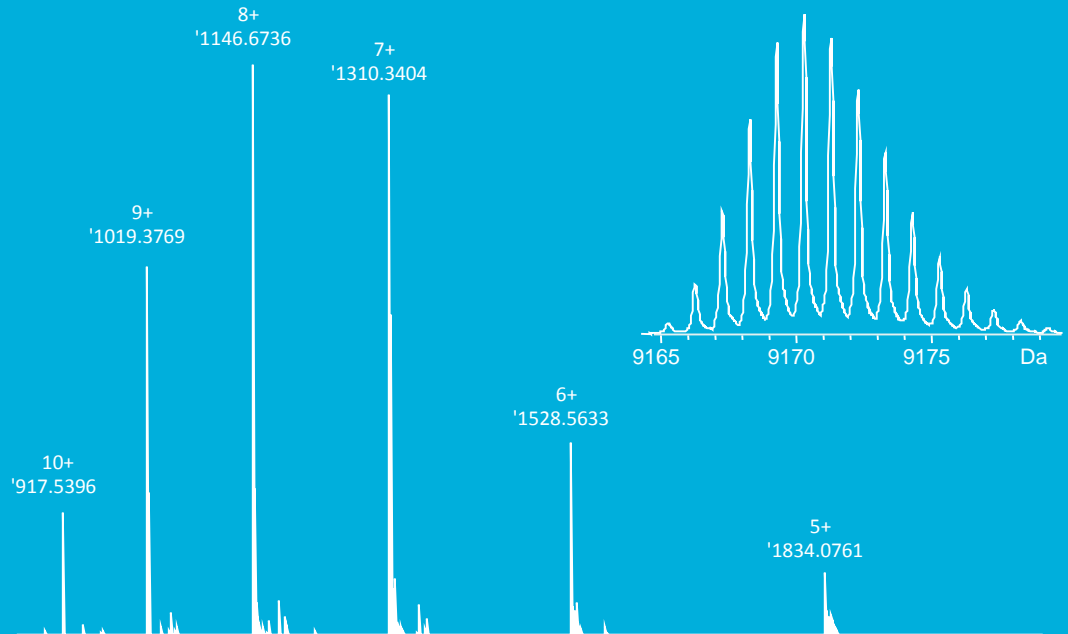# Clinical Glycoproteomics in Research and Diagnostics

X-Omics festival 28-09-2020
Hans JCT Wessels, PhD

Translational Metabolic Laboratory
Department of Laboratory Medicine
Radboudumc
Geert Grooteplein-zuid 10
6525GA Nijmegen,The Netherlands
Hans.wessels@radboudumc.nl

