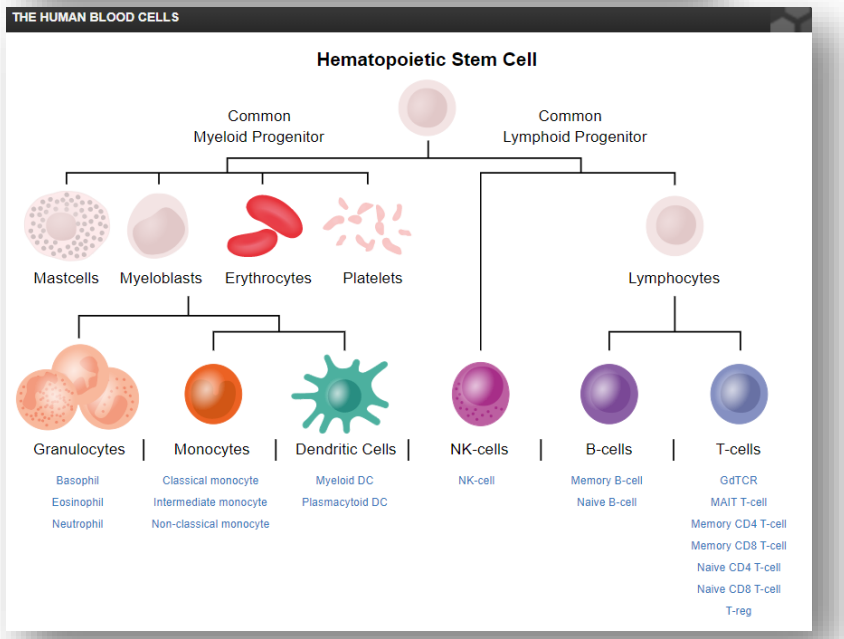
Precise Quantitative Serum LC-MS/MS Profiling: The Impact of Sample Preparation and Sample Source on Biomarker Discovery Studies

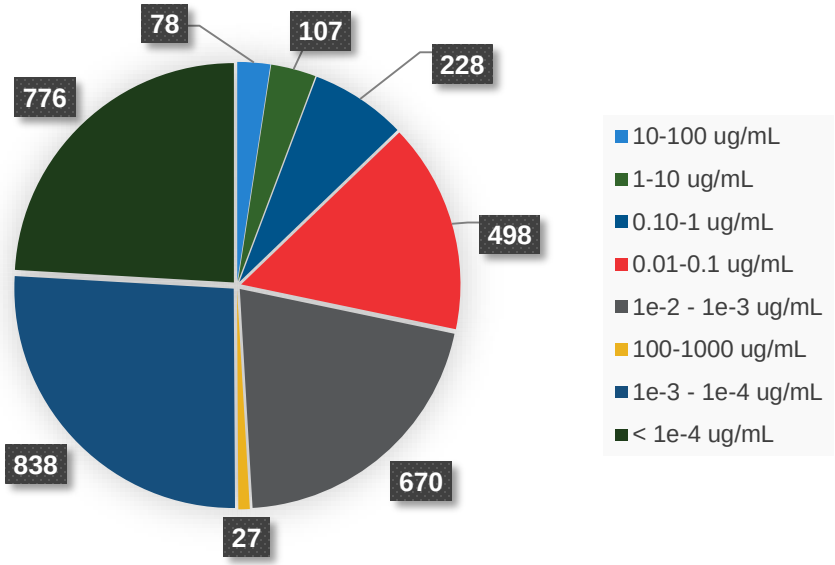
Alexander Boychenko[1], Natalia Govorukhina[2], Runsheng Zheng[1]
[1]Thermo Fisher Scientific, Germering, Germany
[2]University of Groningen, Groningen, The Netherlands

Biomarkers in blood

Blood Complexity

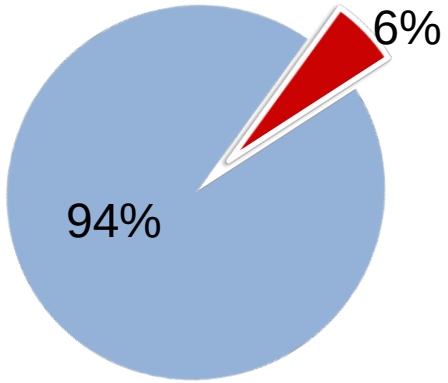


Protein concentrations in blood



https://www.proteinatlas.org/humanproteome/blood

Protein abundances in blood

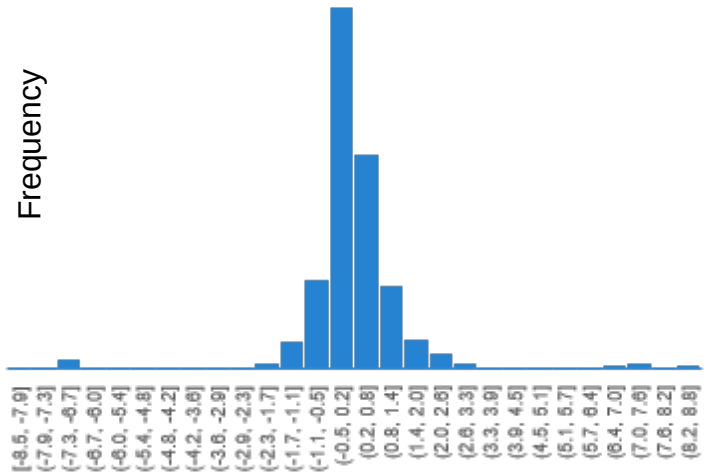


> 3000 proteins identified with MS

> 94 % of total protein amount in serum is occupied by 14 proteins

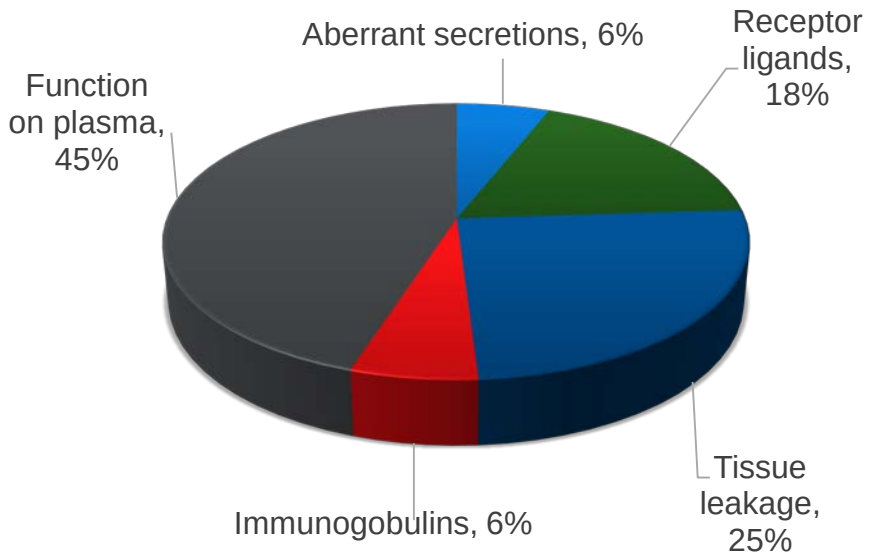
Protein Biomarkers in Blood

Natural biological variation of protein concentration



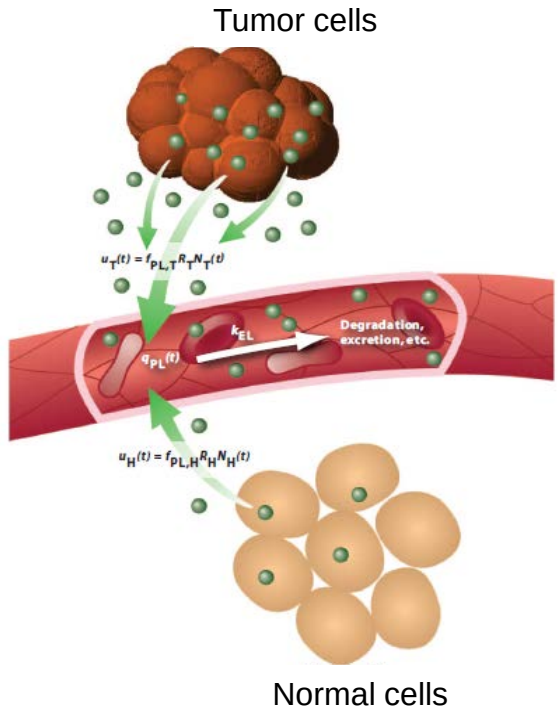
Log2 (Protein Abundance)

