narrow mass tolerances and assumed charge states

TPP Statistical Tools are applied to identify the match probabilities

reSpect is attenuates the matched peaks of PSMs. The resulting spectra are saved in a new mzML/mzXML

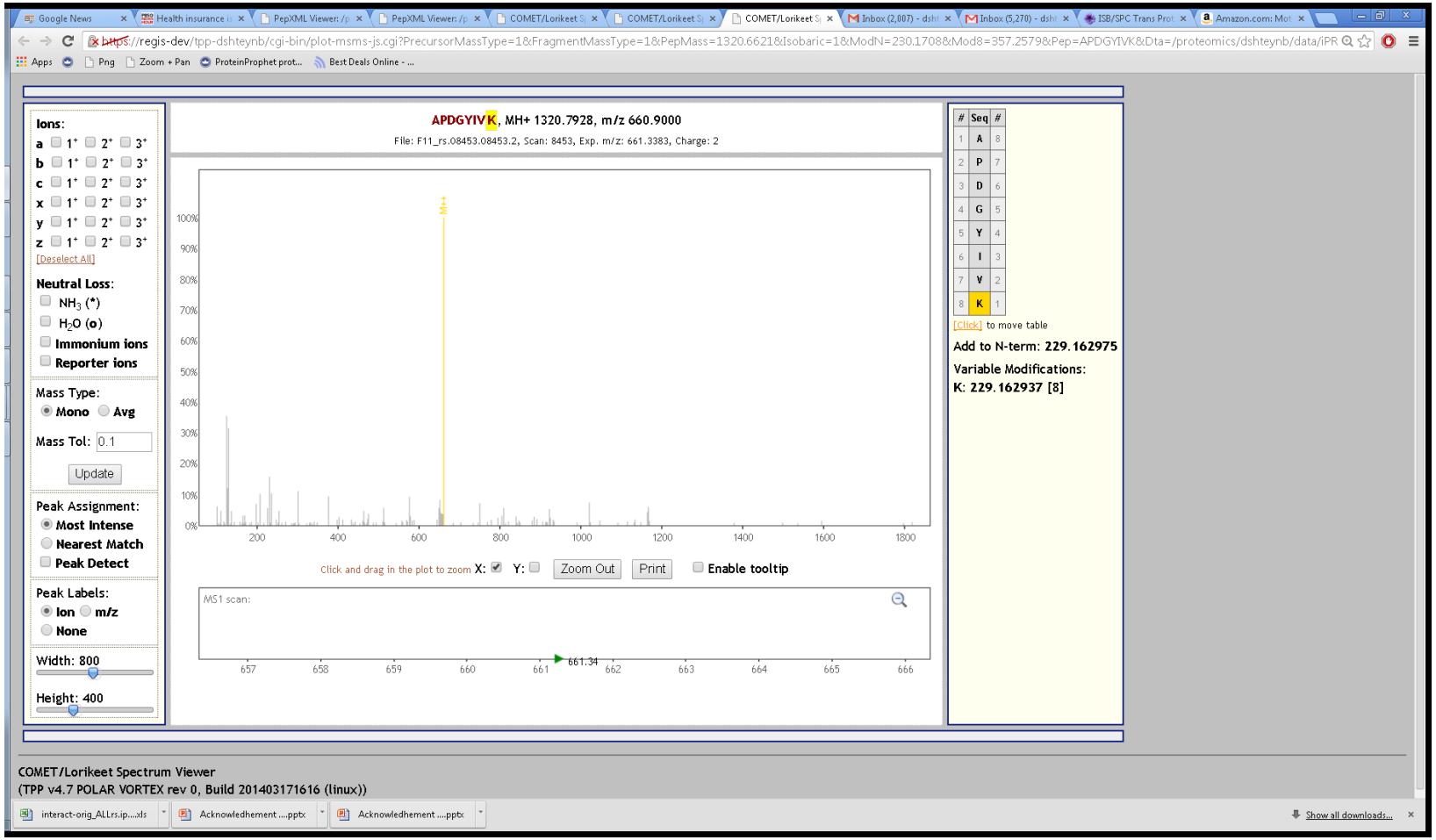
Searches repeated using wide mass tolerance and assuming all charges

reSpect: Attenuation Original *P = 0.999*

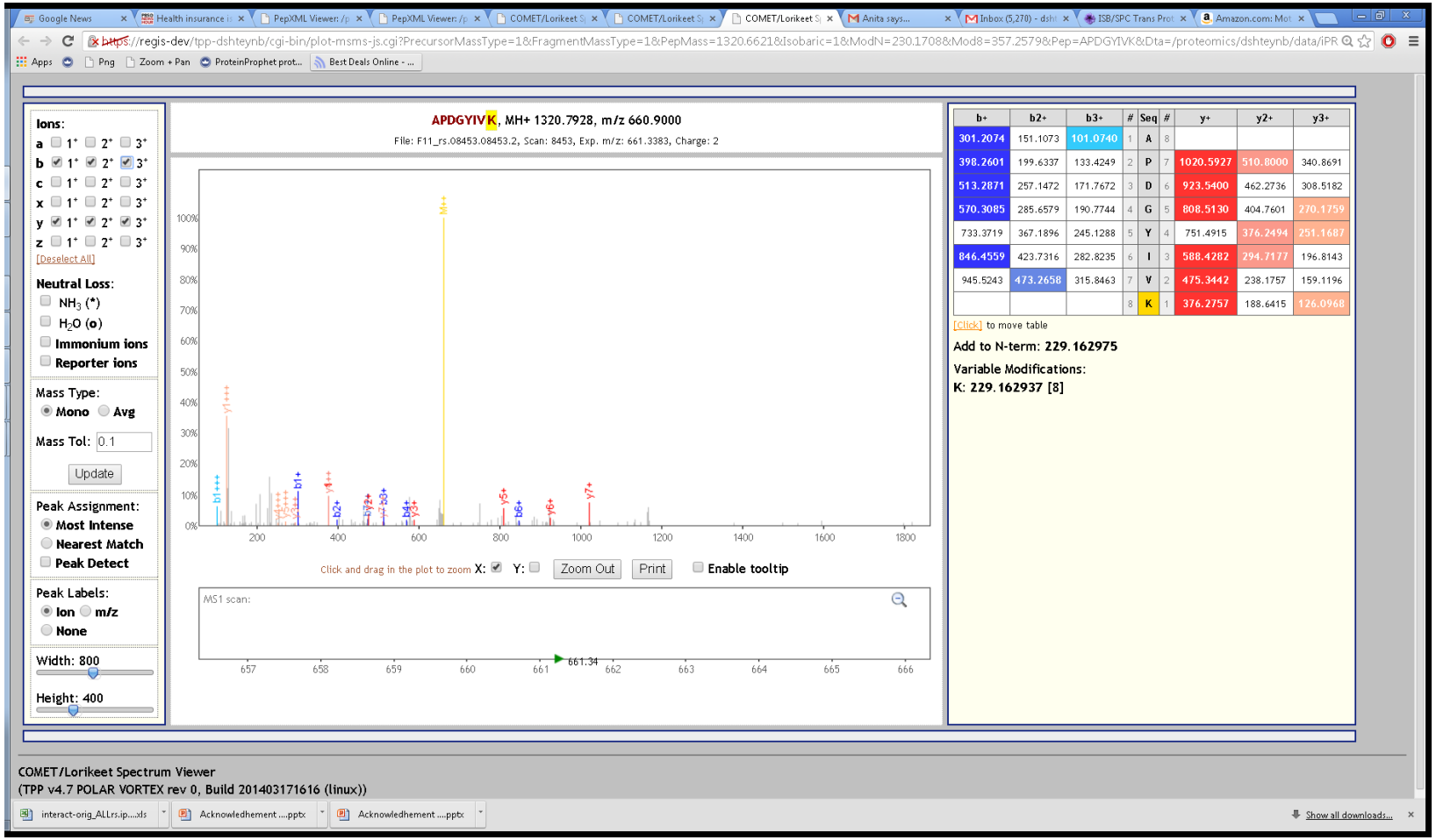


b+	b2+	b3+	#	Seq	#	y+	y2+	y3+
377.2387	189.1230	126.4177	1	F	14			
478.2863	239.6468	160.1003	2	T	13	1605.7656	803.3864	535.9267
606.3449	303.6761	202.7865	3	Q	12	1504.7179	752.8626	502.2442
721.3719	361.1896	241.1288	4	D	11	1376.6593	688.8333	459.5580
822.4195	411.7134	274.8114	5	T	10	1261.6324	631.3198	421.2156
950.4781	475.7427	317.4976	6	Q	9	1160.5847	580.7960	387.5331
1047.5309	524.2691	349.8485	7	P	8	1032.5261	516.7667	344.8469
1184.5898	592.7985	395.5348	8	H	7	935.4734	468.2403	312.4960
1347.6531	674.3302	449.8892	9	Y	6	798.4145	399.7109	266.8097
1460.7372	730.8722	487.5839	10	I	5	635.3511	318.1792	212.4552
1623.8005	812.4039	541.9384	11	Y	4	522.2671	261.6372	174.7605
1710.8325	855.9199	570.9490	12	S	3	359.2037	180.1055	120.4061
1807.8853	904.4463	603.3000	13	P	2	272.1717	136.5895	91.3954
			14	R	1	175.1190	88.0631	59.0445

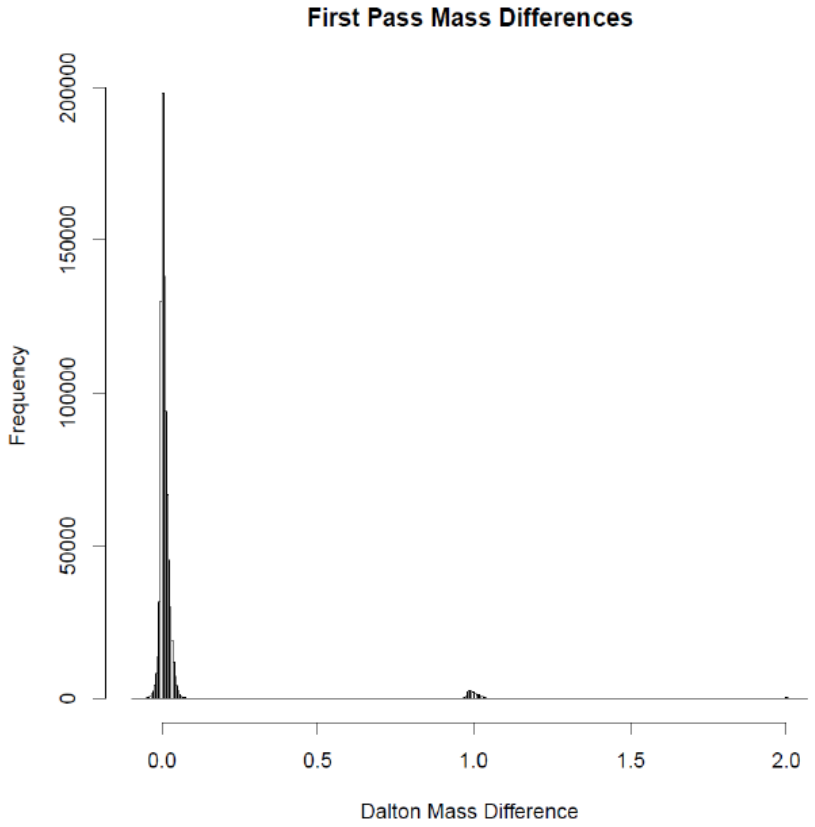
reSpect: Attenuation $I^{rs} = (1-P)*I^{orig}$

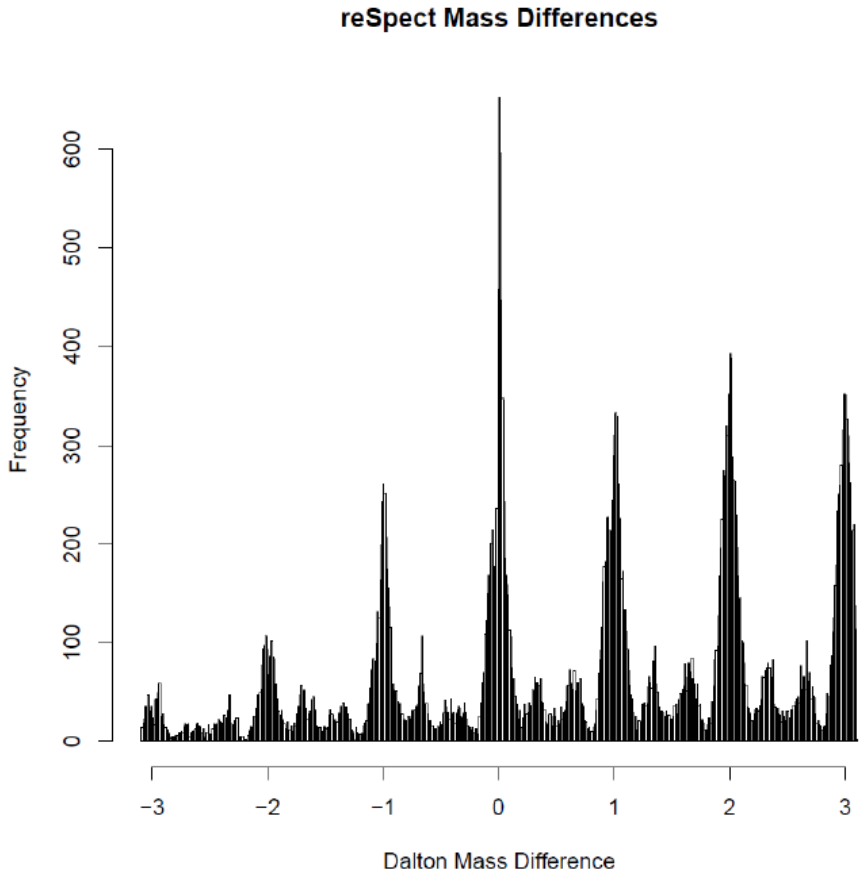


reSpect: Second Match *P = 0.995*



reSpect: Mass Differences

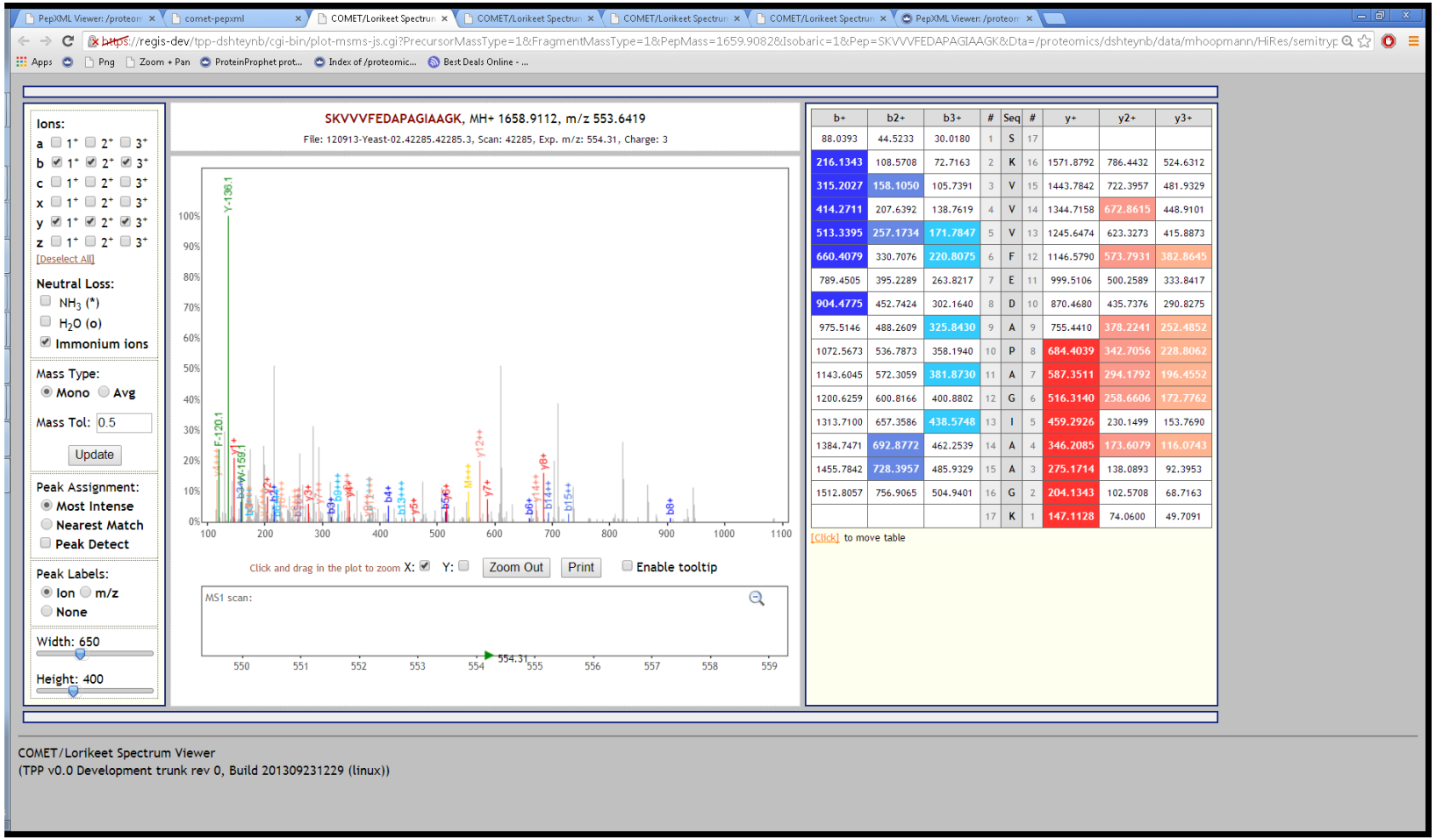




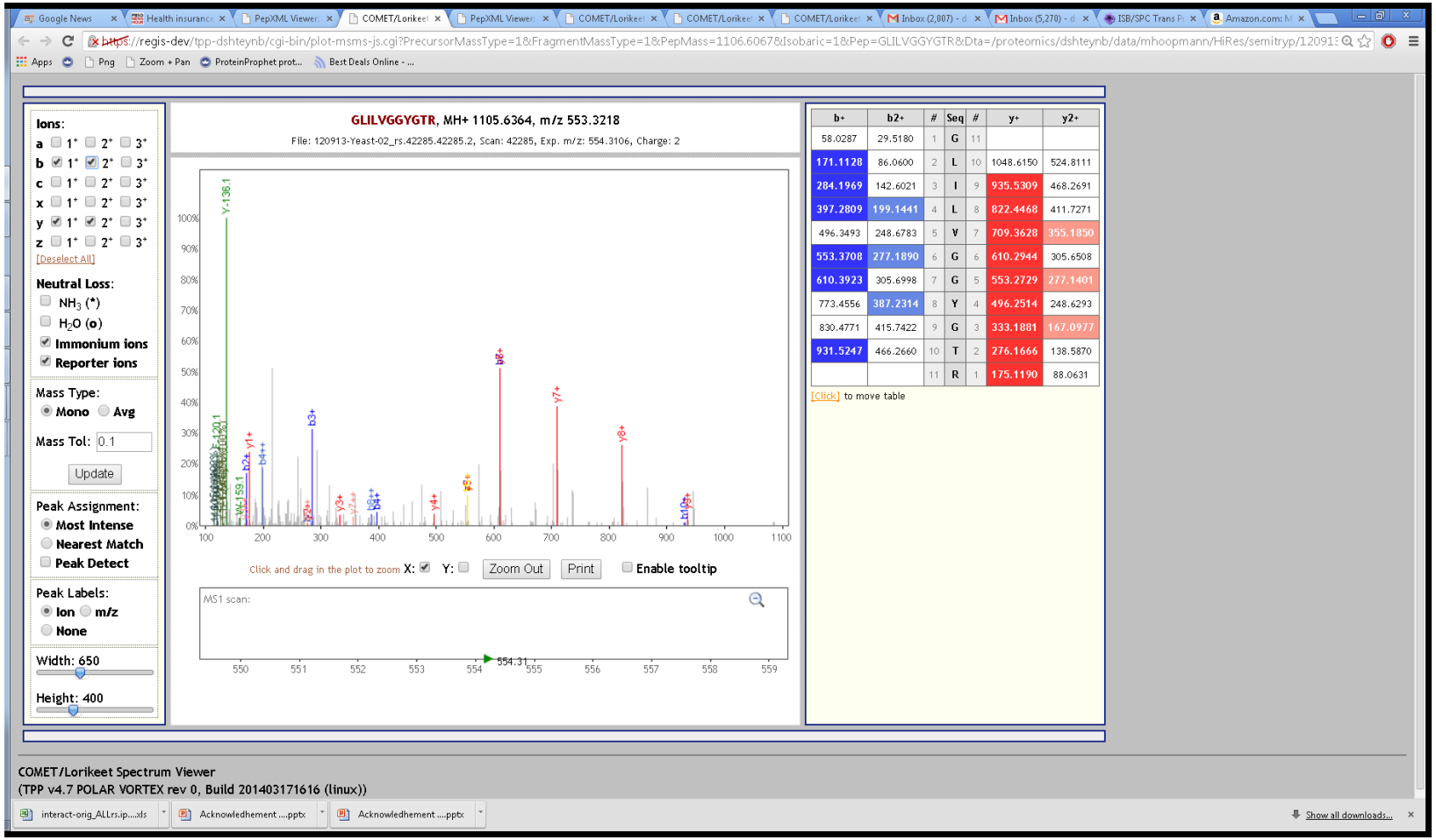
Chimeric Depth

- **Chimeric spectra may contain more than two precursors**
- **reSpect can be iteratively applied until novel PSMs are exhausted**