X-OMICS.nl

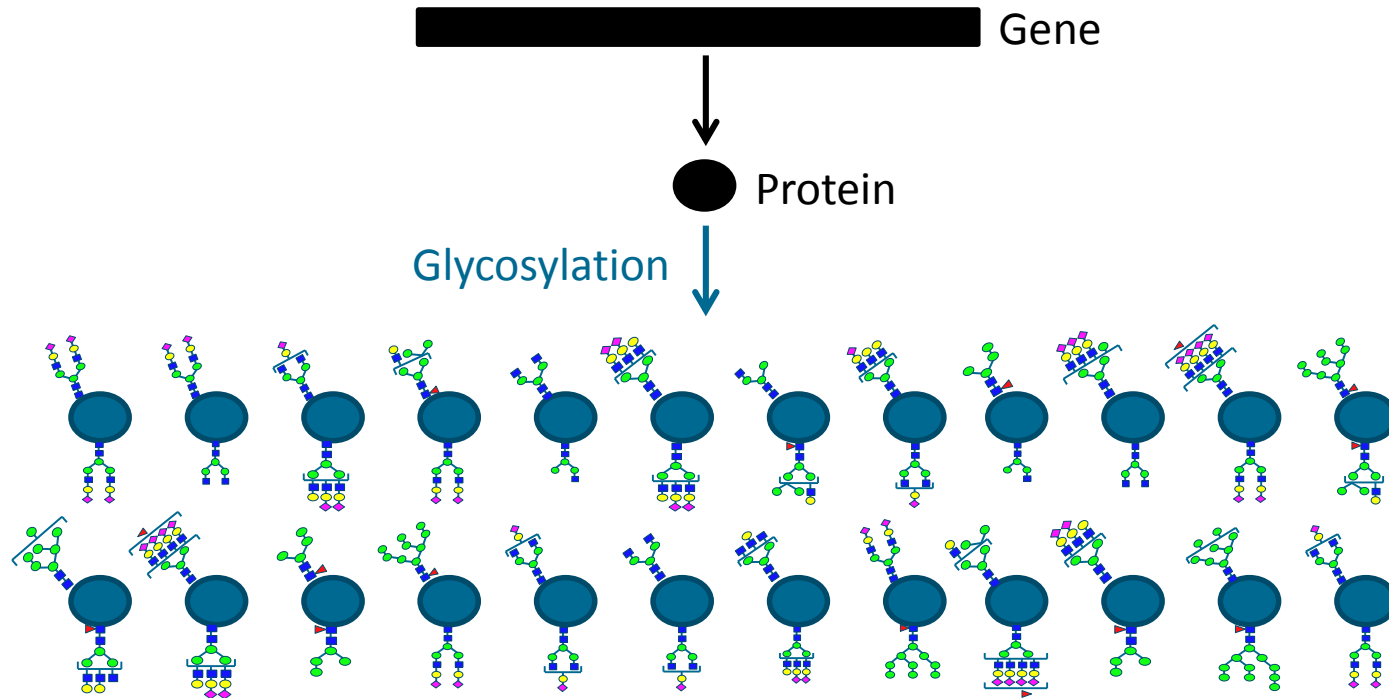Radboud Consortium for Glycoscience

Radboud umc

Protein Glycosylation

More than 90% of all proteins are glycosylated in plasma
Major contributor to proteoform diversity
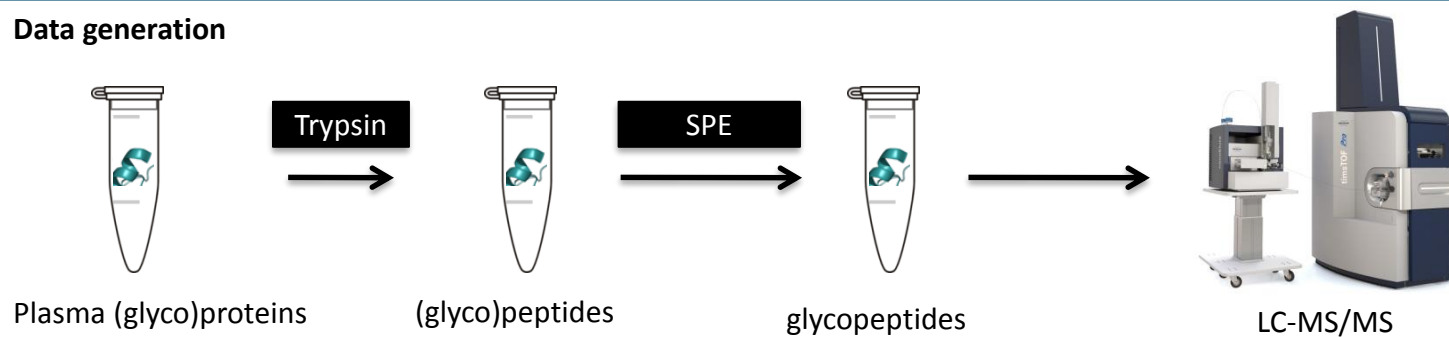Glycosylation is a key modulator of protein biology
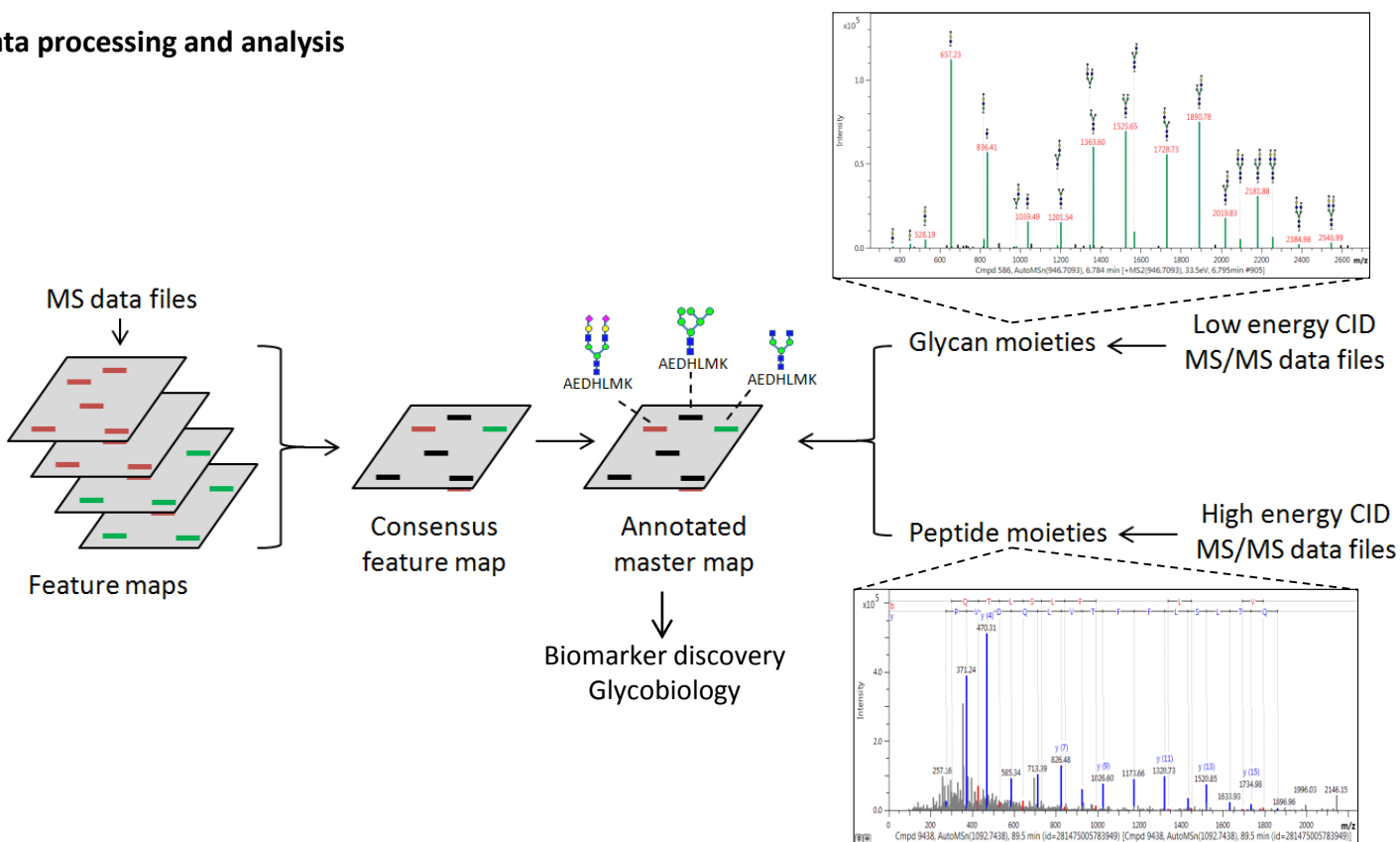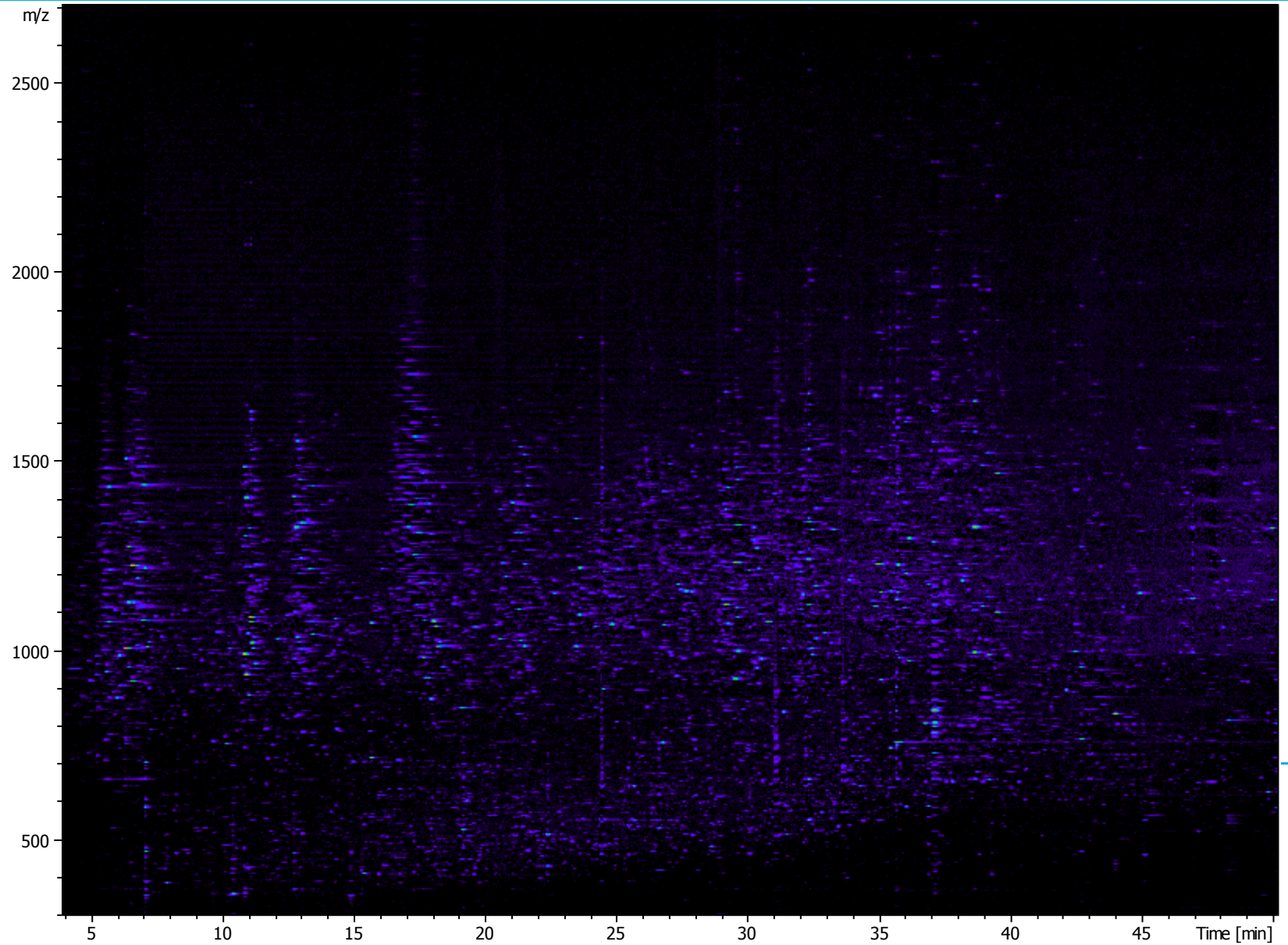Aberrant glycosylation is observed in many (common) human diseases

Gene

Protein

Glycosylation

Glycopeptide profiling LC-MS/MS workflow

Data generation

Plasma (glyco)proteins → Trypsin → (glyco)peptides → SPE → glycopeptides → LC-MS/MS

