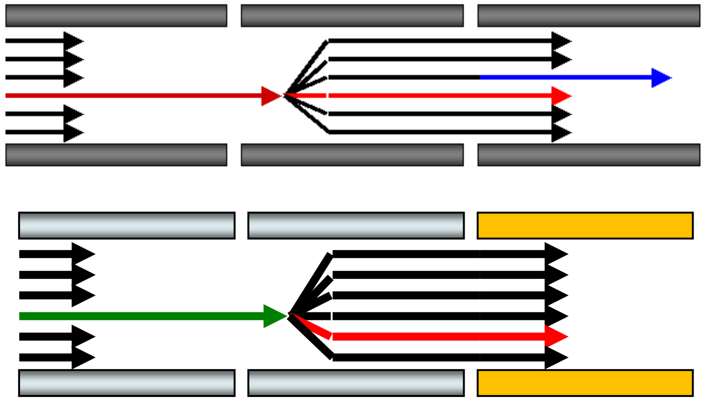
# Skyline integrates the Prosit prediction server

## for proteome-wide DIA data analysis using on-demand fragment intensity and iRT prediction

Brendan MacLean

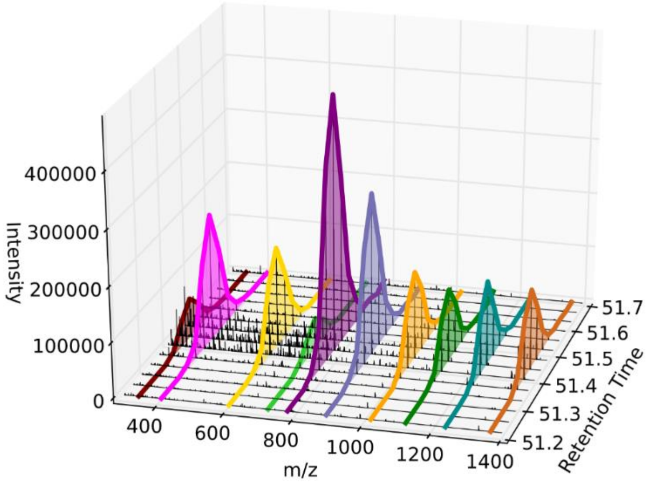
UNIVERSITY *of* WASHINGTON

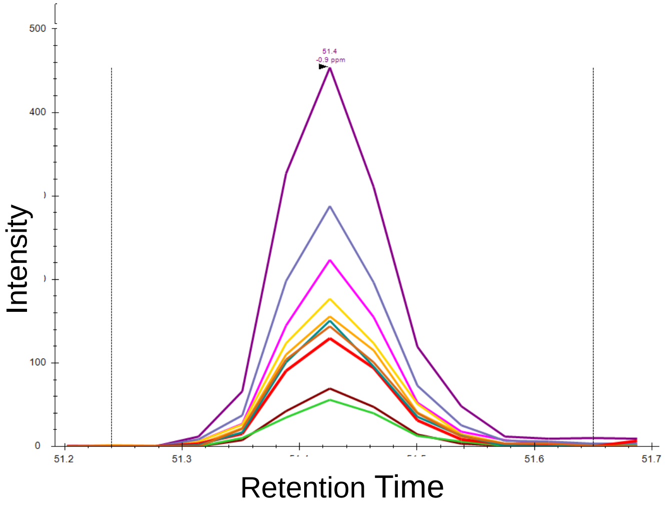


# Chromatography-based Quantification

- **SRM – Selected ion chromatograms**
- **PRM – Extracted ion chromatograms**
- **DIA – Extracted ion chromatograms**
- **DDA – Extracted ion chromatograms from MS1-only**

| Acquisition | Targeted | Survey |
|---|---|---|
| More Selective | PRM | DIA |
| Less Selective | SRM | DDA |





