

# DIA/SWATH -- primer and quantitative benchmarking

**Skyline Webinar #18: DIA/SWATH Data Analysis in Skyline Revisited**

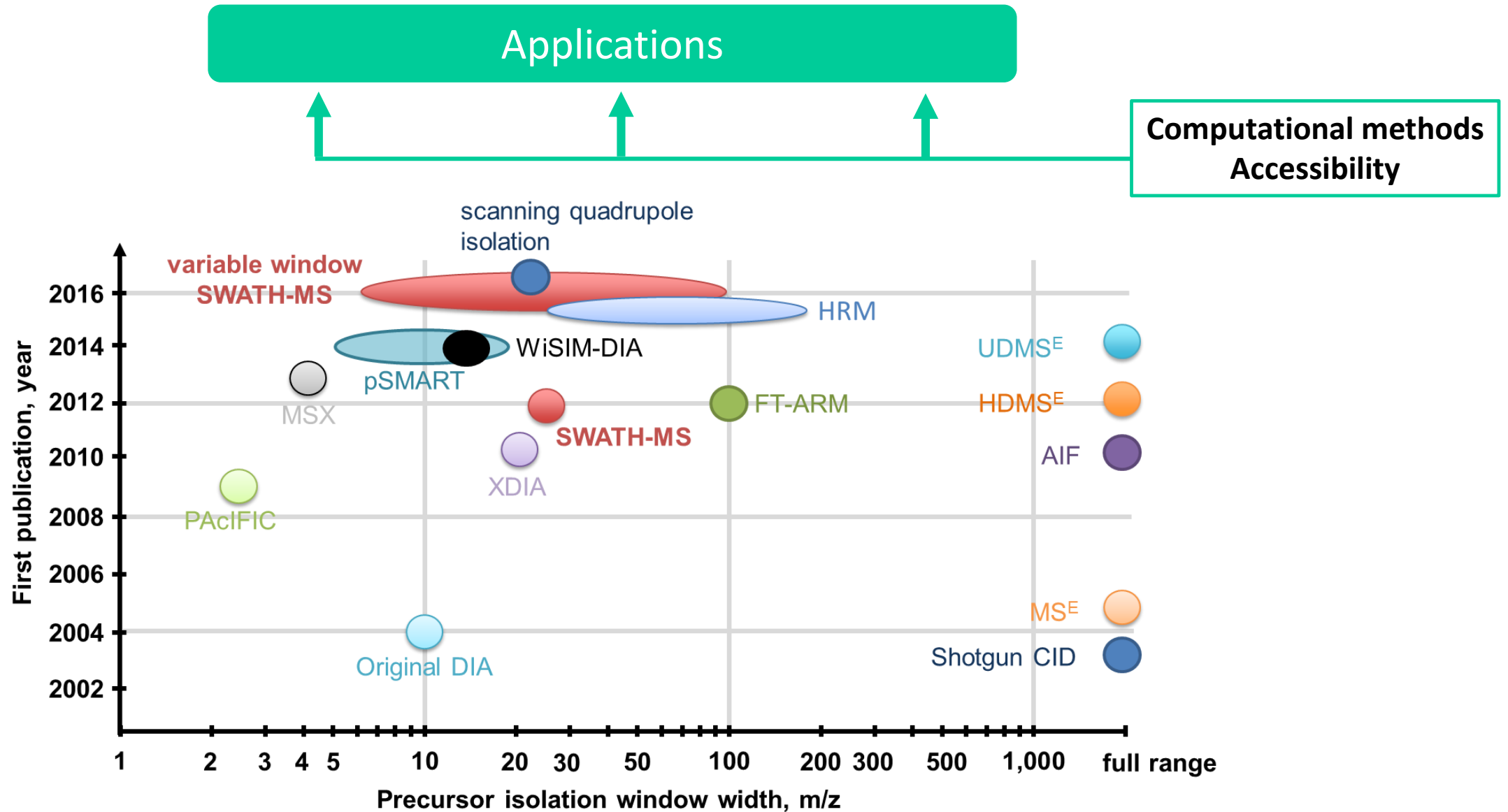
**Ben Collins**

Queen's University Belfast | School of Biological Sciences

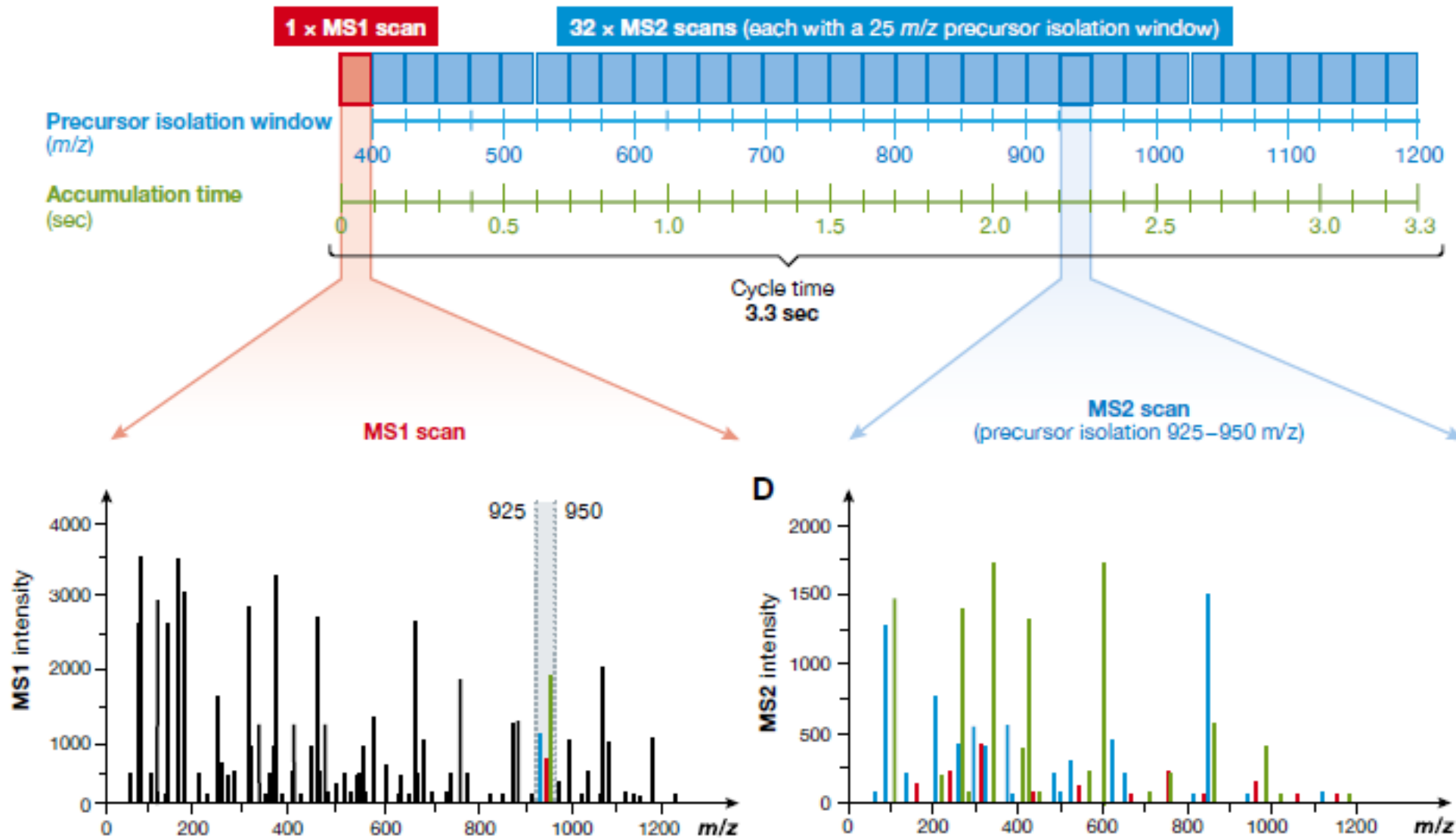
# Outline

- Primer on DIA/SWATH
- Peptide centric or targeted data analysis
- Data set for the tutorial -- and where it came from
  - Inspired by LFQBench (Navarro et al. Nature Biotech 2016)

# A brief history of data independent acquisition



# A typical DIA/SWATH acquisition method



# Tutorial paper – DIA/SWATH

*Review*



