



Tutorial Webinar #15

Optimizing Large Scale DIA with Skyline

With

Brendan MacLean (Principal Developer, Skyline)

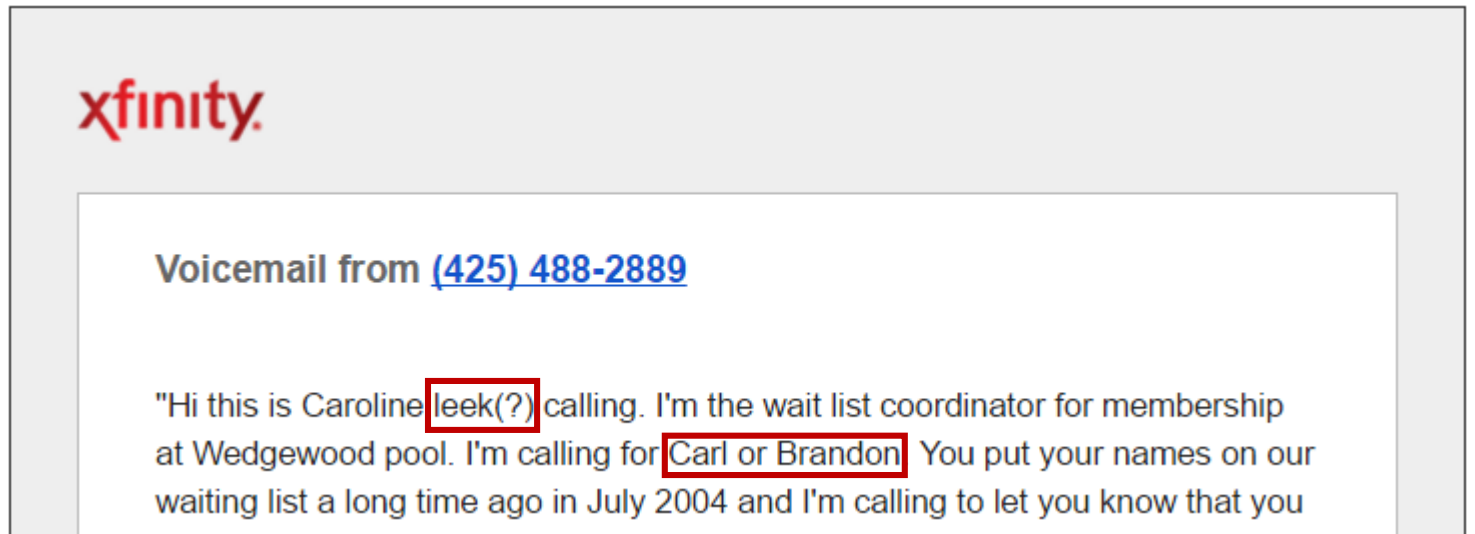
Agenda

- Welcome from the Skyline team!
- Optimizing Large Scale DIA with Skyline
 - Introduction and overview with Brendan MacLean
 - Tutorial with Brendan MacLean
- Audience Q&A – submit questions to Google Form:

<https://skyline.ms/QA4Skyline.url>

On Machine Learning

- Voice recognition (Xfinity and Verizon/Apple)
 - Brendan to Brandon
 - Collette to Carl, Colin, Colleen
 - Alair to Larry
- Statistical analyses
 - Name popularity?
 - Messages left for men v. women?



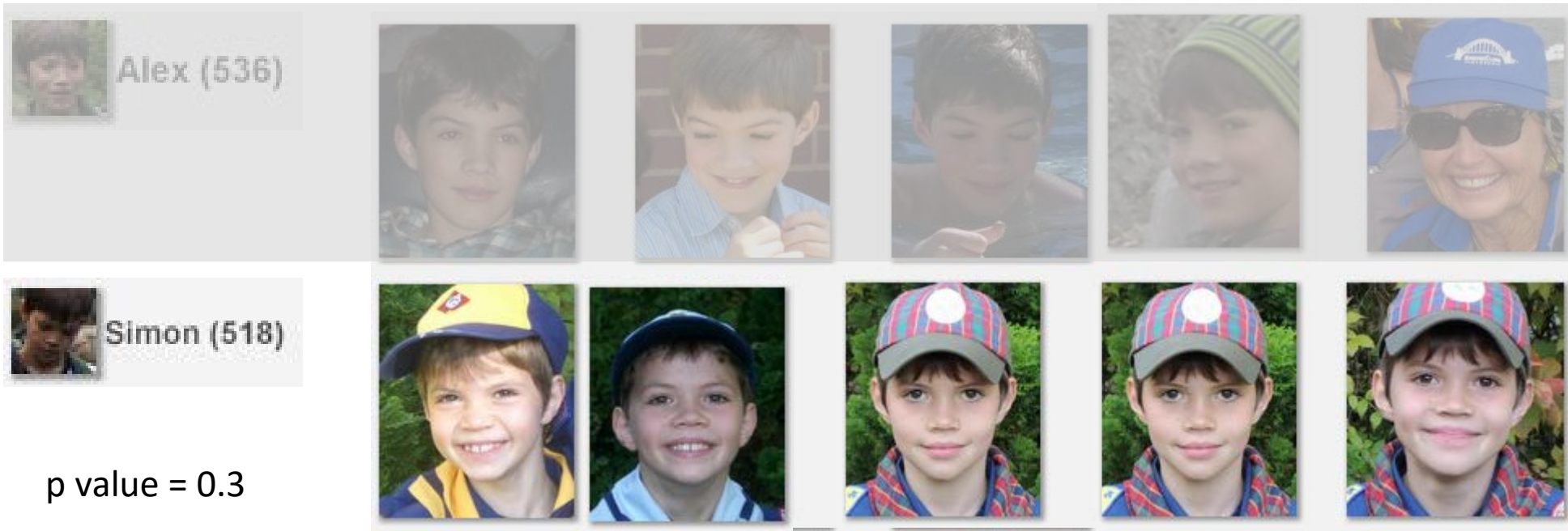
On Machine Learning

- Face Recognition (Google)

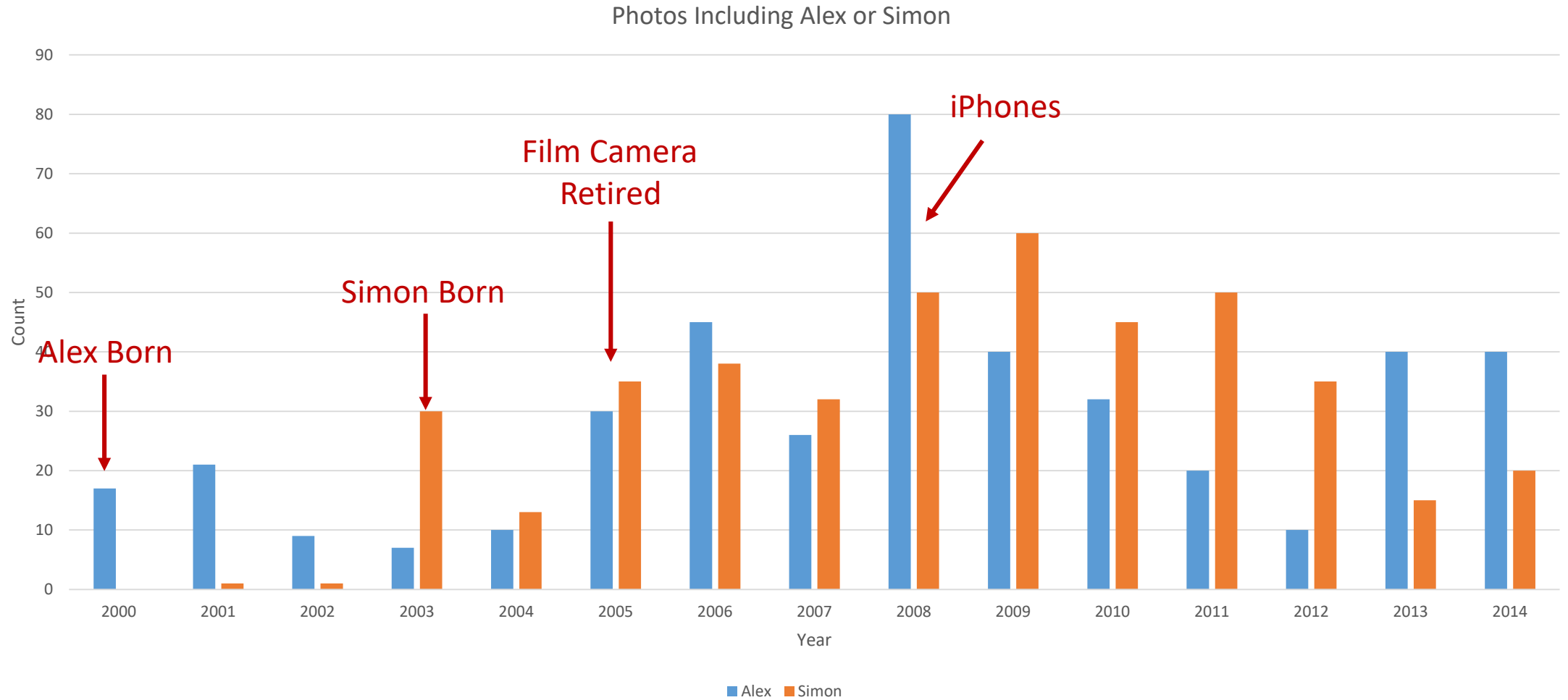


On Machine Learning

- Face Recognition (Google)



Statistical Analysis of Noisy Data



Webinar 8: DDA to Targeted

- ABRF iPRG 2014

	Fake Accession	Name	Origin	Molecular Weight
A	P44015	Ovalbumin	Chicken Egg White	45KD
B	P55752	Myoglobin	Equine Heart	17KD
C	P44374	Phosphorylase b	Rabbit Muscle	97KD
D	P44983	Beta-Galactosidase	Escherichia Coli	116KD
E	P44683	Bovine Serum Albumin	Bovine Serum	66KD
F	P55249	Carbonic Anhydrase	Bovine Erythrocytes	29KD



Sample Preparation

- ABRF iPRG 2014

	A	B	C	D	E	F (fmol)	
Sample 1	65	55	15	2	11	10	+ 200 ng yeast digest
Sample 2	55	15	2	65	0.6	500	+ 200 ng yeast digest
Sample 3	15	2	65	55	10	11	+ 200 ng yeast digest