FDA Approved Biomarkers (mostly nonspecific)



N. Leigh Anderson, Clinical Chemistry 56:2 177–185 (2010)

Specific biomarkers



S.S. Hori, Sci. Transl. Med. 3 109-116 (2011)

> 80% proteins have less than 2-fold abundance variation

51% - normal functions of blood
25% - tissue leakage
18% - receptor ligands

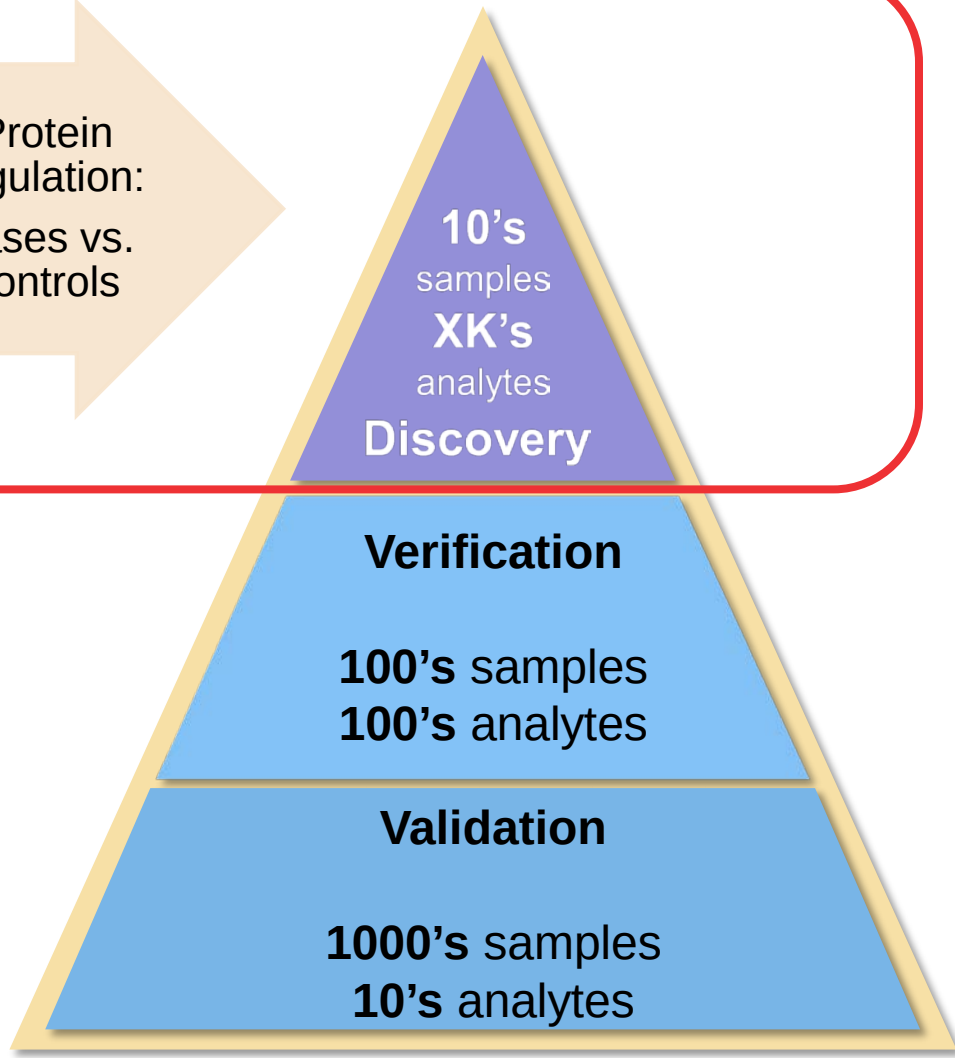
6% - tumor markers

ThermoFisher SCIENTIFIC

Biomarker Discovery

Experimental Design: “Triangular” Biomarker Discovery and Validation Pipeline

1. Analytical Variability
2. Biological Variability
3. Protein regulation: cases vs. controls



- Lower analytical variability gives more confidence in discovered protein abundance changes in cases and controls
- Reliably defined biological variability is the key to unbiased selection of biomarker candidates
- Deeper proteome profiling increases chance of discovering specific biomarkers

“The Devil is in the Detail”: Multicentric Study and Sample Preparation

Lund (SE) – 1 group (Healthy)

Serum tube: BD 367615
30 min - 2 hours clotting; +20 ºC
2,000g, 10 min; RT

UMCG (NL) – 2 groups (CIN0, Healthy)

Serum tube: BD 367953
2 - 8 hours clotting; +20 ºC
3,000g, 10 min; RT

SeraLab (UK) – 1 group (Healthy)

Bleed bag unit (not specified)
O/N clotting; T +4 ºC
2,800g; 20min; 5 ºC

How reliable is your analytical and data processing pipeline?
Can you identify biomarkers in samples from healthy subjects?

Fast, Deep, Quantitative Proteomics Platform

Thermo Scientific™ UltiMate™ 3000 RSLCnano system

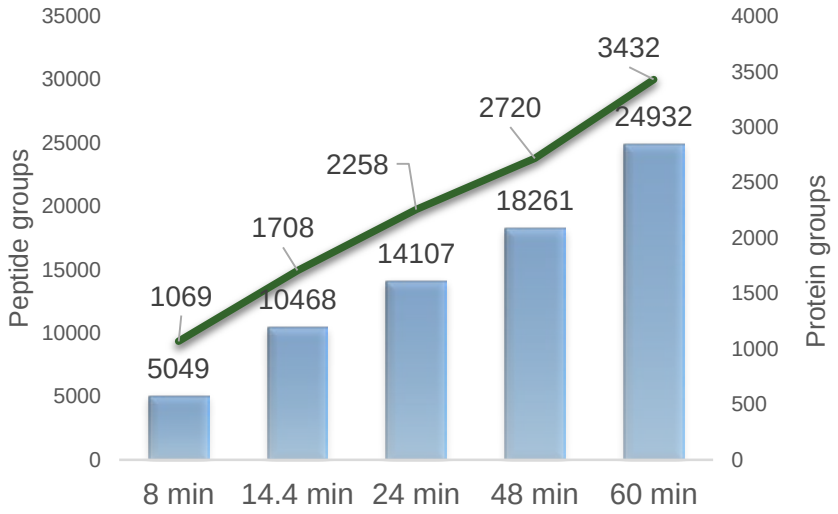
75 µm x 150 mm, 3 µm PN ES800

Thermo Scientific™ Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap™ mass-spectrometer