# Multiple Instrument Vendors

Agilent

SHIMADZU
Excellence in Science

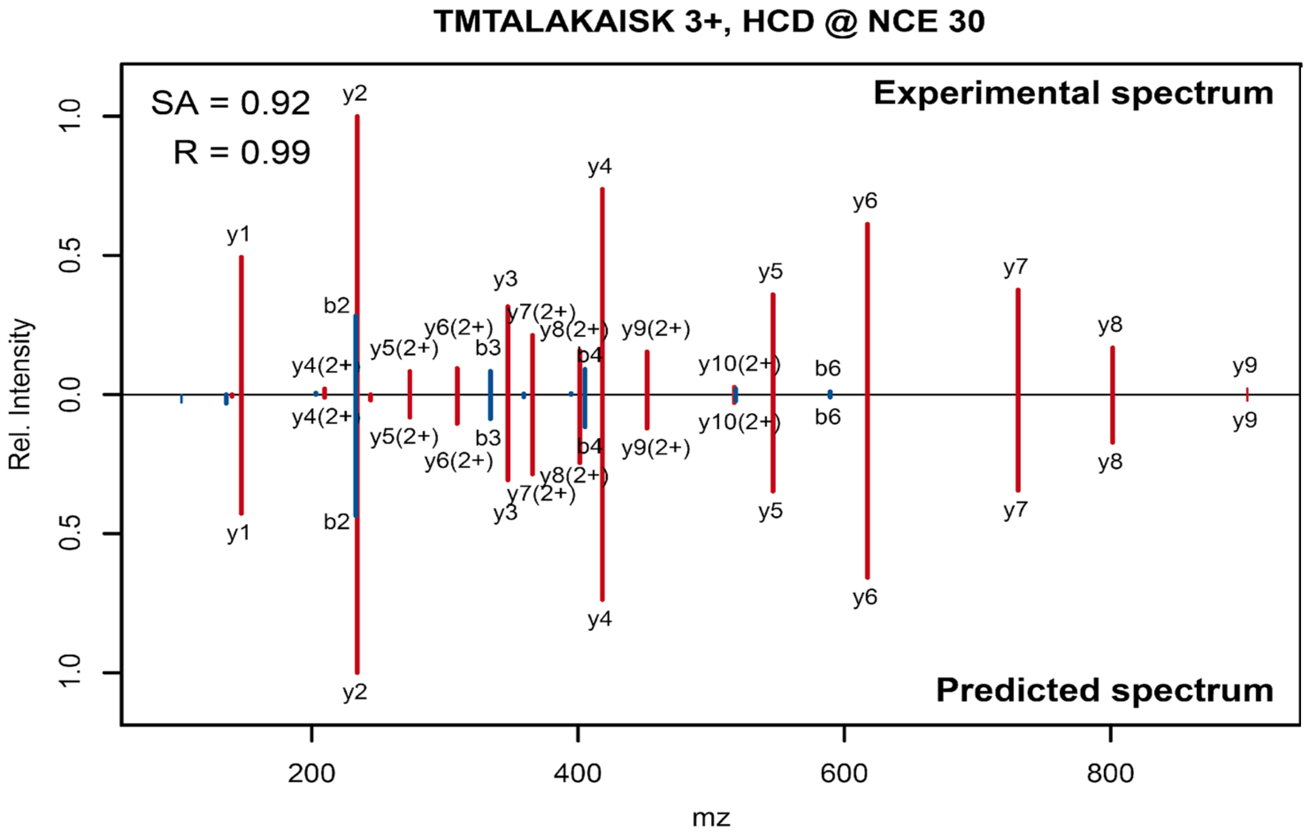
SCIEX

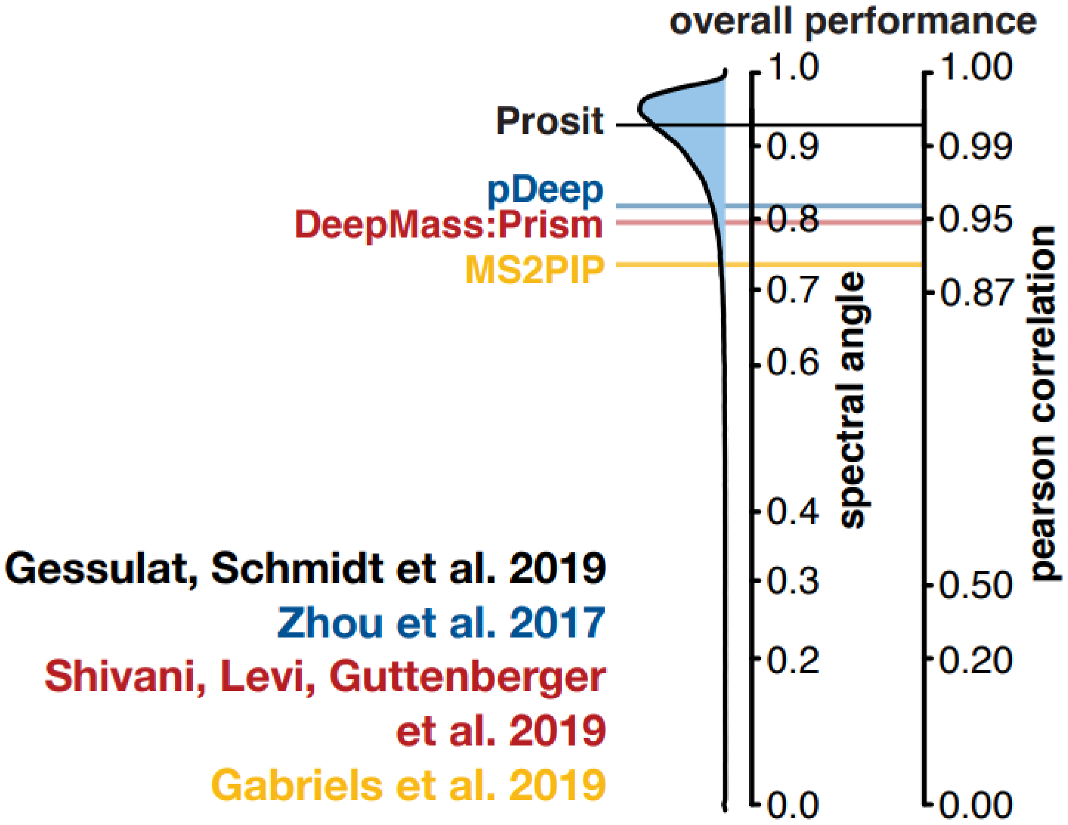
BRUKER

ThermoFisher
SCIENTIFIC

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

# Prosit: Prediction of (nearly) reference-like spectra



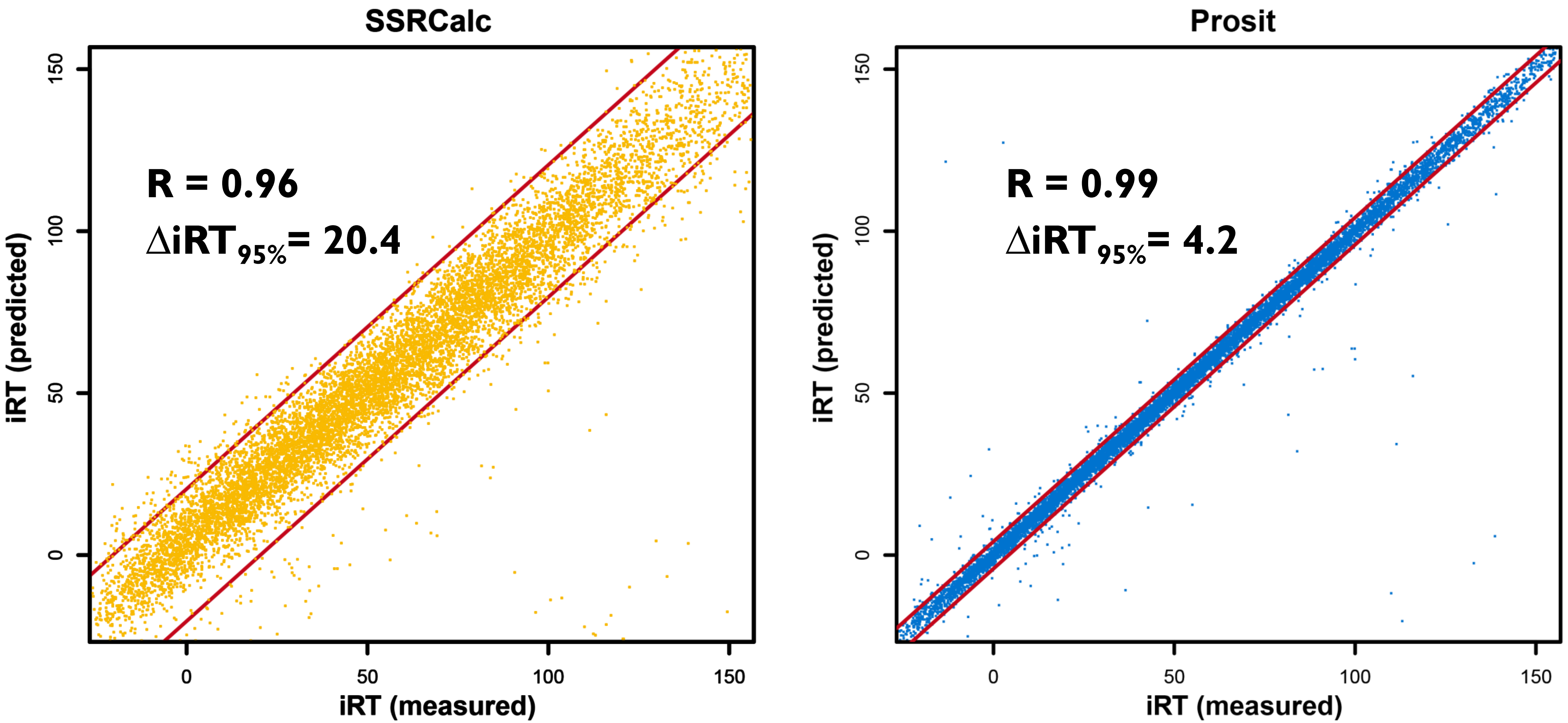


- Trained on Thermo Orbitrap MS/MS spectra

- Thanks to Tobias Schmidt

Gessulat, Schmidt et al. Nat Methods 2019

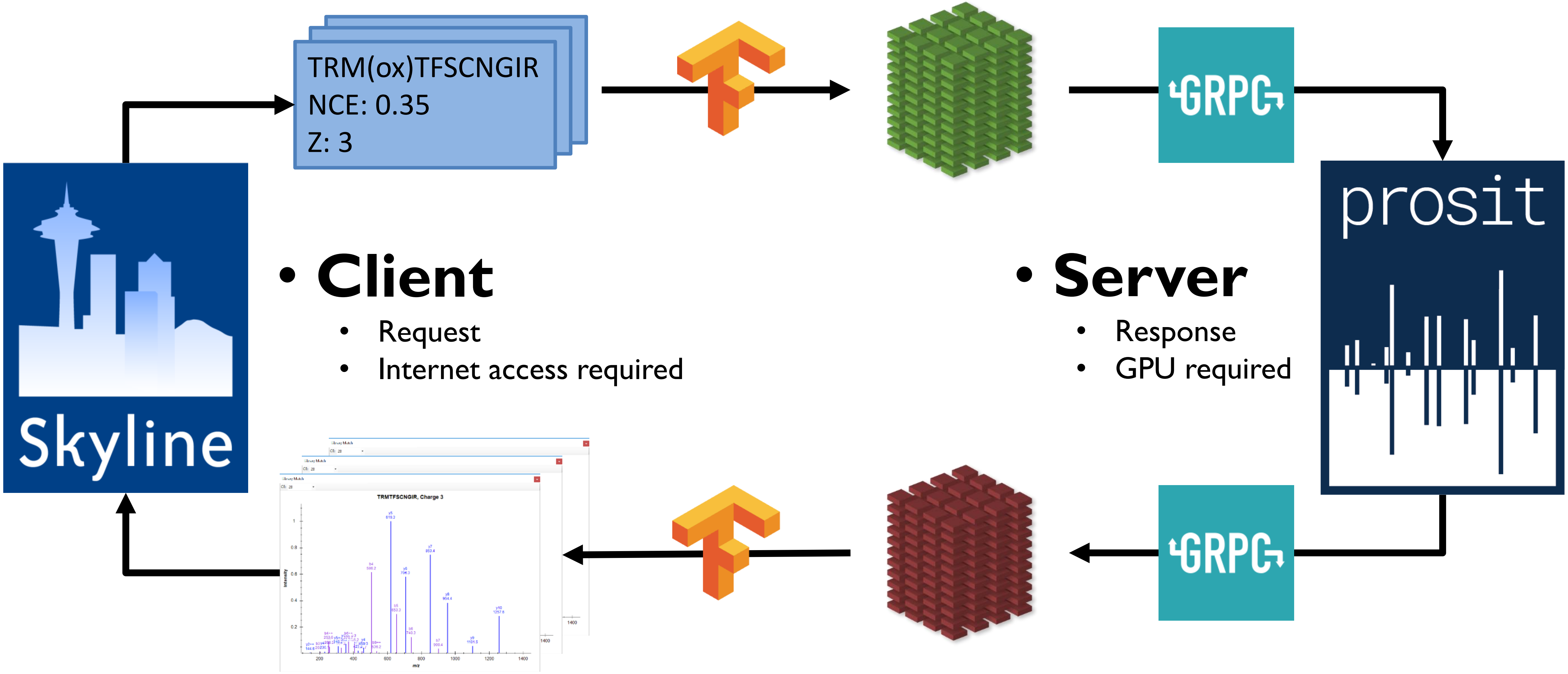
# Prosit: Prediction of (relatively) accurate iRT



- Thanks to Tobias Schmidt

Krokhin et al. Mol Cell Proteomics. 2014; Gessulat, Schmidt et al. Nat Methods 2019