molecular  
systems  
biology

## Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial

Christina Ludwig<sup>1,\*†</sup> , Ludovic Gillet<sup>2,†</sup>, George Rosenberger<sup>2,3</sup> , Sabine Amon<sup>2</sup>, Ben C Collins<sup>2</sup>  &  
Ruedi Aebersold<sup>2,4</sup>

# Outline

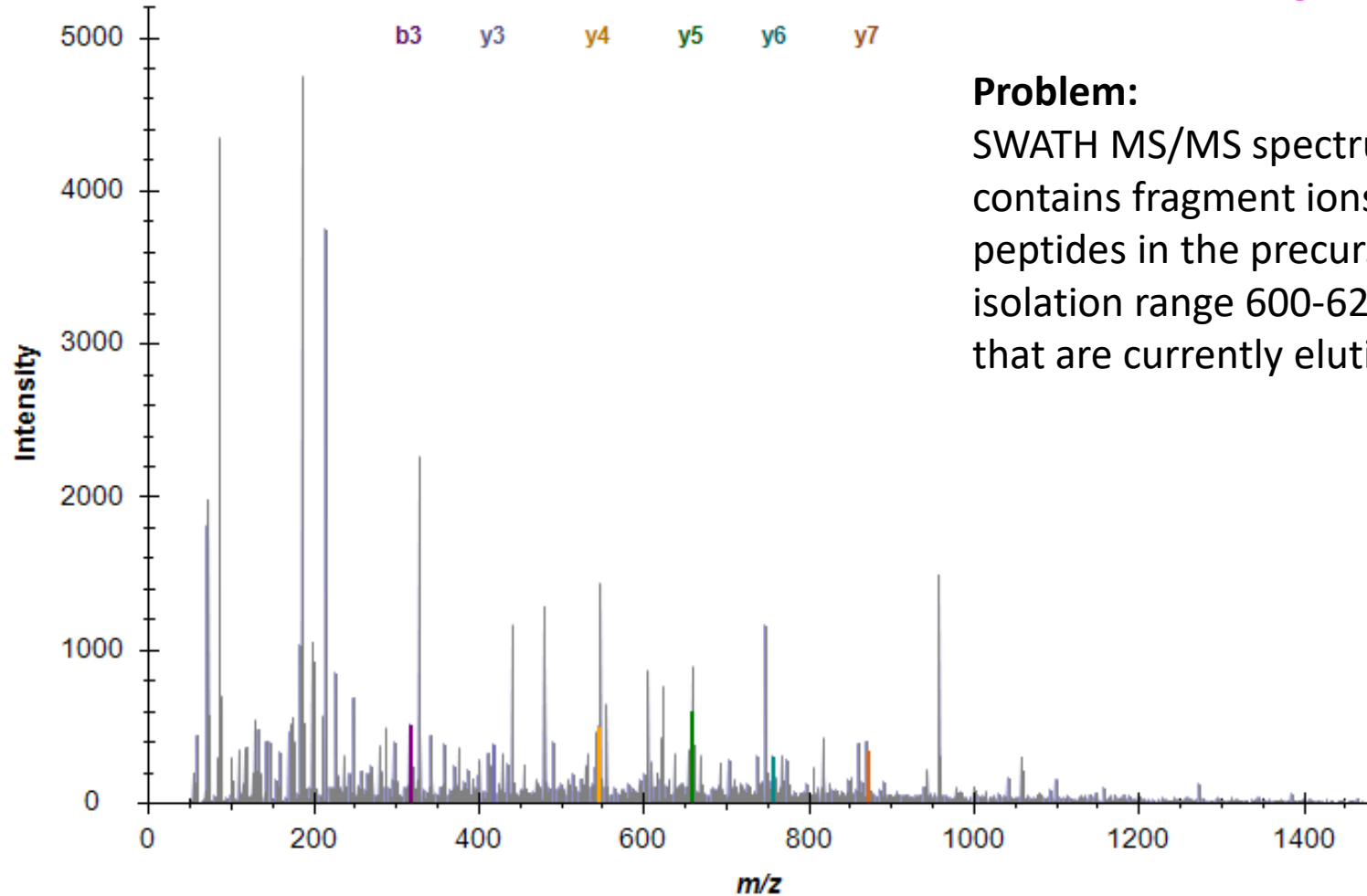
- Primer on DIA/SWATH

- Peptide centric or targeted data analysis

- Data set for the tutorial -- and where it came from
  - Inspired by LFQBench (Navarro et al. Nature Biotech 2019)