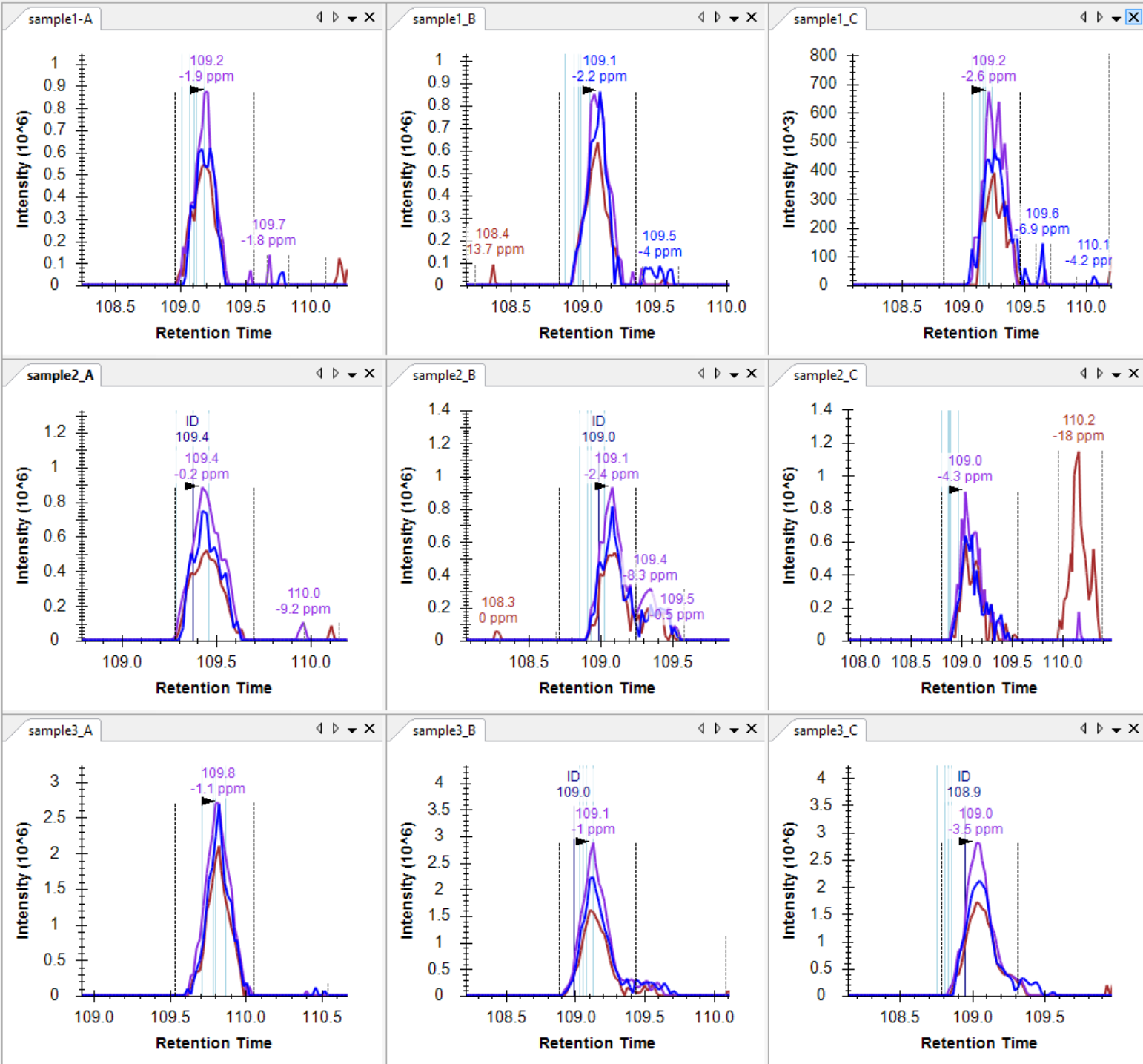
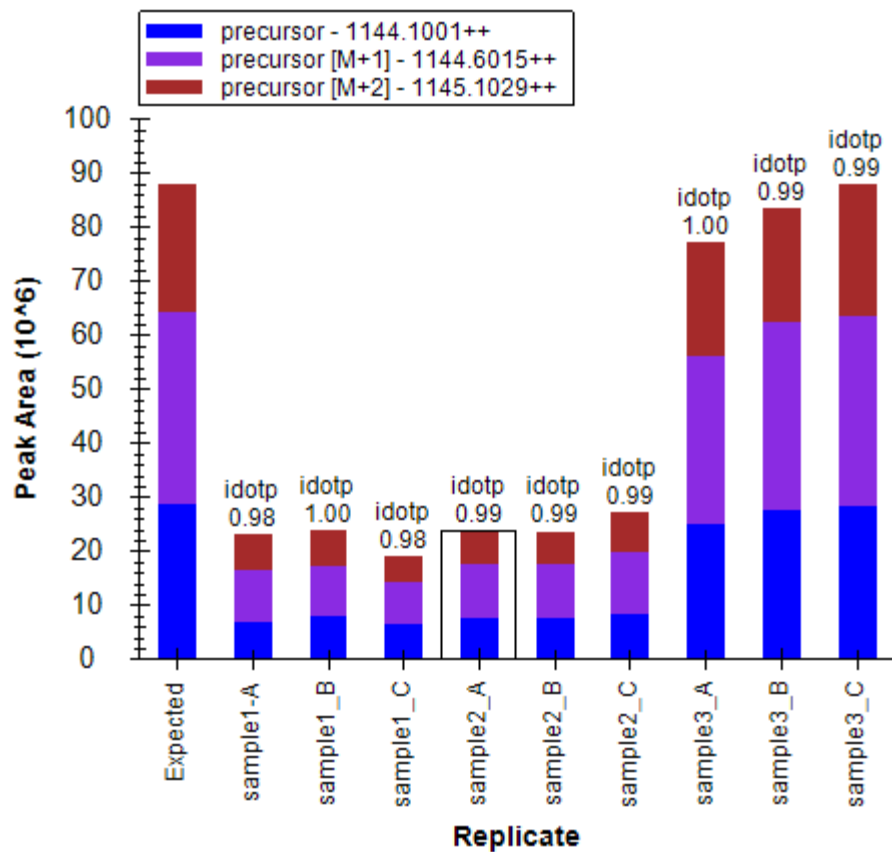
Peptide Differences (q value < 0.05)

- 29 Proteins (5 > 1 peptide)
- 63 Peptides (39 in proteins > 1 peptide)
- 24 Peptides in proteins with one significant peptide

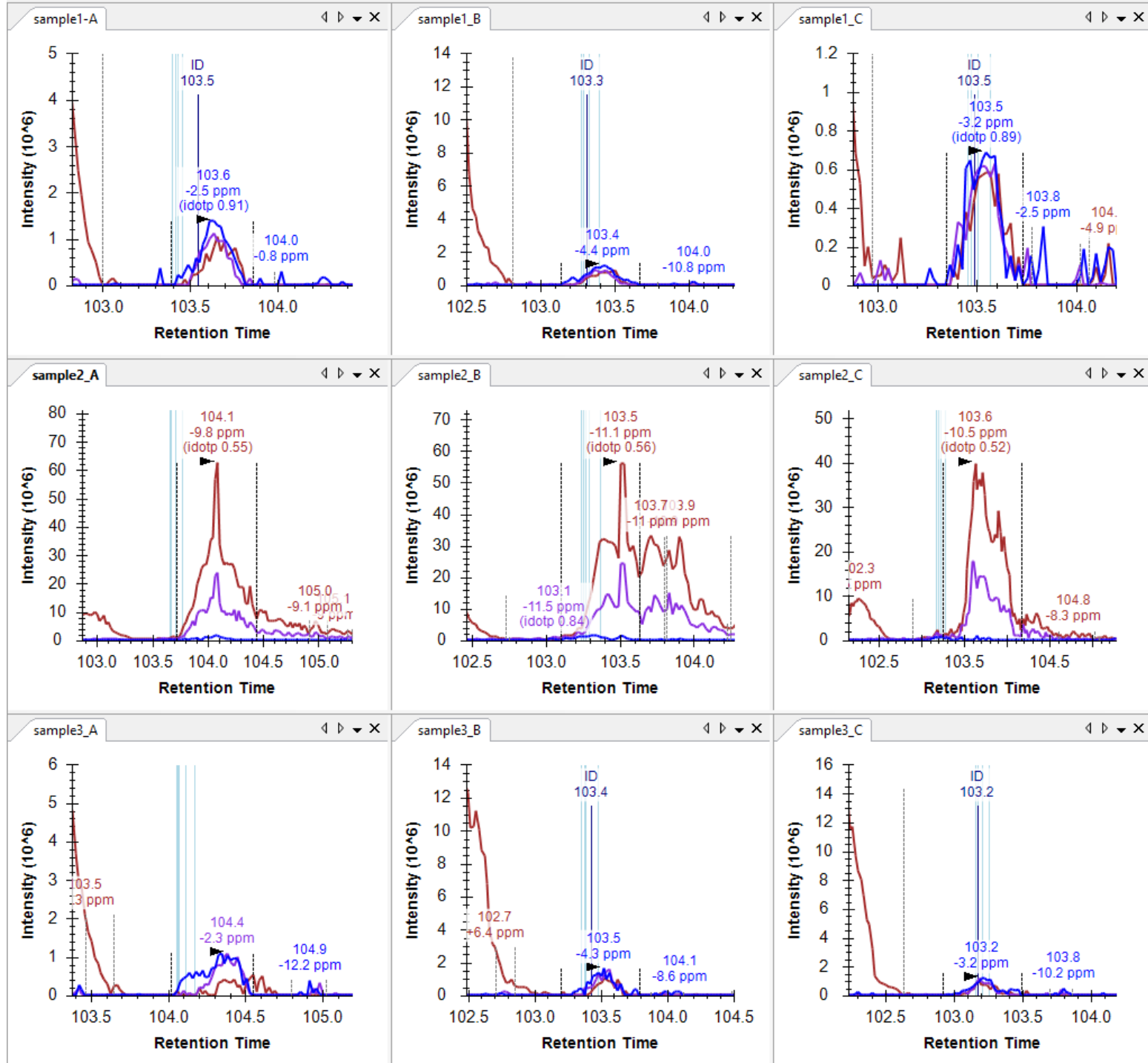
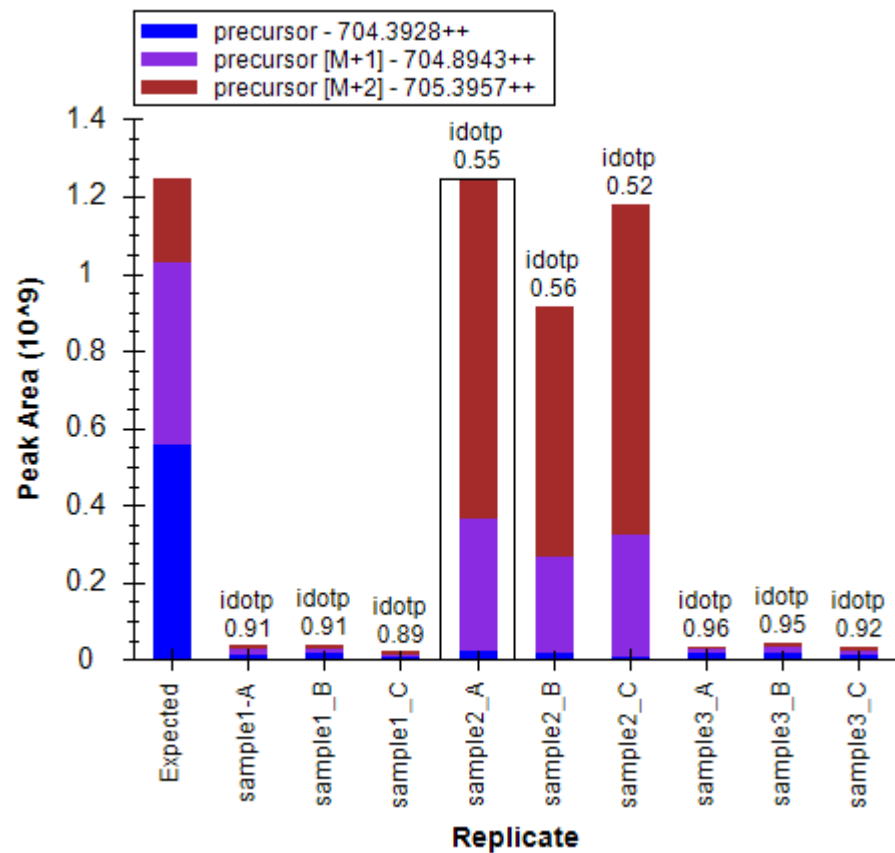
- MSstats protein comparison does better
(see Choi M, J. Proteome Res. 2017)