reSpect: The Original

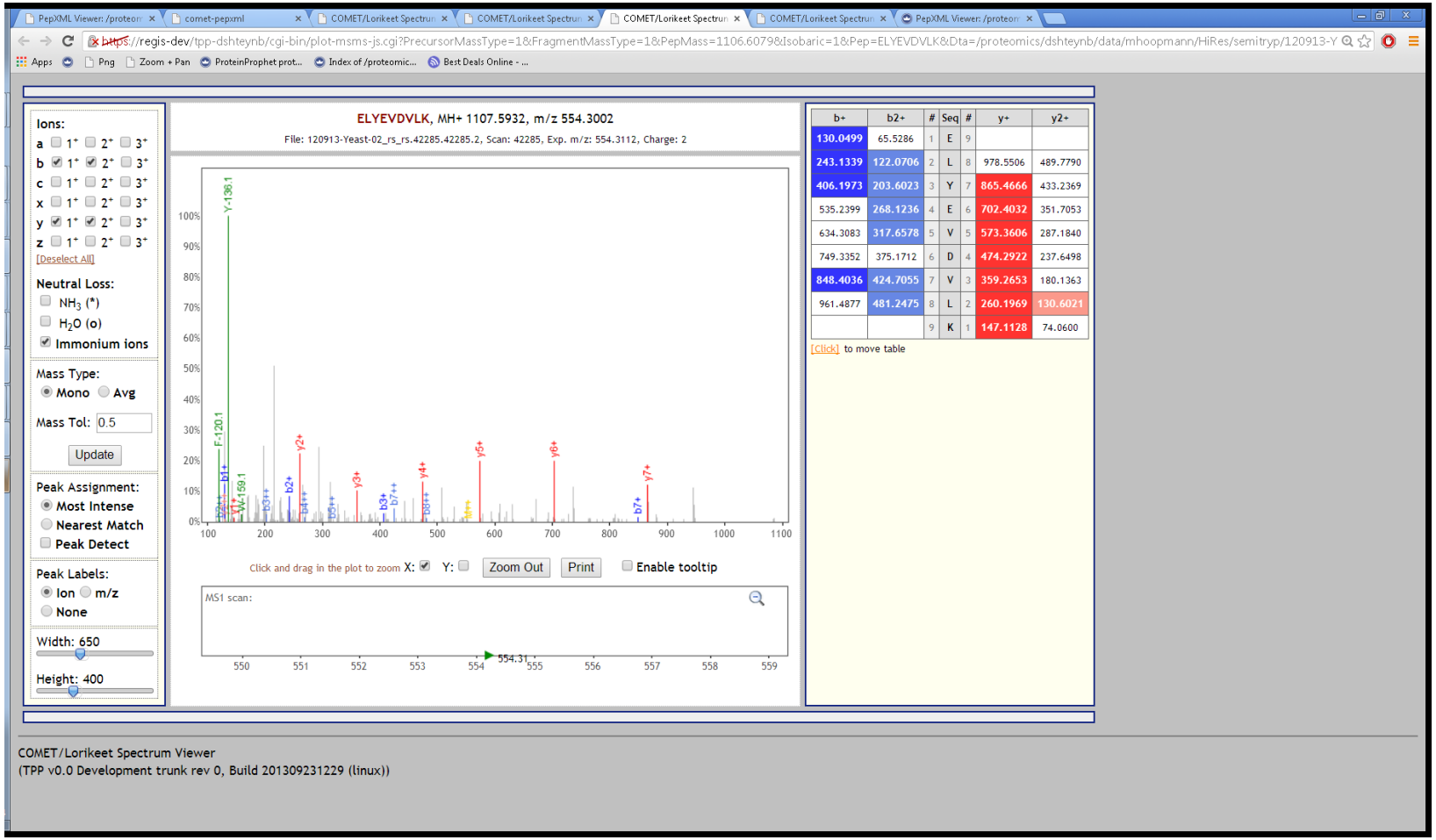


reSpect: The First Iteration

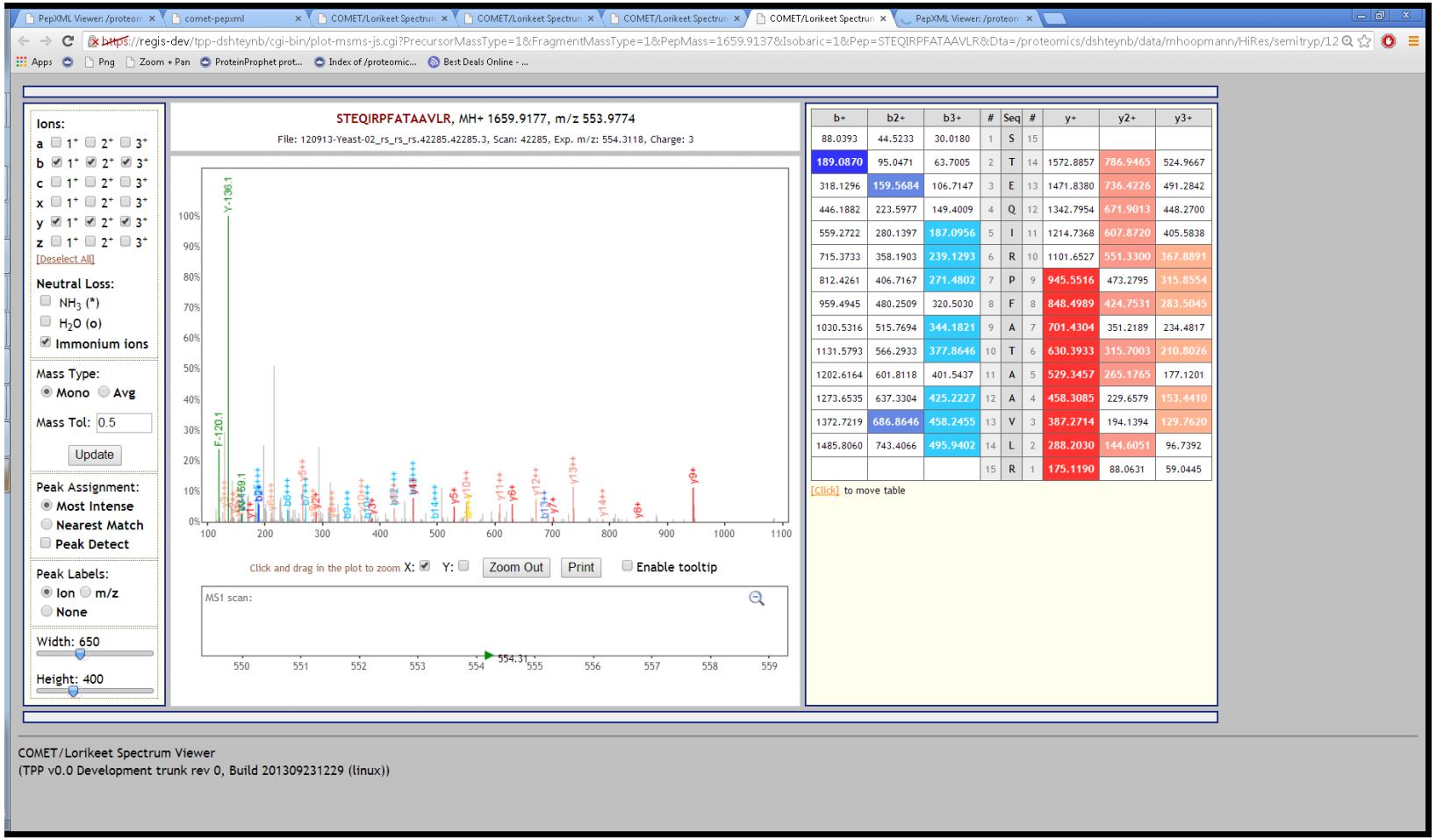


b+	b2+	#	Seq	#	y+	y2+
58.0287	29.5180	1	G	11		
171.1128	86.0600	2	L	10	1048.6150	524.8111
284.1969	142.6021	3	I	9	935.5309	468.2691
397.2809	199.1441	4	L	8	822.4468	411.7271
496.3493	248.6783	5	V	7	709.3628	355.1850
553.3708	277.1890	6	G	6	610.2944	305.6508
610.3923	305.6998	7	G	5	553.2729	277.1401
773.4556	387.2314	8	Y	4	496.2514	248.6293
830.4771	415.7422	9	G	3	333.1881	167.0977
931.5247	466.2660	10	T	2	276.1666	138.5870
		11	R	1	175.1190	88.0631

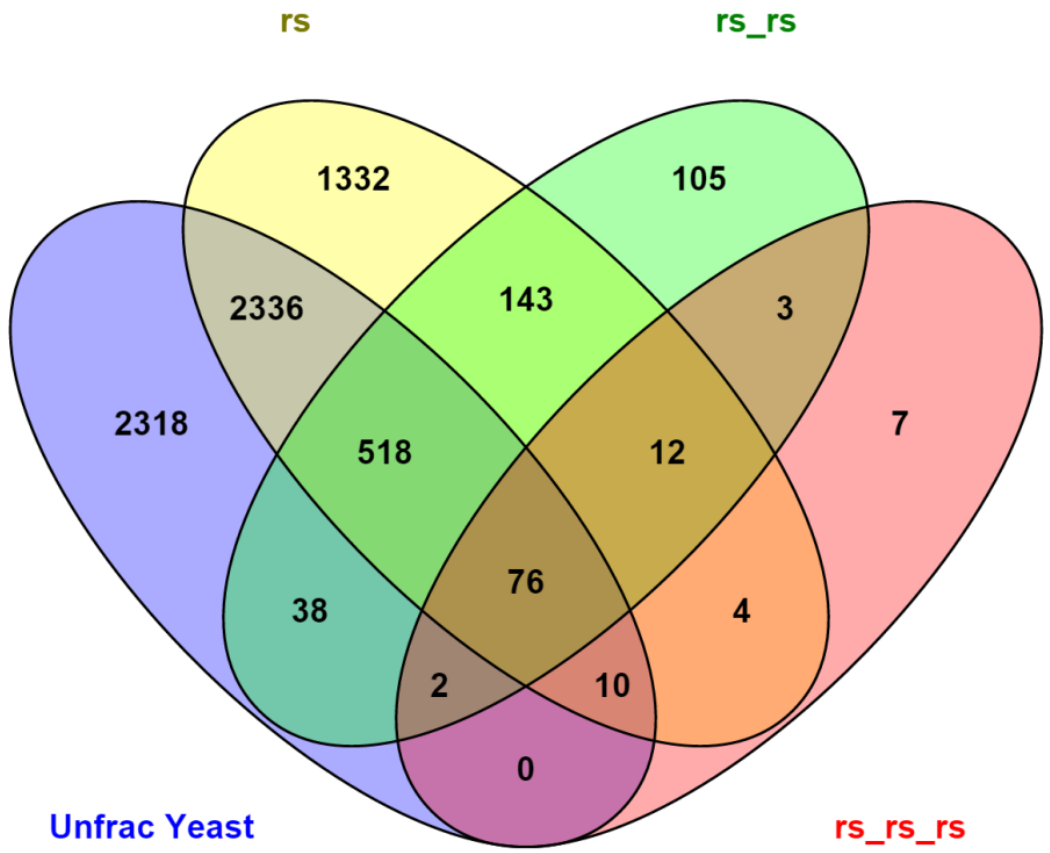
reSpect: The Second Iteration



reSpect: The Third Iteration

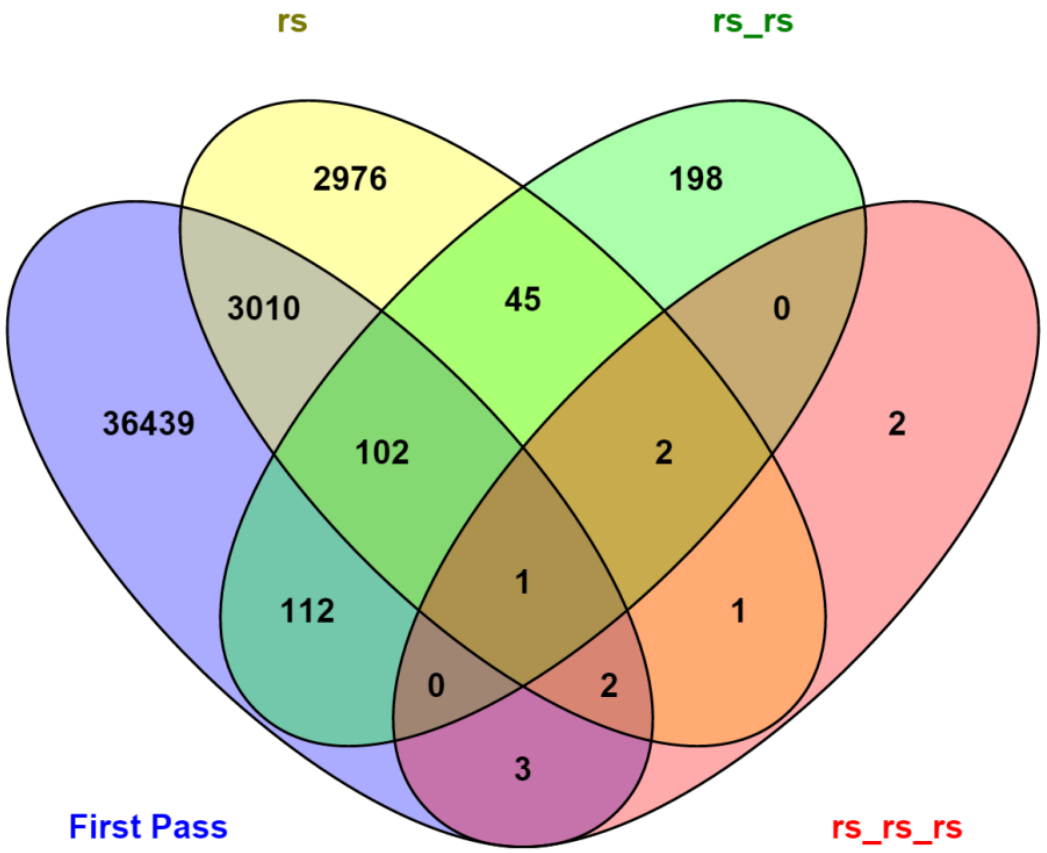


reSpect: No Fractionation



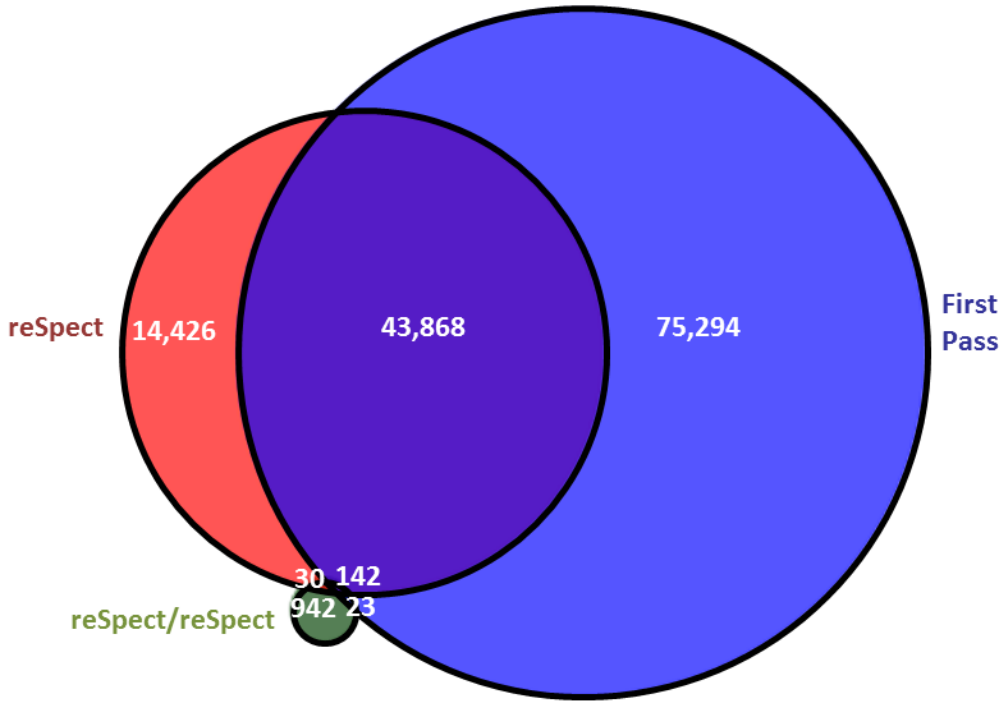
- 4 Replicate Yeast Q-Exactive Runs
- 347,000 spectra
- 30.3% Boost in Distinct Peptide Identification

reSpect: Medium Fractionation



- iPRG2013 study
- 118,000 spectra
- Orbitrap Velos, 14 RP fractions of whole cell lysate of human peripheral blood mononuclear cells
- 8.1% Boost in Distinct Peptide dentifications

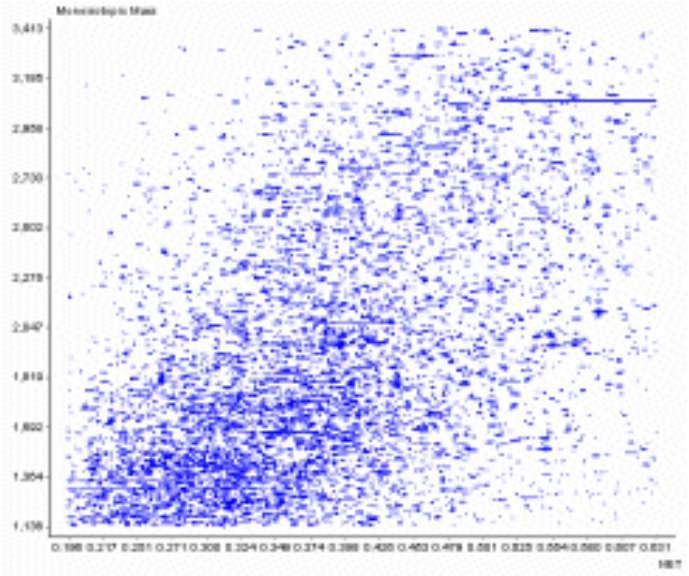
reSpect: High Fractionation



- **Human Cancer Cell Line**
- **Orbitrap Velos, 48 fractions**
- **12.9% Boost in Distinct Peptide Identifications**

Different Datasets, Different Chimeras

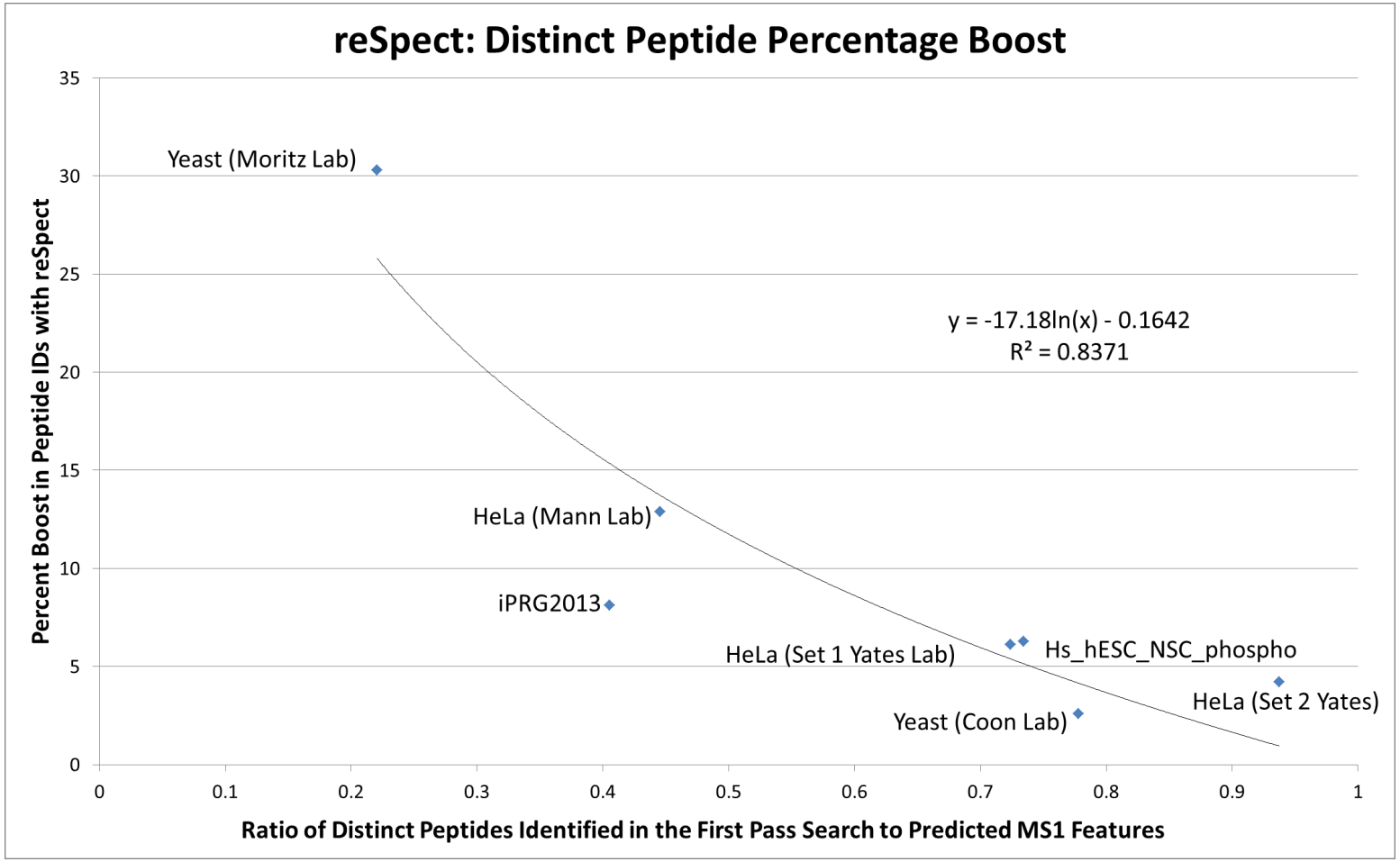
- **Sample complexity and instrument settings have largest impact on chimeric spectra**



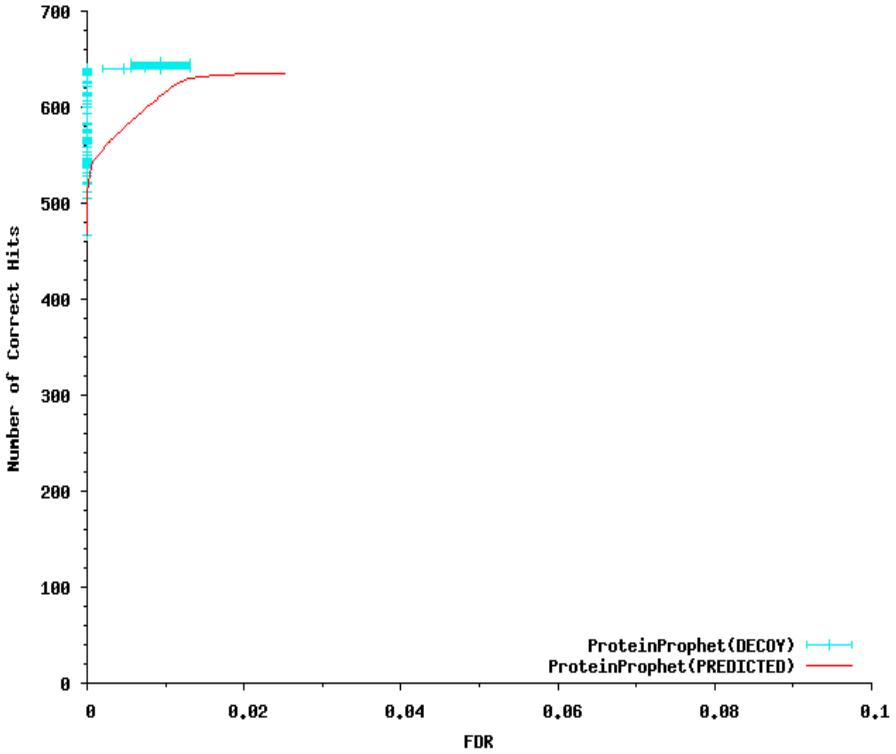
Precursor analysis algorithms can estimate number of precursors to solve.

These numbers correlate with PSMs identified.