Data processing and analysis

MS data files

Feature maps → Consensus feature map → Annotated master map

AEDHLMK   AEDHLMK   AEDHLMK

Glycan moieties ← Low energy CID MS/MS data files

Peptide moieties ← High energy CID MS/MS data files

↓
Biomarker discovery
Glycobiology

Raw LC-MS data

Raw LC-MS data
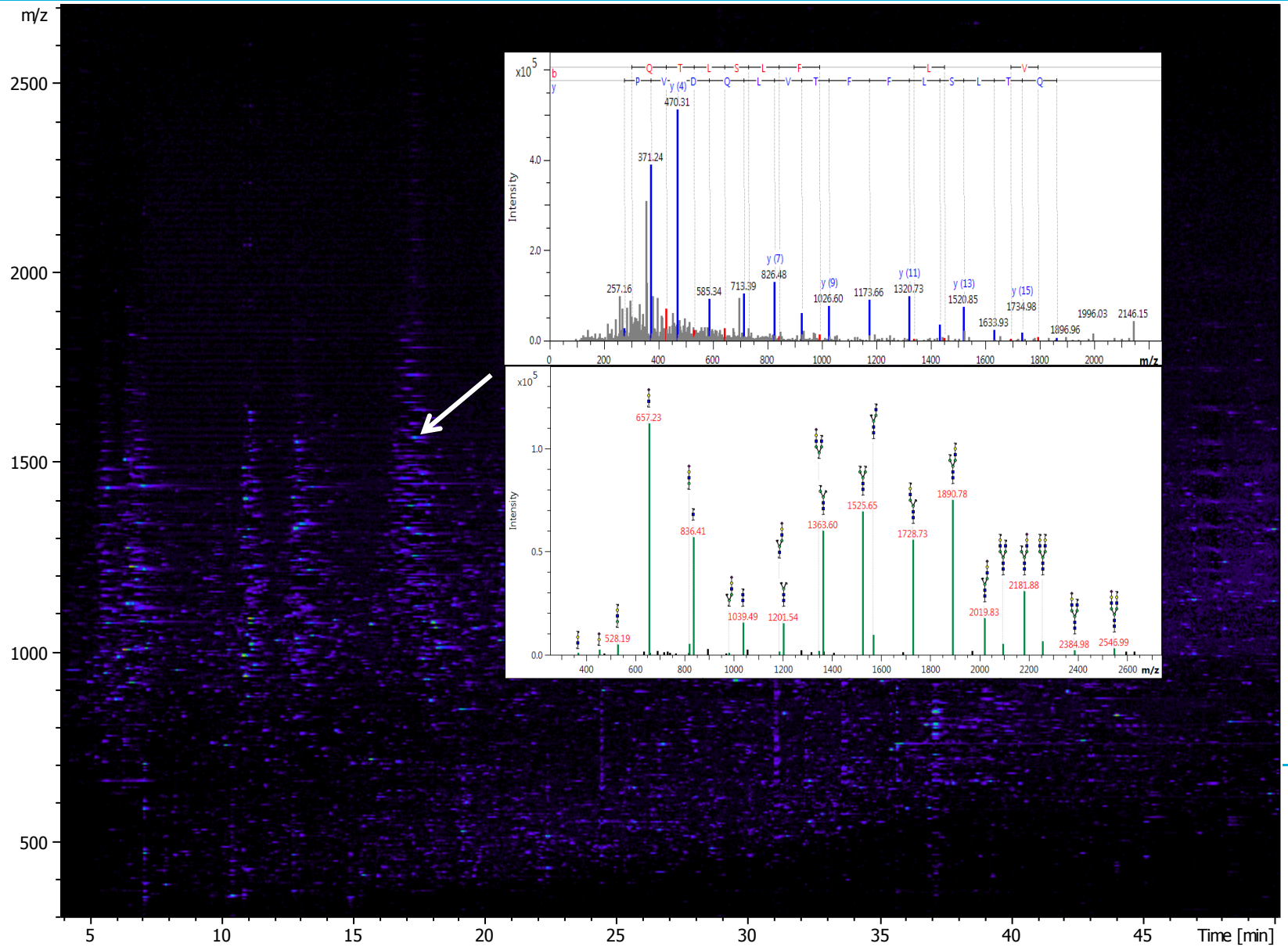
Glycopeptide profiling to examine the plasma glycoproteome

Truncated + fucosylated complex glycans

Complex glycans

CO4B_N226

CO4B_N1328

High-mannose glycans

N226

CO4B

N1328
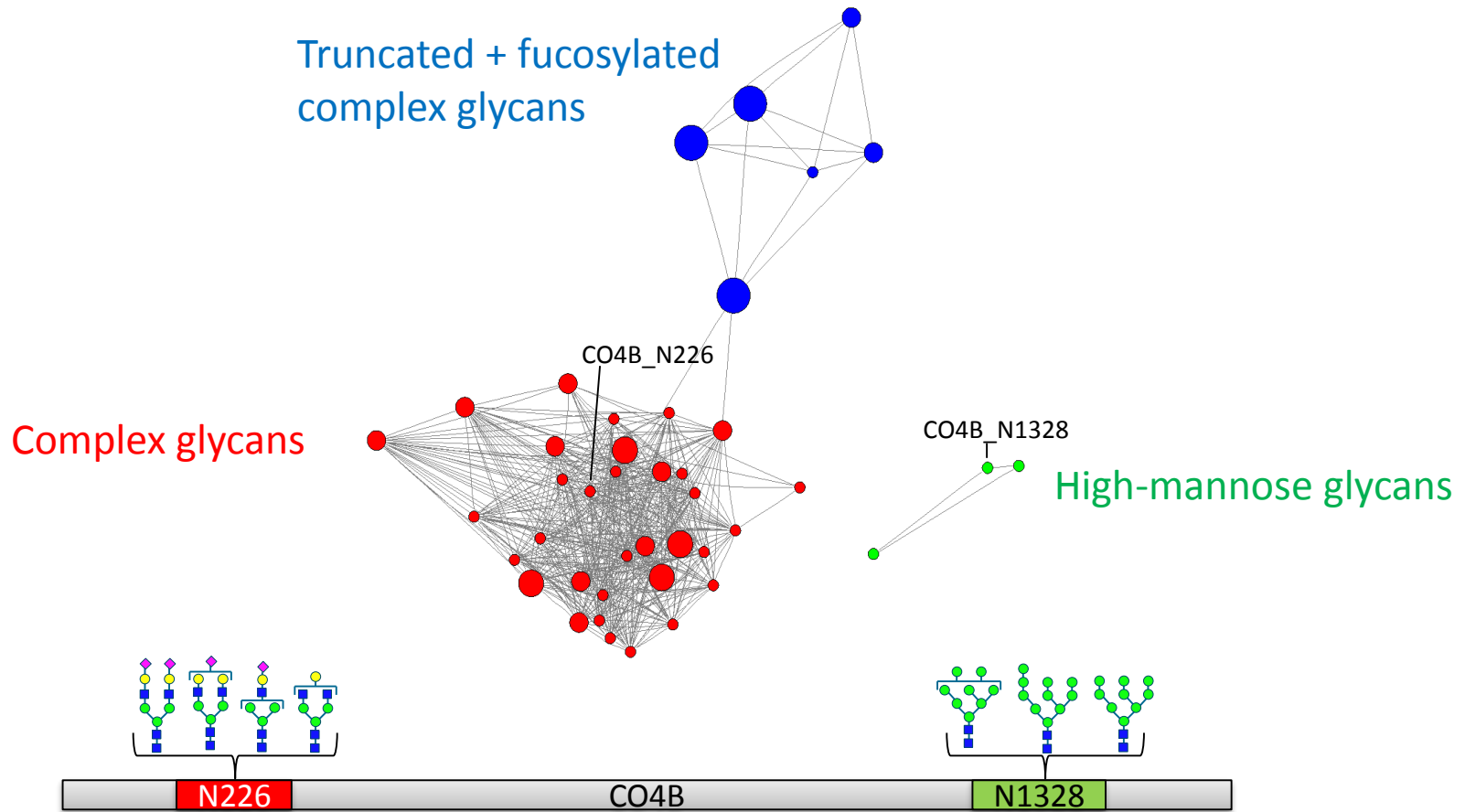
Site-specific glycosylation status for numerous plasma proteins excreted by different cell types can be assessed in a single experiment

Three distinct clusters of glycosylation profiles with varying microheterogeneity are observed in healthy individuals

Genetic defects in protein N-Glycosylation

Left top image: Peter Linders et al, IJMS 21(13):4654
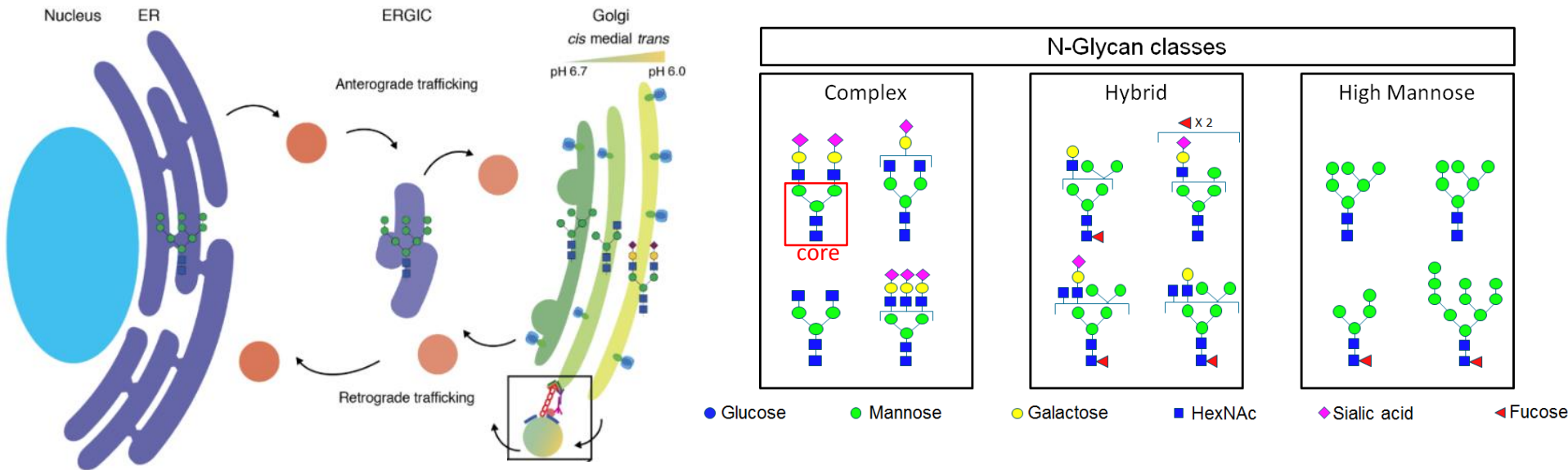
N-glycans are heterogeneous carbohydrate structures with common GlcNAc2Man3 core structure linked to Asparagine amino acid residues