sp|P00330|ADH1_YEAST

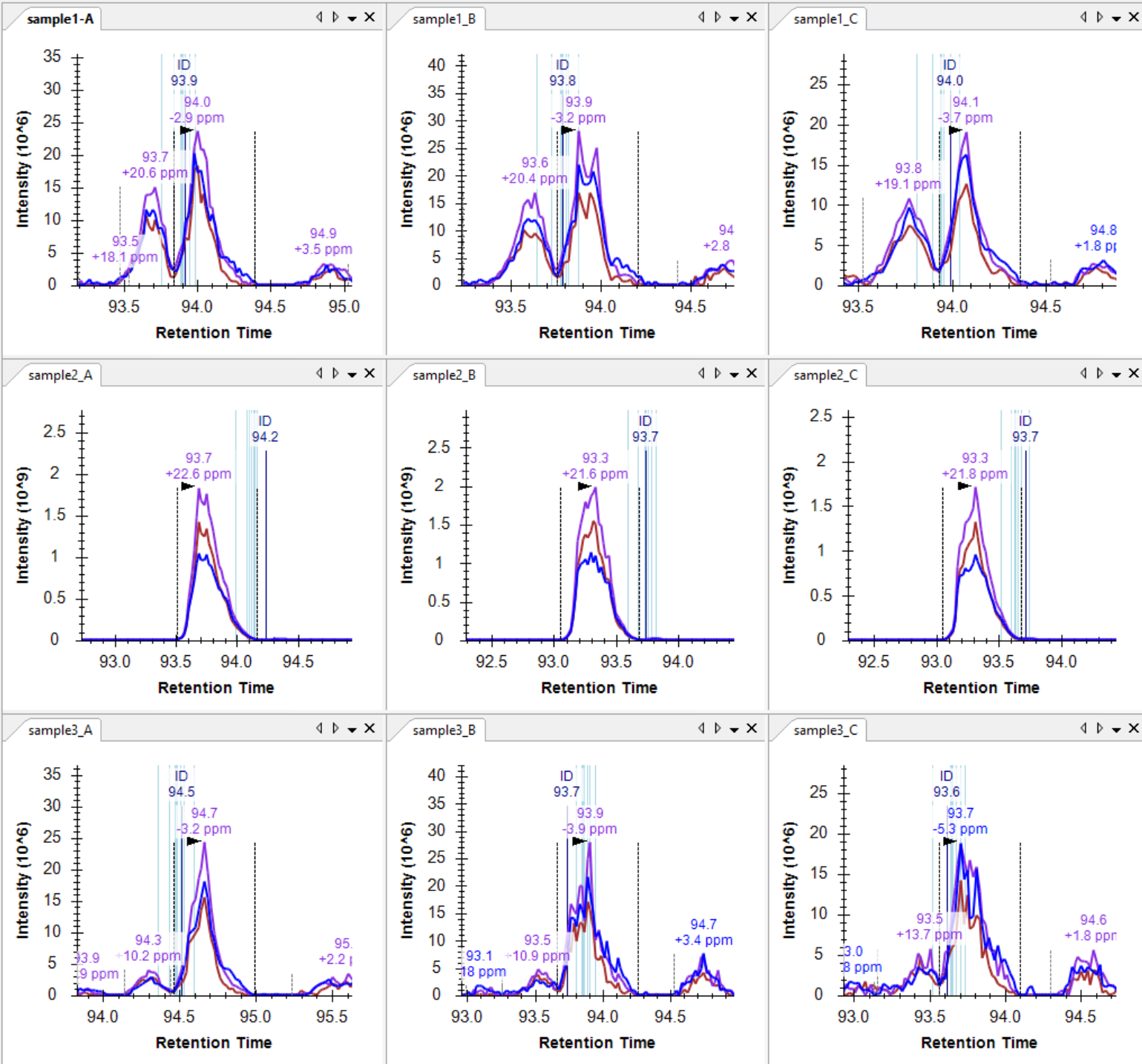
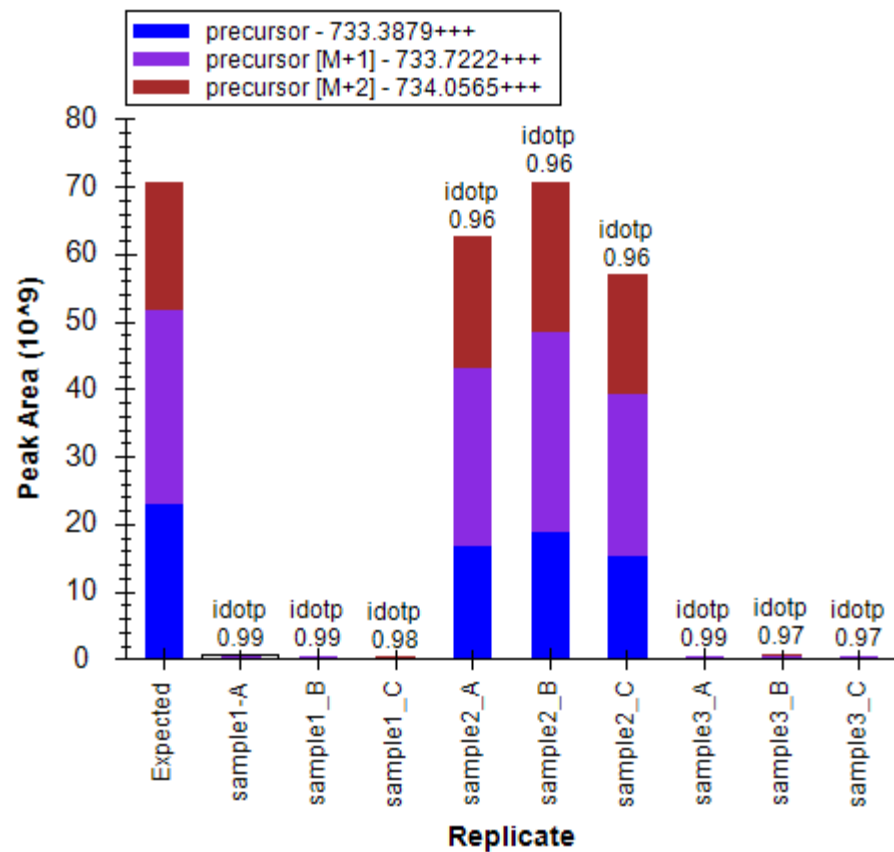
K.SANLMAGHWVAISGAAGGLGSLAVQYAK.A [164, 191]



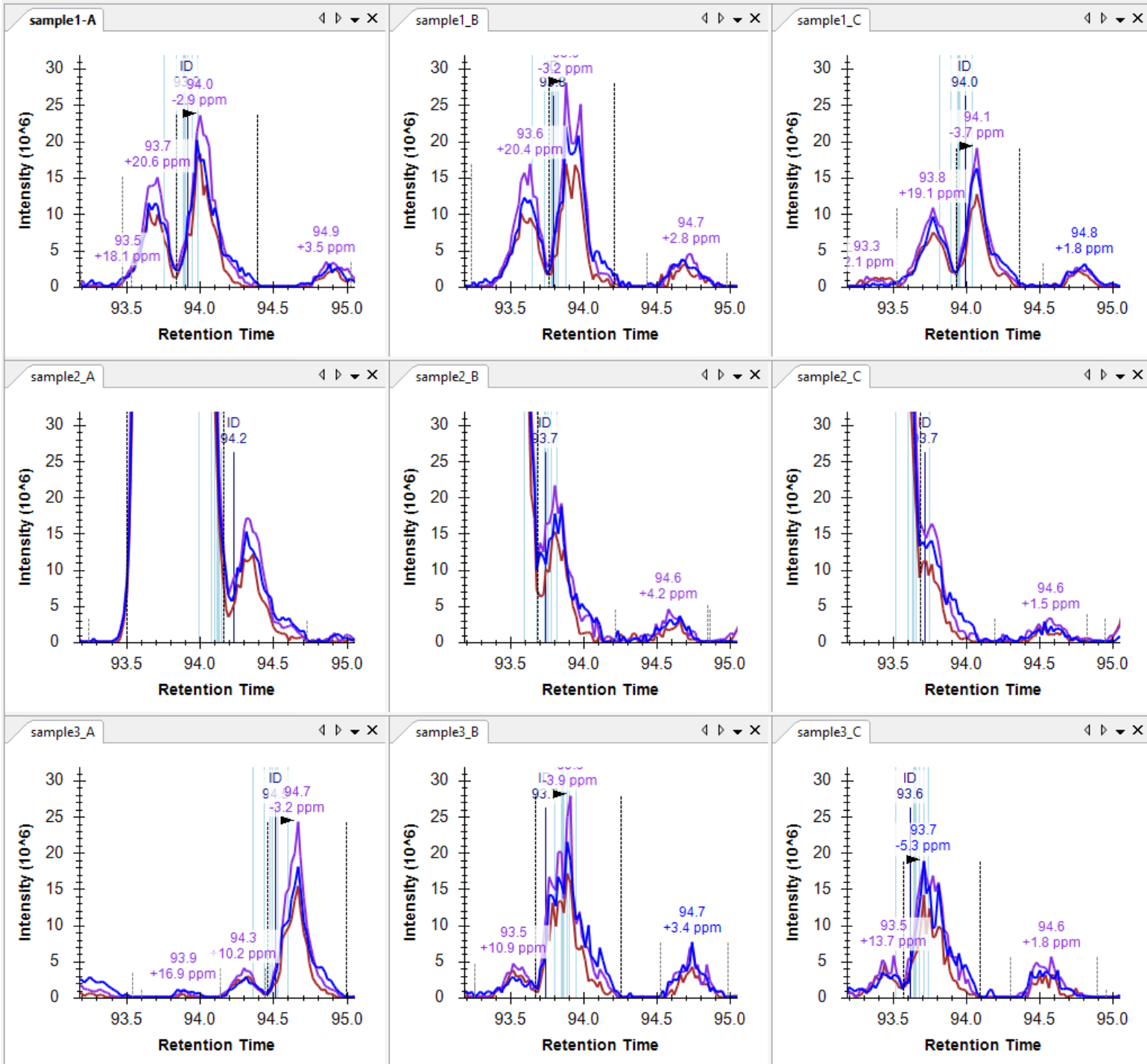
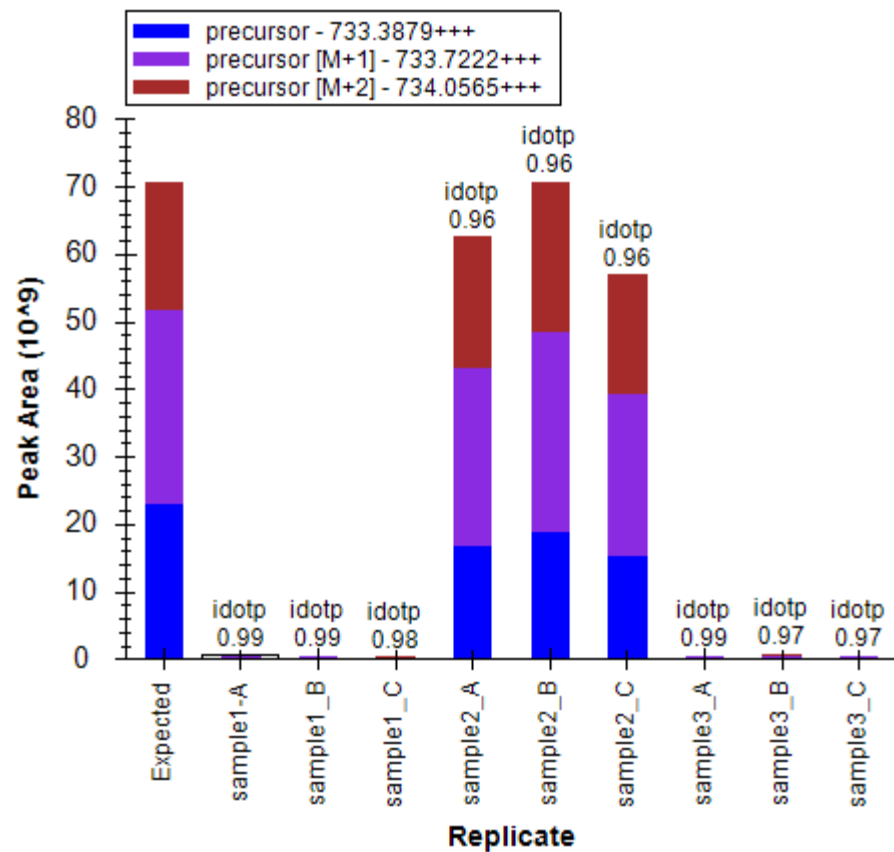
sp|O13525|COQ4_YEAST
R.LYNIYLPWAVR.T [260, 270]



sp|P33734|HIS5_YEAST
 K.IPVIASSGAGVPEHFEEAFLK.T [493, 513]



sp|P33734|HIS5_YEAST
K.IPVIASSGAGVPEHFEEAFLK.T [493, 513]



Skyline Compared (Statistically)

- Webinar 14: Large Scale DIA with Skyline



Navarro, et al. & Tenzer

A multicenter study, Nature Biotech. 2016



- <http://skyline.ms/webinar14.url>

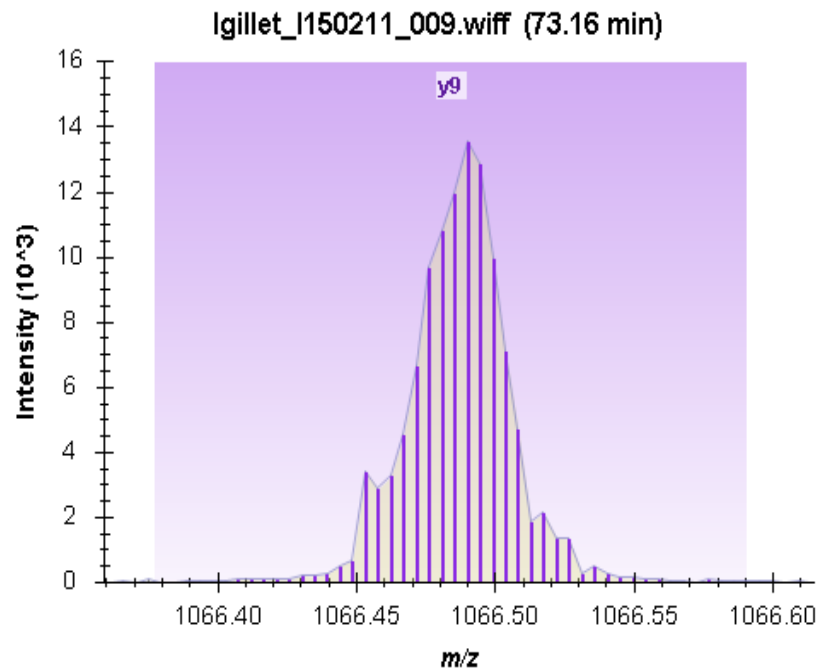
Before (May 2015)