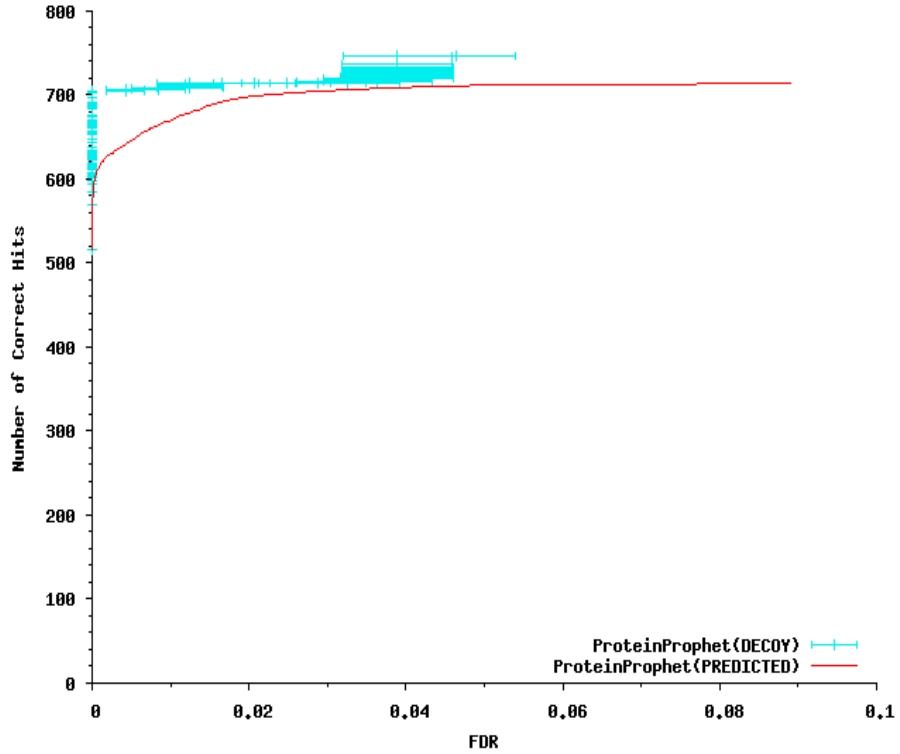
reSpect: Different Datasets



reSpect: Identifies More Proteins

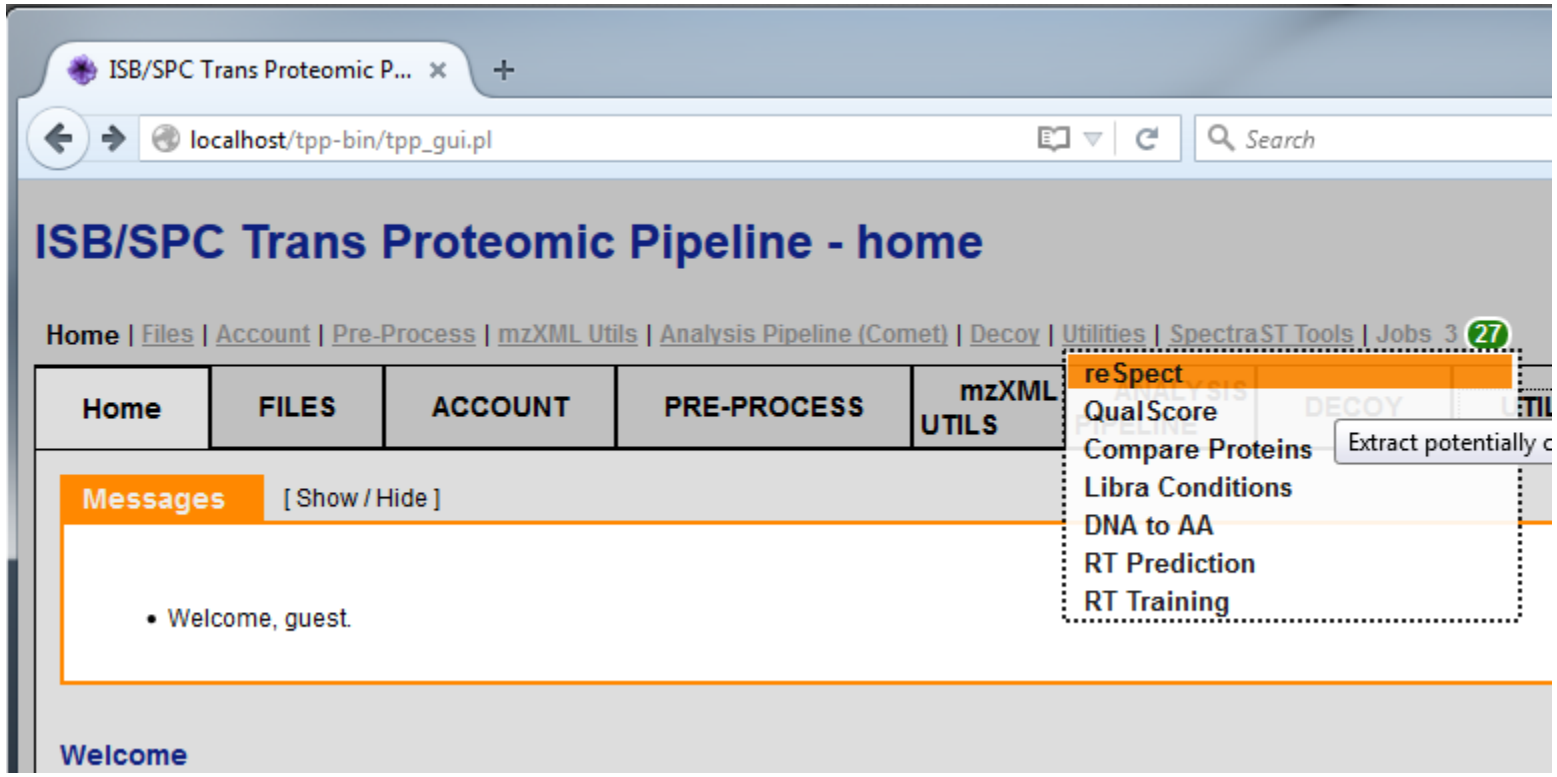


Protein level performance on the Moritz Lab yeast dataset **BEFORE reSpect analysis**.

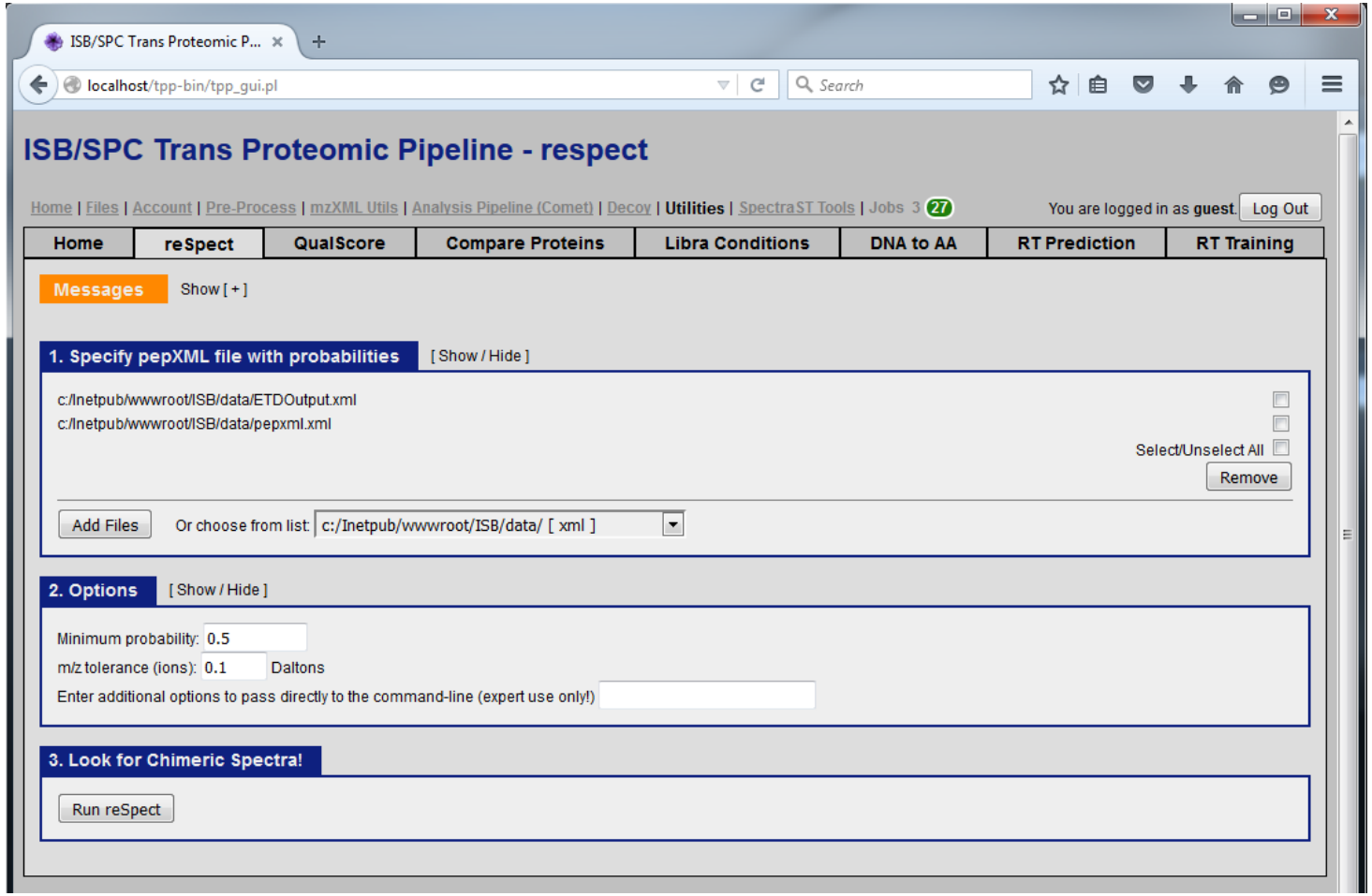


Protein level performance on the Moritz Lab yeast dataset **AFTER reSpect analysis**.

reSpect in the TPP



reSpect in the TPP



reSpect: Summary

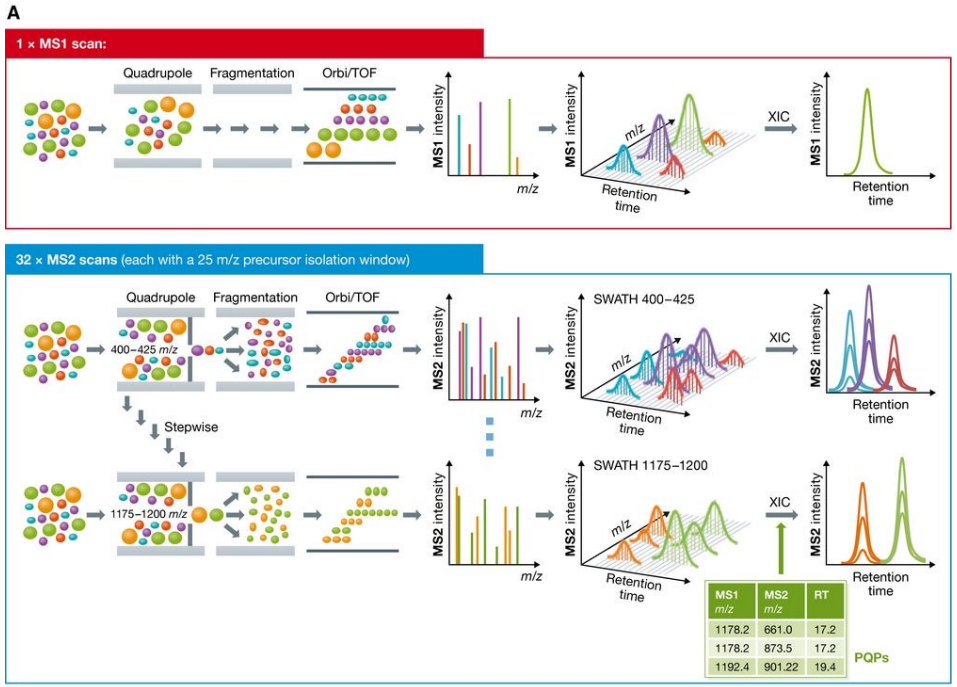
- **Explained peaks are *attenuated* according to the quality of the match amongst medium and high scoring spectra**
- **Attenuated spectra are *researched* with wide mass window and assuming all charges**
- **Uses *same validation* with PeptideProphet and iProphet to estimate error rates**
- ***Automated* within the TPP framework**
- **Identifies new peptide sequences even in highly fractionated data**
- **Works best on *high mass accuracy* MS^2 data**
  - Q-Exactive
  - Orbitrap Velos
  - QTOF
- **Also works nicely on *standard mass accuracy* MS^2 data**
  - LTQ

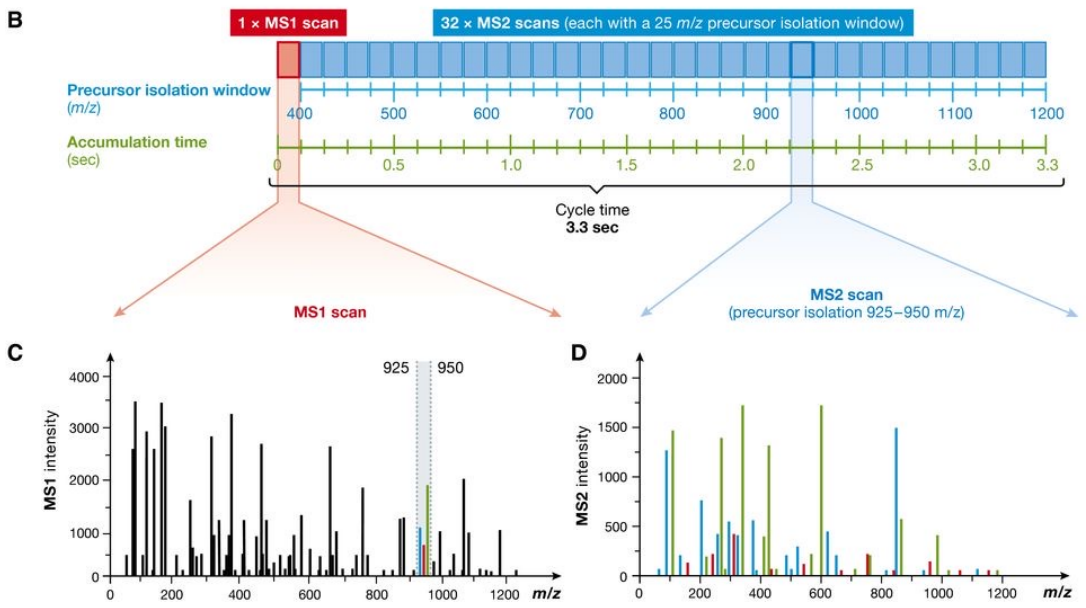
Principles of Data Independent Acquisition

Slides courtesy of Dr. Mukul Midha

ISB

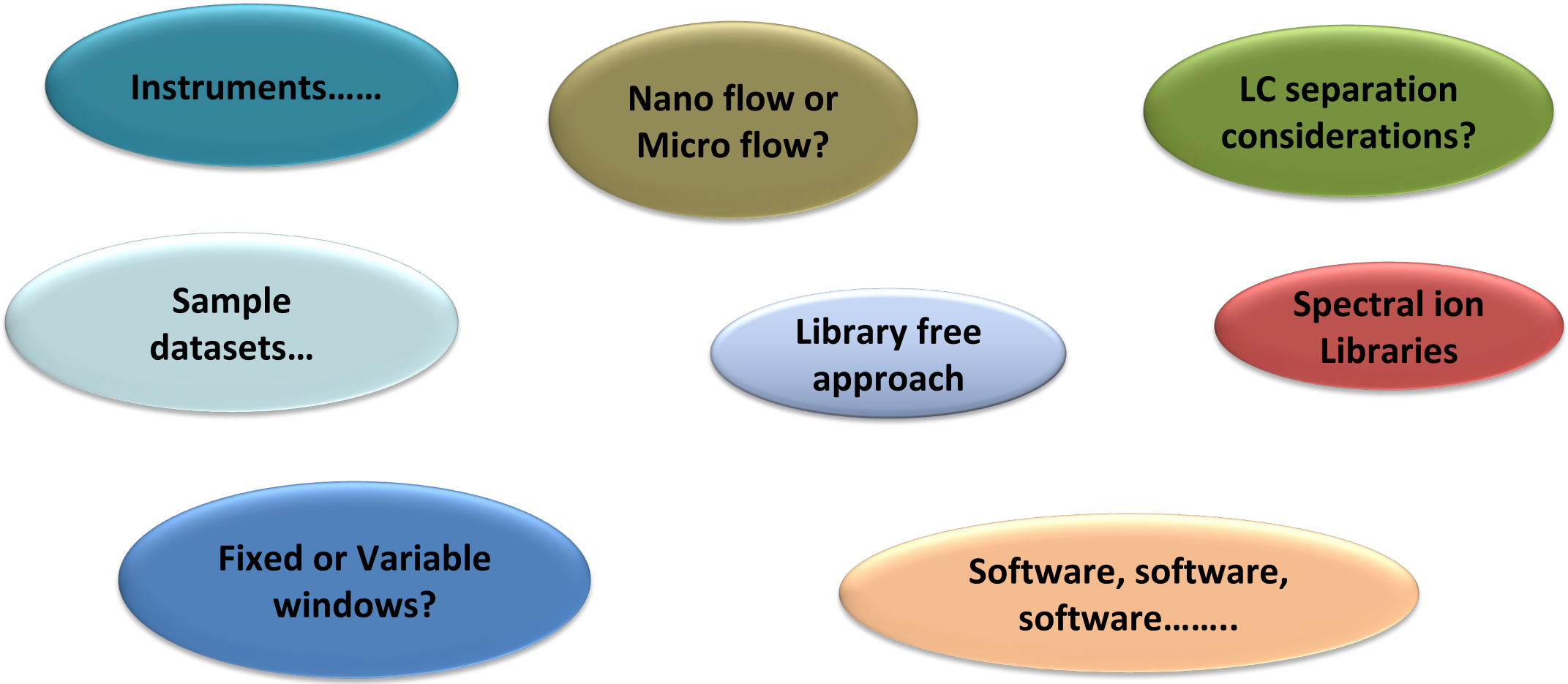
Principle of sequentially windowed acquisition in DIA/SWATH-MS





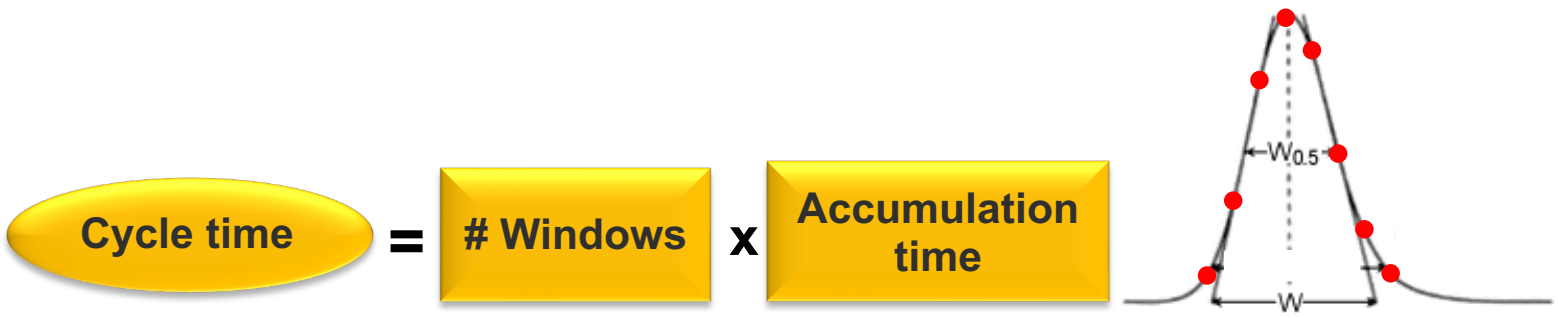
Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial, Ludwig, C. et al Molecular Systems Biology 2018

SWATH-MS/DIA experiment setup…..So much to choose from!!

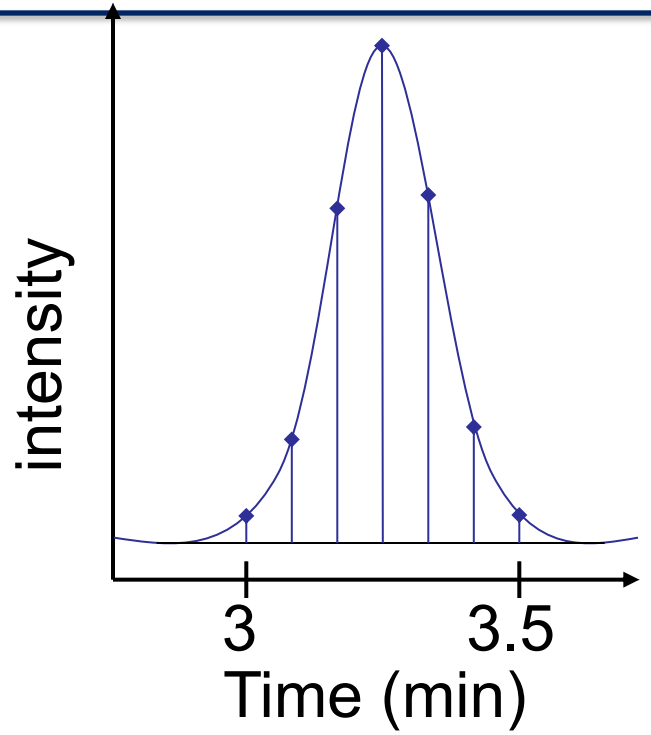


What are the critical acquisition attributes for DIA/SWATH-MS?

