Enhancing limited sample- and single-cell proteomics through tailored Data Independent Acquisition

Valdemaras Petrosius, PhD

Postodoctoral researcher

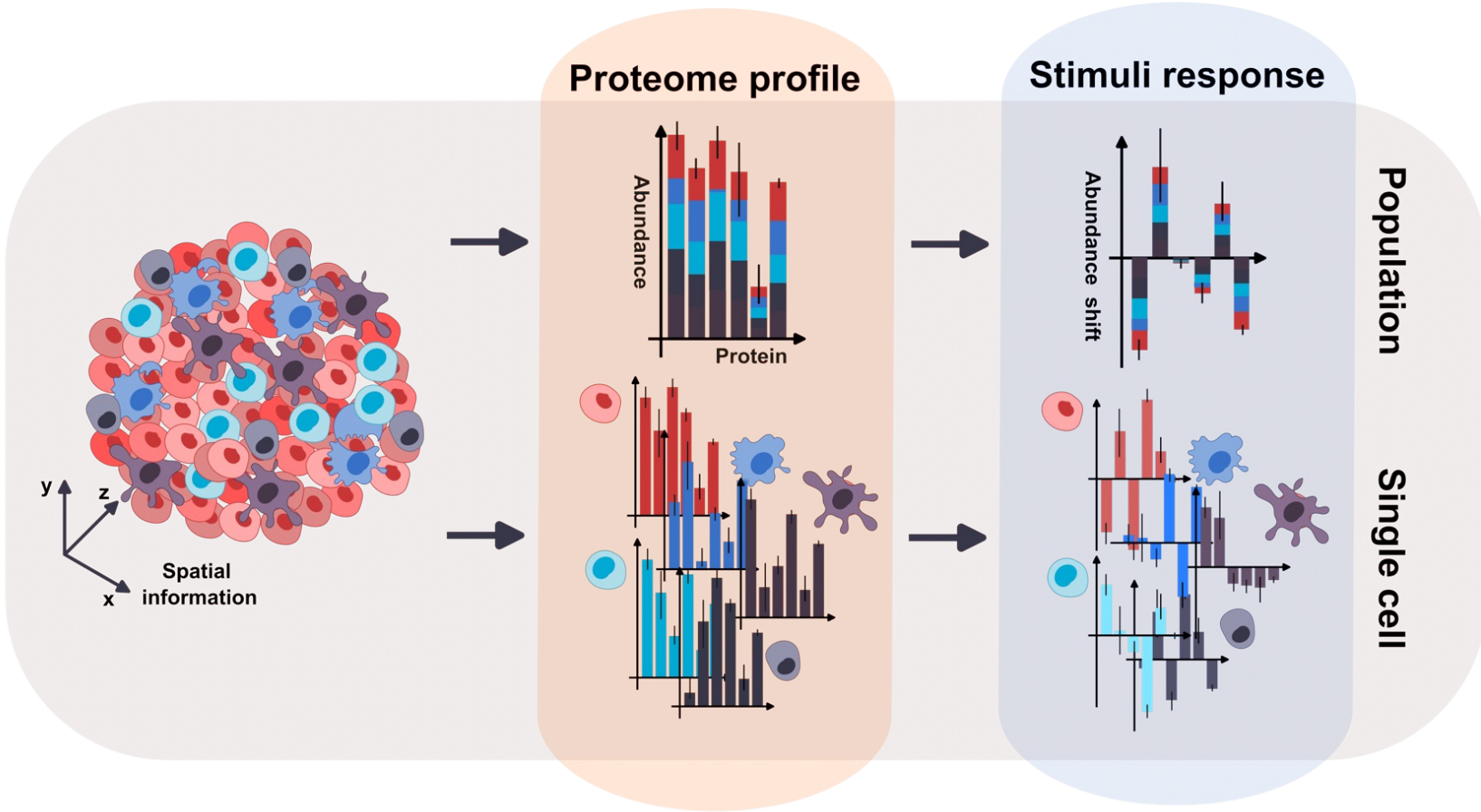
Cell Diversity Lab

Department of Bioengineering

Technical University of Denmark

Why do we need single-cell proteomics?

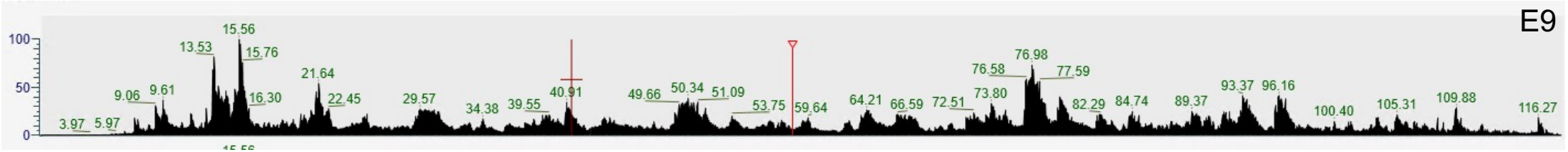
- Cell-state heterogeneity remains opaque with population-based techniques.
- Single-cell resolution can provide deeper insight into developmental programs and disease pathology.
- Single-cell proteomics (scp-MS) is/has reached a stage where it can be leverage for biological application



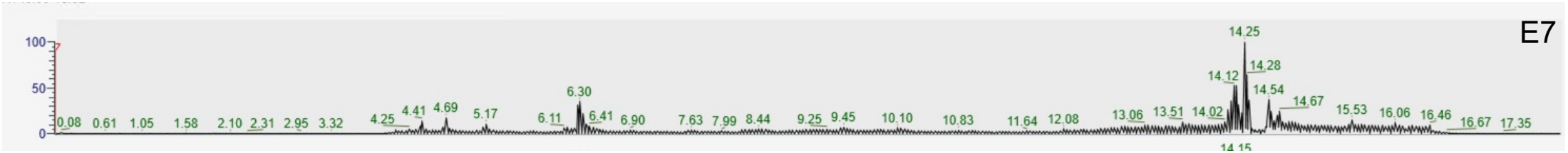
The pivotal challenge in scp-MS

Obtaining deep proteome coverage from the amount of protein encapsulated in a single cell (50-250pg)

Standard sample (250ng)

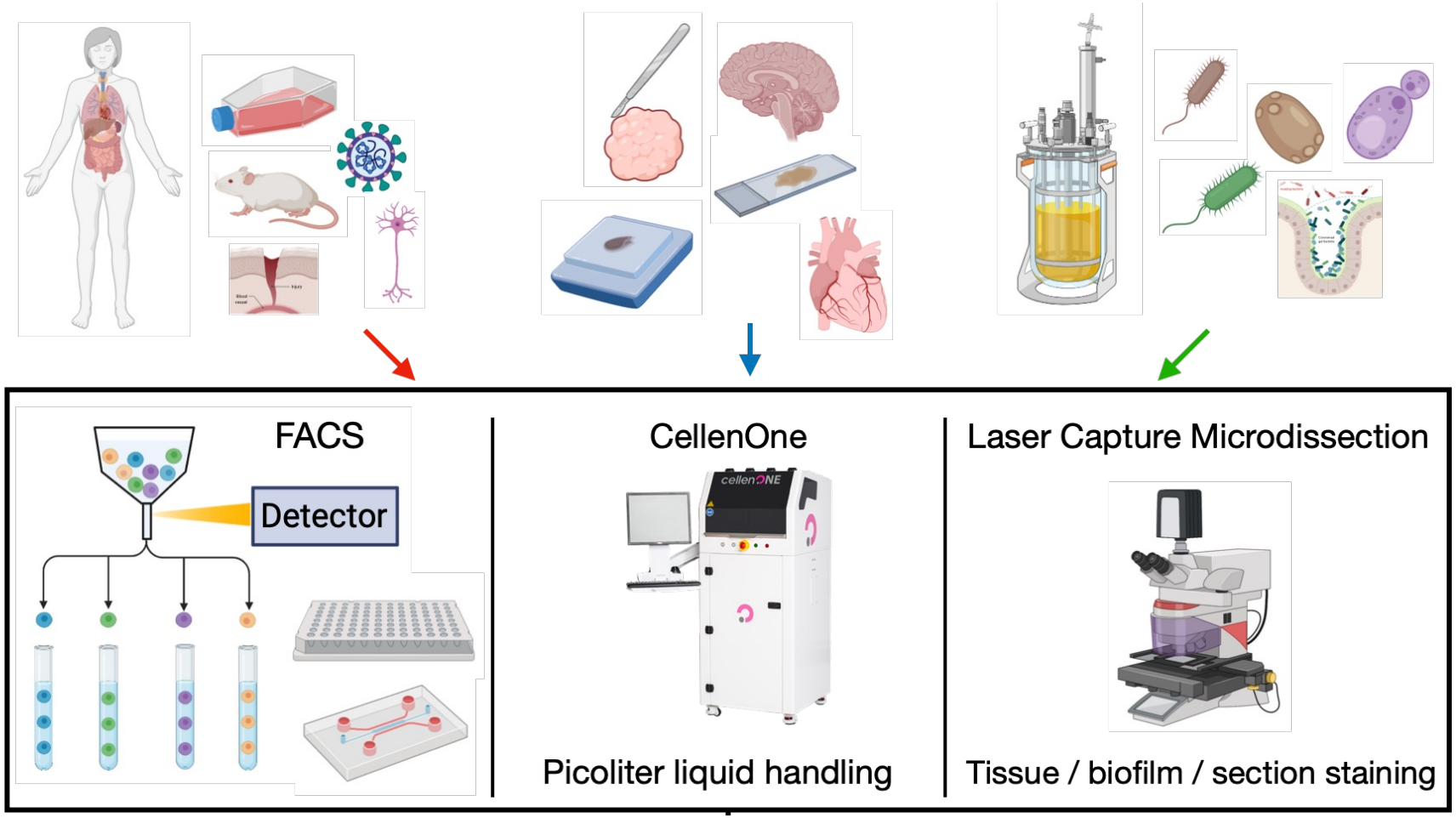


Single HEK293 cell (-- pg)

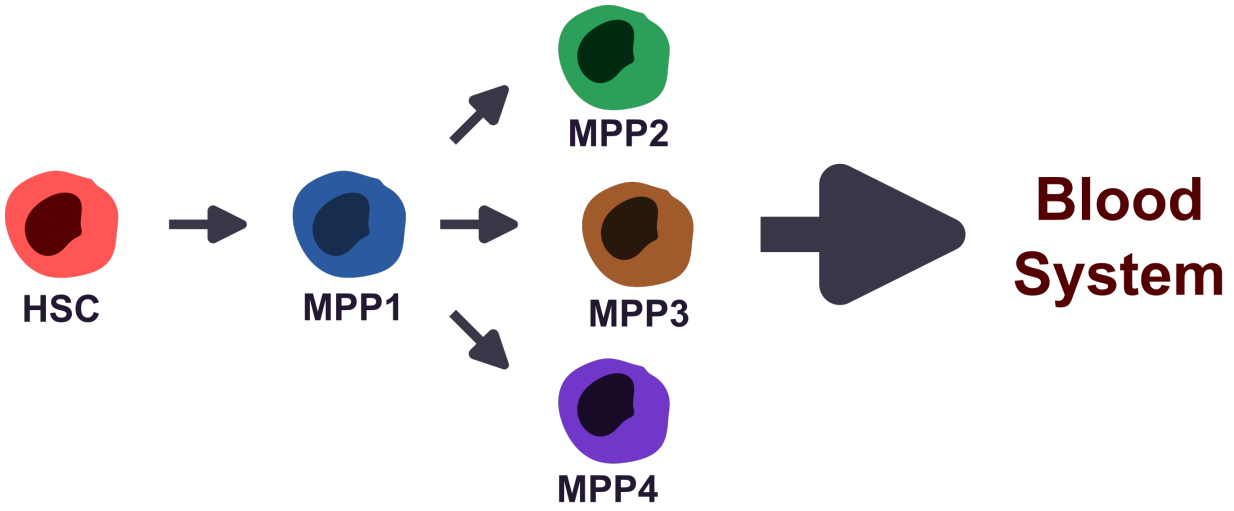


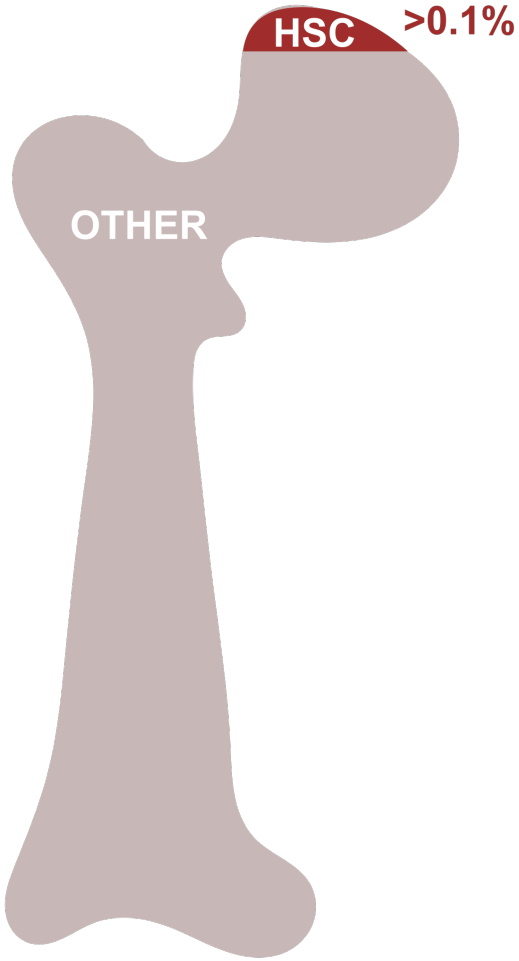
This challenge is slowly being overcome by sample preparation and tailored data acquisition approaches