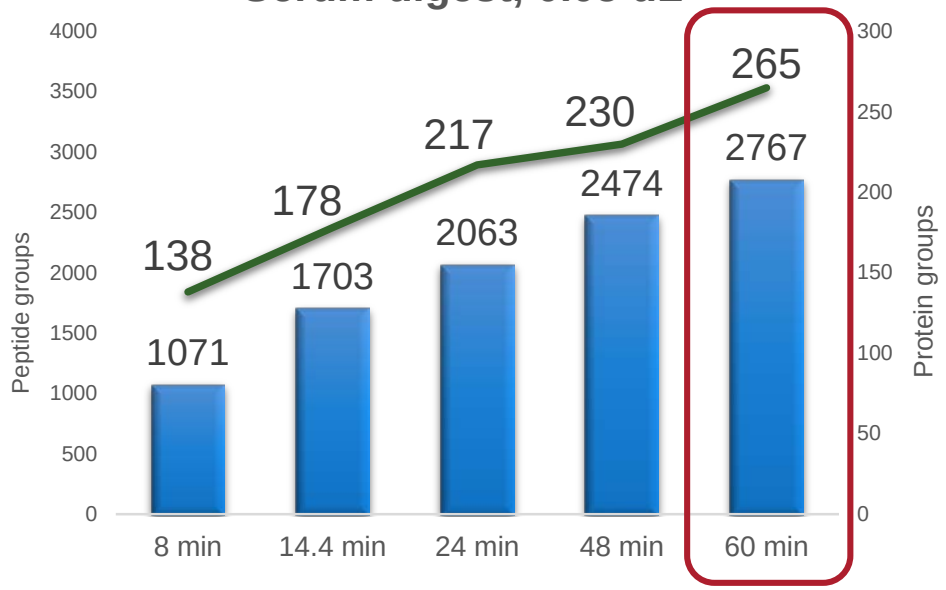
Standardized low-flow LCMS methods

	Flow, µL/min	Samples per 24 hours	MS utilization, %	Average PWHM, sec	Average PW base, sec	Asymmetry
60 min	0.300	24	95	10	19	1.23
48 min	0.600	30	90	9	18	1.21
24 min	0.800	60	87	7	13	1.17
14.4 min	1.000	100	85	4	7	1.13
8 min	1.500	180	75	3	6	1.16

HeLa digest, 200 ng



Serum digest, 0.05 uL



60 min low-flow LCMS method was used for analysis of 15 crude serum samples

For Research Use Only. Not for diagnostic purposes.

ThermoFisher SCIENTIFIC

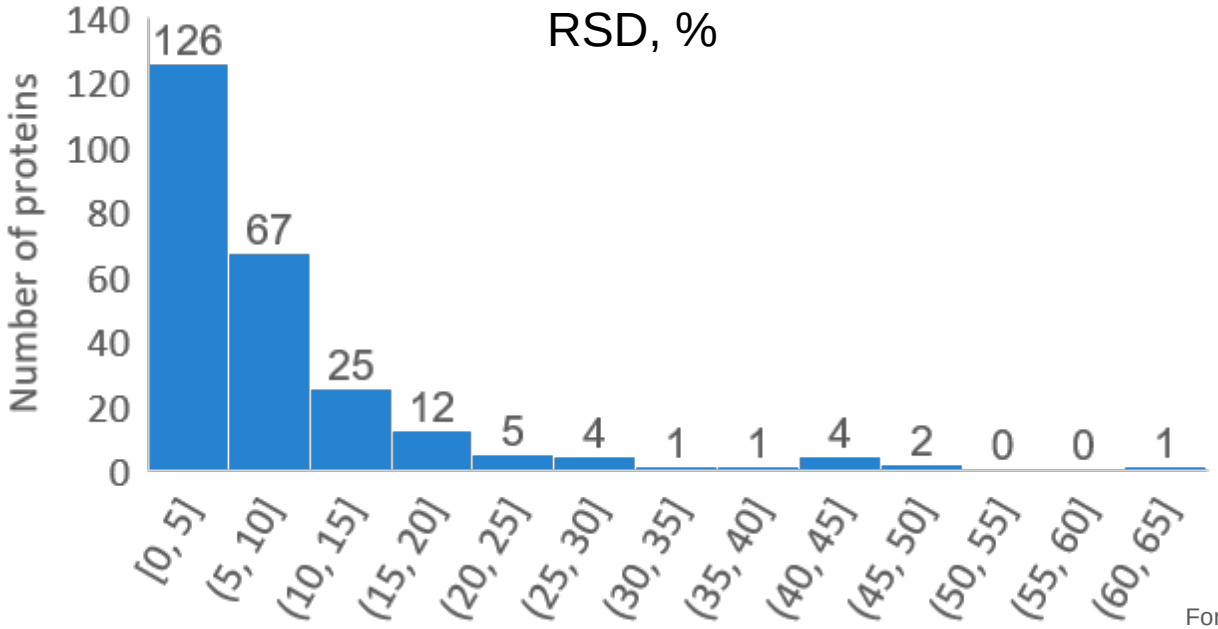
Analytical Variability: Precision of Low-Flow LCMS Analysis

1 Analytical Variability

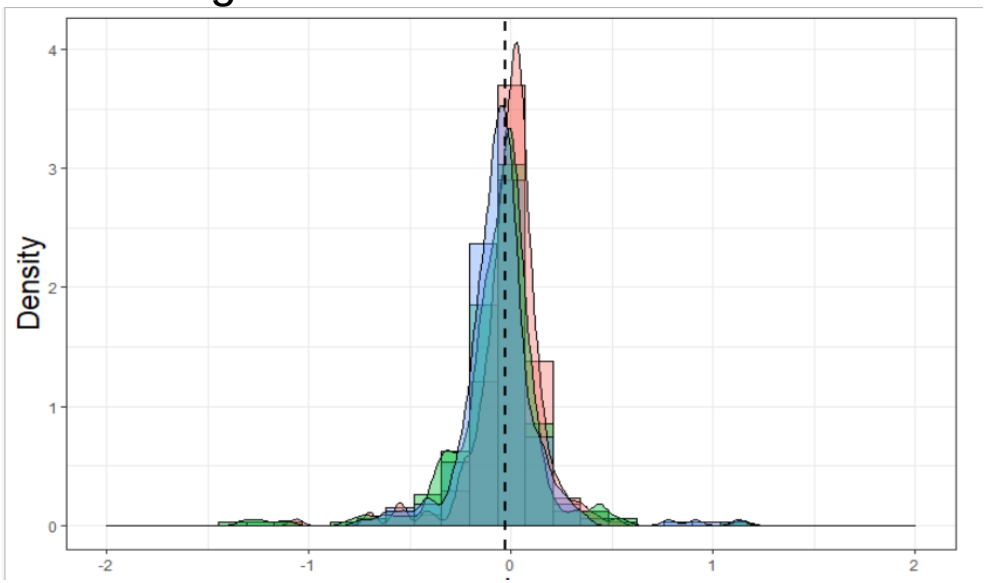
- 3 LCMS replicates per sample
- Label-free quantification

Proteins identified (1% FDR): 265
Proteins quantified: 248
- RSD < 15% 218 (88%)
- 95% of proteins have abundance variation < 50%