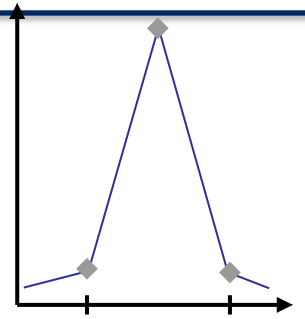
- ✓ High resolution MS/MS
- ✓ Cycle Time
- ✓ Q1 Isolation Windows
- ✓ Dynamic Range

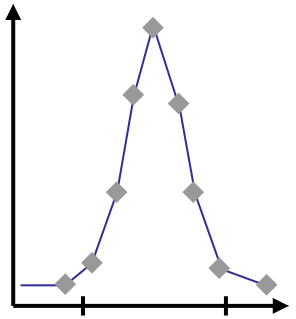


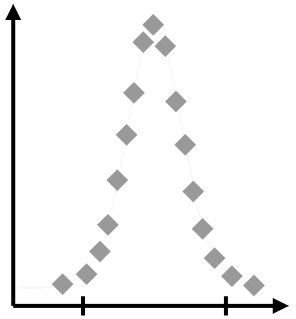
DIA Acquisition- Quantitation



ex. 30 s peak width at base
3 s cycle will collect 10 DPPP
<3 s will over sample
>3 s will under sample







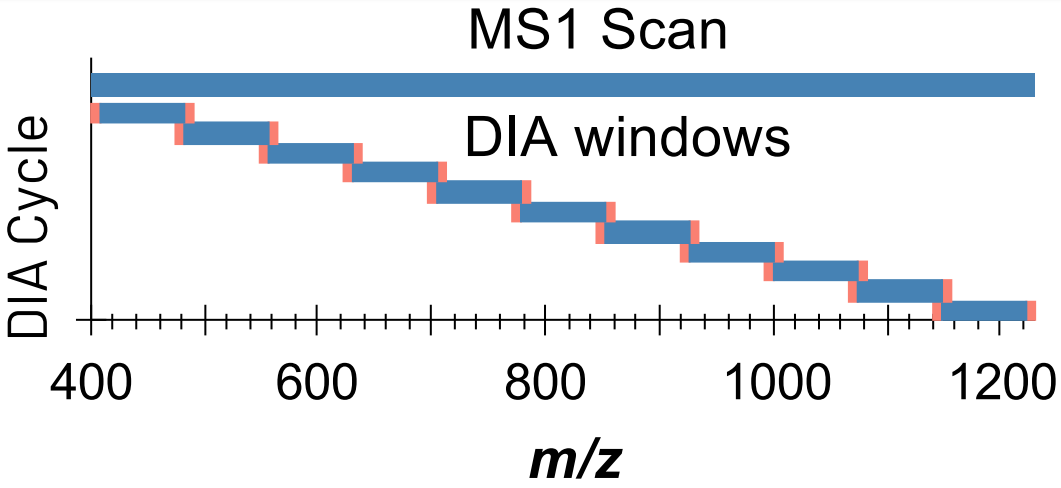
Data Points Per Peak (DPPP)

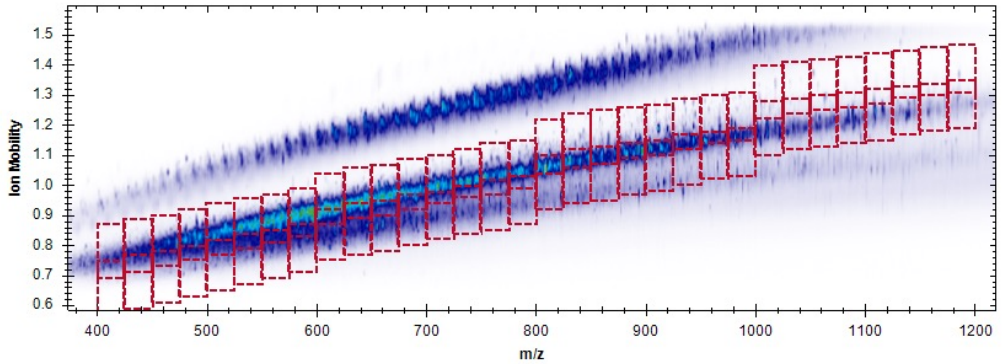
<7 DPPP = under sampling

7-10 DPPP = optimal sampling

>10 DPPP = over sampling

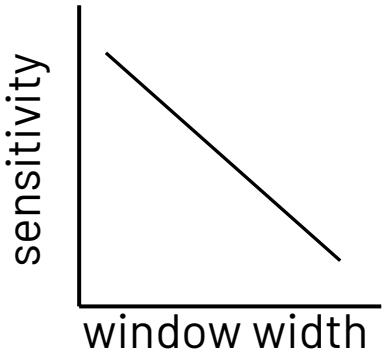
DIA/SWATH-MS precursor isolation windows

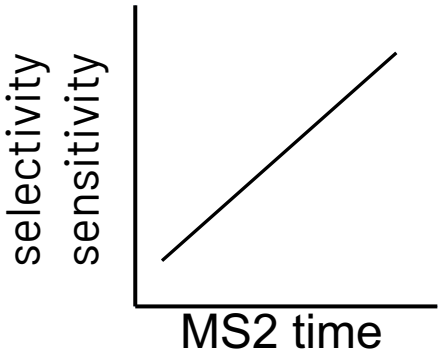




windows x MS2 acquisition time = cycle time

42 (20 *m/z* width, 400-1200 *m/z*, 1 Da overlap) x 60 ms = 3.6 s

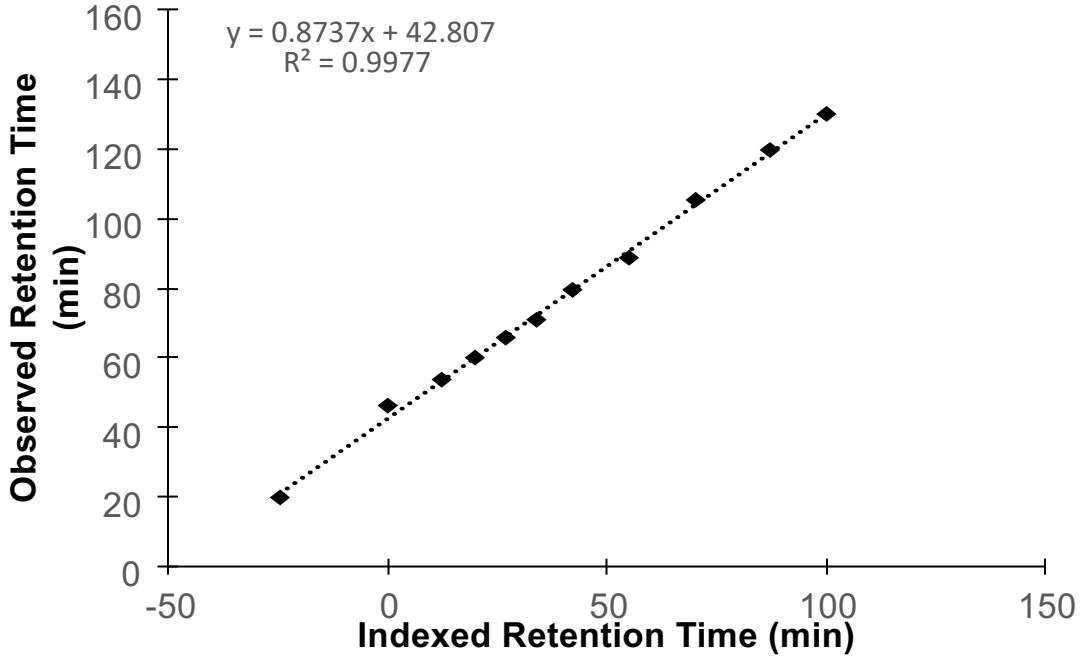


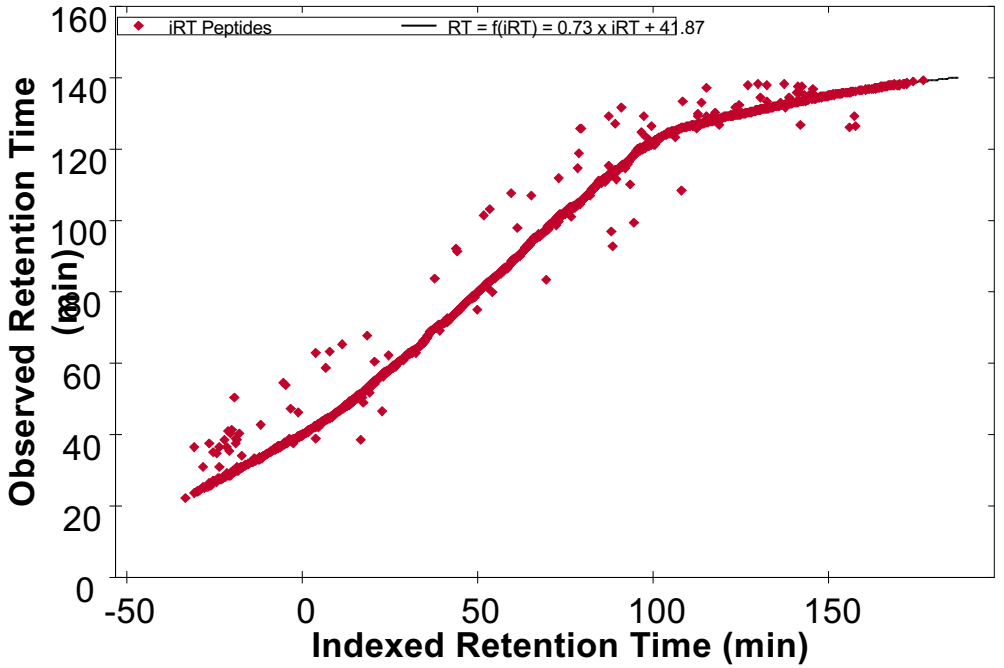


Choices are dependent on chromatography, application and platform.

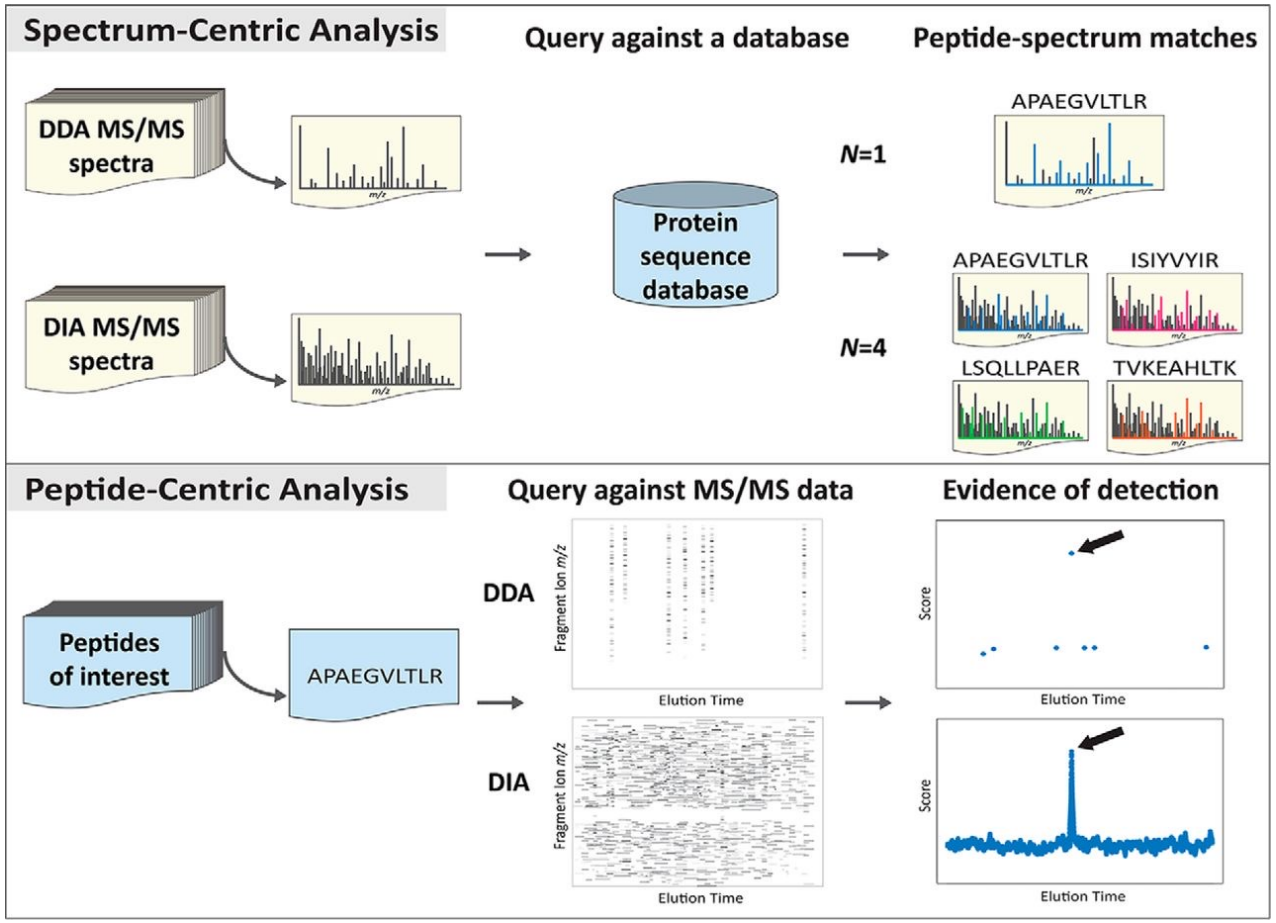
DIA data analysis – retention time normalization

- iRT spike
  - set of non-endogenous peptides
  - Used to convert to iRT scale (linear)
- Observed peptides/features
  - High-precision iRT (non-linear)
  - Anchor points





DIA/SWATH-MS: Library based and library free approach

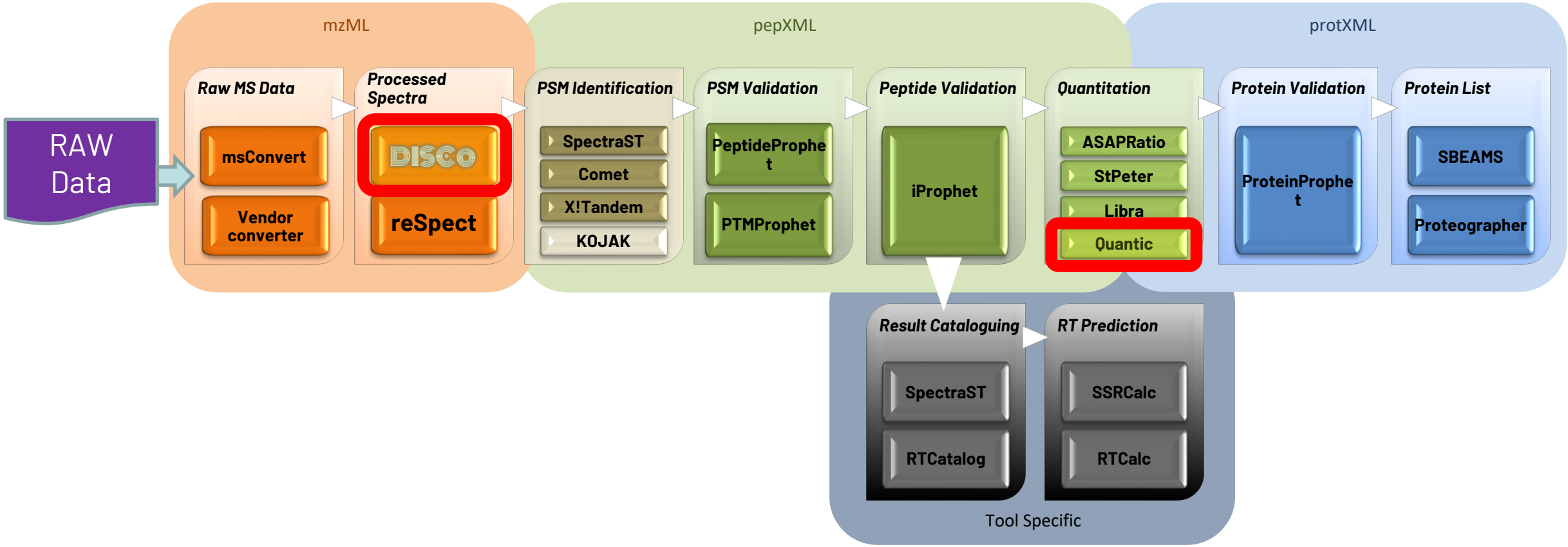


Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data, Ying S. Ting et al , MCP, 2015

DISCO: TPP Software for Data-Driven DIA Analysis and Quantitation

ISB

Trans-Proteomic Pipeline a.k.a. TPP



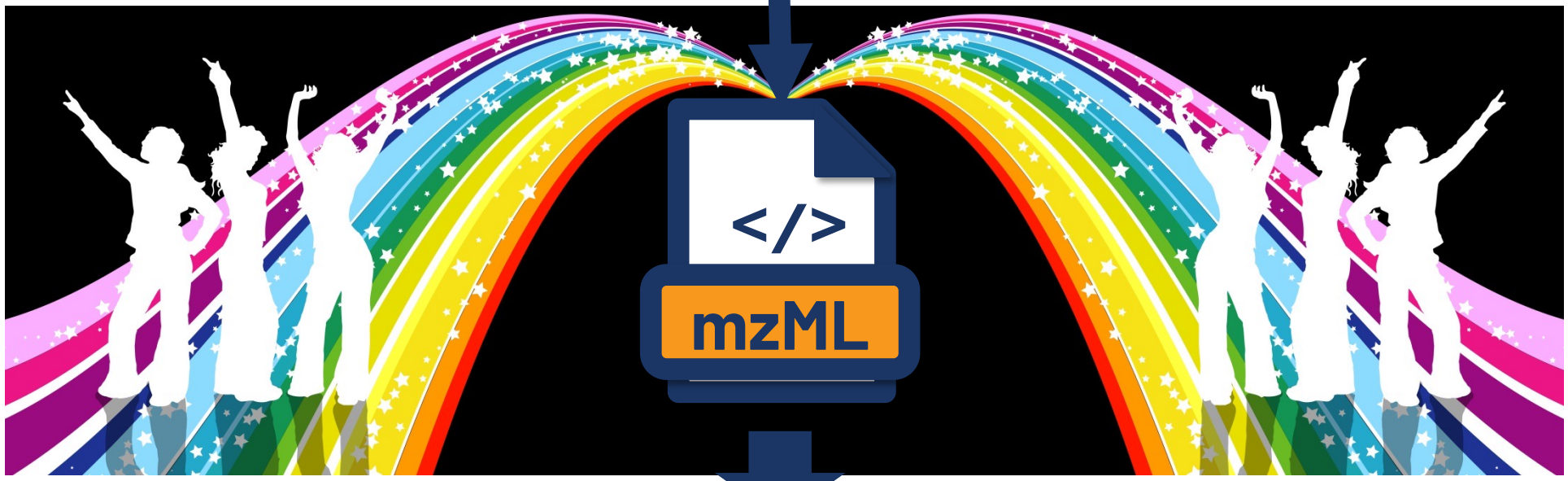
HardKlör Configuration File

DIA Windows

mzML Data

List of DIAUmpire Targets

DISCO



Highlights

- Works on timsTOF PASEF DIA data
- DIA Windows is not required if isolation_window info available
- Multiple scan windows sizes for data point extraction
- Minimum amount of correlation required to select fragments depends on the scan window size (number of data points)
- Kernel Density smoothing of raw signals
- Better utilization of multi-threading resources
- Fail-safe mzML file reading

```
USAGE: DiscoFilter <OPTIONS> <mzML/mzXML_input_file>

OPTIONS:
        TARGETS=<targets_file>   Use specified <target_file> to look for MS1 features  (default: use HardKlör
internally)
        UMPTARGETS=<targets_file>        Use specified <target_file> to look for MS1 features  (default: use
HardKlör internally)
        HKCONFIG=<config_file>   Use specified <config_file> for Hardklör(default: use Hardklor.conf)
        WINDOWS=<DIA_windows_file>
        MAXSCANWINDOW=<number>:  Use specified maximum number of scans over which to track (default=17). This
option applies only when the UMPTARGETS are not specified, when UMPTARGETS are specified the SCANWINDOW is set
to the maximum peak width defined in the UMPTARGETS file.
        MINSCANWINDOW=<number>:  Use specified minimum number of scans over which to track (default=7). This
option applies only when the UMPTARGETS are not specified, when UMPTARGETS are specified the SCANWINDOW is set
to the maximum peak width defined in the UMPTARGETS file.
        AVGSCANHALF=<number>:    Use specified number of +- scans over which to average MS1 scans prior to
feature detection, set to 0 for no averaging. (default=0)
        MZPREC=<number>:         Set mz precision in 'points after the decimal', used for binning and
averaging.  Applies in combination with AVGSCANHALF= greater than 0.  (default=2.1)
        MININTENS=<number>:      Filter out peaks below minimum intensity factor in each spectrum, set as a
non-negative number (default=off)
        MAXPPM=<number>:         Maximum Allowed PPM signal offset best defined in powers of 2 (e.g. 2, 4, 8,
16, etc. default=32)
        PPMFWHM=<number>:        Full Width at Half Maximum expected for Mass PPM profile of peaks (default=16)
        IMFWHM=<number>:         Full Width at Half Maximum expected for Inverse Reduced Ion Mobility profile
of peaks (default=0.05)
        IONMOBBINS=<number>:     Set number of bins to partition Ion Mobility. (default=10)
        SUFFIX=<string>:         Set suffix for output file (default='_ds')
        THREADS=<number>:        Use specified number of threads (default=1).
        STARTSCAN=<number>:      Starting scan to process (default=1).
        ENDSCAN=<number>:        Ending scan to process (default: process until the end of the run).
```

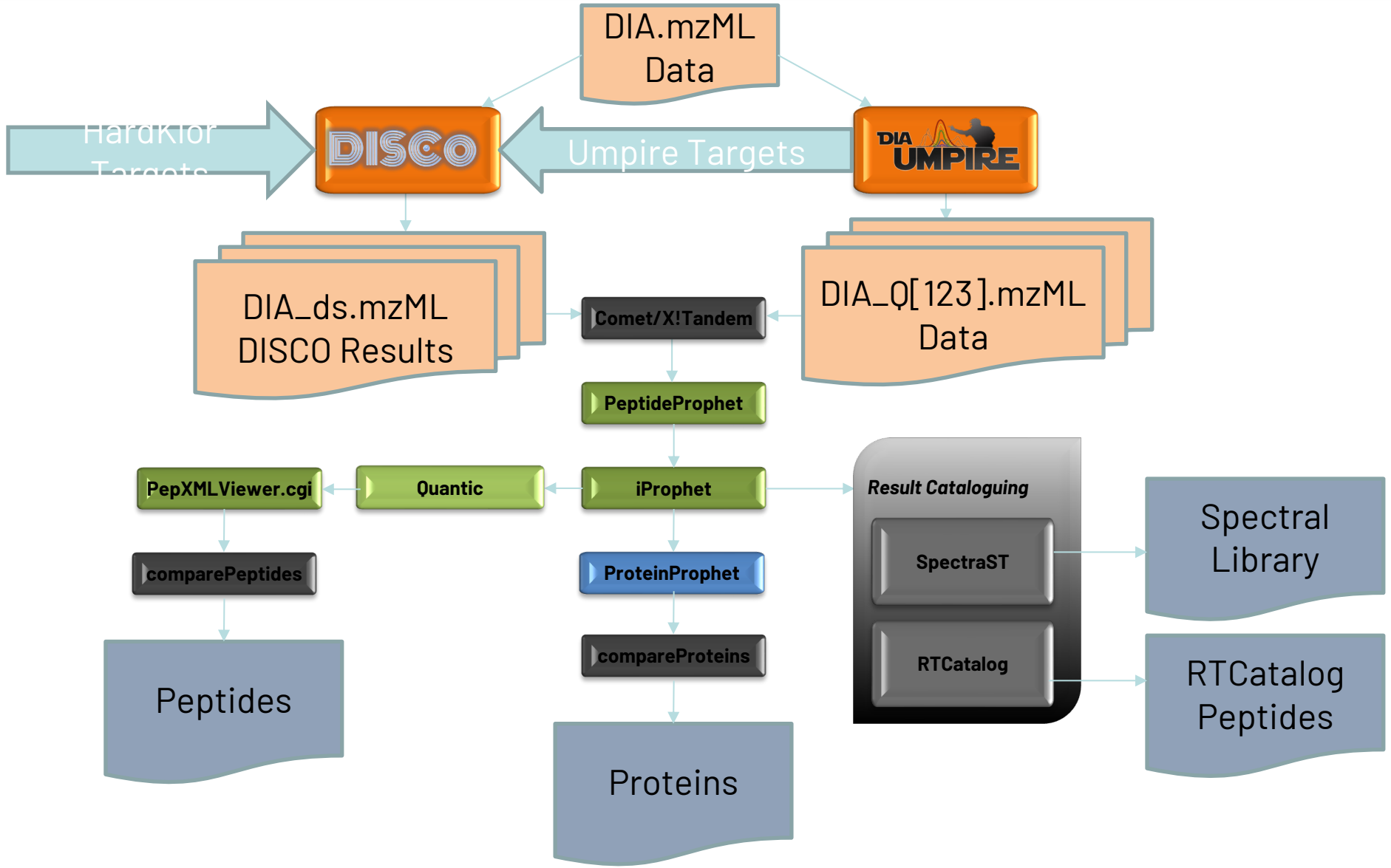
Algorithm

- DISCo matches identified **precursor** signals to fragment signals by looking for those that have a similar shape in Retention Time Space
- Fragment signals are sorted according to the following criteria, in order:

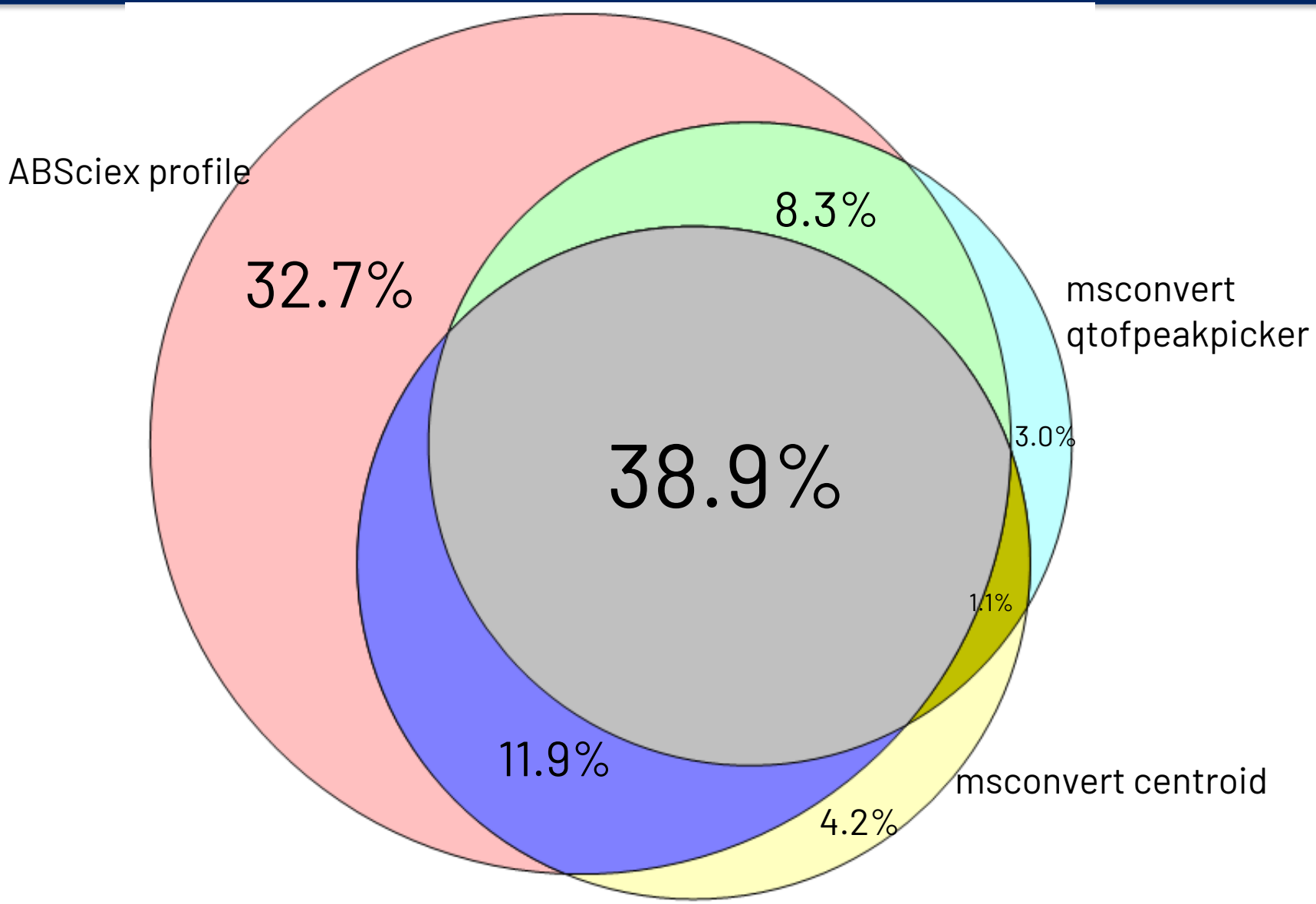
  entropy → distance → correlation → intensity

- Type I Error Rate ***α*** and Type II Error Rate ***β***, ***N*** is sample size (points across peak)
- For each MS^1 target the following MS^2 peaks are selected:
  - Up to 1000 peaks with correlation ≥ ***rMax*** *(α = 0.01, β = 0.05)*
  - Up to 500 peaks with correlation ≥ ***rMid1*** *(α = 0.03, β = 0.1)*
  - Up to 200 peaks with correlation ≥ ***rMid2*** *(α = 0.05, β = 0.2)*
  - Up to 100 peaks with correlation ≥ ***rMin*** *(α = 0.1, β = 0.333)*
  - Peak intensities are scaled by the correlation

Data-Driven DIA Workflow



For TTOF 6600 Converter Choices Matter!

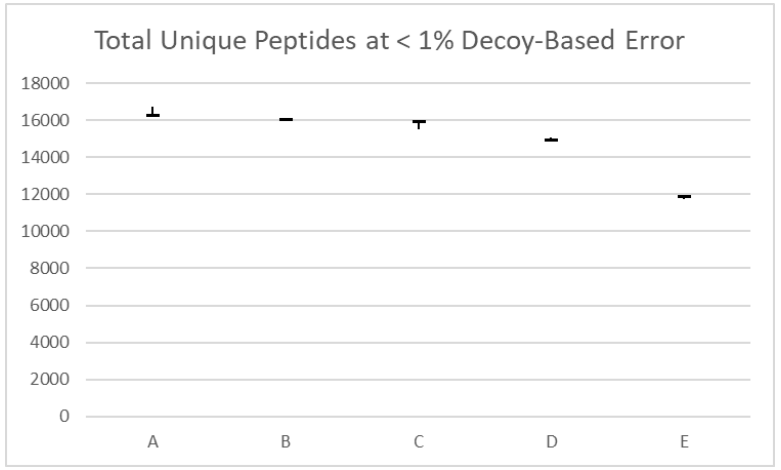


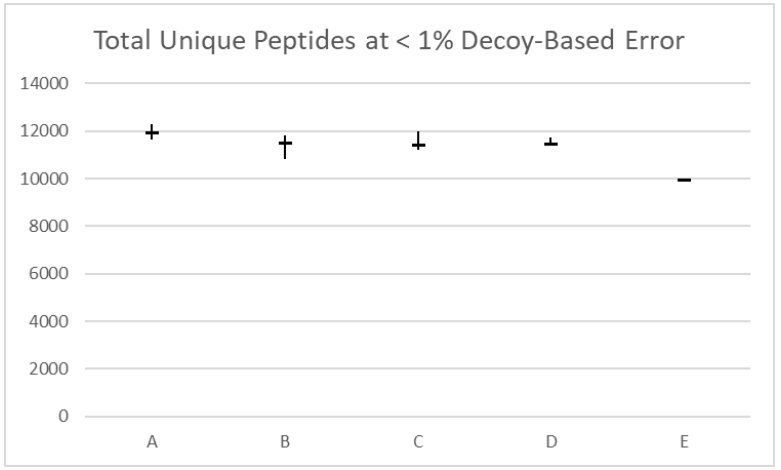
Benchmarking HeLa/Halo Dataset

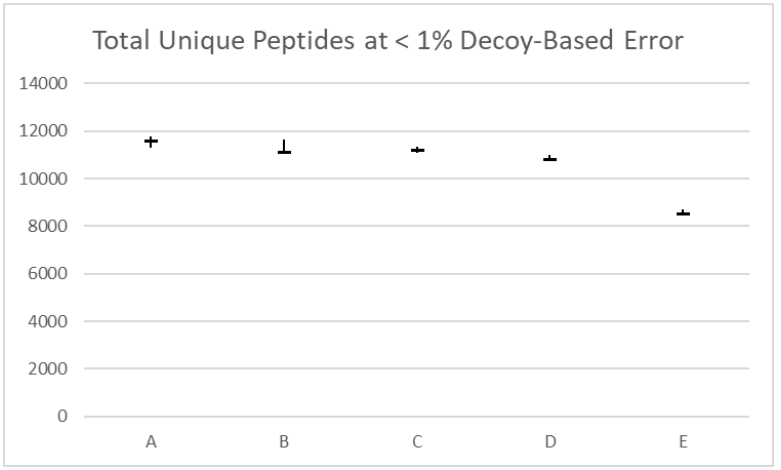
Sample	HeLa Background (ug)	Halo Variable (fmol)	Halo E/x Ratio	Halo log2 (E/x)
A	3.52	1.56	256	E/A= 8
B	3.52	6.25	64	E/B= 6
C	3.52	25	16	E/C= 4
D	3.52	100	4	E/D= 2
E	3.52	400	1	E/E= 0

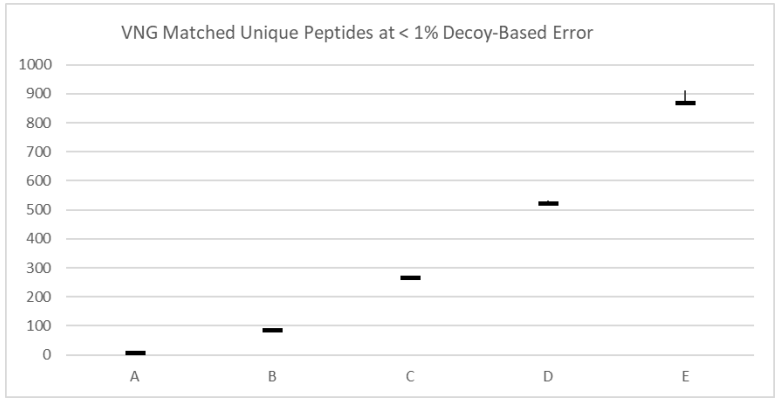
Hela/Halo TTOF6600 :

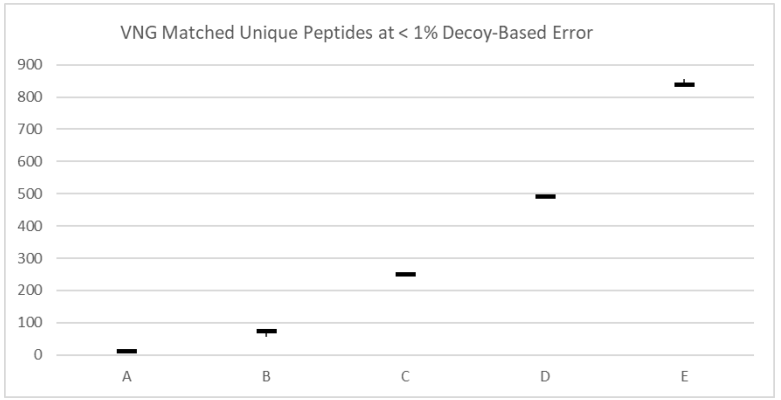
DISCO – DIAUmpire – DISCO w/ Umpire Targets

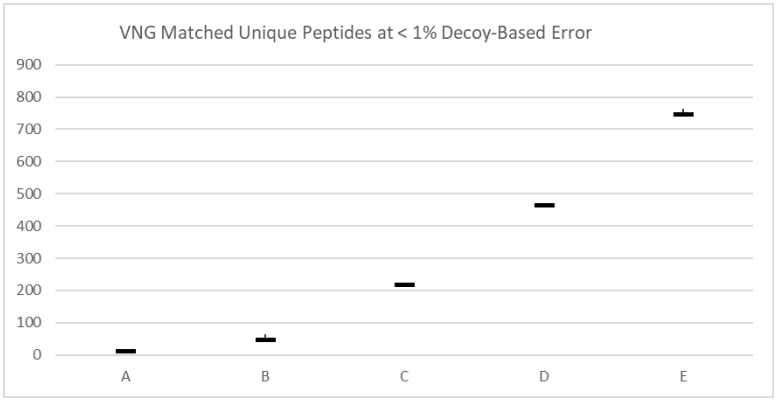




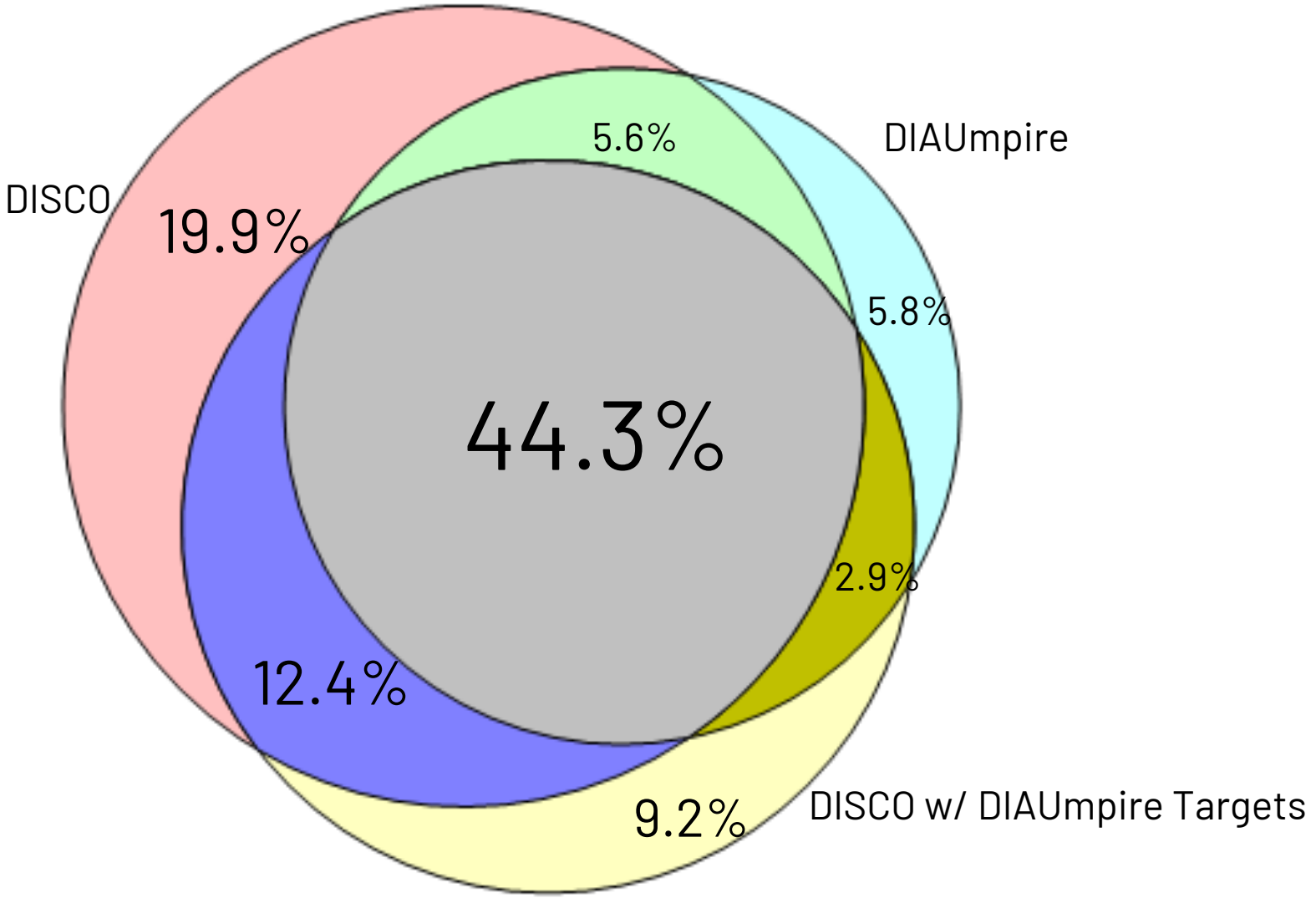








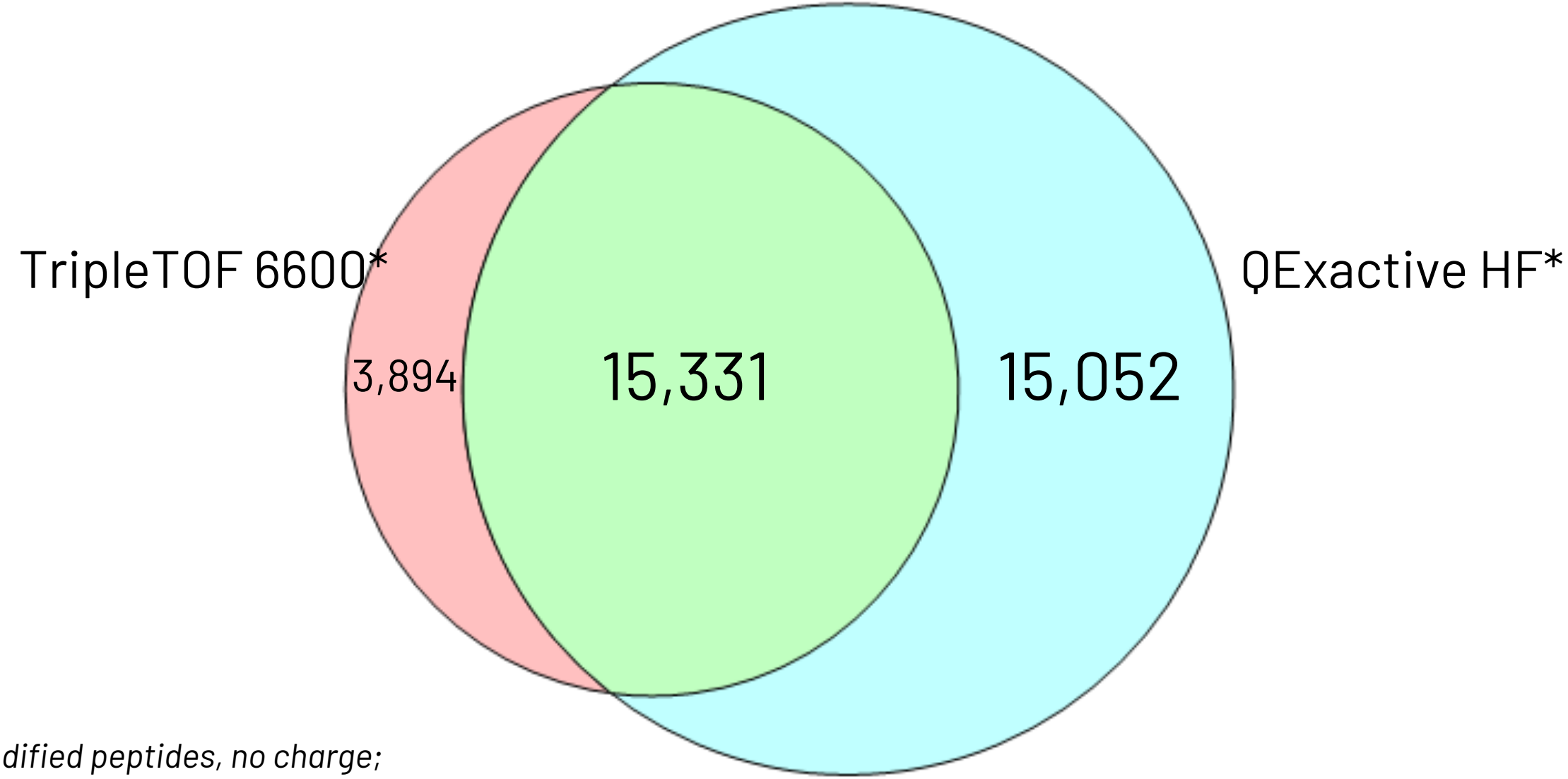
DISCO - DIAUmpire - DISCO w/ DIAUmpire Targets



Mass-Spectrometry Setups

	Triple TOF 6600	Q Exactive HF
Sample Type	HeLa Halo	HeLa Halo
Sample Load (µg)	3.5	0.88
Column Details	500µm X 150mm X 2.4µm	75µm X 500mm X 1.9µm
Flow (nl/min)	5000	300
Gradient length (min)	98	98
Median Peak Width (sec)	22.2	15.6
Acquisition Mode	SWATH	DIA
# Windows	100 variable	100 fixed (6m/z)
MS1 m/z range	400-1250	400-1000
MS1 scans- Accumulation time/Max fill time	250ms	200ms
MS2 scans- Accumulation time/Max fill time	25ms	10ms
Cycle Time (sec)	2.92	4.45

DISCO Peptide Identification: TTOF 6600 vs QE HF



**modified peptides, no charge;*
combining 2 search engines (Comet and X!Tandem)

DISCO Quantitation Options

- **MS1**
  - For each spectrum coming from an MS1 feature, DISCO stores the integrated precursor peak intensity stored as precursor_intensity in mzML
- **MS2 using QuanticParser**
  - For each PSM, Quantic computes the **An**no**t**ated **I**on **C**urrent (antic_score) by summing the intensities of annotated fragments in the spectrum
- **Hybrid – Combined MS1 and MS2 Quantitation**
  - Average of MS1 and MS2 quantitation results

DISCO Peptide Quantitation from PSMs

MS1

- **For each run *r*, for each peptide ion *s* compute**
  - **maximum MS1 precursor_intensity** over all PSMs in run *r* matching peptide ion *s*

$$\mathrm{I}_{\mathrm{s}}^{\mathrm{r}} = \max(\mathrm{I}_{\mathrm{x}} : \forall(\text{PSMs x in r where x matches s}))$$

  - **probability weighted MS1 intensity** over all PSMs in *r* matching *s*,

$$\mathrm{W}_{\mathrm{s}}^{\mathrm{r}} = \frac{\sum_{\forall(\text{PSMs x in r matching s})} \mathrm{p_p}(\mathrm{x})\mathrm{p_i}(\mathrm{x})\mathrm{I_x}}{\sum_{\forall(\text{PSMs x in r matching s})} \mathrm{p_p}(\mathrm{x})\mathrm{p_i}(\mathrm{x})}$$

where:

I_x is the precursor_intensity of PSM x

$p_p(x)$ is PeptideProphet probability and $p_i(x)$ is iProphet probability, of PSM x

DISCO Peptide Quantitation from PSMs

MS2

- **For each run *r*, for each peptide ion *s* compute**
  - **max MS1 antic_score** over all PSMs in run *r* matching peptide ion *s*

$$A_s^r = \max(A_x : \forall(\text{PSMs x in r where x matches s}))$$

  - **probability weighted antic_score** over all PSMs in *r* matching *s*,

$$a_s^r = \frac{\sum_{\forall(\text{PSMs x in r matching s})} p_p(x) p_i(x) A_x}{\sum_{\forall(\text{PSMs x in r matching s})} p_p(x) p_i(x)}$$

where:

A_x is the antic_score of PSM *x*

$p_p(x)$ is PeptideProphet probability and $p_i(x)$ is iProphet probability, of PSM *x*

DISCO Peptide Quantitation from PSMs

Hybrid

- **For each run *r*, for each peptide ion *s* compute**
  - **Compute Hybrid intensity**

$$H_s^r = \frac{1}{2}(I_s^r + a_s^r)$$

where:

I_s^r is the maximum precursor_intensity of peptide ion *s* in run *r*

a_s^r is the weighted antic_score of peptide ion *s* in run *r*

DISCO Peptide Quantitation Sample Means

- **For each peptide ion *s* over *n* replicates $r_1, \ldots, r_n$ of classification *c*, compute sample means:**