N-glycan synthesis is a multi-organellar non-template driven enzymatic process

Analysis of protein-specific N-glycosylation changes at a proteome wide scale is highly anticipated to advance diagnostics and to explore fundamentals of glycobiology

Congenital Disorders of Glycosylation (CDG) as model to evaluate glycoproteomics performance and potential for clinical applications

Potential to improve diagnostic yields in CDG patient care

Gene defect	PLS-DA: AUC	
	GP	TF
ATP6AP1	1.00	0.94
ATP6V0A2	1.00	0.99
B4GALT1	1.00	1.00
CCDC115	1.00	1.00
COG5	1.00	0.82
DYM	1.00	0.87
MAN1B1	1.00	1.00
NANS	1.00	0.83
PGM1	1.00	0.93
TMEM199	1.00	1.00

For all 10 CDG defects every patient is correctly classified using glycopeptide profiling data

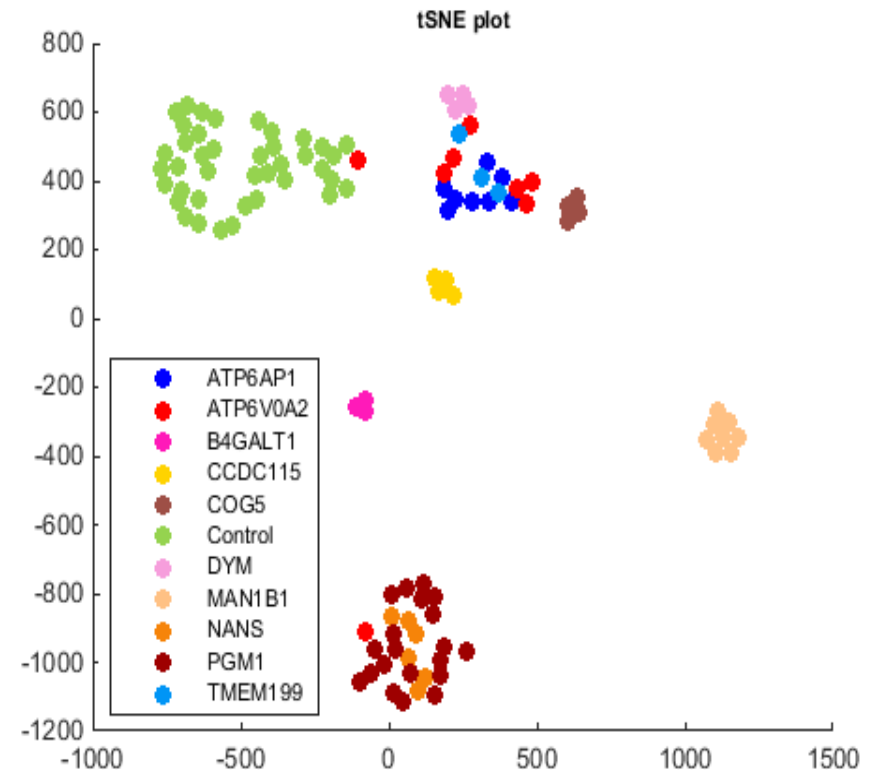Intact TF-MS data correctly classified every patient for 4 CDG defects
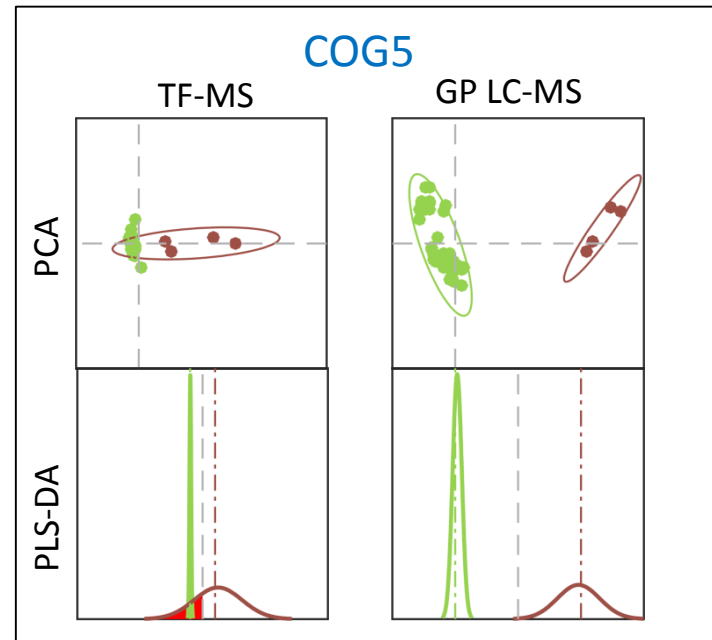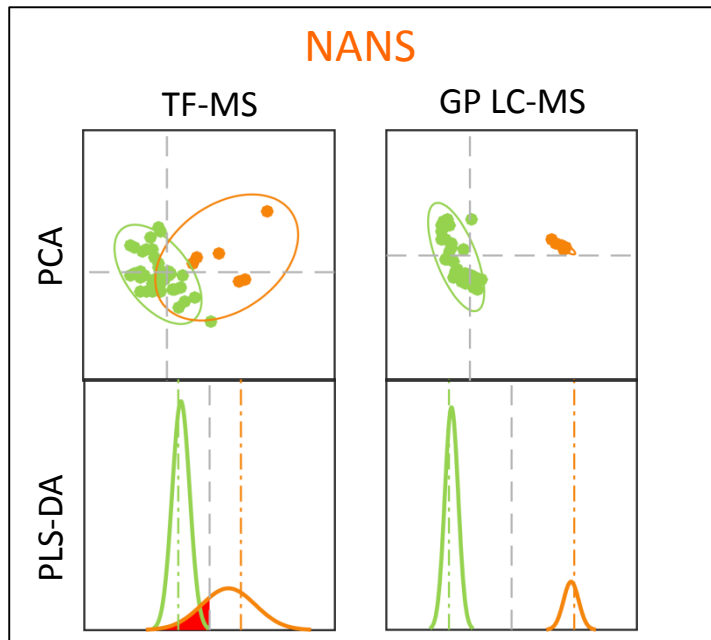
CDG stratification at affected gene- and functional levels

Genetic Algorith – Random Forest (GA-RF)

t-Stochastic Neighbour Embedding (tSNE)

	Control	ATP6AP1	ATP6V0A2	CCDC115	TMEM199	PGM1	NANS	COG5	DYM	B4GALT1	MAN1B1
Control	40	0	0	0	0	0	0	0	0	0	0
ATP6AP1	0	8	0	0	0	0	0	0	0	0	0
ATP6V0A2	0	0	8	0	0	0	0	0	0	0	0
CCDC115	0	0	0	3	0	1	0	0	0	0	0
TMEM199	0	1	2	0	0	0	0	0	0	0	0
PGM1	0	0	0	0	0	20	0	0	0	0	0
NANS	0	0	0	0	0	0	6	0	0	0	0
COG5	0	0	0	0	0	0	0	4	0	0	0
DYM	0	0	0	0	0	0	0	0	4	0	0
B4GALT1	0	0	0	0	0	0	0	0	0	3	0
MAN1B1	0	0	0	0	0	0	0	0	0	0	8

tSNE plot

Legend: ATP6AP1, ATP6V0A2, B4GALT1, CCDC115, COG5, Control, DYM, MAN1B1, NANS, PGM1, TMEM199

Supervised learning by GA-RF sucessfully classified 104 out of 108 samples to their correct sample classes (AUC=0.94)