Single-cell isolation

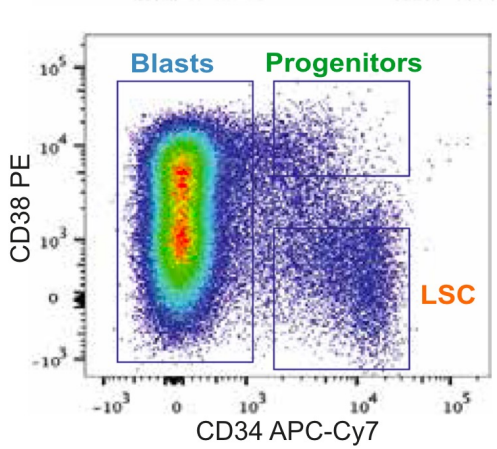


Reaching your population of interest





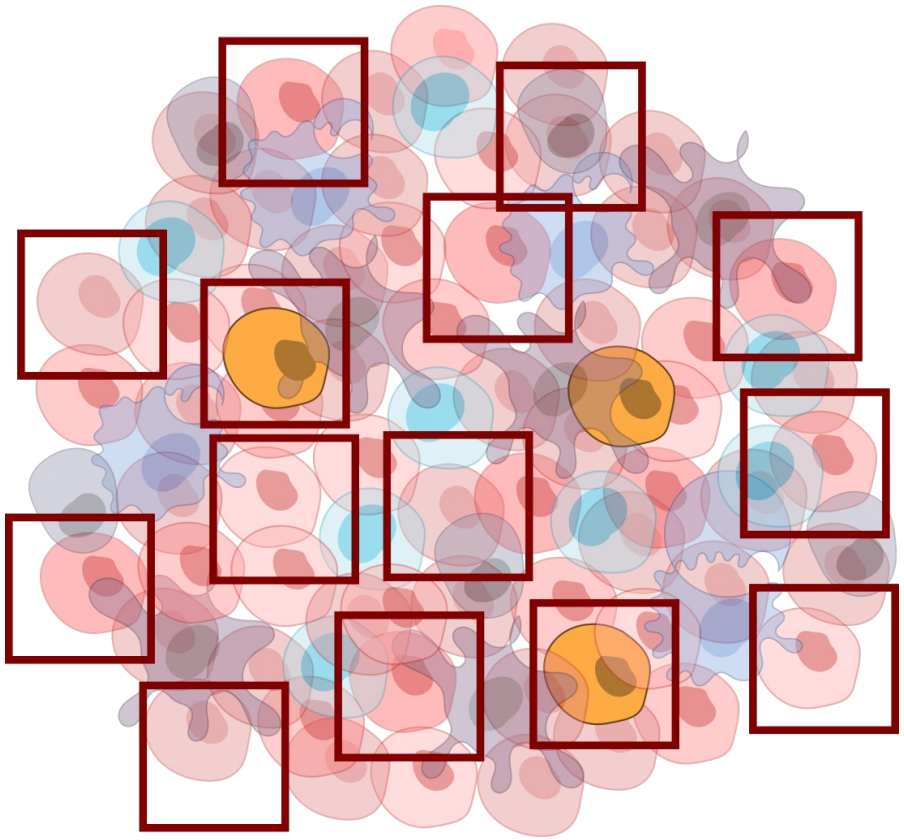
In a 5000 cell dataset only 5 stem cell will be capture!



Cell population specific markers can be used to isolate your population of interest!

What encompasses a good scp-MS acquisition method

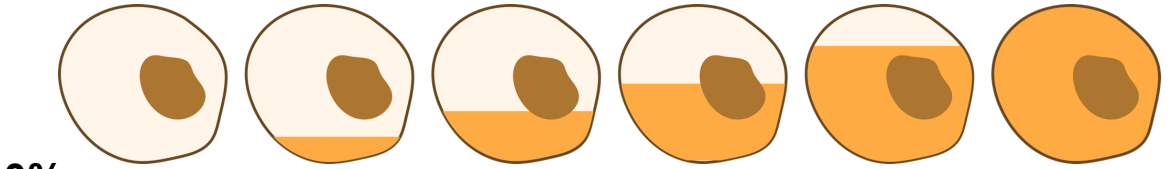
High throughput!



5 000 cell dataset cost:

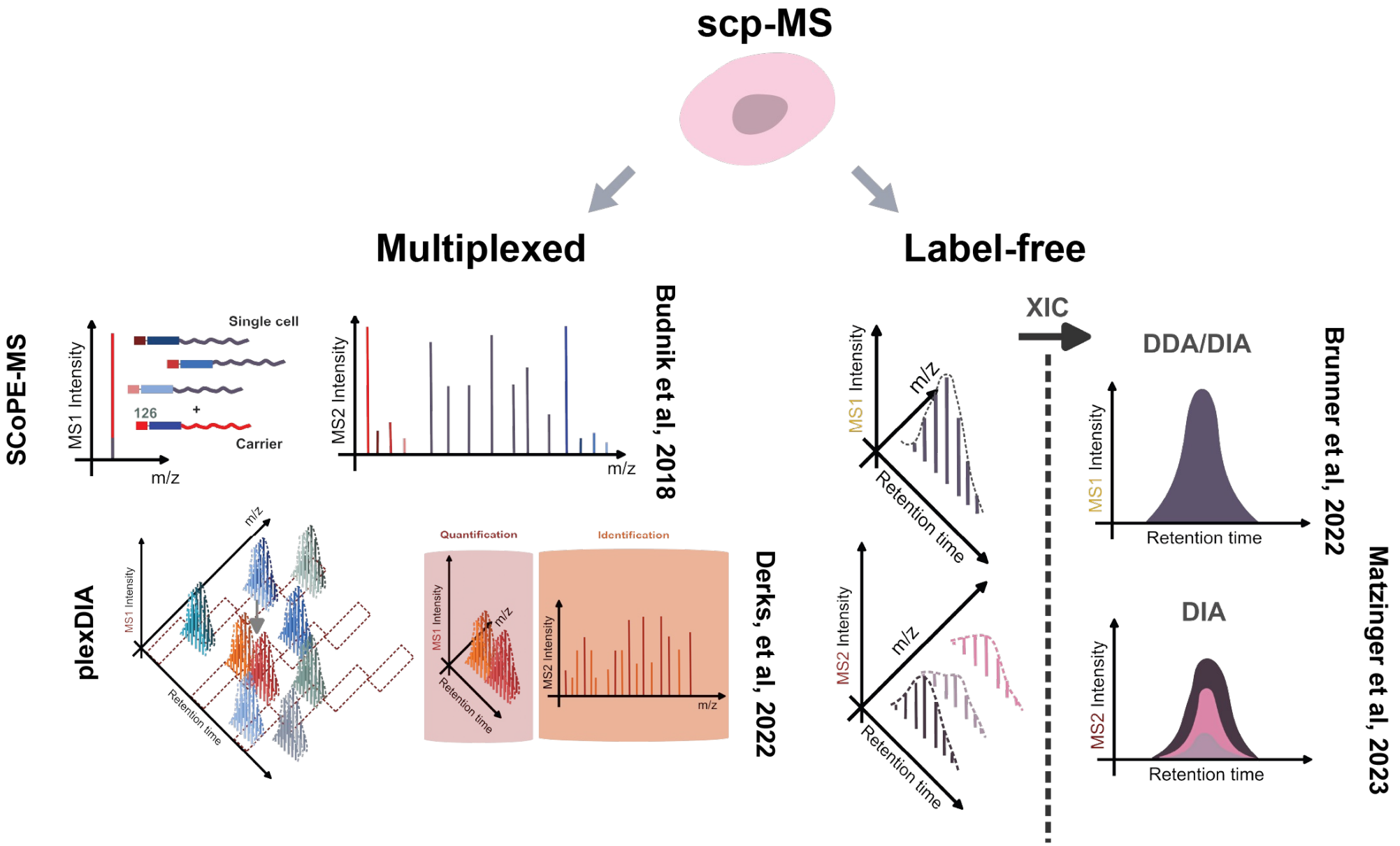
- 20 sample per day (SPD) methods ~ 8 months
- 40 SPD > 4 months
- 80 SPD ~ 2 months
- 160 SPD ~ 1 month

Proteome coverage



- The total number of protein groups expressed in a single-cell is unknown
- In a cell population, 14 000 – 18 000 proteins can be detected (Bekker-Jensen et al, 2017, Sinitcyn et al, 2023)
- In a single cell there are probably ~5000-8000 expressed proteins (SPECULATION!)

Multiplexed vs label-free single-cell proteomics approaches



Principles behind multiplexed scp-MS

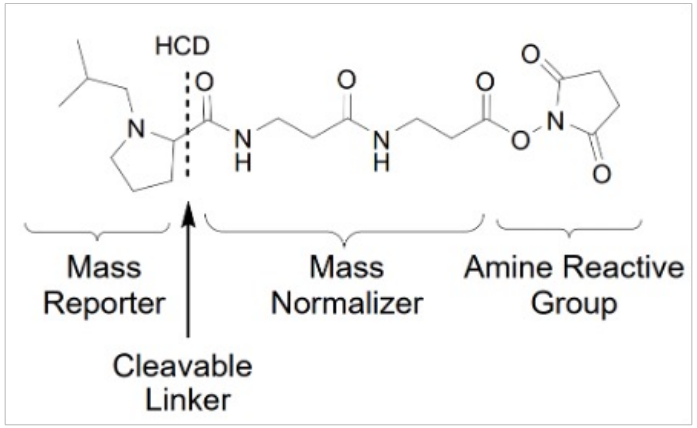
Multiplexed single-cell proteomics enabled through SCoPE-MS

SCoPE-MS – Budnik, ..., Slavov 2018

SCoPE2 – Specht, ..., Slavov 2021

SCoPE2 – Petelski, ..., Slavov 2021

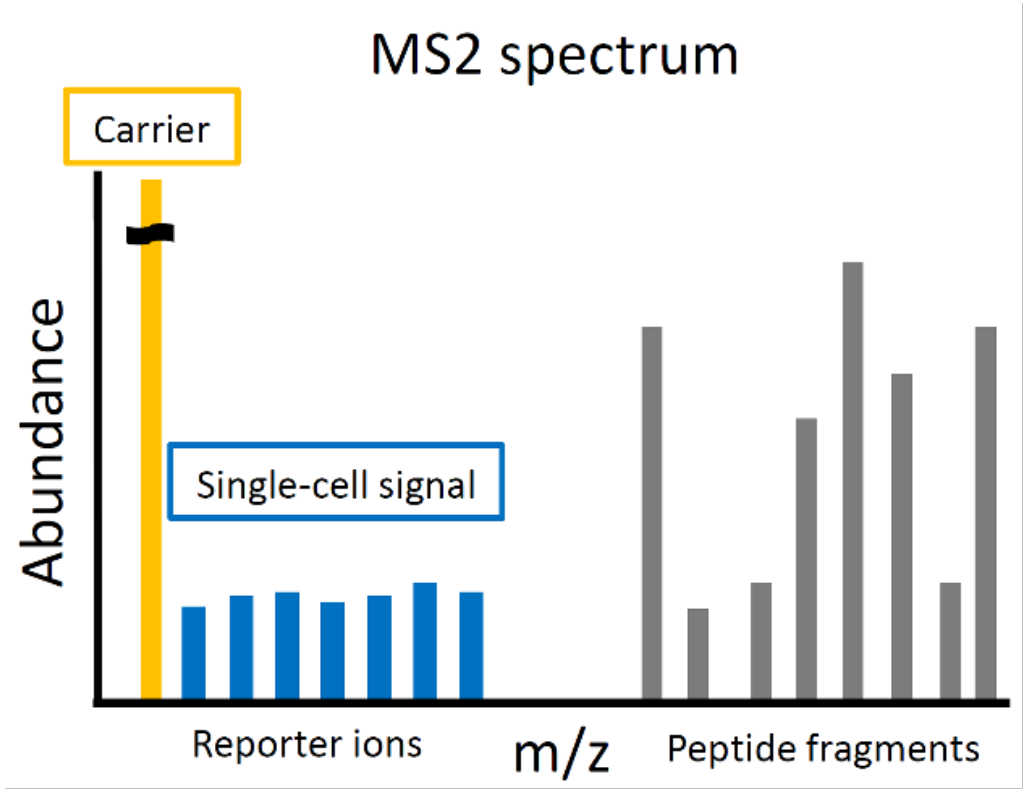
1.Isobaric multiplexing using tandem mass tag (TMTpro)



Example TMTpro labelled peptide:

(N-Term TMTpro)-VSHVSTGGGASLELLEG**K**(**TMTpro**)

2.Addition of carrier channel to facilitate peptide identification (e.g. 200 cell equivalent)



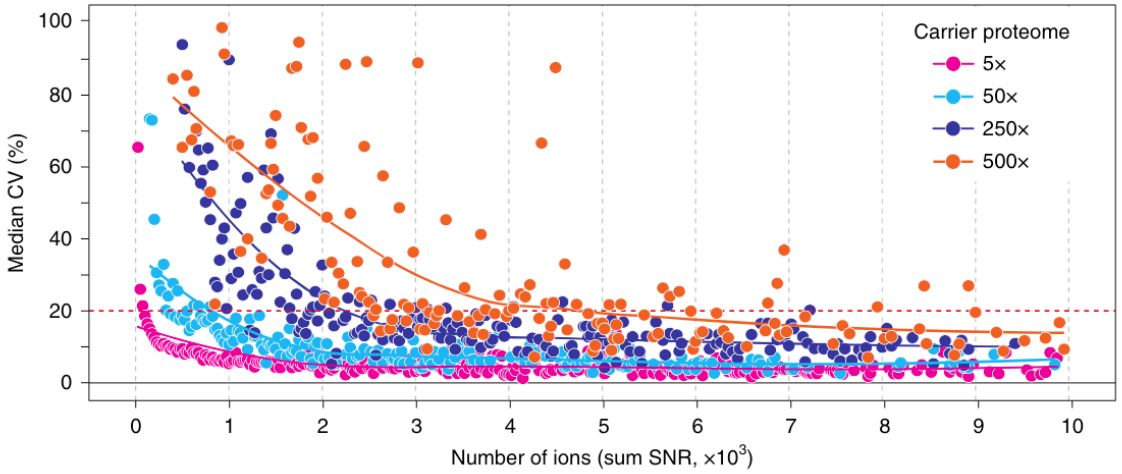
LFQ vs SCoPE-MS

Pros

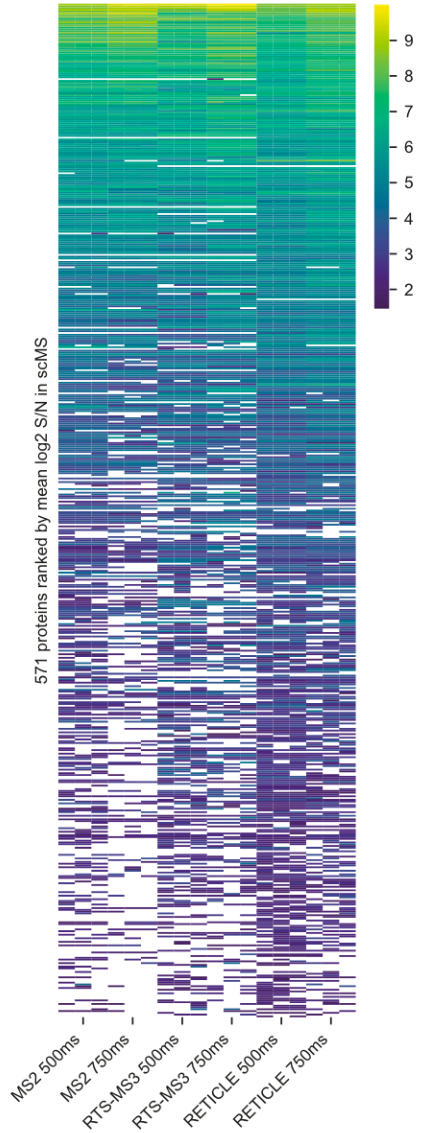
- Higher throughput then label-free methods (320 cells per day vs 40 cells per day)