- **MS1**

$$\overline{W_s^c} = \frac{1}{n}\sum_{i=1}^{n} W_s^{r_i}$$

- **MS2**

$$\overline{a_s^c} = \frac{1}{n}\sum_{i=1}^{n} a_s^{r_i}$$

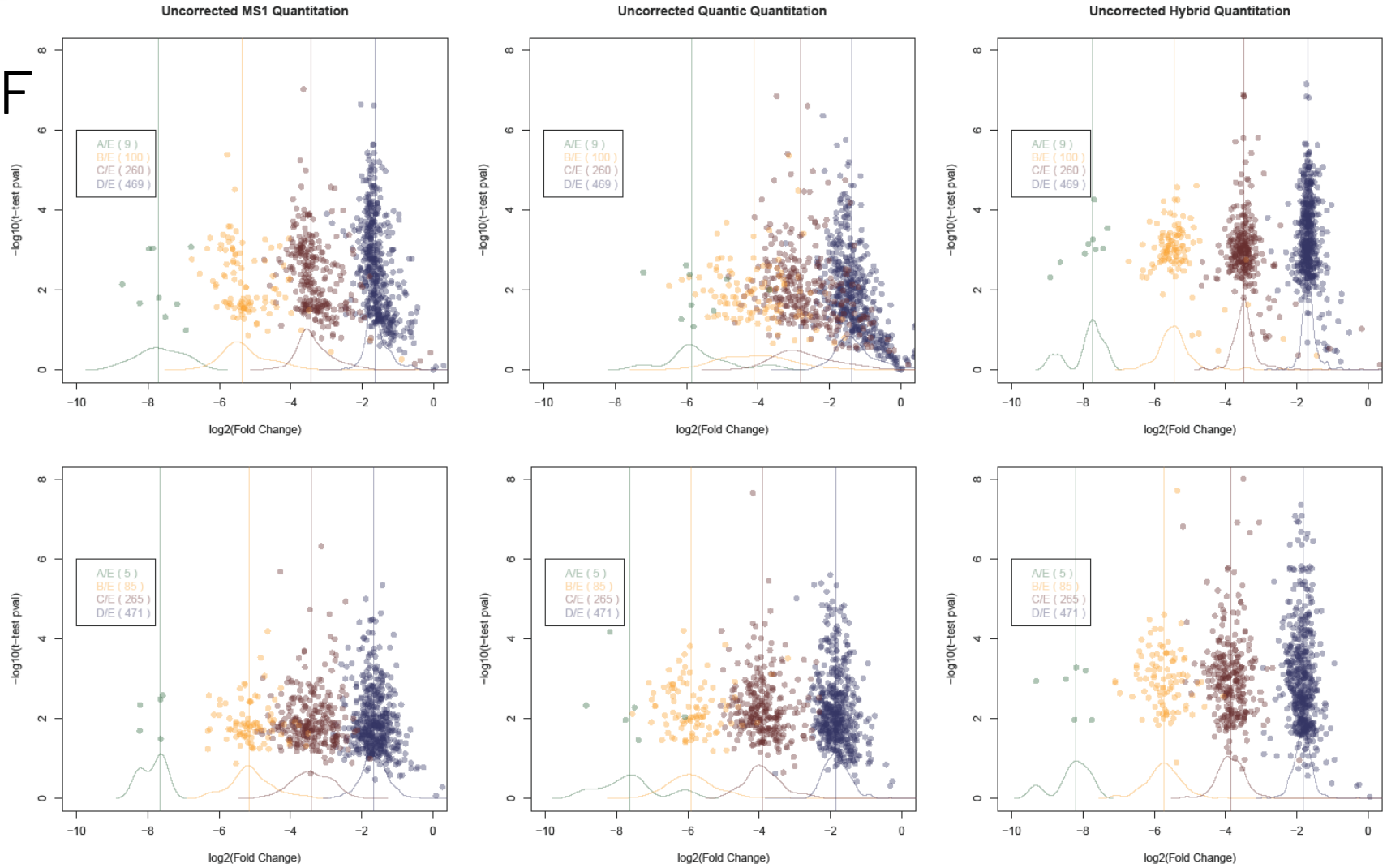
- **Hybrid**

$$\overline{H_s^c} = \frac{1}{n}\sum_{i=1}^{n} H_s^{r_i}$$

DISCO Halo Quantitation (not corrected for background)

QExactive HF

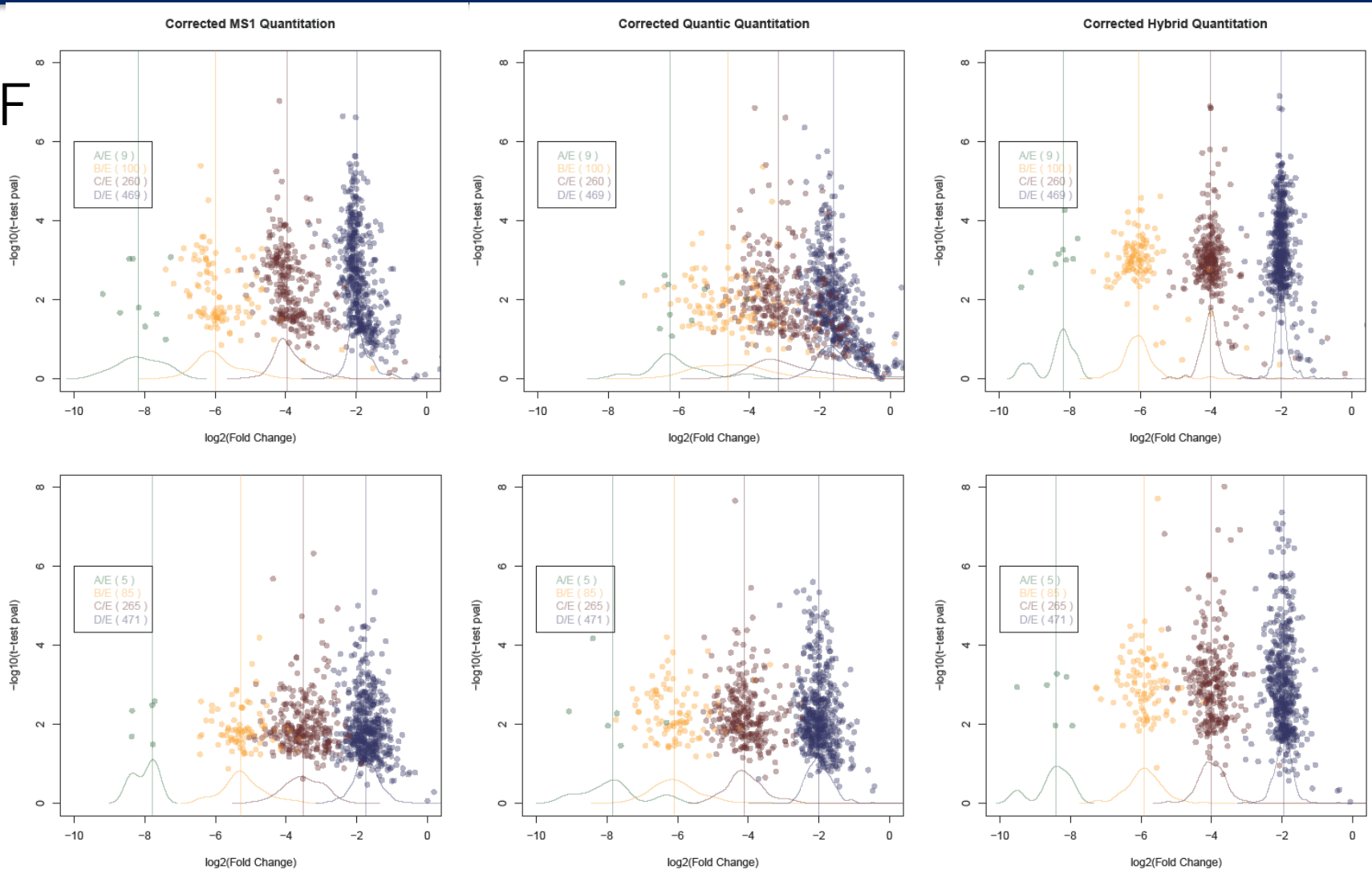
TTOF 6600



DISCO Halo Quantitation (corrected for background)

QExactive HF

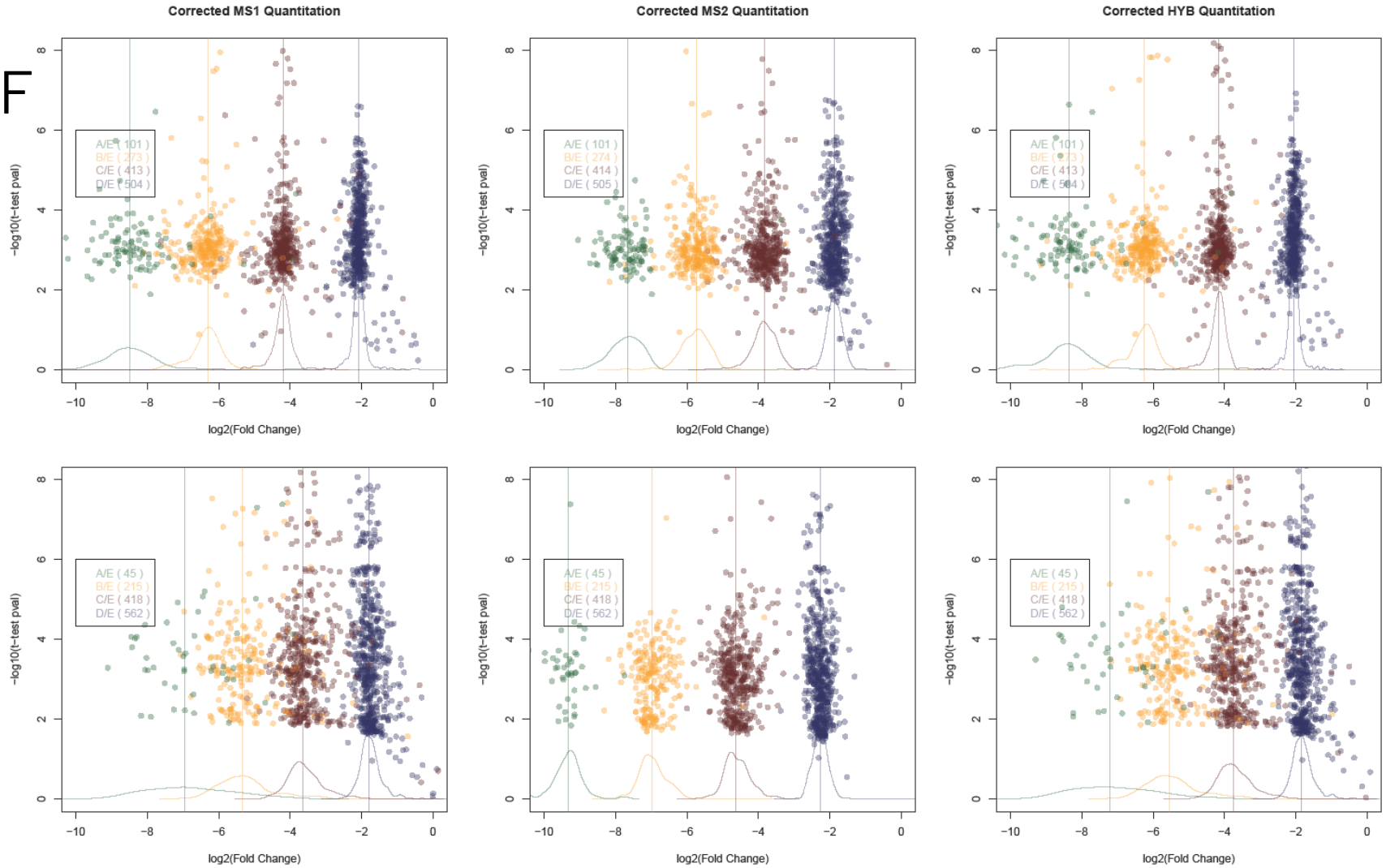
TTOF 6600



Spectronaut Halo Quantitation (corrected for background)

QExactive HF

TTOF 6600



Conclusions

- **Choice of data converter and centroiding options are important to ensure identifications are not lost due to the file converter**
- **DISCO identification rate improves upon DIAUmpire**
  - Combining approaches maximizes data-driven identifications
- **DISCO quantitation computes accurate ratios, reflecting the ground truth of the dataset**
  - Hybrid quantitation reduces variance in ratios
  - Correcting ratios using the median background ratio is important for accurate ratio estimation

QuanTic

Applied to CID Cleavable Isobaric Tag Quantitation

Journal of proteome research

pubs.acs.org/jpr

Article

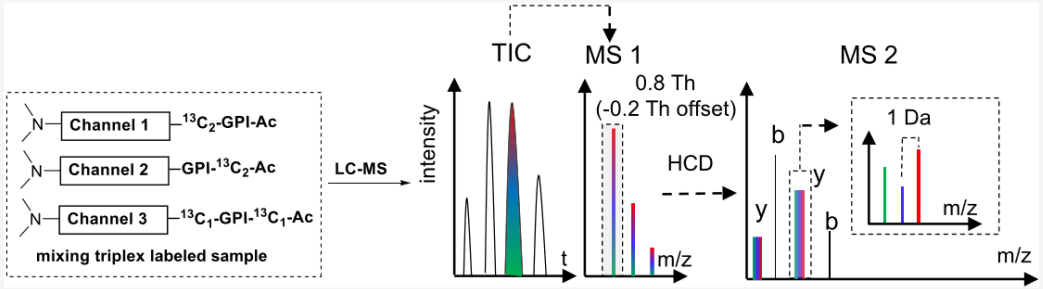
A Collision-Induced Dissociation Cleavable Isobaric Tag for Peptide Fragment Ion-Based Quantification in Proteomics

Xiaobo Tian, Marcel P. de Vries, Hjalmar P. Permentier, and Rainer Bischoff*

Cite This: https://dx.doi.org/10.1021/acs.jproteome.0c00371

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Quantifying peptides based on unique peptide fragment ions avoids the issue of ratio distortion that is commonly observed for reporter ion-based quantification approaches. Herein, we present a collision-induced dissociation-cleavable, isobaric acetyl-isoleucine-proline-glycine (Ac-IPG) tag, which conserves the merits of quantifying peptides based on unique fragments while reducing the complexity of the b-ion series compared to conventional fragment ion-based quantification methods thus facilitating data processing. Multiplex labeling is based on selective N-terminal dimethylation followed by derivatization of the ε-amino group of the C-terminal Lys residue of LysC peptides with isobaric Ac-IPG tags having complementary isotope distributions on Pro-Gly and Ac-Ile. Upon fragmentation between Ile and Pro, the resulting y ions, with the neutral loss of Ac-Ile, can be distinguished between the different labeling channels based on different numbers of isotope labels on the Pro-Gly part and thus contain the information for relative quantification, while b ions of different labeling channels have the same m/z values. The proteome quantification capability of this method was demonstrated by triplex labeling of a yeast proteome spiked with bovine serum albumin (BSA) over a 10-fold dynamic range. With the yeast proteins as the background, BSA was detected at ratios of 1.14:5.06:9.78 when spiked at 1:5:10 ratios. The raw mass data is available on the ProteomeXchange with the identifier PXD 018790.

KEYWORDS: *isobaric labeling, tandem mass spectrometry, fragment ion, quantitative proteomics, stable isotope*

Concept of the Ac-AG-tag

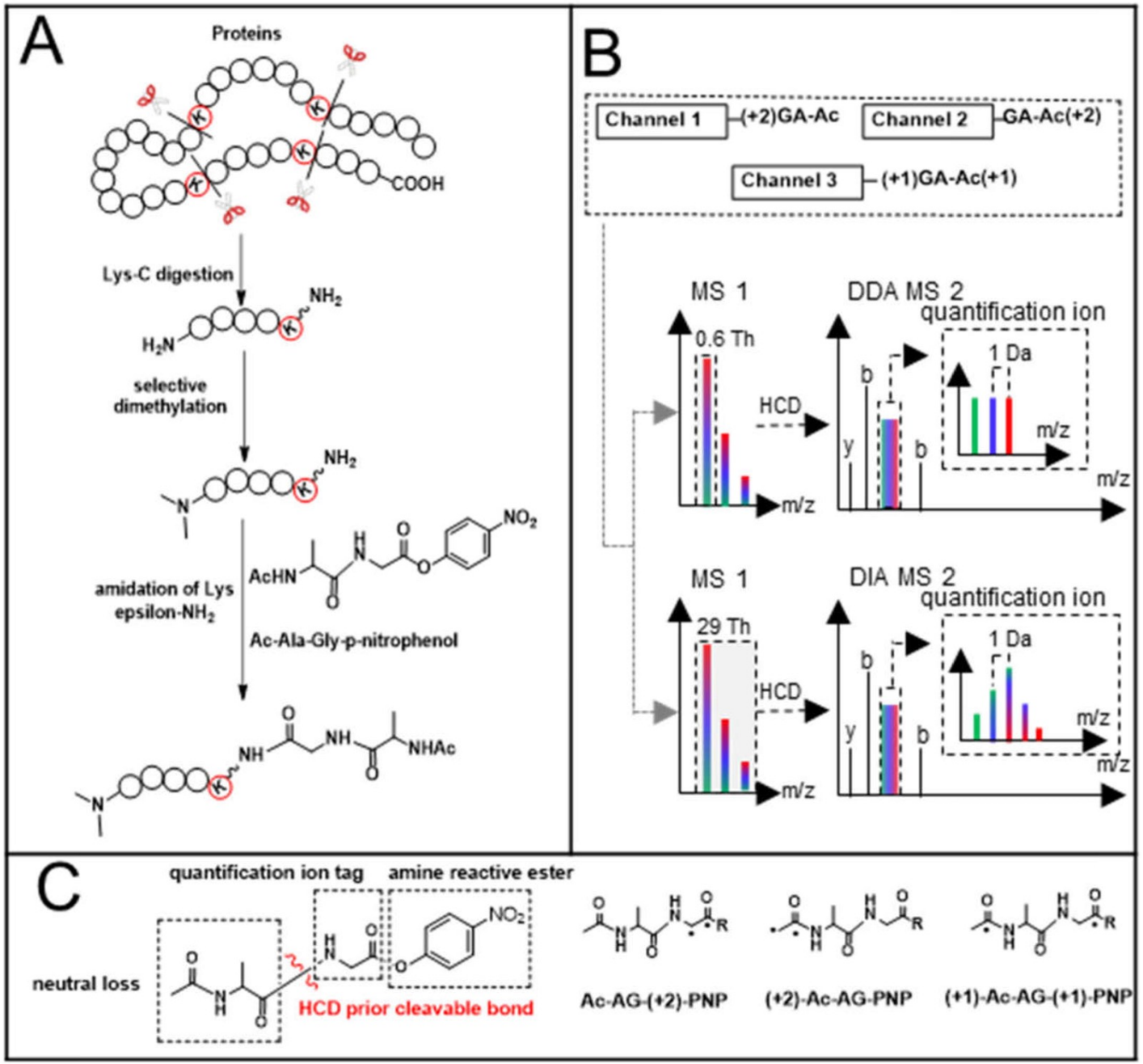


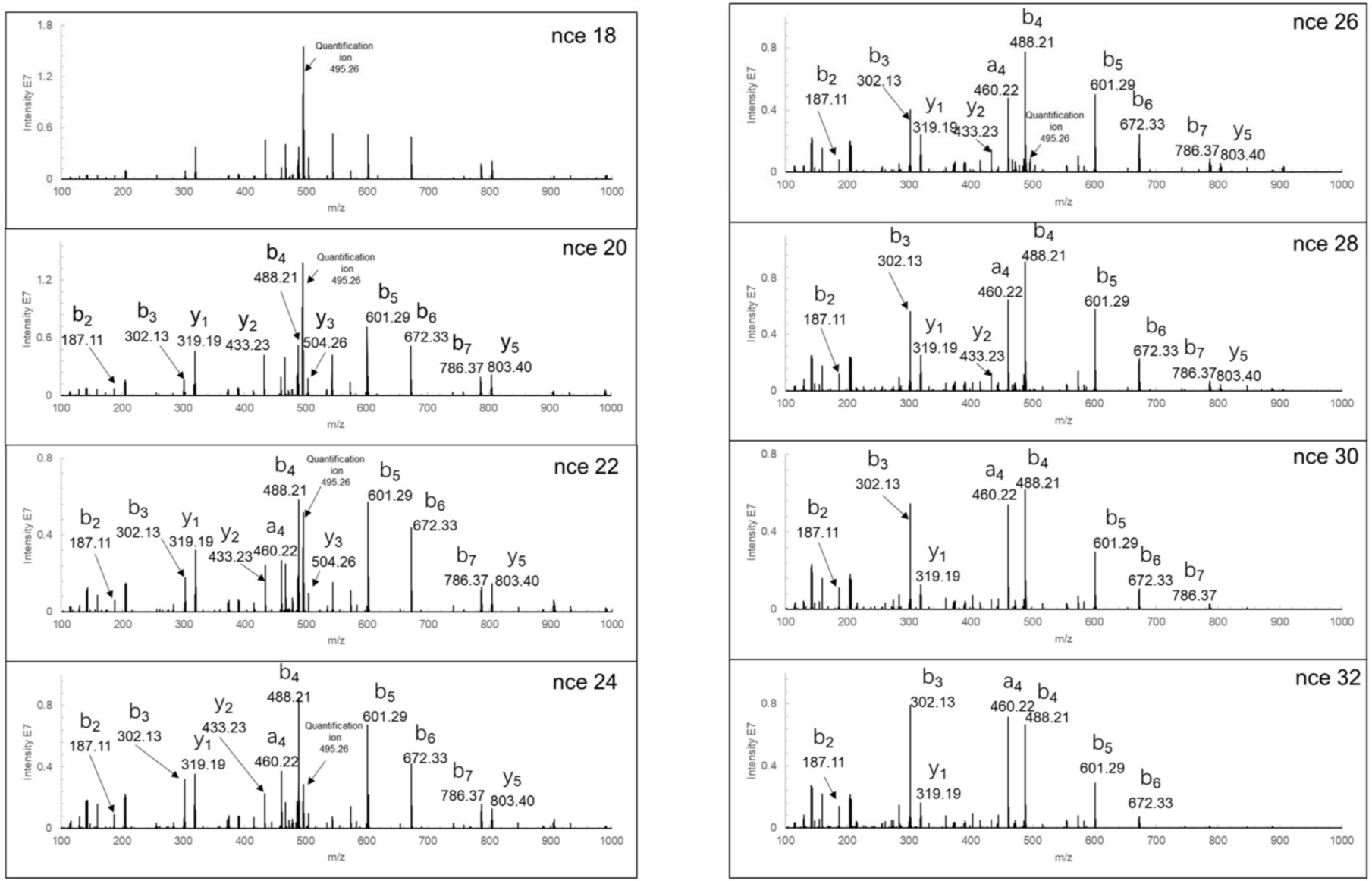
Fig. 1. Schematic view of the Ac-AG approach.