- Spectronaut 42,439 peptides identified (56% more!)
- OpenSWATH 40,387 peptides identified (49% more!)
- DIA Umpire 31,256 peptides identified
- PeakView 28,424 peptides identified
- Skyline 27,121 peptides identified

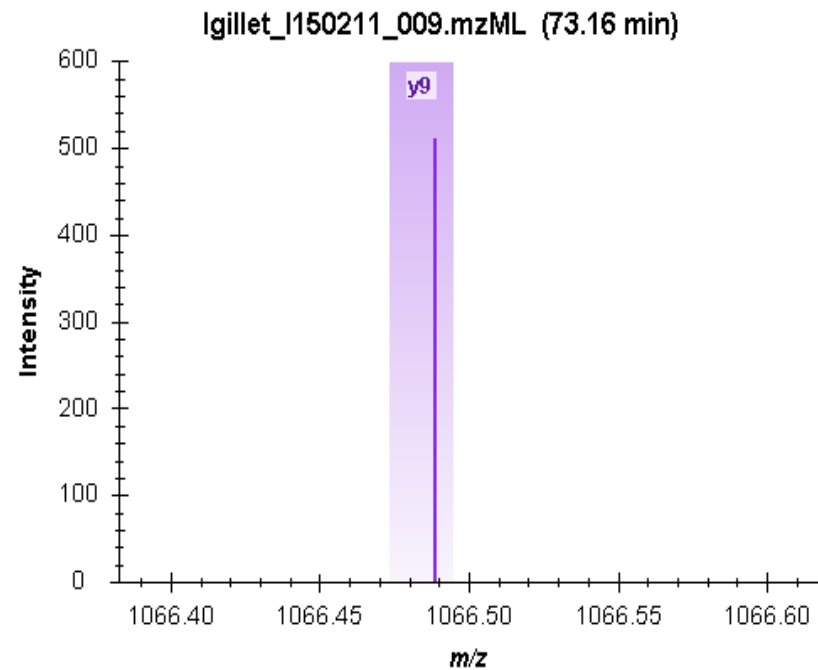
		First Iteration				Second Iteration			
		Identifications	Relative quantifications	separation Y-H	separation E - H	Identifications	Relative quantifications	separation Y-H	separation E - H
DIAumpire	peptides	31256	24456	0.936	0.944				
	proteins (int)	4351	3952	0.952	0.975				
	proteins (TOP3)	3654	3046	0.952	0.957				
openSWATH	peptides	40387	35443	0.966	0.974	40366	35383	0.969	0.978
	proteins (TOP3)	4700	4308	0.967	0.993	4705	4304	0.975	0.984
PeakView	peptides	28424	28424	0.966	0.970	28393	21011	0.984	0.974
	proteins (int)	5749	5749	0.971	0.965				
	proteins (TOP3)	3926	3926	0.971	0.983	3842	3042	0.971	0.989
Skyline	peptides	27121	18505	0.956	0.876				
	proteins (TOP3)	3468	2633	0.946	0.901				
Spectronaut	peptides	42439	37119	0.966	0.969	42325	36292	0.968	0.984
	proteins (TOP3)	4810	4363	0.974	0.962	4801	4314	0.975	0.981

Extraction Width

- 10,000 rp (default)



- ±10 ppm (centroided)



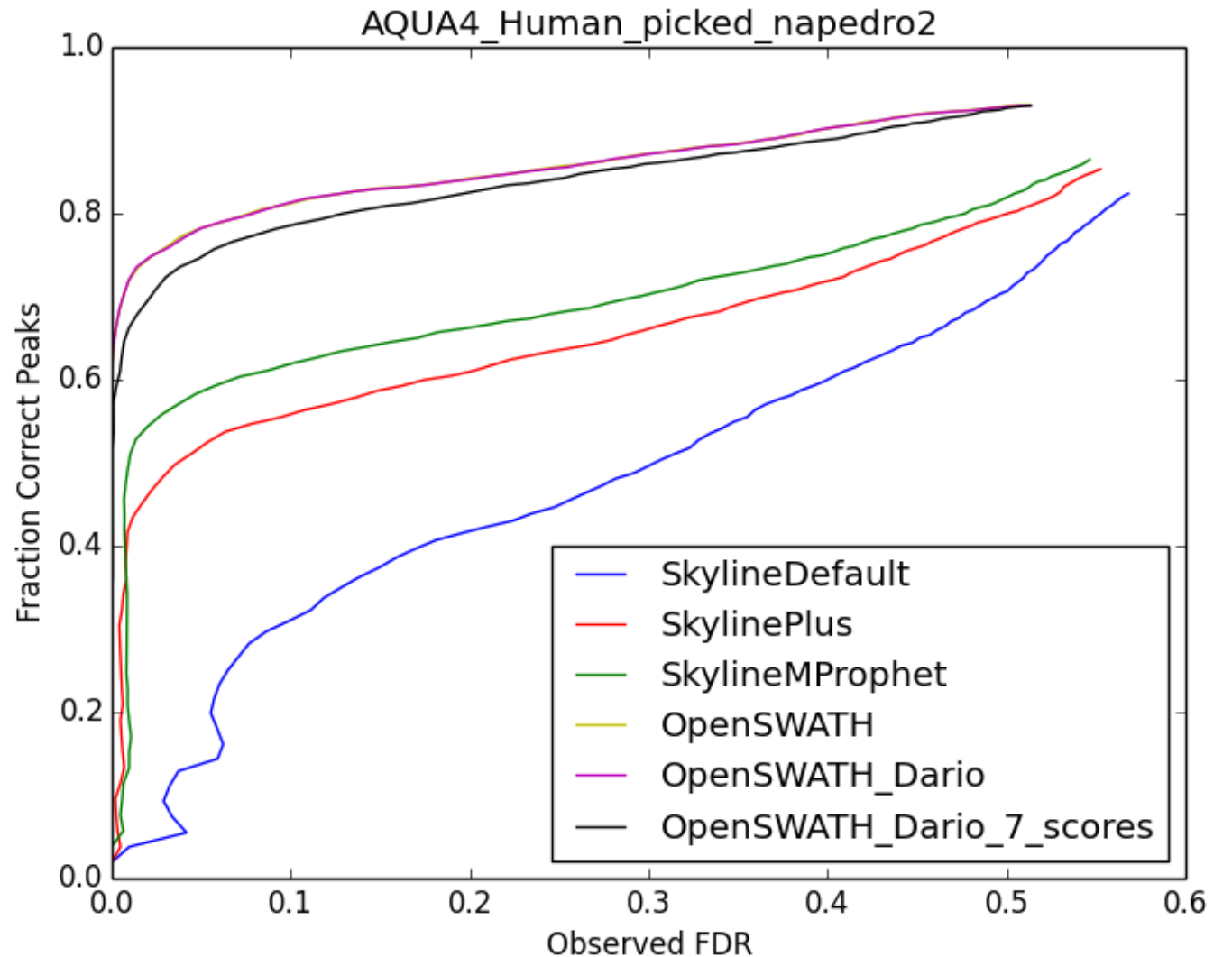
After (October 2016)

- Skyline 42,517 peptides identified (*16,500 human, 13,500 yeast, 12,000 E. coli*)
- Spectronaut 42,325 peptides identified
- OpenSWATH 40,728 peptides identified
- DIA Umpire 36,249 peptides identified
- SWATH 2.0 35,489 peptides identified

		Iteration 1					Iteration 2				
		Median CV of human	Number of IDs	Valid quantification ratios	Overlap yeast-human (arctanh)	Overlap <i>E. coli</i> -human (arctanh)	Median CV of human	Number of IDs	Valid quantification ratios	Overlap yeast-human (arctanh)	Overlap <i>E. coli</i> -human (arctanh)
Peptides	OpenSWATH	6.5%	40,726	36,098	2.12	2.24	8.2%	40,728	35,944	2.10	2.26
	SWATH 2.0	6.6%	35,517	35,517	2.14	2.11	6.1%	35,489	26,303	2.49	2.52
	Skyline	7.4%	40,804	34,103	1.91	1.85	6.9%	42,517	37,977	2.13	2.14
	Spectronaut	5.9%	42,439	37,120	1.97	1.90	6.2%	42,325	36,292	2.11	2.26
	DIA-Umpire	13.2%	36,332	28,785	1.74	1.98	12.9%	36,249	25,677	1.82	2.18
Proteins	OpenSWATH	5.4%	4,632	4,343	2.30	2.56	6.4%	4,636	4,352	2.51	2.60
	SWATH 2.0	5.6%	4,323	4,323	2.37	2.36	6.1%	3,946	3,371	2.42	2.56
	SWATH 2.0 (built-in)	6.0%	6,178	6,178	2.23	2.03					
	Skyline	5.9%	4,518	4,140	2.03	2.15	5.5%	4,692	4,456	2.37	2.43
	Spectronaut	3.3%	4,692	4,346	2.13	2.18	3.3%	4,675	4,300	2.31	2.50
	DIA-Umpire	12.0%	3,795	3,379	2.12	2.30	12.3%	3,673	3,111	2.13	2.85
	DIA-Umpire (built-in)	13.1%	4,849	4,489	1.78	1.94					

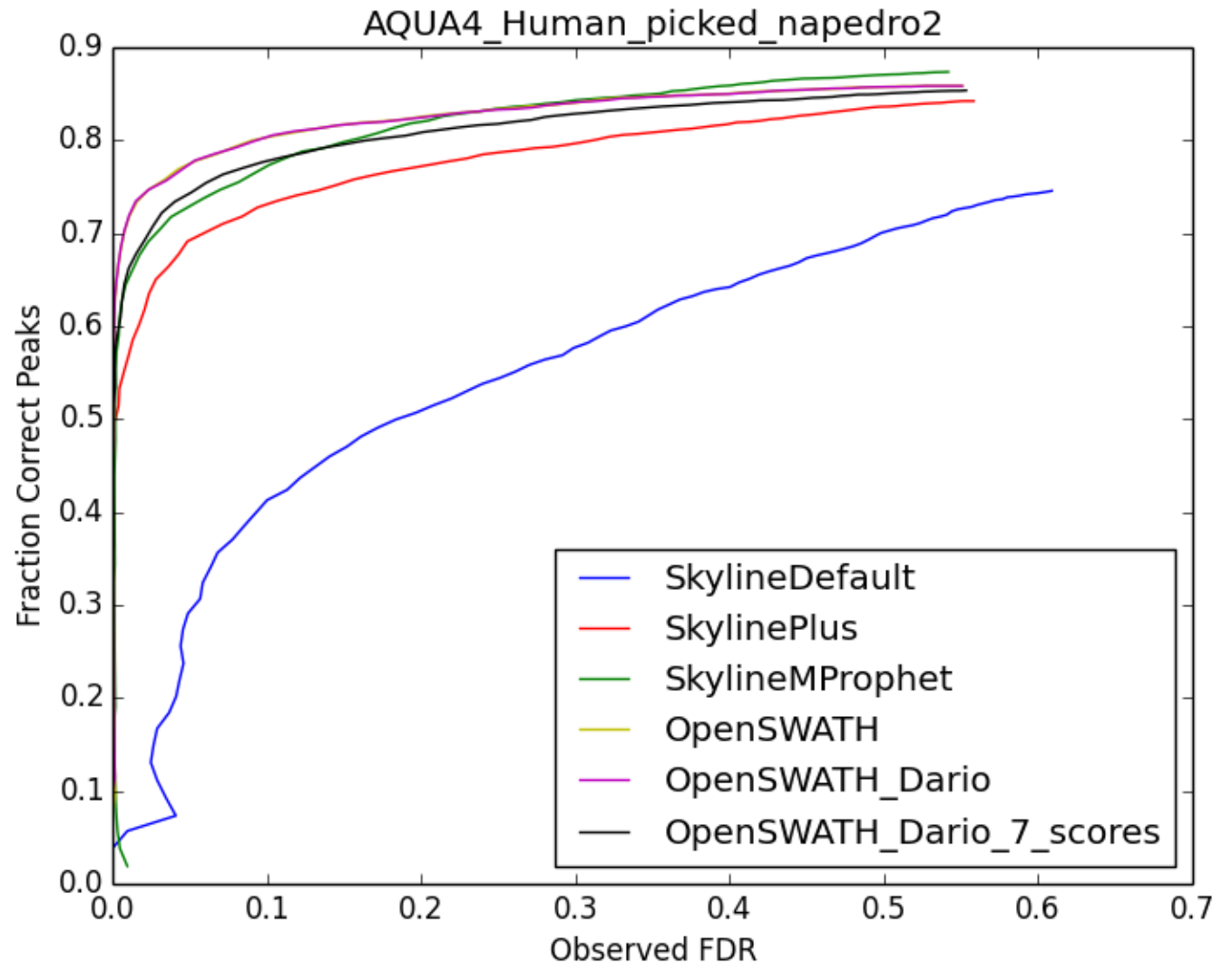
Long Before (January 2014)