# Skyline User Access to Prosit Predictions



- **Client**
  - Request
  - Internet access required

- **Server**
  - Response
  - GPU required

- Thanks to Tobias Rohde

# Testing Prosit on Proteomewide DIA

- Control – Original library
- Replace spectra with Prosit predictions
- Replace iRT with Prosit predictions
- Replace both with Prosit predictions

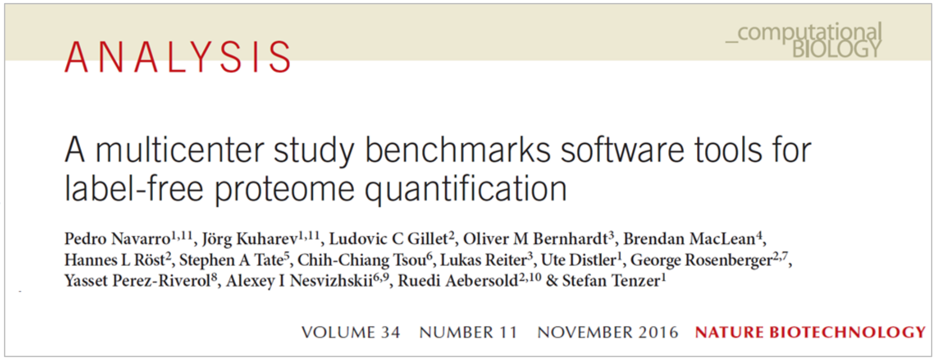
- Does not replace choosing what to target
  - Detectable peptides
  - Optimal precursor charge
- Did not require
  - Ion mobility filtering (dia-PASEF, HDMSe) - MP 253

---

- Wilhelm, M. - MOD am 10:10 Extending Prosit to the prediction of proteotypicity, precursor ion charge and CCS