A) Isobaric labeling steps;

B) LC-MS/MS concept map for a mixture of triplex labeled samples;

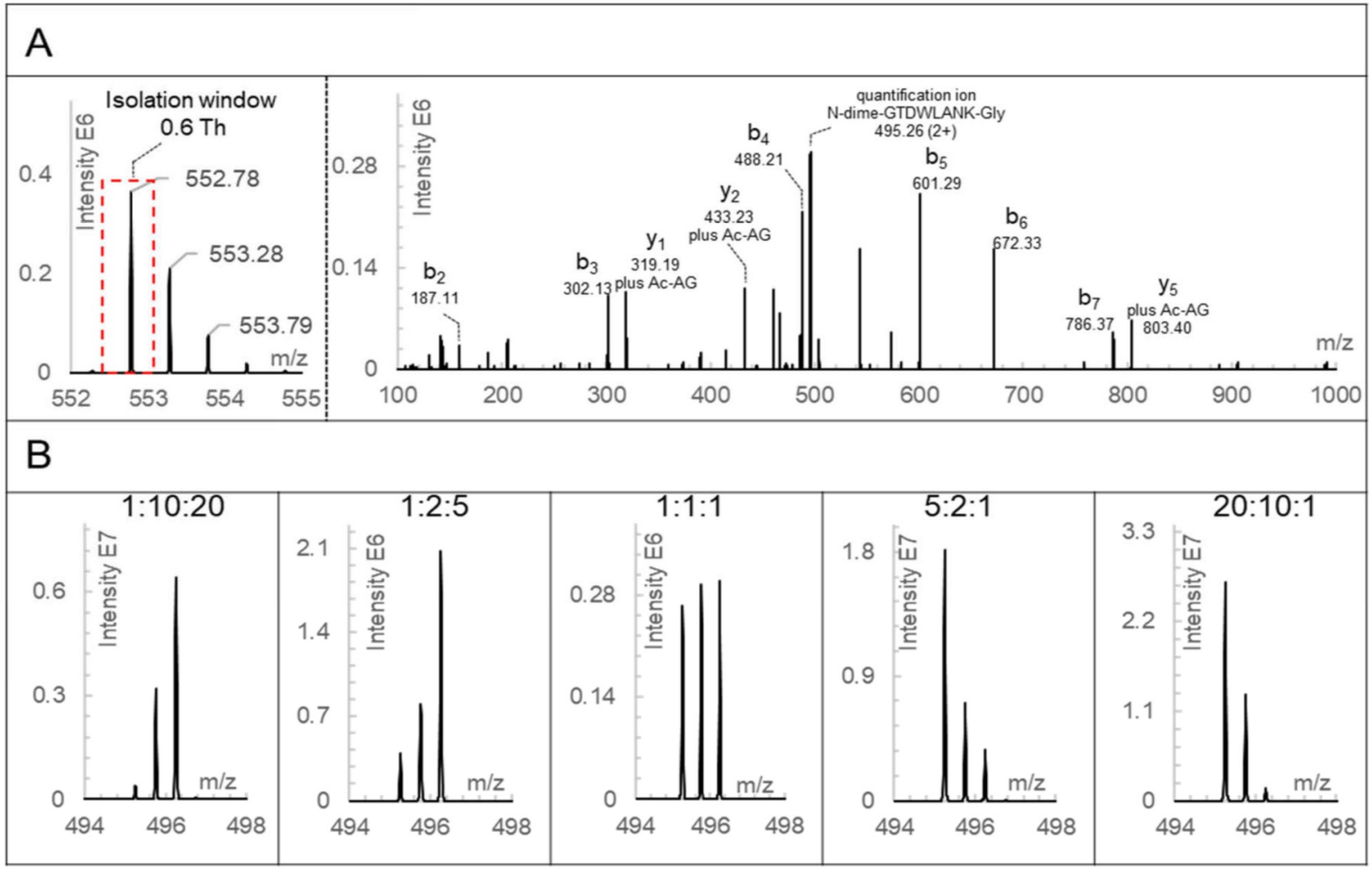
C) Functional design of the Ac-AG-PNP tag (^{13}C isotope locations are marked with"*").

Studying peptide fragmentation with various NCEs



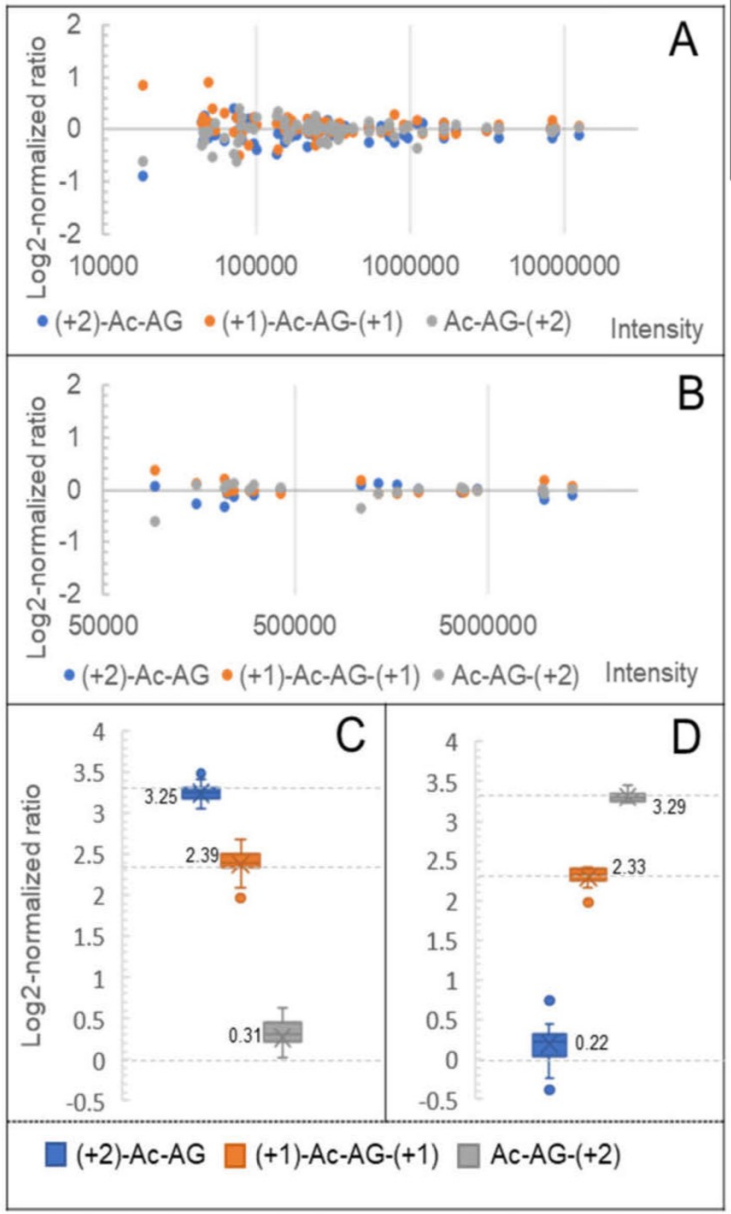
NCE **18** for producing peptide-coupled reporter-ion
NCE **28** for formation of fragments of the peptide backbone

Quantification at the peptide level



A) precursor isolation with a window of 0.6 Th and MS2 spectrum of triplex-labeled N-dime-GTDWLANK-GA-Ac with a combined NCE of 18 and 28. **B)** Peptide-coupled reporter-ions in the DDA MS2 spectra at ratios of 1:10:20, 1:2:5, 1:1:1, 5:2:1 and 20:10:1.

Quantification of triplex labeled LysC BSA



Calculating the ratio for PSM 6134 (AEFVEVTK) from triplex labeled BSA mixed at a ratio of 5:2:1

scan_number	intensity (+0)	intensity (+1)	intensity(+2)	total_intensity (3 channels)	sequence
6134	72202240	28389638	13828063	114419941	AEFVEVTK

Normalizer (weighted average) = (72202240 + 28389638 + 13828063)/(1+2+5) = 14302492.63
Ratio of 'spectrum of 6134' = 72202240 / 14302492.63 : 28389638 / 14302492.63 : 13828063 / 14302492.63
= 5.05 : 1.98 : 0.97

Calculating peptide ratios based on top 3 most intense PSMs.
Calculating protein ratios based on top 3 most intense peptides.

A) Log2-normalized ratio distribution at the PSM level at a mixing ratio of 1:1:1; **B)** Log2-normalized ratio distribution at the peptide level at a mixing ratio of 1:1:1; **C)** Log2-normalized ratio at the peptide level at a mixing ratio of 10:5:1; **D)** Log2-normalized ratio at the peptide level at a mixing ratio of 1:5:10. Expected values for log2-normalized mixing ratios are shown as dotted lines.

QuanTic New Option for Quantitation

K:311.17:-155.1:-156.1:-157.1

<amino acids, n, or c>:<mass_shift>:<neut_loss1>:...:<neut_lossN> Specify modifications with neutral losses for quantitation (default: off)

***Quantify the intensity of neutral loss peaks associated with a modification for all fragments retaining the modification, when all are available for a given fragment ion, skip y1**

PepXML Viewer: **interact-yeast.quant.pep.xml**

Summary ▾ | Display Options ▾ | Pick Columns ▾ | Filtering Options ▾ | Other Actions ▾ | Messages ▾

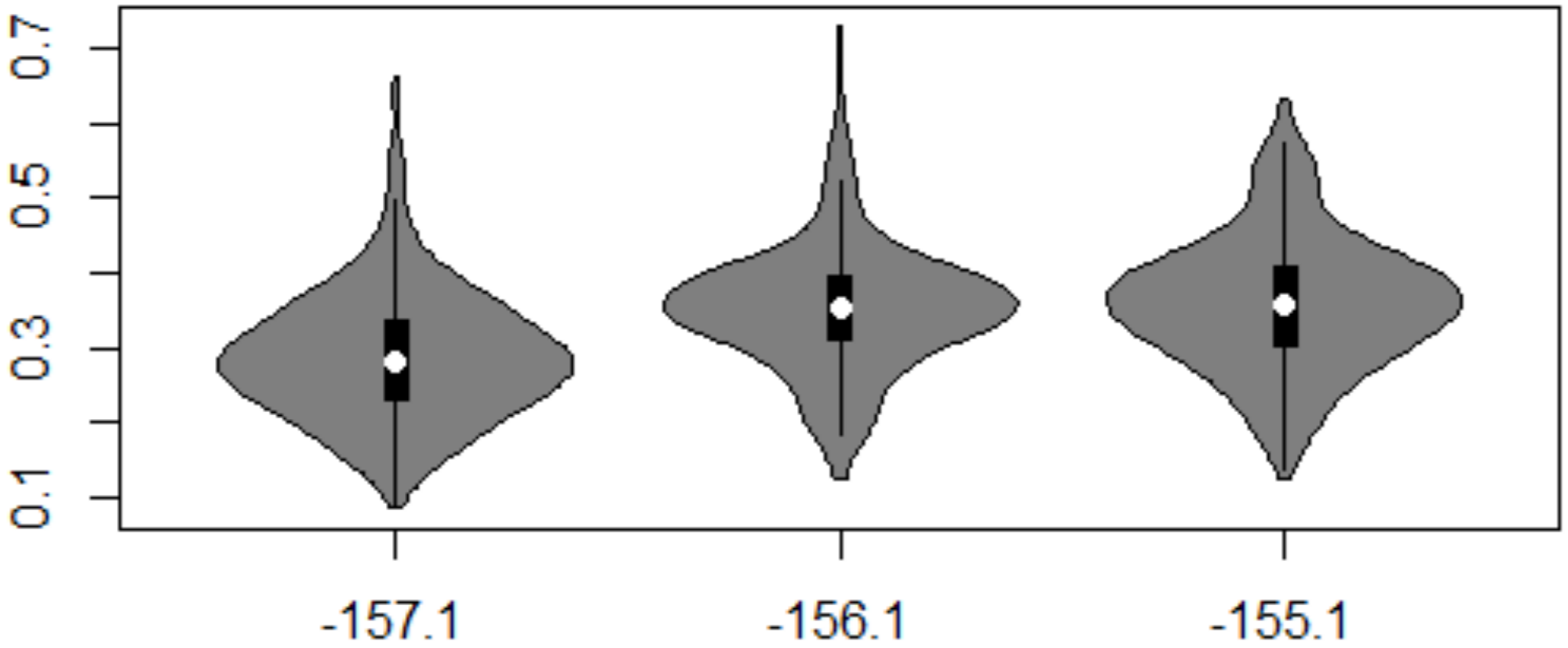
pepXML

Page 1 of 17

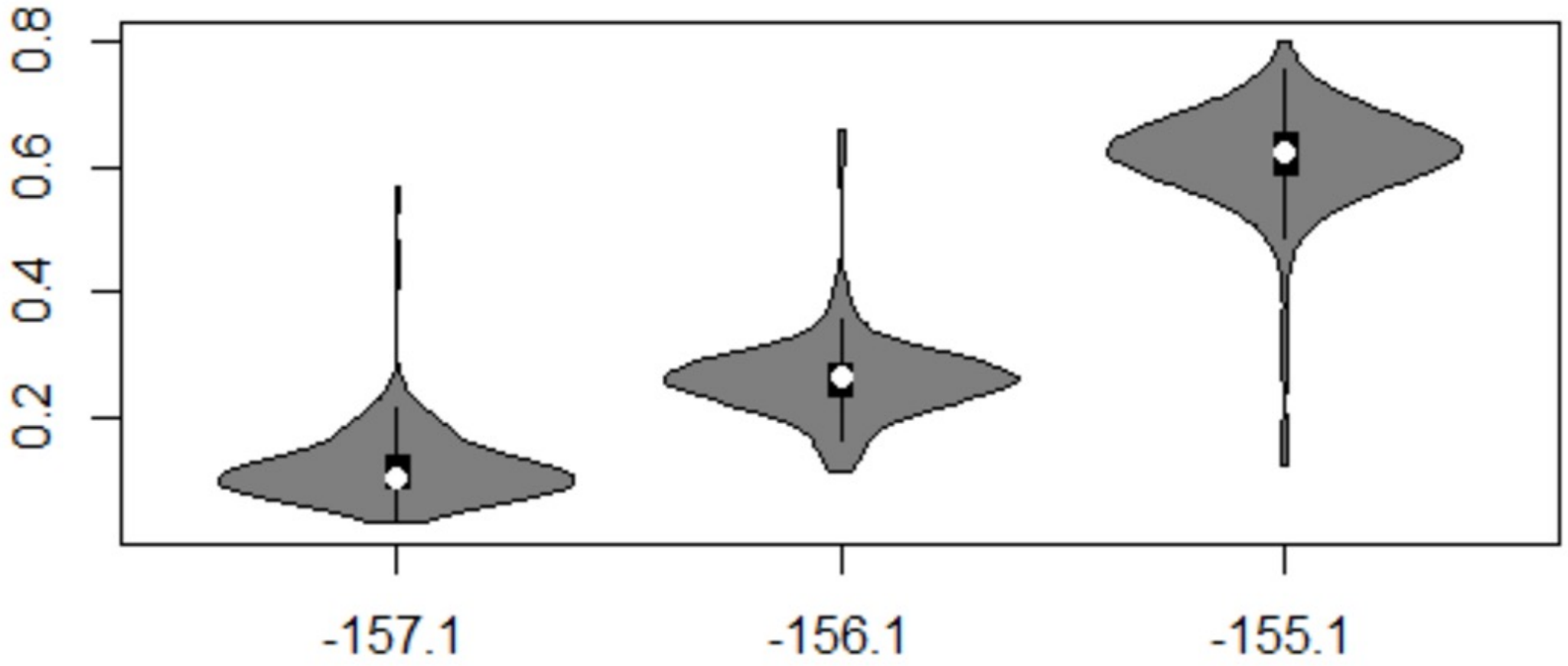
1 FIRST **1** 2 3 4 5 6 11 NEXT LAST

PROB	SPECTRUM	EXPECT	IONS	PEPTIDE	PROTEIN	NPROTS	CALC_MASS	QUANTIC	QUANTIC_TIC	QUANTIC_QUANT	QUANTIC_STRING
1.0000	Ac-IPG-yeast-1-1-1.06237.06237.3	0.00011	10/52	K.n29.04AVYAGENFHHGDK439.26L.-	sp|P00925|ENO2_YEAST	1	1895.9385	3.89711e+06	4.34866e+07	1347818.1,1269763.4,735574.9	M(-155.1),0,581.291,0
1.0000	Ac-IPG-yeast-1-1-1.09376.09376.3	0.000245	11/64	K.n29.04TFAEAMRIGSEVYHNLK439.26.S	sp|P00925|ENO2_YEAST	1	2304.1791	1.30834e+06	1.07953e+07	524884.3,472688.3,315971.1	
1.0000	Ac-IPG-yeast-1-1-1.08330.08330.3	1.2e-07	17/72	R.n29.04SVYDSRGNPTVEVELTTEK439.26.G	sp|P00924|ENO1_YEAST +1	2	2462.2395	1.95827e+06	1.42429e+07	417081.5,289362.8,194558.9	
1.0000	Ac-IPG-yeast-1-1-1.07870.07870.3	0.000231	13/52	K.n29.04LSRAIQTANIALEK439.26.A	sp|P00950|PMG1_YEAST	1	1866.0793	1.70548e+06	1.16365e+07	300725.9,277894.0,107346.1	
1.0000	Ac-IPG-yeast-1-1-1.07563.07563.3	0.000203	12/48	K.n29.04DVVTREYTINLHK439.26.R	sp|P0C2H8|RL31A_YEAST +1	2	1926.0430	702968	1.24334e+07	251100.6,227770.3,126984.0	
0.9988	Ac-IPG-yeast-1-1-1.06329.06329.3	0.696	11/52	K.n29.04LVLVRHGQSEWNEK439.26.N	sp|P00950|PMG1_YEAST	1	2033.0913	681717	4.77765e+06	247265.2,229580.6,123972.0	
0.9994	Ac-IPG-yeast-1-1-1.07926.07926.3	0.00217	10/72	R.n29.04SIAPAYGIPVVLHSDHC160.03AK439.26.K	sp|P14540|ALF_YEAST	1	2373.2370	1.62851e+06	8.71115e+06	243413.1,221734.3,165115.4	M(-155.1),0,740.385,-0.001,6254.6;M(-156.1),0,740.051,-0.0
0.9988	Ac-IPG-yeast-1-1-1.08861.08861.3	0.00522	11/68	K.n29.04IGGIGTVPVGRVETGVIK439.26.P	sp|P02994|EF1A_YEAST	1	2090.2318	633286	6.6591e+06	225457.1,206643.7,131845.3	
0.9999	Ac-IPG-yeast-1-1-1.06736.06736.3	0.00335	9/56	K.n29.04IYPGRGTLFVRGDSK439.26.L	sp|P04449|RL24A_YEAST +1	2	2004.1011	575172	2.72473e+06	222247.7,199969.1,143789.9	
0.9023	Ac-IPG-yeast-1-1-1.11233.11233.2	0.0726	4/34	R.n29.04SFDVPPPPIDASSPFSQK439.26.G	sp|P00950|PMG1_YEAST	1	2254.1376	639804	1.99186e+06	189855.8,148777.9,103828.8	
1.0000	Ac-IPG-yeast-1-1-1.10847.10847.3	0.000426	15/64	K.n29.04ALENPTRPFLAILGGAK439.26.V	sp|P00560|PGK_YEAST	1	2106.2056	482013	3.15874e+06	172527.2,154608.7,110860.0	
1.0000	Ac-IPG-yeast-1-1-1.09442.09442.3	5.7e-07	15/64	K.n29.04IVSNASC160.03TTNC160.03LAPLAK439.26.V	sp|P00358|G3P2_YEAST +3	4	2158.0981	2.00572e+06	1.93312e+07	169416.0,167327.3,118868.3	
1.0000	Ac-IPG-yeast-1-1-1.09266.09266.3	2.14e-06	11/76	K.n29.04TSAVAALTEVRAEDEAALAK439.26.L	sp|P17076|RL8A_YEAST +1	2	2354.2548	815505	7.16469e+06	158161.9,161845.1,89759.4	
0.9975	Ac-IPG-yeast-1-1-1.08020.08020.3	0.385	8/60	K.n29.04YAQDGAGIERELARIK439.26.K	sp|P14126|RL3_YEAST	1	2128.1495	824834	2.78199e+06	132469.4,109744.6,89193.2	
1.0000	Ac-IPG-yeast-1-1-1.07359.07359.3	0.00275	11/60	K.n29.04TVMIAAHGNSLRGLVK439.26.H	sp|P00950|PMG1_YEAST	1	2005.1361	349374	2.41765e+06	108782.1,94074.6,71460.5	
0.9223	Ac-IPG-yeast-1-1-1.06350.06350.3	1.6	3/40	K.n29.04NIVEFHSDHMK439.26.L	sp|P06169|PDC1_YEAST	1	1694.8305	160852	3.35394e+06	104704.5,86711.9,84888.3	
1.0000	Ac-IPG-yeast-1-1-1.05512.05512.3	0.000841	3/40	M.n29.04NFSHGSYEYHK439.26.S	sp|P00549|KPYK1_YEAST	1	1706.7908	208583	4.87777e+06	103630.2,105907.0,71181.5	
1.0000	Ac-IPG-yeast-1-1-1.05694.05694.3	0.0145	9/44	K.n29.04ALNEEAEARRLK439.26.N	sp|P0CX82|RL19A_YEAST +1	2	1737.9592	377602	882798	102408.1,91357.6,76992.9	
0.9999	Ac-IPG-yeast-1-1-1.09738.09738.3	0.0738	7/68	K.n29.04SFLESVIRDSVTYTEHAK439.26.R	sp|P02309|H4_YEAST	1	2420.2442	309493	1.58101e+06	97209.5,78974.3,64076.4	
0.9953	Ac-IPG-yeast-1-1-1.07541.07541.2	0.0568	6/14	R.n29.04EVLGEQGK439.26.D	sp|P00549|KPYK1_YEAST	1	1197.6460	364756	2.79228e+06	88068.0,78618.8,77815.7	
0.9792	Ac-IPG-yeast-1-1-1.08567.08567.3	0.183	5/56	R.n29.04LVIVTDPRSDAQAIK439.26.E	sp|P32905|RSSA1_YEAST	1	1964.1161	280769	1.64864e+06	86562.6,102413.0,56210.7	
1.0000	Ac-IPG-yeast-1-1-1.07839.07839.3	3.3e-05	13/56	F.n29.04AEALRIGSEVYHNLK439.26.S	sp|P00924|ENO1_YEAST	1	2038.1066	280726	2.43851e+06	83421.3,39456.8,23738.8	
0.9892	Ac-IPG-yeast-1-1-1.05824.05824.3	0.958	6/48	K.n29.04LVERAVSEDPRIK439.26.M	sp|P32527|ZUO1_YEAST	1	1850.0480	618037	1.97821e+06	79985.4,76044.3,66386.6	
1.0000	Ac-IPG-yeast-1-1-1.09271.09271.3	3.18e-05	13/68	K.n29.04YVDEQVELADAAPPPEAK439.26.L	sp|P16387|ODPA_YEAST	1	2280.1380	421155	1.87633e+06	77323.9,86621.9,47663.7	
1.0000	Ac-IPG-yeast-1-1-1.04614.04614.3	0.0229	9/48	K.n29.04TAEQVAAERAARK439.26.A	sp|P05737|RL7A_YEAST	1	1738.9545	447613	1.1847e+06	71672.2,71510.2,64837.1	

QuanTic Normalized Results 1:1:1



QuanTic Normalized Results 1:2:5



analytical chemistry

pubs.acs.org/ac

Article

A Versatile Isobaric Tag Enables Proteome Quantification in Data-Dependent and Data-Independent Acquisition Modes

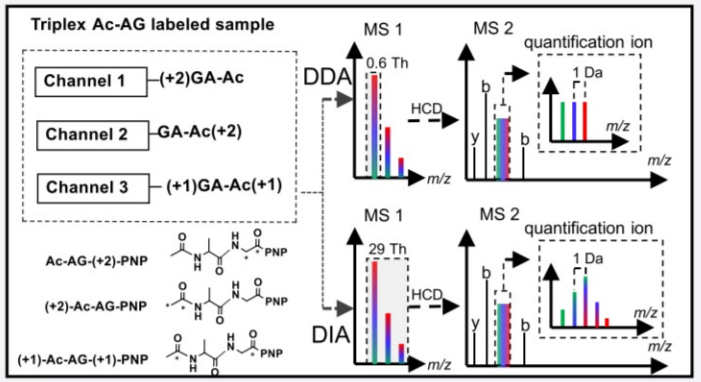
Xiaobo Tian, Marcel P. de Vries, Hjalmar P. Permentier, and Rainer Bischoff*

Cite This: *Anal. Chem.* 2020, 92, 16149–16157

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Quantifying proteins based on peptide-coupled reporter ions is a multiplexed quantitative strategy in proteomics that alleviates the problem of ratio distortion caused by peptide cofragmentation, as commonly observed in other reporter-ion-based approaches, such as TMT and iTRAQ. Data-independent acquisition (DIA) is an attractive alternative to data-dependent acquisition (DDA) due to its better reproducibility. While multiplexed labeling is widely used in DDA, it is rarely used in DIA, presumably because current approaches lead to more complex MS2 spectra, severe ratio distortion, or to a reduction in quantification accuracy and precision. Herein, we present a versatile acetyl-alanine-glycine (Ac-AG) tag that conceals quantitative information in isobarically labeled peptides and reveals it upon tandem MS in the form of peptide-coupled reporter ions. Since the peptide-coupled reporter ion is precursor-specific while fragment ions of the peptide backbone originating from different labeling channels are identical, the Ac-AG tag is compatible with both DDA and DIA. By isolating the monoisotopic peak of the precursor ion in DDA, intensities of the peptide-coupled reporter ions represent the relative ratios between constituent samples, whereas in DIA, the ratio can be inferred after deconvoluting the peptide-coupled reporter ion isotopes. The proteome quantification capability of the Ac-AG tag was demonstrated by triplex labeling of a yeast proteome spiked with bovine serum albumin (BSA) over a 10-fold dynamic range. Within this complex proteomics background, BSA spiked at 1:5:10 ratios was detected at ratios of 1.00:4.87:10.13 in DDA and 1.16:5.20:9.64 in DIA.