- Dario Amodei
(Mallick Lab, Stanford)



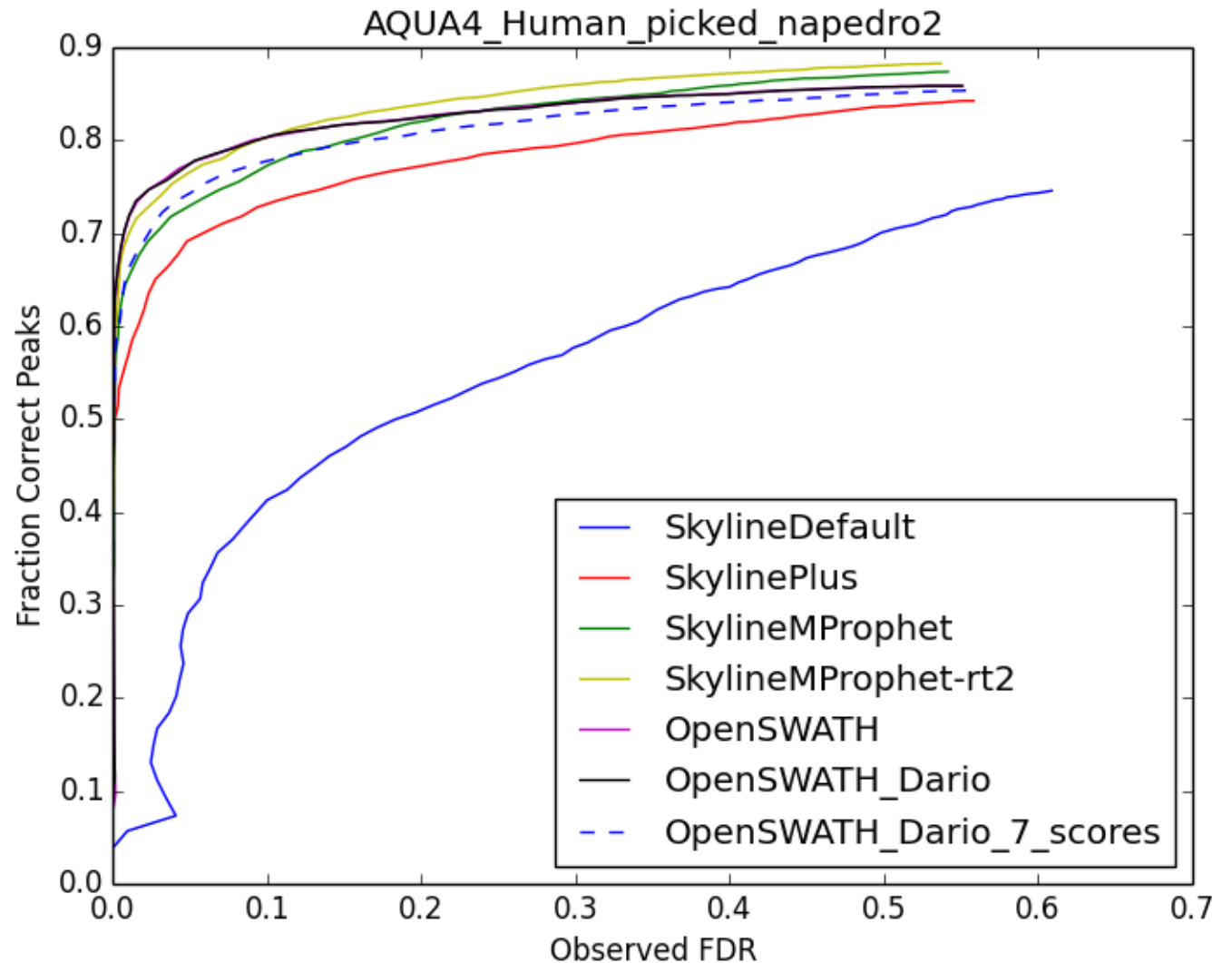
Shortly After (January 2014)

- Incompatibility between
 - New SCIEX wiff reader
 - DIA isolation scheme



The Next Day

- Adding $(RT \text{ delta})^2$

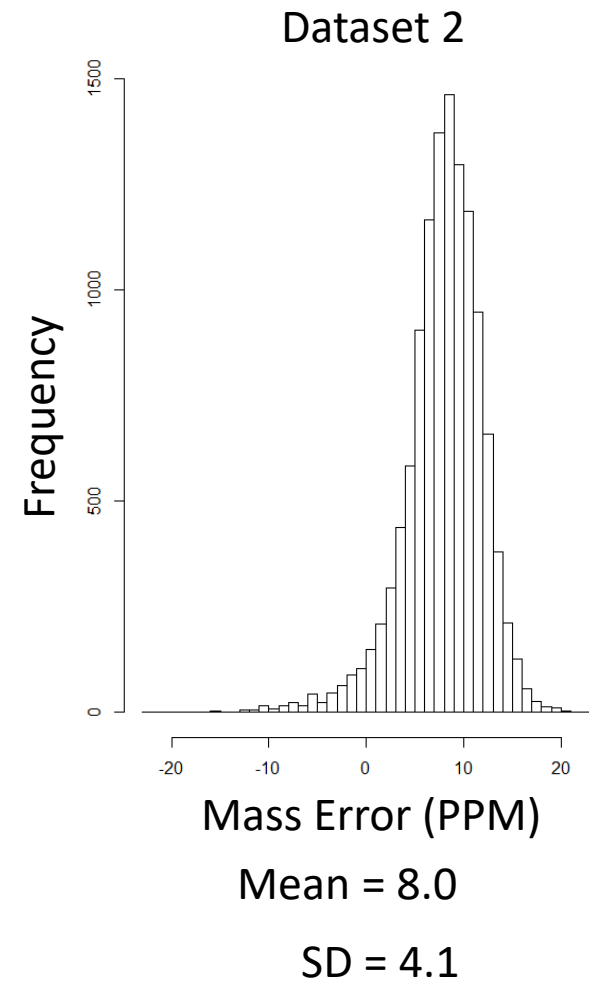
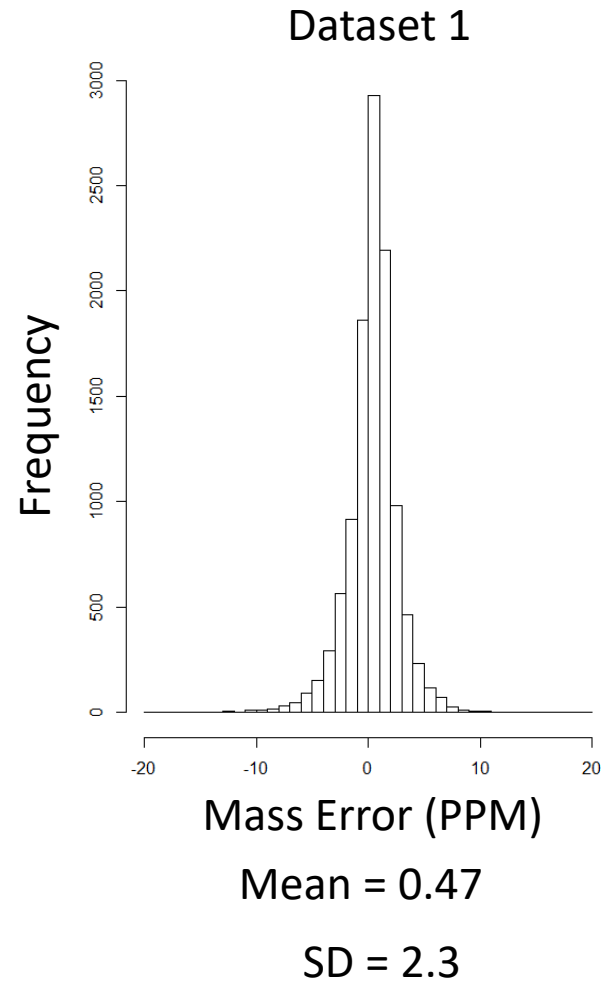


More Recently (June 2016)

- Bruker reports much higher detections with Spectronaut
 - MaxQuant DDA library with 19,500 peptide precursors
- Dataset 1
 - Spectronaut 10,500 peptide precursors detected
- Dataset 2
 - Spectronaut 17,500 peptide precursors detected
 - Skyline 14,000 peptide precursors detected
- 25% more!