# LFQbench Study

- Testing SCIEX TripleTOF – 6600
- Trained on Thermo MS/MS spectra

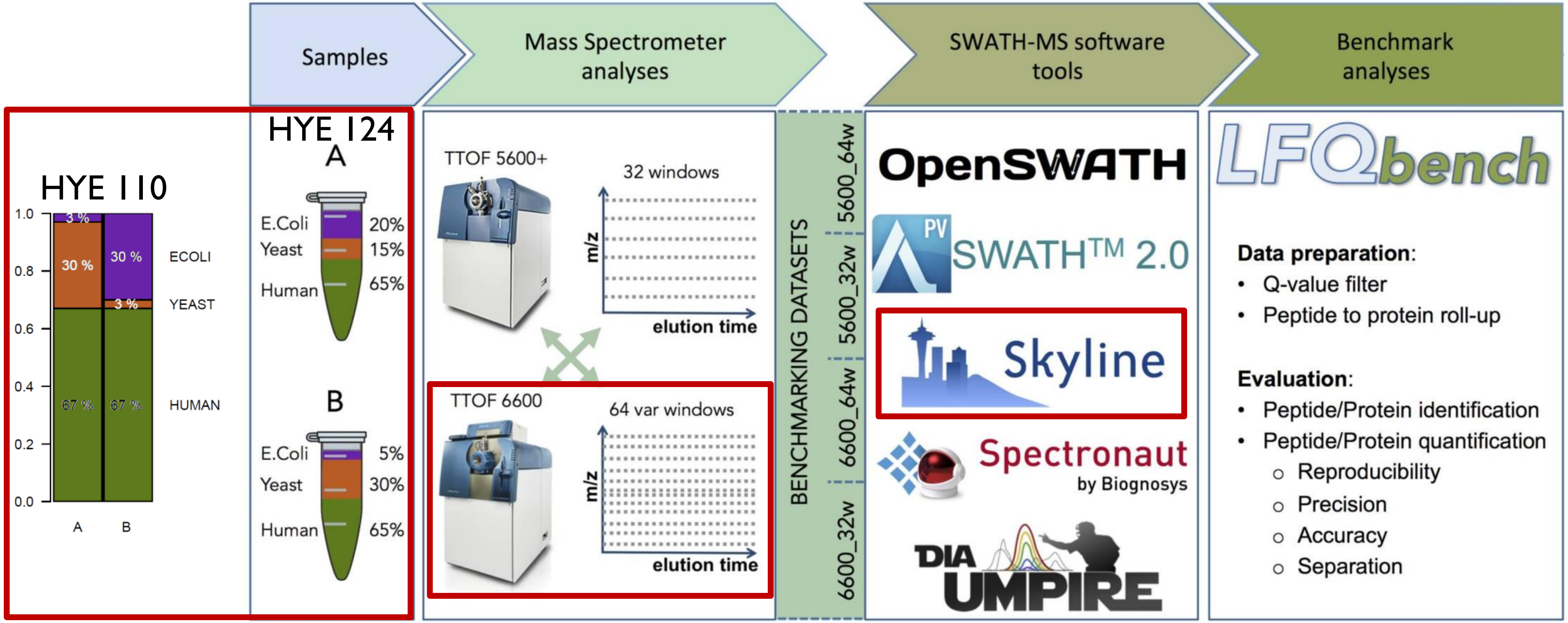


ANALYSIS
computational BIOLOGY

A multicenter study benchmarks software tools for label-free proteome quantification

Pedro Navarro[1,11], Jörg Kuharev[1,11], Ludovic C Gillet[2], Oliver M Bernhardt[3], Brendan MacLean[4], Hannes L Röst[2], Stephen A Tate[5], Chih-Chiang Tsou[6], Lukas Reiter[3], Ute Distler[1], George Rosenberger[2,7], Yasset Perez-Riverol[8], Alexey I Nesvizhskii[6,9], Ruedi Aebersold[2,10] & Stefan Tenzer[1]

VOLUME 34 NUMBER 11 NOVEMBER 2016 NATURE BIOTECHNOLOGY



- Thanks to Ben Collins

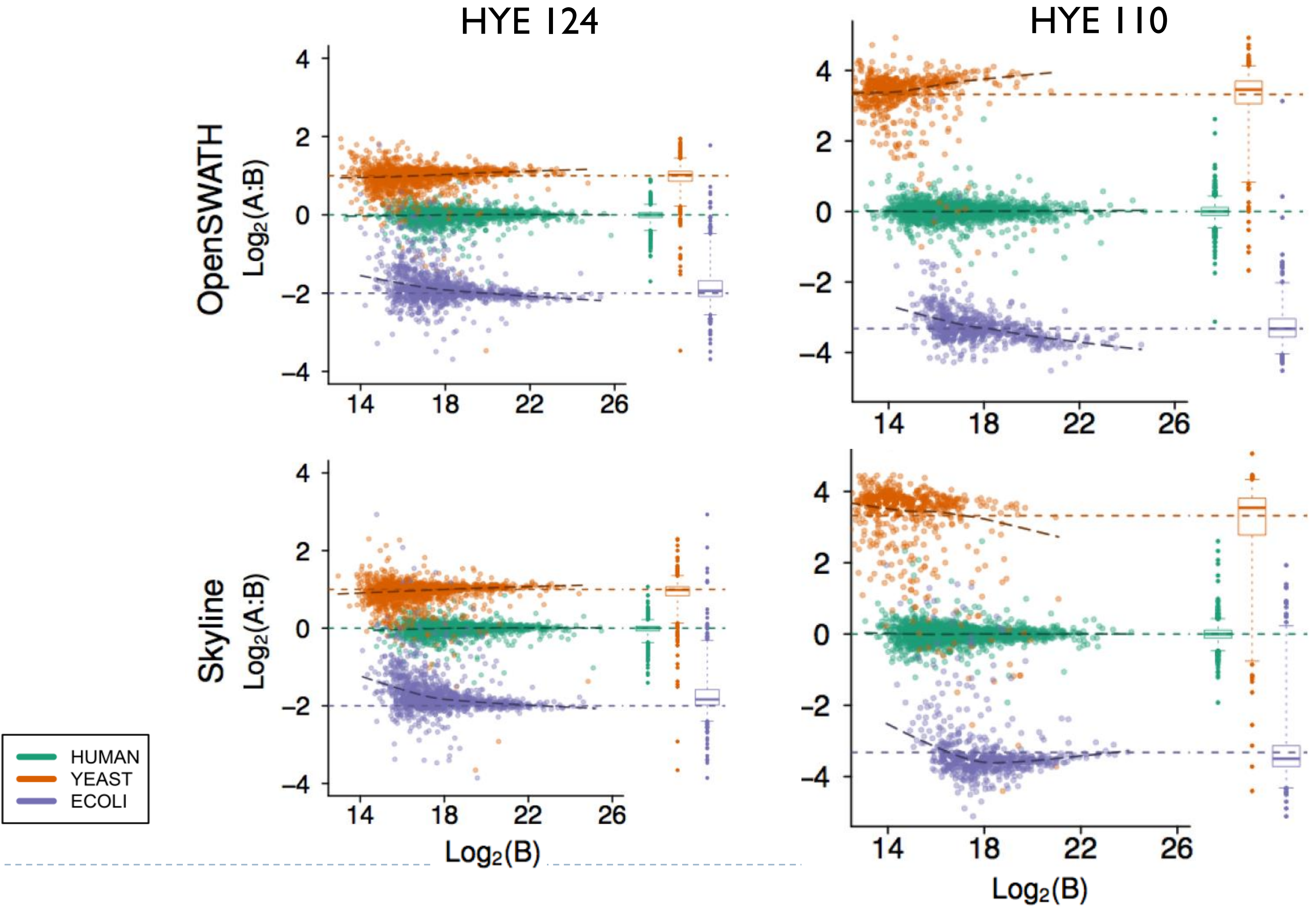
https://skyline.ms/webinar18.url

# LFQbench Study Output

Some questions:

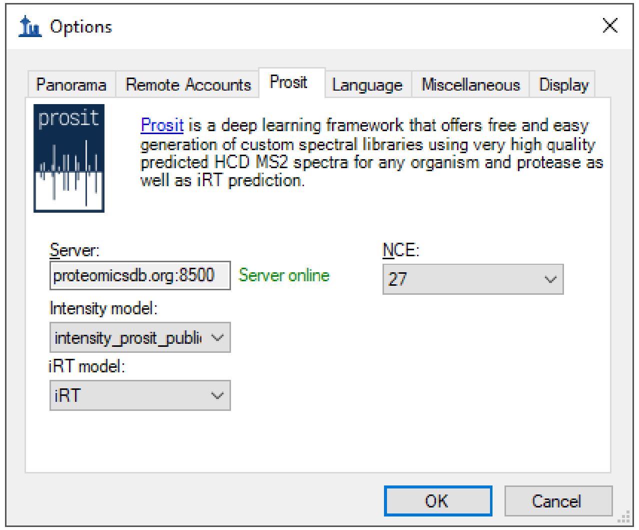
1. Are the quantitative ratios as expected?

2. Are there a lot of data points out of (species) position?



Thanks to Ben Collins

# Skyline-Prosit Integration



Prosit is a deep learning framework that offers free and easy generation of custom spectral libraries using very high quality predicted HCD MS2 spectra for any organism and protease as well as iRT prediction.

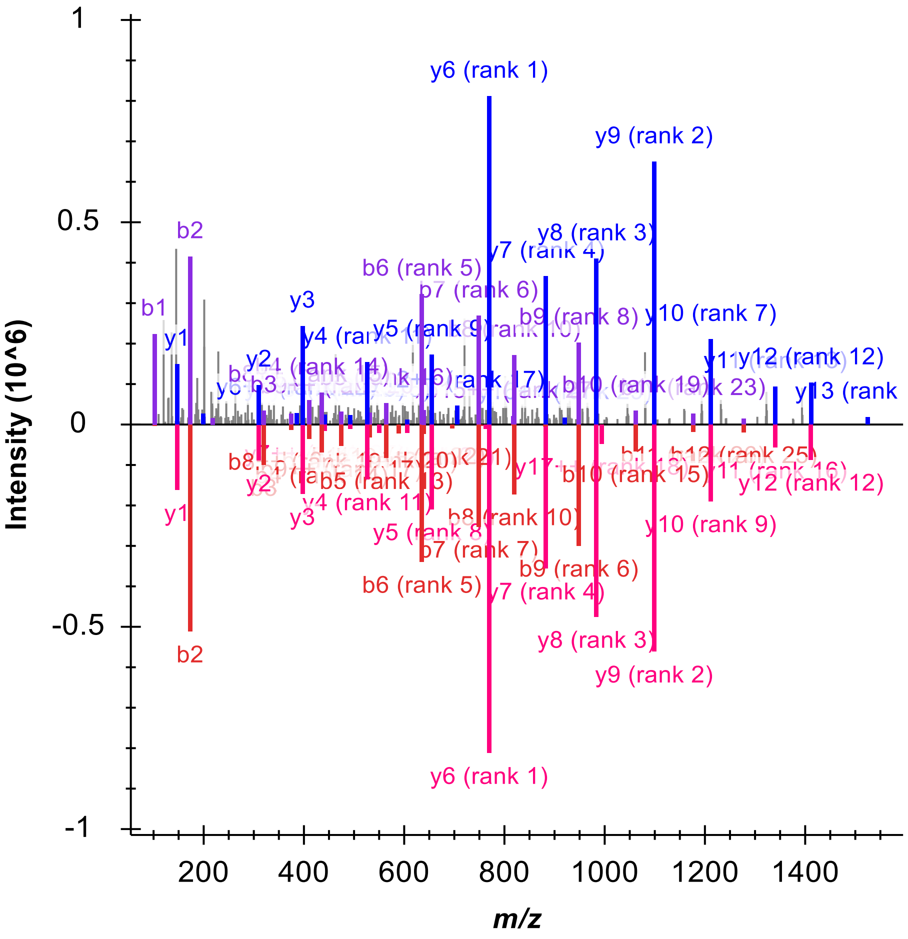
Server: proteomicsdb.org:8500 Server online

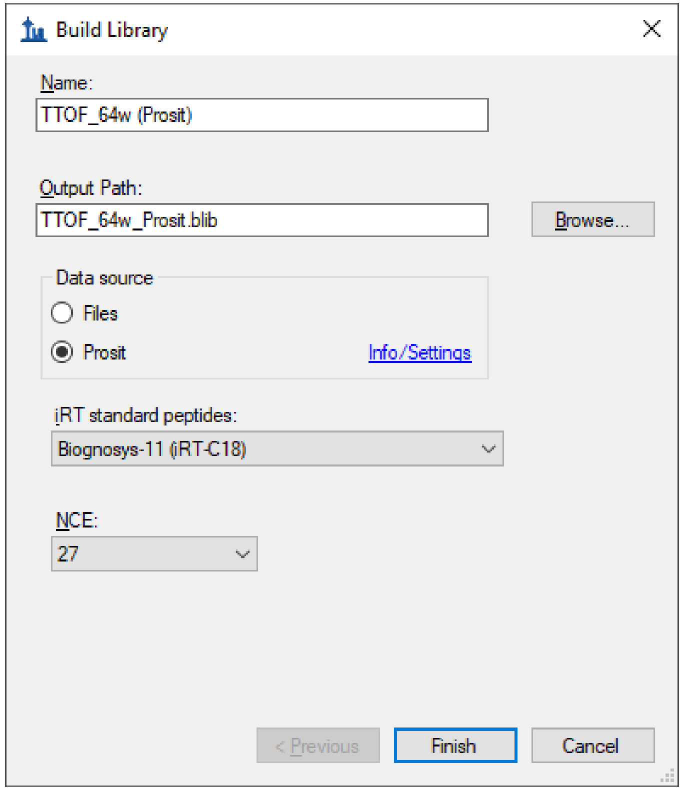
NCE: 27

Intensity model: intensity_prosit_publi

iRT model: iRT

**Tricolor_SuperLibrary (rebuilt) vs. Prosit**
**TAFDEAIAELDTLNEESYK, Charge 3**
**dotp: 0.8014**





Name: TTOF_64w (Prosit)