QuanTic : Flexible Options

K:172.075845:-113.0477:-114.0511:-115.0545

COUPLED

- Option that restricts the quantitation to using only the unfragmented precursor-coupled neutral losses for Quantic

DIAMODE

- Option that allows deconvolves quantitative values for multiple isotope, problem common to DIA data

QuanTic QuanticProteinParser

- **Combine Quantic PSM quantitation to compute protein level quantities using PSM stats weighted by intensity**
- **Fully integrated with ProteinProphet**
  - Input and output are the same protXML file
- **Requires no user input other than the input file**
- **Visualized by ProtXMLViewer.cgi**

QuanTic : ProtXML Protein View

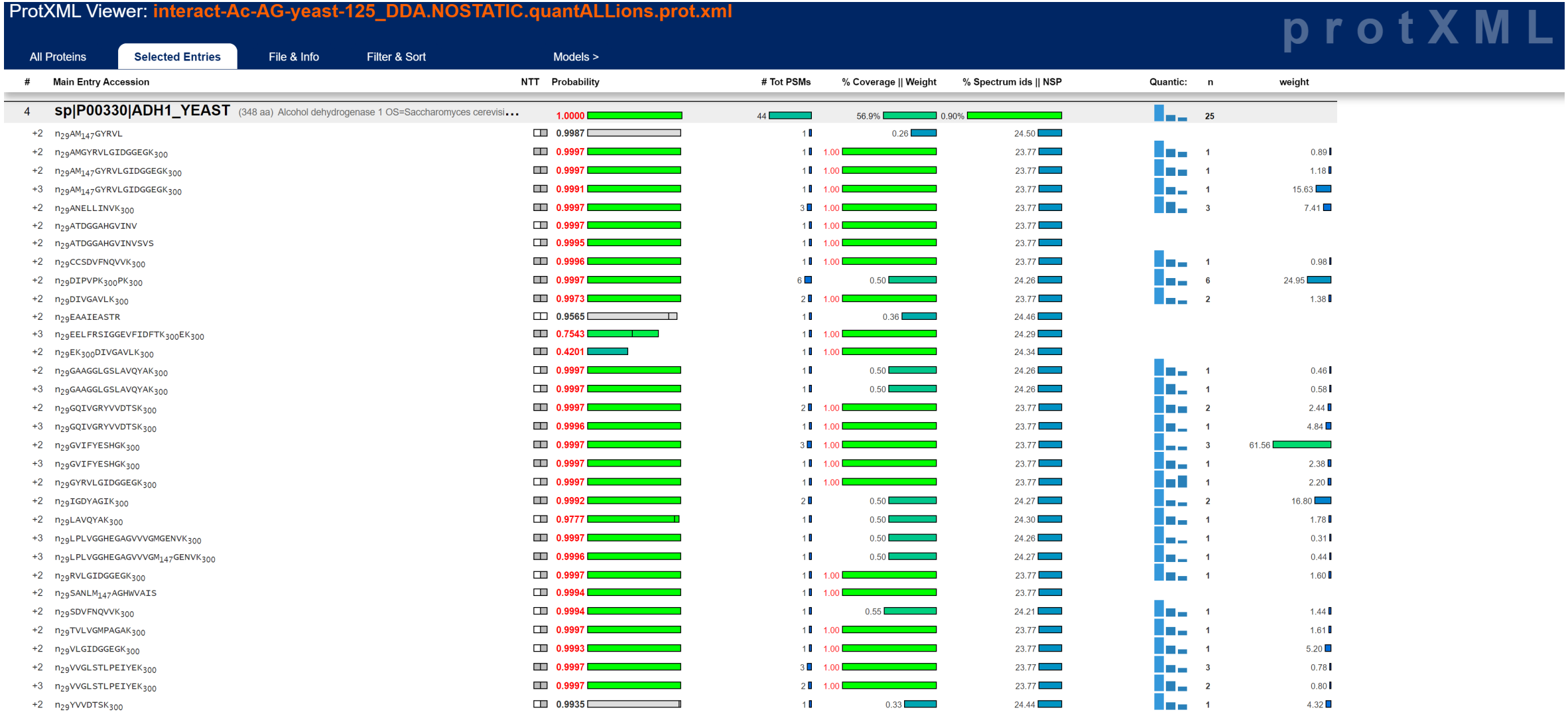
ProtXML Viewer: **interact-Ac-AG-yeast-125_DDA.NOSTATIC.quantALLions.prot.xml**

protXML

All Proteins | Selected Entries | File & Info | Filter & Sort | Models >

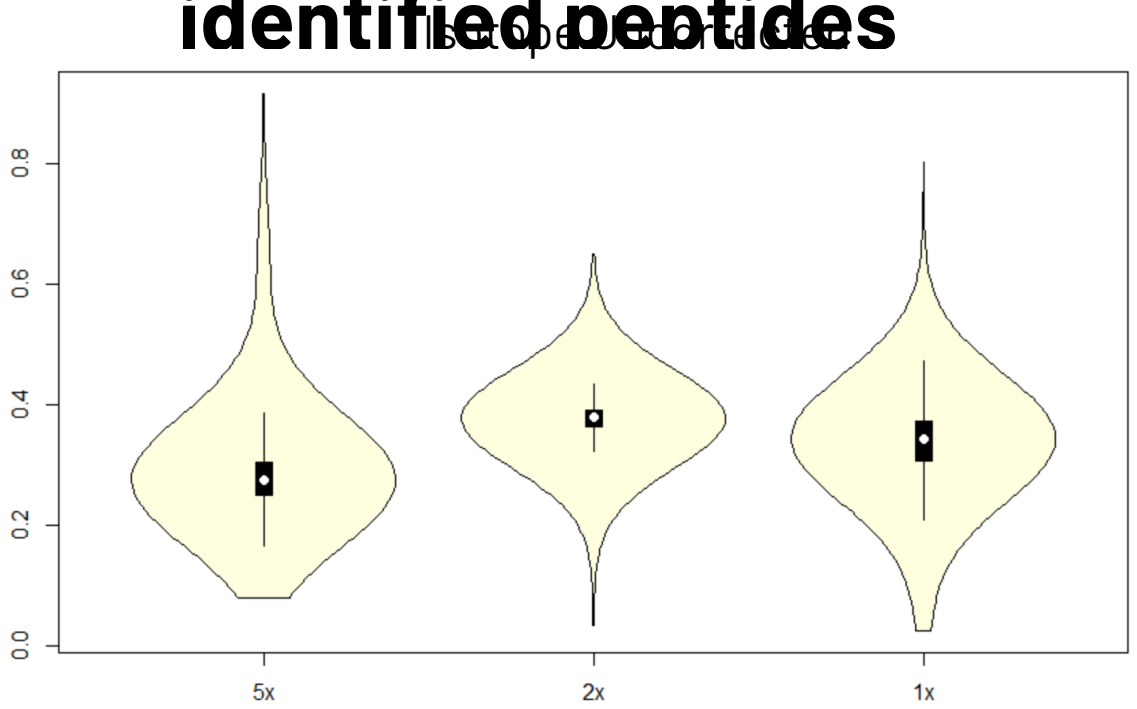
#	Main Entry Accession	NTT	Probability	# Tot PSMs	% Coverage || Weight	% Spectrum ids || NSP	Quantic:	n	weight
1	**gi|21961605|gb|AAH34697.1|** (584 aa) (CONTAMINANT) Keratin 10 [Homo sapiens]		1.0000	3	6.3%	0.07%		3	
2	**sp|O13516|RS9A_YEAST** (197 aa) 40S ribosomal protein S9-A OS=Saccharomyces cere...		1.0000	9	29.9%	0.11%		6	
3	**sp|O14467|MBF1_YEAST** (151 aa) Multiprotein-bridging factor 1 OS=Saccharomyces cer...		1.0000	7	44.4%	0.17%		6	
4	**sp|P00330|ADH1_YEAST** (348 aa) Alcohol dehydrogenase 1 OS=Saccharomyces cerevisi...		1.0000	44	56.9%	0.90%		25	
5	**sp|P00331|ADH2_YEAST** (348 aa) Alcohol dehydrogenase 2 OS=Saccharomyces cerevisi...		1.0000	2	25.3%	0.23%		10	
6	**sp|P00431|CCPR_YEAST** (361 aa) Cytochrome c peroxidase, mitochondrial OS=Sacchar...		1.0000	2	4.4%	0.05%		2	
7	**sp|P00445|SODC_YEAST** (154 aa) Superoxide dismutase [Cu-Zn] OS=Saccharomyces c...		1.0000	9	50.0%	0.21%		6	
8	**sp|P00498|HIS1_YEAST** (297 aa) ATP phosphoribosyltransferase OS=Saccharomyces cer...		1.0000	3	9.4%	0.07%		3	
9	**sp|P00560|PGK_YEAST** (416 aa) Phosphoglycerate kinase OS=Saccharomyces cerevisia...		1.0000	82	74.8%	1.94%		40	
10	**sp|P00815|HIS2_YEAST** (799 aa) Histidine biosynthesis trifunctional protein OS=Saccharo...		1.0000	9	12.4%	0.22%		8	
11	**sp|P00817|IPYR_YEAST** (287 aa) Inorganic pyrophosphatase OS=Saccharomyces cerevis...		1.0000	16	44.9%	0.39%		11	
12	**sp|P00830|ATPB_YEAST** (511 aa) ATP synthase subunit beta, mitochondrial OS=Sacchar...		1.0000	4	9.4%	0.10%		4	
13	**sp|P00931|TRP_YEAST** (707 aa) Tryptophan synthase OS=Saccharomyces cerevisiae (str...		1.0000	4	5.9%	0.09%		3	
14	**sp|P00942|TPIS_YEAST** (248 aa) Triosephosphate isomerase OS=Saccharomyces cerevis...		1.0000	21	51.2%	0.46%		12	
15	**sp|P00950|PMG1_YEAST** (247 aa) Phosphoglycerate mutase 1 OS=Saccharomyces cere...		1.0000	30	58.7%	0.70%		17	
16	**sp|P00958|SYMC_YEAST** (751 aa) Methionine--tRNA ligase, cytoplasmic OS=Saccharom...		1.0000	5	6.8%	0.12%		3	
17	**sp|P02293|H2B1_YEAST** (131 aa) **+1**	IR	1.0000	5	31.3%	0.12%		5	
18	**sp|P02309|H4_YEAST** (103 aa) Histone H4 OS=Saccharomyces cerevisiae (strain ATCC 20...		1.0000	3	36.9%	0.07%		2	
19	**sp|P02400|RLA4_YEAST** (110 aa) 60S acidic ribosomal protein P2-beta OS=Saccharomy...		1.0000	3	31.8%	0.07%		2	
20	**sp|P02406|RL28_YEAST** (149 aa) 60S ribosomal protein L28 OS=Saccharomyces cerevisi...		1.0000	11	43.0%	0.25%		7	
21	**sp|P02407|RS17A_YEAST** (136 aa) **+1**	IR	1.0000	13	58.1%	0.31%		8	
22	**sp|P02829|HSP82_YEAST** (709 aa) ATP-dependent molecular chaperone HSP82 OS=Sa...		1.0000	3	46.3%	0.55%		31	
23	**sp|P02994|EF1A_YEAST** (458 aa) Elongation factor 1-alpha OS=Saccharomyces cerevisi...		1.0000	33	58.5%	0.80%		23	
24	**sp|P04147|PABP_YEAST** (577 aa) Polyadenylate-binding protein, cytoplasmic and nuclea...		1.0000	16	30.0%	0.38%		14	
25	**sp|P04449|RL24A_YEAST** (155 aa) 60S ribosomal protein L24-A OS=Saccharomyces ce...		1.0000	10	33.5%	0.13%		9	
26	**sp|P04456|RL25_YEAST** (142 aa) 60S ribosomal protein L25 OS=Saccharomyces cerevisi...		1.0000	11	49.3%	0.26%		8	
27	**sp|P04801|SYTC_YEAST** (734 aa) Threonine--tRNA ligase, cytoplasmic OS=Saccharomy...		1.0000	10	10.9%	0.24%		9	
28	**sp|P04802|SYDC_YEAST** (557 aa) Aspartate--tRNA ligase, cytoplasmic OS=Saccharomyc...		1.0000	7	14.0%	0.17%		6	
29	**sp|P04806|HXKA_YEAST** (485 aa) Hexokinase-1 OS=Saccharomyces cerevisiae (strain A...		1.0000	24	37.3%	0.48%		16	
30	**sp|P04807|HXKB_YEAST** (486 aa) Hexokinase-2 OS=Saccharomyces cerevisiae (strain A...		1.0000	18	29.2%	0.34%		13	
31	**sp|P04840|VDAC1_YEAST** (283 aa) Mitochondrial outer membrane protein porin 1 OS=S...		1.0000	5	15.2%	0.12%		5	

QuanTic : ProXML Peptide View

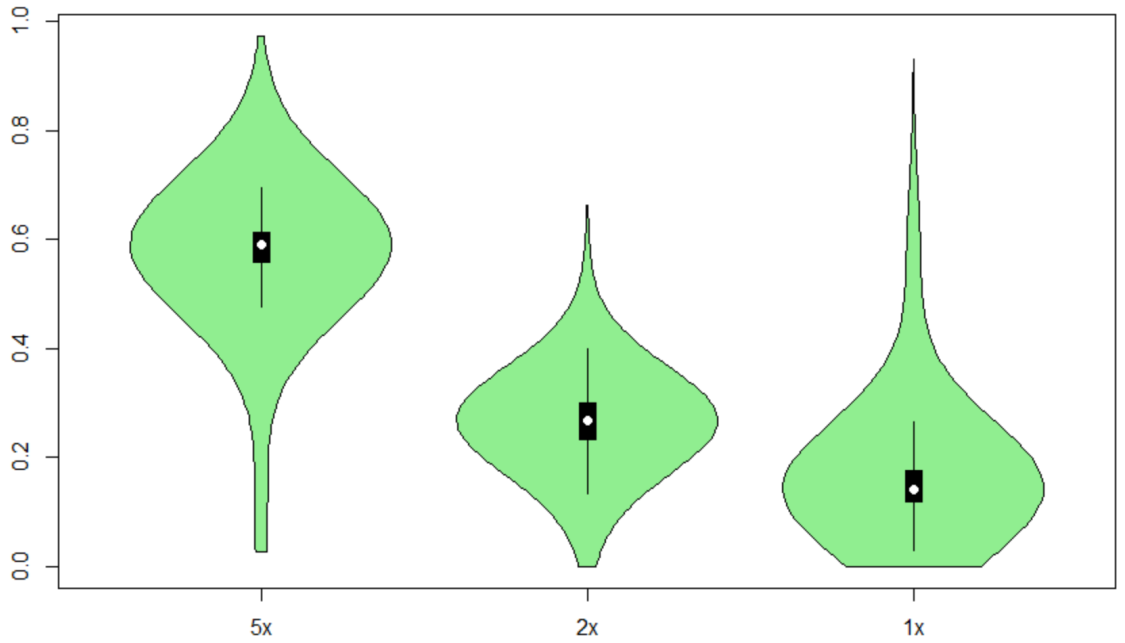


QuanTic : DIAMODE

- **Deconvolves overlapping Isotopes of DIA data to extract quantitative signals, uses Mike's Mercury8 code to predict theoretical isotopic distributions of identified peptides**

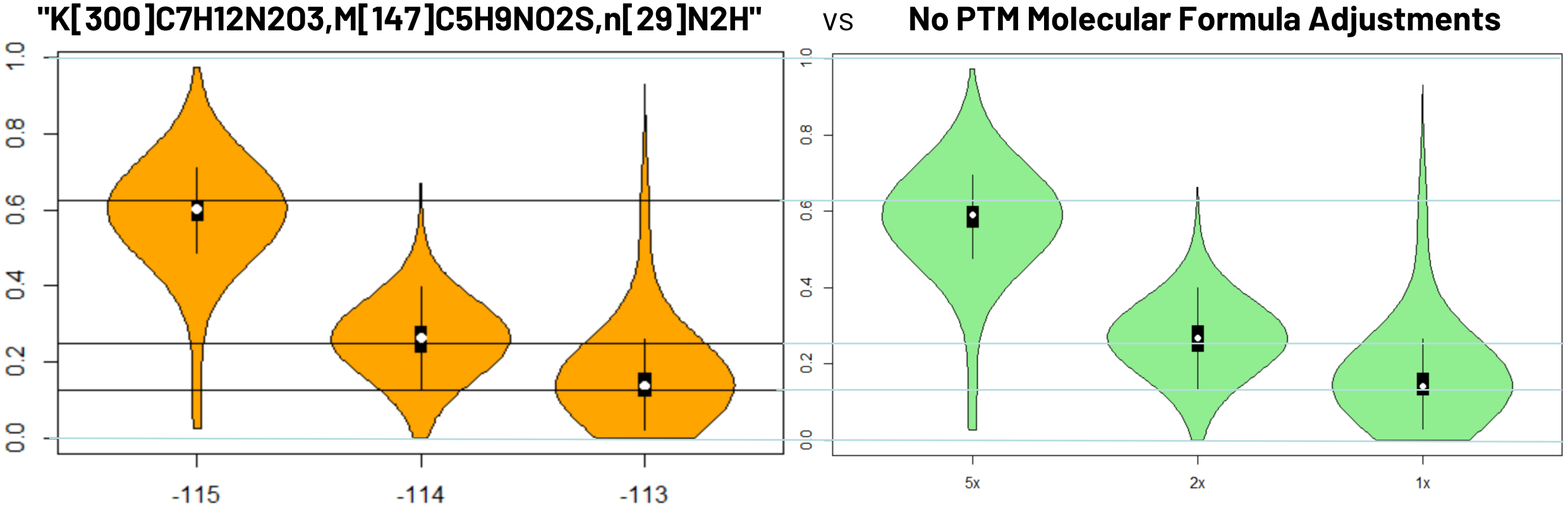


DIAMODE Corrected!

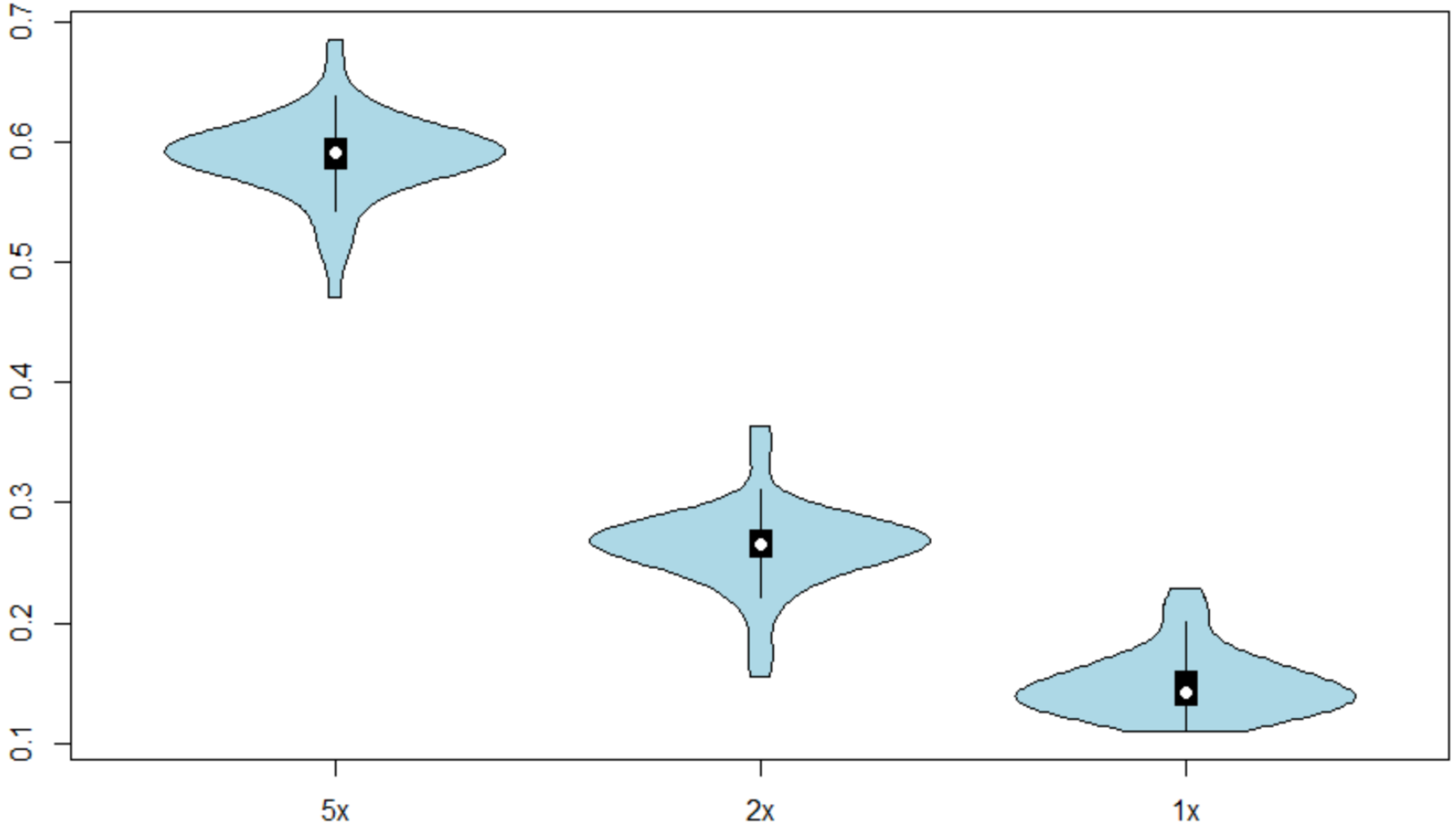


QuanTic : PTM Molecular Formula

- Quantic allows the user to specify PTM molecular formulas for better isotopic distribution estimation with Mercury8



QuanTic



Acknowledgements

- Mukul Midha
- David Campbell
- Zhi Sun
- Mike Hoopmann
- Kristian Swearingen
- Luis Mendoza
- Eric Deutsch
- Rob Moritz

- Thank you for your attention!