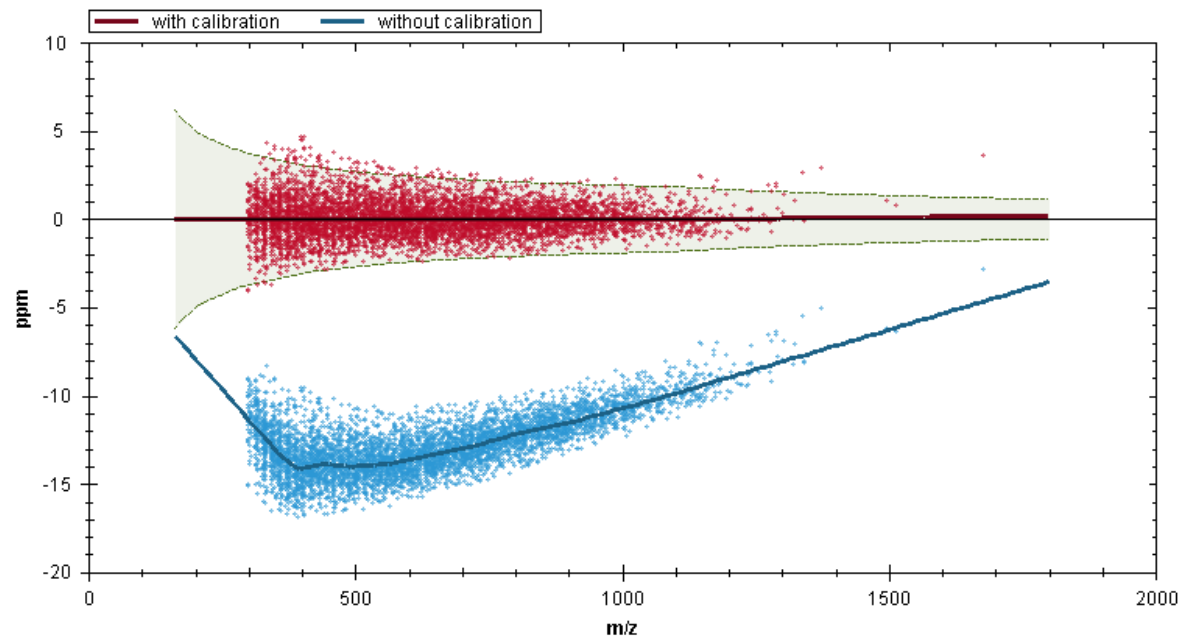
Mass Error Issue

- First clue histograms

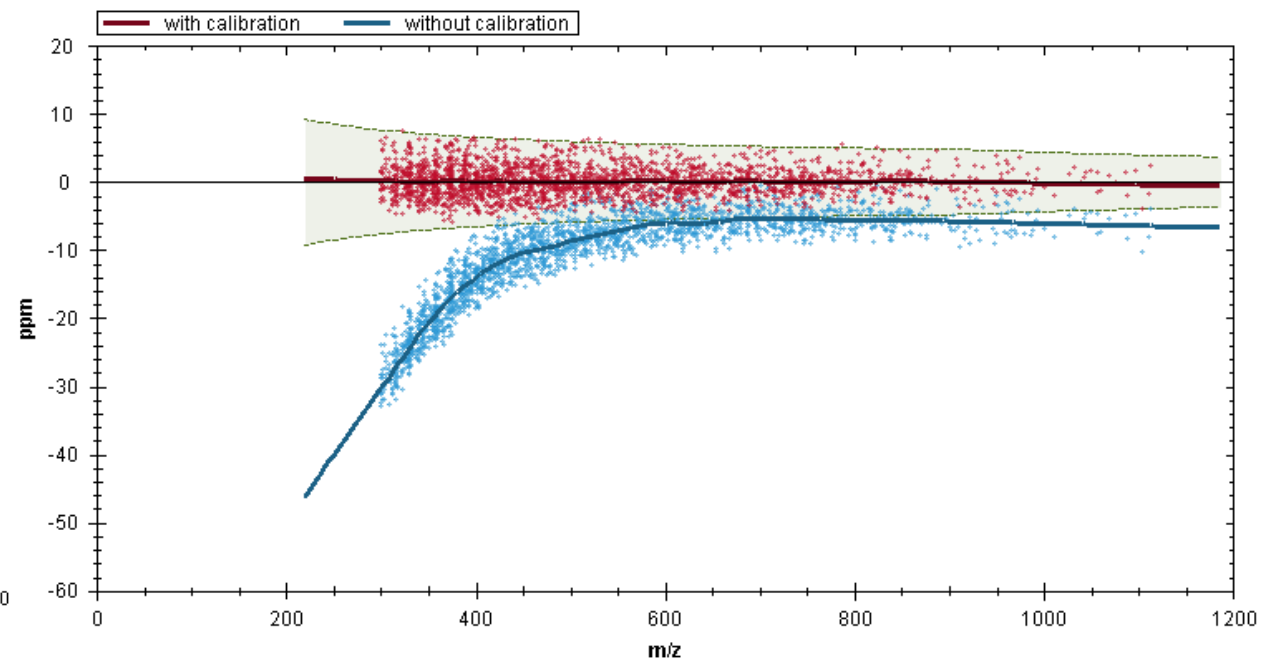


Spectronaut Mass Error Correction

Dataset 1

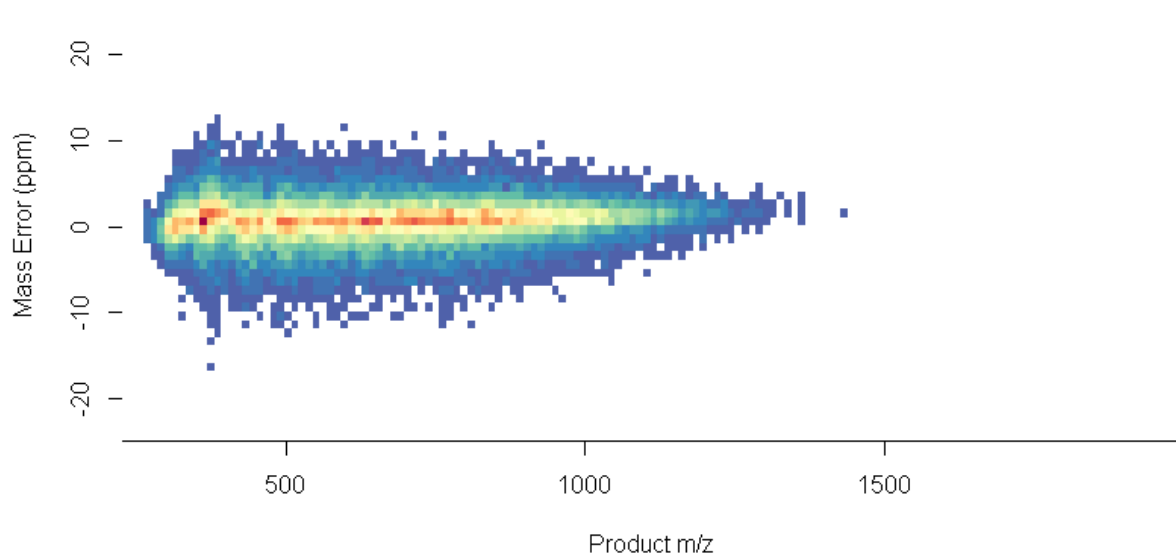


Dataset 2

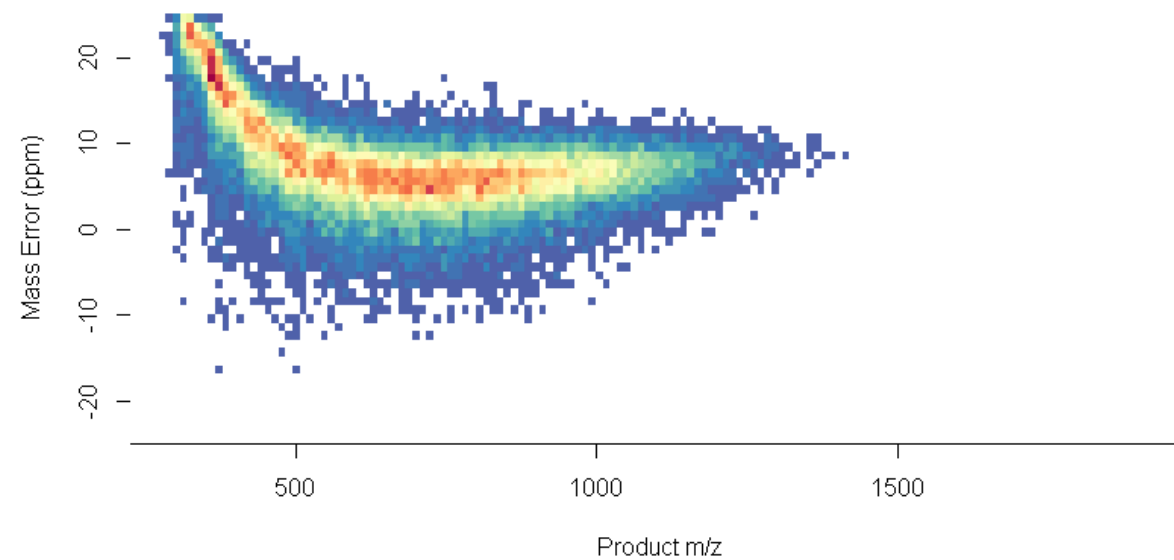


Mass Errors from Skyline Reports (R plots)

Dataset 1



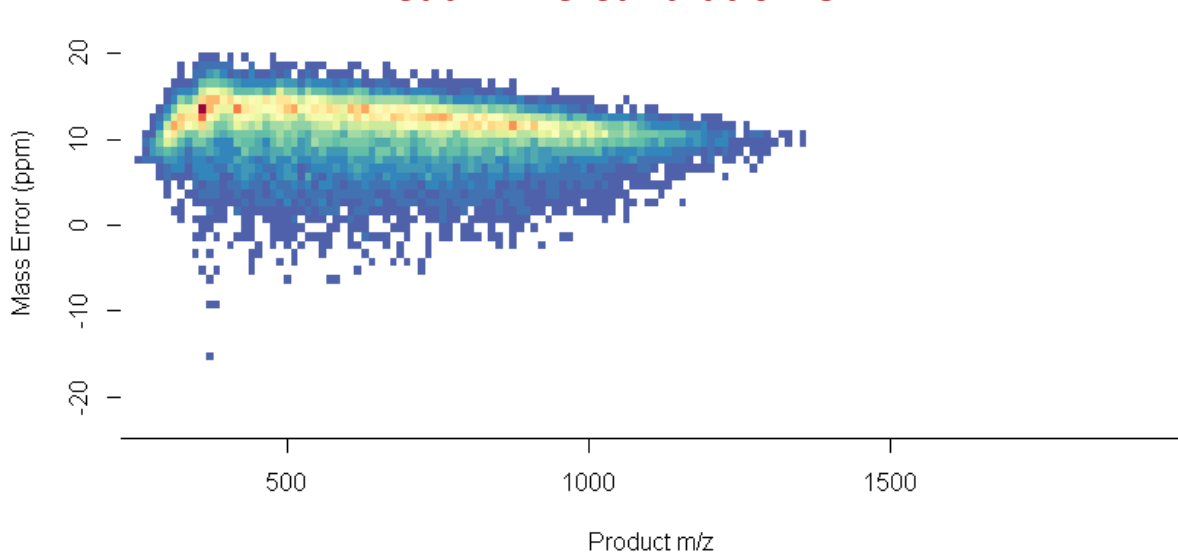
Dataset 2



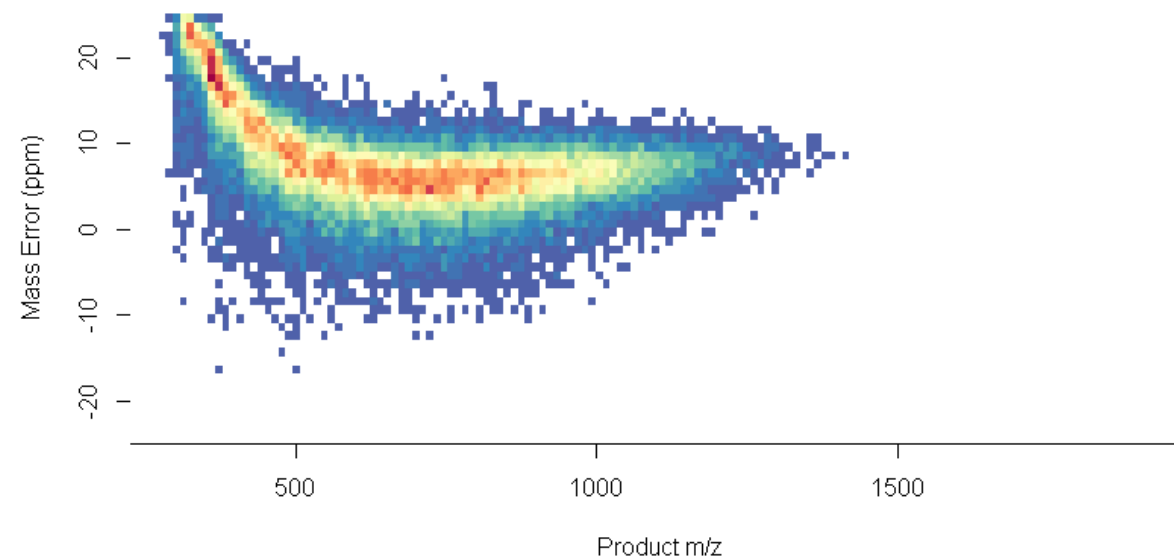
Mass Errors from Skyline Reports (R plots)

Dataset 1

Read-Time Calibration Off



Dataset 2



After Bruker Calibrated Dataset 2 (July 2016)

- Dataset 1
 - Skyline 13,800 peptide precursors detected
 - Spectronaut 10,500 peptide precursors detected
- Dataset 2
 - Skyline 16,700 peptide precursors detected
 - Spectronaut 17,500 peptide precursors detected (correction? 16,400)

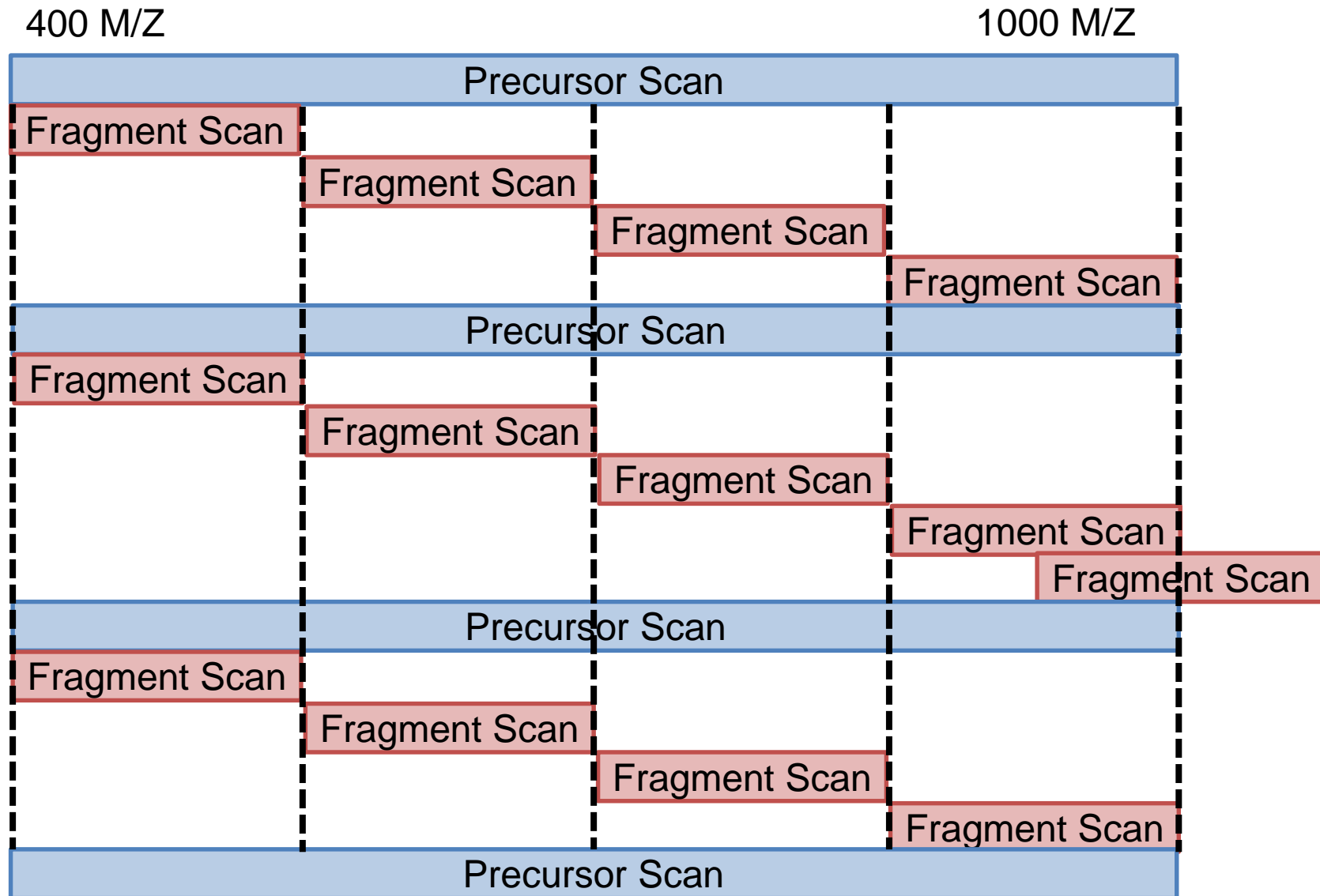
Bruker Dataset 1 and 2 Review in Skyline