Output Path: TTOF_64w_Prosit.blib

Data source
○ Files
● Prosit

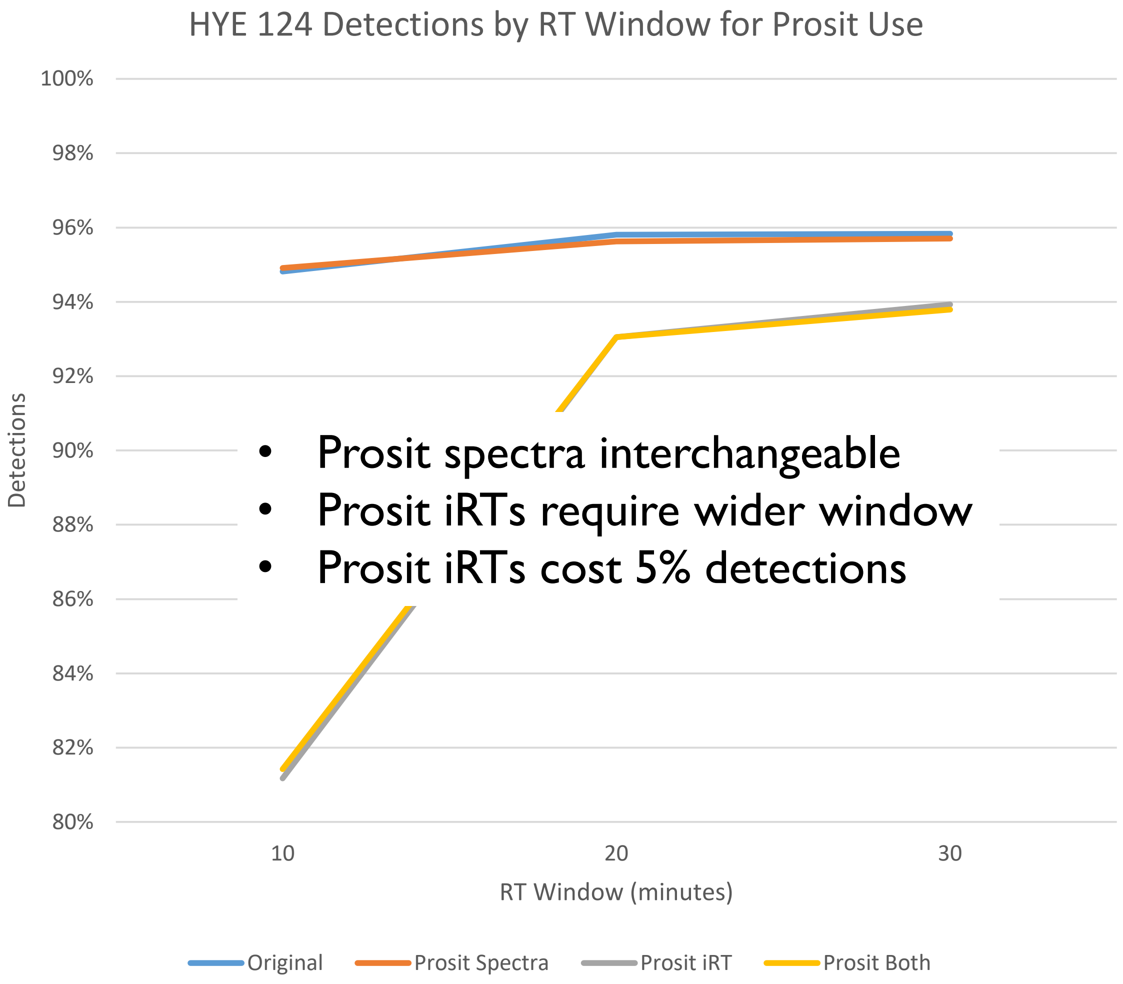
iRT standard peptides: Biognosys-11 (iRT-C18)

NCE: 27

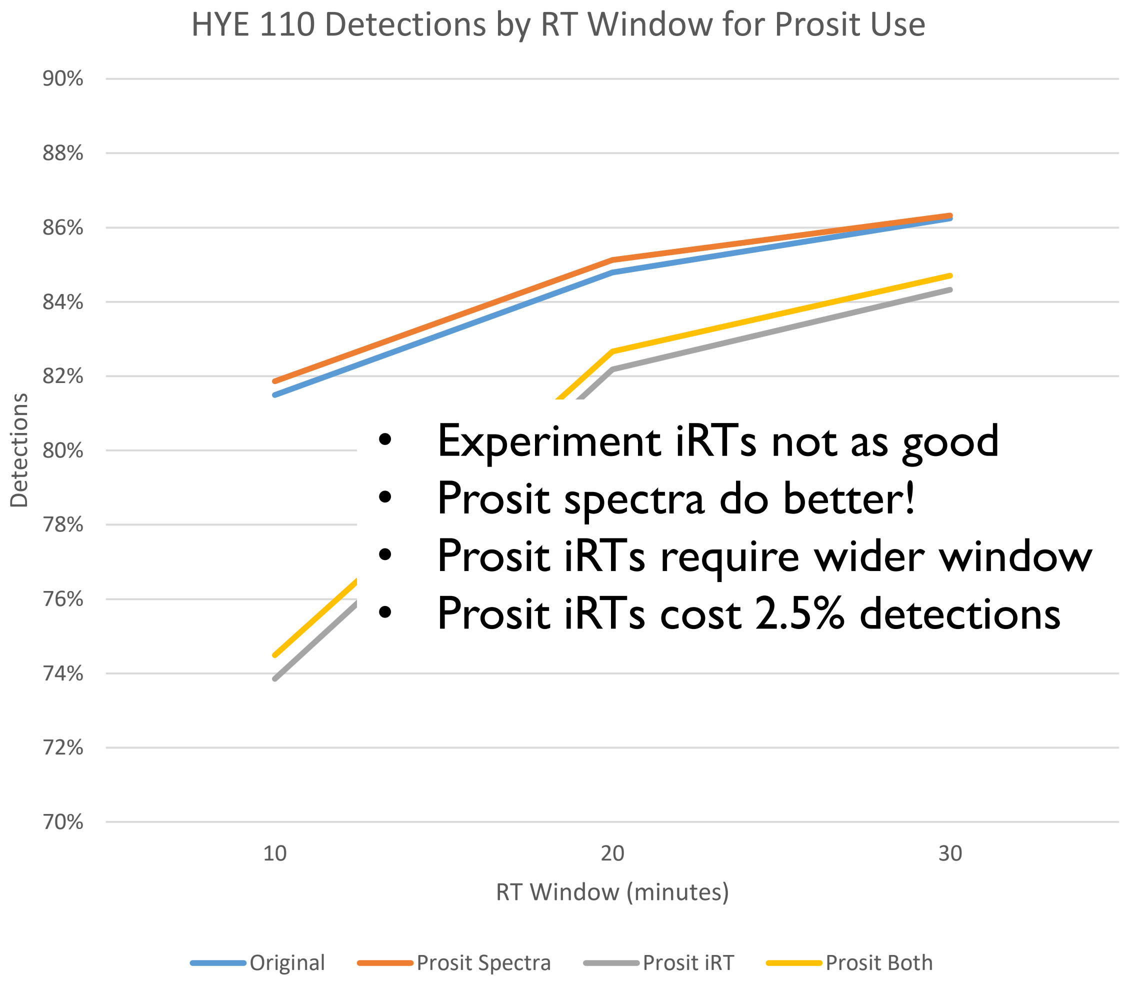
May Institute - https://youtu.be/xm7Niu3lyZA?t=7116