Cons

- The carrier channel introduces a bias in the identification and quantification
- High degree of missing values (~40%)
- Lower proteome coverage compared to LFQ (~1000 vs 2000 protein groups)
- Collecting enough cells for the carrier channel can be challenging when primary patient material is used



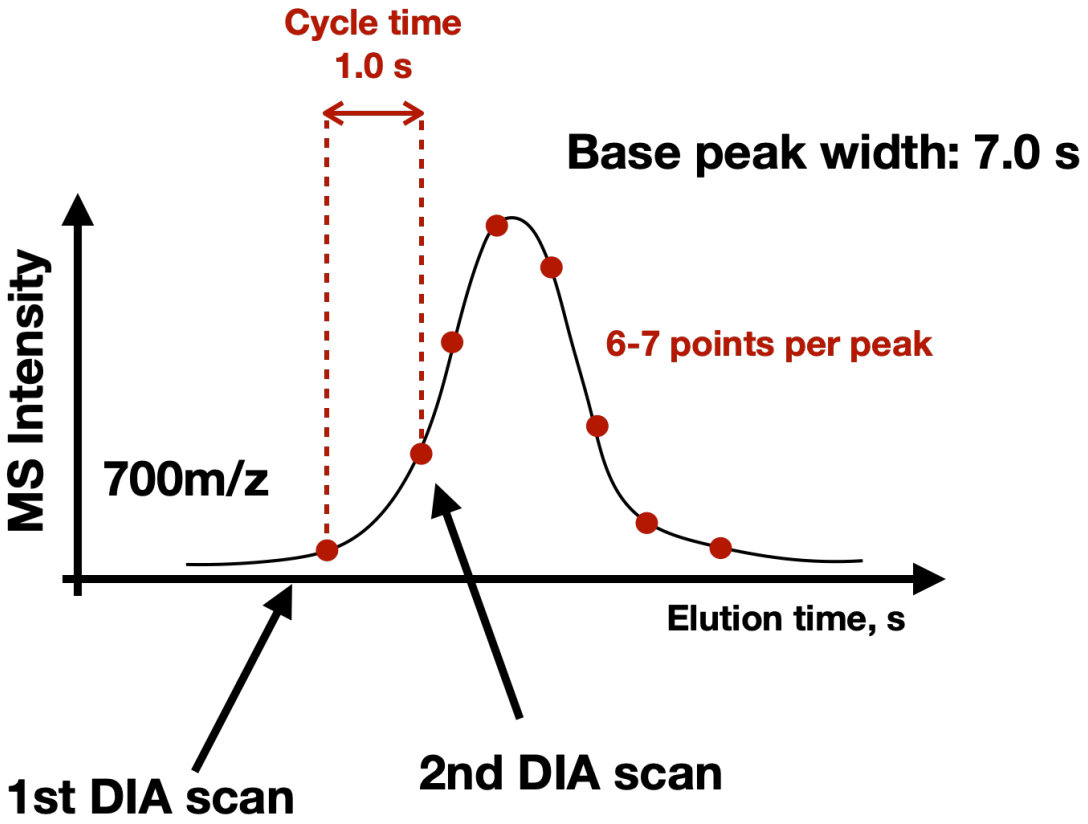
Cheung et al, 2021

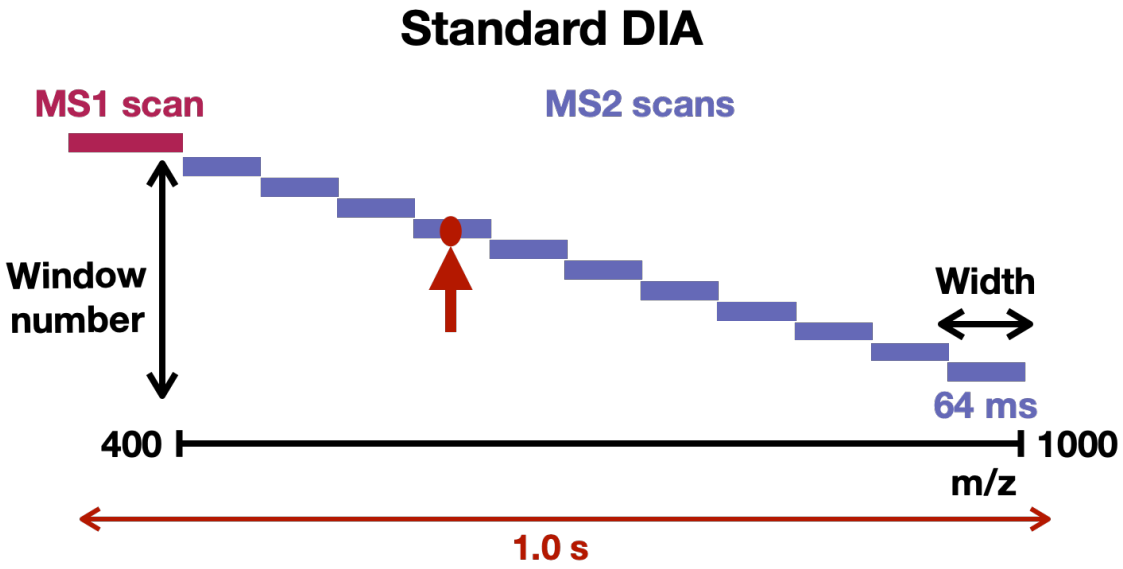


Furtwängler et al, 2022

DIA for single-cell proteomics

The scan cycle time dictates the method optimization space



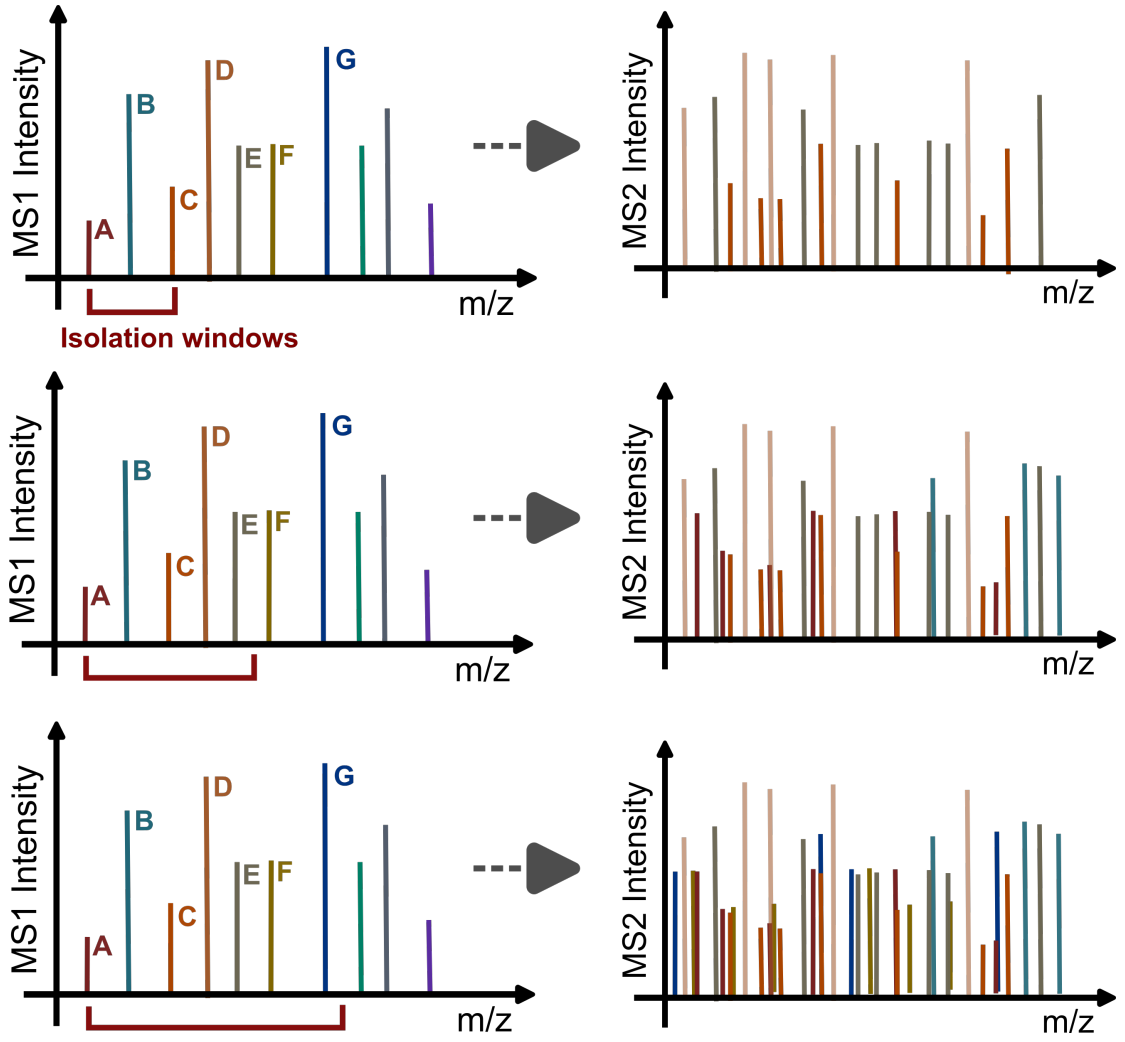


To compensate for increased resolution e.g. 30K -> 60K, the DIA isolation window with has to be accordingly increased to maintain cycle time

What DIA isolation window should I use?

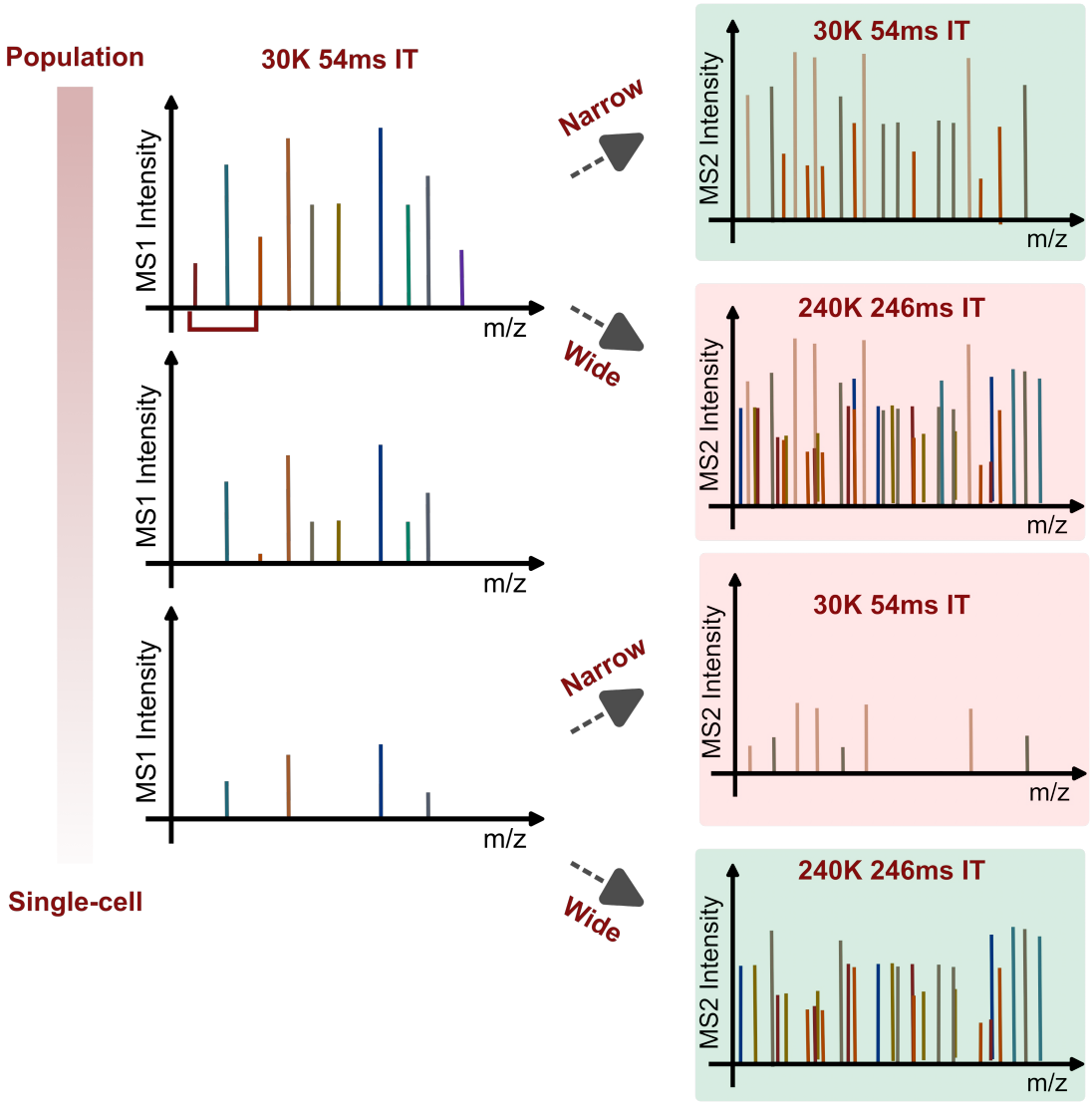
- **Increasing the isolation window size during DIA based acquisition should in theory hamper peptide identification due to more extensive precursor co-isolation.**

- **In high load samples (>10ng) narrow isolation windows with relatively low resolution generally give best performance.**



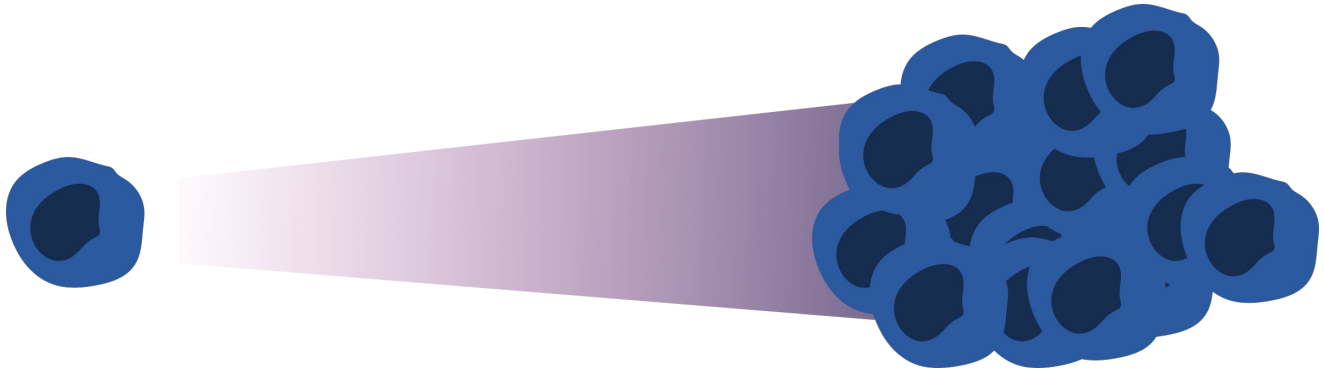
DIA isolation window size for limited input

- DIA method design for limited input samples (< 10ng) has not been comprehensively explored.