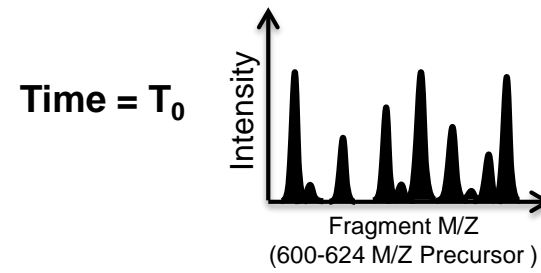
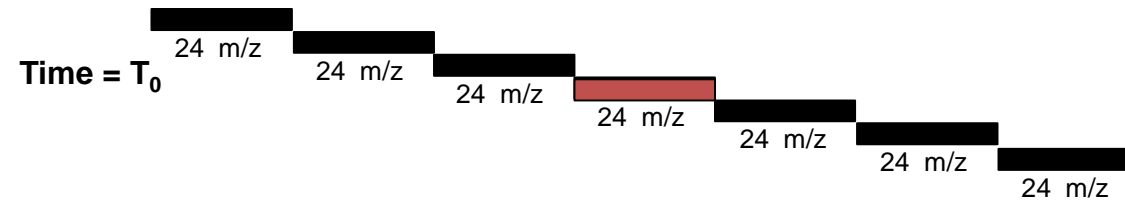
- New mass error plots (thanks to Alex)

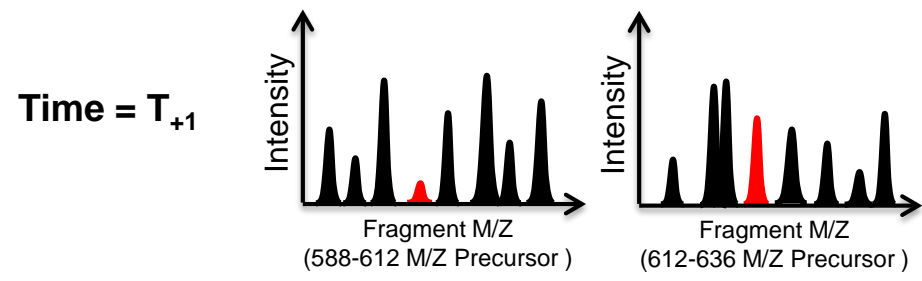
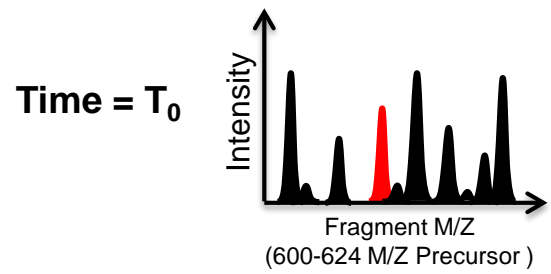
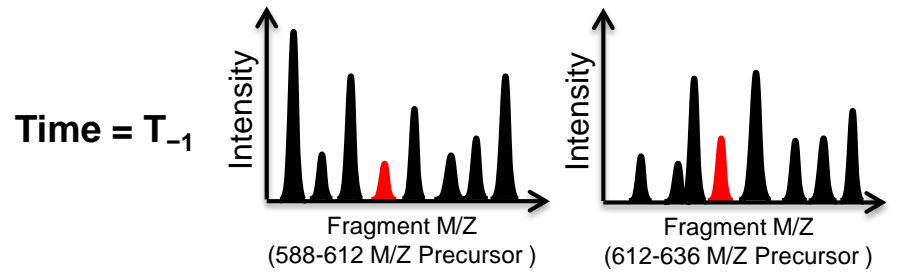
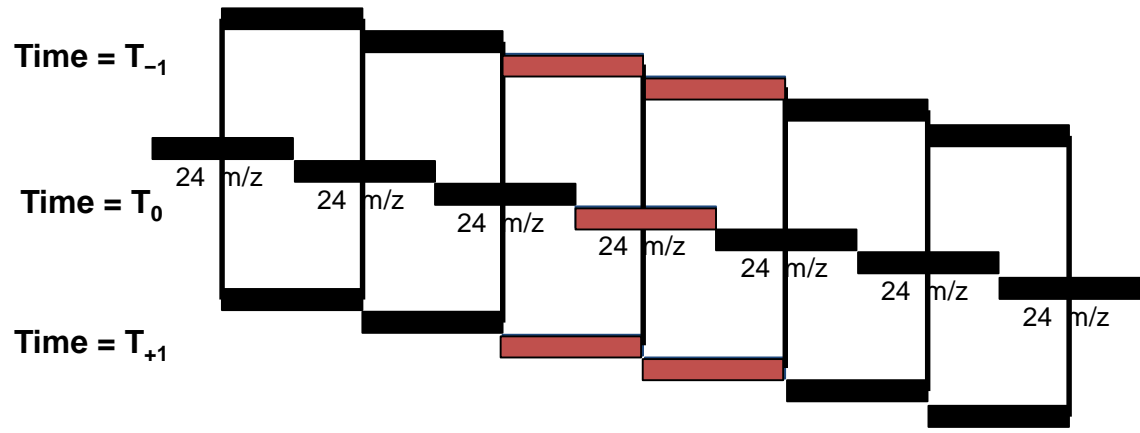
Dataset 3: MacCoss Lab Specified

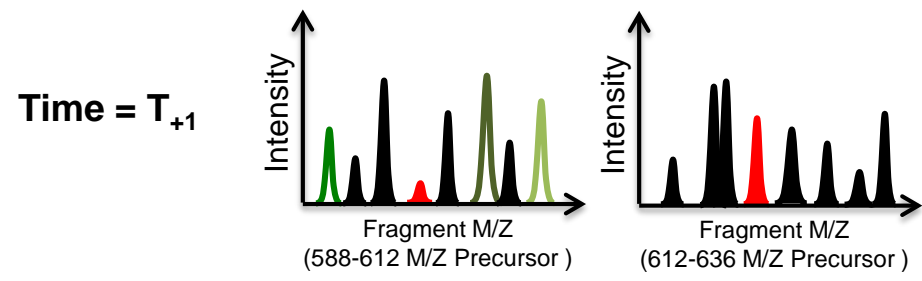
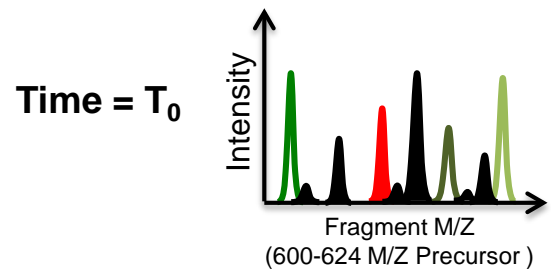
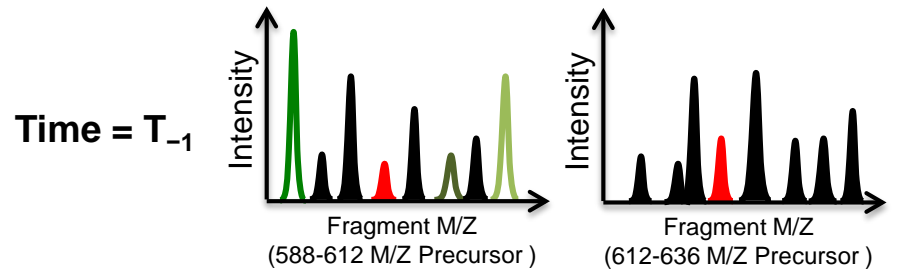
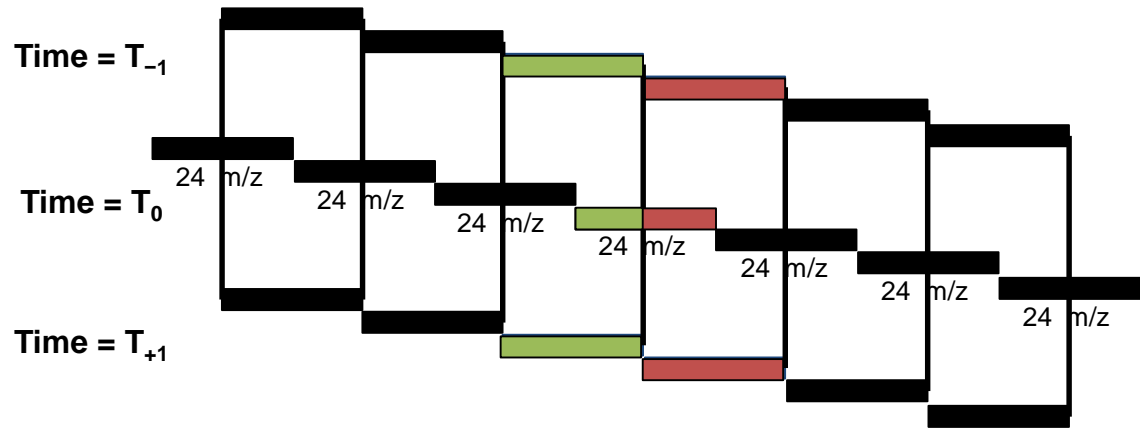
- Optimized window placement (no SWATH 0.5 m/z overlap)
- 2-cycle overlapping isolation scheme for demultiplexing
- 400-1000 m/z (oops 400-964 m/z) by 24 m/z windows
- Library from 12 DDA runs (6 Top20, 6 fIE)
 - 32,860 peptide precursors
 - 28,153 peptides
- Measurable in 400-964 m/z
 - 27,135 peptide precursors
 - 25,023 peptides

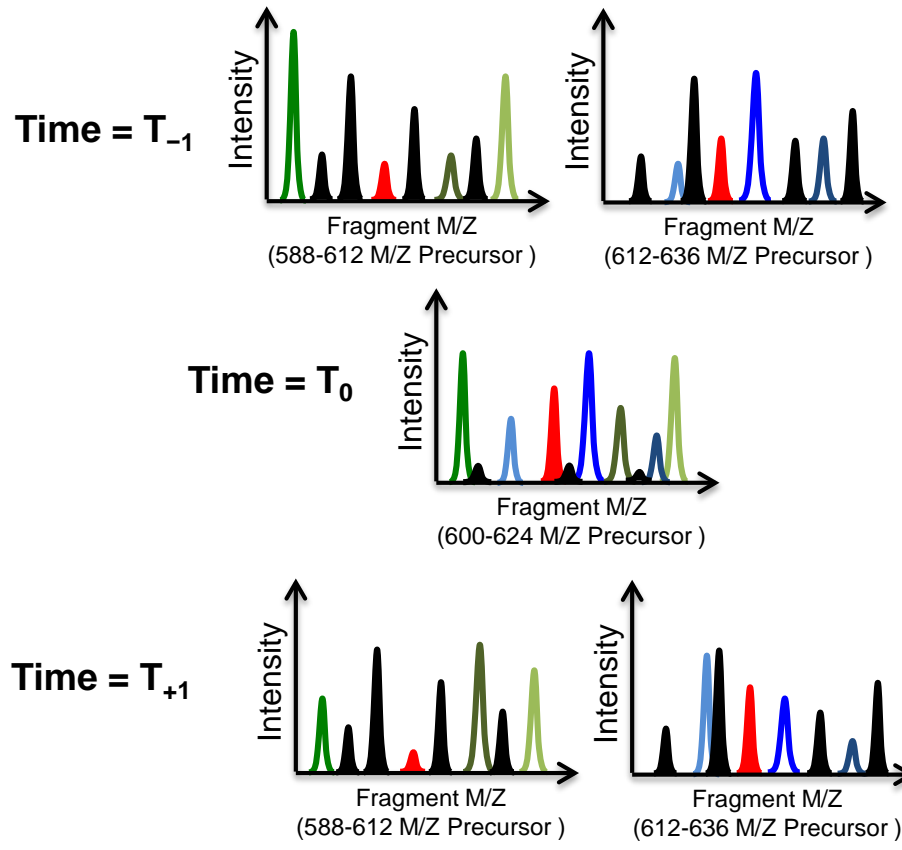
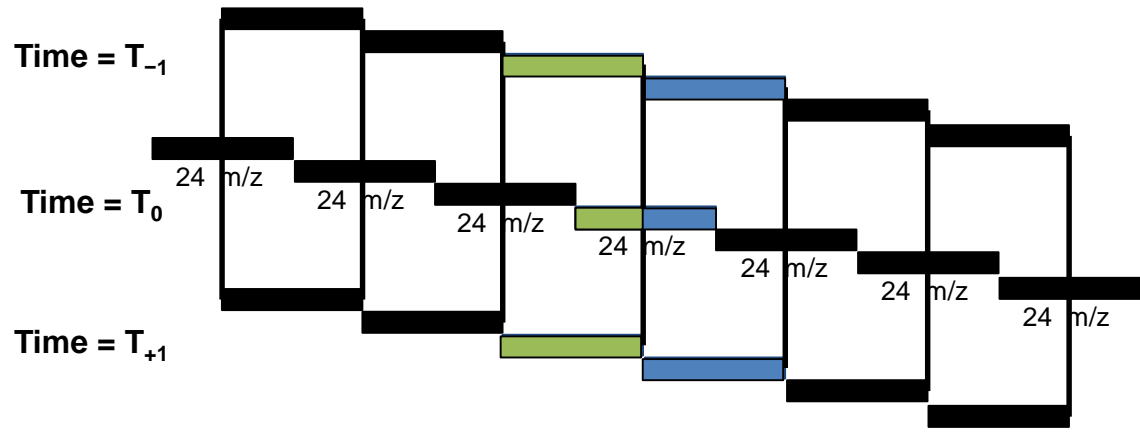
Overlapping Isolation Scheme