# Why Targeted analysis?

fragment ion spectrum 600-625 m/z

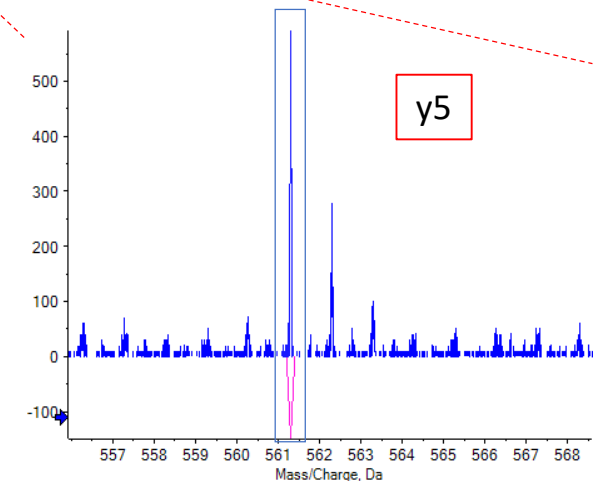
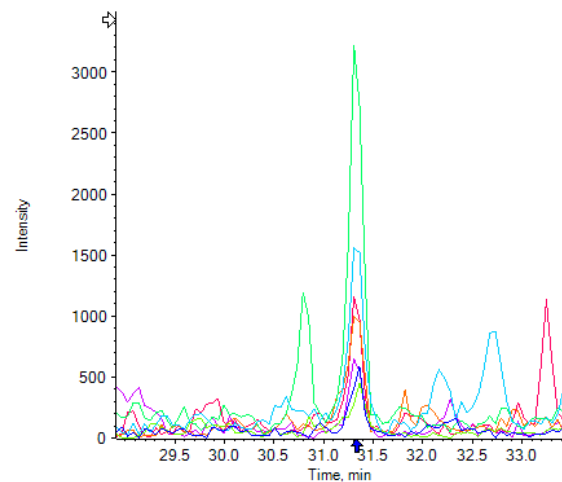
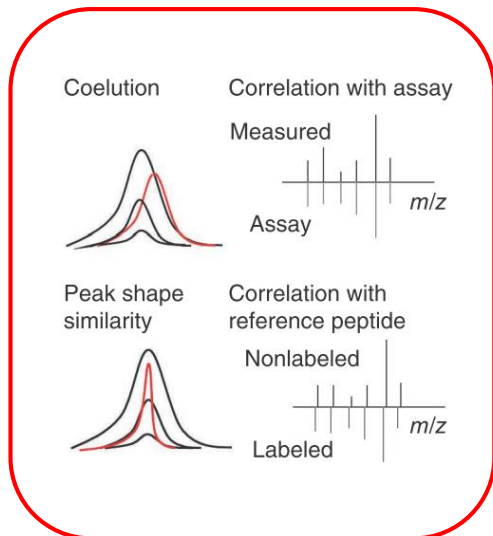
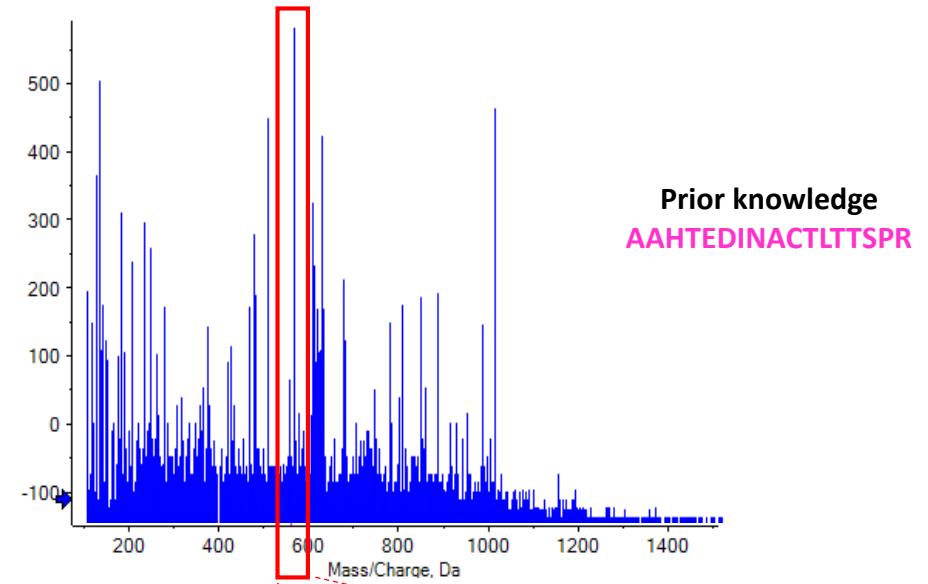
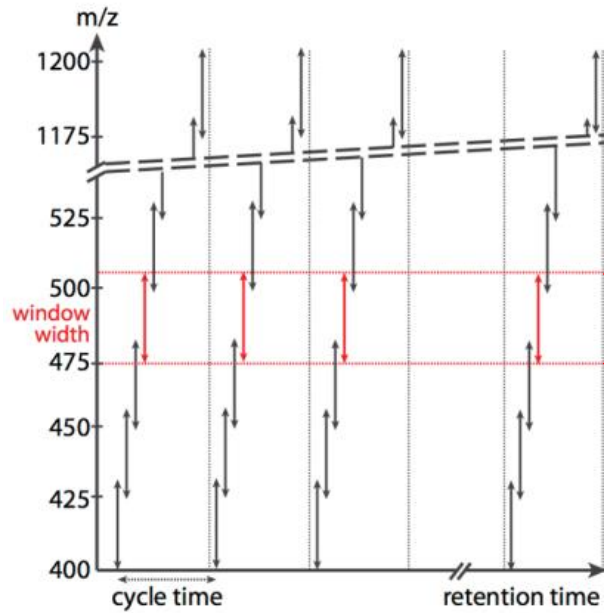