Aim of the project

Aim : Establish limited input tailored LFQ DIA methods for optimal proteome coverage while maintaining quantitative accuracy



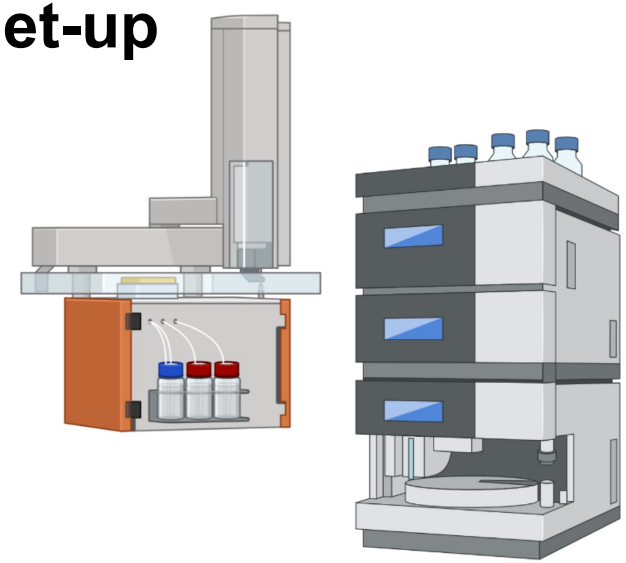
Single-cell

Rare cell population
Minimal-residual disease
FFPE tissue slices

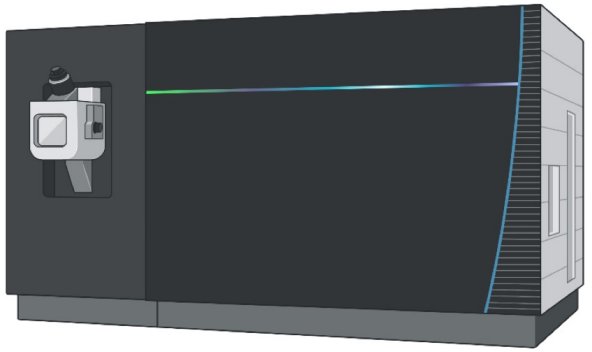
Results

Experimental set-up

Pierce Hela digest and HEK293 single-cells

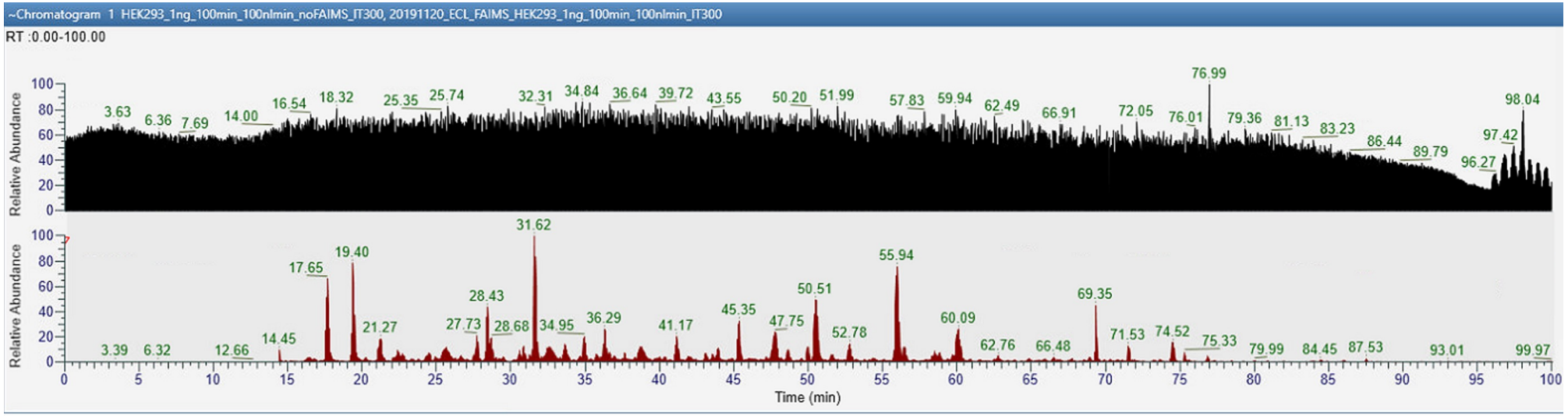


Evosep One or Dionex UltiMate 3000



Orbitrap Eclipse Tribrid + FAIMSPro

Spectronaut™
powered by Pulsar

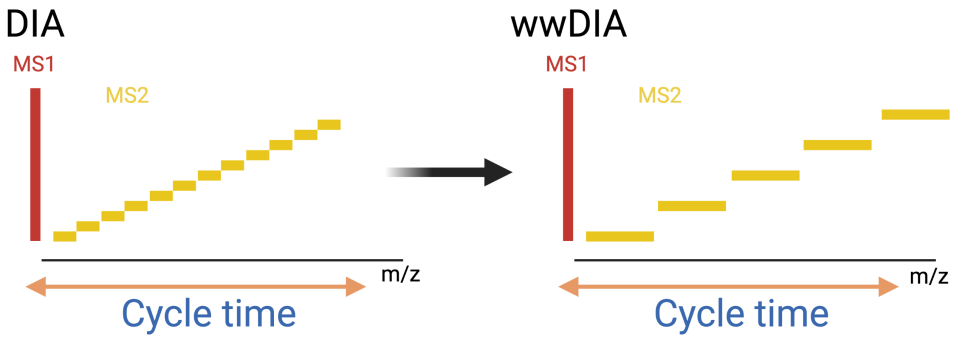


Without FAIMS

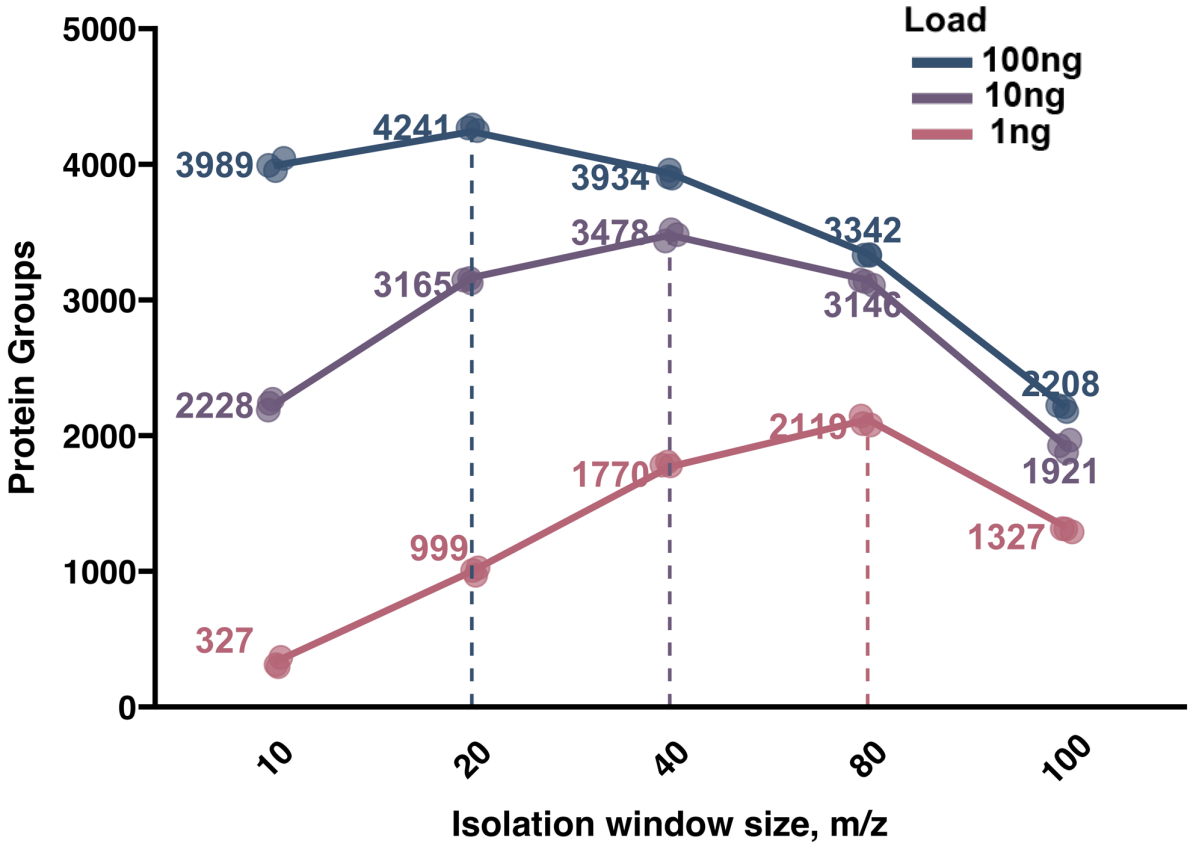
With FAIMS

Inspired by: Karl Mechtler, IMP

Evaluation of wide-windowed DIA for limited sample -and scpMS

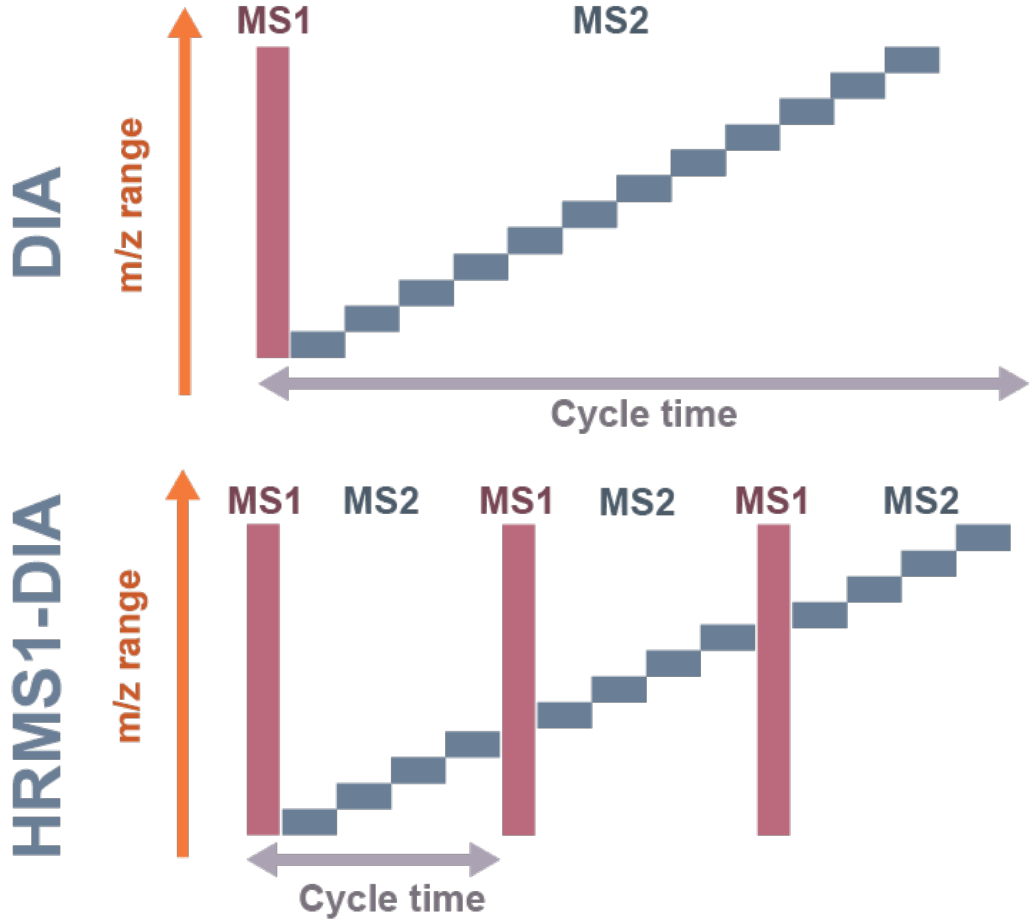


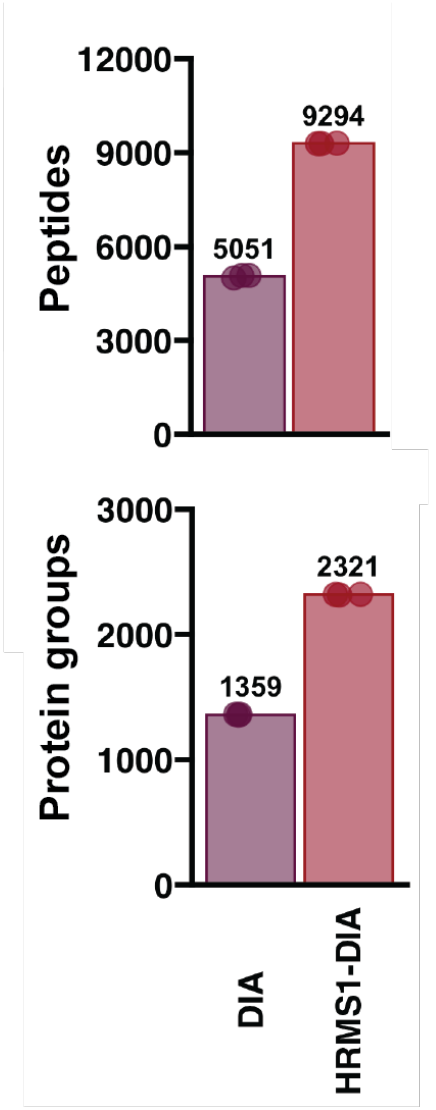
Resolution		Window		
MS1	MS2	Size	Number	Cycle-time
120K	15K	10mz	40	1.536s
120K	30K	20mz	20	1.536s
120K	60K	40mz	10	1.536s
120K	120K	80mz	5	1.536s
120K	120K	100mz	4	1.280s