RSD, %



log2 Protein abundance ratios



Experimental Design: Biological Variability

1 Analytical Variability

2 Biological Variability

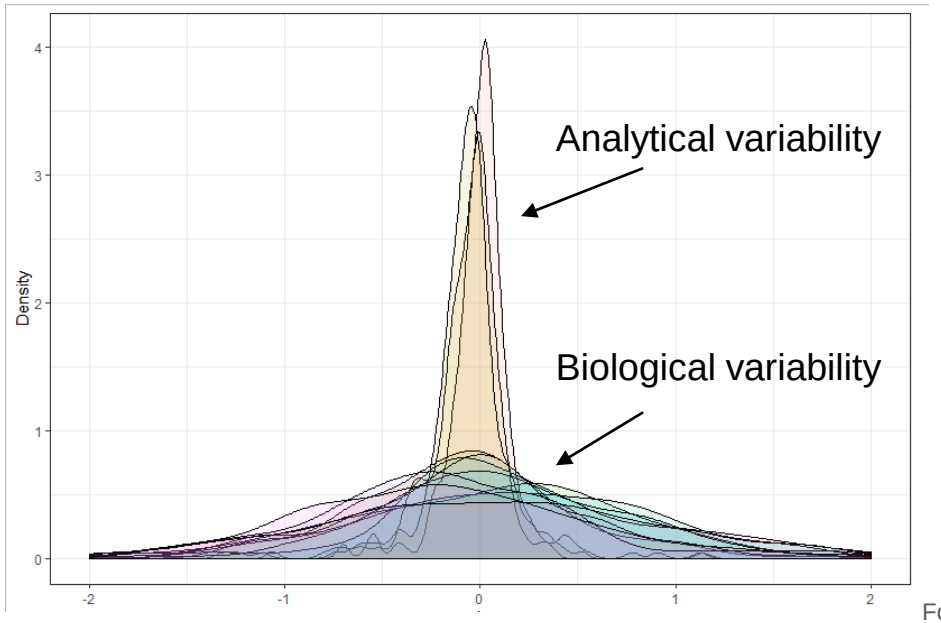
CIN0

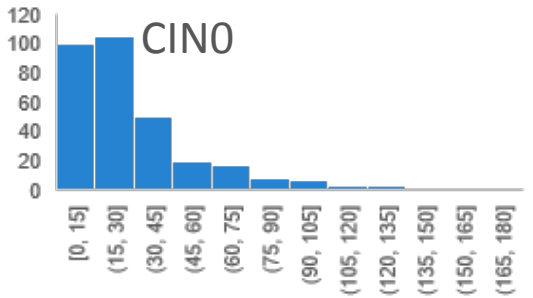
Seralab

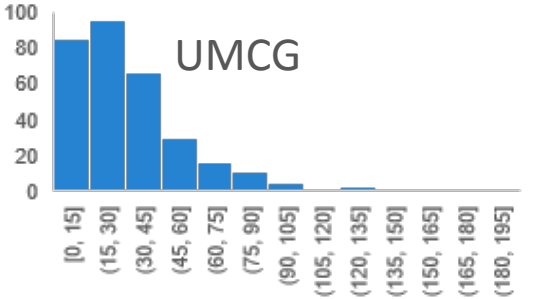
UMCG

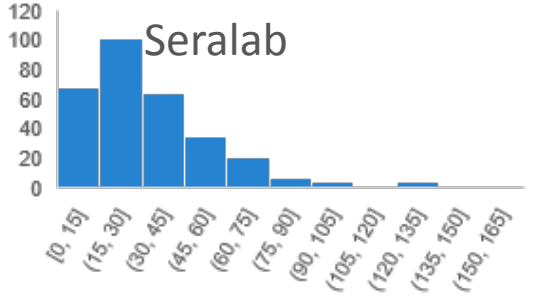
Sweden

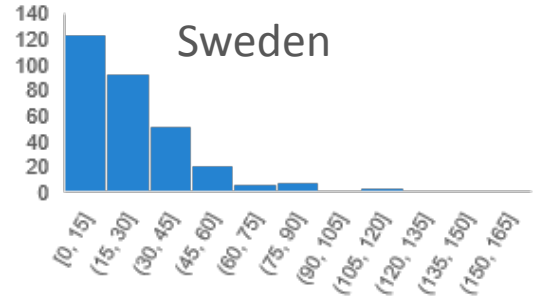
- Biological variability is significantly larger in comparison to analytical variability
- Still 80% of proteins have less than a 2-fold change











Experimental Design: Regulated Proteins

1. Analytical Variability
2. Biological Variability
3. Proteins regulation: cases vs. controls

- 94 proteins had at least one abundance ratio > 2
- 22 proteins had abundance ratio > 4
- 16 of them can be mapped to blood atlas

Data Analysis pipeline

Seralab CIN | UMCG CIN | Sweden CIN | UMCG Seralab | Sweden Seralab | Sweden UMCG

↓

Protein ratio distribution, selection criteria

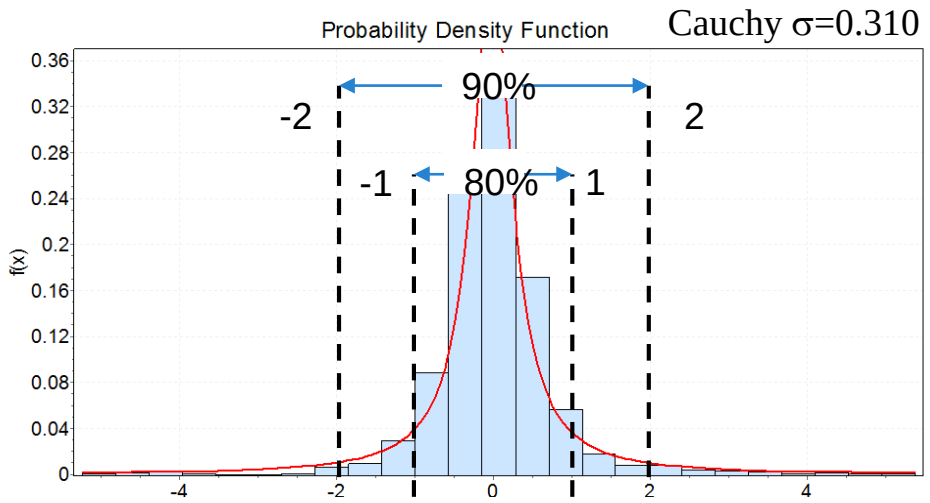
↓

UP and DOWN regulated proteins

↓

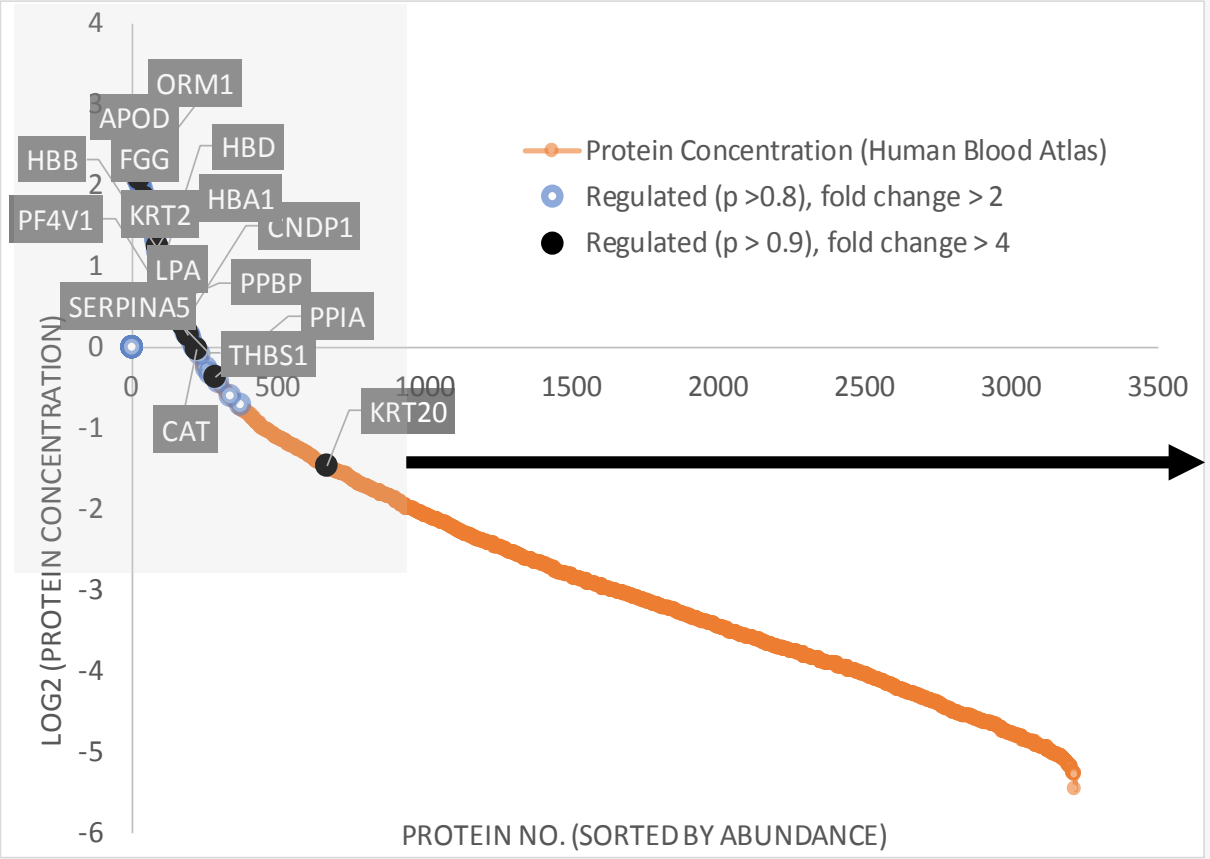
Reliability of quantification, Biological functions