AAHTEDINACTLTTSPR

## Problem:

SWATH MS/MS spectrum contains fragment ions from any peptides in the precursor isolation range 600-625 m/z that are currently eluting

# Peptide centric analysis (or targeted analysis)

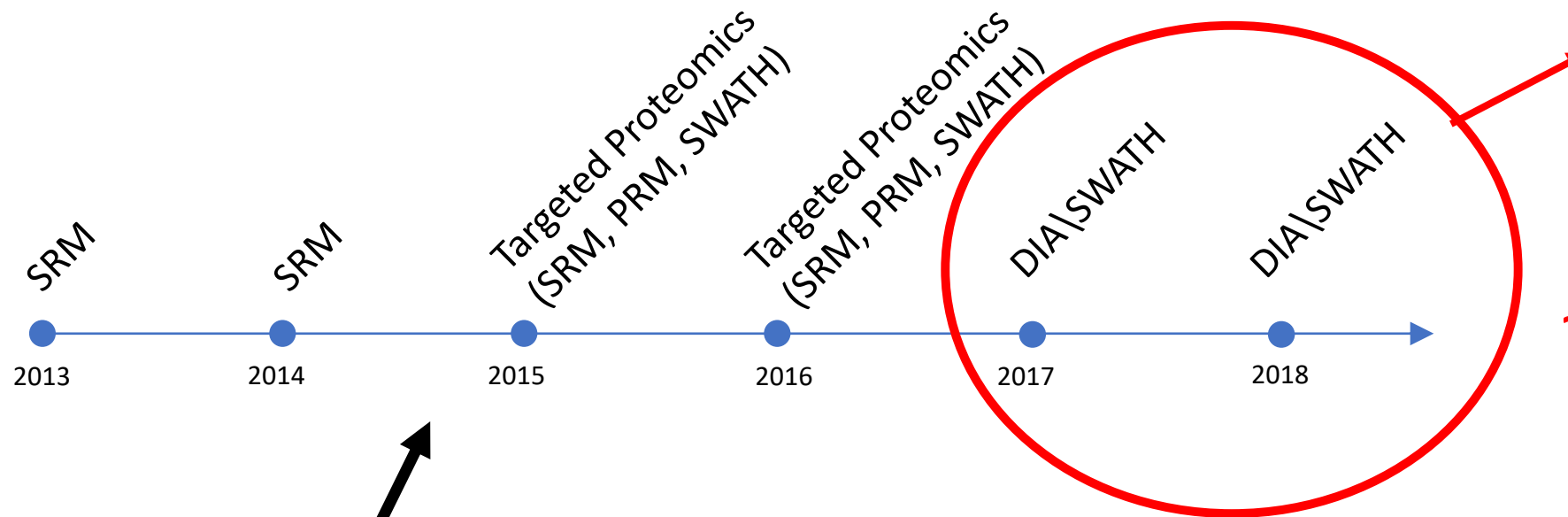




# Outline

- Primer on DIA/SWATH
- Peptide centric or targeted data analysis
- Data set for the tutorial -- and where it came from
  - Inspired by LFQBench (Navarro et al. Nature Biotech 2019)

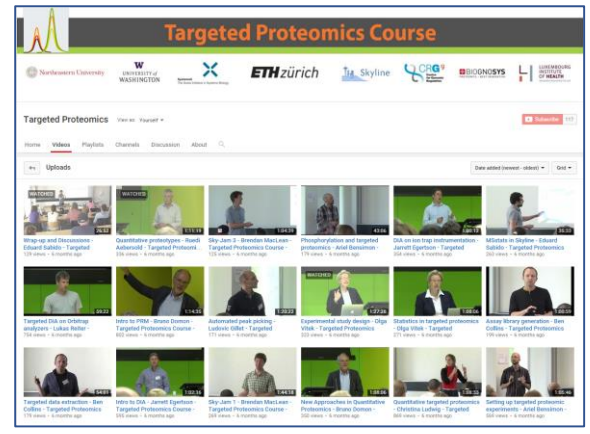
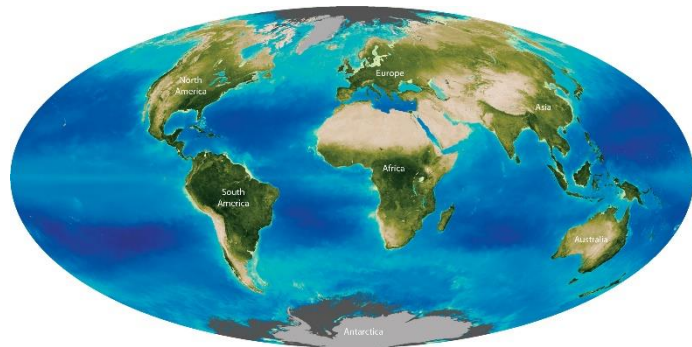
# ETH Course Background