- Wider isolation windows allow for the accommodation of higher resolution/IT.
- Detrimental chimeric spectra effects are overcome by higher resolution/IT

Evaluation of high resolution MS1 (HRMS1) DIA for limited sample -and scpMS

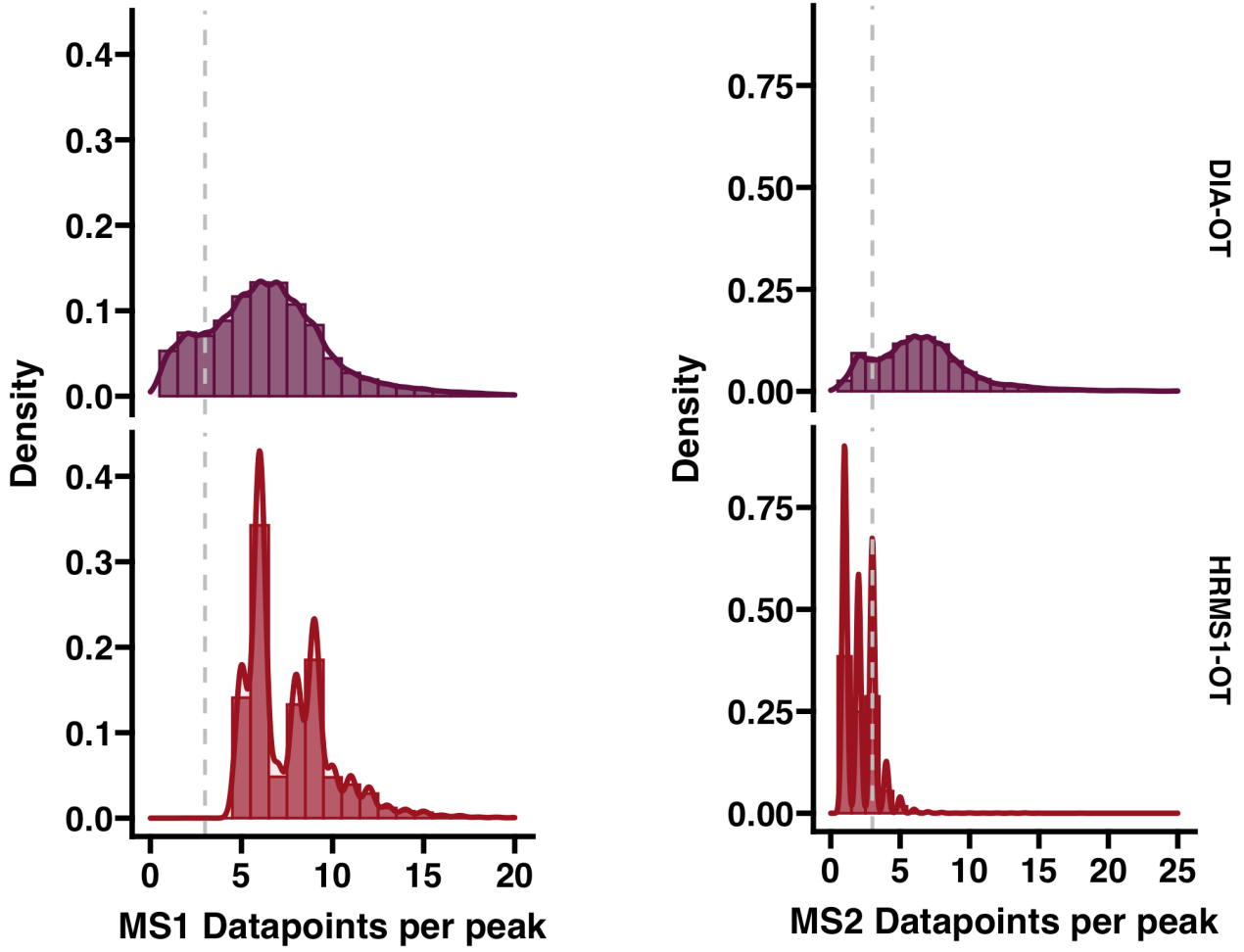


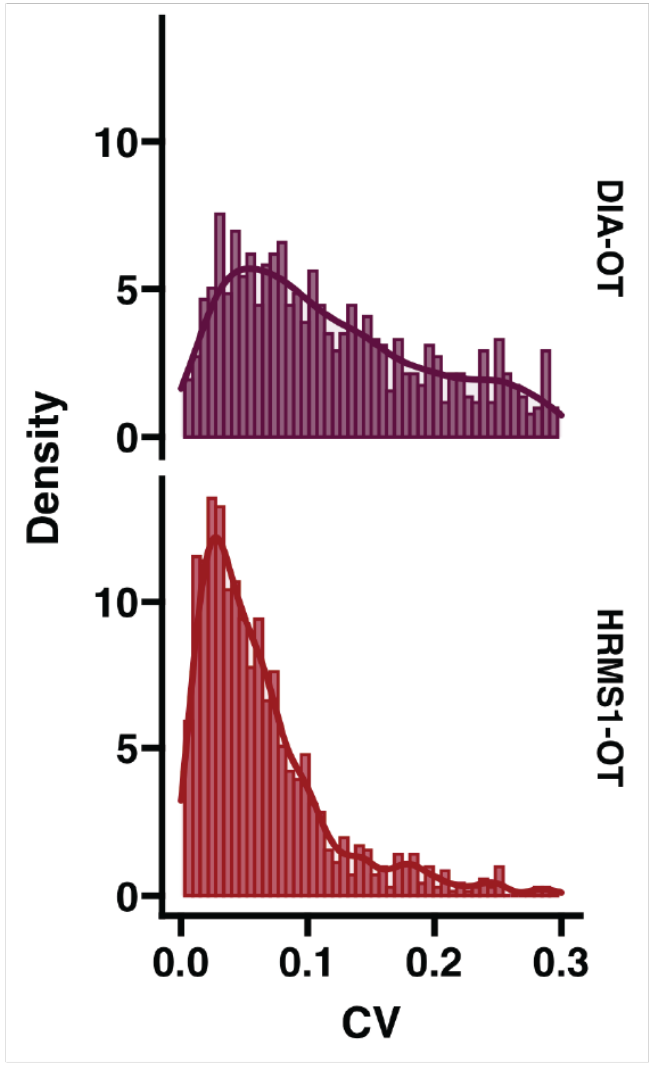


Xuan, Y *et al.* 2020

1 ng injections at 58min gradient (20 sample per day)

Evaluation of high resolution MS1 (HRMS1) DIA for limited sample -and scpMS





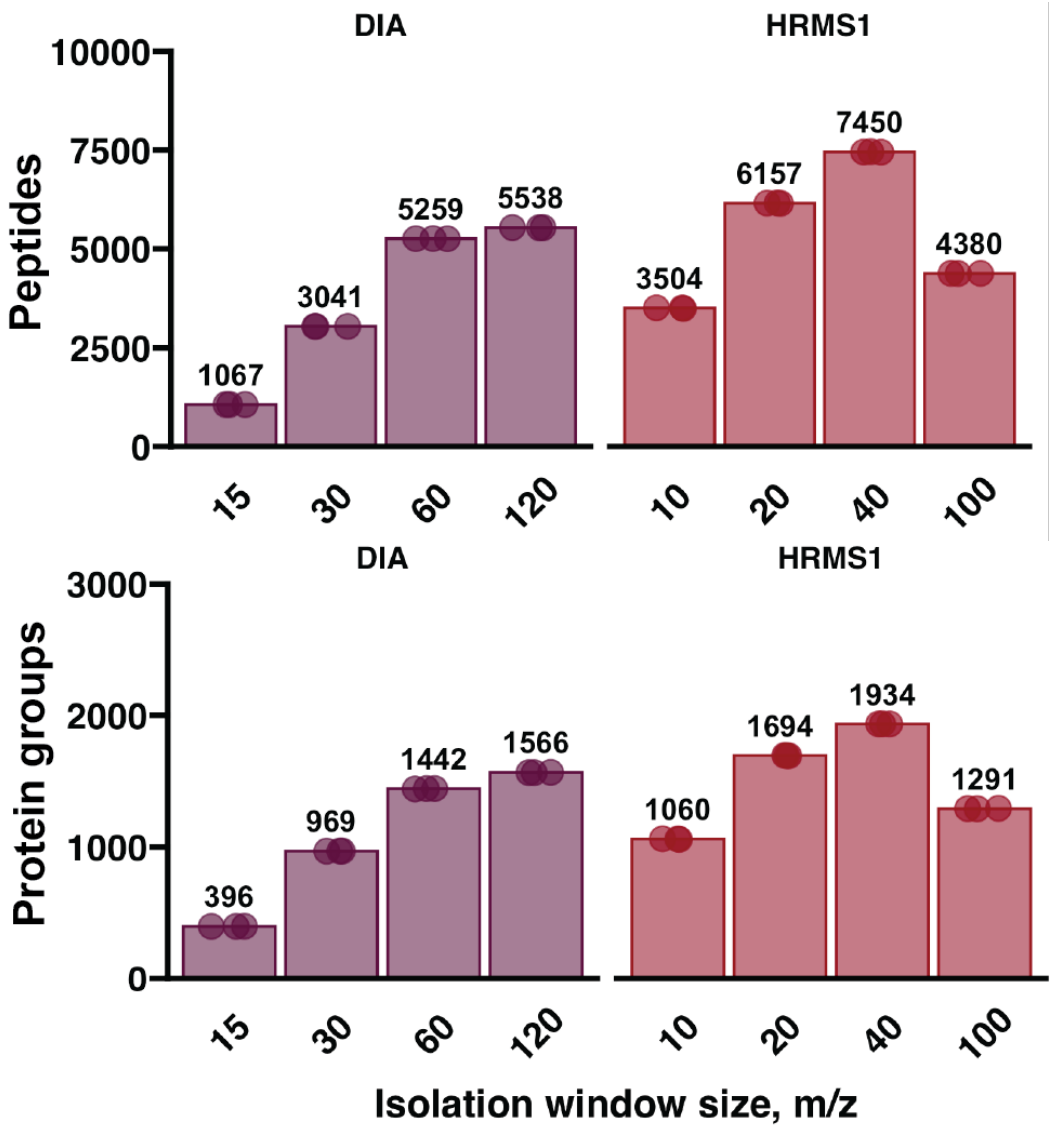
Xuan, Y *et al.* 2020

1 ng injections at 58min gradient (20 sample per day)

Wise-isolation window HRMS1 (WISH) DIA enhanced quantified proteome depth

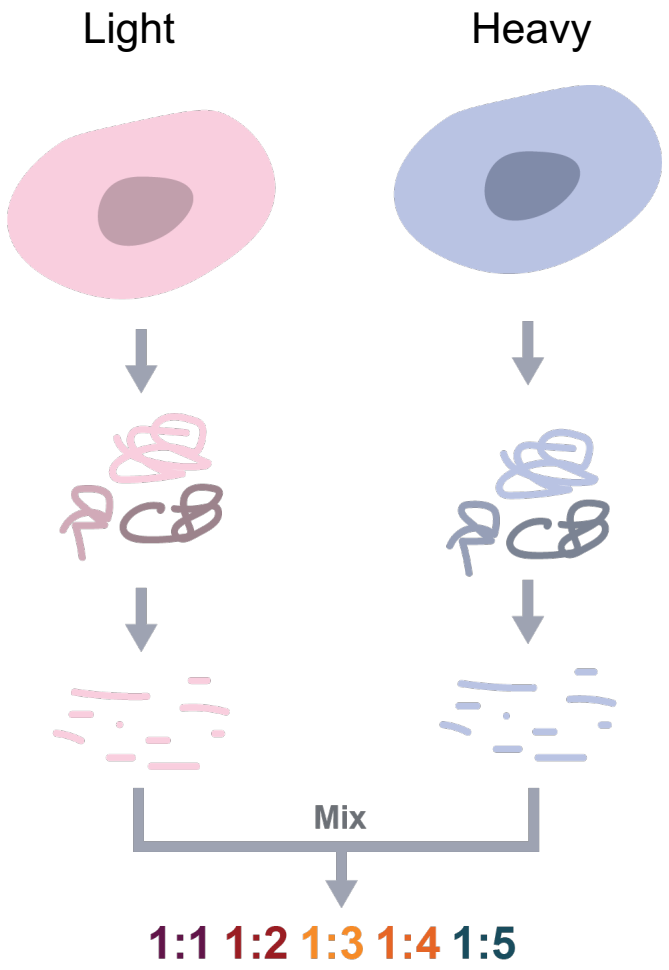
Resolution		Window		DIA
MS1	MS2	Size	Number	Cycle-time
120K	15K	15mz	40	1.536s
120K	30K	30mz	20	1.536s
120K	60K	60mz	10	1.536s
240K	120K	120mz	5	1.536s