3 out of 4 misclassifications resulted from assignment of TMEM199 samples to ATP6AP1 (n=1) and ATP6V0A2 (n=2) classes but all three defects lead to V-ATPase deficiency
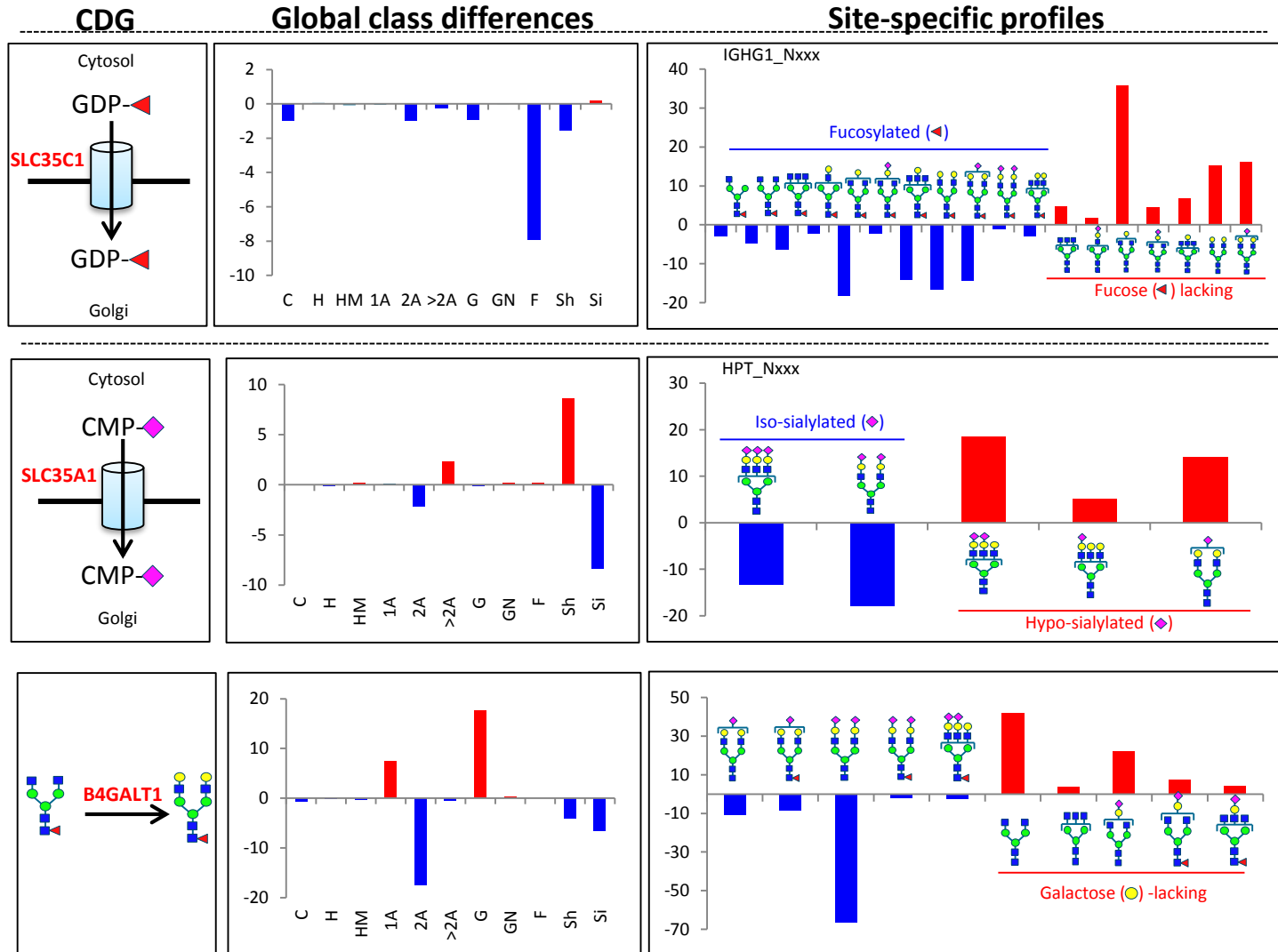
CDG stratification at affected gene- and functional levels

t-Stochastic Neighbour Embedding (tSNE)
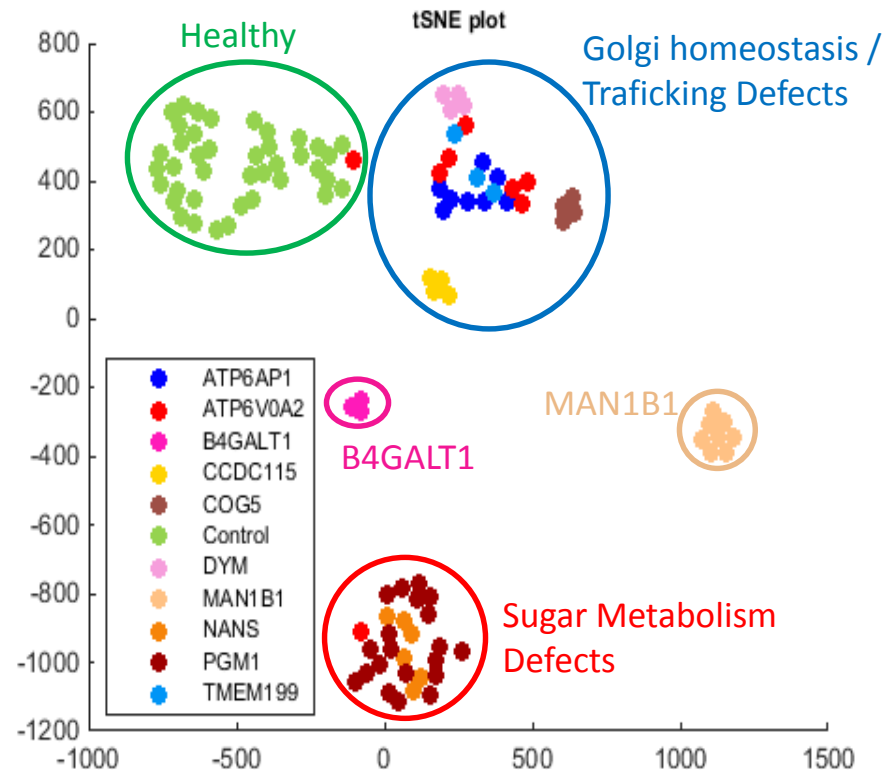
tSNE visualization of GA-RF selected feature data shows grouping of patient samples according to disturbed biological processes or isolated N-glycan synthesis steps

Radboud umc

Observed glycosylation changes are clinically relevant

Observed glycosylation changes correspond with expect increase/decrease of glycoforms according to distinctly disrupted N-glycan biosynthesis steps in patients.

From bench to bedside

Key lessons learned thus far
- Plasma glycoproteomics is analytically robust
- Glycopeptide analysis has high potential to advance diagnostics
- Reference glycosylation of proteins in healthy conditions determines if the protein can be used to monitor specific glycosylation changes
- Glycosylation changes can occur in a protein-, tissue- and site-specific manner

Considerations for diagnostic implementation
- Processing, visualization and evaluation for tens-of-thousands of signals is impractical
- Unambiguous reporting of results is essential
- Complete workflow needs to be complient with ISO-15189 guidelines

Classical biomarker discovery and implementation

Biomarker discovery (holistic)
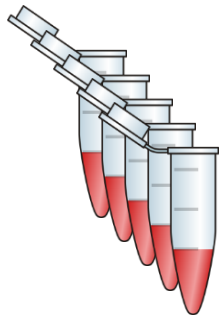
Patient cohort(s)
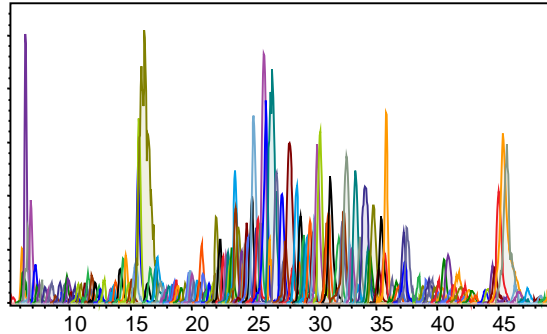
Holistic measurement
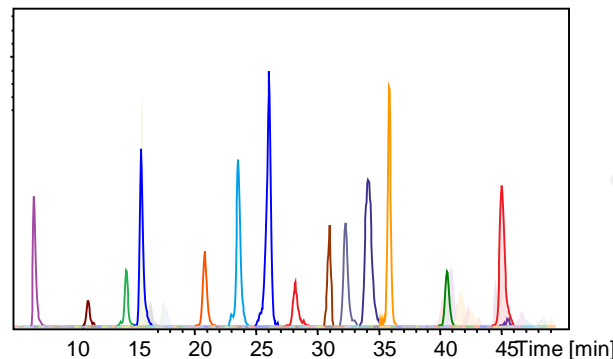(e.g. 250.000 signals)
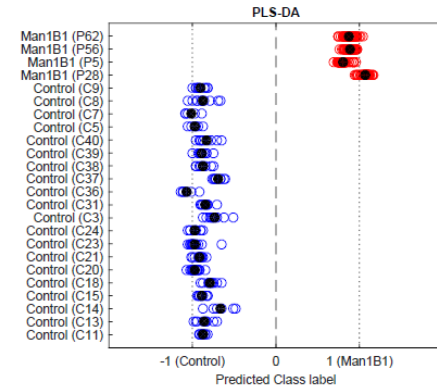
Biomarker discovery and
targeted assay development