Data/Basis for today's tutorial



Strong relationship with lots of courses worldwide



Youtube 'Targeted Proteomics' Channel  
80,000+ views

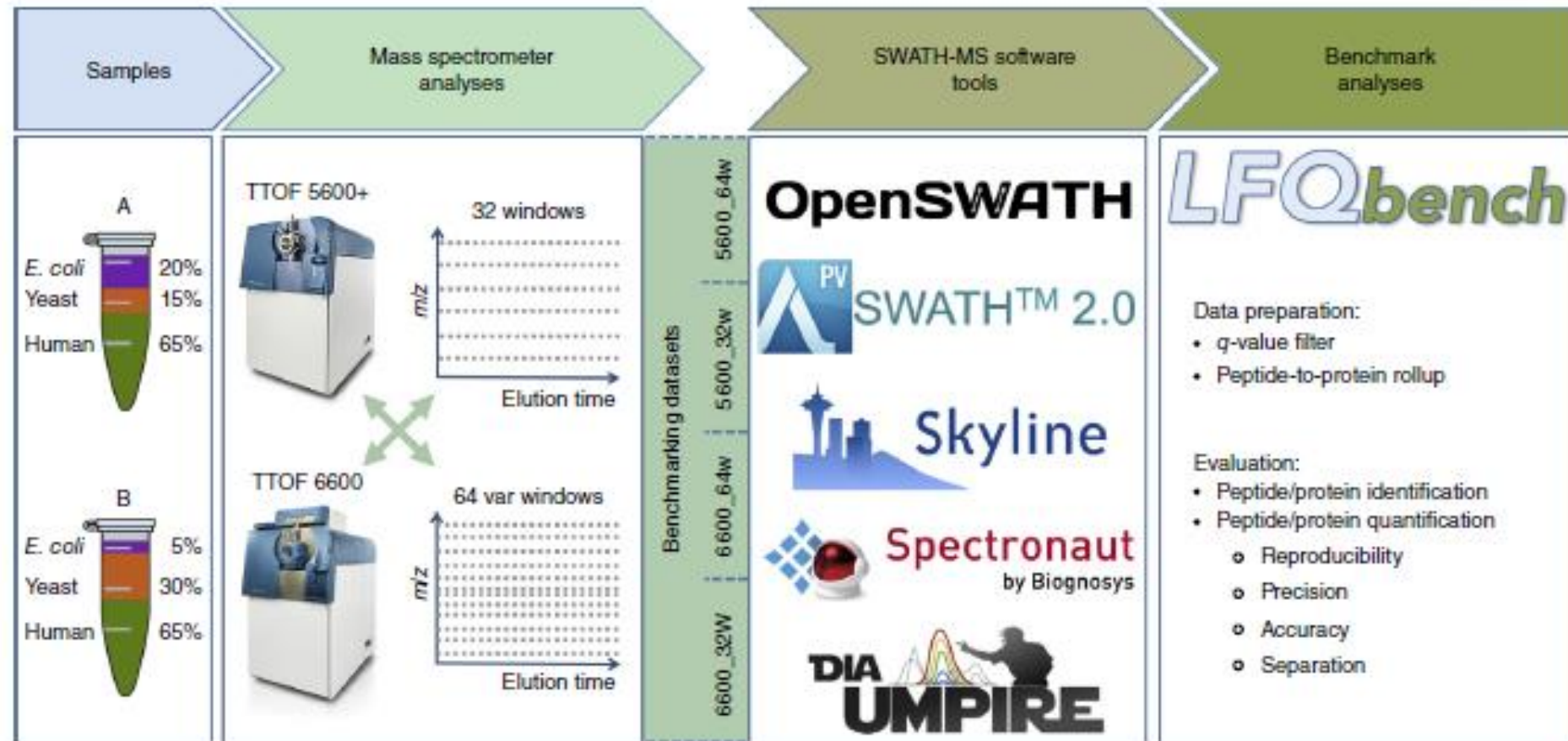
# LFQBench study

## ANALYSIS

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

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

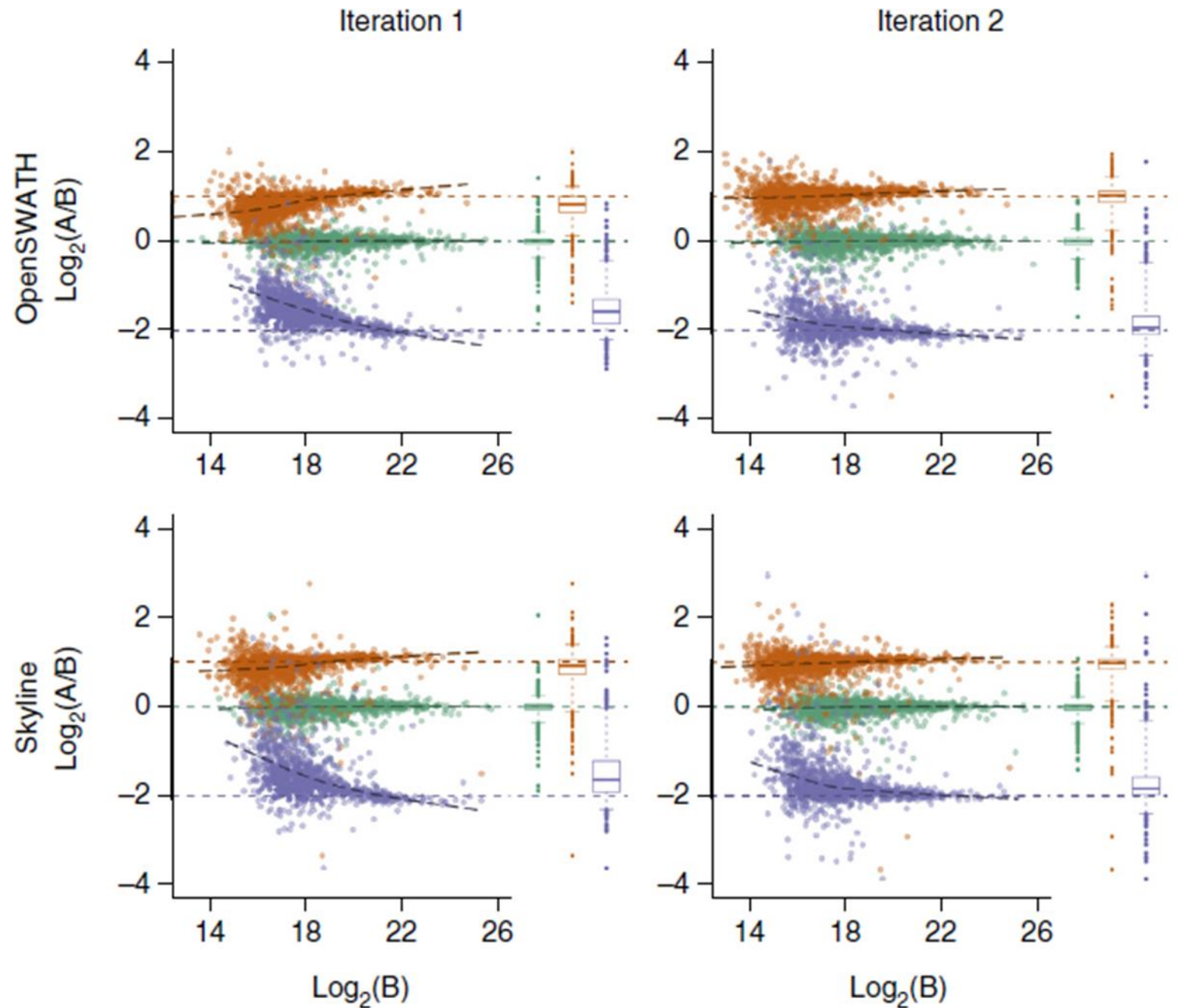
VOLUME 34 NUMBER 11 NOVEMBER 2016 NATURE BIOTECHNOLOGY



# LFQBench study output

Some questions:

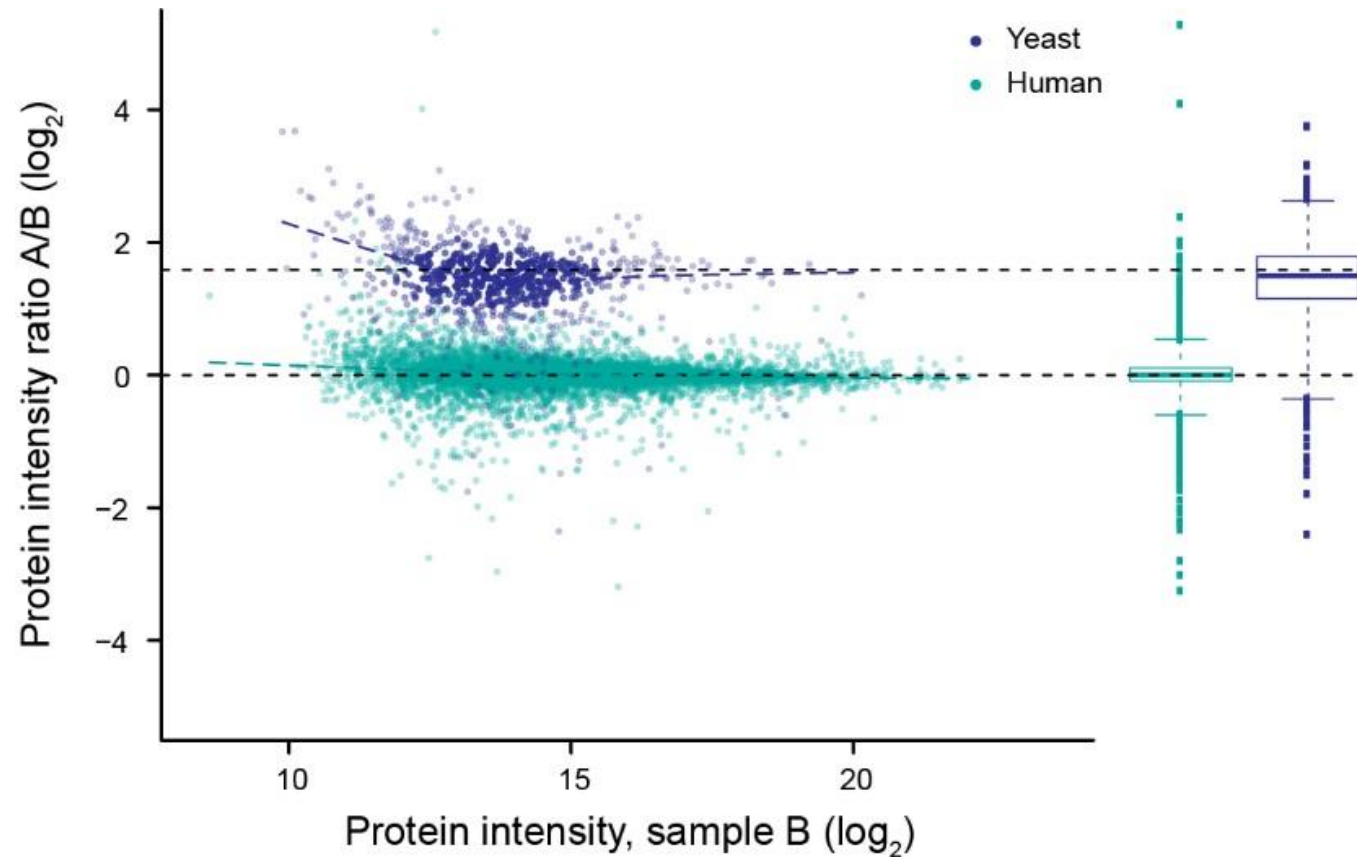
1. Are the quantitative ratios as expected?
2. Are there a lot of data points out of (species) position?



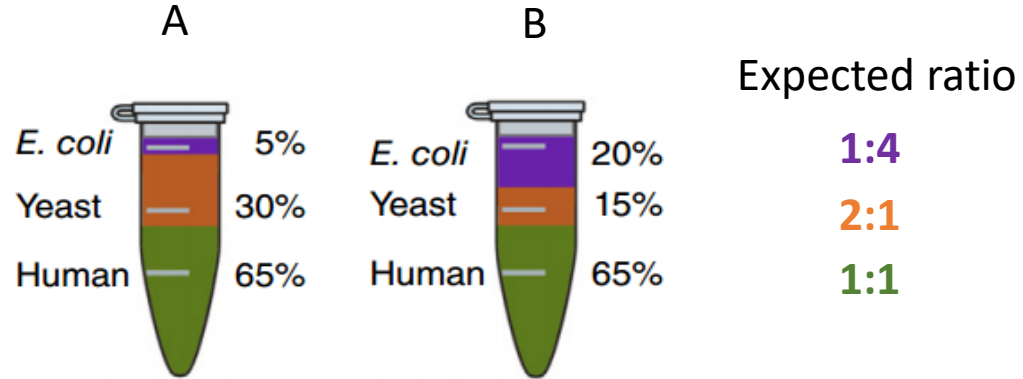
# LFQBench style multispecies mixture proteome is now used frequently for benchmarking of new methods

diaPASEF

Meier, F et al. *BioRxiv* (2020)