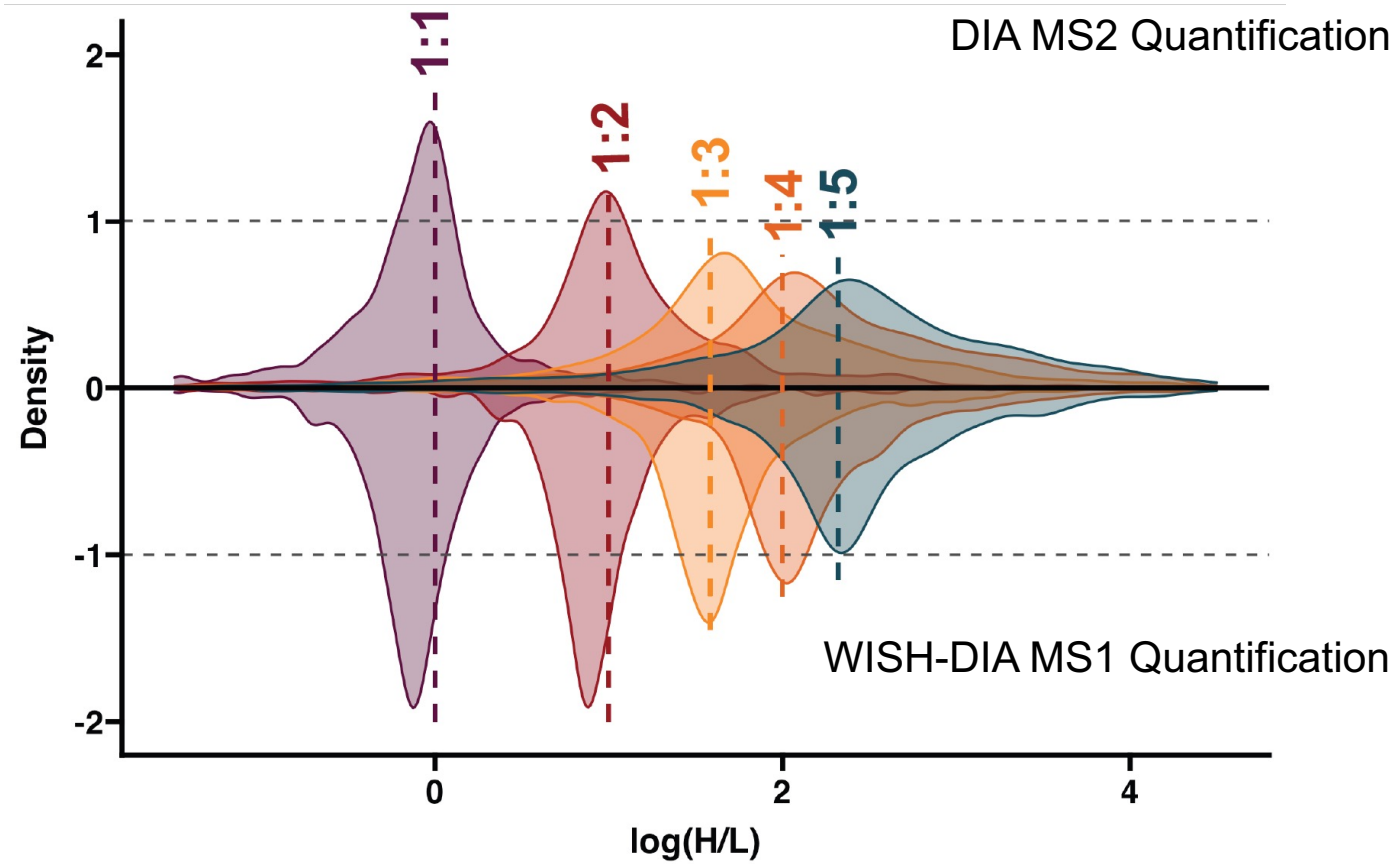
Resolution		Window		HRMS1
MS1	MS2	Size	Number	Cycle-time
120K	30K	10mz	60	1.536s
120K	60K	20mz	30	1.536s
120K	120K	40mz	15	1.536s
240K	240K	100mz	6	1.280s



1 ng injections at 31min gradient (40 sample per day)

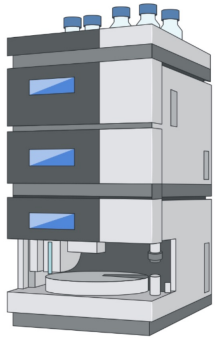
Wise-isolation window HRMS1 (WISH) DIA enhanced quantified proteome depth



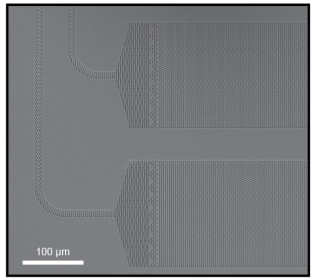


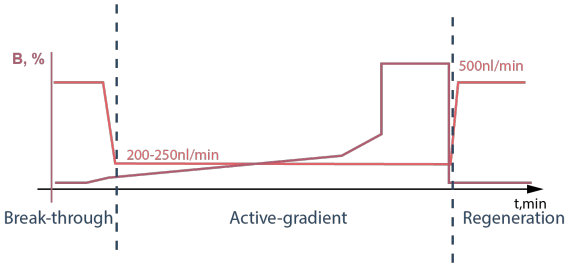
1 ng injections at 31min gradient (40 sample per day)

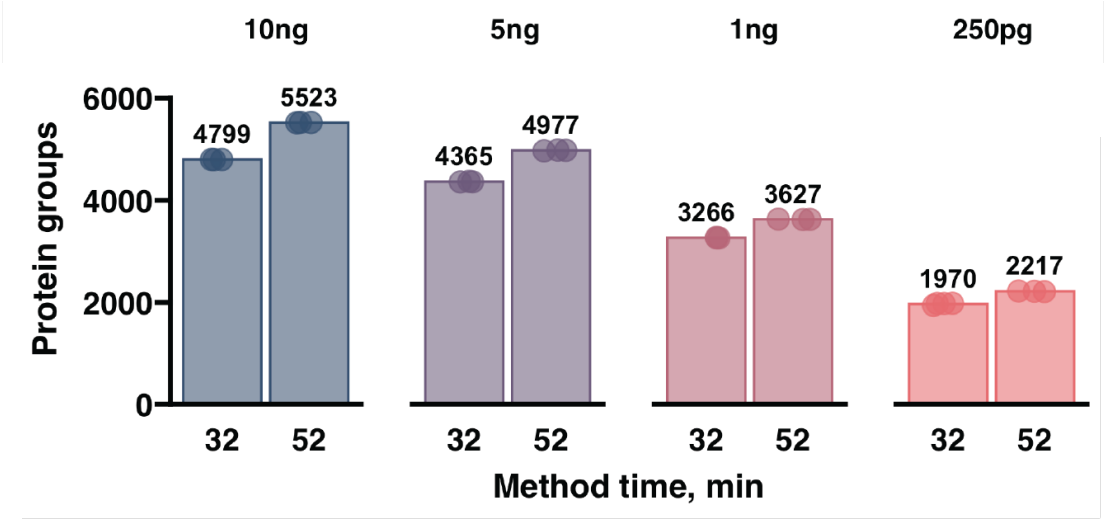
Utilizing the synergy between *µ*PAC Neo Low Load and wide isolation window HRMS1-DIA for low-input proteomics



*µ*PAC Neo Low Load

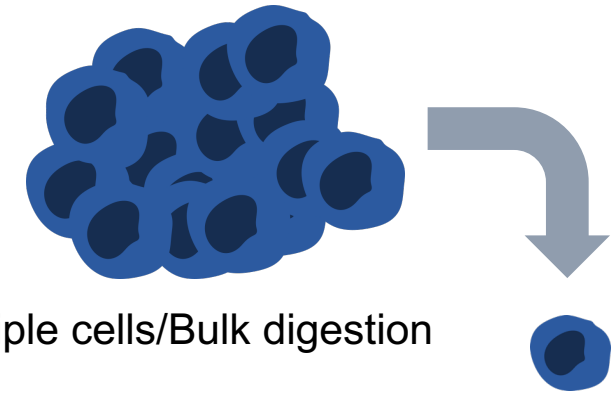






Increasing proteome coverage with the use of spectral libraries

High-load library search

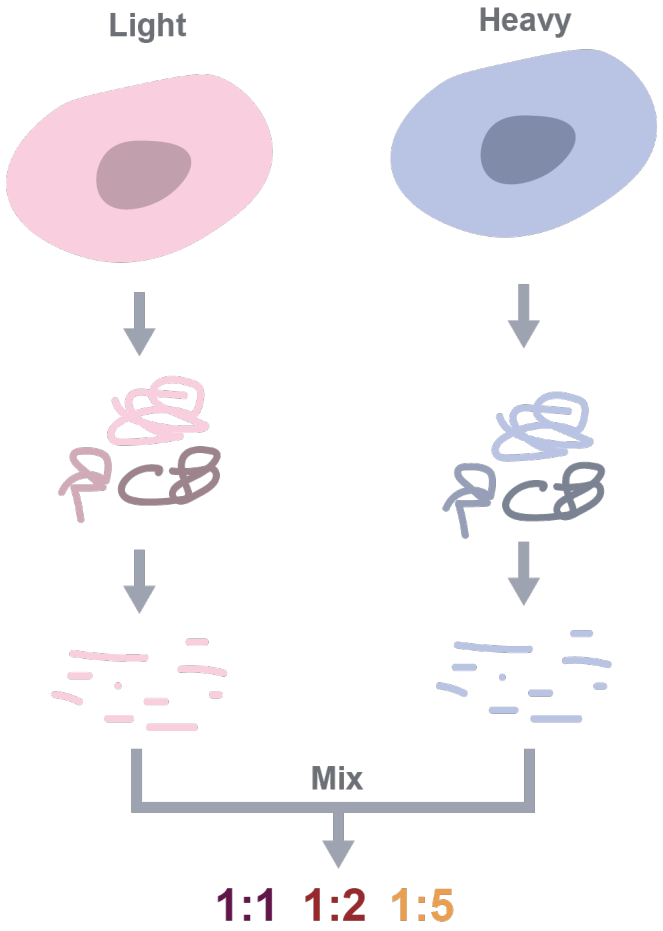


Multiple cells/Bulk digestion

Gas-phase fractionated library search

What is the quantification quality of the additionally identified proteins?

Evaluating the use of libraries to increase proteome coverage of limited input samples



Light + Heavy = 10ng High-Load (HL)

Dilute 10x = 1ng Samples (direct + GPF)

For HL libraries the 10ng sample of each mix was run once

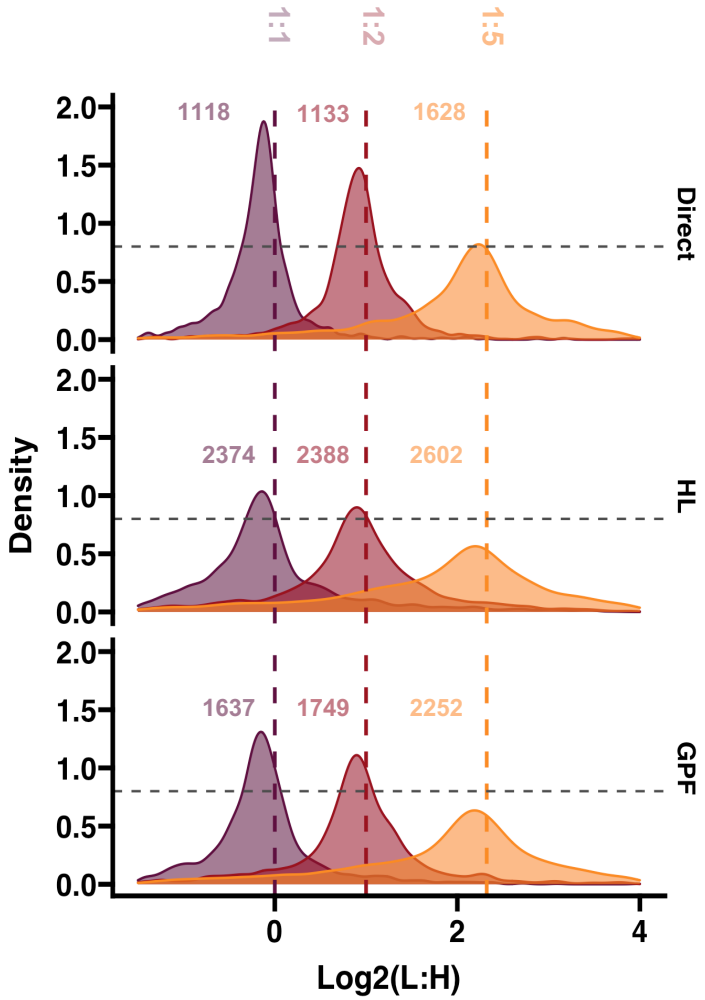
For GPF each mix was run 6 times (100mz each)

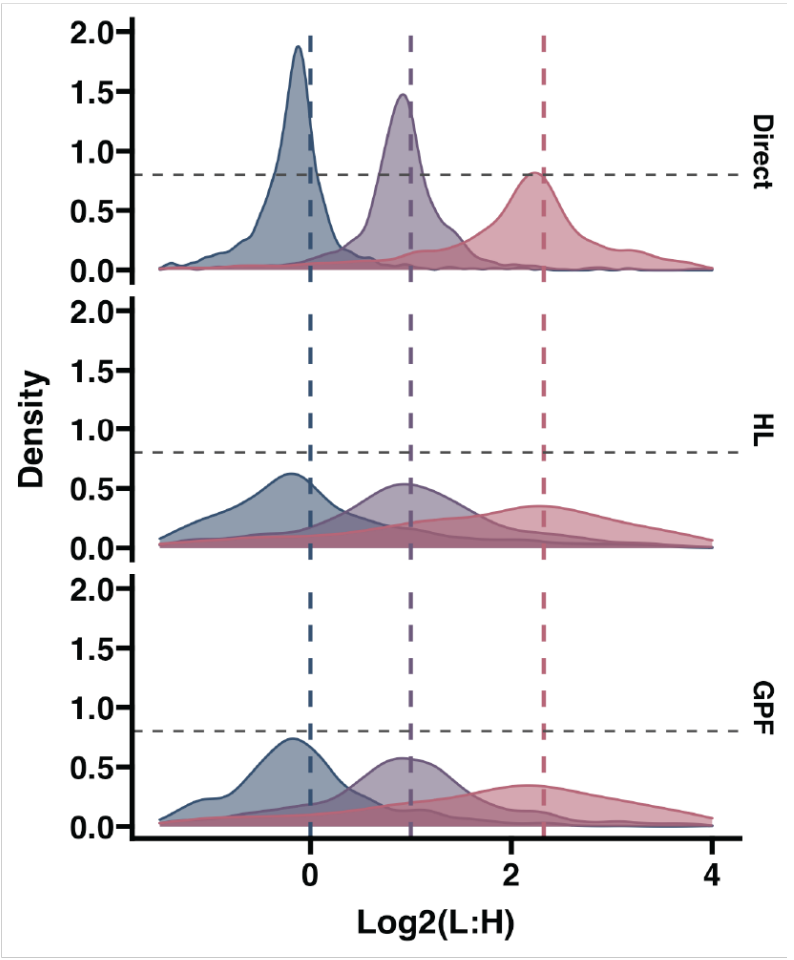
For direct each mix was run in technical triplicate (full mz)

Evaluating the use of libraries to increase proteome coverage of limited input samples