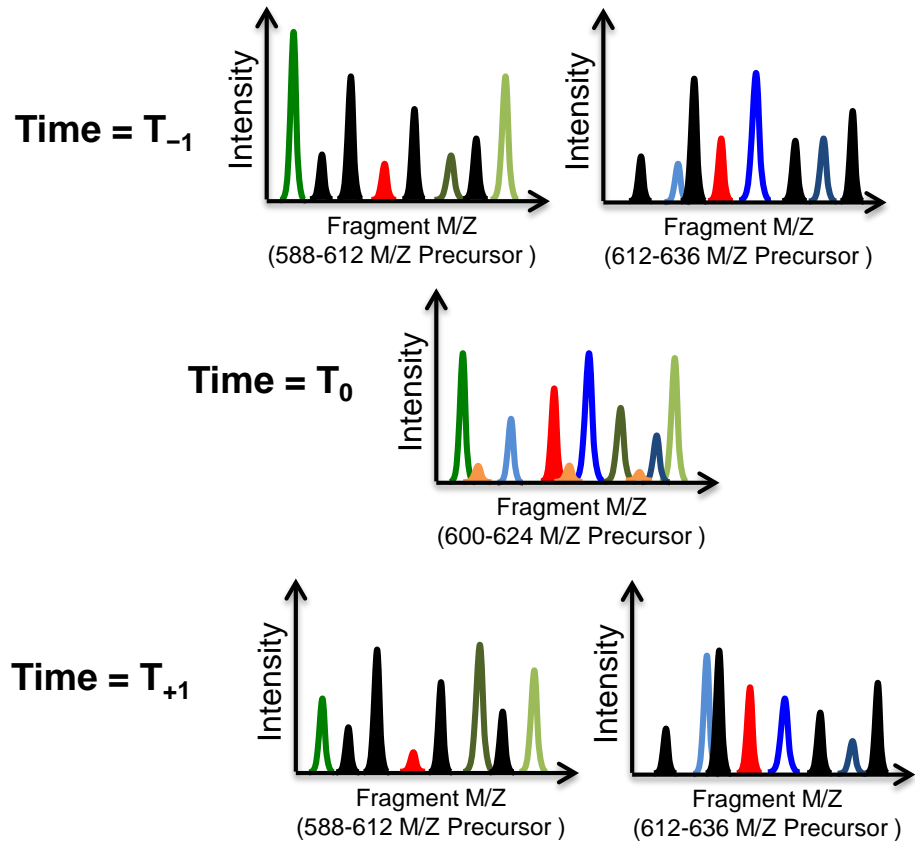
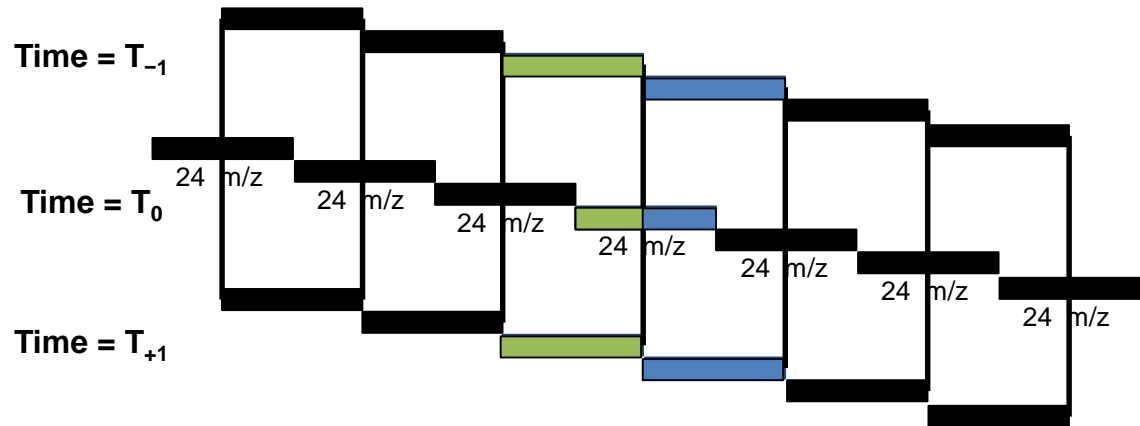


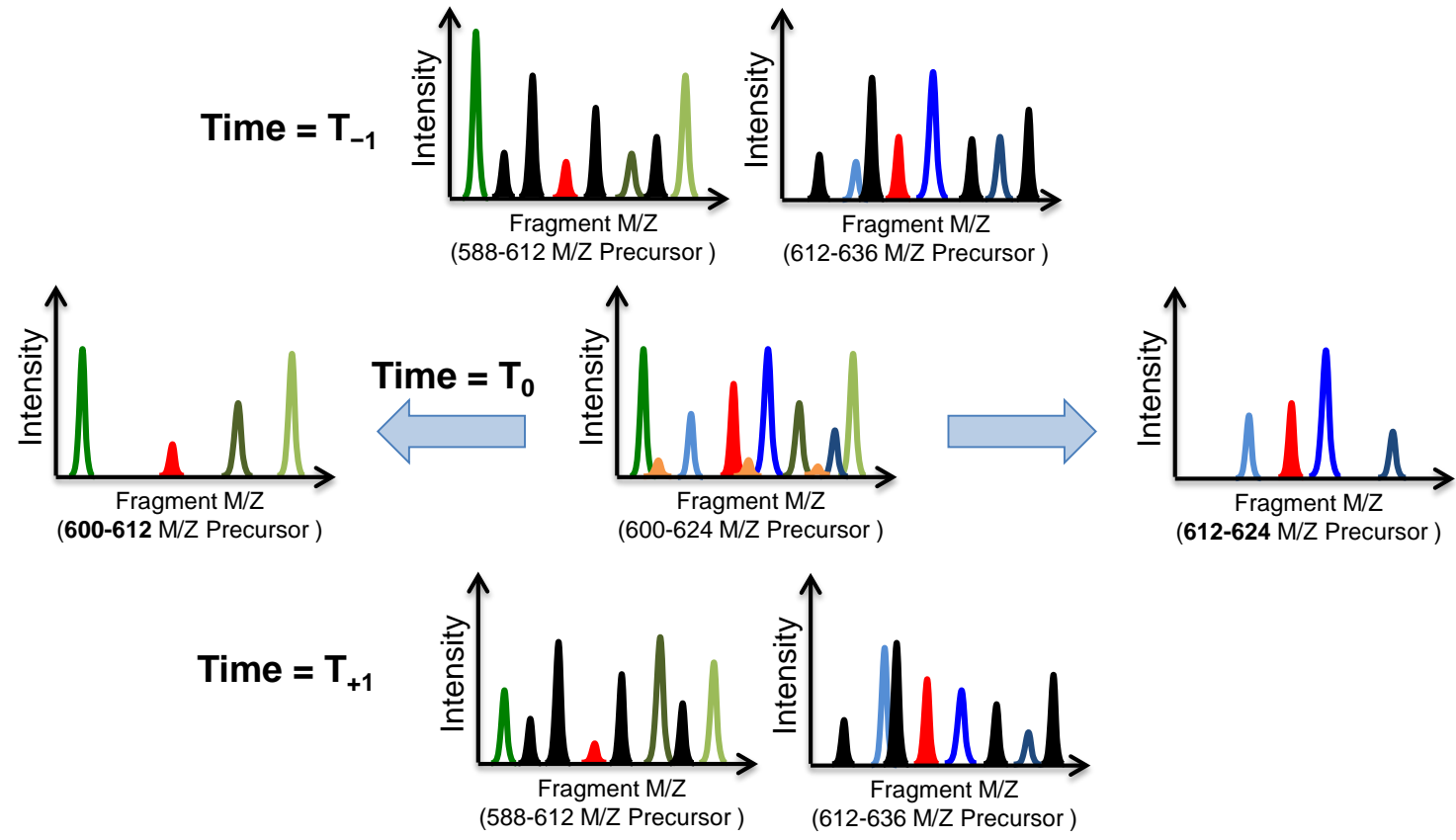
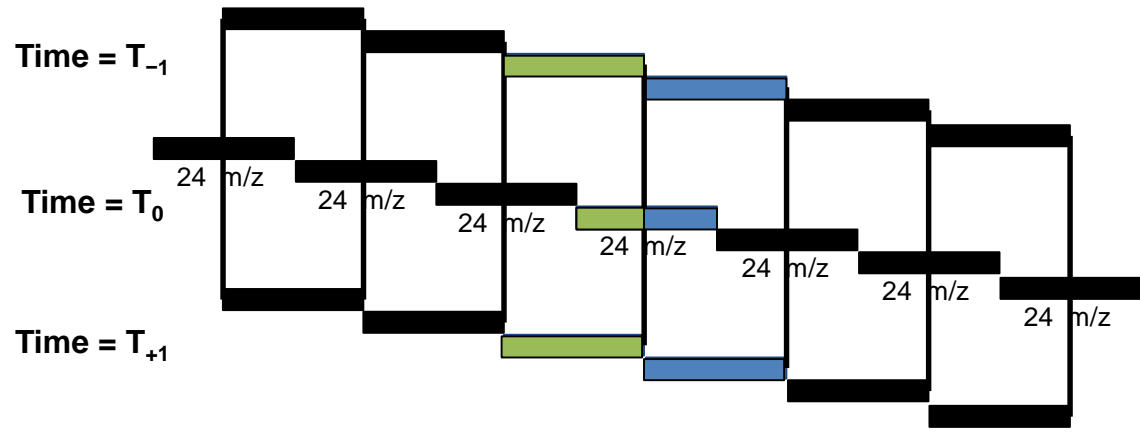












Processing Bruker Data Set 3 in Skyline

Learn More

- Webinar #16: Pan Human DIA with Skyline
 - 1 billion peaks
- Weeklong Courses 2017
 - Northeastern University, Boston – May 1-3 – Full
 - ETH, Zurich, June 26 – 30 – Full
 - University of Washington, Seattle – July 24-28 – Register now!
- Workshops and Conferences 2017
 - Short Course #21: Case Studies in Quantitative Proteomics at ASMS, Indianapolis – June 3-4 – Register now!
 - Skyline User Group Meeting at ASMS, Indianapolis – June 4 – Register now!

Listings updated in **Join Us** section of Skyline homepage:

<https://skyline.ms/Skyline.url>

Questions?

- Ask any questions at the following form:

<https://skyline.ms/QA4Skyline.url>

- Take the post-webinar survey:

<https://skyline.ms/survey4webinar.url>



Skyline Tutorial Webinar #15

This ends this Skyline Tutorial Webinar.

Please give us feedback on the webinar at the following survey:

<https://skyline.ms/survey4webinar.url>

A recording of today's meeting will be available shortly at the Skyline website.

We look forward to seeing you at a future Skyline Tutorial Webinar.