# Data set for the webinar demo

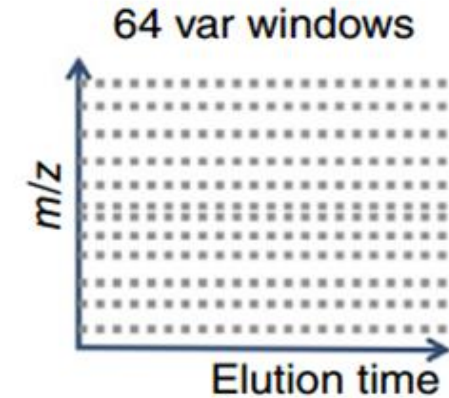


3 x DIA technical replicates  
1 hour gradient

+ 2 DDA files  
(1 of each condition)

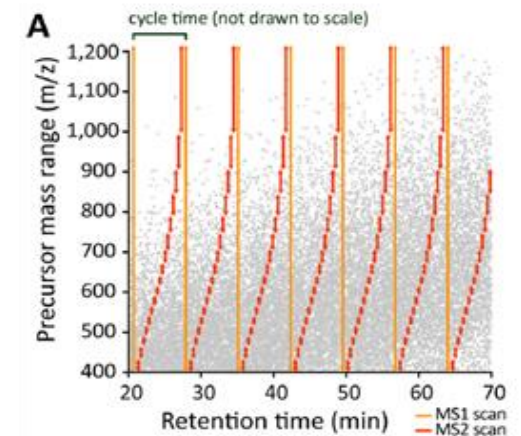
(A/B labelling is opposite from Navarro)

6600 TripleTOF



or

Q-Exactive Plus



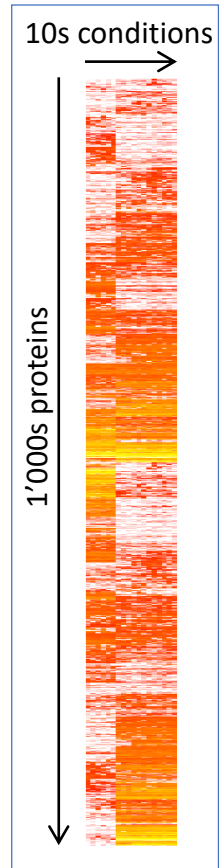


Backup

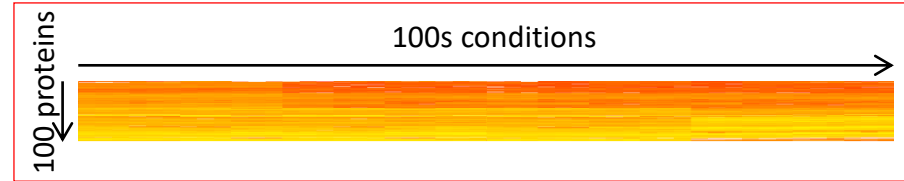


# Motivation for DIA development

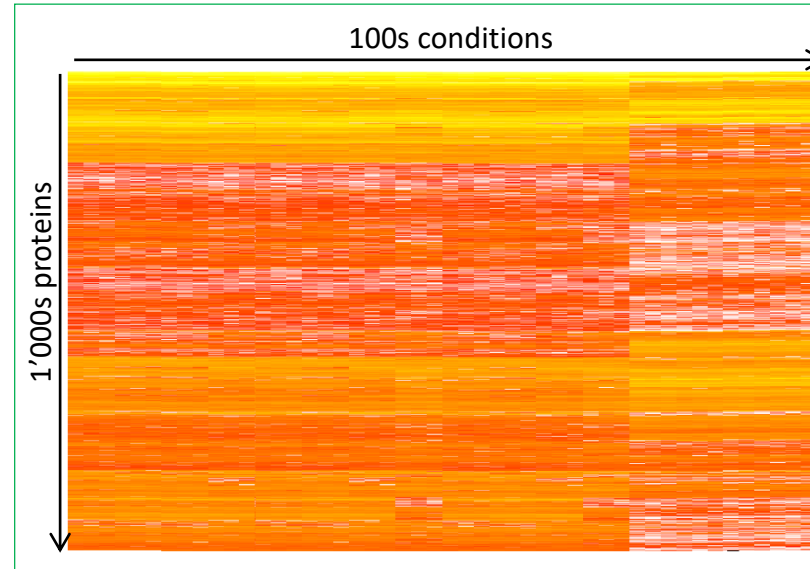
DDA / Discovery proteomics



Targeted (acquisition) proteomics



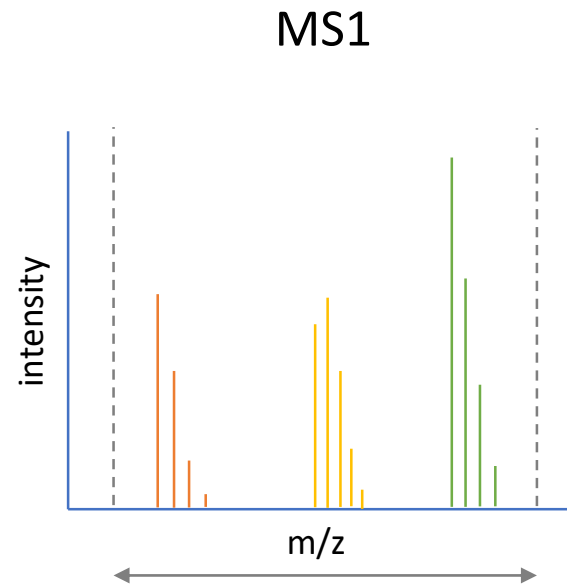
Data Independent Acquisition + "peptide-centric" Targeted Data Extraction



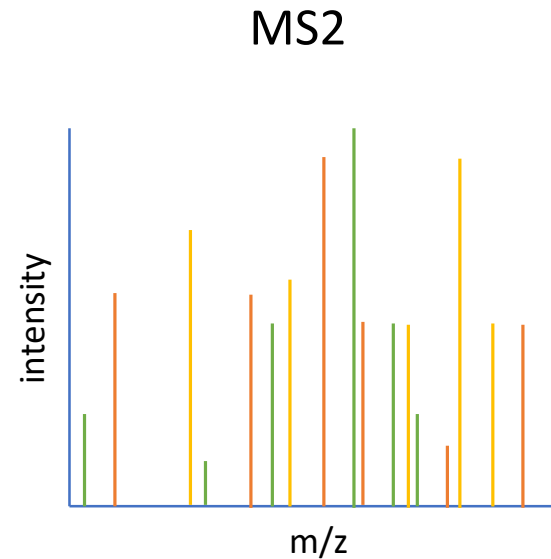
# What is data independent acquisition?