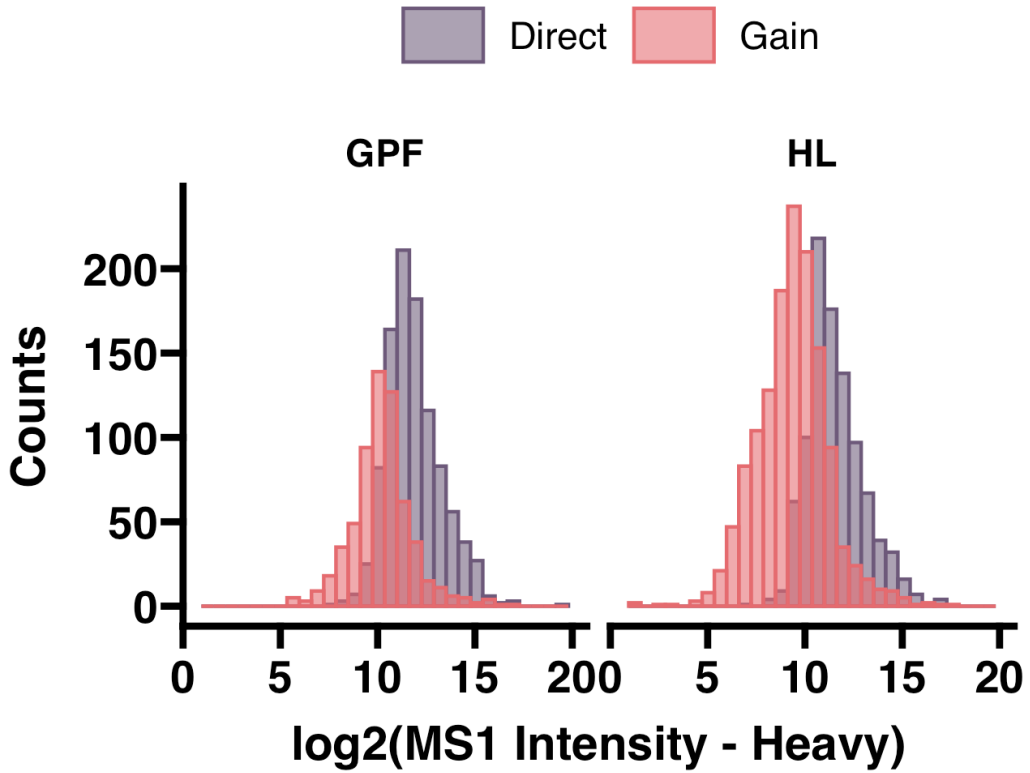
- **Direct** – 1ng searched as directDIA
- **HL** – 1ng searched with 10ng library
- **GPF** – 1ng + GPF runs searches as directDIA

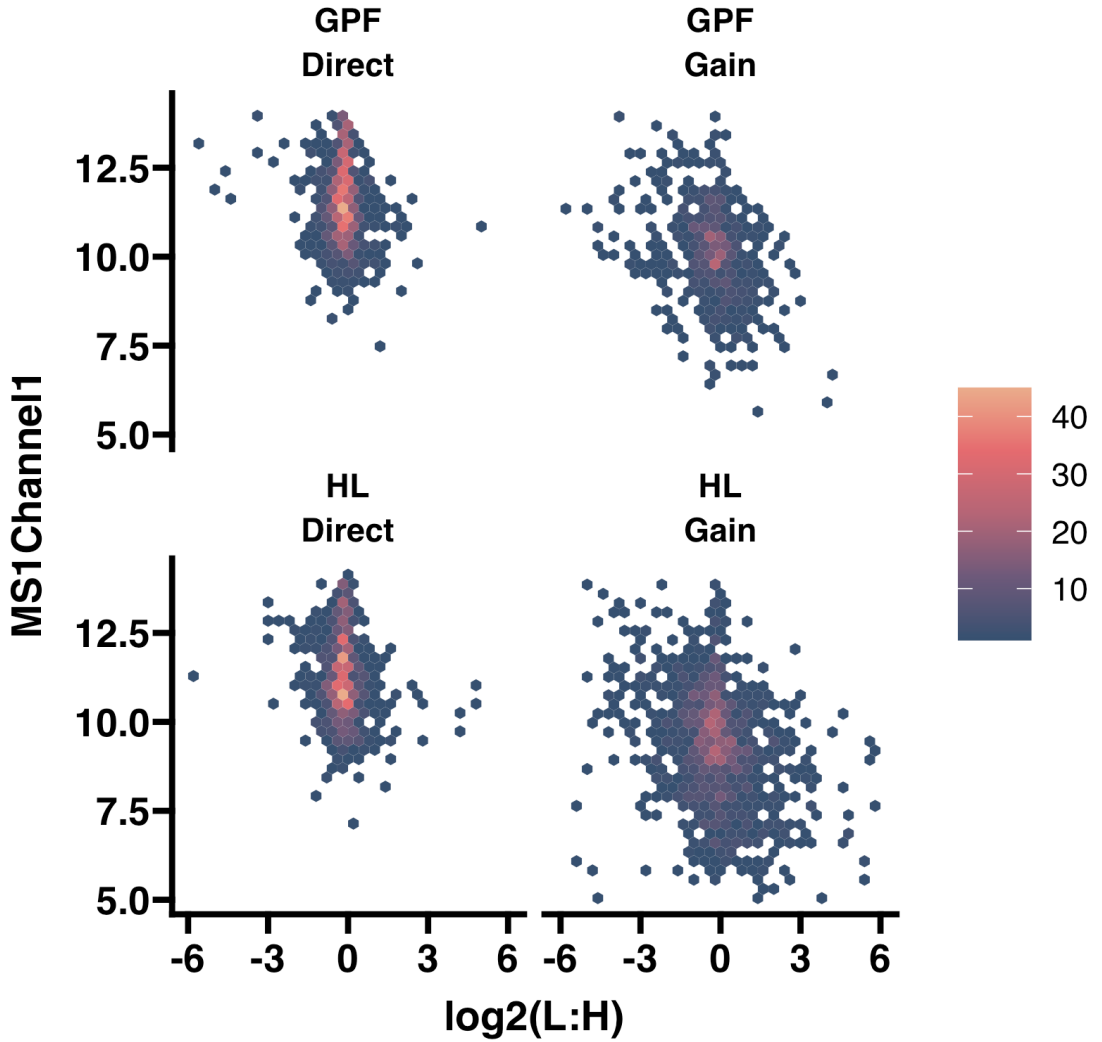
What is the accuracy of just the additional IDs?



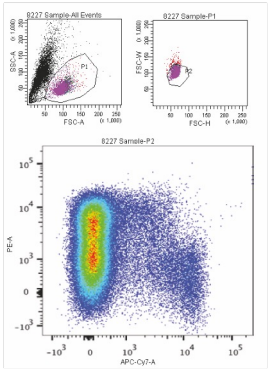


Evaluating the use of libraries to increase proteome coverage of limited input samples

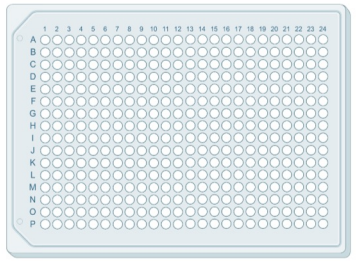




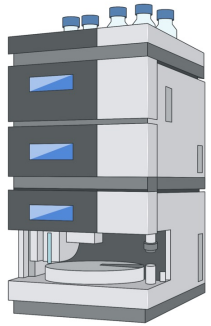
Single-cell DIA method optimization

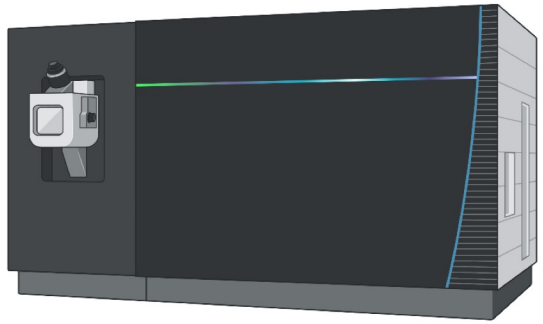


384-well plate



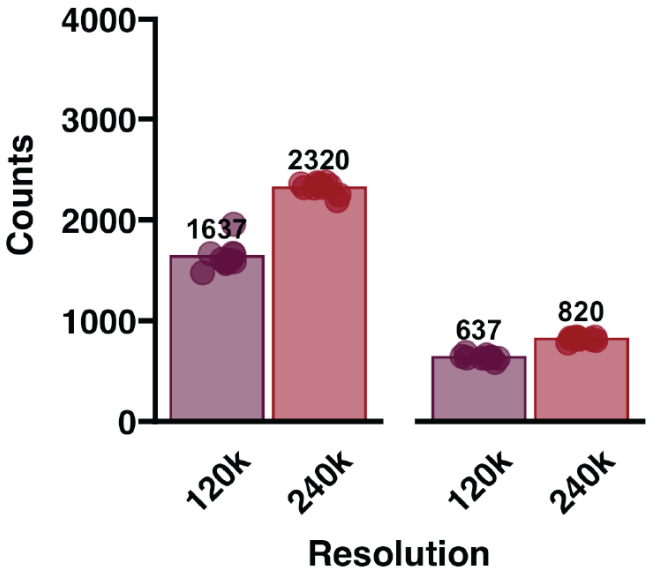
16hr digest (37° C)



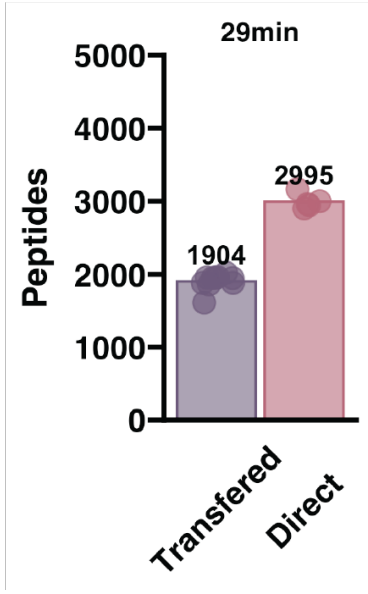


LC-MS analysis of single cells

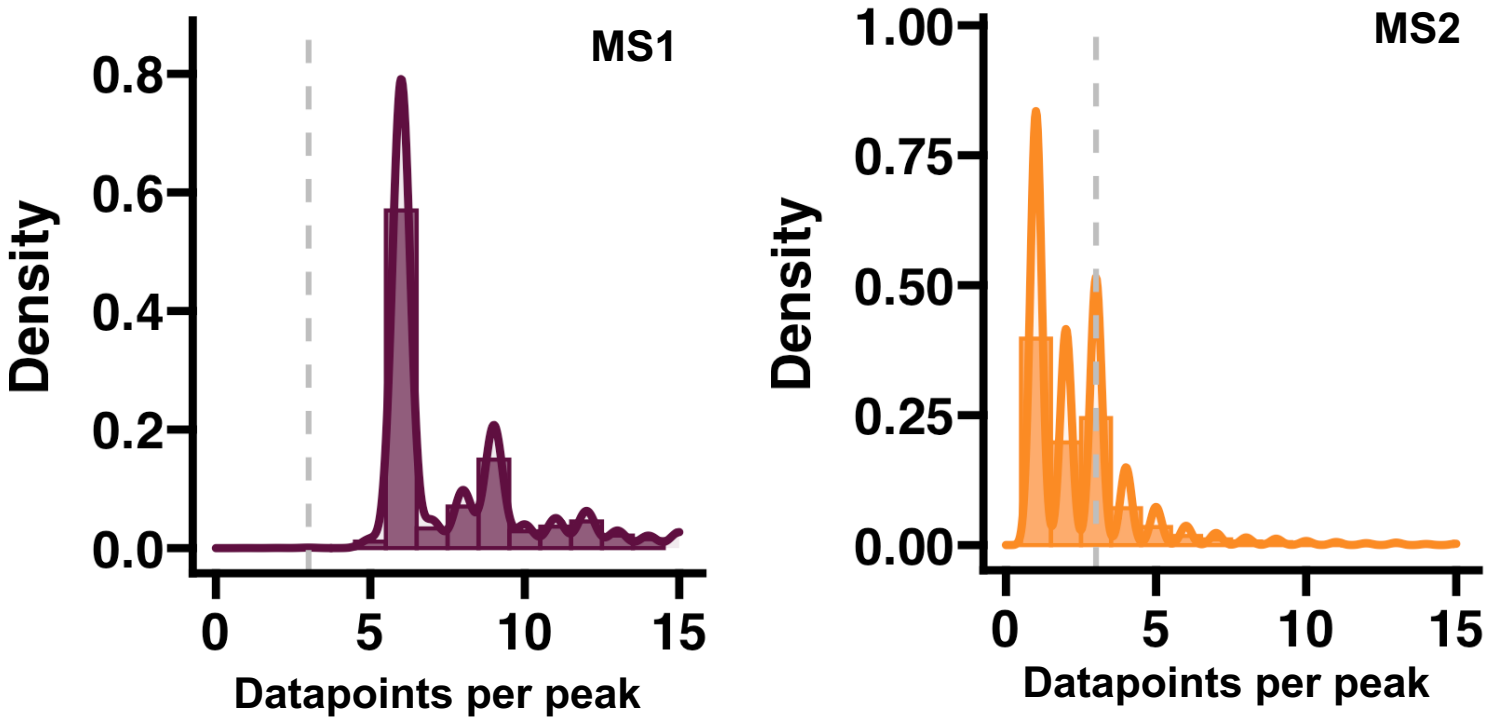
Peptides Protein.Groups



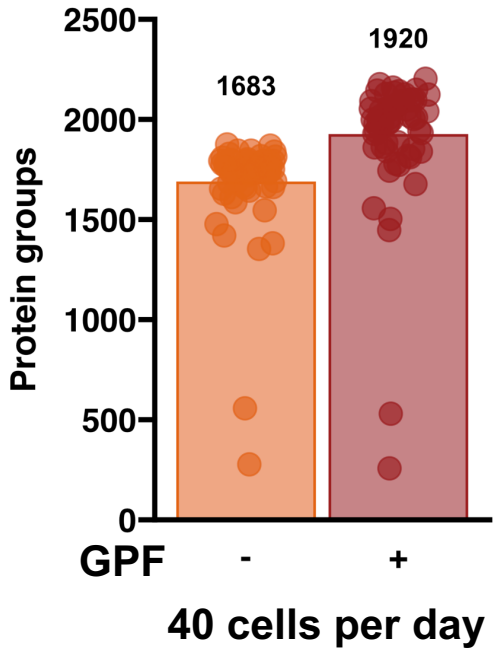
Resolution		Window		HRMS1
MS1	MS2	Size	Number	Cycle-time
120K	120K	40mz	15	1.536s
240K	240K	100mz	6	1.280s