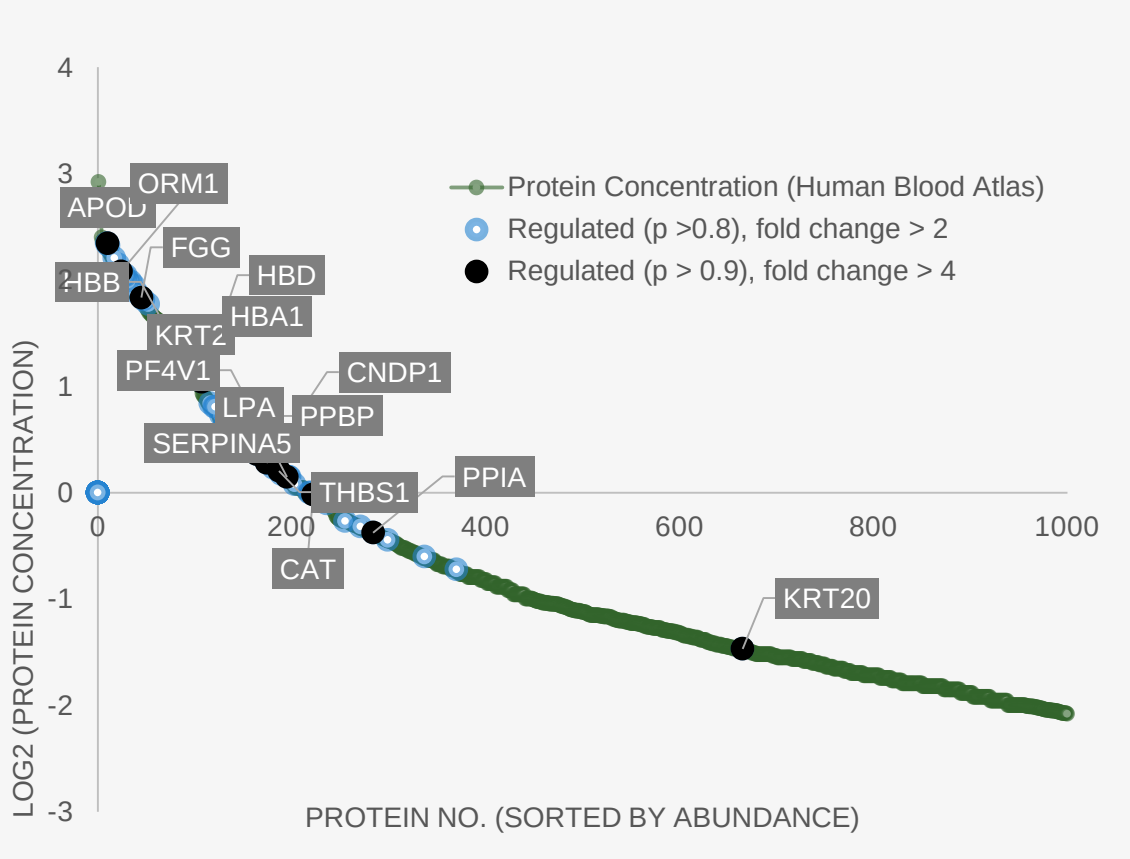
Selection criteria



Boichenko et al. J Proteome Res. 2014 Nov 7;13(11):4995-5007

Why Did We Find Proteins Regulated in Healthy Groups?

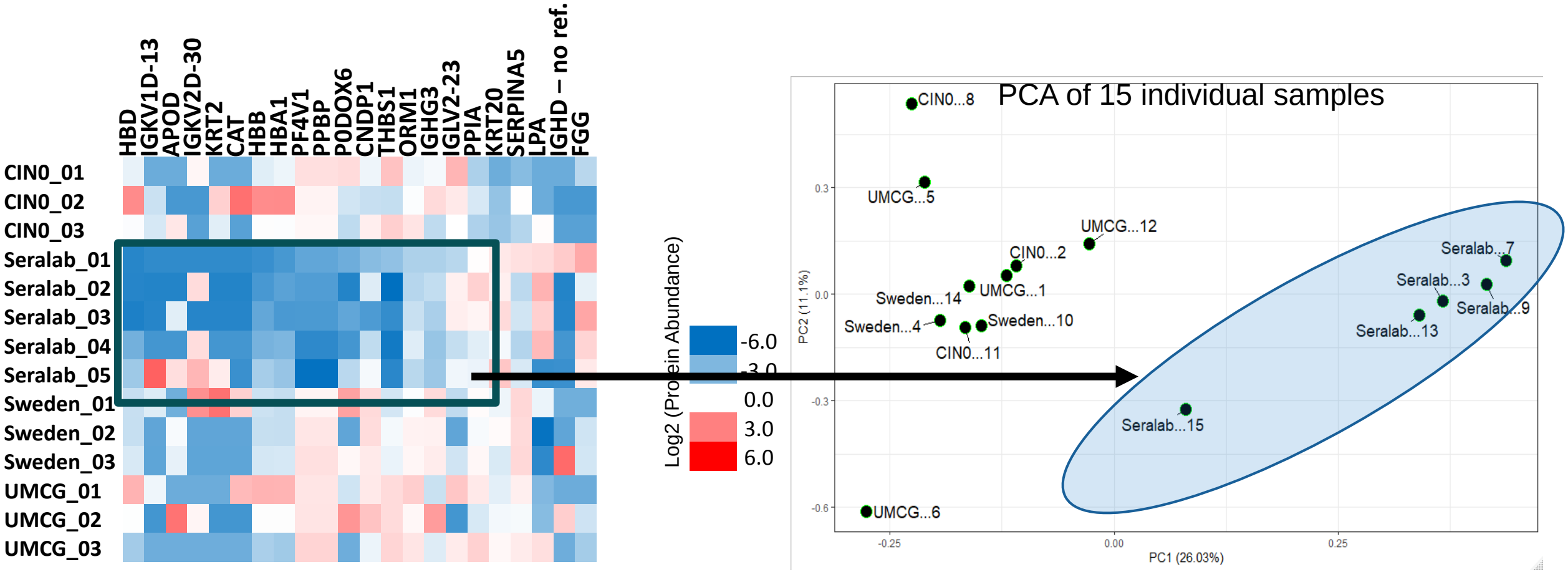




Significant regulation of classical blood proteins in healthy groups is very unlikely and can lead to

- Artificial protein abundance alteration in one or multiple individual samples (storage, analysis error, etc.)
- Artefacts arising from sample preparation

Protein Regulation in Individual Samples



Proteins downregulated in Seralab samples represent classical blood proteins expressed in specific blood cells released during the serum preparation process
A deeper dive into the proteome is required to find specific biomarkers

Deep Profiling of Individual Samples

Deeper Proteome Profiling with Comprehensive Online-2D Low-Flow LC-MS/MS

UltiMate 3000 RSLCnano Online 2D configuration

2nd Dimension

75 µm x 150 mm,
3 µm
PN ES800

1st Dimension

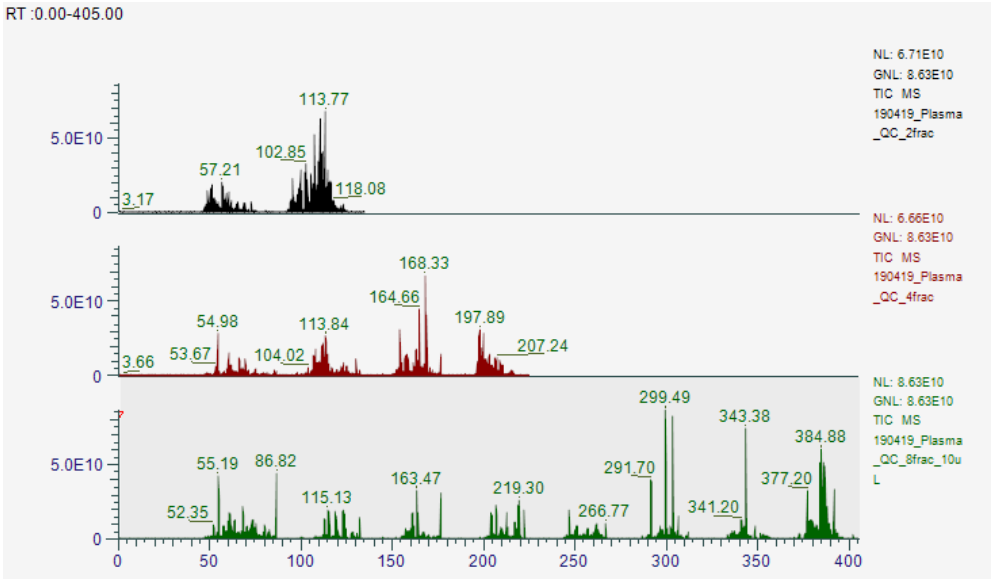
PepSwift, Monolithic
Column 100 µm x 250 mm