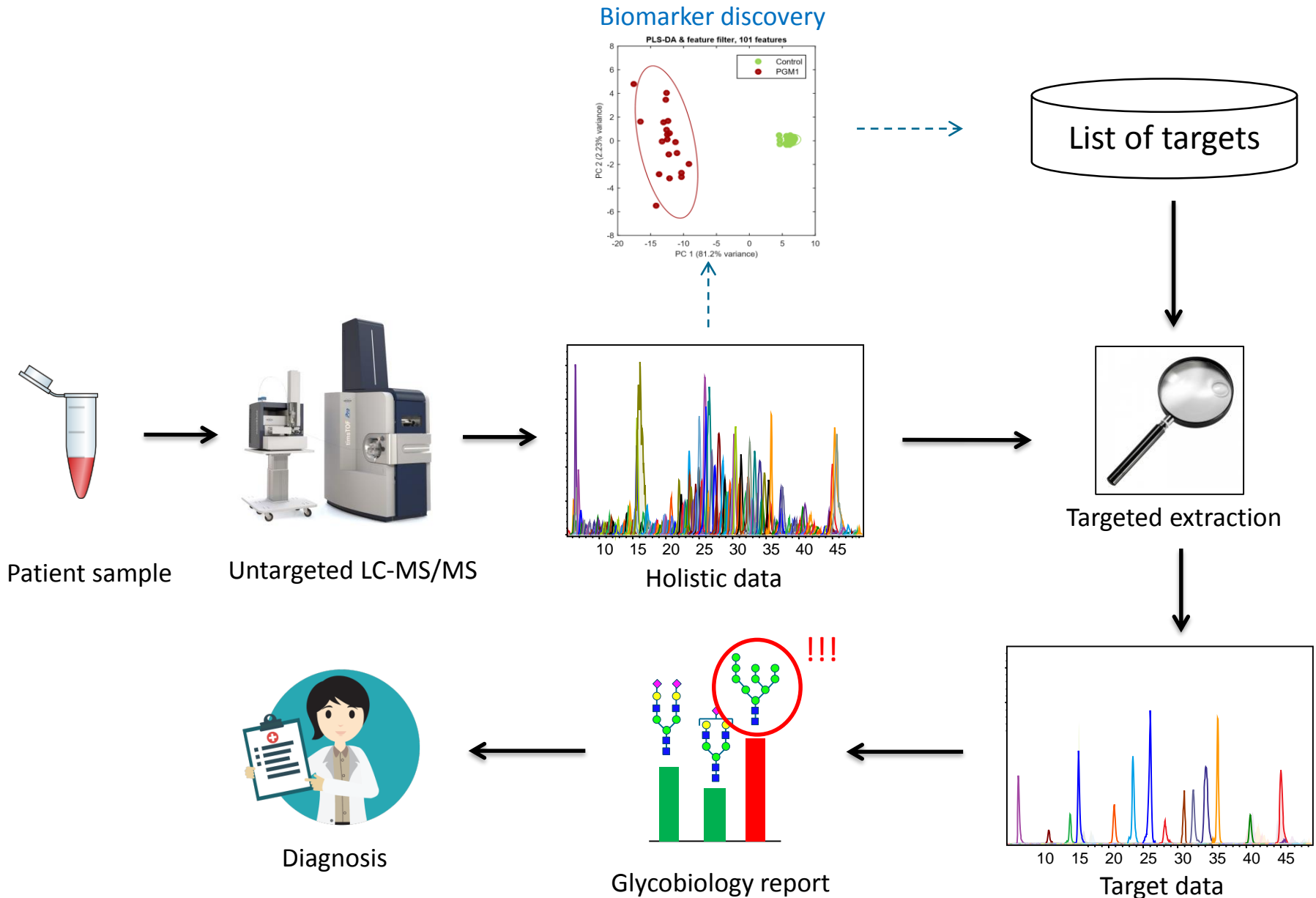## Clinical implementation (targeted)
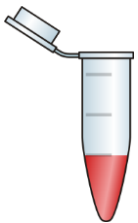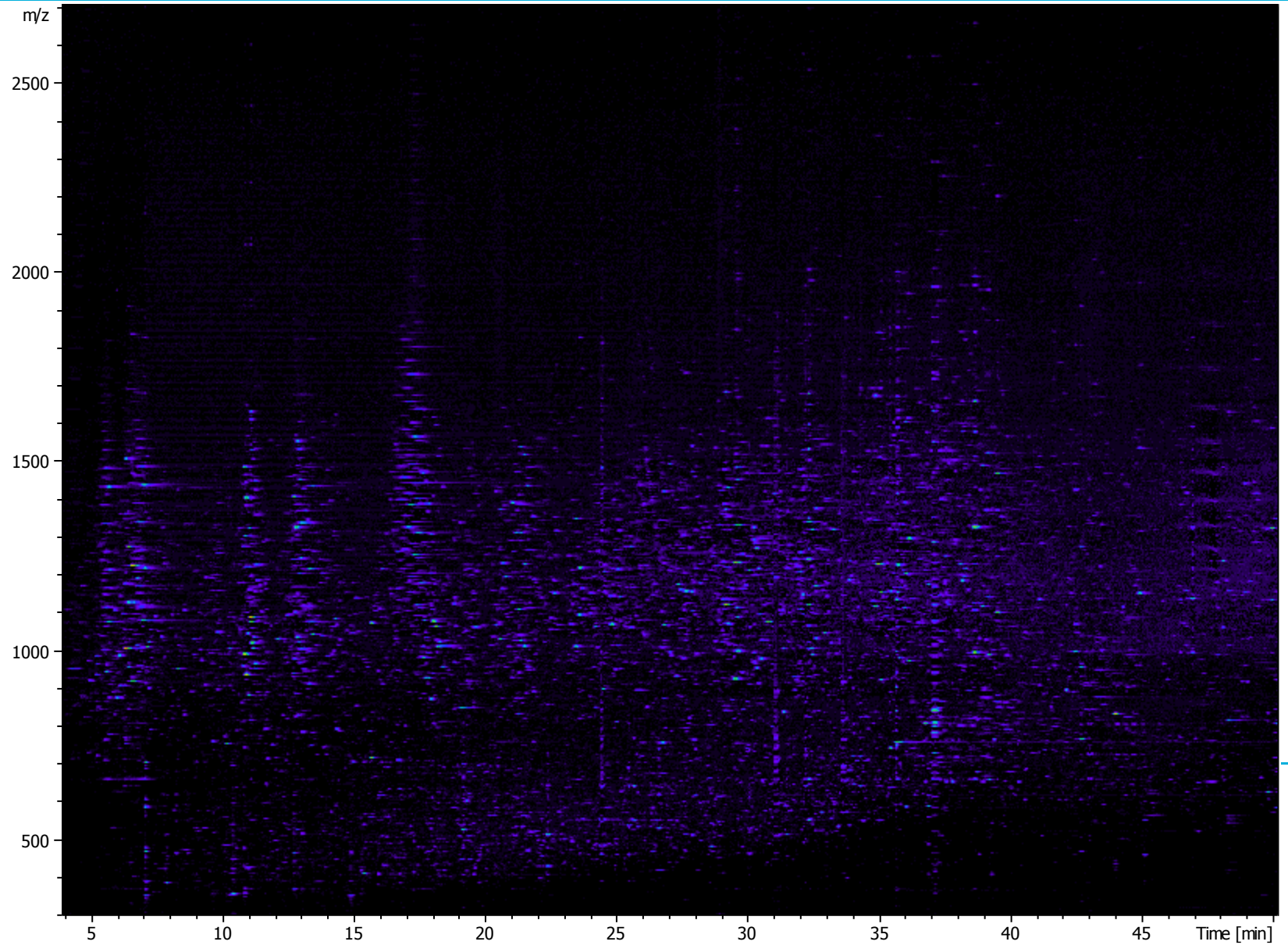