# Prosit Impact on LFQbench Study

44,223 peptides targeted



- Prosit spectra interchangeable
- Prosit iRTs require wider window
- Prosit iRTs cost 5% detections

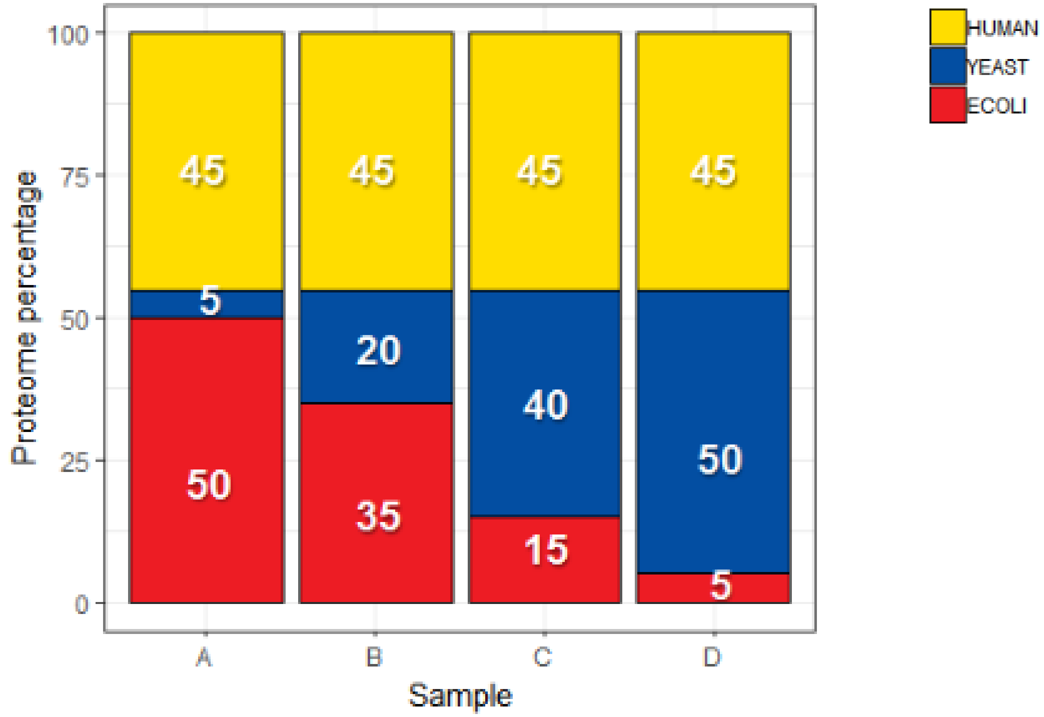


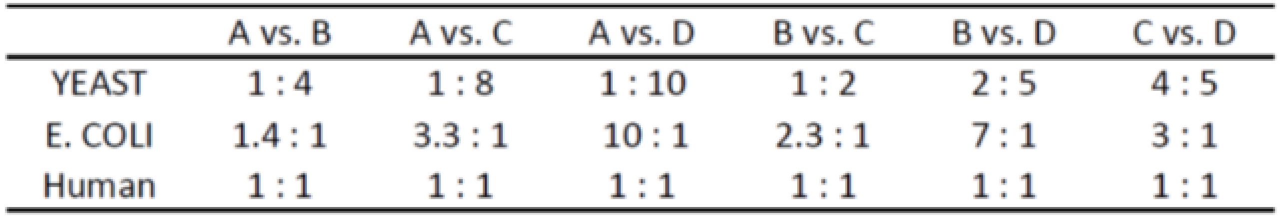
- Experiment iRTs not as good
- Prosit spectra do better!
- Prosit iRTs require wider window
- Prosit iRTs cost 2.5% detections

# Avant-Garde Extended Benchmark

- 4-samples
- 3-organisms
- 6-comparisons

- Thermo Q Exactive
- Narrow window DIA library
  - 12 x gas phase fractions, 2 m/z windows
  - 57,439 peptides
  - 18,000 chosen randomly
  - 16,117 targeted





| | A vs. B | A vs. C | A vs. D | B vs. C | B vs. D | C vs. D |
|---|---|---|---|---|---|---|
| YEAST | 1 : 4 | 1 : 8 | 1 : 10 | 1 : 2 | 2 : 5 | 4 : 5 |
| E. COLI | 1.4 : 1 | 3.3 : 1 | 10 : 1 | 2.3 : 1 | 7 : 1 | 3 : 1 |
| Human | 1 : 1 | 1 : 1 | 1 : 1 | 1 : 1 | 1 : 1 | 1 : 1 |