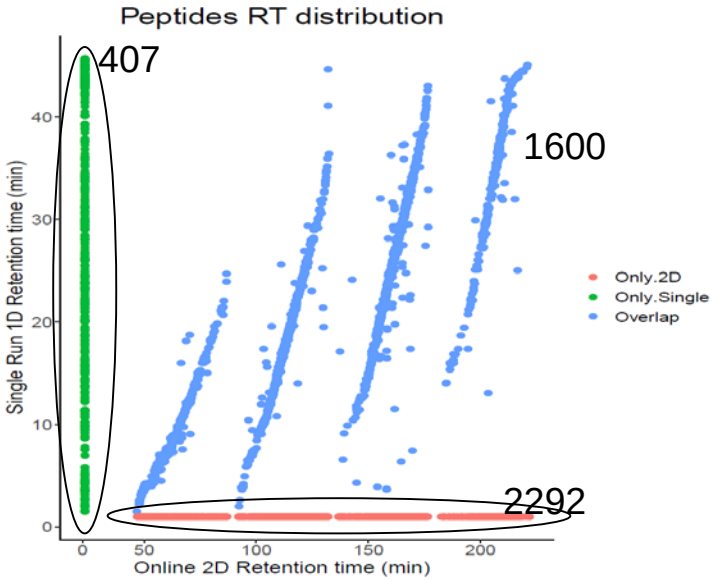
Thermo Scientific™ Orbitrap Exploris™ 480

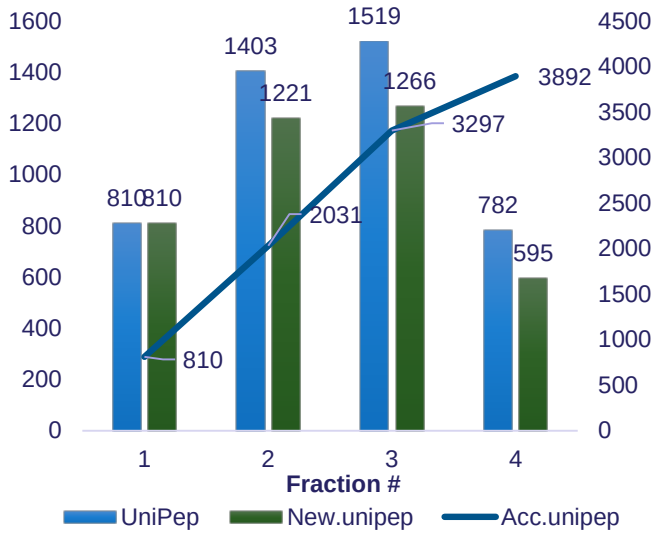
2 fractions (135 min)

4 fractions (225 min)

8 fractions (405 min)



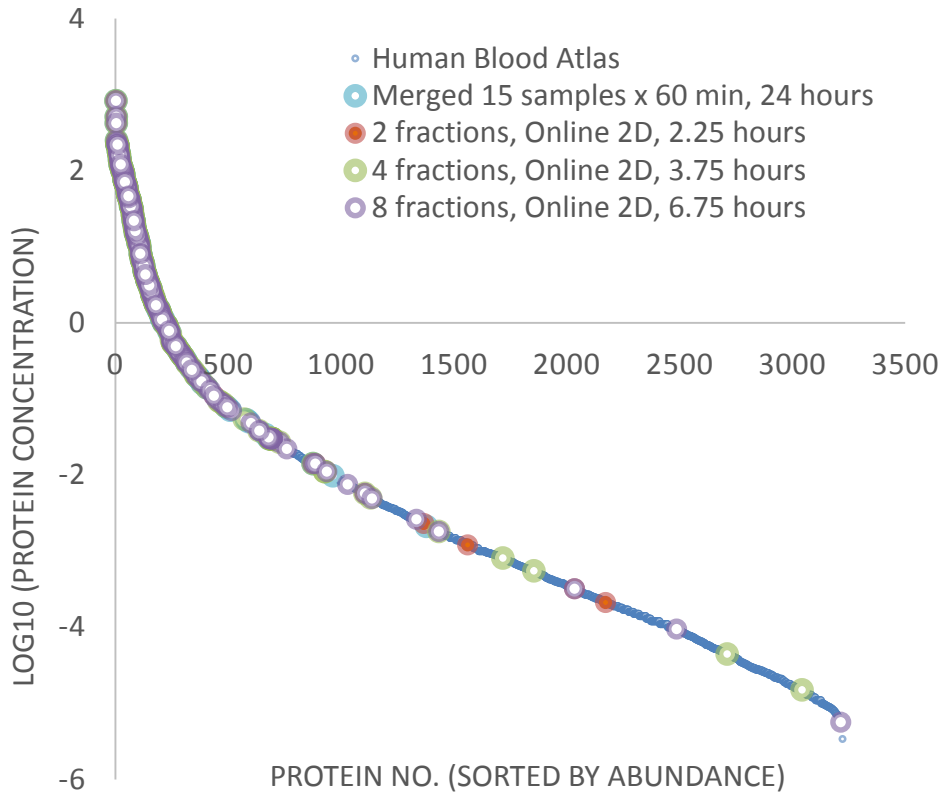


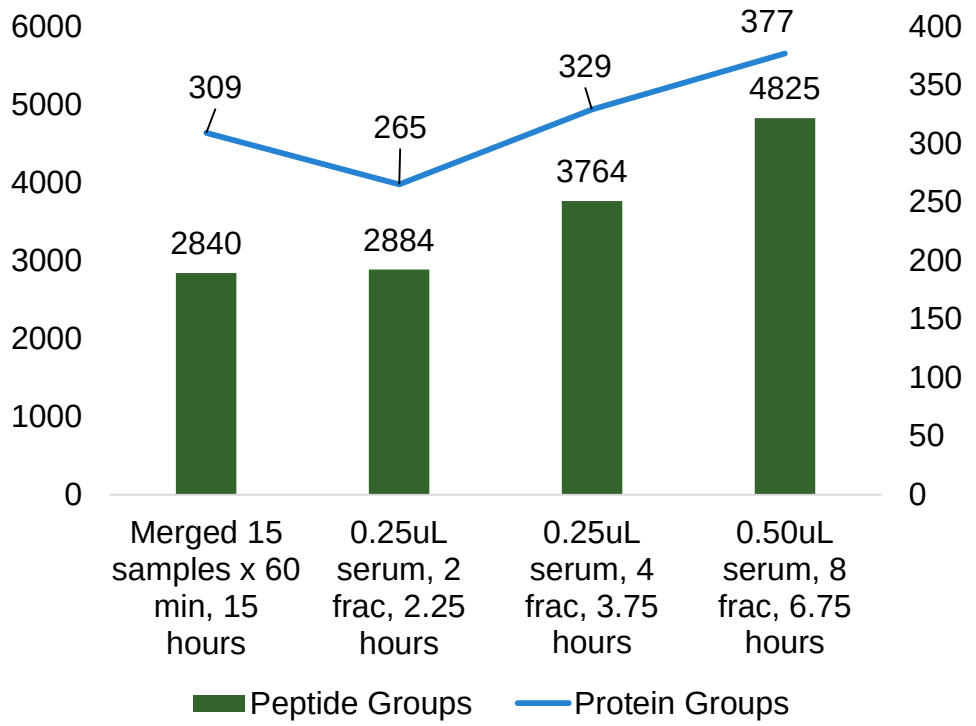


- Automated sample analysis (no manual sample handling)
- Orthogonal selectivity in 1st and 2nd dimension
- Low overlap of peptides elution between fractions
- Deeper complex proteome profiling

For Research Use Only. Not for diagnostic purposes.