Patient sample
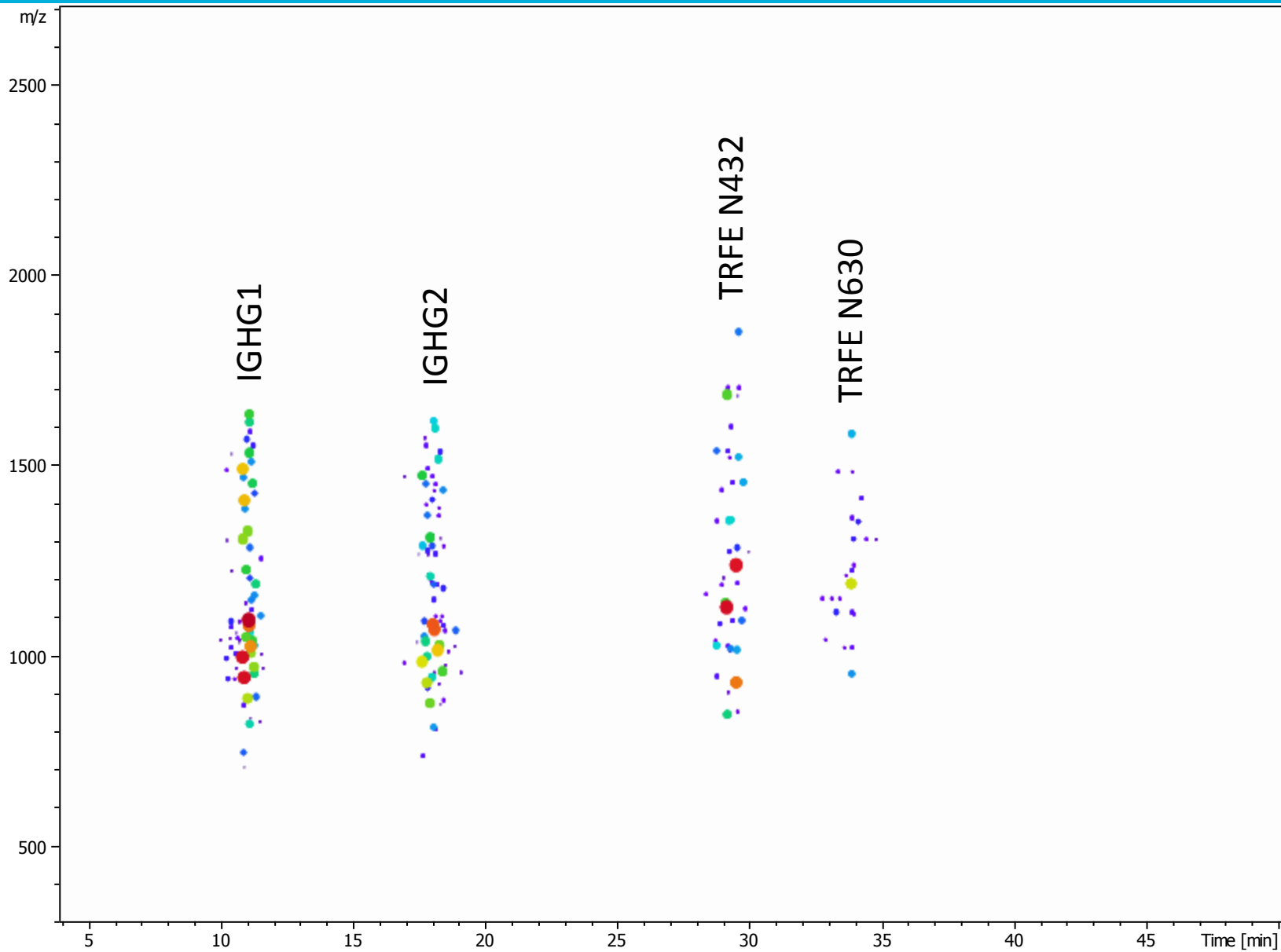
Targeted measurement
(e.g. 15 signals)

Report & Diagnosis

Holistic data acquisition for both research and diagnostics

Biomarker discovery

List of targets

Patient sample

Untargeted LC-MS/MS

Holistic data

Targeted extraction

Target data

Diagnosis

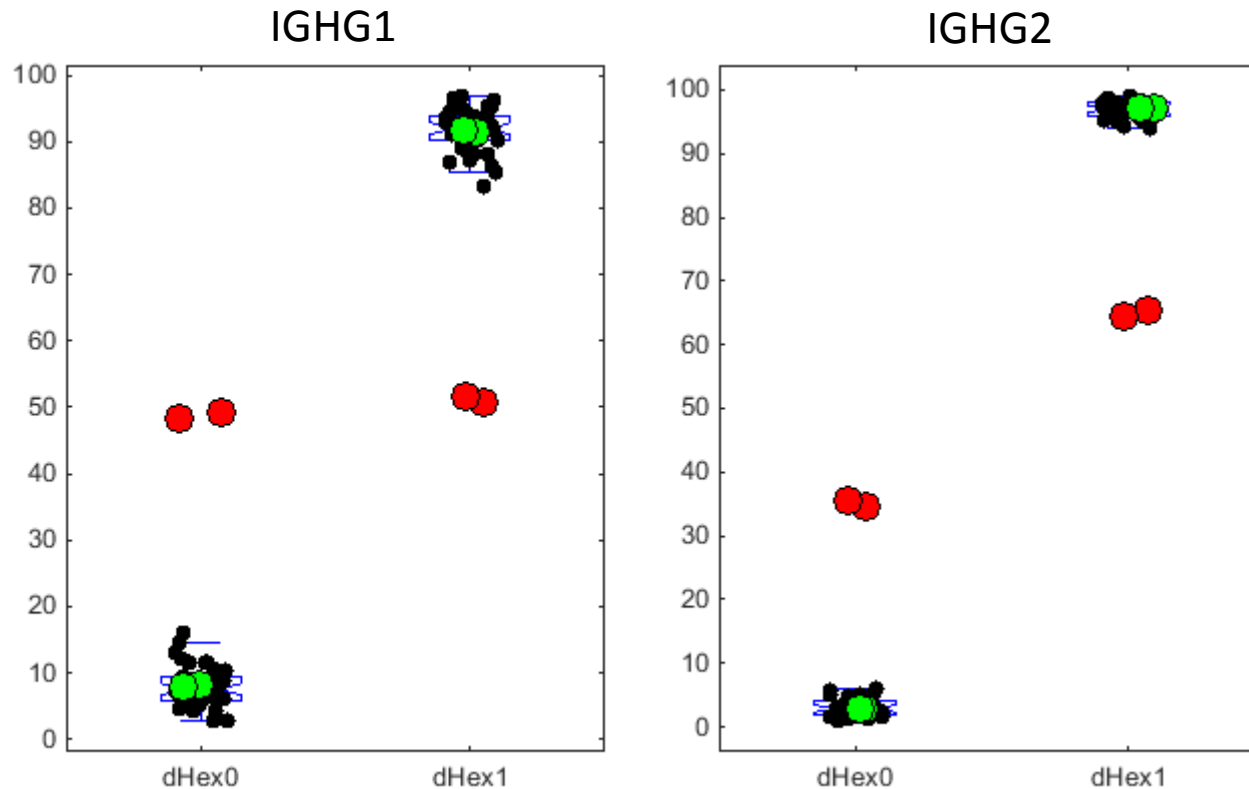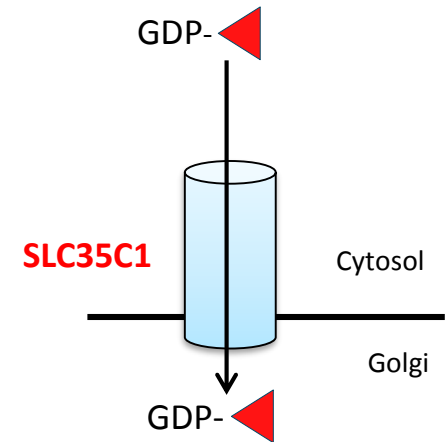Glycobiology report