- DIA term introduced by Venables J, *et al.* Nat Methods (2004)
- Explicitly not data dependent (no stochastic precursor selection)
- Features of DIA:
  - Fixed duty cycle - deterministic (like SRM/PRM but unlike DDA)
  - Untargeted/unbiased data acquisition (like DDA but unlike SRM/PRM)
  - Comprehensive sampling of the precursor space with MS2 spectra
  - (Usually) Wide precursor isolation windows
  - (Usually) complex mixture MS2 spectra

# Problem statement: DIA/SWATH produces highly complex mixture MS2 spectra

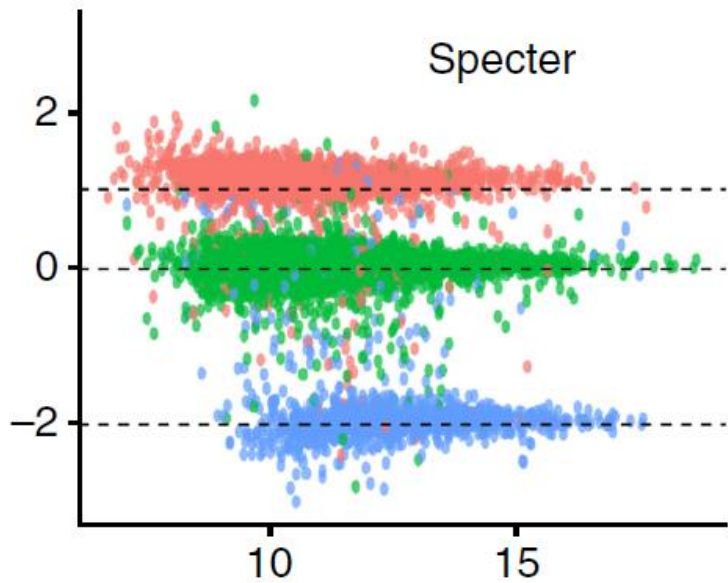


Precursor isolation  
→  
Fragmentation

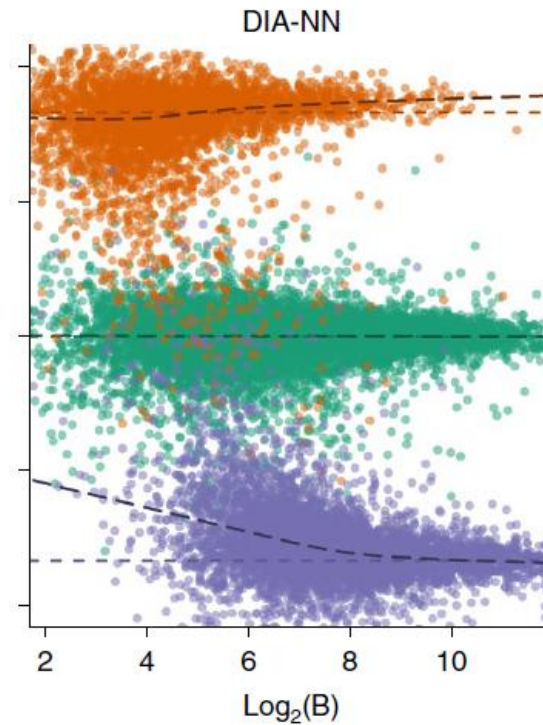


Not directly interpretable by standard  
database searching approaches!

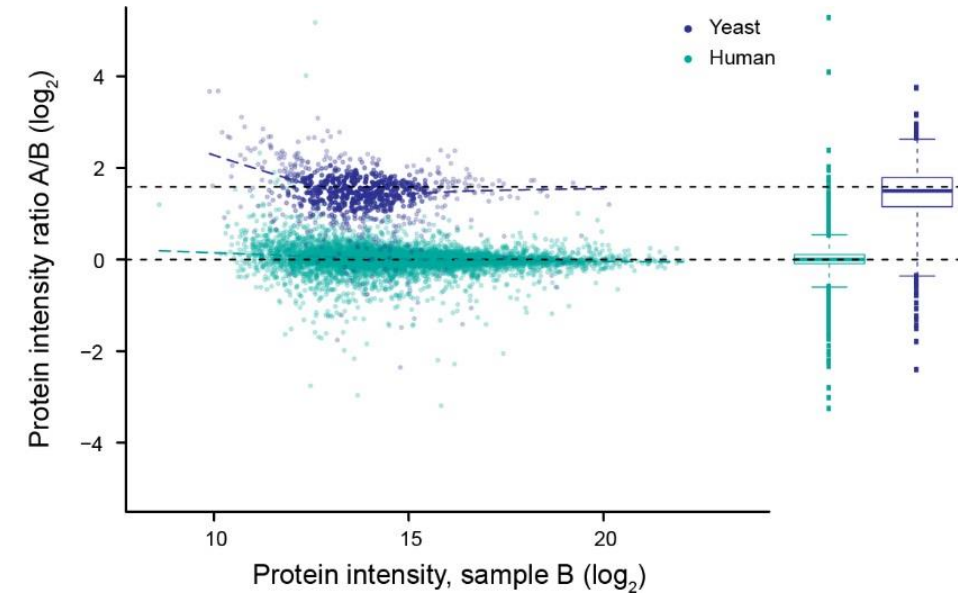
# LFQBench style multispecies mixture proteome is now used frequently for benchmarking of new methods



**Specter**  
Peckner, R et al. *Nature Methods* (2018)



**DIA-NN**  
Demichev, V et al. *Nature Methods* (2020)



**diaPASEF**  
Meier, F et al. *BioRxiv* (2020)