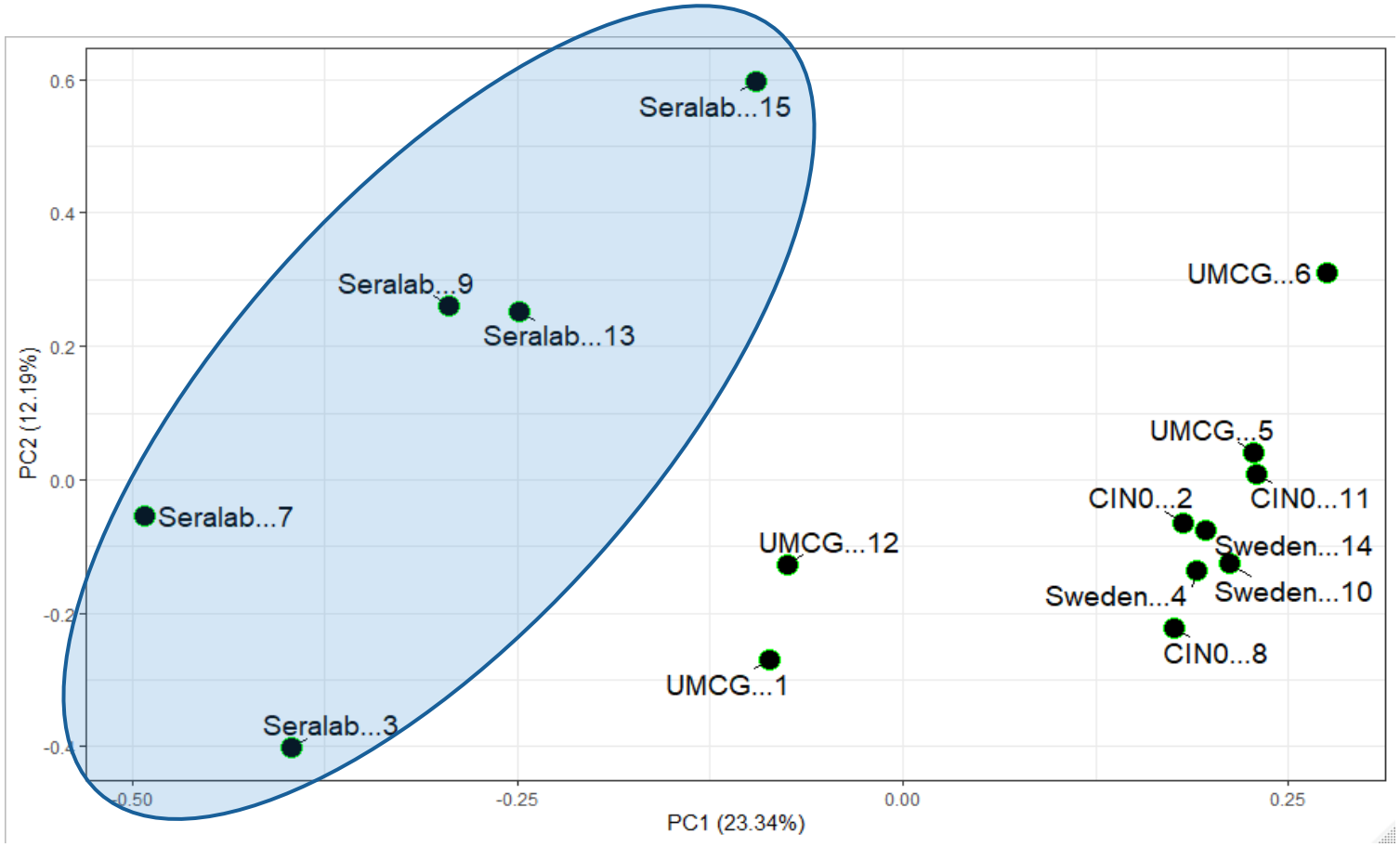
Increased Dynamic Range with Online 2D Low-Flow LC-MS/MS





- Higher loading and peak capacity of online 2D low-flow LC-MS/MS analysis results in deeper serum proteome coverage
- Protein and peptide identifications increase linearly with number of fractions
- Automated on-line 2D LC-MS/MS analysis can be easily adjusted to any number of fractions and combined with DDA and DIA MS acquisition techniques

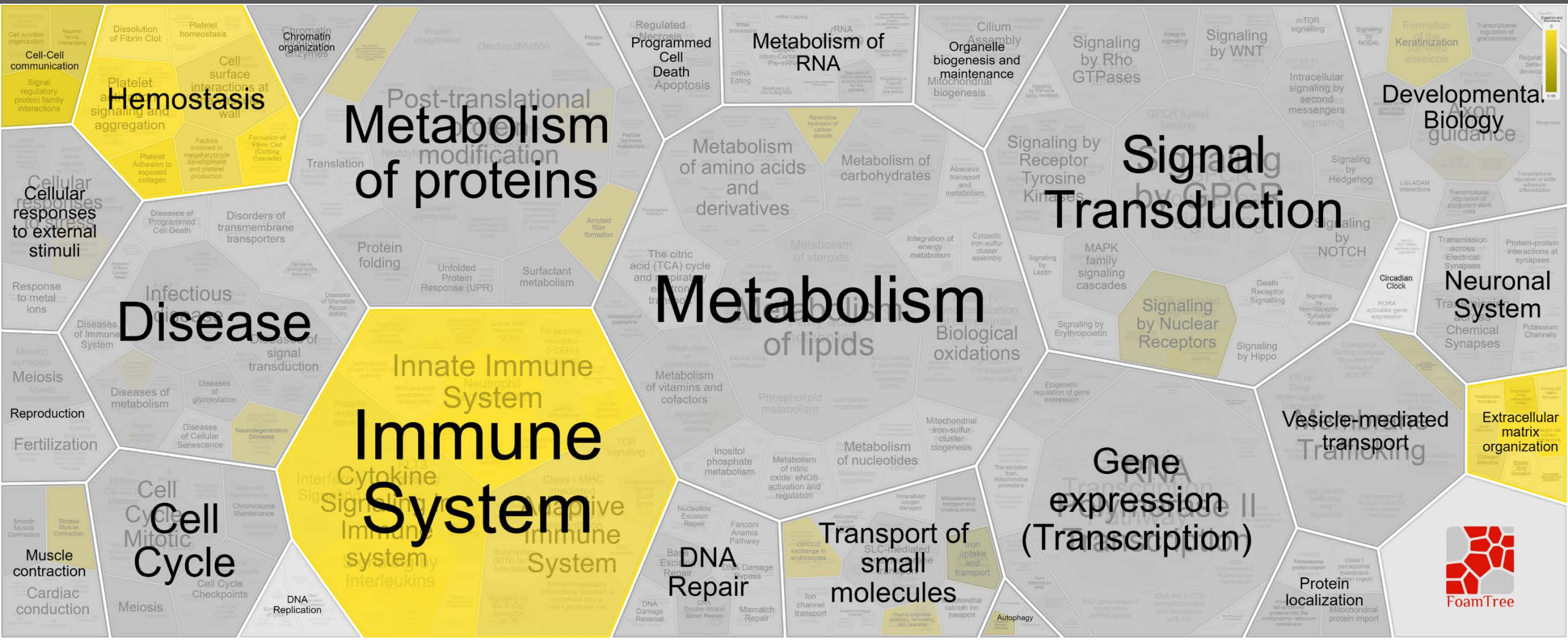
From Groups to Individual Samples: Focus on Precision Medicine



- 15 individual samples were analysed with low-flow online 2D LC-MS/MS
- 442 proteins were quantified based on 3536 unique peptides
- 110 protein ratio combinations were computed for 15 samples
- 237 proteins had at least one abundance ratio > 2
- 151 of these proteins were mapped to the Human Blood Atlas