Raw LC-MS data

Extracted glycopeptide features of interest

Fucose transporter (SLC35C1-CDG) deficiency

IGHG1

IGHG2
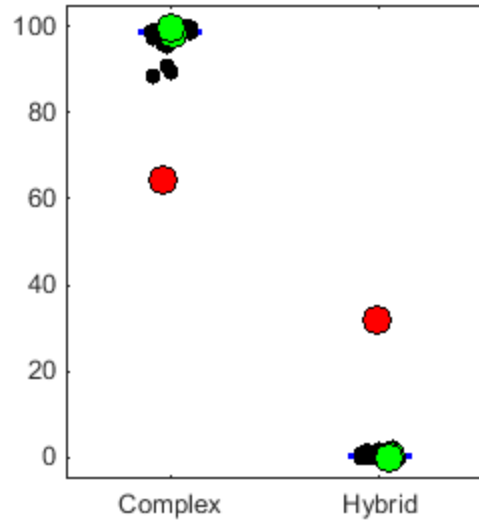
- ● Healthy donors (n=40)
- ● Pools of n=5 random healthy donors
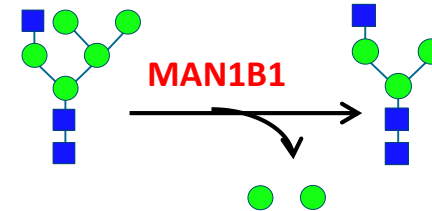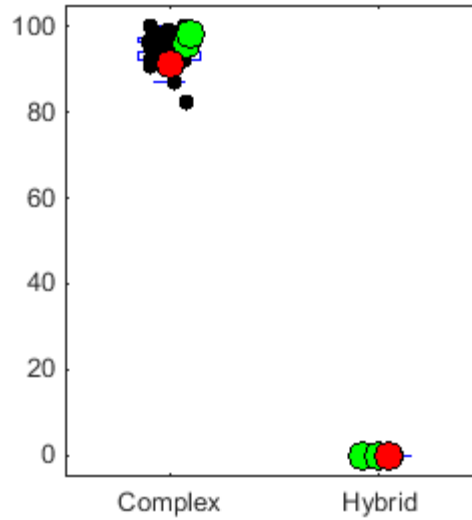- ● SLC35C1 patient samples (n=2; different time points)

GDP-

SLC35C1   Cytosol

Golgi

GDP-

Both IGHG1 and IGHG2 show significant reduction in fucosylation status

Radboud umc

Mannosidase deficiency (MAN1B1-CDG)
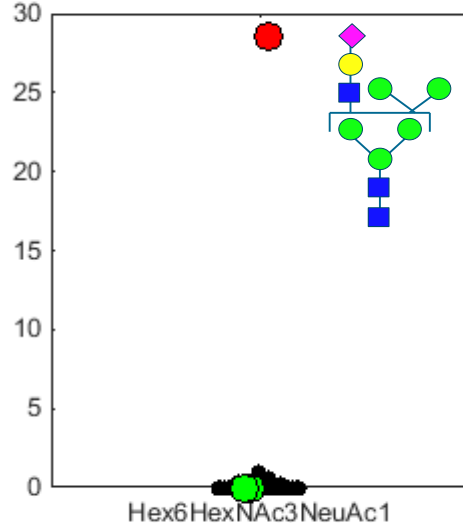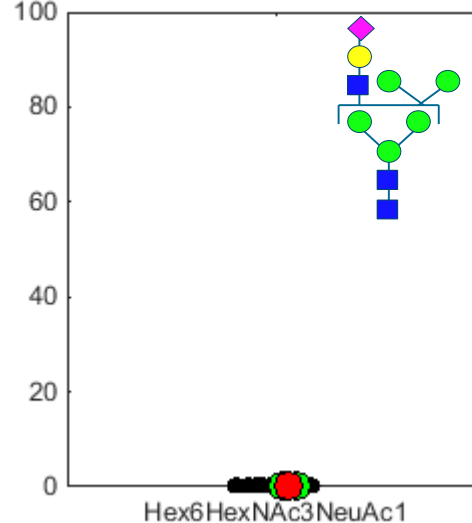

Healthy donors (n=40)
Pools of n=5 random healthy donors
MAN1B1 patient sample

Site-specific presence of diagnostic Hex6HexNAc3NeuAc1 hybrid glycan at N432 is readily observed

Radboud umc

Conclusions and outlook

Conclusions

- Glycoproteomics provides unique opportunities to diagnose, monitor and understand disease.
- Comprehensive plasma glycoproteomics of patient samples provides unique possibilities to explore fundamentals of protein glycobiology
- Plasma glycopeptide profiling can be efficiently implemented in CDG screening for improved patientcare

Outlook

- Extend application of clinical plasma glycoproteomics to diseases other than CDG including IEM, neurodegenerative diseases and cancer
- Develop clinical glycoproteomics for alternative patient materials such as CSF, muscle biopsy, cultured human skin fibroblasts and iPSC-derived cell lines
- Increase glycoproteome coverage via optimized hardware (timsTOF Pro; PASEF) and dedicated glycopeptide analysis software (BSI: Peaks glyco)

Radboud umc

Acknowledgements

Radboud umc
university medical center

Translational Metabolic Laboratory
Alain van Gool
Anouk Suppers
Dirk Lefeber
Esther Willems
Fokje Zijlstra
Hans Wessels
Jolein Gloerich
Koen Rademaker
Maurice van Dael
Melissa Bärenfänger
Merel Post
Monique van Scherpenzeel
Nurulamin Bin Abu Bakar
Purva Kulkarni
Ron Wevers
Thijs Deenik
Yanniek Lagarde

BRUKER

Bruker Daltonics
Gary Kruppa
Kristina Marx
Markus Lubeck
Pierre-Olivier Schmit
Scarlet Beck
Stuart Pengelley

BSi

Bioinformatics Solutions inc.
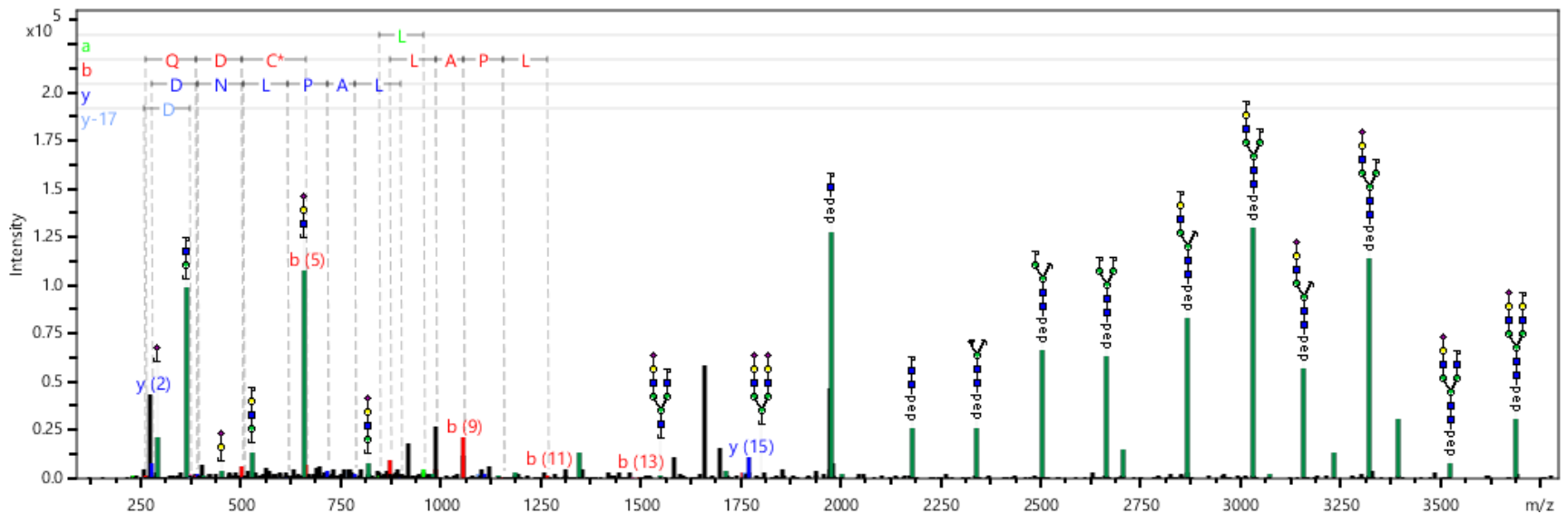Hui Li
Paul Shan
Tony Zhang
Winter Wei

ZonMw   X-OMICS.nl   Radboud Consortium for Glycoscience   umd·nl united for metabolic diseases   METAKIDS Stop metabole ziekten

Current Glycoproteomics: Theory, Methods and Applications

https://www.mdpi.com/journal/ijms/special_issues/glycoproteomics
Submission deadline: 31 December 2020
Edited by: Dr Hans Wessels & Dr Viviana Greco