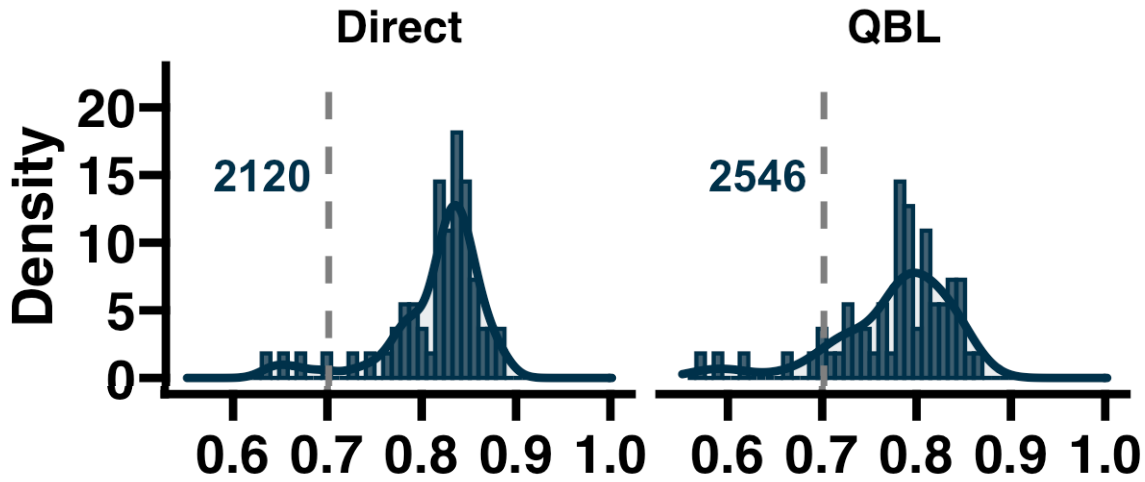


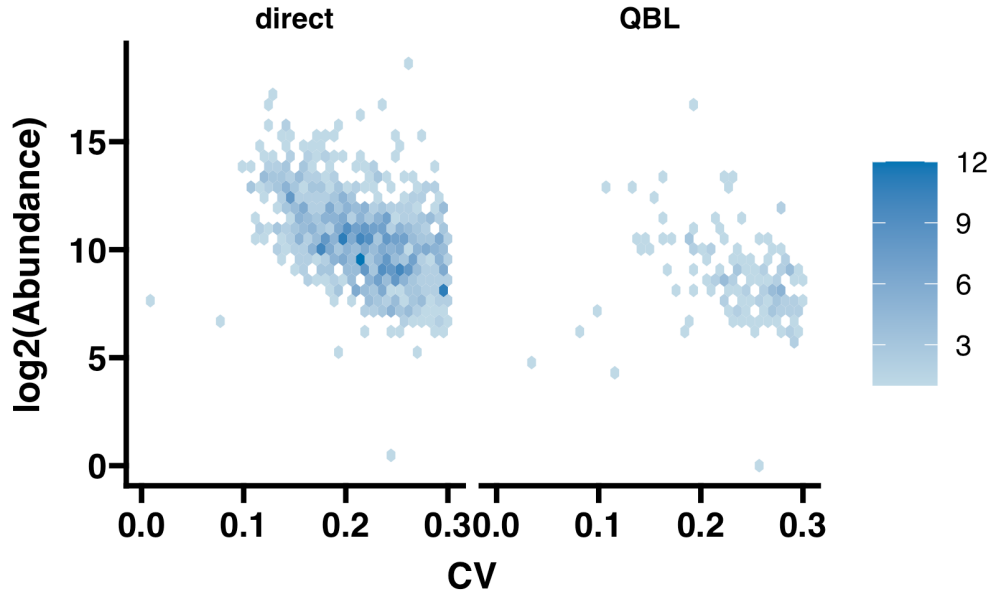
Single-cell DIA method optimization



Showcasing the established workflow for single cell analysis

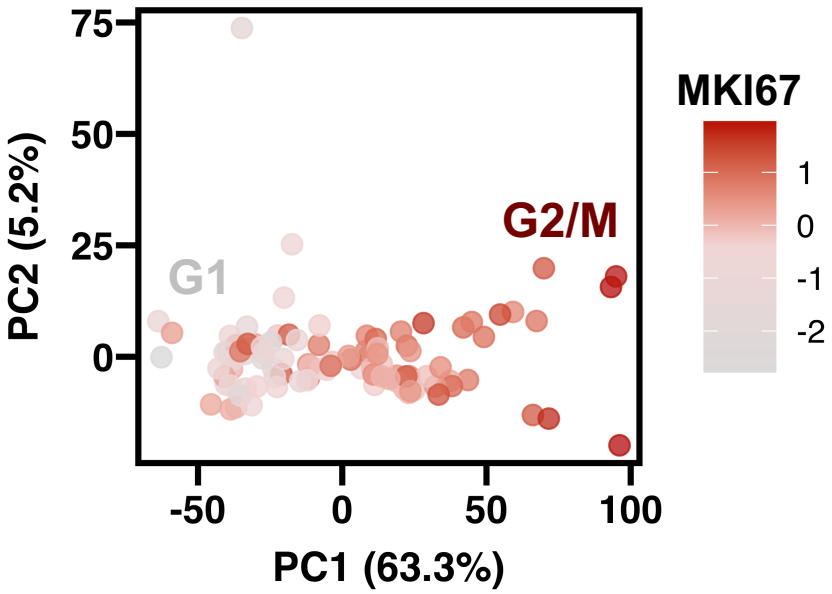


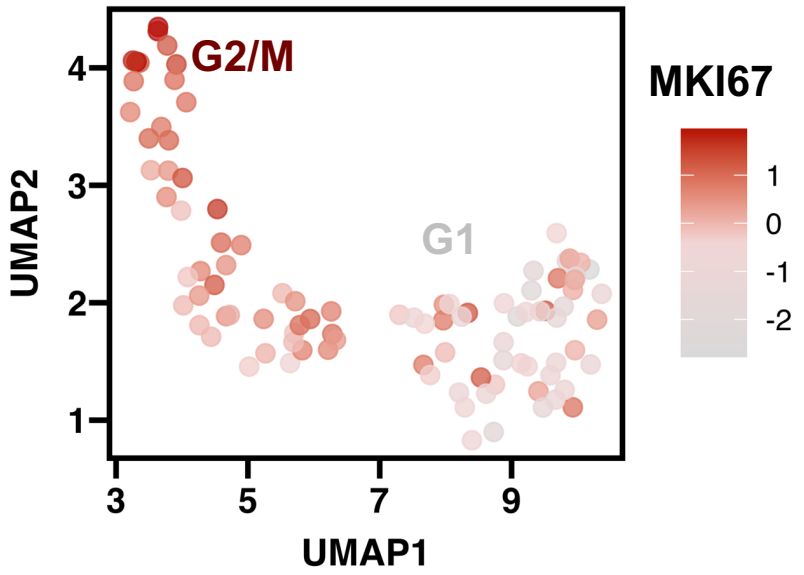


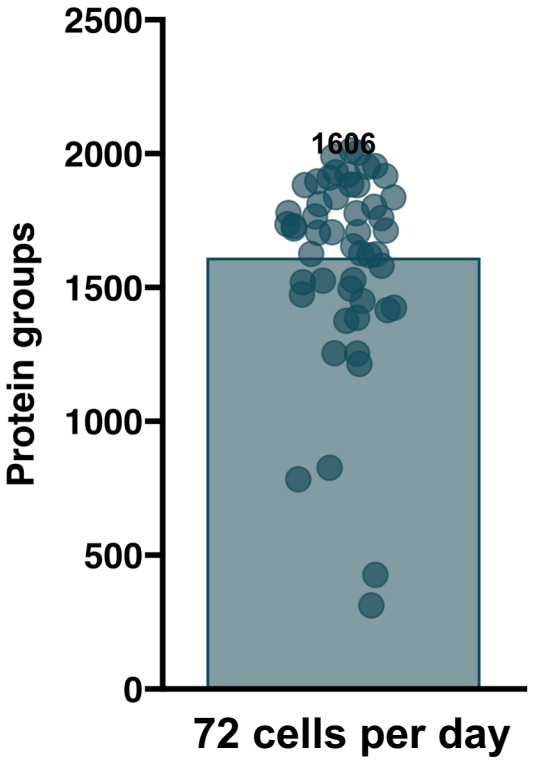


Single-cells were used as input for GPF

Showcasing the established workflow for single cell analysis



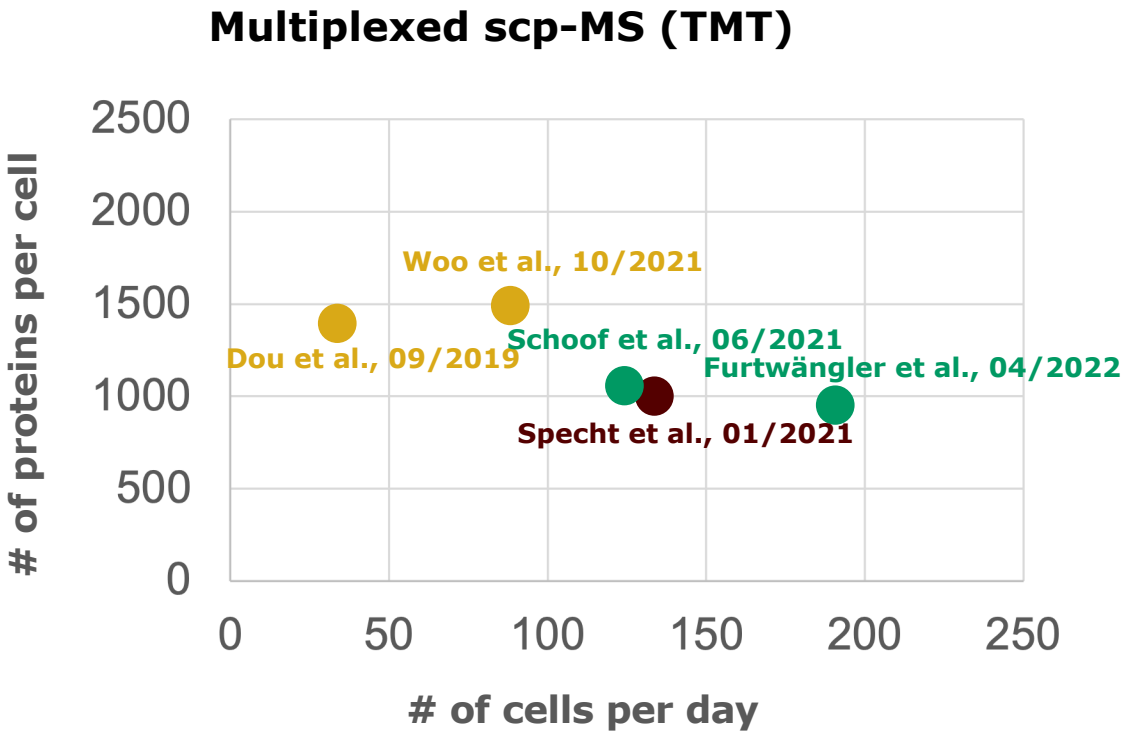


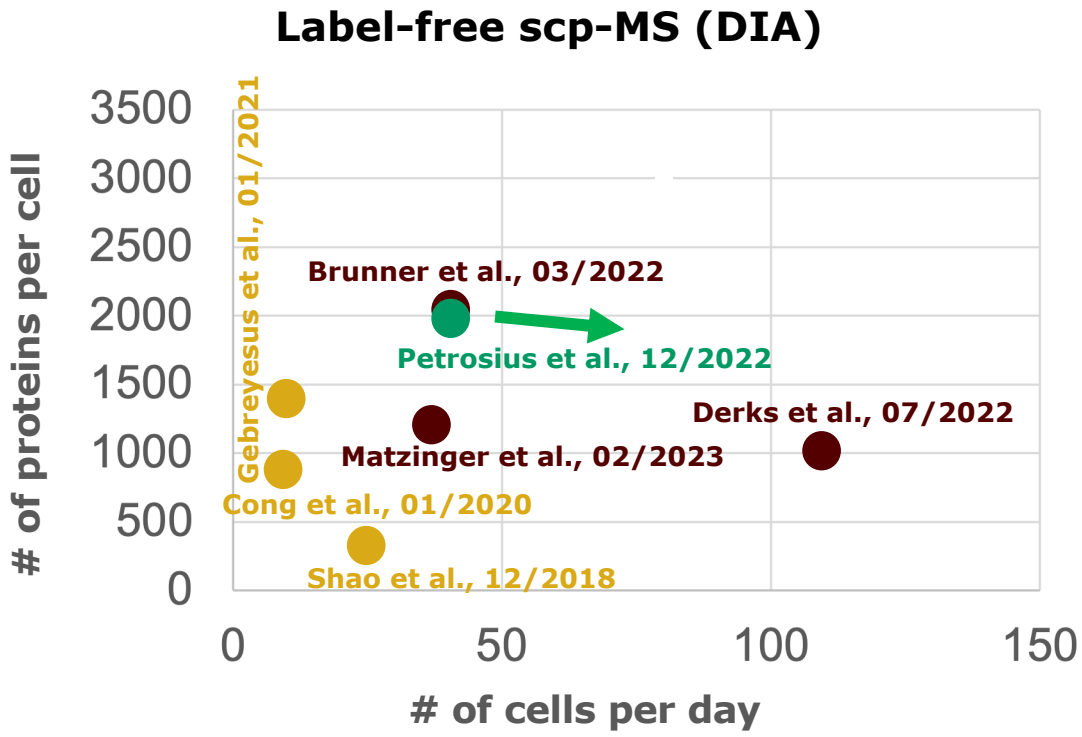


How do we compare to the state-of-the-art?

Low-throughput microfluidics

Our lab contributions





Conclusions

- **DIA method design should be adjusted accordingly to sample load for optimal performance.**
  - Wider windows and higher resolutions is favorable for low input samples
- **Precursor level quantification can be used to enhance analytical sensitivity**
  - WISH-DIA is quantification accuracy similar to standard DIA
- **GPF libraries are a promising alternative to high pH approaches for samples where fractionation is prohibitive**
- **Workflow uses standardized lab equipment and does not require single-cell proteomics designated liquid handling systems**
  - The same proteomic depth should be achievable on on hybrid instruments (e.g. Exploris 480)

Enhancing single-cell proteomics through tailored Data-Independent Acquisition and micropillar array-based chromatography

Valdemaras Petrosius, Pedro Aragon-Fernandez, Nil Üresin, Teeradon Phlairaharn, Benjamin Furtwängler, Jeff op de Beeck, Simon Francis Thomsen, Ulrich auf dem Keller, Bo T. Porse, Erwin M. Schoof

doi: https://doi.org/10.1101/2022.11.29.518366

Acknowledgments

Cell Diversity Lab

Erwin M. Schoof
Pedro Aragon
Susan Ibi Preus
Rune Daucke
Jakob Woessmann

Porse Lab

Bo Porse
Nil Üresin
Benjamin Furtwängler
Rajesh Somasundaram

DTU Proteomics Core

Marie Vestergaard Lukassen
Maike Wennekers Nielsen
Rikke Nord Johansen
Lene Holberg Blicher

ThermoFisher

Jeff Op de Beeck
Robert van Ling

Biognosys

Tejas Gandhi

DTU Technical University of Denmark

Spectronaut™ powered by Pulsar

UNIVERSITY OF COPENHAGEN

LEO

Kræftens Bekæmpelse

BIOGNOSYS

ThermoFisher SCIENTIFIC


DTU