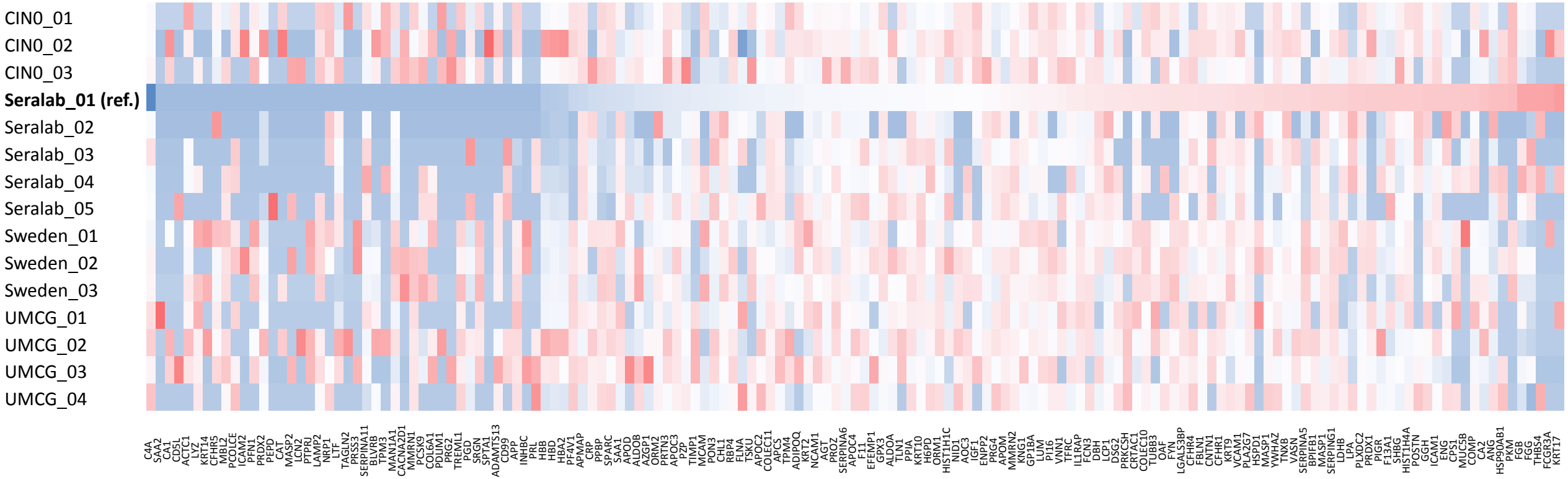
- PCA shows clear separation of samples obtained from Seralab
- There is no clear differentiation between other sample groups

Pathway Mapping for Serum Regulated Proteins https://reactome.org/



Most UP and DOWN regulated proteins in individual samples were related to clotting or generic immune system response

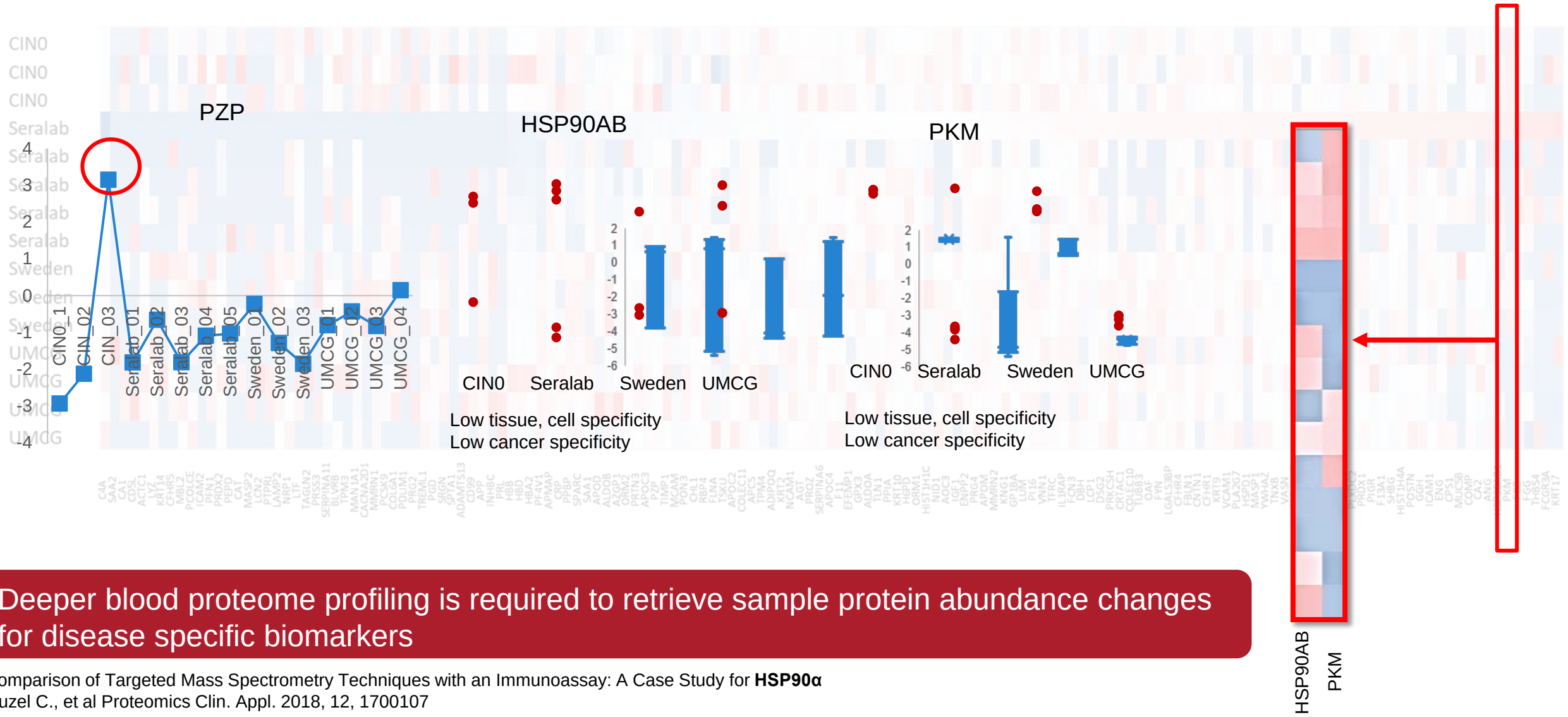
Protein Regulation in Individual Samples



The analysis of individual samples is complicated by a broad variation in expression across healthy individuals, sample preparation, analytical variability

A number of proteins in Seralab samples show significant abundance alternations in comparison with other groups

Significant Sample Specific Regulations: Potential Biomarkers



Deeper blood proteome profiling is required to retrieve sample protein abundance changes for disease specific biomarkers

Comparison of Targeted Mass Spectrometry Techniques with an Immunoassay: A Case Study for **HSP90α**
Guzel C., et al Proteomics Clin. Appl. 2018, 12, 1700107

ThermoFisher SCIENTIFIC

Conclusions

- Standardized high-throughput low-flow LC-MS methods can be used for routine profiling of TOP 200 proteins or development targeted assays with high sensitivity
- Standardized sample preparation of biofluids and critical analysis of potential validation targets is essential for biomarkers discovery studies.
- Profiling of high abundant blood (serum, plasma) proteins has limited benefit for the discovery of new biomarkers or population related changes due to high stability of the blood proteome and non-specific changes related to the immune system response
- Comprehensive high pH RP x low pH RP low-flow online 2D LC-MS/MS analysis shows great potential for deep, automated quantitative blood proteome profiling required to discover cases specific biomarkers

Acknowledgements

University of Groningen

RUG, Groningen, NL

R. Bischoff

N. Govorukhina

UMCG, Groningen NL

A. van der Zee

K. ten Hoor

H. Klip

Erasmus MC

Erasmus MC, Rotterdam

Dr. T. Luider

C. Guzel

Germering, Germany
Bremen, Germany

T. Arrey

P. Jehle

W. Decrop

C. Pynn

R. Zheng

ThermoFisher SCIENTIFIC