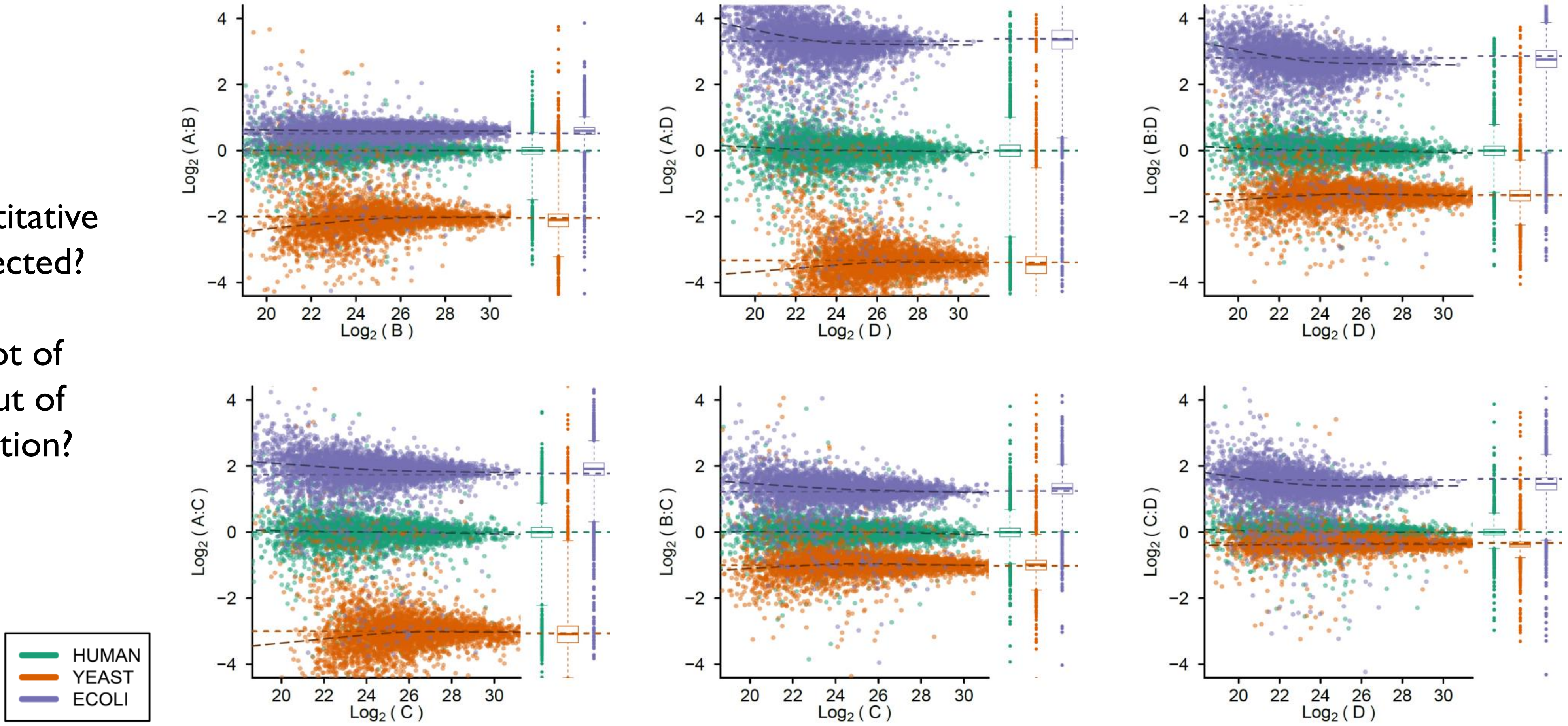
- Vaca, S. Nature Methods, accepted

# Skyline mProphet LFQbench Performance

Some questions:

1. Are the quantitative ratios as expected?

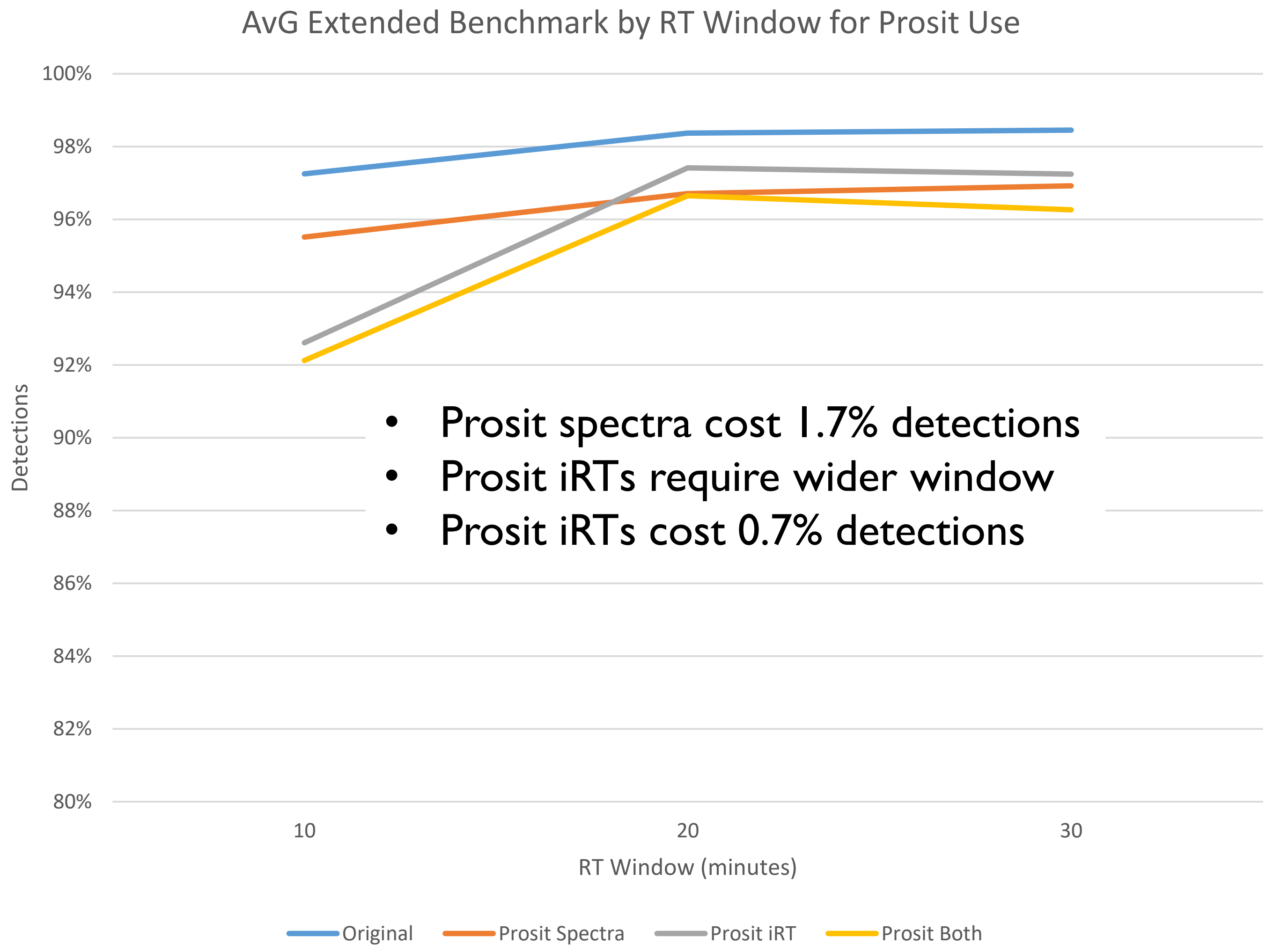
2. Are there a lot of data points out of (species) position?



Vaca, S. Nature Methods, accepted

# Prosit Impact on Extended Benchmark

16,117 peptides targeted



- Prosit spectra cost 1.7% detections
- Prosit iRTs require wider window
- Prosit iRTs cost 0.7% detections

Searle, B. et al. Nat. Communications 2018 & 2020

# Acknowledgments

- Tobias Rohde

- Tobias Schmidt

- Nick Shulman

- Johannes Rank

- Mathias Wilhelm

- Bernhard Kuster

- Michael MacCoss